

Neue und alte Thesen über Wavelets

Report

Author(s):

Buhmann, Martin Dietrich

Publication date:

1995-08

Permanent link:

<https://doi.org/10.3929/ethz-a-004284324>

Rights / license:

In Copyright - Non-Commercial Use Permitted

Originally published in:

SAM Research Report 1995-06

Neue und alte Thesen über Wavelets

M.D. Buhmann

Research Report No. 95-06
August 1995

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Neue und alte Thesen über Wavelets

M.D. Buhmann

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 95-06

August 1995

Abstract

In diesem Artikel soll eine erste, kurze Einführung in die Theorie der Wavelets gegeben werden, die ohne besondere Vorkenntnisse zugänglich ist. Anhand einiger Beispiele werden insbesondere die Zielsetzungen, die mit der Waveletanalyse verfolgt werden, erläutert, so beispielsweise die Kompression grosser Datenmengen oder die Suche nach gegebenen Mustern oder Singularitäten. Die besonderen Eigenschaften der Waveletanalyse, die sie von der Fourieranalyse unterscheiden, und sie für die erwähnten Beispiele prädestinieren, sind eines der Hauptthemen des Artikels.

1. Einführung

Der Zweck dieses Artikels ist, die Theorie der sogenannten Wavelets unter alten und neuen Gesichtspunkten zu beleuchten und dabei insbesondere eine neue Klasse solcher Wavelets vorzustellen. Wavelets ziehen derzeit ausserordentlich viel Interesse auf sich, so dass es in dem Rahmen dieser Arbeit nicht möglich ist, auch nur annähernd alle ihre Aspekte anzusprechen. Eine wesentliche Einschränkung besteht allein darin, dass wir uns nur mit eindimensionalen Wavelets befassen werden. Wir hoffen jedoch, hinreichend Einblick zu geben, damit der Leser die Gründe der Beliebtheit der Wavelets für Anwender und Analytiker sieht und sich vielleicht sogar selbst entschliesst, Wavelets näher zu studieren.

Wavelets dienen der Zerlegung, Bearbeitung und Analyse von Funktionen und Signalen. Letztere können z.B. Ton- oder Bildsignale sein, die gefiltert werden sollen, um Störungen, wie etwa ein Rauschen, zu beseitigen, oder die zerlegt, auf ihre wesentlichen Bestandteile reduziert und - dergestalt komprimiert - später oder an einem anderen Ort zusammengesetzt werden sollen. Insbesondere bei Bildern ist solch eine Kompression, d.h. Verdichtung, der Information, die das Bild vollständig oder mit guter Annäherung beschreibt, oft notwendig zum Zweck der schnellen Übertragung (Bildtelefon) oder Abspeicherung (speichereffiziente Aufzeichnung grosser Datenmengen). Wavelets sind die Bausteine, in die die Signale oder Funktionen zerlegt werden, und anhand derer dann entschieden wird, welche Teile der Zerlegung unerwünscht (Filterung) oder entbehrlich (Datenverdichtung) sind, deswegen fortgelassen werden sollen und bei der späteren Rekonstruktion unberücksichtigt bleiben. Offensichtlich ist es besonders wichtig, ein schnelles und stabiles numerisches Verfahren zu haben, um zu zerlegen und zu rekonstruieren, damit die dem Mathematiker zur Verfügung stehenden Wavelets in der Praxis gebraucht werden können.

Viele der obigen Aspekte weisen bereits auf eine andere, ältere Methode für die gleichen Anwendungen hin, nämlich die Fourierzerlegung. Wir wollen daher nun einen Schritt zurückgehen und die Fourierzerlegung (oder -transformation) beschreiben, die als Vorläufer der Waveletzerlegung betrachtet werden kann.

Sei dazu $f : \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion (ein eindimensionales Signal) endlicher Energie, d.h.

$$f \in L_2(\mathbb{R}) \iff \|f\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(t)|^2 dt < \infty .$$

Dann ist das Integral

$$(1) \quad \hat{f}(x) := \int_{-\infty}^{\infty} e^{-ixt} f(t) dt, \quad x \in \mathbb{R} ,$$

überall bis auf eine Menge mit verschwindendem Mass wohldefiniert, und \hat{f} hat auch endliche Energie (speziell: $\|f\|_2 = \frac{1}{\sqrt{2\pi}} \|\hat{f}\|_2$). Man nennt \hat{f} die Fouriertransformierte von f , und $\hat{f}(x)$ repräsentiert den Anteil der "Frequenz" x an der Funktion f . Die Bedeutung dieser Aussage wird unmittelbar klar, wenn man bedenkt, dass die Umkehrung

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixt} \hat{f}(x) dx, \quad t \in \mathbb{R} ,$$

gilt, die zeigt, dass f wieder aus allen seinen Frequenzbestandteilen zusammengesetzt werden kann. Die Bausteine hier sind also $\{e^{ixt}, x \in \mathbb{R} \mid t \in \mathbb{R}\}$, die Exponentialfunktionen zur Frequenz x . In der Anwendung benutzt man natürlich diskrete Frequenzen $\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots$, und es ist ein fundamentaler Satz der harmonischen Analysis, dass $x_j = j, j \in \mathbb{Z}$, genügen, um alle periodischen Funktionen endlicher Energie, die auf $[-\pi, \pi]$ definiert sind, eindeutig bis auf einer Menge mit Mass 0 darzustellen, und zwar mit Koeffizienten \hat{f}_j ebenfalls endlicher Gesamtenergie, d.h.

$$\sum_{j=-\infty}^{\infty} |\hat{f}_j|^2 < \infty .$$

Die Berechnung jener Koeffizienten erfolgt effizient mit der schnellen Fouriertransformation (FFT [21]), und anschliessend an diese Berechnung kann das Signal gefiltert werden (z.B. werden alle Koeffizienten hoher Frequenzen annulliert) oder komprimiert werden (z.B. werden alle Koeffizienten, die betragsmässig kleiner als eine bestimmte Fehler-toleranz sind, ignoriert). Die Wiederherstellung des Signals erfolgt im Wesentlichen auf die gleiche Art wie seine Zerlegung in die Koeffizienten, denn (1) und die Umkehrformel haben fast die gleiche Form.

Wie aus (1) ersichtlich, muss zur Berechnung von \hat{f} (oder auch \hat{f}_j) stets das gesamte Signal bzw. die Funktion über ihren Definitionsbereich bekannt sein, weswegen sich die Fouriertransformation speziell für stationäre oder periodische Signale eignet. Weil $|e^{ixt}| \equiv 1$, lokalisiert sie weiterhin nicht bezüglich der Zeit t , d.h. das Auftreten einer bestimmten Frequenz und seine Intensität können festgestellt werden, nicht aber die Zeit t , zu welcher sie im Signal f enthalten ist. Diese beiden Hauptnachteile der Fourierzerlegung, nämlich Anwendbarkeit nur auf stationäre Signale und fehlende Lokalität bezüglich der Zeit, will die Wavelettransformation beheben, und wir werden diese neue Transformation im nächsten Abschnitt beschreiben.

