Scientific Computing

Cordula Bernert

Inhaltsverzeichnis

Literaturverzeichnis

[1] Christian Blatter: Wavelets - Eine Einführung, Vieweg Verlag, 2. Auflage 2003

[2] Louis, Maaß, Rieder: Wavelets, BG Teubner, Stuttgart 1998

[3] W. Bäni: Wavelets.Eine Einführung für Ingenieure, Oldenbourg Verlag, 2. Auflage 2005

[4] Janusz Wawrzynowicz: Untersuchung von waveletbasierten Bildkompressionsverfahren unter Anwendung von verschiedenen Wavelets und Codierungen, Diplomarbeit, Hochschule Mittweida 1999

[5] Phong Dieu Nguyen: Beitrag zur Diagnostik der Verzahnungen in Getrieben mittels Zeit-Frequenz-Analyse, Dissertation, TU Chemnitz, Fakultät für Maschinenbau und Verfahrenstechnik, Institut für Mechanik 2002

[6] Metin Akay (Editor): Time frequency and Wavelets in Biomedical Signal Processing, IEEE Press 1998

[5] Sweldens, Piessens; SIAM J. Numer. Anal., 31(4), 1994

[6] I. Daubechies: Ten lectures on Wavelets, CBMS-NSF Regional Conference, Series in Applied Mathematics, SIAM 1992

1 Introduction

1.1 Basics of Approximating Functions

Let f be a mapping from $\mathbb{R}$ to $\mathbb{C}$, i.e. $f : \mathbb{R} \to \mathbb{C}$, $\{e_\alpha(t);\ \alpha \in I;\ t \in \mathbb{R}\}$ be a set of basic functions. We call it a "family" of basis functions. I is a discrete or continuous set of indices.
We look for a representation of f by the functions e_α:

$$\begin{aligned} f(t) &= \sum_{\alpha \in I} c_\alpha e_\alpha(t) \qquad (1.1) \quad \textit{or} \\ f(t) &= \int_I c_\alpha e_\alpha(t) d\alpha \quad (1.2). \end{aligned}$$

$\{e_\alpha(t)\}$ includes so many basis functions that the representations (1.1) and (1.2) are unique.

Analysis with respect to the family $\{e_\alpha(t)\}$:

$$f \stackrel{\{e_\alpha\}}{\rightarrow} (c_\alpha), \quad \alpha \in I$$

Synthesis (inverse operation):

$$(c_\alpha) \stackrel{\alpha \in I}{\rightarrow} f$$

Example 1.1 $\{e_\alpha(t)\} = \{(t-b)^\alpha;\ b \in \mathbb{R},\ \alpha = 0, 1, 2, 3, ...\}$:
Assumption: Let f be arbitrarily often differentiable $\curvearrowright$ f can be represented by a Taylor Series.

$$\begin{aligned} c_\alpha &= \frac{f^{(\alpha)}(b)}{\alpha!} \\ f(t) &= \sum_{\alpha=0}^{\infty} c_\alpha (t-b)^\alpha \end{aligned}$$

Example 1.2 *Let f be periodic with the period 2π, $f \in \mathbb{L}_2(-\pi, \pi)$.
Then $\{e_\alpha(t)\} = \{e^{i\alpha t};\ \alpha \in \mathbb{Z}\}$ is an orthonormal system (ONS) in $\mathbb{L}_2(-\pi, \pi)$ and so*

given

$$\begin{aligned} c_\alpha &= (f, e_\alpha) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-i\alpha t}dt \\ f(t) &= \sum_{\alpha=-\infty}^{\infty} c_\alpha e^{i\alpha t}, \end{aligned}$$

we have the Fourier Series of the function $f(t)$ which converges to $f(t)$ with respect to the $\mathbb{L}_2-$metric.

Example 1.3 *Let $f \in \mathbb{L}_2(\mathbb{R})$. But the family $\{e_\alpha(t)\} = \{e^{i\alpha t};\ \alpha \in \mathbb{R}\}$ is not an ONS in $\mathbb{L}_2(\mathbb{R})$!! However we can develop*

$$\begin{aligned} \widehat{f}(\alpha) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\alpha t}dt \\ f(t) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \widehat{f}(\alpha)e^{-i\alpha t}d\alpha. \end{aligned}$$

$\widehat{f}(\alpha)$ is called the Fourier Transform of $f(t)$, which corresponds to the complex amplitude for the frequency α of the signal $f(t)$.

1.2 Problems in the Analysis and Synthesis of Functions

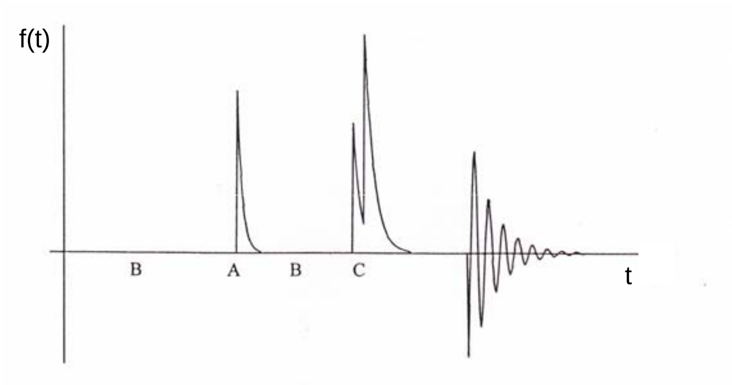
1. During an experiment we get vectors of measured data only, not continuous data.
   $\Longrightarrow$ We need a complete discretisation in practice
   * for the basic functions and
   * for the space of variable t.

   $\Longrightarrow$ the values of all functions are considered at discrete points
   $t = k\tau \;\; \tau > 0;\; k \in \mathbb{Z}$.

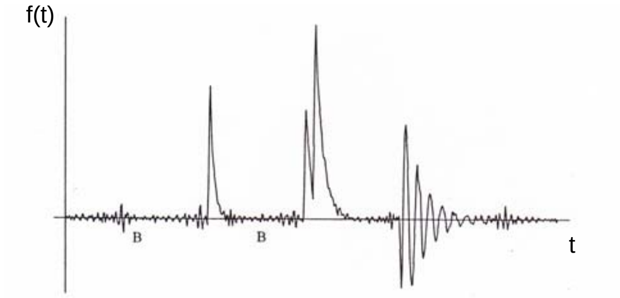
2. The goal of approximation by basis functions is to find a representation which is
   * uniform and
   * compressed as much as possible.

   Application example: image compression.

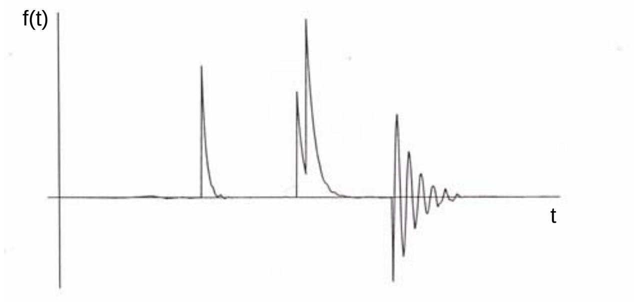
Example 1.4 *Synthesis of the following function*



by Fourier Transform and by Wavelet Transform [3] *:*



Reconstruction by the 100 biggest Fourier coefficients (by modulus)



Reconstruction by the 75 biggest Wavelet coefficients (by modulus)

$\Longrightarrow$ ***Limitations of Fourier Transform Theory:***
**The values of $f(t)$ over the whole domain of t must be reconstructed.*
** There is no localisation of extrema or jump discontinuities (saltus) on the $t-$axis possible.*

Reasons:
On the one hand $y_k = f(\frac{2\pi k}{N}), \quad 0 \leq k < N$ is an exactly located value on the time axis, but on the other hand there are Fourier coefficients c_k which include information from the full axis. That's why in our example we need a lot of coefficients for reconstructing the extinction in one part of the domain and spikes in the other parts of the domain.

We see one manifestation of this problem clearly near discontiuities and we also find residual waviness across the full domain.
$\Longrightarrow$ *bad compression rate because you can't omit something*
$\Longrightarrow$ *bad quality of synthesis*

$\Longrightarrow$ ***Selection of another family of basis functions*** *with the following properties:*

(F1) Representation (1.1) and (1.2) are suitable for a large class of functions. The calculation for analysis and synthesis should be numerically stable and fast.
(F2) The basis functions should be clearly located in time. In the best case they have a bounded domain.
(F3) The transforms of the basis functions should be clearly located in the frequency domain, too.
(F4) The basis functions should form an ONS.

Later we see, by application of the Heisenberg uncertainty principle, that (F2) is inconsistent with (F3). That's why we are looking for a compromize: if possible, local information both about f, and the transform of f should be readily identifiable. (F4) is required for uniqueness.
For avoiding the dilemma of the Heisenberg uncertainty principle there are two possibilities: the windowed Fourier Transform and the Wavelet Transform.

1.3 Short-Time or Windowed Fourier Transform

- Choose a windowing function $g: \quad \mathbb{R} \to \mathbb{R}_{\geq 0}$ such that it is "concentrated with mass one around $t = 0$". In order to achieve this goal we use, for example, functions with a compact domain or functions with a distinct maximum at $t = 0$.

- The window g will be shifted s units to the right:

$$g_s(t) = g(t-s); \quad s > 0,$$

  in order to scan the full $t-$axis.

- The best known example of such a windowing function is the Gaussian distribution, with expectancy $= 0$ and; variance $= \sigma$:

$$g(t) = N(0,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{t^2}{2\sigma^2}), \quad \sigma = const.$$

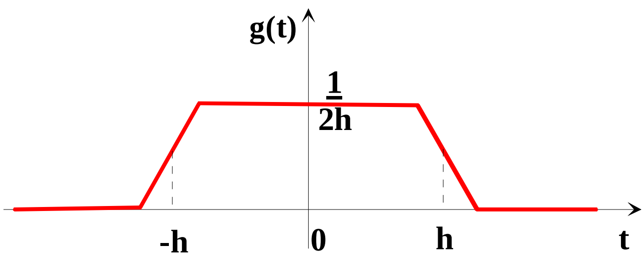
  It leads us to the GABOR Transform (Nobel Prize in Physics 1970).

- Then transforms calculated for example by the basis functions $e^{i\alpha t}$ are the following:

$$
\begin{aligned}
Gf(\alpha, s) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) g(t-s) e^{-i\alpha t} dt \\
Gf &: \mathbb{R} \times \mathbb{R} \to \mathbb{C}
\end{aligned}
$$

- The information about f is very redundant in Gf . $\curvearrowright$ There exist several inversion formulae (GABOR, CALDERON).

Interpretation: We choose $g(t)$ in the following way:



This means that the complex function $Gf(\alpha, s)$ determines which frequencies appear in f inside of the time interval $[s-h; s+h]$ and what is their strength in the signal. If $e^{i\alpha t}$ occurs in the analysed interval $\Longrightarrow$ $|Gf(\alpha, s)|$ is a nonzero number.

Advantage: better time localisation by sliding the window along the time axis

Disadvantage: constant width of the window, because

- you can't locate high-frequency vibrations which occur only in a part of the analysed interval.

- the window is too narrow for detecting a full period of a low-frequency vibration.

$\Longrightarrow$ We expect an improvement if the width of the window varies with α.

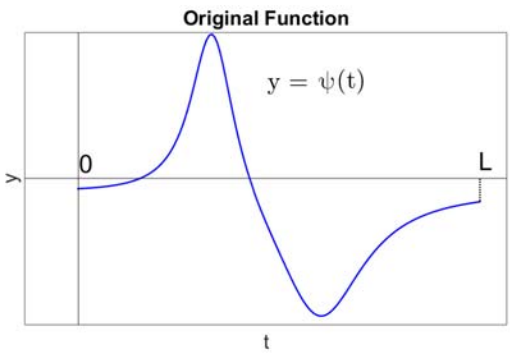
$\Longrightarrow$ The analysing function system has to be changed.

1.4 Wavelet Transform (WT)

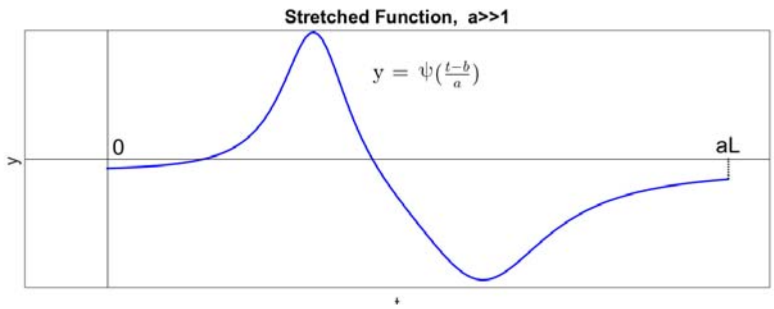
- First we choose an analysing function $\psi(t)$, the so called "mother-wavelet". The mother wavelet is the starting point for the family of derived basis functions, the so called "wavelet functions".
- Wavelet functions are dilated (stretched) and/or shifted copies of the mother wavelets:

$$\begin{aligned} \psi_{ab}(t) &= \frac{1}{|a|^{0.5}}\psi\left(\frac{t-b}{a}\right) \\ \psi &: \mathbb{R} \to \mathbb{C} \\ (a,b) &\in \mathbb{R}^* \times \mathbb{R} = (\mathbb{R}\backslash\{0\}) \times \mathbb{R}. \end{aligned}$$

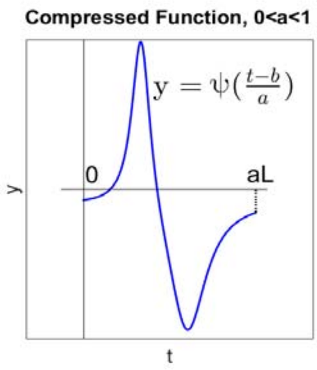
- a : scale parameter, b : shift parameter
  The factor $\frac{1}{|a|^{0.5}}$ normalizes the wavelet functions.
- For example let $\psi(t)$ be the following function :



- With $|a| >> 1$ you will get a wide window for researching slow and low-frequency processes.



