

DIPLOMARBEIT

**Numerische Berechnung  
von komplexen Lichtmasken  
zur atomaren Nanofabrikation**

Angefertigt am Institut für Angewandte Mathematik

Vorgelegt der

Mathematisch–Naturwissenschaftlichen Fakultät der  
Rheinischen Friedrich–Wilhelms–Universität Bonn

Juli 2004

Von

**Jürgen Braun**

aus Bornheim



## Zusammenfassung

Die atomare Nanofabrikation ist eine Methode der Experimentalphysik zur Konstruktion von im Allgemeinen zweidimensionalen Nanostrukturen. Dazu gehört die neutrale Atomlithographie, in der die zweidimensionale Intensitätsverteilung eines Atomstrahls durch ein inhomogenes Lichtfeld kontrolliert und auf einem Substrat abgebildet wird. Das Lichtfeld wird durch Überlagerung stehender Lichtwellen erzeugt und spielt die entscheidende Rolle für die Komplexität des Musters auf dem Substrat.

Die zentrale Aufgabenstellung, die hier behandelt wird, ist die Rekonstruktion eines gegebenen Musters durch geeignete Wahl von  $N$  regelmäßig angeordneten Laserstrahlen. Zunächst wird für die Anordnung der Strahlen ein Verfahren zur Konstruktion von Setups vorgestellt und das in [5] eingeführte mathematische Modell auf den Versuchsaufbau der regelmäßigen Strahlen ausgeweitet. Dies führt ebenfalls zu einem inversen, komplexen, nichtlinearen Minimierungsproblem. Ansätze zur analytischen Bestimmung des Minimums sowie zur Approximation mit Ridge-Functions stellen sich als ungeeignet heraus.

Zur numerischen Rekonstruktion wird, zusätzlich zum Koordinatenabstiegsverfahren aus [5], die Nelder und Mead Simplexmethode vorgestellt. Es werden diese beiden Verfahren durch Rekonstruktion verschiedener Muster mit unterschiedlichen Strahlensetups getestet und anhand berechneter Fehler und Normen verglichen und bewertet.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
<b>2</b>	<b>Das mathematische Modell</b>	<b>7</b>
<b>3</b>	<b>Das inverse Problem</b>	<b>10</b>
3.1	Konstruktion von Strahlensetups . . . . .	10
3.2	Symmetrien und lineare Abhängigkeiten . . . . .	13
3.3	Die diskrete Fouriertransformation . . . . .	15
3.4	Approximationseigenschaften der diskreten Fouriertransformation . . . . .	19
<b>4</b>	<b>Die Rekonstruktion der Koeffizienten</b>	<b>30</b>
4.1	Berücksichtigung der linearen Abhängigkeiten . . . . .	31
4.2	Minimierungsverfahren . . . . .	33
4.2.1	Koordinatenabstiegsverfahren . . . . .	33
4.2.2	Die Nelder und Mead Simplexmethode . . . . .	36
4.2.3	Die Nelder und Mead Simplexmethode auf $\mathbb{S}_r^{2N-1}$ . . . . .	38
4.2.4	Der gesamte Algorithmus . . . . .	39
<b>5</b>	<b>Rang Eins Approximation</b>	<b>41</b>
<b>6</b>	<b>Ridge-Funktionen</b>	<b>45</b>
<b>7</b>	<b>Numerische Ergebnisse</b>	<b>50</b>
7.1	Berechnung verschiedener Muster . . . . .	50
7.1.1	Die Rekonstruktion mit vier Strahlen . . . . .	51
7.1.2	Die Rekonstruktion mit acht Strahlen . . . . .	62
7.1.3	Die Rekonstruktion mit zwölf Strahlen . . . . .	71
7.2	Größe des Fouriergitters . . . . .	82
7.2.1	A priori Fehlerschätzungen für die diskrete Fouriertransformation . . . . .	82
7.2.2	Numerische Untersuchungen . . . . .	83
7.3	Koordinatenabstiegsverfahren mit verschiedenen Startwerten . . . . .	92
7.4	Konvergenzgeschwindigkeit des Koordinatenabstiegsverfahrens . . . . .	93
<b>8</b>	<b>Zusammenfassung und Ausblick</b>	<b>96</b>
	<b>Anhang</b>	<b>98</b>

# 1 Einleitung

Die neutrale Atomlithographie ist eine Methode der atomaren Nanofabrikation zur Generierung von Strukturen durch die direkte Steuerung eines Atomstrahls.

Laserkühlung [14] und Atomoptik [1] haben den Weg zur Manipulation der Bewegung von Atomen in Nanometerbereich geebnet. In der atomaren Nanofabrikation werden atomare Dipolkräfte zur Steuerung von Atomstrahlen in Nanostrukturen verwendet.

Im Gegensatz zur klassischen Lithographie, in der das Licht zur Projektion auf ein Substrat durch dünne Filme und Masken gesandt wird, sind die Rollen von Licht und Material hier in gewisser Weise vertauscht und die inhomogene Intensitätsverteilung des Lichtfeldes verhält sich wie eine immaterielle Maske. Atome werden durch ein Gebiet stehender Lichtwellen gesandt und die Intensitätsverteilung des Lichts wird in das atomare Bild auf dem Substrat transformiert.

In [21] wird ein Überblick über den Themenbereich der Nanofabrikation unter Verwendung neutraler Atomstrahlen gegeben. Des Weiteren führt die Forschungsgruppe von Prof. Meschede am Institut für Angewandte Physik an der Universität Bonn, aber auch Arbeitsgruppen an einigen anderen Universitäten [3, 20], Experimente zu diesem Thema durch. Die Zusammenarbeit mit Prof. Meschede gab Anstoß zu den mathematischen Untersuchungen in dieser Arbeit, sowie in [5] und resultierte in der gemeinsamen Arbeit [15].

Das grundlegende Konzept der neutralen Atomlithographie ist das Folgende: Mittels Laserkühlung werden die Atome eines Atomstrahls auf eine bestimmte Geschwindigkeit gebracht, um anschließend auf die immaterielle Maske der inhomogenen Intensitätsverteilung eines Lichtfeldes zu treffen. Dort erfahren sie durch das elektromagnetische Feld, welches durch das Laserlicht erzeugt wird, Dipolkräfte. Diese Kräfte lenken die Atome zu den lokalen Maxima der Intensitätsverteilung ab. Durch eine andere Wahl der Frequenz der Laserstrahlen, welche die Maske erzeugen, kann dieser Effekt auch umgekehrt werden, so dass die Atome zu den lokalen Minima hin abgelenkt werden. Zur genaueren Analyse dieses Effekts sei auf [14, 15] verwiesen.

Nachdem die Atome diese immaterielle Maske „durchdrungen“ haben, ist ihre Dichteverteilung orthogonal zur Richtung des Atomstrahls fast identisch mit der inhomogenen Intensitätsverteilung des Lichtfeldes.

Die Atome kondensieren dann auf der Oberfläche eines Substrats. Nach einer chemischen Behandlung dieses Substrats erhält man eine Abbildung der atomaren Dichteverteilung und somit der Intensitätsverteilung des Lichtfeldes auf der Oberfläche.

Somit ist das Muster, welches abgebildet wird durch die Wahl des Lichtfeldes beeinflussbar, das durch interferierende Laserstrahlen erzeugt wird. Eine Möglichkeit zur Generierung des Interferenzmusters besteht darin, einen Laserstrahl mit bestimmter Frequenz durch Strahlleiter und Spiegel aufzusplitten [7, 8]. Diese Vorgehensweise wird allerdings sehr schnell unhandlich, wenn mehr als zwei oder drei Strahlen benötigt werden.

Daher verwendet die Bonner Arbeitsgruppe von Prof. Meschede zur Speicherung von Hologrammen Kristalle, welche zur Aufspaltung eines Laserstrahls in mehrere Strahlen geeignet sind und sogar sehr nah beieinander liegen können.

Zur Speicherung eines Hologramms ist ein aufwendiger Prozess notwendig, welcher bis zu zwei Stun-

den in Anspruch nehmen kann und eine Erhitzung des Kristalls auf  $180^{\circ}\text{C}$  benötigt. Dieser Prozess heißt *Schreiben (writing)*, während der Vorgang der Aufspaltung eines Laserstrahls als *Lesen (reading)* bezeichnet wird [15]. Es können auf diese Weise oder mit einem alternativen Beugungsobjekt, Setups mit bis zu 5000 Strahlen gespeichert werden [16]. Da das Schreiben eines Hologramms so aufwendig ist, sind vorherige numerische Simulationen zur Berechnung des zu erwartenden Musters sinnvoll. In diesem Zusammenhang stellt sich unmittelbar die Frage, ob durch eine geeignete Wahl des Strahlensetups eine Intensitätsverteilung des Lichtfeldes erzeugt werden kann, welche ein Muster auf dem Substrat generiert, das eine vorgegebene Struktur hat.

Dies entspricht der folgenden fundamentalen Fragestellung: Finde zu einem gegebenen Muster ein geeignetes Strahlensetup, d.h. im Speziellen Anzahl, Richtung und Koeffizienten der Strahlen, welches eine inhomogene Intensitätsverteilung erzeugt, die, nachdem sie auf das Substrat gebracht wurde, die gegebene Struktur des Musters „möglichst gut“ rekonstruiert.

Zu Beginn dieser Arbeit wird die in [5] vorgeschlagene, mathematische Modellierung noch einmal kurz erläutert. Da in der Praxis periodische Muster erzeugt werden sollen, konzentrieren sich die Untersuchungen anschließend auf einen speziell gewählten Versuchsaufbau von  $N$  regelmäßig angeordneten Strahlen, welcher in der allgemeinen Modellierung aus [5] ausgeschlossen werden musste. Dieser Versuchsaufbau wird zunächst theoretisch erläutert und eine Möglichkeit zur Bestimmung einer passenden Strahlenanordnung vorgestellt. Die Festlegung des Versuchsaufbaus reduziert die Freiheitsgrade auf die Entwicklungskoeffizienten der Laserstrahlen. Dies führt zu einem inversen, nichtlinearen Minimierungsproblem.

Zur Lösung des Problems wird der diskrete Fourierraum als Ansatzraum verwendet. Da dieser den kontinuierlichen Fourierraum nur approximiert, werden in Abschnitt 3.4 Fehlerschätzungen für die diskrete Fouriertransformation der rekonstruierten Muster hergeleitet.

Zu Beginn von Abschnitt 4 wird, aufgrund der zuvor dargestellten Zusammenhänge, das Problem genauer spezifiziert, wodurch in Abschnitt 4.1 unter Berücksichtigung der regelmäßigen Anordnung der Laserstrahlen wie sie in Abschnitt 3.1 konstruiert wurden, ein Fehlerfunktional definiert werden kann. Dieses Kostenfunktional soll nun minimiert werden, was zur endgültigen Aufgabenstellung führt, welche durch numerische Methoden gelöst werden soll, die im Folgenden vorgestellt werden. In Abschnitt 4.2 wird das in [5] suggestiv vorgeschlagene Minimierungsverfahren als *Koordinatenabstiegsverfahren* identifiziert und eine Konvergenzaussage für dieses Verfahren gemacht.

Alternativ wird die *Nelder und Mead Simplexmethode* zur Minimierung des Kostenfunktionals vorgestellt. Diese sollte allerdings nicht mit dem Simplexverfahren zur linearen Optimierung verwechselt werden. In dieser Arbeit ist mit *Simplexmethode* immer die *Nelder und Mead Simplexmethode* gemeint. Das Simplexverfahren eröffnet Möglichkeiten zur Berücksichtigung von Nebenbedingungen, welche bei der Verwendung des Koordinatenabstiegsverfahrens noch vernachlässigt wurden. Dazu werden zwei Strategien verwendet, wovon die eine die Nebenbedingung minimiert und die andere die Nebenbedingung erfüllt.

Zur analytischen Lösung dieses und auch des allgemeinen Problems aus [5] wird in Abschnitt 5 der Ansatz der *Rang Eins Approximation* vorgestellt, welcher die wesentliche Struktur des Problems widerspiegelt, sich aber bei näherer Betrachtung als ungeeignet erweist. Die Gründe hierfür werden in diesem Abschnitt erläutert.

Abschnitt 6 versucht das hier gegebene Minimierungsproblem in den approximationstheoretischen

Rahmen der *Ridge-Functions* einzuordnen. Dazu wird eine kurze Einleitung in die Theorie der Ridge-Functions gegeben, welche die wesentlichen Definitionen und Bezeichnungen bereitstellt. Dieser Ansatz stellt sich allerdings in diesem Zusammenhang als nur sehr bedingt geeignet heraus, da die Definition eines geeigneten Approximationsraumes mit Ridge-Functions als Basis nicht möglich ist. Abschnitt 7 widmet sich nun den numerischen Resultaten und deren Bewertung. In Abschnitt 7.1 werden dazu zunächst die zuvor beschriebenen Verfahren für verschiedene Muster angewandt. Es werden für drei verschiedene Strahlensetups ( $N = 4, 8, 12$ ) Rekonstruktionen für alle Muster berechnet, was einen Vergleich der Minimierungsverfahren erlaubt.

Nachdem diese Resultate vorgestellt wurden, beschäftigt sich Abschnitt 7.3 mit der Abhängigkeit des Koordinatenabstiegsverfahrens von den gewählten Startwerten und Abschnitt 7.2 mit der gewählten Auflösung des Fouriergitters. Es werden sowohl die zuvor in Abschnitt 3.4 gewonnenen a priori Fehlerschätzer angewandt als auch numerisch verschiedene Auflösungen der diskreten Fouriertransformation zur Rekonstruktion getestet.

Als letztes numerisches Resultat wird in Abschnitt 7.4 die Konvergenzgeschwindigkeit des Koordinatenabstiegsverfahrens untersucht.

Zum Abschluss werden in Abschnitt 8 noch einmal alle theoretischen und numerischen Resultate kurz zusammengefasst sowie ein Ausblick über offene Probleme und lohnenswerte Untersuchungen gegeben.

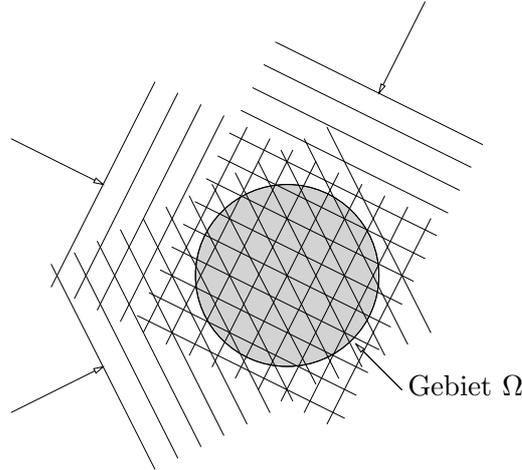


Abbildung 2.1: Schematisches Modell der interferierenden Laserstrahlen

## 2 Das mathematische Modell

Zu Beginn der theoretischen und numerischen Untersuchungen muss das Problem der Rekonstruktion eines vorgegebenen Musters mittels neutraler Atomlithographie mathematisch formuliert werden. Das bedeutet, dass ein Funktionenraum definiert werden muss, der genau die physikalisch konstruierbaren Muster  $f_N$  enthält. Dieser Raum wird durch Vorwärtsberechnung von Mustern identifiziert und mit  $S_N$  bezeichnet. Der Index  $N$  gibt dabei die Anzahl der verwendeten Laserstrahlen an.

Die folgende allgemeine mathematische Modellierung der Funktion  $f_N$  für das auf dem Substrat abgebildete Muster wurde in [5] vorgeschlagen. Abb. 2.1 zeigt das schematische Modell des Interferenzmusters der Laserstrahlen. Die Pfeile geben die Ausbreitungsrichtung der Lichtstrahlen an, die parallelen Linien symbolisieren die Wellenfronten. Das Gebiet  $\Omega \subset \mathbb{R}^2$  stellt das Substrat dar, über dem das Interferenzmuster kontrolliert werden soll. Interferieren  $N$  Laserstrahlen über dem Gebiet  $\Omega$ , so kann eine einzelne Welle mit Index  $j \in I := \{1, \dots, N\}$  als komplexwertige Funktion  $y_j(\mathbf{x}, t; \mathbf{k}_j, a_j, \varphi_j)$  dargestellt werden. Der Funktionswert zum Zeitpunkt  $t \in [0, T]$  am Ort  $\mathbf{x} \in \Omega$  ist abhängig von drei Parametern: Dem *Wellenvektor*  $\mathbf{k}_j \in \mathbb{R}^2$ , welcher die Ausbreitungsrichtung und Wellenlänge der Welle angibt, der *Amplitude*  $a_j \in \mathbb{R}_{\geq 0}$  und der *Phase*  $\varphi_j \in [0, 2\pi)$ . Eine einzelne Welle hat somit die Form

$$\begin{aligned} y_j(\mathbf{x}, t; \mathbf{k}_j, a_j, \varphi_j) &= a_j \exp(i(\langle \mathbf{k}_j, \mathbf{x} \rangle + \varphi_j - \omega t)) \\ &= c_j \exp(i\langle \mathbf{k}_j, \mathbf{x} \rangle) \exp(-i\omega t) , \end{aligned}$$

mit den komplexen Koeffizienten

$$c_j := a_j \exp(i\varphi_j) .$$

$\langle \cdot, \cdot \rangle$  bezeichnet hier sowie im kompletten Verlauf der Arbeit das Standardskalarprodukt des  $\mathbb{R}^2$ . Die Frequenz  $\omega$  ist durch die Wahl des Lasers bestimmt, für alle  $N$  Strahlen gleich und daher nicht mit den Parametern aufgelistet, von denen das Muster abhängt. Verändert man die Wellenlänge  $\lambda = \frac{2\pi c}{\omega}$ , so ändert sich die Struktur des Musters nicht, sondern wird nur mit dem Ursprung als Zentrum gestreckt, wobei  $c$  die Lichtgeschwindigkeit bezeichnet. Sie stellt somit keinen Freiheitsgrad zur Rekonstruktion dar und für die Beträge aller  $N$  Wellenvektoren gilt  $K = |\mathbf{k}_j| = \frac{\omega}{c}$ .

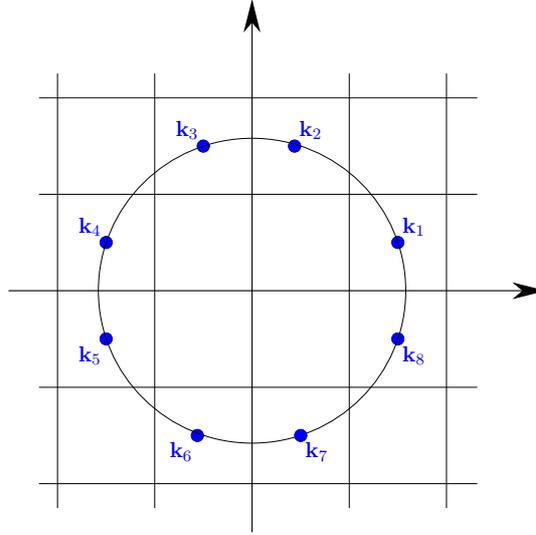


Abbildung 2.2: Lage und Anordnung der Wellenvektoren  $\mathbf{k}_j \in \mathcal{K}$  definiert in (2.2)

Im Folgenden bezeichne  $\mathcal{K}$  die Menge

$$\mathcal{K} := \{\mathbf{k}_j \in \mathbb{S}_K^1 : j \in \mathcal{I} := \{1, \dots, N\}\} \quad (2.1)$$

der  $N$  Wellenvektoren.  $\mathbb{S}_K^1 \subset \mathbb{R}^2$  ist ein Kreis mit Radius  $K \in \mathbb{R}$ . Die präzise Wahl der Vektoren in  $\mathcal{K}$  wird später genauer spezifiziert (siehe Abschnitt 3.1 und Bemerkung 3.2). In Abb. 2.2 ist die Lage der Wellenvektoren aus

$$\mathcal{K} := \left\{ r \begin{pmatrix} 1,5 \\ 0,5 \end{pmatrix}, r \begin{pmatrix} 0,5 \\ 1,5 \end{pmatrix}, r \begin{pmatrix} -0,5 \\ 1,5 \end{pmatrix}, r \begin{pmatrix} -1,5 \\ 0,5 \end{pmatrix}, \right. \\ \left. r \begin{pmatrix} -1,5 \\ -0,5 \end{pmatrix}, r \begin{pmatrix} -0,5 \\ -1,5 \end{pmatrix}, r \begin{pmatrix} 0,5 \\ -1,5 \end{pmatrix}, r \begin{pmatrix} 1,5 \\ -0,5 \end{pmatrix} : r = \frac{2K}{\sqrt{10}} \right\} \quad (2.2)$$

im  $\mathbb{R}^2$  und deren Nummerierung für  $N = 8$  exemplarisch dargestellt.

Die Überlagerung der  $N$  Wellen ergibt sich durch Addition der einzelnen Funktionen  $y_i$  und hat somit die Form

$$y(\mathbf{x}, t) = \sum_{j=1}^N y_j(\mathbf{x}, t) = \sum_{j=1}^N c_j \exp(i\langle \mathbf{k}_j, \mathbf{x} \rangle) \exp(-i\omega t).$$

Setzt man nun

$$\tilde{y}(\mathbf{x}) := \sum_{j=1}^N c_j \exp(i\langle \mathbf{k}_j, \mathbf{x} \rangle), \quad (2.3)$$

so entkoppelt  $y$  in Zeit und Ort und es ist

$$y(\mathbf{x}, t) = \tilde{y}(\mathbf{x}) \exp(-i\omega t).$$

Physikalisch stellt der Realteil der Funktion  $y$  das elektrische Feld dar, welches die Bahnen der Atome beeinflusst. Somit ergibt sich eine Funktion  $f_N$  für das auf dem Substrat abgebildete Muster

als zeitlicher Mittelwert der Intensität durch

$$f_N(\mathbf{x}) = \frac{1}{T} \int_0^T (\operatorname{Re} y(\mathbf{x}, t))^2 dt = \frac{1}{2} |\tilde{y}(\mathbf{x})|^2, \quad \mathbf{x} \in \Omega. \quad (2.4)$$

Hierbei bezeichne  $T = \frac{2\pi}{\omega}$  eine Periode der Oszillation, von der die Funktion allerdings nicht abhängt. Der Index von  $f_N$  soll die Anzahl der verwendeten Strahlen und somit der komplexen Freiheitsgrade des Musters illustrieren.

Die Funktion (2.4)

$$f_N(\mathbf{x}) = \frac{1}{2} \left| \sum_{j=1}^N c_j \exp(i\langle \mathbf{k}_j, \mathbf{x} \rangle) \right|^2, \quad (2.5)$$

kann, die Kenntnis der Größen  $\mathbf{k}_j$ ,  $c_j$  und  $N$  vorausgesetzt, leicht durch Vorwärtssimulation berechnet werden. In [5] werden ein solches Simulationsprogramm und einige berechnete Muster eines allgemeineren Strahlensetups vorgestellt.

### 3 Das inverse Problem

Im Folgenden sei durch  $f$  stets ein beliebiges, aber fest gewähltes Muster bezeichnet. Das Ziel der weiteren Untersuchungen ist, ein solches Muster  $f$  durch  $f_N$  möglichst gut zu rekonstruieren. Gesucht ist somit ein  $f_*$ , für welches

$$f_* = \arg \min_{g \in \mathcal{S}_N} \|g - f\|_{\Omega} , \quad (3.1)$$

mit

$$\mathcal{S}_N = \mathcal{S}_N(\mathbf{k}_1, \dots, \mathbf{k}_N) := \left\{ g = \frac{1}{2} \left| \sum_{j=1}^N c_j \exp(i\langle \mathbf{k}_j, \cdot \rangle) \right|^2 : c_j \in \mathbb{C} \right\} , \quad (3.2)$$

in einer noch genauer zu spezifizierenden Norm  $\|\cdot\|_{\Omega}$  gilt. Der durch (3.2) definierte Raum  $\mathcal{S}_N$  enthält alle physikalisch realisierbaren Muster, bei vorgegebener Anzahl und Richtung der Laserstrahlen. Dabei sind die  $\mathbf{k}_j \in \mathcal{K}$ ,  $j \in \mathcal{I}$  paarweise verschieden und ebenso wie  $N$  fest gewählt, d.h. vorgegebene Größen. Die  $N$  komplexen Koeffizienten  $c_j$  sind somit die einzigen Freiheitsgrade zur Konstruktion einer Funktion aus  $\mathcal{S}_N(\mathbf{k}_1, \dots, \mathbf{k}_N)$ . Im folgenden Verlauf der Arbeit wird die Menge  $\mathcal{S}_N(\mathbf{k}_1, \dots, \mathbf{k}_N)$  zur Abkürzung stets mit  $\mathcal{S}_N$  bezeichnet. Zur genaueren Untersuchung von  $\mathcal{S}_N$  kann die Summe (2.5) ausmultipliziert werden, was zu der Gleichung

$$2f_N(\mathbf{x}) = \sum_{j=1}^N |c_j|^2 + \sum_{\substack{j,l=1 \\ j < l}}^N \left( c_j \bar{c}_l \exp(i\langle \mathbf{k}_{j,l}, \mathbf{x} \rangle) + \text{k. K.} \right) \quad (3.3)$$

führt. Hierbei ist  $\mathbf{k}_{j,l}$  durch  $\mathbf{k}_{j,l} := \mathbf{k}_j - \mathbf{k}_l$  definiert und k. K. bezeichne das komplex Konjugierte des jeweiligen Summanden. Es ist offensichtlich, dass die Vorgabe der  $\mathbf{k}_j$  das Muster wesentlich beeinflusst. Daher wird die Menge  $\mathcal{K}$  der Wellenvektoren im nächsten Abschnitt genauer festgelegt.

#### 3.1 Konstruktion von Strahlensetups

Die Wahl der Anzahl  $N$  der Elemente aus  $\mathcal{K}$  sowie deren Lage auf  $\mathbb{S}_K^1$  wurden bisher noch nicht genauer spezifiziert. Der Radius des Kreises  $\mathbb{S}_K^1$  ist durch den verwendeten Laserstrahl vorgegeben und ergibt sich als Quotient  $K = \frac{\omega}{c}$  aus Frequenz des Lasers und der Lichtgeschwindigkeit. Es bleibt somit noch eine sinnvolle Wahl von  $N$  und  $\mathbf{k}_1, \dots, \mathbf{k}_N$ .

Da die berechneten Muster periodisch sein sollen, müssen die Wellenstrahlen in rationalem Verhältnis zueinander stehen. Dies ist in dem Sinne gemeint, dass alle Differenzen der Wellenvektoren die Bedingung

$$\mathbf{k}_{j,l} \in r\mathbb{Z}^2 \quad \text{für alle } j, l \in \mathcal{I} \quad (3.4)$$

und beliebiges  $r \in \mathbb{R}$  erfüllen müssen, wobei  $\mathcal{I} = \{1, \dots, N\}$  ist. Eine hinreichende Wahl für die  $\mathbf{k}_j$  wäre somit  $\mathbf{k}_j \in r\mathbb{Z}^2$ ,  $j \in \mathcal{I}$ . Dabei ist natürlich die Bedingung  $|\mathbf{k}_j| = K$  für alle  $j \in \mathcal{I}$  zwingend zu erfüllen, was  $r$  von  $K$  abhängig macht. Die im Folgenden vorgestellte Methode zur Konstruktion  $N$  geeigneter Wellenvektoren zeigt, dass es möglich ist ein  $r' \leq r$  zu finden, so dass die Richtungen die Bedingung

$$\mathbf{k}_j \in r'\mathbb{Z}^2 \quad j \in \mathcal{I} \quad (3.5)$$

erfüllen und die daraus resultierenden Differenzen  $\mathbf{k}_{j,l} = \mathbf{k}_j - \mathbf{k}_l$ ,  $j, l \in \mathcal{I}$  der Bedingung (3.4) genügen. Natürlich ist auch  $r' = r'(K)$  durch die Eigenschaft  $|\mathbf{k}_j| = K$  von  $K$  abhängig und muss nicht notwendiger Weise kleiner als  $r$  sein.

Im Folgenden sei eine Methode vorgestellt, mit der ein Strahlensetup generiert werden kann, welches die Bedingungen (3.4) und (3.5) erfüllt und somit die Generierung von periodischen Mustern ermöglicht.

Im ersten Schritt muss ein  $R \in \mathbb{N} \setminus \{0\}$  gewählt werden, für das anschließend alle möglichen 2-Tupel  $(x, y)$  gefunden werden müssen, deren Elemente die Bedingung

$$x^2 + y^2 = R, \quad 0 \leq x \leq y \quad (3.6)$$

erfüllen, wobei auch  $x, y \in \mathbb{N}$  sind. Diese werden im Folgenden auch *gültige Tupel* genannt. Die Bestimmung dieser 2-Tupel ist ein zahlentheoretisches Problem und es ist nicht für jede Wahl von  $R$  garantiert, dass diese überhaupt existieren (z.B.  $R = 3$ ). Dies schränkt obige Wahl von  $R$  natürlich ein, wobei eine exakte und detaillierte Definition dieser Einschränkung hier nicht gegeben wird, da diese eine eingehende zahlentheoretische Untersuchung erfordert. An dieser Stelle sei einfach nur heuristisch gefordert: Wähle  $R \in \mathbb{N} \setminus \{0\}$  so, dass mindestens ein 2-Tupel  $(x, y)$  existiert, welches die Bedingung (3.6) erfüllt.

Ist  $R$  eine Quadratzahl  $R = z^2$ , so entspricht die Suche nach 2-Tupeln  $(x, y)$  der Bestimmung von pythagoräischen Tripeln  $(x, y, z)$  mit gleichem  $z$ .

Die Anzahl der bestimmten 2-Tupel liefert eine Indexmenge  $\mathcal{I}(R)$ , und die gefundenen Tupel seien nun mit  $(x_j, y_j)$ ,  $j \in \mathcal{I}(R)$  bezeichnet, wobei es auf die Reihenfolge nicht ankommt. Man erhält somit die Menge

$$\{(x_j, y_j) \in \mathbb{Z} \times \mathbb{Z} : j \in \mathcal{I}(R) ; x_j^2 + y_j^2 = R ; 0 \leq x_j \leq y_j\} .$$

Die Mächtigkeit  $\#\mathcal{I}(R)$  dieser Menge bestimmt über die Relation

$$N = 8 \#\mathcal{I}(R) - 4 \mathcal{N}(\mathcal{I}(R))$$

die gesamte Anzahl der Strahlen, wobei  $\mathcal{N}(\mathcal{I}(R))$  die Anzahl der 2-Tupel  $(x, y)$  ist, für die  $x = 0$  oder  $x = y$  gilt.  $\mathcal{N}$  kann nur die Werte 0 oder 1 annehmen. Der Faktor rührt daher, dass mit dem gültigen 2-Tupel  $(x_j, y_j)$  die Tupel

$$(y_j, x_j), (-x_j, -y_j), (-y_j, -x_j), (-x_j, y_j), (x_j, -y_j), (-y_j, x_j), (y_j, -x_j), \quad j \in \mathcal{I}(R) \quad (3.7)$$

ebenfalls alle die Bedingung  $x_j^2 + y_j^2 = R$  erfüllen. Ist für ein  $j \in \mathcal{I}(R)$   $x_j = 0$  oder  $x_j = y_j$ , so müssen vier Möglichkeiten subtrahiert werden, da  $-0 = 0$  ist und für  $x_j = y_j$  die Tupel  $(x_j, y_j)$   $(y_j, x_j)$  bzw.  $(-x_j, -y_j)$   $(-y_j, -x_j)$  identisch sind (Bsp.:  $R = 25 \Rightarrow (0, 5)$  und  $(3, 4)$  sind mögliche Tupel  $\Rightarrow \#\mathcal{I}(R) = 2$  und  $\mathcal{N}(\mathcal{I}(R)) = 1 \Rightarrow N = 12$ ).

$N$  ist somit auch von  $R$  abhängig und das Problem der Wahl eines  $R$ 's, welches ein vorgegebenes  $N$  erzeugt, ist ebenfalls von zahlentheoretischer Natur.

Nimmt man die zusätzlichen Möglichkeiten aus (3.7) mit in der Menge auf, so führt dies zur Menge aller gültigen 2-Tupel für das gewählte  $R$ .

Es sei nun  $\mathcal{I}_R$  eine Indexmenge, die alle  $N = 8 \#\mathcal{I}(R) - 4 \mathcal{N}(\mathcal{I}(R))$  Tupel durchläuft, d.h. auch diejenigen in (3.7), wobei die Reihenfolge immer noch keine Rolle spielt. Man erhält nun die Menge

$$\{(x^j, y^j) \in \mathbb{Z} \times \mathbb{Z} : j \in \mathcal{I}_R ; (x^j)^2 + (y^j)^2 = R\}$$

mit  $N$  Elementen.

Weiter sei durch jedes Tupel  $(x^j, y^j)$ ,  $j \in \mathcal{I}_R$  der Vektor

$$\mathbf{k}'_j := (x^j, y^j)^T \in \mathbb{Z}^2$$

definiert. Die Indizes sind hier hochgestellt, um Verwechslungen zwischen den  $\#\mathcal{I}(R)$  2-Tupeln  $(x_j, y_j)$  und den  $N$  2-Tupeln  $(x^j, y^j)$  zu vermeiden, da mit Letzteren auch die weiteren Möglichkeiten aus (3.7) gemeint sind.

Um die Indexmenge  $\mathcal{I}_R$  zu ordnen, wird jedem Vektor  $\mathbf{k}'_j$ ,  $j \in \mathcal{I}_R$  ein Winkel  $\alpha^j$  zugeordnet. Dies geschieht durch die Funktion  $\deg : \mathbb{S}_R^1 \rightarrow [0, 2\pi)$  mit

$$\deg(\mathbf{k}'_j) = \begin{cases} \arccos \frac{x^j}{\sqrt{R}} & \text{falls } x^j > 0 \text{ und } y^j \geq 0 \\ \arccos \frac{y^j}{\sqrt{R}} + \frac{\pi}{2} & \text{falls } x^j \leq 0 \text{ und } y^j > 0 \\ \arccos \frac{-x^j}{\sqrt{R}} + \pi & \text{falls } x^j < 0 \text{ und } y^j \leq 0 \\ \arccos \frac{-y^j}{\sqrt{R}} + \frac{3}{2}\pi & \text{falls } x^j \geq 0 \text{ und } y^j < 0, \end{cases}$$

d.h.  $\alpha^j = \deg(\mathbf{k}'_j)$  für alle  $j \in \mathcal{I}_R$ . Durch die größenmäßige Ordnung der so berechneten Winkel

$$0 \leq \alpha_1 < \dots < \alpha_N < 2\pi$$

erhält man eine geordnete Indexmenge  $\mathcal{I}$ , die alle  $\mathbf{k}'_j$  in mathematisch positiver Reihenfolge durchläuft, d.h. man erhält die Menge

$$\{\mathbf{k}'_j \subset \mathbb{S}_R^1 \cap \mathbb{Z}^2 : j \in \mathcal{I} ; \deg(\mathbf{k}'_j) < \deg(\mathbf{k}'_l) \text{ für } j < l\} .$$

Auch hier sind die Indizes der Winkel nach unten gesetzt worden, um die geänderte Reihenfolge zu betonen. Die Wellenvektoren sind im Folgenden, d.h. im ganzen weiteren Verlauf der Arbeit, in dieser Weise nummeriert und angeordnet.

Der nächste Schritt dient wieder direkt der Bestimmung geeigneter Wellenvektoren  $\mathbf{k}_j$ , während der letzte Schritt die Menge der  $\mathbf{k}'_j$  ja nur ungeordnet hat. Dazu muss die Zahl

$$G(R) := \max_{n \in \mathbb{N}} \{(\mathbf{k}'_j - \mathbf{k}'_l) \in n\mathbb{Z}^2 : j, l \in \mathcal{I}\}$$

bestimmt werden. Damit können alle  $\mathbf{k}'_j$  mit ihrem Kehrwert skaliert, d.h. die Vektoren

$$\mathbf{k}''_j := \frac{1}{G(R)} \mathbf{k}'_j$$

definiert werden, so dass deren Differenzen immer noch ganzzahlig sind.

Somit ergeben sich für das gewählte  $R$  bei vorgegebener Frequenz  $\omega$  und dem daraus resultierenden  $K = \frac{\omega}{c}$  die Wellenvektoren

$$\mathbf{k}_j := r \mathbf{k}''_j \quad \text{für } j \in \mathcal{I} ,$$

mit  $r := \frac{K G(R)}{\sqrt{R}}$ .

Für die so definierten  $\mathbf{k}_j$  gilt wirklich  $|\mathbf{k}_j| = K$  wie sich leicht nachrechnen lässt:

$$|\mathbf{k}_j| = r |\mathbf{k}''_j| = \frac{K G(R)}{\sqrt{R}} |\mathbf{k}''_j| = \frac{K G(R)}{\sqrt{R}} \frac{1}{G(R)} |\mathbf{k}'_j| = \frac{K}{\sqrt{R}} \sqrt{R} = K .$$

Die resultierenden Differenzen erfüllen somit die Bedingung (3.4). Für das  $r'$  aus dieser Gleichung ergibt sich  $r' = \frac{r}{G(R)}$ , d.h. alle  $\mathbf{k}_j$  erfüllen  $\mathbf{k}_j \in r'\mathbb{Z}^2$  für  $j \in \mathcal{I}$ . Somit kann die Menge

$$\mathcal{K} := \{\mathbf{k}_j \in \mathbb{S}_K^1 : j \in \mathcal{I} := \{1, \dots, N\}\}$$

durch obige Vektoren definiert werden. Ein Strahlensetup, welches durch das beschriebene Verfahren bestimmt wird, eignet sich nach Konstruktion zur Erzeugung von periodischen Mustern. In den entsprechenden Abschnitten werden die jeweiligen Mengen  $\mathcal{K}$  für die Rekonstruktion mit vier, acht und zwölf Strahlen durch diese Methode bestimmt. Die genaue Wahl von  $R$  und alle weiteren Schritte bis hin zur konstruierten Menge  $\mathcal{K}$  werden an der jeweiligen Stelle genauer erläutert.

### 3.2 Symmetrien und lineare Abhängigkeiten

Das bisher vorgestellte mathematische Modell für das Muster

$$f_N(\mathbf{x}) = \frac{1}{2} \left| \sum_{j=1}^N c_j \exp(i\langle \mathbf{k}_j, \mathbf{x} \rangle) \right|^2$$

war für allgemeine Wellenvektoren  $\mathbf{k} \in \mathcal{K}$  und beliebiges  $N \in \mathbb{N}$  hergeleitet worden.

Zur Erinnerung: Die Entwicklungskoeffizienten  $c_j$ ,  $j \in \mathcal{I}$  sind die einzigen Freiheitsgrade der Funktion  $f_N$ .

In dieser Arbeit werden allerdings Strahlensetups untersucht, welche durch die Methode aus Abschnitt 3.1 generiert werden. Die so gewonnene Menge weist Symmetrien auf, welche im mathematischen Modell berücksichtigt werden müssen, was im Folgenden erläutert wird.

Zur Konstruktion von  $\mathcal{K}$  wurde zunächst die Menge

$$\{(x_j, y_j) \in \mathbb{Z} \times \mathbb{Z} : j \in \mathcal{I}(R) ; x_j^2 + y_j^2 = R\}$$

der zulässigen Tupel bestimmt, woraus durch die Tatsache, dass die Tupel aus (3.7) ebenfalls  $x^2 + y^2 = R$  erfüllen, alle 2-Tupel bestimmt werden konnten. Zur Definition der  $\mathbf{k}_j$  wurden die daraus definierten Vektoren  $\mathbf{k}'_j$  nur noch skaliert (und ihre Reihenfolge geändert, was aber hier nicht relevant ist).

Betrachtet man die Menge

$$\mathfrak{M}(x, y) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} y \\ x \end{pmatrix}, \begin{pmatrix} -x \\ -y \end{pmatrix}, \begin{pmatrix} -y \\ -x \end{pmatrix}, \begin{pmatrix} -x \\ y \end{pmatrix}, \begin{pmatrix} x \\ -y \end{pmatrix}, \begin{pmatrix} -y \\ x \end{pmatrix}, \begin{pmatrix} y \\ -x \end{pmatrix} \right\} \subset \mathbb{Z}^2,$$

welche alle gültigen 2-Tupel enthält (bzw. die damit definierten Vektoren) die sich aus dem gültigen Tupel  $(x, y) \in \mathbb{Z} \times \mathbb{Z}$  ergeben, so erkennt man sofort ihre Symmetrien. Die Menge besitzt als Symmetrieachsen die x-Achse, die y-Achse sowie die beiden Winkelhalbierenden. In Abb. 3.2 ist die Menge  $\mathfrak{M}(1, 3)$  dargestellt. Der grau unterlegte Bereich ist die Teilmenge von  $\mathbb{Z} \times \mathbb{Z}$  in der alle Tupel liegen, welche die Bedingung  $x^2 + y^2 = R$  mit  $0 \leq x \leq y$  für ein fest gewähltes  $R$  erfüllen, was in Abb. 3.2 durch den Kreis angedeutet wird. Somit setzt sich  $\mathcal{K}$  aus der Vereinigung solcher skalierten Mengen  $\mathfrak{M}(x, y)$  zusammen:

$$\mathcal{K} = \bigcup_{j \in \mathcal{I}(R)} r' \mathfrak{M}(x_j, y_j).$$

Die Elemente in  $\mathcal{K}$  werden zwar alle mit dem gleichen Faktor  $r'$  skaliert und umgeordnet, was die Symmetrie jedoch nicht beeinflusst. Hieraus folgt, dass  $\mathcal{K}$  die gleichen vier Symmetrieachsen hat und aufgrund dieser speziellen Wahl der  $\mathbf{k}_j$  gilt

$$\mathbf{k}_{j,l} = \mathbf{k}_j - \mathbf{k}_l = \mathbf{k}_{(l+\frac{N}{2}) \pmod N} - \mathbf{k}_{(j+\frac{N}{2}) \pmod N},$$

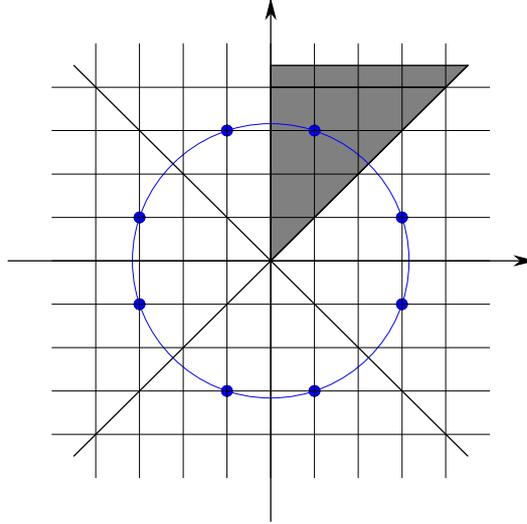


Abbildung 3.1: Die Punkte stellen die Menge  $\mathfrak{M}(1,3)$  dar. Der graue Bereich kennzeichnet die Teilmenge von  $\mathbb{Z} \times \mathbb{Z}$ , für die  $0 \leq x \leq y$  gilt. Zur Bestimmung von  $\mathcal{I}(R)$  werden also alle ganzzahligen Tupel aus dieser Teilmenge bestimmt, welche auf einem Kreis mit dem Radius  $R$  liegen. Die gesamte Menge  $\mathcal{K}$  ergibt sich dann durch Spiegelung an den Symmetrieachsen.

was zu linearen Abhängigkeiten der Summanden führt. Berücksichtigt man diese, so erhält man die Darstellung

$$2f_N(\mathbf{x}) = \sum_{j=1}^N |c_j|^2 + \sum_{j=1}^{\frac{N}{2}} \sum_{l=j+1}^{j+\frac{N}{2}-1} \left( c_{j,l} \exp(i\langle \mathbf{k}_{j,l}, \mathbf{x} \rangle) + \text{k. K.} \right) + \sum_{j=1}^{\frac{N}{2}} \left( \tilde{c}_j \exp(i\langle \mathbf{k}_{j,(j+\frac{N}{2})}, \mathbf{x} \rangle) + \text{k. K.} \right), \quad (3.8)$$

mit

$$\tilde{c}_j := c_j \bar{c}_{j+\frac{N}{2}} \quad \text{und} \quad c_{j,l} := c_j \bar{c}_l + c_{(l+\frac{N}{2}) \pmod{N}} \bar{c}_{(j+\frac{N}{2}) \pmod{N}}. \quad (3.9)$$

Die Rekonstruktion der Koeffizienten  $c_j$  aus (3.8) unter diesen Bedingungen (3.9) führt zu einem inversen Problem, welches einen wesentlichen Teil des Gesamtproblems darstellt.

