

INAUGURAL-DISSERTATION

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathematischen

Gesamtfakultät der Ruprecht-Karls-Universität

Heidelberg

vorgelegt von

Dipl.-Phys. Tim Stenau

aus Paderborn

Tag der mündlichen Prüfung: 24.05.2017

Diffraktive Linsen mit überlappenden Aperturen

in Simulation und Anwendung

Gutachter: Prof. Dr. Karl-Heinz Brenner

Prof. Dr. Jürgen Jahns

Diffraaktive Linsen mit überlappenden Aperturen in Simulation und Anwendung

In dieser Arbeit werden das Konzept und potenzielle Anwendungen diffraktiver Linsen mit überlappenden Aperturen vorgestellt. Die herausragende Eigenschaft dieser Linsenarrays ist, dass die Fokusgröße unabhängig vom lateralen Abstand der Foki in der Fokalebene ist. Skalare Methoden zum Design der Mikrolinsenarrays werden beschrieben, die neben der Erzeugung einfacher punktförmiger und ringförmiger Foki auch die Erzeugung von komplexen Fokalgeometrien ermöglichen. Anhand skalarer wellenoptischer Simulationen werden die Eigenschaften der diffraktiven Linsen mit überlappenden Aperturen untersucht sowie Abweichungen von der idealen Struktur analysiert und quantifiziert. Die Simulationen zweier binärer diffraktiver Linsenarrays mit hohen numerischen Aperturen und integrierter Deckglaskorrektur sowie großem Überlappfaktor werden experimentell bestätigt. Des Weiteren werden Beispielanwendungen aufgezeigt, die von diffraktiven Linsen mit überlappenden Aperturen profitieren, wie z. B. das parallel scannende Mikroskop. Hier ermöglichen diffraktive Linsen mit überlappenden Aperturen gleichzeitig ein großes Sichtfeld, eine hohe Auflösung und einen großen Arbeitsabstand. Experimentell wird ein Demonstrator realisiert, der mit einem maximalen Sichtfeld von ca. $900\ \mu\text{m} \times 1200\ \mu\text{m}$ und einem Arbeitsabstand von 2 mm Strukturen bis zu 450 nm auflöst. Außerdem wird mit rigorosen Simulationen überprüft, inwieweit die skalare Simulation und das skalare Design valide sind. Selbst bei einer hohen numerischen Apertur von 0,9 erzeugen skalar designte Linsenarrays stark lokalisierte Foki.

Investigation of diffractive lenses with overlapping apertures in simulation and application

In this thesis, the concept of diffractive lenses with overlapping apertures and its potential applications are investigated. The outstanding feature of these lens arrays is that the focal dimension and the focal distance can be chosen independently of each other. Scalar design methods are presented and extended to create point-shaped, ring-shaped or arbitrary complex focal arrays. Based on scalar simulations, the characteristic features are analysed and deviations from the ideal structures are quantized. The simulations of binary phase-only diffractive lenses with overlapping apertures and integrated cover slip correction are verified experimentally for two high numerical aperture lens arrays with high overlap factors. Applications like the parallel scanning microscope benefit from the latter, where a high resolution, a large field of view and a large working distance become feasible at the same time. An experimental demonstrator with 2 mm working distance and a $900\ \mu\text{m} \times 1200\ \mu\text{m}$ field of view is realized, which is capable of resolving feature sizes of about 450 nm. Furthermore, the limits of the scalar design and simulation methods are compared to the results of rigorous simulations applying rigorous coupled-wave analysis. Even for high numerical apertures up to 0.9, the scalar design method still produces diffractive lens arrays with highly localised foci.

Inhaltsverzeichnis

Liste der verwendeten Symbole	xi
1 Einleitung	1
1.1 Motivation	1
1.2 Aufbau der Arbeit	1
2 Theoretische Grundlagen	4
2.1 Maxwell-Gleichungen	4
2.1.1 Mikroskopische Maxwell-Gleichungen	4
2.1.2 Makroskopische Maxwell-Gleichungen	5
2.1.3 Makroskopische Maxwell-Gleichungen im Frequenzraum	6
2.2 Reellwertige und monofrequente Felder	7
2.3 Materialgleichungen	9
2.4 Energieerhaltung	10
2.5 Impulserhaltung	13
2.6 Vektorielle Wellenoptik	17
2.6.1 Vektorielle Wellengleichung	17
2.7 Skalare Wellenoptik	18
2.7.1 Skalare Wellengleichung	18
2.7.2 Monochromatische skalare zeitharmonische Felder	19
2.7.3 Intensität eines skalaren Feldes	20
2.7.4 Energiedichte und Energieflussdichte eines skalaren Feldes	21
3 Simulationsmethoden	23
3.1 Lösungsmethoden für die skalare Wellengleichung	23
3.1.1 Exakte Methoden zum Lösen der skalaren Wellengleichung	23
3.1.2 Genäherte Methoden zum Lösen der skalaren Wellengleichung	26
3.2 Lösungsmethoden für die vektorielle Wellengleichung	29
3.2.1 Transversal polarisierte ebene Welle als Lösung der Wellengleichung	30
3.2.2 Exakte Methoden zur Lösung der vektoriellen Wellengleichung	31
3.2.3 Genäherte Methoden zur Lösung der vektoriellen Wellengleichung	32
3.3 Lösungsmethoden für die Maxwell-Gleichungen	33
3.3.1 Rigorous Coupled Wave Analysis	33

4	Design von diffraktiven Linsenarrays mit überlappenden Aperturen	44
4.1	Design-Methoden	44
4.1.1	Überlappende Aperturen durch kohärente Überlagerung	44
4.1.2	Überlappende Aperturen durch Fourier-Ansatz	45
4.2	Design spezieller Linsenarrays mit überlappenden Aperturen	47
4.2.1	Beliebige Beleuchtungsstrukturen	47
4.2.2	Punktförmige Beleuchtungsstrukturen	48
4.2.3	Ringförmige Beleuchtungsstrukturen	54
4.3	Erweiterungen des Design-Prozesses	64
4.3.1	Überlagerung von mehreren Beleuchtungsstrukturen	64
4.3.2	Deckglaskorrektur	65
4.4	Diffraktive Linsen als dünnes Element	68
4.4.1	Modifikationen von dünnen diffraktiven Elementen	68
4.4.2	Dünne diffraktive Linsen für zwei verschiedene Beleuchtungs- wellenlängen	69
4.5	Geometrie des Fokus für diffraktive Linsen mit homogener kreisfö- miger Pupille	70
4.6	Fazit	73
5	Evaluation von diffraktiven Linsen	74
5.1	Theoretische Grenzen für die Effizienz von fokussierenden Lichtver- teilungen	74
5.1.1	Exakte Lösungen für fokussierende Lichtverteilungen in der skalaren Wellenoptik	75
5.1.2	Genäherte Lösungen für fokussierende Lichtverteilungen in der skalaren Wellenoptik	78
5.1.3	Vergleich der analytischen Lichtverteilungen	79
5.1.4	Fokusintensität des idealen DOEs	80
5.1.5	Vergleich der idealen DOE mit der Sommerfeld-Welle	81
5.1.6	Alternative Wahl zur Quantisierung des Energieflusses	83
5.2	Auswirkung von Amplitudennivellierung und Phasendiskretisierung	86
5.2.1	Einfluss von Amplitudennivellierung und Phasenbinarisierung auf die Intensität im Fokus	86
5.2.2	Einfluss der Amplitudennivellierung und Phasendiskretisierung auf das Fokusprofil	90
5.2.3	Analytische Beschreibung der Phasendiskretisierung	95
5.2.4	Spiegelsymmetrie des Intensitätsprofils für binarisierte Pha- senelemente mit niedriger numerischer Apertur	98
5.3	Wählen der optimalen Phasendiskretisierung	100
5.4	Einfluss der Pixelierung des diffraktiven Elements	101
5.5	Auswirkung der Endlichkeit der Periodizität	103
5.6	Abweichungen der Design-Parameter	104
5.6.1	Abweichungen in der Wellenlänge	104
5.6.2	Abweichungen in der Phasenhöhe	106

5.6.3	Abweichungen durch verwaschene Kanten	108
5.7	Leitfaden zur Wahl der Parameter im Design	110
5.8	Fazit	110
6	Vermessung spezieller diffraktiver Elemente mit überlappenden Aperturen	112
6.1	Herstellung von diffraktiven Elementen in Glas	112
6.2	Messung der Pupille	115
6.2.1	Experimenteller Aufbau zur Pupillenmessung	115
6.2.2	Durchführung der Messung	115
6.2.3	Auswertung zur Bestimmung der Phasenhöhe	115
6.3	Messung des Fokusverlaufs	120
6.3.1	Aufbau zur Vermessung des Fokus	120
6.3.2	Ergebnisse der Vermessung des Fokus	122
6.4	Fazit	125
7	Anwendung: Parallel scannendes Mikroskop mit diffraktiven Linsen mit überlappenden Aperturen	127
7.1	Limitationen in der Mikroskopie	127
7.2	Parallel scannendes Mikroskop mit diffraktiven Linsen mit überlappenden Aperturen	130
7.2.1	Funktioneller Aufbau des parallel scannenden Mikroskops . . .	130
7.2.2	Experimentelle Umsetzung des parallel scannenden Mikroskops	130
7.3	Theorie des parallel scannenden Mikroskops	133
7.3.1	Theorie der Bildentstehung	133
7.3.2	Theoretische Auflösung des parallel scannenden Mikroskops .	135
7.4	Aufnahme von Messdaten mit dem parallel scannendem Mikroskop .	143
7.4.1	Ablauf der Messung	143
7.4.2	Dauer einer Messung	144
7.5	Bildrekonstruktion aus Rohbildern	145
7.5.1	Bestimmung des Beleuchtungsrasters	145
7.5.2	Auswertung der Transmission des Beleuchtungsgitters	147
7.6	Experimentelle Bestimmung der maximalen Auflösung des parallel scannenden Mikroskops	152
7.6.1	Bestimmung der maximalen Auflösung durch den Gitterkontrast eines Auflösungstests	152
7.6.2	Vergleich mit der Theorie	156
7.7	Ausblick auf weitere Anwendungsgebiete diffraktiver Linsen mit überlappenden Aperturen in der Mikroskopie	158
7.8	Zusammenfassung und Fazit	158
8	Anwendung: Wellenfrontmessung	160
8.1	Wellenfrontmessung mittels eines Hartmann-Shack-Sensors	160
8.2	Potenzial von diffraktiven Mikrolinsen im Hartmann-Shack-Sensor . .	162

8.3	Fazit	162
9	Aspekte zur nicht skalaren Untersuchung von diffraktiven Linsen	164
9.1	RCWA als Simulationsmethode für diffraktive Elemente mit überlap- penden Aperturen	164
9.2	RCWA Simulation von Linsen mit überlappender Apertur	166
9.2.1	Übergang vom dünnen Element zum dicken Element	166
9.2.2	Simulation von dicken diffraktiven Elementen	167
9.2.3	Wie lange ist die skalare Rechnung für diffraktive Linsen valide?	177
9.3	Beschleunigung der RCWA durch eine Sonderfall-Behandlung für ho- mogene Schichten	183
9.4	Konsistenzüberprüfung der RCWA anhand der Impulserhaltung	184
9.4.1	Beispiel: Homogene Schicht und TM-Polarisation	185
9.4.2	Beispiel: Homogener Block und TM-Polarisation	188
9.5	Fazit	193
10	Zusammenfassung und Ausblick	194
A	Appendix	197
A.1	Koordinatensysteme	197
A.2	Spezielle Funktionen und Distributionen	198
A.3	Komplexe Zahlen	199
A.4	Rechenregeln für Vektoren und Tensoren	199
A.5	Vektoridentitäten	200
A.6	Fourier-Transformations-Paare	200
A.7	Zeit-Mittelung für monofrequente Felder	200
A.8	Gaußscher Integralsatz für Vektorfelder mit Unstetigkeit	202
A.9	Herleitung des Energie-Erhaltungssatzes	202
A.10	Herleitung des Impuls-Erhaltungssatzes	203
A.10.1	Impulserhaltung nach Chu	203
A.10.2	Impulserhaltung nach Minkowski	204
A.11	Herleitung der Wellengleichung	205
A.12	Definition des Brechungsindex	207
A.13	Herleitung der z-Komponenten der Felder und der Modenkopplung im Gitter in der RCWA	207
A.13.1	Bestimmung der z-Komponenten im Gitter	208
A.13.2	Bestimmung der z-Ableitung der lateralen Feldkomponenten	209
A.14	Herleitung der Intensität im Fokus für spezielle analytische Wellen	211
A.15	Bestimmung spezieller Überlappintegrale	214
B	Literatur	216
C	Eigene Publikationsliste	224

Liste der verwendeten Symbole

Notation	Beschreibung
$\delta(\mathbf{r}), \delta(\mathbf{k})$	Delta-Distribution im Ort- bzw. spatialen Frequenzraum
ϵ	Binarisierungsparameter
$\epsilon, \epsilon(\mathbf{r}, \omega), \epsilon(\mathbf{r})$	Relative Permittivität im Ortsraum
$\tilde{\epsilon}, \tilde{\epsilon}(\mathbf{k}, \omega), \tilde{\epsilon}(\mathbf{k})$	Relative Permittivität im spatialen Frequenzraum
$\tilde{\epsilon}_{mn}$	Fourierkoeffizient der relativen Permittivität im spatialen Frequenzraum
$[[\tilde{\epsilon}]]$	Block-Toeplitz-Matrix der Fourierkoeffizienten der relativen Permittivität im spatialen Frequenzraum
ϵ_0	Permittivität des Vakuums
η	Effizienz
$\eta_{\text{Strehl}}^{\text{gen.}}$	Generalisiertes Strehl-Kriterium
η_p	Effizienz der Phasenmodulationsfunktion p
η_{Strehl}	Strehl-Verhältnis
κ	Imaginärteil des komplexen Brechungsindex
Λ	Eigenwert-Diagonalmatrix
Λ	Typische Strukturgröße im Ortsraum
λ	Wellenlänge
λ_j	j -ter Eigenwert
μ	Spatiale Frequenzkoordinate
μ_0	Magnetische Feldkonstante
$\mu, \mu(\mathbf{r}_\perp, \omega), \mu(\mathbf{r}_\perp)$	Relative Permeabilität im Ortsraum
$\tilde{\mu}, \tilde{\mu}(\mathbf{k}_\perp, \omega), \tilde{\mu}(\mathbf{k}_\perp)$	Relative Permeabilität im spatialen Frequenzraum
$\tilde{\mu}_{mn}$	Fourierkoeffizient der relativen Permeabilität im Frequenzraum
$[[\tilde{\mu}]]$	Block-Toeplitz-Matrix der Fourierkoeffizienten der relativen Permeabilität im Frequenzraum
ν	Spatiale Frequenzkoordinate
$\rho(\mathbf{r}, t)$	Elektrische Raumladungsdichte im Ortsraum
$\rho_{\text{frei}}(\mathbf{r}, t)$	Elektrische Raumladungsdichte freier Ladungen im Ortsraum
$\varrho(\mathbf{r}, \omega)$	Elektrische Raumladungsdichte im temporalen Frequenzraum
$\varrho_{\text{frei}}(\mathbf{r}, \omega)$	Elektrische Raumladungsdichte freier Ladungen im temporalen Frequenzraum
ρ	Frequenzvektor im dreidimensionalen spatialen Frequenzraum
ρ_\perp	Frequenzvektor in der $\nu\mu$ -Ebene

Notation	Beschreibung
Φ^+, Φ^-	RCWA-Propagationsmatrizen im Gitter
$\phi, \phi(\mathbf{r}_\perp)$	Phase
φ	Winkelvariable
χ	Polarwinkel der spatialen Frequenzkoordinaten
Ω	Überlappfaktor im Ortsraum
Ω	Raumwinkel
Ω_1, Ω_2	Moden-Kopplungsmatrix in der RCWA
ω	Temporale Frequenzkoordinate
ω_0	Frequenz einer monochromatischen Welle
A	Fläche
$A(\mathbf{r})$	Reelle Amplitude eines skalaren Feldes
$\mathbf{A}, \tilde{\mathbf{A}}$	Transformationsmatrix der tangentialen Feldkomponenten
\mathcal{A}	Globale Absorbtion
\mathbf{a}_l	ETMA-Hilfsmatrix für den l -ten Layer
$\mathbf{B}, \mathbf{B}(\mathbf{r}, t)$	Magnetische Flussdichte im Ortsraum
$\mathcal{B}, \mathcal{B}(\mathbf{r}, \omega)$	Magnetische Flussdichte im temporalen Frequenzraum
\mathbf{b}_l	ETMA-Hilfsmatrix für den l -ten Layer
\vec{C}, C_j	Lösung des H-Felds im Eigenmodenraum des Gitters
c	Vakuumlichtgeschwindigkeit
c_j, c_j^+, c_j^-	RCWA-Eigenmodenkoeffizienten des elektrischen Feldes
c_{mn}	Fourierkoeffizienten
c_{Fokus}	Normierte Intensität im Fokus
D	Durchmesser
$\mathbf{D}, \mathbf{D}(\mathbf{r}, t)$	Elektrische Flussdichte im Ortsraum
$\mathcal{D}, \mathcal{D}(\mathbf{r}, \omega)$	Elektrische Flussdichte im temporalen Frequenzraum
\vec{D}, D_j	Lösung des E-Felds im Eigenmodenraum des Gitters
dk_x, dk_y	Abstand zwischen zwei Abtastpunkten im spatialen Frequenzraum
dx, dy	Abstand zwischen zwei Abtastpunkten im Ortsraum
$\mathbf{E}, \mathbf{E}(\mathbf{r}, t)$	Reelle elektrische Feldstärke als Funktion des Ortes und der Zeit.
$\tilde{\mathbf{E}}, \tilde{\mathbf{E}}(\mathbf{k}, t)$	Komplexe elektrische Feldstärke im spatialen Frequenzraum als Funktion des Wellenvektors und der Zeit
$\mathcal{E}, \mathcal{E}(\mathbf{r}, \omega)$	Komplexe elektrische Feldstärke im Ortsraum als Funktion des Ortes und der Frequenz
$\tilde{\mathcal{E}}, \tilde{\mathcal{E}}(\mathbf{k}, \omega)$	Komplexe elektrische Feldstärke im spatialen Frequenzraum als Funktion des Wellenvektors und der Frequenz
$\hat{\mathbf{e}}$	Einheitsvektor
F_N	Blendenzahl
$\mathbf{F}(\mathbf{r}, t)$	Skalare Energieflussdichte
$\mathbf{F}_{\text{EM}}(\mathbf{r}, t)$	elektromagnetische Kraft

Notation	Beschreibung
\mathcal{F}	Fourier-Transformations-Operator
$\mathbf{f}(\mathbf{r}, t)$	Elektromagnetische Kraftdichte
\mathbf{f}_l	ETMA-Hilfsmatrix für den l -ten Layer
G_α	Fourierkoeffizienten bei der Entwicklung nach der Phase
$\mathbf{G}(\mathbf{r}, t)$	Elektromagnetische Impulsdichte
\mathbf{g}_l	ETMA-Hilfsmatrix für den l -ten Layer
$H(\mathbf{r}, t)$	Skalare Energiedichte
\mathcal{H}	Hankel-Transformations-Operator
$\mathbf{H}, \mathbf{H}(\mathbf{r}, t)$	Magnetische Feldstärke im Ortsraum
$\mathcal{H}, \mathcal{H}(\mathbf{r}, \omega)$	Magnetische Feldstärke im temporalen Frequenzraum
$\tilde{\mathcal{H}}, \tilde{\mathcal{H}}(\mathbf{k}, \omega)$	Magnetische Feldstärke im spatialen und temporalen Frequenzraum
$h(\mathbf{r}), h_z(\mathbf{r})$	Impulsantwort, Punktbildverwaschungsfunktion
$I(\mathbf{r}, t)$	Intensität eines skalaren Feldes
\mathbf{I}	Einheitsmatrix
I_{ges}	Integrierte Gesamtintensität
i	Imaginäre Einheit
J_ν	ν -te Bessel-Funktion erster Gattung
$\mathbf{j}(\mathbf{r}, t)$	Elektrische Stromdichte im Ortsraum
$\mathbf{j}_{\text{frei}}(\mathbf{r}, t)$	Elektrische Stromdichte der freien Ladungsträger im Ortsraum
$\mathbf{j}(\mathbf{r}, \omega)$	Elektrische Stromdichte im temporalen Frequenzraum
$\mathbf{j}_{\text{frei}}(\mathbf{r}, \omega)$	Elektrische Stromdichte der freien Ladungsträger im temporalen Frequenzraum
K	Streifenkontrast
\mathbf{K}	RCWA-Wellenvektor-Matrix
k_0	Wellenzahl im Vakuum
\mathbf{k}	Wellenvektor
$\mathbf{L}_{mn}, \vec{L}_\perp$	Fourier-Koeffizienten des eingestrahltten Lichts in der RCWA
M	Gesamt-Modenzahl in der RCWA
$\mathbf{M}, \mathbf{M}(\mathbf{r}, t)$	Magnetisierung eines Mediums
MO_x, MO_y	Modenzahl in x bzw. y-Richtung in der RCWA
\mathbf{n}	Flächennormalenvektor
$n(\mathbf{r})$	Komplexer Brechungsindex
$n_r(\mathbf{r})$	Realteil des komplexen Brechungsindex
P, P_x, P_y	Periode
$P_z(\boldsymbol{\rho})$	Skalarer Propagator
$\mathbf{P}, \mathbf{P}(\mathbf{r}, t)$	Elektrische Polarisation eines Mediums, elektrisches Dipolmoment pro Volumen
$\mathbf{P}_1, \mathbf{P}_2$	Gittervektoren
\mathcal{P}	Polarisationsvektor
\mathbf{Q}	RCWA-Matrix

Notation	Beschreibung
R	Radius
\mathcal{R}	Globale Reflexion
\mathbf{R}_{mn}	Fourierkoeffizienten des reflektierten Lichts in der RCWA
\mathbf{r}	Positionsvektor im dreidimensionalen Raum
\mathbf{r}_\perp	Positionsvektor in der xy-Ebene
r	Länge des Vektors \mathbf{r}
r_\perp	Länge des Vektors \mathbf{r}_\perp
\mathbf{S}	Poynting-Vektor
$\mathbf{S}_{mn}, \vec{S}_\perp$	Entwicklungskoeffizient des elektrischen Felds im Gitter in der RCWA
\mathbf{s}	Normierter Richtungsvektor
\mathbf{s}_\perp	Zweidimensionaler Richtungsvektor
T	Zeitliche Periode
\mathcal{T}	Globale Transmission
$\mathbf{T}(\mathbf{r}, t)$	Maxwellscher Spannungstensor
\mathbf{T}_{mn}	Fourierkoeffizienten des transmittierten Lichts in der RCWA
t	Zeit
$t(\mathbf{r}_\perp)$	Skalare Transmissionsfunktion für komplexe Amplituden
$U(\mathbf{r}, \omega)$	Komplexe Amplitude eines skalaren zeitharmonischen Feldes
U_0	Einheitenbehaftete Skalierungskonstante der komplexen Amplitude eines skalaren zeitharmonischen Feldes
$\tilde{U}(\mathbf{k}, \omega)$	Komplexe Amplitude eines skalaren zeitharmonischen Feldes im spatialen Frequenzraum
$\mathbf{U}_{mn}, \vec{U}_\perp$	Entwicklungskoeffizienten des magnetischen Feldes im Gitter in der RCWA
V	Volumen
$V(\mathbf{r}, t)$	Amplitude eines skalaren Feldes
v	Geschwindigkeit
W	Energiedichte
\mathbf{W}	Eigenvektormatrix
\mathbf{X}	RCWA-Matrix
x	Spatiale Koordinate
y	Spatiale Koordinate
Z_0	Wellenimpedanz des Vakuums
z	Spatiale Koordinate
z_T	Talbot-Distanz

1 Einleitung

1.1 Motivation

Diffraaktive Linsen mit überlappenden Aperturen sind neue Bausteine in der Werkzeugkiste des optischen Designers. Linsenarrays mit überlappenden Aperturen wurde in der Literatur explizit erstmals im Jahr 2012 von Hulsken, Vossen und Stallinga [40] beschrieben. Die Grundlage der Beschreibung bildet jedoch die kohärente skalare Optik, die spätestens seit der ersten Auflage von Goodmans „Introduction to Fourier Optics“ aus dem Jahr 1968 wohlbekannt ist [34].

Die herausragende Eigenschaft diffraktiver Linsen mit überlappenden Aperturen ist, dass Fokalarrays erzeugt werden können, bei denen die Fokusgröße unabhängig vom Abstand der Foki in der Fokalebene gewählt werden kann. Zu jedem dieser Foki lässt sich ein effektiver Aperturdurchmesser in der Linsenebene definieren. Die Besonderheit ist, dass dieser Aperturdurchmesser größer als der Abstand zwischen zwei Foki sein kann. Dies führt dazu, dass die effektiven Aperturen in der Ebene der diffraktiven Linsen überlappen, woher auch die Bezeichnung „diffraktive Linsen mit überlappenden Aperturen“ stammt.

Dieses Überlappen der Aperturen ist nur mit diffraktiven Linsen möglich. Bei refraktiven Linsen ist die minimale Fokusgröße bei vorgegebener Brennweite vom Aperturdurchmesser der Linse abhängig, gleichzeitig bestimmt der Aperturdurchmesser aber auch den minimalen Abstand der Foki in der Fokalebene. Somit sind bei refraktiven Linsen der Abstand der Foki und die minimale Fokusgröße gekoppelt. Insbesondere eine Beleuchtung mit hochaperturigen Foki auf einem engen Raster ist mit refraktiven Linsen somit nur begrenzt möglich. Diffraktive Linsen mit überlappenden Aperturen erlauben hingegen eine punktförmige Beleuchtung mit hochaperturigen Foki auf einem engen Raster, welches mit refraktiver Optik nicht realisierbar ist.

Das Ziel dieser Arbeit ist, das Design und die Funktionsweise von diffraktiven Linsen näher zu untersuchen, mögliche Anwendungen zu finden und speziell das parallel scannenden Mikroskops mit diffraktiven Linsen und überlappenden Aperturen genauer zu betrachten.

1.2 Aufbau der Arbeit

Zuerst werden im Kapitel 2 und 3 die theoretischen Grundlagen dieser Arbeit sowie die verwendeten Simulationsmethoden vorgestellt. Hierbei wird bei den Grundlagen ein besonderes Augenmerk auf den Übergang von der vektoriellen elektromagnetischen Beschreibung hin zu einer skalaren Beschreibung und die zugrunde liegenden

Näherungen gelegt. Bei der Beschreibung der Simulationsmethoden liegt der Fokus auf Methoden mit diskretem Frequenzraum, da dies eine exakte Simulation von periodischen Problemen erlaubt.

Anschließend geht Kapitel 4 ausführlich auf das skalare Design von diffraktiven Linsen mit überlappender Apertur ein. Zwei Methoden zur Erzeugung des Überlapps der Aperturen und das Design spezieller Linsenarrays mit beliebigen Beleuchtungsstrukturen werden vorgestellt. Insbesondere werden dabei die Spezialfälle punktförmiger und ringförmiger Beleuchtungsstrukturen genauer untersucht. Es werden spezielle Erweiterungen des Design-Prozesses vorgestellt, die die Überlagerung mehrerer Beleuchtungsstrukturen erlauben und die Integration einer Deckglaskorrektur ermöglichen. Abschließend werden Modifikationen erläutert, welche die Herstellung der diffraktiven Linsen mit überlappenden Aperturen erleichtern.

In Kapitel 5 werden die Eigenschaften von verschiedenen Realisierungen von diffraktiven Linsen untersucht. Hierzu wird eine Fokussier-Effizienz für diffraktive Linsen mit überlappenden Aperturen definiert. Anschließend wird der Einfluss von verschiedenen Modifikationen wie einer Phasendiskretisierung oder einer Amplitudennivellierung untersucht. Des Weiteren werden die Auswirkungen der Pixelierung und der Endlichkeit des diffraktiven Elements sowie Abweichungen von den Design-Parametern untersucht. Ein Leitfaden zur Wahl der Parameter für diffraktive Linsen mit überlappenden Aperturen schließt dieses Kapitel ab.

In Kapitel 6 wird anhand zweier binärer Linsenarrays die Funktionsweise der diffraktiven Mikrolinsen mit überlappenden Aperturen experimentell verifiziert. Die numerische Apertur der Linsenarrays ist 0,48 bzw. 0,75. Dies entspricht Überlappfaktoren von ca. 50 bzw. 200. Experimentell wird zuerst die Phasenhöhe der fabrizierten diffraktiven Elemente optisch bestimmt. Anschließend werden sowohl die gemessenen lateralen als auch die axialen Fokalgeometrien mit einer skalaren Simulation der Foki verglichen.

In Kapitel 7 wird ein parallel scannendes Mikroskop als Anwendung vorgestellt, bei dem diffraktive Linsen mit überlappenden Aperturen verwendet werden. Hierbei werden zunächst typische Limitationen der klassischen Mikroskopie wie die Kopplung zwischen Auflösung, Sichtfeld und Arbeitsabstand aufgezeigt. Durch die Verwendung von diffraktiven Linsen mit überlappenden Aperturen werden die Auflösung, das Sichtfeld und der Arbeitsabstand bei einem experimentellen Aufbau entkoppelt. Die Theorie der Bildentstehung, die maximale Auflösung des Aufbaus und die verschiedenen Methoden zur Bildrekonstruktion vorgestellt. Anhand von Beispielmessungen werden die verschiedenen Rekonstruktionsmethoden verglichen und es wird die maximale Auflösung des Aufbaus experimentell bestimmt.

In Kapitel 8 werden diffraktive Linsen mit überlappenden Aperturen zur Wellenfrontmessung in einem Hartmann-Shack-Sensor behandelt.

Kapitel 9 geht auf die nicht skalare Simulation von diffraktiven Linsen mit überlappenden Aperturen ein. Hierbei werden die skalare Rechnung eines dünnen Elements mit rigorosen Rechnungen eines dicken Elements mit Hilfe der „rigorous coupled-wave analysis“ (RCWA) verglichen, um den Gültigkeitsbereich des skalaren Design-Prozesses und der skalaren Simulation abzuschätzen. Zusätzlich wird in

diesem Abschnitt noch eine Beschleunigung der RCWA durch die Sonderbehandlung von homogenen Schichten vorgestellt. Außerdem wird gezeigt, dass die RCWA eine konsistente Beschreibung der Felder liefert, sodass der Impuls in der Simulation erhalten ist.

Abschließend werden im letzten Kapitel die wichtigsten Ergebnisse der Arbeit noch einmal kurz zusammengefasst.

2 Theoretische Grundlagen

In diesem Kapitel werden die theoretischen Grundlagen und die Notation vorgestellt, die in der weiteren Arbeit verwendet werden. Einen guten Überblick über die Grundlagen der klassischen Elektrodynamik und der Optik findet man in den Büchern von Jackson [43], Born und Wolf [9], Goodman [34] sowie Novotny und Hecht [79]. Die in diesem Kapitel vorgestellten Grundlagen lehnen sich an diese Werke an.

Zuerst werden in Abschnitt 2.1 die Maxwell-Gleichungen in ihrer mikroskopischen und makroskopischen Form vorgestellt. Anschließend wird in Abschnitt 2.2 auf die besonderen Eigenschaften reeller und monofrequenter Felder eingegangen. In Abschnitt 2.3 wird der Zusammenhang zwischen den Feldern im Medium und den Materialparametern hergestellt. Anschließend wird in den Abschnitten 2.4 und 2.5 auf die Energie und Impulserhaltung eingegangen, die durch die Maxwell-Gleichungen vorgegeben sind. In den Abschnitten 2.6 und 2.7 werden die vektorielle und skalare Wellengleichung hergeleitet die die Grundlage der Lichtausbreitung im homogenen Medium bilden.

2.1 Maxwell-Gleichungen

Die Maxwell-Gleichungen bilden die Grundlage der klassischen Elektrodynamik und beschreiben die Beziehung von elektrischen und magnetischen Feldern untereinander, sowie deren Zusammenhang mit elektrischen Ladungen und elektrischen Strömen. Im Folgenden werden die Maxwell-Gleichung in ihrer mikroskopischen und makroskopischen Form als System von vier linearen partiellen Differentialgleichungen erster Ordnung vorgestellt.

2.1.1 Mikroskopische Maxwell-Gleichungen

Diskrete Ladungen ρ und die von ihnen hervorgerufenen Ströme \mathbf{j} bilden die Ursache für die magnetische Flussdichte \mathbf{B} und die elektrische Feldstärke \mathbf{E} . Die Kopplung dieser Felder untereinander und mit diskreten Ladungen und Strömen wird durch

die Maxwell-Gleichungen in mikroskopischen Form beschrieben und lauten:

$$\nabla \cdot \mathbf{E}(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t)}{\varepsilon_0} \quad (2.1)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \quad (2.2)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad (2.3)$$

$$\nabla \times \mathbf{B}(\mathbf{r}, t) = \mu_0 \mathbf{j}(\mathbf{r}, t) + \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} \quad (2.4)$$

Die Konstanten ε_0 und μ_0 geben die Permittivität und die magnetische Permeabilität des Vakuums an.

Streng genommen sind die mikroskopischen Maxwell-Gleichungen nicht exakt, sondern nur eine Näherung der allgemeineren Quantenelektrodynamik. Dies hat zur Folge, dass zum Beispiel die experimentell gemessenen Beiträge der Vakuumpolarisation zur Lamb-Verschiebung nicht mit den Maxwell-Gleichungen erklärt werden können. Im Bereich der Optik erweist sich die klassische Beschreibung jedoch als extrem erfolgreich [67].

2.1.2 Makroskopische Maxwell-Gleichungen

Für die Beschreibung makroskopischer Probleme ist ein Übergang zu kontinuierlichen Ladungsdichten und Stromdichten aufgrund der extrem hohen Teilchenzahlen nötig. Hierbei werden die Ladungs- und Stromdichtenbeiträge in einen gebundenen und einen ungebundenen (freien) Anteil aufgeteilt. Die makroskopischen Maxwell-Gleichungen beschreiben nun die Änderungen der elektrischen Flussdichte \mathbf{D} und der magnetischen Feldstärke \mathbf{H} :

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = \rho_{\text{frei}}(\mathbf{r}, t) \quad (2.5)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \quad (2.6)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad (2.7)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \mathbf{j}_{\text{frei}}(\mathbf{r}, t) + \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} \quad (2.8)$$

Maxwell-Gleichungen ohne freie Ladungen und freie Ströme

Für den wichtigen Spezialfall von Materialien ohne freie Ladungen und Ströme vereinfachen sich die Maxwell-Gleichungen folgendermaßen:

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = 0 \quad (2.9)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \quad (2.10)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad (2.11)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} \quad (2.12)$$

2.1.3 Makroskopische Maxwell-Gleichungen im Frequenzraum

Die makroskopischen Maxwell-Gleichungen aus Gleichung 2.5 - 2.8 lassen sich zeitlich in den Frequenzraum transformieren:

$$\nabla \cdot \mathbf{D}(\mathbf{r}, \omega) = \varrho_{\text{frei}}(\mathbf{r}, \omega) \quad (2.13)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, \omega) = 0 \quad (2.14)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, \omega) = +i\omega \mathbf{B}(\mathbf{r}, \omega) \quad (2.15)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, \omega) = \mathbf{j}_{\text{frei}}(\mathbf{r}, \omega) - i\omega \mathbf{D}(\mathbf{r}, \omega) \quad (2.16)$$

Hierbei gilt für die Rücktransformation der Felder in den Zeitbereich:

$$\mathbf{E}(\mathbf{r}, t) = \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega \quad (2.17)$$

$$\mathbf{D}(\mathbf{r}, t) = \int_{-\infty}^{\infty} \mathbf{D}(\mathbf{r}, \omega) e^{-i\omega t} d\omega \quad (2.18)$$

$$\mathbf{H}(\mathbf{r}, t) = \int_{-\infty}^{\infty} \mathbf{H}(\mathbf{r}, \omega) e^{-i\omega t} d\omega \quad (2.19)$$

$$\mathbf{B}(\mathbf{r}, t) = \int_{-\infty}^{\infty} \mathbf{B}(\mathbf{r}, \omega) e^{-i\omega t} d\omega \quad (2.20)$$

Neben den Feldern sind in Gleichung 2.13 und 2.16 auch die Ladungs- und Stromdichten Fourier-Transformiert worden. Hier gilt für die Rücktransformation:

$$\rho_{\text{frei}}(\mathbf{r}, t) = \int_{-\infty}^{\infty} \rho_{\text{frei}}(\mathbf{r}, \omega) e^{-i\omega t} d\omega \quad (2.21)$$

$$\mathbf{j}_{\text{frei}}(\mathbf{r}, t) = \int_{-\infty}^{\infty} \mathbf{j}_{\text{frei}}(\mathbf{r}, \omega) e^{-i\omega t} d\omega \quad (2.22)$$

Die Fourier-Transformationen der Felder sind gegeben durch:

$$\mathbf{E}(\mathbf{r}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r}, t) e^{i\omega t} dt \quad (2.23)$$

$$\mathbf{D}(\mathbf{r}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{D}(\mathbf{r}, t) e^{i\omega t} dt \quad (2.24)$$

$$\mathbf{H}(\mathbf{r}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{H}(\mathbf{r}, t) e^{i\omega t} dt \quad (2.25)$$

$$\mathbf{B}(\mathbf{r}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{B}(\mathbf{r}, t) e^{i\omega t} dt \quad (2.26)$$

Die Fourier-Transformation der Ladungs- und Stromdichten sind

$$\rho_{\text{frei}}(\mathbf{r}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \rho_{\text{frei}}(\mathbf{r}, t) e^{i\omega t} dt \quad (2.27)$$

$$\mathbf{j}_{\text{frei}}(\mathbf{r}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{j}_{\text{frei}}(\mathbf{r}, t) e^{i\omega t} dt. \quad (2.28)$$

Spezialfall: Maxwell-Gleichungen im Frequenzraum ohne freie Ladungen und freie Ströme

Im Spezialfall ohne freie Ladungen und freie Ströme vereinfachen sich die Gleichungen 2.13 bis 2.16 zu

$$\nabla \cdot \mathbf{D}(\mathbf{r}, \omega) = 0 \quad (2.29)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, \omega) = 0 \quad (2.30)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, \omega) = i\omega \mathbf{B}(\mathbf{r}, \omega) \quad (2.31)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, \omega) = -i\omega \mathbf{D}(\mathbf{r}, \omega). \quad (2.32)$$

2.2 Reellwertige und monofrequente Felder

Da das elektrische Feld $\mathbf{E}(\mathbf{r}, t)$ eine reellwertige Messgröße ist, muss Folgendes für die komplexe Amplitude der spektralen Zerlegung gelten:

$$\mathbf{E}(\mathbf{r}, -\omega) = \mathbf{E}^*(\mathbf{r}, \omega) \quad (2.33)$$

$$\mathbf{E}^*(\mathbf{r}, -\omega) = \mathbf{E}(\mathbf{r}, \omega) \quad (2.34)$$

Somit lässt sich die Rekonstruktion des elektrischen Feldes aus Gleichung 2.17 umschreiben:

$$\begin{aligned} \mathbf{E}(\mathbf{r}, t) &= \frac{1}{2} \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega + \frac{1}{2} \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega - \frac{1}{2} \int_{+\infty}^{-\infty} \mathbf{E}(\mathbf{r}, -\omega) e^{+i\omega t} d\omega \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega - \frac{1}{2} \int_{+\infty}^{-\infty} \mathbf{E}^*(\mathbf{r}, \omega) e^{+i\omega t} d\omega \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega + \frac{1}{2} \int_{-\infty}^{+\infty} \mathbf{E}^*(\mathbf{r}, \omega) e^{+i\omega t} d\omega \\ &= \int_{-\infty}^{\infty} \frac{1}{2} (\mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} + \mathbf{E}^*(\mathbf{r}, \omega) e^{+i\omega t}) d\omega \\ &= \int_{-\infty}^{\infty} \operatorname{Re}[\mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t}] d\omega. \end{aligned} \quad (2.35)$$

Analog gilt für das H-Feld:

$$\mathbf{H}(\mathbf{r}, -\omega) = \mathbf{H}^*(\mathbf{r}, \omega) \quad (2.36)$$

Somit folgt für die Rekonstruktion des reellwertigen Feldes:

$$\mathbf{H}(\mathbf{r}, t) = \int_{-\infty}^{\infty} \operatorname{Re}[\mathbf{H}(\mathbf{r}, \omega) e^{-i\omega t}] d\omega \quad (2.37)$$

Monofrequente Felder

Der Spezialfall von monochromatischen Feldern ist in der Optik von besonderem Interesse, da Laser näherungsweise monofrequente Lichtquellen sind. Für monochromatische E- und H-Felder mit Frequenz ω_0 gilt:

$$\mathcal{E}_{\text{mono}}(\mathbf{r}, \omega) = \frac{1}{2} [\mathcal{E}(\mathbf{r}, \omega) \delta(\omega - \omega_0) + \mathcal{E}^*(\mathbf{r}, -\omega) \delta(\omega + \omega_0)] \quad (2.38)$$

$$\mathcal{H}_{\text{mono}}(\mathbf{r}, \omega) = \frac{1}{2} [\mathcal{H}(\mathbf{r}, \omega) \delta(\omega - \omega_0) + \mathcal{H}^*(\mathbf{r}, -\omega) \delta(\omega + \omega_0)] \quad (2.39)$$

Man sieht sofort, dass dieses Felder die Gleichungen 2.34 und 2.36 erfüllen. Einsetzen von Gleichung 2.38 in Gleichung 2.35 rekonstruiert das reellwertige E-Feld:

$$\begin{aligned} \mathbf{E}(\mathbf{r}, t) &= \int_{-\infty}^{\infty} \text{Re} [\mathcal{E}_{\text{mono}}(\mathbf{r}, \omega) e^{-i\omega t}] d\omega \\ &= \int_{-\infty}^{\infty} \text{Re} \left[\frac{1}{2} [\mathcal{E}(\mathbf{r}, \omega) \delta(\omega - \omega_0) + \mathcal{E}^*(\mathbf{r}, -\omega) \delta(\omega + \omega_0)] e^{-i\omega t} \right] d\omega \\ &= \text{Re} \left[\frac{1}{2} [\mathcal{E}(\mathbf{r}, \omega_0) e^{-i\omega_0 t} + \mathcal{E}^*(\mathbf{r}, \omega_0) e^{i\omega_0 t}] \right] \\ &= \text{Re} \left[\frac{1}{2} [\mathcal{E}(\mathbf{r}, \omega_0) e^{-i\omega_0 t} + [\mathcal{E}(\mathbf{r}, \omega_0) e^{-i\omega_0 t}]^*] \right] \\ &= \text{Re} [\mathcal{E}(\mathbf{r}, \omega_0) e^{-i\omega_0 t}]. \end{aligned} \quad (2.40)$$

Analog folgt für die magnetische Feldstärke:

$$\begin{aligned} \mathbf{H}(\mathbf{r}, t) &= \int_{-\infty}^{\infty} \text{Re} [\mathcal{H}_{\text{mono}}(\mathbf{r}, \omega) e^{-i\omega t}] d\omega \\ &= \text{Re} [\mathcal{H}(\mathbf{r}, \omega_0) e^{-i\omega_0 t}]. \end{aligned} \quad (2.41)$$

Zeitliche Mittelung monofrequenter Felder

Man sieht in Gleichung 2.40 und 2.41, dass die Felder mit der Frequenz ω_0 oszillieren. Die Frequenz von sichtbarem Licht liegt zwischen ca. 750 THz und 375 THz. Für eine direkte Messung ist diese hohe Frequenz nicht zugänglich, stattdessen können nur zeitlich gemittelte Felder gemessen werden.

Die zeitliche Mittelung des monochromatischen elektrischen Feldes über die zeitliche Periode $T = \frac{2\pi}{\omega_0}$ ist gegeben durch

$$\langle \mathbf{E}(\mathbf{r}, t) \rangle_T = \frac{1}{T} \int_0^T \mathbf{E}(\mathbf{r}, t) dt. \quad (2.42)$$

In Anhang A.7 sind Beispiele für die Mittelung von Produkten und Ableitungen von beliebigen monochromatischen Feldern aufgelistet.

Maxwell-Gleichungen für monofrequente Felder

Für monofrequente Felder ergeben sich nun die Maxwell-Gleichungen:

$$\nabla \cdot \mathbf{D}_{\text{mono}}(\mathbf{r}, \omega) = \varrho_{\text{frei}}(\mathbf{r}, \omega) \quad (2.43)$$

$$\nabla \cdot \mathbf{B}_{\text{mono}}(\mathbf{r}, \omega) = 0 \quad (2.44)$$

$$\nabla \times \mathbf{E}_{\text{mono}}(\mathbf{r}, \omega) = +i\omega_0 \mathbf{B}_{\text{mono}}(\mathbf{r}, \omega) \quad (2.45)$$

$$\nabla \times \mathbf{H}_{\text{mono}}(\mathbf{r}, \omega) = \mathbf{j}_{\text{frei}}(\mathbf{r}, \omega) - i\omega_0 \mathbf{D}_{\text{mono}}(\mathbf{r}, \omega) \quad (2.46)$$

Man sieht die Gleichungen unterscheiden sich von den Gleichungen 2.13 - 2.16 nur in den Indices.

In der weiteren Arbeit wird oft implizit angenommen, dass es sich um monofrequente Felder handelt. Um jedoch die Formeln übersichtlich zu halten wird im Folgendem der Index „mono“ an den Feldern weggelassen und statt ω_0 wird ω verwendet.

2.3 Materialgleichungen

Über die Materialgleichungen wird ein Zusammenhang zwischen der elektrischen Flussdichte \mathbf{D} , der elektrischen Feldstärke \mathbf{E} und der elektrischen Polarisation \mathbf{P} , sowie der magnetischen Feldstärke \mathbf{H} , der magnetischen Flussdichte \mathbf{B} und der Magnetisierung \mathbf{M} hergestellt.

$$\mathbf{D}(\mathbf{r}, t) = \varepsilon_0 \mathbf{E}(\mathbf{r}, t) + \mathbf{P}(\mathbf{r}, t) \quad (2.47)$$

$$\mathbf{B}(\mathbf{r}, t) = \mu_0 \mathbf{H}(\mathbf{r}, t) + \mu_0 \mathbf{M}(\mathbf{r}, t) \quad (2.48)$$

Diese Beschreibung ist allgemein gültig.

Materialgleichungen für lineare isotrope dispersive Medien

Für lineare isotrope Medien lässt sich mit Hilfe der relativen Permittivität ε und der relativen Permeabilität μ ein Zusammenhang zwischen den Feldstärken und den Flussdichten definieren:

$$\mathbf{D}(\mathbf{r}, \omega) = \varepsilon(\mathbf{r}, \omega) \varepsilon_0 \mathbf{E}(\mathbf{r}, \omega) \quad (2.49)$$

$$\mathbf{B}(\mathbf{r}, \omega) = \mu(\mathbf{r}, \omega) \mu_0 \mathbf{H}(\mathbf{r}, \omega) \quad (2.50)$$

Im Allgemeinen ist die relative Permittivität und relative Permeabilität eine komplexe Größe.

Materialgleichungen für lineare isotrope nicht dispersive Medien

Für lineare, nicht dispersive Medien entfällt die Frequenzabhängigkeit der relativen Permittivität und der relativen Permeabilität:

$$\mathbf{D}(\mathbf{r}, \omega) = \varepsilon(\mathbf{r}) \varepsilon_0 \mathbf{E}(\mathbf{r}, \omega) \quad (2.51)$$

$$\mathbf{B}(\mathbf{r}, \omega) = \mu(\mathbf{r}) \mu_0 \mathbf{H}(\mathbf{r}, \omega) \quad (2.52)$$

Monofrequente Felder

Für monofrequente Felder folgt sowohl für die dispersiven als auch für die nicht dispersiven Medien aus vorherigem Abschnitt folgende Rekonstruktion der reellen Felder:

$$\begin{aligned}\mathbf{D}(\mathbf{r}, t) &= \operatorname{Re}[\mathcal{D}(\mathbf{r}, \omega_0)e^{-i\omega_0 t}] \\ &= \varepsilon_0 \operatorname{Re}[\varepsilon(\mathbf{r}, \omega_0)] \operatorname{Re}[\mathcal{E}(\mathbf{r}, \omega_0)e^{-i\omega_0 t}]\end{aligned}\quad (2.53)$$

$$\begin{aligned}\mathbf{B}(\mathbf{r}, t) &= \operatorname{Re}[\mathcal{B}(\mathbf{r}, \omega_0)e^{-i\omega_0 t}] \\ &= \mu_0 \operatorname{Re}[\mu(\mathbf{r}, \omega_0)] \operatorname{Re}[\mathcal{H}(\mathbf{r}, \omega_0)e^{-i\omega_0 t}]\end{aligned}\quad (2.54)$$

Dies lässt sich auch schreiben als:

$$\mathbf{D}(\mathbf{r}, t) = \varepsilon_0 \operatorname{Re}[\varepsilon(\mathbf{r}, \omega_0)] \mathbf{E}(\mathbf{r}, t) \quad (2.55)$$

$$\mathbf{B}(\mathbf{r}, t) = \mu_0 \operatorname{Re}[\mu(\mathbf{r}, \omega_0)] \mathbf{H}(\mathbf{r}, t) \quad (2.56)$$

Für die Polarisierbarkeit und die Magnetisierung folgen nun:

$$\mathbf{P}(\mathbf{r}, t) = (\operatorname{Re}[\varepsilon(\mathbf{r}, \omega_0)] - 1) \varepsilon_0 \mathbf{E}(\mathbf{r}, t) \quad (2.57)$$

$$\mathbf{M}(\mathbf{r}, t) = (\operatorname{Re}[\mu(\mathbf{r}, \omega_0)] - 1) \mathbf{H}(\mathbf{r}, t) \quad (2.58)$$

2.4 Energieerhaltung

Aus den makroskopischen Maxwell-Gleichungen lässt sich allgemein eine Gleichung für die Leistungsdichte herleiten (siehe Anhang A.9):

$$\nabla \cdot \mathbf{S} = -\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} - \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mathbf{E} \cdot \mathbf{j}_{\text{frei}}. \quad (2.59)$$

Hierbei wird \mathbf{S} als Poynting-Vektor bezeichnet und ist folgendermaßen definiert:

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} \quad (2.60)$$

Betrachtet man die Dimension und Einheit des Poynting-Vektors so sieht man, dass er eine Energie pro Zeit und Fläche, also eine Leistungsflächendichte, beschreibt:

$$[\mathbf{S}] = \frac{\text{V A}}{\text{m m}} = \frac{\text{W}}{\text{m}^2} = \frac{\text{J}}{\text{m}^2 \text{s}}$$

Multipliziert man hingegen den Poynting-Vektor mit $\varepsilon_0 \mu_0 = \frac{1}{c^2}$ so erhält man eine Impulsdichte.

Integriert man nun beide Seiten der Gleichung 2.59 über das Volumen V und wendet den Gaußschen Integralsatz auf die linke Seite der Gleichung an, so erhält man:

$$\int_{\partial V} (\mathbf{E} \times \mathbf{H}) \cdot \mathbf{n} \, dA = - \int_V \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \, dV - \int_V \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} \, dV - \int_V \mathbf{E} \cdot \mathbf{j}_{\text{frei}} \, dV \quad (2.61)$$

Hierbei ist \mathbf{n} der Normalenvektor der Oberfläche A , die zu dem Volumen V gehört. Die linke Seite der Gleichung gibt an, welche Leistungsdichte durch die Fläche $A = \partial V$ fließt. Die rechte Seite der Gleichung gibt an, welche Leistung sich im Volumen dV befindet. Der Term $\int_V \mathbf{E} \cdot \mathbf{j}_{\text{frei}} dV$ wird Joulesche Wärme genannt.

Spezialfall: Energieerhaltung für dispersionsfreie lineare Medien

Im speziellen Fall linearer, isotroper, dispersionsfreier Medien lassen sich die zeitlichen Ableitungen zusammenfassen, somit vereinfacht sich die Gleichung 2.59 zu

$$\nabla \cdot \mathbf{S} + \frac{\partial}{\partial t} W_{\text{Medium}} + \mathbf{E} \cdot \mathbf{j}_{\text{frei}} = 0. \quad (2.62)$$

Diese Gleichung wird als Poynting-Theorem bezeichnet [43]. Hierbei ist W_{Medium} die elektromagnetische Energiedichte im Medium gegeben durch:

$$W_{\text{Medium}} = \frac{1}{2} (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}) \quad (2.63)$$

$$= \frac{1}{2} (\varepsilon \varepsilon_0 E^2 + \mu \mu_0 H^2). \quad (2.64)$$

Durch anwenden des Gaußschen Integralsatzes erhält man das Poynting-Theorem in Integralform:

$$\int_{\partial V} \mathbf{S} \cdot \mathbf{n} dA = - \int_V \frac{\partial}{\partial t} W_{\text{Medium}} dV - \int_V \mathbf{E} \cdot \mathbf{j}_{\text{frei}} dV. \quad (2.65)$$

Die linke Seite der Gleichung gibt an, welche Leistungsdichte durch die Fläche $A = \partial V$ fließt. Der Term mit der zeitlichen Ableitung der Energiedichte auf der rechten Seite der Gleichung gibt an, welche Leistung im Volumen dV erzeugt oder absorbiert wird. Der Term $\int_V \mathbf{E} \cdot \mathbf{j}_{\text{frei}} dV$ wird Joulesche Wärme genannt.

Da zu jeder Zeit die erzeugte bzw. absorbierte Leistung und die Joulesche Wärme im Volumen der Leistungsdichte entspricht, die durch die Oberfläche fließt, folgt aus der Leistungserhaltung eine Energieerhaltung.

Spezialfall: Poynting Vektor einer ebenen Welle

Für eine ebene Welle der Form $\mathbf{E} = \mathbf{E}(\mathbf{r} \cdot \mathbf{s} - vt)$ und $\mathbf{H} = \mathbf{H}(\mathbf{r} \cdot \mathbf{s} - vt)$ mit Richtungsvektor \mathbf{s} und Geschwindigkeit v ergibt sich für den Poyntingvektor [9]:

$$\mathbf{S} = \frac{c}{\sqrt{\varepsilon \mu}} W_{\text{Medium}} \mathbf{s} \quad (2.66)$$

$$= \frac{c}{\sqrt{\varepsilon \mu}} \varepsilon |E|^2 \mathbf{s} \quad (2.67)$$

$$= \frac{c}{\sqrt{\varepsilon \mu}} \mu |H|^2 \mathbf{s} \quad (2.68)$$

Die Leistungsdichte ist proportional zur Energiedichte im Medium, die Richtung wird durch den Richtungsvektor der ebenen Welle bestimmt.

Intensität Im allgemeinen Fall wird angenommen, dass der Betrag des Poyntingvektors das Maß für die messbare Lichtintensität ist [9]. Für die ebene Welle folgt damit, dass die Intensität proportional zum Betragsquadrat des E-Felds ist:

$$I = |\mathbf{S}| = \frac{c}{\sqrt{\varepsilon\mu}} \varepsilon |E|^2 \propto \sqrt{\varepsilon} |E|^2. \quad (2.69)$$

Spezialfall: Energieerhaltung für lineare dispersionsfreie Materialien und monofrequente Felder

Für monofrequente Felder lässt sich die Energieerhaltung in linearen dispersionsfreien Materialien im zeitlichen Mittel betrachten:

$$\left\langle \int_{\partial V} \mathbf{S} \cdot \mathbf{n} \, dA \right\rangle_T = \left\langle - \int_V \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \, dV - \int_V \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} \, dV - \int_V \mathbf{E} \cdot \mathbf{j}_{\text{frei}} \, dV \right\rangle_T \quad (2.70)$$

Es ist zu beachten, dass die Zeit-Mittelung in den Feldern \mathbf{E} , \mathbf{D} , \mathbf{H} und \mathbf{B} durchgeführt wird. Es ergibt sich:

$$\begin{aligned} \int_{\partial V} \frac{1}{2} \operatorname{Re}[\mathcal{E}(\mathbf{r}, \omega) \times \mathcal{H}^*(\mathbf{r}, \omega)] \cdot \mathbf{n} \, dA &= - \frac{1}{2} \int_V \operatorname{Re}[i\omega \mathcal{E}(\mathbf{r}, \omega) \cdot \mathcal{D}^*(\mathbf{r}, \omega)] \, dV \\ &\quad - \frac{1}{2} \int_V \operatorname{Re}[i\omega \mathcal{H}(\mathbf{r}, \omega) \cdot \mathcal{B}^*(\mathbf{r}, \omega)] \, dV \\ &\quad - \frac{1}{2} \int_V \operatorname{Re}[\mathcal{E}(\mathbf{r}, \omega) \cdot \mathbf{j}^*(\mathbf{r}, \omega)] \, dV \end{aligned} \quad (2.71)$$

Nun lassen sich die Materialgleichungen für die linearen dispersionsfreien Medien substituieren, dann erhält man

$$\begin{aligned} \int_{\partial V} \frac{1}{2} \operatorname{Re}[\mathcal{E}(\mathbf{r}, \omega) \times \mathcal{H}^*(\mathbf{r}, \omega)] \cdot \mathbf{n} \, dA &= - \frac{1}{2} \int_V \omega \operatorname{Im}[\varepsilon(\mathbf{r})] \varepsilon_0 \operatorname{Re}[\mathcal{E}(\mathbf{r}, \omega) \cdot \mathcal{E}^*(\mathbf{r}, \omega)] \, dV \\ &\quad - \frac{1}{2} \int_V \omega \operatorname{Im}[\mu(\mathbf{r})] \mu_0 \operatorname{Re}[\mathcal{H}(\mathbf{r}, \omega) \cdot \mathcal{H}^*(\mathbf{r}, \omega)] \, dV \\ &\quad - \frac{1}{2} \int_V \operatorname{Re}[\mathcal{E}(\mathbf{r}, \omega) \cdot \mathbf{j}^*(\mathbf{r}, \omega)] \, dV \end{aligned} \quad (2.72)$$

Mit $\omega \varepsilon_0 = \frac{k_0}{Z_0}$ und $\omega \mu_0 = k_0 Z_0$ lässt sich dies weiter umschreiben. Des Weiteren gilt $\operatorname{Re}[\mathcal{H}(\mathbf{r}, \omega) \cdot \mathcal{H}^*(\mathbf{r}, \omega)] = |\mathcal{H}(\mathbf{r}, \omega)|^2$ und $\operatorname{Re}[\mathcal{E}(\mathbf{r}, \omega) \cdot \mathcal{E}^*(\mathbf{r}, \omega)] = |\mathcal{E}(\mathbf{r}, \omega)|^2$, somit folgt:

$$\begin{aligned} \int_{\partial V} \frac{1}{2} \operatorname{Re}[\mathcal{E}(\mathbf{r}, \omega) \times \mathcal{H}^*(\mathbf{r}, \omega)] \cdot \mathbf{n} \, dA &= - \frac{1}{2} \int_V \frac{k_0}{Z_0} \operatorname{Im}[\varepsilon(\mathbf{r})] |\mathcal{E}(\mathbf{r}, \omega)|^2 \, dV \\ &\quad - \frac{1}{2} \int_V k_0 Z_0 \operatorname{Im}[\mu(\mathbf{r})] |\mathcal{H}(\mathbf{r}, \omega)|^2 \, dV \\ &\quad - \frac{1}{2} \int_V \operatorname{Re}[\mathcal{E}(\mathbf{r}, \omega) \cdot \mathbf{j}^*(\mathbf{r}, \omega)] \, dV. \end{aligned} \quad (2.73)$$

Anwendung der Energieerhaltung: Modellvorstellung einer Kameramessung

Ein Kamerapixel lässt sich in einer idealisiert Modellvorstellung als homogenes stark absorbierendes Medium betrachten. Wendet man nun die Energieerhaltung auf das Kamerapixel an, so muss der Energiefluss durch die Pixelflächen betrachtet werden. Idealisiert lässt sich annehmen, dass nur Energie durch die Frontfläche des Pixels fließt. Da das Medium stark absorbiert gibt es keinen Energiefluss durch die Seitenflächen und durch die hintere Fläche des Pixels. Zur weiteren Vereinfachung wird für das Kamerapixel angenommen, dass der Imaginärteil der Permeabilität null ist und es keine freien Ströme gibt. Dann ergibt sich für die Energieerhaltung:

$$\left\langle \int_{\text{Frontfläche}} \mathbf{S} \cdot \hat{\mathbf{e}}_z dA \right\rangle_T = -\frac{1}{2} \int_{\text{Pixel}} \frac{k_0}{Z_0} \text{Im}[\varepsilon(\mathbf{r})] |\mathcal{E}(\mathbf{r}, \omega)|^2 dV \quad (2.74)$$

Der Energiefluss ist proportional zum Absolutbetrag der Feldstärke. Die Kamera übersetzt die absorbierte Energie in einen Intensitätswert, somit folgt:

$$\langle I(\mathbf{r}) \rangle_T \propto |\mathcal{E}(\mathbf{r}, \omega)|^2 \quad (2.75)$$

Die gemessene Intensität des Kamerapixels ist also proportional zum Absolutbetrag der Feldstärke.

2.5 Impulserhaltung

Aus den Maxwell-Gleichungen lässt sich folgende Gleichung für die Kraftdichte herleiten (siehe Anhang A.10):

$$\begin{aligned} \nabla \cdot \left(\varepsilon_0 \mathbf{E} \mathbf{E} - \varepsilon_0 \frac{1}{2} E^2 \mathbf{I} + \mu_0 \mathbf{H} \mathbf{H} - \mu_0 \frac{1}{2} H^2 \mathbf{I} \right) = \\ + \mathbf{j}_{\text{frei}} \times \mu_0 \mathbf{H} + \varepsilon_0 \mathbf{E} \nabla \cdot \mathbf{E} + \mu_0 \mathbf{H} \nabla \cdot \mathbf{H} - \frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} + \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} \end{aligned} \quad (2.76)$$

Hierbei lässt sich der Maxwellsche Spannungstensor \mathbf{T} zum Vereinfachen benutzen, dieser lässt sich folgendermaßen definieren [43]:

$$\mathbf{T} = \left(\varepsilon_0 \mathbf{E} \mathbf{E} - \varepsilon_0 \frac{1}{2} E^2 \mathbf{I} + \mu_0 \mathbf{H} \mathbf{H} - \mu_0 \frac{1}{2} H^2 \mathbf{I} \right) \quad (2.77)$$

Man beachte, dass der Maxwellsche Spannungstensor \mathbf{T} nicht explizit von den Materialparametern abhängt. Die Definition des äußeren Produkts $\mathbf{E} \mathbf{E}$ sowie die Definition der Divergenz eines Tensors ist in Anhand A.4 gezeigt. Die rechte Seite der Gleichung 2.76 entspricht der elektromagnetischen Kraftdichte \mathbf{f} und der zeitlichen Ableitung der Impulsdichte \mathbf{G} nach Chu [55], die expliziten Definitionen sind in Anhang A.10.1 gezeigt. Mit diesen Ersetzungen erhält man

$$\nabla \cdot \mathbf{T} = \mathbf{f} + \frac{\partial}{\partial t} \mathbf{G}. \quad (2.78)$$

Da diese Kraftdichte-Beziehung zu allen Zeiten und für alle Volumina immer erfüllt ist, ist der Impuls erhalten.

Die Gleichungen 2.76 und 2.78 sind allgemein gültig, jedoch ist es möglich die einzelnen Terme in Gleichung 2.76 anders zu gruppieren, sodass man alternative Definitionen für den Maxwell'schen Spannungstensors, die Kraftdichte und die Impulsdichte erhält. Die unterschiedlichen Definitionen begründen sich zum einen in der Definition der Impulsdichte im Medium, dieses Problem ist auch als Abraham-Minkowski-Dilemma bekannt. Zum anderen unterscheiden sich die einzelnen Formulierungen dadurch, welche Modellvorstellung zur Beschreibung der Ladungsträger im Medium benutzt wird. Dadurch erhält man unterschiedliche Beschreibungen der Kraftdichte [54]. Die Beschreibung nach Chu [55] geht zum Beispiel davon aus, dass die Abraham-Impulsdichte die richtige Beschreibung ist. Eine alternative Beschreibung, die die Minkowski-Impulsdichte als richtig annimmt, wird im Anhang kurz vorgestellt. Welche Modellvorstellung zur Beschreibung der makroskopischen Kraftdichte richtig ist, ist trotz zahlreicher Experimente immer noch nicht abschließend geklärt [36, 55]. Im Folgenden wird die Beschreibung nach Chu verwendet.

Um die elektromagnetische Kraft \mathbf{F}_{EM} auszurechnen, die auf ein Volumen wirkt, lässt sich nun entweder die Kraftdichte über das Volumen integrieren oder man integriert über die Beiträge des Tensors und der zeitlichen Ableitung der Impulsdichte.

$$\mathbf{F}_{\text{EM}} = \int_V \mathbf{f} \, dV \quad (2.79)$$

$$= \int_V \nabla \cdot \mathbf{T} \, dV - \int_V \frac{\partial}{\partial t} \mathbf{G} \, dV \quad (2.80)$$

Für zeitlich harmonische Felder liefert der Beitrag über die zeitliche Ableitung der Impulsdichte im Mittel keinen Kraftbeitrag. Somit ist in diesem Fall die zeitlich gemittelte Kraft über das Integral der Divergenz des Maxwell'schen Spannungstensors bestimmbar.

Impulserhaltung ohne freie Ladungen und freie Ströme

Für den Fall ohne freie Ladungen und ohne freie Ströme vereinfacht sich die Gleichung 2.76 und es ergibt sich

$$\nabla \cdot \mathbf{T} = + \varepsilon_0 \mathbf{E} \nabla \cdot \mathbf{E} + \mu_0 \mathbf{H} \nabla \cdot \mathbf{H} - \frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} + \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H}. \quad (2.81)$$

Nun lässt sich über ein Volumen V integrieren und der Gauß'sche Integralsatz anwenden. Man muss jedoch beachten, dass das Vektorfeld \mathbf{T} Unstetigkeiten besitzt. Der Gauß'sche Integralsatz für Vektorfelder mit einer Unstetigkeit entspricht dem normalen Gauß'schen Integralsatz mit einem zusätzlichen Beitrag, der die Unstetigkeit

berücksichtigt [56], dies ist im Anhang A.8 gezeigt. Man erhält:

$$\begin{aligned} & \int_{\partial V} \mathbf{T} \cdot \mathbf{n} \, dA - \int_{\Sigma} (\mathbf{T}^+ - \mathbf{T}^-) \cdot \mathbf{n}_{\Sigma} \, dA = \\ & \int_V \varepsilon_0 \mathbf{E} \nabla \cdot \mathbf{E} + \mu_0 \mathbf{H} \nabla \cdot \mathbf{H} - \frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} + \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} \, dV \end{aligned} \quad (2.82)$$

Hierbei bedeutet \mathbf{T}^- , dass das Vektorfeld \mathbf{T} an der linken Seite der Sprungstelle Σ ausgewertet wird, und \mathbf{T}^+ , dass das Vektorfeld \mathbf{T} an der rechten Seite der Sprungstelle Σ ausgewertet wird. \mathbf{n} und \mathbf{n}_{Σ} sind die Einheitsnormalen der Grenzflächen bzw. der Sprungflächen.

Das erste Oberflächenintegral erstreckt sich über den äußeren Rand ∂V des Rechengebiets. Vom äußeren Rand werden die Oberflächenintegrale über die Unstetigkeitsflächen Σ abgezogen. Dies beinhaltet implizit Sprungstellen am Anfang und am Ende, sodass man das Innere des Rechengebiets erhält. Das Volumenintegral V bezieht sich jetzt auf das Innere des Rechengebiets.

Impulserhaltung für lineare dispersionsfreie Materialien und monofrequente Felder

Für monofrequente Felder lässt sich in linearen dispersionsfreien Medien eine zeitlich gemittelte Kraft aus der Kraftdichte berechnen.

$$\begin{aligned} & \left\langle \int_{\partial V} \mathbf{T} \cdot \mathbf{n} \, dA \right\rangle_T - \left\langle \int_{\Sigma} (\mathbf{T}^+ - \mathbf{T}^-) \cdot \mathbf{n}_{\Sigma} \, dA \right\rangle_T = \\ & \left\langle \int_V \varepsilon_0 \mathbf{E} \nabla \cdot \mathbf{E} \, dV \right\rangle_T + \left\langle \int_V \mu_0 \mathbf{H} \nabla \cdot \mathbf{H} \, dV \right\rangle_T \\ & - \left\langle \int_V \frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} \, dV \right\rangle_T + \left\langle \int_V \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} \, dV \right\rangle_T \end{aligned} \quad (2.83)$$

Sei $\mathbf{n}_{\Sigma} = \hat{\mathbf{e}}_z$ der Normalenvektor einer Grenzfläche, dann ist zu beachten, dass die z-Komponenten des E- und des H-Feldes unstetig an dieser Grenzfläche sind. Somit schreibt sich der Beitrag über die Sprungstelle als

$$\begin{aligned} & - \left\langle \int_{\Sigma} (\mathbf{T}^+ - \mathbf{T}^-) \cdot \mathbf{n}_{\Sigma} \, dA \right\rangle_T = - \left\langle \int_{\Sigma} \begin{pmatrix} T_{xz}^+ - T_{xz}^- \\ T_{yz}^+ - T_{yz}^- \\ T_{zz}^+ - T_{zz}^- \end{pmatrix} dA \right\rangle_T \\ & = - \frac{1}{2} \int_{\Sigma} \operatorname{Re} \left[\varepsilon_0 \begin{pmatrix} \mathcal{E}_x^+ \mathcal{E}_z^{*+} - \mathcal{E}_x^- \mathcal{E}_z^{*-} \\ \mathcal{E}_y^+ \mathcal{E}_z^{*+} - \mathcal{E}_y^- \mathcal{E}_z^{*-} \\ \frac{1}{2} (\mathcal{E}_z^+ \mathcal{E}_z^{*+} - \mathcal{E}_z^- \mathcal{E}_z^{*-}) \end{pmatrix} + \mu_0 \begin{pmatrix} \mathcal{H}_x^+ \mathcal{H}_z^{*+} - \mathcal{H}_x^- \mathcal{H}_z^{*-} \\ \mathcal{H}_y^+ \mathcal{H}_z^{*+} - \mathcal{H}_y^- \mathcal{H}_z^{*-} \\ \frac{1}{2} (\mathcal{H}_z^+ \mathcal{H}_z^{*+} - \mathcal{H}_z^- \mathcal{H}_z^{*-}) \end{pmatrix} \right] dA. \end{aligned} \quad (2.84)$$

Sei $\mathbf{n} = \hat{\mathbf{e}}_z$ dann gilt

$$\left\langle \int_{\partial V} \mathbf{T} \cdot \mathbf{n} dA \right\rangle_T = \frac{1}{2} \int_{\partial V} \text{Re} \left[\varepsilon_0 \begin{pmatrix} \mathcal{E}_x \mathcal{E}_z^* \\ \mathcal{E}_y \mathcal{E}_z^* \\ \frac{1}{2} (\mathcal{E}_z \mathcal{E}_z^* - \mathcal{E}_x \mathcal{E}_x^* - \mathcal{E}_y \mathcal{E}_y^*) \end{pmatrix} + \mu_0 \begin{pmatrix} \mathcal{H}_x \mathcal{H}_z^* \\ \mathcal{H}_y \mathcal{H}_z^* \\ \frac{1}{2} (\mathcal{H}_z \mathcal{H}_z^* - \mathcal{H}_x \mathcal{H}_x^* - \mathcal{H}_y \mathcal{H}_y^*) \end{pmatrix} \right] dA. \quad (2.85)$$

Für die anderen Terme gilt:

$$\left\langle \int_V \varepsilon_0 \mathbf{E} \nabla \cdot \mathbf{E} dV \right\rangle_T = \frac{1}{2} \varepsilon_0 \int_V \text{Re}[(\nabla \cdot \mathcal{E}(\mathbf{r}, \omega)) \mathcal{E}^*(\mathbf{r}, \omega)] dV \quad (2.86)$$

$$+ \left\langle \int_V \mu_0 \mathbf{H} \nabla \cdot \mathbf{H} dV \right\rangle_T = + \frac{1}{2} \mu_0 \int_V \text{Re}[(\nabla \cdot \mathcal{H}(\mathbf{r}, \omega)) \mathcal{H}^*(\mathbf{r}, \omega)] dV \quad (2.87)$$

$$\begin{aligned} - \left\langle \int_V \frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} dV \right\rangle_T &= + \frac{1}{2} \varepsilon_0 \int_V \text{Re}[i\omega \mathcal{E}(\mathbf{r}, \omega) \times \mathcal{B}^*(\mathbf{r}, \omega)] dV \\ &= + \frac{1}{2} \omega \mu_0 \varepsilon_0 \int_V \text{Re}[i\mu^*(\mathbf{r}, \omega) \mathcal{E}(\mathbf{r}, \omega) \times \mathcal{H}^*(\mathbf{r}, \omega)] dV \end{aligned} \quad (2.88)$$

$$\begin{aligned} + \left\langle \int_V \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} dV \right\rangle_T &= - \frac{1}{2} \mu_0 \int_V \text{Re}[i\omega \mathcal{H}(\mathbf{r}, \omega) \times \mathcal{D}^*(\mathbf{r}, \omega)] dV \\ &= - \frac{1}{2} \omega \mu_0 \varepsilon_0 \int_V \text{Re}[i\varepsilon^*(\mathbf{r}, \omega) \mathcal{H}(\mathbf{r}, \omega) \times \mathcal{E}^*(\mathbf{r}, \omega)] dV \\ &= - \frac{1}{2} \omega \mu_0 \varepsilon_0 \int_V \text{Re}[-i\varepsilon(\mathbf{r}, \omega) \mathcal{H}^*(\mathbf{r}, \omega) \times \mathcal{E}(\mathbf{r}, \omega)] dV \\ &= - \frac{1}{2} \omega \mu_0 \varepsilon_0 \int_V \text{Re}[i\varepsilon(\mathbf{r}, \omega) \mathcal{E}(\mathbf{r}, \omega) \times \mathcal{H}^*(\mathbf{r}, \omega)] dV \end{aligned} \quad (2.89)$$

Die letzten beiden Terme lassen sich zusammenfassen:

$$\begin{aligned} - \left\langle \int_V \frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} dV \right\rangle_T + \left\langle \int_V \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} dV \right\rangle_T &= \\ + \frac{1}{2} \omega \mu_0 \varepsilon_0 \int_V \text{Re}[i(\mu^*(\mathbf{r}, \omega) - \varepsilon(\mathbf{r}, \omega)) \mathcal{E}(\mathbf{r}, \omega) \times \mathcal{H}^*(\mathbf{r}, \omega)] dV & \end{aligned} \quad (2.90)$$

Mit $\mu_0 = Z_0^2 \varepsilon_0$ und $\mu_0 \omega = Z_0 k_0$ folgt für die Impulserhaltung:

$$\begin{aligned}
& \frac{1}{2} \varepsilon_0 \int_{\partial V} \operatorname{Re} \left[\begin{pmatrix} \mathcal{E}_x \mathcal{E}_z^* \\ \mathcal{E}_y \mathcal{E}_z^* \\ \frac{1}{2} (\mathcal{E}_z \mathcal{E}_z^* - \mathcal{E}_x \mathcal{E}_x^* - \mathcal{E}_y \mathcal{E}_y^*) \end{pmatrix} + Z_0^2 \begin{pmatrix} \mathcal{H}_x \mathcal{H}_z^* \\ \mathcal{H}_y \mathcal{H}_z^* \\ \frac{1}{2} (\mathcal{H}_z \mathcal{H}_z^* - \mathcal{H}_x \mathcal{H}_x^* - \mathcal{H}_y \mathcal{H}_y^*) \end{pmatrix} \right] dA \\
& - \frac{1}{2} \varepsilon_0 \int_{\Sigma} \operatorname{Re} \left[\begin{pmatrix} \mathcal{E}_x^+ \mathcal{E}_z^{*+} - \mathcal{E}_x^- \mathcal{E}_z^{*-} \\ \mathcal{E}_y^+ \mathcal{E}_z^{*+} - \mathcal{E}_y^- \mathcal{E}_z^{*-} \\ \frac{1}{2} (\mathcal{E}_z^+ \mathcal{E}_z^{*+} - \mathcal{E}_z^- \mathcal{E}_z^{*-}) \end{pmatrix} + Z_0^2 \begin{pmatrix} +\mathcal{H}_x^+ \mathcal{H}_z^{*+} - \mathcal{H}_x^- \mathcal{H}_z^{*-} \\ +\mathcal{H}_y^+ \mathcal{H}_z^{*+} - \mathcal{H}_y^- \mathcal{H}_z^{*-} \\ \frac{1}{2} (+\mathcal{H}_z^+ \mathcal{H}_z^{*+} - \mathcal{H}_z^- \mathcal{H}_z^{*-}) \end{pmatrix} \right] dA = \\
& + \frac{1}{2} \varepsilon_0 \int_V \operatorname{Re} [(\nabla \cdot \mathbf{E}(\mathbf{r}, \omega)) \mathbf{E}^*(\mathbf{r}, \omega)] dV \\
& + \frac{1}{2} Z_0^2 \varepsilon_0 \int_V \operatorname{Re} [(\nabla \cdot \mathbf{H}(\mathbf{r}, \omega)) \mathbf{H}^*(\mathbf{r}, \omega)] dV \\
& + \frac{1}{2} Z_0 k_0 \varepsilon_0 \int_V \operatorname{Re} [i(\mu^*(\mathbf{r}, \omega) - \varepsilon(\mathbf{r}, \omega)) \mathbf{E}(\mathbf{r}, \omega) \times \mathbf{H}^*(\mathbf{r}, \omega)] dV \tag{2.91}
\end{aligned}$$

2.6 Vektorielle Wellenoptik

2.6.1 Vektorielle Wellengleichung

Aus den Maxwell-Gleichungen erhält man für lineare isotrope Medien ohne freie Ladungen und freie Ströme und für homogene Permeabilitäten die vektorielle Wellengleichung für inhomogene Medien:

$$+\nabla \left(\mathbf{E}(\mathbf{r}, t) \cdot \frac{\nabla \varepsilon}{\varepsilon} \right) + \nabla^2 \mathbf{E}(\mathbf{r}, t) - \mu_0 \mu \varepsilon \varepsilon_0 \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} = \mathbf{0} \tag{2.92}$$

Die Herleitung ist in Anhang A.11 gezeigt. Eine alternative Formulierung der Wellengleichung erhält man durch Verwendung der Identität $\mathbf{E}(\mathbf{r}, t) \cdot \frac{\nabla \varepsilon}{\varepsilon} = \frac{\rho_{\text{frei}}(\mathbf{r}, t)}{\varepsilon} - \nabla \cdot \mathbf{E}(\mathbf{r}, t)$. Für $\rho_{\text{frei}} = 0$ ergibt sich für die vektorielle Wellengleichung

$$-\nabla(\nabla \cdot \mathbf{E}(\mathbf{r}, t)) + \nabla^2 \mathbf{E}(\mathbf{r}, t) - \mu_0 \mu \varepsilon \varepsilon_0 \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} = \mathbf{0} \tag{2.93}$$

Außerdem lässt sich der Brechungsindex $n = \sqrt{\varepsilon \mu}$ zur weiteren Vereinfachung benutzen. Die genaue Beziehung zwischen den Materialparametern ε und μ und dem Brechungsindex n wird in Anhang A.12 gezeigt. Mit dem Brechungsindex erhält man nun:

$$+2\nabla(\mathbf{E}(\mathbf{r}, t) \cdot \nabla \log n(\mathbf{r}, t)) + \nabla^2 \mathbf{E}(\mathbf{r}, t) - \frac{n^2}{c^2} \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} = \mathbf{0} \tag{2.94}$$

Man sieht, dass die einzelnen Komponenten des E-Feld untereinander gekoppelt sind, wenn die Brechungsindexverteilung nicht homogen ist. Dies ist zum Beispiel für dicke Gitter der Fall.

Mit der Näherung $2\nabla(\mathbf{E}(\mathbf{r}, t) \cdot \nabla \log n(\mathbf{r}, t)) \approx \mathbf{0}$ entkoppeln die einzelnen Komponenten des E-Felds und es ergibt sich die vektorielle Wellengleichung:

$$+\nabla^2 \mathbf{E}(\mathbf{r}, t) - \frac{n^2}{c^2} \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} = \mathbf{0}. \quad (2.95)$$

Die angewendete Näherung ist für ein homogenes Medium exakt, somit sind die einzelnen Feldkomponenten im homogenen Medium entkoppelt. Alternativ ist die vektorielle Wellengleichung auch exakt wenn $\nabla(\nabla \cdot \mathbf{E}) = \mathbf{0}$ und $\rho_{\text{frei}} = 0$ gilt.

Spezialfall: Helmholtz-Gleichung

Für dispersionsfreie Medien und zeitharmonische Felder vereinfacht sich die vektorielle Wellengleichung zur vektoriellen Helmholtz-Gleichung:

$$\nabla^2 \mathcal{E}(\mathbf{r}, \omega) + \frac{n^2 \omega^2}{c^2} \mathcal{E}(\mathbf{r}, \omega) = \mathbf{0} \quad (2.96)$$

Zur weiteren Vereinfachungen wird die Wellenzahl k_0 im Vakuum und die Wellenzahl k in einem Medium mit Brechungsindex n folgendermaßen definiert:

$$k_0 = \frac{\omega}{c} \quad (2.97)$$

$$k = nk_0. \quad (2.98)$$

Einsetzen der Definition liefert die vektorielle Helmholtz-Gleichung in ihrer kürzesten Form:

$$\nabla^2 \mathcal{E}(\mathbf{r}, \omega) + k^2 \mathcal{E}(\mathbf{r}, \omega) = \mathbf{0} \quad (2.99)$$

Im Allgemeinen lässt sich die Helmholtz-Gleichung in nur elf Koordinatensystemen durch eine Separation der Variablen analytisch lösen [38].

2.7 Skalare Wellenoptik

Im vorherigen Abschnitt wurde gezeigt, dass wenn die Gradienten der Brechungsindexverteilung klein sind, die einzelnen Komponenten des elektrischen Felds unabhängig voneinander sind. Für die einzelnen Komponenten erhält man dann drei unabhängige skalare Gleichungen der gleichen Form, somit bietet sich eine skalare Beschreibung an. Der Übergang von der vektoriellen Beschreibung zur skalaren Beschreibung erfordert außerdem, dass die Begriffe Intensität, Energiedichte und Energiefluss für ein skalares Feld neu definiert werden müssen [105].

2.7.1 Skalare Wellengleichung

Die Grundlage der skalaren Wellenoptik ist die skalare Wellengleichung. Anstatt der Gleichung für die drei Komponenten des elektrischen Feldes \mathbf{E} betrachtet man nur

eine Gleichung für das skalare Wellenfeld V . Dies ist valide, da die einzelnen Komponenten unter bestimmten Umständen entkoppeln. Man erhält folgende skalare Wellengleichung:

$$\nabla^2 V(\mathbf{r}, t) - \frac{n^2(\mathbf{r}, t)}{c^2} \frac{\partial^2 V(\mathbf{r}, t)}{\partial t^2} = 0 \quad (2.100)$$

Für nicht dispersive Medien lässt sich die Gleichung in den Frequenzraum transformieren:

$$\nabla^2 U(\mathbf{r}, \omega) + \frac{\omega^2 n^2}{c^2} U(\mathbf{r}, \omega) = 0. \quad (2.101)$$

Für die Transformation des skalaren Felds zwischen Zeit und Frequenzraum gilt:

$$V(\mathbf{r}, t) = \int_{-\infty}^{\infty} U(\mathbf{r}, \omega) e^{-i\omega t} d\omega \quad (2.102)$$

$$U(\mathbf{r}, \omega) = \frac{1}{2\pi} V(\mathbf{r}, t) \int_{-\infty}^{\infty} e^{+i\omega t} dt \quad (2.103)$$

Dimension des skalaren Wellenfelds Die skalare Wellengleichung bestimmt nicht die Dimension des skalaren Felds V oder U . Oft wird in skalaren Beugungsproblemen die skalare Amplitude dimensionslos gewählt, sodass gilt:

$$[U(\mathbf{r}, \omega)] = 1 \quad (2.104)$$

Insbesondere bei der Beschreibung von speziellen Wellen wie der Kugelwelle oder der Sommerfeld-Welle muss auf die Skalierung der skalaren Amplitude geachtet werden, hierzu lässt sich die einheitenbehaftete Skalierungskonstante U_0 einführen. Die skalare Amplitude der Kugelwelle ist dann zum Beispiel gegeben durch

$$U(\mathbf{r}, \omega) = U_0 \frac{e^{ik_0 r}}{r}. \quad (2.105)$$

Die Einheit von U_0 wird so gewählt, dass die skalare Amplitude $U(\mathbf{r}, \omega)$ einheitenlos ist. In obigen Beispiel ist die Einheit der Skalierungskonstanten gleich der Einheit des Abstands r zu wählen, damit die skalare Amplitude dimensionslos wird, somit hat U_0 die Dimension einer Länge die zum Beispiel in Metern gemessen werden kann, also ist $[U_0] = [r] = \text{m}$ zu wählen.

2.7.2 Monochromatische skalare zeitharmonische Felder

Analog zu den Vektorfeldern lässt sich ein skalares monochromatisches Feld erzeugen:

$$U_{\text{mono}}(\mathbf{r}, \omega) = \frac{1}{2} (U(\mathbf{r}, \omega_0) \delta(\omega - \omega_0) + U^*(\mathbf{r}, -\omega_0) \delta(\omega + \omega_0)) \quad (2.106)$$

Setzt man dies in Gleichung 2.102 ein erhält man:

$$\begin{aligned}
 V(\mathbf{r}, t) &= \int_{-\infty}^{\infty} U_{\text{mono}}(\mathbf{r}, \omega) d\omega \\
 &= \frac{1}{2} \left(U(\mathbf{r}, \omega_0) e^{-i\omega_0 t} + U^*(\mathbf{r}, \omega_0) e^{+i\omega_0 t} \right) \\
 &= \text{Re} \left(U(\mathbf{r}, \omega_0) e^{-i\omega_0 t} \right).
 \end{aligned} \tag{2.107}$$

Man sieht, dass das skalare Feld im Orts- und Zeitraum wieder eine reelle Größe ist.

Oft wird ein monochromatisches Feld über ein komplexes analytisches Signal beschrieben, das das monochromatische Feld repräsentiert:

$$U_{\text{c. a.}}(\mathbf{r}, \omega) = U(\mathbf{r}, \omega) \delta(\omega - \omega_0) \tag{2.108}$$

Für das Feld im Orts- und Zeitraum ergibt sich

$$\begin{aligned}
 V_{\text{c. a.}}(\mathbf{r}, t) &= \int_0^{\infty} U(\mathbf{r}, \omega) \delta(\omega - \omega_0) e^{-i\omega t} d\omega \\
 &= U(\mathbf{r}, \omega_0) e^{-i\omega_0 t}
 \end{aligned} \tag{2.109}$$

Um hieraus wieder das reelle monochromatische Feld zu erhalten, muss der Realteil gebildet werden.

2.7.3 Intensität eines skalaren Feldes

Die (instantane) Intensität eines skalaren Feldes ist definiert durch [34]

$$I(\mathbf{r}, t) = V(\mathbf{r}, t) V^*(\mathbf{r}, t). \tag{2.110}$$

Für monofrequente Felder ist die zeitlich gemittelte Intensität gegeben durch

$$\langle I(\mathbf{r}, t) \rangle_T = U(\mathbf{r}, \omega) U^*(\mathbf{r}, \omega). \tag{2.111}$$

Für komplexe analytische monofrequente Felder hingegen ist die Intensität nicht mehr eine Funktion der Zeit:

$$I(\mathbf{r}, t) = U(\mathbf{r}, \omega) U^*(\mathbf{r}, \omega). \tag{2.112}$$

Die integrierte Intensität über das gesamte Wellenfeld in einer Ebene ist gegeben durch

$$I_{\text{ges}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\mathbf{r}, \omega) U^*(\mathbf{r}, \omega) d\mathbf{r}_{\perp} \tag{2.113}$$

2.7.4 Energiedichte und Energieflussdichte eines skalaren Feldes

Für skalare Felder lässt sich eine skalare Energiedichte $H(\mathbf{r},t)$ und eine skalare Energieflussdichte $\mathbf{F}(\mathbf{r},t)$ definieren [35, 67]:

$$H(\mathbf{r},t) = \alpha \left(\frac{1}{c^2} \frac{\partial V^*}{\partial t} \frac{\partial V}{\partial t} + \nabla V^* \cdot \nabla V \right) \quad (2.114)$$

$$\mathbf{F}(\mathbf{r},t) = -\alpha \left(\frac{\partial V^*}{\partial t} \nabla V + \frac{\partial V}{\partial t} \nabla V^* \right). \quad (2.115)$$

Hierbei ist α eine Skalierungskonstante die die Einheit der skalaren Energiedichte und Energieflussdichte definiert. Es gilt die skalare Energieerhaltung:

$$\frac{\partial H(\mathbf{r},t)}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{r},t) = 0. \quad (2.116)$$

Jedoch gilt es zu beachten, dass $\mathbf{F}(\mathbf{r},t)$ nicht eindeutig bestimmt ist. Addiert man zur Energieflussdichte die Rotation eines beliebigen Vektorfelds \mathbf{f} , dann erfüllt $\mathbf{F}'(\mathbf{r},t) = \mathbf{F}(\mathbf{r},t) + \nabla \times \mathbf{f}(\mathbf{r},t)$ auch die Gleichung für die skalare Energieerhaltung.

Für komplexe analytische monofrequente skalare Felder folgt:

$$H(\mathbf{r},t) = \alpha \left(\frac{\omega^2}{c^2} U^* U + \nabla U^* \cdot \nabla U \right) \quad (2.117)$$

$$\mathbf{F}(\mathbf{r},t) = -i\omega\alpha(U^* \nabla U - U \nabla U^*) \quad (2.118)$$

Die Energiedichte und die Energieflussdichte sind keine Funktion der Zeit mehr somit vereinfacht sich die Energieerhaltung zu:

$$\nabla \cdot \mathbf{F} = 0. \quad (2.119)$$

Für monofrequente Felder hingegen fällt die Zeitabhängigkeit erst durch die Zeitmittelung weg [105]:

$$\langle \mathbf{F}(\mathbf{r},t) \rangle_T = -i\omega\alpha(U^* \nabla U - U \nabla U^*) \quad (2.120)$$

Im Allgemeinen lässt sich die komplexe Amplitude U in eine reelle Amplitude und eine reelle Phase aufteilen:

$$U = A e^{i\phi} \quad (2.121)$$

Somit gilt für die skalare zeitlich gemittelte Energieflussdichte

$$\langle \mathbf{F}(\mathbf{r},t) \rangle_T = \omega\alpha |A(\mathbf{r})|^2 \nabla \phi(\mathbf{r}). \quad (2.122)$$

Einheit der Skalierungskonstanten Die Einheit der Energiedichte H in SI-Einheiten ist J m^{-3} , die Einheit der Energieflussdichte \mathbf{F} in SI-Einheiten ist $\text{J m}^{-2} \text{s}^{-1}$. Hieraus folgt für die Einheit der Skalierungskonstanten α in Abhängigkeit der Dimension der skalaren Amplitude U :

$$[\alpha] = \begin{cases} \text{J m}^{-1} & \text{falls } [U] = 1 \\ \text{J m}^3 & \text{falls } [U] = \text{m}^{-2} \\ 1 & \text{falls } [U] = \sqrt{\text{J m}^{-1}} \end{cases} \quad (2.123)$$

Totaler Energiefluss

Im Gegensatz zur Energieflussdichte ist das Flächenintegral über die Energieflussdichte eindeutig bestimmt. Diese Leistung die durch eine Ebene mit Normalenvektor $\mathbf{n} = \mathbf{e}_z$ fließt wird totaler Energiefluss genannt und ist gegeben durch:

$$F_z^{\text{total}} = \int_A \mathbf{F} \cdot \mathbf{e}_z \, dA \quad (2.124)$$

$$= -\alpha \int_A \left(\frac{\partial V^*}{\partial t} \frac{\partial V}{\partial z} + \frac{\partial V}{\partial t} \frac{\partial V^*}{\partial z} \right) dA \quad (2.125)$$

Die Einheit des totalen Energieflusses ist $[F_z^{\text{total}}] = \text{J s}^{-1}$. Für monofrequente skalare Felder folgt:

$$F_z^{\text{total}} = -i\alpha\omega \int_A \left(U^* \frac{\partial U}{\partial z} - U \frac{\partial U^*}{\partial z} \right) dA \quad (2.126)$$

Einsetzen der Zerlegung nach ebenen Wellen erlaubt es die z -Ableitung auszuführen:

$$F_z^{\text{total}} = 2\alpha\omega \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{U}_0^* \tilde{U}_0 k_z \frac{d\mathbf{k}_\perp}{4\pi^2} \quad (2.127)$$

Das Ergebnis ist unabhängig von z , somit ist der der totale Energiefluss in einem homogenen nicht absorbierenden Medium erhalten.

Spezialfall: Paraxiale Felder Für paraxiale Felder gilt $k_z \approx k_0$. Somit ist die integrierte Intensität proportional zum totalen Energiefluss.

$$I_{\text{ges}} \propto F_z^{\text{total}} \quad (2.128)$$

Um eine Gleichheit zu bekommen, muss die Skalierungskonstante $\alpha = \frac{1}{2\omega k_0} = \frac{1}{2ck_0^2}$ gewählt werden.

