

Projektionsalgorithmen im wellenoptischen Transmissionsdesign

Dissertation

**zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)**

vorgelegt dem Rat der Physikalisch-Astronomischen Fakultät
der FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

von Dipl.-Phys. Dipl.-Inf. Sven Bühling
geboren am 4. Dezember 1972 in Jena

Gutachter:

1. Prof. Dr. Frank Wyrowski
2. Prof. Dr. Stefan Sinzinger
3. Prof. Dr. Hartmut Bartelt

Zeitraum des Rigorosums: 15. 11. 2002 – 10. 02. 2003

Tag der öffentlichen Verteidigung: 18. 02. 2003

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	3
2.1	Projektionsalgorithmen	3
2.1.1	Projektionsoperatoren in Hilberträumen	4
2.1.2	Problemspezifikation und Gütefunktion	5
2.1.3	Projektionsstärke	6
2.1.4	Serieller Projektionsalgorithmus für zwei Einschränkungsmengen	6
2.1.5	Serieller Projektionsalgorithmus für eine beliebige Anzahl von Einschränkungsmengen	7
2.1.6	Paralleler Multiprojektionsalgorithmus	7
2.2	Wellenpropagation durch optische Systeme	9
2.2.1	Komplexe Felder zur Repräsentation elektromagnetischer Feldkomponenten	9
2.2.2	Diskrete Darstellung von Feldern	9
2.2.3	Sequentielle Propagation durch optische Systeme	10
2.2.4	Transmissionsoperator	11
2.2.5	Ausbreitungsoperatoren für homogene Medien	11
2.2.6	Propagation durch optische Elemente	12
2.2.7	Optische Fouriertransformation	14
2.3	Wellenoptische Spezifikation von Transmissionsdesignproblemen	15
2.3.1	Mono- und multifunktionale Transmissionsdesignprobleme	15
2.3.2	Wichtige Gütefunktionen	17
2.4	Anwendung von Projektionsalgorithmen für monofunktionale Transmissionsdesignprobleme	19
2.4.1	Projektionsoperatoren für typische Transmissionseinschränkungen	19
2.4.2	Projektionsoperatoren für typische Signaleinschränkungen	20
2.4.3	Der Iterative Fouriertransformationsalgorithmus	20
2.4.4	Konvergenznachweis	21
2.4.5	Input-Output-Konzept	23
3	Diffraktive Strahlteilung	24
3.1	Vorbetrachtungen	24

3.1.1	Wellenoptische Modellierung	25
3.1.2	Pixelierungseffekte	26
3.1.3	Wahl der Abtastparameter	28
3.1.4	Iterationssteuerung beim Design überlappungsfreier Strahlteiler	29
3.2	Design für beliebig positionierte Signalordnungen	30
3.2.1	Verfeinerung der Signalebenenraasterung	31
3.2.2	Algorithmus für beliebig positionierte Signalordnungen	34
3.2.3	Designbeispiele	37
3.3	Toleranzeigenschaften von Strahlteilern	39
3.3.1	Höhenfehler bei binären Strahlteilern	39
3.3.2	Höhenfehler bei mehrstufigen Strahlteilern	43
3.3.3	Einfluss von Verschiebungen und Verkippungen	45
3.3.4	Toleranzeigenschaften bezüglich des Eingangsstrahls	46
3.4	Begrenzung der maximalen Rauschintensität	47
3.4.1	Modifikation des Signalprojektionsoperators	47
3.4.2	Erhöhung der Anzahl der zur Verfügung stehenden Rauschordnungen	49
3.4.3	Anzahl der Quantisierungsstufen	50
3.4.4	Anwendungsbeispiel Markierungsoptiken	51
4	Verbesserung des Konvergenzverhaltens durch verstärkte Projektionen	53
4.1	Vorbetrachtungen	53
4.1.1	Grundschema	53
4.1.2	Konvergenzrate	54
4.1.3	Frühere Arbeiten	55
4.2	Konvergenzbeschleunigung beim Design von Phasentransmissionen	56
4.2.1	Abhängigkeit der Konvergenzrate von der Projektionsstärke	57
4.2.2	Steuerung der Projektionsstärke	58
4.2.3	Konvergenzverhalten bei Beispielanwendungen	58
4.2.4	Sicherstellung der Konvergenz des Algorithmus	60
4.3	Überwindung von durch den Quantisierungsoperator verursachten Stagnationen	60
4.3.1	Eigenschaften des Softquantisierungsverfahrens	61
4.3.2	Anwendung verstärkter Projektionen	62
4.3.3	Konvergenzverhalten bei Beispielanwendungen	64
5	Multifunktionale Transmissionsdesignprobleme	66
5.1	Paralleler Multiprojektionsalgorithmus	66
5.1.1	Designprobleme mit in allen optischen Funktionen identischen Eingangsfeldern	67
5.1.2	Verallgemeinerung auf beliebige Eingangsfelder	67
5.2	Paralleler Transmissionsdesignalgorithmus	69
5.2.1	Motivation	69

5.2.2	Grundprinzip	70
5.2.3	Phaseneinschränkung	71
5.2.4	Quantisierungseinschränkung	72
5.2.5	Amplitudeneinschränkung	72
5.3	Vergleich der Konvergenzeigenschaften für eine Beispielanwendung	73
5.4	Frühere Arbeiten	77
6	Verbesserung von Toleranzeigenschaften durch multifunktionale Optimierung	79
6.1	Toleranzeigenschaften bezüglich des Propagationsabstands	80
6.1.1	Modellierung	80
6.1.2	Monofunktionales Design	82
6.1.3	Multifunktionales Design	83
6.2	Toleranzeigenschaften bezüglich globaler Höhenfehler und Wellenlängenfehler	85
6.2.1	Designmethoden	86
6.2.2	Vergleich zwischen simulierten und experimentellen Ergebnissen	89
7	Zusammenfassung	94
A	Beweise und Ableitungen	96
A.1	Zu Kapitel 2	96
A.1.1	Beweis zu Satz 2.16	96
A.1.2	Beweis zu Satz 2.67	97
A.1.3	Beweis zu Satz 2.68	97
A.2	Zu Kapitel 3	98
A.2.1	Ableitung von Gleichung (3.47)	98
A.2.2	Beweis zu Satz 3.26	100
A.3	Zu Kapitel 4	100
A.3.1	Ableitung der Gleichungen (4.15)	100
A.4	Zu Kapitel 5	102
A.4.1	Beweis zu Satz 5.8	102
A.4.2	Beweis zu Satz 5.25	104
A.4.3	Beweis zu Satz 5.28	106
	Literaturverzeichnis	109
	Abbildungsverzeichnis	115
	Tabellenverzeichnis	118
	Index	119

Kapitel 1

Einleitung

Eine aktuelle Tendenz im Bereich der optischen Modellierung sowie des Optikdesigns ist das Zusammenwachsen vormals separat betrachteter Bereiche wie der Holografie, der diffraktiven Optik und der Gittertheorie mit traditionellen Ansätzen aus dem Bereich der geometrischen Optik zu einem vereinheitlichten Konzept, dem *Wave-Optical Engineering* [1]. Die Einbeziehung wellenoptischer Methoden in das Optikdesign wird in einer zunehmenden Zahl von Anwendungen notwendig. Gründe hierfür sind beispielsweise Anforderungen zur Miniaturisierung optischer Systeme oder das Vorliegen von Optimierungszielen, welche – wie beispielsweise die Einkoppeleffizienz in eine Faser – auf wellenoptisch formulierten Gütefunktionen basieren. Begünstigt wird diese Entwicklung durch Fortschritte in der Rechentechnik, welche die typischerweise mit einer höheren Komplexität verbundenen wellenoptischen Berechnungen in einer wachsenden Zahl von Anwendungen praktikabel machen, sowie durch neue Herstellungstechnologien für optische Elemente mit allgemein geformten Oberflächen.

Innerhalb dieser Arbeit werden Aspekte des wellenoptischen Transmissionsdesigns untersucht. Die Optimierung von Transmissionsfunktionen entspricht einer Betrachtung des optischen Systems im Funktionsbild [2]. In einem späterem Schritt – dem Strukturdesign – kann die ermittelte Transmissionsfunktion in die geometrische Beschreibung eines optischen Elements überführt werden [3–5]. Diese Zerlegung in voneinander unabhängige Designschritte führt in vielen Fällen zu einer deutlichen Vereinfachung der Vorgehensweise.

Zur Optimierung von Transmissionsfunktionen betrachten wir die Anwendung sogenannter Projektionsalgorithmen. Dies sind iterativ arbeitende Algorithmen, bei denen Projektionen auf sogenannte Einschränkungsmengen eingesetzt werden, um für die freien Parameter des Optimierungsproblems eine Lösung zu finden, welche die den Einschränkungsmengen entsprechenden Zielanforderungen möglichst gut erfüllt. Dieses Grundschema von Projektionsalgorithmen wurde in einer Arbeit von Gubin et. al. [6] aus dem Jahr 1967 wahrscheinlich erstmalig eingesetzt. Wenige Jahre später wurden die ersten Anwendungen derartiger Algorithmen im Bereich der Optik dokumentiert, wobei sie zunächst zur Rekonstruktion einer Phasenverteilung aus Intensitätsmessungen [7] eingesetzt wurden und wenig später auch im Bereich des Transmissionsdesigns [8].

Im Transmissionsdesign ermöglicht der Einsatz von Projektionsalgorithmen eine effektive Handhabung der typischerweise sehr großen Anzahl freier Parameter (bis zu $\sim 10^8$). Andere Optimie-

rungsverfahren wie *Direct Binary Search* [9–11], *Simulated Annealing* [12–16] und *Optimal Rotation Angle Method* [17, 18] ermöglichen zwar die Optimierung hinsichtlich frei wählbarer Gütefunktionen, besitzen aber den Nachteil, dass in jeder Iteration typischerweise nur *ein* freier Parameter verändert wird. Somit versagen derartige Verfahren, falls bei einer hohen Anzahl freier Parameter der Effekt der Änderung eines Parameters nicht deutlich größer ist als die zu Verfügung stehende numerische Genauigkeit. Genetische Algorithmen [19, 20] sowie der *Simultaneous Perturbation Approach* [21] ermöglichen ebenfalls die Optimierung hinsichtlich beliebiger Gütefunktionen, wobei jedoch innerhalb jeder Iteration eine zufällige Störung auf eine größere Anzahl freier Parameter aufgebracht wird. Da solche Verfahren jedoch typischerweise im Gegensatz zu Projektionsalgorithmen eine sehr hohe Anzahl von Iterationen benötigen und teilweise aufwändig in ihrer Implementierung sind, konnten sie im Bereich des Transmissionsdesigns bisher keine weite Verbreitung finden.

Nach einer kompakten Darstellung der Grundlagen dieser Arbeit in Kapitel 2 wird in Kapitel 3 das Design diffraktiver Strahlteiler untersucht. In Kapitel 4 diskutieren wir die Anwendung verstärkter Projektionen zur Verbesserung des Konvergenzverhaltens von Projektionsalgorithmen. Die in Kapitel 5 betrachteten multifunktionalen Projektionsalgorithmen ermöglichen die Einbeziehung einer beliebigen Anzahl von Systemanforderungen in die Optimierung. Wie in Kapitel 6 gezeigt wird, können derartige Verfahren unter anderem zur Verbesserung der Toleranzeigenschaften optischer Systeme eingesetzt werden. Eine Zusammenfassung der im Rahmen dieser Arbeit gewonnenen wissenschaftlichen Ergebnisse wird in Kapitel 7 gegeben. Im Anhang A werden einige Beweise und Ableitungen dokumentiert, auf die im Verlauf der Arbeit verwiesen wird.

Kapitel 2

Grundlagen

Innerhalb dieses Kapitels werden die für diese Arbeit relevanten mathematischen und physikalischen Grundlagen zusammengefasst. Dabei wird zunächst in Abschnitt 2.1 mit den Projektionsalgorithmen das mathematische Grundkonzept der in dieser Arbeit entwickelten und verwendeten Optimierungsalgorithmen vorgestellt. In dem darauf folgenden Abschnitt 2.2 werden Methoden der wellenoptischen Analyse besprochen, bevor in Abschnitt 2.3 darauf aufbauend die Spezifikation wellenoptischer Designprobleme behandelt wird. Die Konzepte der genannten drei Abschnitte werden schließlich in Abschnitt 2.4 zusammengeführt, in dem die Anwendung von Projektionsalgorithmen bei der Lösung monofunktionaler Transmissionsdesignprobleme dargestellt wird.

2.1 Projektionsalgorithmen

Als *Projektionsalgorithmus* bezeichnen wir einen iterativ arbeitenden Algorithmus, welcher in einer betrachteten Iteration n durch Anwendung von Projektionsoperatoren auf ein Element $u^{(n-1)}$ eines Hilbertraums \mathcal{H} ein Element $u^{(n)} \in \mathcal{H}$ mit dem Ziel ermittelt, N vorgegebene Anforderungen bzw. *Einschränkungen* gleichzeitig zu erfüllen. Das Startelement $u^{(0)} \in \mathcal{H}$ wird hierbei im Allgemeinen von außen vorgegeben.

Das Grundschema von Projektionsalgorithmen wurde wahrscheinlich erstmalig von Gubin et. al. [6] untersucht. Seitdem fanden Projektionsalgorithmen eine breite Anwendung in verschiedenen Bereichen, wie beispielsweise der Rekonstruktion einer Phasenverteilung aus Intensitätsmessungen (*Phase retrieval*) [7, 22, 23], der Rekonstruktion einer Intensitätsverteilung aus Phasenmessungen im Fernfeld [24], der zeitlichen Pulsformung [25], der Interpolation aus beliebig verteilten Abtastwerten [26] sowie im Bereich der Chiffrierung [27]. Vielfältige Anwendungen bestehen weiterhin im Rahmen des Transmissionsdesigns [8, 28–38], welches den Hauptgegenstand dieser Arbeit darstellt.

In Abschnitt 2.1.1 wird zunächst das Konzept von Projektionsoperatoren in Hilberträumen mathematisch definiert. Im Anschluss daran wird in Abschnitt 2.1.2 die Spezifikation von mit Projektionsalgorithmen behandelbaren Optimierungsproblemen betrachtet, und es werden zwei hinsichtlich der Konvergenzeigenschaften relevante Gütefunktionen eingeführt. Das Konzept der Projektionsstärke betrachten wir in Abschnitt 2.1.3. In den darauf folgenden Abschnitten 2.1.4 – 2.1.6

werden schließlich drei verschiedene Formen von Projektionsalgorithmen vorgestellt und diskutiert.

2.1.1 Projektionsoperatoren in Hilberträumen

Ein *Hilbertraum* \mathcal{H} zeichnet sich durch das Vorliegen einer Norm $\|\cdot\|$ sowie eines Skalarproduktes $\langle \cdot | \cdot \rangle$ aus [39], die für alle $u \in \mathcal{H}$ miteinander verknüpft sind über

$$\|u\| = \sqrt{\langle u | u \rangle}. \quad (2.1)$$

Dabei gilt für beliebige Elemente $u, v \in \mathcal{H}$ insbesondere die sogenannte *Dreiecksungleichung*

$$\|u + v\| \leq \|u\| + \|v\|. \quad (2.2)$$

Ein auf einem Hilbertraum \mathcal{H} definierter Projektionsoperator $\Pi_{\mathcal{M}} : \mathcal{H} \mapsto \mathcal{M}$ ordnet jedem Element $u \in \mathcal{H}$ ein Element $v \in \mathcal{M} \subseteq \mathcal{H}$ zu, welches u bezüglich einer *Metrik* $d : \mathcal{H} \times \mathcal{H} \mapsto \mathbb{R}$ am nächsten liegt, d. h.

$$v = \Pi_{\mathcal{M}}u \Leftrightarrow v \in \mathcal{M} \wedge d(u, v) = \min_{v' \in \mathcal{M}} d(u, v'), \quad \text{bzw.} \quad (2.3a)$$

$$v = \underset{v' \in \mathcal{M}}{\operatorname{argmin}} d(u, v'). \quad (2.3b)$$

Innerhalb dieser Arbeit verwenden wir die mit Hilfe der Norm des Hilbertraums definierte euklidische Metrik

$$d(u, v) = \|u - v\|. \quad (2.4)$$

Mit dem für den Hilbertraum der quadratisch Lebesgue-integrierbaren komplexwertigen Funktionen $\mathbb{L}_2^{\mathbb{C}}(\mathbb{R})$ definierten Skalarprodukt

$$\langle u | v \rangle = \int_{-\infty}^{\infty} dx u(x)v^*(x) \quad (2.5)$$

ergibt sich d somit als

$$d(u, v) = \left[\int_{-\infty}^{\infty} dx |u(x) - v(x)|^2 \right]^{1/2}. \quad (2.6)$$

Wir nennen einen Projektionsoperator $\Pi_{\mathcal{M}} : \mathcal{H} \mapsto \mathcal{M}$ *lokal*, wenn er für jede Position x separat, d. h. unabhängig von anderen Positionen, ausgeführt wird [40]. Lokale Projektionsoperatoren besitzen die Eigenschaft der *punktweisen Abstandsminimierung*, welche sich für eine Position x ausdrücken lässt als

$$|u(x) - (\Pi_{\mathcal{M}}u)(x)|^2 = \min_{v \in \mathcal{M}} |u(x) - v(x)|^2. \quad (2.7)$$

Für die in dieser Arbeit betrachteten Metriken folgt aus der punktweisen Abstandsminimierung stets die Abstandsminimierungseigenschaft (2.3).

Eine direkte Folge der Abstandsminimierungseigenschaft sowie der Eigenschaften eines Hilbertraums ist die *Idempotenz* von Projektionsoperatoren,

$$\Pi \Pi u = \Pi u. \quad (2.8)$$

Eine Einschränkung wird als *konvex* bezeichnet, falls die zugehörige Einschränkungsmenge konvex ist. Bei *impliziten Einschränkungen* ist die Einschränkungsmenge $\mathcal{M} \subseteq \mathcal{H}$ indirekt über eine Menge $\widehat{\mathcal{M}}$ definiert, welche in einem Raum $\widehat{\mathcal{H}}$ gegeben ist, in den die Elemente $u \in \mathcal{H}$ über einen linearen Operator $\mathcal{S} : \mathcal{H} \rightarrow \widehat{\mathcal{H}}$ abgebildet werden, d. h.

$$u \in \mathcal{M} \Leftrightarrow \mathcal{S} u \in \widehat{\mathcal{M}}. \quad (2.9)$$

2.1.2 Problemspezifikation und Gütefunktion

Die durch Projektionsalgorithmen behandelbaren Optimierungsprobleme lassen sich durch die Vorgabe von Mengen $\mathcal{M}_j \subseteq \mathcal{H}$, $j = 1 \dots N$ spezifizieren, wobei jede Menge \mathcal{M}_j eine Anforderung bzw. Einschränkung an die Lösung des Optimierungsproblems beschreibt. Die Mengen \mathcal{M}_j bezeichnen wir im Folgenden auch als *Einschränkungsmengen*. Demnach besteht das Ziel eines Projektionsalgorithmus idealerweise darin, ein Element der Durchschnittsmenge

$$\mathcal{M}_0 = \bigcap_{j=1}^N \mathcal{M}_j \quad (2.10)$$

zu finden. Wenn dies theoretisch möglich ist, d. h. wenn die Durchschnittsmenge \mathcal{M}_0 nicht leer ist, sprechen wir von einer *konsistenten* Problemstellung [41], anderenfalls von einem *inkonsistenten* Problem.

Da die in dieser Arbeit betrachteten Probleme häufig inkonsistent sind, ist das oben formulierte ideale Ziel, ein Element von \mathcal{M}_0 zu finden, in vielen Fällen nicht erreichbar. Da auch die Entscheidung, ob ein Problem konsistent oder inkonsistent ist, meist nicht ohne weiteres möglich ist, konzentrieren wir uns auf den allgemeinen Fall nicht notwendigerweise konsistenter Problemstellungen. Zur damit notwendigen Formulierung einer abgeschwächten Zielstellung von Projektionsalgorithmen ist daher die Vorgabe einer Güte- oder Fehlerfunktion notwendig, die während der Optimierung maximiert bzw. minimiert wird.

Ein im Zusammenhang mit Projektionsalgorithmen übliches Fehlermaß ist der *summierte Abstandsfehler* (*Summed Distance Error*, SDE) [42]

$$J(u) = \sum_{j=1}^N d(u, \Pi_j u) = \sum_{j=1}^N \|u - \Pi_j u\|. \quad (2.11)$$

Hierbei bezeichnet Π_j eine Projektion auf die Einschränkungsmenge \mathcal{M}_j .

Ein weiteres im Rahmen dieser Arbeit verwendetes Fehlermaß ist der *summierte quadratische Abstandsfehler* (*Summed Squared Distance Error*, SSDE) [41]

$$\hat{J}(u)^2 = \sum_{j=1}^N d(u, \Pi_j u)^2 = \sum_{j=1}^N \|u - \Pi_j u\|^2. \quad (2.12)$$

Aufgrund der Eigenschaften der Norm sind $J(u)$ und $\hat{J}(u)^2$ nichtnegativ. Für die Lösung $u_0 \in \mathcal{M}_0$ eines konsistenten Problems verschwinden beide Fehlermaße, d. h. $J(u_0) = \hat{J}(u_0)^2 = 0$.

2.1.3 Projektionsstärke

Basierend auf einem Projektionsoperator Π definieren wir den Operator Π^λ als [6]

$$\Pi^\lambda u = u + \lambda(\Pi u - u). \quad (2.13)$$

Der Parameter λ wird hierbei als *Projektionsstärke* bezeichnet. Im Fall von $0 \leq \lambda < 1$ kann der Operator Π^λ als *relaxierte* Version von Π aufgefasst werden; im Fall von $\lambda > 1$ sprechen wir von einer *verstärkten* Projektion. Als Spezialfälle sind in diesem Formalismus die identische Abbildung $\Pi^0 u = u$ für $\lambda = 0$ sowie die ursprüngliche Projektion $\Pi^1 u = \Pi u$ für $\lambda = 1$ enthalten.

2.1.4 Serieller Projektionsalgorithmus für zwei Einschränkungsmengen

Ausgehend von Problemspezifikationen nach Abschnitt 2.1.2 beschränken wir uns in diesem Abschnitt zunächst auf den Fall $N = 2$, d. h. auf zwei Einschränkungsmengen und untersuchen den Projektionsalgorithmus

$$u^{(n)} = \Pi_2 \Pi_1 u^{(n-1)}. \quad (2.14)$$

Diese Algorithmus konvergiert entsprechend des folgenden in Anhang A.1.1 bewiesenen Satzes 2.16 bezüglich des SDEs, d. h. für Iterationen $n > 1$ gilt

$$J(u^{(n)}) \leq J(u^{(n-1)}). \quad (2.15)$$

Die Einschränkung auf Iterationen $n > 1$ entfällt, sofern ein Startelement $u^{(0)} \in \mathcal{M}_1$ gewählt wird. Eine Verallgemeinerung von Satz 2.16 wurde von Levi und Stark [42] gezeigt.

Satz 2.16 *Algorithmus (2.14) konvergiert bezüglich des SDEs aus Gleichung (2.11), d. h. für $n > 1$ gilt die Konvergenzbedingung (2.15).*

Bei Ersetzung der Projektionsoperatoren Π_1 und Π_2 in Algorithmus (2.14) durch ihre zugeordneten relaxierten oder verstärkten Versionen $\Pi_1^{\lambda_1}$ bzw. $\Pi_2^{\lambda_2}$ (vgl. Abschnitt 2.1.3) ist die SDE-Konvergenz nicht für beliebige Projektionsstärken λ_1 und λ_2 gesichert. Levi und Stark [42] untersuchten, welche Bedingungen an λ_1 und λ_2 für eine Absicherung der SDE-Konvergenz zu stellen sind. Wir gehen auf dieses Thema hier nicht näher ein, da in Fällen, in denen in dieser Arbeit relaxierte bzw. verstärkte Projektionen verwendet werden, die Konvergenz durch andere Maßnahmen sichergestellt wird.

Mathematisch gesehen scheint die Beschränkung des in diesem Abschnitt betrachteten Projektionsalgorithmus auf zwei Einschränkungsmengen zunächst nicht gravierend zu sein, da stets mehrere Einschränkungsmengen durch Bildung der Durchschnittsmenge zu einer einzigen Einschränkungsmenge zusammengefasst werden können. In der Praxis ist die effektive Implementierbarkeit der Projektionsoperatoren eine entscheidende Frage, auf Grund derer wir häufig zugunsten effektiverer Projektionsoperatoren auf mehr als zwei Einschränkungen zurückgreifen müssen. Dies ist die Motivation für die in den beiden folgenden Abschnitten betrachteten Projektionsalgorithmen, welche für eine beliebige Anzahl von Einschränkungsmengen angewandt werden können.

2.1.5 Serieller Projektionsalgorithmus für eine beliebige Anzahl von Einschränkungenmengen

Unter Aufhebung der in Abschnitt 2.1.4 getroffenen Einschränkung auf $N = 2$ eröffnet sich eine Reihe von Möglichkeiten zur Formulierung von Projektionsalgorithmen für derartige Problemstellungen. Zwei aus der Literatur bekannte Ansätze werden in diesem sowie in dem nachfolgendem Abschnitt 2.1.6 diskutiert. Der sogenannte *serielle Projektionsalgorithmus* lässt sich darstellen in der Form

$$u^{(n)} = \Pi_N \Pi_{N-1} \dots \Pi_1 u^{(n-1)}, \quad (2.17)$$

wobei wiederum einer oder mehrere Projektionsoperatoren Π_j durch den zugehörigen relaxierten oder verstärkten Operator $\Pi_j^{\lambda_j}$ ersetzt werden können.

Für konvexe Mengen \mathcal{M}_j wurde die SDE-Konvergenz des seriellen Projektionsalgorithmus für $0 < \lambda_j < 2$ ($j = 1 \dots N$) von Youla [43] nachgewiesen. Für den allgemeinen Fall nichtkonvexer Einschränkungen, ist die Konvergenz hingegen nur unter bestimmten Zusatzbedingungen gewährleistet [44]. Ein anschauliches Beispiel für einen Fall, in dem der serielle Projektionsalgorithmus nicht konvergiert, ist in Referenz [42] gegeben.

2.1.6 Paralleler Multiprojektionsalgorithmus

Während beim seriellen Projektionsalgorithmus sämtliche Projektionen innerhalb eines Iterationszyklus nacheinander ausgeführt werden, besteht beim *Parallelen Multiprojektionsalgorithmus* (PMPA) die Idee darin, alle Projektionen gleichzeitig auszuführen und anschließend eine Mittelwertsbildung durchzuführen. Der PMPA wurde eingeführt von Kotzer et. al. [45]. Im Vergleich zu der in dieser Referenz gegebenen Darstellung beschränken wir uns in dieser Arbeit auf die Verwendung euklidischer Normen, womit sich als vereinfachte Darstellung des PMPAs ergibt:

$$u^{(n)} = \frac{1}{N} \sum_{j=1}^N \Pi_j u^{(n-1)}. \quad (2.18)$$

Auch hierbei ist eine Ersetzung der Projektionsoperatoren durch zugehörige relaxierte oder verstärkte Operatoren denkbar. Die Priorität der einzelnen Einschränkungen kann beim PMPA durch Einbeziehung von Wichtungsfaktoren in die Mittelwertsbildung berücksichtigt werden, worauf wir im Interesse einer einfacheren Darstellung jedoch verzichten.

Die SSDE-Konvergenz des PMPAs wurde in einer verallgemeinerten Form des nachfolgenden Satzes 2.22 von Censor et. al. [41] gezeigt. Die Idee zum Beweis dieses Satzes besteht darin, den PMPA auf einen seriellen Projektionsalgorithmus mit $N = 2$ zurückzuführen, welcher auf dem Vektorraum \mathcal{H}^N mit Elementen $\mathbf{u} \in \mathcal{H}^N$ der Form

$$\mathbf{u} = (u_1, u_2, \dots, u_N) \quad (2.19)$$

arbeitet. In \mathcal{H}^N definieren wir die Metrik $\mathbf{d} : \mathcal{H}^N \times \mathcal{H}^N \mapsto \mathbb{R}$

$$\mathbf{d}(\mathbf{u}, \mathbf{v}) = \left[\sum_{j=1}^N \|u_j - v_j\|^2 \right]^{1/2}, \quad (2.20)$$

sowie die Funktion $\tau : \mathcal{H} \mapsto \mathcal{H}^N$

$$\tau(u) = \underbrace{(u, u, \dots, u)}_{N \text{ mal}}. \quad (2.21)$$

Satz 2.22 *Algorithmus (2.18) konvergiert bezüglich des SSDEs aus Gleichung (2.12), d. h. für $n > 1$ gilt*

$$\hat{J}(u^{(n)})^2 \leq \hat{J}(u^{(n-1)})^2. \quad (2.23)$$

Beweis: Darstellung (2.18) ist äquivalent zu

$$\tau(u^{(n)}) = \tilde{\Pi}_2 \tilde{\Pi}_1 \tau(u^{(n-1)}), \quad \text{wobei} \quad (2.24a)$$

$$\tilde{\Pi}_1 \mathbf{u} = (\Pi_1 u_1, \Pi_2 u_2, \dots, \Pi_N u_N) \quad \text{und} \quad (2.24b)$$

$$\tilde{\Pi}_2 \mathbf{u} = \tau \left(\frac{1}{N} \sum_{j=1}^N u_j \right). \quad (2.24c)$$

Algorithmus (2.24a) stellt einen Spezialfall des seriellen Projektionsalgorithmus für zwei Mengen dar. Um Satz 2.22 für diesen Beweis anwenden zu können, prüfen wir zunächst, ob die Operatoren $\tilde{\Pi}_1$ und $\tilde{\Pi}_2$ Projektionen ausführen, welche abstandsminimierend bezüglich der Metrik \mathbf{d} sind.

Lemma 2.25 *Der Operator $\tilde{\Pi}_1$ laut Gleichung (2.24b) definiert eine Projektion auf*

$$\tilde{\mathcal{M}}_1 = \{\mathbf{u} = (u_1, u_2, \dots, u_N) : u_j \in \mathcal{M}_j, j = 1 \dots N\}. \quad (2.26)$$

Der Beweis dieses Lemmas beruht auf der gegebenen Abstandsminimierung bezüglich d der Projektionen Π_j , woraus die Abstandsminimierung bezüglich \mathbf{d} für $\tilde{\Pi}_1$ folgt.

Lemma 2.27 *Der Operator $\tilde{\Pi}_2$ laut Gleichung (2.24c) definiert eine Projektion auf*

$$\tilde{\mathcal{M}}_2 = \{\tau(u) : u \in \mathcal{H}\}. \quad (2.28)$$

Der Beweis von Lemma 2.27 beruht auf der Minimierung der Norm \mathbf{d} aus Gleichung (2.20) durch die Mittelwertbildung in Gleichung (2.24c). Mit Satz 5.8 wird in Anhang A.4.1 eine verallgemeinerte Version von Lemma 2.27 bewiesen.

Aufgrund der Lemmata 2.27 und 2.25, lässt sich Satz 2.16 auf den Algorithmus (2.24a) anwenden. Die daher vorliegende SDE-Konvergenz von Algorithmus (2.24a) stellen wir dar als

$$\mathbf{J}(\tau(u^{(n)})) \leq \mathbf{J}(\tau(u^{(n-1)})) \quad \text{mit} \quad (2.29a)$$

$$\mathbf{J}(\mathbf{u}) = \mathbf{d}(\mathbf{u}, \tilde{\Pi}_1 \mathbf{u}) + \mathbf{d}(\mathbf{u}, \tilde{\Pi}_2 \mathbf{u}). \quad (2.29b)$$

Wegen $\tau(u^{(n)}) \in \tilde{\mathcal{M}}_1$ und Gleichung (2.20) erhalten wir

$$\begin{aligned} \mathbf{J}(\tau(u^{(n)})) &= \mathbf{d}(\mathbf{u}, \tilde{\Pi}_1 \mathbf{u}) \\ &= \left[\sum_{j=1}^N \|u_j - \Pi_j u_j\|^2 \right]^{1/2} \\ &= \hat{J}(u). \end{aligned} \quad (2.30)$$

Somit folgt aus Gleichung (2.29) die SSDE-Konvergenz von Algorithmus (2.18). ■

2.2 Wellenpropagation durch optische Systeme

In diesem Abschnitt werden Grundlagen zur wellenoptischen Propagation durch optische Systeme zusammengefasst. Dabei wird in den Abschnitten 2.2.1 und 2.2.2 die in dieser Arbeit verwendete Feldrepräsentation beschrieben, wobei für eine detailliertere Darstellung dieser Thematik auf Referenz [1] verwiesen wird. In den darauf folgenden Abschnitten 2.2.3 – 2.2.7 werden Methoden zur Berechnung der Propagation durch optische Systeme betrachtet.

2.2.1 Komplexe Felder zur Repräsentation elektromagnetischer Feldkomponenten

Elektromagnetische Felder können durch reellwertige Funktionen $u(x, y, z, t)$ beschrieben werden, wobei der Bezeichner u für eine der skalaren Komponenten des elektrischen oder magnetischen Feldes steht. Bei Beschränkung auf monochromatische Felder mit der Frequenz ν kann der Zeitannteil durch

$$u(x, y, z, t) = \operatorname{Re} \left(u(x, y, z) \exp(-i2\pi\nu t) \right) \quad (2.31)$$

abgespalten werden. Die komplexwertige Funktion $u(x, y, z)$ wird auch als *komplexe Amplitude* bezeichnet.

Die z -Achse stellt im Folgenden eine ausgezeichnete Richtung im optischen System dar: die sogenannte *optische Achse*. Wir betrachten im Folgenden die komplexe Amplitude $u(x, y, z)$ in Ebenen senkrecht zur optischen Achse. Da z somit jeweils konstant ist und aus dem Kontext ersichtlich wird, genügt eine Darstellung der Feldverteilung in der Form $u(x, y)$. Die lateralen Schnitte $u(x, y)$ der komplexen Amplitude werden im Folgenden als *Felder* bezeichnet.

Zur Vereinfachung der mathematischen Schreibweise werden wir uns im Folgenden auf ein-dimensionale Felder $u(x)$ beschränken, wobei die Verallgemeinerung auf den zweidimensionalen Fall stets ohne Schwierigkeiten möglich ist.

2.2.2 Diskrete Darstellung von Feldern

Bei einer numerischen Behandlung von Feldern muss stets eine Abtastung an einer endlichen Anzahl n_x von Positionen erfolgen. Bei einer äquidistanten Abtastung mit dem Abtastabstand δx ergeben sich die Abtastpositionen $x_{\min}, x_{\min} + \delta x, \dots, x_{\min} + (n_x - 1) \delta x$ mit

$$x_{\min} = -\left\lfloor \frac{n_x}{2} \right\rfloor \delta x, \quad (2.32)$$

wobei $\lfloor \cdot \rfloor$ die Abrundung auf eine ganze Zahl bezeichnet.

Bei Verwendung der diskreten Darstellung von Feldern ist auf eine korrekte Abtastung entsprechend des Whittaker-Shannon-Theorems [46] zu achten, d. h. der abgetastete Feldverlauf muss anhand der Abtastwerte fehlerfrei rekonstruierbar sein. In der Praxis kann diese Bedingung im Allgemeinen nur näherungsweise eingehalten werden, da die Bandbegrenztheit einer Funktion laut dem Whittaker-Shannon-Theorem eine Voraussetzung für ihre Abtastbarkeit darstellt, wobei jedoch viele in der Praxis auftretende Felder, wie beispielsweise Gaußfelder, durch Aperturen

begrenzte Felder oder phasenmodulierte Felder mit konstanter Amplitude [47] nicht bandbegrenzt sind. Der hierbei entstehende Fehler kann jedoch im Allgemeinen durch eine ausreichend niedrige Wahl des Abtastabstands δx so stark reduziert werden, dass er physikalisch nicht mehr relevant ist. Im Rahmen dieser Arbeit wird stets von einer ausreichend genauen Abtastung ausgegangen, bzw. diese durch eine geeignete Wahl der Abtastparameter bei den betrachteten Beispielen gewährleistet.

Obwohl bei der praktischen Durchführung der in dieser Arbeit betrachteten Algorithmen stets mit der diskreten Darstellung von Feldern gearbeitet werden muss, verwenden wir im Folgenden zur Vereinfachung der mathematischen Darstellung die kontinuierliche Darstellung. Der Übergang zwischen beiden Darstellungen kann durch Ersetzung der Integrale wie beispielsweise in den Gleichungen (2.5) und (2.6) durch korrespondierende Summen erfolgen.

2.2.3 Sequentielle Propagation durch optische Systeme

Bei der *sequentuellen Propagation*, auf die wir uns im Rahmen dieser Arbeit beschränken möchten, wird vernachlässigt, dass optische Grenzflächen im Allgemeinen das einfallende Licht gleichzeitig teilweise reflektieren und transmittieren. Der einem optischen Element im Rahmen einer sequentiellen Analyse zugeordneten Propagationsoperator beschreibt daher nur entweder den reflektiven oder den transmittiven Teil der optischen Wirkung des betrachteten Elements, wobei wir uns im Folgenden auf den letztgenannten Fall konzentrieren möchten.

Bei einer Zerlegung des optischen Systems in n Propagationsoperatoren P_j ($j = 1 \dots n$) ergibt sich das Ausgangsfeld $u_{\text{out}}(x')$ bei sequentieller Propagation des Eingangsfelds $u_{\text{in}}(x)$ gemäß

$$\begin{aligned} u_{\text{out}}(x') &= P_n \cdots P_2 P_1 u_{\text{in}}(x) \\ &= (P_1 \circ P_2 \circ \cdots \circ P_n) u_{\text{in}}(x). \end{aligned} \quad (2.33)$$

Hier und im Folgenden bezeichnen wir die laterale Koordinate in der Ausgangs- bzw. *Signalebene* des optischen Systems mit x' , da mit der Anwendung eines optischen Operators im Allgemeinen auch eine Koordinatentransformation verbunden sein kann. Wie in Abbildung 2.1 gezeigt wird, beschreibt hierbei jeder Operator P_j die Propagation durch den über die Referenzebenen z_{j-1} und z_j begrenzten Bereich des optischen Systems. Die Zerlegung des optischen Systems in Einzeloperatoren P_j wird anhand der Anwendbarkeit der zur Verfügung stehenden Propagationmethoden auf die verschiedenen Teile des optischen Systems durchgeführt, wobei günstigerweise die Positionen z_j so gewählt werden, dass Bereiche mit homogener Brechzahlverteilung eine maximale Ausdehnung erhalten.

In den folgenden Abschnitten werden Beispiele für mögliche Propagationsoperatoren P_j gegeben.

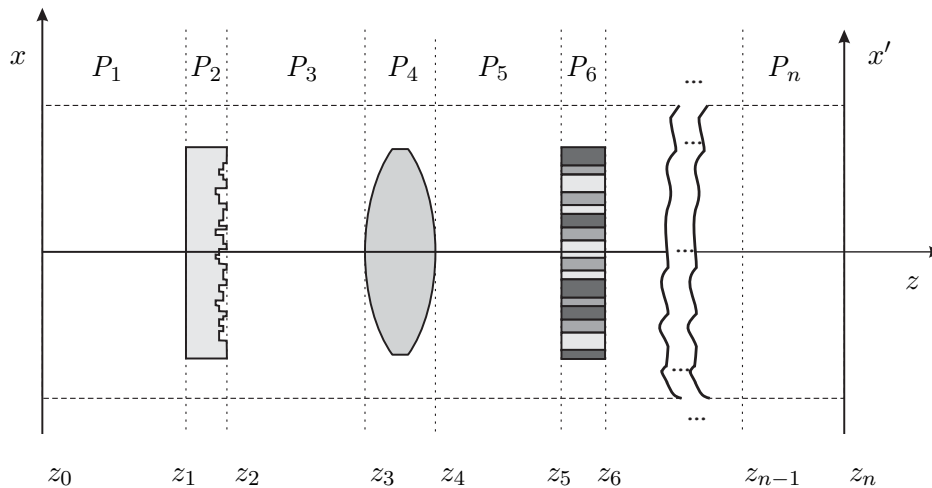


Abbildung 2.1: Zerlegung eines optischen Systems in durch Referenzebenen senkrecht zur z -Achse begrenzte Bereiche, denen jeweils Operatoren P_j ($j = 1 \dots n$) zugeordnet sind

2.2.4 Transmissionsoperator

Die Anwendung des Transmissionsoperators \mathcal{T}_t entspricht dabei einer Multiplikation mit einer Transmissionsfunktion $t(\xi) \in \mathcal{H}$, d. h.

$$\mathcal{T}_t u(x) = t(x)u(x). \quad (2.34)$$

Der Transmissionsoperator \mathcal{T}_t ist ein sogenannter *funktionsbezogener* Operator, welcher im Gegensatz zu *strukturbezogenen* Operatoren nicht direkt einem Bereich des optischen Systems entspricht, sondern lediglich eine optische Wirkung beschreibt. Transmissionsfunktionen resultieren als Ergebnis des in dieser Arbeit betrachteten Transmissionsdesigns und können über Strukturdesignmethoden in reale optische Elemente überführt werden [1]. Während das Strukturdesign im Fall der Anwendbarkeit der in Abschnitt 2.2.6 diskutierten Dünne-Elemente-Approximation eine Transmissionsfunktion mit $|t(x)| \leq 1$ direkt in die Beschreibung eines optischen Elements überführt werden kann, existieren außerhalb dieser Näherung weitere Strukturdesignverfahren [3,5,48], welche im Rahmen dieser Arbeit jedoch nicht weiter betrachtet werden.

2.2.5 Ausbreitungsoperatoren für homogene Medien

Wie beispielsweise in Referenz [49] gezeigt wird, kann der für homogene isotrope dielektrische Medien ohne weitere Näherungen gültige Propagationsoperator Propagationsoperator $P_{\Delta z}^{\text{ASPW}}$,

$$P_{\Delta z}^{\text{ASPW}} u = \mathcal{F}^{-1} \left(\exp[-i\sqrt{k^2 - 4\pi^2 x'^2} \Delta z] \mathcal{F} u \right), \quad (2.35)$$

für die Propagation um einen Abstand Δz über eine Zerlegung des Feldes in das *Spektrum ebener Wellen* (*Angular Spectrum of Plane Waves*, ASPW) abgeleitet werden. Hierbei wurden folgende Definitionen für die Fouriertransformation \mathcal{F} (FT) sowie die inverse Fouriertransformation \mathcal{F}^{-1}

angenommen:

$$[\mathcal{F} u(x)](x') = \int_{-\infty}^{\infty} dx' \exp[-i2\pi xx'] U(x) \quad \text{bzw.} \quad (2.36a)$$

$$[\mathcal{F}^{-1} u(x')](x) = \int_{-\infty}^{\infty} dx \exp[i2\pi xx'] U(x). \quad (2.36b)$$

Bei Anwendung der für paraxiale Felder gültigen Fresnel-Näherung [46] auf Gleichung (2.35) ergibt sich der Fresnel-Operator $P_{\Delta z}^{\text{Fresnel}}$,

$$[P_{\Delta z}^{\text{Fresnel}} u(x)](x') = \frac{A}{i} \exp(ik\Delta z) \exp\left(\frac{ik}{2\Delta z} x'^2\right) \left[\mathcal{F} \left(u(x) \exp\left[\frac{ik}{2\Delta z} x^2\right] \right) \right] \left(\frac{k}{2\pi\Delta z} x' \right), \quad (2.37)$$

wobei A einen reellen Amplitudenfaktor bezeichnet. Falls die Ausdehnung von $u(x)$ in x -Richtung viel kleiner ist als der Propagationsabstand Δz , kann der quadratische Phasenterm innerhalb der Fouriertransformation vernachlässigt werden und es resultiert die Fraunhofersche Fernfeldnäherung [46]

$$[P_{\Delta z}^{\text{Fraunh}} u(x)](x') = \frac{A}{i} \exp(ik\Delta z) \exp\left(\frac{ik}{2\Delta z} x'^2\right) [\mathcal{F} u(x)] \left(\frac{k}{2\pi\Delta z} x' \right). \quad (2.38)$$

Bei diskreten Felddarstellungen entsprechend Abschnitt 2.2.2 können die Fouriertransformationen innerhalb der oben definierten Propagationsoperatoren über die Technik der *schnellen Fouriertransformation* (*Fast Fourier Transform*, FFT) [50] mit einem Zeitaufwand proportional zu $n_x \log(n_x)$ durchgeführt werden. Bei Verwendung der diskreten Darstellung ist jedoch zu beachten, dass sämtliche Zwischenergebnisse sowie die in den Gleichungen (2.35) und (2.37) auftretenden sphärischen bzw. quadratischen Phasenterme korrekt abgetastet werden. Sich hieraus ergebende Forderungen bezüglich des Propagationsabstands und der Abtastparameter werden in Referenz [51] diskutiert.

Aufgrund des Parselvalschen Theorems [50] sowie der eindeutigen Invertierbarkeit der Fouriertransformation erfüllen die in diesem Abschnitt eingeführten Propagationsoperatoren P die Eigenschaften

$$\|Pu\| = \gamma \|u\| \quad (\gamma \in \mathbb{R}, \gamma > 0) \quad \text{sowie} \quad (2.39a)$$

$$PP^{-1}u = u. \quad (2.39b)$$

2.2.6 Propagation durch optische Elemente

Im vorangegangenen Abschnitt wurde mit Gleichung (2.35) ein Ausbreitungsoperator für homogene Regionen abgeleitet, welcher ohne physikalische Näherungen auskommt und daher auch als *rigoros* bezeichnet wird. Im Fall inhomogener Medien existieren analytisch bekannte rigorose Lösungen nur für wenige Spezialfälle. Insbesondere für *Gitter*, d. h. in lateraler Richtung periodische Elemente, existieren jedoch rigorose numerische Näherungsverfahren [52, 53]. Da solche Verfahren häufig mit einem hohen Rechenaufwand verbunden sind, welcher typischerweise in der dritten

Potenz mit der Größe der Gitterperiode Δx steigt, bleibt deren Anwendungsbereich jedoch begrenzt.

Innerhalb zweier Regimes hinsichtlich der auftretenden Strukturgrößen können Näherungen verwendet werden, welche mit einem deutlich geringerem Rechenaufwand verbunden sind. Die *Näherung der effektiven Medien* [54] ist gültig für Elemente mit lokalen Gitterperioden $\Delta x \ll \lambda$. An solchen *Subwellenlängengittern* kann keine Beugung erfolgen, da sämtliche höheren Ordnungen $|m| \geq 1$ entsprechend der Gittergleichung

$$\sin \theta_m = \frac{m\lambda}{\Delta x} \quad (2.40)$$

Winkeln $\theta_m > 90^\circ$ entsprechen und somit *evaneszent* sind. Derartige optische Element wirken somit ähnlich wie brechzahlmodulierte Elemente mit glatten Oberflächen, wobei die resultierende Brechzahlverteilung von dem Füllfaktor der lokalen Gitterstruktur abhängt.

Eine in der Praxis noch höhere Bedeutung besitzt die sogenannte *Dünne-Elemente-Approximation* [55], welche einen besonders einfachen Übergang zwischen der Struktur eines Elements und der damit verbundenen Transmissionsfunktion $t(x)$ erlaubt. Der Transmissionswert an der Position x ergibt sich dabei ausschließlich aus dem Brechzahl- und Absorptionsverlauf entlang eines Strahls, der im Abstand x parallel zur optischen Achse verläuft. Für ein homogenes dielektrisches Element mit einer Brechzahl n und einer Dickenverteilung $d(x)$ ergibt sich somit innerhalb von Vakuum eine Phasenmodulation entsprechend der Transmission

$$t_d(x) = \exp \left[\frac{i2\pi(n-1)}{\lambda} d(x) \right]. \quad (2.41)$$

Die Dünne-Elemente-Approximation kann angewendet werden, wenn die auftretende minimale Strukturgröße δ_{\min} deutlich größer als die Wellenlänge λ ist. Als Strukturgröße wird hierbei anschaulich die laterale Ausdehnung einer Struktur des Höhenprofils bezeichnet, welche einer Phasenmodulation von 2π entspricht. Dieser im Allgemeinen unscharf definierte Begriff kann auch auf Elemente mit modulierter Brechzahl- sowie Absorptionsverteilung übertragen werden.

Es existieren verschiedene Untersuchungen zur genauen Grenze der Anwendbarkeit dieser Approximation [55], wobei sich je nach konkreter Struktur verschiedene Verhältnisse δ_{\min}/λ als Grenzwert ergeben. Während insbesondere bei Gittern mit nur einer Struktur pro Periode Verhältnisse $\delta_{\min}/\lambda > 10$ notwendig sind, um eine gute Übereinstimmung der über die Dünne-Elemente-Approximation resultierenden Beugungseffizienzen mit den rigoros ermittelten Werten zu erhalten [56, 57], ergibt sich für Elemente mit großen Periodenlängen oder auch für nichtperiodische Elemente häufig bereits für $\delta_{\min}/\lambda > 2$ ein akzeptabler Übereinstimmungsgrad [17, 58–60].

Bei Anwendung der Dünne-Elemente-Approximation auf eine (dünne) Linse, deren Grenzflächen die Krümmungsradien R_1 und R_2 besitzen, ergibt sich bis auf einen konstanten Phasenfaktor eine Phasenmodulation entsprechend [46]

$$t_f(x) = \exp \left[\frac{-ik}{2f} x^2 \right] \quad (2.42)$$

mit der *Brennweite* f gemäß

$$f = \left[(n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \right]^{-1}. \quad (2.43)$$

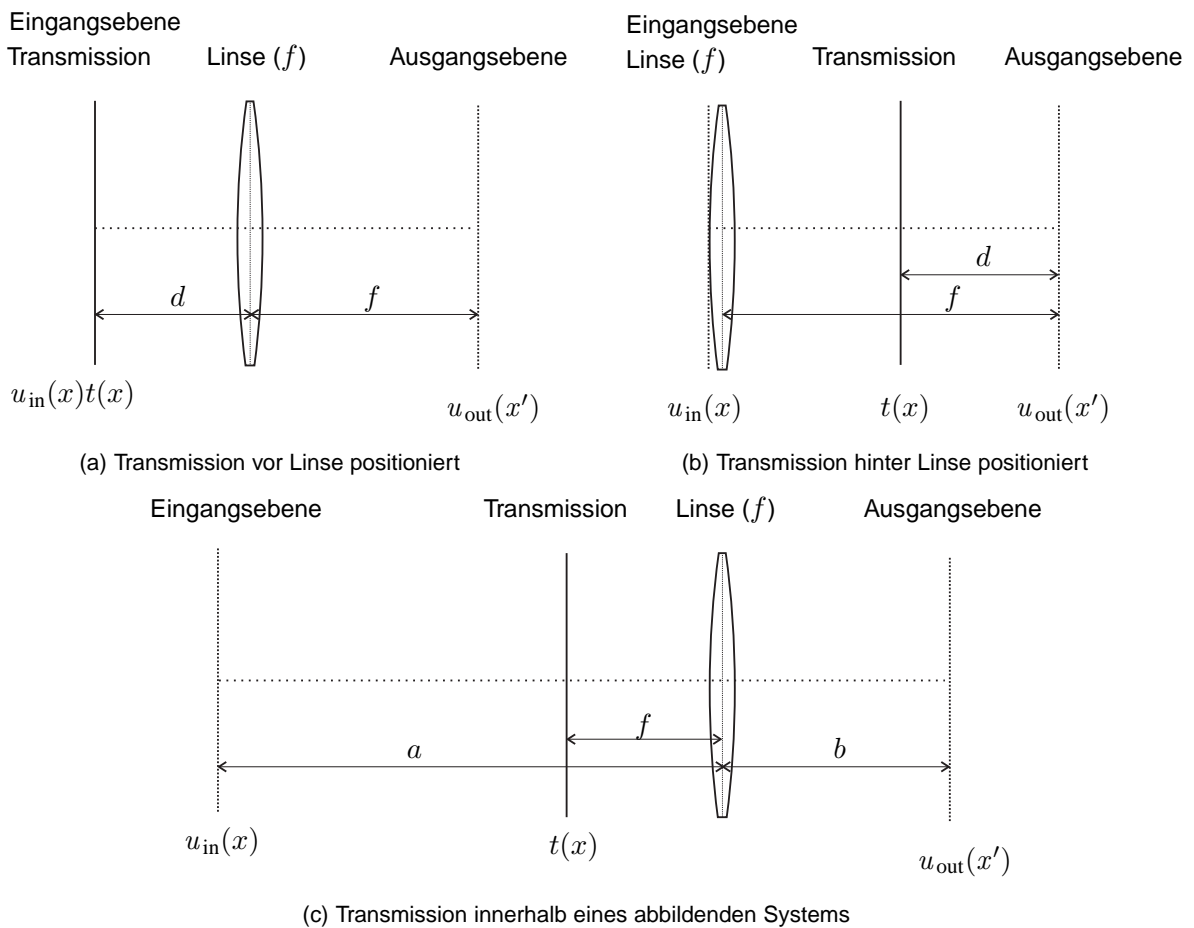


Abbildung 2.2: Beispiele für Systeme zur Realisierung der optischen Fouriertransformation einer Transmissionsfunktion

Zum Abschluss dieses Abschnitts soll noch die bereits am Ende von Abschnitt 2.2.4 erwähnte der Dünne-Elemente-Approximation entsprechende Strukturdesignmethode für den Spezialfall einer Phasentransmission $t(x) = \exp(i\phi(x))$ angegeben werden. Entsprechend Gleichung (2.41) resultiert bei Annahme einer Referenzwellenlänge λ_{ref} sowie einer Brechzahl n dem Höhenprofil

$$d_{\phi}(x) = \frac{\lambda_{\text{ref}}}{2\pi(n-1)}\phi(x). \quad (2.44)$$

2.2.7 Optische Fouriertransformation

Aus den paraxial gültigen Gleichungen (2.37) und (2.42) resultiert für die drei in Abbildung 2.2 gezeigten Systeme eine Fouriertransformationsbeziehung zwischen der Ausgangsebene und der Transmissionsfunktion, wobei jedoch Skalierungen und quadratische Phasenterme auftreten können. Für das System aus Abbildung 2.2(a), in welchem sich eine Transmission in einem Abstand d vor einer Linse befindetet, ergibt sich [46]

$$u_{\text{out}}(x') \sim \exp\left[\frac{ik}{2f}\left(1 - \frac{d}{f}\right)x'^2\right] \int_{-\infty}^{\infty} dx t(x) u_{\text{in}}(x) \exp\left[\frac{-ik}{f}xx'\right]. \quad (2.45)$$

Das sich im Fall $d = f$ ergebende Ausgangsfeld entspricht somit exakt der Fouriertransformation des Produkts $t(x) u_{\text{in}}(x)$. Bei Anordnung der Transmissionsfunktion hinter der Linse entsprechend

Abbildung 2.2(b) ergibt sich der Zusammenhang [46]

$$u_{\text{out}}(x') \sim \exp \left[\frac{ik}{2d} x'^2 \right] \int_{-\infty}^{\infty} dx t(x) u_{\text{in}} \left(\frac{fx}{d} \right) \exp \left[\frac{-ik}{d} xx' \right], \quad (2.46)$$

bei dem im Gegensatz zu Gleichung (2.45) die Skalierung der Fouriertransformation von der Wahl des Abstands d abhängt. Als letzten Fall betrachten wir in Abbildung 2.2(c) die Anordnung einer Transmissionsfunktion innerhalb eines abbildenden Systems, d. h. die Abstände a und b erfüllen die Bedingung

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}. \quad (2.47)$$

Für den Fall $t(x) \equiv 1$ ergibt sich dann eine Abbildungsbeziehung über den Skalierungsfaktor $M = b/a$, d. h.

$$u_{\text{out}}(x') \sim u_{\text{in}}(-x/M) \exp \left(i \frac{\pi x'^2}{\lambda(b-f)} \right), \quad (2.48)$$

wobei wie schon in den vorangegangenen Beispielen die Linsenapertur vernachlässigt wurde. Da entsprechend Gleichung (2.45) zwischen der Ausgangsebene und der Ebene der Transmissionsfunktion eine Fourierbeziehung besteht, resultiert in der Ausgangsebene eine Faltung des entsprechend Gleichung (2.48) skalierten Eingangsfelds mit der fouriertransformierten Transmissionsfunktion.

2.3 Wellenoptische Spezifikation von Transmissionsdesignproblemen

Nachdem in Abschnitt 2.3.1 eine Eingrenzung bezüglich der im Rahmen dieser Arbeit betrachteten Klasse von Designproblemen erfolgt, werden in Abschnitt 2.3.2 einige im Zusammenhang mit wellenoptischen Designproblemen häufig verwendete Gütefunktionen definiert.

2.3.1 Mono- und multifunktionale Transmissionsdesignprobleme

Für die Spezifikation wellenoptischer Designprobleme ist zunächst die Eingangslichtverteilung des betrachteten optischen Systems von Bedeutung. Die Eingangslichtverteilung fällt auf ein optisches System, welches durch eine geeignete Menge von Parametern beschrieben werden kann. Diese Menge kann in zwei Untermengen zerlegt werden: Zum einen die Menge der sogenannten *festen Systemparameter*, welche keine Freiheitsgrade beim Design darstellen, und zum anderen die Menge der *freien Systemparameter*, deren Werte beim Design hinsichtlich einer vorgegebenen Gütefunktion optimiert werden.

Einen in der Praxis wichtigen Aspekt stellen Fehler dar, welche bei der Umsetzung der beim Design ermittelten Werte für die freien Parameter in ein reales optisches System entstehen. Solche sogenannten *Umsetzungsfehler* lassen sich ebenfalls durch eine Menge geeigneter Parameter

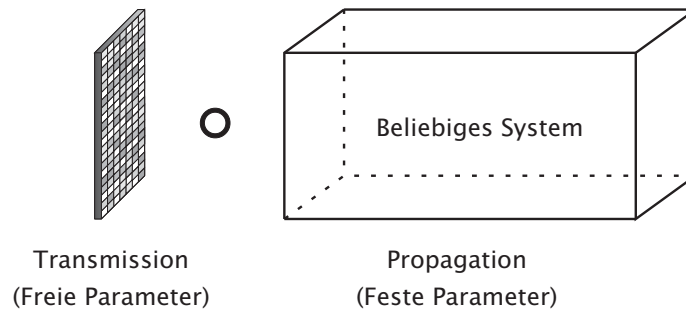


Abbildung 2.3: Zerlegung des bei Transmissionsdesignproblemen betrachteten optischen Systems

beschreiben. Umsetzungsfehler können beispielsweise durch Fabrikationsungenauigkeiten verursacht werden.

