

Wavelet–Analysis
Wintersemester 2014/2015

Joachim Stöckler

Literatur

Der Katalog der Universität Dortmund liefert 85 Treffer auf das Suchwort "Wavelets".

Als Einführung in Wavelets, auch für Nichtmathematiker gedacht, sind folgende Bücher:

- W. Bäni, "Wavelets: eine Einführung für Ingenieure", Oldenbourg, 2005.
- J. Bergh, F. Ekstedt, M. Lindberg, "Wavelets", Studentlitteratur, 1999.
- C. Blatter, "Wavelets: Eine Einführung", Vieweg, 2003.
- A. Boggess, F. Narcowich, "A First Course in Wavelets with Fourier Analysis", Prentice Hall, 2001.
- C. K. Chui, "Wavelets: a mathematical tool for signal analysis", SIAM, 1997.
- M. W. Frazier, "An Introduction to Wavelets through Linear Algebra", Springer, 1999.
- Y. Meyer, "Wavelets: algorithms and applications", SIAM, 1994.
- G. Strang, T. Nguyen, "Wavelets and Filter Banks", Wellesley-Cambridge Press, 1997.

Fundierte mathematische Einführungen in Wavelet-Analyse geben:

- I. Daubechies, "Ten Lectures on Wavelets", CBMS-NSF Reg. Conf. Series in Appl. Math., SIAM, 1992.
- E. Hernández, G. Weiss, "A first course on Wavelets", CRC Press, 1996.
- A. K. Louis, P. Maaß, A. Rieder, "Wavelets: Theorie und Anwendungen", Teubner, 1998.
- S. Mallat, "A Wavelet Tour of Signal Processing", 2nd. Ed., Academic Press, 1999. (Sehr umfangreich, beinhaltet sehr viele Aspekte der Anwendungen.)
- Y. Meyer, "Wavelets and Operators", Cambridge University Press, 1992. Übersetzung von Band 1 "Ondelettes et Opérateurs", Hermann, Paris.

Weiterführende Bücher, zum Teil als Beiträge verschiedener Autoren oder Tagungsbände:

- C. K. Chui, "Wavelets: a tutorial in theory and applications", Academic Press, 1992.

- C. K. Chui, L. Montefusco, L. Puccio, “Wavelets: theory, algorithms, and applications”, Academic Press, 1994.
- Y. Meyer, “Wavelets: Calderón-Zygmund and Multilinear Operators”, Cambridge University Press, 1997. Übersetzung der Bände 2 und 3 von “ondelettes et Opérateurs”, Hermann, Paris.
- L. L. Schumaker, G. Webb, “Recent advances in wavelet analysis”, Academic Press, 1994.
- G. V. Welland, “Beyond Wavelets”, Academic Press, 2003.

Daneben benötigen wir Grundlagen der Fourier-Reihen und Fourier-Transformation. Als Nachschlagewerk zu Fourier-Reihen (und Aussagen zur Lebesgue-Integration) kann z.B. dienen:

- A. Zygmund, “Trigonometric Series”, Cambridge University Press, 1959 (paperback 1988).

Weiterhin bezieht sich die Vorlesung auf einige einschlägige Veröffentlichungen:

- I. Daubechies, Orthonormal bases of compactly supported wavelets, Commun. Pure Appl. Math. 41, No.7, 909-996 (1988), [Zbl 0644.42026].
- R. Q. Jia, Z. Shen, Multiresolution and Wavelets, Proc. Edinburgh Math. Soc. 37 (1994), 271–300.
- W. Lawton, S. L. Lee, Z. Shen, Convergence of the multivariate cascade algorithm, Numer. Math. 78 (1998), 427–438.
- A. Ron, Z. Shen, Affine systems in $L_2(\mathbb{R}^d)$: the analysis of the analysis operator, J. Functional Analysis 148 (1997), 408-447.

KAPITEL 1

Einführung

Wavelet–Analysis ist ein modernes Gebiet der Angewandten Mathematik, das die Theorie der Fourier-Transformation (siehe Analysis 2) erweitert. Die *Wavelet-Transformation* dient in erster Linie der Analyse von Teilinformationen, die eine gegebene Funktion $f : \mathbb{R}^d \rightarrow \mathbb{C}$ enthält. In praktischen Anwendungen ist f zum Beispiel die Wellenform eines akustischen Signals ($d = 1$) oder die Grauwertverteilung eines Bildes ($d = 2$, \mathbb{R}^2 wird ersetzt durch einen rechteckigen Definitionsbereich).

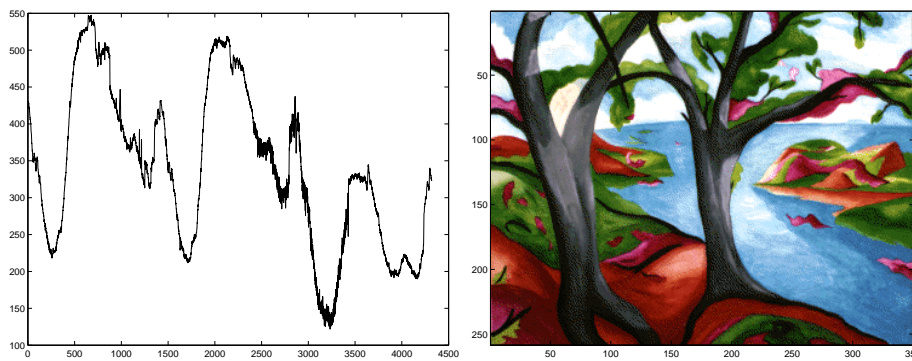


BILD 1.1. Funktionen in der Signal- und Bildverarbeitung

Ziele der Wavelet-Transformation sind unter anderem

- die zeitlich eingegrenzte Frequenzanalyse eines akustischen Signals,
- die Datenkompression (z.B. JPEG2000 bei der Bildverarbeitung),
- das Entrauschen gestörter Messwerte.

Hierbei werden besonders die Eigenschaften der Wavelet-Transformation als *Filter* betont, der Eigenschaften aus einer gegebenen Funktion f “herausfiltert”.

Als mathematisches Werkzeug dient die Wavelet-Transformation weiterhin

- zur Charakterisierung von “lokalen” Eigenschaften von Funktionen, z.B. der lokalen Lipschitz-Stetigkeit von Funktionen,
- zur Darstellung und Definition von wichtigen Funktionenräumen (z.B. Sobolev-Räume),
- zur Diskretisierung von Differentialgleichungen.

Dadurch werden enge Beziehungen zur Funktionalanalysis, Approximationstheorie und Numerischen Mathematik hergestellt.

1. Wichtige Begriffe aus der Linearen Algebra

DEFINITION 1.1. Gegeben sei ein Vektorraum V über dem Körper \mathbb{C} sowie eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$$

mit den folgenden Eigenschaften:

- (i) *Definitheit:* $\langle v, v \rangle \geq 0$ für alle $v \in V$ und $\langle v, v \rangle = 0 \implies v = 0$;
- (ii) *Linearität in der 1. Komponente:* $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$ für alle $\alpha, \beta \in \mathbb{C}$ und $u, v, w \in V$;
- (iii) *Symmetrie:* $\langle v, w \rangle = \overline{\langle w, v \rangle}$ für alle $v, w \in V$.

Dann heißt die Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ ein Skalarprodukt auf V , und der Vektorraum V heißt ein Skalarproduktraum.

Auf einem Skalarproduktraum V wird durch

$$\|v\| := \sqrt{\langle v, v \rangle}, \quad v \in V,$$

eine Norm induziert. Hiermit ist jeder Skalarproduktraum auch ein *normierter Raum*. Die Konvergenz einer Folge $(v_n)_{n \in \mathbb{N}}$ mit Folgengliedern $v_n \in V$ wird wie üblich definiert:

$$\lim_{n \rightarrow \infty} v_n = v \iff \lim_{n \rightarrow \infty} \|v_n - v\| = 0.$$

Ebenso ist der Begriff der Cauchy-Folge in V wie üblich definiert: $(v_n)_{n \in \mathbb{N}}$ ist Cauchy-Folge genau dann, wenn zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass für alle $n, m \geq N$ die Ungleichung $\|v_n - v_m\| < \epsilon$ gilt; dies wird kurz ausgedrückt in der Form

$$\lim_{n, m \rightarrow \infty} \|v_n - v_m\| = 0.$$

DEFINITION 1.2. Ein Skalarproduktraum V heißt vollständig, wenn jede Cauchy-Folge konvergiert, also zu jeder Cauchy-Folge $(v_n)_{n \in \mathbb{N}}$ ein Element $v \in V$ existiert mit $\lim_{n \rightarrow \infty} v_n = v$.

Ein vollständiger Skalarproduktraum V heißt Hilbertraum.

BEMERKUNG 1.3. Der obige Konvergenzbegriff beschreibt die sog. starke Konvergenz. Wir werden später auch den Begriff der schwachen Konvergenz kennenlernen.