3 Simulationsmethoden

In diesem Kapitel werden die Simulationsmethoden vorgestellt, die relevant für die folgende Arbeit sind. Zuerst werden die Lösungsmethoden für die skalare Wellengleichung vorgestellt, anschließend wird die vektorielle Erweiterung betrachtet und zuletzt wird eine Lösungsmethode für die Maxwell-Gleichungen vorgestellt.

3.1 Lösungsmethoden für die skalare Wellengleichung

Die Lösungsmethoden für die skalare Wellengleichung lassen sich in analytisch exakte Lösungen und genäherte Lösungen aufteilen. Insbesondere für die Betrachtung von paraxialen Problemen bieten genäherte Lösungen den Vorteil, dass sie einfache analytische Beschreibungen liefern.

3.1.1 Exakte Methoden zum Lösen der skalaren Wellengleichung

Die skalare Wellengleichung lässt sich im Vakuum und in homogenen Medien exakt lösen. Von besonderer Bedeutung sind die Methoden, bei denen die skalare Amplitude in einer Ebene bekannt ist. Daraus lässt sich die skalare Amplitude im gesamten Raum berechnen. Im Ortsraum geschieht dies durch das Rayleigh-Sommerfeld Integral I, im Frequenzraum nutzt man die Zerlegung nach ebenen Wellen. Mathematisch gesehen sind beide Lösungen äquivalent [67].

Rayleigh-Sommerfeld Integral I

Ist das skalare Feld U in der Ebene $z' = 0$ bekannt, so lässt sich das Feld in jedem Punkt \mathbf{r} im Raum über das Rayleigh-Sommerfeld Integral I berechnen:

$$U(\mathbf{r}) = -\frac{1}{2\pi} \int_{A'} U(\mathbf{r}'_{\perp}, 0) \frac{d}{dz} \left(\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right) d^2\mathbf{r}'_{\perp}. \quad (3.1)$$

Dabei wird über die Apertur A' in der Ebene $z' = 0$ integriert.

Zerlegung nach ebenen Wellen

Ist das skalare Feld U in der Ebene $z = 0$ bekannt, lässt sich das Feld in jedem Punkt \mathbf{r} im Raum auch über die Zerlegung nach ebenen Wellen berechnen:

$$U(\mathbf{r}) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{U}(\mathbf{k}_{\perp}, 0) e^{iz\sqrt{n^2k_0^2 - \mathbf{k}_{\perp} \cdot \mathbf{k}_{\perp}}} e^{i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} \frac{d^2\mathbf{k}_{\perp}}{4\pi^2} \quad (3.2)$$

Hierbei ist $\tilde{U}(\mathbf{k}_\perp, 0)$ die Fourier-Transformierte in der Ebene $z = 0$ und es gilt

$$\tilde{U}(\mathbf{k}_\perp, 0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U(\mathbf{r}_\perp, 0) e^{-i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} d^2\mathbf{r}_\perp \quad (3.3)$$

$$U(\mathbf{r}_\perp, 0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{U}(\mathbf{k}_\perp, 0) e^{i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} \frac{d^2\mathbf{k}_\perp}{4\pi^2}. \quad (3.4)$$

Zerlegung nach ebenen Wellen in Zylinderkoordinaten Für analytische Betrachtungen ist es sinnvoll, die Zerlegung nach ebenen Wellen in Zylinderkoordinaten zu transformieren:

$$U(r_\perp, \varphi, z) = \int_0^\infty \int_0^{2\pi} \tilde{U}(k_\perp, \chi, 0) e^{iz\sqrt{n^2k_0^2 - k_\perp^2}} e^{-ik_\perp r_\perp \cos(\varphi - \chi)} d\chi k_\perp \frac{dk_\perp}{4\pi^2}. \quad (3.5)$$

Die Definition der Zylinderkoordinaten im Orts- und Frequenzraum ist in Anhang A.1 gezeigt. Die Fourier-Transformierte in der Ebene $z = 0$ in Zylinderkoordinaten ist

$$\tilde{U}(k_\perp, \chi, 0) = \int_0^\infty \int_0^{2\pi} U(r_\perp, \varphi, 0) e^{-ik_\perp r_\perp \cos(\varphi - \chi)} d\varphi r_\perp dr_\perp. \quad (3.6)$$

Lässt sich die Funktion $U(r_\perp, \varphi)$ in Polarkoordinaten separieren, sodass $U(r_\perp, \varphi) = U_r(r_\perp)U_\varphi(\varphi)$ ist, dann gilt [34]:

$$\tilde{U}(k_\perp, \chi, 0) = \sum_{m=-\infty}^{\infty} c_m (-i)^m e^{im\chi} \mathcal{H}_m(U_r(r_\perp)) \quad (3.7)$$

$$c_m = \frac{1}{2\pi} \int_0^{2\pi} U_\varphi(\varphi) e^{-im\varphi} d\varphi \quad (3.8)$$

$$\mathcal{H}_m(U_r(r_\perp)) = 2\pi \int_0^\infty U_r(r_\perp) J_m(k_\perp r_\perp) r_\perp dr_\perp. \quad (3.9)$$

Hierbei ist \mathcal{H}_m der Hankel-Transformations-Operator der m -ten Ordnung.

Für den Spezialfall von rotationssymmetrischen Spektren lässt die Zerlegung nach ebenen Wellen weiter vereinfachen:

$$U(r_\perp, z) = \int_0^\infty \tilde{U}(k_\perp, 0) e^{iz\sqrt{n^2k_0^2 - k_\perp^2}} J_0(k_\perp r_\perp) k_\perp \frac{dk_\perp}{2\pi}. \quad (3.10)$$

Hierbei ist $\tilde{U}(k_\perp, 0)$ die Fourier-Bessel-Transformierte und es gilt:

$$\tilde{U}(k_\perp, 0) = 2\pi \int_0^\infty U(r_\perp, 0) J_0(k_\perp r_\perp) r_\perp dr_\perp \quad (3.11)$$

$$U(r_\perp, 0) = \int_0^\infty \tilde{U}(k_\perp, 0) J_0(k_\perp r_\perp) k_\perp \frac{dk_\perp}{2\pi}. \quad (3.12)$$

Rechnerische Implementierung der Lichtausbreitung durch Zerlegung nach ebenen Wellen

Beschränkt man sich in der Simulation auf den Spezialfall von bandbegrenzten periodischen Amplitudenverteilungen in der Anfangsebene, dann ist das dazugehörige Spektrum $\tilde{U}(\mathbf{k}_\perp)$ diskret und besteht aus einer endlichen Anzahl diskreter Beiträge. Jedem Punkt im Frequenzraum lässt sich eine Mode \tilde{U}_{lm} zuordnen:

$$\tilde{U}_{lm} = \tilde{U}(l \delta k_x, m \delta k_y) \quad (3.13)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{U}(\mathbf{k}_\perp) \delta(k_x - l \delta k_x, k_y - m \delta k_y) d\mathbf{k}_\perp. \quad (3.14)$$

Dabei sind $dk_x = \frac{2\pi}{P_x}$ und $dk_y = \frac{2\pi}{P_y}$ die Abtastschrittweiten im Frequenzraum, die durch die Periodenlängen P_x und P_y bestimmt sind.

Bei der numerischen Beschreibung der komplexen Amplitude ist es notwendig, dass die komplexe Amplitude räumlich diskretisiert wird. Die Amplitude U_{pq} lässt sich einem bestimmten Abtastpunkt zuordnen:

$$U_{pq} = U_{\text{diskret}}(p \, dx, q \, dy) \quad (3.15)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\mathbf{r}_\perp) \delta(x - p \, dx, y - q \, dy) d\mathbf{r}_\perp. \quad (3.16)$$

Einsetzen der Definition der Fouriertransformation aus Gleichung 3.3 und Ausführen der Integration über die Delta-Funktion liefert

$$U_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{U}(\mathbf{k}_\perp) e^{i(k_x p \, dx + k_y q \, dy)} \frac{d^2 \mathbf{k}_\perp}{4\pi^2}. \quad (3.17)$$

Da das Spektrum diskret ist, lässt sich das Integral als Riemann-Summe über die einzelnen Beiträge auffassen:

$$U_{pq} = \sum_{l,m} U_{lm} e^{i(l \delta k_x p \, dx + m \delta k_y q \, dy)} \frac{dk_x dk_y}{4\pi^2}. \quad (3.18)$$

Wählt man nun N_x bzw. N_y äquidistante Abtastpunkte mit der Schrittweite $dx = \frac{P_x}{N_x}$ bzw. $dy = \frac{P_y}{N_y}$ und setzt die Abtastschrittweite im Frequenzraum $dk_x = \frac{2\pi}{P_x}$ und $dk_y = \frac{2\pi}{P_y}$ ein, so erhält man

$$U_{pq} = \frac{1}{P_x P_y} \sum_{l,m} \tilde{U}_{lm} e^{2\pi i \left(\frac{lp}{N_x} + \frac{mq}{N_y} \right)}. \quad (3.19)$$

Die abgetastete Amplitude lässt sich also über eine diskrete Fouriertransformation des diskreten Spektrums bestimmen. Umgekehrt lässt sich durch eine inverse diskrete Fourier-Transformation das diskrete Spektrum aus der abgetasteten komplexen Amplitude bestimmen.

Setzt man in Gleichung 3.16 die Zerlegung nach ebenen Wellen aus Gleichung 3.2 ein, so erhält man analog zur obigen Betrachtung die komplexe Amplitude in einer ausgebreiteten z -Ebene:

$$U_{pq}(z) = \frac{1}{P_x P_y} \sum_{l,m} \tilde{U}_{lm}(z=0) e^{2\pi i \left(\frac{lp}{N_x} + \frac{mq}{N_y} \right)} e^{i \sqrt{n^2 k_0^2 - (l dk_x)^2 - (m dk_y)^2} z}. \quad (3.20)$$

Periodische und bandbegrenzte Amplitudenverteilungen lassen sich demnach ohne Näherungen in eine andere Ebene ausbreiten. Deshalb ist die Zerlegung nach ebenen Wellen eine exakte Methode zur Ausbreitung der komplexen Amplitude.

3.1.2 Genäherte Methoden zum Lösen der skalaren Wellengleichung

Im Allgemeinen kann man zwei Arten von Näherungstypen unterscheiden: Die erste Art von Näherung befasst sich mit der Ausbreitung im homogenen Medium und betrachtet Spezialfälle der exakten Ausbreitungsvorschriften, die zweite Art von Näherungen beschreibt die Ausbreitung in inhomogenen Medien.

Genäherte Methoden zur Lichtausbreitung in homogenen Medien

Für die analytische Berechnung der Lichtausbreitung in homogenen Medien lassen sich das Rayleigh-Sommerfeld Integral I und die Zerlegung nach ebenen Wellen für bestimmte Spezialfälle nähern. Im Folgenden werden die wichtigsten Spezialfälle vorgestellt.

Fresnel-Integral im Ortsraum Das Fresnel-Integral im Ortsraum lässt sich aus dem Rayleigh-Sommerfeld Integral I nähern. Man erhält

$$U(\mathbf{r}) = -\frac{i}{\lambda z} e^{ikz} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\mathbf{r}'_{\perp}, 0) e^{ik \frac{(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp})^2}{2z}} d^2 \mathbf{r}'_{\perp} \quad (3.21)$$

$$= -\frac{i}{\lambda z} e^{ikz} e^{ik \frac{r_{\perp}^2}{2z}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\mathbf{r}'_{\perp}, 0) e^{ik \frac{r'_{\perp}{}^2}{2z}} e^{-ik \frac{\mathbf{r}_{\perp} \cdot \mathbf{r}'_{\perp}}{z}} d^2 \mathbf{r}'_{\perp}. \quad (3.22)$$

Sowohl die Phase als auch die Amplitude sind genähert. Die Näherung in der Phase ist

$$\sqrt{(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp})^2 + z^2} = z \sqrt{1 - \frac{(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp})^2}{z^2}} \quad (3.23)$$

$$\approx z + \frac{(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp})^2}{2z} \quad (3.24)$$

und die Näherung in der Amplitude

$$\frac{z}{(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp})^2 + z^2} \left(ik - \frac{1}{\sqrt{(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp})^2 + z^2}} \right) \approx \frac{ik}{z}. \quad (3.25)$$

Diese Näherungen sind gültig, wenn $(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2 \ll z^2$ gilt.

Das Fresnel-Integral lässt sich als Fourier-Transformation und zwei Multiplikationen mit jeweils einer quadratischen Phase auffassen, deshalb wird diese Ausbreitungsmethode manchmal auch als QFQ-Methode bezeichnet [15].

High-NA-Approximation Für die High-NA-Approximation lässt sich das Rayleigh-Sommerfeld-Integral I nähern zu [27]

$$U(\mathbf{r}) = -\frac{iz}{\lambda r^2} e^{ikr} e^{-ik\frac{r_\perp^2}{2r}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\mathbf{r}'_\perp, 0) e^{ik\frac{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2}{2r}} d^2\mathbf{r}'_\perp \quad (3.26)$$

$$= -\frac{iz}{\lambda r^2} e^{ikr} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\mathbf{r}'_\perp, 0) e^{ik\frac{r'^2}{2r}} e^{-ik\frac{\mathbf{r}_\perp \cdot \mathbf{r}'_\perp}{r}} d^2\mathbf{r}'_\perp. \quad (3.27)$$

Im Gegensatz zum Fresnel-Integral im Ortsraum wird hier die Strecke r in der Näherung anstelle der Propagationsdistanz z ausgeklammert. Die Näherung in der Phase ist somit

$$\sqrt{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2 + z^2} = r^2 \sqrt{1 + \frac{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2}{r^2} - \frac{r_\perp^2}{r^2}} \quad (3.28)$$

$$\approx r + \frac{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2}{2r} - \frac{r_\perp^2}{2r} \quad (3.29)$$

$$= r + \frac{\mathbf{r}_\perp \cdot \mathbf{r}'_\perp}{r} + \frac{r'^2}{2r} \quad (3.30)$$

und in der Amplitude

$$\frac{z}{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2 + z^2} \left(ik - \frac{1}{\sqrt{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2 + z^2}} \right) \approx \frac{ik}{r}. \quad (3.31)$$

Diese Näherungen sind somit gültig, wenn $r'^2 - 2\mathbf{r}_\perp \cdot \mathbf{r}'_\perp \ll r^2$ ist. Hiermit lässt sich die Ausbreitung einer in der Aperturebene stark in der Nähe der optischen Achse lokalisierten Ausgangsamplitude berechnen. Oft wird in der numerischen Rechnung im quadratischen Phasenfaktor die Distanz r mit z genähert, damit sich das Integral als Fourier-Transformation schreiben lässt.

Fraunhofer-Integral Für den Spezialfall sehr großer Ausbreitungsdistanzen vereinfacht sich das Rayleigh-Sommerfeld Integral I zum Fraunhofer-Integral

$$U(\mathbf{r}) = -\frac{i}{\lambda z} e^{ikz} e^{ik\frac{r_\perp^2}{2z}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\mathbf{r}'_\perp, 0) e^{-ik\frac{\mathbf{r}_\perp \cdot \mathbf{r}'_\perp}{z}} dx' dy' \quad (3.32)$$

Hierbei ist die Näherung in der Phase:

$$\sqrt{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2 + z^2} = z \sqrt{1 - \frac{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2}{z^2}} \quad (3.33)$$

$$\approx z - \frac{\mathbf{r}_\perp \cdot \mathbf{r}'_\perp}{z} + \frac{r_\perp^2}{2z} \quad (3.34)$$

und die Näherung in der Amplitude

$$\frac{z}{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2 + z^2} \left(ik - \frac{1}{\sqrt{(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2 + z^2}} \right) \approx \frac{ik}{z}. \quad (3.35)$$

Die Näherung ist gültig, wenn $(\mathbf{r}_\perp - \mathbf{r}'_\perp)^2 \ll z^2$ und $r'_\perp \ll z$ gilt, also die Apertur viel kleiner als die Ausbreitungsdistanz ist und der Beobachtungspunkt \mathbf{r}_\perp nahe an der optischen Achse liegt. In diesem Fall entspricht die ausgebreitete Amplitude der Fourier-Transformierten der Verteilung in der Aperturebene.

Zerlegung nach ebenen Wellen in der Fresnel-Näherung Besteht das Spektrum $\tilde{U}(\mathbf{k}_\perp)$ nur aus niederfrequenten Beiträgen, für die gilt $k_\perp^2 \ll n^2 k_0^2$, so lässt sich die Wurzel im Propagator der Zerlegung nach ebenen Wellen nähern

$$\sqrt{n^2 k_0^2 - k_\perp^2} = nk_0 \sqrt{1 - \frac{k_\perp^2}{n^2 k_0^2}} \approx nk_0 - \frac{k_\perp^2}{2nk_0}. \quad (3.36)$$

Für die Zerlegung nach ebenen Wellen in der Fresnel-Näherung erhält man

$$U(\mathbf{r}) = e^{ink_0 z} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{U}(\mathbf{k}_\perp, 0) e^{iz \frac{k_\perp^2}{2nk_0}} e^{i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} \frac{d^2 \mathbf{k}_\perp}{4\pi^2}. \quad (3.37)$$

Mathematisch ist die Zerlegung nach ebenen Wellen in der Fresnel-Näherung äquivalent mit dem Fresnel-Integral im Ortsraum [15].

Stationäre Phasen Näherung Das Integral in der Zerlegung nach ebenen Wellen lässt sich mit der Methode der stationären Phase nähern [9]. Diese Näherung eignet sich besonders gut für die Untersuchung von fokussierenden Feldern [10]. Für den Fall einer idealen konvergierenden Kugelwelle erhält man so das Debye-Integral [23]:

$$U(\vec{r}) = -\frac{i}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_0 \frac{1}{k_z} e^{i\mathbf{k} \cdot \mathbf{r}} dk_x dk_y. \quad (3.38)$$

Dabei ist U_0 wieder der einheitenbehaftete Skalierungsfaktor.

Genäherte Methoden zur Lichtausbreitung in inhomogenen Medien

Die Lichtausbreitung durch optische Komponenten wie Linsen, Prismen, Gitter oder Glasfasern lässt sich als Lichtausbreitung durch ein inhomogenes Medium auffassen. Die Wirkung von Linsen, Prismen und Gittern kann näherungsweise durch die Wirkung eines dünnen Elements beschrieben werden. Für Glasfasern oder allgemein bei optischen Komponenten, bei denen die Dicke nicht vernachlässigbar ist, bietet sich die Beam-Propagation-Method (BPM) oder die Wave-Propagation-Method (WPM) als näherungsweise Lösungsansatz für die Ausbreitung an. Im Folgenden werden die oben genannten Methoden einführend vorgestellt.

Dünne Element Näherung Die Propagation einer skalaren Amplitude durch eine optische Komponente lässt sich in der dünnen Element Näherung als eine Multiplikation mit einer komplexen Transferfunktion $t(\mathbf{r}_\perp)$ schreiben.

$$U'(\mathbf{r}_\perp) = t(\mathbf{r}_\perp)U(\mathbf{r}_\perp) \quad (3.39)$$

Dabei ist $U(\mathbf{r}_\perp)$ die komplexe Amplitude vor und $U'(\mathbf{r}_\perp)$ die komplexe Amplitude nach der optischen Komponente.

Beam Propagation Method Die BPM wurde von Feit und Fleck [29] entwickelt um Gradientenindex-Glasfasern zu simulieren. Die Methode teilt die Propagation in viele kleine Teilpropagation der Schrittweite dz auf. In jedem Teilschritt erfolgt eine Ausbreitung durch ein homogenes effektives Medium \bar{n} und anschließend eine Multiplikation mit einem dünnen Element, das die Brechungsindexabweichung zum homogenen Medium berücksichtigt:

$$U(\mathbf{r}_\perp, dz) = e^{izk_0\Delta n(\mathbf{r}_\perp)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{U}(\mathbf{k}_\perp, 0) e^{idz\sqrt{\bar{n}^2k_0^2 - \mathbf{k}_\perp \cdot \mathbf{k}_\perp}} e^{i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} \frac{d\mathbf{k}_\perp}{4\pi^2} \quad (3.40)$$

Da sich numerisch die Ausbreitung im gemittelten Medium über eine FFT beschreiben lässt, wird diese Methode oft auch als FFT-BPM bezeichnet [106].

Wave Propagation Method Die WPM von Brenner und Singer [17] vermeidet die Aufteilung in ein homogenes Medium und eine Phasenkorrektur durch einen ad-hoc-Ansatz für das um dz ausgebreitete Feld:

$$U(\mathbf{r}_\perp, dz) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{U}(\mathbf{k}_\perp, 0) e^{idz\sqrt{n^2(\mathbf{r}_\perp)k_0^2 - \mathbf{k}_\perp \cdot \mathbf{k}_\perp}} e^{i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} \frac{d\mathbf{k}_\perp}{4\pi^2}. \quad (3.41)$$

3.2 Lösungsmethoden für die vektorielle Wellengleichung

Die Lösungsmethoden für die skalare Wellengleichung werden in diesem Abschnitt zu Lösungsmethoden für die vektorielle Wellengleichung erweitert. Damit diese vektoriellen Lösungen jedoch physikalisch sind, müssen sie weitere Randbedingungen (vorgegeben durch die Maxwell-Gleichungen) erfüllen, sodass die einzelnen Feldkomponenten untereinander gekoppelt sind. Deshalb wird zuerst die Polarisation eingeführt und gezeigt, dass transversal polarisierte ebene Wellen eine physikalische Lösung der vektoriellen Wellengleichung sind. Anschließend werden wichtige vektorielle Lichtausbreitungsmethoden vorgestellt.

3.2.1 Transversal polarisierte ebene Welle als Lösung der Wellengleichung

Die vektorielle Wellengleichung aus Gleichung 2.99 lässt sich in drei unabhängige skalare Wellengleichungen separieren. Mögliche Lösungen für die drei skalaren Gleichungen sind folgende ebene Wellen:

$$\mathcal{E}_x(\mathbf{r}) = \mathcal{E}_{x,0} e^{ink_0 \mathbf{s} \cdot \mathbf{r}} \quad (3.42)$$

$$\mathcal{E}_y(\mathbf{r}) = \mathcal{E}_{y,0} e^{ink_0 \mathbf{s} \cdot \mathbf{r}} \quad (3.43)$$

$$\mathcal{E}_z(\mathbf{r}) = \mathcal{E}_{z,0} e^{ink_0 \mathbf{s} \cdot \mathbf{r}} . \quad (3.44)$$

Hierbei ist \mathbf{s} der auf Eins normierte Richtungsvektor der Ausbreitung

$$\mathbf{s} = \begin{pmatrix} s_x \\ s_y \\ s_z \end{pmatrix} = \begin{pmatrix} \sin(\theta) \cos(\varphi) \\ \sin(\theta) \sin(\varphi) \\ \cos(\theta) \end{pmatrix} \quad (3.45)$$

mit dem Azimutwinkel φ und dem Polarwinkel θ . Für die Lösung der vektoriellen Wellengleichung sind die Amplituden $\mathcal{E}_{x,0}$, $\mathcal{E}_{y,0}$ und $\mathcal{E}_{z,0}$ unabhängig. Soll jedoch diese Lösung gleichzeitig eine Lösung der Maxwell-Gleichung sein, so muss Folgendes gelten:

$$\mathcal{E}(\mathbf{r}) = \mathcal{P} \mathcal{E}_0 e^{ink_0 \mathbf{s} \cdot \mathbf{r}} , \quad (3.46)$$

wobei der Vektor \mathcal{P} die Polarisation beschreibt und sich in zwei Anteile aufteilen lässt:

$$\mathcal{P} = \mathcal{P}_{\text{TE}} \hat{\mathbf{e}}_{\text{TE}} + \mathcal{P}_{\text{TM}} \hat{\mathbf{e}}_{\text{TM}} . \quad (3.47)$$

Die Einheitsvektoren sind folgendermaßen definiert:

$$\hat{\mathbf{e}}_{\text{TE}} = \frac{\hat{\mathbf{e}}_z \times \mathbf{s}}{|\hat{\mathbf{e}}_z \times \mathbf{s}|} = \frac{1}{s_\perp} \begin{pmatrix} -s_y \\ s_x \\ 0 \end{pmatrix} = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{pmatrix} \quad (3.48)$$

$$\hat{\mathbf{e}}_{\text{TM}} = \frac{\hat{\mathbf{e}}_s \times \mathbf{s}}{|\hat{\mathbf{e}}_s \times \mathbf{s}|} = \frac{1}{s_\perp} \begin{pmatrix} s_x s_z \\ s_y s_z \\ -s_\perp^2 \end{pmatrix} = \begin{pmatrix} \cos(\varphi) \cos(\theta) \\ \sin(\varphi) \cos(\theta) \\ -\sin(\theta) \end{pmatrix} . \quad (3.49)$$

Die Einheitsvektoren \mathbf{s} , $\hat{\mathbf{e}}_{\text{TE}}$ und $\hat{\mathbf{e}}_{\text{TM}}$ bilden ein vollständiges Orthonormalsystem. Man sieht, dass die Welle in Gleichung 3.46 keinen Beitrag der Polarisation in \mathbf{s} -Richtung hat, sondern nur Beiträge senkrecht zur Ausbreitungsrichtung. Deshalb nennt man diese Art von Welle eine transversale elektromagnetische Welle. Die Eigenschaft der Transversalität ist eine direkte Konsequenz aus Gleichung 2.29, der Divergenzgleichung für die elektrische Flussdichte.

Aufgrund der Linearität der Wellengleichung lässt sich eine Überlagerung von transversal polarisierten ebenen Wellen erzeugen, die wieder eine physikalische Lösung der vektoriellen Wellengleichung ist. Andererseits lässt sich ein beliebiges Wellenfeld \mathcal{E} aufgrund der Orthogonalität in TE- und TM-Anteile zerlegen.

3.2.2 Exakte Methoden zur Lösung der vektoriellen Wellengleichung

Analog zum skalaren Fall lässt sich die vektorielle Wellengleichung sowohl im Ortsraum, als auch im Frequenzraum für homogene Medien exakt lösen. Im Frequenzraum erhält man die vektorielle Zerlegung nach ebenen Wellen, im Ortsraum das vektorielle Rayleigh-Sommerfeld Integral I.

Vektorielle Zerlegung nach ebenen Wellen

Ist das elektrische Feld in einer Ebene bekannt, so lässt sich das Feld im ganzen Raum berechnen. Die Ausbreitung mit der vektoriellen Zerlegung nach ebenen Wellen ist gegeben durch

$$\mathcal{E}(\mathbf{r}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\mathcal{E}}(\mathbf{k}_{\perp}, 0) e^{iz\sqrt{n^2 k_0^2 - \mathbf{k}_{\perp} \cdot \mathbf{k}_{\perp}}} e^{i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} \frac{d^2 \mathbf{k}_{\perp}}{4\pi^2}. \quad (3.50)$$

Hierbei ist $\tilde{\mathcal{E}}(\mathbf{k}_{\perp}, 0)$ die zweidimensionale Fouriertransformierte von $\mathcal{E}(\mathbf{r}_{\perp}, 0)$ und es gilt:

$$\tilde{\mathcal{E}}(\mathbf{k}_{\perp}, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{E}(\mathbf{r}_{\perp}, 0) e^{-i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} d\mathbf{r}_{\perp} \quad (3.51)$$

$$\mathcal{E}(\mathbf{r}_{\perp}, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\mathcal{E}}(\mathbf{k}_{\perp}, 0) e^{i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} \frac{d\mathbf{k}_{\perp}}{4\pi^2}. \quad (3.52)$$

Für physikalische Wellen ist es ausreichend, wenn nur die x- und y-Komponenten des elektrischen Felds in der Ausgangsebene bekannt sind. Die z-Komponente lässt sich aus der Divergenz-Gleichung der Maxwell-Gleichung 2.29 bestimmen. Somit erhält man für die transversalen Komponenten und die z-Komponente folgende Ausbreitungsvorschriften:

$$\mathcal{E}_{\perp}(\mathbf{r}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\mathcal{E}}_{\perp}(\mathbf{k}_{\perp}, 0) e^{iz\sqrt{n^2 k_0^2 - \mathbf{k}_{\perp} \cdot \mathbf{k}_{\perp}}} e^{i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} \frac{d^2 \mathbf{k}_{\perp}}{4\pi^2} \quad (3.53)$$

$$\mathcal{E}_z(\mathbf{r}) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{k_x}{k_z} \tilde{\mathcal{E}}_x(\mathbf{k}_{\perp}, 0) + \frac{k_y}{k_z} \tilde{\mathcal{E}}_y(\mathbf{k}_{\perp}, 0) \right) e^{iz\sqrt{n^2 k_0^2 - \mathbf{k}_{\perp} \cdot \mathbf{k}_{\perp}}} e^{i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} \frac{d^2 \mathbf{k}_{\perp}}{4\pi^2}. \quad (3.54)$$

Vektorielles Rayleigh-Sommerfeld Integral I

Wendet man die Weyl-Äquivalenz auf die vektorielle Zerlegung nach ebenen Wellen an, so erhält man das vektorielle Rayleigh-Sommerfeld Integral I [62]:

$$\mathcal{E}(\mathbf{r}) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{E}(\mathbf{r}') \frac{\partial}{\partial z} \left(\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right) d\mathbf{r}'_{\perp}. \quad (3.55)$$

Analog zur vektoriellen Zerlegung nach ebenen Wellen reicht es, in der Ausgangsebene nur die x- und y-Komponente des Feldes zu kennen. Somit ergeben sich für die ausgebreiteten Komponenten folgende Integrale:

$$\mathcal{E}_\perp(\mathbf{r}) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{E}_\perp(\mathbf{r}') \frac{\partial}{\partial z} \left(\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right) d\mathbf{r}'_\perp \quad (3.56)$$

$$\mathcal{E}_z(\mathbf{r}) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\mathcal{E}_x(\mathbf{r}') \frac{\partial}{\partial x} \left(\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right) + \mathcal{E}_y(\mathbf{r}') \frac{\partial}{\partial y} \left(\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right) \right) d\mathbf{r}'_\perp. \quad (3.57)$$

Bestimmung der zugehörigen H-Felder

Die beiden vorgestellten exakten Methoden beschreiben nur die Ausbreitung der elektrischen Feldstärke \mathcal{E} . Die dazugehörige magnetische Feldstärke \mathcal{H} lässt sich im homogenen Raum mit Hilfe der Maxwell-Gleichungen im Frequenzraum bestimmen. Für die Fourier-Koeffizienten der magnetischen Feldstärke folgt:

$$\tilde{\mathcal{H}}(\mathbf{k}_\perp, z, \omega) = \frac{1}{\omega\mu\mu_0} \begin{pmatrix} k_x \\ k_y \\ -i\frac{\partial}{\partial z} \end{pmatrix} \times \tilde{\mathcal{E}}(\mathbf{k}_\perp, z, \omega) \quad (3.58)$$

$$= \frac{1}{\omega\mu\mu_0} \begin{pmatrix} k_x \\ k_y \\ -i\frac{\partial}{\partial z} \end{pmatrix} \times \tilde{\mathcal{E}}(\mathbf{k}_\perp, 0) e^{iz\sqrt{n^2k_0^2 - \mathbf{k}_\perp \cdot \mathbf{k}_\perp}} \quad (3.59)$$

$$= \frac{1}{\omega\mu\mu_0} \begin{pmatrix} k_x \\ k_y \\ \sqrt{n^2k_0^2 - \mathbf{k}_\perp \cdot \mathbf{k}_\perp} \end{pmatrix} \times \tilde{\mathcal{E}}(\mathbf{k}_\perp, 0) e^{iz\sqrt{n^2k_0^2 - \mathbf{k}_\perp \cdot \mathbf{k}_\perp}}. \quad (3.60)$$

Die Rücktransformation in den Ortsraum liefert nun die zugehörige magnetische Feldstärke:

$$\mathcal{H}(\mathbf{r}, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\mathcal{H}}(\mathbf{k}_\perp, z, \omega) e^{i\mathbf{k}_\perp \cdot \mathbf{r}} \frac{d\mathbf{k}_\perp}{4\pi^2}. \quad (3.61)$$

3.2.3 Genäherte Methoden zur Lösung der vektoriellen Wellengleichung

Für inhomogene Medien gibt es auch vektorielle Erweiterungen der BPM [98] und der WPM [30]. Die vektoriellen Erweiterungen der BPM basieren meist auf Näherungen des Propagators, für die Simulation von höheren numerischen Aperturen wird meist eine Padé-Näherung durchgeführt [39].

Die meisten dieser vektoriellen Erweiterungen für die Lichtausbreitung im inhomogenen Medium vernachlässigen die polarisationsabhängige Reflexion oder Transmission an den Grenzflächen. Eine erwähnenswerte Ausnahme stellt hier die bidirektionale VWPM dar [31].

3.3 Lösungsmethoden für die Maxwell-Gleichungen

Für die Lösung allgemeiner elektromagnetischer Probleme bietet sich das Lösen der Maxwell-Gleichungen direkt an. Im folgenden Abschnitt wird nur die „rigorous coupled-wave analysis“ (RCWA) als Lösungsmethode ausführlich vorgestellt, da dort, ähnlich wie bei der Zerlegung nach ebenen Wellen, eine Moden-Zerlegung der in x und y-Richtung periodischen elektromagnetischen Felder durchgeführt wird. Dies erlaubt einen einfachen Vergleich mit der skalaren Theorie.

Einen Überblick über die Vielzahl an Methoden zur Lösung spezieller elektromagnetischer Probleme findet sich in den Büchern von Bondeson, Rylander und Ingelström [8], Chew, Michielssen, Song und Jin [19] und Van Rienen [97]. Neben der RCWA sind die Finite-Elemente-Methode (FEM) [49] und die „finite-difference time-domain method“ (FDTD) [93] für die Lösung von allgemeinen elektromagnetischen Problemen von besonderer Bedeutung.

3.3.1 Rigorous Coupled Wave Analysis

In diesem Abschnitt wird eine kurze Einführung in die RCWA gegeben. Die grundlegenden Arbeiten zur Formulierung der RCWA stammen aus den 1980er Jahren von Moharam und Gaylord [73, 74] und beschreiben die Beugung an einem planaren Gitter. Später, im Jahr 1995, wurde die Methode von ihnen zur Beschreibung von binären Gittern auf Gitter mit beliebig vielen in z-Richtung homogenen Schichten erweitert [75, 76]. Ein guter Überblick über die historische Entwicklung der RCWA und Erweiterungen der RCWA findet sich in den Arbeiten von Auer [1, 2]. Die Notation und Theorie des folgenden Abschnitts ist an die Arbeit von Auer [2] angelehnt.

Modenbegrenzte Entwicklung

In der RCWA werden sowohl die Felder als auch die Materialparameter durch modenbegrenzte Fourierreihen beschrieben. Der Modenvektor in x-Richtung ist m , der Modenvektor in y-Richtung ist n :

$$m = -MO_x, -MO_x + 1, \dots, 0, \dots, MO_x - 1, MO_x \quad (3.62)$$

$$n = -MO_y, -MO_y + 1, \dots, 0, \dots, MO_y - 1, MO_y. \quad (3.63)$$

Hierbei beschreibt MO_x die Modenzahl in x-Richtung und MO_y die Modenzahl in y-Richtung.

Unterteilung des Rechengebiets

Bei der Beschreibung des Beugungsproblems wird das Rechengebiet in drei Teile aufgeteilt, den Eingangsbereich (incident), das Gitter und den Ausgangsbereich (transmitted). Das Gitter kann in mehrere Schichten (layer) aufgeteilt werden, dabei muss jede Schicht in z-Richtung konstante Materialparameter haben.

Beschreibung des Gitters

Das Gitter wird durch die lokale Permittivität $\varepsilon(\mathbf{r})$ und die lokale Permeabilität $\mu(\mathbf{r})$ beschrieben. Die Gitterperioden sind P_x und P_y . Die zugehörigen Größen im Frequenzraum erhält man durch Fourier-Transformation:

$$\tilde{\varepsilon}_l(\mathbf{k}_\perp) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varepsilon_l(\mathbf{r}_\perp) e^{-i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} d\mathbf{r}_\perp \quad (3.64)$$

$$\tilde{\mu}_l(\mathbf{k}_\perp) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu_l(\mathbf{r}_\perp) e^{-i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} d\mathbf{r}_\perp. \quad (3.65)$$

Wird das Gitter in der l -ten Schicht an den Stellen $x_i = i dx$ mit $i = (0, 1, \dots, 2MO_x)$ und $dx = \frac{P_x}{2MO_x+1}$, sowie $y_j = j dy$ mit $j = (0, 1, \dots, 2MO_y)$ und $\frac{P_y}{2MO_y+1}$ abgetastet, so erhält man mittels FFT die diskreten Fourier-Koeffizienten des Gitters:

$$\tilde{\varepsilon}_{l;m,n} = \text{fft2}(\varepsilon_{l;i,j}) \quad (3.66)$$

$$\tilde{\mu}_{l;m,n} = \text{fft2}(\mu_{l;i,j}). \quad (3.67)$$

Beschreibung der Beleuchtung

In der RCWA können einzelne ebene Wellen oder Überlagerungen von ebenen Wellen als beleuchtende Eingangsverteilung realisiert werden [3]. Im Folgenden wird nur auf den Spezialfall der konischen Beleuchtung mit einer ebenen Welle eingegangen.

Konische Beleuchtung Für die konische Beleuchtung ist das eingestrahlte Feld folgendermaßen definiert:

$$\mathcal{E}_{\text{inc}}(x, y, z) = \mathcal{P}_{\text{conic}} e^{ik_0 n_I (s_{x,\text{inc}} x + s_{y,\text{inc}} y + s_{z,\text{inc}} z)}. \quad (3.68)$$

Hierbei ist $\mathcal{P}_{\text{conic}}$ der konische Polarisationsvektor mit dem Mischungswinkel ψ , der die TE- und TM-Polarisation überlagert:

$$\mathcal{P}_{\text{conic}} = \sin(\psi) \hat{\mathbf{e}}_{\text{TE,inc}} + \cos(\psi) \hat{\mathbf{e}}_{\text{TM,inc}}. \quad (3.69)$$

Die TE- und TM-Einheitsvektoren sind definiert als

$$\hat{\mathbf{e}}_{\text{TE,inc}} = \frac{\hat{\mathbf{e}}_z \times \mathbf{s}_{\text{inc}}}{|\hat{\mathbf{e}}_z \times \mathbf{s}_{\text{inc}}|} = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{pmatrix} \quad (3.70)$$

$$\hat{\mathbf{e}}_{\text{TM,inc}} = \frac{\hat{\mathbf{e}}_{\text{TE,inc}} \times \mathbf{s}_{\text{inc}}}{|\hat{\mathbf{e}}_{\text{TE,inc}} \times \mathbf{s}_{\text{inc}}|} = \begin{pmatrix} \cos(\varphi) \cos(\theta) \\ \sin(\varphi) \cos(\theta) \\ -\sin(\theta) \end{pmatrix}. \quad (3.71)$$

Man sieht, dass die konische Beleuchtung der Beleuchtung mit einer transversal polarisierten ebenen Welle aus Abschnitt 3.2.1 entspricht.

Im Frequenzraum lässt sich die konische Beleuchtungswelle schreiben als

$$L_{x,mn,\text{inc}} = \begin{cases} \mathcal{P}_{x,\text{conic}} & n = m = 0 \\ 0 & \text{sonst} \end{cases} \quad (3.72)$$

$$L_{y,mn,\text{inc}} = \begin{cases} \mathcal{P}_{y,\text{conic}} & n = m = 0 \\ 0 & \text{sonst} \end{cases}. \quad (3.73)$$

Für die weitere Beschreibung ist es von Vorteil, einen gestapelten Vektor aus der x- und y-Komponente des Feldes einzuführen:

$$\vec{L}_{\text{inc},\perp} = \begin{pmatrix} L_{x,mn,\text{inc}}(\cdot) \\ L_{y,mn,\text{inc}}(\cdot) \end{pmatrix}. \quad (3.74)$$

Die Schreibweise $L_{x,mn,\text{inc}}(\cdot)$ bedeutet, dass die Matrix $L_{x,mn,\text{inc}}$ in einen Spaltenvektor umgeformt wird, bei dem die einzelnen Spalten der Matrix untereinander angeordnet sind.

Ansätze für das E- und H-Feld

Für den Eingangsbereich, das Gitter und den Ausgangsbereich werden folgende modenbegrenzte Ansätze gemacht:

Im Eingangsbereich Im Eingangsbereich werden die Felder als Summe eines eingestrahnten und eines reflektierten Anteils beschrieben. Für das elektrische Feld gilt

$$\mathcal{E}_I = \mathcal{E}_{\text{inc}}(x,y,z) + \sum_{m=-\text{MO}_x}^{+\text{MO}_x} \sum_{n=-\text{MO}_y}^{+\text{MO}_y} \mathbf{R}_{mn} e^{i\phi_{I,mn}(x,y,z)}. \quad (3.75)$$

Der reflektierte Anteil wird durch eine Rayleigh-Entwicklung beschrieben. Hierbei gilt für die Exponentialphase:

$$\phi_{I,mn}(x,y,z) = k_{x,mn}x + k_{y,mn}y + \gamma_{I,mn}z. \quad (3.76)$$

Die Wellenvektoren sind definiert als

$$k_{x,mn} = k_{x,0} + m \frac{2\pi}{P_x} \quad (3.77)$$

$$k_{y,mn} = k_{y,0} + n \frac{2\pi}{P_y} \quad (3.78)$$

$$\gamma_{I,mn} = -\sqrt{k_0^2 n_I^2 - k_{x,mn}^2 - k_{y,mn}^2}. \quad (3.79)$$

Durch einen Offset in der x- und y-Komponente lässt sich eine mögliche schräge Beleuchtung bei der Entwicklung berücksichtigen:

$$k_{x,0} = k_0 n_I s_{x,0} \quad (3.80)$$

$$k_{y,0} = k_0 n_I s_{y,0}. \quad (3.81)$$

Hierbei beschreibt $s_{x,0}$ und $s_{y,0}$ die Richtung der schrägen Beleuchtung. Das magnetische Feld \mathcal{H}_I im Eingangsbereich lässt sich aus dem elektrischen Feld \mathcal{E}_I berechnen.

Im Gitter Im Gitter wird die z-Abhängigkeit der Felder durch eine z-Abhängigkeit der Entwicklungskoeffizienten angenommen:

$$\mathcal{E}_{\text{II}} = \sum_{m=-\text{MO}_x}^{+\text{MO}_x} \sum_{n=-\text{MO}_y}^{+\text{MO}_y} (S_{x,mn}(z)\hat{\mathbf{e}}_x + S_{y,mn}(z)\hat{\mathbf{e}}_y + S_{z,mn}(z)\hat{\mathbf{e}}_z) e^{i(k_{x,mn}x + k_{y,mn}y)} \quad (3.82)$$

$$\mathcal{H}_{\text{II}} = \frac{i}{Z_0} \sum_{m=-\text{MO}_x}^{+\text{MO}_x} \sum_{n=-\text{MO}_y}^{+\text{MO}_y} (U_{x,mn}(z)\hat{\mathbf{e}}_x + U_{y,mn}(z)\hat{\mathbf{e}}_y + U_{z,mn}(z)\hat{\mathbf{e}}_z) e^{i(k_{x,mn}x + k_{y,mn}y)} \quad (3.83)$$

Im Ausgangsbereich Im Ausgangsbereich werden die Felder ähnlich wie im Eingangsbereich durch eine Rayleigh-Entwicklung beschrieben:

$$\mathcal{E}_{\text{III}} = \sum_{m=-\text{MO}_x}^{+\text{MO}_x} \sum_{n=-\text{MO}_y}^{+\text{MO}_y} \mathbf{T}_{mn} e^{i\phi_{\text{III},mn}(x,y,z)}. \quad (3.84)$$

Die Exponentialphase ist hier folgendermaßen definiert :

$$\phi_{\text{III},mn}(x, y, z) = k_{x,mn}x + k_{y,mn}y + \gamma_{\text{III},mn}z. \quad (3.85)$$

Hierbei gilt:

$$\gamma_{\text{III},mn} = \sqrt{k_0^2 n_{\text{III}}^2 - k_{x,mn}^2 - k_{y,mn}^2}. \quad (3.86)$$

Kopplung zwischen E- und H-Feld im Außenraum

Für eine ebene Welle, die über den Richtungsvektor \mathbf{s} oder die Winkel φ und θ beschrieben wird, lassen sich die Tangentialkomponenten des elektrischen Feldes in die Tangentialkomponenten des magnetischen Feldes transformieren:

$$\mathbf{A}\vec{\mathcal{E}}_{\perp} = \vec{\mathcal{H}}_{\perp}. \quad (3.87)$$

Hier ist die Transformationsmatrix gegeben durch

$$\mathbf{A} = \frac{\tau}{\cos(\theta)} \begin{pmatrix} -\sin(\varphi) \cos(\varphi) \sin^2(\theta) & -(1 - \cos^2(\varphi) \sin^2(\theta)) \\ 1 - \sin^2(\varphi) \sin^2(\theta) & \sin(\varphi) \cos(\varphi) \sin^2(\theta) \end{pmatrix} \quad (3.88)$$

$$= \frac{\tau}{s_z} \begin{pmatrix} -s_x s_y & -(1 - s_x^2) \\ 1 - s_y^2 & s_x s_y \end{pmatrix}. \quad (3.89)$$

Wobei $\tau = \frac{n}{Z_0 \mu_0}$ die Skalierungskonstante der Amplitude ist.

Für eine Überlagerung aus ebenen Wellen im Frequenzraum folgt somit für die Feldansätze im Außenraum

$$\tilde{\mathbf{A}}\vec{\mathcal{E}}_{\perp} = \vec{\mathcal{H}}_{\perp}. \quad (3.90)$$

Hierbei sind die Fourier-Koeffizienten der Felder vektorisiert

$$\vec{\mathcal{E}}_{\perp} = \begin{pmatrix} \tilde{\mathcal{E}}_{x,mn}(\cdot) \\ \tilde{\mathcal{E}}_{y,mn}(\cdot) \end{pmatrix} \quad (3.91)$$

$$\vec{\mathcal{H}}_{\perp} = \begin{pmatrix} \tilde{\mathcal{H}}_{x,mn}(\cdot) \\ \tilde{\mathcal{H}}_{y,mn}(\cdot) \end{pmatrix}. \quad (3.92)$$

Die Transformationsmatrizen für die Überlagerung des eingestrahnten, des reflektierten und des transformierten Felds ergeben sich zu

$$\tilde{\mathbf{A}}_{\mathbf{R}} = \frac{-i}{k_0} \begin{pmatrix} \text{diag}\left(\frac{-k_{x,mn}(\cdot)k_{y,mn}(\cdot)}{\gamma_{\text{I},mn}(\cdot)}\right) & \text{diag}\left(\frac{-k_{y,mn}^2(\cdot)+\gamma_{\text{I},mn}^2(\cdot)}{\gamma_{\text{I},mn}(\cdot)}\right) \\ \text{diag}\left(\frac{k_{x,mn}^2(\cdot)+\gamma_{\text{I},mn}^2(\cdot)}{\gamma_{\text{I},mn}(\cdot)}\right) & \text{diag}\left(\frac{k_{y,mn}(\cdot)k_{x,mn}(\cdot)}{\gamma_{\text{I},mn}(\cdot)}\right) \end{pmatrix} \quad (3.93)$$

$$\tilde{\mathbf{A}}_{\mathbf{T}} = \frac{-i}{k_0} \begin{pmatrix} \text{diag}\left(\frac{-k_{x,mn}(\cdot)k_{y,mn}(\cdot)}{\gamma_{\text{III},mn}(\cdot)}\right) & \text{diag}\left(\frac{-k_{y,mn}^2(\cdot)+\gamma_{\text{III},mn}^2(\cdot)}{\gamma_{\text{III},mn}(\cdot)}\right) \\ \text{diag}\left(\frac{k_{x,mn}^2(\cdot)+\gamma_{\text{III},mn}^2(\cdot)}{\gamma_{\text{III},mn}(\cdot)}\right) & \text{diag}\left(\frac{k_{y,mn}(\cdot)k_{x,mn}(\cdot)}{\gamma_{\text{III},mn}(\cdot)}\right) \end{pmatrix} \quad (3.94)$$

$$\tilde{\mathbf{A}}_{\text{inc}} = -\tilde{\mathbf{A}}_{\mathbf{R}}. \quad (3.95)$$

Kopplung der Fourier-Koeffizienten im Gitter

Einsetzen des Ansatzes für die Felder im Gitter von Gleichung 3.83 in die Maxwell-Gleichung führt zu folgender Differentialgleichung. Die Herleitung wird in Anhang A.13 gezeigt.

$$\frac{\partial}{\partial z} \begin{pmatrix} \vec{S}_x \\ \vec{S}_y \\ \vec{U}_x \\ \vec{U}_y \end{pmatrix} = k_0 \begin{pmatrix} 0 & 0 & -\mathbf{K}_x[\tilde{\boldsymbol{\epsilon}}]^{-1}\mathbf{K}_y & \mathbf{K}_x[\tilde{\boldsymbol{\epsilon}}]^{-1}\mathbf{K}_x - [\tilde{\boldsymbol{\mu}}] \\ 0 & 0 & [\tilde{\boldsymbol{\mu}}] - \mathbf{K}_y[\tilde{\boldsymbol{\epsilon}}]^{-1}\mathbf{K}_y & \mathbf{K}_y[\tilde{\boldsymbol{\epsilon}}]^{-1}\mathbf{K}_x \\ -\mathbf{K}_x[\tilde{\boldsymbol{\mu}}]^{-1}\mathbf{K}_y & \mathbf{K}_x[\tilde{\boldsymbol{\mu}}]^{-1}\mathbf{K}_x - [\tilde{\boldsymbol{\epsilon}}] & 0 & 0 \\ [\tilde{\boldsymbol{\epsilon}}] - \mathbf{K}_y[\tilde{\boldsymbol{\mu}}]^{-1}\mathbf{K}_y & \mathbf{K}_y[\tilde{\boldsymbol{\mu}}]^{-1}\mathbf{K}_x & 0 & 0 \end{pmatrix} \begin{pmatrix} \vec{S}_x \\ \vec{S}_y \\ \vec{U}_x \\ \vec{U}_y \end{pmatrix} \quad (3.96)$$

Hierbei wurden die Matrizen für die Modengewichte im Gitter vektorisiert:

$$\vec{S}_x = S_{x,mn}(\cdot) \quad (3.97)$$

$$\vec{S}_y = S_{y,mn}(\cdot) \quad (3.98)$$

$$\vec{U}_x = U_{x,mn}(\cdot) \quad (3.99)$$

$$\vec{U}_y = U_{y,mn}(\cdot). \quad (3.100)$$

Man erhält somit zwei Differentialgleichungen erster Ordnung:

$$\frac{\partial}{\partial z} \vec{S}_\perp = \mathbf{M}_1 \vec{U}_\perp \quad (3.110)$$

$$\frac{\partial}{\partial z} \vec{U}_\perp = \mathbf{M}_2 \vec{S}_\perp. \quad (3.111)$$

Die zwei Differentialgleichung erster Ordnung lassen sich zu einer Differentialgleichung zweiter Ordnung zusammenfassen:

$$\frac{\partial^2}{\partial z^2} \vec{U}_\perp = \mathbf{M}_2 \mathbf{M}_1 \vec{U}_\perp. \quad (3.112)$$

Lösung der Eigenwert-Gleichung im Gitter

Die Differentialgleichung zweiter Ordnung ist eine Eigenwert-Gleichung für \vec{U}_\perp . Mit der in der Literatur typischen Ersetzung $k_0^2 \Omega_1 = \mathbf{M}_2 \mathbf{M}_1$ erhält man:

$$\frac{\partial^2}{\partial z^2} \vec{U}_\perp = k_0^2 \Omega_1 \vec{U}_\perp. \quad (3.113)$$

Ω_1 lässt sich als Produkt von Eigenvektoren und Eigenmatrizen darstellen:

$$[\mathbf{W}, \Lambda] = \text{eig}(\Omega_1) \quad (3.114)$$

$$\Omega_1 = \mathbf{W} \Lambda \mathbf{W}^{-1}. \quad (3.115)$$

Einsetzen in die Ausgangsgleichung ergibt:

$$\frac{\partial^2}{\partial z^2} \vec{U}_\perp = k_0^2 \mathbf{W} \Lambda \mathbf{W}^{-1} \vec{U}_\perp. \quad (3.116)$$

Multipliziert man mit \mathbf{W}^{-1} von links, erhält man:

$$\frac{\partial^2}{\partial z^2} \mathbf{W}^{-1} \vec{U}_\perp = k_0^2 \Lambda \mathbf{W}^{-1} \vec{U}_\perp. \quad (3.117)$$

Mit der Definition von

$$\vec{C} = \mathbf{W}^{-1} \vec{U}_\perp \quad (3.118)$$

und umgekehrt

$$\vec{U}_\perp = \mathbf{W} \vec{C}(z) \quad (3.119)$$

erhält man folgende Wellengleichung für \vec{C} :

$$\frac{\partial^2}{\partial z^2} \vec{C}(z) = k_0^2 \Lambda \vec{C}(z). \quad (3.120)$$

Die allgemein Lösung für $C_j(z)$ ist:

$$C_j(z) := c_{j,-} e^{-zk_0 \sqrt{\lambda_j}} + c_{j,+} e^{zk_0 \sqrt{\lambda_j}} \quad (3.121)$$

Man kann die Koeffizienten der rücklaufenden Welle umdefinieren, indem man eine zusätzliche Phase $e^{-dk_0 \sqrt{\lambda_j}}$ addiert. Dies entspricht einer Propagation um $-d$ für einen nicht strukturierten Layer:

$$c_{j,+} = c_j^- e^{-dk_0 \sqrt{\lambda_j}} \quad (3.122)$$

$$c_{j,-} = c_j^+ . \quad (3.123)$$

Hiermit folgt die für die RCWA typische Lösung für $C_j(z)$:

$$C_j(z) := c_j^+ e^{-zk_0 \sqrt{\lambda_j}} + c_j^- e^{(z-d)k_0 \sqrt{\lambda_j}} . \quad (3.124)$$

Einsetzen in die Gleichung 3.111 liefert

$$\vec{S}_\perp = \Omega_2^{-1} \mathbf{W} \mathbf{Q} \vec{D} = \mathbf{V} \vec{D} . \quad (3.125)$$

Hierbei gilt $\Omega_2^{-1} = k_0 \mathbf{M}_2^{-1}$, \vec{D} wurde definiert als:

$$D_j(z) = \frac{1}{\sqrt{\lambda_j} k_0} \frac{\partial}{\partial z} S_j(z) = -c_j^+ e^{-zk_0 \sqrt{\lambda_j}} + c_j^- e^{(z-d)k_0 \sqrt{\lambda_j}} . \quad (3.126)$$

Somit folgt für \mathbf{V}

$$\mathbf{V} = \Omega_2^{-1} \mathbf{W} \mathbf{Q} , \quad (3.127)$$

wobei \mathbf{Q} die Diagonalmatrix der Wurzel der Eigenvektoren ist:

$$Q_{jj} = \sqrt{\Lambda_j} . \quad (3.128)$$

Die Koeffizienten der E- und H-Felder lassen sich jetzt in Abhängigkeit der Koeffizienten c^+ und c^- schreiben:

$$\begin{pmatrix} \vec{S}_\perp \\ \vec{U}_\perp \end{pmatrix} = \begin{pmatrix} -\mathbf{V} \Phi^-(z) & \mathbf{V} \Phi^+(z) \\ \mathbf{W} \Phi^-(z) & \mathbf{W} \Phi^+(z) \end{pmatrix} \begin{pmatrix} \vec{c}^+ \\ \vec{c}^- \end{pmatrix} . \quad (3.129)$$

Die Entwicklung in axiale Richtung wird über folgende Diagonalmatrizen beschrieben:

$$\Phi_{jj}^-(z) = e^{-zk_0 \sqrt{\lambda_j}} \quad (3.130)$$

$$\Phi_{jj}^+(z) = e^{(z-d)k_0 \sqrt{\lambda_j}} . \quad (3.131)$$

Stetigkeitsbedingungen an den Grenzflächen

Die transversalen Komponenten des elektrischen Felds \mathcal{E}_\perp und die transversalen Komponenten des magnetischen Feldes \mathcal{H}_\perp müssen an den Grenzflächen stetig sein.

Die Stetigkeitsbedingung am linken Rand zwischen Eingangsbereich und Gitter ist

$$\begin{pmatrix} \mathbf{I} \\ \tilde{\mathbf{A}}_{\text{inc}} \end{pmatrix} \vec{L}_{\text{inc},\perp} + \begin{pmatrix} \mathbf{I} \\ \tilde{\mathbf{A}}_{\text{R}} \end{pmatrix} \vec{R}_\perp = \begin{pmatrix} -\mathbf{V} & \mathbf{V}\mathbf{X} \\ \mathbf{W} & \mathbf{W}\mathbf{X} \end{pmatrix} \begin{pmatrix} \vec{c}_1^+ \\ \vec{c}_1^- \end{pmatrix}. \quad (3.132)$$

Hierbei wurde $\Phi^+(0) = \mathbf{X} = \text{diag}\left(e^{-dk_0\sqrt{\lambda_j}}\right)$ verwendet.

Die Stetigkeitsbedingung am rechten Rand zwischen Gitter und Ausgangsbereich ist

$$\begin{pmatrix} \mathbf{I} \\ \tilde{\mathbf{A}}_{\text{T}} \end{pmatrix} \vec{T}_\perp = \begin{pmatrix} -\mathbf{V}\mathbf{X} & \mathbf{V} \\ \mathbf{W}\mathbf{X} & \mathbf{W} \end{pmatrix} \begin{pmatrix} \vec{c}_L^+ \\ \vec{c}_L^- \end{pmatrix}. \quad (3.133)$$

Hierbei wurde $\Phi^-(d) = \mathbf{X} = \text{diag}\left(e^{-dk_0\sqrt{\lambda_j}}\right)$ verwendet.

Die Stetigkeitsbedingung im Gitter an der Layer-Grenze ist

$$\begin{pmatrix} \vec{S}_{\perp,l-1}(d_{l-1}) \\ \vec{U}_{\perp,l-1}(d_{l-1}) \end{pmatrix} = \begin{pmatrix} \vec{S}_{\perp,l}(0) \\ \vec{U}_{\perp,l}(0) \end{pmatrix} \quad (3.134)$$

$$\begin{pmatrix} -\mathbf{V}_{l-1}\mathbf{X}_{l-1} & \mathbf{V}_{l-1} \\ \mathbf{W}_{l-1}\mathbf{X}_{l-1} & \mathbf{W}_{l-1} \end{pmatrix} \begin{pmatrix} \vec{c}_{l-1}^+ \\ \vec{c}_{l-1}^- \end{pmatrix} = \begin{pmatrix} -\mathbf{V}_l & \mathbf{V}_l\mathbf{X}_l \\ \mathbf{W}_l & \mathbf{W}_l\mathbf{X}_l \end{pmatrix} \begin{pmatrix} \vec{c}_l^+ \\ \vec{c}_l^- \end{pmatrix}. \quad (3.135)$$

Einsetzen der Randbedingungen

Bei vorgegebener Eingangsmodenverteilung $\vec{L}_{\text{inc},\perp}$ lässt sich der „partial solution approach“ des „enhanced transmittance matrix approach“ (ETMA) zum Lösen der Randbedingungen verwenden [76]. Diese Lösungsmethode vermeidet die Invertierung der Matrizen \mathbf{X}_l , die sonst numerischen Instabilitäten verursachen. Hierzu werden zuerst folgende Hilfsmatrizen \mathbf{a}_l , \mathbf{b}_l , \mathbf{f}_l und \mathbf{g}_l definiert:

$$\begin{pmatrix} \mathbf{a}_l \\ \mathbf{b}_l \end{pmatrix} = \begin{pmatrix} -\mathbf{V} & \mathbf{f}_{l+1} \\ -\mathbf{W} & \mathbf{g}_{l+1} \end{pmatrix}^{-1} \begin{pmatrix} -\mathbf{V}\mathbf{X} \\ \mathbf{W}\mathbf{X} \end{pmatrix} \quad (3.136)$$

$$\begin{pmatrix} \mathbf{f}_l \\ \mathbf{g}_l \end{pmatrix} = \begin{pmatrix} -\mathbf{V} + \mathbf{V}\mathbf{X}\mathbf{a}_l \\ +\mathbf{W} + \mathbf{W}\mathbf{X}\mathbf{a}_l \end{pmatrix}. \quad (3.137)$$

Die rechte Randbedingungen liefert die Startwerte:

$$\mathbf{f}_{L+1} = \mathbf{I} \quad (3.138)$$

$$\mathbf{g}_{L+1} = \tilde{\mathbf{A}}_{\text{T}}. \quad (3.139)$$

Hiermit lässt sich sukzessiv $\mathbf{a}_l, \mathbf{b}_l, \mathbf{f}_l, \mathbf{g}_l$ für jeden Layer des Gitters berechnen. Die linke Randbedingung lässt sich schreiben als

$$\begin{pmatrix} \vec{c}_1^+ \\ \vec{R}_\perp \end{pmatrix} = \begin{pmatrix} \mathbf{f}_1 & -\mathbf{I} \\ \mathbf{g}_1 & -\tilde{\mathbf{A}}_R \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{I} \\ \tilde{\mathbf{A}}_{\text{inc}} \end{pmatrix} \vec{L}_{\text{inc},\perp}. \quad (3.140)$$

Hiermit lassen sich \vec{c}_1^+ und \vec{R}_\perp berechnen.

Aus den \vec{c}_l^+ lassen sich nun mit der \mathbf{a}_l -Matrix die \vec{c}_l^- berechnen:

$$\vec{c}_l^- = \mathbf{a}_l \vec{c}_l^+. \quad (3.141)$$

Für die Layer $l = 1, \dots, L-1$ lassen sich sukzessiv die \vec{c}_l^+ -Koeffizienten ausrechnen

$$\vec{c}_{l+1}^+ = \mathbf{b}_l \vec{c}_l^+. \quad (3.142)$$

Zu guter Letzt erhält man die Koeffizienten für den Ausgangsbereich

$$\mathbf{T}_\perp = \mathbf{b}_L \vec{c}_L^+. \quad (3.143)$$

z-Komponenten der Fourier-Koeffizienten im Außenraum

Die z-Komponenten der Fourier-Koeffizienten der elektrischen Felder im Außenraum berechnen sich wie folgt:

$$R_{z,mn} = \begin{cases} 0 & m = n = 0 \\ -\frac{1}{\gamma_{I,mn}} (R_{y,mn} k_{y,mn} + R_{x,mn} k_{x,mn}) & \text{sonst} \end{cases} \quad (3.144)$$

$$T_{z,mn} = \begin{cases} 0 & m = n = 0 \\ -\frac{1}{\gamma_{III,mn}} (T_{y,mn} k_{y,mn} + T_{x,mn} k_{x,mn}) & \text{sonst} \end{cases} \quad (3.145)$$

$$L_{z,\text{inc},mn} = \begin{cases} 0 & m = n = 0 \\ -\frac{1}{\gamma_{\text{inc},mn}} (L_{y,\text{inc},mn} k_{y,mn} + L_{x,\text{inc},mn} k_{x,mn}) & \text{sonst} \end{cases}. \quad (3.146)$$

Einsetzen des Ansatzes für die Felder im Gitter von Gleichung 3.83 in die Maxwell-Gleichung führt zu folgenden Gleichungen für die z-Komponenten im Gitter:

$$S_{z,mn} = -\frac{i}{k_0} \sum_{p,q} \varepsilon_{m-p,n-q}^{-1} (k_{x,pq} U_{y,pq} - k_{y,pq} U_{x,pq}) \quad (3.147)$$

$$U_{z,mn} = -\frac{i}{k_0} \sum_{q,p} \mu_{m-p,n-q}^{-1} (k_{x,pq} S_{y,pq} - k_{y,pq} S_{x,pq}). \quad (3.148)$$

Globale Reflexion, globale Transmission und globale Absorption

Aus den Fourier-Koeffizienten im Außenraum lässt sich die globale Reflexion, globale Transmission und die globale Absorption berechnen:

$$\mathcal{R} = \frac{\sum_{mn} \operatorname{Re}[-\gamma_{\text{I},mn}] (|R_{x,mn}|^2 + |R_{y,mn}|^2 + |R_{z,mn}|^2)}{\sum_{mn} \operatorname{Re}[\gamma_{\text{inc},mn}] (|L_{x,\text{inc},mn}|^2 + |L_{y,\text{inc},mn}|^2 + |L_{z,\text{inc},mn}|^2)} \quad (3.149)$$

$$\mathcal{T} = \frac{\sum_{mn} \operatorname{Re}[\gamma_{\text{III},mn}] (|T_{x,mn}|^2 + |T_{y,mn}|^2 + |T_{z,mn}|^2)}{\sum_{mn} \operatorname{Re}[\gamma_{\text{inc},mn}] (|L_{x,\text{inc},mn}|^2 + |L_{y,\text{inc},mn}|^2 + |L_{z,\text{inc},mn}|^2)} \quad (3.150)$$

$$\mathcal{A} = 1 - (\mathcal{T} + \mathcal{R}). \quad (3.151)$$

Anschaulich geben diese Werte an, wie viel Leistung reflektiert, transmittiert und absorbiert wird.

Ablauf einer RCWA-Simulation

Der Ablauf einer RCWA-Simulation lässt sich in drei Teile aufteilen. Im ersten Teil wird allgemein das Problem beschrieben. Hierbei wird die Anzahl der Moden, die Brechungsindexverteilungen im Außenraum, die Periodengrößen, die Permittivitäts- und Permeabilitätsverteilungen im Layer und das eingestrahlte Licht definiert. Im zweiten Teil wird das Eigenwertproblem für die einzelnen Layer gelöst und mit der ETMA werden die Randbedingungen berücksichtigt. Im dritten Teil lassen sich die E- und H-Felder rekonstruieren, sowie bestimmte Eigenschaften wie zum Beispiel die globale Reflexion, Transmission und Absorption berechnen.

4 Design von diffraktiven Linsenarrays mit überlappenden Aperturen

In diesem Kapitel wird das Design von diffraktiven Linsenarrays mit überlappenden Aperturen vorgestellt. Dazu werden in Abschnitt 4.1 zuerst zwei Methoden zur Erzeugung von Linsen mit überlappenden Aperturen vorgestellt. Anschließend wird in Abschnitt 4.2 auf das Design spezieller diffraktiver Linsenarrays eingegangen. Insbesondere wird der allgemeine Fall einer beliebigen Beleuchtungsstruktur vorgestellt und es werden die Spezialfälle punktförmiger und ringförmiger Beleuchtungsstrukturen behandelt. Anschließend werden in Abschnitt 4.3 Erweiterungen des Designprozesses auf mehrere Foki und die Integration einer Deckglaskorrektur behandelt. Im Abschnitt 4.4 wird dann die Realisierung der diffraktiven Linsen als dünnes Element beschrieben. Abschließend wird im Abschnitt 4.5 auf die Geometrie des Fokus für den Spezialfall einer homogenen kreisförmigen Pupille eingegangen.

4.1 Design-Methoden

Diffraktive Linsen mit überlappenden Aperturen lassen sich sowohl im Ortsraum als auch im Frequenzraum designen. Beim Design im Ortsraum lässt sich der Überlapp durch kohärente Addition der komplexen Amplituden realisieren. Im Frequenzraum wird das Spektrum diskretisiert, die Rücktransformation in den Ortsraum erzeugt das Überlappen.

4.1.1 Überlappende Aperturen durch kohärente Überlagerung

Beim Design im Ortsraum spezifiziert man zuerst die Geometrie des Linsenarrays durch die Gittervektoren \mathbf{P}_1 und \mathbf{P}_2 und gibt die komplexe Amplitude $t_A(\mathbf{r}_\perp)$ innerhalb einer Apertur vor. Die Grundperiode $t(\mathbf{r}_\perp)$ des diffraktiven Linsenarrays innerhalb einer Periode erhält man nun durch kohärente Überlagerung der komplexen Amplitude innerhalb der Apertur, die um vielfache des Gitter-Rasters verschoben sind:

$$t(\mathbf{r}_\perp) = \sum_{m,n=-\infty}^{\infty} t_A(\mathbf{r}_\perp + m\mathbf{P}_1 + n\mathbf{P}_2) \quad (4.1)$$

Aufgrund der endlichen Aperturgröße liefern nur endlich viele Terme Beiträge zur Grundperiode $t(\mathbf{r}_\perp)$ bei.

Beispiel: Linsenarray mit quadratischer Apertur Für den Fall eines quadratischen Gitters mit Gittervektoren $\mathbf{P}_1 = P \hat{\mathbf{e}}_x$ und $\mathbf{P}_2 = P \hat{\mathbf{e}}_y$ und einer quadratischen, zentrierten Apertur mit Durchmesser D , ergibt sich für die Grundperiode

$$t(\mathbf{r}_\perp) = \sum_{m,n=-\lceil \frac{D}{2P} \rceil}^{\lceil \frac{D}{2P} \rceil} t_A(\mathbf{r}_\perp + mP\hat{\mathbf{e}}_x + nP\hat{\mathbf{e}}_y). \quad (4.2)$$

Es werden also $(2\lceil \frac{D}{2P} \rceil + 1)^2$ Beiträge für die Grundperiode überlagert.

Beispiel: Linsenarray mit runder Apertur Für den Fall eines quadratischen Gitters mit Gittervektoren $\mathbf{P}_1 = P \hat{\mathbf{e}}_x$ und $\mathbf{P}_2 = P \hat{\mathbf{e}}_y$ und einer runden Apertur mit Radius R ergibt sich

$$t(\mathbf{r}_\perp) = \sum_{m,n} t_A(\mathbf{r}_\perp + mP\hat{\mathbf{e}}_x + nP\hat{\mathbf{e}}_y) \quad \text{für } \sqrt{m^2 + n^2} \leq \frac{R}{P}. \quad (4.3)$$

Definition des Überlapps im Ortsraum

Das Verhältnis zwischen dem Durchmesser der Apertur und der Gitterperiode wird als Definition für den räumlichen Überlapp verwendet:

$$\Omega = \frac{D_{\text{Apertur}}}{P_{\text{Gitter}}} \quad (4.4)$$

Anschaulich gibt der Überlappfaktor Ω an wie viele Perioden zu der Grundperiode beitragen.

In Abbildung 4.1 sieht man ein Beispiel für das Überlappen der Aperturen bei einem quadratischen Gitter und runden Aperturen. Der Überlappfaktor ist genau 2. In rot ist die quadratische Grundperiode eingezeichnet, blau sind die runden Aperturen. Nur die durchgezogenen Aperturen tragen zur Grundperiode bei, die gestrichelten Aperturen liefern hingegen keinen Beitrag zur Grundperiode.

4.1.2 Überlappende Aperturen durch Fourier-Ansatz

Beim Design im Frequenzraum spezifiziert man zuerst die Fouriertransformierte der Grundperiode des Elementes $\tilde{t}(\mathbf{k}_\perp)$ und diskretisiert diese anschließend. Für das diskrete Spektrum \tilde{t}_{mn} erhält man

$$\tilde{t}_{mn} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{t}(\mathbf{k}_\perp) \delta(k_x - m dk_x, k_y - n dk_y) d\mathbf{k}_\perp. \quad (4.5)$$

Die vorgegebenen Gittervektoren des Linsenarrays P_x, P_y definieren die Schrittweiten dk_x und dk_y , mit der bei der Diskretisierung der Frequenzraum abgetastet wird.

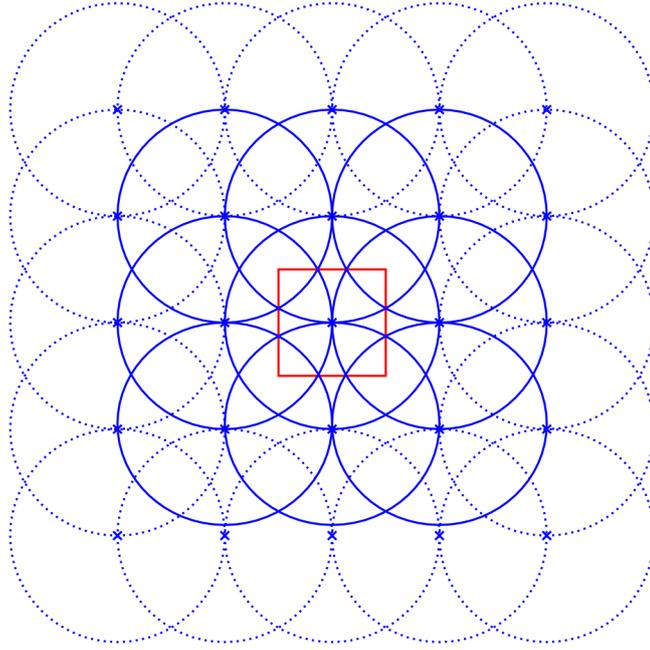


Abbildung 4.1: Visualisierung des Überlapps der Aperturen am Beispiel eines quadratischen Gitters und runden Aperturen. Der Überlappfaktor beträgt genau zwei. Nur die durchgezogenen Aperturen tragen zur Grundperiode bei.

Für die Abtastung im Frequenzraum gilt:

$$dk_x = \frac{2\pi}{P_x} \quad (4.6)$$

$$dk_y = \frac{2\pi}{P_y} \quad (4.7)$$

Durch eine diskrete Fourier-Rücktransformation wie der schnellen inversen Fourier-Transformation (iFFT) oder durch Rücktransformation mit der Fourierreihe erhält man entweder die diskreten Koeffizienten t_{kl} der Grundperiode oder die kontinuierliche Beschreibung der Grundperiode $t(\mathbf{r}_\perp)$ des diffraktiven Elements:

$$t_{kl} = \text{ifft2}[\tilde{t}_{mn}] \frac{dk_x dk_y}{(2\pi)^2} \quad (4.8)$$

$$t(\mathbf{r}_\perp) = \frac{1}{P_x P_y} \sum_{n,m} \tilde{t}_{mn} e^{i2\pi \left(\frac{xm}{P_x} + \frac{yn}{P_y} \right)}. \quad (4.9)$$

Die diskreten Koeffizienten t_{kl} lassen sich anschaulich als die diskretisierte und bandbegrenzte Beschreibung der kontinuierlichen Grundperiode auffassen.

Zum Vergleich, für nicht diskretisierte Spektren $\tilde{t}(\mathbf{k}_\perp)$ erhält man mittels Fourier-

Rücktransformation die aperiodische Verteilung $t'(\mathbf{r}_\perp)$ im Ortsraum:

$$t'(\mathbf{r}_\perp) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{t}(\mathbf{k}_\perp) e^{i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} \frac{d\mathbf{k}_\perp}{4\pi^2}. \quad (4.10)$$

Diese Ortsverteilung ist nicht periodisch, jedoch kann man mit der Methode der kohärenten Überlagerung aus Abschnitt 4.1.1 für ortsbegrenzte Verteilungen $t'(\mathbf{r}_\perp)$ eine periodische Verteilung erzeugen.

Maximale Frequenz beim Ansatz im Fourierraum

Die äquidistante Abtastung im Frequenzraum sorgt dafür, dass bei N_ν Abtastpunkten im Frequenzraum die maximale Frequenz im Frequenzraum $\nu_{\text{Theorie}}^{\max}$ beschränkt ist:

$$\nu_{\text{Theorie}}^{\max} = \frac{N_\nu}{2P}. \quad (4.11)$$

Dies hat zur Folge, dass nur bandbegrenzte Elemente designt werden können. Die maximale Frequenz $\nu_{\text{Design}}^{\max}$ muss kleiner als die maximale Frequenz der Diskretisierung sein:

$$\nu_{\text{Design}}^{\max} \leq \nu_{\text{Theorie}}^{\max}. \quad (4.12)$$

4.2 Design spezieller Linsenarrays mit überlappenden Aperturen

Im Folgenden wird das Design von verschiedenen Linsenarrays mit überlappenden Aperturen aufgezeigt. Zuerst wird der allgemeine Fall beliebiger Beleuchtungsstrukturen behandelt, anschließend werden die Spezialfälle der punktförmigen Beleuchtung und der ringförmigen Beleuchtung näher untersucht.

4.2.1 Beliebige Beleuchtungsstrukturen

Im allgemeinen Fall legt man zuerst die komplexe Amplitude $U_f(\mathbf{r}_\perp)$ fest, die im Abstand f hinter einem diffraktiven optischen Element (DOE) periodisch wiederholt erzeugt werden soll. Um solch ein DOE zu erzeugen, bestimmt man mit der Fourier-Transformation das Spektrum \tilde{U}_f der gewünschten Amplitude. Hierbei ist es von Vorteil das Spektrum in Abhängigkeit der normierten Frequenzkoordinate $\mathbf{s}_\perp = \frac{\mathbf{k}_\perp}{k_0}$ zu schreiben.

$$\tilde{U}_f(\mathbf{s}_\perp) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U(\mathbf{r}_\perp) e^{-ik_0 \mathbf{s}_\perp \cdot \mathbf{r}_\perp} d\mathbf{r}_\perp \quad (4.13)$$

Anschließend wird das Spektrum mittels der Zerlegung nach ebenen Wellen um die Distanz $-f$ ausgebreitet. Dazu wird das Spektrum im Frequenzraum mit dem entsprechenden Propagator multipliziert. Die periodische Wiederholung erreicht man durch eine Diskretisierung im Frequenzraum:

$$\tilde{t}^{\text{DOE}}(\mathbf{s}_{\perp}) = \tilde{U}_f(\mathbf{s}_{\perp}) e^{-ik_0 f \sqrt{n^2 - \mathbf{s}_{\perp} \cdot \mathbf{s}_{\perp}}} \sum_{lm} \delta\left(s_x - \frac{l\lambda}{P_x}, s_y - \frac{m\lambda}{P_y}\right) \quad (4.14)$$

Das diskrete Element lässt sich durch seine Fourier-Koeffizienten beschreiben:

$$\tilde{t}_{lm}^{\text{DOE}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{U}_f(\mathbf{s}_{\perp}) e^{-ik_0 f \sqrt{n^2 - \mathbf{s}_{\perp} \cdot \mathbf{s}_{\perp}}} \delta\left(s_x - \frac{l\lambda}{P_x}, s_y - \frac{m\lambda}{P_y}\right) d\mathbf{s}_{\perp} \quad (4.15)$$

Durch Rücktransformation in den Ortsraum erhält man schließlich die zugehörige Elementarzelle t_{DOE} . Benutzt man hierfür die Fourierreihe, so erhält man eine kontinuierliche Beschreibung für $t_{\text{DOE}}(\mathbf{r}_{\perp})$. Mit einer diskreten Fourier-Rücktransformation wie der iFFT erhält man eine diskrete Beschreibung der Elementarzelle t_{lm}^{DOE} :

$$t_{\text{DOE}}(\mathbf{r}_{\perp}) = \frac{1}{P_x P_y} \sum_{l,m} \tilde{t}_{lm}^{\text{DOE}} e^{i2\pi\left(\frac{x_l}{P_x} + \frac{y_m}{P_y}\right)} \quad (4.16)$$

$$t_{lm}^{\text{DOE}} = \frac{1}{P_x P_y} \text{fft2}\left[\tilde{t}_{lm}^{\text{DOE}}\right] \quad (4.17)$$

Beliebige Beleuchtungsstrukturen mit bandbegrenztem Spektrum

Im Allgemeinen lässt sich das Spektrum des diffraktiven Elements zusätzlich bandbegrenzen:

$$\tilde{t}_{\text{DOE}}^{\text{bandbegrenzt}}(\mathbf{s}_{\perp}) = \tilde{t}_{\text{DOE}}(\mathbf{s}_{\perp}) \text{circ}\left(\frac{|\mathbf{s}_{\perp}|}{s_{\text{max}}}\right) \quad (4.18)$$

Hierbei ist s_{max} die maximale Frequenz, bei der das Spektrum abgeschnitten wird.

4.2.2 Punktförmige Beleuchtungsstrukturen

Der wichtigste Spezialfall der allgemeinen Beleuchtungsstrukturen ist die punktförmige Beleuchtung. Im Allgemeinen gibt es beliebig viele Varianten um eine punktförmige oder näherungsweise punktförmige Beleuchtung zu erzeugen.

Im Folgenden werden die punktförmigen Beleuchtungsstrukturen nach ihrer Beschreibung im Frequenzraum unterschieden. Insbesondere werden die Fälle einer homogenen Pupille, einer kreisförmigen Pupille und einer ringförmigen Pupille näher untersucht. In der Literatur wurden bislang nur diffraktive Linsenarrays mit überlappenden Aperturen und kreisförmigen Pupillen untersucht [40].

Delta-Funktion oder homogene Pupille

Der Idealfall einer punktförmigen Beleuchtung im Ortsraum wird durch die Delta-Funktion beschrieben:

$$U_f(\mathbf{r}_\perp) = U_0 \delta(\mathbf{r}_\perp) \quad (4.19)$$

Über die Skalierung U_0 wird die Einheit der skalaren Amplitude bestimmt. Die nicht-diskretisierte Fouriertransformierte ist eine Konstante:

$$\tilde{U}^f(\mathbf{s}_\perp) = U_0 \quad (4.20)$$

Dieser Idealfall ist in der Praxis nicht zugänglich, da immer nur ein endlicher Beitrag des Frequenzspektrums verwendet werden kann.

Jinc-Funktion oder homogene kreisförmige Pupille

Die Jinc-förmige Amplitudenverteilung (siehe Anhang A.2) kann als Näherung an eine punktförmige Amplitudenverteilung betrachtet werden. Die Amplitudenverteilung im Fokus lässt sich somit definieren als

$$U_f(r_\perp) = U_0 \frac{(k_0 \text{NA})^2}{2\pi} \text{jinc}(k_0 r_\perp \text{NA}) \quad (4.21)$$

$$= U_0 \frac{k_0 \text{NA}}{2\pi} \frac{J_1(k_0 r_\perp \text{NA})}{r_\perp}. \quad (4.22)$$

Dabei ist J_1 die 1. Besselfunktion erster Art. Die Fouriertransformierte dieser Verteilung ist eine Kreisfunktion im Frequenzraum.

$$\tilde{U}^f(s_\perp) = U_0 \text{circ}\left(\frac{s_\perp}{\text{NA}}\right) \quad (4.23)$$

Dies lässt sich leicht zeigen, indem man das Spektrum $\tilde{U}^f(s_\perp)$ in inverse Fourier-Bessel-Transformation aus Gleichung 3.12 einsetzt:

$$U_f(r_\perp) = \int_0^\infty \tilde{U}^f(s_\perp) J_0(k_0 s_\perp r_\perp) \frac{k_0^2}{2\pi} s_\perp ds_\perp \quad (4.24)$$

$$= \frac{k_0^2}{2\pi} \int_0^{\text{NA}} U_0 J_0(k_0 s_\perp r_\perp) s_\perp ds_\perp \quad (4.25)$$

$$= U_0 \frac{k_0 \text{NA}}{2\pi} \frac{J_1(k_0 r_\perp \text{NA})}{r_\perp} \quad (4.26)$$

Eine Jinc-förmige Amplitude erhält man also durch eine rotationssymmetrische Bandbegrenzung des Delta-Funktions-Spektrums. Die numerische Apertur NA bestimmt hierbei die Kante der Bandbegrenzung.

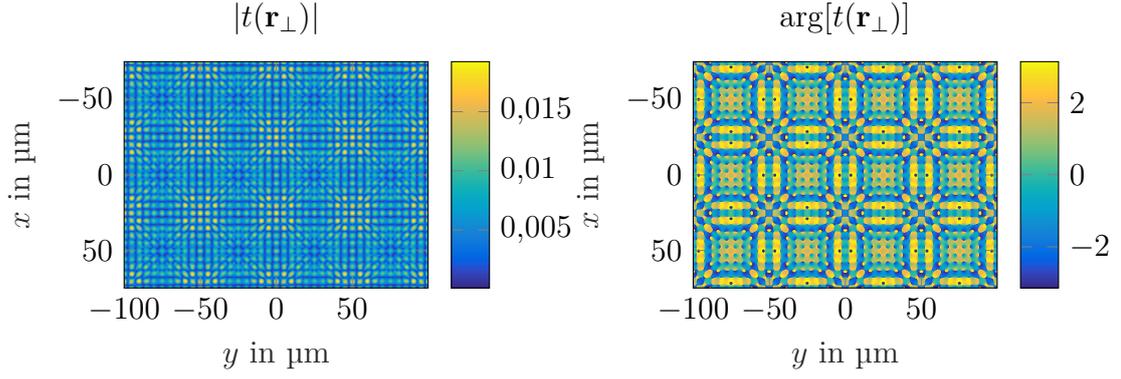


Abbildung 4.2: Phase und Amplitude eines diffraktiven Elements mit kreisförmiger Pupille und einem Überlapp von 2. In x-Richtung sind vier Perioden und in y-Richtung sind drei Perioden der Elementarzelle t gezeigt.

Multipliziert man die Fouriertransformierte im Fokus mit dem Propagator der Zerlegung nach ebenen Wellen, so erhält man die Fouriertransformierte des Elements.

$$\tilde{t}(s_\perp) = U_0 \operatorname{circ}\left(\frac{s_\perp}{\text{NA}}\right) e^{-ik_0 f \sqrt{n^2 - s_\perp \cdot s_\perp}} \quad (4.27)$$

Die Propagationslänge $-f$ entspricht der Fokallänge des diffraktiven Elements. Diese Verteilung im Frequenzraum lässt sich nun diskretisieren. Durch die Rücktransformation in den Ortsraum erhält man die Beschreibung für die Elementarzelle des diffraktiven Elements.

Bestimmung des Überlapps Zu der numerischen Apertur lässt sich unter Kenntnis der Fokallänge f ein effektiver Aperturdurchmesser D_{eff} bestimmen.

$$D_{\text{eff}} = 2f \tan\left(\sin^{-1}(\text{NA})\right) = 2f \frac{\text{NA}}{\sqrt{1 - \text{NA}^2}} \quad (4.28)$$

Über den effektiven Aperturdurchmesser lässt sich der räumliche Überlapp bestimmen:

$$\Omega = \frac{D_{\text{eff}}}{P} = \frac{2f\text{NA}}{P\sqrt{1 - \text{NA}^2}} \quad (4.29)$$

Alternativ lässt sich auch ein Flächen-Überlapp definieren:

$$\Omega_{\text{Fläche}} = \frac{\pi D_{\text{eff}}^2}{4P^2} = \frac{\pi}{4} \Omega^2 = \frac{\pi f^2 \text{NA}^2}{P^2 (1 - \text{NA}^2)} \quad (4.30)$$

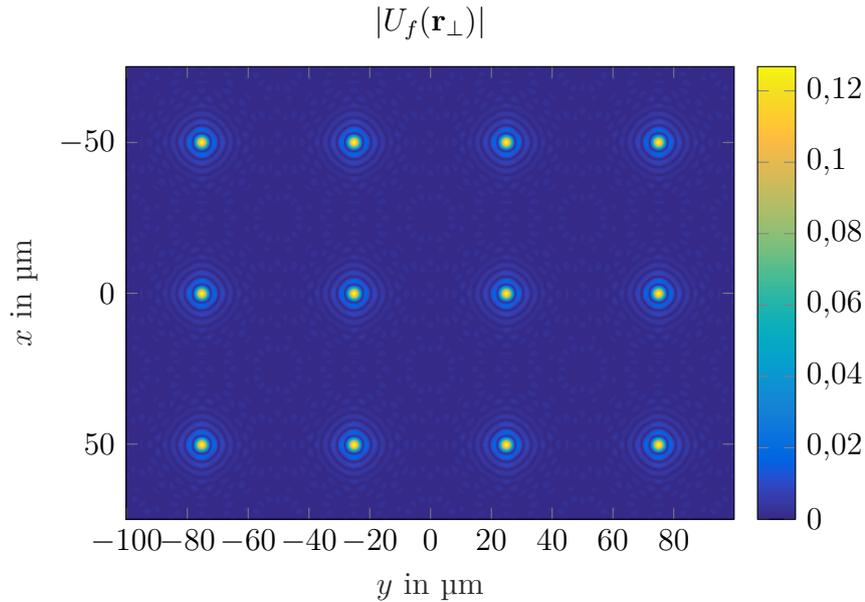


Abbildung 4.3: Amplitude in der Fokusebene eines kontinuierlichen diffraktiven Elements mit überlappenden Aperturen und homogener kreisförmiger Pupille. Die Periode des Elements beträgt $P = 50 \mu\text{m}$ und die numerische Apertur ist $\text{NA} = 0,1$. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$.

Beispiel In Abbildung 4.2 ist ein Beispiel für ein diffraktives Element mit überlappenden Aperturen und kreisförmiger Pupille gezeigt. Die Periode der Elementarzelle beträgt $50 \mu\text{m}$, die Fokusdistanz ist $500 \mu\text{m}$, die numerische Apertur ist $0,1$. Bei der Design-Wellenlänge von $0,5 \mu\text{m}$ entspricht dies einem Überlapp von 2. Mit der Ebenen-Wellen-Zerlegung lässt sich die Ausbreitung in die Fokalebene berechnen. In Abbildung 4.3 sieht man die Amplitude in der Fokusebenen zu diesem kontinuierlichen Element.

Mit Hilfe der Zerlegung nach ebenen Wellen lässt sich auch das axiale Profil der Foki berechnen. In Abbildung 4.4 sieht man einen Schnitt durch die yz -Ebene.

Ideale ringförmige Pupille

Die Ringpupille ist in der Literatur immer wieder von besonderem Interesse, da sie einen schmälere Fokus als die Kreispupille erzeugt. Im Allgemeinen ist dies aber kein Alleinstellungsmerkmal, denn es ist theoretisch möglich die Fokusbreite beliebig klein werden zulassen [24].

Den Idealfall einer unendlich dünnen ringförmige Apertur kann man mit Hilfe der δ -Funktion beschreiben.

$$\tilde{U}_f(s_\perp) = U_0 \delta(\text{NA} - |\mathbf{s}_\perp|) \quad (4.31)$$

Eine direkte Diskretisierung der idealen ringförmigen Pupille ist nicht möglich. Mit Hilfe der Inversen-Fourier-Bessel-Transformation lässt sich aber die Amplitudenver-

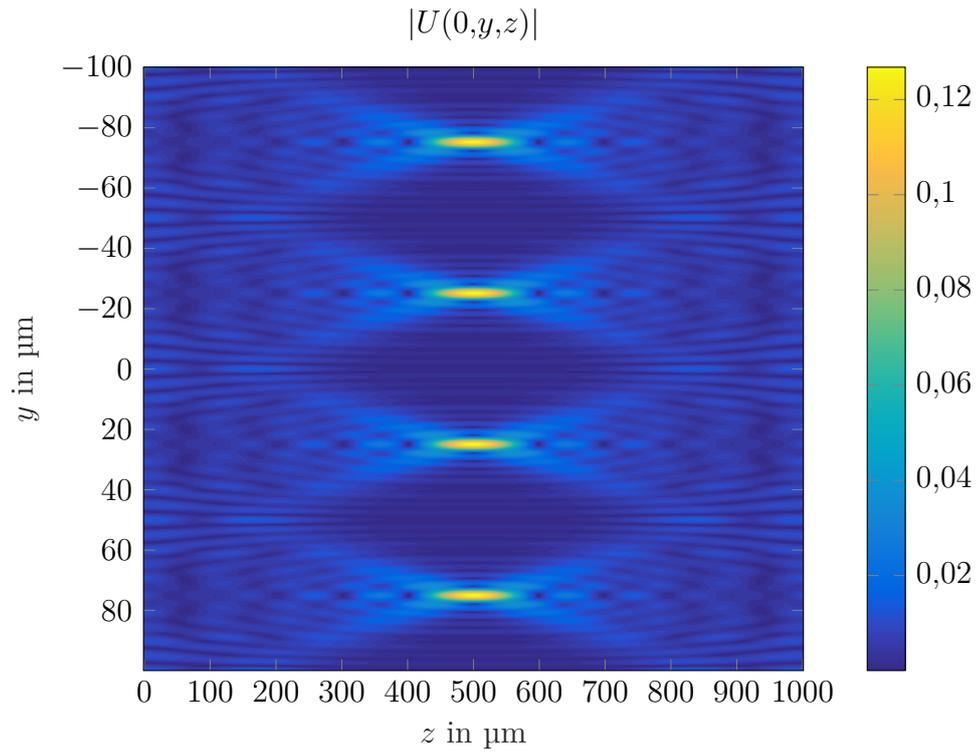


Abbildung 4.4: yz -Schnitt durch die Amplitude in der Fokusebene eines kontinuierlichen diffraktiven Elements mit homogener kreisförmiger Pupille. Die Periode des Elements beträgt $P = 50 \mu\text{m}$ und die numerische Apertur ist $\text{NA} = 0,1$. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$.

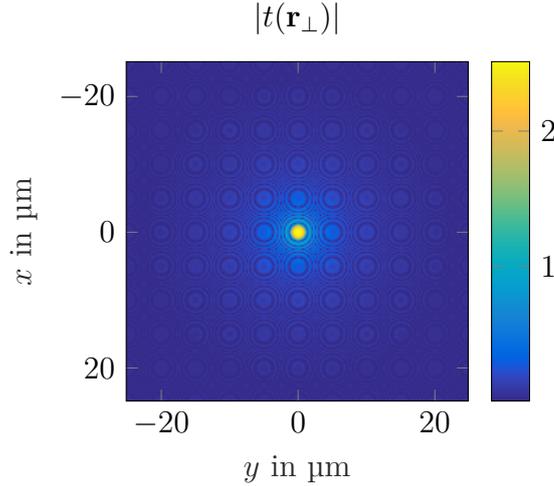


Abbildung 4.5: Amplitudenverteilung einer kontinuierlichen Linse mit idealer ringförmiger Pupille. Die Fokusslänge beträgt $500\ \mu\text{m}$ und die numerische Apertur ist $0,1$. Die Wellenlänge ist $\lambda = 0,5\ \mu\text{m}$.

teilung im Fokus bestimmen.

$$U_f(r_\perp) = U_0 \frac{k_0^2 \text{NA}}{2\pi} J_0(k_0 \text{NA} r_\perp) \quad (4.32)$$

Die Nullstellenbreite des Fokus beträgt $\delta x = 2 \times 0,38 \frac{\lambda}{\text{NA}}$ und ist damit deutlich kleiner als die Breite der homogenen Pupille. Jedoch sind bekanntermaßen die Nebenmaxima dieses Fokus deutlich stärker.

