



Technische Universität München
Zentrum Mathematik



Diplomarbeit

Multi-Coiflets

Andreas Hangauer

Betreuer: Dr. Peter Massopust
und Prof. Brigitte Forster-Heinlein
Zeitraum: 1. Juni 2009 bis 1. Dezember 2009

München, den 1. Dezember 2009

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und ohne unerlaubte Hilfsmittel angefertigt habe.

München, den 1. Dezember 2009

Andreas Hangauer

Inhaltsverzeichnis

1	Einleitung	3
1.1	Darstellung des Problems	3
1.2	Gliederung der Arbeit	5
2	Grundlagen	7
2.1	Allgemeines	7
2.2	Der Raum der unendlichen Linearkombinationen	9
2.3	Allgemeine Multiskalenanalyse	11
2.4	Orthogonale Multiskalenanalyse	15
2.5	Polynome	19
2.5.1	Vektor der Monome	19
2.5.2	Legendre-Polynome	20
3	Vorüberlegungen zur Konstruktion von Multiskalenanalysen	23
3.1	Bedingungen zur Konstruktion von Multiskalenanalysen	23
3.1.1	Verfeinerbarkeit und Verschachtelung	23
3.1.2	Orthogonalität	23
3.1.3	Approximationseigenschaft	23
3.1.4	Konstruktion eines Wavelets zur Skalierungsfunktion	24
3.2	Forderungen an eine Multiskalenanalyse	25
3.2.1	Stetigkeit der Skalierungsfunktion	25
3.2.2	Kompakter Träger	26
3.2.3	Approximationsordnung K	26
3.2.4	Verschwindende Momente der Skalierungsfunktion	27
3.2.5	Symmetrie	29
3.3	Normalformen der Darstellung der Skalierungsfunktion	30
3.3.1	Form mit minimalem Träger	30
4	Allgemeiner Ansatz für Skalierungsfunktionen	33
4.1	Die Basis für die Darstellung von Skalierungsfunktionen	33
4.1.1	Herstellung von Normiertheit aus linearer Unabhängigkeit	34
4.1.2	Definition der Skalierungsfunktion	34
4.1.3	Approximationsordnung	34
4.2	Generelle Eigenschaften	36
4.3	Orthonormalitätsbedingung	36
4.4	Dilatation	37
4.5	Stetigkeitsbedingung	38
4.6	Bedingung für verschwindende Momente	40

4.7	Kriterium, dass eine spezielle Funktion in $\mathbb{S}(\Phi_C)$ liegt	40
4.8	Zusammenfassung der Koeffizientengleichungen für Orthonormalität und Verfeinerbarkeit für Träger $[-1, 1]$	41
4.8.1	Ohne Bedingungen an $C(0)$ und $C(-1)$	41
4.8.2	Falls $C(0)$ vollen Rang besitzt	43
4.9	Hinreichendes Kriterium für Approximationsordnung	44
4.10	Eigenschaften von Skalierungsfunktionen mit verschwindenden Momenten	44
5	Skalierungsfunktionen aus stückweisen Polynomen	49
5.1	Die Basis aus stückweisen Polynomen	49
5.2	Lösen der Koeffizientengleichungen	52
5.3	Numerisch konstruierte Skalierungsfunktionen	53
5.3.1	Skalierungsfunktion mit Länge 4 und einem verschwindendem Moment	54
6	Zusammenfassung	57
	Literaturverzeichnis	61

1 Einleitung

Wavelets haben wichtige Anwendungen in der Signalverarbeitung wie Bildkompression und der numerischen Mathematik, wie im folgenden dargestellt wird.

1.1 Darstellung des Problems

Mittels der klassischen Fouriertransformation ist es möglich ein n -dimensionales Signal in seine Frequenzkomponenten zu zerlegen. Dies liefert eine gleichwertige Darstellung, in dem Sinne, dass eine Rekonstruktion fast überall möglich ist. Der Frequenzbereich liefert zudem eine für viele Probleme natürliche Darstellung, da einerseits einige Mathematische Probleme aus dem Bereich der (partiellen) Differentialgleichungen mittels der Fouriertransformation eine einfachere Gestalt haben (z.B. die lineare Diffusionsgleichung oder die Wellengleichung), andererseits die Eindrücke aus den Sinnen "Hören" und "Sehen" eher auf den Frequenzbereich als den Zeitbereich des wahrgenommenen Signals schließen lassen.

Die Fouriertransformation als mathematisches Hilfsmittel verfügt jedoch über keine zeitliche Lokalisierungseigenschaft, da lokale Änderungen der Funktion im Zeitbereich (also auf einem beschränkten Intervall) sich auf fast alle Punkte des Frequenzbereichs auswirken. So ist es im Rahmen der Fouriertransformation nicht möglich, von einer sich zeitlich verändernden "Tonhöhe" zu sprechen. Abhilfe schafft hier die Kurzzeit-Fouriertransformation, bei der vor der Fouriertransformation von $x(t)$ ein zeitlich verschiebbares "Fenster" $w(t)$ multipliziert wird:

$$\hat{x}(\tau, \omega) = \int_{\mathbb{R}} w(t - \tau)x(t)e^{-i\omega t} dt.$$

Aufgrund der Unschärferelation der Fouriertransformation kann die Frequenzauflösung und Zeitauflösung nicht gleichzeitig beliebig hoch sein. Die Zeitauflösung wird durch die "Länge" von $w(t)$ eingestellt. Im Fall $w(t) = 1$ erhält man die gewöhnliche Fouriertransformation ohne jegliche Zeitauflösung aber maximaler Frequenzauflösung. Mathematisch ist die Unschärferelation gegeben durch die Tatsache, dass das zweite zentrale Moment des Betragsquadrats der Zeitfunktion multipliziert dem zweiten Moment des Betragsquadrats der Fouriertransformierten stets größer als $(4\pi)^{-2}$ ist (Falls die Momente existieren). Insbesondere erhält man für die Kurzzeit-Fouriertransformierte bei festem τ eine gleiche *absolute* Frequenzauflösung, da die

Breite des Fensters w nicht von ω abhängt. Für viele Anwendungen ist es jedoch sinnvoll, eine konstante *relative* Frequenzauflösung zu erhalten. Mathematisch entspricht dies einer Transformation bei der das Fenster $w(t)$ mit ω skaliert wird, also sich proportional mit der Frequenz verkleinert, d.h. eine Transformation der Art

$$\tilde{x}(\tau, \omega) = \int_{\mathbb{R}} w((t - \tau)\omega)x(t)e^{-i\omega t} dt \quad (1.1)$$

Werden Existenz und Bedingungen für die Umkehrbarkeit vernachlässigt, ist erkennbar dass eine solche Transformation die Zeitauflösung bei hohen Frequenzen $|\omega|$ und die Frequenzauflösung bei niedrigen $|\omega|$ verbessern würde. Dies entspricht auch in natürlicher Weise der Empfindlichkeit des menschlichen Ohrs was aus anatomischen Gründen eine relative Frequenzauflösung besitzt. Diese Tatsache manifestiert sich in der logarithmischen Eigenschaft der Tonskala: die Frequenzen der 12 Halbtöne einer Oktave haben ein festes Frequenzverhältnis von $2^{1/12}$, wobei die Oktave gerade einer Frequenzverdopplung entspricht. Für Anwendungen in der Kompression von Audiodaten, wo das Fehlerkriterium durch das menschliche Gehör bestimmt ist, erscheint die Darstellung der Audiosignale mit konstanter relativer Frequenzauflösung ein natürlicher Ansatz und es ist wünschenswert Transformationen mit passenden Eigenschaften zu finden. Falls man Gl. (1.1) mit dem komplexen Phasenfaktor $e^{i\omega\tau}$ multipliziert (der das Betragsspektrum nicht ändert), so erhält man eine Faltung mit einer gefensterter Schwingung, wobei dies dann zu Transformationen von der Gestalt

$$\tilde{x}(\tau, \omega) = \int_{\mathbb{R}} x(t)\phi((t - \tau)\omega)dt$$

verallgemeinert werden kann und die Funktion ϕ eine gefensterter Schwingung oder "kleine Welle" enthält. Diese Beobachtung ist der Ausgangspunkt für die Entwicklung der Wavelettransformation.

Da in der praktischen Anwendung mittels Computer nur zeitdiskrete Signale verarbeitet werden können, sucht man Transformationen der Gestalt

$$\tilde{x}(i, j) = \int_{\mathbb{R}} x(t)\phi(2^j t - i) dt, \quad (1.2)$$

wobei die Zeit $i \in \mathbb{Z}$ linear abgetastet und die Frequenz durch $j \in \mathbb{Z}$ in einer logarithmischen Art und Weise diskretisiert wird. Die Unterschiede zwischen Wavelet- und Kurzzeit-Fouriertransformation sind in Abb. 1.1 versinnbildlicht. Mit der Entwicklung der Theorie der Multiskalenanalyse, wurde das Fundament zur Untersuchung von Transformationen der Art Gl. (1.2) gelegt. Insbesondere Bedingungen für Umkehrbarkeit und Form der Umkehrtransformation sind damit bekannt. Für jedes feste j kann $i \mapsto \tilde{x}(i, j)$ als Tiefpasssignal, das in einem geeigneten Funktionenraum V_j liegt, interpretiert werden. Steigt j werden die Räume größer, d.h. enthalten mehr "hochfrequente

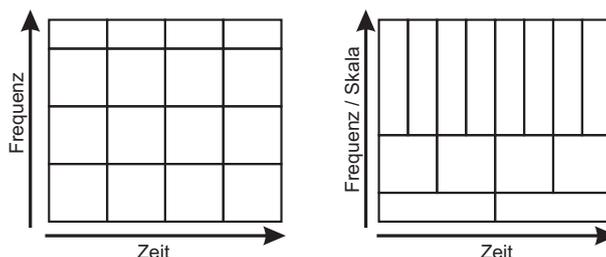


Abbildung 1.1: Schematische Darstellung der Unterschiedlichen Frequenz- und Zeitauflösung der Kurzzeit-Fouriertransformation (links) und der Wavelettransformation (rechts). Da die Unschärfere-lation für beide Transformationen gilt, ist das Produkt aus Zeit- und Frequenzunschärfe (symbolisiert durch die Breite und Höhe der Kästen) konstant.

Anteile" und das in der beschriebenen dyadischen Art und Weise. Praktisch wird aber so vorgegangen, dass ein Signal x nicht in seine verschiedenen Tiefpasskomponenten zerlegt wird, sondern in eine Tiefpasskomponente und eine übrige Anzahl Hochpasskomponenten. Diese enthalten anschaulich interpretiert die Differenzinformation von $\tilde{x}(i, j + 1)$ und $\tilde{x}(i, j)$, also das, was zwar auf Skala $j + 1$ nicht aber auf Skala j dargestellt werden kann. Dies entspricht auch der Darstellungsweise in Abb. 1.1.

Die Entwicklung von Wavelets bzw. Multiskalenanalysen begann mit skalaren Wavelets, d.h. die Räume V_j besitzen nur einen Erzeuger. Ein Durchbruch war die Konstruktion von Skalierungsfunktionen und Wavelets mit kompaktem Träger, beliebiger Approximationsordnung und verschwindenden Momenten der Skalierungsfunktion [Dau92]. Die letzteren werden Coiflets genannt und haben den Vorteil, dass die Darstellungskoeffizienten aus diskret vorliegenden Abtastwerten einer Funktion mit geringem Fehler berechnet werden können (siehe Abschnitt 3.2.4). Der zweite Durchbruch war die Konstruktion der GHM Multi-Skalierungsfunktion [GHM94], bei der zwei symmetrische Generatoren mit kompaktem Träger eine Multiskalenanalyse von Vielfachheit 2 erzeugen. Im Gegensatz zu skalaren Multiskalenanalysen lassen sich Multiskalenanalysen mit einer Vielfachheit R konstruieren, bei denen die Skalierungsfunktionen verfeinerbar, orthogonal und symmetrisch sind, was bei skalaren Skalierungsfunktionen nicht möglich ist (siehe Abschnitt 3.2.5).

1.2 Gliederung der Arbeit

Aufgabe der Arbeit ist es, Untersuchungen zu Multi-Coiflets zu führen, also die Erweiterung der skalaren Coiflets auf Multiskalenanalysen von Vielfachheit R mit dem ultimativen Ziel Multi-Coiflets in möglichst großer All-

gemeinheit zu konstruieren. In der vorliegenden Arbeit werden daher stets Multiskalenanalysen mit einer gewissen Vielfachheit vorausgesetzt. Genauso ist der Begriff Skalierungsfunktion und Wavelet immer vektoriell gemeint.

Im Kapitel 2 werden nötige Grundlagen bereitgestellt: neben Schreibweisen auch wichtige Definitionen wie der ein- und zweiparametrischen Familie auch der Raum \mathbb{S} , der eine wichtige Rolle bei der Charakterisierung der Approximationsordnung spielt. Es werden dafür auch einige später benötigte Eigenschaften formuliert, die teilweise in der Literatur nicht zu finden sind. Es folgt die Definition der Multiskalenanalyse, sowie zugehörige Definitionen wie die der Skalierungsfunktion, des Wavelets und der Tief- und Hochpasskoeffizienten und Beweis einiger ihrer Eigenschaften. Abschließend wird der in der Praxis relevante Mallat Algorithmus hergeleitet, da mit diesem die Zerlegung in Hoch- und Tiefpassanteil praktisch berechnet wird und auch die Interpretation einer Multiskalenanalyse mit den Mitteln der Signalverarbeitung möglich wird.

In Kapitel 3 wird genauer auf die für die Konstruktion von Multiskalenanalysen notwendigen Hilfsmittel eingegangen. Wünschenswerte Zusatzbedingungen werden formuliert und begründet, wie beispielsweise das Vorhandensein einer gewissen Approximationsordnung, verschwindende Momente der Skalierungsfunktion oder Symmetrie der Filterkoeffizienten. Es wird gezeigt, dass die Hauptaufgabe die Konstruktion einer geeigneten Skalierungsfunktion ist und das Wavelets dann daraus berechnet werden kann.

In Kapitel 4 wird ein allgemeiner Ansatz für die zu konstruierende Skalierungsfunktion gewählt und alle nötigen Eigenschaften der Skalierungsfunktion als Koeffizientenbedingungen formuliert. Am Ende von Kapitel 4 werden zwei Hauptergebnisse präsentiert, dies ist zum einen die Charakterisierung der Approximationsordnung, die den experimentellen Befund, dass stückweise polynomielle Skalierungsfunktionen immer eine ihrem Grad entsprechenden Approximationsordnung haben, theoretisch erklärt. Das Resultat gilt auch allgemein für skalare Skalierungsfunktionen. Ein weiteres theoretisches Ergebnis ist, dass stetige Multi-Coiflets einen Träger größer als $[-1, 1]$ haben müssen. Bei vektoriellen Skalierungsfunktionen vergrößert die Approximationsordnung den Träger nicht, wenn die Vielfachheit $R > 1$ entsprechend gewählt wird. Es scheint also bezüglich der verschwindenden Momente der Skalierungsfunktion kein grundsätzlicher Unterschied zu skalaren Coiflets vorzuliegen, da bei diesen der Träger auch mit zunehmender Approximationsordnung und verschwindenden Momenten größer wird. Die hergeleitete untere Schranke für die Länge der Skalierungsfunktion bei verschwindenden Momenten gilt zudem unabhängig von der Vielfachheit R .

In Kapitel 5 wird der allgemeine Ansatz auf stückweise Polynome spezialisiert und eine Lösung der Koeffizientengleichungen angegangen. Da für Multi-Coiflets die "Intertwining Methode" aus [DGH96] nicht anwendbar ist, wird auf eine numerische Lösung der Gleichungen zurückgegriffen.

2 Grundlagen

In diesem Kapitel werden einige nötigen Grundlagen bereitgestellt. Die Definition des Raumes $\mathbb{S}(\Phi)$ stammt aus [Jia97]., wobei einige nötige Eigenschaften selbst hergeleitet werden. Die Definition der Multiskalenanalyse für mehrere Skalierungsfunktionen ist eine direkte Verallgemeinerung der herkömmlichen eindimensionalen Definition wie in [Mal89b].

2.1 Allgemeines

Definition 2.1.1 (Lineare Hülle) Sei \mathcal{B} eine (möglicherweise unendliche) Menge von Funktionen. Die Menge aller (endlichen) Linearkombinationen $\text{span}\{\mathcal{B}\}$ wird lineare Hülle genannt, und mittels

$$\text{span}\{\mathcal{B}\} = \left\{ \sum_{i=1}^n a_i b_i \mid a_i \in \mathbb{R}; b_i \in \mathcal{B}; n \in \mathbb{N} \right\} \quad (2.1)$$

definiert.

Definition 2.1.2 (Träger einer Funktion) Sei Φ eine Funktion $\mathbb{R} \rightarrow \mathbb{R}^R$. Der Abschluss der nicht auf Null abgebildeten Argumente wird als Träger $\text{supp}(\Phi)$ bezeichnet:

$$\text{supp}(\Phi) = \text{cl}_{\mathbb{R}} \{t \mid t \in \mathbb{R}; \Phi(t) \neq \mathbf{0}\}. \quad (2.2)$$

Definition 2.1.3 (Riesz-Basis) Eine abzählbare Menge von Vektoren $\mathcal{B} = \{\mathbf{b}_i \mid i \in I\}$ eines Hilbertraums \mathcal{H} (z.B. $L^2(\mathbb{R})$) heißt Riesz-Basis, falls

$$\text{cl}_{\mathcal{H}} \{\text{span}\{\mathcal{B}\}\} = \mathcal{H} \quad (2.3)$$

und es $A, B > 0$ gibt, so dass für alle Folgen $\mathbf{x} = (x_n) \in l^2$

$$A \|\mathbf{x}\|_2^2 \leq \left| \sum_{i \in I} x_i \mathbf{b}_i \right|^2 \leq B \|\mathbf{x}\|_2^2 \quad (2.4)$$

erfüllt ist.

Bemerkung Die Eigenschaft mit $A > 0$ sichert, dass der Nullvektor eindeutig durch die Nullfolge dargestellt wird, also die Koeffizientenfolge der Darstellung für alle Hilbertraumelemente eindeutig ist. Ist $A = B$ so liegt

eine orthogonale Basis vor und falls $A = B = 1$ eine orthonormale Basis. Grund: Ist $A = B$, so gilt $A \sum_{i \in I} x_i^2 = |\sum_{i \in I} x_i \mathbf{b}_i|^2 = \sum_{i,j \in I} x_i x_j \langle \mathbf{b}_i, \mathbf{b}_j \rangle = \sum_{i \in I} x_i^2 \|\mathbf{b}_i\|^2 + \sum_{i,j \in I, i \neq j} x_i x_j \langle \mathbf{b}_i, \mathbf{b}_j \rangle$. Mit $x_i = \delta_{in}$ folgt $\|\mathbf{b}_n\|^2 = A$. Dann gilt: $\sum_{i,j \in I, i \neq j} x_i x_j \langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0$, was wegen x beliebig, $\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0$ für $i \neq j$ erfordert.

Definition 2.1.4 (Einparametrische Familie der Skalierung p)

Sei $\Phi = (\phi_1 \ \cdots \ \phi_R)^T$ eine Funktion $\mathbb{R} \rightarrow \mathbb{R}^R$. Dann ist die einparametrische Familie $\mathcal{B}_p(\Phi)$ mit $p \in \mathbb{Z}$ definiert als die Menge der $\mathbb{R} \rightarrow \mathbb{R}$ Funktionen

$$\mathcal{B}_p(\Phi) = \{ 2^{p/2} \phi_r(2^p \cdot -k) \mid r = 1, 2, \dots, R; k \in \mathbb{Z} \}. \quad (2.5)$$

Definition 2.1.5 (Zweiparametrische Familie) Sei $\Phi = (\phi_1 \ \cdots \ \phi_R)^T$ eine Funktion $\mathbb{R} \rightarrow \mathbb{R}^R$. Dann ist die zweiparametrische Familie $\mathcal{B}(\Phi)$ definiert als die Menge der $\mathbb{R} \rightarrow \mathbb{R}$ Funktionen

$$\mathcal{B}(\Phi) = \bigcup_{p \in \mathbb{Z}} \mathcal{B}_p(\Phi) = \{ 2^{p/2} \phi_r(2^p \cdot -k) \mid r = 1, \dots, R; k \in \mathbb{Z}; p \in \mathbb{Z} \} \quad (2.6)$$

Definition 2.1.6 (Skalarprodukt) Das Skalarprodukt $\langle \mathbf{a}, \mathbf{b} \rangle$ von zwei vektoriellem Funktionen $\mathbf{a} : \mathbb{R} \rightarrow \mathbb{R}^M$ und $\mathbf{b} : \mathbb{R} \rightarrow \mathbb{R}^N$ wird bezeichnet mit

$$\langle \mathbf{a}, \mathbf{b} \rangle = \int_{\mathbb{R}} \mathbf{a}(t) \mathbf{b}(t)^T dt, \quad (\in \mathbb{R}^{M \times N}). \quad (2.7)$$

Bemerkung In der Definition wird keine Aussage über die Existenz des Skalarprodukts gegeben. Diese hängt auch von dem zugrunde gelegten Maß ab. Für das Lebesgue-Maß existiert das Skalarprodukt beispielsweise, wenn die Komponenten in $L^2(\mathbb{R})$ liegen oder integrierbar sind und mindestens einen kompakten Träger besitzt.

Bemerkung Obwohl das so definierte "Skalarprodukt" eine Matrix liefert, ist die Definition sinnvoll. Das Skalarprodukt bezieht sich auf die Elemente der Vektoren \mathbf{a} und \mathbf{b} die in einem geeigneten Funktionenraum liegen und nicht auf \mathbf{a} und \mathbf{b} selber bezüglich ihrer "Vektor Eigenschaft". Die Matrix enthält daher alle paarweise Skalarprodukte der Komponenten von \mathbf{a} und \mathbf{b} .

Das Skalarprodukt besitzt die folgenden Eigenschaften: (Sei hierzu $\mathbf{a} : \mathbb{R} \rightarrow \mathbb{R}^M$, $\mathbf{b} : \mathbb{R} \rightarrow \mathbb{R}^N$, so dass $\langle \mathbf{a}, \mathbf{b} \rangle \in \mathbb{R}^{M \times N}$ existiert)

$$\langle \mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{b}, \mathbf{a} \rangle^T, \quad (2.8a)$$

$$\langle \mathbf{Aa}, \mathbf{Bb} \rangle = \mathbf{A} \langle \mathbf{a}, \mathbf{b} \rangle \mathbf{B}^T, \quad \mathbf{A} \in \mathbb{R}^{P \times M}, \mathbf{B} \in \mathbb{R}^{Q \times N}, \quad (2.8b)$$

$$\langle \mathbf{a}(\cdot/s), \mathbf{b}(\cdot/s) \rangle = s \langle \mathbf{a}, \mathbf{b} \rangle, \quad s \in \mathbb{R} \setminus \{0\}, \quad (2.8c)$$

$$\langle \mathbf{a}(\cdot - \tau), \mathbf{b}(\cdot - \tau) \rangle = \langle \mathbf{a}, \mathbf{b} \rangle, \quad \tau \in \mathbb{R}. \quad (2.8d)$$

2.2 Der Raum der unendlichen Linearkombinationen

Die lineare Hülle eines Funktionensystems besteht nur aus endlichen Linearkombinationen. "Unendliche Linearkombinationen" sind im Allgemeinen nicht wohldefiniert, da Konvergenz nicht sichergestellt ist. Trotzdem ist es möglich für eine einparametrische Familie einer Funktion mit kompaktem Träger einen Funktionenraum aus unendlichen Linearkombinationen zu definieren.

Bei einer Multiskalenanalyse liegen alle auftretenden Räume in $L^2(\mathbb{R})$ und enthalten daher beispielsweise keine Polynome. Trotzdem können die Polynome bezüglich der Basis dieser Räume dargestellt werden, wenn *beliebige unendliche* Linearkombinationen zugelassen werden. Dies hat erhebliche Vorteile bei der Charakterisierung der Multiskalenanalysen bezüglich ihrer Approximationsordnung.