BEMERKUNG 1.4. In jedem Skalarproduktraum V gilt die Cauchy-Schwarz Ungleichung

$$(1.1) \quad |\langle v, w \rangle| \leq \|v\| \|w\|, \quad v, w \in V,$$

sowie die Dreiecksungleichung

$$(1.2) \quad \|v + w\| \leq \|v\| + \|w\|, \quad v, w \in V.$$

BEISPIEL 1.5.

- a) Die Elemente von \mathbb{C}^N werden als Spaltenvektoren $x = (x_1, \dots, x_N)^T$ mit komplexen Komponenten x_k , $1 \leq k \leq N$, aufgefasst. Versehen mit dem Skalarprodukt

$$\langle x, y \rangle = \sum_{k=1}^N x_k \bar{y}_k$$

ist \mathbb{C}^N ein Hilbertraum, der N -dimensionale Euklidische Raum. Wir identifizieren diesen Raum häufig mit den folgenden Räumen:

- $\ell_N := \{x = (x_0, \dots, x_{N-1}); x_k \in \mathbb{C} \text{ für } 0 \leq k \leq N-1\}$, mit dem Skalarprodukt

$$(1.3) \quad \langle x, y \rangle = \sum_{k=0}^{N-1} x_k \bar{y}_k.$$

- dem Vektorraum der N -periodischen Folgen

$$\tilde{\ell}_N := \{x = (x_k)_{k \in \mathbb{Z}}; x_k \in \mathbb{C}, x_{k+N} = x_k \text{ für } k \in \mathbb{Z}\},$$

mit dem Skalarprodukt

$$\langle x, y \rangle = \sum_{k=0}^{N-1} x_k \bar{y}_k, \quad x = (x_k)_{k \in \mathbb{Z}}, y = (y_k)_{k \in \mathbb{Z}} \in \tilde{\ell}_N,$$

- dem Vektorraum V aller Abbildungen $f : \{0, 1, \dots, N-1\} \rightarrow \mathbb{C}$ mit dem Skalarprodukt

$$\langle f, g \rangle = \sum_{k=0}^{N-1} f(k) \bar{g}(k), \quad f, g \in V.$$

- b) Der Vektorraum der quadrat-summierbaren Folgen

$$\ell^2 = \ell^2(\mathbb{Z}) := \{x = (x_k)_{k \in \mathbb{Z}}; x_k \in \mathbb{C}, \sum_{k \in \mathbb{Z}} |x_k|^2 < \infty\},$$

ist mit dem Skalarprodukt

$$\langle x, y \rangle = \sum_{k=-\infty}^{\infty} x_k \bar{y}_k$$

ein Hilbertraum. Die zugehörige Norm ist die ℓ^2 -Norm

$$\|x\|_2 = \left(\sum_{k=-\infty}^{\infty} |x_k|^2 \right)^{1/2}.$$

Hier lautet die Cauchy-Schwarz Ungleichung

$$\left| \sum_{k=-\infty}^{\infty} x_k \bar{y}_k \right| \leq \left(\sum_{k=-\infty}^{\infty} |x_k|^2 \right)^{1/2} \left(\sum_{k=-\infty}^{\infty} |y_k|^2 \right)^{1/2}.$$

- c) Der Vektorraum der quadrat-integrierbaren Funktionen auf einem Intervall $I \subset \mathbb{R}$ (unbeschränkte Intervalle eingeschlossen)

$$L^2(I) := \{f : I \rightarrow \mathbb{C}; f \text{ messbar und } \int_I |f(x)|^2 dx < \infty\},$$

ist mit dem Skalarprodukt

$$\langle f, g \rangle = \int_I f(x) \overline{g(x)} dx$$

ein Hilbertraum. Die zugehörige Norm ist die L^2 -Norm

$$\|f\|_2 = \left(\int_I |f(x)|^2 dx \right)^{1/2}.$$

Hier lautet die Cauchy-Schwarz Ungleichung

$$\left| \int_I f(x) \overline{g(x)} dx \right| \leq \left(\int_I |f(x)|^2 dx \right)^{1/2} \left(\int_I |g(x)|^2 dx \right)^{1/2}.$$

BEISPIEL 1.6. Der Vektorraum $C([-1, 1])$ mit dem obigen Skalarprodukt

$$\langle f, g \rangle = \int_{-1}^1 f(x) \overline{g(x)} dx$$

ist ein Skalarproduktraum, aber kein Hilbertraum. Die Funktionenfolge

$$f_n : [-1, 1] \rightarrow \mathbb{C}, \quad f_n(x) = \begin{cases} (1+x)^n, & -1 \leq x \leq 0, \\ 1, & 0 < x \leq 1, \end{cases}$$

ist eine Cauchy-Folge bzgl. der L^2 -Norm. Sie konvergiert punktweise (und in der L^2 -Norm) gegen die nichtstetige Funktion

$$g : [-1, 1] \rightarrow \mathbb{C}, \quad g(x) = \begin{cases} 0, & -1 \leq x \leq 0, \\ 1, & 0 < x \leq 1. \end{cases}$$

Die Folge hat also keinen Grenzwert in $C([-1, 1])$.

2. Die diskrete Welt

Wesentliche Aspekte der Wavelet-Analyse lassen sich bereits anhand des Hilbertraums \mathbb{C}^N bzw. des isomorphen Raumes ℓ_N veranschaulichen.

Zunächst aber die folgende Begriffsbildung:

DEFINITION 1.7. Die Vektoren $e_k \in \ell_N$, $0 \leq k \leq n$ (mit $n \leq N-1$), bilden ein Orthonormalsystem, falls

$$(1.4) \quad \langle e_m, e_n \rangle = \begin{cases} 1 & \text{für } m = n, \\ 0 & \text{für } m \neq n \end{cases}$$

gilt. Im Fall $n = N-1$ bilden diese Vektoren eine Orthonormalbasis.

BEMERKUNG 1.8. Zu gegebener Orthonormalbasis $(e_n)_{0 \leq n \leq N-1}$ besitzt jedes Element $v \in \ell_N$ die eindeutige Darstellung

$$(1.5) \quad v = \sum_{k=0}^{N-1} \langle v, e_k \rangle e_k$$

als Linearkombination der Basisvektoren. Die Identität (1.4) liefert die Beziehung

$$(1.6) \quad \|v\|^2 = \sum_{k=0}^{N-1} |\langle v, e_k \rangle|^2, \quad \text{Satz des Pythagoras}$$

d.h. die “Energie” $\|v\|^2$ des Signals v wird exakt durch das Quadrat der ℓ_2 -Norm der Koeffizienten $\langle v, e_k \rangle$ dargestellt. Wir betrachten den einzelnen Koeffizienten $c_k := \langle v, e_k \rangle$ als die “Teilinformation” über den Vektor v , den die Basisfunktion e_k liefert (repräsentiert, herausfiltert).

BEISPIEL 1.9. a) Die “kanonische” Orthonormalbasis von ℓ_N lautet

$$\begin{aligned} e_0 &= (1, 0, 0, \dots, 0), \\ e_1 &= (0, 1, 0, \dots, 0), \\ &\vdots \\ e_{N-1} &= (0, 0, 0, \dots, 1). \end{aligned}$$

Der Koeffizient $c_k = \langle v, e_k \rangle = v_k$ ist die k -te Koordinate von v . Betrachten wir v als zeitdiskretes Signal (mit Amplituden v_k an den Zeitpunkten kT , $T = \text{Zeiteinheit}$), so repräsentiert der Koeffizient c_k die Amplitude zum Zeitpunkt kT .

b) Von der trigonometrischen Interpolation (siehe Numerik I) kennen wir die Orthonormalbasis der komplexen Vektoren

$$e_k = \frac{1}{\sqrt{N}} (1, e^{2\pi i \frac{k}{N}}, e^{2\pi i \frac{2k}{N}}, \dots, e^{2\pi i \frac{(N-1)k}{N}}), \quad 0 \leq k \leq N-1.$$

Der Koeffizient

$$c_k = \langle v, e_k \rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} v_j e^{-2\pi i \frac{jk}{N}}$$

ist (bis auf einen Faktor $\frac{1}{\sqrt{N}}$) der diskrete Fourierkoeffizient¹ von v . Die Identität (1.5) stellt v als (lineare) Überlagerung der “harmonischen Frequenzen” e_k dar. Also ist der Absolutbetrag $|c_k|$ die Amplitude der k -ten Frequenz von v .

Der gesamte Koeffizientenvektor $(c_k)_{0 \leq k \leq N-1}$ wird *diskrete Fourier-Transformierte* von v genannt. Zu seiner Berechnung wird die “Schnelle Fourier-Transformation” (FFT=fast Fourier transform) verwendet. Für $N = 2^K$ werden dazu nur $\mathcal{O}(N \log_2 N)$ Rechenoperationen benötigt.

c) Die *diskrete Walsh-Transformation*² in ℓ_N , mit $N = 2^K$, verwendet die eindeutige Orthonormalbasis $(e_k)_{0 \leq k \leq N-1}$, die durch folgende Festlegungen entsteht:

- $\tilde{e}_k = \sqrt{N} e_k$ besitzt nur die Einträge 1 und -1 und der erste Eintrag ist jeweils 1,
- \tilde{e}_k hat genau k Vorzeichenwechsel,
- die Vektoren sind paarweise orthogonal.