Nach Rückpropagation in die Elementebene erhält man für die kontinuierliche Amplitudenverteilung wieder ein Bessel-Profil, es handelt sich um einen sogenannten „non diffracting beam“ handelt, der sich durch Propagation nicht verbreitert [26]:

$$t'(r_\perp) = U_0 \frac{k_0^2 \text{NA}}{2\pi} e^{-ifk_0 \sqrt{n^2 - \text{NA}^2}} J_0(k_0 \text{NA} r_\perp). \quad (4.33)$$

Das Element ist ein reines Amplitudenelement mit einer konstanten Phase. Durch kohärente Überlagerung im Ortsraum lässt sich die Grundperiode des diffraktiven Elements mit überlappenden Aperturen nur näherungsweise berechnen, da die Besselfunktion nicht ortsbegrenzt ist.

Exemplarisch ist in Abbildung 4.5 die Amplitude eines Elements mit numerischer Apertur von $0,1$ und Fokusslänge $500\ \mu\text{m}$ für die Wellenlänge $0,5\ \mu\text{m}$ gezeigt. Da die Besselfunktion J_0 in Gleichung 4.33 sehr hochfrequent schwingt ist eine sehr feine Abtastung schon bei niedrigen numerischen Aperturen notwendig. Außerdem fällt die Amplitudenverteilung im Fokus und in der Elementebene relativ langsam ab im Vergleich zur homogenen kreisförmigen Pupille. Für große Argumente $x \gg 0,5$ gilt $J_0(x) \approx \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{\pi}{4}\right)$. Dies hat zur Folge, dass sich die einzelnen Foki bei der kohärenten Überlagerung in der Fokalebene eher stören als bei Foki mit homogener kreisförmiger Pupille.

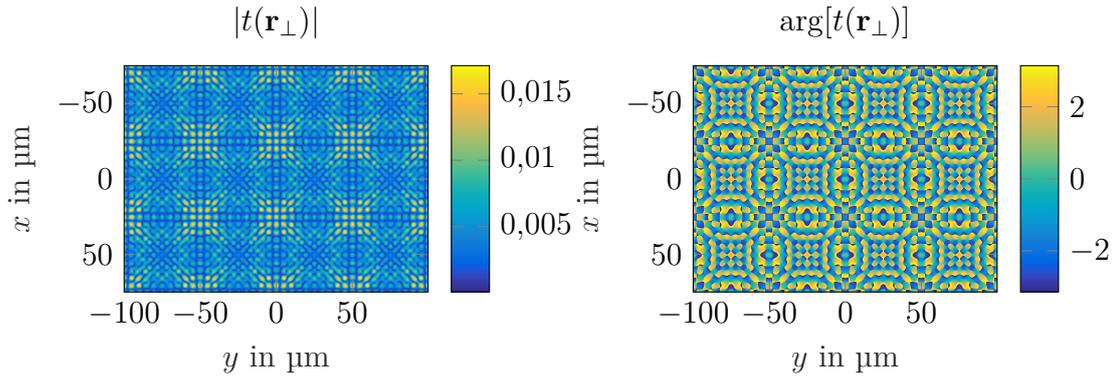


Abbildung 4.6: Phase und Amplitude eines diffraktiven Elements mit ringförmiger Pupille. Die maximale numerische Apertur der Pupille ist $NA_{\max} = 0,1$, die minimale numerische Apertur ist $NA_{\min} = 0,05$. In y -Richtung sind 4 Perioden und in x -Richtung sind 3 Perioden der Elementarzelle t gezeigt. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$.

Ringförmige Pupille mit variabler Ringbreite

Eine ringförmige Pupille mit variabler Ringbreite lässt sich aus der Subtraktion von zwei Kreisaperturen realisieren.

$$\tilde{U}_f(s_\perp) = U_0 \left[\text{circ}\left(\frac{s_\perp}{NA_{\max}}\right) - \text{circ}\left(\frac{s_\perp}{NA_{\min}}\right) \right] \quad (4.34)$$

Der Vorteil dieser Variante ist, dass die Fouriertransformierte des Elements leicht zu diskretisieren ist, somit spart man sich die explizite periodische Fortsetzung.

Exemplarisch ist in Abbildung 4.6 die Amplituden- und Phasenverteilung eines Elements mit $NA_{\max} = 0,1$ und $NA_{\min} = 0,05$ sowie einer Periode von $P = 50 \mu\text{m}$ und einer Wellenlänge $\lambda = 0,5 \mu\text{m}$ gezeigt.

In Abbildung 4.7 sieht man die zugehörige Amplitude in der Fokalebene zu diesem kontinuierlichen Element, Abbildung 4.8 zeigt einen Schnitt durch die yz -Ebene der Foki.

4.2.3 Ringförmige Beleuchtungsstrukturen

Ringförmige Beleuchtungsstrukturen finden oft bei Ausnutzung nicht-linearer Effekte eine große Rolle.

Ideale ringförmige Beleuchtung

Analog zur idealen punktförmigen Beleuchtung kann eine ideale ringförmige Beleuchtung mit Hilfe der Delta-Funktion definiert werden:

$$U_f(r_\perp) = U_0 \delta(|r_\perp| - r_0). \quad (4.35)$$

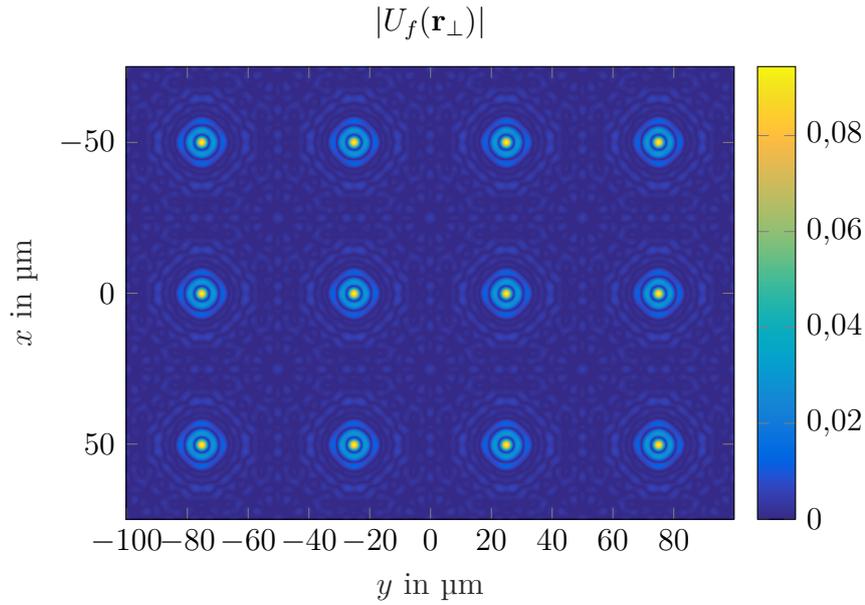


Abbildung 4.7: Amplitude in der Fokusebene eines diffraktiven Elements mit ringförmiger Pupille. Die maximale numerische Apertur der Pupille ist $NA_{\max} = 0,1$, die minimale numerische Apertur ist $NA_{\min} = 0,05$. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$.

Der Parameter r_0 beschreibt den Radius des Beleuchtungsringes. Das Frequenzspektrum lässt sich in diesem Fall analytisch bestimmen:

$$\tilde{U}_f(s_\perp) = 2\pi r_0 U_0 J_0(k_0 r_0 s_\perp). \quad (4.36)$$

Ähnlich zum Spektrum der idealen punktförmigen Beleuchtung ist auch hier das Spektrum nicht bandbegrenzt.

Bandbegrenzte ringförmige Beleuchtung

Das ideale ringförmige Spektrum lässt sich radialsymmetrisch bandbegrenzen:

$$\tilde{U}_f(s_\perp) = 2\pi r_0 U_0 J_0(k_0 r_0 s_\perp) \text{circ}\left(\frac{s_\perp}{NA_{\max}}\right). \quad (4.37)$$

Die Bandbegrenzung des Spektrums bestimmt die Breite des Rings.

In Abbildung 4.9 ist die Amplitude und Phase für ein Beispiелеlement mit $NA = 0,1$ und $r_0 = 10 \mu\text{m}$ gezeigt. Die zugehörigen Amplituden in der Fokalebene sind in Abbildung 4.10 gezeigt. Der YZ-Schnitt in Abbildung 4.11 zeigt, dass auf der optischen Achse vor und hinter dem Fokusring jeweils Maxima existieren.

Da sich die Fokalamplitude als Faltung eines Kreisrings mit der Fokalamplitude einer kreisförmigen homogenen Pupille schreiben lässt, kann man die Nebenmaxima im Ring als die Nebenmaxima der Fokalamplitude aus Gleichung 4.22 identifizieren.

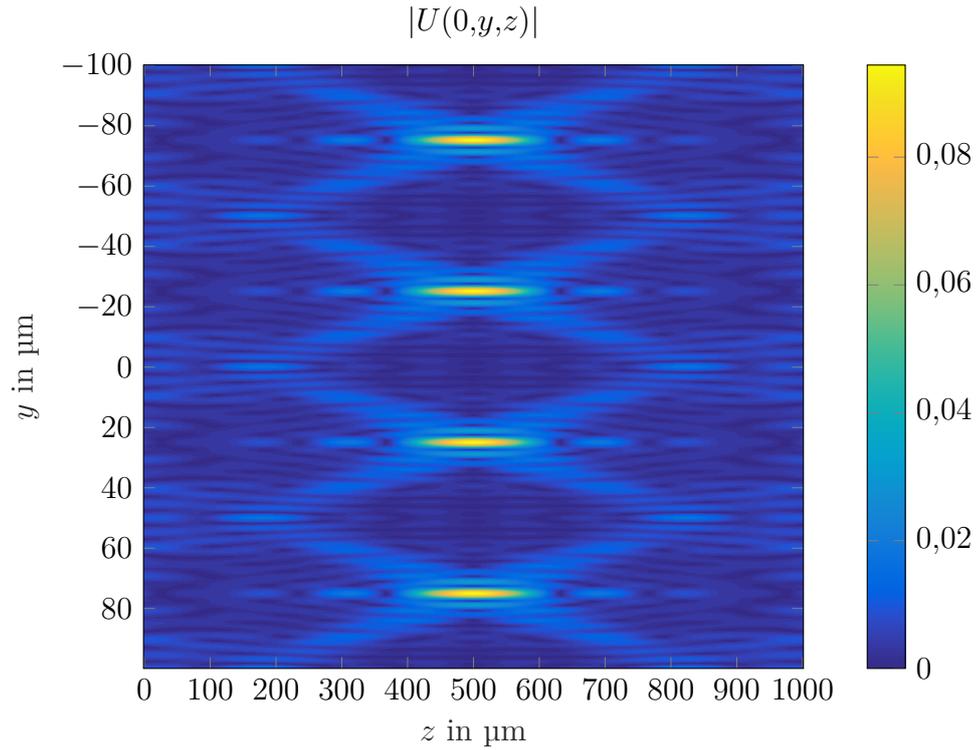


Abbildung 4.8: yz -Schnitt durch die Amplitude in der Fokusebene eines diffraktiven Elements mit ringförmiger Pupille. Die maximale numerische Apertur der Pupille ist $NA_{\max} = 0,1$, die minimale numerische Apertur ist $NA_{\min} = 0,05$. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$.

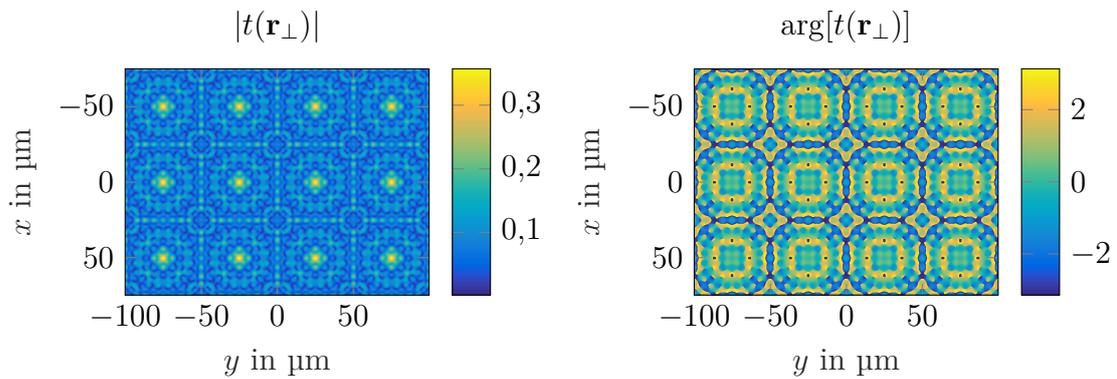


Abbildung 4.9: Phase und Amplitude eines diffraktiven Elements mit ringförmigem Fokus. Die numerische Apertur ist $0,1$. In y -Richtung sind 4 Perioden und in x -Richtung sind 3 Perioden der Elementarzelle t gezeigt. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$.

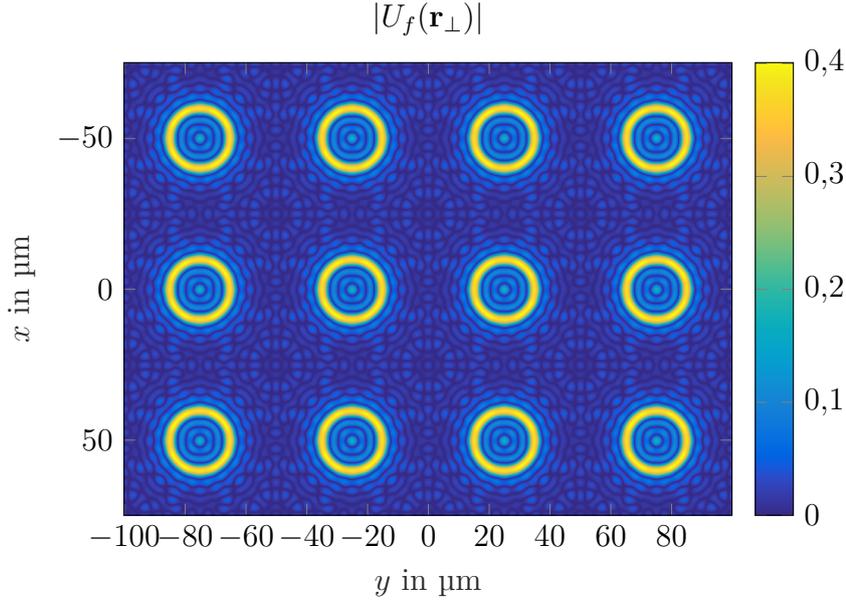


Abbildung 4.10: Amplitude in der Fokusebene eines diffraktiven Elements mit bandbegrenzten kreisförmigen Foki. Der Radius der Kreise beträgt $10\ \mu\text{m}$, die numerische Apertur wurde auf $\text{NA} = 0,1$ begrenzt. Die Wellenlänge ist $\lambda = 0,5\ \mu\text{m}$.

Der Spezialfall $r_0 = 0,61 \frac{\lambda}{\text{NA}}$ ist von besonderem Interesse, da hier ein Minimum in der Mitte auf der optischen Achse erzeugt wird und es keine weiteren Nebenmaxima innerhalb des Kreisrings gibt. Ein Beispiel für die Fokusgeometrie ist in Abbildung 4.12 und 4.13 gezeigt.

Die experimentelle Verifizierung erfolgte in Zusammenarbeit mit Gahbiche [33] im Rahmen einer Bachelorarbeit.

Vortex-Strahlen

Mit spiralen Phasenplatten der Form $e^{im\phi}$ lassen sich Vortex-Strahlen erzeugen [5, 44]. Der Parameter m gibt hier die topologische Ladung des Vortex-Strahls an. Um Vortex-Strahlen mit diffraktive Linsen mit überlappenden Aperturen zu erzeugen, kann man zum Beispiel an jede Periode eines fokussierenden Elements eine Spiral-Phase multiplizieren:

$$t_{\text{DOE}}(\mathbf{r}_{\perp}) = t_{\text{DOE,Kreisupille}}(r_{\perp})e^{im\phi}. \quad (4.38)$$

Dies hat jedoch den Nachteil, dass hier der Überlapp nicht für die Spiral-Phase gilt. Ein Beispiel für die Phase des Elements und die Amplitudenverteilung im Fokus ist in Abbildung 4.14 gezeigt. Es fällt auf, dass die Foki hier keine Rotationssymmetrie haben. Der zugehörige yz -Schnitt ist in Abbildung 4.15.

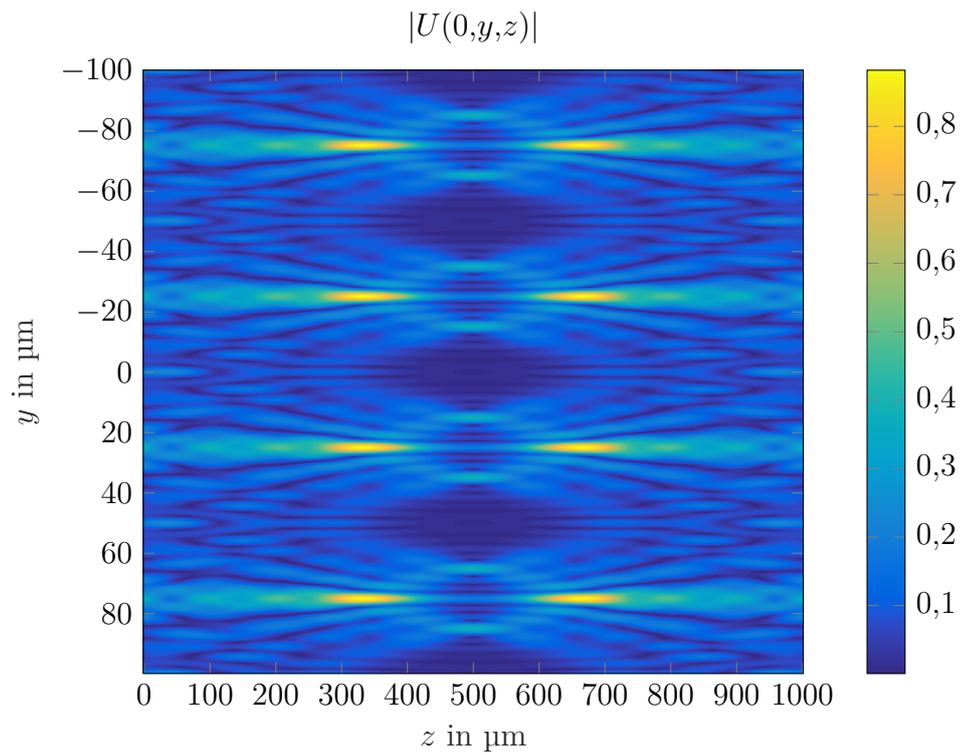


Abbildung 4.11: yz -Schnitt durch die Amplitude in der Fokusebene eines diffraktiven Elements mit bandbegrenzten kreisförmigen Foki. Der Radius der Kreise beträgt $10\ \mu\text{m}$, die numerische Apertur wurde auf $NA = 0,1$ begrenzt. Die Wellenlänge ist $\lambda = 0,5\ \mu\text{m}$.

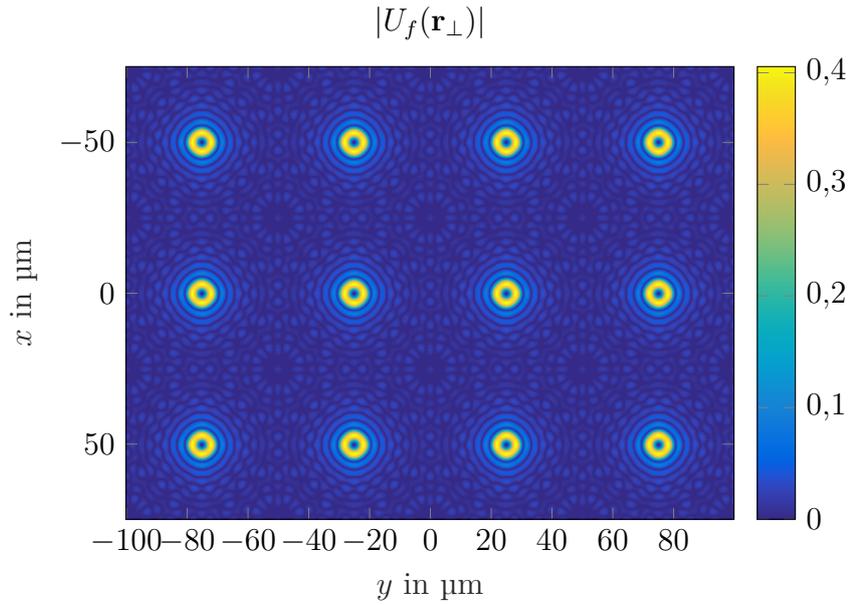


Abbildung 4.12: Amplitude in der Fokusebene eines diffraktiven Elements mit bandbegrenzten kreisförmigen Foki. Die numerische Apertur wurde auf $NA = 0,1$ begrenzt. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$. Der Radius der Kreise beträgt $r_0 = 0,61 \frac{\lambda}{NA} = 3,05 \mu\text{m}$.

Delta-Funktion mit Spiral-Phase Eine alternative Möglichkeit zur Erzeugung einer Fouriertransformierten mit Spiral-Phase ist ein deltaförmiger Fokus mit Spiral-Phase:

$$U_f(\mathbf{r}_\perp) = U_0 e^{im'\phi} \frac{\delta(r_\perp)}{r_\perp}. \quad (4.39)$$

Durch Separation von Radial- und Winkelanteil lässt sich das Fourier-Spektrum berechnen (siehe Gleichung 3.7). Für den Winkelanteil erhält man:

$$c_m = \frac{1}{2\pi} \int_0^{2\pi} U_\phi(\phi) e^{-im\phi} d\phi \quad (4.40)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} e^{im'\phi} e^{-im\phi} d\phi \quad (4.41)$$

$$= \delta_{mm'}. \quad (4.42)$$

Das Kronecker-Delta $\delta_{mm'}$ reduziert somit den Radialbeitrag auf nur eine Hankel-Transformation der Ordnung m' . Setzt man den Radialanteil $U_r(r_\perp) = U_0 \frac{\delta(r_\perp)}{r_\perp}$ ein,

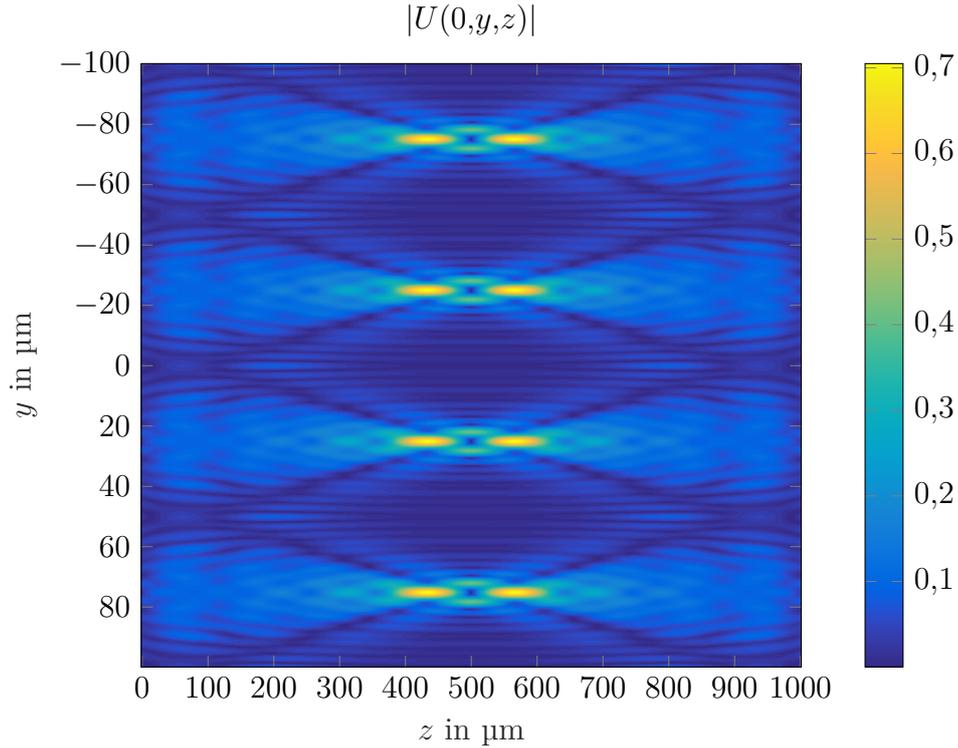


Abbildung 4.13: yz -Schnitt durch die Amplitude in der Fokusebene eines diffraktiven Elements mit bandbegrenzten kreisförmigen Foki. Die numerische Apertur wurde auf $NA = 0,1$ begrenzt. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$. Der Radius der Kreise beträgt $r_0 = 0,61 \frac{\lambda}{NA} = 3,05 \mu\text{m}$.

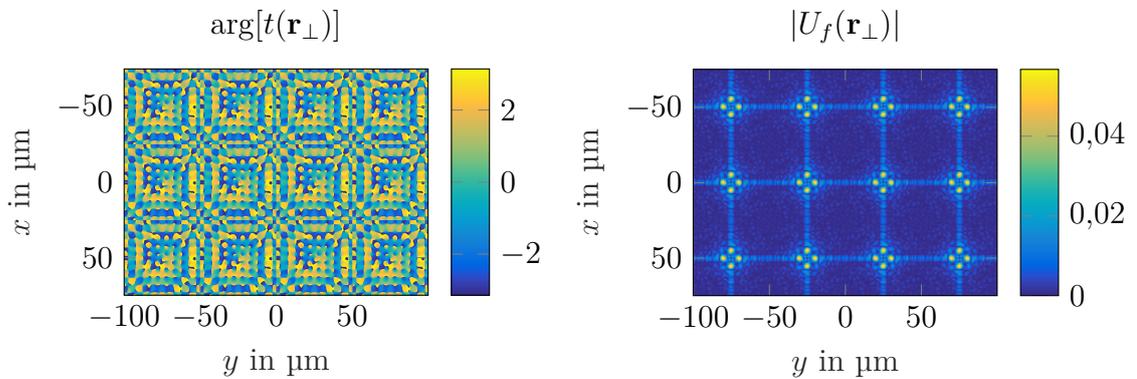


Abbildung 4.14: Links: Phase eines diffraktiven Elements mit homogener Pupille und zusätzlicher Spiralphase pro Gitterperiode. Die Spiralphase hat eine topologische Ladung von $m = 1$. Die numerische Apertur ist $NA = 0,1$. In y -Richtung sind 4 Perioden und in x -Richtung sind 3 Perioden der Elementarzelle t gezeigt. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$. Rechts: Amplitude in der Fokusebene dieses Elements.

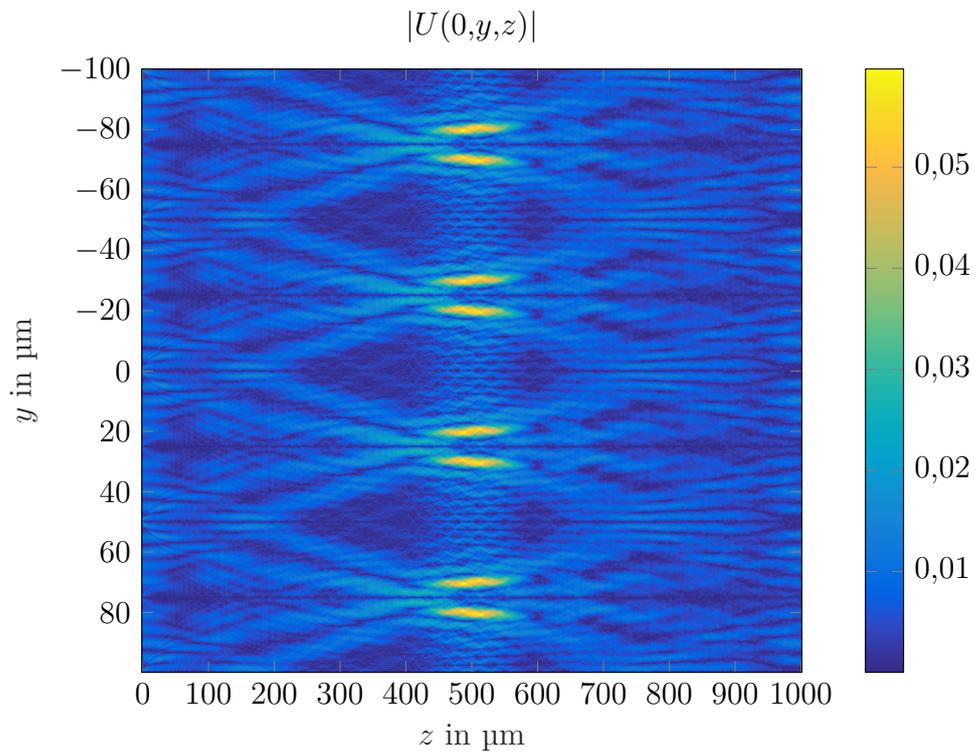


Abbildung 4.15: yz -Schnitt durch die Amplitude in der Fokusebene eines diffraktiven Elements mit homogener Pupille und zusätzlicher Spiralphase pro Gitterperiode. Die numerische Apertur ist $NA = 0,1$. Die Spiralphase hat eine topologische Ladung von $m = 1$. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$.

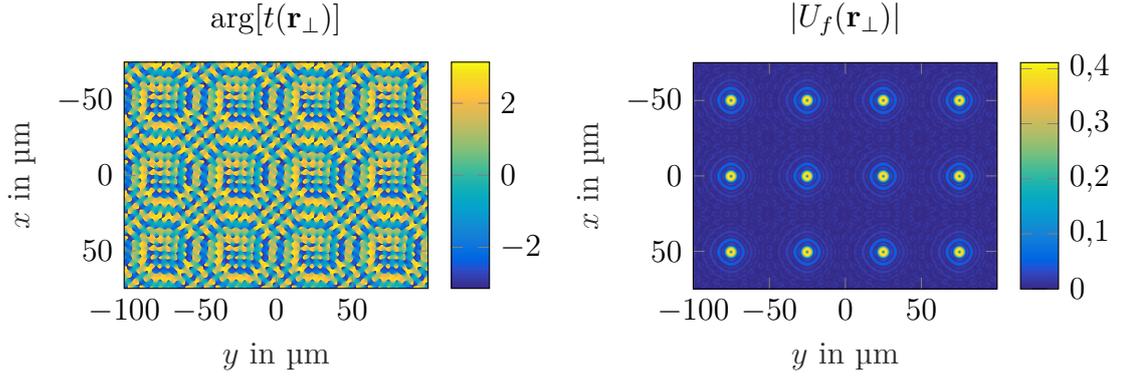


Abbildung 4.16: Links: Phase eines diffraktiven Elements mit homogener kreisförmiger Pupille und zusätzlicher Spiralphase. Die numerische Apertur ist $NA = 0,1$. Die Spiralphase hat eine topologische Ladung von $m = 1$. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$. Rechts: Amplitude in der Fokusebene dieses Elements.

so lässt sich die Hankel-Transformation leicht berechnen:

$$\mathcal{H}_m(U_r(r_\perp)) = 2\pi \int_0^\infty U_r(r_\perp) J_m(k_0 s_\perp r_\perp) r_\perp dr_\perp \quad (4.43)$$

$$= 2\pi \int_0^\infty U_0 \frac{\delta(r_\perp)}{r_\perp} J_m(k_0 s_\perp r_\perp) r_\perp dr_\perp \quad (4.44)$$

$$= 2\pi U_0 \quad (4.45)$$

Insgesamt hat man also ein Spektrum der Form

$$\tilde{U}(s_\perp, \chi) = \sum_{m=-\infty}^{\infty} c_m (-i)^m e^{im\chi} \mathcal{H}_m(U_r(r_\perp)) \quad (4.46)$$

$$= (-i)^{m'} e^{im'\chi} \mathcal{H}_{m'}(U_r(r_\perp)) \quad (4.47)$$

$$= (-i)^{m'} e^{im'\chi} U_0 \quad (4.48)$$

Dieses Spektrum lässt sich wieder bandbegrenzen.

$$\tilde{U}(s_\perp, \chi) = 2\pi U_0 (-i)^{m'} e^{im'\chi} \text{circ}\left(\frac{s_\perp}{NA}\right) \quad (4.49)$$

In Abbildung 4.16 sieht man die Phase in der Elementebene und die Amplitudenverteilung in der Fokalebene für den Fall $m' = 1$. Der zugehörige yz -Schnitt ist in Abbildung 4.17 gezeigt.

Es fällt auf, dass die Vortex-Strahlen keine Intensitätsmaxima auf der optischen Achse haben, dies unterscheidet sie deutlich von den Strahlen mit ringförmiger Beleuchtung. In STED-Anwendungen sind daher die ringförmigen Beleuchtungen aus Abschnitt 4.2.3 den Vortex-Strahlen vorzuziehen, da diese auch eine Einschränkung in z -Richtung erzeugen.

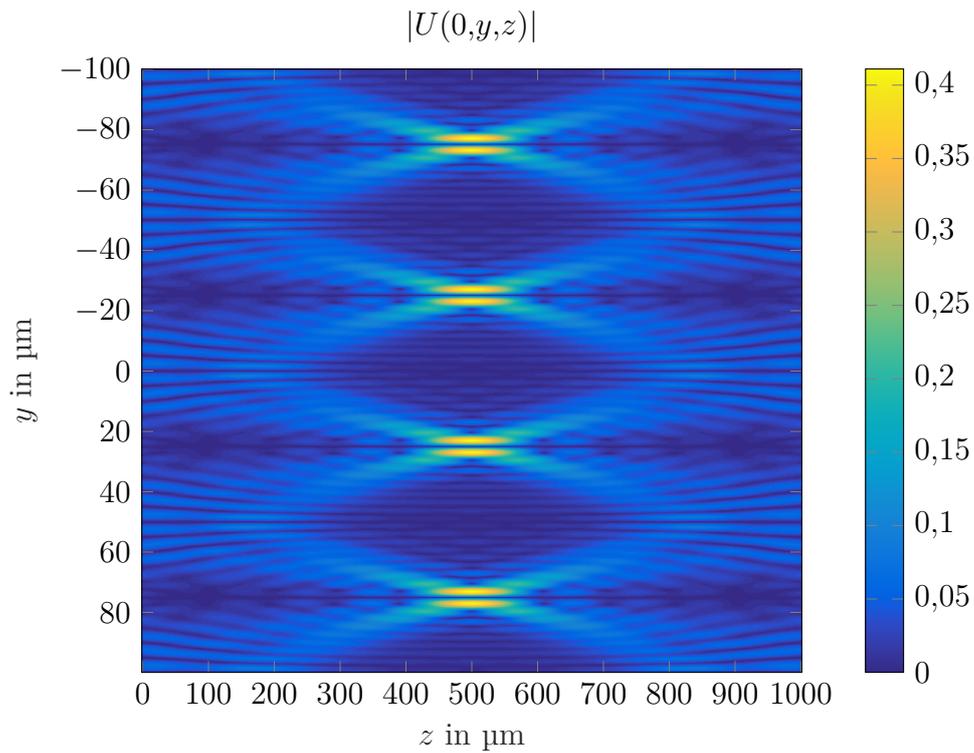


Abbildung 4.17: yz -Schnitt durch die Amplitude in der Fokusebene eines diffraktiven Elements mit homogener kreisförmiger Pupille und zusätzlicher Spiralphase. Die numerische Apertur ist $NA = 0,1$. Die Spiralphase hat eine topologische Ladung von $m = 1$. Die Wellenlänge ist $\lambda = 0,5 \mu\text{m}$.

4.3 Erweiterungen des Design-Prozesses

Der Design-Prozess lässt sich für bestimmte Spezialanwendungen erweitern. Im Folgenden wird die Überlagerung von mehreren Beleuchtungsstrukturen und die Integration einer Deckglaskorrektur behandelt.

4.3.1 Überlagerung von mehreren Beleuchtungsstrukturen

Bei der Überlagerung von mehreren Beleuchtungsstrukturen lässt sich eine laterale Überlagerung in der Fokusebene von einer axialen Überlagerung unterscheiden.

Laterale Überlagerung

Durch Multiplikation einer linearen Phase $e^{i\mathbf{k}_\perp \cdot \Delta \mathbf{r}_\perp}$ an das Spektrum \tilde{t} lässt sich die Beleuchtungsstruktur $U(\mathbf{r}_\perp)$ um $\Delta \mathbf{r}_\perp$ lateral verschieben.

$$\mathcal{F}[t(\mathbf{r}_\perp - \Delta \mathbf{r}_\perp)] = e^{-i\mathbf{k}_\perp \cdot \Delta \mathbf{r}_\perp} \mathcal{F}[t(\mathbf{r}_\perp)] \quad (4.50)$$

Durch Addition von mehreren so modulierten Spektren lassen sich somit Beleuchtungsstrukturen mit mehreren Foki in einer Periode realisieren [40]. Bei der Addition lassen sich zusätzlich Amplitudengewichte A_j und Phasenverschiebungen ϕ_j zwischen den einzelnen Beiträgen festlegen. Aufgrund der Linearität der Fourier-Transformation erhält man somit in der Fokalebene eine Überlagerung der Beleuchtungsstrukturen U_j .

$$U_f(\mathbf{r}_\perp) = \sum_j A_j e^{i\phi_j} U_j(\mathbf{r}_\perp) \quad (4.51)$$

Für das Gesamtspektrum gilt

$$\tilde{t} = \sum_j A_j \tilde{t}_j e^{i\phi_j} \quad (4.52)$$

Für den Spezialfall zweier gleicher rotationssymmetrischer Spektren mit relativer Phasenverschiebung von $\Delta\phi = \pi$ lässt sich ein Strahl mit einer Nullamplitude auf der Elementachse erzeugen

$$U_f(\mathbf{r}_\perp) = U_1(\mathbf{r}_\perp - \Delta \mathbf{r}_\perp) + e^{i\Delta\phi} U_1(\mathbf{r}_\perp + \Delta \mathbf{r}_\perp) \quad (4.53)$$

$$= U_1(\mathbf{r}_\perp - \Delta \mathbf{r}_\perp) - U_1(\mathbf{r}_\perp + \Delta \mathbf{r}_\perp) \quad (4.54)$$

$$\Rightarrow U_f(\mathbf{0}) = 0 \quad (4.55)$$

Axiale Überlagerung

Eine axiale Überlagerung der Beiträge \tilde{U}_j ist durch separate Rückpropagation um die Distanz f_j in die Elementebene und anschließende Addition der Spektren oder der komplexen Amplituden möglich.

$$\tilde{t} = \sum_j \tilde{U}_j(s_\perp) e^{-ik_0 f_j \sqrt{n^2 - s_\perp^2}} \quad (4.56)$$

Der Fokalverlauf $U(\mathbf{r})$ eines so erzeugten Elements ist die kohärente Addition der komplexen Amplituden der Fokalverläufe der einzelnen Beiträge $U_j(\mathbf{r})$.

$$U(\mathbf{r}) = \sum_j U_j(\mathbf{r}) \quad (4.57)$$

Ein Beispiel für die axiale Überlagerung ist in Abbildung 4.18 gezeigt.

4.3.2 Deckglaskorrektur

Eine Deckglaskorrektur ist relativ leicht in den Design-Prozess integrierbar [40]. Hierzu wird die Propagation zwischen dem diffraktiven Element und der Beleuchtungsstruktur in Teilpropagationen in den jeweiligen homogenen Medien unterteilt. Die Fokusdistanz f setzt sich nun aus Teildistanzen z_j im Medium mit Brechungsindex n_j zusammen.

$$f = \sum_j z_j \quad (4.58)$$

Für das diffraktive Element mit Deckglaskorrektur ergibt sich:

$$\tilde{t}_{\text{DOE}}(\mathbf{s}_{\perp}) = \tilde{U}_f(\mathbf{s}_{\perp}) \prod_j e^{-ik_0 z_j \sqrt{n_j^2 - \mathbf{s}_{\perp}^2}}. \quad (4.59)$$

In Abbildung 4.19 sieht man die Auswirkungen der Propagation durch ein Deckglas der Dicke $170 \mu\text{m}$. Das erzeugende diffraktive Element hat eine numerische Apertur von $\text{NA} = 0,5$. Der Fokus des nicht für das Deckglas korrigierten Elements zeigt deutliche Abberationen und eine allgemeine laterale Verschiebung in z-Richtung. Der Fokus des korrigierten Elements zeigt hingegen ein beugungsbegrenztes Profil. Die Asymmetrie des Fokus um die Fokalebene ist dadurch zu begründen, dass der linke Teil des Fokus im Deckglas befindet.

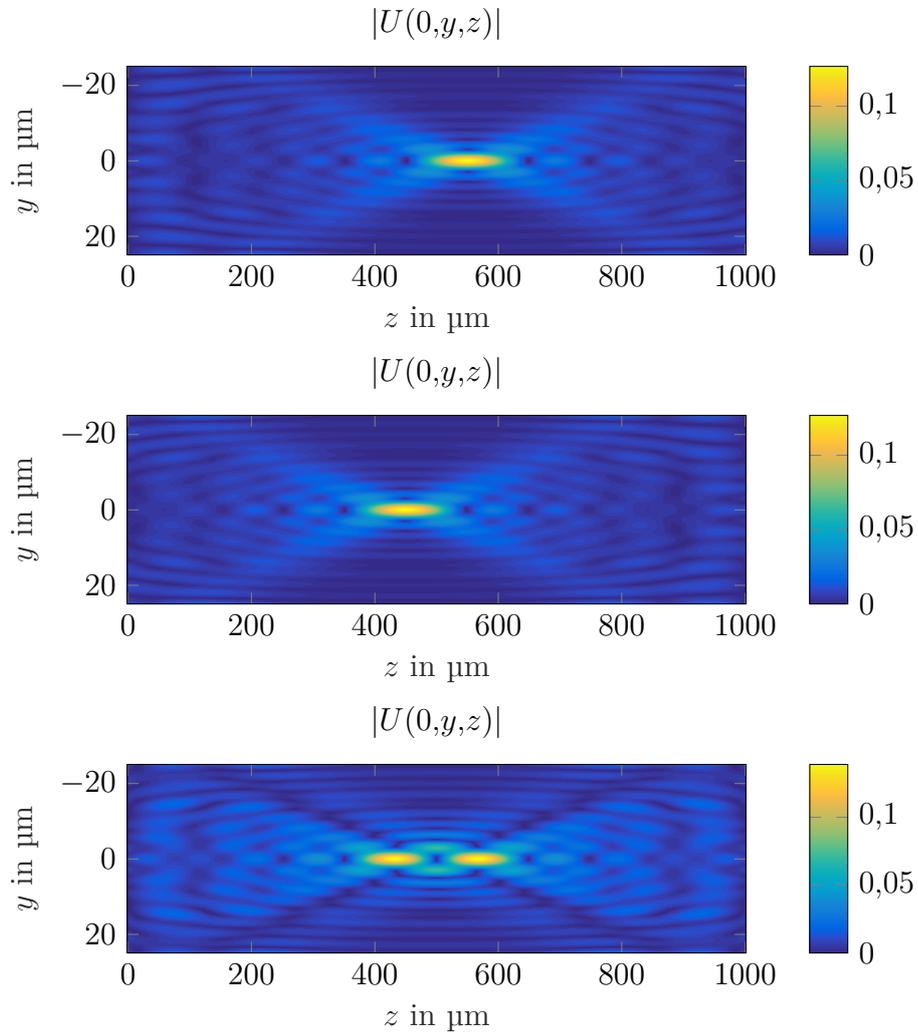


Abbildung 4.18: Oben: Fokalverlauf eines diffraktiven Elements mit Brennweite $f = 450 \mu\text{m}$. Mitte: Fokalverlauf eines diffraktiven Elements mit $f = 550 \mu\text{m}$. Unten: Fokalverlauf eines Elements dass die beiden Brennweiten $f_1 = 450 \mu\text{m}$ und $f_2 = 450 \mu\text{m}$ in einem Element kombiniert.

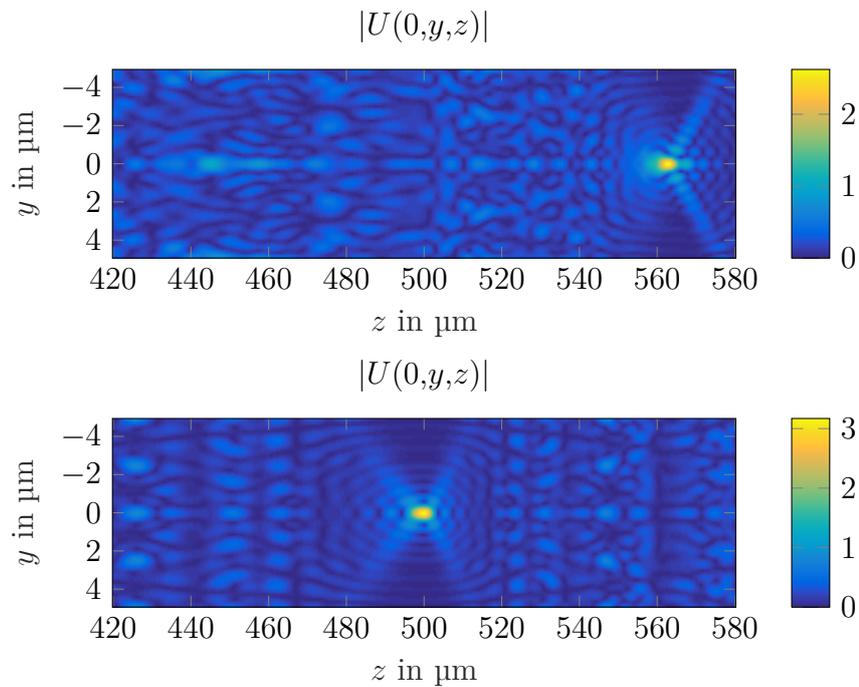


Abbildung 4.19: Oben: Fokalverlauf eines diffraktiven Elements ohne Deckglas Korrektur mit $NA = 0,5$ und Brennweite $f = 500 \mu\text{m}$. Im Bereich $330 \mu\text{m}$ und $500 \mu\text{m}$ befindet sich ein Deckglas der Dicke $170 \mu\text{m}$. Unten: Fokalverlauf eines diffraktiven Elements mit Deckglaskorrektur mit $NA = 0,5$ und Brennweite $f = 500 \mu\text{m}$. Wieder befindet sich im Bereich $330 \mu\text{m}$ und $500 \mu\text{m}$ befindet sich ein Deckglas der Dicke $170 \mu\text{m}$.

4.4 Diffraktive Linsen als dünnes Element

Die Wirkung der diffraktiven Linsen lässt sich über die Transmissionsfunktion eines dünnen Elements beschreiben. Diese Transmissionsfunktion t_{DOE} ist eine komplexe Größe die sich in eine positive reelle Amplitudenverteilung A_{DOE} und eine Phasenverteilung ϕ_{DOE} aufsplitten lässt.

$$t_{\text{DOE}}(\mathbf{r}_{\perp}) = A_{\text{DOE}}(\mathbf{r}_{\perp})e^{i\phi_{\text{DOE}}(\mathbf{r}_{\perp})} \quad (4.60)$$

Die Amplitudenverteilung $A_{\text{DOE}}(\mathbf{r}_{\perp})$ ist ein Maß für die Abschwächung am Ort \mathbf{r}_{\perp} . Die Phasenverteilung $\phi(\mathbf{r})$ lässt sich auf einen Wertebereich der Breite 2π begrenzt.

4.4.1 Modifikationen von dünnen diffraktiven Elementen

Insbesondere für die Herstellung eines diffraktiven Elements sind Modifikationen der Phase und Amplitude interessant, da durch bestimmte Modifikationen der Aufwand bei der Herstellung deutlich reduziert werden kann. Das modifizierte Phasenelement wird nun beschrieben durch

$$t_{\text{DOE}}^{\text{mod}}(\mathbf{r}_{\perp}) = A_{\text{DOE}}^{\text{mod}}(\mathbf{r}_{\perp})e^{i\phi_{\text{DOE}}^{\text{mod}}(\mathbf{r}_{\perp})}. \quad (4.61)$$

Von besonderem Interesse sind hierbei die Nivellierung der Amplitude und die Diskretisierung der Phase.

Nivellierung der Amplitude

Beim Nivellieren der Amplitude wird $A_{\text{DOE}} = 1$ angenommen. Dies entspricht dem Fall, dass kein Licht beim Durchgang durch das diffraktive Element absorbiert wird.

Diskretisierung der Phase

Bei der Diskretisierung der Phase wird die Phase $\phi_{\text{DOE}}(\mathbf{r}_{\perp})$ in N konstante Phasen ϕ_j aufgeteilt:

$$\phi_{\text{DOE}}^N(\mathbf{r}_{\perp}) = \begin{cases} \phi_1 & \text{für } \phi_{1,\min} \leq \phi_{\text{DOE}}(\mathbf{r}_{\perp}) < \phi_{1,\max} \\ \vdots & \\ \phi_j & \text{für } \phi_{j,\min} \leq \phi_{\text{DOE}}(\mathbf{r}_{\perp}) < \phi_{j,\max} \\ \vdots & \\ \phi_N & \text{für } \phi_{N,\min} \leq \phi_{\text{DOE}}(\mathbf{r}_{\perp}) < \phi_{N,\max} \end{cases} \quad (4.62)$$

Der einfachsten Fall sind zwei Diskretisierungsstufen, die jeweils das arithmetische Mittel der oberen und unteren Grenze sind:

$$\phi_{\text{DOE}}^{\text{binär}}(\mathbf{r}_{\perp}) = \begin{cases} -\frac{\pi}{2} & \text{für } -\pi \leq \phi_{\text{DOE}}(\mathbf{r}_{\perp}) < 0 \\ +\frac{\pi}{2} & \text{für } 0 \leq \phi_{\text{DOE}}(\mathbf{r}_{\perp}) < +\pi \end{cases} \quad (4.63)$$

Durch Multiplikation der Transmissionsfunktion t_{DOE} mit der Phase $e^{i2\pi\epsilon}$ vor dem Ausführen der Phasendiskretisierung lässt sich der Sprung in der Phasenverteilung verschieben.

$$t_{\text{DOE},\epsilon} = t_{\text{DOE}}e^{i2\pi\epsilon} \quad (4.64)$$

Somit ergibt sich

$$\phi_{\text{DOE}}^{\text{binär}, \epsilon}(\mathbf{r}_{\perp}) = \begin{cases} -\frac{\pi}{2} & \text{für } -\pi \leq \phi_{\text{DOE},\epsilon}(\mathbf{r}_{\perp}) < 0 \\ +\frac{\pi}{2} & \text{für } 0 \leq \phi_{\text{DOE},\epsilon}(\mathbf{r}_{\perp}) < +\pi \end{cases}. \quad (4.65)$$

4.4.2 Dünne diffraktive Linsen für zwei verschiedene Beleuchtungswellenlängen

Neben der Kombination von mehreren Beleuchtungsstrukturen der gleichen Wellenlänge ist es auch möglich ein diffraktives Element mit zwei unterschiedlichen Wellenlängen zu betreiben, wobei für jede Wellenlänge eine eigene Beleuchtungsstruktur definiert wird. Genutzt wird diese Technik um zum Beispiel diffraktive Linsen in der konfokalen Fluoreszenzmikroskopie zu benutzen, denn hier unterscheiden sich das Anregungs- und das Fluoreszenzlicht in der Wellenlänge [47]. Ein weitere Anwendung ist, dass eine Linse gleichzeitig zur Erzeugung einer optischen Pinzette und als Abbildungslinse verwendet wird [48].

Die Methode zur Überlagerung zwei diffraktiver Elemente für unterschiedliche Wellenlängen beruht darauf, dass eine Phasenverzögerung, durch eine bestimmte Ausbreitungsstanz d in einem homogenen Medium mit Brechungsindex n gegenüber der Ausbreitung in Luft, erzeugt wird und dass hierbei vielfache der Dicke $d_0 = \frac{\lambda_1}{n_{\lambda_1}-1}$ zu d addiert werden können ohne dass sich das Beugungsbild für die Wellenlänge λ_1 ändert [78]. Für eine zweite Wellenlänge λ_2 hingegen wird durch eine Verlängerung der Propagation im Medium n um die Distanz d_0 eine Phasenverschiebung ungleich 2π induziert. Durch geschicktes Auswählen der vielfachen der Dicke d_0 lässt sich somit näherungsweise eine bestimmte Phasendifferenz für eine Wellenlänge λ_2 einstellen.

Sei zum Beispiel d_1 die Propagationslänge im Medium n um die Designphase ϕ_1 des ersten diffraktiven Elements für die Wellenlänge λ_1 zu erzeugen, dann gilt $\phi_1 = \frac{(n_{\lambda_1}-1)d_1 2\pi}{\lambda_1}$. Andererseits sei ϕ_2 die Designphase des zweiten diffraktiven Elements für die Wellenlänge λ_2 . Eine Ausbreitung von Licht mit der Wellenlänge λ_2 um die Länge $d_1 + m d_0$ durch ein Medium mit Brechungsindex n erzeugt Phasendifferenz von $\frac{(n_{\lambda_2}-1)(d_1+m d_0)2\pi}{\lambda_2}$. Für das Beugungsmuster sind jedoch nur Phasenwerte modulo 2π relevant. Nun wird das m gesucht bei dem $\phi_2 = \text{mod}_{2\pi}\left(\frac{(n_{\lambda_1}-1)(d_1+m d_0)2\pi}{\lambda_2}\right)$ gilt oder die Abweichung $\left|\phi_2 - \text{mod}_{2\pi}\left(\frac{(n_{\lambda_1}-1)(d_1+m d_0)2\pi}{\lambda_2}\right)\right|$ minimal ist.

4.5 Geometrie des Fokus für diffraktive Linsen mit homogener kreisförmiger Pupille

Die weitere Arbeit beschäftigt sich mit diffraktiven Linsen mit homogener kreisförmiger Pupille, deshalb wird in diesem Abschnitt die Geometrie für diesen Spezialfall genauer untersucht.

Amplitude in der Fokalebene

Nach Konstruktion ist die Amplitude in der Fokalebene gegeben durch

$$U_f(r_\perp) = U_0 \frac{k_0 \text{NA}}{2\pi} \frac{J_1(k_0 r_\perp \text{NA})}{r_\perp}. \quad (4.66)$$

Wenn bei der Diskretisierung im Frequenzraum die Periode groß genug gewählt ist, dann ist der Fokus nicht durch Beiträge von Nachbarperioden gestört und die ± 1 . Nullstelle um den Fokus liegen bei $\pm 3,83 \frac{1}{k_0 \text{NA}} = 0,61 \frac{\lambda}{\text{NA}}$. Die Breite auf halber Höhe ist $\delta x_{\text{FWHM}} = 0,514 \frac{\lambda}{\text{NA}}$.

Amplitude auf der optischen Achse

Das Profil auf der optischen Achse lässt sich analytisch bestimmen. Allgemein gilt

$$U(r_\perp, z) = \frac{k_0^2}{2\pi} \int_0^\infty U_0 \text{circ}\left(\frac{s_\perp}{\text{NA}}\right) e^{-ik_0 f \sqrt{n^2 - s_\perp^2}} e^{ik_0 z \sqrt{n^2 - s_\perp^2}} J_0(k_0 s_\perp r_\perp) s_\perp ds_\perp. \quad (4.67)$$

Auf der optischen Achse vereinfacht sich dies zu

$$U(0, z) = \frac{k_0^2}{2\pi} \int_0^\infty U_0 \text{circ}\left(\frac{s_\perp}{\text{NA}}\right) e^{ik_0(z-f)\sqrt{n^2 - s_\perp^2}} s_\perp ds_\perp \quad (4.68)$$

$$= \frac{k_0^2}{2\pi} \int_0^{\text{NA}} U_0 e^{ik_0(z-f)\sqrt{n^2 - s_\perp^2}} s_\perp ds_\perp. \quad (4.69)$$

Mit den Ersetzungen $s' = \sqrt{n^2 - s_\perp^2}$, $s_\perp = \sqrt{n^2 - s'^2}$, $\frac{ds_\perp}{ds'} = -\frac{s'}{\sqrt{n^2 - s'^2}}$, $a = k_0(z - f)$, $s'_{\min} = n$ und $s'_{\max} = \sqrt{n^2 - \text{NA}^2}$ lässt sich das Integral umschreiben und analytisch für $f \neq z$ lösen:

$$U(0, z) = -\frac{k_0^2}{2\pi} U_0 \int_{s'_{\min}}^{s'_{\max}} e^{ias'} s' ds' \quad (4.70)$$

$$= \frac{k_0^2}{2\pi a^2} U_0 \left((ias'_{\max} - 1) e^{ias'_{\max}} - (ias'_{\min} - 1) e^{ias'_{\min}} \right) \quad (4.71)$$

$$= \frac{U_0}{2\pi(z-f)^2} \left(ik_0(z-f)\sqrt{n^2 - \text{NA}^2} - 1 \right) e^{ik_0(z-f)\sqrt{n^2 - \text{NA}^2}} - \frac{U_0}{2\pi(z-f)^2} (ik_0(z-f)n - 1) e^{ik_0(z-f)n}. \quad (4.72)$$

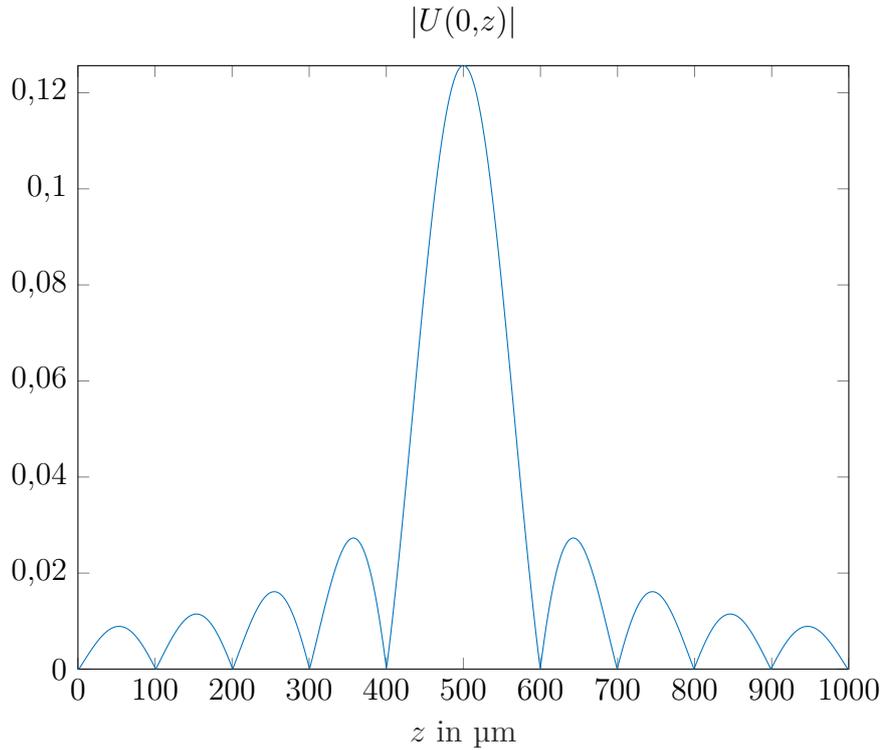


Abbildung 4.20: Analytisches xz -Profil durch den Fokus eines diffraktiven Elements mit homogener Pupille. Die Elementarperiode ist quadratisch und hat eine Kantenlänge von $50\ \mu\text{m}$. Der Fokus hat eine numerische Apertur von $0,1$ und liegt für die Wellenlänge $0,5\ \mu\text{m}$ bei $500\ \mu\text{m}$.

Für $f = z$ ergibt sich

$$U(0,f) = U_0 \frac{k_0^2}{4\pi} \text{NA}^2. \quad (4.73)$$

In Abbildung 4.20 ist der Betrag der Amplitude für ein Beispiel-Element mit numerischer Apertur von $\text{NA} = 0,1$, Fokusslänge $f = 500\ \mu\text{m}$ bei einer Wellenlänge von $\lambda = 0,5\ \mu\text{m}$ gezeigt. Für die Intensität auf der optischen Achse gilt:

$$I(0,z) = I_0 \left[\frac{k_0^4}{2\pi^2 a^4} + \frac{k_0^4}{4\pi^2 a^2} (s_{\max}'^2 + s_{\min}'^2) - \frac{k_0^4}{2\pi^2 a^4} (1 + a^2 s_{\max}' s_{\min}') \cos(a(s_{\max}' - s_{\min}')) + \frac{k_0^4}{2\pi^2 a^3} (s_{\min}' - s_{\max}') \sin(a(s_{\max}' - s_{\min}')) \right] \quad (4.74)$$

Dabei ist $I_0 = |U_0|^2$.

Axiale Länge des Fokus

Im Fall kleiner numerischer Aperturen liegen das ± 1 . Minimum um den Fokus bei $z \approx f \pm \frac{n4\pi}{k_0 \text{NA}^2} = f \pm \frac{2n\lambda}{\text{NA}^2}$. Für kleine numerische Aperturen sind die Minima gleichzeitig Nullstellen. Hauptsächlich verantwortlich hierfür sind folgende Terme:

$$\frac{I'(0,z)}{I_0} \approx \frac{k_0^4}{4\pi^2 a^2} (s_{\max}'^2 + s_{\min}'^2) - \frac{k_0^4}{2\pi^2 a^4} a^2 s_{\max}' s_{\min}' \cos(a(s_{\max}' - s_{\min}')) \quad (4.75)$$

Setzt man diesen Term gleich Null, erhält man nach Rücksubstitution

$$n^2 - \frac{1}{2} \text{NA}^2 = \sqrt{n^2 - \text{NA}^2} n \cos\left(k_0(z-f)\left(\sqrt{n^2 - \text{NA}^2} - n\right)\right) \quad (4.76)$$

$$\frac{n^2 - \frac{1}{2} \text{NA}^2}{\sqrt{n^2 - \text{NA}^2} n} = \cos\left(k_0(z-f)\left(\sqrt{n^2 - \text{NA}^2} - n\right)\right). \quad (4.77)$$

Für kleine numerische Aperturen gilt $\sqrt{n^2 - \text{NA}^2} \approx n - \frac{\text{NA}^2}{2n}$, somit vereinfacht sich die Gleichung zu

$$1 = \cos\left(k_0(z-f)\left(-\frac{\text{NA}^2}{2n}\right)\right). \quad (4.78)$$

Diese Gleichung ist für $z = f \pm \frac{n4\pi}{k_0 \text{NA}^2}$ erfüllt, dies entspricht den Nullstellen um den Fokus.

Für große numerische Aperturen sind die Minima keine Nullstellen. Insbesondere im Grenzfall $\text{NA} = n$ werden aus den aus den Minima der Amplitude Wendepunkte. Die normierte Intensität ist somit gegeben durch

$$\frac{I(0,z)}{I_0} = \frac{k_0^4}{2\pi^2 a^4} + \frac{k_0^4}{4\pi^2 a^2} n^2 - \frac{k_0^4}{2\pi^2 a^4} \cos(an) - \frac{k_0^4}{2\pi^2 a^3} n \sin(an). \quad (4.79)$$

Leitet man diese Gleichung nach $\Delta z = z - f$ ab und setzt das Ergebnis Null, so erhält man

$$\frac{\partial I}{\partial \Delta z} = -\frac{\left(k_0 n \Delta z \cos\left(\frac{k_0 n \Delta z}{2}\right) - 2 \sin\left(\frac{k_0 n \Delta z}{2}\right)\right)^2}{\pi^2 \Delta z^5} = 0. \quad (4.80)$$

Diese Gleichung ist für $z \neq f$ nur erfüllt, wenn die Differenz in der Klammer Null ist, also muss gelten

$$k_0 n \Delta z \cos\left(\frac{k_0 n \Delta z}{2}\right) = 2 \sin\left(\frac{k_0 n \Delta z}{2}\right) \quad (4.81)$$

$$\left(\frac{k_0 n \Delta z}{2}\right) = \tan\left(\frac{k_0 n \Delta z}{2}\right). \quad (4.82)$$

Dies ist erfüllt für $\Delta z = \pm 1,430 \frac{\lambda}{n}$. Der Abstand zwischen den zwei Nullstellen definiert somit eine axiale Länge des Fokus für den Grenzfall $\text{NA} = n$. Vergleicht man diese axiale Länge mit der Nullstellenbreite der paraxialen Rechnung so fällt auf, dass die paraxiale Rechnung für den nicht paraxialen Grenzfall eine größere axiale Länge des Fokus liefert.

Energiefluss durch Fokusellipsoid

Speziell für niedrige numerische Aperturen lässt sich der Energiefluss durch das Fokusellipsoid über die Nullstellen der Amplitude bestimmen:

$$P_{\text{Fokusellipsoid}} = \int_V |U(\mathbf{r}_\perp, z)|^2 \text{circ}\left(\frac{|\mathbf{r}_\perp|}{0.61 \frac{\lambda}{\text{NA}}}\right) \text{circ}\left(\frac{|z - f|}{2 \frac{\lambda}{\text{NA}^2}}\right) dV \quad (4.83)$$

Für diffraktive Linsen mit höherer Apertur muss die Begrenzung in axialer Richtung angepasst werden.

4.6 Fazit

In diesem Abschnitt wurde das Design von diffraktiven Linsenarrays mit überlappenden Aperturen vorgestellt, hierbei wurde explizit auf zwei Methoden zur Erzeugung des Überlapps der Aperturen aufgezeigt.

Eine allgemeine Methode zur Erzeugung von beliebigen Beleuchtungsstrukturen wurde vorgestellt. Die Spezialfälle punktförmiger und ringförmiger Beleuchtungsstrukturen wurden ausführlich betrachtet. In skalaren Simulationen wurden dabei jeweils Beispiele für die lateralen und axialen Fokalprofile, sowie die erzeugende Transmissionsfunktion aufgezeigt. Explizit wurden diffraktive Linsen mit überlappenden Aperturen und ringförmiger Pupille hier zum ersten Mal beschrieben. Auch die explizite Beschreibung von ringförmigen Beleuchtungsstrukturen mit überlappenden Aperturen ist neu.

Erweiterungen des Design-Prozesses wurden vorgestellt, die eine Überlagerung von mehreren Beleuchtungsstrukturen erlauben. Insbesondere wurde hier die axiale Überlagerung zum ersten Mal explizit beschrieben. Des Weiteren wurde eine Methode zur Deckglaskorrektur vorgestellt, die auf die Korrektur von beliebigen Schichtsystemen mit in lateraler Richtung konstantem Brechungsindex erweiterbar ist.

Anschließend wurde auf die Realisierung der diffraktiven Linsen als dünnes Element eingegangen und es wurden die typische Modifikationen vorgestellt, die die Herstellung erleichtern.

Zum Schluss wurde die Geometrie des Fokus für die ideale diffraktive Linse mit homogener kreisförmiger Pupille betrachtet. Im Folgenden Kapitel wird der Einfluss von Abweichungen und bestimmten Modifikationen der idealen diffraktiven Linse auf den Fokus untersucht.

5 Evaluation von diffraktiven Linsen

Um die Eigenschaften der im vorherigen Kapitel vorgestellten diffraktiven Linsen mit überlappenden Aperturen quantitativ evaluieren zu können werden in Abschnitt 5.1 zunächst die Eigenschaften von idealen fokussierenden Lichtverteilungen betrachtet und anschließend mit den diffraktiven Linsenarrays verglichen. Anschließend werden in den Abschnitten 5.2 und 5.4 die Auswirkungen von bestimmten Modifikationen wie der Phasendiskretisierung und Amplitudennivellierung untersucht und der Einfluss der Pixelierung auf den Fokus betrachtet. Da im Design eine unendliche Periode vorausgesetzt wurde, wird der Einfluss der Endlichkeit des Elements in Abschnitt 5.5 untersucht. Um Fehler von real fabrizierten diffraktiven Linsen evaluieren zu können, wird die Robustheit gegen Abweichungen von den Design-Parametern in Abschnitt 5.6 untersucht. Zum Abschluss werden in Abschnitt 5.7 noch einmal die Auswirkungen und Limitationen der Parameterwahl für die diffraktiven Elemente diskutiert und ein Leitfaden für das Design diffraktiver Linsen vorgeschlagen.

5.1 Theoretische Grenzen für die Effizienz von fokussierenden Lichtverteilungen

Viele nicht lineare optische Prozesse wie Photo-Polymerisation oder Fluoreszenzanregung benötigen einen spezifischen lokalen Energieeintrag bzw. Intensität [72]. Hierbei muss man häufig einen lokalen Schwellwert erreichen, aber gleichzeitig soll der Energieeintrag bzw. der Energiefluss durch das restliche Medium möglichst gering gehalten werden, sodass zum Beispiel eine Schädigung der Probe durch Aufheizen vermieden wird.

Um das Verhältnis von nützlicher Intensität im Fokus und der Energie im Hintergrund zu quantifizieren, definieren wir c_{Fokus} als die Intensität im Fokus $I_f = I(0, f)$ im Verhältnis zum Energiefluss durch die Fokusebene. Als Maß für den Energiefluss P wird die integrierte Intensität $P = I_{\text{ges}} = \int_{\text{Fokusebene}} I(\mathbf{r}_{\perp}) d\mathbf{r}_{\perp}$ verwendet.

$$c_{\text{Fokus}} = \frac{I_f}{P} \quad (5.1)$$

Die so definierte Größe ist nicht dimensionslos. Durch das Verwenden der Fläche der Pupille in der Frequenzdomäne $S = \int_{\text{Pupille}} d\boldsymbol{\rho}_{\perp}$ als weiteren Normierungsfaktor, erhält man das generalisierte Strehl-Kriterium [65]. Dies ist eine dimensionslose Größe und kann als Fokussier-Effizienz aufgefasst werden.

$$\eta_{\text{Strehl}}^{\text{gen.}} = \frac{I_f}{S P} \quad (5.2)$$

Im Falle der Lichtausbreitung in einem nicht absorbierenden Medium entspricht der Energiefluss durch die Fokusebene P_f dem Energiefluss durch die Aperturebene P_{Apertur} , deshalb kann man eine beliebige Ebene zur Bestimmung des Energieflusses verwenden.

5.1.1 Exakte Lösungen für fokussierende Lichtverteilungen in der skalaren Wellenoptik

In diesem Abschnitt wird die Intensität im Fokus und der Energiefluss durch die Aperturebene von verschiedenen analytischen Lichtverteilungen wie der Sommerfeld-Welle, der Kugel-Welle, der sphärischen Phase und der quadratischen Phase exakt bestimmt.

Sommerfeld-Welle Die Sommerfeld-Welle, auch perfekte Welle [90] genannt, wird beschrieben durch:

$$U(r_{\perp}) = -\frac{1}{2\pi} U_0 \frac{\partial}{\partial f} \frac{e^{-ik\sqrt{r_{\perp}^2+f^2}}}{\sqrt{r_{\perp}^2+f^2}} \quad (5.3)$$

$$= \frac{f}{2\pi} U_0 \frac{e^{-ik\sqrt{r_{\perp}^2+f^2}}}{\sqrt{f^2+r_{\perp}^2}} \left(\frac{1}{f^2+r_{\perp}^2} + \frac{ik}{\sqrt{f^2+r_{\perp}^2}} \right). \quad (5.4)$$

Durch das Einführen einer Apertur mit Radius R lässt sich die numerische Apertur dieser Lichtverteilung begrenzen. Für die Amplitude im Fokus ergibt sich (siehe Anhang A.14):

$$U(0,f) = U_0 \left[\frac{1}{8\pi f^2} - \frac{f^2}{8\pi(f^2+R^2)^2} + k^2 \frac{R^2}{4\pi(f^2+R^2)} \right]. \quad (5.5)$$

Mit den Ersetzungen $\text{NA} = \frac{R}{\sqrt{R^2+f^2}}$, $R = \frac{f\text{NA}}{\sqrt{1-\text{NA}^2}}$ und $\frac{f^4}{(f^2+R^2)^2} = (1-\text{NA}^2)^2$ folgt

$$U(0,f) = U_0 \left[\frac{1}{8\pi f^2} - \frac{1}{8\pi f^2} (1-\text{NA}^2)^2 + \frac{k^2}{4\pi} \text{NA}^2 \right] \quad (5.6)$$

$$= U_0 \left[+\frac{\text{NA}^2}{4\pi f^2} - \frac{\text{NA}^4}{8\pi f^2} + \frac{k^2}{4\pi} \text{NA}^2 \right]. \quad (5.7)$$

Die zugehörige Energie, die durch die Apertur fließt, ist gegeben durch:

$$P = \int_A |U(r_\perp)|^2 d^2 \mathbf{r}_\perp \quad (5.8)$$

$$= I_0 \int_A \frac{f^2}{4\pi^2} \left(\frac{1}{(f^2 + r_\perp^2)^3} + \frac{k^2}{(f^2 + r_\perp^2)^2} \right) d^2 \mathbf{r}_\perp \quad (5.9)$$

$$= I_0 \int_0^R \frac{f^2}{2\pi} \left(\frac{1}{(f^2 + r_\perp^2)^3} + \frac{k^2}{(f^2 + r_\perp^2)^2} \right) r_\perp dr_\perp \quad (5.10)$$

$$= I_0 \frac{f^2}{2\pi} \left(\frac{1}{4f^4} - \frac{1}{4(f^2 + R^2)^2} + \frac{k^2 R^2}{2f^2(f^2 + R^2)} \right) \quad (5.11)$$

$$= I_0 \left[\frac{1}{8\pi f^2} - \frac{f^2}{8\pi(f^2 + R^2)^2} + \frac{k^2 R^2}{4\pi(f^2 + R^2)} \right], \quad (5.12)$$

mit $\text{NA} = \frac{R}{\sqrt{R^2 + f^2}}$, $R = \frac{f \text{NA}}{\sqrt{1 - \text{NA}^2}}$ und $\frac{f^4}{(f^2 + R^2)^2} = (1 - \text{NA}^2)^2$ folgt

$$P = I_0 \left[\frac{1}{8\pi f^2} - \frac{1}{8\pi f^2} (1 - \text{NA}^2)^2 + \frac{k^2 \text{NA}^2}{4\pi} \right] \quad (5.13)$$

$$= I_0 \left[\frac{\text{NA}^2}{2\pi f^2} - \frac{\text{NA}^4}{8\pi f^2} + \frac{k^2 \text{NA}^2}{4\pi} \right]. \quad (5.14)$$

Die Fläche in der Pupillenebene im Frequenzraum ist durch die numerische Apertur begrenzt:

$$S = \int_{\text{Pupille}} d\boldsymbol{\rho}_\perp = \int_0^{\frac{\text{NA}}{\lambda}} \int_0^{2\pi} \rho d\phi d\rho_\perp = \frac{k^2 \text{NA}^2}{4\pi}. \quad (5.15)$$

Für den Fall $\frac{1}{f} \ll k$ lassen sich die Ergebnisse weiter vereinfachen. Die Forderung $\frac{1}{f} \ll k$ ist äquivalent dazu, dass die evaneszenten Wellen keinen nennenswerten Beitrag zum Fokus liefern.

$$U(0, f) = U_0 \frac{k^2}{4\pi} \text{NA}^2 \quad (5.16)$$

$$I(0, f) = I_0 \frac{k^4}{16\pi^2} \text{NA}^4 \quad (5.17)$$

$$P = I_0 \frac{k^2 \text{NA}^2}{4\pi} \quad (5.18)$$

$$S = \frac{k^2 \text{NA}^2}{4\pi} \quad (5.19)$$

$$c_{\text{Fokus}} = \frac{k^2 \text{NA}^2}{4\pi} \quad (5.20)$$

Insbesondere für das generalisierte Strehl-Kriterium erhält man

$$\eta_{\text{Strehl}}^{\text{gen.}} = 1 \quad (5.21)$$

Dies lässt sich so interpretieren, dass die perfekte Welle für alle numerischen Aperturen ein ideales generalisiertes Strehl-Verhältnis hat.

Kugel-Welle Analog zum vorherigen Abschnitt lässt sich die komplexe Amplitude einer Kugelwelle untersuchen:

$$U(r_{\perp}) = U_0 f \frac{e^{-ik\sqrt{r_{\perp}^2+f^2}}}{\sqrt{r_{\perp}^2+f^2}} \quad (5.22)$$

Die Kugelwelle ist so normiert, dass im Ursprung in der Elementebene $|U(\mathbf{0})| = U_0$ gilt. Hinter einer Apertur mit Radius R in der Ebene $z = 0$ folgt für die Amplitude und Intensität im Fokus:

$$U(0,f) = U_0 \left[-ikf \left(1 - \frac{1}{\sqrt{\frac{R^2}{f^2} + 1}} \right) + \frac{R^2}{2(R^2 + f^2)} \right] \quad (5.23)$$

$$I(0,f) = I_0 \left[k^2 f^2 \left(1 - \frac{1}{\sqrt{\frac{R^2}{f^2} + 1}} \right)^2 + \frac{R^4}{4(R^2 + f^2)^2} \right] \quad (5.24)$$

Die Herleitung ist in Anhang A.14 gezeigt. Das Ergebnis ist konsistent mit der analytischen Lösung für das Fokusprofil von Osterberg und Smith [81] und Bertilone [6, 7]. Die zugehörige Energie, die durch die Apertur fließt, ist gegeben durch:

$$P = I_0 \int_0^{2\pi} \int_0^R \frac{f^2}{f^2 + r_{\perp}^2} r_{\perp} dr_{\perp} d\phi = I_0 f^2 \pi \log \left(\frac{f^2 + R^2}{f^2} \right) \quad (5.25)$$

Mit den Ersetzungen $\text{NA} = \frac{R}{\sqrt{R^2+f^2}}$ und $\frac{R}{f} = \frac{\text{NA}}{\sqrt{1-\text{NA}^2}}$ folgt:

$$U(0,f) = U_0 \left[-ikf \left(1 - \frac{1}{\sqrt{\frac{\text{NA}^2}{1-\text{NA}^2} + 1}} \right) + \frac{1}{2} \text{NA}^2 \right] \quad (5.26)$$

$$P = I_0 f^2 \pi \log \left(f^2 + \frac{1 - \text{NA}^2}{\text{NA}^2} \right). \quad (5.27)$$

Sphärische Phase Die sphärische Phase ist folgendermaßen definiert:

$$U(r_{\perp}) = U_0 e^{-ik\sqrt{r_{\perp}^2+f^2}}. \quad (5.28)$$

Hinter einer Apertur mit Radius R folgt

$$U(0,f) = U_0 \left[1 - \frac{f}{\sqrt{R^2+f^2}} + ikf \log \left(\frac{f}{\sqrt{R^2+f^2}} \right) \right] \quad (5.29)$$

$$I(0,f) = I_0 \left[1 - \frac{2f}{\sqrt{R^2+f^2}} + \frac{f^2}{R^2+f^2} + k^2 f^2 \log^2 \left(\frac{f}{\sqrt{R^2+f^2}} \right) \right]. \quad (5.30)$$

Die Herleitung ist in Anhang A.14 zu finden. Das Ergebnis ist konsistent mit der analytischen Lösung auf der optischen Achse von Hutton [41].

Die zugehörige Energie, die durch die Apertur fließt, ist gegeben durch:

$$P = I_0 \pi R^2 \quad (5.31)$$

Dies lässt sich als Funktion der numerischen Apertur schreiben:

$$P = I_0 \frac{\pi f^2 \text{NA}^2}{1 - \text{NA}^2} \quad (5.32)$$

Für die Intensität im Fokus in Abhängigkeit der numerischen Apertur ergibt sich für die Intensität im Fokus:

$$I(0, f) = I_0 \left[1 - \frac{2}{\sqrt{\frac{1}{1 - \text{NA}^2}}} + \frac{1}{\frac{1}{1 - \text{NA}^2}} + k^2 f^2 \log^2 \left(\frac{1}{\sqrt{\frac{1}{1 - \text{NA}^2}}} \right) \right] \quad (5.33)$$

$$= I_0 \left[2 - \text{NA}^2 - \frac{2}{\sqrt{\frac{1}{1 - \text{NA}^2}}} + k^2 f^2 \log^2 \left(\frac{1}{\sqrt{\frac{1}{1 - \text{NA}^2}}} \right) \right]. \quad (5.34)$$

5.1.2 Genäherte Lösungen für fokussierende Lichtverteilungen in der skalaren Wellenoptik

Im nächsten Abschnitt wird die quadratische Phase untersucht. Die quadratische Phase lässt sich als paraxiale Näherung der sphärischen Phase auffassen. Für die Lichtausbreitung wird die Fresnel-Näherung verwendet.

Quadratische Phase Die komplexe Amplitude der quadratischen Phasenverteilung ist gegeben durch

$$U(r_\perp) = U_0 e^{-ik \frac{r_\perp^2}{2f}} \quad (5.35)$$

Hinter einer Apertur mit Radius R erhält man für die Amplitudenverteilung im Fokus:

$$U(r_\perp, f) = -\frac{2\pi i}{\lambda f} e^{ikf} e^{ik \frac{r_\perp^2}{2f}} U_0 \int_0^R J_0 \left(\frac{kr'_\perp r_\perp}{f} \right) r'_\perp dr'_\perp \quad (5.36)$$

$$= -\frac{2\pi i}{\lambda f} e^{ikf} e^{ik \frac{r_\perp^2}{2f}} U_0 \frac{Rf J_1 \left(\frac{kr_\perp R}{f} \right)}{kr_\perp} \quad (5.37)$$

$$= -iR e^{ikf} e^{ik \frac{r_\perp^2}{2f}} U_0 \frac{J_1 \left(\frac{kRr_\perp}{f} \right)}{r_\perp}. \quad (5.38)$$

Im Grenzfall $r_\perp \rightarrow 0$ geht $\frac{J_1 \left(\frac{kRr_\perp}{f} \right)}{r_\perp} \rightarrow \frac{kR}{2f}$. Deshalb erhält man:

$$U(0, f) = -ie^{ikf} U_0 \frac{kR^2}{2f} \quad (5.39)$$

$$I(0, f) = I_0 \frac{k^2 R^4}{4f^2}. \quad (5.40)$$

Der Energiefluss durch die Apertur ist der gleiche wie bei der sphärischen Phase.

5.1.3 Vergleich der analytischen Lichtverteilungen

In Abbildung 5.1 werden die Ergebnisse der vorherigen Abschnitte zusammengefasst, indem das Verhältnis der Intensität im Fokus normiert auf die integrierte Intensität für verschiedene numerische Aperturen verglichen wird. Das Variieren der numeri-

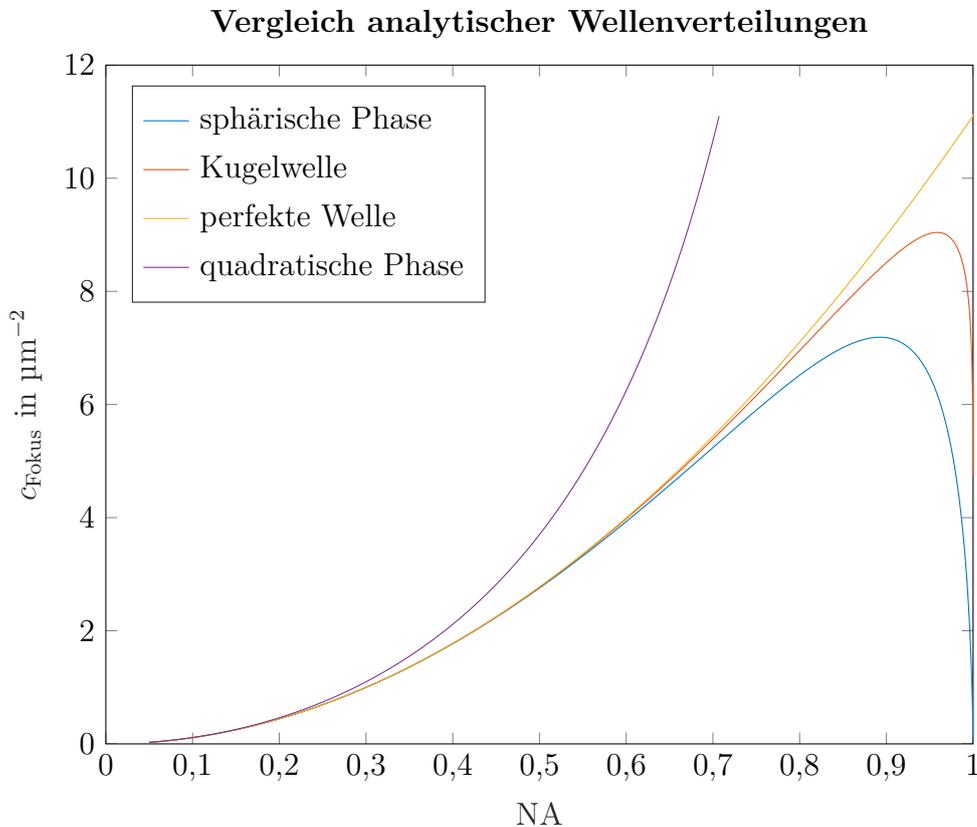


Abbildung 5.1: Fokussiereffizienz η für verschiedene Lichtverteilungen.

schen Aperturen lässt sich durch ein variieren der Brennweite bei gleichem Aperturdurchmesser oder durch variieren des Aperturdurchmessers bei gleicher Brennweite erreichen. Im gezeigten Beispiel wurde ein fester Radius für die Apertur von $100 \mu\text{m}$ angenommen und die Brennweite f als Funktion der numerischen Apertur verwendet.

Für die perfekte Welle gilt: je größer die numerische Apertur, desto größer ist die Intensität im Fokus bei gleichbleibendem Energiefluss. Zwischen der perfekten Welle, der Kugelwelle und der sphärischen Phase sieht man in den skalar exakten Rechnungen erst bei einer numerischen Apertur von $NA > 0,75$ Unterschiede. Bei sehr hohen numerischen Aperturen ändert sich das Skalierungsverhalten für die sphärische Phase und die Kugelwelle, ab einer $NA > 0,9$ sinkt die Intensität

Die paraxiale Rechnung unterscheidet sich ab einer NA von 0,3 deutlich von den skalar exakten Rechnungen. Im Bereich NA gegen 1 divergiert dieses Kriterium, hier ist jedoch die paraxiale Näherung auch nicht mehr gültig.

In Abbildung 5.2 ist das generalisierte Strehl-Kriterium für die verschiedenen Wellen als Funktion der numerischen Apertur gezeigt. Man sieht deutliche Unter-

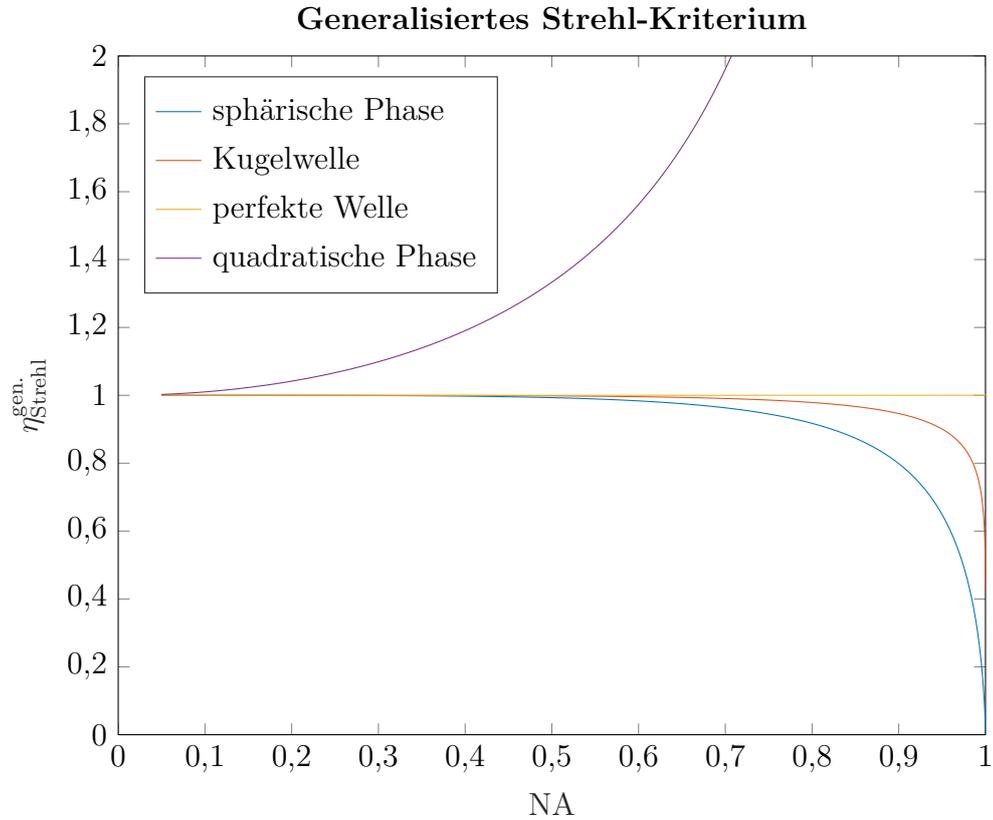


Abbildung 5.2: Generalisiertes Strehl-Kriterium $\eta_{\text{Strehl}}^{\text{gen.}}$ für verschiedene Lichtverteilungen.

schiede zwischen den einzelnen Wellen. Die paraxiale Rechnung mit der quadratischen Phase hat auch bei niedriger numerischer Apertur ein anderes Skalierungsverhalten als die drei anderen Wellen. Auch die Abweichungen der Kugelwelle und der sphärischen Phase gegenüber der perfekten Welle sind schon ab der numerischen Apertur von 0,8 bzw. 0,6 deutlich erkennbar.

5.1.4 Fokusintensität des idealen DOEs

Für ein ideales DOE mit homogener Pupille lässt sich die Intensität im Fokus anhand der gewünschten Amplitude leicht bestimmen:

$$\lim_{r_{\perp} \rightarrow 0} U_f(r_{\perp}) = \lim_{r_{\perp} \rightarrow 0} U_0 \frac{k_0 \text{NA}}{2\pi} \frac{J_1(k_0 r_{\perp} \text{NA})}{r_{\perp}} \quad (5.41)$$

$$= U_0 \frac{k_0 \text{NA}}{2\pi} \frac{k_0 \text{NA}}{2}. \quad (5.42)$$

Für die Intensität im Fokus gilt dann:

$$I(0,f) = I_0 \frac{k_0^4 \text{NA}^4}{16\pi^2}. \quad (5.43)$$

Mathematisch entspricht die Annahme einer homogenen Pupille, dass die Periodenlänge dieses Elements unendlich ist.

Wie in Kapitel 4 gezeigt entspricht die Intensität im Fokus für die diskretisierte Pupille der Intensität im Fokus der homogenen Pupille, wenn die Foki sich nicht gegenseitig stören.

Energiefluss durch die Elementperiode Der Energiefluss durch die Grundperiode lässt sich aus dem Integral über das Absolutbetragsquadrat der komplexen Transmissionsfunktion des Elements berechnen:

$$P^{\text{DOE}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |t^{\text{DOE}}(\mathbf{r}_{\perp})|^2 d\mathbf{r}_{\perp}. \quad (5.44)$$

Dieses Integral lässt sich am einfachsten durch Anwenden des Theorems von Parseval bestimmen:

$$P^{\text{DOE}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |t(\mathbf{r}_{\perp})|^2 d\mathbf{r}_{\perp} \quad (5.45)$$

$$= \frac{1}{2\pi} \int_0^{\infty} |\tilde{t}(s_{\perp})|^2 k_0^2 s_{\perp} ds_{\perp} \quad (5.46)$$

$$= \frac{1}{2\pi} \int_0^{\text{NA}} |U_0|^2 k_0^2 s_{\perp} ds_{\perp} \quad (5.47)$$

$$= I_0 \frac{k_0^2}{4\pi} \text{NA}^2. \quad (5.48)$$

Hiermit ergibt sich für das Verhältnis von Fokus-Intensität und Energiefluss

$$c_{\text{Fokus}} = \frac{k_0^2}{4\pi} \text{NA}^2. \quad (5.49)$$

In Abbildung 5.3 sieht man dieses Verhältnis als Funktion der numerischen Apertur.

5.1.5 Vergleich der idealen DOE mit der Sommerfeld-Welle

Die Ergebnisse für die Fokalintensität und den Energiefluss pro Periode der idealen DOE stimmen mit den Ergebnissen für die Sommerfeld-Welle im Grenzfall $\frac{1}{k} \ll f$ überein. Anschaulich bedeutet dieser Grenzfall, dass keine evaneszenten Wellen zur Intensität im Fokus beitragen. Diese Annahme ist bei sichtbaren optischen Wellenlängen quasi immer erfüllt. Die ideale DOE hat somit genauso wie die Sommerfeld-Welle ein generalisiertes Strehl von 1 für alle numerischen Aperturen.