### Bemerkung 3.1

Für die Indizes  $(j + \frac{N}{2}) \pmod{N}$  der Vektoren  $\mathbf{k}$  und Koeffizienten  $c$  ist zu beachten, dass sich im Falle  $j = \frac{N}{2}$  der Index  $(\frac{N}{2} + \frac{N}{2}) \pmod{N} = 0$  ergibt. Daher sei für die gesamte Arbeit in einer solchen Situation  $\mathbf{k}_0 := \mathbf{k}_N$  und  $c_0 := c_N$  definiert.

Die spezielle Lage  $\mathbf{k}_{j,l} \in r\mathbb{Z}^2$  der Differenzen impliziert für die Funktion  $f_N$  eine Periode  $q := \frac{2\pi}{r}$ , was eine Formulierung dieses inversen Problems über die diskrete Fourieranalyse ermöglicht. Es soll nun der diskrete Fourierraum kurz als Ansatzraum motiviert werden.

### 3.3 Die diskrete Fouriertransformation

Der durch die Problemstellung (3.1) natürlicherweise gegebene Ansatzraum ist der Raum  $\mathcal{S}_N$ , welcher ausschließlich von Exponentialfunktionen aufgespannt wird.

Dieser Raum hat die schöne Eigenschaft, dass sich die kontinuierliche Fouriertransformation von Elementen aus  $\mathcal{S}_N$  sehr einfach berechnen lässt, da sie lediglich eine Linearkombination von Dirac  $\delta$ -Distributionen ist. Damit hat auch die zweidimensionale kontinuierliche Fouriertransformation  $\mathcal{F}[f_N]$  (vgl. (3.20)) des Musters  $f_N \in \mathcal{S}_N$  die einfache Gestalt

$$\begin{aligned} 2\mathcal{F}[f_N](\mathbf{k}) = & \sum_{j=1}^N |c_j|^2 \delta(\mathbf{k}) + \sum_{j=1}^{\frac{N}{2}} \sum_{l=j+1}^{\frac{N}{2}-1} \left( c_{j,l} \delta(\mathbf{k} - \mathbf{k}_{j,l}) + \bar{c}_{j,l} \delta(\mathbf{k} + \mathbf{k}_{j,l}) \right) \\ & + \sum_{j=1}^{\frac{N}{2}} \left( \tilde{c}_j \delta(\mathbf{k} - \mathbf{k}_{j,(j+\frac{N}{2})}) + \bar{\tilde{c}}_j \delta(\mathbf{k} + \mathbf{k}_{j,(j+\frac{N}{2})}) \right). \end{aligned} \quad (3.10)$$

Zur weiteren **numerischen** Behandlung des Problems muss nun die kontinuierliche durch die diskrete Fouriertransformation approximiert werden.

Diese bisher beschriebene Motivation des diskreten Fourierraumes als Ansatzraum wurde auch schon in [5] erläutert, soll an dieser Stelle allerdings noch etwas vertieft werden. Dazu wird verdeutlicht, dass auch die diskrete Fouriertransformation von Funktionen aus  $\mathcal{S}_N$ , speziell von  $f_N \in \mathcal{S}_N$ , Eigenschaften aufweist, welche der Form (3.10) entsprechen.

Zur Vereinfachung betrachten wir dazu zunächst die univariate Situation, welche dann auf den zweidimensionalen Fall übertragen wird. Die eindimensionale **diskrete Fouriertransformation** einer  $q$ -periodischen Funktion  $g : \mathbb{R} \rightarrow \mathbb{C}$  mit Gittergröße  $L := 2^J$  ist gegeben durch

$$\hat{g}(k) := \frac{1}{L} \sum_{j=0}^{L-1} g(x_j) \exp\left(-\frac{2\pi i}{L} k j\right). \quad (3.11)$$

Dabei wird in der Summe die Funktion  $g$  an den  $L$  äquidistanten Stützstellen  $x_l := l\Delta x$  mit der Schrittweite  $\Delta x := \frac{q}{L}$  ausgewertet.

Betrachtet man nun für  $k \in \mathbb{Z}$  eine Funktion der Form

$$g_k(x) := d \exp(ir k x),$$

welche die Struktur eines Summanden aus (3.8) für den eindimensionalen Fall besitzt, da  $rk$  hier die Rolle eines eindimensionalen Wellenvektors aus  $r\mathbb{Z}$  spielt, so erhält man mit  $d \in \mathbb{C}$ ,  $j, l \in [0, L-1] \cap \mathbb{Z}$  und der erkennbaren Periode  $q = \frac{2\pi}{r}$  von  $g_k$  zunächst

$$g_k(x_j) = d \exp\left(\frac{2\pi i}{L} k j\right)$$

und damit für die Fouriertransformierte

$$\hat{g}_k(l) = \frac{1}{L} d \sum_{j=0}^{L-1} \exp\left(\frac{2\pi i}{L} j(k-l)\right) = d \delta_{kl}.$$

Diese Eigenschaft der diskreten Fouriertransformation überträgt sich auf den zweidimensionalen Fall  $g : \mathbb{R}^2 \rightarrow \mathbb{C}$  durch einen einfachen Tensorproduktansatz. Die zweidimensionale Fouriertransformation ist gegeben durch

$$\widehat{g}(\mathbf{k}) := \frac{1}{L^2} \sum_{j,l=0}^{L-1} g(\mathbf{x}_{j,l}) \exp\left(-\frac{2\pi i}{L} \langle \mathbf{k}, (j,l)^T \rangle\right), \quad (3.12)$$

mit  $\mathbf{x}_{j,l} := \Delta x \begin{pmatrix} j \\ l \end{pmatrix}$ . Definiert man nun die Menge der Differenzen aller  $N$  Wellenvektoren

$$\mathcal{K}_d := \{\mathbf{k}_{j,l} = \mathbf{k}_j - \mathbf{k}_l : \mathbf{k}_j, \mathbf{k}_l \in \mathcal{K} ; j, l = 1, \dots, N\}, \quad (3.13)$$

so ergibt sich aufgrund der Linearität der diskreten Fouriertransformation für das Muster (3.8), dass für gegebene Wellenvektoren  $\mathbf{k}_j \in \mathcal{K}$  nach der diskreten Fouriertransformation  $\widehat{f}_N(r^{-1}\mathbf{k}) \neq 0$  nur für die resultierenden Differenzen  $\mathbf{k} \in \mathcal{K}_d$  gilt und speziell an diesen Stellen die Werte

$$\widehat{f}_N(r^{-1}\mathbf{k}_{j,l}) = \begin{cases} \tilde{c}_j & \text{falls } \mathbf{k}_{j,l} \in \mathcal{K}_d \text{ und } l - j = \frac{N}{2} \\ \bar{\tilde{c}}_j & \text{falls } \mathbf{k}_{j,l} \in \mathcal{K}_d \text{ und } j - l = \frac{N}{2} \\ c_{jl} & \text{falls } \mathbf{k}_{j,l} \in \mathcal{K}_d \text{ und } (j,l) \in \mathcal{D} \\ \bar{c}_{jl} & \text{falls } \mathbf{k}_{j,l} \in \mathcal{K}_d \text{ und } (l,j) \in \mathcal{D} \\ \sum_{j=1}^N |c_j|^2 & \text{falls } \mathbf{k}_{j,l} \in \mathcal{K}_d \text{ und } j = l \\ 0 & \text{sonst} \end{cases} \quad (3.14)$$

angenommen werden, wobei

$$\mathcal{D} := \left\{ (j,l) : j = 1, \dots, \frac{N}{2}; l = j + 1, \dots, j + \frac{N}{2} - 1 \right\}$$

die Indexmenge der Doppelsumme aus (3.8) ist. Diese Eigenschaft (3.14) entspricht somit der speziellen Form (3.10) der kontinuierlichen Fouriertransformation von  $f_N$ .

### Bemerkung 3.2

Die diskrete Fouriertransformierte des Musters  $f_N$  ist also nur an den skalierten Stellen  $r^{-1}\mathbf{k}_{j,l} \in r^{-1}\mathcal{K}_d$  von 0 verschieden und liefert dort die Koeffizienten der Summanden in (3.8). Abb. 3.2 zeigt die Menge  $\mathcal{K}_d$ , welche sich aus der Wahl der Menge  $\mathcal{K}$  gemäß (2.2) ergibt. Die Wellenvektoren  $\mathbf{k}_j$  in (2.1) werden in Abschnitt 3.1 gerade so gewählt, dass  $\mathcal{K}_d$  eine Teilmenge von  $r\mathbb{Z}^2$  ist. Diese Wahl macht die in [5] (Abschnitt 4.2) beschriebene approximative Bestimmung der Fourierkoeffizienten auf dem Gitter überflüssig. Die Menge  $\mathcal{K}_d$  ist für festes  $N$  nicht notwendiger Weise eindeutig.

Somit ergibt sich folgende Aufgabe:

### Aufgabe 3.3

Berechne die Koeffizienten  $c_{j,l}$  und  $\tilde{c}_j$  aus (3.8) durch Auswerten der diskreten Fouriertransformierten  $\widehat{f}$  des Musters  $f$  an den Stellen  $r^{-1}\mathbf{k}_{j,l} \in r^{-1}\mathcal{K}_d$ .

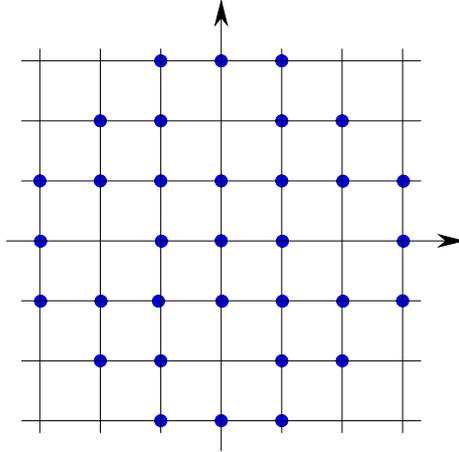


Abbildung 3.2: Die Menge  $\mathcal{K}_d \subset \mathbb{Z}^2$ ;  $\#\mathcal{K}_d = 33$  für  $\mathcal{K}$  gemäß (2.2)

**Definition 3.4**

Für eine Funktion  $g : \mathbb{R}^2 \rightarrow \mathbb{C}$  sei die Seminorm

$$\|g\|_{\ell_2} := \sqrt{\sum_{j,l=0}^{L-1} |g(\mathbf{x}_{j,l})|^2} \quad \text{mit} \quad \mathbf{x}_{j,l} := \begin{pmatrix} j \\ l \end{pmatrix} \in \mathbb{Z}^2$$

durch punktweise Auswertung definiert.

**Bemerkung 3.5**

Für die zweidimensionale diskrete Fouriertransformation (3.12) und die **inverse** diskrete Fouriertransformation  $g : \mathbb{R}^2 \rightarrow \mathbb{C}$ , welche mit  $\mathbf{w}_{j,l} := (j,l)^T \in \mathbb{Z}^2$  durch

$$g(\mathbf{x}) := \sum_{j,l=0}^{L-1} \widehat{g}(\mathbf{w}_{j,l}) \exp\left(\frac{2\pi i}{L} \langle \mathbf{x}, \mathbf{w}_{j,l} \rangle\right) \tag{3.15}$$

definiert ist, gilt die Relation

$$\|g\|_{L_2} := \|g\|_{L_2([0,L]^2)} = L \|\widehat{g}\|_{\ell_2} .$$

Speziell folgt für  $f_N$  gemäß (3.3), dass sich ihre  $L_2$ -Norm einfach durch

$$\|f_N\|_{L_2} = L \sqrt{\sum_{\mathbf{k}_{j,l} \in \mathcal{K}_d} \left| \widehat{f}_N(r^{-1}\mathbf{k}_{j,l}) \right|^2}$$

berechnen lässt.

**Bemerkung 3.6**

Man beachte, dass die diskrete Fouriertransformierte eines beliebigen Musters  $f \notin \mathcal{S}_N$  im Allgemeinen auch auf  $\mathbb{Z}^2 \setminus r^{-1}\mathcal{K}_d$  von 0 verschieden ist. Dies führt bei einer Approximation einer beliebigen

Funktion  $f$  zunächst durch Elemente  $g \in \mathcal{S}_d$  in der Norm  $\|\cdot\|_\Omega := \|\widehat{\cdot}\|_{\ell_2}$  zu einem **unvermeidlichen** Approximationsfehler

$$E_{\text{approx}} := \inf_{g \in \mathcal{S}_d} \left\| \widehat{f} - \widehat{g} \right\|_{\ell_2}, \quad (3.16)$$

mit

$$\mathcal{S}_d := \left\{ g : g(\mathbf{x}) = \sum_{\mathbf{k}_{j,l} \in \mathcal{K}_d} c_{j,l} \exp(i\langle \mathbf{x}, \mathbf{k}_{j,l} \rangle); c_{j,l} \in \mathbb{C} \right\}. \quad (3.17)$$

Hierbei unterstreicht der Index die Abhängigkeit von  $\mathcal{K}_d$ . Die Anzahl der Elemente in  $\mathcal{K}_d$  bestimmt die Anzahl der Freiheitsgrade von Objekten in  $\mathcal{S}_d$ .

### Bemerkung 3.7

Der Raum  $\mathcal{S}_d$  wird von den gleichen Funktionen aufgespannt wie  $\mathcal{S}_N$ . Allerdings ist der bedeutende Unterschied zwischen diesen beiden Räumen der, dass die komplexen Entwicklungskoeffizienten in  $\mathcal{S}_d$  beliebig sind und der Raum somit linear ist, während die in  $\mathcal{S}_N$  den Bedingungen (3.9) unterliegen, was die Linearität des Raumes zerstört. Es ist  $\mathcal{S}_N \subset \mathcal{S}_d$ , denn für jedes  $g \in \mathcal{S}_N$  gilt

$$g(\mathbf{x}) = \sum_{j,l=1}^N \frac{1}{2} c_j \bar{c}_l \exp(i\langle \mathbf{k}_{j,l}, \mathbf{x} \rangle) \in \mathcal{S}_d,$$

da  $\frac{1}{2} c_j \bar{c}_l \in \mathbb{C}$  für  $j, l = 1, \dots, N$ . (vgl. (3.3) (3.13) (3.17))

Eine Approximation aus  $\mathcal{S}_N$  kann also höchstens so gut sein wie die aus  $\mathcal{S}_d$ .

Der Fehler  $E_{\text{approx}}$  ist im Allgemeinen vom Muster  $f$  und der Wahl der  $\mathbf{k}_j$  sowie ihrer Anzahl  $N$  abhängig. Eine geeignete Wahl der Wellenvektoren trägt somit wesentlich zur Qualität der Rekonstruktion bei.

Zusätzlich macht man bei beliebigen Mustern auch durch das Ersetzen der kontinuierlichen durch die diskrete Fouriertransformation einen Approximationsfehler, welcher von der Auflösung der diskreten Fouriertransformation abhängt. Dies wird in Abschnitt 3.4 genauer untersucht.

Für ein beliebiges  $f \notin \mathcal{S}_d$  sei über die diskrete Fouriertransformation von  $f$  die Funktion

$$f_d(\mathbf{x}) := \sum_{\mathbf{k}_{j,l} \in \mathcal{K}_d} \widehat{f}(r^{-1} \mathbf{k}_{j,l}) \exp(i\langle \mathbf{x}, \mathbf{k}_{j,l} \rangle) \in \mathcal{S}_d \quad (3.18)$$

definiert (vgl. Aufgabe 3.3) und man kann leicht nachrechnen, dass

$$\left\| \widehat{f} - \widehat{f}_d \right\|_{\ell_2} = E_{\text{approx}} \quad (3.19)$$

gilt und somit der minimale Fehler erreicht wird. Die beste Approximation  $f_d$  aus  $\mathcal{S}_d$  an  $f$  kann also direkt aus der diskreten Fouriertransformation von  $f$  berechnet werden.

### Bemerkung 3.8

$\mathcal{S}_d$  bzw.  $f_d$  ist nur von  $\mathcal{K}_d$  und somit von der Wahl von  $\mathcal{K}$  abhängig.  $E_{\text{approx}}$  kann also direkt durch die Wahl der verwendeten Strahlen verbessert werden.

### 3.4 Approximationseigenschaften der diskreten Fouriertransformation

Zu einer Funktion  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $n \in \mathbb{N}$  wird in dieser Arbeit die **diskrete** Fouriertransformation mit  $\hat{g}$  bezeichnet. Zusätzlich soll nun mit  $\mathcal{F}[g]$  ein Symbol für die **kontinuierliche** Fouriertransformation der Funktion  $g$  eingeführt werden wie es schon zuvor verwendet wurde und welche für  $\mathbf{x}, \mathbf{k} \in \mathbb{R}^n$ ,  $n \in \mathbb{N}$  durch

$$\mathcal{F}[g](\mathbf{k}) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} g(\mathbf{x}) \exp(-i\langle \mathbf{k}, \mathbf{x} \rangle) d\mathbf{x} \quad (3.20)$$

definiert ist.

Wie in Abschnitt 3.3 und [5] erläutert wird, muss bei numerischen Rechnungen  $\mathcal{F}[g]$  durch  $\hat{g}$  approximiert werden, was auch zur Lösung des hier gegebenen Problems getan wird und zum diskreten Fourierraum als Ansatzraum führt. Die Güte dieser Approximation soll in diesem Abschnitt untersucht werden.

Dazu wird im Folgenden kurz die Interpretation der diskreten Fouriertransformation als (kontinuierliche) Transformierte einer diskreten Approximation von  $g$  erläutert, wie sie in [9] beschrieben wird. Weiter wird in diesem Kontext eine Abschätzung der Approximationsgüte der diskreten Fouriertransformation an die (kontinuierliche) Fouriertransformation gegeben. Diese Betrachtungen werden in [9] für stetige stückweise glatte Funktionen  $g : \mathbb{R} \rightarrow \mathbb{R}$  gemacht und hier auf den für die Problemstellung in dieser Arbeit benötigten Fall von Funktionen  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  ausgeweitet.

Zunächst sollen einige Konventionen und Begriffe festgelegt werden.  $L$  bezeichne in diesem Abschnitt lediglich eine positive ganze Zahl. Man betrachtet nun eine stetige, stückweise glatte Funktion  $g$  über einem reellen Intervall.

Stückweise glatt bedeutet, dass das Gebiet, über dem die Funktion definiert ist, in (disjunkte) Teilgebiete zerlegt werden kann, in denen sie mindestens einmal stetig differenzierbar ist.

Ein reguläres *sampling* der Ordnung  $L$  besteht nun aus zwei geordneten Mengen von  $L$  Zahlen. Die erste ist die Menge der Punkte

$$\{x_0, \dots, x_{L-1}\}$$

mit

$$x_j := j\Delta x \quad \text{für} \quad j = 0, \dots, L-1$$

und der festen Schrittweite  $\Delta x$ . Das kleinste Intervall  $\Omega := [0, q)$ , mit  $q := L\Delta x$ , welches diese Punkte enthält, heißt minimales Fenster. Die zweite geordnete Menge besteht aus den zur Menge  $\{x_0, \dots, x_{L-1}\}$  zugehörigen Funktionswerten

$$\{g_0, \dots, g_{L-1}\}$$

mit

$$g_j := g(x_j) \quad \text{für} \quad j = 0, \dots, L-1.$$

Damit ist zur Funktion  $g : \mathbb{R} \rightarrow \mathbb{R}$  eine diskrete Approximation gegeben durch eine Linearkombination von Dirac-Distributionen

$$\tilde{g}(x) := \sum_{j=0}^{L-1} \tilde{g}_j \delta(x - x_j) \quad (3.21)$$

mit  $\tilde{g}_j := g_j \Delta x$ , wenn  $\tilde{g}$  außerhalb des minimalen Fensters durch 0 fortgesetzt wird oder man wählt eine periodische Fortsetzung mit der Periode  $q$  und erhält

$$\tilde{g}(x) := \sum_{j=-\infty}^{\infty} \tilde{g}_j \delta(x - x_j) \quad (3.22)$$

mit  $\tilde{g}_{j+LZ} := \tilde{g}_j$ .

Eine dieser Approximationen wird verwendet, wenn die Funktion  $g$  nur an den Stellen  $\{x_0, \dots, x_{L-1}\}$  bekannt ist, d.h. die Funktionswerte  $\{g_0, \dots, g_{L-1}\}$  gegeben sind. Diese Situation ist bei numerischen Rechnungen natürlich immer gegeben. Die Wahl der Approximation ist sinnvoll, da die Fouriertransformationen der  $\delta$ -Funktionen einfach zu berechnen sind und das Integral von  $g$  durch das Integral von  $\tilde{g}$  approximiert werden kann, was im Folgenden erläutert wird und eine Fehlerschätzung liefert. Wir betrachten nun also die Fouriertransformierte einer periodischen diskreten Approximation, d.h. die Funktion

$$\mathcal{F}[\tilde{g}](k) = \frac{1}{2\pi} \int_{\mathbb{R}} \tilde{g}(x) \exp(-ikx) dx .$$

Da  $\tilde{g}$  periodisch ist, gilt dies auch für  $\mathcal{F}[\tilde{g}]$  mit der Schrittweite  $\Delta k = \frac{2\pi}{L\Delta x}$  und die Transformierte hat die Form

$$\mathcal{F}[\tilde{g}](k) = \sum_{l=-\infty}^{\infty} \hat{g}(l) \delta(k - l\Delta k)$$

mit  $\hat{g}(l+LZ) = \hat{g}(l)$ . Nach weiterer Rechnung (siehe Anhang) erhält man für die Koeffizienten, dass diese durch die Summen

$$\hat{g}(l) = \frac{1}{L} \sum_{j=0}^{L-1} g_j \exp\left(-\frac{2\pi i}{L} jl\right) \quad (3.23)$$

für  $l = 0, \dots, L-1$  und umgekehrt die ursprünglichen Funktionswerte durch

$$g_j = \sum_{l=0}^{L-1} \hat{g}(l) \exp\left(\frac{2\pi i}{L} lj\right) \quad (3.24)$$

für  $j = 0, \dots, L-1$  berechnet werden. Hier findet sich die Definition der diskreten Fouriertransformation aus Abschnitt 3.3 wieder, was die Wahl der Bezeichnung  $\hat{g}$  rechtfertigt.

Wir betrachten nun ganz allgemein den Fehler, welchen man macht, wenn man anstatt des Integrals

$$\int_{\Omega} h(x) dx$$

die Riemannsche Summe

$$\sum_{j=0}^{L-1} h(x_j) \Delta x$$

berechnet. Dabei sei  $h$  eine reellwertige, stetige und stückweise glatte Funktion über  $\Omega$ . Weiter existiere eine positive Konstante  $M$ , so dass

$$|h'(x)| \leq M$$

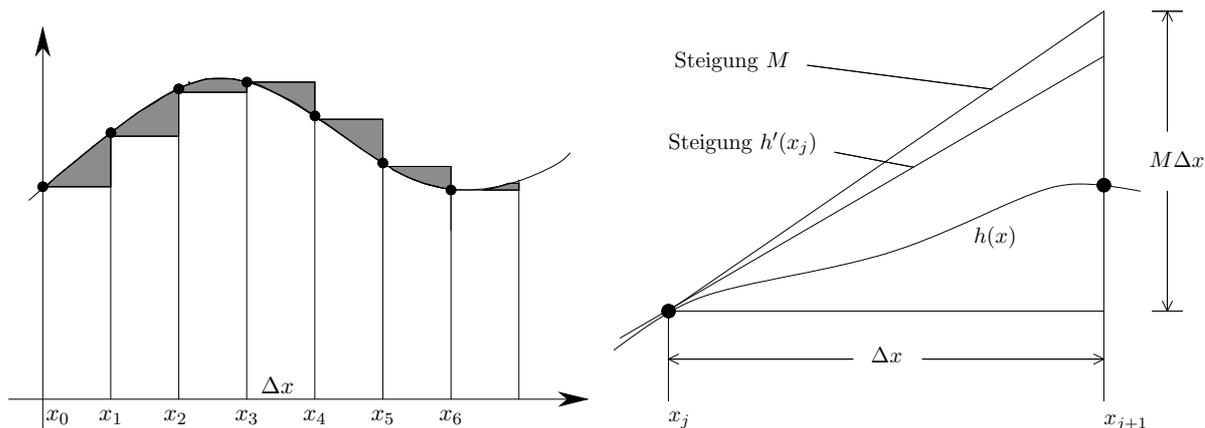


Abbildung 3.3: Links ist der Unterschied zwischen der Fläche unter dem Graphen der Funktion  $h$  und der Riemanschen Summe grau dargestellt. In der rechten Skizze ist ein Teilintervall dargestellt, in dem die Fläche zwischen Graphen und Summe durch ein rechtwinkliges Dreieck mit Katheten der Länge  $\Delta x$  und  $M\Delta x$  überdeckt werden kann. Die Höhe  $M\Delta x$  ergibt sich daraus, dass die Steigung von  $h$  niemals größer als  $M$  wird und  $h$  somit unter der Geraden durch  $h(x_j)$  mit Steigung  $M$  liegt.

für alle  $x$  aus den jeweiligen Teilintervallen von  $\Omega$  ist, in denen die Funktion  $h$  differenzierbar ist. Geometrisch kann man obiges Integral als Fläche unter dem Graphen von  $h$  in  $\Omega$  interpretieren, während die Riemansche Summe als Gesamtfläche von Rechtecken mit Breiten  $\Delta x$  und Höhen  $g_j$  für  $j = 0, \dots, L - 1$  aufgefasst werden kann.

Der Unterschied zwischen diesen beiden Flächen ist der Fehler, welcher beim Ersetzen des Integrals durch die Riemansche Summe begangen wird und auf der linken Seite von Abb. 3.3 grau dargestellt ist. Dieser soll nun abgeschätzt werden. Dazu betrachtet man jedes der  $L$  Teilintervalle  $[x_j, x_{j+1})$  für  $j = 0, \dots, L - 1$  separat und erkennt, dass jedes „Teilgebiet des Fehlers“ durch ein rechtwinkliges Dreieck mit Katheten der Länge  $\Delta x$  und  $M\Delta x$  überdeckt werden kann und somit kleiner oder gleich  $\frac{1}{2}M(\Delta x)^2$  ist. Dies ist auf der rechten Seite von Abb. 3.3 illustriert. Ist  $h := h_1 + ih_2$  nun eine komplexwertige Funktion, deren Ableitung durch  $h' := h'_1 + ih'_2$  definiert ist, so gilt obige Fehlerschranke jeweils für den Real- und den Imaginärteil separat. Berücksichtigt man die Definition des Betrages einer komplexen Zahl  $|a + ib| := \sqrt{a^2 + b^2}$ , so erhält man den zusätzlichen Faktor  $\sqrt{2}$  und insgesamt die Abschätzung

$$\left| \int_{\Omega} h(x) \, dx - \sum_{j=0}^{L-1} h(j\Delta x)\Delta x \right| \leq \frac{1}{2}\sqrt{2}LM(\Delta x)^2. \quad (3.25)$$

Betrachtet man nun wieder die diskrete Approximation

$$\tilde{g}(x) := \sum_{j=-\infty}^{\infty} \tilde{g}_j \delta(x - x_j)$$

mit  $\tilde{g}_j = g(x_j)\Delta x$ , so kann man das Integral über das Produkt  $g\phi$  bzw.  $\tilde{g}\phi$  mit einer beliebigen,

stetigen, stückweise glatten Funktion  $\phi$  berechnen. Für Letzteres ergibt sich dann

$$\begin{aligned}
\int_{\Omega} \tilde{g}(x) \phi(x) dx &= \int_{\Omega} \sum_{j=-\infty}^{\infty} \tilde{g}_j \delta(x - x_j) \phi(x) dx \\
&= \sum_{j=-\infty}^{\infty} g(x_j) \Delta x \int_{\Omega} \delta(x - x_j) \phi(x) dx \\
&= \sum_{j=0}^{L-1} g(x_j) \Delta x \phi(x_j),
\end{aligned} \tag{3.26}$$

da  $\int_{\Omega} \delta(x - x_j) \phi(x) dx = 0$  für  $j < 0$  und  $j \geq L$  ist.

Damit kann obige Abschätzung (3.25), gemeinsam mit der Produktregel für die Ableitung, auf die Funktion  $h = g\phi$  angewendet werden und man erhält

$$\left| \int_{\Omega} g(x) \phi(x) dx - \int_{\Omega} \tilde{g}(x) \phi(x) dx \right| \leq \frac{1}{2} \sqrt{2} L (\Delta x)^2 \left( \|g'\|_{L_{\infty}} \|\phi\|_{L_{\infty}} + \|g\|_{L_{\infty}} \|\phi'\|_{L_{\infty}} \right), \tag{3.27}$$

mit der üblichen Definition der Supremumsnorm

$$\|g\|_{L_{\infty}} := \|g\|_{L_{\infty}(\Omega)} := \operatorname{ess\,sup}_{\Omega} |g|. \tag{3.28}$$

Anschaulich ist die Norm der nur stückweise differenzierbaren Funktion so zu verstehen, dass das Supremum über die „glatten Teilgebiete“ gebildet und anschließend das Maximum über diese Suprema gewählt wird.

Um die gewonnenen Fehlerschätzungen nun auf die Fouriertransformation einer allgemeinen Funktion  $g$  anwenden zu können, muss man berücksichtigen, dass obige Fehlerschätzungen auf der Integration über  $\Omega$  beruhen, während bei der kontinuierlichen Fouriertransformation über ganz  $\mathbb{R}$  integriert wird. Die folgenden Überlegungen zeigen allerdings, dass diese Abschätzungen trotzdem sinnvoll sind.

Für eine  $q$ -periodische Funktion  $g$  und  $z \in \mathbb{Z}$  gilt  $g(x) = g(zq + x)$  und damit

$$\begin{aligned}
\mathcal{F}[g](k) &= \frac{1}{2\pi} \int_{\mathbb{R}} g(x) \exp(-ikx) dx \\
&= \frac{1}{2\pi} \sum_{z \in \mathbb{Z}} \int_{zq}^{zq+q} g(x) \exp(-ikx) dx \\
&= \frac{1}{2\pi} \sum_{z \in \mathbb{Z}} \int_0^q g(zq + y) \exp(-ik(zq + y)) dy \\
&= \frac{1}{2\pi} \sum_{z \in \mathbb{Z}} \int_0^q g(y) \exp(-ikzq) \exp(-iky) dy \\
&= \frac{1}{2\pi} \sum_{z \in \mathbb{Z}} c(k, z) \int_0^q g(y) \exp(-iky) dy
\end{aligned}$$

mit  $c(k, z) := \exp(-ikzq)$ . Obige Fehlerschätzungen gelten somit für jeden Summanden, wobei jeweils der Faktor  $c(k, z)$  zusätzlich berücksichtigt werden muss.

Daher wird im Folgenden immer nur eine Abschätzung über eine Periode  $q$  betrachtet und man erhält als endgültiges Resultat für die Einschränkungen der kontinuierlichen Fouriertransformation

$$\mathcal{F}[g](k)|_{\Omega} := \frac{1}{2\pi} \int_{\Omega} g(x) \exp(-ikx) dx$$

und der diskreten Fouriertransformation

$$\mathcal{F}[\tilde{g}](k)|_{\Omega} := \frac{1}{2\pi} \int_{\Omega} \tilde{g}(x) \exp(-ikx) dx$$

in der Integration von  $\mathbb{R}$  auf  $\Omega$  die Abschätzung mit  $\phi(x) := \frac{1}{2\pi} \exp(-ikx)$ :

### Satz 3.9

Für die Einschränkungen der kontinuierlichen und diskreten Fouriertransformation einer  $q$ -periodischen Funktion  $g$  auf eine Periode  $q = L\Delta x$  gilt an den Stellen  $k \in \mathbb{N} \Delta k \cap \Omega$  mit  $\Delta k = \frac{2\pi}{q}$  die Abschätzung

$$\begin{aligned} |\mathcal{F}[g](k)|_{\Omega} - \mathcal{F}[\tilde{g}](k)|_{\Omega}| &\leq \frac{1}{2} \sqrt{2} \left( \|g'\|_{L_{\infty}} + |k| \|g\|_{L_{\infty}} \right) \frac{L}{2\pi} (\Delta x)^2 \\ &= \frac{1}{2} \sqrt{2} \left( \|g'\|_{L_{\infty}} + |k| \|g\|_{L_{\infty}} \right) \frac{q}{2\pi} \Delta x. \end{aligned} \quad (3.29)$$

Für festes  $k$  und  $\Delta x \rightarrow 0$  geht dieser Fehler von der Ordnung 1 gegen Null, was auch zu erwarten ist. Damit wird die Approximation der diskreten an die kontinuierliche Fouriertransformation bei höherer Auflösung besser.

Dieses Resultat aus [9] soll nun auf den hier benötigten Fall, d.h. für Funktionen  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  ausgeweitet werden.

Dazu beginnt man analog zum eindimensionalen Fall mit einer Menge von Funktionswerten

$$\{g_{\mathbf{z}} : \mathbf{z} \in \mathcal{I}\}$$

an den  $(L_x L_y)$  Stellen

$$\mathcal{I} := \left\{ \begin{pmatrix} j\Delta x \\ l\Delta y \end{pmatrix} \in \mathbb{R}^2 : j = 0, \dots, L_x - 1 ; l = 0, \dots, L_y - 1 \right\}.$$

Dabei sind  $\Delta x$  und  $\Delta y$  die Schrittweiten sowie  $L_x$  und  $L_y$  deren Anzahl in  $x$ - bzw.  $y$ -Richtung. Das Gebiet, welches nun betrachtet wird, ist das Rechteck

$$\Omega := [0, L_x \Delta x) \times [0, L_y \Delta y) \subset \mathbb{R}^2.$$

Die Funktion  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  sei wiederum stetig, stückweise glatt und wir nehmen nun direkt an, dass  $g(\mathbf{x}) = 0$  für  $\mathbf{x} \notin \Omega$  ist.

Zu dieser Funktion erhält man dann eine diskrete Approximation durch die Definition von

$$\tilde{g}(\mathbf{x}) := \sum_{\mathbf{z} \in \mathcal{I}} \tilde{g}_{\mathbf{z}} \delta(\mathbf{x} - \mathbf{z})$$

und den Koeffizienten  $\tilde{g}_{\mathbf{z}} := g_{\mathbf{z}}\Delta x\Delta y$ .

Ruft man sich nun den eindimensionalen Fall ins Gedächtnis zurück, so ergab sich die diskrete Fouriertransformation aus der (kontinuierlichen) Transformation  $\mathcal{F}[\tilde{g}]$  dieser Approximation und man erhält damit die Darstellung

$$\begin{aligned}\mathcal{F}[\tilde{g}](\mathbf{k}) &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \tilde{g}(\mathbf{x}) \exp(-i\langle \mathbf{k}, \mathbf{x} \rangle) \, d\mathbf{x} \\ &= \sum_{\mathbf{z} \in \mathcal{I}} \tilde{g}_{\mathbf{z}} \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \delta(\mathbf{x} - \mathbf{z}) \exp(-i\langle \mathbf{k}, \mathbf{x} \rangle) \, d\mathbf{x} \\ &= \sum_{\mathbf{z} \in \mathcal{I}} g(\mathbf{z})\Delta x\Delta y \frac{1}{(2\pi)^2} \exp(-i\langle \mathbf{k}, \mathbf{z} \rangle) .\end{aligned}$$

Betrachtet man nun wieder das Produkt  $h(\mathbf{x}) := g(\mathbf{x}) \exp(-i\langle \mathbf{k}, \mathbf{x} \rangle)$ , so ergibt sich der Fehler, welchen man durch die diskrete Fouriertransformation im Vergleich zur kontinuierlichen Fouriertransformation begeht als der Fehler, welcher durch Ersetzen des Integrals

$$\int_{\Omega} h(\mathbf{x}) \, d\mathbf{x}$$

durch die Riemannsche Summe

$$\sum_{\mathbf{z} \in \mathcal{I}} h(\mathbf{z})\Delta x\Delta y$$

entsteht. Dazu sei  $h$  zunächst wieder eine reellwertige Funktion. Diese Differenz kann nun, analog zum eindimensionalen Fall, graphisch abgeschätzt werden, da sie als Volumen zwischen dem Graphen von  $g$  und dem Graphen der Treppenfunktion

$$\sum_{\mathbf{z} \in \mathcal{I}} g_{\mathbf{z}}\chi_{\Omega_{\mathbf{z}}}$$

interpretiert werden kann.  $\chi_{\Omega_{\mathbf{z}}}$  ist die charakteristische Funktion über dem Rechteck

$$\Omega_{\mathbf{z}} := \left[ \langle \mathbf{z}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle, \langle \mathbf{z}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle + \Delta x \right) \times \left[ \langle \mathbf{z}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle, \langle \mathbf{z}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle + \Delta y \right) \quad (3.30)$$

mit den Seitenlängen  $\Delta x$  bzw.  $\Delta y$  und dem Punkt  $\mathbf{z} \in \mathbb{R}^2$  als linke untere Ecke. Damit hat das Gebiet die Zerlegung  $\Omega = \bigcup_{\mathbf{z} \in \mathcal{I}} \Omega_{\mathbf{z}}$ .

Betrachtet man nun das „Fehlervolumen“ über jedem Teilgebiet  $\Omega_{\mathbf{z}}$  separat, so lässt sich dieses nach oben abschätzen durch das Volumen des Würfels mit der Grundfläche  $\Omega_{\mathbf{z}}$  und der Höhe

$$\Delta h := \|\nabla h\|_{L^\infty} \sqrt{(\Delta x)^2 + (\Delta y)^2} .$$

Das Supremum in der Definition der Norm

$$\|\nabla h\|_{L^\infty} := \|\nabla h\|_{L^\infty(\Omega)} := \operatorname{ess\,sup}_{\Omega} |\nabla h|$$

ist hier grob gesprochen zunächst über den Teilgebieten zu bilden, in denen  $h$  differenzierbar ist. Anschließend wird das Maximum über die Suprema als Norm des Gradienten von  $h$  über dem ganzen Gebiet definiert.

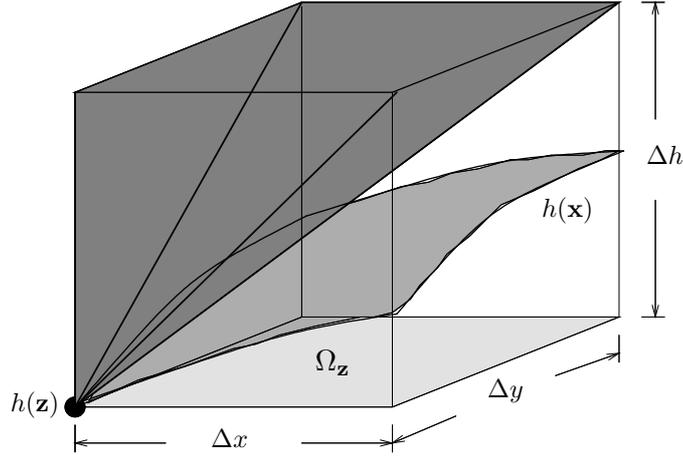


Abbildung 3.4: In dieser Abbildung ist der Würfel skizziert, durch welchen sich das „Fehlvolumen“ über einem Teilgebiet abschätzen lässt. Dieses Teilgebiet  $\Omega_{\mathbf{z}}$  ist hellgrau, der Graph von  $h$  mittelgrau dargestellt. Das dunkelgrau gekennzeichnete Volumen ist das der Pyramide, welches zur Verschärfung der Abschätzung abgezogen werden kann. Die von  $(\mathbf{z}, h(\mathbf{z})) \in \mathbb{R}^3$  ausgehende Diagonale des Würfels hat die Steigung  $\|\nabla h\|_{L_\infty}$ .

Die Größe  $\Delta h$  ergibt sich als maximale Veränderung einer Geraden durch  $(\mathbf{z}, h(\mathbf{z}))^T \in \mathbb{R}^3$  mit der Steigung  $\|\nabla h\|_{L_\infty}$  über dem Gebiet  $\Omega_{\mathbf{z}}$ . Somit liegt der Graph der Funktion  $h$  in  $\Omega_{\mathbf{z}}$  unter dem Kegel, welcher durch Rotation dieser Geraden um die  $z$ -Achse entsteht auf der die Funktionswerte aufgetragen werden. Somit kann zur Verschärfung der Abschätzung des Fehlers das Schnittvolumen des nach oben geöffneten Kegels mit der Spitze  $(\mathbf{z}, h(\mathbf{z}))^T$  und dem Würfel ausgeschnitten werden. Um die Rechnung zu vereinfachen wird stattdessen das Volumen der Pyramide mit Grundfläche  $\Omega_{\mathbf{z}}$  und Höhe  $\Delta h$  abgezogen, was eine unwesentliche Verschlechterung der Abschätzung zur Folge hat. Die hier beschriebenen Flächen und Volumina sind in Abb. 3.4 skizziert.

Damit lässt sich das „Fehlvolumen“  $V$  zwischen den beiden Graphen über dem Teilgebiet  $\Omega_{\mathbf{z}}$  insgesamt durch

$$\begin{aligned}
 V &= V_{\text{Würfel}} - V_{\text{Pyramide}} \\
 &= \frac{2}{3} \Delta x \Delta y \Delta h \\
 &= \frac{2}{3} \|\nabla h\|_{L_\infty} \Delta x \Delta y \sqrt{(\Delta x)^2 + (\Delta y)^2}
 \end{aligned}$$

nach oben abschätzen. Man erhält somit für das gesamte Gebiet  $\Omega$  durch Summation über alle Teilgebiete  $\Omega_{\mathbf{z}}$ ,  $\mathbf{z} \in \mathcal{I}$  die Abschätzung

$$\left| \int_{\Omega} h(\mathbf{x}) \, d\mathbf{x} - \sum_{\mathbf{z} \in \mathcal{I}} h(\mathbf{z}) \Delta x \Delta y \right| \leq \frac{2}{3} L_x L_y \|\nabla h\|_{L_\infty} \Delta x \Delta y \sqrt{(\Delta x)^2 + (\Delta y)^2}. \quad (3.31)$$

Erinnert man sich nun, dass  $h = g\phi$  war und wendet die Produktregel an, so erhält man mit  $\phi(\mathbf{x}) := \frac{1}{(2\pi)^2} \exp(-i\langle \mathbf{k}, \mathbf{x} \rangle)$  auch für den zweidimensionalen Fall eine Fehlerschätzung.

**Satz 3.10**

Für die Einschränkungen der (kontinuierlichen) Fouriertransformation  $\mathcal{F}[g]$  und der diskreten Fouriertransformation  $\mathcal{F}[\tilde{g}]$  auf das Gebiet  $\Omega = [0, L_x \Delta x) \times [0, L_y \Delta y)$  gilt an den Stellen

$$\mathbf{k} \in \left\{ \begin{pmatrix} j \Delta k_x \\ l \Delta k_y \end{pmatrix} \in \mathbb{R}^2 : j = 0, \dots, L_x - 1 ; l = 0, \dots, L_y - 1 \right\} .$$

mit den Schrittweiten  $\Delta k_x := \frac{2\pi}{L_x \Delta x}$  und  $\Delta k_y := \frac{2\pi}{L_y \Delta y}$  die Abschätzung

$$\begin{aligned} |\mathcal{F}[g](\mathbf{k})|_{\Omega} - \mathcal{F}[\tilde{g}](\mathbf{k})|_{\Omega} | &\leq \frac{2}{3} \sqrt{2} \frac{L_x L_y}{2\pi 2\pi} \Delta x \Delta y \left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_2 (\|\nabla g\|_{L_\infty} + |\mathbf{k}| \|g\|_{L_\infty}) \\ &= \frac{2}{3} \sqrt{2} \frac{|\Omega|}{(2\pi)^2} \left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_2 (\|\nabla g\|_{L_\infty} + |\mathbf{k}| \|g\|_{L_\infty}) . \end{aligned} \quad (3.32)$$

Dabei sei mit

$$\left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_2 := \sqrt{(\Delta x)^2 + (\Delta y)^2}$$

die euklidische Norm des Vektors und mit  $|\Omega|$  der Flächeninhalt des Gebietes bezeichnet.