Die Zuordnung zwischen je einer Eingangslichtverteilung, einer Belegung der festen Systemparameter, einer Belegung der Parameter zur Beschreibung von Umsetzungsfehlern, sowie einer Gütefunktion bezeichnen wir als *optische Funktion*. Die Zahl der bei einem Designverfahren berücksichtigten optischen Funktionen wird im Folgenden mit B bezeichnet. Je nachdem, ob $B = 1$ oder $B \geq 1$ vorliegt, sprechen wir von einem *monofunktionalen* oder von einem *multifunktionalen* Design. Obwohl B für praktisch anwendbare Designalgorithmen im Allgemeinen endlich gewählt werden muss ($B < \infty$), kann ein multifunktionales Design auch zur Optimierung des Systemverhaltens in einem kontinuierlichen Parameterbereich benutzt werden, welcher sich anschaulich *zwischen* den einzelnen optischen Funktionen erstreckt.

Der Schwerpunkt dieser Arbeit liegt auf sogenannten *Transmissionsdesignproblemen*. Hierbei werden die Abtastwerte einer Transmissionsfunktion optimiert, während der Rest des optischen Systems beim Design fixiert ist. Teile des optischen Systems, welche sich vor der Transmissionsfunktion befinden, können durch Propagation der Eingangslichtverteilung in die Ebene unmittelbar vor der Transmissionsfunktion berücksichtigt werden, und es genügt somit, optische Systeme entsprechend des in Abbildung 2.3 gezeigten Schemas zu betrachten.

Innerhalb der einzelnen optischen Funktionen beschränken wir uns auf Eingangslichtverteilungen, welche sich durch die in Abschnitt 2.2.2 eingeführten abgetasteten komplexen Amplituden repräsentieren lassen und daher monochromatisch und kohärent sind. Das der optischen Funktion mit dem Index β ($\beta = 1 \dots B$) zugeordnete Eingangsfeld bezeichnen wir im Folgenden mit $u_{\text{in},\beta}(x)$. Die auf die Transmissionsfunktion $t(x)$ wirkenden Umsetzungsfehler beschreiben wir durch einen Operator E_β , welcher im Allgemeinen für jede optische Funktion verschieden sein kann. Das Ausgangsfeld $u_{\text{out},\beta}(x')$ ergibt sich somit gemäß

$$u_{\text{out},\beta}(x') = P_\beta ([E_\beta t(x)] u_{\text{in},\beta}(x))(x'). \quad (2.49)$$

Falls keine Umsetzungsfehler vorliegen, ergibt sich als Vereinfachung

$$u_{\text{out},\beta}(x') = P_\beta (t(x) u_{\text{in},\beta}(x))(x'). \quad (2.50)$$

Um Projektionsalgorithmen auf Transmissionsdesignprobleme anwenden zu können, werden entsprechend Abschnitt 2.1.2 Einschränkungsmengen benötigt. Eine Einschränkungsmenge resultiert

hierbei aus der Menge der als Lösung in Frage kommenden Transmissionsfunktionen, welche wir im Folgenden mit $\mathcal{M}_{\text{tran}}$ bezeichnen. Beispiele hierfür werden später in Abschnitt 2.4.1 diskutiert.

Neben dieser sogenannten *Transmissionseinschränkung* betrachten wir für jede optische Funktion eine Einschränkungsmenge $\mathcal{M}_{\text{sig},\beta}$ der für eine Lösung idealen Ausgangsfelder, welche im Folgenden auch als *Signalfelder* bezeichnet werden. Die *Signaleinschränkung* $\mathcal{M}_{\text{sig},\beta}$ sollte für jede optische Funktion passend zu der für diese optische Funktion spezifizierten Gütefunktion gewählt werden. Im folgenden Abschnitt werden verschiedene im Bereich des wellenoptischen Designs wichtige Gütefunktionen vorgestellt. Beispiele für Signaleinschränkungen mit zugehörigen Projektionsoperatoren werden in Abschnitt 2.4.1 diskutiert.

2.3.2 Wichtige Gütefunktionen

Das *Signal-Rausch-Verhältnis* (SNR) ist ein Maß für die Übereinstimmung zwischen einer (komplexwertigen) Signalfunktion $s(x)$ mit einer (fehlerbehafteten) Funktion $f(x)$, welches definiert ist über

$$\text{SNR}(f, s) = \frac{\int_{-\infty}^{\infty} dx' |f(x')|^2}{\int_{-\infty}^{\infty} dx' |f(x') - s(x')|^2} = \frac{\|f\|^2}{\|f - s\|^2}, \quad (2.51)$$

wobei der numerische Wert dieses Verhältnisses typischerweise in dB angegeben wird. Das SNR für ein Ausgangsfeld u_{out} im Vergleich zu einem Signalfeld u_{sig} kann demnach über die Substitutionen $f \rightarrow u_{\text{out}}$ und $s \rightarrow u_{\text{sig}}$ nach oben stehender Formel berechnet werden. Bei mehrelementigen Mengen \mathcal{M}_{sig} ist im Allgemeinen das bestmögliche SNR bezüglich aller Signalfelder $u_{\text{sig}} \in \mathcal{M}_{\text{sig}}$ von Interesse. Die Mengen \mathcal{M}_{sig} werden entsprechend sogenannter *Designfreiheiten* konstruiert, für die wir im Folgenden als Beispiele die Phasen-, Skalierungs- und Amplitudenfreiheit [57] betrachten.

Phasenfreiheit bedeutet, dass nur der Amplitudenbetrag des Ausgangsfelds von Interesse ist, was sich in den Substitutionen $f \rightarrow |u_{\text{out}}|$ und $s \rightarrow |u_{\text{sig}}|$ zur Berechnung des SNRs gemäß Gleichung (2.51) widerspiegelt.

Bei Vorhandensein von *Skalierungsfreiheit* kann die Signalfunktion $s(x')$ mit einem beliebigen Faktor $\alpha \in \mathbb{C}$ skaliert werden. Der optimale Skalierungsfaktor α zur Maximierung des SNRs entsprechend Gleichung (2.51) mit $s \rightarrow \alpha s$ ergibt sich als

$$\alpha = \frac{\int_{-\infty}^{\infty} dx' f(x') s^*(x')}{\int_{-\infty}^{\infty} dx' |s(x')|^2} = \frac{\langle f | s \rangle}{\|s\|^2}. \quad (2.52)$$

Im Fall von *Amplitudenfreiheit* ist nur das Ausgangsfeld in einem bestimmten lateralen Bereich $x' \in \mathcal{W}_{\text{sig}}$ von Interesse, wobei die Menge \mathcal{W}_{sig} auch als *Signalfenster* bezeichnet wird. Die Integrationen in den Gleichungen (2.51) und (2.52) erstrecken sich dann nur noch über Positionen $x' \in \mathcal{W}_{\text{sig}}$.

Bei Vorliegen aller drei genannten Typen von Designfreiheiten lässt sich demnach das SNR für

ein Ausgangsfeld u_{out} schreiben als

$$\text{SNR} = \frac{\int_{x' \in \mathcal{W}_{\text{sig}}} dx' |u_{\text{out}}(x')|^2}{\int_{x' \in \mathcal{W}_{\text{sig}}} dx' (|u_{\text{out}}(x')| - \alpha |u_{\text{sig}}(x')|)^2} \quad (2.53)$$

mit dem optimalen Skalierungsfaktor

$$\alpha = \frac{\int_{x' \in \mathcal{W}_{\text{sig}}} dx' |u_{\text{out}}(x')| |u_{\text{sig}}(x')|}{\int_{x' \in \mathcal{W}_{\text{sig}}} dx' |u_{\text{sig}}(x')|^2}. \quad (2.54)$$

Während bei der Berechnung des SNRs alle Positionen $x' \in \mathcal{W}_{\text{sig}}$ einfließen, bestimmt sich das insbesondere bei Strahlteilern häufig verwendete Maß des *Uniformitätsfehlers* lediglich aus den größten auftretenden Abweichungen zwischen Ausgangs- und Signalfeld. Bezeichnen wir die größten und kleinsten auftretenden Quotienten zwischen Ausgangsintensität und skaliertes Signalintensität mit \hat{I}_{min} bzw. \hat{I}_{max} mit

$$\hat{I}_{\text{min}} = \min_{x' \in \mathcal{W}_{\text{sig}}} \frac{|u_{\text{out}}(x')|^2}{|\alpha u_{\text{sig}}(x')|^2} \quad \text{und} \quad (2.55a)$$

$$\hat{I}_{\text{max}} = \max_{x' \in \mathcal{W}_{\text{sig}}} \frac{|u_{\text{out}}(x')|^2}{|\alpha u_{\text{sig}}(x')|^2}, \quad (2.55b)$$

so berechnet sich der Uniformitätsfehler E_{unif} gemäß Referenz [17]

$$E_{\text{unif}} = \frac{\hat{I}_{\text{max}} - \hat{I}_{\text{min}}}{\hat{I}_{\text{max}} + \hat{I}_{\text{min}}}. \quad (2.56)$$

In den Gleichungen (2.55) wurde angenommen, dass die Ausgangsintensität an einer Position x' proportional zu $|u_{\text{out}}(x')|^2$ ist, was nur in paraxialer Näherung gültig ist. Bei Verwendung dieser Näherung wird im Folgenden nicht mehr explizit auf die damit verbundene Einschränkung auf paraxiale Systeme hingewiesen.

Während mit Hilfe der beiden oben genannten Gütefunktionen die Abweichung zwischen einem Ausgangsfeld und der gewünschten Signalverteilung gemessen werden kann, steht bei den nun folgenden Effizienzmaßen der Wirkungsgrad des optischen Systems im Vordergrund. Die *Fenster-effizienz* η_{win} wird berechnet als das Verhältnis zwischen dem Leistungsanteil des Ausgangsfelds im Signalfenster und der Leistung des Eingangsfelds,

$$\eta_{\text{win}} = \frac{\int_{x' \in \mathcal{W}_{\text{sig}}} dx' |u_{\text{out}}(x')|^2}{\int_{-\infty}^{\infty} dx' |u_{\text{in}}(\xi)|^2}. \quad (2.57)$$

Bei der Berechnung der *Konversionseffizienz* besteht das Ziel darin, den Leistungsanteil des Ausgangsfelds zu ermitteln, welcher in das Signal *fließt* [61]. Mit Hilfe des optimalen Skalierungsfaktors α aus Gleichung (2.54) kann die Konversionseffizienz als Verhältnis zwischen skaliertes Signalleistung und Eingangsleistung dargestellt werden:

$$\eta_{\text{conv}} = |\alpha|^2 \frac{\int_{x' \in \mathcal{W}_{\text{sig}}} dx' |u_{\text{sig}}(x')|^2}{\int_{-\infty}^{\infty} dx' |u_{\text{in}}(x')|^2}. \quad (2.58)$$

Für fehlerfreie Ausgangsfelder sind Konversions- und Fenstereffizienz miteinander identisch. In Fällen, in denen die Unterscheidung zwischen Konversions- und Fenstereffizienz irrelevant ist, wird somit in dieser Arbeit auch die verkürzte Bezeichnung *Effizienz* verwendet werden.

2.4 Anwendung von Projektionsalgorithmen für monofunktionale Transmissionsdesignprobleme

In diesem Abschnitt wird gezeigt, wie sich das in Abschnitt 2.1 eingeführte Konzept der Projektionsalgorithmen zur Lösung monofunktionaler Transmissionsdesignprobleme einsetzen lässt. Die Aufhebung der hierbei getroffenen Einschränkung auf Monofunktionalität ist Hauptgegenstand der Kapitel 5 und 6 dieser Arbeit. In Abschnitt 2.3 wurden bereits die Begriffe der Transmissions- und Signaleinschränkung eingeführt, für die wir in den beiden folgenden Abschnitten 2.4.1 bzw. 2.4.2 konkrete Beispiele betrachten werden.

Da die zugehörigen Einschränkungsmengen jedoch in verschiedenen Domänen definiert werden, welche beispielsweise über Gleichung (2.50) miteinander verbunden sind, müssen für eine Anwendung der in Abschnitt 2.1 eingeführten Projektionsalgorithmen die Einschränkungsmengen geeignet umformuliert werden, was zu dem in Abschnitt 2.4.3 eingeführten Konzept des Iterativen Fouriertransformationsalgorithmus (IFTA) führt. In Abschnitt 2.4.4 untersuchen wir, unter welchen Bedingungen der IFTA bezüglich des in Abschnitt 2.1.2 eingeführten SDEs konvergiert. Abschnitt 2.4.5 stellt schließlich mit dem Input-Output-Konzept eine für Kapitel 4 dieser Arbeit wichtige Verallgemeinerung des IFTAs dar.

2.4.1 Projektionsoperatoren für typische Transmissionseinschränkungen

Wir betrachten im Folgenden die in der Praxis am häufigsten auftretenden Transmissionseinschränkungen. Dazu gehört die sogenannte *Phaseneinschränkung*, welche besagt, dass die Transmissionsfunktion eine reine Phasenfunktion ist. Diese Anforderung ist durch die Überführbarkeit der Transmissionsfunktion in ein absorptionsfreies höhen- oder brechzahlmoduliertes optisches Element mittels der Dünne-Elemente-Approximation entsprechend Abschnitt 2.2.6 motiviert. Für die Phaseneinschränkung stellt sich die Einschränkungsmenge $\mathcal{M}_{\text{tran}}$ dar als $\mathcal{M}_{\text{tran}} = \mathcal{M}_{\text{phase}}$ mit

$$\mathcal{M}_{\text{phase}} = \{u \in \mathcal{H} : \forall x \exists \phi : u(x) = \exp(i\phi)\}. \quad (2.59)$$

Der entsprechend Abschnitt 2.1.1 punktweise abstandsminimierende Projektionsoperator für $\mathcal{M}_{\text{phase}}$ ergibt sich als

$$\Pi_{\text{phase}} t(x) = t(x)/|t(x)|. \quad (2.60)$$

Während bei Vorliegen einer Phaseneinschränkung die Transmissionsfunktion an jeder Position x einen beliebigen Phasenwert $t(x) = \exp(i\phi(x))$ annehmen kann, bestehen in der Praxis – ebenfalls häufig motiviert durch die Herstellbarkeit der korrespondierenden optischen Elemente – in vielen

Fällen Einschränkungen an den Wertebereich von $\phi(x)$. Ein Beispiel hierfür stellt die Beschränkung auf Q Phasenwerte zur äquidistanten Abdeckung des Phasenintervalls $[0, 2\pi)$ dar, was der Einschränkungsmenge

$$\mathcal{M}_{\text{quant},Q} = \left\{ u \in \mathcal{H} : \forall x : u(x) \in \widetilde{\mathcal{M}}_{\text{quant},Q} \right\} \quad \text{mit} \quad (2.61a)$$

$$\widetilde{\mathcal{M}}_{\text{quant},Q} = \left\{ \exp(i2\pi j/Q) : j \in \mathbb{Z} \right\} \quad (2.61b)$$

entspricht, wobei wir den zugehörigen punktweise abstandsminimierenden Quantisierungsoperator mit $\Pi_{\text{quant},Q}$ bezeichnen.

2.4.2 Projektionsoperatoren für typische Signaleinschränkungen

Die in Abschnitt 2.3.2 eingeführten Konzepte der Phasen-, Amplituden- und Skalierungsfreiheit spiegeln sich nicht nur in den dort aufgeführten Gütefunktionen wider, sondern auch in den zugehörigen Signalprojektionsoperatoren. Im Fall von Phasenfreiheit besteht die punktweise abstandsminimierende Projektion darin, an jeder Position x die Amplitude durch die Signalamplitude zu ersetzen, wobei die Phase unangetastet bleibt. Der zugehörige Projektionsoperator Π_P lässt sich schreiben als

$$\Pi_P u_{\text{out}}(x') = \begin{cases} |u_{\text{sig}}(x')| \exp [i \arg (u_{\text{out}}(x'))] & \text{falls } x' \in \mathcal{W}_{\text{sig}} \\ 0 & \text{sonst} \end{cases}, \quad (2.62)$$

wobei u_{out} und u_{sig} für Ausgangs- bzw. Signalfeld stehen.

Falls zusätzlich Amplituden- und Skalierungsfreiheit bestehen, ergibt sich der ebenfalls punktweise abstandsminimierende Projektionsoperator Π_{PAS} ,

$$\Pi_{\text{PAS}} u_{\text{out}}(x') = \begin{cases} |\alpha u_{\text{sig}}(x')| \exp [i \arg (u_{\text{out}}(x'))] & \text{falls } x' \in \mathcal{W}_{\text{sig}} \\ u_{\text{out}}(x') & \text{sonst} \end{cases}, \quad (2.63)$$

mit α gemäß Gleichung (2.54).

Während die Berechnung von α entsprechend Gleichung (2.54) dem Vorliegen vollständiger Skalierungsfreiheit entspricht, ist in Teilen dieser Arbeit auch das Konzept der *eingeschränkten Skalierungsfreiheit* von Bedeutung. So kann beispielweise eine untere Schranke α_l für den Betrag des Skalierungsfaktors vorgegeben werden, auf welche der Skalierungsfaktor gesetzt wird, sofern die Berechnung entsprechend Gleichung (2.54) einen kleineren Wert ergibt. Das Ansetzen einer unteren Schranke α_l für den Skalierungsfaktor entspricht der Vorgabe einer Zieleffizienz, welche der Effizienz des angenommenen Ausgangsfelds $u_{\text{out}} = \alpha_l u_{\text{sig}}$ entspricht.

2.4.3 Der Iterative Fouriertransformationsalgorithmus

Der *Iterative Fouriertransformationsalgorithmus* (IFTA) ist ein Projektionsalgorithmus, der zur Lösung monofunktionaler Transmissionsdesignprobleme eingesetzt werden kann. Das Grundsche-ma dieses Algorithmus, der von Liu und Gallagher [8] in das Optikdesign eingeführt wurde, ist in

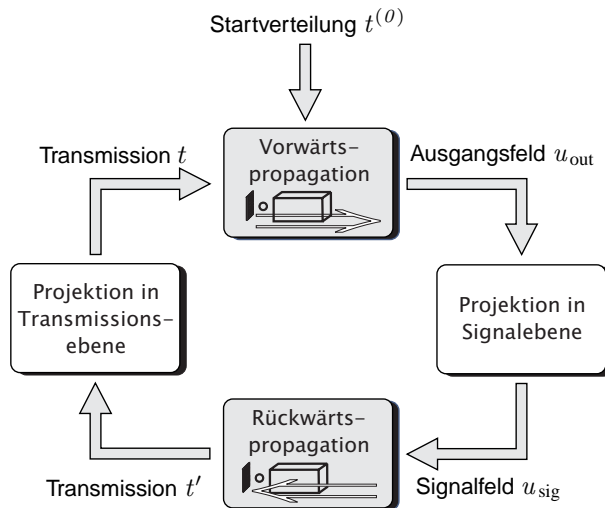


Abbildung 2.4: Schritte innerhalb eines Zyklus des Iterativen Fouriertransformationsalgorithmus

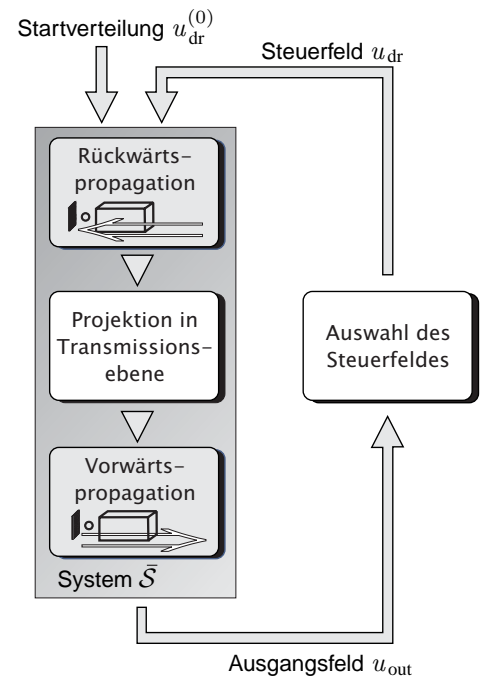


Abbildung 2.5: Input-Output-Konzept als Verallgemeinerung des IFTAs aus Abbildung 2.4

Abbildung 2.4 dargestellt. In Iteration n ergeben sich Ausgangsfeld $u_{\text{out}}^{(n)}$ sowie Transmissionsfunktion $t^{(n)}$ gemäß

$$u_{\text{out}}^{(n)} = P u_{\text{in}} t^{(n-1)} \quad \text{sowie} \quad (2.64a)$$

$$t^{(n)} = \Pi_{\text{tran}} \frac{P^{-1} \Pi_{\text{sig}} u_{\text{out}}^{(n)}}{u_{\text{in}}}. \quad (2.64b)$$

Jeder Iterationszyklus besteht demnach aus der abwechselnden Anwendung von Propagations- und Projektionsschritten, wobei die beiden Projektionsoperatoren zum Übergang zwischen Transmissions- und Signalebene zueinander invers sind. Beispiele für Propagations- und Projektionsoperatoren wurden in den Abschnitten 2.2 bzw. 2.4.1 und 2.4.2 gegeben.

Für den Begriff des Iterativen Fouriertransformationsalgorithmus gibt es keine allgemein akzeptierte scharfe Abgrenzung. Im Allgemeinen wird von einem IFTA gesprochen, falls die Propagationsoperatoren auf Fouriertransformationen basieren, was insbesondere auf die Fraunhofer- und Fresnel-Ausbreitung aus Abschnitt 2.2.5 zutrifft. Wie wir im Folgenden Abschnitt 2.4.4 zeigen werden, kann der IFTA als Spezialfall des in Abschnitt 2.1.4 eingeführten seriellen Projektionsalgorithmus mit zwei Mengen dargestellt werden, wobei beide Propagationsoperatoren mit einer der beiden Projektionen zusammengefasst werden.

2.4.4 Konvergenznachweis

Zum Nachweis der Konvergenz des IFTAs betrachten wir die folgenden beiden Einschränkungen:

$$\mathcal{M}_1 = \{P^{-1} u_{\text{sig}} : u_{\text{sig}} \in \mathcal{M}_{\text{sig}}\} \quad \text{und} \quad (2.65a)$$

$$\mathcal{M}_2 = \{u_{\text{in}} t : t \in \mathcal{M}_{\text{tran}}\}, \quad (2.65b)$$

wobei \mathcal{M}_1 entsprechend Abschnitt 2.1.1 eine implizite Einschränkung darstellt. Nach den in Anhang A.1 gezeigten Sätzen 2.67 und 2.68 definieren die auf Basis der Projektionsoperatoren Π_{tran} und Π_{sig} konstruierten Operatoren

$$\Pi_1 u = P^{-1} \Pi_{\text{sig}} P u \quad \text{und} \quad (2.66a)$$

$$\Pi_2 u = u_{\text{in}} \Pi_{\text{tran}} u_{\text{in}}^{-1} u \quad (2.66b)$$

Projektionen auf die durch die Gleichungen (2.65) gegebenen Mengen \mathcal{M}_1 bzw. \mathcal{M}_2 .

Satz 2.67 *Der Operator Π_1 nach Gleichung (2.66a) definiert eine Projektion auf \mathcal{M}_1 gemäß Gleichung (2.65a), sofern der Propagationsoperator P die Eigenschaften (2.39) erfüllt.*

Satz 2.68 *Der Operator Π_2 nach Gleichung (2.66b) definiert eine Projektion auf \mathcal{M}_2 gemäß Gleichung (2.65b), sofern Π_{tran} ein punktweise abstandsminimierender Projektionsoperator auf $\mathcal{M}_{\text{tran}}$ ist.*

Betrachten wir nun den Algorithmus

$$u^{(n)} = \Pi_2 \Pi_1 u^{(n-1)} \quad (2.69)$$

mit den durch die Gleichungen (2.66a) und (2.66b) gegebenen Projektionsoperatoren. Für diesen Algorithmus ist wegen Satz 2.16 die Konvergenzeigenschaft (2.15) erfüllt.

Da die Algorithmen (2.64) und (2.69) für $u^{(n)} = u_{\text{in}} t^{(n)}$ miteinander äquivalent sind, folgt für den IFTA die Konvergenzeigenschaft

$$J(u_{\text{in}} t^{(n)}) \leq J(u_{\text{in}} t^{(n-1)}). \quad (2.70)$$

Mit Einführung der modifizierten Form des SDEs

$$J_{\text{in}}(t) = J(u_{\text{in}} t) \quad (2.71)$$

erhalten wir

$$J_{\text{in}}(t^{(n)}) \leq J_{\text{in}}(t^{(n-1)}). \quad (2.72)$$

Zum Abschluss dieses Abschnitts betrachten wir eine zu Gleichung (2.65) alternative Form der Definition der Einschränkungsmengen

$$\mathcal{M}'_1 = \{u_{\text{in}}^{-1} P^{-1} u_{\text{sig}} : u_{\text{sig}} \in \mathcal{M}_{\text{sig}}\}, \quad (2.73a)$$

$$\mathcal{M}'_2 = \mathcal{M}_{\text{tran}}. \quad (2.73b)$$

In diesem Fall kann der zu \mathcal{M}'_1 gehörige Projektionsoperator Π'_1 nicht ohne weiteres aus Π_{sig} konstruiert werden, bzw. die Ausführung von Π'_1 wäre mit einem deutlich höheren Rechenaufwand verbunden. Der Operator

$$\Pi'_1 u = u_{\text{in}}^{-1} P^{-1} \Pi_{\text{sig}} P u_{\text{in}} u \quad (2.74)$$

erfüllt die Abstandsminimierungseigenschaft (2.3) für nichtkonstante Eingangsfelder u_{in} im Allgemeinen nicht.

Diese Betrachtung zeigt, dass der IFTA im Allgemeinen nur im Fall konstanter Eingangsfelder $u_{\text{in}}(x)$ bezüglich $J(t^{(n)})$ konvergiert.

2.4.5 Input-Output-Konzept

Zur Beschreibung des Input-Output-Konzepts betrachten wir einen zunächst abstrakten Systemoperator \bar{S} , welcher eine Eingangsinformation f_{in} in eine Ausgangsinformation f_{out} transformiert:

$$f_{\text{out}} = \bar{S} f_{\text{in}}. \quad (2.75)$$

Bei Input-Output-Algorithmen besteht das Ziel darin, ein f_{in} zu finden, sodass das zugehörige f_{out} bestimmte Eigenschaften erfüllt, die beispielsweise durch eine Gütefunktion spezifiziert werden. Unter Kenntnis bestimmter Eigenschaften von \bar{S} ist es häufig möglich, einen iterativen Algorithmus zu formulieren, der jeder Iteration ein f_{in} mit dem Ziel wählt, die f_{out} hinsichtlich der Optimierung der vorgegebenen Gütefunktion zu steuern.

Betrachten wir das Grundschemata des Iterativen Fouriertransformationsalgorithmus in Abbildung 2.4, so bieten sich innerhalb des Input-Output-Konzepts zwei Möglichkeiten zur Verallgemeinerung dieses Algorithmus an. Die erste Möglichkeit, welche in Abbildung 2.5 dargestellt ist, ergibt sich aus der Zusammenfassung von Rückpropagation, Transmissionsprojektion und Vorwärtspropagation zu dem Operator \bar{S} . In diesem Fall wird die Signalprojektion durch eine verallgemeinerte Operation ersetzt. Das aus einer Anwendung dieser Operation resultierende Feld bezeichnen wir jetzt nicht mehr als Signalfeld, sondern als *Steuerfeld* u_{dr} (*Driving Field*). Die Motivation für diese Begriffswahl wurde im vorangegangenen Absatz gegeben. Aus der Anwendung von \bar{S} auf ein Steuerfeld u_{dr} , resultiert ein Ausgangsfeld u_{out} entsprechend

$$u_{\text{out}}^{(n)} = \bar{S} u_{\text{dr}}^{(n)} = P u_{\text{in}} \Pi_{\text{tran}} \frac{(P^{-1} u_{\text{dr}}^{(n)})}{u_{\text{in}}}, \quad (2.76)$$

wobei der IFTA entsprechend Gleichung (2.64) der Wahl

$$u_{\text{dr}}^{(n)} = \Pi_{\text{sig}} u_{\text{out}}^{(n-1)} \quad (2.77)$$

entspricht.

Alternativ könnte der IFTA auch durch Ersetzen der Transmissionsprojektion in einen Input-Output-Algorithmus verallgemeinert werden. Die erstgenannte Variante hat jedoch den Vorteil, dass nach jeder Iteration die Transmissionseinschränkungen erfüllt sind, was im Allgemeinen eine Voraussetzung für die Überführbarkeit der erhaltenen Transmission in ein reales optisches Element darstellt.

In Kapitel 4 wird das Konzept des Input-Output-Algorithmus benutzt, um einen Projektionsalgorithmus mit im Vergleich zum IFTA verbesserten Konvergenzeigenschaften zu konstruieren.

Kapitel 3

Diffraktive Strahlteilung

Diffraktive Strahlteiler werden zur Replizierung eines Eingangstrahls unter vorgegebenen Winkeln und relativen Wichtungen berechnet, wobei die Strahlvervielfältigung auf Beugungseffekten beruht. Im Folgenden bezeichnen wir diffraktive Strahlteiler auch verkürzt als Strahlteiler, da nichtdiffraktive Strahlteiler, welche beispielsweise über halbdurchlässige Schichtensysteme oder über Prismenspiegel realisiert werden können, im Rahmen dieser Arbeit nicht betrachtet werden. Die ersten Anwendungen von Strahlteilern erfolgten zur Replikation von Belichtungsmustern bei der Herstellung elektronischer Schaltkreise [62, 63]. Weitere Anwendungsfelder bestehen im Rahmen der optischen Informationsverarbeitung und -rechentechnik [64–66], der optischen Messtechnik [67], der Materialbearbeitung [15, 68–70], der Lasertechnik [71, 72] und der Telekommunikation. Nicht zuletzt werden Strahlteiler auch für visuelle Anwendungen eingesetzt, wie beispielsweise zur Markierung von Positionen oder Regionen in Mess- oder Justiersystemen.

Nach Darstellung einiger grundlegender Konzepte für das Design von Strahlteilern in Abschnitt 3.1 zeigen wir in Abschnitt 3.2, wie die mit *konventionellen* Designtechniken verbundenen Einschränkungen hinsichtlich der Ablenkwinkel der replizierten Strahlen aufgehoben werden können. In dem darauf folgenden Abschnitt 3.3 werden die für die praktische Anwendung wichtigen Toleranzeigenschaften von Strahlteilern diskutiert. Da bei diffraktiven Strahlteilern neben den gewünschten Replikaten des Eingangstrahls, welche auch als *Signalordnungen* bezeichnet werden, stets auch als *Rauschordnungen* bezeichnete ungewünschte Replikationen auftreten, kann in manchen Anwendungen die Intensität solcher Rauschordnungen eine kritische Eigenschaft darstellen. In Abschnitt 3.4 werden daher Möglichkeiten zur Begrenzung der maximalen Rauschintensität diskutiert.

3.1 Vorbetrachtungen

Nach einer Betrachtung der wellenoptischen Modellierung von Strahlteilungsproblemen in Abschnitt 3.1.1 werden in Abschnitt 3.1.2 Pixelierungseffekte diskutiert, welche bei der Wahl der Abtastparameter entsprechend Abschnitt 3.1.3 berücksichtigt werden müssen. In Abschnitt 3.1.4 wird schließlich ein in der Praxis bewährtes Schema zur Anwendung des aus Abschnitt 2.4.3 bekannten iterativen Fouriertransformationsalgorithmus beim Strahlteilerdesign dargestellt.

3.1.1 Wellenoptische Modellierung

Wir betrachten im Folgenden Strahlteilungsprobleme, welche sich entsprechend Abschnitt 2.3.1 als Transmissionsdesignprobleme mit der in Abbildung 2.3 gezeigten Grundzerlegung des optischen Systems modellieren lassen. Für den auf die Transmissionsfunktion folgenden Teil des optischen Systems kommen neben homogenen Regionen, welche über die Ausbreitungsoperatoren aus Abschnitt 2.2.5 modelliert werden können, auch Systeme mit Linsen in Frage. In Abschnitt 2.2.7 wurden bereits Beispiele für einfache Linsensysteme dargestellt, welche insbesondere auch im Zusammenhang mit Strahlteilern eingesetzt werden.

Bei der Berechnung des Ausgangsfelds nach Gleichung (2.50) ergibt sich für den Fall der Fresnel-Ausbreitung aus Gleichung (2.37) unter Zuhilfenahme des Faltungstheorems [46] die Beziehung

$$u_{\text{out}}(x') = P_{\Delta z, n}^{\text{Fresnel}}(u_{\text{in}} t) \sim \exp\left[\frac{ik}{2\Delta z}x'^2\right] \left(\mathcal{F} t \star \mathcal{F} \left(u_{\text{in}}(x) \exp\left[\frac{ik}{2\Delta z}x^2\right] \right) \right). \quad (3.1)$$

Für die Ausgangsintensitätsverteilung ergibt sich somit der schematische Zusammenhang

$$|u_{\text{out}}|^2 \sim |\mathcal{F} t \star u_{\text{transf}}|^2, \quad (3.2)$$

wobei entsprechend Gleichung (2.37) zwischen den Transmissions- und Signalebenekoordinaten eine Skalierungsbeziehung besteht. Aus der Betrachtung des Spezialfalls $t(x) \equiv 1$ wird klar, dass die Intensitätsverteilung $|u_{\text{transf}}|^2$ der Ausgangsintensität des Systems ohne Transmission entspricht.

Eine Faltungsbeziehung entsprechend Gleichung (3.2) ergibt sich ebenfalls für die Fraunhofer-Ausbreitung (Abschnitt 2.2.5), für die paraxiale Modellierung der in Abbildung 2.2 gezeigten Linsensysteme, sowie allgemein für die Propagation durch beliebige innerhalb der paraxialen Matrixoptik beschreibbare Systeme unter Benutzung des Collins-Integrals [73]. Die Skalierung der Signalebenekoordinaten hängt dabei von dem konkreten optischen System ab. Neben der betrachteten Gültigkeit von Gleichung (3.2) für Ortskoordinaten im paraxialen Fall, ergibt sich im nichtparaxialen Fall eine Gültigkeit dieser Faltungsbeziehung im Winkelspektrum für Wellenzahlkoordinaten.

Da bei Betrachtung des Produkts $u_{\text{in}} t$ eine Apertur in der Transmissionsebene stets dem Eingangsfeld u_{in} zugeordnet werden kann, gehen wir im Folgenden ohne Beschränkung der Allgemeinheit von einer periodischen Transmissionsfunktion aus. Bei einer Periodenlänge Δx lässt sich $t(x)$ somit darstellen als

$$t(x) = \text{comb}(x/\Delta x) \star t_p(x), \quad (3.3)$$

wobei $t_p(x)$ eine nur über den Intervall $[0, \Delta x)$ ausgedehnte Funktion bezeichnet und die Kammfunktion definiert ist über [46]

$$\text{comb}(x) = \sum_{n=-\infty}^{\infty} \delta(x - n). \quad (3.4)$$

Somit besteht $\mathcal{F}t$ aus diskreten Delta-Funktionen,

$$[\mathcal{F}t](x') = \text{comb}(x'/\delta x')[\mathcal{F}t_p](x'), \quad (3.5)$$

welche auf einem äquidistanten Raster mit Abstand

$$\delta x' = 1/\Delta x \quad (3.6)$$

liegen. Diese Delta-Funktionen führen zu Replikationen des transformierten Eingangsfelds an ihren jeweiligen Positionen und entsprechen somit den einzelnen *Ordnungen* eines Strahlteilers. Je nachdem, ob eine Ordnung innerhalb oder außerhalb des Signalfensters \mathcal{W}_{sig} (vgl. Abschnitt 2.3.2) liegt, wird sie als *Signal-* oder *Rauschordnung* bezeichnet.

Von einem *überlappungsfreien Strahlteiler* sprechen wir, wenn Interferenzeffekte zwischen verschiedenen Ordnungen vernachlässigbar sind, d. h.

$$\text{width}(u_{\text{transf}}) < \delta x', \quad (3.7)$$

wobei $\text{width}(u_{\text{transf}})$ ein geeignetes Maß für den Durchmesser des transformierten Eingangsfelds bezeichnet. Bei Erfüllung der Bedingung (3.7) bleibt die optische Wirkung eines Strahlteilers unabhängig vom konkreten Eingangsfeld sowie vom optischen System. Insbesondere hängen bei vorliegender Überlappungsfreiheit Gütefunktionen wie Effizienz und Uniformität (s. Abschnitt 2.3.2) nicht von u_{transf} ab. Beim Design sowie der Analyse derartiger Strahlteiler ist es demnach ausreichend, nur den Anteil $\mathcal{F}t$ der Faltungsbeziehung (3.2) zu betrachten.

Die beim Design relevanten Signalebenenkoordinaten x' ergeben sich somit aus der Fouriertransformationsbeziehung (2.36a). Hieraus kann über

$$k_x = 2\pi x' \quad (3.8)$$

die x -Komponente des Wellenzahlvektors berechnet werden. Der Winkel ϕ bezüglich der optischen Achse resultiert für eine Ordnung x' gemäß

$$\sin \phi = \frac{k_x}{k} = x'\lambda. \quad (3.9)$$

Die Aufteilung der Leistung des Eingangsstrahls auf die einzelnen Ordnungen erfolgt entsprechend Gleichung (3.5) bei überlappungsfreien Strahlteilern proportional zu dem Modulationsfaktor $|\mathcal{F}t_p(x')|^2$. Zur Bezeichnung des Leistungsanteils wird im Folgenden häufig vereinfacht von der Intensität von Ordnungen gesprochen, wobei bei Vorliegen von Überlappungsfreiheit eine Proportionalität zwischen dem Leistungsanteil und der Maximalintensität einer Ordnung besteht.

Im Folgenden wird zunächst das Design überlappungsfreier Strahlteiler näher betrachtet. Designmethoden für nichtüberlappungsfreie Strahlteiler werden in Abschnitt 3.2 dargestellt.

3.1.2 Pixelierungseffekte

Laut vorangegangenem Abschnitts basiert das Design überlappungsfreier Strahlteiler auf der Berechnung der Fouriertransformation (FT) der Transmissionsfunktion. Bei abgetasteten Transmissionsfunktionen bietet sich hierzu die schon in Abschnitt 2.2.5 erwähnte schnelle Fouriertransformation (FFT) an, bei der als Ergebnis die FT auf einem Ausgangsraster entsprechend Gleichung (3.6)

resultiert. Da – wie schon in Abschnitt 2.2.2 erwähnt – insbesondere Phasentransmissionen wegen der fehlenden Bandbegrenzung [47] streng mathematisch gesehen nicht korrekt abtastbar sind, können sich auch für die aus der FFT resultierenden Abtastwerte der Fouriertransformation Fehler ergeben. Im Fall pixelierter Transmissionsfunktionen können jedoch – wie im Folgenden gezeigt wird – diese Fehler größtenteils korrigiert werden.

Wir stellen hierzu die Transmission $t(x)$ als Faltung einer modulierten Kammfunktion mit einer für die Form eines einzelnen Pixels beschreibenden Rechteckfunktion dar:

$$t(x) = [t_s(x) \times \text{comb}(x/\delta x)] \star \text{rect}(x/\delta x), \quad \text{wobei} \quad (3.10a)$$

$$\text{rect}(x) = \begin{cases} 1 & \text{falls } x \in [-\frac{1}{2}, \frac{1}{2}) \\ 0 & \text{sonst} \end{cases}. \quad (3.10b)$$

Die FT hiervon ergibt sich als

$$[\mathcal{F}t](x') \sim [[\mathcal{F}t_s](x') \star \text{comb}(x'/\Delta x')] \times \text{sinc}(\pi x'/\Delta x') \quad \text{mit} \quad (3.11a)$$

$$\text{sinc}(x) = \sin x/x. \quad (3.11b)$$

Die Faltung mit der Kammfunktion ergibt hierbei eine periodische Wiederholung des beispielsweise über eine FFT ermittelbaren Spektrums $[\mathcal{F}t_s](x')$, wobei das Ergebnis global über eine sinc-Funktion moduliert wird.

Aufgrund dieser Modulation ergibt sich einerseits eine relative Abschwächung der äußeren Ordnungen, welche durch Einbeziehung des Modulationsfaktors in den beim IFTA benutzten Propagationsoperator beim Design berücksichtigt werden kann. Alternativ können auch die Zielintensitäten der Signalordnungen entsprechend dieser sinc-Modulation präkompensiert werden.

Einen zweiten Effekt der Pixelierung stellen abgeschwächte Wiederholungen der zentralen n_x Ordnungen des Strahlteilers dar, welche auf Grund der in Gleichung (3.11a) vorliegenden Faltung mit der Kammfunktion auftreten. Dies führt einerseits zu einer Effizienzreduktion um den Leistungsanteil η_{pix} , welcher in den zentralen n_x Ordnungen auf Grund der sinc-Modulation verloren geht:

$$\eta_{\text{pix}} = \frac{\int_{-\Delta x'/2}^{\Delta x'/2} dx' |[\mathcal{F}t_s](x')|^2 |1 - \text{sinc}(\pi x'/\Delta x')|^2}{\int_{-\Delta x'/2}^{\Delta x'/2} dx' |[\mathcal{F}t_s](x')|^2}. \quad (3.12)$$

Des Weiteren kann die Intensität der durch die periodische Wiederholung von Signalordnungen entstehenden Rauschordnungen in manchen Anwendungen kritisch sein. Da jede an einer Position x' lokalisierte Signalordnung auf einem Raster mit dem Abstand $\Delta x'$ wiederholt wird, ergibt sich die Wiederholung dieser Ordnung mit der maximalen relativen Intensität im Fall von $x' > 0$ an der Position $x' - \Delta x$ und im Fall von $x' < 0$ an der Position $x' + \Delta x$. Durch Quadrierung des Verhältnisses der entsprechenden sinc-Faktoren kann somit die maximale relative Intensität der durch die periodische Wiederholung für jede Signalordnung resultierenden Rauschordnungen abgeschätzt werden.

3.1.3 Wahl der Abtastparameter

Bei der Wahl der Abtastparameter der Transmissionsfunktion empfiehlt es sich, zunächst den Abtastabstand δx zu ermitteln, bevor die Anzahl n_x der Abtastwerte in einer Periode der Transmissionsfunktion bestimmt wird. Für die Wahl beider Parameter werden im Folgenden Kriterien angegeben.

Der Abtastabstand δx muss in jedem Fall klein genug gewählt werden, sodass der Ablenkwinkel ϕ_{\max} der äußersten Signalordnung kleiner ist als der Ablenkwinkel der ersten Ordnung eines binären Gitters mit einer Periodenlänge von zwei Abtastwerten. Es ergibt sich somit die Forderung

$$\sin \phi_{\max} < \frac{\lambda}{2\delta x}, \quad (3.13)$$

welche sich entsprechend Gleichung (3.9) für die Position x'_{\max} der äußersten Signalordnung umschreiben lässt als

$$x'_{\max} < \frac{1}{2\delta x}. \quad (3.14)$$

Als zweites Kriterium bei der Wahl von δx müssen im Fall einer späteren pixelierten Realisierung der Transmissionsfunktion die entsprechend Abschnitt 3.1.2 entstehenden Effizienzverluste η_{pix} sowie die Intensität der durch Pixelierung entstehenden Rauschordnungen berücksichtigt werden. Beide Effekte können durch eine Verringerung von δx vermindert werden, sodass sich bei Vorgabe eines maximal akzeptablen Effizienzverlusts η_{pix} oder einer maximalen Rauschintensität wiederum eine obere Schranke für δx ergibt, welche typischerweise stärker ist, als die durch Gleichung (3.14) resultierende Beschränkung.

Falls der Abtastabstand δx nicht deutlich größer als die Wellenlänge ist, kann entsprechend Abschnitt 2.2.6 die Gültigkeit der Dünne-Elemente-Approximation verletzt werden. Wenn die Transmissionsfunktion später unter Benutzung dieser Approximation in ein optisches Element überführt werden soll, ist es demnach notwendig, durch geeignete Maßnahmen während des Designs [74] sicherzustellen, dass die entstehenden Strukturgrößen trotz kleinerem Abtastabstands nicht zu einer Verletzung der Dünne-Elemente-Approximation führen.

Nicht zuletzt müssen bei der Wahl von δx auch Einschränkungen berücksichtigt werden, welche aus der für das korrespondierende optische Element gewählten Herstellungstechnologie resultieren.

Bei der Wahl von n_x entstehen Beschränkungen nach oben einerseits durch eine eventuell vorgegebene maximale laterale Ausdehnung des optischen Elements, welche durch

$$\Delta x = n_x \delta x \quad (3.15)$$

nicht überschritten werden sollte, sowie durch die Begrenzung der bei der numerischen Optimierung zur Verfügung stehenden Rechenkapazitäten. Eine weitere Beschränkung nach oben kann durch die Forderung nach Überlappungsfreiheit entsprechend Abschnitt 3.1.1 entstehen, wobei jedoch durch den Übergang zum Design eines nichtüberlappungsfreien Strahlteilers entsprechend Abschnitt 3.2 eine derartige Beschränkung in vielen Fällen aufgehoben werden kann.

Je größer n_x gewählt wird, desto stärker kann – wie in Abschnitt 3.4 gezeigt wird – die maximale Rauschintensität reduziert werden und desto feiner wird entsprechend der Gleichungen (3.6) und (3.15) die Rasterung in der Signalebene. In vielen Fällen bietet sich eine Optimierung von n_x innerhalb des in Frage kommenden Bereichs an, um eine Positionierung der Signalordnungen innerhalb der vorgegebenen Toleranzen zu erreichen. Falls dies nicht möglich ist, kann mit Hilfe des später in Abschnitt 3.2.2 vorgestellten Verfahrens auch eine rasterunabhängige Positionierung der Signalordnungen realisiert werden.

3.1.4 Iterationssteuerung beim Design überlappungsfreier Strahlteiler

In Abschnitt 3.1.1 wurde gezeigt, dass beim Design überlappungsfreier Strahlteiler Eingangsfeld und konkreter optischer Systemaufbau nicht berücksichtigt werden müssen. Bei Anwendung des IFTAs entsprechend Abschnitt 2.4.3 zur Lösung monofunktionaler Strahlteilungsprobleme genügt die Betrachtung der Abtastwerte aus einer Periode der Transmissionsfunktion, wobei FFT bzw. inverse FFT als Propagationsoperatoren eingesetzt werden können. Im Fall pixelierter Transmissionsfunktionen kann dabei die sinc-Modulation entsprechend Abschnitt 2.2.7 berücksichtigt werden.

Die im IFTA-Iterationszyklus nach Abbildung 2.4 vorkommenden Projektionsoperatoren werden dabei passend zu dem gegebenen Designproblem gewählt, wobei einige Beispiele in den Abschnitten 2.4.1 und 2.4.2 gegeben wurden. In der Praxis erweist es sich als vorteilhaft, die Projektionsoperatoren innerhalb der IFTA-Iterationen nicht sofort entsprechend der eigentlichen Einschränkungen bzw. Designfreiheiten des Designproblems zu wählen, sondern diese stattdessen Schritt für Schritt einzuführen [30].

Für das Design von Strahlteilern mit Phaseneinschränkung in der Transmissionsebene sowie Phasen-, Amplituden- und Skalierungsfreiheit in der Signalebene hat sich dabei das in Abbildung 3.1 gezeigte Ablaufschema bewährt. Die Startverteilung $t^{(0)}$ wird demnach aus der Rückpropagation eines Feldes ermittelt, welches ausschließlich Signalordnungen mit einer idealen Amplitudenverteilung und zufällig gewählten Phasen enthält. Danach erfolgen innerhalb der *Phasensynthese* zunächst einige Iterationen ohne Nutzung von Amplitudenfreiheit, bevor in der *SNR-Optimierung* die vollen Designfreiheiten zugelassen werden. Zur Erzielung eines optimalen Ergebnisses ist es notwendig, das Ablaufschema nach Abbildung 3.1 mehrfach für verschiedene Startverteilungen durchzuführen.

Für den Initialisierungsschritt in Abbildung 3.1 sind natürlich auch andere Varianten denkbar. In Referenz [33] wird beispielsweise ein analytischer Ansatz zur Ermittlung einer guten Startverteilung für das Design von Strahlteilern vorgestellt.

Die Motivation für den Schritt der Phasensynthese liegt darin, diejenige Phasenverteilung in der Signalebene zu finden, welche zu einer Maximierung der für jede komplexe Signalverteilung abhängig von den Transmissionseinschränkungen ableitbaren oberen Grenze der Konversionseffizienz [75–78] führt. Während der SNR-Optimierung ändern sich die Signalphasen typischerweise nur noch geringfügig.

Im Fall von Quantisierungseinschränkungen ergibt sich das Problem, dass bei Einsatz des

(1) <i>Initialisierung</i> : Setze $t^{(0)}$ durch Rückwärtspropagation eines Signalfelds mit zufällig gewählten Phasen der Signalordnungen
(2) <i>Phasensynthese</i> : Iterationen mit Π_{phase} und Π_{P} , d. h. ohne Amplitudenfreiheit
(3) <i>SNR-Optimierung</i> : Iterationen mit Π_{phase} and Π_{PAS} , d. h. unter voller Ausnutzung der vorhandenen Designfreiheiten

Abbildung 3.1: Schritte beim Design von Strahlteilern mit Phaseneinschränkung

(1) <i>Initialisierung</i> : Setze $t^{(0)}$ durch Rückwärtspropagation eines Signalfelds mit zufällig gewählten Phasen der Signalordnungen
(2) <i>Phasensynthese</i> : Iterationen mit Π_{phase} und Π_{P} , d. h. ohne Amplitudenfreiheit und Quantisierung
(3) <i>SNR-Optimierung</i> : Iterationen mit Π_{phase} and Π_{PAS} , d. h. ohne Quantisierung
(4) <i>Softquantisierung</i> : Iterationen mit $\Pi_{\text{softquant}, Q}^{\lambda}$ für linear ansteigende Projektionsstärke λ ($\lambda = 0 \dots 1$) sowie Π_{PAS}

Abbildung 3.2: Schritte beim Design quantisierter Strahlteiler unter Verwendung der Softquantisierung

Quantisierungsoperators $\Pi_{\text{quant}, Q}$ aus Abschnitt 2.4.1 typischerweise bereits nach wenigen Iterationen eine Stagnation eintritt [79, 80], d. h. die sich nach Anwendung des Quantisierungsoperators ergebende Transmissionsfunktion ändert sich im Verlauf der Iterationen nicht mehr. Daher empfiehlt es sich, zunächst wie in Abbildung 3.1 vorzugehen, d. h. für eine Phaseneinschränkung zu optimieren. Im Anschluss daran wird in einem zusätzlichen Schritt die Quantisierung durchgeführt, wobei sich bei Verwendung des *Softquantisierungsverfahrens* [79] das in Abbildung 3.2 gezeigte Ablaufschema ergibt.

Zur Softquantisierung auf Q äquidistante Phasenwerte wird der Operator $\Pi_{\text{softquant}, Q}^{\lambda}$,

$$(\Pi_{\text{softquant}, Q}^{\lambda} t)(x) = \begin{cases} z & \text{falls } |(\arg[t(x)] - \arg[z]) \bmod 2\pi| \leq \lambda\pi/Q \\ t(x) & \text{sonst} \end{cases} \quad (3.16a)$$

mit

$$z = (\Pi_{\text{quant}, Q} t)(x) \quad \text{und} \quad (3.16b)$$

$$a \bmod b = c \Leftrightarrow c \in [-b/2, b/2) \wedge c = a + jb \wedge j \in \mathbb{Z}$$

verwendet. Der Parameter λ wird während dieser Iterationsphase typischerweise linear von $\lambda = 0$ (keine Quantisierung) auf $\lambda = 1$ (vollständige Quantisierung) erhöht, wobei jedoch auch andere Varianten denkbar sind [38].

In Abschnitt 4.3 wird ein alternatives Quantisierungsverfahren eingeführt, welches gegenüber der Softquantisierung einige Vorteile aufweist.

3.2 Design für beliebig positionierte Signalordnungen

Innerhalb dieses Abschnittes, dessen Ergebnisse auch in Referenz [81] publiziert sind, werden Möglichkeiten zur Verbesserung der Positionierungsgenauigkeit von Signalordnungen untersucht.

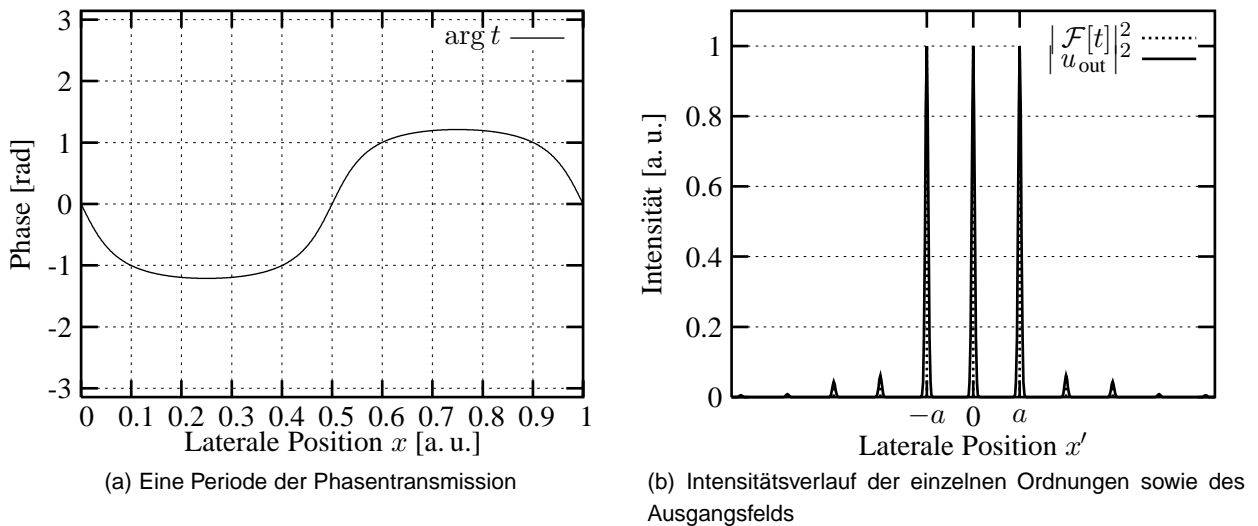


Abbildung 3.3: Optimale Lösung des regulären 1-zu-3-Strahlteilers mit Phaseneinschränkung

Bei einer Positionierung auf einem entsprechend Abschnitt 3.1 aus der Periodenlänge der Transmissionsfunktion resultierenden äquidistanten Raster kann bei Verfeinerung dieses Rasters die Eigenschaft der Überlappungsfreiheit verloren gehen. In Abschnitt 3.2.1 wird gezeigt, dass hieraus entstehende Probleme über eine Modifikation des Designalgorithmus umgangen werden können. In den Abschnitten 3.2.2 und 3.2.3 betrachten wir einen Designalgorithmus, welcher eine freie Positionierung der Signalordnungen erlaubt.

3.2.1 Verfeinerung der Signalebene rasterung

Zur Illustration der in diesem Abschnitt eingeführten Konzepte betrachten wir im Folgenden verschiedene Designs für 1-zu-3-Strahlteiler, wobei die eingeführten Konzepte natürlich auch auf andere Strahlteilungsprobleme anwendbar sind. Das Beispiel des 1-zu-3-Strahlteilers wurde gewählt, da für den regulären 1-zu-3-Strahlteiler die optimale Phasentransmissionslösung bekannt ist, für welche in Abbildung 3.3 eine Periode der Phasentransmission sowie die sich bei einem als gaußförmig angenommenen Feld u_{transf} entsprechend Gleichung (3.2) ergebende Ausgangsintensitätsverteilung dargestellt sind. Der Strahlteiler aus Abbildung 3.3 besitzt eine ideale Uniformität bei einer Effizienz von 92.6 %.

In Abbildung 3.4 betrachten wir ein Design für einen irregulären 1-zu-3-Strahlteiler, welches mit $\delta x' = 0.2 a$ eine im Vergleich zu Abbildung 3.3 um den Faktor 5 verfeinerte Signalebene rasterung besitzt. Die drei Signalordnungen befinden sich an den Positionen

$$\mathcal{W}_{\text{sig}} = \{-a; 0; 1.4 a\}. \quad (3.17)$$

Bei Ermittlung des Designs aus Abbildung 3.4 über das Verfahren aus Abschnitt 3.1.4 wurden für die drei Signalordnungen dieselben Zielintensitäten vorgegeben. Aufgrund der Definition des Signalfensters entsprechend Gleichung (3.17) erlaubt der während der SNR-Optimierung angewendete Operator Π_{PAS} aus Gleichung (2.63) auch Rauschen zwischen den Signalordnungen, was – wie in Abbildung 3.4(b) erkennbar ist – auch ausgenutzt wird.

