

ADINA GUIAŞ

---

**Eine analytische Methode  
zur Punktereduktion und  
Flächenrekonstruktion**

DISSERTATION

---

UNIVERSITÄT DORTMUND

**2005**



# Eine analytische Methode zur Punktereduktion und Flächenrekonstruktion

DISSERTATION

zur Erlangung des Grades  
eines Doktors der Naturwissenschaften  
der Universität Dortmund

Dem Fachbereich Mathematik  
der Universität Dortmund  
vorgelegt von:

Adina Guiaş

2005

Tag der mündlichen Prüfung: 4. Juli 2005

Vorsitzender:	Prof. Dr. Wolfgang Achtziger
Erster Gutachter:	Prof. Dr. Herbert Koch
Zweiter Gutachter:	Prof. Dr. Joachim Stöckler
Prüfer:	Prof. Dr. Winfried Kaballo
Prüfer:	PD Dr. Jörn Mehnen

# Vorwort

Diese Dissertation entstand im Rahmen der *Graduate School for Production Engineering and Logistics* an der Universität Dortmund, unterstützt durch ein dreijähriges Promotionsstipendium, das ich ab November 2001 auf Empfehlung von Prof.Dr.Dr. h.c.mult. Willi Jäger (Heidelberg) und Prof.Dr. Herbert Koch (Dortmund) erhielt. Sie befasst sich mit der Anwendung mathematischer Methoden auf eine praktische Fragestellung, welche relevant für die Ingenieurwissenschaften ist. Die hier benutzte mathematische Theorie der Wavelets und  $BMO^{-1}$ -Funktionen ermöglicht die Erkennung von Strukturen auf einer gegebenen Fläche.

Mein Dank geht vor allem an Herrn Prof.Dr. Herbert Koch für die Chance zu promovieren und für die permanente und unermüdliche Unterstützung. Seine Anregungen hatten entscheidenden Einfluss auf meine Arbeit. Ebenfalls danke ich Herrn Prof.Dr. Joachim Stöckler für seine fachliche Unterstützung bei ihrer Fertigstellung. Geprägt hat diese Arbeit auch die enge Zusammenarbeit mit Herrn P.D.Dr.-Ing. Jörn Mehnen vom Institut für Spanende Fertigung der Universität Dortmund, der mir neben der Lieferung der Daten immer mit praktischen Ratschlägen und permanenter Aufmunterung in schwierigen Situationen zur Seite stand und dem ich dafür recht herzlich danke. Darüber hinaus geht mein Dank an die *Graduate School for Production Engineering and Logistics* für das Stipendium. Für die Unterstützung bei der Korrektur dieser Arbeit danke ich ganz herzlich Frau Dr.-Ing. Martina Köster.

Diese Arbeit widme ich meinem Vater, Marcel Cazacu, der mich über jedwelche erdenkliche Grenze hinaus geliebt hat und der zum Beginn dieser Arbeit verstorben ist... als auch meiner Mutter Doreti, die mir von Kind an die Liebe zur Mathematik mitgegeben hat und mir stets mit viel Liebe und Geduld beistand.

Adina Guiaş

Dortmund, im Mai 2005

Nu ești înfrânt atunci când sângeri  
Și nici când ochii'n lacrimi ți-s.  
Adevăratele înfrângeri  
Sunt renunțările la vis!

(R.Gyr)



# Inhaltsverzeichnis

<b>1</b>	<b>Motivation und Problemstellung</b>	<b>1</b>
<b>2</b>	<b>Wavelets, <math>BMO</math>- und <math>BMO^{-1}</math>-Funktionen</b>	<b>7</b>
2.1	Wavelets . . . . .	7
2.1.1	Eindimensionale Wavelets . . . . .	8
2.1.2	Die schnelle Wavelet-Transformation . . . . .	14
2.1.3	Wavelets in mehreren Dimensionen . . . . .	17
2.1.4	Wavelet-Analyse diskreter Signale . . . . .	18
2.2	Die Räume $BMO(\mathbb{R}^n)$ und $BMO^{-1}(\mathbb{R}^n)$ . . . . .	21
2.2.1	Der Raum $BMO(\mathbb{R}^n)$ . . . . .	21
2.2.2	Dualität zwischen $BMO$ und $H^1$ . . . . .	22
2.2.3	$H^1(\mathbb{R}^n)$ , $BMO(\mathbb{R}^n)$ und Wavelets . . . . .	24
2.2.4	Der Raum $BMO^{-1}(\mathbb{R}^n)$ . . . . .	29
2.2.5	$BMO$ und $BMO^{-1}$ als homogene Triebel-Lizorkin Räume . . . . .	30
2.2.6	Anwendungen des Raumes $BMO^{-1}$ . . . . .	37
<b>3</b>	<b>Algorithmus zur Punktereduktion und Flächenrekon- struktion</b>	<b>43</b>
3.1	Berechnung der $BMO^{-1}$ -Koeffizienten . . . . .	44
3.1.1	Lineare Interpolation einer gegebenen Punk- temenge . . . . .	45
3.1.2	Waveletzerlegung und Berechnung der $BMO^{-1}$ -Koeffizienten . . . . .	49
3.2	Auswahl der Punkte . . . . .	50
3.3	Rekonstruktion der Fläche . . . . .	51
3.3.1	Glättende Splines für eine gegebene Knoten- menge . . . . .	52

---

3.3.2	Eine adaptive Strategie zur Platzierung der Knoten . . . . .	56
3.3.3	Glättung durch Vorgabe des Glättungsparameters . . . . .	58
3.4	Auswahl der Parameter . . . . .	59
3.5	Praktische Aspekte . . . . .	60
<b>4</b>	<b>Anwendungen</b>	<b>65</b>
4.1	Rekonstruktion eines Testobjekts . . . . .	66
4.2	Rekonstruktion des Motorkolbens . . . . .	69
4.3	Rekonstruktion des Doms . . . . .	71
4.4	Exportieren der Daten im <i>IGES</i> -Format . . . . .	74
<b>5</b>	<b>Zusammenfassung und Ausblick</b>	<b>77</b>

# Kapitel 1

## Motivation und Problemstellung

In dieser Arbeit wird das Problem der Flächenrekonstruktion analytisch betrachtet. Im Gegensatz zu ingenieurwissenschaftlichen Lösungswegen wird in dieser Arbeit ein mathematischer Ansatz verfolgt, um die aktuellen Schwierigkeiten, die in der Praxis entstehen, zu lösen.

Der Flächenrekonstruktion kommt im Maschinenbau eine große Bedeutung zu. Während der Entwicklungsphase eines Produktes wird oft mit Prototypen und konkreten Modellen gearbeitet, deren Oberflächenform sich im Verlauf dieses Prozesses ständig ändert. Da es sich meistens um Unikate handelt, ist es oft wünschenswert, die Zwischenstufen dieser Entwicklung in rechneradäquater Form zu speichern, um gegebenenfalls auf ein vorheriges Profil zurückgreifen zu können. Andere mögliche Anwendungen sind unter anderem das Duplizieren von Werkstücken oder die Kontrolle der Abweichungen eines gegebenen Werkstückes von den vorgeschriebenen Parametern.

Die Abbildung einer physischen Flächenform in eine CAD-kompatible Darstellung basierend auf diskreten Punktmengen wird als *Flächenrekonstruktion (reverse engineering)* bezeichnet, [5] S.3. Im Maschinenbau werden die Werkstücke meistens in Form von CAD-Modellen konzipiert, anhand derer die Fräsmaschinen programmiert werden können, um das entsprechende Teil zu realisieren. *Reverse engineering* bezeichnet genau den umgekehrten Prozess: Ausgehend von einem Werkstück wird eine passende CAD-Darstellung gesucht, wel-

---

che einerseits kompakt sein soll, andererseits jedoch die reale Fläche genau und detailgetreu nachbilden soll.

Ausgangspunkt der Flächenrekonstruktion ist eine Menge von gemessenen Punkten auf der Oberfläche des Objektes, deren Koordinaten durch Bewegung eines kugelförmigen Rubinsensors auf äquidistanten Niveaulinien berechnet werden. Die Koordinaten der Punkte werden durch dieses Verfahren sehr genau erfasst. In der Praxis beträgt der Abstand zwischen den gemessenen Linien Bruchteile von Millimetern. Ein Beispiel einer solchen Punktemenge ist in Abbildung 1.1 dargestellt.

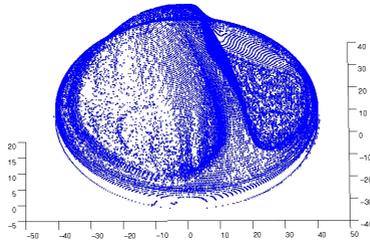


Abbildung 1.1: Digitalisierung eines Motorkolbens

Im Hinblick auf die erwähnten Anwendungen im technischen Bereich (CAD-Modelle) ist das Ziel eine Darstellung der Fläche durch Tensorprodukt-Splines, basierend auf möglichst wenig Knoten, die aber eine möglichst genaue Approximation der gemessenen Punktemenge realisiert. Diese zwei Forderungen stehen im Widerspruch zueinander. Der Begriff “gute Approximation” ist in dieser Situation problematisch. Zum Einen gibt es keine universellen Qualitätsmerkmale der Rekonstruktion. Die Bewertung von Abweichungen, Rauheit und Details, die reproduziert werden müssen, hängt sehr stark vom Problem ab. Das Verfahren kann z.B. auf die Rekonstruktion von Werkstücken, die Analyse von Oberflächen, auf digitale Speicherung von Flächen, usw. angewendet werden. Es wird deutlich, dass viele Anforderungen existieren, daher kann es kein einheitlicher Begriff “gute Approximation” geben. Eine zu hohe Genauigkeit in der Annäherung an die einzelnen Punkte kann die analytische Darstellung zu sehr starken Oszillationen zwischen den Punkten zwingen. Diese können durch die Anforderung an die Glattheit ausbalanciert

werden. Eine zu starke Glättung würde jedoch auf Kosten mancher Details des Objektes erfolgen, die durch die dargestellte Fläche nicht mehr richtig reproduziert werden würden. Die Güte der Approximation ist also sehr stark kontextabhängig. Da die Anzahl der gemessenen Punkte in der Praxis sehr groß ist, zwischen 30.000-200.000, ist es zunächst erforderlich, eine geeignete Auswahl der Punkte zu treffen, um die *modellspezifische Redundanz* zu minimieren. Diese wird in [5], S.34, definiert als “Zahl der Mengen der Punkte, die durch Approximation mit einem oder einer Menge vorgegebener Flächenmodelle keinen wesentlichen Beitrag zur eindeutigen Festlegung der Flächenstruktur leisten”. Mit anderen Worten, es sollen nur diejenigen Punkte ausgewählt werden, die relevante Informationen über die Struktur der Fläche liefern können. Die bisher angewandten Methoden, z.B. das *TDV-Verfahren* (tangenciales Distanzschrittverfahren), sind jedoch eher intuitiv als analytisch begründet.

Das Ziel dieser Arbeit ist es also, eine flexible und zuverlässige Methode für die beschriebene Problematik zu entwickeln, welche für verschiedene Anforderungen einsetzbar ist. Die wichtigsten analytischen Werkzeuge dieses Verfahrens sind die *Wavelets* und der *Raum der  $BMO^{-1}$ -Funktionen*.

Wavelets sind Familien von Funktionen, die durch Dilatationen und Translationen einer sogenannten *Mutter-Wavelet* entstehen. Sie ermöglichen die Zerlegung eines Signals in elementare Bausteine, die sowohl in der Zeit als auch in der Frequenz definiert sind. Diese Eigenschaft, zusammen mit der Orthogonalität der Waveletfunktionen führt dazu, dass Wavelets auch in der harmonischen Analysis eine bedeutende Rolle spielen. Man kann verschiedene Funktionenräume, wie z.B. die Klassen der *Triebel-Lizorkin Räume* und der *Besov-Räume* (welche die meisten üblichen Räume wie Sobolev-Räume, Hardy-Räume und *BMO* enthalten), mit Hilfe abzählbar vieler Waveletkoeffizienten  $c_\lambda = \langle f, \psi_\lambda \rangle$  charakterisieren. Speziell wird hier für den Raum  $BMO^{-1}$  (bestehend aus verallgemeinerten Divergenzen von *BMO*-Vektorfeldern) eine solche Charakterisierung hergeleitet.

Der Raum  $BMO^{-1}$  wird in einigen Anwendungen der Analysis und in der Bildverarbeitung dazu benutzt, um Texturen oder schnell oszillierende Muster zu modellieren, vgl. [3]. Man ist bestrebt, ein Signal (Bild)  $f$  als eine Summe  $u + v$  darzustellen, wobei  $u$  die Objekte (Strukturen) aus dem Bild modellieren soll, während  $v$  für die Texturen und das Rauschen steht. Dies wird realisiert, indem die Norm  $\|u\|_{\mathcal{B}}$  in einem geeigneten Banachraum  $\mathcal{B}$  beschränkt ist, während eine andere Norm  $\|v\|_*$  “klein” sein soll, um den stark oszillierenden

---

Charakter der  $v$ -Komponente auszudrücken.

Im Unterschied zur Bildverarbeitung sind die Funktionen (Flächen), die bei der Flächenrekonstruktion auftreten, in der Regel stetig, oft glatt. Der wichtigste Unterschied zwischen der hier präsentierten Anwendung und der Problematik der Bildverarbeitung ist jedoch die Forderung nach einer viel höheren Genauigkeit in der Approximation der Funktion. Wenn wir eine Fläche als Graph einer Funktion betrachten, enthält der sogenannte  $BMO^{-1}$ -Koeffizient im Punkt  $x$  Informationen über die lokale Struktur der Fläche in der Umgebung des Punktes  $(x, f(x))$ . Dieser wird erhalten durch eine gewichtete Mittelung der Waveletkoeffizienten  $c_\lambda = \langle f, \psi_\lambda \rangle$  entsprechend der Waveletfunktionen  $\psi_\lambda$ , die um diesen Punkt lokalisiert sind. Große  $BMO^{-1}$ -Koeffizienten weisen auf "Strukturen" auf der Oberfläche hin, während kleine Werte auf flache Stellen oder auf schnell oszillierende Werte der Funktion (welche z.B. auf Rauheiten der Oberfläche zurückzuführen sind) hinweisen. Diese Koeffizienten liefern Informationen über die ganze Oberfläche, auch dort, wo die Waveletkoeffizienten klein sind. Die Größe des  $BMO^{-1}$ -Koeffizienten in einem Punkt wird also eine bedeutende Rolle spielen bei der Entscheidung, ob dieser Punkt gewählt wird oder nicht.

Im ersten Teil der Arbeit werden alle theoretischen Begriffe erläutert, die zur Herleitung der Formel für die Berechnung der  $BMO^{-1}$ -Koeffizienten notwendig sind: Wavelets, Hardy-Räume,  $BMO$  (der in die Definition von  $BMO^{-1}$  eingeht) und schließlich Triebel-Lizorkin Räume, aus deren Little-wood-Paley Charakterisierung in [15] hier eine Charakterisierung mit Hilfe der Wavelets hergeleitet wird. Der Raum  $BMO^{-1}$  wird mit einem Raum dieser Familie identifiziert und somit eine Waveletcharakterisierung dieses Raumes entwickelt.

Im zweiten Teil der Arbeit wird die eigentliche Methode zur Punktereduktion und Flächenrekonstruktion dargestellt. Ausgehend von einer großen Punktemenge, gemessen auf einer (unbekannten) Oberfläche, erfolgt eine analytische Darstellung der Fläche, die auf möglichst wenigen Parametern (Knoten, Koeffizienten) basiert. Zunächst wird eine Interpolation der Fläche in den Punkten eines regulären Gitters ausgewertet. Somit kann man eine diskrete Waveletzerlegung durchführen und die  $BMO^{-1}$ -Koeffizienten in jedem Gitterpunkt berechnen. Dies hat zum Ziel, die Fläche gemäß der Größenordnung der  $BMO^{-1}$ -Koeffizienten in Bereiche aufzuteilen, in denen dann die gegebenen Punkte mehr oder weniger ausgedünnt werden. Durch Änderung gewisser threshold-Parameter können die entstandenen Bereiche geändert werden, um somit die Punktedichten stärker an die Struk-

turen der Fläche anzupassen. Zu den ausgewählten Punkten werden zusätzlich auch die entsprechenden  $BMO^{-1}$ -Koeffizienten übertragen, da sie im späteren Verlauf, bei der Berechnung der Gewichte der Spline-Rekonstruktion, mit einbezogen werden.

Als nächstes ist es erforderlich, anhand der reduzierten Punkte menge eine Spline-Fläche zu bestimmen, welche den ursprünglichen Punktesatz so gut wie möglich approximiert und auch gewisse Anforderungen an die Glattheit erfüllt. In der Approximationstheorie wird dieses Problem als *surface fitting* (Flächenanpassung) bezeichnet. Die wie oben beschrieben ausgewählten Punkte (eine Teilmenge der ursprünglich gemessenen Punkte) bilden eine unstrukturierte Punktmenge in  $\mathbb{R}^3$ . In diesem Fall ist die Flächenanpassung mit Hilfe von Spline-Funktionen ein nichttriviales Problem der Approximationstheorie und wird u.a. in [7], [8] und [9] behandelt. Im Unterschied zum eindimensionalen Fall oder zu der Situation, in welcher den gegebenen Punkten ein Gitter zugrunde liegt, gibt es hier kein Analogon der Schoenberg-Whitney Bedingung, die die eindeutige Lösbarkeit des least-squares Problems gewährleistet. Das System, welches die Koeffizienten einer approximierenden least-squares Spline-Funktion liefern soll, ist in dieser Situation oft rangdefizient. Auch wenn man eine spezielle Lösung auswählt (z.B. mit minimaler Norm), hat die rekonstruierte Fläche einen 'rauen' Aspekt und weist oft starke Oszillationen zwischen den Punkten auf.

Unter diesen Umständen ist eine Glättung erforderlich, die die erwähnten Probleme der Nichteindeutigkeit und Rauheit behebt. Das Ziel ist es, unter vorgegebenen Genauigkeits- und Glattheitsanforderungen eine Darstellung der Fläche mit Hilfe von Tensorprodukt-Splines mit möglichst wenig Knoten zu erreichen. Das Problem wird zunächst für eine gegebene Knotenmenge als ein Minimierungsproblem formuliert und gelöst. Es wird entweder eine geeignete Glattheitsnorm minimiert mit der Forderung, dass der quadratische Fehler unter einer vorgegebenen Schranke liegt, oder eine lineare Kombination aus dem quadratischen Fehler und der Glattheitsnorm verwendet.

Um eine passende Knotenmenge zu finden, wird eine adaptive Strategie verwendet. Die Knoten werden sukzessive gewählt mit dem Ziel, den quadratischen Fehler zu minimieren. Die Gewichte, welche in die Berechnung dieses Fehlers eingehen, kann man in Abhängigkeit von den  $BMO^{-1}$ -Koeffizienten wählen, und zwar invers proportional. Dies gewährleistet bereits bei der least-squares Approximation eine lokale Glättung an den problematischen Stellen der Fläche, ohne die Details (Kanten) zu zerstören.

---

Schließlich werden einige Anwendungen vorgestellt. Zunächst wird das Verfahren anhand eines künstlich generierten Objekts getestet. Es handelt sich um einen Kegelstumpf, auf dem eine Halbkugel liegt. Die Koordinaten dieses Testobjektes wurden analytisch berechnet, und die Schritte des Algorithmus sind einzeln dargestellt: die Punktere-  
duktion mit analytischer Begründung und die Splinerekonstruktion mit adaptiver Knotenplatzierung. Anhand dieses Testobjektes erfolgt dann auch eine Diskussion der Glättung und deren Konsequenzen (z.B. die Existenz eventueller Texturen). Anschließend werden einige praktische Beispiele berechnet. Beim ersten Beispiel handelt es sich um die Rekonstruktion eines Motorkolbens, dessen Koordinaten in Abbildung 1.1 dieses Kapitels dargestellt wurden. Dieses Beispiel ist interessant für unser Vorhaben, da es sowohl sehr glatte Bereiche aufweist als auch den 'Kamm', der nicht auf einer geraden Linie verläuft und eher einer schiefen Kante entspricht, und nicht zuletzt an steile Kanten angrenzende raue Bereiche. Als ein weiteres Beispiel wird der 'Dom' aus Abbildung 4.10 (S. 68) präsentiert. Es handelt sich um ein wesentlich größeres Objekt( ca. 200.000 ursprünglich gemessene Punkte, im Vergleich zu ca. 35.000 ursprünglich gemessenen Punkten beim ersten praktischen Beispiel). Der Dom hat sehr steile Bereiche und scharfe Kanten und Rillen.

Der hier vorgestellte Algorithmus reagiert flexibel auf alle Anforderungen und ist auf alle Punktedaten des beschriebenen Typs anwendbar. Durch Steuerung der threshold-Parameter kann man die Einteilung der Fläche in Bereiche beeinflussen, und in jedem dieser Bereiche kann man bei Bedarf die ursprünglich gemessenen Punkte mehr oder weniger ausdünnen, bis man eine optimale Rekonstruktion erhält.

Die erhaltene Spline-Fläche wird schließlich im IGES-Format (*Initial Graphics Exchange Specification*) dargestellt, um mit den gängigen CAD-Programmen kompatibel zu sein. Ein solches Format eignet sich auch für die Weiterleitung dieser Daten an Fräsmaschinen, die dann das entsprechende Objekt herstellen können.

# Kapitel 2

## Wavelets, $BMO$ - und $BMO^{-1}$ -Funktionen

In diesem Kapitel werden die Eigenschaften von Wavelets, des Raumes  $BMO(\mathbb{R}^n)$  und des Raumes  $BMO^{-1}(\mathbb{R}^n)$  vorgestellt sowie die Charakterisierung dieser Räume mit Hilfe von orthogonalen Basen von Wavelets erläutert.

### 2.1 Wavelets

Wavelets sind eines der bedeutendsten mathematischen Hilfsmittel in vielen Bereichen der Signalverarbeitung oder Bildverarbeitung, finden aber oft Anwendungen auch innerhalb der Analysis. Wavelets sind Funktionen mit lokalisiertem Träger und oszillatorischem Charakter. Sie werden durch Dilatationen und Translationen einer sogenannten *Mutter-Wavelet* erhalten. Durch deren Charakteristiken und Konstruktionsweise ergibt sich eine der wichtigsten Eigenschaften der Wavelets, die man in der Sprache der Signaltheorie wie folgt darstellen kann: sie ermöglichen die Zerlegung eines Signals in Zeit *und* Frequenz, im Unterschied zur Fouriertransformation, die eine Frequenz-Zerlegung erzeugt, welche jedoch keinen lokalen Charakter hat.

Mit Hilfe der Wavelets kann man eine sogenannte *Multiresolutionsanalyse* eines Signals durchführen, d.h. das Signal wird entsprechend der Information die es enthält, in verschiedene Resolutions-

Skalen zerlegt. Einerseits wird gemittelt und diese Mittelung wird in einer Basis mit Hilfe der sogenannten Skalierungsfunktionen dargestellt. Somit wird die Resolution grober, aber der dabei entstandene Fehler wird berücksichtigt, indem er in der Waveletbasis der entsprechenden Skala dargestellt wird. Diese Zerlegung ermöglicht die Ableitung des Informationsgehaltes des Signals auf verschiedenen Niveaus, aber gleichzeitig auch eine vollständige Rekonstruktion des Signals.

Eine andere Anwendung der Multiresolutionsanalyse liegt im Bereich der Signalverarbeitung. Man kann dadurch in einen Abschnitt eines Signals "hineinzoomen", um z.B. die unnötigen Details zu entfernen (entrauschen) oder um eine Kompression zu erreichen, indem man nur noch diejenigen Waveletfunktionen für die Darstellung des Signals berücksichtigt, deren Koeffizienten oberhalb einer Grenze liegen (thresholding).

Als Referenzen aus der umfangreichen Literatur zu diesem Thema seien [10] [11] und [17] erwähnt, aus denen die Elemente der Wavelet-Theorie stammen, die in dieser Arbeit verwendet werden.

Die o.g. Eigenschaften der Wavelets -der lokalisierte (oft kompakte) Träger und der oszillatorische Charakter- finden einen bedeutenden Anwendungsbereich auch in der harmonischen Analysis, z.B. in [2]. Hierauf basiert der theoretische Abschnitt dieser Arbeit. Deswegen werden im folgenden die grundlegenden Begriffe und Konzepte präsentiert, welche zur Charakterisierung der Räume  $BMO$  und  $BMO^{-1}$  mit Hilfe von Wavelets erforderlich bzw. in der praktischen Implementierung der Wavelet-Transformation für unsere Anwendung relevant sind.

## 2.1.1 Eindimensionale Wavelets

Sei  $f \in L^2(\mathbb{R})$  und  $\hat{f}$  die Fouriertransformierte:

$$\mathcal{F}f = \hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)e^{-ix\omega} dx. \quad (2.1)$$

**Definition 1** Eine Funktion  $\psi \in L^2(\mathbb{R})$ , welche die Zulässigkeitsbedingung

$$0 < c_\psi := 2\pi \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (2.2)$$

erfüllt, heißt Wavelet.

Die Wavelet-Transformierte einer Funktion  $f \in L^2(\mathbb{R})$  zum Wavelet  $\psi$  ist durch

$$L_\psi f(a, b) = \frac{1}{\sqrt{c_\psi}} |a|^{-1/2} \int_{\mathbb{R}} f(x) \psi\left(\frac{x-b}{a}\right) dx, \quad (2.3)$$

$a > 0, b \in \mathbb{R}$  definiert.

Die Zulässigkeitsbedingung impliziert eine notwendige Bedingung, damit eine Funktion  $\psi \in L^1(\mathbb{R})$  ein Wavelet ist. Der Satz von Riemann-Lebesgue besagt, dass die Fourier-Transformierte  $\hat{\psi}$  stetig auf  $\mathbb{R}$  ist. Damit (2.2) erfüllt ist, muß also

$$0 = \hat{\psi}(0) = (2\pi)^{-1/2} \int_{\mathbb{R}} \psi(x) dx$$

gelten. Daher ist der Mittelwert einer Waveletfunktion (Moment der Ordnung 0) immer gleich Null. Man kann auch Wavelets konstruieren, für welche die Momente beliebiger Ordnung:  $\int x^m \psi(x) dx$  verschwinden.

**Beispiele:**

(i) Haar-Wavelet

$$\psi(x) = \begin{cases} 1 & : 0 \leq x < 1/2 \\ -1 & : 1/2 \leq x \leq 1 \\ 0 & : \text{sonst} \end{cases}$$

(ii) Mexikanischer Hut

$$\psi(x) = -\frac{d^2}{dx^2} e^{-x^2/2} = (1 - x^2) e^{-x^2/2}.$$

(iii) Meyer-Wavelet

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2\pi}} e^{i\omega/2} (\theta(\omega) + \theta(-\omega))$$

mit

$$\theta(\omega) = \begin{cases} \sin\left(\frac{\pi}{2} \nu\left(\frac{3\omega}{2\pi} - 1\right)\right) & \text{für } \frac{2\pi}{3} \leq \omega \leq \frac{4\pi}{3} \\ \cos\left(\frac{\pi}{2} \nu\left(\frac{3\omega}{4\pi} - 1\right)\right) & \text{für } \frac{4\pi}{3} \leq \omega \leq \frac{8\pi}{3} \\ 0 & \text{sonst} \end{cases}$$

wobei  $\nu: \mathbb{R} \rightarrow [0, 1]$  eine glatte Funktion ist mit folgenden Eigenschaften:  $\nu(\omega) = 0$  für  $\omega \leq 0$ ,  $\nu(\omega) = 1$  für  $\omega \geq 1$  und  $\nu(\omega) + \nu(1-\omega) = 1$ .