Definition 2.2.1 (Der Raum $\mathbb{S}(\Phi)$) Sei $\Phi = (\phi_1 \ \cdots \ \phi_R)^T$ eine Funktion $\mathbb{R} \rightarrow \mathbb{R}^R$ mit kompaktem Träger, dann ist der Raum $\mathbb{S}(\Phi)$ definiert als die Menge aller "unendlichen Linearkombinationen" von $\mathcal{B}_0(\Phi)$:

$$\mathbb{S}(\Phi) = \left\{ \sum_{k \in \mathbb{Z}} \sum_{r=1}^R a_{k,r} \phi_r(\cdot - k) \mid a_{k,r} \in \mathbb{R}; k \in \mathbb{Z}; r = 1, \dots, R \right\}. \quad (2.9)$$

Bemerkung Da die Funktionen ϕ_r ebenfalls kompakten Träger besitzen, liegen stets Summen mit nur endlich vielen nichttrivialen Summanden vor, was die Wohldefiniertheit zeigt.

Selbst wenn $\phi_r \in L^2(\mathbb{R})$ gilt, ist im Allgemeinen $\mathbb{S}(\Phi) \not\subset L^2(\mathbb{R})$. Es gilt auch $\text{cl}_{L^2(\mathbb{R})} \{\text{span} \{\mathcal{B}_0(\Phi)\}\} \subsetneq \mathbb{S}(\Phi)$. Man hat jedoch den folgenden Satz:

Satz 2.2.1 (Zusammenhang zwischen $\mathbb{S}(\Phi)$ und $\text{cl}_{L^2(\mathbb{R})} \{\text{span} \{\mathcal{B}_0(\Phi)\}\}$) Ist $\Phi : \mathbb{R} \rightarrow \mathbb{R}^R$ eine Funktion mit kompaktem Träger, dann gilt

$$\text{cl}_{L^2(\mathbb{R})} \{\text{span} \{\mathcal{B}_0(\Phi)\}\} = \mathbb{S}(\Phi) \cap L^2(\mathbb{R}). \quad (2.10)$$

Beweis Siehe Theorem 2 in [Jia97]. (Es gilt sogar in den Räumen $L^p(\mathbb{R})$, $1 < p < \infty$, nicht jedoch für $p = 1$). \square

Satz 2.2.2 (Bestimmung der Koeffizienten in $\mathbb{S}(\Phi)$) Ist $\Phi : \mathbb{R} \rightarrow \mathbb{R}^R$ eine integrierbare Funktion mit kompaktem Träger, $\mathbb{S}(\Phi)$ der Raum der "unendlichen Linearkombinationen" von $\mathcal{B}_0(\Phi)$ und $\mathcal{B}_0(\Phi)$ orthonormal, dann gilt für die eindeutig bestimmten Koeffizienten $\mathbf{f} : \mathbb{Z} \rightarrow \mathbb{R}^R$ der Darstellung von $f \in \mathbb{S}(\Phi)$ bezüglich $\mathcal{B}_0(\Phi)$:

$$\mathbf{f}(k) = \langle f, \Phi(\cdot - k) \rangle \quad (2.11)$$

mit

$$f = \sum_k \mathbf{f}(k) \Phi(\cdot - k). \quad (2.12)$$

Beweis Nach Voraussetzung $f \in \mathbb{S}(\Phi)$ existiert $f : \mathbb{Z} \rightarrow \mathbb{R}^R$, so dass $f = \sum_k f(k) \Phi(\cdot - k)$ gilt. Wird das Skalarprodukt mit $\Phi(\cdot - l)$ auf beiden Seiten von Gl. (2.12) angewendet, so ergibt sich unter Ausnutzung der Orthonormalität $f(l) = \langle f, \Phi(\cdot - l) \rangle$. Zudem ist $f(l)$ eindeutig bestimmt. \square

Bemerkung Obwohl die Form der Koeffizienten es suggeriert, handelt es sich hierbei nicht um eine Projektion. Nur wenn bekannt ist, dass die Funktion f bezüglich $\mathcal{B}(\Phi)$ darstellbar ist, also $f \in \mathbb{S}(\Phi)$, sind ihre Koeffizienten nach Gl. (2.11) gegeben. Es wird nicht für allgemeines f die Projektion auf den Raum $\text{cl}_{L^2(\mathbb{R})} \{\text{span} \{\mathcal{B}_0(\Phi)\}\}$ berechnet. Dies ist nur dann der Fall, wenn $f \in L^2(\mathbb{R})$ ist.

Satz 2.2.3 (Kriterium, dass eine Funktion in $\mathbb{S}(\Phi)$ liegt) *Sei Φ eine integrierbare Funktion mit kompaktem Träger und $\mathbb{S}(\Phi)$ der Raum der "unendlichen Linearkombinationen" von $\mathcal{B}_0(\Phi)$. Wenn weiter $\mathcal{B}_0(\Phi)$ orthonormal ist, dann liegt eine lokal integrierbare Funktion f genau dann in $\mathbb{S}(\Phi)$, wenn gilt*

$$f = \sum_k \langle f, \Phi(\cdot - k) \rangle \Phi(\cdot - k). \quad (2.13)$$

Beweis Falls $f \in \mathbb{S}(\Phi)$, dann gilt die Aussage nach Satz 2.2.2. Gilt dagegen Gl. (2.13), so existiert das Skalarprodukt weil f lokal integrierbar ist und Φ kompakten Träger besitzt. Es ist dann f trivialerweise in $\mathbb{S}(\Phi)$, da es eine unendliche Linearkombination aus $\mathcal{B}_0(\Phi)$ ist. \square

Durch die spezielle Struktur der Polynome, liegen sie in dem Raum $\mathbb{S}(\Phi)$ schon dann, wenn sie nur auf einem Einheitsintervall z.B. $[0, 1)$ darstellbar sind, allerdings nur falls die richtigen Koeffizienten vorliegen. Falls beliebige Koeffizienten vorliegen, dann muss der Bereich mindestens doppelt so lang sein wie die Länge der Skalierungsfunktion um Enthaltensein der Polynome in $\mathbb{S}(\Phi)$ sicherzustellen.

Satz 2.2.4 (Kriterium, dass $\Pi_{N-1} \subset \mathbb{S}(\Phi)$) (I) *Sei Φ eine integrierbare Funktion mit kompaktem Träger und $\mathbb{S}(\Phi)$ der Raum der "unendlichen Linearkombinationen" von $\mathcal{B}_0(\Phi)$. Wenn weiter $\mathcal{B}_0(\Phi)$ orthonormal ist, dann liegen die Polynome vom Grad $< N$ genau dann in $\mathbb{S}(\Phi)$, wenn*

$$\mathbf{p}(t) = \sum_k \langle \mathbf{p}, \Phi(\cdot - k) \rangle \Phi(t - k), \quad \text{für } t \in [I, I + 1) \quad (2.14)$$

und einem $I \in \mathbb{Z}$. \mathbf{p} bezeichnet den Vektor der Monome vom Grad $< N$ (siehe Abschnitt 2.5).

Beweis Die Richtung " \Rightarrow " folgt aus Satz 2.2.3.

" \Leftarrow ": Bezeichne $\mathbf{q}(t) = \sum_k \langle \mathbf{p}, \Phi(\cdot - k) \rangle \Phi(t - k)$. Es gilt $\mathbf{q}(t) = \mathbf{T}(-l)\mathbf{q}(t - l)$ mit der Matrix aus Gl. (2.49d) für $\mathbf{T}(l)\mathbf{p}(t) = \mathbf{p}(t - l)$. Dies wird wie folgt gezeigt: Sei

$l \in \mathbb{Z}$, dann ist $\mathbf{q}(t) = \sum_k \langle \mathbf{p}, \Phi(\cdot - l - k) \rangle \Phi(t - l - k) = \sum_k \langle \mathbf{p}(\cdot + l), \Phi(\cdot - k) \rangle \Phi(t - l - k) = \mathbf{T}(-l) \sum_k \langle \mathbf{p}, \Phi(\cdot - k) \rangle \Phi(t - l - k) = \mathbf{T}(-l) \mathbf{q}(t - l)$. Gilt nun $\mathbf{p}(t) = \mathbf{q}(t)$ für $t \in [I, I + 1)$, so folgt nach Multiplikation mit $\mathbf{T}(l)$: $\mathbf{T}(l)\mathbf{p}(t) = \mathbf{T}(l)\mathbf{q}(t) \Leftrightarrow \mathbf{p}(t - l) = \mathbf{q}(t - l)$, ebenfalls für $t \in [I, I + 1)$. Also stimmt \mathbf{p} und \mathbf{q} auf jedem Intervall $[I - l, I - l + 1)$ überein, also auf ganz \mathbb{R} . \square

Satz 2.2.5 (Kriterium, dass $\Pi_{N-1} \subset \mathbb{S}(\Phi)$, (II)) *Sei Φ eine integrierbare Funktion mit kompaktem Träger höchstens der Länge $W \in \mathbb{Z}$ und $\mathbb{S}(\Phi)$ der Raum der "unendlichen Linearkombinationen" von $\mathcal{B}_0(\Phi)$. Wenn weiter $\mathcal{B}_0(\Phi)$ orthonormal ist, dann liegen die Polynome vom Grad $< N$ genau dann in $\mathbb{S}(\Phi)$, wenn es Koeffizientenmatrizen $\mathbf{X}(k)$ gibt mit*

$$\mathbf{p}(t) = \sum_k \mathbf{X}(k) \Phi(t - k), \quad t \in [I - W + 1, I + W) \quad (2.15)$$

für ein $I \in \mathbb{Z}$, mit \mathbf{p} dem Vektor der Monome vom Grad $< N$ (siehe Abschnitt 2.5).

Beweis Die Richtung " \Rightarrow " folgt aus Satz 2.2.3.

" \Leftarrow ": Sei der Träger von Φ oBdA in $[0, W]$ enthalten. Dies ist keine Einschränkung, da die Funktion Φ beliebig verschoben werden kann und immer noch der gleiche Raum $\mathbb{S}(\Phi)$ erzeugt wird. Für $t \in [I, I + 1)$ läuft die Summation in Gl. (2.15) dann von $k = I - W + 1, \dots, I$. Nun wird auf beiden Seiten das Skalarprodukt mit $\Phi(\cdot - k)$ und $k = I - W + 1, \dots, I$ berechnet. Diese Operation ist wohldefiniert, da der Integrationsbereich in $\langle \mathbf{p}, \Phi(\cdot - k) \rangle$ durch $[k, k + W]$ gegeben ist, d.h. für das geforderte k ist dieser $[I - W + 1, I + W]$. Für diesen Bereich stimmen aber nach Voraussetzung beide Seiten der Gleichung überein. Man hat dann also für $k = I - W + 1, \dots, I$, $\mathbf{X}(k) = \langle \mathbf{p}, \Phi(\cdot - k) \rangle$. Nun lässt sich für $t \in [I, I + 1)$ Satz 2.2.4 anwenden, womit die Behauptung folgt. \square

2.3 Allgemeine Multiskalenanalyse

Mit dem Begriff der Multiskalenanalyse werden die theoretischen Grundlagen für die rekursive Aufspaltung von Funktionen in eine grobe Näherung und immer feiner werdende Details in einer dyadischen Art und Weise gelegt. Die Axiome für die skalare Multiskalenanalysen stammen aus [Mal89b]. Sie sind ohne Probleme vektorieell erweiterbar, wie auch umfangreiche Literatur über Multi-Wavelets zeigt [GHM94, SS94, DGH96, Roa97, DGH99].

Das für die Anwendung sehr wichtige Ergebnis ist der Mallat Algorithmus, mit dem die sukzessiven Zerlegungen in Hoch- und Tiefpassanteil in V_p und W_p sowohl berechnet werden als auch wieder rückgängig gemacht werden können. Hierzu müssen lediglich die Koeffizienten der Approximation einer Funktion in einem gewissen Raum V_n bezüglich der üblichen Basis bekannt sein. Von dort ausgehend, kann eine Zerlegung in Details und Grobe Näherung in den W_p und V_p mit $p \leq n$ berechnet werden. In [Mal89a] ist der

Mallat Algorithmus für skalare Multiskalenanalysen beschrieben, er verallgemeinert sich aber ohne weiteres auf vektorielle Multiskalenanalysen, wie sie hier untersucht werden.

Definition 2.3.1 (Multiskalenanalyse der Vielfachheit R) *Eine Folge von abgeschlossenen Unterräumen (V_p) mit $V_p \subset L^2(\mathbb{R})$ heißt Multiskalenanalyse der Vielfachheit R mit $R \in \{1, 2, \dots\}$ falls gilt:*

$$\textbf{Verschachtelung} \quad \dots \subset V_{p-1} \subset V_p \subset V_{p+1} \subset \dots, \quad (2.16a)$$

$$\textbf{Approximation} \quad \bigcup_{p \in \mathbb{Z}} V_p \text{ ist dicht in } L^2(\mathbb{R}) \text{ und } \bigcap_{p \in \mathbb{Z}} V_p = \emptyset, \quad (2.16b)$$

$$\textbf{Verfeinerbarkeit} \quad f \in V_p \Leftrightarrow f(2^{-p} \cdot) \in V_0 \text{ für } p \in \mathbb{Z}, \quad (2.16c)$$

$$\textbf{Basis} \quad \exists \Phi = (\phi_1 \ \dots \ \phi_R)^T : \mathcal{B}_0(\Phi) \text{ ist Riesz Basis von } V_0. \quad (2.16d)$$

Bemerkung Falls die Funktionen aus der Basis ϕ_r einen kompakten Träger haben, so ergeben sich einige Implikationen im obigen Axiomensystem, denn es folgt beispielsweise die Approximationseigenschaft Gl. (2.16b) aus den übrigen Axiomen (Theorem 2.2 in [JS94] sowie [dBDR92]).

Definition 2.3.2 (Projektionsoperator auf V_p) *Der Operator $L^2(\mathbb{R}) \rightarrow V_p$, der einer Funktion $f \in L^2(\mathbb{R})$ ihre beste Approximation aus dem Raum V_p einer Multiskalenanalyse zuordnet, ist linear und wird mit P_p bezeichnet. Es gilt:*

$$P_p f \in V_p, \quad (2.17a)$$

$$P_p^2 = P_p, \quad (2.17b)$$

$$\|f - P_p f\|_2 = \inf_{h \in V_p} \|f - h\|_2. \quad (2.17c)$$

Bemerkung zu Existenz: Da die V_p als abgeschlossene Teilräume eines Hilbertraums vorausgesetzt werden, gilt der Projektionssatz, d.h. jedes Element des Hilbertraums ist eindeutig zerlegbar in V_p und V_p^\perp , wobei das Element in V_p eindeutig bestimmt ist und eine beste Approximation darstellt [Alt06].

Definition 2.3.3 (Hoch- und Tiefpassräume) *Bei einer Multiskalenanalyse werden die V_p als Tiefpassräume und die W_p als Hochpassräume bezeichnet. Hierbei ist W_p das orthogonale Komplement von V_p in V_{p+1} :*

$$W_p = V_{p+1} \ominus V_p. \quad (2.18)$$

Bemerkung Da das orthogonale Komplement in einem Hilbertraum abgeschlossen ist, ist W_p ebenfalls abgeschlossen ($W_p = V_{p+1} \cap V_p^\perp$). Die Eigenschaft der Abgeschlossenheit ist wichtig, um die Existenz der Projektion auf W_p und V_p bzw. die eindeutige Zerlegbarkeit einer Funktion aus V_{p+1} in Elemente in W_p und V_p zu sichern.

Interpretation: Die eindeutige Zerlegung einer Funktion $f_{p+1} \in V_{p+1}$ mit $f_{p+1} = f_p + w_p$ (mit $f_p \in V_p$ und $w_p \in W_p$) wird als Aufspaltung in Tief- und Hochpass Komponenten interpretiert, also Zerlegung in eine "grobe Näherung" (f_p) und "Details" (w_p).

Satz 2.3.1 Sei (V_p) eine Multiskalenanalyse und bezeichne $\mathcal{B}_0(\Phi)$ eine zugehörige Riesz-Basis von V_0 . Dann ist $\mathcal{B}_p(\Phi)$ eine Riesz-Basis für V_p , und es gilt

$$V_p = \text{cl}_{L^2(\mathbb{R})} \{ \text{span} \{ \mathcal{B}_p(\Phi) \} \}. \quad (2.19)$$

Beweis Wegen Gl. (2.16c) gilt: $f \in V_p \Leftrightarrow f(2^{-p}\cdot) \in V_0 \Leftrightarrow f(2^{-p}t) = \sum_k \mathbf{f}_k \Phi(t - k) \Leftrightarrow f(t) = \sum_k \mathbf{f}_k \Phi(2^p t - k)$. Ähnlich kann man zeigen, dass die Konstanten A und B aus der Definition Gl. (2.4) für alle $\mathcal{B}_p(\Phi)$ gleich sind, also $A, B > 0$ gilt. \square

Bemerkung Dieser Satz zeigt, dass alleine durch die Angabe einer geeigneten Funktion Φ die Multiskalenanalyse festgelegt ist.

Definition 2.3.4 (Skalierungsfunktion und Wavelets) Der Vektor Φ zu einer Riesz-Basis $\mathcal{B}_0(\Phi)$ von V_0 , wird Skalierungsfunktion (oder Skalierungsvektor) genannt, d.h.

$$\Phi(t) = (\phi_1(t) \ \dots \ \phi_R(t))^T \text{ mit } \text{cl}_{L^2(\mathbb{R})} \{ \text{span} \{ \mathcal{B}_0(\Phi) \} \} = V_0. \quad (2.20)$$

Der Vektor Ψ zu einer Riesz-Basis $\mathcal{B}_0(\Psi)$ von W_0 wird mit Wavelet (oder Multiwavelet) bezeichnet, d.h.

$$\Psi(t) = (\psi_1(t) \ \dots \ \psi_R(t))^T \text{ mit } \text{cl}_{L^2(\mathbb{R})} \{ \text{span} \{ \mathcal{B}_0(\Psi) \} \} = W_0. \quad (2.21)$$

Bemerkung Da $W_0 \subseteq V_1$ gilt, existieren Funktionen $\psi_r \in V_1$ die eine Riesz-Basis von W_0 bilden. Diese sind als geeignete Linearkombination der Riesz-Basis $\mathcal{B}_1(\Phi)$ von V_1 gegeben.

Die Skalierungsfunktion zu einem Raum V_0 bzw. das Wavelet zu W_0 ist keineswegs eindeutig: orthogonale Transformationen oder ganzzahlige Verschiebungen der einzelnen Komponenten der Skalierungsfunktion bilden wieder eine Skalierungsfunktion.

Definition 2.3.5 (Hoch- und Tiefpass Filterkoeffizienten, Verfeinerungskoeffizienten) Sei Φ eine Skalierungsfunktion und Ψ ein zugehöriges Wavelet. Dann heißt die eindeutige Koeffizientenfolge $\mathbf{G} : \mathbb{Z} \rightarrow \mathbb{R}^{R \times R} : k \mapsto \mathbf{G}(k)$ der Darstellung von Φ bezüglich $\mathcal{B}_1(\Phi)$ Tiefpass-Filterkoeffizienten oder Verfeinerungskoeffizienten. Die ebenfalls eindeutige Matrixfolge $\mathbf{H} : \mathbb{Z} \rightarrow \mathbb{R}^{R \times R} : k \mapsto \mathbf{H}(k)$ zur Darstellung von Ψ bezüglich $\mathcal{B}_1(\Phi)$ wird mit Hochpass-Filterkoeffizienten bezeichnet. Es gilt also:

$$\Phi(t) = \sum_k \mathbf{G}(k) \sqrt{2} \Phi(2t - k), \quad (2.22)$$

$$\Psi(t) = \sum_k \mathbf{H}(k) \sqrt{2} \Phi(2t - k). \quad (2.23)$$

Bemerkung Die Folgen \mathbf{G} und \mathbf{H} sind eindeutig bestimmt, da die Darstellung bezüglich der Riesz-Basis $\mathcal{B}_1(\Phi)$ eindeutig ist. Sie hängen aber von der Wahl von Φ und Ψ ab, die bei einer gegebenen Multiskalenanalyse keineswegs eindeutig sind.

Hat Φ einen kompakten Träger $\text{supp}(\Phi) = [-P, P]$ mit $P \in \mathbb{N} \setminus \{0\}$, so ist die Länge von \mathbf{G} höchstens $6P - 1$. Genauer sind $\mathbf{G}(k) = 0$ für $k < -3P + 1$ oder $k > 3P - 1$.

Satz 2.3.2 (Mallat Algorithmus Teil 1: Rekonstruktion) *Sei Φ eine Skalierungsfunktion, (V_p) die zugehörige Multiskalenanalyse, Ψ ein zugehöriges Wavelet und $f \in V_p$ sowie $w \in W_p$. Dann gilt für die Koeffizientenfolgen von f , w und $w + f \in V_{p+1}$ bezüglich der Basen $\mathcal{B}_p(\Phi)$ und $\mathcal{B}_p(\Psi)$ der Zusammenhang*

$$\mathbf{f}_{p+1}(k) = \sum_l \mathbf{f}_p(l) \mathbf{G}(k - 2l) + \sum_l \mathbf{w}_p(l) \mathbf{H}(k - 2l), \quad (2.24)$$

$$\text{bzw.} \quad \mathbf{f}_{p+1} = (2 \uparrow \mathbf{f}_p) * \mathbf{G} + (2 \uparrow \mathbf{w}_p) * \mathbf{H}, \quad (2.25)$$

wobei $\mathbf{G}(k)$, $\mathbf{H}(k)$ die Hoch- und Tiefpassfilterkoeffizienten aus Def. 2.3.5 sind, sowie

$$f(t) = \sum_k \mathbf{f}_p(k) 2^{p/2} \Phi(2^p t - k) \quad (\in V_p), \quad (2.26)$$

$$w(t) = \sum_k \mathbf{w}_p(k) 2^{p/2} \Psi(2^p t - k) \quad (\in W_p), \quad (2.27)$$

$$f(t) + w(t) = \sum_k \mathbf{f}_{p+1}(k) 2^{(p+1)/2} \Phi(2^{p+1} t - k) \quad (\in V_{p+1}), \quad (2.28)$$

gegeben ist der Auftaktungsoperator $2 \uparrow$ bzw. der Faltungsoperator $*$ formal definiert ist als

$$(2 \uparrow f)(k) = \begin{cases} 0 & k \text{ ungerade} \\ f(k/2) & k \text{ gerade} \end{cases}, \quad (2.29)$$

$$(\mathbf{a} * \mathbf{B})(k) = \sum_l \mathbf{a}(l) \mathbf{B}(k - l) = \sum_l \mathbf{a}(k - l) \mathbf{B}(l). \quad (2.30)$$

Beweis Wird die Verfeinerungsgleichung Gl. (2.22) in Gl. (2.26) und Gl. (2.23) in Gl. (2.27) eingesetzt, so erhält man nach einer Variablensubstitution und Koeffizientenvergleich mit Gl. (2.28) das Ergebnis Gl. (2.24). \square

Die vektorielle Version des Mallat Algorithmus unterscheidet sich zu der skalaren Version [Mal89a] nur in der Beachtung der Reihenfolge der Multiplikations bei der Faltung mit den Filterkoeffizienten.

Bemerkung: Orthogonalität wird für die Rekonstruktion nicht vorausgesetzt. Der Syntheseteil des Mallat Algorithmus ist in Abb. 2.1 dargestellt, wo die

rekursive Struktur sichtbar wird. Sind die Koeffizienten für die Hochpassräume $w_n(k)$ für $n = p - D + 1, p - D + 2, \dots, p - 1$ und die Koeffizienten für den Tiefpassraum, der das "größte Detail" enthält bekannt, so kann daraus die Koeffizientenfolge in $f_p(k)$ – und daher auch die zugehörige Funktion in V_p – wieder exakt rekonstruiert werden.

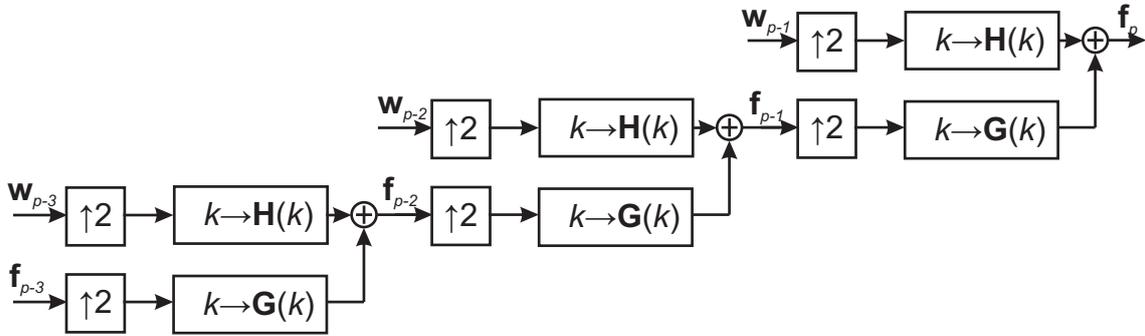


Abbildung 2.1: Der Signalflussgraph des Mallat Algorithmus Teil 1: Synthese. Da hier vektorielle Signale vorausgesetzt sind, ist die mit den Kästen dargestellte Faltung nicht mehr kommutativ. Sie symbolisieren eine Faltung "von rechts" wie in Gl. (2.30)

2.4 Orthogonale Multiskalenanalyse

Definition 2.4.1 (Orthogonale Multiskalenanalyse) Ist (V_p) eine Multiskalenanalyse von Vielfachheit R und existiert eine Funktion $\Phi = (\phi_1 \dots \phi_R)^T$, so dass die einparametrische Familie $\mathcal{B}_0(\Phi)$ eine orthogonale Riesz-Basis von V_0 ist, so heißt (V_p) orthogonale Multiskalenanalyse.