Beispiel für $N = 4$:

$$\tilde{e}_0 = (1, 1, 1, 1)^T, \quad \tilde{e}_1 = (1, 1, -1, -1)^T, \quad \tilde{e}_2 = (1, -1, -1, 1)^T, \quad \tilde{e}_3 = (1, -1, 1, -1)^T.$$

Die Gesamtheit der Koeffizienten $c_k = \langle v, e_k \rangle$ nennt man die *Walsh-Transformierte* von v . Da die Vektoren e_k , $0 \leq k \leq N-1$, mit wachsendem k stärker oszillieren, wird diese Transformation manchmal als “Poor man’s FFT” bezeichnet. Ihre Berechnung ist effizienter als die der FFT, da nur Additionen und Subtraktionen auftreten.

¹Jean Baptiste Joseph Fourier, 1768–1830

²Joseph Leonard Walsh, 1895–1973

Hier noch ein kleines Matlab-Programm, das die Vektoren $\tilde{e}_n = \sqrt{N}e_n$, $0 \leq n \leq N-1$ für $N = 2^K$ berechnet (Ausgabe als Zeilen der Matrix w):

```
function w=walsh(K)
% generate 2^K Walsh functions
w=[1];
for i=1:K
    w=[w,w;w,-w];
end
% for sequence ordering according to number of sign changes
c=sum(abs(diff(w')))/2; % number of sign changes in rows
[c,i]=sort(c);
w=w(i,:);
```

Weitere Informationen (incl. der obigen Rekursion) enthält die Datei walsh.m auf der Vorlesungs-Webseite.

ÜBUNG 1.10. In Übungsaufgabe 3 kann man die Walsh-Basis herstellen, indem man über “Wavelet-Packet 1D” sich die Funktionen des “Haar”-Typs anschaut. Einziger Unterschied: die Reihenfolge der Nummerierung der Basisfunktionen ist anders.

Wie wir gesehen haben, liefert die Wahl verschiedener ONBs $(e_k)_{0 \leq k \leq N-1}$ anhand der Koeffizienten $c_k = \langle v, e_k \rangle$ auch unterschiedlichen Informationsgehalt. Die am stärksten “lokalisierte” Information erhalten wir von der kanonischen Basis der Einheitsvektoren (Beispiel a), die beste Frequenz-Information von der diskreten Fouriertransformation (Beispiel b). Andererseits liefert die kanonische Basis keinerlei Information über die Frequenzen des Vektors v , denn dazu müssten ja mehrere Werte v_k miteinander verglichen werden. Analog liefern die diskreten Fourierkoeffizienten von v keine direkte Information über die Signalwerte v_k selbst, da ein Mittel über den gesamten Zeitbereich gebildet wird.

In gewisser Weise stellt die Wavelet-Analyse einen Zwischenweg zwischen diesen beiden Extremfällen dar. Die erhaltenen Koeffizienten erlauben sowohl Rückschlüsse auf die Komponenten v_k selbst als auch auf die diskreten Fourierkoeffizienten. Wir machen dies am Beispiel der (diskreten) Haar-Wavelets³ deutlich.

Wir betrachten den Fall $N = 8$. Die ersten beiden Basisvektoren der Orthonormalbasis sind

$$e_0 = \frac{1}{2\sqrt{2}}(1, 1, 1, 1, 1, 1, 1, 1), \quad e_1 = \frac{1}{2\sqrt{2}}(1, 1, 1, 1, -1, -1, -1, -1),$$

stimmen also mit denen der Walsh-Basis überein. Alle weiteren Elemente entstehen durch “Stauen” und “Verschieben” von e_1 , z.B. sind

$$e_2 = \frac{1}{2}(1, 1, -1, -1, 0, 0, 0, 0), \quad e_3 = \frac{1}{2}(0, 0, 0, 0, 1, 1, -1, -1),$$

$$e_4 = \frac{1}{\sqrt{2}}(1, -1, 0, 0, 0, 0, 0, 0), \quad e_5 = \frac{1}{\sqrt{2}}(0, 0, 1, -1, 0, 0, 0, 0), \quad \text{etc.}$$

³Alfréd Haar, 1885–1933

Für allgemeines $N = 2^K$ ist e_0 der konstante Vektor $2^{-K/2}(1, 1, \dots, 1)$ und für jedes weitere $n = 2^j + k$, $0 \leq j \leq K-1$, $0 \leq k < 2^j$, ist

$$e_n = 2^{(j-K)/2} \left(\underbrace{0, \dots, 0}_{2^{K-j-k}}, \underbrace{1, \dots, 1}_{2^{K-j-1}}, \underbrace{-1, \dots, -1}_{2^{K-j-1}}, \underbrace{0, \dots, 0}_{2^{K-j}(2^j-k-1)} \right).$$

Man erkennt hieran, dass für festes $0 \leq j \leq K-1$ der Bereich der nicht verschwindenden Koordinaten von e_{2^j+k} die Länge $N/2^j$ besitzt, sich also mit wachsendem j verkleinert. Der weitere Index $0 \leq k < 2^j$ bewirkt, dass dieser Bereich der Länge $N/2^j$ von links nach rechts wandert, ohne dass dabei Überlappungen auftreten. Mit wachsendem j werden also immer kürzere Bereiche des Vektors v “gefiltert”. Die Positionierung dieser Bereiche erfolgt über den Index k .

Analog zum Walsh-Beispiel hier ein Matlab-Programm, das die Vektoren $\tilde{e}_n = \sqrt{N}e_n$, $0 \leq n \leq N-1$ für $N = 2^K$ berechnet (Ausgabe als Zeilen der Matrix H):

```
function H = haar(J)
% HAAR computes the Haar matrix of size 2^J
H = 1;
for j = 1:J
    H = [kron(H, [1,1]); kron(eye(2^(j-1)), [1,-1])] / sqrt(2);
end
```

BEMERKUNG 1.11. *Man zeigt, dass die Vektoren e_n , $0 \leq n \leq 2^K - 1$, eine Orthonormalbasis von ℓ_{2^K} bilden. Dazu ersetzt man den Basisvektor e_n für $n \geq 1$ durch die Treppenfunktion*

$$\psi_n = 2^{(j-K)/2} \left(\chi_{[2^k, 2^{K-j-1}, (2k+1)2^{K-j-1})} - \chi_{[(2k+1)2^{K-j-1}, (2k+2)2^{K-j-1})} \right)$$

und betrachtet die inneren Produkte

$$\langle e_n, e_m \rangle = \int_0^{2^K} \psi_n(x) \psi_m(x) dx.$$

Um den Informationsgehalt der Koeffizienten $c_k = \langle v, e_k \rangle$ zur Haar-Basis zu interpretieren, führen wir die folgenden Bezeichnungen ein: für gerades $M \in \mathbb{N}$ setze

$$a(x_k, \dots, x_{k+M-1}) := \frac{1}{M} \sum_{i=k}^{k+M-1} x_i,$$

$$d(x_k, \dots, x_{k+M-1}) := \frac{1}{M} \left(\sum_{i=k}^{k+M/2-1} x_i - \sum_{i=k+M/2}^{k+M-1} x_i \right)$$

(“a” für “average” und “d” für “detail”). Dann gilt

- $c_0 = \langle v, e_0 \rangle = \sqrt{N}a(v_0, \dots, v_{N-1})$; dies ist der Mittelwert des gesamten Vektors v ,
- $c_1 = \langle v, e_1 \rangle = \sqrt{N}d(v_0, \dots, v_{N-1})$; dies ist die Differenz der beiden Mittel, die jeweils über die Hälfte der Koordinaten von v gebildet werden.
- $c_{2^j+k} = d(v_{kN/2^j}, \dots, v_{(k+1)N/2^j-1})$ enthält die Differenz der beiden Mittel gebildet über einen Abschnitt von v der Länge $N/2^j$. Dieser Abschnitt ergibt bei festem j und wachsendem k eine disjunkte Partition des Vektors.

Mit anderen Worten: Die Stauchung (Parameter 2^j) bewirkt die Analyse immer kürzerer Abschnitte des Vektors v . Wir betrachten daher 2^j (oder einfach j) als die Verfeinerungsstufe der Wavelet-Analyse. Die zusätzliche Verschiebung (Parameter k) stellt sicher, dass auf jeder Verfeinerungsstufe der gesamte Vektor in die Analyse einbezogen wird. Die einzelnen Koeffizienten geben daher eine (grobe) Frequenzinformation über einen bestimmten Teilbereich des gegebenen Vektors v , dessen Länge mit einer Zweierpotenz (“dyadisch”) abnimmt.

