

Dissertation

Spotzuordnung und
Wellenfront-Rekonstruktion für
Shack-Hartmann-Sensoren

Sascha Groening
Innstr.: 33
94032 Passau
groening@forwiss.uni-passau.de

3. Dezember 2001

Inhaltsverzeichnis

1	Einleitung	5
2	Shack-Hartmann-Sensor	6
2.1	Funktionsprinzip	6
2.2	Einsatzgebiete	9
3	Mathematische Grundlagen	13
3.1	Splines	13
3.1.1	Univariate Splines	13
3.1.2	Bivariate Tensorprodukt-Splines	25
3.2	Beste Approximation im Quadratmittel	30
3.2.1	Lösung linearer Ausgleichsprobleme	31
3.2.2	Lösung linearer Ausgleichsprobleme mit linearen Nebenbedingungen	34
4	Spotdetektion	36
4.1	Optische Grundlagen	36
4.1.1	Grundlagen Fourier-Optik	36
4.1.2	Das Winkelspektrum	37
4.1.3	Propagation des Winkelspektrums	38
4.1.4	Effekt einer beugenden Struktur auf das Winkelspektrum	40
4.1.5	Das Propagations-Phänomen als linearer räumlicher Filter	41
4.1.6	Fresnel Beugung	42
4.1.7	Fraunhofer Beugung	44
4.1.8	Fraunhofer Beugungsmuster einer kreisrunden Apertur	45
4.1.9	Eine Dünne Linse als Phasentransformation	46
4.1.10	Elektromagnetische Grundlagen	49
4.2	Mathematische Modellierung der Spotentstehung	52
4.3	Schwerpunktmethode	55
4.3.1	Schwerpunktmethode mit statischem Schwellwert	55

4.3.2	Schwerpunktmethode mit Zwei-Niveau-Passung	55
4.4	Funktionenpassung	55
4.5	Vergleich der vorgestellten Verfahren	55
5	Spotzuordnung	56
5.1	Bekannte Verfahren zur erweiterten Spotzuordnung	56
5.1.1	Hardware-basierte Methoden	56
5.1.2	Software-basierte Methoden	58
5.2	Spotzuordnung durch iterative Funktionen-Passung	59
5.2.1	Prinzip	60
5.2.2	Mathematische Grundlagen	65
5.2.3	Basisfunktionen	68
5.2.4	Praktische Umsetzung	69
5.2.5	Spot-Zuordnung bei nicht stetig differenzierbaren Wellenfronten	70
5.2.6	Spot-Zuordnung bei unstetigen Wellenfronten	70
5.2.7	Komplexität und Laufzeit	70
5.3	Vergleich mit dem Stand der Technik	72
5.3.1	Statische Spotzuordnung	72
5.3.2	Kontinuierung	74
5.3.3	Spotzuordnung durch iterative Funktionenpassung	75
6	Wellenfront-Rekonstruktion	77
6.1	Bekannte Verfahren zur Wellenfront-Rekonstruktion	77
6.2	Wellenfront-Rekonstruktion durch Spline-Passung	78
6.2.1	Mathematische Grundlagen	78
6.2.2	Praktische Umsetzung	79
6.2.3	Polarkoordinaten-Rekonstruktion	79
6.2.4	Rekonstruktion von nicht stetig differenzierbaren Wellenfronten	82
6.2.5	Rekonstruktion von unstetigen Wellenfronten	82
6.2.6	Komplexität und Laufzeit	82
6.3	Vergleich der vergestellten Verfahren	82
7	Zusammenfassung	94
8	Ergebnisse und Ausblick	95
A	Laufzeiten der implementierten Algorithmen	96

Kapitel 1

Einleitung

Symbolverzeichnis

\mathbb{N}	Menge der natürlichen Zahlen $\{0, 1, \dots\}$
$\mathbb{N}_>$	Menge der positiven natürlichen Zahlen $\{1, 2, \dots\}$
\mathbb{Z}	Menge der ganzen Zahlen
\mathbb{R}	Menge der reellen Zahlen
\mathbb{R}_{\geq}	Menge der nicht negativen reellen Zahlen $\{x \in \mathbb{R} x \geq 0\}$
$\mathbb{R}_{>}$	Menge der positiven reellen Zahlen $\{x \in \mathbb{R} x > 0\}$
$\langle \cdot \cdot \rangle$	Standardskalarprodukt $\langle \cdot \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ definiert als $\langle x y \rangle = \sum_{i=1}^n x_i y_i$ für alle $x, y \in \mathbb{R}^n$
$\ \cdot\ $	Euklidische Norm $\ \cdot\ : \mathbb{R}^n \rightarrow \mathbb{R}$ definiert als $\ x\ = \sqrt{\langle x x \rangle}$ für alle $x \in \mathbb{R}^n$
$\ \cdot\ _{\infty}$	Maximumsnorm $\ \cdot\ _{\infty} : \mathbb{R}^n \rightarrow \mathbb{R}$ definiert als $\ x\ _{\infty} = \max\{ x_i 1 \leq i \leq n\}$ für alle $x \in \mathbb{R}^n$
\mathcal{P}_k	Raum der Polynome vom Grad höchstens k
$\mathcal{C}^k[a, b]$	Raum der auf dem Intervall $[a, b]$ k -mal stetig differenzierbaren Funktionen
$\mathcal{C}(\mathbb{R})$	Raum der auf ganz \mathbb{R} stetigen Funktionen
$\mathcal{F}(\mathbb{R}, \mathbb{R})$	Menge aller Funktionen, die von \mathbb{R} nach \mathbb{R} abbilden
$\text{Aff}(\mathbb{R}^2, \mathbb{R}^2)$	Menge aller affin-linearen Funktionen, die von \mathbb{R}^2 nach \mathbb{R}^2 abbilden
Π_k	Projektion auf die k -te Komponente eines Vektors $\Pi_k : \mathbb{K}^n \rightarrow \mathbb{K}$ für $\mathbb{K} \in \{\mathbb{R}, \mathbb{Z}\}$ und $k \in \{1, \dots, n\}$
$\mathcal{G}_{m,n,g_{max}}$	Menge aller digitalisierten und quantisierten Grauwertbilder auf dem Bildrechteck $\{1, \dots, m\} \times \{1, \dots, n\}$ mit maximalem Grauwert g_{max}
$\mathbb{P}_{k,\lambda}$	Spliner Raum vom Grad k mit Knotenvektor λ
$\mathbb{P}_{k,l,\lambda,\mu}$	Tensorprodukt-Spline-Raum vom Grad k in x - und vom Grad l in y -Richtung mit Knotenvektor λ in x - und μ in y -Richtung
$\lfloor \cdot \rfloor$	$\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{Z}$ definiert als $\lfloor x \rfloor := \max\{n \in \mathbb{Z} n \leq x\}$
$\lceil \cdot \rceil$	$\lceil \cdot \rceil : \mathbb{R} \rightarrow \mathbb{Z}$ definiert als $\lceil x \rceil := \min\{n \in \mathbb{Z} n \geq x\}$
rd	rd: $\mathbb{R} \rightarrow \mathbb{Z}$ Rundung auf die nächste ganze Zahl definiert als $\text{rd}(x) = \lfloor x + \frac{1}{2} \rfloor$ für alle $x \in \mathbb{R}$
$ A $	Mächtigkeit der Menge A
$\{a, \dots, b\}$	Kurzschreibweise für $\{a, a + 1, \dots, b\}$
$\lim_{x \rightarrow a-} f(x)$	Kurzschreibweise für $\lim_{\substack{x \rightarrow a \\ x < a}} f(x)$
$\lim_{x \rightarrow a+} f(x)$	Kurzschreibweise für $\lim_{\substack{x \rightarrow a \\ x > a}} f(x)$
$f(x+)$	Kurzschreibweise für $\lim_{h \rightarrow 0+} f(x+h)$
$f(x-)$	Kurzschreibweise für $\lim_{h \rightarrow 0+} f(x-h)$

Kapitel 2

Shack-Hartmann-Sensor

2.1 Funktionsprinzip

Ein Shack-Hartmann-Sensor dient zur Vermessung von Wellenfronten. Als **Wellenfront** bezeichnet man eine Oberfläche im Raum, die von Punkten gleicher optischer Weglänge auf den von einer Lichtquelle ausgehenden Strahlen gebildet wird. Bei einer idealisierten, punktförmigen Lichtquelle breiten sich die Lichtstrahlen radial und gleichförmig in alle Richtungen aus, damit ergeben sich konzentrische, im Durchmesser anwachsende Kugeln, als Wellenfronten.

Eine Linse dient dazu eine einfallende Wellenfront in einer bestimmten Art und Weise zu verformen. Durch geeignete Linsen kann erreicht werden, daß die Strahlen einer sphärischen oder ebenen Wellenfront in einem Punkt konvergieren. Dieser Punkt wird als **Fokuspunkt** oder **Spot** bezeichnet. Die zu den Achsen der Linse parallele Ebene, in der der Fokuspunkt liegt, nennt man **paraxiale Fokalebene**. Als **optische Achse** einer Linse bezeichnet man eine Gerade durch den Mittelpunkt der Linse, die senkrecht zur Linsenebene steht. Der Auftreffpunkt der optischen Achse in der Fokalebene wird im Folgenden als **Referenzpunkt** bezeichnet.

Will man die Steigung einer einfallenden Wellenfront lokal erfassen, so kann man das durch die Auswertung der lateralen Verschiebung des Fokus einer im Strahlengang positionierten Linse, erreichen. Diese Fokus-Verschiebung wird allerdings nur Aussagen über die mittlere Steigung der Wellenfront über der Apertur der Linse zulassen. Man wird also die Apertur der Linse möglichst klein wählen, um die einfallende Wellenfront möglichst fein aufzulösen. Dabei ist abgesehen von den physikalischen Problemen der Herstellung einer solchen

Linse ein Kompromiß zwischen möglichst kleiner Apertur und ausreichender Helligkeit des Fokuspunkts zu finden. Zur Vermessung einer einfallenden Welle an verschiedenen Punkten kann eine in einer Ebene verschiebbare Linse verwendet werden. Vorteil ist dabei, daß die Wellenfront sehr feinaufgelöst werden kann. Nachteil ist die Notwendigkeit einer hochgenauen Verfahrenheit, die Notwendigkeit mehrerer Messungen, was natürlich eine sich in ihrer Form möglichst konstante einfallende Welle verlangt. Eine Alternative bietet die Verwendung einer festen Anordnung von mehreren Linsen gleicher Brennweite in einer Ebene. Einen solchen Sensor nennt man **Shack-Hartmann-Sensor** oder kurz **SHS**. Wird anstelle eines Mikrolinsen-Arrays ein Array von Löchern gleichen Durchmessers verwendet, so spricht man von einem **Hartmann-Sensor**. Er wird dann verwendet, wenn die Energie des einfallenden Lichts hoch genug, oder vielleicht sogar zu hoch ist. Dies ist beispielsweise bei der Vermessung von Hoch-Energie-Lasern der Fall.

Die Anordnung der Linsen beim SHS ist im Prinzip frei wählbar. Im Interesse schneller Rekonstruktions-Algorithmen hat sich im FMSII-Projekt Wellensensor eine Anordnung der Linsen auf den Gitterpunkten eines rechtwinkligen, äquidistanten Gitters bewährt. Ein solches Array von kleinen Linsen wird als **Mikrolinsenarray** bezeichnet. Im Projekt wurden Mikrolinsen-Arrays mit 42 mal 42 kreisrunden Linsen mit Durchmesser 0.15mm verwendet. Die Zwischenräume zwischen den Linsen und die das Mikrolinsen-Array umgebende Fläche sollte durch Blenden abgedeckt sein, um eine Verfälschung des Meßergebnisses durch das Licht, das nicht durch eine Linse fällt, zu vermeiden. Dadurch wird eine einfallende Wellenfront in bestimmte einzelne Bereiche aufgeteilt. Man erhält so gleich ein ganzes Array lokaler Fokalverschiebungen (Siehe Abb. 2.1.), die Aussagen über das mittlere Steigungsverhalten der einfallenden Wellenfront über den Mikrolinsen zulassen. Zum Zwecke einfacher Digitalisierung kann man das Mikrolinsen-Array auf einen CCD-Chip aufkleben. Dies zusammen mit der Tatsache, daß keine bewegliche Teile notwendig sein, gewährleistet die erwähnte hohe mechanische Stabilität des SHS. Jeder Mikrolinse ist ein zwei-dimensionaler Index (**Mikrolinsenindex**) zugeordnet, der beschreibt in welcher Zeile und Spalte in dem Array sich diese Mikrolinse befindet. Eine Linse in einem Mikrolinsen-Array bezeichnet man auch als Subapertur. Im Folgenden wird, wie in der Literatur üblicherweise auch, der durch die Projektion einer Mikrolinse in die Fokalebene definierte Bereich als **Subapertur** bezeichnet.

Als Resultat der Hindurchleitung einer Wellenfront durch ein Mikrolinsenarray erhält man je Mikrolinse einen Spot. Diese Anhäufung von Spots wird als **Spotfeld** oder **Spotbild** bezeichnet. Im Falle großer Wellenfrontsteigungen

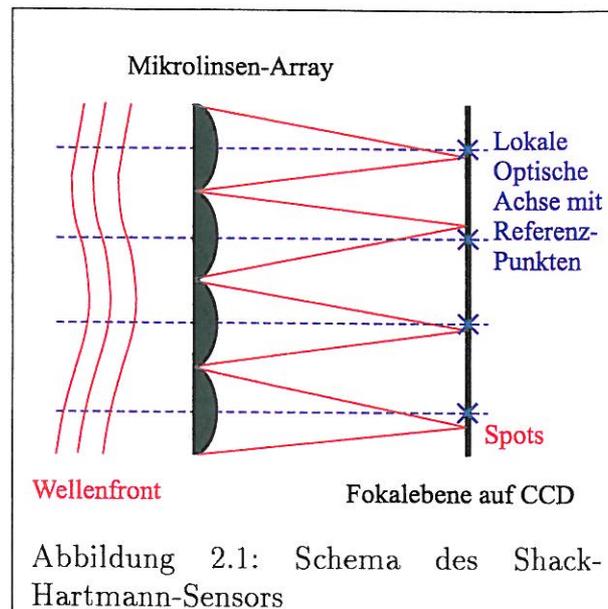
können Spots ihre Subapertur verlassen. Das Zuordnungsproblem von Spot zu Mikrolinse ist dadurch nicht mehr trivial zu lösen.

Ein Shack-Hartmann-Sensor mißt eine Wellenfront, in dem er sie mit einem Mikrolinsen-Array sampelt. In der Fokalebene der Mikrolinsen zeigt sich ein charakteristisches Spotmuster. Die horizontale und vertikale Ablenkung $\sigma_i \in \mathbb{R}^2$ eines Spots von der optischen Achse der jeweiligen Mikrolinse gibt ein Maß für die mittlere Wellenfrontsteigung $w_i \in \mathbb{R}^2$ über dieser Mikrolinse an. Sie ist definiert als

$$w_i := \frac{\sigma_i}{\sqrt{f^2 + \Pi_1(\sigma_i) + \Pi_2(\sigma_i)}},$$

wobei $f \in \mathbb{R}$ die Fokallänge der Mikrolinsen ist und $\Pi_j : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert sei als die Projektion auf die j -te Komponente eines Vektors. Die Herleitung dieser Formel kann in [LPS00] nachgeschlagen werden. Es wird angenommen, daß die lokalen Krümmungen der Wellenfront über einer Mikrolinse klein sind. Die mittleren Wellenfrontsteigungen w_i werden dem optischen Zentrum einer Mikrolinse zugeordnet. Aus dem sich ergebenden Feld von partiellen Ableitungen kann die einfallende Wellenfront entweder lokal oder global rekonstruiert werden. Verfahren zur lokalen Rekonstruktion werden in [MM92] beschrieben. Ein Vergleich zwischen lokalen und globalen Rekonstruktions-Algorithmien ist in [Sou80] zu finden. In dieser Arbeit beschränken wir uns auf globale Wellenfront-Rekonstruktions-Algorithmien.

Durch das Anfitten eines Splines oder Polynoms an das Feld von partiellen Ableitungen kann wiederum eine stetige oder sogar k -mal differenzierbare Wellenfront rekonstruiert werden, die je nach Wahl des Funktionenraums eine mehr oder weniger gute Näherung der realen Wellenfront darstellt. Zunächst ist allerdings eine eindeutige Zuordnung eines jeden Spots zu seiner jeweiligen Mikrolinse notwendig.



Wenn die lokalen Steigungen einer Wellenfront über dem Mikrolinsen-Array zu groß sind, verlassen die Spots ihre jeweiligen Subaperturen. Der oben beschriebene Effekt stellt sich um so leichter ein, je empfindlicher der Sensor eingestellt ist, also je größer die Fokallänge der Mikrolinsen im Verhältnis zum Subapertur-Durchmesser gewählt wird. Bei Vergrößerung der Fokallänge der Mikrolinse und gleichbleibendem Subapertur-Durchmesser, vergrößert sich auch der „Hebel“ einer Mikrolinse, damit werden kleinere Wellenfrontsteigungsunterschiede meßbar. In diesem Fall ist eine eindeutige Zuordnung von Spot zu Mikrolinse nur mit speziellen Techniken durchführbar.

2.2 Einsatzgebiete

In den letzten Jahren hat sich der Shack-Hartmann-Sensor (SHS) in der optischen Meßtechnik etabliert. Die Fortschritte in der Mikrooptik und der Kamera-Technik haben dem SHS ein weites Anwendungsgebiet erschlossen. Der SHS zeichnet sich dabei vor allem durch die Einfachheit des Meßprinzips und eine hohe mechanische Stabilität aus (Siehe [PLS⁺00]).

- Zunächst fand der SHS seine Anwendung in der Kompensation atmosphärischer Störungen in der Astronomie. Die durch die Atmosphäre

verursachten Wellenfront-Deformationen wurden mit einem SHS gemessen und mittels Algorithmen oder adaptiver Optiken eliminiert. Die hohe Auswerte-Geschwindigkeit des SHS war hier mitentscheidend für seinen Erfolg. Die Anzahl der Mikrolinsen war allerdings klein (ca. 100) (Siehe [RSW94]).

- In [Man92] wird ein Verfahren zur Ausrichtung der Elemente zusammengesetzter Spiegel („Segmented Mirrors“) skizziert, die vor allem in großen Teleskopen eingesetzt werden. Jedes Spiegel-Element wird durch traditionelle Shack-Hartmann-Sensorik vermessen. In einer zweiten Stufe werden Beugungsmuster von Teilen der Wellenfront, die von der Stufe zwischen zwei Spiegel-Elementen stammen ausgewertet, um die Ausrichtung der Elemente entlang der optischen Achse durchzuführen.
- Messung des Brechungs-Index und tomographische Rekonstruktion von durchsichtigen Körpern oder Gasen z.B. zur Messung von Gas-Strömungen [RWG⁺95] oder von durchsichtigen, mikroskopisch kleinen Objekten [ZH97].
- Vermessung des Querschnitts von Laserstrahlen (Siehe [LPS99]). Ein zunehmendes Qualitätsbewußtsein in der produzierenden Industrie erfordert eine ständige Überwachung der Produktionsmittel. Dies gilt insbesondere auch für die Lasermaterialbearbeitung, bei der eine fortwährende Kontrolle der Laserstrahlparameter für eine hohe Qualität des Bearbeitungsergebnisses notwendig ist.
- Vermessung rauher Oberflächen in streifender Inzidenz (Siehe [PLS99]).

In den letzten Jahren fand der SHS durch die Entwicklung von CCD-Kameras mit einer großen Anzahl von Pixeln zunehmend Anwendung in der Prüfung optischer Oberflächen. Dabei können Mikrolinsen-Arrays mit 1000 Linsen und mehr verwendet werden. In der optischen Prüftechnik wurde der SHS nicht zuletzt wegen seiner hohen mechanischen Stabilität und der Einfachheit des Meßprinzips angewendet. Im Rahmen des Forschungsverbunds FORMIKROSYS II wurden im Teilprojekt „Entwicklung eines Meßverfahrens mit hohem Dynamikbereich für die Qualitätssicherung von optischen Asphären und für die in-situ Messung von Wellenfronten (kurz Wellensensor)“ drei verschiedene optische Aufbauten zur Messung von rotationssymmetrischen Asphären, Toroiden und Gleitsichtgläsern entwickelt und getestet.

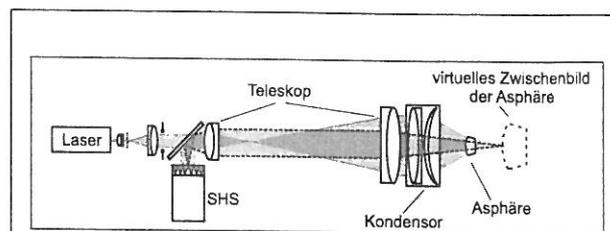


Abbildung 2.2: Asphären-Meßplatz - Prüfung von rotationssymmetrischen Asphären unter Verwendung einer sphärischen Kompensation

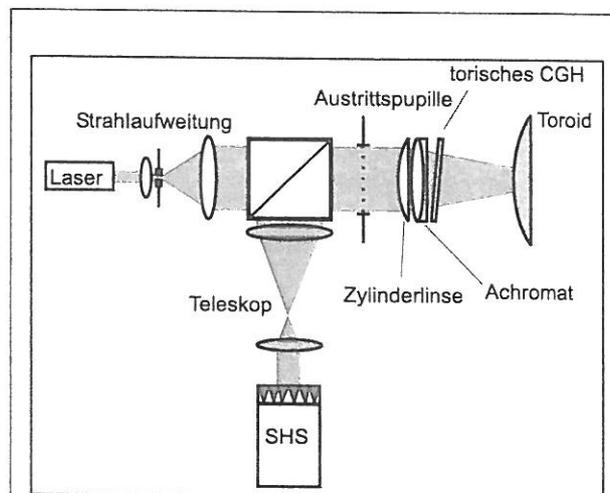
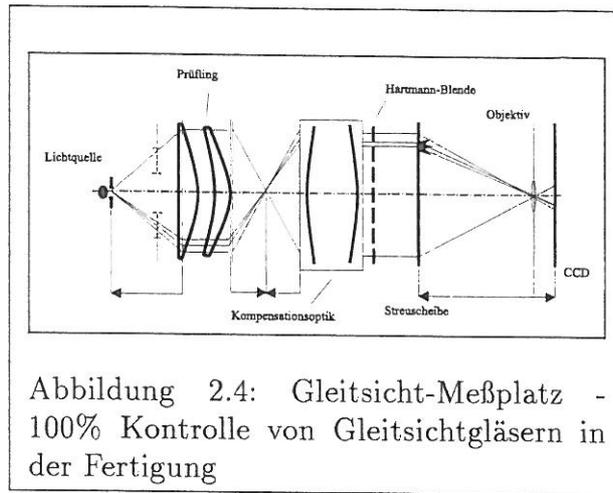


Abbildung 2.3: Toroid-Meßplatz - Prüfung von Toroiden ohne passendes computer-generiertes Hologramm



Kapitel 3

Mathematische Grundlagen

3.1 Splines

3.1.1 Univariate Splines

Definition 3.1

Eine Funktion $s : I \rightarrow \mathbb{R}$ definiert auf einem kompakten Intervall $I := [a, b]$ ($a, b \in \mathbb{R}$) mit streng monoton steigender **Break-Point-Sequenz** $\tau_0, \dots, \tau_n \in \mathbb{R}$ mit $\tau_0 := a, \tau_n := b, n \in \mathbb{N}_{>}$ und **Glattheitsbedingungen** $\nu_1, \dots, \nu_{n-1} \in \{0, \dots, k\}$ wird **univariater polynomialer Spline** (im folgenden kurz **Spline**) vom Grad $k \in \mathbb{N}$ genannt, wenn die folgenden Bedingungen erfüllt sind:

1. Auf jedem Knotenintervall $[\tau_j, \tau_{j+1}[$ für $j = 0, \dots, n - 1$ ist s gegeben als ein Polynom vom Grad höchstens k .

$$s|_{[\tau_j, \tau_{j+1}[} \in \mathcal{P}_k, \quad \text{für } j = 0, \dots, n - 1. \quad (3.1)$$

2. Für $i \in \{1, \dots, n - 1\}$ ist die Funktion s an der Stelle τ_i $(\nu_i - 1)$ -mal stetig differenzierbar, falls $\nu_i \geq 2$ und stetig, falls $\nu_i = 1$.

◇

Für $k = 0$ handelt es sich bei s um eine stückweise konstante Funktion. Dieser Fall interessiert uns aber nicht weiter. Deshalb haben Splines im folgenden implizit alle einen Grad $k > 0$.

Sei $g := \sum_{i=1}^{n-1} ((k + 1) - \nu_i)$. In der Regel werden die Glattheitsbedingungen durch Knotenvervielfachungen in einem **Knotenvektor** $\lambda = (\lambda_0, \dots, \lambda_{g+1}) \in \mathbb{R}^{g+2}$ kodiert. Die Elemente dieses Vektors heißen **Knoten**, wobei $\lambda_1, \dots, \lambda_g$ **innere Knoten** genannt werden.

Der Einfachheit halber werden Knotenvektoren im folgenden immer als Zeilenvektoren notiert. Der Knotenvektor λ wird folgendermaßen aus den Break-Points τ_0, \dots, τ_n und den Glattheitsbedingungen ν_1, \dots, ν_{n-1} generiert: Sei $(r_1 \dots, r_n)^T \in \mathbb{R}^n$ rekursiv definiert als

1. $r_1 := 1$
2. $r_{i+1} := r_i + ((k + 1) - \nu_i)$ für $i \in \{1, \dots, n - 1\}$.

Dabei entspricht r_i für $i \in \{1, \dots, n\}$ dem ersten Index, an dem der Knoten τ_i im Knotenvektor λ auftaucht. Dann wird der Knotenvektor λ durch

1. $\lambda_0 := \tau_0$
2. $\lambda_{r_i} = \dots = \lambda_{r_{i+1}-1} := \tau_i$ für $i \in \{1, \dots, n - 1\}$
3. $\lambda_{g+1} := \tau_n$

definiert. Siehe [Die93] und [?].

Im folgenden sei, falls nicht explizit anders angegeben, das Definitions-Intervall I , der Splinegrad k , die Anzahl der inneren Knoten g und der Knotenvektor $\lambda = (\lambda_0, \dots, \lambda_{g+1})$ wie oben definiert.

Definition 3.2

Der Vektorraum der Funktionen, die die Bedingungen aus Definition 3.1 erfüllen, wird **Splinerraum** vom Grad k mit Knotenvektor λ genannt. Er wird mit $\mathbb{P}_{k,\lambda}$ bezeichnet.

◇

Satz 3.3

Der Splinerraum $\mathbb{P}_{k,\lambda}$ hat die Dimension

$$\dim(\mathbb{P}_{k,\lambda}) = g + k + 1. \quad (3.2)$$

Beweis siehe [Die93].

□

Beispiel 3.4

Seien $k = 2$, $n = 3$, die Break-Points τ_0, \dots, τ_3 durch $(\tau_0, \tau_1, \tau_2, \tau_3) := (0, 1, 2, 3)$ und die Glattheitsbedingungen $\nu_1 := 0$ und $\nu_2 := 1$ gegeben. Wir berechnen $g := \sum_{i=1}^{n-1} ((k + 1) - \nu_i) = 5$ und den Vektor $(r_1, r_2, r_3) := (1, 4, 6)$. Damit ist der Knotenvektor $\lambda = (\lambda_0, \dots, \lambda_{g+1}) := (0, 1, 1, 1, 2, 2, 3)$. Man sieht also, daß eine parabolische Splinefunktion $s \in \mathbb{P}_{2,\lambda}$ wegen der Knotenverdreifung bei Eins eine Unstetigkeitsstelle bei Eins haben muß.

◇

In der Praxis verwendet man meist die B-Spline-Basis. B-Splines werden in der Literatur üblicherweise über dividierte Differenzen definiert.

Definition 3.5

Die k -te **dividierte Differenz** einer genügend oft (siehe unten) differenzierbaren Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ an den Stellen $s_0, \dots, s_k \in \mathbb{R}$ ist definiert als der führende Koeffizient (d.h. der Koeffizient von x^k) des eindeutig bestimmten Polynoms $p \in \mathcal{P}_k(\mathbb{R})$ für das gilt:

$$f^{(j)}(s_i) = p^{(j)}(s_i) \quad \text{für alle } i \in \{0, \dots, k\} \\ j \in \{0, \dots, |\{l \in \{0, \dots, k\} \mid s_l = s_i\}| - 1\}$$

Hierbei bezeichnet $f^{(j)}$ die j -te Ableitung der Funktion f für $j > 0$ und $f^{(0)} = f$.

Die angegebene dividierte Differenz wird mit $[s_0, \dots, s_k]f$ bezeichnet. Siehe [?].

◇

Definition 3.6

Sei $k, n \in \mathbb{N}$ mit $n > k$ und $\lambda = (\lambda_0, \dots, \lambda_n) \in \mathbb{R}^{n+1}$ mit $\lambda_i < \lambda_{i+k+1}$ für $i \in \{0, \dots, n-k-1\}$. Weiter ist $f_x : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $f_x(y) = (\max\{0, y-x\})^k$. Dann ist der (normalisierte) **B-Spline**¹ $N_{i,k+1} : \mathbb{R} \rightarrow \mathbb{R}$ vom Grad k mit der **Knotensequenz** $\lambda_i, \dots, \lambda_{i+k+1}$ für $i \in \{0, \dots, n-k-1\}$ definiert als

$$N_{i,k+1}(x) = (\lambda_{i+k+1} - \lambda_i)[\lambda_i, \dots, \lambda_{i+k+1}]f_x$$

◇

Um die Notation knapp zu halten, wird die zu einem B-Spline $N_{i,k+1}$ gehörige Knotensequenz $\lambda_i, \dots, \lambda_{i+k+1}$ nicht explizit notiert. Eine solche Funktion $N_{i,k+1}$ ist ein Spline vom Grad k mit dem Knotenvektor $(\lambda_i, \dots, \lambda_{i+k+1})$. Beweis siehe [Die93]. Es handelt sich also streng genommen einfach um einen Vektor eines Funktionenraums, der üblicherweise keinen eigenen Namen trägt. Da die Bezeichnung B-Spline aber in der Literatur durchwegs gebraucht wird, soll sie auch hier verwendet werden.

Will man einen B-Spline an einer bestimmten Stelle $x \in \mathbb{R}$ auswerten, wird man aus Rücksicht auf die numerische Stabilität des Verfahrens nicht die obige Definition heranziehen.

¹B-Spline ist eine Abkürzung für Basis-Spline.

Satz 3.7

Sei k, n, λ wie in Definition 3.6. Für $i \in \{0, \dots, n - k - 1\}$ kann ein B-Spline $N_{i,k+1} : \mathbb{R} \rightarrow \mathbb{R}$ vom Grad $k \in \mathbb{N}$ mit der streng monoton steigenden **Knotensequenz** $\lambda_i, \dots, \lambda_{i+k+1}$ an einer Stelle $x \in \mathbb{R}$ durch die Rekursionsformel von DE BOOR und COX

$$\begin{aligned} N_{i,1}(x) &= \begin{cases} 1, & \text{falls } x \in [\lambda_i, \lambda_{i+1}[\\ 0, & \text{falls } x \notin [\lambda_i, \lambda_{i+1}[\end{cases} \\ N_{i,k+1}(x) &= \frac{x-\lambda_i}{\lambda_{i+k}-\lambda_i} N_{i,k}(x) + \frac{\lambda_{i+k+1}-x}{\lambda_{i+k+1}-\lambda_{i+1}} N_{i+1,k}(x) \end{aligned} \quad (3.3)$$

berechnet werden. Beweis siehe [Die93].