Bemerkung: Dies impliziert, dass die Konstanten A und B aus der Definition der Riesz-Basis gleich 1 sind.

Definition 2.4.2 (Orthonormale Skalierungsfunktion/Wavelet) Ist (V_p) eine orthogonale Multiskalenanalyse von Vielfachheit R , dann heißt eine zugehörige Skalierungsfunktion bzw. Wavelet orthonormal, falls die einparametrische Familie $\mathcal{B}_0(\Phi)$ bzw. $\mathcal{B}_0(\Psi)$ eine orthonormale Basis ist. $\mathcal{B}_0(\Phi)$ ist orthonormal genau dann, wenn gilt: (Bedingung für $\mathcal{B}_0(\Psi)$ is analog)

$$\langle \Phi(\cdot - k), \Phi(\cdot - l) \rangle = \int_{\mathbb{R}} \Phi(t - k) \Phi(t - l)^T dt = \delta_{kl} \mathbf{I}, \quad k, l \in \mathbb{Z}. \quad (2.31)$$

Bemerkung Diese Definition ist notwendig, denn nicht jede einparametrische Familie, die eine orthogonale Multiskalenanalyse erzeugt, ist auch selber orthonormal oder orthogonal. Der Begriff orthogonale Multiskalenanalyse stellt lediglich sicher, dass eine solche Basis existiert.

Satz 2.4.1 (Orthonormalitätsbedingungen an die Filterkoeffizienten (I)) Sei Φ eine orthonormale Skalierungsfunktion einer orthogonalen Multiskalenanalyse und \mathbf{G} die in Def. 2.3.5 definierten Filterkoeffizienten. Dann ist

$$\sum_l \mathbf{G}(l)\mathbf{G}(l+2k)^T = \delta_{0k}\mathbf{I}, \quad k \in \mathbb{Z}. \quad (2.32)$$

Beweis $\langle \Phi, \Phi(\cdot - k) \rangle = 2 \sum_{l,m} \mathbf{G}(l) \langle \Phi(2 \cdot -l), \Phi(2 \cdot -2k - m) \rangle \mathbf{G}(m)^T$. Also gilt $\langle \Phi, \Phi(\cdot - k) \rangle = \delta_k \mathbf{I} \Rightarrow \sum_m \mathbf{G}(2k + m)\mathbf{G}(m)^T = \delta_k \mathbf{I}$. \square

Satz 2.4.2 (Orthonormalitätsbedingungen an die Filterkoeffizienten (II)) Sei Φ eine orthonormale Skalierungsfunktion einer orthogonalen Multiskalenanalyse (V_p), $\Psi \in V_1$ und \mathbf{H} und \mathbf{G} die Folgen entsprechend Gl. (2.22) und Gl. (2.23). Bezeichne W_0 den von $\mathcal{B}_0(\Psi)$ erzeugten Teilraum von V_1 . Sowohl $\mathcal{B}_0(\Psi)$ ist orthonormal als auch $V_1 = W_0 \oplus V_0$ genau dann, wenn gilt:

$$\sum_l \mathbf{H}(l)\mathbf{H}(l+2k)^T = \delta_{0k}\mathbf{I}, \quad k \in \mathbb{Z}, \quad (2.33)$$

$$\sum_l \mathbf{H}(l)\mathbf{G}(l+2k)^T = \mathbf{0}, \quad k \in \mathbb{Z}, \quad (2.34)$$

$$\sum_l \mathbf{G}(n+2l)^T \mathbf{G}(m+2l) + \mathbf{H}(n+2l)^T \mathbf{H}(m+2l) = \delta_{nm}\mathbf{I} \quad n, m \in \mathbb{Z}. \quad (2.35)$$

Beweis Zu Gl. (2.33): $\mathcal{B}_0(\Psi)$ orthonormal $\Leftrightarrow \langle \Psi, \Psi(\cdot - k) \rangle = \delta_k \mathbf{I}, k \in \mathbb{Z}$. Es gilt aber: $\langle \Psi, \Psi(\cdot - k) \rangle = 2 \sum_{l,m} \mathbf{H}(l) \langle \Phi(2 \cdot -l), \Phi(2 \cdot -2k - m) \rangle \mathbf{H}(m)^T = \sum_m \mathbf{H}(2k + m)\mathbf{H}(m)^T$.

Zu Gl. (2.34): $W_0 \perp V_0 \Leftrightarrow \mathcal{B}_0(\Psi) \perp \mathcal{B}_0(\Phi) \Leftrightarrow \langle \Psi, \Phi(\cdot - k) \rangle = \mathbf{0}, k \in \mathbb{Z}$. Es gilt aber: $\langle \Psi, \Phi(\cdot - k) \rangle = 2 \sum_{l,m} \mathbf{H}(l) \langle \Phi(2 \cdot -l), \Phi(2 \cdot -2k - m) \rangle \mathbf{G}(m)^T = \sum_m \mathbf{H}(2k + m)\mathbf{G}(m)^T$.

Zu Gl. (2.35): Setzt man die Synthese- und Analysegleichung aus dem Mallat Algorithmus ineinander ein, so erhält man Gl. (2.35). Beim Beweis der Rückrichtung, kann man annehmen, dass Gl. (2.34) und Gl. (2.33) gelten, was nach obigen Argumenten die Orthonormalität von Ψ und die Orthogonalität von Φ und Ψ bedeutet. Dann folgt wieder aus dem Einsetzen der Analyse- in die Synthesegleichung des Mallat Algorithmus, dass $V_1 = W_0 \oplus V_0$. \square

Satz 2.4.3 (Gestalt des Projektionsoperators) Der Projektionsoperator $P_p : L^2(\mathbb{R}) \rightarrow V_p$ der Funktionen aus $L^2(\mathbb{R})$ auf den Unterraum V_p einer orthogonalen Multiskalenanalyse projiziert, hat für eine zugehörige orthonormale Skalierungsfunktion Φ die Gestalt

$$(P_p f)(t) = \sum_k \mathbf{f}_p(k) 2^{p/2} \Phi(2^p t - k), \quad (2.36)$$

wobei die Koeffizientenfolge $\mathbf{f}_p(k)$ von $P_p f$ in der Basis $\mathcal{B}_p(\Phi)$ gegeben ist durch

$$\mathbf{f}_p(k) = \langle f, 2^{p/2} \Phi(2^p \cdot -k) \rangle. \quad (2.37)$$

Beweis Da $P_p f \in V_p$ gilt $P_p f = \sum_k \mathbf{f}_p(k) 2^{p/2} \Phi(2^p \cdot -k)$ mit geeigneten $\mathbf{f}_p(k) \in \mathbb{R}^{1 \times R}$. Es folgt: $\langle P_p f, 2^{p/2} \Phi(2^p \cdot -l) \rangle = \mathbf{f}_p(l)$, da $\langle 2^{p/2} \Phi(2^p \cdot -k), 2^{p/2} \Phi(2^p \cdot -l) \rangle = \delta_{kl} \mathbf{I}$. Gleichzeitig ist aber $\langle P_p f, \Phi(2^p \cdot -l) \rangle = \langle f, \Phi(2^p \cdot -l) \rangle$, da $P_p f - f$ nach Definition orthogonal zu V_p ist, also insbesondere auch zu $\Phi(2^p \cdot -l)$. \square

Bemerkung Die Voraussetzung, dass $f \in {}^2(\mathbb{R})$ liegt, ist dabei wesentlich. Durch Vergleich mit Satz 2.2.2 ist ersichtlich, dass die Formel zur Bestimmung der Koeffizienten in $\mathbb{S}(\Phi)$ gleich zu Gl. (2.37) für $p = 0$ ist. Ist also anders gesagt, $f \in \mathbb{S}(\Phi)$ aber nicht in $L^2(\mathbb{R})$ so, werden also die Koeffizienten der Darstellung in $\mathbb{S}(\Phi)$ berechnet, was nicht der Projektion auf V_0 entspricht.

Satz 2.4.4 (Hoch- und Tiefpass Filterkoeffizienten für orthogonale Multiskalenanalysen) Sei Φ eine orthonormale Skalierungsfunktion und Ψ ein zugehöriges Wavelet. Dann sind die Filterkoeffizienten aus Def. 2.3.5 gegeben durch

$$\mathbf{G}(k) = 2^{-1/2} \langle \Phi(\cdot/2), \Phi(\cdot - k) \rangle = \sqrt{2} \langle \Phi, \Phi(2 \cdot -k) \rangle, \quad (2.38)$$

$$\mathbf{H}(k) = 2^{-1/2} \langle \Psi(\cdot/2), \Phi(\cdot - k) \rangle = \sqrt{2} \langle \Psi, \Phi(2 \cdot -k) \rangle. \quad (2.39)$$

Beweis Da $\Phi(\cdot/2) \in V_{-1} \subset V_0$ liegt, gilt nach Gl. (2.37): $\Phi(t/2) = \sum_k \langle \Phi(\cdot/2), \Phi(\cdot - k) \rangle \Phi(t - k)$. Also folgt durch die Eindeutigkeit der $\mathbf{G}(k)$ durch Vergleich mit Gl. (2.22) direkt die erste Form von Gl. (2.38). Die zweite Form ergibt sich durch eine einfache Variablensubstitution. Die Herleitung für $\mathbf{H}(k)$ ist analog. \square

Satz 2.4.5 (Mallat Algorithmus Teil 2: Analyse) Sei Φ eine orthonormale Skalierungsfunktion, (V_p) die zugehörige orthogonale Multiskalenanalyse, Ψ ein zugehöriges orthonormales Wavelet. Sei weiter $f + w \in V_{p+1}$ mit $f \in V_p$ sowie $w \in W_p$. Dann gilt für die Koeffizientenfolgen von f und w bezüglich der Basis $\mathcal{B}_0(\Phi)$ und $\mathcal{B}_0(\Psi)$ abhängig von der Koeffizientenfolge von $w + f$ der Zusammenhang

$$\mathbf{f}_p(k) = \sum_l \mathbf{f}_{p+1}(l) \mathbf{G}(l - 2k)^T, \quad (2.40)$$

$$\mathbf{w}_p(k) = \sum_l \mathbf{f}_{p+1}(l) \mathbf{H}(l - 2k)^T, \quad (2.41)$$

$$\text{bzw. } \mathbf{f}_p = 2 \downarrow (\mathbf{f}_{p+1} * \mathbf{G}(-\cdot)^T), \quad (2.42)$$

$$\mathbf{w}_p = 2 \downarrow (\mathbf{f}_{p+1} * \mathbf{H}(-\cdot)^T), \quad (2.43)$$

wobei $G(k)$, $H(k)$ die Hoch- und Tiefpassfilterkoeffizienten aus Def. 2.3.5 sind, sowie

$$f(t) + w(t) = \sum_k \mathbf{f}_{p+1}(k) 2^{(p+1)/2} \Phi(2^{p+1}t - k), \quad (\in V_{p+1}) \quad (2.44)$$

$$f(t) = \sum_k \mathbf{f}_p(k) 2^{p/2} \Phi(2^p t - k), \quad (\in V_p) \quad (2.45)$$

$$w(t) = \sum_k \mathbf{w}_p(k) 2^{p/2} \Psi(2^p t - k), \quad (\in W_p) \quad (2.46)$$

gegeben ist, und der formalen Definition des Abtastoperators $2 \downarrow$

$$(2 \downarrow f)(k) = f(2k). \quad (2.47)$$

Beweis Es gilt $f = P_p(w + f)$, also folgt unter Ausnutzung der Orthogonalität Gl. (2.37) und Verfeinerbarkeit Gl. (2.22): $\mathbf{f}_p(k) = \langle f + w, 2^{p/2} \Phi(2^p \cdot -k) \rangle = \sum_l \langle f + w, 2^{(p+1)/2} \Phi(2^{p+1} \cdot -2k - l) \rangle G(l)^T$. Wird nun $f + w$ durch Gl. (2.44) ersetzt, dann erhält man unter Ausnutzung der Orthonormalität $\mathbf{f}_p(k) = \sum_l \mathbf{f}_{p+1}(2k + l) G(l)^T$. Da W_p ein abgeschlossener Teilraum von $L^2(\mathbb{R})$ mit orthonormaler Basis $\mathcal{B}_p(\Psi)$ ist, gilt $\mathbf{w}_p(k) = \langle f + w, 2^{p/2} \Psi(2^p \cdot -k) \rangle = \sum_l \langle f + w, 2^{(p+1)/2} \Phi(2^{p+1} \cdot -2k - l) \rangle H(l)^T$. Wird nun $f + w$ wieder durch Gl. (2.44) ersetzt, dann erhält man unter Ausnutzung der Orthonormalität $\mathbf{w}_p(k) = \sum_l \mathbf{f}_{p+1}(2k + l) H(l)^T$ \square

Die vektorielle Version des Mallat Algorithmus unterscheidet sich zu der skalaren Version [Mal89a] nur in der Beachtung der Reihenfolge und der zusätzlichen Transposition bei der Faltung mit den Filterkoeffizienten.

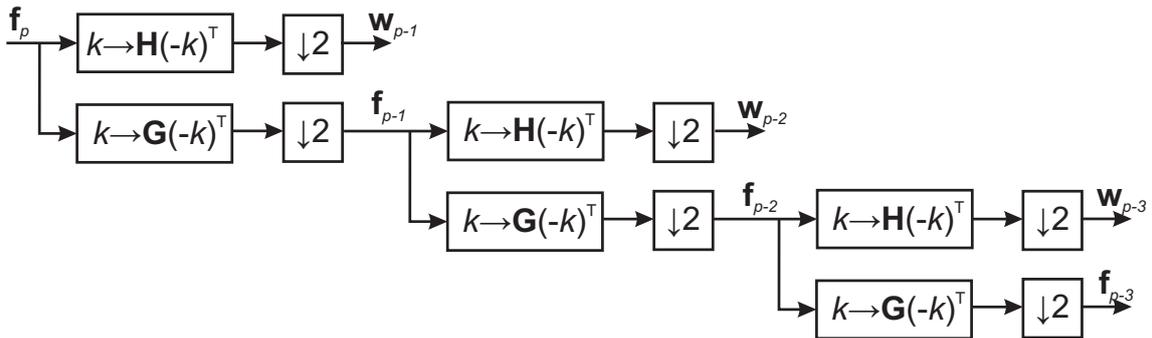


Abbildung 2.2: Der Signalflussgraph des Mallat Algorithmus Teil 2: Analyse. Wie auch in Abb. 2.1 symbolisieren die Kästen eine "Faltung von rechts". Im Gegensatz zur Synthese wird eine orthonormale Multiskalenanalyse vorausgesetzt.

Der Syntheseteil des Mallat Algorithmus erlaubt es, eine Funktion f aus einem hinreichend großen Raum V_n sukzessive in Hoch- und Tiefpassanteile

(in V_p und W_p mit $p < n$) zu zerlegen. Dazu wird jeweils der sich ergebende Tiefpassanteil ein weiteres Mal in Hoch- und Tiefpassanteil aufgespalten. Die endgültige Zerlegung besteht dann aus einem einzigen (übrigen) Tiefpassanteil, der das "größte Detail" symbolisiert und den Hochpassanteilen der verschiedenen Skalen.

Eine wichtige Eigenschaft für die Anwendung ist, dass die Filterkoeffizienten die einzig benötigte Information über die Multiskalenanalyse sind. Kenntnis der Skalierungsfunktion Φ und des Wavelets Ψ ist nicht notwendig.

Bemerkenswert ist ebenfalls die Beziehung zu dem Abtasttheorem. Obwohl die Voraussetzung des Abtasttheorems bei der stattfindenden Abtastung um 2 verletzt ist, ermöglicht die spezielle Struktur von Skalierungsfunktion und Wavelet eine ideale Rekonstruktion. Eine Aufspaltung eines Signals in Hoch- und Tiefpassanteil mit einem idealen Hoch- und Tiefpass würde den Frequenzbereich gerade bei $f_a/4$ (mit f_a der Abtastfrequenz) teilen. Bei einer idealen Filterung ist es dann möglich, die Datenrate zu dezimieren, d.h. eine Abtastung mit Faktor 2 vorzunehmen, wie es hier auch geschieht. Dies ist nur durch die vorherige ideale Filterung im Einklang mit dem Abtasttheorem, das besagt das eine ideale Rekonstruktion möglich ist, wenn die sich ergebende Signalbandbreite kleiner gleich $f_a/2$. Andererseits verhindert die Dezimation so Redundanz in den Abtastwerten. Das Bemerkenswerte ist, dass bei einer Multiskalenanalyse sich diese Struktur automatisch ergibt, wenn nur die Axiome der Multiskalenanalyse erfüllt sind. Anschaulich ermöglicht die Beziehung zwischen der Skalierungsfunktion und dem Wavelet die ideale Rekonstruktion, obwohl keine Filterung vorliegt die die Bandbreite der Signale genau halbiert.

2.5 Polynome

Von besonderer Wichtigkeit in dieser Arbeit sind die Polynome, weswegen dafür eine einheitliche Schreibweise eingeführt wird.

Sei der Raum der Polynome vom Grad kleiner oder gleich $N - 1$ mit Π_{N-1} bezeichnet. Es gilt $\Pi_{N-1} \not\subset L^2(\mathbb{R})$ für alle $N \in \{1, 2, \dots\}$.

2.5.1 Vektor der Monome

Die Monome vom Grad 0 bis $N - 1$ werden in dem Vektor $\mathbf{p} : \mathbb{R} \rightarrow \mathbb{R}^N : t \mapsto \mathbf{p}(t)$ zusammengefasst:

$$\mathbf{p}(t) = \begin{pmatrix} 1 \\ t \\ t^2 \\ \vdots \\ t^{N-1} \end{pmatrix}. \quad (2.48)$$

Die Dimension N wird in der Schreibweise \mathbf{p} unterdrückt, da es entweder aus dem Kontext klar oder nicht relevant ist, da zudem die Komponenten des Vektors nicht von N abhängen. Mithilfe dieser Vektorschreibweise lassen sich viele Eigenschaften der Polynome kompakt ausdrücken.

Satz 2.5.1 (Verschiebung und Skalierung des Polynomvektors). *Es gilt für die Definition Gl. (2.48) folgende Verschiebungs- und Skalierungseigenschaft:*

$$\mathbf{p}(t - \tau) = \mathbf{T}(\tau)\mathbf{p}(t), \quad \tau \in \mathbb{R}, \quad (2.49a)$$

$$\mathbf{p}(at) = \mathbf{S}(a)\mathbf{p}(t), \quad a \in \mathbb{R}, \quad (2.49b)$$

mit den Definitionen

$$\mathbf{S} : \mathbb{R} \rightarrow \mathbb{R}^{N \times N} : a \rightarrow \mathbf{S}(a) := \text{diag}(1, a, \dots, a^{N-1}) \quad (2.49c)$$

$$\mathbf{T} : \mathbb{R} \rightarrow \mathbb{R}^{N \times N} : \tau \rightarrow (\mathbf{T}(\tau))_{ij} := \begin{cases} 0 & j > i \\ \binom{i}{j} (-\tau)^{i-j} & j \leq i \end{cases} \quad i, j = 0, \dots, N-1. \quad (2.49d)$$

Es folgt direkt aus der Definition:

$$\mathbf{S}(a)^{-1} = \mathbf{S}(1/a), \quad a \in \mathbb{R} \setminus \{0\}, \quad (2.49e)$$

$$\mathbf{S}(ab) = \mathbf{S}(a)\mathbf{S}(b), \quad a, b \in \mathbb{R}, \quad (2.49f)$$

$$\mathbf{S}(1) = \mathbf{I}, \quad (2.49g)$$

$$\mathbf{T}(\tau)^{-1} = \mathbf{T}(-\tau), \quad \tau \in \mathbb{R}, \quad (2.49h)$$

$$\mathbf{T}(\tau + \sigma) = \mathbf{T}(\tau)\mathbf{T}(\sigma), \quad \tau, \sigma \in \mathbb{R}, \quad (2.49i)$$

$$\mathbf{T}(0) = \mathbf{I}, \quad (2.49j)$$

$$\mathbf{S}(a)\mathbf{T}(\tau) = \mathbf{T}(a\tau)\mathbf{S}(a), \quad a, \tau \in \mathbb{R}. \quad (2.49k)$$

Wie auch in $\mathbf{p}(t)$ wird der maximale Grad $N - 1$ der Polynome in den Matrizen $\mathbf{S}(a)$ und $\mathbf{T}(\tau)$ unterdrückt, da die Einträge der Matrizen nicht von N abhängen. So sind $\mathbf{S}(a)$ und $\mathbf{T}(\tau)$ immer Teilmatrizen von den Matrizen für größeres N .

Beispiel Für Grad 3, also $N = 4$, haben die Matrizen $\mathbf{T}(\tau)$ und $\mathbf{S}(a)$ die folgende Gestalt:

$$\mathbf{T}(\tau) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\tau & 1 & 0 & 0 \\ \tau^2 & -2\tau & 1 & 0 \\ -\tau^3 & 3\tau^2 & -3\tau & 1 \end{pmatrix} \quad \mathbf{S}(a) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a^2 & 0 \\ 0 & 0 & 0 & a^3 \end{pmatrix} \quad (2.50)$$

2.5.2 Legendre-Polynome

Die auf dem Intervall $[0, 1]$ bezüglich des Lebesgue-Maßes orthogonalen Polynome sind die verschobenen Legendre Polynome. Sie werden üblicherweise

mit P_n^* bezeichnet [Wei99]. Versehen mit einer Normierung werden diese Polynome vom Grad kleiner N in dem Vektor $\mathbf{l} : \mathbb{R} \rightarrow \mathbb{R}^N : t \mapsto \mathbf{l}(t)$ zusammengefasst:

$$\mathbf{l}(t) = \begin{pmatrix} 1 \\ \sqrt{3}(2t - 1) \\ \sqrt{5}(6t^2 - 6t + 1) \\ \vdots \\ \sqrt{2N - 1} \sum_{i=0}^{N-1} (-1)^{i+N-1} \binom{N-1}{i} \binom{i+N-1}{i} t^i \end{pmatrix}. \quad (2.51)$$

Die ersten 4 verschobenen Legendre-Polynome sind in Abb. 2.3 dargestellt.

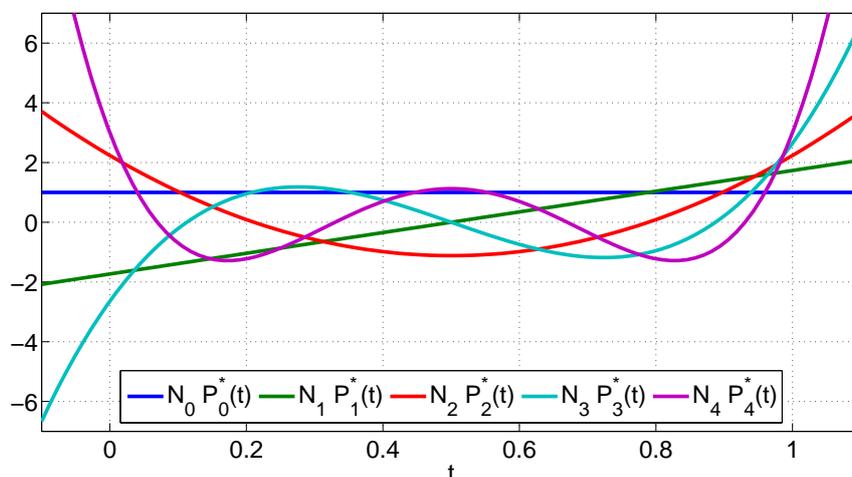


Abbildung 2.3: Die verschobenen Legendre Polynome vom Grad ≤ 4 (also $N = 5$) mit entsprechender Normierung $N_i = \sqrt{2i + 1}$, damit sie ein Orthonormalsystem auf $[0, 1]$ bilden.