Diese Neuentwicklung ist ganz und gar nicht nur von theoretischem Interesse: Schon heute wird die Wavelettransformation, bzw. wird ihre Diskretisierung, die diskrete Wavelettransformation oder, analog der Fast Fourier Transform, “Fast Wavelet Transform” genannt, angewendet, um sehr grosse Karteien von Bilddaten zu komprimieren und elektronisch abzuspeichern, um Bilddaten ohne grossen Qualitätsverlust in Echtzeit über lange Distanzen zu übermitteln, um Singularitäten in Signalen zu finden und gegebenenfalls herauszufiltern oder um Muster in Bildern automatisch zu erkennen. Weitere Anwendungen sind Rekonstruktion historischer Photographien oder solcher, die aus wissenschaftlichen Experimenten unscharf oder sonst irgendwie beeinträchtigt hervorkommen. Allen diesen sehr praktischen Anwendungen ist gemeinsam, dass eine Filterungs- oder Kompressionsmethode gefragt ist, die sich verschiedenen Frequenzen zu verschiedenen Zeiten anpassen vermag, und auf diese Art berücksichtigen kann, dass es z.B. in einem Portrait an manchen Orten im Hintergrund über einen grossen Bereich kaum speicherungswürdige Information gibt, während es woanders viel Detail gibt, das durch zahlreiche Koeffizienten in der Entwicklung getreu repräsentiert werden muss. Speziell Kontouren in einem Bild verlangen genaue Wiedergabe, weil sich das menschliche Auge besonders anhand dieser orientiert. Obwohl die Fouriertransformation von vornherein weniger für alle diese Zwecke

geeignet ist, weil sie nicht lokal ist, ist es vor allem die Fast Wavelet Transform, die durch einfache, diskrete Faltungen der Koeffizientenvektoren einer ersten Approximation des Signals die Waveletzerlegung zu berechnen erlaubt, die den Wavelets den entscheidenden Vorteil in der Praxis gibt.

2. Die Wavelettransformation

Die ‘‘Bausteine’’, die der Wavelettransformation zu Grunde liegen, sind oszillierende, quadratintegrierbare Funktionen (also *kleine Wellen*), die lokal sind und durch Verschiebung und Skalierung verschiedene Frequenzen **zeitabhangig** abtasten konnen. Eine typische Waveletfunktion ist das Haar-Wavelet, so benannt nach dem Haarschen Mass, aus dem es hergeleitet wurde

$$\psi(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}), \\ -1, & x \in [\frac{1}{2}, 1), \\ 0, & \text{sonst,} \end{cases} \quad x \in \mathbb{R} .$$

Die Funktion ψ hat verschwindendes Integral uber \mathbb{R} , oszilliert also. Sie ist lokal, weil sie sogar kompakten Trager hat. Die Wavelettransformation ist fur dieses oder ein allgemeines, quadratintegrierbares ψ und $f \in L^2(\mathbb{R})$

$$W f(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \bar{\psi}\left(\frac{t-b}{a}\right) dt, \quad b \in \mathbb{R}, \quad a > 0 .$$

Im Gegensatz zur Fouriertransformation treten hier also zwei Parameter auf. Durch a wird ψ skaliert, oszilliert also schneller (wenn $a \ll 1$) oder langsamer (wenn $a \gg 1$), und mittels b wird es gleichzeitig entlang der reellen Achse verschoben. Im Vergleich zur Fouriertransformation ersetzt ψ die Exponentialfunktion, und es ist nun das Ziel, f auf ahnliche Weise zu rekonstruieren aus seiner Wavelettransformation und einen der FFT verwandten Algorithmus zur Verfugung zu stellen.

Beide Ziele konnen leichter erreicht werden, indem wir fordern, dass $\{2^{k/2}\psi(2^k \cdot -j) \mid k, j \in \mathbb{Z}\}$ eine orthonormale Basis von $L^2(\mathbb{R})$ bilden, denn dann gilt

$$(2) \quad f = \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} 2^{k/2} W f(2^{-k}, 2^{-k}j) \psi(2^k \cdot -j)$$

fur alle $f \in L^2(\mathbb{R})$. Es sind also die diskreten Auswertungen $W f(2^{-k}, 2^{-k}j)$ der Wavelettransformation die Koeffizienten der Entwicklung. Die Darstellung (2) erfullt einerseits unser Anliegen, f in verschiedene Frequenzbestandteile zu zerlegen (so geschehen durch die Skalierung mit 2^k), und zwar lokal bezuglich der Zeitachse. Das wird garantiert durch die Verschiebung mit j und die Lokalitat von ψ , die fur das Haar-Wavelet besonders eklatant ist, aber auch fur die spater vorzustellenden Wavelets gilt. Andererseits genugt sie der Forderung nach einem schnellen Algorithmus; denn die $W f(2^{-k}, 2^{-k}j)$ konnen

schnell, rekursiv mit der Fast Wavelet Transform berechnet werden. Es ist aus diesen Gründen üblich, ψ nur dann ein Wavelet zu nennen, wenn $\{2^{k/2}\psi(2^k \cdot -j) \mid k, j \in \mathbb{Z}\}$ eine orthonormale Basis bilden. Wir wollen uns an diese Konvention halten. Warum verlangen wir Orthogonalität? Die Orthogonalität

$$\int_{-\infty}^{\infty} \psi(2^k t - j) \overline{\psi}(2^{k'} t - j') dt = 0, \quad \forall k \neq k', j \neq j',$$

erinnert an die Orthogonalität der Exponentialfunktionen

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ijt} e^{-ij't} dt = 0, \quad \forall j \neq j',$$

und erfüllt in der Tat ähnliche Aufgaben wie bei der Fourieranalyse periodischer Funktionen.