With $|a| << 1$ you will get a narrow window for researching fast and high-frequency processes.



- The wavelet transform

$$Wf(a,b) = (f, \psi_{\alpha b}) = \frac{1}{\sqrt{c_\psi}} \int_{-\infty}^{\infty} f(t)\overline{\psi_{ab}(t)}dt = \frac{1}{\sqrt{c_\psi}|a|^{0.5}} \int_{-\infty}^{\infty} f(t)\overline{\psi\left(\frac{t-b}{a}\right)}dt$$

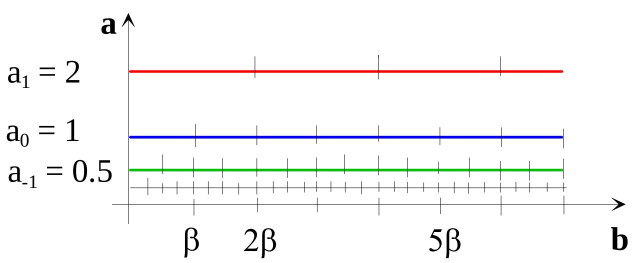
delivers data sets $\{Wf(a,b) \mid (a,b) \in \mathbb{R}^* \times \mathbb{R}\}$ wich have a high redundancy. c_ψ is a specific constant parameter of the mother wavelet.

- $\exists$ inversion formula

$$f = \frac{1}{\sqrt{c_\psi}} \int_{\mathbb{R}^* \times \mathbb{R}} Wf(a,b)\psi_{ab}(t)\frac{dadb}{a^2}.$$

- For calculation we need a discretisation of the index set $\mathbb{R}^* \times \mathbb{R}$ which conforms to the formula above. In the literature you find the following useful definition with a zoom step $\sigma > 1$

$$\begin{aligned} a_r &= \sigma^r \quad with\ \sigma = 2,\ r \in \mathbb{Z} \\ b_k &= ka_r\beta = k\sigma^r\beta \quad with\ \beta > 0,\ k \in \mathbb{Z}. \end{aligned}$$



Thus the scale of discretisation adapts to the width of the dilated wavelet functions.

$\Longrightarrow$ The corresponding wavelet functions are self-similar.
$\Longrightarrow$ Multi Scale Analysis
$\Longrightarrow$ Fast Wavelet Transform

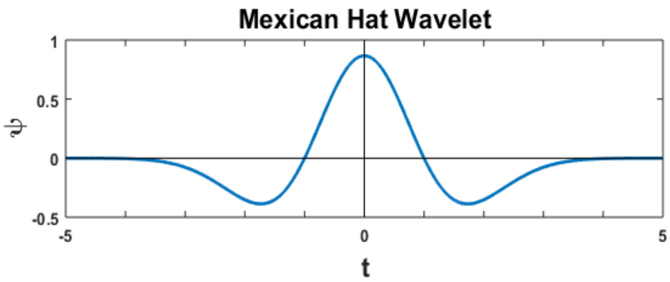
$\Longrightarrow$ Further wavelets can be designed such that:

- they have a compact domain,
- they are orthonormal and
- they allow fast numerical algorithms.

The theory of wavelets was developed in the $1980'$s and $1990'$s.
Examples for well-known mother wavelets are :

- the HAAR-wavelet,
- the Mexikan hat wavelet
- the MEYER-wavelet
- the DAUBECHIES-wavelet
- the BATTLE-LEMARIE-wavelet.

In the picture you see the Mexikan Hat mother wavelet.



The goal of our course is to illuminate the mathematical background of the wavelet transform as a foundation for easy application of the technique. Requirements are:

- basics of functional analysis,
- theory of the Fourier transform and
- properties of the HAAR mother wavelet as an easy illustration for the general theory.

2 Basics of Fourier Analysis

2.1 Fourier Series

We consider the space
$\mathbb{L}_2(\mathbb{R}/2\pi) = \{f : \mathbb{R} \to \mathbb{C},\ measurable,\ 2\pi - periodic,\ \frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2 dt < \infty\}$

with the inner product $(f,g) = \frac{1}{2\pi}\int\limits_0^{2\pi} f(t)\overline{g(t)}dt$, and

the induced norm $\|f\|^2 = \frac{1}{2\pi}\int\limits_0^{2\pi} |f(t)|^2\, dt.$

Attention: This is a set of equivalent classes of functions. Functions of one class may be different on a set of measure zero!
With respect to the metric function $d(f,g) = \|f-g\|$ this space is complete and is therefore a Hilbert space.
We use the complete ONS:

$$e_k(t) = \frac{1}{\sqrt{2\pi}} \exp(ikt) \qquad k = 0, \pm 1, \pm 2, ...$$

and get the (general) Fourier series

$$\begin{aligned} f(t) &= \sum_{k=-\infty}^{\infty} c_k e_k(t)\ , \quad \text{with} \\ c_k &= (f, e_k) = \widehat{f}(k) \end{aligned}$$

The following theorems say something about the character of convergence of these Fourier series.

Theorem 2.1 *RIEMANN-LEBESGUE-Lemma:* $\lim_{k\to\pm\infty} |c_k| = 0$.

Theorem 2.2 *PARSEVAL Formula:* $\sum\limits_{k=-\infty}^{\infty} \widehat{f}(k)\overline{\widehat{g}(k)} = (f,g) \quad \forall f,g \in \mathbb{L}_2(\mathbb{R}/2\pi)$.

Especially we have:

$$\sum_{k=-\infty}^{\infty} \left|\widehat{f}(k)\right|^2 = \sum_{k=-\infty}^{\infty} |c_k|^2 = \sum_{k=-\infty}^{\infty} |(f,e_i)|^2 = \|f\|^2 .$$

Theorem 2.3 *The (general) Fourier Series of a function $f \in \mathbb{L}_2(\mathbb{R}/2\pi)$ converges to f (with respect to the $\mathbb{L}_2-$metric).*

Theorem 2.4 *CARLESON 1966*

$$s_N(t) = \sum_{k=-N}^{N} c_k e_k \overset{almost\ everywhere}{\underset{N\to\infty}{\longrightarrow}} f \in \mathbb{L}_2(\mathbb{R}/2\pi)$$

Definition 2.1 $V(f) = \sup\limits_T \sum\limits_{k=1}^{n} |f(t_k) - f(t_{k-1})|$ *is called variation over all segmentations* $T: \quad 0 \le t_0 < t_1 < ... < t_n = 2\pi$ *of the interval* $[0, 2\pi]$.

Theorem 2.5 $f \in \mathbb{R}/2\pi, \quad continuos, \quad V(f) < \infty \quad \Longrightarrow \quad s_N(t) \overset{glm.}{\underset{N\to\infty}{\longrightarrow}} f$

Theorem 2.6 $f \in \mathbb{R}/2\pi,\ r \ge 0,\ f^{(r)}$ *continuos*, $\ V(f^{(r)}) = V < \infty$
$\Longrightarrow \quad |c_k| \le \dfrac{V}{2\pi |k|^{r+1}} \quad \forall k \ne 0$

This means that the smoother the function, the faster the convergence of the absolute value of the coefficients c_k to zero as $|k| \to \infty$.

Theorem 2.7 *If* $c_k = O\left(\dfrac{1}{|k|^{r+1+\varepsilon}}\right)$, $|k| \to \infty$, *for certain* $\varepsilon > 0$
$\Longrightarrow \quad f(t) = \sum\limits_{k=-\infty}^{\infty} c_k e_k$ *is at least $r-$times continuously differentiable.*

Theorem 2.8 $f \in \mathbb{L}_2(\mathbb{R}/L) \quad \curvearrowright$

$$f(t) \sim \sum_{k=-\infty}^{\infty} c_k e_k(t\frac{2\pi}{L}) \quad with$$

$$c_k = \frac{1}{\sqrt{L}} \int_0^T f(t) \exp\left(-ik\frac{2\pi}{L}t\right) dt = \left(f, e_k(t\frac{2\pi}{L})\right),$$

and

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{1}{L} \int_0^T |f(t)|^2 \, dt.$$

2.2 Fourier Transform

Let f be a function with $f \in \mathbb{L}_1(\mathbb{R})$, i.e. $\int_{-\infty}^{\infty} |f(t)|\, dt = \|f\|_1 = I < \infty$.

Definition 2.2 *Fourier Transform (FT) of $f(t)$:*

$$\widehat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt, \quad \omega \in \mathbb{R}$$

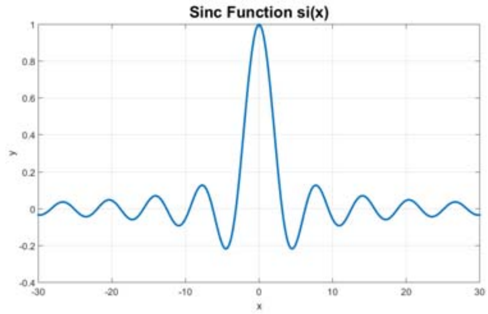
Theorem 2.9 $f \in \mathbb{L}_1(\mathbb{R}) \quad \Longrightarrow \quad \widehat{f}(\omega)$ *is continuous;* $\lim_{|\omega| \to \infty} |\widehat{f}(\omega)| = 0$

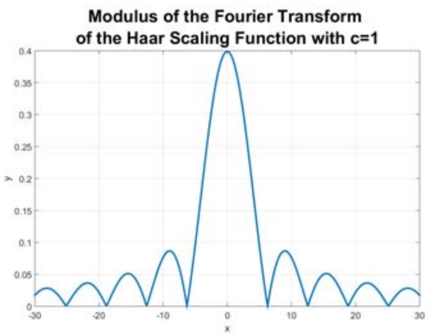
The Fourier Transform is necessary for analysing nonperiodic signals because for the synthesis of a nonperiodic function we need all frequencies, not only the multiples of the first harmonic. In the literature you can find the Fourier transform in different forms. Especially the factors in front of the formulas are different, in a manner analogous to the Fourier series.
Given $\omega \in \mathbb{R}$. Then $\widehat{f}(\omega)$ is the complex amplitude of the vibration e_ω in f.

Example 2.1 *The HAAR scaling function is the following function with parameter $c = 1$:*

$$f(t) = \boldsymbol{\phi}_c(t) = \begin{cases} 1 & 0 \le t \le c \\ 0 & otherwise \end{cases}$$

$$\begin{aligned}
\widehat{\boldsymbol{\phi}_c}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \boldsymbol{\phi}_c(t) e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} \int_{0}^{c} e^{-i\omega t} dt \\
&= -\frac{1}{\sqrt{2\pi} i\omega} \left[e^{-i\omega c} - 1 \right] \\
&= \frac{1}{\sqrt{2\pi} i\omega} e^{-ic\omega/2} \left[e^{ic\omega/2} - e^{-ic\omega/2} \right] \\
&= e^{-ic\omega/2} \frac{2}{\sqrt{2\pi}\omega} \frac{e^{ic\omega/2} - e^{-ic\omega/2}}{2i} \\
&= e^{-ic\omega/2} \frac{c}{\sqrt{2\pi}} \frac{2}{\omega c} \sin(\frac{\omega c}{2}) \\
&= \frac{c}{\sqrt{2\pi}} e^{-ic\omega/2} si(\frac{\omega c}{2})
\end{aligned}$$





Example 2.2 *GAUSSian distribution curve (bell-shaped curve):*

$$\begin{aligned} f(t) &= N_{0,1}(t) = \frac{1}{\sqrt{2\pi}} \exp(\frac{-t^2}{2}) \\ \widehat{f}(\omega) &= \frac{1}{\sqrt{2\pi}} \exp(\frac{-\omega^2}{2}) \end{aligned}$$

The FT of the GAUSSian distribution curve is again the GAUSSian distribution curve.

If $f \in \mathbb{L}_2(\mathbb{R})$, then there must not exist the FT of f because the functions $\{e_\omega\}$ are not a basis in $\mathbb{L}_2(\mathbb{R})$:
If $e_\omega \notin \mathbb{L}_2(\mathbb{R}) \quad \curvearrowright \quad \widehat{f}(\omega) \neq (f,\, e_\omega)$.
On the other hand if you consider $\overline{\mathbb{X}} = \overline{\mathbb{L}_1 \cap \mathbb{L}_2} = \mathbb{L}_2$, you can expand the FT to the full space $\mathbb{L}_2(\mathbb{R})$.