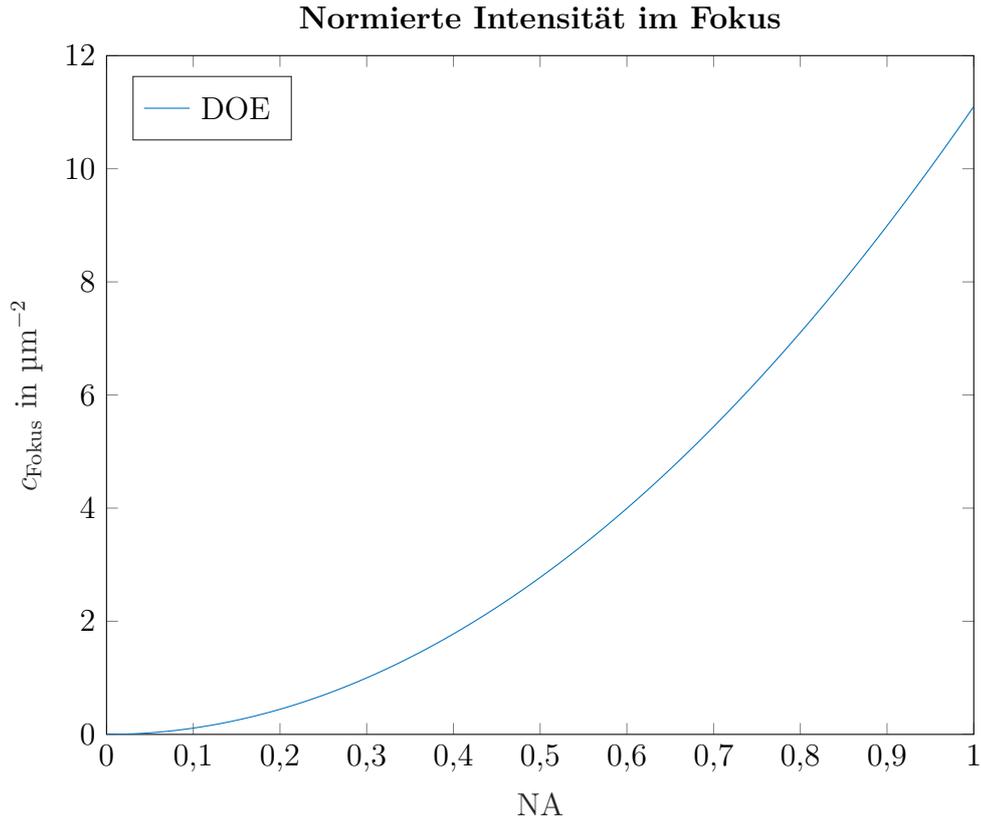


Abbildung 5.3: Normierte Intensität im Fokus für diffraktive Elemente.

Plausibel wird das gleiche Skalierungsverhalten von Sommerfeld-Welle und idealer DOE, wenn man sich das Spektrum der Sommerfeld-Welle anschaut, dies lässt sich aus der Weyl-Äquivalenz [102] herleiten:

$$\frac{e^{-ikr}}{\mathbf{r}} = -2\pi i \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{k_z} e^{ik_z|z|} e^{i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} \frac{d\mathbf{k}_{\perp}}{4\pi^2}$$

Somit gilt für die Ableitung:

$$\begin{aligned} \frac{\partial}{\partial z} \frac{e^{-ikr}}{\mathbf{r}} &= -\frac{\partial}{\partial z} \frac{i}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{k_z} e^{ik_z|z|} e^{i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} d\mathbf{k}_{\perp} \\ &= 2\pi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{ik_z|z|} e^{i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} \frac{d\mathbf{k}_{\perp}}{4\pi^2} \end{aligned}$$

Das Spektrum der Sommerfeld-Welle in jeder z -Ebene ist somit gegeben durch

$$\tilde{U}(k_{\perp}, z) = 2\pi U_0 e^{ik_z|z|} \tag{5.50}$$

Begrenzt man das Spektrum indem man nur noch Frequenzen mit $k_{\perp} \leq k_0 \text{NA}$ zulässt, so erhält man

$$\tilde{U}(k_{\perp}, z) = 2\pi U_0 e^{ik_z|z|} \text{circ}\left(\frac{k_{\perp}}{k_0 \text{NA}}\right) \tag{5.51}$$

Man sieht, dass das bandbegrenzte Spektrum der Sommerfeld-Welle in der Fokalebene $z = 0$ proportional zu dem Spektrum einer Linse mit homogener Pupille ist.

5.1.6 Alternative Wahl zur Quantisierung des Energieflusses

Als Maß für den Energiefluss wurde im vorherigen Abschnitten die integrierte Intensität verwendet. Für skalare Felder lässt sich auch der totale Energiefluss F_z^{total} als Maß zur Bestimmung des Energieflusses verwenden. Hiermit lässt sich die Intensität im Fokus wieder folgendermaßen skalieren:

$$c_{\text{Fokus}}^{\text{totale Energie}} = \frac{I_f}{F_z^{\text{total}}} \quad (5.52)$$

Durch zusätzliches Normieren auf das Quadrat der Wellenzahl k_0 erhält man eine dimensionslose Größe ähnlich der Effizienz [57]:

$$\eta_{\text{Fokus}}^{\text{totale Energie}} = \frac{I_f}{F_z^{\text{total}} k_0^2}. \quad (5.53)$$

Für die homogene bandbegrenzte Pupille und die Kugelwelle werden nun im Folgenden die Werte für den totalen Energiefluss analytisch bestimmt.

Homogene bandbegrenzte Pupille oder bandbegrenzte perfekte Welle

Für den Fall einer homogenen bandbegrenzten Pupille ist der totale Energiefluss gegeben durch:

$$F_z^{\text{total}} = 2\alpha\omega \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |U_0|^2 \text{circ}\left(\frac{k_{\perp}}{k_0 \text{NA}}\right) k_z \frac{d\mathbf{k}_{\perp}}{4\pi^2} \quad (5.54)$$

$$= \frac{2\alpha\omega}{4\pi^2} I_0 \int_0^{2\pi} \int_0^{k_0 \text{NA}} k_z k_{\perp} dk_{\perp} d\phi \quad (5.55)$$

$$= \frac{\alpha\omega}{\pi} I_0 \int_0^{k_0 \text{NA}} \sqrt{k_0^2 - k_{\perp}^2} k_{\perp} dk_{\perp} \quad (5.56)$$

$$= -\frac{\alpha\omega}{\pi} I_0 \frac{1}{3} \left(k_0^2 - k_{\perp}^2\right)^{\frac{3}{2}} \Big|_0^{k_0 \text{NA}} \quad (5.57)$$

$$= +\frac{\alpha\omega k_0^3}{3\pi} I_0 \left(1 - \left(1 - \text{NA}^2\right)^{\frac{3}{2}}\right). \quad (5.58)$$

Dieser Fall ist äquivalent zur Betrachtung einer bandbegrenzten perfekten Welle.

Abschätzung des Fehlers auf dem Gesamtenergiefluss bei einer bandbegrenzten perfekten Welle

Für paraxiale Wellenfelder wurde in Abschnitt 2.7.4 gezeigt, dass der totale Energiefluss proportional zur integrierten Intensität ist. Für homogene Pupillen lässt sich der relative Fehler zwischen der integrierten Intensität aus Gleichung 5.18 und des totalen Energieflusses aus Gleichung 5.58 betrachten. Das

Relativer Fehler des Energieflusses

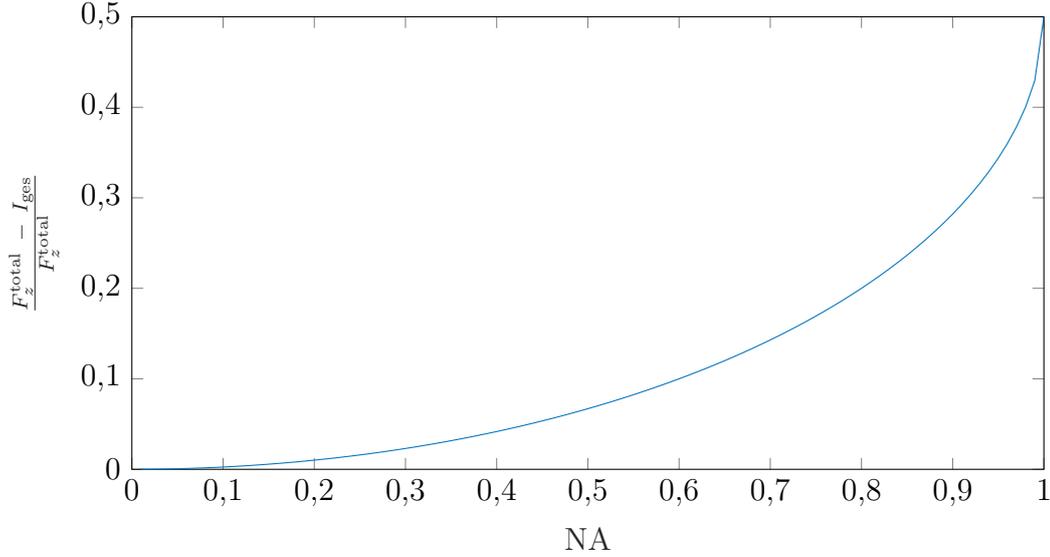


Abbildung 5.4: Relative Abweichung zwischen integrierter Intensität und totalem Energiefluss für eine homogene Pupille. Für den Vergleich wurde die Skalierungskonstante $\alpha = \frac{1}{2\omega k_0}$ gesetzt.

Ergebnis ist in Abbildung 5.4 aufgetragen, hierbei wurde $\alpha = \frac{1}{2c_0 k_0^2}$ verwendet. Man sieht bei einer NA von 0,5 ist die Abweichung ca. 6 %, bei NA = 0,75 ca. 17 % und bei NA = 1,0 genau 50 %.

Kugelwelle

Zum Vergleich lässt sich auch der totale Energiefluss einer Kugelwelle berechnen. Das Spektrum der konvergierenden Kugel-Welle lässt sich aus der Weyl-Äquivalenz [67, 102] herleiten:

$$\frac{e^{-ikr}}{\mathbf{r}} = -2\pi i \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{k_z} e^{ik_z|z|} e^{i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp}} \frac{d\mathbf{k}_{\perp}}{4\pi^2} \quad (5.59)$$

Das Spektrum in einer beliebigen z-Ebene ist also $\tilde{U} = -2\pi i U_0 \frac{e^{ik_z z}}{k_z}$. Hiermit folgt für den totalen Energiefluss:

$$F_z^{\text{total}} = 2\alpha\omega \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |U_0|^2 \frac{4\pi^2}{k_z^2} k_z \frac{d\mathbf{k}_{\perp}}{4\pi^2} \quad (5.60)$$

$$= 4\alpha\omega\pi I_0 \int_0^{k_0 \text{NA}} \frac{1}{\sqrt{k_0^2 - k_{\perp}^2}} k_{\perp} dk_{\perp} \quad (5.61)$$

$$= 4\alpha\omega\pi I_0 \left(k_0 - \sqrt{k_0^2 - k_0^2 \text{NA}^2} \right) \quad (5.62)$$

$$= 4\alpha\omega\pi k_0 I_0 \left(1 - \sqrt{1 - \text{NA}^2} \right). \quad (5.63)$$

Vergleich der analytischen Wellen

In Abbildung 5.5 sieht man die Fokusintensität der bandbegrenzten perfekten Welle und der Kugelwelle bei unterschiedlichen numerischen Aperturen aber gleichem totalen Energiefluss. Der Kurvenverlauf für die Kugelwelle und die perfekte Welle

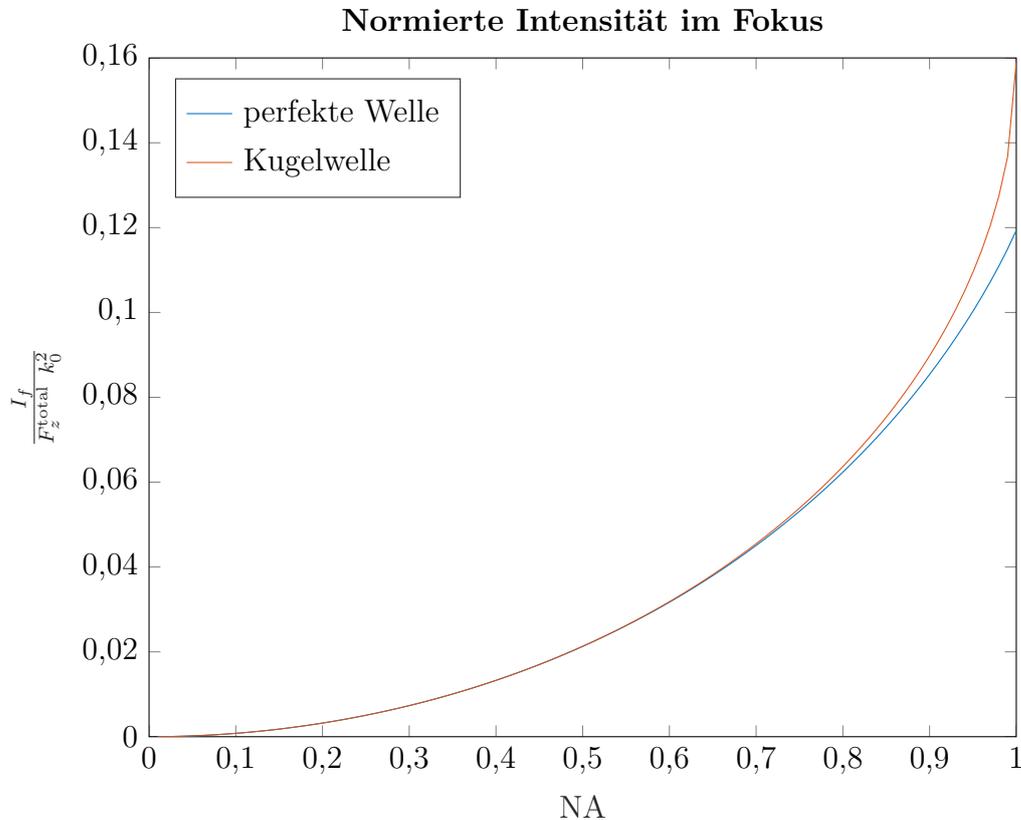


Abbildung 5.5: Normierte Intensität im Fokus mit totalem Energiefluss als Referenz. Für die Skalierungskonstante wurde $\alpha = \frac{1}{2\omega k_0}$ angenommen.

ist bei niedrigen und moderaten numerischen Aperturen gleich, für hohe numerische Aperturen zeigen sich Unterschiede.

Dieses Kriterium zeigt im Gegensatz zum generalisiertem Strehl-Kriterium, dass die Kugelwelle eine bessere Lichtkonzentration als die perfekte Welle hat [57]. Der Unterschied wird jedoch erst bei sehr hohen numerischen Aperturen deutlich.

Fazit

Ob der totale Energiefluss oder die integrierte Intensität als Maß für den Energiefluss genommen wird, entscheidet darüber, welches Wellenfeld die bessere Lichtkonzentration hat. Im Allgemeinen zeigen sich, dass erst bei sehr hohen numerischen Aperturen Unterschiede in der normierten Intensität im Fokus deutlich werden. Somit bleibt die normierte Intensität im Fokus und das daraus abgeleitete generalisierte

Strehl-Kriterium ein valides Instrument um die Fokussiereffizienz zu beurteilen. Erst für numerische Aperturen größer als ca. 0,8 ist ein Design mit einer Kugelwelle einem Design mit einer homogenen Pupille vorzuziehen. In diesem Bereich sind jedoch die mögliche Abweichungen aufgrund der Verletzung der dünnen Element Näherung zu berücksichtigen.

5.2 Auswirkung von Amplitudennivellierung und Phasendiskretisierung

In diesem Abschnitt werden verschiedene Realisierungen der Linsen mit überlappender Apertur untersucht. Die Linsentypen unterscheiden sich dadurch, dass die Phase kontinuierlich ist oder diskretisiert wurde bzw. dass die Amplitude kontinuierlich oder nivelliert ist.

Zuerst wird in Abschnitt 5.2.1 und 5.2.2 die Intensität und die Geometrie des Fokus für die verschiedenen Linsentypen untersucht. Anschließend wird in Abschnitt 5.2.3 analytisch die Effizienz der Phasendiskretisierung in Abhängigkeit der Anzahl der Phasenstufen hergeleitet. Des Weiteren wird in Abschnitt 5.2.4 der Spezialfall von binären Phasenelementen mit niedriger numerischer Apertur betrachtet.

5.2.1 Einfluss von Amplitudennivellierung und Phasenbinarisierung auf die Intensität im Fokus

In Abbildung 5.6 ist die normierte Intensität im Fokus für vier verschiedene Linsentypen als Funktion der numerischen Apertur aufgetragen. Man sieht, dass die normierte Intensität im Fokus für den Fall kontinuierlicher Phasen und Amplituden am höchsten ist. Eine Amplitudennivellierung vermindert die normierte Intensität im Fokus weniger als eine Phasenbinarisierung. Die Kombination von Amplitudennivellierung und Phasenbinarisierung erzeugt in der Regel noch geringere normierte Intensitäten im Fokus.

Um den Einfluss der Phasen und Amplituden Modifikation quantisieren zu können, lässt sich das Strehl-Verhältnis benutzen [65]. Hierbei wird das erzeugte Wellenfeld des modifizierten Elements als ein Wellenfeld mit Aberrationen betrachtet und es wird die Intensität im Fokus mit der eines idealen Wellenfelds ohne Aberrationen verglichen. Zur Vergleichbarkeit wird die Intensität im Fokus auf die integrierte Intensität des Wellenfelds normiert. Aufgrund der periodischen Fortsetzung des Wellenfelds wird bei den diffraktiven Linsen mit überlappenden Aperturen die integrierte Intensität pro Periode verwendet:

$$\eta_{\text{Strehl}} = \frac{I(0,f)}{I_{\text{ideal}}(0,f)} \frac{P_{\text{ideal}}}{P}. \quad (5.64)$$

In Abbildung 5.7 ist das Strehl-Verhältnis als Maß für die Fokussiereffizienz aufgetragen. Hierbei wurde die normierte Intensität eines idealen DOEs als Referenz benutzt.

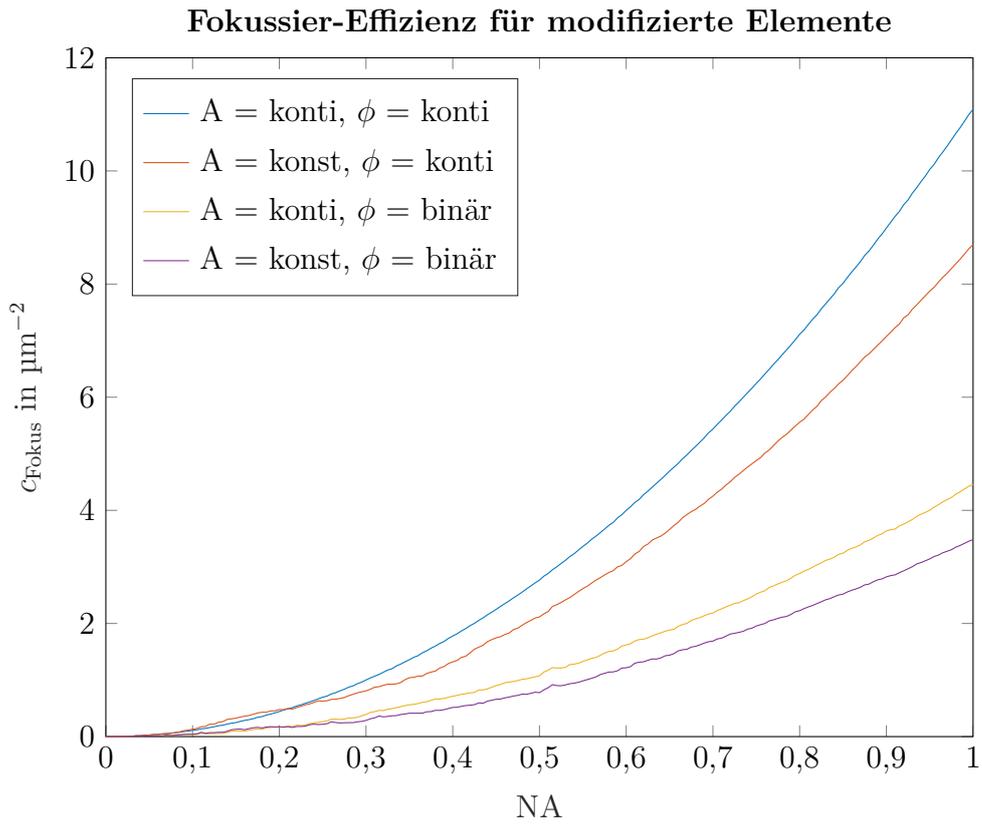


Abbildung 5.6: Einfluss der Amplitudennivellierung und Phasendiskretisierung auf die normierte Intensität im Fokus.

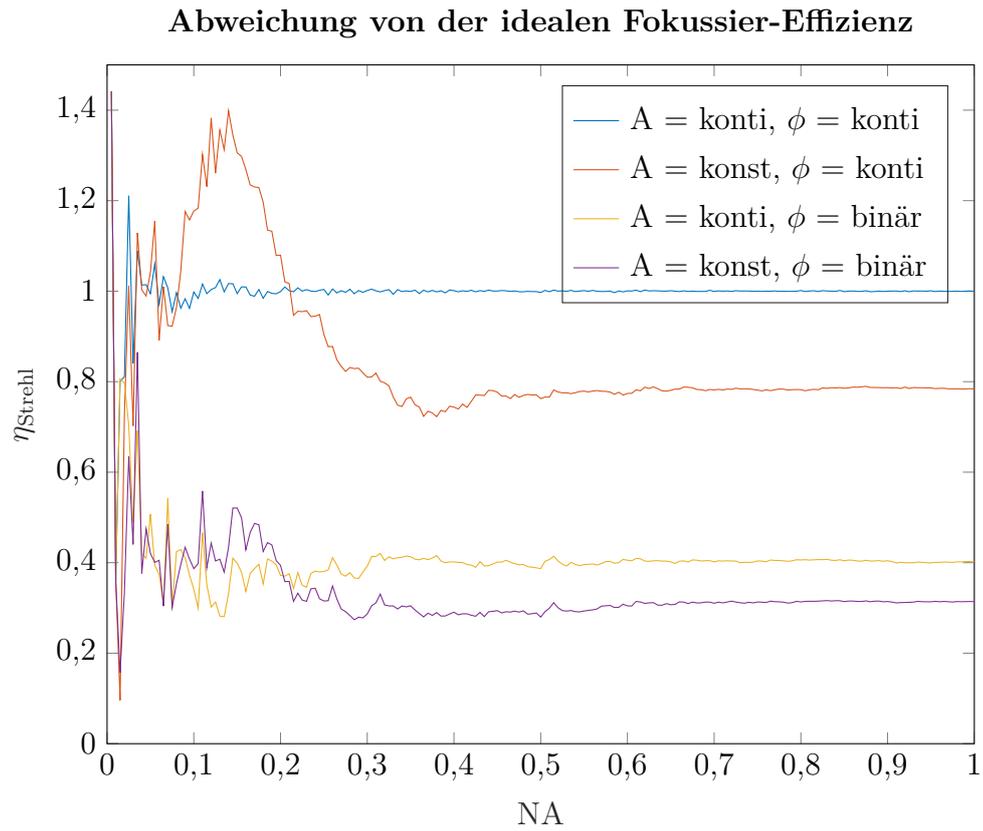


Abbildung 5.7: Einfluss der Amplitudennivellierung und Phasendiskretisierung auf die Fokussier-Effizienz, beschrieben durch das Strehl-Verhältnis.

Abweichung von der idealen Fokussier-Effizienz

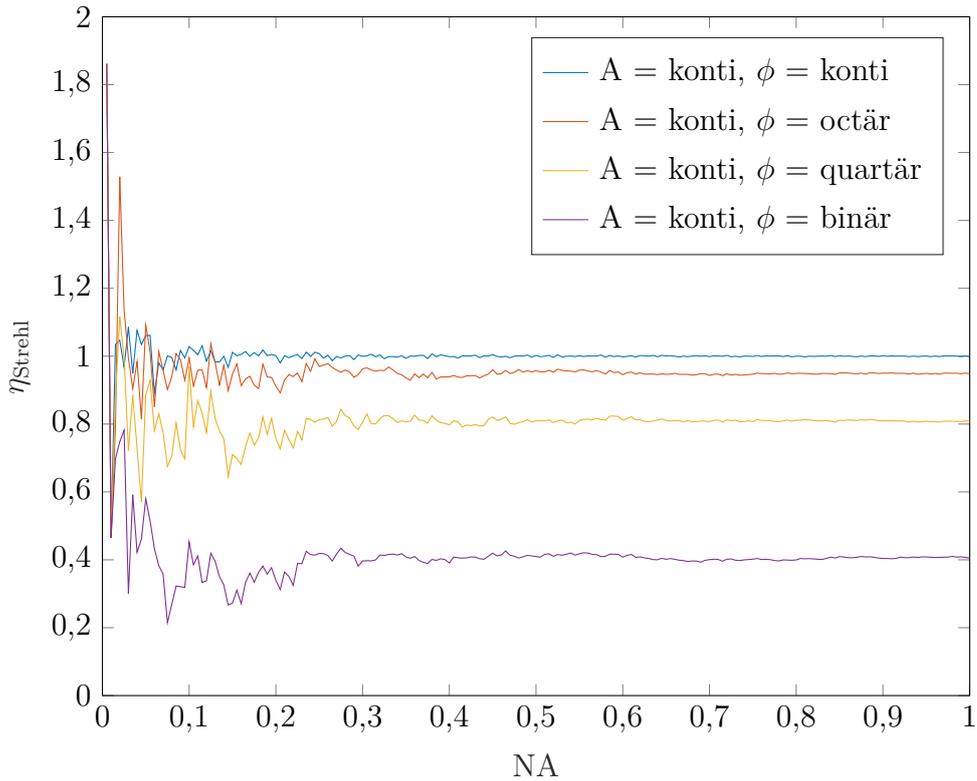


Abbildung 5.8: Effizienz der Phasenquantisierung für 2, 4 oder 8 Phasenstufen.

Für mittlere und hohe numerische Aperturen lässt sich der Einfluss der Amplitudennivellierung als eine Abschwächung um den Faktor $\frac{\pi}{4} \approx 0,785$ schreiben. Eine Phasenbinarisierung hingegen bewirkt eine Abschwächung um einen Faktor $\frac{4}{\pi^2} \approx 0,405$. Diese beiden Effekte scheinen unabhängig von einander zu sein. Man erhält für ein amplitudennivelliertes und phasenbinarisiertes Element bei hohen numerischen Aperturen eine Abschwächung von $\frac{1}{\pi} \approx 0,318$, dies entspricht dem Produkt der Abschwächfaktoren der einzelnen Effekte.

Einfluss der Anzahl der Stufen der Phasendiskretisierung

Neben der Phasenbinarisierung kann eine allgemeine Phasendiskretisierung betrachtet werden. In Abbildung 5.8 wird der Einfluss der Phasenquantisierung mit zwei, vier oder acht Phasenstufen bei nicht modifizierter variabler Amplitude untersucht. Für den Fall von vier Quantisierungsstufen ist die Abschwächung in der Größenordnung der Abschwächung der Amplitudennivellierung.

Für mittlere und große numerische Aperturen konvergieren die Werte zu je einem konstanten Wert, dieser entspricht dem theoretischen Optimum von phasendiskre-

N_{Level}	η
2	$\frac{4}{\pi^2} \approx 0.405$
4	$\frac{8}{\pi^2} \approx 0.811$
8	≈ 0.950
16	≈ 0.987

Tabelle 5.1: Tabellarische Auflistung der Fokussier-Effizienz nach Gleichung 5.65.

tisierten Elementen gegeben durch [46, 66, 100]:

$$\eta = \left(\frac{N_{\text{Level}} \sin\left(\frac{\pi}{N_{\text{Level}}}\right)}{\pi} \right)^2. \quad (5.65)$$

Eine Herleitung wird in Abschnitt 5.2.3 gezeigt. In Tabelle 5.1 ist die Effizienz beispielhaft für eine unterschiedliche Anzahl an Phasenstufen aufgelistet. Eine Verdoppelung von zwei auf vier Phasenlevel verdoppelt die Effizienz. Weiteres verdoppeln der Anzahl der Phasenlevel ergibt Effizienzen, die schon sehr nahe bei Eins liegt.

5.2.2 Einfluss der Amplitudennivellierung und Phasendiskretisierung auf das Fokusprofil

Neben der Punkt-Intensität im geometrischen Fokus lässt sich allgemeiner die Intensitätsverteilung in einem Bereich um den geometrischen Fokus betrachten. Hierfür erweisen sich insbesondere xz-Schnitte durch den Fokus als geeignet, da hier sowohl das laterale als auch das axiale Fokusprofil sichtbar ist.

Einfluss der Amplitudennivellierung

Im Folgenden wird zuerst der Einfluss der Amplitudennivellierung auf das Fokusprofil untersucht. Hierbei lassen sich bei niedriger und hoher numerischer Apertur unterschiedliche Effekte beobachten.

Beobachtung für kleine NA: In Abbildung 5.9 sieht man die Auswirkung der Amplitudennivellierung bei niedriger numerischer Apertur. Die Intensität im Fokus des amplitudennivellierten Elements unterscheidet sich deutlich von dem des kontinuierlichen Elements. Im xz-Verlauf treten Modulationen mit deutlich höherer Frequenz auf. Die maximale Intensität liegt nicht mehr im theoretischen Fokus. Für niedrige numerische Aperturen ist die Intensitätsverteilung ausgedehnt und stört deshalb die Nachbarperioden.

Beobachtung für mittlere und höhere NA: In Abbildung 5.10 sieht man die Auswirkung der Amplitudennivellierung für ein Element mit mittlerer numerischer

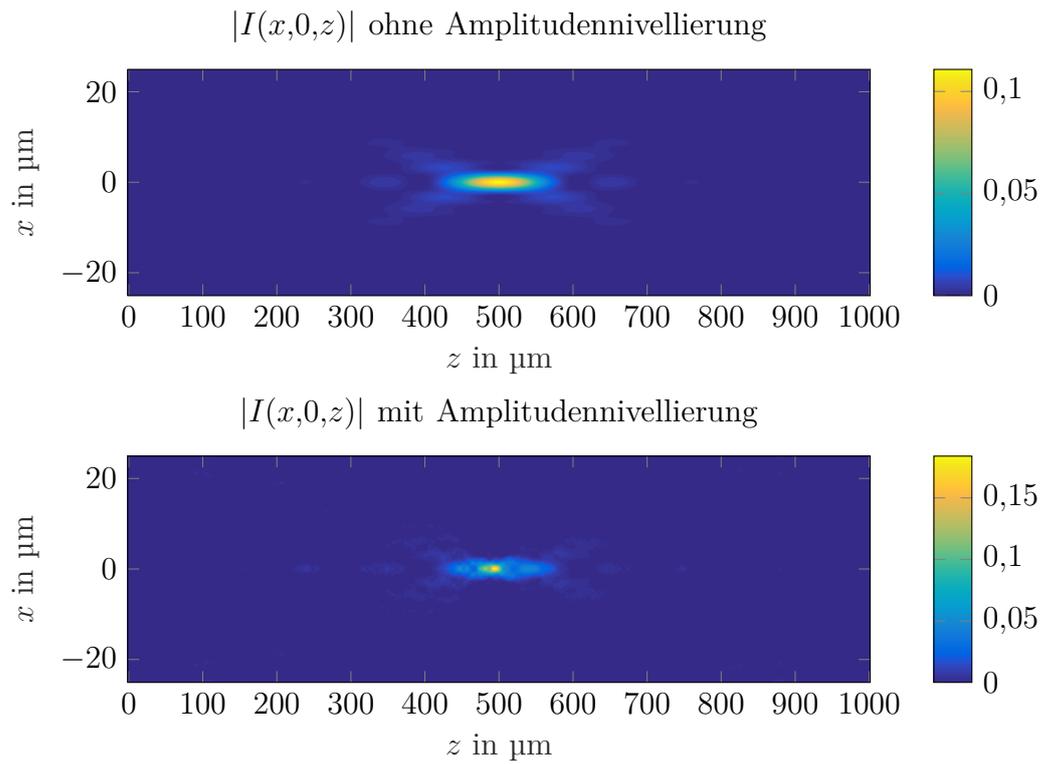


Abbildung 5.9: Einfluss der Amplitudennivellierung auf den Fokusverlauf eines Elements mit Periode $50 \mu\text{m}$ und Brennweite $500 \mu\text{m}$ bei einer numerischer Apertur von $0,1$ und der Wellenlänge von $0,532 \mu\text{m}$. Die Parameterwahl entspricht einem Überlapp von ca. 2 .

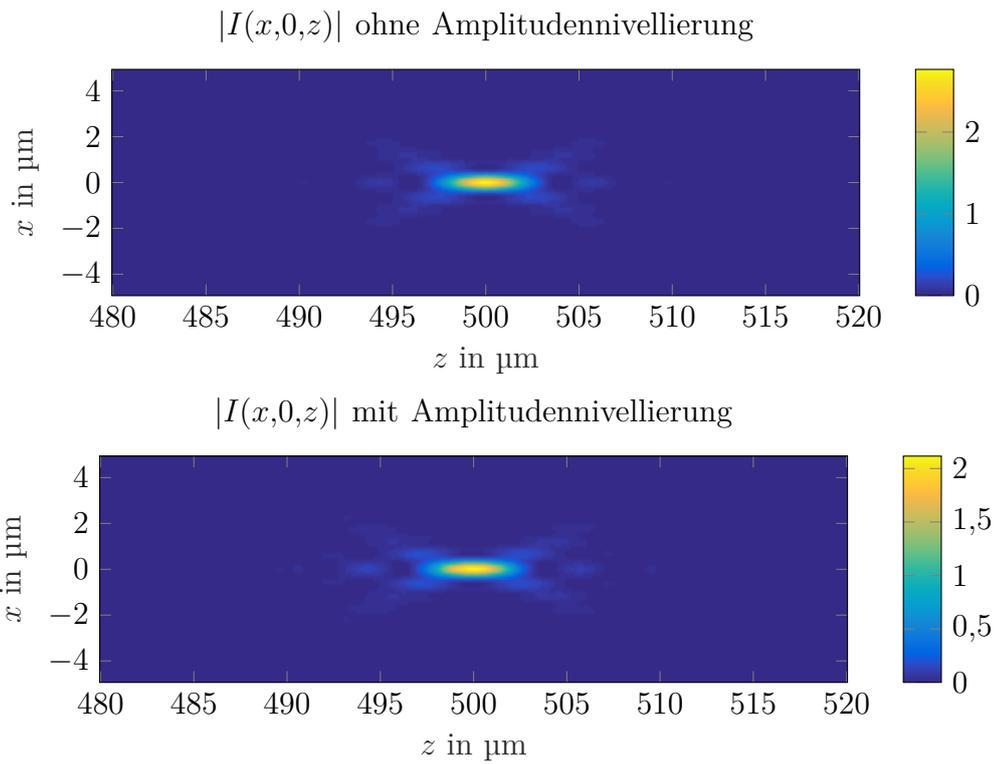


Abbildung 5.10: Einfluss der Amplitudennivellierung auf den Fokusverlauf eines Elements mit Periode $50 \mu\text{m}$ und Brennweite $500 \mu\text{m}$ bei einer numerischen Apertur von $0,5$ und der Wellenlänge von $0,532 \mu\text{m}$. Die Parameterwahl entspricht einem Überlapp von ca. $11,5$.

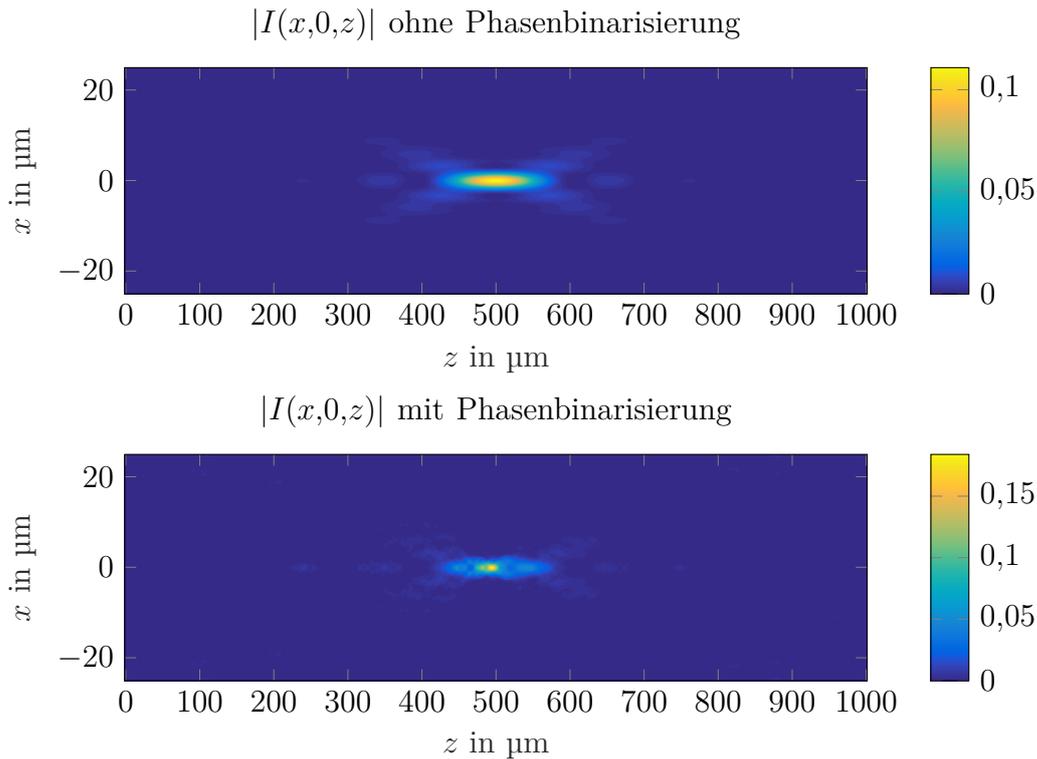


Abbildung 5.11: Einfluss der Phasenbinarisierung auf den Fokusverlauf eines Elements mit Periode $50\ \mu\text{m}$ und Brennweite $500\ \mu\text{m}$. Die numerische Apertur ist $0,1$ und entspricht einem Überlapp von ca. 2 .

Apertur. Der xz -Verlauf ist qualitativ gleich, nur die Intensität im Fokus ist um den Faktor $\frac{\pi}{4}$ skaliert.

Einfluss der Phasendiskretisierung

In Abbildung 5.11 und 5.12 sieht man den Einfluss der Phasenbinarisierung für den Fall einer niedrigen und einer mittleren numerischen Apertur. Es zeigt sich analog zur Amplitudennivellierung, dass bei kleiner numerischer Apertur der Fokalverlauf modelliert wird und dass bei mittlerer und großer numerischer Apertur die Intensität im Fokus nur um einen Faktor $\frac{4}{\pi^2}$ skaliert ist.

Schiebt man beim Design den Phasensprung durch eine Addition einer globalen Phase vor der Phasenquantisierung (siehe Abschnitt 4.4), so erhält man im Fall kleiner numerischer Aperturen eine leicht veränderte Modulation der Intensitätsverteilung im Fokus.

Benutzt man mehrere Phasenstufen bei der Phasendiskretisierung, so konvergiert die Intensitätsverteilung gegen die kontinuierliche Verteilung. Beispiele für den Fokus mit vier und acht Phasenstufen bei einer kleinen numerischen Apertur sind in Abbildung 5.13 zu sehen. Schon bei vier Phasenstufen ist eine deutliche Reduktion der Modulation im Fokus sichtbar. Bei acht Phasenstufen zeigt sich kein qualitativer

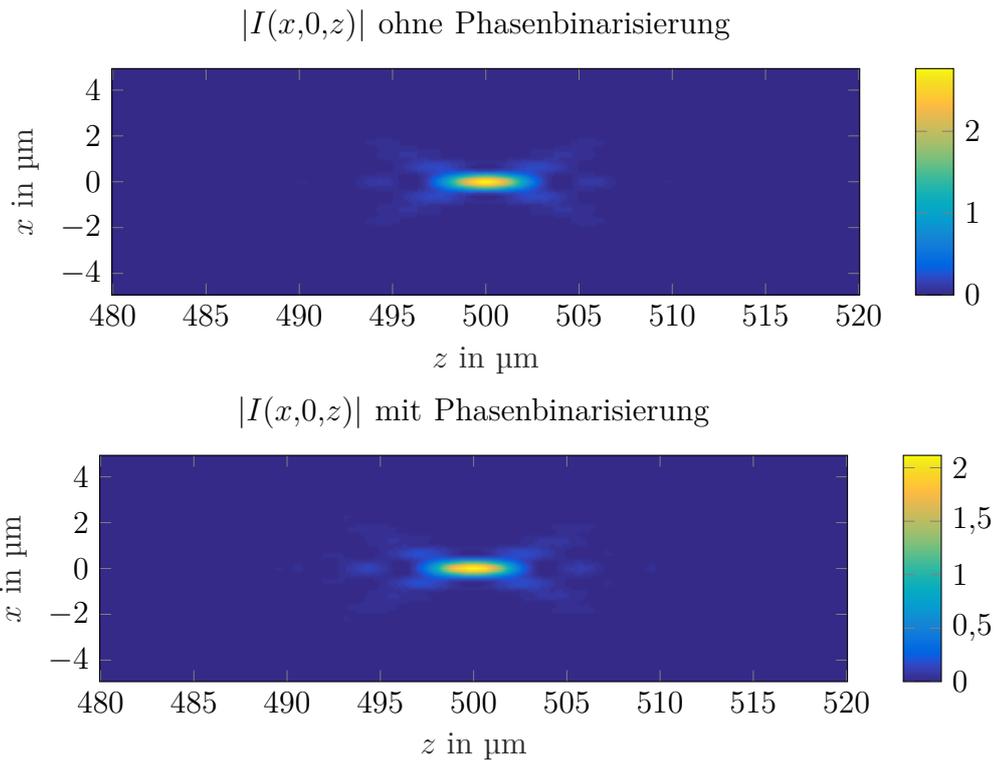


Abbildung 5.12: Einfluss der Phasenbinarisierung auf den Fokusverlauf eines Elements mit Periode $50\ \mu\text{m}$ und Brennweite $500\ \mu\text{m}$. Die numerische Apertur ist $0,5$ und entspricht einem Überlapp von ca. $12,5$.

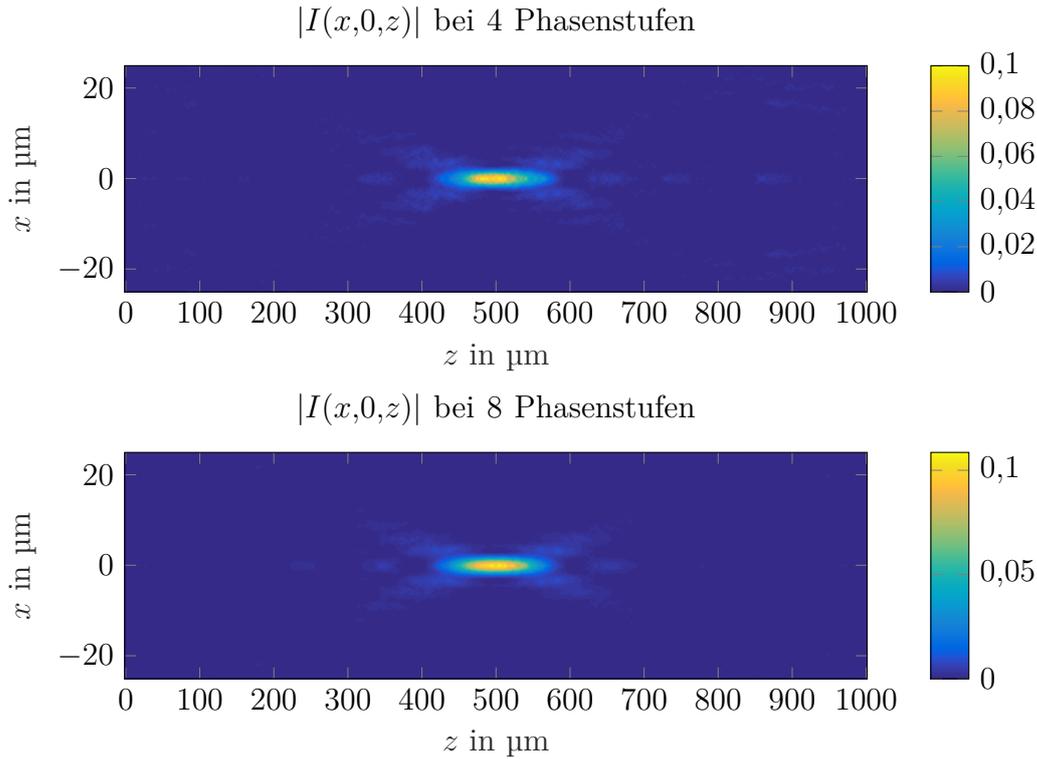


Abbildung 5.13: Einfluss der Anzahl der Phasenstufen auf den Fokusverlauf eines Elements mit Periode $50\ \mu\text{m}$ und Brennweite $500\ \mu\text{m}$. Die numerische Apertur ist $0,1$ und entspricht einem Überlapp von ca. 2 .

Unterschied zum Profil des nicht phasendiskretisierten Elements.

5.2.3 Analytische Beschreibung der Phasendiskretisierung

Um die Phasendiskretisierung genauer zu untersuchen wird im Folgenden der Phasenanteil einer allgemeinen Transmissionsfunktion genauer betrachtet. Der Einfachheit halber wird die Amplitude auf 1 gesetzt. Um die Beeinflussung der Phase quantifizieren zu können wird die Phasenmodulationsfunktion $p(\phi)$ eingeführt, die eine modulierte Phasenverteilung aus einer Ausgangsphase ϕ erzeugt [77]. Die Transmissionsfunktion eines solchen Elements ist gegeben durch

$$t(\mathbf{r}_\perp) = t(\phi(\mathbf{r}_\perp)) = e^{ip(\phi(\mathbf{r}_\perp))}. \quad (5.66)$$

Die Funktion $t(\phi(\mathbf{r}_\perp))$ ist periodisch in ϕ mit Periodenlänge 2π , somit lässt sie sich für jeden Punkt \mathbf{r}_\perp als Fourierreihe darstellen:

$$t(\phi) = \sum_{\alpha=-\infty}^{\infty} G_\alpha e^{i\alpha\phi}. \quad (5.67)$$

Die Fourier-Koeffizienten G_α sind folgendermaßen definiert:

$$G_\alpha = \frac{1}{2\pi} \int_0^{2\pi} t(\phi) e^{-i\alpha\phi} d\phi \quad (5.68)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} e^{ip(\phi)} e^{-i\alpha\phi} d\phi. \quad (5.69)$$

Setzt man zum Beispiel eine konvergierende quadratische Phase $\phi = -\frac{\pi r_\perp^2}{\lambda f}$ ein, so sieht man, dass das Element $t(\phi)$ aus Wellen mit den Phasen $\phi_\alpha = -\alpha \frac{\pi r_\perp^2}{\lambda f}$ zusammengesetzt ist, die jeweils ihren Fokus bei $z = \alpha f$ haben. Der Koeffizient G_α gibt die Stärke der Teilwelle an. Im Speziellen gibt der Koeffizient G_1 immer die Stärke der nicht modellierten Ausgangswelle an, somit lässt sich die Effizienz der Modulationsfunktion $p(\phi)$ über das Absolutquadrat des ersten Koeffizienten definieren:

$$\eta_p = |G_1|^2. \quad (5.70)$$

Spezialfall einer kontinuierlichen Phase Für die nicht diskretisierte Phase erhält man folgende Modulationsfunktion:

$$p(\phi(\mathbf{r}_\perp)) = \phi(\mathbf{r}_\perp). \quad (5.71)$$

Setzt man dies in die Fourier-Reihe ein, so sieht man, dass nur die erste Ordnung einen Beitrag liefert:

$$G_\alpha = \frac{1}{2\pi} \int_0^{2\pi} t(\phi) e^{-i\alpha\phi} d\phi \quad (5.72)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} e^{i\phi} e^{-i\alpha\phi} d\phi \quad (5.73)$$

$$= \delta_{\alpha 1}. \quad (5.74)$$

Somit hat diese Modulationsfunktion eine Effizienz von $\eta_p = 1$.

Spezialfall für eine diskretisierte Phase Eine Modulationsfunktion, die die Ausgangsphase in N äquidistante Phasenlevel diskretisiert ist gegeben durch:

$$p(\phi) = \begin{cases} 0 & \text{für } 0 < \text{mod}(\phi, 2\pi) \leq \frac{1}{2} \frac{2\pi}{N} \\ 1 \frac{2\pi}{N} & \text{für } \frac{1}{2} \frac{2\pi}{N} < \text{mod}(\phi, 2\pi) \leq \left(\frac{1}{2} + 1\right) \frac{2\pi}{N} \\ \vdots & \\ m \frac{2\pi}{N} & \text{für } \left(\frac{1}{2} + m - 1\right) \frac{2\pi}{N} < \text{mod}(\phi, 2\pi) \leq \left(\frac{1}{2} + m\right) \frac{2\pi}{N} \\ \vdots & \\ (N-1) \frac{2\pi}{N} & \text{für } \left(\frac{1}{2} + N - 2\right) \frac{2\pi}{N} < \text{mod}(\phi, 2\pi) \leq \left(\frac{1}{2} + N - 1\right) \frac{2\pi}{N} \\ 0 & \text{für } \left(\frac{1}{2} + N - 1\right) \frac{2\pi}{N} < \text{mod}(\phi, 2\pi) \leq 2\pi \end{cases} \quad (5.75)$$

Die dazu passende Transmissionsfunktion $t'(\phi)$ lässt sich erzeugen, indem man zuerst die kontinuierliche Funktion $t(\phi)$ abtastet und anschließend mit einer Rechteckfunktion faltet. So erhält man:

$$t'(\phi) = \left[t(\phi) \sum_{m=0}^{N-1} \delta\left(m \frac{2\pi}{N} - \phi\right) \right] \otimes_{\phi} \text{rectp}\left(\frac{\phi}{\frac{2\pi}{N}}\right) \quad (5.76)$$

$$= \sum_{m=0}^{N-1} \left(\left[t(\phi) \delta\left(m \frac{2\pi}{N} - \phi\right) \right] \otimes_{\phi} \text{rectp}\left(\frac{\phi}{\frac{2\pi}{N}}\right) \right) \quad (5.77)$$

Hierbei ist $\text{rectp}(\phi)$ die 2π -periodisch fortgesetzte Rechteckfunktion. Setzt man dies nun in die Definition des Fourier-Koeffizienten ein, so erhält man:

$$\begin{aligned} G_{\alpha} &= \frac{1}{2\pi} \int_0^{2\pi} t'(\phi) e^{-i\alpha\phi} d\phi \\ &= \frac{1}{2\pi} \int_0^{2\pi} \sum_{m=0}^{N-1} \left(\left[t(\phi) \delta\left(m \frac{2\pi}{N} - \phi\right) \right] \otimes_{\phi} \text{rectp}\left(\frac{\phi}{\frac{2\pi}{N}}\right) \right) e^{-i\alpha\phi} d\phi \\ &= \sum_{m=0}^{N-1} \frac{1}{2\pi} \int_0^{2\pi} \left(\left[t(\phi) \delta\left(m \frac{2\pi}{N} - \phi\right) \right] \otimes_{\phi} \text{rectp}\left(\frac{\phi}{\frac{2\pi}{N}}\right) \right) e^{-i\alpha\phi} d\phi \end{aligned} \quad (5.78)$$

Anwenden des Faltungstheorems für die Fourierreihe erlaubt es die Faltung in eine Multiplikation zweier Integrale zu überführen, die dann separat bestimmt werden können:

$$\begin{aligned} G_{\alpha} &= \sum_{m=0}^{N-1} 2\pi \frac{1}{2\pi} \int_0^{2\pi} \left[t(\phi) \delta\left(m \frac{2\pi}{N} - \phi\right) \right] e^{-i\alpha\phi} d\phi \times \frac{1}{2\pi} \int_0^{2\pi} \text{rectp}\left(\frac{\phi}{\frac{2\pi}{N}}\right) e^{-i\alpha\phi} d\phi \\ &= \sum_{m=0}^{N-1} t\left(\phi = \frac{2\pi m}{N}\right) e^{-i2\pi\alpha \frac{m}{N}} \times \frac{1}{2\pi} \frac{2\pi}{N} \text{sinc}\left(\frac{2\pi}{N} \frac{\alpha}{2\pi}\right) \\ &= \sum_{m=0}^{N-1} t\left(\phi = \frac{2\pi m}{N}\right) e^{-i2\pi\alpha \frac{m}{N}} \times \frac{1}{N} \frac{\sin\left(\frac{\alpha\pi}{N}\right)}{\frac{\alpha\pi}{N}} \\ &= \sum_{m=0}^{N-1} t\left(\phi = \frac{2\pi m}{N}\right) e^{-i2\pi\alpha \frac{m}{N}} \frac{\sin\left(\frac{\pi\alpha}{N}\right)}{\alpha\pi} \\ &= \frac{\sin\left(\frac{\pi\alpha}{N}\right)}{\alpha\pi} \sum_{m=0}^{N-1} t\left(\phi = \frac{2\pi m}{N}\right) e^{-i2\pi\alpha \frac{m}{N}} \end{aligned} \quad (5.79)$$

Die Effizienz eines Elements, mit einer in N äquidistante Level diskretisierten Phase ist durch das Amplituden-Betragsquadrat des ersten Koeffizienten gegeben:

$$\eta_p = |G_1|^2 \quad (5.80)$$

$$= \frac{\sin^2\left(\frac{\pi}{N}\right)}{\pi^2} \left| \sum_{m=0}^{N-1} e^{i\frac{2\pi m}{N}} e^{-i2\pi \frac{m}{N}} \right|^2 \quad (5.81)$$

$$= \frac{\sin^2\left(\frac{\pi}{N}\right)}{\pi^2} N^2 \quad (5.82)$$

Spezialfall eines binären Element mit unterschiedlicher Phasenhöhe Der Spezialfall eines binären Phasenelements, bei dem die Phasenlevel nicht äquidistant sind, lässt sich folgendermaßen betrachten:

$$p(\phi) = \begin{cases} 0 & \text{für } 0 < \text{mod}(\phi, 2\pi) \leq \pi \\ \pi c & \text{für } \pi < \text{mod}(\phi, 2\pi) \leq 2\pi \end{cases}. \quad (5.83)$$

Typische Gründe für so eine Abweichung sind zum Beispiel die Verwendung einer anderen Wellenlänge zur Beleuchtung oder eine Abweichung in der Fabrikation der Phasenhöhe.

Für $N = 2$ ergibt sich für die Fourier-Koeffizienten

$$G_\alpha = \frac{\sin\left(\frac{\pi\alpha}{N}\right)}{\alpha\pi} \sum_{m=0}^{N-1} t\left(p\left(\phi = \frac{2\pi m}{N}\right)\right) e^{-i2\pi\alpha\frac{m}{N}} \quad (5.84)$$

$$= \frac{\sin\left(\frac{\pi\alpha}{2}\right)}{\alpha\pi} \left(t(\phi = 0) e^{-i2\pi\alpha\frac{0}{N}} + t(\phi = \pi c) e^{-i2\pi\alpha\frac{1}{2}} \right) \quad (5.85)$$

$$= \frac{\sin\left(\frac{\pi\alpha}{2}\right)}{\alpha\pi} \left(1 + e^{i\pi c} e^{-i\pi\alpha} \right). \quad (5.86)$$

Insbesondere für $\alpha = 1$ ist der erste Koeffizient gegeben durch

$$G_1 = \frac{1}{\pi} \left(1 - e^{i\pi c} \right). \quad (5.87)$$

Somit ergibt sich für die Modulationsfunktion folgende Effizienz in Abhängigkeit von der relativen Phasenhöhe c

$$\eta_p = |G_1|^2 = \left(\frac{2}{\pi} \right)^2 \sin^2\left(\frac{\pi c}{2}\right). \quad (5.88)$$

Man sieht, dass das Ergebnis für die Effizienz eines binären Phasenelements mit zusätzlicher \sin^2 Modulation entspricht.

5.2.4 Spiegelsymmetrie des Intensitätsprofils für binarisierte Phasenelemente mit niedriger numerischer Apertur

Für binäre Phasenelemente mit quadratischer Grundperiode P und Phasenhöhen-differenz $\Delta\phi = \pi$ sowie einer niedrigen numerischen Apertur gilt, dass die Intensitätsverteilung an der Ebene der halben Talbot-Distanz gespiegelt ist:

$$I(x, y, z) = I(x, y, z_T - z). \quad (5.89)$$

Hierbei ist die Talbot-Distanz z_T gegeben durch [94]

$$z_T = \frac{2P^2}{\lambda}. \quad (5.90)$$

Beweis Sei $t(x,y)$ darstellbar als Fourier-Reihe $t(x,y) = \sum_{mn} c_{mn} e^{2\pi i(x\frac{m}{P} + y\frac{n}{P})}$ mit $c_{mn} = \int_0^P \int_0^P t(x,y) e^{-2\pi i(x\frac{m}{P} + y\frac{n}{P})} dx dy$. Dann ist die komplexe Amplitude in der Ebene z gegeben durch:

$$U(x,y,z) = e^{ikz} \sum_{m,n} c_{mn} e^{2\pi i(\frac{mx}{P} + \frac{ny}{P})} e^{-\pi iz\lambda \left(\left(\frac{m}{P}\right)^2 + \left(\frac{n}{P}\right)^2 \right)}. \quad (5.91)$$

Da das Phasenelement nur eine niedrige numerische Apertur hat, ist die verwendete Beschreibung in Fresnel-Näherung exakt. In einer weiteren Ebene $z_T - z$ ist die komplexe Amplitude gegeben durch:

$$\begin{aligned} U(x,y,z_T - z) &= e^{i\frac{2\pi}{\lambda} \frac{2P^2}{\lambda}} e^{-ikz} \sum_{m,n} c_{mn} e^{2\pi i(\frac{mx}{P} + \frac{ny}{P})} e^{-2\pi i(m^2 + n^2)} e^{+\pi iz\lambda \left(\left(\frac{m}{P}\right)^2 + \left(\frac{n}{P}\right)^2 \right)} \\ &= e^{i\frac{4\pi P^2}{\lambda^2}} e^{-ikz} \sum_{m,n} c_{mn} e^{2\pi i(\frac{mx}{P} + \frac{ny}{P})} e^{+\pi iz\lambda \left(\left(\frac{m}{P}\right)^2 + \left(\frac{n}{P}\right)^2 \right)} \end{aligned} \quad (5.92)$$

Für Elemente mit Phasenhöhe 0 und π ist $t(x)$ rein reell, somit gilt $c_{m\ n} = c_{-m\ -n}^*$:

$$\begin{aligned} U(x,y,z_T - z) &= e^{i\frac{4\pi P^2}{\lambda^2}} e^{-ikz} \sum_{m,n} c_{-m-n}^* e^{+2\pi i(\frac{mx}{P} + \frac{ny}{P})} e^{+\pi iz\lambda \left(\left(\frac{m}{P}\right)^2 + \left(\frac{n}{P}\right)^2 \right)} \\ &= e^{i\frac{4\pi P^2}{\lambda^2}} e^{-ikz} \sum_{m,n} c_{mn}^* e^{-2\pi i(\frac{mx}{P} + \frac{ny}{P})} e^{+\pi iz\lambda \left(\left(\frac{m}{P}\right)^2 + \left(\frac{n}{P}\right)^2 \right)} \\ &= e^{i\frac{4\pi P^2}{\lambda^2}} U^*(x,y,z). \end{aligned} \quad (5.93)$$

Im allgemeinen Fall ist nur die Phasendifferenz $\Delta\phi = \pi$ vorgegeben. Durch herausziehen einer konstanten Phase ϕ_{offset} aus der Definition der Koeffizienten c_{mn} lässt sich der allgemeine Fall auf den obigen Spezialfall abbilden:

$$\begin{aligned} U(x,y,z_T - z) &= e^{i\frac{4\pi P^2}{\lambda^2}} e^{i\phi_{\text{offset}}} e^{-ikz} \sum_{m,n} c_{-m-n}^* e^{+2\pi i(\frac{mx}{P} + \frac{ny}{P})} e^{+\pi iz\lambda \left(\left(\frac{m}{P}\right)^2 + \left(\frac{n}{P}\right)^2 \right)} \\ &= e^{i\frac{4\pi P^2}{\lambda^2}} e^{i\phi_{\text{offset}}} e^{-ikz} \sum_{m,n} c_{mn}^* e^{-2\pi i(\frac{mx}{P} + \frac{ny}{P})} e^{+\pi iz\lambda \left(\left(\frac{m}{P}\right)^2 + \left(\frac{n}{P}\right)^2 \right)} \\ &= e^{i\frac{4\pi P^2}{\lambda^2}} e^{i\phi_{\text{offset}}} U^*(x,y,z). \end{aligned} \quad (5.94)$$

Somit gilt für die Intensität

$$I(x,y,z_T - z) = I(x,y,z). \quad (5.95)$$

Man sieht, dass die Intensität an der Talbot-Ebene gespiegelt ist. Dies bedeutet insbesondere auch, dass die Foki an der Talbot-Ebene gespiegelt sind.

Berücksichtigt man zusätzlich, dass die Intensitätsverteilung im Abstand der halben Talbot-Distanz um eine halbe Periode in x und y-Richtung verschoben ist [45], dann gilt insbesondere für die Ebene $z = f + \frac{z_T}{2}$ dass $I(x, y, f + \frac{z_T}{2}) = I(x - \frac{P}{2}, y - \frac{P}{2}, f)$ gilt. Innerhalb einer Talbot-Distanz erhält man bei einem binären Phasenelement nun in vier Ebenen die designten Foki und zwar bei f , $\frac{z_T}{2} - f$, $\frac{z_T}{2} + f$ und $z_T - f$, wenn $f \leq \frac{z_T}{4}$ ist.

5.3 Wählen der optimalen Phasendiskretisierung

Im vorherigen Abschnitt haben wir gesehen, dass die Phasendiskretisierung einen starken Einfluss auf die Intensität im Fokus hat. Für die Wahl einer guten Binarisierung lässt sich zum Beispiel die Intensität im Fokus, die Intensität innerhalb der ersten Nullstelle oder die Energie, die durch das Fokusellipsoid fließt, in Abhängigkeit des ϵ -Parameters bestimmen. Da jedoch besonders bei geringen numerischen Aperturen eine starke Modulation innerhalb des Fokus auftritt (vgl. Abschnitt 5.2.2), scheint die Beurteilungsform nicht optimal. Auch für komplexere Fokalgeometrien würde diese Art der Optimierung versagen.

Anstatt die Binarisierung über die Eigenschaften des Fokus zu definieren lässt sich die Binarisierung durch die Ähnlichkeit zu der nicht binarisierten Verteilung beurteilen. Dies ist in sofern auch besonders sinnvoll, da die Verteilung so gewählt wurde, dass ein optimaler Fokus erzeugt wird (vgl. Abschnitt 5.1.5).

Hierzu herrscht in der Literatur die Meinung [51, 52, 77], dass für nur phasenmodulierende Elemente die Binarisierung die beste ist, bei der der Abstand in der komplexen Ebene zwischen dem ursprünglichen Element und dem binarisierten Element am geringsten ist.

Dieses Optimierungsproblem ist zweidimensional und leicht durch einen Brute-Force-Ansatz zu lösen. Um das optimale Element zu finden variiert man in der ersten Dimension die Phasenhöhe, in der zweiten Dimension schiebt man mit dem ϵ -Parameter die Phasenpunkte auf dem Einheitskreis, (vgl. 4.64).

Für ein binäres Phasenelement mit den Phasenhöhen $\phi_0 = 0 - 2\pi\epsilon$ und $\phi_1 = \Delta\phi - 2\pi\epsilon$ lässt sich für jeden Pixel j die euklidische Distanz berechnen zwischen den ursprünglichen Werten t_j^{DOE} und den Projektionswerten $e^{i\phi_0}$ und $e^{i\phi_1}$:

$$d_j^{\phi_0} = \sqrt{(\text{Re}[t_j^{\text{DOE}}] - \cos \phi_0)^2 + (\text{Im}[t_j^{\text{DOE}}] - \sin \phi_0)^2} \quad (5.96)$$

$$d_j^{\phi_1} = \sqrt{(\text{Re}[t_j^{\text{DOE}}] - \cos \phi_1)^2 + (\text{Im}[t_j^{\text{DOE}}] - \sin \phi_1)^2} \quad (5.97)$$

Anschließend wird der Pixel auf den Wert gesetzt, bei dem die euklidische Distanz am geringsten war. Für das gesamte Element lassen sich nun die minimalen euklidischen Distanzen addieren:

$$d_{\text{Euklid}} = \sum_j \min(d_j^{\phi_0}, d_j^{\phi_1}). \quad (5.98)$$

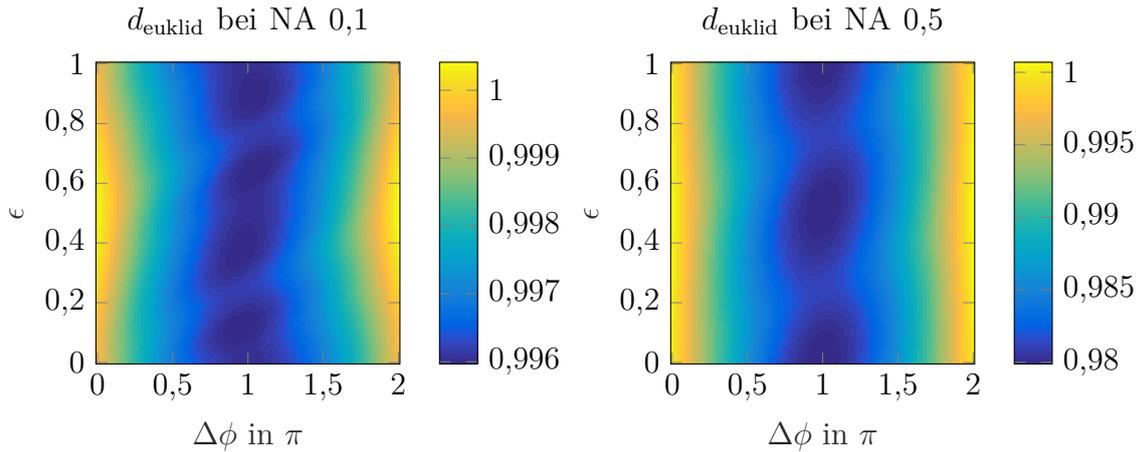


Abbildung 5.14: Einfluss der Phasenhöhe und des ϵ -Parameters bei der Binarisierung auf die durchschnittliche euklidische Distanz zweier binärer Phasenelemente mit Periode $50\ \mu\text{m}$ und Brennweite $500\ \mu\text{m}$. Die numerische Apertur beträgt oben $0,1$ und unten $0,5$. Dies entspricht einem Überlapp von ca. 2 bzw. $11,5$.

Nun kann für jede Phasenhöhe $\Delta\phi$ und jede Binarisierungsschwelle ϵ dieses Maß ausgerechnet werden.

In Abbildung 5.14 sieht man ein Beispiel für die durchschnittliche euklidische Distanz d_{Euklid} als Funktion der Phasenhöhe und der Binarisierungsschwelle ϵ für je ein Element mit numerischer Apertur von $0,1$ und $0,5$. Das optimale Element ist nun das mit der minimalen euklidischen Distanz.

Für kleine numerische Aperturen ist die durchschnittliche Amplitude fast Null, deshalb ist die durchschnittliche Abweichung nahezu immer 1 . Für höhere numerische Aperturen ist die durchschnittliche Amplitude größer als bei kleinen Amplituden, deshalb ist hier eine kleine durchschnittliche euklidische Distanz möglich.

Man sieht, dass die optimale Phasenhöhe meist bei π für alle Epsilonwerte liegt. Jedoch zeigt sich, dass manchmal eine kleine Abweichung von π vorteilhaft sein kann.

5.4 Einfluss der Pixelierung des diffraktiven Elements

Bei der Herstellung eines diffraktiven Elements wird meist eine pixelierte Version fabriziert. Mathematisch lässt sich die Pixelierung des Elements als Kombination aus Abtastung und Faltung beschreiben.

$$t_{\text{DOE}}^{\text{pixeliert}}(\mathbf{r}_{\perp}) = t_{\text{DOE}}^{\text{abgetastet}}(\mathbf{r}_{\perp}) \otimes_{\mathbf{r}_{\perp}} \text{rect}\left(\frac{x}{dx}\right) \text{rect}\left(\frac{y}{dy}\right) \quad (5.99)$$

$$t_{\text{DOE}}^{\text{abgetastet}}(\mathbf{r}_{\perp}) = t_{\text{DOE}}(\mathbf{r}_{\perp}) \sum_{n,m} \delta(x - n dx, y - n dy) \quad (5.100)$$

Hierbei beschreibt t_{DOE} die Ausgangsverteilung und dx bzw. dy die Breite der Pixel.

Einfluss der Pixelierung auf den Fokus

Im Ortsraum lässt sich die Ausbreitung von der Elementebene bis in die Fokusebene als Faltung mit der Punktbildverwaschungsfunktion $h_z(\mathbf{r}_\perp)$ beschreiben.

$$U_f(\mathbf{r}_\perp) = t_{\text{DOE}}^{\text{pixeliert}}(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} h_z(\mathbf{r}_\perp) \quad (5.101)$$

Diese Gleichung lässt sich in den Frequenzraum transformieren. Im Frequenzraum bedeutet die Pixelierung, dass die Fourier-Transformierte mit einer sinc-Funktion moduliert ist. Die Ausbreitung in die Fokalebene wird durch eine Modulation mit dem Propagator P_z beschrieben. Man erhält somit für das Spektrum in der Fokusebene

$$\tilde{U}_f(\boldsymbol{\rho}_\perp) = \tilde{t}_{\text{DOE}}^{\text{abgetastet}}(\boldsymbol{\rho}_\perp) dx dy \text{sinc}(dx \nu) \text{sinc}(dy \mu) P_z(\boldsymbol{\rho}_\perp) \quad (5.102)$$

Vor der Rücktransformation in den Ortsraum lässt sich die Reihenfolge von Multiplikation mit dem Propagator und den sinc-Termen vertauschen.

$$\tilde{U}_f(\boldsymbol{\rho}_\perp) = \tilde{U}_f^{\text{abgetastet}}(\boldsymbol{\rho}_\perp) dx dy \text{sinc}(dx \nu) \text{sinc}(dy \mu) \quad (5.103)$$

$$\tilde{U}_f^{\text{abgetastet}}(\boldsymbol{\rho}_\perp) = \tilde{t}_{\text{DOE}}^{\text{abgetastet}}(\boldsymbol{\rho}_\perp) P_z(\boldsymbol{\rho}_\perp) \quad (5.104)$$

Das Spektrum in der Fokalebene lässt sich also als das ausgebreitete Spektrum des abgetasteten diffraktiven Elements betrachten, das mit sinc-Funktionen moduliert ist. Die Rücktransformierte im Ortsraum liefert

$$U_f(\mathbf{r}_\perp) = U_f^{\text{abgetastet}}(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} \text{rect}\left(\frac{x}{dx}\right) \text{rect}\left(\frac{y}{dy}\right) \quad (5.105)$$

$$U_f^{\text{abgetastet}}(\mathbf{r}_\perp) = t_{\text{DOE}}^{\text{abgetastet}}(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} h_z(\mathbf{r}_\perp). \quad (5.106)$$

Man sieht, dass die Pixelierung des Elements eine Faltung der Amplitude mit der Breite des Pixels bedingt.

Auswirkungen für die Simulation und das Design diffraktiver Linsen

Aus dieser Untersuchung der Pixelierung lassen sich mehrere Dinge für die Simulation und das Design von diffraktiven Linsen folgern:

- Der Einfluss der Pixelierung lässt sich in der Simulation durch eine Sinc-Modulation des Spektrums abbilden.
- Die minimale Fokusbreite ist durch die Pixelgröße beschränkt, da die Amplitude im Fokus mit der Pixelgröße gefaltet ist.

Mit Hilfe des Abbeschen Auflösungskriterium lässt sich hieraus eine Beschränkung der numerischen Apertur ableiten. Die maximale numerische Apertur bei einer Pixelgröße von dx ist somit durch $\text{NA}_{\text{max}} = \frac{\lambda}{2dx}$ gegeben.

- Die Pixelierung des Elements verursacht eine Reduzierung der Amplitude im Fokus. Vergleicht man die Amplitude $U_f^{\text{pixeliert}}$ im Fokus des pixelierten Elements mit der Amplitude $U_f^{\text{nicht pixeliert}}$ des Fokus eines nicht pixelierten Elements, so bemerkt man eine Reduzierung durch die zusätzliche Faltung mit der Pixel-Breite:

$$U_f^{\text{pixeliert}}(\mathbf{r}_\perp) = t_{\text{DOE}}^{\text{abgetastet}}(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} h_z(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} \text{rect}\left(\frac{x}{dx}\right) \text{rect}\left(\frac{y}{dy}\right) \quad (5.107)$$

$$U_f^{\text{nicht pixeliert}}(\mathbf{r}_\perp) = t_{\text{DOE}}(\mathbf{r}_\perp) \otimes_{\mathbf{r}_\perp} h_z(\mathbf{r}_\perp). \quad (5.108)$$

Dieser Effekt ist vernachlässigbar, wenn die Pixelbreite viel kleiner ist als die Breite der Punktbildverwaschungsfunktion h_z .

5.5 Auswirkung der Endlichkeit der Periodizität

Im Unterschied zu den idealen diffraktiven Elementen besitzen hergestellte diffraktive Elemente nur eine endliche Anzahl an Wiederholungen der Grundperiode. Dies hat zur Folge, dass die Randperioden gestört sind. Ein Maß dafür, wie viele Randperioden gestört sind, ist durch den räumlichen Überlappfaktor Ω vorgegeben.

Um dies zu verdeutlichen ist in Abbildung 5.15 ein Ausschnitt eines Linescans durch die Foki eines diffraktiven Elements mit 100x100 Perioden gezeigt. Der Überlappfaktor des Elements beträgt 48. Man sieht, dass links und rechts die ersten 48 Foki in der Intensität vermindert sind. Die Foki 49 bis 100 sind Intensitätsmäßig alle sehr ähnlich und nicht gestört. Die Anzahl der nicht gestörten Foki lässt sich ermitteln, indem man von der Gesamtzahl der Perioden den räumlichen Überlapp abzieht.

Auswirkungen für die Herstellung von diffraktiver Linsen mit überlappenden Aperturen

Ein reales diffraktives Element hat nur eine begrenzte Anzahl an periodischen Wiederholungen der Grundperiode. Aus der Anzahl der gewünschten ungestörten Foki n_{Foki} , der Periodengröße P und dem effektiven Aperturdurchmesser D_{eff} lässt sich die minimale Elementbreite D_{Element} für das diffraktive Linsenarray mit überlappenden Aperturen bestimmen:

$$D_{\text{Element}} \approx D_{\text{eff}} + n_{\text{Foki}} P. \quad (5.109)$$

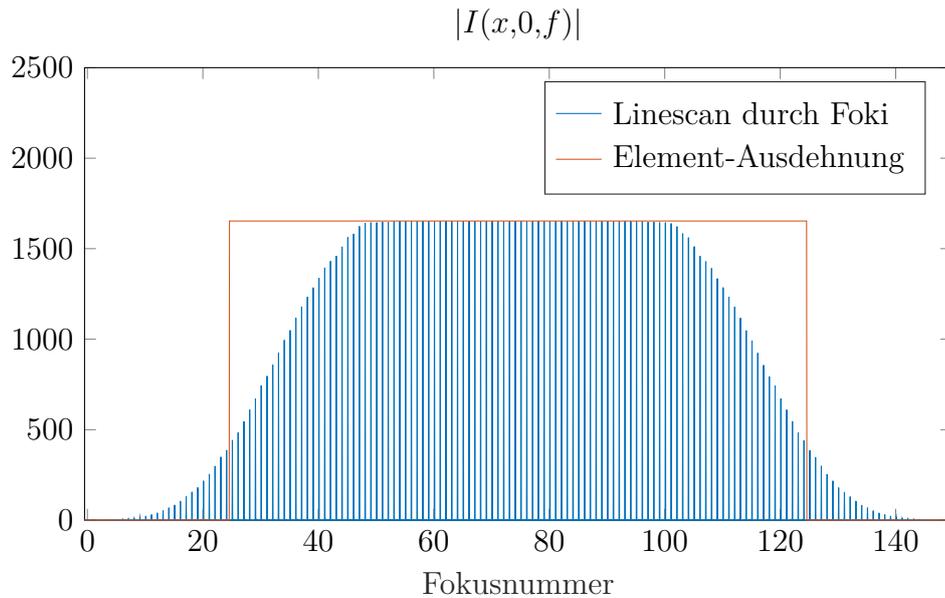


Abbildung 5.15: Linescan durch die Foki eines diffraktiven Elements mit 100 Wiederholungen der Grundperioden. Der Überlappfaktor des Elements beträgt 48.

5.6 Abweichungen der Design-Parameter

In der Praxis kommt es oft vor, dass die diffraktiven Elemente nicht genau mit den im Design spezifizierten Parametern betrieben werden. Dies kann zum Beispiel dadurch begründet sein, dass man das diffraktive Element mit einer anderen Lichtquelle benutzen will, die sich jedoch in der Wellenlänge unterscheidet. Neben diesen bewusst in Kauf genommenen Abweichungen gibt es auch Abweichungen, die durch Fehler oder Toleranzen im Fertigungsprozess der diffraktiven Elementen entstehen, wie zum Beispiel eine Abweichung in der Phasenhöhe des diffraktiven Elements. Solche Abweichungen werden nun in den folgenden Abschnitten untersucht.

5.6.1 Abweichungen in der Wellenlänge

Beschreibung für eine ideale Phasenverteilung Für eine ideale diffraktive Linse mit kontinuierlichem Frequenzspektrum lässt sich eine Abweichung von Beleuchtungs- und Design-Wellenlänge analytisch untersuchen. Betreibt man das diffraktive Element mit einer Wellenlänge λ die von der Designwellenlänge λ_{Design} unterscheidet,

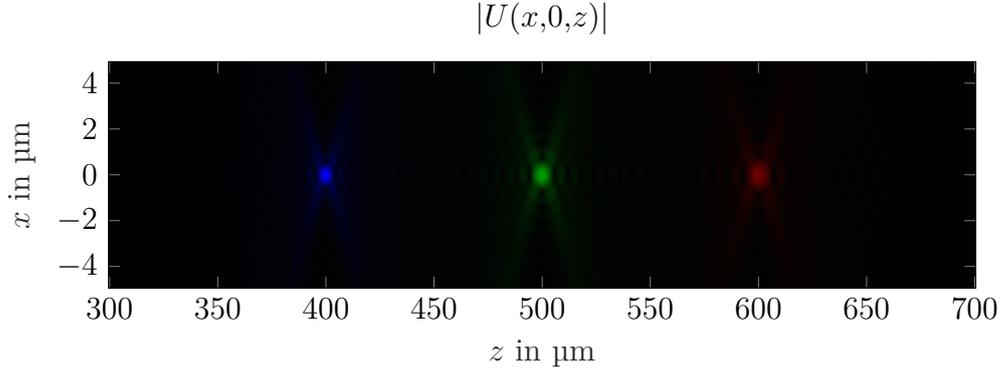


Abbildung 5.16: Ausschnitt eines xz-Schnitts für ein DOE mit $NA = 0,4$, $P = 50 \mu\text{m}$, $f = 500 \mu\text{m}$ bei $\lambda_{\text{Design}} = 0,5 \mu\text{m}$. Der xz-Schnitt wurde für Beleuchtung des DOE mit Licht der Wellenlänge $\lambda_{\text{blau}} = 0,4 \mu\text{m}$, $\lambda_{\text{grün}} = 0,5 \mu\text{m}$ und $\lambda_{\text{rot}} = 0,6 \mu\text{m}$ berechnet.

dann gilt für die ausgebreitete Amplitude

$$U(r_{\perp}, z) = \frac{k^2}{2\pi} \int_0^{\infty} U_0 \text{circ}\left(\frac{s_{\perp}}{NA}\right) e^{-ik_{\text{Design}} f \sqrt{n^2 - s_{\perp}^2}} e^{ikz \sqrt{n^2 - s_{\perp}^2}} J_0(k s_{\perp} r_{\perp}) s_{\perp} ds_{\perp} \quad (5.110)$$

$$= \frac{k^2}{k_{\text{Design}}^2} \frac{k_{\text{Design}}^2}{2\pi} U_0 \int_0^{\infty} \text{circ}\left(\frac{s_{\perp}}{NA}\right) e^{i\left(\frac{k}{k_{\text{Design}}} z - f\right) k_{\text{Design}} \sqrt{n^2 - s_{\perp}^2}} J_0(k s_{\perp} r_{\perp}) s_{\perp} ds_{\perp} \quad (5.111)$$

Hierbei ist $k_{\text{Design}} = \frac{2\pi}{\lambda_{\text{Design}}}$ die Wellenzahl der Design-Wellenlänge. Ersetzt man $\frac{k}{k_{\text{Design}}} z = z'$ so erhält man

$$U(r_{\perp}, z') = \frac{k^2}{k_{\text{Design}}^2} \frac{k_{\text{Design}}^2}{2\pi} U_0 \int_0^{\infty} \text{circ}\left(\frac{s_{\perp}}{NA}\right) e^{i(z' - f) k_{\text{Design}} \sqrt{n^2 - s_{\perp}^2}} J_0(k s_{\perp} r_{\perp}) s_{\perp} ds_{\perp} \quad (5.112)$$

Auf der optischen Achse $r_{\perp} = 0$ sieht man, dass bei einer Abweichung von der Wellenlänge nur die Amplitude mit $\frac{k^2}{k_{\text{Design}}^2}$ und die z-Achse mit $\frac{k}{k_{\text{Design}}}$ skaliert wird. Im Fokus gilt $\frac{k}{k_{\text{Design}}} z - f_{\text{Design}} = 0$, somit liegt der Fokus dieses Elements jetzt bei $f = f_{\text{Design}} \frac{\lambda}{\lambda_{\text{Design}}}$. Ein Beispiel hierzu ist in Abbildung 5.16 gezeigt.

Das hier in der Rechnung benutzte analytische diffraktive Element verhält sich wie eine refraktive Linse mit normaler Dispersion, da eine kontinuierliche und nicht gewrappte Phasenverteilung für das diffraktive Element angenommen wurde.

Bei einem fabrizierten diffraktiven Element ist die kontinuierliche Phase für die Design-Wellenlänge auf den Bereich zwischen 0 und 2π projiziert.

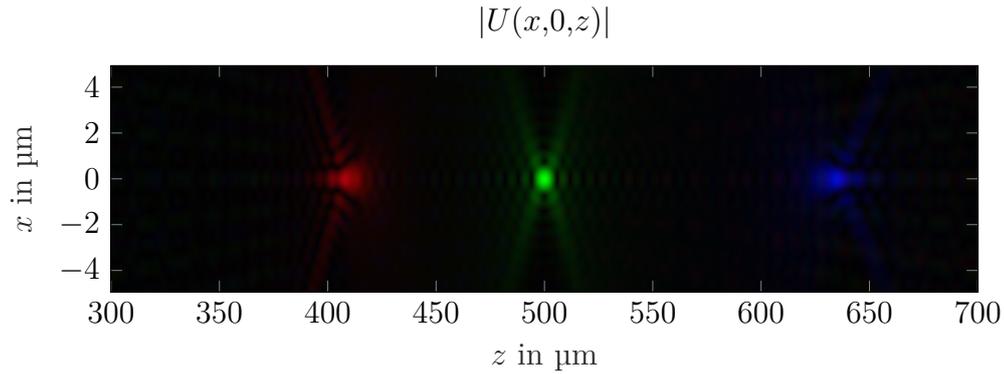


Abbildung 5.17: Ausschnitt eines xz-Scan für ein DOE mit $NA = 0,4$, $P = 50 \mu\text{m}$, $f = 500 \mu\text{m}$ bei $\lambda_{\text{Design}} = 0,5 \mu\text{m}$. Der xz-Schnitt wurde für Beleuchtung des DOE mit Licht der Wellenlänge $\lambda_{\text{blau}} = 0,4 \mu\text{m}$, $\lambda_{\text{grün}} = 0,5 \mu\text{m}$ und $\lambda_{\text{rot}} = 0,6 \mu\text{m}$ berechnet.

Beschreibung für ein real fabriziertes Element Die ideale analytische Beschreibung aus dem vorherigen Abschnitt ist für ein real fabriziertes Element aus mehreren Gründen nicht gültig. Erstens ist die Diskretisierung im Frequenzraum von der Design-Wellenlänge abhängig. Zweitens wird in der idealen Beschreibung eine kontinuierliche Designphase ϕ^{Design} vorausgesetzt, in einem realen Element wird die Phasenverteilung jedoch durch eine unterschiedliche lange Propagationsstrecke d im Brechungsmedium n verursacht. Die so erzeugten Phasenverteilung wird meistens auf einen Bereich von 0 bis 2π eingeschränkt:

$$\phi^{\text{Design}}(\mathbf{r}_{\perp}) = \text{mod} \left((n-1)k_0^{\text{Design}} d(\mathbf{r}_{\perp}), 2\pi \right). \quad (5.113)$$

Weicht man bei einem real fabrizierten Element von der Designwellenlänge ab, so zeigt sich im Gegensatz zur idealen Phasenverteilung eine anormale Dispersion. Dies bedeutet, dass rot verschobenes Licht die Foki näher am Element und blau verschobenes Licht weiter weg von der Design-Fokusslänge erzeugt. In Abbildung 5.17 ist der Fokalverlauf eines diffraktiven Elements mit anormaler Dispersion gezeigt. Man sieht außerdem, dass die Foki, bei denen die Propagationswellenlänge von der Designwellenlänge abweicht, aberriert sind.

5.6.2 Abweichungen in der Phasenhöhe

Durch den Fertigungsprozess können Abweichungen zur Design-Phasenhöhe entstehen. Im Folgenden werden die Abweichungen der Phasenhöhe für ein binarisiertes diffraktives Element mit konstanter Amplitude untersucht. Hierzu lässt sich die Intensität im Fokus oder der Energiefluss durch das Fokusellipsoid betrachten. In Abbildung 5.18 ist die Fokusintensität als Funktion der Phasenhöhe für ein binäres Phasenelement mit einer numerischen Apertur von $0,5$ gezeigt.

Man sieht eine Abhängigkeit der Form \sin^2 . Dies entspricht der theoretischen Vorhersage aus Gleichung 5.88. Die maximale relative Intensität ist konsistent mit dem

Auswirkung der Phasenhöhe auf die Intensität im Fokus

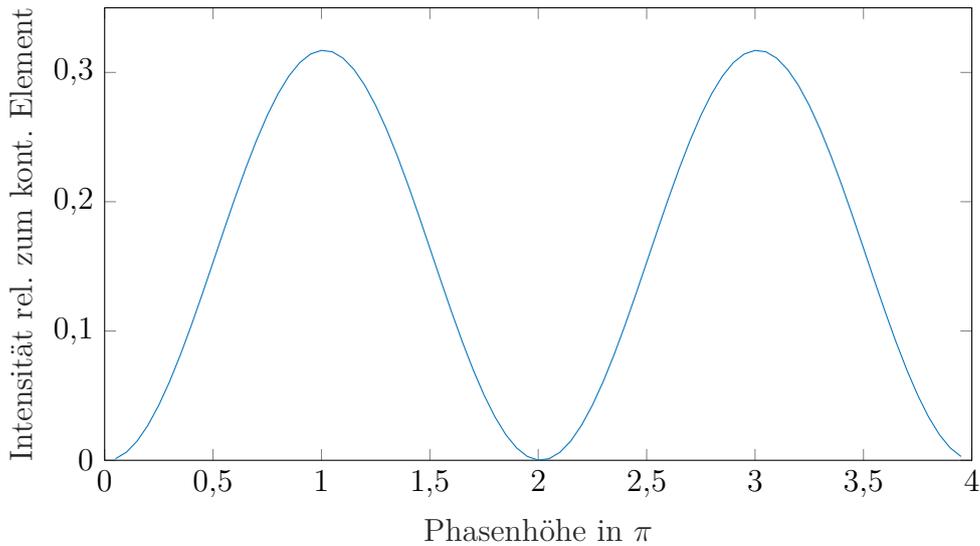


Abbildung 5.18: Auswirkung der Phasenhöhen eines binären Phasenelements auf die Intensität im Fokus. Die numerische Apertur des diffraktiven Elements ist 0,5, die Periode beträgt $50 \mu\text{m}$, die Pixelgröße ist $0,125 \mu\text{m}$, die Fokallänge $2000 \mu\text{m}$ und die benutzte Wellenlänge ist $0,532 \mu\text{m}$.

Ergebnis aus Abschnitt 5.2.1 für ein phasenbinarisierendes und amplitudennivelliertes Element. Kleine Änderungen im Vergleich zur optimalen Phasenhöhe ändern die Intensität im Fokus nur geringfügig. Die 2π -Periodizität ist durch die dünne Elementnäherung begründet, die dafür sorgt, dass die Elemente mit Phasenhöhe $\Delta\phi$ und $\Delta\phi + 2m\pi$ mit $m \in \mathbb{Z}$ durch dieselben Werte im Intervall von 0 bis 2π beschrieben werden.

Untersucht man hingegen die Intensität im Fokus für kleinere numerische Aperturen, so sieht man, dass die \sin^2 -Abhängigkeit bezüglich der Phasenhöhe leicht lateral verschoben sein kann. Dies bedeutet, dass bei einer Phasenhöhe, die leicht von π abweicht, das Maximum der Intensität im Fokus liegt. Ein Beispiel für die Intensität im Fokus eines binären diffraktiven Elements mit numerischer Apertur von 0,1 ist in Abbildung 5.19 gezeigt.

Diese Beobachtung ist konsistent mit den Ergebnissen aus Abschnitt 5.3. Dort ist gezeigt, dass die ideale Phasenhöhe bei binarisierten diffraktiven Elementen von π abweichen kann. Die deutlichen Abweichungen der relativen Intensität, im Vergleich zum Beispiel mit der mittleren numerischen Apertur, sind durch die Störungen des Fokalprofils zu erklären. Die Abweichung ist so groß, da die numerische Apertur von 0,1 aus diesem Beispiel sich in der Nähe eines Minimums des Strehl-Verhältnisses befindet.

Auswirkung der Phasenhöhe auf die Intensität im Fokus

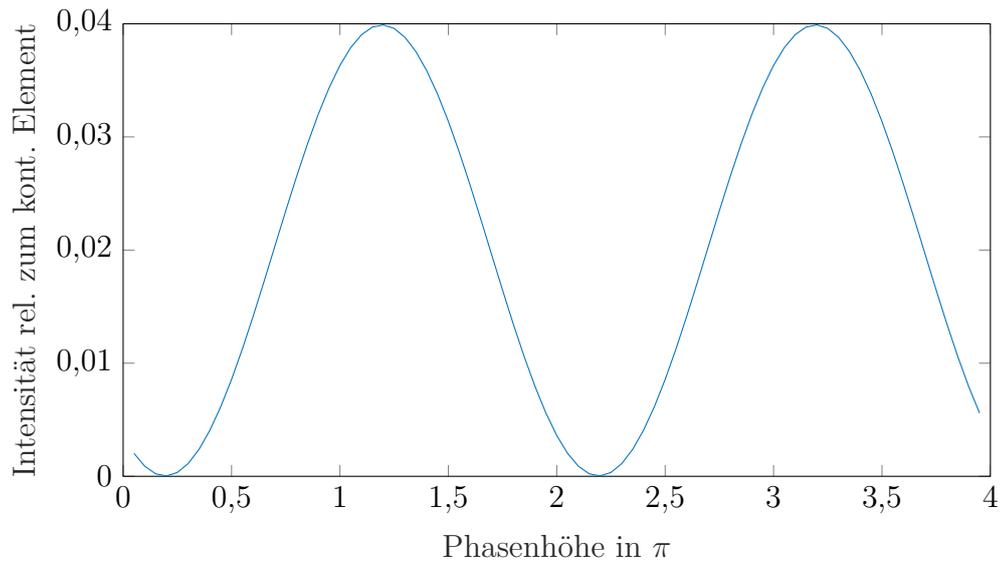


Abbildung 5.19: Auswirkung der Phasenhöhen eines binären diffraktiven Elements auf die Intensität im Fokus. Die numerische Apertur des diffraktiven Elements ist 0,1, die Periode beträgt $50 \mu\text{m}$, die Pixelgröße ist $0,125 \mu\text{m}$, die Fokusslänge ist $2000 \mu\text{m}$ und die benutzte Wellenlänge ist $0,532 \mu\text{m}$.

5.6.3 Abweichungen durch verwaschene Kanten

Für die Herstellung wird das diffraktive Element pixeliert, hierbei werden harte Kanten angenommen, die im realen Fertigungsprozess nur näherungsweise realisiert werden können. Um den Einfluss dieses Effektes abschätzen zu können wurde die binäre Phasenverteilung mit einer Gaußfunktion $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{r_1^2}{2\sigma^2}}$ gefaltet und der Einfluss auf den Energiefluss durch den Fokus analysiert.

In Abbildung 5.20 ist ein Beispiel für den Einfluss der Verwaschung auf den Energiefluss durch das Fokusellipsoid gezeigt. Man sieht, dass eine leichte Verwaschung nur einen geringen Einfluss auf den Energiefluss im Fokusellipsoid hat.