Weiterhin zeigen die Formeln

$$\begin{aligned} 2a(x_0, \dots, x_{M-1}) &= a(x_0, \dots, x_{M/2-1}) + a(x_{M/2}, \dots, x_{M-1}), \\ 2d(x_0, \dots, x_{M-1}) &= a(x_0, \dots, x_{M/2-1}) - a(x_{M/2}, \dots, x_{M-1}), \end{aligned}$$

dass sehr einfache Rekursionsformeln (ausgehend von der Detailinformation $a(x_k) = x_k$ der Länge 1) zur Berechnung aller Koeffizienten der Haar-Basis führen. Weiterhin erhält man die einfachen Umkehrformeln

$$\begin{aligned} a(x_0, \dots, x_{M/2-1}) &= a(x_0, \dots, x_{M-1}) + d(x_0, \dots, x_{M-1}), \\ a(x_{M/2}, \dots, x_{M-1}) &= a(x_0, \dots, x_{M-1}) - d(x_0, \dots, x_{M-1}). \end{aligned}$$

Hiermit lassen sich aus allen Koeffizienten $c_n = \langle v, e_n \rangle$, $0 \leq n \leq 2^K - 1$, zur Haar-Basis wieder die Folgenglieder v_k rekursiv berechnen.

Beispiel: Betrachtet man $x \in \ell_8$ mit $x = (1, 3, 2, 1, -1, 4, 2, 1)$, so ergibt sich das Schema

$$\begin{array}{l} \left. \begin{array}{l} a(x_0, x_1) = \frac{1+3}{2} = 2 \\ a(x_2, x_3) = \frac{3}{2} \\ a(x_4, x_5) = \frac{3}{2} \\ a(x_6, x_7) = \frac{3}{2} \end{array} \right\} \begin{array}{l} d(x_0, x_1) = \frac{1-3}{2} = -1 \\ d(x_2, x_3) = \frac{1}{2} \\ d(x_4, x_5) = -\frac{5}{2} \\ d(x_6, x_7) = \frac{1}{2} \end{array} \begin{array}{l} \cdot \sqrt{2} \\ \cdot \sqrt{2} \\ \cdot \sqrt{2} \\ \cdot \sqrt{2} \end{array} \left. \begin{array}{l} c_4 = -\sqrt{2} \\ c_5 = 1/\sqrt{2} \\ c_6 = -5/\sqrt{2} \\ c_7 = 1/\sqrt{2} \end{array} \right. \\ \downarrow \quad \searrow \\ \left. \begin{array}{l} a(x_0, \dots, x_3) = \frac{7}{4} \\ a(x_4, \dots, x_7) = \frac{3}{2} \end{array} \right\} \begin{array}{l} d(x_0, \dots, x_3) = \frac{1}{4} \\ d(x_4, \dots, x_7) = 0 \end{array} \begin{array}{l} \cdot \sqrt{4} \\ \cdot \sqrt{4} \end{array} \left. \begin{array}{l} c_2 = 1/2 \\ c_3 = 0 \end{array} \right. \\ \downarrow \quad \searrow \\ a(x_0, \dots, x_7) = \frac{13}{8} \quad d(x_0, \dots, x_7) = \frac{1}{8} \left. \begin{array}{l} \cdot \sqrt{8} \\ \cdot \sqrt{8} \end{array} \right\} c_1 = 1/\sqrt{8} \end{array}$$

Den Koeffizienten c_0 erhält man als $c_0 = \sqrt{8}a(x_0, \dots, x_7) = 13/\sqrt{8}$.

Das abschließende numerische Beispiel soll die 4 Transformationen gegenüberstellen.

3. Die kontinuierliche Welt

Die eigentliche Struktur der Wavelet-Analyse zeigt sich noch besser durch Betrachtung von Funktionenräumen wie z.B. den Hilberträumen $L^2([0, 1])$ oder $L^2(\mathbb{R})$. Wir gehen vor wie im vorherigen Abschnitt und definieren zunächst den Begriff der Orthonormalbasis.

DEFINITION 1.12. H sei ein Hilbertraum mit $\dim H = \infty$. Die Vektoren $e_n \in H$, $n \in \mathbb{N}$, bilden ein Orthonormalsystem in H , falls

$$\langle e_m, e_n \rangle = \begin{cases} 1 & \text{für } m = n, \\ 0 & \text{für } m \neq n \end{cases}$$

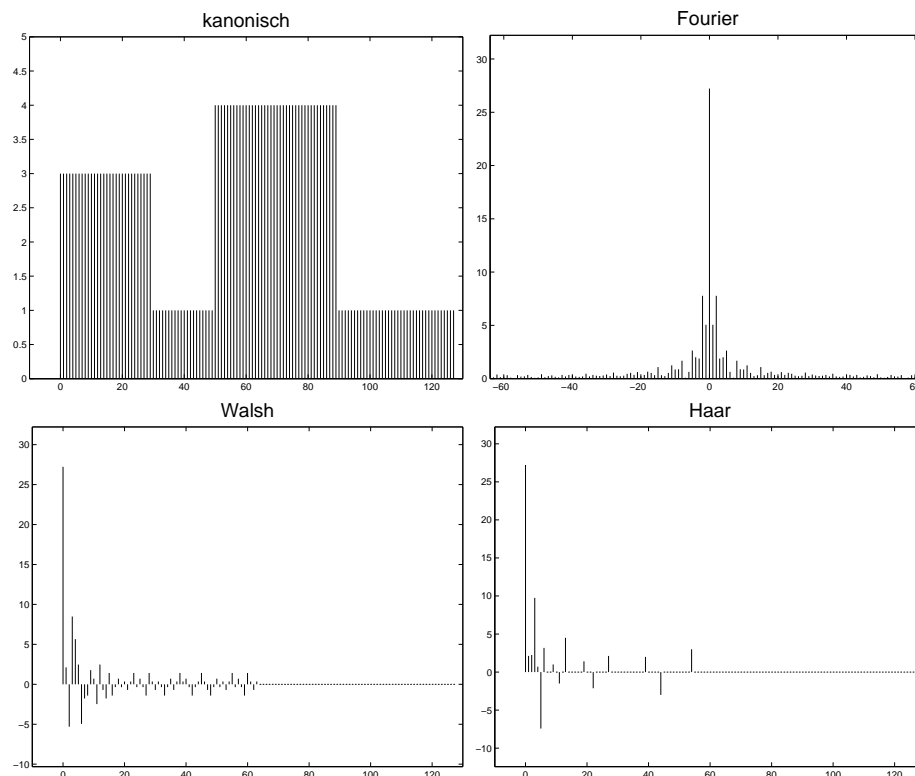


BILD 1.2. Darstellung eines einfachen Signals in vier verschiedenen Orthonormalbasen

gilt. Falls zusätzlich für jedes $h \in H$

$$\lim_{N \rightarrow \infty} \left\| h - \sum_{n=1}^N \langle h, e_n \rangle e_n \right\| = 0.$$

gilt, so ist $(e_n)_{n \in \mathbb{N}}$ eine Orthonormalbasis von H .

BEMERKUNG 1.13. a) Die zweite Eigenschaft besagt, dass

$$(1.7) \quad h = \sum_{n=1}^{\infty} \langle h, e_n \rangle e_n$$

gilt, wobei die Reihe in der Norm von H konvergiert (sog. "starke Konvergenz"). Für $H = L^2(\mathbb{R})$ ist dies die Konvergenz im quadratischen Mittel, die von der Behandlung der Fourier-Reihen in der Analysis bekannt sein sollte. Man nennt die Reihe auch eine Orthogonalreihe.

Orthogonalreihen sind "unbedingt konvergent", d.h. sie dürfen beliebig umgeordnet werden, ohne dass der Grenzwert sich ändert. Deshalb können wir o.E. andere abzählbare Indexmengen, wie z.B. \mathbb{Z} oder $\mathbb{Z} \times \mathbb{Z}$ verwenden. Dies ist manchmal hilfreich, um gewisse Eigenschaften der Basisfunktion e_n durch die Indizierung auszudrücken, z.B. $e_{j,k}$ mit $j, k \in \mathbb{Z}$ (s. Beispiele unten).

b) *Es gelten die Besselsche Ungleichung*

$$(1.8) \quad \sum_{n=1}^N |\langle h, e_n \rangle|^2 \leq \|h\|^2, \quad h \in H,$$

sowie die Parseval-Identität

$$(1.9) \quad \sum_{n=1}^{\infty} |\langle h, e_n \rangle|^2 = \|h\|^2, \quad h \in H.$$

Diese Identität ist wiederum äquivalent zur Plancherel-Identität für das Skalarprodukt von zwei Elementen,

$$(1.10) \quad \langle g, h \rangle = \sum_{n=1}^{\infty} \langle g, e_n \rangle \overline{\langle h, e_n \rangle}, \quad g, h \in H.$$

Da auf der rechten Seite das ℓ^2 -Skalarprodukt der Koeffizientenfolgen

$$c = (\langle g, e_n \rangle)_{n \in \mathbb{N}}$$

von g sowie

$$d = (\langle h, e_n \rangle)_{n \in \mathbb{N}}$$

von h steht, wird durch jede Orthonormalbasis $(e_n)_{n \in \mathbb{N}}$ ein Isomorphismus der Hilberträume H und $\ell^2(\mathbb{N})$ erzeugt.