Zum Beispiel kann man  $\nu(\omega) = \omega^4(35 - 84\omega + 70\omega^2 - 20\omega^3)$  für  $\omega \in [0, 1]$  nehmen.

Das Meyer-Wavelet ist unendlich oft differenzierbar und besitzt zudem die Eigenschaft  $\sum_{j \in \mathbb{Z}} |\hat{\psi}(2^j \omega)|^2 = 1/2\pi$ , vgl.[11], S.100. Die Eigenschaften der Fouriertransformierten des Meyer-Wavelets erlauben es, eine Wavelet-Ortho-normalbasis von  $L^2(\mathbb{R}^n)$  zu konstruieren, welche im Rahmen der Littlewood-Paley Charakterisierung der Triebel-Lizorkin Räume genau passt. Die Orthogonalität der Wavelets hebt jedoch diese Funktionenfamilie aus den anderen Wahlmöglichkeiten hervor. Somit kann man eine Beschreibung der Triebel-Lizorkin Räume (welche als Spezialfälle den Hardy-Raum  $H^1$  und die Räume  $BMO$  und  $BMO^{-1}$  enthalten) mit Hilfe der Wavelets erreichen. Auf diese Aspekte wird in dieser Arbeit ausführlicher eingegangen.

Die kontinuierliche Wavelet-Transformation (2.3) liefert Informationen über das Signal  $f$ , lokalisiert um den Punkt  $x = b$ . Der Parameter  $a$  steuert dabei die Größe des Einflubereiches: für  $a \rightarrow 0$  "zoomt" die Wavelet-Transformation immer schärfer auf  $x = b$ .

In vielen Anwendungen, wie z.B. bei der Multiresolutionsanalyse, ist man jedoch daran interessiert, diskrete Werte für die Parameter  $a$  und  $b$  zu betrachten. Dabei ist man bestrebt, die Details einer Funktion auf einer diskreten Hierarchie von Skalen zu erfassen. Die Unterräume, welche den verschiedenen Skalen entsprechen, besitzen Basen, die mit Hilfe einer *Skalierungsfunktion* konstruiert sind, während das orthogonale Komplement eines Raumes im nächsten Element dieser Hierarchie durch eine Wavelet-Basis aufgespannt wird.

**Definition 2** Eine Multiresolutionsanalyse von  $L^2(\mathbb{R})$  ist eine aufsteigende Folge abgeschlossener Unterräume  $\{0\} \subset \dots V_j \subset V_{j-1} \subset \dots \subset L^2(\mathbb{R})$ ,  $j \in \mathbb{Z}$ , mit folgenden Eigenschaften:

$$\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}) \quad (2.4)$$

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\} \quad (2.5)$$

$$f(\cdot) \in V_j \Leftrightarrow f(2^j \cdot) \in V_0 \quad (2.6)$$

Es existiert eine Skalierungsfunktion  $\phi \in V_0$ , so dass deren ganzzahlige Translate eine Riesz-Basis von  $V_0$  erzeugen, d.h.

$$V_0 = \overline{\text{span}\{\phi(\cdot - k) | k \in \mathbb{Z}\}} \quad (2.7)$$

und es existieren Konstanten  $A$  und  $B$ , so dass

$$A \sum_{k \in \mathbb{Z}} |c_k|^2 \leq \left\| \sum_{k \in \mathbb{Z}} c_k \phi(\cdot - k) \right\|_{L^2}^2 \leq B \sum_{k \in \mathbb{Z}} |c_k|^2 \quad (2.8)$$

für alle Folgen  $\{c_k\}_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$ .

Das folgende Resultat gibt eine notwendige und hinreichende Bedingung an die Skalierungsfunktion  $\phi$  an, für welche diese eine Orthonormalbasis von  $V_0$  erzeugt.

**Lemma 1** Für  $\phi \in L^2(\mathbb{R})$  ist  $\{\phi(\cdot - k) | k \in \mathbb{Z}\}$  ein orthonormales System genau dann, wenn

$$\sum_{k \in \mathbb{Z}} |\hat{\phi}(\omega + 2k\pi)|^2 = \frac{1}{2\pi} \quad (2.9)$$

fast überall gilt.

**Beweis.** “ $\Rightarrow$ ” Angenommen für alle  $n \in \mathbb{Z}$  gilt

$$\langle \phi(\cdot - m), \phi(\cdot - n) \rangle = \delta_{mn}.$$

Somit folgt:

$$\begin{aligned} \delta_{0n} &= \langle \phi(\cdot), \phi(\cdot - n) \rangle = \langle \hat{\phi}(\cdot), \hat{\phi}(\cdot) e^{-in\cdot} \rangle \\ &= \int_{\mathbb{R}} |\hat{\phi}(\omega)|^2 e^{in\omega} d\omega \\ &= \int_0^{2\pi} \sum_{k \in \mathbb{Z}} |\hat{\phi}(\omega + 2k\pi)|^2 e^{in\omega} d\omega \end{aligned} \quad (2.10)$$

wobei der Übergang zur letzten Gleichung auf dem Satz von der monotonen Konvergenz beruht. Der  $n$ -te Fourierkoeffizient der  $2\pi$ -periodischen Funktion  $\sum_{k \in \mathbb{Z}} |\hat{\phi}(\omega + 2k\pi)|^2$  ist also  $\delta_{0n}$ . Dies impliziert die Beziehung (2.9).

“ $\Leftarrow$ ” Falls (2.9) erfüllt ist, so gilt:

$$\delta_{0n} = \frac{1}{2\pi} \int_0^{2\pi} e^{in\omega} d\omega = \int_0^{2\pi} \sum_{k \in \mathbb{Z}} |\hat{\phi}(\omega + 2k\pi)|^2 e^{in\omega} d\omega = \langle \phi(\cdot), \phi(\cdot - n) \rangle$$

Nach einer Translation folgt dann auch:

$$\langle \phi(\cdot - m), \phi(\cdot - n) \rangle = \delta_{mn}$$

für alle  $m, n \in \mathbb{Z}$ . □

Die Räume  $V_j$  sind skalierte Versionen des Grundraumes  $V_0$ , der durch Translationen der Skalierungsfunktion  $\phi$  aufgespannt wird. Aus (2.6), (2.7) und (2.8) folgen unmittelbar einige Eigenschaften dieser skalierten Räume:

Der Raum  $V_j$  wird von den Funktionen

$$\phi_{j,k}(x) = 2^{-j/2} \phi(2^{-j}x - k), \quad (2.11)$$

$k \in \mathbb{Z}$  aufgespannt. Es gilt auch  $\|\phi_{j,k}\|_{L^2} = \|\phi\|_{L^2}$ .

$V_0$  ist translationsinvariant, im Sinne dass

$$f \in V_0 \Leftrightarrow f(\cdot - k) \in V_0$$

für  $k \in \mathbb{Z}$ . Somit gilt auch

$$f \in V_j \Leftrightarrow f(\cdot - 2^j k) \in V_j$$

für  $k \in \mathbb{Z}$ .

Die orthogonale Projektion  $P_j f$  einer  $L^2$ -Funktion  $f$  auf dem Raum  $V_j$  wird auch als Darstellung von  $f$  auf der "Skala"  $V_j$  interpretiert, und enthält alle "Details" von  $f$  bis zur Größe  $2^j$ . Je größer  $j$ , desto grober ist die Resolution, in der die Funktion  $f$  dargestellt wird. Der Übergang vom Raum  $V_{j-1}$  zum Raum  $V_j$ , d.h. von der Projektion  $P_{j-1} f$  zur  $P_j f$ , ist also eine Mittelung, eine Darstellung der glatten, niederfrequenten Anteile von  $P_{j-1} f$  in der Basis des Raumes  $V_j$ . Die rauen, hochfrequenten Anteile von  $P_{j-1} f$ , genannt "Details", gehören zum Orthogonalkomplement  $W_j$ , d.h. es gilt  $V_{j-1} = V_j \oplus W_j$ ,  $V_j \perp W_j$ .

Eine Orthonormalbasis des Raumes  $W_j$  besteht genau aus den Wavelets

$$\psi_{j,k} = 2^{-j/2} \psi(2^{-j}x - k) \quad (2.12)$$

mit  $k \in \mathbb{Z}$ , wobei sich diese Wavelets aus einer geeigneten Gleichung mit Hilfe der Skalierungsfunktion  $\phi$  konstruieren lassen.

Sei  $Q_j$  die orthogonale Projektion von  $L^2(\mathbb{R})$  auf  $V_j$ , d.h.

$$P_{j-1} = Q_j + P_j.$$

Es gilt offensichtlich:

$$V_j = \bigoplus_{i \geq j+1} W_i \quad (2.13)$$

und damit

$$L^2(\mathbb{R}) = \bigoplus_{i \in \mathbb{Z}} W_i. \quad (2.14)$$

Die Waveletfamilie  $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$  bildet damit eine Orthonormalbasis von  $L^2(\mathbb{R})$ .

Die Räume  $W_j$  erben die Skalierungseigenschaft (2.6) der Räume  $V_j$ :

$$f \in W_j \Leftrightarrow f(2^j \cdot) \in W_0.$$

Aus (2.13) und (2.14) folgt, dass eine Funktion  $f \in L^2(\mathbb{R})$  wie folgt zerlegt werden kann:

$$\begin{aligned} f &= \sum_{i \in \mathbb{Z}} Q_i f = \sum_{i \geq j+1} Q_i f + \sum_{i \leq j} Q_i f \\ &= P_j f + \sum_{i \leq j} Q_i f. \end{aligned} \quad (2.15)$$

Dies entspricht einer Projektion des Signals  $f$  auf die Skala  $V_j$ . Der verbleibende Rest aus dem Hochfrequenzbereich wird in seine Anteile in verschiedenen Frequenzbändern  $Q_i f$ ,  $i \leq j$  aufgeteilt. Dabei enthält  $Q_i f$  nur die Details, die  $P_{i-1} f$  von  $P_i f$  unterscheiden, d.h.  $Q_i = P_i - P_{i-1}$ .

Das folgende Theorem beinhaltet die Konstruktion der Wavelets, die einer Multiresolutionsanalyse mit gegebener Skalierungsfunktion entsprechen.

**Theorem 1** *Für die Skalierungsfunktion  $\phi$  existiert eine Folge  $\{h_k\}_{k \in \mathbb{Z}}$  reeller Zahlen mit*

$$\phi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} h_k \phi(2x - k) \quad (2.16)$$

(Skalierungsgleichung).

*Falls die Funktionen  $\phi(\cdot - k)$ ,  $k \in \mathbb{Z}$  eine Orthonormalbasis von  $V_0$  bilden, so besitzt die Funktion  $\psi \in V_{-1}$ , definiert durch*

$$\psi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} (-1)^k h_{1-k} \phi(2x - k) = \sum_{k \in \mathbb{Z}} (-1)^k h_{1-k} \phi_{-1,k}(x) \quad (2.17)$$

folgende Eigenschaften:

(i)  $\{\psi_{j,k}(\cdot) = 2^{-j/2} \psi(2^{-j} \cdot - k) \mid k \in \mathbb{Z}\}$  ist eine Orthonormalbasis für  $W_j$ .

(ii)  $\{\psi_{j,k}\}_{j,k \in \mathbb{Z}}$  ist eine Orthonormalbasis für  $L^2(\mathbb{R})$ .

(iii)  $\psi$  ist ein Wavelet mit  $c_\psi = 2\pi \int_{\mathbb{R}} |\omega|^{-1} |\hat{\psi}(\omega)|^2 d\omega = 2 \ln 2$ .

Zum Beweis des Satzes wird auf [11], S. 122-123, verwiesen.

**Bemerkung:** Ein Resultat von Daubechies ermöglicht, für jede gegebene Anzahl  $r$  von verschwindenden Momenten:  $\int_{\mathbb{R}} x^\alpha \psi(x) dx = 0$  für  $\alpha \leq r$  Wavelets mit kompaktem Träger von minimaler Länge zu konstruieren. Zur Bestimmung der entsprechenden Koeffizienten der Skalierungsgleichung werden trigonometrische Polynome eingesetzt. Hierbei wird auf [17] oder [10], S. 249-254 verwiesen. Zum Beispiel erfüllen die Filterkoeffizienten  $\{h_k\}$  im Falle der Daubechies 'db-2'-Wavelets das Gleichungssystem:

$$\begin{aligned} h_0^2 + h_1^2 + h_2^2 + h_3^2 &= 1 \\ h_1 h_0 + h_2 h_1 + h_3 h_2 &= 9/16 \\ h_2 h_0 + h_3 h_1 &= 0 \\ h_3 h_0 &= -1/16 \end{aligned}$$

Alle anderen  $h_k$  sind in diesem Fall gleich Null, s. [11], S.169.

**Bemerkung:** Das Verschwinden der Momente  $\int_{\mathbb{R}} x^\alpha \psi(x) dx = 0$  für  $\alpha \leq r$  impliziert, dass die Wavelets orthogonal zu Polynomen des Grades  $\leq r$  sind. Die Waveletkoeffizienten  $\langle f, \psi_{j,k} \rangle$  sind also klein an den Stellen, an denen die Funktion  $f$  gut durch solche Polynome approximierbar ist. Große Waveletkoeffizienten in der Umgebung eines Punktes deuten entweder auf Singularitäten hin, z.B. Unstetigkeiten oder Nicht-Differenzierbarkeit, oder auf große Werte der Ableitung.

## 2.1.2 Die schnelle Wavelet-Transformation

Im folgenden wird der Algorithmus zur Berechnung der diskreten Wavelet-Transformation vorgestellt.

Gegeben sei eine Funktion  $f \in V_0$ , d.h. aus dem Grundraum der Multiresolutionsanalyse zu der Skalierungsfunktion  $\phi$ . Es gilt also die Darstellung

$$f(x) = \sum_{k \in \mathbb{Z}} c_k^0 \phi(x - k)$$

mit der Koeffizientenfolge  $c^0 = \{c_k^0\}_{k \in \mathbb{Z}} \in \ell_2(\mathbb{Z})$ . Mit der entsprechenden Wavelet-Funktion  $\psi$  bildet  $\{\psi_{j,k} = 2^{-j/2} \psi(2^{-j} \cdot -k) \mid k, j \in \mathbb{Z}\}$  eine Orthonormalbasis von  $L^2(\mathbb{R})$ .

Das Ziel der *schnellen Wavelet-Transformation* ist die Berechnung der Waveletkoeffizienten  $\langle f, \psi_{j,k} \rangle$ ,  $j \in \mathbb{N}$ ,  $k \in \mathbb{Z}$ , wobei  $\langle \cdot, \cdot \rangle$  das Skalarprodukt in  $L^2(\mathbb{R})$  bezeichnet.

Wir führen zunächst folgende Notationen ein:

$$c_k^j = \langle f, \phi_{j,k} \rangle, \quad c^j = \{c_k^j\}_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z}) \quad (2.18)$$

$$d_k^j = \langle f, \psi_{j,k} \rangle, \quad d^j = \{d_k^j\}_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z}) \quad (2.19)$$

Mit den Skalierungsgleichungen (2.16) und (2.17) und mit  $g_i = (-1)^i h_{1-i}$  erhalten wir:

$$c_k^j = \sum_{i \in \mathbb{Z}} h_i \langle f, \phi_{j-1, 2k+i} \rangle = \sum_{i \in \mathbb{Z}} h_{i-2k} c_i^{j-1} \quad (2.20)$$

$$d_k^j = \sum_{i \in \mathbb{Z}} g_i \langle f, \phi_{j-1, 2k+i} \rangle = \sum_{i \in \mathbb{Z}} g_{i-2k} c_i^{j-1} \quad (2.21)$$

Aus dieser Rekursion wird der Algorithmus für die Berechnung der Approximationskoeffizienten  $c^j$  und der Detailkoeffizienten  $d^j$  auf dem Niveau  $j$  hergeleitet, der von einer Darstellung aus dem Grundraum  $V_0$  ausgeht.

Auf dem Folgenraum  $\ell^2(\mathbb{Z})$  definieren wir die Operatoren  $H$  und  $G$  durch:

$$Hc = c *_2 h := \{(Hc)_k = \sum_{i \in \mathbb{Z}} h_{i-2k} c_i\} \quad (2.22)$$

$$Gc = c *_2 g := \{(Gc)_k = \sum_{i \in \mathbb{Z}} g_{i-2k} c_i\} \quad (2.23)$$

für alle Folgen  $c \in \ell^2(\mathbb{Z})$ . Das Symbol  $*_2$  deutet auf eine Analogie mit der üblichen Faltung  $c * h = \sum_{i \in \mathbb{Z}} h_{k-i} c_i$  hin. In der Sprache der Signalverarbeitung bezeichnet  $*_2$  eine Faltung mit anschließendem "Down-Sampling" um den Faktor 2.

Ausgehend von dem gegebenen Signal  $f$ , charakterisiert durch die Koeffizientenfolge  $c^0 = \{c_k\}_{k \in \mathbb{Z}}$ , kann man eine Waveletzerlegung auf  $M$  Skalen wie folgt durchführen:

Berechne für  $j = \overline{1, M}$

$$\begin{aligned} c^j &= Hc^{j-1} \\ d^j &= Gc^{j-1} \end{aligned}$$

Der Algorithmus der schnellen Wavelet-Transformation entspricht also folgendem Schema:

$$\begin{array}{ccccccc} c^0 & \xrightarrow{H} & c^1 & \xrightarrow{H} & c^2 & \dots & c^{M-1} & \xrightarrow{H} & c^M \\ & \searrow G & & \searrow G & & & & \searrow G & \\ & & d^1 & & d^2 & & \dots & & d^M \end{array}$$

Das Ergebnis der Waveletzerlegung besteht aus den Approximationskoeffizienten  $c^M$  auf der Skala  $M$  und den Detailkoeffizienten auf allen Skalen  $d^j$ ,  $j = \overline{1, M}$ . Im Falle endlicher Folgen wird durch das Down-Sampling die Länge der Koeffizientenfolgen  $c^j$  in jedem Schritt etwa halbiert, lediglich an den Rändern können zusätzliche Koeffizienten in Abhängigkeit von den Längen der Filter  $\{h_k\}$  und  $\{g_k\}$  entstehen.

Ausgehend von  $c^M$  und  $d^j$ ,  $j = \overline{1, M}$ , kann das ursprüngliche Signal  $c^0$  konstruiert werden. Wir nehmen zunächst an, dass die Zerlegung auf einem Niveau durchgeführt wurde. Unter Verwendung der Orthogonalität der Zerlegung  $V_0 = V_1 \oplus W_1$  lässt sich  $c^0$  aus  $c^1$  und  $d^1$  rekonstruieren. Wegen der Skalierungsgleichungen gilt:

$$\begin{aligned} \sum_{k \in \mathbb{Z}} c_k^0 \phi_{0,k} &= \sum_{i \in \mathbb{Z}} c_i^1 \phi_{1,i} + \sum_{i \in \mathbb{Z}} d_i^1 \psi_{1,i} \\ &= \sum_{i \in \mathbb{Z}} c_i^1 \sum_{\ell \in \mathbb{Z}} h_\ell \phi_{0,2i+\ell} + \sum_{i \in \mathbb{Z}} d_i^1 \sum_{\ell \in \mathbb{Z}} g_\ell \phi_{0,2i+\ell}. \end{aligned}$$

Ein Koeffizientenvergleich ergibt:

$$c_k^0 = \sum_{\ell \in \mathbb{Z}} c_\ell^1 h_{k-2\ell} + \sum_{\ell \in \mathbb{Z}} d_\ell^1 g_{k-2\ell}.$$

Nach dem gleichen Prinzip wird für  $j = \overline{1, M}$  die Koeffizientenfolge  $c^{j-1}$  aus  $c^j$  und  $d^j$  rekonstruiert:

$$c^{j-1} = H^* c^j + G^* d^j$$

wobei  $H^*$  und  $G^*$  die adjungierten Operatoren zu  $H$  bzw  $G$  sind und durch

$$H^* c := \{(H^* c)_k = \sum_{\ell \in \mathbb{Z}} h_{k-2\ell} c_\ell\} \quad (2.24)$$

$$G^* c := \{(G^* c)_k = \sum_{\ell \in \mathbb{Z}} g_{k-2\ell} c_\ell\} \quad (2.25)$$

definiert sind, s. [11] S. 134.

Die Rekonstruktion erfolgt nach folgendem Schema:

$$\begin{array}{ccccccc} c^M & \xrightarrow{H^*} & c^{M-1} & \xrightarrow{H^*} & c^{M-2} & \dots & c^1 & \xrightarrow{H^*} & c^0 \\ & \nearrow G^* & & \nearrow G^* & & & & \nearrow G^* & \\ d^M & & d^{M-1} & & \dots & & d^1 & & \end{array}$$

In diesem Prozess wird die Länge der Folgen  $c^j$  etwa verdoppelt. Allerdings kann die Länge der Eingabefolge nicht exakt rekonstruiert werden, da bei dem Zerlegungsprozeß Folgen  $c^0$  unterschiedlicher Längen zu gleichlangen Folgen  $c^j, d^j$  führen können. Die rekonstruierte Folge  $c^0$  enthält eventuell einige zusätzliche Nullen.

### 2.1.3 Wavelets in mehreren Dimensionen

Gegeben sei eine Familie von Wavelets, die eine Orthonormalbasis in  $L^2(\mathbb{R})$  bildet, mit Skalierungsfunktion  $\phi$  und ‘‘Mutter-Wavelet’’  $\psi$ , welche folgende Regularitätseigenschaft erfüllen:

$$|\partial^\alpha \phi(x)| \leq C_m(1 + |x|)^{-m} \quad (2.26)$$

$$|\partial^\alpha \psi(x)| \leq C_m(1 + |x|)^{-m} \quad (2.27)$$

für alle  $m \in \mathbb{N}$  und alle Multiindizes  $\alpha = (\alpha_1, \dots, \alpha_n)$  mit  $0 \leq |\alpha| \leq r$ , wobei  $r \in \mathbb{N}$  eine feste natürliche Zahl ist. Diese Funktionen entsprechen einer sogenannten  $r$ -regulären Multiresolutionsanalyse, [2], S.21.

Basierend auf dieser Familie kann eine orthonormale Waveletbasis für  $L^2(\mathbb{R}^n)$  nach folgender Tensorproduktmethode konstruiert werden. Mit  $Q = \{x \in \mathbb{R}^n : 2^j x - k \in [0, 1)^n\}$  für  $k \in \mathbb{Z}^n, j \in \mathbb{Z}$  wird der (halb-offene) dyadische Würfel mit Mittelpunkt  $2^{-j-1}(2k + 1)$  und Kantenlänge  $2^{-j}$  bezeichnet. Sei weiterhin  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$  mit  $\varepsilon_i \in \{0, 1\}$ , wobei  $\varepsilon \neq 0$ .

Definiere

$$\psi_Q^\varepsilon(x) := 2^{nj/2} \psi^{\varepsilon_1}(2^j x_1 - k_1) \dots \psi^{\varepsilon_n}(2^j x_n - k_n) \quad (2.28)$$

wobei  $\psi^0 = \phi$  und  $\psi^1 = \psi$ . Man betrachtet also Produkte von eindimensionalen Funktionen in jeder Variable, wobei mindestens ein Wavelet dabei sein muß. Dies wird durch die Bedingung  $\varepsilon \neq 0$  erreicht, welche die Produkte, die nur aus Dilationen und Translationen der Skalierungsfunktion  $\phi$  bestehen, ausschließt.

Die so konstruierte Familie  $\{\psi_Q^\varepsilon\}$  bildet eine Orthonormalbasis in  $L^2(\mathbb{R}^n)$  und besitzt folgende Eigenschaften, vgl. [2], S.108:

- Regularitätsbedingung:

$$|D^\alpha \psi_Q^\varepsilon| \leq C 2^{j|\alpha|} 2^{nj/2} \quad \text{für } \alpha \in \mathbb{N}_0^n, |\alpha| \leq r. \quad (2.29)$$

- Momentbedingung:

$$\int_{\mathbb{R}^n} x^\alpha \psi_Q^\varepsilon dx = 0 \quad \text{für } |\alpha| \leq r. \quad (2.30)$$

Sei

$$\Lambda = \{(Q, \varepsilon) : Q \text{ dyadischer Würfel, } \varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \text{ mit } \varepsilon_i \in \{0, 1\}, \varepsilon \neq 0\}.$$

Die Elemente  $\lambda \in \Lambda$  kann man daher mit Punkten  $\lambda = 2^{-j}k + 2^{-j-1}\varepsilon$ , ( $k \in \mathbb{Z}^n, \varepsilon \in \{0, 1\}^n, \varepsilon \neq 0$ ) des dyadischen Würfels  $Q = Q(\lambda) = \{x \in \mathbb{R}^n : 2^j x - k \in [0, 1]^n\}$  identifizieren. Für ein dyadischer Würfel  $Q$  bezeichne  $l(Q)$  dessen Kantenlänge. Diese Notationen vereinfachen das Kalkül beim Einsatz dieser multidimensionalen Wavelets in der Charakterisierung der  $BMO$ - und  $BMO^{-1}$ -Funktionen.

Im Falle von Wavelets mit kompaktem Träger sei  $m \geq 1$ , so dass die Träger von  $\psi$  und  $\phi$  in dem Intervall  $m[0, 1] := [1/2 - m/2, 1/2 + m/2]$  enthalten sind. Der Träger  $\psi_Q^\varepsilon$  ist damit im Würfel  $mQ$  enthalten, d.h. in dem Würfel mit demselben Mittelpunkt wie  $Q$ , aber  $m$ -facher Kantenlänge.

### 2.1.4 Wavelet-Analyse diskreter Signale

Die Signale, die man in der Praxis antrifft, sind meistens Daten, die als diskrete Punktemengen vorliegen, z.B. digitalisierte akustische Signale (1-dimensional) oder digitalisierte Bilder (2-dimensional).

Im **eindimensionalen Fall** sind also die diskretisierten Werte

$$s_k = s(kh), \quad k \in \mathbb{Z}$$

gegeben. Um eine Wavelet-Transformation durchzuführen, werden die  $s_k$  als Koeffizienten einer Funktion  $\tilde{f}$  interpretiert:

$$\tilde{f}(t) = \sum_{k \in \mathbb{Z}} s_k \phi(h^{-1}t - k),$$

die nach einer Skalierungsfunktion  $\phi$  entwickelt wurde. Falls  $\phi$  die Interpolationseigenschaft

$$\phi(j) = \delta_{0,j}$$

besitzen würde, dann hätten wir  $\tilde{f}(kh) = s(kh) = s_k$ . Im allgemeinen ist dies jedoch nicht richtig, es gilt aber

$$\tilde{f}(kh) = s_k + O(h^\alpha) \text{ für } h \rightarrow 0$$

falls  $s$  hölderstetig mit dem Exponent  $\alpha \in (0, 1)$  ist, die Skalierungsfunktion  $\phi$  einen kompakten Träger besitzt und die Gleichung  $\sum_{k \in \mathbb{Z}} \phi(x - k) = 1$  für fast alle  $x \in \mathbb{R}$  erfüllt, vgl.[11], S. 238.