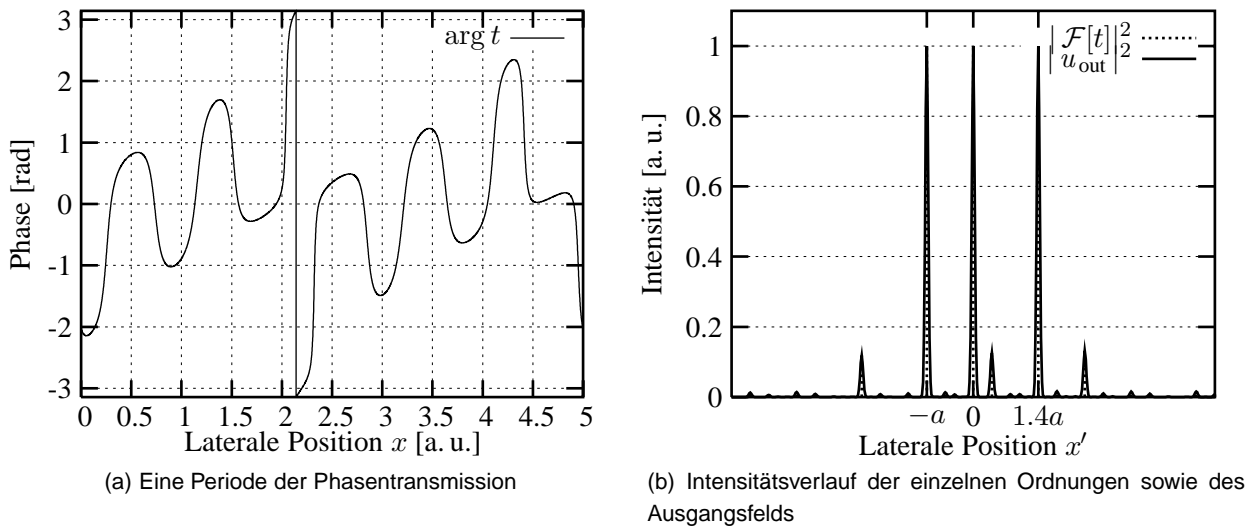


Abbildung 3.4: Irregulärer 1-zu-3-Strahlteiler mit Signalordnungen auf äquidistantem Raster mit $\delta x' = 0.2a$

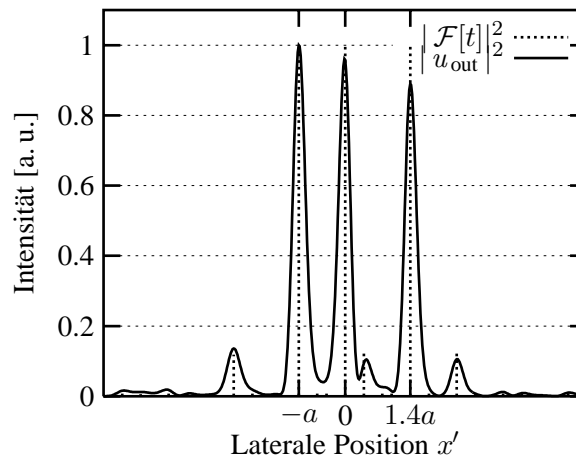


Abbildung 3.5: Aus Faltung mit einem im Vergleich zu Abbildung 3.4(b) um den Faktor 4 breiteren Feld u_{transf} resultierende Ausgangsintensitätsverteilung

Für das in Abbildung 3.4 gezeigte Design resultiert eine Effizienz von 82.6 % bei einem Uniformitätsfehler von weniger als 0.1 %. Die im Vergleich zum regulären 1-zu-3-Strahlteiler auftretende Effizienzreduktion ist eine Folge der veränderten Position der Signalordnungen. Wie aus Betrachtungen zur oberen Grenze der Konversionseffizienz [76] sowie aus Vergleichen tatsächlich erreichter Effizienzen für verschiedene Strahlteilungsprobleme [78] bekannt ist, hat insbesondere bei kleinen Teilerverhältnissen die Konfiguration der Signalordnungen erheblichen Einfluss auf die erzielbare Effizienz von Strahlteilern.

Während in Abbildung 3.4 von einem gaußförmigen Feld u_{transf} ausgegangen wird, für welches Überlappungsfreiheit resultiert, betrachten wir in Abbildung 3.5 die aus der Faltung mit einem um den Faktor 4 breiteren Gauß-Feld resultierende Ausgangsintensitätsverteilung. Die Verletzung von Bedingung (3.7) führt in diesem Fall zu einem deutlich sichtbaren Uniformitätsfehler von 5.7 %. Die Ursache hierfür stellen in diesem Beispiel im Wesentlichen Überlagerungen der Signalordnungen mit den benachbarten Rauschordnungen dar.

Der auftretende Uniformitätsfehler kann einerseits durch die Einbeziehung von $\mathcal{F}^{-1}u_{\text{transf}}$ als

Eingangsfeld während der IFTA-Iterationen vermieden werden. Damit würde jedoch die Insensitivität des Strahlteilers bezüglich der Wahl des Eingangsfelds verloren gehen, weshalb wir im Folgenden einen anderen Lösungsweg verwenden, welcher darin besteht, Rauschordnungen in der direkten Nachbarschaft von Signalordnungen zu unterdrücken.

Bei Unterdrückung der Rauschordnungen innerhalb der mit dem Durchmesser d_{frame} um sämtliche Signalordnungen aufgespannten kreisförmigen Bereiche ergibt sich der im Vergleich zu Gleichung (2.63) modifizierte Signalprojektionsoperator Π_{PAS} ,

$$(\Pi_{\text{PAS}} u_{\text{out}})(x') = \begin{cases} |\alpha u_{\text{sig}}(x')| \exp [i \arg (u_{\text{out}}(x'))] & \mathbf{falls} \ x' \in \mathcal{W}_{\text{sig}} \\ 0 & \mathbf{falls} \ \exists x'' \in \mathcal{W}_{\text{sig}} : |x' - x''| \leq \frac{d_{\text{frame}}}{2} \\ u_{\text{out}}(x') & \mathbf{sonst} \end{cases} \quad (3.18)$$

Falls die durch Verwendung dieser Signalprojektion bei der SNR-Optimierung angestrebte Unterdrückung der in Nachbarschaft zu Signalordnungen befindlichen Rauschordnungen gelingt und in Abwandlung von Bedingung (3.7) die Beziehung

$$\text{width}(u_{\text{transf}}) < d_{\text{frame}} / 2, \quad (3.19)$$

erfüllt ist, können Überlagerungen zwischen Signalordnungen und benachbarten Rauschordnungen vermieden werden. Demzufolge sollte der Parameter d_{frame} in der Praxis so gewählt werden, dass Bedingung (3.19) für alle relevanten Eingangsfelder erfüllt wird.

Leider beeinflusst die Modifikation des Signalprojektionsoperators in Gleichung (3.18) auch die beim Design erzielbare Effizienz. Eine Erhöhung von d_{frame} entspricht einer Verringerung der zur Verfügung stehenden Designfreiheiten, was zu einer Verringerung der erreichbaren Effizienz führt, wie durch das folgende Beispiel demonstriert wird. In Abbildung 3.6 ist ein mit Hilfe des Signalprojektionsoperators aus Gleichung (3.18) mit $d_{\text{frame}} = 0.8 a$ durchgeführtes Design dargestellt. Hierbei wurde ein Uniformitätsfehler von 0.1 % erreicht, wobei die erzielte Effizienz jedoch nur 71.5 % beträgt.

Während mit dem Projektionsoperator aus Gleichung (3.18) Überlagerungen zwischen Signalordnungen und benachbarten Rauschordnungen beim Design unterdrückt werden können, ist diese Methode hingegen nicht zur Unterdrückung von Überlagerungen zwischen unterschiedlichen Signalordnungen einsetzbar. Im Fall dicht beieinander liegender Signalordnungen und langsam abfallender Strahlprofile – wie beispielsweise sinc- oder Airy-förmige Profile – können derartige Interferenzeffekte jedoch ebenfalls einen nicht vernachlässigbaren Einfluss auf die Uniformität eines Strahlteilers besitzen. So kann beispielsweise ein Seitenausläufer einer Signalordnung, welcher an der Position der benachbarten Signalordnung bereits auf 0.1 % seiner Maximalintensität abgefallen ist, bei kohärenter Überlagerung zu einem Uniformitätsfehler von mehr als 6 % führen. In derartigen Fällen ist eine Vermeidung von Uniformitätsfehlern nur durch die bereits im Zusammenhang mit Abbildung 3.5 erwähnte Einbeziehung eines Eingangsfelds während der IFTA-Iteration möglich, wobei jedoch die Insensitivität bezüglich der Wahl des Eingangsfelds für den resultierenden Strahlteiler verloren geht.

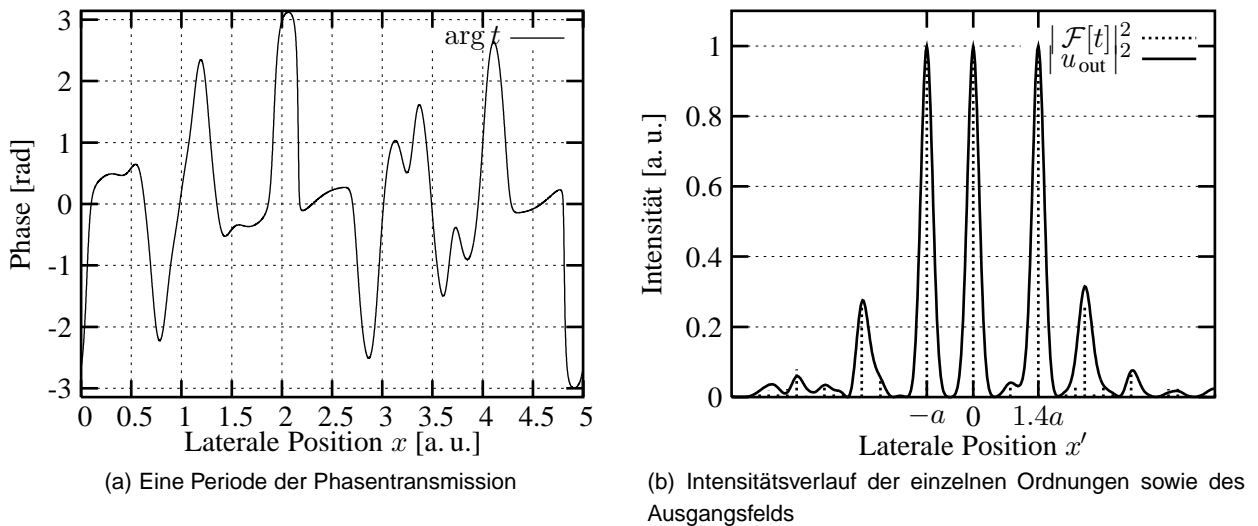


Abbildung 3.6: Irregulärer 1-zu-3-Strahlteiler mit Signalordnungen auf einem äquidistantem Raster mit $\delta x' = 0.2 a$. Während des Designs wurde das Rauschen innerhalb der um sämtliche Signalordnungen mit einem Durchmesser $d_{\text{frame}} = 0.8 a$ aufgespannten kreisförmigen Bereichen unterdrückt.

3.2.2 Algorithmus für beliebig positionierte Signalordnungen

Insbesondere bei industriellen Anwendungen von Strahlteilern liegt die gewünschte Verteilung der Signalordnungen nicht immer *a priori* auf einem äquidistanten Raster. Im Bereich der Materialbearbeitung werden beispielsweise Strahlteiler zur parallelen Durchführung mehrerer Bohrungen oder Schnitte eingesetzt, wobei in vielen Fällen eine freie Positionierbarkeit der Bohrungs- oder Schnittpositionen gewünscht wird.

Falls für die Positionierung der Signalordnungen Toleranzbereiche spezifiziert sind, kann eine Anwendung des Ansatzes aus Abschnitt 3.2 möglich sein, obwohl die ursprünglichen Zielpositionen nicht auf einem äquidistanten Raster liegen. So wären die Designs aus den Abbildungen 3.4 und 3.6 beispielsweise Lösungen eines Strahlteilungsproblems mit einer gewünschten Anordnung der Signalordnungen entsprechend

$$\mathcal{W}_{\text{sig}} = \{-a; 0; \sqrt{2}a\}, \quad (3.20)$$

sofern die Toleranz für die Positionierung der Signalordnungen wenigstens $\pm 1.5\%$ beträgt.

Dennoch ist dieser auf einem äquidistanten Raster basierende Ansatz in vielen Anwendungsfällen nicht zufriedenstellend, da einerseits mit einer Verfeinerung des Signalebenenrasters typischerweise die Zahl der während der IFTA-Iterationen zu betrachtenden Abtastwerte ansteigt und andererseits bei einigen Anwendungen das Ziel besteht, die Signalordnungen so genau wie möglich zu positionieren.

Der in diesem Abschnitt entwickelte Designalgorithmus ermöglicht eine beliebige Positionierung der Signalordnungen unter Aufgabe der Periodizität der Transmissionsfunktion innerhalb der beleuchteten Apertur. Diese Problemstellung wurde bereits von Bengtsson [17] untersucht, dessen Ansatz deutlich von dem im Folgenden vorgestellten Verfahren abweicht. Bei dem von Bengtsson verwendeten Optimierungsverfahren wird pro Iteration nur ein Abtastwert der Transmissionsfunktion verändert, was bei großen n_x – und insbesondere im Fall zweidimensionaler Abtastung – zu

Problemen führen kann, wenn der Effekt der Änderung eines Abtastwerts der Transmissionsfunktion nicht deutlich größer ist als die bei der Berechnung der Propagation vorliegende numerische Genauigkeit. In Projektionsalgorithmen können demgegenüber in jeder Iteration viele Abtastwerte gleichzeitig verändert werden, was zu einem deutlich besseren Konvergenzverhalten führen kann. Gegenüber dem Ansatz von Bengtsson besitzt das im Folgenden dargestellte Verfahren weiterhin den Vorteil, dass Störungen der Uniformität der Signalordnungen durch die umgebende Rauschverteilung über eine Adaption des Ansatzes aus Abschnitt 3.2.1 minimiert werden können.

Bei Zulassung einer rasterunabhängigen Positionierung der Signalordnungen ist es nicht mehr möglich, FFT-Operationen zur Ermittlung der komplexen Amplitude von Signalordnungen einzusetzen. Da die Fouriertransformationsintegrale entsprechend der Gleichungen (2.36) bzw. die hierzu korrespondierenden Summen jedoch für jede beliebige Signalebenenposition auswertbar sind, kann über die Berechnung einer sogenannten *punktweisen Fouriertransformation* für jede Signalordnung die komplexe Amplitude ermittelt werden. Die hierzu inverse Operation liefert in der Transmissionsebene für jede Signalordnung x' eine lineare Phasenfunktion, deren Steigung von x' abhängt. Die für die einzelnen Signalordnungen resultierenden linearen Phasenfunktionen werden bei Anwendung der inversen Propagationsoperation in der Transmissionsebene komplex aufaddiert.

Betrachten wir erneut den Designansatz aus Abbildung 3.1, so stellen wir fest, dass in dem Initialisierungsschritt sowie in der Phasensynthese somit lediglich die bisher verwendeten FFT-Operationen durch punktweise Fouriertransformationen ersetzt werden müssen. Bezüglich des notwendigen Rechenaufwands ist diese Modifikation des Algorithmus im Allgemeinen unkritisch: Während bei der FFT für n_x Abtastwerte ein zu $n_x \log n_x$ proportionaler Zeitaufwand vorliegt, resultiert bei der punktweisen FT für n_{sig} Signalordnungen ein Aufwand proportional zu $n_x n_{\text{sig}}$.

Sobald im Rahmen der SNR-Optimierung jedoch Amplitudenfreiheit beim Design zugelassen wird, ist die Verwendung punktweiser Fouriertransformationen nicht mehr ohne weiteres möglich, da dann während der Propagationsschritte zur Erfassung der Rauschverteilung wesentlich mehr Signalebenenpositionen betrachtet werden müssen, was zu einem für viele praktische Anwendungen inakzeptabel hohen Rechenaufwand führen würde. Als Ansatz für ein Verfahren, in welchem bei Zulassung von Amplitudenfreiheit wiederum FFT-Operationen eingesetzt werden können, zerlegen wir die Transmissionsfunktion $t(x)$ in zwei additive Anteile:

$$t(x) = t_{\text{sig}}(x) + t_{\text{noise}}(x). \quad (3.21)$$

Der Anteil t_{sig} resultiert dabei als Ergebnis der Phasensynthese bei einer Beendigung der IFTA-Iteration vor Anwendung des Transmissionsprojektionsoperators. Demnach führt t_{sig} zwar zur Erfüllung der Signaleinschränkungen, wobei jedoch die Transmissionseinschränkung im Allgemeinen verletzt wird.

Bei der Ermittlung des zweiten Anteils t_{noise} bestehen die folgenden beiden Ziele:

(A) Die Summierung entsprechend Gleichung (3.21) zu einer Erfüllung der Transmissionseinschränkung führen.

(B) Die gute Uniformität der von t_{sig} erzeugten Ausgangsverteilung darf hierbei nicht zerstört

werden.

Da eine additive Zerlegung in der Transmissionsebene entsprechend Gleichung (3.21) auf Grund der Linearität des Ausbreitungsoperators zu einer linearen Superposition der korrespondierenden Ausgangsfelder in der Signalebene führt, entspricht Ziel (B) der Anforderung, dass die zu t_{sig} und t_{noise} gehörigen Ausgangsfelder nicht überlappen dürfen. Diese Anforderung spiegelt sich in dem Signalprojektionsoperator

$$\Pi_{\text{noise}} u(x') = \begin{cases} 0 & \mathbf{falls} \ \exists x'' \in \mathcal{W}_{\text{sig}} : |x' - x''| \leq \frac{d_{\text{frame}}}{2} \\ u(x') & \mathbf{sonst} \end{cases} \quad (3.22)$$

wider, wobei d_{frame} wiederum den Durchmesser der um die Signalordnungen aufgespannten kreisförmigen Bereiche bezeichnet, innerhalb derer das Rauschen unterdrückt wird. Für den Operator Π_{noise} lässt sich ohne weiteres die punktweise Abstandsminimierung bezüglich der Einschränkungsmenge $\mathcal{M}_{\text{noise}}$ nachweisen, welche definiert ist über

$$\mathcal{M}_{\text{noise}} = \{u \in \mathcal{H} : \forall x' : [u(x') = 0 \vee \nexists x'' \in \mathcal{W}_{\text{sig}} : |x' - x''| \leq d_{\text{frame}}/2]\} \quad (3.23)$$

Um die Transmissionsprojektion anhand von Ziel (A) durchführen zu können, führen wir im Folgenden das Konzept der sogenannten *Summenprojektionen* ein. Ein Summenprojektionsoperator Π^{+t_d} wird basierend auf einem beliebigen Projektionsoperator Π über

$$\Pi^{+t_d} t = \Pi[t + t_d] - t_d \quad (3.24)$$

definiert, wobei t_d ein beliebiges Element des zu Grunde liegenden Hilbertraums \mathcal{H} bezeichnet.

Der folgende in Anhang A.2.2 bewiesene Satz 3.26 besagt, dass der über Gleichung (3.24) definierte Operator Π^{+t_d} eine Projektion auf die Menge

$$\mathcal{M}_2 = \{t_1 + t_d : t_1 \in \mathcal{M}_1\}. \quad (3.25)$$

darstellt. Demnach geht bei Verwendung der dem Ziel (A) entsprechenden Transmissionssummenprojektion $\Pi_{\text{phase}}^{+t_{\text{sig}}}$ die SDE-Konvergenzeigenschaft entsprechend Abschnitt 2.4.4 nicht verloren.

Satz 3.26 *Der auf Basis eines bezüglich der Norm d aus Gleichung (2.6) abstandsminimierenden Projektionsoperators $\Pi_1 : \mathcal{M} \rightarrow \mathcal{M}_1$ entsprechend Gleichung (3.24) konstruierte Operator $\Pi_2 = \Pi_1^{+t_d}$ definiert eine bezüglich d abstandsminimierende Projektion auf die Menge \mathcal{M}_2 aus Gleichung (3.25).*

Auf Basis der oben eingeführten Konzepte lässt sich der in Abbildung 3.7 dargestellte Algorithmus für das Design von Strahlteilern mit beliebiger Positionierung der Signalordnungen konstruieren. Hierbei sind die ersten beiden Schritte – wie schon erwähnt – äquivalent zum Ablaufschema aus Abbildung 3.1, wobei lediglich veränderte Propagationsoperatoren verwendet werden. Innerhalb des dritten Schrittes wird t_{sig} entsprechend der beim Design angestrebten Effizienz η_{goal} skaliert. Dies ist notwendig, weil es in der darauf folgenden SNR-Optimierung kein Äquivalent zu der in das Ablaufschema 3.1 durch Verwendung des Operators Π_{PAS} einbezogenen Skalierungsfreiheit

(1)	<i>Initialisierung</i> : Setze $t^{(0)}$ durch Rückwärtspropagation eines Signalfelds mit zufällig gewählten Phasen der Signalordnungen über inverse punktweise FT
(2)	FT-Operationen zur Optimierung von t_{sig}
(3)	Skalierung von t_{sig} entsprechend der vorgegebenen Zieleffizienz η_{goal}
(4)	<i>SNR-Optimierung</i> : Iterationen mit $\Pi_{\text{phase}}^{+t_{\text{sig}}}$ and Π_{noise} , unter Verwendung von FFT-Operationen zur Optimierung von t_{noise}
(5)	Falls notwendig: <i>Quantisierung</i>
(6)	Berechnung der resultierenden Transmission gemäß $t(x) = t_{\text{sig}}(x) + t_{\text{noise}}(x)$

Abbildung 3.7: Schritte beim Design von Strahlteilern mit beliebiger Positionierung der Signalordnungen

besteht. Innerhalb von SNR-Optimierung und Quantisierung werden dann Summenprojektion in der Transmissionsebene sowie der Operator Π_{noise} aus Gleichung (3.22) in der Signalebene verwendet. Der abschließende Summierungsschritt entspricht der Anwendung von Gleichung (3.21).

Eine mehrfache Durchführung des Ablaufschemas von Abbildung 3.7 ist einerseits wie schon in Abschnitt 3.1.4 notwendig, um die Auswirkung der Wahl der Startverteilung einzuschränken. Andererseits kann die mehrfache Durchführung auch erforderlich sein, um eine geeignete Wahl für den Parameter η_{goal} treffen zu können. Hierbei werden verschiedene Werte für η_{goal} mit dem Ziel getestet, eine höchstmögliche Effizienz bei einem noch akzeptierbaren Uniformitätsfehler zu erreichen. Falls η_{goal} dabei zu groß gewählt wurde, kann eine Überlappung der aus t_{sig} und t_{noise} resultierenden Ausgangsfelder während der SNR-Optimierung nicht beseitigt werden, was zu einer schlechten Uniformität führen kann. Bei einer zu niedrig getroffenen Wahl für η_{goal} kann typischerweise keine optimale Effizienz erzielt werden.

In den im folgenden Abschnitt betrachteten Designbeispielen wurde η_{goal} stets etwas höher als die tatsächlich erreichte Effizienz gewählt, was zu Designs mit guten Effizienzen führte, während die Uniformitätsfehler in einer Größenordnung bleiben, in der sie experimentell auch auf Grund von Fabrikationsungenauigkeiten auftreten.

3.2.3 Designbeispiele

Abbildung 3.8 zeigt ein Design für einen 1-zu-3-Strahlteiler mit Signalordnungspositionen entsprechend Beziehung (3.20), welches mit dem Verfahren aus dem vorangegangenen Abschnitt berechnet wurde. Die nichtperiodische Transmissionsfunktion wurde beim Design mit 5120 Werten abgetastet, wobei in Abbildung 3.8(a) nur die zentralen 512 Abtastwerte dargestellt werden.

Als Zieleffizienz wurde bei dem betrachteten Design $\eta_{\text{goal}} = 75\%$ vorgegeben, wobei letztendlich eine Effizienz von 71.3% bei einem Uniformitätsfehler von 0.6% erreicht wurde. Über ein Design mit $\eta_{\text{goal}} = 85\%$ wurde eine Effizienz von 75.9% erreicht, wobei sich jedoch der Uniformitätsfehler auf 1.2% verschlechterte. Bei einem niedriger gewählten Wert von $\eta_{\text{goal}} = 71\%$ wurde demgegenüber eine Effizienz von 69.8% bei einem im Vergleich zum Design aus Abbildung 3.8 verringerten Uniformitätsfehler von 0.3% erreicht. Diese Ergebnisse bestätigen die Bemerkungen zur Wahl von η_{goal} , welche am Ende des vorangegangenen Abschnitts getroffen wurden.

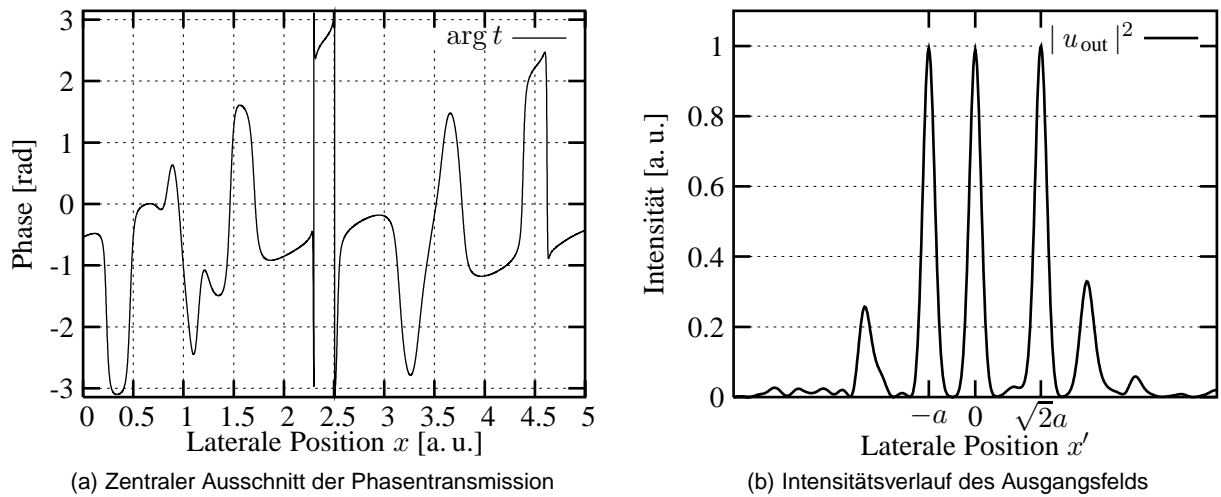
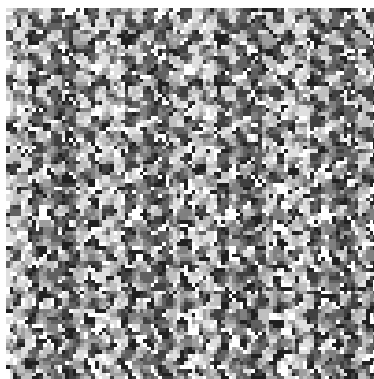
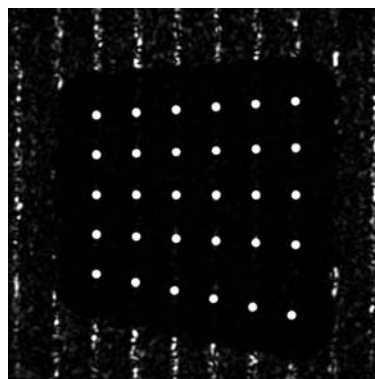


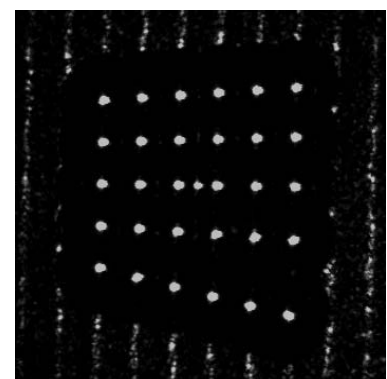
Abbildung 3.8: Irregulärer 1-zu-3-Strahlteiler mit Signalordnungen im Abstandsverhältnis $1:\sqrt{2}$



(a) Phase der zentralen 100×100 Abtastwerte der mit 1000×1000 Abtastwerten berechneten Transmissionsfunktion



(b) Ausgangsintensitätsverteilung mit Normierung auf 1% des Maximalwerts



(c) CCD-Kameraaufnahme der experimentell resultierenden Ausgangsintensitätsverteilung

Abbildung 3.9: Irregulärer 1-zu-30-Strahlteiler mit auf geneigten Linien positionierten Signalordnungen

Als weiteres Beispiel zur Anwendung des Designverfahrens aus Abschnitt 3.2.2 betrachten wir in Abbildung 3.9 ein Design für einen speziellen 1-zu-30-Strahlteiler, bei dem die Signalordnungen auf fünf Linien mit Neigungswinkeln von 4° , 2° , 0° , -3° , und -12° angeordnet sind. Bei dem gezeigten Design, welches für 1000×1000 Abtastwerte mit acht Phasenquantisierungswerten durchgeführt wurde, resultiert in der Simulation eine Effizienz von 59.9% bei einem Uniformitätsfehler von 5.4%. Die relativ niedrige Effizienz dieses Designs ist eine Folge der auftretenden Pixelierungsverluste (vgl. Abschnitt 3.1.2). In Abbildung 3.9(b), in der auf Grund der gewählten Normierung Teile der Rauschverteilung sichtbar sind, zeigt sich, dass die Rauschverteilung in diesem Beispiel durch eine Wahl von d_{frame} in der Größenordnung des Abstands zwischen den Signalordnungen gut von den Signalordnungen separiert werden konnte.

Der berechnete Strahlteiler wurde mit Hilfe der Dünne-Elemente-Approximation in ein Höhenprofil überführt, welches mit einer Pixelgröße von $5 \mu\text{m}$ in Quarz hergestellt wurde. Abbildung 3.9(c) zeigt die bei einer Beleuchtung mit $\lambda_{\text{ref}} = 632.8 \text{ nm}$ im Fernfeld aufgenommene Intensitätsverteilung. Abgesehen von dem Auftreten eines zusätzlichen Spots auf der Position der optischen Achse, welcher eine Folge der bei der Herstellung dieses Strahlteilers entstandenen rela-

tiven Höhenfehler von etwa 5 % ist, besteht visuell eine gute Übereinstimmung mit der simulierten Intensitätsverteilung aus Abbildung 3.9(b). Durch Vermessung des Leistungsanteils in den einzelnen Signalordnungen mit einer geeigneten Blende wurde ein Uniformitätsfehler von 10.4 % ermittelt. Für die Abweichungen zwischen simulierten und experimentellen Ergebnisse kommen die in Abschnitt 6.2.2 diskutierten Ursachen in Betracht.

3.3 Toleranzeigenschaften von Strahlteilern

Bei der Betrachtung der Toleranzeigenschaften von Strahlteilern legen wir den Schwerpunkt auf den in den Abschnitten 3.3.1 und 3.3.2 untersuchten Einfluss globaler Höhenfehler bei binären bzw. mehrstufigen Strahlteilern. In dem darauf folgenden Abschnitt 3.3.3 wird die Auswirkung von Verschiebungen und Verkippen des Strahlteilerelements dargestellt. In Abschnitt 3.3.4 erfolgt schließlich eine Diskussion der Toleranzeigenschaften bezüglich Veränderungen der Wellenlänge sowie der Feldverteilung des Eingangsstrahls.

3.3.1 Höhenfehler bei binären Strahlteilern

In diesem und dem folgenden Abschnitt wird die Auswirkung *globaler Höhenfehler* auf die optische Wirkung eines Strahlteilerelements untersucht. Wir bezeichnen einen Höhenfehler als global, sofern die Transformation des gewollten Profils $d(x)$ in das tatsächlich realisierte Profil $d'(x)$ ortsunabhängig ist. Andere Arten von Höhenfehlern, wie beispielsweise lokale Fehler, werden im Rahmen dieser Arbeit nicht untersucht. Übersetzt in die entsprechend der Dünne-Elemente-Approximation resultierenden Transmissionsfunktionen stellen globale Höhenfehler einen Spezialfall der in Gleichung (2.49) eingeführten Operatoren E_β zur Beschreibung von Umsetzungsfehlern dar.

Einen wichtigen Spezialfall globaler Höhenfehler stellen *lineare Höhenfehler* dar, welche durch

$$d'(x) = \gamma_{\text{lin}} d(x) \quad (3.27)$$

mit dem Skalierungsfaktor $\gamma_{\text{lin}} > 0$ beschrieben werden. Lineare Höhenfehler sind geeignet zur Modellierung verschiedener Arten von Herstellungsfehlern. Beispiele hierfür sind falsche Belichtungsdosen oder Entwicklungsdauern bei der Resiststrukturierung, falsche Ätzdauern oder Ätzraten bei der Strukturierung von Quarz oder Höhenfehler durch Schrumpfung des Elements, wie sie während des Abkühlungsprozesses bei der Replikation über Plastikabformung auftreten können. Eine weitere Ursache für lineare Höhenfehler stellen durch Temperatureinflüsse verursachte Ausdehnungseffekte dar. Änderungen der Brechzahl des Strahlteilerelements sind entsprechend Gleichung (2.41) ebenfalls äquivalent mit linearen Höhenfehlern.

In diesem Abschnitt betrachten wir Strahlteiler mit binären Transmissionen

$$t(x) : \mathbb{R} \rightarrow \{1, \exp(i \gamma_{\text{bin}} \pi)\}. \quad (3.28)$$

Weiterhin gehen wir von der Gültigkeit der Dünne-Elemente-Approximation aus, was entsprechend Gleichung (2.41) zu einer Proportionalität zwischen dem Parameter γ_{bin} und der Modulati-

onstiefe des binären Höhenprofils führt. Die Wirkung linearer Höhenfehler kann somit korrespondierend zu Gleichung (3.27) auch über

$$\gamma_{\text{bin}} = \gamma_{\text{lin}} \gamma_{\text{goal}} \quad (3.29)$$

beschrieben werden, wobei γ_{goal} der angestrebten Modulationstiefe entspricht.

Bezeichnen wir mit I_j die Intensitäten der einzelnen Ordnungen von $t(x)$ und mit \tilde{I}_j die Intensitäten der einzelnen Ordnungen von $\tilde{t}(x)$ gemäß

$$\tilde{t}(x) = t(x)^{1/\gamma_{\text{bin}}} \quad \text{mit} \quad \tilde{t}(x) : \mathbb{R} \rightarrow \{1, -1\}, \quad (3.30)$$

so ergibt sich bei Verwendung der Normierung

$$\sum_j I_j = \sum_j \tilde{I}_j = 1 \quad (3.31)$$

der Zusammenhang [82]

$$I_j(\gamma_{\text{bin}}) = \frac{1 - \cos \gamma_{\text{bin}} \pi}{2} \tilde{I}_j \quad (j \neq 0), \quad (3.32a)$$

$$I_0(\gamma_{\text{bin}}) = 1 - \sum_{j \neq 0} I_j(\gamma_{\text{bin}}) = 1 - \frac{1 - \cos \gamma_{\text{bin}} \pi}{2} \sum_{j \neq 0} \tilde{I}_j, \quad (3.32b)$$

Hierbei bezeichnet \sum_j jeweils die Summe über alle Ordnungen, und $\sum_{j \neq 0}$ steht für eine Summierung, bei der die nullte Ordnung ausgenommen wird.

Anhand der Gleichungen (3.32) können wir folgende Schlussfolgerungen hinsichtlich der Sensitivität binärer Strahlteiler bezüglich linearer Höhenfehler ziehen:

(A) Die zentrale Ordnung $j = 0$ verhält sich grundsätzlich verschieden von allen anderen Ordnungen. Falls sie nicht zu den Signalordnungen gehört, resultiert nach Gleichung (3.32a) bei Veränderung von γ_{bin} durch einen Höhenfehler $\gamma_{\text{lin}} \neq 1$ lediglich eine Skalierung der Effizienz, wohingegen der Uniformitätsfehler konstant bleibt. Falls keine Forderungen hinsichtlich der Rauschverteilung gestellt werden, besitzen somit binäre Strahlteiler mit $0 \notin \mathcal{W}_{\text{sig}}$ im Vergleich zu binären Strahlteilern mit $0 \in \mathcal{W}_{\text{sig}}$ deutlich verbesserte Toleranzeigenschaften bezüglich von Höhenfehlern.

(B) Betrachten wir die Änderungen der Intensitäten I_j bei Aufbringung eines durch γ_{lin} beschriebenen Höhenfehlers, d. h.

$$\Delta I_j(\gamma_{\text{lin}}) = I_j(\gamma_{\text{bin}}) - I_j(\gamma_{\text{goal}}), \quad (3.33)$$

so stellen wir fest, dass für $j \neq 0$ die ΔI_j wegen Gleichung (3.32a) jeweils dasselbe Vorzeichen besitzen. Da auf Grund der Energieerhaltung

$$\sum_j \Delta I_j = 0 \quad (3.34)$$

gilt, ergibt sich ΔI_0 über

$$\Delta I_0 = \sum_{j \neq 0} \Delta I_j. \quad (3.35)$$

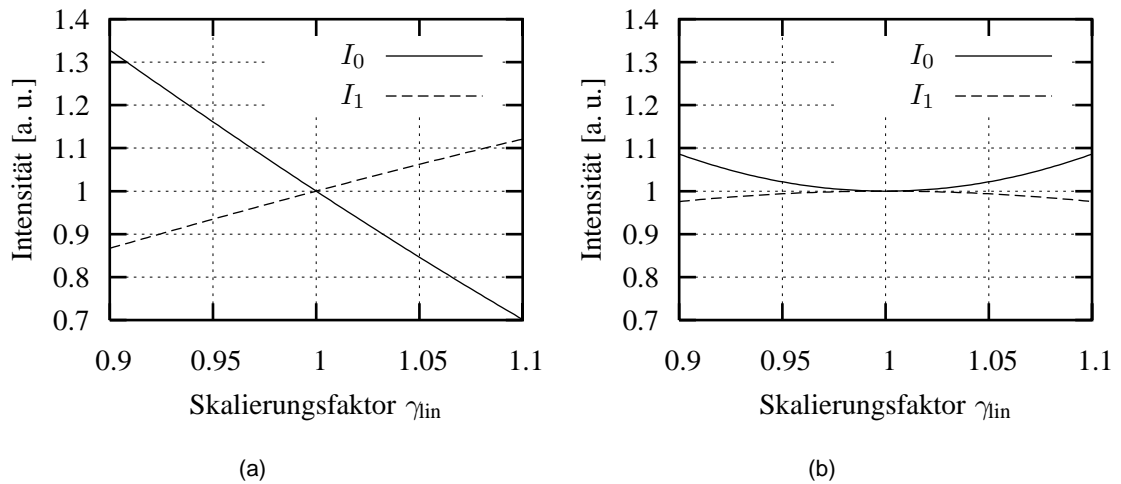


Abbildung 3.10: Abhängigkeit der Intensität der 0. sowie ± 1 . Ordnung vom Höhengskalierungsfehler für das effizienzoptimierte (a) und toleranzoptimierte Design (b) des binären 1-zu-3-Strahlteilers

Somit ist ΔI_0 betragsmäßig größer als die Änderung der Intensitäten aller anderen Ordnungen. Bei binären Strahlteilern mit $0 \in \mathcal{W}_{\text{sig}}$ ist daher die Zunahme des Uniformitätsfehlers bei Höhenfehlern im Wesentlichen durch ΔI_0 bestimmt.

(C) Wegen $I_0 \geq 0$ resultiert aus Gleichung (3.32b) eine Minimierung von I_0 für $\gamma_{\text{bin}} = 1$. Da für binäre Strahlteiler mit $0 \in \mathcal{W}_{\text{sig}}$ der Uniformitätsfehler entsprechend Punkt (B) im Wesentlichen aus der Änderung von I_0 resultiert, besitzen somit Lösungen mit $\gamma_{\text{goal}} = 1$ ein im Vergleich zu Lösungen mit $\gamma_{\text{goal}} \neq 1$ verbessertes Toleranzverhalten.

(D) Aufgrund von Gleichung (3.35) und da die ΔI_j wegen (B) für $j \neq 0$ dieselben Vorzeichen besitzen, resultiert anschaulich eine Proportionalität zwischen ΔI_0 und der Anzahl der Signalordnungen. Für Designs mit einer geringen Anzahl n_{sig} von Signalordnungen kann demnach im Allgemeinen eine geringere Empfindlichkeit des Uniformitätsfehlers bezüglich von Höhenfehlern erreicht werden als für Designs mit großem n_{sig} .

Die Auswirkung von Schlussfolgerung (C) wird in den Abbildungen 3.10 und 3.11 am Beispiel des binären 1-zu-3-Strahlteilers demonstriert. Wir betrachten dabei zwei verschiedene Lösungen für dieses Designproblem, welche über eine entsprechend Gleichung (3.3) periodische Transmissionsfunktion mit

$$t_p(x) = \begin{cases} \exp(i \gamma_{\text{bin}} \pi) & \mathbf{falls} \ x \in [0, a \Delta x) \\ 1 & \mathbf{falls} \ x \in [a \Delta x, 1) \\ 0 & \mathbf{sonst} \end{cases} \quad (3.36)$$

beschrieben werden. Dies ist einerseits die effizienzoptimierte Lösung dieses Designproblems mit $\gamma_{\text{bin}} = 0.6391$ und $a = 0.5$, für welche sich eine Effizienz von 86.5 % ergibt. Zum Vergleich hierzu betrachten wir die toleranzoptimierte Lösung mit $\gamma_{\text{bin}} = 1$ und $a = 0.2647$, für welche eine Effizienz von 66.4 % resultiert.

In Abbildung 3.10 wird der Verlauf der Intensitäten I_0 und I_1 für beide Designs in Abhängigkeit vom Skalierungsfaktor γ_{lin} dargestellt. In Abbildung 3.10(b) zeigt sich dabei, dass $I_0(\gamma_{\text{lin}})$ entsprechend Schlussfolgerung (C) für $\gamma_{\text{lin}} = 1$ ein Minimum besitzt. Das daraus resultierende verbesserte

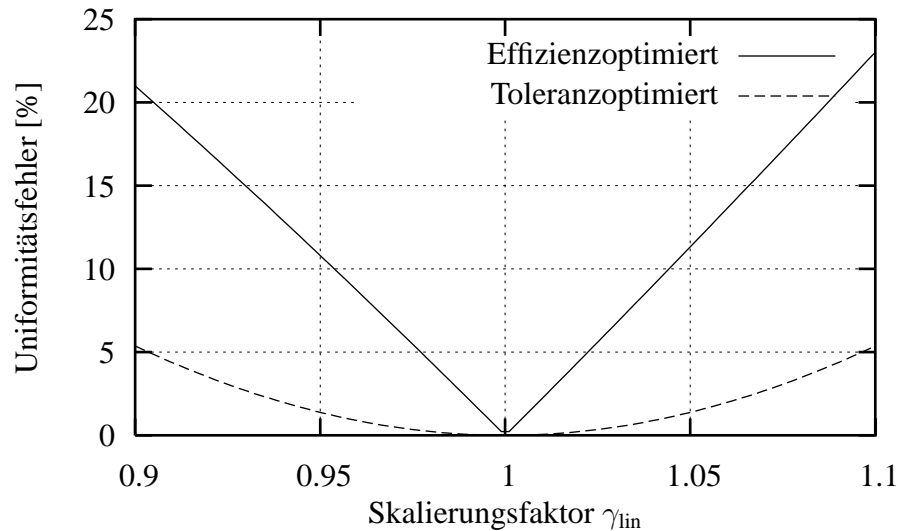


Abbildung 3.11: Uniformitätsfehler in Abhängigkeit vom Höhenskalierungsfehler für die beiden betrachteten Lösungen des binären 1-zu-3-Strahlteilers

Toleranzverhalten wird im Vergleich zu Abbildung 3.10(a) sowie aus der Darstellung der Abhängigkeit des Uniformitätsfehlers vom Skalierungsfaktor γ_{lin} für beide Designs in Abbildung 3.11 sichtbar.

In dem betrachteten Beispiel weist das toleranzoptimierte Design eine im Vergleich zum effizienzoptimierten Design deutlich verringerte Effizienz auf. Diese Beobachtung lässt sich teilweise auch für binäre Strahlteiler mit anderen Konfigurationen der Signalordnungen verallgemeinern. Wie in Referenz [78] gezeigt wird, bestehen abhängig von n_{sig} zum Teil deutliche Unterschiede zwischen Effizienzen, welche für Designs mit $\gamma_{bin} = 1$ (toleranzoptimiert) sowie für Designs mit frei wählbaren γ_{bin} resultieren. In dem folgenden Abschnitt 3.3.2 sowie in Abschnitt 6.2 wird jedoch gezeigt, dass im Fall von $Q > 2$ in vielen praktischen Fällen keine bzw. deutlich geringere Effizienzverluste zur Erlangung eines Designs mit guten Toleranzeigenschaften in Kauf genommen werden müssen.

Um die im Zusammenhang mit den Gleichungen (3.32) getroffene Schlussfolgerung (D) zu veranschaulichen, betrachten wir im Folgenden beliebige binäre Strahlteiler mit $0 \in \mathcal{W}_{sig}$, deren n_{sig} Signalordnungen bei $\gamma_{bin} = 1$ exakt die gleiche Intensität besitzen. Bei einer angenommenen Effizienz η bedeutet dies auf Grund der Normierungsbedingung (3.31)

$$\tilde{I}_0 = \frac{\eta}{n_{sig}}. \quad (3.37)$$

Für I_0 lässt sich hierbei die Beziehung

$$I_0 = 1 - \frac{1 - \cos \gamma_{lin} \pi}{2} \left(1 - \frac{\eta}{n_{sig}} \right) \quad (3.38)$$

ableiten und für eine beliebige Signalordnung $j \neq 0$ erhalten wir

$$I_j = \frac{1 - \cos \gamma_{lin} \pi}{2} \frac{\eta}{n_{sig}}. \quad (3.39)$$

Der im Fall von $\gamma_{lin} \neq 1$ resultierende Uniformitätsfehler ergibt sich entsprechend Gleichung (2.56)

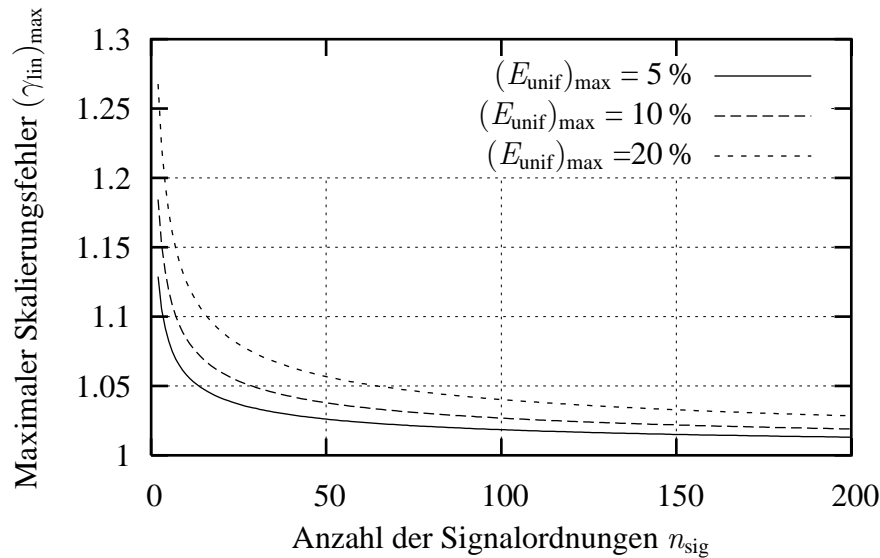


Abbildung 3.12: Abhängigkeit des zulässigen Höhenfehlers $(\gamma_{\text{lin}})_{\text{max}}$ von der Anzahl der Signalordnungen n_{sig} für unterschiedliche Werte des maximal tolerierbaren Uniformitätsfehlers $(E_{\text{unif}})_{\text{max}}$

als

$$E_{\text{unif}} = \frac{I_0 - I_j}{I_0 + I_j}. \quad (3.40)$$

Durch Einsetzen der Gleichungen (3.38) und (3.39) in Gleichung (3.40) kann bei Vorgabe eines maximal tolerierbaren Uniformitätsfehlers $(E_{\text{unif}})_{\text{max}}$ numerisch der hierfür erlaubte maximale Höhenfehler $(\gamma_{\text{lin}})_{\text{max}}$ bestimmt werden. Das Diagramm 3.12 zeigt die so ermittelte Abhängigkeit zwischen $(\gamma_{\text{lin}})_{\text{max}}$ und n_{sig} für verschiedene Werte von $(E_{\text{unif}})_{\text{max}}$, wobei für $\eta = 0.8$ angenommen wurde. Es ist erkennbar, dass für größere Werte von n_{sig} die Profilhöhe typischerweise bis auf wenige Prozent genau realisiert werden muss, wobei die Sensitivität des Uniformitätsfehlers gegen Höhenskalierungsfehler mit wachsendem n_{sig} zunimmt.

3.3.2 Höhenfehler bei mehrstufigen Strahlteilern

Nachdem in dem vorangegangenen Abschnitt die Auswirkung von Höhenfehlern bei binären Strahlteilern diskutiert wurde, soll nun die Betrachtung auf mehrstufige Strahlteiler erweitert werden. Dabei besteht das Ziel, eine Abschätzung zu treffen, inwieweit bei Zulassung von $Q > 2$ Quantisierungsstufen das Toleranzverhalten gegenüber dem binären Fall verbessert werden kann. Innerhalb dieses Abschnitts beschränken wir uns auf die Betrachtung linearer Höhenfehler.

Das Verhalten der Ordnungen $j \neq 1$ kann im Fall $Q > 2$ nicht mehr über eine einfache Beziehung wie Gleichung (3.32a) beschrieben werden. Insbesondere kann im Gegensatz zu binären Strahlteilern mit $\gamma_{\text{goal}} = 1$ die Intensität einzelner Ordnungen $j \neq 0$ auch für $\gamma_{\text{lin}} \neq 1$ zunehmen. Dennoch ist die Intensitätsänderung von Ordnungen $j \neq 0$ typischerweise auch im Fall $Q > 2$ deutlich kleiner als ΔI_0 , sodass über die alleinige Betrachtung des Verhaltens $\Delta I_0(\gamma_{\text{lin}})$,

$$\Delta I_0(\gamma_{\text{lin}}) = I_0(\gamma_{\text{lin}}) - I_0(1), \quad (3.41)$$

eine gute Abschätzung für das Verhalten des Uniformitätsfehlers für Strahlteiler mit $0 \in \mathcal{W}_{\text{sig}}$ erlangt werden kann.

Durch Anwendung eines linearen Skalierungsfaktors γ_{lin} auf die Q Phasenwerte entsprechend Gleichung (2.61a) erhalten wir die modifizierten Quantisierungswerte p_j ($j = 0, \dots, Q-1$) mit

$$p_j(\gamma_{\text{lin}}) = \exp(i \gamma_{\text{lin}} 2\pi j/Q). \quad (3.42)$$

Mit n_j bezeichnen wir im Folgenden die Anzahl der Abtastwerte der Transmissionsfunktion, welche den Quantisierungswert p_j besitzen ($0 \leq j < Q$). Die Diskretisierung des Integrals (2.36a) für $x' = 0$ führt für die Berechnung von I_0 eine Summierung über alle Abtastwerte der Transmissionsfunktion, d. h. bei entsprechender Normierung ergibt sich

$$I_0(\gamma_{\text{lin}}) = \left| \sum_{j=0}^{Q-1} n_j p_j(\gamma_{\text{lin}}) \right|^2. \quad (3.43)$$

Zur Vereinfachung der Darstellung beschränken wir uns im Folgenden auf den Fall $Q = 2^q$ mit $q \in \mathbb{N}$. In diesem Fall gibt es $Q/2$ Paare von Quantisierungswerten $(p_j, p_{j+Q/2})$, die sich für $\gamma_{\text{lin}} = 1$ jeweils zu Null addieren. Die Summe in Gleichung (3.43) kann somit wie folgt in zwei Terme aufgespalten werden:

$$I_0(\gamma_{\text{lin}}) = |\tilde{u}_0 + \hat{u}_0|^2 \quad \text{mit} \quad (3.44a)$$

$$\tilde{u}_0 = \sum_{j=0}^{Q-1} \tilde{n}_j p_j(\gamma_{\text{lin}}), \quad (3.44b)$$

$$\hat{u}_0 = (1 + \exp[i \gamma_{\text{lin}} \pi]) \sum_{j=0}^{Q/2-1} \hat{n}_j p_j(\gamma_{\text{lin}}), \quad \text{wobei} \quad (3.44c)$$

$$\hat{n}_j = \min(n_j, n_{j+Q/2}) \quad \text{und} \quad (3.44d)$$

$$\tilde{n}_j = \begin{cases} n_j - \hat{n}_j & \text{falls } n_j - \hat{n}_j > 0 \\ 0 & \text{sonst} \end{cases}. \quad (3.44e)$$

Der Anteil \hat{u}_0 verschwindet für $\gamma_{\text{lin}} = 1$. Für $\gamma_{\text{lin}} \approx 1$ ändert sich die Länge des durch \hat{u}_0 in der komplexen Zahlenebene beschriebenen Vektors näherungsweise proportional zu

$$|1 + \exp(i \gamma_{\text{lin}} \pi)| = \sqrt{2(1 + \cos \gamma_{\text{lin}} \pi)}. \quad (3.45)$$

Der Einfluss dieser Längenänderung auf I_0 entsprechend Gleichung (3.44a) ist anschaulich natürlich dann am größten, wenn die Vektoren \tilde{u}_0 und \hat{u}_0 zueinander parallel sind, während der Einfluss bei aufeinander senkrecht stehenden Vektoren minimiert wird. Die Richtung des Vektors \hat{u}_0 resultiert entsprechend Gleichung (3.44c) aus zwei additiven Anteilen:

$$\arg \hat{u}_0 = \arg(1 + \exp[i \gamma_{\text{lin}} \pi]) + \arg \sum_{j=0}^{Q/2-1} \hat{n}_j p_j(\gamma_{\text{lin}}). \quad (3.46)$$

Der erste Summand auf der rechten Seite von Gleichung (3.46) ergibt sich bei $\gamma_{\text{lin}} \approx 1$ je nach Vorzeichen von $(\gamma_{\text{lin}} - 1)$ näherungsweise zu $\pi/2$ oder $3/2\pi$. Demgegenüber kann der zweite Summand sowie die Phase $\arg \tilde{u}_0$ durch Multiplikation eines konstanten Phasenterms $\exp(i\phi)$ zu der

betrachteten Transmissionsfunktion beliebig um jeweils ϕ verschoben werden. Durch eine geeignete Wahl von ϕ kann daher im Allgemeinen erreicht werden, dass die Vektoren \tilde{u}_0 und \hat{u}_0 zueinander senkrecht werden, was – wie oben erwähnt – zu einer Verbesserung der Toleranzeigenschaften von Strahlteilern mit $0 \in \mathcal{W}_{\text{sig}}$ führt. Im Gegensatz zu dem in Abschnitt 3.3.1 diskutierten Fall binärer Strahlteiler müssen hierbei keine Effizienzeinbußen im Vergleich zu einem effizienzoptimierten Design in Kauf genommen werden. In Abschnitt 6.2 wird diese Form der Verbesserung der Toleranzeigenschaften anhand eines Beispiels demonstriert.

Im Fall von binären Transmissionen mit $\gamma_{\text{goal}} = 1$ stehen für $\gamma_{\text{lin}} \approx 1$ die Vektoren \tilde{u}_0 und \hat{u}_0 stets zueinander näherungsweise senkrecht, was eine weitere Erklärung für die in Abschnitt 3.3.1 für diesen Fall festgestellten guten Toleranzeigenschaften darstellt. Wir haben somit im vorangegangenen Absatz gezeigt, dass durch Multiplikation eines geeigneten konstanten Phasenwerts für eine mehrstufige Transmissionsfunktion ein ähnlich gutes Toleranzverhalten wie bei binären Strahlteilern mit $\gamma_{\text{goal}} = 1$ erreicht werden kann. Im verbleibenden Teil dieses Abschnittes wird dargestellt, inwieweit die Toleranzeigenschaften eines Strahlteilers mit $0 \in \mathcal{W}_{\text{sig}}$ bei einem mehrstufigen Design gegenüber denen eines binären Designs verbessert werden können.

Wir betrachten hierzu die Quotienten $(\Delta I_0)_{q=1}/(\Delta I_0)_{q=q'}$, welche gemäß oben stehender Diskussion für Designs mit $0 \in \mathcal{W}_{\text{sig}}$ näherungsweise die relative Verringerung der Sensitivität des Uniformitätsfehlers bezüglich Höhenfehlern γ_{lin} für $Q = 2^{q'}$ gegenüber dem binären Fall aus Abschnitt 3.3.1 beschreiben. Wie Anhang A.2.1 gezeigt wird, kann zur Berechnung dieser Quotienten die approximativ gültige Beziehung

$$\frac{(\Delta I_0)_{q=q'-1}}{(\Delta I_0)_{q=q'}} = \frac{2}{1 + \cos(2^{-q'+1} \gamma_{\text{lin}} \pi)} \quad (3.47)$$

abgeleitet werden.

Tabelle 3.1 zeigt in der dritten Spalte den Wert dieses Quotienten für $q' = 2 \dots 4$. Die vierte Spalte zeigt die hieraus rekursiv berechenbaren Quotienten $(\Delta I_0)_{q=1}/(\Delta I_0)_{q=q'}$, welche – wie schon erwähnt – anschaulich die Verbesserung der Toleranzeigenschaften gegenüber dem binären Fall beschreiben. Es zeigt sich, dass bei Übergang von $Q = 2$ zu $Q = 4$ eine Verbesserung um den Faktor 2 möglich ist, d. h. die Kurven für $(E_{\text{unif}})_{\text{max}} = 10\%$ und 20% in Abbildung 3.12 entsprechen näherungsweise 5% bzw. 10% für $Q = 4$. Eine weitere Erhöhung von Q hat jedoch gemäß der letzten Spalte von Tabelle 3.1 nur noch eine geringe Auswirkung auf die für Strahlteiler mit $0 \in \mathcal{W}_{\text{sig}}$ erzielbaren Toleranzeigenschaften bezüglich linearer Höhenfehler.

3.3.3 Einfluss von Verschiebungen und Verkippungen

Bei der Betrachtung der Auswirkung von Verschiebungen eines Strahlteilerelements muss zwischen lateralen und longitudinalen Verschiebungen unterschieden werden. Laterale Verschiebungen der Transmissionsfunktion $t(x)$ bewirken in $\mathcal{F}t$ das Auftreten eines multiplikativen linearen Phasenterms und sind daher auf Grund der bei Strahlteilerdesigns typischerweise bestehenden Phasenfreiheit in den meisten Fällen irrelevant. Erst wenn die Apertur des Strahlteilerelements den Eingangsstrahl beschneidet, resultiert ein Einfluss auf die Intensitätsverteilung in der Ausgangsebene.

q'	$Q = 2^{q'}$	$\frac{(\Delta I_0)_{q=q'-1}}{(\Delta I_0)_{q=q'}}$	$\frac{(\Delta I_0)_{q=1}}{(\Delta I_0)_{q=q'}}$
2	4	2.000	2.000
3	8	1.172	2.243
4	16	1.040	2.436

Tabelle 3.1: Verbesserung der Toleranzeigenschaften gegenüber Höhenfehlern durch Verwendung von $Q > 2$ Quantisierungsstufen. Die dargestellten Werte ergeben sich gerundet aus Anwendung von Gleichung (3.47).