BEMERKUNG 1.14. *Ein Hilbertraum H mit $\dim H = \infty$ heißt separabel, wenn es eine abzählbare Teilmenge $M \subset H$ gibt, die dicht in H ist. Die folgende Aussage ist gar nicht schwer zu beweisen: Der Hilbertraum H ist genau dann separabel, wenn er eine abzählbare Orthonormalbasis besitzt.*

Beweis von " \implies ": Die Gram-Schmidt Orthogonalisierung macht aus einer abzählbaren dichten Teilmenge $(y_n)_{n \in \mathbb{N}}$ (unter Weglassen evtl. entstehender Nullvektoren) eine Orthonormalbasis.

Beweis von " \impliedby ": Zu gegebener Orthonormalbasis $(e_n)_{n \in \mathbb{N}}$ ist die Menge

$$M = \left\{ \sum_{n=1}^N c_n e_n; N \in \mathbb{N}, c_n \in \mathbb{Q} + i\mathbb{Q} \text{ für } 1 \leq n \leq N \right\}$$

abzählbar und dicht in H .

Die von uns verwendeten Hilberträume sind sämtlich separabel.

BEISPIEL 1.15. a) Der Hilbertraum $L^2([0, 1])$ besitzt die Orthonormalbasis der Funktionen $e_n, n \in \mathbb{Z}$, mit

$$e_n(x) = e^{2\pi i n x}, \quad x \in [0, 1].$$

(Hier wird als Indexmenge \mathbb{Z} anstatt \mathbb{N} verwendet, weil dadurch die Basis-elemente e_n klarer beschrieben werden.) Anders ausgedrückt: jede quadratintegrierbare Funktion $f \in L^2([0, 1])$ ist der Grenzwert ihrer Fourier-Reihe

$$(1.11) \quad f = \sum_{n=-\infty}^{\infty} c_n(f) e_n, \quad c_n(f) = \langle f, e_n \rangle = \int_0^1 f(x) e^{-2\pi i n x} dx.$$

Die Konvergenz erfolgt in der Norm des Hilbertraums $L^2([0, 1])$ und wird *Konvergenz im quadratischen Mittel* genannt,

$$\lim_{N \rightarrow \infty} \int_0^1 \left| f(x) - \sum_{n=-N}^N c_n(f) e^{2\pi i n x} \right|^2 dx = 0.$$

Die Koeffizienten $c_n(f) = \langle f, e_n \rangle$ sind die *Fourierkoeffizienten* der Funktion f . Die Darstellung von f in (1.11) besagt, dass jede Funktion $f \in L^2([0, 1])$ eine (unendliche) Überlagerung harmonischer Schwingungen e_n , $n \in \mathbb{Z}$, ist. Der Fourierkoeffizient gibt die Amplitude und Phasenverschiebung der n -ten periodischen Schwingung an.

Der Hilbertraum $L^2([0, 1])$ wird mit dem Raum der 1-periodischen Funktionen $f: \mathbb{R} \rightarrow \mathbb{C}$ mit $f(x) = f(x+1)$, $x \in \mathbb{R}$, identifiziert.

Physikalische Messungen liefern häufig nicht die zugrunde liegende Funktion f (in ihrer "Wellenform"), sondern die Werte $c_n(f)$ ihres Spektrums (s. Spektroskopie). Um daraus die Funktion f zu erhalten, bildet man die Fourierreihe ("Fourier-Synthese").

- b) Der Hilbertraum $L^2(\mathbb{R})$ der (nicht-periodischen) quadrat-integrierbaren Funktionen ist separabel. Eine Orthonormalbasis (die aber nicht weiter verwendet wird) ist durch die Familie $e_{n,k}$, $n, k \in \mathbb{Z}$, mit

$$e_{n,k}(x) = e^{2\pi i n(x-k)} \chi_{[k, k+1)}(x)$$

gegeben. Hierbei bezeichnet χ_I die charakteristische Funktion zur Menge I .

- c) Die Walsh-Funktionen werden ähnlich wie im diskreten Fall gebildet. Die n -te Walsh-Funktion $w_n: [0, 1) \rightarrow \mathbb{R}$, $n \geq 0$, nimmt nur die Werte 1 oder -1 an, erfüllt $w_n(0) = 1$, hat genau n Sprungstellen in "dyadischen" Punkten der Form $k2^{-j} \in (0, 1)$ und ist (bzgl. des gewöhnlichen Skalarprodukts) orthogonal zu w_0, \dots, w_{n-1} (und damit zu allen w_k , $k \neq n$). Eine explizite Formel der w_n , bis auf Vertauschung der Reihenfolge innerhalb einer "Oktave" $2^{m-1} \leq n \leq 2^m - 1$, lautet

$$w_n(x) = (-1)^{k_0 x_0 + k_1 x_1 + \dots + k_m x_m},$$

wobei $n = \sum_{i=0}^m k_i 2^i$ und $x = \sum_{i \geq 0} x_i 2^{-1-i}$.

Wir kommen nun zur Vorstellung der einfachsten Wavelet-Basis, der so genannten *Haar-Basis* von $L^2([0, 1])$. Die Indexmenge zur Bezeichnung der Basiselemente wird wieder nicht durch \mathbb{N} , sondern durch

$$\mathbb{J} = \{0\} \cup \{(j, k); j \in \mathbb{N}_0, 0 \leq k \leq 2^j - 1\}$$

festgelegt. Die Funktionen der Orthonormalbasis lauten

(1.12)

$$e_0(x) = 1, \quad e_{j,k}(x) = \begin{cases} 2^{j/2}, & x \in [k2^{-j}, (k + \frac{1}{2})2^{-j}), \\ -2^{j/2}, & x \in [(k + \frac{1}{2})2^{-j}, (k + 1)2^{-j}), \\ 0, & \text{sonst,} \end{cases} \quad x \in [0, 1).$$

Der Träger $\text{supp } e_{j,k}$ ist der Abschluss des dyadischen Intervalls $I_{j,k} = [k2^{-j}, (k + 1)2^{-j})$. Dieses Intervall hat die Länge 2^{-j} .

THEOREM 1.16. *Die Funktionen $e_0, e_{j,k}$ mit $j \in \mathbb{N}_0$ und $0 \leq k \leq 2^j - 1$ bilden eine Orthonormalbasis von $L^2([0, 1])$.*

Beweis: Orthonormalität: Es gilt $\|e_0\| = 1$ sowie

$$\langle e_0, e_{j,k} \rangle = \int_0^1 e_{j,k}(x) dx = 0, \quad j \in \mathbb{N}_0, \quad 0 \leq k \leq 2^j - 1.$$

Weiterhin ist für alle $j \in \mathbb{N}_0$ und $0 \leq k \leq 2^j - 1$

$$\|e_{j,k}\|^2 = \int_0^1 |e_{j,k}(x)|^2 dx = \int_{I_{j,k}} 2^j dx = 1.$$

Für $0 \leq k' \leq 2^j - 1$, $k \neq k'$, gilt $\langle e_{j,k'}, e_{j,k} \rangle = 0$, weil die Träger der Funktionen disjunkt sind (genauer: der Durchschnitt ist höchstens ein Punkt, also eine Nullmenge). Für $j' < j$ und $0 \leq k' \leq 2^{j'} - 1$ ist

$$\langle e_{j',k'}, e_{j,k} \rangle = \int_{I_{j,k}} e_{j',k'}(x) e_{j,k}(x) dx = 0,$$

weil $e_{j',k'}$ auf $I_{j,k}$ konstant ist (entweder 0, 1 oder -1) und das Integralmittel von $e_{j,k}$ Null ist. Damit ist die paarweise Orthogonalität aller Funktionen gezeigt.

Vollständigkeit: Wir zeigen, dass für beliebiges $f \in L^2([0, 1))$ aus

$$(1.13) \quad \langle f, e_0 \rangle = \langle f, e_{j,k} \rangle = 0 \quad \text{für alle } j, k$$

auch $f = 0$ folgt.

1. Schritt: Für ein dyadisches Intervall $I_{j,k} = [k2^{-j}, (k+1)2^{-j}]$ (mit $j \in \mathbb{N}_0$ und $0 \leq k \leq 2^j - 1$) sei

$$a_{j,k} := 2^{j/2} \int_{I_{j,k}} f(x) dx.$$

Mittels vollständiger Induktion nach j zeigen wir, dass $a_{j,k} = 0$ für alle j, k gilt. Zunächst folgt aus $I_{0,0} = [0, 1)$ und der ersten Identität in (1.13) sofort $a_{0,0} = 0$. Das dyadische Intervall $I_{j,k}$ ist Vereinigung der disjunkten Intervalle $I_{j+1,2k}$ und $I_{j+1,2k+1}$. Einfaches Nachrechnen ergibt für die charakteristischen Funktionen

$$2^{j/2} \chi_{I_{j,k}} + e_{j,k} = 2^{(j+2)/2} \chi_{I_{j+1,2k}}, \quad 2^{j/2} \chi_{I_{j,k}} - e_{j,k} = 2^{(j+2)/2} \chi_{I_{j+1,2k+1}}.$$

Daher folgt aus $a_{j,k} = 0$ und der zweiten Identität in (1.13)

$$a_{j+1,2k} = a_{j+1,2k+1} = 0.$$

Damit ist der Induktionsschritt von j nach $j+1$ gezeigt.