Sie ist nämlich wesentlich für das effiziente Funktionieren der FFT, und das gleiche gilt für die FWT, die Fast Wavelet Transform, die aus Platzgründen hier nicht näher beschrieben wird. Anstatt einer ausführlichen Besprechung in diesem Artikel wird auf die hervorragende Beschreibung in [9] verwiesen und auf die kurze Darstellung eines Spezialfalles der FWT in Abschnitt 3. Aber wir gehen näher auf eine andere Aufgabe der Orthogonalität ein. Sie garantiert, dass die Zerlegung in (2) stabil ist, d.h. dass die Grösse eines Koeffizienten der Zerlegung an der Stelle $j2^{-k}$ bei "Frequenz" 2^k tatsächlich das Verhalten von f an dieser Stelle quantitativ darstellt. Insbesondere ist die Quadratsumme der Koeffizienten gleich einem konstanten Vielfachen der Energie zum Quadrat von f . Weil Orthogonalität nur bezüglich k für ein Funktionieren der FWT ausreicht und die eben besprochene Stabilität auch erfüllt ist, falls für alle $c \in \ell^2(\mathbb{Z})$, $c = (c_j)_{j=-\infty}^{\infty}$, die **Rieszsche Bedingung**

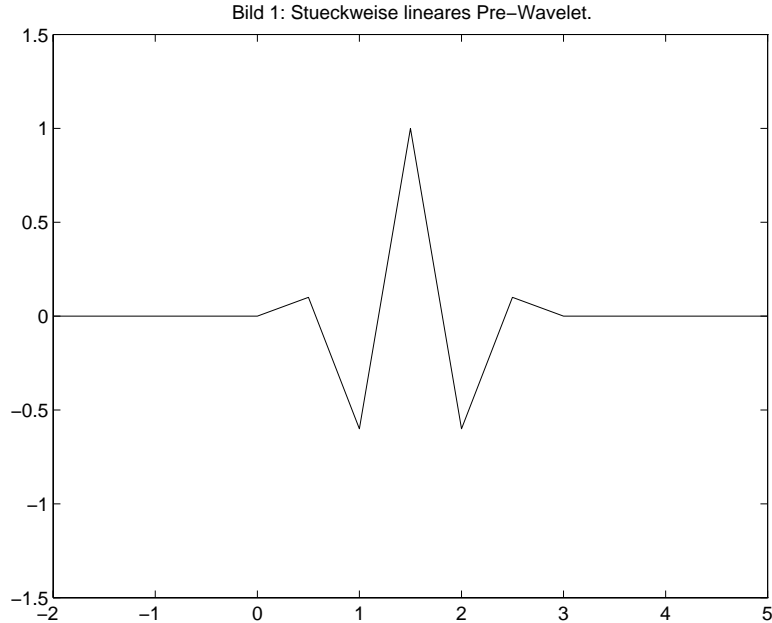
$$(3) \quad \lambda \|c\|_2 \leq \left\| \sum_{j=-\infty}^{\infty} c_j \psi(2^k \cdot -j) \right\|_2 \leq \Lambda \|c\|_2,$$

gilt, wobei $0 < \lambda \leq \Lambda < \infty$ unabhängig von k und c sind, lässt man die Bedingung $\forall j \neq j'$ in der Orthogonalitätsbedingung oft aus, fordert stattdessen (3) und dass $L^2(\mathbb{R})$ von den $\psi(2^k \cdot -j)$ erzeugt wird. Man spricht dann von Prä-Wavelets. Hierbei bezeichnet $\|\cdot\|_2$ sowohl die Norm auf $\ell^2(\mathbb{Z})$ als auch auf $L^2(\mathbb{R})$, je nach Argument.

Wir werden uns in diesem Artikel von nun an nur noch mit Prä-Wavelets befassen. Es ist anzumerken, dass die Darstellung (2) andere Koeffizienten hat, wenn ψ nur ein Prä-Wavelet ist, was aber Nutzen und Berechenbarkeit der Darstellung keinen Abbruch tut. Es soll ferner hervorgehoben werden, dass in der Praxis nur ein endlicher Teil von (2) berechnet wird.

Das Haar-Wavelet ist das einfachste Beispiel sogenannter Spline (Prä-) Wavelets, die aus stückweisen Polynomen bestehen und eine besonders zweckmässige Klasse von (Prä-) Wavelets sind. Wir werden weitere Beispiele geben. Warum sind Spline Prä-Wavelets nützlich? Eine Zerlegung (2) von glatten Funktionen ist umso geeigneter, sagen wir für

die Kompression einer Funktion, desto weniger Koeffizienten gebraucht werden, um sie in einer endlichen Summe nahezu getreu darzustellen. Es gibt aber eine reiche Theorie der Splines ([19] z.B.), die zeigt, dass glatte Funktionen in der Tat effizient durch Splinefunktionen approximiert werden können. Daher sind diese Räume besonders für Prä-Wavelets geeignet. Ein weiterer Grund ist, dass Spline Prä-Wavelets einfache Darstellungen haben, die auch leicht in numerischen Verfahren implementiert werden können. Ein stückweise lineares Prä-Wavelet sieht so aus:



Spline Prä-Wavelets sind mehrfach stetig differenzierbar, aber nicht in $C^\infty(\mathbb{R})$. Es wird bei Splines gefordert, dass sie in $C^{k-1}(\mathbb{R})$ liegen, wenn sie stückweise Polynome vom Grad k sind. Es existieren Spline Prä-Wavelets von beliebigem Grad mit kompaktem Träger. Sie können mit einer expliziten Formel angegeben werden [9]. Viele andere wichtige Klassen von (Prä-) Wavelets existieren, siehe [8,9,11-16] und insbesondere die Daubechies Wavelets [11]. Im nächsten Abschnitt zeigen wir, wie Spline Prä-Wavelets regularisiert werden können, damit sie C^∞ sind.

Wir stellen hier nur noch die Daubechies Wavelets vor, die besonders oft verwendet und in der Literatur besprochen werden. Ihre Konstruktion basiert auf der Existenz einer Funktion $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ mit kompaktem Träger, die die sogenannte Verfeinerungsgleichung

$$\varphi(x) = \sqrt{2} \sum_{k=0}^N h_k \varphi(2x - k), \quad x \in \mathbb{R},$$

für geeignete reelle Koeffizienten h_k erfüllt und quadratintegrierbar ist. Wenn die ganzzahligen Translate von φ nämlich ausserdem orthonormal sind, ist

$$\psi(x) = \sqrt{2} \sum_{k=1-N}^1 (-1)^k h_{1-k} \varphi(2x - k), \quad x \in \mathbb{R},$$

ein Wavelet mit kompaktem Träger [11]. Für das Daubechies Wavelet beginnt man mit der Wahl eines trigonometrischen Polynoms $q(x) = \frac{1}{2} + \sum_{k=1}^N \alpha_k \cos(2k-1)x$ und eines anderen $H(x) = \sum_{k=0}^N h_k \exp(ikx)$ mit reellen Koeffizienten derart, dass $q = |H|^2$. Dann muss eine Lösung der Verfeinerungsgleichung berechnet werden. Das Resultat ist ein orthogonales Wavelet ψ mit kompaktem Träger, das Daubechies Wavelet D_N . Die Funktion φ kann beispielsweise dadurch berechnet werden, dass man seine Fouriertransformierte durch das unendliche Produkt

$$\hat{\varphi}(x) = \prod_{j=1}^{\infty} H(2^{-j}x), \quad x \in \mathbb{R},$$

bestimmt. Bei der Wahl von H muss natürlich besonderes Augenmerk auf die Konvergenz dieses Produkts gerichtet werden.