Die Auswirkung longitudinaler Verschiebungen hängt hingegen von dem verwendeten Systemaufbau ab. In dem Aufbau 2.2(a) wirkt sich die Wahl des Abstand d gemäß Gleichung (2.45) wiederum nur auf die resultierende Phasenverteilung aus. Im Gegensatz hierzu tritt für den Aufbau 2.2(b) entsprechend Gleichung (2.46) zusätzlich eine lineare Skalierung der Signalebenenkoordinaten auf, d. h. der Abstand zwischen zwei betrachteten Ordnungen ist proportional zu d . Aufgrund der im Zusammenhang mit Abbildung 2.2(a) getroffenen Diskussion, hat eine longitudinale Verschiebung des Strahlteilerelements im Aufbau 2.2(c) keine Auswirkung auf die Intensitätsverteilung in der Brennebene hinter der Linse. Falls sich der Strahlteiler jedoch nicht in der vorderen Brennebene der Linse befindet, verlaufen die replizierten Eingangsstrahlen hinter der Linse nicht parallel, wodurch sich wiederum eine von der longitudinalen Verschiebung des Strahlteilerelements abhängige Skalierung des Abstands zwischen den Ordnungen in der Ausgangsebene ergibt.

Im Rahmen der Dünne-Elemente-Approximation resultieren bei Verkippungen von Strahlteilern bezüglich der optischen Achse drei Effekte. Zum einen skaliert bei einer Verkippung um einen Winkel θ der optisch wirksame Abtastabstand mit dem Faktor $\cos \theta$, was zu einer Skalierung der Signalebenenkoordinaten mit dem Faktor $(\cos \theta)^{-1}$ führt.

Eine zweiter auftretender Effekt ist die Änderung des effektiven Höhenprofils $d'(x)$ im Vergleich zu dem gemäß Gleichung (2.41) zu $t(x)$ korrespondierenden Profil $d(x)$. Das effektive Höhenprofil $d'(x)$ ergibt sich über die Beziehung

$$d'(x) = (\cos \theta)^{-1} d(x). \quad (3.48)$$

und stellt somit einen Spezialfall der linearen Höhenfehler gemäß Gleichung (3.27) dar.

Einen dritten Effekt stellen *Verwaschungen* des Phasenverlaufs dar. Die im Rahmen der Dünne-Elemente-Approximation betrachteten Strahlen parallel zur optischen Achse können beim Auftreten von Verkippungen Stufen im Höhenprofil $d(x)$ schräg schneiden, wodurch die ursprünglichen Sprungstellen linear geglättet werden. Bei pixelierten Designs entspricht dieser Effekt einer Änderung der Pixelform, was zu einer Änderung der globalen Modulation in Gleichung (3.11a) führt.

3.3.4 Toleranzeigenschaften bezüglich des Eingangsstrahls

Bei der Betrachtung der Toleranzeigenschaften bezüglich des Eingangsstrahls unterscheiden wir zwischen der Auswirkung von Wellenlängenabweichungen sowie dem Einfluss der komplexen Amplitude. Eine Abweichung zwischen der beim Design angenommenen und der tatsächlich vor-

liegenden Wellenlänge, welche wir auch als *Wellenlängenfehler* bezeichnen, ist einerseits entsprechend Gleichung (2.41) äquivalent mit den bereits in den Abschnitten 3.3.1 und 3.3.2 betrachteten linearen Höhenfehlern. Des Weiteren ergibt sich eine Skalierung der Signalebenenkoordinaten, welche abhängig von dem gewählten optischen Systemaufbau ist. In den Gleichungen (2.37), (2.45) und (2.46) kann diese Skalierung proportional zu $k = 2\pi/\lambda$ direkt abgelesen werden.

Die Toleranzeigenschaften bezüglich der Wahl der komplexen Amplitude des Eingangsstrahls wurden bereits in den Abschnitten 3.1.1 und 3.2.1 angesprochen. Bei Vorliegen von Überlappungsfreiheit entsprechend Bedingung (3.7) sind Gütefunktionen wie Effizienz und Uniformitätsfehler unabhängig von der Wahl des Eingangsfelds. Bei Verwendung der Designmethoden aus den Abschnitten 3.2.1 und 3.2.2 kann eine Unabhängigkeit des Uniformitätsfehlers von der Wahl des Eingangsfelds erreicht werden, sofern Bedingung (3.19) erfüllt ist.

3.4 Begrenzung der maximalen Rauschintensität

In vielen praktischen Anwendungen von Strahlteilern ist die Intensität der im Allgemeinen nicht vermeidbaren Rauschordnungen eine kritische Größe und muss entsprechend vorgegebener Spezifikationen nach oben begrenzt werden. Die Begrenzung der Intensität von Rauschordnungen, welche durch Pixelierung entstehen, wurde bereits in Abschnitt 3.1.2 betrachtet. In Abschnitt 3.4.1 wird eine Modifikation des Signalprojektionsoperators diskutiert, welche eine Reduktion der maximalen Rauschintensität bis zu einer bestimmten Grenze möglich macht. Diese Grenze hängt – wie in den Abschnitten 3.4.2 und 3.4.3 gezeigt wird – von der gewählten Abtastung der Transmissionsfunktion bzw. von der Anzahl der Quantisierungsstufen ab. In Abschnitt 3.4.4 demonstrieren wir einen praktischen Anwendungsfall für Strahlteiler mit reduzierter Rauschintensität.

3.4.1 Modifikation des Signalprojektionsoperators

Zur Einbeziehung einer oberen Begrenzung der Rauschintensität in eine IFTA-basierte Optimierung muss die Wirkung der Signalprojektionsoperation innerhalb des Bereichs der Amplitudenfreiheit zur Begrenzung auf einen vorgegebenen Maximalwert \bar{u}_{\max} modifiziert werden. Für den Operator Π_{PAS} aus Abschnitt 2.4.2 ergibt sich

$$\Pi_{\text{PAS}} u_{\text{out}}(x') = \begin{cases} |\alpha u_{\text{sig}}(x')| \exp [i \arg (u_{\text{out}}(x'))] & \text{falls } x' \in \mathcal{W}_{\text{sig}} \\ \text{clip}(u_{\text{out}}(x'), \bar{u}_{\max}) & \text{sonst} \end{cases} \quad (3.49)$$

mit

$$\text{clip}(u, \bar{u}_{\max}) = \begin{cases} u & \text{falls } |u| \leq \bar{u}_{\max} \\ u \bar{u}_{\max} / |u| & \text{sonst} \end{cases} . \quad (3.50)$$

Für den Operator Π_{noise} aus Abschnitt 3.2.2 kann diese Modifikation analog durchgeführt werden:

$$\Pi_{\text{noise}} u(x') = \begin{cases} 0 & \text{falls } \exists x'' \in \mathcal{W}_{\text{sig}} : |x' - x''| \leq d_{\text{frame}} / 2 \\ \text{clip}(u(x'), \bar{u}_{\max}) & \text{sonst} \end{cases} . \quad (3.51)$$

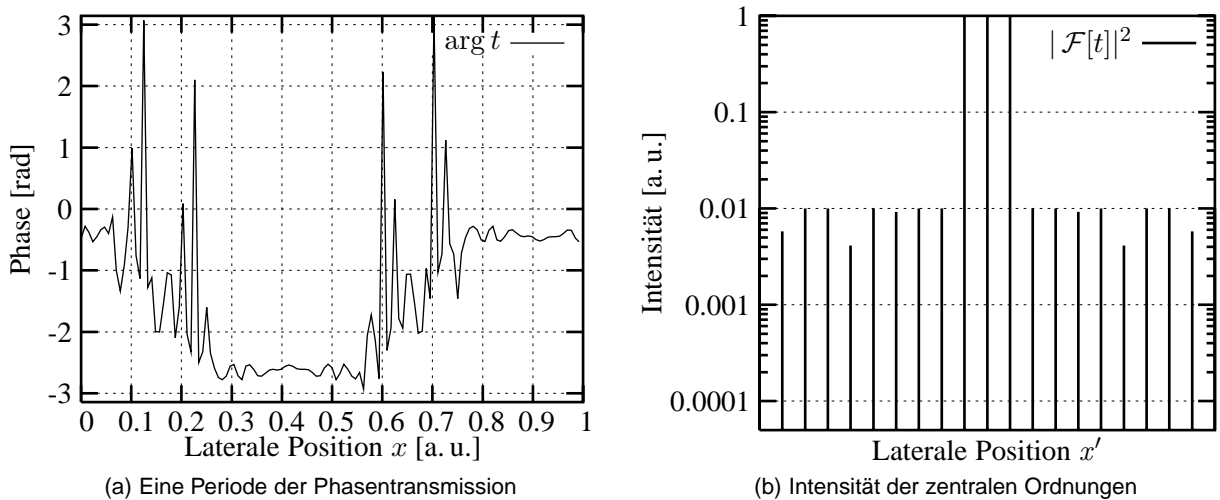


Abbildung 3.13: Lösung für einen regulären 1-zu-3-Strahlteiler mit einer auf 1 % reduzierten relativen Rauschintensität

Für die Wahl von \bar{u}_{\max} besteht ein sinnvoller unterer Grenzwert \bar{u}_{\lim} , welcher der minimal erreichbaren Rauschintensität entspricht. Solange \bar{u}_{\max} größer als \bar{u}_{\lim} gewählt wird, kann bei der Iteration die angestrebte Rauschreduktion auf \bar{u}_{\max} im Wesentlichen erreicht werden, wobei jedoch die beim Design erreichte Effizienz im Allgemeinen umso geringer ausfällt, je tiefer \bar{u}_{\max} gewählt wurde. Dies liegt an der durch die Beschränkung der Amplitudenfreiheit hervorgerufenen Verringerung der Designfreiheiten. Bei einer Wahl von \bar{u}_{\max} kleiner als \bar{u}_{\lim} kann die gewünschte Rauschreduktion nicht vollständig realisiert werden, was auch dazu führt, dass die Signaleinschränkung hinsichtlich der Signalordnungen ebenfalls schlechter erfüllt wird. Dabei resultiert häufig ein erhöhter Uniformitätsfehler. In vielen Fällen kann bei weiterer Reduktion von $\bar{u}_{\max} < \bar{u}_{\lim}$ wiederum ein Ansteigen der erzielten Rauschintensität beobachtet werden. In der Praxis kann es daher notwendig sein, das Design mehrfach für verschiedene Werte von \bar{u}_{\max} durchzuführen, um eine ausreichende Rauschreduktion bzw. einen geeigneten Kompromiss zwischen Rauschintensität und Effizienz zu erzielen.

Im Folgenden demonstrieren wir die Rauschreduktion durch Anwendung des Signalprojektionsoperators Π_{PAS} nach Gleichung (3.49) am Beispiel des Designs eines regulären 1-zu-3-Strahlteilers mit Phaseneinschränkung. Die in Abbildung 3.3 gezeigte Lösung besitzt die für dieses Designproblem optimale Effizienz von 92.6 %, wobei in der ± 3 . Ordnung eine relative Rauschintensität von etwa 8 % vorliegt. Die hier und im Folgenden angegebenen relativen Rauschintensitäten beziehen sich auf die durchschnittlich erreichte Intensität der Signalordnungen.

Die Durchführung des Designverfahrens nach Abbildung 3.1 mit einem Wert von \bar{u}_{\max} , welcher einer relativen Rauschintensität von 1 % entspricht, führt zu dem in Abbildung 3.13 gezeigten Design. Hierbei wurde die Transmissionsfunktion mit $n_x = 128$ abgetastet. Das gezeigte Design besitzt eine Effizienz von 81.4 % bei einer relativen Rauschintensität von 1 %, was die erwähnte Effizienzreduktion durch Einführung der Rauschbeschränkung demonstriert. In dem betrachteten Beispiel entspricht \bar{u}_{\max} näherungsweise \bar{u}_{\lim} , d. h. es ist keine weitere deutliche Reduktion der Rauschintensität möglich.

Die Einführung einer Rauschbegrenzung führt typischerweise zu einer stärkeren Ausnutzung

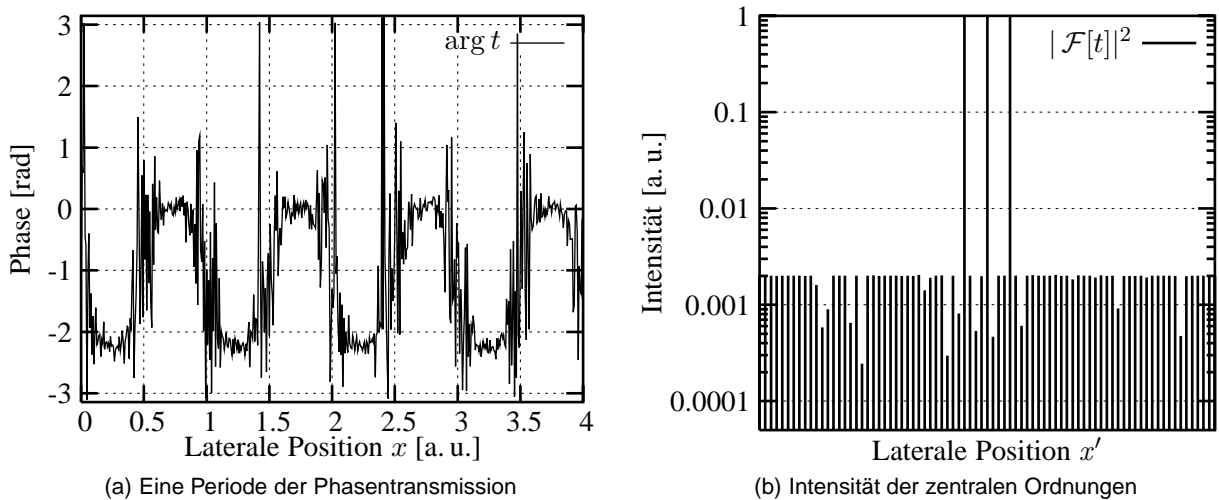


Abbildung 3.14: Lösung für einen regulären 1-zu-3-Strahlteiler mit einer auf 0.2% reduzierten relativen Rauschintensität

höherer Rauschordnungen, was an den in Abbildung 3.13(a) im Vergleich zu Abbildung 3.3(a) auftretenden hohen Ortsfrequenzen erkennbar ist. In der Praxis ist die Einführung hoher Ortsfrequenzen im Allgemeinen durch die Forderung nach Gültigkeit der Dünne-Elemente-Approximation beschränkt (vgl. Diskussion in Abschnitt 3.1.3). Des Weiteren können Höhenprofile entsprechend Abbildung 3.13(a) auf Grund der auftretenden Sprünge häufig deutlich schwerer mit ausreichender Genauigkeit realisiert werden als glatte Profile wie in Abbildung 3.3(a). Aus den genannten Gründen ergibt sich somit in vielen Fällen eine Beschränkung in der Ausnutzung höherer Rauschordnungen zur Reduzierung der Rauschintensität.

3.4.2 Erhöhung der Anzahl der zur Verfügung stehenden Rauschordnungen

Da bei einem Design mit einer Effizienz η im Fall einer Phasentransmission der Leistungsanteil $(1 - \eta)$ auf die Rauschordnungen verteilt werden muss, kann es zur Einführung einer Begrenzung für den maximalen Leistungsanteil pro Rauschordnung notwendig werden, die Zahl der zur Verfügung stehenden Rauschordnungen zu erhöhen.

Da jedoch entsprechend der Diskussion am Ende des vorangegangenen Abschnitts der maximale Ablenkwinkel der als Amplitudenfremie verfügbaren Rauschordnungen im Allgemeinen begrenzt ist, kann es notwendig werden, die Zahl der Rauschordnungen durch eine Verfeinerung der Signalebenenraasterung zu erhöhen. In Abschnitt 3.2.1 wurde bereits die Verfeinerung der Signalebenenraasterung zur Erhöhung der Positioniergenauigkeit der Signalordnungen diskutiert. Die dabei dargestellten Aspekte im Zusammenhang mit der Überlappungsfremie sind auch bei einer Verfeinerung der Signalebenenraasterung zur Rauschreduktion zu berücksichtigen.

Als Beispiel betrachten wir in Abbildung 3.14 ein Design mit einer im Vergleich zu Abbildung 3.13 um den Faktor vier vergrößerten Periodenlänge, d. h. $n_x = 512$. Die resultierende Verfeinerung der Signalebenenraasterung ist in Abbildung 3.14(b) erkennbar. Da für dieses Design Überlappungsfremie angenommen wurde, erfolgte keine Unterdrückung der Rauschordnungen

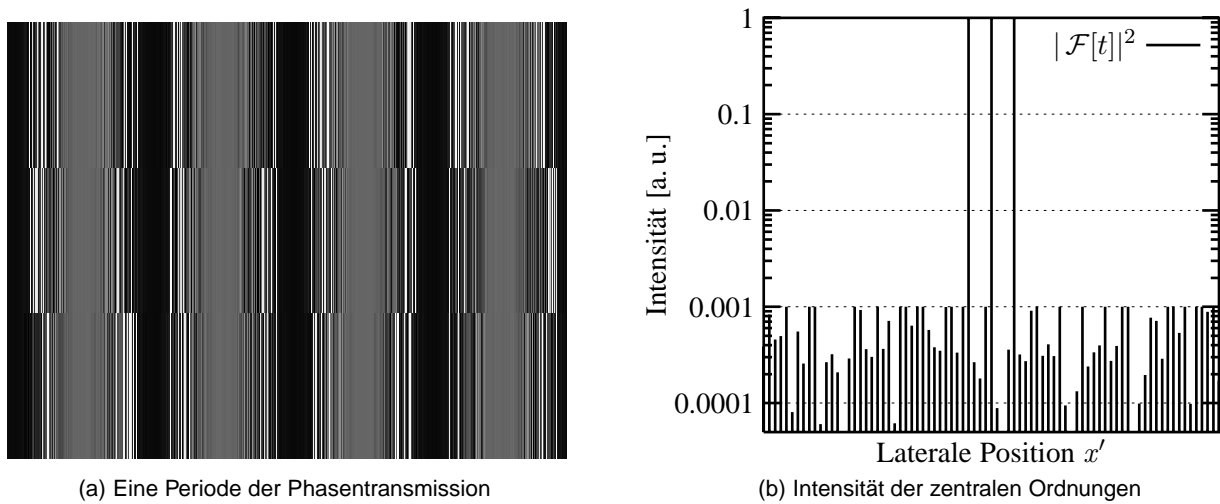


Abbildung 3.15: Lösung für einen regulären 1-zu-3-Strahlteiler mit einer auf 0.1 % reduzierten relativen Rauschintensität

in der Umgebung der Signalordnungen entsprechend Gleichung (3.18). Für das gezeigte Design wurde eine Effizienz von 78.0 % bei einer relativen Rauschintensität von 0.2 % erreicht.

Die in Abbildung 3.14(a) gezeigte Periode der Transmissionsfunktion weist eine Grundstruktur auf, welche einer vierfachen Wiederholung des Verlaufs aus Abbildung 3.13(a) ähnelt. Daran, dass in Abbildung 3.14(b) im Vergleich zu Abbildung 3.13(b) Zwischenordnungen entstehen und ausgenutzt werden zeigt sich jedoch, dass Abweichungen zwischen den vier *Quasi-Unterperioden* in Abbildung 3.14(a) bestehen.

Die Ausnutzung der Zwischenordnungen muss bei Anwendung des Designschemas aus Abbildung 3.1 durch Einfügung eines zusätzlichen Störschrittes erzwungen werden. Dieser wurde in dem gezeigten Beispiel implementiert, indem vor Beginn der SNR-Optimierung in der Signalebene ein zufälliges Rauschen mit einer im Vergleich zu \bar{u}_{\max} kleinen Amplitude aufaddiert wurde.

Bei eindimensionalen Strahlteilungsproblemen stellt der Übergang zu einem zweidimensionalen Design ($n_y \neq 1$) eine weitere Möglichkeit dar, die Zahl der zur Verfügung stehenden Rauschordnungen zu erhöhen. Das in Abbildung 3.15 gezeigte Designbeispiel demonstriert, dass dieses Konzept sehr nützlich zur Reduktion der Rauschintensität sein kann. Bei dem betrachteten Beispiel ermöglichte ein Übergang zu einem zweidimensionalen Design mit $n_y = 3$ eine Reduktion des Rauschens auf 0.1 % bei einer im Vergleich zum Beispiel aus Abbildung 3.14 sogar leicht erhöhten Effizienz von 80.8 %. In Abbildung 3.15(a) ist erkennbar, dass die in den drei Zeilen innerhalb einer Periode der Transmissionsfunktion entstehenden *Quasi-Unterperioden* wiederum nur geringfügig voneinander abweichen.

3.4.3 Anzahl der Quantisierungsstufen

Bei Einführung einer Quantisierung entsteht im Allgemeinen zusätzliches Rauschen, welches die Reduzierung der Rauschintensität erschwert. Die Abhängigkeit der Reduzierbarkeit der Rauschintensität von der Anzahl der Quantisierungsstufen Q wird in Tabelle 3.2 für die Quantisierung des Beispiels aus Abbildung 3.14 demonstriert. Dabei wird das Softquantisierungsschema entspre-

Q	Rauschintensität		Effizienz	Uniformitätsfehler	Abbildung
	(angestrebt)	(erreicht)			
∞	1 %	1.0 %	81.4 %	0.0 %	3.14
16	1 %	1.2 %	77.2 %	0.8 %	–
8	1 %	1.3 %	75.3 %	1.3 %	3.16
4	1 %	3.9 %	62.8 %	4.8 %	–
4	2 %	2.8 %	58.5 %	3.3 %	–
2	1 %	5.2 %	50.2 %	1.0 %	–
2	3 %	4.4 %	49.8 %	0.4 %	–

Tabelle 3.2: Abhängigkeit für Reduzierbarkeit der Rauschintensität von der Anzahl der Quantisierungsstufen Q

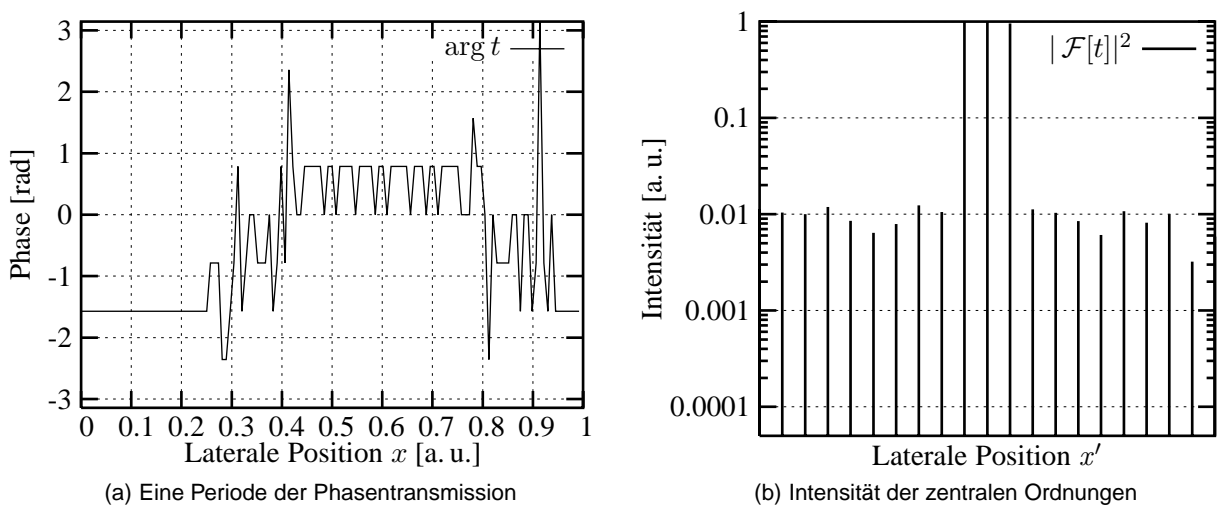


Abbildung 3.16: Achtstufig quantisierte Lösung für regulären 1-zu-3-Strahlteiler mit einer auf 1.3% reduzierten relativen Rauschintensität

chend Abbildung 3.2 angewendet, wobei für jeden Wert von Q jeweils 10 Versuche mit unterschiedlichen Startverteilungen durchgeführt wurden. Für Tabelle 3.2 wurde davon jeweils die Lösung mit der geringsten Rauschintensität ausgewählt. Abbildung 3.16 zeigt eine auf diese Weise für $Q = 8$ resultierende Lösung mit einer Effizienz von 75.3% bei einem Uniformitätsfehler von 1.3%.

Aus Tabelle 3.2 wird ersichtlich, dass für geringere Q die minimal erreichbare Rauschintensität zunimmt, wobei die Effizienz abnimmt. Weiterhin wird das in Abschnitt 3.4.1 erwähnte Ansteigen des Uniformitätsfehlers im Fall von $\bar{u}_{\max} < \bar{u}_{\lim}$ demonstriert. Die bei $Q = 2$ beobachteten relativ geringen Uniformitätsfehler resultieren aus der Symmetrie der Intensitätsverteilung der Ausgangsordnungen im Fall einer reellen Transmissionsfunktion.

3.4.4 Anwendungsbeispiel Markierungsoptiken

In modernen Pyrometern werden vielfach Markierungsoptiken eingesetzt, welche dem Anwender für die berührungslose Temperaturmessung anvisierten Messbereich veranschaulichen. Solche Markierungsoptiken basieren vielfach auf diffraktiven Strahlteilern, die den kollimierten Strahl

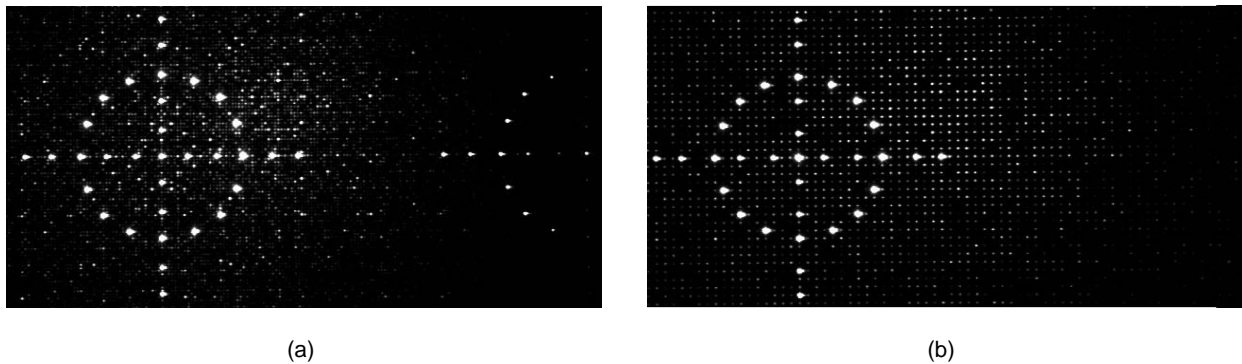


Abbildung 3.17: CCD-Kameraaufnahme der für zwei verschiedene Markierungsoptiken resultierenden Fernfeldintensitätsverteilung. Im Gegensatz zu Design (a) ist bei Design (b) kein Pixelierungsrausch sichtbar und die maximale Intensität der Rauschordnungen ist auf einen niedrigeren Wert begrenzt.

einer Laserdiode zur Markierung des Messbereichs replizieren.

Abbildung 3.17(a) zeigt einen Ausschnitt der Fernfeldintensitätsverteilung welches qualitativ näherungsweise einem derzeit mit großen Stückzahlen hergestellten Pyrometer entspricht. Das zu Abbildung 3.17(a) gehörende Strahlteiler-element wurde mit 128×128 Abtastwerten in einer Periode ohne Beschränkung der maximalen Rauschintensität beim Design ermittelt und mit einer Pixelgröße von $3 \mu\text{m}$ für $\lambda_{\text{ref}} = 632.8 \text{ nm}$ hergestellt. Die äußersten Signalordnungen entstehen für $\lambda = \lambda_{\text{ref}}$ in einem Winkel von 4° zur optischen Achse, wobei der für die gewählte Abtastung entsprechend Gleichung (3.13) resultierenden maximalen Ablenkwinkel ϕ_{max} in diesem Fall 6.1° beträgt. Die somit resultierenden starken Pixelierungseffekte sind Abbildung 3.17(a) deutlich erkennbar und führen bei dem kommerziell erhältlichen Gerät zu relativen Intensitäten der durch Pixelierung entstandenen Rauschordnungen von etwa 9 %, wobei die maximale relative Intensität der in der direkten Umgebung der Signalordnungen liegenden Rauschordnungen etwa 3 % beträgt. Die angegebenen relativen Intensitätswerte beziehen sich hierbei auf die durchschnittliche gemessene Intensität der Signalordnungen.

Abbildung 3.17(b) zeigt die für ein für ein hinsichtlich der Rauschverteilung deutlich verbessertes Design resultierende Fernfeldintensitätsverteilung. Eine Periode der Transmissionsfunktion besitzt hierbei 256×256 Abtastwerte bei einer Pixelgröße von $0.6 \mu\text{m}$. Da der für ϕ_{max} in diesem Fall resultierende Wert mit 31.8° deutlich größer ist als die Ablenkwinkel der Signalordnungen, sind Pixelierungseffekte vernachlässigbar. Unter Ausnutzung der Methoden aus den Abschnitten 3.17(b) und 3.17(b) wurde für die relative Intensität der in der direkten Umgebung der Signalordnungen liegenden Rauschordnungen der Wert von 1 % simuliert, wobei sich experimentell ein Wert von 1.3 % ergeben hat.

Kapitel 4

Verbesserung des Konvergenzverhaltens durch verstärkte Projektionen

Mit dem zunehmenden Einsatz von Projektionsalgorithmen zur Lösung in der Praxis auftretender optischer Designprobleme gewinnen Fragen nach dem Rechenaufwand sowie dem Konvergenzverhalten solcher Verfahren an Bedeutung. Innerhalb dieses Kapitels, dessen Ergebnisse auch in Referenz [36] publiziert wurden, untersuchen wir Möglichkeiten zur Konvergenzbeschleunigung, welche sich auf Basis des in Abschnitt 2.1.3 eingeführten Konzepts der Projektionsstärke ergeben. Nach einigen Vorbetrachtungen in Abschnitt 4.1 betrachten wir in Abschnitt 4.2 die Konvergenzbeschleunigung beim Design von Phasentransmissionen. Im Anschluss daran wird in Abschnitt 4.3 ein Verfahren zur Überwindung von beim Design quantisierter Transmissionen auftretender Stagnationen diskutiert.

4.1 Vorbetrachtungen

Nach Darstellung des Grundschemas der in diesem Abschnitt betrachteten Projektionsalgorithmen in Abschnitt 4.1.1 wird in Abschnitt 4.1.2 ein Maß zur Beschreibung der Konvergenzgeschwindigkeit definiert. Frühere Arbeiten zum Einsatz von verstärkten Projektionsoperatoren werden in Abschnitt 4.1.3 diskutiert.

4.1.1 Grundschema

Im Folgenden betrachten wir Transmissionsdesignprobleme mit der in Abbildung 2.3 gezeigten Struktur des optischen Systems. Wir beschränken uns in der Darstellung auf monofunktionale Designprobleme, wobei die eingeführten Konzepte auch auf multifunktionale Designprobleme verallgemeinerbar sind.

Die Einbeziehung von verstärkten Projektionen in den IFTA erfolgt durch Anwendung eines Operators Π_{sig}^λ in der Signalebene. Ein derart modifizierter IFTA kann als Spezialfall des in Abschnitt 2.4.5 eingeführten Input-Output-Konzepts aufgefasst werden. Wir betrachten somit im Folgenden einen Input-Output-Algorithmus entsprechend Gleichung (2.76), bei dem das Steuerfeld in

Iteration n gemäß

$$u_{\text{dr}}^{(n)} = \Pi_{\text{sig}}^\lambda u_{\text{out}}^{(n-1)} = u_{\text{out}}^{(n-1)} + \lambda (\Pi_{\text{sig}} u_{\text{out}}^{(n-1)} - u_{\text{out}}^{(n-1)}) \quad (4.1)$$

gewählt wird.

In den Abschnitten 4.2 und 4.3 werden verschiedene Strategien zur Wahl der Projektionsstärke während der Iteration diskutiert. Bei der Ableitung in Abschnitt 4.2 wird davon ausgegangen, dass leichte Änderungen des Steuerfelds u_{dr} zu ähnlichen Änderungen des Ausgangsfelds führen [83], d. h.

$$\bar{\mathcal{S}}(u_{\text{dr}} + \Delta u_{\text{dr}}) = \bar{\mathcal{S}} u_{\text{dr}} + \bar{\alpha} \Delta u_{\text{dr}} + u_{\text{noise}}, \quad (4.2)$$

wobei $\bar{\alpha}$ einen reellen Durchkoppelfaktor und u_{noise} einen Rauschterm bezeichnet. Günstigerweise wird $\bar{\alpha}$ zur Minimierung von $\|u_{\text{noise}}\|$ gewählt als

$$\bar{\alpha} = \text{Re} \left[\frac{\langle \bar{\mathcal{S}}(u_{\text{dr}} + \Delta u_{\text{dr}}) - \bar{\mathcal{S}} u_{\text{dr}} | \Delta u_{\text{dr}} \rangle}{\|\Delta u_{\text{dr}}\|^2} \right]. \quad (4.3)$$

Unter Anwendung von Gleichung (4.2) resultiert durch Einsetzen von Gleichung (4.1) in Gleichung (2.76)

$$\begin{aligned} u_{\text{out}}^{(n)} &= \bar{\mathcal{S}} \left(u_{\text{out}}^{(n-1)} + \lambda (\Pi_{\text{sig}} u_{\text{out}}^{(n-1)} - u_{\text{out}}^{(n-1)}) \right) \\ &= u_{\text{out}}^{(n-1)} + \bar{\alpha} \lambda (\Pi_{\text{sig}} u_{\text{out}}^{(n-1)} - u_{\text{out}}^{(n-1)}) + u_{\text{noise}} \end{aligned} \quad (4.4)$$

$$= (1 - \bar{\alpha} \lambda) u_{\text{out}}^{(n-1)} + \bar{\alpha} \lambda \Pi_{\text{sig}} u_{\text{out}}^{(n-1)} + u_{\text{noise}}. \quad (4.5)$$

Hierbei wurde die Idempotenzeigenschaft

$$\bar{\mathcal{S}} \bar{\mathcal{S}} = \bar{\mathcal{S}}. \quad (4.6)$$

des Operators $\bar{\mathcal{S}}$ ausgenutzt.

4.1.2 Konvergenzrate

Zur Bewertung der Konvergenzgeschwindigkeit definieren wir ein auf dem in Abschnitt 2.1.2 mit Gleichung (2.11) eingeführten summierten Abstandsfehler J basierendes Maß v_{conv} als

$$v_{\text{conv}}^{(n)} = J(u_{\text{out}}^{(n-1)})^2 - J(u_{\text{out}}^{(n)})^2, \quad (4.7)$$

wobei v_{conv} im Folgenden als *Konvergenzrate* bezeichnet wird. Die Quadrierung der SDE-Terme wurde zur Vereinfachung der folgenden mathematischen Betrachtungen eingeführt.

Im Gegensatz zur Konvergenzrate verbinden wir mit dem Begriff der *Konvergenzgeschwindigkeit* einen zeitlichen Bezug. Sie kann beispielsweise als Quotient von v_{conv} mit der für eine Iteration benötigten Rechenzeit definiert werden.

Die Konvergenzrate in Iteration n hängt natürlich von der Wahl der Projektionsstärke λ bei der Ermittlung des Steuerfelds entsprechend Gleichung (4.1) ab. Um näheren Aufschluss über

diese Abhängigkeit zu erhalten, betrachten wir zunächst den Term $J(u_{\text{out}}^{(n)})^2$ aus Gleichung (4.7). Aufgrund der Abstandsminimierungseigenschaft von Π_{sig} erhalten wir

$$\begin{aligned} J(u_{\text{out}}^{(n)})^2 &= \|u_{\text{out}}^{(n)} - \Pi_{\text{sig}} u_{\text{out}}^{(n)}\|^2 \\ &\leq \|u_{\text{out}}^{(n)} - \Pi_{\text{sig}} u_{\text{out}}^{(n-1)}\|^2. \end{aligned} \quad (4.8)$$

Durch Einsetzen von Gleichung (4.5) erhalten wir

$$J(u_{\text{out}}^{(n)})^2 \leq \|(1 - \bar{\alpha} \lambda) u_{\text{out}}^{(n-1)} - (1 - \bar{\alpha} \lambda) \Pi_{\text{sig}} u_{\text{out}}^{(n-1)} + u_{\text{noise}}\|^2, \quad (4.9)$$

wobei u_{noise} im Allgemeinen natürlich auch von λ abhängt. Unter Verwendung der Dreiecksungleichung (2.2) folgt

$$\begin{aligned} J(u_{\text{out}}^{(n)})^2 &\leq \|(1 - \bar{\alpha} \lambda) u_{\text{out}}^{(n-1)} - (1 - \bar{\alpha} \lambda) \Pi_{\text{sig}} u_{\text{out}}^{(n-1)}\|^2 + \|u_{\text{noise}}\|^2 \\ &= |1 - \bar{\alpha} \lambda|^2 J(u_{\text{out}}^{(n-1)})^2 + \|u_{\text{noise}}\|^2. \end{aligned} \quad (4.10)$$

Durch Einsetzen der Beziehung (4.10) in Gleichung (4.7) erhalten wir schließlich als untere Abschätzung für $v_{\text{conv}}^{(n)}(\lambda)$ die Beziehung

$$v_{\text{conv}}^{(n)}(\lambda) \geq (1 - |1 - \bar{\alpha} \lambda|^2) J(u_{\text{out}}^{(n-1)})^2 - \|u_{\text{noise}}\|^2, \quad (4.11)$$

auf deren Basis in Abschnitt 4.2 ein Verfahren zur Steuerung von λ entwickelt wird.

4.1.3 Frühere Arbeiten

Modifizierte Projektionsoperatoren wurden bereits in früheren Arbeiten zum Transmissionsdesign eingesetzt. Beim Design von Strahlformern sind relaxierte Projektionen der Form (2.13) hilfreich zur Vermeidung von *Speckles* [84]. Johansson und Bengtsson [85] setzten zur Strahlformung verstärkte Projektionen ein, welche im Wesentlichen der Form (2.13) mit $\lambda = 2$ entsprechen, wobei jedoch die durch den Projektionsoperator lokal eingeführte Veränderung des Feldes beschränkt wurde, um Speckles zu vermeiden. Der bei der Softquantisierung verwendete Operator $\Pi_{\text{softquant}, Q}^\lambda$ aus Gleichung (3.16) stellt eine abgeschwächte Version des Quantisierungsoperators $\Pi_{\text{quant}, Q}$ dar, wobei die Abschwächung jedoch nicht der Form (2.13) entspricht.

Dass durch Modifikation von Projektionsoperatoren die Konvergenzgeschwindigkeit von Projektionsalgorithmen erhöht werden kann, wurde bereits von Fienup [83, 86] demonstriert. Neben speziellen Formen von Input-Output-Algorithmen betrachtete er auch den Ansatz von Abschnitt 4.1.1 und unterbreitete den Vorschlag, die Projektionsstärke entsprechend

$$\lambda = \bar{\alpha}^{-1} \quad (4.12)$$

zu wählen. Dieser Ansatz liegt zunächst einmal nahe, da er zu einem Verschwinden des ersten Terms in Gleichung (4.5) führt. Des Weiteren führt er zu einer Maximierung des ersten Terms in der Abschätzung (4.11). Im Allgemeinen ist $\bar{\alpha}$ jedoch nicht konstant und hängt insbesondere von λ ab, was dazu führt, dass anstelle von Gleichung (4.12) eigentlich eine Lösung von

$$\bar{\alpha}(\lambda)\lambda = 1 \quad (4.13)$$

gefunden werden müsste, wobei jedoch die Beziehung $\bar{\alpha}(\lambda)$ im Allgemeinen nicht bekannt ist. Ein weiteres mit diesem Ansatz verbundenes Problem besteht darin, dass für steigende Werte von λ die mit u_{noise} verbundenen Terme in Gleichung (4.5) und in Abschätzung (4.11) an Einfluss gewinnen, was wiederum zu einer Verschlechterung der Konvergenzrate führt. Für eine optimale Konvergenzrate müsste demnach für λ ein Wert gewählt werden, der etwas kleiner als die kleinste Lösung von Gleichung (4.13) ist.

Levi und Stark [42] untersuchten ebenfalls das Problem der Verwendung der optimalen Projektionsstärke λ_{opt} . Sie wiesen die Beziehung

$$\lambda_{\text{opt}} \geq 1 \quad (4.14)$$

nach und schlugen vor, λ_{opt} über die direkte Suche nach einem Maximum von $v_{\text{conv}}(\lambda)$ zu bestimmen. Dieser Ansatz ist praktikabel, wenn die involvierten Projektionsoperatoren linear sind, da nur in diesem Fall der Aufwand der Suche nach λ_{opt} klein gegenüber dem Aufwand einer Iteration ist. Die im Transmissionsdesign typischerweise verwendeten Projektionsoperatoren (vgl. Abschnitte 2.4.1 und 2.4.2) besitzen diese Eigenschaft jedoch im Allgemeinen nicht.

Nashold und Saleh [87] untersuchten den Einfluss der Projektionsstärke auf die Konvergenzgeschwindigkeit von Projektionsalgorithmen bei Verwendung von entsprechend des Schemas (2.13) verstärkten Projektionen. Für die Wahl der während der Iteration konstant gehaltenen Projektionsstärke λ wurde dabei keine Strategie angegeben. In der genannten Arbeit wurde gezeigt, dass Verbesserungen des Konvergenzverhaltens für bestimmte Werte $\lambda > 1$ erzielt werden können, wobei jedoch zu große Werte λ auch zu einer Verschlechterung des Konvergenzverhaltens führen können.

4.2 Konvergenzbeschleunigung beim Design von Phasentransmissionen

In diesem Abschnitt entwickeln wir ein Verfahren zur Veränderung der Projektionsstärke während der Iteration, welches bei Vorliegen einer Phaseneinschränkung in der Transmissionsebene und Amplitudenfreiheit in der Signalebene typischerweise zu einer deutlichen Erhöhung der Konvergenzgeschwindigkeit im Vergleich zum IFTA führt. Im Rahmen der in den Abbildungen 3.1 und 3.2 gezeigten Ablaufschemen für das Design von Strahlteilern ist das im Folgenden vorgestellte Verfahren somit während der SNR-Optimierung einsetzbar.

Zur Illustration der entwickelten Konzepte betrachten wir in diesem Abschnitt zwei Strahlteileranwendungen als Beispiele: Zum einen untersuchen wir das Design von regulären eindimensionalen 1-zu-3-Strahlteilern, motiviert durch das Vorhandensein einer optimalen analytischen Lösung für dieses Designproblem [88]. Hierbei werden 1024 Abtastwerte innerhalb einer Periode der Transmission berechnet. Da der für eindimensionale Strahlteilerdesigns notwendige Rechenaufwand im Allgemeinen sehr niedrig ist, betrachten wir als zweites Beispiel mit dem Design eines Strahlteilers zur Erzeugung eines regulären Feldes von 300×300 Signalordnungen ein etwas anspruchsvolleres Designproblem, wobei 1000×1000 Abtastwerte innerhalb einer Periode berechnet

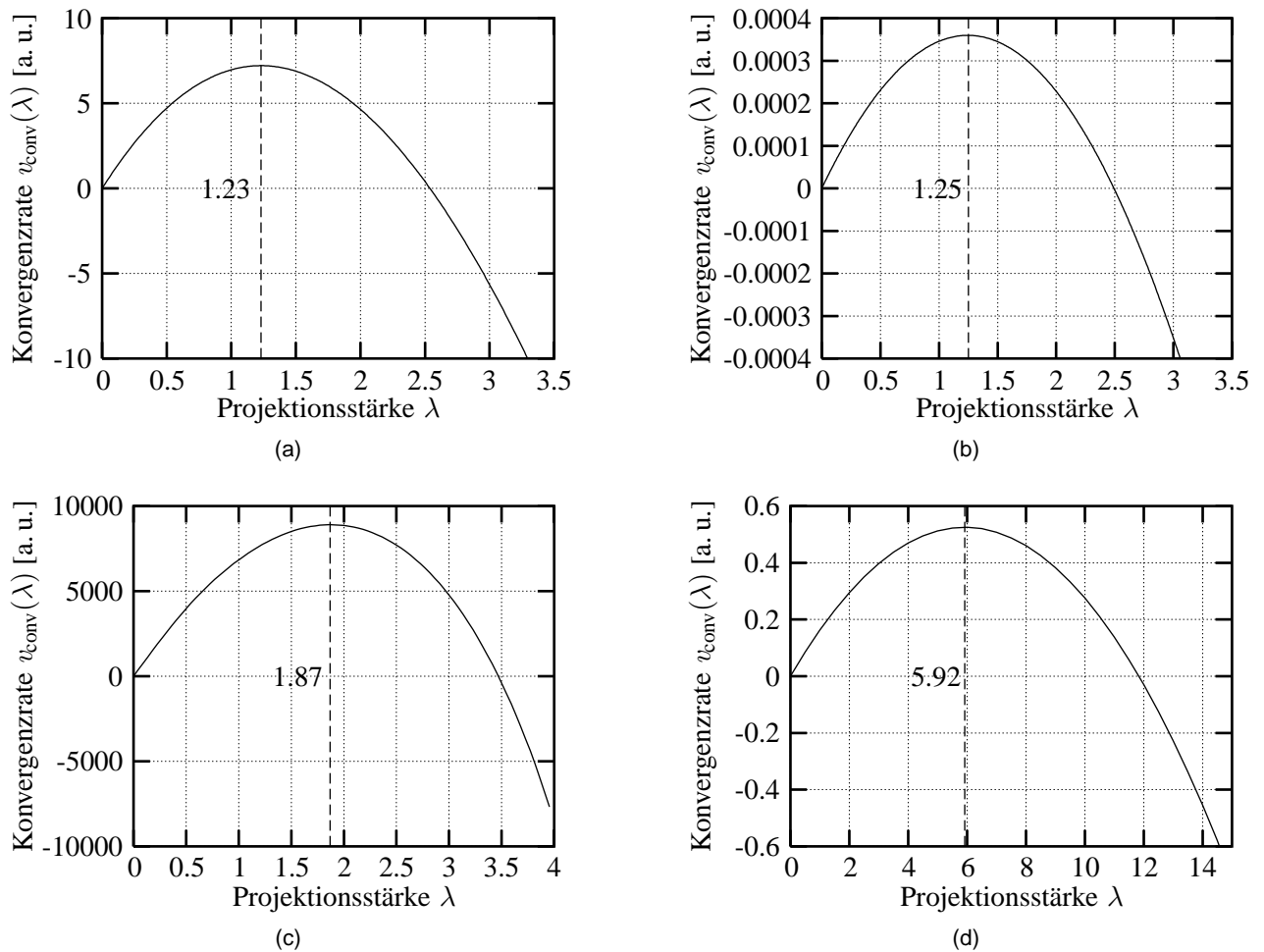


Abbildung 4.1: Abhängigkeit der Konvergenzrate von der Projektionsstärke für das Design eines regulären 1-zu-3-Strahlteilers vor der ersten (a) bzw. vierten (b) Iteration der SNR-Optimierung; sowie für das Design eines Strahlteilers zur Erzeugung eines regulären Feldes von 300×300 Signalordnungen ebenfalls vor der ersten (c) bzw. vierten (d) Iteration. Vor der SNR-Optimierung wurde entsprechend des in Abbildung 3.1 gezeigten Ablaufs eine Phasensynthese mit jeweils 10 Iterationen durchgeführt.

werden.

4.2.1 Abhängigkeit der Konvergenzrate von der Projektionsstärke

In Abschnitt 4.1.2 wurde mit Gleichung (4.11) eine untere Abschätzung für den Zusammenhang $v_{\text{conv}}^{(n)}(\lambda)$ abgeleitet. In Abbildung 4.1 ist der tatsächliche Verlauf $v_{\text{conv}}^{(n)}(\lambda)$ für die beiden betrachteten Beispielanwendungen an jeweils zwei verschiedenen Iterationszuständen dargestellt. Die optimale Projektionsstärke λ_{opt} , d. h. die Position des Maximums von $v_{\text{conv}}^{(n)}(\lambda)$, wurde in den Diagrammen von Abbildung 4.1 jeweils durch gestrichelte Linien gekennzeichnet. Es ist erkennbar, dass λ_{opt} im Allgemeinen sowohl vom Designproblem als auch vom Status innerhalb der Iteration abhängt: Während sich bei dem in den Abbildungen 4.1(a) und 4.1(b) betrachteten Designproblem nur eine leichte Änderung von λ_{opt} zwischen den beiden betrachteten Iterationszuständen ergibt, kann im Vergleich von Abbildung 4.1(c) und 4.1(d) ein relativ starker Anstieg beobachtet werden. Des Weiteren bestätigt sich in Abbildung 4.1 die Gültigkeit der Ungleichung (4.14).

4.2.2 Steuerung der Projektionsstärke

Besonders in Diagramm 4.1(d) besteht für λ_{opt} eine deutlich höhere Konvergenzrate $v_{\text{conv}}(\lambda_{\text{opt}})$ als in dem beim IFTA vorliegenden Fall mit $v_{\text{conv}}(1)$. Eine dynamische Anpassung der Projektionsstärke während der Iteration kann demnach zu einer deutlichen Verbesserung des Konvergenzverhaltens führen.

Der in diesem Abschnitt vorgestellte Algorithmus verfolgt das Ziel, vor jeder Iteration diejenige Projektionsstärke λ zu finden, welche die Kurve $v_{\text{conv}}(\lambda)$ für die darauf folgende Iteration maximiert. Damit aus der Erhöhung der Konvergenzrate im Vergleich zum Fall $\lambda = 1$ auch eine höhere Konvergenzgeschwindigkeit resultiert, muss die Ermittlung von λ mit einem relativ geringen Rechenaufwand verbunden sein.

Leider benötigt die Ermittlung eines Punktes der Kurve $v_{\text{conv}}(\lambda)$ einen Rechenaufwand, der etwa dem Aufwand einer Iteration entsprechend Gleichung (2.76) entspricht. Der in Abschnitt 4.1.3 erwähnte Ansatz von Levi und Stark [42], den Wert λ_{opt} durch Suche innerhalb von $v_{\text{conv}}(\lambda)$ zu bestimmen, erscheint daher nicht praktikabel. Das Wissen über den Verlauf der Kurven $v_{\text{conv}}(\lambda)$ muss somit während der Iteration gesammelt werden, wobei wir für eine Abschätzung von λ_{opt} annehmen müssen, dass sich der Kurvenverlauf – oder zumindest die Position des Kurvenmaximums – zwischen aufeinanderfolgenden Iterationen nur leicht verändert.

In Abschnitt A.3.1 wird auf Basis der in Abschnitt 4.1.2 gezeigten Beziehung (4.11) folgendes Schema zur Berechnung von $\lambda^{(n+1)}$ als Approximation der in Iteration $(n+1)$ optimalen Projektionsstärke auf Basis der Projektionsstärke $\lambda^{(n)}$ sowie der Ausgangsfelder $u_{\text{out}}^{(n)}$ und $u_{\text{out}}^{(n-1)}$ abgeleitet:

$$\lambda^{(n+1)} = \frac{\bar{\alpha} J(u_{\text{out}}^{(n-1)})^2}{\bar{\alpha}^2 J(u_{\text{out}}^{(n-1)})^2 + \beta} \quad \text{mit} \quad (4.15a)$$

$$\bar{\alpha} = \text{Re} \left[\frac{\langle u_{\text{out}}^{(n)} - u_{\text{out}}^{(n-1)} \mid \lambda^{(n)} (\Pi_{\text{sig}} u_{\text{out}}^{(n-1)} - u_{\text{out}}^{(n-1)}) \rangle}{\| \lambda^{(n)} (\Pi_{\text{sig}} u_{\text{out}}^{(n-1)} - u_{\text{out}}^{(n-1)}) \|^2} \right] \quad \text{und} \quad (4.15b)$$

$$\beta = \frac{|1 - \bar{\alpha} \lambda^{(n)}|^2 J(u_{\text{out}}^{(n-1)})^2 - J(u_{\text{out}}^{(n)})^2}{(\lambda^{(n)})^2}. \quad (4.15c)$$

4.2.3 Konvergenzverhalten bei Beispielanwendungen

Um die mit dem in Abschnitt 4.2.2 vorgestellten Verfahren zur Projektionsstärkensteuerung mögliche Verbesserung des Konvergenzverhaltens zu demonstrieren, setzen wir dieses Verfahren innerhalb der SNR-Optimierung für die beiden am Anfang von Abschnitt 4.2 vorgestellten Beispielanwendungen ein. Das resultierende Konvergenzverhalten wird verglichen mit dem IFTA sowie einem Algorithmus, der die optimale Projektionsstärke vor jeder Iteration über Suche innerhalb der Kurve $v_{\text{conv}}(\lambda)$ ermittelt. Dabei wird die Iteration bei den drei betrachteten Verfahren mit jeweils gleicher Startverteilung durchgeführt. Der letztgenannte Algorithmus führt zwar zu einer bestmöglichen Konvergenzrate innerhalb des betrachteten Input-Output-Schemas, ist jedoch auf Grund des hohen Rechenaufwandes zur Ermittlung von λ_{opt} nur von theoretischem Interesse.

Abbildung 4.2(a) zeigt die Entwicklung des Signal-Rausch-Verhältnisses für das Design eines regulären 1-zu-3-Strahlteilers. Im Vergleich zum IFTA kann für den Algorithmus nach Ab-

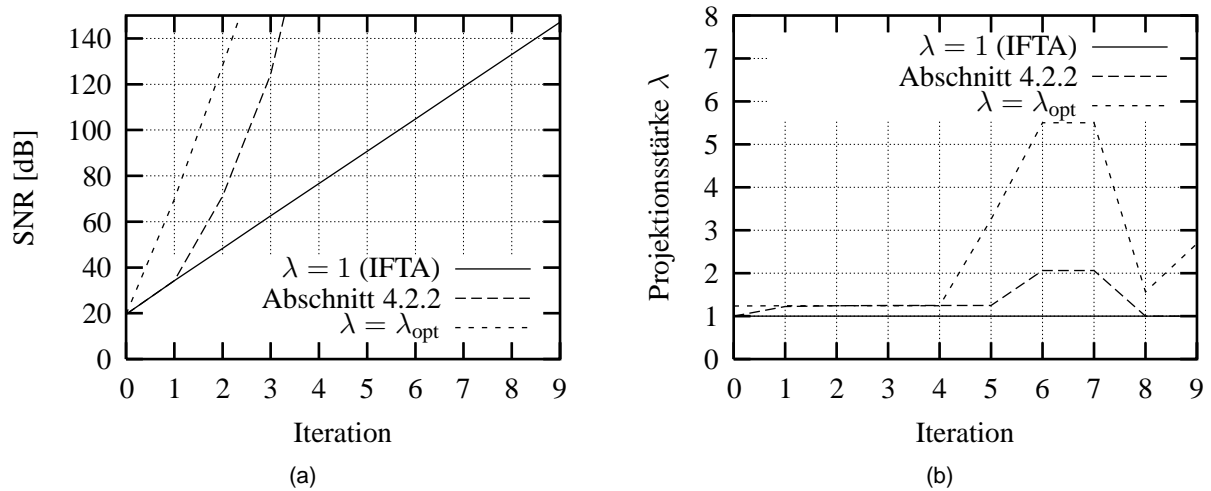


Abbildung 4.2: Verlauf von SNR (a) und Projektionsstärke (b) während der Iteration zur SNR-Optimierung beim Design eines regulären 1-zu-3-Strahlteilers für IFTA (durchgängige Linie), für den Algorithmus mit Projektionsstärkenveränderung entsprechend Abschnitt 4.2.2 (gestrichelte Linie) sowie für einen Algorithmus, der die optimale Projektionsstärke vor jeder Iteration über Suche innerhalb der Kurve $v_{conv}(\lambda)$ ermittelt (gepunktete Linie)

schnitt 4.2.2 ein deutlich schnelleres Ansteigen des SNRs beobachtet werden. Im Vergleich zu dem in der gepunkteten Kurve gezeigten maximal möglichen SNR-Anstieg ergibt sich für den Algorithmus nach Abschnitt 4.2.2 im Wesentlichen eine Verzögerung von einer Iteration, welche darauf zurückzuführen ist, dass in der ersten Iteration die Projektionsstärke $\lambda^{(0)} = 1$ verwendet wird.

Aus dem Verlauf der gepunkteten Kurve in Abbildung 4.2(b) wird ersichtlich, dass die optimale Projektionsstärke während der ersten fünf Iterationen annähernd konstant ist mit dem Wert $\lambda_{opt} \approx 1.25$. Dieses Verhalten, welches sich auch schon im Vergleich der Abbildungen 4.1(a) und 4.1(b) zeigte, stellt eine spezielle Eigenschaft dieses Designproblems dar. Für alle drei betrachteten Designalgorithmen wird eine Konversionseffizienz im Bereich $92.50\% \leq \eta_{conv} \leq 92.52\%$ erreicht, was sehr nah an die Effizienz der optimalen Lösung dieses Designproblems [88] von 92.56% heranreicht. Im Verlauf der SNR-Optimierung ist jeweils ein leichter Abfall der Konversionseffizienz zu beobachten, wobei sich das Verhalten für die drei betrachteten Ansätze ähnelt.