Es ist klar, dass sich $\mathbf{l}(t)$ als lineare Transformation des Polynomvektors $\mathbf{p}(t)$ darstellen lässt:

$$\mathbf{l}(t) = \mathbf{L}\mathbf{p}(t) \quad (2.52)$$

mit

$$(\mathbf{L})_{ij} = \begin{cases} 0 & j > i \\ \sqrt{2i + 1} (-1)^{i+j} \binom{i}{j} \binom{i+j}{j} & j \leq i \end{cases} \quad i, j = 0..N. \quad (2.53)$$

\mathbf{L} ist eine untere Dreiecksmatrix und invertierbar, da auf der Diagonalen keine Nullen vorhanden sind. Die Inverse Matrix \mathbf{L}^{-1} hat dann ebenfalls eine Dreiecksgestalt.

Beispiel Die Matrix L hat für Grad 4, also $N = 5$ folgende Gestalt

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -\sqrt{3} & 2\sqrt{3} & 0 & 0 & 0 \\ \sqrt{5} & -6\sqrt{5} & 6\sqrt{5} & 0 & 0 \\ -\sqrt{7} & 12\sqrt{7} & -30\sqrt{7} & 20\sqrt{7} & 0 \\ 3 & -60 & 270 & -420 & 210 \end{pmatrix} \quad (2.54)$$

und enthält in der i -ten Zeile, die Koeffizienten vor t^j in dem normierten Legendrepolynom $P_i^*(t)$.

3 Vorüberlegungen zur Konstruktion von Multiskalenanalysen

3.1 Bedingungen zur Konstruktion von Multiskalenanalysen

Aus Abschnitt 2.3 ist bekannt, dass eine geeignete Funktion Φ eine Multiskalenanalyse vollständig beschreibt, da durch die Angabe von Φ die Basen von allen V_p festgelegt sind. Bei der Wahl von Φ ist lediglich sicherzustellen, dass die Axiome der Multiskalenanalyse erfüllt sind (siehe Def. 2.3.1). Um den Mallat Algorithmus anwenden zu können, muss Φ zudem orthonormal sein. (Diese Bedingung kann abgeschwächt werden, falls statt Orthonormalität Biorthonormalität gefordert wird).

Als Ergebnis dieses Unterkapitels kann festgehalten werden, dass falls Φ einen kompakten Träger hat, es genau dann eine orthogonale Multiskalenanalyse erzeugt, wenn die Verfeinerbarkeitsgleichung und Orthonormalitätsbedingung für Φ gilt.

3.1.1 Verfeinerbarkeit und Verschachtelung

Gilt für eine Skalierungsfunktion Φ die Verfeinerungsgleichung Gl. (2.22)

$$\Phi(\cdot/2) = \sqrt{2} \sum_k \mathbf{G}(k) \Phi(\cdot - k), \quad (3.1)$$

so gilt mit $V_p := \text{cl}_{L^2(\mathbb{R})} \{\text{span} \{\mathcal{B}_p(\Phi)\}\}$ das Verschachtelungsaxiom Gl. (2.16a) und das Verfeinerbarkeitsaxiom Gl. (2.16c). Das Verschachtelungsaxiom gilt, weil die Basis $\mathcal{B}_p(\Phi)$ von V_p in $\mathcal{B}_{p+1}(\Phi) \subset V_{p+1}$ darstellbar ist, also $V_p \subset V_{p+1}$ ist. Aufgrund der Definition $V_p := \text{cl}_{L^2(\mathbb{R})} \{\text{span} \{\mathcal{B}_p(\Phi)\}\}$, gilt auch die Verfeinerungsgleichung $f \in V_p \Leftrightarrow f(2^{-p}\cdot) \in V_0$ für $p \in \mathbb{Z}$.

3.1.2 Orthogonalität

Ist $\mathcal{B}_0(\Phi)$ orthonormal, also gilt

$$\langle \Phi, \Phi(\cdot - k) \rangle = \delta_k \mathbf{I}, \quad (3.2)$$

so ist $\mathcal{B}_0(\Phi)$ eine Riesz-Basis von V_0 .

3.1.3 Approximationseigenschaft

Hat Φ einen kompakten Träger, so folgt die Approximationseigenschaft Gl. (2.16b) der Multiskalenanalyse aus den übrigen Forderungen der Multiskalenanalyse (siehe Kommentar bei Def. 2.3.1).

3.1.4 Konstruktion eines Wavelets zur Skalierungsfunktion

Liegt eine orthogonale Multiskalenanalyse mit orthonormaler Skalierungsfunktion Φ vor, so gibt es allgemeine Verfahren, um ein zugehöriges orthonormales Wavelet zu konstruieren [SS95, DGHM96, LLS96]. Diese Verfahren basieren darauf eine paraunitäre Erweiterung einer Matrixfunktion zu finden. Der Ansatz wird hier kurz skizziert.

In Anlehnung an die für skalare Folgen definierte z -Transformation, wird für $m \in \mathbb{Z}$ die von den Tiefpasskoeffizienten $G(k)$ abhängige komplexe Funktion

$$\mathbf{G}^{(m)} : \mathbb{C} \rightarrow \mathbb{R}^{R \times R} : z \mapsto \mathbf{G}^{(m)}(z) = \sum_k \mathbf{G}(m + 2k)z^n \quad (3.3)$$

eingeführt. Diese Funktionen bilden die Polyphasenmatrix

$$\mathbf{G} : \mathbb{C} \rightarrow \mathbb{R}^{R \times 2R} : z \mapsto \mathbf{G}(z) = \left(\mathbf{G}^{(0)}(z) \mid \mathbf{G}^{(1)}(z) \right). \quad (3.4)$$

Es fällt dann auf, dass diese Matrix aufgrund der Orthonormalität von Φ paraunitär ist, d.h.

$$\mathbf{G}(z)\mathbf{G}(z)^H = \mathbf{I}, \quad \text{für } |z| = 1 \text{ mit } z \in \mathbb{C} \quad (3.5)$$

gilt. Grund: (Hier ohne Berücksichtigung von Konvergenz, die aber bei Skalierungsfunktionen mit kompaktem Träger stets gesichert ist). Es gilt

$\mathbf{G}(z)\mathbf{G}(z)^H = \mathbf{G}^{(0)}(z)\mathbf{G}^{(0)}(z)^H + \mathbf{G}^{(1)}(z)\mathbf{G}^{(1)}(z)^H$ was sich für $|z| = 1$ zu $\mathbf{G}(z)\mathbf{G}(z)^H = \sum_l (\sum_k \mathbf{G}(k)\mathbf{G}(k+2l)) z^l$ auflöst. Es ist also erkennbar, dass $\mathbf{G}(z)\mathbf{G}(z)^H = \mathbf{I}$, $|z| = 1$ äquivalent zur notwendigen Orthonormalitätsbedingung Gl. (2.32) ist.

Falls nun die paraunitäre Matrix $\mathbf{G}(z)$ zu einer quadratischen paraunitären Matrix erweitert werden kann, so lässt sich aus den hinzugekommenen Einträgen eine Koeffizientenfolge $\mathbf{H}(k)$ extrahieren, die auch die anderen Bedingungen Gl. (2.33) und Gl. (2.34) für Orthonormalität des Wavelets und Orthogonalität von Wavelet und Skalierungsfunktion erfüllen.

Satz 3.1.1 (Konstruktion des Wavelets aus Φ) *Ist Φ eine orthonormale Skalierungsfunktion zu einer Multiskalenanalyse und bezeichne $\mathbf{G}(z)$ die in Gl. (3.4) definierte paraunitäre Polyphasenmatrix. Die Erweiterung auf die quadratische paraunitäre Matrix $\tilde{\mathbf{G}}(z)$ ergibt dann das zugehörige Wavelet Ψ , wobei*

$$\tilde{\mathbf{G}}(z) = \begin{pmatrix} \mathbf{G}^{(0)}(z) & \mathbf{G}^{(1)}(z) \\ \mathbf{H}^{(0)}(z) & \mathbf{H}^{(1)}(z) \end{pmatrix} \quad (3.6)$$

gilt. Die Funktionen $\mathbf{H}^{(0)}(z)$, $\mathbf{H}^{(1)}(z)$ sind dabei in analoger Weise wie $\mathbf{G}^{(0)}(z)$ und $\mathbf{G}^{(1)}(z)$ definiert und liefern die geraden und ungeraden Elemente der Folge \mathbf{H} (und so auch Ψ), welche dann Gl. (2.33) und Gl. (2.34) erfüllt. Diese sind aufgrund der Orthonormalität von Φ notwendig und hinreichend für die Orthonormalität von $\mathcal{B}_0(\Psi)$ und der Orthogonalität von $\mathcal{B}_0(\Psi)$ und $\mathcal{B}_0(\Phi)$.

Beweis Für den Beweis und Konstruktion von $\tilde{G}(z)$ siehe [LLS96]. \square .

Bemerkung Die Rekonstruktionsbedingung Gl. (2.35) muss für das sich ergebende Wavelet noch geprüft werden, oder aus anderen Eigenschaften abgeleitet werden.

3.2 Forderungen an eine Multiskalenanalyse

Wie gesehen, ist die Konstruktion einer Multiskalenanalyse durch die Konstruktion einer geeigneten Skalierungsfunktion bestimmt. Die gewünschten Eigenschaften der Multiskalenanalyse werden dabei durch entsprechende Forderungen an die Skalierungsfunktion sichergestellt. Das sind im Wesentlichen:

- Stetigkeit der Skalierungsfunktion (impliziert Stetigkeit des Wavelets),
- Kompaktheit des Trägers der Skalierungsfunktion,
- Verschwindende Momente des Wavelets,
- Verschwindende Momente der Skalierungsfunktion,
- Symmetrie der Filterkoeffizientenfolge.

3.2.1 Stetigkeit der Skalierungsfunktion

Ist die Skalierungsfunktion Φ einer Multiskalenanalyse stetig, so sind die Basisfunktionen $\mathcal{B}_p(\Phi)$ ebenfalls stetig. Haben die Skalierungsfunktionen kompakten Träger, so sind auch alle Funktionen in $V_p \subset \mathcal{S}(\Phi)$ stetig, da auf einem festen Intervall $[I_1, I_2]$ nur endlich viele Basiselemente einen Beitrag zu $f = \sum_k f(k)\Phi(t - k)$ liefern. Zudem impliziert dann die Stetigkeit von Φ auch die Stetigkeit des Wavelets Ψ , da die einzelnen Komponenten von $\psi_r \in V_1$ liegen. Zusammenfassend gilt also:

Satz 3.2.1 *Ist Φ eine stetige Skalierungsfunktion mit kompaktem Träger, so ist die Multiskalenanalyse ebenfalls stetig, d.h. es gilt:*

$$f \in V_p \Rightarrow f \text{ stetig.} \quad (3.7)$$

Zudem ist ein zugehöriges Wavelet Ψ ebenfalls stetig.

3.2.2 Kompakter Träger

Falls die Skalierungsfunktion einer Multiskalenanalyse einen kompakten Träger besitzt, so ergeben sich einige Vereinfachungen in der Theorie, da die meisten auftauchenden Summen endliche Summen werden und Konvergenzbetrachtungen nicht gemacht werden müssen. Weiterhin ist die in Def. 2.3.1 erwähnte Eigenschaft zu nennen, nach der die Approximationseigenschaft einer Multiskalenanalyse aus den übrigen Axiomen der Multiskalenanalyse folgt. Die für die Praxis wichtigste Eigenschaft ist jedoch, dass die Filterkoeffizienten $G(k)$ und $H(k)$ für Skalierungsfunktionen mit kompaktem Träger dann endlich lange Folgen sind, wodurch die Faltung mit diesen Folgen auch mit einem Computer berechenbar wird. Es werden dann FIR (*Finite Impulse Response*) Filter implementiert. Diese sind immer stabil und können auch die Eigenschaft der Linearphasigkeit haben (siehe folgenden Abschnitt über Symmetrie). Falls unendlich lange Folgen vorliegen, gelten stärkere Bedingungen an die Berechenbarkeit. So muss die z -Transformierte der Folgen G und H eine (in z) rationale Funktion sein, damit das Filter in dem herkömmlichen Differenzenschema für IIR (*Infinite Impulse Response*) Filter berechnet werden kann (Dies ist das diskrete Analogon zu einer Differentialgleichung). Siehe hierzu Kap. 3 in [Vai93]. Zudem haben diese IIR Filter auch nur eine für $k \rightarrow \infty$ nichttriviale Impulsantwort (diese wäre gleich G oder H); für kleine k bricht die Folge irgendwann ab. Falls die Folge $G(k)$ beidseitig unendlich lang ist, so ist Berechenbarkeit im Computer nur unter sehr einschränkenden Bedingungen möglich, zudem werden bei der Zerlegung die Tief- und Hochpasssignale unendlich lang, auch wenn das zu analysierende Signal endlich lang ist.

3.2.3 Approximationsordnung K

Ein wichtiger Begriff ist die Approximationsordnung einer Multiskalenanalyse. Sie gibt an, wie "schnell" die Approximationen einer Funktion f in V_p besser wird, wenn die Räume V_p größer werden ($p \rightarrow \infty$).

Definition 3.2.1 (Approximationsordnung einer Multiskalenanalyse) Sei Φ eine Skalierungsfunktion und (V_p) die zugehörige Multiskalenanalyse. Diese hat Approximationsordnung K , falls es für jedes genügend glatte $f \in L^2(\mathbb{R})$ eine von p unabhängige Konstante $C = C(f)$ gibt, mit

$$\inf_{g \in V_p} \|f - g\|_2 \leq C(2^{-p})^K \quad p \in \mathbb{Z}. \quad (3.8)$$

Dabei bezeichnet $\|\cdot\|_2$ die Norm im $L^2(\mathbb{R})$.

Satz 3.2.2 (Charakterisierung der Approximationsordnung) Eine Multiskalenanalyse von Vielfachheit R mit einer Skalierungsfunktion Φ mit kompaktem

Träger hat genau dann Approximationsordnung K , wenn

$$\Pi_{K-1} \subset \mathbb{S}(\Phi) \quad (3.9)$$

gilt. Dies bedeutet, dass alle Polynome vom Grad $< K$ in der Basis $\mathbb{B}_0(\Phi)$ dargestellt werden können, also Koeffizienten $\mathbf{p}_n(k) \in \mathbb{R}^{1 \times R}$ existieren, sodass

$$t^n = \sum_k \mathbf{p}_n(k) \Phi(t - k) \quad (3.10)$$

gilt.

Beweis Siehe [Jia97]. □

Bemerkung Nur dadurch, dass Φ kompakten Träger hat, ist der Raum der "unendlichen Linearkombinationen" $\mathbb{S}(\Phi)$ wohldefiniert, und die auftretende Summe Gl. (3.10) ist endlich, was zeigt, dass dieses nützliche Kriterium nur gilt, weil ein kompakter Träger von Φ vorliegt.

3.2.4 Verschwindende Momente der Skalierungsfunktion

Eine ebenfalls wichtige Eigenschaft ist die Forderung von möglichst vielen verschwindenden Momenten der Skalierungsfunktion. Das nullte Moment wird im Gegensatz zu der Definition der verschwindenden Momente des Wavelets hierbei nicht berücksichtigt.

Trotzdem ist es nützlich zur Bestimmung der Koeffizienten in V_p eines Initialen f , möglichst viele verschwindende höhere Momente zu haben, wie der untenstehende Satz zeigt. Bevor der Mallat Algorithmus angewendet werden kann, muss ein f bezüglich der Basis $\mathcal{B}_n(\Phi)$ dargestellt werden. Mathematisch ist hierzu Kenntnis aller kontinuierlichen Skalarprodukte zwischen f und allen Elementen der Basis notwendig. Dies ist mit einem Computer nicht zu berechnen, da ein kontinuierliches f in voller Allgemeinheit nicht sinnvoll abgespeichert werden kann. Üblicherweise werden äquidistante "Abtastwerte" $f(t_i)$ von f vorliegen, aus denen dann die Koeffizienten in V_n berechnet werden müssen. Um den Fehler gering zu halten, benötigt man dafür verschwindende Momente der Skalierungsfunktion, wie Satz 3.2.3 zeigt.

Definition 3.2.2 (Momentenvektoren und Momentenmatrix der Skalierungsfunktion) Das l -te Moment $\mathbf{m}_l \in \mathbb{R}^R$ der Skalierungsfunktion Φ einer Multiskalenanalyse mit Vielfachheit R wird wie folgt definiert:

$$\mathbf{m}_l = \langle \Phi, (\cdot)^l \rangle = \int_{\mathbb{R}} \Phi(t) t^l dt. \quad (3.11)$$

Die Momentenmatrix fasst die ersten N Momente zusammen:

$$\mathbf{M} = \begin{pmatrix} \mathbf{m}_0^T \\ \mathbf{m}_1^T \\ \vdots \\ \mathbf{m}_{N-1}^T \end{pmatrix} = \langle \mathbf{p}, \Phi \rangle. \quad (3.12)$$

Bemerkung (*Matrix der nichtzentralen Momente*) Die Matrix der nichtzentralen Momente M_τ einer Funktion Φ lässt sich durch die Matrix der zentralen Momente M ausdrücken. Mit

$$M_\tau = \langle \mathbf{p}(\cdot - \tau), \Phi \rangle, \quad (3.13)$$

gilt

$$M_\tau = \mathbf{T}(\tau) \langle \mathbf{p}, \Phi \rangle = \mathbf{T}(\tau)M. \quad (3.14)$$

Wenn die Skalierungsfunktion Φ einer orthogonalen Multiskalenanalyse L verschwindende Momente besitzt, so gilt folgende Aussage über das Wachstum der Koeffizienten bezüglich einer Basisentwicklung. Der Satz ist eine direkte vektorielle Verallgemeinerung aus [BCR91], wo er für skalare Skalierungsfunktionen formuliert ist.

Satz 3.2.3 (Wachstum der Koeffizienten in V_p) *Ist (V_p) eine orthogonale Multiskalenanalyse und besitzt eine zugehörige orthonormale Skalierungsfunktion Φ L verschwindende (nicht notwendigerweise zentrale) Momente, d.h. gilt $\mathbf{m}_{l,\tau} = \mathbf{0}$ für $l = 1 \dots L$ mit $\mathbf{m}_{l,\tau} = \int_{\mathbb{R}} \Phi(t)(t - \tau)^l dt$, und ist $f \in L^2(\mathbb{R})$ eine Funktion die mit ihrer Taylorreihe übereinstimmt, so gilt für die Koeffizienten $\mathbf{f}_p(k)$ der Approximation von f in V_p bezüglich der Basis $\mathcal{B}_p(\Phi)$:*

$$\mathbf{f}_p(k) = 2^{-p/2} f(2^{-p}k + \tau) \mathbf{m}_0^T + O(2^{-p(L+1/2)}), \quad (3.15)$$

mit

$$P_p f = \sum_k \mathbf{f}_p(k) \Phi(\cdot - k). \quad (3.16)$$

Beweis Nach Gl. (2.37) gilt $\mathbf{f}_p(k) = \langle f, 2^{p/2} \Phi(2^p \cdot - k) \rangle = 2^{p/2} \langle f(\cdot + 2^{-p}k), \Phi(2^p \cdot) \rangle = 2^{p/2} \sum_{l=0}^{\infty} \frac{f^{(l)}(2^{-p}k + \tau)}{l!} \langle (\cdot - \tau)^l, \Phi(2^p \cdot) \rangle = \sum_{l=0}^{\infty} \frac{f^{(l)}(2^{-p}k + \tau)}{l!} 2^{-p(l+1/2)} \mathbf{m}_{l,\tau}^T$. Mit $\mathbf{m}_{l,\tau} = \mathbf{0}$ für $l = 1..L$ und $\mathbf{m}_{0,\tau} = \mathbf{m}_0$ folgt Gl. (3.15). \square

Bemerkung Dieser Satz besagt also, dass die Approximation erster Ordnung für die Koeffizienten $\mathbf{f}_p(k)$ einer $L^2(\mathbb{R})$ Funktion f in V_p im Wesentlichen durch die diskrete Abtastwerte $f_k = f(2^{-p}k + \tau)$ gegeben ist. Je größer p (also umfangreicheres V_p) ist und je mehr verschwindende Momente die Skalierungsfunktion besitzt, desto kleiner ist der Fehler durch diese Approximation. Zudem ist kein wesentlicher Unterschied vorhanden, ob zentrale verschwindende Momente oder nichtzentrale verschwindende Momente vorliegen: es verschiebt sich lediglich der Abtastpunkt.

3.2.5 Symmetrie

Die Symmetrie der Filterkoeffizienten bzw. der Skalierungsfunktion und des Wavelets hat zwei Vorteile:

- Reduzierter Rechenaufwand, da in der Faltungsoperation im Mallat Algorithmus dann der Rechenaufwand halbiert wird. Dies gilt zumindest dann, wenn die Symmetrie – bei sonstigen gleichen Eigenschaften – nicht die Länge der Skalierungsfunktion verdoppelt.
- Linearphasigkeit der Filter G und H . Auf diese Eigenschaft wird näher eingegangen.

Definition 3.2.3 (linearphasiges Filter) *Ein Filter mit der Impulsantwort $\mathbb{Z} \rightarrow \mathbb{R}^{N \times M} : k \rightarrow \mathbf{H}(k)$ heißt linearphasig, falls der Phasengang linear mit der normierten Frequenz $\Omega \in [-\pi, \pi]$ ist:*

$$\angle \mathbf{H}(e^{i\Omega}) = \text{linear bezüglich } \Omega \quad \Omega \in [-\pi, \pi] \quad (3.17)$$

mit der Übertragungsfunktion des Filters

$$\mathbf{H}(z) = \sum_k \mathbf{H}(k) z^k \quad (3.18)$$

Dabei bezeichnet $\angle w$ das Argument bzw. den Winkel der komplexen Zahl w .

Bemerkung: Ein Filter mit der Impulsantwort \mathbf{H} implementiert eine Faltung mit \mathbf{H} , d.h. dem Eingangssignal $\mathbf{x} : \mathbb{Z} \rightarrow \mathbb{R}^M : k \mapsto \mathbf{x}(k)$ wird das Ausgangssignal $\mathbf{y} : \mathbb{Z} \rightarrow \mathbb{R}^N : k \mapsto \mathbf{y}(k) := \sum_l \mathbf{H}(k-l)\mathbf{x}(l)$ zugeordnet.

Aus der Systemtheorie ist bekannt, dass eine Sinusschwingung der Frequenz Ω am Eingang eines Filters immer eine Sinusschwingung am Ausgang eines Filters produziert. Die Amplitude und Phase des Ausgangssignals, sind dann gegeben durch den Betrag und Phase des Frequenzganges $\mathbf{H}(e^{i\Omega})$. Genauer folgt für

$$\mathbf{x}(k) = \mathbf{x}_0 e^{i\Omega k} \quad (3.19)$$

mit $\mathbf{y}(k) = \sum_l \mathbf{H}(k-l)\mathbf{x}(l)$:

$$\mathbf{y}(k) = \mathbf{A}(\Omega) e^{i(\Omega k + \mathbf{P}(\Omega))} \mathbf{x}_0. \quad (3.20)$$

Hierbei ist die Zerlegung des Frequenzganges $\mathbf{H}(e^{i\Omega})$ in Amplitude \mathbf{A} und Phase \mathbf{P}

$$\mathbf{A}(\Omega) e^{i\mathbf{P}(\Omega)} = \mathbf{H}(e^{i\Omega}). \quad (3.21)$$

bekannt. Falls nun $P(\Omega)$ linear bezüglich Ω ist, also $P(\Omega) = T_0\Omega$, so haben alle Schwingungen die durch das Filter laufen, eine *konstante* Laufzeit. Sie erfahren also nur eine Frequenzabhängige Dämpfung aber keine frequenzabhängige Verzögerung:

$$\mathbf{y}(k) = \mathbf{A}(\Omega)e^{i\Omega(k+T_0)} \mathbf{x}_0 \quad (3.22)$$

Daher werden linearphasige Filter auch *verzerrungsfrei* genannt.