2. Schritt: Jede Funktion $f \in L^2([0, 1))$ ist Lebesgue-integrierbar. Aus $\int_{I_{j,k}} f(x) dx = 0$ für jedes dyadische Intervall $I_{j,k}$ folgt mit elementaren Argumenten der Integrationstheorie, dass $\int_I f(x) dx = 0$ für jede messbare Menge $I \subset [0, 1)$ gilt. Daraus folgt wiederum $f = 0$. \square

BEMERKUNG 1.17. *Welchen Informationsgehalt hat nun der einzelne "Wavelet-Koeffizient" $b_{j,k} := \langle f, e_{j,k} \rangle$ der Haar-Basis? Wir stellen sofort fest, dass dieser Koeffizient nur vom Verhalten von f im dyadischen Intervall $I_{j,k}$ abhängt. Mit wachsendem k "zoomen" die Koeffizienten also in immer kleinere Bereiche der Funktion f . Mit $x_{j,k}$ bezeichnen wir den Mittelpunkt des Intervalls $I_{j,k}$. Wir betrachten nun die folgenden 2 speziellen Situationen:*

1. Fall: f erfüllt im Intervall $I_{j,k} = [2^{-j}k, 2^{-j}(k+1)]$ die Lipschitz-Bedingung

$$|f(x) - f(x_{j,k})| \leq M|x - x_{j,k}|^\alpha, \quad x \in I_{j,k},$$

mit einem $0 < \alpha \leq 1$. Dann ist wegen $\int_{I_{j,k}} e_{j,k}(x) dx = 0$ der Koeffizient

$$d_{j,k} = \int_{I_{j,k}} f(x)e_{j,k}(x) dx = \int_{I_{j,k}} (f(x) - f(x_{j,k}))e_{j,k}(x) dx,$$

also

$$|d_{j,k}| \leq M2^{j/2} \int_{I_{j,k}} |x - x_{j,k}|^\alpha dx = \frac{2^{-\alpha}M}{\alpha + 1} 2^{-j(\alpha+1/2)}.$$

2. Fall: f erfüllt in den beiden Teilen $[2^{-j}k, x_{j,k})$ und $(x_{j,k}, 2^{-j}(k+1)]$ die Lipschitz-Bedingung der Ordnung $0 < \alpha \leq 1$, hat aber in $x_{j,k}$ einen Sprung der Höhe $2c$; genauer: es existiert ein $a \in \mathbb{R}$ (der Mittelwert der links- und rechtsseitigen Grenzwerte von f in $x_{j,k}$) mit

$$|f(x) - (a + c)| \leq M|x - x_{j,k}|^\alpha \quad \text{für } x \in (x_{j,k}, 2^{-j}(k+1)],$$

$$|f(x) - (a - c)| \leq M|x - x_{j,k}|^\alpha \quad \text{für } x \in [2^{-j}k, x_{j,k}).$$

Dann ist

$$d_{j,k} = \int_{I_{j,k}} f(x)e_{j,k}(x) dx = a \underbrace{\int_{I_{j,k}} e_{j,k}(x) dx}_{=0} - c \underbrace{\int_{I_{j,k}} |e_{j,k}(x)| dx}_{=2^{-j/2}} + R = c2^{-j/2} + R$$

mit

$$\begin{aligned} |R| &\leq \int_{2^{-j}k}^{x_{j,k}} 2^{j/2}|f(x) - (a - c)| dx + \int_{x_{j,k}}^{2^{-j}(k+1)} 2^{j/2}|f(x) - (a + c)| dx \\ &\leq M2^{j/2} \int_{I_{j,k}} |x - x_{j,k}|^\alpha dx \\ &\leq \frac{2^{-\alpha}M}{\alpha + 1} 2^{-j(\alpha+1/2)}. \end{aligned}$$

Man stellt also den Unterschied des asymptotischen Abklingverhaltens von $|d_{j,k}|$ für $j \rightarrow \infty$ zum 1. Fall fest: Die Abnahmerate $2^{-j/2}$ deutet auf die Unstetigkeitsstelle hin.

Die Größe der Wavelet-Koeffizienten $d_{j,k}$ mit wachsendem j , die zu Intervallen $I_{j,k}$ in der Nähe einer festen Stelle $x \in [0, 1)$ gehören, lässt tatsächlich eine Aussage zur "Glattheit" der Funktion f an der Stelle x zu. Wir werden dies im folgenden Kapitel genau untersuchen.

BEMERKUNG 1.18. Die Struktur der Haar-Wavelets ist besonders eindrucksvoll: Wir definieren die Funktion $\psi : \mathbb{R} \rightarrow \mathbb{C}$ durch

$$\psi(x) = \chi_{[0,1/2)}(x) - \chi_{[1/2,1)}(x).$$

Dann gilt für alle $j \in \mathbb{N}_0$ und $0 \leq k \leq 2^j - 1$ die Identität

$$e_{j,k}(x) = 2^{j/2}\psi(2^j x - k) \quad (\text{insbesondere } e_{0,0} = \psi).$$

D.h. alle Basisfunktionen (mit Ausnahme von $e_0 = \chi_{[0,1]}$) werden von einer einzigen Funktion ψ durch Skalierung des Arguments (mit 2^j) und Verschiebung (um $2^{-j}k$) bestimmt. Man nennt die Funktion ψ daher das Mutter-Wavelet.

Die erste Basisfunktion e_0 wird durch die charakteristische Funktion $\phi : \mathbb{R} \rightarrow \mathbb{C}$, $\phi(x) = \chi_{[0,1]}(x)$ definiert. Diese Funktion spielt wegen der beiden Gleichungen

$$(1.14) \quad \phi(x) = \phi(2x) + \phi(2x - 1), \quad \psi(x) = \phi(2x) - \phi(2x - 1)$$

eine ebenso wichtige Rolle. Sie wird manchmal das Vater-Wavelet oder auch die Skalierungsfunktion des Haar-Wavelets genannt. Ihre Bedeutung steckt hauptsächlich in der effizienten Berechnung der Waveletkoeffizienten. Hierauf gehen wir erst später ein.

BEMERKUNG 1.19. Ebenso wie im “diskreten Regime” ist auch der Zusammenhang der Wavelet-Koeffizienten zu den Integralmitteln von f einfach herzuleiten. Wie oben setzen wir

$$\begin{aligned} a_{j,k} &= 2^{j/2} \int_{I_{j,k}} f(x) dx \\ d_{j,k} &= \langle f, e_{j,k} \rangle = 2^{j/2} \left(\int_{I_{j+1,2k}} f(x) dx - \int_{I_{j+1,2k+1}} f(x) dx \right), \end{aligned}$$

Dann gelten einerseits die “Zerlegungs-Relationen”

$$\begin{aligned} a_{j,k} &= 2^{-1/2}(a_{j+1,2k} + a_{j+1,2k+1}), \\ d_{j,k} &= 2^{-1/2}(a_{j+1,2k} - a_{j+1,2k+1}). \end{aligned}$$

Hiermit können die Wavelet-Koeffizienten $d_{j,k}$ für $j < J$ rekursiv aus den Integralmitteln $a_{J,k}$ berechnet werden. Umgekehrt gelten die “Synthese-Relationen”

$$\begin{aligned} a_{j+1,2k} &= 2^{-1/2}(a_{j,k} + d_{j,k}), \\ a_{j+1,2k+1} &= 2^{-1/2}(a_{j,k} - d_{j,k}). \end{aligned}$$

Hiermit lassen sich die Integralmittel zu den kürzeren Intervallen (nämlich $I_{j+1,2k}$ und $I_{j+1,2k+1}$) aus dem Integralmittel $I_{j,k}$ und dem zugehörigen Waveletkoeffizienten berechnen.

BEISPIEL 1.20. Das Signal `nbump3.mat` in der Wavelet-Toolbox von Matlab (siehe Ordner `wavedemo`) wird mit dem Haar-Wavelet analysiert. Wir rufen hierzu die Wavelet-Toolbox auf mit `wavemenu`.

Kontinuierliche Wavelet–Transformation

Bevor wir zur Konstruktion weiterer Wavelet-Basen (etwa von $L^2(\mathbb{R})$) gehen, wollen wir genauer auf den Informationsgehalt der inneren Produkte $\langle f, e_{j,k} \rangle$ wie in Bemerkung 1.17 eingehen. Dazu führen wir den folgenden Begriff ein.

DEFINITION 2.1. *Gegeben sei ein Intervall $I \subset \mathbb{R}$ und eine (Lebesgue-)integrierbare Funktion $f : I \rightarrow \mathbb{C}$.*

a) *Falls für $\nu \in \mathbb{N}_0$ das Integral*

$$M_\nu(f) := \int_I t^\nu f(t) dt$$

existiert, heißt es das ν -te Moment von f .

b) *Falls für ein $L \in \mathbb{N}$ alle Momente $M_k(f)$, $0 \leq k \leq L - 1$, existieren und gleich Null sind, so besitzt f mindestens L verschwindende Momente.*

HILFSSATZ 2.2. *Sei $I = (a, b)$ ein (beschränktes oder unbeschränktes) Intervall. Eine stetige integrierbare Funktion $f : I \rightarrow \mathbb{C}$ besitzt genau dann ein verschwindendes Moment, wenn $f = g'$ mit einer stetig differenzierbaren Funktion $g : I \rightarrow \mathbb{C}$ gilt, die*

$$\lim_{t \rightarrow a} g(t) = \lim_{t \rightarrow b} g(t) = 0$$

erfüllt.