Es wurde auch hier die Fehlerschranke, welche für reellwertige Funktionen  $h$  gefunden wurde, auf Real- und Imaginärteil angewendet, woraus der Faktor  $\sqrt{2}$  resultiert. Auch dieser Fehler geht für festes  $\mathbf{k}$  und  $\Delta x, \Delta y \rightarrow 0$  von der Ordnung 1 gegen Null, was zu erwarten ist.

Alle bisherigen Überlegungen dieses Abschnitts haben die Stetigkeit der Funktion  $g$  vorausgesetzt. In Abschnitt 7 werden allerdings Muster rekonstruiert, welche *nicht* durch eine stetige Funktion zu beschreiben sind. Die dort untersuchten Funktionen  $f : \Omega \rightarrow \{0, 1\}$  nehmen lediglich die Werte Null und Eins an und haben insbesondere die Form

$$f(\mathbf{x}) = \chi_{\Omega_M}(\mathbf{x}) , \quad (3.33)$$

oder

$$f(\mathbf{x}) = 1 - \chi_{\Omega_M}(\mathbf{x}) = \chi_{\Omega \setminus \Omega_M}(\mathbf{x}) ,$$

wenn  $\Omega_M \subset \Omega$  das zu rekonstruierende Muster ist. Man kann sich somit o.B.d.A. auf die erste Darstellung (3.33) der Funktion  $f$  einschränken. Diese ist in  $\Omega$  nicht stetig und die bisher gewonnenen Abschätzungen können hier nicht angewandt werden.

Es sollen daher, analog zur bisherigen Vorgehensweise, auch Fehlerschätzungen für die diskrete Fouriertransformation, speziell dieser charakteristischen Funktionen, hergeleitet werden. Dazu wird nun wieder graphisch abgeschätzt, welchen Fehler man begeht, wenn man mit einer stetigen Funktion  $\phi$  und  $h(\mathbf{x}) := \chi_{\Omega_M}(\mathbf{x})\phi(\mathbf{x})$  das Integral

$$\int_{\Omega} h(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} \chi_{\Omega_M}(\mathbf{x}) \phi(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega_M} \phi(\mathbf{x}) \, d\mathbf{x}$$

durch die Riemannsche Summe

$$\sum_{\mathbf{z} \in \mathcal{I}} h(\mathbf{z}) \Delta x \Delta y = \sum_{\mathbf{z} \in \mathcal{I}} \chi_{\Omega_M}(\mathbf{z}) \phi(\mathbf{z}) \Delta x \Delta y = \sum_{\substack{\mathbf{z} \in \mathcal{I} \\ \mathbf{z} \in \Omega_M}} \phi(\mathbf{z}) \Delta x \Delta y$$

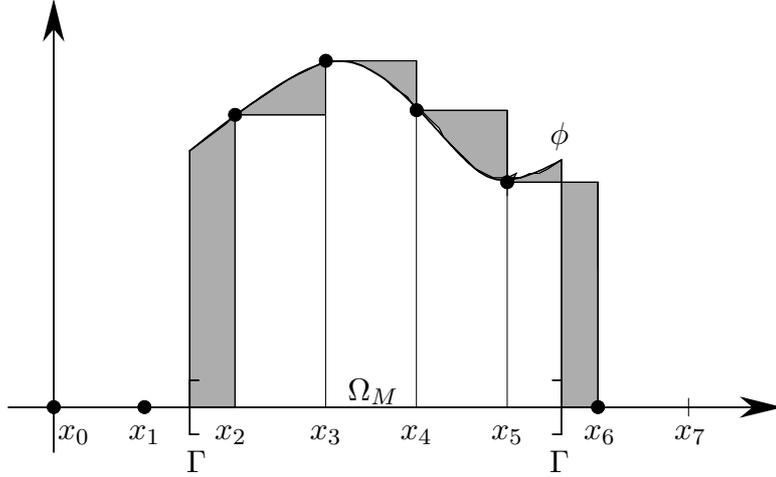


Abbildung 3.5: Die Abbildung zeigt die unstetige Funktion  $h(x) := \chi_{\Omega_M}(x)\phi(x)$  und den Fehler welchen man begeht, wenn das Integral über  $\Omega$  durch eine Riemannsche Summe ersetzt wird. Dieser ist graphisch durch die graue Fläche dargestellt. Die Punkte sind die Stellen an denen  $\phi$  in der Riemannschen Summe ausgewertet wird. Für die inneren Teilgebiete kann obige Abschätzung (3.25) durch Dreiecke verwendet werden. Die Fehleranteile über Teilgebieten, die den Rand  $\Gamma$  des Gebietes  $\Omega_M$  enthalten, können jeweils durch ein Rechteck der Höhe  $\|\phi\|_{L_\infty}$  und Breite  $\Delta x$  überdeckt werden.

ersetzt.

In Abb. 3.5 ist der eindimensionale Fall  $\Omega_M \subset \Omega \subset \mathbb{R}$  skizziert, wobei dann in der Riemannschen Summe  $\Delta y$  und  $L_y$  natürlich durch Eins zu ersetzen sind.

Für diejenigen Teilgebiete  $\Omega_{\mathbf{z}}$ , die vollständig in  $\Omega_M$  enthalten sind, kann man obige Abschätzungen (3.25) bzw. (3.31) für  $\phi$  anwenden. Die Teilgebiete  $\Omega_{\mathbf{z}}$ , welche nichtleeren Schnitt mit dem Rand  $\Gamma$  von  $\Omega_M$  haben, erfordern eine etwas genauere Betrachtung. Der Fehler über diesen Gebieten kann allerdings auch relativ einfach durch  $\|\phi\|_{L_\infty} \Delta x$  im eindimensionalen bzw.  $\|\phi\|_{L_\infty} \Delta x \Delta y$  im zweidimensionalen Fall nach oben abgeschätzt werden. Um den gesamten Fehler behandeln zu können, muss man also diese beiden Fälle unterscheiden und definiert zunächst die Menge aller Eckpunkte der Rechtecke  $\Omega_{\mathbf{z}}$ , welche vollständig in  $\Omega_M$  liegen

$$\mathcal{I}_{\Omega_M} := \{ \mathbf{z} \in \mathcal{I} : \Omega_{\mathbf{z}} \subset \bar{\Omega}_M \}$$

und die Eckpunkte der Teilgebiete, deren Inneres  $\overset{\circ}{\Omega}_{\mathbf{z}}$  nichtleeren Schnitt mit dem Rand  $\Gamma$  haben:

$$\mathcal{I}_\Gamma := \left\{ \mathbf{z} \in \mathcal{I} : \overset{\circ}{\Omega}_{\mathbf{z}} \cap \Gamma \neq \emptyset \right\}.$$

$\mathcal{I}_{\Omega_M}$  und  $\mathcal{I}_\Gamma$  sind disjunkt und die Menge  $\mathcal{I}_K$  enthalte nun alle restlichen Punkte aus  $\mathcal{I}$ , d.h. sie ist die Komplementmenge

$$\mathcal{I}_K := \mathcal{I} \setminus (\mathcal{I}_{\Omega_M} \cup \mathcal{I}_\Gamma).$$

Weiter ergibt sich mit diesen Definitionen, dass die Mengen

$$\bigcup_{\mathbf{z} \in \mathcal{I}_{\Omega_M}} \Omega_{\mathbf{z}} \subset \Omega_M \subset \bigcup_{\mathbf{z} \in (\mathcal{I}_{\Omega_M} \cup \mathcal{I}_\Gamma)} \Omega_{\mathbf{z}}$$

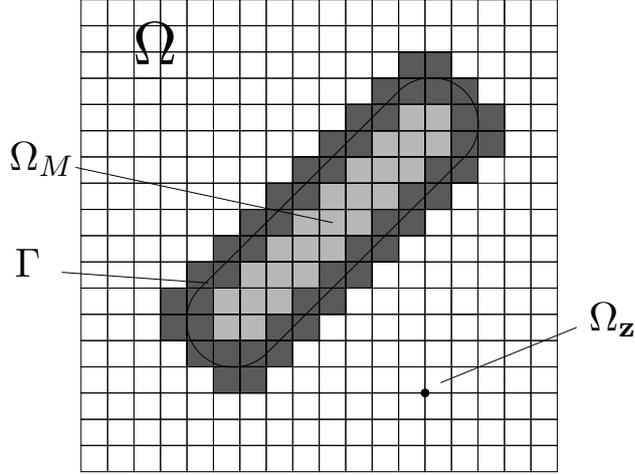


Abbildung 3.6: Hier ist die Unterteilung des Gebietes  $\Omega$  in die Teilgebiete  $\Omega_{\mathbf{z}}$  dargestellt. Exemplarisch wurde ein Gebiet  $\Omega_M$  gewählt, welches dem Muster  $f$  in Abb. 7.2 entspricht. Die dunkle Linie ist der Rand  $\Gamma$  von  $\Omega_M$ . Die hellgrauen Rechtecke sind diejenigen mit den linken unteren Eckpunkten in  $\mathcal{I}_{\Omega_M}$ , die dunkelgrauen in  $\mathcal{I}_{\Gamma}$ . Somit sind die Ecken der weißen Teilgebiete in der Menge  $\mathcal{I}_K$  enthalten.

geschachtelt sind. Bezeichnet man nun die Anzahl der Elemente in den jeweiligen Eckpunkt mengen mit  $L_{\Omega_M} := \#\mathcal{I}_{\Omega_M}$ ,  $L_{\Gamma} := \#\mathcal{I}_{\Gamma}$  und  $L_K := \#\mathcal{I}_K$ , so gilt

$$L_x L_y = L_{\Omega_M} + L_{\Gamma} + L_K . \quad (3.34)$$

Diese Definitionen wurden nur für den zweidimensionalen Fall gemacht, sind aber direkt auf die univariate Situation übertragbar, wenn man anstatt den linken unteren Ecken von Rechtecken die linken Randpunkte  $x_j$  von Intervallen  $[x_j, x_{j+1})$  verwendet, was dann zu

$$L = L_{\Omega_M} + L_{\Gamma} + L_K$$

führt.

Mit diesen Bezeichnungen kann man nun den gesamten Fehler abschätzen, wenn man sich erinnert, dass für den ersten, eindimensionalen Fall  $\phi(x) := \frac{1}{2\pi} \exp(-ikx)$  zu wählen war und  $L_{\Omega_M} \leq L$  ist.

### Korollar 3.11

Für die Einschränkungen der kontinuierlichen und diskreten Fouriertransformation einer  $q$ -periodischen Funktion der speziellen Form  $g = \chi_{\Omega_M}$  auf eine Periode  $q = L\Delta x$  gilt an den Stellen  $k \in \mathbb{N}\Delta k \cap \Omega$  mit  $\Delta k = \frac{2\pi}{q}$  die Abschätzung

$$\begin{aligned} |\mathcal{F}[g](k)|_{\Omega} - \mathcal{F}[\tilde{g}](k)|_{\Omega} | &\leq \frac{1}{2}\sqrt{2}L_{\Omega_M} \|\phi'\|_{L_{\infty}} (\Delta x)^2 + L_{\Gamma} \|\phi\|_{L_{\infty}} \Delta x \\ &= \frac{1}{2}\sqrt{2} \frac{L_{\Omega_M}}{2\pi} |k| (\Delta x)^2 + \frac{L_{\Gamma}}{2\pi} \Delta x \\ &\leq \frac{1}{2}\sqrt{2} \frac{q}{2\pi} |k| \Delta x + \frac{L_{\Gamma}}{2\pi} \Delta x . \end{aligned} \quad (3.35)$$

In zwei Dimensionen ergibt sich mit der Wahl  $\phi(\mathbf{x}) := \frac{1}{(2\pi)^2} \exp(-i\langle \mathbf{k}, \mathbf{x} \rangle)$  und wegen  $L_{\Omega_M} \leq L_x L_y$  das entsprechende Resultat:

**Korollar 3.12**

Für die Einschränkungen der (kontinuierlichen) Fouriertransformation  $\mathcal{F}[g]$  und der diskreten Fouriertransformation  $\mathcal{F}[\tilde{g}]$  auf das Gebiet  $\Omega = [0, L_x \Delta x) \times [0, L_y \Delta y)$  gilt für Funktionen der Gestalt  $g = \chi_{\Omega_M}$  an den Stellen

$$\mathbf{k} \in \left\{ \begin{pmatrix} j \Delta k_x \\ l \Delta k_y \end{pmatrix} \in \mathbb{R}^2 : j = 0, \dots, L_x - 1 ; l = 0, \dots, L_y - 1 \right\} .$$

mit den Schrittweiten  $\Delta k_x := \frac{2\pi}{L_x \Delta x}$  und  $\Delta k_y := \frac{2\pi}{L_y \Delta y}$  die Abschätzung

$$\begin{aligned} | \mathcal{F}[g](\mathbf{k})|_{\Omega} - \mathcal{F}[\tilde{g}](\mathbf{k})|_{\Omega} | &\leq \frac{2}{3} \sqrt{2} L_{\Omega_M} \|\nabla \phi\|_{L_{\infty}} \Delta x \Delta y \left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_2 + L_{\Gamma} \|\phi\|_{L_{\infty}} \Delta x \Delta y \\ &= \frac{2}{3} \sqrt{2} \frac{L_{\Omega_M}}{(2\pi)^2} |\mathbf{k}| \Delta x \Delta y \left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_2 + \frac{L_{\Gamma}}{(2\pi)^2} \Delta x \Delta y \\ &\leq \frac{2}{3} \sqrt{2} \frac{|\Omega|}{(2\pi)^2} |\mathbf{k}| \left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_2 + \frac{L_{\Gamma}}{(2\pi)^2} \Delta x \Delta y . \end{aligned} \quad (3.36)$$

Damit gehen die ersten Summanden in beiden Dimensionen bei festem  $k$  bzw.  $\mathbf{k}$  für wachsende  $L$ ,  $L_x$  und  $L_y$  von der Ordnung 1 gegen Null. Obwohl auch  $L_{\Gamma}$  durch  $L$  bzw.  $(L_x L_y)$  beschränkt ist, wird dies für die zweiten Summanden nicht direkt klar, sondern erfordert eine weitere Argumentation, welche nur für 2D geliefert wird und sich dann auf die eindimensionale Situation übertragen lässt. Dazu seien mit  $|\Omega|$ ,  $|\Omega_M|$  und  $|\Omega \setminus \Omega_M|$  die Flächen der jeweilige Gebiete bezeichnet, d.h. es ist zunächst mit (3.34)

$$\begin{aligned} |\Omega| &= L_x L_y \Delta x \Delta y \\ &= (L_{\Omega_M} + L_{\Gamma} + L_K) \Delta x \Delta y \\ &= L_{\Omega_M} \Delta x \Delta y + L_{\Gamma} \Delta x \Delta y + L_K \Delta x \Delta y . \end{aligned}$$

Betrachtet man nun den Grenzwert  $L_x, L_y \rightarrow \infty$  und berücksichtigt, dass  $|\Omega| = |\Omega_M| + |\Omega \setminus \Omega_M|$  ist, so erkennt man die Tatsache, dass

$$|\Omega| = \underbrace{L_{\Omega_M} \Delta x \Delta y}_{\rightarrow |\Omega_M|} + \underbrace{L_{\Gamma} \Delta x \Delta y}_{\rightarrow 0} + \underbrace{L_K \Delta x \Delta y}_{\rightarrow |\Omega \setminus \Omega_M|} .$$

Somit geht auch der gesamte Fehler in (3.36) für die nichtstetige Funktion  $f(\mathbf{x}) = \chi_{\Omega_M}(\mathbf{x})$  bei festem  $\mathbf{k}$  und wachsenden  $L_x, L_y$  von der Ordnung 1 gegen Null wie man bei höherer Auflösung der diskreten Fouriertransformation erwartet. Das gleiche Resultat gilt auch für den eindimensionalen Fall (3.35).

## 4 Die Rekonstruktion der Koeffizienten

Nach diesen approximationstheoretischen Untersuchungen soll nun das Augenmerk wieder auf die Rekonstruktion der Muster gerichtet werden. Sei im Folgenden  $f$  stets ein beliebiges fest gewähltes Muster. Weiter sei durch  $f_d$  die in (3.18) bestimmte beste Approximation aus  $\mathcal{S}_d$  an dieses  $f$  bezeichnet. Dieser lineare Raum war durch

$$\mathcal{S}_d := \left\{ g : g(\mathbf{x}) = \sum_{\mathbf{k}_{j,l} \in \mathcal{K}_d} c_{j,l} \exp(i\langle \mathbf{x}, \mathbf{k}_{j,l} \rangle); c_{j,l} \in \mathbb{C} \right\}$$

definiert. Außerdem wurde in (3.2) der nichtlineare Teilraum

$$\mathcal{S}_N := \left\{ g = \frac{1}{2} \left| \sum_{j=1}^N c_j \exp(i\langle \mathbf{k}_j, \cdot \rangle) \right|^2 : c_j \in \mathbb{C} \right\} \subset \mathcal{S}_d$$

definiert, welcher alle physikalisch realisierbaren Muster enthält.

In Abschnitt 3.3 wurde gezeigt, dass die diskrete Fouriertransformation einer Funktion  $g \in \mathcal{S}_N$  eine besonders einfache Gestalt hat: Ihre Fourierkoeffizienten an den Stellen  $r^{-1}\mathbf{k}_{j,l} \in r^{-1}\mathcal{K}_d$  berechnen sich direkt aus den Entwicklungskoeffizienten  $c_j \in \mathbb{C}$ ,  $j \in \mathcal{I}$  der Funktion  $g$ , wobei dort die linearen Abhängigkeiten (3.9)

$$\tilde{c}_j := c_j \bar{c}_{j+\frac{N}{2}} \quad \text{und} \quad c_{j,l} := c_j \bar{c}_l + c_{(l+\frac{N}{2}) \pmod{N}} \bar{c}_{(j+\frac{N}{2}) \pmod{N}}$$

berücksichtigt werden müssen. An allen anderen Stellen verschwinden die Fourierkoeffizienten.

Bei der Formulierung des Problems (3.1) wurde bisher noch keine Norm spezifiziert, sondern zunächst einfach ein  $f_* \in \mathcal{S}_N$  gesucht, für das

$$\|g - f\|_{\Omega}$$

minimiert wird. Die „schönen“ Eigenschaften der Funktionen  $g \in \mathcal{S}_N$  bezüglich der diskreten Fouriertransformation können durch eine Wahl der Seminorm

$$\|\cdot\|_{\Omega} := \|\hat{\cdot}\|_{\ell_2}$$

unmittelbar ausgenutzt werden. Zunächst ergibt sich damit das folgende Lemma:

### Lemma 4.1

Sei  $f$  beliebig und fest,  $f_d$  die beste Approximation aus  $\mathcal{S}_d$  an  $f$  wie in (3.18) definiert. Dann gilt für  $g \in \mathcal{S}_N$

$$\|\hat{g} - \hat{f}\|_{\ell_2}^2 = \|\hat{g} - \hat{f}_d\|_{\ell_2}^2 + \|\hat{f}_d - \hat{f}\|_{\ell_2}^2. \quad (4.1)$$

### Beweis

Nach Definition von  $f_d$  gilt für  $\mathbf{k} \in r^{-1}\mathcal{K}_d$ , dass  $\widehat{f}(\mathbf{k}) = \widehat{f}_d(\mathbf{k})$  ist und für  $g \in \mathcal{S}_N$  sowie für  $\mathbf{k} \notin r^{-1}\mathcal{K}_d$ , dass  $\widehat{g}(\mathbf{k}) = \widehat{f}_d(\mathbf{k}) = 0$  ist. Damit ergibt sich, dass

$$\begin{aligned} \left\| \widehat{g} - \widehat{f} \right\|_{\ell_2}^2 &= \sum_{\mathbf{k} \in \Omega \cap \mathbb{Z}^2} \left| \widehat{g}(\mathbf{k}) - \widehat{f}(\mathbf{k}) \right|^2 \\ &= \sum_{\substack{\mathbf{k} \in \Omega \cap \mathbb{Z}^2 \\ \mathbf{k} \in r^{-1}\mathcal{K}_d}} \left| \widehat{g}(\mathbf{k}) - \widehat{f}(\mathbf{k}) \right|^2 + \sum_{\substack{\mathbf{k} \in \Omega \cap \mathbb{Z}^2 \\ \mathbf{k} \notin r^{-1}\mathcal{K}_d}} \left| \widehat{g}(\mathbf{k}) - \widehat{f}(\mathbf{k}) \right|^2 \\ &= \sum_{\substack{\mathbf{k} \in \Omega \cap \mathbb{Z}^2 \\ \mathbf{k} \in r^{-1}\mathcal{K}_d}} \left| \widehat{g}(\mathbf{k}) - \widehat{f}_d(\mathbf{k}) \right|^2 + \sum_{\substack{\mathbf{k} \in \Omega \cap \mathbb{Z}^2 \\ \mathbf{k} \notin r^{-1}\mathcal{K}_d}} \left| \widehat{f}(\mathbf{k}) \right|^2 \\ &= \left\| \widehat{g} - \widehat{f}_d \right\|_{\ell_2}^2 + \sum_{\substack{\mathbf{k} \in \Omega \cap \mathbb{Z}^2 \\ \mathbf{k} \notin r^{-1}\mathcal{K}_d}} \left| \widehat{f}(\mathbf{k}) - \widehat{f}_d(\mathbf{k}) \right|^2 \\ &= \left\| \widehat{g} - \widehat{f}_d \right\|_{\ell_2}^2 + \left\| \widehat{f} - \widehat{f}_d \right\|_{\ell_2}^2, \end{aligned}$$

was die Behauptung beweist. □

Die Lösung  $f_*$  von (3.1) liegt per Definition in  $\mathcal{S}_N$  und Lemma 4.1 gilt somit für  $f_*$ . Möchte man nun  $f$  durch ein  $g \in \mathcal{S}_N$  in der Norm  $\|\cdot\|_{\Omega} := \|\widehat{\cdot}\|_{\ell_2}$  approximieren, so bedeutet obige Gleichung, dass

$$\left\| \widehat{g} - \widehat{f} \right\|_{\ell_2}^2 = \left\| \widehat{g} - \widehat{f}_d \right\|_{\ell_2}^2 + \text{const}(f)$$

ist. Die von  $f$  abhängige Konstante in dieser Gleichung ist das Quadrat des unvermeidlichen Approximationsfehlers  $E_{\text{approx}}$ , welcher durch  $f_d$  berechnet werden kann. Dies führt direkt zu folgendem Problem.

#### Problem 4.2

Finde ein  $f_* \in \mathcal{S}_N$ , welches den Fehler

$$E(g) := \left\| \widehat{f}_d - \widehat{g} \right\|_{\ell_2}, \tag{4.2}$$

zur besten Approximation  $f_d$  aus  $\mathcal{S}_d$  an das gegebene Muster  $f \notin \mathcal{S}_d$  minimiert.

### 4.1 Berücksichtigung der linearen Abhängigkeiten

Abschnitt 3.2 hat gezeigt, dass sich für die Differenzen  $\mathbf{k}_{j,l} \in \mathcal{K}_d$  der Wellenvektoren  $\mathbf{k}_j \in \mathcal{K}$  lineare Abhängigkeiten ergeben, wenn  $\mathcal{K}$  durch die Methode aus Abschnitt 3.1 konstruiert wird. Diese wurden in der Darstellung (3.3) einer Funktion  $g \in \mathcal{S}_N$  berücksichtigt und nun sollen die Entwicklungskoeffizienten  $c_j$ ,  $j \in I$  der Funktion  $f_*$  bestimmt werden, für die das Funktional  $E$  minimiert wird.

Dazu kann die Darstellung des Fehlerfunktionals  $E$  aus (4.2) in eine geeignetere Form umgeschrieben werden. Sei dazu  $g \in \mathcal{S}_N$  mit der Darstellung (3.3)

$$g(\mathbf{x}) = \sum_{j=1}^N |c_j|^2 + \sum_{\substack{j,l=1 \\ j < l}}^N \left( c_j \bar{c}_l \exp(i\langle \mathbf{k}_{j,l}, \mathbf{x} \rangle) + \text{k. K.} \right).$$

Für die Fouriertransformierte einer Funktion  $g : \mathbb{R}^2 \rightarrow \mathbb{C}$  folgt, dass  $\widehat{g}(-\mathbf{k}) = \overline{\widehat{g}(\mathbf{k})}$  ist und somit ergibt sich wegen (3.14) zusammen mit der besten Approximation (3.18) aus  $\mathcal{S}_d$  für  $E$

$$\begin{aligned} \frac{1}{2} E(g)^2 &= \sum_{j=1}^{\frac{N}{2}} \sum_{l=j+1}^{j+\frac{N}{2}-1} \left| c_j \bar{c}_l + c_{(l+\frac{N}{2}) \pmod{N}} \bar{c}_{(j+\frac{N}{2}) \pmod{N}} - \widehat{f}(r^{-1}\mathbf{k}_{j,l}) \right|^2 \\ &+ \sum_{j=1}^{\frac{N}{2}} \left| c_j \bar{c}_{j+\frac{N}{2}} - \widehat{f}(r^{-1}\mathbf{k}_{j,j+\frac{N}{2}}) \right|^2 + \left| \sum_{j=1}^N |c_j|^2 - \widehat{f}(0) \right|^2. \end{aligned} \quad (4.3)$$

**Bemerkung 4.3**

Der letzte Summand in (4.3) kontrolliert lediglich den Mittelwert des Musters. Da das Hauptaugenmerk allerdings auf der Rekonstruktion der Struktur des Musters  $f$  liegt, ist es aus physikalischen Gründen legitim, zur analytischen Vereinfachung des Problems diesen Summanden zu vernachlässigen und das Fehlerfunktional

$$\begin{aligned} \mathcal{J}(g) &:= \sum_{j=1}^{\frac{N}{2}} \sum_{l=j+1}^{j+\frac{N}{2}-1} \left| c_j \bar{c}_l + c_{(l+\frac{N}{2}) \pmod{N}} \bar{c}_{(j+\frac{N}{2}) \pmod{N}} - \widehat{f}(r^{-1}\mathbf{k}_{j,l}) \right|^2 \\ &+ \sum_{j=1}^{\frac{N}{2}} \left| c_j \bar{c}_{j+\frac{N}{2}} - \widehat{f}(r^{-1}\mathbf{k}_{j,j+\frac{N}{2}}) \right|^2 \end{aligned} \quad (4.4)$$

zu definieren. Die numerischen Resultate in Abschnitt 7 werden zeigen, dass diese Konzentration auf die Rekonstruktion der Struktur sogar von Vorteil ist.

Mit einem Operator  $N : \mathcal{S}_d \rightarrow \mathcal{S}_d$

$$(Ng)(\mathbf{x}) := \begin{cases} g(\mathbf{x}) & \text{falls } \mathbf{x} \neq 0 \\ 0 & \text{falls } \mathbf{x} = 0 \end{cases}, \quad (4.5)$$

gilt der einfache Zusammenhang

$$2 \mathcal{J}(g) = \left\| N(\widehat{f}_d) - N(\widehat{g}) \right\|_{\ell_2}^2.$$

**Bemerkung 4.4**

Das Fehlerfunktional  $\mathcal{J}$  kann als Funktion  $\mathcal{J} : \mathbb{C}^N \rightarrow \mathbb{R}$  betrachtet werden, da die komplexen Koeffizienten die Freiheitsgrade einer Funktion  $g \in \mathcal{S}_N$  darstellen.

Aufgrund dieser Tatsachen kann Problem 4.2 durch das folgende Problem ersetzt werden.

**Problem 4.5**

Finde ein  $f_* \in \mathcal{S}_N$ , welches das Funktional  $\mathcal{J}$ , definiert in (4.4), minimiert. Dies ist äquivalent dazu, dass zu gegebenen  $\hat{f}(\mathbf{k}) \in \mathbb{C}$ , mit  $\mathbf{k} \in r^{-1}\mathcal{K}_d$ , für alle  $j = 1, \dots, N$  die Koeffizienten  $c_j \in \mathbb{C}$  zu finden sind, so dass  $\mathcal{J}$  minimiert wird.

**Bemerkung 4.6**

Das Minimum  $f_*$  ist nicht eindeutig, da eine Phasenverschiebung aller  $c_j$  um ein beliebiges  $\varphi \in \mathbb{R}$  den Wert von  $\mathcal{J}$  unverändert lässt.

**Bemerkung 4.7**

Im Allgemeinen ist **nicht** zu erwarten, dass

$$\inf_{g \in \mathcal{S}_N} \left\| N(\hat{f}_d) - N(\hat{g}) \right\|_{\ell_2} = 0 \quad (4.6)$$

erreicht wird.

**4.2 Minimierungsverfahren**

Das Minimum des Funktionals  $\mathcal{J}$  aus (4.4) soll nun numerisch bestimmt werden. Dabei wird  $\mathcal{J}$  nach Bemerkung 4.4 als Funktion  $\mathcal{J} : \mathbb{C}^N \rightarrow \mathbb{R}$  behandelt.

**4.2.1 Koordinatenabstiegsverfahren**

Zur Lösung von Problem 4.5 wurde in [5] ein Minimierungsverfahren vorgestellt, welches sich als effizient herausgestellt hat. Dabei wird nacheinander das Minimum von  $\mathcal{J}$  bezüglich jeder der  $N$  Koordinatenrichtungen analytisch berechnet und die jeweilige Koordinate direkt durch das berechnete Minimum ersetzt. Sind die Minima in allen Richtungen berechnet, wiederholt man diesen Schritt, bis ein Abbruchkriterium erfüllt ist. Diese Vorgehensweise kann durch eine Folge komplexer Vektoren in  $\mathbb{C}^N$  beschrieben werden.

Diese Folge  $\{\mathbf{c}^k\} \subset \mathbb{C}^N$  mit  $\mathbf{c} := (c_1, \dots, c_N)^T$  ist durch

$$c_j^{k+1} := \arg \min_{\xi \in \mathbb{C}} \mathcal{J}(c_1^{k+1}, \dots, c_{j-1}^{k+1}, \xi, c_{j+1}^k, \dots, c_N^k) \quad ; \quad j = 1, \dots, N \quad (4.7)$$

definiert, wofür in [4] folgender Satz bewiesen wird:

**Satz 4.8**

Sei  $g : X \rightarrow \mathbb{R}$  eine stetig differenzierbare Funktion auf  $X = X_1 \times X_2 \times \dots \times X_m$ , wobei  $X_j$  für alle  $j = 1, \dots, m$  eine abgeschlossene Teilmenge von  $\mathbb{R}^{n_j}$  ist. Weiter sei für alle  $j = 1, \dots, m$  und  $x \in X$  das Minimum

$$\min_{\xi \in X_j} g(x_1, \dots, x_{j-1}, \xi, x_{j+1}, \dots, x_m)$$

eindeutig erreicht und  $\{x^k\}$  die Folge, welche durch

$$x_j^{k+1} := \arg \min_{\xi \in X_j} g(x_1^{k+1}, \dots, x_{j-1}^{k+1}, \xi, x_{j+1}^k, \dots, x_m^k) \quad ; \quad j = 1, \dots, m \quad (4.8)$$

definiert ist.

Dann ist jeder Häufungspunkt von  $\{x^k\}$  ein stationärer Punkt von  $g$ .

Das durch die Folge  $\{x^k\}$  definierte Verfahren heißt **Koordinatenabstiegsverfahren**.

Dieser Satz kann nun auf obige Folge  $\{\mathbf{c}^k\}$  angewandt werden und liefert eine Konvergenzaussage für das Verfahren aus [5].

**Lemma 4.9**

Die Folge  $\{\mathbf{c}^k\}$ , welche durch (4.7) definiert ist, konvergiert gegen einen stationären Punkt des Fehlerfunktionals  $\mathcal{J}$ , welches in (4.4) definiert wurde (vgl. Bemerkung 4.4).

**Beweis**

$\mathbb{C} = \mathbb{R}^2$  ist abgeschlossen und  $\mathcal{J} : \mathbb{C}^N \rightarrow \mathbb{R}$  ist reell stetig differenzierbar. Zu zeigen ist nun, dass das Minimum von  $\mathcal{J}$  in einer komplexen Koordinatenrichtung eindeutig erreicht wird, da dann obiger Satz 4.8 angewendet werden kann. Das Minimum (4.7) in einer Koordinatenrichtung  $j$  kann sogar analytisch berechnet werden. Dazu definiere zunächst das Funktional  $\tilde{\mathcal{J}}_j(g)$  für  $j = 1, \dots, \frac{N}{2}$  durch

$$\begin{aligned} \tilde{\mathcal{J}}_j(g) := & \sum_{\substack{l=j+1 \\ l-j \neq \frac{N}{2}}}^N \left| c_j \bar{c}_l + c_{(l+\frac{N}{2}) \pmod{N}} \bar{c}_{(j+\frac{N}{2}) \pmod{N}} - \hat{f}(r^{-1} \mathbf{k}_{j,l}) \right|^2 \\ & + \sum_{l=1}^{j-1} \left| c_l \bar{c}_j + c_{(j+\frac{N}{2}) \pmod{N}} \bar{c}_{(l+\frac{N}{2}) \pmod{N}} - \hat{f}(r^{-1} \mathbf{k}_{l,j}) \right|^2 \end{aligned}$$

und

$$\begin{aligned} \tilde{\mathcal{J}}_j(g) := & \sum_{l=\frac{N}{2}+1}^{j-1} \left| c_j \bar{c}_l + c_{(l+\frac{N}{2}) \pmod{N}} \bar{c}_{(j+\frac{N}{2}) \pmod{N}} - \hat{f}(r^{-1} \mathbf{k}_{j,l}) \right|^2 \\ & + \sum_{\substack{l=1 \\ j-l \neq \frac{N}{2}}}^{\frac{N}{2}} \left| c_l \bar{c}_j + c_{(j+\frac{N}{2}) \pmod{N}} \bar{c}_{(l+\frac{N}{2}) \pmod{N}} - \hat{f}(r^{-1} \mathbf{k}_{l,j}) \right|^2 \\ & + \sum_{l=j+1}^N \left| c_l \bar{c}_j + c_{(j+\frac{N}{2}) \pmod{N}} \bar{c}_{(l+\frac{N}{2}) \pmod{N}} - \hat{f}(r^{-1} \mathbf{k}_{l,j}) \right|^2 \end{aligned}$$

für  $j = \frac{N}{2} + 1, \dots, N$  und damit

$$\mathcal{J}_j(g) := \tilde{\mathcal{J}}_j(g) + \left| c_j \bar{c}_{(j+\frac{N}{2}) \pmod{N}} - \hat{f}(r^{-1} \mathbf{k}_{j,(j+\frac{N}{2}) \pmod{N}}) \right|^2. \quad (4.9)$$

Es gilt für das Fehlerfunktional (4.4) der Zusammenhang

$$4 \mathcal{J}(g) = \sum_{l=1}^N \tilde{\mathcal{J}}_l(g) + 2 \sum_{l=1}^N \left| c_l \bar{c}_{(l+\frac{N}{2}) \pmod{N}} - \hat{f}(r^{-1} \mathbf{k}_{l,(l+\frac{N}{2}) \pmod{N}}) \right|^2.$$

Um  $\mathcal{J}$  gemäß (4.7) in einer Koordinatenrichtung  $j$  zu minimieren, genügt es,  $\mathcal{J}_j$  bezüglich  $c_j$  zu minimieren, da  $\mathcal{J}_j$  alle Summanden von  $\mathcal{J}$  zusammenfasst, welche  $c_j$  enthalten. Das absolute Minimum von  $\mathcal{J}_j$  und somit von  $\mathcal{J}$  in dieser Richtung kann analytisch berechnet werden. Dazu identifiziert man  $\mathbb{C}$  mit der reellen Ebene  $\mathbb{R}^2$  mittels  $c_j = x + iy \mapsto (x, y)$  und betrachtet  $\mathcal{J}_j$  als Funktion von  $(x, y)$ . Berechnet man nun den Gradienten  $\nabla \mathcal{J}_j = (\partial_x \mathcal{J}_j, \partial_y \mathcal{J}_j)^T$  des Funktionals und identifiziert umgekehrt  $\mathbb{R}^2$  mit  $\mathbb{C}$  mittels  $(x, y) \mapsto x + iy$ , so erkennt man, dass die notwendige Bedingung  $\partial_x \mathcal{J}_j + i \partial_y \mathcal{J}_j = 0$  für Extrema bei

$$c_j^{min} := \frac{\sum_{\substack{l=1 \\ l-j \neq \frac{N}{2} \\ l \neq j}}^N \left( \widehat{f}(r^{-1} \mathbf{k}_{j,l}) - c_{(l+\frac{N}{2}) \pmod{N}} \bar{c}_{(j+\frac{N}{2}) \pmod{N}} \right) c_l + \widehat{f}(r^{-1} \mathbf{k}_{j,(j+\frac{N}{2}) \pmod{N}}) c_{(j+\frac{N}{2}) \pmod{N}}}{\sum_{\substack{l=1 \\ l \neq j}}^N |c_l|^2} \quad (4.10)$$

eindeutig erfüllt ist. Weiterhin ist die Hesse-Matrix von  $\mathcal{J}_j(x, y)$  positiv definit, was die Behauptung beweist. □

Es sei an dieser Stelle noch bemerkt, dass das Funktional  $\mathcal{J}_j : \mathbb{C} \rightarrow \mathbb{R}$  zwar reell stetig differenzierbar, aber nicht komplex differenzierbar ist, da die *Cauchy-Riemann'schen Differentialgleichungen* nicht erfüllt sind.

Damit kann das Koordinatenabstiegsverfahren (4.7) gemeinsam mit einem Abbruchkriterium als Algorithmus formuliert werden:

**Algorithmus 4.10**

Für einen Startwert  $\mathbf{c} \in \mathbb{C}^N$  und einen geeigneten Abbruchparameter  $\mu > 0$  sowie eine maximale Iterationszahl  $N_0 \in \mathbb{N}$  berechne:

```

m = 0
 $\mathcal{J}^m(g) := \mathcal{J}(c_1, \dots, c_N)$ 
do
  for (l = 1, \dots, N)
     $c_l \leftarrow c_l^{min}$ 
   $\mathcal{J}^{m+1}(g) := \mathcal{J}(c_1, \dots, c_N)$ 
  m ← m + 1
while ( | $\mathcal{J}^m(g) - \mathcal{J}^{m-1}(g)$ | >  $\mu$  && m <  $N_0$  )

```

Das Verfahren bricht also ab, sobald im Vergleich zum vorherigen Iterationsschritt keine wesentliche Verbesserung des Kostenfunktionals erreicht werden kann oder eine maximale Iterationszahl überschritten wird.

### Bemerkung 4.11

Für Algorithmus 4.10 ist allerdings nur die Konvergenz gegen einen stationären Punkt von  $\mathcal{J}$  garantiert. Über die Abhängigkeit von Startwerten, die Konvergenzrate und das absolute Minimum kann a priori keine Aussage getroffen werden. Diese Punkte werden in Abschnitt 7 genauer untersucht.

### 4.2.2 Die Nelder und Mead Simplexmethode

In obigem Koordinatenabstiegsverfahren wurde entlang einer festen Menge von Koordinatenrichtungen minimiert, wobei deren lineare Unabhängigkeit eine Verbesserung des Kostenfunctionals  $\mathcal{J}$  garantierte. Diese Idee kann durch Verwendung einer anderen Menge von Richtungen generalisiert werden. Es existieren verschiedene Verfahren dieses Typs der direkten Suchmethoden, deren Grundprinzip allerdings meist nur auf heuristischen Argumenten basiert. Da diese Methoden aber im Allgemeinen sehr leicht zu implementieren sind und keine Berechnung von Gradienten erfordern, soll in diesem Kontext exemplarisch das Nelder und Mead Simplexverfahren wie es in [4] beschrieben ist, vorgestellt und angewendet werden. Dieses Verfahren sollte nicht mit dem Simplexverfahren zur linearen Optimierung verwechselt werden.

Es soll eine Funktion  $g : X_1 \times X_2 \times \dots \times X_m =: X \rightarrow \mathbb{R}$  minimiert werden. Der Algorithmus startet mit einem *Simplex*, d.h. mit der konvexen Hülle von  $n + 1$  Punkten  $x^0, x^1, \dots, x^n \in X$  und generiert in jedem Iterationsschritt ein neues Simplex. Im Folgenden wird die Erzeugung des neuen Simplex, d.h. im Speziellen einer neuen Simplexecke beschrieben. Dazu sei mit  $x_{min}$  und  $x_{max}$  die „beste“ und „schlechteste“ Ecke bezeichnet, d.h. es ist

$$x_{min} := \arg \min_{j=0, \dots, n} g(x^j)$$
$$x_{max} := \arg \max_{j=0, \dots, n} g(x^j).$$

Weiter bezeichne  $\hat{x}$  den Mittelpunkt des Simplex, welches ohne  $x_{max}$  gebildet und somit durch

$$\hat{x} := \frac{1}{n} \left( -x_{max} + \sum_{j=0}^n x^j \right)$$

definiert wird. Die Iteration ersetzt nun im Simplex die „schlechteste“ Ecke  $x_{max}$  durch einen „besseren“ Wert  $x_{neu}$ . Im ersten Schritt zur Berechnung der neuen Ecke wird der *Reflektionspunkt*

$$x_{ref} := \hat{x} + \beta(\hat{x} - x_{max})$$

für einen Parameter  $\beta > 0$  berechnet. Der Punkt  $x_{ref}$  liegt also auf der Geraden, welche durch  $x_{max}$  und  $\hat{x}$  bestimmt ist. Der nächste Schritt besteht aus der *Kostenberechnung* des Reflektionspunktes. Dabei werden drei Fälle unterschieden:

- (1)  $x_{ref}$  hat **minimale Kosten**, wenn  $g(x_{min}) > g(x_{ref})$
- (2)  $x_{ref}$  hat **mittlere Kosten**, wenn  $\max\{g(x^i) : x^i \neq x_{max}\} > g(x_{ref}) \geq g(x_{min})$
- (3)  $x_{ref}$  hat **maximale Kosten**, wenn  $g(x_{ref}) \geq \max\{g(x^i) : x^i \neq x_{max}\}$ .

Die Aufdatierung der Ecke  $x_{neu}$  erfolgt nun in Abhängigkeit dieser Kategorisierung.

(I) (versuche Expansion)

Hat  $x_{ref}$  minimale Kosten, so berechne für einen Parameter  $\gamma > 0$

$$x_{exp} = x_{ref} + \gamma(x_{ref} - \hat{x})$$

und definiere

$$x_{neu} = \begin{cases} x_{exp} & \text{wenn } g(x_{exp}) < g(x_{ref}) \\ x_{ref} & \text{sonst.} \end{cases}$$

(II) (Reflektion)

Hat  $x_{ref}$  mittlere Kosten, so definiere  $x_{neu} = x_{ref}$ .

(III) (Kontraktion)

Hat  $x_{ref}$  maximale Kosten, so berechne für einen Parameter  $\theta \in (0, 1)$

$$x_{neu} = \begin{cases} \theta x_{max} + (1 - \theta)\hat{x} & \text{wenn } g(x_{max}) \leq g(x_{ref}) \\ \theta x_{ref} + (1 - \theta)\hat{x} & \text{sonst.} \end{cases}$$

Dieser Iterationschritt stellt den Kern des Nelder und Mead Simplexverfahrens dar. Damit soll nun das Funktional (4.4) minimiert werden, wobei  $\mathcal{J}$  wieder als Funktion  $\mathcal{J} : \mathbb{C}^N \rightarrow \mathbb{R}$  betrachtet wird. Für  $\beta = \gamma = 1$  und  $\theta = \frac{1}{2}$  erhält man die von Nelder und Mead ursprünglich entwickelte Methode, welche dem Verfahren ihren Namen gibt.

#### Algorithmus 4.12

Definiere ein komplexes Startsimplex  $\Delta_0 \subset \mathbb{C}^N$  als konvexe Hülle von  $N + 1$  Punkten  $\mathbf{c}^j \in \mathbb{C}^N$ ,  $j = 0, \dots, N$  und wähle die Parameter  $\beta, \gamma > 0$  und  $\theta \in (0, 1)$  sowie die Abbruchkriterien  $\mu > 0$  und  $N_0 \in \mathbb{N}$ .