Theorem 2.10 *PARSEVAL-PLANCHEREL Formula /1/*
$f,g \in \mathbb{L}_2(\mathbb{R}) \quad \Longrightarrow \quad$ *The FT is an isometry, i.e.* $(\widehat{f}, \widehat{g}) = (f,g)$,

$$\textit{especially } \left\| \widehat{f} \right\|_{\mathbb{L}_2}^2 = \|f\|_{\mathbb{L}_2}^2 .$$

If your operator is an isometry then the invers operator is equal to the adjoint operator, i.e.: $(FT)^{-1} = (FT)^* \quad \curvearrowright$

Theorem 2.11 $f \in \mathbb{L}_1(\mathbb{R}),\ \widehat{f} \in \mathbb{L}_1(\mathbb{R})$ *(irreducible conditions!)* $\curvearrowright$

$$f(t) = (FT)^{-1} \widehat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(\omega) e^{i\omega t} d\omega \quad \textit{almost everywhere,}$$

especially in regions of $f(t)$ where $f(t)$ is continuous. This means that the original signal f is a linear combination of vibrations of all possible frequencies ω, which are represented by the amplitude $\widehat{f}(\omega)$

Properties of the FT

1) Translation	$T_h f(t) = f(t-h)$	
	$\widehat{(T_h f)}(\omega) = e^{-i\omega h}\widehat{f}(\omega)$	$(R1)$
	$\widehat{(e^{ith} f)}(\omega) = \widehat{f}(\omega - h)$	$(R2)$
2) Dilatation	$D_a f(t) = f(\frac{t}{a})$	
	$\widehat{(D_a f)}(\omega) = \lvert a\rvert D_{\frac{1}{a}}\widehat{f}(\omega)$	$(R3)$
3) Convolution	$(f * g)(x) = \int_{-\infty}^{\infty} f(x-t)g(t)dt$	
	$\widehat{(f * g)}(\omega) = \sqrt{2\pi}\cdot\widehat{f}(\omega)\widehat{g}(\omega)$	
4) Differentiation	$f(t), f'(t) \in \mathbb{L}_1(\mathbb{R})$	
	$\widehat{f'}(\omega) = i\omega\widehat{f}(\omega)$	$(R4)$
	$\widehat{(tf)}(\omega) = i\left(\widehat{f}\right)'(\omega)$	$(R5)$
5)	FT is linear.	
6) If f is a real function:	$\overline{\widehat{f}(\omega)} = \widehat{f}(-\omega)$	$(R6)$

Theorem 2.12 *$f \in \mathbb{L}_1(\mathbb{R})$; $\int_{-\infty}^{\infty} |t|^r |f(t)| dt < \infty$ for $r \geq 1$*
$\Longrightarrow$ *$\widehat{f}(\omega)$ is at least $r-$times countinuously differentiable and*

$$\left(\widehat{f}\right)^{(r)}(\omega) = (-i)^r\,\widehat{(t^r f)}(\omega).$$

2.3 HEISENBERG Uncertainty Principle

From property (R3) of the Fourier transform we get:

$$\widehat{(D_a f)}(\omega) = \widehat{f(\frac{t}{a})} = |a|\widehat{f}(a\omega).$$

That means:
Compression of the function f in the original domain corresponds to a stretching of $\widehat{f}$ in the transformed domain.This is accompanied by a corresponding reduction in the absolut values of $\widehat{f}$ (consider for example $a = 0.5$). So you can qualitativly see that a time signal f and its Fourier transform can't be well located in a little domain of the $t-$axis respectively the $\omega-$axis at the same time. Now we are looking for a quantification of this situation:

Theorem 2.13 *Given* $\psi \in \mathbb{L}_2(\mathbb{R}) \quad \curvearrowright \quad \|t\psi\| \cdot \left\|\omega\widehat{\psi}\right\| \geq \frac{1}{2}\|\psi\|^2$.

Theorem 2.14 *Given* $\psi \in \mathbb{L}_2(\mathbb{R});\ t_0 \in \mathbb{R};\ \omega_0 \in \mathbb{R} \quad \curvearrowright$

$$\|(t-t_0)\psi\| \cdot \left\|(\omega-\omega_0)\widehat{\psi}\right\| \geq \frac{1}{2}\|\psi\|^2 .$$

(Proofs see. [1], p. 45 - 46)

Interpretation:
For example given $\psi \in \mathbb{L}_2(\mathbb{R})$ with

$$\begin{aligned} \|\psi\|^2 &= \int_{-\infty}^{\infty} |\psi|^2 \, dt = 1 \\ &\stackrel{!}{=} \int_{-\infty}^{\infty} f_X dt \end{aligned}$$

Then $f_X = |\psi(t)|^2$ can be interpreted as a probability density for the function X with the meaning „the signal is substantially different from zero".

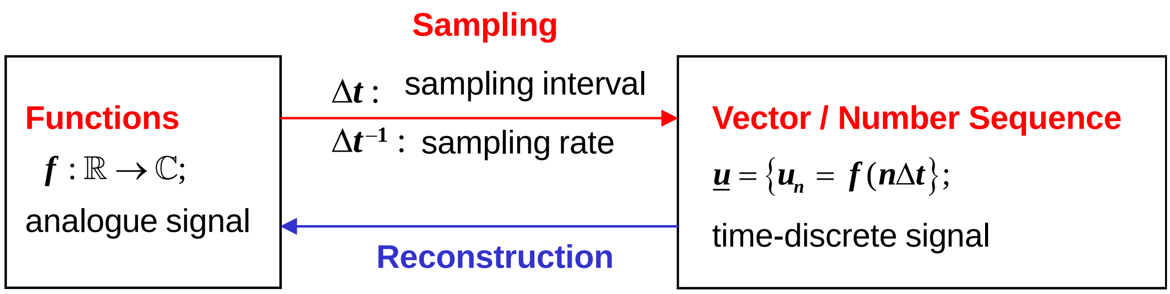
$$\|t\psi\|^2 = \int_{-\infty}^{\infty} t^2 |\psi|^2 \, dt = \int_{-\infty}^{\infty} t^2 f_X dt = \sigma_X^2$$

But this is the central moment of order two, i.e. a measure of the spread of the signal on the t-axis, of the „width of the signal". The points of reference for the „width of the signal" are $t = 0$ and $t = t_0$, i.e. the expectancy. Therefore the inequalities from the theorems above say :

$$\text{„width of the signal"} \cdot \text{„width of the spectrum"} \geq \frac{1}{2}\|\psi\|^2 = \frac{1}{2}.$$

Thus both "widths"can't be similarly small.
(Connection to physics:
$|\psi(t)|^2$: probability density for the location of a particle,
$\left|\widehat{\psi}(\omega)\right|^2$: probability density for the momentum of a particle)

2.4 Sampling Theorem (SHANNON)



If $f(t)$ is T-periodic then f can be represented by a discrete Fourier series using the sampled values $\underline{u} = \{u_0, ..., u_{N-1}\}$ with $T = N\triangle t$.
If $f \in \mathbb{L}_2(\mathbb{R})$, then by sampling we get a number sequence of $\mathbf{l}_2$: $\underline{u} = \{u_n\}_{n\in\mathbb{Z}}$ with $\sum_{n\in\mathbb{Z}} |u_n|^2 < \infty$.

Theorem 2.15 *Sampling Theorem by SHANNON:*
Let be $\triangle t \leq \frac{\pi}{\Omega}$; $\widehat{f}(\omega) = 0$ *for* $|\omega| > \Omega$, *i.e. if* f *is* Ω*-bandwidth limited, and* $f \in \mathbb{L}_2(\mathbb{R})$ *or* $f(t) = O\left(\frac{1}{|t|^{1+\varepsilon}}\right)$; $\varepsilon > 0$, *then:*

$$f(t) = \sum_{k=-\infty}^{\infty} f(k\triangle t) si(\Omega(t - k\triangle t)); \quad t \in \mathbb{R}.$$

In German this series is called "Kardinalreihe von f*". It tends uniformly to* f*. (Proof see* [1] *p. 48-49)*

Notation 2.1 *All functions* e_ω *which occur in the spectrum of* f *have a period* $T \geq \frac{2\pi}{\Omega}$*.*
If $\triangle t \leq \frac{\pi}{\Omega}$ *then there are two samples in every period.*

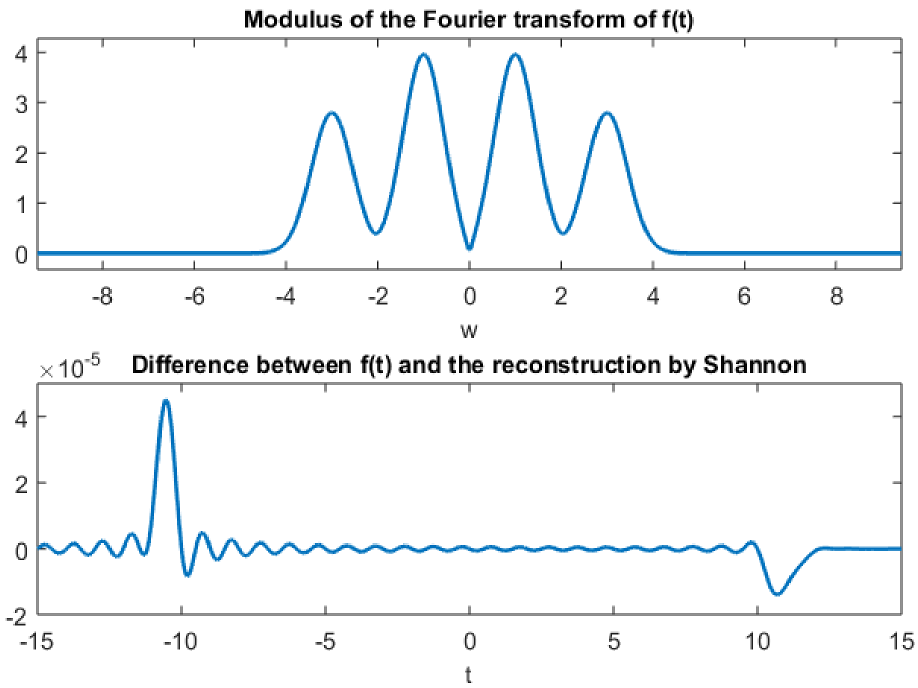
Notation 2.2 $\Omega = \frac{\pi}{\triangle t}$ *is called the NYQUIST frequency or the cut-off frequency for the sampling interval* $\triangle t$*.*

Notation 2.3 $\frac{1}{\triangle t} = \frac{\Omega}{\pi}$ *is the number of samples per time unit. It is called the NYQUIST-rate.*

Example 2.3 [3] *p. 254:*

$$f(t) = \exp(-0.1t^2)(2\sin t + \cos 3t)$$

This function is bandwith limited by $\Omega = 2\pi \quad \curvearrowright \quad \triangle t \leq \frac{\pi}{\Omega} = \frac{\pi}{2\pi} = \frac{1}{2}$, *which you can see in the following pucture:*



Because of $|f(t)| < 10^{-4}$ *for* $|t| > 10$ *we choose* $-20 \leq k \leq 20$ *for the reconstruction series in connection with* $\triangle t = 0.5$ *and get very small differences between the reconstructed and the original function.*

Notation 2.4 *Given* $(\triangle t)^{-1}$. $\quad \curvearrowright \quad \Omega = \pi \cdot (\triangle t)^{-1}$
When $\Omega < \Omega'$ *(real limitation), the high-frequency part of the signal with* $\Omega \leq \omega \leq \Omega'$ *is not ignored or filtered out. Instead it is undersampled and will appear at a lower frequency in the spectrum.*
$\Longrightarrow$ *Aliasing by „Undersampling"*
On the other hand if $\Omega > \Omega'$ *(real limitation), then the so called „Oversampling" improves the convergence by reconstruction.*

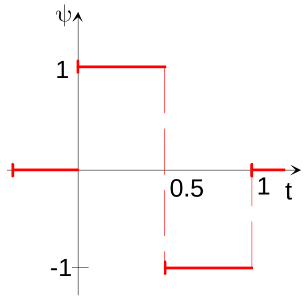
3 HAAR Wavelet

3.1 HAAR Basis

In 1910 Haar described a complete ONS for the space $\mathbb{L}_2(\mathbb{R})$. Today we write it down as a set of dilated and shifted copies of a special function, the mother wavelet:

Definition 3.1 *HAAR Wavelet*

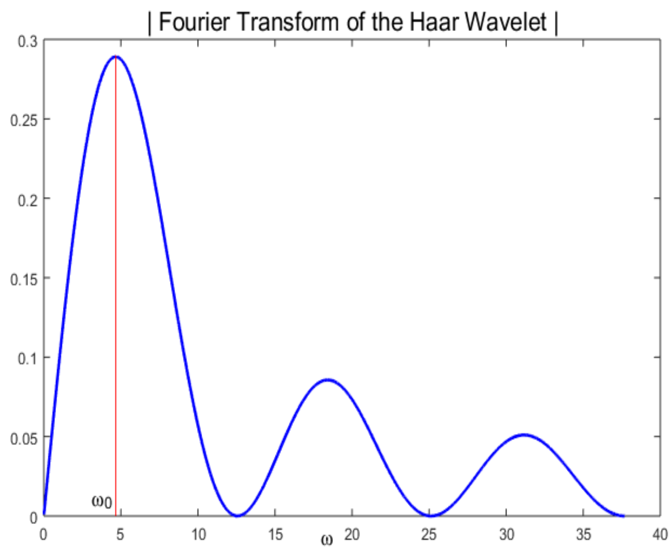
$$\psi_H(t) = \begin{cases} 1 & 0 \le t < 0.5 \\ -1 & 0.5 \le t < 1 \\ 0 & otherwise \end{cases}$$



Properties of the HAAR Wavelet:

- Its domain is bounded.
- It is discontinuous, and so not differentiable.
- $\int_{-\infty}^{\infty} \psi_H(t)dt = 0; \quad \int_{-\infty}^{\infty} |\psi_H(t)|^2 \, dt = 1,$

- $$
\begin{aligned}
\widehat{\psi}(\omega) &= \frac{1}{\sqrt{2\pi}}\left(\int_0^{0.5} e^{-i\omega t}dt - \int_{0.5}^{1} e^{-i\omega t}dt\right)\\
&= \left(\frac{1}{\sqrt{2\pi}}\right)\left(-\frac{1}{i\omega}\right)\left(\left[e^{-i\omega t}\right]_0^{0.5} - \left[e^{-i\omega t}\right]_{0.5}^{1}\right)\\
&= \frac{i}{\sqrt{2\pi}\omega}\left(e^{-i\omega/2} - 1 - e^{-i\omega} + e^{-i\omega/2}\right)\\
&= \frac{i}{\sqrt{2\pi}\omega}e^{-i\omega/2}\left(2 - e^{i\omega/2} - e^{-i\omega/2}\right)\\
&= \frac{i}{\sqrt{2\pi}\omega}e^{-i\omega/2}\left(2 - 2\cos(\frac{\omega}{2})\right)\\
&= \frac{i}{\sqrt{2\pi}\omega}e^{-i\omega/2}\cdot 2\left(2\sin^2(\frac{\omega}{4})\right)\\
&= \frac{i}{\sqrt{2\pi}}e^{-i\omega/2}\cdot\frac{4}{\omega}\sin^2(\frac{\omega}{4})\\
&= \frac{i}{\sqrt{2\pi}}e^{-i\omega/2}\cdot si(\frac{\omega}{4})\sin(\frac{\omega}{4})\\
\left|\widehat{\psi}(\omega)\right| &= \frac{1}{\sqrt{2\pi}}\frac{4}{|\omega|}\sin^2(\frac{\omega}{4}).
\end{aligned}
$$

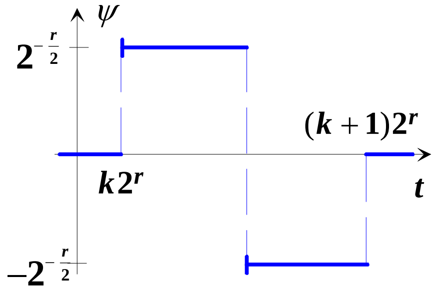


- $\curvearrowright$ As $|\omega| \to \infty$, $|\widehat{\psi}(\omega)|$ tends to zero like $\frac{1}{|\omega|}$.
- $\left|\widehat{\psi}(\omega)\right|$ is even.
- We find the maximum of $|\widehat{\psi}(\omega)|$ at $\omega_0 = 4.6622$.[1] $\curvearrowright$ $\widehat{\psi}(\omega)$ is relatively well located at ω_0.