Mit der Substitution  $x = h^{-1}t$  erhalten wir

$$f(x) = \tilde{f}(hx) = \sum_{k \in \mathbb{Z}} s_k \phi(x - k)$$

und  $f$  liegt somit im Raum  $V_0$  der Multiresolutionsanalyse zur Skalierungsfunktion  $\phi$ .

Nun kann die schnelle Wavelet-Transformation angewandt werden. Mit  $c^0 = \{s_k\}_{k \in \mathbb{Z}}$  berechnet man

$$c^m = Hc^{m-1}, \quad d^m = Gc^{m-1},$$

also die Koeffizienten der Approximation und die Detailkoeffizienten auf dem Niveau  $m$ , für  $m = 1, M$ .

Falls das Signal eine **2-dimensionale** Struktur hat, z.B. ein digitalisiertes Bild oder eine digitalisierte Oberfläche, liegen die Punktedaten meist in Form einer Matrix  $c^0 = \{c_{ij}\}_{i,j=1}^n$  vor. Mit Hilfe der Operatoren  $H$  und  $G$  werden folgende Operatoren, die auf Matrizen angewandt werden können, definiert:

$$H_Z c := \{(H_Z c)_{ij} = \sum_{\ell \in \mathbb{Z}} h_{\ell-2j} c_{i\ell}\} \quad (2.31)$$

$$H_S c := \{(H_S c)_{ij} = \sum_{\ell \in \mathbb{Z}} h_{\ell-2i} c_{\ell j}\} \quad (2.32)$$

$$G_Z c := \{(G_Z c)_{ij} = \sum_{\ell \in \mathbb{Z}} g_{\ell-2j} c_{i\ell}\} \quad (2.33)$$

$$G_S c := \{(G_S c)_{ij} = \sum_{\ell \in \mathbb{Z}} g_{\ell-2i} c_{\ell j}\} \quad (2.34)$$

$H_Z$  bedeutet die Anwendung des Operators  $H$  auf die Zeilen der Matrix  $c$ , während  $H_S$  die Anwendung von  $H$  auf die Spalten der Matrix darstellt. Die Operatoren  $G_S$  und  $G_Z$  haben eine analoge Bedeutung.

Im ersten Schritt der 2-dimensionalen Waveletzerlegung erhält man also folgende Matrizen:

$$c^1 = H_S H_Z c^0, \quad d_h^1 = G_S H_Z c^0, \quad d_v^1 = H_S G_Z c^0, \quad d_d^1 = G_S G_Z c^0$$

$c^1$  enthält die Approximationskoeffizienten auf der ersten Skala, während  $d_h^1, d_v^1, d_d^1$  die Detailkoeffizienten (horizontal, vertikal und diagonal) enthalten.

Die Prozedur kann wiederholt werden durch eine rekursive Anwendung auf die im vorherigen Schritt erhaltenen Approximationskoeffizienten. Nach einer Zerlegung auf  $M$  Niveaus erhält man so die Approximationskoeffizienten  $c^M$  auf der Skala  $M$ , sowie alle Detailkoeffizienten  $d_h^j, d_v^j, d_d^j$ ,  $j = \overline{1, M}$ . Ausgehend von diesen Daten kann auch hier, analog zum eindimensionalen Fall, das Originalsignal  $c^0$  rekonstruiert werden. Ein Beispiel für eine 2-dimensionale Waveletzerlegung ist in Abbildung 2.1 dargestellt.

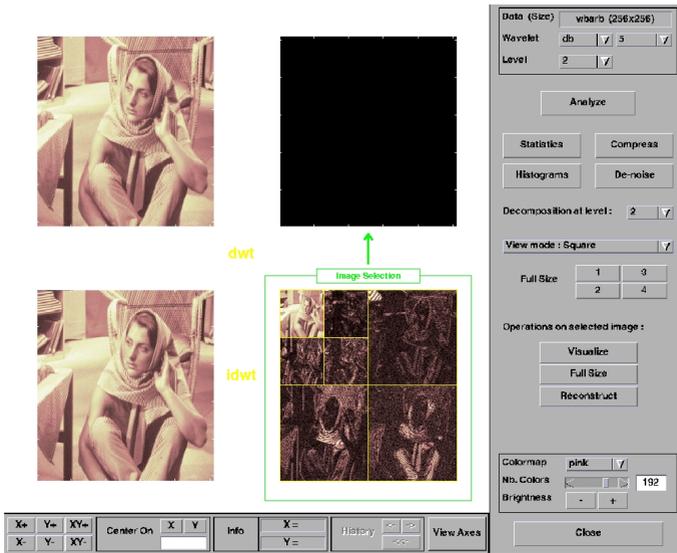


Abbildung 2.1: 2-Dimensionale Waveletzerlegung

Das Originalbild (oben links) wird auf zwei Skalen mit Hilfe der Daubechies 'db5' -Wavelets zerlegt (unten rechts). Die Zerlegung besteht aus der Matrix  $c^2$  (Approximiertes Bild auf der Skala 2) und den Detailkoeffizienten  $d_h^2, d_v^2, d_d^2, d_h^1, d_v^1, d_d^1$ , d.h. die anderen Matrizen, welche horizontal, vertikal oder diagonal zu  $c^2$  auf den entsprechenden Skalen positioniert sind. Unten links ist das rekonstruierte Bild dargestellt.

## 2.2 Die Räume $BMO(\mathbb{R}^n)$ und $BMO^{-1}(\mathbb{R}^n)$

### 2.2.1 Der Raum $BMO(\mathbb{R}^n)$

Der Raum der  $BMO$ -Funktionen auf  $\mathbb{R}^n$  wird in der harmonischen Analysis durch folgende Eigenschaft definiert:

$$\|f\|_{BMO} := \sup_B \left( \frac{1}{|B|} \int_B |f(x) - f_B| dx \right) < \infty \quad (2.35)$$

wobei  $f \in L^1_{loc}(\mathbb{R}^n)$ .  $B$  ist eine Kugel in  $\mathbb{R}^n$  mit Volumen  $|B|$  und  $f_B = |B|^{-1} \int_B f(x) dx$  bezeichnet den Mittelwert der Funktion  $f$  über  $B$ . Der Name  $BMO$  kommt aus dem englischen 'bounded mean oscillation', welches die Eigenschaft (2.35) ausdrückt.

**Bemerkung:** Die obige "Norm" verschwindet für konstante Funktionen.  $\|\cdot\|_{BMO}$  wird also zur echten Norm nur, wenn man alle  $BMO$ -Funktionen, die sich durch eine Konstante unterscheiden, miteinander identifiziert.

Der Raum  $BMO$  dient oft auch als 'Ersatz' für  $L^\infty$ . Beschränkte Funktionen sind offensichtlich Elemente aus  $BMO$ . Die Klasse der  $BMO$ -Funktionen ist jedoch etwas ausgedehnter, ein Beispiel einer unbeschränkten  $BMO$ -Funktion ist  $f(x) = \log|x|$ .

Die *sharp*-Funktion einer Funktion  $f \in L^1_{loc}(\mathbb{R}^n)$  ist definiert durch:

$$f^\sharp(x) = \sup_{B \ni x} \frac{1}{|B|} \int_B |f(y) - f_B| dy \quad (2.36)$$

wobei das Supremum über alle Kugeln  $B$ , die den Punkt  $x$  enthalten, betrachtet wird. Es gilt offensichtlich:  $f \in BMO(\mathbb{R}^n) \Leftrightarrow f^\sharp \in L^\infty(\mathbb{R}^n)$ .

**Lemma 2** Falls  $f, g \in BMO$ , so sind  $|f|, \min(f, g), \max(f, g)$  Elemente aus  $BMO$ .

**Beweis.** Wegen  $2 \min(f, g) = f + g - |f - g|$  bzw.  $2 \max(f, g) = f + g + |f - g|$ , genügt es zu zeigen, dass für  $f \in BMO$  auch  $|f| \in BMO$ . Wegen  $\| |f| - |f_B| \| \leq |f - f_B|$  ist aber die Behauptung offensichtlich.  $\square$

Die  $BMO$ -Funktionen erfüllen zusätzliche Ungleichungen, welche deren definierende Eigenschaft (2.35) verallgemeinern, vgl. [1], S.144.

**Theorem 2** (Die John-Nirenberg Ungleichungen)

Die  $BMO$ -Funktionen  $f$  erfüllen folgende Ungleichungen:

(i)

$$\frac{1}{|B|} \int_B |f(x) - f_B|^p dx \leq c_p \|f\|_{BMO}^p \quad (2.37)$$

für alle  $p < \infty$  und alle Kugeln  $B$ .

(ii) Es existieren positive Konstanten  $c_1$  und  $c_2$ , so dass für alle  $\alpha > 0$  und alle Kugeln  $B$  gilt:

$$|\{x \in B : |f(x) - f_B| > \alpha\}| \leq c_1 e^{-c_2 \alpha / \|f\|_{BMO}} |B| \quad (2.38)$$

## 2.2.2 Dualität zwischen $BMO$ und $H^1$

Die Funktionen aus dem Hardyraum  $H^1$  sind durch deren *atomare Zerlegung* charakterisiert:  $g = \sum_{k=0}^{\infty} \lambda_k a_k$  ( $L^1$ -Konvergenz), mit

$\sum_{k=0}^{\infty} |\lambda_k| < \infty$ , wobei  $a_k$  sogenannte *Atome* sind, d.h. Funktionen charakterisiert durch folgende Eigenschaften:  $\text{supp } a_k \subset B_k$ , (der Träger ist Teilmenge einer Kugel),  $|a_k| \leq |B_k|^{-1}$  fast überall und  $\int a_k(x) dx = 0$ .

Die  $H^1$ -Norm ist definiert durch:  $\|g\|_{H^1} = \inf \sum_{k=0}^{\infty} |\lambda_k|$ , wobei das Infimum über alle atomaren Zerlegungen von  $f$  genommen wird.

Der Dualraum von  $H^1$  ist  $BMO$ . Dies wird ausgedrückt durch die Darstellung der stetigen linearen Funktionalen  $l$  auf  $H^1$  in der Form:

$$\ell(g) = \sum_{k=0}^{\infty} \lambda_k \int_{\mathbb{R}^n} f(x) a_k(x) dx \quad \forall g = \sum_k \lambda_k a_k \in H^1(\mathbb{R}^n) \quad (2.39)$$

mit  $f \in BMO(\mathbb{R}^n)$ .

Es gilt zunächst folgende Ungleichung:

**Theorem 3** Für alle  $f \in BMO$  und  $g = \sum_k \lambda_k a_k \in H^1$  gilt die Abschätzung:

$$\left| \sum_{k=0}^{\infty} \lambda_k \int_{\mathbb{R}^n} f(x) a_k(x) dx \right| \leq C \|f\|_{BMO} \|g\|_{H^1}. \quad (2.40)$$

Folglich entspricht jeder  $BMO$ -Funktion  $f$  ein stetiges lineares Funktional  $\ell$  auf  $H^1$ , gegeben durch (2.39).

**Beweis.** Sei zunächst  $f \in L^\infty$ . Wegen der  $L^1$ -Konvergenz der Reihe  $g = \sum_k \lambda_k a_k$  und wegen  $\int a_k = 0$  gilt:

$$\begin{aligned} \int_{\mathbb{R}^n} f(x)g(x)dx &= \sum_k \lambda_k \int_{\mathbb{R}^n} f(x)a_k(x)dx \\ &= \sum_k \lambda_k \int_{\mathbb{R}^n} [f(x) - f_{B_k}]a_k(x)dx, \end{aligned}$$

wobei  $\text{supp } a_k \subset B_k$ . Wegen  $|a_k| \leq |B_k|^{-1}$  gilt weiterhin:

$$\left| \int_{\mathbb{R}^n} f(x)g(x)dx \right| \leq \sum_k \frac{|\lambda_k|}{|B_k|} \int_{B_k} |f(x) - f_{B_k}|dx \leq \|f\|_{BMO} \sum_k |\lambda_k|.$$

Diese Abschätzung gilt für alle Darstellungen  $g = \sum_k a_k \lambda_k$ . Die Ungleichung (2.40) folgt nun aus  $\|g\|_{H^1} = \inf \sum_k |\lambda_k|$ .

Sei nun  $f \in BMO$  beliebig. Für die  $L^\infty$ -Funktionen  $f^{(N)}$  definiert durch

$$f^{(N)}(x) = \begin{cases} -N & \text{für } f(x) \leq -N \\ f(x) & \text{für } -N \leq f(x) \leq N \\ N & \text{für } f(x) \geq N \end{cases}$$

gilt nach Lemma 2:  $\|f^{(N)}\|_{BMO} \leq c\|f\|_{BMO}$ , wobei  $c$  eine Konstante (unabhängig von  $N$ ) ist. Nach dem bereits bewiesenen Schritt gilt also:

$$\left| \int_{\mathbb{R}^n} f^{(N)}(x)g(x)dx \right| \leq C\|f\|_{BMO}\|g\|_{H^1} \quad (2.41)$$

für alle  $g \in H^1$ . Wegen  $f \in L^1_{loc}$  und  $a_k \in L^\infty$  folgt:

$$|f^{(N)}(x)a_k(x)| \leq |f(x)a_k(x)| \in L^1(B_k).$$

Der Satz von der dominierten Konvergenz impliziert daher:

$$\lim_{N \rightarrow \infty} \int_{B_k} f^{(N)}(x)a_k(x)dx = \int_{B_k} f(x)a_k(x)dx$$

für alle  $k$ . Wegen der Ungleichung (2.41) mit  $g = \lambda_k a_k$  liefert eine weitere Anwendung des Satzes von der dominierten Konvergenz auf  $l^1(\mathbb{N})$ :

$$\lim_{N \rightarrow \infty} \sum_{k \in \mathbb{N}} \lambda_k \int_{\mathbb{R}^n} f^{(N)}(x)a_k(x)dx = \sum_{k \in \mathbb{N}} \lambda_k \int_{\mathbb{R}^n} f(x)a_k(x)dx.$$

Zusammen mit (2.41) impliziert dies die gewünschte Behauptung.  $\square$

Die Umkehrung gilt auch, vgl. [1], S.142-144:

**Theorem 4** *Jedes stetige lineare Funktional  $\ell$  auf  $H^1$  besitzt eine Darstellung (2.39) mit  $f \in BMO$ , so dass die Abschätzung*

$$\|f\|_{BMO} \leq C\|\ell\|$$

*gilt.*

### 2.2.3 $H^1(\mathbb{R}^n)$ , $BMO(\mathbb{R}^n)$ und Wavelets

Wir betrachten nun Funktionen  $f$ , für die folgende Waveletzerlegung gilt:

$$f = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda = \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \psi_\lambda \quad (2.42)$$

wobei  $\{\psi_\lambda\}_{\lambda \in \Lambda}$  eine nach (2.28) konstruierte Waveletbasis von  $L^2(\mathbb{R}^n)$  mit Regularität  $r \geq 1$  ist. In diesem Kapitel betrachten wir Wavelets mit *kompaktem Träger*, obwohl die gleichen Ergebnisse auch mit Wavelets ohne diese Eigenschaft gelten, z.B. die Meyer-Wavelets. Dies wird später bei der Charakterisierung der Triebel-Lizorkin Räume erläutert.

Für  $f \in L^p(\mathbb{R}^n)$  mit  $1 < p < \infty$  konvergiert die Reihe (2.42) *unbedingt*, d.h.  $\sum_{\lambda} c_{\sigma(\lambda)} \psi_{\sigma(\lambda)}$  konvergiert in der  $L^p$ -Norm für beliebige

Permutationen  $\sigma$  der Elemente  $\lambda \in \Lambda$ . Diese Eigenschaft gilt nicht für den Fall  $p = 1$ . Die  $L^1$ -Funktionen  $f$ , für welche die Reihe (2.42) unbedingt in der  $L^1$ -Norm konvergiert, bilden genau den Hardyraum  $H^1(\mathbb{R}^n)$ , vgl. [2], S.138-141.

Das nächste Ergebnis beinhaltet die Darstellung der  $BMO$ -Funktionen mit Hilfe der Reihe (2.42) und drückt die Dualität zwischen  $H^1$  und  $BMO$  im Kontext der Waveletbasis aus, vgl. [2], S.154:

**Theorem 5** (i) *Sei  $f \in BMO(\mathbb{R}^n)$ . Dann erfüllen die Waveletkoeffizienten  $c_\lambda = \langle f, \psi_\lambda \rangle$  folgende Bedingung: es existiert eine Konstante  $C > 0$ , so dass für jeden dyadischen Würfel  $Q$  gilt:*

$$\sum_{Q(\lambda) \subset Q} |c_\lambda|^2 \leq C|Q| \quad (2.43)$$

wobei  $Q(\lambda)$  die Menge aller dyadischen Sub-Würfel des Würfels  $Q$  durchläuft. (Für jeden solcher Sub-Würfel gibt es also  $2^n - 1$  Werte für  $\lambda$ , daher auch für  $\psi_\lambda$  und  $c_\lambda$ , entsprechend zu allen möglichen Werten von  $\varepsilon$ ).

(ii) *Falls die Waveletkoeffizienten die Bedingung (2.43) erfüllen, konvergiert die Reihe  $\sum c_\lambda \psi_\lambda(x)$  gegen eine  $BMO$ -Funktion  $f$  im*

*Sinne der schwach-\* Topologie auf  $BMO$ , betrachtet als Dualraum von  $H^1$ .*

**Beweis.** Im folgenden bezeichnen  $C, C', \dots$  generische Konstanten, welche nur von der Raumdimension und von der Waveletfamilie abhängen.

(i) Sei  $f \in BMO$  und  $Q$  ein beliebiger dyadischer Würfel. Sei  $B$  die Kugel mit demselben Mittelpunkt wie  $Q$  und Radius  $\sqrt{n}l(Q)$ , wobei  $l(Q)$  die Kantenlänge des Würfels  $Q$  bezeichnet. Die Bezeichnungen  $mQ, mB$  stehen für den Würfel, bzw. die Kugel mit demselben Mittelpunkt  $Q$  oder  $B$ , aber mit  $m$ -facher Kantenlänge bzw. Radius. Wir zerlegen  $f = f_1 + f_2 + f_{mB}$ , wobei  $f_{mB} = |mB|^{-1} \int_{mB} f$  wie üblich für den Mittelwert steht, während  $f_1, f_2$  definiert sind durch:

$$f_1(x) = \begin{cases} f(x) - f_{mB} & x \in mB \\ 0 & \text{sonst} \end{cases}$$

$$f_2(x) = \begin{cases} 0 & x \in mB \\ f(x) - f_{mB} & \text{sonst} \end{cases}$$

Wir haben somit:

$$\sum_{Q(\lambda) \subset Q} |c_\lambda|^2 = \sum_{Q(\lambda) \subset Q} |\langle f, \psi_\lambda \rangle|^2 = \sum_{Q(\lambda) \subset Q} |\langle f_1 + f_2 + f_{mB}, \psi_\lambda \rangle|^2 \quad (2.44)$$

Für die Waveletfunktionen  $\psi_\lambda$  in der obigen Summe gilt durch deren Konstruktion  $\text{supp } \psi_\lambda \subset mQ(\lambda) \subset mQ \subset mB$ . Aus der Definition von  $f_2$  folgt also, dass  $\langle f_2, \psi_\lambda \rangle = 0$ . Wegen der Momentbedingung der Wavelets gilt weiterhin:  $\langle f_{mB}, \psi_\lambda \rangle = f_{mB} \int \psi_\lambda(x) dx = 0$ .

Wir erhalten also:

$$\begin{aligned} \sum_{Q(\lambda) \subset Q} |c_\lambda|^2 &= \sum_{Q(\lambda) \subset Q} |\langle f_1, \psi_\lambda \rangle|^2 \leq \sum_{\lambda \in \Lambda} |\langle f_1, \psi_\lambda \rangle|^2 \\ &= \|f_1\|_{L^2(\mathbb{R}^n)}^2 = \|f_1\|_{L^2(mB)}^2 \end{aligned}$$

wobei die letzte Gleichheit auf den Eigenschaften des Trägers von  $f_1$  beruht. Die John-Nirenberg Ungleichung (2.37) impliziert daher

$$\|f_1\|_{L^2(mB)}^2 = \int_{mB} |f(x) - f_{mB}|^2 dx \leq C \|f\|_{BMO}^2 |mB| \leq C \|f\|_{BMO}^2 |Q| \quad (2.45)$$

wobei die letzte Konstante  $C$  nur von  $n$  und  $m$  abhängt.

(ii) Angenommen die Waveletkoeffizienten  $\{c_\lambda\}$  erfüllen die Bedingung (2.43).

Sei  $B$  die Kugel mit Mittelpunkt  $x_0$  und Radius  $r$  und  $q \in \mathbb{Z}$ , so dass  $2^{-q} \leq r < 2^{-q+1}$  gilt. Mit  $Q_j$  bezeichnet man im folgenden einen generischen dyadischen Würfel mit Kantenlänge  $2^{-j}$ .

Für den Moment nehmen wir an, dass nur endlich viele Koeffizienten  $c_\lambda$  ungleich Null sind. Wir zerlegen  $\sum c_\lambda \psi_\lambda(x) =: f(x) = f_1(x) + f_2(x) + f_3(x)$  mit

$$f_1(x) = \sum_{\substack{Q_j(\lambda): j \geq q \\ mQ_j(\lambda) \cap B \neq \emptyset}} c_\lambda \psi_\lambda(x) \quad (2.46)$$

$$f_2(x) = \sum_{\substack{Q_j(\lambda): j < q \\ mQ_j(\lambda) \cap B \neq \emptyset}} c_\lambda \psi_\lambda(x) \quad (2.47)$$

$$f_3(x) = \sum_{mQ(\lambda) \cap B = \emptyset} c_\lambda \psi_\lambda(x) \quad (2.48)$$

In der ersten Summe wird über alle dyadischen Würfel mit einer Kantenlänge von höchstens  $2^{-q}$  summiert, in der zweiten über alle Würfel mit einer Kantenlänge größer oder gleich  $2^{-q+1}$ . Beide Summen entsprechen denjenigen Waveletfunktionen, deren Träger (bzw. die Würfel  $mQ(\lambda)$  welche den Träger enthalten) die Kugel  $B$  schneiden.  $f_3$  steht schließlich für die restlichen Terme, welche den Waveletfunktionen entsprechen, deren Träger disjunkt zu der Kugel  $B$  sind, da sie Teilmengen der Würfel  $mQ(\lambda)$  sind. Aus diesem Grund gilt  $f_3(x) = 0$  für  $x \in B$ .

Wir betrachten nun den Fall  $mQ_j(\lambda) \cap B \neq \emptyset$  für  $j \geq q$ . Dann gibt es eine Konstante  $M$  (abhängig nur von  $n$  und  $m$ ), so dass  $Q_j(\lambda) \subset MB$  gilt. Die Bedingung (2.43) (mit einer geeigneten Konstante, so dass das Ergebnis auch für Kugeln gültig ist) impliziert:

$$\|f_1\|_{L^2(B)}^2 \leq \sum_{Q(\lambda) \subset MB} |c_\lambda|^2 \leq C|B| \quad (2.49)$$

Als nächstes betrachten wir den Fall  $mQ_j(\lambda) \cap B \neq \emptyset$  für  $j < q$ . Es gilt:

$$\begin{aligned}
 |f_2(x) - f_2(x_0)| &\leq \sum_{\substack{Q_j(\lambda): j < q \\ mQ_j(\lambda) \cap B \neq \emptyset}} |c_\lambda| |\psi_\lambda(x) - \psi_\lambda(x_0)| \\
 &\leq \sum_{\substack{Q_j(\lambda): j < q \\ mQ_j(\lambda) \cap B \neq \emptyset}} C |Q_j(\lambda)|^{1/2} |\psi_\lambda(x) - \psi_\lambda(x_0)| \\
 &\leq \sum_{\substack{Q_j(\lambda): j < q \\ mQ_j(\lambda) \cap B \neq \emptyset}} C 2^{-nj/2} C' 2^j 2^{nj/2} |x - x_0| \\
 &= \sum_{\substack{Q_j(\lambda): j < q \\ mQ_j(\lambda) \cap B \neq \emptyset}} C'' 2^j |x - x_0|
 \end{aligned} \tag{2.50}$$

$$\begin{aligned}
 &\leq \sum_{\substack{Q_j(\lambda): j < q \\ mQ_j(\lambda) \cap B \neq \emptyset}} C 2^{-nj/2} C' 2^j 2^{nj/2} |x - x_0| \\
 &= \sum_{\substack{Q_j(\lambda): j < q \\ mQ_j(\lambda) \cap B \neq \emptyset}} C'' 2^j |x - x_0|
 \end{aligned} \tag{2.51}$$

In der Herleitung dieser Ungleichungen wurde sowohl die Bedingung (2.43) benutzt (wobei in der linken Seite nur  $|c_\lambda|$  steht), als auch die Regularitätseigenschaft  $r \geq 1$  der Wavelets. Für jedes  $j < q$  gibt es höchstens  $M^n$  dyadische Würfel  $mQ_j(\lambda)$  (mit Kantenlänge  $m2^{-j}$ ), für welche  $mQ_j(\lambda) \cap B \neq \emptyset$  gilt. Mit  $C' = M^n C''$  kann man aus Gleichung (2.51) schließen, dass für  $x \in B$

$$|f_2(x) - f_2(x_0)| \leq \sum_{Q_j(\lambda): j < q} C' 2^j |x - x_0| = 2^q C' |x - x_0| \leq 2C' \tag{2.52}$$

gilt, da  $|x - x_0| \leq r < 2^{-q+1}$ .

Aus (2.49) und (2.52) folgt:

$$\begin{aligned}
 \int_B |f(x) - f_2(x_0)|^2 dx &\leq 2 \int_B |f_1(x)|^2 dx + 2 \int_B |f_2(x) - f_2(x_0)|^2 dx \\
 &\leq 2C|B| + 8C'^2|B| = C''|B|
 \end{aligned} \tag{2.53}$$

Es gilt also:

$$\begin{aligned}
 \int_B |f(x) - f_2(x_0)| dx &\leq \left( \int_B |f(x) - f_2(x_0)|^2 dx \right)^{1/2} |B|^{1/2} \\
 &\leq (C''|B|)^{1/2} |B|^{1/2} = C'''|B|.
 \end{aligned}$$

Als Konsequenz gilt also, dass  $f \in BMO$  mit  $\|f\|_{BMO} \leq C'''$ , mit  $C'''$  eine Konstante, die unabhängig von der Anzahl der von Null verschiedenen Waveletkoeffizienten  $c_\lambda$  ist. Es gilt zusätzlich  $C''' \leq$

$C' C^{1/2}$ , wobei die Konstante  $C'$  nur von  $n$  und  $m$  abhängt, mit der Konstante

$$C := \sup_Q |Q|^{-1} \sum_{Q(\lambda) \subset Q} |c_\lambda|^2$$

aus (2.43). Dies folgt aus (2.49) und (2.50).