□

Bemerkung 3.8

Sei k, n, λ wie in Definition 3.6. Will man statt der in Satz 3.7 geforderten streng monoton steigenden Knotensequenz $\lambda_i, \dots, \lambda_{i+k+1}$ für $i \in \{0, \dots, n - k - 1\}$ auch monoton steigende Knotensequenzen mit $\lambda_i < \lambda_{i+k+1}$ zulassen, so muß man die Rekursionsformel 3.3 leicht abwandeln. Um dies einzusehen, versuche man zum Beispiel mit obiger Rekursionsformel für $k := 2$ und $\lambda = (\lambda_0, \lambda_1, \lambda_2, \lambda_3) := (0, 1, 1, 2)$ den Wert $N_{0,3}(\frac{1}{2})$ zu berechnen. Das Ergebnis ist ein „Division durch Null“-Fehler. Eine Tatsache die so explizit in der Literatur nicht erwähnt wird.

Folgendermaßen modifiziert kann die obige Rekursionsformel aber auch für die Auswertung von B-Splines mit monoton steigenden Knotensequenzen $\lambda_i, \dots, \lambda_{i+k+1}$ mit $\lambda_i < \lambda_{i+k+1}$ verwendet werden.

$$N_{i,k+1}(x) = \begin{cases} 1 & , \text{ falls } k = 0 \text{ und } x \in [\lambda_i, \lambda_{i+1}[\\ 0 & , \text{ falls } k = 0 \text{ und } x \notin [\lambda_i, \lambda_{i+1}[\\ c_{i,k}(x) + d_{i,k}(x) & , \text{ sonst} \end{cases} \quad (3.4)$$

Dabei sei $c_{i,k} : \mathbb{R} \rightarrow \mathbb{R}$ und $d_{i,k} : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$c_{i,k}(x) := \begin{cases} 0 & , \text{ falls } N_{i,k}(x) = 0 \\ \frac{x-\lambda_i}{\lambda_{i+k}-\lambda_i} N_{i,k}(x) & , \text{ sonst } ^2. \end{cases}$$

und

$$d_{i,k}(x) := \begin{cases} 0 & , \text{ falls } N_{i+1,k} = 0 \\ \frac{\lambda_{i+k+1}-x}{\lambda_{i+k+1}-\lambda_{i+1}} N_{i+1,k}(x) & , \text{ sonst } ^2. \end{cases}$$

Im folgenden seien als Knotensequenz $\lambda_i, \dots, \lambda_{i+k+1}$ für einen B-Spline $N_{i,k+1}$ immer auch monoton steigende Knotensequenzen mit $\lambda_i < \lambda_{i+k+1}$ zugelassen.

²Hier definieren wir also $\frac{0}{0} = 0$.

Sei $\lambda_i = \dots = \lambda_{i+k} < \lambda_{i+k+1}$, dann hat $N_{i,k+1}$ an λ_i eine Unstetigkeitsstelle. Es gilt $N_{i,k+1}(\lambda_i+) = N_{i,k+1}(\lambda_i)$.

Sei $\lambda_i < \lambda_{i+1} \dots = \lambda_{i+k+1}$, dann hat $N_{i,k+1}$ an λ_{i+1} eine Unstetigkeitsstelle. Es gilt $N_{i,k+1}(\lambda_{i+1}-) = N_{i,k+1}^-(\lambda_{i+1})$ mit

$$N_{i,k+1}^-(x) = \begin{cases} 1 & , \text{ falls } k = 0 \text{ und } x \in]\lambda_i, \lambda_{i+1}] \\ 0 & , \text{ falls } k = 0 \text{ und } x \notin]\lambda_i, \lambda_{i+1}] \\ c_{i,k}^-(x) + d_{i,k}^-(x) & , \text{ sonst} \end{cases} \quad (3.5)$$

Dabei sei $c_{i,k}^- : \mathbb{R} \rightarrow \mathbb{R}$ und $d_{i,k}^- : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$c_{i,k}^-(x) := \begin{cases} 0 & , \text{ falls } N_{i,k}^-(x) = 0 \\ \frac{x-\lambda_i}{\lambda_{i+k}-\lambda_i} N_{i,k}^-(x) & , \text{ sonst }^3. \end{cases}$$

und

$$d_{i,k}^-(x) := \begin{cases} 0 & , \text{ falls } N_{i+1,k}^- = 0 \\ \frac{\lambda_{i+k+1}-x}{\lambda_{i+k+1}-\lambda_{i+1}} N_{i+1,k}^-(x) & , \text{ sonst }^3. \end{cases}$$

◇

Satz 3.9

Mit dem Knotenvektor $\lambda = (\lambda_0, \dots, \lambda_{g+1}) \in \mathbb{R}^{g+2}$ können wir $(g - k + 1)$ linear unabhängige Splines vom Grad k erzeugen, nämlich: $N_{i,k+1}$ für $i \in \{0, \dots, g - k\}$.

Beweis siehe [dB78].

□

Bemerkung 3.10

Der Vektorraum $\mathbb{P}_{k,\lambda}$ hat nach Satz 3.3 die Dimension $g + k + 1$. Für eine vollständige Basis des Vektorraums $\mathbb{P}_{k,\lambda}$ fehlen also noch $2k$ linear unabhängige Splines.

Durch Definition zusätzlicher Knoten $\lambda_{-k}, \dots, \lambda_{-1} \in \mathbb{R}$ und $\lambda_{g+2}, \dots, \lambda_{g+k+1} \in \mathbb{R}$ am Rand von λ , mit

$$\begin{aligned} \lambda_{-k} &\leq \lambda_{-k+1} \leq \dots \leq \lambda_{-1} \leq \lambda_0 = a, \\ b = \lambda_{g+1} &\leq \lambda_{g+2} \leq \dots \leq \lambda_{g+k} \leq \lambda_{g+k+1}, \end{aligned} \quad (3.6)$$

die aber ansonsten beliebig sind, können die fehlenden Basisfunktionen als $N_{i,k+1}$ für $i = -k, \dots, -1$ und $i = g - k + 1, \dots, g$ erzeugt werden [Die93].

In den meisten Fällen wird in der Literatur

$$\lambda_{-k} = \dots = \lambda_{-1} = \lambda_0 = a \quad \text{und} \quad b = \lambda_g = \lambda_{g+1} = \dots = \lambda_{g+k+1}$$

gewählt [?].

◇

Im folgenden seien $\lambda_{-k}, \dots, \lambda_{-1}$ und $\lambda_{g+2}, \dots, \lambda_{g+k+1}$ wie in Bemerkung 3.10 definiert.

Satz 3.11**B-Spline-Basis**

Mit den wie in Bemerkung 3.10 definierten zusätzlichen Knoten, erhält jeder Spline $s \in \mathbb{P}_{k,\lambda}$ eine eindeutig bestimmte Darstellung der Form

$$s(x) = \sum_{i=-k}^g \xi_i N_{i,k+1}(x). \quad (3.7)$$

Beweis siehe [Die93].

□

Definition 3.12

Die Menge $\{N_{i,k+1} \mid i = -k, \dots, g\}$ nennt man **B-Spline-Basis**.

Die ξ_i ($\xi_i \in \mathbb{R}$ für $i = -k, \dots, g$) werden **B-Spline-Koeffizienten** genannt.

◇

Streng genommen ist ein Spline $s \in \mathbb{P}_{k,\lambda}$ nur auf $[\lambda_0, \lambda_{g+1}]$ definiert. Also müßten die Basisfunktionen ebenfalls auf dieses Intervall eingeschränkt werden. Um die Notation kurz zu halten, identifizieren wir aber einen Spline mit der Fortsetzung auf ganz \mathbb{R} , die über obige Formel gegeben wird, und behalten im Gedächtnis, daß mit der B-Spline-Basis streng genommen die Menge der obigen Funktionen eingeschränkt auf I gemeint ist.

Bemerkung 3.13

Diese Basis hat die interessante Eigenschaft eine Partition der Eins auf I zu bilden.⁴

$$\sum_{i=-k}^g N_{i,k+1}(x) = 1 \quad \text{für alle } x \in I \quad (3.8)$$

Siehe [Die93].

Die Zusicherung dieser Eigenschaft, kann unter Umständen auch Probleme aufwerfen. Siehe Bemerkung 3.17.

◇

Bemerkungen 3.14

Wir wollen noch einige grundlegende Eigenschaften der B-Splines anführen. Die zugehörigen Beweise finden sich in [dB78] und [dB90].

Sei jeweils $i \in \{-k, \dots, g\}$.

⁴Deswegen heißen diese Basisfunktionen auch **normalisiert**.

1. **Positivität**

$$N_{i,k+1}(x) \geq 0 \text{ für alle } x \in \mathbb{R}. \quad (3.9)$$

 2. **Beschränkter Träger**

$$N_{i,k+1}(x) = 0 \text{ für alle } x \notin [\lambda_i, \lambda_{i+k+1}]. \quad (3.10)$$

 3. **Randwerte**

Sei $1 \leq r < k + 1$, dann gilt für $l = 0, \dots, k - r$

$$\begin{aligned} N_{i,k+1}^{(l)}(\lambda_i) &= 0, \text{ falls } \lambda_i = \dots = \lambda_{i+r-1} < \lambda_{i+r} \\ N_{i,k+1}^{(l)}(\lambda_{i+k+1}) &= 0, \text{ falls } \lambda_{i+k+1-r} < \lambda_{i+k+2-r} = \dots = \lambda_{i+k+1} \end{aligned} \quad (3.11)$$

Speziell gilt, falls $\lambda_i < \dots < \lambda_{i+k+1}$ streng monoton steigend ist,

$$N_{i,k+1}^{(l)}(\lambda_i) = N_{i,k+1}^{(l)}(\lambda_{i+k+1}) = 0, \text{ für } l \in 0, \dots, k - 1. \quad (3.12)$$

Die Funktionen $N_{i,k+1}$ fallen also am Rand von $[\lambda_i, \lambda_{i+k+1}]$ entsprechend glatt gegen Null ab.

 4. **Erste Ableitung**

Stetige Differenzierbarkeit an der Stelle $x \in \mathbb{R}$ vorausgesetzt gilt

$$N'_{i,k+1}(x) = \begin{cases} 0 & , \text{ falls } k = 0 \\ k(c_{i,k}(x) - d_{i,k}(x)) & , \text{ sonst} \end{cases} \quad (3.13)$$

mit $c_{i,k} : \mathbb{R} \rightarrow \mathbb{R}$ und $d_{i,k} : \mathbb{R} \rightarrow \mathbb{R}$ definiert als

$$c_{i,k}(x) := \begin{cases} 0 & , \text{ falls } N_{i,k}(x) = 0 \\ \frac{N_{i,k}(x)}{\lambda_{i+k} - \lambda_i} & , \text{ sonst}^5 \end{cases}$$

und

$$d_{i,k}(x) := \begin{cases} 0 & , \text{ falls } N_{i+1,k}(x) = 0 \\ \frac{N_{i+1,k}(x)}{\lambda_{i+k+1} - \lambda_{i+1}} & , \text{ sonst}^5 \end{cases}$$

Die erste Ableitung eines B-Spline vom Grad k ist ein Spline vom Grad $k - 1$ mit demselben Knotenvektor.

Ist $N_{i,k+1}$ an der Stelle x nicht differenzierbar erhält man den rechten

⁵Hier definieren wir also $\frac{0}{0} = 0$.

Auswertungspunkt $x \in [\lambda_{-k}, \lambda_{g+k+1}]$
 $i \in \{-k, \dots, g+k\}$ mit $x \in [\lambda_i, \lambda_{i+1}[$, bzw. $x \in]\lambda_i, \lambda_{i+1}]$,
siehe Bemerkungen weiter unten.

Ausgabe: Alle Basisfunktionen am Punkt x ausgewertet

$$(N_{-k,k+1}(x), \dots, N_{g,k+1}(x)) \in \mathbb{R}^{g+k+1}$$

Sei $A = (a_{j,l})_{\substack{j \in \{-k, \dots, g+k\} \\ l \in \{0, \dots, k\}}}$ eine Familie reeller Zahlen.

Initialisiere $a_{j,l} := 0$ für alle $j \in \{-k, \dots, g-k\}$ und $l \in \{0, \dots, k\}$

Setze $a_{i,0} := 1$

forall ($l \in \{1, \dots, k\}$) **do**

Berechne Spalte l des Dreieckschemas.

forall ($j \in \{\min\{i, g+k-l\}, \dots, \max\{i-l, -k\}\}$) **do**

Berechne Eintrag j in Spalte l des Dreieckschemas.

Teste, ob erster Summand überhaupt berechnet werden muß:

if ($j \geq (i-l+1)$ und $j \leq i$) **then**

Berechne ersten Summanden gemäß der Rekursionsformel in Bemerkung 3.8.

if ($a_{j,l-1} \neq 0$) **then**

$$temp1 := \frac{x-\lambda_j}{\lambda_{j+l}-\lambda_j} \cdot a_{j,l-1}$$

else

$$temp1 := 0$$

end if

else

$$temp1 := 0$$

end if

Teste, ob zweiter Summand überhaupt berechnet werden muß:

if ($j \geq (i-l)$ und $j \leq (i-1)$) **then**

Berechne zweiten Summanden gemäß der Rekursionsformel in Bemerkung 3.8

if ($a_{j+1,l-1} \neq 0$) **then**

$$temp2 := \frac{\lambda_{j+l+1}-x}{\lambda_{j+l+1}-\lambda_{j+1}} \cdot a_{j+1,l-1}$$

else

$$temp2 := 0$$

end if

else

$$temp2 := 0$$

end if

$$a_{j,l} := temp1 + temp2$$

end while

end while

return $(a_{-k,k}, \dots, a_{g,k})$

Möchte man einen Spline $s \in \mathbb{P}_{k,\lambda}$ an einer Stelle auswerten, an der $k + 1$ Knoten zusammenfallen, so hat der Spline dort eventuell eine Unstetigkeitsstelle. Sei $k + 1 \leq l \leq g$ und $\lambda_{l-k-1} < \lambda_{l-k} = \dots = \lambda_l < \lambda_{l+1}$, dann gilt wegen Eigenschaft 3.10

$$s(\lambda_l+) = \sum_{j=l-k}^l \xi_j N_{j,k+1}(\lambda_l).$$

Der obige Algorithmus angewendet mit $i = l$ liefert alle Werte, die zur Berechnung des Grenzwertes von rechts $s(\lambda_l+)$ notwendig sind.

Für den Grenzwert von links gilt wegen Eigenschaft 3.10

$$s(\lambda_l-) = \sum_{j=l-2k-1}^{l-k-1} \xi_j N_{j,k+1}^-(\lambda_l).$$

Obiger Algorithmus angewendet mit $i = l - k - 1$ liefert alles zur Berechnung des Grenzwertes von links $s(\lambda_l-)$ notwendige.

Es gibt noch andere unter Umständen auch effizientere Algorithmen zur Auswertung eines Splines an einer bestimmten Stelle. Für unsere Zwecke ist eine Implementierung obigen Schemas aber hinreichend effizient.

◇

Beispiel 3.16

Sehen wir uns die Basisfunktionen eines Splineraums etwas näher an.

Sei $k := 2, g := 3, (\lambda_{-2}, \lambda_{-1}, \lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) := (-2, -1, 0, 1, 1, 1, 2, 3, 4)$, $a := \lambda_0 = 0$ und $b := \lambda_4 = 2$, dann ist $\{N_{i,3} | i \in \{-2, \dots, 3\}\}$ eine Basis von $\mathbb{P}_{2,(0,1,1,1,2)}$.

Abbildung 3.1: Basisfunktionen von $\mathbb{P}_{2,(0,1,1,1,2)}$

In Bild 3.1 sieht man sehr gut, daß die Basisfunktionen aus Parabeln zusammengesetzt sind.

Der Spline $N_{-2,3}$ hat $(-2, -1, 0, 1)$ als Knotenvektor. Es gilt also $N_{-2,3}(x) \geq 0$ für alle $x \in [-2, 1]$ und $N_{-2,3}(x) = 0$ sonst. $N_{-2,3}$ ist auf ganz \mathbb{R} einmal stetig differenzierbar.

Der Spline $N_{-1,3}$ hat $(-1, 0, 1, 1)$ als Knotenvektor. Es gilt also $N_{-1,3}(x) \geq 0$ für alle $x \in [-1, 1]$ und $N_{-1,3}(x) = 0$ sonst. Er ist an der Stelle 1 zwar stetig, aber nicht differenzierbar. Der Graph der Funktion hat an dieser Stelle also einen Knick.

Der Spline $N_{0,3}$ hat $(0, 1, 1, 1)$ als Knotenvektor. Es gilt also $N_{0,3}(x) \geq 0$ für alle $x \in [0, 1]$ und $N_{0,3}(x) = 0$ sonst. Diese Funktion ist an der Stelle 1 unstetig. Es gilt $N_{0,3}(1) = 0$. Der Grenzwert von links $N_{0,3}(1-)$ ist 1 der Grenzwert von rechts $N_{0,3}(1+)$ ist 0.

Bei der Funktion $N_{1,3}$ ist das genau umgekehrt. Auf diese Weise ist die Eigenschaft der Partition der Eins aus Bemerkung 3.13 auf ganz $[a, b]$ gewährleistet. Es gilt $N_{1,3}(x) \geq 0$ für alle $x \in [1, 2]$ und $N_{1,3}(x) = 0$ sonst. Diese Funktion ist an der Stelle 1 ebenfalls unstetig. Es gilt $N_{1,3}(1) = 0$. Der Grenzwert von links $N_{1,3}(1-)$ ist 0 der Grenzwert von rechts $N_{1,3}(1+)$ ist 1.

$N_{2,3}$ ist an der Stelle 1 zwar stetig aber nicht differenzierbar.

$N_{3,3}$ ist auf ganz \mathbb{R} einmal stetig differenzierbar.

Nehmen wir nun einen Spline $s \in \mathbb{P}_{2,(0,1,1,1,2)}$ definiert als

$$s(x) = \sum_{i=-2}^3 \xi_i N_{i,3}(x)$$

mit $\xi_{-2} = \dots = \xi_3 := 1$.

Abbildung 3.2: Beispiel eines $s \in \mathbb{P}_{2,(0,1,1,1,2)}$

In Abbildung 3.2 sieht man sehr gut, daß $N_{-2,3}, \dots, N_{3,3}$ auf $[0, 2]$ eine Partition der Eins bilden - außerhalb dieses Intervalls natürlich nicht. Siehe Bemerkung 3.13.

Außerdem sieht man, daß ein Spline $s \in \mathbb{P}_{2,(0,1,1,1,2)}$ an der Stelle 1 eine Unstetigkeit aufweisen kann, er muß aber nicht.

Definiert man beispielsweise $\xi_{-2} = \xi_{-1} = \xi_0 := 1$ und $\xi_1 = \xi_2 = \xi_3 := \frac{1}{2}$, dann weist der Spline s an der Stelle 1 tatsächlich eine Unstetigkeit auf. Siehe Abbildung 3.3

Abbildung 3.3: Beispiel eines $s \in \mathbb{P}_{2,(0,1,1,1,2)}$

Betrachten wir nun die ersten Ableitungen $N'_{-2,3}, \dots, N'_{3,3}$ in Bild 3.4 und 3.5.

Abbildung 3.4: $N'_{-2,3}$, $N'_{-1,3}$ und $N'_{0,3}$

Abbildung 3.5: $N'_{1,3}$, $N'_{2,3}$ und $N'_{3,3}$

Wie man sieht, ist die erste Ableitung eines B-Spline vom Grad k wieder ein Spline vom Grad $k - 1$ mit demselben Knotenvektor. In unserem Fall handelt es sich also um stückweise affine Funktionen.

$N'_{-2,3}$ hat den Knotenvektor $(-2, -1, 0, 1)$ und ist damit stetig.

$N'_{-1,3}$ hat den Knotenvektor $(-1, 0, 1, 1)$ und ist deshalb unstetig an der Stelle 1. Für den Grenzwert von links gilt $N'_{-1,3}(1-) = -2$, für den Grenzwert von rechts gilt $N'_{-1,3}(1+) = 0$. $N'_{0,3}$ hat den Knotenvektor $(0, 1, 1, 1)$ und ist an der Stelle 1 nicht definiert, da $N_{0,3}$ dort nicht differenzierbar ist.

Für den Grenzwert von links gilt $N'_{0,3}(1-) = 2$, für den Grenzwert von rechts gilt $N'_{0,3}(1+) = 0$.

Die zweiten Ableitungen $N''_{-2,3}, \dots, N''_{3,3}$ sind stückweise konstante Funktionen. Wie die Graphen dieser Funktionen aussehen, wird wohl bei der Betrachtung der ersten Ableitungen hinreichend klar. Deshalb wollen wir hier auf eine eingehendere Betrachtung verzichten.

◇

Bemerkung 3.17

Auf ein Problem, das in der Literatur nicht erwähnt wird, sei an dieser Stelle noch hingewiesen. Sei $k := 2$, $g := 0$, $(\lambda_{-2}, \lambda_{-1}, \lambda_0, \lambda_1, \lambda_2, \lambda_3) := (-2, -1, 0, 1, 1, 1)$ und $I = [0, 1]$ dann gilt $N_{0,3}(1) = 0$ gemäß der Rekursionsformel in Bemerkung 3.8, weil $1 \notin [0, 1[$. Aufgrund der Bemerkung 3.13 über die Bildung einer Partition der Eins und der Randwerteeigenschaft 3.11 muß aber gelten $N_{0,3}(1) = 1$. Das heißt, in diesem Fall muß man den Limes von links - für diesen gilt nämlich $N_{0,3}(1-) = 1$ - berechnen, wenn man auf die Gültigkeit von Bemerkung 3.13 wert legt. Das ist zum Beispiel bei bestimmten Problemen in der CAD der Fall. Diese Problematik berührt uns aber nicht, weil wir unsere Knoten am Rand äquidistant mit einem Abstand größer Null wählen, den Algorithmus aus Bemerkung 3.15 verwenden und im Einzelfall sehr genau unterscheiden, welchen Funktionswert wir gerade ausrechnen und ob wir im Zweifelsfall den Grenzwert von rechts oder links benötigen.

◇

3.1.2 Bivariate Tensorprodukt-Splines

Die Theorie der univariaten Splines (also Splines mit einer Veränderlichen) kann auf verschiedene Arten auf Funktionen mit mehr als einer Variablen ausgedehnt werden. Unter all diesen Generalisierungen ist mit Sicherheit das Konzept der sogenannten **Tensorprodukt-Splines** dasjenige, welches am einfachsten zu handhaben ist. Tensorprodukt-Splines haben den Nachteil, daß sie auf rechteckige Definitionsbereiche beschränkt sind. Das beeinträchtigt uns bei unserer Anwendung aber keineswegs.

Definition 3.18

Sei $\mathcal{F}(I, \mathbb{R})$ die Menge aller Funktionen, die auf einem Intervall $I \subset \mathbb{R}$ definiert sind und nach \mathbb{R} abbilden. Seien $F \subset \mathcal{F}(I, \mathbb{R})$ und $G \subset \mathcal{F}(J, \mathbb{R})$ zwei endlich-dimensionale Vektorräume von Funktionen mit $I, J \subset \mathbb{R}$.

1. Für $f \in F$ und $g \in G$ ist die **Tensorprodukt-Funktion** $f \otimes g : I \times J \rightarrow \mathbb{R}$ definiert durch

$$(f \otimes g)(x, y) = f(x) \cdot g(y).$$

2. Der **Tensorprodukt-Raum** $F \otimes G$ ist definiert durch

$$F \otimes G = \left\{ \sum_{j=1}^k f_j \otimes g_j \mid k \in \mathbb{N}, f_1, \dots, f_k \in F, g_1, \dots, g_k \in G \right\}.$$

[?]

◇

Satz 3.19

Sei $\{f_1, \dots, f_m\}$ eine Basis von F und $\{g_1, \dots, g_n\}$ eine Basis von G , dann ist die Menge

$$M = \{f_i \otimes g_j : I \times J \rightarrow \mathbb{R} \mid i \in \{1, \dots, m\}, j \in \{1, \dots, n\}\}$$

eine Basis von $F \otimes G$.

Ein Beweis findet sich in [dB78].

□

Definition 3.20

Seien $I := [a, b] \subset \mathbb{R}$, $J := [c, d] \subset \mathbb{R}$ zwei kompakte Intervalle. Sei $R := I \times J$ das **Definitions-Rechteck**. Eine Funktion $s : R \rightarrow \mathbb{R}$ mit streng monoton steigender Break-Point-Sequenz $\tau_0, \dots, \tau_m \in \mathbb{R}$ mit $\tau_0 := a$, $\tau_m := b$, $m \in \mathbb{N}_{>}$ und Glattheitsbedingungen $\nu_1, \dots, \nu_{m-1} \in \{0, \dots, k\}$ in x-Richtung, sowie

mit streng monoton steigender Break-Point-Sequenz $\kappa_0, \dots, \kappa_n \in \mathbb{R}$ mit $\kappa_0 := c$, $\kappa_n := d$, $n \in \mathbb{N}_>$ und Glattheitsbedingungen $\omega_1, \dots, \omega_{n-1} \in \{0, \dots, l\}$ in y-Richtung wird **bivariater polynomialer Tensorprodukt-Spline** (im folgenden kurz Spline) vom Grad $k \in \mathbb{N}$ in x- und $l \in \mathbb{N}$ in y-Richtung genannt, wenn sie folgende Bedingungen erfüllt:

1. Auf jedem **Knotenrechteck** $R_{i,j} = [\tau_i, \tau_{i+1}] \times [\kappa_j, \kappa_{j+1}]$ für $i \in \{0, \dots, m-1\}$ und $j \in \{0, \dots, n-1\}$ ist s gegeben durch ein Polynom vom Grad höchstens k in x und vom Grad höchstens l in y

$$s|_{R_{i,j}} \in \mathcal{P}_k \otimes \mathcal{P}_l, \text{ für } i \in \{0, \dots, m-1\} \text{ und } j \in \{0, \dots, n-1\}. \quad (3.15)$$

2. Für $i \in \{1, \dots, m-1\}$ ist die Funktion s an der Stelle (τ_i, y) für alle $y \in J$ $(\nu_i - 1)$ -mal stetig partiell nach x differenzierbar, falls $\nu_i \geq 2$, und stetig in x-Richtung, falls $\nu_i = 1$.

Für $i \in \{1, \dots, n-1\}$ ist die Funktion s an der Stelle (x, κ_i) für alle $x \in I$ $(\omega_i - 1)$ -mal stetig partiell nach y differenzierbar, falls $\omega_i \geq 2$ und stetig in y-Richtung, falls $\omega_i = 1$.

Speziell gilt, falls $\nu_1 = \dots = \nu_{m-1} = k$ und $\omega_1 = \dots = \omega_{n-1} = l$, so ist die Funktion s und alle ihre partiellen Ableitungen $\partial^{i+j}s(x, y)/\partial x^i \partial y^j$ mit $0 \leq i < k$ und $0 \leq j < l$ stetig auf R .

$$\frac{\partial^{i+j}s(x, y)}{\partial x^i \partial y^j} \in \mathcal{C}(R) \text{ mit } 0 \leq i < k \text{ und } 0 \leq j < l. \quad (3.16)$$

◇

Ganz analog zum univariaten Fall kodiert man die Glattheitsbedingungen durch Knotenvervielfachungen in den Knotenvektoren.

Sei $g := \sum_{i=1}^{m-1} ((k+1) - \nu_i)$ und $h := \sum_{i=1}^{n-1} ((l+1) - \omega_i)$.

Dann kann der **Knotenvektor in x-Richtung** $\lambda = (\lambda_0, \dots, \lambda_{g+1}) \in \mathbb{R}^{g+2}$ und der **Knotenvektor in y-Richtung** $\mu = (\mu_0, \dots, \mu_{h+1}) \in \mathbb{R}^{h+2}$ gemäß dem Schema in Definition 3.1 aus den entsprechenden Break-Point-Sequenzen und Glattheitsbedingungen gewonnenen werden.

Im folgenden sei das Definitions-Rechteck R , der Splinegrad in x-Richtung k , die Anzahl der inneren Knoten g des Knotenvektors $\lambda = (\lambda_0, \dots, \lambda_{g+1})$ in x-Richtung, der Splinegrad in y-Richtung l , die Anzahl der inneren Knoten h des Knotenvektors $\mu = (\mu_0, \dots, \mu_{h+1})$ in y-Richtung wie oben definiert.

Definition 3.21

Alle Funktionen, die die Bedingungen aus Satz 3.20 erfüllen, bilden einen Vektorraum der mit $\mathbb{P}_{k,l,\lambda,\mu}$ bezeichnet wird.

◇

Satz 3.22

$\mathbb{P}_{k,l,\lambda,\mu}$ hat die Dimension

$$\begin{aligned} \dim(\mathbb{P}_{k,l,\lambda,\mu}) &= \dim(\mathbb{P}_{k,\lambda}) \cdot \dim(\mathbb{P}_{l,\mu}) \\ &= (g+k+1)(h+l+1). \end{aligned} \quad (3.17)$$

Der Beweis folgt direkt aus den Sätzen 3.3 und 3.19.

□

Wir benötigen nun wieder eine Basis für diesen Funktionenraum.

Analog zum univariaten Fall definieren wir am Rand der Knotenvektoren λ und μ zusätzliche Knoten

$$\begin{aligned} \lambda_{-k} \leq \lambda_{-k+1} \leq \cdots \leq \lambda_{-1} \leq \lambda_0 = a, \\ b = \lambda_{g+1} \leq \lambda_{g+2} \leq \cdots \leq \lambda_{g+k} \leq \lambda_{g+k+1} \end{aligned} \quad (3.18)$$

und

$$\begin{aligned} \mu_{-l} \leq \mu_{-l+1} \leq \cdots \leq \mu_{-1} \leq \mu_0 = c, \\ d = \mu_{h+1} \leq \mu_{h+2} \leq \cdots \leq \mu_{h+l} \leq \mu_{h+l+1}. \end{aligned} \quad (3.19)$$

Satz 3.23

Damit erhält jeder Spline $s \in \mathbb{P}_{k,l,\lambda,\mu}$ eine eindeutig bestimmte Darstellung der Form

$$s(x, y) = \sum_{i=-k}^g \sum_{j=-l}^h c_{i,j} N_{i,k+1}(x) M_{j,l+1}(y), \quad (3.20)$$

wobei $N_{i,k+1}$ für $i \in \{-k, \dots, g\}$ die gemäß Definition 3.6 definierten B-Splines auf den Knotensequenzen $\lambda_i, \dots, \lambda_{i+k+1}$ und $M_{j,l+1}$ für $j \in \{-l, \dots, h\}$ die B-Splines auf den Knotensequenzen $\mu_j, \dots, \mu_{j+l+1}$ sind.

Die $c_{i,j} \in \mathbb{R}$ für $i \in \{-k, \dots, g\}$ und $j \in \{-l, \dots, h\}$ werden als **Tensorprodukt-B-Spline-Koeffizienten** bezeichnet.

Der Beweis folgt direkt aus Satz 3.19.

□

Hier gelten natürlich wieder die gleichen Anmerkungen zum Definitionsbereich des Tensorproduktsplines und der Basisfunktionen wie im univariaten Fall. Siehe Satz 3.11.