Man kann zeigen, dass die Eigenschaft der Linearphasigkeit eine Symmetrie der Filterkoeffizienten $H(k)$ erfordert. Es gibt dabei 2 (bzw. 4) verschiedene Typen von Symmetrie, die dabei alle auf linearphasige Filter führen (siehe 2.4.2. in [Vai93]). Die einzelnen Komponenten $h_{nm}(k)$ von $H(k)$ sind linearphasig falls gilt:

- $h_{nm}(k) = 0$, $k < 0$ oder $k \geq N$,
- und $h_{nm}(k) = h_{nm}(N - k)$,
- oder $h_{nm}(k) = -h_{nm}(N - k)$,

wobei N gerade und ungerade sein kann. Wenn diese beiden Fälle noch mit unterschieden werden, dann erhält man 4 Symmetriebedingungen. Eine ganzzahlige Verschiebung von h_{nm} ändert die Eigenschaft der Linearphasigkeit nicht. Abhängig von der Parität der Länge des Trägers von h_{nm} kann T_0 ganz- oder halbzahlige Einträge haben.

3.3 Normalformen der Darstellung der Skalierungsfunktion

Wie vorher betont, ist eine zu einer Multiskalenanalyse gehörige Skalierungsfunktion keineswegs eindeutig festgelegt, denn orthogonale Transformationen und ganzzahlige Verschiebungen bilden wieder eine gültige Skalierungsfunktion. Daher ist es wünschenswert, Normalformen zu haben, die diese Mehrdeutigkeit ausschließen oder verringern.

3.3.1 Form mit minimalem Träger

Entsprechend [DGH96] kann eine Normalform mit minimalem Träger (*engl. minimally supported*) für Skalierungsfunktionen $\Phi \in L^2(\mathbb{R})^R$ mit $\text{supp}(\Phi) \subset [-1, 1]$ eingeführt werden.

Satz 3.3.1 (Normalform auf $[-1, 1]$). *Sei Φ eine orthogonale Skalierungsfunktion zu einer Multiskalenanalyse von Vielfachheit R , Φ bestehe aus linear unabhängigen Komponenten und $\text{supp}(\Phi) \subset [-1, 1]$. Dann existiert eine orthogonale Skalierungsfunktion $\tilde{\Phi}$ für die Multiskalenanalyse mit Träger $\subset [-1, 1]$*

und

$$\tilde{\Phi}_{\chi_{[0,1]}} \text{ hat } R \text{ l.u. Komponenten, sowie} \quad (3.23)$$

$$\tilde{\Phi}_{\chi_{[-1,0]}} \text{ hat } k \text{ l.u. Komponenten und } R - k \text{ Nullfunktionen.} \quad (3.24)$$

l.u. bedeutet hierbei "linear unabhängig".

Beweis Sei $W = \langle \Phi_{\chi_{[0,1]}}, \Phi_{\chi_{[0,1]}} \rangle$. Da W symmetrisch ist, existiert die Eigenwertzerlegung $W = Q^T D Q$ mit $Q^{-1} = Q^T$ und D die Diagonalmatrix der Eigenwerte. Betrachte $\Phi^{(1)} = Q\Phi$. Es ist $W^{(1)} = \langle \Phi^{(1)}_{\chi_{[0,1]}}, \Phi^{(1)}_{\chi_{[0,1]}} \rangle = D$. Klar, dass zu den verschwindenden Diagonalelementen von D Nullfunktionen in $\Phi^{(1)}_{\chi_{[0,1]}}$ gehören. Diese zugehörigen Funktionen $\phi_r^{(1)}$ auf $[-1, 1]$ werden nun nach rechts verschoben und bilden mit den übrigen unverschobenen Funktionen aus $\Phi^{(1)}$ die Funktion $\Phi^{(2)}$. Klar, dass $\phi_r^{(1)}_{\chi_{[-1,0]}} \neq 0$ ist, weil Φ als linear unabhängig auf $[-1, 1]$ vorausgesetzt ist. Zudem ist klar, dass $\Phi^{(1)}$ und $\Phi^{(2)}$ wie Φ auch, orthogonal sind und den gleichen Raum erzeugen, da orthogonale Transformationen und Verschiebungen den erzeugten Raum nicht ändern. Nun hat $\Phi^{(2)}_{\chi_{[0,1]}}$ linear unabhängige Komponenten, da die hinzukommenden verschobenen Funktionen auf $[0, 1]$ orthogonal, also auch linear unabhängig zu den vorherigen sind. Als letzten Schritt bringt man $\Phi^{(2)}$ wieder durch eine orthogonale Transformation Q auf die endgültige Form $\tilde{\Phi} = Q\Phi^{(2)}$, sodass sich entsprechend des Ranges von $W^{(2)} = \langle \Phi^{(2)}_{\chi_{[-1,0]}}, \Phi^{(2)}_{\chi_{[-1,0]}} \rangle$ $\tilde{\Phi}$ auf $[-1, 0]$ in linear unabhängige Komponenten und Nullfunktionen sortiert wird. \square

4 Allgemeiner Ansatz für Skalierungsfunktionen

Allgemeine (vektorielle) Multiskalenanalysen lassen die Konstruktion von stetigen Skalierungsfunktionen und Wavelets aus stückweisen Polynomen zu, die sowohl Approximationsordnung haben, als auch orthogonal mit kompaktem Träger sind [DGH99, DGH96, Roa97]. Zu deren Untersuchung wird ein allgemeiner Ansatz für die Skalierungsfunktion Φ gewählt.

Es werden zuerst Überlegungen allgemeiner Art aufgeführt, die nicht nur für stückweise Polynome gelten. Im nachfolgenden Kapitel werden diese dann für Skalierungsfunktionen aus stückweisen Polynomen spezialisiert.

4.1 Die Basis für die Darstellung von Skalierungsfunktionen

Um Skalierungsfunktion effizient darstellen zu können, sei hierzu eine M -dimensionale vektorielle Funktion $\mathbf{B} : \mathbb{R} \rightarrow \mathbb{R}^M : t \mapsto \mathbf{B}(t)$ mit Träger auf $[0, 1]$ gegeben, deren Komponenten aus $L^2(\mathbb{R})$ und orthonormal sind und mit geeigneten $\mathbf{p}_n(k) \in \mathbb{R}^{1 \times M}; k \in \mathbb{Z}$ die Polynome vom Grad $< N$ darstellen können:

$$\text{supp}(\mathbf{B}) = [0, 1], \quad (4.1a)$$

$$\mathbf{B}(1) = \mathbf{0}, \quad (4.1b)$$

$$\langle \mathbf{B}, \mathbf{B} \rangle = \mathbf{I} \quad (\in \mathbb{R}^{M \times M}), \quad (4.1c)$$

$$t^n = \sum_k \mathbf{p}_n(k) \mathbf{B}(t - k), \quad n = 0, \dots, N - 1. \quad (4.1d)$$

Die letzte Eigenschaft kann kompakt als $\mathbf{p} \in \mathbb{S}(\mathbf{B})$ mit $\mathbf{p}(t) \in \mathbb{R}^N$ geschrieben werden. Weiterhin soll das Dilat als Summe darstellbar sein:

$$\mathbf{B}(t/2) = \mathbf{D}_0 \mathbf{B}(t) + \mathbf{D}_1 \mathbf{B}(t - 1). \quad (4.1e)$$

Es ist klar, dass die einparametrische Familie $\mathcal{B}_0(\mathbf{B})$ der Funktion \mathbf{B} orthonormal ist, da die Länge des Trägers von \mathbf{B} 1 ist und zudem Gl. (4.1c) gilt. Nach Abschnitt 3.1 erzeugt eine Funktion \mathbf{B} , die sowohl die Verfeinerbarkeitseigenschaft Gl. (4.1e) besitzt, als auch orthonormal ist, eine orthogonale Multiskalenanalyse. Insbesondere gilt auch die Approximationseigenschaft Gl. (2.16b). Aus der Orthonormalität folgt auch die Eindeutigkeit der Darstellung einer Funktion aus $\text{cl}_{L^2(\mathbb{R})} \{\text{span} \{\mathcal{B}_0(\mathbf{B})\}\}$ bezüglich $\mathcal{B}_0(\mathbf{B})$, da diese einparametrische Familie eine Riesz-Basis ist.

4.1.1 Herstellung von Normiertheit aus linearer Unabhängigkeit

Die Eigenschaft der Normiertheit Gl. (4.1c) ist keine wesentliche Einschränkung: Ist B' mit Träger auf $[0, 1]$, $B'(1) = 0$, linear unabhängigen Komponenten gegeben und erfüllt Gl. (4.1e) mit D'_0 und D'_1 , so gelten für $B = MB'$ alle Gleichungen inklusive der "Normiertheit", wobei $M = (\langle B', B' \rangle^{-1})^{1/2}$ und $D_i = M^{-1}D'_iM$, $i = 0, 1$. Für die Definition der Matrixwurzel siehe Abschnitt 4.3.

4.1.2 Definition der Skalierungsfunktion

Um Multiskalenanalysen zu konstruieren, ist es nun die Aufgabe, eine geeignete zwischenliegende Multiskalenanalyse mit gewünschten Eigenschaften zu finden.

Dieses führt auf die folgende Definition:

Definition 4.1.1 (Allgemeiner Ansatz für die Skalierungsfunktion) *Ist die eine Funktion B mit den Eigenschaften Gl. (4.1a) bis Gl. (4.1d) festgelegt, so wird mit einer beliebige Matrixfolge $C : \mathbb{Z} \rightarrow \mathbb{R}^{R \times M} : p \mapsto C(p)$ eine Funktion Φ_C mit Komponenten aus $\mathbb{S}(B)$ definiert und mit Φ_C bezeichnet:*

$$\Phi_C : \mathbb{R} \rightarrow \mathbb{R}^R : t \mapsto \Phi_C(t) := \sum_{p \in \mathbb{Z}} C(p) B(t - p). \quad (4.2)$$

Bemerkung Man hat für den wichtigen Spezialfall einer Skalierungsfunktion mit Träger auf $[-1, 1]$: $\Phi_C(t) = C(-1)B(t) + C(0)B(t - 1)$.

Bemerkung Falls die Koeffizientenfolge C unendlich lang ist, ist im Allgemeinen nicht gesichert, dass die Komponenten von Φ_C in $L^2(\mathbb{R})$ liegen.

4.1.3 Approximationsordnung

Damit eine Approximationsordnung von Φ_C realisiert werden kann, ist es nötig, dass Polynome von hinreichend großem Grad bezüglich $\mathcal{B}_0(B)$ dargestellt werden können. Daher die Forderung Gl. (4.1d). Es gibt daher einige Implikationen an die Eigenwerte von D_0 und D_1 die nun hergeleitet werden. Sei $P : \mathbb{Z} \rightarrow \mathbb{R}^{N \times M} : k \mapsto P(k)$ die Koeffizientenfolge für den Polynomvektor $p(k)$ vom Grad $K - 1$ bezüglich B . Die Zeilen von $P(k)$ sind die $p_n(k) \in \mathbb{R}^{1 \times M}$ aus Gl. (4.1d):

$$P(k) = \begin{pmatrix} p_0(k) \\ p_1(k) \\ \vdots \\ p_{N-1}(k) \end{pmatrix} \quad (4.3)$$

Klar, dass $\Phi_{\mathbf{P}}(t) = \mathbf{p}(t) = (1 \ t \ \dots \ t^{N-1})^T$ ist.

Man findet nun allgemein, dass $\mathbf{p}_n(0)$ und $\mathbf{p}_n(-1)$ Eigenvektoren von \mathbf{D}_0 und \mathbf{D}_1 zum Eigenwert 2^{-n} sind. Zudem gelten noch verschiedene Rekursionsbeziehungen:

Satz 4.1.1 *Ist \mathbf{B} wie in Abschnitt 4.1 mit \mathbf{D}_0 und \mathbf{D}_1 aus der Verfeinerungsgleichung definiert und sind die Polynome vom Grad $< N$ bezüglich $\mathcal{B}_0(\mathbf{B})$ darstellbar, dann gelten für die Koeffizienten $\mathbf{p}_n(k)$ der Darstellung $t^n = \sum_k \mathbf{p}_n(k)\mathbf{B}(t-k)$ mit $n = 0, 1, \dots, N-1$ die folgenden Eigenschaften:*

$$\mathbf{p}_n(k)\mathbf{D}_0 = 2^{-n}\mathbf{p}_n(2k), \quad (4.4a)$$

$$\mathbf{p}_n(k)\mathbf{D}_1 = 2^{-n}\mathbf{p}_n(2k+1), \quad (4.4b)$$

bzw. in Matrixschreibweise:

$$\mathbf{P}(k)\mathbf{D}_0 = \mathbf{S}(1/2)\mathbf{P}(2k), \quad (4.4c)$$

$$\mathbf{P}(k)\mathbf{D}_1 = \mathbf{S}(1/2)\mathbf{P}(2k+1). \quad (4.4d)$$

Insbesondere gelten die Eigenvektorbeziehungen:

$$\mathbf{p}_n(0)\mathbf{D}_0 = 2^{-n}\mathbf{p}_n(0), \quad (4.4e)$$

$$\mathbf{p}_n(-1)\mathbf{D}_1 = 2^{-n}\mathbf{p}_n(-1), \quad (4.4f)$$

sowie

$$\mathbf{P}(k) = \mathbf{T}(-k)\mathbf{P}(0). \quad (4.4g)$$

Die Matrizen $\mathbf{T}(k)$ und $\mathbf{S}(a)$ sind in Abschnitt 2.5 definiert und explizit angegeben.

Beweis Ist $\mathbf{p}(t) = \sum_k \mathbf{P}(k)\mathbf{B}(t-k)$ dann folgt $\mathbf{p}(t/2) = \sum_k \mathbf{P}(k)\mathbf{B}((t-2k)/2) = \sum_k \mathbf{P}(k)\mathbf{D}_0\mathbf{B}(t-2k) + \sum_k \mathbf{P}(k)\mathbf{D}_1\mathbf{B}(t-2k-1)$. Andererseits gilt $\mathbf{p}(t/2) = \mathbf{S}(1/2)\mathbf{p}(t) = \sum_k \mathbf{S}(1/2)\mathbf{P}(k)\mathbf{B}(t-k)$. Aufgrund der Eindeutigkeit der Darstellung führt man einen Koeffizientenvergleich durch und erhält so Gl. (4.4c) und Gl. (4.4d). Die Gleichungen Gl. (4.4a) und Gl. (4.4b) folgen mit der Definition von $\mathbf{P}(k)$ direkt daraus. Um Gl. (4.4g) zu zeigen, betrachtet man $\mathbf{p}(t-l)$, was einerseits gegeben ist als $\mathbf{T}(l)\mathbf{p}(t) = \sum_k \mathbf{T}(l)\mathbf{P}(k)\mathbf{B}(t-k)$, andererseits gleich $\sum_k \mathbf{P}(k)\mathbf{B}(t-l-k) = \sum_k \mathbf{P}(k-l)\mathbf{B}(t-k)$ ist. Der Koeffizientenvergleich für den Summanden $k=0$ führt dann auf Gl. (4.4g). \square

Nun werden nacheinander die nötigen Eigenschaften einer Skalierungsfunktion $\Phi_{\mathbf{C}}$ anhand von äquivalenten Koeffizientenbedingungen, d.h. Forderungen an die Folge \mathbf{C} formuliert.

4.2 Generelle Eigenschaften

Satz 4.2.1 (Skalarprodukt) Falls das Skalarprodukt zweier beliebiger Funktion Φ_C und Φ_F entsprechend Def. 4.1.1 mit gleicher Basisfunktion existiert, ist es gegeben durch das Skalarprodukt der Koeffizientenfolgen C und D :

$$\langle \Phi_C, \Phi_F \rangle = \sum_k C(k)F(k)^T. \quad (4.5)$$

Beweis $\langle \Phi_C, \Phi_F \rangle = \left\langle \sum_p C(p) \mathbf{B}(t-p), \sum_q F(q) \mathbf{B}(t-q) \right\rangle$
 $= \sum_{p,q} C(p) \langle \mathbf{B}(t-p), \mathbf{B}(t-q) \rangle F(q)^T = \sum_p C(p)F(p)^T \quad \square$

Bemerkung Das Skalarprodukt existiert stets, wenn eine der beiden Funktionen kompakten Träger besitzt, d.h. eine der beiden Koeffizientenfolgen endlich lang ist oder beide Funktionen Φ_C und Φ_F Komponenten aus $L^2(\mathbb{R})$ haben.

Satz 4.2.2 (Ganzzahlige Verschiebung) Die ganzzahlige Verschiebung einer beliebigen Funktion Φ_C entsprechend Def. 4.1.1 ist gegeben durch die Verschiebung der Koeffizientenfolge:

$$\Phi_C(\cdot - k) = \Phi_{C(\cdot - k)}, \quad k \in \mathbb{Z} \quad (4.6)$$

Beweis $\Phi_C(t - k) = \sum_p C(p) \mathbf{B}(t - k - p) = \sum_p C(p - k) \mathbf{B}(t - p) \quad \square$

Satz 4.2.3 (Lineare Transformation) Die lineare Transformation einer beliebigen Funktion Φ_C entsprechend Def. 4.1.1 ist gegeben durch die entsprechende Transformation der Koeffizientenfolge und die Addition von zwei entsprechenden Funktionen Φ_C und Φ_F mit gleicher Basis durch Addition der Koeffizienten:

$$M\Phi_C = \Phi_{MC} \quad (4.7)$$

$$\Phi_C + \Phi_F = \Phi_{C+F} \quad (4.8)$$

Beweis $M\Phi_C(t) = \sum_p MC(p) \mathbf{B}(t-p)$ und $\Phi_C(t) + \Phi_F(t) = \sum_p (C(p) + F(p)) \mathbf{B}(t-p)$
 \square

4.3 Orthonormalitätsbedingung

Satz 4.3.1 (Kriterium für Orthonormalität) Die Funktion Φ_C aus Def. 4.1.1 mit der Matrixfolge C ist eine orthonormale Skalierungsfunktion genau dann, wenn

$$\sum_l C(l)C(l-k)^T = \delta_k \mathbf{I}, \quad k \in \mathbb{Z}. \quad (4.9)$$

Beweis Wendet man die Eigenschaft für das Skalarprodukt Gl. (4.5) und die Verschiebungseigenschaft Gl. (4.6) an, so sieht man dass Gl. (2.31) und Gl. (4.9) äquivalent sind. \square

Bemerkung Es ist klar, dass für die Erfüllung von Gl. (4.9) es aufgrund der Symmetrie ausreicht $k \geq 0$ zu betrachten. Zudem kann es sein, dass für unendliche lange C das Skalarprodukt nicht existiert, was aber dann bedeutet, dass Φ_C nicht orthonormal ist.

Zudem gilt: Falls die "Orthogonalität" (Gl. (4.9) für $k \neq 0$) erfüllt ist, kann "Normiertheit" im Allgemeinen (Gl. (4.9) für $k = 0$) hergestellt werden:

Satz 4.3.2 (Herstellung der Normierung) *Falls mit einer Matrixfolge $C : \mathbb{Z} \rightarrow \mathbb{R}^{R \times M} : k \mapsto C(k)$ die Beziehung $\sum_l C(l)C(l-k)^T = 0$ für $k > 0$ gilt, und existiert $\sum_l C(l)C(l)^T$ und besitzt maximalen Rang R (\Leftrightarrow die Komponenten der Funktion Φ_C sind linear unabhängig), so existiert eine invertierbare Matrix M , sodass für $\tilde{C} = MC$ die Gleichung $\sum_l \tilde{C}(l)\tilde{C}(l-k)^T = \delta_k \mathbf{I}$ für alle $k \in \mathbb{Z}$ gilt. Die Matrix ist gegeben durch*

$$M = \left(\left(\sum_l C_l C_l^T \right)^{-1} \right)^{1/2}. \quad (4.10)$$

Bemerkung zur Existenz: Da die Matrix $\sum_l C(l)C(l)^T$ maximalen Rang hat, ist sie invertierbar. Zudem ist sie als Summe von symmetrisch positiv semidefiniten Matrizen wieder symmetrisch positiv semidefinit. (Definition von A positiv semidefinit: $\forall x : x^T A x \geq 0$). Die Invertierbarkeit sichert positive Definitheit und daher auch Existenz der Matrixwurzel $X = A^{1/2}$ aus symmetrisch positiv definitem A mit $X^T X = A$. Diese kann beispielsweise mit der Cholesky-Zerlegung berechnet werden [GvL96].

Beweis Es ist nach Definition $\sum_l (MC(l))(MC(l))^T = M(\sum_l C(l)C(l))M^T$. Da M invertierbar ist gilt: $\sum_l (MC(l))(MC(l+k))^T = 0 \Leftrightarrow \sum_l C(l)C(l+k)^T = 0$. Bleibt zu zeigen: $M \sum_l C(l)C(l)^T M^T = M(M^T M)^{-1} M^T = (M^{-1})^{-1} (M^T M)^{-1} (M^{T^{-1}})^{-1} = ((M^T)^{-1} M^T M M^{-1})^{-1} = \mathbf{I}$. \square

4.4 Dilatation

Satz 4.4.1 (Dilatation) *Die Dilatation $\Phi_C(\cdot/2)$ einer beliebigen Funktion entsprechend Def. 4.1.1 ist gegeben durch die folgende Transformation der Koeffizientenfolge, wobei die Matrizen D_0 und D_1 aus der Verfeinerungsgleichung Gl. (4.1e) von B stammen:*

$$C'(p) = C(p \operatorname{div} 2) D_{p \bmod 2}, \quad (4.11)$$

$$= \begin{cases} C(p/2) D_0 & p \text{ gerade} \\ C((p-1)/2) D_1 & p \text{ ungerade} \end{cases}, \quad (4.12)$$

mit $p \in \mathbb{Z}$ und

$$\Phi_{C'} = \Phi_C(\cdot/2), \quad (4.13)$$

$$p \operatorname{div} 2 = \lfloor p/2 \rfloor, \quad (4.14)$$

$$p \operatorname{mod} 2 = p - 2\lfloor p/2 \rfloor. \quad (4.15)$$

Beweis Es ist ja $\mathbf{B}(t/2) = \mathbf{D}_0\mathbf{B}(t) + \mathbf{D}_1\mathbf{B}(t-1)$. Also folgt für $\Phi_C(t/2) = \sum_p \mathbf{C}(p)\mathbf{B}(t/2-p) = \sum_p \mathbf{C}(p)(\mathbf{D}_0\mathbf{B}(t-2p) + \mathbf{D}_1\mathbf{B}(t-2p-1)) = \sum_p \mathbf{C}(p \operatorname{div} 2)\mathbf{D}_{p \operatorname{mod} 2}\mathbf{B}(t-p)$ \square

Beispiel Ist die Funktion Φ_C maximal von Länge 2, d.h. höchstens $\mathbf{C}(-1)$ und $\mathbf{C}(0)$ sind ungleich Null, so ist das Dilat höchstens von Länge 4 und entsprechenden Koeffizienten von $\Phi'_C(t) := \Phi_C(t/2)$ sind gegeben durch

$$\mathbf{C}'(-2) = \mathbf{C}(-1)\mathbf{D}_0, \quad \mathbf{C}'(-1) = \mathbf{C}(-1)\mathbf{D}_1, \quad \mathbf{C}'(0) = \mathbf{C}(0)\mathbf{D}_0, \quad \mathbf{C}'(1) = \mathbf{C}(0)\mathbf{D}_1. \quad (4.16)$$

4.5 Stetigkeitsbedingung

Damit Φ_C stetig ist, muss es an den ganzzahligen Stellen stetig sein, d.h. die Koeffizienten $\mathbf{C}(p)$ und $\mathbf{C}(p+1)$ müssen eine bestimmte Beziehung erfüllen. Falls zudem die Basis \mathbf{B} selber nicht stetig ist, so muss die Wahl von $\mathbf{C}(p)$ auch sicherstellen, dass $\mathbf{C}(p)\mathbf{B}$ eine auf $(0, 1)$ stetige Funktion ist.