Definition 3.2 *Wavelet functions (child wavelets) of the HAAR wavelet:*

$$\psi_{r,k}(t) = 2^{-\frac{r}{2}}\psi_H\left(\frac{t-k2^r}{2^r}\right)$$

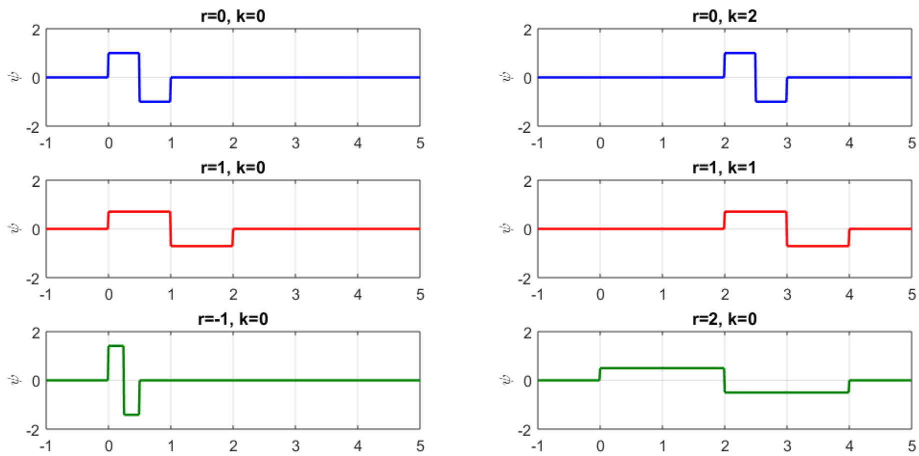
$$= 2^{-\frac{r}{2}} \cdot \begin{cases} 1 & 0 \le \dfrac{t-k2^r}{2^r} < \frac{1}{2} \\ -1 & \frac{1}{2} \le \dfrac{t-k2^r}{2^r} < 1 \\ 0 & \mathit{otherwise} \end{cases}$$



Calculation of the domain:

$$\begin{array}{ccccccccccc} 0 & \le & t-k2^r & < & \frac{1}{2}2^r & \Longrightarrow & k2^r & \le & t & < & (k+\frac{1}{2})2^r \\ \frac{1}{2}2^r & \le & t-k2^r & < & 2^r & \Longrightarrow & (k+\frac{1}{2})2^r & \le & t & < & (k+1)2^r \end{array}$$

Examples of wavelet functions:



Properties of the wavelet functions:

- Their domain is bounded.
- The bigger r the longer the domain intervals, i.e. the corresponding wavelet functions have longer wave lengths, they are long-wave.
- k is the shift parameter.
- $\|\psi_{rk}(t)\|^2 = 1$ (proof: homework)

Theorem 3.1 *The functions $\psi_{rk}(t)$ are an ONS in $\mathbb{L}_2(\mathbb{R})$.*

Proof:

1a) $\|\psi_{rk}(t)\|^2 = 1$ *(see above, properties of $\psi_{rk}(t)$)*

1b) We want to show: $(\psi_{rk}, \psi_{sl}) = 0$ *for* $\overline{(r = s \wedge k = l)}$ *i.e.* $r \neq s \vee k \neq l$

Case 1: Consider ψ_{rk} and ψ_{rl} with $k \neq l$:

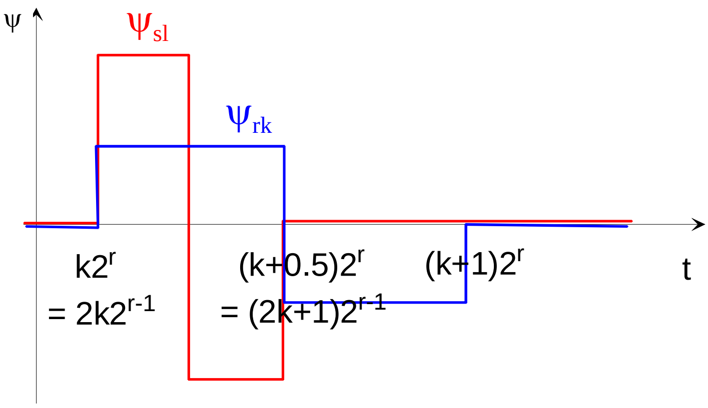
supp $\psi_{rk} \cap$ *supp* $\psi_{rl} = [k2^r;\ (k+1)2^r) \cap [l2^r;\ (l+1)2^r) = \emptyset$

$\Longrightarrow \quad (\psi_{rk}, \psi_{rl}) = 0$ *for* $k \neq l$

Case 2: Consider ψ_{rk} and ψ_{sl}; Let be $s < r$

supp $\psi_{rk} = [k2^r\,;(k+1)2^r)$; *supp* $\psi_{sl} = [l2^s;\ (l+1)2^s)$

Multiples of different powers of 2, which differ only by "1" in the exponent are lying in the "first half of the domain corresponding to the next higher power of 2":



$\Longrightarrow \quad (\psi_{rk}, \psi_{sl}) = 0$ *for* $s \neq r \quad \forall k, l$

2) We construct a basis by induction

- $f \in \mathbb{L}_2(\mathbb{R}); \quad n \in N.$
- *We consider the step approximation* $T_{-n}f$ *of* f *by intervals of the width* 2^{-n}.
- *Over the interval* $I_{-n,k} = [k2^{-n}; (k+1)2^{-n})$, $T_{-n}f$ is a constant function with the value
$$f_{-n,k} = 2^n \int_{k2^{-n}}^{(k+1)2^{-n}} f(t)dt = MW_{-n,k}$$
But this is the mean value of f *at* $I_{-n,k}$.
- If n *is big enough then* $T_{-n}f$ *appoximates* f *with any accuracy. (proof see Analysis, 1st semester)*
- We have to show *that we can approximate* f *by a finite linear combination of* ψ_{rk} *with respect to the* $\mathbb{L}_2-$*metric with any accuracy.*
- Because of the *approximation of functions of* $\mathbb{L}_2(\mathbb{R})$ *by step functions it is sufficient to consider a function* f of the following kind (instead of an arbitrary function $f \in \mathbb{L}_2(\mathbb{R})$) *:*
$$\exists m, n \quad | \qquad f(t) \equiv 0 \qquad for \quad |t| \geq 2^m$$
$$\wedge \quad f(t) \text{ is a step function, constant on } I_{-n,k}$$
- *Idea behind the proof:*
a) We construct a sequence of wavelet polynomials by induction
$$\{\psi_r\}_{r \geq -n} \quad \text{with} \quad \psi_r = \sum_{j=-n+1}^{r} \left(\sum_k c_{jk} \psi_{jk} \right)$$
b) We start with the smallest details i.e. with the most short-wave details of length 2^{-n+1}
c) Double the length of the domain. Look for details of this length until the length of the domain is bigger than 2^m *and the remainder* $f_r \to 0$ *as* $r \to \infty$.
- *Ansatz:*
$$\begin{aligned} f &= \psi_r + f_r \\ f_r &= const. \\ &= f_{rk} = 2^{-r} \int_{I_{rk}} f(t)dt = MW_{r,k} \quad (3.1.1) \\ \text{over the interval } I_{rk} &= [k2^r; (k+1)2^r) \end{aligned}$$

Base clause:

$$r = -n: \quad \psi_{-n} = 0; \quad f_{-n} = f$$

Induction hypothesis:

$$f = \psi_r + f_r \text{ for the index } r$$

Induction step:

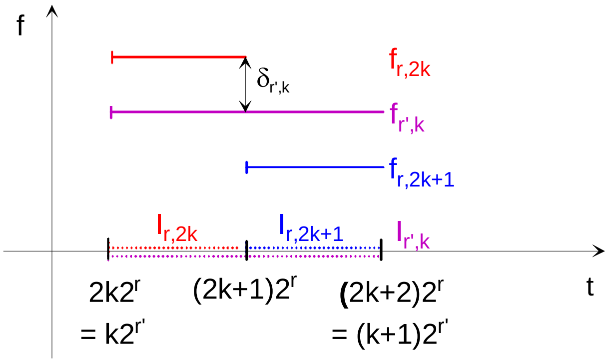
$$r' = r + 1$$

We calculate:

$$\delta_{r'k} = \frac{1}{2}\left(f_{r,2k} - f_{r,2k+1}\right)$$

$$f_{r',k} = \frac{1}{2}\left(f_{r,2k} + f_{r,2k+1}\right)$$

$$c_{r',k} = 2^{r'/2}\delta_{r',k}$$



$\delta_{r',k}$ *is a half of the step height and* $f_{r',k}$ *is the new mean value over the doubled interval.*

$$f_{I_{r'k}} = \left\{\begin{array}{ll} f_{r,2k} & t \in I_{r,2k} \\ f_{r,2k+1} & t \in I_{r,2k+1} \\ 0 & otherwise \end{array}\right\} = f_{r'k} + 2^{+\frac{r'}{2}}\delta_{r'k} \cdot \psi_{r'k} \quad with$$

$$\psi_{r'k} = 2^{-\frac{r'}{2}}\left\{\begin{array}{ll} +1 & t \in I_{r,2k} \\ -1 & t \in I_{r,2k+1} \\ 0 & otherwise \end{array}\right\}$$

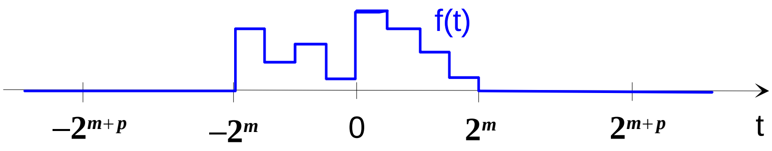
Corresponding to the induction hypothesis we get*:*

$$\begin{aligned} f &= \psi_r + f_r \\ &= \psi_r + \sum_k \left(c_{r'k}\psi_{r'k} + f_{r'k}\right) \\ &= \quad \psi_{r'} \quad + \quad f_{r'} \end{aligned}$$

- *After* $n+m$ *steps we get :*

$$f = \psi_{-n+n+m} + f_m = \sum_{j=-n+1}^{m}\left(\sum_k c_{jk}\psi_{jk}\right) + f_m$$

f_m is a constant function over the intervals $I_{m,k} = [k2^m; (k+1)2^m)$ with the length 2^m.