Auswirkung der Verwaschung auf den Energiefluss im Fokusellipsoid

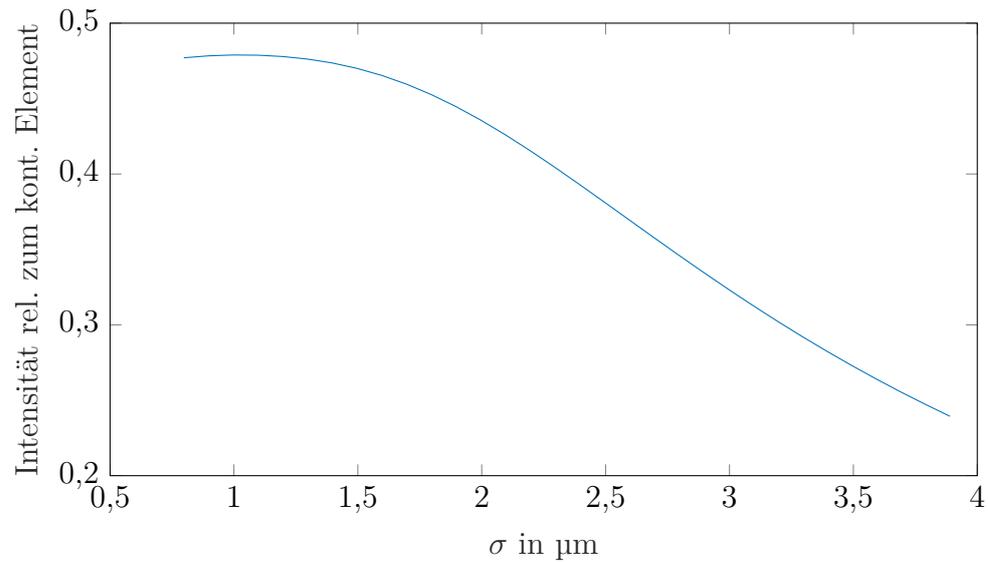


Abbildung 5.20: Auswirkung der Verwaschung eines binären diffraktiven Elements auf den Energiefluss durch das Fokusellipsoid. Die numerische Apertur des diffraktiven Elements ist 0,07, die Periode beträgt $52,8 \mu\text{m}$, die Pixelgröße ist $1,6 \mu\text{m}$, die Fokusslänge ist $3000 \mu\text{m}$ und die Wellenlänge ist $0,532 \mu\text{m}$. Beim Design wurde der Schwellwert $\epsilon = 0,07$ benutzt.

5.7 Leitfaden zur Wahl der Parameter im Design

Aus den vorherigen Abschnitten lassen sich zusammenfassend folgende allgemeine Aussagen zur Wahl der verschiedenen Design-Parameter der diffraktiven Linsen mit überlappenden Aperturen herleiten:

- Die Fokuslänge und die Periode des Grundelements sind frei wählbar und sollten auf das jeweilige Problem angepasst werden. Bei der Wahl der Periode sollte jedoch darauf geachtet werden, dass die Ausdehnung des Fokus deutlich kleiner als die Periodengröße ist.
- Im Elementdesign sollte die Pixelgröße möglichst klein gewählt werden, da sonst die Pixelgröße die minimale Fokusgröße limitiert.
- Für ein beugungsbegrenztes Fokalprofil ist eine höhere numerische Apertur notwendig, da die Amplituden-Nivellierung und die Phasen-Diskretisierung eine Störung des Fokusprofils erzeugen. Allgemein ist diese Fokusstörung für höhere numerische Aperturen geringer als für niedrige.
- Die maximale numerische Apertur ist durch die Elementgröße und die Anzahl der parallel zu nutzenden Foki, sowie die Ausdehnung der erzeugbaren ebenen Welle limitiert.
- Die optimale Binarisierungsschwelle und Phasenhöhe lässt sich durch einen Brute-Force-Ansatz ermitteln.

5.8 Fazit

In diesem Kapitel wurde gezeigt, dass das Design von diffraktiven Linsen mit überlappenden Aperturen mit homogener Pupille im Strehlschen Sinn optimal ist, da die idealen, nicht modifizierten Linsen ein generalisiertes Strehl-Kriterium von eins haben. Für sehr hohe numerische Aperturen bietet jedoch ein Design basierend auf Kugelwellen eine leicht bessere Fokussier-Effizienz, wenn man den totalen Energiefluss als Normierungsparameter benutzt. Für numerische Aperturen bis $NA = 0,8$ ist der Unterschied zwischen beiden Ansätzen in Bezug auf die Fokussier-Effizienz jedoch minimal.

Für niedrige numerische Aperturen wurde gezeigt, dass Modifikationen wie eine Phasendiskretisierung und eine Amplitudennivellierung das Fokusprofil stören. Für mittlere und hohe numerische Aperturen wird die Störung des Fokusprofils durch diese Modifikationen minimal und es lässt sich mit dem Strehl-Verhältnis eine Abweichung von der idealen Fokussier-Effizienz bestimmen. Gleichzeitig wurde analytisch die Effizienz für bestimmte Phasenmodulationsfunktionen betrachtet. Die so ermittelten Effizienzen für eine Phasendiskretisierung mit äquidistanten Phasenniveaus ist konsistent mit der Effizienz aus dem Strehl-Verhältnis für mittlere und hohe numerische Aperturen. Für niedrige Aperturen wurden Abweichungen gefunden, diese

zeigen auf, dass das Strehl-Verhältnis aufgrund der Störungen im Fokusprofil ungeeignet ist um die Intensität im Fokus zu beurteilen.

Des Weiteren wurde die Robustheit der diffraktiven Linsen bezüglich der Abweichung von einem idealen Design untersucht. Durch die Wahl einer geringen Pixelgröße des Elements lässt sich der Einfluss der Pixelierung des Elements reduzieren. Die Auswirkungen der Endlichkeit der Periodizität lässt sich minimieren, indem die Anzahl der Perioden im fabrizierten Element größer als der Überlapp-Faktor ist. Auch gegen Abweichungen von der Design-Phasenhöhe und der Design-Wellenlänge zeigen sich die diffraktiven Linsen robust.

6 Vermessung spezieller diffraktiver Elemente mit überlappenden Aperturen

Um die Funktion der diffraktiven Linsen mit überlappenden Aperturen experimentell zu bestätigen, wird in diesem Kapitel die Vermessung von zwei diffraktiven Linsenarrays mit mittlerer beziehungsweise hoher numerischer Apertur vorgestellt. Zuerst wird in Abschnitt 6.1 auf die Herstellung und die Designparameter der fabrizierten diffraktiven Linsenarrays eingegangen. Anschließend werden in Abschnitt 6.2 die Pupillen der diffraktiven Linsen gemessen, um eventuelle Abweichungen in der Phasenhöhe zu quantifizieren. Danach werden in Abschnitt 6.3 die Messungen der Fokusprofile für die beiden diffraktiven Elemente vorgestellt und mit einer skalaren Simulation verglichen. Hiermit kann die ordnungsgemäße Funktion der diffraktiven Linsen und der integrierten Deckglaskorrektur auf direkte Weise untersucht werden.

6.1 Herstellung von diffraktiven Elementen in Glas

In Zusammenarbeit mit der Karlsruhe Nano Micro Facility (KNMF) wurden zwei diffraktive Elemente auf einem Quarzglas-Substrat hergestellt. Die Strukturierung des Glases wurde durch Elektronenstrahl-Lithographie und reaktives Ionenätzen realisiert. Das Element mit der mittleren numerischen Apertur von $NA = 0,48$ wird im folgenden DOE48 genannt, analog dazu bezeichnet DOE75 das Element mit der hohen numerischen Apertur von $NA = 0,75$.

In Tabelle 6.1 sind die Designparameter der beiden diffraktiven Elemente zusammengefasst. Die Periode von $44\ \mu\text{m}$ wurde so groß gewählt, damit die Foki selbst bei einer Messung ohne Vergrößerungsoptik deutlich voneinander unterscheidbar sind. Bei einer Pixelgröße von $2,2\ \mu\text{m}$ wären die Foki genau 20 Pixel auseinander. Die Wellenlänge ist auf einen Nd:YAG-Festkörperlaser mit Kaliumtitanylphosphatkristall (KTP) angepasst, der eine Emissionswellenlänge von $\lambda = 532\ \text{nm}$ besitzt. Die numerische Apertur wurde beim Element DOE48 so gewählt, dass das Element einen Überlappfaktor von 50 hat. Die Pixelgröße dieses Elements wurde über das Abbesche Auflösungskriterium bestimmt. Für das Element DOE75 hingegen wurde die hohe numerische Apertur von $NA = 0,75$ festgelegt, dies entspricht einem Überlappfaktor von ca. 103. Die Pixelgröße von $0,4\ \mu\text{m}$ ist in diesem Fall etwas größer als der durch das Abbesche Auflösungskriterium gegebene Wert von $0,35\ \mu\text{m}$, damit die Periode ein ganzzahliges Vielfaches der Pixelgröße ist. Gleichzeitig gewährleistet diese Wahl, dass ein Pixel auch ein Vielfaches der Elektronenstrahl-Lithographie-Pixel von $5\ \text{nm}$ ist. Die Deckglas Parameter entsprechen einem relativ dünnen Standarddeckglas. Mit den gewählten Epsilon-Parametern wurde die Intensität im Fokus maximiert.

	DOE48	DOE75
Numerische Apertur	0,48	0,75
Periode	$44 \mu\text{m} \times 44 \mu\text{m}$	$44 \mu\text{m} \times 44 \mu\text{m}$
dx_{Pixel}	$0,55 \mu\text{m} \times 0,55 \mu\text{m}$	$0,4 \mu\text{m} \times 0,4 \mu\text{m}$
$N_{\text{Wiederholungen}}$	300×300	300×300
D_{Element}	$13,2 \text{ mm} \times 13,2 \text{ mm}$	$13,2 \text{ mm} \times 13,2 \text{ mm}$
λ_{Design}	$0,532 \mu\text{m}$	$0,532 \mu\text{m}$
n_{Deckglas}	1,5263	1,5263
d_{Deckglas}	$150 \mu\text{m}$	$150 \mu\text{m}$
ϵ	0,48	0,05
Ω	50	103

Tabelle 6.1: Design-Parameter der diffraktiven Linsenarrays DOE48 und DOE75.

Mit 300 Wiederholungen soll sichergestellt werden, dass in der Mitte des Elementes keine Störungen durch die Endlichkeit des Elements auftreten, da diese Anzahl an Wiederholungen deutlich größer als der Überlappfaktor ist.

Mit diesen Designparametern erhält man wie im Kapitel 4 beschrieben die binäre Phasenverteilung für die Grundperiode. In Abbildung 6.1 sind die Muster der Designphasen gezeigt.

Der Prozess der Herstellung der binären diffraktiven Elemente verlief folgendermaßen: Zuerst wurde eine Chromschicht auf das Glassubstrat gesputtert, anschließend wurde durch Spincoaten eine Resistschicht aufgetragen. Die Resistschicht wurde mittels Elektronenstrahl-Lithographie belichtet, hierbei wurden das gewünschte Phasenmuster aus Abbildung 6.1 verwendet, anschließend wurde der Resist entwickelt. Danach wurde die Struktur in Chrom geätzt und der restliche Resist entfernt. Zuletzt wurde das Quarzglas geätzt und die Reste der Chromschicht entfernt.

In Abbildung 6.2 sind Mikroskopbilder von je einer Elementarperiode der hergestellten Elemente gezeigt. Man sieht eine gute Übereinstimmung zwischen den Designmustern und den in Glas hergestellten Mustern.

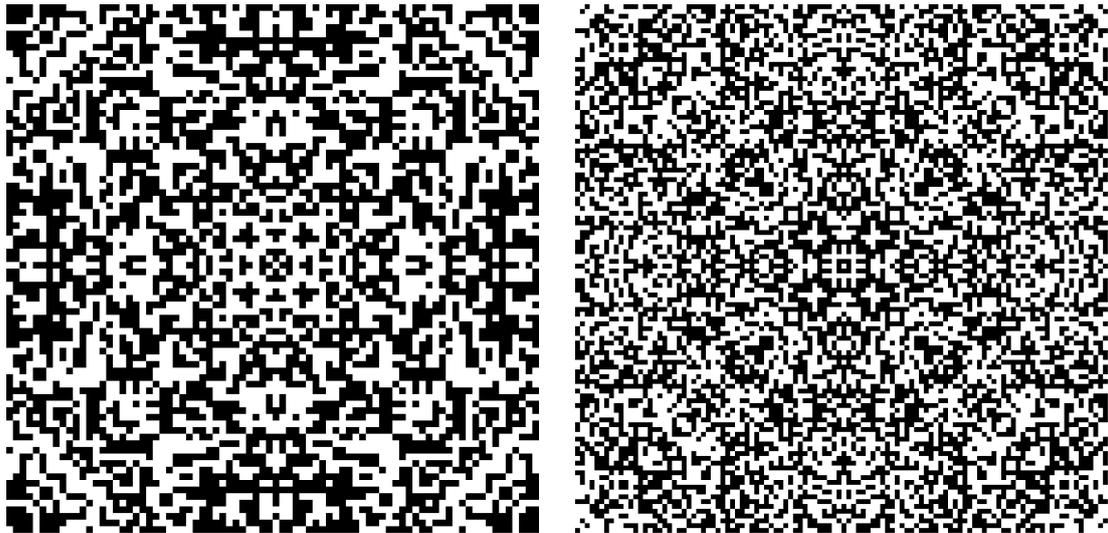


Abbildung 6.1: Binäre Muster der Designphasen. Links ist DOE48 und rechts ist DOE75 gezeigt.

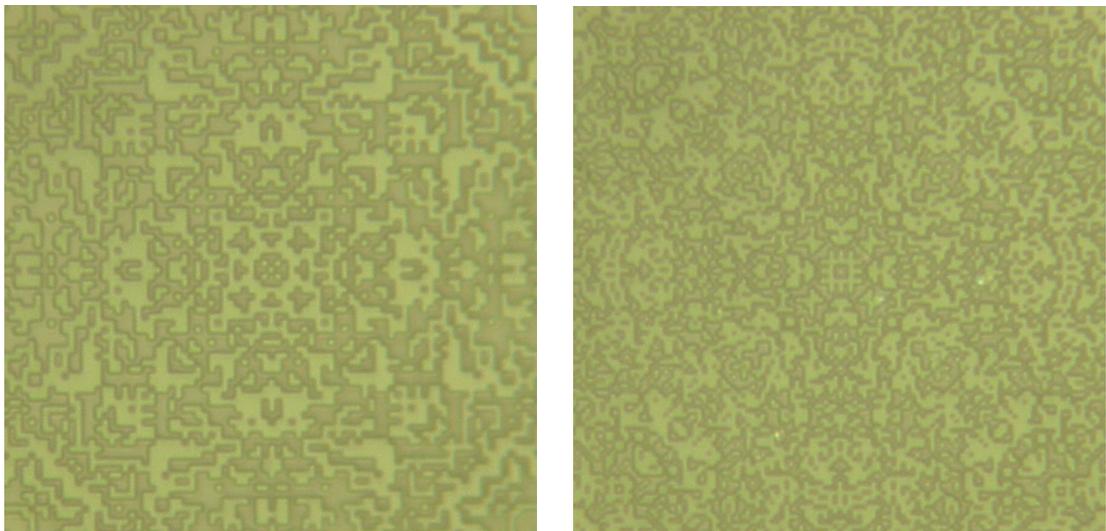


Abbildung 6.2: Mikroskopbilder einer Grundperiode der Größe $44\ \mu\text{m} \times 44\ \mu\text{m}$ der hergestellten diffraktiven Elemente. Links ist DOE48 und rechts ist DOE75 zu sehen.

6.2 Messung der Pupille

Aus der Messung der Pupille lässt sich die Phasenhöhe des diffraktiven Elements abschätzen. Hierbei nutzt man aus, dass bei binären diffraktiven Elementen die Stärke der nullten Beugungsordnung stark von der Phasenhöhe abhängt. Das Verhältnis der Stärken zweier anderen Beugungsordnungen, die nicht null sind, ist nicht von der Phasenhöhe abhängig [28]. Die Phasenhöhe bestimmt nur das Energieverhältnis zwischen der nullten und den anderen von null verschiedenen Beugungsordnungen [28]. Bei der Messung der Pupille lässt sich die relative Stärke der nullten Beugungsordnung zu einer anderen Beugungsordnung bestimmen. Dieses experimentell bestimmte Verhältnis lässt sich anschließend mit dem Verhältnis aus skalaren Simulationen vergleichen, bei denen die Phasenhöhe variiert wurde. Somit lässt sich auf die Phasenhöhe des experimentell vermessenen Elements schließen.

6.2.1 Experimenteller Aufbau zur Pupillenmessung

Der experimentelle Aufbau zur Messung der Pupille ist in Abbildung 6.3 gezeigt. Der Nd:YVO₄-Festkörperlaser ($\lambda = 532 \text{ nm}$) wird durch eine Kombination von Neutralfiltern abgeschwächt und beleuchtet das diffraktive Element. Hinter dem diffraktiven Element steht ein Achromat mit Brennweite 40 mm, der so positioniert ist, dass die Kamera im Fokus der Linse steht. Vor dem Achromat ist zusätzlich eine Irisblende angebracht, mit der man bei Bedarf den effektiven Durchmesser der Linse einschränken kann.

6.2.2 Durchführung der Messung

Der Aufbau bietet zwei Messstrategien an, um den dynamischen Umfang der Intensitäten in der Pupille zu messen. Entweder kann die Belichtungszeit oder die Stärke der Abschwächung verändert werden. Die Variation der Belichtungszeit erweist sich als vorteilhaft, da die optischen Filter einen relativ hohen Fehler in Bezug auf die Abschwächung haben.

Für die Messung der Pupille wurde die Linse abgeblendet, sodass die Beugungsordnungen auf der Kamera auf ca. 7 Pixel im Durchmesser verteilt sind. Anschließend wurde die Pupille mit unterschiedlicher Belichtungszeit aufgenommen und jeweils eine Dunkelmessung zu den Belichtungszeiten aufgenommen. Mit der Dunkelmessung lässt sich der Kamerauntergrund und systematische Fehler korrigieren. Unter der Annahme, dass sich die gemessene Kameraintensität linear zur Belichtungszeit verhält, lassen sich nun die gemessenen Energie in der (0,0)-ten Beugungsordnung und anderen Beugungsordnungen vergleichen.

6.2.3 Auswertung zur Bestimmung der Phasenhöhe

Für das diffraktive Element DOE48 und DOE75 wurde jeweils die Stärke verschiedener Beugungsordnungen in der Pupille vermessen. Hierbei wurde die Intensität

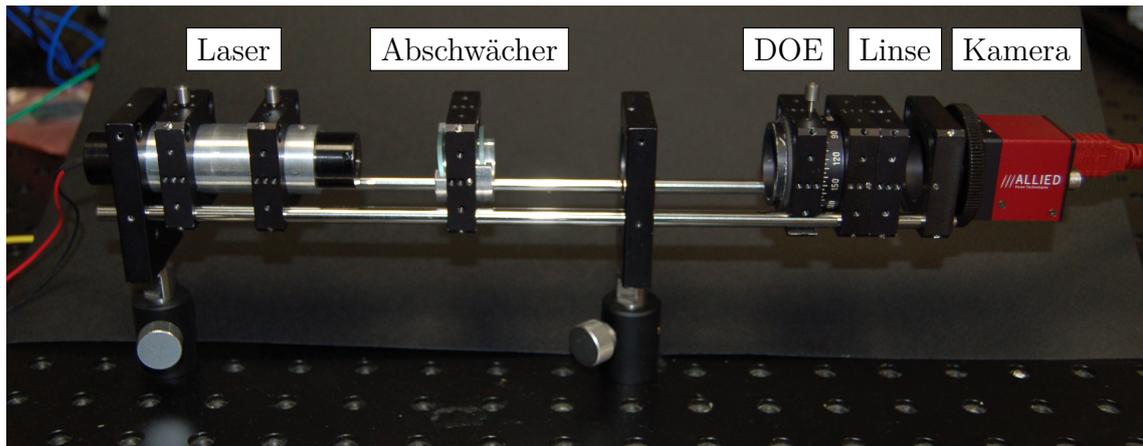


Abbildung 6.3: Aufbau zur Messung der Pupille von diffraktiven Linsen.

der einzelner Beugungsordnungen innerhalb eines Radius von 4 Pixeln um den Mittelpunkt integriert. Wie bereits in der Einleitung des Kapitels erwähnt, wird das Verhältnis zwischen der nullten und einer anderen Beugungsordnung gebildet und mit einer skalaren Simulation verglichen, um die Phasenhöhe abzuschätzen.

Bestimmung der Phasenhöhe für DOE48

Ausschnitte des Beugungsmusters, die verschiedene Beugungsordnungen in der Pupille zeigen und bei denen der Untergrund korrigiert ist, sind in Abbildung 6.4 zu sehen. Die Belichtungszeit bei der nullten Beugungsordnung war $531 \mu\text{s}$, bei der Aufnahme der zweiten Beugungsordnungen war die Belichtungszeit $50\,040 \mu\text{s}$. Die Energien in den einzelnen Beugungsordnungen erhält man, indem man die über die Intensität integriert anschließend mit der Quanteneffizienz der Kamera multipliziert und durch die Belichtungszeit teilt. Mittelt man zuerst die Energien der zweiten Beugungsordnung zu \bar{E}_2 und bildet danach das Energie Verhältnis zur nullten Beugungsordnung E_0 so erhält man $E_0/\bar{E}_2 = 112 \pm 12$. Hierbei wurde angenommen, dass die relativen Fehler auf E_0 und E_2 gleich groß sind.

Zum Vergleich lässt sich mit einer skalaren Simulation das Verhältnis von nullter und zweiter Beugungsordnung für unterschiedliche Phasenhöhen berechnen. Da sich in der Simulation die Beugungsordnungen nur über einen Pixel erstrecken entspricht das Intensitätsverhältnis dem Energieverhältnis aus der Messung. In Abbildung 6.5 ist das Ergebnis dieser Simulation aufgetragen. Für eine Phasenhöhe von $(1.14 \pm 0.01)\pi$ erhält man das experimentell ermittelte Verhältnis der Beugungsstärken. Somit weicht die Phasenhöhe um ca. 14 % von der spezifizierten Phasenhöhe von π ab.

Bestimmung der Phasenhöhe für DOE75

Analog zum vorherigen Abschnitt kann wird die Pupille von DOE75 untersucht. Hierbei wird nun das Verhältnis zwischen der (0,0)-ten Beugungsordnung und der

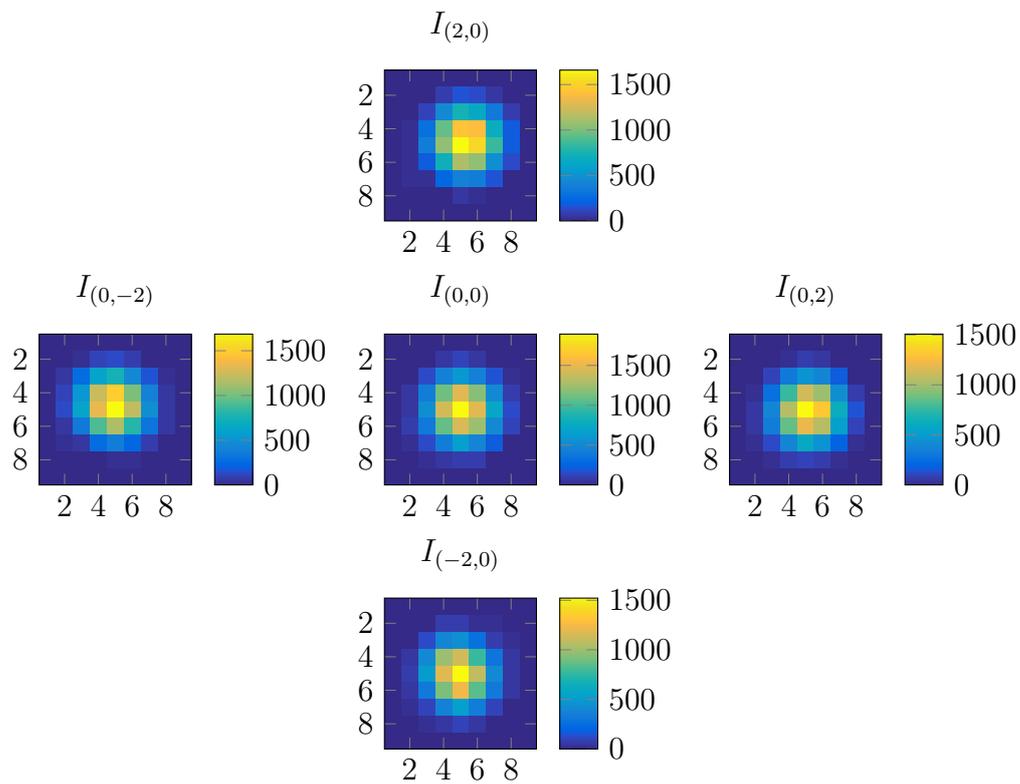


Abbildung 6.4: Ausschnitte aus zwei Bildern der Pupillemessung des Elements DOE48. Die Belichtungszeit bei der Aufnahme der (0,0)-ten Beugungsordnung ist ca. 100 mal länger als die Belichtungszeit bei der Aufnahme der zweiten Beugungsordnungen, somit werden ähnliche Aussteuerungen der Kamera erreicht.

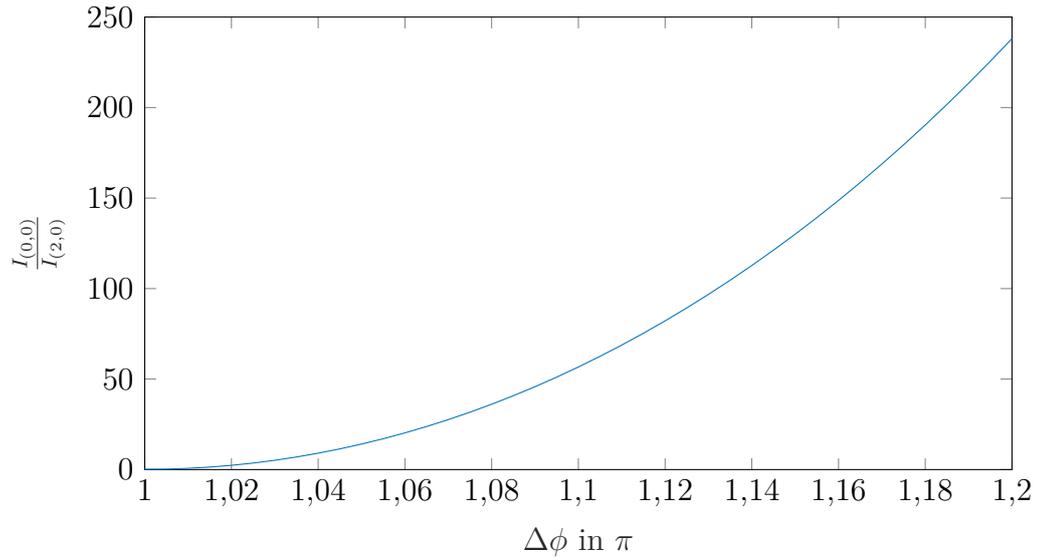


Abbildung 6.5: Simulation des Verhältnisses der Intensität in der (0,0)-ten Beugungsordnung und der (2,0)-ten Beugungsordnung als Funktion der Phasenhöhe.

diagonal dazu liegenden $(\pm 1, \pm 1)$ -ten Beugungsordnung benutzt. In Abbildung 6.6 sind die Intensitäten der Beugungsordnung abgebildet, der Untergrund wurde mit der Dunkelmessung korrigiert. Die Belichtungszeit bei der Messung der nullten Beugungsordnung war $531 \mu\text{s}$, bei der Aufnahme der ersten Beugungsordnungen war die Belichtungszeit $50\,040 \mu\text{s}$.

Bildet man zuerst den Mittelwert der Energie in der ersten Beugungsordnung und betrachtet dann das Verhältnis mit der Energie der nullten Beugungsordnung, so erhält man $E_0/\bar{E}_1 = 332 \pm 32$.

Die Ergebnisse der skalaren Simulation für das Intensitätsverhältnis der Beugungsordnungen bei unterschiedlichen Phasenhöhen sind in Abbildung 6.7 gezeigt. Vergleicht man das experimentell ermittelte Verhältnis mit den Verhältnissen der Beugungsstärken aus der skalaren Simulation, so erhält man bei einer Phasenhöhe von $(1.12 \pm 0.01)\pi$ eine gute Übereinstimmung. Somit weicht die Phasenhöhe des fabrizierten Elements um ca. 12 % von der spezifizierten idealen Phasenhöhe von π ab.

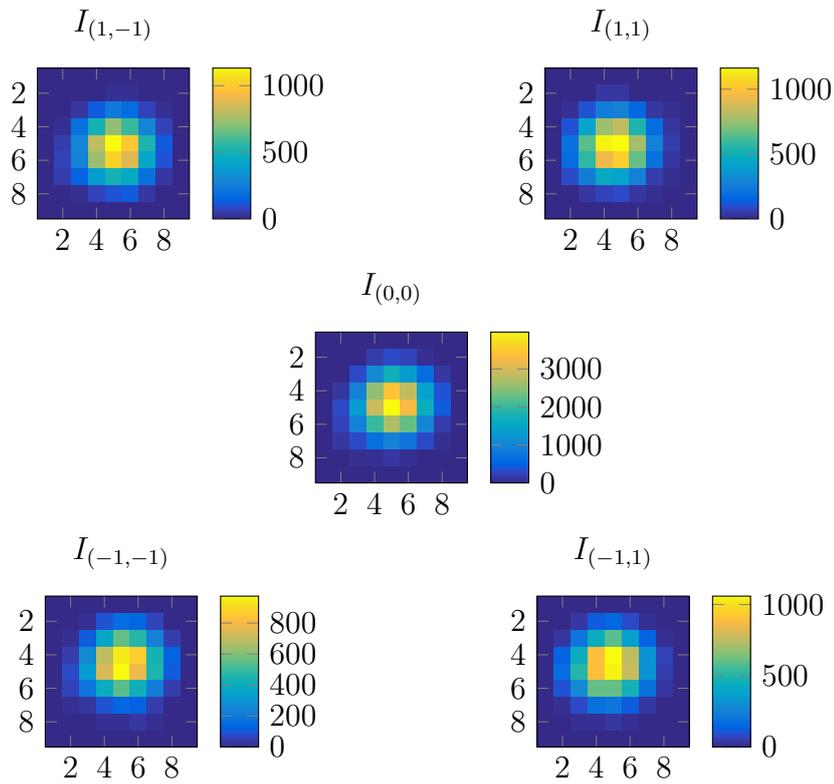


Abbildung 6.6: Ausschnitte aus zwei Bildern der Pupillemessung des Elements DOE75. Die Belichtungszeit bei der Aufnahme der (0,0)-ten Beugungsordnung ist ca. 100 mal länger als die Belichtungszeit bei der Aufnahme der ersten Beugungsordnungen, somit wird in beiden Fällen eine gute Aussteuerung der Kamera erreicht.

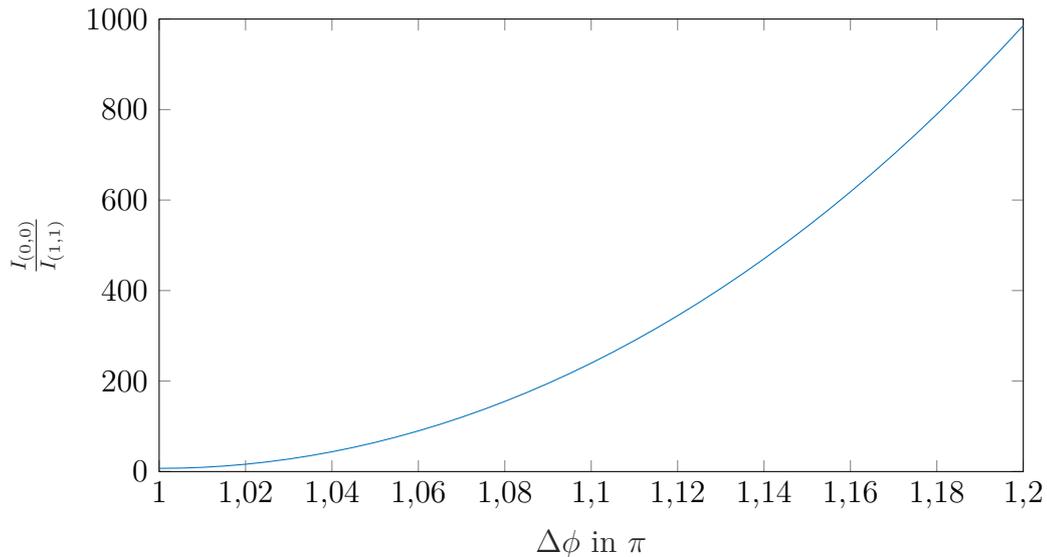


Abbildung 6.7: Simulation des Verhältnisses der Intensität in der (0,0)-ten Beugungsordnung und der (1,1)-ten Beugungsordnung als Funktion der Phasenhöhe.

6.3 Messung des Fokusverlaufs

Um die Funktion der diffraktiven Linsen mit überlappenden Aperturen direkt verifizieren zu können, werden die Foki untersucht.

6.3.1 Aufbau zur Vermessung des Fokus

In Abbildung 6.8 sieht man den Aufbau zur Messung des Fokalverlaufs der diffraktiven optischen Elemente. Hierbei wurde eine Abbildungsoptik bestehend aus einem Mikroskopobjektiv, Tubuslinse und einer Kamera auf einen xyz-Piezo-Tisch montiert. Das diffraktive Element liegt auf einem Objektträger, der von unten beleuchtet wird. Zur Beobachtung wird ein Mikroskopobjektiv mit 40-facher Vergrößerung und einer numerischen Apertur von 0,75 verwendet.

Da sowohl das diffraktive Element als auch das Mikroskopobjektiv eine Deckglaskorrektur integriert haben, sind auf dem diffraktiven Element zwei Deckgläser aufgelegt. Das erste Deckglas hat eine Dicke von 150 μm und ist für die Deckglaskorrektur des diffraktiven Elements, das zweite hat eine Dicke von 170 μm und ist für die Deckglaskorrektur des Mikroskopobjektivs. Dass die Deckgläser direkt auf dem diffraktiven Element aufliegen ist kein Problem, denn es lässt sich durch Betrachtung der Ebenenwellenzerlegung zeigen, dass die Positionen der Deckgläser keinen Einfluss auf den aufgenommenen Fokusverlauf haben.

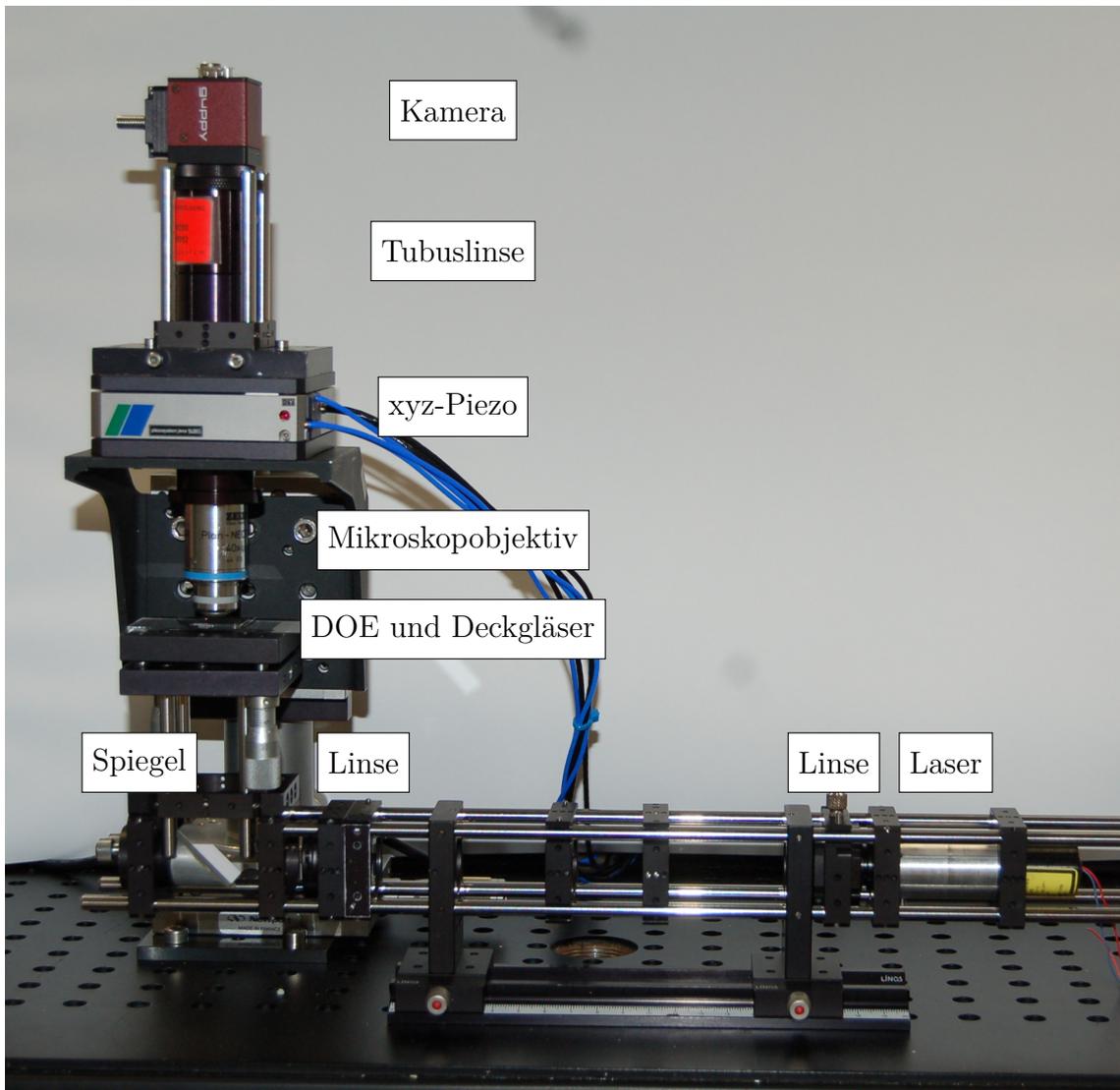


Abbildung 6.8: Aufbau zur Messung des Fokalverlaufs von diffraktiven Linsen.

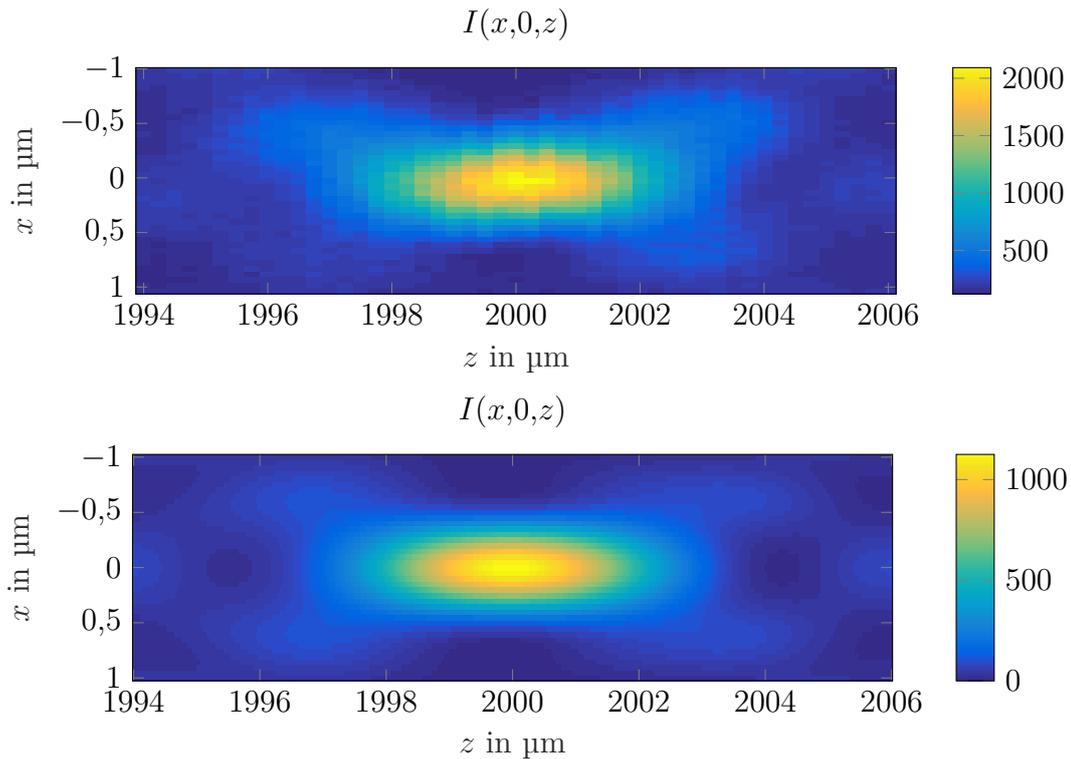


Abbildung 6.9: Der gemessene xz-Schnitt durch den Fokus des DOE48 Elements ist oben zu sehen, zum Vergleich ist unten die skalare Simulation des xz-Schnitts gezeigt. Die absoluten Skalierungen unterscheiden sich, da sich die Intensitätswerte der Messung auf die Aussteuerung der Kamera beziehen. In der Simulation beziehen sich die Intensitätswerte auf eine Beleuchtung mit einer ebenen Welle der Stärke eins.

6.3.2 Ergebnisse der Vermessung des Fokus

Mit Hilfe des Piezos lässt sich ein laterales Fokusprofil in unterschiedlichen z-Ebenen aufnehmen. Aus diesen xy-Schnitten lässt sich dann ein xz-Schnitt durch die Foki der diffraktiven Elementen erzeugen. Zum Vergleich wird jeweils die entsprechende skalare Simulation des Fokusprofils gezeigt, die die numerische Apertur des Mikroskopobjektivs berücksichtigt. In den Simulationen wurde die Grundperiode des diffraktiven Elements so abgetastet, dass elf Rechenpixel einem Elementpixel entsprechen.

Fokalprofil für DOE48

In Abbildung 6.9 sieht man einen gemessenen xz-Schnitt durch die Intensität des Fokus und zum Vergleich einen simulierten xz-Schnitt. Für den gemessenen xz-Schnitt wurden im Abstand von $0,25 \mu\text{m}$ 49 Bilder durch den Fokus aufgenommen. Die Lage der z-Achse wurde so angepasst, dass sie mit der Simulation übereinstimmt.

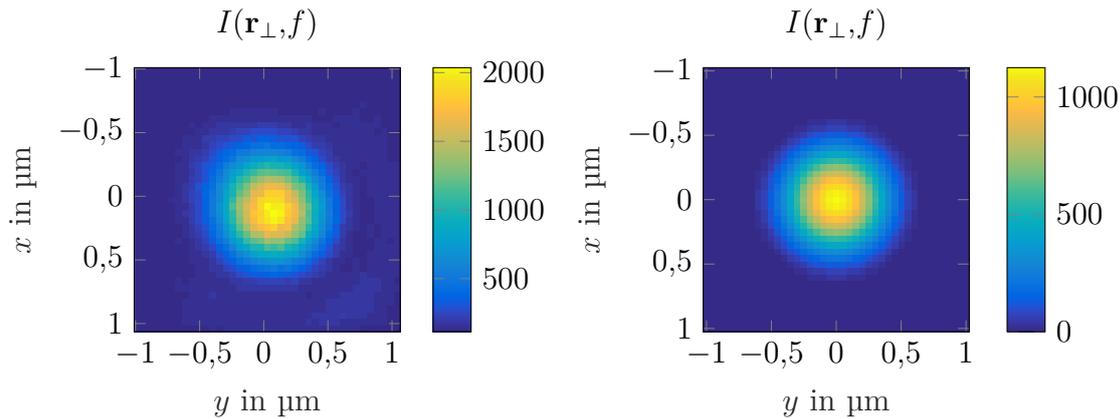


Abbildung 6.10: Das gemessene xy-Profil im Fokus des DOE48 Elements ist links zu sehen, zum Vergleich ist die Simulation des xy-Profiles rechts gezeigt. Die absoluten Skalierungen unterscheiden sich, da sich die Intensitätswerte der Messung auf die Aussteuerung der Kamera beziehen. In der Simulation beziehen sich die Intensitätswerte auf eine Beleuchtung mit einer ebenen Welle der Stärke eins.

Man sieht, dass sowohl die laterale als auch die axiale Länge des Fokus sehr gut mit der Simulation übereinstimmen. In der Messung lassen sich auf der optischen Achse auch das ± 1 . Minimum und ± 1 . Nebenmaximum identifizieren. Qualitativ stimmen die simulierten und gemessenen Fokalschnitte sehr gut überein.

Zum quantitativen Vergleich ist in Abbildung 6.10 das gemessene xy-Profil durch den Fokus, sowie die Simulation des Profils in der xy-Ebene des Fokus gezeigt. Die FWHM im Fokus beträgt in der Messung $(0,625 \pm 0,025) \mu\text{m}$ und in der Simulation $(0,606 \pm 0,025) \mu\text{m}$. Die Fehler auf die FWHM-Messung resultieren aus der Ablesegenauigkeit der jeweiligen Achse.

Für das diffraktive Linsenarray DOE48 stimmen die Simulation und die experimentellen Ergebnisse qualitativ sehr gut und quantitativ im Bereich ihrer Fehler überein.

Vergleicht man die experimentell ermittelte Breite jedoch mit der FWHM-Breite $570 \mu\text{m}$ der idealen kreisförmigen Pupille mit einer numerischen Apertur von 0,48 und einer Wellenlänge von $0,532 \mu\text{m}$ (vgl. Abschnitt 4.5), dann sieht man, dass sowohl die Simulation als auch die Messung ein größeres Ergebnis liefern. Diese Abweichung ist aber leicht durch die numerische Apertur des Beobachtungsobjektivs sowie der finiten Pixelgröße des diffraktiven Elements zu erklären. Die Simulation und die Messung stimmen überein, da in der Simulation die numerische Apertur des Objektivs durch eine Filterung des Spektrums und die finite Pixelgröße durch eine Überabtastung von mehreren Rechenpixeln pro Elementpixel berücksichtigt sind.

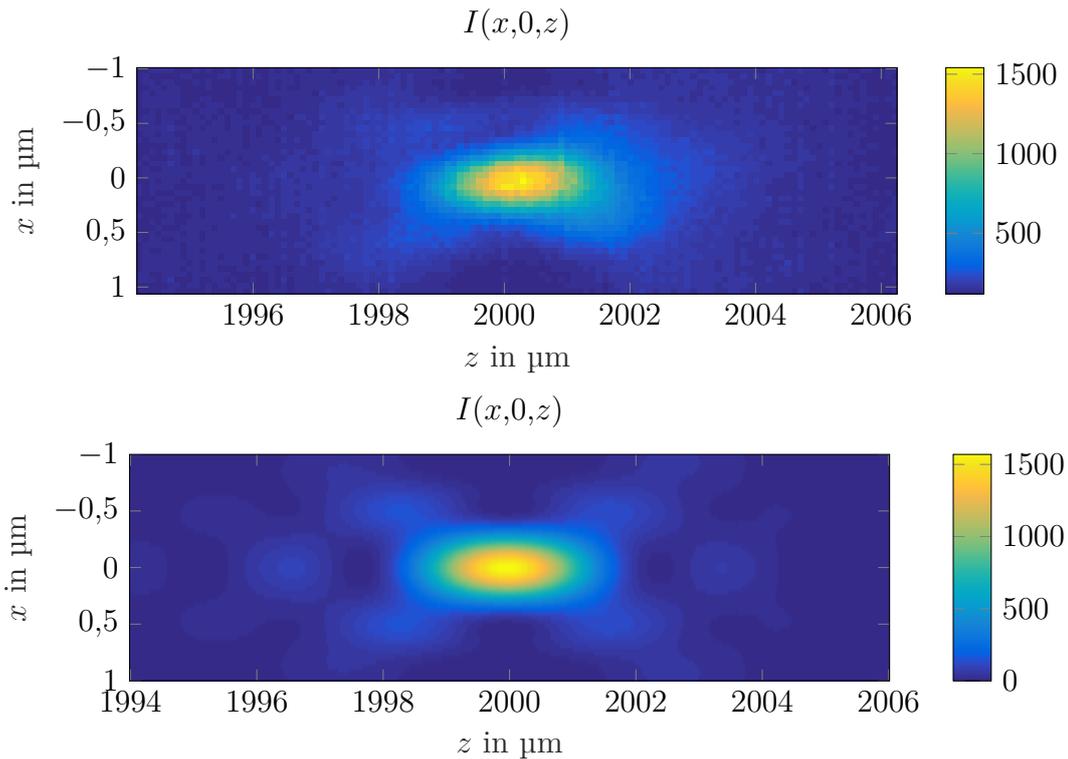


Abbildung 6.11: Der gemessene xz-Schnitt durch den Fokus des DOE75 Elements ist oben zu sehen, zum Vergleich ist unten die skalare Simulation des xz-Schnitts gezeigt. Die Intensitätswerte der Messung beziehen sich auf die Aussteuerung der Kamera. In der Simulation beziehen sich die Intensitätswerte auf eine Beleuchtung mit einer ebenen Welle der Stärke eins.

Fokalprofil für DOE75

In Abbildung 6.11 sieht man einen gemessenen xz-Schnitt durch die Intensität des Fokus und einen simulierten xz-Schnitt durch das Fokusprofil. Für den gemessenen xz-Schnitt wurde im Abstand von $0,10 \mu\text{m}$ 121 Bilder durch den Fokus aufgenommen. Die Lage der z-Achse wurde wieder so angepasst, dass sie mit der Simulation übereinstimmt.

Die laterale und axiale Ausdehnung des Fokus stimmt qualitativ mit der skalaren Simulation überein. Die Nebenmaxima und Nebenminima sind in der Messung jedoch nicht mehr erkennbar. Für den quantitativen Vergleich ist in Abbildung 6.12 das gemessene xy-Profil durch den Fokus, sowie die skalare Simulation des Fokusprofils in der xy-Ebene gezeigt. In der Fokusebene stimmen das gemessene und das simulierte xy-Profil qualitativ und quantitativ sehr gut überein. Die FWHM im Fokus beträgt in der Messung $(0,443 \pm 0,025) \mu\text{m}$ und in der Simulation $(0,450 \pm 0,020) \mu\text{m}$. Die Fehler auf die FWHM-Messung resultieren aus der Ablesegenauigkeit der jeweiligen Achsen.

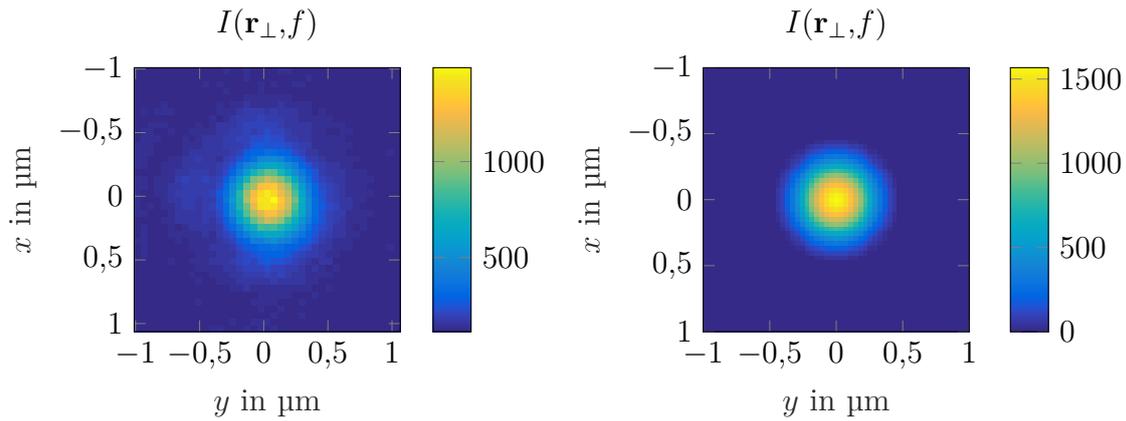


Abbildung 6.12: Das gemessene xy-Profil im Fokus des Elements DOE75 ist links zu sehen, zum Vergleich ist die skalare Simulation des xy-Profiles rechts gezeigt. Die Intensitätswerte der Messung beziehen sich auf die Aussteuerung der Kamera. In der Simulation beziehen sich die Intensitätswerte auf eine Beleuchtung mit einer ebenen Welle der Stärke eins.

Vergleicht man die experimentell ermittelte Breite mit der FWHM-Breite von $365 \mu\text{m}$ einer idealen kreisförmigen Pupille mit einer numerischen Apertur von $0,75$ und einer Wellenlänge von $0,532 \mu\text{m}$ (vgl. Abschnitt 4.5), dann sieht man, dass sowohl die Simulation als auch die Messung ein größeres Ergebnis liefern. Der Grund dieser Abweichung ist wieder durch die numerische Apertur des Beobachtungsobjektivs sowie der finiten Pixelgröße des diffraktiven Elements zu erklären.

Zusammenfassen lässt sich für das Element DOE75 sagen, dass die Fokusbreite gut mit der skalaren Simulation übereinstimmt. Die weitere Geometrie der Nebenmaxima stimmt bei dieser hohen numerischen Apertur jedoch nicht mehr überein. Ein möglicher Grund hierfür ist, dass die skalare Simulation des diffraktiven Elements als dünnes Element nicht mehr valide ist.

6.4 Fazit

In diesem Kapitel wurde die Funktionsweise von zwei diffraktiven Elementen experimentell verifiziert. Trotz einer Abweichung von 12 bzw. 14 % von der Designphasenhöhe erzeugen die diffraktiven Linsen mit überlappenden Aperturen Foki, deren laterale und axiale Ausdehnung der skalaren Simulation entsprechen. Dass diese Abweichung von der idealen Phasenhöhe nur einen geringen Einfluss hat, ist konsistent mit der Beobachtung, dass geringe Abweichungen in der Phasenhöhe nur eine geringe Verminderung der Effizienz des diffraktiven Elements bewirken (vgl. Gleichung 5.88).

Die Messung von beugungsbegrenzten Fokalprofilen ohne deutliche Aberrationen bestätigt die Funktion der integrierten Deckglaskorrektur der diffraktiven Linsen.

Im gemessenen xz -Schnitt des Fokus des diffraktiven Elements DOE48, das eine numerische Apertur von 0,48 hat, ist in den Nebenmaxima eine deutliche Übereinstimmung mit der skalaren Rechnung zu sehen.

Hingegen ist bei der Messung des xz -Profils der Foki des diffraktiven Elements DOE75, das eine numerische Apertur von 0,75 hat, keine gute Übereinstimmung in den Nebenmaxima im Vergleich zur skalaren Simulation festzustellen. Diese Abweichungen für hohe numerische Aperturen scheinen dadurch begründet zu sein, dass die dünne Element Näherung, die im skalaren Design verwendet wird, nur noch bedingt gültig ist. Insbesondere wird in Abschnitt 9.2 gezeigt, dass in der rigorosen Rechnung des Fokusprofils eines skalar designten diffraktiven Elements mit hoher numerischer Apertur die Nullstellen und Nebenmaxima nicht mehr so deutlich im axialen Profil ausgeprägt sind.

Allgemein kann festgehalten werden, dass trotz dieser Abweichungen für hohe numerische Aperturen das skalare Design der diffraktiven Linsen mit überlappenden Aperturen, auch für hohe numerische Aperturen, zu diffraktiven Elementen führt, die Foki mit beugungsbegrenzter Ausdehnung haben.

7 Anwendung: Parallel scannendes Mikroskop mit diffraktiven Linsen mit überlappenden Aperturen

Nachdem in den vorherigen Kapiteln das Design und die Eigenschaften von diffraktiven Linsen mit überlappenden Aperturen vorgestellt wurden, soll in diesem Kapitel das parallel scannende Mikroskop als Anwendung vorgestellt werden, das von der Verwendung diffraktiver Linsen mit überlappenden Aperturen profitiert.

In Abschnitt 7.1 wird diese Anwendung durch die Limitationen in der klassischen Mikroskopie motiviert. In Abschnitt 7.2 werden das Prinzip des parallel scannenden Mikroskops und die experimentelle Umsetzung vorgestellt. Im folgenden Abschnitt 7.3 werden die Theorie der Bildentstehung und die maximale Auflösung bei einem idealen Systemen untersucht. Anschließend wird die Bildaufnahme und Bildrekonstruktion anhand von einer Beispielmessung eingegangen (Kapitel 7.4 und 7.5). Danach wird experimentell die maximale Auflösung des realisierten parallel scannenden Mikroskops ermittelt (Kapitel 7.6).

7.1 Limitationen in der Mikroskopie

Die klassische Mikroskopie ist bezüglich der Kopplung von Auflösung, Sichtfeld und Arbeitsabstand limitiert. Abscannende Verfahren können die starre Kopplung zwischen Auflösung und Sichtfeld auf Kosten der Messzeit auflösen. Hierzu werden verschiedene Möglichkeiten zur Realisierung des Abscannens der Probe vorgestellt, anschließend wird auf Möglichkeiten zur Parallelisierung eingegangen. Der Einsatz von diffraktiven Linsen mit überlappenden Aperturen erlaubt zusätzlich die Verwendung von großen Arbeitsabständen. Im Folgenden wird detailliert auf diese Punkte eingegangen.

Kopplung zwischen Auflösung und Sichtfeld

Bei der lichtmikroskopischen Untersuchung von Proben gibt es klassischerweise eine Abwägung zwischen der Größe des Sichtfeldes und der Auflösung. Im einfachsten Fall wird durch die Auswahl des Mikroskopobjektivs die Auflösung und das Sichtfeld bestimmt. Hohe Auflösungen sind nur über einem kleinen Sichtfeld zu erreichen, wohingegen große Sichtfelder in der Regel eine geringere Auflösung haben. Gleichzeitig bedingt eine hohe Auflösung einen relativ geringen Arbeitsabstand zwischen Objektiv und Probe.

Bei der Untersuchung der Skalierungseigenschaften von Linsensystemen hat Lohmann [64] ein empirisches Gesetz gefunden, das diesen Zusammenhang beschreibt:

$$f \text{ [mm]} \cong F_N^3 \quad (7.1)$$

Hierbei ist f die Brennweite des Systems und F_N die Blendenzahl. Da die Auflösung invers proportional zur Blendenzahl und das Sichtfeld sowie der Arbeitsabstand proportional zur Brennweite sind, ergeben sich empirische Beziehungen zwischen Auflösung und Sichtfeld sowie Auflösung und Arbeitsabstand.

Systeme, die von diesem empirischen Gesetz abweichen, sind Lithographieobjektive und moderne astronomische Teleskope [11]. In diesen Speziaisystemen ist die Abweichung jedoch mit einem erheblichen finanziellen Aufwand verbunden.

Neben diesen klassischen Systemen mit refraktiver Optik existieren in der Literatur auch viele Mikroskopieverfahren ohne Linsen, bei denen durch die Verwendung von holographischen Methoden und einer strukturierten kohärenten Beleuchtung dieses empirische Gesetz umgangen wird [84, 107].

Entkopplung von Auflösung und Sichtfeld durch Scanning-Mikroskopie

Ein allgemeiner Lösungsansatz um dennoch großflächig eine Probe in hoher Auflösung untersuchen zu können, ist das Abrastern der Probe.

Die Verfahren, die die Probe abscannen, lassen sich in punktscannende und flächenscannende Verfahren aufteilen. Bei flächenscannenden Verfahren wird die Probe mit einem normalen Weitfeldmikroskop beobachtet und während der Beobachtung lateral verschoben. Anschließend werden die Einzelbilder zu einem großen Bild zusammengefügt [20]. Als Detektor wird eine Kamera verwendet. Bei punktscannenden Verfahren wird die Probe mit einem Fokus punktweise abgerastert. Aus diesen Punktmessungen wird ein Gesamtbild rekonstruiert [87]. Anstatt einer Kamera wird in der Regel ein Detektor mit nur einem Pixel benutzt.

In beiden Fällen muss ein hochaperturiges Objektiv benutzt werden um eine hohe Auflösung zu erreichen, sodass der Arbeitsabstand wieder gering ist. Das große Sichtfeld ist kein Resultat des optischen Systems, sondern wird durch eine Kombination mehrerer Einzelmessungen erreicht, sodass eine längere Messzeit benötigt wird.

Realisierung des Scannens

Flächenscannende System realisieren das Scannen in der Regel dadurch, dass die Probe mechanisch bewegt wird [20]. Dies kann für mechanisch instabile Proben zu Artefakten führen. Punktscannende Verfahren unterscheiden sich durch die Art, wie das Abrastern der Probe realisiert wird. Im Allgemeinen unterscheidet man zwischen Strahlscannenden und Objektscannenden Verfahren. Eine Auflistung mit Diskussion der Vor- und Nachteile der einzelnen Verfahren findet man bei Wilson und Sheppard [103].

Insbesondere für die Realisierung schneller Systeme ist es relevant, dass beim Scannen nur eine geringe Masse bewegt und dadurch beschleunigt wird. Von besonderer Bedeutung ist die Realisierung des Scannens durch laterale Bewegung des Mikrolinsenarrays oder der Probe, da solche Systeme sehr einfach zu realisieren sind. Der Nachteil ist eine relativ lange Scannzeit. Alternativ kann das Scannen durch eine verkippte Wellenfront, die auf eine fokussierende Optik trifft, realisiert werden, wodurch ein lateral von der optischen Achse verschobener Fokus erzeugt wird. Die Verkipfung der Wellenfront kann zum Beispiel durch Umlenkspiegel oder akustooptische Modulatoren realisiert werden.

Im Allgemeinen ist das Scannen durch eine verkippte Beleuchtungswelle schneller [104], jedoch sollte man für große Scanwinkel f - θ -Linsen zur Fokussierung benutzen, da sonst die Probe in unterschiedlichen z -Ebenen abgetastet wird [32].

Parallelisierung der Scanning-Mikroskopie

Im Allgemeinen lässt sich durch Parallelisierung die Messzeit in der Scanning-Mikroskopie verkürzen.

Die Parallelisierung durch einfache Replikation einer Weitfeld-Mikroskopieoptik ist durch ihre räumliche Ausdehnung beziehungsweise durch den maximalen Durchmesser der Optik begrenzt. Nur durch gleichzeitige Miniaturisierung lässt sich eine hohe Parallelisierung erreichen [88]. Problematisch ist, dass durch die Miniaturisierung erhöhte Anforderungen an die Herstellung der optischen Komponenten gestellt werden, da die Wellenlänge bei der Skalierung eine Invariante ist [64].

In punktwise scannenden Mikroskopieverfahren lässt sich eine Parallelisierung durch die Verwendung von Mikrolinsenarrays erreichen [80]. Insbesondere der Spezialfall der konfokalen Mikroskopie mit Mikrolinsen ist in der Literatur gut untersucht [95]. Anstatt eines Punktdetektors wird aufgrund der Parallelisierung nun ein Flächendetektor benutzt.

Einsatz von diffraktiven Linsen mit überlappenden Aperturen

Neben der Benutzung von refraktiven Linsen ist auch der Einsatz von diffraktiven Linsen mit überlappenden Aperturen zur Parallelisierung möglich. Dies wurde zuerst von Hulsken, Vossen und Stallinga [40] vorgestellt.

Durch den Überlapp der Aperturen lässt sich ein dichtes Beleuchtungsgitter erzeugen, wodurch eine sehr hohe Parallelisierung möglich ist. Insbesondere erlauben die diffraktiven Linsen hohe Parallelisierungsgrade, die mit refraktiver Optik nicht möglich sind. Ein weiterer Vorteil der Verwendung von diffraktiven Linsen mit überlappenden Aperturen ist, dass ein relativ großer Arbeitsabstand zwischen dem diffraktiven Element und der Probe möglich ist. Dieser Vorteil der diffraktiven Optik in der Scanning-Mikroskopie wurde auch schon von Cremer und Cremer [22] erkannt.

Das Abrastern der Probe lässt sich bei den diffraktiven Linsen sowohl über eine Verkipfung der Beleuchtungswellenfront oder durch eine mechanische Bewegung des Linsenarrays realisieren.

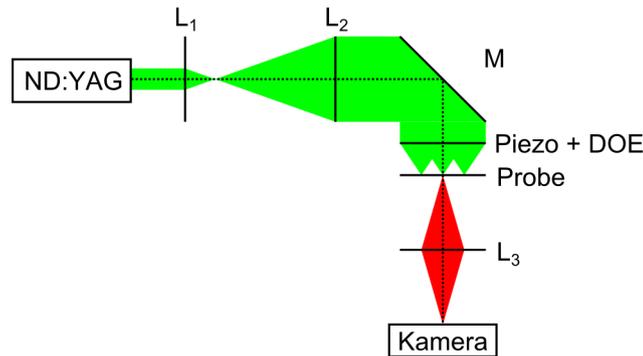


Abbildung 7.1: Schematische Zeichnung des Scanning Mikroskops. In grün ist der Strahlengang im Beleuchtungsarm und in rot ist der Strahlengang im Abbildungsarm gezeigt.

7.2 Parallel scannendes Mikroskop mit diffraktiven Linsen mit überlappenden Aperturen

In diesem Abschnitt wird der konzeptionelle Aufbau des parallel scannenden Mikroskops vorgestellt und dessen experimentelle Realisierung vorgestellt.

7.2.1 Funktioneller Aufbau des parallel scannenden Mikroskops

Der schematische Aufbau des parallel scannenden Mikroskops mit diffraktiven Mikrolinsen ist in Abbildung 7.1 gezeigt. Funktional lässt sich das parallele scannende Mikroskop in einen Teil zur Beleuchtung und einen Teil zur Beobachtung aufteilen.

Der Beleuchtungsteil besteht aus einer Kepler-Teleskopoptik mit zwei Linsen, die den Laserstrahl aufweitet um eine homogene Beleuchtungswelle für das diffraktive Element zu erzeugen. Das diffraktive Element bestehend aus den diffraktiven Linsen mit überlappenden Aperturen erzeugt eine punktweise Beleuchtung der Probe. Die Fassung des diffraktiven Elements mit einem Piezo-Tisch ermöglicht ein Abrastern der Probe in allen drei Raumdimensionen.

Der Beobachtungsteil besteht aus einer einfachen Abbildungsoptik, bestehend aus Kamera und Objektiv. Dieser Teil bildet die Foki auf die Kamera ab, somit lässt sich lokal die Transmission der Probe messen.

Durch die Verwendung hochaperturiger Foki zur Beleuchtung und einer niedera-perturigen Optik zur Beobachtung ist die Kombination einer hohen Auflösung und eines relativ großen Sichtfeldes gleichzeitig möglich.

7.2.2 Experimentelle Umsetzung des parallel scannenden Mikroskops

In Abbildung 7.2 sieht man die experimentelle Umsetzung des parallel scannenden Mikroskops.

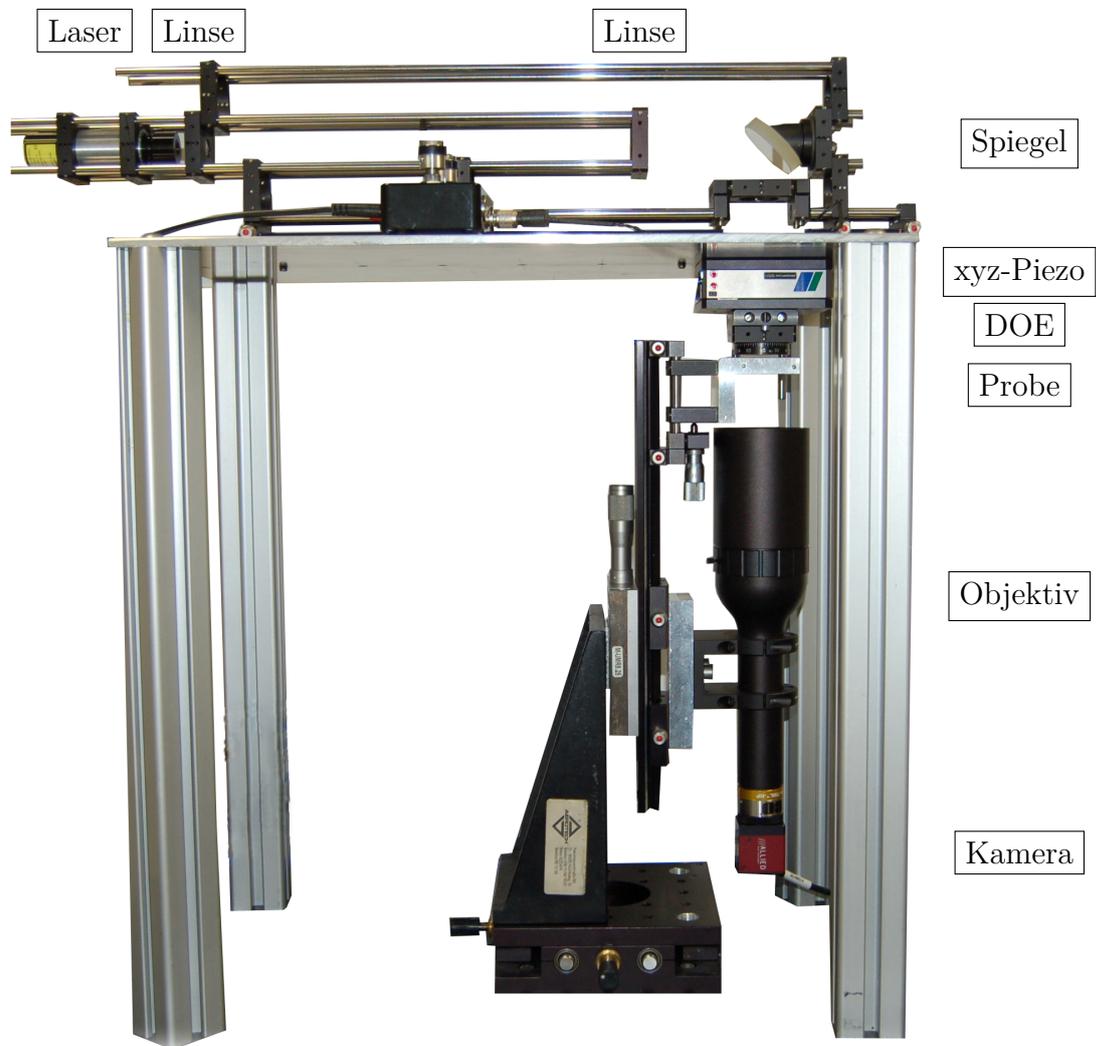


Abbildung 7.2: Experimenteller Aufbau des parallel scannenden Mikroskops.

Der Beleuchtungsarm des Mikroskops besteht aus einem waagrecht montiertem Laser und einer anschließenden Aufweitungsoptik, der eine ebene Welle zur Beleuchtung des diffraktiven Elements erzeugt. Diese ebene Welle wird dann über einen Spiegel umgelenkt und trifft senkrecht auf das diffraktive optische Element, das in einer Halterung an einem Piezo-Tisch befestigt ist.

Der Beobachtungsarm ist senkrecht montiert. Das Beobachtungsobjektiv und die Kamera bilden eine Einheit, die sich in xyz-Richtung verschieben lässt. Zusätzlich ist am Abbildungsarm noch ein Objektisch montiert, der es ermöglicht, den Abstand zwischen dem diffraktiven Element und der zu untersuchenden Probe unabhängig vom Abstand des Objektivs und des diffraktiven Elements einzustellen.

Im Folgenden sind die wichtigsten Komponenten und ihre charakteristischen Eigenschaften aufgelistet:

Lichtquelle Als Lichtquelle wird ein 1 mW Nd:YVO₄+KTP Festkörperlaser von Picotronic (DA532-1-3(20x95)) mit der Wellenlänge $\lambda = 532$ nm verwendet. Die Divergenz des Lasers ist auf 0,1 mrad spezifiziert.

Aufweitungsoptik Zur Strahlaufweitung wird ein Kepler-Teleskop aus zwei Achromaten mit $f_1 = 25$ mm, $D_1 = 12,5$ mm und $f_2 = 250$ mm, $D_2 = 31,5$ mm verwendet. Dies entspricht einer Strahlaufweitung um den Faktor 10 und einer Winkelverkleinerung um den Faktor $\frac{1}{10}$.

Umlenkspiegel Der Laserstrahl wird mit einem Planspiegel von Qioptik (Artikel-Nr. G340010400) mit Silber-Reflexcoating und einer Planarität von $\frac{1}{10}\lambda$ umgelenkt.

Piezo Zur Positionierung des diffraktiven Elements wird ein xyz-Positioniertisch von Piezosystems Jena (Tritor102CAP) mit integriertem Positionsmesssystem verwendet. Die Wiederholbarkeit ist mit ± 10 nm spezifiziert.

Beobachtungsobjektiv Das Objektiv ist von Edmund Optics (Techspec Silver Series Telecentric Lens # 58-432) und hat einen Vergrößerungsfaktor von 4, einen Arbeitsabstand von 44 mm, eine variable Blende und einen C-Mount-Anschluss.

Kamera Die Bilder werden mit einer CMOS-Kamera von Allied Vision (Guppy F503) mit 2592×1944 Pixeln und einer Pixelgröße von $2,2 \mu\text{m} \times 2,2 \mu\text{m}$ aufgenommen.

Im Vergleich zum Scanning-Mikroskop von Hulsken, Vossen und Stallinga [40] wird in dieser Umsetzung das Scannen durch eine laterale Bewegung des Mikrolinsenarrays realisiert, um ein äquidistantes Abtasten der Probe zu gewährleisten. Gleichzeitig bietet dies die Möglichkeit die Probe in verschiedenen z-Ebenen zu vermessen. Die gewählte Kombination von Beobachtungsobjektiv und Kamera erlaubt es im Vergleich eine ca. 8-mal größere Probenfläche zu untersuchen.

7.3 Theorie des parallel scannenden Mikroskops

Zuerst wird die Theorie der Bildentstehung im Scanning-Mikroskop in Abhängigkeit der Größe der konfokalen Blende vorgestellt. Anschließend wird daraus die theoretische Auflösung des parallel scannenden Mikroskops für die Spezialfälle einer Punktblende und einer sehr großen Blende hergeleitet.

7.3.1 Theorie der Bildentstehung

Die Abbildungseigenschaften des vorgestellten scannenden Mikroskops lassen sich mit der spatial kohärenten Abbildungstheorie bestimmen [69, 86, 103]. Das Abrastern der Probe lässt sich durch eine laterale Verschiebung der Probe um den Vektor \mathbf{r}_\perp^s beschreiben. Damit folgt für die komplexe Amplitude auf dem Detektor

$$U(\mathbf{r}_\perp, \mathbf{r}_\perp^s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [h_1(\mathbf{r}'_\perp) t(\mathbf{r}'_\perp - \mathbf{r}_\perp^s)] h_2(\mathbf{r}_\perp - \mathbf{r}'_\perp) d\mathbf{r}'_\perp \quad (7.2)$$

$$= [h_1(\mathbf{r}_\perp) t(\mathbf{r}_\perp - \mathbf{r}_\perp^s)] \otimes h_2(\mathbf{r}_\perp). \quad (7.3)$$

Hierbei beschreibt h_1 die Punktbildfunktion der Beleuchtungsoptik und h_2 ist die Punktbildfunktion der Abbildungsoptik, t ist die komplexe Transmissionsfunktion der Probe. Die Intensität auf dem Detektor lässt sich aus der komplexen Amplitude bestimmen:

$$I(\mathbf{r}_\perp, \mathbf{r}_\perp^s) = |[h_1(\mathbf{r}_\perp) t(\mathbf{r}_\perp - \mathbf{r}_\perp^s)] \otimes h_2(\mathbf{r}_\perp)|^2 \quad (7.4)$$

$$= \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\tilde{h}_1(\boldsymbol{\rho}_\perp) \otimes (\tilde{t}(\boldsymbol{\rho}_\perp) e^{-2\pi i \boldsymbol{\rho}_\perp \cdot \mathbf{r}_\perp^s})] \tilde{h}_2(\boldsymbol{\rho}_\perp) e^{+2\pi i \boldsymbol{\rho}_\perp \cdot \mathbf{r}_\perp} d\boldsymbol{\rho}_\perp \right|^2. \quad (7.5)$$

Einführen einer konfokalen Blende vor dem Detektor

Die Betrachtung der Intensität im Ursprung ist equivalent zu der Einführung einer idealen und damit infinitesimal kleinen konfokalen Blende. Für die Intensität im Ursprung erhält man

$$I(\mathbf{0}, \mathbf{r}_\perp^s) = \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\tilde{h}_1(\boldsymbol{\rho}_\perp) \otimes (\tilde{t}(\boldsymbol{\rho}_\perp) e^{-2\pi i \boldsymbol{\rho}_\perp \cdot \mathbf{r}_\perp^s})] \tilde{h}_2(\boldsymbol{\rho}_\perp) d\boldsymbol{\rho}_\perp \right|^2 \quad (7.6)$$

Das rekonstruierte Bild, das man erhält, ist somit

$$I_{\text{Rek}}(\mathbf{r}_\perp) = \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\tilde{h}_1(\boldsymbol{\rho}_\perp) \otimes (\tilde{t}(\boldsymbol{\rho}_\perp) e^{-2\pi i \boldsymbol{\rho}_\perp \cdot \mathbf{r}_\perp})] \tilde{h}_2(\boldsymbol{\rho}_\perp) d\boldsymbol{\rho}_\perp \right|^2. \quad (7.7)$$

Alternativ lässt sich der Einfluss der idealen konfokalen Blende auf die Amplitude betrachten, hierzu untersucht man komplexe Amplitude im Ursprung auf dem

Detektor:

$$U(\mathbf{0}, \mathbf{r}_\perp^s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [h_1(\mathbf{r}'_\perp) t(\mathbf{r}'_\perp - \mathbf{r}_\perp^s)] h_2(-\mathbf{r}'_\perp) d\mathbf{r}'_\perp \quad (7.8)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1(\mathbf{r}'_\perp) h_2(-\mathbf{r}'_\perp) t(\mathbf{r}'_\perp - \mathbf{r}_\perp^s) d\mathbf{r}'_\perp. \quad (7.9)$$

Die Faltung reduziert sich auf eine Multiplikation. Für gerade Funktionen $h_2(-\mathbf{r}'_\perp) = h_2(\mathbf{r}'_\perp)$ ergibt sich eine neue Faltung bezüglich der Verschiebungsordinate \mathbf{r}_\perp^s :

$$U(\mathbf{0}, \mathbf{r}_\perp^s) = [h_1(\mathbf{r}_\perp) h_2(\mathbf{r}_\perp)] \otimes t(\mathbf{r}_\perp). \quad (7.10)$$

Somit folgt für die Intensität im Ursprung und das rekonstruierte Bild:

$$I(\mathbf{0}, \mathbf{r}_\perp^s) = |[h_1(\mathbf{r}_\perp) h_2(\mathbf{r}_\perp)] \otimes t(\mathbf{r}_\perp)|^2 \quad (7.11)$$

$$I_{\text{Rek}}(\mathbf{r}_\perp) = |[h_1(\mathbf{r}_\perp) h_2(\mathbf{r}_\perp)] \otimes t(\mathbf{r}_\perp)|^2. \quad (7.12)$$

Man sieht, dass das abgescannte Bild für eine ideale konfokale Blende eine Faltung mit dem Produkt der Beleuchtungs- und Abbildungs-Punktbildfunktion mit der Transmissionsfunktion der Probe ist.

Spezialfall für niederaperturige Beobachtung Für den Fall, dass die numerische Apertur der Beobachtung viel kleiner ist als die numerische Apertur der Beleuchtung, lässt sich das Produkt der Punktbildfunktionen nähern:

$$h_1(\mathbf{r}_\perp) h_2(\mathbf{r}_\perp) \approx h_1(\mathbf{r}_\perp). \quad (7.13)$$

Somit ergibt sich für das rekonstruierte Bild bei der Verwendung einer idealen konfokalen Apertur

$$I_{\text{Rek}}(\mathbf{r}_\perp) \approx |h_1(\mathbf{r}_\perp) \otimes t(\mathbf{r}_\perp)|^2. \quad (7.14)$$

Dies entspricht der Intensität, die man bei einer homogenen kohärenten Beleuchtung und einer Abbildung mit einer Beobachtungsoptik erwartet. Dabei entspricht h_1 der Punktbildfunktion der Beobachtungsoptik.

Einfluss der finiten Größe der konfokalen Blende Für finite konfokale Blenden muss über die Blendenebene integriert werden, somit ist das rekonstruierte Bild gegeben durch:

$$I_{\text{Rek}}(\mathbf{r}_\perp^s) = \iint_{\text{Blende}} I(\mathbf{r}_\perp, \mathbf{r}_\perp^s) d\mathbf{r}_\perp. \quad (7.15)$$

Cox und Sheppard [21] haben herausgefunden, dass die finite Größe der konfokalen Blende der Grund ist, warum in der Praxis bei konfokalen Mikroskopen meinst die theoretische Auflösungsgrenze nicht erreicht wird.

Für den Grenzfall sehr große Blenden erhält man

$$I_{\text{Rek}}(\mathbf{r}_{\perp}^s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(\mathbf{r}_{\perp}, \mathbf{r}_{\perp}^s) d\mathbf{r}_{\perp} \quad (7.16)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |[h_1(\mathbf{r}_{\perp})t(\mathbf{r}_{\perp} - \mathbf{r}_{\perp}^s)] \otimes h_2(\mathbf{r}_{\perp})|^2 d\mathbf{r}_{\perp}. \quad (7.17)$$

Anschaulich wird somit über das gesamte transmittierte Licht integriert. Dies wird schon näherungsweise erreicht, wenn man über einen lokalisierten Peak auf dem Detektor integriert und die Intensität im Untergrund vernachlässigbar klein ist. Durch anwenden des Theorems von Parseval lässt sich das rekonstruierte Bild auch alternativ beschreiben als

$$I_{\text{Rek}}(\mathbf{r}_{\perp}^s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |[\tilde{h}_1(\boldsymbol{\rho}_{\perp}) \otimes (\tilde{t}(\boldsymbol{\rho}_{\perp})e^{-2\pi i \boldsymbol{\rho}_{\perp} \cdot \mathbf{r}_{\perp}^s})] \tilde{h}_2(\boldsymbol{\rho}_{\perp})|^2 d\boldsymbol{\rho}_{\perp}. \quad (7.18)$$

Vergleicht man dies mit Gleichung 7.7, so sieht man, dass durch die Integration über die gesamte Detektorebene die Reihenfolge zwischen Integration und Absolutquadratsbildung vertauscht wird.

7.3.2 Theoretische Auflösung des parallel scannenden Mikroskops

Die Theorie der Bildentstehung aus dem vorherigem Abschnitt lässt sich nutzen, um die theoretische Auflösung des parallel scannenden Mikroskops zu untersuchen. Hierzu wird zuerst die Auflösung des parallel scannenden Mikroskops mit einer idealen konfokalen Blende untersucht. Zum Vergleich wird auch der Grenzfall einer unendlich großen Blende betrachtet, bei dem über die Intensität in der Kameraebene integriert wird.

Theoretische Auflösung des parallel scannenden Mikroskops mit idealer konfokaler Blende

Um die Auflösung des parallel scannenden Mikroskops zu quantifizieren, betrachten wir ein in x-Richtung ausgerichtetes Kosinusgitter mit Periode P , dass um Δx verschoben ist.

$$U_{\text{Gitter}}(\mathbf{r}_{\perp}) = \frac{1}{2} + \frac{1}{2} \cos\left(2\pi \frac{x - \Delta x}{P}\right) \quad (7.19)$$

Im Frequenzraum lässt sich dieses Gitter durch drei diskrete Frequenzbeiträge darstellen.

$$\tilde{U}_{\text{Gitter}}(\boldsymbol{\rho}_{\perp}) = \frac{1}{2}\delta(\nu, \mu) + \frac{1}{4}\delta\left(\nu - \frac{1}{P}, \mu\right)e^{-2\pi i \nu \Delta x} + \frac{1}{4}\delta\left(\nu + \frac{1}{P}, \mu\right)e^{-2\pi i \nu \Delta x} \quad (7.20)$$

Die komplexe Amplitude im Ortsraum in der Gitterebene lässt sich als Multiplikation der Beleuchtungsamplitude und des Gitters beschreiben. Im Frequenzraum entspricht dies einer Faltung. Für einen idealen NA-begrenzten Fokus erhält man

$$\begin{aligned}\tilde{U}_{\text{Gitter+Bel}}(\boldsymbol{\rho}_{\perp}) &= \frac{1}{2} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) \\ &\quad + \frac{1}{4} \text{circ} \left(\frac{\sqrt{\left(\nu - \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) e^{-2\pi i \frac{\Delta x}{P}} \\ &\quad + \frac{1}{4} \text{circ} \left(\frac{\sqrt{\left(\nu + \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) e^{+2\pi i \frac{\Delta x}{P}}.\end{aligned}\quad (7.21)$$

Die Abbildung in die Sensorebene lässt sich idealisiert als multiplikativer Filter im Frequenzraum schreiben. Im Ursprung erhält man für die komplexe Amplitude:

$$\begin{aligned}U_{\text{Kamera}}(\mathbf{0}) &= \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\max(\text{NA}_{\text{Bel}}, \text{NA}_{\text{Beob}})} \right) d\nu d\mu \\ &\quad + \frac{1}{4} e^{-2\pi i \frac{\Delta x}{P}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right) \text{circ} \left(\frac{\sqrt{\left(\nu - \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) d\nu d\mu \\ &\quad + \frac{1}{4} e^{+2\pi i \frac{\Delta x}{P}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right) \text{circ} \left(\frac{\sqrt{\left(\nu + \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) d\nu d\mu\end{aligned}\quad (7.22)$$

Der erste Term beschreibt die Fläche der Überlappung zwischen zwei Kreisen, die den gleichen Mittelpunkt haben. Zur Verkürzung der Formeln wird für das Ergebnis dieses Überlappintegrals folgende Substitution eingeführt:

$$A = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\max(\text{NA}_{\text{Bel}}, \text{NA}_{\text{Beob}})} \right) d\nu d\mu \quad (7.23)$$

In den letzten zwei Integralen wird die Fläche der Überlappung zweier Kreise mit den Radien $\frac{\text{NA}_{\text{Beob}}}{\lambda}$ und $\frac{\text{NA}_{\text{Bel}}}{\lambda}$ bestimmt, deren Mittelpunkte $\frac{1}{P}$ von einander entfernt sind. Bis auf den Vorfaktor sind die zwei Terme gleich. Zur Verkürzung der Formeln wird auch hier für das Ergebnis dieses Überlappintegrals eine Substitution eingeführt:

$$\begin{aligned}B &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right) \text{circ} \left(\frac{\sqrt{\left(\nu - \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) d\nu d\mu \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right) \text{circ} \left(\frac{\sqrt{\left(\nu + \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) d\nu d\mu\end{aligned}\quad (7.24)$$

Streifenkontrast am Kosinusgitter

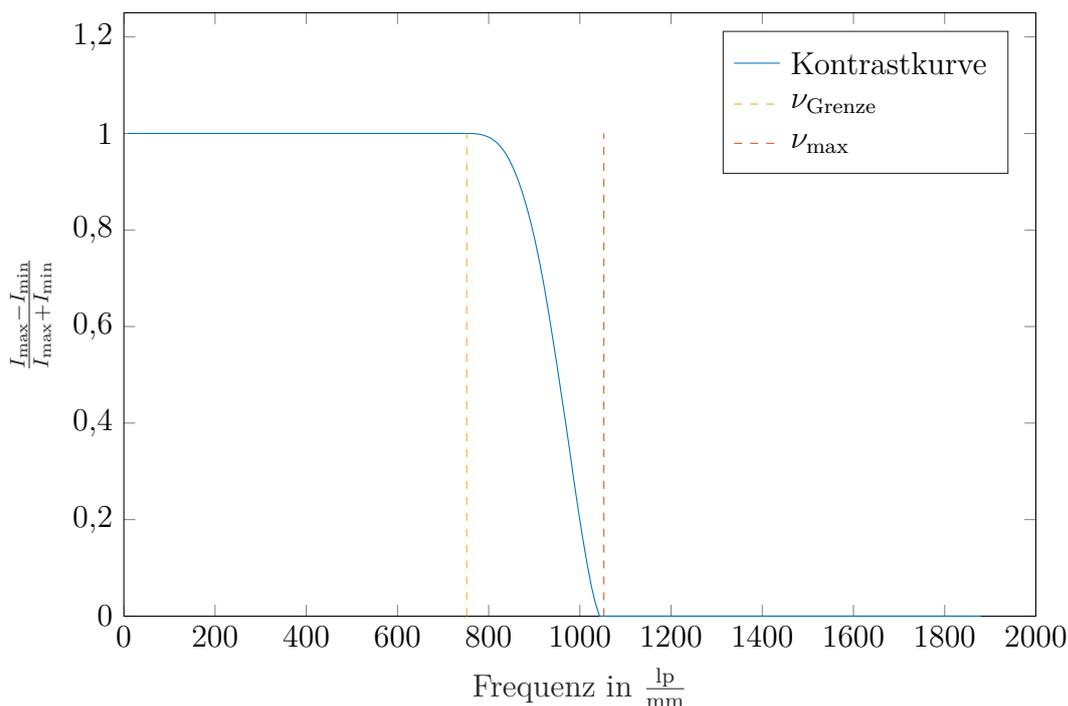


Abbildung 7.3: Streifenkontrast eines Kosinusgitters für eine Beleuchtungs-NA von 0,48 und eine Beobachtungs-NA von 0,08 bei der Wellenlänge $\lambda = 0,532 \mu\text{m}$.

Die Integration über die Fläche der Überlappung B lässt sich analytisch lösen, die Rechnung dazu ist in Anhang A.15 gezeigt. Die Intensität im Ursprung auf der Kamera lässt sich nun als Funktion der Überlappintegrale A und B schreiben:

$$I(0,0) = \left(\frac{1}{2} \left[A + B \cos \left(2\pi \frac{\Delta x}{P} \right) \right] \right)^2 \quad (7.25)$$

Die Intensität ist minimal bzw. maximal wenn der Kosinus die Werte -1 bzw. 1 annimmt. Hiermit lässt sich der Kontrast in Abhängigkeit der Überlappintegrale A und B schreiben:

$$K = \frac{I^{\max} - I^{\min}}{I^{\max} + I^{\min}} = 2 \frac{AB}{A^2 + B^2}. \quad (7.26)$$

Da das Überlappintegral B von der Periode P abhängt, ist der Kontrast auch Abhängig von der Periode des Kosinus. In Abbildung 7.3 ist ein Beispiel für den Streifenkontrast gezeigt. Bis zu einer Grenzfrequenz ν_{Grenze} ist der Streifenkontrast konstant. Bis zu dieser Grenze sind die Überlappintegrale A und B gleich groß, danach sinkt der Streifenkontrast.

$$\nu_{\text{Grenze}} = \left| \frac{\text{NA}_{\text{Bel}} - \text{NA}_{\text{Beob}}}{\lambda} \right| \quad (7.27)$$

Die maximale Frequenz ν_{\max} ist durch die mit der Wellenlänge gewichtete Summe der numerischen Aperturen der Beleuchtung und der Beobachtung gegeben.

$$\nu_{\max} = \frac{\text{NA}_{\text{Bel}} + \text{NA}_{\text{Beob}}}{\lambda} \quad (7.28)$$

Somit ist das kleinste aufzulösende Detail mit Hilfe des Nyquist-Shannon-Abtasttheorems gegeben durch

$$\delta x = \frac{1}{2\nu_{\max}} = \frac{\lambda}{2(\text{NA}_{\text{Bel}} + \text{NA}_{\text{Beob}})}. \quad (7.29)$$

Theoretische Auflösung des parallel scannenden Mikroskops mit idealer, unendlich großer Blende

Der Übergang von einer infinitesimalen Blende zu einer unendlich großen Blende lässt sich durch die Vertauschung der Integration mit der Absolutquadratsbildung in Gleichung 7.7 erreichen. In diesem Fall lässt sich wieder analytisch der Streifenkontrast eines Kosinusgitters in Abhängigkeiten von bestimmten Überlappintegralen angeben. Die maximale Frequenz lässt sich aus dem minimalen Kontrast bestimmen. Dies führt zu der gleichen maximalen Frequenz wie im Falle einer idealen konfokalen Blende.