Wir betrachten nun den allgemeinen Fall.

Sei

$$f_{\rho_1}^{\rho_2} = \sum_{\rho_1 \leq l(Q(\lambda)) \leq \rho_2} c_\lambda \psi_\lambda = f_{\rho_1}^1 + f_1^{\rho_2}$$

für  $\rho_1 \leq 1 < \rho_2$ . Wir haben eben gezeigt, dass  $f_{\rho_1}^1$  und  $f_1^{\rho_2}$  im Raum  $BMO$  liegen, gleichmäßig bezüglich  $\rho_1, \rho_2$ . Um die Behauptung des Theorems zu erhalten, muss die Existenz der Limiten

$$\lim_{\rho_1 \rightarrow 0} \int f_{\rho_1}^1(x) g(x) dx \quad \text{und} \quad \lim_{\rho_2 \rightarrow \infty} \int f_1^{\rho_2}(x) g(x) dx$$

für alle Funktionen  $g$  aus einem dichten Teilraum von  $H^1$  nachgewiesen werden. Zu diesem Zweck kann man den Raum  $H_a^1$  der endlichen linearen Kombinationen von  $H^1$ -Atomen wählen, der aus Funktionen mit kompaktem Träger und verschwindendem Integral besteht.

Sei  $g \in H_a^1$  mit  $\text{supp } g \subset B$ . Da  $g \in L^2$ , um die Existenz von  $\lim_{\rho_1 \rightarrow 0} \langle f_{\rho_1}^1, g \rangle$  zu beweisen, ist es ausreichend die  $L^2(B)$ -Konvergenz von  $f_{\rho_1}^1$  für  $\rho_1 \rightarrow 0$  nachzuweisen. Aber dies folgt aus (2.49), wo endliche Summen mit beliebig vielen Termen gleichmäßig in  $L^2(B)$  abgeschätzt wurden.

Es bleibt zu zeigen, dass :

$$\lim_{\rho_1, \rho_2 \rightarrow \infty} \left[ \int_B f_{\rho_1}^{\rho_2}(x) g(x) dx - \int_B f_1^{\rho_2}(x) g(x) dx \right] = 0.$$

Es gilt:

$$\begin{aligned} & \left| \int_B f_{\rho_1}^{\rho_2}(x) g(x) dx - \int_B f_1^{\rho_2}(x) g(x) dx \right| \leq \\ & \leq \sum_{l(Q(\lambda)) \geq \min(\rho_1, \rho_2)} |c_\lambda| \left| \int_B \psi_\lambda(x) g(x) dx \right| \\ & \leq C \sum_{l(Q(\lambda)) \geq \min(\rho_1, \rho_2)} |c_\lambda| \int_B |\psi_\lambda(x) - \psi_\lambda(\bar{x})| dx \end{aligned}$$

wobei die Eigenschaften  $g \in L^\infty$ ,  $\int g = 0$  ausgenutzt wurden, mit  $\bar{x}$  als beliebigem Punkt aus  $B$ . Analog zu den Abschätzungen (2.50),

(2.51) und (2.52), wo die Regularitätseigenschaft der Wavelets und die Ungleichung  $|c_\lambda| \leq C|Q(\lambda)|^{1/2}$  verwendet wurden, erhalten wir weiterhin:

$$\begin{aligned}
 & \left| \int_B f_1^{\rho_1}(x)g(x)dx - \int_B f_1^{\rho_2}(x)g(x)dx \right| \leq \\
 & \leq C \sum_{\substack{mQ_j(\lambda) \cap B \neq \emptyset \\ 2^{-j} \geq \min(\rho_1, \rho_2)}} 2^{-nj/2} 2^{j(n/2+1)} |x - \bar{x}| \\
 & = C \sum_{\substack{mQ_j(\lambda) \cap B \neq \emptyset \\ 2^{-j} \geq \min(\rho_1, \rho_2)}} 2^j |x - \bar{x}| \\
 & \leq C_B \sum_{2^{-j} \geq \min(\rho_1, \rho_2)} 2^j \longrightarrow_{\rho_1, \rho_2 \rightarrow \infty} 0
 \end{aligned}$$

Der Beweis des Satzes ist somit beendet. □

**Bemerkung.** Der hier präsentierte Beweis kombiniert die Argumente aus [2], S.154-155 und [1], S.166-169.

Als Konsequenz ergibt sich folgendes Resultat:

**Korollar 1** Die  $BMO$ -Norm der Funktion  $f = \sum c_\lambda \psi_\lambda$  ist vergleichbar mit der Größe

$$\sup_Q \left( \frac{1}{|Q|} \sum_{Q(\lambda) \subset Q} |c_\lambda|^2 \right)^{1/2}. \tag{2.54}$$

Der Supremum wird hier über alle dyadischen Würfel  $Q$  genommen.

### 2.2.4 Der Raum $BMO^{-1}(\mathbb{R}^n)$

Der Raum  $BMO^{-1}$  wird in der angewandten Analysis benutzt, um Texturen oder schnell oszillierende Muster zu modellieren. Er wird im Kontext der Anwendungen in der Bildverarbeitung, vgl. [3], oder der Navier-Stokes Gleichungen, vgl. [4] und [3], vorgestellt.

Eine Distribution  $f$  ist Element von  $BMO^{-1}(\mathbb{R}^n)$  genau dann, wenn  $g_i \in BMO(\mathbb{R}^n)$  existiert, mit  $f = \operatorname{div} (g_1, \dots, g_n) = \sum_{i=1}^n \partial_i g_i$ . Die Ableitungen werden hier im Distributionssinn betrachtet.

Es gilt:  $\|f\|_{BMO^{-1}} = \inf \sum_{i=1}^n \|g_i\|_{BMO}$  wobei das Infimum über alle Zerlegungen  $f = \sum_{i=1}^n \partial_i g_i$  genommen wird. Im Unterschied zu  $BMO$  ist hier  $\|\cdot\|_{BMO^{-1}}$  eine echte Norm.

## 2.2.5 $BMO$ und $BMO^{-1}$ als homogene Triebel-Lizorkin Räume

Die bisher betrachteten Räume:  $H^1$ ,  $BMO$  und  $BMO^{-1}$  sind Spezialfälle der sogenannten *homogenen Triebel-Lizorkin Räume*  $\dot{F}_{p,q}^\alpha$ . Die Eigenschaften dieser Räume werden z.B. in [12] und [15] erläutert.

Sei  $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$  der Schwartz-Raum der schnell abfallenden  $C^\infty$ -Funktionen und sei  $\phi \in \mathcal{S}$  mit folgenden Eigenschaften:

$$\text{supp}\hat{\phi} \subset \{\omega \in \mathbb{R}^n : 1/2 \leq |\omega| \leq 2\} \quad (2.55)$$

$$|\hat{\phi}(\omega)| \geq c > 0 \text{ für } 3/5 \leq |\omega| \leq 5/3. \quad (2.56)$$

Für  $\alpha \in \mathbb{R}$ ,  $0 < p < +\infty$ ,  $0 < q \leq +\infty$  und eine beliebige Funktion  $\phi \in \mathcal{S}$ , welche (2.55) und (2.56) erfüllt, wird der *homogene Triebel-Lizorkin Raum*  $\dot{F}_{p,q}^\alpha$  als die Menge aller  $f \in \mathcal{S}'/\mathcal{P}$  (temperierte Distributionen modulo Polynome) definiert, für welche

$$\|f\|_{\dot{F}_{p,q}^\alpha} := \left\| \left( \sum_{j \in \mathbb{Z}} (2^{j\alpha} |\phi_j * f|)^q \right)^{1/q} \right\|_{L^p} < +\infty \quad (2.57)$$

gilt, wobei  $\phi_j(x) = 2^{nj} \phi(2^j x)$ . Für  $q = \infty$  ersetzt man die  $l^q$ -Norm mit der sup-Norm.

Für den Fall  $p = \infty$ ,  $q < +\infty$  kann man die  $L^p$ -Norm in (2.57) nicht einfach durch die  $L^\infty$ -Norm ersetzen, weil die so erhaltene quasi-Norm von der Funktion  $\phi$  abhängig wäre, s.[15]. Eine Erweiterung der Skala der Triebel-Lizorkin Räume wäre aber auch für  $p = \infty$ ,  $q < +\infty$  wünschenswert, unter anderem auch aus Dualitätsgründen. Man definiert daher:

$$\|f\|_{\dot{F}_{\infty,q}^\alpha} := \left( \sup_Q \frac{1}{|Q|} \int_Q \sum_{j=-\log_2 l(Q)}^{\infty} (2^{j\alpha} |(\phi_j * f)(x)|)^q dx \right)^{1/q} \quad (2.58)$$

wobei  $Q$  die Menge aller dyadischen Würfeln durchläuft. Der entsprechende Raum  $\dot{F}_{\infty,q}^\alpha(\mathbb{R}^n)$  definiert sich als Menge aller  $f \in \mathcal{S}'/\mathcal{P}$  mit  $\|f\|_{\dot{F}_{\infty,q}^\alpha} < +\infty$ .

Die Eigenschaften (2.57) bzw. (2.58) sind unabhängig von der Wahl der Funktion  $\phi$ , vgl. [15] oder [12], S.240.

Für  $p, q \geq 1$  ist  $\dot{F}_{p,q}^\alpha$  ein Banachraum, sonst ist  $\|\cdot\|_{\dot{F}_{p,q}^\alpha}$  für  $0 < p < +\infty$ ,  $0 < q \leq +\infty$ , bzw.  $p = +\infty$ ,  $1 < q \leq +\infty$  nur eine *quasi-Norm*, d.h. die Dreiecksungleichung wird durch eine Ungleichung der Art  $\|f + g\| \leq C(\|f\| + \|g\|)$  mit  $C \geq 1$  ersetzt, vgl. [12], S.240.

Für verschiedene zulässige Funktionen  $\phi$  in der Charakterisierung der Triebel-Lizorkin Räume erhält man somit äquivalente Normen, bzw. quasi-Normen.

Die Familie der homogenen Triebel-Lizorkin Räume enthält unter anderem auch die Räume  $H^1(\mathbb{R}^n) = \dot{F}_{1,2}^0(\mathbb{R}^n)$  und  $BMO(\mathbb{R}^n) = \dot{F}_{\infty,2}^0$ , s. [12], S.244 oder [15], S.70. Allgemein gilt die Dualität:

$$(\dot{F}_{p,q}^\alpha)^* = \dot{F}_{p',q'}^{-\alpha} \quad (2.59)$$

für  $\alpha \in \mathbb{R}$  und  $p, q \in [1, \infty)$ , mit  $1/p + 1/p' = 1$  und  $1/q + 1/q' = 1$ , vgl. [12], S.178 und 239.

**Bemerkung:** Falls man in (2.57) die  $L^p$ - und die  $l^q$ -Normen untereinander vertauscht, erhält man die entsprechenden homogenen Besov-Räume  $\dot{B}_{p,q}^\alpha$ . Alle diese Räume besitzen auch inhomogene Versionen. Für eine Funktion  $\Phi \in \mathcal{S}$  mit  $\text{supp } \hat{\Phi} \subset \{\omega : |\omega| \leq 2\}$  und  $|\hat{\Phi}(\omega)| \geq c > 0$  für  $|\omega| \leq 5/3$  wird der *inhomogene Triebel-Lizorkin Raum*  $F_{p,q}^\alpha$  für  $0 < p < +\infty$  durch die Eigenschaft

$$\|f\|_{F_{p,q}^\alpha} = \|\Phi * f\|_{L^p} + \left\| \left( \sum_{j=1}^{\infty} 2^{j\alpha} |\phi_j * f|^q \right)^{1/q} \right\|_{L^p} < +\infty \quad (2.60)$$

definiert, wobei diesmal  $f \in \mathcal{S}'$  (anstelle von  $f \in \mathcal{S}'/\mathcal{P}$ ). Im Unterschied zu (2.57) wurden hier die niedrigen Frequenzen von  $f$  in einem einzigen Term zusammengefasst. Die inhomogenen Triebel-Lizorkin Räume enthalten als Spezialfälle u.a. die Sobolevräume:  $W^{m,p}(\mathbb{R}^n) = F_{p,2}^m(\mathbb{R}^n)$ , [12], S.51.

Für  $p = +\infty$  lautet die charakterisierende Eigenschaft:

$$\begin{aligned} \|f\|_{F_{\infty,q}^\alpha} := & \\ & \|\Phi * f\|_{L^\infty} + \sup_{l(Q) < 1} \left( \frac{1}{|Q|} \int_Q \sum_{j=-\log_2 l(Q)}^{\infty} (2^{j\alpha} |(\phi_j * f)(x)|)^q dx \right)^{1/q} \\ & < +\infty. \end{aligned} \quad (2.61)$$

Der Raum  $BMO^{-1}(\mathbb{R}^n)$  kann auch als Element der Skala der Triebel-Lizorkin Räume betrachtet werden.

**Satz 1** *Es gilt  $BMO^{-1}(\mathbb{R}^n) = \dot{F}_{\infty,2}^{-1}(\mathbb{R}^n)$ .*

**Beweis.** Die Argumentation geht über die Verwendung des Riesz-Potentials. Dieses wird für  $s \in \mathbb{R}$  durch

$$I_s f := \mathcal{F}^{-1}(|\omega|^{-s} \cdot \hat{f}(\omega))$$

definiert, wobei  $\mathcal{F}^{-1}$  die inverse Fouriertransformierte darstellt und  $f \in \mathcal{S}'$  eine temperierte Distribution ist. Falls  $f \in E$  (Element eines Funktionenraumes  $E$ ), dann wird das Bild  $I_s(E)$  als Raum der Äquivalenzklassen modulo Polynome betrachtet, vgl.[13], S.549 oder [12], S.242.

Laut [12], S.242 gilt:  $I_s(\dot{F}_{p,q}^\alpha) = \dot{F}_{p,q}^{\alpha+s}$ . Es folgt also:

$$\dot{F}_{\infty,2}^{-1}(\mathbb{R}^n) = I_{-1}(\dot{F}_{\infty,2}^0(\mathbb{R}^n)) = I_{-1}(BMO(\mathbb{R}^n)).$$

Schließlich ist  $I_{-1}(BMO(\mathbb{R}^n)) = BMO^{-1}(\mathbb{R}^n)$ , vgl. [13], S.549. □

Die Arbeit [15] behandelt die Darstellung der Elemente der Triebel-Lizorkin Räume  $F_{p,q}^\alpha$  in der Form

$$f = \sum_Q \langle f, \phi_Q \rangle \psi_Q \tag{2.62}$$

wobei  $Q$  die Menge aller dyadischen Würfel durchläuft.  $\phi_Q$  und  $\psi_Q$  sind Translationen und Dilatationen von Funktionen  $\phi, \psi$ , welche neben (2.55),(2.56) auch die Eigenschaft

$$\sum_{j \in \mathbb{Z}} \widehat{\phi(2^j \omega)} \hat{\psi}(2^j \omega) = 1/2\pi \tag{2.63}$$

erfüllen. Diese zusätzliche Eigenschaft ermöglicht die Darstellung einer Distribution  $f \in \mathcal{S}'/\mathcal{P}$  als  $f = \sum_Q c_Q \psi_Q$  mit  $c_Q = \langle f, \phi_Q \rangle$ . Die Eindeutigkeit der Koeffizienten  $c_Q$  ist jedoch nicht gewährleistet. Die Konvergenz erfolgt im Sinne der Dualität. Sei  $\mathcal{S}_0 = \{u \in \mathcal{S} : \int x^\alpha u(x) dx = 0 \text{ für alle } \alpha \in \mathbb{N}\}$  der Raum der Schwartzfunktionen, für welche alle Momente verschwinden.

Es gilt  $\sum_Q c_Q \psi_Q = f \in \mathcal{S}'/\mathcal{P}$  (welches man mit dem Dualraum  $\mathcal{S}'_0$  identifizieren kann, vgl. [16]) genau dann, wenn für jede Testfunktion  $u \in \mathcal{S}_0$  die Identität  $\langle f, u \rangle = \sum_Q c_Q \langle \psi_Q, u \rangle$  gilt. Die Funktionenfamilie  $\{\phi_Q, \psi_Q\}$  ist 'fast orthogonal', im Sinne dass  $\langle \phi_Q, \psi_{Q'} \rangle = 0$  falls  $Q, Q'$  sich nicht auf gleichen (oder benachbarten) Skalen befinden, und sonst klein im Betrag ist, falls  $Q$  und  $Q'$  weit weg voneinander liegen.

Eines der wichtigsten Ergebnisse aus [15] ist die Charakterisierung der Triebel-Lizorkin Räume durch passende Normen der Koeffizientenfolgen. Für  $\alpha \in \mathbb{R}$ ,  $0 < p < +\infty$  und  $0 < q \leq +\infty$  definiert man den Raum  $\dot{f}_{p,q}^\alpha$  als die Menge der Folgen  $\{c_Q\}_Q$  indiziert nach allen dyadischen Würfeln  $Q$ , für welche

$$\|c\|_{\dot{f}_{p,q}^\alpha} := \left\| \left( \sum_Q \left( |Q|^{-\frac{\alpha}{n}} |c_Q| \tilde{\chi}_Q(\cdot) \right)^q \right)^{1/q} \right\|_{L^p} < +\infty \quad (2.64)$$

gilt, mit der üblichen Modifikation für  $q = \infty$ . Mit  $\tilde{\chi}_Q = |Q|^{-1/2} \chi_Q$  wird hier die  $L^2$ -normierte charakteristische Funktion des dyadischen Würfels  $Q$  bezeichnet.

Für  $p = +\infty$ ,  $q < +\infty$  definiert man den Raum  $\dot{f}_{\infty,q}^\alpha$  durch die Eigenschaft

$$\|c\|_{\dot{f}_{\infty,q}^\alpha} := \sup_Q \left( \frac{1}{|Q|} \int_Q \sum_{Q' \subset Q} \left( |Q'|^{-\frac{\alpha}{n}} |c_{Q'}| \tilde{\chi}_{Q'}(x) \right)^q dx \right)^{1/q} < +\infty \quad (2.65)$$

wobei  $Q'$  die Menge aller dyadischen Subwürfel von  $Q$  durchläuft. Eine Integration in 2.65 liefert

$$\|c\|_{\dot{f}_{\infty,q}^\alpha} = \sup_Q \left( \frac{1}{|Q|} \sum_{Q' \subset Q} \left( |Q'|^{-\frac{\alpha}{n} - \frac{1}{2} + \frac{1}{q}} |c_{Q'}| \right)^q \right)^{1/q}. \quad (2.66)$$

Laut [15] gilt folgende Charakterisierung der Triebel-Lizorkin Räume:

**Theorem 6** *Seien  $\phi, \psi$  zwei Funktionen mit den Eigenschaften (2.55), (2.56) und sei  $Q = Q_{j,k}$  der dyadische Würfel mit Kantenlänge  $2^{-j}$  und linke untere Ecke in  $2^{-j}k$ , mit  $k \in \mathbb{Z}$ . Man definiert  $\phi_Q(x) = 2^{nj/2} \phi(2^j x - k)$  und  $\psi_Q = 2^{nj/2} \psi(2^j x - k)$ . Für alle  $f \in \mathcal{S}'/\mathcal{P}$  gelte die Darstellung  $f = \sum_Q \langle f, \phi_Q \rangle \psi_Q$ , dies ist insbesondere der Fall wenn (2.63) gilt. Dann ist eine Distribution  $f \in \mathcal{S}'/\mathcal{P}$  genau dann in  $\dot{F}_{p,q}^\alpha$ , wenn für  $c_Q = \langle f, \phi_Q \rangle$  die Eigenschaft  $c = \{c_Q\}_Q \in \dot{f}_{p,q}^\alpha$  erfüllt ist. Es gilt die Normäquivalenz  $\|f\|_{\dot{F}_{p,q}^\alpha} \sim \|c\|_{\dot{f}_{p,q}^\alpha}$ .*

Für dieses Resultat wird unter anderem auch folgendes Lemma benötigt, welches in [15] bewiesen wird.

**Lemma 3** *Für eine Funktion  $\phi \in \mathcal{S}$  welche (2.55) und (2.56) erfüllt sei  $\tilde{\phi}(x) := \phi(-x)$ . Zu  $f \in \mathcal{S}'/\mathcal{P}$  konstruiert man die Folge  $\sup(f) =$*

$\{sup_Q(f)\}_Q$  indiziert nach dyadischen Würfeln mit Kantenlänge  $2^{-j}$  durch:

$$sup_Q(f) = 2^{-nj/2} \sup_{y \in Q} |\tilde{\phi}_j * f(y)|.$$

Dann gilt  $\|f\|_{\dot{F}_{p,q}^\alpha} \sim \|\sup(f)\|_{\dot{F}_{p,q}^\alpha}$ .

Die Arbeit [15] entstand fast gleichzeitig mit der Entwicklung der orthogonalen Waveletbasen durch Lemarié und Meyer, vgl. [16]. In einer Raumdimension kann man speziell  $\phi = \psi$  gleich einem Wavelet von Meyer-Typ wie in 2.1.1, Beispiel (iii) wählen. Dann sind die Eigenschaften (2.55), (2.56) und (2.63) erfüllt und die Wavelets bilden zudem auch eine Orthonormalbasis von  $L^2(\mathbb{R})$ . In einer Raumdimension kann man also das Ergebnis aus Theorem 6 direkt als Waveletdarstellung interpretieren.

Ausgehend von den eindimensionalen Meyer-Wavelets kann man in höheren Dimensionen Tensorprodukt-Wavelets wie in Sektion 2.1.4 konstruieren, welche eine Orthonormalbasis von  $L^2(\mathbb{R}^n)$  bilden, vgl. [16]. Einem dyadischen Würfel entspricht jedoch nicht nur eine Funktion  $\psi_Q$ , sondern mehrere Wavelets  $\psi_Q^\varepsilon$ . Für eine Distribution  $f \in \mathcal{S}'/\mathcal{P}$  gilt die Darstellung  $f = \sum_\lambda \langle f, \psi_\lambda \rangle \psi_\lambda$ , mit  $\lambda = (Q, \varepsilon)$  und  $\psi_\lambda = \psi_Q^\varepsilon$ . Die Konvergenz erfolgt in  $\mathcal{S}'/\mathcal{P}$ , vgl. [16]. Dann gilt folgendes Resultat:

**Theorem 7** Eine Distribution  $f \in \mathcal{S}'/\mathcal{P}$  mit der Waveletdarstellung  $f = \sum_\lambda c_\lambda \psi_\lambda$  liegt genau dann in  $\dot{F}_{p,q}^\alpha$ , wenn

$$\left\| \left( \sum_\lambda \left( |Q(\lambda)|^{-\frac{\alpha}{n}} |c_\lambda| |\tilde{\chi}_{Q(\lambda)}(\cdot)|^q \right)^{1/q} \right) \right\|_{L^p} < +\infty \quad (2.67)$$

für  $p < +\infty$ , mit der üblichen Modifikation falls  $q = +\infty$ , bzw.

$$\sup_Q \left( \frac{1}{|Q|} \sum_{Q(\lambda) \subset Q} \left( |Q(\lambda)|^{-\frac{\alpha}{n} - \frac{1}{2} + \frac{1}{q}} |c_\lambda| \right)^q \right)^{1/q} < +\infty \quad (2.68)$$

falls  $p = +\infty, q < +\infty$ . Das Supremum in (2.68) wird über allen dyadischen Würfeln  $Q$  gebildet und die Summe läuft über alle dyadische Subwürfel von  $Q$ . Die Ausdrücke in (2.67) bzw. (2.68) sind vergleichbar zu  $\|f\|_{\dot{F}_{p,q}^\alpha}$ .

**Beweis.**

“ $\Rightarrow$ ” Sei zunächst  $f \in \dot{F}_{p,q}^\alpha$ . Gemäß deren Konstruktion erfüllen die Wavelets  $\psi^\varepsilon$  für jedes  $\varepsilon$  die Bedingungen (2.55) und (2.56), vgl. [16]. Das gleiche gilt aber auch für  $\tilde{\psi}^\varepsilon(x) = \psi^\varepsilon(-x)$ . Somit gilt nach Definition

$$\|f\|_{\dot{F}_{p,q}^\alpha} \sim \left\| \left( \sum_{j \in \mathbb{Z}} (2^{j\alpha} |\tilde{\psi}_j^\varepsilon * f|)^q \right)^{1/q} \right\|_{L^p} < +\infty$$

für  $p < +\infty$ , bzw.

$$\|f\|_{\dot{F}_{\infty,q}^\alpha} \sim \left( \sup_Q \frac{1}{|Q|} \int_Q \sum_{j=-\log_2 l(Q)}^{\infty} (2^{j\alpha} |\tilde{\psi}_j^\varepsilon * f|(x))^q dx \right)^{1/q} < +\infty$$

wobei  $\tilde{\psi}_j^\varepsilon(x) = 2^{nj} \psi^\varepsilon(-2^j x)$ . Es gilt weiterhin:

$$|c_Q^\varepsilon| = |\langle f, \psi_Q^\varepsilon \rangle| = 2^{-nj/2} |\tilde{\psi}_j^\varepsilon * f(2^{-j} k)| \leq 2^{-nj/2} \sup_{y \in Q} |\tilde{\psi}_j^\varepsilon * f(y)| =: \text{sup}_Q(f)^\varepsilon,$$

wobei nach Lemma 3:

$$\|\text{sup}(f)^\varepsilon\|_{\dot{f}_{p,q}^\alpha} \sim \|f\|_{\dot{F}_{p,q}^\alpha} < +\infty.$$

Folglich gilt  $\|\{c_Q^\varepsilon\}_Q\|_{\dot{f}_{p,q}^\alpha} \leq \|\text{sup}(f)^\varepsilon\|_{\dot{f}_{p,q}^\alpha} < +\infty$ . Unter Verwendung der Ungleichung  $(\sum_\varepsilon |a_\varepsilon|)^\beta \leq C_{n,\beta} \sum_\varepsilon |a_\varepsilon|^\beta$  für  $\beta > 0$  und der Dreiecksungleichung bezüglich der  $L^p$ -Norm ergibt sich:

$$\begin{aligned} & \left\| \left( \sum_\varepsilon \sum_Q (|Q|^{-\frac{\alpha}{n}} |c_Q^\varepsilon| \tilde{\chi}_Q(\cdot))^q \right)^{1/q} \right\|_{L^p} \leq \\ & \leq C_{n,q} \sum_\varepsilon \left\| \left( \sum_Q (|Q|^{-\frac{\alpha}{n}} |c_Q^\varepsilon| \tilde{\chi}_Q(\cdot))^q \right)^{1/q} \right\|_{L^p} \\ & = C_{n,q} \sum_\varepsilon \|\{c_Q^\varepsilon\}_Q\|_{\dot{f}_{p,q}^\alpha} \leq C_{n,q} \sum_\varepsilon \|\text{sup}(f)^\varepsilon\|_{\dot{f}_{p,q}^\alpha} \sim \|f\|_{\dot{F}_{p,q}^\alpha} < +\infty. \end{aligned}$$

Falls  $p = +\infty$  gilt analog:

$$\begin{aligned} & \sup_Q \left( \frac{1}{|Q|} \sum_{\varepsilon} \sum_{Q' \subset Q} \left( |Q'|^{-\frac{\alpha}{n} - \frac{1}{2} + \frac{1}{q}} |c_{Q'}^{\varepsilon}| \right)^q \right)^{1/q} \leq \\ & \leq C_{n,q} \sum_{\varepsilon} \sup_Q \left( \frac{1}{|Q|} \sum_{Q' \subset Q} \left( |Q'|^{-\frac{\alpha}{n} - \frac{1}{2} + \frac{1}{q}} |c_{Q'}^{\varepsilon}| \right)^q \right)^{1/q} \\ & = C_{n,q} \sum_{\varepsilon} \|\{c_Q^{\varepsilon}\}_Q\|_{\dot{f}_{\infty,q}^{\alpha}} \leq C_{n,q} \sum_{\varepsilon} \|\sup(f)^{\varepsilon}\|_{\dot{f}_{\infty,q}^{\alpha}} \sim \|f\|_{\dot{F}_{\infty,q}^{\alpha}} < +\infty. \end{aligned}$$

Der erste Teil des Satzes ist somit bewiesen.