Definition 4.5.1 (Links- und Rechtsseitiger Grenzwert) *Ist \mathbf{B} eine reelle Funktion, so bezeichnet die Schreibweise $\mathbf{B}(t_0+)$ den rechtsseitigen und $\mathbf{B}(t_0-)$ den linksseitigen Grenzwert an der Stelle t :*

$$\mathbf{B}(t_0-) = \lim_{t \rightarrow t_0, t < t_0} \mathbf{B}(t), \quad (4.17)$$

$$\mathbf{B}(t_0+) = \lim_{t \rightarrow t_0, t > t_0} \mathbf{B}(t). \quad (4.18)$$

Satz 4.5.1 (Stetigkeitsbedingung für die Basis \mathbf{B}) *Besitzt $\mathbf{B} : \mathbb{R} \rightarrow \mathbb{R}^M$ mit Träger $[0, 1]$ genau $I \in \mathbb{N}$ Unstetigkeitsstellen auf $(0, 1)$, dann ist $\mathbf{B}' : \mathbb{R} \rightarrow \mathbb{R}^{M'}$: $t \mapsto \mathbf{B}'(t) = \mathbf{K}\mathbf{B}(t)$ stetig auf $(0, 1)$ genau dann, wenn für $\mathbf{K} \in \mathbb{R}^{M' \times M}$ gilt:*

$$\mathbf{K}\Delta = \mathbf{0}. \quad (4.19)$$

Dabei bezeichnet Δ die Matrix mit den Differenzvektoren Δ_i der Unstetigkeitsstellen t_i , mit $i = 1, \dots, I$.

$$\Delta = (\Delta_1 \mid \cdots \mid \Delta_I) \quad (\in \mathbb{R}^{M \times I}) \quad (4.20)$$

$$\Delta_i = \mathbf{B}(t_i-) - \mathbf{B}(t_i+). \quad (\in \mathbb{R}^M) \quad (4.21)$$

Beweis Es ist notwendig, dass $B'(t) = KB(t)$ stetig bei t_i ist, also $B'(t_i+) = B'(t_i-)$ für $i = 1, 2, \dots, I$ gilt. Dies ist gleichbedeutend mit $K\Delta_i = 0$ für $i = 1, 2, \dots, I$. K muss also orthogonal zu jedem Δ_i sein. \square

Bemerkung Durch Beseitigen von I Unstetigkeitsstellen geht also höchstens die Dimension I verloren. Zudem ist es auch nicht entscheidend, wo genau die Sprungstellen sind, lediglich die Sprunghöhe ist von Belang.

Satz 4.5.2 Sei C und Φ_C entsprechend Def. 4.1.1 gegeben, und $C : \mathbb{Z} \rightarrow \mathbb{R}^{R \times M}$ mit $\text{supp}(\Phi_C) \subseteq [-P, P]$ und habe die zugehörige Basis B genau I Unstetigkeitsstellen auf $(0, 1)$. Φ_C ist genau dann stetig, wenn gilt:

$$\sum_{p=-P}^{P-1} C(p)S(p)^T = 0. \quad (4.22)$$

Die konstante Matrixfolge $S : \mathbb{Z} \rightarrow \mathbb{R}^{2P(I+1)+1 \times M}$ ist definiert als

$$S(-P)^T = (\Delta \ 0 \ \cdots \ 0 \ -b_0 \ b_1 \ 0 \ \cdots \ 0 \ 0) \quad (4.23)$$

$$S(-P+1)^T = (0 \ \Delta \ \cdots \ 0 \ 0 \ -b_0 \ b_1 \ \cdots \ 0 \ 0) \quad (4.24)$$

$$\vdots$$

$$S(P-1)^T = (0 \ 0 \ \cdots \ \Delta \ 0 \ 0 \ 0 \ \cdots \ -b_0 \ b_1), \quad (4.25)$$

wobei die Sprünge an den Unstetigkeitsstellen in B in $\Delta \in \mathbb{R}^{M \times I}$ entsprechend Gl. (4.20) zusammengefasst sind, und $b_0 = B(0+)$ bzw. $b_0 = B(1-)$ als Kurzschreibweise verwendet ist.

Beweis Außerhalb von $[-P, P]$ ist $\Phi_C \equiv 0$, also stetig. Auf den offenen Intervallen $(p, p+1)$, $p \in \mathbb{Z}$ ist Φ_C nach Satz 4.5.1 stetig genau dann, wenn $C_p \Delta = 0$ gilt. Dies bildet die ersten $2P$ Blockeinträge in $S(n)$. Damit Φ_C stetig bei $p = -P, -P+1, \dots, P-1$ muss $\Phi_C(p+) = \Phi_C(p-)$ sein. Dafür ist nötig: $C_{-P}B(0+) = 0$, $C_p B(1-) = C_{p+1}B(0+)$, $p = -P, \dots, P-2$ und $C_{P-1}B(1-) = 0$. Dies bildet die letzten $2P+1$ Einträge in $S(n)$. Man erhält so zudem, dass $S(n)^T$ aus $2P \cdot I + 2P + 1$ Spalten besteht. \square

Bemerkung Anschaulich interpretiert definiert die Folge S eine Funktion Φ_S deren beliebige Linearkombinationen immer unstetig an den kritischen Stellen sind. Da der Raum der Funktionen Φ_C endlichdimensional ist, und der Raum der stetigen Funktionen ein ebenfalls (endlichdimensionaler) Unterraum ist, existiert auch dessen orthogonales Komplement in. Dieses enthält gerade die stets unstetigen Funktionen und wird von Φ_S erzeugt.

4.6 Bedingung für verschwindende Momente

Verschwindende Momente bedeutet, dass die Funktion orthogonal zu Polynomen gewisser Ordnung ist. Dies ist mit der Theorie problemlos darstellbar.

Satz 4.6.1 Sei C und Φ_C entsprechend Def. 4.1.1 gegeben, also $C : \mathbb{Z} \rightarrow \mathbb{R}^{R \times M}$, und habe Φ kompakten Träger. Weiterhin sei in der zugehörigen Basis $t^l = \sum \mathbf{p}_l(k) \mathbf{B}(t - k)$ gegeben. Dann hat Φ_C genau dann ein verschwindendes l -tes Moment, wenn

$$\sum_p C(p) \mathbf{p}_l(p)^T = \mathbf{0}. \quad (4.26)$$

Beweis Die Eigenschaft folgt direkt aus der Definition des verschwindenden Moments und der Eigenschaft des Skalarprodukts für die Koeffizientenfolge.

□

4.7 Kriterium, dass eine spezielle Funktion in $\mathbb{S}(\Phi_C)$ liegt

Satz 4.7.1 Sei Φ_C eine Funktion nach Def. 4.1.1 mit orthonormaler einparametrischer Familie $\mathcal{B}_0(\Phi_C)$ und kompaktem Träger. Eine zweite Funktion Φ_F bezüglich der gleichen Basis \mathbf{B} liegt genau dann in dem aufgespannten Raum $\mathbb{S}(\Phi_C)$, wenn für die Koeffizientenfolgen gilt:

$$\mathbf{F}(p) = \sum_k \sum_l \mathbf{F}(l) C(l - k)^T C(p - k), \quad \forall p \in \mathbb{Z}, \quad (4.27)$$

oder äquivalent

$$\mathbf{0} = \sum_l \mathbf{F}(p + l) \left(\sum_k C(k + l)^T C(k) - \delta_l \mathbf{I} \right), \quad \forall p \in \mathbb{Z}. \quad (4.28)$$

Beweis Es ist $\Phi_F \in \mathbb{S}(\Phi_C) \Leftrightarrow \Phi_F = \sum_k \langle \Phi_F, \Phi_C(\cdot - k) \rangle \Phi_C(\cdot - k) = \sum_k \sum_l \mathbf{F}_l C(l - k)^T \Phi_C(\cdot - k)$. Mit der Eindeutigkeit der Darstellung bzgl. \mathbf{B} , erhält man die äquivalente Formulierung $\mathbf{F}_p = \sum_k \sum_l \mathbf{F}_l C(l - k)^T C(p - k)$, also Gl. (4.27). Bringt man $\mathbf{F}(p)$ auf die andere Seite und ersetzt $p - k \rightarrow k$ und $l \rightarrow p + l$ so ergibt sich Gl. (4.28). □

Bemerkung Soll geprüft werden, ob die Polynome im Raum liegen, so ist es ausreichend, dass Gl. (4.28) für $p = 0$ erfüllt ist. Grund: Für die Polynome ist nach Gl. (4.4g) $\mathbf{F}(p + l) = \mathbf{P}(p + l) = \mathbf{T}(p)\mathbf{P}(l)$. Nach Multiplikation mit der invertierbaren Matrix $\mathbf{T}(-p)$ ist die Gleichung dann von der Form Gl. (4.28) mit $p = 0$.

Bemerkung Falls Φ_F kompakten Träger hat, so liefert das Kriterium sogar eine Aussage ob $\Phi_F \in \text{cl}_{L^2(\mathbb{R})} \{\text{span} \{\mathcal{B}_0(\Phi_C)\}\}$ ist, oder nicht. Dies liegt an Satz 2.2.1 mit $\Phi_F \in L^2(\mathbb{R})$.

4.8 Zusammenfassung der Koeffizientengleichungen für Orthonormalität und Verfeinerbarkeit für Träger $[-1, 1]$

Wenn die Skalierungsfunktion Φ_C maximal den Träger $[-1, 1]$ besitzt ist die Länge der Koeffizientenfolge höchstens 2. Es gibt nur die nichttrivialen Koeffizienten $C(-1)$ und $C(0)$. Φ_C hat dann die Form

$$\Phi_C(t) = C(-1)B(t+1) + C(0)B(t). \quad (4.29)$$

Um sicherzustellen, dass mit Φ_C eine orthonormale Multiskalenanalyse erzeugt wird, ist es notwendig die Verfeinerbarkeits- und Orthonormalitätsbedingung zu erfüllen. Verfeinerbarkeit bedeutet dass die Dilatation $\Phi_C(\cdot/2)$ in $cl_{L^2(\mathbb{R})} \{ \text{span} \{ \mathcal{B}_0(\Phi_C) \} \}$ liegt.

4.8.1 Ohne Bedingungen an $C(0)$ und $C(-1)$

Die Orthogonalität ist mit Gl. (4.9) äquivalent zu

$$\mathbf{I} = C(-1)C(-1)^T + C(0)C(0)^T \quad (4.30)$$

$$\mathbf{0} = C(-1)C(0)^T \quad (4.31)$$

Verfeinerbarkeit ist mit Gl. (4.28) und Gl. (4.11) äquivalent zu

$$\mathbf{0} = C(0)D_1C(-1)^TC(0) \quad (4.32)$$

$$\mathbf{0} = C(0)D_0C(-1)^TC(0) + C(0)D_1(C(-1)^TC(-1) + C(0)^TC(0) - \mathbf{I}) \quad (4.33)$$

$$\begin{aligned} \mathbf{0} = & C(-1)D_1C(-1)^TC(0) + C(0)D_0(C(-1)^TC(-1) + C(0)^TC(0) - \mathbf{I}) \\ & + C(0)D(1)C(0)^TC(-1) \end{aligned} \quad (4.34)$$

$$\begin{aligned} \mathbf{0} = & C(-1)D_0C(-1)^TC(0) + C(-1)D_1(C(-1)^TC(-1) + C(0)^TC(0) - \mathbf{I}) \\ & + C(0)D_0C(0)^TC(-1) \end{aligned} \quad (4.35)$$

$$\mathbf{0} = C(-1)D_0(C(-1)^TC(-1) + C(0)^TC_0 - \mathbf{I}) + C(-1)D_1C(0)^TC(-1) \quad (4.36)$$

$$\mathbf{0} = C(-1)D_0C(0)^TC(-1) \quad (4.37)$$

Bei der Herleitung ist es Nützlich sich zu überlegen, dass die Koeffizientenfolge F für das Dilat aus 4 statt 2 (wie C) nichttrivialen Elementen besteht und nach Gl. (4.11) gegeben ist durch

$$F(-2) = C(-1)D_0 \quad F(-1) = C(-1)D_1 \quad F(0) = C(0)D_0 \quad F(1) = C(-1)D_0 \quad (4.38)$$

mit

$$\Phi_F(t) = \Phi_C(t/2) \quad \text{bzw.} \quad \sum_l F(l)B(t-l) = \sum_l C(l)B(t/2-l). \quad (4.39)$$

Die Summe in Gl. (4.28) hat für $k = -1, 0, 1$ die folgenden Werte:

$$\sum_k \mathbf{C}(k+l)^T \mathbf{C}(k) - \delta_l \mathbf{I} = \begin{cases} \mathbf{C}(-1)^T \mathbf{C}(0) & k = -1 \\ \mathbf{C}(-1)^T \mathbf{C}(-1) + \mathbf{C}(0)^T \mathbf{C}(0) - \mathbf{I} & k = 0 \\ \mathbf{C}(0)^T \mathbf{C}(-1) & k = 1 \\ \mathbf{0} & \text{sonst} \end{cases} \quad (4.40)$$

Eine andere gleichwertige Formulierung der Verfeinerbarkeit ist: Mit

$$\mathbf{G}(-2) = \mathbf{C}(-1) \mathbf{D}_0 \mathbf{C}(0)^T, \quad (4.41)$$

$$\mathbf{G}(-1) = \mathbf{C}(-1) \mathbf{D}_0 \mathbf{C}(-1)^T + \mathbf{C}(-1) \mathbf{D}_1 \mathbf{C}(0)^T, \quad (4.42)$$

$$\mathbf{G}(0) = \mathbf{C}(-1) \mathbf{D}_1 \mathbf{C}(-1)^T + \mathbf{C}(0) \mathbf{D}_0 \mathbf{C}(0)^T, \quad (4.43)$$

$$\mathbf{G}(1) = \mathbf{C}(0) \mathbf{D}_0 \mathbf{C}(-1)^T + \mathbf{C}(0) \mathbf{D}_1 \mathbf{C}(0)^T, \quad (4.44)$$

$$\mathbf{G}(2) = \mathbf{C}(0) \mathbf{D}_1 \mathbf{C}(-1)^T, \quad (4.45)$$

muss gelten:

$$\mathbf{0} = \mathbf{G}(-2) \mathbf{C}(-1), \quad (4.46)$$

$$\mathbf{C}(-1) \mathbf{D}_0 = \mathbf{G}(-1) \mathbf{C}(-1) + \mathbf{G}(-2) \mathbf{C}(0), \quad (4.47)$$

$$\mathbf{C}(-1) \mathbf{D}_1 = \mathbf{G}(0) \mathbf{C}(-1) + \mathbf{G}(-1) \mathbf{C}(0), \quad (4.48)$$

$$\mathbf{C}(0) \mathbf{D}_0 = \mathbf{G}(1) \mathbf{C}(-1) + \mathbf{G}(0) \mathbf{C}(0), \quad (4.49)$$

$$\mathbf{C}(0) \mathbf{D}_1 = \mathbf{G}(2) \mathbf{C}(-1) + \mathbf{G}(1) \mathbf{C}(0), \quad (4.50)$$

$$\mathbf{0} = \mathbf{G}(2) \mathbf{C}(0). \quad (4.51)$$

Bemerkung Falls Stetigkeit und verschwindende Momente noch zusätzlich gefordert werden, ist der folgende Ansatz zweckmäßig:

$$\mathbf{C}(0) = \mathbf{K} \mathbf{R}(0) \quad \mathbf{C}(-1) = \mathbf{K} \mathbf{R}(-1) \quad (4.52)$$

wobei die Matrix \mathbf{K} nun die unbekanntenen Koeffizienten enthält. Dabei ist die Folge \mathbf{R} so, dass $\Phi_{\mathbf{R}}$ gerade das orthogonale Komplement der Vereinigung von unstetigen Funktionen (siehe Abschnitt 4.5) und Polynomen von geeignetem Grad ist. Praktisch heißt dies, dass die Folge \mathbf{B} orthogonal zur Folge \mathbf{S} und \mathbf{P} sein muss. Da ein kompakter Träger vorliegt ist \mathbf{R} mit linearer Algebra leicht zu berechnen, z.B. mittels der Singulärwertzerlegung.

Bemerkung Falls in Gl. (4.46) bis Gl. (4.51) die $\mathbf{G}(k)$ bekannt sind, so kann für die Matrizen $\mathbf{C}(0)$ und $\mathbf{C}(-1)$ aufgelöst werden, da ein lineares Gleichungssystem vorliegt. Dazu werden je zwei Gleichungen zusammengefasst:

$$\begin{pmatrix} \mathbf{C}(-1) \\ \mathbf{C}(0) \end{pmatrix} \mathbf{D}_0 = \begin{pmatrix} \mathbf{G}(-1) & \mathbf{G}(-2) \\ \mathbf{G}(1) & \mathbf{G}(0) \end{pmatrix} \begin{pmatrix} \mathbf{C}(-1) \\ \mathbf{C}(0) \end{pmatrix} \quad (4.53)$$

$$\begin{pmatrix} \mathbf{C}(-1) \\ \mathbf{C}(0) \end{pmatrix} \mathbf{D}_1 = \begin{pmatrix} \mathbf{G}(0) & \mathbf{G}(-1) \\ \mathbf{G}(2) & \mathbf{G}(1) \end{pmatrix} \begin{pmatrix} \mathbf{C}(-1) \\ \mathbf{C}(0) \end{pmatrix} \quad (4.54)$$

$$\mathbf{0} = \begin{pmatrix} \mathbf{G}(-2) & \mathbf{0} \\ \mathbf{0} & \mathbf{G}(2) \end{pmatrix} \begin{pmatrix} \mathbf{C}(-1) \\ \mathbf{C}(0) \end{pmatrix} \quad (4.55)$$

Es handelt sich bei den ersten beiden Gleichungen um Sylvester Gleichungen, deren zu deren Lösung das Problem am besten mit dem Kronecker-Matrixprodukt \otimes und dem Vektorisierungsoperator vec in eine Vektor-Matrix Schreibweise umgeformt wird [Bre78]:

$$\text{Gl. (4.53)} \Rightarrow \left((\mathbf{D}_0^T \otimes \mathbf{I}) - \left(\mathbf{I} \otimes \begin{pmatrix} \mathbf{G}(-1) & \mathbf{G}(-2) \\ \mathbf{G}(1) & \mathbf{G}(0) \end{pmatrix} \right) \right) \text{vec} \begin{pmatrix} \mathbf{C}(-1) \\ \mathbf{C}(0) \end{pmatrix} = \mathbf{0} \quad (4.56)$$

$$\text{Gl. (4.54)} \Rightarrow \left((\mathbf{D}_1^T \otimes \mathbf{I}) - \left(\mathbf{I} \otimes \begin{pmatrix} \mathbf{G}(0) & \mathbf{G}(-1) \\ \mathbf{G}(2) & \mathbf{G}(1) \end{pmatrix} \right) \right) \text{vec} \begin{pmatrix} \mathbf{C}(-1) \\ \mathbf{C}(0) \end{pmatrix} = \mathbf{0} \quad (4.57)$$

Dabei ist das Kronecker-Produkt \otimes für beliebige Matrizen definiert als

$$\otimes : (\mathbb{R}^{M \times N}, \mathbb{R}^{P \times Q}) \rightarrow \mathbb{R}^{MP \times NQ} : \mathbf{A} \otimes \mathbf{B} := \begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1N}\mathbf{B} \\ \vdots & & \vdots \\ a_{M1}\mathbf{B} & \cdots & a_{MN}\mathbf{B} \end{pmatrix} \quad (4.58)$$

und der Vektorisierungsoperator gegeben durch das Untereinanderfügen der Spalten einer Matrix

$$\text{vec} : \mathbb{R}^{N \times M} \rightarrow \mathbb{R}^{NM} : \mathbf{A} \mapsto \text{vec}(\mathbf{A}) = \begin{pmatrix} \mathbf{A}_{*1} \\ \vdots \\ \mathbf{A}_{*M} \end{pmatrix}. \quad (4.59)$$

Falls das Matrixprodukt \mathbf{ABC} wohlgeformt ist, gilt die wichtige Eigenschaft:

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}). \quad (4.60)$$

4.8.2 Falls $\mathbf{C}(0)$ vollen Rang besitzt

Aufgrund der Überlegungen in Abschnitt 3.3.1 ist es ohne Einschränkung möglich \mathbf{C}_0 mit vollem Rang anzunehmen. In diesem Fall hat die Pseudoinverse \mathbf{C}_0^+ von \mathbf{C}_0 die folgende Gestalt:

$$\mathbf{C}(0)^+ = \mathbf{C}(0)(\mathbf{C}(0)\mathbf{C}(0)^T)^{-1} \quad (4.61)$$

und ist eine Rechtsinverse der Matrix $\mathbf{C}(0)$: $\mathbf{C}(0)\mathbf{C}(0)^+ = \mathbf{I}$.

Dann sind die Verfeinerungskoeffizienten gegeben durch (dies ergibt sich durch Multiplikation von Gl. (4.46-4.51) mit \mathbf{C}_0^+ von rechts):

$$\mathbf{G}(-2) = \mathbf{C}(-1)\mathbf{D}_0\mathbf{C}(0)^+ \quad (4.62)$$

$$\mathbf{G}(-1) = \mathbf{C}(-1)\mathbf{D}_1\mathbf{C}(0)^+ \quad (4.63)$$

$$\mathbf{G}(0) = \mathbf{C}(0)\mathbf{D}_0\mathbf{C}(0)^+ \quad (4.64)$$

$$\mathbf{G}(1) = \mathbf{C}(0)\mathbf{D}_1\mathbf{C}(0)^+ \quad (4.65)$$

$$\mathbf{G}(2) = \mathbf{0} \quad (4.66)$$

Um Verfeinerbarkeit von $\Phi_{\mathbf{C}}$ sicherzustellen, ist es natürlich nötig, dass diese $\mathbf{G}(k)$ auch Gl. (4.46-4.51) erfüllen. Zudem muss auch Orthonormalität gelten.

4.9 Hinreichendes Kriterium für Approximationsordnung

Notwendige oder hinreichende Kriterien für eine gewisse Approximationsordnung von Multiskalenanalysen gibt es anhand des Wavelets [Jia91] oder der Verfeinerungskoeffizienten [JRZ96].

Aufgrund des experimentellen Befundes, dass Multiskalenanalysen aus stückweisen Polynomen immer eine gewisse Approximationsordnung besitzen, unabhängig davon, ob die entsprechenden Gleichungen, die die Approximationsordnung erzwingen sollen, beteiligt sind oder nicht, lag die Vermutung nahe, dass die Approximationsordnung schon allein durch die stückweisen polynomiellen Abschnitte der Skalierungsfunktion erzwungen wird. Tatsächlich gilt folgende allgemeine Eigenschaft.

Satz 4.9.1 (Approximationsordnung durch stückweise polynomielle Skalierungsfunktion) *Sei Φ eine Skalierungsfunktion mit kompaktem Träger zu einer Multiskalenanalyse (V_p) . Existiert eine Komponente ϕ_r der Skalierungsfunktion, die auf einem endlichen Intervall polynomiell vom Grad $N - 1$ ist, d.h. gibt es $I_1, I_2 \in \mathbb{R}, I_2 > I_1$ und $a_i \in \mathbb{R}, i = 0..N - 1, a_{N-1} \neq 0$ mit*

$$\phi_r(t) = \sum_{i=0}^{N-1} a_i t^i, \quad \text{für } t \in [I_1, I_2], \quad (4.67)$$

dann gilt

$$\Pi_{N-1} \subset \mathbb{S}(\Phi), \quad (4.68)$$

d.h. (V_p) hat Approximationsordnung N .

Beweis Sei W die Länge des Trägers von Φ . Es gibt ein $n \in \mathbb{N}$, so dass $[i_1, i_2] \subseteq [I_1 2^n, I_2 2^n]$ mit $i_1, i_2 \in \mathbb{Z}$ und $i_2 - i_1 \geq N + 2W$. Aufgrund der Verfeinerungseigenschaft existieren Vektoren $\mathbf{x}(k)$ mit $\phi_r(t/2^n) = \sum_k \mathbf{x}(k) \Phi(t - k)$. Man erhält dann: $\sum_k \mathbf{x}(k) \Phi(t - k) = \sum_{i=0}^{N-1} \tilde{a}_i t^i$ für $t \in [i_1, i_2]$ und $\tilde{a}_i := 2^{-in} a_i$. Alle Funktionen $\sum_{i=0}^{N-1} \tilde{a}_i (t+l)^i$ mit $l = 0..N - 1$ sind für $t \in [i_1, i_1 + 2W]$ also durch eine Linearkombination in $\mathcal{B}_0(\Phi)$ exakt darstellbar. (Die Koeffizientenfolge lautet $(\mathbf{x}(\cdot + l))$). Da die N verschobenen Polynome $\sum_{i=0}^{N-1} \tilde{a}_i (t+l)^i$ alle vom Grad $N - 1$ sind, sind sie linear unabhängig. Es existieren dann Koeffizienten f_{lm} , sodass $t^m = \sum_l f_{lm} \sum_{i=0}^{N-1} \tilde{a}_i (t+l)^i$ für $m = 0..N - 1$. Dies zeigt, dass auf dem Intervall $[i_1, i_1 + 2W]$ alle Monome $t^m = \sum_k (\sum_l f_{lm} \mathbf{x}(k+l)) \Phi(t - k)$, $m = 0..N - 1$ dargestellt werden können. Dann können sie aber nach Satz 2.2.5 für alle $t \in \mathbb{R}$ dargestellt werden. Also gilt $\Pi_{N-1} \subset \mathbb{S}(\Phi)$ \square

4.10 Eigenschaften von Skalierungsfunktionen mit verschwindenden Momenten

Satz 4.10.1 (Koeffizientenfolge der Polynome bei Approximationsordnung) *Ist (V_p) eine orthogonale Multiskalenanalyse von Vielfachheit R und Φ eine or-*

thonormale Skalierungsfunktion mit kompaktem Träger und Approximationsordnung N , d.h. $\Pi_{N-1} \subset \mathbb{S}(\Phi)$, dann gilt für die Koeffizienten des Polynomvektors $\mathbf{p}(t)$ vom Grad $N - 1$ bezüglich $\mathcal{B}_0(\Phi)$:

$$\mathbf{p}(t) = \sum_k \mathbf{T}(-k) \mathbf{M} \Phi(t - k). \quad (4.69)$$

Dabei ist $\mathbf{M} \in \mathbb{R}^{N \times R}$ die in Def. 3.2.2 definierte Momentenmatrix mit den ersten N Momenten von Φ .