Für die Intensität des rekonstruierten Bildes folgt explizit

$$\begin{aligned} I_{\text{Rek}}(\mathbf{r}_{\perp}^s) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \left[\tilde{h}_1(\boldsymbol{\rho}_{\perp}) \otimes (\tilde{t}(\boldsymbol{\rho}_{\perp}) e^{-2\pi i \boldsymbol{\rho}_{\perp} \cdot \mathbf{r}_{\perp}^s}) \right] \tilde{h}_2(\boldsymbol{\rho}_{\perp}) \right|^2 d\boldsymbol{\rho}_{\perp} \quad (7.30) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \frac{1}{2} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right) \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) \right. \\ &\quad \left. + \frac{1}{4} e^{-2\pi i \frac{\Delta x}{P}} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right) \text{circ} \left(\frac{\sqrt{\left(\nu - \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) \right. \\ &\quad \left. + \frac{1}{4} e^{+2\pi i \frac{\Delta x}{P}} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right) \text{circ} \left(\frac{\sqrt{\left(\nu + \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) \right|^2 d\boldsymbol{\rho}_{\perp}. \quad (7.31) \end{aligned}$$

Mit den Ersetzungen $A = \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right)$ und $B = \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right)$ für die Kreisfunktionen und $B^+ = \text{circ} \left(\frac{\sqrt{\left(\nu + \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right)$ und $B^- = \text{circ} \left(\frac{\sqrt{\left(\nu - \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right)$ für die verschobenen Kreisfunktionen vereinfacht sich dies zu

$$I_{\text{Rek}}(\mathbf{r}_{\perp}^s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \frac{1}{2} AB + \frac{1}{4} e^{-2\pi i \frac{\Delta x}{P}} AB^- + \frac{1}{4} e^{+2\pi i \frac{\Delta x}{P}} AB^+ \right|^2 d\boldsymbol{\rho}_{\perp}. \quad (7.32)$$

Im Ursprung erhält man bei $\Delta x = 0$ die maximale Intensität

$$I_{\text{Rek}}^{\max}(\mathbf{r}_{\perp}^s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \frac{1}{2}AB + \frac{1}{4}AB^- + \frac{1}{4}AB^+ \right|^2 d\boldsymbol{\rho}_{\perp} \quad (7.33)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{A^2}{4} \left(B^2 + BB^+ + BB^- - \frac{B^-B^-}{4} - \frac{B^+B^+}{4} + \frac{B^+B^-}{2} \right) d\boldsymbol{\rho}_{\perp} \quad (7.34)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{A^2}{4} \left(B^2 + 2BB^+ - \frac{B^+B^+}{2} + \frac{B^+B^-}{2} \right) d\boldsymbol{\rho}_{\perp}. \quad (7.35)$$

Hierbei wurde angenommen, dass das Integral über A^2BB^+ und A^2BB^- , sowie $A^2B^+B^+$ und $A^2B^-B^-$ identisch sind. Die minimale Intensität ist bei $\Delta x = \frac{P}{2}$

$$I_{\text{Rek}}^{\min}(\mathbf{r}_{\perp}^s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \frac{1}{2}AB - \frac{1}{4}AB^- - \frac{1}{4}AB^+ \right|^2 d\boldsymbol{\rho}_{\perp} \quad (7.36)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{A^2}{4} \left(B^2 - BB^+ - BB^- + \frac{B^-B^-}{4} + \frac{B^+B^+}{4} + \frac{1}{2}B^+B^- \right) d\boldsymbol{\rho}_{\perp} \quad (7.37)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{A^2}{4} \left(B^2 - 2BB^+ + \frac{B^+B^+}{2} + \frac{1}{2}B^+B^- \right) d\boldsymbol{\rho}_{\perp}. \quad (7.38)$$

Somit lässt sich der Kontrast schreiben als

$$K = \frac{I^{\max} - I^{\min}}{I^{\max} + I^{\min}} \quad (7.39)$$

$$= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{A^2}{4} (+4BB^+ - B^+B^+) d\boldsymbol{\rho}_{\perp}}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{A^2}{4} (2B^2 + B^+B^-) d\boldsymbol{\rho}_{\perp}}. \quad (7.40)$$

Der Kontrast wird Null, wenn die beiden Überlappintegrale $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A^2BB^+ d\boldsymbol{\rho}_{\perp}$ und $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{A^2B^+B^+}{4} d\boldsymbol{\rho}_{\perp}$ Null sind oder ihre Differenz Null wird. Da nur B^+ von der Periode abhängig ist, werden die beiden Überlappintegrale anschaulicher Weise Null, wenn $\frac{1}{P} > \frac{\text{NA}_{\text{Bel}} + \text{NA}_{\text{Beob}}}{\lambda}$ ist. Dies entspricht der gleichen maximalen Frequenz ν_{\max} wie im Fall einer infinitesimalen Blende.

$$\nu_{\max} = \frac{\text{NA}_{\text{Bel}} + \text{NA}_{\text{Beob}}}{\lambda} \quad (7.41)$$

Spezialfall einer sehr niederaperturigen Beobachtung

Für den Spezialfall, dass die numerische Apertur der Beobachtung viel kleiner ist als die numerische Apertur der Beleuchtung ergibt sich für die maximale Frequenz

$$\nu_{\max} \approx \frac{\text{NA}_{\text{Bel}}}{\lambda}. \quad (7.42)$$

Somit ist das kleinste aufzulösende Detail gegeben durch

$$\delta x = \frac{1}{2\nu_{\max}} = \frac{\lambda}{2\text{NA}_{\text{Bel}}}. \quad (7.43)$$

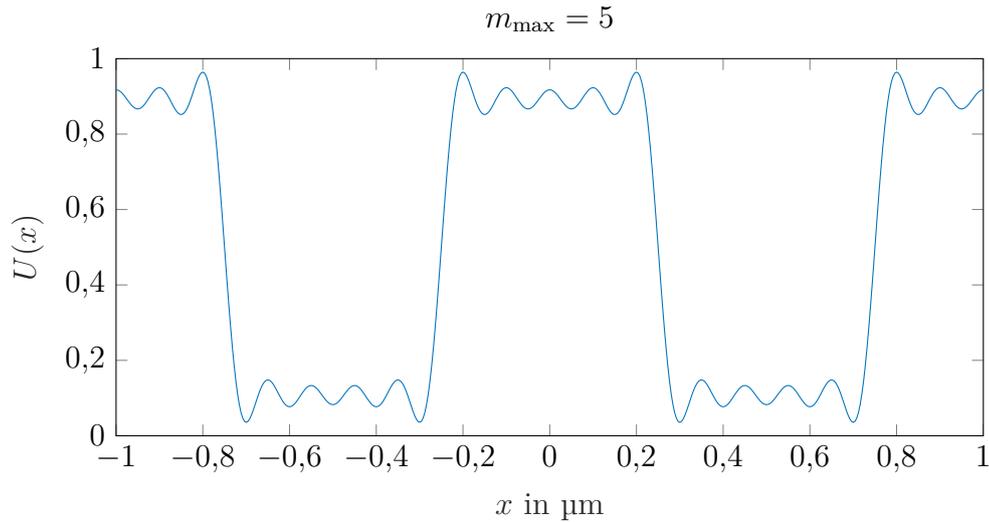


Abbildung 7.4: Plot des modenbeschränkten Gitters aus Gleichung 7.45 für $P = 1$ und $\Delta x = 0$.

Unterschiede zum realen endlichen Gitter

Das Kosinusgitter aus vorherigem Abschnitt ist ein ideales Gitter. Reale Gitter sind endlich und haben oft ein binäres Profil. Im Folgenden wird der Einfluss dieser zwei Effekte betrachtet.

Binäres Profil Näherungsweise lässt sich ein binäres Profil durch eine Überlagerung aus Kosinusfunktionen beschreiben:

$$U_{\text{Gitter}}^{\text{binär}}(x) \approx \frac{1}{2} + \frac{1}{2} \sum_{m=1}^{m_{\text{max}}} (-1)^{m-1} \frac{\cos\left(2\pi(2m-1)\frac{x-\Delta x}{P}\right)}{(2m-1)} \quad (7.44)$$

In Abbildung 7.4 ist die Amplitude für ein Gitter mit $m_{\text{max}} = 5$ gezeigt.

Die Fourier-Transformation dieses Gitters ist gegeben durch

$$\begin{aligned} \tilde{U}_{\text{Gitter}}^{\text{binär}} &= \frac{1}{2} \delta(\nu, \mu) \\ &+ \frac{1}{4} \delta(\mu) e^{-2\pi i \Delta x \nu} \sum_{m=1}^{m_{\text{max}}} \frac{(-1)^{m-1}}{2m-1} \left[\delta\left(\nu - \frac{2m-1}{P}\right) + \delta\left(\nu + \frac{2m-1}{P}\right) \right] \end{aligned} \quad (7.45)$$

Man sieht zu jeder Ordnung m zwei Delta-Funktionsbeiträge, die im Frequenzraum um $\delta\nu = \pm \frac{2m-1}{P}$ verschoben sind. Nach der Faltung mit dem Beleuchtungsspektrum und Multiplikation mit dem Beobachtungsfiler erhält man bei der Berechnung der komplexen Amplitude im Ursprung wieder Überlappintegrale. Für die m -te Ordnung

lässt sich der Beitrag des Überlappintegrals folgendermaßen definieren:

$$B_m \propto \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{circ}\left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}}\right) \text{circ}\left(\frac{\sqrt{\left(\nu + \frac{2m-1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}}\right) d\boldsymbol{\rho}_{\perp} \quad (7.46)$$

Die maximal beobachtbare Frequenz ist durch $\nu_{\text{max}} = \frac{\text{NA}_{\text{Bel}} + \text{NA}_{\text{Beob}}}{\lambda}$ gegeben, da die Überlappintegrale B_m für größere Frequenzen Null sind. Die maximal beobachtbare Frequenz an einem Kosinusgitter und einem binären Gitter unterscheiden sich nicht.

Endlichkeit des Gitters Die Auswirkungen der Endlichkeit des Gitters lässt sich durch eine räumliche Beschränkung mit einer Rechteckfunktion im Ortsraum modellieren.

$$U_{\text{Gitter endl.}}(x, y) = \left[\frac{1}{2} + \frac{1}{2} \cos\left(2\pi \frac{x - \Delta x}{P}\right) \right] \text{rect}\left(\frac{x - \Delta x}{n_P P}\right) \quad (7.47)$$

Das zugehörige Spektrum ist gegeben durch:

$$\begin{aligned} \tilde{U}_{\text{Gitter endl.}}(\nu, \mu) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\frac{1}{2} \delta(\nu', \mu') + \frac{1}{4} \delta\left(\nu' - \frac{1}{P}, \mu'\right) e^{-2\pi i \nu' \Delta x} + \frac{1}{4} \delta\left(\nu' + \frac{1}{P}, \mu'\right) e^{-2\pi i \nu' \Delta x} \right] \\ & n_P P e^{-2\pi i (\nu - \nu') \Delta x} \text{sinc}(n_P P (\nu - \nu')) d\nu' d\mu' \end{aligned} \quad (7.48)$$

$$\begin{aligned} &= \frac{1}{2} \delta(\mu) n_P P e^{-2\pi i \nu \Delta x} \text{sinc}(n_P P \nu) \\ &+ \frac{1}{4} \delta(\mu) e^{-2\pi i \frac{1}{P} \Delta x} n_P P e^{-2\pi i (\nu - \frac{1}{P}) \Delta x} \text{sinc}\left(n_P P \left(\nu - \frac{1}{P}\right)\right) \\ &+ \frac{1}{4} \delta(\mu) e^{+2\pi i \frac{1}{P} \Delta x} n_P P e^{-2\pi i (\nu + \frac{1}{P}) \Delta x} \text{sinc}\left(n_P P \left(\nu + \frac{1}{P}\right)\right) \end{aligned} \quad (7.49)$$

$$\begin{aligned} &= \frac{1}{2} \delta(\mu) n_P P e^{-2\pi i \nu \Delta x} \text{sinc}(n_P P \nu) \\ &+ \frac{1}{4} \delta(\mu) n_P P e^{-2\pi i \nu \Delta x} \text{sinc}(n_P (P\nu - 1)) \\ &+ \frac{1}{4} \delta(\mu) n_P P e^{-2\pi i \nu \Delta x} \text{sinc}(n_P (P\nu + 1)) \end{aligned} \quad (7.50)$$

Mit Hilfe dieses Spektrums und Gleichung 7.6 lässt sich wieder der Streifenkontrast zwischen der Intensität im Ursprung für $\Delta x = 0$ und $\Delta x = \frac{P}{2}$ berechnen. Die neue Kontrastfunktion K ist nun gegeben durch

$$K = \frac{I^{\Delta x=0}(0,0) - I^{\Delta x=\frac{P}{2}}(0,0)}{I^{\Delta x=0}(0,0) + I^{\Delta x=\frac{P}{2}}(0,0)}. \quad (7.51)$$

Für Gitter mit 5, 7 und 9 Perioden ist der Streifenkontrast in Abbildung 7.5 gezeigt. Bis zur Auflösungsgrenze des unendlichen Kosinusgitters verhalten sich die Kontrastfunktion wie die des unendlichen Gitters. Bei höheren Frequenzen kommt es zu einer Kontrastinvertierung.

Streifenkontrast am Kosinusgitter

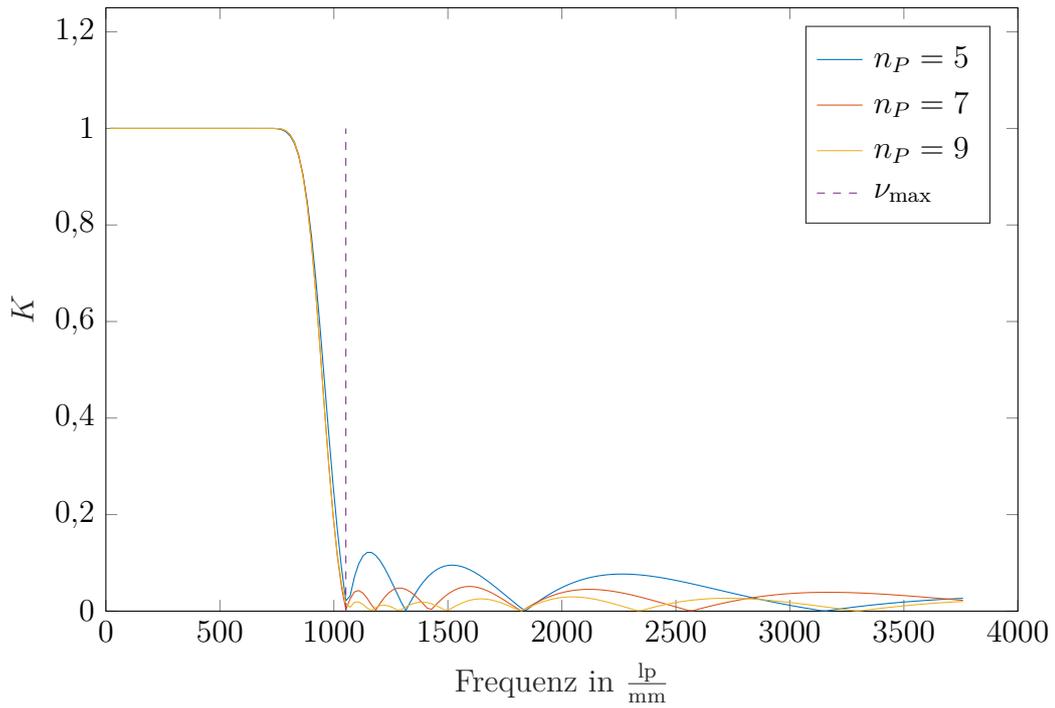


Abbildung 7.5: Streifenkontrast eines endlichen Kosinusgitters für eine Beleuchtungsoptik mit einer numerischen Apertur von 0,48 und einer Beobachtungsoptik mit numerischer Apertur von 0,08 bei der Wellenlänge $\lambda = 0,532 \mu\text{m}$. Die Gitterlänge beträgt 5, 7 oder 9 Perioden.

7.4 Aufnahme von Messdaten mit dem parallel scannendem Mikroskop

Im folgenden Abschnitt wird der Prozess der Aufnahme von Messdaten beschrieben. Bevor Messdaten mit dem parallel scannenden Mikroskop aufgenommen werden können, muss das Mikroskop auf die Probe justiert werden. Vor der Aufnahme lässt sich das Sampling der Grundperiode und die Größe des Sichtfelds auswählen. Für eine Aufnahme mit nicht eingeschränktem Sichtfeld wird die Dauer der Aufnahme gemessen. Anschließend werden Möglichkeiten zur Beschleunigungen dieser Aufnahme vorgestellt.

7.4.1 Ablauf der Messung

Justierung der Probe und der Abbildungsoptik

Bevor mit dem parallelen scannenden Mikroskop eine Probe gemessen werden kann, muss das Mikroskop für diese Probe justiert werden. Hierzu legt man die Probe mit Deckglas auf den Objektstisch, dann wird der Abstand zwischen der Kamera und Foki angepasst, damit die Foki auf die Kamera abgebildet werden. Anschließend lässt sich am Objektstisch die Höhe verstellen, um die gewünschte z-Ebene scharf zu stellen.

Festlegung der Abtastung der Probe

Um ein lückenloses Bild der Probe zu erhalten wird die Grundperiode des Gitters P_x bzw. P_y in N_x bzw. N_y gleich große Teile unterteilt. Durch laterale Bewegungen des Piezo um $dx_{\text{Piezo}} = \frac{P_x}{N_x}$ oder $dy_{\text{Piezo}} = \frac{P_y}{N_y}$ zwischen zwei Bildern wird die Probe vollständig abgetastet. Die Piezo-Schrittweite bestimmt die Pixelauflösung des rekonstruierten Bildes.

Um die Einstellzeit des Piezos möglichst gering zu halten ist es von Vorteil die Probe in Schlangenlinien abzurastern. Zwischen zwei Aufnahmen muss das diffraktive Element dann nur entlang einer Piezoachse bewegt werden. Die relative Bewegung zwischen den Aufnahmen ist außerdem sehr klein, sodass die Einschwingzeit des Piezos auf die neue Position auch relativ klein ist.

Festlegung des Sichtfelds

Durch die Einschränkung der „Region of Interest“ wird nur noch ein Ausschnitt des Kamerasensors zur Bildaufnahme benutzt. Dadurch lässt sich das Sichtfeld beschränken. Insbesondere bei räumlich begrenzten Proben lässt sich so die Menge der von der Kamera zu übertragenden Daten reduzieren.

Aufnahme der Messdaten

Pro Piezoposition wird bei der Messung ein Bild aufgenommen, an einen PC übertragen und gespeichert. Anschließend lässt sich aus diesen Daten das Bild eines z-Schnitts der Probe rekonstruieren.

Da im vorgestellten Mikroskopaufbau die Schärfentiefe des Abbildungssystems viel schlechter als die laterale Ausdehnung der einzelnen Foki ist, lässt sich die gewünschte z-Ebene mit dem Piezo auch noch nachträglich nach der initialen Justierung leicht anpassen. Insbesondere lassen sich so 3-D Volumina aufnehmen. Somit eignet sich das parallel scannende Mikroskop auch zur Tomographie dünner Schichten.

7.4.2 Dauer einer Messung

Die zeitliche Dauer einer Messung ist bei dem aufgebauten Testsystem hauptsächlich durch drei zeitintensive Schritte bestimmt, die für jedes Einzelbild durchzuführen sind. Die Anzahl der Einzelbilder ist durch das Sampling der Grundperiode vorgegeben.

Pro Einzelbild braucht der Piezo ca. 0,2s um das Linsenarray auf eine neuen Position, mit einer Genauigkeit von ≤ 50 nm, zu bewegen und auszuschwingen. Die Übertragung eines Bilds mit 12 Bit und voller Auflösung dauert ca. 0,23s, das Abspeichern eines Bildes ca. 0,23s. Insgesamt braucht das System ca. 0,65s pro Bild, wenn diese drei Aufgaben hintereinander abgearbeitet werden. Dies entspricht ca. 1,5 fps.

Durch geschicktes parallelisieren der zeitintensiven Schritte konnten experimentell Übertragungsraten von 2,8 fps realisiert werden. Hierbei wird zum Beispiel das Bild von der Kamera, in der Zeit auf den PC übertragen und gespeichert, in der sich der Piezo auf die gewünschte Position einstellt.

Zum Vergleich, die maximal mögliche Übertragungsrates der Kamera bei 12 Bit und voller Auflösung beträgt 4,33 fps und ist technisch durch die Bandbreite der FireWire-Anbindung begründet [37].

Um die Aufnahmegeschwindigkeit weiter zu erhöhen bieten sich folgende Ansätze an:

Eine Verkleinerung des Sichtfeldbereichs sowie eine Reduzierung der Bit-Tiefe ermöglichen eine schnellere Übertragung des Bildes von der Kamera zum Computer.

Eine direkte Verarbeitung des Bildes ist möglich, damit nicht mehr das ganze Bild sondern nur noch die für das Ergebnis relevante Information für das zu rekonstruierende Bild gespeichert wird.

Alternativ könnte ein spezieller Sensor, der nur die Intensität in einer bestimmten Umgebung um die einzelnen Foki aufnimmt, die Aufnahme beschleunigen, da dann nur die örtlich relevanten Informationen übertragen werden müssen.

7.5 Bildrekonstruktion aus Rohbildern

Um aus den Rohbildern ein Bild der Probe zu rekonstruieren wird zuerst die Geometrie und Lage des Beleuchtungsrasters bestimmt. Hierzu wird die von Brenner [12] beschriebene Grid-Fit-Methode benutzt. Anschließend ist es möglich die Transmission der einzelnen Beleuchtungsfoki zu evaluieren und daraus ein rekonstruiertes Bild zu erzeugen. Im Folgenden werden die einzelnen Schritte genauer am Beispiel einer Messung mit 80 x 80 Abtastpunkten pro Periode und Verwendung des Elements DOE48 beschrieben. Gleichzeitig wird gezeigt, dass das parallel scannende Mikroskop sich zur Untersuchung eines großen Sichtfeldes eignet.

7.5.1 Bestimmung des Beleuchtungsrasters

Der Grid-Fit Algorithmus benutzt einen Schwellwert um zu entscheiden, ob ein Pixel möglicherweise zu einem Fokus gehört. Anschließend werden benachbarte Pixel zu einem Lichtfleck zusammengefasst, von jedem dieser Lichtflecken wird der Schwerpunkt bestimmt. Aus drei Lichtflecken am Rand wird dann zuerst eine Schätzung der Gittervektoren und des Gitter-Ursprungs ermittelt und auf Plausibilität geprüft. Damit lassen sich die einzelnen Peaks gemäß ihrer Position sortieren. Anhand der sortierten Peaks kann man die Gitterparameter durch einen Least-Square-Fit bestimmen.

Für die spätere Auswertung bieten sich drei Strategien an:

1. Die Bestimmung des Beleuchtungsrasters folgt aus dem ersten Bild der Messung. Die Position des Beleuchtungsrasters bei anderen Piezoeinstellungen kann mit Hilfe der Piezo-Verschiebung berechnet werden.
2. Das Beleuchtungsraster wird für jedes Bild der Messung bestimmt. Anschließend wird mit Hilfe der Piezo-Verschiebung der Offset jedes einzelnen Bilds korrigiert. Durch Mittlung erhält man ein Gitter, das am Besten zu allen Messungen gleichzeitig passt. Falls die Bestimmung des Gitters bei einem Bild fehlschlägt wird dieses Bild nicht bei der Bestimmung des gemittelten Gitters verwendet. Die Position des Beleuchtungsrasters wird für jedes Bild nun mit diesem gemittelten Gitter und mit Hilfe der Piezo-Verschiebung in der Auswertung bestimmt.
3. Das Beleuchtungsraster wird für jedes Bild der Messung bestimmt. Anschließend wird bei der Rekonstruktion das individuell bestimmte Gitter zur Rekonstruktion verwendet.

Der Vorteil der ersten Strategie ist, dass die Rekonstruktion schon parallel zur laufenden Messung erfolgen kann. Im Gegensatz dazu kann bei der Verwendung der zweiten Strategie die Auswertung erst nach Beendigung der Aufnahme starten, wenn alle Bilder aufgenommen wurden. Die dritte Strategie ist die genaueste, jedoch kann es bei der Messung von tausenden von Bildern vorkommen, dass die automatische

Abweichung in Pixel

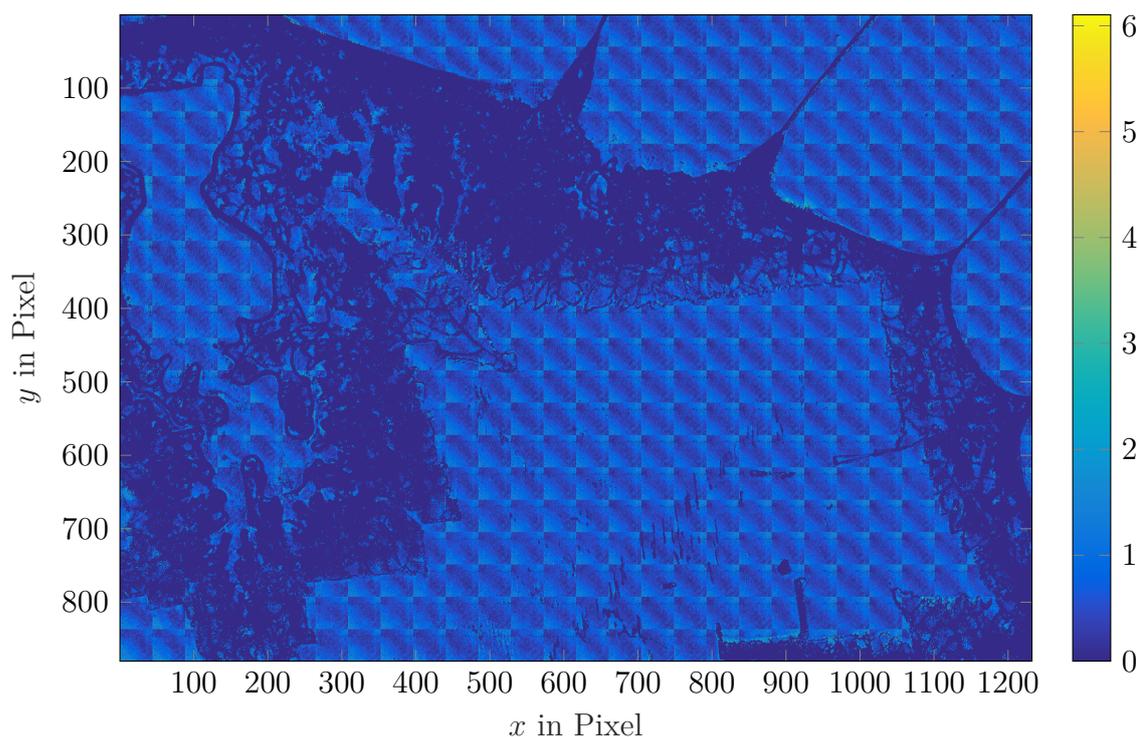


Abbildung 7.6: Bestimmung der Abweichung der einzelnen identifizierten Peaks gegenüber eines Rasters, das aus allen Einzelmessungen bestimmt wurde. Eine Abweichung von einem Pixel entspricht eine Abweichung von $0,55 \mu\text{m}$.

Gitterbestimmung fehlschlägt. Hier ist es ratsam Strategie eins oder zwei als Ersatzlösung parat zu haben.

Um den Fehler der zweiten Strategie abzuschätzen, ist in Abbildung 7.6 der Betrag der Abweichung für jeden detektierten Peak bei einer realen Messung gegenüber dem gemittelten Gitter aus allen Messungen gezeigt. Im Mittel ergibt sich eine Abweichung von $(0,71 \pm 0,41)$ Pixel pro detektiertem Peak. In Bereichen mit hoher Absorption wurden aufgrund der Schwellwertbedingung keine Peaks detektiert. Im Allgemeinen sind die größten Abweichungen (größer 2 Pixel) an den Rändern der absorbierenden Strukturen zu finden. Die Absorption stört hier den Fokus und verändert somit auch den Schwerpunkt des Fokus.

In den Bereichen, in denen die Probe homogen und nicht absorbierend ist, sieht man, dass die Abweichungen am größten sind, wenn die beiden Piezo-Achsen unterschiedlich weit gefahren sind.

Insbesondere lässt sich die Abweichung der Peaks des ersten Bilds von dem gemittelten Grid, das aus allen Bildern bestimmt wurde, untersuchen. Es ergibt sich eine durchschnittliche Abweichung von $(0,29 \pm 0,20)$ Pixel pro detektiertem Peak des Gitters.

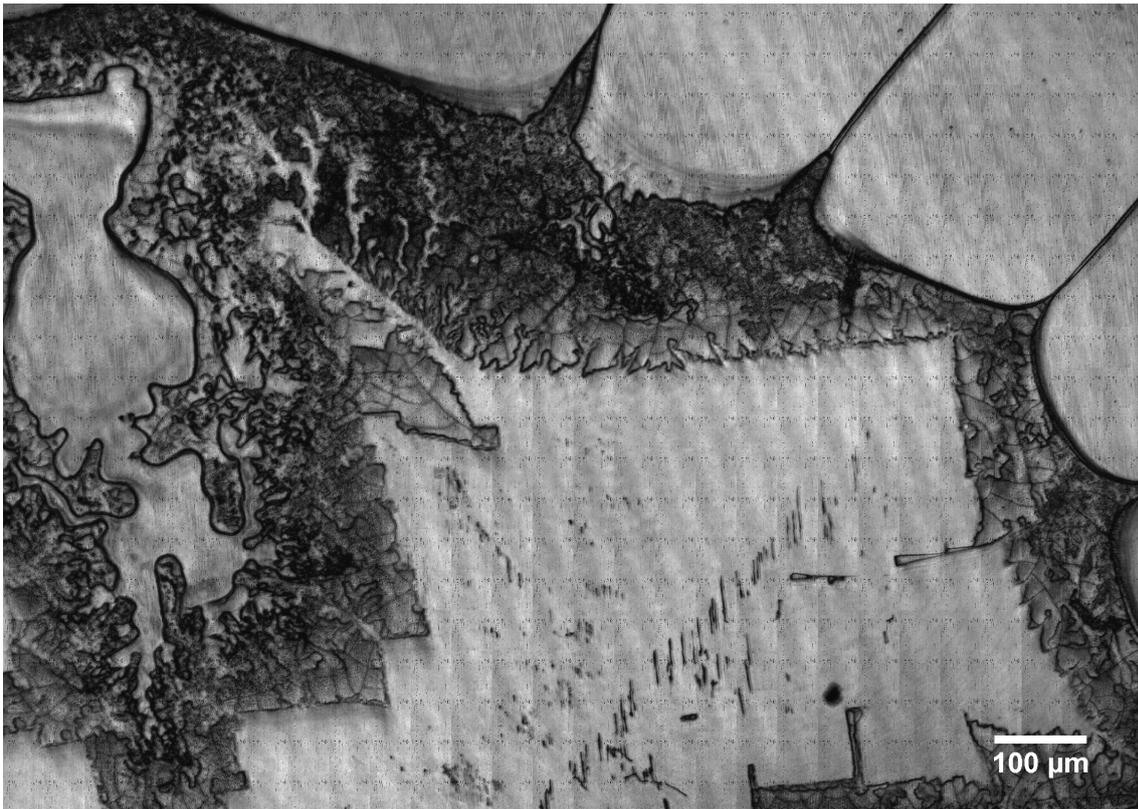


Abbildung 7.7: Rekonstruktion eines Testobjekts mit individueller Gitterbestimmung auf jedem Bild. In die Rekonstruktion wurde die Intensität den jeweiligen Gitterpunkten eingetragen. Für die Gitterpunkte, die nicht auf dem Pixelraster liegen, wurde die Intensität linear interpoliert.

7.5.2 Auswertung der Transmission des Beleuchtungsgitters

Zur Rekonstruktion wird die transmittierte Intensität an den Beleuchtungsgitterpunkten ermittelt. Hierzu wird die Intensität über eine bestimmte Pixelumgebung um einen Gitterpunkt integriert.

Zuerst werden Rekonstruktionen gezeigt, bei der das Auswertungsgitter für die Peakposition für jedes Bild einzeln bestimmt wurde, dies entspricht der dritten Strategie aus dem vorherigen Kapitel. Anschließend werden Rekonstruktionen gezeigt, bei der die Auswertung auf einem Gitter beruht, das aus allen Bildern erzeugt wurde.

Rekonstruktion mit individuellem Auswertungsgitter

In Abbildung 7.7 sieht man das rekonstruierte Bild für ein individuelles Auswertungsgitter pro aufgenommenen Bild. Die Intensität wurde hier nur auf dem Gitterpunkt ausgewertet und in das Ergebnisbild abgespeichert. Für Gitterpunkte, die

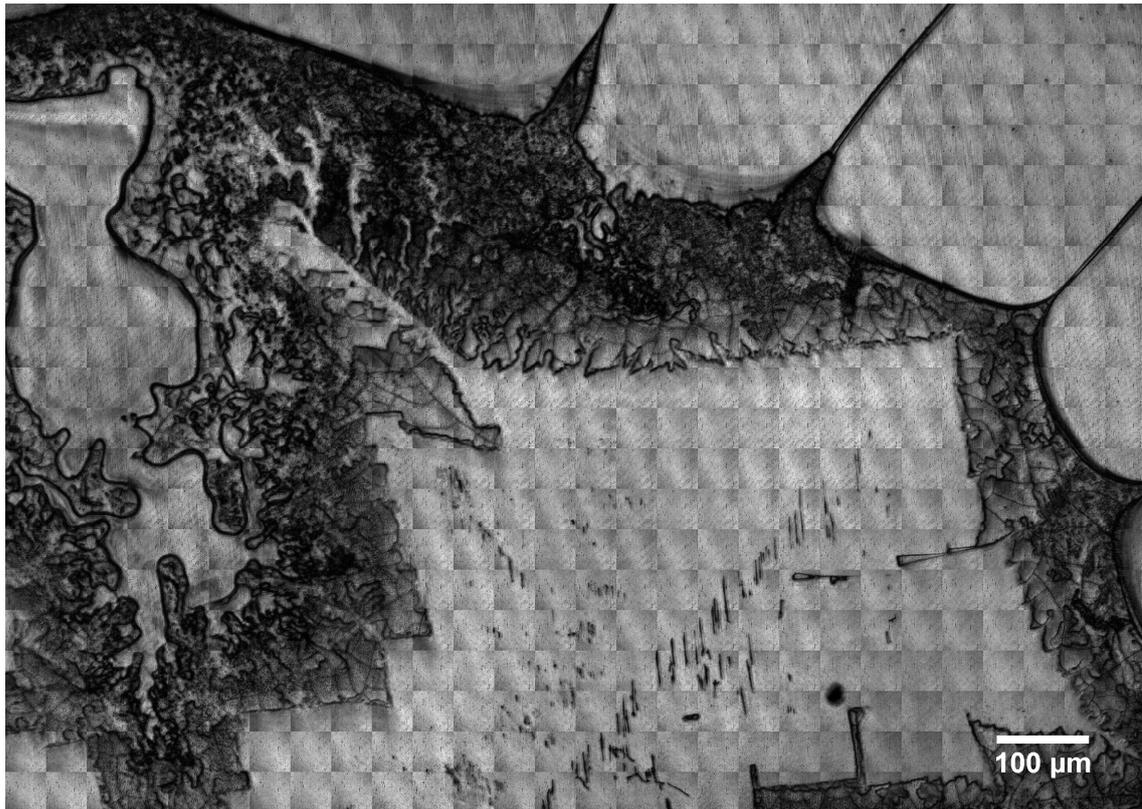


Abbildung 7.8: Rekonstruktion eines Testobjekts. Für die Rekonstruktion wurde die Intensität an den jeweiligen Gitterpunkten eingetragen. Das Gitter zur Rekonstruktion wurde aus allen Aufnahmen bestimmt.

nicht auf dem Pixelraster liegen, wurde die Intensität bilinear interpoliert.

Pro Periode sieht man, dass immer die gleichen Pixel in einem regelmäßigen Gitter schwarz sind. Das kommt daher, dass hier die Gitterbestimmung fehlschlägt. Bei der Betrachtung der Stetigkeit über die Periodengrenze fallen Unstetigkeiten durch laterale Verschiebung auf, diese sind durch die Abweichung der Piezoposition erklärbar oder über eine Verdrehung zwischen den Piezo-Achsen und den Achsen des diffraktiven Elements.

Rekonstruktion mit gemitteltem Auswertungsgitter

Im Folgenden werden Rekonstruktionen für verschieden große Pixelumgebungen gezeigt. In Abbildung 7.8 wurde nur die Intensität an den Gitterpunkten in das Rekonstruktionsbild gespeichert, dies kann man als Integration über eine Pixelumgebung von null Pixeln auffassen. Die Intensitätswerte von Gitterpunkten, die zwischen dem Pixelraster liegen, wurden durch bilineare Interpolation bestimmt.

Man sieht deutlich zwei Bildartefakte in der Rekonstruktion. In jeder Periode fällt die Intensität von der Mitte der Periode zu zwei gegenüberliegenden Ecken ab. Das Muster des Intensitätsabfalls pro Periode stimmt mit dem Muster der Abwei-

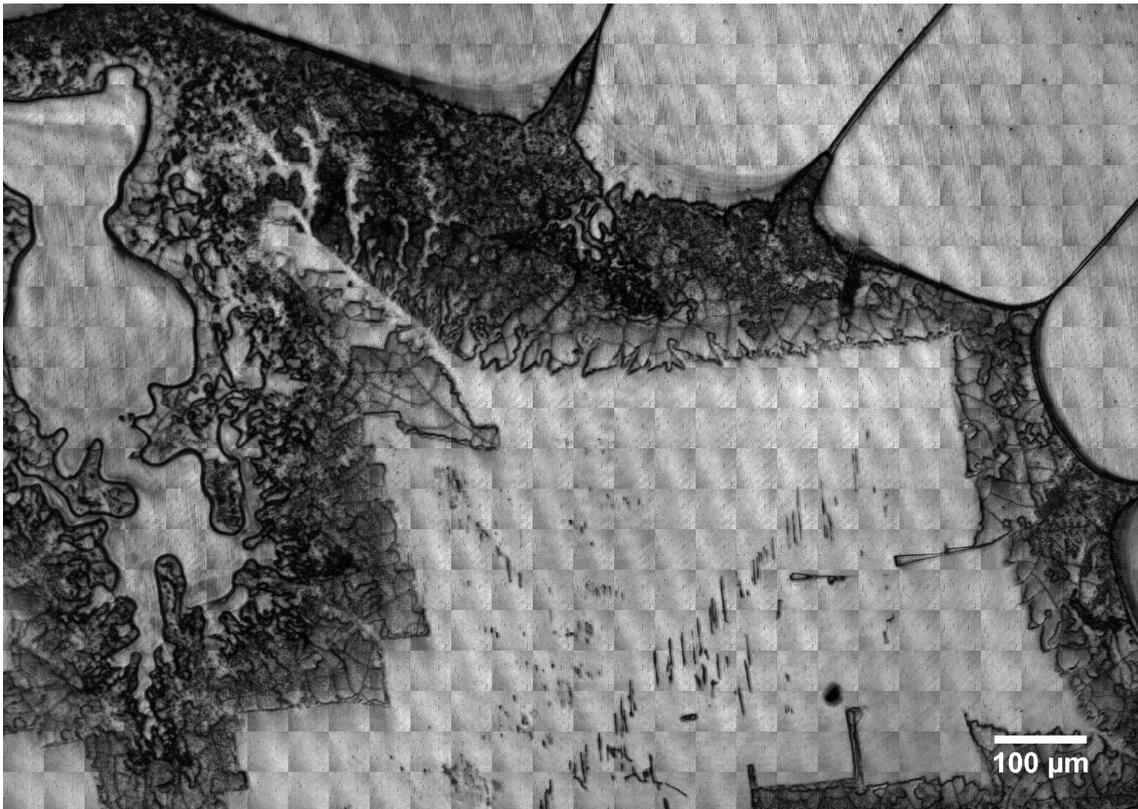


Abbildung 7.9: Rekonstruktion eines Testobjekts. Für die Rekonstruktion wurde die Intensität an den jeweiligen Gitterpunkten in einer 3×3 Pixel großen Umgebung aufsummiert und in das Ergebnisbild eingetragen. Das Gitter zur Rekonstruktion wurde aus allen Aufnahmen bestimmt.

chung aus 7.6 überein, somit ist die Abweichung vom Gitter hier der Grund für diese Intensitätsmodulation. Außerdem fällt bei genauerer Betrachtung eine weitere höherfrequente Modulation, die ein diagonales Streifenmuster erzeugt, innerhalb der Periode auf. Auch diese Modulation ist schon in Abbildung 7.6 zu erkennen, daher sind beide Bildartefakte auf Abweichungen des Gitters zurückzuführen.

Durch Integration über eine größere Detektorfläche lässt sich die Abweichung des Beleuchtungsrasters kompensieren. In Abbildung 7.9 sieht man ein Beispiel für eine Rekonstruktion, bei der über eine Fläche von 3×3 Pixel integriert wird und in Abbildung 7.10 eine Rekonstruktion, bei der über eine Fläche von 9×9 Pixel integriert wird.

Man sieht, dass das Vergrößern der Integrationsumgebung auf 3×3 Pixel die Artefakte nur leicht verringert. Die Vergrößerung auf 9×9 Pixel verringert die Artefakte jedoch deutlich. Sowohl der Intensitätsabfall zum Rand der Periode, als auch die höherfrequente Modulation innerhalb der Periode sind in Abbildung 7.10 nicht mehr vorhanden.

Das Wählen einer Integrationsumgebung von 9×9 Pixel stellt sicher, dass auch im

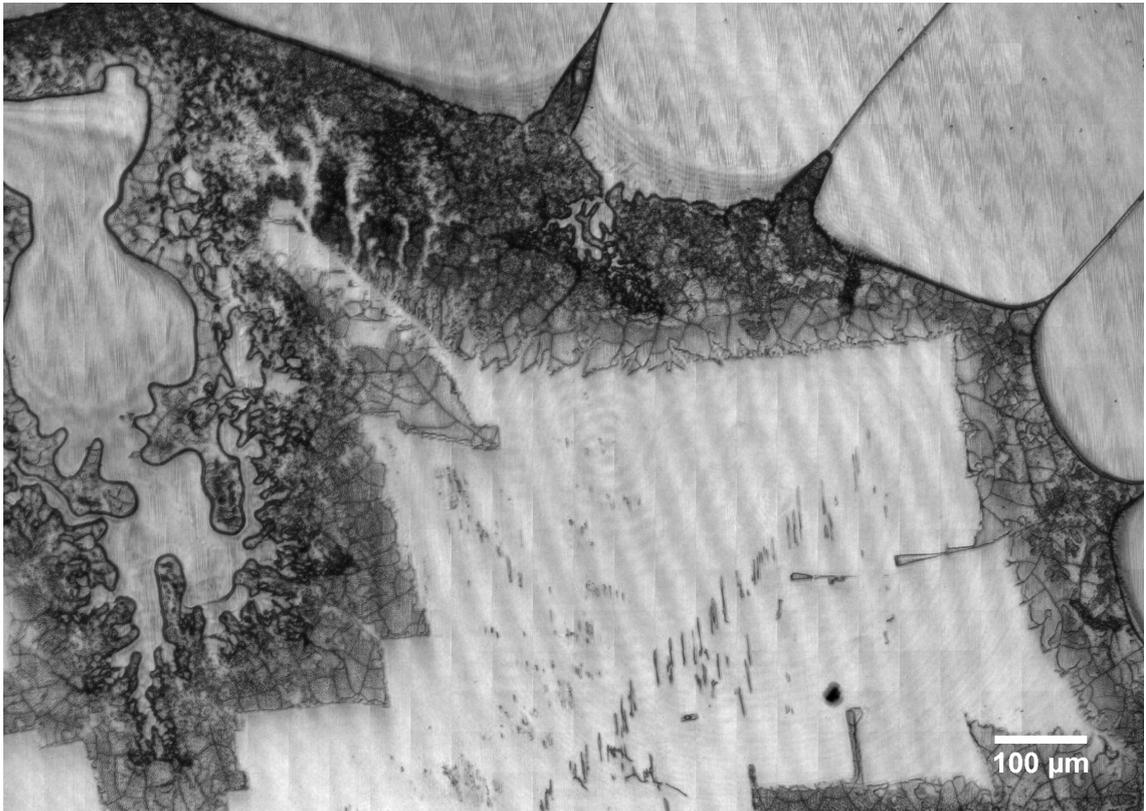
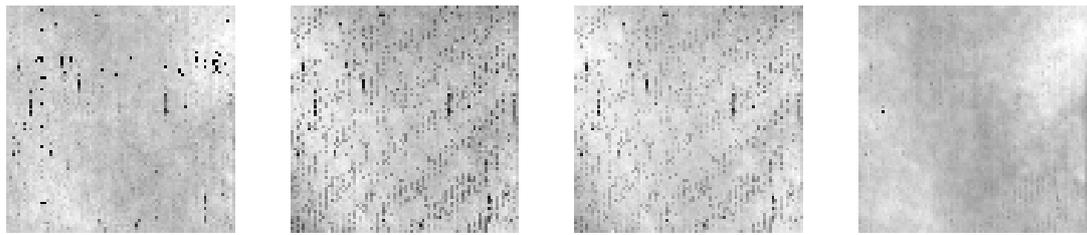


Abbildung 7.10: Rekonstruktion eines Testobjekts. Für die Rekonstruktion wurde die Intensität an den jeweiligen Gitterpunkten in einer 9×9 Pixel großen Umgebung aufsummiert und in das Ergebnisbild eingetragen. Das Gitter zur Rekonstruktion wurde aus allen Aufnahmen bestimmt.



a) Individuelles Grid, Peak Intensität b) Gemitteltes Grid, Peak Intensität c) Gemitteltes Grid, 3×3 Pixel Umgebung d) Gemitteltes Grid, 9×9 Pixel Umgebung

Abbildung 7.11: Vergleich der einzelnen Rekonstruktionsmethoden anhand einer homogenen, nicht absorbierenden Fläche der Probe von der Größe einer Periode.

Fall einer Verschiebung von bis zu zwei Pixeln immer noch über den ganzen Peak in der Sensorebene integriert wird. Die kleinere Integrationsumgebung von 3×3 Pixel schafft es hingegen nicht, die Abweichung der Peaks vom Gitter zu kompensieren.

Ähnlich wie bei der Rekonstruktion mit individuellem Auswertungs-Gitter sind auch hier Unstetigkeiten an der Periodengrenze sichtbar. Mögliche Gründe hierfür sind die in Abbildung 7.6 beobachteten Abweichungen des Beleuchtungsrasters, sowie eine mögliche Verdrehung zwischen den Piezo-Achsen und den DOE-Achsen.

Rekonstruktion mit initialer Bestimmung des Auswertungsgitters

Im vorherigen Abschnitt wurde gezeigt, dass zwischen den Gitterpunkten aus dem ersten Bild und den Gitterpunkten, die aus allen Bildern bestimmt wurden, nur eine geringe durchschnittliche Abweichung besteht. Beachtet man zusätzlich, dass eine Integration über eine größere Pixelumgebung vorteilhaft ist um die Rekonstruktionsartefakte, die durch Abweichungen vom Gitter entstehen, zu minimieren, so lässt sich Schlussfolgern, dass die Bestimmung des Gitters aus allen Bildern unnötig ist. Für eine Rekonstruktion mit 9×9 Pixelumgebung ist eine initiale Gitterbestimmung aus dem ersten Bild völlig ausreichen.

Fazit zur Auswertung der transmittierten Intensität

Anhand der Qualität der Rekonstruktion der homogenen nicht absorbierenden Flächen lassen sich die einzelnen Auswertungsmethoden beurteilen. In Abbildung 7.11 sieht man Ausschnitte von der Größe einer Periode aus den Abbildungen 7.7 - 7.10, die den gleichen homogenen, nicht absorbierenden Bereich aus der Mitte der Rekonstruktionen zeigen.

Die Auswertung mit dem gemittelten Gitterraster und einer großer Integrationsumgebung liefern das Rekonstruktionsergebnis mit den wenigsten Artefakten. Qualitativ entspricht diese Rekonstruktion der Rekonstruktion mit individuellem Grid und der Auswertung der Peak-Intensität, welche dem idealen konfokalen Scanning-

Prinzip am nächsten kommt. Jedoch enthält Letztere deutliche Artefakte, die durch eine fehlgeschlagene Bestimmung des Gitterrasters hervorgerufen werden.

Da für eine homogene Rekonstruktion ein großes Integrationsgebiet gewählt werden muss, ist es möglich, statt des gemittelten Gitters die Gitterparameter nur anhand des ersten Bilds zu bestimmen, denn die durchschnittliche Abweichung ist sehr viel kleiner als die Breite der Integrationsumgebung. Im Folgendem werden deshalb immer die Gitterparameter aus der ersten Messung bestimmt und für die Rekonstruktion eine große Integrationsumgebung verwendet.

7.6 Experimentelle Bestimmung der maximalen Auflösung des parallel scannenden Mikroskops

In diesem Abschnitt wird zuerst experimentell die maximale Auflösung des parallel scannenden Mikroskops bestimmt. Anschließend wird das experimentelle Ergebnis mit den theoretischen Vorhersagen aus Abschnitt 7.3.2 verglichen.

7.6.1 Bestimmung der maximalen Auflösung durch den Gitterkontrast eines Auflösungs-tests

Um die Auflösung des parallel scannenden Mikroskops experimentell zu quantifizieren, wurde ein Auflösungs-test von Zeiss benutzt. Der Auflösungs-test besteht aus einem Glassubstrat mit einer Chromstrukturierung. Die Chromstrukturierung besteht aus mehreren Gittern mit einer unterschiedlichen Anzahl an Linienpaaren pro Millimeter. Jedes Gitter besteht aus jeweils fünf Gitterbalken in vertikaler und horizontaler Richtung und einer Beschriftung, die die Anzahl an Linienpaaren pro Millimeter angibt.

Zur Bestimmung der maximalen Auflösung wird quasi eine tomographische Aufnahme des Auflösungs-tests gemacht, indem mit dem Piezo die Beleuchtungsebene in verschiedenen z-Ebenen verschoben wird und jeweils eine komplette Messung gemacht wird. Anschließend wird indirekt auf die optimale z-Einstellung geschlossen, bei der in der Rekonstruktion die maximale Auflösung erreicht wird.

In Abbildung 7.12 sind Rekonstruktionen für verschiedene Bereiche des Auflösungs-tests bei der optimalen z-Einstellung gezeigt. Zur besseren Visualisierung sind die einzelnen Teile der Rekonstruktion nebeneinander abgebildet, auf dem Auflösungs-test sind die Gitter vertikal untereinander angeordnet.

In Abbildung 7.13 sieht man eine Vergrößerung der Gitterstrukturen mit 900, 1000, 1100 und 1250 Linienpaaren pro Millimeter. Man sieht, dass der Kontrast über den vertikalen und horizontalen Gittern unterschiedlich ist. Im Allgemeinen ist der Kontrast über dem horizontalen Gitter besser als über dem vertikalen Gitter. Aus dem Kontrast lässt sich die maximale Auflösung des Mikroskops ermitteln.

Zur besseren Visualisierung des Kontrastes sind in Abbildung 7.14 Linescans durch die Gitterstrukturen gezeigt. Je nach Orientierung des Gitters kann man die fünf

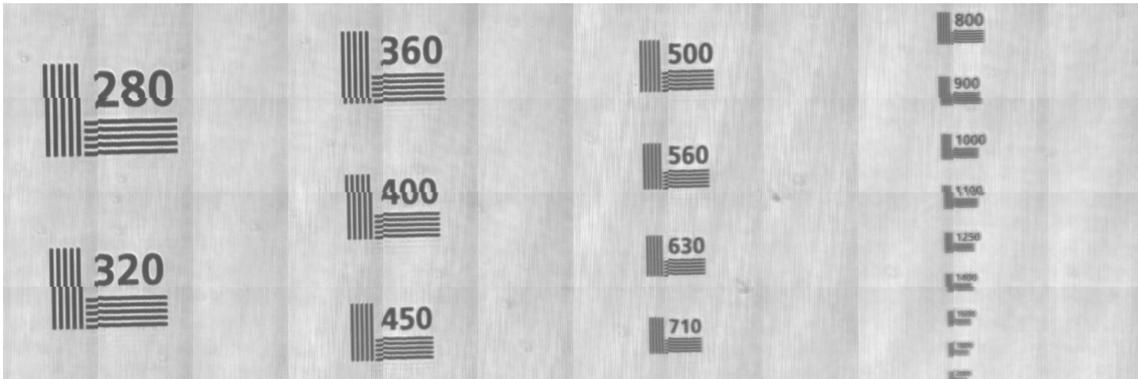


Abbildung 7.12: Rekonstruktion des Auflösungstests. Zur besseren Darstellbarkeit sind die einzelnen Teilbereiche des Auflösungstest nebeneinander angeordnet.

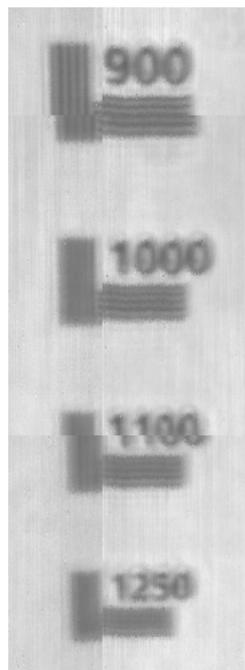
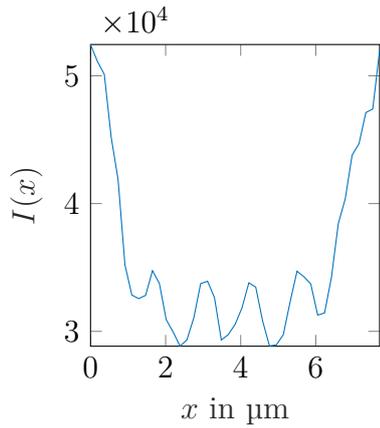
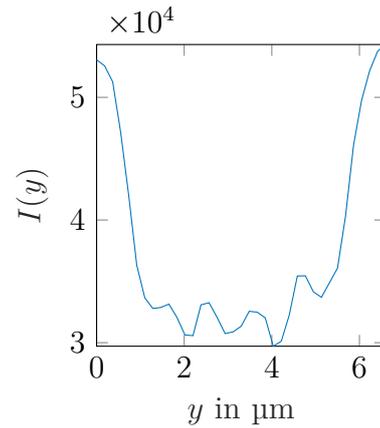


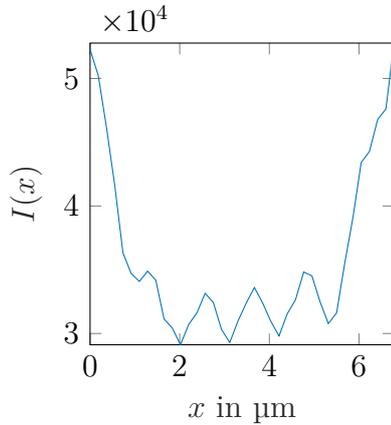
Abbildung 7.13: Vergrößerung der Rekonstruktion des Auflösungstest aus Abbildung 7.12 um die maximale Auflösung abzuschätzen.



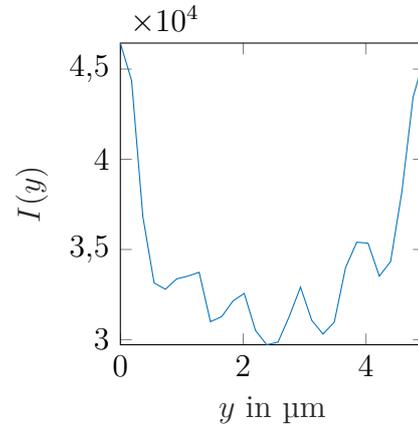
a) Horizontaler Linescan bei 800 lp/mm.



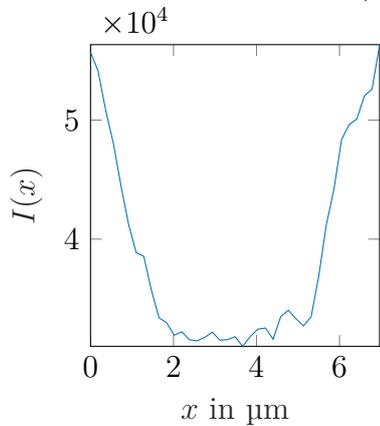
b) Vertikaler Linescan bei 1000 lp/mm.



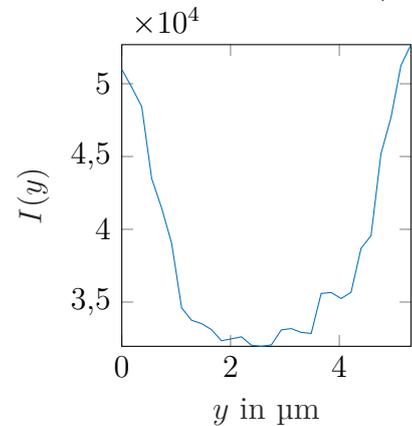
c) Horizontaler Linescan bei 900 lp/mm.



d) Vertikaler Linescan bei 1100 lp/mm.



e) Horizontaler Linescan bei 1000 lp/mm.



f) Vertikaler Linescan bei 1250 lp/mm.

Abbildung 7.14: In a), c) und e) sind Linescans durch das vertikale Gitter bei 800, 900 und 1000 Linienpaaren pro Millimeter gezeigt. In b), d) und f) sind Linescans durch das horizontale Gitter bei 1000, 1100 und 1250 Linienpaaren pro Millimeter gezeigt. Die einzelnen Balken des Gitters zeigen sind in den Fällen a) bis d) als separate Minima.

Balken bis zu einer anderen Liniendichte noch auflösen. In Abbildung 7.14c sieht man, dass im Linescan des vertikalen Gitters mit 900 Linienpaaren pro Millimeter die fünf Balken sich noch deutlich voneinander trennen lassen. Bei 1000 Linienpaaren pro Millimeter hingegen sind die einzelnen Balken nicht mehr voneinander unterscheidbar. Für das horizontale Gitter lassen sich sogar im Linescan durch die Gitterstruktur mit 1100 Linienpaaren pro Millimeter die fünf Balken noch deutlich voneinander unterscheiden. Der zugehörige Linescan ist in Abbildung 7.14d zu sehen. Hingegen lassen sich bei 1250 Linienpaaren pro Millimeter die einzelnen Balken bei dieser Orientierung des Gitters nicht unterscheiden.

Die Asymmetrie bezüglich der maximalen Auflösung der horizontalen und vertikalen Gitter lässt sich durch die Asymmetrie beim Abscannen der Grundperiode erklären. Durch das Verwenden eines Schlangenrasters werden die Messpunkte über das horizontale Gitter direkt hintereinander aufgenommen, die Aufnahme der Messpunkte über das vertikale Gitter dauert hingegen deutlich länger. Somit wirken sich zeitliche Störungen wie zum Beispiel mechanische Drifts des Aufbaus oder Hysterese-Effekte deutlicher auf die vertikale Auflösung auf.

Experimentell lassen sich in der Rekonstruktion also Strukturen bis zu einer Größe $\delta y = \frac{1}{2 \times 1100} \text{mm} \approx 450 \text{ nm}$ und $\delta x = \frac{1}{2 \times 900} \text{mm} \approx 550 \text{ nm}$ auflösen.

Diskussion zur Messmethode Diese Methode zur Bestimmung der Auflösungsgrenze liefert ein einfach auszuwertendes Ergebnis, aber sie ist sehr abhängig von der richtigen Justierung des Aufbaus. Insbesondere muss sichergestellt werden, dass die Foki das Gitter in der richtigen z-Ebene abtasten. In der Tomographie ist diese Sensitivität bezüglich der z-Ebene der große Vorteil der Scanning-Mikroskopie gegenüber Weitfeld-Verfahren [21], da Streulichtbeiträge aus Ebenen die nicht im Fokus sind nur einen geringen niederfrequenten Beitrag liefern. Bei der Bestimmung der maximalen Auflösung mit Hilfe eines dünnen Gitters erschwert diese Sensitivität die Messung. Dies führt dazu, dass die Messung sehr zeitintensiv ist. Diese direkte Methode zur Bestimmung der maximalen Auflösung anhand eines fünf Balkengitters ist von mehreren Standardisierungsorganisationen empfohlen (NBS 1010a testchart [99], ANSI/ISO 3334:2006 [42]).

Neben der direkten Bestimmung der Auflösung über den Kontrast eines Gitters lässt sich die Auflösung des Scanning-Mikroskops über die Analyse einer stark absorbierenden Kante bestimmen [4]. Ähnlich wie bei der Untersuchung des Gitters muss auch hier eine tomographische Aufnahme der Kante zur Bestimmung der optimalen z-Justierung aufgenommen werden. Neben der maximalen Auflösung lässt sich hier die Modulations-Transfer-Funktion (MTF) des Systems bestimmen. Streng genommen ist die Berechnung der MTF jedoch nur für den Fall einer idealen konfokalen Blende möglich und nicht für Systeme mit einer größeren konfokalen Blende. Die Bestimmung der MTF aus nur einem Linescan stellt jedoch sehr hohe Anforderungen an die Genauigkeit der Abtastung der Kante und eine rauscharme Aufnahme, sodass in der Praxis sich die Abbildung einer schrägen Kante durchgesetzt hat [18, 85].

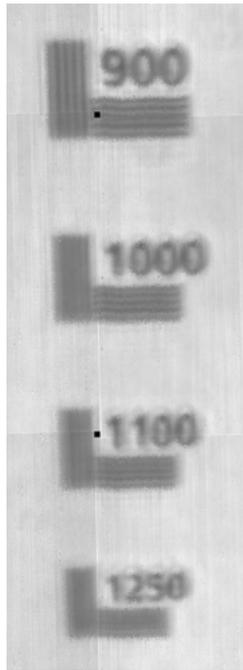


Abbildung 7.15: Ausschnitt der um 1° gedrehten Rekonstruktion.

Korrektur der Unstetigkeiten in der Rekonstruktion an den Periodengrenzen in der Messung zur Bestimmung der maximalen Auflösung

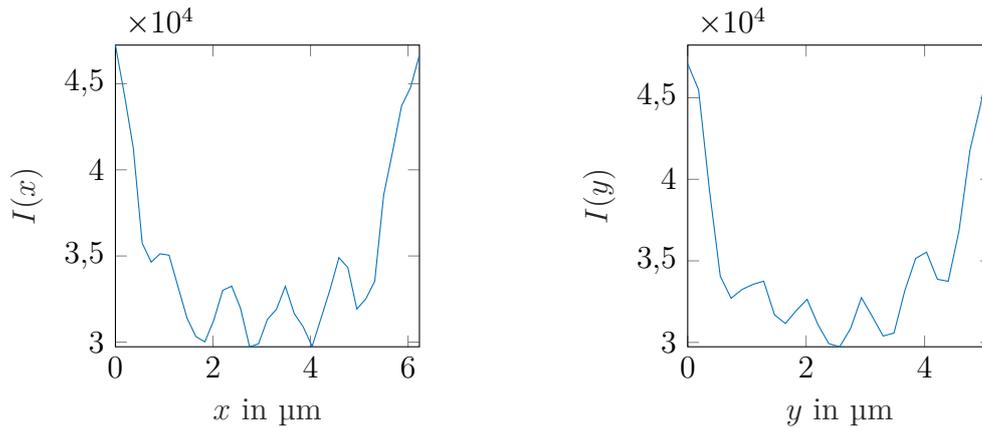
Insbesondere bei sehr hohen Pixelauflösungen lassen sich in der Rekonstruktion Unstetigkeiten an den Periodengrenzen feststellen (vgl. Abbildung 7.13). Ein möglicher Grund für diese Unstetigkeit könnte ein Verdrehungswinkel zwischen Piezoachse und Gitterachse des diffraktiven Elements sein.

Durch eine Rotation der Teilrekonstruktion innerhalb der einzelnen Periode des diffraktiven Elements lässt sich diese Unstetigkeit beseitigen. In Abbildung 7.15 sieht man die um 1° korrigierte Rekonstruktion. Die Übergänge an den Periodengrenzen sind nun stetig. Bei der Berechnung der Drehung wurde eine bilineare Interpolation verwendet.

Um den Einfluss der Korrektur auf die Auflösung beurteilen zu können sind in Abbildung 7.16 wieder Linescans durch das vertikale Gitter mit 900 Linienpaaren pro Millimeter und das horizontale Gitter mit 1100 Linienpaaren pro Millimeter gezeigt. In beiden Linescans lassen sich auch hier die fünf Balken als deutlich voneinander unterscheiden. Auch der Kontrast entspricht in etwa dem der Originalrekonstruktion. Somit lässt sich festhalten, dass die Korrektur der Unstetigkeit an der Periodengrenze die Auflösung der Rekonstruktion nicht stark beeinflusst.

7.6.2 Vergleich mit der Theorie

Die experimentell bestimmte Auflösung lässt sich mit der Theorie aus Abschnitt 7.3.2 vergleichen. Die numerische Apertur des diffraktiven Mikrolinsenarrays der



a) Horizontaler Linescan bei 900 lp/mm b) Vertikaler Linescan bei 1100 lp/mm

Abbildung 7.16: Linescans durch die vertikale Gitterstruktur und durch die horizontale Gitterstruktur der um 1° gedrehten Rekonstruktion.

Beleuchtung ist $NA_{\text{Bel}} = 0,48$, die verwendete Wellenlänge ist $\lambda = 0,532 \mu\text{m}$. Nimmt man an, dass die numerische Apertur der Beobachtung sehr klein ist, ergibt sich die maximale Auflösung aus Gleichung 7.43 zu $\delta x = \frac{\lambda}{2NA_{\text{Bel}}} = 554 \text{ nm}$.

Die beobachtete maximale Auflösung von 1100 Linienpaaren pro Millimeter bzw. die korrespondierende minimale noch aufzulösende Strukturgröße von $\delta x \approx 454 \text{ nm}$, ist kleiner als die Auflösungsgrenze in der Näherung für vernachlässigbar kleine numerische Aperturen bei der Beobachtungsoptik.

Für nicht vernachlässigbar kleine numerische Aperturen in der Beobachtungsoptik bestimmt sich die maximale Auflösung nach Gleichung 7.29. In Tabelle 7.1 ist die Auflösungsgrenze für verschiedene numerische Aperturen der Beobachtungsoptik NA_{Beob} aufgelistet.

Die experimentell bestimmte Auflösungsgrenze liefert eine korrespondierende effektive numerischen Apertur von ca. $NA_{\text{Beob}} \approx 0,1$. Dieser Wert ist deutlich kleiner als die vom Hersteller spezifizierte numerische Apertur von $NA_{\text{Beob}} = 0,32$. Der Unterschied zwischen der effektiven numerischen Apertur und der spezifizierten numerischen Apertur sind durch die starken Aberrationen des Objektivs bei voll geöffneter Blende zu erklären.

NA_{Beob}	$NA_{\text{Bel}} + NA_{\text{Beob}}$	$\delta x = \frac{532 \text{ nm}}{2 \times (NA_{\text{Bel}} + NA_{\text{Beob}})}$ in nm
0	0,56	554
0,08	0,56	475
0,10	0,58	459
0,12	0,60	443
0,16	0,64	416
0,24	0,72	369
0,32	0,80	333

Tabelle 7.1: Tabellarische Auflistung der theoretischen Auflösungsgrenze des scannenden Mikroskops nach Gleichung 7.29 für eine Wellenlänge von $\lambda = 532 \text{ nm}$, einer numerischen Apertur von $NA_{\text{Bel}} = 0,48$ für die diffraktiven Elemente der Beleuchtung und für verschiedenen numerischen Aperturen für die Beobachtungsoptik.

7.7 Ausblick auf weitere Anwendungsgebiete diffraktiver Linsen mit überlappenden Aperturen in der Mikroskopie

Weitere potentielle Anwendungsgebiete der diffraktiven Linsen mit überlappenden Aperturen lassen sich auch in den weiteren punktscannenden Mikroskopieverfahren finden. Anstatt die Transmission der Foki der diffraktiven Linsen direkt zu untersuchen wäre es möglich, die Foki zur Anregung von fluoreszenzmarkierten Proben zu verwenden und somit ein parallel scannendes Fluoreszenzmikroskop zu realisieren. Baulich würde sich so ein Aufbau nur durch einen zusätzlichen Filter in der Abbildungsoptik unterscheiden, der das Anregungslicht von dem Fluoreszenzlicht trennt. Auch parallele Varianten der STED-Mikroskopie ließen sich realisieren, hierbei wäre es insbesondere auch möglich die Ausschaltfoki (engl. *depletion foci*) durch ein diffraktives Element zu erzeugen. Um einen räumlichen Zugang zum Fluoreszenzlicht zu erhalten wäre es von Vorteil die Anregungsfoki und die Ausschaltfoki durch dasselbe diffraktive Profil zu erzeugen. Problematisch bei den Fluoreszenzverfahren ist jedoch, dass zur Beobachtung immer eine Optik mit hoher numerischer Apertur eingesetzt wird, um möglichst viel der isotrop abgestrahlten Fluoreszenz zu detektieren. Dies schränkt das Sichtfeld ein und somit lässt sich nur eine geringe Parallelität erreichen.

7.8 Zusammenfassung und Fazit

In diesem Kapitel wurde das parallel scannende Mikroskop als eine Anwendung für die diffraktiven Linsen mit überlappenden Aperturen vorgestellt. Die Verwendung der diffraktiven Linsen erlaubt die gleichzeitige Untersuchung eines großen Sichtfelds

durch eine hohe Parallelität.

Analytisch wurde das Prinzip der Bildgebung des parallel scannenden Mikroskops untersucht und die Auflösungsgrenze bestimmt. Insbesondere wurde der Spezialfall einer Beobachtungsoptik mit niedriger numerischer Apertur betrachtet.

Es wurde ein Testaufbau realisiert, bei dem die Verwendung eines hochaperturigen diffraktiven Mikrolinsenarrays mit $NA = 0,48$ zur Beleuchtung eine hohe Auflösung und die Verwendung einer Abbildungsoptik mit nur geringer Vergrößerung gleichzeitig ein großes Sichtfeld von ca. $900 \mu\text{m} \times 1200 \mu\text{m}$ ermöglicht.

Mit der Untersuchung eines großflächigen Testobjekts wurde das Prinzip und die Bildrekonstruktion aus den Messdaten erfolgreich demonstriert.

Bei der experimentellen Bestimmung der Auflösung anhand eines Auflösungstests konnten Strukturen bis zu einer Größe von 450 nm bei einer Wellenlänge von $\lambda = 532 \text{ nm}$ aufgelöst werden. Die erreichte Auflösung des Testsystems ist besser als die Auflösung in einem Weitfeldmikroskop, bei dem die Probe homogen kohärent beleuchtet wird und das Objektiv die gleiche numerischen Apertur von $NA = 0,48$ wie die im Mikrolinsen im Testsystem hat.

8 Anwendung: Wellenfrontmessung

Typische Anwendungsfelder für die Wellenfrontmessung finden sich in der Bildgebung in Astronomie [68] oder der Ophthalmologie [61]. Durch die Kenntnis der Wellenfront lassen sich die dynamischen Störungen der Wellenfront durch die Atmosphäre korrigieren, sodass erdgebundene Teleskope mit großen Spiegeldurchmessern sinnvoll betrieben werden können. Außerdem lassen sich die statischen Abbildungsfehler des Auges quantifizieren, sodass diese durch refraktive Chirurgie vermindert werden können. Im Allgemeinen lässt sich die Qualität von abbildenden optischen Systemen durch die Messung der Verzerrungen einer ebenen Wellenfront quantitativ bestimmen. Die Messung mittels eines Hartmann-Shack-Sensors ist eine Standardmethode in der Optik [82]. Dieses Prinzip wird im folgenden Kapitel erklärt. Anschließend wird auf die Limitationen bei der Wellenfrontmessung mittels Hartmann-Shack-Sensor eingegangen.

Da diffraktive Linsen mit überlappenden Aperturen die Erzeugung von Fokusarrays mit geringem Fokus-Pitch und einer langen Brennweite erlauben, wird untersucht, ob diese Eigenschaft die Wellenfrontmessung verbessern kann.

8.1 Wellenfrontmessung mittels eines Hartmann-Shack-Sensors

Prinzipiell besteht der Hartmann-Shack-Sensor nur aus einem Mikrolinsenarray und einer Kamera [82]. Das Mikrolinsenarray ist so zur Kamera ausgerichtet, dass bei senkrechter Beleuchtung die Foki direkt in der Sensorebene der Kamera erzeugt werden. Jede Mikrolinse schneidet aus einer ankommenden Wellenfront ein Teil aus und fokussiert diesen Beitrag auf die Kamera. Aus den Verschiebungen δ_x bzw. δ_y des Fokus auf dem Sensor gegenüber der Linsenachse lässt sich der mittlere lokale Gradient der Wellenfront $\nabla\phi_{\text{lokal}}$ in der Mikrolinsenebene unter der Mikrolinsenapertur bestimmen:

$$\nabla_x\phi_{\text{lokal}} = \frac{\partial\phi_{\text{lokal}}}{\partial x} = k_0 \frac{\delta_x}{\sqrt{\delta_x^2 + \delta_y^2 + f^2}} \quad (8.1)$$

$$\nabla_y\phi_{\text{lokal}} = \frac{\partial\phi_{\text{lokal}}}{\partial y} = k_0 \frac{\delta_y}{\sqrt{\delta_x^2 + \delta_y^2 + f^2}}. \quad (8.2)$$

Dabei ist f die Brennweite der Mikrolinsen, δ_x und δ_y sind die Abweichung vom Schwerpunkt gegenüber der optischen Achse und k_0 ist die Wellenzahl des Lichts. Man erhält also Informationen über die gemittelten lokalen Gradienten der Wellenfront. Hieraus kann anschließend die Wellenfront rekonstruiert werden [14, 59, 89, 101].

Bestimmung der Einflüsse auf die Genauigkeit der Gradientenmessung

Sowohl die Brennweite als auch der Mikrolinsendurchmesser haben einen direkten Einfluss auf die Gradientenmessung. Der Mikrolinsendurchmesser bestimmt die Abtastung der Wellenfront. Je kleiner der Durchmesser ist, desto mehr Abtastpunkte sind möglich. Die Brennweite der Mikrolinse bestimmt die Länge des optischen Hebels und schränkt dadurch den messbaren Wertebereich der Gradienten ein. Die Unsicherheiten auf die lokalen Phasengradienten ergeben sich zu:

$$\Delta \nabla_x \phi_{\text{lokal}} = k_0 \sqrt{\left(\frac{(\delta_y^2 + f^2) \Delta x_{\text{Schwerpunkt}}}{(\delta_x^2 + \delta_y^2 + f^2)^{\frac{3}{2}}} \right)^2 + \left(\frac{\delta_x \delta_y \Delta y_{\text{Schwerpunkt}}}{(\delta_x^2 + \delta_y^2 + f^2)^{\frac{3}{2}}} \right)^2} \quad (8.3)$$

$$\Delta \nabla_y \phi_{\text{lokal}} = k_0 \sqrt{\left(\frac{\delta_x \delta_y \Delta x_{\text{Schwerpunkt}}}{(\delta_x^2 + \delta_y^2 + f^2)^{\frac{3}{2}}} \right)^2 + \left(\frac{(\delta_x^2 + f^2) \Delta y_{\text{Schwerpunkt}}}{(\delta_x^2 + \delta_y^2 + f^2)^{\frac{3}{2}}} \right)^2} \quad (8.4)$$

Hierbei ist $\Delta x_{\text{Schwerpunkt}}$ und $\Delta y_{\text{Schwerpunkt}}$ die Unsicherheit der Schwerpunktsbestimmung in x- und y-Richtung. Oft gilt $\delta_x \ll f$ und $\delta_y \ll f$, sodass die Unsicherheit im Phasengradienten weiter genährt werden können:

$$\Delta \nabla_x \phi_{\text{lokal}} \approx \frac{k_0}{f} \Delta x_{\text{Schwerpunkt}} \quad (8.5)$$

$$\Delta \nabla_y \phi_{\text{lokal}} \approx \frac{k_0}{f} \Delta y_{\text{Schwerpunkt}} \quad (8.6)$$

Man sieht, dass die Fokusslänge neben der Genauigkeit der Schwerpunktsbestimmung die limitierende Größe für die Genauigkeit der Gradientenmessung ist.

Anforderungen an das Mikrolinsenarray des Hartmann-Shack-Sensors

Insbesondere für die Vermessung von Wellenfronten mit kleinen Gradienten ergeben sich folgende Anforderungen an das Mikrolinsenarray:

- Um eine hohe räumliche Abtastung der Wellenfront zu gewährleisten, soll der Mikrolinsendurchmesser möglichst klein sein.
- Um geringe Unsicherheiten auf die Phasengradienten zu erreichen, soll die Mikrolinsenfokusslänge möglichst lang sein.
- Die Fokusbreite muss deutlich kleiner als der Mikrolinsendurchmesser sein, damit eine Schwerpunktsbestimmung möglich ist.

Diese drei Forderungen sind bei refraktiven Linsen nicht voneinander unabhängig, da die minimale Fokusgröße durch den Mikrolinsendurchmesser und die Brennweite bestimmt sind. Fixiert man zum Beispiel den Mikrolinsendurchmesser auf $D = 100 \mu\text{m}$ und die numerische Apertur auf $NA = 0,05$, dann ist für Licht der Wellenlänge $\lambda = 0,5 \mu\text{m}$ der Fokus ca. $\delta x = 10 \mu\text{m}$ breit. Die maximale Fokusdistanz für diese Linse folgt aus $f \approx \frac{D}{2\delta x_{\text{Fokus}}}$ zu $f \approx 1000 \mu\text{m}$.

8.2 Potenzial von diffraktiven Mikrolinsen im Hartmann-Shack-Sensor

Diffraktive Mikrolinsenarrays mit überlappenden Aperturen haben im Gegensatz zu refraktiven Mikrolinsenarrays keine starre Beziehung zwischen Fokusgröße, Brennweite und lateralem Fokusabstand. Somit ist es möglich Mikrolinsenarrays mit längeren Brennweiten zu konstruieren, als es im refraktiven Fall möglich ist. Das Potential, die Genauigkeit der Gradientenmessung und damit die Genauigkeit der Wellenfrontrekonstruktion mit diffraktiven Mikrolinsen mit überlappenden Aperturen zu steigern wurde zuerst von Stenau und Brenner [91] erkannt. Die längeren Brennweiten sind durch das Überlappen der Aperturen in der Mikrolinsenebene möglich. Das hat jedoch den Nachteil, dass der nun gemessene Gradient über einen größeren Bereich als im refraktiven Fall gemittelt wird. Somit sind die Gradienten nicht mehr unabhängig voneinander.

Benutzt man die ermittelten Gradienten zur Rekonstruktion der Wellenfront ohne den Überlapp zu berücksichtigen, so erhält man keine bessere Genauigkeit als im Vergleich zu einem Mikrolinsenarray ohne Überlapp.

Um das Potential der diffraktiven Mikrolinsen mit überlappender Apertur ausnutzen zu können, muss eine Wellenfrontrekonstruktion konstruiert werden, die berücksichtigt, dass die Mittelung der Wellenfront über einen Bereich erfolgt, der größer als das Sampling der Wellenfront ist.

Dieser Rekonstruktionsalgorithmus wäre nicht nur für den Fall von diffraktiven Linsen mit überlappenden Aperturen sinnvoll, sondern könnte auch zu verbesserten Wellenfrontrekonstruktionen mit klassischen Hartmann-Shack-Sensoren führen. Nimmt man mit einem klassischen Hartmann-Shack-Sensor zum Beispiel vier Bilder einer Wellenfront auf und verschiebt zwischen den einzelnen Bildern den Wellenfront Sensor um die Distanzen $(\Delta x, \Delta y) = (P_x/2, 0)$, $(\Delta x, \Delta y) = (0, P_y/2)$ und $(\Delta x, \Delta y) = (P_x/2, P_y/2)$ und speichert die ermittelten Gradienten in einer gemeinsamen Karte für die lokalen Gradienten, dann entspricht dieser Datensatz einer Messung mit diffraktiven Linsen und einem Überlappfaktor von 2.

Die Entwicklung eines solchen Rekonstruktions-Algorithmus stellt eine spannende Aufgabe dar, die in nachfolgenden Projekten untersucht werden kann. Ein möglicher Ansatz ist, die gemessenen Phasengradienten $\nabla_{\perp} \phi_{\text{lokal}}^{\text{gem.}}$ als zweidimensionale Faltung einer Fensterfunktion h mit dem originalen Phasengradienten $\nabla_{\perp} \phi_{\text{lokal}}^{\text{orig.}}$ darzustellen. So kann $\nabla \phi_{\text{lokal}}^{\text{gem.}} = h \otimes \nabla_{\perp} \phi_{\text{lokal}}^{\text{orig.}}$ durch eine Entfaltung vor der Wellenfrontrekonstruktion durchgeführt werden.

8.3 Fazit

Diffraktive Linsen mit überlappenden Aperturen haben das Potenzial, die Wellenfrontrekonstruktion im Hartmann-Shack-Sensor zu verbessern, indem längere Brennweiten verwendet werden. Für die so gewonnenen Daten muss jedoch noch ein Rekonstruktionsalgorithmus gefunden werden, der die Kopplung zwischen den benach-

barten Gradienten berücksichtigt. Eine naive Benutzung der üblichen Wellenfrontrekonstruktionsalgorithmen auf Messdaten von Linsen mit überlappenden Aperturen führt zu keiner Verbesserung, sondern zu einer Verschlechterung der Wellenfrontrekonstruktion im Vergleich zu normalen Mikrolinsen. Die Entwicklung dieses neuen Algorithmus kann auch die Wellenfrontrekonstruktion für den klassischen Hartmann-Shack-Sensor verbessern, da durch laterale Bewegung des Sensors die Funktionsweise von überlappenden Aperturen simuliert werden kann, um eine höhere räumliche Abtastung des Wellenfelds zu erreichen.

9 Aspekte zur nicht skalaren Untersuchung von diffraktiven Linsen

Das Design der vorgestellten diffraktiven Linsen mit überlappenden Aperturen beruht auf einer skalaren Beschreibung des Lichts und der dünnen Element Näherung. In diesem Kapitel soll untersucht werden, ob die Funktion der diffraktiven Linsen noch gewährleistet ist, wenn man den Vektorcharakter des Lichts und die reale Ausdehnung der diffraktiven Elemente berücksichtigt. Hierzu wird die Lichtausbreitung durch die diffraktiven Elemente mit der „rigorous coupled-wave analysis“ (RCWA) berechnet. In Abschnitt 9.1 werden zuerst die Eignung und die Limitationen der RCWA für Simulationen von Linsen mit überlappenden Aperturen diskutiert. Anschließend wird der Übergang von einem dünnen Element zu einem dicken Element mit einer komplexen Permittivitätsverteilung betrachtet (Abschnitt 9.2). Dabei wird die Intensität im Fokus aus der skalaren Simulation des dünnen Elements mit der RCWA Simulation des dicken Elements verglichen. In Abschnitt 9.3 wird eine Sonderfallbehandlung für homogene Schichten in der RCWA vorgestellt, die es zum Beispiel erlaubt, eine zusätzliche Lichtausbreitung durch ein Deckglas performant zu simulieren. Abschließend wird noch gezeigt, dass die RCWA in sich konsistente Ergebnisse liefert, sodass die Impulserhaltung gilt (Abschnitt 9.4).

9.1 RCWA als Simulationsmethode für diffraktive Elemente mit überlappenden Aperturen

Die RCWA als allgemeine Lösungsmethode der Maxwell-Gleichungen für periodische Probleme wurde in Abschnitt 3.3.1 vorgestellt. Es handelt sich um eine Methode, die den Vektorcharakter des Lichts und die räumliche Ausdehnung von Gitterstrukturen berücksichtigt. In diesem Sinn ist die RCWA Simulation physikalischer als die skalare Simulation mit dünner Element Näherung. Da diffraktive Linsen mit überlappenden Aperturen im Idealfall nur über eine Grundperiode beschrieben werden, die sich unendlich periodisch wiederholt, eignen sie sich für eine Simulation mittels RCWA.

Unterschiede zwischen RCWA und skalarer Simulation mit dünner Element Näherung

Im Gegensatz zur skalaren Simulation können mit der RCWA verschiedene Polarisierungen der Beleuchtungswelle simuliert werden. Bei refraktiven Linsen spielt die Polarisation bei sehr hohen numerischen Aperturen eine wichtige Rolle und beeinflusst die Fokusgeometrie [25].

Das Gitter wird nun anstatt durch eine relative Abschwächung und Phasenverzögerung durch eine dreidimensionale Permeabilitäts- und Permittivitätsverteilung beschrieben. An den Grenzflächen werden in der RCWA die polarisationsspezifische Reflexion und Transmission berücksichtigt. Außerdem werden Mehrfachreflexionen an und innerhalb des Gitters berücksichtigt. Dies ist insbesondere für die Simulation mit einer verkippten Beleuchtungswelle relevant.

Hinter dem Gitter erhält man eine vektorielle Feldverteilung, die Intensität ist durch das Amplitudenbetragsquadrat der elektrischen Feldstärke gegeben.

Praktische Limitationen bei der Simulation mit der RCWA

Im Allgemeinen lassen sich mit der RCWA nur modenbegrenzte Probleme exakt lösen [1]. Für praktische Rechnungen bedeutet dies, dass man bei der Beschreibung des dreidimensionalen Gitters auf wenige Moden begrenzt ist (deutlich weniger als 100×100). Diese Limitierung kommt durch den enormen Speicherbedarf beim Lösen der Randbedingungen mittels des „enhanced transmittance matrix approach“ (ETMA). Des Weiteren gibt es eine Limitierung durch die Rechenzeit, da die Matrixinversionen einer quadratischen Matrix der Größe $N \times N$ sowie die Berechnung der Eigenvektoren und Eigenwerte typischerweise mit $\mathcal{O}(N^3)$ (gaußsches Eliminationsverfahren) bis $\mathcal{O}(N^{2,807})$ (Strassen Algorithmus) skalieren [92].

Die Begrenzung der Modenanzahl schränkt die Periode der Größe des zu untersuchenden Gitters ein, da die Schrittweite des Samplings im Richtungsraum durch die Periodengröße definiert ist (vgl. Gleichung 3.77). Andererseits ist eine hohe numerische Apertur der Simulation $\text{NA}_{\text{Sim}} = \frac{\lambda \text{MO}_x}{P_x}$ meist größer 1 erwünscht.

Beispiel Für eine Simulation mit $\text{NA}_{\text{Sim}} = 1$ und $\text{MO}_x = \text{MO}_y = 30$ Moden bei einer Wellenlänge von $\lambda = 0,5 \mu\text{m}$ ist die maximale Periode des Gitters gegeben durch $P_x = \frac{\lambda \text{MO}_x}{\text{NA}_{\text{Sim}}} = 15 \mu\text{m}$ gegeben.

Für die Simulation diffraktiver optischer Elemente eignet sich insbesondere die „localized input field RCWA“ (LIF-RCWA) [3]. Neben der Eigenschaft, dass diese RCWA Variante strukturierte Eingangslichtverteilungen simulieren kann, ist das Besondere an dieser RCWA Variante, dass die Felder als diskret angenommen werden und sich somit der Frequenzraum periodisch wiederholt. Dies hat zyklische Größen für die Permittivität und die Permeabilität zur Folge. Gleichzeitig ist in der LIF-RCWA die Abtastung des Gitters an die Modenzahl gekoppelt, somit können nur Gitter mit einer ungeraden Anzahl an Abtastpunkten simuliert werden.

Rekonstruiert man das Gitter an den Abtastpunkten erhält man wieder die ursprüngliche modenbegrenzte Verteilung. Wenn man von diesem „LIF-RCWA Sampling“ abweicht, dann ist nicht sichergestellt, dass die Simulationsergebnisse in sich konsistent sind [1]. Deshalb werden im folgenden Abschnitt nicht die diffraktiven Elemente DOE48 und DOE75 simuliert, die eine gerade Anzahl an Abtastpunkten haben, sondern es werden Elemente untersucht, die sich durch eine ungerade Anzahl an Abtastpunkten von den Elementen DOE48 und DOE75 unterscheiden.

9.2 RCWA Simulation von Linsen mit überlappender Apertur

In diesem Abschnitt werden RCWA Simulationen von diffraktiven Linsen mit überlappenden Aperturen vorgestellt und mit skalaren Simulationen verglichen. Zuerst wird jedoch der Übergang von der skalaren Transmissionsfunktion des dünnen Elements hin zum dicken Volumenelement vorgestellt. Anschließend werden für niedrige, mittlere und hohe numerische Aperturen die Fokalprofile aus der RCWA Simulation mit einer korrespondierenden skalaren Simulation verglichen. Abschließend wird diskutiert, bis zu welcher numerischen Apertur eine skalare Betrachtung mit dünner Element Näherung für die diffraktiven Linsen mit überlappenden Aperturen valide ist, wenn das diffraktive Elemente die Dicke der Wellenlänge hat.

9.2.1 Übergang vom dünnen Element zum dicken Element

Der Übergang von der skalaren Transmissionsfunktion in dünner Element Näherung zu einem Element der Dicke d lässt sich erreichen, indem man die Amplitude und Phase des dünnen Elements in eine komplexe Brechungsindexverteilung umrechnet:

$$t_{\text{DOE}}(\mathbf{r}_{\perp}) = A(\mathbf{r}_{\perp})e^{i\phi(\mathbf{r}_{\perp})} \quad (9.1)$$

$$= e^{ik_0n(\mathbf{r}_{\perp})d} \quad (9.2)$$

$$= e^{-\kappa(\mathbf{r}_{\perp})k_0d}e^{ik_0n_r(\mathbf{r}_{\perp})d}. \quad (9.3)$$

Hierbei wird mit n_r der Realteil und mit κ der Imaginärteil des Brechungsindex n bezeichnet. Löst man nun nach Real- und Imaginärteil des Brechungsindex auf, erhält man:

$$\kappa(\mathbf{r}_{\perp}) = \frac{1}{k_0d} \log\left(\frac{1}{A(\mathbf{r}_{\perp})}\right) \quad (9.4)$$

$$n_r(\mathbf{r}_{\perp}) = \frac{\phi(\mathbf{r}_{\perp})}{k_0d}. \quad (9.5)$$

Da für die Transmissionsfunktion nur eine relative Phasendifferenz relevant ist, lässt sich durch einen Offset in der Phase der minimale Wert des Realteils des Brechungsindex auf den Wert 1 schieben:

$$n_r(\mathbf{r}_{\perp}) = 1 + \frac{\phi(\mathbf{r}_{\perp}) - \min[\phi(\mathbf{r}_{\perp})]}{k_0d}. \quad (9.6)$$

Für $\mu = 1$ lässt sich der komplexe Brechungsindex in eine komplexe Epsilon-Verteilung umrechnen:

$$\epsilon(\mathbf{r}_{\perp}) = \left(n_r^2(\mathbf{r}_{\perp}) - \kappa^2(\mathbf{r}_{\perp})\right) + i2n_r(\mathbf{r}_{\perp})\kappa(\mathbf{r}_{\perp}). \quad (9.7)$$

Mit dieser Epsilon-Verteilung der Dicke d lässt sich nun näherungsweise die Funktion eines skalar designten diffraktiven Element in einer RCWA Simulation untersuchen.

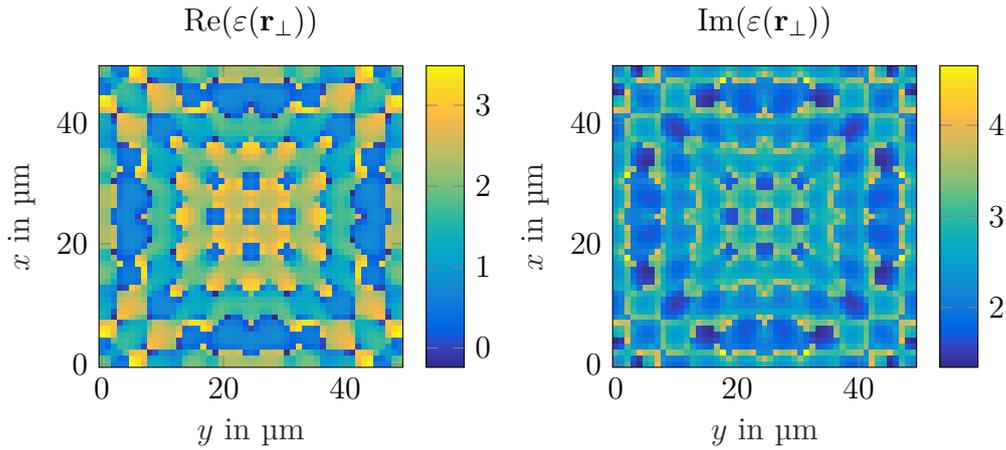


Abbildung 9.1: Epsilon-Verteilung berechnet aus der Phasenverteilung eines dünnen Elements mit numerischer Apertur $NA = 0,1$. Für die Beschreibung wurde $MO_x \times MO_y = 25 \times 25$ gewählt.

Eine skalare Transmissionsfunktion, die nur eine Phasenmodulation und keine Amplitudenmodulation bewirkt, lässt sich durch ein dickes Element mit einer rein reellen Epsilon-Verteilung beschreiben.

9.2.2 Simulation von dicken diffraktiven Elementen

Ähnlich wie im Abschnitt 5.2 bei der skalaren Simulation werden im Folgenden Elemente mit kontinuierlicher Amplituden- und Phasenverteilung sowie amplitudennivellierte und phasenbinarisierte Elemente für unterschiedliche numerische Aperturen untersucht.

Diffraktive Linsen mit niedriger numerischer Apertur

Zuerst werden ein kontinuierliches Element sowie ein modifiziertes Element, bei dem die Phase binarisiert und die Amplitude nivelliert wurde, mit einer numerischen Apertur von 0,1 untersucht. Die Periode beträgt jeweils $P_x = P_y = 50 \mu\text{m}$. Die Design- und Ausbreitungswellenlänge ist $\lambda = 0,532 \mu\text{m}$. Der Fokus liegt bei $f = 500 \mu\text{m}$. Die Dicke der Elemente entspricht der Wellenlänge $d = \lambda$. Die skalare Transmissionsfunktion hat $N_x = N_y = 51$ Pixel, in der RCWA lässt sich dieses Element durch die Verwendung von $MO_x \times MO_y = 25 \times 25$ Moden beschreiben. Somit ist die numerische Apertur der Simulation $NA_{\text{Sim}} \approx 0,25$.

Kontinuierliches Element In Abbildung 9.1 sieht man ein Beispiel für die Epsilon-Verteilung eines kontinuierlichen Elements mit einer numerischen Apertur von $NA = 0,1$. Der Realteil der Epsilon-Verteilung besitzt negative Anteile und der Imaginärteil der Epsilon-Verteilung ist relativ groß.