Im folgenden seien $\lambda_{-k}, \dots, \lambda_{-1}$ und $\lambda_{g+2}, \dots, \lambda_{g+k+1}$ sowie $\mu_{-l}, \dots, \mu_{-1}$ und $\mu_{h+2}, \dots, \mu_{h+l+1}$ wie oben definiert.

Definition 3.24

Eine Menge $\{(\lambda_i, y) | y \in \mathbb{R}\}$ für $i \in \{-k, \dots, g+k+1\}$ bzw. $\{(x, \mu_j) | x \in \mathbb{R}\}$ für $j \in \{-l, \dots, h+l+1\}$ wird als **Knotenlinie** in x- bzw. y-Richtung bezeichnet. Die Knotenlinie zu einem inneren Knoten wird als innere Knotenlinie bezeichnet.

◇

Bemerkungen 3.25

Die folgenden Eigenschaften der **Tensorprodukt-B-Splines** $N_{i,k+1} \otimes M_{j,l+1}$ für $i \in \{-k, \dots, g\}$ und $j \in \{-l, \dots, h\}$ können unmittelbar aus den entsprechenden Eigenschaften der univariaten B-Splines hergeleitet werden.

1. Positivität und beschränkter Träger

$$\begin{aligned} N_{i,k+1}(x)M_{j,l+1}(y) &\geq 0 \quad \text{für alle } (x, y) \in R \\ &= 0 \quad \text{für alle } (x, y) \notin [\lambda_i, \lambda_{i+k+1}] \times [\mu_j, \mu_{j+l+1}] \end{aligned} \quad (3.21)$$

2. Partition der Eins

$$\sum_{i=-k}^g \sum_{j=-l}^h N_{i,k+1}(x)M_{j,l+1}(y) = 1 \quad \text{für alle } (x, y) \in R. \quad (3.22)$$

◇

Satz 3.26

Ist ein Tensorprodukt-Spline $s : R \rightarrow \mathbb{R}$ mit

$$s(x, y) = \sum_{i=-k}^g \sum_{j=-l}^h c_{i,j} N_{i,k+1}(x)M_{j,l+1}(y),$$

an einer Stelle $(x, y) \in R$ p -mal stetig partiell nach x und q -mal stetig partiell nach y differenzierbar für $p, q \in \mathbb{N}$, so gilt

$$\frac{\partial^{p+q}s(x, y)}{\partial x^p \partial y^q} = \sum_{i=-k}^g \sum_{j=-l}^h c_{i,j} N_{i,k+1}^{(p)}(x)M_{j,l+1}^{(q)}(y). \quad (3.23)$$

Der Beweis ist offensichtlich.

□

Beispiel 3.27

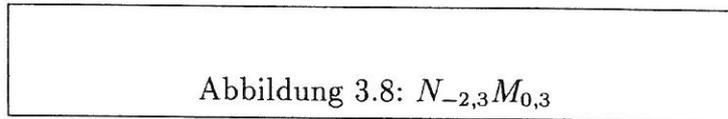
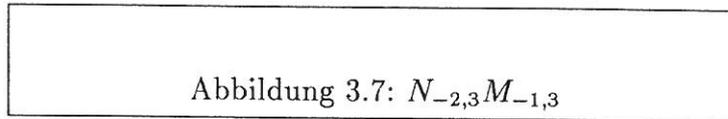
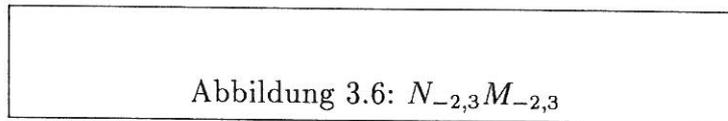
Nun wollen wir noch ein Beispiel eines biquadratischen Tensorprodukt-Spline-Raums betrachten, wie wir ihn auch später benötigen werden.

Sei $k = l := 2, g := 1, h := 3, (\lambda_{-2}, \lambda_{-1}, \lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4) := (-2, -1, 0, 1, 2, 3, 4)$ und $(\mu_{-2}, \mu_{-1}, \mu_0, \mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) := (-2, -1, 0, 1, 1, 1, 2, 3, 4)$, dann bildet $\{N_{i,3} \otimes M_{j,3} | i \in \{-2, \dots, 1\}, j \in \{-2, \dots, 3\}\}$ eine Basis von $\mathbb{P}_{2,2,(0,1,2),(0,1,1,1,2)}$.
 Sehen wir uns drei dieser Basisfunktionen einmal an.

$N_{-2,3} \otimes M_{-2,3}$ (Abbildung 3.6) ist überall stetig differenzierbar.

$N_{-2,3} \otimes M_{-1,3}$ (Abbildung 3.7) hat bereits eine Unstetigkeit in der ersten partiellen Ableitung nach y für $y = 1$.

$N_{-2,3} \otimes M_{0,3}$ (Abbildung 3.8) ist unstetig in y -Richtung bei $y = 1$.



Sehen wir uns noch ein Beispiel für einen Spline $s \in \mathbb{P}_{2,2,(0,1,2),(0,1,1,1,2)}$ an definiert als

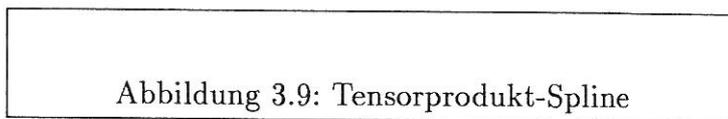
$$s(x, y) := \sum_{i=-2}^1 \sum_{j=-2}^3 c_{i,j} N_{i,3}(x) M_{j,3}(y) \quad \text{für } (x, y) \in \mathbb{R}$$

mit

$$c_{i,j} := \begin{cases} \frac{1}{2} & , \text{ für } i \geq 0 \text{ und } j \geq 1 \\ 1 & , \text{ sonst} \end{cases}$$

für $(i, j) \in \{-2, \dots, 1\} \times \{-2, \dots, 3\}$.

In Abbildung 3.9 sieht man die Unstetigkeit von s in der ersten partiellen Ableitung nach y bei $y = 1$.



3.2 Beste Approximation im Quadratmittel

Die folgenden Ausführungen sind sinngemäß aus [Don95] entnommen.

Gegeben seien Messungen $y_1, \dots, y_m \in \mathbb{R}$ mit $m \in \mathbb{N}_{>}$ aus einer Versuchsreihe zu den Zeitpunkten $t_1, \dots, t_m \in I$ innerhalb des Meßintervalls $I := [a, b]$ mit $a, b \in \mathbb{R}$ und $a < b$. Diese Meßresultate mögen durch inhärente Ungenauigkeiten nur bedingt zuverlässig sein. Andererseits sei - etwa aus physikalischen Erwägungen, oder weil wir es so bestimmen - klar, daß die zeitabhängige Beschreibung des Meßvorgangs nur einer Funktion $f : I \rightarrow \mathbb{R}$ gehorchen könne mit $f = \sum_{i=1}^n \xi_i g_i$ für das Meßintervall I , bekannte Basisfunktionen $g_i : I \rightarrow \mathbb{R}$ eines endlich dimensionalen Funktionenraums F mit $\dim(F) = n$ und unbekannte Koeffizienten $\xi_i \in \mathbb{R}$ für $i \in \{1, \dots, n\}$.

Gesucht sind dann ξ_1, \dots, ξ_n , sodaß

$$\sum_{i=1}^n \xi_i g_i(t_j) = f(t_j) = y_j \quad \text{für } j = 1, \dots, m$$

gilt. Setzt man für $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ und $y \in \mathbb{R}^m$

$$A := \begin{pmatrix} g_1(t_1) & \cdots & g_n(t_1) \\ \vdots & \ddots & \vdots \\ g_1(t_m) & \cdots & g_n(t_m) \end{pmatrix}, x := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \quad \text{und } y := \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

lautet diese Gleichung:

$$Ax = y, \tag{3.24}$$

ist also ein lineares Gleichungssystem.

In den meisten Anwendungsfällen wird man mit vielen Messungen rechnen können, während die Anzahl n der Basisfunktionen eher klein sein wird. Ist aber $m > n$, so handelt es sich bei 3.24 um ein **überbestimmtes** lineares Gleichungssystem, das nur in Ausnahmefällen noch lösbar ist. Man sucht deshalb ein $x_0 \in \mathbb{R}^n$, für das die Differenz

$$Ax_0 - y$$

dem Nullvektor möglichst nahekommt. Mißt man die „Nähe“ zur Null durch die euklidische Norm, gelangt man zum **euklidischen linearen Ausgleichsproblem**:

Gesucht ist ein Vektor $x_0 \in \mathbb{R}^n$ mit

$$\|Ax_0 - y\| = \inf\{\|Ax - y\| \mid x \in \mathbb{R}^n\}.$$

Mit anderen Worten man sucht eine **bestapproximierende Abbildung** $f \in F$, sodaß die **Quadratmittelabweichung** $\sum_{j=1}^m (f(t_j) - y_j)^2$ minimal wird. Sei $f = \sum_{i=1}^n \xi_i g_i$, dann gilt

$$\sum_{j=1}^m (f(t_j) - y_j)^2 = \sum_{j=1}^m \left(\sum_{i=1}^n \xi_i g_i(t_j) - y_j \right)^2 = \|Ax - y\|^2.$$

Die Bestimmung einer derartigen bestapproximierenden Abbildung über die Lösung eines linearen Ausgleichsproblems, wird als **Passung** bezeichnet.

Es gibt stets ein $x_0 \in \mathbb{R}^n$ mit

$$\|Ax_0 - y\| = \inf\{\|Ax - y\| \mid x \in \mathbb{R}^n\}.$$

Das Infimum kann also als Minimum geschrieben werden. Beweis siehe [Don95].

3.2.1 Lösung linearer Ausgleichsprobleme

Die folgenden Ausführungen sind sinngemäß aus [Don95] entnommen.

Ein Vektor $x_0 \in \mathbb{R}^n$ ist genau dann Lösung eines linearen Ausgleichsproblems

$$\|Ax_0 - y\| = \min\{\|Ax - y\| \mid x \in \mathbb{R}^n\} \quad (3.25)$$

oder kurz

$$\min_{x \in \mathbb{R}^n} \|Ax - y\|, \quad (3.26)$$

wenn er die **Normalengleichung**

$$\wedge \quad A^T Ax_0 = A^T y$$

erfüllt. Beweis siehe [Don95].

Das Kernproblem eines Lösungsverfahrens über die Normalengleichung ist seine gelegentlich hohe numerische Instabilität. Ist die Matrix $A^T A$ sehr schlecht konditioniert (was in der Praxis nicht selten vorkommt), so erhält man nur ein sehr ungenaues Resultat.

Die am häufigsten verwendete Methode zur Lösung eines linearen Ausgleichsproblems stützt sich auf orthogonale Transformationen.

Da orthogonale Transformationen normerhaltend sind, reicht es eine geeignete unitäre Matrix Q zu suchen und das äquivalente Problem der Minimierung von $\|(Q^T Ax) - (Q^T b)\|$ zu lösen. Die Matrix Q sollte so gewählt werden,

daß $Q^T A$ eine möglichst einfache Struktur bekommt und somit das neue Minimierungsproblem einfacher zu lösen ist [Don95].

Die folgenden Ausführungen sind sinngemäß aus [?] entnommen. Eine Null-Matrix wird im folgenden jeweils durch eine $\mathbf{0}$ symbolisiert. Die Dimensionen werden jeweils aus dem Kontext klar.

Sei $A \in \mathbb{R}^{m \times n}$ mit $m \geq n$ und $y \in \mathbb{R}^m$ sowie eine bereits berechnete orthogonale Matrix $Q \in \mathbb{R}^{m \times m}$ gegeben, sodaß

$$Q^T A = R = \begin{pmatrix} R_1 \\ \mathbf{0} \end{pmatrix}$$

eine obere Dreiecksmatrix mit $R \in \mathbb{R}^{m \times n}$ und $R_1 \in \mathbb{R}^{n \times n}$ ist. Sei

$$Q^T y = \begin{pmatrix} c \\ d \end{pmatrix}$$

mit $c \in \mathbb{R}^n$ und $d \in \mathbb{R}^{m-n}$, dann gilt

$$\|Ax - y\|^2 = \|Q^T Ax - Q^T b\|^2 = \|R_1 x - c\|^2 + \|d\|^2 \quad \text{für alle } x \in \mathbb{R}^n.$$

Wenn also $\text{Rang}(A) = \text{Rang}(R_1) = n$ dann ist unsere gesuchte Lösung x_0 des linearen Ausgleichsproblems 3.25 gegeben als die Lösung des oberen Dreieckssystems

$$R_1 x_0 = c$$

Im folgenden seien Q und R jeweils wie oben definiert. Die Berechnung einer solchen QR-Zerlegung kann zum Beispiel über Householder-Transformationen erfolgen. Siehe [?], [ÄB96] oder [Don95].

Der obige Satz bietet eine Lösung für lineare Ausgleichsprobleme, falls die Systemmatrix A vollen Rang hat. In der Praxis muß das aber keineswegs immer so sein.

Falls die Systemmatrix $A \in \mathbb{R}^{m \times n}$ einen Rangdefekt aufweist, also $\text{Rang}(A) = r < n$ gilt, dann hat das zugehörige lineare Ausgleichsproblem unendlich viele Lösungen. Es existiert aber eine Lösung x_0 mit minimaler Norm. Beweis siehe [?].

QR-Zerlegung mit Spaltenpivotisierung

Sei $A \in \mathbb{R}^{m \times n}$ mit $\text{Rang}(A) = r < n$, dann existiert eine QR-Zerlegung von A und eine Permutationsmatrix $\Pi \in \mathbb{R}^{n \times n}$, sodaß $A\Pi = QR$, wobei

$$R = \begin{pmatrix} R_{11} & R_{12} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

mit $R_{11} \in \mathbb{R}^{r \times r}$, $R_{12} \in \mathbb{R}^{r \times (n-r)}$. Damit kann das lineare Ausgleichsproblem gelöst werden.

Für alle $x \in \mathbb{R}^n$ gilt

$$\begin{aligned} \|Ax - b\|^2 &= \|(Q^T A \Pi)(\Pi^T x) - (Q^T b)\|^2 \\ &= \|R_{11}y - (c - R_{12}z)\|^2 + \|d\|^2, \end{aligned}$$

mit

$$\Pi^T x = \begin{pmatrix} y \\ z \end{pmatrix} \quad \text{und} \quad Q^T b = \begin{pmatrix} c \\ d \end{pmatrix}$$

für $y \in \mathbb{R}^r$, $z \in \mathbb{R}^{n-r}$, $c \in \mathbb{R}^r$ und $d \in \mathbb{R}^{m-r}$. Für eine Lösung x des linearen Ausgleichsproblems, muß also gelten

$$x = \Pi \begin{pmatrix} R_{11}^{-1}(c - R_{12}z) \\ z \end{pmatrix}.$$

Falls man in diesem Ausdruck z auf Null setzt, erhält man die sogenannte **Basislösung** $x_B \in \mathbb{R}^n$

$$x_B = \Pi \begin{pmatrix} R_{11}^{-1}c \\ \mathbf{0} \end{pmatrix}.$$

Die Basislösung ist nicht notwendigerweise auch die Lösung mit minimaler euklidischer Norm.

Für die Lösung mit minimaler euklidischer Norm $x_{LS} \in \mathbb{R}^n$ gilt

$$\|x_{LS}\| = \min_{z \in \mathbb{R}^{n-r}} \left\| x_B - \Pi \begin{pmatrix} R_{11}^{-1}R_{12} \\ -I_{n-r} \end{pmatrix} z \right\|,$$

wobei $I_k \in \mathbb{R}^{k \times k}$ für $k \in \mathbb{N}_>$ die Einheitsmatrix bezeichne. Also ist die Basislösung nur dann auch die Lösung mit minimaler euklidischer Norm, wenn R_{12} die Null-Matrix ist.

Man erreicht eine solche **vollständige orthogonale Zerlegung**, indem man die Matrix R von rechts mit einem geeigneten Produkt aus orthogonalen Matrizen multipliziert. Genauer gesagt man berechnet eine QR-Zerlegung von R^T . Seien $Z_1, \dots, Z_r \in \mathbb{R}^{n \times n}$ orthogonale Matrizen, sodaß

$$Z_r \cdots Z_1 R^T = Z_r \cdots Z_1 \begin{pmatrix} R_{11}^T & \mathbf{0} \\ R_{12}^T & \mathbf{0} \end{pmatrix} = \begin{pmatrix} T_{11}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

mit einer oberen Dreiecksmatrix $T_{11} \in \mathbb{R}^{r \times r}$ ist, dann gilt für $Z \in \mathbb{R}^{n \times n}$ definiert als $Z = \Pi Z_1 \cdots Z_r$ und $T \in \mathbb{R}^{m \times n}$

$$Q^T A Z = T = \begin{pmatrix} T_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Die gesuchte Lösung mit minimaler euklidischer Norm ist dann

$$x_{LS} = Z \begin{pmatrix} T_{11}^{-1} c \\ \mathbf{0} \end{pmatrix}.$$

Ein geeigneter Algorithmus, der eine QR-Zerlegung einer Matrix und eine geeignete Permutationsmatrix Π berechnet, findet sich in [?], [ÅB96] oder in [Don95].

3.2.2 Lösung linearer Ausgleichsprobleme mit linearen Nebenbedingungen

Die folgenden Ausführungen sind sinngemäß aus [?] entnommen.

Wir haben in Kapitel 3.1.2 gesehen, wie über den Knotenentwurf Unstetigkeiten an bestimmten Stellen eines Tensorprodukt-Splines zugelassen werden können. Will man in bestimmten Bereichen Stetigkeit garantieren und in anderen Unstetigkeit und läßt dies der Knotenentwurf so nicht zu, kann dieses Ziel nachträglich bei der Passung durch das Aufstellen von linearen Nebenbedingungen erreicht werden. Deshalb führen wir nun die Theorie zur Lösung linearer Ausgleichsprobleme mit linearen Nebenbedingungen ein. Wie diese Nebenbedingungen in der Praxis aussehen müssen, wird in Kapitel ?? näher erläutert.

Gegeben sei $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times n}$ und $d \in \mathbb{R}^p$. Gesucht ist ein Vektor $x \in \mathbb{R}^n$, für den $Bx = d$ gilt und der gleichzeitig eine Lösung des linearen Ausgleichsproblems

$$\|Ax - b\| = \min\{\|Ay - b\| \mid y \in \mathbb{R}^n\}$$

ist. Sei $\mathcal{M} = \{x \mid x \in \mathbb{R}^n, Bx = d\}$. Gesucht ist also ein Vektor $x \in \mathcal{M}$ mit

$$\|Ax - b\| = \min\{\|Ay - b\| \mid y \in \mathcal{M}\} \quad (3.27)$$

Das Problem 3.27 besitzt eine Lösung, falls die linearen Nebenbedingungen **konsistent** sind, mit anderen Worten falls $\text{Rang}(B) = p$ ist, also das lineare

Gleichungssystem $Bx = d$ eine Lösung besitzt. Die Lösung ist genau dann eindeutig bestimmt, wenn

$$\text{Rang}(B) = p \quad \text{und} \quad \text{Rang}\begin{pmatrix} A \\ B \end{pmatrix} = n$$

ist. Falls $\text{Rang}(A) = p$ und $\text{Rang}\begin{pmatrix} A \\ B \end{pmatrix} < n$ ist die Lösung nicht eindeutig bestimmt. Es existiert aber eine Lösung mit minimaler euklidischer Norm.

Eine einfache Möglichkeit zur Lösung von 3.27 ist es ein äquivalentes niedriger dimensionales Ausgleichsproblem ohne Nebenbedingungen zu berechnen. Siehe [?]. Dazu eignet sich zum Beispiel die Methode der „direkten Eliminierung“, die im folgenden kurz vorgestellt werden soll.

Vorausgesetzt sei der Einfachheit halber, daß die Matrizen A und B vollen Rang haben. Man berechnet zuerst eine QR-Zerlegung der Matrix B^T .

Sei

$$Q^T B^T = \begin{pmatrix} R \\ \mathbf{0} \end{pmatrix}$$

mit $Q \in \mathbb{R}^{n \times n}$, $R \in \mathbb{R}^{p \times p}$ und setzt man

$$AQ = (A_1 \ A_2) \quad \text{und} \quad Q^T x = \begin{pmatrix} y \\ z \end{pmatrix}$$

mit $A_1 \in \mathbb{R}^{m \times p}$ und $A_2 \in \mathbb{R}^{m \times (n-p)}$, dann genügt es einen Vektor $y \in \mathbb{R}^p$ mit $R^T y = d$ und einen Vektor $z \in \mathbb{R}^{n-p}$ zu bestimmen, der das lineare Ausgleichsproblem ohne Nebenbedingungen

$$\|A_2 z - (b - A_1 y)\| = \min\{\|A_2 z^* - (b - A_1 y)\| \mid z^* \in \mathbb{R}^{n-p}\} \quad (3.28)$$

löst.

$$x = Q \begin{pmatrix} y \\ z \end{pmatrix}$$

ist dann der gesuchte Lösungsvektor für 3.27. Beweis siehe [?]. Ist $\text{Rang}(B) = p$ und $\text{Rang}(A) < n$ so erhält man mit Hilfe des in Abschnitt 3.2.1 vorgestellten Verfahrens der QR-Zerlegung mit Spaltenpivotisierung eine Lösung mit minimaler euklidischer Norm.

Kapitel 4

Spotdetektion

In Folgenden wird die Entstehung der Spots in der Detektorebene eines SHS mathematisch modelliert. Die optischen Grundlagen können in [Hec99] nachgeschlagen werden. Eine Einführung in die Fourier-Optik ist in [Goo96] zu finden. Die zum Verständnis dieses Kapitels nötigen Grundbegriffe werden im Folgenden angeführt. Falls nicht explizit anders vermerkt gehen wir im Folgenden von monochromatischem, kohärentem Licht aus, da die im Projekt verwendete Lichtquelle ein Laser ist.

4.1 Optische Grundlagen

Ein Mikrolinsenarray besteht aus einzelnen Linsen. Über einer Linse wird eine einfallende Wellenfront zunächst durch die Linse transformiert. Die Mikrolinsen haben typischerweise eine Breite von $150\mu m$ und eine Dicke von $1\mu m$ (?), bei einer Fokallänge von $f = 4mm$. Das bedeutet, daß wir mit recht von dünnen Linsen ausgehen können. Im Anschluß an die Linse, wird die Wellenfront noch um die Fokallänge weit durch den freien Raum propagiert. Wir werden uns also zunächst um den Effekt dieser beiden Transformationen auf eine einfallende Wellenfront kümmern und anschließend die Umsetzung auf ein ganzes Mikrolinsenarray besprechen.

4.1.1 Grundlagen Fourier-Optik

Definition 4.1

Seien f und g zwei Funktionen $f, g : \mathbb{R}^2 \rightarrow \mathbb{C}$. Dann heißt

$$(f * g)(x, y) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t, u)g(x - t, y - u)du dt$$

Faltung von f und g .

◇

Definition 4.2

Seien Lebesgue-integrierbare Funktionen $g, G : \mathbb{R}^2 \rightarrow \mathbb{C}$ gegeben, dann ist die zweidimensionale Fourier-Transformierte als $\mathcal{F}\{g\}$ und die inverse Transformierte als $\mathcal{F}^{-1}\{G\}$ notiert.

$$\mathcal{F}\{g\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \exp[-2\pi i(f_x x + f_y y)] dx dy$$

$$\mathcal{F}^{-1}\{G\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(f_x, f_y) \exp[2\pi i(f_x x + f_y y)] df_x df_y$$

◇

Satz 4.3

Für zwei Lebesgue-integrierbare Funktionen $f, g : \mathbb{R}^2 \rightarrow \mathbb{C}$ gilt

$$\mathcal{F}\{f * g\} = \mathcal{F}\{f\} \cdot \mathcal{F}\{g\}.$$

Der Beweis kann in [Gra92] nachgeschlagen werden.

◇

4.1.2 Das Winkelspektrum

Angenommen eine monochromatische Welle, die sich mit einer Komponente in positiver z -Richtung fortpflanzt, trifft auf eine (x, y) -Ebene. Sei das komplexe Feld entlang der $z = 0$ -Ebene durch $U(x, y, 0)$ repräsentiert. Dann ist das Feld $U(x, y, z)$ in einer dazu parallelen Ebene in Ausbreitungsrichtung der Welle gesucht. In der $z = 0$ -Ebene hat die Funktion U eine zweidimensionale Fourier-Transformierte

$$A(f_x, f_y, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x, y, 0) \exp[-2\pi i(f_x x + f_y y)] dx dy$$

Die Fourier-Transformation kann als Zerlegung einer komplizierten Funktion in eine Ansammlung einfacherer komplexer Exponential-Funktionen gesehen werden. Schreiben wir also die Funktion U als eine inverse Fouriertransformierte ihres Spektrums.

$$U(x, y, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(f_x, f_y, 0) \exp[2\pi i(f_x x + f_y y)] df_x df_y$$

Der Integrand dieses Integrals kann als Planwelle gesehen werden, die sich mit Wellenvektor \vec{k} ausbreitet, wobei \vec{k} den Betrag $2\pi/\lambda$ und den Richtungs-Cosinus (α, β, γ) hat. Eine Planwelle hat in komplexer Schreibweise demnach folgende Form:

$$p(x, y, z, t) = \exp \left[i \left(\langle \vec{k} | \vec{r} \rangle - 2\pi vt \right) \right],$$

wobei $\vec{r} \in \mathbb{R}^3$ ein Positionsvektor und $\vec{k} = \frac{2\pi}{\lambda}(\alpha, \beta, \gamma)^T \in \mathbb{R}^3$ der Wellenvektor ist. Läßt man die Zeitabhängigkeit weg, so erhält man entlang einer Ebene mit konstanter z-Koordinate den komplexen Phasor-Term

$$P(x, y, z) = \exp \left[i \langle \vec{k} | \vec{r} \rangle \right] = \exp \left[i \frac{2\pi}{\lambda} (\alpha x + \beta y) \right] \exp \left[i \frac{2\pi}{\lambda} \gamma z \right]$$

Die Richtungs-Cosinusse hängen über die Gleichung

$$\gamma = \sqrt{1 - \alpha^2 - \beta^2}$$

zusammen. Also kann entlang einer $z = 0$ -Ebene die komplexe Exponential-Funktion $\exp [2\pi i (f_x x + f_y y)]$ als Planwelle, die sich in Richtung der Richtungs-Cosinusse

$$\alpha = \lambda f_x, \quad \beta = \lambda f_y, \quad \gamma = \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2}$$

ausbreitet. Die Funktion

$$A \left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, 0 \right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x, y, 0) \exp \left[-2\pi i \left(\frac{\alpha}{\lambda} x + \frac{\beta}{\lambda} y \right) \right] dx dy$$

nennt man Winkel-Spektrum der Störung $U(x, y, 0)$.

4.1.3 Propagation des Winkelspektrums

Das Winkel-Spektrum in einer konstanten z-Ebene kann als

$$A \left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z \right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x, y, z) \exp \left[-2\pi i \left(\frac{\alpha}{\lambda} x + \frac{\beta}{\lambda} y \right) \right] dx dy$$

geschrieben werden. Die Störung U in dieser Ebene kann über die inverse Fourier-Transformierte ihres Winkel-Spektrums geschrieben werden:

$$U(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A \left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z \right) \left[2\pi i \left(\frac{\alpha}{\lambda} x + \frac{\beta}{\lambda} y \right) \right] d\frac{\alpha}{\lambda} d\frac{\beta}{\lambda}$$

Zusätzlich muß U die Helmholtz Gleichung erfüllen:

$$\nabla^2 U + k^2 U = 0.$$

Also muß die Differential-Gleichung

$$\frac{d^2}{dz^2} A\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z\right) + \left(\frac{2\pi}{\lambda}\right)^2 [1 - \alpha^2 - \beta^2] A\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z\right) = 0$$

gelten. Eine elementare Lösung dieser Gleichung kann in der Form

$$A\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z\right) = A\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, 0\right) \exp\left[i\frac{2\pi}{\lambda}\sqrt{1 - \alpha^2 - \beta^2}z\right] \quad (4.1)$$

geschrieben werden. Dieses Ergebnis zeigt, daß wenn die Richtungs-Cosinusse die Bedingung

$$\alpha^2 + \beta^2 < 1 \quad (4.2)$$

erfüllen, was alle Richtungs-Cosinusse tun, der Effekt einer Propagation einer Störung über eine Entfernung z einfach eine Änderung der einzelnen Phasenwerte des Winkel-Spektrums ist. Jede Planwellen-Komponente der Welle pflanzt sich in einem anderen Winkel fort. Demnach legt jede Komponente zwischen zwei parallelen Ebenen eine andere Entfernung zurück.

Falls

$$\alpha^2 + \beta^2 > 1$$

gilt, sind α und β nicht mehr als Richtungs-Cosinusse interpretierbar. Die Wurzel in Gleichung 4.1 ist nun imaginär. Also kann die Gleichung geschrieben werden als

$$A\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, z\right) = A\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, 0\right) \exp(-\mu z)$$

wobei

$$\mu = \frac{2\pi}{\lambda}\sqrt{\alpha^2 + \beta^2 - 1}$$

ist. Da μ eine positive reelle Zahl ist, werden diese Wellenkomponenten schnell durch die Wellenpropagation abgeschwächt. Solche Komponenten werden **schwindende Wellen** genannt.

Zum Schluß bemerken wir, daß die Störung, die in (x, y, z) zu beobachten ist, durch ihr anfängliches Winkel-Spektrum ausgedrückt werden kann, indem wir

Gleichung 4.1 auf den Gültigkeitsbereich von Gleichung 4.2 beschränken und invers Fourier-transformieren:

$$U(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}, 0\right) \exp\left[i\frac{2\pi}{\lambda}\sqrt{1-\alpha^2-\beta^2}z\right] \\ \times \text{circ}(\sqrt{\alpha^2+\beta^2}) \exp\left[2\pi i\left(\frac{\alpha}{\lambda}x + \frac{\beta}{\lambda}y\right)\right] d\frac{\alpha}{\lambda}d\frac{\beta}{\lambda} \quad (4.3)$$

Dabei ist die circ-Funktion definiert als $\text{circ} : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$\text{circ}(\sqrt{\alpha^2+\beta^2}) = \begin{cases} 1, & \sqrt{\alpha^2+\beta^2} < 1 \\ \frac{1}{2}, & \sqrt{\alpha^2+\beta^2} = 1 \\ 0, & \text{sonst} \end{cases}$$

4.1.4 Effekt einer beugenden Struktur auf das Winkelspektrum

Angenommen es befindet sich ein undurchsichtiger, unendlich ausgedehnter Schirm mit einer beugenden Struktur in der Ebene $z = 0$. Wir untersuchen nun den Effekt dieser beugenden Struktur auf das Winkelspektrum einer Störung. Wir definieren die **Amplituden-Durchlässigkeits-Funktion** der Apertur als das Verhältnis der Amplitude des durchgelassenen Feldes $U_t(x, y, 0)$ und der Amplitude des eintreffenden Feldes $U_i(x, y, 0)$ an jedem Punkt (x, y) in der $z = 0$ -Ebene

$$t_A(x, y) = \frac{U_t(x, y, 0)}{U_i(x, y, 0)}$$

Daraus folgt, daß

$$U_t(x, y, 0) = U_i(x, y, 0)t_A(x, y)$$

ist. Nun kann der Faltungs-Satz angewendet werden um das Winkel-Spektrum $A_t(\alpha/\lambda, \beta/\lambda)$ des eintreffenden Feldes mit dem Winkelspektrum $A_i(\alpha/\lambda, \beta/\lambda)$ des durchgelassenen Feldes in Beziehung zu setzen,

$$A_t\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) = \left[A_i\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) \otimes T\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right)\right],$$

wobei

$$T\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t_A(x, y) \exp\left[2\pi i\left(\frac{\alpha}{\lambda}x + \frac{\beta}{\lambda}y\right)\right] dx dy$$

ist. Das Winkelspektrum des durchgelassenen Feldes ist demnach eine Faltung des Winkelspektrums des einfallenden Feldes mit einem Winkelspektrum, das die beugende Apertur charakterisiert. Im Falle einer einfallenden Planwelle mit Amplitude 1 ist das Winkelspektrum des durchgelassenen Feldes gerade die Fouriertransformierte der Amplituden-Durchlässigkeits-Funktion der beugenden Apertur.