Beweis Da $\Pi_{N-1} \subset \mathbb{S}(\Phi)$, sei $\mathbf{p}(t) = \sum_k \mathbf{P}(k) \Phi(t - k)$. Einerseits ist $\mathbf{p}(t - l) = \mathbf{T}(l) \mathbf{p}(t) = \sum_k \mathbf{T}(l) \mathbf{P}(k) \Phi(t - k)$ und andererseits $\mathbf{p}(t - l) = \sum_k \mathbf{P}(k - l) \Phi(t - k)$. Ein Koeffizientenvergleich für $k = 0$ ergibt $\mathbf{P}(l) = \mathbf{T}(-l) \mathbf{P}(0)$. Wegen der Orthonormalität von Φ ist $\mathbf{P}(0) = \langle \mathbf{p}, \Phi \rangle$. Dies ist gerade die Definition der Momentenmatrix \mathbf{M} der Skalierungsfunktion. \square

Bemerkung Dieser Satz zeigt, dass die Koeffizienten der Polynome nur durch die Momentenmatrix bestimmt sind, da die $\mathbf{T}(k)$ feste, von Φ unabhängige Matrizen sind. Insbesondere wird hier ein Zusammenhang zwischen der Approximationsordnung und den verschwindenden Momenten der Skalierungsfunktion deutlich. Zudem ist ersichtlich, dass wenn die Skalierungsfunktion – inklusive dem ersten Moment – verschwindende Momente besitzt, sie keine Approximationsordnung haben kann, da dann die ersten Komponenten von Gl. (4.69) sich alle zu Null ergeben. Dies ist durch die Dreieckstruktur von $\mathbf{T}(k)$ begründet.

Satz 4.10.2 (Koeffizientenfolge der Polynome bei verschwindenden Momenten der Skalierungsfunktion) Sei (V_p) eine orthogonale Multiskalenanalyse von Vielfachheit R mit einer orthonormale Skalierungsfunktion mit kompaktem Träger und Approximationsordnung N , d.h. $\Pi_{N-1} \subset \mathbb{S}(\Phi)$. Besitzt Φ L nichtzentrale verschwindende höhere Momente, d.h. $\mathbf{m}_{l,\tau} = \mathbf{0}$, $l = 1 \dots L$, dann gilt für die Koeffizienten des Polynomvektors $\mathbf{p}(t)$ vom Grad $\min(N - 1, L)$:

$$\mathbf{p}(t) = \sum_k \mathbf{p}(k + \tau) \mathbf{m}_0^T \Phi(t - k). \quad (4.70)$$

Beweis Die Anzahl der Zeilen von \mathbf{M} bzw. \mathbf{M}_τ und die Dimension von \mathbf{T} ist $\min(N, L + 1)$. Die nichtzentrale Momentenmatrix \mathbf{M}_τ ist durch die Matrix der zentralen Momente \mathbf{M} durch $\mathbf{M}_\tau = \mathbf{T}(\tau) \mathbf{M}$ gegeben, siehe Gl. (3.2.4). Dann hat Gl. (4.69) die Form $\mathbf{p}(t) = \sum_k \mathbf{T}(-k) \mathbf{T}(-\tau) \mathbf{M}_\tau \Phi(t - k)$. Es gilt $\mathbf{T}(-k) \mathbf{T}(-\tau) = \mathbf{T}(-k - \tau)$ und die Matrix \mathbf{M}_τ enthält bis auf die erste Zeile (\mathbf{m}_0^T) nur Nullzeilen. $\mathbf{T}(-k - \tau) \mathbf{M}_\tau$ ist daher gleich der ersten Spalte von $\mathbf{T}(-k - \tau)$ multipliziert mit \mathbf{m}_0^T . Aus der Definition von \mathbf{T} Gl. (2.49d) ist ersichtlich, dass diese erste Spalte die Potenzen von $k + \tau$ enthält. Also folgt $\mathbf{p}(t) = \sum_k \mathbf{p}(k + \tau) \mathbf{m}_0^T \Phi(t - k)$. \square

Satz 4.10.3 (Minimale Länge der Skalierungsfunktion bei verschwindenden Momenten) Sei (V_p) eine orthogonale Multiskalenanalyse mit einer orthonormalen Skalierungsfunktion mit kompaktem Träger und Approximationsordnung N , d.h. $\Pi_{N-1} \subset \mathbb{S}(\Phi)$. Besitzt Φ L verschwindende höhere Momente um τ , d.h. $\mathbf{m}_{l,\tau} = 0$, $l = 1 \dots L$, dann hat Φ mindestens eine Länge von $\min(N, L + 1)$.

Beweis Die linke Seite von Gl. (4.70) besteht auf $t \in [0, t]$ aus $\min(N, L + 1)$ linear unabhängigen Komponenten. Dann muss auch $\{\mathbf{m}_0^T \Phi(\cdot - k) \chi_{[0,1]} | k \in \mathbb{Z}\}$ aus mindestens $\min(N, L + 1)$ linear unabhängigen Elementen bestehen. Insbesondere muss die Länge von Φ mindestens $\min(N, L + 1)$ sein, da sonst weniger als $\min(N, L + 1)$ nicht verschwindende Elemente vorhanden sind. \square

Beispiel Falls Gleichheit gilt, so ist $\mathbf{m}_0^T \Phi$ vollständig festgelegt. Dieser Sachverhalt lässt sich auch durch ein Beispiel verdeutlichen. Sei der Träger von Φ mit Approximationsordnung 2 gleich $[-1, 1]$ mit $\tau = 0$ und $L = 1$: es liege also ein zentrales verschwindendes Moment vor. Dann ist für $t \in [0, 1)$ Gl. (4.70) gegeben durch

$$\begin{pmatrix} 1 \\ t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0(t) \\ x_1(t) \end{pmatrix} \quad (4.71)$$

mit der "Rechten und Linken Hälfte" von $\mathbf{m}_0 \Phi(t)$: $x_0(t) = \mathbf{m}_0^T \Phi(t) \chi_{[0,1]}(t)$ und $x_1(t) = \mathbf{m}_0^T \Phi(t - 1) \chi_{[0,1]}(t)$. Aufgelöst ergibt sich für $t \in [0, 1)$ $x_0(t) = 1 - t$ und $x_1(t) = t$. Rückübersetzt in $\mathbf{m}_0^T \Phi(t)$ heißt dies

$$\mathbf{m}_0^T \Phi(t) = \begin{cases} 1 + t & -1 \leq t < 0 \\ 1 - t & 0 \leq t < 1 \\ 0 & \text{sonst} \end{cases} \quad (4.72)$$

Bei skalaren Skalierungsfunktionen gibt es für L verschwindende Momente der Skalierungsfunktion die folgende äquivalente Bedingung an die Verfeinerungskoeffizienten $g(k)$

$$\sum_k k^l g(k) = \sqrt{2} \delta_l, \quad l = 0, \dots, L. \quad (4.73)$$

Diese Bedingung bedeutet gerade, dass $g(k)$ verschwindende Momente hat. Ein solch schöner Zusammenhang kann für Multi-Skalierungsfunktionen allerdings nur in einer Richtung gezeigt werden.

Satz 4.10.4 Sei Φ eine Skalierungsfunktion zu einer Multiskalenanalyse. Dann gilt folgender Zusammenhang zwischen der (nichtzentralen) Momentenmatrix \mathbf{M}_τ (falls die Momente existieren) und den Verfeinerungskoeffizienten $\mathbf{G}(k)$:

$$\mathbf{M}_\tau = 2^{-1/2} \mathbf{S}(1/2) \sum_k \mathbf{T}(\tau - k) \mathbf{M}_\tau \mathbf{G}(k)^T. \quad (4.74)$$

Die Dimension von \mathbf{T} und \mathbf{S} muss zur Anzahl der betrachteten Momente passend gewählt sein (was aber stets möglich ist).

Beweis Es gilt für die ersten N Momente von Φ zusammengefasst in der Matrix $M_\tau = \langle \mathbf{p}(\cdot - \tau), \Phi \rangle$. Mit der Verfeinerbarkeit ist

$$\begin{aligned} M_\tau &= \sqrt{2} \sum_k \langle \mathbf{p}(\cdot - \tau), \Phi(2\cdot - k) \rangle \mathbf{G}(k)^T = \sqrt{2} \sum_k \langle \mathbf{p}(\cdot + k/2 - \tau), \Phi(2\cdot) \rangle \mathbf{G}(k)^T \\ &= 2^{-1/2} \sum_k \langle \mathbf{p}(\cdot/2 + k/2 - \tau), \Phi \rangle \mathbf{G}(k)^T = 2^{-1/2} \mathbf{S}(1/2) \sum_k \langle \mathbf{p}(\cdot + k - 2\tau), \Phi \rangle \mathbf{G}(k)^T \\ &= 2^{-1/2} \mathbf{S}(1/2) \sum_k \mathbf{T}(\tau - k) \langle \mathbf{p}(\cdot - \tau), \Phi \rangle \mathbf{G}(k)^T = 2^{-1/2} \mathbf{S}(1/2) \sum_k \mathbf{T}(\tau - k) M_\tau \mathbf{G}(k)^T \quad \square \end{aligned}$$

Satz 4.10.5 Sei Φ eine Skalierungsfunktion zu einer Multiskalenanalyse und besitze L verschwindende höhere Momente um τ , d.h. $\mathbf{m}_{l,\tau} = \mathbf{0}$, $l = 1, \dots, L$. Dann gilt:

$$\left(\sum_k (k - \tau)^l \mathbf{G}(k) \right) \mathbf{m}_0 = \sqrt{2} \delta_l \mathbf{m}_0, \quad l = 0, \dots, L. \quad (4.75)$$

Beweis Betrachte die ersten L Momente von Φ in M_τ . Die Matrix besteht nur als Nullzeilen, außer der ersten Zeile (dem nullten zentralen Moment \mathbf{m}_0^T). Setzt man dies in Gl. (4.74) ein, so erkennt man, dass $\mathbf{T}(\tau - k) M_\tau$ nur aus dem Produkt der ersten Zeile von $\mathbf{T}(\tau - k)$ und \mathbf{m}_0^T besteht. Die erste Zeile sind aber gerade die Potenzen von $k - \tau$, also $\mathbf{p}(k - \tau)$. Es ergibt sich also: $M_\tau = 2^{-1/2} \mathbf{S}(1/2) \sum_k \mathbf{p}(k - \tau) \mathbf{m}_0^T \mathbf{G}(k)^T$. Wird dieses zeilenweise zerlegt geschrieben, so ergibt sich nach einer Transposition und Multiplikation mit 2^n Gl. (4.75). \square

Bemerkung Es gibt eine deutliche Ähnlichkeit von Gl. (4.75) zu der Eigenschaft im skalaren Fall Gl. (4.73). Im vektoriellen Fall ist das nullte Moment jedoch ein Vektor und keine Zahl mehr, weswegen es nicht wegkürzen lässt. Dies ist auch der Grund dafür, dass die Umkehrung nicht gilt. Analog dem Beweis der Umkehrung für den skalaren Fall würde dafür $\sum_k (k - \tau)^l \mathbf{G}(k) \mathbf{m}_l = \mathbf{0}$ für $l > 0$ benötigt. In Analogie dazu müsste dann $\sum_k (k - \tau)^l \mathbf{G}(k) = \mathbf{0}$ gelten, was aber im vektoriellen Fall nicht folgt.

Bemerkung Für $l = 0$ wird auch deutlich, dass das nullte Moment der Skalierungsfunktion immer ein Eigenvektor zum Eigenwert $\sqrt{2}$ von der Summe der Verfeinerungskoeffizienten $\sum_k \mathbf{G}(k)$ ist. Dies gilt für allgemeine Skalierungsfunktionen unabhängig von Orthonormalität und verschwindenden Momenten.

5 Skalierungsfunktionen aus stückweisen Polynomen

5.1 Die Basis aus stückweisen Polynomen

Um Skalierungsfunktionen aus stückweisen Polynomen darstellen zu können, wird nun eine Basis aus orthogonalen Polynomen definiert. Sie besteht aus allen Polynomen vom Grad $< N$ die jeweils orthogonal auf den $Q \in \{1, 2, \dots\}$ äquidistanten Unterteilungen des Einheitsintervalls sind. Es ergeben sich dann QN Funktionen in der Basis B . Es lassen sich alle stückweisen Polynome mit endlich vielen Teilintervallen darstellen. Sind die darzustellenden Teilintervalle allerdings nicht äquidistant, wird eine sehr große Zahl an äquidistanten Teilintervallen in dieser Darstellung benötigt. Solche Skalierungsfunktionen wie in [SS94], wo die Intervalle unendlich fein werden, können mit der vorliegenden Darstellung nicht behandelt werden. Die Skalierungsfunktionen mit Länge 1 und $Q = 1$, also $\Phi = 1$, ist bereits von Alpert diskutiert [Alp92]. Wie alle Skalierungsfunktionen deren Träger auf $[0, 1]$ beschränkt ist, ist sie unstetig. Grund: Ist eine Skalierungsfunktion an allen ganzzahligen Stellen gleich Null, so ist sie aufgrund der Verfeinerbarkeit auch an allen Stellen $2^{-k}p$ mit $k, p \in \mathbb{Z}$ gleich Null.

Definition 5.1.1 (Basis für stückweise Polynome) Sei $Q \in \mathbb{N}$, $Q > 0$ die Anzahl der Unterteilungen des Einheitsintervalls und $N - 1$ der maximale Grad der stückweisen Polynome. In diesem Kapitel wird die folgende Funktion $B : \mathbb{R} \rightarrow \mathbb{R}^{NQ} : t \mapsto B(t)$ als Basis im Sinne von Abschnitt 4.1 benutzt:

$$B(t) = \sqrt{Q} \begin{pmatrix} 1(Qt) \chi_{[0,1]}(Qt) \\ 1(Qt - 1) \chi_{[0,1]}(Qt - 1) \\ \vdots \\ 1(Qt - Q + 1) \chi_{[0,1]}(Qt - Q + 1) \end{pmatrix}. \quad (5.1)$$

Hierbei ist $1 : \mathbb{R} \rightarrow \mathbb{R}^N : t \mapsto 1(t)$ der Vektor der auf $[0, 1]$ orthonormalen Legendre Polynome vom Grad $< N$ (Siehe Abschnitt 2.5.2).

Für $Q = 2$ und $N = 3$ ist B in Abb. 5.1 dargestellt. Es ist klar, dass die so definierte Basis auch die Verfeinerungsgleichung erfüllt, da die Komponenten alle polynomiell vom Grad $< N$ sind.

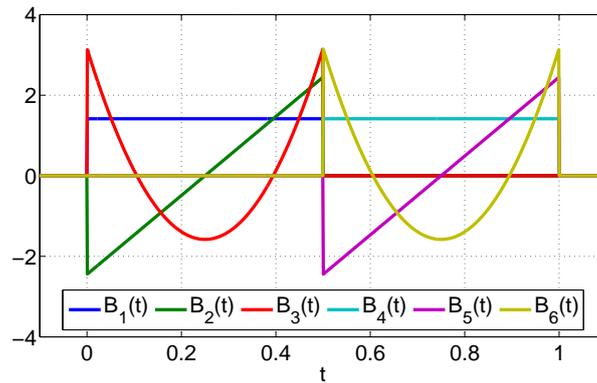


Abbildung 5.1: Die NQ Komponenten der Basisfunktion $B(t)$, gezeigt für $Q = 2$ (also zwei Teilintervalle) und $N = 3$ (Polynome von Grad 0 bis Grad 2). Die Funktionen bilden ein Orthonormalsystem.

Satz 5.1.1 (Verfeinerung der Basis aus stückweisen Polynomen). Die Matrizen $D_0 \in \mathbb{R}^{NQ \times NQ}$ und $D_1 \in \mathbb{R}^{NQ \times NQ}$ der Verfeinerungsgleichung aus Gl. (4.1e) für die Basis B Gl. (5.1) mit maximalen Grad $N - 1$ und Q Teilintervallen haben die folgende Blockmatrix Gestalt, wobei die Teilmatrizen aus $\mathbb{R}^{N \times N}$ sind. Für gerades Q gilt

$$D_0 = \begin{pmatrix} A & B & & & & \\ & A & B & & & \\ & & \ddots & \ddots & & \\ & & & A & B & \\ & & & & & \end{pmatrix}, \quad D_1 = \begin{pmatrix} & & & & & \\ & A & B & & & \\ & & A & B & & \\ & & & \ddots & \ddots & \\ & & & & A & B \end{pmatrix} \quad (5.2)$$

und für ungerades Q ist

$$D_0 = \begin{pmatrix} A & B & & & \\ & A & B & & \\ & & \ddots & \ddots & \\ & & & A & \\ & & & & \end{pmatrix}, \quad D_1 = \begin{pmatrix} & & & & \\ & B & & & \\ & & A & B & \\ & & & \ddots & \ddots & \\ & & & & A & B \end{pmatrix} \quad (5.3)$$

mit der Kurzschreibweise

$$A = LS(1/2)L^{-1}, \quad B = LS(1/2)T(-1)L^{-1}. \quad (5.4)$$

Beweis Es gilt: $\chi_{[0,1]}(Qt/2-k) = \chi_{[0,1]}(Qt-2k) + \chi_{[0,1]}(Qt-2k-1)$. Zudem ist $l(Qt/2-k) = l((Qt-2k)/2) = \mathbf{LS}(1/2)\mathbf{L}^{-1}(Qt-2k) = \mathbf{LS}(1/2)\mathbf{T}(-1)\mathbf{L}^{-1}(Qt-2k-1)$. Es gilt also: $l(Qt/2-k)\chi_{[0,1]}(Qt/2-k) = \mathbf{LS}(1/2)\mathbf{L}^{-1}(Qt-2k)\chi_{[0,1]}(Qt-2k) + \mathbf{LS}(1/2)\mathbf{T}(-1)\mathbf{L}^{-1}((Qt-2)/2)\chi_{[0,1]}(Qt-2k-1)$. Schreibt man dieses für $k = 0, 1, \dots, Q-1$ in Matrixform so erhält man das angegebene \mathbf{D}_0 und \mathbf{D}_1 . \square

Bemerkung Für den wichtigen Spezialfall $Q = 2$ ist

$$\mathbf{D}_0 = \begin{pmatrix} \mathbf{LS}(1/2)\mathbf{L}^{-1} & \mathbf{LS}(1/2)\mathbf{T}(-1)\mathbf{L}^{-1} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{D}_1 = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{LS}(1/2)\mathbf{L}^{-1} & \mathbf{LS}(1/2)\mathbf{T}(-1)\mathbf{L}^{-1} \end{pmatrix}. \quad (5.5)$$

Beispiel Für $Q = 2$ und $N = 3$, d.h. die Basis besteht aus Polynomen höchstens von Grad 2, ist:

$$\mathbf{D}_0 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -1/2\sqrt{3} & 1/2 & 0 & 1/2\sqrt{3} & 1/2 & 0 \\ 0 & -1/4\sqrt{15} & 1/4 & 0 & 1/4\sqrt{15} & 1/4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (5.6)$$

$$\mathbf{D}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ -1/2\sqrt{3} & 1/2 & 0 & 1/2\sqrt{3} & 1/2 & 0 \\ 0 & -1/4\sqrt{15} & 1/4 & 0 & 1/4\sqrt{15} & 1/4 \end{pmatrix}. \quad (5.7)$$

Nach Konstruktion ist es klar, dass mittels $\mathbf{B}(t)$ die Polynome vom Grad $< N$ dargestellt werden können. Die zugehörigen Koeffizienten sind Eigenvektoren von \mathbf{D}_i zu Eigenwerten 2^{-n} (siehe Abschnitt 4.1.3). Die Matrizen \mathbf{D}_0 und \mathbf{D}_{-1} haben also mindestens die Eigenwerte 2^{-k} ($k = 0, 1, \dots, N-1$) und 0. Ein Eigenvektor zur Matrix \mathbf{D}_0^T zum Eigenwert 2^{-n} ist der Vektor $\mathbf{p}_n(0)$ aus der eindeutigen Darstellung $\Phi_{\mathbf{p}_0}(t) = t^n$. Der analoge Eigenvektor zu \mathbf{D}_1 ist $\mathbf{p}_n(-1)$ aus der gleichen Zerlegung. Die Eigenvektoren zu \mathbf{D}_0 und \mathbf{D}_{-1} zum Eigenwert Null sind die Einheitsvektoren. Man prüft leicht nach, dass die Matrizen nicht normal sind, d.h. eine unitäre (orthogonale) Diagonalisierung existiert nicht. Für den Fall $Q = 2$ sind jedoch alle bekannten $2N$ Eigenwerte einfach, wodurch eine Basis aus den genannten Eigenvektoren existiert.

Satz 5.1.2 (Diagonalisierung von \mathbf{D}_0 und \mathbf{D}_1 für den Fall $Q = 2$). Für $Q = 2$ gilt für die Verfeinerungskoeffizienten \mathbf{D}_0 und \mathbf{D}_1 aus Gl. (5.5):

$$\mathbf{U}_i \mathbf{D}_i \mathbf{U}_i^{-1} = \Sigma, \quad i = 1, 2 \quad (5.8)$$

$$U_0 = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{S}(1/2)\mathbf{L}^{-1} & \mathbf{S}(1/2)\mathbf{T}(-1)\mathbf{L}^{-1} \end{pmatrix}, \quad U_0^{-1} = \begin{pmatrix} -\mathbf{L}\mathbf{T}(-1)\mathbf{L}^{-1} & \mathbf{L}\mathbf{S}(2) \\ \mathbf{I} & \mathbf{0} \end{pmatrix}, \quad (5.9)$$

$$U_1 = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{S}(1/2)\mathbf{T}(2)\mathbf{L}^{-1} & \mathbf{S}(1/2)\mathbf{T}(1)\mathbf{L}^{-1} \end{pmatrix}, \quad U_1^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{L}\mathbf{T}(1)\mathbf{L}^{-1} & \mathbf{L}\mathbf{T}(-1)\mathbf{S}(2) \end{pmatrix}, \quad (5.10)$$

und

$$\Sigma = \text{diag}(0, \dots, 0, 1, 2^{-1}, \dots, 2^{-N+1}) = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}(1/2) \end{pmatrix}. \quad (5.11)$$

Beweis Durch Nachrechnen zeigt man unmittelbar, dass $U_0 U_0^{-1} = \mathbf{I}$ sowie $U_1 U_1^{-1} = \mathbf{I}$, bei letzterem muss beachtet werden, dass sowohl $\mathbf{T}(0) = \mathbf{I}$ (Gl. (2.49j)) und $\mathbf{S}(1/2)\mathbf{S}(2) = \mathbf{I}$ (Gl. (2.49e)), als auch $\mathbf{T}(a)\mathbf{T}(b) = \mathbf{T}(a+b)$ für alle a, b (Gl. (2.49i)) gilt. Beim Nachrechnen von Gl. (5.8) muss zusätzlich $\mathbf{T}(-2)\mathbf{S}(2) = \mathbf{S}(2)\mathbf{T}(-1)$ (Gl. (2.49k)) berücksichtigt werden. \square

5.2 Lösen der Koeffizientengleichungen

Um Skalierungsfunktionen Φ_C zu einer Multiskalenanalyse zu konstruieren, müssen die Koeffizientengleichungen für (C_p) mit $C_p \in \mathbb{R}^{R \times NQ}$ und $\text{supp}(C) = [-P, P-1]$ gelöst werden. Dazu geht man wie folgt vor:

1. Wahl der Parameter der Basis:

- a) $Q \in \mathbb{N} \setminus \{0\}$: Anzahl der Unterteilungen pro Einheitsintervall auf dem jeweils ein Teilabschnitt der stückweisen Polynome definiert ist.
- b) $N \in \mathbb{N} \setminus \{0\}$: Maximaler Grad der stückweisen Polynome in der Basis.