$m = 0$

**do**

$$\mathbf{c}_{max} \leftarrow \mathbf{c}_{neu} \quad \rightsquigarrow \quad \Delta_{m+1}$$

$$m \leftarrow m + 1$$

**while** (  $\max\{|\mathbf{c}^j - \mathbf{c}^k| : j, k = 0, \dots, N\} > \mu \quad \&\& \quad m < N_0$  )

$\mathbf{c} \leftarrow \mathbf{c}_{min}$

Das Verfahren bricht also ab, wenn der Durchmesser des Simplex  $\Delta_m$ , d.h. der maximale Abstand zweier Ecken, kleiner als  $\mu$  wird oder eine maximale Iterationszahl  $N_0$  überschritten wird.

#### Bemerkung 4.13

Die Wahl des Startsimplex für Algorithmus 4.12 ist nicht beliebig. Um über den ganzen Raum  $\mathbb{C}^N$  zu minimieren, muss das Simplex ebenfalls die Dimension  $N$  haben. Daher werden mindestens  $N + 1$  Ecken benötigt, deren Wahl somit auch durch die Dimension des Simplex beeinflusst ist. In [11] wird ein Simplex  $\Delta_m$ , welches diese Eigenschaft besitzt, als **nicht degeneriert** bezeichnet. Ist das  $N$ -dimensionale Volumen von Null verschieden, d.h.  $\text{vol}(\Delta_m) > 0$ , so ist  $\Delta_m$  nicht degeneriert. In dieser Arbeit wird ebenfalls gezeigt, dass alle „nachfolgenden“ Simplizes ebenfalls nicht degeneriert sind. Das Nelder und Mead Simplexverfahren minimiert also über ganz  $\mathbb{C}^N$ . Das Resultat  $\mathbf{c}$  hängt ebenfalls von der Wahl der Parameter  $\beta, \gamma$  und  $\theta$  ab. Dies wird in Abschnitt 7 genauer untersucht.

Des Weiteren wird in [11] die Konvergenz für strikt konvexe Funktionen in 1D bewiesen und es werden verschiedene limitierte Konvergenzaussagen für strikt konvexe Funktionen in 2D gemacht. In [13] wird aber auch eine Familie von strikt konvexen Funktionen und eine Startkonfiguration vorgestellt, für welche das Verfahren gegen einen instationären Punkt konvergiert. Es sind keine allgemeinen Konvergenzaussagen bekannt. Allerdings hat sich das Verfahren in der Praxis meist als effizient erwiesen, dabei für „niedere“ Dimensionen mehr als für höherdimensionale Probleme ([11]). Die numerischen Resultate in Abschnitt 7 sind für  $N = 8$  noch zufriedenstellend, werden aber schon bei zwölf komplexen Freiheitsgraden unzuverlässig.

Hier kann allerdings auch eine elementare Eigenschaft des Simplexverfahren genutzt werden. In der Definition (4.4) von  $\mathcal{J}$  wurde der letzte Summand aus (4.3) vernachlässigt (vgl. Bemerkung 4.3). Diese Veränderung wurde zur analytischen Vereinfachung des Problems gemacht, d.h. im Speziellen, um Ableitungen und somit die Minima (4.10) in den einzelnen Koordinatenrichtungen explizit berechnen zu können. Diese Berechnung ist für die Simplexmethode nicht mehr notwendig und eine Minimierung von

$$\mathcal{J}(\mathbf{c}) + \left| \sum_{j=1}^N |c_j|^2 - \widehat{f}(0) \right|^2 \quad (4.11)$$

birgt somit keinen Verlust von Effizienz. Dies führt an Stelle von Problem 4.5 zu einem alternativen Problem, welches auch den Mittelwert berücksichtigt. Der rechte Term in (4.11) wird im weiteren Verlauf der Arbeit auch als *Nebenbedingung* bezeichnet, was in der Praxis der Kontrolle des mittleren Grauwertes eines Bildes entspricht.

#### Problem 4.14

Finde ein  $f_* \in \mathcal{S}_N$  mit Koeffizientenvektor  $\mathbf{c}^* \in \mathbb{C}^N$  welches (4.11) minimiert (vgl. Bemerkung 4.4).

#### 4.2.3 Die Nelder und Mead Simplexmethode auf $\mathbb{S}_r^{2N-1}$

Eine weitere Möglichkeit zur Berücksichtigung der Grauwerte besteht darin, den letzten Summanden aus (4.3) bei der Minimierung als Nebenbedingung zu betrachten, was zu einer Minimierung über einem Unterraum führt.

Die in Bemerkung 4.13 erwähnte Restriktion der Simplexwahl kann somit auch genutzt werden. Ist  $\mathcal{U} \subset \mathbb{C}^N$  ein beliebiger Unterraum von  $\mathbb{C}^N$  und wählt man ein Startsimplex  $\Delta_0 \subset \mathcal{U}$ , so liegt jedes durch die Nelder und Mead Simplexmethode iterierte Simplex  $\Delta_m$  in  $\mathcal{U}$ , da das Verfahren lediglich Vektorraumoperationen verwendet. Dies entspricht einer Minimierung von  $\mathcal{J} : \mathbb{C}^N \rightarrow \mathbb{R}$  auf dem Unterraum  $\mathcal{U}$ .

Obige Nebenbedingung hat in dem hier vorliegenden Fall eine besonders einfache Gestalt, welche bei näherer Betrachtung sofort zu erkennen ist. Die Gleichung

$$\sum_{j=1}^N |c_j|^2 - \widehat{f}(0)$$

ist die Darstellung einer komplexen Sphäre mit Radius  $\sqrt{\widehat{f}(0)} \in \mathbb{R}$  um den Ursprung. Dies motiviert die Minimierung von  $\mathcal{J}$  auf dieser Sphäre, da somit die Nebenbedingungen automatisch erfüllt sind. Unter Verwendung von reellen Polarkoordinaten in der Nelder und Mead Simplexmethode ist

dies ohne wesentliche Erhöhung des Aufwands möglich. Die Methode kann wie in Abschnitt 4.2.2 beschrieben auf die Polarkoordinaten (betrachtet als Vektoren in  $\mathbb{R}^{2N-1}$ ) angewendet werden, es muss lediglich vor der Auswertung von  $\mathcal{J}$  eine Transformation  $T_r : \mathbb{R}^{2N-1} \rightarrow \mathbb{C}^N$  der (reellen) Polarkoordinaten in (komplexe) kartesische Koordinaten vorgenommen werden.  $T_r : (\phi_1, \dots, \phi_{2N-1}) \mapsto (c_1, \dots, c_N)$  ist durch

$$\begin{array}{rcl}
x_1 & := & r \begin{pmatrix} \sin \phi_1 & \sin \phi_2 & \cdots & \cdots & \cdots & \cdots & \sin \phi_{2N-1} \end{pmatrix} \\
x_2 & := & r \begin{pmatrix} \sin \phi_1 & \sin \phi_2 & \cdots & \cdots & \cdots & \sin \phi_{2N-2} & \cos \phi_{2N-1} \end{pmatrix} \\
x_3 & := & r \begin{pmatrix} \sin \phi_1 & \sin \phi_2 & \cdots & \cdots & \sin \phi_{2N-3} & \cos \phi_{2N-2} & \end{pmatrix} \\
\vdots & := & \vdots \\
x_j & := & r \begin{pmatrix} \sin \phi_1 & \cdots & \sin \phi_{2N-j} & \cos \phi_{2N-j+1} & \end{pmatrix} \\
\vdots & := & \vdots \\
x_{2N-1} & := & r \begin{pmatrix} \sin \phi_1 & \cos \phi_2 & \end{pmatrix} \\
x_{2N} & := & r \begin{pmatrix} \cos \phi_1 & \end{pmatrix}
\end{array}$$

und

$$c_j := x_{2j-1} + i x_{2j} \quad \text{für } j = 1, \dots, N$$

definiert. Die Anwendung der Nelder und Mead Simplexmethode auf  $(\phi_1, \dots, \phi_{2N-1})$  und die Wahl  $r := \widehat{f}(0)^{-\frac{1}{2}}$  minimiert  $\mathcal{J}$  auf  $\mathbb{S}_r^{2N-1}$  und berücksichtigt somit die Nebenbedingung nicht nur approximativ, sondern erfüllt sie sogar exakt. Dies ist äquivalent zur Lösung von folgendem Problem:

**Problem 4.15**

Finde ein  $\Phi^* := (\phi_1^*, \dots, \phi_{2N-1}^*) \in \mathbb{R}^{2N-1}$ , welches das Funktional  $\mathcal{J} \circ T_r$  für  $r = \sqrt{\widehat{f}(0)}$  minimiert.  $T_r(\Phi^*) =: \mathbf{c}^* \in \mathbb{C}^N$  ist dann der Koeffizientenvektor von  $f_* \in \mathcal{S}_N$ .

Dazu muss in Algorithmus 4.12 lediglich  $\mathcal{J}$  durch  $\mathcal{J} \circ T_r$  ersetzt und ein nicht degeneriertes Startsimplex  $\Delta_0$  aus  $\mathbb{R}^{2N-1}$  gewählt werden.

**4.2.4 Der gesamte Algorithmus**

Es sollen nun die einzelnen Schritte der vorherigen Abschnitte zur Rekonstruktion zusammengefasst und ein allgemeines Verfahren zur Lösung der generellen Aufgabenstellung (3.1) und des daraus resultierenden Problems 4.2 formuliert werden.

Es wurden in Abschnitt 4 insgesamt drei verschiedene Kostenfunktionale definiert, um den Koeffizientenvektor von  $f_*$  berechnen zu können. Diese unterschieden sich nur unwesentlich und führten zu Problem 4.5, Problem 4.14 und Problem 4.15. Dabei kann mit dem Koordinatenabstiegsverfahren (Algorithmus 4.10) allerdings nur Ersteres gelöst werden, während sich die Nelder und Mead Simplexmethode (Algorithmus 4.12) hingegen zur Lösung all dieser Varianten eignet. Damit ergeben sich nun vier verschiedene Möglichkeiten zur Bestimmung des Koeffizientenvektors von  $f_*$ .

Insgesamt können die wesentlichen Schritte zur Rekonstruktion eines vorgegebenen Musters  $f \notin \mathcal{S}_d$  durch ein  $f_* \in \mathcal{S}_N$  wie folgt zusammengefasst werden:

**Algorithmus 4.16**

Zu gegebenem, beliebigem und festem Muster  $f$

1. berechne die diskrete Fouriertransformation  $\widehat{f}$  von  $f$ .

2. Berechne die beste Approximation  $f_d$  durch Auswerten von  $\hat{f}$  an den Stellen  $\mathbf{k} \in r^{-1}\mathcal{K}_d$  (vgl. Aufgabe 3.3).

3. Führe Algorithmus 4.10 oder Algorithmus 4.12 zur Berechnung von  $f_*$  aus Problem 4.2 aus.

Schematisch kann dieser Algorithmus in einem Diagramm dargestellt werden.

$$f \xrightarrow{1.} \hat{f} \xrightarrow{2.} f_d \xrightarrow{3.} f_*$$

**Bemerkung 4.17**

Schritt 3. in Algorithmus 4.16 kann also nach obigen Überlegungen in vier verschiedenen Varianten durchgeführt werden.

## 5 Rang Eins Approximation

In diesem Abschnitt soll zunächst noch einmal eine Untersuchung der in [5] vorausgesetzten Situation der linearen Unabhängigkeit der Differenzen  $\mathbf{k}_{j,l} \in \mathcal{K}_d$  eingeschoben werden. Diese lässt eine bis auf diese Einschränkung beliebige Wahl von  $\mathcal{K}$  zu. Es muss allerdings insbesondere die in Abschnitt 3.1 konstruierte Form von  $\mathcal{K}$  ausgeschlossen werden, da dort lineare Abhängigkeiten zwischen den  $\mathbf{k}_{j,l}$  auftreten.

Definiert man nun den Vektor  $\mathbf{c} := (c_1, \dots, c_N)^T \in \mathbb{C}^N$  und die Matrix  $\mathbf{F} := \left( \widehat{f}(r^{-1}\mathbf{k}_{j,l}) \right)_{j,l=1}^N \in \mathbb{C}^{N \times N}$ , so kann man ein mit Problem 4.2 verwandtes, vereinfachtes Problem folgendermaßen definieren.

### Problem 5.1

Finde ein  $\mathbf{c} \in \mathbb{C}^N$ , welches das Funktional

$$E_F(\mathbf{c}) := \|\mathbf{c}\mathbf{c}^* - \mathbf{F}\|_F^2 \quad (5.1)$$

minimiert. Dabei bezeichne  $\|\mathbf{A}\|_F := \sqrt{\text{Spur}\{\mathbf{A}^*\mathbf{A}\}}$  die Frobeniusnorm der Matrix  $\mathbf{A}$ .

$E_F$  unterscheidet sich von  $E^2$  lediglich durch den Summanden  $|\sum_{j=1}^N |c_j|^2 - \widehat{f}(0)|^2$  aus (4.3), welcher in  $E_F$  die Gestalt  $\sum_{j=1}^N ||c_j|^2 - \widehat{f}(0)|^2$  hat und aus der Diagonalen von  $\mathbf{c}\mathbf{c}^* - \mathbf{F}$  in (5.1) resultiert. Es gilt also keine Gleichheit, sondern nur  $E_F \leq E^2$  und die Minimierung von  $E_F$  garantiert kein minimales  $E$ .

Die Darstellung (5.1) sei hier trotzdem erwähnt, da sie eine exakte analytische Berechnung ihres Minimums zulässt, welche im Folgenden kurz erläutert wird.

Zu der hermiteschen Matrix  $\mathbf{F}$  existiert eine unitäre Matrix  $\mathbf{U}$ , so dass  $\mathbf{U}\mathbf{F}\mathbf{U}^* = \mathbf{S}$  eine Diagonalmatrix mit den reellen Eigenwerten  $s_1 \geq s_2 \geq \dots \geq s_N$  von  $\mathbf{F}$  als Diagonalelementen ist. Somit ergibt sich

$$E_F(\mathbf{c}) = \|\mathbf{c}\mathbf{c}^* - \mathbf{F}\|_F^2 = \|\mathbf{U}\mathbf{c}\mathbf{c}^*\mathbf{U}^* - \mathbf{U}\mathbf{F}\mathbf{U}^*\|_F^2 = \|\mathbf{U}\mathbf{c}(\mathbf{U}\mathbf{c})^* - \mathbf{S}\|_F^2.$$

Setzt man nun  $\mathbf{d} = (d_1, \dots, d_N)^T := \mathbf{U}\mathbf{c}$ , so erhält man

$$E_F(\mathbf{d}) = \|\mathbf{d}\mathbf{d}^* - \mathbf{S}\|_F^2$$

und für die Ableitung dieses Funktionals in der Richtung  $\mathbf{v} := (v_1, \dots, v_N)^T \in \mathbb{C}^N$

$$\begin{aligned} D_{\mathbf{v}}(E_F(\mathbf{d})) &= 2 \text{Spur}\{(\mathbf{d}\mathbf{d}^* - \mathbf{S})(\mathbf{d}\mathbf{v}^* + \mathbf{v}\mathbf{d}^*)\} \\ &= 2 \sum_{j=1}^N \left[ (|d_j|^2 - s_j)(d_j \bar{v}_j + v_j \bar{d}_j) + \sum_{\substack{l=1 \\ l \neq j}}^N d_j \bar{d}_l (d_l \bar{v}_j + v_l \bar{d}_j) \right]. \end{aligned}$$

Ist die Matrix  $\mathbf{F}$  zusätzlich positiv definit, so sind alle reellen Eigenwerte größer als Null. In diesem Fall lassen sich Nullstellen von  $D_{\mathbf{v}}(E_F(\mathbf{d}))$  einfach bestimmen. Wählt man  $\mathbf{d}$  dermaßen, dass  $d_l = 0$  für  $l = 1, \dots, (j-1), (j+1), \dots, N$  und  $|d_j|^2 = s_j$ , so sind diese komplexen Kreise für  $j = 1, \dots, N$  und jede beliebige Wahl von  $\mathbf{v}$  Nullstellen der Ableitung und somit mögliche Minima. Die Ableitung von  $D_{\mathbf{v}}(E_F(\mathbf{d}))$  in Richtung  $\mathbf{w} := (w_1, \dots, w_N)^T \in \mathbb{C}^N$  ist gegeben durch

$$D(\mathbf{v}, \mathbf{w})(E_F(\mathbf{d})) = 2 \text{Spur}\{(\mathbf{d}\mathbf{d}^* - \mathbf{S})(\mathbf{w}\mathbf{v}^* + \mathbf{v}\mathbf{w}^*) + (\mathbf{d}\mathbf{w}^* + \mathbf{w}\mathbf{d}^*)(\mathbf{d}\mathbf{v}^* + \mathbf{v}\mathbf{d}^*)\}$$

und man erhält somit für die oben gewählten Vektoren  $\mathbf{d}$ , dass

$$\begin{aligned} D(\mathbf{v}, \mathbf{v})(E_F(\mathbf{d})) &= 2 \operatorname{Spur} \{ 2(\mathbf{d}\mathbf{d}^* - \mathbf{S})\mathbf{v}\mathbf{v}^* + (\mathbf{d}\mathbf{v}^* + \mathbf{v}\mathbf{d}^*)^2 \} \\ &= 2s_j(v_j + \bar{v}_j)^2 + 4 \sum_{\substack{l=1 \\ l \neq j}}^N (s_j - s_l) |v_l|^2 \end{aligned}$$

gilt. Für  $j = 1$  ist die zweite Ableitung also positiv definit, da  $s_1$  der größte Eigenwert von  $\mathbf{F}$  ist und  $\mathbf{d}_{min} := (d_1, 0, \dots, 0)^T$  mit  $|d_1|^2 = s_1$  ist ein Minimum von  $E_F(\mathbf{d})$ . Hieraus ergibt sich ein Minimum  $\mathbf{c}_{min}$  von  $E_F(\mathbf{c})$  aus der Rücktransformation  $\mathbf{c}_{min} = \mathbf{U}^*\mathbf{d}_{min}$ .

Problem 5.1 lässt sich für diesen ganz speziellen Fall einer positiv definiten Matrix  $\mathbf{F}$  also exakt lösen, entspricht aber nicht der wirklichen Problemstellung in diesem Kontext. Außerdem wurden keine Forderungen an das gegebene Muster  $f$  gestellt, wodurch  $\mathbf{F}$  im Allgemeinen sicherlich nicht positiv definit ist.

Um die hier gegebene Problemstellung exakt zu beschreiben, kann die Diagonale von  $\mathbf{c}\mathbf{c}^*$  durch Subtraktion von geeigneten Matrizen eliminiert werden. Damit würden auch in einem solchen Fall die mittleren Grauwerte des Musters in der Rekonstruktion nicht berücksichtigt, was aber wie schon des Öfteren erläutert physikalisch legitim ist.

Zur Elimination der  $N$  Diagonalelemente definiere zunächst die Matrizen  $\mathbf{H}_j := (\delta_{j,\xi} \delta_{j,\mu})_{\xi,\mu=1}^N \in \mathbb{C}^{N \times N}$  für  $j = 1, \dots, N$  und mit deren Hilfe die Blockmatrix  $\mathbf{H} := (\mathbf{H}_1 \dots \mathbf{H}_N) \in \mathbb{C}^{N \times N^2}$ . Weiter sei mit  $\mathbf{C} := \operatorname{diag}(\mathbf{c}\mathbf{c}^*) \in \mathbb{C}^{N^2 \times N^2}$  eine Blockdiagonalmatrix mit den Diagonalenblöcken  $\mathbf{c}\mathbf{c}^*$  definiert. Damit hat das Funktional

$$\tilde{E}_F(\mathbf{c}) := \|\mathbf{c}\mathbf{c}^* - \mathbf{H}\mathbf{C}\mathbf{H}^* - \mathbf{F}_0\|_F^2 \quad (5.2)$$

die gewünschte Form, wobei  $\mathbf{F}_0$  so definiert sei, dass die Diagonale nur Nullen enthält und  $\mathbf{F}_0$  abseits der Diagonalen mit  $\mathbf{F}$  identisch ist. Da  $\mathbf{F}_0$  ebenfalls hermitesch ist, existiert wieder eine (andere) unitäre Matrix  $\mathbf{U}$ , so dass  $\mathbf{U}\mathbf{F}_0\mathbf{U}^* =: \mathbf{S}_0$  eine Diagonalmatrix ist.  $\tilde{E}_F$  kann analog zu obiger Vorgehensweise umgeformt werden. Es ist

$$\tilde{E}_F(\mathbf{c}) = \|\mathbf{U}\mathbf{c}\mathbf{c}^*\mathbf{U}^* - \mathbf{U}\mathbf{H}\mathbf{C}\mathbf{H}^*\mathbf{U}^* - \mathbf{S}_0\|_F^2$$

und mit der Substitution  $\mathbf{d} := \mathbf{U}\mathbf{c}$  erhält man, dass

$$\tilde{E}_F(\mathbf{d}) = \|\mathbf{d}\mathbf{d}^* + \check{\mathbf{H}}\mathbf{D}\check{\mathbf{H}}^* - \mathbf{S}_0\|_F^2$$

ist, wobei  $\check{\mathbf{H}} := (\mathbf{U}\mathbf{H}_1\mathbf{U}^* \dots \mathbf{U}\mathbf{H}_N\mathbf{U}^*)$  und  $\mathbf{D} := \operatorname{diag}(\mathbf{d}\mathbf{d}^*)$  abgekürzt wurden. Definiert man nun eine weitere Blockdiagonalmatrix  $\mathbf{D}_\mathbf{v} := \operatorname{diag}(\mathbf{d}\mathbf{v}^*)$  mit den Diagonalböcken  $\mathbf{d}\mathbf{v}^*$ , so hat die Ableitung von  $\tilde{E}_F$  in Richtung  $\mathbf{v} \in \mathbb{C}^N$  die Gestalt

$$D_\mathbf{v}\tilde{E}_F(\mathbf{d}) = 2 \operatorname{Spur} \{ (\mathbf{d}\mathbf{d}^* - \check{\mathbf{H}}\mathbf{D}\check{\mathbf{H}}^* - \mathbf{S}_0)[(\mathbf{d}\mathbf{v}^* + \mathbf{v}\mathbf{d}^*) - \check{\mathbf{H}}(\mathbf{D}_\mathbf{v} + \mathbf{D}_\mathbf{v}^*)\check{\mathbf{H}}^*] \} . \quad (5.3)$$

Die Bestimmung einer exakten Nullstelle dieser Ableitung ist nicht direkt möglich, sondern müsste approximativ durch numerische Verfahren ermittelt werden.

Das Funktional  $\tilde{E}_F$  kommt der Aufgabenstellung, welche in dieser Arbeit untersucht wird, allerdings auch noch nicht sehr nahe, da es die lineare Abhängigkeit der  $\mathbf{k} \in \mathcal{K}_d$  nicht berücksichtigt. Dies ist jedoch durch einige Veränderungen möglich. Dazu definiere mit der Einheitsmatrix  $\mathbf{I} \in \mathbb{R}^{\frac{N}{2} \times \frac{N}{2}}$  für  $j = 1, \dots, \frac{N}{2}$  die Matrizen

$$\mathbf{L}_j := (\delta_{j,\xi} \delta_{\frac{N}{2}+j,\mu} + \delta_{\frac{N}{2}+j,\xi} \delta_{j,\mu})_{\xi,\mu=1}^N \quad \text{und} \quad \mathbf{R}_j := \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} - \mathbf{L}_j$$

aus  $\mathbb{C}^{N \times N}$ . Mit Hilfe dieser Matrizen können nun die Blockmatrizen

$$\mathbf{L} := (\mathbf{L}_1 \dots \mathbf{L}_{\frac{N}{2}} \mathbf{0} \dots \mathbf{0}) \in \mathbb{R}^{N \times N^2}$$

sowie

$$\mathbf{R} := (\mathbf{R}_1 \dots \mathbf{R}_{\frac{N}{2}} \mathbf{0} \dots \mathbf{0})^T \in \mathbb{R}^{N^2 \times N}$$

definiert werden. Damit berücksichtigt das Funktional

$$\check{E}_F(\mathbf{c}) := \|\mathbf{c}\mathbf{c}^* - \mathbf{H}\mathbf{C}\mathbf{H}^* + \mathbf{L}\bar{\mathbf{C}}\mathbf{R} - \mathbf{F}_0\|_F^2 \quad (5.4)$$

wie gewünscht die linearen Abhängigkeiten der Differenzen  $\mathbf{k} \in \mathcal{K}_d$ , wenn  $\mathcal{K}$  wie in (2.1) gewählt wurde. Es ist  $\bar{\mathbf{C}} = \text{diag}(\bar{\mathbf{c}}\mathbf{c}^T) \in \mathbb{C}^{N^2 \times N^2}$  die komplex konjugierte Matrix von  $\mathbf{C} \in \mathbb{C}^{N^2 \times N^2}$ .

Analog zur bisherigen Vorgehensweise wird  $\mathbf{F}_0$  wiederum durch eine unitäre Matrix  $\mathbf{U}$  auf Diagonalform gebracht. Es ist also

$$\check{E}_F(\mathbf{c}) = \|\mathbf{U}\mathbf{c}\mathbf{c}^*\mathbf{U}^* - \mathbf{U}\mathbf{H}\mathbf{C}\mathbf{H}^*\mathbf{U}^* + \mathbf{U}\mathbf{L}\bar{\mathbf{C}}\mathbf{R}\mathbf{U}^* - \mathbf{S}_0\|_F^2$$

und durch die erneute Substitution  $\mathbf{d} := \mathbf{U}\mathbf{c}$  erhält man

$$\check{E}_F(\mathbf{d}) = \|\mathbf{d}\mathbf{d}^* - \check{\mathbf{H}}\mathbf{D}\check{\mathbf{H}}^* + \check{\mathbf{L}}\bar{\mathbf{D}}\check{\mathbf{R}} - \mathbf{S}_0\|_F^2,$$

wobei nun

$$\check{\mathbf{L}} := (\mathbf{U}\mathbf{L}_1\mathbf{U}^T \dots \mathbf{U}\mathbf{L}_{\frac{N}{2}}\mathbf{U}^T \mathbf{0} \dots \mathbf{0}) \in \mathbb{R}^{N \times N^2}$$

sowie

$$\check{\mathbf{R}} := (\bar{\mathbf{U}}\mathbf{R}_1\mathbf{U}^* \dots \bar{\mathbf{U}}\mathbf{R}_{\frac{N}{2}}\mathbf{U}^* \mathbf{0} \dots \mathbf{0})^T \in \mathbb{R}^{N \times N^2}$$

abgekürzt wurde und auch hier  $\bar{\mathbf{D}} = \text{diag}(\bar{\mathbf{D}}\mathbf{D}^T)$  die komplex konjugierte Matrix von  $\mathbf{D}$  ist. Dies gilt auch für die in der Ableitung auftretende Matrix  $\bar{\mathbf{D}}_{\mathbf{v}} = \text{diag}(\bar{\mathbf{d}}\mathbf{v}^T)$ .

Leitet man nun  $\check{E}_F$  in Richtung  $\mathbf{v} \in \mathbb{C}^N$  ab, so ergibt sich für die Ableitung

$$D_{\mathbf{v}}\check{E}_F(\mathbf{d}) = 2 \text{Spur} \left\{ (\mathbf{d}\mathbf{d}^* - \check{\mathbf{H}}\mathbf{D}\check{\mathbf{H}}^* + \check{\mathbf{L}}\bar{\mathbf{D}}\check{\mathbf{R}} - \mathbf{S}_0) \right. \\ \left. [(\mathbf{d}\mathbf{v}^* + \mathbf{v}\mathbf{d}^*) - \check{\mathbf{H}}(\mathbf{D}_{\mathbf{v}} + \mathbf{D}_{\mathbf{v}}^*)\check{\mathbf{H}}^* + \check{\mathbf{L}}(\bar{\mathbf{D}}_{\mathbf{v}} + \bar{\mathbf{D}}_{\mathbf{v}}^*)\check{\mathbf{R}}] \right\}.$$

Die Nullstellen von  $D_{\mathbf{v}}\check{E}_F$  sind auch hier nicht direkt theoretisch bestimmbar, was wiederum die Verwendung numerischer Verfahren notwendig macht.

Die Anpassung der einfachen Darstellung  $\|\mathbf{c}\mathbf{c}^* - \mathbf{F}\|_F^2$  an die eigentliche Problemstellung sowohl für linear unabhängige als auch linear abhängige  $\mathbf{k} \in \mathcal{K}_d$  hat also in beiden Fällen zur Folge, dass die Nullstellen der Ableitung nicht mehr exakt bestimmt werden können.

Die Haupteigenschaft, welche Anlass zur Betrachtung dieses Ansatzes gab, geht somit verloren. Dies liegt daran, dass in den Funktionalen  $\check{E}_F$  und  $\check{E}_F$  sowie deren Ableitungen die weiteren Summanden notwendig sind. Die Diagonalgestalt der Matrix  $\mathbf{S}_0$  kann dadurch nicht mehr ausgenutzt werden wie

es noch beim Funktional  $E_F$  der Fall war.

Es besteht natürlich trotzdem die Möglichkeit, die Nullstellen der jeweiligen Ableitungen durch ein numerisches Verfahren wie z.B. das Newton-Verfahren zu bestimmen, was eine Alternative zu den in Abschnitt 4.2 vorgeschlagenen Verfahren wäre. Diese Alternative wird in dieser Arbeit allerdings nicht weiter verfolgt, sondern das numerische Hauptaugenmerk auf die Verfahren in Abschnitt 4.2 gelegt.

## 6 Ridge-Funktionen

Die Funktionen  $\exp(i\langle \mathbf{k}, \mathbf{x} \rangle)$ , welche die Basis zur Darstellung von  $f_N$  in (3.3), aber auch des Raumes  $\mathcal{S}_d$  in (3.17) bilden, sind in der Literatur im Zusammenhang mit partiellen Differentialgleichungen als *ebene Wellen* bekannt [6, 10]. In einer früheren Arbeit über Computer-Tomographie wurden diese Basisfunktionen *Ridge-Functions* genannt [12], treten aber auch auf natürliche Weise in der harmonischen Analysis und speziellen Funktionentheorie sowie bei neuronalen Netzwerken auf [17]. Zur Approximation durch Linearkombinationen von *Ridge-Functions* existieren einige Resultate, welche Algorithmen und deren Approximations- und Konvergenzeigenschaften betreffen [17, 19]. Dieser Ansatz wurde auch schon in [5] angesprochen, jedoch nicht weiter verfolgt. In diesem Abschnitt soll nun untersucht werden, inwiefern *Ridge-Functions* in den hier betrachteten Rahmen passen, d.h. ob die Räume  $\mathcal{S}_N$  und  $\mathcal{S}_d$  *Ridge-Function*-Räume sind, ob Algorithmen zur Approximation verwendet werden können und ob man weitere approximationstheoretische Erkenntnisse gewinnen kann.

Zunächst wird eine kurze Einführung in den Themenbereich der *Ridge-Functions* gegeben.

Als *Ridge-Function* bezeichnet man eine Funktion  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  der Form

$$h(x_1, \dots, x_n) = \phi(\langle \mathbf{a}, \mathbf{x} \rangle),$$

mit einer Funktion  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  und den Vektoren  $\mathbf{a} := (a_1, \dots, a_n)^T \in \mathbb{R}^n \setminus \{\mathbf{0}\}$  sowie  $\mathbf{x} := (x_1, \dots, x_n)^T \in \mathbb{R}^n$ . Es handelt sich also um eine multivariate Funktion, welche auf den parallelen Hyperebenen  $\langle \mathbf{a}, \mathbf{x} \rangle = \alpha$ ,  $\alpha \in \mathbb{R}$ , konstant ist. Der Vektor  $\mathbf{a}$  wird üblicherweise (und in Einklang mit der Bezeichnung in den vorherigen Abschnitten) als *Richtung* bezeichnet.

Es gibt verschiedene Möglichkeiten, mit Hilfe dieser Basisfunktionen Approximationsräume zu definieren, wobei hier der Raum

$$\mathcal{R}(\mathbf{a}^1, \dots, \mathbf{a}^s) := \left\{ \sum_{j=1}^s \phi_j(\langle \mathbf{a}^j, \mathbf{x} \rangle) : \phi_j : \mathbb{R} \rightarrow \mathbb{R} \right\} \quad (6.1)$$

mit  $s \in \mathbb{N}$  gegebenen festen Richtungen  $\mathbf{a}^j$  betrachtet wird. Die Funktionen  $\phi_j$  stellen hier die Freiheitsgrade dar. Für diesen Raum werden in [19] einige Resultate bewiesen, während für andere Räume, mit z.B. zusätzlich variablen Richtungen  $\mathbf{a}^j$ , wenige Ergebnisse bekannt sind.

Der Raum passt insofern in den Rahmen dieser Arbeit, als dass die Strahlenrichtungen  $\mathbf{k} \in \mathcal{K}$  und somit die Differenzen  $\mathbf{k} \in \mathcal{K}_d$  hier auch fest gewählt sind.

Es gibt eine wesentliche Methode, welche aus  $\mathcal{R}(\mathbf{a}^1, \dots, \mathbf{a}^s)$  in einer bestimmten Norm über einem Gebiet im  $\mathbb{R}^n$  approximiert. Diese ist unter anderem als *Von Neumann Alternating Algorithm*, *Cyclic Coordinate Algorithm*, *Schwarz Decomposition Method* oder *Diliberto-Strauss Algorithm* bekannt. Die Idee, welche dieser Methode zu Grunde liegt, besteht darin, den Approximationsraum als Summe

$$\mathcal{R}(\mathbf{a}^1, \dots, \mathbf{a}^s) = \mathcal{M}_1 + \dots + \mathcal{M}_s$$

der Räume

$$\mathcal{M}_j := \{g(\langle \mathbf{a}^j, \mathbf{x} \rangle) : g \in X_j\} \quad (6.2)$$

und geeigneten Funktionenräumen  $X_j$  für  $j = 1, \dots, s$  darzustellen.

Sei nun  $P_j$  ein Operator der besten Approximation an  $\mathcal{M}_j$ , d.h.  $P_j g$  ist für  $j = 1, \dots, s$  die beste Approximation aus  $\mathcal{M}_j$  an die (beliebige) Funktion  $g : \mathbb{R}^n \rightarrow \mathbb{R}$ . Definiert man nun den Operator

$$E := (I - P_s) \dots (I - P_1), \quad (6.3)$$

so iteriert der Algorithmus diesen Operator  $E$ , d.h. man betrachtet den Grenzwert

$$\lim_{\ell \rightarrow \infty} E^\ell g.$$

Unter gewissen Voraussetzungen konvergiert dieser Algorithmus für  $s = 2$  (*Diliberto-Strauss Algorithm*) wie gewünscht gegen die beste Approximation aus dem Raum  $\mathcal{M}_1 + \dots + \mathcal{M}_s$ . Für  $s \geq 3$  ist dies jedoch nicht mehr ohne Weiteres der Fall. In [19] sind für diese Situation zwei Theoreme angegeben, wovon eines die Konvergenz unter gewissen Voraussetzungen garantiert und das zweite die Konvergenzrate angibt. Weiter werden Aussagen über den Grad der Approximation gemacht.

In dem hier vorgestellten mathematischen Modell sind die beiden geschachtelten Räume  $\mathcal{S}_N \subset \mathcal{S}_d$  die Ansatzräume, aus denen Approximationen gesucht sind. Zur Erinnerung: Es soll eine Funktion  $f \notin \mathcal{S}_d$  durch ein  $g \in \mathcal{S}_N$  approximiert werden. Dazu wird zunächst die beste Approximation  $f_d \in \mathcal{S}_d$  bestimmt, welche sich direkt aus der Funktion  $f$  berechnen lässt. Die Hauptaufgabe besteht dann darin, die beste Approximation  $f_*$  aus  $\mathcal{S}_N$  an  $f_d$  zu finden.

Es stellt sich nun im Zusammenhang mit den eben vorgestellten *Ridge-Functions* die Frage, ob die Räume  $\mathcal{S}_d$  und  $\mathcal{S}_N$  von der gleichen Gestalt wie  $\mathcal{R}(\mathbf{a}^1, \dots, \mathbf{a}^s)$ , d.h. die Methoden und Aussagen aus [19] hier anwendbar sind. Dabei soll zunächst der Raum  $\mathcal{S}_d$  betrachtet werden, da aus diesem Raum die „erste“ Approximation  $f_d$  berechnet wird.

Eine Funktion  $g \in \mathcal{S}_d$  hat die Form

$$g(\mathbf{x}) = \sum_{\mathbf{k} \in \mathcal{K}_d} c_{\mathbf{k}} \exp(i\langle \mathbf{k}, \mathbf{x} \rangle),$$

was durch eine geeignete Nummerierung der  $\mathbf{k} \in \mathcal{K}_d$  und unter Ausnutzung der Tatsache, dass aus  $\mathbf{k} \in \mathcal{K}_d$  folgt, dass auch  $-\mathbf{k} \in \mathcal{K}_d$  ist, als

$$g(\mathbf{x}) = \sum_{j=1}^s (c_{\mathbf{k}^j} \exp(i\langle \mathbf{k}^j, \mathbf{x} \rangle) + k.K.) \quad (6.4)$$

geschrieben werden kann. Hier wurde allerdings wie schon des Öfteren in dieser Arbeit, das Element  $\mathbf{0}$  aus  $\mathcal{K}_d$  entfernt (vgl. Bemerkung 4.3). Es ist somit  $s = \#\mathcal{K}_d/2$  und die hochgestellten Indizes der  $\mathbf{k}$  verdeutlichen, dass hier im Allgemeinen eine andere Indizierung vorliegt, als sie bisher verwendet wurde ( $\mathbf{k}_j \neq \mathbf{k}^j$ ). Die Indizes der komplexen Koeffizienten  $c_{\mathbf{k}}$  und  $c_{\mathbf{k}^j}$  sollen hier lediglich andeuten, dass sie zur jeweiligen Richtung  $\mathbf{k}$  bzw.  $\mathbf{k}^j$  gehören. Diese Darstellung des Raumes  $\mathcal{S}_d$  wurde gewählt, da hierdurch die Gestalt von  $\mathcal{R}(\mathbf{a}^1, \dots, \mathbf{a}^s)$  zu erkennen ist. Es ist nämlich

$$\begin{aligned} \mathcal{S}_d = \mathcal{R}(\mathbf{k}^1, \dots, \mathbf{k}^s) &= \left\{ \sum_{j=1}^s \phi_j(\langle \mathbf{k}^j, \mathbf{x} \rangle) : \phi_j : \mathbb{R} \rightarrow \mathbb{R} \right\} \\ &= \mathcal{M}_1 + \dots + \mathcal{M}_s \end{aligned}$$

mit

$$\mathcal{M}_j := \{ \phi(\langle \mathbf{k}^j, \mathbf{x} \rangle) : \phi \in X \}$$

und

$$X := \{ \phi(\varphi) = c \exp(i\varphi) + \bar{c} \exp(-i\varphi) : c \in \mathbb{C} \} .$$

Man beachte, dass die Elemente von  $X$  Funktionen  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  sind.  $\mathcal{S}_d$  hat also die gewünschte Form eines *Ridge-Function*-Raumes, wie er in [19] untersucht wird. Betrachtet man nun als Nächstes die Approximation aus  $\mathcal{S}_d$  an  $f$  bezüglich der Norm  $\| \cdot \|_{\Omega} := \| \hat{\cdot} \|_{\ell_2}$ , so ist leicht nachzurechnen, dass für  $j = 1, \dots, s$

$$P_j f(\mathbf{x}) := \hat{f}(r^{-1} \mathbf{k}^j) \exp(i \langle \mathbf{k}^j, \mathbf{x} \rangle) + \hat{f}(-r^{-1} \mathbf{k}^j) \exp(i \langle -\mathbf{k}^j, \mathbf{x} \rangle)$$

die beste Approximation aus  $\mathcal{M}_j$  an  $f$  ist.  $P_j$  ist also der in [19] geforderte Operator der besten Approximation, welcher zur Definition von (6.3) benötigt wird. Diese Definition führt direkt zu der Gleichung

$$\begin{aligned} Ef &= f - \sum_{j=1}^s \left( \hat{f}(r^{-1} \mathbf{k}^j) \exp(i \langle \mathbf{k}^j, \mathbf{x} \rangle) + \hat{f}(-r^{-1} \mathbf{k}^j) \exp(i \langle -\mathbf{k}^j, \mathbf{x} \rangle) \right) \\ &= f - f_d . \end{aligned}$$

Weiter gilt für  $\ell \geq 1 \in \mathbb{N}$ ,  $E^\ell f = Ef$  und somit folgt, dass sich der Grenzwert

$$\lim_{\ell \rightarrow \infty} E^\ell f = f - f_d$$

direkt ergibt.

Die Untersuchungen in diesem Rahmen führen also ebenfalls zur besten Approximation  $f_d$  aus  $\mathcal{S}_d$  wie sie schon in (3.18) definiert wurde und direkt berechnet werden kann.

In [19] wird auch ein Resultat über den Grad der Approximation mit *Ridge-Functions* vorgestellt. Leider kann dieses hier nicht verwendet werden, da zur Herleitung dieser Abschätzung für den Raum  $\pi_m^n$  der Polynome auf  $\mathbb{R}^n$  mit Höchstgrad  $m$  gefordert wird, dass

$$\pi_m^n \subset \mathcal{R}(\mathbf{a}^1, \dots, \mathbf{a}^s)$$

gilt. Dies ist für  $\mathcal{R}(\mathbf{k}^1, \dots, \mathbf{k}^s) = \mathcal{S}_d$  nicht erfüllt, weshalb die Abschätzung für den Approximationsgrad nicht auf  $\mathcal{S}_d$  angewendet werden kann. Diese Untersuchungen für  $\mathcal{S}_d$  bieten also lediglich einen alternativen Zugang zur besten Approximation  $f_d \in \mathcal{S}_d$ .

Es soll nun untersucht werden, inwieweit sich der Raum  $\mathcal{S}_N$  in die Form  $\mathcal{R}(\mathbf{a}^1, \dots, \mathbf{a}^s)$  bringen lässt. Dies ist sicherlich die wesentlich interessantere Aufgabe, da die Approximation von  $f_d$  durch ein Element aus  $\mathcal{S}_N$  ja den Kern der Rekonstruktion des Musters  $f$  darstellt. Es stellt sich allerdings heraus, dass es auch die viel schwierigere Aufgabe ist, welche sich wie im Folgenden erläutert, nicht lösen lässt.

Zur Vereinfachung soll hier nur der Fall der linear unabhängigen Differenzen  $\mathbf{k} \in \mathcal{K}_d$  betrachtet werden, da hierdurch die auftretenden Probleme ausreichend klar werden und sich diese Situation

auf den linear abhängigen Fall durch analoge Vorgehensweise übertragen lässt.

Eine Funktion  $g \in \mathcal{S}_N$  hat die Form

$$g(\mathbf{x}) = \frac{1}{2} \sum_{l=1}^N |c_l|^2 + \frac{1}{2} \sum_{\substack{j,l=1 \\ j < l}}^N \left( c_j \bar{c}_l \exp(i \langle \mathbf{k}_{j,l}, \mathbf{x} \rangle) + k.K. \right),$$

wobei wieder  $\mathbf{k}_{j,l} := \mathbf{k}_j - \mathbf{k}_l$  definiert ist und die  $\mathbf{k}_j$  bzw.  $\mathbf{k}_l$  die Richtungen der Strahlen und die  $c_j$  bzw.  $c_l$  die komplexen Koeffizienten der jeweiligen Strahlen sind. Mit  $k.K.$  sei wiederum das komplex Konjugierte bezeichnet. Beschränkt man sich auch hier nur auf die Menge  $\mathcal{K}_d \setminus \{\mathbf{0}\}$ , so genügt es die zweite Summe betrachten (vgl. Bemerkung 4.3). Da die einzigen Freiheitsgrade, welche man zur Konstruktion der Funktion  $g$  zur Verfügung hat, die  $N$  komplexen Koeffizienten  $c_j$  für  $j = 1, \dots, N$  sind, was einem komplexen Vektor  $\mathbf{c} := (c_1, \dots, c_N)^T \in \mathbb{C}^N$  entspricht, muss hier  $\mathcal{M}_j$  wie folgt definiert werden. Für  $j = 1, \dots, N$  sei der Raum

$$\mathcal{M}_j = \left\{ \phi \left( \langle \mathbf{k}_{j,1}, \mathbf{x} \rangle, \dots, \langle \mathbf{k}_{j,N}, \mathbf{x} \rangle \right) : \phi \in X_j \right\}$$

mit multivariaten Funktionen  $\phi$  definiert. Weiter sei  $\mathbf{e}_j$  der  $j$ -te Einheitsvektor des  $\mathbb{R}^N$ ,  $\mathbf{I}_j$  die Einheitsmatrix des  $\mathbb{R}^N$ , bei der das  $j$ -te Diagonalelement durch 0 ersetzt wurde und  $\mathbf{e}^{i\varphi} := (e^{i\varphi_1}, \dots, e^{i\varphi_N})^T$  für  $\varphi_1, \dots, \varphi_N \in \mathbb{R}$ . Der Raum  $X_j$  sei dann für  $j = 1, \dots, N$  durch

$$X_j := \left\{ \frac{1}{4} (\mathbf{e}_j^T \mathbf{c} \mathbf{c}^* \mathbf{I}_j \mathbf{e}^{i\varphi} + k.K.) : c_j \in \mathbb{C} \right\}$$

definiert.