Then, a*t most, A and B are not equal zero:*

$$\begin{aligned} A &= f_{m,-1} = MW_{m,-1} \neq 0 \text{ over } I_{m,-1} = [-2^m; 0) \\ B &= f_{m,0} = MW_{m,0} \neq 0 \text{ over } I_{m,0} = [0; 2^m). \end{aligned}$$

Continue with the method. *After p steps the remainder will be:*

$$f_m = \sum_{j=m+1}^{m+p} \left(\sum_k c_{jk} \psi_{jk} \right) + f_{m+p},$$

with $f_{m+p} = const.$ over $[-2^{m+p}; 0)$ and over $[0; 2^{m+p})$. Otherwise $f_{m+p} = 0$. Because of the calculation we get the mean values: $f_{m+p,-1} = 2^{-p}A$ and $f_{m+p,0} = 2^{-p}B$. $\curvearrowright$

$$\begin{aligned} \|f_{m+p}\|^2 &= \int_{-\infty}^{\infty} |f_{m+p}(t)|^2 \, dt \\ &= 2^{m+p} \left((2^{-p}A)^2 + (2^{-p}B)^2 \right) \\ \|f_{m+p}\| &= \left\| f - \psi_{m+p} \right\| \\ &= \sqrt{2^{m-p}(A^2 + B^2)} \\ &= C \cdot 2^{-\frac{p}{2}} \quad \xrightarrow{p \to \infty} \quad 0 \end{aligned}$$

■

3.2 Fast HAAR Transform

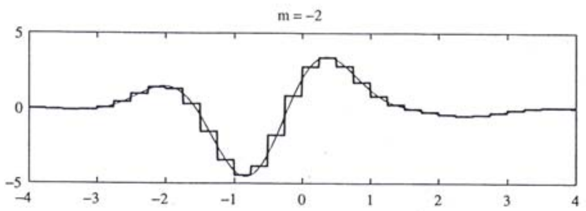
Every function $f \in \mathbb{L}_2(\mathbb{R})$ can be represented by the ONS of HAAR wavelet functions as:

$$f \simeq \sum_{r=r_0+1}^{r_1} \sum_{k \in Z} c_{rk} \psi_{rk}.$$

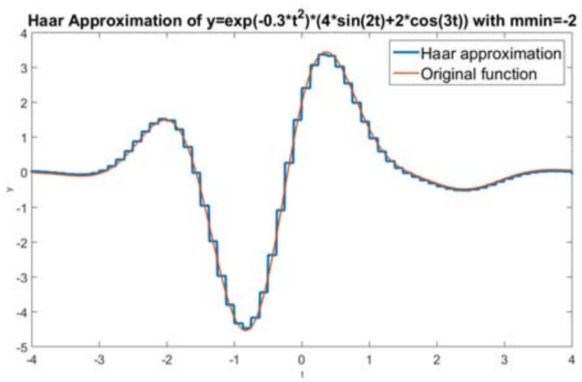
with the coefficients

$$
\begin{aligned}
c_{rk} &= (f, \psi_{rk}) = \int_{k2^r}^{(k+1)2^r} f\psi_{rk} dt \\
&= 2^{-\frac{r}{2}} \left[\int_{k2^r}^{(k+0.5)2^r} f(t)dt - \int_{(k+0.5)2^r}^{(k+1)2^r} f(t)dt \right].
\end{aligned}
$$

Example 3.1 $f(t) = e^{-0.3t^2}(4\sin(2t) + 2\cos(3t))$; $\quad D_f = [-4; 4]$ *(see [3])*
In the formula above we choose $-2 \le r \le 6$ and k such that $\psi_{rk} \neq 0$ in the domain D_f. Thus 70 coefficients c_{rk} must be calculated by numerical integration! This procedure is too expensive for the moderate-quality result.



step approximation $T_{-2}f$



approximation of f by linear combination of wavelet functions

But the proof of Theorem 4.3.1 gives us a faster algorithm:

- **The interesting point** is the following formula:

$$
\left.
\begin{aligned}
f_{r+1,k} &= \tfrac{1}{2}\left(f_{r,2k} + f_{r,2k+1}\right) \\
c_{r+1,k} &= 2^{\frac{r+1}{2}} \tfrac{1}{2}\left(f_{r,2k} - f_{r,2k+1}\right)
\end{aligned}
\right\} \quad (3.2.1)
$$

- Together with the **initial values**

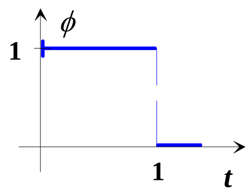
$$
f_{r_0,k} \approx f((k+0.5)2^{r_0}),
$$

  and formula (3.2.1) we get a

- **recurrence scheme** for calculating the values $c_{r,k}$ without integration. If f is sufficiently smooth and r_0 small enough the initial values $f_{r_0,k}$ can be interpreted as values of a step function in the interval $I_{r_0,k} = [k2^{r_0}; (k+1)2^{r_0})$.
- **Complexity**: $N = 2 \cdot 2^r \cdot 2^k$ initial values, the algorithm stops after $r+k$ steps, During the first step there are $\frac{N}{2}$ pairs of intervals, and per pair we do 2 additions. During the following step the number of pairs will be halved.
$\curvearrowright$ $\quad Complexity = 2 \cdot \frac{N}{2}(1 + \frac{1}{2} + \frac{1}{4} + ...) = 2N$ additions
- Therefore the algorithm (3.2.1.) is extremly fast. But the unbalance in the factors in front of the formulas is bad for the synthesis of the function.

Thus we are looking for an **improvement by a scaling function**. The HAAR Scaling function

$$\phi(t) = \begin{cases} 1 & 0 \leq t < 1 \\ 0 & otherwise \end{cases}$$



leads us to another family of functions:

$$\phi_{rk}(t) = 2^{-\frac{r}{2}}\phi(2^{-r}t - k), \quad k, r \in \mathbb{Z}.$$

with the domain:

$$0 \;\leq\; 2^{-r}t - k \;<\; 1 \qquad \Longrightarrow \qquad k2^r \;\leq\; t \;<\; (k+1)2^r.$$

$$\|\phi_{rk}(t)\|^2 = \int_{-\infty}^{\infty} |\phi_{rk}(t)|^2 \, dt = \int_{k2^r}^{(k+1)2^r} \left(2^{-\frac{r}{2}}\right)^2 dt = 2^{-r} \cdot 2^r = 1$$

Let r be given. Then the domains of ϕ_{rk} with $k \in \mathbb{Z}$ are disjunct and the functions ϕ_{rk} are an ONB of $\mathbb{L}_2(\mathbb{R})$.
$\curvearrowright$

$$\begin{aligned} T_r f &= \sum_k u_{rk}\phi_{rk} \quad with \\ u_{rk} &= (f, \phi_{rk}) = \int_{k2^r}^{(k+1)2^r} f \cdot 2^{-\frac{r}{2}} dt \\ &= 2^{\frac{r}{2}} \cdot 2^{-r} \int_{k2^r}^{(k+1)2^r} f(t) dt \\ &= 2^{\frac{r}{2}} \cdot f_{rk} \quad (from\ (3.1.1.)) \end{aligned}$$

Replacement in formula (3.2.1) results in

$$\begin{aligned} f_{r+1,k} &= \frac{1}{2}\left(f_{r,2k} + f_{r,2k+1}\right) \\ 2^{-\frac{r+1}{2}} u_{r+1,k} &= \frac{1}{2}\left(2^{-\frac{r}{2}} u_{r,2k} + 2^{-\frac{r}{2}} u_{r,2k+1}\right) \\ u_{r+1,k} &= \frac{\sqrt{2}}{2}\left(u_{r,2k} + u_{r,2k+1}\right) \quad (3.2.2.a) \end{aligned}$$

and

$$\begin{aligned} c_{r+1,k} &= 2^{\frac{r+1}{2}} \frac{1}{2}\left(f_{r,2k} - f_{r,2k+1}\right) \\ &= 2^{\frac{r+1}{2}} \frac{1}{2} 2^{-\frac{r}{2}}\left(u_{r,2k} - u_{r,2k+1}\right) \\ &= \frac{\sqrt{2}}{2}\left(u_{r,2k} - u_{r,2k+1}\right) \quad (3.2.2.b). \end{aligned}$$

Now formulas (3.2.2) are symmetric. The algorithm is analogous to (3.2.1) with the initial values

$$u_{r_0,k} = 2^{\frac{r_0}{2}} \cdot f_{r_0 k} \approx 2^{\frac{r_0}{2}} \cdot f((k+0.5)2^{r_0})$$

and stops at $r = M$.

<u>**Inverse transform/synthesis:**</u>

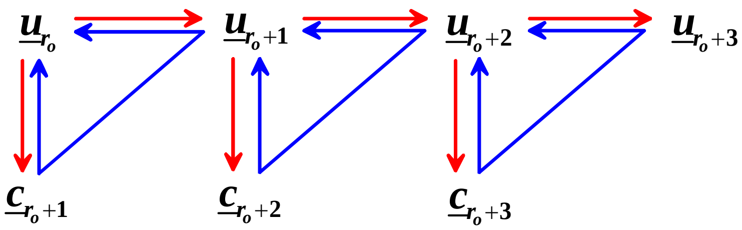
$$\begin{aligned} u_{r+1,k} + c_{r+1,k} &= \frac{\sqrt{2}}{2} u_{r,2k} \cdot 2 \\ u_{r+1,k} - c_{r+1,k} &= \frac{\sqrt{2}}{2} u_{r,2k+1} \cdot 2 \end{aligned}$$

$\curvearrowright$

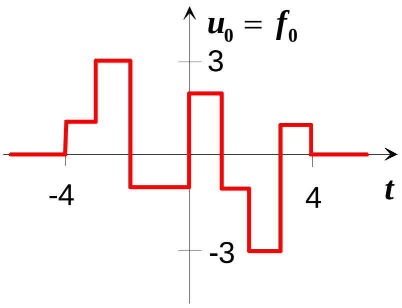
$$\left.\begin{array}{r} u_{r,2k} = \frac{\sqrt{2}}{2}\left(u_{r+1,k} + c_{r+1,k}\right) \\ u_{r,2k+1} = \frac{\sqrt{2}}{2}\left(u_{r+1,k} - c_{r+1,k}\right) \end{array}\right\} (3.2.3)$$

- Initial values for synthesis are $\underline{u}_M$ and $\underline{c}_M$.
- After calculation (3.2.3) we know the values $f_{r_0,k} = 2^{-r_0/2} \cdot u_{r_0,k}$, i.e. a step approximation of f.
- Complexity is the same as in analysis.

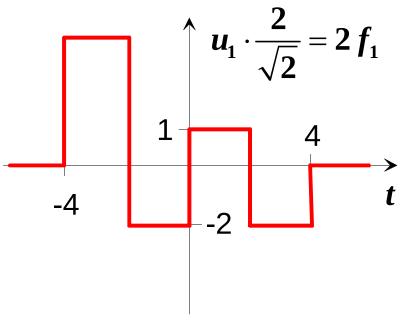
Structure of the algorithm:



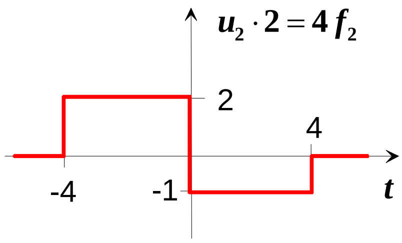
Example 3.2



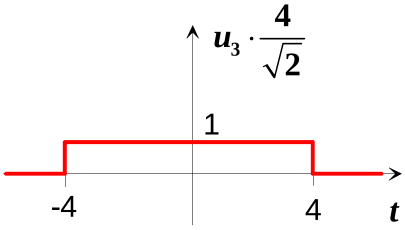
$$r_0 = 0,$$
$$u_0 = f_0 = (..., 0, 1, 3, -1, -1, 2, -1, -3, 1, 0...)$$
$$k = -4, ..., 0, ..., 3$$
(3.2.2) results:



$$r = 1$$
$$c_1 = \frac{\sqrt{2}}{2}(..., 0, -2, 0, 3, -4, 0, ...)$$
$$u_1 = \frac{\sqrt{2}}{2}(..., 0, 4, -2, 1, -2, 0, ...)$$
$$k = -2, ..., 0, 1$$
(3.2.2) results further:



$$r = 2$$
$$c_2 = \frac{1}{2}(..., 0, 6, 3, 0, ...)$$
$$u_2 = \frac{1}{2}(..., 0, 2, -1, 0, ...)$$
$$k = -1, 0$$
(3.2.2) results further:



$$\left[c_3 = \frac{\sqrt{2}}{4}(..., 0, 3, 0, ...)\right]$$
$$\left[u_3 = \frac{\sqrt{2}}{4}(..., 0, 1, 0, ...)\right]$$
But there are not more ψ − functions!
Use the rest!

$$
\begin{aligned}
T_0 f &= c_{1,-2}\psi_{1,-2} + c_{1,-1}\psi_{1,-1} + c_{1,0}\psi_{1,0} + c_{1,1}\psi_{1,1} + \\
&\quad + c_{2,-1}\psi_{2,-1} + \ c_{2,0}\psi_{2,0} + \\
&\quad + u_{2,-1}\phi_{2,-1} + u_{2,0}\phi_{2,0} \\
&= -\sqrt{2}\psi_{1,-2} + \frac{3\sqrt{2}}{2}\psi_{1,0} - 2\sqrt{2}\psi_{1,1} + \\
&\quad + 3\psi_{2,-1} + \frac{3}{2}\psi_{2,0} + \\
&\quad + 1\phi_{2,-1} - \frac{1}{2}\phi_{2,0} \\
&= \textit{wavelet polynomial by } c_1 \textit{ and } c_2 + \textit{rest by } u_2.
\end{aligned}
$$

4 Continuous Wavelet Transform

4.1 Wavelets - Definition and Properties

Definition 4.1 *Let ψ be a mapping from $\mathbb{R}$ to $\mathbb{C}$ with*

1. $\psi \in \mathbb{L}_2(\mathbb{R}); \quad \|\psi\| = 1 \qquad (4.1.1)$ *and*
2. $$0 < c_\psi = 2\pi \int_{\mathbb{R}^*=\mathbb{R}\setminus\{0\}} \frac{\left|\widehat{\psi}(\omega)\right|^2}{|\omega|} d\omega < \infty \qquad (4.1.2).$$

ψ is then called (mother) wavelet.

- (4.1.1) and (4.1.2) are minimal requirements.
- In practice all wavelets ψ belong to $\mathbb{L}_1(\mathbb{R})$.