3. Neue Klassen von Prä-Wavelets

In diesem Abschnitt werden wir eine neue Klasse von Prä-Wavelets vorstellen, die zu effizienter Zerlegung von Funktionen geeignet sind. Im einfachsten Fall sind sie eine Verallgemeinerung der im letzten Abschnitt beschriebenen linearen Prä-Wavelets. Jene können nämlich als Linearkombinationen der stückweise linearen Funktionen $|\cdot - \frac{1}{2}j|$, $j \in \mathbb{Z}$, dargestellt werden, und wir befassen uns nun mit Linearkombinationen von

$$(4) \quad \sqrt{(\cdot - \tau j)^2 + c^2}, \quad j \in \mathbb{Z},$$

wobei $\tau \in (0, 1]$ die Verschiebungslänge, z.B. $\tau = \frac{1}{2}$, und c ein fester nichtnegativer Parameter ist. Die Funktion (4) ist für $c = 0$ wieder obige Betragsfunktion, die bei $\frac{1}{2}j$ keine stetige Ableitung besitzt, während sie für $c > 0$ in $C^\infty(\mathbb{R})$ ist. Daher rührt das Interesse an (4), denn glatte Funktionen können gut mit glatten Wavelets zerlegt werden, wobei trotzdem auch für weniger glatte oder singuläre Funktionen durch die Skalierung, also lokale detailreiche Approximation, Sorge getragen wird. Potenzen von (4) entsprechen übrigens Splines höherer Ordnung, so ist $((\cdot - \tau j)^2 + c^2)^{3/2}$ etwa eine Regularisierung von $|\cdot - \tau j|^3$. Die Funktionen (4) heissen ‘‘Multiquadric-Funktionen’’, und ihre den Splines ähnlichen Approximationseigenschaften in einer Dimension werden z.B. in [3] untersucht. Es gibt Verallgemeinerungen auf höhere Dimensionen, die besonders interessante Approximationseigenschaften haben [4], deren Besprechung allerdings den Rahmen unserer Ausführungen hier leider sprengen würde. Die Grössen τj , um die die Funktionen verschoben sind, werden üblicherweise als Zentren oder Stützstellen bezeichnet.

Das folgende Resultat gibt Auskunft über die Existenz und die Eigenschaften von Multiquadric Prä-Wavelets. Dazu ist es hilfreich, einige vereinfachende Notationen einzuführen. Wir schreiben von nun an

$$\phi(r) = \sqrt{r^2 + c^2} \quad \text{und} \quad \Phi(r) = \phi'' * \phi''(r),$$

wobei

$$f * g(r) = \int_{-\infty}^{\infty} f(r-t)g(t) dt$$

Faltung bedeutet (für geeignete f, g , beispielsweise in $L^2(\mathbb{R})$); beachte hierbei, dass

$$\phi''(r) = \frac{c^2}{(\sqrt{r^2 + c^2})^3}$$

in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ ist und dass diese kubische Abklingrate von $\phi''(r)$ für $r \rightarrow \pm\infty$ durch Faltung mit sich selbst erhalten bleibt, d.h. $|\Phi(r)| = O(|r|^{-3})$ für $r \rightarrow \pm\infty$.

Weiterhin sei $\Psi_\ell(x) = \phi[-\ell\tau, 0, \ell\tau](x)$ die zweite dividierte Differenz von $\phi(x - \cdot)$ an den Punkten $-\ell\tau, 0, \ell\tau$, wobei $\ell \in \mathbb{R}$ und $\ell > 0$. Man beachte dabei, dass $\Psi_\ell(x)$ genauso schnell wie $\phi''(x)$ abklingt. Das folgende Resultat ist nicht nur für die Schrittweite $\tau = \frac{1}{2}$ wie oben, sondern für allgemeine Schrittweiten formuliert.

Satz 1 [4,5]: Sei $0 < \tau \leq 1$ die Schrittweite äquidistanter Punkte $\{j\tau \mid j \in \mathbb{Z}\}$. Seien τ und c derart, dass die Bedingung

$$\max \left[\frac{2}{\tau}, \sup_{t \in [0, \tau]} \sum_{k=-\infty}^{\infty} \phi''(t - k\tau) \right] < \inf_{\vartheta \in (0, 4\tau)} \Phi(\vartheta)$$

erfüllt ist. Dann gibt es $d = (d_j)_{j=-\infty}^{\infty} \in \ell^2(\mathbb{Z})$, so dass

$$(5) \quad \psi(x) = \sum_{j=-\infty}^{\infty} d_j \Psi_1(x - j\tau), \quad x \in \mathbb{R},$$

die Abschätzung $|\psi(x)| \leq C(1 + |x|)^{-3}$ für ein von x unabhängiges C und die Orthogonalitätsbedingung

$$(6) \quad \int_{-\infty}^{\infty} \psi(x) \Psi_2(x - 2\tau j) dx = 0 \quad \forall j \in \mathbb{Z}$$

erfüllt. Ausserdem gibt es für alle $e = (e_j)_{j=-\infty}^{\infty} \in \ell^2(\mathbb{Z})$ Koeffizientenfolgen $a = (a_j)_{j=-\infty}^{\infty}$ und $b = (b_j)_{j=-\infty}^{\infty}$ in $\ell^2(\mathbb{Z})$, so dass die folgende Beziehung gilt:

$$(7) \quad \sum_{j=-\infty}^{\infty} e_j \Psi_1(x - j\tau) = \sum_{j=-\infty}^{\infty} a_j \psi(x - 2j\tau) + \sum_{j=-\infty}^{\infty} b_j \Psi_2(x - 2j\tau).$$

□

Wir sollten uns erst einmal davon überzeugen, dass die Bedingung des Satzes für geeignet kleine τ und c erfüllt werden kann. Dazu bemerken wir, dass ihre linke Seite asymptotisch für $\tau \rightarrow 0$ durch $\frac{2}{\tau}$ repräsentiert ist, weil die Summe links eine Riemann-Summe ist und $\int_{-\infty}^{\infty} \phi''(t) dt = 2$. Die rechte Seite ist für $\tau \rightarrow 0$ im Grenzwert

$$\int_{-\infty}^{\infty} \phi''(t)^2 dt = \int_{-\infty}^{\infty} \frac{c^4}{(t^2 + c^2)^3} dt = \frac{1}{c} \int_{-\infty}^{\infty} \frac{dt}{(1 + t^2)^3}.$$

Für ein zuerst gewähltes kleines τ brauchen wir also nun $c > 0$ hinreichend klein zu wählen, um die Bedingung zu erfüllen.