Beweis: Die Stammfunktion $g(t) = \int_a^t f(s) ds$ besitzt die geforderten Eigenschaften. \square

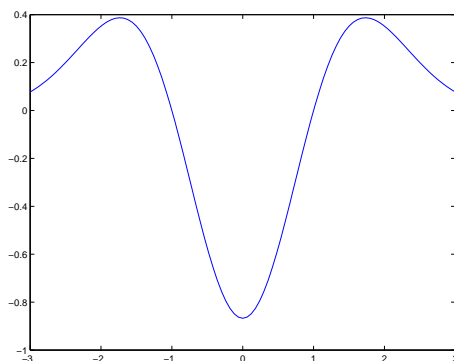
ÜBUNG 2.3. *Zeigen Sie, dass die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$,*

$$f(t) = \frac{2}{\pi^{1/4} \sqrt{3\sigma}} \left(\frac{t^2}{\sigma^2} - 1 \right) \exp\left(\frac{-t^2}{2\sigma^2} \right)$$

zwei verschwindende Momente besitzt. Diese Funktion heißt im Englischen “Mexican hat” (deutsch: Sombrero?), da der Graph bei Rotation um die y -Achse ein entsprechendes Bild liefert.

Eine weitere Möglichkeit, Funktionen auf $I = \mathbb{R}$ mit verschwindendem Moment zu erhalten, wird in der folgenden Aufgabe behandelt.

ÜBUNG 2.4. *Gegeben sei eine integrierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$, für die die ersten L Momente existieren und die $M_0(f) \neq 0$ erfüllt. Weiter seien $c_k \in \mathbb{C}$, $s_k \in \mathbb{R}$,*

BILD 2.1. Mexican hat mit $\sigma = 1$

$1 \leq k \leq K$. Dann gilt: Die Funktion

$$g(t) = \sum_{k=1}^K c_k f(t - s_k)$$

besitzt genau dann L verschwindende Momente, wenn

$$\sum_{k=1}^K c_k s_k^\nu = 0, \quad 0 \leq \nu \leq L - 1,$$

gilt.

BEMERKUNG 2.5. Für integrierbare Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$ (also $f \in L^1(\mathbb{R})$) ist die Fourier-Transformierte

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(t) e^{-it\xi} dt, \quad \xi \in \mathbb{R},$$

definiert. Wegen der Integrierbarkeit von f ist \hat{f} gleichmäßig stetig und beschränkt mit $|\hat{f}(\xi)| \leq \|f\|_1$. Die Eigenschaft $M_0(f) = 0$ ist gleichwertig mit der Bedingung $\hat{f}(0) = 0$.

Im nächsten Kapitel behandeln wir weitere Rechenregeln für die Fourier-Transformation. Z.B. folgt aus der Integrierbarkeit von $t^\nu f(t)$, $0 \leq \nu \leq L - 1$, dass die Fourier-Transformierte \hat{f} mindestens $L - 1$ -mal stetig differenzierbar ist. Die Eigenschaft, L verschwindende Momente zu haben, ist dann äquivalent zu der Beziehung

$$\frac{d^\nu}{d\xi^\nu} \hat{f}(0) = 0, \quad 0 \leq \nu \leq L - 1.$$

Wir kommen nun zur ersten, sehr allgemeinen Definition eines Wavelets.

DEFINITION 2.6. (Wavelet)

Eine Funktion $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ heißt Wavelet, wenn $M_0(\psi) = 0$ gilt.

Der englische Begriff “Wavelet” bezieht sich auf die Form des Grafen von ψ , den man sich als “kleine Welle” vorstellen kann. Wie beim Haar-Wavelet wird die

Stauchung und Verschiebung der Funktion ψ eine wichtige Rolle spielen. Dazu die folgende Kurzschreibweise.

BEZEICHNUNG 2.7. Gegeben sei eine Funktion $\psi : \mathbb{R} \rightarrow \mathbb{C}$. Für $a \in \mathbb{R} \setminus \{0\}$ und $b \in \mathbb{R}$ verwenden wir die Bezeichnung

$$\psi_{a,b} := |a|^{-1/2} \psi \left(\frac{\cdot - b}{a} \right).$$

Ohne die Gefahr von Verwechslungen werden wir später für $j, k \in \mathbb{Z}$ auch

$$\psi_{j,k} := 2^{j/2} \psi(2^j \cdot -k)$$

verwenden. Dies entspricht der Wahl $a = 2^{-j}$ und $b = 2^{-j}k$.

DEFINITION 2.8. Sei $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ ein Wavelet. Die kontinuierliche Wavelet-Transformation (CWT) ist der lineare Operator W_ψ auf $L_2(\mathbb{R})$ mit

$$(W_\psi f)(a, b) = \langle f, \psi_{a,b} \rangle = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \overline{\psi \left(\frac{t-b}{a} \right)} dt, \quad a \neq 0, b \in \mathbb{R}.$$

ÜBUNG 2.9. Man berechne (mit Matlab) die kontinuierliche Wavelet-Transformation $W_\psi f$ zum Mexican-Hat Wavelet ψ und der Funktion

$$f(t) = \begin{cases} 1 - |t|, & t \in [-1, 1], \\ 2/3, & t \in (1, \pi/2), \\ 2 \sin(50t) + 3 \cdot \text{rand}(1), & t \in [\pi/2, 2\pi]. \end{cases}$$

Zeichnen Sie den Graphen der Funktion f . Zum Zeichnen der Wavelet-Transformation $W_\psi f$ verwende man die Darstellung der Funktionswerte auf einem Rechteck mit Achsen $-10 \leq \log_2 a \leq 0$ und $-5 \leq b \leq 10$ durch Grau- oder Farbwerte.

BEMERKUNG 2.10. (i) Für festes $a \neq 0$ ist die Funktion $b \mapsto (W_\psi f)(a, b)$ die Faltung von f mit der Funktion

$$\tilde{\psi}_a(x) = |a|^{-1/2} \overline{\psi \left(-\frac{x}{a} \right)},$$

d.h.

$$(W_\psi f)(a, b) = \langle f, \psi_{a,b} \rangle = f * \tilde{\psi}_a(b) = \int_{-\infty}^{\infty} f(t) \tilde{\psi}_a(b-t) dt.$$

Für die L_1 - und die L_2 -norm von $\tilde{\psi}_a$ gilt

$$\|\tilde{\psi}_a\|_2 = \|\psi\|_2, \quad \|\tilde{\psi}_a\|_1 = |a|^{1/2} \|\psi\|_1.$$

Die Cauchy-Schwarz Ungleichung liefert die punktweise Abschätzung

$$(2.1) \quad |(W_\psi f)(a, b)| \leq \|\tilde{\psi}_a\|_2 \|f\|_2.$$

Weiterhin gilt mit den üblichen Abschätzungen für die Faltung von Funktionen

$$\|(W_\psi f)(a, \cdot)\|_2 = \|f * \tilde{\psi}_a\|_2 \leq \|\tilde{\psi}_a\|_1 \|f\|_2 = |a|^{1/2} \|\psi\|_1 \|f\|_2.$$

Insbesondere ist also die Funktion $(W_\psi f)(a, \cdot) : \mathbb{R} \rightarrow \mathbb{C}$ quadrat-integrierbar. Zur Integrierbarkeit über den gesamten Parameterbereich $(a, b) \in \mathbb{R}^* \times \mathbb{R}$ werden wir später eine genaue Aussage kennen lernen.

(ii) Der Wert $(W_\psi f)(a, b)$ bestimmt (im Fall $\|\psi\|_2 = 1$) die Korrelation von f und $\tilde{\psi}_a$.

DEFINITION 2.11. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ heißt Hölder- (oder Lipschitz-) stetig zum Exponenten $0 < \alpha \leq 1$, geschrieben $f \in \text{Lip}^\alpha(\mathbb{R})$, wenn $K > 0$ existiert mit

$$(2.2) \quad |f(x) - f(y)| \leq K \cdot |x - y|^\alpha$$

für alle $x, y \in \mathbb{R}$. f heißt Hölder-stetig zum Exponenten α an der Stelle x_0 , falls $K > 0$ existiert mit

$$(2.3) \quad |f(x) - f(x_0)| \leq K \cdot |x - x_0|^\alpha$$

für alle $x \in \mathbb{R}$.

BEISPIEL 2.12. a) $f : \mathbb{R} \rightarrow \mathbb{C}$ mit $f(x) = \sqrt{|x|}$ ist in jedem Punkt $x_0 \neq 0$ Hölder-stetig zum Exponenten 1, in $x_0 = 0$ jedoch nur Hölder-stetig zum Exponenten $\frac{1}{2}$.

b) $f : \mathbb{R} \rightarrow \mathbb{C}$ differenzierbar und f' beschränkt impliziert $f \in H^1(\mathbb{R})$ (Beweis mit MWS).