Theorem 4.1 $\psi(t) \in \mathbb{L}_2(\mathbb{R}); \quad t\psi \in \mathbb{L}_1(\mathbb{R}); \quad \curvearrowright$
(4.1.2) $\iff \int_{-\infty}^{\infty} \psi(t)dt = 0 \iff \widehat{\psi}(0) = 0$ *(4.1.3).*

Theorem 4.2 *Given a k-times differentiable function ϕ; $k \geq 1$ with $\phi^{(k)} \in \mathbb{L}_2(\mathbb{R}); \quad \phi^{(k)} \neq 0; \implies$ After normalising $\psi(\omega) = \phi^{(k)}(\omega)$ is a wavelet.*

Theorem 4.3 *If $\psi \in \mathbb{L}_2(\mathbb{R}); \quad \|\psi\| = 1; \quad \int_{-\infty}^{\infty} \psi(t)dt = 0;$ and if the domain of ψ is compact then ψ is a wavelet.*

Theorem 4.4 *If $0 \neq \psi \in \mathbb{L}_2(\mathbb{R}) \cap \mathbb{L}_1(\mathbb{R}); \quad \int_{-\infty}^{\infty} \psi(t)dt = 0;$ and if there exists a number $\beta > 0.5 \quad | \quad \int_{\mathbb{R}} |t|^\beta |\psi(t)|dt = k < \infty$ then ψ is a wavelet.*

Theorem 4.5 *The collection of all wavelets $\Psi = \{\psi : \mathbb{R} \to \mathbb{C} \quad | \quad (4.1.1),\ (4.1.2)\}$ is dense in $\mathbb{L}_2(\mathbb{R})$.*

(Proofs see. [2])

Example 4.1 *HAAR Wavelet:*

$$\psi_H(t) = \begin{cases} 1 & 0 \leq t < 0.5 \\ -1 & 0.5 \leq t < 1 \\ 0 & \text{otherwise} \end{cases}$$

$\psi_H \in \mathbb{L}_2(\mathbb{R})$ *because* $\int_0^1 1dt = 1 < \infty$, *thus requirement (4.1.1) is satisfied.*
$t\psi_H \in \mathbb{L}_1(\mathbb{R})$ *because*

$$\int_0^{0.5} tdt - \int_{0.5}^1 tdt = -\frac{1}{4}.$$

Furthermore: $\psi_H = \phi_{0.5} - \phi_{0.5}(t-0.5) \quad \curvearrowright$

$$\begin{aligned}
\widehat{\psi_H} &= \widehat{\phi_{0.5}} - e^{-i\omega/2}\widehat{\phi_{0.5}} \\
&= \frac{e^{-i\omega/4}}{\sqrt{2\pi}2} si\left(\frac{\omega}{4}\right)\left[1 - e^{-i\omega/2}\right] \\
&= \frac{ie^{-i\omega/4}}{\sqrt{2\pi}} si\left(\frac{\omega}{4}\right) e^{-i\omega/4}\left[\frac{e^{i\omega/4} - e^{-i\omega/4}}{2i}\right] \\
&= \frac{ie^{-i\omega/2}}{\sqrt{2\pi}} si\left(\frac{\omega}{4}\right)\sin\left(\frac{\omega}{4}\right).
\end{aligned}$$

$\Longrightarrow \quad \widehat{\psi_H}(0) = 0$ *and so the requirements of theorem 4.1 are satisfied. Thus* ψ_H *is a wavelet (*$c_\psi = 2\ln 2$ *see* [2], *p. 17).*

Example 4.2 *Modulated GAUSS Function:*
Choose a fundamental frequency, for example $\omega = 5$.
But $\chi(t) = e^{i\omega t}e^{-t^2/2}$ *is not a wavelet because* $\widehat{\chi}(0) \neq 0$. $\curvearrowright$ *Thus we take the ansatz:*

$$\psi(t) = (e^{i\omega t} - A)e^{-t^2/2}$$

From the properties (R2) and example 2.2 of the Fourier transform we get:

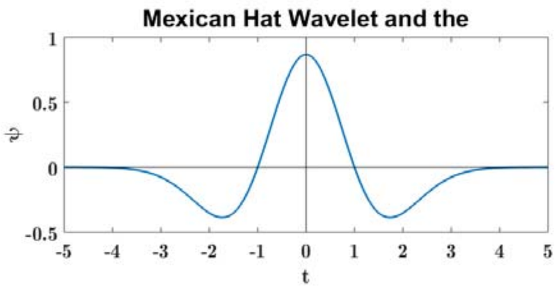
$$\begin{aligned}
\widehat{\psi}(\xi) &= \exp\left(-\frac{(\xi-\omega)^2}{2}\right) - A\exp\left(-\frac{\xi^2}{2}\right) \\
\widehat{\psi}(0) &= \exp\left(-\frac{\omega^2}{2}\right) - A \stackrel{!}{=} 0 \quad \curvearrowright \quad A = \exp\left(-\frac{\omega^2}{2}\right).
\end{aligned}$$

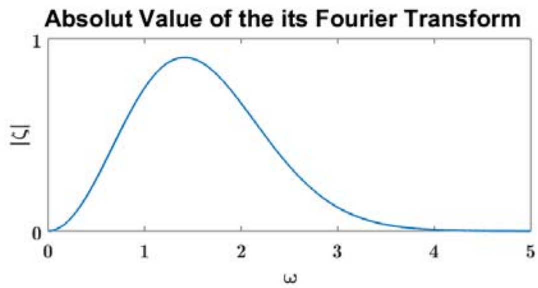
After nomalising, the resulting function

$$\psi(t) = (\exp(i\omega t) - \exp(-\frac{\omega^2}{2}))\exp(\frac{-t^2}{2})$$

will be a wavelet.

Example 4.3 *Mexican Hat*





$$\psi(t) = \frac{2}{\sqrt{3}\sqrt[4]{\pi}}(1-t^2)\exp(-\frac{t^2}{2}) = \gamma(1-t^2)\exp(-\frac{t^2}{2})$$
$$\psi(t) = -\gamma g''(t) \quad \text{with} \quad g(t) = \exp(-\frac{t^2}{2})$$

From property (R4) of the Fourier transform we get:

$$\widehat{\psi}(\omega) = -\gamma(i\omega)^2\widehat{g}(\omega) = \widetilde{\gamma}\omega^2\exp(-\frac{\omega^2}{2}) \quad \curvearrowright$$
$$\widehat{\psi}(0) = 0$$

Therefore $\psi(t)$ is a wavelet.

4.2 Continuous Wavelet Transform

Definition 4.2 *Given the wavelet ψ, a function $f \in \mathbb{L}_2(\mathbb{R})$ and a number $a \neq 0$. Then*

$$Wf(a,b) = \frac{1}{\sqrt{c_\psi}}\frac{1}{|a|^{0.5}}\int_{-\infty}^{\infty} f(t)\overline{\psi(\frac{t-b}{a})}dt \qquad (4.2.1)$$

is called a ***wavelet transform*** *of f with respect to ψ.*

- $D_{Wf} = \mathbb{R}^2_- = \{(a,b)^T \mid a \in \mathbb{R}^*,\ b \in \mathbb{R}\}$
- For calculating a wavelet transform it is necessary to know what is the corresponding wavelet.
- $a \neq 0 \quad \curvearrowright \quad \psi_a(t) = \frac{1}{|a|^{0.5}}\psi\left(\frac{t}{a}\right).$
  This function corresponds to a stretched ($|a| > 1$) or a compressed ($|a| < 1$) function. The function is normalized because
  $$\int_{\mathbb{R}} |\psi_a(t)|^2 dt = \frac{1}{|a|}\int_{\mathbb{R}} |\psi\left(\frac{t}{a}\right)|^2 dt = \frac{1}{|a|}\int_{\mathbb{R}} |\psi(z)|^2 |a| dz = 1.$$
- Shifting ψ_a by $b > 0$ to the right results:
  $$\psi_{a,b}(t) = \psi_a(t-b) = \frac{1}{|a|^{0.5}}\psi\left(\frac{t-b}{a}\right)$$
  with $\left\|\psi_{a,b}\right\| = 1$.
- $\curvearrowright$
  $$\begin{aligned} Wf(a,b) &= (f,\psi_{a,b}) \cdot \frac{1}{\sqrt{c_\psi}} \\ |Wf(a,b)| \overset{Schwarz}{\leq} \frac{1}{\sqrt{c_\psi}}\|f\| \cdot \left\|\psi_{a,b}\right\| &= \frac{1}{\sqrt{c_\psi}}\|f\| \quad \forall (a,b)^T \in \mathbb{R}^2_- \end{aligned}$$

Example 4.4 *HAAR Wavelet: $a > 0$*

$$\psi\left(\frac{t-b}{a}\right) = \begin{cases} 1 & 0 \leq \frac{t-b}{a} < 0.5 \quad \curvearrowright \quad b \leq t < b + \frac{a}{2} \\ -1 & 0.5 \leq \frac{t-b}{a} < 1 \quad \curvearrowright \quad b + \frac{a}{2} \leq t < b + a \\ 0 & otherwise \end{cases}$$

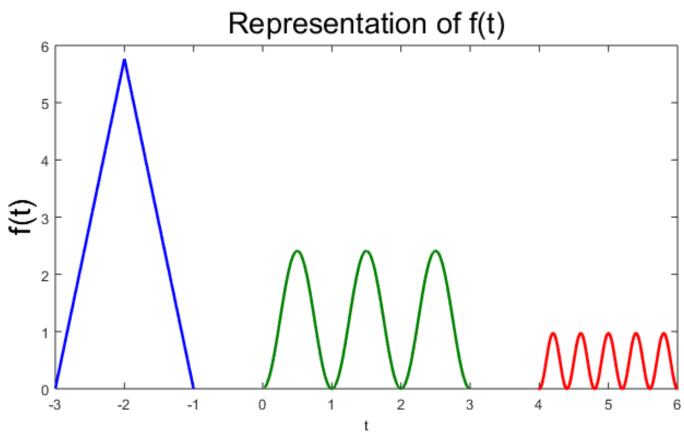
$\curvearrowright$

$$\begin{aligned} Wf(a,b) &= \frac{1}{\sqrt{c_\psi}}\frac{1}{|a|^{0.5}}\left[\int_b^{b+a/2} f(t)\,dt - \int_{b+a/2}^{b+a} f(t)\,dt\right] \\ &= \frac{1}{\sqrt{c_\psi}}\frac{|a|^{0.5}}{2}\left[\frac{2}{a}\int_b^{b+a/2} f(t)\,dt - \frac{2}{a}\int_{b+a/2}^{b+a} f(t)\,dt\right] \end{aligned}$$

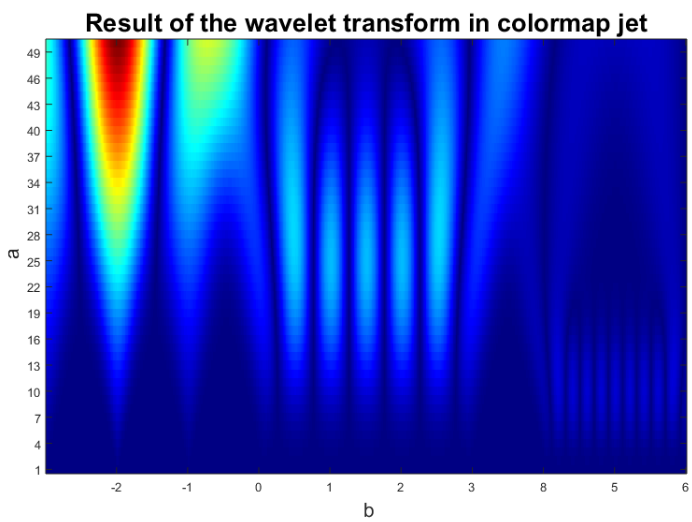
Interpretation:
This is the difference of two mean values of the function f which are taken over two adjacent intervals of the length $\frac{a}{2}$ centered at $\left(b + \frac{a}{2}\right)$ multiplied by a normalising factor, i.e. it is a „floating difference". (see. digital filter in [3] *p. 29)*

Example 4.5 *Analysis of the function $f(t)$ by the Mexican hat*

$$
\begin{array}{rcl}
f\left(t\right) & = & 2.883 f_1\left(t\right) + 1.205 f_2\left(t\right) + 0.968 f_3\left(t\right) \\
& & \begin{array}{lcc}
f_1\left(t\right) = 2 - 2|t+2| & for & -3 \leq t \leq -1 \\
f_2\left(t\right) = 1 - \cos(2\pi t) & for & 0 \leq t \leq 3 \\
f_3\left(t\right) = \frac{1}{2}(1 - \cos(5\pi t)) & for & 4 \leq t \leq 6 \\
f_i\left(t\right) = 0 & i = 1, 2, 3 & otherwise
\end{array}
\end{array}
$$



Representation of the wavelet transform by colors:



- $Wf : \mathbb{R}^2_- \to \mathbb{C}^1$: For the calculation of an inverse formula for the wavelet transform we need an inner product of functions $u : \mathbb{R}^2_- \to \mathbb{C}^1$ and therefore we need a measure on $\mathbb{R}^2_- = \mathbb{R}^* \times \mathbb{R}$, $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$.
- But the meaning of the variables a and b is mathematically different. That's why the LEBESGUE measure $d\mu = dadb$ is not applicable.
- $(a,b)^T \in \mathbb{R}^2_-$ defines an affine stretching $S_{a,b}(\tau) : a\tau + b = t$, such that $|a|$ is the parameter of stretching and b is the parameter of shifting only.
- Therefore we use the HAAR measure $d\mu = \frac{1}{a^2} dadb$
- and the corresponding Hilbert space $\mathbb{H} = \mathbb{L}_2\left(\mathbb{R}^2_-, d\mu\right)$ with the inner product

$$(u,v)_{\mathbb{H}} = \int_{\mathbb{R}^2_-} u(a,b)\overline{v(a,b)}\frac{1}{a^2}dadb.$$

  This is a weighted Hilbert space (important for consideration when using group theory).
- Our goal is to calculate the inverse operator of $Wf(a,b)$. First we show that the wavelet transform is an isometry and then we use $Wf^{-1} = Wf^*$.