2. Wahl der Parameter der Skalierungsfunktion:

- a) $P \in \mathbb{N} \setminus \{0\}$: Maximale Länge der Skalierungsfunktion. Es gilt $\text{supp}(\Phi_C) \subseteq [-P, P]$ wobei $C(p)$ nur für $p = -P, \dots, P-1$ ungleich Null sein kann.
- b) $R \in \mathbb{N} \setminus \{0\}$: Vielfachheit der Skalierungsfunktion. Für $R = 1$ ergeben sich herkömmliche skalare Skalierungsfunktionen.

3. Feste Eigenschaften

- a) Orthonormalität: Für Orthonormalität von Φ_C muss C Gl. (4.9) für $k = 0, 1, \dots, P$ erfüllen.
- b) Verfeinerbarkeit: Für Verfeinerbarkeit muss Gl. (4.28) gelten, wobei für F Gl. (4.11) eingesetzt wird.

4. Wahl der zusätzlichen Eigenschaften

- a) Stetigkeit: Um Stetigkeit von Φ_C zu gewährleisten, muss die Koeffizientenfolge C Gl. (4.22) erfüllen.
- b) Verschwindende Momente: Um L verschwindende Momente von Φ_C zu realisieren, muss die Koeffizientenfolge C Gl. (4.26) erfüllen. Bei verschwindenden Momenten muss nach Satz 4.10.3 Φ_C aber eine Mindestlänge $2P \geq \min(N, L + 1)$ besitzen.
- c) Approximationsordnung: Nach Satz 4.10.2 haben Skalierungsfunktionen die polynomiell sind, automatisch die Approximationsordnung vom Grad + 1 des polynomiellen Abschnitts. Daher kann die Approximationsordnung sehr leicht über den Ansatz, also den maximalen Grad N eingestellt werden. Es muss lediglich sichergestellt werden, dass Φ_C dann auch diesen maximalen Grad N hat, d.h. die entsprechenden Koeffizienten dürfen nicht alle verschwinden. Die Approximationsordnung lässt sich jedoch auch über Gl. (4.28) mit Gl. (4.3) erzwingen.

In der gezeigten Formulierung ergibt sich typischerweise ein überbestimmtes Gleichungssystem. Doch falls Lösungen existieren, ist die Lösung nicht eindeutig, da beispielsweise eine orthogonale Transformation der C_p auch Lösung ist. Falls der Träger einer *Komponente* sich nicht über den ganzen Bereich von $[-P, P]$ erstreckt, so ist diese Funktion ganzzahlig "verschiebar" ohne, dass Stetigkeit, Verfeinerbarkeit, Approximationsordnung und Orthogonalität verletzt werden.

Schreibt man die Gleichungen als skalares Gleichungssystem in den Elementen der Matrizen C_p so erhält man ein System von multivariaten Polynomen. Die Ordnung der Gleichungen für Verfeinerbarkeit ist 3, die für Orthogonalität und Approximationsordnung 2 und die Bedingungen für Stetigkeit und verschwindende Momente sind linear.

5.3 Numerisch konstruierte Skalierungsfunktionen

Obwohl symbolische Verfahren zur Lösung dieser Gleichungen existieren (mittels Gröbner Basen), sind diese leider nicht anwendbar, da der Rechen- und Speicheraufwand bei der vorhandenen Zahl von Gleichungen und Variablen zu hoch für heutige Computer ist.

Die oben angegebenen Gleichungen wurden in Matlab programmiert und mittels einem numerischen Optimierungsalgorithmus eine Approximation einer Nullstelle bestimmt. Das verwendete Verfahren ist eine Hybridlösung aus dem Gauss-Newton und Levenberg-Marquard Algorithmus (Matlab Kommando *fsolve*). Die Anfangswerte für die zu bestimmenden Koeffizienten wurden zufällig bestimmt. Falls der Algorithmus dann für mehrere Durchläufe keine Lösung findet, d.h. gegen ein nicht verschwindendes (lokales) Mini-

imum der Norm des Gleichungssystems konvergiert, gibt dies ein Hinweis darauf, dass eventuell keine Lösung existiert.

Während es problemlos möglich ist, auf dem Bereich $[-1, 1]$ orthonormale und verfeinerbare Skalierungsfunktionen mit mehr als 2 Komponenten aus stückweisen Polynomen zu konstruieren, gelingt dies wegen Satz 4.10.3 nicht, falls noch verschwindende Momente der Skalierungsfunktion gefordert werden. Daher geht man zur Länge 4 über, wobei die numerische Lösung mit hinreichender Genauigkeit eine Nullstelle darstellt. Die Gleichungen sind hierzu alle so formuliert, dass eine Seite Null lautet. Das numerische Abbruchkriterium ist so eingestellt, dass die Frobeniusnorm der jeder einzelnen Matrixgleichung für Orthonormalität und Verfeinerbarkeit kleiner als 10^{-6} ist. Die Frobeniusnorm $\|\mathbf{A}\|_F^2$ einer Matrix $\mathbf{A} = (a_{ij})_{ij}$ ist definiert als der quadratische Mittelwert der Einträge:

$$\|\mathbf{A}\|_F^2 = \sum_{i,j} a_{ij}^2 \quad (5.12)$$

5.3.1 Skalierungsfunktion mit Länge 4 und einem verschwindendem Moment

Der gewählte Ansatz lautet

- $N = 3$: Polynome bis zum Grad 2 in der Basis.
- $Q = 2$: Basis mit Unterteilung der Einheitsintervalle in 2 Teile.
- $R = 3$: Komponenten der Skalierungsfunktion.
- $P = 2$: Länge der Skalierungsfunktion 4 mit $\text{supp}(\Phi_{\mathbf{C}}) \subset [-2, 2]$.
- $L = 1$: Verschwindendes erstes Moment der Skalierungsfunktion.

Die Skalierungsfunktion wird dann mit $\mathbf{C}(k) \in \mathbb{R}^{3 \times 6}$ so dargestellt:

$$\Phi_{\mathbf{C}}(t) = \sum_{k=-2}^1 \mathbf{C}(k) \mathbf{B}(t-k) \quad (5.13)$$

mit $\mathbf{B}(t) \in \mathbb{R}^6$ und

$$\mathbf{B}(t) = \sqrt{2} \begin{pmatrix} \chi_{[0,1]}(2t) \\ \chi_{[0,1]}(2t) \sqrt{3}(4t-1) \\ \chi_{[0,1]}(2t) \sqrt{5}(24t^2-12t+1) \\ \chi_{[0,1]}(2t-1) \\ \chi_{[0,1]}(2t-1) \sqrt{3}(4t-3) \\ \chi_{[0,1]}(2t-1) \sqrt{5}(24t^2-36t+13) \end{pmatrix}. \quad (5.14)$$

Das angegebene \mathbf{B} ist in Abb. 5.1 dargestellt.

Die numerisch gefundene Skalierungsfunktion Φ_C ist in Abb. 5.2 dargestellt und besitzt die folgenden nichttrivialen Koeffizientenmatrizen:

$$C(-2) \approx 10^{-3} \begin{pmatrix} 3.17 & -30.80 & -25.28 & -163.73 & -280.33 & -191.64 \\ -44.90 & -42.89 & -13.14 & -31.53 & -209.87 & -214.91 \\ 20.52 & 6.62 & -4.05 & -49.27 & -17.35 & 18.85 \end{pmatrix}, \quad (5.15a)$$

$$C(-1) \approx 10^{-3} \begin{pmatrix} 195.22 & 598.41 & -105.79 & 517.47 & -157.28 & 91.79 \\ 99.86 & 417.99 & -112.47 & -470.25 & -355.13 & 191.19 \\ -135.01 & -49.05 & 5.76 & 580.37 & -160.13 & -476.20 \end{pmatrix}, \quad (5.15b)$$

$$C(0) \approx 10^{-3} \begin{pmatrix} 363.47 & -56.15 & -4.66 & 63.55 & -83.88 & 21.00 \\ -533.00 & 80.20 & 6.29 & -128.96 & 120.94 & -18.60 \\ -586.12 & 109.81 & 6.49 & 4.60 & 136.25 & -67.08 \end{pmatrix}, \quad (5.15c)$$

$$C(1) \approx 10^{-3} \begin{pmatrix} -29.91 & 3.34 & 0.41 & -13.99 & 6.70 & 1.07 \\ 33.50 & -3.73 & -0.46 & 15.68 & -7.51 & -1.19 \\ 77.93 & -8.68 & -1.06 & 36.47 & -17.47 & -2.77 \end{pmatrix} \quad (5.15d)$$

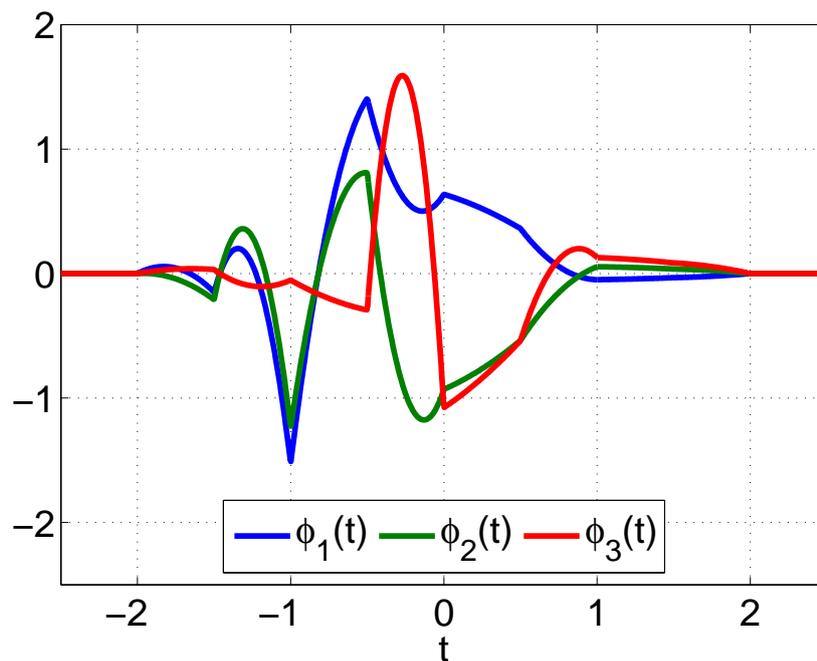


Abbildung 5.2: Die numerisch bestimmte orthonormale Skalierungsfunktion mit 3 Komponenten, Approximationsordnung 3, und verschwindendem ersten Moment.

Die Verfeinerungskoeffizienten $10^3 \cdot G(k)$ für $k = -3P + 1, \dots, 3P - 1$, also $k = -5, \dots, 5$ sind gegeben durch: (Die Aufzählung geht dabei von links nach

rechts und von oben nach unten)

$$\begin{pmatrix} -0.57 & 0.64 & 1.49 \\ 0.79 & -0.88 & -2.06 \\ -0.56 & 0.63 & 1.46 \end{pmatrix}, \quad \begin{pmatrix} 9.41 & -12.91 & -17.56 \\ -5.71 & 9.66 & 5.28 \\ 6.06 & -9.11 & -9.01 \end{pmatrix}, \quad (5.16a)$$

$$\begin{pmatrix} 24.49 & 2.87 & -70.31 \\ 14.09 & -27.49 & -116.12 \\ 7.20 & 8.59 & 12.87 \end{pmatrix}, \quad \begin{pmatrix} -167.26 & 356.12 & 112.76 \\ -16.50 & 267.47 & 69.61 \\ -94.44 & 79.50 & 63.73 \end{pmatrix}, \quad (5.16b)$$

$$\begin{pmatrix} 662.37 & -407.45 & 63.93 \\ 304.91 & 39.31 & 204.41 \\ 156.89 & -307.36 & -425.03 \end{pmatrix}, \quad \begin{pmatrix} 150.36 & -358.93 & -106.06 \\ -712.53 & 387.74 & -87.06 \\ 306.80 & 463.99 & 491.67 \end{pmatrix}, \quad (5.16c)$$

$$\begin{pmatrix} 160.56 & -160.85 & 24.13 \\ -222.71 & 234.44 & -6.20 \\ -224.09 & 211.90 & -78.04 \end{pmatrix}, \quad \begin{pmatrix} -46.20 & 0.07 & -6.16 \\ 39.99 & 8.93 & -0.51 \\ 151.42 & -27.15 & 36.08 \end{pmatrix}, \quad (5.16d)$$

$$\begin{pmatrix} -13.72 & 15.37 & 2.82 \\ 15.17 & -17.05 & -3.31 \\ 36.49 & -40.67 & -6.81 \end{pmatrix}, \quad \begin{pmatrix} -2.63 & 2.21 & -1.69 \\ 2.95 & -2.47 & 1.89 \\ 6.86 & -5.75 & 4.40 \end{pmatrix}, \quad (5.16e)$$

$$\begin{pmatrix} -0.06 & 0.05 & -0.04 \\ 0.06 & -0.05 & 0.05 \\ 0.15 & -0.12 & 0.11 \end{pmatrix}. \quad (5.16f)$$

6 Zusammenfassung

Das Ziel der vorliegenden Arbeit war es, Multi-Coiflets zu konstruieren, d.h. eine vektorielle Skalierungsfunktion zu finden, die verschwindende höhere Momente besitzt und eine Multiskalenanalyse erzeugt. Es bot sich an diese aus stückweisen Polynomen zu konstruieren. Ausgehend von der Beobachtung, dass eine Skalierungsfunktion mit kompaktem Träger nach der Wahl einer geeigneten polynomiellen Basis durch endlich viele Koeffizienten dargestellt werden kann, wurde eine geeignete Theorie für diese Darstellung entwickelt, in der Hoffnung, dass diese Gleichungen lösbar sind. Die üblichen Forderungen an die Skalierungsfunktion, wie Orthonormalität, Verfeinerbarkeit, Approximationsordnung, Stetigkeit und verschwindende Momente wurden anhand dieser Koeffizienten formuliert. Es zeigte sich, dass diese Koeffizienten am besten durch eine Matrixfolge dargestellt werden.

Um größtmögliche Allgemeinheit zu gewährleisten, wurden die Eigenschaften zuerst für eine allgemeine Basis hergeleitet (Kapitel 4) und erst in Kapitel 5 auf stückweise Polynome spezialisiert. Es zeigte sich zudem, dass sich die Wahl der Basis ausschließlich in den Matrizen D_0 und D_1 zur Verfeinerung der Basis widerspiegelt. Für die stückweisen Polynome ist ein expliziter Ausdruck für diese Matrizen angegeben. Für den Fall von zwei Unterteilungen des Einheitsintervalls existiert sogar eine Eigenwertzerlegung dieser Matrizen, die auch gefunden und explizit angegeben wurde. Zudem wurde ein Zusammenhang zwischen den Eigenwerten und Eigenvektoren dieser Matrizen und den Koeffizienten für die Darstellung der Polynome gefunden und angegeben.

Es zeigte sich in der Arbeit, dass das Zusammenfassen der Polynome $1, t, t^2, \dots$ kleiner als ein bestimmter Grad, erhebliche Vorteile in der Darstellung, Herleitung und kompakten Formulierung von Zusammenhängen hat. Die Matrizen die eine Verschiebung oder Skalierung des Polynomvektors realisieren, werden häufig benutzt und sind von großem Nutzen.

Eine weitere technisch sehr wichtiges Konstrukt ist der Raum Φ der "unendlichen Linearkombinationen" der ganzzahligen Verschiebungen von Φ mit kompaktem Träger. Dieser Raum ist umfangreicher als $L^2(\mathbb{R})$ und kann daher auch die Polynome enthalten, was eine Charakterisierung der Approximationsordnung von Φ erheblich vereinfacht. Als weitere Hilfsmittel zur Charakterisierung dieser Räume wurden die Sätze 2.2.3 bis 2.2.5 hergeleitet.

Die Koeffizientengleichungen für Orthogonalität, Verfeinerbarkeit, Approximationsordnung, Stetigkeit und verschwindende Momente der Skalierungsfunktion bilden ein multivariates Polynomielles Gleichungssystem von 3 Ord-

nung, d.h. es treten neben den Produkten von 2 oder 3 Koeffizienten nur Summen auf. Obwohl symbolische Verfahren zur Lösung dieser Gleichungen existieren (mittels Gröbner Basen), sind diese leider nicht anwendbar, da der Rechen- und Speicheraufwand bei der vorhandenen Zahl von Gleichungen und Variablen zu hoch für heutige Computer ist. Zudem konnte kein anderes analytisches Verfahren zur Lösung der Gleichungen gefunden werden. Es wurde daher auf klassische numerische Algorithmen zurückgegriffen. Es zeigte sich, dass mit dem Ansatz einer Skalierungsfunktion mit Träger in $[-1, 1]$ keine Multi-Coiflets konstruiert werden können. Weiterführende theoretische Untersuchungen brachten den Grund dafür hervor. Es konnte bewiesen werden, dass für verschwindende Momente einer Skalierungsfunktion eine minimale Länge existiert (Satz 4.10.3). Dieser Satz gilt sowohl für skalare und vektorielle Skalierungsfunktionen, wobei die minimale Länge nicht von der Vielfachheit der Multiskalenanalyse abhängt, sondern nur von der Anzahl der verschwindenden Momente und der Approximationsordnung.

Durch diese Erkenntnis konnte dann numerisch eine Skalierungsfunktion mit Länge 4, Approximationsordnung 3 und einem verschwindenden Moment konstruiert werden (siehe Abschnitt 5.3.1). Es ist auf diese Weise ebenso möglich, Skalierungsfunktionen mit mehr verschwindenden Momenten und Approximationsordnung zu finden, vorausgesetzt genügend Rechenzeit ist vorhanden und die Länge der Skalierungsfunktion ist groß genug gewählt. Durch empirische Beobachtungen konnte eine weitere Vermutung formuliert und auch bewiesen werden. So ist die Approximationsordnung einer Multiskalenanalyse mit einer Skalierungsfunktion mit kompakten Träger festgelegt, falls eine Komponente der Skalierungsfunktion auf einem endlichen Intervall polynomiell ist (siehe Satz 4.9.1). Dieser Satz erklärt die Beobachtung, warum Skalierungsfunktionen aus stückweisen Polynomen immer eine gewisse Approximationsordnung besitzen, unabhängig davon ob die Gleichungen zum erzwingen der Approximationsordnung aktiv sind oder nicht. Für den Beweis ist der Raum Φ mit den vorher abgeleiteten Eigenschaften elementar.

Ein weiteres wichtiges Resultat ist Satz 4.10.5, der den Zusammenhang zwischen verschwindenden Momenten der Skalierungsfunktion und verschwindenden Momenten der Verfeinerungskoeffizientenfolge auf den vektoriellen Fall verallgemeinert. Während im skalaren Fall eine notwendige und hinreichende Beziehung gilt, ist dies bei vektoriellen Skalierungsfunktionen im Allgemeinen nicht so.

Die wichtigste Erkenntnis der Arbeit ist, dass für Multi-Coiflets ebenso eine minimale Länge für die Skalierungsfunktion existiert, wie für skalare Coiflets. Dies verkompliziert die Konstruktion von Multi-Coiflets, da eine Konstruktion mittels der "Intertwining" Theorie nicht durchführbar ist.

Als Ausblick bleibt die Untersuchung, ob Multi-Coiflets existieren, die zwar die hergeleitete Ungleichung für die Länge erfüllen, aber dennoch kürzer als

die skalaren Coiflets von Daubechies sind. Es ist dabei eventuell nicht sinnvoll, ausschließlich einen Ansatz mit einer stückweise Polynomiellen Skalierungsfunktion zu verfolgen, da die sich die ergebenden Koeffizientengleichungen nicht einfach geschlossen lösbar sind. Möglicherweise bringt eine Konstruktion bei der auch fraktaler Funktionen zugelassen sind mehr Erfolg.

Literaturverzeichnis

- [Alp92] ALPERT, BRADLEY K.: *Wavelets: a tutorial in theory and applications*, Kapitel Wavelets and other bases for fast numerical linear algebra, Seiten 181–216. Academic Press Professional, Inc., San Diego, CA, USA, 1992.
- [Alt06] ALT, HANS WILHELM: *Lineare Funktionalanalysis*. Springer-Verlag Berlin Heidelberg, 2006.
- [BCR91] BEYLKIN, G., R. R. COIFMAN und V. ROKHLIN: *Fast Wavelet Transforms and Numerical Algorithms*. Communications on Pure and Applied Mathematics, 44(2):141–183, 1991.
- [Bre78] BREWER, JOHN W.: *Kronecker products and matrix calculus in system theory*. IEEE Transactions on Circuits and Systems, pages 772–781, 1978.
- [Dau92] DAUBECHIES, INGRID: *Ten Lectures on Wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- [dBDR92] BOOR, CARL DE, RONALD A. DEVORE, and AMOS RON: *On the construction of multivariate (pre)wavelets*. Constructive Approximation, 1992.
- [DGH96] DONOVAN, GEORGE C., JEFFREY S. GERONIMO, and DOUGLAS P. HARDIN: *Intertwining multiresolution analyses and the construction of piecewise-polynomial wavelets*. SIAM Journal on Mathematical Analysis, 27(6):1791–1815, 1996.
- [DGH99] DONOVAN, G. C., J. S. GERONIMO, and D. P. HARDIN: *Orthogonal polynomials and the construction of piecewise polynomial smooth wavelets*. SIAM Journal on Mathematical Analysis, 30(5):1029–1056, 1999.
- [DGHM96] DONOVAN, GEORGE C., JEFFREY S. GERONIMO, DOUGLAS P. HARDIN, and PETER R. MASSOPUST: *Construction of orthogonal wavelets using fractal interpolation functions*. SIAM Journal on Mathematical Analysis, 27(4):1158–1192, 1996.

- [GHM94] GERONIMO, J. S., D. P. HARDIN, and P. R. MASSOPUST: *Fractal functions and wavelet expansions based on several scaling functions*. Journal of Approximation Theory, 78(3):373 – 401, 1994.
- [GvL96] GOLUB, G. H. and C. F. VAN LOAN: *Matrix Computations*. Johns Hopkins University Press, 3rd edition, 1996.
- [Jia91] JIA, RONG-QING: *A characterization of the approximation order of translation invariant spaces of functions*. Proceedings of the American Mathematical Society, 111(1):61–70, 1991.
- [Jia97] JIA, RONG-QING: *Shift-invariant spaces on the real line*. Proc. Amer. Math. Soc., 125:785–793, 1997.
- [JRZ96] JIA, RONGQING, S. D. RIEMENSCHNEIDER, and DINGXUAN ZHOU: *Approximation by multiple refinable functions*. Canadian J. Math, 49:944–962, 1996.
- [JS94] JIA, RONGQING and ZUOWEI SHEN: *Multiresolution and wavelets*. Proc. Edinburgh Math. Soc, 37:271–300, 1994.
- [LLS96] LAWTON, W., S. L. LEE, and ZUOWEI SHEN: *An algorithm for matrix extension and wavelet construction*. Mathematics of Computation, 65(214):723–737, 1996.
- [Mal89a] MALLAT, S.G.: *A theory for multiresolution signal decomposition: the wavelet representation*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 11(7):674–693, Jul 1989.
- [Mal89b] MALLAT, STEPHANE G.: *Multiresolution approximations and wavelet orthonormal bases of $l^2(\mathbb{R})$* . Transactions of the American Mathematical Society, 315(1):69–87, 1989.
- [Roa97] ROACH, DAVID W.: *Multiwavelet Prefilters: Orthogonal Prefilters Preserving Approximation Order $p \leq 3$* . PhD thesis, Vanderbilt University, 1997.
- [SS94] STRANG, GILBERT and VASILY STRELA: *Orthogonal multiwavelets with vanishing moments*. J. Optical Eng, 33:2104–2107, 1994.
- [SS95] STRANG, G. and V. STRELA: *Short wavelets and matrix dilation equations*. Signal Processing, IEEE Transactions on, 43(1):108–115, Jan 1995.
- [Vai93] VAIDYANATHAN, P. P.: *Multirate Systems And Filter Banks*. Prentice Hall, September 1993.
- [Wei99] WEISSTEIN, ERIC W.: *CRC Concise Encyclopedia of Mathematics*. CRC Press, Boca Raton, Fla. :, 1999.