Da der zum Design von eindimensionalen Strahlteilern notwendige Rechenaufwand ohnehin in der Regel sehr gering ist, ergibt sich durch den Einsatz des Verfahrens nach Abschnitt 4.2.2 bei dem ersten Beispiel noch kein nennenswerter praktischer Nutzen. Für das zweite betrachtete Beispiel, bei dem mit 1000×1000 Abtastwerten gerechnet wird, ergibt sich jedoch eine andere Situation. Da bei diesem Beispiel der Aufwand zur Ermittlung der optimalen Projektionsstärke schon relativ hoch ist, vergleichen wir in Abbildung 4.3 nur IFTA und Algorithmus nach Abschnitt 4.2.2 miteinander. In Abbildung 4.3(a) zeigt sich, dass für den Algorithmus nach Abschnitt 4.2.2 zur Erreichung eines bestimmten SNR-Werts nur etwa die halbe Anzahl von Iterationen notwendig ist. Da der zusätzliche Rechenaufwand zur Ermittlung der Projektionsstärke bei großen Feldern klein wird gegen den Rechenaufwand der Propagationsschritte, zeigt sich, dass in der Praxis durch den Algorithmus nach Abschnitt 4.2.2 eine deutliche Verbesserung der Konvergenzgeschwindigkeit erreicht werden kann.

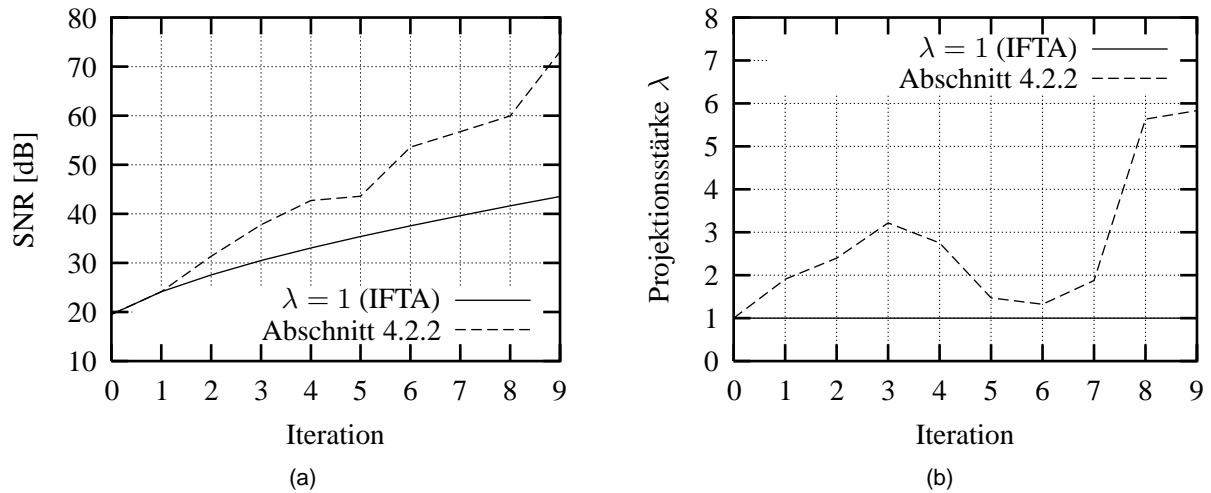


Abbildung 4.3: Verlauf von SNR (a) und Projektionsstärke (b) während der Iteration zur SNR-Optimierung beim Design eines Strahlteilers zur Erzeugung eines regulären Feldes von 300×300 Signalordnungen für den IFTA (durchgängige Linie) und den Algorithmus mit Projektionsstärkenveränderung entsprechend Abschnitt 4.2.2 (gestrichelte Linie)

4.2.4 Sicherstellung der Konvergenz des Algorithmus

Im Fall $\lambda = 1$ stimmt der durch die Gleichungen (2.76) und (4.1) gegebene Input-Output-Algorithmus mit dem IFTA überein, womit gemäß Abschnitt 5.1.1 die SDE-Konvergenz für diesen Fall sichergestellt ist. Obwohl die Wahl von λ entsprechend Abschnitt 4.2.2 für die betrachtete Art von Designproblemen im Allgemeinen zu einer Verbesserung der Konvergenzeigenschaften im Vergleich zum Fall $\lambda = 1$ führt, kann jedoch der Fall $v_{\text{conv}}(\lambda) < 0$ im Allgemeinen nicht ausgeschlossen werden. Für konvexe Einschränkungsmengen wurde von Youla und Webb die Konvergenz für $0 < \lambda < 2$ nachgewiesen [43]. Da jedoch insbesondere die Transmissionseinschränkung typischerweise nicht konvex ist (vgl. Abschnitt 2.4.1) und eine Einschränkung auf $\lambda < 2$ nicht wünschenswert erscheint (vgl. Abbildung 4.1(d)), kann die Konvergenz des Algorithmus nur durch eine Modifikation des Schemas (4.1) zur Ermittlung des Steuerfelds sichergestellt werden.

Wir prüfen hierzu nach jeder Iteration n die Erfüllung der Konvergenzbedingung $v_{\text{conv}}^{(n)} > 0$, welche äquivalent ist mit der Bedingung $J(u_{\text{out}}^{(n)}) < J(u_{\text{out}}^{(n-1)})$. Falls sie erfüllt ist, wird die entsprechend Abschnitt 4.2.2 ermittelte Projektionsstärke $\lambda^{(n+1)}$ verwendet. Anderenfalls wird Iteration n verworfen und mit $\lambda = 1$ basierend auf dem Ergebnis der vorangegangenen Iteration fortgefahren. In Abwandlung zu Gleichung (4.1) ergibt sich somit

$$u_{\text{dr}}^{(n+1)} = \begin{cases} u_{\text{out}}^{(n)} + \lambda^{(n+1)} (\Pi_{\text{sig}} u_{\text{out}}^{(n)} - u_{\text{out}}^{(n)}) & J(u_{\text{out}}^{(n)}) < J(u_{\text{out}}^{(n-1)}) \\ \Pi_{\text{sig}} u_{\text{out}}^{(n-1)} & \text{sonst} \end{cases} \quad (4.16)$$

4.3 Überwindung von durch den Quantisierungsoperator verursachten Stagnationen

Nach einer Betrachtung des Softquantisierungsverfahrens in Abschnitt 4.3.1 entwickeln wir in Abschnitt 4.3.2 ein für quantisierte Transmissionen geeignetes Steuerungsverfahren für die Projekti-

onsstärke. In Abschnitt 4.3.3 wird dieses Verfahren anhand einer Bespielanwendung demonstriert.

4.3.1 Eigenschaften des Softquantisierungsverfahrens

Eine weit verbreitete Methode zum wellenoptischen Design quantisierter Transmissionsfunktionen ist das in Abschnitt 3.1.4 eingeführte Softquantisierungsverfahren. Obwohl dieses Verfahren in vielen Fällen ausreichend gute Ergebnisse liefert, besitzt es einige negative Eigenschaften, welche im Folgenden diskutiert werden.

Da während der Iterationen der Softquantisierungsparameter schrittweise von $\lambda = 0$ (keine Quantisierung) auf $\lambda = 1$ (vollständige Quantisierung) erhöht wird, muss die Anzahl $n_{\text{softquant}}$ der dafür vorgesehenen Iterationen vorher festgelegt werden. Wegen der mit dem Quantisierungsoperator $\Pi_{\text{quant},Q} \equiv \Pi_{\text{softquant},Q}^1$ verbundenen Stagnationsprobleme kann durch weitere Iterationen mit $\lambda = 1$ typischerweise keine Verbesserung des Ergebnisses erzielt werden. Wie in Referenz [36] gezeigt wird, führen größere Werte von $n_{\text{softquant}}$ nur tendenziell zu besseren Ergebnissen. Die wünschenswerte Eigenschaft von Optimierungsverfahren, dass mit einer Erhöhung des Rechenaufwands eine Verbesserung des Ergebnisses erzielt werden kann, ist somit für das Softquantisierungsverfahren nicht erfüllt.

Als weitere negative Eigenschaft stellt sich die relativ starke Abhängigkeit des Ergebnisses der Softquantisierung von der gewählten Startverteilung dar. Zur Illustration hiervon betrachten wir als Beispiel die Softquantisierung für reguläre 1-zu-3-Strahlteiler auf vier Quantisierungsstufen. Die schwarzen Balken in Abbildung 4.4 zeigen die Verteilung der bei der Softquantisierung mit 100 verschiedenen Startverteilungen erzielten Konversionseffizienzen. Dabei wurde Designschemas aus Abbildung 3.2 mit 10 Iterationen für die Signalphasensynthese, 50 Iterationen für die die SNR-Optimierung und 100 Iterationen für die Softquantisierung angewendet.

Da typischerweise das Ziel besteht, gute Ergebnisse sowohl für Effizienz als auch für das Signal-Rausch-Verhältnis zu erzielen, werden in den Histogrammen von Abbildung 4.4 untere Schwellwerte für das SNR vorgegeben: 50 dB für Abbildung 4.4(a) und 30 dB für Abbildung 4.4(b). Designs, deren SNR unter dem betrachteten Schwellwert liegt, werden in den jeweils ersten Balken erfasst, während die weiteren Balken die Verteilung der erreichten Konversionseffizienzen für Designs mit einem SNR oberhalb des Schwellwertes darstellen.

Wie aus der Ergebnisverteilung in Abbildung 4.4 zu sehen ist, kann über die Betrachtung einer hohen Anzahl von Startverteilung mit der Softquantisierung in vielen Fällen ein gutes Ergebnis erzielt werden. Dennoch ist ein in einem stärkerem Maße deterministisches Designverfahren erstrebenswert.

Eine weitere Motivation zur Entwicklung eines verbesserten Verfahrens ist durch die mit der Softquantisierung verbundene Begrenzung des erzielbaren SNRs gegeben: Während in dem zuletzt betrachteten Beispiel in einigen Fällen SNR-Werte über 50 dB erreicht werden konnten, ist es für Strahlteilungsprobleme mit einer höheren Anzahl von Signalordnungen teilweise schwierig, mit Hilfe der Softquantisierung überhaupt Lösungen oberhalb von 30 dB zu finden.

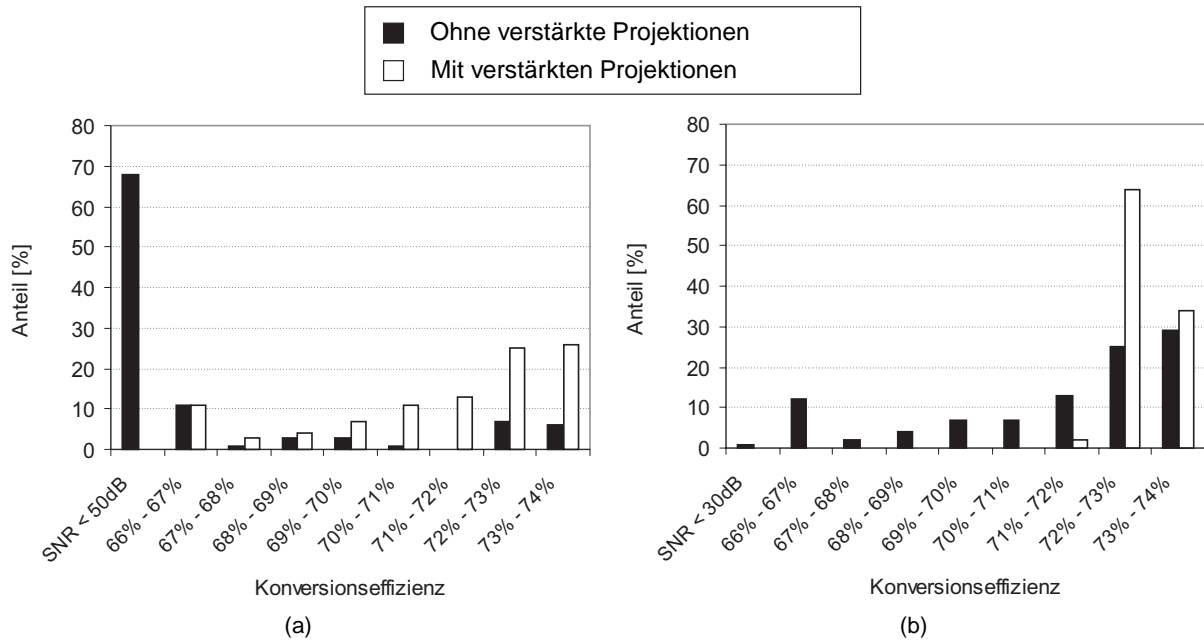


Abbildung 4.4: Verteilung der erreichten Konversionseffizienzen für Designs mit einem Signal-Rausch-Verhältnis oberhalb von 50 dB (a) bzw. 30 dB für die über die Abbildungen 3.2 und 4.5 beschriebenen Designverfahren ohne bzw. mit verstärkten Projektionen. Zur Ermittlung der Ergebnisverteilung wurden jeweils 100 Designs mit zufällig gewählten Startverteilungen durchgeführt.

4.3.2 Anwendung verstärkter Projektionen

Zur Überwindung der durch Quantisierungsoperatoren hervorgerufenen Stagnationen greifen wir wiederum auf den in Abschnitt 4.1.1 entwickelten Grundansatz zurück. Die im Folgenden beschriebene Iterationsphase mit verstärkten Projektionen kann – wie in Abbildung 4.5 gezeigt wird – im Anschluss an die Softquantisierung durchgeführt werden. Insbesondere bei einer relativ hohen Anzahl von Quantisierungswerten ($m \gtrsim 8$) kann der Softquantisierungsschritt jedoch auch übersprungen werden.

Zur Entwicklung eines Verfahrens für die Steuerung der Projektionsstärke λ während der Iterationen gehen wir ähnlich vor wie schon in Abschnitt 4.2.1 und betrachten zunächst die Abhängigkeit der Konvergenzrate von der Projektionsstärke. Abbildung 4.6 zeigt einen typischen Verlauf der Kurve $v_{\text{conv}}^{(n)}(\lambda)$, welche entlang der x-Achse in drei Bereiche unterteilt werden kann: (1) den Stagnationsbereich mit $v_{\text{conv}}(\lambda) = 0$ (im Beispiel für $0 \leq \lambda \leq 63.390$); (2) den Konvergenzbereich mit $v_{\text{conv}}(\lambda) > 0$ (im Beispiel für $63.390 \leq \lambda \leq 63.404$) und (3) den Divergenzbereich mit $v_{\text{conv}}(\lambda) < 0$ (im Beispiel für $63.404 \leq \lambda$). Durch Iterationen mit Projektionsstärken innerhalb des Stagnationsbereiches wird die Transmissionsfunktion nicht verändert. Dies trifft in unserem Beispiel auf herkömmliche IFTA-Iterationen zu, da $\lambda = 1$ zum Stagnationsbereich gehört. Für zunehmende Werte für λ außerhalb des Stagnationsbereiches treten mehr und mehr Änderungen an der Transmission auf. Die in der Kurve von Abbildung 4.6 beobachtbaren Sprünge entsprechen jeweils der Änderung von resultierenden Quantisierungswerten an einzelnen Positionen der Transmissionsfunktion.

Die Steuerung der Projektionsstärke wird erschwert durch den Umstand, dass sich die Lage des Konvergenzbereiches durch Iterationen außerhalb des Stagnationsbereiches verändern. Da der Konvergenzbereich zudem typischerweise sehr schmal im Vergleich zum Stagnationsbereich ist,

(1)	<i>Initialisierung:</i> Setze $t^{(0)}$ durch Rückwärtspropagation eines Signalfelds mit zufällig gewählten Phasen der Signalordnungen
(2)	<i>Signalphasensynthese:</i> Iterationen mit Π_{phase} und Π_{P} , d. h. ohne Amplitudenfreiheit und Quantisierung
(3)	<i>SNR-Optimierung:</i> Iterationen mit Π_{phase} and Π_{PAS} , d. h. ohne Quantisierung
(4)	<i>Softquantisierung:</i> Iterationen mit $\Pi_{\text{softquant}, Q}^{\lambda}$ für linear ansteigende Projektionsstärke λ sowie Π_{PAS}
(5)	<i>Verstärkte Projektionen:</i> Iterationen mit $\Pi_{\text{quant}, Q}$ und verstärkten Signalprojektionen $\Pi_{\text{sig}}^{\lambda}$

Abbildung 4.5: Schritte beim Design quantisierter Strahlteiler unter Ausnutzung verstärkter Projektionen

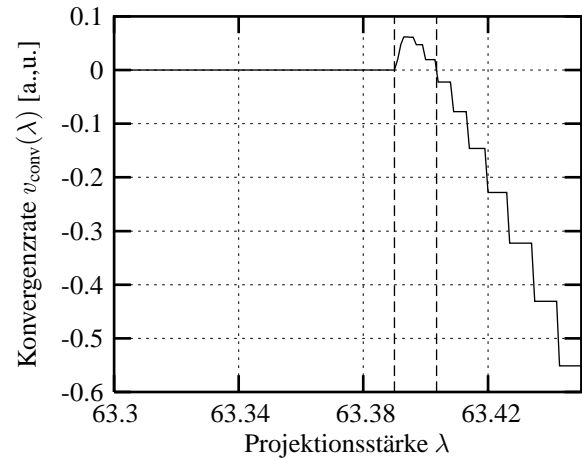


Abbildung 4.6: Abhängigkeit der Konvergenzrate von der Projektionsstärke für einen typischen Iterationszustand während des Design eines regulären 1-zu-3-Strahlteilers mit $Q = 4$

erscheint es nicht länger möglich, wie in Abschnitt 4.2.2 ein Modell für die Kurve $v_{\text{conv}}^{(n)}(\lambda)$ aufzustellen, wobei ja das Ziel bestand, eine Approximation für λ_{opt} über dieses Modell zu erhalten. Leichte Abweichungen von λ_{opt} würden häufig zu Iterationen im Stagnations- oder Divergenzbereich führen.

Andererseits erscheint es jedoch nicht sehr schwer, einen Algorithmus mit verstärkten Projektionen zu konstruieren, welcher eine höhere Konvergenzrate v_{conv} als IFTA mit $\lambda = 1$ besitzt, da – wie schon erwähnt – der IFTA typischerweise nach wenigen Iterationen stagniert, was $v_{\text{conv}} = 0$ bedeutet. Wir konstruieren im Folgenden einen Algorithmus zur Steuerung der Projektionsstärke, welcher die Strategie der binären Suche [89] zum Auffinden des Konvergenzbereichs benutzt.

Zur Speicherung des Wissens über die Lage des Konvergenzbereichs benutzt der Algorithmus zwei Variablen λ_l und λ_u , welche untere bzw. obere Schranken für die im Allgemeinen unbekannt Grenzen des Konvergenzbereichs darstellen. In jeder Iteration n ist der Konvergenzbereich demnach ein Unterbereich von $[\lambda_l^{(n)}, \lambda_u^{(n)}]$. Da vor Beginn der Iterationen kein Wissen über die Lage des Konvergenzbereichs vorhanden ist, wird die Iteration mit den Werten $\lambda^{(0)} = 1$, $\lambda_l^{(0)} = 0$ und $\lambda_u^{(0)} = \infty$ gestartet. Nach jeder Iteration n wird geprüft, zu welchem der drei Bereiche die vorher gewählte Projektionsstärke $\lambda^{(n)}$ gehört. Falls $\lambda^{(n)}$ zum Stagnationsbereich gehört, kennen wir eine neue untere Schranke für den Konvergenzbereich. Im Gegensatz hierzu ergibt sich für Iterationen im Divergenzbereich eine neue obere Schranke, wobei derartige Iterationen rückgängig gemacht werden müssen, um eine Konvergenz des Algorithmus sicherzustellen.

Iterationen innerhalb des Konvergenzbereichs stellen das eigentliche Ziel dar, wobei in diesem Fall jedoch das Wissen über die Lage des Konvergenzbereichs verloren geht. Für die nächste Iteration $(n+1)$ benutzen wir die Projektionsstärke $\lambda^{(n+1)} = (\lambda_l^{(n+1)} + \lambda_u^{(n+1)})/2$, sofern $\lambda_u^{(n+1)} < \infty$. Falls noch keine endliche obere Schranke $\lambda_u^{(n+1)}$ bekannt ist, wird die Projektionsstärke um einen konstanten Faktor $\lambda_{\text{incr}} > 1$ erhöht.

Dieser Algorithmus wird durch die folgenden Formeln zusammengefasst:

$$\lambda_l^{(n+1)} = \begin{cases} 0 & \text{falls } \lambda^{(n)} \text{ im Konvergenzbereich} \\ \max(\lambda_l^{(n)}, \lambda^{(n)}) & \text{falls } \lambda^{(n)} \text{ im Stagnationsbereich} \\ \lambda_l^{(n)} & \text{sonst} \end{cases} \quad (4.17a)$$

$$\lambda_u^{(n+1)} = \begin{cases} \infty & \text{falls } \lambda^{(n)} \text{ im Konvergenzbereich} \\ \lambda_u^{(n)} & \text{falls } \lambda^{(n)} \text{ im Stagnationsbereich} \\ \min(\lambda_u^{(n)}, \lambda^{(n)}) & \text{sonst} \end{cases} \quad (4.17b)$$

$$\lambda^{(n+1)} = \begin{cases} \lambda^{(n)} & \text{falls } \lambda^{(n)} \text{ im Konvergenzbereich} \\ (\lambda_l^{(n+1)} + \lambda_u^{(n+1)})/2 & \text{falls } \lambda^{(n)} \text{ außerhalb des Konvergenz-} \\ & \text{bereichs und } \lambda_u^{(n+1)} \neq \infty \\ \lambda_{\text{incr}} \lambda^{(n)} & \text{sonst} \end{cases} \quad (4.17c)$$

Die Wahl von λ_{incr} in Gleichung (4.17c) ist nach unserer Erfahrung nicht sehr kritisch: Die Konstante λ_{incr} muss in jedem Fall größer als 1 gewählt werden, wobei kleine Werte für λ_{incr} das Auffinden des Konvergenzbereichs aus dem Stagnationsbereich heraus verlangsamen und große Werte zu einem häufigeren Auftreten des Divergenzbereiches führen. Der im Folgenden verwendete Wert $\lambda_{\text{incr}} = 1.3$ stellt einen geeigneten Kompromiss dar.

Für den beschriebenen Iterationsprozess können verschiedene Abbruchkriterien gewählt werden. Theoretisch kann die Iteration solange fortgesetzt werden, bis $\lambda_l^{(n)}$ und $\lambda_u^{(n)}$ innerhalb der begrenzten Rechengenauigkeit gleich sind, d. h. so lange, bis kein Konvergenzbereich mehr gefunden werden kann.

Von den ersten 250 Iterationen eines typischen Iterationsprozesses fielen 58 % in den Konvergenzbereich und nur 5 % in den Stagnationsbereich. Mit dem Verlauf der Iterationen nimmt der Anteil der Iterationen innerhalb des Konvergenzbereichs jedoch typischerweise ab.

4.3.3 Konvergenzverhalten bei Beispielanwendungen

Bei der Durchführung von 100 Designs entsprechend des Ablaufschemas aus Abbildung 4.5 resultierte die in den weißen Balken des Histogramms in Abbildung 4.4(a) dargestellte Ergebnisverteilung. Die Iterationsphase mit verstärkten Projektionen wurde dabei beendet, sobald der SNR-Schwellwert von 50 dB erreicht wurde. Bei dem betrachteten Beispiel war dies stets nach weniger als 100 Iterationen der Fall. Im Vergleich der schraffierten Balken mit den weißen Balken in Abbildung 4.4(a) erkennen wir, dass durch Anwendung der Iterationsphase mit verstärkten Projektionen einerseits stets ein SNR oberhalb von 50 dB erreicht werden konnte und andererseits die Wahrscheinlichkeit der Erzielung einer Lösung mit gleichzeitig hohem SNR und hoher Effizienz deutlich gesteigert werden konnte: Die Wahrscheinlichkeit der Erzielung einer Effizienz oberhalb 72 % bei einem SNR oberhalb von 50 dB konnte von etwa 10 % auf etwa 50 % gesteigert werden.

Falls niedrigere SNR-Werte akzeptabel sind, kann die Wahrscheinlichkeit der Erzielung hoher Effizienzen durch Einschränkung der Skalierungsfreiheit in der Signalprojektion (vgl. Abschnitt 2.4.2) weiter gesteigert werden. Zur Erzielung höchstmöglicher Effizienzen stellt es sich

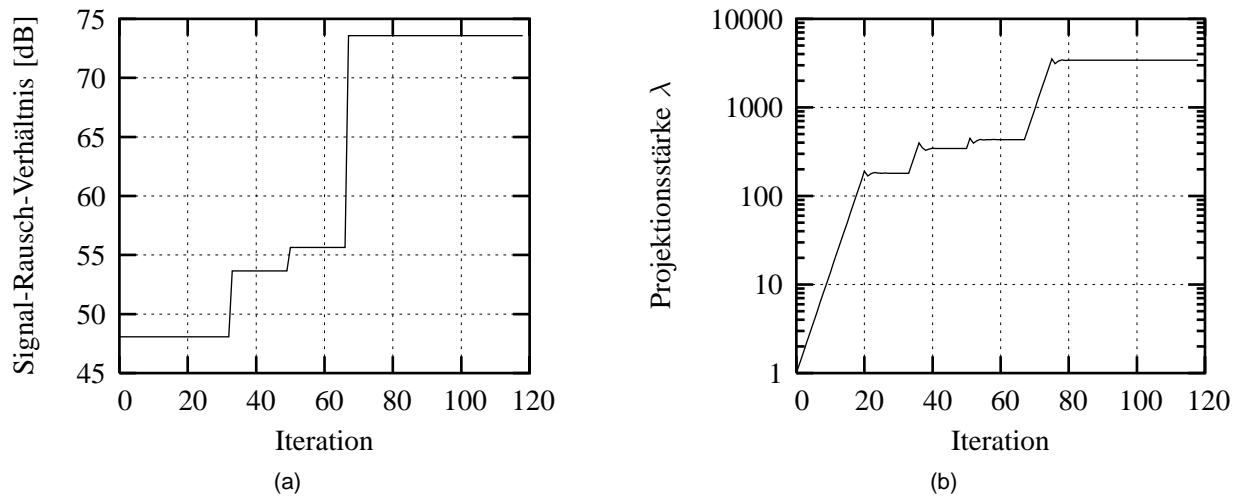


Abbildung 4.7: Verlauf von SNR (a) und Projektionsstärke während der Iterationen des Algorithmus aus Abschnitt 4.3.2 beim Design eines regulären 1-zu-3-Strahlteilers mit $Q = 4$.

im Allgemeinen als günstig heraus, die untere Grenze für den Skalierungsfaktor entsprechend einer Effizienz zu wählen, welche höher als der tatsächlich erreichbare Wert ist. Da sich hiermit die Problemstellung jedoch stärker von einer Konsistenz entfernt (vgl. Abschnitt 2.1.2), ist mit Einbußen hinsichtlich des erzielten SNRs zu rechnen. In dem betrachteten Beispiel führte die Einführung einer unteren Schranke für den Skalierungsfaktor, welche einer Effizienz von 74 % entspricht, zu der in den weißen Balken des Histogramms aus Abbildung 4.4(b) dargestellten Ergebnisverteilung. Für fast alle Startverteilungen wurden hierbei Effizienzen oberhalb von 72 % erreicht, wobei jedoch die erzielten SNRs in vielen Fällen unterhalb von 50 dB blieben.

Die Abbildungen 4.7(a) und 4.7(b) zeigen den Verlauf von SNR und Projektionsstärke während der Iteration des Algorithmus aus Abschnitt 4.3.2 mit einer für das oben betrachtete Beispiel über Softquantisierung ermittelten Startverteilung. Innerhalb der betrachteten 118 Iterationen kann das SNR deutlich gesteigert werden. Danach ist ein Auffinden des Konvergenzbereichs nicht mehr möglich. Die Konversionseffizienz steigt in diesem Beispiel während der Iteration von 73.18 % auf 73.22 % leicht an.

Als ein weiteres Beispiel betrachten wir das Design eines zweidimensionalen Strahlteilers zur Erzeugung eines regulären Feldes von 300×300 Signalordnungen mit $Q = 4$ unter Verwendung von 1000×1000 Abtastwerten. Ein Design über Softquantisierung entsprechend des Schemas von Abbildung 3.1 führte zu dem Ergebnis $\text{SNR} = 15.9 \text{ dB}$ und $\eta_{\text{conv}} = 59.8 \%$. Nehmen wir diese Lösung als Startverteilung für den Algorithmus aus Abschnitt 4.3.2, erhalten wir nach 500 Iterationen bei voller Skalierungsfreiheit die Werte $\text{SNR} = 28.5 \text{ dB}$ und $\eta_{\text{conv}} = 54.7 \%$. Das SNR wurde in diesem Fall demnach auf Kosten der Konversionseffizienz erhöht. Um ein Abfallen der Konversionseffizienz während der Iteration zu vermeiden, ist eine Beschränkung der Skalierungsfreiheit notwendig. Die Einführung einer unteren Grenze für den Skalierungsfaktor, welche einer Effizienz von 65 % entspricht, führte bei erneuter Durchführung von 500 Iterationen beginnend mit dem Ergebnis der Softquantisierung zu $\text{SNR} = 28.6 \text{ dB}$ und $\eta_{\text{conv}} = 60.2 \%$. Es konnte somit sowohl hinsichtlich des SNRs als auch hinsichtlich der Konversionseffizienz eine Verbesserung gegenüber dem Resultat der Softquantisierung erreicht werden.

Kapitel 5

Multifunktionale Transmissionsdesignprobleme

In Abschnitt 2.4 wurde mit dem Iterativen Fouriertransformationsalgorithmus (IFTA) ein Verfahren zur Lösung monofunktionaler Transmissionsdesignprobleme vorgestellt. Daran anknüpfend werden in den Abschnitten 5.1 und 5.2 zwei für multifunktionale Transmissionsdesignprobleme geeignete Ansätze vorgestellt. In Abschnitt 5.3 werden die Konvergenzeigenschaften beider Verfahren anhand einer Beispielanwendung untersucht und diskutiert. Weitere Beispielanwendungen für multifunktionale Designverfahren werden in Kapitel 6 betrachtet. Den Bezug der vorgestellten Verfahren zu früheren Arbeiten stellen wir in Abschnitt 5.4 dar.

5.1 Paralleler Multiprojektionsalgorithmus

Bei der Spezifikation von Transmissionsdesignproblemen laut Abschnitt 2.3.1 erfolgt eine Beschreibung der Transmissions- und Signaleinschränkungen durch Einschränkungsmengen $\mathcal{M}_{\text{tran}}$ bzw. $\mathcal{M}_{\text{sig},\beta}$ ($\beta = 1 \dots B$). Beispiele für häufig in diesem Zusammenhang verwendete Einschränkungsmengen $\mathcal{M}_{\text{tran}}$ und $\mathcal{M}_{\text{sig},\beta}$ sowie zugehörige Projektionsoperatoren Π_{tran} bzw. $\Pi_{\text{sig},\beta}$ können in den Abschnitten 2.4.1 und 2.4.2 gefunden werden. Diese Einschränkungsmengen und Projektionsoperatoren eignen sich jedoch nicht für die Anwendung des in Abschnitt 2.1.6 vorgestellten parallelen Multiprojektionsalgorithmus (PMPA), da die Eingangsfelder $u_{\text{in},\beta}$ sowie die Propagationsoperatoren P_β in den Einschränkungsmengen keine Berücksichtigung finden. Wie die folgende Darstellung zeigt, können die gegebenen Einschränkungsmengen jedoch hinsichtlich einer Anwendbarkeit des PMPAs umformuliert werden.

In Abschnitt 5.1.1 betrachten wir Designprobleme, bei denen für sämtliche optischen Funktionen dasselbe Beleuchtungsfeld spezifiziert ist. Diese Einschränkung wird in dem darauf folgenden Abschnitt 5.1.2 aufgehoben.

5.1.1 Designprobleme mit in allen optischen Funktionen identischen Eingangsfeldern

In diesem Abschnitt betrachten wir multifunktionale Designprobleme, bei denen die den einzelnen optischen Funktionen zugeordneten Eingangsfelder $u_{\text{in},\beta}$ ($\beta = 1 \dots B$) jeweils gleich sind, d. h.

$$u_{\text{in},1} = u_{\text{in},2} = \dots = u_{\text{in},B} = u_{\text{in}}. \quad (5.1)$$

In diesem Fall können die Einschränkungsmengen \mathcal{M}_j ($j = 1 \dots N$) analog zu Gleichung (2.65) definiert werden:

$$\mathcal{M}_\beta = \{P_\beta^{-1} u : u \in \mathcal{M}_{\text{sig},\beta}\} \quad (\beta = 1 \dots B), \quad (5.2a)$$

$$\mathcal{M}_N = \{u_{\text{in}} t : t \in \mathcal{M}_{\text{tran}}\}. \quad (5.2b)$$

Als zugehörige Projektionsoperatoren definieren wir in Analogie zu den Gleichungen (2.66a) und (2.66b)

$$\Pi_\beta u = P_\beta^{-1} \Pi_{\text{sig},\beta} P_\beta u \quad (\beta = 1 \dots B) \quad (5.3a)$$

$$\Pi_N u = u_{\text{in}} \Pi_{\text{tran}} u_{\text{in}}^{-1} u. \quad (5.3b)$$

Falls die in den Sätzen 2.67 und 2.68 genannten Voraussetzungen erfüllt sind, resultiert somit bei Anwendung des PMPAs mit den Projektionsoperatoren (5.3) entsprechend Satz 2.22 eine SSDE-Konvergenz.

5.1.2 Verallgemeinerung auf beliebige Eingangsfelder

Im vorangegangenen Abschnitt erfolgte die Berücksichtigung der Eingangsfelder entsprechend Gleichung (5.2b) innerhalb der Menge \mathcal{M}_N . Die Definition der Menge \mathcal{M}_N entsprechend Gleichung (5.2b) ist jedoch an die Bedingung (5.1) geknüpft, da hierbei nur *ein* Eingangsfeld u_{in} berücksichtigt werden kann. Aufgrund der im Zusammenhang mit Gleichung (2.73a) getroffenen Diskussion können die Eingangsfelder $u_{\text{in},\beta}$ auch nicht ohne weiteres innerhalb der Einschränkungsmengen \mathcal{M}_β ($\beta = 1 \dots B$) erfasst werden, da in diesem Fall die effektive Durchführung der zugehörigen Projektion im Allgemeinen nicht möglich ist. Mit Hilfe der im Beweis zu Satz 2.22 verwendeten äquivalenten Formulierung des PMPAs im Produktraum \mathcal{H}^N ist eine Übertragung des Konzepts des PMPAs auf den in diesem Abschnitt betrachteten allgemeinen Fall multifunktionaler Transmissionsdesignprobleme jedoch möglich, wie im Folgenden gezeigt wird.

Wir betrachten hierzu in Abwandlung zu den Gleichungen (5.2) die Einschränkungsmengen

$$\mathcal{M}_\beta = \{P_\beta^{-1} u : u \in \mathcal{M}_{\text{sig},\beta}\} \quad (\beta = 1 \dots B), \quad (5.4a)$$

$$\mathcal{M}_N = \mathcal{M}_{\text{tran}} \quad (5.4b)$$

mit den in Analogie zu den Gleichungen (5.3) gegebenen Projektionsoperatoren

$$\Pi_\beta u = P_\beta^{-1} \Pi_{\text{sig},\beta} P_\beta u \quad (\beta = 1 \dots B), \quad (5.5a)$$

$$\Pi_N u = \Pi_{\text{tran}} u. \quad (5.5b)$$

In Abwandlung der Gleichungen (2.26) und (2.28) können hieraus unter Berücksichtigung der Eingangsfelder $u_{\text{in},\beta}$ die Mengen $\widetilde{\mathcal{M}}_1 \subseteq \mathcal{H}^N$ bzw. $\widetilde{\mathcal{M}}_2 \subseteq \mathcal{H}^N$ definiert werden:

$$\widetilde{\mathcal{M}}_1 = \{\mathbf{u} = (u_1, u_2, \dots, u_N) : u_j \in \mathcal{M}_j, j = 1 \dots N\}, \quad (5.6a)$$

$$\widetilde{\mathcal{M}}_2 = \{\mathbf{u} = (u_{\text{in},1}t_{\mathbf{u}}, u_{\text{in},2}t_{\mathbf{u}}, \dots, u_{\text{in},B}t_{\mathbf{u}}, t_{\mathbf{u}}) : t_{\mathbf{u}} \in \mathcal{H}\}. \quad (5.6b)$$

Die Mengen $\widetilde{\mathcal{M}}_1$ und $\widetilde{\mathcal{M}}_2$ entsprechend dieser Definition spiegeln die gegebenen Einschränkungen korrekt wider und ermöglichen – wie im Folgenden gezeigt wird – die Definition von effektiv durchführbaren zugehörigen Projektionsoperatoren. Für die Projektion auf $\widetilde{\mathcal{M}}_1$ ist der zugehörige Projektionsoperator $\widetilde{\Pi}_1$ durch Gleichung (2.24b) gegeben.

Zur Projektion auf $\widetilde{\mathcal{M}}_2$ betrachten wir den folgenden Operator:

$$\widetilde{\Pi}_2 \mathbf{u} = \mathbf{v} = (u_{\text{in},1}t_{\mathbf{v}}, u_{\text{in},2}t_{\mathbf{v}}, \dots, u_{\text{in},B}t_{\mathbf{v}}, t_{\mathbf{v}}) \quad \text{mit} \quad (5.7a)$$

$$t_{\mathbf{v}} = \frac{\sum_{j=1}^N u_{\text{in},j}^* u_j}{\sum_{j=1}^N |u_{\text{in},j}|^2}, \quad (5.7b)$$

wobei in Gleichung (5.7b) zur Vereinfachung der Notation $u_{\text{in},N} \equiv 1$ gesetzt wurde.

Satz 5.8 *Der über die Gleichungen (5.7) definierte Operator $\widetilde{\Pi}_2$ bewirkt eine Projektion auf die durch Gleichung (5.6b) gegebene Menge $\widetilde{\mathcal{M}}_2$.*

Mit Satz 5.8, dessen Beweis in Anhang A.4.1 dokumentiert ist, folgt für den Algorithmus

$$\mathbf{u}^{(n)} = \widetilde{\Pi}_2 \widetilde{\Pi}_1 \mathbf{u}^{(n-1)} \quad (5.9)$$

analog zu den Betrachtungen in Abschnitt 2.1.6 die Konvergenzeigenschaft

$$\mathbf{J}(\mathbf{u}^{(n)}) \leq \mathbf{J}(\mathbf{u}^{(n-1)}). \quad (5.10)$$

Motiviert durch die in Gleichung (5.6b) gegebene Darstellung der Elemente $\mathbf{u}^{(n)} \in \widetilde{\mathcal{M}}_2$ definieren wir in Analogie zu Gleichung (2.71) eine im Vergleich zu Gleichung (2.12) modifizierte Form des SSDEs mit

$$\hat{J}_{\text{in}}(t_{\mathbf{u}}^{(n)})^2 = \mathbf{J}^2((u_{\text{in},1}t_{\mathbf{u}}^{(n)}, u_{\text{in},2}t_{\mathbf{u}}^{(n)}, \dots, u_{\text{in},B}t_{\mathbf{u}}^{(n)}, t_{\mathbf{u}}^{(n)})), \quad (5.11)$$

wodurch sich die Konvergenzeigenschaft (5.10) darstellen lässt als

$$\hat{J}_{\text{in}}(t_{\mathbf{u}}^{(n)})^2 \leq \hat{J}_{\text{in}}(t_{\mathbf{u}}^{(n-1)})^2. \quad (5.12)$$

Zur Veranschaulichung dieser Konvergenzeigenschaft soll im Folgenden $\hat{J}_{\text{in}}(t_{\mathbf{u}}^{(n)})$ über die Propagations- und Projektionsoperatoren P_{β} bzw. $\Pi_{\text{sig},\beta}$ und Π_{tran} ausgedrückt werden. Wegen $\mathbf{u}^{(n)} \in \widetilde{\mathcal{M}}_2$ folgt

$$\begin{aligned} \hat{J}_{\text{in}}(t_{\mathbf{u}}^{(n)})^2 &= \mathbf{J}^2(\mathbf{u}^{(n)}) \\ &= \mathbf{d}^2(\mathbf{u}^{(n)}, \widetilde{\Pi}_1 \mathbf{u}^{(n)}) \\ &= \sum_{j=1}^N \|u_{\text{in},j}t_{\mathbf{u}}^{(n)} - \Pi_j u_{\text{in},j}t_{\mathbf{u}}^{(n)}\|^2. \end{aligned} \quad (5.13)$$

Durch Einsetzen der Gleichungen (5.5) resultiert hieraus unter Verwendung von Gleichung (2.39a)

$$\hat{J}_{\text{in}}(t_{\mathbf{u}}^{(n)})^2 = \hat{J}_{\text{sig}}(t_{\mathbf{u}}^{(n)})^2 + \hat{J}_{\text{tran}}(t_{\mathbf{u}}^{(n)})^2 \quad \text{mit} \quad (5.14a)$$

$$\begin{aligned} \hat{J}_{\text{sig}}(t_{\mathbf{u}}^{(n)})^2 &= \sum_{\beta=1}^B \| u_{\text{in},\beta} t_{\mathbf{u}}^{(n)} - P_{\beta}^{-1} \Pi_{\text{sig},\beta} P_{\beta} u_{\text{in},\beta} t_{\mathbf{u}}^{(n)} \|^2, \\ &= \sum_{\beta=1}^B \gamma_{\beta}^{-1} \| P_{\beta} u_{\text{in},\beta} t_{\mathbf{u}}^{(n)} - \Pi_{\text{sig},\beta} P_{\beta} u_{\text{in},\beta} t_{\mathbf{u}}^{(n)} \|^2 \quad \text{und} \end{aligned} \quad (5.14b)$$

$$\hat{J}_{\text{tran}}(t_{\mathbf{u}}^{(n)})^2 = \| t_{\mathbf{u}}^{(n)} - \Pi_{\text{tran}} t_{\mathbf{u}}^{(n)} \|^2. \quad (5.14c)$$

Die Anwendbarkeit des PMPAs für beliebige multifunktionale Transmissionsdesignprobleme wurde somit nachgewiesen und seine Konvergenzeigenschaften untersucht. Im folgenden Abschnitt 5.2 wird ein hierzu alternativer Projektionsalgorithmus betrachtet, welcher – wie die folgende Diskussion zeigt – Vorteile gegenüber dem PMPA aufweist.

5.2 Paralleler Transmissionsdesignalgorithmus

Der *Parallele Transmissionsdesignalgorithmus* (PTDA) stellt eine Alternative gegenüber dem im vorangegangenen Abschnitt betrachteten PMPA dar, wobei die Motivation zur Konstruktion dieses alternativen Verfahrens in Abschnitt 5.2.1 gegeben wird. Bei der Darstellung des Grundprinzips des PTDA in Abschnitt 5.2.2 zeigt sich, dass hierbei der Algorithmus nicht wie bisher für beliebige Transmissionseinschränkungen allgemein formuliert werden kann, sofern eine Konvergenzeigenschaft erfüllt sein soll. Daher werden in den Abschnitten 5.2.3 – 5.2.5 die in der Praxis wichtigsten Typen von Transmissionseinschränkungen separat betrachtet.

5.2.1 Motivation

Unter den $N = B + 1$ Einschränkungen eines multifunktionalen Transmissionsdesignproblems nimmt die Transmissionseinschränkung eine Sonderstellung ein, da ihre Erfüllung im Allgemeinen eine Voraussetzung für die Überführbarkeit der Transmissionsfunktion in ein real herstellbares optische Element darstellt. Demnach muss für eine Lösung des Designproblems die Transmissionseinschränkung in jedem Fall erfüllt sein. Im Gegensatz hierzu ist die Erfüllung der den optischen Funktionen zugeordneten Einschränkungen laut Abschnitt 2.1.2 jedoch im allgemeinen Fall inkonsistenter Problemstellungen nicht möglich.

Bei monofunktionalen Transmissionsdesignproblemen ($B = 1$) kann die Erfüllung der Transmissionseinschränkung sichergestellt werden, indem – wie bei dem in Abschnitt 2.4.3 vorgestellten IFTA – als letzte Operation innerhalb jeder Iteration der Transmissionsprojektionsoperator angewendet wird. Wie in Abschnitt 2.4.4 gezeigt, erhalten wir in diesem Fall entsprechend Gleichung (2.72) eine Konvergenz bezüglich des modifizierten SDEs $J_{\text{in}}(t^{(n)})$, wobei wegen der Erfüllung der Transmissionseinschränkung gilt

$$J_{\text{in}}(t^{(n)}) = J_{\text{in}}(\Pi_{\text{tran}} t^{(n)}). \quad (5.15)$$

Bei der Anwendung des PMPAs laut Abschnitt 5.1 ist die Transmissionseinschränkung nach Anwendung des Projektionsoperators $\tilde{\Pi}_2$ aus Gleichung (2.24c) im Allgemeinen nicht erfüllt. Aus der in Abschnitt 5.1.2 entsprechend Gleichung (5.12) nachgewiesenen Konvergenz bezüglich des modifizierten SSDEs $\hat{J}_{\text{in}}(t^{(n)})^2$ lässt sich daher im Allgemeinen keine Konvergenz bezüglich

$$\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2 = \hat{J}_{\text{sig}}(\Pi_{\text{tran}} t^{(n)})^2 \quad (5.16)$$

ableiten, wobei \hat{J}_{sig} durch Gleichung (5.14b) gegeben ist. Eine Ausnahme stellt der Fall dar, dass $\hat{J}_{\text{in}}(t^{(n)})^2$ gegen Null konvergiert, was nur für konsistente Optimierungsprobleme möglich ist. In diesem Fall resultiert auch für $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$ eine Konvergenz gegen Null.

Das Konvergenzverhalten des PMPAs bezüglich $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$ kann somit tendenziell umso schlechter werden, je stärker die dem Optimierungsproblem auferlegten Einschränkungen von einer Konsistenz abweichen. Von einem Verfahren mit einer Konvergenz bezüglich $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$ können insbesondere für stark inkonsistente Designprobleme bessere Ergebnisse erwartet werden, was die Motivation für die im Folgenden beschriebene Konstruktion des PTDA's darstellt.

5.2.2 Grundprinzip

Die Konstruktion des PTDA's erfolgt im Vergleich zur Definition (5.6) der Einschränkungsmengen $\tilde{\mathcal{M}}_1$ und $\tilde{\mathcal{M}}_2$ des PMPAs über eine Verlagerung der Transmissionseinschränkung aus der Menge $\tilde{\mathcal{M}}_1$ nach $\tilde{\mathcal{M}}_2$. Somit ist bei Anwendung des Algorithmus

$$\mathbf{u}^{(n)} = \tilde{\Pi}_2 \tilde{\Pi}_1 \mathbf{u}^{(n-1)} \quad (5.17)$$

nach jeder Iteration die Transmissionseinschränkung erfüllt, was der im vorangegangenen Abschnitt gegebenen Motivation für die Konstruktion des PMPAs entspricht. Diese Modifikation der Definition (5.6) führt zu folgenden Einschränkungsmengen für den PTDA:

$$\tilde{\mathcal{M}}_1 = \{\mathbf{u} = (u_1, u_2, \dots, u_B) : u_\beta \in \mathcal{M}_\beta, \beta = 1 \dots B\}, \quad (5.18a)$$

$$\tilde{\mathcal{M}}_2 = \{\mathbf{u} = (u_{\text{in},1} t_{\mathbf{u}}, u_{\text{in},2} t_{\mathbf{u}}, \dots, u_{\text{in},B} t_{\mathbf{u}}) : t_{\mathbf{u}} \in \mathcal{M}_{\text{tran}}\}. \quad (5.18b)$$

Zur Formulierung des Algorithmus benötigen wir die zu $\tilde{\mathcal{M}}_1$ und $\tilde{\mathcal{M}}_2$ gehörigen Projektionsoperatoren $\tilde{\Pi}_1$ bzw. $\tilde{\Pi}_2$. Während $\tilde{\Pi}_1$ analog zu Gleichung (2.24b) aus den Projektionsoperatoren Π_β ($\beta = 1 \dots B$) konstruiert wird über

$$\tilde{\Pi}_1 \mathbf{u} = (\Pi_1 u_1, \Pi_2 u_2, \dots, \Pi_B u_B), \quad (5.19)$$

kann $\tilde{\Pi}_2$ demgegenüber nicht ohne weiteres für beliebige Transmissionseinschränkungen einheitlich formuliert werden. Die zunächst auf Grund von Gleichung (5.7) naheliegende Konstruktion

$$\tilde{\Pi}_2 \mathbf{u} = \mathbf{v} = (u_{\text{in},1} t_{\mathbf{v}}, u_{\text{in},2} t_{\mathbf{v}}, \dots, u_{\text{in},B} t_{\mathbf{v}}), \quad \text{wobei} \quad (5.20a)$$

$$t_{\mathbf{v}} = \Pi_{\text{tran}} \frac{\sum_{\beta=1}^B u_{\text{in},\beta}^* u_\beta}{\sum_{\beta=1}^B |u_{\text{in},\beta}|^2}, \quad (5.20b)$$

ist nicht notwendigerweise abstandsminimierend, da bei Betrachtung dreier Mengen $\mathcal{M}_a \supseteq \mathcal{M}_b \supseteq \mathcal{M}_c$ mit zugehörigen Projektionsoperatoren $\Pi_{a \rightarrow b} : \mathcal{M}_a \mapsto \mathcal{M}_b$, $\Pi_{b \rightarrow c} : \mathcal{M}_b \mapsto \mathcal{M}_c$, und $\Pi_{a \rightarrow c} : \mathcal{M}_a \mapsto \mathcal{M}_c$ die Beziehung

$$\Pi_{a \rightarrow c} = \Pi_{b \rightarrow c} \Pi_{a \rightarrow b} \quad (5.21)$$

nicht notwendigerweise erfüllt ist. Der Operator $\tilde{\Pi}_2$ gemäß der Definition (5.20) entspricht hierbei einem Operator $\Pi_{b \rightarrow c} \Pi_{a \rightarrow b}$ mit

$$\mathcal{M}_a = \mathcal{H}^B, \quad (5.22a)$$

$$\mathcal{M}_b = \{\mathbf{u} = (u_{\text{in},1}t, u_{\text{in},2}t, \dots, u_{\text{in},B}t) : t \in \mathcal{H}\} \quad \text{und} \quad (5.22b)$$

$$\mathcal{M}_c = \tilde{\mathcal{M}}_2 = \{\mathbf{u} = (u_{\text{in},1}t, u_{\text{in},2}t, \dots, u_{\text{in},B}t) : t \in \mathcal{M}_{\text{tran}}\}, \quad (5.22c)$$

und bewirkt somit im Allgemeinen keine Projektion auf $\mathcal{M}_c = \tilde{\mathcal{M}}_2$.

In den folgenden Abschnitten 5.2.3 und 5.2.4 werden wir mit der Phasen- bzw. Phasenquantisierungseinschränkung zwei Fälle betrachten, in denen das Schema (5.20) zu Konstruktion von $\tilde{\Pi}_2$ dennoch erfolgreich ist. Für die in Abschnitt 5.2.5 betrachtete Amplitudeneinschränkung, kann demgegenüber die Abstandsminimierung von $\tilde{\Pi}_2$ bei einer Definition entsprechend der Gleichungen (5.20) nicht sichergestellt werden.

5.2.3 Phaseneinschränkung

Als Spezialfall der Mengendefinitionen (5.18) ergibt im Fall der Phaseneinschränkung

$$\tilde{\mathcal{M}}_1 = \{\mathbf{u} = (u_1, u_2, \dots, u_B) : u_\beta \in \mathcal{M}_\beta, \beta = 1 \dots B\}, \quad (5.23a)$$

$$\tilde{\mathcal{M}}_2 = \{\mathbf{u} = (u_{\text{in},1}t_{\mathbf{u}}, u_{\text{in},2}t_{\mathbf{u}}, \dots, u_{\text{in},B}t_{\mathbf{u}}) : t_{\mathbf{u}} \in \mathcal{M}_{\text{phase}}\}, \quad (5.23b)$$

wobei $\mathcal{M}_{\text{phase}}$ durch Gleichung (2.59) gegeben ist.

Zur Definition der zugehörigen Projektionsoperatoren betrachten wir

$$\tilde{\Pi}_1 \mathbf{u} = (\Pi_1 u_1, \Pi_2 u_2, \dots, \Pi_B u_B) \quad \text{und} \quad (5.24a)$$

$$\tilde{\Pi}_2 \mathbf{u} = \mathbf{v} = (u_{\text{in},1}t_{\mathbf{v}}, u_{\text{in},2}t_{\mathbf{v}}, \dots, u_{\text{in},B}t_{\mathbf{v}}) \quad \text{mit} \quad (5.24b)$$

$$t_{\mathbf{v}} = \Pi_{\text{phase}} \frac{\sum_{\beta=1}^B u_{\text{in},\beta}^* u_\beta}{\sum_{\beta=1}^B |u_{\text{in},\beta}|^2} = \Pi_{\text{phase}} \sum_{\beta=1}^B u_{\text{in},\beta}^* u_\beta. \quad (5.24c)$$

Während die Abstandsminimierungseigenschaft von $\tilde{\Pi}_1$ in Lemma 2.25 nachgewiesen wurde, zeigt der nachfolgende Satz 5.25, dessen Beweis in Anhang A.4.2 dokumentiert wird, dass es sich bei dem nach dem Schema (5.20) konstruierten Operator $\tilde{\Pi}_2$ ebenfalls um einen Projektionsoperator handelt. Somit folgt unter den genannten Voraussetzungen eine Konvergenz des PTDA's bezüglich $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$.

Satz 5.25 *Der Operator $\tilde{\Pi}_2$ gemäß Gleichung (5.24b) definiert einen punktweise abstandsminimierenden Projektionsoperator auf $\tilde{\mathcal{M}}_2$ aus Gleichung (5.23b).*

5.2.4 Quantisierungseinschränkung

Für die aus Abschnitt 2.4.1 bekannte Einschränkung auf Q äquidistante Phasenwerte ergibt sich $\widetilde{\mathcal{M}}_2$ entsprechend Gleichung (5.18b) als

$$\widetilde{\mathcal{M}}_2 = \{\mathbf{u} = (u_{\text{in},1}t_{\mathbf{u}}, u_{\text{in},2}t_{\mathbf{u}}, \dots, u_{\text{in},B}t_{\mathbf{u}}) : t_{\mathbf{u}} \in \mathcal{M}_{\text{quant},Q}\}. \quad (5.26)$$

Der folgende Satz zeigt, dass der nach dem Schema (5.20) konstruierte Operator $\widetilde{\Pi}_2$,

$$\widetilde{\Pi}_2 \mathbf{u} = \mathbf{v} = (u_{\text{in},1}t_{\mathbf{v}}, u_{\text{in},2}t_{\mathbf{v}}, \dots, u_{\text{in},B}t_{\mathbf{v}}) \quad \text{mit} \quad (5.27a)$$

$$t_{\mathbf{v}} = \Pi_{\text{quant},Q} \frac{\sum_{\beta=1}^B u_{\text{in},\beta}^* u_{\beta}}{\sum_{\beta=1}^B |u_{\text{in},\beta}|^2}, \quad (5.27b)$$

auch in diesem Fall einen Projektionsoperator darstellt, womit für den PTDA eine Konvergenz hinsichtlich $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$ folgt.

Satz 5.28 *Der Operator $\widetilde{\Pi}_2$ aus Gleichung (5.27) definiert eine punktweise abstandsminimierende Projektion auf $\widetilde{\mathcal{M}}_2$ nach Gleichung (5.26).*

Da in dem in Anhang A.4.3 dargestellten Beweis dieses Satzes keinen Gebrauch von der Äquidistanz der Phasenwerte in $\widetilde{\mathcal{M}}_{\text{quant},Q}$ gemacht wird, lässt sich das Schema (5.20) auch für die Quantisierung auf nichtäquidistante Phasenwerte anwenden. Falls die Quantisierungswerte jedoch auch Amplitudenvariationen besitzen, führt Schema (5.20) im Allgemeinen nicht mehr zu einem Projektionsoperator, wie aus dem folgenden Abschnitt klar wird. In solchen Fällen kann jedoch auf Grund der Endlichkeit der Menge der Quantisierungswerte ein Projektionsoperator konstruiert werden, der für jede Position x denjenigen Quantisierungswert wählt, welcher zu dem geringsten Beitrag zur Metrik $d(\mathbf{u}, \mathbf{v})$ führt. Eine derartige Konstruktion ist allerdings im Vergleich zur Auswertung von Gleichung (5.27) mit einem erhöhten Rechenaufwand verbunden.

5.2.5 Amplitudeneinschränkung

Die Projektion $\widetilde{\Pi}_2$ im Fall von Amplitudeneinschränkungen kann über die im Beweis von Satz 5.8 in Anhang A.4.1 gezeigte Vorgehensweise konstruiert werden als

$$\widetilde{\Pi}_2 \mathbf{u} = \mathbf{v} = (u_{\text{in},1}t_{\mathbf{v}}, u_{\text{in},2}t_{\mathbf{v}}, \dots, u_{\text{in},B}t_{\mathbf{v}}) \quad \text{mit} \quad (5.29a)$$

$$t_{\mathbf{v}}(x) = \frac{\sum_{\beta=1}^B \text{Re}(u_{\text{in},\beta}^*(x)u_{\beta}(x))}{\sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2}. \quad (5.29b)$$

Demgegenüber würde das Schema (5.20) zu einer Transmissionsfunktion

$$t'_{\mathbf{v}}(x) = \left| \frac{\sum_{\beta=1}^B u_{\text{in},\beta}^* u_{\beta}}{\sum_{\beta=1}^B |u_{\text{in},\beta}|^2} \right| \quad (5.30)$$

führen. Da im Allgemeinen

$$t_{\mathbf{v}}(x) \neq t'_{\mathbf{v}}(x) \quad (5.31)$$

gilt, ist das Schema (5.20) zur Konstruktion von $\widetilde{\Pi}_2$ in diesem Fall also nicht erfolgreich.