Mit diesen Definitionen von  $\mathcal{M}_j$  und  $X_j$  erreicht man zwar die Darstellung

$$\mathcal{S}_N = \mathcal{M}_1 + \dots + \mathcal{M}_N$$

von  $\mathcal{S}_N$  als Summe, aber der Raum  $\mathcal{M}_j$  hat nicht die gewünschte Gestalt: Die Funktionen  $\phi \in \mathcal{M}_j$  sind nicht auf  $\mathbb{R}$  definiert, d.h.  $\phi : \mathbb{R} \rightarrow \mathbb{R}$ , sondern auf  $\mathbb{R}^N$ , was bedeutet, dass  $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$  ist. Eine alternative Definition von  $\mathcal{M}_j$  ist nicht möglich, da der Koeffizient  $c_j$  für alle  $j = 1, \dots, N$  in der Summe

$$g(\mathbf{x}) = \frac{1}{2} \sum_{\substack{j,l=1 \\ j < l}}^N \left( c_j \bar{c}_l \exp(i \langle \mathbf{k}_{j,l}, \mathbf{x} \rangle) + k.K. \right),$$

in  $(N - 1)$  Summanden mit verschiedenen Richtungen (im *Ridge-Function*-Sinne) auftaucht.

Der Raum  $\mathcal{S}_N$  kann also nicht in die Form eines *Ridge-Function*-Raumes gebracht werden. Somit können auch die Methoden, welche für *Ridge-Functions* bekannt sind, in diesem Kontext nicht verwendet werden.

Für  $\mathcal{S}_N$  wurde bei obiger Darstellung die lineare Unabhängigkeit der Differenzen  $\mathbf{k} \in \mathcal{K}_d$  vorausgesetzt. Die Probleme, die hier auftauchen, treten ebenso auf wenn die  $\mathbf{k} \in \mathcal{K}_d$  linear abhängig sind. Die Koeffizienten  $c_j$  tauchen in diesem Fall, zusätzlich zu den  $(N - 1)$  Summanden wie oben, in

$(N - 1)$  weiteren Termen in komplex konjugierter Form auf. Da eine genaue Darstellung dieser Problematik nur zu größerer Unübersichtlichkeit führt und das Resultat das Gleiche ist, wird hierauf verzichtet.

Insgesamt hat dieser Abschnitt gezeigt, dass  $\mathcal{S}_d$  zwar als *Ridge-Function*-Raum  $\mathcal{R}(\mathbf{a}^1, \dots, \mathbf{a}^s)$  dargestellt werden kann, aber die Approximationstechniken für *Ridge-Functions* das gleiche Resultat  $f_d \in \mathcal{S}_d$  liefern, wie schon bekannt und die Aussagen über den Grad der Approximation nicht verwendet werden können. Da die beste Approximation  $f_d$  (und somit auch  $E_{\text{approx}}$ ) ohne Probleme direkt zu berechnen ist, bietet dies lediglich eine andere Betrachtungsweise. *Ridge-Functions* bieten somit zwar einen alternativen Definitionsrahmen, aber keine weiteren Ansatzmöglichkeiten zur Approximation.

Für den Raum  $\mathcal{S}_N$  scheitert sogar schon der Versuch, eine Darstellung zu finden, welche die gleiche Struktur wie  $\mathcal{R}(\mathbf{a}^1, \dots, \mathbf{a}^s)$  aufweist. Dies macht es natürlich unmöglich, irgendein Theorem oder einen Algorithmus aus [19] oder dem allgemeinen Themenbereich der *Ridge-Functions* anzuwenden. P. P. Petrushev bezeichnet das approximationstheoretische Problem als schwer [18].

Die Approximation mit *Ridge-Functions* ist also nicht der richtige Ansatz zur Behandlung des Problems der Rekonstruktion eines gegebenen Muster.

## 7 Numerische Ergebnisse

Nachdem sich alle bisherigen Abschnitte mit theoretischen Untersuchungen und Berechnungen beschäftigt haben, sollen die beschriebenen Verfahren nun auch numerisch angewandt werden. Dabei werden verschiedene Aspekte untersucht:

Die Effizienz der verschiedenen Verfahren, die Auflösung des Fouriergitters, die Abhängigkeit des Koordinatenabstiegsverfahrens von den gewählten Startwerten sowie seine Konvergenzgeschwindigkeit.

Zur Untersuchung der Effizienz von Koordinatenabstiegsverfahren und der Nelder und Mead Simplexmethode werden die Resultate für verschiedene Strahlensetups und mehrere Muster angegeben. Die jeweiligen Strahlensetups werden mit der in Abschnitt 3.1 beschriebenen Methode erzeugt, was zu Beginn des jeweiligen Unterabschnitts vorgeführt wird. Sie eignen sich somit per Konstruktion zur Generierung von periodischen Mustern.

### 7.1 Berechnung verschiedener Muster

Zur Evaluation wurde Algorithmus 4.16 in all seinen Varianten für verschiedene Muster und Strahlensetups durchgeführt. Das bedeutet konkret, dass Problem 4.5 mit Koordinatenabstiegsverfahren und Simplexverfahren gelöst wurde sowie Letzteres zusätzlich zur Lösung von Problem 4.14 und Problem 4.15 Verwendung fand.

Zur Unterscheidung der Verfahren sollen nun Bezeichnungen vereinbart werden. Im Folgenden ist mit *Koordinatenabstiegsverfahren* die Anwendung von Algorithmus 4.10 und mit *Simplexverfahren* die Anwendung von Algorithmus 4.12 auf das Funktional  $\mathcal{J}$  in (4.4) bezeichnet (siehe (i) bzw. (ii) in folgender Aufzählung). Entsprechend ist mit *Simplexverfahren (mit NB)*, NB steht für Nebenbedingung, die Anwendung von Algorithmus 4.12 auf (4.11) gemeint (siehe (iii) in folgender Aufzählung) und mit *polarem Simplexverfahren* die Anwendung auf  $\mathcal{J} \circ T_r$ . (siehe (iv) und Abschnitt 4.2.3) Dementsprechend ist die Definition des Kostenfunktionals  $\mathcal{J}_c$  im jeweiligen Zusammenhang mit diesen Bezeichnungen zu wählen.

In anderen Worten heißt dies, dass für jedes Strahlensetup jedes Muster auf vier verschiedene Arten rekonstruiert wurde (vgl. Bemerkung 4.17): Dies geschah durch Minimierung des Kostenfunktionals  $\mathcal{J}_c$  mit

- (i)  $\mathcal{J}_c := \mathcal{J}$  durch das *Koordinatenabstiegsverfahren*,
- (ii)  $\mathcal{J}_c := \mathcal{J}$  durch die Nelder und Mead Simplexmethode, (*Simplexverfahren*)
- (iii)  $\mathcal{J}_c(\mathbf{c}) := \mathcal{J}(\mathbf{c}) + \left| \sum_{j=1}^N |c_j|^2 - \widehat{f}(0) \right|^2$  durch die Nelder und Mead Simplexmethode (*Simplexverfahren (mit NB)*) sowie
- (iv)  $\mathcal{J}_c := \mathcal{J} \circ T_r$  durch die Nelder und Mead Simplexmethode (*polares Simplexverfahren*).

Dabei bezeichnet  $\mathbf{c} := (c_1, \dots, c_N)^T \in \mathbb{C}^N$  wieder den Koeffizientenvektor von  $f_*$ . Zur Bewertung der Verfahren wurde das jeweils minimierte Kostenfunktional berechnet, welches in diesem Abschnitt immer mit  $\mathcal{J}_c$  bezeichnet wird. Weiter wurde der Fehler  $E$  zwischen  $\widehat{f}_d$  und  $\widehat{f}_*$  in der  $\ell_2$ -Norm sowie der Fehler  $E_{\text{approx}}$  zwischen  $\widehat{f}$  und  $\widehat{f}_d$  berechnet, wobei  $f$  immer das gegebene und  $f_*$  das berechnete Muster ist. Mit  $f_d$  ist wieder die beste Approximation aus  $\mathcal{S}_d$  an  $f$  gemeint wie sie in (3.18) definiert wurde. Diese Fehler wurden zusätzlich „ohne den Ursprung“, d.h. für  $N(\cdot)$  aus (4.5)

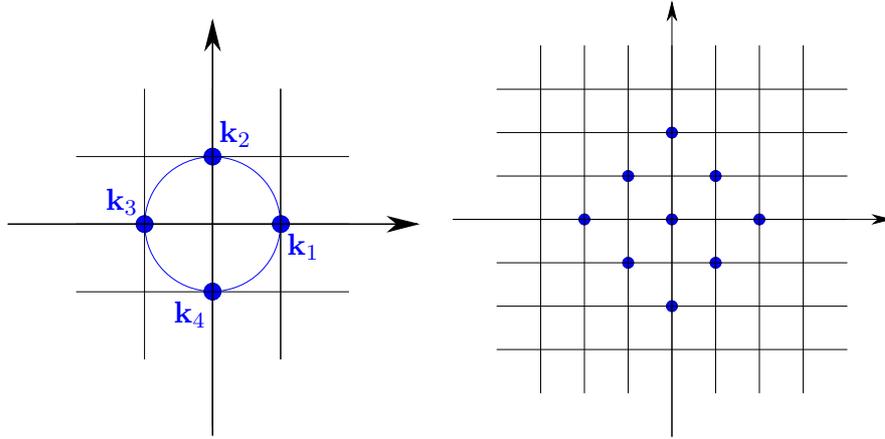


Abbildung 7.1: Lage und Anordnung der Wellenvektoren  $\mathbf{k}_j \in \mathcal{K}$  definiert in (7.1) und die resultierenden Differenzen.

berechnet, was den Funktionswert  $\hat{f}(0)$  ausschließt. Dies wurde gemacht, da nur zwei Verfahren die Nebenbedingungen einbeziehen, was einer Berücksichtigung dieses Wertes entspricht. Die Resultate werden im Folgenden vorgestellt.

Das gegebene Muster  $f$  ist jeweils ein Bild der Größe  $256 \times 256$  Pixel mit Grauwerten zwischen 0 (schwarz) und 255 (weiß) und entspricht einer Periode eines periodischen Musters. Es wird eine Skalierung vorgenommen, so dass  $f : [0, 256]^2 \rightarrow [0, 1]$  abbildet. Weiter ist bei allen Untersuchungen die Auflösung des Fouriergitters  $L = 256$  gewählt. Abschnitt 7.2 wird zeigen, dass diese Wahl sinnvoll ist. In Algorithmus 4.10 wurde der Abbruchparameter  $\mu = 10^{-5}$  gewählt. Alle weiteren Parameter werden in den jeweiligen Abschnitten, Tabellen und Bildunterschriften angegeben.

Im Folgenden sei mit  $f_*$  die mittels der beschriebenen Verfahren berechnete Lösung bezeichnet.

### 7.1.1 Die Rekonstruktion mit vier Strahlen

Das Strahlensetup, welches in diesem Abschnitt untersucht wird ergibt sich durch die Wahl  $R = 1$  in der Methode zur Generierung von  $\mathcal{K}$  aus Abschnitt 3.1. Somit ist  $(0, 1)$  das einzige zulässige 2-Tupel, woraus sich  $\#\mathcal{I}(R) = 1$  und  $\mathcal{N}(\mathcal{I}(R)) = 1$  ergibt und man insgesamt  $N = 4$  Elemente in  $\mathcal{K}$  hat. Durch die weitere Konstruktion ergibt sich  $r = r' = K$  und damit die Menge

$$\mathcal{K} := \left\{ K \begin{pmatrix} 1 \\ 0 \end{pmatrix}, K \begin{pmatrix} 0 \\ 1 \end{pmatrix}, K \begin{pmatrix} -1 \\ 0 \end{pmatrix}, K \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\}. \quad (7.1)$$

Die Menge ist in Abb. 7.1 graphisch dargestellt. Die resultierenden Differenzen liegen alle nah am Ursprung. Die beste Approximation  $f_d \in \mathcal{S}_d$  entsteht also aus den Fourierkoeffizienten von  $f$  mit „niedrigen Frequenzen“ (vergl (3.18)). Es ist somit zu erwarten, dass nur grobe Strukturen rekonstruiert werden können, da keine Fourierkoeffizienten mit „höheren Frequenzen“ zur Verfügung stehen. Die folgenden Ergebnisse zeigen fast alle nur unzureichende Annäherungen an das gewünschte Muster.

Für jedes Muster ist, neben den vier Rekonstruktionen, das Originalmuster  $f$  und die beste Approximation  $f_d$  dargestellt. Die numerischen Ergebnisse der Rekonstruktionen und die Parameter der Simplexmethode sind jeweils in einer Tabelle angegeben. In der Summe sind dies also sechs

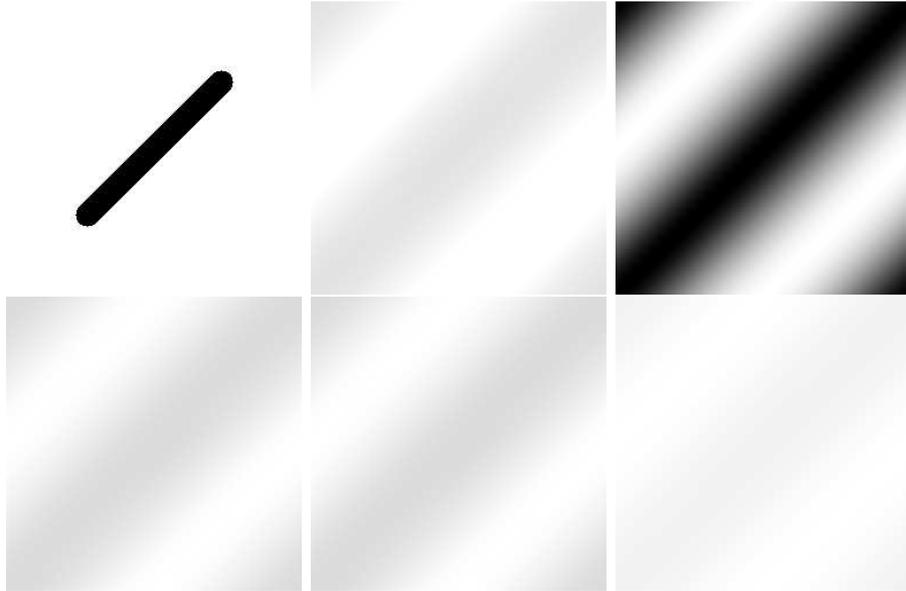


Abbildung 7.2: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.1 zu finden.

Abbildungen und zwei Tabellen pro Muster. Die a priori berechenbaren Fehler und Normen aller Muster sind in der Tabelle 7.15 zusammengefasst.

Da die Verfahren aus Abschnitt 4.2 zur Rekonstruktion von  $f_d$  angewendet werden, muss zur Bewertung der Verfahren das Resultat  $f_*$  mit diesem Muster verglichen werden (vgl. Bemerkung 3.6).

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$1.243e-06$	$7.780e-02$	$1.347e-01$	$1.089e-03$	$7.807e-01$
Simplexverfahren	$3.421e-07$	$7.781e-02$	$2.719e+00$	$6.244e-04$	$1.827e+00$
Simplexverfahren (mit NB)	$3.487e-07$	$7.781e-02$	$9.470e-01$	$6.514e-04$	$5.301e-02$
polares Simplexverfahren	$3.432e-07$	$7.781e-02$	$9.470e-01$	$6.516e-04$	$5.301e-02$

Tabelle 7.1: Numerische Resultate der Rekonstruktion von Muster Abb. 7.2 mit vier Strahlen und den beschriebenen Verfahren.

<b>Parameter</b>	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	1.7	1.3	0.9
$\gamma$	1.6	0.1	1.2
$\theta$	0.6	0.8	0.6

Tabelle 7.2: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.2.

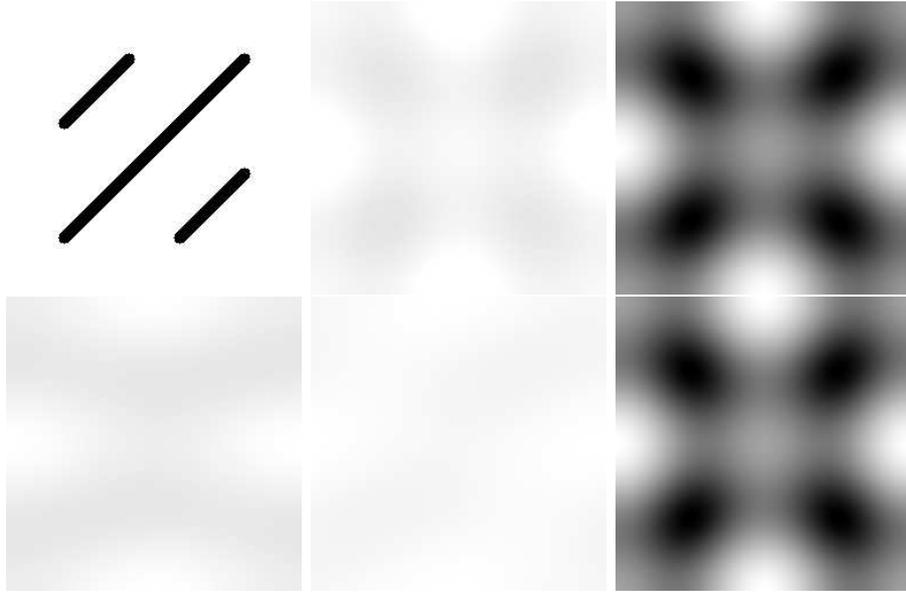


Abbildung 7.3: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.3 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$1.232e-05$	$4.829e-02$	$9.294e-02$	$4.880e-03$	$7.947e-01$
Simplexverfahren	$6.324e-06$	$4.999e-02$	$9.797e-02$	$3.344e-03$	$7.898e-01$
Simplexverfahren (mit NB)	$4.361e-04$	$4.066e-02$	$9.358e-01$	$2.952e-02$	$6.760e-02$
polares Simplexverfahren	$8.083e-04$	$1.763e-02$	$9.351e-01$	$4.021e-02$	$7.291e-02$

Tabelle 7.3: Numerische Resultate der Rekonstruktion von Muster Abb. 7.3 mit vier Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.8	1.5	0.8
$\gamma$	2.3	2.3	0.5
$\theta$	0.7	0.7	0.2

Tabelle 7.4: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.3.

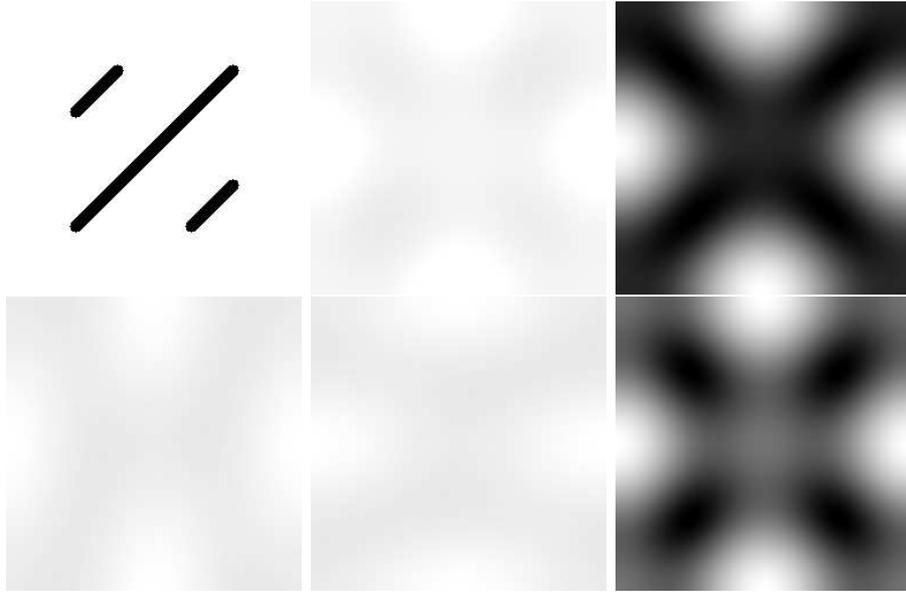


Abbildung 7.4: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.5 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$7.334e-05$	$4.116e-02$	$5.909e-02$	$1.211e-02$	$8.577e-01$
Simplexverfahren	$1.193e-05$	$4.298e-02$	$7.435e-02$	$4.785e-03$	$8.393e-01$
Simplexverfahren (mit NB)	$2.234e-04$	$3.792e-02$	$9.494e-01$	$2.112e-02$	$5.307e-02$
polares Simplexverfahren	$2.234e-04$	$3.792e-02$	$9.494e-01$	$2.112e-02$	$5.307e-02$

Tabelle 7.5: Numerische Resultate der Rekonstruktion von Muster Abb. 7.4 mit vier Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	1.1	1.3	0.9
$\gamma$	0.4	0.1	2.9
$\theta$	0.6	0.6	0.7

Tabelle 7.6: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.4.

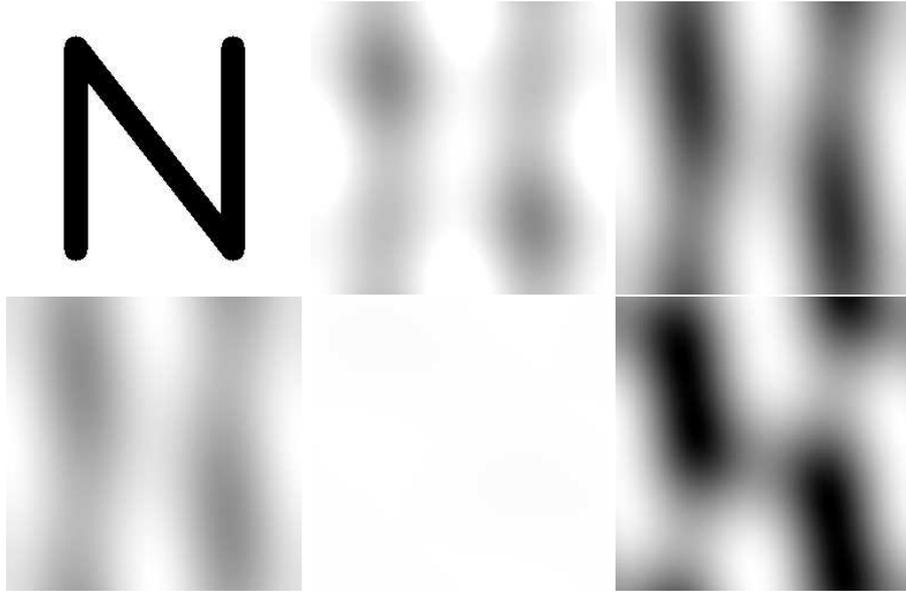


Abbildung 7.5: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.7 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$1.303e-03$	$1.672e-01$	$4.269e-01$	$5.105e-02$	$2.815e-01$
Simplexverfahren	$5.141e-03$	$1.474e-01$	$2.953e-01$	$1.014e-01$	$4.259e-01$
Simplexverfahren (mit NB)	$1.137e-03$	$1.698e-01$	$8.357e-01$	$4.769e-02$	$1.562e-01$
polares Simplexverfahren	$1.545e-02$	$3.253e-03$	$8.183e-01$	$1.758e-01$	$2.303e-01$

Tabelle 7.7: Numerische Resultate der Rekonstruktion von Muster Abb. 7.5 mit vier Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	1.1	1.4	2.9
$\gamma$	1.1	0.4	2.0
$\theta$	0.6	0.7	0.8

Tabelle 7.8: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.5.

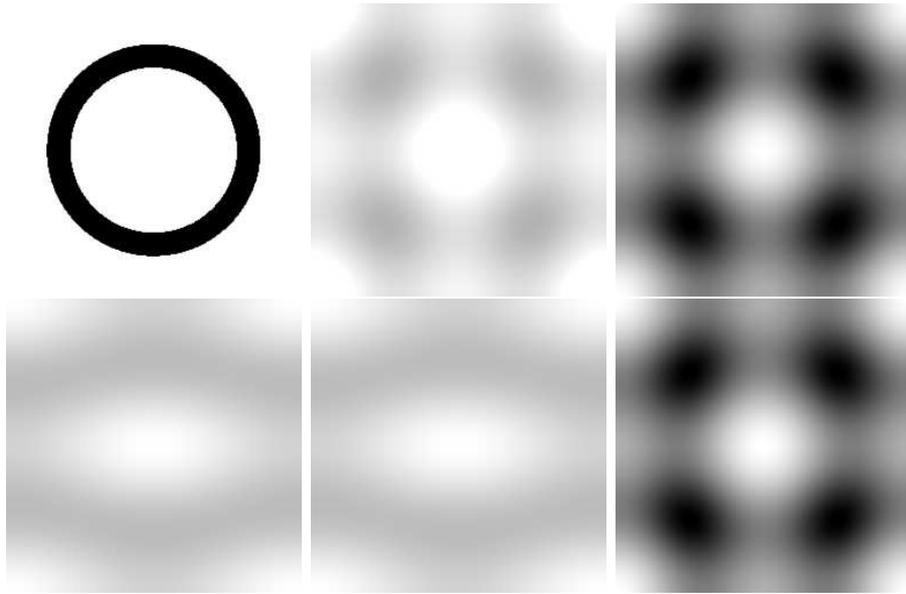


Abbildung 7.6: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.9 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$3.555e-07$	$1.405e-01$	$2.838e-01$	$7.543e-04$	$4.539e-01$
Simplexverfahren	$3.554e-07$	$1.403e-01$	$2.839e-01$	$1.871e-04$	$4.538e-01$
Simplexverfahren (mit NB)	$3.652e-03$	$1.126e-01$	$8.444e-01$	$8.547e-02$	$1.609e-01$
polares Simplexverfahren	$3.652e-03$	$1.126e-01$	$8.445e-01$	$8.547e-02$	$1.610e-01$

Tabelle 7.9: Numerische Resultate der Rekonstruktion von Muster Abb. 7.6 mit vier Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.9	0.8	1.1
$\gamma$	2.0	0.6	2.7
$\theta$	0.7	0.4	0.5

Tabelle 7.10: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.6.

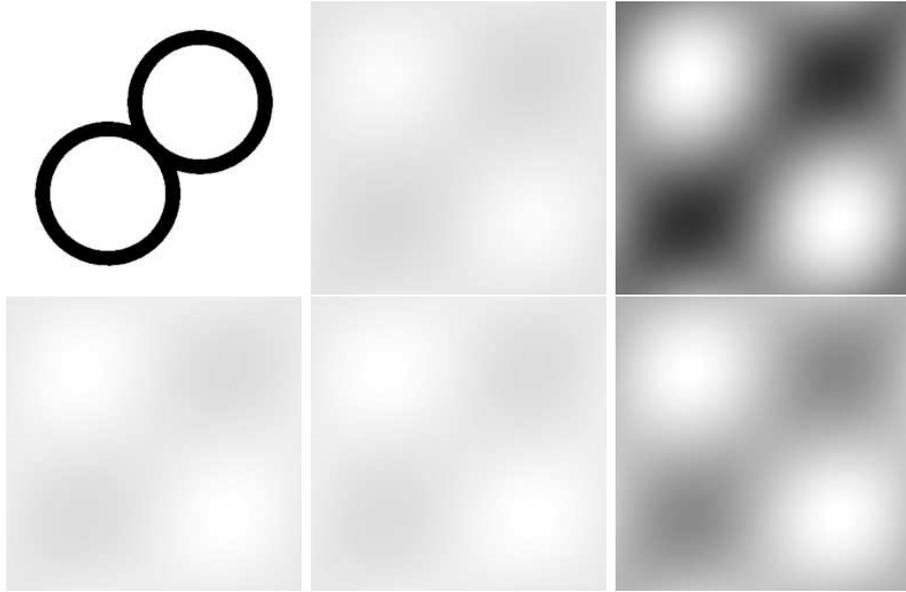


Abbildung 7.7: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.11 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$1.043e-05$	$4.877e-02$	$1.154e-01$	$4.367e-03$	$6.454e-01$
Simplexverfahren	$2.764e-07$	$4.871e-02$	$2.293e-01$	$2.334e-05$	$5.259e-01$
Simplexverfahren (mit NB)	$3.256e-06$	$4.869e-02$	$8.674e-01$	$2.628e-03$	$1.160e-01$
polares Simplexverfahren	$4.792e-06$	$4.869e-02$	$8.674e-01$	$2.703e-03$	$1.160e-01$

Tabelle 7.11: Numerische Resultate der Rekonstruktion von Muster Abb. 7.7 mit vier Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.9	1.4	0.8
$\gamma$	1.1	2.4	1.3
$\theta$	0.4	0.7	0.5

Tabelle 7.12: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.7.

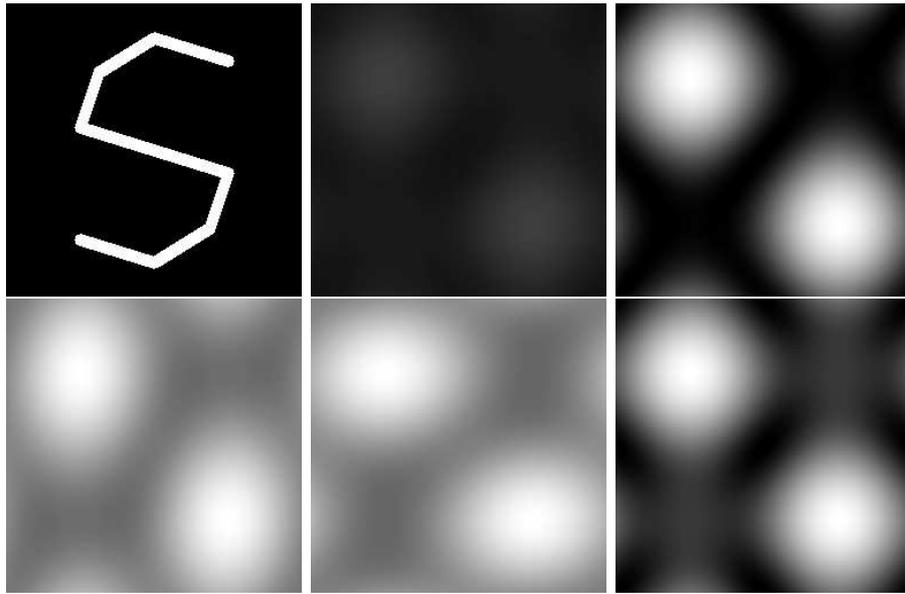


Abbildung 7.8: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.13 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$9.681e-06$	$2.950e-02$	$3.998e-02$	$4.267e-03$	$2.135e-02$
Simplexverfahren	$3.235e-06$	$2.965e-02$	$4.430e-02$	$2.395e-03$	$2.696e-02$
Simplexverfahren (mit NB)	$3.122e-05$	$2.930e-02$	$8.306e-02$	$7.896e-03$	$7.210e-02$
polares Simplexverfahren	$5.397e-05$	$2.886e-02$	$8.301e-02$	$1.032e-02$	$7.251e-02$

Tabelle 7.13: Numerische Resultate der Rekonstruktion von Muster Abb. 7.8 mit vier Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.9	0.9	1.1
$\gamma$	2.5	0.5	1.1
$\theta$	0.7	0.3	0.5

Tabelle 7.14: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.8.

Abbildung	7.2	7.3	7.4	7.5	7.6	7.7	7.8
$\ f\ _{L_2(\Omega)}$	2.487e+02	2.475e+02	2.493e+02	2.315e+02	2.339e+02	2.377e+02	7.142e+01
$\ \hat{f}\ _{\ell_2}$	9.715e-01	9.669e-01	9.740e-01	9.046e-01	9.137e-01	9.287e-01	2.790e-01
$\ N(\hat{f}_d)\ _{\ell_2}$	7.781e-02	5.010e-02	4.325e-02	1.769e-01	1.403e-01	4.871e-02	2.975e-02
$\ \hat{f}_d\ _{\ell_2}$	9.470e-01	9.363e-01	9.497e-01	8.371e-01	8.486e-01	8.674e-01	8.332e-02
$E_{\text{approx}}$	2.167e-01	2.415e-01	2.164e-01	3.427e-01	3.387e-01	3.317e-01	2.665e-01

Tabelle 7.15: Die mit vier Strahlen zu rekonstruierenden Muster mit den jeweiligen Musterabhängigen Größen.

Abbildung	7.2	7.3	7.4	7.5	7.6	7.7	7.8
Koordinaten- abstiegsverfahren	2.17e-01	2.42e-01	2.17e-01	3.47e-01	3.39e-01	3.32e-01	2.66e-01
Simplex- verfahren	2.17e-01	2.41e-01	2.16e-01	3.57e-01	3.39e-01	3.32e-01	2.66e-01
Simplexverfahren (mit NB)	2.17e-01	2.43e-01	2.17e-01	3.46e-01	3.49e-01	3.32e-01	2.66e-01
polares Simplexverfahren	2.17e-01	2.43e-01	2.17e-01	3.85e-01	3.49e-01	3.32e-01	2.66e-01

Tabelle 7.16: Rekonstruktionsgüte  $\|N(\hat{f}_* - \hat{f})\|_{\ell_2}$  der Rekonstruktion mit vier Strahlen.

In Tabelle 7.15 sind die „musterabhängigen“ Größen angegeben, d.h. die Normen und Fehler, welche bei einem fest vorgegebenen Strahlensetup für das jeweilige Muster konstant sind und a priori berechnet werden können. Die Tabellen 7.1, 7.3, 7.5, 7.7, 7.9, 7.11 und 7.13 listen für die entsprechenden Muster alle berechneten Fehler und Normen der Rekonstruktionen für das jeweilige Verfahren auf. Diese Werte sind von den Minimierungsverfahren und deren Güte abhängig.

Zusätzlich sind in den Tabellen 7.2, 7.4, 7.6, 7.8, 7.10, 7.12 und 7.14 die Parameter  $\beta$ ,  $\gamma$  und  $\theta$  aufgelistet, welche in den Varianten der Nelder und Mead Simplexmethode verwendet wurden. Dies ist notwendig, da die Ergebnisse für jedes Muster erheblich von der Wahl dieser Parameter abhängig sind wie die Praxis gezeigt hat. Für jede Kombination von  $\beta, \gamma$  und  $\theta$  wurden das jeweilige Verfahren und Muster mit  $N_0 = 1000$  getestet. Die aufgelisteten Parameter lieferten das beste Resultat, d.h. den kleinsten Wert des Kostenfunktional und wurden anschließend mit  $N_0 = 10000$  zur Rekonstruktion verwendet. Dabei wurden, um den numerischen Aufwand nicht zu groß zu gestalten,  $\beta, \gamma \in [0, 3]$  und  $\theta \in (0, 1)$  mit der Genauigkeit von einer Nachkommastelle getestet.

Zusammenfassend ist in Tabelle 7.16 die Rekonstruktionsgüte  $\|N(\widehat{f}_* - \widehat{f})\|_{\ell_2}$  für alle Muster und Verfahren berechnet worden. Diese Größe wurde zur Übersichtlichkeit noch einmal als Gesamtmaß angegeben, ergibt sich aber auch aus dem Zusammenhang

$$\|N(\widehat{f}_* - \widehat{f})\|_{\ell_2}^2 = E(N(f_*))^2 + (E_{\text{approx}})^2.$$

Die Resultate zeigen, dass dieses Strahlensetup nur zur Rekonstruktion einfacherer Strukturen ausreichend ist. Abbildungen 7.2 und 7.6 weisen noch Ähnlichkeiten zu ihren Rekonstruktionen auf, während einfachere Kombinationen dieser Muster (Abbildungen 7.3, 7.4 und 7.5) nur noch schwer zu erkennen sind. Die Abbildungen 7.8 und 7.7 können mit diesem Strahlensetup nicht ausreichend rekonstruiert werden. Dies liegt allerdings an dem gewählten Strahlensetup wie ein Vergleich von  $f$  und  $f_d$  zeigt.

Zur Erinnerung: Die Approximationsgüte von  $f_d \in \mathcal{S}_d$  an  $f$  ist vom (nun festgelegten) Strahlensetup abhängig und man *berechnet* eine Approximation  $f_* \in \mathcal{S}_N$  an  $f_d$ .

Aufgrund der geringen Anzahl von Freiheitsgraden ( $N = 4$ ) hat  $\mathcal{J}_c$  allerdings noch eine recht einfache Gestalt. In allen Fällen (bis auf Abb. 7.5) liegt das errechnete Minimum sowohl für Koordinatenabstiegsverfahren als auch Simplexverfahren nahe bei 0. Dies gilt auch für das Simplexverfahren (mit NB) und das polare Simplexverfahren, Abbildungen 7.5 und 7.6 bilden hier die Ausnahmen. Koordinatenabstiegs- und Simplexverfahren liefern im Wesentlichen die gleichen Ergebnisse.  $\|f_*\|_{\ell_2}$  und  $E(f_*)$  weichen zwar teilweise voneinander ab, was allerdings am minimierten Funktional  $\mathcal{J}_c$  liegt (vgl. Bemerkung 4.3).

Die Originalmuster in Abb. 7.3 und Abb. 7.4 unterscheiden sich nur unwesentlich, die Rekonstruktionen weisen jedoch einige Unterschiede auf. Die Güte der Rekonstruktion scheint also zusätzlich vom Muster  $f$  abzuhängen.

Auch das Simplexverfahren (mit NB) und das polare Simplexverfahren liefern für Abb. 7.2, 7.6, 7.7, 7.8 fast gleiche Resultate. In Abb. 7.5 sind die großen Abweichungen zwischen  $\|N(\widehat{f}_*)\|_{\ell_2}$  bzw.  $E(N(f_*))$ , bei diesen beiden Verfahren und das sehr „helle“ Muster (es ist nicht 1!) des polaren Simplexverfahren durch eine zu hohe Gewichtung der Nebenbedingung zu erklären. Das polare Simplexverfahren erfüllt die Nebenbedingung exakt und ist daher in seiner Suche nach einem Minimum auf einen Unterraum eingeschränkt ( $\mathcal{J}_c(f_*)$  und  $E(N(f_*))$  groß). Das Simplexverfahren (mit NB) erfüllt die Nebenbedingung lediglich approximativ und kann somit über dem ganzen Raum  $\mathbb{C}^N$  minimieren ( $\mathcal{J}_c(f_*)$  und  $E(N(f_*))$  kleiner).

Somit bestätigen die Resultate obige Erwartungen. Zu einer detaillierteren Rekonstruktion sind also mehr als vier Strahlen und somit mehr Differenzen in  $\mathcal{K}_d$  notwendig.

### Bemerkung 7.1

*Die letzte Abb. 7.8 unterscheidet sich von den übrigen Abbildungen, abgesehen von der anderen Form des Musters, zusätzlich wesentlich dadurch, dass der Hintergrund des Originalbildes schwarz ist. Hier tritt somit der Fall ein, dass  $f = 1 - \chi_{\Omega_M}$  ist. Aufgrund der Linearität der Fouriertransformation ändert sich durch die Konstante 1 betragsmäßig nur der Wert  $\widehat{f}(0)$ , während alle anderen Werte der Fouriertransformierten lediglich ihr Vorzeichen wechseln.*

*Das Funktional ist per Konstruktion von dem gegebenen Muster abhängig, d.h. man hat die zusätzliche Abhängigkeit  $\mathcal{J}_c(\mathbf{c}, f)$  von  $f$ . Dies wird nur an dieser Stelle explizit erwähnt, da das Muster sonst immer als fest gegeben vorausgesetzt wird.*

*Hier bewirkt die Änderung des Musters allerdings eine wesentliche Änderung des Funktionals  $\mathcal{J}_c$ , welche nicht durch eine einfache Transformation des Koeffizientenvektors  $\mathbf{c}$  kompensiert werden kann, obwohl sich die Muster in ihrer Struktur nicht unterscheiden. Das Minimum von  $\mathcal{J}_c(\cdot, 1 - \chi_{\Omega_M})$*

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$6.048e-03$	$9.854e-02$	$1.437e+00$	$1.100e-01$	$5.543e-01$
Simplexverfahren	$5.789e-03$	$1.034e-01$	$1.535e+01$	$1.076e-01$	$1.446e+01$
Simplexverfahren (mit NB)	$5.839e-03$	$9.916e-02$	$9.492e-01$	$1.081e-01$	$1.205e-01$
polares Simplexverfahren	$9.224e-03$	$3.096e-02$	$9.443e-01$	$1.358e-01$	$1.458e-01$

Tabelle 7.17: Numerische Resultate der Rekonstruktion von Muster Abb. 7.9 mit acht Strahlen und den beschriebenen Verfahren.

ergibt sich nicht durch eine einfache Drehung, Spiegelung oder Translation aus dem Minimum von  $\mathcal{J}_c(\cdot, \chi_{\Omega_M})$ .

Es macht also einen wesentlichen Unterschied, ob man ein schwarzes Muster auf weißem Hintergrund oder ein weißes auf schwarzem Hintergrund rekonstruiert.

### 7.1.2 Die Rekonstruktion mit acht Strahlen

In diesem Abschnitt soll das in (2.2) schon vorgestellte Strahlensetup mit  $N = 8$  untersucht werden. Dieser Wert für  $N$  ergibt sich durch die Wahl  $R = 10$  in Abschnitt 3.1. Damit erhält man wieder  $\#\mathcal{I}(R) = 1$ , da  $(1, 3)$  das einzige gültige Tupel ist. Allerdings ist hier  $\mathcal{N}(\mathcal{I}(R)) = 0$ , woraus  $N = 8$  folgt. Weiter kann man leicht nachrechnen, dass  $r = \frac{2K}{\sqrt{10}}$  und  $r' = \frac{K}{\sqrt{10}}$  ist, was zu

$$\mathcal{K} := \left\{ r \begin{pmatrix} 1, 5 \\ 0, 5 \end{pmatrix}, r \begin{pmatrix} 0, 5 \\ 1, 5 \end{pmatrix}, r \begin{pmatrix} -0, 5 \\ 1, 5 \end{pmatrix}, r \begin{pmatrix} -1, 5 \\ 0, 5 \end{pmatrix}, \right. \\ \left. r \begin{pmatrix} -1, 5 \\ -0, 5 \end{pmatrix}, r \begin{pmatrix} -0, 5 \\ -1, 5 \end{pmatrix}, r \begin{pmatrix} 0, 5 \\ -1, 5 \end{pmatrix}, r \begin{pmatrix} 1, 5 \\ -0, 5 \end{pmatrix} : r = \frac{2K}{\sqrt{10}} \right\} \quad (7.2)$$

führt. Abb. 2.2 und Abb. 3.2 zeigen die Nummerierung der acht  $\mathbf{k}_j$  und die Menge  $\mathcal{K}_d$ . Wiederum liegen die Differenzen nah am Ursprung. In  $\mathcal{K}_d$  liegen allerdings auch Differenzen mit „größerem“ Betrag, wodurch in  $f_d$  auch „höhere“ Frequenzen einspielen (vgl.(3.18)). Es ist also eine deutliche Verbesserung der Rekonstruktionen aus Abschnitt 7.1.1 zu erwarten. Es wurden die gleichen Originalmuster wie in Abschnitt 7.1.1 rekonstruiert. Die Ergebnisse der Berechnungen sind in der gleichen Weise dargestellt wie im vorigen Abschnitt, d.h. man hat wieder sechs Abbildungen und zwei Tabellen pro Muster. Die a priori berechenbaren Werte sind in Tabelle 7.31 zu finden.