"  $\Leftarrow$  " Sei  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$  fest, mit  $\varepsilon_i \in \{0, 1\}$ ,  $\varepsilon \neq 0$  und betrachte die Darstellungen

$$f^{\varepsilon} = \sum_Q \langle f, \psi_Q^{\varepsilon} \rangle \psi_Q^{\varepsilon} = \sum_Q c_Q^{\varepsilon} \psi_Q^{\varepsilon} \quad (2.69)$$

mit Konvergenz in  $\mathcal{S}'/\mathcal{P}$ . Es gilt offensichtlich  $f = \sum_{\varepsilon} f^{\varepsilon}$ . Falls (2.67) bzw. (2.68) gelten, folgt sofort dass  $\|\{c_Q^{\varepsilon}\}_Q\|_{\dot{f}_{p,q}^{\alpha}} < +\infty$ . Nach Theorem 6 gilt daher  $f^{\varepsilon} \in \dot{F}_{p,q}^{\alpha}$  für alle  $\varepsilon$ , also auch  $f = \sum_{\varepsilon} f^{\varepsilon} \in \dot{F}_{p,q}^{\alpha}$ . Es gilt weiterhin:

$$\|f\|_{\dot{F}_{p,q}^{\alpha}} \leq C \sum_{\varepsilon} \|f^{\varepsilon}\|_{\dot{F}_{p,q}^{\alpha}} \sim \sum_{\varepsilon} \|\{c_Q^{\varepsilon}\}_Q\|_{\dot{f}_{p,q}^{\alpha}} \leq (2^n - 1) \| \|(c_{\lambda})_{\lambda} \|$$

wobei  $\| \|(c_{\lambda})_{\lambda} \|$  durch (2.67) bzw. (2.68) gegeben ist, je nachdem ob  $p < +\infty$  oder  $p = +\infty$ . □

Als Folgerung erhalten wir die gleichen Charakterisierungen der Hardy-Räume  $H^p$  und des Raumes  $BMO$  (alle sind Spezialfälle der Triebel-Lizorkin Räume), welche in [2], Kapitel 5 und 6, bewiesen sind (im Fall von  $BMO$  auch in dieser Arbeit). Theorem 7 liefert zusätzlich auch eine ähnliche Charakterisierung des Raumes  $BMO^{-1}(\mathbb{R}^n) = \dot{F}_{\infty,2}^{-1}(\mathbb{R}^n)$ :

**Korollar 2** *Eine Distribution  $f \in \mathcal{S}'$  mit der Waveletdarstellung  $f = \sum_{\lambda} c_{\lambda} \psi_{\lambda}$  liegt genau dann in  $BMO^{-1}(\mathbb{R}^n)$ , wenn die Waveletkoeffizienten  $(c_{\lambda})_{\lambda}$  die Ungleichung*

$$\sup_Q \left( \frac{1}{|Q|} \sum_{Q(\lambda) \subset Q} |Q(\lambda)|^{\frac{2}{n}} |c_{\lambda}|^2 \right)^{1/2} < +\infty \quad (2.70)$$

erfüllen. Der Ausdruck (2.70) ist vergleichbar mit der  $BMO^{-1}$ -Norm von  $f$ .

**Bemerkung 1:** Im Fall des Raumes  $BMO^{-1}$  kann man die Forderung  $f \in \mathcal{S}'/\mathcal{P}$  durch  $f \in \mathcal{S}'$  ersetzen, weil die Äquivalenzklassen modulo Polynome hier nur aus einem einzigen Element bestehen. Dies folgt aus der Tatsache, dass die  $BMO$ -Funktionen Äquivalenzklassen modulo Konstanten bilden, also für der Raum der Ableitungen,  $BMO^{-1}$ , eine ähnliche Quotientenbildung nicht erforderlich ist.

**Bemerkung 2:** Eine andere Möglichkeit, eine Darstellung der Form (2.62) zu realisieren, wäre die Benutzung von *biorthogonalen Wavelets*. Im Unterschied zu der üblichen Waveletdarstellungen sind hier die Waveletfunktionen  $\phi$  und  $\psi$  verschieden, aber symmetrisch. Die Orthogonalität der Waveletfamilie bedeutet in diesem Fall die Beziehung  $\langle \phi_Q, \psi_{Q'} \rangle = \delta_{Q, Q'}$  für alle dyadischen Würfel  $Q, Q'$ .

### 2.2.6 Anwendungen des Raumes $BMO^{-1}$

Im folgenden werden einige Eigenschaften des Raumes  $BMO^{-1}$  präsentiert, deren Motivation aus der Bildverarbeitung stammt. Hier ist man bestrebt, ein Signal (Bild)  $f$  als eine Summe  $u + v$  darzustellen, wobei  $u$  die Objekte aus dem Bild modellieren soll, während  $v$  für die Texturen und das Rauschen steht. Dies wird realisiert, indem die Norm  $\|u\|_{\mathcal{B}}$  in einem geeigneten Banachraum  $\mathcal{B}$  beschränkt ist, während eine andere Norm  $\|v\|_*$  "klein" sein soll, um den stark oszillierenden Charakter der  $v$ -Komponente auszudrücken. Eine Möglichkeit einer solchen Zerlegung ist z.B. im *Osher-Rudin* Modell dargestellt, vgl. [3], Kap.1.13. Hier ist  $\mathcal{B} = BV$  (Funktionen mit beschränkter Variation) und der Raum, welcher die oszillierende Komponente modellieren soll, wird analog zu  $BMO^{-1}(\mathbb{R}^2)$  definiert, wobei  $L^\infty(\mathbb{R}^2)$  anstelle von  $BMO(\mathbb{R}^2)$  genommen wird.

Die nächsten Ergebnisse zeigen dass die auf  $BMO^{-1}$  basierende Zerlegung sich ähnlich zu dem Osher-Rudin Modell verhält, dessen Eigenschaften in [3], Kap.1.13-1.15 dargestellt sind.

Zunächst zur Auswahl des Banachraumes  $\mathcal{B}$ , welcher die Strukturen (Objekte) aus dem Bild modellieren soll. In Hinsicht auf die Verwendung von Dualitätsargumenten wählt man hier  $\mathcal{B} = I_1(H^1(\mathbb{R}^n)) = \dot{F}_{1,2}^1(\mathbb{R}^n)$ .  $\mathcal{B}$  ist also das Bild des Hardyraumes  $H^1(\mathbb{R}^n)$  unter dem Riesz-Potential  $I_1$ . Es gilt auch die Charakterisierung:  $\mathcal{B} = \{f \in L_{loc}^1(\mathbb{R}^n) \mid \partial_i f \in H^1(\mathbb{R}^n), i = \overline{1, n}\}$ , vgl.[14].

(2.59) impliziert  $(\dot{F}_{1,2}^1(\mathbb{R}^n))^* = \dot{F}_{\infty,2}^{-1}(\mathbb{R}^n) = BMO^{-1}(\mathbb{R}^n)$ . Es gilt also die Dualität  $\mathcal{B}^* = BMO^{-1}(\mathbb{R}^n)$ . Angesichts dieser Eigenschaft

wählen wir auf  $\mathcal{B}$  eine Norm  $\|\cdot\|_{\mathcal{B}}$  so dass gilt:

$$\left| \int_{\mathbb{R}^n} f(x)g(x)dx \right| \leq \|f\|_{\mathcal{B}} \cdot \|g\|_{BMO^{-1}} \quad (2.71)$$

für alle  $f \in \mathcal{B}$ ,  $g \in BMO^{-1}(\mathbb{R}^n)$ .

Laut [12], S.129 und der Bemerkung auf S.238, falls  $s_0 - n/p_0 = s_1 - n/p_1$ , gilt die stetige Einbettung:  $\dot{F}_{p_0,q}^{s_0}(\mathbb{R}^n) \subset \dot{F}_{p_1,r}^{s_1}(\mathbb{R}^n)$ . Für  $n = 2$  gilt also:

$$\mathcal{B} = \dot{F}_{1,2}^1(\mathbb{R}^2) \subset \dot{F}_{2,2}^0 = L^2(\mathbb{R}^2) \subset \dot{F}_{\infty,2}^{-1}(\mathbb{R}^2) = BMO^{-1}(\mathbb{R}^2).$$

Für einen Parameter  $\lambda > 0$  wählen wir diejenige Zerlegung  $f = u_\lambda + v_\lambda$  einer Funktion  $f \in L^2(\mathbb{R}^2)$ , welche folgendes Minimierungsproblem löst:

$$\inf_{f=u+v} \{ \|u\|_{\mathcal{B}} + \lambda \|v\|_2^2 \}.$$

Zu  $f \in L^2(\mathbb{R}^2)$  und  $\lambda > 0$  kann man daher das Funktional

$$J_{f,\lambda}(u) := \|u\|_{\mathcal{B}} + \lambda \|f - u\|_2^2$$

auf  $\mathcal{B}$  assoziieren. Wegen der Struktur gegeben durch die zwei Normen ist das Funktional  $J_{f,\lambda}$  konvex und schwach unterhalbstetig auf  $\mathcal{B}$ . Wegen der eingebauten  $L^2$ -Norm ist es sogar strikt konvex, so dass ein eindeutiges Minimum existiert. Das entsprechende Paar  $(u_\lambda, v_\lambda)$  mit  $v_\lambda = f - u_\lambda$  wird auch als *optimale Zerlegung* von  $f$  bezeichnet.

Der Parameter  $\lambda$  beschreibt die Größe der Details, die im Bild erscheinen: die kleinsten Objekte, welche detektiert werden können, haben die Größenordnung  $O(1/\lambda)$ . Dies wird auch durch folgendes Resultat erläutert:

**Lemma 4** *Sei  $f \in L^2(\mathbb{R}^2)$ . Es gilt  $\|f\|_{BMO^{-1}} \leq 1/2\lambda$  genau dann, wenn die optimale Zerlegung von  $f$  durch  $u_\lambda = 0, v_\lambda = f$  gegeben ist.*

**Beweis.** Es gilt  $u_\lambda = 0, v_\lambda = f$  genau dann, wenn

$$\lambda \|f\|_2^2 \leq \|w\|_{\mathcal{B}} + \lambda \|f - w\|_2^2$$

für alle  $w \in \mathcal{B}$ . Wenn man  $\varepsilon w$  für  $w$  einsetzt, mit  $\varepsilon > 0$ , ist dies äquivalent zu:

$$\lambda \|f\|_2^2 \leq \varepsilon \|w\|_{\mathcal{B}} + \lambda \|f\|_2^2 - 2\lambda \varepsilon \langle f, w \rangle + \lambda \varepsilon^2 \|w\|_2^2.$$

Dies ist gleichbedeutend mit

$$\langle f, w \rangle \leq \frac{1}{2\lambda} \|w\|_{\mathcal{B}} + \frac{\varepsilon}{2} \|w\|_2^2.$$

Mit  $\varepsilon \rightarrow 0$  erhält man also für alle  $w \in \mathcal{B}$ :

$$\langle f, w \rangle \leq \frac{1}{2\lambda} \|w\|_{\mathcal{B}}.$$

Für  $\varepsilon < 0$  erhält man analog

$$\langle f, w \rangle \geq -\frac{1}{2\lambda} \|w\|_{\mathcal{B}}$$

also insgesamt

$$|\langle f, w \rangle| \leq \frac{1}{2\lambda} \|w\|_{\mathcal{B}}.$$

Angesichts der Dualität (2.71) ist dies äquivalent zu  $\|f\|_{BMO^{-1}} \leq 1/2\lambda$ .

□

Dieses Lemma besagt nichts anderes als die Tatsache dass die  $L^2$ -Funktionen mit  $BMO^{-1}$ -Norm kleiner als  $1/2\lambda$  als Rauschen oder Texturen interpretiert werden. Falls die  $BMO^{-1}$ -Norm des Signals diesen kritischen Wert überschreitet, dann erläutert das folgende Theorem die Eigenschaften der optimalen Zerlegung.

**Theorem 8** *Seien  $f, u, v \in L^2(\mathbb{R}^2)$ . Falls  $\|f\|_{BMO^{-1}} > 1/2\lambda$  gilt, so ist  $(u, v)$  die optimale Zerlegung von  $f$  genau dann, wenn*

$$\|v\|_{BMO^{-1}} = \frac{1}{2\lambda} \text{ und } \langle u, v \rangle = \frac{1}{2\lambda} \|u\|_{\mathcal{B}}.$$

**Beweis.** " $\Rightarrow$ " Falls  $(u, v)$  die optimale Zerlegung von  $f$  ist, so gilt für alle  $w \in \mathcal{B}$  und  $\varepsilon \in \mathbb{R}$ :

$$\|u\|_{\mathcal{B}} + \lambda \|v\|_2^2 \leq \|u + \varepsilon w\|_{\mathcal{B}} + \lambda \|v - \varepsilon w\|_2^2 \leq \|u\|_{\mathcal{B}} + |\varepsilon| \|w\|_{\mathcal{B}} + \lambda \|v - \varepsilon w\|_2^2.$$

Dies impliziert also

$$\lambda \|v\|_2^2 \leq |\varepsilon| \|w\|_{\mathcal{B}} + \lambda \|v - \varepsilon w\|_2^2$$

und analog zu dem Beweis von Lemma 1 erhält man

$$\|v\|_{BMO^{-1}} \leq \frac{1}{2\lambda}. \tag{2.72}$$

Mit  $w = u$  und  $|\varepsilon| < 1$  folgt:

$$\begin{aligned} \|u\|_{\mathcal{B}} + \lambda \|v\|_2^2 &\leq \|u + \varepsilon u\|_{\mathcal{B}} + \lambda \|v - \varepsilon u\|_2^2 \\ &= (1 + \varepsilon) \|u\|_{\mathcal{B}} + \lambda \|v\|_2^2 - 2\lambda \varepsilon \langle u, v \rangle + \lambda \varepsilon^2 \|u\|_2^2, \end{aligned}$$

also

$$2\lambda\varepsilon\langle u, v \rangle \leq \varepsilon\|u\|_{\mathcal{B}} + \lambda\varepsilon^2\|u\|_2^2.$$

Für  $\varepsilon \rightarrow 0, \varepsilon > 0$  erhält man

$$\langle u, v \rangle \leq \frac{1}{2\lambda}\|u\|_{\mathcal{B}}$$

während  $\varepsilon \rightarrow 0, \varepsilon < 0$

$$\langle u, v \rangle \geq \frac{1}{2\lambda}\|u\|_{\mathcal{B}}$$

impliziert. Folglich ist

$$\langle u, v \rangle = \frac{1}{2\lambda}\|u\|_{\mathcal{B}}. \quad (2.73)$$

Angesichts (2.72) folgt dass  $\|v\|_{BMO^{-1}} = 1/2\lambda$ , sonst könnte in (2.73) keine Gleichheit gelten.

"  $\Leftarrow$  " Da  $\|v\|_{BMO^{-1}} = 1/2\lambda$ , gilt laut (2.71):

$$\langle u + \varepsilon h, v \rangle \leq \|u + \varepsilon h\|_{\mathcal{B}} \cdot \|v\|_{BMO^{-1}} = \frac{1}{2\lambda}\|u + \varepsilon h\|_{\mathcal{B}}.$$

Für eine andere Zerlegung  $f = (u + \varepsilon h) + (v - \varepsilon h)$  gilt also:

$$\begin{aligned} \|u + \varepsilon h\|_{\mathcal{B}} + \lambda\|v - \varepsilon h\|_2^2 &= \|u + \varepsilon h\|_{\mathcal{B}} + \lambda\|v\|_2^2 - 2\lambda\varepsilon\langle v, h \rangle + \lambda\varepsilon^2\|h\|_2^2 \\ &\geq 2\lambda\langle u + \varepsilon h, v \rangle + \lambda\|v\|_2^2 - 2\lambda\varepsilon\langle v, h \rangle + \lambda\varepsilon^2\|h\|_2^2 \\ &= 2\lambda\langle u, v \rangle + \lambda\|v\|_2^2 + \lambda\varepsilon^2\|h\|_2^2 \\ &= \|u\|_{\mathcal{B}} + \lambda\|v\|_2^2 + \lambda\varepsilon^2\|h\|_2^2 \\ &\geq \|u\|_{\mathcal{B}} + \lambda\|v\|_2^2 \end{aligned}$$

wobei im vorletzten Schritt die Gleichung  $\langle u, v \rangle = \|u\|_{\mathcal{B}}/2\lambda$  aus der Voraussetzung benutzt wurde. Somit bildet  $(u, v)$  die optimale Zerlegung von  $f$ . □

Angesichts der Dualität (2.71) folgt mit Theorem 8, dass die Funktion  $u$  aus der optimalen Zerlegung  $f = u + v$  ein Element aus  $\mathcal{B}$  ist, wo die Linearform  $w \rightarrow \int w(x)v(x)dx$  ihr Maximum annimmt.

Die Verwendung des Raumes  $BMO^{-1}$  oder des Osher-Rudin Modells (basierend auf dem Raum  $BV$ ), für die Beschreibung von Texturen und dessen Anwendung in der Bildverarbeitung hat auch einige Nachteile. Wird zum Beispiel ein Bild, dargestellt als  $f(x) = \chi_{\Omega}(x)$ , gewählt, wobei  $\Omega$  ein beschränktes Gebiet mit glattem Rand ist, dann wird es nach der optimalen Zerlegung nicht beibehalten, obwohl kein

erkennbares "Rauschen" vorhanden ist. Laut Theorem 8 muß die  $v$ -Komponente die Norm  $\|v\|_{BMO^{-1}} = 1/2\lambda$  haben, welches nicht unbedingt klein ist. Der Wunsch, dass eine solche Zerlegung immer nur das tatsächliche Rauschen eliminiert, während es die Strukturen des Bildes beibehält, wird nicht immer realisiert. Neben dem Rauschen werden manchmal auch echte Details weggefiltert. Da wir jedoch nicht an diesem Anwendungsgebiet des Raumes  $BMO^{-1}$  interessiert sind, verweisen wir für weitere Details zu diesem Thema auf [3].



## Kapitel 3

# Algorithmus zur Punktereduktion und Flächenrekonstruktion

In diesem Kapitel wird eine Methode vorgestellt, die eine Lösung zu folgendem praktischen Problem liefert: ausgehend von einer großen Anzahl diskreter Punkte, die auf einer (unbekannten) Fläche liegen, soll eine analytische Darstellung dieser Fläche gegeben werden, welche auf möglichst wenigen Parametern (Knoten, Koeffizienten) basiert. Die große Anzahl dieser Messpunkte verhindert zunächst die Anwendung von Interpolations- oder Approximationsmethoden. Daher muß zunächst eine *Punktereduktion* erfolgen. Ein Kriterium, an welchen Stellen die Punkte mehr ausgedünnt werden können und an welchen nicht, liefert die Größe der  $BMO^{-1}$ -Koeffizienten:

$$b_x = \left( \sup_{Q \ni x} \frac{1}{|Q|} \sum_{Q(\lambda) \subset Q} |Q(\lambda)| \cdot |c_\lambda|^2 \right)^{1/2} \quad (3.1)$$

Diese werden für eine erste Interpolation, basierend auf allen Punkten berechnet. Laut der Ergebnisse des vorherigen Kapitels sind diese Koeffizienten groß dort, wo “Strukturen” auftauchen, und klein an den “flachen” Stellen, auch wenn hier durch mögliche Rauheiten eventuell große Gradienten entstehen können. Dies ist z.B. der Fall, wenn die

vorhandenen Punktedaten verrauscht sind und die erste Interpolante der Fläche somit schnelle Oszillationen von kleiner Amplitude aufweist.

Im Unterschied zur Bildverarbeitung, sind die Funktionen, die in unserer Problematik auftauchen, in der Regel stetig. Unser Ziel ist nicht, die “Strukturen” von dem “Rauschen” zu separieren, sondern die Fläche in Bereiche aufzuteilen, gemäss der Größenordnung der  $BMO^{-1}$ -Koeffizienten, wo dann die gegebenen Punkte mehr oder weniger ausgedünnt werden. In diesem Kapitel wird das dazu benutzte Verfahren ausführlich beschrieben.

Nach der Punktereduktion erfolgt im nächsten Schritt die Rekonstruktion der Fläche. Es muss das Problem der “Flächenanpassung” gelöst werden, d.h. die Bestimmung einer Spline-Fläche, welche die reduzierten Punkte gut approximiert und auch gewisse Anforderungen an die Glattheit erfüllt. Der Begriff “gute Approximation” ist in dieser Situation schwierig zu definieren. Eine zu hohe Genauigkeit in der Annäherung an die einzelnen Punkte kann die Spline-Funktion zu sehr starken Oszillationen zwischen den Punkten zwingen. Diese können durch die Anforderung an die Glattheit ausbalanciert werden. Eine zu starke Glättung würde jedoch auf Kosten mancher Details des Objektes gehen, die durch die darstellende Fläche nicht mehr richtig reproduziert werden. Den besten Kompromiss zwischen diesen Aspekten erreicht man am schnellsten auf heuristische Weise. In Kapitel 3.5 wird erläutert, wie man dies mit minimalem Rechenaufwand erreichen kann.

### 3.1 Erste Interpolation, Berechnung und thresholding der $BMO^{-1}$ -Koeffizienten

Die Berechnung der Wavelet- und implizit auch der  $BMO^{-1}$ -Koeffizienten setzt voraus, dass wir eine Fläche vom Typ einer Funktion  $z = f(x, y)$  haben. Parametrisch definierte Flächen werden hier nicht betrachtet. Die gegebenen Daten liegen zunächst nur als eine diskrete, unstrukturierte Punktemenge vor. Um eine 2-dimensionale diskrete Waveletzerlegung durchführen zu können, ist es erforderlich, dass die  $(x, y)$ -Koordinaten der Punkte auf einem Gitter strukturiert sind. Als Erstes muss also die Fläche anhand der gegebenen Punkte interpoliert werden und in den Punkten eines Gitters ausgewertet werden. Eine

Richtung des Gitters kann gleich der Richtung der gemessenen Linien angenommen werden und wir nehmen an, dass das gemessene Objekt auf der Ebene  $z = 0$  liegt. Die Auswertung in den Gitterpunkten erfolgt durch lineare Interpolation.

### 3.1.1 Lineare Interpolation einer gegebenen Punktemenge

Gegeben sei die Punktemenge  $\mathcal{P} = \{P_i(x_i, y_i, z_i) : i = 1, \dots, n\} \subset \mathbb{R}^3$ , wobei keine zwei Punkte mit identischen  $(x, y)$ -Koordinaten vorkommen. Gesucht ist die Auswertung einer linearen Interpolante  $Z(i_k, j_l)$  in den Knoten  $(i_k, j_l)$  eines gegebenen Gitters, wobei  $(i_k)_{k=0}^{n_1}$  eine äquidistante Unterteilung eines Intervalls  $[a, b]$  und  $(j_l)_{l=0}^{n_2}$  eine äquidistante Unterteilung eines Intervalls  $[c, d]$  ist. Es gelte zusätzlich die Bedingung:

$$\text{conv}(\{(x_i, y_i) : i = 1, \dots, n\}) \subset [a, b] \times [c, d],$$

wobei  $\text{conv}(A)$  die konvexe Hülle der Menge  $A$  darstellt.

Die lineare Interpolation basiert auf der Delaunay-Triangulation  $\mathcal{T}$  der Menge  $\mathcal{P}_{xy} = \{(x_i, y_i) : i = 1, \dots, n\} \subset \mathbb{R}^2$ . Für die Punkte des Gitters außerhalb von  $\text{conv}(\mathcal{P}_{xy})$  setzen wir  $Z(i_k, j_l) = 0$ . Sei nun  $P = (i_k, j_l)$  ein Punkt des Gitters und  $\Delta = (ABC)$  das entsprechende Dreieck aus der Triangulation  $\mathcal{T}$  mit  $P \in \Delta$ . Geometrisch erhält man den gesuchten  $z$ -Wert, indem die Senkrechte auf der  $(xy)$ -Ebene im Punkt  $P$  die Ebene  $(P_A P_B P_C)$  schneidet, wie in Abbildung 3.1 illustriert wird.

Rechnerisch erfolgt die Bestimmung des  $z$ -Wertes mit Hilfe der baryzentrischen Koordinaten. Die *baryzentrischen Koordinaten* (z.B. [6], S.2-3) des Punktes  $P$  bezüglich des Dreiecks  $(ABC)$  sind proportional zu den Flächeninhalten der Dreiecke  $(PBC)$ ,  $(PAC)$ ,  $(PAB)$  und berechnen sich durch:

$$w_A = \frac{\begin{vmatrix} P_x & P_y & 1 \\ B_x & B_y & 1 \\ C_x & C_y & 1 \end{vmatrix}}{\begin{vmatrix} A_x & A_y & 1 \\ B_x & B_y & 1 \\ C_x & C_y & 1 \end{vmatrix}}, w_B = \frac{\begin{vmatrix} A_x & A_y & 1 \\ P_x & P_y & 1 \\ C_x & C_y & 1 \end{vmatrix}}{\begin{vmatrix} A_x & A_y & 1 \\ B_x & B_y & 1 \\ C_x & C_y & 1 \end{vmatrix}}, w_C = \frac{\begin{vmatrix} A_x & A_y & 1 \\ B_x & B_y & 1 \\ P_x & P_y & 1 \end{vmatrix}}{\begin{vmatrix} A_x & A_y & 1 \\ B_x & B_y & 1 \\ C_x & C_y & 1 \end{vmatrix}}.$$

Es gilt immer  $w_A + w_B + w_C = 1$ . Mit diesen Koordinaten ergibt  $Z(P) = Z(i_k, j_l) := w_A z_A + w_B z_B + w_C z_C$  den Wert der gesuchten

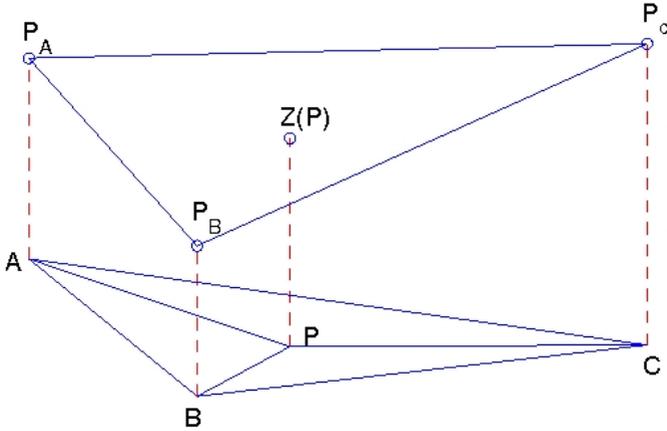


Abbildung 3.1: Lineare Interpolation durch baryzentrische Koordinaten

linearen Interpolation in dem Punkt  $P = (i_k, j_l)$ , wobei  $P_I(I_x, I_y, z_I)$  mit  $I \in \{A, B, C\}$  die entsprechenden Punkte in  $\mathbb{R}^3$  sind. Auf diese Art und Weise erhält man die entsprechenden Werte der Interpolation in allen Gitterpunkten.