4.1.5 Das Propagations-Phänomen als linearer räumlicher Filter

Eine Welle pflanzt sich von der $z = 0$ Ebene bis zu einer dazu parallelen Ebene mit Entfernung > 0 fort. Dabei wird das Feld der einfallenden Störung $U(x, y, 0)$ in ein anderes Feld $U(x, y, z)$ transformiert. Dies erfüllt die Definition eines Systems. Daß es sich um ein lineares System handelt folgt direkt aus der Linearität der Wellengleichung. Die Rauminvarianz kann am leichtesten durch die tatsächliche Herleitung einer Übertragungsfunktion gezeigt werden. Sei das Raumspektrum einer Störung $U(x, y, z)$ wieder repräsentiert durch $A(f_x, f_y, z)$ und das Raumspektrum von $U(x, y, 0)$ durch $A(f_x, f_y, 0)$, dann kann $U(x, y, z)$ ausgedrückt werden als:

$$U(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(f_x, f_y, z) \exp[2\pi i(f_x x + f_y y)] df_x df_y$$

Zusätzlich kann man es wie aus Gleichung 4.3 folgt auch schreiben als:

$$U(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(f_x, f_y, 0) \exp \left[i \frac{2\pi}{\lambda} \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2} z \right] \\ \times \text{circ}(\sqrt{(\lambda f_x)^2 + (\lambda f_y)^2}) \exp [2\pi i (f_x x + f_y y)] df_x df_y$$

Ein Vergleich der beiden Gleichungen zeigt, daß

$$A(f_x, f_y, z) = A(f_x, f_y, 0) \text{circ} \left(\sqrt{(\lambda f_x)^2 + (\lambda f_y)^2} \right) \\ \times \exp \left[2\pi i \frac{z}{\lambda} \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2} \right]$$

Die Übertragungsfunktion des Wellen-Propagations-Phänomens ist also

$$H(f_x, f_y) = \begin{cases} \exp \left[2\pi i \frac{z}{\lambda} \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2} \right], & \text{falls } \sqrt{f_x^2 + f_y^2} < \frac{1}{\lambda}, \\ 0, & \text{sonst.} \end{cases} \quad (4.4)$$

Die Intensität einer skalaren, monochromatischen Welle in einem Punkt P ist definiert als das Quadrat der komplexen Phasor Darstellung der Störung $U(P)$:

$$I(P) = \|U(P)\|^2$$

Sie ist direkt proportional zur Energiedichte und ist die Größe, die wir tatsächlich physikalisch messen können.

4.1.6 Fresnel Beugung

Angenommen man hat eine beugende Öffnung in einem unendlich ausgedehnten Schirm Σ in der $z = 0$ -Ebene, die in positiver z -Richtung beleuchtet wird und im Abstand z dahinter eine Beobachtungsebene. Das Huygens-Fresnel-Prinzip besagt, daß die Störung in einem Punkt der Beobachtungsebene gegeben ist als

$$U(P_0) = \frac{1}{\lambda i} \int \int_{\Sigma} U(P_1) \frac{\exp(kr_{01}i)}{r_{01}} \cos \Theta ds \quad (4.5)$$

wobei P_1 ein bestimmter Punkt in der Beobachtungsebene, Θ der Winkel zwischen Normalenvektor von Σ und Vektor \vec{r}_{01} von P_0 nach P_1 ist. Es handelt sich also um das Integral aller von Punkten im Schirm Σ ausgesandten Kugelwellen, die im Punkt P_1 ankommen. Der Winkel $\cos \Theta$ ist exakt gegeben durch

$$\cos \Theta = \frac{z}{r_{01}},$$

deshalb kann das Huygen-Fresnel Prinzip umgeschrieben werden als:

$$U(x, y) = \frac{z}{\lambda i} \int \int_{\Sigma} U(\xi, \eta) \frac{\exp(kr_{01}i)}{r_{01}^2} d\xi d\eta,$$

wobei der Abstand r_{01} exakt gegeben ist als

$$r_{01} = \sqrt{z^2 + (x - \xi)^2 + (y - \eta)^2}. \quad (4.6)$$

In diesem Ausdruck stecken nur zwei Approximationen: die Annahme, daß die skalare Theorie anwendbar ist und daß sich die Beobachtungsebene viele Wellenlängen hinter der beugenden Apertur befindet, daß also $r_{01} \gg \lambda$ ist. Man kann nun noch weitere Approximationen einführen um des Fresnel-Huygens-Prinzips in eine einfachere und leichter benutzbarere Form zu überführen.

Sei $b \in \mathbb{R}$ mit $|b| \leq 1$, dann kann die Wurzel aus Gleichung 4.6 über die Binomialerweiterung angenähert werden.

$$\sqrt{1+b} = 1 + \frac{1}{2}b - \frac{1}{8}b^2 + \dots,$$

wobei die für eine bestimmte Genauigkeit nötige Anzahl Terme von der Größe von b abhängt. r_{01} kann geschrieben werden als

$$r_{01} = z \sqrt{1 + \left(\frac{x-\xi}{z}\right)^2 + \left(\frac{y-\eta}{z}\right)^2}$$

Setzen wir b mit dem zweiten und dritten Term unter der Wurzel gleich und behalten wir nur die ersten beiden Terme der Binomialerweiterung der Wurzel, dann erhalten wir

$$r_{01} \approx z \left[1 + \frac{1}{2} \left(\frac{x-\xi}{z}\right)^2 + \frac{1}{2} \left(\frac{y-\eta}{z}\right)^2 \right].$$

Der Fehler der durch diese Approximation eingeführt wird, hängt vom Erscheinungsort von r_{01} ab. Für das r_{01}^2 im Nenner ist der Fehler durch das Ersetzen von r_{01} durch den ersten Term seiner Binomialerweiterung in der Regel akzeptabel klein. Das r_{01} im Exponenten der exp-Funktion ist dabei viel kritischer, da es zum ersten mit einer sehr großen Zahl k multipliziert wird, die im Falle von sichtbarem Licht größer als 10^7 sein kann. Zweitens können Phasen Verschiebungen von einem Bruchteil eines Rads den Wert der Exponential-Funktion signifikant verändern. Hier behalten wir also die ersten beiden Terme der Erweiterung. Der resultierenden Term für das Feld ergibt sich also zu

$$U(x, y) = \frac{e^{ikz}}{i\lambda z} \int \int_{-\infty}^{\infty} U(\xi, \eta) \exp \left\{ i \frac{k}{2z} [(x-\xi)^2 + (y-\eta)^2] \right\} d\xi d\eta. \quad (4.7)$$

Diese Gleichung kann leicht als eine Faltung der Form

$$U(x, y) = \int \int_{-\infty}^{\infty} U(\xi, \eta) h(x-\xi, y-\eta) d\xi d\eta$$

gesehen werden, wobei der Faltungskern

$$h(x, y) = \frac{e^{ikz}}{i\lambda z} \exp \left[\frac{ik}{2z} (x^2 + y^2) \right] \quad (4.8)$$

ist. Eine andere Form der Gleichung 4.7 kann durch Ausklammern des Faktors $\exp\left[\frac{ik}{2z}(x^2 + y^2)\right]$ gefunden werden:

$$U(x, y) = \frac{e^{ikz}}{i\lambda z} e^{\frac{ik}{2z}(x^2 + y^2)} \int \int_{-\infty}^{\infty} \left\{ U(\xi, \eta) e^{i\frac{k}{2z}(\xi^2 + \eta^2)} \right\} e^{-i\frac{2\pi}{\lambda z}(x\xi + y\eta)} d\xi d\eta. \quad (4.9)$$

Dies erkennen wir abgesehen von multiplikativen Faktoren als Fourier-Transformation des Produkts des komplexen Feldes unmittelbar rechts der Apertur mit einem quadratischen Phasen-Exponentialfunktion. Wir bezeichnen beide Formeln 4.7 und 4.9 als Fresnel-Beugungs-Integrale.

Die Übertragungsfunktion für die Fresnel-Beugung ist die Fourier-Transformierte des Faltungskerns 4.8.

$$H(f_x, f_y) = \mathcal{F}(h(x, y)) = \mathcal{F} \left\{ h(x, y) = \frac{e^{ikz}}{i\lambda z} \exp \left[\frac{ik}{2z}(x^2 + y^2) \right] \right\} = e^{ikz} \exp \left[-\pi \lambda z i (f_x^2 + f_y^2) \right] \quad (4.10)$$

Dieser Ausdruck besteht aus einer konstanten Phasenverzögerung die vom Abstand z der beiden Ebenen voneinander abhängt und einem quadratischen Phasenverzögerungsfaktor, der vom Winkel abhängt in dem die Planwellenkomponenten sich von der Aperturebene wegbewegen. Approximiert man in der Übertragungsfunktion für die Winkelspektrums-Propagation 4.4 den Exponenten durch die ersten beiden Glieder seiner Binomial-Erweiterung

$$\sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2} \approx 1 - \frac{(\lambda f_x)^2}{2} - \frac{(\lambda f_y)^2}{2}$$

erhält man die Übertragungsfunktion für Fresnel-Beugung. Obige Erweiterung ist dann zulässig, wenn $|\lambda f_x| \ll 1$ und $|\lambda f_y| \ll 1$ ist. Damit schränkt man f_x und f_y einfach auf kleine Winkel ein.

4.1.7 Fraunhofer Beugung

Gilt für die Entfernung z zwischen Apertur-Ebene und Beobachtungsebene, daß

$$z \gg \frac{k(\xi^2 + \eta^2)_{\max}}{2} \quad (4.11)$$

dann wird aus dem quadratischen Phasenterm $\exp[ik/(2z)(\xi^2 + \eta^2)]$ ungefähr 1 innerhalb des Beobachtungsausschnitts. Damit kann das Feld der Störung

in der Beobachtungsebene abgesehen von einem multiplikativen, konstanten Phasenterm einfach durch die Fourier-Transformation des Feldes in der Aperturebene erhalten werden. Es gilt dann

$$U(x, y) = \frac{e^{ikz}}{i\lambda z} e^{\frac{ik}{2z}(x^2+y^2)} \int \int_{-\infty}^{\infty} U(\xi, \eta) e^{-i\frac{2\pi}{\lambda z}(x\xi+y\eta)} d\xi d\eta. \quad (4.12)$$

Abgesehen von multiplikativen Faktoren, die dem Integral vorangestellt sind, ist dieser Ausdruck gerade die Fouriertransformierte der Verteilung in der Aperturebene, ausgewertet an den Frequenzen

$$f_x = \frac{x}{\lambda z} \quad f_y = \frac{y}{\lambda z}.$$

4.1.8 Fraunhofer Beugungsmuster einer kreisrunden Apertur

Sei der Radius der beugenden Apertur w gegeben und sei q eine Radiuskoordinate in der Apertur, dann ist die Amplituden-Durchlässigkeitsfunktion

$$t_A(q) = \text{circ} \left(\frac{q}{w} \right)$$

Die Kreissymmetrie des Problems impliziert, daß die Fourier-Transformation in der Fraunhofer-Beugungs-Formel 4.12 als eine Fourier-Bessel-Transformation geschrieben werden kann. Sei also r eine Radius-Koordinate in der Beobachtungsebene, dann haben wir

$$U(r) = \frac{e^{ikz}}{i\lambda z} \exp \left(i \frac{kr^2}{2z} \right) \mathcal{B} \{ U(q) \} \Big|_{\rho=r/(\lambda z)}$$

wobei $q = \sqrt{\xi^2 + \eta^2}$ den Radius in der Apertur-Ebene repräsentiert und $\rho = \sqrt{f_x^2 + f_y^2}$ den Radius in der Raumfrequenz-Domäne. Für eine ebene Welle mit Amplitude 1, deren Fortpflanzungsrichtung senkrecht auf der Aperturebene steht, ist das Feld das durch die Apertur hindurchgelassen wird gleich der Amplituden-Durchlässigkeitsfunktion. Es gilt

$$\mathcal{B} \left\{ \text{circ} \left(\frac{q}{w} \right) \right\} = A \frac{J_1(2\pi w \rho)}{\pi w \rho}$$

mit $A = \pi w^2$. Die Amplituden-Verteilung des Fraunhofer-Beugungsmusters ist demnach mit $k = \frac{2\pi}{\lambda}$

$$U(r) = e^{ikz} e^{i\frac{kr^2}{2z}} \frac{A}{\lambda z i} \left[2 \frac{J_1(kwr/z)}{kwr/z} \right].$$

Die Intensitätsverteilung kann als

$$I(r) = \left(\frac{A}{\lambda z}\right)^2 \left[2 \frac{J_1(kwr/z)}{kwr/z}\right]^2.$$

geschrieben werden.

4.1.9 Eine Dünne Linse als Phasentransformation

Eine Linse besteht aus optisch dichtem Material, meist Glas mit einem refraktivem Index von ungefähr 1.5, in dem sich Licht mit einer anderen Geschwindigkeit ausbreitet als in der Luft. Eine Linse wird als dünne Linse bezeichnet, wenn Licht das an den Koordinaten (x, y) an einer Seite in die Linse eintritt ungefähr an den gleichen Koordinaten an der anderen Seite der Linse wieder austritt, daß die Translation innerhalb der Linse vernachlässigbar klein ist. Bei den im Projekt verwendeten Mikrolinsen ist dies sicher der Fall. Eine Linse verzögert also eine einfallende Wellenfront an einem bestimmten Punkt um einen von der Dicke des Glases in diesem Punkt abhängigen Faktor.

Betrachten wir eine Linse zwischen zwei Ebenen mit Abstand Δ_0 . Sei Δ_0 auch die maximale Dicke dieser Linse in axialer Richtung und sei $\Delta(x, y)$ die Dicke in einem bestimmten Punkt (x, y) . Dann läßt sich die Phasenverzögerung von Ebene zu Ebene mit Durchschreitung einer dazwischenliegenden Linse in einem bestimmten Punkt (x, y) schreiben als

$$\Phi(x, y) = kn\Delta(x, y) + k[\Delta_0 - \Delta(x, y)].$$

Dabei sei n der refraktive Index des Materials der Linse. $kn\Delta(x, y)$ ist die Phasenverzögerung durch die Linse. $k[\Delta_0 - \Delta(x, y)]$ ist die Phasenverzögerung durch die freie Luft zwischen den beiden Ebenen. Äquivalent dazu kann die Phasentransformation durch einen multiplikativen Phasentransformationsterm geschrieben werden

$$t_l(x, y) = \exp[ik\Delta_0] \exp[ik(n-1)\Delta(x, y)]$$

Das komplexe Feld unmittelbar hinter der Linse steht also in folgender Beziehung zum komplexen Feld unmittelbar vor der Linse:

$$U_l'(x, y) = t_l(x, y)U_l(x, y)$$

Nun müssen wir nur noch eine Dickenfunktion für $\Delta(x, y)$ finden. Zunächst eine Konvention: Wenn ein Lichtstrahl von links nach rechts geht und auf

eine konvexe Oberfläche trifft, legen wir einen positiven Krümmungsradius fest, während wir einen negativen Krümmungsradius annehmen, wenn er auf eine konkave Oberfläche trifft. Die Dicken-Funktion einer Linse setzen wir wie in Abbildung ?? abgebildet aus drei Termen zusammen

$$\Delta(x, y) = \Delta_1(x, y) + \Delta_2(x, y) + \Delta_3(x, y).$$

Dabei ist die Dickenfunktion $\Delta_1(x, y)$ gegeben durch

$$\begin{aligned} \Delta_1(x, y) &= \Delta_{01} - \left(R_1 - \sqrt{R_1^2 - x^2 - y^2} \right) \\ &= \Delta_{01} - R_1 \left(1 - \sqrt{1 - \frac{x^2 + y^2}{R_1^2}} \right). \end{aligned}$$

Die zweite Komponente besteht aus Region mit konstanter Dicke. Also ist

$$\Delta_2(x, y) = \Delta_{02}.$$

Die dritte Komponente wird wiederum gegeben durch den Ausdruck

$$\begin{aligned} \Delta_3(x, y) &= \Delta_{03} - \left(-R_2 - \sqrt{R_2^2 - x^2 - y^2} \right) \\ &= \Delta_{03} + R_2 \left(1 - \sqrt{1 - \frac{x^2 + y^2}{R_2^2}} \right). \end{aligned}$$

Damit ergibt sich die Gesamtdicke zu

$$\Delta(x, y) = \Delta_0 - \Delta_{01} - R_1 \left(1 - \sqrt{1 - \frac{x^2 + y^2}{R_1^2}} \right) + \Delta_{03} + R_2 \left(1 - \sqrt{1 - \frac{x^2 + y^2}{R_2^2}} \right),$$

wobei $\Delta_0 = \Delta_{01} + \Delta_{02} + \Delta_{03}$ ist. Approximiert man jetzt wieder die beiden Wurzeln in obigem Ausdruck durch die ersten beiden Glieder ihrer Binomialentwicklung

$$\sqrt{1 - \frac{x^2 + y^2}{R_1^2}} \approx 1 - \frac{x^2 + y^2}{2R_1^2},$$

$$\sqrt{1 - \frac{x^2 + y^2}{R_2^2}} \approx 1 - \frac{x^2 + y^2}{2R_2^2},$$

approximiert man also wiederum sphärische Flächen durch parabolische, so ergibt sich die Linsendickenfunktion zu

$$\Delta(x, y) = \Delta_0 - \frac{x^2 + y^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right).$$

Setzt man diesen Ausdruck in die Gleichung für die Phasentransformationsfunktion der Linse ein, so erhält man

$$t_l(x, y) = \exp[ikn\Delta_0] \exp \left[-ik(n-1) \frac{x^2 + y^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \right]$$

Die physikalischen Eigenschaften einer Linse (n, R_1, R_2) können in eine einzelne Formel kombiniert werden, nämlich die Fokallänge der Linse, die als

$$\frac{1}{f} = (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \quad (4.13)$$

definiert ist. Vernachlässigt man den konstanten Phasenterm, also die Dicke der Linse, kann die Phasentransformation der Linse nun als

$$t_l(x, y) = \exp \left[-i \frac{k}{2f} (x^2 + y^2) \right]$$

geschrieben werden. Wir gehen also von einer dünnen Linse aus.

Sei eine planare Eingabe mit Durchlässigkeitsfunktion $t_A(x, y)$ unmittelbar vor eine Sammellinse mit Fokallänge f gesetzt. Die Eingabe wird also von einer monochromatischen Planwelle mit Amplitude A senkrecht beleuchtet. In diesem Fall ist das Eingabefeld über der Linse gleich

$$U_l(x, y) = At_A(x, y)$$

Die endliche Ausdehnung der Linse kann durch eine Pupillen-Funktion $P(x, y)$ definiert als

$$P(x, y) = \begin{cases} 1, & \text{innerhalb der Linsen-Apertur,} \\ 0, & \text{sonst.} \end{cases}$$

modelliert werden. Demnach wäre die komplexe Amplitudenverteilung unmittelbar hinter der Linse gegeben durch

$$U'_l(x, y) = U_l(x, y)P(x, y) \exp \left[-i \frac{k}{2f} (x^2 + y^2) \right] \quad (4.14)$$

Um die Verteilung $U_f(x, y)$ in der Fokalebene der Linse zu finden wenden wir die Fresnel-Beugungs-Formel 4.9 an, wobei wir $z = f$ setzen.

$$U_f(u, v) = \frac{\exp[i\frac{k}{2f}(u^2 + v^2)]}{i\lambda f} \times \int \int_{-\infty}^{\infty} U_l'(x, y) e^{i\frac{k}{2f}(x^2 + y^2)} e^{-i\frac{2\pi}{\lambda f}(xu + yv)} dx dy, \quad (4.15)$$

wobei ein konstanter Phasenterm vernachlässigt wurde. Ersetzt man $U_l'(x, y)$ durch den Term aus Gleichung 4.14, so heben sich die quadratischen Phasenterme im Integranden gegeneinander auf, wodurch übrig bleibt

$$U_f(u, v) = \frac{\exp[i\frac{k}{2f}(u^2 + v^2)]}{i\lambda f} \times \int \int_{-\infty}^{\infty} U_l(x, y) P(x, y) e^{-i\frac{2\pi}{\lambda f}(xu + yv)} dx dy, \quad (4.16)$$

Somit ist das Feld nach der Linse proportional zur zweidimensionalen Fourier-Transformierten des Eintrittsfeldes innerhalb der Linsen-Apertur. Damit sehen wir, daß die komplexe Amplituden Verteilung des Feldes in der Fokalebene der Linse genau das Fraunhofer-Beugungsmuster des Eingabefeldes über der Linse ist, obwohl der Abstand von der Apertur- zur Beobachtungsebene nur der Fokallänge f entspricht und nicht unbedingt das übliche Kriterium für die Validität der Fraunhofer-Approximation 4.11 erfüllt.

4.1.10 Elektromagnetische Grundlagen

Ein Lichtstrahl kann man als Weg im \mathbb{R}^3 sehen, entlang dem sich ein elektromagnetisches Feld mit Stärke $\psi : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}$ in sinus-förmiger Art und Weise mit Geschwindigkeit $v \in \mathbb{R}$ ausbreitet. In einem homogenen Medium entspricht dieser Weg einer geraden Linie. $\psi(x, y, z, t)$ gehorcht der dreidimensionalen Wellendifferentialgleichung

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2}$$

Eine Welle kann man als eine beliebige Lösung dieser Wellendifferentialgleichung definieren. Mit der Phasenfunktion $\varphi : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}$ und der maximalen Amplitude der Welle $A \in \mathbb{R}$ läßt sich ψ auch schreiben als:

$$\psi(r, t) = A \sin(\varphi(r, t)).$$

In komplexer Schreibweise, seien ψ und φ jeweils Funktionen die in die komplexen Zahlen abbilden. Damit läßt sich ψ schreiben als:

$$\psi(r, t) = A e^{i\varphi(r, t)}.$$

Eine ebene Welle im \mathbb{R}^3 mit Anfangsphase $\epsilon \in \mathbb{R}$, Wellenlänge $\lambda \in \mathbb{R}$ und Richtungsvektor $k \in \mathbb{R}^3$ hat z.B. die Phasenfunktion

$$\varphi(r, t) = \langle k | r \rangle \mp \omega t + \epsilon$$

wobei $\omega := v \|k\|$ und $\|k\| = \frac{2\pi}{\lambda}$ ist.

Die Fläche, die zu einem bestimmten Zeitpunkt t_0 alle Punkte gleicher Phase verbindet, bezeichnet man als **Wellenfront**.

Sobald Licht auf ein Hindernis trifft und sich dadurch die Amplitude eines Teils der Wellenfront ändert, tritt **Beugung** auf. Die geometrische Optik, oder Theorie der achsennahen Strahlen, vernachlässigt jegliche Beugungseffekte. Stellt man sich eine Anordnung mit einer Linse mit Fokallänge f in der Linsenebene Σ und einer Fokalebene σ parallel zu dieser Ebene mit Abstand f vor, so müssten sich gemäß der Theorie der geometrischen Optik die Strahlen einer eintreffenden ebenen Welle (mit Wellenfronten parallel zu Σ) in der Fokalebene in einem Punkt bündeln. In der Realität beobachtet man allerdings, daß die „Fokuspunkte“ eine gewisse räumliche Ausdehnung haben. Zur Wellenfrontrekonstruktion ist die Ermittlung von Wellenfrontsteigungen in bestimmten Punkten notwendig. Das heißt man muß aus der Form dieser kleinen hellen Scheiben eine Koordinate extrahieren, der dann eine bestimmte Wellenfrontsteigung zugeordnet wird. Für eine gute mathematische Modellierung ist also ein genaues Verständnis der physikalischen Vorgänge notwendig, die zur Entstehung dieser Fokus-Scheibchen führt. Wir begeben uns also auf das Gebiet der physikalischen Optik und der Fourier-Optik. In unserem Fall sind die zu beobachtenden Spots in der Detektorebene des SHS aufgrund der Mikrolinsen mit umgebender Blende als Fernfeld-Beugungsmustern der Subaperturen zu sehen. Es tritt also der Fall von Fraunhoferbeugung auf. Gehen wir zunächst auf die physikalischen Vorgänge vor, nach und hinter einer einzelnen (Mikro-)Linse ein. Die elektrische Feldverteilung kann innerhalb der Blendenöffnung einer Linse mittels der Blendenöffnungsfunktion $\mathcal{A} : \mathbb{R}^2 \rightarrow \mathbb{C}$ mit

$$\mathcal{A}(x, y) := \mathcal{A}_0(x, y) e^{i\phi(x, y)}.$$

modelliert werden. Dabei beinhaltet $\mathcal{A}_0 : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Amplitudenverteilung und $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Phasenunterschiede von Punkt zu Punkt innerhalb

der Blendenöffnung. In der Annahme gleichmäßiger Ausleuchtung der Linse können wir \mathcal{A}_0 als konstant annehmen.

Die elektrische Feldverteilung in der Fokalebene der Linse entspricht der Fourier-Transformierten der Blendenöffnungsfunktion. Sei f_l die Fokallänge der Linsen, dann ist die elektrische Feldverteilung in der Fokalebene gegeben als

$$\psi(Y, Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{A}(y, z) e^{i\|k\|(Yy+Zz)/f} dydz.$$

Mit

$$k_y := \frac{\|k\| Y}{f_l} \quad \text{und} \quad k_z := \frac{\|k\| Z}{f_l}$$

wird daraus

$$\begin{aligned} \psi(k_y, k_z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{A}(y, z) e^{i(k_y y + k_z z)} dydz \\ &= \mathcal{F}[\mathcal{A}(k_y, k_z)]. \end{aligned}$$

k_y und k_z werden auch als Raumfrequenzen bezeichnet. Jedem Punkt in der Bildebene entspricht eine Raumfrequenz.

Die Intensitätsverteilung I ergibt sich bis auf Konstanten, die wir vernachlässigen wollen, über das Quadrat des zeitlichen Mittels der Amplitude des elektrischen Feldes ψ .

$$I := \langle \text{RE}(\psi)^2 \rangle$$

wobei das zeitliche Mittel der Funktion f definiert sei als:

$$\langle f(t) \rangle = \frac{1}{T} \int_t^{t+T} f(x) dx.$$

Die Länge der Mittelwertbildung hängt in der Realität z.B. von der Integrationszeit der verwendeten CCD-Kamera ab.

Als Beispiel sei der Fall einer kreisförmigen Apertur aufgeführt, auf die eine ebene Welle derart fällt, daß sich innerhalb der Blendenöffnung keine Phasenunterschiede ergeben. Damit ist $\varphi(x, y) = 0 \forall (x, y) \in \mathbb{R}^2$. Im Fall einer kreisförmigen Öffnung mit Radius $a \in \mathbb{R}^+$ mit umgebender undurchsichtiger Blende entspricht die zugehörige Blendenöffnungsfunktion der Zylinderfunktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$f(x, y) = \begin{cases} 1, & \sqrt{x^2 + y^2} \leq a, \\ 0, & \text{sonst.} \end{cases}$$

Die elektrische Feldverteilung in der Fokalebene der Linse entspricht der Fourier-Transformierten der Blendenöffnungsfunktion, also $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ (Abb. 4.2) definiert als

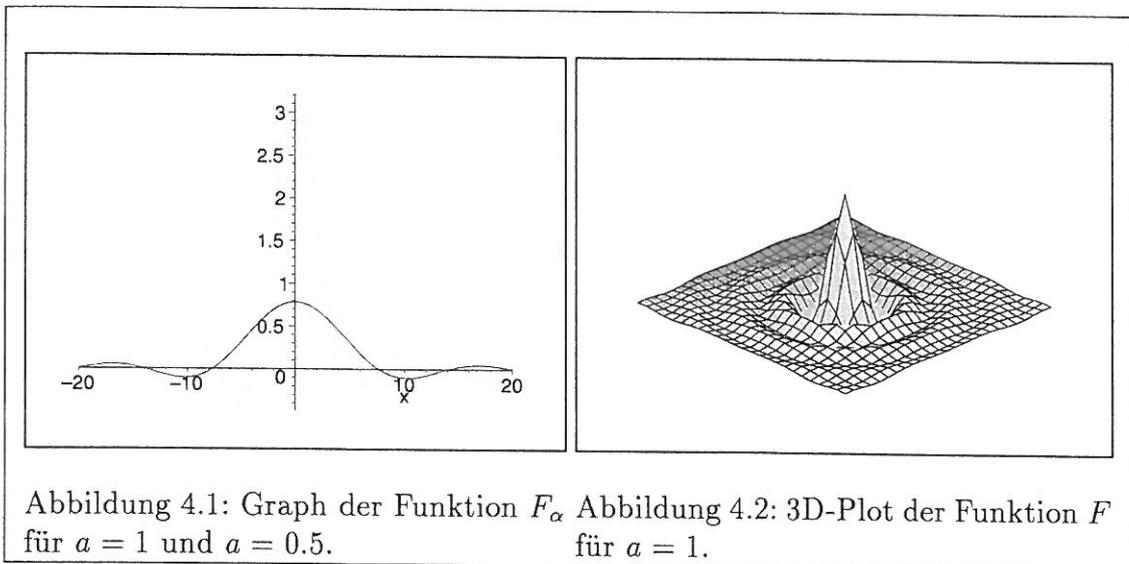
$$F(x, y) := F_\alpha(\sqrt{x^2 + y^2})$$

mit $F_\alpha : \mathbb{R} \rightarrow \mathbb{R}$ (Abb. 4.1) definiert als

$$F_\alpha(k_\alpha) = 2\pi a^2 \left[\frac{J_1(k_\alpha a)}{k_\alpha a} \right],$$

wobei $J_m : \mathbb{R} \rightarrow \mathbb{R}$ für $m \in \mathbb{N}$ die Bessel-Funktion (der ersten Art) m -ter Ordnung bezeichnet. Sie ist definiert als

$$J_m(u) = \frac{i^{-m}}{2\pi} \int_0^{2\pi} e^{iu \cos v} dv.$$



In Abb. 4.1 kann man eine Folge der Eigenschaft der Linse als Fourier-Transformator ersehen: Je kleiner der Durchmesser der Blendenöffnung, desto größer wird die räumliche Ausdehnung des Beugungsmusters.