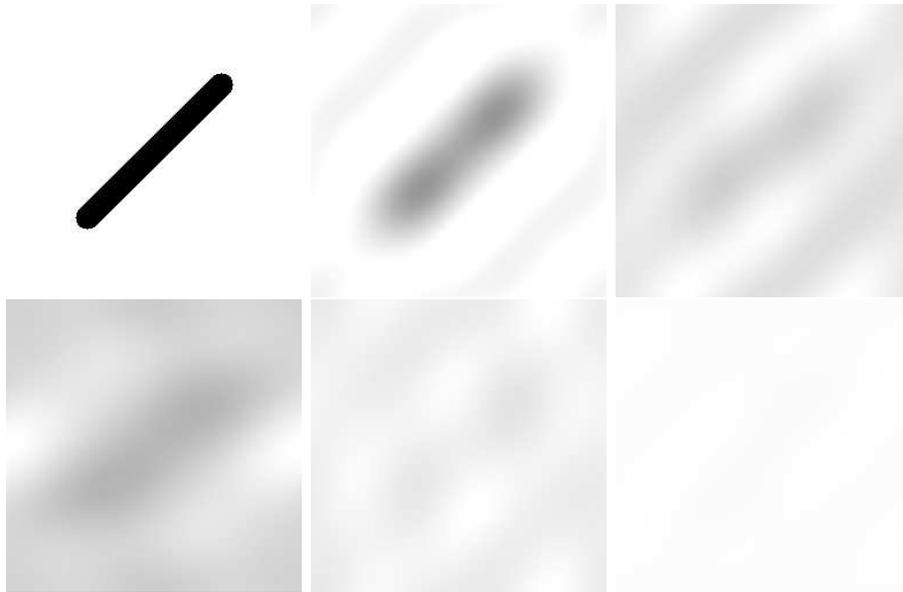


Abbildung 7.9: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.17 zu finden.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.9	1.0	1.7
$\gamma$	0.6	0.1	1.2
$\theta$	0.3	0.6	0.4

Tabelle 7.18: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.9.

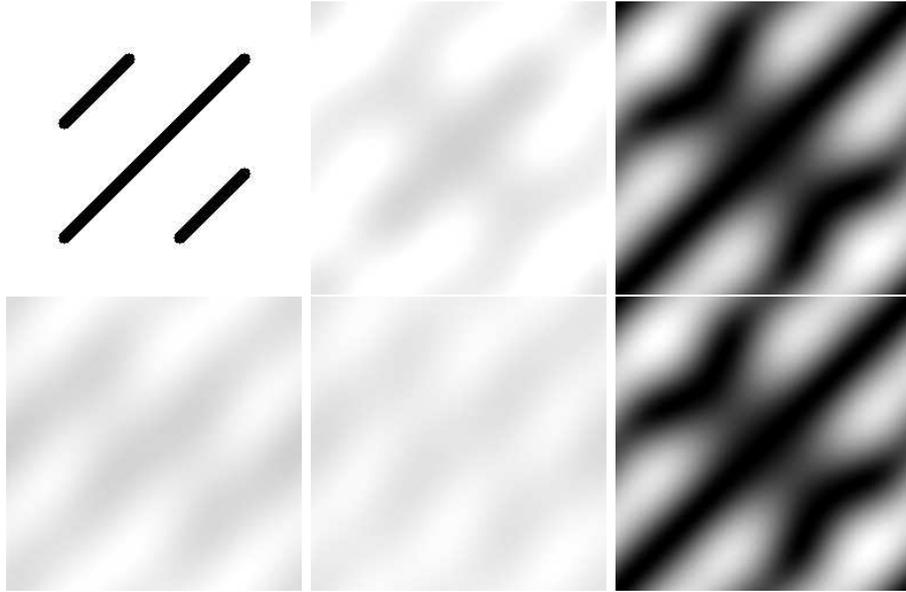


Abbildung 7.10: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.19 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$7.163e-04$	$7.754e-02$	$1.250e-01$	$3.783e-02$	$7.770e-01$
Simplexverfahren	$7.159e-04$	$7.778e-02$	$1.253e-01$	$3.781e-02$	$7.768e-01$
Simplexverfahren (mit NB)	$1.056e-03$	$7.368e-02$	$9.378e-01$	$4.595e-02$	$7.621e-02$
polares Simplexverfahren	$1.601e-03$	$4.605e-02$	$9.361e-01$	$5.658e-02$	$8.307e-02$

Tabelle 7.19: Numerische Resultate der Rekonstruktion von Muster Abb. 7.10 mit acht Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	1.1	1.1	0.8
$\gamma$	0.3	0.3	0.9
$\theta$	0.6	0.6	0.3

Tabelle 7.20: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.10.

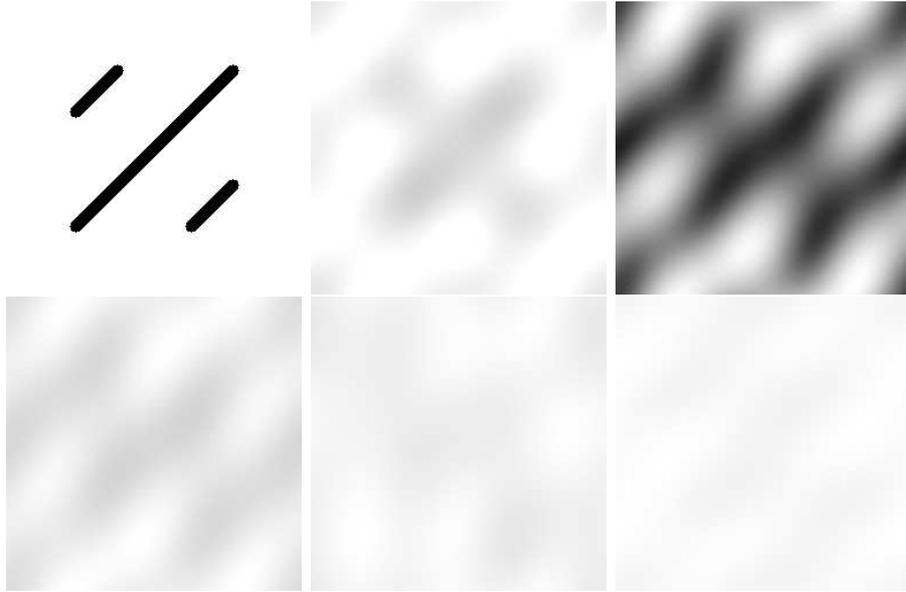


Abbildung 7.11: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.21 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$1.426e-03$	$6.475e-02$	$1.447e-01$	$5.341e-02$	$7.725e-01$
Simplexverfahren	$1.325e-03$	$6.779e-02$	$3.421e+00$	$5.149e-02$	$2.521e+00$
Simplexverfahren (mit NB)	$1.334e-03$	$6.697e-02$	$9.511e-01$	$5.167e-02$	$7.099e-02$
polares Simplexverfahren	$2.880e-03$	$3.260e-02$	$9.492e-01$	$7.590e-02$	$9.017e-02$

Tabelle 7.21: Numerische Resultate der Rekonstruktion von Muster Abb. 7.11 mit acht Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	1.7	1.2	0.4
$\gamma$	2.9	0.0	0.6
$\theta$	0.8	0.7	0.8

Tabelle 7.22: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.11.

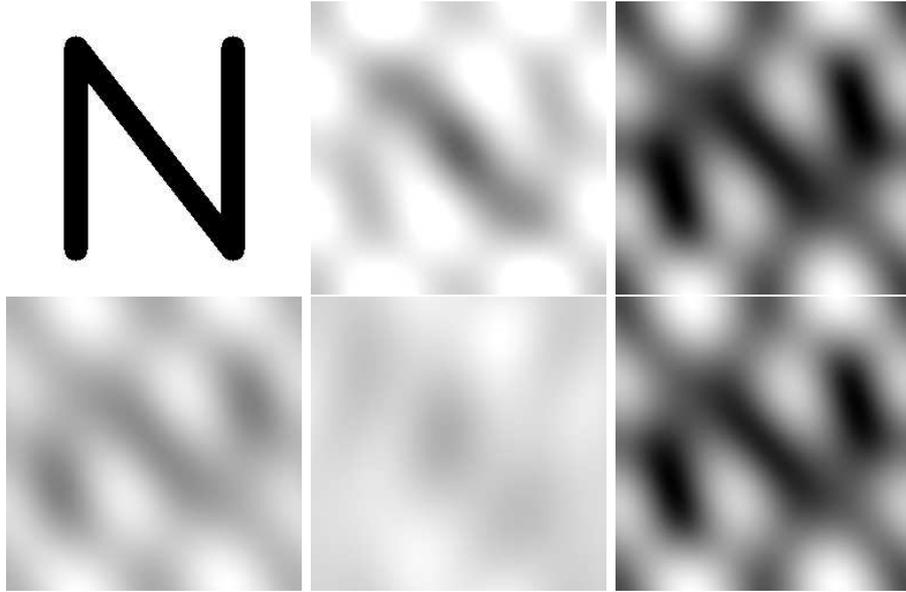


Abbildung 7.12: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.23 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$4.739e-03$	$1.634e-01$	$3.312e-01$	$9.735e-02$	$3.937e-01$
Simplexverfahren	$4.731e-03$	$1.633e-01$	$3.506e-01$	$9.727e-02$	$3.722e-01$
Simplexverfahren (mit NB)	$4.792e-03$	$1.629e-01$	$8.343e-01$	$9.790e-02$	$1.781e-01$
polares Simplexverfahren	$1.251e-02$	$8.416e-02$	$8.226e-01$	$1.582e-01$	$2.171e-01$

Tabelle 7.23: Numerische Resultate der Rekonstruktion von Muster Abb. 7.12 mit acht Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	1.1	1.2	1.4
$\gamma$	0.2	1.5	0.9
$\theta$	0.6	0.6	0.1

Tabelle 7.24: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.12.

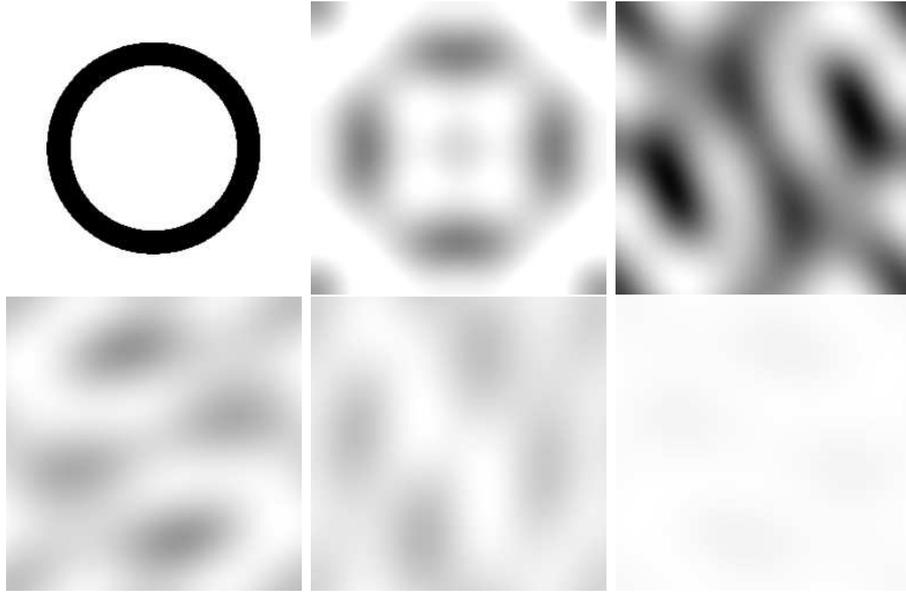


Abbildung 7.13: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.25 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$1.326e-02$	$1.241e-01$	$3.073e-01$	$1.628e-01$	$4.498e-01$
Simplexverfahren	$1.157e-02$	$1.319e-01$	$7.869e+00$	$1.521e-01$	$7.169e+00$
Simplexverfahren (mit NB)	$1.214e-02$	$1.242e-01$	$8.465e-01$	$1.558e-01$	$2.074e-01$
polares Simplexverfahren	$1.501e-02$	$8.103e-02$	$8.409e-01$	$1.732e-01$	$2.205e-01$

Tabelle 7.25: Numerische Resultate der Rekonstruktion von Muster Abb. 7.13 mit acht Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.7	1.1	0.7
$\gamma$	0.2	0.0	1.2
$\theta$	0.6	0.8	0.3

Tabelle 7.26: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.13.

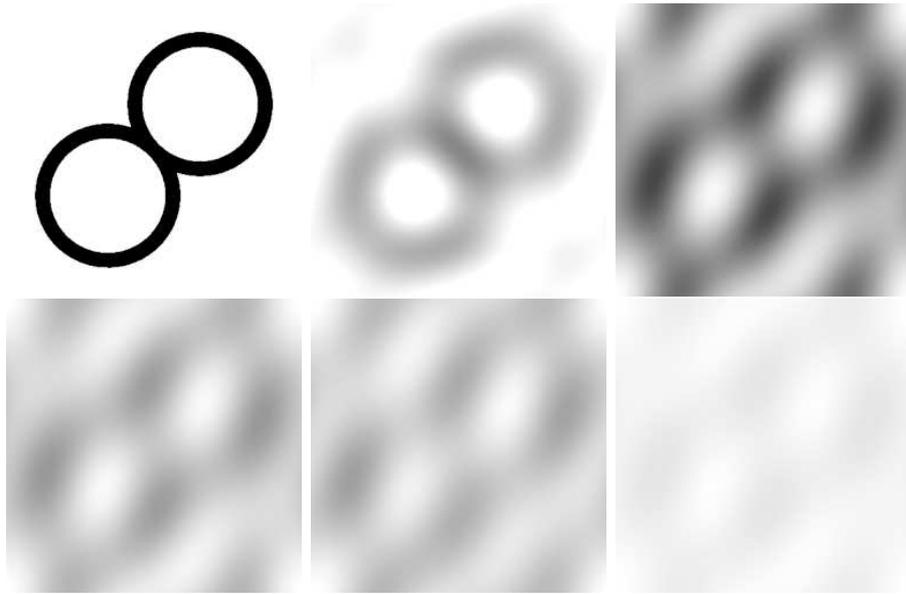


Abbildung 7.14: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.27 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$8.809e-03$	$1.384e-01$	$4.328e-01$	$1.327e-01$	$3.649e-01$
Simplexverfahren	$8.653e-03$	$1.374e-01$	$3.718e+00$	$1.315e-01$	$2.968e+00$
Simplexverfahren (mit NB)	$8.686e-03$	$1.369e-01$	$8.768e-01$	$1.318e-01$	$1.756e-01$
polares Simplexverfahren	$9.930e-03$	$1.127e-01$	$8.733e-01$	$1.409e-01$	$1.825e-01$

Tabelle 7.27: Numerische Resultate der Rekonstruktion von Muster Abb. 7.14 mit acht Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	1.3	1.2	0.6
$\gamma$	0.2	2.9	0.8
$\theta$	0.6	0.8	0.5

Tabelle 7.28: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.14.

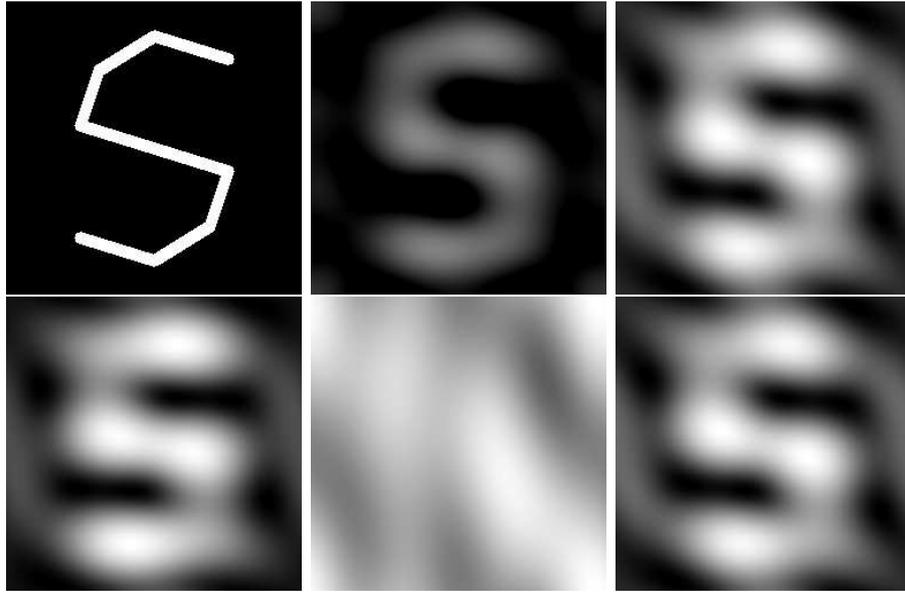


Abbildung 7.15: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.29 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$8.711e-04$	$1.250e-01$	$2.041e-01$	$4.173e-02$	$1.608e-01$
Simplexverfahren	$8.711e-04$	$1.250e-01$	$2.041e-01$	$4.173e-04$	$1.608e-01$
Simplexverfahren (mit NB)	$2.291e-03$	$8.147e-02$	$1.256e-01$	$6.286e-02$	$1.094e-01$
polares Simplexverfahren	$8.510e-03$	$2.235e-02$	$8.098e-02$	$1.305e-01$	$1.489e-01$

Tabelle 7.29: Numerische Resultate der Rekonstruktion von Muster Abb. 7.15 mit acht Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	1.1	1.0	2.7
$\gamma$	0.1	0.0	1.9
$\theta$	0.7	0.6	0.9

Tabelle 7.30: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.15.

Abbildung	7.9	7.10	7.11	7.12	7.13	7.14	7.15
$\ f\ _{L_2(\Omega)}$	2.487e+02	2.475e+02	2.493e+02	2.316e+02	2.339e+02	2.377e+02	7.142e+01
$\ \hat{f}\ _{\ell_2}$	9.715e-01	9.669e-01	9.740e-01	9.046e-01	9.137e-01	9.287e-01	2.790e-01
$\ N(\hat{f}_d)\ _{\ell_2}$	1.494e-01	8.649e-02	8.473e-02	1.901e-01	2.022e-01	1.904e-01	1.318e-01
$\ \hat{f}_d\ _{\ell_2}$	9.556e-01	9.389e-01	9.525e-01	8.400e-01	8.610e-01	8.867e-01	1.531e-01
$E_{\text{approx}}$	1.752e-01	2.309e-01	2.037e-01	3.355e-01	3.058e-01	2.759e-01	2.333e-01

Tabelle 7.31: Musterabhängige Normen und Fehler der Rekonstruktion mit acht Strahlen.

Abbildung	7.9	7.10	7.11	7.12	7.13	7.14	7.15
Koordinatenabstiegsverfahren	2.07e-01	2.34e-01	2.11e-01	3.49e-01	3.46e-01	3.06e-01	2.37e-01
Simplexverfahren	2.06e-01	2.35e-01	2.10e-01	3.49e-01	3.42e-01	3.06e-01	2.37e-01
Simplexverfahren (mit NB)	2.06e-01	2.35e-01	2.10e-01	3.50e-01	3.43e-01	3.06e-01	2.46e-01
polares Simplexverfahren	2.22e-01	2.38e-01	2.17e-01	3.71e-01	3.51e-01	3.10e-01	2.67e-01

Tabelle 7.32: Rekonstruktionsgüte  $\|N(\hat{f}_* - \hat{f})\|_{\ell_2}$  der Rekonstruktion mit acht Strahlen.

Wie in Abschnitt 7.1.1 sind in Tabelle 7.15 die Normen und Fehler angegeben, welche nicht von den numerischen Minimierungsverfahren, sondern von den Mustern abhängen. Die numerischen Ergebnisse sind in den Tabellen 7.17, 7.19, 7.21, 7.23, 7.25, 7.27 und 7.29 aufgelistet.

In den Tabellen 7.18, 7.20, 7.22, 7.24, 7.26, 7.28 und 7.30 sind wiederum die Parameter der Nelder und Mead Simplexmethode aufgelistet, welche in Tests die besten Ergebnisse lieferten. Da der Rechenaufwand der Simplexmethode durch die größere Dimension  $N$  steigt, wurden die Tests hier mit  $N_0 = 100$  durchgeführt. Alle anderen Parameter der Tests blieben gleich.

Tabelle 7.32 listet die Rekonstruktionsgüte  $\|N(\hat{f}_* - \hat{f})\|_{\ell_2}$  für alle Muster und Verfahren auf.

Die Resultate in diesem Abschnitt weisen schon erheblich mehr Ähnlichkeiten zu den Originalmustern auf. Herauszuheben sind hier wohl Abb. 7.13 im negativen und Abb. 7.15 im positiven Sinne. Diese Verbesserung ist ausschließlich auf das umfangreichere Strahlensetup zurückzuführen, was die Muster  $f_d$  für alle Abbildungen beweisen. Hierdurch wird allerdings auch das Funktional  $\mathcal{J}_c$  „komplizierter“ in dem Sinne, dass die berechneten Minima größer sind.

Auch hier liefern Koordinatenabstiegsverfahren und Simplexverfahren fast die gleichen Resultate.

Es weichen zwar in den Abb. 7.9, 7.11, 7.13 und 7.14 die Werte von  $\|\hat{f}_*\|_{\ell_2}$  und  $E(f_*)$  teilweise sogar erheblich voneinander ab, dies ist allerdings wieder darauf zurückzuführen, dass die Nebenbedingung beim Koordinatenabstiegs- sowie Simplexverfahren nicht berücksichtigt und beim Simplexverfahren (mit NB) sowie polaren Simplexverfahren berücksichtigt wurde.

Die Ergebnisse von Simplexverfahren (mit NB) und polarem Simplexverfahren gehen in diesem Abschnitt auseinander. Zwar liegen die Werte von  $\|\hat{f}_*\|_{\ell_2}$  und  $E(f_*)$  noch „relativ“ nah beieinander, doch tritt hier der Effekt der zu hohen Gewichtung der Nebenbedingung wie in Abb. 7.12 fast überall auf, wie der Vergleich von  $\|N(\hat{f}_*)\|_{\ell_2}$  und  $E(N(f_*))$  für die beiden Verfahren zeigt. Ausnahmen bilden hier noch die Muster in Abb. 7.10, 7.13 bzw. 7.14. Somit ist zur Berücksichtigung der Nebenbedingung hier das Simplexverfahren (mit NB) dem polaren Simplexverfahren vorzuziehen. Es ist aber insgesamt eine deutliche Verbesserung zur Rekonstruktion mittels vier Strahlen erreicht worden.

### 7.1.3 Die Rekonstruktion mit zwölf Strahlen

Die Verbesserung der Rekonstruktion durch Hinzunahme von Elementen in  $\mathcal{K}$  aus dem vorherigen Abschnitt motiviert eine weitere Vergrößerung dieser Menge, d.h. die Verwendung von mehr Strahlen. Daher soll nun Algorithmus 4.16 mit einem  $N > 8$  durchgeführt werden. Wieder wird  $\mathcal{K}$  durch die in Abschnitt 3.1 beschriebene Methode konstruiert, wozu nun  $R = 50$  gewählt wird. Für diesen Wert kann man die gültigen Tupel  $(1, 7)$  und  $(5, 5)$  bestimmen, woraus folgt, dass  $\#\mathcal{I}(R) = 2$  und  $\mathcal{N}(\mathcal{I}(R)) = 1$  ist. Somit ergeben sich  $N = 12$  Strahlen und  $r = \frac{2K}{\sqrt{50}}$  sowie  $r' = \frac{K}{\sqrt{50}}$ , womit die Menge

$$\mathcal{K} := \left\{ r \begin{pmatrix} 3, 5 \\ 0, 5 \end{pmatrix}, r \begin{pmatrix} 2, 5 \\ 2, 5 \end{pmatrix}, r \begin{pmatrix} 0, 5 \\ 3, 5 \end{pmatrix}, r \begin{pmatrix} -0, 5 \\ 3, 5 \end{pmatrix}, r \begin{pmatrix} -2, 5 \\ 2, 5 \end{pmatrix}, r \begin{pmatrix} -3, 5 \\ 0, 5 \end{pmatrix}, \right. \\ \left. r \begin{pmatrix} -3, 5 \\ -0, 5 \end{pmatrix}, r \begin{pmatrix} -2, 5 \\ -2, 5 \end{pmatrix}, r \begin{pmatrix} -0, 5 \\ -3, 5 \end{pmatrix}, r \begin{pmatrix} 0, 5 \\ -3, 5 \end{pmatrix}, r \begin{pmatrix} 2, 5 \\ -2, 5 \end{pmatrix}, r \begin{pmatrix} 3, 5 \\ -0, 5 \end{pmatrix} : r = \frac{2K}{\sqrt{50}} \right\} \quad (7.3)$$

definiert wird.

In Abb. 7.16 ist die Lage und Nummerierung der  $\mathbf{k}_j \in \mathcal{K}$  und  $\mathbf{k} \in \mathcal{K}_d$  dargestellt. In diesem Fall enthält  $\mathcal{K}_d$  ebenfalls Werte die nah am Ursprung liegen sowie betragsmäßig größere Elemente, welche allerdings weiter im  $\mathbb{R}^2$  verteilt, d.h. nicht unbedingt benachbart sind. Dies bedeutet, dass die beste Approximation  $f_d$  aus  $\mathcal{S}_d$  an  $f$  sowohl „hohe“ als auch „niedrige“ Frequenzen berücksichtigt (vgl. (3.18)). Es wurden wieder die selben Muster wie in Abschnitt 7.1.1 und Abschnitt 7.1.2 rekonstruiert und die Resultate in sechs Abbildungen und zwei Tabellen pro Muster dargestellt. Die a priori Normen und Fehler sind für dieses Strahlensetup in Tabelle 7.47 aufgelistet.

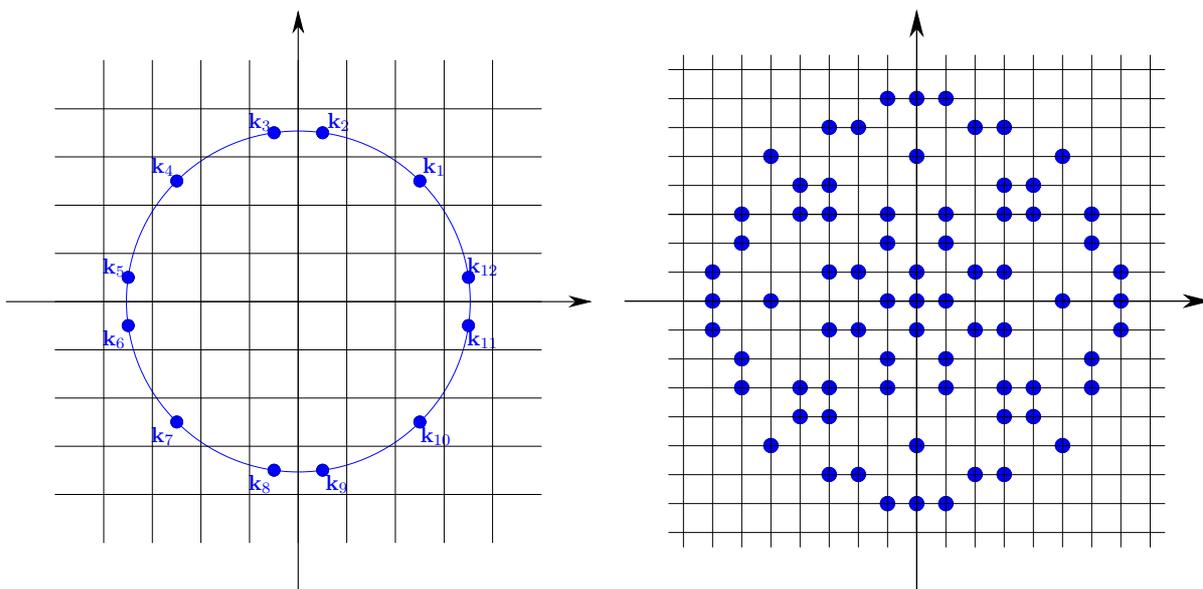


Abbildung 7.16: Lage und Anordnung der Wellenvektoren  $\mathbf{k}_j \in \mathcal{K}$  definiert in (7.3) und die resultierenden Differenzen.

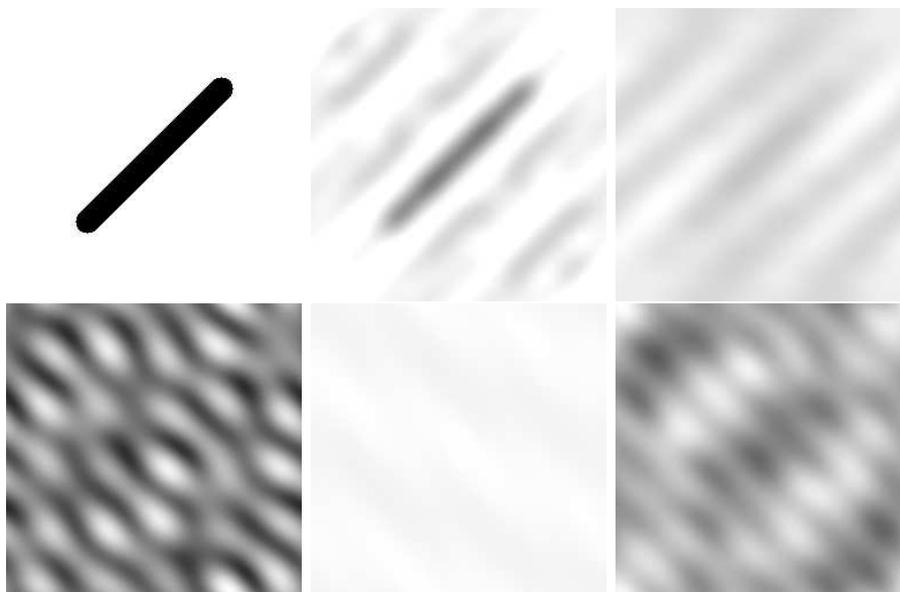


Abbildung 7.17: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.33 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$6.556e-03$	$1.028e-01$	$1.381e+00$	$1.145e-01$	$4.995e-01$
Simplexverfahren	$9.797e-03$	$6.459e-02$	$2.783e-01$	$1.400e-01$	$6.357e-01$
Simplexverfahren (mit NB)	$7.433e-02$	$3.590e-01$	$1.016e+00$	$3.855e-01$	$3.901e-01$
polares Simplexverfahren	$1.153e-02$	$1.857e-02$	$9.440e-01$	$1.518e-01$	$1.608e-01$

Tabelle 7.33: Numerische Resultate der Rekonstruktion von Muster Abb. 7.17 mit zwölf Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.9	0.4	1.9
$\gamma$	0.1	1.4	0.4
$\theta$	0.3	0.5	0.3

Tabelle 7.34: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.17.

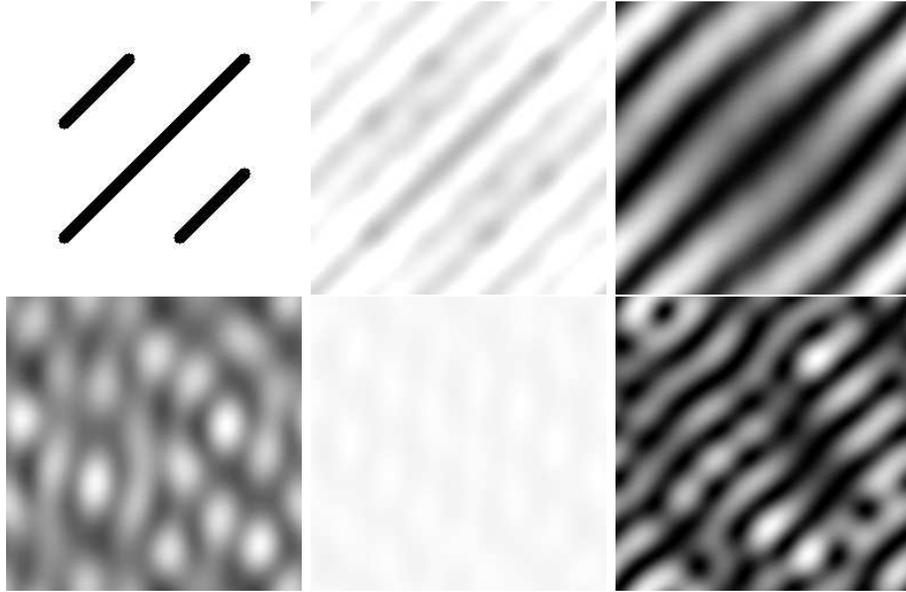


Abbildung 7.18: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.35 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$4.141e-03$	$7.751e-02$	$1.348e-01$	$9.100e-02$	$7.693e-01$
Simplexverfahren	$5.124e-03$	$5.180e-02$	$8.725e-02$	$1.012e-01$	$8.103e-01$
Simplexverfahren (mit NB)	$4.933e-02$	$2.934e-01$	$9.581e-01$	$3.124e-01$	$3.147e-01$
polares Simplexverfahren	$7.035e-03$	$1.461e-02$	$9.351e-01$	$1.186e-01$	$1.333e-01$

Tabelle 7.35: Numerische Resultate der Rekonstruktion von Muster Abb. 7.18 mit zwölf Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.8	0.4	1.5
$\gamma$	2.3	1.7	0.0
$\theta$	0.7	0.4	0.1

Tabelle 7.36: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.18.

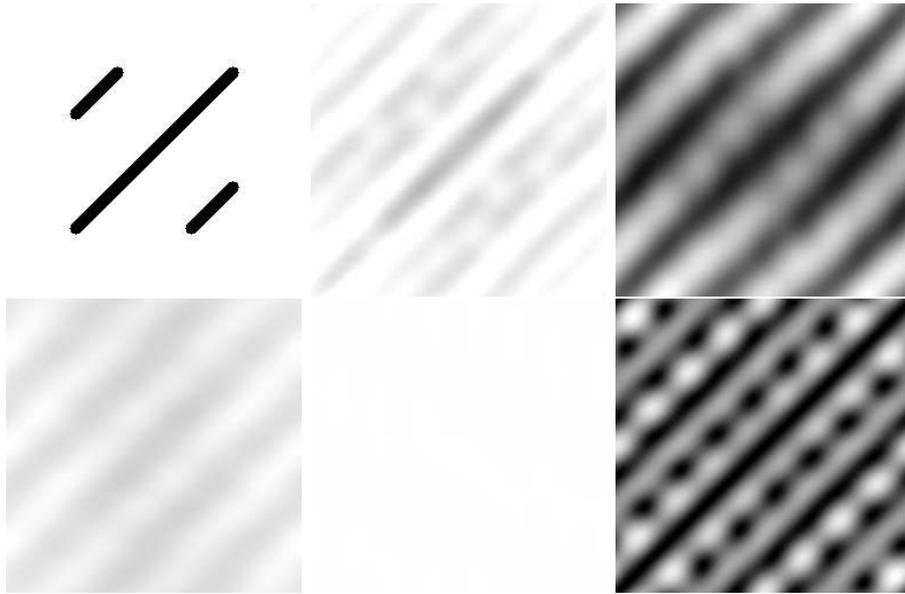


Abbildung 7.19: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.37 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$3.172e-03$	$6.248e-02$	$1.342e-01$	$7.964e-02$	$7.853e-01$
Simplexverfahren	$2.424e-03$	$7.450e-02$	$1.437e-01$	$6.963e-02$	$7.802e-01$
Simplexverfahren (mit NB)	$2.999e-03$	$6.585e-02$	$9.510e-01$	$7.744e-02$	$9.148e-02$
polares Simplexverfahren	$5.193e-03$	$1.894e-03$	$9.487e-01$	$1.019e-01$	$1.129e-01$

Tabelle 7.37: Numerische Resultate der Rekonstruktion von Muster Abb. 7.19 mit zwölf Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	1.1	2.1	0.8
$\gamma$	0.4	0.8	2.8
$\theta$	0.6	0.4	0.5

Tabelle 7.38: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.19.

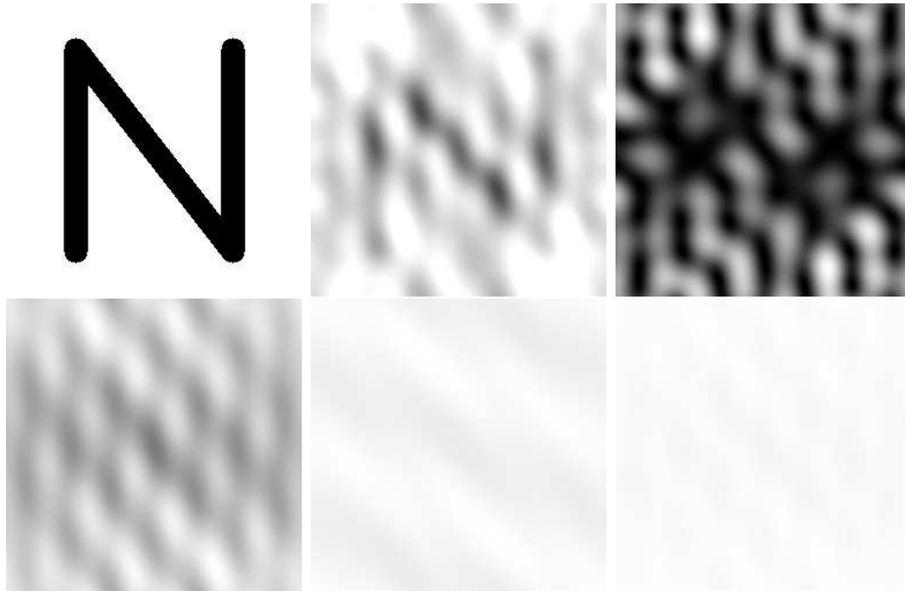


Abbildung 7.20: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.39 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$1.070e-02$	$1.229e-01$	$2.010e-01$	$1.464e-01$	$5.310e-01$
Simplexverfahren	$7.173e-03$	$1.420e-01$	$1.181e+01$	$1.198e-01$	$1.114e+01$
Simplexverfahren (mit NB)	$7.468e-03$	$1.408e-01$	$8.306e-01$	$1.222e-01$	$1.928e-01$
polares Simplexverfahren	$1.680e-02$	$2.713e-02$	$8.187e-01$	$1.833e-01$	$2.360e-01$

Tabelle 7.39: Numerische Resultate der Rekonstruktion von Muster Abb. 7.20 mit zwölf Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.7	2.7	2.4
$\gamma$	0.2	0.2	0.7
$\theta$	0.9	0.4	0.3

Tabelle 7.40: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.20.

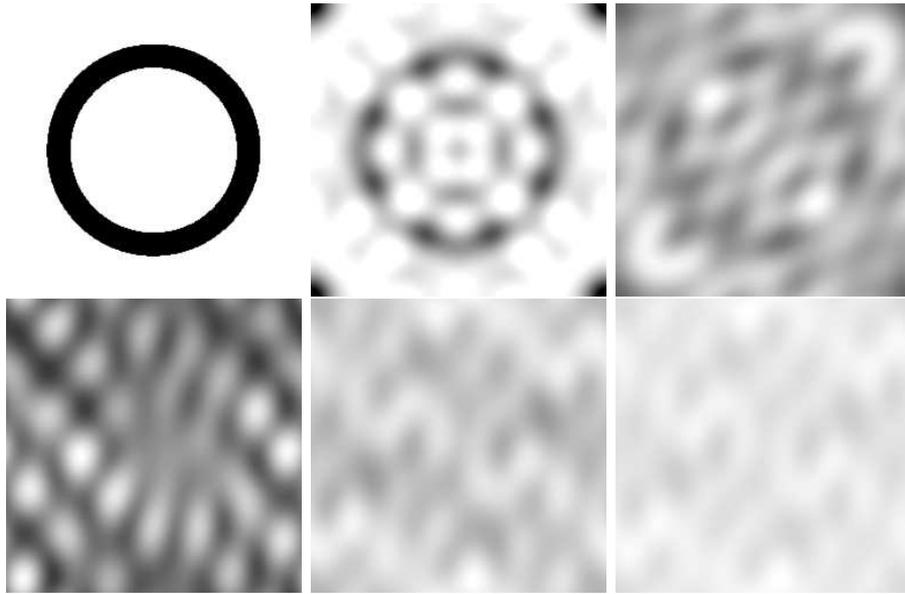


Abbildung 7.21: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.41 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$1.653e-02$	$1.450e-01$	$7.017e-01$	$1.818e-01$	$1.823e-01$
Simplexverfahren	$1.899e-02$	$1.227e-01$	$2.233e+00$	$1.949e-01$	$1.542e+00$
Simplexverfahren (mit NB)	$7.396e-02$	$2.885e-01$	$9.011e-01$	$3.839e-01$	$4.133e-01$
polares Simplexverfahren	$1.966e-02$	$1.022e-01$	$8.432e-01$	$1.983e-01$	$2.407e-01$

Tabelle 7.41: Numerische Resultate der Rekonstruktion von Muster Abb. 7.21 mit zwölf Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.8	0.4	0.9
$\gamma$	0.1	1.8	0.5
$\theta$	0.9	0.4	0.3

Tabelle 7.42: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.21.

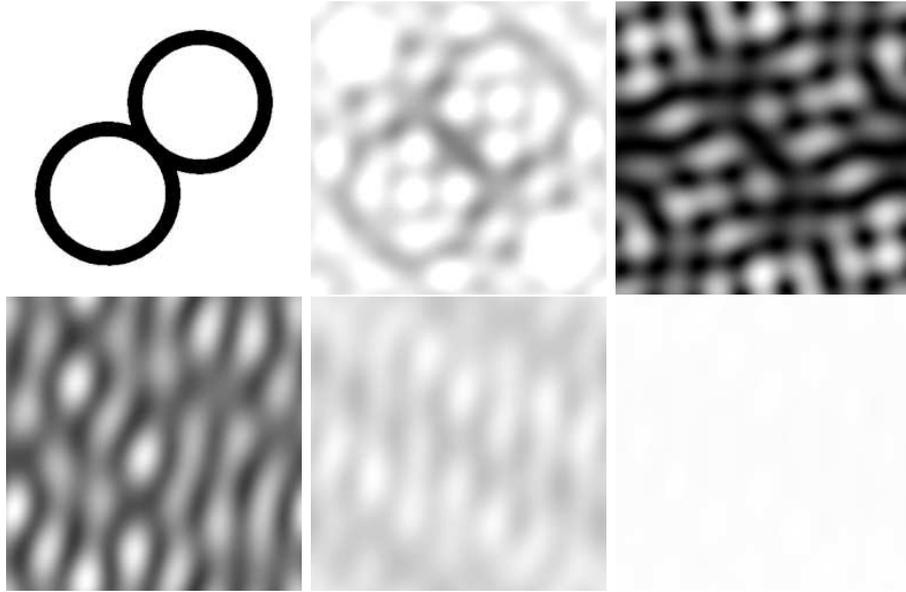


Abbildung 7.22: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.43 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$1.065e-02$	$7.21e-02$	$1.647e-01$	$1.459e-01$	$6.344e-01$
Simplexverfahren	$9.826e-03$	$1.020e-01$	$1.637e+01$	$1.402e-01$	$1.567e+01$
Simplexverfahren (mit NB)	$7.060e-02$	$9.764e-02$	$9.262e-01$	$3.758e-01$	$3.938e-01$
polares Simplexverfahren	$1.210e-02$	$3.236e-01$	$8.690e-01$	$1.556e-01$	$1.941e-01$

Tabelle 7.43: Numerische Resultate der Rekonstruktion von Muster Abb. 7.22 mit zwölf Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.9	0.4	0.8
$\gamma$	0.1	1.7	2.9
$\theta$	0.3	0.4	0.1

Tabelle 7.44: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.22.