5.3 Vergleich der Konvergenzeigenschaften für eine Beispielanwendung

In diesem Abschnitt werden die Konvergenzeigenschaften des PTDA für Phaseneinschränkungen mit denen des PMPAs verglichen. In die Betrachtungen wird zusätzlich eine Abwandlung des PTDA mit einbezogen, welche eine im Vergleich zu Gleichung (5.24c) veränderte Mittelwertbildung benutzt: Gleichung (5.24c) entspricht einer gewichteten Mittelung der den einzelnen optischen Funktionen entsprechenden Transmissionen, wobei die Wichtung über die Intensitätsverteilung der Eingangsfelder erfolgt. In der abgeänderten Variante erfolgt die Mittelung der einzelnen Transmissionen ungewichtet, was in Abwandlung zu Gleichung (5.24b) folgendem Projektionsoperator $\tilde{\Pi}_2$ entspricht:

$$\tilde{\Pi}_2 \mathbf{u} = \mathbf{v} = (u_{\text{in},1} t_{\mathbf{v}}, u_{\text{in},2} t_{\mathbf{v}}, \dots, u_{\text{in},B} t_{\mathbf{v}}) \quad \text{mit} \quad (5.32a)$$

$$t_{\mathbf{v}} = \Pi_{\text{phase}} \sum_{\beta=1}^B \frac{u_{\beta}}{u_{\text{in},\beta}}. \quad (5.32b)$$

Für den Fall konstanter Eingangsfelder ist der PTDA mit der oben beschriebenen Abwandlung äquivalent.

Zum Vergleich der genannten Designverfahren betrachten wir analog zu Referenz [35] ein Strahlformungsproblem zur Umwandlung eines gaußförmigen Eingangsfelds in eine Top-Hat-Verteilung, wobei die Strahlformungsoptik tolerant gegen Schwankungen der Breite der gaußförmigen Eingangsverteilung sein soll. Die Umwandlung von Gauß- nach Top-Hat-Verteilungen gehört zu den Standardanwendungen im Bereich der Strahlformung. Die Forderung nach Toleranz bezüglich Schwankungen in der Breite der Eingangsverteilung ist motiviert durch die insbesondere bei Laserdioden auftretenden fabrikationsbedingten Schwankungen der Tailenbreite.

Wir gehen in diesem Beispiel von einer gaußförmigen Eingangsverteilung in der Tailenposition mit Tailenbreiten w_0 zwischen 1.5 mm und 2.5 mm bei einer Wellenlänge von 632.8 nm aus, wobei w_0 den Radius bezeichnet, innerhalb dessen die Intensität um den Faktor $1/e^2$ abgefallen ist. Eine Beschränkung auf zwei optische Funktionen ($B = 2$) für $w_0 = 1.5$ mm bzw. $w_0 = 2.5$ mm erweist sich im Folgenden als ausreichend. In der Signalebene bestehen Amplituden-, Phasen- und Skalierungsfreiheit. Dabei wird das Signalfenster so gewählt, dass es ausschließlich den Bereich enthält, innerhalb dessen die Intensitätsverteilung zu homogenisieren ist. Die Skalierungsfreiheit wird bei Anwendung der Signalprojektion hinsichtlich einer Zieleffizienz von 90 % beschränkt (vgl. Abschnitt 2.3.2).

In Referenz [35] wurde bereits nachgewiesen, dass mit Hilfe eines multifunktionalen Designs die Sensitivität eines derartigen Strahlformers gegenüber Schwankungen der Tailenbreite w_0 deutlich verringert werden kann. Die Verbesserung des Toleranzverhaltens wurde dabei im Vergleich zu einem monofunktionalen IFTA-Design (vgl. Abschnitt 2.4.3) sowie im Vergleich zu einem Strahlformerdesign basierend auf Methoden der geometrischen Optik [30] nachgewiesen. Wir beschränken uns daher im Folgenden auf einen Vergleich der am Anfang dieses Abschnitts genannten multifunktionalen Designmethoden untereinander.

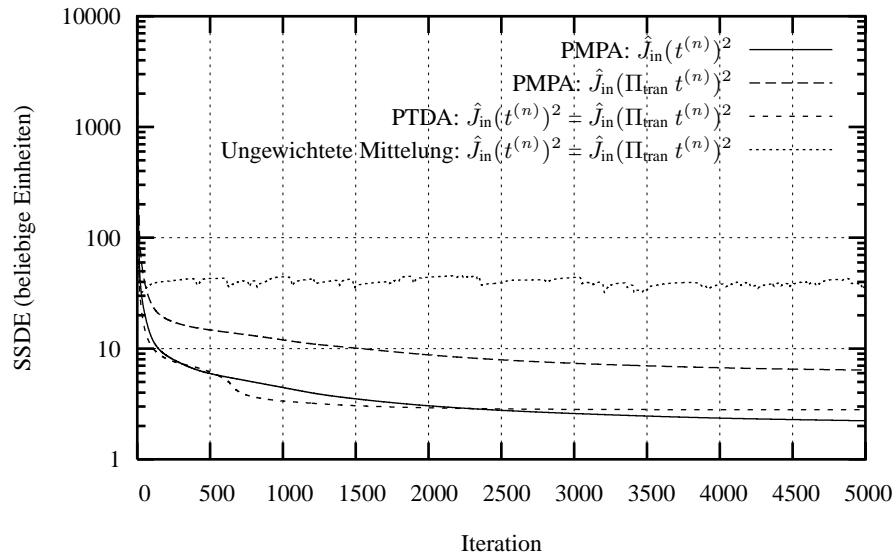


Abbildung 5.1: Verlauf von $\hat{J}_{\text{in}}(t^{(n)})^2$ bzw. $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$ während der Iteration für den PMPA, den PTDA sowie für das über die Gleichungen (5.32) beschriebene abgewandelte Verfahren mit ungewichteten Mittelungen

Abbildung 5.1 zeigt den Verlauf der entsprechend Abschnitt 5.2.1 interessanten Gütefunktion $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$ für die drei betrachteten Designverfahren. Für den PMPA ist zusätzlich der Verlauf $\hat{J}_{\text{in}}(t^{(n)})^2$ dargestellt, welcher in Übereinstimmung mit der in Abschnitt 5.2.1 geführten Diskussion ein deutlich besseres Konvergenzverhalten besitzt als $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$. Es zeigt sich, dass für den bei einem Vergleich der drei Verfahren hauptsächlich relevanten Verlauf von $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$ über den PTDA eine stärkere Fehlerreduktion möglich ist als über den PMPA, was eine Bestätigung für die in Abschnitt 5.2.1 durchgeführten Betrachtungen darstellt. Der auf ungewichteten Mittelungen basierende Algorithmus weist im Gegensatz zu den anderen beiden Verfahren keine Konvergenz auf. Da auch der verbleibende Restfehler deutlich größer ausfällt als bei den anderen beiden Verfahren, zeigt sich, dass die gewichtete Mittelung im Fall nichtkonstanter Eingangsfelder essentiell ist.

In den Abbildungen 5.2 und 5.3 werden die mit den drei Algorithmen nach jeweils 5000 Iterationen erzielten Lösungen hinsichtlich ihrer Sensitivität gegen Veränderung der Tailenbreite w_0 miteinander verglichen. Um eine bessere Vergleichbarkeit zu gewährleisten, wurde bei der betrachteten Lösung des PMPAs die Phase extrahiert, um eine Erfüllung der Transmissionseinschränkung zu gewährleisten. Das in Abbildung 5.2 gezeigte Verhalten des SNRs weist im Vergleich zwischen PMPA und PTDA keine starken Unterschiede auf: Mit beiden Algorithmen konnten Lösungen gefunden werden, welche über einen breiten Bereich von w_0 eine gute Homogenität ($\text{SNR} > 25$ dB) besitzen. Im Gegensatz hierzu erreicht die über ungewichtete Mittelung erzielte Lösung nicht das gewünschte tolerante Verhalten.

Hinsichtlich des in Abbildung 5.3 dargestellten Verhaltens der Konversionseffizienz bestehen deutliche Unterschiede zwischen PMPA und PTDA: Die mit dem PTDA erzielte Konversionseffizienz ist in Abhängigkeit von w_0 etwa 5 – 6 % höher. Der durch die untere Beschränkung der Skalierungsfreiheit vorgegebene Zielwert für die Effizienz von 90 % konnte für alle drei Algorithmen nicht erreicht werden. Entsprechend der in Abschnitt 4.3.3 geführten Diskussion ist es jedoch

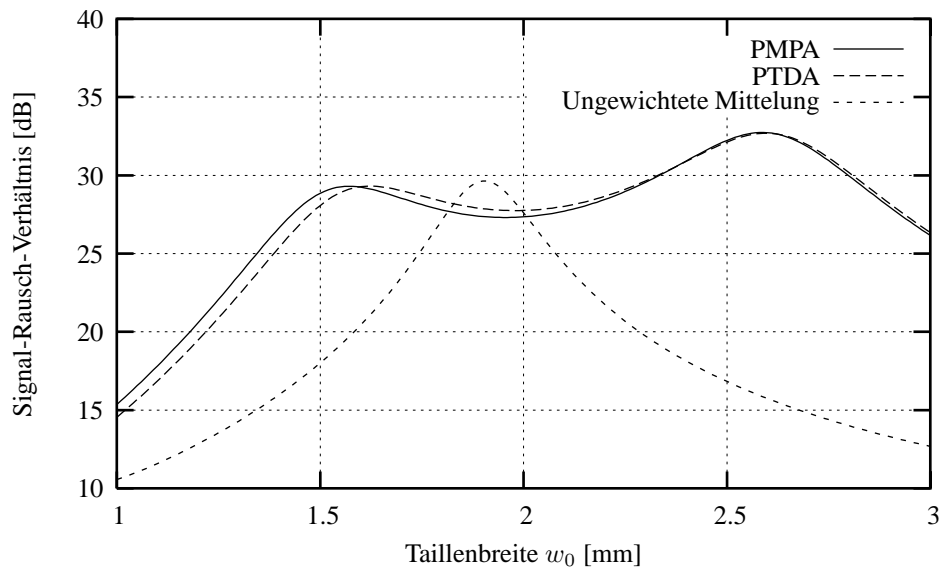


Abbildung 5.2: Abhängigkeit des Signal-Rausch-Verhältnisses von der Tailenbreite w_0 für die Ergebnisse der Anwendung von jeweils 5000 Iterationen des PMPAs (durchgezogen), des PTDA (gestrichelt) sowie der PTDA-Variante mit ungewichteter Transmissionsmittelung entsprechend Gleichung (5.32) (gepunktet)

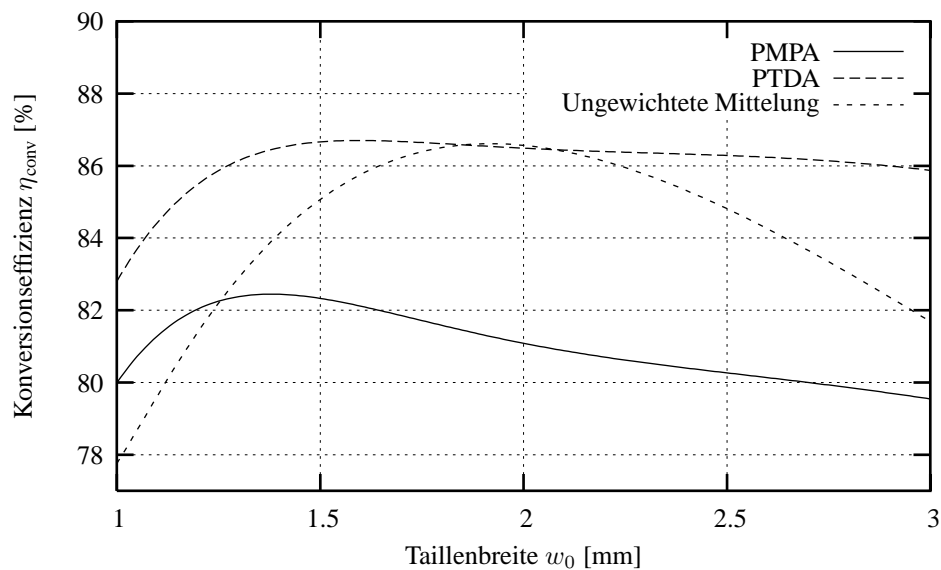


Abbildung 5.3: Abhängigkeit der Konversionseffizienz η_{conv} von der Tailenbreite w_0 für die drei bereits in Abbildung 5.2 untersuchten Designs

zur Erzielung von Lösungen mit hoher Effizienz im Allgemeinen günstig, die untere Beschränkung der Skalierungsfreiheit entsprechend einer höheren als der tatsächlich erreichbaren Effizienz zu wählen.

Abbildung 5.4 zeigt die für $w_0 = 1.5$ mm, $w_0 = 2.0$ mm und $w_0 = 2.5$ mm für die drei betrachteten Designs resultierenden Ausgangsintensitätsverteilungen. Obwohl die Lösungen für PMPA und PTDA laut Abbildung 5.2 eine ähnliche SNR-Charakteristik besitzen, stellen sich die in der mittleren Spalte von Abbildung 5.4 gezeigten Ergebnisse des PTDA als visuell besser dar.

Die in einigen Ausgangsverteilungen stark sichtbare Asymmetrie ist durch numerische Ungenauigkeiten während der Iteration entstanden. Solche leichten Ungenauigkeiten sind im Übrigen

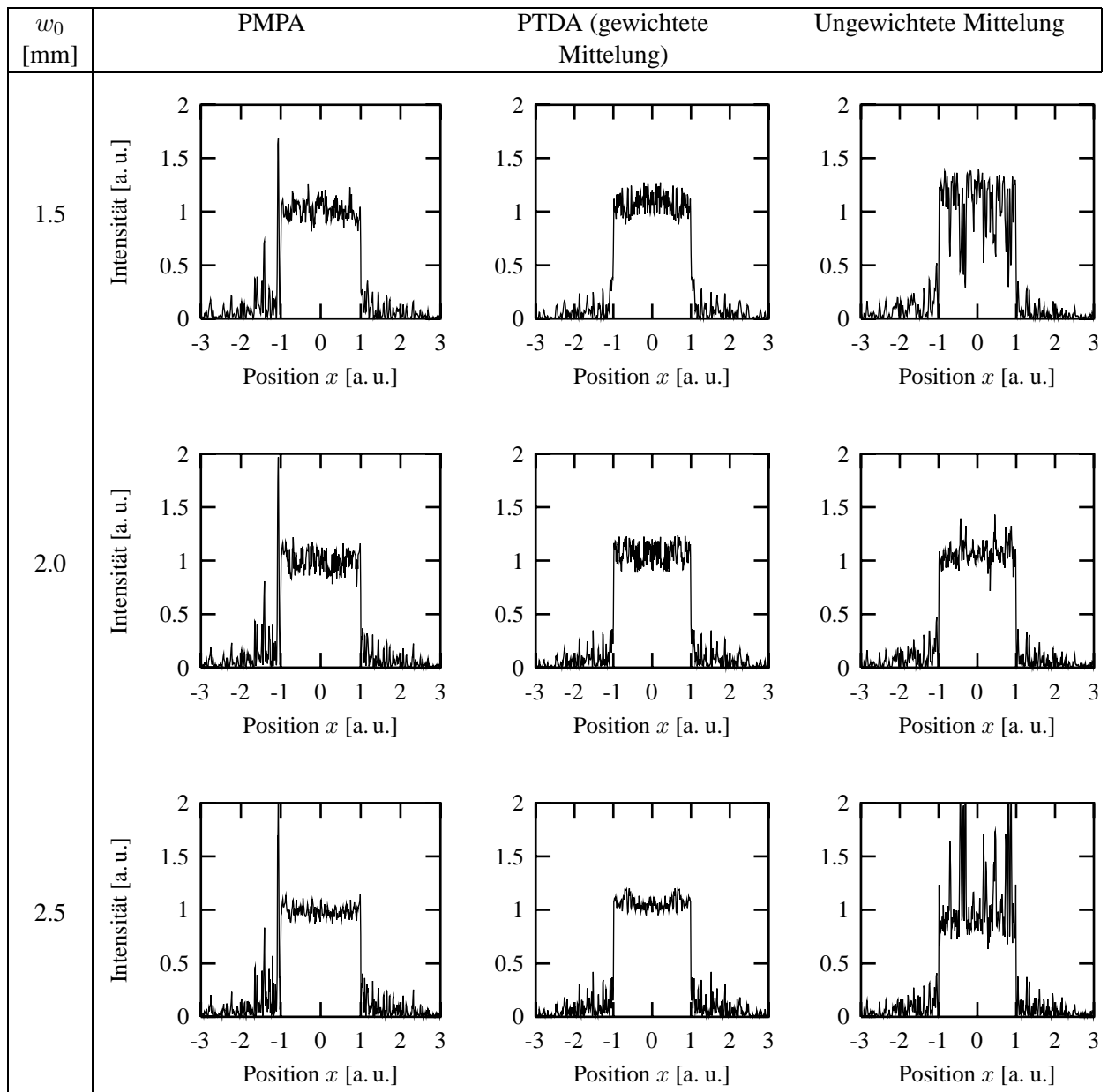


Abbildung 5.4: Laterale Ausgangsintensitätsverteilungen der drei bereits in Abbildung 5.2 untersuchten Designs für $w_0 = 1.5$ mm, $w_0 = 2.0$ mm und $w_0 = 2.5$ mm

auch die Ursache dafür, dass trotz der spiegelsymmetrischen Eingangsfelder mit konstanter Phase sowie der konstanten Startverteilung $t^{(0)}$ eine nicht binäre Phasenfunktion als Lösung gefunden wird. Da die Fouriertransformation von symmetrischen reellen Funktionen ebenfalls symmetrisch und reell ist, würde die Iteration bei idealer Genauigkeit in diesem Fall zu keinem sinnvollen Ergebnis führen.

Am Ende von Abschnitt 5.2.1 wurde begründet, warum der Vorteil des PTDA gegenüber dem PMPA tendenziell zunimmt, je stärker die Einschränkungen von einer Konsistenz abweichen. Für das betrachtete Beispiel können wir die Inkonsistenz durch Herabsetzung der unteren Schranke des Skalierungsfaktors verringern. Bei einer Beschränkung der Skalierungsfreiheit entsprechend einer Zieleffizienz von 85 % resultieren für den PTDA innerhalb für $w_0 \in [1.5 \text{ mm}, 2.5 \text{ mm}]$ Effizienzen von etwa 83 %, während sich für den PTDA eine im Vergleich zu Abbildung 5.3 leicht reduzierte

Effizienz ergibt. Der Unterschied zwischen den für beide Algorithmen erzielten Ergebnissen hat sich hierbei also verringert.

5.4 Frühere Arbeiten

Frühere Arbeiten zu optischen Problemstellungen, welche sich als multifunktionale Designprobleme entsprechend Abschnitt 2.3.1 formulieren lassen, fanden zum einen im Bereich der *3D-Strahlformung* statt. Eine Übersicht zu Arbeiten in diesem Bereich kann in Referenz [90] gefunden werden. Ein hierbei weit verbreiteter Grundansatz besteht darin, das Ausgangsfeld nur in einer endlichen Anzahl von lateralen Ebenen zu betrachten. Bei einer multifunktionalen Modellierung entsprechend Abschnitt 2.3.1 entspricht hierbei jede betrachtete Ausgangsebene einer optischen Funktion.

3D-Strahlformungsverfahren, welche auf einer Mittelung der für die einzelnen optischen Funktionen resultierenden Transmissionen basieren, können teilweise als Spezialfälle der Designansätze aus den Abschnitten 5.1 und 5.2 angesehen werden. Im Gegensatz zu einfachen Ansätzen, wie beispielsweise dem von Leseberg [91], enthält der Algorithmus von Cumme [92] auch Projektionsoperatoren zur Einbeziehung der Transmissionseinschränkung und der Signalebenenfreiheiten sowie eine Iterationssteuerung, welche Ähnlichkeiten mit dem PTDA aus Abschnitt 5.2 besitzt. Da in der genannten Arbeit die Mittelwertbildung der einzelnen Transmissionen im Gegensatz zu Gleichung (5.24c) keine Wichtung bezüglich der Eingangsfelder umfasst, entstehen bei diesem Ansatz die in Abschnitt 5.3 demonstrierten Konvergenzprobleme, sofern von nichtkonstanten Eingangsfeldern ausgegangen wird.

Auch der in Abschnitt 2.1.5 betrachtete serielle Projektionsalgorithmus für mehr als zwei Mengen wurde zur 3D-Strahlformung eingesetzt [93, 94]. Da dieser Algorithmus jedoch im Allgemeinen nur für konvexe Einschränkungsmengen konvergiert, bleibt der praktische Anwendungsbereich dieses Ansatzes stark begrenzt.

Bei Problemstellungen zur *axialen Strahlformung* wird nur das Intensitätsprofil entlang der optischen Achse betrachtet. Die hierfür vorgeschlagenen Verfahren [95, 96] basieren nicht auf der oben erwähnten multifunktionalen Modellierung.

Ein weiteres Anwendungsfeld multifunktionaler Designverfahren ist die Optimierung von Transmissionen unter Berücksichtigung globaler Höhenfehler oder damit äquivalenter Wellenlängenfehler. Das Verfahren von Levy et. al. [97] zur Berechnung eines Strahlformers, welcher für drei diskrete Wellenlängen jeweils verschiedene Strahlumformungen bewirkt, stellt einen Spezialfall des Verfahrens aus Abschnitt 5.2.3 dar. Der Ansatz von Fuse et. al. [70] zur Verbesserung der Toleranzeigenschaften von Strahlteilern gegenüber Höhenfehlern basiert demgegenüber auf einer Optimierung über *Direct Binary Search* und besitzt somit gegenüber Projektionsalgorithmen die in Kapitel 1 diskutierten prinzipiellen Nachteile.

Ein anderes Optimierungsverfahren, die sogenannte *Optimal Rotation Angle Method (ORA)*, wurde unter anderem zur Berechnung von multifunktionalen Strahlteilern vorgeschlagen. Es wurden sowohl Beispiele demonstriert, in denen für verschiedene Wellenlängen jeweils unterschied-

liche Intensitätsverteilungen resultieren [18, 98], als auch ein Beispiel aus dem Bereich des toleranzoptimierten Designs, wobei das Ziel darin bestand, den Uniformitätsfehler eines speziellen 1-zu-5-Strahlteilers über einen breiten Bereich von linearen Höhenfehlern bzw. Wellenlängenfehlern klein zu halten [99]. Die Nachteile von ORA-basierten Ansätze gegenüber Projektionsalgorithmen wurden bereits in Kapitel 1 diskutiert.

Einen zu den oben genannten Transmissionsdesignverfahren alternativen Ansatz stellen Strukturdesignmethoden dar, welche die Einbeziehung von Systemanforderungen für mehrere diskrete Beleuchtungswellenlängen oder ausgedehnte Wellenlängenbereiche ermöglichen. Als Erweiterung des in Abschnitt 2.2.6 diskutierten Strukturdesigns auf der Basis der Dünne-Elemente-Approximation existieren zum einen Ansätze, welche zwei für verschiedene Wellenlängen berechnete Transmissionsfunktionen in ein Höhenprofil überführen, dessen Tiefe einer Phasenmodulation um mehrere Vielfache von 2π entspricht [100–103]. Die diskreten Höhenstufen werden dabei so gewählt, dass sich für jeden Abtastwert bei einer Wellenlänge λ_1 mehrere im Rahmen der Dünne-Elemente-Approximation äquivalente Höhenwerte möglich sind, die sich in ihrer Phasenmodulation jeweils um ganze Vielfache von 2π voneinander unterscheiden. Für Wellenlängen $\lambda_2 \neq \lambda_1$ führen diese Höhenwerte jedoch zu unterschiedlichen Phasenmodulationen, was eine gezielte Wahl der sich für λ_2 ergebenden Transmissionsfunktion möglich macht. Zum anderen bestehen Strukturdesignverfahren, welche eine Transmissionsfunktion über zwei oder mehr direkt benachbarte höhenmodulierte Elemente realisieren, wobei über die Wahl der Modulationstiefen beider Elemente unter Ausnutzung der Dispersionseigenschaften die Wellenlängenabhängigkeit der von diesen Elementen bewirkten Phasenmodulationen deutlich reduziert werden kann [104–107].

Zum Abschluss möchten wir noch auf Verfahren verweisen, welche auf der Idee der linearen Optimierung zur Einbeziehung mehrerer optischer Funktionen beruhen [35, 108, 109]. Wie im Beweis zu Satz 5.8 gezeigt wurde, entspricht die gewichtete Mittelwertbildung der Lösung eines linearen Optimierungsproblems. Da sich dieses lineare Optimierungsproblem für den SDE bei Propagationsoperatoren, welche Gleichung (2.39a) erfüllen, äquivalent zwischen Transmissions- und Signalebene übertragen lässt, stellt der Ansatz aus Abschnitt 5.2.3 eine effizienter implementierbare Verallgemeinerung des Verfahrens aus Referenz [35] dar.

Kapitel 6

Verbesserung von Toleranzeigenschaften durch multifunktionale Optimierung

Die Toleranzeigenschaften optischer Systeme stellen in vielen Anwendungsfällen eine kritische Anforderung dar, welche beim Design berücksichtigt werden muss. Es kann zwischen zwei verschiedenen Motivationen zur Durchführung toleranzoptimierter Designs unterschieden werden. Zum einen bestehen Anwendungen, in denen die betroffenen System- und Beleuchtungsparameter unvermeidbar zeitlich schwanken oder mit verschiedenen Werten überlagert auftreten. Beispiele hierfür stellen Anwendungen mit starken Temperatureinflüssen auf Brechzahl, Wellenlänge der Lichtquelle oder Ausdehnung der optischen Elemente dar sowie Anwendungen mit polychromatischer Beleuchtung, welche beispielsweise auch bei kurzen Pulslängen vorliegt.

Während bei der eben genannten Klasse von Anwendungen die Erfüllung der vorgegebenen Toleranzeigenschaften unabdingbar für die Realisierung der gewünschten optischen Wirkung ist, existieren auch Anwendungen, bei denen bestimmte Toleranzeigenschaften zur Vereinfachung der Herstellbarkeit oder Justierbarkeit des optischen Systems gefordert werden. So ist das im folgenden Abschnitt 6.1 betrachtete Verfahren zur Erhöhung der Schärfentiefe beim *Proximity Printing* durch die Forderung nach Justierunempfindlichkeit hinsichtlich des Abstands zwischen Maske und Substrat motiviert. Ebenso können die in den Abschnitt 6.2 betrachteten Höhenfehler bei der lithografischen Herstellung von Strahlteilern in vielen Fällen durch aufwändigere Herstellungsverfahren deutlich reduziert werden [110], und fallen somit in diese Kategorie von Anwendungen.

Das ebenfalls in Abschnitt 6.2 untersuchte Design eines gegen Wellenlängenabweichungen toleranten Strahlteilers sowie das bereits in Abschnitt 5.3 betrachtete Beispiel zum Design eines Strahlformers, welcher tolerant gegen Schwankungen der Breite des Beleuchtungsfelds ist, stellen hingegen Beispiele für die erstgenannte Kategorie von Designproblemen dar.

Bei den in diesem Kapitel vorgestellten Beispielanwendungen erfolgt die Optimierung der Toleranzeigenschaften unter Benutzung des in Abschnitt 5.2 eingeführten PTDA. Einige Anwendungsfälle alternativer Methoden im Bereich des toleranzoptimierten Designs wurden bereits in Abschnitt 5.4 aufgeführt.

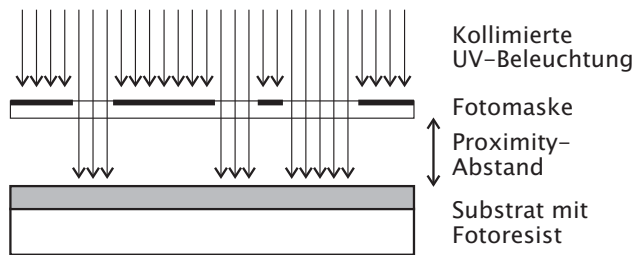


Abbildung 6.1: Systemaufbau bei Fotolithografie mit Mask-Aligner



Abbildung 6.2: Betrachtete Zielstruktur mit einer minimalen Strichbreite von $3 \mu\text{m}$

6.1 Toleranzeigenschaften bezüglich des Propagationsabstands

In diesem Abschnitt betrachten wir die Berechnung von komplexwertigen Transmissionsfunktionen für Fotomasken, welche zur Belichtung von Substraten mit Hilfe von *Mask-Alignern* eingesetzt werden können. Entsprechend Abbildung 6.1 besteht zwischen der Maske und dem zu belichtenden Substrat ein sogenannter *Proximity-Abstand* d_{prox} , welcher insbesondere beim industriellen Einsatz dieses Verfahrens notwendig ist, um die Maske vor Beschädigung zu schützen. Ein typischer Wert beträgt $d_{\text{prox}} = 50 \mu\text{m}$. Da typischerweise mit UV-Licht gearbeitet wird, ist d_{prox} deutlich größer als die Wellenlänge und Propagationseffekte können die auf dem Substrat erzielbare Auflösung begrenzen.

In diesem Abschnitt nehmen wir Bezug auf Referenz [111], in welcher das monofunktionale Design einer komplexwertigen Maske zur Generierung der in Abbildung 6.2 gezeigten Belichtungsstruktur mit einer minimalen Strichbreite $3 \mu\text{m}$ diskutiert wurde. Dabei wurde gezeigt, dass bei Verwendung einer binären Amplitudenmaske, welche exakt der Zielstruktur aus Abbildung 6.2 entspricht, auf Grund der Propagationseffekte keine ausreichende Auflösung in der Substratebene erreicht wird. Demgegenüber konnte das Pattern bei Verwendung einer Maske mit zwei Amplituden- und vier Phasenstufen aufgelöst werden, wobei jedoch eine relativ hohe Empfindlichkeit gegenüber Abweichungen des Proximity-Abstands besteht.

Die für das in Referenz [111] erhaltenen Ergebnisse werden in Abschnitt 6.1.2 dargestellt. In Abschnitt 6.1.3 demonstrieren wir, dass die Toleranzeigenschaften gegenüber Abweichungen des Proximity-Abstands mit Hilfe eines multifunktionalen Designs erhöht werden können. Die beiden Designs zu Grunde liegende wellenoptische Modellierung wird in dem zunächst folgenden Abschnitt 6.1.1 zusammengefasst.

6.1.1 Modellierung

Die Modellierung der Beleuchtungswelle erfolgt beim Design als monochromatische ebene Welle mit senkrechtem Einfall und einer Wellenlänge $\lambda = 365 \text{ nm}$. Bei der anschließenden Analyse erfolgt eine genauere Approximation der tatsächlichen Beleuchtung, welche eine spektrale Verteilung mit 60 % Leistungsanteil bei 365 nm, 15 % bei 405 nm und 25 % bei 436 nm besitzt. Die

Ausdehnung der Lichtquelle wird unter Annahme einer gleichförmigen Verteilung des Einfallswinkels $\theta \in [0, 1^\circ]$ modelliert, wobei θ in eine beliebige Richtung aufgespannt werden kann. Zur Simulation der Intensitätsverteilung auf dem Substrat werden 100 verschiedene ebene Wellen in die Substratebene propagiert und dort durch Addition der Betragsquadrate inkohärent überlagert. Wellenlänge und Winkel der einzelnen ebenen Wellen werden dabei entsprechend der vorgegebenen Wahrscheinlichkeitsverteilung zufällig gewählt.

Die Lichtpropagation durch die Fotomaske wird über die Dünne-Elemente-Approximation modelliert, wobei sich entsprechend Abschnitt 2.2.6 für die bei der Analyse betrachtete polychromatische Beleuchtung trotz Vernachlässigung der Dispersionsbeziehung $n(\lambda)$ eine von der Wellenlänge abhängige Transmission ergibt. Die Propagation zwischen Maske und Substrat wird über den Operator $P_{\Delta z}^{\text{ASPW}}$ aus Abschnitt 2.2.5 mit $\Delta z = d_{\text{prox}}$ simuliert. Im Fall des in Abschnitt 6.1.3 betrachteten multifunktionalen Designs werden den einzelnen optischen Funktionen jeweils verschiedene Werten für Δz zugeordnet.

Den Zusammenhang zwischen Resistprofil $R(x, y)$ und der über

$$D(x, y) = I_{\text{out}}(x, y)t \quad (6.1)$$

aus der Intensitätsverteilung $I_{\text{out}}(x, y)$ und der Belichtungsdauer t resultierende Dosisverteilung $D(x, y)$ modellierenden wir über

$$R(x, y) = \begin{cases} 1 - \left(\frac{D(x, y)}{D_0} \right)^\eta & \text{falls } D(x, y) \leq D_0 \\ 0 & \text{sonst} \end{cases}, \quad (6.2)$$

D_0 bezeichnet hierbei die Dosis, welche notwendig ist um den Resist komplett zu entfernen, und η den Kontrastparameter des Resists. In dem betrachteten Beispiel werden die Werte $D_0 = 1500 \text{ J/m}^{-2}$ sowie $\eta = 5$ angenommen.

In dem in Abbildung 6.2 weiß dargestellten Bereich, welchen wir auch als *Belichtungsbereich* bezeichnen, soll der Resist komplett entfernt werden, d. h. entsprechend Gleichung (6.2) muss $I(x, y)$ in diesem Bereich einen von der Belichtungsdauer t abhängigen unteren Schwellwert überschreiten. Außerhalb des Belichtungsbereichs besteht ein unterer Schwellwert für die Resistdicke, welcher bei der betrachteten Beispielanwendung auf $R(x, y) \geq 0.5$ gesetzt wird. Hiermit resultiert ein ebenfalls von t abhängiger oberer Schwellwert für $I(x, y)$. Entsprechend dieser beiden Schwellwertbedingungen kann für jeden betrachteten Abtastwert demnach ermittelt werden, ob er entsprechend eines simulierten Resistprofils $R(x, y)$ korrekt implementiert wird. Zur Bewertung eines vorgegebenen Designs wird die Belichtungsdauer t hinsichtlich einer Minimierung der Anzahl der falsch implementierten Abtastwerte gewählt.

Beim Design wird demgegenüber ein geeigneter Wert für t fest vorgegeben, womit über den gesamten Bereich des Substrats eine entsprechend der oben beschriebenen Schwellwertbedingungen eingeschränkte Amplitudenfreiheit resultiert. Zusätzlich kann beim Design Phasenfreiheit ausgenutzt werden.

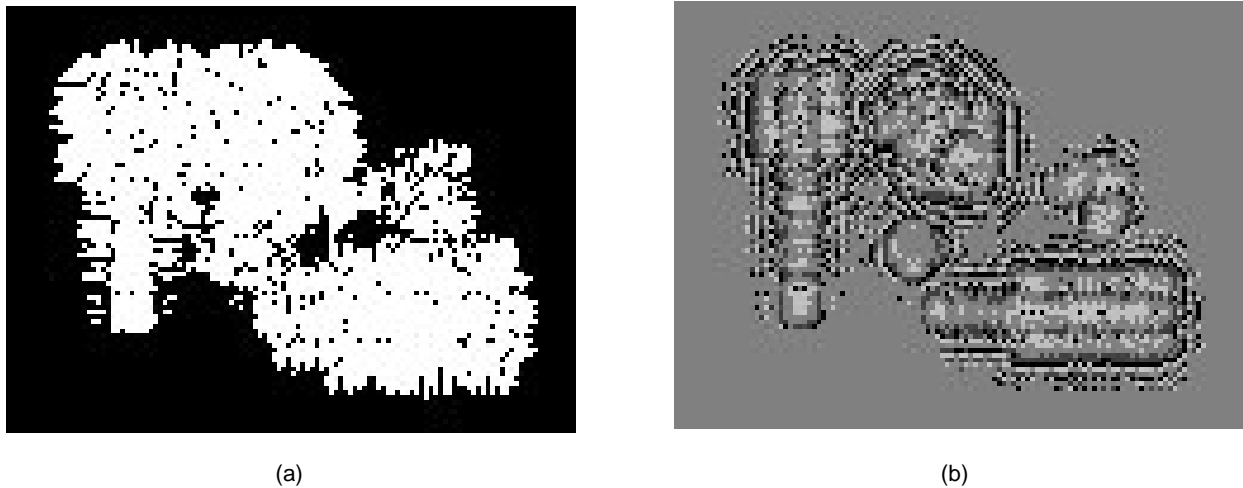


Abbildung 6.3: Amplitude (a) und Phase (b) der bei einem monofunktionalen Design resultierenden Transmission

6.1.2 Monofunktionales Design

Die in Referenz [111] angewendete monofunktionale Designstrategie basiert auf dem IFTA, welcher in Abschnitt 2.4.3 diskutiert wurde. Dabei wird die Signalprojektion entsprechend der im vorangegangenen Abschnitt beschriebenen Modellierung mit Phasenfreiheit und über Schwellwertbedingungen eingeschränkter Amplitudenfreiheit durchgeführt. Die Transmissionsprojektion wird mit einer entsprechend Referenz [112] modifizierten Version des Softquantisierungsoperators durchgeführt, wobei die Modifikation notwendig war, um neben den betrachteten $Q = 4$ Phasenwerten zusätzlich eine binäre Amplitudenmodulation zu erreichen.

Die Abtastung der Transmissionsfunktion erfolgt beim Design mit $\delta x = \delta y = 1 \mu\text{m}$ und $n_x = n_y = 128$. Bei der Analyse wird zur Einbeziehung von Pixelierungseffekten demgegenüber mit $\delta x = \delta y = 0.5 \mu\text{m}$ gerechnet, wobei einen Abtastwert während des Designs 2×2 Abtastwerte bei der Analyse zugeordnet sind.

Abbildung 6.3 zeigt Amplitude und Phase einer aus einem monofunktionalen Design nach 10 Iterationen resultierenden Transmissionsfunktion. Die hierzu korrespondierende Maske wurde lithografisch hergestellt, wobei die Phasenmodulation über vier Höhenstufen in Quarz und die Amplitudenmodulation über eine partielle Chrom-Beschichtung realisiert wurde. Eine detailliertere Beschreibung des Herstellungsprozesses wird in Referenz [111] gegeben.

Abbildung 6.4(a) zeigt einen zentralen Ausschnitt der über die in Abschnitt 6.1.1 beschriebene inkohärente Simulation berechneten Ausgangsintensitätsverteilung. In Abbildung 6.4(b) wird die über Gleichung (6.2) aus der simulierten Intensitätsverteilung berechenbare Resistprofil anhand der in Abschnitt 6.1.1 genannten Schwellwertbedingungen *ternärisiert*: Positionen mit $R(x, y) = 0$ werden dabei schwarz dargestellt, Positionen mit $0 < R(x, y) < 0.5$ grau und Positionen mit $R(x, y) \geq 0.5$ weiß. Die Abbildungen 6.4(c) und 6.4(d) zeigen das experimentell erhaltene Resistprofil bzw. die resultierende Struktur in Chrom, wobei für eine ausführliche Beschreibung der durchgeführten Experimente auf Referenz [111] verwiesen wird.

Im Vergleich zwischen den Abbildungen 6.4(b) und 6.4(d) ist eine in vielen Details gute Über-

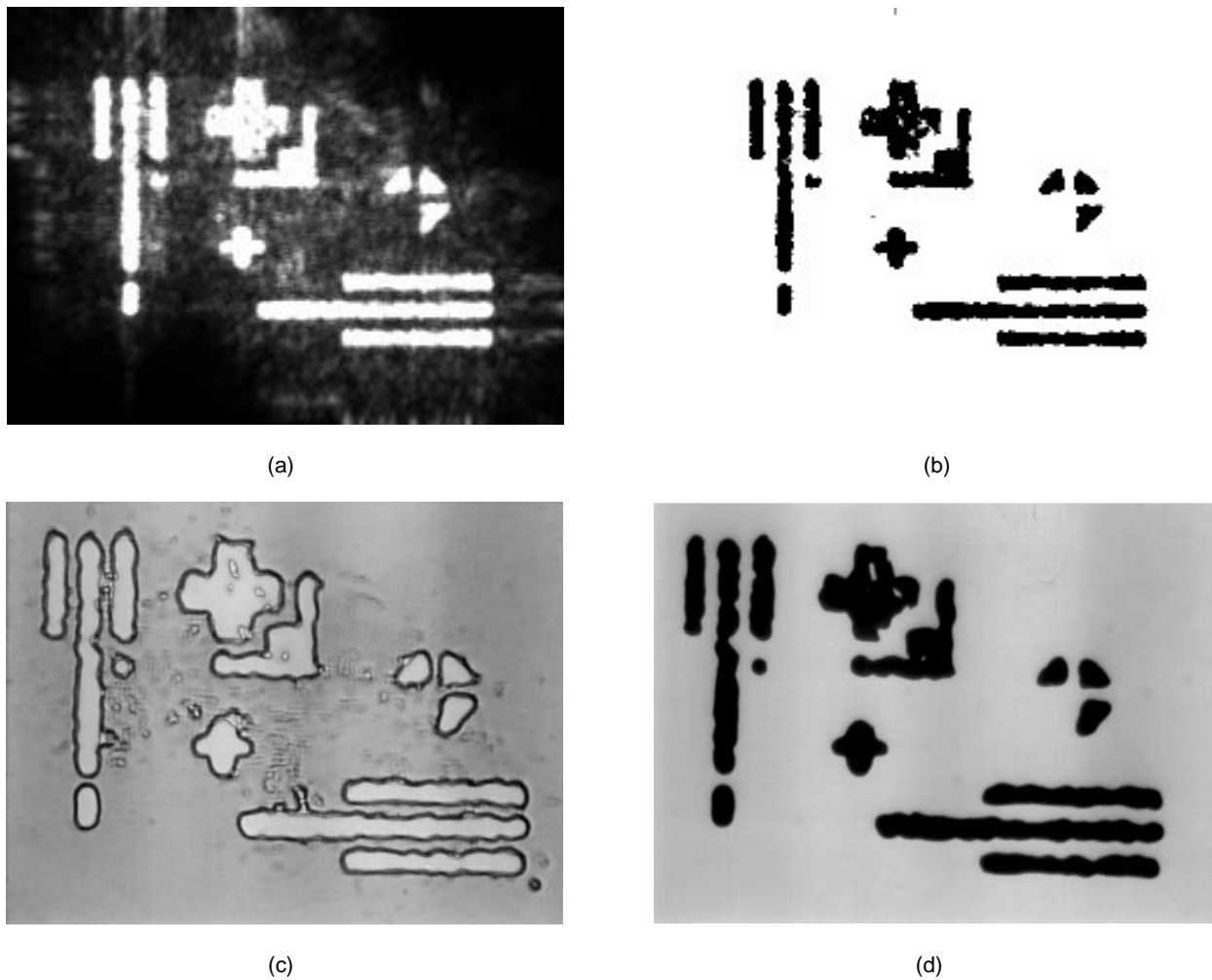


Abbildung 6.4: Ausgangsintensitätsverteilung aus inkohärenter Simulation (a), ternärisiertes simuliertes Resistprofil (b), Transmissionsfotografie des Resistprofils (c) und resultierende Strukturierung in Chrom (d) für das monofunktionale Design aus Abbildung 6.3 bei dem Proximity-Abstand $d_{\text{prox}} = 50 \mu\text{m}$

einstimmung zwischen simulierten und experimentellen Ergebnissen zu erkennen. Viele der in der Struktur von Abbildung 6.4(d) auftretenden Fehler lassen sich auch in Abbildung 6.4(b) erkennen. Als systematischer Unterschied zwischen beiden Abbildungen ist eine leichte *Verschmierung* der Struktur in Abbildung 6.4(d) zu beobachten, welche auf nichtlokale Effekte zurückzuführen ist, die in der Modellierung des Zusammenhangs zwischen Intensitätsverteilung und Resistprofil entsprechend Gleichung (6.2) nicht berücksichtigt werden. Abweichungen zwischen simulierten und experimentellen Ergebnissen können auch darin begründet sein, dass der Proximity-Abstand oder die optimale Belichtungszeit im Experiment nicht ausreichend genau realisiert wurden.

6.1.3 Multifunktionales Design

Wie aus dem Verlauf der ersten Kurve in Abbildung 6.5 ersichtlich wird, reagiert das in Abschnitt 6.1.2 vorgestellte monofunktionale Design empfindlich auf Änderungen des Proximity-Abstands. Innerhalb des praktisch interessanten Toleranzbereichs $d_{\text{prox}} = (50 \pm 10) \mu\text{m}$ ergibt sich für das monofunktionale Design eine deutliche Verschlechterung der resultierenden Struktur, wie auch an den Abbildungen 6.6(a) und 6.7(a), welche für $d_{\text{prox}} = 40 \mu\text{m}$ bzw. $d_{\text{prox}} = 60 \mu\text{m}$

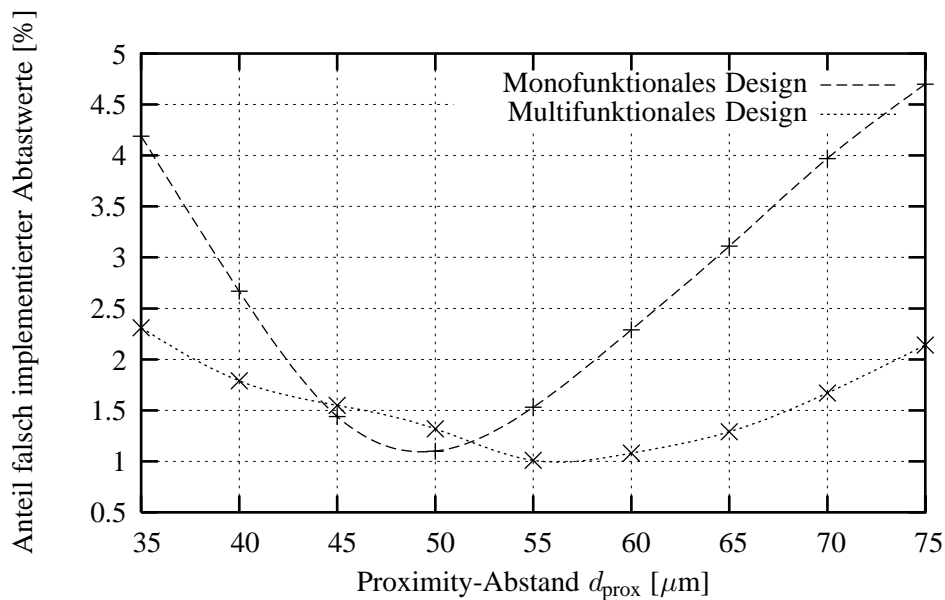


Abbildung 6.5: Abhängigkeit des bei optimal gewählter Belichtungszeit resultierenden Anteils der falsch implementierten Abtastwerte vom Proximity-Abstand für mono- und multifunktionales Design

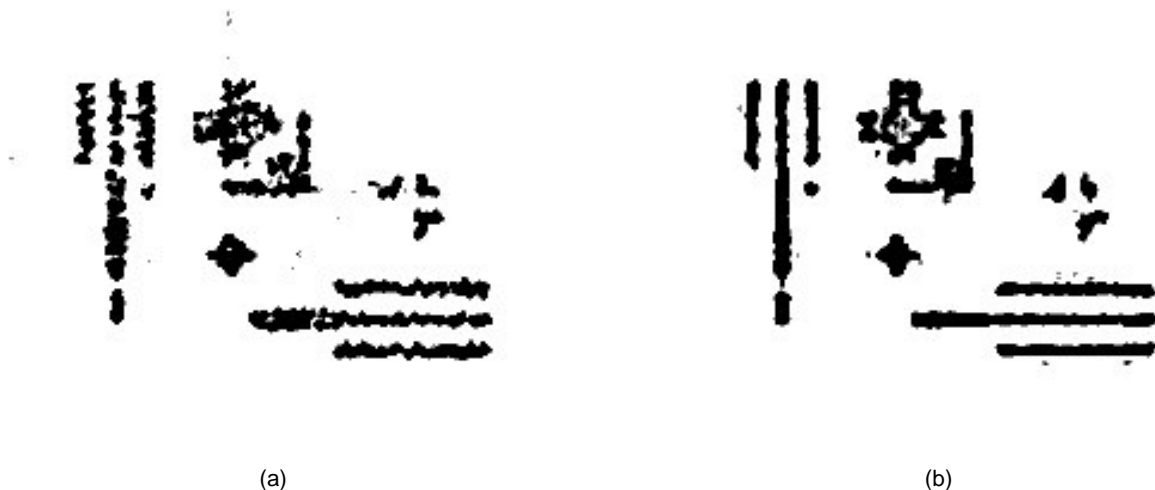


Abbildung 6.6: Ternärisiertes Resistprofil für monofunktionales (a) und multifunktionales Design (b) bei einem Proximity-Abstand $d_{\text{prox}} = 40 \mu\text{m}$

resultieren, im Vergleich zu Abbildung 6.4(b) deutlich wird.

Zur Verbesserung des Toleranzverhaltens wurde entsprechend der multifunktionalen Modellierung aus Abschnitt 6.1.1 ein Design für drei optische Funktionen für die Propagationsabstände $40 \mu\text{m}$, $50 \mu\text{m}$ und $60 \mu\text{m}$ über den PTDA aus Abschnitt 5.2 durchgeführt, bei dem die Mittelwertbildung zwischen den einzelnen optischen Funktionen entsprechend des Schemas aus Gleichung (5.20b) durchgeführt wurde. Dabei wurde wie beim monofunktionalen Design eine Softquantisierung auf zwei Amplituden- und vier Phasenstufen innerhalb von 10 Iterationen durchgeführt. Als Startverteilung wurde eine binäre Amplitudenmodulation entsprechend der Zielstruktur von Abbildung 6.2 gewählt. Abbildung 6.8 zeigt die auf diese Weise erhaltene Transmissionsfunktion. In Abbildung 6.9 sind die simulierte Intensitätsverteilung sowie das ternärisierte Resistprofil für $d_{\text{prox}} = 50 \mu\text{m}$ dargestellt.

Anhand der Abbildungen 6.5 – 6.7 können die Toleranzeigenschaften beider Designs mit-

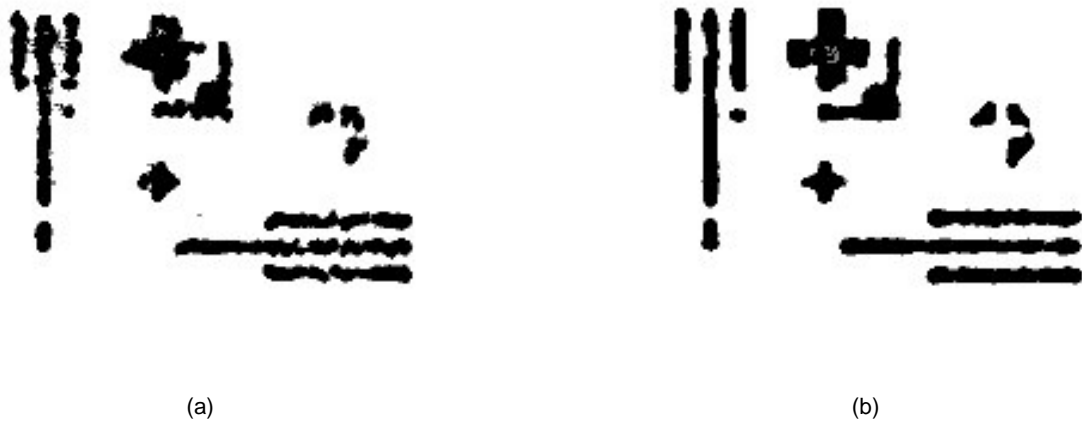
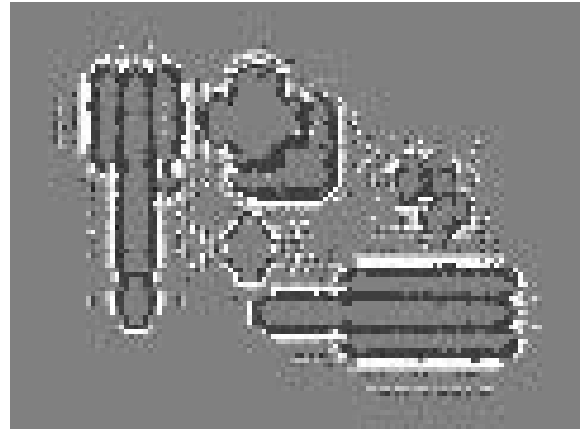


Abbildung 6.7: Ternärisiertes Resistprofil für monofunktionales (a) und multifunktionales Design (b) bei einem Proximity-Abstand $d_{\text{prox}} = 60 \mu\text{m}$



(a)



(b)

Abbildung 6.8: Amplitude (a) und Phase (b) der bei einem multifunktionalen Design resultierenden Transmission

einander verglichen werden. Es zeigt sich, dass die drei optischen Funktionen unterschiedlich gut realisiert werden, wodurch sich für das multifunktionale Design ein optimaler Proximity-Abstand $d_{\text{prox}} \approx 55 \mu\text{m}$ ergibt. Dieser Effekt kann über eine Änderung der den optischen Funktionen zugeordneten Propagationsabstände oder durch eine Priorisierung der optischen Funktionen bei der Mittelwertbildung entsprechend Gleichung (5.20b) bei Bedarf korrigiert werden.

6.2 Toleranzeigenschaften bezüglich globaler Höhenfehler und Wellenlängenfehler

In Abschnitt 3.3 wurden verschiedene Ursachen für globale Höhenfehler bei der Herstellung und Anwendung von Strahlteilern genannt und ihre Auswirkung auf die optische Wirkung von Strahlteilern diskutiert. Laut Abschnitt 3.3.4 entsprechen Wellenlängenabweichungen in ihrer Wirkung auf die Intensitätsverteilung zwischen den Ordnungen des Strahlteilers linearen Höhenskalierungs-



Abbildung 6.9: Ausgangsintensitätsverteilung aus inkohärenter Simulation (a) und ternärisiertes Resistprofil (b) für das multifunktionale Design aus Abbildung 6.8 bei dem Proximity-Abstand $d_{\text{prox}} = 50 \mu\text{m}$

fehlern, wobei jedoch im Fall von Wellenlängenabweichungen zusätzlich eine Skalierung der Signalebenenkoordinaten auftritt.

In diesem Abschnitt möchten wir anhand eines speziellen zweidimensionalen Strahlteilungsproblems demonstrieren, inwieweit die in dieser Arbeit vorgestellten Designtechniken zur Berechnung von Strahlteilern eingesetzt werden können, welche ein optimiertes Toleranzverhalten gegenüber globalen Höhenfehlern oder Wellenlängenabweichungen besitzen. Nach Betrachtung der verwendeten Designmethoden in Abschnitt 6.2.1 untersuchen wir in Abschnitt 6.2.2 die resultierenden Toleranzeigenschaften anhand einer Beispielanwendung, wobei experimentelle und numerisch simulierte Ergebnisse verglichen und diskutiert werden.

6.2.1 Designmethoden

Laut Abschnitt 3.3.2 kann durch Multiplikation eines konstanten Phasenterms $\exp(i\phi)$ zur Transmissionsfunktion $t(x)$ eines Strahlteilers mit $0 \in \mathcal{W}_{\text{sig}}$ in vielen Fällen eine deutliche Verbesserung des Toleranzverhaltens gegenüber linearen Höhenfehlern erreicht werden. Durch einen zusätzlichen Designschritt, in welchem der optimale Phasenterm durch einen einfachen Suchalgorithmus ermittelt wird, können daher beispielsweise die Designverfahren aus den Abschnitten 3.1.4 und 4.3.2 hinsichtlich einer zusätzlichen Optimierung der Toleranzeigenschaften erweitert werden. Eine weitere im Folgenden betrachtete Möglichkeit zur Toleranzoptimierung stellt die Verwendung eines multifunktionalen Designverfahrens dar.

Globale Höhenfehler stellen einen Spezialfall der entsprechend Gleichung (2.49) durch einen Operator E_β beschriebenen Umsetzungsfehler dar. Obwohl bei der Beschreibung der beiden multifunktionalen Designverfahren in Kapitel 5 von der vereinfachten Modellierung (2.50) ausgegangen wird, welche $E_\beta \equiv 1$ impliziert, können diese Algorithmen auch für die Optimierung der Toleranzeigenschaften überlappungsfreier Strahlteiler hinsichtlich globaler Höhenfehler übertragen werden, da in diesem Fall beim Design von $u_{\text{in},\beta} \equiv 1$ ausgegangen werden kann. Gleichung (2.49)

vereinfacht sich somit zu

$$u_{\text{out},\beta}(x') = \underbrace{(P_\beta E_\beta t(x))}_{= P'_\beta}(x'), \quad (6.3)$$

was bei Substitution des modifizierten Propagationsoperators

$$P'_\beta = P_\beta E_\beta \quad (6.4)$$

exakt dem Schema von Gleichung (2.50) entspricht und somit eine direkte Anwendung der Algorithmen aus Kapitel 5 ermöglicht.

Bei Annahme eines linearen Höhenfehlers $\gamma_{\text{lin},\beta}$ für die optische Funktion mit Index β ergibt sich

$$E_\beta t = t^{\gamma_{\text{lin},\beta}}. \quad (6.5)$$

Dabei kann $\gamma_{\text{lin},\beta}$ entsprechend der Gleichungen (2.41) und (2.44) auch aus einer Abweichung zwischen der für die optische Funktion angenommenen Wellenlänge λ_β und der beim Strukturdesign verwendeten Referenzwellenlänge λ_{ref} resultieren. Unter Vernachlässigung der Dispersionsbeziehung $n(\lambda)$ folgt in diesem Fall die Beziehung

$$\gamma_{\text{lin},\beta} = \frac{\lambda_{\text{ref}}}{\lambda_\beta}. \quad (6.6)$$

Neben linearen Höhenfehlern sind für mehrstufige Elemente auch globale Höhenfehler von Bedeutung, welche bei Anwendung der *binären Maskentechnik* resultieren. Dieser Begriff steht für Herstellungsverfahren, bei denen Höhenprofile mit $Q = 2^q$ Stufen in q Herstellungsschritten realisiert werden, wobei in jedem Schritt eine binäre Strukturierung aufgebracht wird, deren Tiefe jeweils pro Schritt halbiert (oder verdoppelt) wird. Auch für Höhenfehler, die bei Anwendung der binären Maskentechnik entstehen können, ist es möglich, Operatoren E_β für die einzelnen optischen Funktionen zu konstruieren, welche eine Abbildung zwischen den beim Design angenommenen Phasenquantisierungswerten und den auf Grund der Höhenfehler verfälschten Phasenwerten bewirken.