In diesem Satz werden also Linearkombinationen von Multiquadric Funktionen mit (Translations-) Schrittweite τ gebildet, die zu jenen mit doppelter Schrittweite orthogonal sind und nach (7) die Differenz der beiden, von den unterschiedlich translatierten Multiquadric Funktionen aufgespannten Räumen erzeugen. Wir bezeichnen jene beiden linearen Räume mit V_0 bzw. V_1 und die Differenz mit W_0

$$(8) \quad \begin{aligned} V_j &= \left\{ \sum_{k=-\infty}^{\infty} c_k \Psi_{2^{1-j}}(\cdot - 2^{1-j} k \tau) \mid \sum_{k=-\infty}^{\infty} c_k^2 < \infty \right\}, \\ W_0 &= \left\{ \sum_{k=-\infty}^{\infty} c_k \psi(\cdot - 2k\tau) \mid \sum_{k=-\infty}^{\infty} c_k^2 < \infty \right\}. \end{aligned}$$

Es gilt also gemäss Satz 1 $\psi \in V_1$, $\psi \perp V_0$, $W_0 \perp V_0$ und

$$(9) \quad V_1 = V_0 + W_0.$$

Beginnend mit einem c und einem τ derart, dass die Bedingung des Satzes für τ , $\frac{1}{2}\tau, \dots, 2^{1-k}\tau$ erfüllt ist, kann man diese Zerlegung rekursiv zu

$$(10) \quad V_k = W_{k-1} + W_{k-2} + \dots + W_0 + V_0$$

erweitern, wobei V_k und V_0 durch (8) definiert sind und die $W_0, W_1, \dots, W_{k-2}, W_{k-1}$ die entsprechenden orthogonalen Differenzräume sind. In der Darstellung (10) sind die W_j 's genau die Räume, die von den in Satz 1 beschriebenen Prä-Wavelets (stabil, s. die Bemerkungen in [4] und [5]) erzeugt werden und somit einen endlichen Teil der gewünschten Zerlegung in Wavelet-Komponenten geben, wobei V_0 ein Restglied (das sogenannte "blurred signal") enthält und V_k der Raum mit der kleinsten Schrittweite ist, aus dem wir eine zu zerlegende Funktion $f \in L^2(\mathbb{R})$ zunächst approximieren, bevor wir mit der Dekomposition beginnen. Dass diese Vorgehensweise sinnvoll ist, beweisen die folgenden zwei Resultate. Sie zeigen nämlich, dass die Vereinigung aller V_j dicht in $L^2(\mathbb{R})$ ist, und so kann man für ein geeignetes k jedes $f \in L^2(\mathbb{R})$ erst durch ein $\tilde{g}_k \in V_k$ approximieren und dann die Zerlegung (10) anwenden. Der folgende Satz ist für eine allgemeine Funktion Ψ , deren Dilatationen die V_j erzeugen, und nicht nur für obige Ψ_ℓ formuliert.

Ausserdem werden für die Räume V_j auch Indices $j < 0$ in Satz 2 zugelassen, obwohl wir später keinen Gebrauch davon machen werden. Um zu demonstrieren, wie bei derartigen Aussagen vorgegangen werden kann, geben wir den Beweis an.

Satz 2: Sei $\Psi : \mathbb{R} \rightarrow \mathbb{R}$ derart, dass für eine fixe Konstante K gilt $|\Psi(x)| \leq K(1+|x|)^{-1-\epsilon}$, wobei $\epsilon > 0$ fest ist. Gilt ausserdem $\sum_{j=-\infty}^{\infty} \Psi(x - \tau j) \equiv 1$, so erfüllen die Räume (8) (wobei wir nur für diesen Satz $\Psi_{2^{1-j}} := \Psi(2^{j-1}\cdot)$ wählen) für jedes $J \in \mathbb{Z}$

$$(11) \quad \overline{\bigcup_{j=J}^{\infty} V_j} = L^2(\mathbb{R}).$$

Beweis: Ohne Einschränkung der Allgemeinheit nehmen wir $\tau = 1$ an. Wir zeigen zunächst, dass eine stetige Funktion mit stetiger Ableitung und kompaktem Träger für $j \rightarrow \infty$ beliebig nahe durch eine Summe

$$(12) \quad S_j(x) = \sum_{k=-\infty}^{\infty} f(2^{1-j}k) \Psi(2^{j-1}x - k), \quad x \in \mathbb{R},$$

approximiert werden kann. Diese Form der Approximation wird oft als Quasi-Interpolation bezeichnet. Es gilt für diese Quasi-Interpolation

$$\begin{aligned} |S_j(x) - f(x)| &= \left| \sum_{k=-\infty}^{\infty} (f(2^{1-j}k) - f(x)) \Psi(2^{j-1}x - k) \right| \\ &\leq K_0 \sum_{|x-2^{1-j}k| \leq 2^{1-j}} |f(x) - f(2^{1-j}k)| \\ &\quad + K_1 \sum_{2^{1-j} \leq |x-2^{1-j}k| \leq 1} |x - 2^{1-j}k| |\Psi(2^{j-1}x - k)| \\ &\quad + K_2 \sum_{|x-2^{1-j}k| \geq 1} |\Psi(2^{j-1}x - k)| \\ &\leq o(1) K_0 + K K_1 \sum_{2^{1-j} \leq |x-2^{1-j}k| \leq 1} |x - 2^{1-j}k|^{-\epsilon} 2^{(1-j)(1+\epsilon)} \\ &\quad + K K_2 \sum_{|x-2^{1-j}k| \geq 1} |2^{j-1}x - k|^{-1-\epsilon} \\ &\leq o(1) K_0 + K_3 2^{(1-j)\epsilon} + K_4 2^{1-j} + K_5 2^{(1-j)\epsilon} = o(1), \quad j \rightarrow \infty, \end{aligned}$$

wobei K_0, K_1, \dots, K_5 feste, von x unabhängige Konstanten sind.