4.2 Mathematische Modellierung der Spotentstehung

Die folgenden Ausführungen sind aus [RSNK99] entnommen und wurden um die notwendigen Änderungen für kreisrunde Mikrolinsen ergänzt. Sei $f_{zyl} :$

$\mathbb{R}^2 \rightarrow \mathbb{R}$ die Zylinderfunktion definiert als

$$f_{zyl}(x, y) = \begin{cases} 1, & \sqrt{x^2 + y^2} \leq 1, \\ 0, & \text{sonst.} \end{cases}$$

Die komplexe Blendenöffnungsfunktion eines Mikrolinsen-Arrays mit kreisrunden Einzellinsen mit der Fokallänge $f_i \in \mathbb{R}$ und mit Durchmesser $d \in \mathbb{R}$ für eine senkrecht einfallende ebene Wellenfront ist gegeben durch $t_H : \mathbb{R}^2 \rightarrow \mathbb{C}$

$$t_H(x) := \sum_{s \in \{1 \dots S\}} t(x - x_c(s)),$$

wobei die Blendenöffnungsfunktion einer einzelnen Linse $t : \mathbb{R}^2 \rightarrow \mathbb{C}$ definiert ist als

$$t(x) := e^{-i \frac{k_w}{2f_i} \|x\|^2} f_{zyl} \left[\frac{x}{d} \right].$$

Dabei sei $k_w := 2\pi/\lambda$ und $\lambda \in \mathbb{R}$ die mittlere Wellenlänge des verwendeten Lichts. $S \in \mathbb{N}$ ist die Anzahl der Subaperturen des Mikrolinsenarrays und $x_c : \{1 \dots S\} \rightarrow \mathbb{R}^2$ das Zentrum der jeweiligen Subapertur.

Die auf den Shack-Hartmann-Sensor eintreffende Wellenfront gelangt vor dem Auftreffen auf den Sensor durch eine kreisrunde Blende. Die Blendenöffnungsfunktion dieser Blende sei gegeben als $U_{ab} : \mathbb{R}^2 \rightarrow \mathbb{C}$

$$U_{ab}(x) := P(x) e^{i\Phi_{ab}(x)},$$

wobei $\Phi_{ab} : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Phasenunterschiede der eintreffenden Wellenfronten innerhalb der Blende modelliert und $P : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Zylinderfunktion mit Radius der Eintrittsblende ist. Diese Modellierung geht von einer gleichmäßigen Ausleuchtung der Eintrittsblende aus.

Das elektromagnetische Feld unmittelbar hinter den Mikrolinsen ist demnach gegeben als $U_L : \mathbb{R}^2 \rightarrow \mathbb{C}$

$$U_L(x) := U_{ab}(x) t_H(x).$$

Die Intensität des Feldes in der Fokalebene der Mikrolinsen erhalten wir durch Propagation des Feldes U_l eine Fokallänge weit in die Detektorebene. Die Raumfrequenz-Propagations-Funktion $H_A : \mathbb{R}^2 \rightarrow \mathbb{C}$ für diese Propagationslänge ist gegeben durch

$$H_A(f) := e^{2\pi i \frac{f_l}{\lambda} (1 - \lambda \|f\|^2)^{1/2}},$$

wobei f eine Raumfrequenz ist. Also kann das Feld in der Detektorebene des SHS modelliert werden durch $U_D : \mathbb{R}^2 \rightarrow \mathbb{C}$ definiert als

$$U_D(x) := \mathcal{F}^{-1} \{ \mathcal{F}[U_L(x)] H_A(f) \}.$$

f ist dabei wieder eine Raumfrequenz. Die Intensität in der Detektorebene kann man dann durch $I_D : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert als

$$I_D(x) := |U_D(x)|^2$$

berechnen.

Sei die zweidimensionale Diracsche Delta-Funktion $\delta : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben als

$$\delta(x, y) := \begin{cases} \infty, & x = 0 \wedge y = 0, \\ 0, & \text{sonst,} \end{cases}$$

mit der Zusatzbedingung

$$\int_{-\infty}^{\infty} \delta(x) dx = 1.$$

Es gilt

$$t_H(x) = \left(\sum_{s \in \{1, \dots, S\}} \delta(x - x_c(s)) \right) * t(x).$$

Läßt man den Radius der Eintrittsblende gegen unendlich gehen, so ist $U_L = t_H$ und es gilt nach dem Faltungssatz 4.3

$$\mathcal{F}[U_L(x)] = \mathcal{F} \left[\sum_{s \in \{1, \dots, S\}} \delta(x - x_c(s)) \right] \cdot \mathcal{F}[t(x)].$$

Das Fernfeldbeugungsbild des gesamten Mikrolinsenarrays ist also gleich der Fouriertransformation der Blendenöffnungsfunktion einer einzelnen Mikrolinse multipliziert mit der Fourier-Transformierten einer Summe aus Diracschen Delta-Funktionen mit Peaks an den Stellen der Subaperturzentren.

Klassische rechteckige bzw. runde Mikrolinsen haben eine rechteckförmige bzw. zylinderförmige Blendenöffnungsfunktion. Das Beugungsbild einer solchen Subapertur entspricht der Fourier-Transformierten der Blendenöffnungsfunktion.

Im Falle einer rechteckförmigen Subapertur ist dies eine sinc^2 -Funktion, im Falle einer runden Mikrolinse im wesentlichen das Quadrat einer Bessel-Funktion (der ersten Art) erster Ordnung (Siehe [Hec99].). Diese Funktionen sind sehr wenig lokal. Hinter einem Mikrolinsen-Array entspricht die Wellenfront einer Überlagerung (kohärente Addition) dieser Einzelfunktionen. Man kann sich leicht vorstellen, daß die Wellenfront hinter einer bestimmten Mikrolinse durch Wellenfronten benachbarter Mikrolinsen nicht unwesentlich beeinflußt wird. In [RH93] wird zwar dargelegt, daß dieser Effekt auf die Spotzentren in der Fokalebene nur untergeordnete Auswirkungen hat, jedoch wird in [MBN00] auch eine Möglichkeit beschrieben Mikrolinsen mit einer gaußförmigen Blendenöffnungsfunktion herzustellen. Das Beugungsbild einer solchen Linse ist wiederum eine gaußförmige Funktion, die erheblich lokaler als eine sinc^2 -Funktion oder Quadrat einer Bessel-Funktion erster Ordnung ist.

4.3 Schwerpunktmethode

4.3.1 Schwerpunktmethode mit statischem Schwellwert

4.3.2 Schwerpunktmethode mit Zwei-Niveau-Passung

4.4 Funktionenpassung

4.5 Vergleich der vorgestellten Verfahren

Kapitel 5

Spotzuordnung

Der traditionelle Ansatz zur Spotzuordnung wie er beispielsweise in [?] verwendet wird, setzt voraus, daß die Spots in ihren jeweiligen Subaperturen verbleiben müssen. Jeder Subapertur ist damit genau ein Spot zugeordnet. Das Zuordnungsproblem ist in diesem Fall trivial lösbar. Der maximal meßbare Wellenfronthub über dem Sensor unterliegt bei Verwendung dieses Algorithmus allerdings einer sehr harten Einschränkung. Im Projekt Wellensensor sollten mittels sphärischer Kompensation Asphären vermessen werden. Dies war mit diesem Algorithmus nicht möglich, da bei den auftretenden stark asphärischen Wellenfronten, die Spots zum Teil weit aus ihren Subaperturen auswandern.

In diesem Kapitel werden nun verschiedene bekannte Techniken zur Dynamik-Erweiterung beschrieben und ihre Vor- und Nachteile erläutert. Anschließend wird ein neuer Algorithmus beschrieben, der erheblich besser ist als der Stand der Technik. Er ist in der Lage den Dynamikbereich des Sensors hinsichtlich der meßbaren Wellenfrontsteigungen erheblich zu vergrößern. Die vorgestellte Methode ist rein softwarebasiert und erhält als Eingabe nur ein Spotbild.

5.1 Bekannte Verfahren zur erweiterten Spotzuordnung

5.1.1 Hardware-basierte Methoden

- **Zwei-Messungs-Methode:**

In [WBJ93] und [SL96] wird ein Verfahren präsentiert, das die oben angeführte Einschränkung ebenfalls umgehen kann. Allerdings werden

für diesen Algorithmus zwei Messungen benötigt um eine Auswertung zu erhalten. Es ist eine Verfahrenseinheit notwendig, die den Abstand zwischen Mikrolinsenarray und CCD-Kamera variiert. Dieses Verfahren ist aufwendiger und langsamer, als das hier vorgestellte Verfahren, da sowohl der Preis für das Meßgerät, als auch die benötigte Rechenzeit pro Auswertung ansteigen. Für jede Messung ist das Verfahren der Mikrolinsen nötig. Dabei kann es leicht zu Fehlern bei der Einstellung des exakten Abstands zwischen CCD und Kamera kommen. Das wiederum führt zu fehlerhaften Meßdaten. Dieses Verfahren wird in den WaveScope-Sensoren der Firma Applied Optics Associates eingesetzt.

- **Zwei-Sensor-Methode:**

In [Sch98] wird ein Verfahren beschrieben mittels zweier Sensoren unterschiedlicher Empfindlichkeit das Zuordnungsproblem zu lösen. Der Sensor mit der groben Empfindlichkeit dient dazu im Spot-Bild des Sensors mit der feineren Empfindlichkeit Suchbereiche für Spots zu extrapolieren.

- **Signatur-Methode:**

In [Sch99] und [LPS00] wird ein Verfahren beschrieben mittels unterschiedlich gekennzeichnete Spots die Lösung des Zuordnungsproblems zu erleichtern. Dabei werden benachbarte Spots deutlich unterscheidbar gekennzeichnet. Dies kann beispielsweise durch das Versehen der Mikrolinsen mit astigmatischen Aberrationen deutlich unterscheidbarer Hauptachsen geschehen.

- **Verwendung von Nebenwissen über eintreffende Wellenfronten:**

Nebenwissen über eintreffende Wellenfronten kann dazu genutzt werden die Anordnung der Mikrolinsen so zu wählen, daß im Falle des Eintreffens der erwarteten Wellenfront die Foki der Mikrolinsen auf bestimmte Detektoren fallen. Pro Mikrolinse ist dann nur ein Detektor notwendig. Die Abmessungen der Detektoroberfläche bestimmen den Grad an Übereinstimmung zwischen erwarteter und einfallender Welle. Diese Methode wird in [Inc93] beschrieben.

- **Variationen der Linsenanordnung:**

In [Opt00] wird ein Verfahren beschrieben, in einem SHS bestimmte Mikrolinsenarrays einzusetzen, die bestimmte exakt vorherbestimmte Abweichungen der Anordnung der Linsen vom Gitter enthalten. Diese lokalen Abweichungen von der Gitterstruktur müssen in einer Art

und Weise beschaffen sein, daß sie im resultierenden Spotbild wiedererkannt werden können. Die universelle Einsetzbarkeit des Sensors zur Vermessung beliebiger Wellenfrontformen wird dadurch eingeschränkt. Der Dynamikbereich aber über den mit der statischen Suchbereichsmethode erreichbaren hinaus ausgedehnt.

5.1.2 Software-basierte Methoden

- **Verwendung von Nebenwissen über eintreffende Wellenfronten:**

Ist Nebenwissen über die zu erwartenden Wellenfronten vorhanden, so kann dieses bei der Spot-Zuordnung entsprechend berücksichtigt werden. Erwartet man z.B. sphärische Wellenfronten, so können die Spot-Suchbereiche durch eine entsprechende Skalierung des Referenzrasters derart gewählt werden, daß höchstens ein Spot in einen Suchbereich fällt.

- **Auswertung von Spotbild + traditionellem Bild:**

Ein anderer Ansatz ist die zusätzliche Verwendung eines traditionellen Grauwert-Bildes der einfallenden Wellenfront, das mit einer einzigen Linse gebildet wird und mit einer zweiten CCD-Kamera aufgenommen wird. Siehe [RS98] and [RSNK99]. Die Bildaufnahme wird für beide Bilder (das traditionelle sowie das Spotbild des SHS) mathematisch modelliert. Die einfallende Wellenfront wird als Linearkombination von Zernike-Polynomen modelliert. Die Koeffizienten der Zernike-Polynome werden mittels nichtlinearer Optimierung iterativ so verbessert, daß die aufgenommenen Bilder möglichst wenig von den mathematisch berechneten abweichen. Dieses Verfahren ist allerdings recht langsam. Außerdem stellt sich das Problem gute Startwerte für die Optimierung zu finden. Im Falle von Subapertur-Überschreitungen und wenig Nebenwissen über die eintreffenden Wellenfronten wird es schwierig sein geeignete Startwerte für die Optimierung zu finden. Dieses Verfahren könnte als Ergänzung zu dem in dieser Arbeit beschriebenen Algorithmus der Spotzuordnung durch iterative Funktionenpassung gesehen werden. Unklar ist, wofür das traditionell aufgenommene Grauwertbild eigentlich gut ist. Dies wird in den oben erwähnten Papers nicht näher ausgeführt(?).

- **Zuordnung über bewegliche 3x3-Fenster:**

In [Ltd98] wird ein Verfahren ähnlich zu dem in [Gro98] beschriebenen

verwendet. Die Zuordnung erfolgt lokal in 3x3-Spot-großen Fenstern. Die Fenster werden jeweils überlappend weiterbewegt. Die Menge konsistent zugeordneter Spots wird damit iterativ vergrößert.

- **Unwrapping Algorithmus:**

Der nach dem bisherigen Stand der Technik beste Algorithmus der auf einem modifizierten Unwrapping-Verfahren aus der Interferometrie beruht, wird in [PLS98] beschrieben. Er wurde ebenfalls im Projekt Wellensensor entwickelt. Dieses Verfahren funktioniert allerdings nur bis zu einer bestimmten maximalen lokalen Wellenfrontkrümmung.

5.2 Spotzuordnung durch iterative Funktionen-Passung

Das hier vorgestellte Verfahren wurde in [GS99] veröffentlicht und in [GSD⁺00] mit dem in [PLS98] beschriebenen Kontinuierungs-Verfahren verglichen. Ein vom Grundsatz her ähnliches Verfahren wird in [Gre87] für die Dynamikerweiterung von Interferometern vorgestellt.

Im Folgenden wird ein Verfahren beschrieben, das ausgehend von einer konsistent zugeordneten Menge von Spots Positionen benachbarter Spots extrapoliert und diese dann in einem iterativen Prozeß ebenfalls zuordnet. Mittels einer Spline-Passung wird eine Funktion berechnet, die die jeweiligen Mikrolinsen-Indizes möglichst gut auf die Koordinaten der bereits konsistent zugeordneten Spots abbildet. Mittels dieser Funktion werden für benachbarte Mikrolinsen-Indizes Spotpositionen extrapoliert. Anschließend wird in lokalen Umgebungen um diese extrapolierten Spotpositionen nach noch nicht zugeordneten Spots gesucht. Wird in einem Suchbereich ein Spot gefunden, so wird er dem entsprechenden Mikrolinsenindex zugeordnet. Sind alle Suchbereiche abgearbeitet wird eine erneute Spline-Passung auf der nun um einige zuordenbare Spots erweiterten Menge der konsistent parametrisierten Spots durchgeführt. Dieser iterative Prozeß wird beendet, sobald in einer Iteration kein neuer Spot mehr zugeordnet werden kann.

5.2.1 Prinzip

Sei ein Grauwertbild von Spots gegeben, dann wird jede zusammenhängende Menge von Pixeln, deren Grauwerte über einem bestimmten statisch vorgegebenen Schwellwert liegen als Spotpixel betrachtet. Es wird für zusammenhängende Menge von Spotpixeln der Schwerpunkt berechnet. Dieser Schwerpunkt stellt im Falle von kleinen Wellenfrontsteigungen eine gute Näherung für den Strahl dar, der genau durch das optische Zentrum einer Mikrolinse geht. Speziell im Randbereich können allerdings aufgrund von Aberrationen ellipsoide und kometenschweifartige Spots auftreten. Für diesen Bereich kann mittels des in [Can95] beschriebenen Verfahrens ausgehend vom Spotschwerpunkt und durch lokale Simulation des Spotbildes und Vergleich mit dem real vorhandenen eine verbesserte Modellierung des Zentralstrahls gefunden werden. Da es sich bei dem verwendeten Laser um einen kohärente Lichtquelle handelt treten Interferenzeffekte auf, d.h. die Spots beeinflussen sich gegenseitig. Diese Effekte wurden in [RH93] untersucht. Unter bestimmten Umständen sind sie allerdings vernachlässigbar klein.

Im Folgenden wird zunächst von stetigen Wellenfronten und zusammenhängenden Feldern von Subaperturen, die ein vollständiges äquidistantes Gitter bilden ausgegangen. In unserem Fall handelt es sich eigentlich immer um kreisförmige, ellipsoide oder rechteckige Subaperturfelder. Für jede Zusammenhangskomponente von Spotpixeln wird der Schwerpunkt berechnet. Diese Menge von Spotschwerpunkten ergibt unsere Eingabemenge von Spot-Koordinaten. Der Schwerpunkt eines Spots gibt im Falle einer relativ kleinen lokalen Wellenfrontsteigung über einer Mikrolinse eine gute Näherung für den Zentralstrahl. Im Falle starker Wellenfrontsteigungen sorgen sphärische Aberrationen, Koma und Tilt für eine zunehmende Verformung des ursprünglich kreisrunden Spots hin zu einem ellipsoiden oder kometenförmigen Spot. In diesem Fall ist der Schwerpunkt der Zusammenhangskomponente keine gute Näherung mehr. Abhilfe für diesen Fall wird in einem gesonderten Kapitel behandelt. In diesem Kapitel wollen wir zunächst davon ausgehen, daß die Wellenfrontsteigungen so sind, daß der Schwerpunkt eine gute Modellierung des Zentralstrahls darstellt.

Für die initiale Zuordnung einer kleinen Menge von Spots müssen wir geeignetes Nebenwissen verwenden. Bei der Vermessung von rotationssymmetrischen Asphären mit einem sphärischen Kondensator, dessen Scheitelkrümmungsradius dem der Asphäre im zentralen Bereich entspricht, kann dabei das Wissen herangezogen werden, daß die sich ergebenden Wellenfronten im zentralen Bereich eines Spotfeldes relativ eben sein werden. Damit ist in diesem Be-

reich eine initiale Zuordnung mit der weiter unten beschriebenen Startheuristik problemlos möglich. Für die Verarbeitung von anders gearteten Spotfeldern, muß zunächst durch Analyse ihrer Entstehung geeignetes Nebenwissen ermittelt werden. Für unseren Fall hat sich die Vorgehensweise, die initiale Zuordnung im zentralen Bereich der Spotfelder zu beginnen und sich dann iterativ nach außen zu arbeiten als sehr zuverlässig und effizient erwiesen.

Wir gehen nun von einer vorhandenen konsistent parametrisierten Spot- und Mikrolinsenindexmenge aus. Hierbei hat sich eine rechteckige Menge von neun Spots als sehr geeignet erwiesen. Einerseits ist die Aussicht im zentralen Bereich eine geeignete Menge von neun Spots zu finden sehr groß, andererseits reicht sie auch aus um mit Splines von niedrigem Grad (stückweise affin lineare Funktionen) eine genügend stabile Extrapolation benachbarter Spotpositionen zu erhalten.

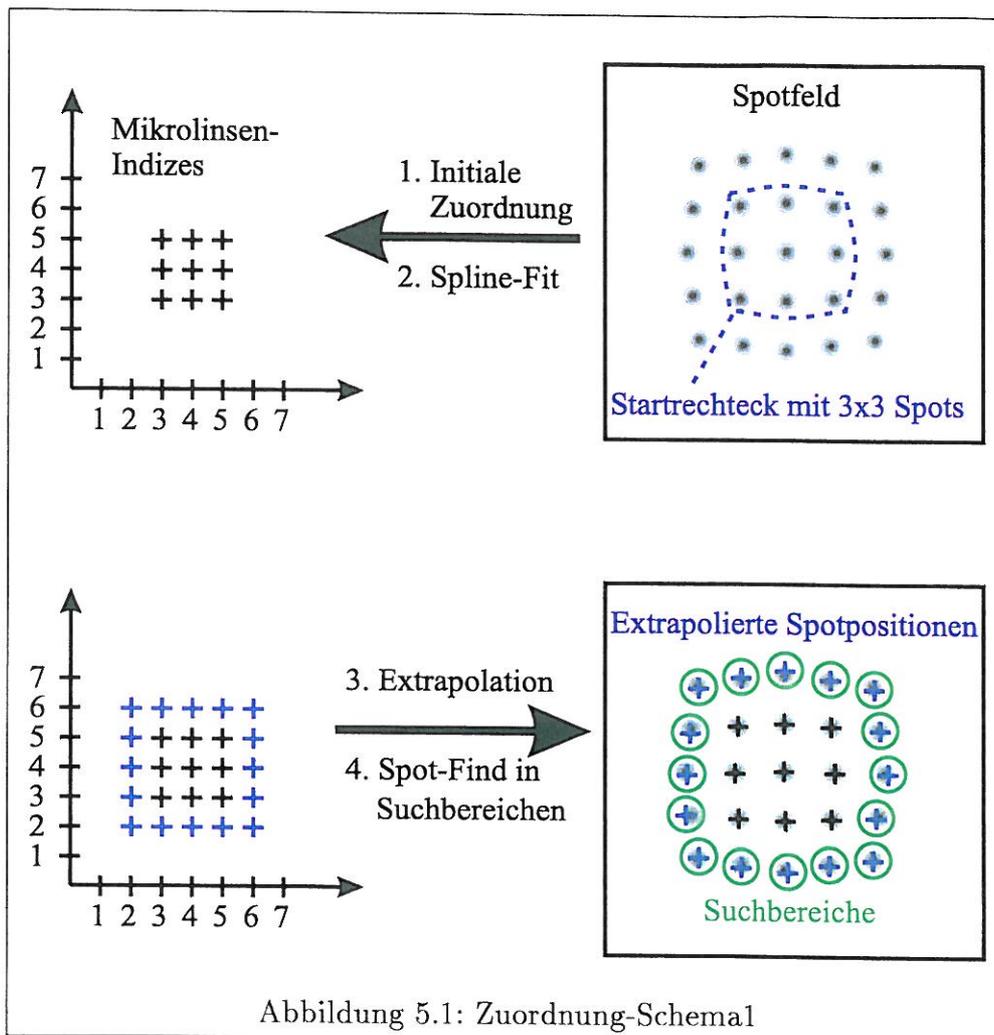
Es seien Spotzentren $y_1, \dots, y_9 \in \mathbb{R}^2$ und zugeordnete Parameterpunkte $x_1, \dots, x_9 \in \mathbb{Z}^2$ gegeben. Ein Parameterpunkt ist eine zweidimensionale Nummer, die sich vom Index der zugeordneten Mikrolinse nur um einen bestimmten für alle Spots gleichen Faktor $z \in \mathbb{Z}^2$ unterscheidet. Der genaue Wert von z ist für den Zuordnungsalgorithmus zunächst unerheblich. Wichtig ist lediglich, daß die Zuordnung konsistent erfolgt. Daß also Spots von benachbarten Mikrolinsen auch den entsprechenden benachbarten Parameterpunkten zugeordnet werden. Die Zuordnung von Spotzentrum zu Parameterpunkt wird auch Parametrisierung genannt. In dem Beispiel in Abb. 5.1 (Schritt 1) wird eine Menge von neun Spots in der Mitte des Spotfeldes als Startmenge ausgewählt. Die Startparametrisierung muß konsistent sein, d.h. durch Addition des unbekanntes Translationsvektors z erhält man Zeile und Spalte der Mikrolinse im Mikrolinsenarray und durch Skalierung mit dem Durchmesser der Subaperturen erhält man die Koordinaten des zugehörigen Referenzpunktes.

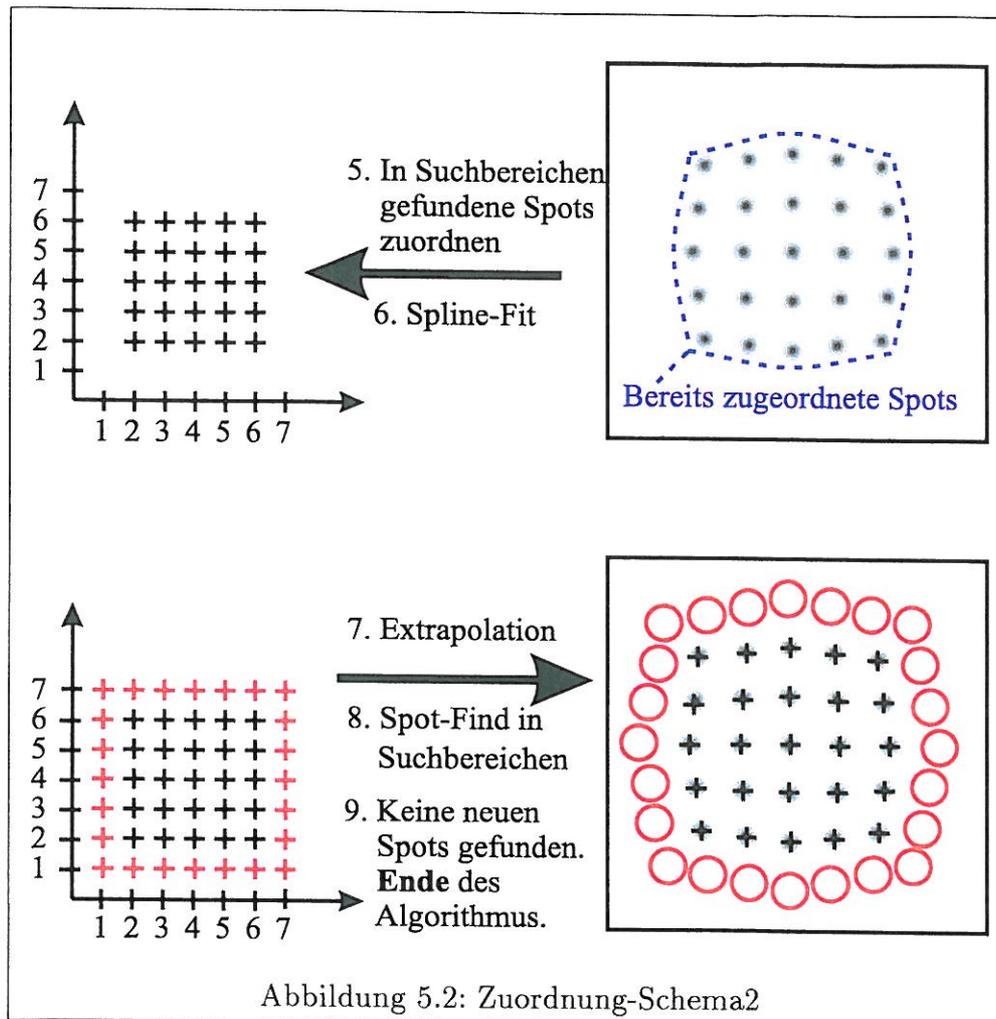
Die horizontale und vertikale Ablenkung $\sigma_i = (\Pi_1(y_i) - \Pi_1(r_i), \Pi_2(y_i) - \Pi_2(r_i))$ eines bestimmten Spots von der lokalen optischen Achse einer Mikrolinse (also dem Referenzpunkt r_i) wird später zur Berechnung der partiellen Ableitungen der Wellenfront an dem Referenzpunkt r_i verwendet. Eine gute Strategie zur Bestimmung einer konsistent parametrisierten Startmenge ist existentiell wichtig für den Erfolg der gesamten Zuordnung.

Ein Startspot $p \in \mathbb{R}^2$ wird ausgewählt und sein nächster Nachbar $q \in \mathbb{R}^2$ wird gesucht. Der nächste Nachbar q habe von p den Abstand $d \in \mathbb{R}$. Nun wird ein Gitter mit Seitenlänge $3d$ und Mittelpunkt p erzeugt. Eine gute Startmenge von Spots wurde gefunden, wenn jedes $d \times d$ großes Quadrat

des Gitters genau ein Spotzentrum enthält (siehe Abb. xx). Nebenwissen über die zu vermessenden Wellenfront kann dazu dienen Bereiche des Spotfeldes auszuwählen in dem die Suche nach einer Startmenge aussichtsreich ist. Steht solches Wissen nicht zur Verfügung müssen nacheinander alle Spots als Mittelpunkte eines solchen Gitters ausgewählt werden, bis schließlich einer obiges Kriterium erfüllt. Die Parametrisierung der Startmenge ist einfach. Das Spotzentrum im oberen linken Quadrat des Gitters in Abb xx mit den Koordinaten y_1 wird dem Parameterpunkt $x_1 = (a, b)$ zugeordnet. Das Spotzentrum im unteren rechten Quadrat mit Koordinaten y_9 wird dem Parameterpunkt $x_9 = (a + 2, b + 2)$ zugeordnet. Die Wahl von $(a, b) \in \mathbb{Z}^2$ ist beliebig. In dieser Stufe des Algorithmus ist es nicht wichtig genau zu wissen, welcher Spot in der Startmenge exakt welcher Mikrolinse zugeordnet ist. Es reicht aus die Spots zu Parameterpunkten zuzuordnen und sie später bei der Tiltkorrektur auf die Mikrolinsen-Numerierung zu verschieben.

In einer Iteration hat man Parameterpunkte x_1, \dots, x_m und Spotzentren y_1, \dots, y_m . Nun wird eine Funktion $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ berechnet, die die Spotzentren y_i an den Parameterpunkten x_i möglichst gut approximiert (siehe Abb. xx Schritt 2) indem wir zwei bivariate Tensorprodukt-Splines an die Eingabedaten fitten (einen für jede Komponente der y_i). Es ist ratsam den Funktionenraum aus dem die Funktion Φ ausgewählt wird nach jeder Iteration der wachsenden Zahl zugeordneter Punktepaare anzupassen. Nach dem Spline-Fit wird die resultierende Funktion Φ an die Eingabemenge umgebenden Parameterpunkten x_i ausgewertet (Abb 5.1 Schritt 3) um extrapolierte Spotpositionen $\tilde{y}_i := \Phi(x_i)$ für $i \in \{m + 1, \dots, m + l\}$ zu erhalten in deren Umgebung dann nach noch nicht zugeordneten Spotzentren gesucht wird. Im Beispiel in Abb. 5.1 hat man $m = 9$ Eingabespots und erhält $l = 16$ extrapolierte Spotpositionen in der ersten Iteration und $m = 25$ Eingabespots und $l = 24$ extrapolierte Spotpositionen in der zweiten Iteration. Nun wird in jeder Iteration für jedes $i \in \{m + 1, \dots, m + l\}$ das tatsächliche Spotzentrum y_i in einer kleinen lokalen Umgebung von \tilde{y}_i gesucht (Abb. 5.1 Schritt 4). Das Spotzentrum das innerhalb dieser lokalen Umgebung liegt und kleinsten Abstand zum extrapolierten Spotzentrum hat wird ausgewählt. Dadurch wird die Menge der Spotzentren mit zugeordneten Parameterpunkten in jeder Iteration um mindestens einen Spot erweitert. Der Algorithmus wird abgebrochen, wenn in einer Iteration keine neuen Spots mehr zugeordnet werden konnten (Abb 5.2 Schritt 8 und 9). In einem letzten Schritt muß noch der Verschiebungsfaktor $z \in \mathbb{Z}^2$ ermittelt werden, der die Menge der Parameterpunkte auf die Numerierung der Mikrolinsen verschiebt. Dies kann in einem Schritt mit der Tiltkorrektur erfolgen.