Theorem 4.6 *The wavelet transform by the wavelet $\psi : \mathbb{L}_2(\mathbb{R}) \to \mathbb{L}_2\left(\mathbb{R}^2_-, d\mu\right) = \mathbb{H}$ is an isometry.* [2]

Proof. *$\psi \in \mathbb{L}_2(\mathbb{R})$ implies $\psi\left(\frac{\cdot - b}{a}\right) \in \mathbb{L}_2(\mathbb{R}) \implies Wf$ is an operater from $\mathbb{L}_2(\mathbb{R})$ to $\mathbb{L}_2(\mathbb{R})$, i.e. $|Wf(a,b)| < \infty$.*

$$\begin{aligned}
\|Wf(a,b)\|^2_{\mathbb{H}} &= \int_{\mathbb{R}}\int_{\mathbb{R}^*} |Wf(a,b)|^2 \frac{1}{a^2} dadb \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}^*} \left|\widehat{Wf}(a,\omega)\right|^2 \frac{1}{a^2} dad\omega \quad \text{(PARSEVAL-PLANCHEREL Formula, respect to } b)
\end{aligned}$$

Calculation of $\widehat{Wf(a,\omega)}$ by $Wf(a,b)$:

$$\begin{aligned}
Wf(a,b) &= \frac{1}{\sqrt{c_\psi}}\frac{1}{|a|^{0.5}}\int_{-\infty}^{\infty} f(t)\overline{\psi(\frac{t}{a} - \frac{b}{a})}dt \\
&= \frac{1}{\sqrt{c_\psi|a|}}\int_{-\infty}^{\infty} f(t)\overline{\psi(\frac{b}{-a} - \frac{t}{-a})}dt \\
&= \frac{1}{\sqrt{c_\psi|a|}}\left(f * \overline{\psi(\frac{\cdot}{-a})}\right) b \\
&= \frac{1}{\sqrt{c_\psi|a|}}\left(f * \overline{D_{-a}\psi}\right) b
\end{aligned}$$

$$\begin{aligned}
\widehat{Wf}\,(a,\omega) &= \frac{1}{\sqrt{c_\psi |a|}}\sqrt{2\pi}\left[\widehat{f}(\omega)\cdot\widehat{D_{-a}\overline{\psi}}\right] \quad \textit{(rule of convolution)} \\
&= \frac{\sqrt{2\pi}}{\sqrt{c_\psi |a|}}\left[\widehat{f}(\omega)|a|D_{\frac{1}{-a}}\widehat{\overline{\psi}}\right] \quad \textit{(R3)} \\
&= \frac{\sqrt{2\pi}|a|}{\sqrt{c_\psi |a|}}\left[\widehat{f}(\omega)\widehat{\overline{\psi}}\left(\frac{\omega}{\frac{1}{-a}}\right)\right] \\
&= \frac{\sqrt{2\pi|a|}}{\sqrt{c_\psi}}\left[\widehat{f}(\omega)\widehat{\overline{\psi}}\,(-a\omega)\right] \quad \textit{(R6)}
\end{aligned}$$

Thus:

$$\|Wf\,(a,b)\|^2_{\mathbb{H}} = \int_{\mathbb{R}}\int_{\mathbb{R}^*}\frac{2\pi|a|}{c_\psi}\left|\widehat{\psi}\,(a\omega)\right|^2\left|\widehat{f}\,(\omega)\right|^2\frac{1}{a^2}da\,d\omega$$

We use the coordinate transformation $r = a|\omega|$; $dr = |\omega|da$, *split the integral in 2 summands, calculate the integrals, sum up and get:*

$$\begin{aligned}
\|Wf\,(a,b)\|^2_{\mathbb{H}} &= \int_{\mathbb{R}}\int_{\mathbb{R}^*}\frac{2\pi|r|}{c_\psi|\omega|}\left|\widehat{\psi}\,(r)\right|^2\left|\widehat{f}\,(\omega)\right|^2\frac{|\omega|^2}{|r|^2}\frac{dr}{|\omega|}d\omega \\
&= \frac{1}{c_\psi}\cdot\int_{\mathbb{R}^*}2\pi\left|\widehat{\psi}\,(r)\right|^2\frac{dr}{|r|}\cdot\int_{\mathbb{R}}\left|\widehat{f}\,(\omega)\right|^2 d\omega \\
&= \frac{1}{c_\psi}\cdot c_\psi\cdot\|f\|^2_{\mathbb{L}_2}
\end{aligned}$$

■

Now the inverse operator is the adjoint operator W^* on the transformed domain (range) ([2], p. 52):

$$\begin{aligned}
(f, W^*g)_{\mathbb{L}_2} &= (Wf, g)_{\mathbb{L}_2\left(\mathbb{R}^2_-,\ d\mu\right)} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}^*} Wf\,(a,b)\,\overline{g(a,b)}\frac{1}{a^2}da\,db \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}^*}\frac{1}{\sqrt{c_\psi}}\int_{\mathbb{R}}\frac{1}{\sqrt{|a|}}f\,(t)\,\overline{\psi\left(\frac{t-b}{a}\right)}dt\;\overline{g(a,b)}\frac{1}{a^2}da\,db \\
&= \int_{\mathbb{R}} f\,(t)\left[\frac{1}{\sqrt{c_\psi}}\int_{\mathbb{R}}\int_{\mathbb{R}^*}\frac{1}{\sqrt{|a|}}\overline{\psi\left(\frac{t-b}{a}\right)}\,\overline{g(a,b)}\frac{1}{a^2}da\,db\right]dt \\
&= \int_{\mathbb{R}} f\,(t)\left[\ \overline{W^*g}\ \right]dt
\end{aligned}$$

$\curvearrowright$

$$f = W^*\,(Wf) = \frac{1}{\sqrt{c_\psi}}\int_{\mathbb{R}}\int_{\mathbb{R}^*}\frac{1}{\sqrt{|a|}}Wf\,(a,b)\,\psi\left(\frac{t-b}{a}\right)\frac{1}{a^2}da\,db \qquad (4.2.2)$$

- It is possible to distribute the factors in front of the integral in another way (see Fourier transform). For example Blatter [1] works with the factor 1 in the wavelet transform and with the factor c_ψ in the inverse transform. But then the wavelet transform is not an isometry and we loose the possibility to invert Wf easily.

- Exchanging the order of integration is a delicate operation because of the improper integrals.

- Even it is possible to use different wavelets in analysing and synthesising. (see [2], prewavelet transform)

4.3 Properties of the Continuous Wavelet Transform

Now we describe the wavelet transform by the dilatation operator D_a and the translation operator T_h (see 2.2):

$$\begin{aligned} Wf(a,b) &= \frac{1}{\sqrt{c_\psi}} \frac{1}{|a|^{0.5}} \int_{\mathbb{R}} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \\ &= \frac{1}{\sqrt{c_\psi}} \frac{1}{|a|^{0.5}} \left(f(t), T_b D_a \psi\right)_{\mathbb{L}_2}, \end{aligned}$$

i.e. $Wf(a,b)$ is an inner product of f by the function system $\{T_b D_a \psi \mid a \neq 0 \ ; \ b \in \mathbb{R}\}$ which is complete if ψ satisfies the conditions (4.1.1) and (4.1.2).

4.3.1 Filter properties

Filters in signal processing are used to

- reduce data errors,

- separate high- and low- frequency parts of the signal and

- emphasize certain frequences.

The commonly used filters are filters by linear convolution in connection to the Fourier transform.

$$f_\phi = f * \phi \xrightarrow{FT} \widehat{f_\phi} = \sqrt{2\pi}\ \widehat{f} \cdot \widehat{\phi}$$

Example 4.6 *Kinds of filters are:*
low pass filter: $\widehat{\phi} \sim \chi_{[-B;B]}$ *: high frequencies are attenuated*
band pass filter: $\widehat{\phi} \sim \chi_{a\leq\omega\leq b}$ *: spectrum between* a *and* b *is selected*
high pass filter: $\widehat{\phi} \sim 1 - \chi_{[-B;B]}$*: low frequencies are attenuated*
($\chi_{[-B;B]}$ *is the characteristic function of the interval* $[-B;B]$.
Interpretation:
Let f *be a function which consists of details bigger than* L *only.*

$$f = \sum \alpha_j \chi_{I_j}(t) \quad with \quad |I_j| > L$$

$$\begin{aligned}
|\widehat{\chi_I}(\omega)| &= \frac{1}{\sqrt{2\pi}} \left| \exp\left(-i\frac{A+B}{2}\omega\right) \frac{2}{\omega} \sin\left(\frac{\omega|I|}{2}\right) \right| \quad with \quad I = [A;B],\ h = \frac{A+B}{2} \\
&= \frac{|I|}{\sqrt{2\pi}} \left| \frac{2}{|I|\omega} \sin\left(\frac{\omega|I|}{2}\right) \right| \\
&= \frac{|I|}{\sqrt{2\pi}} \left| si\left(\frac{\omega|I|}{2}\right) \right| \quad with \quad |I| = B - A > L
\end{aligned}$$

Because of $si(\cdot) >> 0$ in $[-\pi;\pi]$

$$\frac{\omega|I|}{2} \leq \pi \text{ implies } \omega \leq \frac{2\pi}{|I|}.$$

I.e. the domain of $\widehat{\chi_I}(\omega)$ is the interval $\left[-\frac{2\pi}{|I|};\frac{2\pi}{|I|}\right]$ essentially. Therefore details of a size $|I| \geq L$ correspond to a frequency $\omega \leq \frac{2\pi}{L}$.
On the other hand: $\left|\widehat{f}(\omega_0)\right| >> 0$, corresponds to details of size $\frac{2\pi}{\omega_0}$ in f.
Interpretation of the Wavelet Transform as a Filter:
We write $Wf(a,b)$ by convolution (see proof of theorem 4.6):

$$Wf(a,b) = \frac{1}{\sqrt{c_\psi}} \frac{1}{|a|^{0.5}} (f * \overline{D_{-a}\psi}) b$$

This corresponds to a filter by $\overline{\psi\left(\frac{\cdot}{-a}\right)}$ for given a.
Theorem 4.1 says $\widehat{\psi}(0) = 0$. $\psi \in \mathbb{L}_1(\mathbb{R}) \cap \mathbb{L}_2(\mathbb{R})$ implies $\lim_{\omega\to\infty} \left|\widehat{\psi}(\omega)\right| = 0$. Therefore the wavelet transform is a band pass filter. Because of the isometry (theorem 4.6) we get:

Theorem 4.7

$$Wf(a,b) = \frac{|a|^{0.5}}{\sqrt{c_\psi}} \int_{\mathbb{R}} \widehat{\psi}(a\omega)\, \overline{\widehat{f}(\omega)} e^{-i\omega b} d\omega$$

Proof: [2], *p. 28*

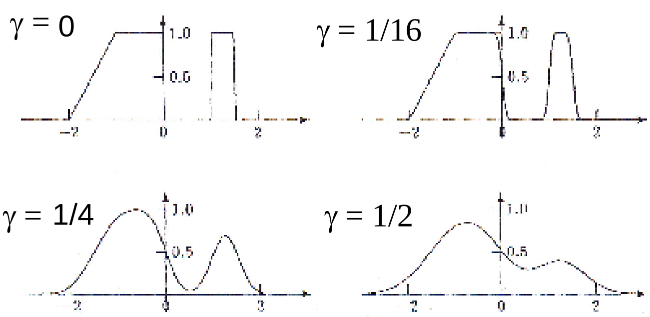
Conclusions:

- If $\widehat{\psi}(\omega)$ is concentrated at ω_0 (HAAR, Mexican hat), then $\widehat{\psi}(a\omega)$ is concentrated at $\omega = \frac{\omega_0}{a}$.
- For given a the wavelet transform $Wf(a,b)$ mostly includes information about the signal f at frequency $\frac{\omega_0}{a}$.
- Therefore a is called the frequency parameter.
- For given a the wavelet transform $Wf(a,b)$ includes information about details of size $2\pi/\frac{\omega_0}{a}$ in f.

Example 4.7 *In the synthesis of a function you must see a behavior with respect to details like that written above. That's why we give the inverse wavelet transform a parameter γ which controls the calculation of the improper integral:*

$$f_\gamma(t) = \int_{a>|\gamma|} \int_{\mathbb{R}} Wf(a,b)) \frac{1}{\sqrt{c_\psi}|a|^{0.5}} \psi\left(\frac{t-b}{a}\right) \frac{1}{a^2} dbda; \quad \lim_{\gamma \to 0} f_\gamma(t) = f$$

We get the following picture from the Mexican hat $\curvearrowright$ $\omega_0 = \sqrt{2},\ c_\psi = 1$ *([2], p. 18, fig. 1.4)*
In the original function you find two objects of size $L_I = 2$ and $L_{II} = 0.5$, which have „sharp edges". These edges are details of a very short length:



The wavelet transform includes information about details of size

$$L = \frac{2\pi}{\frac{\omega_0}{a}} = \frac{2\pi}{\sqrt{2}} a \xrightarrow{a \to 0} 0.$$

Because of $L = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{2}} a$ we get

$$\omega = \frac{\sqrt{2}}{a} \xrightarrow{a \to 0} \infty.$$

Thus we get a sharp mapping of edges only as $a \to 0$. What happens if you want to map objects of size $L = \frac{2\pi}{\sqrt{2}} a \geq \frac{2\pi}{\sqrt{2}} \gamma = \sqrt{2}\pi\gamma$ in f_γ?

$$\gamma_1 = \tfrac{1}{2} \quad \curvearrowright \quad L_1 = \sqrt{2}\pi \cdot \tfrac{1}{2} \approx 2.22$$

$$\gamma_2 = \tfrac{1}{4} \quad \curvearrowright \quad L_2 = \sqrt{2}\pi \cdot \tfrac{1}{4} \approx 1.11$$

$$\gamma_3 = \tfrac{1}{16} \quad \curvearrowright \quad L_3 = \sqrt{2}\pi \cdot \tfrac{1}{16} \approx 0.28$$

4.3.2 Phase Space Representation

In physics and in signal processing scientists are interested in the frequency distribution of a function at the point t_0 or over the interval $[t_0, t_e]$. That means they are looking for a function $Df(t, \omega)$ corresponding to $f(t)$ which indicates what is the contribution of the frequency ω at time t to the signal f.