Wir müssen nun noch zeigen, dass $S_j \in L^2(\mathbb{R})$ und dass $\|S_j\|_2$ von einer Konstanten, die von $j \in \mathbb{Z}$ unabhängig ist, beschränkt werden kann. Ist dieses gezeigt, so folgt die Aussage des Satzes aus der Tatsache, dass stetig differenzierbare Funktionen mit kompaktem Träger dicht in $L^2(\mathbb{R})$ liegen. Sei $\text{supp } f \subset [-M, M]$. Dann gelten die Abschätzungen

$$\begin{aligned} \|S_j\|_2^2 &\leq \int_{-2M}^{2M} |S_j(x)|^2 dx + \int_{|x|>2M} |S_j(x)|^2 dx \\ &\leq 4M \sup_{x \in [-M, M]} |f(x)|^2 \left[\sup_{x \in [0, 1]} \sum_{k=-\infty}^{\infty} |\Psi(x - k)| \right]^2 \\ &\quad + K^2 \sup_{x \in [-M, M]} |f(x)|^2 \int_{|x|>2M} \left(\sum_{|k| < 2^{j-1}M} (1 + |2^{j-1}x - k|)^{-1-\epsilon} \right)^2 dx \\ &\leq K_6 + K_7 \sup_{x \in [-M, M]} |f(x)|^2 \int_{|x|>2M} |x|^{-2-2\epsilon} dx \\ &\leq K_8, \end{aligned}$$

wobei K_6, K_7 und K_8 feste, von j unabhängige Konstanten sind. □

Satz 3 [5,18]: Die vor Satz 1 definierten Funktionen erfüllen $\sum_{k=-\infty}^{\infty} \Psi_{2^{1-j}}(x - 2^{1-j}\tau k) \equiv 1$, für alle $j \in \mathbb{Z}$ und $\tau \in (0, 1]$. □

Weil $|\Psi_{2^{1-j}}(x)| = O(|x|^{-3})$, $|x| \rightarrow \infty$, folgt aus Satz 3, dass die Voraussetzungen des Satzes 2 erfüllt sind. Es können somit, für gegebenes $f \in L^2(\mathbb{R})$ und $\epsilon > 0$, k und $\tilde{g}_k \in V_k$ mit $\|f - \tilde{g}_k\|_2 < \epsilon$ gefunden werden. Dieses \tilde{g}_k kann beispielsweise von der Form (12) sein: $\tilde{g}_k = S_k$ mit $\Psi := \Psi_{2^{1-k}}(2^{1-k} \cdot)$. Anschliessend kann \tilde{g}_k in

$$\tilde{g}_k = g_{k-1} + g_{k-2} + \dots + g_0 + \tilde{g}_0$$

zerlegt werden, wobei $g_j \in W_j$ und $\tilde{g}_0 \in V_0$, gemäss (10). Diese Zerlegung findet schrittweise statt, d.h. man beginnt

$$\tilde{g}_k = g_{k-1} + \tilde{g}_{k-1}, \quad \tilde{g}_{k-1} \in V_{k-1},$$

und setzt

$$\tilde{g}_k = g_{k-1} + g_{k-2} + \tilde{g}_{k-2}, \quad \tilde{g}_{k-2} \in V_{k-2},$$

fort, usw. Man muss für unsere Prä-Wavelets fragen, wie in jedem Schritt g_j und \tilde{g}_j bzw. ihre Koeffizienten berechnet werden (damit sind die Koeffizienten bezüglich der Prä-Wavelets, die W_j erzeugen, bzw. die bezüglich der Translate von $\Psi_{2^{1-j}}$ gemeint). Die Antwort gibt (7), denn gemäss dieser Identität erfährt man, wie in jedem Schritt von V_j zu W_{j-1} und V_{j-1} zu gelangen ist. Da insbesondere aus (7) ersichtlich ist, wie die Erzeugenden von V_j zu zerlegen sind, handelt es sich in jedem Zerlegungsschritt nur um eine Multiplikation von Koeffizienten in V_j mit geeigneten (bi-infiniten) Matrizen, um zu den Koeffizientenvektoren in V_{j-1} und W_{j-1} zu kommen. Im Wesentlichen ist damit die FWT beschrieben. Es lässt sich übrigens zeigen, dass die Einträge der genannten Matrizen kubisch abklingen und daher die Fehler, die sich durch Abschneiden der unendlichen Summen ergeben, kontrollierbar bleiben.

Prä-Wavelets, die beispielsweise $f \in L^2([0, 1])$, also ein auf einem kompakten Intervall definiertes f , zu zerlegen vermögen, sind auch von Interesse. Wir werden einen möglichen Ansatz nur kurz beschreiben; er beruht auf der Tatsache, dass mittels Multiquadric-Funktionen auch stetige Funktionen auf $[0, 1]$ beliebig nahe approximiert werden, wobei nur Zentren der Multiquadric-Funktionen innerhalb oder nahe bei $[0, 1]$ verwendet werden. Zuerst wollen wir einen weiteren Satz über Approximation mit Multiquadric-Funktionen angeben. Das folgende Resultat ist als eine Präzisierung von Satz 2 für Multiquadric Approximationen zu verstehen, weil es angibt, mit welcher Präzision dreimal stetig differenzierbare Funktionen mit beschränkter Ableitung aus den Räumen V_j durch Quasi-Interpolation angenähert werden können. Da solche Funktionen dicht in $L^2(\mathbb{R})$ liegen, bedeutet dieses insbesondere, dass $\overline{\bigcup_{j=J}^{\infty} V_j} = L^2(\mathbb{R})$ für alle $J \in \mathbb{Z}$.

Der Satz ist von zusätzlichem Interesse, weil die Qualität der anfänglichen Approximation $\tilde{g}_k \approx f$ offensichtlich von besonderer Bedeutung ist. Sie ist umso besser, je kleiner wir k

wählen können, um für f eine Approximation in V_k mit vorgegebener Toleranz zu finden. Das bedeutet auch, dass wir mit weniger Koeffizienten bei der Approximation von f auskommen.

Satz 4 [3]: Sei $f \in C^3(\mathbb{R})$ und habe f beschränkte Ableitungen. Definiere

$$S_k(x) = \sum_{j=-\infty}^{\infty} f(2^{1-k}\tau j) \Psi_{2^{1-k}}(x - 2^{1-k}\tau j), \quad x \in \mathbb{R}.$$

Die Summe ist wohldefiniert, da $\Psi_{2^{1-k}}$ kubisch abklingt. Dann gibt es eine von k , c und τ unabhängige Konstante K , so dass

$$\sup_{x \in \mathbb{R}} |f(x) - S_k(x)| \leq K (4^{1-k}\tau^2 + c^2 \min[|\log c|, |\log \tau| + k]). \quad \square$$

Mit anderen Worten kann ein beliebiges $f \in C^3(\mathbb{R})$ mit beschränkten Ableitungen beliebig nahe gleichmässig durch Quasi-Interpolation approximiert werden, wenn k gross und c klein genug ist (für festes τ). Es gibt nun ein ähnliches Resultat für $f : [0, 1] \rightarrow \mathbb{R}$:

Satz 5 [1]: Sei $f : [0, 1] \rightarrow \mathbb{R}$ stetig und ω_f sein Stetigkeitsmodul. Definiere die Quasi-Interpolante

$$(13) \quad S_k(x) = f(0) \sum_{j=-\infty}^{-1} \Psi_{2^{1-k}}(x - 2^{1-k}\tau j) + \sum_{j=0}^{[2^k-1/\tau]} f(2^{1-k}\tau j) \Psi_{2^{1-k}}(x - 2^{1-k}\tau j) + f(1) \sum_{j=[2^k-1/\tau]+1}^{\infty} \Psi_{2^{1-k}}(x - 2^{1-k}\tau j).$$

Dabei ist $[\cdot]$ die Gauss-Klammer.