Der Erfolg der oben beschriebenen Strategie hängt von der Wahl eines geeigneten Funktionenraums ab. Splines vom Grad zwei mit Knotenrechtecken die maximal 6×6 Parameterpunkte einschließen haben sich im FORMIKROSYS II Projekt Wellensensor als eine geeignete Wahl zur Vermessung der von AGFA zur Verfügung gestellten Asphären herausgestellt. Desweiteren hängt der Erfolg auch von der Art ab, wie die Menge der parametrisierten Spotzentren in jeder Iteration erweitert wird. Bei der Vermessung von Wellenfronten mit großen lokalen Krümmungen ist es ratsam in einer Iteration von einer $m \times m$ großen Menge von Parameterpunkten maximal auf eine Menge von $(m + 2) \times (m + 2)$ Punkten zu erweitern. Im Falle von kleinen lokalen Krümmungen ist es möglich in einer Iteration von einer $m \times m$ großen Menge von Parameterpunkten auf eine $(m + r) \times (m + r)$ großen Menge von Punkten auszudehnen.

Für den Fall von nicht rechteckigen Spotsfeldern (z.B. kreisförmige) ist es in einer Iteration nicht immer möglich zu allen extrapolierten Spotzentren auch zuordenbare Spotzentren zu finden. Im Interesse eines schnellen Algorithmus ist allerdings ein Arbeiten auf vollständigen Gittern (Tensorprodukt-Spline-Passung auf Daten, die auf einem vollständigen Gitter liegen) anzuraten. Deshalb besteht die Möglichkeit die berechneten extrapolierten Spotzentren mit zugeordneten Parameterpunkt wieder als Eingabe für die Splinepassungen der nächsten Iteration zu verwenden. Diese müssen nur entsprechend markiert werden, damit die an die Zuordnung anschließende Wellenfrontrekonstruktion nur auf den real vorhandenen Spots arbeitet.

5.2.2 Mathematische Grundlagen

Sei $\mathbb{N}_{>} := \{1, 2, \dots\}$ die Menge der natürlichen Zahlen ohne die 0.

Sei \mathbb{K} ein Körper und $\Pi_i : \mathbb{K}^n \rightarrow \mathbb{K}$ für $i \in \{1, \dots, n\}$ die Projektion auf die i -te Komponente eines Vektors.

Sei $\langle \cdot | \cdot \rangle : \mathbb{R}^n \rightarrow \mathbb{R}$ das Standardskalarprodukt auf dem \mathbb{R}^n definiert als $\langle x | y \rangle := \sum_{i=1}^n x_i y_i$ für alle $x, y \in \mathbb{R}^n$. Sei $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ die euklidische Norm auf dem \mathbb{R}^n definiert als $\|x\| = \sqrt{\langle x | x \rangle}$ für alle $x \in \mathbb{R}^n$.

Ein parametrisierter Punkt $p \in \mathbb{R}^2 \times \mathbb{R}^2$ ist ein Paar bestehend aus dem Parameterpunkt $\Pi_1(p)$ und dem Datenpunkt $\Pi_2(p)$.

Gegeben sei eine Startmenge $P := \{(x_1, y_1), \dots, (x_m, y_m)\} \subset \mathbb{Z}^2 \times \mathbb{R}^2$ mit $m \in \mathbb{N}_{>}$ von konsistent parametrisierten Punkten und eine endliche Menge $Q \subset \mathbb{R}^2$ von noch nicht parametrisierten Punkten. Auf geeignete Strategien zur Ermittlung einer konsistent parametrisierten Startmenge wird unten näher eingegangen.

Es wird vorausgesetzt, daß die Parameterpunkte x_1, \dots, x_m ein regelmäßiges, vollständiges Gitter bilden. Sei $\{x_1, \dots, x_m\} = \{x_S, \dots, x_S + d_x\} \times \{y_S, \dots, y_S + d_y\}$ für ein $(x_S, y_S) \in \mathbb{Z}^2$ und $d_x, d_y \in \mathbb{N}_{>}$.

Sei $t \in \mathbb{N}$. Anfangs setzen wir $t := 0$. Sei $\epsilon_{Max} \in \mathbb{R}_{>}$ der Abstand, den ein berechneter Punkt von einem noch nicht parametrisierten Datenpunkt haben darf, um mit ihm identifiziert zu werden¹. Sei $t_{Step} \in \mathbb{N}_{>}$ die Schrittweite mit der, die „rechteckige“ Menge auf der eine Passung durchgeführt wird in

¹Es hat sich gezeigt, daß $\epsilon_{Max} := 6$ auf dem vorliegenden Bildmaterial gute Ergebnisse liefert.

jeder Iteration vergrößert wird ²

1. Sei $P = \{(x_1, y_1), \dots, (x_m, y_m)\}$ für ein $m \in \mathbb{N}_{>}$. Bestimme zu gegebenen Parameterpunkten x_1, \dots, x_m , Datenpunkten y_1, \dots, y_m und Basisfunktionen $\varphi_1, \dots, \varphi_n$ mit $n \in \mathbb{N}_{>}$ und $m > n$ eines geeigneten n -dimensionalen Funktionenraums $\mathcal{F}(\mathbb{R}^2, \mathbb{R})$ eine im Least-Squares-Sinne bestapproximierende Abbildung $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Zu diesem Zweck werden $A \in \mathbb{R}^{m \times n}$ und $b_1, b_2 \in \mathbb{R}^m$ wie folgt definiert:

$$A := \begin{pmatrix} \varphi_1(x_1) & \dots & \varphi_n(x_1) \\ \vdots & \ddots & \vdots \\ \varphi_1(x_m) & \dots & \varphi_n(x_m) \end{pmatrix},$$

$$b_1 := \begin{pmatrix} \Pi_1(y_1) \\ \vdots \\ \Pi_1(y_m) \end{pmatrix}, b_2 := \begin{pmatrix} \Pi_2(y_1) \\ \vdots \\ \Pi_2(y_m) \end{pmatrix}.$$

Durch Lösung der beiden linearen Ausgleichsprobleme

$$\min_{\xi \in \mathbb{R}^n} \|A\xi - b_1\|, \min_{\nu \in \mathbb{R}^n} \|A\nu - b_2\|, \quad (5.1)$$

kann man die Koeffizienten $\xi, \nu \in \mathbb{R}^n$ der gesuchten Abbildung $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definiert als

$$\Phi := \begin{pmatrix} \sum_{i=1}^n \xi_i \varphi_i \\ \sum_{i=1}^n \nu_i \varphi_i \end{pmatrix},$$

finden, die die Fehlerfunktion

$$(\xi, \nu) \rightarrow \sum_{j=1}^m \left(\sum_{i=1}^n \xi_i \varphi_i(x_j) - \Pi_1(y_j) \right)^2 + \sum_{j=1}^m \left(\sum_{i=1}^n \nu_i \varphi_i(x_j) - \Pi_2(y_j) \right)^2.$$

minimiert. Lösungsmethoden für lineare Ausgleichsprobleme können in [GL89] und in [ÅB96] nachgeschlagen werden.

2. Sei $R, R' \subset \mathbb{Z}^2$ definiert als

$$R := \{x_S - t, \dots, x_S + d_x + t\} \times \{y_S - t, \dots, y_S + d_y + t\}$$

²*t_{Step}* kann bei der vorliegenden Anwendung problemlos auf 2 gesetzt werden.

bzw.

$$R' := \{x_S - (t + 1), \dots, x_S + d_x + t + 1\} \times \{y_S - (t + 1), \dots, y_S + d_y + t + 1\}.$$

Für alle $x \in R' \setminus R$ teste, ob ein noch nicht parametrisierter Punkt $q \in Q$ existiert mit $\|f(x) - q\| < \epsilon_{Max}$. Falls ja, setze $Q := Q \setminus \{q\}$, $P := P \cup \{(x, q)\}$ ansonsten setze $P := P \cup \{(x, \Phi(x))\}$.

3. Falls im letzten Schritt kein noch nicht parametrisierter Datenpunkt parametrisiert werden konnte, beende Algorithmus.
Ansonsten setze $t := t + t_{Step}$. Adaptiere eventuell Funktionenraum $\mathcal{F}(\mathbb{R}^2, \mathbb{R})$ entsprechend der nun gestiegenen Anzahl an parametrisierten Punkten, die als Eingabe für eine erneute Passung dienen.
Gehe zu Schritt 1.

Eine geeignete Startmenge von konsistent parametrisierten Punkten kann zum Beispiel über die im folgenden beschriebene Heuristik gewonnen werden. Sei $Q \subset \mathbb{R}^2$ die Menge aller noch nicht parametrisierten Datenpunkte. Sei $P \subset \mathbb{Z}^2 \times \mathbb{R}^2$ die Menge der parametrisierten Punkte. Anfangs ist $P = \{\}$. Eine „rechteckige“ Menge von 3x3 Punkten wird als Startmenge ausgewählt und wie in Abbildung 5.3 parametrisiert.

1. Bestimme Schwerpunkt des Spotfeldes. Suche Punkt $q_1 \in Q$ mit minimalem Abstand zum Spotfeldschwerpunkt. Bestimme zweiten Punkt $q_2 \in Q \setminus \{q_1\}$ mit minimalem Abstand zu q_1 .
2. Sei $d \in \mathbb{R}$ und $(x, y) \in \mathbb{R}^2$ definiert als $d := \|q_1 - q_2\|$ und

$$(x, y) := (\Pi_1(q_1) - \lfloor d/2 \rfloor, \Pi_2(q_1) - \lfloor d/2 \rfloor).$$

Für alle $(i, j) \in \{-1, 0, 1\} \times \{-1, 0, 1\}$ teste, ob genau ein Punkt $q \in Q$ existiert, der im Rechteck $R_{ij} \subset \mathbb{R}^2$ definiert als

$$R_{ij} := [x + id, x + (i + 1)d] \times [y + id, y + (i + 1)d]$$

liegt.

Falls nein, suche anderen Startpunkt q_1 bzw. beende Algorithmus falls bereits alle Punkte auf ihre Fähigkeit als Startpunkt zu dienen überprüft wurden.

Falls ja, füge für alle $(i, j) \in \{-1, 0, 1\} \times \{-1, 0, 1\}$ den Punkt $q \in Q \cap R_{ij}$ zur Menge der konsistent parametrisierten Punkte hinzu: Setze $P := P \cup \{(i, j), q\}$.

Lösche Punkt aus Menge der noch nicht parametrisierten Punkte: Setze $Q := Q \setminus \{q\}$.

definiert werden. Wir benötigen allerdings eine Basis für Abbildungen von $\mathbb{R}^2 \rightarrow \mathbb{R}$. Das Tensorprodukt $f \otimes g : \mathbb{R}^2 \rightarrow \mathbb{R}$ zweier Funktionen $f, g : \mathbb{R} \rightarrow \mathbb{R}$ ist definiert als $(f \otimes g)(x, y) := f(x) \cdot g(y)$. Nun können wir den Funktionenraum der von der Tensorprodukt-Spline-Basis

$$M := \{N_{i,k+1} \otimes N_{j,k+1} \mid i \in \{-g, \dots, k\}, j \in \{-g, \dots, k\}\}$$

erzeugt wird für die Lösung des Approximationsproblems heranziehen. Für jedes $i \in \{-k, \dots, g+k+1\}$ werden die Mengen $\{(x, \tau_i) \mid x \in \mathbb{R}\}$ und $\{(\tau_i, y) \mid y \in \mathbb{R}\}$ horizontale bzw. vertikale Knotenlinien genannt. Die Zwischenräume zwischen den Knotenlinien werden Knotenrechtecke genannt. Die Tensorprodukt-Splines sind stückweise bivariate Polynome die entlang der Knotenlinien aneinandergestückt sind. Zur Vereinfachung der Notation definieren wir $n := |M| = (g+k+1)^2$, eine bijektive Abbildung $\sigma : \{-k, \dots, g\} \times \{-k, \dots, g\}$ und $\varphi_{\sigma(i,j)} := N_{i,k+1} \otimes N_{j,k+1}$ für alle $i, j \in \{-k, \dots, g\}$. Wegen der lokalen Trägereigenschaft der Splines (Siehe [Die93].) muß das Rechteck $[\tau_0, \tau_{g+1}] \times [\tau_0, \tau_{g+1}]$ alle Parameterpunkte in einer Iteration enthalten. Es muß also die folgende Ungleichung für alle Knoten gelten $\tau_0 < \Pi_j(x_i) < \tau_g$ mit $j = 1, 2$ und $i = 1, \dots, m+l$. Es ist wichtig, daß diese Ungleichung auch für die Parameterpunkte der extrapolierten Spotzentren gelten muß.

Durch Ausnutzung der speziellen Gitterstruktur der Parameterpunkte (Sie bilden in jeder Iteration ein vollständiges Gitter.) können die beiden linearen Ausgleichsprobleme in Gleichung [xxx] in vier kleinere zerlegt werden, was den Algorithmus in seiner Laufzeit erheblich beschleunigt. Die Details können in [Die93] nachgeschlagen werden.

5.2.4 Praktische Umsetzung

Für die Wahl der Funktionenräume $\mathcal{F}(\mathbb{R}^2, \mathbb{R})$ haben sich Tensorprodukte stückweise affin-linearer Funktionen oder parabolische Splines bewährt. Manchmal ist es auch ratsam für die ersten paar Iterationen stückweise affin-lineare Funktionen zu verwenden und später den Grad der Splines um eins zu erhöhen. Wir verwenden einen äquidistanten Knotenlinienentwurf, bei dem bei einer Passung maximal 6x6 Parameterpunkte in einem Knotenrechteck liegen. Dieser Entwurf erscheint geeignet, da die Parameterpunkte gleichmäßig verteilt sind und ein vollständiges Gitter bilden. Er hat sich auch in der Praxis bewährt. Die Knotenlinien werden in jeder Iteration neu bestimmt. Die Theorie der Tensorprodukt-Splines kann zum Beispiel in [Die93] nachgeschlagen werden.

5.2.5 Spot-Zuordnung bei nicht stetig differenzierbaren Wellenfronten

5.2.6 Spot-Zuordnung bei unstetigen Wellenfronten

Ein Vergleich dieser drei Verfahren ist in [CTO00] zu finden.

Broadband-Phasing

Siehe [CTD⁺98].

Narrowband Phasing

Siehe [COT00].

Phase Discontinuity Sensing

Siehe [CTS99].

5.2.7 Komplexität und Laufzeit

In jeder Iteration des Algorithmus sind die beiden linearen Ausgleichsprobleme aus 5.1 zu lösen. Da die Parameterpunkte ein vollständiges Gitter bilden und wir einen Tensorprodukt-Funktionenraum für die Passung verwenden, können diese zwei linearen Ausgleichsprobleme, wie in [Zim97] dargestellt wird, in vier kleinere zerlegt werden. Dadurch verringert sich die Anzahl der benötigten Floating-Point-Operationen (**Flops**) erheblich. Zur Vereinfachung gehen wir von quadratischen Parameterpunktmenge aus. Seien in einer Iteration m Eingabepunkte gegeben, dann hat das zugehörige Parameterpunktmengequadrat die Kantenlänge \sqrt{m} . Gemäß dem in Kapitel 5.2.4 gemachten Anmerkungen verwenden wir einen äquidistanten Knotenlinienentwurf und lassen eine bestimmte maximale $l \cdot l$ mit $l \in \mathbb{N}$ Parameterpunkte in einem Knotenrechteck zu. Das heißt die Anzahl der inneren Knoten g in horizontaler und vertikaler Richtung hängt von der Anzahl der Parameterpunkte $g = \lfloor \frac{\sqrt{m}}{l} \rfloor$ ab. Für die Lösung eines linearen Ausgleichsproblems mit Householder-Spiegelungen werden nach [Zim97]

$$\mathcal{O}(2m(g+k+1) + \sqrt{m}(\sqrt{m} + g + k + 1)(g + k + 1))$$

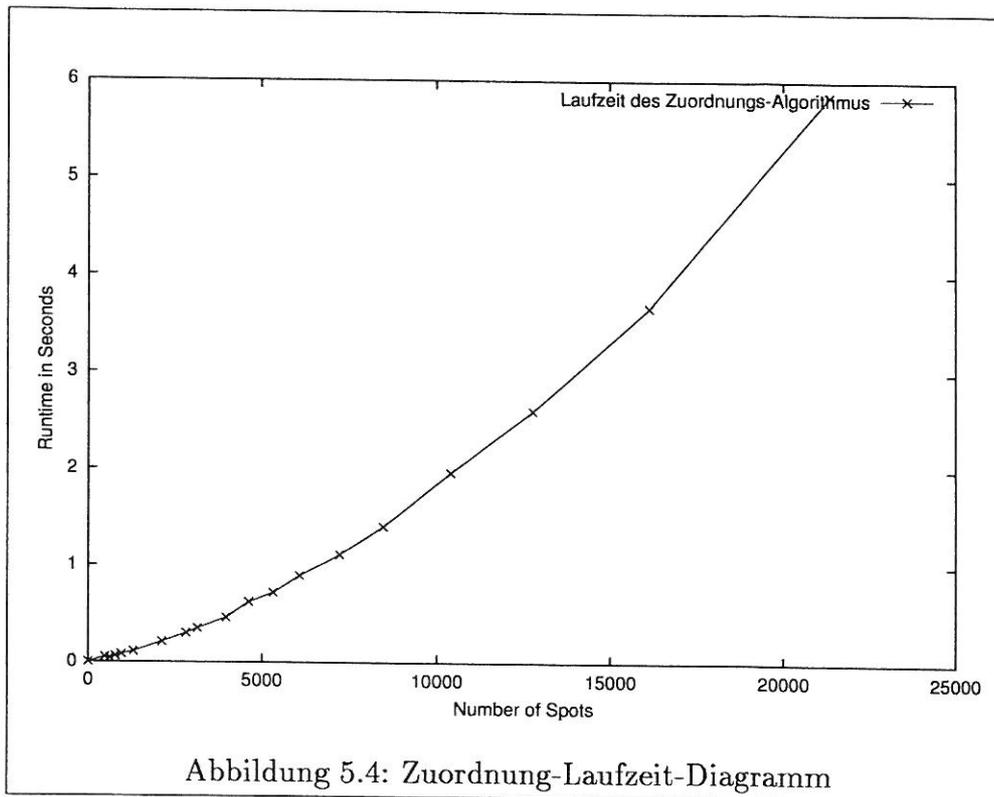
Flops benötigt. Für den Aufwand S_i in der i -ten Iteration gilt also

$$\begin{aligned}
 S_i &= \mathcal{O}(8m(g+k+1) + 4\sqrt{m}(\sqrt{m}+g+k+1)(g+k+1)) \\
 &= \mathcal{O}\left(8m\left(\left\lfloor\frac{\sqrt{m}}{l}\right\rfloor + k + 1\right) + \right. \\
 &\quad \left. + 4\sqrt{m}\left(\sqrt{m} + \left\lfloor\frac{\sqrt{m}}{l}\right\rfloor + k + 1\right)\left(\left\lfloor\frac{\sqrt{m}}{l}\right\rfloor + k + 1\right)\right) \\
 &= \mathcal{O}\left(m^{\frac{3}{2}}\right). \tag{5.2}
 \end{aligned}$$

Sei ein quadratisches Spotfeld mit n Spots gegeben und wird das Parameterpunktmengenquadrat in jeder Iteration auf jeder Seite um genau eine Punktreihe erweitert, ergibt sich der Gesamtaufwand

$$\begin{aligned}
 S &= (3 \cdot 3)^{\frac{3}{2}} + (5 \cdot 5)^{\frac{3}{2}} + \cdots + (\sqrt{n} \cdot \sqrt{n})^{\frac{3}{2}} \\
 &\leq (1 \cdot 1)^{\frac{3}{2}} + (2 \cdot 2)^{\frac{3}{2}} + \cdots + (\sqrt{n} \cdot \sqrt{n})^{\frac{3}{2}} \\
 &= \sum_{i=1}^{\sqrt{n}} i^3 \quad (\text{Siehe [GR81].}) \\
 &= \frac{1}{2}(n^2 + 2n^{\frac{2}{3}} + n) \\
 &= \mathcal{O}(n^2). \tag{5.3}
 \end{aligned}$$

In Abb. 5.4 ist die Laufzeit bei gegebener Anzahl Spots für quadratische Spotfelder aufgetragen. Es ist gut ersichtlich, daß der konstante Faktor, der sich in der Komplexitäts-Abschätzung $\mathcal{O}(n^2)$ verbirgt, relativ klein ist. Für kreisförmige Spotfelder mit Durchmesser d Spots ergibt sich der gleiche Rechenaufwand wie für quadratische Spotfelder mit Kantenlänge d , da der Algorithmus grundsätzlich auf rechteckigen Spotfeldern arbeitet. Bei Nichtvorhandensein realer Spots in den Randbereichen werden an ihrer Stelle extrapolierte Spotpositionen als Eingabe für die nächste Iteration verwendet, damit sichergestellt ist, daß der Algorithmus immer auf vollständigen Gittern von Parameterpunkten arbeitet.



5.3 Vergleich mit dem Stand der Technik

Im Folgenden wird der vorgestellte Algorithmus mit dem Stand der Technik verglichen. Es wird der nach dem Stand der Technik beste Algorithmus (Kontinuierungs-Algorithmus) als Referenz angenommen. Der nach der statischen Spotzuordnungsmethode mögliche Dynamikbereich wird ebenfalls angegeben.

5.3.1 Statische Spotzuordnung

Die folgenden Ausführungen sind sinngemäß aus [Lin99] entnommen. Wenn keine Subapertur-Überschreitungen zugelassen werden, kann der maximal meßbare Wellenfront-Steigungs-Winkel φ_{\max} über einer Mikrolinse durch

$$\varphi_{\max} \approx \sin \varphi_{\max} \approx \tan \varphi_{\max} = \frac{D}{2f}$$

abgeschätzt werden. Dabei wurden kleine Winkel angenommen, so daß der Tangens gleich dem Winkel selbst im Bogenmaß gesetzt werden darf. $D \in \mathbb{R}$

bezeichnet den Durchmesser der Mikrolinsen und $f \in \mathbb{R}$ deren Fokallänge. Damit ergibt sich für verschiedene Wellenfronttypen ein bestimmter maximal meßbarer **Wellenfronthub**, im Folgenden als Peak-To-Valley-Wert oder kurz **PTV-Wert** bezeichnet.

Eine ebene Welle, die unter einem Winkel $\alpha \in \mathbb{R}$ auf das Linsenarray trifft, kann dann eindeutig gemessen werden, wenn gilt: $\alpha \leq \varphi_{\max}$. Da die Phasenfunktion einer ebenen Welle gleich $W(x) := ax$ ist und die Ableitung nach x gleich a , gilt wegen der Eikonalgleichung:

$$a_{\max} = \frac{\partial W_{\max}}{\partial x} = \frac{\sin \varphi_{\max}}{\lambda} \approx \frac{\varphi_{\max}}{\lambda} \approx \frac{D}{2f\lambda}$$

λ bezeichnet die Wellenlänge des einfallenden Lichts. Da sich die Referenzpunkt in der Mitte jeder Subapertur befinden wird für einen Sensor mit $N \times N$ Mikrolinsen ein auswertbarer Bereich von $-\frac{(N-1)D}{2}$ bis $\frac{(N-1)D}{2}$ angenommen. Der maximale Wellenfront-Hub für ebene Wellen ist damit

$$\Delta W_{\max} := \frac{(N-1)D^2}{2f\lambda}.$$

In paraxialer Näherung läßt sich eine sphärische Welle mit Krümmungsradius R als parabolische Näherung schreiben mit $W(x) := x^2/(2\lambda R)$. Die maximale Steigung dieser parabolischen Welle ergibt sich am Rand des Sensors, d.h. bei $x_{\max} = \frac{(N-1)D}{2}$:

$$\frac{\varphi_{\max}}{\lambda} \approx \frac{\partial W_{\max}}{\partial x} = \frac{x_{\max}}{\lambda R_{\min}} = \frac{(N-1)D}{2\lambda R_{\min}}.$$

Der minimal noch meßbare Krümmungsradius R_{\min} ist also:

$$R_{\min} = (N-1)f.$$

Der maximal meßbare Wellenfront-Hub für eine sphärische (parabolische) Welle beträgt damit also:

$$\Delta W_{\max} = \frac{x_{\max}^2}{2\lambda R_{\min}} = \frac{(N-1)^2 D^2}{8\lambda(N-1)f} = \frac{(N-1)D^2}{8f\lambda}.$$

Er ist also 1/4 des Werts für eine ebene Welle.

Für den im FORMIKROSYS II Projekt Wellensensor verwendeten Sensor

ergibt sich damit für $N = 42$, $D = 150\mu m$, $f = 4mm$ und $\lambda = 0.6328\mu m$ ein maximaler Wellenfronthub von

$$\Delta W_{\max} = \frac{(N-1)D^2}{2f\lambda} \approx 182\lambda$$

für eine ebene Welle und

$$\Delta W_{\max} = \frac{(N-1)D^2}{8f\lambda} \approx 46\lambda$$

für eine sphärische bzw. parabolische Welle.

5.3.2 Kontinuierung

Der Kontinuierungs-Algorithmus stellt an den Spotverschiebungs-Unterschied zwischen benachbarten Subaperturen folgende Bedingung:

$$|\sigma_1 - \sigma_2| < \frac{D}{2}.$$

Für ebene Wellen gibt es demnach von algorithmischer Seite keine Beschränkung. Für sphärische und parabolische Wellenfronten können wir uns wegen Rotationssymmetrie auf einen Schnitt entlang der x-Achse beschränken. Sei also $W_0 : \mathbb{R} \rightarrow \mathbb{R}$ definiert als $W_0(x) = W(x, 0)$; Für benachbarte Subaperturen auf der x-Achse am Rand des Sensors also bei $x_{\max} = \frac{(N-1)D}{2}$ muß also gelten

$$\begin{aligned} |fW_0'(x_{\max}) - fW_0'(x_{\max} - D)| &< \frac{D}{2} \\ |W_0'(x_{\max}) - W_0'(x_{\max} - D)| &< \frac{D}{2f} \end{aligned}$$

Für parabolische Wellen ist W_0'' konstant und darf nicht größer sein als $\frac{1}{2f}$. Damit ergibt sich ein maximaler lokaler Krümmungsradius $r = \frac{1}{W_0''(0)} = 2f$. Für den im Projekt verwendeten Sensor ist damit die Vermessung von Wellenfronten mit lokalen Krümmungsradien bis hinunter auf 8mm möglich. Für eine parabolische Wellenfront der Form $W_0(x) = ax^2$ mit Krümmungsradius $r = 8[mm]$ gilt $W_0''(x) = 2a$ und wegen $r = \frac{1}{2a} = 8$ ist $a = \frac{1}{16}$. Der maximal meßbare Wellenfronthub für parabolische Wellenfronten ergibt sich also zu

$$\Delta W_{\max} = \frac{|W_0(0) - W_0(x_{\max})|}{\lambda} \approx 934\lambda.$$

Sei der maximale Wellenfronthub $\Delta W_{\max} \in \mathbb{R}$ einer sphärischen Wellenfront gegeben, dann ist mit $r = \frac{(N-1)D}{2}$, $r_S = (r^2 + (\Delta W_{\max})^2)/(2\Delta W_{\max})$, $C = \left\{ (x, y) \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} \leq \frac{(N-1)D}{2} \right\}$, die Funktion $W : C \rightarrow \mathbb{R}$

$$W(x, y) = \sqrt{r_S^2 - x^2 - y^2}$$

eine sphärische Wellenfront mit der Eigenschaft

$$|W_0(x_{\max}) - W_0(0)| = \Delta W_{\max}$$

für $W_0 : [-(N-1)D/2, (N-1)D/2] \rightarrow \mathbb{R}$ mit $W_0(x) = W(x, 0)$. Man kann nun die erste Ableitung W'_0 berechnen und zu gegebener maximaler Spotablenkung den zugehörigen Wellenfronthub ausrechnen, den eine sphärische Wellenfront haben darf, damit eine Spotzuordnung durch den Kontinuierungs-Algorithmus noch möglich ist. Dieser beträgt 817λ .

Bei der Asphären-Vermessung mit sphärischer Kompensation treten Wellenfronten der Form

$$W(x, y) = \frac{aCr^2}{1 + \sqrt{1 - (1 + K)C^2r^2}} - \frac{aCr^2}{1 + \sqrt{1 - C^2r^2}}$$

auf, mit $r = \sqrt{x^2 + y^2}$, mit dem Skalierungsfaktor $a \in \mathbb{R}$, dem Krümmungsfaktor $C \in \mathbb{R}$ und dem konischen Faktor $K \in \mathbb{R}$. Für $K = 0$ erhält man eine Sphäre. Für den Krümmungsradius $R \in \mathbb{R}$ gilt $C = 1/R$.

$$a_4 = 0.000420517$$

$$a_6 = 4.00605e - 005$$

$$a_8 = -1.10217e - 005$$

$$a_{10} = 6.91263e - 007$$

$$C = 0.23474$$

$$K = -0.485124$$

Läßt man den Skalierungsfaktor a variieren und erzeugt damit Wellenfronten mit verschiedenen PTV-Werten, so beträgt der maximale Wellenfronthub, der noch mit dem Kontinuierungs-Algorithmus vermessbar ist, 78λ .

5.3.3 Spotzuordnung durch iterative Funktionenpassung

Für ebene Wellen gibt es genau wie bei der Kontinuierung keine Beschränkungen. Für sphärische Wellenfronten kann eine Spotzuordnung bis zu einem PTV-

Wert von 1000λ , für die oben angegebenen asphärischen Wellenfronten bis zu einem PTV-Wert von 110λ durchgeführt werden. Dies bedeutet einen erheblichen Fortschritt gegenüber dem Stand der Technik.

Kapitel 6

Wellenfront-Rekonstruktion

6.1 Bekannte Verfahren zur Wellenfront-Rekonstruktion

Definition Zernike-Polynome

$$Z_n^l(r, \phi) := R_n^l(r) e^{il\phi} \quad (6.1)$$

mit

$$n \geq 0, n \geq |l| \quad \text{und} \quad n - |l| = \text{gerade Zahl.}$$

Definition Radial-Polynome:

$$R_n^{\pm m}(r) = \sum_{s=0}^{\frac{n-m}{2}} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+m}{2} - s\right)! \left(\frac{n-m}{2} - s\right)!} r^{n-2s}, \quad (6.2)$$

mit

$$m = |l|.$$

Eigenschaften:

Orthogonalität der Zernike-Polynome auf dem Einheitskreis:

$$\int_0^1 \int_0^{2\pi} r \cdot (Z_n^l(r, \phi))^* \cdot Z_{n'}^{l'}(r, \phi) dr d\phi = \frac{\pi}{n+1} \delta_{nn'} \delta_{ll'} \quad (6.3)$$

Orthogonalität der Radial-Polynome auf dem Einheitsradius:

$$\int_0^1 r \cdot R_n^m(r) \cdot R_{n'}^m(r) dr = \frac{1}{2(n+1)} \delta_{nn'}. \quad (6.4)$$

6.2 Wellenfront-Rekonstruktion durch Spline-Passung

6.2.1 Mathematische Grundlagen

Seien das Definitionsrechteck $R \subset \mathbb{R}^2$ definiert als $R := [a, b] \times [c, d]$ mit $a, b, c, d \in \mathbb{R}$ und $a < b$ sowie $c < d$, die Meßstellen $x_1, \dots, x_m \in R$, Meßwerte $y_1, \dots, y_m \in \mathbb{R}^2$ mit $m \in \mathbb{N}_{>}$ gegeben. Die Meßwerte y_i sollen als partielle Ableitungen einer zu approximierenden Abbildung $W : R \rightarrow \mathbb{R}$ interpretiert werden. Dabei sei

$$y_i = \begin{pmatrix} \frac{\partial W}{\partial x}(x_i) \\ \frac{\partial W}{\partial y}(x_i) \end{pmatrix} \quad \text{für alle } i \in \{1, \dots, m\}.$$

Gegeben sei des weiteren ein endlich-dimensionaler Vektorraum von Funktionen, die von R nach \mathbb{R} abbilden, $\mathcal{F}(R, \mathbb{R})$ und eine Basis $\{\varphi_1, \dots, \varphi_n \in \mathcal{F}(R, \mathbb{R})\}$ mit $n \in \mathbb{N}_{>}$. Es sei sichergestellt, daß $m > n$ gilt.