Ein mögliches Kriterium für die Wahl der Feinheit des Gitters wäre, dass der Fehler auf jeder gemessenen Linie  $j_l$ , berechnet als

$$\sum_{P_i: y_i = j_l} (z_i - g(x_i, y_i))^2,$$

eine vorgegebene Toleranz  $tol$  nicht überschreitet, wobei  $g$  die lineare Interpolante durch alle Punkte der Form  $(i_k, j_l, Z(i_k, j_l))$ ,  $j_l$  fest, bezeichnet. Ein feineres Gitter bedeutet eine höhere Auflösung, d.h. einen kleineren Fehler, aber auch einen größeren Rechenaufwand. Deswegen soll die Anzahl der Gitterpunkte nicht von einer höheren Größenordnung als die Anzahl der ursprünglich gemessenen Punkte sein. Diese Aspekte werden im folgenden Beispiel erläutert.

**Beispiel:** Auf der Oberfläche des Kolbens hat man etwa 35.000 gemessene Punkte, deren  $(x, y)$ -Koordinaten sich innerhalb einer Kreisscheibe mit einem Durchmesser von 80mm befinden. Der Abstand zwischen den gemessenen Linien beträgt 0.5mm. Eine Gitterweite von 0.5mm (in beiden Richtungen) führt zu einer Anzahl von etwa 25.000 Gitterpunkten, von denen etwa ein Anteil von  $\pi/4$ , also ca. 19.000, innerhalb der Kreisscheibe liegen. D.h. man verliert schon in diesem Schritt fast die Hälfte der Information, welche die ursprünglich gemessenen Punkten liefern könnten. Falls man eine Gitterweite von 0.25mm in beiden Richtungen annimmt, dann entstehen ca. 100.000 Gitterpunkte, von denen ca. 74.000 innerhalb der Kreisscheibe liegen. Die Punktzahl hat sich also verdoppelt, aber die lineare Approximation ist in diesem Fall deutlich besser.

Im Falle des Kolbens kann man zur Kontrolle des Fehlers die neuen Linien (Polygonzüge durch die Werte in äquidistanten Punkten entlang einer Gitterlinie) mit den ursprünglichen Linien (Polygonzüge durch die ursprünglich gemessenen Punkte auf einer Linie) vergleichen. Dies wird in den Abbildungen 3.2 und 3.3 für eine Niveaulinie  $\{y = const\}$  des Kolbens dargestellt.

Der Fehler ist sehr klein mit Ausnahme der sehr steilen Stellen, wo er im Inneren des Intervalls unter 0.5mm liegt und am Rande etwa 2mm beträgt. Das Profil am Rand ist in Wirklichkeit senkrecht (daher kein Graph einer Funktion) und wurde in diesem Fall durch eine sehr steile Gerade approximiert. Dadurch entsteht der Fehler am Rand.

Ein Kriterium zur Auswahl der Gitterweite ist natürlich auch die Bedingung, dass die gemessenen Linien Teil des Gitters sind. Wegen der im nächsten Schritt erfolgenden Waveletzerlegung und wegen der Symmetrie des Objektes, wäre es andererseits wünschenswert, dass die Dimensionen der entsprechenden Matrix möglichst gleiche Werte haben, d.h. die Matrix sollte 'fast' quadratisch sein. Aus diesen beiden Gründen ergibt sich im Falle des Kolbens eine optimale Gitterweite von 0.25mm in beiden Richtungen.

Bei anderen Objekten, bei denen der Abstand zwischen den gemessenen Linien kleiner oder gleich 0.25mm ist, kann dieser Abstand auch als Gitterweite in beiden Richtungen genommen werden.

**Schlussfolgerung:** Weil die Originalpunkte mit einer sehr kleinen Schrittweite gemessen wurden, ist die Genauigkeit der linearen Interpolation völlig ausreichend. Eine kubische Interpolation basierend auf Triangulationen der ursprünglichen Punkte bringt keinen wesentlichen Unterschied, außerdem kann dies manchmal zu Artefak-

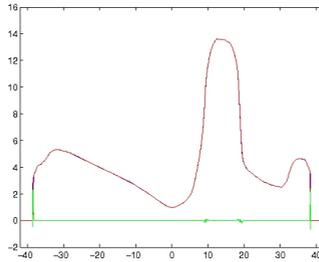


Abbildung 3.2: Originalkurve (blau), linear interpolierte Kurve (rot), Fehler als Differenz (grün)

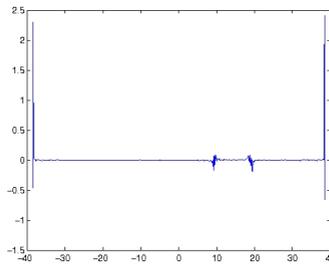


Abbildung 3.3: Fehler als Differenz

ten ('Zacken') an den steilen Stellen der Oberfläche führen. Im linearen Fall treten solche Effekte nicht auf. Aus diesem Grund ist für unsere Zwecke die lineare Interpolation die erste Wahl, auch weil schließlich Punkte aus der ursprünglich gemessenen Menge ausgewählt werden, und diese lineare Interpolation nur als Hilfsmittel für das im folgenden benutzte Auswahlverfahren dient. Für die Objekte, die in der Praxis auftreten (deren Größenordnung ca. 10cm beträgt), erweist sich eine Gitterweite von 0.25mm in beide Richtungen als ausreichend, sowohl angesichts der Genauigkeit der linearen Interpolation, als auch angesichts der Anzahl der entstandenen Gitterpunkte. Dieser Wert entspricht meistens dem Abstand zwischen den tatsächlich gemessenen Linien, oder kann als Vorgabe für das Messverfahren dienen.

### 3.1.2 Waveletzerlegung und Berechnung der $BMO^{-1}$ -Koeffizienten

Für die Matrix  $Z = (Z(i, j))_{i=1, m}^{j=1, n}$  wird nun eine Waveletzerlegung durchgeführt. Je nach der Dimension der Matrizen, die in praktischen Anwendungen auftreten (200x200 bis 600x600), erfolgt diese Zerlegung auf drei bis vier Niveaus. Bei einer Resolution in dieser Größenordnung würde eine weitere Zerlegung eine zu grobe Approximation auf dem letzten Niveau bringen und das Ergebnis würde damit nicht verbessert. Analog zur Formel (3.1) werden (approximierende)  $BMO^{-1}$  Koeffizienten in jedem Punkt  $x = (i, j)$  durch

$$b_{(i, j)} = \left( \sup_{Q \ni (i, j)} \frac{1}{|Q|} \sum_{Q(\lambda) \subset Q} |Q(\lambda)| \cdot |c_\lambda|^2 \right)^{1/2}$$

berechnet, wobei in diesem Fall die dyadischen Würfel  $Q$  und  $Q(\lambda)$  nur den berechneten Skalen angehören, an denen die Waveletkoeffizienten  $c_\lambda$  (horizontal, vertikal und diagonal) bestimmt wurden.

Um den analytischen Inhalt der Punkte besser nutzen zu können, führen wir nun ein 'Binning' der Punkte des Gitters anhand der Größe der  $BMO^{-1}$ -Koeffizienten durch, also eine Einteilung gemäss ihrer Größenordnung

('bin'=Schublade, engl.). Für gegebene threshold Parameter  $0 = \alpha_0 < \alpha_1 < \alpha_2 < \dots < \alpha_M < \alpha_{M+1} := \infty$  wurden die Gitterpunkte in disjunkte Teilmengen  $A_k$ ,  $k = \overline{0, M}$  eingeteilt, wobei  $(i, j) \in A_k \Leftrightarrow b_{(i, j)} \in [\alpha_k, \alpha_{k+1})$ .

Die Flächen, die in der Praxis vorkommen, sind meist durch stetige Funktionen darstellbar. Darauf entstehen durch diese Einteilung verschiedene Gebiete (Teilgebiete), ausgegrenzt durch die Niveaukurven  $\{x : b_x = \alpha_k\}$ . Die Auswahl der threshold Parameter  $\alpha_i$  muss so erfolgen, dass die unterschiedlichen Teile der Fläche (glatte, raue, kantige, usw) voneinander abgegrenzt werden, um in jeder dieser Bereiche die Punkte mehr oder weniger auszudünnen, abhängig von der Größe der  $BMO^{-1}$ -Koeffizienten. Nach einigen heuristischen Versuchen und ohne viel Zeitaufwand erhält man meistens das gewünschte Ergebnis. Dies wird in der Abbildung 3.4 im Falle des Kolbens illustriert.

Der Einsatz der  $BMO^{-1}$ -Koeffizienten liefert also Informationen über die lokale Struktur der Fläche und ermöglicht die Auswahl der Punkte über die ganze Oberfläche, mit unterschiedlichen Dichten in den einzelnen Bereiche.

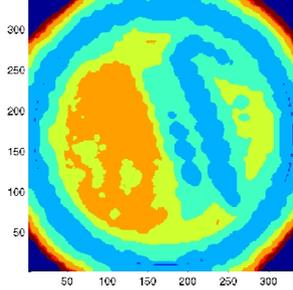


Abbildung 3.4: Einteilung der Gitterpunkte entsprechend der Größe der  $BMO^{-1}$ -Koeffizienten

## 3.2 Auswahl der Punkte

Gegeben seien die reellen Parameter  $0 = \alpha_0 < \alpha_1 < \alpha_2 < \dots < \alpha_M$  und die ganzzahligen Parameter  $d_0 \geq d_1 \geq \dots \geq d_M \geq 1$ .

Als erstes werden die Punkte aus  $A_M := \{(i, j) : b_{(i,j)} \geq \alpha_M\}$  wie folgt ausgedünnt: die Matrix  $Z = (Z(i, j))_{i=1, m}^{j=1, n}$  wird Element für Element durchlaufen. Falls der laufende Gitterpunkt in  $A_M$  liegt, wird er zu der bereits ausgewählten Menge nur dann hinzugefügt, falls der Abstand zu den Punkten dieser Menge (berechnet als  $|i_1 - i_2| + |j_1 - j_2|$ ) größer oder gleich  $d_M$  ist. Die Auswahl wird Schritt für Schritt erweitert, bis man die ganze Matrix durchlaufen hat. Nach dieser Prozedur erhält man somit eine Auswahl von Punkten aus  $A_M$ .

Im folgenden werden zu der bereits vorhandenen Auswahl nach dem gleichen Prinzip sukzessive Punkte aus den Mengen  $A_k$ ,  $k = M - 1, M - 2, \dots, 0$  hinzugefügt: ein Punkt aus  $A_k$  wird dann zu der bisherigen Selektion hinzugefügt (welche aus Punkten aus  $\bigcup_{j=k}^M A_j$  besteht), wenn der Abstand des Punktes zu dieser Menge größer oder gleich  $d_k$  ist.

Falls den Punkten aus den Mengen  $A_k$  ganze Bereiche entsprechen, dann bedeutet diese Prozedur im Grunde nichts anderes als eine Ausdünnung der Punkte in der Region  $A_k$  entsprechend dem Distanzparameter  $d_k$ .

Für jeden Parametersatz  $\gamma = (\alpha_1, \alpha_2, \dots, \alpha_M, d_1, d_2, \dots, d_M)$  erhalten wir also eine reduzierte Punktmenge  $\mathcal{S}'_\gamma$ , deren  $(x, y)$ -Koordinaten auf dem betrachteten Gitter liegen. Ausgehend von  $\mathcal{S}'_\gamma$  wählen wir

diejenigen ursprünglich gemessenen Punkte aus, welche dieser Punktmenge am nächsten liegen. Die Werte der  $BMO^{-1}$ -Koeffizienten der ausgewählten Gitterpunkte werden auf diese ursprünglich gemessenen Punkte übertragen, da sie in die bei der Spline-Interpolation benötigten Gewichte mit einbezogen werden.

Der reduzierte Punktesatz  $\mathcal{S}_\gamma$  ist also eine Teilmenge des Originalpunktesatzes, zur deren Bestimmung die Gitterinterpolation und das Binning der  $BMO^{-1}$ -Koeffizienten als Hilfsmittel benutzt werden. Er wird durch die Auswahl der Parameter  $\alpha_i$  beeinflusst, welche die Einteilung der Fläche in verschiedenen Gebieten steuern, so wie von der Wahl der Distanzparameter  $d_i$ , welche die Dichte der ausgewählten Punkte innerhalb dieser Gebiete festlegen.

Die ausgewählte Punktmenge ist unstrukturiert, daher muss ein geeignetes Verfahren benutzt werden, um eine gut approximierende Spline-Fläche zu bestimmen. Dieses wird in der nächsten Sektion vorgestellt.

### 3.3 Rekonstruktion der Fläche und Berechnung des Fehlers

Im folgenden wird eine Methode vorgestellt, die ein *surface-fitting* einer unstrukturierten Punktmenge in  $\mathbb{R}^3$  realisiert. Der Versuch, eine least-square Splineapproximation durchzuführen, scheitert oft an dem Rangdefizit der Matrix des Gleichungssystems, das die gesuchten Koeffizienten der Splines liefern soll. Die Lösung des least-square Problems ist also oft nicht eindeutig, im Unterschied zu der Situation, wenn die Punktmenge auf einem Gitter strukturiert ist. Hier gibt es kein Analogon der Schoenberg-Whitney Bedingung, welche die eindeutige Konstruktion einer least-squares Spline-Funktion gewährleistet. Auch wenn man eine spezielle least-squares Lösung des linearen Gleichungssystems zur Bestimmung der Koeffizienten auswählt (z.B. mit minimaler Norm, gegeben durch die Moore-Penrose Inverse oder Pseudoinverse), hat die rekonstruierte Fläche einen 'rauen' Aspekt und weist oft starke Oszillationen zwischen den Punkten auf.

Unter diesen Umständen wird eine sogenannte *Glattheitsnorm* eingeführt. Ziel dabei ist es, diese Norm zu minimieren unter der Bedingung, dass der quadratische Fehler unter einer vorgeschriebenen Toleranz liegt, oder man minimiert eine Linearkombination zwischen dem quadratischen Fehler und der Glattheitsnorm.

Zunächst wird eine Approximation durch glättende Splines für

eine gegebene Knotenmenge diskutiert, anschließend auch eine adaptive Strategie zur Platzierung der Knoten.

### 3.3.1 Glättende Splines für eine gegebene Knotenmenge

Seien also die gegebenen Daten  $P_r(x_r, y_r, z_r)$ ,  $r = 1, \dots, m$  mit  $(x_r, y_r) \in R := [a, b] \times [c, d]$ . Die vorhandenen Knoten seien  $\lambda_{-3} = \lambda_{-2} = \lambda_{-1} = \lambda_0 = a < \lambda_1 < \lambda_2 < \dots < b = \lambda_{g+1} = \lambda_{g+2} = \lambda_{g+3} = \lambda_{g+4}$  in der  $x$ -Richtung und  $\mu_{-3} = \mu_{-2} = \mu_{-1} = \mu_0 = c < \mu_1 < \mu_2 < \dots < d = \mu_{h+1} = \mu_{h+2} = \mu_{h+3} = \mu_{h+4}$  in der  $y$ -Richtung. Die Randknoten werden jeweils mit Vielfachheit 4 betrachtet, da eine Approximation mit bikubischen Splines angestrebt wird. Seien  $B_{i,4}(x)$  die kubischen Spline-Basisfunktionen mit Träger im Intervall  $[\lambda_i, \lambda_{i+4}]$ ,  $i = -3, \dots, g$  und  $\tilde{B}_{j,4}(y)$  die entsprechenden Basisfunktionen mit Träger im Intervall  $[\mu_j, \mu_{j+4}]$ ,  $j = -3, \dots, h$ .

Das Tensorprodukt bikubische Spline-Funktion bezüglich dieser Knoten hat also die Form

$$s(x, y) = \sum_{i=-3}^g \sum_{j=-3}^h c_{i,j} B_{i,4}(x) \tilde{B}_{j,4}(y). \quad (3.2)$$

Für eine Spline-Funktion der Form (3.2) mit entsprechendem Koeffizientenvektor  $\underline{c} = (c_{i,j})$  und für die Gewichte  $w_r$ ,  $r = 1, \dots, m$ , definiert man den quadratischen Fehler bezüglich der gegebenen Punkte  $P_r(x_r, y_r, z_r)$ ,  $r = 1, \dots, m$  als

$$\delta(\underline{c}) = \sum_{r=1}^m (w_r (z_r - s(x_r, y_r)))^2. \quad (3.3)$$

Das *least-square Problem* entspräche in diesem Falle der Bestimmung des Vektors  $\underline{c}$ , welcher  $\delta(\underline{c})$  minimiert. Dies entspricht der least-squares Lösung des überbestimmten linearen Gleichungssystems:

$$w_r \sum_{i=-3}^g \sum_{j=-3}^h c_{i,j} B_{i,4}(x_r) \tilde{B}_{j,4}(y_r) = w_r z_r, \quad r = 1, \dots, m, \quad (3.4)$$

oder in Matrix-Schreibweise:

$$E \underline{c} = \underline{z}. \quad (3.5)$$

Falls die Matrix  $E$  nicht vollen Rang hat, ist die least-squares Lösung jedoch nicht eindeutig. Dies ist oft der Fall bei einer unstrukturierten

Punktmenge, z.B. wenn in einem Patch  $R_{i,j} := [\lambda_i, \lambda_{i+1}] \times [\mu_j, \mu_{j+1}]$  keiner der gegebenen Punkte  $P_r$  liegt. Aber auch wenn die Knoten so gewählt werden, dass diese Situation nicht eintritt, kann trotzdem ein Rangdefizit vorhanden sein, welches nicht apriori erkennbar ist (siehe [7], S.150-151). Daher wählt man üblicherweise die Lösung mit minimaler Norm  $\|\underline{c}\|$ .

Die so erhaltenen Spline-Funktionen können jedoch einen ziemlich 'rauen' Aspekt haben, deswegen strebt man durch Hinzufügen zusätzlicher Gleichungen eine Regularisierung des Systems (3.4) an, so dass man eine eindeutige least-squares Lösung erhält, zu der die entsprechende Spline-Fläche ein glatteres Aussehen vorweist.

Für  $q = 1, \dots, g$  bezeichne  $a_{i,q} = B'''_{i,4}(\lambda_q+) - B'''_{i,4}(\lambda_q-)$  den Sprung der dritten Ableitung der Basisfunktion  $B_{i,4}$  in dem Knoten  $\lambda_q$ . Falls dieser Knoten nicht im Träger der Funktion  $B_{i,4}$  liegt, ist der entsprechende Wert gleich Null. Für  $r = 1, \dots, h$  definiert man analog  $b_{j,r} = \tilde{B}'''_{j,4}(\mu_r+) - \tilde{B}'''_{j,4}(\mu_r-)$ .

Für den Koeffizientenvektor  $\underline{c}$  definiert man folgende *Glätteitsnorm* ([7], p.159):

$$\eta(\underline{c}) = \sum_{q=1}^g \sum_{j=-3}^h \left( \sum_{i=-3}^g c_{i,j} a_{i,q} \right)^2 + \sum_{r=1}^h \sum_{i=-3}^g \left( \sum_{j=-3}^h c_{i,j} b_{i,r} \right)^2. \quad (3.6)$$

Für Polynome gilt immer  $\eta(\underline{c}) = 0$ , da die dritten Ableitungen stetig sind. Allgemein gilt: für kleine Werte von  $\eta(\underline{c})$  ist die Spline-Funktion 'nahe' an einen Polynom, also die entsprechende Fläche erscheint 'glatter' als in der Situation eines großen  $\eta(\underline{c})$ , d.h. wenn die dritten Ableitungen starke Oszillationen vorweisen.

Das Problem der *glättenden Spline* formuliert man daher als

$$\text{minimiere } \eta(\underline{c}) \quad (3.7)$$

$$\text{unter der Bedingung } \delta(\underline{c}) \leq S, \quad (3.8)$$

für einen gegebenen Parameter  $S > 0$ .

Durch Einführung eines Lagrange-Multiplikators  $p$  erhält man nach [7], S.159-160, folgendes überbestimmte Gleichungssystem für

die gesuchten Koeffizienten  $\underline{c}$  der glättender Spline  $s_p(x, y)$ :

$$w_r \sum_{i=-3}^g \sum_{j=-3}^h c_{i,j} B_{i,4}(x_r) \tilde{B}_{j,4}(y_r) = w_r z_r, \quad r = 1, \dots, m, \quad (3.9)$$

$$\frac{1}{\sqrt{p}} \sum_{i=-3}^g c_{i,j} a_{i,q} = 0, \quad q = 1, \dots, g; \quad j = -3, \dots, h, \quad (3.10)$$

$$\frac{1}{\sqrt{p}} \sum_{j=-3}^h c_{i,j} b_{j,r} = 0, \quad r = 1, \dots, h; \quad i = -3, \dots, g, \quad (3.11)$$

wobei  $p$  die positive Lösung der Gleichung  $F(p) = S$  ist, mit

$$F(p) = F_{g,h}(p) := \sum_{r=1}^m (w_r(z_r - s_p(x_r, y_r)))^2. \quad (3.12)$$

Die Indizes  $g, h$  deuten die Abhängigkeit von den gegebenen (inneren) Knoten an, deren Anzahl in die  $x$ -Richtung gleich  $g$  und in die  $y$ -Richtung gleich  $h$  ist. In Matrix-Schreibweise erhält man also das System:

$$P \underline{c} = \begin{bmatrix} E \\ \frac{1}{\sqrt{p}} D \\ \frac{1}{\sqrt{p}} H \end{bmatrix} \underline{c} = \begin{bmatrix} \underline{z} \\ \underline{0} \\ \underline{0} \end{bmatrix} \quad (3.13)$$

mit  $E$  gegeben aus (3.5).

Für moderate Werte von  $p$  hat die Matrix  $P$  vollen Rang, insbesondere wegen des regularisierenden Effektes der Glattheitsnorm (3.6). Die Werte von  $p$  hängen von der Wahl der Toleranz  $S$  ab. Die gesuchte Lösung der Gleichung  $F(p) = S$  erhält man durch folgende iterative Prozedur aus [7], 5.2.4, S.82-83. Diese liefert in der Praxis auch im Fall einer rangdefizienten Observationsmatrix  $E$  gute Ergebnisse bei der Bestimmung einer (nicht notwendig eindeutigen) Lösung  $p^*$ .

Mit den Startwerten  $p_1 = 0, p_2 = -\frac{F(0) - S}{F(\infty) - S}, p_3 = \infty$  bestimmt man die entsprechenden Funktionen  $s_{p_1}, s_{p_2}, s_{p_3}$  als least-squares Lösungen des Systems (3.9)-(3.11). Der Interpolationsansatz

$$R(p) = \frac{\alpha p + \beta}{p + \gamma}$$

liefert nach [8] eine gute Reproduktion der Eigenschaften von  $F(p)$ : falls  $\gamma > 0$  ist  $R(p)$  stetig auf  $(0, \infty)$  und für  $\alpha\gamma > \beta$  (d.h. falls  $R(0) >$

$R(\infty)$  folgt, dass  $R'(p) < 0$  und  $R''(p) > 0$  gilt. Die Parameter  $\alpha, \beta, \gamma$  werden durch die Bedingung  $R(p_i) = F(p_i)$  bestimmt. Die approximative Lösung der Gleichung  $R(p) = S$  wird dann nach einer iterativen Prozedur berechnet. Sei  $F_i = R(p_i)$ . Der nächste Wert  $p_*$  berechnet sich dann durch:

$$p_* = \frac{p_1(F_1 - F_3)(F_2 - S) - p_2(F_2 - F_3)(F_1 - S)}{(F_1 - F_2)(F_3 - S)}$$

für  $p_3 = \infty$ , oder  $p_* =$

$$\frac{p_1 p_2 (F_3 - S)(F_1 - F_2) + p_1 p_3 (F_2 - S)(F_3 - F_1) + p_2 p_3 (F_1 - S)(F_2 - F_3)}{p_3 (F_3 - S)(F_1 - F_2) + p_2 (F_2 - S)(F_3 - F_1) + p_1 (F_1 - S)(F_2 - F_3)}$$

für  $p_3 \neq \infty$ .

Falls  $|F(p_*) - S|/S$  kleiner als eine vorgeschriebene Toleranz ist, wähle  $p_*$  als gesuchten Wert und die entsprechende Spline-Funktion  $s_{p_*}$  als gesuchte Interpolante. Sonst setze  $p_3 = p_2, p_2 = p_*$ , falls  $F_2 < S$ , oder setze  $p_1 = p_2, p_2 = p_*$ , falls  $F_2 > S$ . Die Formel zur Berechnung von  $p_*$  gewährleistet immer, dass dieser Wert im Intervall  $(p_1, p_3)$  liegt. Aus den vier Punkten  $\{p_1, p_*, p_2, p_3\}$  bestimmt man nun die nächsten Werte von  $p_1, p_2, p_3$  nach dem angegebenen Kriterium. Die Länge des Intervalls  $[p_1, p_3]$  wird immer kleiner und strebt in den meisten Situationen gegen 0, somit ist die iterative Prozedur konvergent. Dies geschieht mit Sicherheit, wenn die sogenannte *Observationsmatrix*  $E$  vollen Rang hat. Dann weisen  $s_p(x, y)$  und  $F_{g,h}(p)$  folgende Eigenschaften auf ([7], p.161):

- Für jedes positive  $p$  existiert eine eindeutige Spline-Funktion  $s_p(x, y)$ , deren Koeffizientenvektor  $\underline{c}$  die least-squares Lösung des Systems (3.13) ist.
- Für  $p \rightarrow \infty$  konvergiert  $s_p(x, y)$  gegen die least-square Spline-Funktion  $S_{g,h}(x, y)$  bezüglich der gegebenen Knoten, deren Koeffizienten durch die least-squares Lösung des Systems (3.5) bestimmt sind, d.h. die Gleichungen (3.10),(3.11) werden einfach vernachlässigt. Sei

$$F_{g,h}(\infty) = \sum_{r=1}^m (w_r (z_r - S_{g,h}(x_r, y_r)))^2. \quad (3.14)$$

- Für  $p \rightarrow 0$  haben die Gleichungen (3.10),(3.11) ein immer größeres Gewicht im Gesamtsystem und die least-squares Lösung neigt dazu, diese exakt zu erfüllen, d.h. die dritte Ableitung

der Spline-Funktion darf im Limes  $p \rightarrow 0$  keine Sprünge aufweisen. Daher konvergiert  $s_p(x, y)$  gegen das entsprechende least-square bikubische Polynom  $P_{3,3}(x, y)$  bezüglich der Punkte  $P_r$ ,  $r = 1, \dots, m$ , definiert auf das Gebiet  $[a, b] \times [c, d]$ . Sei

$$F_{g,h}(0) = \sum_{r=1}^m (w_r(z_r - P_{3,3}(x_r, y_r)))^2. \quad (3.15)$$

- Für  $p > 0$  ist  $F_{g,h}(p)$  eine stetige, streng monoton fallende und konvexe Funktion.