Die relative Genauigkeit der Höhenrealisierung ist besonders in dem Maskenschritt mit der höchsten Strukturierungstiefe für die optische Wirkung relevant. Den Maskenschritt mit der größten Strukturierungstiefe, welche bei idealer Realisierung einer Phasenmodulation um π entspricht, bezeichnen wir im Folgenden mit Schritt #1. Schritt # j entspricht demnach einer Phasenmodulation um π/j .

Die in Kapitel 5 geführten Konvergenzbeweise basierten unter anderem auf der Gültigkeit von Satz 2.67, welcher die Erfüllung der Eigenschaften (2.39) für die verwendeten Propagationsoperatoren verlangt. Falls ein Propagationsoperator P_β diese beiden Eigenschaften erfüllt, folgt für den entsprechend Gleichung (6.4) konstruierten Operator P'_β die Gültigkeit der Eigenschaft (2.39a), sofern E_β nur auf die Phase von t wirkt. Aufgrund der für Exponenten kleiner eins auftretenden

Vieldeutigkeit der Exponierung im Bereich der komplexen Zahlen kann jedoch für E_β nach Gleichung (6.5) für $\gamma_{\text{lin},\beta} \neq 1$ eine eindeutige Invertierbarkeit des Propagationsoperators P'_β entsprechend Eigenschaft (2.39b) nicht mehr gefolgert werden. Da Höhenfehler aus der binären Maskentechnik eine Verallgemeinerung von linearen Höhenfehlern darstellen, kann von einer Erfüllung der Eigenschaft (2.39b) auch in diesem Fall nicht ausgegangen werden.

Obwohl somit die Konvergenz der multifunktionalen Designalgorithmen aus Kapitel 5 nicht sichergestellt ist, kann ihr Einsatz beim Design von gegen Höhenfehlern toleranten Strahlteilern dennoch sinnvoll sein, wie die folgenden Betrachtungen zeigen. Wir untersuchen die Anwendung von drei verschiedenen Designalgorithmen beim Design eines speziellen zweidimensionalen 1-zu-21-Strahlteilers, dessen Signalordnungen ein Quadrat mit fehlenden Eckpunkten beschreiben, welches sich in x - und in y -Richtung jeweils zwischen der -2. und +2. Ordnung erstreckt (vgl. Abbildung 6.11). Die Designs werden auf einem Raster von 128×128 Abtastwerten für acht äquidistante Phasenquantisierungswerte unter Ausnutzung von Phasen-, Amplituden und Skalierungsfreiheit durchgeführt.

Für das genannte Strahlteilungsproblem betrachten wir im Folgenden zwei Designs mit jeweils verschiedenen Toleranzanforderungen: Das ist zum einen die Forderung nach einer Insensitivität gegenüber Höhenfehlern bei der binären Maskentechnik. Die multifunktionale Modellierung erfolgt hierbei über fünf optische Funktionen, wovon eine für den Fall einer idealen Höhenrealisierung steht und die anderen vier dem Auftreten eines Höhenskalierungsfehlers von $\pm 3\%$ entweder in Maskenschritt #1 oder in Schritt #2 entsprechen.

Als Beispiel für Insensitivität bezüglich Wellenlängenabweichungen betrachten wir die Anforderung, dass ein für $\lambda_{\text{ref}} = 632.8$ nm realisierter Strahlteiler in einem Wellenlängenbereich von 594 nm bis 650 nm eine gute Uniformität aufweisen soll. Hierbei verwenden wir eine Modellierung mit drei optischen Funktionen für die Wellenlängen 594 nm, 632.8 nm und 650 nm, wobei die korrespondierenden Skalierungsfaktoren $\gamma_{\text{lin},\beta}$ entsprechend Gleichung (6.6) berechnet werden.

Das erste betrachtete Designverfahren entspricht dem Ablaufschema aus Abbildung 4.5 unter Auslassung des Softquantisierungsschritts. Für Signalphasensynthese, SNR-Optimierung und Iteration mit verstärkten Projektionen werden 20, 100 bzw. bis zu 1000 Iterationen verwendet, wobei die letztgenannte Iterationsphase entsprechend der Diskussion am Ende von Abschnitt 4.3.2 abgebrochen wird, sobald kein Konvergenzbereich mehr gefunden werden kann.

In dem zweiten untersuchten Designverfahren wird zusätzlich zu den Iterationsschritten des ersten Verfahrens eine Suche nach einem die Toleranzeigenschaften optimierenden konstanten Phasenterm durchgeführt, wobei als Suchkriterium der innerhalb des betrachteten Toleranzbereichs auftretende maximale Uniformitätsfehler verwendet wird. Dabei werden im Sinne einer Rechenzeit sparenden Approximation nur die für die oben beschriebenen optischen Funktionen resultierenden Uniformitätsfehler betrachtet.

Das dritte betrachtete Designverfahren ist identisch mit dem zweiten bis auf Anwendung des PTDA in den Iterationsphasen der SNR-Optimierung sowie der Anwendung verstärkter Projektionen, wobei die oben beschriebene multifunktionale Modellierung verwendet wird.

Abbildung 6.10 zeigt die bei Durchführung von jeweils 100 Designs für die drei betrachteten

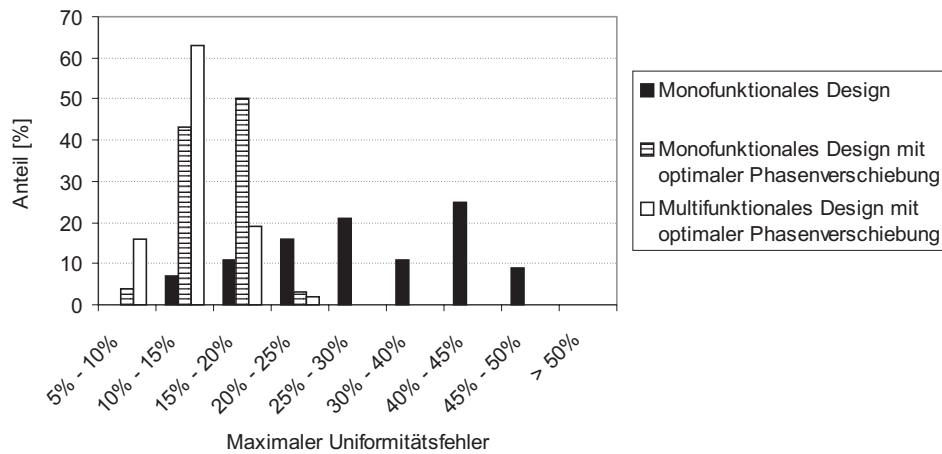


Abbildung 6.10: Verteilung des innerhalb des Wellenlängenbereichs $\lambda \in [594 \text{ nm}, 650 \text{ nm}]$ erreichten maximalen Uniformitätsfehlers. Es werden die drei im Text beschriebenen Designstrategien untersucht, wobei zur Ermittlung der Ergebnisverteilung jeweils 100 Designs mit zufällig gewählten Startverteilungen durchgeführt werden.

Designverfahren resultierende Ergebnisverteilung hinsichtlich des innerhalb des Wellenlängenbereichs $\lambda \in [594 \text{ nm}, 650 \text{ nm}]$ auftretenden maximalen Uniformitätsfehlers. Es zeigt sich, dass schon durch Multiplikation eines toleranzoptimierenden Phasenterms eine deutliche Verbesserung der Ergebnisverteilung erreicht werden kann, wobei die Hinzunahme der beschriebenen multifunktionalen Optimierung eine weitere Verbesserung der Ergebnisverteilung ermöglicht und somit ein deterministischeres Auffinden toleranzoptimierter Lösungen ermöglicht.

6.2.2 Vergleich zwischen simulierten und experimentellen Ergebnissen

In diesem Abschnitt betrachten wir drei verschiedene Lösungen für das im vorangegangenen Abschnitt beschriebene 1-zu-21-Strahlteilungsproblem. Design A stellt hierbei eine effizienzoptimierte Lösung des monofunktionalen Designs ohne Berücksichtigung von Toleranzeigenschaften dar. Design B wurde über ein multifunktionales Design mit anschließender Aufmultiplikation eines optimalen Phasenfaktors für die im vorangegangenen Abschnitt beschriebenen Toleranzanforderungen bezüglich Höhenfehlern bei der binären Maskentechnik durchgeführt. Zur Ermittlung von Design C wurde dieselbe Designmethode zur Erlangung eines gegenüber Wellenlängenfehlern toleranten Designs angewandt, wobei Wellenlängenbereich und multifunktionale Modellierung der im vorangegangenen Abschnitt gegebenen Spezifikation entsprechen. Für die Designs B und C wurden jeweils Lösungen mit besonders guten Toleranzeigenschaften und gleichzeitig relativ guten Effizienzen ausgewählt.

Tabelle 6.1 zeigt in der 5. – 7. Spalte die für den Fall einer idealen pixelierten Realisierung der drei Designs resultierenden Werte für Fenstereffizienz η_{win} , Uniformitätsfehler E_{unif} bzw. relativer Intensität der nullten Ordnung I_0 . Bei der Simulation der Fenstereffizienz wurden hierbei die bei zwei unbeschichteten Grenzflächen entstehenden Reflexionsverluste berücksichtigt, was in diesem Fall zu einer Effizienzreduktion um etwa 6 % führt.

Die drei Strahlteilerdesigns wurden jeweils lithografisch unter Verwendung der binären Mas-

Design	Höhenfehler Maskenschritt			Simulation ohne Höhenfehler			Simulation mit Höhenfehlern			Labor- messungen		
	#1	#2	#3	η_{win} [%]	E_{unif} [%]	I_0	η_{win} [%]	E_{unif} [%]	I_0	η_{win} [%]	E_{unif} [%]	I_0
A	1.05	1.04	1.02	82.4	0.7	0.99	82.2	30.7	1.62	82.6	36.4	1.71
B	1.05	1.05	0.98	81.3	0.4	0.99	81.0	9.1	1.07	80.9	8.9	1.12
C	1.01	1.00	1.01	78.7	3.4	0.95	78.7	2.4	0.97	78.6	6.7	0.93

Tabelle 6.1: Simulierte und gemessene Werte für Fenstereffizienz η_{win} , Uniformitätsfehler E_{unif} und relative Intensität der nullten Ordnung I_0 für die drei betrachteten Designs. Die Simulationen und Messungen wurden für die beim Design angenommene Referenzwellenlänge $\lambda_{\text{ref}} = 632.8 \text{ nm}$ durchgeführt.

kentechnik mit einer Pixelgröße von $\delta x = \delta y = 3 \mu\text{m}$ in Quarz hergestellt. Die dabei für die einzelnen Maskenschritte aufgetretenen Höhenfehler sind in den Spalten 2 – 4 der Tabelle 6.1 als Quotienten zwischen Ist- und Soll-Höhe angegeben. Die angegebenen Messungen wurden nach jedem Maskenschritt an genau einem Messpunkt innerhalb jedes Strahlteilers durchgeführt und besitzen eine relative Genauigkeit von ca. 5 %. Unter Einbeziehung dieser Höhenfehler in die Simulation resultieren die in den Spalten 6 – 8 angegebenen Werte für η_{win} , E_{unif} und I_0 .

Die drei letzten Spalten von Tabelle 6.1 enthalten die für die hergestellten Strahlteilerelemente gemessenen Ergebnisse. Zur Ermittlung dieser Werte wurde die absolute Leistung in den einzelnen Signalordnungen sowie die Leistung des Laserstrahls über ein Photometer mit einer geeigneten Lochblende gemessen. Dabei wurde für jede Leistungsmessung die Blendenposition zur Maximierung der gemessenen Leistung feinjustiert, was zu einem zufälligen relativen Fehler für die einzelnen Leistungsmessungen von etwa 2 % führt.

Die in Tabelle 6.1 auftretenden Unterschiede zwischen gemessenen und simulierten Ergebnisse beschränken sich primär auf die Werte für E_{unif} und I_0 . Als Ursache hierfür kommen neben den im vorangegangenen Abschnitt diskutierten Messfehlern vor allem Herstellungsfehler in Betracht. Aufgrund der endlichen Überdeckungsgenauigkeit zwischen den einzelnen Maskenschritten entstehen in der Höhenstruktur ungewollte Stege mit kleiner lateraler Ausdehnung, welche die optische Wirkung verfälschen können. Die relativen Intensitäten der einzelnen Signalordnungen wird auch durch die nicht ideale Realisierung der Flankenwinkel der Strukturen beeinflusst. Eine Hauptursache für die beobachteten Abweichungen zwischen experimentellen und simulierten Werten für I_0 stellt die inhomogene laterale Verteilung der Ätztiefe innerhalb der einzelnen Maskenschritte dar, deren Auftreten auch daran erkennbar ist, dass die drei Strahlteiler trotz gemeinsamer Herstellung auf einer Quarzplatte die in Tabelle 6.1 angegebenen unterschiedlichen Höhenfehler aufweisen.

Abbildung 6.11 zeigt die mit einer CCD-Kamera aufgenommene Intensitätsverteilung für den nach Design A hergestellten Strahlteiler. Der auf Grund von Herstellungsungenauigkeiten entstandene Uniformitätsfehler von etwa 36 % ist visuell in diesem Fall kaum sichtbar. Die teilweise zwischen den auf einem äquidistanten Raster liegenden Ordnungen beobachtbaren zusätzlichen Spots resultieren aus Mehrfachreflexionen innerhalb des unbeschichteten Strahlteilerelements.

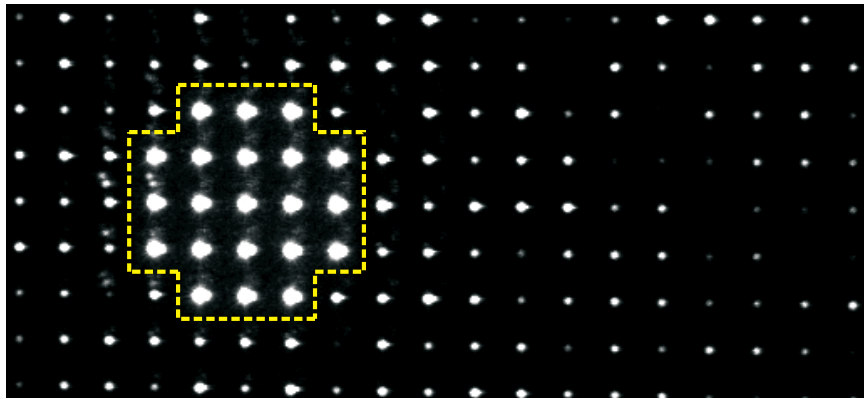


Abbildung 6.11: Mit einer CCD-Kamera aufgenommene Intensitätsverteilung für den entsprechend Design A hergestellten Strahlteiler. Innerhalb des schraffiert markierten Bereichs befinden sich die 21 Signalordnungen.

In Abbildung 6.12 ist die für die Designs A und B resultierende Verteilung des Uniformitätsfehlers in Abhängigkeit von Höhenfehlern in den Maskenschritten #1 und #2 dargestellt. Die von den einzelnen Konturlinien umschlossene Fläche stellt ein Maß für die Toleranzeigenschaften des jeweiligen Designs dar. Für das Design B ist somit eine deutliche Verbesserung der Toleranzeigenschaften im Vergleich zu Design A erkennbar.

Da bei der Herstellung der Strahlteilerelemente für die Designs A und B entsprechend Tabelle 6.1 in den hauptsächlich relevanten Maskenschritten #1 und #2 etwa gleiche Höhenfehler entstanden sind, ist eine gute Vergleichbarkeit der für beide Strahlteiler gewonnenen experimentellen Ergebnisse gewährleistet. Es ergibt sich somit entsprechend Tabelle 6.1 durch Anwendung eines toleranzoptimierten Designs in diesem Beispiel eine praktische Verbesserung des Uniformitätsfehlers von 36.4 % auf 8.9 % bei einer Effizienzeinbuße von 1.5 %. Bei einer genaueren Realisierung der Profilhöhen in den einzelnen Maskenschritten kann auch ein deutlich niedrigerer Uniformitätsfehler praktisch erreicht werden.

Für das hinsichtlich der Toleranzeigenschaften gegenüber Wellenlängenabweichungen optimierte Design C sind in Abbildung 6.13 die simulierten Verläufe $E_{\text{unif}}(\lambda)$ und $I_0(\lambda)$ mit den für die Wellenlängen 594 nm (He-Ne-Laser gelb), 632.8 nm (He-Ne-Laser rot) und 650 nm (Laserdiodenmodul PM03/5384 von Power Technology Inc.) gemessenen Werten gegenübergestellt, wobei eine sehr gute Übereinstimmung zwischen den unter Berücksichtigung der aufgetretenen Höhenfehler gewonnenen Simulationsergebnissen und den Messresultaten besteht.

Die aus einer Simulation unter Annahme einer fehlerfreien Realisierung für das Design A resultierenden Kurven sind ebenfalls in Abbildung 6.13 gezeigt und verdeutlichen die für das Design C im Vergleich zu Design A deutlich verbesserten Toleranzeigenschaften hinsichtlich Wellenlängenabweichungen: Während für das Design A in dem betrachteten Wellenlängenbereich Uniformitätsfehler von über 35 % simuliert werden, liegt der für Design C auftretende Maximalwert unter 8 %. Aufgrund der oben diskutierten Herstellungsfehler liegen die für Design C tatsächlich gemessenen Uniformitätsfehler von bis zu 14.3 % etwas über den simulierten Werten, wobei jedoch selbst diese Messergebnisse noch eine klare Verbesserung des Toleranzverhaltens gegenüber den

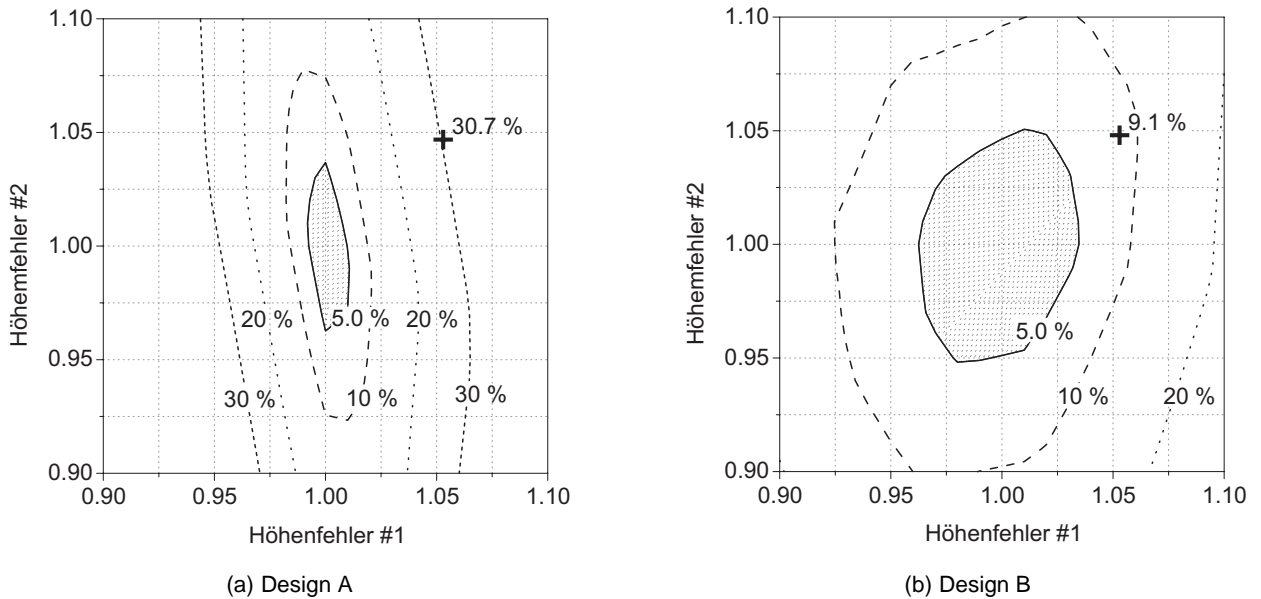
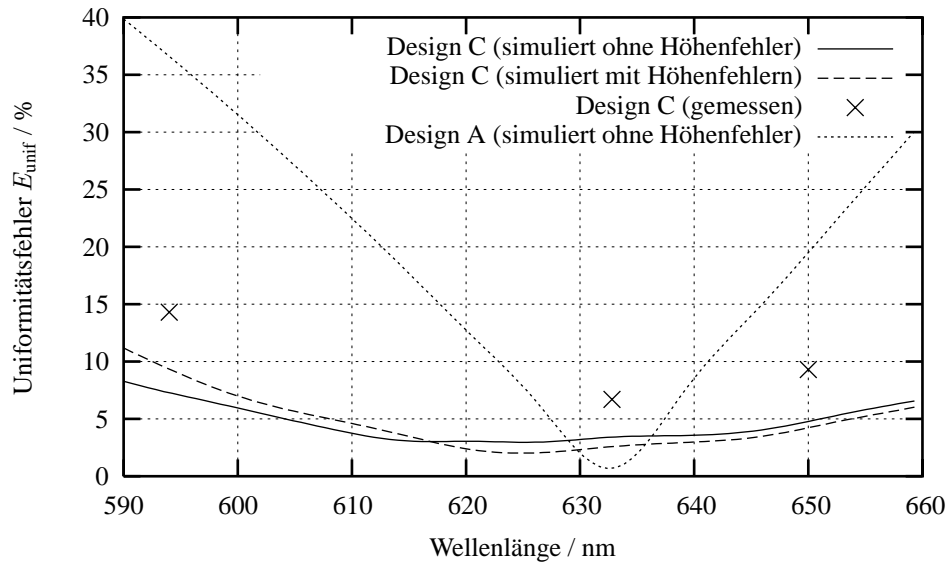


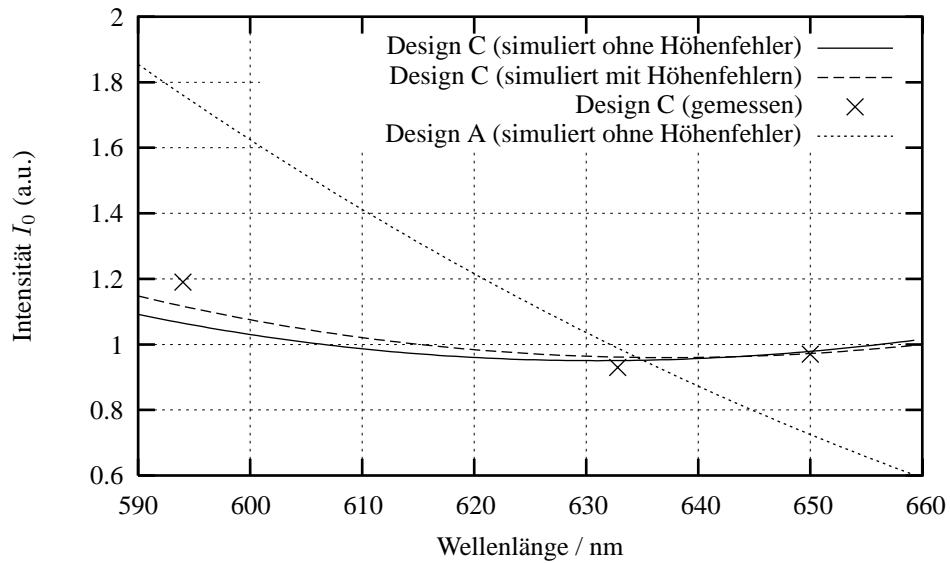
Abbildung 6.12: Verteilung des Uniformitätsfehlers in Abhängigkeit von Höhenfehlern in den Maskenschritten #1 und #2. Der schraffierte Bereich entspricht dabei dem Wertebereich beider Höhenfehler, innerhalb dessen ein Uniformitätsfehler von weniger als 5 % resultiert. Die Kreuze markieren jeweils die Positionen in beiden Diagrammen, welche den bei der Herstellung der Strahlteiler entstandenen Höhenfehlern entsprechen.

Simulationsergebnissen für Design A darstellen. Da bei der praktischen Realisierung von Design A entsprechend Tabelle 6.1 deutliche Höhenfehler aufgetreten sind, können hiermit erzielte Messergebnisse nicht sinnvoll in diesen Vergleich einbezogen werden.

In Übereinstimmung mit der Diskussion der Kurven $I_0(\gamma_{\text{lin}})$ in Abbildung 3.10 wird in Abbildung 6.13(b) wiederum deutlich, dass für die Erreichung guter Toleranzeigenschaften bei Strahlteilern mit $0 \in \mathcal{W}_{\text{sig}}$ der Verlauf $I_0(\lambda)$ innerhalb des betrachteten Toleranzbereichs ein Minimum besitzen sollte. Für das Design C liegt dieser Minimalwert von I_0 sogar unter dem Sollwert, was den für $\lambda = \lambda_{\text{ref}}$ simulierten Uniformitätsfehler von 3.4 % erklärt, welcher auf Kosten eines höheren Uniformitätsfehlers an den Grenzen des betrachteten Bereichs vermeidbar gewesen wäre.



(a)



(b)

Abbildung 6.13: Abhängigkeit zwischen Uniformitätsfehler (a) und skaliertes Intensität der nullten Ordnung (b) simuliert für die Designs A und C im Vergleich zu den an drei diskreten Wellenlängen für Design C gemessenen Werten

Kapitel 7

Zusammenfassung

In der vorliegenden Arbeit wurde der Einsatz von Projektionsalgorithmen im wellenoptischen Transmissionsdesign untersucht, wobei der Schwerpunkte entsprechend der Aufteilung in die Kapitel 3 – 6 gesetzt wurden auf

- das Design diffraktiver Strahlteiler,
- die Verbesserung von Konvergenzeigenschaften durch Anwendung verstärkter Projektionen,
- die Entwicklung multifunktionaler Designstrategien sowie
- die Verbesserung von Toleranzeigenschaften über multifunktionale Optimierung.

Die wichtigsten Ergebnisse dieser Arbeit können in den folgenden Punkten zusammengefasst werden:

1. Design diffraktiver Strahlteiler für beliebig positionierte Signalordnungen: Nachdem in Abschnitt 3.1 das im Rahmen dieser Arbeit ergänzte Konzept zur Modellierung diffraktiver Strahlteilungsprobleme dargestellt wurde, wobei in Abschnitt 3.1.3 insbesondere die für die Wahl der Abtastparameter relevanten Kriterien diskutiert wurden, erfolgte in Abschnitt 3.2 eine schrittweise Erweiterung dieses Konzepts hinsichtlich einer Aufhebung der für die Positionierung der Signalordnungen bestehenden Einschränkungen. Es wurde dabei zunächst in Abschnitt 3.2.1 der Übergang zu nichtüberlappenden Strahlteilern diskutiert, bevor in Abschnitt 3.2.2 ein neues Designverfahren entwickelt wurde, welches die völlig freie Positionierung der Signalordnungen erlaubt. Da dieses Verfahren ebenfalls auf Projektionsalgorithmen basiert und den Einsatz von schnellen Fouriertransformationen in einem Großteil seiner Iterationen erlaubt, bleibt der damit verbundene Rechenaufwand in einem Rahmen, der vielfältige praktische Anwendungen erlaubt. Ein derart berechneter Strahlteiler wurde entsprechend Abschnitt 3.2.3 auch experimentell überprüft.

2. Toleranzeigenschaften von Strahlteilern: Die Toleranzeigenschaften von Strahlteilern wurden in Abschnitt 3.3 diskutiert. Insbesondere erfolgte in den Abschnitten 3.3.1 und 3.3.2 eine Untersuchung der Fragestellung, in welchem Maße eine Insensitivität hinsichtlich globaler Höhenfehler idealerweise erreicht werden kann. In Abschnitt 6.2 werden die im Rahmen dieser Arbeit

entwickelten Designansätze verglichen, wobei die für die betrachteten Beispielanwendungen berechneten und hergestellten Strahlteiler in Abschnitt 6.2.2 auch experimentell verifiziert wurden.

3. Begrenzbarkeit der maximalen Intensität von Rauschordnungen: In Abschnitt 3.4 wurden Ansätze zur Begrenzung der maximalen Intensität der Rauschordnungen demonstriert. Die Anwendbarkeit der gezeigten Methoden wurde in Abschnitt 3.4.4 auch experimentell nachgewiesen.

4. Ausnutzung verstärkter Projektionen zur Verbesserung von Konvergenzeigenschaften: In Kapitel 4 wurden Möglichkeiten zur Verbesserung der Konvergenzeigenschaften von Projektionsalgorithmen untersucht und zwei Verfahren zur Steuerung der Projektionsstärke während der Iteration entwickelt. Die in dem für die beiden Verfahren jeweils geeigneten Anwendungsbereich typischerweise resultierende Verbesserung der Konvergenzeigenschaften wurde anhand von Beispielanwendungen demonstriert.

5. Verfahren zum Multifunktionalen Transmissionsdesign: In Abschnitt 5.1 wurde gezeigt, wie mit dem parallelen Multiprojektionsalgorithmus ein bekannter Grundansatz für die Konstruktion von Projektionsalgorithmen zur Lösung multifunktionaler Transmissionsdesignprobleme angepasst werden kann. In Abschnitt 5.2 wurde mit dem parallelen Transmissionsdesignalgorithmus ein hierzu alternatives Verfahren entwickelt, dessen Vorteile gegenüber dem parallelen Multiprojektionsalgorithmus in den Abschnitten 5.2.1 und 5.3 gezeigt wurden.

6. Proximity Printing: In Abschnitt 6.1 wird mit der Berechnung komplexwertiger Transmissionsmasken für die Fotolithografie eine weitere experimentell verifizierte Beispielanwendung für den Einsatz von Projektionsalgorithmen im Optikdesign betrachtet. Es wurde dabei gezeigt, dass durch den Einsatz wellenoptisch berechneter Belichtungsmasken mit komplexwertigen Transmissionsfunktionen eine deutliche Verbesserung der erzielbaren Auflösung gegenüber konventionellen binären Amplitudenmasken erreicht werden kann. In Abschnitt 6.1.3 wurde weiterhin die Möglichkeit einer Verringerung der Sensitivität gegen Abweichungen des Proximity-Abstands über den Einsatz des parallelen Transmissionsdesignalgorithmus nachgewiesen.

Anhang A

Beweise und Ableitungen

A.1 Zu Kapitel 2

A.1.1 Beweis zu Satz 2.16

Satz 2.16 Algorithmus (2.14) konvergiert bezüglich des SDEs aus Gleichung (2.11), d. h. für $n > 1$ gilt die Konvergenzbedingung (2.15).

Beweis: Aus Gleichung (2.14) folgt für $n > 0$

$$u^{(n)} \in \mathcal{M}_2. \tag{A.1}$$

Entsprechend Gleichung (2.11) erhalten wir

$$J(u^{(n-1)}) = d(\Pi_1 u^{(n-1)}, u^{(n-1)}) + d(\Pi_2 u^{(n-1)}, u^{(n-1)}), \tag{A.2a}$$

$$\begin{aligned} J(\Pi_1 u^{(n-1)}) &= d(\Pi_1 \Pi_1 u^{(n-1)}, \Pi_1 u^{(n-1)}) + d(\Pi_2 \Pi_1 u^{(n-1)}, \Pi_1 u^{(n-1)}) \\ &= d(\Pi_2 \Pi_1 u^{(n-1)}, \Pi_1 u^{(n-1)}) \quad [\text{wegen Gleichung (2.8)}] \end{aligned} \tag{A.2b}$$

sowie

$$\begin{aligned} J(u^{(n)}) &= d(\Pi_1 \Pi_2 \Pi_1 u^{(n-1)}, \Pi_2 \Pi_1 u^{(n-1)}) + d(\Pi_2 \Pi_2 \Pi_1 u^{(n-1)}, \Pi_2 \Pi_1 u^{(n-1)}) \\ &= d(\Pi_1 \Pi_2 \Pi_1 u^{(n-1)}, \Pi_2 \Pi_1 u^{(n-1)}). \end{aligned} \tag{A.2c}$$

Betrachten wir zunächst den ersten Summanden auf der rechten Seite von Gleichung (A.2a) sowie Gleichung (A.2b). Aufgrund der Abstandsminimierungseigenschaft (2.3) des Projektionsoperators Π_2 , welcher auf das Element $\Pi_1 u^{(n-1)}$ angewendet wird, folgt mit Gleichung (A.1) für $n > 1$

$$d(\Pi_2 \Pi_1 u^{(n-1)}, \Pi_1 u^{(n-1)}) \leq d(\Pi_1 u^{(n-1)}, u^{(n-1)}) \tag{A.3}$$

und damit

$$J(\Pi_1 u^{(n-1)}) \leq J(u^{(n-1)}). \tag{A.4}$$

Betrachten wir nun die Gleichungen (A.2b) und (A.2c). Aufgrund der Abstandsminimierungseigenschaft (2.3) des Projektionsoperators Π_1 , welcher auf das Element $\Pi_2 \Pi_1 u^{(n-1)}$ angewendet wird, folgt mit $\Pi_1 u^{(n-1)} \in \mathcal{M}_1$ für $n > 1$

$$J(u^{(n)}) \leq J(\Pi_1 u^{(n-1)}). \quad (\text{A.5})$$

Zusammen mit Gleichung (A.4) folgt somit die Gültigkeit von Satz 2.16. ■

A.1.2 Beweis zu Satz 2.67

Satz 2.67 *Der Operator Π_1 nach Gleichung (2.66a) definiert eine Projektion auf \mathcal{M}_1 gemäß Gleichung (2.65a), sofern der Propagationsoperator P die Eigenschaften (2.39) erfüllt.*

Beweis: Wir betrachten hierzu ein Element $u' \in \mathcal{M}_1$ und treffen für einen indirekten Beweis die Annahme

$$\|u - u'\| < \|u - \Pi_1 u\|. \quad (\text{A.6})$$

Wegen $u' \in \mathcal{M}_1$ existiert ein Element $u'_{\text{sig}} \in \mathcal{M}_{\text{sig}}$ mit

$$u' = P^{-1} u'_{\text{sig}}. \quad (\text{A.7})$$

Durch Anwendung der Gleichungen (2.39) ergibt sich somit

$$\|u - u'\| = \gamma^{-1} \|Pu - u'_{\text{sig}}\| \quad \text{und} \quad (\text{A.8a})$$

$$\|u - \Pi_1 u\| = \gamma^{-1} \|Pu - \Pi_{\text{sig}} Pu\|. \quad (\text{A.8b})$$

Aufgrund der Abstandsminimierungseigenschaft (2.3) des Projektionsoperators Π_{sig} , welcher auf das Element Pu angewendet wird, folgt mit $u'_{\text{sig}} \in \mathcal{M}_{\text{sig}}$

$$\|Pu - \Pi_{\text{sig}} Pu\| \leq \|Pu - u'_{\text{sig}}\|. \quad (\text{A.9})$$

Wegen $\gamma > 0$ resultiert somit aus den Gleichungen (A.8) und (A.9) im Gegensatz zur Annahme (A.6)

$$\|u - \Pi_1 u\| \leq \|u - u'\|. \quad (\text{A.10})$$

■

A.1.3 Beweis zu Satz 2.68

Satz 2.68 *Der Operator Π_2 nach Gleichung (2.66b) definiert eine Projektion auf \mathcal{M}_2 gemäß Gleichung (2.65b), sofern Π_{tran} ein punktweise abstandsminimierender Projektionsoperator auf $\mathcal{M}_{\text{tran}}$ ist.*

Beweis: Wir betrachten hierzu ein Element $u' \in \mathcal{M}_2$ und treffen für einen indirekten Beweis die Annahme

$$\|u - u'\| < \|u - \Pi_2 u\|. \quad (\text{A.11})$$

Wegen $u' \in \mathcal{M}_2$ existiert ein Element $t' \in \mathcal{M}_{\text{tran}}$ mit

$$u' = u_{\text{in}} t'. \quad (\text{A.12})$$

Unter Ausnutzung der Gleichungen (A.12) und (2.66b) erhalten wir

$$\|u - u'\| = \|u_{\text{in}}(u_{\text{in}}^{-1} u - t')\| \quad \text{und} \quad (\text{A.13a})$$

$$\|u - \Pi_2 u\| = \|u_{\text{in}}(u_{\text{in}}^{-1} u - \Pi_{\text{tran}} u_{\text{in}}^{-1} u)\|. \quad (\text{A.13b})$$

Aus $t' \in \mathcal{M}_{\text{tran}}$ folgt

$$\|u_{\text{in}}^{-1} u - \Pi_{\text{tran}} u_{\text{in}}^{-1} u\| \leq \|u_{\text{in}}^{-1} u - t'\|. \quad (\text{A.14})$$

Da Π_{tran} laut Voraussetzung punktweise abstandsminimierend ist, folgt mit Gleichung (2.7) und $t' \in \mathcal{M}_{\text{tran}}$ für alle Positionen x

$$|u_{\text{in}}(x)^{-1} u(x) - \Pi_{\text{tran}} u_{\text{in}}^{-1}(x) u(x)|^2 \leq |u_{\text{in}}(x)^{-1} u(x) - t'(x)|^2. \quad (\text{A.15})$$

Betrachten wir die Gleichung (A.13a) für die Position x , erhalten wir

$$|u_{\text{in}}(x)(u_{\text{in}}(x)^{-1} u(x) - t'(x))|^2 = |u_{\text{in}}(x)|^2 |u_{\text{in}}(x)^{-1} u(x) - t'(x)|^2 \quad (\text{A.16})$$

Wegen $|u_{\text{in}}(x)|^2 \geq 0$ folgt mit Gleichung (A.15)

$$\begin{aligned} |u_{\text{in}}(x)(u_{\text{in}}(x)^{-1} u(x) - t'(x))|^2 &= |u_{\text{in}}(x)|^2 |u_{\text{in}}(x)^{-1} u(x) - t'(x)|^2 \\ &\geq |u_{\text{in}}(x)|^2 |u_{\text{in}}(x)^{-1} u(x) - (\Pi_{\text{tran}} u_{\text{in}}^{-1} u)(x)|^2 \\ &= |u_{\text{in}}(x)(u_{\text{in}}(x)^{-1} u(x) - (\Pi_{\text{tran}} u_{\text{in}}^{-1} u)(x))|^2. \end{aligned} \quad (\text{A.17})$$

Aus dem Vergleich der Gleichungen (A.13) mit Gleichung (A.17) folgt im Widerspruch zur Annahme (A.11)

$$\|u - u'\| \geq \|u - \Pi_2 u\|. \quad (\text{A.18})$$

■

A.2 Zu Kapitel 3

A.2.1 Ableitung von Gleichung (3.47)

Laut Abschnitt 3.3.2 stehen bei mehrstufigen Strahlteiler-elementen mit $0 \in \mathcal{W}_{\text{sig}}$ die über die Gleichungen (3.44b) und (3.44c) gegebenen Vektoren \tilde{u}_0 bzw. \hat{u}_0 näherungsweise senkrecht aufeinander. Bei Vernachlässigung der Abhängigkeit $\tilde{u}_o(\gamma_{\text{lin}})$ folgt somit aus Gleichung (3.44a) näherungsweise die Beziehung

$$\frac{(\Delta I_0)_{q=1}}{(\Delta I_0)_{q=q'}} = \frac{(|\hat{u}_0|^2)_{q=1}}{(|\hat{u}_0|^2)_{q=q'}}. \quad (\text{A.19})$$

Betrachten wir zunächst den Fall $q' = 2$, so ergibt sich bei Berücksichtigung von Gleichung (3.45)

$$\frac{(\Delta I_0)_{q=1}}{(\Delta I_0)_{q=2}} = \frac{2(1 + \cos \gamma_{\text{lin}} \pi)[(\hat{n}_0)_{q=1}]^2}{2(1 + \cos \gamma_{\text{lin}} \pi)(\hat{n}_0^2 + \hat{n}_1^2 + 2\hat{n}_0\hat{n}_1 \cos \frac{\gamma_{\text{lin}} \pi}{2})}. \quad (\text{A.20})$$

Dabei beziehen sich die \hat{n}_j im Nenner auf das Design mit $q = 2$. Wenn wir im Folgenden davon ausgehen, dass Zahl der Abtastwerte n_x in beiden betrachteten Designs gleich sind und $\tilde{n}_j \ll n_x$ gilt, kann

$$(\hat{n}_0)_{q=1} = \hat{n}_0 + \hat{n}_1 \quad (\text{A.21})$$

angenommen werden. Somit folgt

$$\frac{(\Delta I_0)_{q=1}}{(\Delta I_0)_{q=2}} = \frac{(\hat{n}_0 + \hat{n}_1)^2}{\bar{n}_0^2 + \bar{n}_1^2 + 2\bar{n}_0\bar{n}_1 \cos \frac{\gamma_{\text{lin}} \pi}{2}} \quad (\text{A.22a})$$

$$= \left[1 + 2\alpha(\alpha - 1)(1 - \cos \frac{\gamma_{\text{lin}} \pi}{2}) \right]^{-1} \quad \text{mit} \quad (\text{A.22b})$$

$$\alpha = \frac{\bar{n}_0}{\bar{n}_0 + \bar{n}_1}. \quad (\text{A.22c})$$

Gleichung (A.22b) wird für $\alpha = 1/2$ minimiert, was dem Fall $\bar{n}_0 = \bar{n}_1$ entspricht. Demnach sind die Toleranzeigenschaften eines vierstufigen Designs tendenziell dann am besten, wenn die vier Quantisierungswerte gleichmäßig genutzt werden. Zur Charakterisierung der maximal erreichbaren Verbesserung der Toleranzeigenschaften durch ein vierstufiges Design mit $\gamma_{\text{lin}} \approx 1$ setzen wir in Gleichung (A.22b) $\alpha = 1/2$ und $\gamma_{\text{lin}} = 1$ ein und erhalten

$$\frac{(\Delta I_0)_{q=1}}{(\Delta I_0)_{q=2}} = 2. \quad (\text{A.23})$$

Im Folgenden besteht das Ziel, derartige Quotienten für beliebige Werte $q = q'$ abzuleiten. Hiezu stellen wir zunächst fest, dass der Nenner von Gleichung (A.22a) der Addition zweier Vektoren mit Längen proportional zu \bar{n}_0 und \bar{n}_1 entspricht, zwischen denen ein Winkel von $\gamma_{\text{lin}} \pi/2$ aufgespannt wird. Während bei $q = 2$ die beiden Vektoren für die Quantisierungswerte $\{p_0, p_2\}$ und $\{p_1, p_3\}$ stehen, kann bei $q = 3$ ähnlich vorgegangen werden, indem Vektoren für $\{p_0, p_2, p_4, p_6\}$ und $\{p_1, p_3, p_5, p_7\}$ konstruiert werden, die dann einen Winkel $\gamma_{\text{lin}} \pi/4$ aufspannen. Für beliebige $q = q'$ ergibt sich somit als Verallgemeinerung von Gleichung (A.22b)

$$\frac{(\Delta I_0)_{q=q'-1}}{(\Delta I_0)_{q=q'}} = \left[1 + 2\alpha(\alpha - 1)(1 - \cos(2^{-q'+1} \gamma_{\text{lin}} \pi)) \right]^{-1}, \quad (\text{A.24})$$

wobei α analog zu Gleichung (A.22c) das Verhältnis zwischen der Länge des ersten Vektors und der Summe der Längen beider Vektoren bezeichnet. Da dieser Ausdruck wiederum für $\alpha = 1/2$ minimiert wird, lässt sich die im Zusammenhang mit Gleichung (A.22c) getroffene Aussage verallgemeinern, d. h. um gute Toleranzeigenschaften zu erreichen sollten die verschiedenen Quantisierungswerte möglichst gleichmäßig genutzt werden.

Durch Einsetzen von $\alpha = 1/2$ sowie $\gamma_{\text{lin}} = 1$ in Gleichung (A.24) erhalten wir in Verallgemeinerung von Gleichung (A.23) die in Abschnitt 3.1.3 verwendete Gleichung (3.47).

A.2.2 Beweis zu Satz 3.26

Satz 3.26 *Der auf Basis eines bezüglich der Norm d aus Gleichung (2.6) abstandsminimierenden Projektionsoperators $\Pi_1 : \mathcal{M} \rightarrow \mathcal{M}_1$ entsprechend Gleichung (3.24) konstruierte Operator $\Pi_2 = \Pi_1^{+t_d}$ definiert eine bezüglich d abstandsminimierende Projektion auf die Menge \mathcal{M}_2 aus Gleichung (3.25).*

Beweis: Wir betrachten ein Element $t_2 \in \mathcal{M}_2$ mit

$$t_2 = \Pi_2 t = \Pi_1[t + t_d] - t_d. \quad (\text{A.25})$$

Zum Nachweis der Abstandsminimierung genügt es zu zeigen, dass für ein beliebiges Element $t'_2 \in \mathcal{M}_2$ die Bedingung

$$d(t_2, t) \leq d(t'_2, t). \quad (\text{A.26})$$

gilt. Wegen $t'_2 \in \mathcal{M}_2$ existiert ein Element $t'_1 \in \mathcal{M}_1$ mit

$$t'_2 = t'_1 - t_d. \quad (\text{A.27})$$

Des Weiteren betrachten wir ein Element $t_1 \in \mathcal{M}_1$ mit

$$t_1 = \Pi_1[t + t_d], \quad (\text{A.28})$$

wobei wegen Gleichung (A.25)

$$t_2 = t_1 - t_d \quad (\text{A.29})$$

folgt. Aufgrund der Definition der Metrik d in Gleichung (2.6) resultiert aus den Gleichungen (A.29) und (A.27)

$$d(t_2, t) = d(t_1 - t_d, t) = d(t_1, t + t_d) \quad \text{bzw.} \quad (\text{A.30})$$

$$d(t'_2, t) = d(t'_1 - t_d, t) = d(t'_1, t + t_d). \quad (\text{A.31})$$

Wegen der Abstandsminimierungseigenschaft von Π_1 folgt aus Gleichung (A.28)

$$d(t_1, t + t_d) \leq d(t'_1, t + t_d). \quad (\text{A.32})$$

Unter Benutzung der Gleichungen (A.30) und (A.31) zeigt sich die Äquivalenz der Ungleichungen (A.26) und (A.32) und somit die Gültigkeit der Behauptung. ■

A.3 Zu Kapitel 4

A.3.1 Ableitung der Gleichungen (4.15)

Wir stellen im Folgenden einen Ansatz vor, welcher nach Iteration n eine Approximation

$$v_{\text{appr}}^{(n)}(\lambda) \approx v_{\text{conv}}^{(n)}(\lambda) \quad (\text{A.33})$$

ermittelt und die Projektionsstärke $\lambda^{(n+1)}$ der darauf folgenden Iteration anhand des Maximums von $v_{\text{appr}}^{(n)}(\lambda)$ bestimmt. Da dies für die erste Iteration noch nicht möglich ist, setzen wir hierbei

$$\lambda^{(0)} = 1. \quad (\text{A.34})$$

Mit der Beziehung (4.11) wurde in Abschnitt 4.1.2 eine untere Abschätzung für $v_{\text{conv}}^{(n)}(\lambda)$ abgeleitet. Die approximative Modellierung $v_{\text{appr}}^{(n)}(\lambda)$ konstruieren wir durch Übernahme des ersten Terms auf der rechten Seite der Ungleichung (4.11) unter Hinzunahme eines Korrekturterms proportional zu λ^2 :

$$v_{\text{appr}}^{(n)}(\lambda) = (1 - |1 - \bar{\alpha} \lambda|^2) J(u_{\text{out}}^{(n-1)})^2 - \beta \lambda^2. \quad (\text{A.35})$$

Der Korrekturterm $\beta \lambda^2$ ist motiviert durch die Annahme, dass der Term $\|u_{\text{noise}}\|^2$ in Ungleichung (4.11) näherungsweise quadratisch mit λ zunimmt. Für $\lambda = 0$ ergibt sich die exakte Übereinstimmung

$$v_{\text{conv}}^{(n)}(0) = v_{\text{appr}}^{(n)}(0) = 0. \quad (\text{A.36})$$

Die gesuchte Projektionsstärke $\lambda^{(n+1)}$ wird als Maximalstelle der Kurve $v_{\text{appr}}^{(n)}(\lambda)$ gewählt, d. h.

$$\frac{\partial}{\partial \lambda} v_{\text{appr}}^{(n)}(\lambda) = 0 \Big|_{\lambda=\lambda^{(n+1)}}. \quad (\text{A.37})$$

Durch Einsetzen von Gleichung (A.35) ergibt sich Gleichung (4.15a).

Die Parameter $\bar{\alpha}$ und β können – wie im Folgenden gezeigt wird – aus $u_{\text{out}}^{(n)}$, $u_{\text{out}}^{(n-1)}$ und $\lambda^{(n)}$ berechnet werden. Zur Ermittlung von $\bar{\alpha}$ schreiben wir zunächst Gleichung (4.4) um in

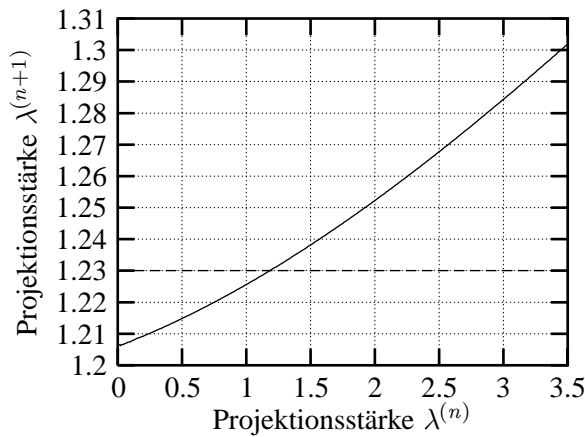
$$u_{\text{noise}} = u_{\text{out}}^{(n)} - u_{\text{out}}^{(n-1)} - \bar{\alpha} \lambda^{(n)} (\Pi_{\text{sig}} u_{\text{out}}^{(n-1)} - u_{\text{out}}^{(n-1)}). \quad (\text{A.38})$$

Ähnlich wie bei der Ermittlung des optimalen Skalierungsfaktors α in Gleichung (2.52) lässt sich der Parameter $\bar{\alpha}$ zur Minimierung von $\|u_{\text{noise}}\|$ ableiten, wobei Gleichung (4.15b) resultiert.

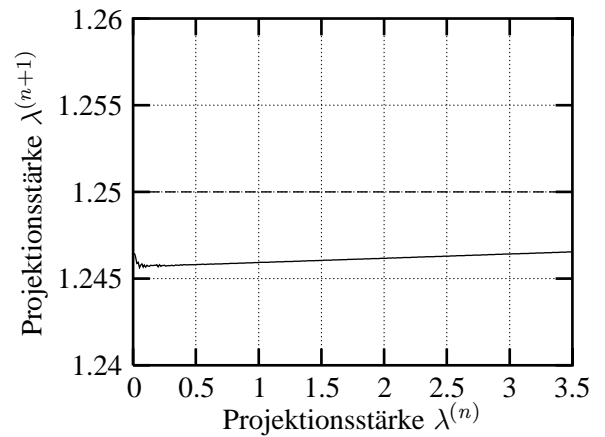
Durch Einsetzen von $\bar{\alpha}$ in Gleichung (A.35) resultiert unter der Annahme von $v_{\text{conv}}^{(n)} = v_{\text{appr}}^{(n)}$ Gleichung (4.15c) zur Berechnung des Parameters β .

Mit den Gleichungen (4.15) haben wir somit ein Schema zur Berechnung von $\lambda^{(n+1)}$ als Approximation der in Iteration $(n + 1)$ optimalen Projektionsstärke konstruiert. Dabei wird auf die Ausgangsfelder $u_{\text{out}}^{(n)}$, $u_{\text{out}}^{(n-1)}$ sowie auf die Projektionsstärke $\lambda^{(n)}$ zurückgegriffen.

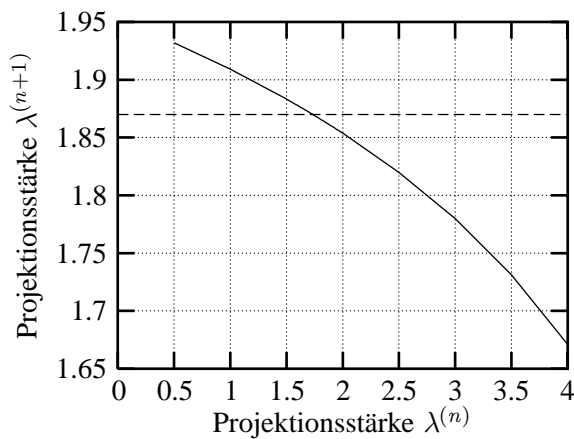
Zur Überprüfung der zu Grunde liegenden Modellierung (A.35) betrachten wir in Abbildung A.1 die Abhängigkeit zwischen dem entsprechend Gleichung (4.15a) ermittelten Wert $\lambda^{(n+1)}$ und der Projektionsstärke $\lambda^{(n)}$ der vorangegangenen Iteration für die vier bereits in Abbildung 4.1 untersuchten Iterationszustände. Idealerweise sollten sich unabhängig von der Wahl $\lambda^{(n)}$ für $\lambda^{(n+1)}$ die jeweilige optimale Projektionsstärke λ_{opt} ergeben, welche sowohl in den Diagrammen von Abbildung 4.1 als auch in Abbildung A.1 als gestrichelte Linie eingezeichnet wurde. In den Diagrammen von Abbildung A.1 treten jeweils nur relativ leichte Abweichungen von diesem Idealwert auf, was eine Bestätigung der Modellierung (A.35) darstellt.



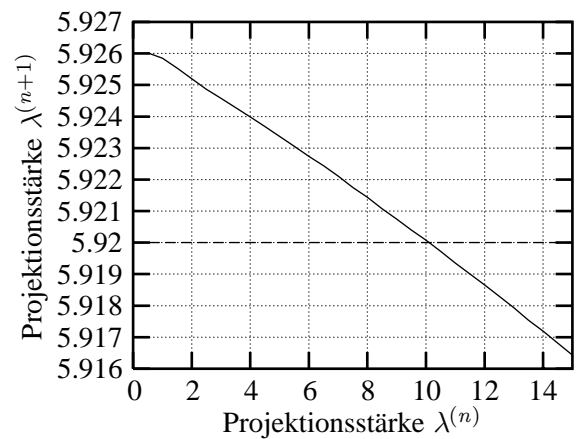
(a) Beispiel aus Abbildung 4.1(a)



(b) Beispiel aus Abbildung 4.1(b)



(c) Beispiel aus Abbildung 4.1(c)



(d) Beispiel aus Abbildung 4.1(d)

Abbildung A.1: Entsprechend Gleichung (4.15a) ermittelte Approximation der optimalen Projektionsstärke $\lambda^{(n+1)}$ in Abhängigkeit von der Projektionsstärke $\lambda^{(n)}$ der vorangegangenen Iteration für die in Abbildung 4.1 betrachteten Iterationszustände

A.4 Zu Kapitel 5

A.4.1 Beweis zu Satz 5.8

Satz 5.8 *Der über die Gleichungen (5.7) definierte Operator $\tilde{\Pi}_2$ bewirkt eine Projektion auf die durch Gleichung (5.6b) gegebene Menge $\tilde{\mathcal{M}}_2$.*

Beweis: Wir betrachten

$$\mathbf{v} = \tilde{\Pi}_2 \mathbf{u}. \quad (\text{A.39})$$

Im Vergleich von Gleichung (5.7a) mit der Mengendefinition (5.6b) folgt

$$\mathbf{v} \in \tilde{\mathcal{M}}_2. \quad (\text{A.40})$$

Es bleibt daher gemäß Gleichung (2.3a) zu zeigen, dass

$$\mathbf{d}(\mathbf{u}, \mathbf{v}) = \min_{\mathbf{v}' \in \tilde{\mathcal{M}}_2} \mathbf{d}(\mathbf{u}, \mathbf{v}'), \quad (\text{A.41})$$

wobei die Metrik $\mathbf{d} : \mathcal{H}^N \times \mathcal{H}^N \mapsto \mathbb{R}$ durch Gleichung (2.20) gegeben ist.

Entsprechend der im Zusammenhang mit Gleichung (2.7) getroffenen Diskussion stellt die punktweise Abstandsminimierung, welche sich für eine Position x darstellen lässt als

$$\sum_{j=1}^N |u_j(x) - v_j(x)|^2 = \min_{\mathbf{v}' \in \widetilde{\mathcal{M}}_2} \sum_{i=1}^N |u_j(x) - v'_j(x)|^2, \quad (\text{A.42})$$

eine hinreichende Voraussetzung für die Erfüllung der Abstandsminimierungseigenschaft (A.41) dar. Unter Berücksichtigung der Darstellungsform (5.6b) der Elemente $\mathbf{v} \in \widetilde{\mathcal{M}}_2$ können wir Gleichung (A.42) umformen in

$$\sum_{j=1}^N |u_j(x) - t_{\mathbf{v}}(x) u_{\text{in},j}(x)|^2 = \min_{z \in \mathbb{C}} \sum_{j=1}^N |u_j(x) - z u_{\text{in},j}(x)|^2 \quad \text{bzw.} \quad (\text{A.43})$$

$$t_{\mathbf{v}}(x) = \operatorname{argmin}_{z \in \mathbb{C}} \sum_{j=1}^N |u_j(x) - z u_{\text{in},j}(x)|^2. \quad (\text{A.44})$$

Wir konstruieren im Folgenden

$$z' = \operatorname{argmin}_{z \in \mathbb{C}} \sum_{j=1}^N |u_j(x) - z u_{\text{in},j}(x)|^2 \quad (\text{A.45})$$

für eine Position x , um zu zeigen, dass hierbei mit

$$t_{\mathbf{v}}(x) = z' \quad (\text{A.46})$$

die Behauptung (A.44) und somit die Gültigkeit von Satz 5.8 folgt.