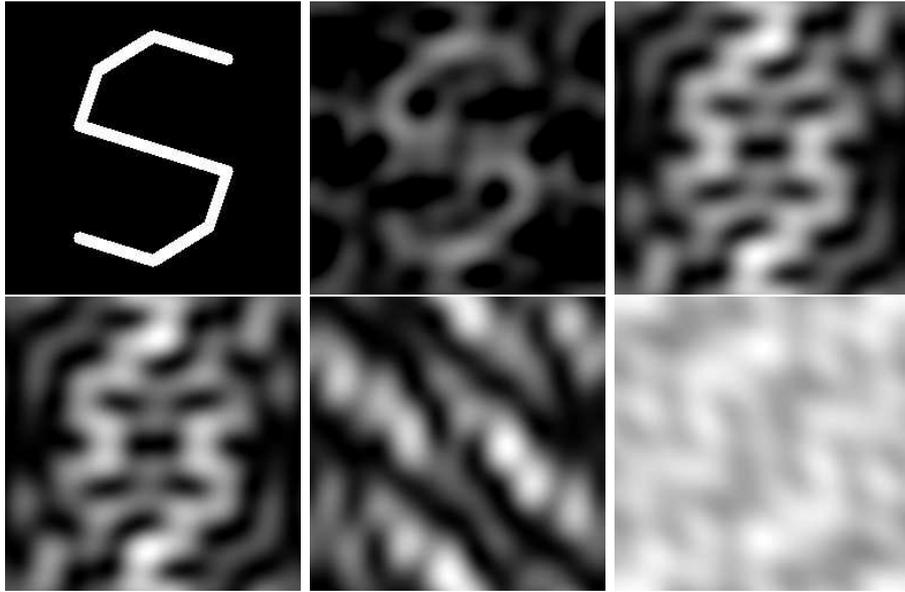


Abbildung 7.23: In der oberen Reihe sind von links die Muster  $f$  und  $f_d$  abgebildet. Das Bild rechts oben wurde mittels Koordinatenabstiegsverfahren, das darunter mit dem Simplexverfahren rekonstruiert. Links unten ist das Resultat des Simplexverfahren (mit NB) und rechts daneben des polaren Simplexverfahren dargestellt. Die zugehörigen numerischen Werte sind in Tabelle 7.45 zu finden.

Verfahren	$\mathcal{J}_c(f_*)$	$\ N(\hat{f}_*)\ _{\ell_2}$	$\ \hat{f}_*\ _{\ell_2}$	$E(N(f_*))$	$E(f_*)$
Koordinatenabstiegsverfahren	$4.008e-03$	$9.016e-02$	$1.437e-01$	$8.953e-02$	$1.386e-01$
Simplexverfahren	$5.799e-03$	$6.810e-02$	$4.214e-01$	$1.077e-01$	$4.237e-01$
Simplexverfahren (mit NB)	$4.255e-03$	$7.184e-02$	$1.116e-01$	$9.163e-02$	$1.212e-01$
polares Simplexverfahren	$6.939e-03$	$6.534e-02$	$1.016e-01$	$1.178e-01$	$1.379e-01$

Tabelle 7.45: Numerische Resultate der Rekonstruktion von Muster Abb. 7.23 mit zwölf Strahlen und den beschriebenen Verfahren.

Parameter	Simplexverfahren	Simplexverfahren (mit NB)	polares Simplexverfahren
$\beta$	0.9	2.9	2.4
$\gamma$	0.1	2.9	1.9
$\theta$	0.3	0.3	0.6

Tabelle 7.46: Parameter für die verschiedenen Formen der Nelder und Mead Simplexmethode zur Rekonstruktion des Musters aus Abb. 7.23.

Abbildung	7.17	7.18	7.19	7.20	7.21	7.22	7.23
$\ f\ _{L_2(\Omega)}$	2.487e+02	2.475e+02	2.493e+02	2.316e+02	2.339e+02	2.377e+02	7.142e+01
$\ \hat{f}\ _{\ell_2}$	9.715e-01	9.669e-01	9.740e-01	9.046e-01	9.137e-01	9.287e-01	2.790e-01
$\ N(\hat{f}_d)\ _{\ell_2}$	1.530e-01	1.195e-01	1.020e-01	1.895e-01	2.308e-01	1.756e-01	1.270e-01
$\ \hat{f}_d\ _{\ell_2}$	9.561e-01	9.426e-01	9.541e-01	8.399e-01	8.682e-01	8.837e-01	1.490e-01
$E_{\text{approx}}$	1.721e-01	2.157e-01	1.957e-01	3.358e-01	2.849e-01	2.856e-01	2.359e-01

Tabelle 7.47: Musterabhängige Normen und Fehler Rekonstruktion mit zwölf Strahlen.

Abbildung	7.17	7.18	7.19	7.20	7.21	7.22	7.23
Koordinatenabstiegsverfahren	2.07e-01	2.34e-01	2.11e-01	3.66e-01	3.38e-01	3.21e-01	2.52e-01
Simplexverfahren	2.22e-01	2.38e-01	2.08e-01	3.57e-01	3.45e-01	3.18e-01	2.59e-01
Simplexverfahren (mit NB)	4.22e-01	3.77e-10	2.10e-01	3.57e-01	4.78e-01	4.72e-01	2.53e-01
polares Simplexverfahren	2.30e-01	2.46e-01	2.21e-01	3.83e-01	3.47e-01	3.25e-01	2.64e-01

Tabelle 7.48: Rekonstruktionsgüte  $\|N(\hat{f}_* - \hat{f})\|_{\ell_2}$  der Rekonstruktion mit zwölf Strahlen.

Wiederum sind in Tabelle 7.47 die musterabhängigen und in den Tabellen 7.33, 7.35, 7.37, 7.39, 7.41, 7.43 und 7.45 die verfahrensspezifischen Fehler und Normen aufgelistet.

Die Tabellen 7.34, 7.36, 7.38, 7.40, 7.42, 7.44 und 7.46 listen die Parameter der Nelder und Mead Simplexmethode auf, welche in Tests mit diesem Strahlensetup die besten Ergebnisse lieferten. Auch hier wurde mit  $N_0 = 100$  getestet. Tabelle 7.48 listet der Vollständigkeit halber noch einmal die Rekonstruktionsgüte  $\|N(\hat{f}_* - \hat{f})\|_{\ell_2}$  für alle Muster und Verfahren auf.

Durch dieses Strahlensetup konnten die Ergebnisse nicht sonderlich verbessert werden. Keines der rekonstruierten Muster weist mehr Ähnlichkeiten zum entsprechenden Muster  $f$  auf, als bei der Verwendung von acht Strahlen. Zwar ist die Approximation an  $f$  aus  $\mathcal{S}_d$  visuell besser, was in den Abb. 7.17, 7.18, 7.19 und 7.21 auch dadurch bestätigt wird, dass  $E_{\text{approx}}$  im Allgemeinen kleiner ist als in den vorigen Setups, aber nun entstehen größere Abweichungen zwischen  $f_*$  und  $f_d$ .

Es fällt auf, dass trotz der visuellen Verbesserung von  $f_d$ ,  $E_{\text{approx}}$  für alle Muster relativ groß und in drei Fällen sogar größer als in Abschnitt 7.1.2 (Abb. 7.20, 7.22 und 7.23). Dies ist darauf zurückzuführen, dass in  $\mathcal{K}_d$  im Vergleich zu Abschnitt 7.1.2 nur sehr wenige Elemente direkt um den Ursprung verteilt liegen. Dort liegen aber im Allgemeinen die betragsmäßig größten Fourierkoeffizienten der Funktion  $f$ . In Abb. 7.24 stellen die Punkte die Elemente der Menge  $\mathcal{K}_d$  aus Abschnitt 7.1.3

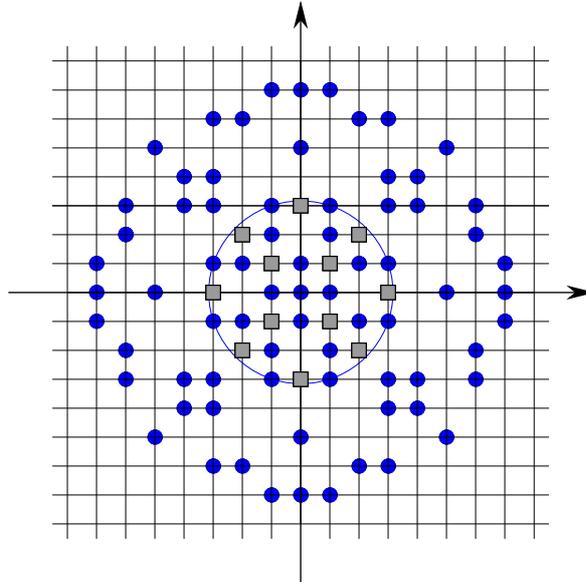


Abbildung 7.24: Die Mengen  $\mathcal{K}_d$  aus Abschnitt 7.1.2 und Abschnitt 7.1.3 sind hier gemeinsam dargestellt. Die Punkte stellen die Differenzen dar, welche sich durch die in (7.3) definierte Menge  $\mathcal{K}$  ergeben, während die durch (2.2) resultierende Menge innerhalb des Kreises liegt (Punkte und Quadrate). Letztere hat ihre Elemente um den Ursprung konzentriert, die erste Menge verteilt sich auch in größerer Distanz zum Ursprung.

dar, welche teilweise um den Ursprung, aber auch weiter entfernt liegen. Im Kreis sind zusätzlich zu den Elementen dieser Menge auch die Elemente der Differenzenmenge (2.2) aus Abschnitt 7.1.2 durch Quadrate hinzugefügt, soweit sie noch nicht in der Menge  $\mathcal{K}_d$  enthalten waren. In diesem Kreis liegen somit alle Elemente von  $\mathcal{K}_d$  aus Abschnitt 7.1.2. Bis auf vier vereinzelte Stellen hat diese Menge keine Lücken, während das größere  $\mathcal{K}_d$  aus Abschnitt 7.1.3 sehr große Lücken aufweist. Die bessere Konzentration der Differenzen aus Abschnitt 7.1.2 um den Ursprung ist hier deutlich zu erkennen. Dies ist auch für größer werdendes  $N$  durch eine bessere Wahl des Strahlensetups in  $\mathcal{K}$  anzustreben.

Wieder wurden die berechneten Minima des Kostenfunktional durch die weiteren Freiheitsgrade größer. Die Resultate von Koordinatenabstiegs- und Simplexverfahren weichen nun auch voneinander ab, während bei den vorherigen Berechnungen fast identische Ergebnisse erzielt wurden. Für Abb. 7.17, 7.18, 7.21 und 7.23 lieferte das Koordinatenabstiegsverfahren die besseren Minima, für die restlichen Muster das Simplexverfahren. Die gleichen Diskrepanzen zeigen Simplexverfahren (mit NB) und das polare Simplexverfahren auf. Visuell ist hier das Koordinatenabstiegsverfahren dem Simplexverfahren und Simplexverfahren (mit NB) bzw. polarem Simplexverfahren vorzuziehen. Insgesamt ist das Setup aus Abschnitt 7.1.2 sogar noch vorzuziehen, obwohl hier  $N$  kleiner gewählt war.

Die Frage nach der Ursache für die großen Abweichungen in den Resultaten gibt Anlass zu weiteren Untersuchungen wie z.B.:

- Erreicht Algorithmus 4.16 ein absolutes Minimum? (Das absolute Minimum ist nach Bemerkung 4.6 nicht eindeutig.)
- Wird für  $N \rightarrow \infty$  das Minimum von  $\mathcal{J}_c$  größer?

- Wird  $\mathcal{S}_N$  für  $N \rightarrow \infty$  größer,
- oder wird durch mehr Abhängigkeiten in (3.9)  $\mathcal{S}_N$  für  $N \rightarrow \infty$  sogar kleiner?
- Kann  $\mathcal{K}$  besser gewählt werden?

Betrachtet man die Ergebnisse von Abschnitt 7.1 im Ganzen, so ist festzustellen, dass die Hinzunahme der Nebenbedingung durch das Simplexverfahren (mit NB) und das polare Simplexverfahren sowohl in den numerischen als auch visuellen Ergebnissen keine Verbesserung bringen konnte. Dies bestätigt Bemerkung 4.3, in der die Nebenbedingung ja aus analytischen Gründen noch vernachlässigt wurde, nun auch numerisch.

Für die einfacheren Setups ( $N = 4$  und  $N = 8$ ) bestätigen Simplexverfahren und Koordinatenabstiegsverfahren noch gegenseitig ihre Ergebnisse, was keinen Anlass gibt ein Verfahren zu bevorzugen. Für  $N = 12$  fällt eine Aussage über das bessere Verfahren zunächst schwer. Es ist aber allgemein bekannt, dass das Nelder und Mead Simplexverfahren für größer werdende Dimension schlechter wird [11]. Zur Verbesserung der Rekonstruktionen ist aber sicherlich noch die Verwendung von mehr als acht oder zwölf Strahlen notwendig. Ob dann die Verwendung der Simplexmethode noch sinnvoll ist, bleibt fraglich.

Zusätzlich geht bei der Minimierung von  $\mathcal{J}$  aber auch ein wesentlicher Vorteil der Simplexmethode gegenüber dem Koordinatenabstiegsverfahren verloren. Dieser Vorteil der Simplexmethode lag darin, dass keine Ableitungen berechnet werden mussten und es nur mit Funktionsauswertungen auskommt. Da hier aber das Minimum in jeder Koordinatenrichtung analytisch berechnet werden kann (vgl. Beweis von Lemma 4.9), benötigt das Koordinatenabstiegsverfahren auch nur eine „Funktions-“ Auswertung von (4.10). Es ist an dieser Stelle kein weiteres Verfahren zur Berechnung des Minimums von  $\mathcal{J}$  in einer Koordinatenrichtung notwendig.

Diese Tatsachen lassen den Schluss zu, dass das in [5] vorgeschlagene Koordinatenabstiegsverfahren eine sinnvolle Wahl zur Minimierung von  $\mathcal{J}$  ist.

## 7.2 Größe des Fouriergitters

Zur Berechnung der im obigen Abschnitt exemplarisch dargestellten Muster wurde als Größe für das Fouriergitter  $L = 256$  gewählt.

In diesem Abschnitt soll nun untersucht werden, ob durch eine andere Wahl von  $L$  die Rekonstruktion verbessert werden kann, da durch eine feinere Auflösung der diskreten Fouriertransformation die kontinuierliche Fouriertransformation besser approximiert werden kann wie in Abschnitt 3.4 erläutert wurde. Da die theoretischen Resultate aus diesem Abschnitt einen sehr hohen Wert für  $L$  liefern wie im folgenden Abschnitt 7.2.1 gezeigt wird, werden in Abschnitt 7.2.2 noch einmal Tests mit verschiedenen Werten von  $L$  durchgeführt, um nachzuprüfen ob dadurch auch eine Verbesserung des Endresultates erreicht werden kann.

### 7.2.1 A priori Fehlerschätzungen für die diskrete Fouriertransformation

Im Allgemeinen kann das Rekonstruktionsverfahren auf beliebige Muster angewandt werden, welche alle Grauwerte annehmen. Dies muss dann in der Theorie der Fehlerschätzer berücksichtigt werden. In Abschnitt 7 wurden allerdings nur Schwarz-Weiß-Muster rekonstruiert, weshalb sich die weitere Betrachtung auf Muster dieses Typs beschränkt.

Für die Approximation der diskreten an die kontinuierliche Fouriertransformation wurde in Abschnitt 3.4 die Fehlerschätzung (3.36)

$$\begin{aligned} |\mathcal{F}[g](\mathbf{k})|_{\Omega} - \mathcal{F}[\tilde{g}](\mathbf{k})|_{\Omega} | &\leq \frac{2}{3}\sqrt{2} \frac{L_{\Omega_M}}{(2\pi)^2} |\mathbf{k}| \Delta x \Delta y \sqrt{(\Delta x)^2 + (\Delta y)^2} + \frac{L_{\Gamma}}{(2\pi)^2} \Delta x \Delta y \\ &\leq \frac{2}{3}\sqrt{2} \frac{|\Omega|}{(2\pi)^2} |\mathbf{k}| \left\| \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right\|_2 + \frac{L_{\Gamma}}{(2\pi)^2} \Delta x \Delta y, \end{aligned}$$

speziell für die hier verwendeten, nichtstetigen Funktionen  $f = \chi_{\Omega_M}$  berechnet.

Da die Originalmuster  $f$  Bilder der Größe  $256 \times 256$  Pixel sind, ist das Gebiet  $\Omega_M$  eine Vereinigung von Rechtecken, d.h. genauer der schwarzen bzw. weißen Pixel. Betrachtet man nun das Gesamtgebiet  $\Omega = [0, 256]^2$ , so liegt der Rand  $\partial\Omega_M = \Gamma$  für jede Auflösung der Fouriertransformation  $L = 2^J > 256$  auf den Gitternetzlinien. Hieraus folgt, dass kein Rechteck  $\Omega_{\mathbf{z}}$  (vgl. (3.30)) existiert, dessen Inneres nichtleeren Schnitt mit dem Rand von  $\Omega_M$  hat, d.h. dass  $L_{\Gamma} = 0$  ist. Es wird hier die schärfere der beiden Abschätzungen verwendet, da  $L_{\Omega_M}$  für jedes Muster exakt berechnet werden kann. Die weitere Abschätzung wird lediglich gemacht, um die Fehlerordnung 1 herauszuheben. Da  $L_{\Gamma} = 0$  ist, reduziert sich die rechte Seite der Abschätzung (3.36) auf den ersten Summanden, welcher sich aus dem Faktor

$$E := \frac{2}{3}\sqrt{2} \frac{L_{\Omega_M}}{(2\pi)^2} \Delta x \Delta y \sqrt{(\Delta x)^2 + (\Delta y)^2} \quad (7.4)$$

und dem Betrag von  $\mathbf{k}$  zusammensetzt. Die Schrittweiten  $\Delta x = \Delta y = \frac{256}{L}$  sind in beiden Richtungen gleich.

Obiger Faktor  $E$  kann für jedes Muster der Form  $f = \chi_{\Omega_M}$  a priori berechnet werden. In Abb. 7.25 ist für das Gebiet  $\Omega_M = [64, 192]^2$  das Muster  $f$  dargestellt. Dieses spezielle Muster wurde hier gewählt, da seine kontinuierliche Fouriertransformation  $\mathcal{F}[f]$  sehr leicht auszurechnen ist und sich somit auch die linke Seite der Ungleichung (3.36) berechnen lässt. In Abb. 7.25 sind die Resultate dieser Berechnungen sowie der Fehler  $E$  für ein  $\mathbf{k} \neq 0$  graphisch dargestellt. Dabei wurde auch die linke Seite der Ungleichung mit  $|\mathbf{k}|^{-1}$  multipliziert, um einen Vergleich der Werte zu ermöglichen. Das bedeutet, dass auf der  $x$ -Achse die Schrittweite  $\Delta x = \Delta y$  und auf der  $y$ -Achse die Fehler  $E$  bzw.  $|\mathcal{F}[f](\mathbf{k})|_{\Omega} - \mathcal{F}[\tilde{f}](\mathbf{k})|_{\Omega} | |\mathbf{k}|^{-1}$  aufgetragen sind. Die vorhergesagte Fehlerordnung 1 wird hier bestätigt. In Tabelle 7.49 ist der Fehlerfaktor  $E$  für noch feinere Auflösungen der diskreten Fouriertransformation angegeben.

Es fällt auf, dass man nach Abschätzung (3.36) eine sehr hohe Auflösung des Fouriergitters benötigt, um den Fehlerfaktor  $E$  klein zu machen. Dies kann an einer schlechten Approximation der kontinuierlichen durch die diskrete Fouriertransformation für Funktionen  $f = \chi_{\Omega_M}$  liegen oder daran, dass die Ungleichung (3.36) nicht scharf genug ist, um Rückschlüsse auf eine geeignete Auflösung zuzulassen. Da die Muster  $f$  eine sehr einfache Struktur haben und der exakte Fehler in Abb. 7.25 viel kleiner als der geschätzte Fehler ist, liegt die zweite Möglichkeit nahe, was die Untersuchungen in Abschnitt 7.2.2 motiviert.

## 7.2.2 Numerische Untersuchungen

Um zu verifizieren, dass der in allen numerischen Tests gewählte Wert  $L = 256$  trotz der Resultate aus Abschnitt 7.2.1 sinnvoll ist, sollen nun Tests mit unterschiedlichen Größen des Fouriergitters durchgeführt werden.

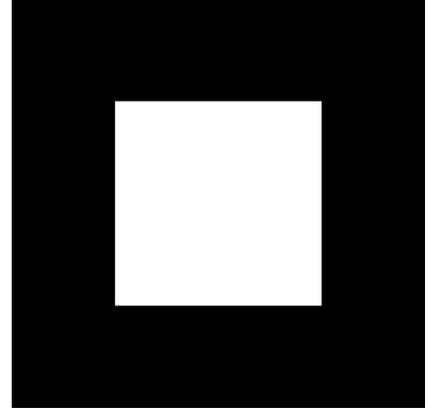
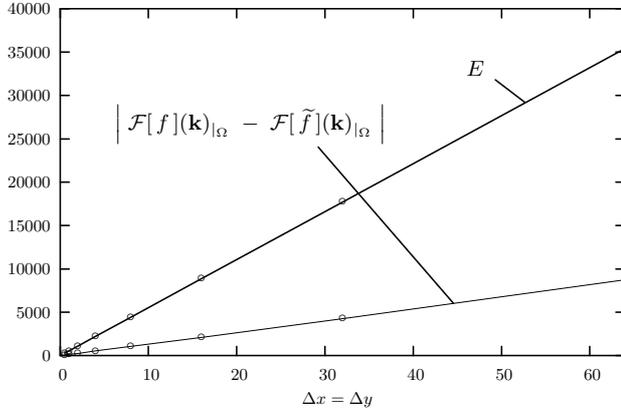


Abbildung 7.25: Rechts ist das Muster dargestellt, für welches exemplarisch der Fehler zwischen kontinuierlicher und diskreter Fouriertransformation bei wachsender Auflösung untersucht wurde. Die linke Graphik stellt die Differenz zwischen der exakt berechneten kontinuierlichen Fouriertransformation und der diskreten Fouriertransformation an der Stelle  $\mathbf{k} = (0, \Delta k_x)^T \in \mathbb{R}^2$ , d.h. die linke Seite aus (3.36) sowie den Faktor  $E$  aus (7.4) gegenüber. Dabei wurde die linke Seite der Abschätzung mit  $|\mathbf{k}|^{-1}$  skaliert, um einen Vergleich mit  $E$  zu ermöglichen.

Auflösung $L$	256	512	1024	2048	4096	8192
$J = \log_2 L$	8	9	10	11	12	13
$\Delta x = \Delta y$	1.000e+00	5.000e-01	2.500e-01	1.250e-01	6.250e-02	3.125e-02
$L_{\Omega_M}$	16 384	65 536	262 144	1 048 576	4 194 304	16 777 216
$E$	5.5335e+02	2.7667e+02	1.3834e+02	6.9169e+01	3.4584e+01	1.7292e+01

Tabelle 7.49: A priori Fehlerschätzungen für verschiedene Auflösungen des Fouriergitters und das Originalmuster aus Abb. 7.25.

Dazu wurde für alle Strahlensetups aus Abschnitt 7.1 Algorithmus 4.16 für Koordinatenabstiegsverfahren, Simplexverfahren, Simplexverfahren (mit NB) und polares Simplexverfahren mit unterschiedlicher Gittergröße noch einmal durchgeführt. Es wurden wieder die gleichen Muster mit  $L = 2^J$  für  $J = 4, 5, \dots, 12$  rekonstruiert und die berechneten Werte des Kostenfunktionals verglichen. Ein großes  $L$  ist gleichbedeutend mit einer hohen Auflösung der Fouriertransformation.

Im Folgenden werden die Ergebnisse in Graphen dargestellt, wobei das Kostenfunktional  $\mathcal{J}$  auf der y-Achse in Abhängigkeit von der Gittergröße  $L$  auf der x-Achse angegeben ist. An den jeweiligen Kurven stehen die Abbildungsnummern der zugehörigen Muster.

In den Abbildungen 7.26, 7.27 und 7.28 sind die Ergebnisse des Koordinatenabstiegsverfahren für die Strahlensetups aus Abschnitt 7.1.1, Abschnitt 7.1.2 und Abschnitt 7.1.3 angegeben. In Abb. 7.26 liegen die Kurven der Abbildungen 7.2, 7.6, 7.7 und 7.8 genau auf der x-Achse, weshalb auf deren Bezeichnung verzichtet wurde, um die Abbildung übersichtlich zu halten.

In fast allen Fällen hat eine Vergrößerung des Fouriergitters über  $L = 2^7 = 128$  hinaus keinen Einfluss mehr auf den Wert des Kostenfunktionals  $\mathcal{J}$ . Ausnahmen bilden hier lediglich 7.3 und 7.4 in Abb. 7.26, wobei hier sogar größere Werte auftreten.

Die Resultate für das Simplexverfahren sind in den Abbildungen 7.29, 7.30 und 7.31 zu sehen. Wie in Abschnitt 7.1 bestätigen die Ergebnisse die des Koordinatenabstiegsverfahrens. In Abb. 7.29 und 7.30 sind alle Kurven fast waagrecht und auch in Abb. 7.31 sind nur kleine Schwankungen in den Funktionswerten zu erkennen. Im Simplexverfahren wurden für alle  $L$  die Parameter  $\beta$ ,  $\gamma$  und  $\theta$  verwendet, welche sich für  $L = 256$  als günstig erwiesen hatten, was eine Ursache für diese Schwankungen sein kann.

Die Abbildungen 7.32, 7.33 und 7.34 zeigen die Testergebnisse des Simplexverfahren (mit NB). In Abb. 7.32 liegen die Kurven der Abb. 7.2, 7.3, 7.7 und 7.8 wieder auf der x-Achse und wurden zur besseren Übersicht nicht mit Bezeichnungen versehen. Teilweise treten für dieses Verfahren erheblich große Schwankungen in den Resultaten auf. Diese zeigen allerdings nicht den Trend, dass für größer werdendes  $L$  die Ergebnisse besser werden, sondern treten in Sprüngen auf. Dies legt eine Optimierung der Parameter  $\beta$ ,  $\gamma$  und  $\theta$  für jedes  $L$  nahe.

Gleiches gilt für die Resultate des polaren Simplexverfahren in den Abbildungen 7.35, 7.36 und 7.37. In Abb. 7.35 sind wieder die Kurven 7.2, 7.3, 7.4, 7.7 und 7.8, welche alle nah an der x-Achse liegen, extra nicht mit Bezeichnungen versehen.

Insgesamt zeigen die Untersuchungen dieses Abschnitts, dass für die beiden wesentlichen Verfahren, d.h. Koordinatenabstiegsverfahren und Simplexverfahren, die Rekonstruktion durch die Wahl eines größeren  $L$  nicht verbessert werden kann. Die Auflösung der Fouriertransformation in Abschnitt 7.1 ist also absolut ausreichend, was wohl unter anderem an der einfachen Struktur der zu rekonstruierenden Muster  $f$  liegt, welche lediglich schwarz- und weiß-Werte annehmen.

Um den Einfluss der Wahl von  $L$  auf das Simplexverfahren (mit NB) und das polare Simplexverfahren vernünftig bewerten zu können, müssten hier wohl zunächst für jede Wahl von  $L$  die geeigneten Parameter  $\beta$ ,  $\gamma$  und  $\theta$  des Simplexverfahrens bestimmt werden. Diese aufwendige Bestimmung wird hier nicht gemacht, da Abschnitt 7.1 schon gezeigt hat, dass die Hinzunahme der Nebenbedingungen eher von Nachteil ist.

Insgesamt kann man aus Abschnitt 7.1 aber auch für das Simplexverfahren schließen, dass es die Ergebnisse des Koordinatenabstiegsverfahrens im Allgemeinen höchstens bestätigt und nicht verbessert hat. Daher beschränken sich die Untersuchungen in den weiteren Abschnitten ebenfalls auf das Koordinatenabstiegsverfahren.

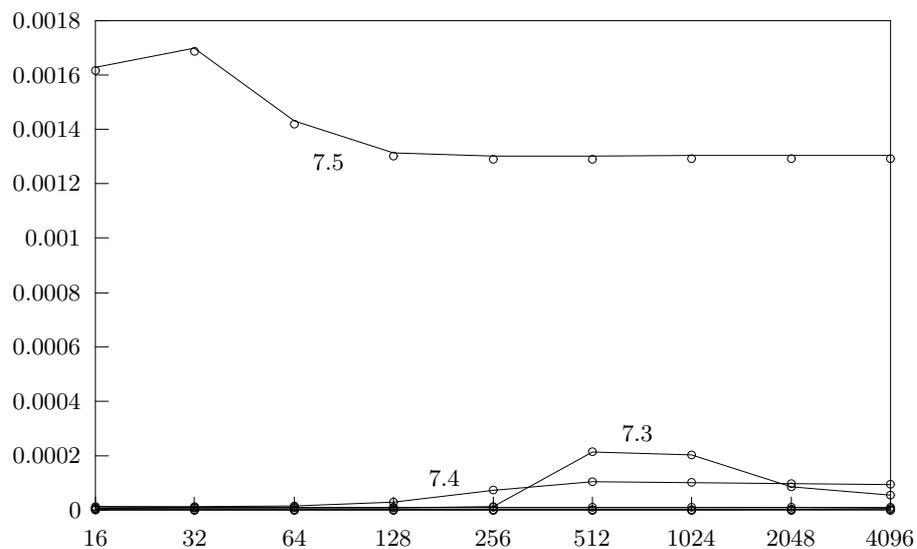


Abbildung 7.26: Abhängigkeit des Kostenfunktionals  $\mathcal{J}$  von der Auflösung  $L$  der diskreten Fouriertransformation unter Verwendung des Koordinatenabstiegsverfahrens für das Setup aus Abschnitt 7.1.1 mit vier Strahlen.

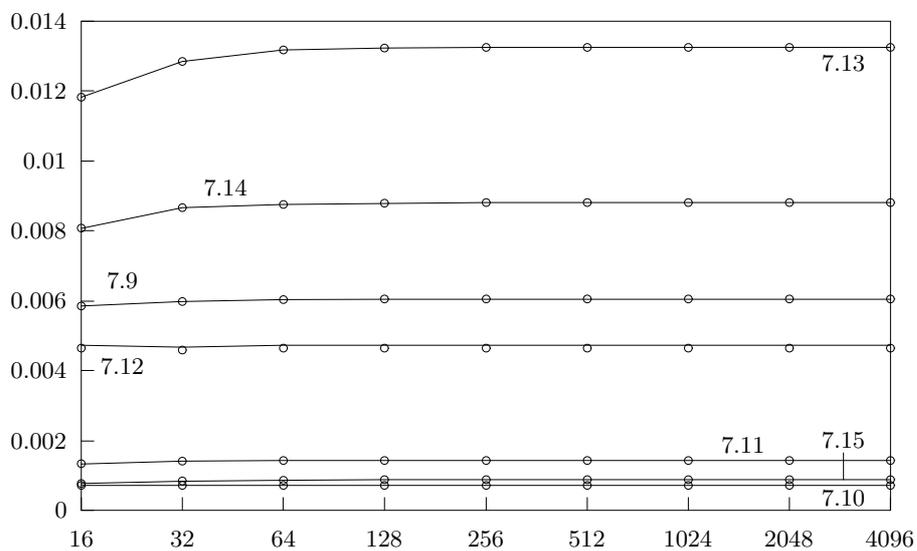


Abbildung 7.27: Werte des Kostenfunktionals  $\mathcal{J}$  für verschiedene Auflösungen  $L$  der diskreten Fouriertransformation. Es wurde das Koordinatenabstiegsverfahren für das Setup aus Abschnitt 7.1.2 verwendet (acht Strahlen).

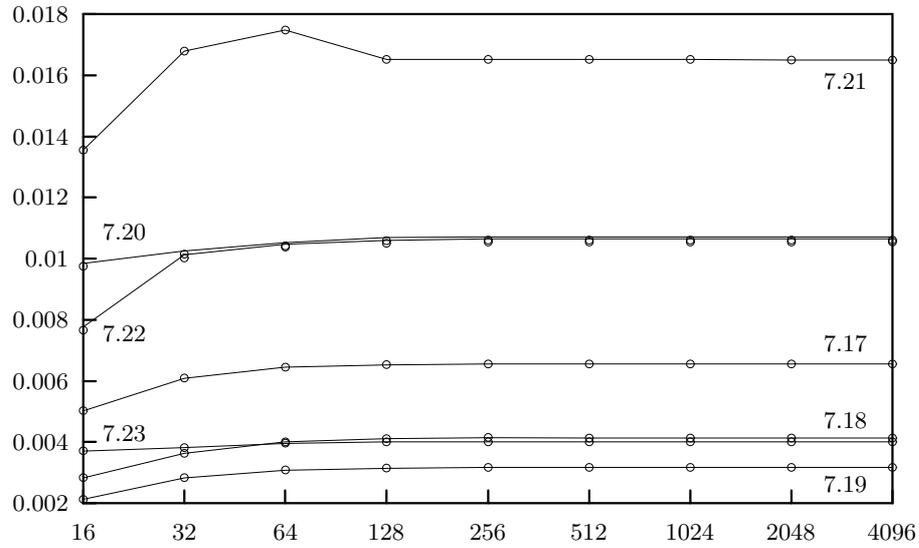


Abbildung 7.28: Abhängigkeit des mittels Koordinatenabstiegsverfahren berechneten Minimums des Kostenfunktional  $\mathcal{J}$  von der Größe des Fouriergitters  $L$  für das Strahlensetup aus Abschnitt 7.1.3 (zwölf Strahlen).

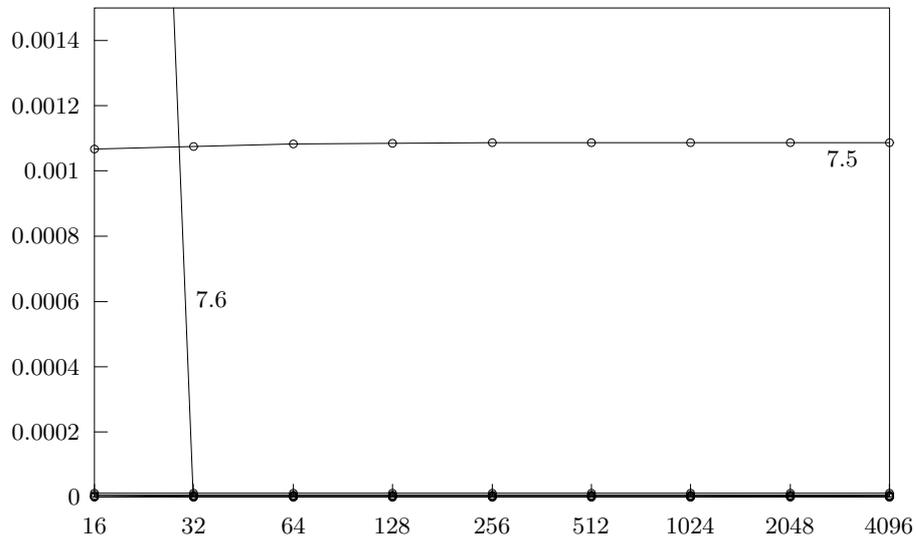


Abbildung 7.29: Werte des Kostenfunktional  $\mathcal{J}$  für verschiedene Auflösungen  $L$  der diskreten Fouriertransformation bei Verwendung des Simplexverfahrens für das Setup aus Abschnitt 7.1.1 (vier Strahlen).

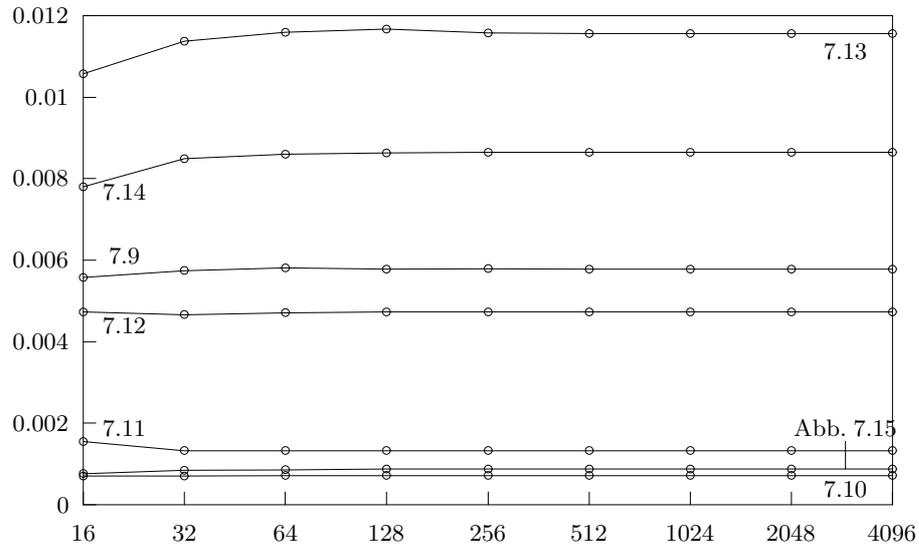


Abbildung 7.30: Für verschiedene Größen  $L$  des Fouriergitters ist der Wert des Kostenfunktionals  $\mathcal{J}$  dargestellt. Die Rekonstruktion erfolgte für das Setup mit acht Strahlen (Abschnitt 7.1.2) mit dem Simplexverfahren.

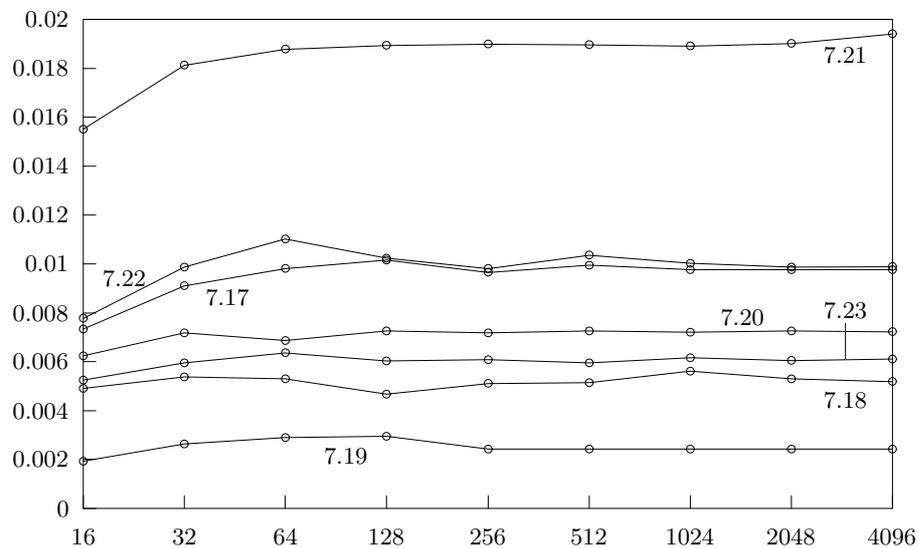


Abbildung 7.31: Werte des Kostenfunktionals  $\mathcal{J}$  in Abhängigkeit von der Größe  $L$  des Fouriergitters unter Anwendung des Simplexverfahrens für zwölf Strahlen (Abschnitt 7.1.3).

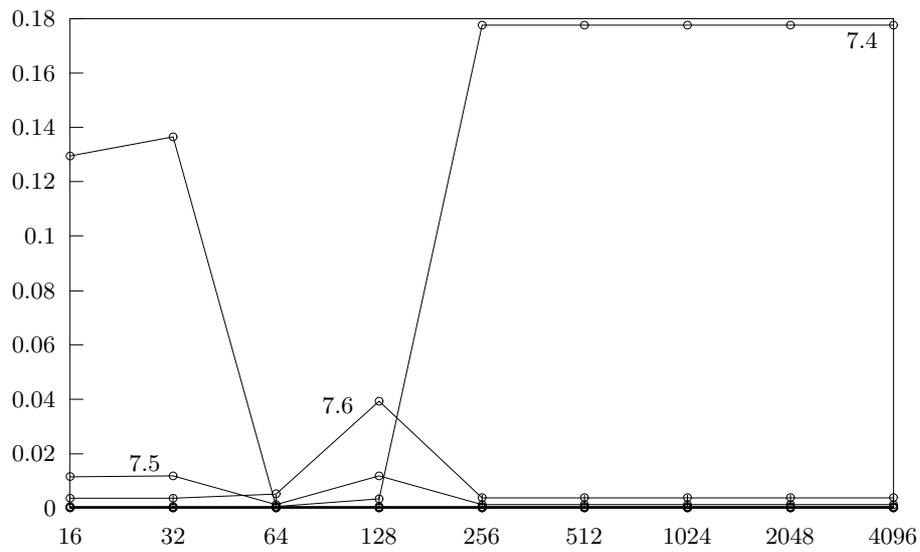


Abbildung 7.32: Abhängigkeit der Werte des Kostenfunktional  $\mathcal{J}$  von der Auflösung  $L$  der diskreten Fouriertransformation. Hier wurde das Simplexverfahren (mit NB) und das Strahlensetup aus Abschnitt 7.1.1 (vier Strahlen) zur Rekonstruktion verwendet.

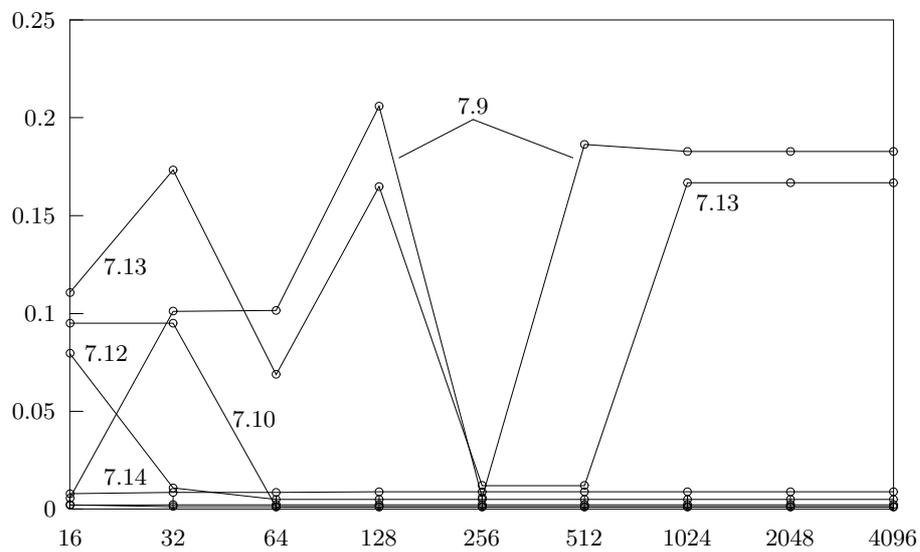


Abbildung 7.33: Abhängigkeit des Kostenfunktional  $\mathcal{J}$  von der Auflösung  $L$  der diskreten Fouriertransformation unter Verwendung des Simplexverfahrens (mit NB) für das Setup aus Abschnitt 7.1.2 mit acht Strahlen.

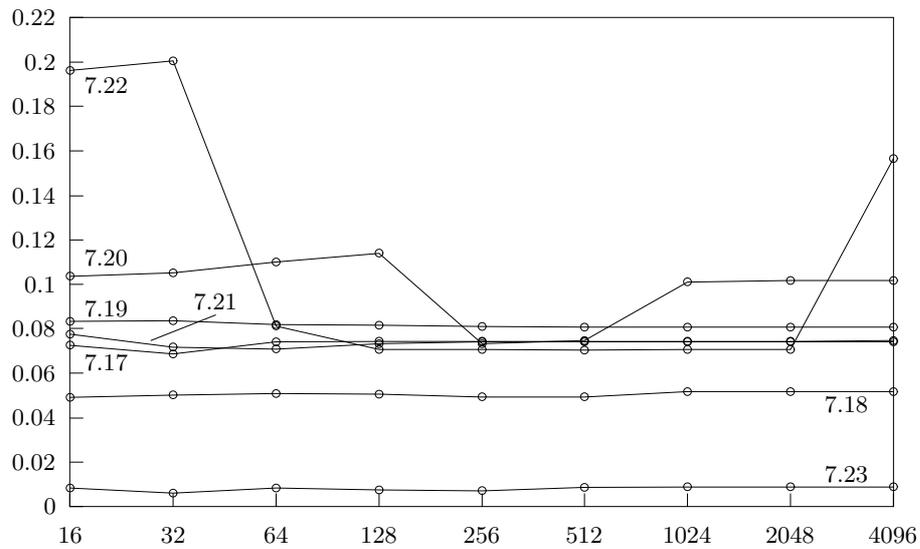


Abbildung 7.34: Werte des Kostenfunktional  $\mathcal{J}$  für verschiedene Auflösungen  $L$  der diskreten Fouriertransformation bei Verwendung des Simplexverfahrens (mit NB) für das Setup aus Abschnitt 7.1.3 (zwölf Strahlen).