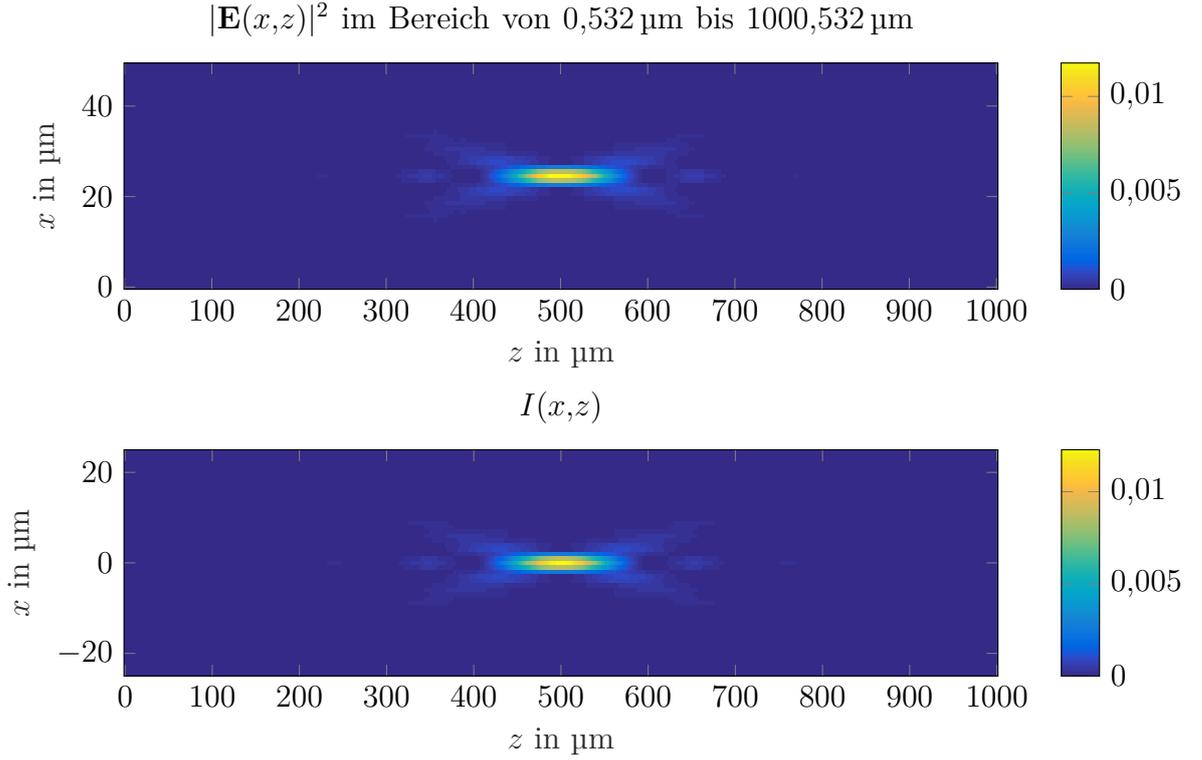


Abbildung 9.2: Intensitätsverteilung hinter dem diffraktiven Element aus Abbildung 9.1. Oben: Berechnung der Intensität mit der RCWA mit $\text{MO}_x = \text{MO}_y = 25$. Unten: Berechnung der Intensität mit der skalaren ebenen Wellen Zerlegung.

In der RCWA Simulation entspricht das Absolutquadrat des E-Felds der Intensität. Für einen xz -Schnitt des Fokus ist in Abbildung 9.2 die zugehörige Intensitätsverteilung gezeigt. Zum Vergleich wird auch die Intensitätsverteilung der skalaren Rechnung mit dünner Element Näherung und Ausbreitung mit der ebenen Wellen Zerlegung bei gleichem Sampling gezeigt. Die Fokalprofile der RCWA Simulation und der skalaren Rechnung stimmen sehr gut qualitativ und quantitativ überein.

Betrachtet man den Linescan durch den Fokus, so sieht man eine kleine Abweichung im Fokus (siehe Abbildung 9.3). Vergleicht man die maximale Intensität des skalaren Fokusprofils mit der maximalen Intensität der RCWA Rechnung, so erhält man ein Verhältnis von $\frac{I_{\max}^{\text{RCWA}}}{I_{\max}^{\text{AS}}} = \frac{0,117}{0,0123} = 0,95$.

Um die Abweichung genauer zu quantifizieren wird in Abbildung 9.4 die Differenz zwischen der Intensität aus RCWA Rechnung und skalarer Rechnung gebildet. Die Amplitude der Differenz ist deutlich geringer als die typische Amplitude des Beugungsmusters.

Die geringe Abweichung in der Intensität ist erstaunlich, wenn man die globale Transmission $\mathcal{T}_{\text{TEA}} = \frac{1}{P_x P_y} \sum_{m,n} |t_{mn}|^2 dx dy = 0,0000443$ und globale Absorption $\mathcal{A}_{\text{TEA}} = 1 - \mathcal{T}_{\text{TEA}} = 0,9999557$ des abgetasteten dünnen Elements mit

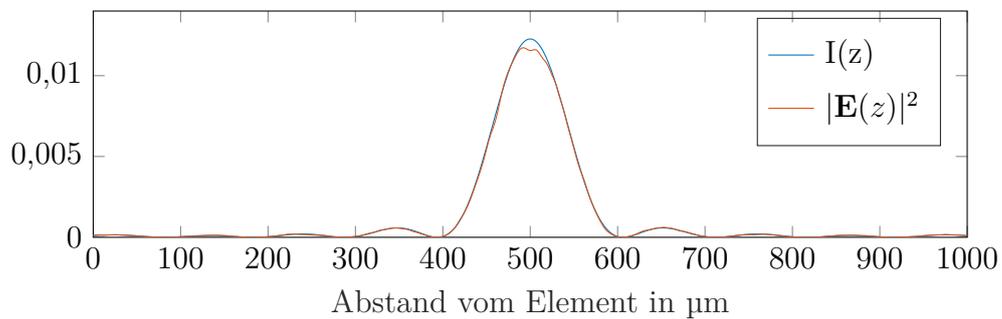


Abbildung 9.3: Intensität auf der optischen Achse durch den Fokus eines kontinuierlichen diffraktiven Elements mit numerischer Apertur von 0,1 für die RCWA Rechnung (rot) und die skalare Simulation (blau).

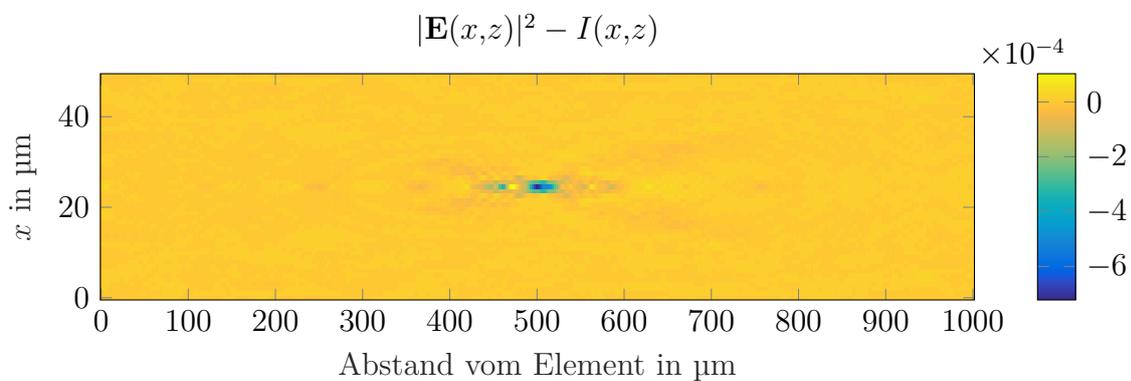


Abbildung 9.4: Abweichung zwischen der RCWA Rechnung und der skalaren Rechnung.

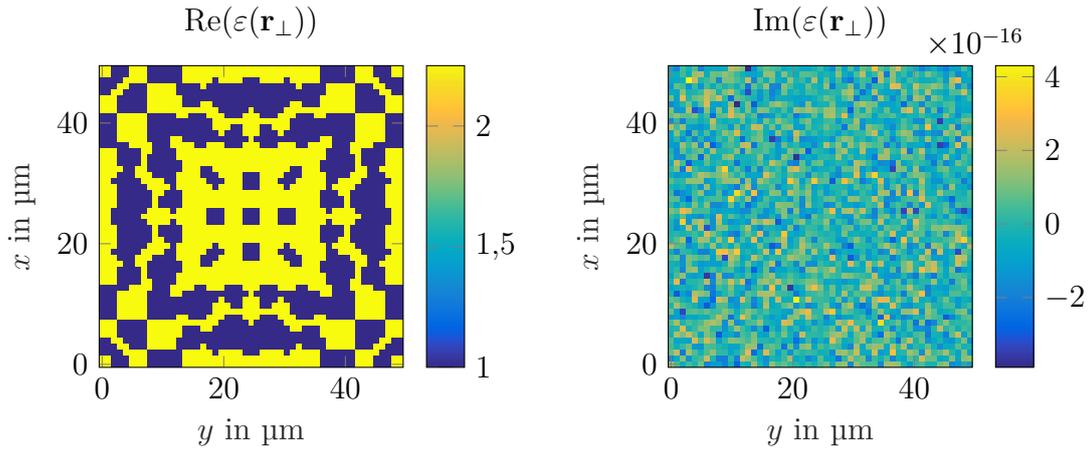


Abbildung 9.5: Rekonstruktion der Epsilon-Verteilung berechnet aus einer Transmissionsfunktion, bei der die Phase binarisiert und die Amplitude auf Eins gesetzt wurde. Für die Beschreibung wurde $MO_x \times MO_y = 25 \times 25$ gewählt.

den Werten aus der RCWA vergleicht. In der RCWA ist die globale Transmission $\mathcal{T} = 0,000\,043\,5$, die globale Reflexion $\mathcal{R} = 0,149\,990\,0$ und die globale Absorption $\mathcal{A} = 0,849\,966\,5$. Man sieht die globale Transmission stimmt gut überein, obwohl die Reflexion in RCWA Rechnung einen relativ starken Beitrag hat. Betrachtet man das Verhältnis der globalen Transmission aus der skalaren und der RCWA Rechnung erhält man $\frac{\mathcal{T}}{\mathcal{T}_{\text{TEA}}} = 0,98$. Dies Verhältnis ist etwas größer als das Verhältnis der maximalen Intensitäten aber in derselben Größenordnung.

Modifiziertes Element In Abbildung 9.5 sieht man die Rekonstruktion der Epsilon-Verteilung eines phasenbinarisierten und amplitudennivellierten Elements. Die Parameter entsprechen denen des kontinuierlichen Elements. Durch die Amplitudennivellierung erhält man nur positive Epsilon-Werte und der Imaginärteil der Epsilon-Verteilung ist im Rahmen der numerischen Genauigkeit null. Durch die Phasenbinarisierung erhält man nur zwei Werte für den Realteil der Epsilon-Verteilung. Der niedrigere Wert entspricht der relativen Permittivität des Vakuums, somit beschreibt die erzeugte Verteilung ein binäres Gitter.

Die RCWA Simulation liefert einen xz-Schnitt durch den Fokus, dieser ist in Abbildung 9.6 gezeigt. Zum Vergleich ist auch die skalare Simulation des Fokalprofils gezeigt.

Die Intensität der beiden Simulationen stimmt sehr gut überein, dies wird auch insbesondere in Abbildung 9.7 beim Vergleich der Linescans auf der optischen Achse durch den Fokus deutlich. Die gute Übereinstimmung kommt unter anderem daher, dass die globale Transmission $\mathcal{T} = 0,9998$ nahezu Eins und die globale Reflexion $\mathcal{R} = 0,0002$ vernachlässigbar ist.

Betrachtet man die Differenz zwischen RCWA Rechnung und skalarer Rechnung

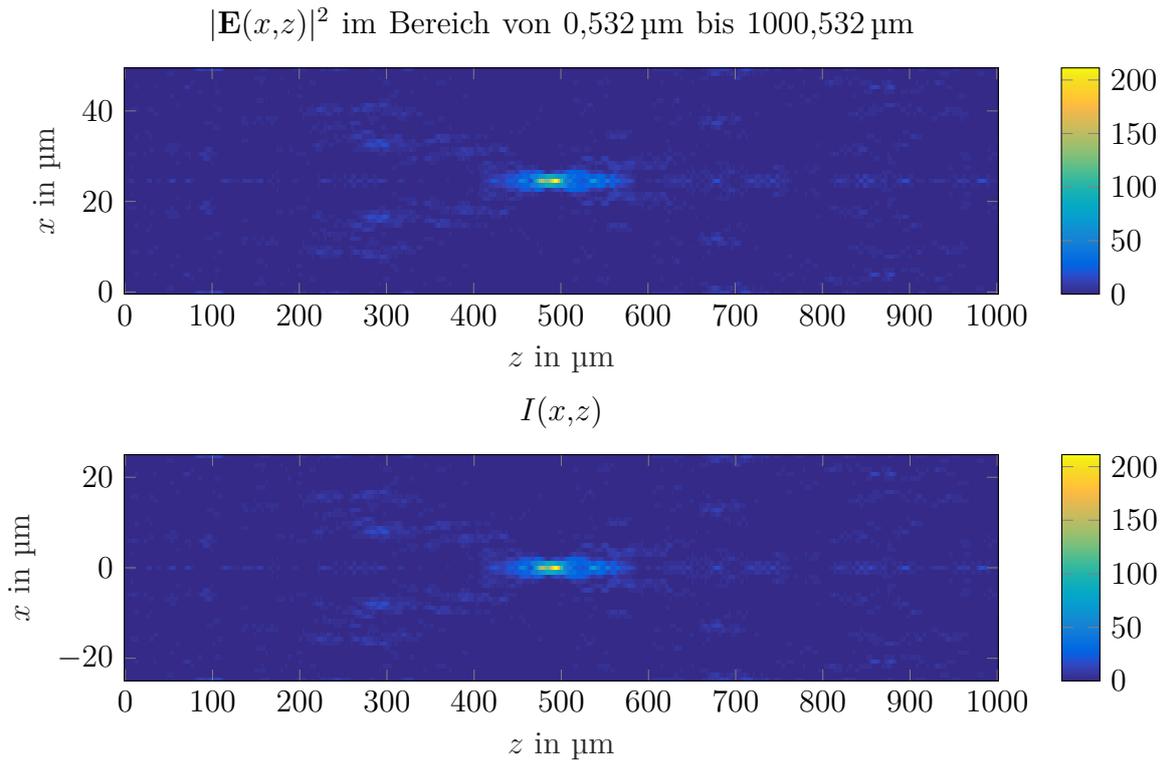


Abbildung 9.6: Intensitätsverteilung hinter einem phasenbinarisierten diffraktiven Element mit einer numerischen Apertur von $NA = 0,1$ berechnet mit RCWA und $MO_x = MO_y = 25$ (oben) und skalar mit der ebenen Wellenzerlegung (unten).

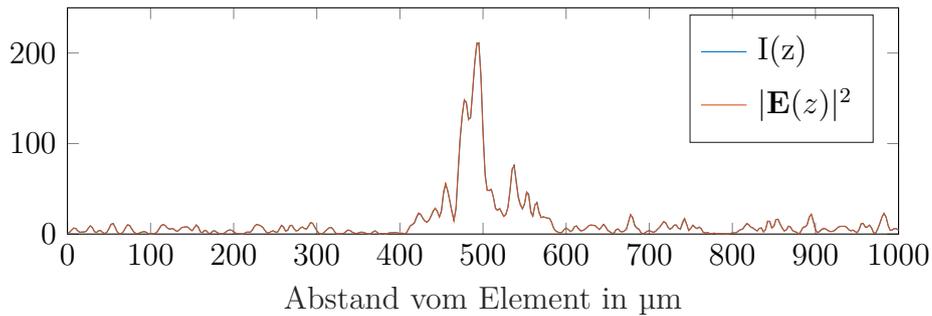


Abbildung 9.7: Intensität auf der optischen Achse durch den Fokus eines phasenbinarisierten und amplitudennivellierten diffraktiven Elements mit numerischer Apertur von $0,1$ für die RCWA Rechnung (rot) und die skalare Simulation (blau).

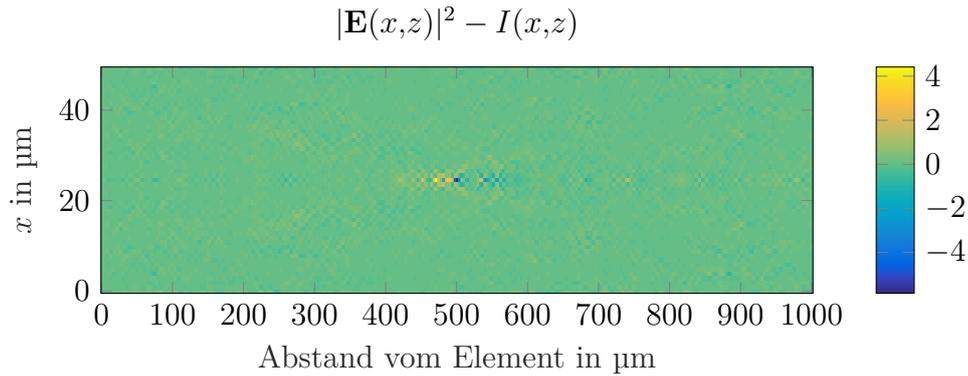


Abbildung 9.8: Abweichung zwischen der RCWA Rechnung und der skalaren Rechnung.

in Abbildung 9.8 so sieht man, dass auch hier die Amplitude der Modulation deutlich geringer ist als die typische Amplitude des Beugungsmusters.

Fazit für diffraktive Linsen mit niedriger numerischer Apertur Für die diffraktiven Elemente mit niedriger numerischer Apertur lässt sich festhalten, dass die skalare Rechnung sehr gut mit der RCWA Rechnung übereinstimmt. Dies bedeutet, dass die skalare Rechnung mit dünner Element Näherung für diesen Fall valide ist.

Diffraktive Linsen mit mittlerer numerischer Apertur

In diesem Abschnitt werden zwei diffraktive Linsenarrays mit einer mittleren numerischen Apertur von $NA = 0.48$ untersucht. Es wird wieder eine Element mit kontinuierlicher Amplituden und Phasenverteilung untersucht und ein Element, das phasenbinarisiert und amplitudennivelliert wurde. Beide Elemente haben die Dicke $d = \lambda$ und die Periode beträgt jeweils $P_x = P_y = 50 \mu\text{m}$. Die Design und Ausbreitungswellenlänge ist $\lambda = 0,532 \mu\text{m}$ und der Fokus liegt bei $f = 500 \mu\text{m}$. Die 101×101 Pixel des Elements werden in der Simulation mit $MO_x \times MO_y = 50 \times 50$ Moden beschrieben. Dadurch ist die maximale numerische Apertur der RCWA Simulation $NA_{\text{Sim}} \approx 0.5$.

Kontinuierliches Element In Abbildung 9.9 sieht man die Epsilon-Verteilung des kontinuierlichen Elements mit einer numerischen Apertur von $NA = 0,48$.

Die zugehörige Intensität im xz-Schnitt des Fokusprofils aus der RCWA Simulation ist in Abbildung 9.10 gezeigt. Zum Vergleich ist auch die Intensität aus der skalaren Rechnung mit der dünnen Element Näherung und der Ausbreitung mittels der Zerlegung nach ebenen Wellen bei gleichem Sampling gezeigt. Man sieht, dass die Geometrie der Fokalprofile gut übereinstimmen, die Intensität ist jedoch in der RCWA Simulation geringer.

Für einen direkten Vergleich ist in Abbildung 9.11 jeweils ein Linescan durch den Fokus der RCWA Simulation und der skalaren Rechnung gezeigt. Die maxi-

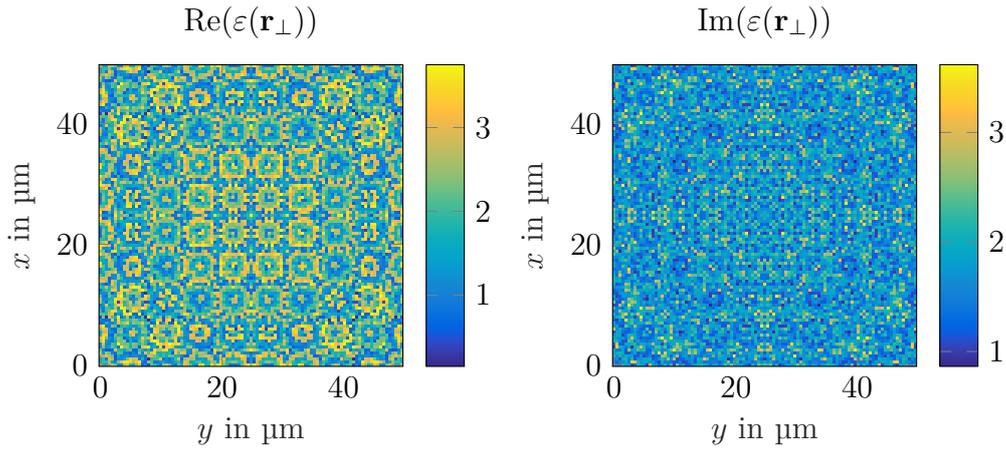


Abbildung 9.9: Epsilon-Verteilung erzeugt aus der Phasenverteilung eines dünnen Elements mit numerischer Apertur $NA = 0,48$. Für die Beschreibung wurde $MO_x \times MO_y = 50 \times 50$ gewählt.

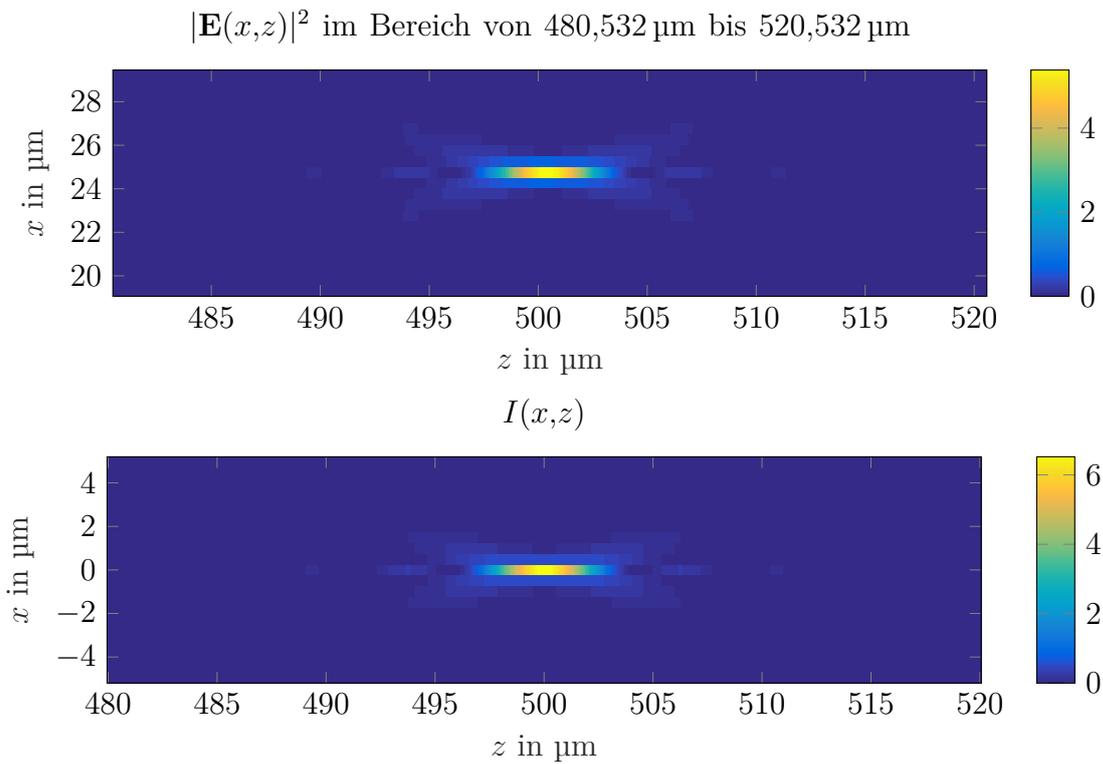


Abbildung 9.10: Intensitätsverteilung hinter dem diffraktiven Element aus Abbildung 9.13. Oben: Berechnung der Intensität mit der RCWA mit $MO_x = MO_y = 50$. Unten: Berechnung der Intensität mit der skalaren ebenen Wellen Zerlegung.

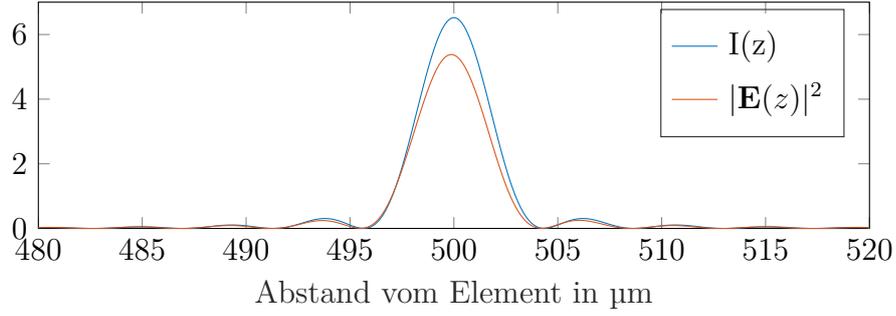


Abbildung 9.11: Intensität auf der optischen Achse durch den Fokus eines kontinuierlichen diffraktiven Elements mit numerischer Apertur von 0,48 für die RCWA Rechnung (rot) und die skalare Simulation (blau).

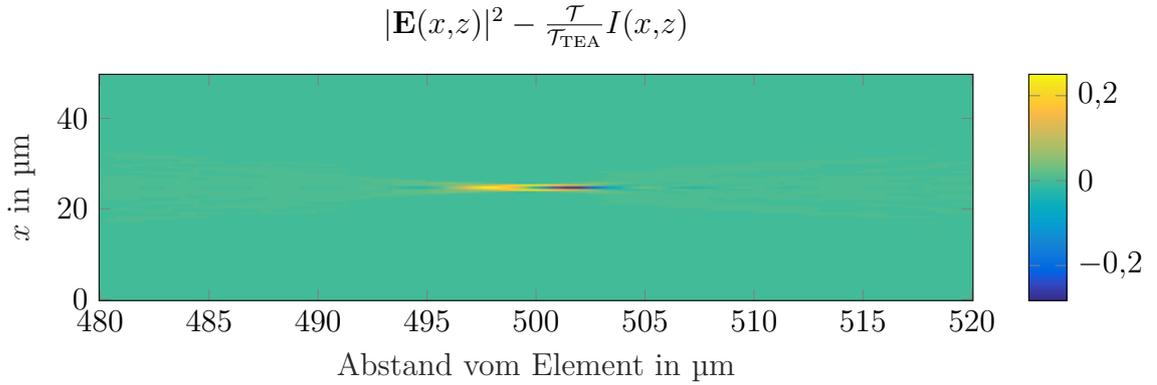


Abbildung 9.12: Abweichung zwischen der RCWA Rechnung und der skalaren Rechnung. Für einen fairen Vergleich wurde die Intensität der skalaren Rechnung auf die gleiche Transmission hinter dem diffraktiven Element wie in der RCWA Rechnung skaliert.

malen Intensitäten unterscheiden sich um einen Faktor $\frac{I_{\text{max}}^{\text{RCWA}}}{I_{\text{max}}^{\text{AS}}} = \frac{5,3812}{6,5229} = 0,825$. Außerdem sieht man, dass der Fokus in der RCWA Rechnung minimal um $\Delta z = (0,125 \pm 0,025) \mu\text{m}$ vor dem Fokus der skalaren Rechnung liegt.

Die RCWA liefert für die globale Transmission $\mathcal{T} = 0,00085$, die globale Reflexion $\mathcal{R} = 0,10046$ und die globale Absorption $\mathcal{A} = 0,89869$. Aus der skalaren Rechnung folgt für die globale Transmission $\mathcal{T}_{\text{TEA}} = 0,00102$. Hieraus folgt für das Verhältnis $\frac{\mathcal{T}}{\mathcal{T}_{\text{TEA}}} = 0,834$. Man sieht, dass die Abweichung in den maximalen Intensitäten ungefähr der Abweichung in der Transmission entspricht.

Um die Abweichungen im Beugungsmuster zu verdeutlichen, wird in Abbildung 9.12 zuerst die skalare Intensität mit dem Verhältnis der Transmission skaliert und dann wird die Differenz zwischen der skalierten skalaren Rechnung und RCWA Rechnung gebildet. Auf der optischen Achse sieht man zuerst ein Maximum, dann ein Minimum. Dies spiegelt die Tatsache wieder, dass der RCWA Fokus leicht vor dem skalaren Fokus liegt.

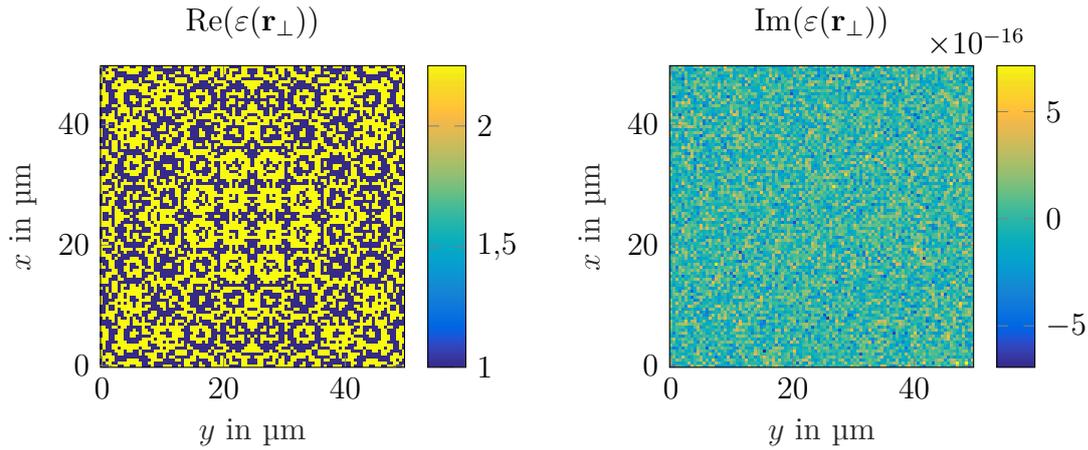


Abbildung 9.13: Rekonstruktion der Epsilon-Verteilung berechnet aus einer Transmissionsfunktion, bei der die Phase binarisiert und die Amplitude auf Eins gesetzt wurde. Für die Beschreibung wurde $MO_x \times MO_y = 50 \times 50$ gewählt.

Modifiziertes Element In Abbildung 9.13 sieht man die Rekonstruktion der Epsilon-Verteilung eines phasenbinarisierten und amplitudennivellierten Elements. Die Parameter entsprechen denen des kontinuierlichen Elements. Der Realteil der Permittivität ist binär, der Imaginärteil ist bis auf numerisches Rauschen null.

Die RCWA Simulation des zugehörige Fokalverlaufs ist in Abbildung 9.14 gezeigt. Zum Vergleich ist auch die skalare Simulation des Fokalprofils gezeigt. Die Intensität der beiden Simulationen stimmt sowohl qualitativ als auch quantitativ sehr gut überein.

In Abbildung 9.15 sieht man, dass auch hier der Fokus in der RCWA Simulation vor dem Fokus aus der skalaren Rechnung liegt. Die axiale Abweichung des Fokus entspricht $\Delta z = (0,275 \pm 0,025) \mu\text{m}$, also ca. der Hälfte der Wellenlänge bzw. der Hälfte der Dicke des diffraktiven Elements. Das Verhältnis der Maximal-Intensitäten beträgt $\frac{I_{RCWA}}{I_{AS}} = \frac{1879,9}{1890,2} = 0,9946$.

In Abbildung 9.16 sieht man die Abweichung zwischen der skalaren und der RCWA Simulation.

Die globale Transmission $\mathcal{T} = 0,9928$ ist nahezu Eins, die globale Reflexion ist $\mathcal{R} = 0,0072$ und damit vernachlässigbar. Auch in diesem Beispiel entspricht die Abweichung in der maximalen Intensität ungefähr der globalen Transmission.

Fazit für diffraktive Linsen mit mittlerer numerischer Apertur Für die beiden diffraktiven Elemente mit numerischer Apertur von $NA = 0,48$ ist die skalare Rechnung und die dünne Element Näherung sehr valide. Die Fokusgeometrie ist in beiden Simulationen sehr ähnlich. Der Intensitätsunterschied zwischen skalarer Rechnung und RCWA ist durch eine geringere Transmission in der RCWA Rechnung zu erklären, der durch die Berücksichtigung von Polarisation entsteht. Außerdem ist eine

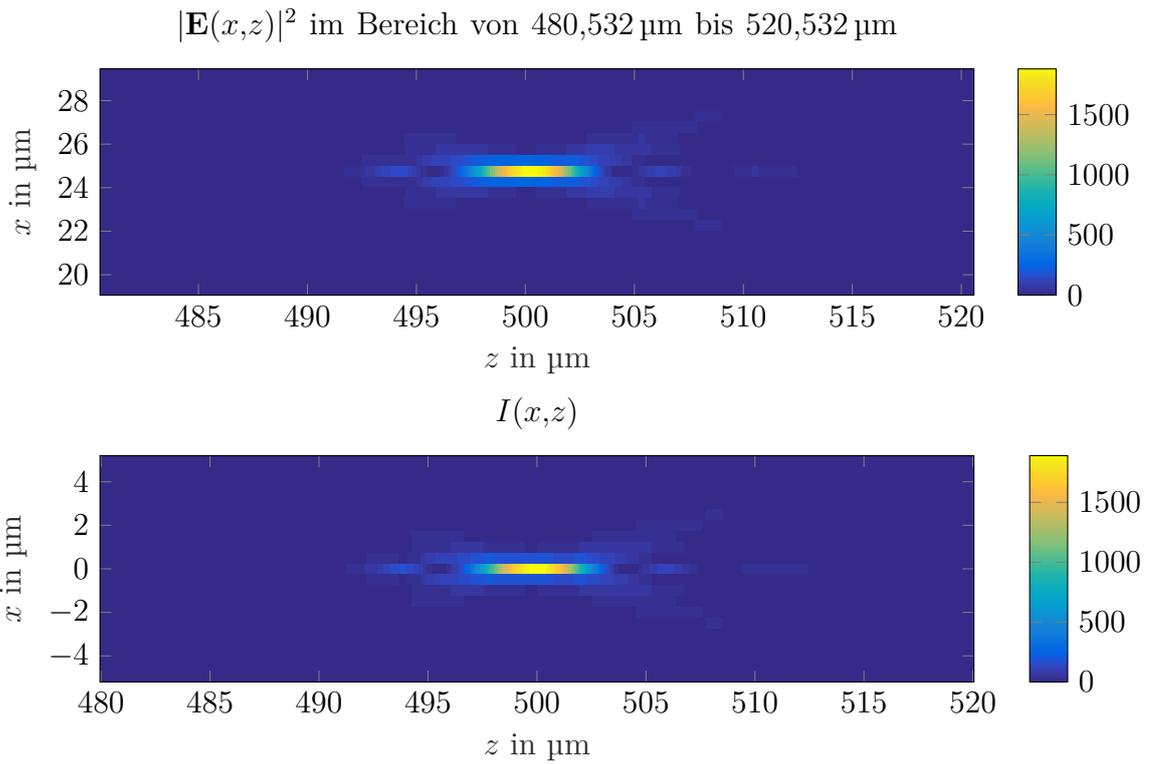


Abbildung 9.14: Intensitätsverteilung hinter einem phasenbinarisierte und amplitudennivellierten Element berechnet mit RCWA und $MO_x = MO_y = 50$ (oben) und skalar mit der Zerlegung nach ebenen Wellen (unten).

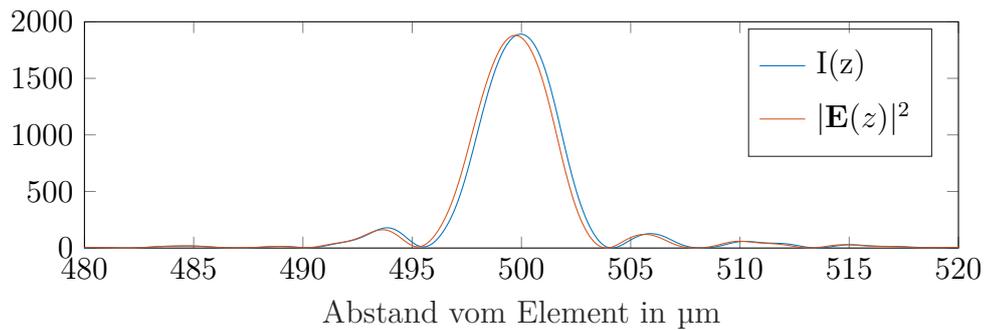


Abbildung 9.15: Intensität auf der optischen Achse durch den Fokus eines phasenbinarisierten und amplitudennivellierten diffraktiven Elements mit numerischer Apertur von 0,48 für die RCWA Rechnung (rot) und die skalare Simulation (blau).

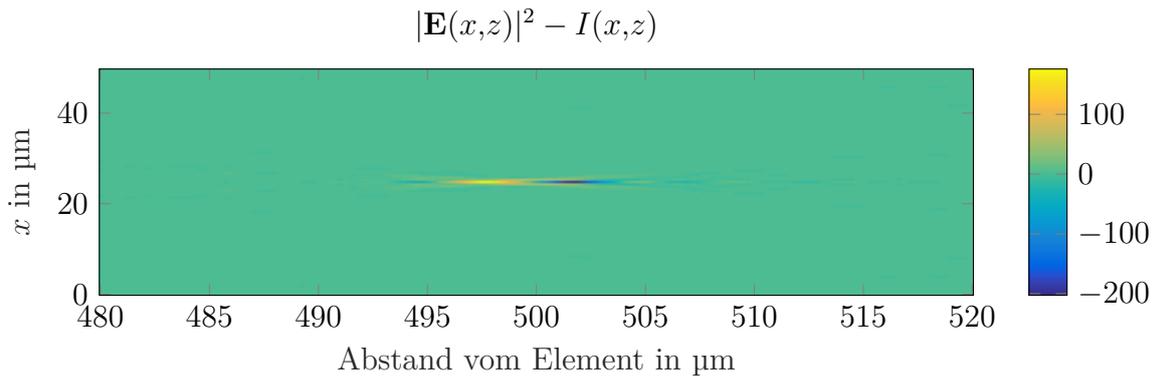


Abbildung 9.16: Abweichung zwischen der RCWA Rechnung und der skalaren Rechnung.

Fokusverschiebung hin zu dem diffraktiven Element beobachtbar, diese Distanz der Verschiebung ist jedoch nicht einheitlich beim kontinuierlichen und modifizierten Element.

Fazit

Anhand mehrerer Beispiele wurde gezeigt, dass die skalare Simulation von diffraktiven Linsen mit überlappenden Aperturen bis zu einer numerischen Apertur von $NA = 0,48$ noch vertrauenswürdig ist und die skalare Näherung und die dünne Element Näherung valide sind. Da die Parameter der simulierten diffraktiven Linsen mit mittlerer numerischer Apertur ungefähr den Parametern des DOE48 entsprechen, ist davon auszugehen, dass auch die skalaren Simulationen zu dem fabrizierten Element noch valide sind.

Die Überprüfung der skalaren Simulation für ein diffraktives Element mit ähnlichen Parameter wie vom DOE75 ist aufgrund der Kombination von hoher numerischer Apertur und großer Periode mit der RCWA nicht möglich, da der Speicherbedarf für die Simulation sehr groß ist. Außerdem wird hier auch die Rechenzeit zu einem großen Problem.

Im folgenden Abschnitt wird deshalb anhand von Beispielen mit kleinerer Periodengröße der Gültigkeitsbereich der skalaren Rechnung untersucht.

9.2.3 Wie lange ist die skalare Rechnung für diffraktive Linsen valide?

Um den Gültigkeitsbereich der skalaren Rechnung für diffraktive Linsen zu klären, werden zuerst ein paar Literaturmeinungen zu diesem Thema vorgestellt. Anschließend wird anhand diffraktiver Linsen mit sehr hohen numerischen Aperturen Abweichungen zwischen skalarer und RCWA Rechnung aufgezeigt. Anhand von qualitativen Vergleichen des axialen Fokalprofils der RCWA Rechnung und der skalaren

Rechnung wird dann die Gültigkeit der skalaren Rechnung für die diffraktive Linsen abgeschätzt.

Literaturmeinungen zur Gültigkeit der skalaren Theorie

In der Literatur wurde die Gültigkeit der skalaren Rechnung an verschiedenen Beispielsystemen wie Phasengittern [50], Blazegittern [83] und diffraktive Strahlteilern [70, 71] untersucht und mit rigorosen Rechnungen wie der FDTD und RCWA verglichen. Je nach System kommen die Autoren zu unterschiedlichen Gültigkeitskriterien der skalaren Rechnung.

Pommet, Moharam und Grann [83] finden bei binären Gittern signifikante Fehler, wenn die Gitterperiode $\Lambda < 14\lambda$ ist. Wenn beim Gitter der Füllfaktor 0,5 beträgt, werden die Fehler allerdings minimal, sodass für Gitterperioden von bis zu $\Lambda = 2\lambda$ die skalare Rechnung akkurat ist. Jing und Jin [50] bestätigen den von Pommet, Moharam und Grann [83] gefundenen den Gültigkeitsbereich an binären Phasengittern mit einem Füllfaktor von 0,5.

Mellin und Nordin [71] finden, dass bei diskretisierten Phasenverteilungen die Zerlegung nach ebenen Wellen als valide angenommen werden kann, wenn die typische Strukturgröße, gegeben durch $\Lambda = \frac{\lambda}{\text{NA}}$, größer als das doppelte der Wellenlänge ist. Deshalb hält er die skalare Rechnung bis zu einer numerischen Apertur von 0,5 valide.

Abweichung zwischen skalarer Simulation und RCWA Simulation bei sehr hohen numerischen Aperturen

Geringe Abweichungen zwischen skalarer Simulation und RCWA Simulation wurden schon im Abschnitt 9.2.2 deutlich. Geht man zu sehr hohen numerischen Aperturen, so sind die Abweichungen zwischen skalarer und rigoroser Rechnung noch deutlicher. Dies wird hier anhand diffraktiver Linsen mit einer numerischen Apertur von $\text{NA} = 0,9$ veranschaulicht. Die Periode beträgt $P_x = P_y = 10 \mu\text{m}$, die Design und Ausbreitungswellenlänge ist $\lambda = 0,532 \mu\text{m}$ und der Fokus liegt bei $f = 500 \mu\text{m}$.

In Abbildung 9.17 ist das simulierte Fokalprofil eines kontinuierlichen diffraktiven Elements, sowohl aus der skalaren als auch aus der RCWA Simulation gezeigt.

Man sieht leichte Unterschiede im Fokalprofil, insbesondere im Bereich der ersten Nullstelle, die vor dem Fokus liegt. Die Abweichung wird auch beim Vergleich der beiden Linescans durch den Fokus in Abbildung 9.18 deutlich. Man sieht, dass die RCWA Rechnung nicht mehr symmetrisch bezüglich des Fokus ist. Insbesondere ist das Nebenmaximum vor dem Fokus stark unterdrückt. Außerdem erkennt man deutliche Unterschiede in der Lage der maximalen Intensität.

Analysiert man für das kontinuierliche Element die Transmission, so erhält man aus der RCWA für die globale Transmission $\mathcal{T} = 0,05977$. Aus der skalaren Rechnung folgt für die globale Transmission $\mathcal{T}_{\text{TEA}} = 0,0885$. Hieraus folgt für das Verhältnis $\frac{\mathcal{T}}{\mathcal{T}_{\text{TEA}}} = 0,6754$. Die maximalen Intensitäten unterscheiden sich jedoch durch

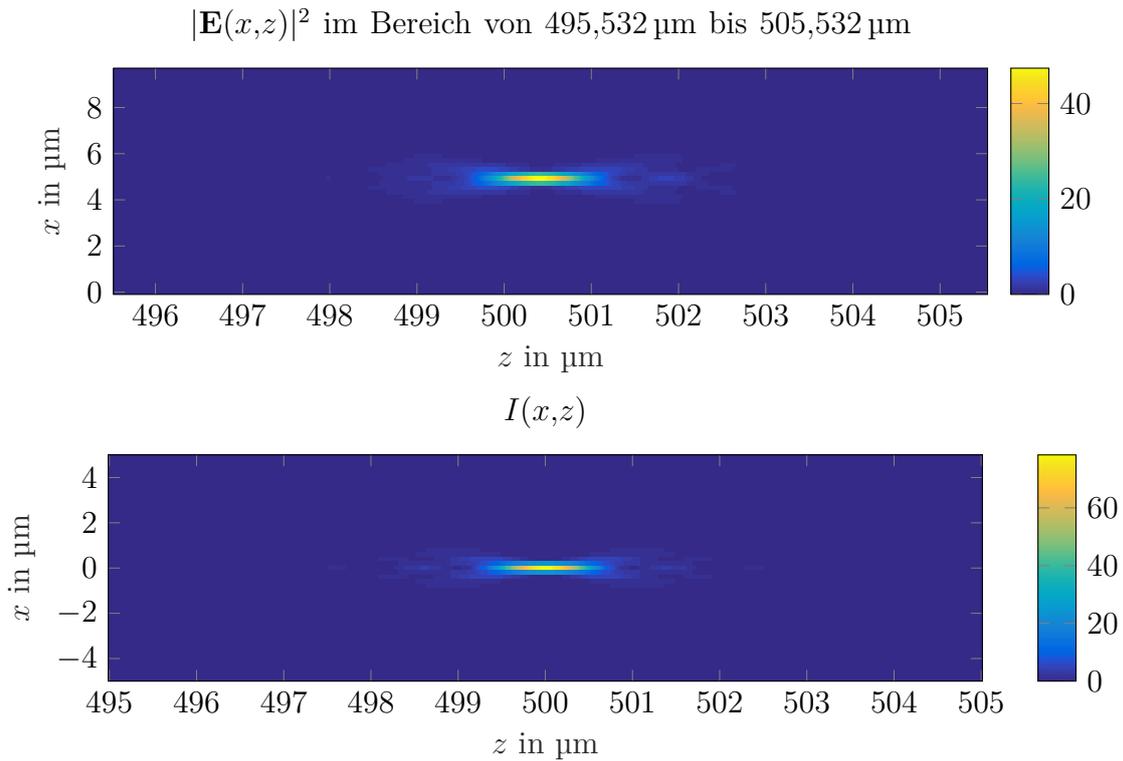


Abbildung 9.17: Intensitätsverteilung hinter einem phasenbinarisierte Element berechnet mit RCWA und $MO_x = MO_y = 25$ (oben) und skalar mit der ebenen Wellenzerlegung (unten).

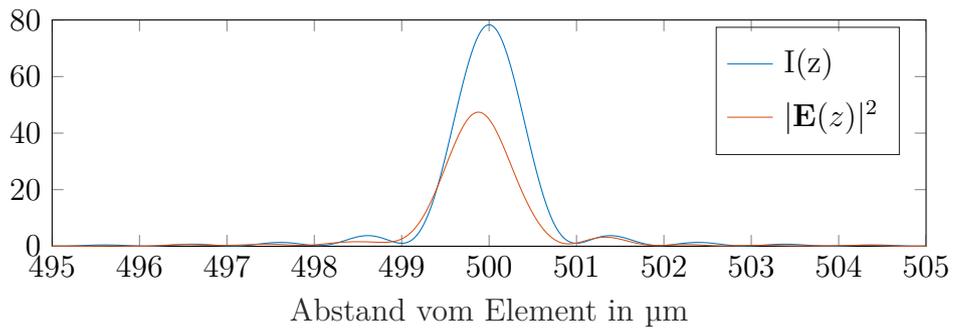


Abbildung 9.18: Intensität auf der optischen Achse durch den Fokus eines kontinuierlichen diffraktiven Elements mit numerischer Apertur von 0,9 für die RCWA Rechnung (rot) und die skalare Simulation (blau).

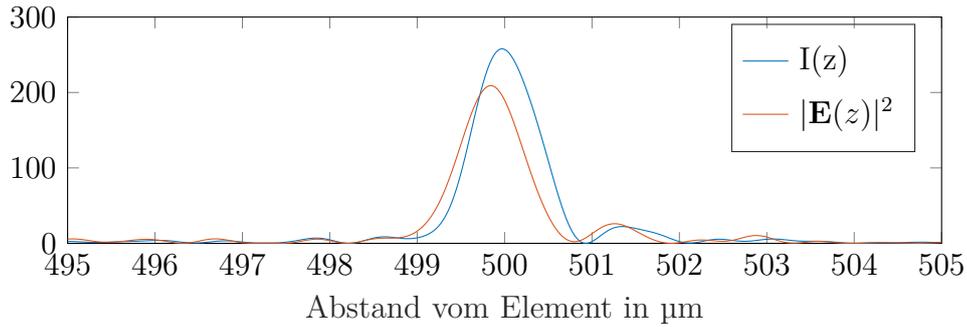


Abbildung 9.19: Intensität auf der optischen Achse durch den Fokus eines phasenbinarisierten und amplitudennivellierten diffraktiven Elements mit numerischer Apertur von 0,9. In rot ist das Ergebnis der RCWA Simulation gezeigt, in blau das der skalaren Simulation.

einen Faktor von $\frac{I_{\max}^{\text{RCWA}}}{I_{\max}^{\text{AS}}} = \frac{47,4224}{78,3225} = 0,6055$. Somit lässt sich die geringere Fokusintensität nicht alleine auf die geringere Transmission zurückführen.

Eine ähnliche Beobachtung lässt sich auch bei der Untersuchung eines phasenbinarisierten und amplitudennivellierten Elements mit sonst gleichen Parametern machen. Hierzu ist in Abbildung 9.19 wieder ein Linescan durch die den Fokus der RCWA Simulation und der skalaren Simulation gezeigt. Auch hier sieht man die Asymmetrie und eine deutliche Abweichung in der maximalen Intensität.

Fazit für diffraktive Linsen mit sehr hohen numerischen Aperturen Bei sehr hohen numerischen Aperturen erzeugen die diffraktiven Linsen auch in der rigorosen Rechnung noch einen Fokus, jedoch weichen sowohl Geometrie und Position als auch die maximale Intensität von der skalaren Simulation ab. Hierbei ist die Abweichung in der maximalen Intensität nicht allein durch die Abweichung in der Transmission zu erklären.

Systematischer Vergleich zur Abschätzung der Validität der skalaren Rechnung

Um die Validität der skalaren Rechnung abzuschätzen sind in Abbildung 9.20 und 9.21 die Intensitäten auf der optischen Achse durch den Fokus aus der skalaren Rechnung und der RCWA Rechnung für kontinuierliche Elemente, sowie phasenbinarisierte und amplitudennivellierte Elemente mit unterschiedlichen numerischen Aperturen gezeigt. Die Periode des Elements ist $P_x = P_y = 10 \mu\text{m}$ und es wurde eine Wellenlänge von $\lambda = 0,532 \mu\text{m}$ verwendet. Die Fokusslänge beträgt $f = 100 \mu\text{m}$. Die rigorose Simulation wurde mit $\text{MO}_x \times \text{MO}_y = 20 \times 20$ Moden durchgeführt, die maximale numerische Apertur der Simulation ist somit $\text{NA}_{\text{Simulation}} = \frac{P_x \text{MO}_x}{\lambda} = 1$.

Bei den Simulationen des kontinuierlichen Elements sieht man, dass die Unterschiede in der maximalen Intensität umso deutlicher werden, je höher die numerische

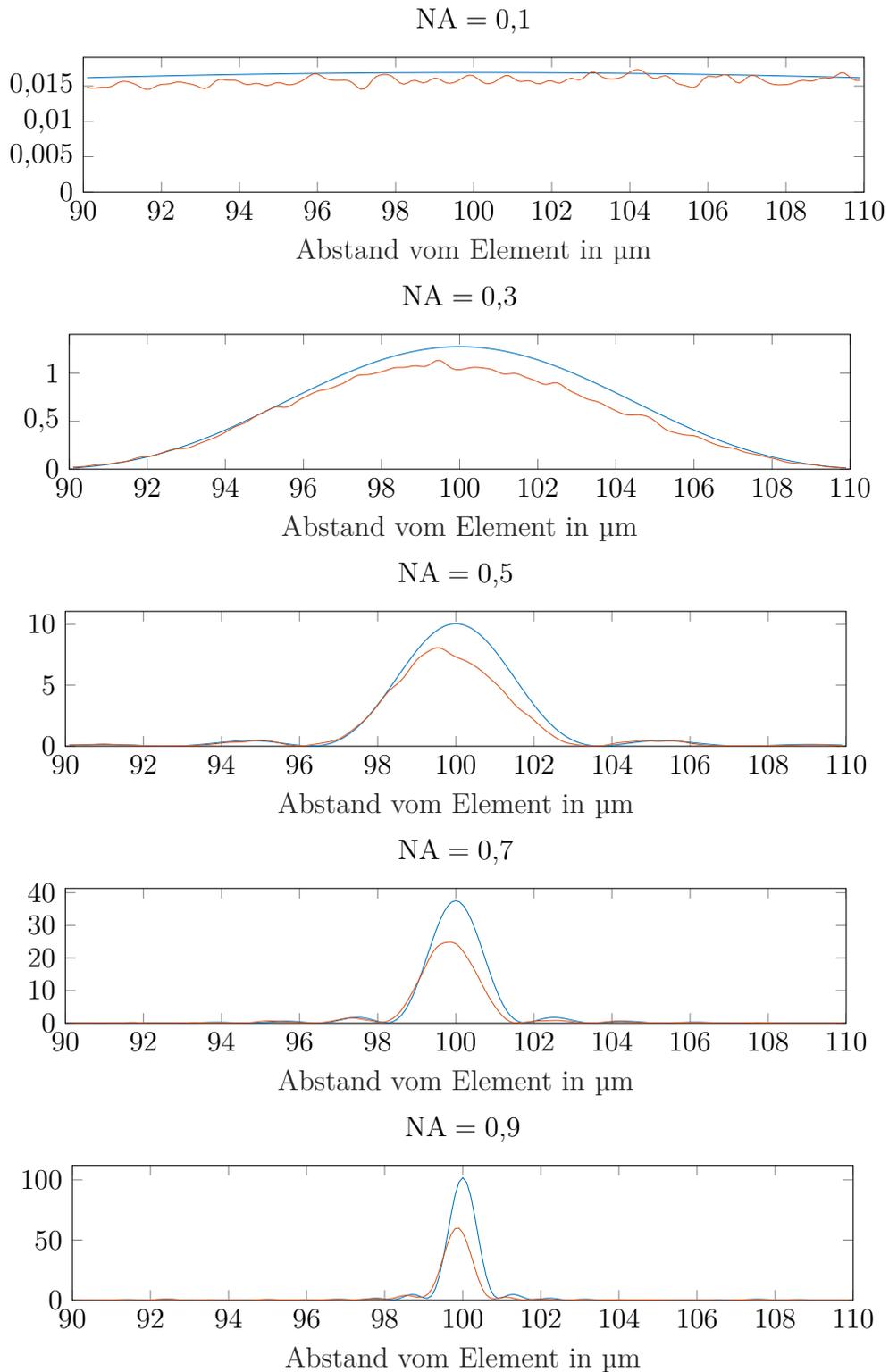


Abbildung 9.20: Intensität auf der optischen Achse durch den Fokus für kontinuierliche Elemente mit verschiedenen numerischen Aperturen. Blau entspricht der skalaren Rechnung und rot der RCWA Rechnung.

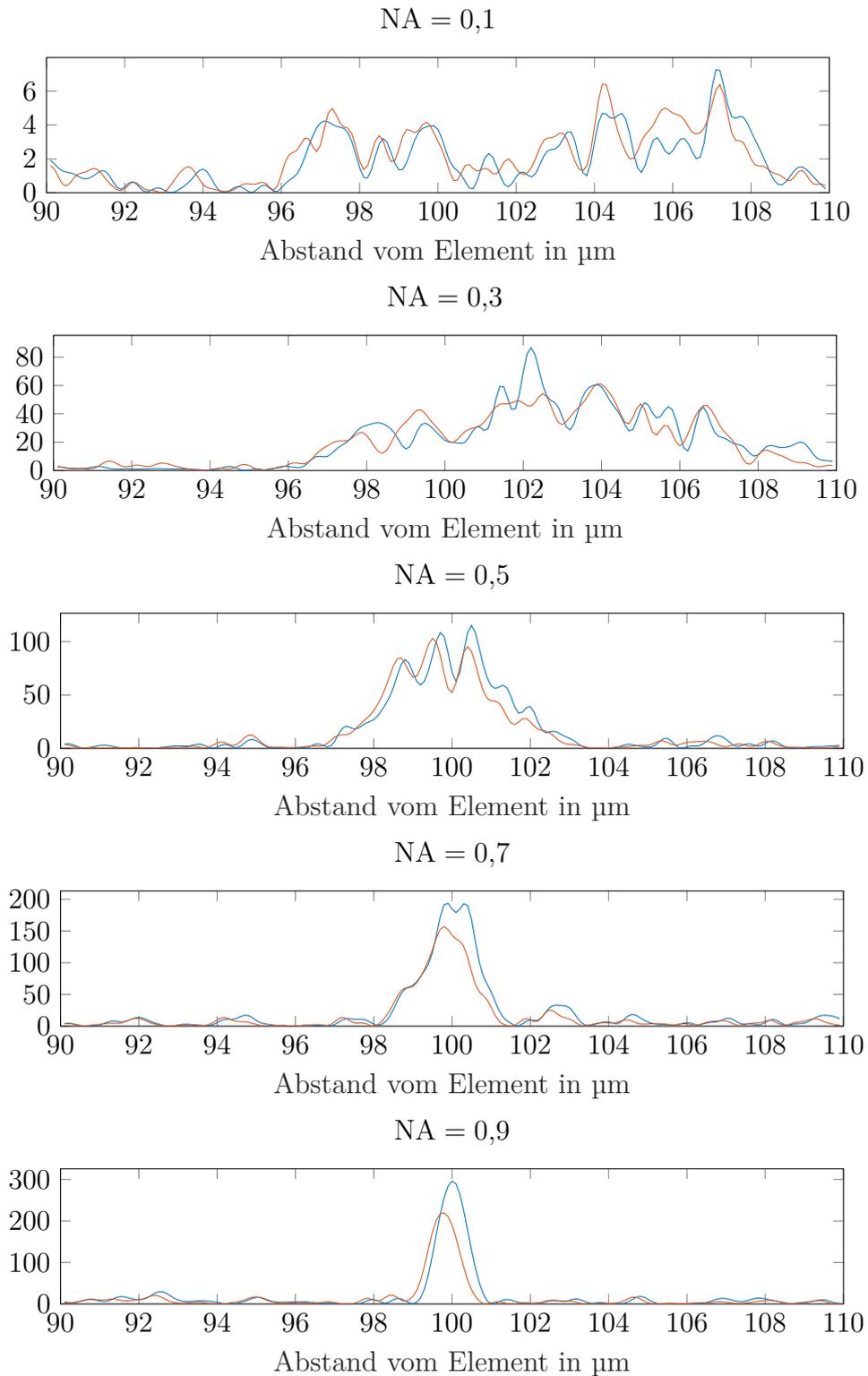


Abbildung 9.21: Linescan auf der optischen Achse durch den Fokus für phasenbina-
risierte und amplitudennivellierte Elemente mit verschiedenen nume-
rischen Aperturen. Blau entspricht der skalaren Rechnung und rot
der RCWA Rechnung.

Apertur ist. Insgesamt ist die Kurvenform in den Rechnungen für die kontinuierlichen Elemente sehr ähnlich.

Bei den Simulationen des phasenbinarisierten und amplitudennivellierten Elements sieht man auch schon bei niedrigen numerischen Aperturen die typischen Störungen des fokalen Profils. Hierbei gibt es deutliche Abweichungen zwischen rigoroser und skalarer Rechnung, allerdings haben beide Kurven eine ähnliche Einhüllende. Bei hohen numerischen Aperturen sieht man wieder einen deutlichen Unterschied in der maximalen Intensität, sowie eine leichte Verschiebung des Fokus hin zum diffraktiven Element.

Fazit Für die Gültigkeit der skalaren Rechnung bei den diffraktiven Linsen mit überlappenden Aperturen lässt sich keine harte Grenze ziehen, bis wann die skalare Näherung mit dünner Element Näherung valide ist. Insbesondere für die phasenbinarisierten und amplitudennivellierten Elemente lässt sich festhalten, dass bis zu einer numerischen Apertur von $NA = 0,5$ die Abweichungen relativ gering sind. Bei höheren numerischen Aperturen sieht man deutliche Abweichungen sowohl in der Geometrie des Fokus als auch in der maximalen Amplitude. Beim Vergleich mit den Literaturmeinungen sieht man, dass der Gültigkeitsbereich für die skalare Rechnung an phasenbinarisierten und amplitudennivellierten Elementen konsistent zu den Ergebnissen von Mellin und Nordin [71] ist.

Für kontinuierliche diffraktive Elemente treten selbst bei einer numerischen Apertur von $NA = 0,5$ deutliche Abweichungen in der maximalen Intensität auf.

9.3 Beschleunigung der RCWA durch eine Sonderfall-Behandlung für homogene Schichten

Bei der Simulation von diffraktiven Elementen mit Deckglas-Korrektur folgt nach der strukturierten Schicht, die das diffraktive Element beschreibt, meist noch mehrere homogene Schichten, die die Ausbreitung bis zum Deckglas und die Ausbreitung im Deckglas berücksichtigen. Für diese homogenen Schichten ist es vorteilhaft eine Sonderfall-Behandlung für manche der in Abschnitt 3.3.1 definierten Größen in die RCWA zu integrieren. Insbesondere lässt sich der Speicherbedarf für die homogenen Schichten reduzieren indem man bestimmte Größen für diese Schichten als dünnbesetzte Matrizen (engl. *sparse matrix*) definiert. Im Einzelnen bedeutet dies:

- Die Toeplitz-Matrizen für die Beschreibung der Materialparameter werden zu skalierten Einheitsmatrizen. Somit erreicht man eine Reduktion von $(2 MO_x + 1)^2 \times (2 MO_y + 1)^2$ auf nur noch einen relevanten Koeffizienten pro Matrix:

$$[\tilde{\epsilon}] = \epsilon \mathbf{I} \tag{9.8}$$

$$[\tilde{\mu}] = \mu \mathbf{I} \tag{9.9}$$

- Die Eigenwert- und Eigenvektor-Berechnung wird trivial, da Ω_1 schon eine Diagonalmatrix ist, somit gilt:

$$\Lambda = \Omega_1 \quad (9.10)$$

$$\mathbf{W} = \mathbf{I} \quad (9.11)$$

- Ω_2^{-1} lässt sich schnell als Blockdiagonalmatrix analytisch invertieren und ist wieder eine Blockdiagonalmatrix.

Weitere Vorteile ergeben sich durch Sonderfall-Behandlungen beim Lösen der Randbedingung in der ETMA. Da die ETMA zuerst von hinten nach vorne durch die Schichten geht kann man folgende Vorteile nutzen:

- Die Matrix $\begin{pmatrix} -\mathbf{V} & \mathbf{f} \\ -\mathbf{W} & \mathbf{g} \end{pmatrix}$ ist für die homogenen Schichten eine 4×4 Block-Diagonalmatrix und lässt sich analytisch invertieren. Hierdurch kann die Zeit für die Invertierung deutlich beschleunigt werden.
- Die Hilfsmatrizen \mathbf{a} , \mathbf{b} , \mathbf{f} und \mathbf{g} sind für die homogenen Schichten nicht voll besetzt und lassen sich als Sparse-Matrix behandeln.

Zusammenfassend lässt sich feststellen, dass durch die homogenen Schichten, die hinter der strukturierten Schicht sind, das zu lösende Problem nicht schwieriger ist als die Simulation mit nur einer strukturierten Schicht. Sowohl speicherseitig als auch rechenzeitmäßig sind die zusätzlichen homogenen Schichten hinter dem diffraktiven Element nicht relevant. Dass die homogenen Schichten die Rechenzeit nicht merklich erhöhen liegt daran, dass die zwei zeitintensiven Schritte, die Eigenwertberechnung in der Schicht und die Matrix-Inversion in der ETMA, umgangen beziehungsweise in ihrer Komplexität erheblich reduziert wurden.

Für den Fall, dass sich vor der strukturierten Schicht homogene Schichten befinden, hat Junker [53] eine Erweiterung der ETMA vorgeschlagen, in der für die homogenen Schichten alle Hilfsmatrizen dünnbesetzt sind und sich somit die Invertierungen auch schnell durchführen lassen.

9.4 Konsistenzüberprüfung der RCWA anhand der Impulserhaltung

Ein typisches Problem der klassischen RCWA ist, dass die modenbegrenzte Beschreibung nur im Idealfall von unendlich vielen Moden gegen die kontinuierliche Beschreibung konvergiert [60]. Dies hat zur Folge, dass bei der Simulation nicht modenbegrenzter Strukturen sowohl die Felder als auch die Reflexions- und Transmissionskoeffizienten noch nicht konvergiert sind, sodass die klassische RCWA nicht konsistente Ergebnisse liefert und die allgemeinen physikalischen Erhaltungssätze verletzt sind [13].

Durch die Behandlung von modenbegrenzten Systemen und die Einführung von zyklisch fortgesetzten Materialparametern und durch Wählen der konsistenten Abtastung der Felder wurde dieses Problem von Auer und Brenner [3] gelöst. Diese Form der RCWA wird LIF-RCWA genannt. Es wurde gezeigt, dass die Energieerhaltung in der LIF-RCWA erfüllt ist [1].

Brenner und Auer [16] haben die Idee der Überprüfung der Konsistenz der Felder und der Reflexions- und Transmissions-Koeffizienten auf die Impulserhaltung erweitert und für TE-polarisierte Wellen verifiziert.

Für TM-polarisierte Wellen muss der Ansatz erweitert werden. Hier existiert ein nicht zu vernachlässigbarer Beitrag, der durch die Unstetigkeiten der Felder im Maxwell'schen Spannungstensor begründet ist (vgl. Abschnitt 2.5).

Im folgenden Abschnitt wird anhand zweier Beispiele gezeigt, dass die LIF-RCWA die Impulserhaltung erfüllt. Dies wird zuerst an einer homogenen Schicht gezeigt, die mit einer TM-polarisierten ebenen Welle beleuchtet wird. Anschließend wird ein homogener Block mit einer ebenen Welle mit TM-Polarisation beleuchtet.

9.4.1 Beispiel: Homogene Schicht und TM-Polarisation

Im Folgenden wird eine homogene Schicht der Dicke $0,3 \mu\text{m}$ mit komplexer Epsilon-Verteilung angenommen, die mit einer TM-polarisierten ebenen Welle der Wellenlänge $\lambda = 0,633 \mu\text{m}$ unter einem Winkel von 30° beleuchtet wird. Sowohl vor der Schicht als auch hinter der Schicht befindet sich Luft. Die Periode wurde so gewählt, dass die beleuchtende Welle an der Periodengrenze keine Unstetigkeit besitzt. In Tabelle 9.1 sind alle relevanten Parameter der Simulation aufgelistet.

Die Rekonstruktion der Epsilon-Verteilung ist in Abbildung 9.22 gezeigt. Aufgrund des LIF-RCWA Samplings wird der Ausgangswert der Permittivität $\varepsilon_{\text{Layer}} = 2,882\,775 + 0,289i$ exakt rekonstruiert (vgl. Tabelle 9.1).

Mit der RCWA lassen sich nun die zugehörigen \mathcal{E} und \mathcal{H} -Felder berechnen. Die von null verschiedenen Komponenten sind in Abbildung 9.23 gezeigt. Man sieht deutlich die Unstetigkeit in der \mathcal{E}_z -Komponente.

Nun lässt sich mit Gleichung 2.73 die Energieerhaltung und mit Gleichung 2.83 die Impulserhaltung aus Kapitel 2.5 überprüfen.

Für eine ebene Welle mit einer Feldstärke von $|\mathcal{E}| = 1 \text{ V } \mu\text{m}^{-2}$ erhält man für das zeitlich gemittelte Integral des Maxwell'schen Spannungstensors:

$$\left\langle \int_{\partial V} \mathbf{T} \cdot \mathbf{n} \, dA \right\rangle_T = \begin{pmatrix} 920,494 \\ 0 \\ 2509,599 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.12)$$

Der zeitlich gemittelte Beitrag des Integrals über die Unstetigkeit des Tensors ergibt:

$$-\left\langle \int_{\Sigma} (\mathbf{T}^+ - \mathbf{T}^-) \cdot \mathbf{n}_{\Sigma} \, dA \right\rangle_T = \begin{pmatrix} -594,327 \\ 0 \\ 768,876 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.13)$$

	Parameter	Wert
Umgebungsmedium	n_I	1
	n_{III}	1
Beleuchtungsparameter	λ	$0,633 \mu\text{m}$
	ψ	0°
	θ	30°
	φ	0°
Gitterdimension	P_x	$\frac{\lambda}{n_I \sin \theta} = 1,266 \mu\text{m}$
	P_y	$1 \mu\text{m}$
Schichtparameter	n_{Layer}	$1,7 + 0,085i$
	$\varepsilon_{\text{Layer}}$	$n_{\text{Layer}}^2 = 2,882\,775 + 0,289i$
	d_{Layer}	$0,3 \mu\text{m}$
Fouriermoden	MO_x	20
	MO_y	0
Abtastung in z	z_{min}	$-0,1 \mu\text{m}$
	z_{max}	$0,4 \mu\text{m}$
	N_z	1001

Tabelle 9.1: Parameter der RCWA Simulation der homogenen Schicht mit TM-polarisierter Beleuchtungswelle.

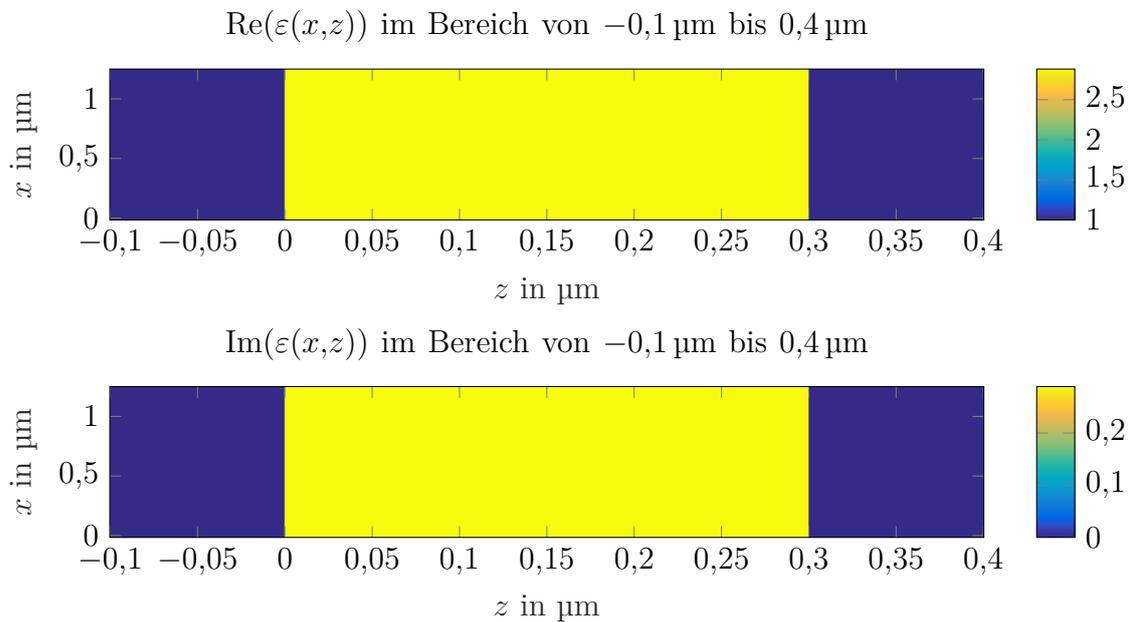


Abbildung 9.22: Rekonstruktion des Realteils (oben) und des Imaginärteils (unten) der Permittivitätsverteilung für eine homogene Schicht mit der Dicke $d = 0,3 \mu\text{m}$.

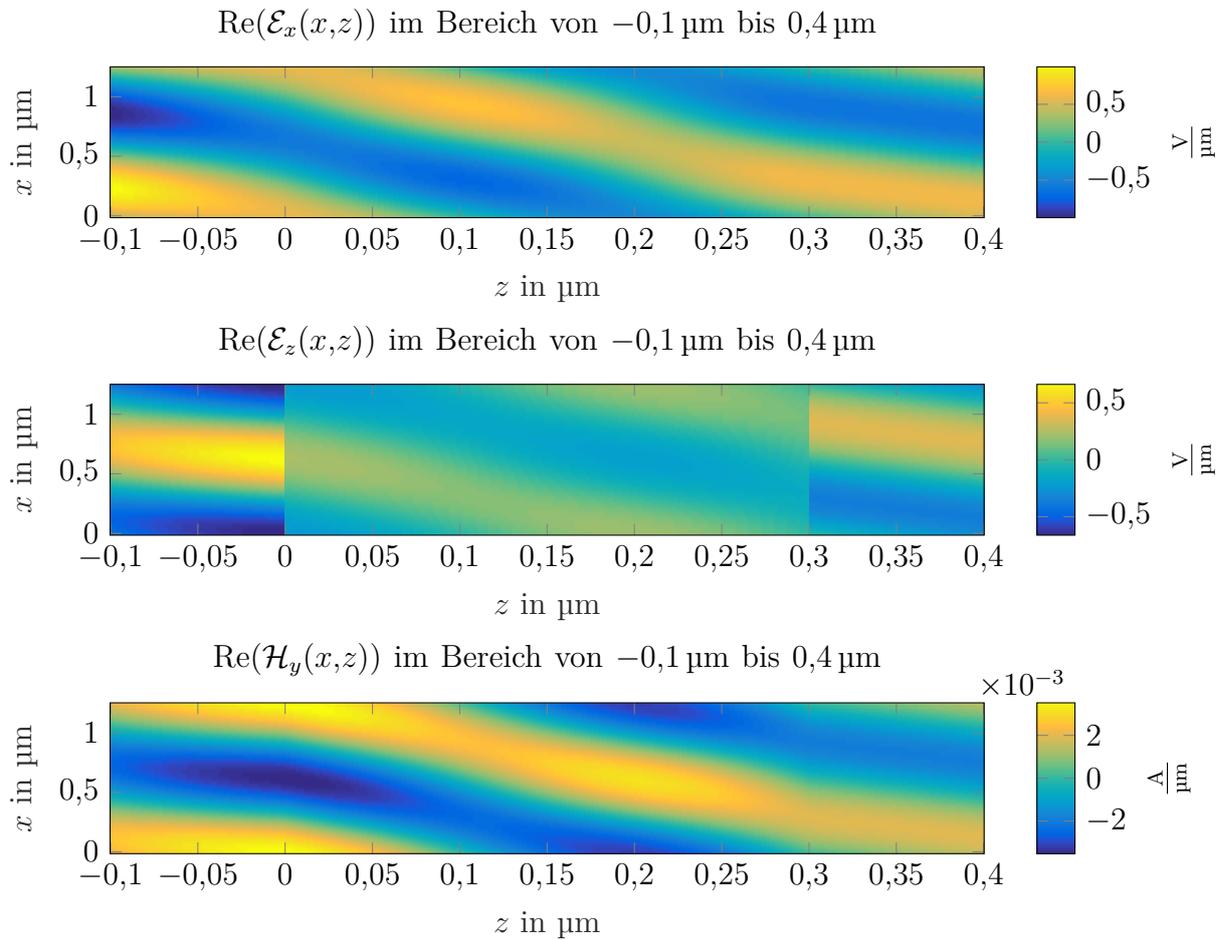


Abbildung 9.23: Rekonstruktion der von null verschiedenen Feldkomponenten für eine homogene Schicht.

Die Beiträge der Unstetigkeit sind für den Fall der TM-Polarisation in der gleichen Größenordnung wie die Beiträge des Tensors selbst. Die zeitlich gemittelten Beiträge der Integrale über die Volumenterme sind:

$$\left\langle \int_V +\epsilon_0 \mathbf{E} \nabla \cdot \mathbf{E} dV \right\rangle_T = \begin{pmatrix} 0,006 \\ 0 \\ -0,411 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.14)$$

$$+ \left\langle \int_V \mu_0 \mathbf{H} \nabla \cdot \mathbf{H} dV \right\rangle_T = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.15)$$

$$- \left\langle \int_V \frac{\partial \mathbf{B}}{\partial t} \times \epsilon_0 \mathbf{E} dV \right\rangle_T = \begin{pmatrix} 326,465 \\ 0 \\ -211,518 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.16)$$

$$+ \left\langle \int_V \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} dV \right\rangle_T = \begin{pmatrix} 0 \\ 0 \\ 3488,562 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.17)$$

Die expliziten Berechnungsvorschriften sind in Abschnitt 2.5 gezeigt. Die Beiträge der Volumenterme ergeben somit folgende Kraftdichte:

$$\begin{aligned} & + \left\langle \int_V +\epsilon_0 \mathbf{E} \nabla \cdot \mathbf{E} dV \right\rangle_T + \left\langle \int_V \mu_0 \mathbf{H} \nabla \cdot \mathbf{H} dV \right\rangle_T \\ & - \left\langle \int_V \frac{\partial \mathbf{B}}{\partial t} \times \epsilon_0 \mathbf{E} dV \right\rangle_T + \left\langle \int_V \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} dV \right\rangle_T = \begin{pmatrix} 326,471 \\ 0 \\ 3276,633 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \end{aligned} \quad (9.18)$$

Analog dazu ergeben die Beiträge des Tensors zusammengefasst:

$$\left\langle \int_{\partial V} \mathbf{T} \cdot \mathbf{n} dA \right\rangle_T - \left\langle \int_{\Sigma} (\mathbf{T}^+ - \mathbf{T}^-) \cdot \mathbf{n}_{\Sigma} dA \right\rangle_T = \begin{pmatrix} 326,167 \\ 0 \\ 3278,475 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.19)$$

Vergleicht man die Kraftdichten, so sieht man eine gute Übereinstimmung. Der relative Fehler beträgt 0,093 % für die x-Komponente und 0,056 % für die z-Komponente.

9.4.2 Beispiel: Homogener Block und TM-Polarisation

Nun wird anstatt einer homogenen Schicht eine strukturierte Schicht untersucht. Innerhalb der Periode ist ein homogener Block mit derselben Permittivität wie im Beispiel zuvor. Die Größe des Blocks entspricht der halben Periodengröße, somit wird ein Gitter mit einem Füllfaktor von 0,5 simuliert. Die sonstigen Parameter entsprechen der vorherigen Simulation und sind in Tabelle 9.2 zusammengefasst.

	Parameter	Wert
Umgebungsmedium	n_I	1
	n_{III}	1
Beleuchtungsparameter	λ	0,633 μm
	ψ	0°
	θ	30°
	φ	0°
Gitterdimension	P_x	$\frac{\lambda}{n_I \sin \theta} = 1,266 \mu\text{m}$
	P_y	1 μm
Schichtparameter	$n_{\text{Layer}}^{\text{Block}}$	1,7 + 0,085i
	$\varepsilon_{\text{Layer}}^{\text{Block}}$	$(n_{\text{Layer}}^{\text{Block}})^2 = 2,882\,775 + 0,289i$
	$\varepsilon_{\text{Layer}}^{\text{Umgebung}}$	1
	d_{Layer}	0,3 μm
Fouriermoden	D_x^{Block}	$\frac{P_x}{2} = 0,633 \mu\text{m}$
	MO_x	20
	MO_y	0
Abtastung in z	z_{min}	-0,1 μm
	z_{max}	0,4 μm
	N_z	1001

Tabelle 9.2: Parameter der Simulation des homogenen Blocks mit TM-Beleuchtung.

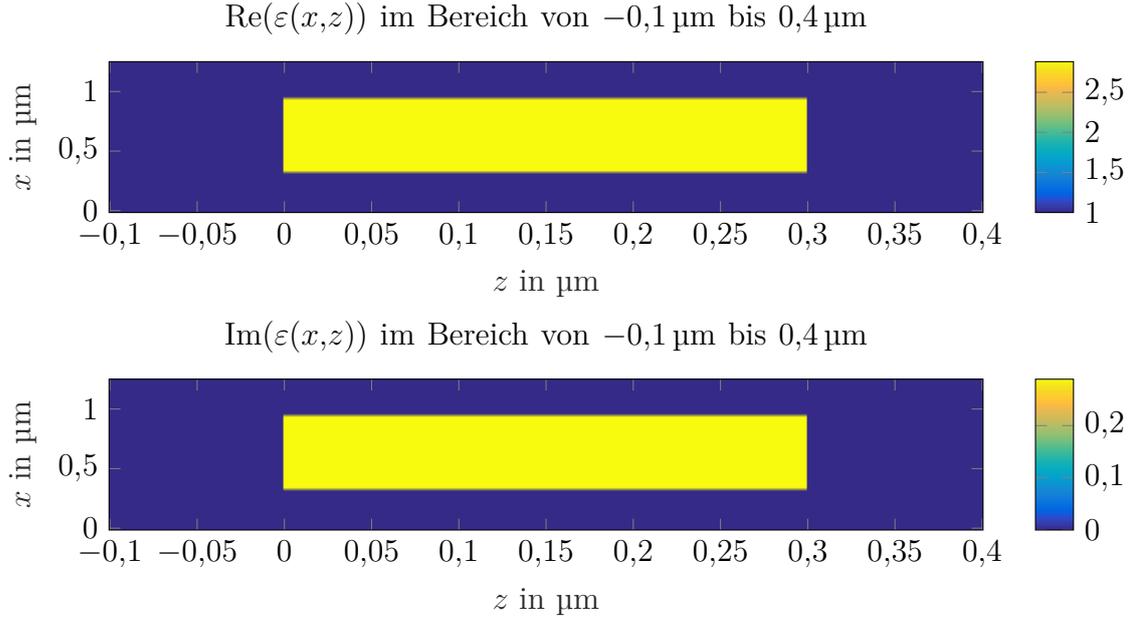


Abbildung 9.24: Rekonstruktion des Realteils (oben) und des Imaginärteils (unten) der Permittivitätsverteilung für einen homogenen Block mit der Dicke $d = 0,3 \mu\text{m}$, die Breite des Blocks entspricht der Hälfte der Periode.

Die Rekonstruktion der Epsilon-Verteilung ist in Abbildung 9.24 gezeigt. Mit der RCWA lassen sich nun die zugehörigen \mathcal{E} und \mathcal{H} -Felder berechnen. Die von null verschiedenen Komponenten sind in Abbildung 9.25 gezeigt. Man sieht wieder die Unstetigkeit in der \mathcal{E}_z Komponente.

Für eine ebene Welle mit einer Feldstärke von $|\mathcal{E}| = 1 \text{ V } \mu\text{m}^{-2}$ erhält man für das zeitlich gemittelte Integral des Tensors:

$$\left\langle \int_{\partial V} \mathbf{T} \cdot \mathbf{n} dA \right\rangle_T = \begin{pmatrix} 1476,461 \\ 0 \\ 1987,709 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.20)$$

Der zeitlich gemittelte Beitrag des Integrals über die Unstetigkeit des Tensors ergibt:

$$-\left\langle \int_{\Sigma} (\mathbf{T}^+ - \mathbf{T}^-) \cdot \mathbf{n}_{\Sigma} dA \right\rangle_T = \begin{pmatrix} -505,974 \\ 0 \\ -696,309 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.21)$$

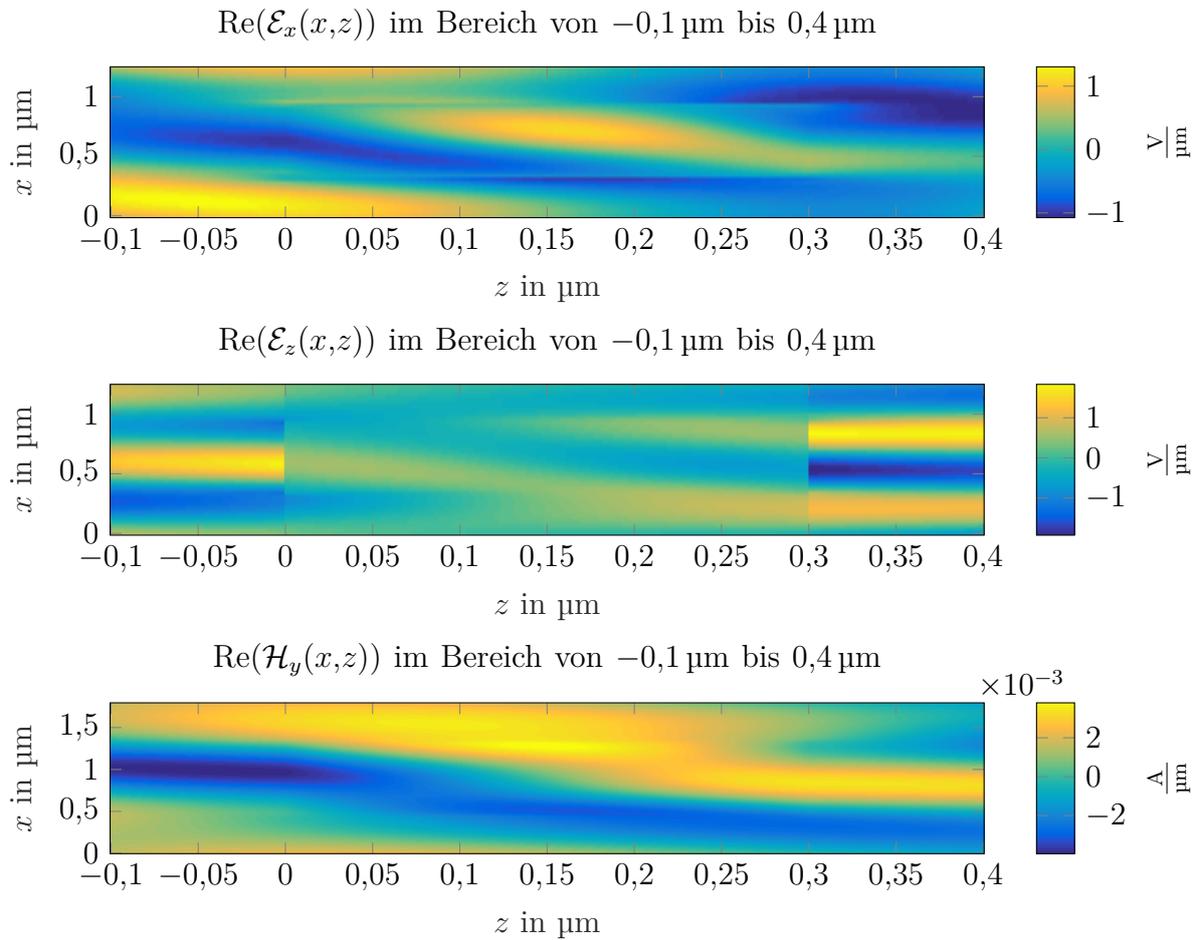


Abbildung 9.25: Rekonstruktion der von null verschiedenen Feldkomponenten für einen homogenen Block.

Für die Beiträge der Volumenterme und dem Term über die Unstetigkeit erhält man:

$$\left\langle \int_V +\epsilon_0 \mathbf{E} \nabla \cdot \mathbf{E} dV \right\rangle_T = \begin{pmatrix} 774,346 \\ 0 \\ 184,42 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.22)$$

$$+ \left\langle \int_V \mu_0 \mathbf{H} \nabla \cdot \mathbf{H} dV \right\rangle_T = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.23)$$

$$- \left\langle \int_V \frac{\partial \mathbf{B}}{\partial t} \times \epsilon_0 \mathbf{E} dV \right\rangle_T = \begin{pmatrix} 193,861 \\ 0 \\ 1372,282 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.24)$$

$$+ \left\langle \int_V \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} dV \right\rangle_T = \begin{pmatrix} 0 \\ 0 \\ -272,117 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.25)$$

Die Beiträge der Volumenterme ergeben somit folgende Kraftdichte:

$$\begin{aligned} & + \left\langle \int_V +\epsilon_0 \mathbf{E} \nabla \cdot \mathbf{E} dV \right\rangle_T + \left\langle \int_V \mu_0 \mathbf{H} \nabla \cdot \mathbf{H} dV \right\rangle_T \\ & - \left\langle \int_V \frac{\partial \mathbf{B}}{\partial t} \times \epsilon_0 \mathbf{E} dV \right\rangle_T + \left\langle \int_V \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} dV \right\rangle_T = \begin{pmatrix} 968,207 \\ 0 \\ 1285,106 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \end{aligned} \quad (9.26)$$

Analog dazu ergeben die Beiträge des Tensors zusammengefasst:

$$\left\langle \int_{\partial V} \mathbf{T} \cdot \mathbf{n} dA \right\rangle_T - \left\langle \int_{\Sigma} (\mathbf{T}^+ - \mathbf{T}^-) \cdot \mathbf{n}_{\Sigma} dA \right\rangle_T = \begin{pmatrix} 970,487 \\ 0 \\ 1291,400 \end{pmatrix} \times 10^{-21} \text{ N } \mu\text{m}^{-2} \quad (9.27)$$

Vergleicht man die Kraftdichten so sieht man auch für dieses Beispiel eine gute Übereinstimmung. Der relative Fehler beträgt 0,24 % für die x-Komponente und 0,49 % für die z-Komponente.

Fazit

Neben der Energieerhaltung erfüllt die LIF-RCWA auch die Impulserhaltung. Sowohl die Integration über die Kraftdichte, als auch die Integration über die Divergenz des Maxwell'schen Spannungstensors liefern konsistente Ergebnisse. Beim Anwenden des Gauß'schen Integralsatzes muss jedoch der Maxwell'sche Spannungstensor auf Unstetigkeiten überprüft werden. Insbesondere bei Grenzflächen mit Normalenvektor parallel zur z-Achse und einer TM-polarisierten Beleuchtung existiert ein Kraftdichtebeitrag von der Unstetigkeitsfläche. In zwei Beispielen wurde die Impulserhaltung verifiziert.

9.5 Fazit

Die LIF-RCWA eignet sich gut für die Simulation diffraktive Linsenarrays mit überlappenden Aperturen. Die periodischen Randbedingungen der RCWA erlauben die Reduzierung des Arrays auf eine Grundperiode. Im Gegensatz zur skalaren Rechnung mit dünner Element Näherung wird in der RCWA die Dicke des diffraktiven Elements berücksichtigt, außerdem werden der Vektorcharakter der Lichts sowie polarisationsspezifische Reflexionen an den Grenzflächen korrekt behandelt. Aufgrund des hohen Speicher- und Rechenzeitaufwandes eignet sich die Methode jedoch nur für die Simulation diffraktiver Elemente mit relativ kleinen Perioden bei gleichzeitig hoher numerischer Apertur.

Für die Simulation von homogenen Schichten wurde eine Sonderfallbehandlung vorgestellt. Damit ist die Simulation von Deckgläsern im Bereich zwischen dem diffraktiven Element und dem Fokus performant möglich. Des Weiteren wurde für die LIF-RCWA an zwei Beispielen gezeigt, dass konsistente Ergebnisse berechnet werden und die Impulserhaltung erfüllt ist.

Der Vergleich zwischen der Intensität aus der skalaren Rechnung und der RCWA Simulation zeigt, dass bis zu einer numerischen Apertur von $NA = 0,5$ die skalare Rechnung für phasenbinarisierte und amplitudennivellierte Elemente sehr gut übereinstimmen, sodass hier die skalare Rechnung angewendet werden kann. Diese Beobachtung ist konsistent zur der von Mellin und Nordin [71], die eine ähnliche Gültigkeitsgrenze bei diffraktiven Strahlteilern gefunden haben.

Für diffraktive Elemente mit höheren numerischen Aperturen haben sich Abweichungen zwischen rigoroser und skalarer Rechnung gezeigt, sodass hier für eine quantitative Analyse eine rigorose Simulation erforderlich ist. Die Abweichungen bei höheren numerischen Aperturen äußern sich in leichten Abweichungen des Fokalprofils. Hierbei wird die maximale Intensität in der skalaren Rechnung im Vergleich zur RCWA Rechnung überschätzt.

Allgemein sollte festgehalten werden, dass auch im Fall sehr hoher numerischer Aperturen die Permittivitätsverteilung, die aus einer skalaren Rechnung berechnet wurde, in der rigorosen Simulation näherungsweise beugungsbegrenzte Foki liefert.

Um das Verhalten bei sehr hohen numerischen Aperturen zu optimieren, ist eine Erweiterung des skalaren Design-Prozess auf einen vektoriellen Design-Prozess möglich. Hierzu könnte zum Beispiel die vektorielle Feldverteilung einer elektrischen Dipol-Welle in der Fokusebene angenommen werden. Mit Hilfe der vektoriellen Zerlegung nach ebenen Wellen ist die Rückpropagation bis zur der Grenzfläche des diffraktiven Elements möglich. Anschließend müsste das inverse Problem gelöst werden, um die passende Permittivitätsverteilung zu finden, die diese Feldverteilung erzeugt. Ansätze hierzu finden sich zum Beispiel bei Trauter [96].

10 Zusammenfassung und Ausblick

In dieser Arbeit wurden das Design und die Funktionsweise von diffraktiven Linsen mit überlappenden Aperturen vorgestellt. Die Unabhängigkeit von Fokusgröße und Arbeitsabstand dieser Linsenarrays wird durch die kohärente Überlagerung fokussierender Teilwellen in einem diffraktiven optischen Element erreicht, bei dem die Aperturen der Teilwellen in der Elementebene überlappen.

Zur Erzeugung dieser diffraktiven Linsen wurden Design-Methoden im Ortsraum und im Frequenzraum vorgestellt. Neben dem Design einfacher punktförmiger oder ringförmiger Foki wurde auch der allgemeine Fall komplexer Fokalstrukturen untersucht. Erweiterungen des Design-Prozesses erlauben die Überlagerung mehrerer Beleuchtungsstrukturen sowie eine Deckglaskorrektur.

Anhand skalaren Simulationen wurde die normierte Intensität im Fokus für verschiedene abgeblendete Wellenarten wie die Sommerfeld-Welle, die Kugelwelle oder die sphärische Phase untersucht. Hierbei zeigte sich, dass die abgeblendete Sommerfeld-Welle ein ideales generalisiertes Strehl besitzt und die gleiche normierte Fokusintensität erzeugt wie ein ideales diffraktives Element. In diesem Sinne fokussieren auch die diffraktiven Elemente mit überlappenden Aperturen optimal.

Da oft anstatt diffraktiver Elemente mit kontinuierlicher Amplituden- und Phasenmodulation modifizierte Elemente mit quantisierter Phase und konstanter Amplitude hergestellt werden, wurden die Auswirkungen dieser typischen Modifikationen auf die Foki der diffraktiven Elemente mit überlappenden Aperturen untersucht. Bei niedrigen numerischen Aperturen wurden starke Störungen im Fokusprofil festgestellt, bei mittleren und hohen numerischen Aperturen hingegen erhält man weiterhin beugungsbegrenzte Fokalprofile, deren normierte Intensität im Vergleich zum nicht modifizierten Element vermindert ist.