Falls wir eine Knotenmenge gefunden haben, für welche

$$F_{g,h}(\infty) \leq S < F_{g,h}(0) \quad (3.16)$$

gilt, dann gibt es ein eindeutiges  $p_*$  und eine eindeutige glättende Spline-Funktion  $s_{p_*}(x, y)$  zu diesen Knoten, so dass  $F_{g,h}(p_*) = S$ . Der Wert für  $p_*$  wird numerisch durch die oben beschriebene iterative Prozedur bestimmt. Diese glättende Spline-Funktion  $s_{p_*}(x, y)$  liegt also zwischen den zwei Extremfällen: einerseits das least-squares bikubische Polynom bezüglich der gegebenen Punkte (maximale Glattheit) und andererseits das entsprechende least-squares Spline (beste Approximation, aber auf Kosten der Glattheit). Der Kompromiß zwischen der Approximation und der Glattheit wird durch den Parameter  $S$  gesteuert.

Falls die erste Ungleichung in (3.16) nicht erfüllt ist, dann kann für diese Knotenmenge die Nebenbedingung (3.8) nicht gelten. Daher muss zunächst eine passende Knotenmenge bestimmt werden. Die Methode dazu wird im folgenden beschrieben.

### 3.3.2 Eine adaptive Strategie zur Platzierung der Knoten

Falls die Knoten so ausgewählt werden, dass in jedem Patch  $R_{i,j} := [\lambda_i, \lambda_{i+1}] \times [\mu_j, \mu_{j+1}]$  genau ein Punkt  $P_r$  liegt, kann man direkt das interpolierende Spline auswählen (d.h.  $F_{g,h}(\infty) = 0$ ), um die Bedingung (3.16) zu erfüllen. Ziel ist jedoch eine Spline-Funktion mit möglichst wenig Knoten interessiert, daher wird zunächst nur mit den Randknoten das entsprechende least-squares Polynom  $P_{3,3}(x, y) = S_{0,0}(x, y)$  zu den Punkten  $P_r$  berechnet. Falls die Matrix  $E$  nicht vollen Rang hat, betrachtet man hier und im folgenden immer die

Lösung mit minimaler Norm des entsprechenden linearen Gleichungssystems. Falls  $F_{0,0}(\infty) \leq S$ , dann ist die gewünschte Funktion bereits gefunden. Gilt dies nicht, werden sukzessive least-squares Spline-Funktionen  $S_{g,h}(x, y)$  mit einer immer wachsenden Knotenzahl berechnet, bis die Bedingung (3.16) erfüllt ist. Bei jedem Iterationsschritt berechnet man folgende Größen:

$$R_r = w_r(z_r - S_{g,h}(x_r, y_r)), \quad r = 1, \dots, m \quad (3.17)$$

$$\delta_i^x = \sum_{\lambda_i \leq x_r < \lambda_{i+1}} R_r^2, \quad i = 0, \dots, g \quad (3.18)$$

$$\delta_j^y = \sum_{\mu_j \leq y_r < \mu_{j+1}} R_r^2, \quad j = 0, \dots, h \quad (3.19)$$

$$\delta_u^x = \max\{\delta_i^x : i = 0, \dots, g\} \quad (3.20)$$

$$\delta_v^y = \max\{\delta_j^y : j = 0, \dots, h\}. \quad (3.21)$$

Falls  $\delta_u^x \geq \delta_v^y$ , fügen wir einen Knoten bezüglich der  $x$ -Variable ein, da der Fehler in dieser Richtung größer ist. Sonst wird ein Knoten in  $y$ -Richtung eingefügt. Die Position des neuen Knotens ist gegeben durch

$$X = \frac{\sum_{\lambda_u \leq x_r < \lambda_{u+1}} R_r^2 x_r}{\sum_{\lambda_u \leq x_r < \lambda_{u+1}} R_r^2}, \quad (3.22)$$

bzw.

$$Y = \frac{\sum_{\mu_v \leq y_r < \mu_{v+1}} R_r^2 y_r}{\sum_{\mu_v \leq y_r < \mu_{v+1}} R_r^2}, \quad (3.23)$$

entsprechend der gewählten Richtung. Man fügt also zu demjenigen Streifen (vertikal  $[\lambda_u, \lambda_{u+1}] \times [c, d]$  oder horizontal  $[a, b] \times [\mu_v, \mu_{v+1}]$ ) einen Knoten hinzu, in dem der Fehler der Approximation am größten ist, d.h. zwischen  $\lambda_u$  und  $\lambda_{u+1}$ , bzw. zwischen  $\mu_v$  und  $\mu_{v+1}$ . Die genaue Positionierung innerhalb dieses Intervalls ist durch die Formel (3.22) gegeben. Die gewichtete Mittelung bewirkt, dass der Knoten dort positioniert wird, wo die Gewichte (d.h. Fehler in den einzelnen Punkten) am größten sind. Es garantiert auch, dass jeder der Streifen  $[\lambda_u, X] \times [c, d]$  und  $[X, \lambda_{u+1}] \times [c, d]$  Punkte  $P_r$  enthält. Dies ist nicht immer gewährleistet, z.B. wenn der neue Knoten einfach in der Mitte des Intervalls platziert wird.

Durch Wiederholung dieser Prozedur erhält man schließlich eine Spline-Funktion  $S_{g,h}(x, y)$  bezüglich der so konstruierten Knotenmenge, für welche  $F_{g,h}(\infty) \leq S$  gilt. Für diese Knotenmenge kann man also das Problem der glättenden Spline lösen.

Nun wird eine numerische Lösung  $p_*$  der Gleichung  $F(p_*) = S$  nach der bereits beschriebenen iterativen Prozedur bestimmt. Mit diesem  $p_*$  bestimmt man anschließend die least-squares Lösung des Systems (3.13), das im allgemeinen vollen Rang hat. Somit werden die entsprechenden Koeffizienten der gesuchten glättenden Spline-Funktion  $s_{p_*}(x, y)$  berechnet, die das Problem (3.7)-(3.8) löst.

### 3.3.3 Glättung durch Vorgabe des Glättungsparameters

Die bisher beschriebene Methode findet Anwendung, wenn der quadratische Fehler der Approximation unter einer vorgeschriebenen Schranke liegen muss. Sonst kann man die oben beschriebene adaptive Strategie der Knotenplatzierung anwenden, um die entsprechenden least-square Splines zu berechnen, während die Glättung durch verschiedene Auswahlmöglichkeiten von  $p$  gesteuert wird. In der Praxis, wenn keine expliziten Vorgaben an den Fehler der Approximation gegeben sind, wird diese Variante bevorzugt, da man mit viel weniger Aufwand zu einem akzeptablen Ergebnis kommt.

Das Problem lautet also jetzt:

$$\text{minimiere } \delta(\underline{c}) + \frac{1}{p}\eta(\underline{c}) \quad (3.24)$$

und führt ebenfalls zu dem Gleichungssystem (3.9)-(3.11), wobei diesmal  $p$  vorgegeben ist. Die least-square Lösung dieses Systems liefert dann die Koeffizienten der gesuchten Spline-Funktion.

Ist eine Schranke  $S$  für den Fehler der glättenden Spline vorgesehen, sind etliche Iterationen nötig, wobei in jedem Schritt eine neue glättende Spline-Funktion zu berechnen ist. Außerdem muß ein passender Wert für  $S$  gewählt werden, der von Fall zu Fall verschieden sein kann. Die richtige Wahl des Parameters  $S$  hängt von der Struktur der gegebenen Punktedaten ab. Werte die sehr nahe an Null liegen, sind aus numerischen Gründen nicht zu empfehlen, zu große Werte führen zu einer ungenauen Approximation. Der optimale Wert für  $S$  wird heuristisch bestimmt, daher sind auch hier einige Versuche notwendig. Dieser Aufwand lohnt sich nur, wenn man eine bereits

berechnete, akzeptable Flächenanpassung durch Spline-Funktionen optimieren möchte.

Wenn man in erster Linie nur an dieser 'akzeptablen Flächenanpassung' interessiert ist, reicht es aus, ca. 3-4 Werte von  $p$  zu testen und anschließend denjenigen auszuwählen, der den besten Kompromiss zwischen Glattheit und Approximationsfehler realisiert.

### 3.4 Auswahl der Parameter

Die Schritte **2.** und **3.** des Algorithmus liefern ein Ergebnis für vorgegebene Werte mehrerer Parameter. Die ausgewählte Punktmenge  $\mathcal{S}_\gamma$  hängt zunächst von den threshold Werten  $\alpha_i$  ab sowie von den Ausdünnungsentfernungen  $d_i$ . Durch Variation dieser Parameter kann man die Auswahl der reduzierten Punkte steuern, bis deren Anzahl in einer bestimmten Größenordnung liegt, oder um gewisse Bereiche der Fläche je nach Bedarf durch dichtere oder weniger dichte Punktmenge abzudecken.

Die entsprechende *smoothing-spline* Funktion, die die ausgewählten Punkte gemäß Schritt **4.** approximiert, hängt wiederum von den Werten der Gewichte  $w_r$ ,  $r = 1, \dots, m = |\mathcal{S}_\gamma|$  und des Glättungsparameters  $p$  ab, der entweder frei gewählt ist oder bestimmt wird durch die Forderung, dass der quadratische Fehler die Schranke  $S$  nicht überschreiten soll. Der Wert von  $S$  in Kombination mit der Anzahl der ausgewählten Punkte beeinflusst natürlich auch die Anzahl der Knoten der approximierenden Spline-Funktion.

Zur Auswahl der Gewichte bestehen verschiedene Möglichkeiten: Die einfachste Möglichkeit besteht darin, alle Gewichte gleich 1 zu setzen. Eine bessere Alternative ist es jedoch, die Gewichte nach folgender Funktion zu definieren:

$$w_r(x_r, y_r, z_r) = \frac{1}{\varepsilon + \beta \cdot b_{(i_k, j_l)}} \quad (3.25)$$

somit sind die Gewichte abhängig von dem  $BMO^{-1}$ -Koeffizient des Gitterpunktes  $(i_k, j_l)$ , der dem Punkt  $(x_r, y_r)$  am nächsten liegt, wobei  $\beta$  und  $\varepsilon$  positive Parameter sind. Der Parameter  $\varepsilon$  wird eingeführt, damit die Gewichte  $w_r$  im Falle sehr kleiner  $BMO^{-1}$ -Koeffizienten nicht beliebig groß werden können. Typischerweise wählt man  $\varepsilon = 0.1$ , um Gewichte im Wertebereich  $(1/(\varepsilon + \beta b_{max}), 10)$  zu erhalten, wobei  $b_{max}$  das Maximum der berechneten  $BMO^{-1}$ -Koeffizienten darstellt.

Diese Auswahl der Gewichte verbessert die optische Darstellung der erhaltenen Spline-Fläche erheblich. Die Erklärung ist folgende: Werden alle Punkte gleich gewichtet, tendiert die least-squares Spline-Fläche dazu, sich im gleichen Maße allen Punkten anzunähern. Dies führt in den Bereichen, in denen die ausgewählten Punkte dichter liegen (typischerweise bei Kanten), zu sichtbaren 'Zerknitterungen' der Fläche wegen der zu starken Forderung sich den ausgewählten Punkten anzupassen, die in diesen Bereichen dichter liegen. Wenn man diese Forderung abschwächt durch die Wahl kleinerer Gewichte für solche Punkte, erhält die Fläche einen natürlicheren Aspekt. Es erfolgt eine Art lokale Glättung, die bei geeigneter Wahl von  $\varepsilon$  und  $\beta$  die Details der Fläche nicht zerstört. Im nächsten Schritt, bei der globalen Glättung gemäß des Parameters  $p$ , ist nun viel weniger Glättung erforderlich (dies entspricht zu großen Werten von  $p$ ) als in dem Fall, wo alle Punkte gleich gewichtet würden. Diese Art Glättung hat einen Effekt über die ganze Fläche, während die passende Auswahl der Gewichte nur gewisse Teile der Fläche beeinflusst. Wo die Fläche genügend glatt ist, bringt die Wahl größerer Gewichte mehr Genauigkeit, aber keine Zerstörung des glatten Aussehens.

## 3.5 Praktische Aspekte der Interpolation und Berechnung des Fehlers

Im folgenden wird die praktische Umsetzung der Interpolationsmethode beschrieben, die von dem gegebenen reduzierten Punktesatz ausgeht.

**1. Adaptive Platzierung der Knoten.** Nach jeder neuen Knotenlegung wird die least-squares Spline-Funktion entsprechend der reduzierten Punktemenge berechnet. Dies erfolgt so lange, bis der quadratische Fehler der Approximation unter einer vorgegebenen Schranke liegt. Mit jedem neu gelegten Knoten wird der Fehler immer kleiner, man bemerkt jedoch, dass ab einem gewissen Wert diese Abnahme immer geringer ausfällt. Dies ist der Effekt des sogenannten '*overfitting*', wenn eine Punktemenge durch eine Splinefunktion approximiert wird, die auf zu vielen Knoten basiert, wie in den Abbildungen 3.5 und 3.6 illustriert wird.

### KAPITEL 3. ALGORITHMUS ZUR PUNKTEREDUKTION UND FLÄCHENREKONSTRUKTION

---

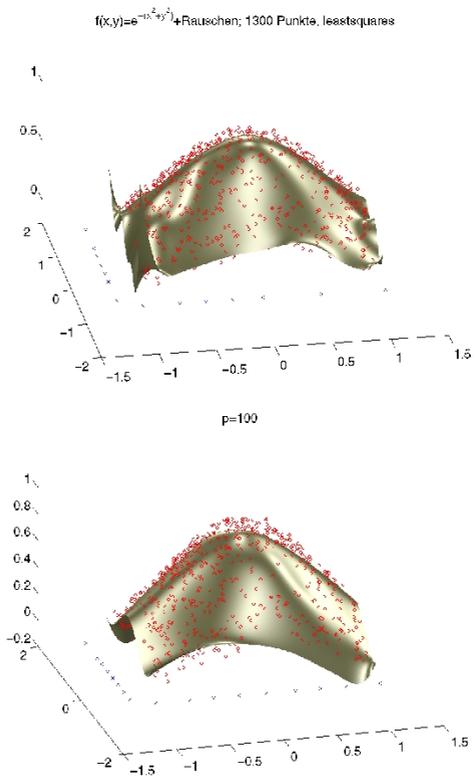


Abbildung 3.5: Approximierung durch 7x7 Knoten

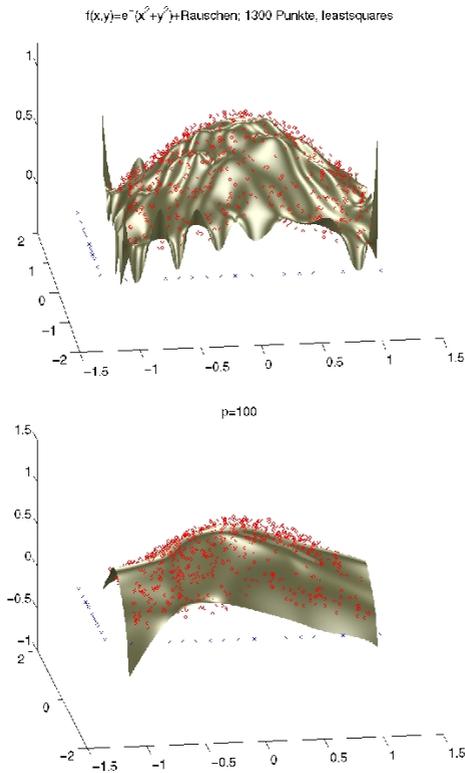


Abbildung 3.6: OVERFITTING: Approximierung durch  $13 \times 12$  Knoten

Oben sind die least-squares Splines dargestellt, unten die glättenden Splines. Man bemerkt, dass im Falle des overfitting auch nicht durch die Glättung das richtige Profil wiederhergestellt wird.

Wenn also die Abnahmerate des quadratischen Fehlers zu klein wird, wird die Prozedur gestoppt und ergibt somit die gewünschte Knotenmenge, ohne im voraus eine Schranke für diesen Fehler zu haben.

**2. Berechnung der glättenden Splines ab einer bestimmten Knotenzahl.** Wird durch sukzessive Platzierung eine bestimmte

Anzahl von Knoten erreicht, für die eine akzeptable Approximation zu erwarten ist, kann man zu jeder Knotenmenge parallel zu den least-squares Splines auch die glättenden Splines für 3-4 Werte des Glättungsparameters  $p$  berechnen und die entstandenen Fehler sowie die laufende Knotenmenge speichern. Nachdem die Platzierung der Knoten beendet ist, kann anhand der gespeicherten Ergebnisse eventuell auf eine frühere Knotenmenge zurückgegriffen werden, wenn für diese Menge der Fehler der glättenden Spline kleiner sein sollte, oder wenn der gleichen Fehler mit einer wesentlich kleineren Knotenmenge erreichbar ist.

**3. Optimierung der Approximation.** Anhand der bereits berechneten Fehler in jedem Schritt kann nun ein passender Wert für die Schranke  $S$  gewählt werden, die den quadratischen Fehler der glättenden Splines majorisiert. Zu dieser Schranke kann, wie bereits beschrieben, auch der optimale Wert von  $p$  berechnet werden.

Das ursprüngliche Problem war jedoch die Approximation der ursprünglich gemessenen Punktemenge  $P^{orig}(x_i, y_i, z_i)$ ,  $i = 1, \dots, n$ . Deswegen ist folgendes Maß für den Fehler interessant:

$$e(s_p) := \sum_{i=1}^n (z_i - s_p(x_i, y_i))^2 \quad (3.26)$$

Hier werden alle Gewichte auf 1 gesetzt, da die gemessenen Punkte die Fläche im Prinzip gleichmäßig abdecken. Deswegen ist es nicht notwendig, in einigen Bereiche eine stärkere Anpassung der Fläche an die Punkte zu fordern als in anderen.

Da in der Praxis die Anzahl der Punkte unterschiedlich sein kann, macht es Sinn, die *normierten* Fehler zu betrachten, nämlich

$$err_{sel}^w = \left( \frac{\sum_{r=1}^m (w_r (z_r - s_p(x_r, y_r)))^2}{\sum_{r=1}^m w_r} \right)^{\frac{1}{2}} \quad (3.27)$$

$$err_{orig} = \left( \frac{\sum_{i=1}^n (z_i - s_p(x_i, y_i))^2}{n} \right)^{\frac{1}{2}} \quad (3.28)$$

Diese Werte entsprechen einem 'mittleren Fehler' pro Punkt.



# Kapitel 4

## Anwendungen

In diesem Teil der Arbeit werden die Ergebnisse der Implementierung des beschriebenen Algorithmus auf praktische Beispiele präsentiert. Diese wurden so ausgewählt, dass sie besonders problematische Stellen aus Hinsicht einer Spline-Approximation aufweisen: Unstetigkeiten der ersten Ableitung, die entlang gewisser Kurven verlaufen, oder hohe Werte dieser Ableitungen, welche auf steile Profile hinweisen.

Als Erstes wurde das Verfahren anhand eines künstlich generierten Objektes getestet, an einem Kegelstumpf, auf dem eine Halbkugel liegt. Anhand der gleichen reduzierten Punktmenge werden mehrere Rekonstruktionen miteinander verglichen. Man beobachtet den Effekt des Glättungsparameters und den Einfluss auf die Rekonstruktion, die die Wahl der Gewichte der Spline-Approximation in Abhängigkeit von den  $BMO^{-1}$ -Koeffizienten mit sich bringt. Es stellt sich heraus, dass in vielen Fällen eine Verwendung von solchen Gewichten (beschrieben in Kapitel 3.4.) eine optisch bessere Rekonstruktion liefert.

Die nächsten Objekte, anhand derer das Verfahren getestet wurde, wurden am Institut für Spanende Fertigung der Universität Dortmund digitalisiert und von Herrn P.D.Dr.-Ing. Jörn Mehnen zur Verfügung gestellt. Es handelt sich um einen Motorkolben und um einen sogenannten "Dom". Beide Objekte weisen problematische Stellen im oben beschriebenen Sinne auf.

Schließlich wird auch die Exportierung der rekonstruierten Flächen in das IGES-Format diskutiert und anhand eines Beispiels verdeutlicht.

Die Programmierung des gesamten Punktereduktions- und Rekonstruktionsverfahrens erfolgte in MATLAB.

## 4.1 Rekonstruktion eines Testobjekts

Als erstes wird die beschriebene Methode auf ein Testobjekt angewandt. Es handelt sich um einen Kegelstumpf, auf dem eine Halbkugel liegt. Auf der Oberfläche des Objektes wurden die Koordinaten von 20.069 Punkten auf äquidistanten Linien berechnet, analog zu den reellen gemessenen Punktedaten. Abbildung 4.1 zeigt das Objekt

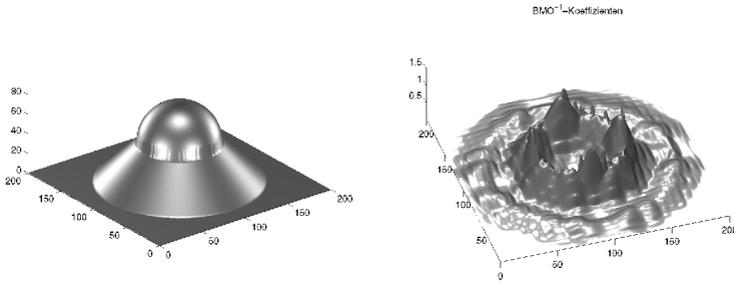


Abbildung 4.1: Testobjekt (links) und  $BMO^{-1}$ -Koeffizienten (rechts)

(linear interpoliert zwischen allen 20.069 Punkten) und die  $BMO^{-1}$ -Koeffizienten, während Abbildung 4.2 die Waveletkoeffizienten auf Level 1 horizontal (links) und Level 2 diagonal (rechts) darstellt.

Die Waveletzerlegung wurde auf drei Niveaus mit Hilfe der Daubechies 'db5'-wavelets durchgeführt. Man bemerkt, dass die Waveletkoeffizienten nur an den singulären Stellen (Kanten) eine relevante Größe haben, während die  $BMO^{-1}$ -Koeffizienten, wegen der Mittelung, auch in den anderen Regionen der Fläche relevant sind. Die anhand der Größe der  $BMO^{-1}$ -Koeffizienten und der beschriebenen thresholding Prozedur ausgewählten Punkte sind in Abbildung 4.3 dargestellt.

In den durch thresholding entstandenen Regionen wurden die Punkte dichter oder weniger dicht ausgewählt, je nachdem, ob die  $BMO^{-1}$ -Koeffizienten größer oder kleiner sind. Obwohl die berechneten Punkte nur auf der Oberfläche des Objektes liegen, wurden auch

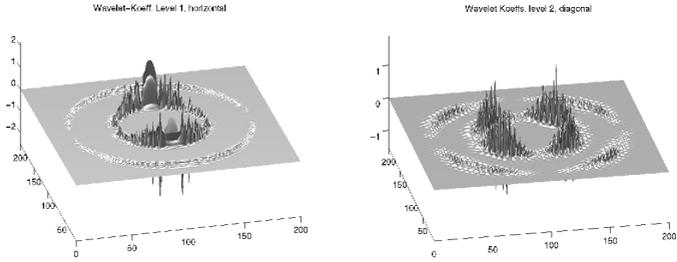


Abbildung 4.2: Waveletkoeffizienten: Level 1 horizontal und Level 2 diagonal

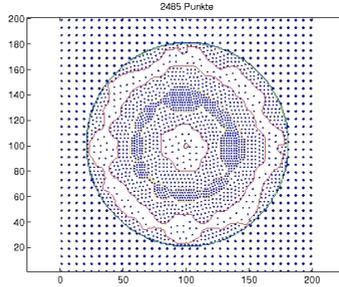


Abbildung 4.3: 2485 ausgewählte Punkte

Punkte auf der Platte ausgewählt, auf der ursprünglich keine Punkte betrachtet wurden. Dies geschieht auch aus rechnerischen Gründen, damit man Punkte im ganzen rechteckigen Gebiet  $[a, b] \times [c, d]$  erhält, aber diese Wahl führt auch zu einem optisch 'angenehmeren' Anblick der flachen Platte, wo sonst starke Oszillationen möglich gewesen wären.

Als erstes wurde die Flächenrekonstruktion für die Gewichte  $w_r = 1$  durchgeführt. Die Ergebnisse sind in Abbildung 4.4 dargestellt. Die Anzahl der Knoten beträgt  $31 \times 31$ . Der Wert  $p = 100$  (Bild links) bedeutet wenig Glättung, während der Wert  $p = 0.1$  (Bild rechts) mehr Glättung entspricht. Für  $p = 100$  erhält man die Fehler (gemäß (3.27) und (3.28)):  $err_{sel}^{w=1} = 0.40$ ,  $err_{orig} = 0.42$ , während man für  $p = 0.1$  die Werte  $err_{sel}^{w=1} = 0.74$ ,  $err_{orig} = 0.42$  erhält.

#### 4.1. REKONSTRUKTION EINES TESTOBJEKTS

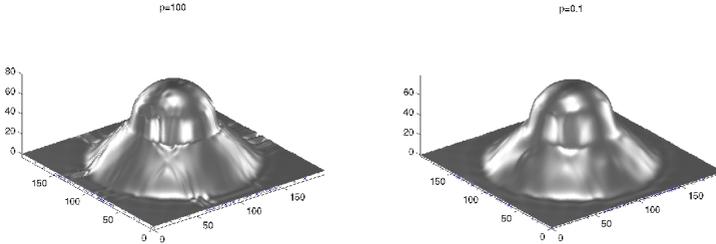


Abbildung 4.4: Rekonstruktion für  $w_r = 1$ ;  $p = 100$  (links) und  $p = 0.1$  (rechts)

Man bemerkt also folgenden Effekt: die Glättung erhöht zwar den Fehler bezüglich der ausgewählten Punkte, aber beeinflusst nur unwesentlich den Fehler bezüglich der ursprünglich berechneten Punkte (deren Anzahl fast zehn Mal so groß ist, wenn man die auf der Platte gewählten Punkten nicht in Betracht zieht).