Dann gilt $\sup_{x \in [0,1]} |f(x) - S_k(x)| \leq (1 + 2^{k-1}c/\tau) \omega_f(\tau \cdot 2^{1-k})$ und, wenn f sogar stetig differenzierbar ist, $\omega_f(\tau \cdot 2^{1-k}) \leq \sup |f'(x)| \tau \cdot 2^{1-k}$. \square

Somit kann man $f \in L^2([0, 1])$ durch $\tilde{g}_k \in V_k$ mit genügend grossem k (wobei, wie schon zuvor, c auch klein genug gewählt werden muss) hinreichend nahe approximieren, und zwar auf die im vorhergehenden Satz gegebene Art. Dann kann f wie vorher zerlegt werden, denn die Basisfunktionen der genannten Approximation sind die gleichen wie oben, nur ist (13) eine effizientere Art, für ein $f \in L^2([0, 1])$ die anfängliche Annäherung nur unter Verwendung von Zentren in $[0,1]$ vorzunehmen. Hierbei sind zwei Bemerkungen angebracht. Erstens sichert die in Satz 1 angegebene Abklingeigenschaft der Prä-Wavelets, dass die Zerlegung weiterhin im Interval $[0,1]$ konzentriert bleibt. Zweitens ist die Näherung (13) für $x \in [0, 1]$ **unabhängig** von den Stützstellen ausserhalb von $[0,1]$. Mehr sogar: man kann die hier und zuvor (insbesondere in Satz 1) verwendeten dividierten Differenzen auf allgemeinen, nicht notwendig äquidistanten Stützstellen basieren lassen und erhält dann Ψ 's und Prä-Wavelets auf nicht gleichmässig verteilten Stützstellen. Dabei wird die Rolle von τ durch die maximale bzw. minimale Schrittweite der neuen, beliebig verteilten

Stützstellen übernommen. Auf dieser Idee basieren die Artikel [8,5], bei denen Spline Prä-Wavelets und Multiquadric Prä-Wavelets auf nicht äquidistanten Stützstellen eingeführt und analysiert wurden. Diese zu besprechen würde im Rahmen dieser Arbeit zu weit führen, aber, um noch einmal auf (13) und seine Zerlegung zurückzukommen, was zählt, ist f 's Verhalten an den Stützstellen in $[0, 1]$.

Es gibt viele andere sinnvolle Verallgemeinerungen dieser Prä-Wavelets, genannt seien die aus [7], welche Orthogonalität bezüglich inneren Produkten vom Sobolev-Typ aufweisen, d.h. im einfachsten Falle

$$(f, g) = \int_{-\infty}^{\infty} f(x) g(x) dx + \lambda \int_{-\infty}^{\infty} f'(x) g'(x) dx, \quad \lambda > 0,$$

wobei f, g, f' und g' quadratintegrierbar sein sollen. Diese sind nützlich, um nicht-lineare hyperbolische Differentialgleichungen mittels Spektralmethoden numerisch zu lösen, denn die entstehenden Steifigkeitsmatrizen aus dem Galerkin-Ansatz sind bei einer dem Differentialoperator korrespondierenden Wahl des inneren Produkts trivial [7].

Es soll noch bemerkt sein, dass für die Berechnung der Zerlegung einer quadratintegrierbaren Funktion die Form des Prä-Wavelets eigentlich selbst keine Rolle spielt. Es sind vielmehr die Koeffizienten aus (7) wichtig. Dennoch können beispielsweise die Prä-Wavelets aus Satz 1 leicht direkt berechnet werden. Sie sind von der Form

$$\sum_{k=-\infty}^{\infty} d_k^{2\ell} \Psi_k^*(x), \quad x \in \mathbb{R},$$

wobei $\Psi_k^* = \phi[-2\tau, -\tau, 0, \tau, 2\tau]$ eine vierte dividierte Differenz ist und $(d_k^\ell)_{k,\ell=-\infty}^{\infty}$ die Einträge der Inversen der Matrix

$$D = (\phi * \Psi_k^*(\tau\ell))_{k,\ell=-\infty}^{\infty}$$

sind.

4. Weitere Beispiele, offene Fragen und abschliessende Bemerkungen

Viele neue Anwendungen von Wavelets und Prä-Wavelets sind denkbar und viele werden fast täglich entdeckt. Zu denken ist beispielsweise an medizinische Anwendungen in der Tomographie, wo auf Röntgen- oder Kernresonanzbildern Singularitäten, die in der Form von Tumoren u.ä. auftreten, identifiziert werden sollen. In einer technischen Anwendung ist es notwendig, beim angestrebten automatisierten Ablesen der Postleitzahlen von Paketen, Störungen wie z.B. Paketschnüre, die das Adressfeld überqueren, aus dem Bild auszufiltern. Alle Anwendungen, die bereits in der Einleitung erwähnt wurden, lassen Erweiterungen zu. Beispielsweise kann die Datenkompression auch auf die Verdichtung der Information in einer grossen Steifigkeitsmatrix, die bei der numerischen Lösung einer partiellen Differentialgleichung auftritt, bezogen werden. Filterungsaufgaben treten bei Anwendungen von der Übertragung analog gespeicherter Musikaufnahmen bis zur

Verbesserung der Bildqualität der Übertragungen des Hubbleteleskops oder anderer astronomischer, etwa durch atmosphärische Störungen verschlechterter Bilder auf. Die numerische Lösung partieller Differentialgleichungen bietet ein weites Feld der Anwendung für Wavelets, da die Dilatationen und Translate wegen ihrer Adaptationsfähigkeit gut als Erzeugende der der numerischen Lösung unterliegenden Räume geeignet sind.