Gesucht ist nun eine Abbildung $\Phi \in \mathcal{F}(R, \mathbb{R})$, deren partielle Ableitungen an den gegebenen Stützstellen die gegebenen Meßwerte möglichst gut approximieren, die also den Quadratmittelfehler

$$\sum_{i=1}^m \left(\frac{\partial \Phi}{\partial x}(x_i) - \Pi_1(y_i) \right)^2 + \sum_{i=1}^m \left(\frac{\partial \Phi}{\partial y}(x_i) - \Pi_2(y_i) \right)^2 \quad (6.5)$$

minimieren.

Seien $A \in \mathbb{R}^{2m \times n}$ und $b \in \mathbb{R}^{2m}$ definiert als

$$A := \begin{pmatrix} \frac{\partial \varphi_1}{\partial x}(x_1) & \frac{\partial \varphi_2}{\partial x}(x_1) & \dots & \frac{\partial \varphi_n}{\partial x}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varphi_1}{\partial x}(x_m) & \frac{\partial \varphi_2}{\partial x}(x_m) & \dots & \frac{\partial \varphi_n}{\partial x}(x_m) \\ \frac{\partial \varphi_1}{\partial y}(x_1) & \frac{\partial \varphi_2}{\partial y}(x_1) & \dots & \frac{\partial \varphi_n}{\partial y}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varphi_1}{\partial y}(x_m) & \frac{\partial \varphi_2}{\partial y}(x_m) & \dots & \frac{\partial \varphi_n}{\partial y}(x_m) \end{pmatrix} \quad \text{und } b := \begin{pmatrix} \Pi_1(y_1) \\ \vdots \\ \Pi_1(y_m) \\ \Pi_2(y_1) \\ \vdots \\ \Pi_2(y_m) \end{pmatrix},$$

dann können wir die Koeffizienten der gesuchten bestapproximierenden Abbildung durch die Lösung des linearen Ausgleichsproblems

$$\min_{\xi \in \mathbb{R}^n} \|A\xi - b\| \quad (6.6)$$

bestimmen. Die gesuchte Abbildung Φ ist dann definiert als

$$\Phi := \sum_{i=1}^n \xi_i \varphi_i. \quad (6.7)$$

6.2.2 Praktische Umsetzung

Seien Spotverschiebungen $\sigma_1, \dots, \sigma_m \in \mathbb{R}^2$ and den Stellen $x_1, \dots, x_m \in \mathbb{R}^2$ gegeben, wobei die erste Koordinate jeweils die Verschiebung in horizontaler und die zweite Koordinate die Verschiebung in vertikaler Richtung angibt. Ferner sei der Abstand $f \in \mathbb{R}$ der Fokalebene vom Mikrolinsenarray bekannt. Dann können durch $y_i \in \mathbb{R}^2$ definiert als

$$y_i := \frac{1}{\sqrt{f^2 + \|\sigma\|^2}} \sigma_i \quad \text{für alle } i \in \{1, \dots, m\}$$

die Steigungen der Wellenfront an den Stellen x_1, \dots, x_m angenähert werden. Für nähere Erläuterungen des Funktionsprinzips des Shack-Hartmann-Sensors sei an dieser Stelle auf [Gea95] verwiesen.

6.2.3 Polarkoordinaten-Rekonstruktion

Ein Nachteil von Tensorprodukt-Splines ist, daß ihr Definitionsbereich auf rechteckige Bereiche beschränkt ist. Da es sich bei den im Rahmen des Projekts Wellensensor zu vermessenden Linsen aber oft um runde Objekte handelt, sind auch die Spotfelder eher rund und damit ist auch die Menge der Stützstellen und der Definitionsbereich der zu rekonstruierenden Wellenfront eher kreisförmig. Natürlich kann man immer einen rechteckigen Bereich finden, der diesen runden Definitionsbereich enthält und innerhalb dieses rechteckigen Bereichs Knotenlinien wählen, jedoch werden dabei je nach Wahl des Knotenliniengitters mehr oder weniger viele und mehr oder weniger große Knotenrechtecke keine Stützstellen enthalten. Dies erhöht das Risiko eine rangdefekte Systemmatrix zu erhalten und damit auch das Risiko numerischer Instabilität der Approximationsalgorithmen. Eine andere Möglichkeit ist die Transformation der Stützstellen aus dem kartesischen Koordinatensystem in ein Polarkoordinatensystem. Damit der approximierende Spline immer noch bestimmte Glattheitseigenschaften erfüllt ist die Einführung bestimmter Nebenbedingungen unumgänglich.

Sei $z : C \rightarrow \mathbb{R}$ eine Funktion mit beschränkten Ableitungen auf der Einheitskreisscheibe $C := \{(x, y) \mid x^2 + y^2 \leq 1\}$. Wird sie in Polarkoordinaten

dargestellt, so hat die Funktion $s : R \rightarrow \mathbb{R}$ definiert als

$$s(u, v) := z(u \cos v, u \sin v) \quad (6.8)$$

ebenfalls beschränkte Ableitungen auf dem Rechteck $R := [0, 1] \times [-\pi, \pi]$ und erfüllt zusätzlich einige Randwerteigenschaften.

Sei $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definiert als

$$\sigma(x, y) := \left(\sqrt{x^2 + y^2}, \operatorname{atn}(x, y) \right)$$

mit $\operatorname{atn}(x, y) : \mathbb{R}^2 \rightarrow [-\pi, \pi]$ definiert als

$$\operatorname{atn}(x, y) := \begin{cases} \pi/2 & , x = 0, y \geq 0, \\ -\pi/2 & , x = 0, y < 0, \\ \arctan(y/x) & , x > 0, \\ \arctan(y/x) + \pi & , x < 0, y > 0, \\ \arctan(y/x) - \pi & , x < 0, y \leq 0. \end{cases}$$

Leider ist σ keine bijektive Abbildung, da sie nicht surjektiv ist. Auf die Menge der Punkte $\{(0, y) \mid y \in [-\pi, \pi] \setminus \{\pi/2\}\}$ wird nicht abgebildet. Sie ist allerdings injektiv.

Zunächst sieht man sehr einfach, daß s periodisch in der Variable v sein muß, da man gerne die Stetigkeit von z entlang der Halbgeraden $\{(x, 0) \mid x \in \mathbb{R}, x \leq 0\}$ erhalten möchte.

Desweiteren kann man beweisen, daß wenn z stetige partielle Ableitungen $g_{i,j} \in \mathbb{R}$ mit

$$g_{i,j} := \frac{\partial^{i+j} z(0, 0)}{\partial x^i \partial y^j} \quad \text{mit } i, j \in \mathbb{N}, i + j \leq 2$$

bis zur zweiten Ordnung im Ursprung $(0, 0)$ hat, folgende Identitäten gelten:

$$s(0, v) = g_{0,0}, \quad -\pi \leq v \leq \pi, \quad (6.9)$$

$$\frac{\partial s(0, v)}{\partial u} = g_{1,0} \cos(v) + g_{0,1} \sin(v), \quad -\pi \leq v \leq \pi, \quad (6.10)$$

$$\frac{\partial^2 s(0, v)}{\partial u^2} = g_{0,2} \cos^2(v) + g_{0,2} \sin^2(v) + g_{1,1} \sin(2v), \quad -\pi \leq v \leq \pi \quad (6.11)$$

Die Bedingung 6.9 ist trivial. Die Kettenregel

$$\begin{aligned}\frac{\partial s}{\partial u} &= \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial u}, \\ \frac{\partial^2 s}{\partial u^2} &= \frac{\partial^2 z}{\partial x^2} \cdot \left(\frac{\partial x}{\partial u}\right)^2 + \frac{\partial^2 z}{\partial y^2} \cdot \left(\frac{\partial y}{\partial u}\right)^2 + 2 \frac{\partial^2 z}{\partial x \partial y} \cdot \frac{\partial x}{\partial u} \cdot \frac{\partial y}{\partial u} + \\ &\quad + \frac{\partial z}{\partial x} \cdot \frac{\partial^2 x}{\partial u^2} + \frac{\partial z}{\partial y} \cdot \frac{\partial^2 y}{\partial u^2}\end{aligned}$$

angewendet auf 6.8 und für $u = 0$ ausgewertet, ergibt automatisch die beiden Bedingungen 6.10 und 6.11. Später werden wir versuchen diese Bedingungen auf einen Tensorprodukt-Spline vom Grad k mit $k \geq 3$ anzuwenden. Deshalb ist es wichtig, daß diese Bedingungen ausreichen die C^2 -Stetigkeit für z im Ursprung zu garantieren. Für die Existenz der ersten partiellen Ableitungen ist es notwendig, daß die Gleichungen

$$\begin{aligned}g_{1,0} &= \lim_{x \rightarrow 0_+} \frac{z(x, 0) - z(0, 0)}{x} = \lim_{x \rightarrow 0_-} \frac{z(x, 0) - z(0, 0)}{x}, \\ &= \lim_{u \rightarrow 0_+} \frac{s(u, 0) - s(0, 0)}{u} = \lim_{u \rightarrow 0_+} \frac{s(u, 0) - s(0, 0)}{-u}, \\ &= \frac{\partial s(0, 0)}{\partial u} = -\frac{s(0, \pi)}{\partial u},\end{aligned}\tag{6.12}$$

$$g_{0,1} = \frac{\partial s(0, \frac{\pi}{2})}{\partial u} = -\frac{s(0, -\frac{\pi}{2})}{\partial u}\tag{6.13}$$

gelten. Um die Stetigkeit dieser Ableitungen zu zeigen müssen wir überprüfen, daß

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\partial z(x, y)}{\partial x} = g_{1,0},\tag{6.14}$$

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\partial z(x, y)}{\partial y} = g_{0,1}\tag{6.15}$$

gilt. Durch erneute Anwendung der Kettenregel, können wir für $(x, y) \neq (0, 0)$ schreiben

$$\begin{aligned}\frac{\partial z(x, y)}{\partial x} &= \frac{\partial s(u, v)}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial s(u, v)}{\partial v} \cdot \frac{\partial v}{\partial x} \\ &= \cos(v) \cdot \frac{\partial s(u, v)}{\partial u} - \frac{\sin(v)}{u} \cdot \frac{\partial s(u, v)}{\partial v},\end{aligned}\tag{6.16}$$

$$\frac{\partial z(x, y)}{\partial y} = \sin(v) \cdot \frac{\partial s(u, v)}{\partial u} + \frac{\cos(v)}{u} \cdot \frac{\partial s(u, v)}{\partial v}\tag{6.17}$$

und folgendes erweitern

$$\frac{\partial s(u, v)}{\partial u} = \frac{\partial s(0, v)}{\partial u} + u \frac{\partial^2 s(0, v)}{\partial u^2} + \dots, \quad (6.18)$$

$$\frac{\partial s(u, v)}{\partial v} = \frac{\partial s(0, v)}{\partial v} + u \frac{\partial^2 s(0, v)}{\partial u \partial v} + \frac{u^2}{2} \frac{\partial^3 s(0, v)}{\partial u^2 \partial v} + \dots \quad (6.19)$$

Setzt man 6.10-6.11 in 6.18 ein und setzt man die Ableitungen nach v in 6.19, und bildet dann den Grenzwert von 6.16-6.17 für $u \rightarrow 0+$, so erhält man sofort 6.14-6.15. Auf ähnliche Art und Weise können wir die auch die Stetigkeit der Ableitungen der zweiten Ordnung beweisen. Deren Wert ist beispielsweise gegeben durch

$$g_{2,0} = \frac{\partial^2 s(0, 0)}{\partial u^2} = \frac{\partial^2 s(0, \pi)}{\partial u^2}, \quad (6.20)$$

$$g_{0,2} = \frac{\partial^2 s(0, \frac{\pi}{2})}{\partial u^2} = \frac{\partial^2 s(0, -\frac{\pi}{2})}{\partial u^2}, \quad (6.21)$$

$$\begin{aligned} g_{1,1} &= \frac{1}{2} \frac{\partial^3 s(0, 0)}{\partial u^2 \partial v} = \frac{1}{2} \frac{\partial^3 s(0, \pi)}{\partial u^2 \partial v} \\ &= -\frac{1}{2} \frac{\partial^3 s(0, \frac{\pi}{2})}{\partial u^2 \partial v} = -\frac{1}{2} \frac{\partial^3 s(0, -\frac{\pi}{2})}{\partial u^2 \partial v}. \end{aligned} \quad (6.22)$$

Wir wollen nun untersuchen, wie diese verschiedenen Randbedingungen auf einen Tensorprodukt-Spline $s : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$s(u, v) = \sum_{i=-k}^g \sum_{j=-l}^h c_{i,j} N_{i,k+1}(u) N_{j,l+1}(v) \quad (6.23)$$

angewendet werden können.

6.2.4 Rekonstruktion von nicht stetig differenzierbaren Wellenfronten

6.2.5 Rekonstruktion von unstetigen Wellenfronten

6.2.6 Komplexität und Laufzeit

6.3 Vergleich der vorgestellten Verfahren

Im Folgenden werden vier verschiedene Funktionenräume auf ihre spezielle Eignung für verschiedene Rekonstruktionsprobleme untersucht. Die Spaltenkungen für die verwendeten Wellenfronten wurden mathematisch berechnet. Als Definitionsbereich wird eine Kreisscheibe mit Radius 3.075mm

verwendet, wie sie durch die im Projekt verwendete CCD-Kamera mit quadratischen Pixeln vorgegeben war. Die Referenzpunkte ergeben sich durch das im Projekt verwendete Mikrolinsen-Array mit $n_x \times n_y = 42 \times 42$ Mikrolinsen mit Durchmesser $d := 0.15$ mm, die auf einem regelmäßigen Gitter liegen. Die Stützstellen $S := \{s_{kl} \in \mathbb{R}^2 | k \in \{0, 1, \dots, n_x - 1\}, l \in \{0, 1, \dots, n_y - 1\}\}$ haben also die Koordinaten

$$s_{kl} := (-3.075 + kd, -3.075 + ld).$$

Eine kreisförmige Blende mit Radius 3.15mm wird durch die Auswahl bestimmter Stützstellen berücksichtigt. Das heißt wir verwenden nur die Menge der Stützstellen $S_B := \{s_{kl} | \|s_{kl}\| < 3.15\}$. Seien an einer Stützstelle partielle Ableitungen d_x und d_y gegeben, dann berechnen sich die Spotablenkungen σ_x und σ_y als

$$\sigma_x = f \cdot \frac{d_x}{\sqrt{1 - d_x^2 - d_y^2}} \quad \text{und} \quad \sigma_y = f \cdot \frac{d_y}{\sqrt{1 - d_x^2 - d_y^2}}$$

Nun kann gezielt die Eignung bestimmter Funktionenräume zur Rekonstruktion verschiedener Wellenfront-Typen aufgezeigt werden.

Die Laufzeit für die Wellenfront-Rekonstruktion setzt sich im wesentlichen aus der benötigten Zeit für das Aufstellen der Systemmatrix und dem Lösen des linearen Ausgleichsproblems zusammen. Experimentell konnte ermittelt werden, daß dabei der letzte Faktor klar dominiert. Wir können uns daher auf die Betrachtung des Zusammenhangs der Anzahl der Spalten der Systemmatrix mit dem Rekonstruktionsfehler beschränken.

Als erstes wird eine rotationssymmetrische, asphärische Wellenfront, wie sie im Projekt Wellensensor vorherrschte mit einem PTV-Wert von hier 200λ untersucht. Die ideale Wellenfront hat die Form

$$W_{ideal}(x, y) := \frac{dCr^2}{1 + \sqrt{1 - (1 + K)C^2r^2}} + a_4r^4 + a_6r^6 + a_8r^8 + a_{10}r^{10} \quad (6.24)$$

mit

$$\begin{aligned}
 a_4 &:= 0.000420517 \\
 a_6 &:= 4.00605e - 005 \\
 a_8 &:= -1.10217e - 005 \\
 a_{10} &:= 6.91263e - 007 \\
 C &:= 0.23474 \\
 K &:= -0.485124 \\
 d &:= 0.0102736 \\
 r &:= \sqrt{x^2 + y^2}
 \end{aligned}$$

Die partiellen Ableitungen berechnen sich als

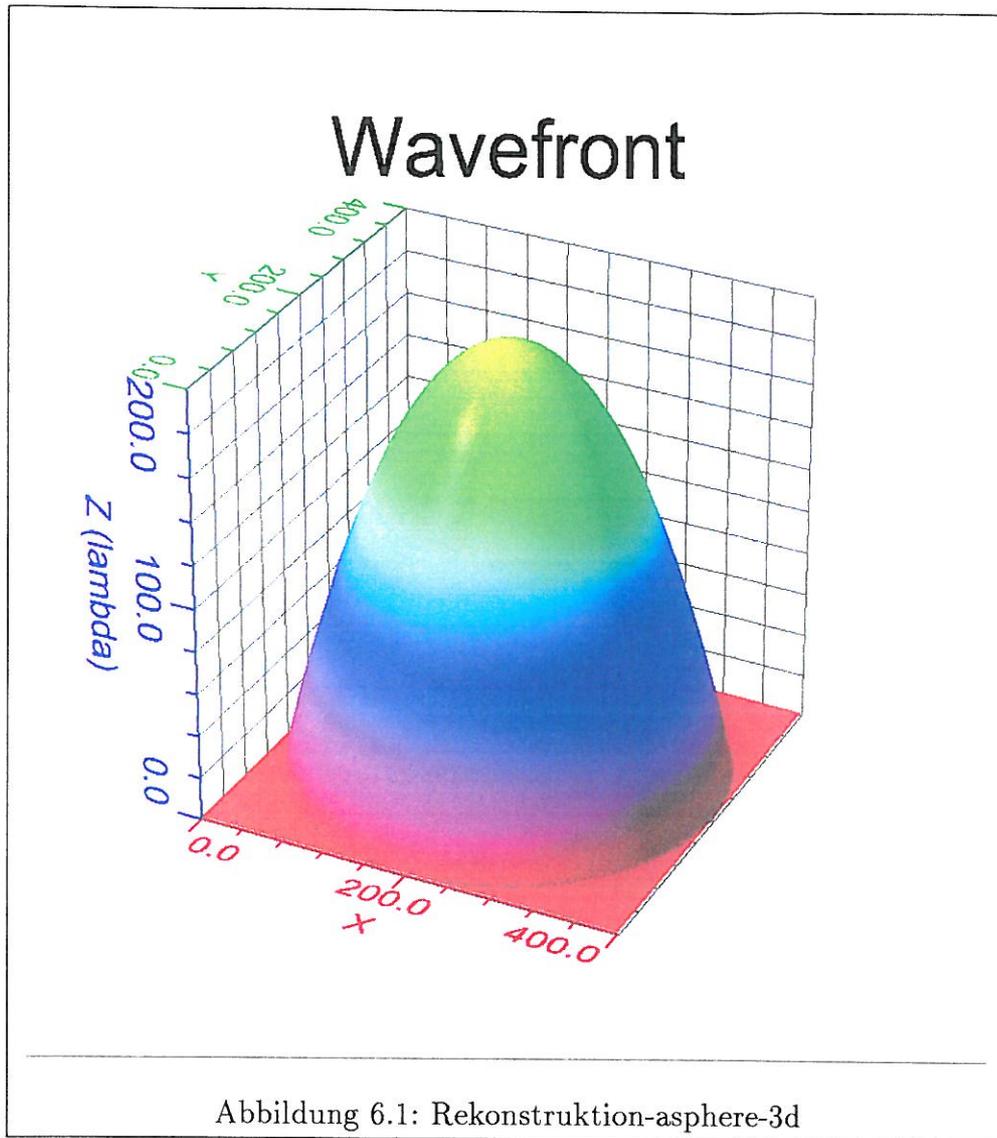
$$\begin{aligned}
 \frac{\partial W_{ideal}}{\partial x}(x, y) &= \frac{2dCx}{1 + T(x, y)} - \frac{Cr^2(-2C^2x - 2C^2Kx)}{2(1 + T(x, y))^2T(x, y)} + \\
 &\quad + 4a_4r^2x + 6a_6r^4x + 8a_8r^6x + 10a_{10}r^8x
 \end{aligned} \tag{6.25}$$

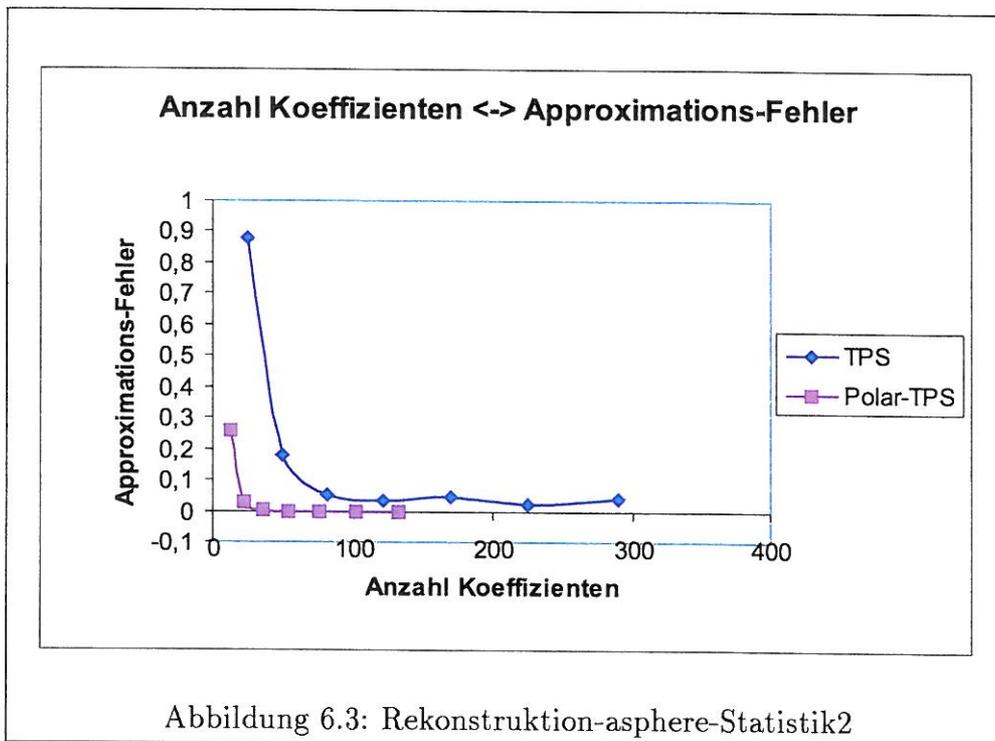
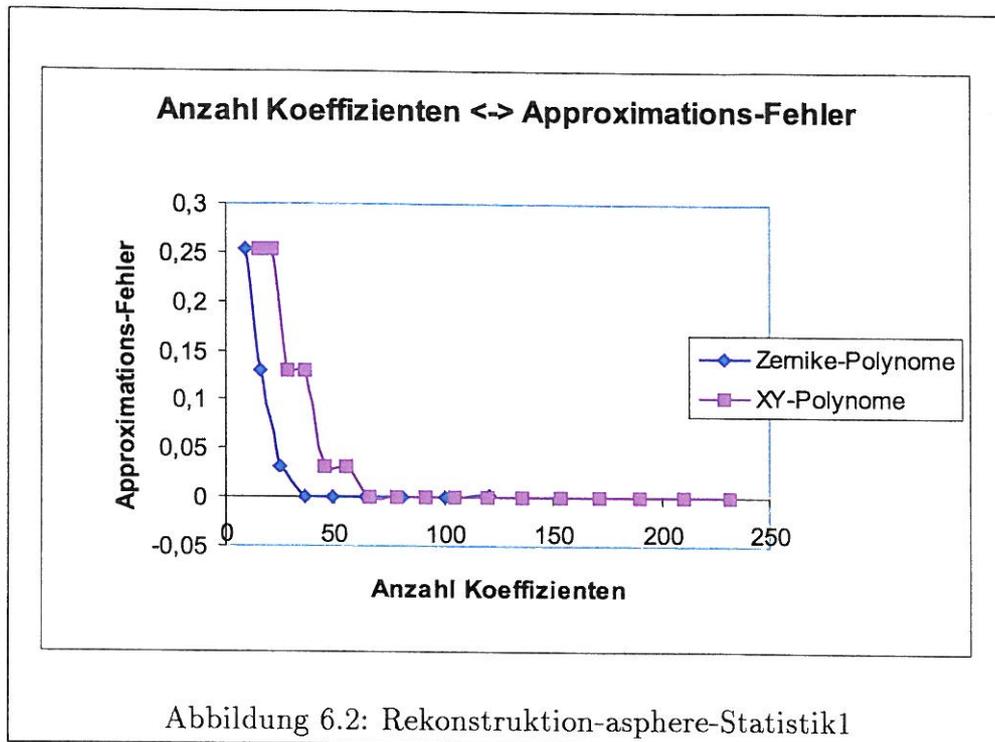
$$\begin{aligned}
 \frac{\partial W_{ideal}}{\partial y}(x, y) &= \frac{2dCy}{1 + T(x, y)} - \frac{Cr^2(-2C^2y - 2C^2Ky)}{2(1 + T(x, y))^2T(x, y)} + \\
 &\quad + 4a_4r^2y + 6a_6r^4y + 8a_8r^6y + 10a_{10}r^8y
 \end{aligned} \tag{6.26}$$

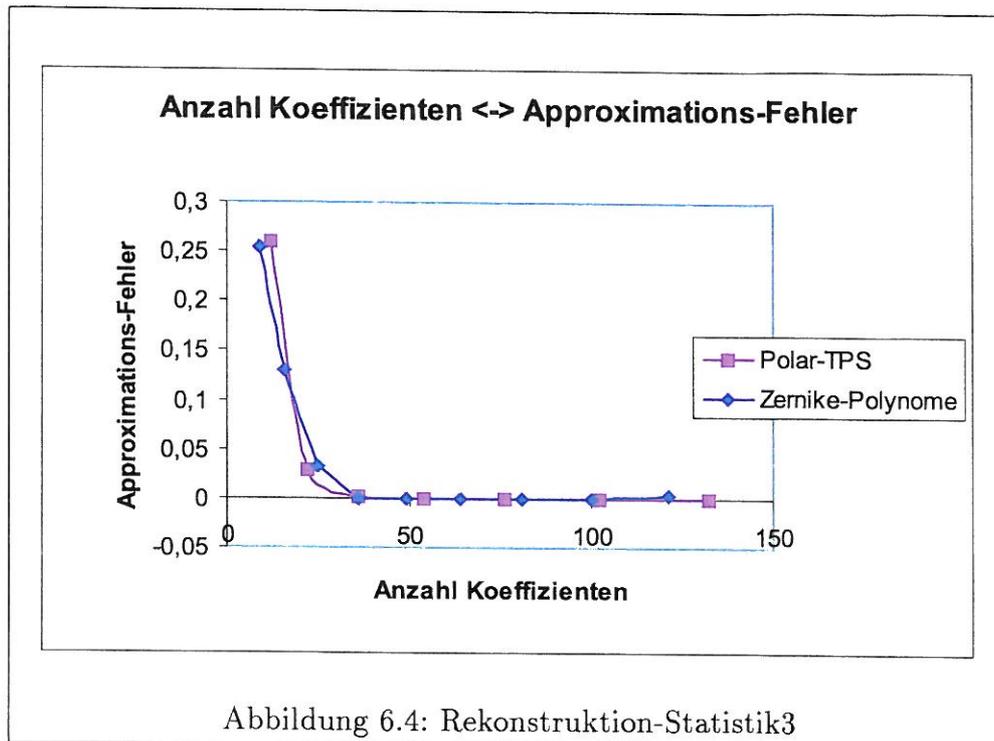
mit $T : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert als

$$T(x, y) := \sqrt{1 - C^2x^2 - C^2y^2 - C^2Kx^2 - C^2Ky^2}.$$

Die rekonstruierte Wellenfront ist in Abb. 6.1 abgebildet.