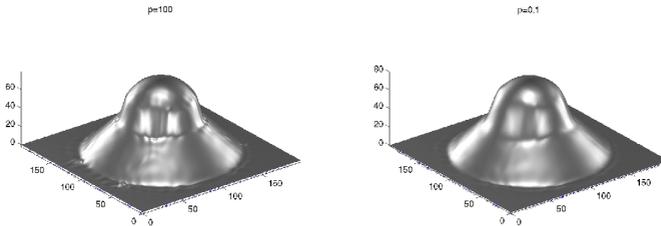


Abbildung 4.5: Rekonstruktion für  $w_r = (0.1 + 5b_r)^{-1}$ ;  $p = 100$  (links) und  $p = 0.1$  (rechts)

Als nächstes wurde die Flächenrekonstruktion für die Gewichte  $w_r = (0.1 + 5b_r)^{-1}$  durchgeführt, wobei  $b_r$  der  $BMO^{-1}$ -Koeffizient des Punktes  $P_r(x_r, y_r, z_r)$  ist (übertragen von dem nächstliegenden Gitterpunkt). Da die  $BMO^{-1}$ -Koeffizienten der Punkte auf der Platte, die weiter weg von dem Objekt liegen, gleich Null sind, ist nun der Grund der Einführung des Parameters  $\varepsilon$  in (3.25) einleuchtend.

Abbildung 4.5 zeigt die Ergebnisse der Rekonstruktion bei der gleichen Anzahl von Knoten ( $31 \times 31$ ) wie in der vorherigen Berechnung. Es sind jedoch nicht die gleichen Knoten, da die adaptive Prozedur bei jeder neuen Knotenplatzierung den quadratischen Fehler bezüglich der Gewichte verringert, und in diesem Fall sind andere Gewichte vorhanden. Für  $p = 100$  erhält man die Fehler  $err_{sel}^w = 0.12$ ,  $err_{sel}^{w=1} = 0.60$ ,  $err_{orig} = 0.49$  und für  $p = 0.1$  erhält man die Werte  $err_{sel}^w = 0.17$ ,  $err_{sel}^{w=1} = 0.90$ ,  $err_{orig} = 0.49$ . Im Vergleich zu den Gewichten  $w_r = 1$  ist hier die Fläche wesentlich glatter und optisch besser aussehend, wobei der Fehler bezüglich der ursprünglich berechneten Punkte nur etwas größer ist. Zur Erinnerung: bei der Berechnung des Fehlers bezüglich der Originalpunkte werden alle Gewichte auf 1 gesetzt. Hier wurde auch der Fehler bezüglich der ausgewählten Punkte mit den Gewichten 1 berechnet, um einen Vergleich mit dem vorherigen Beispiel zu haben. Wie erwartet, sind die Fehler in diesem Beispiel (für  $p = 100$  und  $p = 0.1$ ) wenig größer als im ersten Beispiel (mit Gewichten gleich 1) für dieselben Werte von  $p$ .

Die Maßeinheiten aus diesem Beispiel kann man als Millimeter interpretieren. Die Platte hat die Maße  $200\text{mm} \times 200\text{mm}$  und die Höhe des Objektes ist  $80\text{mm}$ . Der Fehler pro Punkt ist also in beiden Beispielen im Schnitt unter  $0.5\text{mm}$ .

## 4.2 Rekonstruktion des Motorkolbens

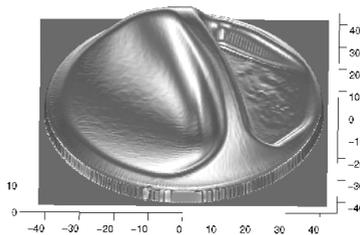


Abbildung 4.6: Motorkolben

Im folgenden wird die Rekonstruktion des Motorkolbens aus Ab-

bildung 4.6 durchgeführt.

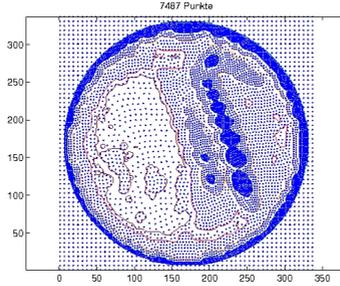


Abbildung 4.7: 7487 ausgewählte Punkte

Auf der Oberfläche des Kolbens wurden 34.493 Punkte gemessen. Nach der Berechnung der  $BMO^{-1}$ -Koeffizienten, thresholding und Punkteauswahlverfahren erhält man 7487 Punkte, wie in Abbildung 4.7 dargestellt. Wenn man von den Punkten auf der Platte absieht, ist dies etwa  $1/5$  der Anzahl der Originalpunkte.

Zur Rekonstruktion wurden die Gewichte  $w_r = (0.1 + 5b_r)^{-1}$  benutzt und die Knotenzahl beträgt  $48 \times 48$ . Wie im Testbeispiel ist auch hier die Ansicht der rekonstruierten Fläche für diese Auswahl der Gewichte deutlich besser, als wenn  $w_r = 1$  gesetzt worden wäre.

Die Ergebnisse der Rekonstruktion für  $p = 100$  und  $p = 10$  sind in den Abbildungen 4.8 und 4.9 dargestellt.

Für  $p = 100$  erhält man die Fehler  $err_{sel}^w = 0.10$ ,  $err_{sel}^{w=1} = 0.32$ ,  $err_{orig} = 0.47$ , während für  $p = 10$  die Werte  $err_{sel}^w = 0.12$ ,  $err_{sel}^{w=1} = 0.37$ ,  $err_{orig} = 0.47$  berechnet wurden.

Im Unterschied zum Testobjekt, welches mit Ausnahme der Kanten eine sehr glatte Fläche aufweist, ist im Falle des Kolbens der Fehler bezüglich der Originalpunkte stets größer als der Fehler bezüglich der ausgewählten Punkte. Für eine komplexe Fläche mit vielen problematischen Stellen wie in diesem Beispiel ist diese Tatsache auch zu erwarten.

Das Endergebnis ist also eine Rekonstruktion des Kolbens (mit 34483 ursprünglich gemessenen Punkte) mit Hilfe einer Spline-Fläche basierend auf  $48 \times 48$  Knoten. Dies entspricht einem Gitter mit ca. 2300 Punkten. Zur Berechnung dieser Spline-Funktion wurde eine reduzierte Punktemenge von 7487 Punkten benutzt. Weil diese Funk-

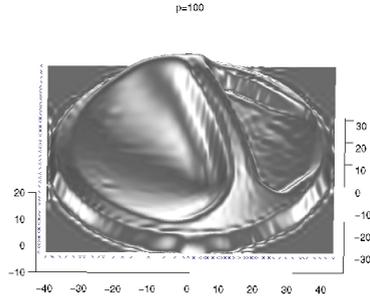


Abbildung 4.8: Rekonstruktion des Kolbens,  $p=100$

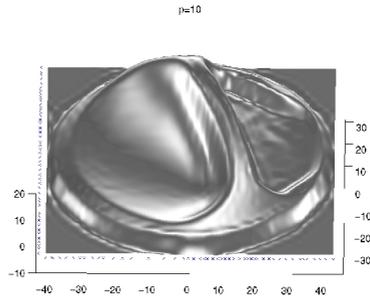


Abbildung 4.9: Rekonstruktion des Kolbens,  $p=10$

tion schließlich in das IGES-Format exportiert wird, sind jedoch in Hinsicht auf die Enddarstellung nur die Größe der Knotenmenge und der Grad der Spline-Funktion relevant.

### 4.3 Rekonstruktion des Doms

Ein anderes Objekt, worauf das hier entwickelte Punktereduktions- und Flächenrekonstruktionsverfahren angewendet wird, ist der ‐Dom‐, dargestellt in Abbildung 4.10.

Auf der Oberfläche des Doms wurden 204.686 Punkte gemessen, auf parallelen Linien im Abstand von 0.2mm. Die lineare Interpolation erfolgte auf ein Gitter der Größe  $606 \times 606$ . Es ist zu bemerken, dass der Dom deutlich größer als der Kolben ist (Platte ca.  $130 \times 130$ mm



Abbildung 4.10: Dom

statt  $80 \times 80$  mm und Höhe von über 50 mm, anstatt 18 mm). Er hat etwa die gleichen Größenordnung (in Millimeter) wie das erste Testobjekt, aber weist deutlich mehr Details als dieses auf. Es ist daher eine höhere Auflösung notwendig, sowohl was die gemessenen Punkte betrifft, als auch bezüglich der Anzahl der ausgewählten Punkte. Wegen der größeren Dimensionen des Gitters wurde hier eine Waveletzerlegung auf 4 Niveaus durchgeführt, so dass man nach der Punktereduktion eine Anzahl von 16.138 Punkten erhält (Abbildung 4.11).

16 138 ausgewählte Punkte

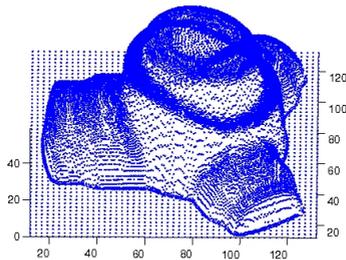


Abbildung 4.11: 16.138 ausgewählte Punkte

Diese Anzahl wurde infolge einiger Versuche bestimmt: Durch Variieren der Threshold- und Distanzparameter kann sowohl die Anzahl der ausgewählten Punkte als auch deren lokale Dichte gesteuert werden. Es stellt sich heraus, dass bei einer zu kleinen Anzahl von ausgewählten Punkten, z.B. bei etwa 11.000, die Rekonstruktion nicht optimal ist, insbesondere aus optischer Sicht. Glättung und Steuerung der Gewichte können in diesem Fall nicht verhindern, dass gewisse Details nicht originalgetreu dargestellt werden. Die Rekonstruktion erscheint entweder zu rau, oder aber eine zu starke Glättung führt zum Verlust von wichtigen Details. Zu viele ausgewählte Punkte, etwa 20.000 oder mehr, bedeuten gleichzeitig einen höheren Rechenaufwand. In diesem Beispiel wird deutlich, dass im Falle des Doms eine Anzahl von 16.138 Punkten einen akzeptablen Kompromiss zwischen Rechenaufwand und Qualität der rekonstruierten Fläche darstellt.

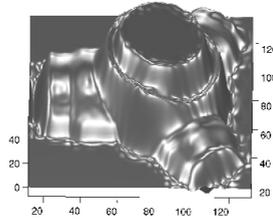


Abbildung 4.12: Rekonstruktion des Doms,  $p=5$ ,  $37 \times 37$  Knoten

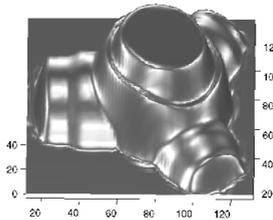


Abbildung 4.13: Rekonstruktion des Doms,  $p=5$ ,  $73 \times 73$  Knoten

Wegen der Größe des Objektes ist hier zu erwarten, dass man für die Darstellung mehr Knoten als in den vorherigen Beispielen benötigt. Im Falle des Kolbens, mit einem Gitter von  $337 \times 337$  Punkten, benötigte man  $48 \times 48$  Knoten. Wegen der großen Anzahl von ausgewählten Punkten würde der Algorithmus zu lange Zeit brauchen, um Schritt für Schritt bei einer Knotenzahl in der Ordnung von etwa  $80 \times 80$  anzukommen, da nach jeder neuen Knotenlegung jedesmal eine andere least-squares Spline berechnet werden muss, welche auf immer mehr Knoten basiert. Deswegen wird folgendermaßen vorgegangen: Der Algorithmus wird angehalten, sobald z.B. 50 Knoten in einer Richtung erreicht werden, und man sucht unter den gespeicherten vergangenen Knotenkonfigurationen die optimale aus, d.h. diejenige, für die der Fehler der geglätteten Spline-Fläche am kleinsten ist. Anschließend verdoppelt man die Knotenzahl direkt, indem man neue Knoten in der Mitte der vorhandenen Intervalle hinzufügt. Diese erste Phase der sukzessiven Berechnung mehrerer least-squares Splines und der adaptiven Knotenplatzierung kann man als eine gröbere Approximation des Objektes interpretieren, die im Großen alle relevanten Details erfasst, während die Verdoppelung der Knotenzahl als eine Verfeinerung angesehen werden kann, um ein glatteres Aussehen der Kanten und der sonstigen Details zu erhalten.

In Abbildung 4.12 ist eine Rekonstruktion des Doms basierend auf  $37 \times 37$  Knoten dargestellt. Als Glättungsparameter wurde  $p = 5$  gewählt und die Gewichte haben die Form:  $w_r = (0.1 + 3b_r)^{-1}$ . Als Fehler erhält man die Werte  $err_{sel}^w = 0.50$ ,  $err_{sel}^{w=1} = 1.14$ ,  $err_{orig} = 1.54$ . Nach der Verdopplung der Knotenzahl erhält man für die gleichen Werte des Glättungsparameters und der Gewichte die Rekonstruktion aus Abbildung 4.13. Die entstandenen Fehler sind jetzt:  $err_{sel}^w = 0.43$ ,  $err_{sel}^{w=1} = 0.98$ ,  $err_{orig} = 1.33$ . Wie man sehen kann, sind die Fehlerunterschiede nicht allzu groß, das Aussehen der Rekonstruktion mit doppelter Knotenzahl ist aber deutlich besser.

## 4.4 Exportieren der Daten im *IGES*-Format

IGES (Initial Graphics Exchange Specification) ist eine der verbreitetsten Standardisierungen für die Darstellung verschiedener geometrischer Objekte in CAD-Systemen. Die einzelnen Teile dieser Objekte können z.B. verschiedene Kurven oder Flächen sein, die durch Splines oder NURBS (Non-uniform Rational B-Splines) definiert sind.

KAPITEL 4. ANWENDUNGEN

Wegen seiner Verbreitung und Kompatibilität mit vielen CAD-Programmen, ist es in Hinsicht auf die Anwendbarkeit des Verfahrens wünschenswert, die rekonstruierte Fläche auch im IGES-Format darzustellen. Somit kann diese auch in den Verarbeitungsmaschinen eingegeben werden, um das gewünschte Objekt herzustellen. Aus diesem Grund ist eine Darstellung, die möglichst wenig Parameter enthält, notwendig, um die Größe der Datei in akzeptablen Grenzen zu halten.

Eine IGES-Datei muss genaue Formatierungsvorschriften erfüllen (z.B. jede Zeile muss genau 80 Charaktere enthalten und muss mit einer entsprechenden Zeilennummer enden) und die Datei muss in verschiedene Sektionen eingeteilt werden, welche Informationen über das dargestellte Objekt enthalten ("global section", "directory entry section", "parameter data section", "terminal section"). Diese Konvertierung wurde ebenfalls in MATLAB programmiert, und sieht z.B. im Falle des Kolbens wie folgt aus:

```

IGES-conversion                               S 1
1H,,1H;,23HIGESCONVERT,10Htest.igs,45H by Adina Guias,           G 1
11HVersion 3.2,32.308,15,308,15,5HNURBs,1,2,2HMM,1,0,1,          G 2
14H1000504.182300,1E-008,1000,18HAdina Guias,17HUniDO,10,0;     G 3
128 1 0 0 0 0 0                                         000000000D 1
128 0 16 1102 0                                         OD 2
128,49,49,3,3,0,0,0,0,0,-42.05,-42.05,-42.05,-42.05,-40.2598,  1P 1
-39.8759,-37.7944,-36.0661,-33.8477,-31.9941,-30.8117,-28.2276, 1P 2
-26.4307,-25.1834,-22.1971,-19.5434,-16.9831,-12.0153,-7.4557,  1P 3
-4.208,-1.6242,-0.28751,1.2697,3.118,4.2778,5.2879,6.8085,     1P 4
8.2705,9.3694,10.4734,11.6252,12.9759,14.5287,16.1555,17.0463, 1P 5
17.9953,19.0503,19.7599,21.6146,23.3039,24.6561,26.311,28.2049, 1P 6
30.1638,32.5264,33.3882,34.3468,36.6275,38.2281,39.7635,42.05, 1P 7
42.05,42.05,42.05,-42.05,-42.05,-42.05,-42.05,-39.6735,-38.4749, 1P 8
-36.768,-35.2466,-32.8136,-31.4298,-29.1231,-27.0241,-25.4001, 1P 9
-23.958,-20.1739,-18.6729,-17.3413,-14.8485,-12.394,-10.3838, 1P 10
-8.4921,-6.5407,-4.3002,-2.1314,-0.5661,1.373,3.3143,5.1467,   1P 11
6.6855,8.2003,10.7578,12.64,14.6268,16.7808,18.268,19.9103,    1P 12
21.3891,22.9092,23.9462,25.4386,26.8721,27.7529,28.6366,29.9521, 1P 13
31.4853,32.9264,34.7616,36.7979,38.9034,39.657,42.05,42.05,    1P 14
42.05,42.05,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1, 1P 15
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1, 1P 16
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1, 1P 17
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1, 1P 18
.....
41.5714,0.0094546,21.677,41.5714,0.0094577,23.1291,41.5714,      1P 1095
0.002352,24.8181,41.5714,-0.0013615,26.5279,41.5714,          1P 1096
28.3724,41.5714,0.0013971,30.1189,41.5714,0.00041129,31.726,  1P 1097
41.5714,-0.00086799,33.4105,41.5714,0.00011788,35.0234,       1P 1098
0.00017066,36.4708,41.5714,-0.0001318,38.2032,41.5714,       1P 1099
3.2269e-05,39.7438,41.5714,1.6328e-05,40.8283,41.5714,        1P 1100
-3.5274e-05,41.5927,41.5714,2.2519e-06,                        1P 1101
-42.05,42.05,-42.05,42.05;                                     1P 1102
S 1G 3D 2P 1102                                             T 1

```

#### 4.4. EXPORTIEREN DER DATEN IM *IGES*-FORMAT

In Abbildung 4.14 kann man einige Screenshots von der Darstellung des Objektes in einem IGES-Viewer beobachten.

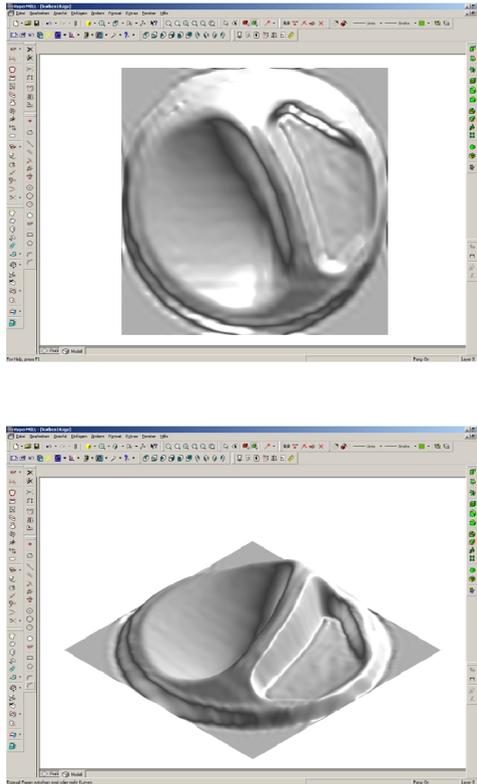


Abbildung 4.14: IGES-Darstellung des rekonstruierten Kolbens

# Kapitel 5

## Zusammenfassung und Ausblick

In dieser Arbeit wird das Problem der Flächenrekonstruktion, d.h. die Abbildung einer physischen Flächenform in ein mathematisches CAD-kompatibles Modell, analytisch betrachtet. Ausgangspunkt ist eine große Menge digital gemessener Punkte auf der Oberfläche eines Objektes. Ziel ist es, verschiedene (in der Praxis entwickelte) Prototypen derart zu beschreiben, dass einerseits eine kompakte Darstellung vorliegt, andererseits die gegebene Oberfläche genau und detailgetreu nachgebildet wird. Hierbei werden zahlreiche praktische Aspekte einbezogen, z.B.

- eine Abbildungsform, welche die Form-, Lage- oder Messfehler in einer erlaubten Toleranz hält;
- qualitativ hochwertige, z.B. glatte, scharfkantige Flächendarstellungen;
- CAD/IGES-adäquate Weiterverarbeitbarkeit;
- schnelle und sichere Berechnung der Rekonstruktionen.

Zur analytischen Lösung des Problems sind außerdem theoretische Aspekte zu berücksichtigen, z.B.

- die Wahl von passenden Kriterien zur Charakterisierung einer guten Approximation der Oberfläche;

- 
- geeignete Anwendung der Wavelets und  $BMO^{-1}$ -Funktionen, um die Anzahl der ursprünglich gemessenen Punkte zu reduzieren;
  - adaptive Strategie zur Platzierung der Knoten und Berechnung der Splinekoeffizienten;
  - die Wahl einer speicherplatz-, rechenzeitsparenden und CAD-adäquaten mathematischen Flächendarstellung.

Ausgangspunkt des Verfahrens ist eine Menge von sehr genau gemessenen Punkten auf der Oberfläche eines Objektes. Diese Vermessungen erfolgen digital, sehr genau und liefern sehr viele Punktedaten, die in dieser Form für eine Rekonstruktion durch CAD-kompatible Formate für eine Fräsmaschine nicht verwendet werden können. Das Ziel dieser Arbeit ist eine analytische Darstellung derartiger Oberflächen in Spline-Form, die auf so wenig Parametern wie möglich (Knoten, Koeffizienten) basiert. Der Einsatz von Wavelets und  $BMO^{-1}$ -Funktionen ermöglicht die Erkennung von Strukturen auf einer gegebenen Fläche, die nicht auf ein "Rauschen" zurückzuführen sind. Dies spielt eine wichtige Rolle in der Entwicklung einer Methode für Punktereduktion und Flächenrekonstruktion.

Wegen der großen Anzahl von ursprünglich gemessenen Punkten muß zunächst eine Punktereduktion erfolgen, sonst wäre der erforderliche Rechenaufwand untragbar. Das Problem, das es hier zu lösen gilt, ist, die gegebene Fläche in Bereiche aufzuteilen, welche relevante Informationen über die vorhandenen Strukturen enthalten. Dies wird durch die  $BMO^{-1}$ -Koeffizienten erreicht, deren Berechnung durch eine gewichtete Mittelung der Waveletkoeffizienten in der Umgebung jedes Punktes erfolgt. Die vorhandenen Punkte werden danach in den entstandenen Bereiche mehr oder weniger ausgedünnt, und als zusätzliche Information werden die entsprechenden  $BMO^{-1}$ -Koeffizienten der ausgewählten Punkte aufgenommen.

Anhand der reduzierten Punktemenge wird die Fläche durch die sogenannten glättenden Splines rekonstruiert. Die Gewichte, die dabei eingehen, werden in Abhängigkeit von den entsprechenden  $BMO^{-1}$ -Koeffizienten, invers proportional, ausgewählt. Dadurch erreicht man eine lokale Glättung an den Stellen mit größerer Punktedichte, wo die Spline-Funktionen ausgeprägte Artefakte aufweisen können, während man an den Stellen mit kleinerer Punktedichte eine große Genauigkeit erreicht. Die geschilderte Methode ermöglicht die Rekonstruktion einer Oberfläche samt vorhandenen größeren Unregelmäßigkeiten (Rauheiten). Die eventuellen Texturen (feinere Unregelmäßigkeiten)

werden jedoch wegen der Punktereduktion nicht mehr wiederhergestellt, insbesondere wenn man zusätzlich glätten muß.

Der entwickelte Algorithmus wird anhand einiger Objekte erfolgreich getestet. Er reagiert flexibel auf alle Anforderungen und ist auf alle Punktedaten, des gegebenen Typs anwendbar. Durch Steuerung der threshold-Parameter kann man die Einteilung der Fläche in Gebiete beeinflussen, und in jedem solchen Gebiet kann man bei Bedarf die ursprünglich gemessenen Punkte entsprechend der vorgestellten Kriterien mehr oder weniger ausdünnen, bis man eine optimale Rekonstruktion erhält. Anschließend werden die Daten der rekonstruierten Fläche in das IGES-Format exportiert, um Kompatibilität mit den gängigen CAD-Programmen und der Software der Bearbeitungsmaschinen zu erhalten.

In dieser Arbeit wird das Problem nur für Flächen, die als Graph einer gegebenen Funktion dargestellt sind, angegangen und gelöst. Es wäre jedoch wünschenswert, die Methode auch auf parametrisch definierte Flächen anwenden zu können. Dies könnte man durch eine Zerlegung des Objektes in mehrere Teilobjekte anhand der ursprünglich gemessenen Punkte erreichen (wenn die Struktur es auf natürliche Weise erlaubt). Jedes Teilobjekt könnte dann separat rekonstruiert werden, eventuell auch nach einer Rotation, um eine bessere Positionierung für die Anwendung der Methode zu erreichen.

Ein weiteres Ziel zur Abrundung dieser Arbeit wäre die Erstellung einer graphischen Benutzeroberfläche, um die Anwendung des Verfahrens und die notwendigen Versuche zur optimalen Wahl der Parameter einfacher zu gestalten.

Schlußfolgernd kann man sagen, dass der Einsatz analytischer Methoden eine wesentliche qualitative Verbesserung bei der Lösung der betrachteten anwendungsorientierten Problematik mit sich bringt.

---

# Literaturverzeichnis

- [1] E.M.Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality and Oscillatory Integrals*, Princeton University Press, 1993
- [2] Y.Meyer, *Wavelets and Operators I.*, Cambridge University Press, 1992
- [3] Y.Meyer, *Oscillating patterns in image processing and in some nonlinear evolution equations. The Fifteenth Dean Jaqueline B. Lewis Memorial Lectures*, University Lecture Series. 22. Providence, RI: American Mathematical Society (AMS), 2001
- [4] H.Koch, D.Tataru, *Well-posedness for the Navier-Stokes Equations*, Adv.Math. **157**, (22-35), 2001
- [5] J.Mehnen, *Evolutionäre Flächenrekonstruktion*, Schriftenreihe des ISF Dortmund, Hrsg. Klaus Weinert, Vulkan Verlag Essen, 2000
- [6] W.Dahmen, *Mathematische Methoden in der Geometrischen Datenverarbeitung (CAGD)*, Vorlesungsskript, Freie Universität Berlin, 1990
- [7] P.Dierckx, *Curve and Surface Fitting with Splines*, Oxford University Press, 1993
- [8] P.Dierckx, *An Algorithm for Surface-Fitting with Spline Functions*, IMA J.Num.Anal. **1**, (267-283), 1981
- [9] L.L.Schumaker, *Fitting Surfaces to Scattered Data*, in *Approximation Theory II* (ed. C.K.Chui, L.L.Schumaker, G.G.Lorentz), (203-268), Wiley, New York, 1976
- [10] S.Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999
- [11] A.Louis, P.Maaß, A.Rieder, *Wavelets*, Teubner-Verlag, Stuttgart 1998

- [12] H.Triebel, *Theory of Function Spaces*, Birkhäuser, 1983
- [13] R.Strichartz, *Bounded Mean Oscillation and Sobolev Spaces*, Indiana Univ.Math.J. **29**(4), (539-558), 1980
- [14] S.Janson, *On Functions with Derivatives in  $H^1$* , in *Harmonic Analysis and Partial Differential Equations*, Proceedings of the International Conference held in El Escorial, Spain, June 1987, J.Garcia-Cuerva(ed.) Springer, 1989
- [15] M.Frazier, B.Jawerth, *A Discrete Transform and Decomposition of Distribution Spaces*, J.Funct.Anal. **93**, (34-170), 1990
- [16] P.G.Lemarié, Y.Meyer, *Ondelettes et bases hilbertiennes*, Rev.Math.Iberoamericana. **2**, (1-18), 1986
- [17] I.Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992