Offene Fragen über Wavelets, die die Mathematiker beschäftigen, beziehen sich in erster Linie auf mehrdimensionale Wavelets und Prä-Wavelets. Es wurden vor allem mehrdimensionale (Prä-) Wavelets in der Literatur diskutiert, die auf Gitterstützstellen definiert sind, aber die Frage nach der Existenz auf beliebig verteilten Stützstellen ist für mehr als eine Dimension noch weit offen. Dabei ist besonders zu beachten, dass in mehreren Dimensionen selbst bei äquidistanten Stützstellen nicht mehr ein (Prä-) Wavelet genügt, das dann entlang dem Gitter verschoben wird, sondern mehrere (im einfachsten Falle der Skalierung aller Komponenten mit 2 sind es $2^n - 1$, wobei n die Dimension des unterliegenden Raumes ist) gefunden werden müssen. Besondere Schwierigkeiten ergeben sich, wenn auf beschränkten Mengen $\Omega \subset \mathbb{R}^n$ statt auf dem ganzen Raum gearbeitet werden soll, weil schon der der Zerlegung vorausgehenden Approximation, die im Falle von $n = 1$ (Satz 4) noch recht einfach war, spezielle Beachtung geschenkt werden muss. Umso schwieriger ist die Konstruktion geeigneter (Prä-) Wavelets für diesen Fall.

Als mögliche Verallgemeinerungen des ein- oder mehr-dimensionalen Ansatzes seien nur die Multiwavelets, die Wavelet-Packets [9] und die Windowed Radon Transformation genannt. Obwohl also in den Anwendungen bereits viel und erfolgreich mit Wavelets gearbeitet wird, ist der Bereich der Theorie und mathematischen Analyse bei weitem noch nicht ausgeschöpft. Es sei bei allen diesen Erweiterungsmöglichkeiten nur beispielhaft erwähnt, dass Wavelet-Packets die Zerlegung (9) dahingehend weiterverarbeiten, dass nicht nur für W_0 orthogonale Basen gefordert werden, sondern dass auch der "Rest" in V_0 orthogonal zerlegt werden soll. Dabei sei noch darauf hingewiesen, dass sowohl aus Prä-Wavelets als auch aus Basen in V_0 , die die Riesz-Stabilitätsbedingung erfüllen, Orthonormalsysteme durch Gram-Schmidt Orthogonalisierung gewonnen werden können. Alternativ kann man auch ohne orthonormale Systeme $\psi(2^k \cdot -j)$ eine Zerlegung (2) haben, indem man bi-orthogonale Wavelets berechnet, wo der Wf -Anteil in (2) von einem Orthonormalsystem zu den $\{2^{k/2}\psi(2^k \cdot -j) \mid k, j \in \mathbb{Z}\}$ abhängt.

Es sollte zum Abschluss dieses Artikels hervorgehoben werden, dass er nicht den Anspruch erhebt, eine vollständige oder endgültige Sicht der Wavelets anzugeben. Weit davon entfernt! Wie bereits am Anfang des Artikels erwähnt, ist die derzeitige Entwicklung viel zu rasant, als dass man mehr als einen kleinen Einblick in einem kurzen Aufsatz geben kann, zumal mehrdimensionale Theorie vollständig ausser Acht gelassen wurde und des Autors persönliche Sicht als Approximationstheoretiker und im Hinblick auf Anwendungen gegeben wurde. Der geneigte Leser fasse dieses vielleicht als Ansporn auf, die zitierte und andere Literatur zu Rate zu ziehen und die vielen anderen Sichtweisen zu erfahren.

Literaturverzeichnis

1. Beatson, R.K. und M.J.D. Powell, "Univariate multiquadric approximation: quasi-interpolation to scattered data", *Constructive Approximation* 8 (1992), 275-288.
2. de Boor, C., R.A. DeVore und A. Ron, "On the construction of multivariate (pre) wavelets", *Constructive Approximation* 9 (1993), 123-166.
3. Buhmann, M.D., "Convergence of univariate quasi-interpolation using multiquadrics", *IMA J. Numerical Analysis* 8 (1988), 365-383.
4. Buhmann, M.D., "Discrete least-squares approximation and pre-wavelets from radial function spaces", *Math. Proc. of the Cambridge Phil. Soc.* 114 (1993), 533-558.
5. Buhmann, M.D., "Multiquadric pre-wavelets on non-equally spaced knots in one dimension", *Mathematics of Computation* 10 (1995).
6. Buhmann, M.D., "Pre-wavelets on scattered knots and from radial function spaces: a review", in *den Proceedings of the 6th Conference on Mathematics of Surfaces*, herausgegeben von R.R. Martin, *IMA Conf. Proc. Ser.*, Oxford University Press.
7. Buhmann, M.D., A. Iserles, S.P. Nørsett und J.M. Sanz-Serna, "Sobolev-orthogonal functions and spectral methods", in Vorbereitung.
8. Buhmann, M.D. und C.A. Micchelli, "Spline pre-wavelets for non-uniform knots", *Numerische Mathematik* 61 (1992), 455-474.
9. Chui, C.K., "An Introduction to Wavelets", Academic Press, Boston 1992.
10. Dahmen, W. und C.A. Micchelli, "On stationary subdivision and the construction of compactly supported orthonormal wavelets", in: *Multivariate Approximation and Interpolation*, herausgegeben von K. Jetter und W. Haussmann, Birkhäuser, Basel 1990, 69-89.
11. Daubechies, I., "Ten Lectures on Wavelets", *SIAM-CBMS Lecture Notes* 61, Philadelphia 1992.
12. DeVore, R.A. und B. Lucier, "Wavelets", *Acta Numerica* 1 (1992), 1-56.
13. Feichtinger, H.G. und K. Gröchenig, "Gabor Wavelets and the Heisenberg group: Gabor expansions and short time Fourier transform from the group theoretical point of view", in: *Wavelets*, herausgegeben von C.K. Chui, Academic Press, Boston 1992, 359-397.
14. Louis, A.K., P. Maass und A. Rieder, "Wavelets", Teubner, Stuttgart 1994.
15. Meyer, Y., "Ondelettes et Opérateurs", Hermann, Paris 1990.

16. Micchelli, C.A., "Using the refinement equation for the construction of pre-wavelets", Numer. Algorithms 1 (1991), 75-116.
17. Oswald, P., "Multilevel Finite Element Approximation", Teubner, Stuttgart 1994.
18. Powell, M.J.D., "Univariate multiquadric approximation: reproduction of linear polynomials", in: Multivariate Approximation and Interpolation, herausgegeben von K. Jetter und W. Haussmann, Birkhäuser, Basel 1990, 227-240.
19. Powell, M.J.D., "Approximation Theory and Methods", Cambridge University Press, Cambridge 1981.
20. Riemenschneider, S.D. und Z. Shen, "Box-splines, cardinal series, and wavelets", in: Approximation Theory and Functional Analysis, herausgegeben von C.K. Chui, Academic Press, New York 1991.
21. Stoer, J., "Einführung in die numerische Mathematik I", Springer, Berlin 1983.
22. Strang, G., "Wavelets and dilation equations: a brief introduction", SIAM Review 31 (1989), 614-627.