Unter Zuhilfenahme der Zerlegung von $z \in \mathbb{C}$ in Real- und Imaginärteil, $z = a + ib$ ($a, b \in \mathbb{R}$), führen wir für den auf der rechten Seite von Gleichung (A.45) zu minimierenden Term die Abkürzung $\tilde{d}(a, b)$ ein, d. h.

$$\tilde{d}(a, b) = \sum_{j=1}^N |u_j(x) - (a + ib) u_{\text{in},j}(x)|^2. \quad (\text{A.47})$$

Zur Berechnung der Minimierung in Gleichung (A.45) betrachten wir die partiellen Ableitungen von $\tilde{d}(a, b)$ nach a und b . Für die ersten Ableitungen ergibt sich

$$\frac{\partial}{\partial a} \tilde{d}(a, b) = \sum_{j=1}^N [2a |u_{\text{in},j}(x)|^2 - 2 \operatorname{Re}(u_{\text{in},j}^*(x) u_j(x))] \quad \text{und} \quad (\text{A.48a})$$

$$\frac{\partial}{\partial b} \tilde{d}(a, b) = \sum_{j=1}^N [2b |u_{\text{in},j}(x)|^2 - 2 \operatorname{Im}(u_{\text{in},j}^*(x) u_j(x))]. \quad (\text{A.48b})$$

Für die zweiten Ableitungen resultiert

$$\frac{\partial^2}{\partial a^2} \tilde{d}(a, b) = 2 \sum_{j=1}^N |u_{\text{in},j}(x)|^2 \quad \text{und} \quad (\text{A.49a})$$

$$\frac{\partial^2}{\partial b^2} \tilde{d}(a, b) = 2 \sum_{j=1}^N |u_{\text{in},j}(x)|^2. \quad (\text{A.49b})$$

Da die in den Gleichungen (A.49) gegebenen zweiten Ableitungen nichtnegativ sind, kann durch Nullsetzen der ersten Ableitungen in den Gleichungen (A.48) die Minimierung in Gleichung (A.45) gelöst werden, d. h.

$$0 = \frac{\partial}{\partial a} \tilde{d}(a, b) \Big|_{a=a'}, \quad (\text{A.50a})$$

$$0 = \frac{\partial}{\partial b} \tilde{d}(a, b) \Big|_{b=b'}. \quad (\text{A.50b})$$

Somit ergibt sich

$$a' = \frac{\sum_{j=1}^N \operatorname{Re}(u_{\text{in},j}^*(x) u_j(x))}{\sum_{j=1}^N |u_{\text{in},j}(x)|^2}, \quad (\text{A.51a})$$

$$b' = \frac{\sum_{j=1}^N \operatorname{Im}(u_{\text{in},j}^*(x) u_j(x))}{\sum_{j=1}^N |u_{\text{in},j}(x)|^2}. \quad (\text{A.51b})$$

Aus der Äquivalenz der Gleichungen (A.51) mit Gleichung (5.7b) folgt die Gültigkeit der Behauptung (A.46) und damit Satz 5.8. ■

A.4.2 Beweis zu Satz 5.25

Satz 5.25 *Der Operator $\tilde{\Pi}_2$ gemäß Gleichung (5.24b) definiert einen punktweise abstandsminimierenden Projektionsoperator auf $\tilde{\mathcal{M}}_2$ aus Gleichung (5.23b).*

Beweis: Wir betrachten

$$\mathbf{v} = \tilde{\Pi}_2 \mathbf{u}. \quad (\text{A.52})$$

Durch Vergleichen der Definitionen (5.24b) und (5.23b) erkennen wir

$$\mathbf{v} \in \tilde{\mathcal{M}}_2. \quad (\text{A.53})$$

Wir weisen im Folgenden die punktweise Abstandsminimierung von $\tilde{\Pi}_2$ nach, woraus zusammen mit der Beziehung (A.53) die Behauptung von Satz 5.25 folgt.

In Abwandlung von Gleichung (A.44) lässt sich die Forderung nach punktweiser Abstandsminimierung für eine Position x schreiben als

$$t_{\mathbf{v}}(x) = \operatorname{argmin}_{z \in \mathbb{C}: |z|=1} \sum_{j=1}^N |u_j(x) - z u_{\text{in},j}(x)|^2. \quad (\text{A.54})$$

Wie schon im Beweis zu Satz 5.8 führen wir eine Abkürzung \tilde{d} für den auf der rechten Seite von Gleichung (A.54) zu minimierenden Term ein, wobei \tilde{d} diesmal jedoch in Abhängigkeit der Polarkoordinaten r und ϕ , welche sich aus der Zerlegung $z = r \exp(i\phi)$ ($r, \phi \in \mathbb{R}$) ergeben, dargestellt wird:

$$\tilde{d}(r, \phi) = \sum_{\beta=1}^B |u_{\beta}(x) - r \exp(i\phi) u_{\text{in},\beta}(x)|^2. \quad (\text{A.55})$$

Durch Anwendung elementarer Umformungen folgt aus Gleichung (A.55)

$$\begin{aligned} \tilde{d}(r, \phi) = & \sum_{\beta=1}^B \left[|u_{\beta}(x)|^2 + r^2 |u_{\text{in},\beta}(x)|^2 + \right. \\ & - 2r \left(\cos \phi \operatorname{Re}[u_{\beta}(x)] \operatorname{Re}[u_{\text{in},\beta}(x)] + \sin \phi \operatorname{Im}[u_{\beta}(x)] \operatorname{Re}[u_{\text{in},\beta}(x)] + \right. \\ & \left. \left. + \cos \phi \operatorname{Im}[u_{\beta}(x)] \operatorname{Im}[u_{\text{in},\beta}(x)] - \sin \phi \operatorname{Re}[u_{\beta}(x)] \operatorname{Im}[u_{\text{in},\beta}(x)] \right) \right]. \end{aligned} \quad (\text{A.56})$$

Gleichung (A.56) kann umgeformt werden in

$$\begin{aligned} \tilde{d}(r, \phi) = & \tilde{d}_r(r) + r \tilde{d}_{\phi}(\phi) \quad \text{mit} \quad (\text{A.57a}) \\ \tilde{d}_r(r) = & \sum_{\beta=1}^B \left[|u_{\beta}(x)|^2 + r^2 |u_{\text{in},\beta}(x)|^2 \right] \quad \text{und} \\ \tilde{d}_{\phi}(\phi) = & -2 \sum_{\beta=1}^B \left[\cos \phi \operatorname{Re}[u_{\beta}(x)] \operatorname{Re}[u_{\text{in},\beta}(x)] + \sin \phi \operatorname{Im}[u_{\beta}(x)] \operatorname{Re}[u_{\text{in},\beta}(x)] + \right. \\ & \left. + \cos \phi \operatorname{Im}[u_{\beta}(x)] \operatorname{Im}[u_{\text{in},\beta}(x)] - \sin \phi \operatorname{Re}[u_{\beta}(x)] \operatorname{Im}[u_{\text{in},\beta}(x)] \right]. \end{aligned}$$

Betrachten wir nun die Berechnung der Transmissionsfunktion $t_{\mathbf{v}}(x)$ gemäß Gleichung (5.24c). Wenn sich vor Anwendung von Π_{phase} an der Position x der Wert

$$r_1 \exp(i\phi_1) = \frac{\sum_{\beta=1}^B u_{\text{in},\beta}^* u_{\beta}}{\sum_{\beta=1}^B |u_{\text{in},\beta}|^2} \quad (\text{A.58})$$

ergibt, so resultiert nach Anwendung von Π_{phase} entsprechend Gleichung (2.60) der Wert

$$\exp(i\phi_1) = \Pi_{\text{phase}}[r_1 \exp(i\phi_1)] = t_{\mathbf{v}}(x). \quad (\text{A.59})$$

Analog zu der Vorgehensweise im Beweis zu Satz 5.8 kann leicht gezeigt werden, dass für beliebige $z_2 = r_2 \exp(i\phi_2)$

$$\tilde{d}(r_2, \phi_2) \geq \tilde{d}(r_1, \phi_1) \quad (\text{A.60})$$

folgt. Insbesondere gilt somit

$$\tilde{d}(r_1, \phi_2) \geq \tilde{d}(r_1, \phi_1). \quad (\text{A.61})$$

Nehmen wir nun im Sinne eines indirekten Beweises an, $\tilde{\Pi}_2$ wäre nicht punktweise abstandsminimierend. Demnach müsste es einen Phasenwert ϕ_2 geben mit

$$\tilde{d}(1, \phi_2) < \tilde{d}(1, \phi_1). \quad (\text{A.62})$$

Wegen der Zerlegung (A.57) folgt daraus für einen beliebigen Radius $r_1 > 0$

$$\tilde{d}(r_1, \phi_2) < \tilde{d}(r_1, \phi_1), \quad (\text{A.63})$$

was im Widerspruch zu Ungleichung (A.61) steht. Somit ist für

$$t_{\mathbf{v}}(x) = \exp(i\phi_1) \quad (\text{A.64})$$

also die Behauptung (A.54) erfüllt, womit die Gültigkeit von Satz 5.25 folgt. ■

A.4.3 Beweis zu Satz 5.28

Satz 5.28 *Der Operator $\tilde{\Pi}_2$ aus Gleichung (5.27) definiert eine punktweise abstandsminimierende Projektion auf $\tilde{\mathcal{M}}_2$ nach Gleichung (5.26).*

Beweis: Wir betrachten

$$\mathbf{v} = \tilde{\Pi}_2 \mathbf{u}, \quad (\text{A.65})$$

wobei aus den Gleichungen (5.26) und (5.27)

$$\mathbf{v} \in \tilde{\mathcal{M}}_2 \quad (\text{A.66})$$

folgt.

Wir gehen im Folgenden ähnlich wie im Beweis zu Satz 5.25 vor und zeigen die punktweise Abstandsminimierung, welche sich in Abwandlung zu Gleichung (A.54) ausdrücken lässt als

$$t_{\mathbf{v}}(x) = \operatorname{argmin}_{z \in \tilde{\mathcal{M}}_{\text{quant}, Q}} \sum_{j=1}^N |u_j(x) - z u_{\text{in},j}(x)|^2. \quad (\text{A.67})$$

Wie schon in Gleichung (A.55) betrachten wir den zu minimierenden Term auf der rechten Seite von Gleichung (A.67) bezüglich der Polarkoordinatenzerlegung von $z = r \exp(i\phi)$:

$$\tilde{d}(r, \phi) = \sum_{\beta=1}^B |u_{\beta}(x) - r \exp(i\phi) u_{\text{in},\beta}(x)|^2. \quad (\text{A.68})$$

Den aus Gleichung (5.27b) an der Position x resultierenden Phasenwert bezeichnen wir im Folgenden mit ϕ_1 , d. h.

$$t_{\mathbf{v}}(x) = \exp(i\phi_1). \quad (\text{A.69})$$

Nehmen wir nun im Sinne eines indirekten Beweises an, $\tilde{\Pi}_2$ wäre nicht punktweise abstandsminimierend. Demnach müsste es einen Phasenwert $\phi_2 \neq \phi_1$ mit $\exp(i\phi_2) \in \tilde{\mathcal{M}}_{\text{quant}, Q}$ geben mit

$$\tilde{d}(1, \phi_2) < \tilde{d}(1, \phi_1). \quad (\text{A.70})$$

In Übertragung der Argumentation aus dem Beweis zu Satz 5.25 würde daraus für einen beliebigen Radius $r > 0$ folgen

$$\tilde{d}(r, \phi_2) < \tilde{d}(r, \phi_1). \quad (\text{A.71})$$

Den Radius r wählen wir in Analogie zu Gleichung (A.58) entsprechend des Transmissionswertes z_3 , welcher sich vor Anwendung der Transmissionseinschränkung ergibt:

$$z_3 = r \exp(i\phi_3) = \frac{\sum_{\beta=1}^B u_{\text{in},\beta}^*(x) u_{\beta}(x)}{\sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2}. \quad (\text{A.72})$$

Bei der Anwendung von $\Pi_{\text{quant},Q}$ bei der Ermittlung von ν wird der resultierende Phasenwert ϕ_1 als derjenige Phasenwert innerhalb von $\widetilde{\mathcal{M}}_{\text{quant},Q}$ gewählt, welcher am dichtesten an ϕ_3 liegt. Demnach folgt

$$|(\phi_1 - \phi_3) \bmod 2\pi| \leq |(\phi_2 - \phi_3) \bmod 2\pi|, \quad \text{wobei} \quad (\text{A.73a})$$

$$a \bmod b = c \Leftrightarrow c \in [-b/2, b/2) \wedge c = a + jb \wedge j \in \mathbb{Z}. \quad (\text{A.73b})$$

Analog zu der Vorgehensweise im Beweis zu Satz 5.8 können leicht die Beziehungen

$$\tilde{d}(r, \phi_1) \geq \tilde{d}(r, \phi_3) \quad \text{und} \quad (\text{A.74a})$$

$$\tilde{d}(r, \phi_2) \geq \tilde{d}(r, \phi_3). \quad (\text{A.74b})$$

gezeigt werden.

Durch Einführung der formalen Größen

$$\Delta u_\beta = \frac{u_\beta(x)}{u_{\text{in},\beta}(x)} - z_3 \quad (\text{A.75})$$

können wir unter Benutzung von

$$u_\beta(x) = u_{\text{in},\beta}(x)(z_3 + \Delta u_\beta) \quad \text{und} \quad (\text{A.76})$$

$$r \exp(i\phi) = \exp[i(\phi - \phi_3)]z_3 \quad (\text{A.77})$$

Gleichung (A.68) umformen in

$$\begin{aligned} \tilde{d}(r, \phi) &= \sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 |z_3 + \Delta u_\beta - \exp[i(\phi - \phi_3)]z_3|^2 \\ &= \sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 |(1 - \exp[i(\phi - \phi_3)])z_3 + \Delta u_\beta|^2. \end{aligned} \quad (\text{A.78})$$

Für den Fall $\phi = \phi_3$ folgt damit

$$\tilde{d}(r, \phi_3) = \sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 |\Delta u_\beta|^2. \quad (\text{A.79})$$

Durch weiteres Umformung von Gleichung (A.78) erhalten wir

$$\begin{aligned} \tilde{d}(r, \phi) &= \sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 \left(|1 - \exp[i(\phi - \phi_3)]|^2 |c|^2 \right. \\ &\quad \left. - 2 \operatorname{Re} [(1 - \exp[i(\phi - \phi_3)])c\Delta u_\beta^*] + |\Delta u_\beta|^2 \right). \end{aligned} \quad (\text{A.80})$$

Unter Berücksichtigung von Gleichung (A.79) können wir folgende Zerlegung bezüglich der drei Summanden innerhalb der großen Klammern durchführen:

$$\tilde{d}(r, \phi) = \tilde{d}_1(r, \phi) + \tilde{d}_2(r, \phi) + \tilde{d}(r, \phi_3), \quad \text{wobei} \quad (\text{A.81a})$$

$$\tilde{d}_1(r, \phi) = |1 - \exp[i(\phi - \phi_3)]|^2 |c|^2 \sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 \quad \text{und} \quad (\text{A.81b})$$

$$\tilde{d}_2(r, \phi) = -2 \sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 \operatorname{Re} [(1 - \exp[i(\phi - \phi_3)])c\Delta u_\beta^*]. \quad (\text{A.81c})$$

Um zu zeigen, dass hierbei der Term $\tilde{d}_2(r, \phi)$ verschwindet, formen wir Gleichung (A.81c) weiter um in

$$\begin{aligned}
-\frac{1}{2}\tilde{d}_2(r, \phi) &= \sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 \left(\operatorname{Re} [(1 - \exp[i(\phi - \phi_3)])c] \operatorname{Re} [\Delta u_\beta] \right. \\
&\quad \left. + \operatorname{Im} [(1 - \exp[i(\phi - \phi_3)])c] \operatorname{Im} [\Delta u_\beta] \right) \\
&= \operatorname{Re} [(1 - \exp[i(\phi - \phi_3)])c] \sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 \operatorname{Re} [\Delta u_\beta] \\
&\quad + \operatorname{Im} [(1 - \exp[i(\phi - \phi_3)])c] \sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 \operatorname{Im} [\Delta u_\beta]. \tag{A.82}
\end{aligned}$$

Da mit Hilfe der Gleichungen (A.72) und (A.75) leicht gezeigt werden kann, dass

$$\sum_{\beta=1}^B |u_{\text{in},\beta}(x)|^2 \Delta u_\beta = 0 \tag{A.83}$$

gilt, folgt aus Gleichung (A.82)

$$\tilde{d}_2(r, \phi) = 0 \tag{A.84}$$

und somit als Vereinfachung von Gleichung (A.81a)

$$\tilde{d}(r, \phi) = \tilde{d}_1(r, \phi) + \tilde{d}(r, \phi_3). \tag{A.85}$$

Aus der Annahme (A.71) würde somit folgen

$$\tilde{d}_1(r, \phi_2) < \tilde{d}_1(r, \phi_1). \tag{A.86}$$

Betrachten wir nun die ϕ -Abhängigkeit von $\tilde{d}_1(r, \phi)$. Wegen

$$|1 - \exp[i(\phi - \phi_3)]|^2 = 2 - 2 \cos(\phi - \phi_3) \tag{A.87}$$

folgt aus Gleichung (A.81b) die Proportionalität

$$\tilde{d}_1(r, \phi) \sim 1 - \cos(\phi - \phi_3). \tag{A.88}$$

Für $\phi = \phi_3$ verschwindet also der Beitrag $\tilde{d}_1(r, \phi)$ und für $|\phi - \phi_3| \leq \pi$ wächst $\tilde{d}_1(r, \phi)$ monoton mit $|\phi - \phi_3|$. Durch Verwendung der modulo-Operation aus Gleichung (A.73b) kann die Einschränkung für ϕ aufgehoben werden, d. h. $\tilde{d}_1(r, \phi)$ wächst monoton mit $|(\phi - \phi_3) \bmod 2\pi|$. Aufgrund von Ungleichung (A.73a) folgt somit im Widerspruch zu Annahme (A.86)

$$\tilde{d}_1(r, \phi_1) \leq \tilde{d}_1(r, \phi_2), \tag{A.89}$$

womit die Gültigkeit von Satz 5.28 bewiesen ist. ■

Literaturverzeichnis

- [1] WYROWSKI, F. und J. TURUNEN: *Wave-optical engineering*. In: GUENTHER, A. H. (Herausgeber): *International Trends in Applied Optics*, Kapitel 21, Seiten 471 – 496. SPIE Press, London, 2002.
- [2] WYROWSKI, F. und W.T. RHODES: *Use of holographic concepts for the design of optical systems*. In: FOURNIER, JEAN-MARC (Herausgeber): *50 Years of Holography*. Springer-Verlag, Berlin, 2000.
- [3] PFEIL, A. v. und F. WYROWSKI: *Wave-optical structure design with the local plane-interface approximation*. *J. Mod. Optics*, 47(13):2335 – 2350, 2000.
- [4] PFEIL, A. v.: *Analyse und Design optischer Systeme mit der Näherung lokal ebener Grenzflächen*. Doktorarbeit, Friedrich-Schiller-Universität, Jena, 2000.
- [5] DRAUSCHKE, A.: *Die Näherung lokaler Elementarstrukturen in Analyse und Design optischer Elemente*. Doktorarbeit, Friedrich-Schiller-Universität, Jena, Germany, 2002.
- [6] GUBIN, L. G., B. T. POLYAK und E. RAIK: *The method of projections for finding the common point of convex sets*. *Comp. Math. & Math. Phys.*, 7:1211–1228, 1967.
- [7] GERCHBERG, R. W. und W. O. SAXTON: *A practical algorithm for the determination of phase from image and diffraction plane pictures*. *Optik*, 35(2):237–246, 1972.
- [8] LIU, B. und N. C. GALLAGHER: *Convergence of a spectrum shaping algorithm*. *Appl. Opt.*, 13(11):2470–2471, 1974.
- [9] SELDOWITZ, M. A., J. P. ALLEBACH und D. W. SWEENEY: *Synthesis of digital holograms by direct binary search*. *Appl. Opt.*, 26(14):2788 – 2798, 1987.
- [10] JENNISON, B. K., J. P. ALLEBACH und D. W. SWEENEY: *Efficient design of direct-binary-search computer-generated holograms*. *J. Opt. Soc. Am. A*, 8:652 – 660, 1991.
- [11] CHHETRI, B. B., S. YANG und T. SHIMOMURA: *Stochastic approach in the efficient design of the direct-binary-search algorithm for hologram synthesis*. *Appl. Opt.*, 39:5956 – 5964, 2000.
- [12] KIRCKPATRICK, S., JR. C. D. GELATT und M. P. VECCHI: *Optimization by simulated annealing*. *Science*, 220:671 – 680, 1983.
- [13] KIM, M. S. und C. C. GUEST: *Simulated annealing algorithm for binary phase only filters in pattern classification*. *Appl. Opt.*, 29:1203 – 1208, 1990.
- [14] YOSHIKAWA, N. und T. YATAGAI: *Phase optimization of a kinoform by simulated annealing*. *Appl. Opt.*, 33:863 – 868, 1994.
- [15] GILLET, J.-N. und Y. SHENG: *Iterative simulated quenching for designing irregular-spot-array generators*. *Appl. Opt.*, 39(20):3456 – 3465, 2000.
- [16] MEISTER, M. und R. J. WINFIELD: *Novel approaches to direct search algorithms for the design of diffractive optical elements*. *Optics Commun.*, 203:39 – 49, 2002.
- [17] BENGTTSSON, J.: *Design of fan-out kinoforms in the entire scalar diffraction regime with an optimal-rotation-angle method*. *Appl. Opt.*, 36(32):8435 – 8444, 1997.

- [18] OGURA, Y., N. SHIRAI, J. TANIDA und Y. ICHIOKA: *Wavelength-multiplexing diffractive phase elements: design, fabrication, and performance evaluation*. J. Opt. Soc. Am. A, 18(5):1082 – 1092, 2001.
- [19] MICHALEWICZ, Z.: *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag, Berlin, 1992.
- [20] JOHNSON, E. G. und M. A. G. ABUSHAGUR: *Microgenetic-algorithm optimization methods applied to dielectric gratings*. J. Opt. Soc. Am. A, 12(5):1152 – 1160, 1995.
- [21] MAEDA, Y., H. HIRANO und Y. KANATA: *A learning rule of neuronal networks via simultaneous perturbation and its hardware implementation*. Neuronal Networks, 8(2):251 – 259, 1995.
- [22] FIENUP, J. R.: *Phase retrieval algorithms: a comparison*. Appl. Opt., 21(15):2758–2769, 1982.
- [23] BAUSCHKE, H. H., P. L. COMBETTES und D. R. LUKE: *Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization*. J. Opt. Soc. Am. A, 19(7):1334 – 1345, 2002.
- [24] LEVI, A. und H. STARK: *Signal restoration from phase by projections onto convex sets*. J. Opt. Soc. Am. A, 73(6):810 – 822, 1983.
- [25] HACKER, M., G. STOBRAWA und T. FEURER: *Iterative Fourier transform algorithm for phase-only pulse shaping*. Optics Express, 9(4):191 – 199, 2001.
- [26] YEH, S. und H. STARK: *Iterative and one-step reconstruction from nonuniform samples by convex projections*. J. Opt. Soc. Am. A, 7(3):491 – 499, 1990.
- [27] JAVIDI, B., N. TOWGHI, N. MAGHZI und S. C. VERRALL: *Error-reduction techniques and error analysis for fully phase- and amplitude-based encryption*. Appl. Opt., 39(23):4117 – 4130, 2000.
- [28] PRONGUE, D., H. P. HERZIG, R. DÄNDLIKER und M. T. GALE: *Optimized kinoform structures for highly efficient fan-out elements*. Appl. Opt., 31(26):5706 – 5711, 1992.
- [29] ROSEN, J. und A. YARIV: *Synthesis of an arbitrary axial field profile by computer-generated holograms*. Optics Lett., 19(11):843 – 845, 1994.
- [30] AAGEDAL, H., F. WYROWSKI und M. SCHMID: *Paraxial beam splitting and shaping*. In: TURUNEN, J. und F. WYROWSKI (Herausgeber): *Diffractive optics for industrial and commercial applications*, Kapitel 6, Seiten 165–188. Akademie Verlag, Berlin, 1997.
- [31] CATINO, W. C., J. L. LOCICERO und H. STARK: *Design of continuous and quantized amplitude holograms by generalized projections*. J. Opt. Soc. Am. A, 15(1):68 – 76, 1998.
- [32] LEVY, U., D. MENDLOVIC, Z. ZALEVSKY, G. SHABTAY und E. MAROM: *Iterative algorithm for determining optimal beam profiles in a three-dimensional space*. Appl. Opt., 38(32):6732 – 6736, 1999.
- [33] ARRIZÓN, V., M. TESTORF, S. SINZINGER und J. JAHNS: *Iterative optimization of phase-only diffractive optical elements based on a lenslet array*. J. Opt. Soc. Am. A, 17(12):2157 – 2164, 2000.
- [34] WANG, W., T. LI und Y. LI: *A hybrid algorithm for the design of DOE in uniform illumination*. Optics Commun., 181:261 – 265, 2000.
- [35] BÜHLING, S. und F. WYROWSKI: *Solving tolerancing and three-dimensional beam shaping problems by multifunctional wave optical design*. Opt. Eng., 40(8):1590 – 1597, August 2001.
- [36] BÜHLING, S. und F. WYROWSKI: *Improving performance of iterative Fourier transform algorithm by utilizing variable strength projections*. J. Mod. Optics, 49(11):1871 – 1892, September 2002.
- [37] DI, F., Y. YINGBAI, L. SI, K. TIAN und J. GUOFAN: *Designing diffractive phase plates for beam smoothing in the fractional Fourier domain*. J. Mod. Optics, 49(7):1125 – 1133, 2002.

- [38] ŠKEREN, M., I. RICHTER und P. FIALA: *Iterative Fourier transform algorithm: Comparison of various approaches*. J. Mod. Optics, accepted, 2002.
- [39] BRONSTEIN, I. N. und K. A. SEMENDJAJEW: *Ergänzende Kapitel zu Taschenbuch der Mathematik*. BSB B. G. Teubner Verlagsgesellschaft, 1990.
- [40] KOTZER, T., N. COHEN und J. SHAMIR: *Generalized projection algorithms with applications to optics and signal restoration*. Optics Commun., 156(1-3):77–91, 1998.
- [41] CENSOR, Y., N. COHEN, T. KUTSCHER und J. SHAMIR: *Summed squared distance error reduction by simultaneous multiprojections and applications*. Applied Mathematics and Computation, 126(2–3):157–179, 2002.
- [42] LEVI, A. und H. STARK: *Image restoration by the method of generalized projections with application to restoration from magnitude*. J. Opt. Soc. Am. A, 1(9):932–943, 1984.
- [43] YOULA, D. C. und H. WEBB: *Image restoration by the method of convex projections. I. Theory*. IEEE Transactions on Medical Imaging, MI-1(2):81–94, 1982.
- [44] COMBETTES, P. L. und H. J. TRUSSEL: *Method of successive projections for finding a common point of sets in metric spaces*. Journal of Optimization Theory and Applications, 67:487–507, 1990.
- [45] KOTZER, T., J. ROSEN und J. SHAMIR: *Application of serial- and parallel-projection methods to correlation filter design*. Appl. Opt., 34(20):3883–3895, 1984.
- [46] GOODMAN, J. W.: *Introduction to Fourier optics*. McGraw-Hill, New York, 1968.
- [47] WYROWSKI, F.: *Considerations on convolutions and phase factors*. Optics Commun., 81(6):353–358, 1991.
- [48] FARN, M. W.: *Binary gratings with increased efficiency*. Appl. Opt., 31:4453–4458, 1992.
- [49] BÜHLING, S.: *Wellenoptisches Design unter Berücksichtigung multifunktionaler Systemanforderungen*. Diplomarbeit, Friedrich-Schiller-Universität, Jena, Germany, 2000.
- [50] BRIGHAM, E. O.: *The fast Fourier transform*. Prentice-Hall, Englewood Cliffs, N. J., 1974.
- [51] SCHIMMEL, H.: *Strategien zum wellenoptischen Systemdesign*. Diplomarbeit, Friedrich-Schiller-Universität, Jena, Germany, 1999.
- [52] PETIT, R. (Herausgeber): *Electromagnetic Theory of Gratings*. Springer, Berlin, 1980.
- [53] MAYSTRE, D.: *Rigorous vector theories of diffraction gratings*. Progress in Optics, 21:1–67, 1984.
- [54] KUITTINEN, M., J. TURUNEN und P. VAHIMAA: *Subwavelength-structured elements*. In: TURUNEN, J. und F. WYROWSKI (Herausgeber): *Diffraction optics for industrial and commercial applications*, Kapitel 11, Seiten 303–323. Akademie Verlag, Berlin, 1997.
- [55] WYROWSKI, F. und O. BRYNGDAHL: *Digital holography as part of diffractive optics*. Rep. Prog. Phys., 54:1481–1571, 1991.
- [56] POMMET, D. A., M. G. MAHARAM und E. B. GRANN: *Limits of scalar diffraction theory for diffractive phase elements*. J. Opt. Soc. Am. A, 11(6):1827–1834, 1994.
- [57] TURUNEN, J. und F. WYROWSKI: *Introduction to diffractive optics*. In: TURUNEN, J. und F. WYROWSKI (Herausgeber): *Diffraction optics for industrial and commercial applications*, Kapitel 1, Seiten 1–57. Akademie Verlag, Berlin, 1997.
- [58] BENGTTSSON, J., N. ERIKSSON und A. LARSSON: *Small-feature-size fan-out kinoform etched in GaAs*. Appl. Opt., 35(5):801–806, 1996.

- [59] LEVY, U., D. MENDLOVIC und E. MAROM: *Efficiency analysis of diffractive lenses*. J. Opt. Soc. Am. A, 18(1):86 – 93, 2001.
- [60] MELLIN, S. D. und G. P. NORDIN: *Limits of scalar diffraction theory and an iterative angular spectrum algorithm for finite aperture diffractive optical element design*. Optics Express, 8(13):705 – 722, 2001.
- [61] WYROWSKI, F. und H. AAGEDAL: *Wave transformation by physical-optics system design*. International Journal of Optoelectronics, 12(4):127 – 143, 1998.
- [62] DAMMANN, H. und K. GÖRTLER: *High-efficiency in-line multiple imaging by means of multiple phase holograms*. Optics Commun., 3:312 – 315, 1971.
- [63] DAMMANN, H. und E. KLOTZ: *Coherent optical generation and inspection of two-dimensional periodic structures*. Optica Acta, 24:505 – 515, 1977.
- [64] MENDLOVIC, D., I. OUZIELI, I. KIRYUSCHEV und E. MAROM: *Two-dimensional wavelet transform achieved by computer-generated multireference matched filter and Damman grating*. Appl. Opt., 34:8213 – 8219, 1995.
- [65] MAIT, J. N. und K. BRENNER: *Optical symbolic substitution: system design using phase-only holograms*. Appl. Opt., 27:1692 – 1700, 1988.
- [66] WEN, Z., P. YEH und X. YANG: *Modified two-dimensional hamming neuronal network and its optical implementation using Damman gratings*. Opt. Eng., 35:2136 – 2144, 1996.
- [67] BLOM, H., M. JOHANSSON, A.-S. HEDMAN, L. LUNDBERG, A. HANNING, S. HÅRD und R. RIGLER: *Parallel fluorescence detection of single biomolecules in microarrays by a diffractive-optical-designed 2×2 fan-out element*. Appl. Opt., 41(16):3336 – 3342, 2002.
- [68] FUSE, K., K. EBATA, T. OKADA K. KURISU und M. SHIOZAKI: *Optical Properties of ZnSe Diffractive Optical Elements for Spot Array Generation*. In: Proc. SPIE, Band 3888 der Reihe High-Power Lasers in Manufacturing, Seiten 497 – 508, 2000.
- [69] AMAKO, J., K. UMETSU und H. NAKAO: *Laser soldering with light-intensity patterns reconstructed from computer-generated holograms*. Appl. Opt., 40(31):5643 – 5649, 2001.
- [70] FUSE, K., T. HIRAI, T. USHIRO, T. OKADA, K. KURISU und K. EBATA: *Design and Performance of Multilevel Phase Fan-Out Diffractive Optical Elements for Laser Materials Processing*. In: Proc. ICALEO, 2002.
- [71] TONDUSSON, M., C. FROEHLI, V. KERMENE und M. VAMPOUILLE: *Coherent combination of four laser beams in a multi-axis Fourier cavity using a diffractive optical element*. Journal of Optics A: Pure and Applied Optics, 3:521 – 526, 2001.
- [72] VANDAMME, E. und P. MUYS: *Ultrasonic beamsplitter for high-power CO₂ lasers*. Journal of Optics A: Pure and Applied Optics, 3:S68 – S70, 2001.
- [73] COLLINS, S. A.: *Lens-System Diffraction Integral Written in Terms of Matrix Optics*. J. Opt. Soc. Am., 60(9):1168 – 1177, 1970.
- [74] AAGEDAL, H. und F. WYROWSKI: *On pixel-oriented structure parametrization for design of diffractive elements*. J. Mod. Optics, 45(7):1451 – 1464, 1998.
- [75] LÜPKEN, H., T. PETER, F. WYROWSKI und O. BRYNGDAHL: *Phase synthesis for array illuminator*. Optics Commun., 91:163 – 167, 1992.
- [76] WYROWSKI, F.: *Upper bound of the diffraction efficiency of diffractive phase elements*. Optics Lett., 16(24):1915–1917, 1991.

- [77] WYROWSKI, F.: *Efficiency of quantized diffractive phase elements*. Optics Commun., 92(41):119–126, 1992.
- [78] KRACKHARDT, U., J. N. MAIT und N. STREIBL: *Upper bound on the diffraction efficiency of phase-only fanout elements*. Appl. Opt., 31(1):27 – 37, 1992.
- [79] WYROWSKI, F.: *Diffractive optical elements: Iterative calculation of quantized, blazed phase structures*. J. Opt. Soc. Am. A, 7:961–969, 1990.
- [80] CATINO, W. C., J. L. LOCICERO und H. STARK: *Design of continuous and quantized phase holograms by generalized projections*. J. Opt. Soc. Am. A, 14(10):2715 – 2725, 1997.
- [81] BÜHLING, S. und F. WYROWSKI: *Arbitrary spot location diffractive beam splitting elements*. J. Opt. Soc. Am. A, 19(12):2414 – 2423, December 2002.
- [82] JAHNS, J., M. M. DOWNS, M. E. PRISE, N. STREIBL und S. J. WALKER: *Dammann gratings for laser beam shaping*. Opt. Eng., 28(12):1267 – 1275, 1989.
- [83] FIENUP, J. R.: *Iterative method applied to image reconstruction and to computer-generated holograms*. Opt. Eng., 19(3):297 – 305, 1980.
- [84] AAGEDAL, H., M. SCHMID, T. BETH, S. TEIWES und F. WYROWSKI: *Theory of speckles in diffractive optics and its application to beam shaping*. J. Mod. Optics, 43:1409 – 1421, 1996.
- [85] JOHANSSON, M. und J. BENGTSSON: *Robust design method for highly efficient beam-shaping diffractive optical elements using an iterative-Fourier-transform algorithm with soft operations*. J. Mod. Optics, 47(8):1385 – 1398, 2000.
- [86] FIENUP, J. R.: *Reconstruction of an object from the modulus of its Fourier transform*. Optics Lett., 3(1):27 – 29, 1978.
- [87] NASHOLD, K. M. und B. E. A. SALEH: *Image construction through diffraction-limited high-contrast imaging systems: an iterative approach*. J. Opt. Soc. Am. A, 2(5):635 – 643, 1985.
- [88] GORI, F., M. SANTARSIERO, S. VICALVI, R. BORGHI, G. CINCOTTI, E. DI FABRIZIO und M. GENTILI: *Analytical derivation of the optimum triplicator*. Optics Commun., 157:13 – 18, 1998.
- [89] CORMEN, T. H., C. E. LEISERSON und R. L. RIVEST: *Introduction to algorithms*. McGraw-Hill, 1989.
- [90] BRYNGDAHL, O. und F. WYROWSKI: *Digital Holography - Computer-generated Holograms*. Progress in Optics, 28:1–86, 1990.
- [91] LESEBERG, D.: *Computer-generated three-dimensional image holograms*. Appl. Opt., 31(2):223 – 229, 1992.
- [92] CUMME, M.: *Untersuchungen zur Strahlformung mit diffraktiven Phasenelementen*. Diplomarbeit, Friedrich-Schiller-Universität, Jena, Germany, 1998.
- [93] PIESTUN, R. und J. SHAMIR: *Control of wave-front propagation with diffractive elements*. Optics Lett., 19(11):771–773, 1994.
- [94] PIESTUN, R., B. SPEKTOR und J. SHAMIR: *Unconventional light distributions in three-dimensional domains*. J. Mod. Optics, 43(7):1495–1507, 1996.
- [95] SALIK, B., J. ROSEN und A. YARIV: *One-dimensional beam shaping*. J. Opt. Soc. Am. A, 12(8):1702 – 1706, 1995.
- [96] LIU, R., B.-Y. GU, B.-Z. DONG und G.-Z. YANG: *Design of diffractive phase elements that realize axial intensity modulation based on conjugate gradient method*. J. Opt. Soc. Am. A, 15(3):689–694, 1998.

- [97] LEVY, U., E. MAROM und D. MENDLOVIC: *Simultaneous multicolor image formation with a single diffractive optical element*. Optics Lett., 26(15):1149 – 1151, 2001.
- [98] BENGTTSSON, J.: *Kinoforms designed to produce different fan-out patterns for two wavelengths*. Appl. Opt., 37(11):2011 – 2020, 1998.
- [99] BENGTTSSON, J. und M. JOHANSSON: *Fan-out diffractive optical elements designed for increased fabrication tolerances to linear relief depth errors*. Appl. Opt., 41(2):281 – 289, 2002.
- [100] DAMMANN, H.: *Color separation gratings*. Appl. Opt., 17:2273 – 2279, 1978.
- [101] DAMMANN, H.: *Spectral characteristics of stepped-phase gratings*. Optik, 53:409 – 417, 1979.
- [102] BARTON, I. M., P. BLAIR und M. R. TAGHIZADEH: *Dual-wavelength operation diffractive phase elements for pattern formation*. Optics Express, 1(2):54–58, 1997.
- [103] DOSKOLOVICH, L. L. und P. M. REPETTO: *Design of DOEs for wavelength demultiplexing and spatial focusing*. Journal of Optics A: Pure and Applied Optics, 2:488 – 493, 2000.
- [104] EBSTEIN, S. M.: *Nearly index-matched optics for aspherical, diffractive, and achromatic-phase diffractive elements*. Optics Lett., 21(18):1454 – 1456, 1996.
- [105] ARIELI, Y., S. OZERI, N. EISENBERG und S. NOACH: *Design of a diffractive optical element for wide spectral bandwidth*. Optics Lett., 23(11):823–824, 1998.
- [106] ARIELI, Y., S. NOACH, S. OZERI und N. EISENBERG: *Design of diffractive optical elements for multiple wavelengths*. Appl. Opt., 37(26):6174 – 6177, 1998.
- [107] NOACH, S., Y. ARIELI und N. EISENBERG: *Wave-front control and aberration correction with a diffractive optical element*. Optics Lett., 24(5):333 – 335, 1999.
- [108] CEDERQUIST, J. N. und J. R. FIENUP: *Analytic design of optimum holographic optical elements*. J. Opt. Soc. Am. A, 4(4):699–705, 1987.
- [109] SHABATAY, G., Z. ZALEVSKY, U. LEVY und D. MENDLOVIC: *Optimal synthesis of three-dimensional complex amplitude distributions*. Optics Lett., 25(6):363 – 365, 2000.
- [110] LARSSON, M., M. EKBERG, F. NIKOLAJEFF und S. HARD: *Successive development optimization of resist kinoforms manufactured with direct-writing, electron-beam lithography*. Appl. Opt., 33(7):1176 – 1179, 1994.
- [111] BÜHLING, S., F. WYROWSKI, E.-B. KLEY, A.J.M. NELLISSEN, L. WANG und M. DIRKSWAGER: *Resolution enhanced proximity printing by phase and amplitude modulating masks*. J. Micromech. Microeng., 11(5):603 – 611, September 2001.
- [112] WYROWSKI, F.: *Iterative quantization of digital amplitude holograms*. Appl. Opt., 28:3864–3870, 1989.

Abbildungsverzeichnis

2.1	Zerlegung eines optischen Systems in durch Referenzebenen senkrecht zur z -Achse begrenzte Bereiche, denen jeweils Operatoren P_j ($j = 1 \dots n$) zugeordnet sind . . .	11
2.2	Beispiele für Systeme zur Realisierung der optischen Fouriertransformation einer Transmissionsfunktion	14
2.3	Zerlegung des bei Transmissionsdesignproblemen betrachteten optischen Systems .	16
2.4	Schritte innerhalb eines Zyklus des Iterativen Fouriertransformationsalgorithmus .	21
2.5	Input-Output-Konzept als Verallgemeinerung des IFTAs aus Abbildung 2.4	21
3.1	Schritte beim Design von Strahlteilern mit Phaseneinschränkung	30
3.2	Schritte beim Design quantisierter Strahlteiler unter Verwendung der Softquantisierung	30
3.3	Optimale Lösung des regulären 1-zu-3-Strahlteilers mit Phaseneinschränkung . . .	31
3.4	Irregulärer 1-zu-3-Strahlteiler mit Signalordnungen auf äquidistantem Raster mit $\delta x' = 0.2 a$	32
3.5	Aus Faltung mit einem im Vergleich zu Abbildung 3.4(b) um den Faktor 4 breiteren Feld u_{transf} resultierende Ausgangsintensitätsverteilung	32
3.6	Irregulärer 1-zu-3-Strahlteiler mit Signalordnungen auf einem äquidistantem Raster mit $\delta x' = 0.2 a$. Während des Designs wurde das Rauschen innerhalb der um sämtliche Signalordnungen mit einem Durchmesser $d_{\text{frame}} = 0.8 a$ aufgespannten kreisförmigen Bereichen unterdrückt.	34
3.7	Schritte beim Design von Strahlteilern mit beliebiger Positionierung der Signalordnungen	37
3.8	Irregulärer 1-zu-3-Strahlteiler mit Signalordnungen im Abstandsverhältnis $1:\sqrt{2}$. .	38
3.9	Irregulärer 1-zu-30-Strahlteiler mit auf geneigten Linien positionierten Signalordnungen	38
3.10	Abhängigkeit der Intensität der 0. sowie ± 1 . Ordnung vom Höhengskalierungsfehler für das effizienzoptimierte (a) und toleranzoptimierte Design (b) des binären 1-zu-3-Strahlteilers	41
3.11	Uniformitätsfehler in Abhängigkeit vom Höhengskalierungsfehler für die beiden betrachteten Lösungen des binären 1-zu-3-Strahlteilers	42
3.12	Abhängigkeit des zulässigen Höhenfehlers $(\gamma_{\text{lin}})_{\text{max}}$ von der Anzahl der Signalordnungen n_{sig} für unterschiedliche Werte des maximal tolerierbaren Uniformitätsfehlers $(E_{\text{unif}})_{\text{max}}$	43
3.13	Lösung für einen regulären 1-zu-3-Strahlteiler mit einer auf 1 % reduzierten relativen Rauschintensität	48
3.14	Lösung für einen regulären 1-zu-3-Strahlteiler mit einer auf 0.2 % reduzierten relativen Rauschintensität	49

3.15	Lösung für einen regulären 1-zu-3-Strahlteiler mit einer auf 0.1 % reduzierten relativen Rauschintensität	50
3.16	Achtstufig quantisierte Lösung für regulären 1-zu-3-Strahlteilers mit einer auf 1.3 % reduzierten relativen Rauschintensität	51
3.17	CCD-Kameraaufnahme der für zwei verschiedene Markierungsoptiken resultierenden Fernfeldintensitätsverteilung. Im Gegsatz zu Design (a) ist bei Design (b) kein Pixelierungsrausch sichtbar und die maximale Intensität der Rauschordnungen ist auf einen niedrigeren Wert begrenzt.	52
4.1	Abhängigkeit der Konvergenzrate von der Projektionsstärke für das Design eines regulären 1-zu-3-Strahlteilers vor der ersten (a) bzw. vierten (b) Iteration der SNR-Optimierung; sowie für das Design eines Strahlteilers zur Erzeugung eines regulären Feldes von 300×300 Signalordnungen ebenfalls vor der ersten (c) bzw. vierten (d) Iteration. Vor der SNR-Optimierung wurde entsprechend des in Abbildung 3.1 gezeigten Ablaufs eine Phasensynthese mit jeweils 10 Iterationen durchgeführt. . .	57
4.2	Verlauf von SNR (a) und Projektionsstärke (b) während der Iteration zur SNR-Optimierung beim Design eines regulären 1-zu-3-Strahlteilers für IFTA (durchgängige Linie), für den Algorithmus mit Projektionsstärkenveränderung entsprechend Abschnitt 4.2.2 (gestrichelte Linie) sowie für einen Algorithmus, der die optimale Projektionsstärke vor jeder Iteration über Suche innerhalb der Kurve $v_{\text{conv}}(\lambda)$ ermittelt (gepunktete Linie)	59
4.3	Verlauf von SNR (a) und Projektionsstärke (b) während der Iteration zur SNR-Optimierung beim Design eines Strahlteilers zur Erzeugung eines regulären Feldes von 300×300 Signalordnungen für den IFTA (durchgängige Linie) und den Algorithmus mit Projektionsstärkenveränderung entsprechend Abschnitt 4.2.2 (gestrichelte Linie)	60
4.4	Verteilung der erreichten Konversionseffizienzen für Designs mit einem Signal-Rausch-Verhältnis oberhalb von 50 dB (a) bzw. 30 dB für die über die Abbildungen 3.2 und 4.5 beschriebenen Designverfahren ohne bzw. mit verstärkten Projektionen. Zur Ermittlung der Ergebnisverteilung wurden jeweils 100 Designs mit zufällig gewählten Startverteilungen durchgeführt.	62
4.5	Schritte beim Design quantisierter Strahlteiler unter Ausnutzung verstärkter Projektionen	63
4.6	Abhängigkeit der Konvergenzrate von der Projektionsstärke für einen typischen Iterationszustands während des Design eines regulären 1-zu-3-Strahlteilers mit $Q = 4$	63
4.7	Verlauf von SNR (a) und Projektionsstärke während der Iterationen des Algorithmus aus Abschnitt 4.3.2 beim Desing eines regulären 1-zu-3-Strahlteilers mit $Q = 4$	65
5.1	Verlauf von $\hat{J}_{\text{in}}(t^{(n)})^2$ bzw. $\hat{J}_{\text{in}}(\Pi_{\text{tran}} t^{(n)})^2$ während der Iteration für den PMPA, den PTDA sowie für das über die Gleichungen (5.32) beschriebene abgewandelte Verfahren mit ungewichteten Mittelungen	74
5.2	Abhängigkeit des Signal-Rausch-Verhältnisses von der Tailenbreite w_0 für die Ergebnisse der Anwendung von jeweils 5000 Iterationen des PMPAs (durchgezogen), des PTDA (gestrichelt) sowie der PTDA-Variante mit ungewichteter Transmissionsmittelung entsprechend Gleichung (5.32) (gepunktet)	75

5.3	Abhängigkeit der Konversionseffizienz η_{conv} von der Taillenbreite w_0 für die drei bereits in Abbildung 5.2 untersuchten Designs	75
5.4	Laterale Ausgangsintensitätsverteilungen der drei bereits in Abbildung 5.2 untersuchten Designs für $w_0 = 1.5 \text{ mm}$, $w_0 = 2.0 \text{ mm}$ und $w_0 = 2.5 \text{ mm}$	76
6.1	Systemaufbau bei Fotolithografie mit Mask-Aligner	80
6.2	Betrachtete Zielstruktur mit einer minimalen Strichbreite von $3 \mu\text{m}$	80
6.3	Amplitude (a) und Phase (b) der bei einem monofunktionalen Design resultierenden Transmission	82
6.4	Ausgangsintensitätsverteilung aus inkohärenter Simulation (a), ternärisiertes simuliertes Resistprofil (b), Transmissionsfotografie des Resistprofils (c) und resultierende Strukturierung in Chrom (d) für das monofunktionale Design aus Abbildung 6.3 bei dem Proximity-Abstand $d_{\text{prox}} = 50 \mu\text{m}$	83
6.5	Abhängigkeit des bei optimal gewählter Belichtungszeit resultierenden Anteils der falsch implementierten Abtastwerte vom Proximity-Abstand für mono- und multifunktionales Design	84
6.6	Ternärisiertes Resistprofil für monofunktionales (a) und multifunktionales Design (b) bei einem Proximity-Abstand $d_{\text{prox}} = 40 \mu\text{m}$	84
6.7	Ternärisiertes Resistprofil für monofunktionales (a) und multifunktionales Design (b) bei einem Proximity-Abstand $d_{\text{prox}} = 60 \mu\text{m}$	85
6.8	Amplitude (a) und Phase (b) der bei einem multifunktionalen Design resultierenden Transmission	85
6.9	Ausgangsintensitätsverteilung aus inkohärenter Simulation (a) und ternärisiertes Resistprofil (b) für das multifunktionale Design aus Abbildung 6.8 bei dem Proximity-Abstand $d_{\text{prox}} = 50 \mu\text{m}$	86
6.10	Verteilung des innerhalb des Wellenlängenbereichs $\lambda \in [594 \text{ nm}, 650 \text{ nm}]$ erreichten maximalen Uniformitätsfehlers. Es werden die drei im Text beschriebenen Designstrategien untersucht, wobei zur Ermittlung der Ergebnisverteilung jeweils 100 Designs mit zufällig gewählten Startverteilungen durchgeführt werden.	89
6.11	Mit einer CCD-Kamera aufgenommene Intensitätsverteilung für den entsprechend Design A hergestellten Strahlteiler. Innerhalb des schraffiert markierten Bereichs befinden sich die 21 Signalordnungen.	91
6.12	Verteilung des Uniformitätsfehlers in Abhängigkeit von Höhenfehlern in den Maskenschritten #1 und #2. Der schraffierte Bereich entspricht dabei dem Wertebereich beider Höhenfehler, innerhalb dessen ein Uniformitätsfehler von weniger als 5 % resultiert. Die Kreuze markieren jeweils die Positionen in beiden Diagrammen, welche den bei der Herstellung der Strahlteiler entstandenen Höhenfehlern entsprechen.	92
6.13	Abhängigkeit zwischen Uniformitätsfehler (a) und skaliertes Intensität der nullten Ordnung (b) simuliert für die Designs A und C im Vergleich zu den an drei diskreten Wellenlängen für Design C gemessenen Werten	93
A.1	Entsprechend Gleichung (4.15a) ermittelte Approximation der optimalen Projektionsstärke $\lambda^{(n+1)}$ in Abhängigkeit von der Projektionsstärke $\lambda^{(n)}$ der vorangegangenen Iteration für die in Abbildung 4.1 betrachteten Iterationszustände	102

Tabellenverzeichnis

3.1	Verbesserung der Toleranzeigenschaften gegenüber Höhenfehlern durch Verwendung von $Q > 2$ Quantisierungsstufen. Die dargestellten Werte ergeben sich gerundet aus Anwendung von Gleichung (3.47).	46
3.2	Abhängigkeit für Reduzierbarkeit der Rauschintensität von der Anzahl der Quantisierungsstufen Q	51
6.1	Simulierte und gemessene Werte für Fenstereffizienz η_{win} , Uniformitätsfehler E_{unif} und relative Intensität der nullten Ordnung I_0 für die drei betrachteten Designs. Die Simulationen und Messungen wurden für die beim Design angenommene Referenzwellenlänge $\lambda_{\text{ref}} = 632.8 \text{ nm}$ durchgeführt.	90

Index

- 3D-Strahlformung, 77
- Abstandsfehler
 - summierter, 5
 - summierter quadratischer, 5
- Abstandsminimierung, 4
 - punktweise, 4
- Abtastabstand, 9
- Abtastung, 9
- Achse
 - optische, 9
- Amplitude
 - komplexe, 9
- ASPW, 11

- Dünne-Elemente-Approximation, 11, 13, 19, 28, 38, 39, 46, 49, 78, 81
- Design
 - monofunktionales, 16
 - multifunktionales, 16
- Designfreiheit, 17
- Direct Binary Search, 2, 77
- Dreiecksungleichung, 4

- Effizienz, 19
- Einschränkung, 3
 - implizite, 5
 - konvexe, 5
- Einschränkungsmenge, 5

- Feld, 9
- Fenstereffizienz, 18
- FFT, 12, 26
- Fouriertransformation, 11
 - inverse, 11
 - punktweise, 35
 - schnelle, 12, 26
- Fraunhofersche Fernfeldnäherung, 12
- Fresnel-Näherung, 12
- FT, 11
- Funktion
 - optische, 16
- Funktionsbild, 1

- Gitter, 12
- Gittergleichung, 13

- Höhenfehler, 77
 - globaler, 39, 85
 - linearer, 39
- Hilbertraum, 4

- Idempotenz, 5
- IFTA, 20
- Iterativer Fouriertransformationalgorithmus, 20

- Kammfunktion, 25
- Konsistenz, 5
- Konvergenzgeschwindigkeit, 54
- Konvergenzrate, 54
- Konversionseffizienz, 18

- Mask-Aligner, 80
- Maskentechnik
 - binäre, 87
- Metrik, 4

- Näherung der effektiven Medien, 13
- Norm, 4

- Operator
 - funktionsbezogen, 11
 - strukturbezogen, 11
- Optimal Rotation Angle Method, 2, 77
- ORA, 2, 77
- Ordnung, 26

- Paralleler Projektionsalgorithmus, 7
- Parselvalsches Theorem, 12
- Periodenlänge, 25
- Phaseneinschränkung, 19
- Phasenfreiheit, 17

- PMPA, 7
- Projektionsalgorithmus, 1, 3
 - serieller, 7
- Projektionsoperator, 4
 - lokaler, 4
 - relaxierter, 6
 - verstärkter, 6
- Propagation
 - sequentielle, 10
- Propagationsoperator
 - rigoroser, 12
- Proximity printing, 79
- Proximity-Abstand, 80
- PTDA, 69

- Quantisierungsoperator, 20

- Rauschordnung, 24, 26
- Rechteckfunktion, 27
- Referenzebenen, 10

- SDE, 5
- SDE-Konvergenz, 6
- Signal-Rausch-Verhältnis, 17
- Signalebene, 10
- Signaleinschränkung, 17
- Signalfeld, 17
- Signalordnung, 24, 26
- Simulated Annealing, 2
- Simultaneous Perturbation Approach, 2
- sinc-Funktion, 27
- sinc-Modulation, 27
- Skalarprodukt, 4
- Skalierungsfaktor, 39
 - optimaler, 17
- Skalierungsfehler
 - linearer, 78
- Skalierungsfreiheit, 17
- SNR, 17
- Softquantisierung, 61
- Speckles, 55
- Spektrum ebener Wellen, 11
- SSDE, 5
- Steuerfeld, 23
- Strahlformung, 73
 - 3D, 77
 - axiale, 77
- Strahlteiler, 24
 - überlappungsfreier, 26, 49
- Strahlteilung, 61, 86
- Strukturdesign, 1, 11, 78
- Subwellenlängengitter, 13
- Summenprojektionen, 36
- Systemparameter
 - fester, 15
 - freier, 15

- Ternärisierung, 82
- Toleranzeigenschaften, 39
- Transmissionsdesign, 1, 3, 11, 16
- Transmissionsdesignalgorithmus
 - Paralleler, 69
- Transmissionseinschränkung, 17
- Transmissionsfunktion, 11
- Transmissionsoperator, 11

- Überlappungsfreiheit, 26, 49
- Umsetzungsfehler, 15, 86
- Uniformitätsfehler, 18

- Verkippungen, 46
- Verschiebung
 - laterale, 45
 - longitudinale, 46

- Wave-Optical Engineering, 1
- Wellenlängenfehler, 47, 77, 78, 85

Danksagung

Allen, die zum Gelingen dieser Arbeit beigetragen haben, möchte ich an dieser Stelle meinen Dank aussprechen.

Prof. Dr. Frank Wyrowski danke ich für seine Unterstützung im Rahmen der Betreuung dieser Arbeit sowie für die bei den regelmäßigen fachlichen Gesprächen erhaltenen hilfreichen Anregungen.

Auch Hagen Schimmel danke für die zahlreichen in fachlichen Diskussionen gegebenen Anregungen, die den Fortgang dieser Arbeit beförderten.

Petra Wyrowski danke ich für die Möglichkeit, diese Dissertation im Rahmen meiner Anstellung bei der LightTrans GmbH anfertigen zu können.

Für ihren Einsatz bei der Herstellung der in dieser Arbeit untersuchten Strahlteiler möchte ich den Mitgliedern der Arbeitsgruppe Mikrostrukturtechnik und Mikrooptik am Institut für Angewandte Physik der Friedrich-Schiller-Universität Jena danken.

Ganz besonderen Dank verdient Daniela Stephan für die vielfältige und kontinuierliche Unterstützung im privaten Umfeld.

Lebenslauf

Name Bühling
Vorname Sven
Geburtsdatum 4. Dezember 1972
Familienstand ledig

1979 – 1987 Polytechnischen Oberschule “Theodor Neubauer” in Hermsdorf
1987 – 1991 Spezialschule mathematisch-naturwissenschaftlich-technischer Richtung “Carl Zeiss” in Jena
Juni 1991 Abitur
1991 – 1992 Zivildienst
1992 – 2000 Studium in den Fachrichtungen Physik und Informatik an der Friedrich-Schiller-Universität Jena
Oktober 1998 Diplom Informatik
Februar 2000 Diplom Physik
1999 Wissenschaftlicher Mitarbeiter der Friedrich-Schiller-Universität Jena im Institut für Angewandte Physik
seit 2000 Angestellter der LightTrans GmbH Jena

Jena, den 2. November 2002

Sven Bühling

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung des dargestellten Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen:

1. Herr Prof. Dr. Frank Wyrowski als Betreuer dieser Arbeit;
2. Mitglieder der Arbeitsgruppe Mikrostrukturtechnik und Mikrooptik am Institut für Angewandte Physik der Friedrich-Schiller-Universität Jena mit der Herstellung der Strahlteiler-elemente, auf denen die in den Abschnitten 3.2, 3.4 und 6.2 dargestellten experimentellen Ergebnisse basieren, mit der Durchführung der Strukturiefenmessungen dieser Elemente sowie mit der Herstellung der in Abschnitt 6.1.2 untersuchten komplexwertigen Fotomaske;
3. A. J. M. Nellissen, Lingli Wang und Maarten Dirkswager von Philips Centre for Manufacturing Technology Eindhoven mit der Durchführung der in Abschnitt 6.1.2 dargestellten Experimente.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Person) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für die Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die geltende Promotionsordnung der Physikalisch-Astronomischen Fakultät ist mir bekannt.

Ich versichere ehrenwörtlich, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Jena, den

Sven Bühling