Als nächstes wird eine Wellenfront mit einem PTV-Wert von 20λ untersucht, die eine Knickstelle entlang der x-Achse aufweist. Sei

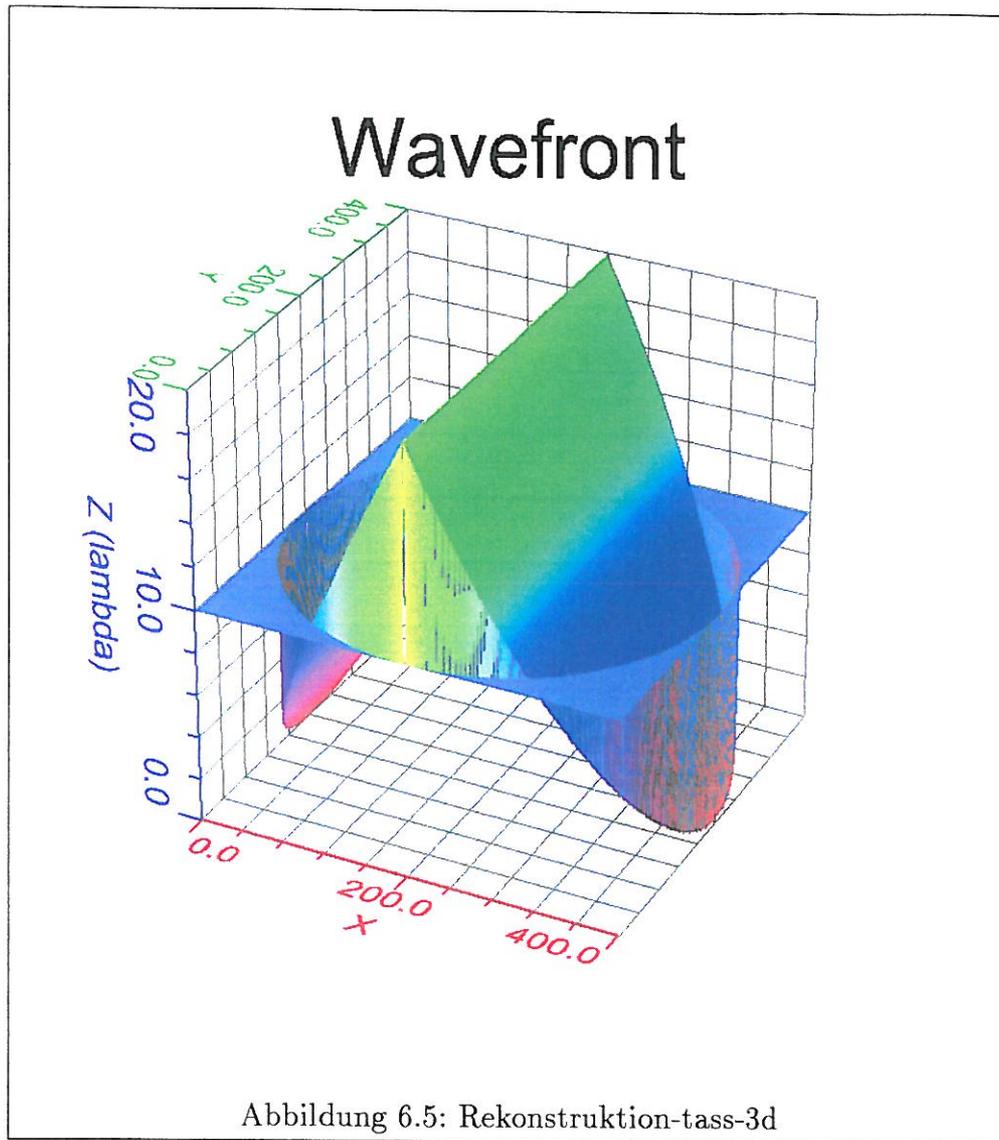
$$W_{ideal}(x, y) := \begin{cases} -d_x x & x < 0, \\ d_x x & \text{sonst.} \end{cases}$$

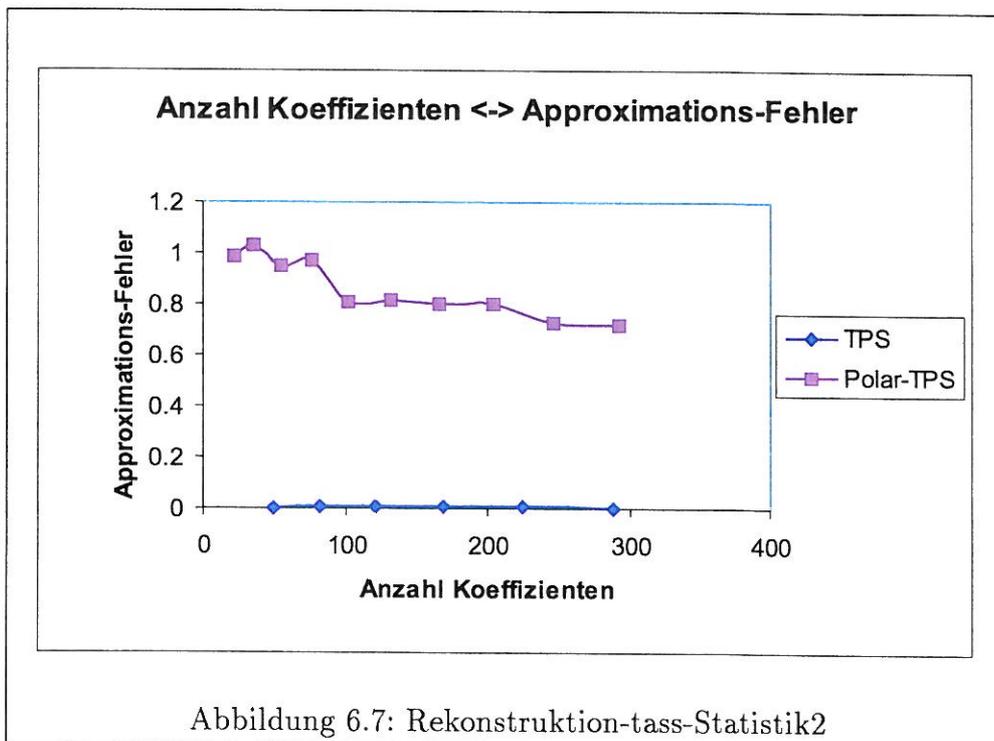
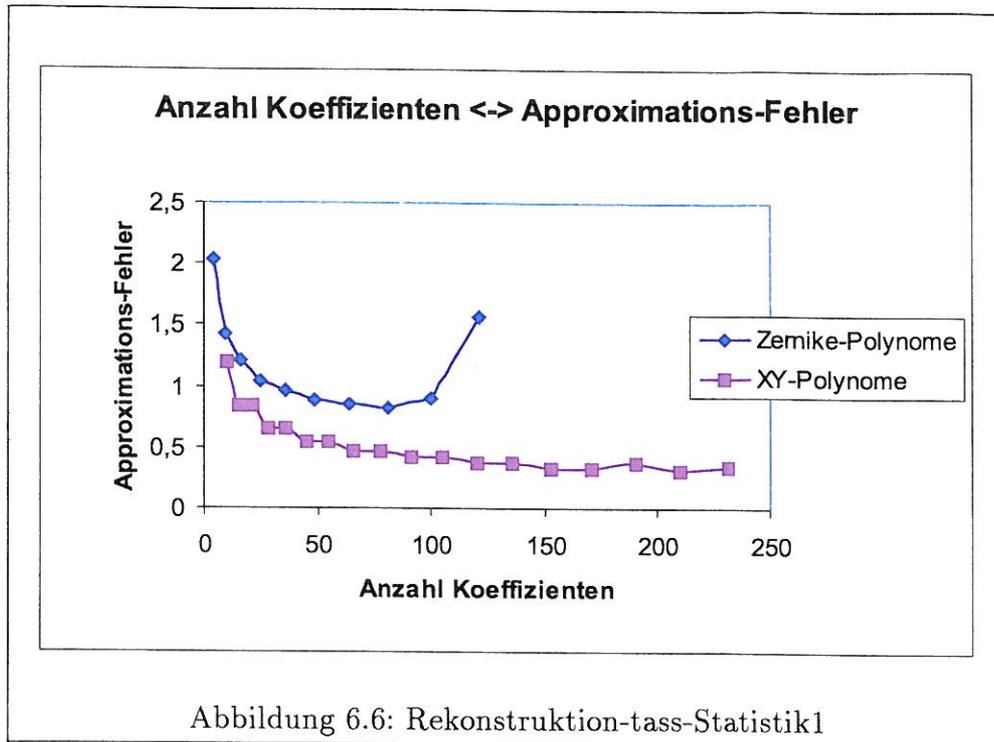
mit $d_x := 20/3.075$. Die partiellen Ableitungen sind

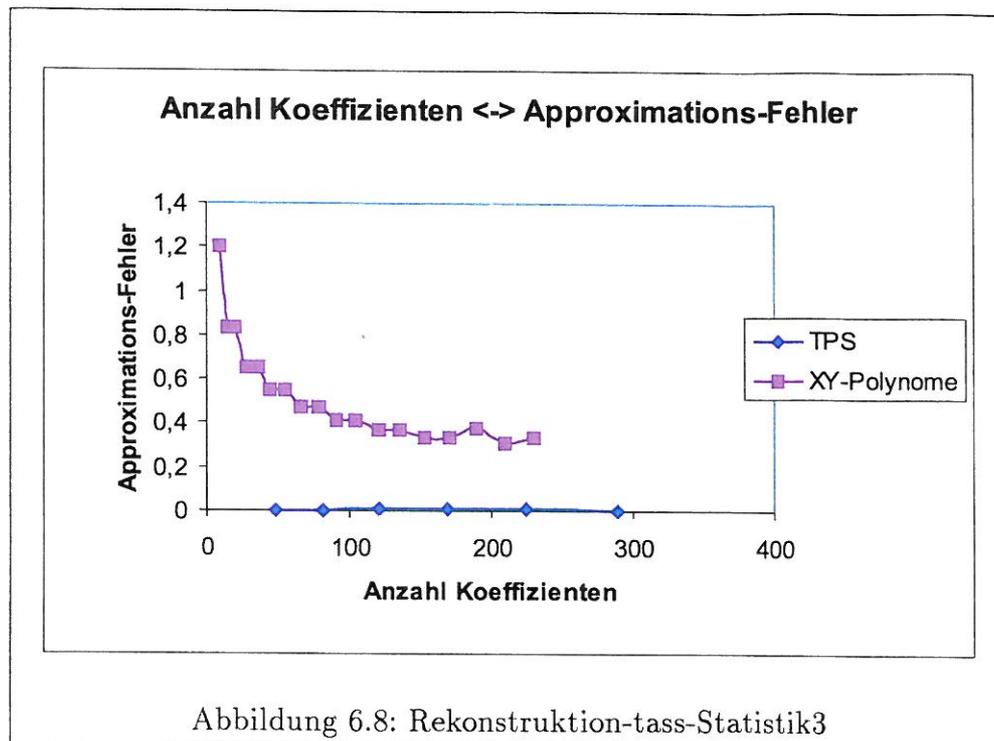
$$\frac{\partial W_{ideal}(x, y)}{\partial x} = \begin{cases} -d_x & x < 0, \\ d_x & \text{sonst,} \end{cases}$$

$$\frac{\partial W_{ideal}(x, y)}{\partial y} = 0.$$

Die rekonstruierte Wellenfront ist in Abb. 6.5 abgebildet. Zur Rekonstruktion eignen sich hierbei besonders Tensorprodukt-Splines, mit entsprechenden Knotenvervielfachungen entlang der Knickstelle.







Als letztes wird eine Wellenfront mit einem PTV-Wert von 20λ untersucht, die eine Knickstelle entlang der Linie $\sqrt{x^2 + y^2} = 2$ aufweist. Sei

$$W_{ideal}(x, y) := \begin{cases} 0 & , r \leq 2, \\ -d(r - 2)^2 & , \text{sonst.} \end{cases}$$

mit $d := 20/1.075^2$ und $r := \sqrt{x^2 + y^2}$. Der Graph dieser Funktion ist in Abb. 6.9 abgebildet. Die partiellen Ableitungen sind

$$\frac{\partial W_{ideal}(x, y)}{\partial x} = \begin{cases} 0 & , r \leq 2, \\ -2d(r - 2)x/r & , \text{sonst,} \end{cases}$$

$$\frac{\partial W_{ideal}(x, y)}{\partial y} = \begin{cases} 0 & , r \leq 2, \\ -2d(r - 2)y/r & , \text{sonst,} \end{cases}$$

Zur Rekonstruktion dieser Wellenfront eignen sich besonders Polarkoordinaten-TPS mit einer entsprechenden Knotenvervielfachung entlang der Knickstelle.

WAVEGUIT

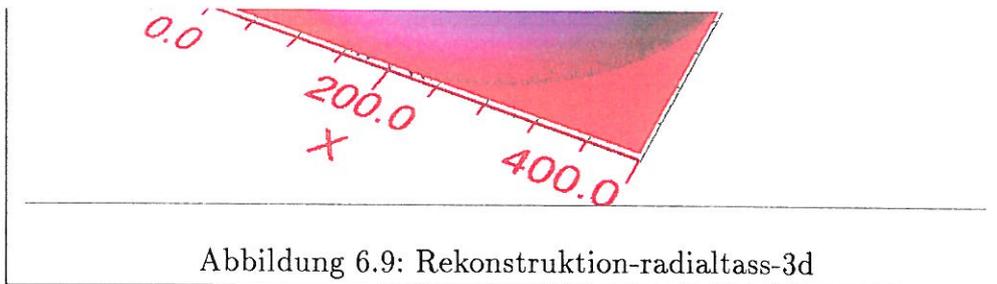
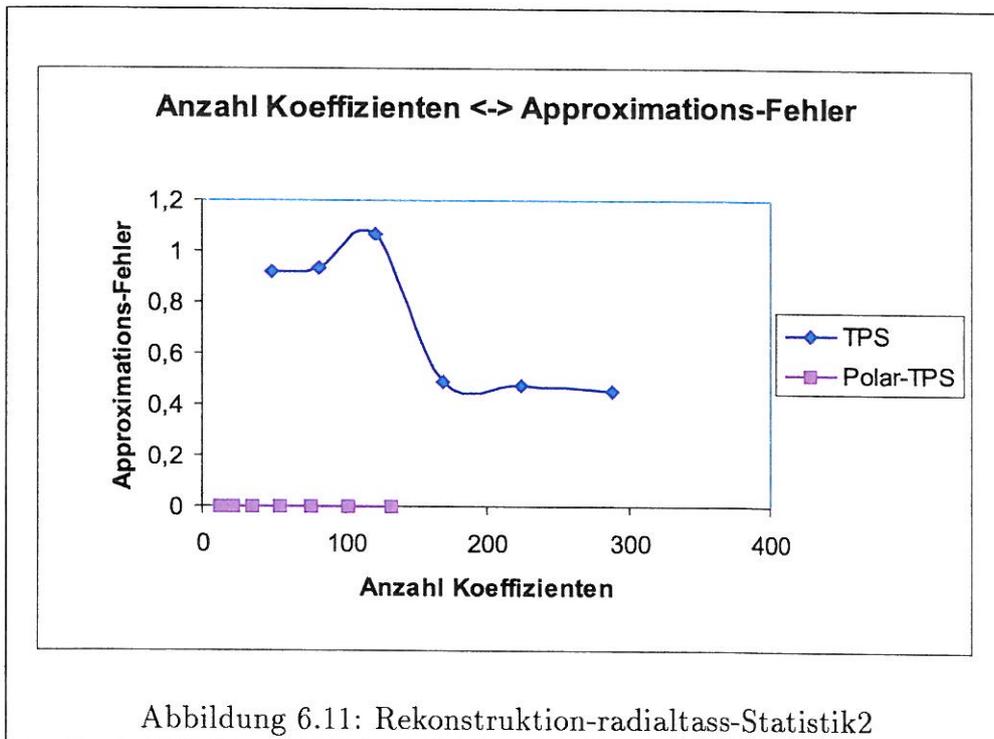
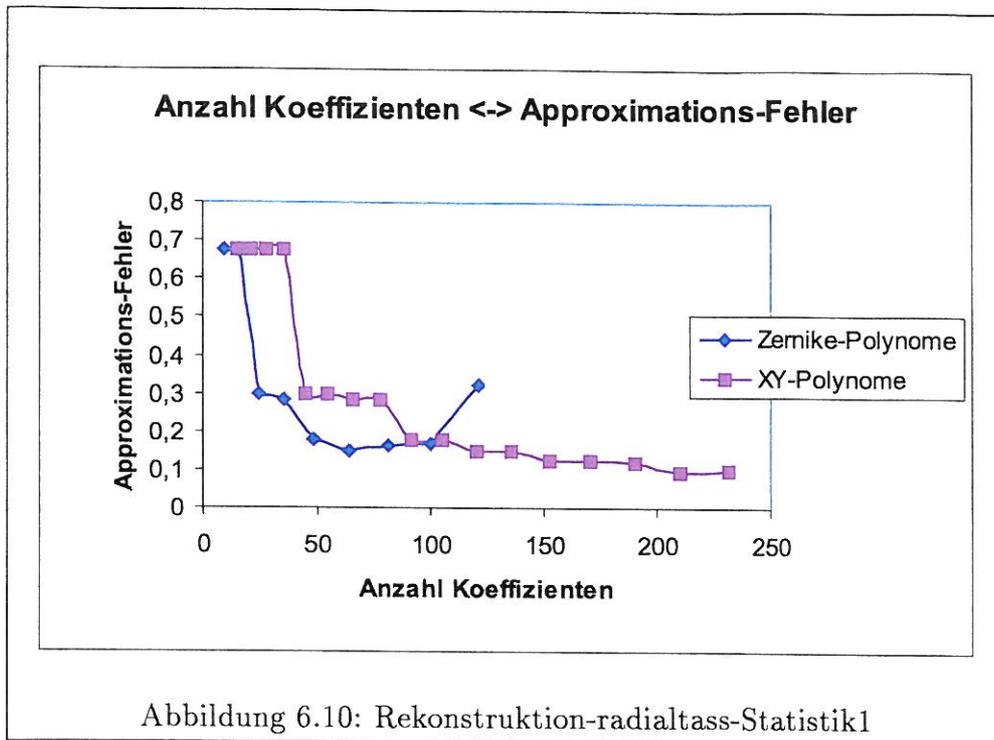
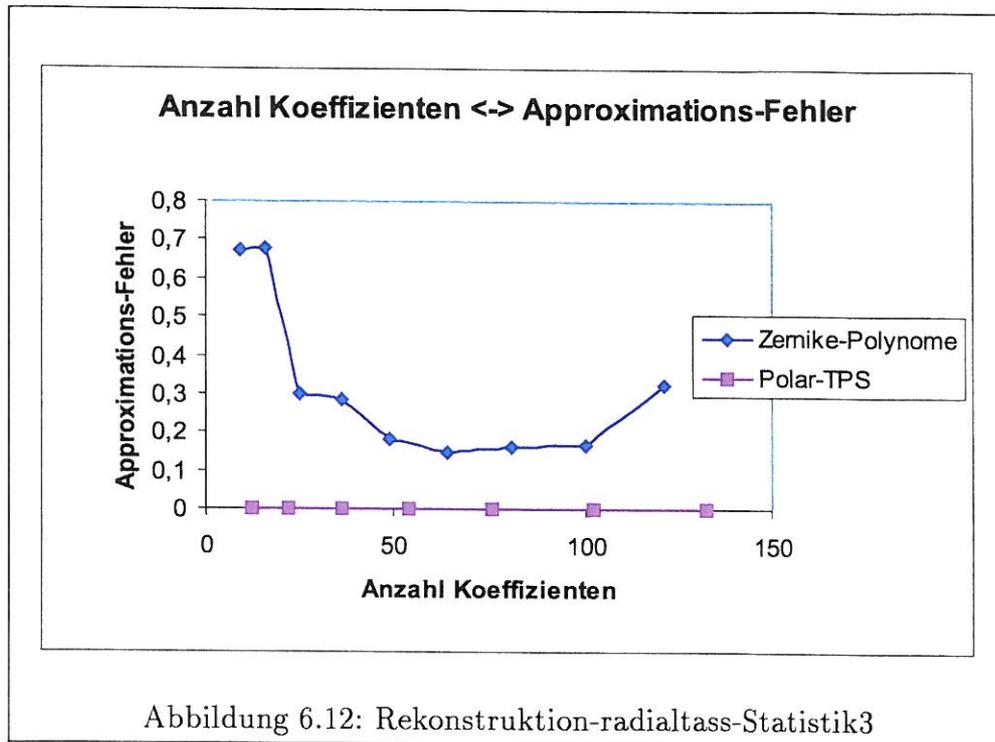


Abbildung 6.9: Rekonstruktion-radialtass-3d





Es konnte gezeigt werden, daß durch Einbeziehung von ausgewähltem Nebenwissen über die eintreffende Wellenfront das Rekonstruktionsergebnis gezielt verbessert werden kann. Die in dieser Arbeit besonders untersuchten Tensorprodukt-Splines bzw. Polarkoordinaten-TPS erwiesen sich als besonders geeignet für die Rekonstruktion von Wellenfronten mit Knickstellen.

Kapitel 7

Zusammenfassung

Kapitel 8

Ergebnisse und Ausblick

Die vorgestellten Verfahren ergeben insgesamt ein ermutigendes Resultat. Durch verschiedene Vorverarbeitungsstufen wurden mögliche Kreuzungspunkte des in den Bildern sichtbaren Visiorasters bestimmt. Das eingesetzte mehrstufige Parametrisierungsverfahren leistet zusätzlich die gewünschte Selektion korrekt erkannter Kreuzungspunkte. Die abschließende Tensorprodukt-Spline-Passung berechnet nicht in der Vorverarbeitung erkannte Kreuzungspunkte und gibt Aufschluß über Parametrisierungsfehler und nicht korrekt erkannte Kreuzungspunkte. Auf Kritikpunkte und Verbesserungsmöglichkeiten der eingesetzten Verfahren wurde in den einzelnen Kapiteln bereits eingegangen.

Das Problem der zuverlässigen automatischen Erkennung der Kreuzungspunkte eines deformierten Visiorasters erscheint vor diesem Hintergrund als lösbar. Es wurde auch deutlich, daß dieses Problem sicherlich einen gewissen Aufwand hinsichtlich der Algorithmik erfordert, der aber auch nicht zu hoch ist. Da es sich nicht um ein in Echtzeit zu lösendes Problem handelt, sind auch durchaus komplexere Algorithmen zur Segmentierung und Parametrisierung in Betracht zu ziehen. Für eine zuverlässige und schnelle Segmentierung von Vorteil wäre eine gezielte Bereichsselektion des in einem Bild sichtbaren Werkstücks durch den Benutzer. Weitere Arbeit sollte auch noch in die Erhöhung der Genauigkeit der in der Vorverarbeitung ermittelten Kreuzungspunkte investiert werden. Das vordringlichste Ziel wäre aber die gezielte Verbesserung der Parametrisierung der Kreuzungspunkte. Dabei wäre eine iterative Verbesserung der Parametrisierung durch Analyse des Ergebnisses der Tensorprodukt-Spline-Passung und anschließender erneuter Passung denkbar.

Anhang A

Laufzeiten der implementierten Algorithmen

Die folgende Tabelle soll einen Eindruck von den Laufzeiten der vorgestellten Algorithmen vermitteln. Sie wurden auf einer Sun Ultra Sparc Creator-1 unter SunOS 5.6 ermittelt. Da die Algorithmen in Bezug auf ihre Laufzeit kaum optimiert sind und das System während des Tests auch noch von anderen Benutzern genutzt wurde, sind diese Meßwerte eher als grobe Anhaltspunkte zu verstehen. Die Implementierungssprache ist C++. Als Eingabe dienen die Bilder aus Abbildung ?? bis ?? auf Seite ??ff. Die Angaben haben das Format „Minuten:Sekunden“.

	219-03	219-04	219-05	219-06
Hochpaßfilterung	0:01	0:01	0:01	0:02
Maskengenerierung	0:40	0:42	0:44	0:42
Geradenpassung	0:07	0:07	0:07	0:06
Skelettierung	0:02	0:02	0:03	0:02
Polygonzugapproximation	10:17	11:19	9:54	9:37
Kreuzungspunktsuche incl. Parametrisierung	6:57	5:30	10:48	9:13
Iterative Passung	4:25	6:47	5:47	8:15
Konsistenz-Test	1:31	2:46	1:40	2:16
Tensorprodukt-Spline-Passung	0:19	0:11	0:17	0:15
Laufzeit insgesamt	24:19	27:25	29:21	30:28

ANHANG A. LAUFZEITEN DER IMPLEMENTIERTEN ALGORITHMEN 97

	219-07	219-08	219-09	219-10
Hochpaßfilterung	0:01	0:01	0:01	0:01
Maskengenerierung	0:40	0:41	0:40	0:41
Geradenpassung	0:06	0:07	0:07	0:07
Skelettierung	0:02	0:02	0:03	0:03
Polygonzugapproximation	9:00	8:44	8:46	9:05
Kreuzungspunktsuche incl.				
Parametrisierung	9:39	10:24	14:38	10:07
Iterative Passung	11:39	11:57	14:11	7:43
Konsistenz-Test	2:06	2:24	3:00	1:39
Tensorprodukt-Spline-Passung	0:09	0:12	0:15	0:14
Laufzeit insgesamt	33:22	34:32	41:41	29:41

Literaturverzeichnis

- [ÅB96] Åke and Bjørk. *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia, 1996.
- [Can95] Robert C. Cannon. Global wavefront reconstruction using shack-hartmann sensors. *Journal of the Optical Society of America A*, 12(9):2031–2039, 1995.
- [COT00] Gary A. Chanan, Catherine Ohara, and Mitchell Troy. Phasing the primary mirror segments of the keck telescopes ii: the narrow-band phasing algorithm. *Applied Optics-OT*, 39(25):4706–4714, Sept. 2000.
- [CTD⁺98] Gary A. Chanan, Mitchell Troy, Frank G. Dekens, Scott Michaels, Jerry Nelson, Terry Mast, and David Kirkman. Phasing the primary mirror segments of the keck telescopes: the broadband phasing algorithm. *Applied Optics-OT*, 37(1):140–155, Jan. 1998.
- [CTO00] Gary A. Chanan, Mitchell Troy, and Catherine Ohara. Phasing the primary mirror segments of the keck telescopes: a comparison of different techniques. *Proc. SPIE - Optical Design, Materials, Fabrication, and Maintenance*, 4003:188–202, Jul. 2000.
- [CTS99] Gary A. Chanan, Mitchell Troy, and Edwin Sirko. Phase discontinuity sensing: a method for phasing segmented mirrors in the infrared. *Applied Optics-OT*, 38(4):704–713, Feb. 1999.
- [dB78] Carl de Boor. *A Practical Guide to Splines*. Springer Verlag, 1978.
- [dB90] Carl de Boor. *Splinefunktionen*. Birkhäuser Verlag, 1990.
- [Die93] Paul Dierckx. *Curve and Surface Fitting with Splines*. Clarendon Press, Oxford, 1993.
- [Don95] Prof. Dr. Klaus Donner. Vorlesung Numerik I, 1995.

- [Gea95] J. M. Geary. *Introduction to Wavefront Sensors, Tutorial Texts in Optical Engineering*, volume 18. SPIE - The International Society for Optical Engineering, 1995.
- [GL89] Gene H. Golub and Charles F. van Loan. *Matrix Computations*. Johns Hopkins University Press, 1989.
- [Goo96] Joseph W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, New York, second edition, 1996.
- [GR81] I.S. Gradstein and I.M. Ryshik. *Tables of Series, Products and Integrals*. Harri Deutsch Verlag, 1981.
- [Gra92] Prof. Dr. Graf. Vorlesung Digitale Bildverarbeitung I, 1992. Skriptum von Gernot Schreib.
- [Gre87] John E. Greivenkamp. Sub-nyquist interferometry. *Applied Optics*, 26(24):5245–5258, 1987.
- [Gro98] Sascha Groening. Tensorprodukt-Spline-Approximation deformierter Raster in der Visoplastizität. Diplomarbeit am Lehrstuhl für numerische Mathematik Prof. Dr. Donner, Universität Passau, 1998.
- [GS99] S. Groening and B. Sick. Wave-front reconstruction with a shackhartmann sensor using an iterative spline fitting method. In *9th Int. Trade Fair and Conference for Sensors, Transducers and Systems, Proceedings 2*, pages 457–462, Mai 1999.
- [GSD⁺00] S. Groening, B. Sick, K. Donner, J. Pfund, N. Lindlein, and J. Schwider. Wave-front reconstruction with a shackhartmann sensor with an iterative spline fitting method. *Applied Optics*, 39(4):561–567, 2000.
- [Hec99] Eugene Hecht. *Optik*. Oldenbourg Verlag, München, 1999.
- [HLK77] J. W. Hardy, J. E. Lefebvre, and C. L. Koliopoulos. Real-time atmospheric compensation. *Journal of the Optical Society of America*, 67(3):360–369, 1977.
- [Hud77] Richard H. Hudgin. Wave-front reconstruction for compensated imaging. *Journal of the Optical Society of America*, 67(3):375–378, 1977.

- [Inc93] Hughes Danbury Optical Systems Inc. Wavefront sensor having a lenslet array as a null corrector. Patent US 5 233 174, 1993. Offenlegungstag 3.8.1993.
- [Lin99] N. Lindlein. Anmerkungen zu dynamik-bereich und genauigkeit eines shack-hartmann sensors. Lehrstuhl für Optik, Universität Erlangen-Nürnberg - Interner Bericht, ??? 1999.
- [LPS99] N. Lindlein, J. Pfund, and J. Schwider. Wellenfrontsensor für den optical workshop. *Laser*, 5:28–31, 1999.
- [LPS00] N. Lindlein, J. Pfund, and J. Schwider. Expansion of the dynamic range of a shack-hartmann sensor by using astigmatic microlenses. *Optical Engineering*, 39(8):2220–2225, 2000.
- [Ltd98] Visionix Ltd. Apparatus for mapping optical elements. Patent US 5 825 476, 1998. Offenlegungstag 20.10.1998.
- [Man92] Paul K. Manhart. Full surface phasing of a segmented mirror. *Applied Optics*, 31(27):5744–5745, 1992.
- [MBN00] Justin D. Mansell, Robert L. Byer, and Daniel R. Neal. Apodized micro-lenses for hartmann wavefront sensing. *World Scientific*, ?(?):0000–0000, 2000.
- [MM92] Daniel Malacara and Zacarias Malacara. ??? *Optical Engineering*, 31(7):1551–0, 1992.
- [Opt00] Imagine Optic. Method and device for analyzing a highly dynamic wavefront. Patent WO 00/42401, 2000. Offenlegungstag 14.1.2000.
- [PLS98] J. Pfund, N. Lindlein, and J. Schwider. Dynamic range expansion of a shack-hartmann sensor by using a modified unwrapping algorithm. *Optical Letters*, 23(13):995–997, 1998.
- [PLS99] J. Pfund, N. Lindlein, and J. Schwider. Streifende inzidenz. Lehrstuhl für Optik, Universität Erlangen-Nürnberg - Interner Bericht, Mai 1999.
- [PLS+00] J. Pfund, Dr. N. Lindlein, Prof. Dr. J. Schwider, S. Groening, Prof. Dr. K. Donner, Dr. A. Bode, F. Gruna, Dr. K.-E. Elßner, Dr. W. Holota, and K. Eckerl. TP: 3 Entwicklung eines Messverfahrens mit hohem Dynamikbereich für die Qualitätsprüfung von

- optischen Asphären und für die in-situ Messung von Wellenfronten (kurz: Wellensensor). Technical report, Lehrstuhl für Optik (LfO) / Universität Erlangen-Nürnberg, 2000.
- [RH93] G. Roblin and D. Horville. Étude de l'aberration propre a une trame de microlentilles. *J. Optics(Paris)*, 24:77–87, 1993.
- [RS98] M. C. Roggemann and T. J. Schulz. Algorithm to increase the largest aberration that can be reconstructed from hartmann sensor measurements. *Applied Optics*, 37(20):4321–4329, 1998.
- [RSNK99] Michael C. Roggemann, Timothy J. Schulz, Chee W. Ngai, and Jason T. Kraft. Joint processing of hartmann sensor and conventional image measurements to estimate large aberrations: theory and experimental results. *Applied Optics*, 38(11):2249–2255, 1999.
- [RSW94] Michael C. Roggemann, Craig A. Stoudt, and Byron M. Welsh. Image-spectrum signal-to-noise-ratio improvements by statistical frame selection for adaptive-optics imaging through atmospheric turbulence. *Optical Engineering*, 33(10):3254–3264, 1994.
- [RWG⁺95] Michael C. Roggemann, Byron M. Welsh, Patrick J. Gardner, Robert L. Johnson, and Byron L. Pedersen. Sensing three-dimensional index-of-refraction variations by means of optical wavefront sensor measurements and tomographic reconstruction. *Optical Engineering*, 34(5):1374–1384, 1995.
- [Sch98] Prof. Dr. Johannes Schwider. Verfahren zur Erhöhung der Meßdynamik des Shack-Hartmann Sensors. Patent DE 197 05 119 A 1, 1998. Offenlegungstag 13.8.1998.
- [Sch99] Prof. Dr. Johannes Schwider. Shack-Hartmann Sensor mit ortsvarianter Linsenordnung. Patent DE 198 00 844.9, 1999. Offenlegungstag 15.7.1999.
- [SL96] L. E. Schmutz and B. M. Levine. Hartmann sensors detect optical fabrication errors. *Laser Focus World*, 4:111, 1996.
- [Sou80] W. H. Southwell. Wave front estimation from wave front slope measurements. *Journal of the Optical Society of America*, 70:998–1006, 1980.

- [WBJ93] A. Wirth, T. L. Bruno, and A.J. Jankevics. Applying hartmann wavefront-sensing technology to precision optical testing of the hubble space telescope correctors. *Proc. SPIE - Active and Adaptive Optical Components and Systems II*, 1920:328–336, 1993.
- [ZH97] Fa. Carl Zeiss Heidenheim. Vorrichtung und Verfahren zur orts-aufgelösten Vermessung von Wellenfrontdeformationen. Patent DE 196 24 421 A 1, 1997. Offenlegungstag 2.1.1997.
- [Zim97] Alexander Zimmermann. *Parametrisierungs- und Lokalisierungstechniken in der Spline-Oberflächenapproximation*. Berichte aus der Informatik. Shaker Verlag, Aachen, 1997.