Die Einschränkung des Hölder-Exponenten α auf $(0, 1]$ in der obigen Definition ist sinnvoll, da aus der Bedingung 2.2 mit $\alpha > 1$ sofort die Differenzierbarkeit von f sowie die Beziehung $f' \equiv 0$ folgt, f also eine konstante Funktion ist. Um trotzdem Hölder-Klassen zu beliebigem Exponenten $\alpha > 0$ zu betrachten, verwendet man Ableitungen von f .

DEFINITION 2.13. Für $\alpha > 0$, $\alpha = r + \beta$ mit $r \in \mathbb{N}_0$ und $\beta \in (0, 1]$, definieren wir

$$\text{Lip}^\alpha(\mathbb{R}) := \{f \in C^r(\mathbb{R}); f^{(r)} \in \text{Lip}^\beta(\mathbb{R})\}.$$

f heißt Hölder-stetig zum Exponenten α an der Stelle x_0 , falls f in einer Umgebung von x_0 r -mal differenzierbar ist und $f^{(r)}$ an der Stelle x_0 Hölder-stetig zum Exponenten β ist.

Sowohl für die globale wie auch für die lokale Hölder-Stetigkeit liefert die kontinuierliche Wavelet-Transformation (fast) scharfe Charakterisierungen.

SATZ 2.14. (Globale Hölder-Stetigkeit) Gegeben sei ein Wavelet ψ mit $r + 1$ verschwindenden Momenten sowie beschränktem Moment $M_{r+1}(\psi)$. Weiter sei $f \in L^2(\mathbb{R})$ beschränkt und stetig.

(i) Falls f Hölder-stetig zum Exponenten $r < \alpha \leq r + 1$ ist und $f^{(r)}$ beschränkt ist, so existiert $K > 0$ mit

$$(2.4) \quad |W_\psi f(a, b)| \leq K \cdot |a|^{\alpha + \frac{1}{2}}, \quad a, b \in \mathbb{R}, a \neq 0.$$

(ii) Die Umkehrung gilt für $r < \alpha < r + 1$ und ψ mit kompaktem Träger: Falls eine Konstante $K > 0$ so existiert, dass (2.4) gilt, so ist f Hölder-stetig zum Exponenten α .

Die Aussage über die lokale Hölder-Stetigkeit wird etwas komplizierter.

SATZ 2.15. Gegeben sei ein Wavelet ψ mit kompaktem Träger und $r + 1$ verschwindenden Momenten (sowie beschränktem Moment $M_{r+1}(\psi)$). Weiter sei $f \in L^2(\mathbb{R})$ sowie $x_0 \in \mathbb{R}$.

- (i) Falls f Hölder-stetig zum Exponenten $r < \alpha = r + \beta \leq r + 1$ an der Stelle x_0 ist, so existieren $a_0 > 0$ und $K > 0$ so, dass für alle $a, b \in \mathbb{R}$ mit $0 < |a| < a_0$ gilt

$$(2.5) \quad |W_\psi f(a, b)| \leq K \cdot |a|^{r+\frac{1}{2}} (|a|^\beta + |b - x_0|^\beta).$$

- (ii) Es gilt die folgende Umkehrung: Falls es Konstanten $a_0, K > 0$ sowie $\gamma > r$ so gibt, dass

$$(2.6) \quad |W_\psi f(a, b)| \leq K \cdot |a|^{\gamma+\frac{1}{2}},$$

für alle $a, b \in \mathbb{R}$ mit $0 < |a| < a_0$ gilt, und falls darüber hinaus $0 < \beta < 1$ existiert mit

$$(2.7) \quad |W_\psi f(a, b)| \leq K \cdot |a|^{r+\frac{1}{2}} \left(|a|^\beta + \frac{|b - x_0|^\beta}{1 + |\log |b - x_0||} \right)$$

für alle $a, b \in \mathbb{R}$, $0 < |a| < a_0$, so ist f Hölder-stetig zum Exponenten $\alpha = r + \beta$ an der Stelle x_0 .

Die Beweise der Umkehrungen (ii) der obigen Sätze können wir erst nach weiteren Vorbereitungen (insb. Hilfsmittel aus der Fourier-Transformation) in Kapitel 3 angeben. Ohne weitere Vorbereitung können aber die Teile (i) schon mit einfachen Hilfsmitteln der Analysis bewiesen werden.

Beweis. Zu $x_0 \in \mathbb{R}$ wählen wir eine Umgebung U so, dass f in U mindestens r -mal differenzierbar und die r -te Ableitung Hölder-stetig zum Exponenten β ist. Unter den Voraussetzungen von Satz 2.14 kann $U = \mathbb{R}$ gewählt werden, ansonsten werde $U = (x_0 - \delta, x_0 + \delta)$ mit $\delta > 0$ gewählt. Die Taylorentwicklung von f mit Entwicklungspunkt $b \in U$ lautet

$$f(x) = \sum_{j=0}^{r-1} \frac{f^{(j)}(b)}{j!} (x-b)^j + \int_b^x \frac{f^{(r)}(t)}{(r-1)!} (x-t)^{r-1} dt, \quad x \in U.$$

Die Voraussetzungen an ψ liefern für beliebige $a, b \in \mathbb{R}$, $a \neq 0$, die Identität

$$(2.8) \quad \int_{\mathbb{R}} \psi_{a,b}(x) (x-b)^j dx = 0, \quad 0 \leq j \leq r.$$

Hieraus folgt insbesondere für $j = r$ und $U = \mathbb{R}$

$$(2.9) \quad \int_{\mathbb{R}} \left(\int_b^x \frac{f^{(r)}(t)}{(r-1)!} (x-t)^{r-1} dt \right) \psi_{a,b}(x) dx = 0.$$

Dann erhalten wir weiter für $U = \mathbb{R}$ aus den Voraussetzungen von Satz 2.14(i)

$$\begin{aligned} |W_\psi f(a, b)| &= \left| \int_{\mathbb{R}} f(x) \overline{\psi_{a,b}(x)} dx \right| \\ &= \left| \int_{\mathbb{R}} \left(\int_b^x \frac{f^{(r)}(t) - f^{(r)}(b)}{(r-1)!} (x-t)^{r-1} dt \right) \overline{\psi_{a,b}(x)} dx \right| \\ &\leq K |a|^{-1/2} \int_{\mathbb{R}} \underbrace{\left(\int_b^x \frac{|t-b|^\beta}{(r-1)!} (x-t)^{r-1} dt \right)}_{\leq |x-b|^{\beta+r}/(r-1)!} \left| \psi \left(\frac{x-b}{a} \right) \right| dx \\ &= \tilde{K} |a|^{\alpha+\frac{1}{2}} \int_{\mathbb{R}} |y|^\alpha |\psi(y)| dy. \end{aligned}$$

Die Existenz von $M_{r+1}(\psi)$ sichert zu, dass das letzte Integral endlich ist. Die Konstante K in (2.4) hängt also nur von r , ψ und der Lipschitz-Konstante K der r -ten Ableitung von f ab. Damit ist Satz 2.14(i) bewiesen.

In Satz 2.15 wird vorausgesetzt, dass ψ kompakten Träger besitzt. Also existiert $a_0 > 0$ so, dass für alle $|a| \leq a_0$ und $b \in [x_0 - \delta/2, x_0 + \delta/2]$ die Beziehung $\text{supp } \psi_{a,b} \subset U$ gilt. Wegen $0 < \beta \leq 1$ gilt die allgemeine Ungleichung

$$|x - y|^\beta \leq |x - z|^\beta + |z - y|^\beta.$$

Für alle $|a| \leq a_0$ und $b \in [x_0 - \delta/2, x_0 + \delta/2]$ folgt dann wie beim vorherigen Beweisteil

$$\begin{aligned} |W_\psi f(a, b)| &= \left| \int_U f(x) \overline{\psi_{a,b}(x)} dx \right| \\ &= \left| \int_U \left(\int_b^x \frac{f^{(r)}(t) - f^{(r)}(x_0)}{(r-1)!} (x-t)^{r-1} dt \right) \overline{\psi_{a,b}(x)} dx \right| \\ &\leq K |a|^{-1/2} \int_{\mathbb{R}} \underbrace{\left(\int_b^x \frac{|t-x_0|^\beta}{(r-1)!} (x-t)^{r-1} dt \right)}_{\leq (|x-b|^\beta + |b-x_0|^\beta) |x-b|^{r/(r-1)!}} \left| \psi \left(\frac{x-b}{a} \right) \right| dx \\ &= \tilde{K} \left(|a|^{\alpha+\frac{1}{2}} \int_{\mathbb{R}} |y|^\alpha |\psi(y)| dy + |b-x_0|^\beta |a|^{r+\frac{1}{2}} \int_{\mathbb{R}} |y|^r |\psi(y)| dy \right). \end{aligned}$$

Hieraus folgt die Behauptung von Satz 2.15(i). \square