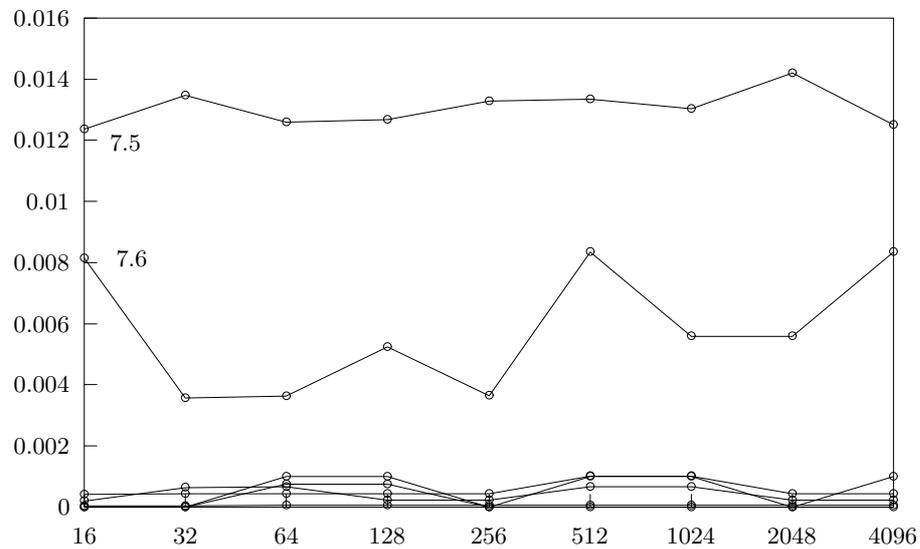


Abbildung 7.35: Für verschiedene Größen  $L$  des Fouriergitters ist der Wert des Kostenfunktional  $\mathcal{J}$  dargestellt. Die Rekonstruktion erfolgte für das Setup mit vier Strahlen (Abschnitt 7.1.1) mit dem polaren Simplexverfahren.

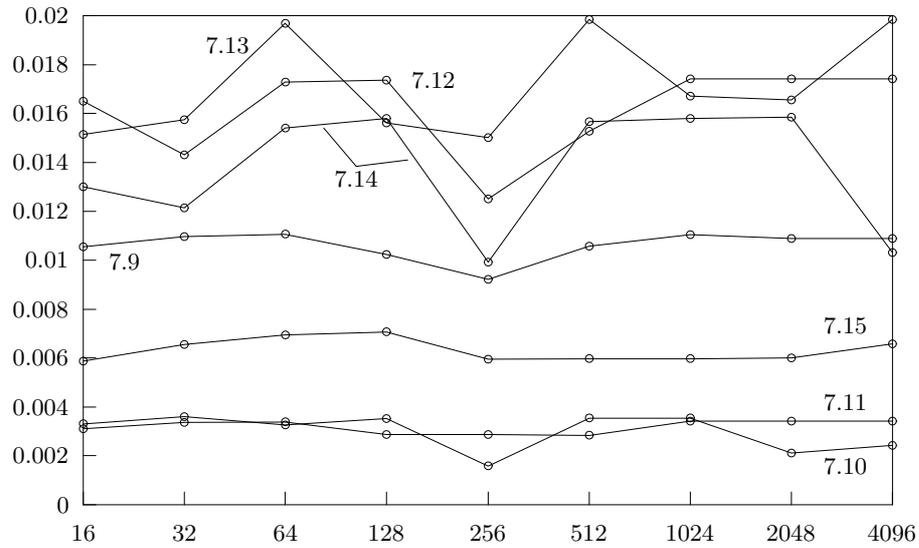


Abbildung 7.36: Abhängigkeit des mittels polarem Simplexverfahren berechneten Minimums des Kostenfunktional  $\mathcal{J}$  von der Größe des Fouriergitters  $L$  für das Strahlensetup aus Abschnitt 7.1.2 (acht Strahlen).

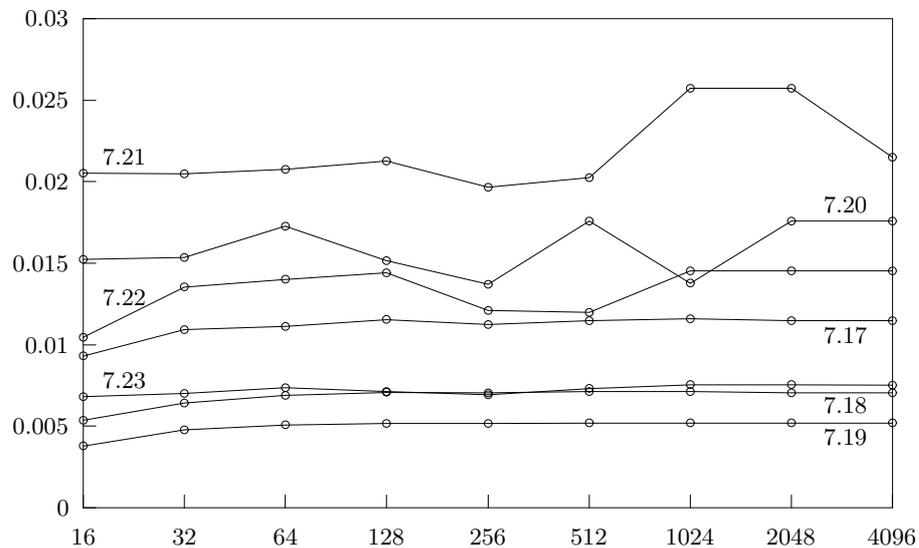


Abbildung 7.37: Abhängigkeit des Kostenfunktional  $\mathcal{J}$  von der Auflösung  $L$  der diskreten Fouriertransformation unter Verwendung des polaren Simplexverfahrens für das Setup aus Abschnitt 7.1.3 mit zwölf Strahlen.

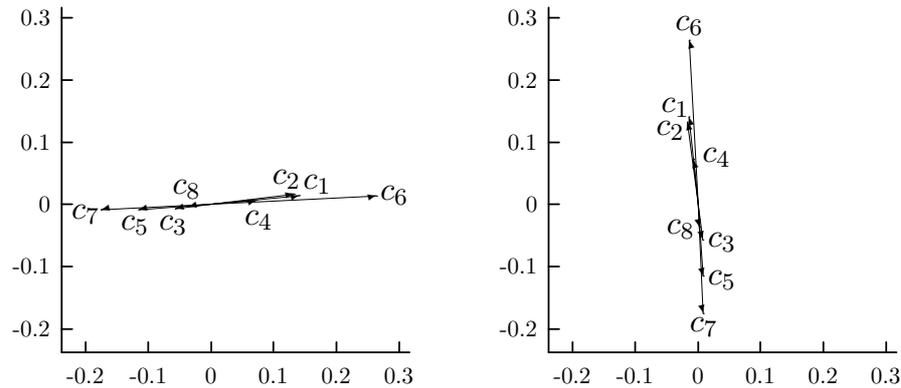


Abbildung 7.38: Entwicklungskoeffizienten von  $f_*$  in der komplexen Ebene  $\mathbb{C}$ . Die Koeffizienten wurden für Abb. 7.15 mit dem Koordinatenabstiegsverfahren berechnet. Dabei wurden links die Startwerte  $c_1 = \dots = c_8 = 1$  und rechts  $c_1 = \dots = c_8 = i$  verwendet.

### 7.3 Koordinatenabstiegsverfahren mit verschiedenen Startwerten

Das Koordinatenabstiegsverfahren wie es in Algorithmus 4.10 definiert wurde, benötigt Startwerte für die Iteration. Es stellt sich nun die Frage, ob der Grenzwert der Folge  $\{\mathbf{c}^k\}$  von diesen Startwerten abhängt. Lemma 4.9 garantierte, dass dieser Grenzwert ein stationärer Punkt des Funktionals  $\mathcal{J}$  ist. Da das Minimum von  $\mathcal{J}$  nach Bemerkung 4.6 nicht eindeutig ist, lässt dies vermuten, dass die berechneten Koeffizienten  $c_j$  und somit  $f_*$  von den Startwerten abhängen. Sollte der Algorithmus das absolute Minimum nicht erreichen (vgl. (4.7)), ist dies sogar wahrscheinlich.

Die Abbildungen 7.38 und 7.39 zeigen die mittels Koordinatenabstiegsverfahren berechneten Koeffizienten von  $f_*$  für das Muster aus Abb. 7.15 in der komplexen Ebene  $\mathbb{C}$ . Die zugehörigen Startwerte kann man den Bildunterschriften entnehmen.

Die Koeffizienten sind in allen vier Fällen jeweils nur um ein Phase  $\varphi$  verschoben, wobei in Abb. 7.39 die Nummerierung der Koeffizienten im Vergleich zu Abb. 7.38 zusätzlich zyklisch um  $\frac{N}{2}$  vertauscht ist. Diese Vertauschung hat aber aufgrund des symmetrischen Strahlensetups und der daraus folgenden Tatsache, dass

$$\mathbf{k}_{j,l} = \mathbf{k}_j - \mathbf{k}_l = \mathbf{k}_{(l+\frac{N}{2}) \pmod{N}} - \mathbf{k}_{(j+\frac{N}{2}) \pmod{N}}$$

gilt, ebenso wie die Phasenverschiebung keinen Einfluss auf das Muster.

Es sind an dieser Stelle nur die Koeffizientenvektoren für das Muster aus Abb. 7.15 dargestellt, da sich für die anderen Muster die Koeffizienten fast nur durch ihre Beträge und nicht durch ihre Richtungen unterscheiden. Dies ist natürlich für den hier untersuchten Sachverhalt unerheblich, macht aber eine graphische Darstellung unmöglich, da alle Pfeile übereinander liegen würden und somit nicht mehr unterschieden werden könnten. Es sei daher an dieser Stelle nur erwähnt, dass sich bei den anderen Mustern die berechneten Koeffizienten, bei Veränderung von Startwerten ebenfalls durch eine Phasenverschiebung ergeben und somit zum gleichen Muster führen. Es ergibt sich also das gleiche Resultat wie hier dargestellt wurde.

Eine Phasenverschiebung erhält die Struktur des Musters  $f_*$ , was die Vermutung zulässt, dass (bis auf diese Phasenverschiebung) ein absolutes Minimum erreicht wird.

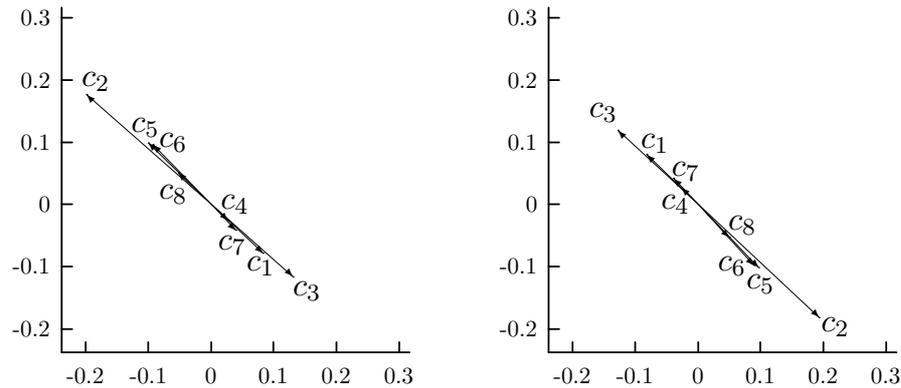


Abbildung 7.39: Wiederum sind die Koeffizienten von  $f_*$  für Abb. 7.15 in der komplexen Ebene  $\mathbb{C}$  dargestellt. Hier wurden links die Startwerte  $c_1 = \dots = c_4 = 1$  und  $c_5 = \dots = c_8 = i$ , rechts  $c_1 = \dots = c_4 = i$  und  $c_5 = \dots = c_8 = 1$  verwendet.

#### 7.4 Konvergenzgeschwindigkeit des Koordinatenabstiegsverfahrens

Abschnitt 7.1 endete mit dem Schluss, dass das Koordinatenabstiegsverfahren eine sinnvolle Wahl zu Rekonstruktion ist. Aus diesem Grunde wird in diesem Abschnitt das Konvergenzverhalten des Verfahrens genauer untersucht. Dazu wird im Folgenden der Wert des Kostenfunktionals  $\mathcal{J}$  nach jedem Iterationsschritt des Koordinatenabstiegsverfahren betrachtet.

Die Abbildungen 7.40, 7.41 und 7.42 geben den Wert von  $\mathcal{J}$  auf der y-Achse in Abhängigkeit vom Iterationsschritt  $m$  auf der x-Achse an. Wieder sind die zugehörigen Abbildungsnummern an den Kurven angegeben. In Abb. 7.40 liegen die Kurven der Muster in den Abb. 7.2, 7.3, 7.4, 7.7 und 7.8 fast übereinander, weshalb sie nicht einzeln beschriftet wurden.

Für alle Strahlensetups sowie Muster tritt eine Saturation schon nach wenigen Iterationsschritten ein. Die maximale Iterationszahl  $N_0$  in Algorithmus 4.10 wird also in keinem der Fälle erreicht, was dieses Abbruchkriterium nahezu überflüssig macht.

Wesentlicher Grund für die schnelle Saturation ist, dass in Algorithmus 4.10 das absolute Minimum von  $\mathcal{J}_j$ , d.h. das Minimum des Kostenfunktionals  $\mathcal{J}$  in  $j$ -ter Koordinatenrichtung, analytisch berechnet werden kann und somit keine weitere innere Iteration zur numerischen Berechnung dieses Minimums nötig ist.

Die Ergebnisse zeigen die hohe Konvergenzgeschwindigkeit des Verfahrens und unterstreichen die Aussage aus Abschnitt 7.1, dass es eine sinnvolle und effiziente Wahl zur Rekonstruktion der  $c_j$  ist.

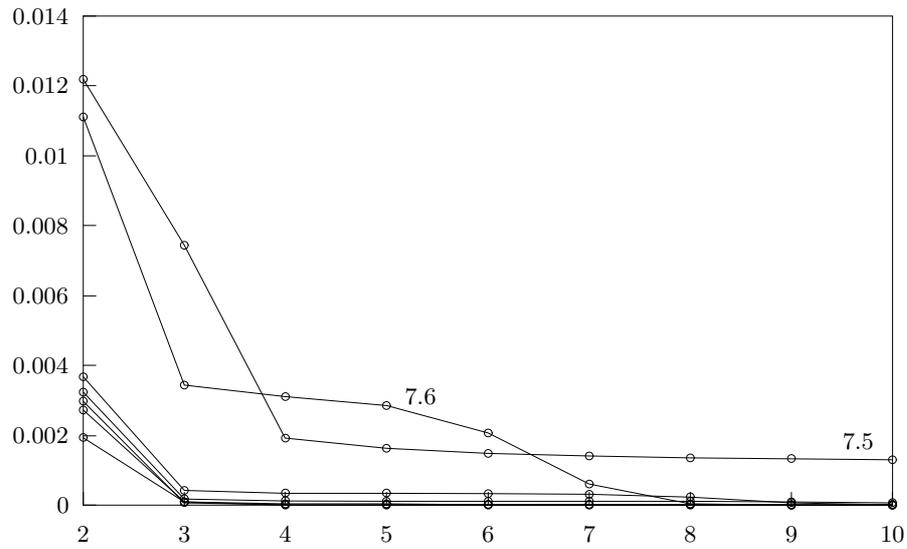


Abbildung 7.40: Wert des Kostenfunktional  $\mathcal{J}$  für jeden Iterationsschritt des Koordinatenabstiegsverfahrens unter Verwendung von vier Strahlen (Abschnitt 7.1.1). Eine Saturation tritt schon nach wenigen Iterationen ein.

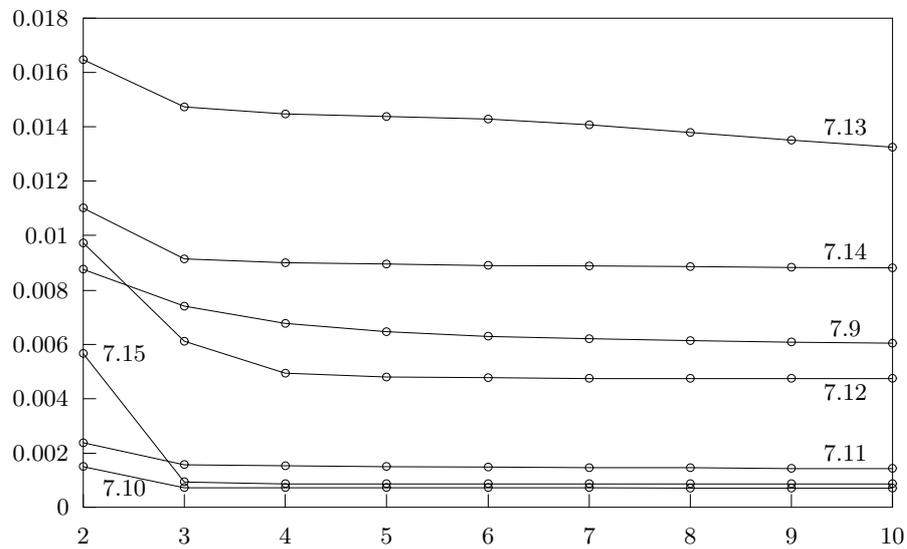


Abbildung 7.41: Es ist der Fehler  $\mathcal{J}$  für jeden Iterationsschritt des Koordinatenabstiegsverfahrens dargestellt. Hier wurde das Setup aus Abschnitt 7.1.2 verwendet (acht Strahlen). Eine Saturation tritt auch hier schon nach wenigen Iterationen ein.

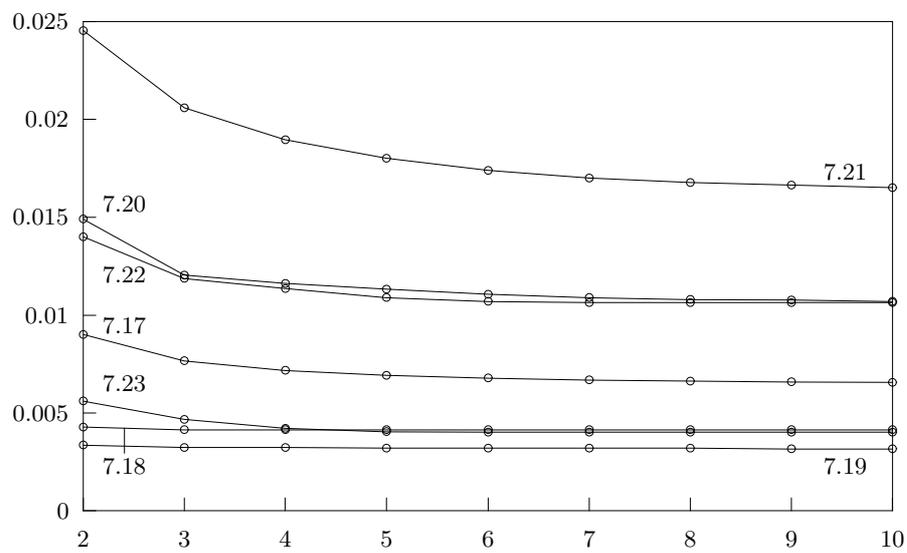


Abbildung 7.42: Für jeden Iterationsschritt des Koordinatenabstiegsverfahrens ist der Wert von  $\mathcal{J}$  dargestellt. Es wurden zwölf Strahlen gemäß Abschnitt 7.1.3 verwendet. Wiederum tritt eine Saturation schon nach wenigen Iterationen ein.

## 8 Zusammenfassung und Ausblick

In dieser Arbeit sind die Ergebnisse systematischer Tests zur *numerischen Berechnung von komplexen Lichtmasken in der atomaren Nanofabrikation*, aber auch die dazu benötigten theoretischen Grundlagen dargestellt.

Dabei wurde zu Beginn das mathematische Modell zur Behandlung dieser Aufgabe vorgestellt, was zu einem komplexen, nichtlinearen inversen Problem führte. Dies geschah in starker Anlehnung an [5], wobei nun die Approximationsräume  $\mathcal{S}_d$  und  $\mathcal{S}_N$  definiert wurden, was in dieser Arbeit noch nicht explizit gemacht wurde.

Der wesentliche Unterschied liegt jedoch in der Konstruktion und anschließenden Betrachtung von regelmäßigen Strahlensetups, was zu den linearen Abhängigkeiten der Differenzen in der Menge  $\mathcal{K}_d$  und damit zu den Bedingungen (3.9) an die Entwicklungskoeffizienten führte.

Des Weiteren schlossen die theoretischen Untersuchungen Fehlerschätzungen für die diskrete Fouriertransformation ein.

Um das Koordinatenabstiegsverfahren im Hinblick auf die Güte der berechneten Minima des Kostenfunktionalen bewerten zu können, wurde zum Vergleich das Nelder und Mead Simplexverfahren eingeführt und implementiert. Dabei eröffnete Letzteres sogar Möglichkeiten zur Berücksichtigung von Nebenbedingungen.

Die theoretische Grundlage der linearen Abhängigkeit von Differenzen der Richtungsvektoren  $\mathbf{k} \in \mathcal{K}$  in Abschnitt 3 und Abschnitt 4.1 wurden zwar in [5] noch ausgeschlossen, sind aber von C. Burstedde am Institut für Angewandte Mathematik der Universität Bonn schon entwickelt, allerdings noch nicht in einer Arbeit niedergeschrieben worden. Ebenso wurde für diesen Fall der Kern des Rekonstruktionsprogramms mit dem Koordinatenabstiegsverfahren als Minimierer von ihm implementiert, was die Berechnung von Normen und Fehlern sowie das Nelder und Mead Simplexverfahren in seinen Varianten allerdings nicht einschließt.

Der Versuch zur theoretischen Bestimmung eines Minimums des Kostenfunktionalen mittels der *Rang Eins Approximation* führte zu keinem positiven Resultat und scheint auch für zukünftige Untersuchungen kein fruchtbarer Ansatz zu sein.

Ebenso führte der Versuch der Approximation mit *Ridge-Functions* zu keinem Fortschritt in approximationstheoretischer Hinsicht. Der Raum  $\mathcal{S}_N$  kann nicht mit einer Basis von Ridge-Functions in herkömmlichen Sinne dargestellt werden. Es stellt sich aber die Frage, ob nicht Techniken aus diesem Themenbereich auf die hier benötigte Basis, welche ja noch Ähnlichkeiten zu einer Ridge-Function-Basis hat, angewandt werden können. So wird in [19] auch ein Ridge-Function-Raum mit variablen Richtungen definiert.

Um den Einfluss der Originalbilder zu bewerten, wurden verschiedene Muster rekonstruiert.

Die numerischen Tests haben die Effizienz des Koordinatenabstiegsverfahrens gezeigt. Das alternative Nelder und Mead Simplexverfahren hat die Ergebnisse im Wesentlichen bestätigt und nur in Ausnahmefällen wirklich verbessert. Durch die beiden zusätzlichen Varianten der Simplexmethode zur Berücksichtigung der Nebenbedingungen konnte Bemerkung 4.3 auch numerisch bestätigt werden, d.h. bei zukünftigen Betrachtungen macht es keinen Sinn, großen Wert auf die Nebenbedingung zu legen. Das Koordinatenabstiegsverfahren ist somit eine geeignete Wahl zur Minimierung des Kostenfunktionalen, insbesondere wenn man berücksichtigt, dass das Nelder und Mead Simplexverfahren für größer werdende Dimensionen schlechter wird, wie Erfahrungen zeigen.

Ebenso ist die Konvergenzrate des Koordinatenabstiegsverfahrens sehr hoch, wie die schnelle Saturation in Abschnitt 7.4 zeigt. Die Wahl unterschiedlicher Startwerte führte lediglich zu einer Phasenverschiebung und in einigen Fällen zu einer zyklischen Vertauschung der Koeffizienten, was

allerdings keinen Einfluss auf das Resultat hat.

Alle Minimierungsverfahren wurden unter Verwendung unterschiedlicher Strahlensetups zur Rekonstruktion der gleichen Originalmuster verwendet. Damit konnte der Einfluss der Richtungen  $\mathbf{k} \in \mathcal{K}$  auf das gesamte Verfahren getestet werden.

In diesem Zusammenhang hat es sich als vorteilhaft herausgestellt, wenn die resultierenden Differenzen  $\mathbf{k} \in \mathcal{K}_d$  direkt um den Ursprung liegen. Somit stellt sich hier offensichtlich die Aufgabe, Strahlensetups zu finden, die diese Bedingung mit möglichst vielen Elementen erfüllen. Allerdings ist dabei zu beachten, dass die  $\mathbf{k} \in \mathcal{K}_d$  in rationalem Verhältnis zueinander stehen müssen (vgl.(3.4)). Dies führt zu einer Verkleinerung des Fehlers  $E_{\text{approx}}$  in (3.19), ist aber aufgrund dieser Bedingung ein zahlentheoretisches Problem.

Insgesamt haben die Untersuchungen in dieser Arbeit gezeigt, dass diese zahlentheoretische Problemstellung, welche sich durch die Konstruktion von Strahlensetups unter obigen Bedingungen ergibt, der wesentliche Bereich ist in dem sich Möglichkeiten zur Verbesserung der Rekonstruktion eröffnen. Das bedeutet, dass die Richtungen der einzelnen Lichtstrahlen zusätzliche Freiheitsgrade im Approximationsraum  $\mathcal{S}_N$  darstellen sollten. Ob in diesem Bereich zahlentheoretische Ergebnisse existieren, ist bis zu diesem Zeitpunkt nicht klar, da eine umfangreiche Literaturrecherche sowie eine Anfrage an Zahlentheoretiker ohne Ergebnis blieben.

## Danksagung

An dieser Stelle möchte ich mich bei Frau Prof. Dr. A. Kunoth bedanken, die mir die Möglichkeit zur Erstellung dieser Diplomarbeit bot und mich mit vielen Anregungen unterstützt hat.

Ein weiterer Dank gilt Herrn Prof. Dr. D. Meschede vom Institut für Angewandte Physik an der Universität Bonn für den Hinweis auf das Problem und die Aufgabenstellung aus der Atomphysik sowie für die Übernahme des Korreferats.

Ebenso gebührt C. Burstedde Dank, der mir bei Problemen und Fragen sehr geholfen hat und zusätzlich den C-Code seines Rekonstruktionsprogrammes zur Verfügung gestellt hat.

Diese Arbeit wurde teilweise vom Sonderforschungsbereich 611 der DFG an der Universität Bonn unterstützt.

Bonn, im Juli 2004

Jürgen Braun

## Anhang

Die *Fouriertransformation* einer Funktion  $g \in L_1(\mathbb{R}^n)$  mit  $n \in \mathbb{N}$  ist definiert durch

$$\mathcal{F}[g](\mathbf{k}) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} g(\mathbf{x}) \exp(-i\langle \mathbf{k}, \mathbf{x} \rangle) \, d\mathbf{x}.$$

Ist weiter  $\mathcal{F}[g] \in L_1(\mathbb{R}^n)$ , so ist die *inverse Fouriertransformation* durch

$$\mathcal{F}^{-1}[\mathcal{F}[g]](\mathbf{x}) := \int_{\mathbb{R}^n} \mathcal{F}[g](\mathbf{k}) \exp(i\langle \mathbf{k}, \mathbf{x} \rangle) \, d\mathbf{k}$$

definiert und es gilt

$$\mathcal{F}^{-1}[\mathcal{F}[g]] = \mathcal{F}[\mathcal{F}^{-1}[g]] = g.$$

Die *Fourierreihenentwicklung* einer  $q$ -periodischen Funktion  $g : \mathbb{R} \rightarrow \mathbb{C}$  mit  $q \in \mathbb{R}$  ist gegeben durch

$$g(x) = \sum_{l=-\infty}^{\infty} c_l \exp\left(\frac{2\pi i}{q} l x\right),$$

mit den komplexen Koeffizienten

$$c_l = \frac{1}{q} \int_a^{a+q} g(x) \exp\left(-\frac{2\pi i}{q} l x\right) \, dx.$$

Dabei ist  $a \in \mathbb{R}$  beliebig.

Es seien nun lediglich die  $L$  diskreten Funktionswerte

$$g_j := g(x_j) \quad \text{für} \quad j = 0, \dots, L-1$$

einer Funktion  $g : \mathbb{R} \rightarrow \mathbb{R}$  an den äquidistanten Stützstellen

$$x_j := j\Delta x \quad \text{für} \quad j = 0, \dots, L-1$$

bekannt. Dabei ist  $\Delta x \in \mathbb{R}$  die Schrittweite zwischen zwei Stützstellen und  $L \in \mathbb{N}$  beliebig. Mit Hilfe dieser Funktionswerte kann man nun eine diskrete Approximation

$$\tilde{g}(x) := \sum_{j=-\infty}^{\infty} \tilde{g}_j \delta(x - x_j)$$

an die Originalfunktion  $g$  mit den skalierten Funktionswerten  $\tilde{g}_j := g_j \Delta x$  definieren. Dabei wurde die Funktion  $\tilde{g}$  periodisch fortgesetzt, d.h. es gilt  $\tilde{g}_{j+L} := \tilde{g}_j$  und  $\tilde{g}$  hat die Periode  $q := L\Delta x$ .

Diese Betrachtungsweise ist z.B. bei numerischen Berechnungen angebracht, da auch hier die Funktion  $g$  nur an diskreten Punkten ausgewertet werden kann.

Da  $\tilde{g}$  eine  $q$ -periodische Funktion ist, kann sie in eine Fourierreihe entwickelt werden, d.h. sie hat die Darstellung

$$\tilde{g}(x) = \sum_{l=-\infty}^{\infty} \hat{g}(l) \exp\left(\frac{2\pi i}{q} l x\right),$$

mit den Fourierkoeffizienten  $\hat{g}(l)$  für  $l \in \mathbb{Z}$ . Man kann nun die Fouriertransformierte dieser Reihenentwicklung direkt berechnen und erhält mit  $\Delta k := \frac{2\pi}{L\Delta x} = \frac{2\pi}{q}$

$$\begin{aligned}\mathcal{F}[\tilde{g}](k) &= \mathcal{F}\left[\sum_{l=-\infty}^{\infty} \hat{g}(l) \exp\left(\frac{2\pi i}{q}lx\right)\right] \\ &= \sum_{l=-\infty}^{\infty} \hat{g}(l) \mathcal{F}[\exp(il\Delta kx)] \\ &= \sum_{l=-\infty}^{\infty} \hat{g}(l) \delta(k - l\Delta k) .\end{aligned}\tag{*}$$

Andererseits kann man die Fouriertransformation der diskreten Approximation  $\tilde{g}$  in der ursprünglichen Darstellung ausrechnen und erhält mit  $p := \frac{2\pi}{\Delta x} = L\Delta k$

$$\begin{aligned}\mathcal{F}[\tilde{g}](k) &= \sum_{j=-\infty}^{\infty} \tilde{g}_j \mathcal{F}[\delta(x - x_j)](k) \\ &= \sum_{j=-\infty}^{\infty} \tilde{g}_j \frac{1}{2\pi} \int_{\mathbb{R}} \delta(x - j\Delta x) \exp(-ikx) dx \\ &= \sum_{j=-\infty}^{\infty} \frac{1}{2\pi} \tilde{g}_j \exp(-ikj\Delta x) \\ &= \sum_{j=-\infty}^{\infty} \frac{1}{2\pi} \tilde{g}_{-j} \exp\left(\frac{2\pi i}{p}kj\right) ,\end{aligned}$$

worin man die Fourierreihenentwicklung der Funktion  $\mathcal{F}[\tilde{g}]$  erkennt. Anhand dieser Darstellung ist die Periodizität der Fouriertransformation  $\mathcal{F}[\tilde{g}]$  und die Periode  $p = \frac{2\pi}{\Delta x}$  direkt ersichtlich, woraus man für die Fourierkoeffizienten von  $\tilde{g}$  die Eigenschaft  $\hat{g}(l) = \hat{g}(l + L\mathbb{Z})$  folgern kann. Dies muss gelten, da ansonsten obige Darstellung von  $\mathcal{F}[\tilde{g}]$  mit  $\delta$ -Funktionen nicht  $p$ -periodisch ist.

Die Berechnung der Koeffizienten  $\hat{g}(l)$  sowie der ursprünglichen Funktionswerte  $g_j = g(x_j)$  wird im Folgenden erläutert.

Die Koeffizienten  $\hat{g}(l)$  der Fourierreihenentwicklung einer  $q$ -periodischen Funktion  $\tilde{g}$  berechnen sich für  $l \in \mathbb{Z}$  bei beliebigem  $a \in \mathbb{R}$  durch die Integrale

$$\hat{g}(l) = \frac{1}{q} \int_a^{a+q} \tilde{g}(x) \exp\left(-\frac{2\pi i}{q}lx\right) dx .$$

Wählt man nun  $a := -\frac{1}{2}\Delta x$ , so folgt mit  $q = L\Delta x$ , dass  $a + q = (L - \frac{1}{2})\Delta x$  ist und die Koeffizienten

berechnet werden durch

$$\begin{aligned}
\widehat{g}(l) &= \frac{1}{q} \int_{-\frac{1}{2}\Delta x}^{(L-\frac{1}{2})\Delta x} \left[ \sum_{j=-\infty}^{\infty} \widetilde{g}_j \delta(x - j\Delta x) \right] \exp\left(-\frac{2\pi i}{q}lx\right) dx \\
&= \frac{1}{q} \sum_{j=-\infty}^{\infty} \int_{-\frac{1}{2}\Delta x}^{(L-\frac{1}{2})\Delta x} \widetilde{g}_j \delta(x - j\Delta x) \exp\left(-\frac{2\pi i}{q}lx\right) dx \\
&= \frac{1}{q} \sum_{j=0}^{L-1} \widetilde{g}_j \exp\left(-\frac{2\pi i}{q}lj\Delta x\right) .
\end{aligned}$$

Erinnert man sich weiterhin, dass die Koeffizienten der diskreten Approximation die mit  $\Delta x$  skalierten Funktionswerte waren, d.h.  $\widetilde{g}_j = g_j \Delta x$  so erhält man

$$\begin{aligned}
\widehat{g}(l) &= \frac{1}{L\Delta x} \sum_{j=0}^{L-1} g_j \Delta x \exp\left(-\frac{2\pi i}{q}lj\frac{q}{L}\right) \\
&= \frac{1}{L} \sum_{j=0}^{L-1} g_j \exp\left(-\frac{2\pi i}{L}lj\right) .
\end{aligned}$$

Die kontinuierliche Fouriertransformation der diskreten Approximation einer  $q$ -periodischen Funktion  $g$  führt somit zur Definition der *diskreten Fouriertransformation*

$$\widehat{g}(l) := \frac{1}{L} \sum_{j=0}^{L-1} g(x_j) \exp\left(-\frac{2\pi i}{L}lj\right) \quad (**)$$

für  $l \in \mathbb{Z}$ , mit den äquidistanten Stützstellen  $x_j := j\Delta x$ , der Schrittweite  $\Delta x = \frac{q}{L}$  und der Auflösung  $L \in \mathbb{N}$ .

Betrachtet man andererseits die Fourierreihenentwicklung von  $\mathcal{F}[\widetilde{g}]$ , so berechnen sich die Fourierkoeffizienten  $\frac{1}{2\pi}\widetilde{g}_{-j}$  dieser Reihe wiederum für  $j \in \mathbb{Z}$  und ein beliebiges  $a \in \mathbb{R}$  durch die Integrale

$$\frac{1}{2\pi}\widetilde{g}_{-j} = \frac{1}{p} \int_a^{a+p} \mathcal{F}[\widetilde{g}](k) \exp\left(-\frac{2\pi i}{p}jk\right) dk ,$$

woraus sich für  $a = -\frac{1}{2}\Delta k$  und  $\Delta k = \frac{p}{L}$  sowie  $p = \frac{2\pi}{\Delta x}$  die Darstellung

$$\begin{aligned}
\frac{1}{2\pi}\tilde{g}_{-j} &= \frac{1}{p} \int_{-\frac{1}{2}\Delta k}^{(L-\frac{1}{2})\Delta k} \sum_{l=-\infty}^{\infty} \hat{g}(l)\delta(k-l\Delta k) \exp\left(-\frac{2\pi i}{p}jk\right) dk \\
&= \frac{1}{p} \sum_{l=-\infty}^{\infty} \int_{-\frac{1}{2}\Delta k}^{(L-\frac{1}{2})\Delta k} \hat{g}(l)\delta(k-l\Delta k) \exp\left(-\frac{2\pi i}{p}jk\right) dk \\
&= \frac{1}{p} \sum_{l=0}^{L-1} \hat{g}(l) \exp\left(-\frac{2\pi i}{p}jl\Delta k\right) \\
&= \frac{\Delta x}{2\pi} \sum_{l=0}^{L-1} \hat{g}(l) \exp\left(-\frac{2\pi i}{L}jl\right)
\end{aligned}$$

ergibt. Damit gilt für die diskreten Funktionswerte

$$g_j = \frac{\tilde{g}_j}{\Delta x} = \sum_{l=0}^{L-1} \hat{g}(l) \exp\left(\frac{2\pi i}{L}jl\right),$$

was zur Definition der *inversen diskreten Fouriertransformation* durch

$$g(x_j) := \sum_{l=0}^{L-1} \hat{g}(l) \exp\left(\frac{2\pi i}{L}jl\right) \quad (***)$$

mit der Schrittweite  $\Delta k = \frac{p}{L}$  und der Auflösung  $L \in \mathbb{N}$  führt.

Geht man nun in zwei Dimensionen analog zur univariaten Situation vor und betrachtet die Funktion  $g$  mit den Schrittweiten  $\Delta x$  in  $x$ -Richtung und  $\Delta y$  in  $y$ -Richtung an den  $L_x L_y$  Stützstellen

$$\mathbf{x}_{\xi,\nu} := (\xi\Delta x, \nu\Delta y)^T \quad \text{für } \xi = 0, \dots, L_x - 1 \quad \text{und} \quad \nu = 0, \dots, L_y - 1,$$

so kann man diese wiederum periodisch fortsetzen durch

$$g(\mathbf{x}_{\xi,\nu}) =: g(\mathbf{x}_{\xi+L_x\mathbb{Z}, \nu+L_y\mathbb{Z}}).$$

Damit ergibt sich in  $x$ -Richtung die Periode  $q_x := L_x\Delta x$  bzw. in  $y$ -Richtung die Periode  $q_y := L_y\Delta y$  und man erhält mit den Schrittweiten  $\Delta k_x := \frac{2\pi}{q_x}$  sowie  $\Delta k_y := \frac{2\pi}{q_y}$  anstatt (\*) mit  $\mathbf{w}_{j,l} := (j, l)^T \in \mathbb{Z}^2$  entsprechend die Darstellung

$$\mathcal{F}[\tilde{g}](\mathbf{k}) = \sum_{j,l=-\infty}^{\infty} \hat{g}(\mathbf{w}_{j,l}) \delta\left(\mathbf{k} - \begin{pmatrix} j\Delta k_x \\ l\Delta k_y \end{pmatrix}\right).$$

Hieraus kann man wie im eindimensionalen Fall die Perioden  $p_x = \frac{2\pi}{\Delta x}$  bzw.  $p_y = \frac{2\pi}{\Delta y}$  der Funktion  $\mathcal{F}[\tilde{g}]$  in  $x$ - bzw.  $y$ -Richtung im zweidimensionalen Fall folgern und damit, dass

$$\hat{g}(\mathbf{w}_{j,l}) = \hat{g}(\mathbf{w}_{j+L_x\mathbb{Z}, l+L_y\mathbb{Z}})$$

gilt. Es ergibt sich somit anstatt (\*\*) die Definition der *zweidimensionalen diskreten Fouriertransformation* durch

$$\widehat{g}(\mathbf{w}_{j,l}) := \frac{1}{L^2} \sum_{\xi, \nu=0}^{L-1} g(\mathbf{x}_{\xi, \nu}) \exp\left(-\frac{2\pi i}{L} \langle \mathbf{w}_{j,l}, (\xi, \nu)^T \rangle\right)$$

für  $\mathbf{w}_{j,l} \in \mathbb{Z}^2$  und umgekehrt anstatt (\*\*\*) die *zweidimensionale inverse diskrete Fouriertransformation*

$$g(\mathbf{x}_{\xi, \nu}) := \sum_{j, l=0}^{L-1} \widehat{g}(\mathbf{w}_{j,l}) \exp\left(\frac{2\pi i}{L} \langle (\xi, \nu)^T, \mathbf{w}_{j,l} \rangle\right)$$

für  $\mathbf{x}_{\xi, \nu} := (\xi \Delta x, \nu \Delta y)^T \in \mathbb{R}^2$ ,  $\xi, \nu \in \mathbb{Z}$ .

## Literatur

- [1] C. S. Adams, M. Siegel, J. Mlynek, *Atom optics*, Phys. Rep. **240**, 143 (1994).
- [2] H. W. Alt, *Lineare Funktionalanalysis*, Springer-Verlag, New York - Heidelberg - Berlin, 3. Auflage, 1999.
- [3] A. S. Bell, B. Brezger, U. Drodowsky, S. Nowak, T. Pfau, J. Stuhler, Th. Schulze, J. Mlynek, *Nano-Lithography with atoms*, Surf. Sci. **433-435** (1999), p. 40-47.
- [4] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 2nd ed., 1999.
- [5] C. Burstedde, *Computing light masks in neutral atom lithography*, Manuscript, April 2002, Submitted for publication.
- [6] R. Courant, D. Hilbert, *Methods of Mathematical Physics. Volume II: Partial Differential Equations*, Wiley Classics Edition (John Wiley & Sons/Interscience Publishers, New York, 1989).
- [7] U. Drodofsky, J. Stuhler, T. Schulze, M. Drewsen, B. Brezger, T. Pfau, J. Mlynek, *Hexagonal nanostructures generated by light masks for neutral atoms*, Appl. Phys. B **65**, 755 (1997).
- [8] R. Guptka, J. J. McClelland, Z. J. Jabbour, R. Celotta, *Nanofabrication of a two-dimensional array using laser-focused atomic deposition*, Appl. Phys. Lett. **67**, 1378 (1995).
- [9] K. B. Howell, *Principles of Fourier Analysis*, Chapman & Hall / CRC (2001).
- [10] F. John, *Plane Waves and Spherical Means. Applied to Partial Differential Equations*, Springer-Verlag, New York - Heidelberg - Berlin, 1981.
- [11] J. C. Lagarias, J. A. Reeds, M. H. Wright, P. E. Wright, *Convergence properties of the Nelder-Mead simplex algorithm in low dimensions*, SIAM J. Optim. 9 No.1, 112-147 (1998).
- [12] B. S. Logan, L. A. Shepp, *Optimal reconstruction of a function from its projections*, Duke Math. J. **42** (1975), p. 645-659.
- [13] K. I. M. McKinnon, *Convergence of the Nelder-Mead simplex method to a non-stationary point*, May 1996, Submitted for publication.
- [14] H. Metcalf, P. van der Straten, *Laser Cooling and Trapping*, Springer-Verlag, New York - Heidelberg, 1999.
- [15] M. Mützel, U. Rasbach, D. Meschede, C. Burstedde, J. Braun, A. Kunothe, K. Peithmann, K. Buse, *Atomic Nanofabrication with complex light fields*, Appl. Phys. **B 77**, 1-9 (2003).
- [16] F. H. Mok, *Angle-multiplexed storage of 5000 holograms in lithium niobate*, Opt. Lett. **18**, 915 (1993).
- [17] P. P. Petrushev, *Approximation by ridge functions and neural networks*, SIAM J. Math. Anal. **30**, No.1 155-189 (1999).
- [18] P. P. Petrushev, *mündliche Mitteilung*.

- [19] A. Pinkus, *Approximation by ridge functions*, in *Surface fitting and multiresolution methods*, edited by Le Méhauté, Alain et al. (Vanderbilt University Press, Nashville, TN, 1997), p. 279-292.
- [20] Special Issue on nanomanipulation of atoms, *Appl. Phys. B* **70**, 5 (2000), edited by D. Meschede and J. Mlynek.
- [21] J. H. Thywissen, K. S. Johnson, R. Younkin, N. H. Dekker, K. K. Bergren, A. P. Chu, M. Prentiss, S. A. Lee, *Nanofabrication using neutral atomic beams*, *J. Vac. Sci. Technol. B* **15**, 6 (1997).