Neben diesen bewussten Modifikationen wurden auch der Einfluss der Pixelierung des diffraktiven Elements sowie die Auswirkungen der endlichen Wiederholungszahl der Grundperiode und diverse Abweichungen von den Design-Parametern untersucht. Hierbei zeigt sich, dass die diffraktiven Elemente mit überlappenden Aperturen sehr robust gegenüber Abweichungen in der Wellenlänge, der Phasenhöhe des Elements oder gegenüber verwaschenen Kanten bei der Herstellung sind.

Für ein diffraktives Linsenarray mit überlappenden Aperturen mit integrierter Deckglaskorrektur und einer numerischen Apertur von 0,48 und ein diffraktives Linsenarray mit einer numerischen Apertur von 0,75 wurde ein dichtes Beleuchtungsrastrer mit einer Periode von $44\ \mu\text{m}$ experimentell gut bestätigt. In der Elementebene entspricht dies Überlappfaktoren von ca. 50 bzw. 103.

Als Anwendungen, die von diffraktiven Mikrolinsen mit überlappenden Aperturen profitieren können, wurden die parallel scannende Mikroskopie und die Wellenfront-

messung vorgestellt. Bei der parallel scannenden Mikroskopie erlaubt die Verwendung von diffraktiven Linsen mit überlappenden Aperturen eine dichte Beleuchtung der Probe mit hochaperturigen Foki und gleichzeitig einen großen Arbeitsabstand, eine hohe Parallelität, ein großes Sichtfeld und hohe Auflösungen.

Experimentell wurde ein Testaufbau eines parallel scannenden Mikroskops mit einem maximalen Sichtfeld von ca. $900\ \mu\text{m} \times 1200\ \mu\text{m}$ realisiert, das Strukturen bis zu einer Größe von $450\ \text{nm}$ bei der Wellenlänge $0,532\ \mu\text{m}$ auflösen kann.

Eine weitere Anwendung, die potentiell von diffraktiven Linsen mit überlappenden Aperturen profitiert, ist die Wellenfrontmessung mittels Hartmann-Shack-Sensor, für die ein Ansatz für einen neuen Rekonstruktions-Algorithmus vorgeschlagen wurde.

Abschließend wurde die Validität des skalaren Design-Prozesses anhand von rigorosen vektoriellen Simulationen mit der „rigorous coupled-wave analysis“ (RCWA) überprüft. Hierbei zeigte sich, dass skalar designte diffraktive Elemente in der vektoriellen Rechnung auch bei sehr hohen numerischen Aperturen Foki erzeugen. Aus dem Vergleich zwischen skalarer und vektorieller Rechnung wurde die Gültigkeit der skalaren Rechnung abgeschätzt. Für phasenbinarisierte und amplitudennivelierte Elemente ähnelt die skalare Rechnung der vektoriellen Rechnung bis zu einer numerischen Apertur von ca. 0,5 und kann deshalb als valide angenommen werden.

Zusätzlich wurde gezeigt, dass die LIF-RCWA konsistente Ergebnisse liefert, da der Impuls in der elektromagnetischen Rechnung erhalten bleibt.

Ausblick

Prinzipiell lassen sich diffraktive Linsen mit überlappenden Aperturen immer dann einsetzen, wenn eine punktförmige Beleuchtung auf einem regelmäßigen Raster verwendet wird. Insbesondere birgt Verwendung diffraktiver Linsen mit überlappenden Aperturen zur Wellenfrontvermessung ein großes Potential. Die Entwicklung eines Rekonstruktionsalgorithmus, der die überlappenden Aperturen berücksichtigt, stellt eine weitere spannende Anwendung für diffraktive Linsen mit überlappenden Aperturen dar. Solch ein Algorithmus könnte auch dazu benutzt werden, verbesserte Wellenfrontrekonstruktionen aus den Daten eines normalen Hartmann-Shack-Sensors zu erhalten.

In diesem Zusammenhang ist es weiterhin möglich, neben dem Spezialfall von diffraktiven Linsen auf einem quadratischen Gitterraster Fokusarrays mit anderer Geometrie wie zum Beispiel einem hexagonalen Gittern zu erzeugen. Dies erreicht man, wenn bei der Diskretisierung im Frequenzraum ein entsprechendes Gitter verwendet wird.

Im Bereich der punktscannenden Mikroskopie bietet sich der Einsatz von diffraktiven Linsen mit überlappenden Aperturen bei der Realisierung paralleler Fluoreszenzmikroskope an. Insbesondere parallele Varianten der STED-Mikroskopie, bei denen die Ausschaltfoki auch durch ein diffraktives Linsenarray erzeugt werden, bieten sich an.

Weitere Einsatzbereiche von Linsen mit überlappenden Aperturen ergeben sich in

Anwendungen, bei denen refraktive Mikrolinsenarrays durch ihre harten Aperturkanten Mehrdeutigkeiten erzeugen. Ein Beispiel hierfür findet sich bei einer speziellen Form der iterativen Phasenrekonstruktion, hier lässt sich durch den Überlapp der Aperturen eine Phasenmehrdeutigkeit auflösen [63].

A Appendix

A.1 Koordinatensysteme

Kartesische Koordinaten

Position im Ortsraum in drei Dimensionen:

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (\text{A.1})$$

Position im Ortsraum in zwei Dimensionen:

$$\mathbf{r}_\perp = \begin{pmatrix} x \\ y \end{pmatrix} \quad (\text{A.2})$$

Position im Frequenzraum in drei Dimensionen:

$$\boldsymbol{\rho} = \begin{pmatrix} \nu \\ \mu \\ \xi \end{pmatrix} \quad (\text{A.3})$$

Position im Frequenzraum in zwei Dimensionen:

$$\boldsymbol{\rho}_\perp = \begin{pmatrix} \nu \\ \mu \end{pmatrix} \quad (\text{A.4})$$

Kugelkoordinaten

Position im Ortsraum in drei Dimensionen:

$$\mathbf{r} = \begin{pmatrix} r \sin \theta \cos \varphi \\ r \sin \theta \sin \varphi \\ r \cos \theta \end{pmatrix} \quad (\text{A.5})$$

Zylinderkoordinaten, Polarkoordinaten

Position im Ortsraum in drei Dimensionen:

$$\mathbf{r} = \begin{pmatrix} r_\perp \cos \varphi \\ r_\perp \sin \varphi \\ z \end{pmatrix} \quad (\text{A.6})$$

Position im Ortsraum in zwei Dimensionen:

$$\mathbf{r}_\perp = \begin{pmatrix} r_\perp \cos \varphi \\ r_\perp \sin \varphi \end{pmatrix} \quad (\text{A.7})$$

Position im Frequenzraum in drei Dimensionen:

$$\boldsymbol{\rho} = \begin{pmatrix} \rho_\perp \cos \chi \\ \rho_\perp \sin \chi \\ \xi \end{pmatrix} \quad (\text{A.8})$$

Position im Frequenzraum in zwei Dimensionen:

$$\boldsymbol{\rho}_\perp = \begin{pmatrix} \rho_\perp \cos \chi \\ \rho_\perp \sin \chi \end{pmatrix} \quad (\text{A.9})$$

A.2 Spezielle Funktionen und Distributionen

Definition und Eigenschaften von speziellen Funktionen:

Delta-Distribution

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0) \quad (\text{A.10})$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_0) d\mathbf{r} = f(\mathbf{r}_0) \quad (\text{A.11})$$

Rechteck-Funktion

$$\text{rect}(x) = \begin{cases} 1 & \text{für } |x| < \frac{1}{2} \\ 0 & \text{sonst} \end{cases} \quad (\text{A.12})$$

Sinc-Funktion

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \quad (\text{A.13})$$

Circ-Funktion

$$\text{circ}(\mathbf{r}_\perp) = \text{rect}\left(\frac{r_\perp}{2}\right) \quad (\text{A.14})$$

Jinc-Funktion

$$\text{jinc}(x) = \frac{J_1(x)}{x} \quad (\text{A.15})$$

A.3 Komplexe Zahlen

Seien a und b komplexe Zahlen der Form $a = a_r + ia_i$ und $b = b_r + ib_i$, dann gilt:

$$\operatorname{Re}[ia] = -\operatorname{Im}[a] \quad (\text{A.16})$$

$$\operatorname{Re}[ab^*] = \operatorname{Re}[a^*b] \quad (\text{A.17})$$

$$\operatorname{Re}[iab^*a^*] = \operatorname{Im}[b]\operatorname{Re}[aa^*] \quad (\text{A.18})$$

A.4 Rechenregeln für Vektoren und Tensoren

Seien \mathbf{A} und \mathbf{B} dreidimensionale Vektoren und \mathbf{T} eine quadratische Matrix bzw. ein Tensor zweiter Stufe, dann gilt:

$$\mathbf{A} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix} \quad (\text{A.19})$$

Der Einheits-Tensor oder die Einheits-Matrix ist gegeben durch:

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{A.20})$$

Das Skalar-Produkt zwischen zwei Vektoren ist definiert als:

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{A}^T \mathbf{B} = A_x B_x + A_y B_y + A_z B_z \quad (\text{A.21})$$

Das äußeres Produkt zweier Vektoren bilden einen Tensor:

$$\mathbf{AB} = \mathbf{AB}^T = \begin{pmatrix} A_x B_x & A_x B_y & A_x B_z \\ A_y B_x & A_y B_y & A_y B_z \\ A_z B_x & A_z B_y & A_z B_z \end{pmatrix} \quad (\text{A.22})$$

Das Produkt zwischen Tensor und Vektor ist definiert als:

$$\mathbf{TA} = \begin{pmatrix} T_{xx}A_x + T_{xy}A_y + T_{xz}A_z \\ T_{yx}A_x + T_{yy}A_y + T_{yz}A_z \\ T_{zx}A_x + T_{zy}A_y + T_{zz}A_z \end{pmatrix} \quad (\text{A.23})$$

Für das Skalar-Produkt zwischen Vektor und Tensor gilt:

$$\mathbf{A} \cdot \mathbf{T} = \mathbf{AT}^T \quad (\text{A.24})$$

Die Divergenz eines Tensors ist somit folgendermaßen definiert:

$$\nabla \cdot \mathbf{T} = \begin{pmatrix} \frac{\partial T_{xx}}{\partial x} + \frac{\partial T_{xy}}{\partial y} + \frac{\partial T_{xz}}{\partial z} \\ \frac{\partial T_{yx}}{\partial x} + \frac{\partial T_{yy}}{\partial y} + \frac{\partial T_{yz}}{\partial z} \\ \frac{\partial T_{zx}}{\partial x} + \frac{\partial T_{zy}}{\partial y} + \frac{\partial T_{zz}}{\partial z} \end{pmatrix} \quad (\text{A.25})$$

In Komponentenschreibweise ergibt sich für die Divergenz des Tensors:

$$(\nabla \cdot \mathbf{T})_i = \sum_j \frac{\partial T_{ij}}{\partial e_j} e_i \quad (\text{A.26})$$

A.5 Vektoridentitäten

Seien \mathbf{A} und \mathbf{B} dreidimensionale Vektorfelder und sei c eine skalare Konstante, dann gelten folgende allgemeine Vektoridentitäten:

$$\nabla \cdot (c\mathbf{A}) = \mathbf{A} \cdot \nabla c + c \nabla \cdot \mathbf{A} \quad (\text{A.27})$$

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = (\nabla \times \mathbf{A}) \cdot \mathbf{B} - \mathbf{A} \cdot (\nabla \times \mathbf{B}) \quad (\text{A.28})$$

$$\nabla \cdot (\mathbf{A}\mathbf{B}) = (\mathbf{A} \cdot \nabla)\mathbf{B} + (\nabla \cdot \mathbf{A})\mathbf{B} \quad (\text{A.29})$$

$$\nabla \times (c\mathbf{A}) = \nabla c \times \mathbf{A} + c \nabla \times \mathbf{A} \quad (\text{A.30})$$

$$(\nabla \times \mathbf{A}) \times \mathbf{A} = -\frac{1}{2} \nabla A^2 + (\mathbf{A} \cdot \nabla)\mathbf{A} \quad (\text{A.31})$$

$$(\nabla \times \mathbf{A}) \times \mathbf{A} = \nabla \cdot \left(\mathbf{A}\mathbf{A} - \frac{1}{2} A^2 \mathbf{I} \right) - (\nabla \cdot \mathbf{A})\mathbf{A} \quad (\text{A.32})$$

$$\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \quad (\text{A.33})$$

A.6 Fourier-Transformations-Paare

Für bestimmte zweidimensionale analytische Funktionen $f(x,y)$ lässt sich die Fouriertransformierte $\tilde{f}(\nu,\mu)$ analytisch bestimmen. Im Folgenden sind ein paar wichtige Fourier-Transformations-Paare aufgelistet:

$f(x,y)$	$\circ \bullet$	$\tilde{f}(\nu,\mu) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) e^{-2\pi i(x\nu+y\mu)} dx dy$
$\cos\left(2\pi \frac{x-\Delta x}{P}\right)$	$\circ \bullet$	$\frac{1}{2} \delta\left(\nu - \frac{1}{P}, \mu\right) e^{-2\pi i\nu\Delta x} + \frac{1}{2} \delta\left(\nu + \frac{1}{P}, \mu\right) e^{-2\pi i\nu\Delta x}$
$\text{rect}\left(\frac{x-\Delta x}{P}\right)$	$\circ \bullet$	$P e^{-2\pi i\nu\Delta x} \text{sinc}(P\nu) \delta(\mu)$

A.7 Zeit-Mittlung für monofrequente Felder

Für zeitlich monofrequente Felder $\mathbf{A}(\mathbf{r},t)$ und $\mathbf{B}(\mathbf{r},t)$, die über die komplexen analytischen Signale der Form $\mathbf{A}(\mathbf{r},t) = \frac{1}{2}(\mathbf{A}(\mathbf{r},\omega)e^{-i\omega t} + \mathbf{A}^*(\mathbf{r},\omega)e^{i\omega t})$ und $\mathbf{B}(\mathbf{r},t) = \frac{1}{2}(\mathbf{B}(\mathbf{r},\omega)e^{-i\omega t} + \mathbf{B}^*(\mathbf{r},\omega)e^{i\omega t})$ beschrieben werden, gelten folgende zeitliche Mittelwerte:

$$\langle \mathbf{A}(\mathbf{r},t) \rangle_T = 0 \quad (\text{A.34})$$

$$\begin{aligned}
& \langle \mathbf{A}(\mathbf{r},t) \cdot \mathbf{A}(\mathbf{r},t) \rangle_T \\
&= \langle A^2(\mathbf{r},t) \rangle_T \\
&= \left\langle \frac{1}{2} (\mathbf{A}(\mathbf{r},\omega)e^{-i\omega t} + \mathbf{A}^*(\mathbf{r},\omega)e^{i\omega t}) \cdot \frac{1}{2} (\mathbf{A}(\mathbf{r},\omega)e^{-i\omega t} + \mathbf{A}^*(\mathbf{r},\omega)e^{i\omega t}) \right\rangle_T \\
&= \frac{1}{4} \langle 2\mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{A}^*(\mathbf{r},\omega) + \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{A}^*(\mathbf{r},\omega)e^{2i\omega t} + \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{A}^*(\mathbf{r},\omega)e^{-2i\omega t} \rangle_T \\
&= \frac{1}{2} \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{A}^*(\mathbf{r},\omega) \\
&= \frac{1}{2} |\mathbf{A}(\mathbf{r},\omega)|^2
\end{aligned} \tag{A.35}$$

$$\begin{aligned}
& \langle \mathbf{A}(\mathbf{r},t) \cdot \mathbf{B}(\mathbf{r},t) \rangle_T \\
&= \left\langle \frac{1}{2} (\mathbf{A}(\mathbf{r},\omega)e^{-i\omega t} + \mathbf{A}^*(\mathbf{r},\omega)e^{i\omega t}) \cdot \frac{1}{2} (\mathbf{B}(\mathbf{r},\omega')e^{-i\omega' t} + \mathbf{B}^*(\mathbf{r},\omega')e^{i\omega' t}) \right\rangle_T \\
&= \frac{1}{4} \langle \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}(\mathbf{r},\omega')e^{-i(\omega+\omega')t} + \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega')e^{-i(\omega-\omega')t} \\
&\quad + \mathbf{A}^*(\mathbf{r},\omega) \cdot \mathbf{B}(\mathbf{r},\omega')e^{i(\omega-\omega')t} + \mathbf{A}^*(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega')e^{i(\omega+\omega')t} \rangle_T \\
&= \frac{1}{4} [\mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega) + \mathbf{A}^*(\mathbf{r},\omega) \cdot \mathbf{B}(\mathbf{r},\omega)] \\
&= \frac{1}{4} [\mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega) + (\mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega))^*] \\
&= \frac{1}{2} \text{Re}[\mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega)]
\end{aligned} \tag{A.36}$$

$$\begin{aligned}
& \left\langle \mathbf{A}(\mathbf{r},t) \cdot \frac{\partial}{\partial t} \mathbf{B}(\mathbf{r},t) \right\rangle_T \\
&= \left\langle \frac{1}{2} (\mathbf{A}(\mathbf{r},\omega)e^{-i\omega t} + \mathbf{A}^*(\mathbf{r},\omega)e^{i\omega t}) \cdot \frac{1}{2} (-i\omega \mathbf{B}(\mathbf{r},\omega')e^{-i\omega' t} + i\omega \mathbf{B}^*(\mathbf{r},\omega')e^{i\omega' t}) \right\rangle_T \\
&= \frac{1}{4} \langle -i\omega \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}(\mathbf{r},\omega')e^{-i(\omega+\omega')t} + i\omega \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega')e^{-i(\omega-\omega')t} \\
&\quad - i\omega \mathbf{A}^*(\mathbf{r},\omega) \cdot \mathbf{B}(\mathbf{r},\omega')e^{i(\omega-\omega')t} + i\omega \mathbf{A}^*(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega')e^{i(\omega+\omega')t} \rangle_T \\
&= \frac{1}{4} [i\omega \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega) - i\omega \mathbf{A}^*(\mathbf{r},\omega) \cdot \mathbf{B}(\mathbf{r},\omega)] \\
&= \frac{1}{4} [i\omega \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega) + (i\omega \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega))^*] \\
&= \frac{1}{2} \text{Re}[i\omega \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{B}^*(\mathbf{r},\omega)]
\end{aligned} \tag{A.37}$$

$$\begin{aligned}
\left\langle \mathbf{A}(\mathbf{r},t) \cdot \frac{\partial}{\partial t} \mathbf{A}(\mathbf{r},t) \right\rangle_T &= \frac{1}{2} \text{Re}[i\omega \mathbf{A}(\mathbf{r},\omega) \cdot \mathbf{A}^*(\mathbf{r},\omega)] \\
&= 0
\end{aligned} \tag{A.38}$$

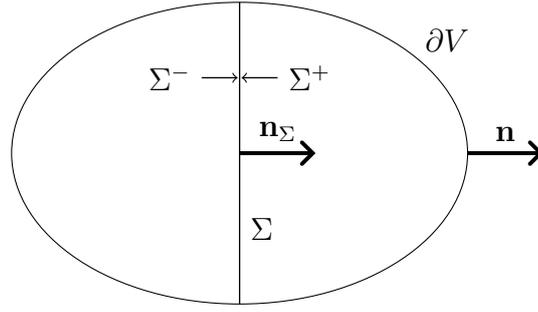


Abbildung A.1: Schematische Zeichnung eines Volumens V mit Oberfläche ∂V und einer Sprungebene Σ mit den dazugehörigen Normalenvektoren. Die Vorderseite der Sprungebene ist Σ^+ , die Rückseite Σ^- .

A.8 Gaußscher Integralsatz für Vektorfelder mit Unstetigkeit

Sei \mathbf{T} ein Vektorfeld, das stetig differenzierbar auf dem Volumen V ist, außer an der Sprungebene Σ mit der Einheitsnormalen \mathbf{n}_Σ . Dann lässt sich das Volumenintegral über die Divergenz des Vektorfeldes in einen Beitrag über die Oberfläche ∂V und in einen Beitrag über die Sprungebene Σ darstellen.

$$\int_V \nabla \cdot \mathbf{T} dV = \int_{\partial V} \mathbf{T} \cdot \mathbf{n} dA - \int_\Sigma (\mathbf{T}^+ - \mathbf{T}^-) \cdot \mathbf{n}_\Sigma dA \quad (\text{A.39})$$

Die Geometrie des Problems ist in Abbildung A.1 gezeigt. Hierbei ist \mathbf{n} der nach außen gerichteter Normalenvektor der Oberfläche ∂V . Der Normalenvektor \mathbf{n}_Σ definiert eine Vorderseite Σ^+ und eine Rückseite Σ^- der Sprungebene. Bewegt man sich von der Rückseite zur Vorderseite der Sprungebene, bewegt man sich in Richtung des Normalenvektors \mathbf{n}_Σ . Der Beitrag des Vektorfeldes ausgewertet an der Rückseite der Sprungebene wird als \mathbf{T}^- bezeichnet, der Beitrag auf der Vorderseite wird \mathbf{T}^+ genannt. Ein Beweis für diesen Satz ist zum Beispiel in [56] gegeben.

A.9 Herleitung des Energie-Erhaltungssatzes

Aus den makroskopischen Maxwell-Gleichungen erhält man durch Multiplikation der Rotationsgleichungen mit $\mathbf{H} \cdot$ und $-\mathbf{E} \cdot$ und anschließender Addition

$$\mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H}) = -\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mathbf{E} \cdot \mathbf{j}_{frei} - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}. \quad (\text{A.40})$$

Zusammenfassen der beiden Divergenzterme auf der linken Seite mit der Identität $\mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H}) = \nabla \cdot (\mathbf{E} \times \mathbf{H})$ führt zu:

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = -\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mathbf{E} \cdot \mathbf{j}_{frei} - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}. \quad (\text{A.41})$$

A.10 Herleitung des Impuls-Erhaltungssatzes

Aus den makroskopischen Maxwell-Gleichungen erhält man durch Multiplikation der Rotations-Gleichungen mit $\times \varepsilon_0 \mathbf{E}$ und $\times \mu_0 \mathbf{H}$ folgende Gleichungen:

$$(\nabla \times \mathbf{E}) \times \varepsilon_0 \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} \quad (\text{A.42})$$

$$(\nabla \times \mathbf{H}) \times \mu_0 \mathbf{H} = \mathbf{j}_{frei} \times \mu_0 \mathbf{H} + \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} \quad (\text{A.43})$$

Nun wird die Vektor-Identität für das doppelte Kreuzprodukt $(\nabla \times \mathbf{A}) \times \mathbf{A} = \nabla \cdot (\mathbf{A}\mathbf{A} - \frac{1}{2}A^2\mathbf{I}) - \mathbf{A}\nabla \cdot \mathbf{A}$ verwendet:

$$\nabla \cdot \left(\varepsilon_0 \mathbf{E}\mathbf{E} - \varepsilon_0 \frac{1}{2} E^2 \mathbf{I} \right) - \varepsilon_0 \mathbf{E}\nabla \cdot \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} \quad (\text{A.44})$$

$$\nabla \cdot \left(\mu_0 \mathbf{H}\mathbf{H} - \mu_0 \frac{1}{2} H^2 \mathbf{I} \right) - \mu_0 \mathbf{H}\nabla \cdot \mathbf{H} = \mathbf{j}_{frei} \times \mu_0 \mathbf{H} + \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} \quad (\text{A.45})$$

Addieren der Gleichungen liefert:

$$\begin{aligned} & \nabla \cdot \left(\varepsilon_0 \mathbf{E}\mathbf{E} - \varepsilon_0 \frac{1}{2} E^2 \mathbf{I} \right) - \varepsilon_0 \mathbf{E}\nabla \cdot \mathbf{E} \\ & + \nabla \cdot \left(\mu_0 \mathbf{H}\mathbf{H} - \mu_0 \frac{1}{2} H^2 \mathbf{I} \right) - \mu_0 \mathbf{H}\nabla \cdot \mathbf{H} = \\ & -\frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} + \mathbf{j}_{frei} \times \mu_0 \mathbf{H} + \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} \end{aligned} \quad (\text{A.46})$$

Durch umsordieren der Terme ergibt sich die finale Impulserhaltung:

$$\begin{aligned} & \nabla \cdot \left(\varepsilon_0 \mathbf{E}\mathbf{E} - \varepsilon_0 \frac{1}{2} E^2 \mathbf{I} + \mu_0 \mathbf{H}\mathbf{H} - \mu_0 \frac{1}{2} H^2 \mathbf{I} \right) = \\ & + \varepsilon_0 \mathbf{E}\nabla \cdot \mathbf{E} + \mu_0 \mathbf{H}\nabla \cdot \mathbf{H} - \frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} + \mathbf{j}_{frei} \times \mu_0 \mathbf{H} + \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H} \end{aligned} \quad (\text{A.47})$$

Für den Fall keiner freien Ladungen und keiner freien Ströme erhält man

$$\begin{aligned} & \nabla \cdot \left(\varepsilon_0 \mathbf{E}\mathbf{E} - \varepsilon_0 \frac{1}{2} E^2 \mathbf{I} + \mu_0 \mathbf{H}\mathbf{H} - \mu_0 \frac{1}{2} H^2 \mathbf{I} \right) = \\ & + \varepsilon_0 \mathbf{E}\nabla \cdot \mathbf{E} + \mu_0 \mathbf{H}\nabla \cdot \mathbf{H} - \frac{\partial \mathbf{B}}{\partial t} \times \varepsilon_0 \mathbf{E} + \frac{\partial \mathbf{D}}{\partial t} \times \mu_0 \mathbf{H}. \end{aligned} \quad (\text{A.48})$$

A.10.1 Impulserhaltung nach Chu

Setzt man die Materiegleichungen (2.47 2.48) in die Divergenzgleichungen der Maxwell-Gleichungen ein und multipliziert mit \mathbf{E} und \mathbf{H} , so erhält man:

$$+\varepsilon_0 \mathbf{E}\nabla \cdot \mathbf{E} - \rho_{frei} \mathbf{E} = -(\nabla \cdot \mathbf{P})\mathbf{E} \quad (\text{A.49})$$

$$+\mu_0 \mathbf{H}\nabla \cdot \mathbf{H} - (\nabla \cdot \mathbf{B})\mathbf{H} = -(\nabla \cdot \mu_0 \mathbf{M}) \quad (\text{A.50})$$

Setzt man diese Beziehung in Gleichung A.47 ein, so erhält man die Formulierung der Impulserhaltung nach Chu:

$$\nabla \cdot \mathbf{T} = \mathbf{f}_{\text{Chu}} + \frac{\partial}{\partial t} \mathbf{G}_{\text{Chu}} \quad (\text{A.51})$$

$$\mathbf{T} = -\frac{1}{2}(\varepsilon_0 E^2 + \mu_0 H^2) \mathbf{I} + \varepsilon_0 \mathbf{E} \mathbf{E} + \mu_0 \mathbf{H} \mathbf{H} \quad (\text{A.52})$$

$$\begin{aligned} \mathbf{f}_{\text{Chu}} = & [\rho_{\text{frei}} - (\nabla \cdot \mathbf{P})] \mathbf{E} - (\nabla \cdot \mu_0 \mathbf{M}) \mathbf{H} + \left(\mathbf{j}_{\text{frei}} + \frac{\partial \mathbf{P}}{\partial t} \right) \times \mu_0 \mathbf{H} \\ & - \mu_0 \frac{\partial \mathbf{M}}{\partial t} \times \varepsilon_0 \mathbf{E} \end{aligned} \quad (\text{A.53})$$

$$\mathbf{G}_{\text{Chu}} = \varepsilon_0 \mu_0 \mathbf{E} \times \mathbf{H} \quad (\text{A.54})$$

A.10.2 Impulserhaltung nach Minkowski

Eine alternative Formulierung der Impulserhaltung nach Minkowski ist:

$$\nabla \cdot \mathbf{T}_{\text{Minkowski}} = \mathbf{f}_{\text{Minkowski}} + \frac{\partial}{\partial t} \mathbf{G}_{\text{Minkowski}} \quad (\text{A.55})$$

Hierbei sind die einzelnen Terme folgendermaßen definiert:

$$\mathbf{G}_{\text{Minkowski}} = \mathbf{D} \times \mathbf{B} \quad (\text{A.56})$$

$$\mathbf{T}_{\text{Minkowski}} = -\frac{1}{2}(\mathbf{D} \cdot \mathbf{E} + \mathbf{B} \cdot \mathbf{H}) \mathbf{I} + \mathbf{D} \mathbf{E} + \mathbf{B} \mathbf{H} \quad (\text{A.57})$$

$$\mathbf{f}_{\text{Minkowski}} = -\frac{1}{2} E^2 \varepsilon_0 \nabla \varepsilon - \frac{1}{2} H^2 \mu_0 \nabla \mu + \rho_{\text{frei}} \mathbf{E} + \mathbf{j}_{\text{frei}} \times \mathbf{B} \quad (\text{A.58})$$

Ausgangspunkt der Herleitung sind wieder die Rotations-Gleichung der makroskopischen Maxwell-Gleichungen, nun wird jedoch mit $\times \varepsilon \varepsilon_0 \mathbf{E}$ und $\times \mu \mu_0 \mathbf{H}$ multipliziert:

$$(\nabla \times \mathbf{E}) \times \varepsilon \varepsilon_0 \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \times \varepsilon \varepsilon_0 \mathbf{E} \quad (\text{A.59})$$

$$(\nabla \times \mathbf{H}) \times \mu \mu_0 \mathbf{H} = \mathbf{j}_{\text{frei}} \times \mu \mu_0 \mathbf{H} + \frac{\partial \mathbf{D}}{\partial t} \times \mu \mu_0 \mathbf{H} \quad (\text{A.60})$$

Nun lässt sich in der oberen Gleichung das Kreuzprodukt umdrehen und auf der rechten Seite der Gleichungen lässt sich die Definition des \mathbf{D} und \mathbf{B} -Felds einsetzen. So erhält man:

$$(\nabla \times \mathbf{E}) \times \varepsilon \varepsilon_0 \mathbf{E} = \mathbf{D} \times \frac{\partial \mathbf{B}}{\partial t} \quad (\text{A.61})$$

$$(\nabla \times \mathbf{H}) \times \mu \mu_0 \mathbf{H} = \mathbf{j}_{\text{frei}} \times \mathbf{B} + \frac{\partial \mathbf{D}}{\partial t} \times \mu \mu_0 \mathbf{H} \quad (\text{A.62})$$

Nimmt man zusätzlich die zwei Divergenzgleichung der Maxwell-Gleichungen und multipliziert beide Seiten jeweils mit \mathbf{E} und \mathbf{H} , dann erhält man folgende vier Glei-

chungen, die addiert die Impulserhaltung ergeben:

$$(\nabla \cdot \varepsilon \varepsilon_0 \mathbf{E}) \mathbf{E} = \rho_{frei} \mathbf{E} \quad (\text{A.63})$$

$$(\nabla \cdot \mu \mu_0 \mathbf{H}) \mathbf{H} = 0 \quad (\text{A.64})$$

$$(\nabla \times \mathbf{E}) \times \varepsilon \varepsilon_0 \mathbf{E} = \mathbf{D} \times \frac{\partial \mathbf{B}}{\partial t} \quad (\text{A.65})$$

$$(\nabla \times \mathbf{H}) \times \mu \mu_0 \mathbf{H} = \mathbf{j}_{frei} \times \mathbf{B} + \frac{\partial \mathbf{D}}{\partial t} \times \mu \mu_0 \mathbf{H} \quad (\text{A.66})$$

Die Terme auf der rechten Seiten der Gleichung lassen sich zur zeitlichen Ableitung der Impulsdichte zusammenfassen, die anderen Terme bilden den Beitrag einer Lorentz-Kraft-Dichte. Die linke Seite der Gleichung liefert die Divergenz des Tensors, sowie die Helmholtz-Kraft-Dichte $-\frac{1}{2} E^2 \varepsilon_0 \nabla \varepsilon - \frac{1}{2} H^2 \mu_0 \nabla \mu$.

A.11 Herleitung der Wellengleichung

Nimmt man die makroskopischen Maxwell-Gleichungen und multipliziert $\nabla \times$ an die Rotationsgleichungen, so erhält man folgende Gleichungen:

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = \rho_{frei}(\mathbf{r}, t) \quad (\text{A.67})$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \quad (\text{A.68})$$

$$\nabla \times (\nabla \times \mathbf{E}(\mathbf{r}, t)) = -\frac{\partial \nabla \times \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad (\text{A.69})$$

$$\nabla \times (\nabla \times \mathbf{H}(\mathbf{r}, t)) = \nabla \times \mathbf{j}_{frei}(\mathbf{r}, t) + \frac{\partial \nabla \times \mathbf{D}(\mathbf{r}, t)}{\partial t} \quad (\text{A.70})$$

Einsetzen der linearen Materie-Gleichungen und Anwenden der Vektoridentität für die doppelte Rotation $\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$ ergibt:

$$\nabla \cdot \varepsilon \varepsilon_0 \mathbf{E}(\mathbf{r}, t) = \rho_{frei}(\mathbf{r}, t) \quad (\text{A.71})$$

$$\nabla \cdot \mu \mu_0 \mathbf{H}(\mathbf{r}, t) = 0 \quad (\text{A.72})$$

$$\nabla(\nabla \cdot \mathbf{E}(\mathbf{r}, t)) - \nabla^2 \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \nabla \times \mu \mu_0 \mathbf{H}(\mathbf{r}, t)}{\partial t} \quad (\text{A.73})$$

$$\nabla(\nabla \cdot \mathbf{H}) - \nabla^2 \mathbf{H} = \nabla \times \mathbf{j}_{frei}(\mathbf{r}, t) + \frac{\partial \nabla \times \varepsilon \varepsilon_0 \mathbf{E}(\mathbf{r}, t)}{\partial t} \quad (\text{A.74})$$

Ausführen der Produkt-Regeln $\nabla \cdot (c\mathbf{A}) = \mathbf{A} \cdot \nabla c + c\nabla \cdot \mathbf{A}$ und $\nabla \times (c\mathbf{A}) = \nabla c \times \mathbf{A} + c\nabla \times \mathbf{A}$ liefert:

$$\nabla \cdot \mathbf{E}(\mathbf{r}, t) = \frac{\rho_{frei}(\mathbf{r}, t)}{\varepsilon\varepsilon_0} - \mathbf{E}(\mathbf{r}, t) \cdot \frac{\nabla\varepsilon}{\varepsilon} \quad (\text{A.75})$$

$$\nabla \cdot \mathbf{H}(\mathbf{r}, t) = -\mathbf{H}(\mathbf{r}, t) \cdot \frac{\nabla\mu}{\mu} \quad (\text{A.76})$$

$$\nabla(\nabla \cdot \mathbf{E}(\mathbf{r}, t)) - \nabla^2 \mathbf{E}(\mathbf{r}, t) = -\mu_0 \frac{\partial \nabla \mu \times \mathbf{H}(\mathbf{r}, t)}{\partial t} - \mu_0 \mu \frac{\partial \nabla \times \mathbf{H}(\mathbf{r}, t)}{\partial t} \quad (\text{A.77})$$

$$\begin{aligned} \nabla(\nabla \cdot \mathbf{H}) - \nabla^2 \mathbf{H} &= \nabla \times \mathbf{j}_{frei}(\mathbf{r}, t) + \varepsilon_0 \frac{\partial \nabla \varepsilon \times \mathbf{E}(\mathbf{r}, t)}{\partial t} \\ &+ \varepsilon \varepsilon_0 \frac{\partial \nabla \times \mathbf{E}(\mathbf{r}, t)}{\partial t} \end{aligned} \quad (\text{A.78})$$

Nun lassen sich die Divergenz-Gleichungen in die Rotationsgleichungen einsetzen:

$$\nabla \cdot \mathbf{E}(\mathbf{r}, t) = \frac{\rho_{frei}(\mathbf{r}, t)}{\varepsilon\varepsilon_0} - \mathbf{E}(\mathbf{r}, t) \cdot \frac{\nabla\varepsilon}{\varepsilon} \quad (\text{A.79})$$

$$\nabla \cdot \mathbf{H}(\mathbf{r}, t) = -\mathbf{H}(\mathbf{r}, t) \cdot \frac{\nabla\mu}{\mu} \quad (\text{A.80})$$

$$\begin{aligned} \nabla(\nabla \cdot \mathbf{E}(\mathbf{r}, t)) - \nabla^2 \mathbf{E}(\mathbf{r}, t) &= -\mu_0 \frac{\partial \nabla \mu \times \mathbf{H}(\mathbf{r}, t)}{\partial t} - \mu_0 \mu \varepsilon \varepsilon_0 \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} \\ &- \mu_0 \mu \frac{\partial \mathbf{j}_{frei}(\mathbf{r}, t)}{\partial t} \end{aligned} \quad (\text{A.81})$$

$$\begin{aligned} \nabla(\nabla \cdot \mathbf{H}) - \nabla^2 \mathbf{H} &= \nabla \times \mathbf{j}_{frei}(\mathbf{r}, t) + \varepsilon_0 \frac{\partial \nabla \varepsilon \times \mathbf{E}(\mathbf{r}, t)}{\partial t} \\ &- \varepsilon \varepsilon_0 \mu \mu_0 \frac{\partial^2 \mathbf{H}(\mathbf{r}, t)}{\partial t^2} \end{aligned} \quad (\text{A.82})$$

Somit erhält man die Wellengleichungen:

$$\begin{aligned} \frac{\nabla \rho_{frei}(\mathbf{r}, t)}{\varepsilon \varepsilon_0} + \mu_0 \mu \frac{\partial \mathbf{j}_{frei}(\mathbf{r}, t)}{\partial t} = \\ + \nabla \left(\mathbf{E}(\mathbf{r}, t) \cdot \frac{\nabla \varepsilon}{\varepsilon} \right) + \nabla^2 \mathbf{E}(\mathbf{r}, t) - \mu_0 \frac{\partial \nabla \mu \times \mathbf{H}(\mathbf{r}, t)}{\partial t} - \mu_0 \mu \varepsilon \varepsilon_0 \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} \end{aligned} \quad (\text{A.83})$$

Ohne freie Ladungen und freie Ströme vereinfacht sich dies zu

$$\nabla \left(\mathbf{E}(\mathbf{r}, t) \cdot \frac{\nabla \varepsilon}{\varepsilon} \right) + \nabla^2 \mathbf{E}(\mathbf{r}, t) - \mu_0 \frac{\partial \nabla \mu \times \mathbf{H}(\mathbf{r}, t)}{\partial t} - \mu_0 \mu \varepsilon \varepsilon_0 \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} = \mathbf{0} \quad (\text{A.84})$$

Für homogene Permeabilitäten entkoppelt das E-Feld vom H-Feld. Man erhält die vektorielle Wellengleichung für inhomogene Medien:

$$\nabla \left(\mathbf{E}(\mathbf{r}, t) \cdot \frac{\nabla \varepsilon}{\varepsilon} \right) + \nabla^2 \mathbf{E}(\mathbf{r}, t) - \mu_0 \mu \varepsilon \varepsilon_0 \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} = \mathbf{0} \quad (\text{A.85})$$

A.12 Definition des Brechungsindex

Für lineare isotrope Medien lassen sich Materialgleichungen mit komplexen Permittivitäten und Permeabilitäten definieren:

$$\varepsilon = \varepsilon_r + i\varepsilon_i \quad (\text{A.86})$$

$$\mu = \mu_r + i\mu_i \quad (\text{A.87})$$

Der Brechungsindex n ist auch eine komplexe Größe:

$$n = n_r + i\kappa \quad (\text{A.88})$$

Allgemein gilt:

$$n = \sqrt{\varepsilon\mu} \quad (\text{A.89})$$

Für $\mu = 1$ ergibt sich [58]

$$\varepsilon_r = n_r^2 - \kappa^2 \quad (\text{A.90})$$

$$\varepsilon_i = 2n_r\kappa \quad (\text{A.91})$$

$$n = \sqrt{\frac{1}{2}\left(\varepsilon_r + \sqrt{\varepsilon_r^2 + \varepsilon_i^2}\right)} \quad (\text{A.92})$$

$$\kappa = \sqrt{\frac{1}{2}\left(-\varepsilon_r + \sqrt{\varepsilon_r^2 + \varepsilon_i^2}\right)}. \quad (\text{A.93})$$

A.13 Herleitung der z-Komponenten der Felder und der Modenkopplung im Gitter in der RCWA

Einsetzen der Feldansätze in die Maxwell-Gleichung liefert für homogene Permeabilitäten und Permittivitäten folgende Grundgleichungen:

$$\frac{\partial}{\partial y} S_{z,mn} e^{i(k_x, mn x + k_y, mn y)} - \frac{\partial}{\partial z} S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} = i k_0 \mu U_{x,mn} e^{i(k_x, mn x + k_y, mn y)} \quad (\text{A.94})$$

$$\frac{\partial}{\partial z} S_{x,mn} e^{i(k_x, mn x + k_y, mn y)} - \frac{\partial}{\partial x} S_{z,mn} e^{i(k_x, mn x + k_y, mn y)} = i k_0 \mu U_{y,mn} e^{i(k_x, mn x + k_y, mn y)} \quad (\text{A.95})$$

$$\frac{\partial}{\partial x} S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} - \frac{\partial}{\partial y} S_{x,mn} e^{i(k_x, mn x + k_y, mn y)} = i k_0 \mu U_{z,mn} e^{i(k_x, mn x + k_y, mn y)} \quad (\text{A.96})$$

$$\frac{i}{Z_0} \frac{\partial}{\partial y} U_{z,mn} e^{i(k_x, mn x + k_y, mn y)} - \frac{i}{Z_0} \frac{\partial}{\partial z} U_{y,mn} e^{i(k_x, mn x + k_y, mn y)} = -i\omega\varepsilon\varepsilon_0 S_{x,mn} e^{i(k_x, mn x + k_y, mn y)} \quad (\text{A.97})$$

$$\frac{i}{Z_0} \frac{\partial}{\partial z} U_{x,mn} e^{i(k_x, mn x + k_y, mn y)} - \frac{i}{Z_0} \frac{\partial}{\partial x} U_{z,mn} e^{i(k_x, mn x + k_y, mn y)} = -i\omega\varepsilon\varepsilon_0 S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} \quad (\text{A.98})$$

$$\frac{i}{Z_0} \frac{\partial}{\partial x} U_{y,mn} e^{i(k_x, mn x + k_y, mn y)} - \frac{i}{Z_0} \frac{\partial}{\partial y} U_{x,mn} e^{i(k_x, mn x + k_y, mn y)} = -i\omega\varepsilon\varepsilon_0 S_{z,mn} e^{i(k_x, mn x + k_y, mn y)} \quad (\text{A.99})$$

Hierbei wurde die Ersetzung $\frac{\omega\mu_0}{Z_0} = k_0$ vorgenommen. Multipliziert man die letzten drei Gleichungen mit $-iZ_0$ und ersetzt $\omega\varepsilon_0 Z_0 = k_0$, dann sieht man, dass die Grundgleichungen symmetrisch bezüglich der gleichzeitigen Vertauschung von $S \leftrightarrow U$ und $\varepsilon \leftrightarrow \mu$ sind.

A.13.1 Bestimmung der z-Komponenten im Gitter

Für $S_{z,mn}$ folgt aus den Grundgleichungen

$$S_{z,mn} e^{i(k_x, mn x + k_y, mn y)} = \frac{i}{\omega\varepsilon_0} \varepsilon^{-1} \left(k_{x,mn} \frac{\text{ii}}{Z_0} U_{y,mn} - k_{y,mn} \frac{\text{ii}}{Z_0} U_{x,mn} \right) e^{i(k_x, mn x + k_y, mn y)} \quad (\text{A.100})$$

$$= \frac{i}{k_0} \varepsilon^{-1} (\text{ii} k_{x,mn} U_{y,mn} - \text{ii} k_{y,mn} U_{x,mn}) e^{i(k_x, mn x + k_y, mn y)}. \quad (\text{A.101})$$

Hierbei ist $\frac{1}{\omega\varepsilon_0 Z_0} = \frac{1}{k_0}$ eingesetzt worden.

Nun wird eine Entwicklung für die Epsilon-Verteilung eingesetzt, somit können inhomogene Permittivitäten beschrieben werden:

$$\sum_{mn} S_{z,mn} e^{i(k_x, mn x + k_y, mn y)} = \frac{i}{k_0} \left(\sum_{m,n} \tilde{\varepsilon}_{mn}^{-1} e^{2\pi i \left(\frac{m}{P_x} x + \frac{n}{P_y} y \right)} \right) \cdot \left(\sum_{m,n} (\text{ii} k_{x,mn} U_{y,mn} - \text{ii} k_{y,mn} U_{x,mn}) e^{i(k_x, mn x + k_y, mn y)} \right) \quad (\text{A.102})$$

$$\sum_{mn} S_{z,mn} e^{i(k_x, 0x + k_y, 0y)} e^{i \left(m \frac{2\pi}{P_x} x + n \frac{2\pi}{P_y} y \right)} = \frac{i}{k_0} \left(\sum_{m,n} \tilde{\varepsilon}_{mn}^{-1} e^{2\pi i \left(\frac{m}{P_x} x + \frac{n}{P_y} y \right)} \right) \cdot \left(\sum_{m,n} (\text{ii} k_{x,mn} U_{y,mn} - \text{ii} k_{y,mn} U_{x,mn}) e^{i \left(m \frac{2\pi}{P_x} x + n \frac{2\pi}{P_y} y \right)} \right) e^{i(k_x, 0x + k_y, 0y)} \quad (\text{A.103})$$

$$\begin{aligned}
& \sum_{mn} S_{z,mn} e^{i(k_x, mn x + k_y, mn y)} = \\
& \frac{i}{k_0} \sum_{m,n} \left(\sum_{p,q} \tilde{\varepsilon}_{m-pn-q}^{-1} (\text{ii} k_{x,pq} U_{y,pq} - \text{ii} k_{y,pq} U_{x,pq}) \right) e^{i \left(m \frac{2\pi}{P_x} x + n \frac{2\pi}{P_y} y \right)} e^{i(k_x, 0x + k_y, 0y)}
\end{aligned} \tag{A.104}$$

Das Endergebnis für $S_{z,mn}$ ist:

$$\begin{aligned}
S_{z,mn} e^{i(k_x, mn x + k_y, mn y)} = \\
\frac{i}{k_0} \left(\sum_{p,q} \tilde{\varepsilon}_{m-pn-q}^{-1} (\text{ii} k_{x,pq} U_{y,pq} - \text{ii} k_{y,pq} U_{x,pq}) \right) e^{i(k_x, mn x + k_y, mn y)}
\end{aligned} \tag{A.105}$$

Analog zur Bestimmung von $S_{z,mn}$ folgt für $U_{z,mn}$:

$$U_{z,mn} e^{i(k_x, mn x + k_y, mn y)} \tag{A.106}$$

$$= \frac{\text{ii}}{k_0} \mu^{-1} \left(\frac{\partial}{\partial x} S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} - \frac{\partial}{\partial y} S_{x,mn} e^{i(k_x, mn x + k_y, mn y)} \right) \tag{A.107}$$

$$= \frac{\text{ii}}{k_0} \mu^{-1} \left(\text{ii} k_{x,mn} S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} - \text{ii} k_{y,mn} S_{x,mn} e^{i(k_x, mn x + k_y, mn y)} \right) \tag{A.108}$$

$$= -\frac{i}{k_0} \sum_{q,p} \tilde{\mu}_{m-p,n-q}^{-1} (k_{x,pq} S_{y,pq} - k_{y,pq} S_{x,pq}) e^{i(k_x, mn x + k_y, mn y)} \tag{A.109}$$

Die Ergebnisse von $S_{z,mn}$ und $U_{z,mn}$ sind symmetrisch bezüglich der Vertauschung $S \leftrightarrow U$ und $\varepsilon \leftrightarrow \mu$.

A.13.2 Bestimmung der z-Ableitung der lateralen Feldkomponenten

Bestimmung von $\frac{\partial}{\partial z} S_{y,mn}$

$$\begin{aligned}
\frac{\partial}{\partial z} S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} = -\text{ii} \mu k_0 U_{x,mn} e^{i(k_x, mn x + k_y, mn y)} \\
+ \frac{\partial}{\partial y} S_{z,mn} e^{i(k_x, mn x + k_y, mn y)}
\end{aligned} \tag{A.110}$$

Durch Einsetzen von $S_{z,mn}$ erhält man:

$$\begin{aligned}
\frac{\partial}{\partial z} S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} = -\text{ii} \mu k_0 U_{x,mn} e^{i(k_x, mn x + k_y, mn y)} \\
+ \frac{\partial}{\partial y} \frac{i}{k_0} \left(\sum_{p,q} \tilde{\varepsilon}_{m-pn-q}^{-1} (\text{ii} k_{x,pq} U_{y,pq} - \text{ii} k_{y,pq} U_{x,pq}) \right) e^{i(k_x, mn x + k_y, mn y)}
\end{aligned} \tag{A.111}$$

Ausführen der Ableitung mit Produktregel führt zu:

$$\begin{aligned}
\frac{\partial}{\partial z} S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} = -\text{ii} \mu k_0 U_{x,mn} e^{i(k_x, mn x + k_y, mn y)} \\
+ \frac{\text{ii}}{k_0} k_{y,mn} \sum_{p,q} \tilde{\varepsilon}_{m-pn-q}^{-1} (\text{ii} k_{x,pq} U_{y,pq} - \text{ii} k_{y,pq} U_{x,pq}) e^{i(k_x, mn x + k_y, mn y)}
\end{aligned} \tag{A.112}$$

Durch Vereinfachen mit $ii = -1$ ergibt sich:

$$\begin{aligned} \frac{\partial}{\partial z} S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} &= -ii k_0 \mu U_{x,mn} e^{i(k_x, mn x + k_y, mn y)} \\ &\quad - \frac{ii}{k_0} k_{y,mn} \sum_{p,q} \tilde{\varepsilon}_{m-pn-q}^{-1} (k_{x,pq} U_{y,pq} - k_{y,pq} U_{x,pq}) e^{i(k_x, mn x + k_y, mn y)} \end{aligned} \quad (\text{A.113})$$

Mit Einführung der Entwicklung von μ erhält man:

$$\begin{aligned} \frac{\partial}{\partial z} S_{y,mn} e^{i(k_x, mn x + k_y, mn y)} &= -ii k_0 \sum_{p,q} \tilde{\mu}_{m-p, n-q} U_{x,pq} e^{i(k_x, mn x + k_y, mn y)} \\ &\quad - \frac{ii}{k_0} k_{y,mn} \sum_{p,q} \tilde{\varepsilon}_{m-pn-q}^{-1} (k_{x,pq} U_{y,pq} - k_{y,pq} U_{x,pq}) e^{i(k_x, mn x + k_y, mn y)} \end{aligned} \quad (\text{A.114})$$

In Matrix-Schreibweise mit den normierten k -Vektoren kann die z -Ableitung von $S_{y,mn}$ geschrieben werden als:

$$\frac{\partial}{\partial z} \vec{S}_y = -ii k_0 [\tilde{\boldsymbol{\mu}}] \vec{U}_x - ii k_0 (\mathbf{K}_y [\tilde{\boldsymbol{\varepsilon}}]^{-1} \mathbf{K}_x \vec{U}_y - \mathbf{K}_y [\tilde{\boldsymbol{\varepsilon}}]^{-1} \mathbf{K}_y \vec{U}_x) \quad (\text{A.115})$$

$$\frac{\partial}{\partial z} \vec{S}_y = k_0 [\tilde{\boldsymbol{\mu}}] \vec{U}_x + k_0 (\mathbf{K}_y [\tilde{\boldsymbol{\varepsilon}}]^{-1} \mathbf{K}_x \vec{U}_y - \mathbf{K}_y [\tilde{\boldsymbol{\varepsilon}}]^{-1} \mathbf{K}_y \vec{U}_x). \quad (\text{A.116})$$

Bestimmung von $\frac{\partial}{\partial z} S_{x,mn}$

$$\frac{\partial}{\partial z} S_{x,mn} e^{i(k_x, mn x + k_y, mn y)} = ii k_0 \mu U_{y,mn} e^{i(k_x, mn x + k_y, mn y)} + \frac{\partial}{\partial x} S_{z,mn} e^{i(k_x, mn x + k_y, mn y)} \quad (\text{A.117})$$

Durch Einsetzen von $S_{z,mn}$ erhält man:

$$\begin{aligned} \frac{\partial}{\partial z} S_{x,mn} e^{i(k_x, mn x + k_y, mn y)} &= ii k_0 \mu U_{y,mn} e^{i(k_x, mn x + k_y, mn y)} \\ &\quad + \frac{\partial}{\partial x} \frac{i}{k_0} \sum_{p,q} \tilde{\varepsilon}_{m-p, n-q}^{-1} (ii k_{x,pq} U_{y,pq} - ii k_{y,pq} U_{x,pq}) e^{i(k_x, mn x + k_y, mn y)} \end{aligned} \quad (\text{A.118})$$

Ausführen der Ableitung mit Produktregel, vereinfachen mit $ii = -1$ und Einsetzen der Entwicklung von μ ergibt:

$$\begin{aligned} \frac{\partial}{\partial z} S_{x,mn} e^{i(k_x, mn x + k_y, mn y)} &= ii k_0 \sum_{p,q} \tilde{\mu}_{m-p, n-q} U_{y,pq} e^{i(k_x, mn x + k_y, mn y)} \\ &\quad - \frac{ii}{k_0} k_{x,mn} \sum_{p,q} \tilde{\varepsilon}_{m-p, n-q}^{-1} (k_{x,pq} U_{y,pq} - k_{y,pq} U_{x,pq}) e^{i(k_x, mn x + k_y, mn y)} \end{aligned} \quad (\text{A.119})$$

In Matrix-Schreibweise ergibt sich:

$$\frac{\partial}{\partial z} \vec{S}_x = +ii k_0 [\tilde{\boldsymbol{\mu}}] \vec{U}_y + ii k_0 (\mathbf{K}_x [\tilde{\boldsymbol{\varepsilon}}]^{-1} \mathbf{K}_y \vec{U}_x - \mathbf{K}_x [\tilde{\boldsymbol{\varepsilon}}]^{-1} \mathbf{K}_x \vec{U}_y) \quad (\text{A.120})$$

$$\frac{\partial}{\partial z} \vec{S}_x = -k_0 [\tilde{\boldsymbol{\mu}}] \vec{U}_y - k_0 (\mathbf{K}_x [\tilde{\boldsymbol{\varepsilon}}]^{-1} \mathbf{K}_y \vec{U}_x - \mathbf{K}_x [\tilde{\boldsymbol{\varepsilon}}]^{-1} \mathbf{K}_x \vec{U}_y). \quad (\text{A.121})$$

Bestimmung von $\frac{\partial}{\partial z} U_{x,mn}$ und $\frac{\partial}{\partial z} U_{y,mn}$

Aufgrund der Symmetrie der Grundgleichungen lassen sich die Vertauschungsoperationen $S \leftrightarrow U$ und $\varepsilon \leftrightarrow \mu$ auf die z-Ableitung anwenden, somit werden aus

$$\frac{\partial}{\partial z} \vec{S}_x = -k_0 [\tilde{\mu}] \vec{U}_y - k_0 (\mathbf{K}_x [\tilde{\varepsilon}]^{-1} \mathbf{K}_y \vec{U}_x - \mathbf{K}_x [\tilde{\varepsilon}]^{-1} \mathbf{K}_x \vec{U}_y) \quad (\text{A.122})$$

$$\frac{\partial}{\partial z} \vec{S}_y = k_0 [\tilde{\mu}] \vec{U}_x + k_0 (\mathbf{K}_y [\tilde{\varepsilon}]^{-1} \mathbf{K}_x \vec{U}_y - \mathbf{K}_y [\tilde{\varepsilon}]^{-1} \mathbf{K}_y \vec{U}_x) \quad (\text{A.123})$$

folgende Gleichungen:

$$\frac{\partial}{\partial z} \vec{U}_x = -k_0 [\tilde{\varepsilon}] \vec{S}_y - k_0 (\mathbf{K}_x [\tilde{\mu}]^{-1} \mathbf{K}_y \vec{S}_x - \mathbf{K}_x [\tilde{\mu}]^{-1} \mathbf{K}_x \vec{S}_y) \quad (\text{A.124})$$

$$\frac{\partial}{\partial z} \vec{U}_y = k_0 [\tilde{\varepsilon}] \vec{S}_x + k_0 (\mathbf{K}_y [\tilde{\mu}]^{-1} \mathbf{K}_x \vec{S}_y - \mathbf{K}_y [\tilde{\mu}]^{-1} \mathbf{K}_y \vec{S}_x). \quad (\text{A.125})$$

Die Gleichungen A.122 - A.125 beschreiben die Kopplung der tangentialen Feldkomponenten im Gitter.

A.14 Herleitung der Intensität im Fokus für spezielle analytische Wellen

Sommerfeld-Welle

Gegeben sei die Sommerfeld-Welle

$$U(r_\perp) = \frac{f}{2\pi} U_0 \frac{e^{-ik\sqrt{r_\perp^2+f^2}}}{\sqrt{f^2+r_\perp^2}} \left(\frac{1}{f^2+r_\perp^2} + \frac{ik}{\sqrt{f^2+r_\perp^2}} \right) \quad (\text{A.126})$$

$$= \frac{f}{2\pi} U_0 \frac{e^{-ik\sqrt{r_\perp^2+f^2}}}{f^2+r_\perp^2} \left(\frac{1}{\sqrt{f^2+r_\perp^2}} + ik \right). \quad (\text{A.127})$$

Hinter einer Apertur mit Radius R folgt für das ausgebreitete Feld mit dem Rayleigh-Sommerfeld Integral I:

$$U(\mathbf{r}) = -\frac{1}{2\pi} \int_{A'} U(\mathbf{r}'_\perp, 0) \frac{d}{dz} \left(\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right) d^2\mathbf{r}'_\perp \quad (\text{A.128})$$

$$= -\frac{1}{2\pi} \int_{A'} U(\mathbf{r}'_\perp, 0) \frac{z e^{+ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|^2} \left(ik - \frac{1}{|\mathbf{r}-\mathbf{r}'|} \right) d^2\mathbf{r}'_\perp \quad (\text{A.129})$$

Auf der optischen Achse vereinfacht sich dies für radial symmetrische Anfangsverteilungen zu:

$$U(0, z) = -\int_0^R U(r'_\perp, 0) \frac{z e^{+ik\sqrt{z^2+r'^2_\perp}}}{z^2+r'^2_\perp} \left(ik - \frac{1}{\sqrt{z^2+r'^2_\perp}} \right) r'_\perp dr'_\perp \quad (\text{A.130})$$

Im Punkt $z = f$ gilt für die Sommerfeld-Welle

$$\begin{aligned}
U(0,f) &= - \int_0^R U(r'_\perp,0) \frac{f e^{+ik\sqrt{f^2+r'^2_\perp}}}{f^2 + r'^2_\perp} \left(ik - \frac{1}{\sqrt{f^2 + r'^2_\perp}} \right) r'_\perp dr'_\perp \\
&= - \int_0^R \frac{f}{2\pi} U_0 \frac{e^{-ik\sqrt{r'^2_\perp+f^2}}}{f^2 + r'^2_\perp} \left(\frac{1}{\sqrt{f^2 + r'^2_\perp}} + ik \right) \\
&\quad \frac{f e^{+ik\sqrt{f^2+r'^2_\perp}}}{f^2 + r'^2_\perp} \left(ik - \frac{1}{\sqrt{f^2 + r'^2_\perp}} \right) r'_\perp dr'_\perp \\
&= - \int_0^R \frac{f^2}{2\pi} U_0 \frac{1}{(f^2 + r'^2_\perp)^2} \left(\frac{1}{\sqrt{f^2 + r'^2_\perp}} + ik \right) \left(ik - \frac{1}{\sqrt{f^2 + r'^2_\perp}} \right) r'_\perp dr'_\perp \\
&= \int_0^R \frac{f^2}{2\pi} U_0 \frac{1}{(f^2 + r'^2_\perp)^2} \left(\frac{1}{f^2 + r'^2_\perp} + k^2 \right) r'_\perp dr'_\perp \\
&= \int_0^R \frac{f^2}{2\pi} U_0 \frac{1}{(f^2 + r'^2_\perp)^3} r'_\perp dr'_\perp + \int_0^R \frac{f^2 k^2}{2\pi} U_0 \frac{1}{(f^2 + r'^2_\perp)^2} r'_\perp dr'_\perp \\
&= \frac{f^2}{2\pi} U_0 \frac{1}{4} \left(\frac{1}{f^4} - \frac{1}{(f^2 + R^2)^2} \right) + \frac{f^2 k^2}{2\pi} U_0 \frac{R^2}{2f^2(f^2 + R^2)} \\
&= U_0 \left(\frac{1}{8\pi f^2} - \frac{f^2}{8\pi(f^2 + R^2)^2} + k^2 \frac{R^2}{4\pi(f^2 + R^2)} \right). \tag{A.131}
\end{aligned}$$

Kugel-Welle

Die Kugel-Welle ist definiert als:

$$U(r_\perp) = U_0 f \frac{e^{-ik\sqrt{r_\perp^2+f^2}}}{\sqrt{r_\perp^2 + f^2}} \tag{A.132}$$

Hinter einer Apertur mit Radius R folgt für das ausgebreitete Feld mit dem Rayleigh-Sommerfeld Integral I:

$$\begin{aligned}
U(0,f) &= - \int_0^R U(r'_\perp,0) \frac{f e^{+ik\sqrt{f^2+r'^2_\perp}}}{f^2 + r'^2_\perp} \left(ik - \frac{1}{\sqrt{f^2 + r'^2_\perp}} \right) r'_\perp dr'_\perp \\
&= - \int_0^R U_0 f \frac{1}{\sqrt{r'^2_\perp + f^2}} \frac{f}{r'^2_\perp + f^2} \left(ik - \frac{1}{\sqrt{r'^2_\perp + f^2}} \right) r'_\perp dr'_\perp \\
&= -U_0 \int_0^R \frac{f^2}{(r'^2_\perp + f^2)^{\frac{3}{2}}} \left(ik - \frac{1}{\sqrt{r'^2_\perp + f^2}} \right) r'_\perp dr'_\perp \\
&= -U_0 \int_0^R \frac{ik f^2}{(r'^2_\perp + f^2)^{\frac{3}{2}}} r'_\perp dr'_\perp + U_0 \int_0^R \frac{f^2}{(r'^2_\perp + f^2)^{\frac{3}{2}}} r'_\perp dr'_\perp \\
&= U_0 \left(-ikf \left(1 - \frac{1}{\sqrt{\frac{R^2}{f^2} + 1}} \right) + \frac{R^2}{2(R^2 + f^2)} \right) \tag{A.133}
\end{aligned}$$

Somit gilt für die Intensität im Fokus

$$I(0,f) = I_0 \left(k^2 f^2 \left(1 - \frac{1}{\sqrt{\frac{R^2}{f^2} + 1}} \right)^2 + \frac{R^4}{4(R^2 + f^2)^2} \right). \tag{A.134}$$

Sphärische Phase

Die sphärische Phase ist definiert als:

$$U(r_\perp) = U_0 e^{-ik\sqrt{r_\perp^2 + f^2}} \tag{A.135}$$

Hinter einer Apertur mit Radius R folgt für das ausgebreitete Feld mit dem Rayleigh-Sommerfeld Integral I:

$$\begin{aligned}
U(0,f) &= - \int_0^R U(r'_\perp,0) \frac{f e^{+ik\sqrt{f^2+r'^2_\perp}}}{f^2 + r'^2_\perp} \left(ik - \frac{1}{\sqrt{f^2 + r'^2_\perp}} \right) r'_\perp dr'_\perp \\
&= -U_0 \int_0^R \frac{f}{f^2 + r'^2_\perp} \left(ik - \frac{1}{\sqrt{f^2 + r'^2_\perp}} \right) r'_\perp dr'_\perp \\
&= -U_0 \int_0^R \frac{ikf}{f^2 + r'^2_\perp} r'_\perp dr'_\perp + \int_0^R \frac{f}{(f^2 + r'^2_\perp)^{\frac{3}{2}}} r'_\perp dr'_\perp \\
&= U_0 \left(-ikf \frac{1}{2} \log(f^2 + R^2) + ikf \frac{1}{2} \log(f^2) + 1 - \frac{f}{\sqrt{f^2 + R^2}} \right) \\
&= U_0 \left(+ikf \frac{1}{2} \log\left(\frac{f^2}{f^2 + R^2}\right) + 1 - \frac{f}{\sqrt{f^2 + R^2}} \right) \\
&= U_0 \left(+ikf \log\left(\frac{f}{\sqrt{f^2 + R^2}}\right) + 1 - \frac{f}{\sqrt{f^2 + R^2}} \right) \tag{A.136}
\end{aligned}$$

Somit gilt für die Intensität im Fokus

$$I(0,f) = I_0 \left[1 - \frac{2f}{\sqrt{R^2 + f^2}} + \frac{f^2}{R^2 + f^2} + k^2 f^2 \log^2 \left(\frac{f}{\sqrt{R^2 + f^2}} \right) \right]. \quad (\text{A.137})$$

A.15 Bestimmung spezieller Überlappintegrale

Gegeben sei folgendes Flächenintegral:

$$\begin{aligned} B &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right) \text{circ} \left(\frac{\sqrt{\left(\nu - \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) d\nu d\mu \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{circ} \left(\frac{\sqrt{\nu^2 + \mu^2 \lambda}}{\text{NA}_{\text{Beob}}} \right) \text{circ} \left(\frac{\sqrt{\left(\nu + \frac{1}{P}\right)^2 + \mu^2 \lambda}}{\text{NA}_{\text{Bel}}} \right) d\nu d\mu \end{aligned} \quad (\text{A.138})$$

Die Integration über die Überlappfläche B lässt sich in drei Fälle aufteilen. Hierzu ist es von Vorteil, $\text{NA}_{\min} = \min(\text{NA}_{\text{Beob}}, \text{NA}_{\text{Bel}})$ und $\text{NA}_{\max} = \max(\text{NA}_{\text{Beob}}, \text{NA}_{\text{Bel}})$ zu setzen, um die numerischen Aperturen nach ihrer Größe zu unterscheiden.

1. $\frac{\text{NA}_{\min}}{\lambda} + \frac{1}{P} < \frac{\text{NA}_{\max}}{\lambda}$
2. $\frac{\text{NA}_{\min}}{\lambda} + \frac{1}{P} > \frac{\text{NA}_{\max}}{\lambda}$ und $\frac{1}{P} < \frac{\text{NA}_{\max}}{\lambda}$
3. $\frac{1}{P} > \frac{\text{NA}_{\max}}{\lambda}$ und $\frac{1}{P} < \frac{\text{NA}_{\max}}{\lambda} + \frac{\text{NA}_{\min}}{\lambda}$

Für den ersten Fall ist das Überlappintegral eine Kreisfläche und es gilt:

$$B^{\text{Fall 1}} = \frac{\text{NA}_{\min}^2}{\lambda^2} \pi. \quad (\text{A.139})$$

Für den zweiten Fall besteht die Fläche aus zwei Kreissegmenten und einem Vollkreis:

$$B^{\text{Fall 2}} = B_0^{\text{Fall 2}} - B_1^{\text{Fall 2}} + B_2^{\text{Fall 2}} \quad (\text{A.140})$$

$$B_0^{\text{Fall 2}} = \frac{\text{NA}_{\min}^2}{\lambda^2} \pi \quad (\text{A.141})$$

$$B_1^{\text{Fall 2}} = \frac{\text{NA}_{\min}^2}{\lambda^2} (\pi - \theta_{\text{NA}_{\min}} - \sin(\pi - \theta_{\text{NA}_{\min}}) \cos(\pi - \theta_{\text{NA}_{\min}})) \quad (\text{A.142})$$

$$B_2^{\text{Fall 2}} = \frac{\text{NA}_{\max}^2}{\lambda^2} (\theta_{\text{NA}_{\max}} - \sin(\theta_{\text{NA}_{\max}}) \cos(\theta_{\text{NA}_{\max}})). \quad (\text{A.143})$$

Hierbei sind $\theta_{\text{NA}_{\max}}$ und $\theta_{\text{NA}_{\min}}$ die Winkel im Dreieck aus den Kreismittelpunkten und dem Schnittpunkt der beiden Kreise.

Im dritten Fall lässt sich die Fläche in zwei Teilsegmente aufteilen. Die Fläche jedes Teilsegments lässt sich aus der Differenz eines Kreissektors und der Fläche zweier gleich großen rechtwinkligen Dreiecke berechnen:

$$B^{\text{Fall 3}} = B_1^{\text{Fall 3}} + B_2^{\text{Fall 3}} \quad (\text{A.144})$$

$$B_1^{\text{Fall 3}} = \frac{NA_{\min}^2}{\lambda^2} (\theta_{NA_{\min}} - \sin(\theta_{NA_{\min}}) \cos(\theta_{NA_{\min}})) \quad (\text{A.145})$$

$$B_2^{\text{Fall 3}} = \frac{NA_{\max}^2}{\lambda^2} (\theta_{NA_{\max}} - \sin(\theta_{NA_{\max}}) \cos(\theta_{NA_{\max}})). \quad (\text{A.146})$$

Die Winkel $\theta_{NA_{\max}}$ und $\theta_{NA_{\min}}$ lassen sich mit dem Kosinussatz aus dem Dreieck $\frac{1}{P}$, $\frac{NA_{\max}}{\lambda}$, $\frac{NA_{\min}}{\lambda}$ bestimmen:

$$\frac{NA_{\min}^2}{\lambda^2} = \frac{1}{P^2} + \frac{NA_{\max}^2}{\lambda^2} - 2 \frac{1}{P} \frac{NA_{\max}}{\lambda} \cos(\theta_{NA_{\max}}) \quad (\text{A.147})$$

$$\frac{NA_{\max}^2}{\lambda^2} = \frac{1}{P^2} + \frac{NA_{\min}^2}{\lambda^2} - 2 \frac{1}{P} \frac{NA_{\min}}{\lambda} \cos(\theta_{NA_{\min}}). \quad (\text{A.148})$$

B Literatur

- [1] M. Auer: „Numerical treatment of localized fields in rigorous diffraction theory and its application to light absorption in structured layers“. Diss. Ruprecht-Karls-Universität Heidelberg, 2016.
- [2] M. Auer: „Numerische Analyse optischer Nanostrukturen mit dreidimensionalen rigorosen Methoden“. Diplomarbeit. Ruprecht-Karls-Universität Heidelberg, 2009.
- [3] M. Auer und K.-H. Brenner: „Localized input fields in rigorous coupled-wave analysis“. In: *J. Opt. Soc. Am. A* Jg. 31, Nr. 11 (2014), S. 2385–2393.
- [4] R. Barakat: „Determination of the optical transfer function directly from the edge spread function“. In: *J. Opt. Soc. Am.* Jg. 55, Nr. 10 (1965), S. 1217–1219.
- [5] M. W. Beijersbergen, R. P. C. Coerwinkel, M. Kristensen und J. P. Woerdman: „Helical-wavefront laser beams produced with a spiral phaseplate“. In: *Opt. Commun.* Jg. 112, Nr. 5 (1994), S. 321–327.
- [6] D. C. Bertilone: „On the exact Kirchhoff and Rayleigh-Sommerfeld theories for the focusing of an infinite scalar spherical wave-field“. In: *Opt. Commun.* Jg. 85, Nr. 2 (1991), S. 153–158.
- [7] D. C. Bertilone: „Wave theory for a converging spherical incident wave in an infinite-aperture system“. In: *J. Mod. Opt.* Jg. 38, Nr. 8 (1991), S. 1531–1536.
- [8] A. Bondeson, T. Rylander und P. Ingelström: *Computational electromagnetics*. Bd. 51. Springer Science & Business Media, 2005.
- [9] M. Born und E. Wolf: *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Cambridge University Press, 2000.
- [10] J. J. M. Braat, S. van Haver, A. J. E. M. Janssen und P. Dirksen: „Assessment of optical systems by means of point-spread functions“. In: *Progress in Optics* Jg. 51 (2008), S. 349–468.
- [11] D. J. Brady und N. Hagen: „Multiscale lens design“. In: *Opt. Express* Jg. 17, Nr. 13 (2009), S. 10659–10674.
- [12] K.-H. Brenner: *Accuracy of Fitting an array of spots to a model (GRIDFIT)*. Techn. Ber. Chair of Optoelectronics, Institute of Computer Engineering, Heidelberg University, 2013.

- [13] K.-H. Brenner: „Aspects for calculating local absorption with the rigorous coupled-wave method“. In: *Opt. Express* Jg. 18, Nr. 10 (2010), S. 10369–10376.
- [14] K.-H. Brenner: „Shifted base functions: an efficient and versatile new tool in optics“. In: *Journal of Physics: Conference Series*. Bd. 139. IOP Publishing. 2008, S. 012002.
- [15] K.-H. Brenner: *Skript zur Vorlesung Computational Optics*. 2014.
- [16] K.-H. Brenner und M. Auer: „Verification of near-field calculations by conservation laws“. In: *2016 15th Workshop on Information Optics (WIO)*. IEEE. 2016, S. 1–3.
- [17] K.-H. Brenner und W. Singer: „Light propagation through microlenses: a new simulation method“. In: *Appl. Opt.* Jg. 32, Nr. 26 (1993), S. 4984–4988.
- [18] P. D. Burns und D. Williams: „Refined slanted-edge measurement for practical camera and scanner testing“. In: *IS and TS PICS Conference*. Society for Imaging Science & Technology. 2002, S. 191–195.
- [19] W. C. Chew, E. Michielssen, J. Song und J.-M. Jin: *Fast and efficient algorithms in computational electromagnetics*. Artech House Publishers, 2001.
- [20] S. K. Chow, H. Hakozaki, D. L. Price, N. A. MacLean, T. J. Deerinck, J. C. Bouwer, M. E. Martone, S. T. Peltier und M. H. Ellisman: „Automated microscopy system for mosaic acquisition and processing“. In: *J. Microsc.* Jg. 222, Nr. 2 (2006), S. 76–84.
- [21] G. Cox und C. J. R. Sheppard: „Practical limits of resolution in confocal and non-linear microscopy“. In: *Microsc. Res. Tech.* Jg. 63, Nr. 1 (2004), S. 18–22.
- [22] C. Cremer und T. Cremer: „Considerations on a laser-scanning-microscope with high resolution and depth of field“. In: *Microsc. Acta* (1974), S. 31–44.
- [23] P. Debye: „Das Verhalten von Lichtwellen in der Nähe eines Brennpunktes oder einer Brennlinie“. In: *Ann. Phys. (Leipzig)* Jg. 335, Nr. 14 (1909), S. 755–776.
- [24] G. T. Di Francia: „Super-gain antennas and optical resolving power“. In: *Il Nuovo Cimento (1943-1954)* Jg. 9 (1952), S. 426–438.
- [25] R. Dorn, S. Quabis und G. Leuchs: „Sharper focus for a radially polarized light beam“. In: *Phys. Rev. Lett.* Jg. 91, Nr. 23 (2003), S. 233901.
- [26] J. Durnin: „Exact solutions for nondiffracting beams. I. The scalar theory“. In: *J. Opt. Soc. Am. A* Jg. 4, Nr. 4 (1987), S. 651–654.
- [27] Y. M. Engelberg und S. Ruschin: „Fast method for physical optics propagation of high-numerical-aperture beams“. In: *J. Opt. Soc. Am. A* Jg. 21, Nr. 11 (2004), S. 2135–2145.

- [28] M. W. Farn und W. B. Veldkamp: „Binary optics“. In: *Handbook of Optics* Jg. 2 (1995), S. 8–1.
- [29] M. D. Feit und J. A. Fleck: „Light propagation in graded-index optical fibers“. In: *Appl. Opt.* Jg. 17, Nr. 24 (1978), S. 3990–3998.
- [30] M. Fertig und K.-H. Brenner: „Vector wave propagation method“. In: *J. Opt. Soc. Am. A* Jg. 27, Nr. 4 (2010), S. 709–717.
- [31] M. Fertig: „Vector Wave Propagation Method: Ein Beitrag zum elektromagnetischen Optikrechnen“. Diss. Universität Mannheim, 2011.
- [32] B. Furlong und S. Motakef: „Scanning lenses and systems“. In: *Photonik International* Jg. 2 (2008), S. 20–23.
- [33] H. Gahbiche: „Array von diffraktiven optischen Elementen zur Erzeugung ringförmiger Foki und Ortsbestimmung“. Diplomarbeit. Ruprecht-Karls-Universität Heidelberg, 2016.
- [34] J. W. Goodman: *Introduction to Fourier optics*. Roberts und Company Publishers, 2005.
- [35] H. S. Green und E. Wolf: „A scalar representation of electromagnetic fields“. In: *Pro. Phys. Soc. London, Sec. A* Jg. 66, Nr. 12 (1953), S. 1129.
- [36] D. J. Griffiths: „Resource letter EM-1: electromagnetic momentum“. In: *Am. J. Phys.* Jg. 80, Nr. 1 (2012), S. 7–18.
- [37] *GUPPY Technical Manual V.7.1.0*. Allied Vision Technologies GmbH, Taschenweg 21, D-07646 Stadtroda, 2009.
- [38] J. C. Gutiérrez-Vega, M. D. Iturbe-Castillo und S. Chávez-Cerda: „Alternative formulation for invariant optical fields: Mathieu beams“. In: *Opt. Lett.* Jg. 25, Nr. 20 (2000), S. 1493–1495.
- [39] G. R. Hadley: „Wide-angle beam propagation using Padé approximant operators“. In: *Opt. Lett.* Jg. 17, Nr. 20 (1992), S. 1426–1428.
- [40] B. Hulsken, D. Vossen und S. Stallinga: „High NA diffractive array illuminators and application in a multi-spot scanning microscope“. In: *J. Eur. Opt. Soc., Rapid. Publ.* Jg. 7 (2012).
- [41] G. D. Hutton: „An exact solution for the on-axis amplitude of a circularly symmetric aperture function“. In: *Opt. Commu.* Jg. 37, Nr. 6 (1981), S. 379–382.
- [42] *ISO 3334:2006, Micrographics — ISO resolution test chart No. 2 — Description and use*. Standard. Geneva, CH, 2006.
- [43] J. D. Jackson: *Classical electrodynamics*. Wiley, 1999.
- [44] J. Jahns: „Continuous and discrete diffractive elements with polar symmetries“. In: *App. Opt.* Jg. 56, Nr. 1 (2017), A1–A7.

- [45] J. Jahns und S. Helfert: *Introduction to micro-and nanooptics*. John Wiley & Sons, 2012.
- [46] J. Jahns und S. J. Walker: „Two-dimensional array of diffractive microlenses fabricated by thin film deposition“. In: *App. Opt.* Jg. 29, Nr. 7 (1990), S. 931–936.
- [47] A. Jesacher, S. Bernet und M. Ritsch-Marte: „Colored point spread function engineering for parallel confocal microscopy“. In: *Opt. Express* Jg. 24, Nr. 24 (2016), S. 27395–27402.
- [48] A. Jesacher, S. Bernet und M. Ritsch-Marte: „Combined holographic optical trapping and optical image processing using a single diffractive pattern displayed on a spatial light modulator“. In: *Opt. Lett.* Jg. 39, Nr. 18 (2014), S. 5337–5340.
- [49] J.-M. Jin: *The finite element method in electromagnetics*. John Wiley & Sons, 2014.
- [50] X. Jing und Y. Jin: „Transmittance analysis of diffraction phase grating“. In: *App. Opt.* Jg. 50, Nr. 9 (2011), S. C11–C18.
- [51] R. D. Juday: „Generality of matched filtering and minimum Euclidean distance projection for optical pattern recognition“. In: *J. Opt. Soc. Am. A* Jg. 18, Nr. 8 (2001), S. 1882–1896.
- [52] R. D. Juday: „Optimal realizable filters and the minimum Euclidean distance principle“. In: *App. Opt.* Jg. 32, Nr. 26 (1993), S. 5100–5111.
- [53] A. Junker. private Kommunikation. 2016.
- [54] B. A. Kemp: „Resolution of the Abraham-Minkowski debate: Implications for the electromagnetic wave theory of light in matter“. In: *J. Appl. Phys.* Jg. 109, Nr. 11 (2011), S. 111101.
- [55] B. A. Kemp: „Macroscopic Theory of Optical Momentum“. In: *Progress in Optics* Jg. 60 (2015), S. 437–488.
- [56] A. R. Khoei: *Extended finite element method: theory and applications*. John Wiley & Sons, 2014.
- [57] M. W. Kowarz: „Energy constraints in optimum apodization problems“. In: *Opt. Commu.* Jg. 110, Nr. 3-4 (1994), S. 274–278.
- [58] L. D. Landau, J. S. Bell, M. J. Kearsley, L. P. Pitaevskii, E. M. Lifshitz und J. B. Sykes: *Electrodynamics of continuous media*. Bd. 8. Elsevier, 2013.
- [59] R. G. Lane und M. Tallon: „Wave-front reconstruction using a Shack–Hartmann sensor“. In: *App. Opt.* Jg. 31, Nr. 32 (1992), S. 6902–6908.
- [60] L. Li: „Mathematical reflections on the Fourier modal method in grating theory“. In: *Mathematical modeling in optical science* Jg. 22 (2001), S. 111–139.

- [61] J. Liang, B. Grimm, S. Goelz und J. F. Bille: „Objective measurement of wave aberrations of the human eye with the use of a Hartmann–Shack wave-front sensor“. In: *J. Opt. Soc. Am. A* Jg. 11, Nr. 7 (1994), S. 1949–1957.
- [62] P. Liu und B. Lü: „The vectorial angular-spectrum representation and Rayleigh-Sommerfeld diffraction formulae“. In: *Opt. Laser Technol.* Jg. 39, Nr. 4 (2007), S. 741–744.
- [63] X. Liu und K.-H. Brenner: „Phase Retrieval with a Diffractive Micro Lens Array“. In: *Int. J. Optomechatronics* Jg. 8, Nr. 4 (2014), S. 251–259.
- [64] A. W. Lohmann: „Scaling laws for lens systems“. In: *App. Opt.* Jg. 28, Nr. 23 (1989), S. 4996–4998.
- [65] V. N. Mahajan: „Luneberg apodization problem I“. In: *Opt. Lett.* Jg. 5, Nr. 6 (1980), S. 267–269.
- [66] J. N. Mait: „Understanding diffractive optic design in the scalar domain“. In: *J. Opt. Soc. Am. A* Jg. 12, Nr. 10 (1995), S. 2145–2158.
- [67] L. Mandel und E. Wolf: *Optical coherence and quantum optics*. Cambridge University Press, 1995.
- [68] C. Max: „Introduction to adaptive optics and its history“. In: *NSF Center for Adaptive Optics, University of California at Santa Cruz, DOE Lawrence Livermore National Laboratory, American Astronomical Society 197th Meeting*. 2001.
- [69] C. W. Mccutchen: „Superresolution in microscopy and the Abbe resolution limit“. In: *J. Opt. Soc. Am.* Jg. 57, Nr. 10 (1967), S. 1190–1190.
- [70] S. D. Mellin: „Design and analysis of finite aperture diffractive optical elements“. Diss. The University of Alabama in Huntsville, 2001.
- [71] S. Mellin und G. Nordin: „Limits of scalar diffraction theory and an iterative angular spectrum algorithm for finite aperture diffractive optical element design“. In: *Opt. Express* Jg. 8, Nr. 13 (2001), S. 705–722.
- [72] *Microposit S1800 series photoresists*. Rohm und Haas Company, Philadelphia, PA, USA, 2006.
- [73] M. G. Moharam und T. K. Gaylord: „Rigorous coupled-wave analysis of planar-grating diffraction“. In: *J. Opt. Soc. Am.* Jg. 71, Nr. 7 (1981), S. 811–818.
- [74] M. G. Moharam und T. K. Gaylord: „Three-dimensional vector coupled-wave analysis of planar-grating diffraction“. In: *J. Opt. Soc. Am.* Jg. 73, Nr. 9 (1983), S. 1105–1112.
- [75] M. G. Moharam, T. K. Gaylord, E. B. Grann und D. A. Pommet: „Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings“. In: *J. Opt. Soc. Am. A* Jg. 12, Nr. 5 (1995), S. 1068–1076.

- [76] M. G. Moharam, T. K. Gaylord, D. A. Pommet und E. B. Grann: „Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach“. In: *J. Opt. Soc. Am. A* Jg. 12, Nr. 5 (1995), S. 1077–1086.
- [77] I. Moreno, C. Iemmi, A. Márquez, J. Campos und M. J. Yzuel: „Modulation light efficiency of diffractive lenses displayed in a restricted phase-mostly modulation display“. In: *App. Opt.* Jg. 43, Nr. 34 (2004), S. 6278–6284.
- [78] S. Noach, A. Lewis, Y. Arieli und N. Eisenberg: „Integrated diffractive and refractive elements for spectrum shaping“. In: *App. Opt.* Jg. 35, Nr. 19 (1996), S. 3635–3639.
- [79] L. Novotny und B. Hecht: *Principles of nano-optics*. Cambridge University Press, 2012.
- [80] A. Orth und K. Crozier: „Microscopy with microlens arrays: high throughput, high resolution and light-field imaging“. In: *Opt. Express* Jg. 20, Nr. 12 (2012), S. 13522–13531.
- [81] H. Osterberg und L. W. Smith: „Closed solutions of Rayleigh’s diffraction integral for axial points“. In: *J. Opt. Soc. Am.* Jg. 51, Nr. 10 (1961), S. 1050–1054.
- [82] B. C. Platt und R. Shack: „History and principles of Shack-Hartmann wave-front sensing“. In: *J. Refract. Surg.* Jg. 17, Nr. 5 (2001), S573–S577.
- [83] D. A. Pommet, M. G. Moharam und E. B. Grann: „Limits of scalar diffraction theory for diffractive phase elements“. In: *J. Opt. Soc. Am. A* Jg. 11, Nr. 6 (1994), S. 1827–1834.
- [84] J. M. Rodenburg, A. C. Hurst und A. G. Cullis: „Transmission microscopy without lenses for objects of unlimited size“. In: *Ultramicroscopy* Jg. 107, Nr. 2 (2007), S. 227–231.
- [85] J. K. M. Roland: „A study of slanted-edge MTF stability and repeatability“. In: *SPIE/IS&T Electronic Imaging*. International Society for Optics und Photonics. 2015, S. 93960L–93960L.
- [86] C. J. R. Sheppard und A. Choudhury: „Image formation in the scanning microscope“. In: *J. Mod. Opt.* Jg. 24, Nr. 10 (1977), S. 1051–1073.
- [87] D. M. Shotton: „Confocal scanning optical microscopy and its applications for biological specimens“. In: *J. Cell Sci.* Jg. 94, Nr. 2 (1989), S. 175–206.
- [88] E. Slognat: „Aspects and implementations for accelerating image acquisition in microscopy“. Diss. Ruprecht-Karls-Universität Heidelberg, 2013.
- [89] W. H. Southwell: „Wave-front estimation from wave-front slope measurements“. In: *J. Opt. Soc. Am.* Jg. 70, Nr. 8 (1980), S. 998–1006.
- [90] J. J. Stamnes: „Focusing of a perfect wave and the Airy pattern formula“. In: *Opt. Commu.* Jg. 37, Nr. 5 (1981), S. 311–314.

- [91] T. Stenau und K.-H. Brenner: *Design of a diffractive micro lens array for Shack-Hartmann wavefront sensing*. Techn. Ber. Chair of Optoelectronics, Institute of Computer Engineering, Heidelberg University, 2012.
- [92] V. Strassen: „Gaussian elimination is not optimal“. In: *Numer. Math.* Jg. 13, Nr. 4 (1969), S. 354–356.
- [93] A. Taflove und S. C. Hagness: *Computational electrodynamics*. Artech House Publishers, 2000.
- [94] H. F. Talbot: „LXXVI. Facts relating to optical science. No. IV“. In: *Philos. Mag.* Jg. 9, Nr. 56 (1836), S. 401–407.
- [95] H. J. Tiziani, R. Achi, R. Krämer und L. Wieggers: „Theoretical analysis of confocal microscopy with microlenses“. In: *App. Opt.* Jg. 35, Nr. 1 (1996), S. 120–125.
- [96] B. Trauter: „Untersuchungen zum inversen, elektromagnetischen Gitterbeugungsproblem mit Anwendungen in der Gittercharakterisierung“. Diss. Universität Mannheim, 2009.
- [97] U. Van Rienen: *Numerical methods in computational electrodynamics: linear systems in practical applications*. Bd. 12. Springer Science & Business Media, 2012.
- [98] J. Wangüemert-Pérez, I. Molina-Fernández und M. Luque-Nieto: „A novel Fourier based 3D full-vectorial beam propagation method“. In: *Opt. Quant. Electron.* Jg. 36, Nr. 1-3 (2004), S. 285–301.
- [99] F. E. Washer und I. C. Gardner: *Method for determining the resolving power of photographic lenses*. Bd. 13. US Department of Commerce, National Bureau of Standards, 1973.
- [100] W. H. Welch, J. E. Morris und M. R. Feldman: „Iterative discrete on-axis encoding of radially symmetric computer-generated holograms“. In: *J. Opt. Soc. Am. A* Jg. 10, Nr. 8 (Aug. 1993), S. 1729–1738.
- [101] A. Weyer: „Untersuchung zur Erhöhung der Genauigkeit eines Hartmann-Shack-Sensors mit diffraktiven Linsen“. Diplomarbeit. Ruprecht-Karls-Universität Heidelberg, 2015.
- [102] H. Weyl: „Ausbreitung elektromagnetischer Wellen über einem ebenen Leiter“. In: *Annalen der Physik* Jg. 365, Nr. 21 (1919), S. 481–500.
- [103] T. Wilson und C. Sheppard: *Theory and practice of scanning optical microscopy*. Bd. 180. Academic Press London, 1984.
- [104] P. W. Winter und H. Shroff: „Faster fluorescence microscopy: advances in high speed biological imaging“. In: *Curr. Opin. Chem. Biol.* Jg. 20 (2014), S. 46–53.
- [105] E. Wolf: „A scalar representation of electromagnetic fields: II“. In: *Proc. Phys. Soc. London* Jg. 74, Nr. 3 (1959), S. 269.

- [106] C. L. Xu und W. P. Huang: „Finite-difference beam propagation method for guide-wave optics“. In: *Progress In Electromagnetics Research* Jg. 11 (1995), S. 1–49.
- [107] G. Zheng, R. Horstmeyer und C. Yang: „Wide-field, high-resolution Fourier ptychographic microscopy“. In: *Nat. Photon.* Jg. 7, Nr. 9 (2013), S. 739–745.

C Eigene Publikationsliste

Folgende Publikationen sind während der Promotionszeit entstanden:

- [A1] X. Liu, T. Stenau und K.-H. Brenner: „Diffractive micro lens arrays with overlapping apertures“. In: *2012 11th Euro-American Workshop on Information Optics*. 2012.
- [A2] K.-H. Brenner, T. Stenau und M. Azizian: „Entwicklung eines scannenden Mikroskops mit diffraktiven Mikrolinsen“. In: *DGaO-Proceedings*. 2013.
- [A3] T. Stenau und K.-H. Brenner: „Analyse der Designparameter von binären diffraktiven Mikrolinsenarrays in Hinblick auf den Herstellungsprozess“. In: *DGaO-Proceedings*. 2013.
- [A4] A. Junker, T. Stenau und K.-H. Brenner: „Scalar wave-optical reconstruction of plenoptic camera images“. In: *App. Opt.* Jg. 53, Nr. 25 (2014), S. 5784–5790.
- [A5] A. Junker, T. Stenau und K.-H. Brenner: „Wave-Optical Reconstruction of Plenoptic Camera Images“. In: *Fringe 2013*. Springer, 2014, S. 117–122.
- [A6] T. Stenau und K.-H. Brenner: „Diffractive Lenses with Overlapping Aperture – A New Tool in Scanning Microscopy“. In: *Imaging Systems and Applications*. Optical Society of America. 2016, IT1F–1.
- [A7] T. Stenau und K.-H. Brenner: „Light concentration efficiency of diffractive lenses with overlapping apertures“. In: *DGaO-Proceedings*. 2016.

Einzelne Aspekte des Designs und der Evaluation von diffraktiven Linsen mit überlappenden Aperturen wurden in den Arbeiten [A2, A1, A3, A6, A7] publiziert. Die Ergebnisse der experimentellen Vermessung wurden in [A6] veröffentlicht. Der Aufbau und die Funktionsweise des parallel scannenden Mikroskops mit diffraktiven Linsen mit überlappenden Aperturen wurde in den Arbeiten [A2, A7] vorgestellt. In den Publikationen [A4, A5] wurden Teile der für das Scanning-Mikroskop entwickelten Software zur Justierung und zur Bildaufnahme der plenoptischen Kamera verwendet.

Danksagung

An erster Stelle möchte ich mich bei Herrn Prof. Dr. Karl-Heinz Brenner für die Betreuung und die umfassende Unterstützung während meiner Promotion bedanken. Als Zweites möchte ich mich bei Herrn Prof. Dr. Jürgen Jahns für die Übernahme des Zweitgutachtens der Arbeit bedanken.

Ein weiterer Dank geht an die aktuellen und ehemaligen Mitarbeiter des Lehrstuhls für Optoelektronik am ZITI, die dafür gesorgt haben, dass ich mich dort immer wohl gefühlt habe. Insbesondere sind hier Sabine Volk für ihre Unterstützung bei Verwaltungsangelegenheiten und Wolfgang Stumpfs für seine Hilfe bei mechanischen Problemen jeglicher Art zu erwähnen. Außerdem möchte ich mich bei André Junker, Maximilian Auer und Xiyuan Liu für die interessanten Diskussionen und ihre Hilfe bedanken.