Definition 4.3 *The set $\{(t, \omega) \mid t, \omega \in \mathbb{R}\}$ is called phase space.*

Definition 4.4 *The function $Df(t, \omega)$ is called the phase space representation of f.*

The phase space representation of f is not unique. For example the inner product $Df(t_0, \omega_0) = (g_{t_0,\omega_0}, f)$ is such a phase space representation with a function g concentrated around t_0 and $\widehat{g}$ concentrated around ω_0.
A very precise localisation of g around t_0 and $\widehat{g}$ around ω_0 at the same time is not possible because of the HEISENBERG uncertainty principle.

$$\|(t - t_0)g\| \cdot \|(\omega - \omega_0)\widehat{g}\| \geq \frac{1}{2} \|g\|^2 \quad (s.\ 2.3)$$

Definition 4.5 *$g \in \mathbb{L}_2(\mathbb{R})$, $\|g\|_{\mathbb{L}_2} = 1$; If*

$$\begin{aligned} -\infty &< t_0 = \int_{\mathbb{R}} t\, |g(t)|^2 \, dt < \infty \text{ and if} \\ -\infty &< \omega_0 = \int_{\mathbb{R}} \omega\, |\widehat{g}(\omega)|^2 \, d\omega < \infty \end{aligned}$$

then g is located around the phase point (t_0, ω_0) with uncertainty

$$\mu(g) = \|(t - t_0)g\|^2 \cdot \|(\omega - \omega_0)\widehat{g}\|^2 \geq \frac{1}{4}.$$

Interpretation of the Wavelet transform as a Phase Space Representation:

Let ψ be a wavelet, i.e. $\|\psi\|_{\mathbb{L}_2} = 1$, with $\int_{\mathbb{R}} t \left|\psi(t)\right|^2 dt = 0 = t_0$. Usually $|\psi|$ is an even function with two distinct maxima and so we have to adapt the above concept :

$$\begin{aligned} \omega_0^+ &= \int_0^\infty \omega \left|\widehat{\psi}(\omega)\right|^2 d\omega \\ \omega_0^- &= \int_{-\infty}^0 \omega \left|\widehat{\psi}(\omega)\right|^2 d\omega \end{aligned}$$

Then ψ is located around $\left(t_0, \omega_0^\pm\right)$. Therefore $\psi_{a,b} = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$ is located around $\left(t_0^{ab}, \omega_0^{ab}\right)$ with

$$\begin{aligned} t_0^{ab} &= \frac{1}{a}\int_{\mathbb{R}} t \left|\psi\left(\frac{t-b}{a}\right)\right|^2 dt \\ &= \frac{1}{a}\int_{\mathbb{R}} (b+as)\left|\psi(s)\right|^2 a ds \\ &= a\int_{\mathbb{R}} s\left|\psi(s)\right|^2 ds + b\int_{\mathbb{R}} \left|\psi(s)\right|^2 ds \\ &= b \end{aligned}$$

$$\begin{aligned} \omega_0^{ab} &= \int_{0\le\pm\omega<\infty} \omega \left|\widehat{\psi}_{ab}\right|^2 d\omega \\ &= \int_{0\le\pm\omega<\infty} \omega \frac{1}{|a|}\left|e^{-i\omega b}\right|^2 |a|^2 \left|\widehat{\psi}(a\omega)\right|^2 d\omega \\ &= \int_{0\le\pm\omega<\infty} a\omega \left|\widehat{\psi}(a\omega)\right|^2 d\omega \\ &= \int_{0\le\pm\omega<\infty} s \left|\widehat{\psi}(s)\right|^2 \frac{1}{a} ds \\ &= \frac{\omega_0^\pm}{a}. \end{aligned}$$

$\curvearrowright$ With $(a,b)\in\mathbb{R}^2$; $a\neq 0$ the point $\left(t_0^{ab}, \omega_0^{ab}\right)$ passes through the phase space.

$\curvearrowright$

$$Wf(a,b) = Df\left(b, \frac{\omega_0^\pm}{a}\right)$$

Given a.$\curvearrowright$

$$Wf(a,\cdot) = Df\left(\cdot, \frac{\omega_0^\pm}{a}\right)$$

corresponds to the frequency change around $\frac{\omega_0^\pm}{a}$.
Given b. $\curvearrowright$

$$Wf(\cdot, b) = Df\left(b, \frac{\omega_0^\pm}{\cdot}\right)$$

corresponds to the frequency distribution at b.

4.3.3 Approximation Properties

We look for a system of classification for the set of all wavelets which are important for signal processing:
$\Longrightarrow$ classification by high frequency behavior
$\Longrightarrow$ looking for properties of Wf as $|a| \to 0$

Definition 4.6 *ψ is called a wavelet of order $N \in \mathbb{N}$, if*

1. *the mean value and the first $N-1$ moments of ψ vanish:* $\int_{\mathbb{R}} t^k \psi dt = 0,\ 0 \le k \le N-1$
2. *the N-th moment is finite and not equal to zero:* $\int_{\mathbb{R}} t^N \psi dt = c < \infty$; $c \neq 0$

Theorem 4.8 $f \in H^s$; $s \in \mathbb{R}$; *Let $\psi \in \mathbb{L}_2(\mathbb{R})$ be a wavelet of order N ;*
$\mu = \frac{(-1)^N}{N!} \cdot c;\ \ c \in \mathbb{R}$

$$\Longrightarrow \left\| \frac{sgn^N(-a)}{|a|^{N+0.5}} \sqrt{c_\psi} Wf(a,.) - \mu f^{(N)}(.) \right\|_{H^{s+N}} \stackrel{a\to 0}{\longrightarrow} 0$$

This means that the high frequency behavior of two wavelet transforms with respect to different wavelets of the same order differ only in the factor $\gamma = \frac{c}{\sqrt{c_\psi} N!}$. The order of a wavelet defines the behavior of the wavelet transform in the case of $|a| << 1$. Analogous theorems are satisfied in case of wavelets with a compact domain.
Further: (proofs see [2])

Theorem 4.9 *The order of wavelets with a compact domain is finite.*

Theorem 4.10 $\exists$ *wavelets* $\psi \in S(\mathbb{R})$ | $\int_{\mathbb{R}} t^k \psi dt = 0 \quad \forall k \in \mathbb{N}_0$ *(example: MEYER-wavelet).*

In addition to this classification the order of a wavelet gives some information about decaying of the wavelet transform as $a \to 0$

Theorem 4.11 *Let $\psi \in \mathbb{L}_1(\mathbb{R})$ be a wavelet of order N;*

1.) $\int_{\mathbb{R}} t^k \psi dt = 0$ *for* $k = 0, 1, ..., N-1$; $\int_{\mathbb{R}} t^N \psi dt \in \mathbb{R}$;

2.) $f \in \mathbb{L}_2(\mathbb{R})$; $\exists k \in \{1, ..., N\} \mid f^{(k)} \in \mathbb{L}^\infty(\mathbb{R})$

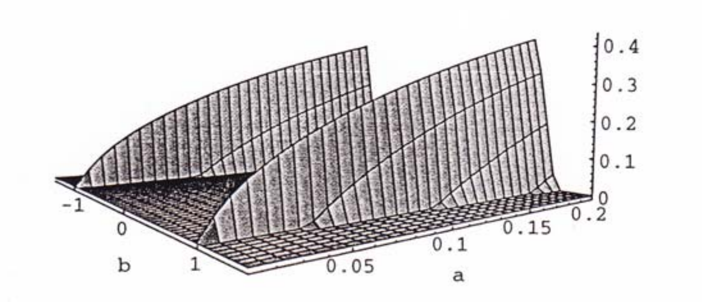
$$\Longrightarrow |Wf(a,b)| \leq \|Wf(a,.)\|_{\mathbb{L}^\infty} = O(|a|^{k+0.5}); \quad a \to 0; \quad \textit{for almost all } b \in \mathbb{R}.$$

That means: in high frequencies the wavelet transform falls the faster the smoother the transformed function and the smoother the transforming wavelet. The number of vanishing moments of the wavelet limits the accessible decay rate. If all the moments of a wavelet vanish then only the function f defines the decay rate.
$\Longrightarrow$ Choose wavelets with an order as high as possible.
There are many analogous theorems for this fact. An interesting example is the following:

Definition 4.7 *An isolated saltus b of the r^{th} derivative of f with $f^{(r)}(b+0) - f^{(r)}(b-0) = \delta$ is called "r-Knackpunkt"(crucial point of order r).*

Theorem 4.12 *Let ψ be a wavelet of order N with compact domain, $f \in \mathbb{L}_2(\mathbb{R})$ has a crucial point at b, $r < N$.*
Therefore: $Wf(a,b) = |a|^{r+0.5}(C\delta + o(1))$; $a \to 0$; $C \neq C(f)$

Example 4.8 *Wavelet transform of the characteristic function over the interval* $[-1;1]$ *(by the HAAR wavelet, see* [2]*):*



Conclusions:

- If f is a smooth function then $Wf(a,b)$ tends to zero very quickly as $|a| \to 0$
- At crucial points you can find a lot of high frequencies because there exists detail of very small size.

- Additionally at these points the decay rate is smaller compared with other points $\Longrightarrow$ good detectability
- Therefore the data compression rate is very high: Only these values of $Wf(a,b)$ will be saved which are bigger than a threshold (We are not interested in small values of $Wf(a,b)$ because they are not needed for synthesis.)

5 Discrete Wavelet Transform

The understanding and correct interpretation of the continuous wavelet transform are starting points of our consideration. In practice there are two problems:

1. efficient calculation of the wavelet transform
2. efficient reconstruction of signals, i.e. the efficient calculation of the inverse wavelet transform.

First have a look at the 2^{nd} problem: Find discrete subsets $\{\begin{pmatrix} a_i \\ b_i \end{pmatrix}\} \subset D_{Wf}$, which are sufficient for the reconstruction of f. In this way we come to the theory of „frames" and the „multi scale analysis". These two things enable an efficient calculation both of the wavelet and of the inverse wavelet transform. Thus the first problem is solved too.
In the derivation of the Wavelet Transform we need a higher redundance in describing the elements of a Hilbert space than we have by an ONS. That's why frames were introduced, first by Richard Duffin and Albert Charles Schaeffer in 1952. Later in the 1980s frames were used in the theory of wavelets by Stephane Mallat, Ingrid Daubechies and Yves Meyer.
A frame is a generalisation of an ONS at least in an inner product space $\mathbb{X}$ or in our case in a separable Hilbert space. There a frame is a subset $\{e_\alpha \in \mathbb{X}\}_{\alpha \in I}$, such that the frame condition is satified:

$$A\,\|v\| \leq \|Tv\| \leq B\,\|v\| \qquad \forall v \in \mathbb{X}; \quad A, B \in \mathbb{R}; \quad 0 < A \leq B$$

with a special frame operator T. The constants A and B give some information about the redundance of the frame. If $A = B$ then the frame ist called tight and the frame operator is an isometry. An ONS is a tight frame with $A = B = 1$. In the case $\mathbb{X} = \mathbb{L}_2(\mathbb{R})$ the set $\{e_\alpha \in \mathbb{X}\}_{\alpha \in I}$ is a family of functions

$$h. = \{h_m(t) | m \in M, h_m(t) \in \mathbb{L}_2(\mathbb{R})\}$$

with $Tf(m) = (f, h_m(t))$ for $f(t) \in \mathbb{L}_2(\mathbb{R}),\ m \in M$, such that

$$\|Tf\|^2 = \int_M |Tf(m)|^2\, d\mu(m)$$

and

1. $Tf(m)$ is $\mu-$measurable $\forall f(t) \in \mathbb{L}_2(\mathbb{R})$

2. $A\,\|f(t)\| \leq \|Tf(m)\| \leq B\,\|f(t)\| \qquad \forall f(t) \in \mathbb{L}_2(\mathbb{R}); \quad A, B \in \mathbb{R}; \quad 0 < A \leq B.$

In the theory of wavelets we choose $M := \mathbb{R}^2_- := \{(a,b) \mid \; a, b \in \mathbb{R};\; a \neq 0\}$ with the measure $d\mu = \dfrac{dadb}{a^2}$.
With the mother wavelet ψ we generate the family
$\psi. := \{\psi_{ab} \mid \psi_{ab} = \dfrac{1}{\sqrt{c_\psi}} \dfrac{1}{|a|^{0.5}} \psi(\dfrac{t-b}{a}); \quad a, b \in \mathbb{R};\; a \neq 0\}$

Then the frame operator corresponding to the family ψ is the wavelet transform, i.e. $Tf(a,b) = (f, \psi_{ab}) = Wf(a,b)$. The wavelet transform is an isometry and therefore we obtain the following formula

$$\|Wf(a,b)\|^2_{\mathbb{H}} = \|f\|^2_{\mathbb{L}_2}$$

and theorem 5.5: $\psi.$ is a tight frame with bound 1 for every wavelet ψ.
We also get the formula of the inverse wavelet transform

$$f = \int_{\mathbb{R}^2_-} Wf(a,b)\psi_{ab} \frac{dadb}{a^2} \qquad \forall f \in \mathbb{L}_2(\mathbb{R}).$$

for the synthesis of $f \in \mathbb{L}_2(\mathbb{R})$ analogous to chapter 4.

5.1 Wavelet Frames

We recall: A subset of a Hilbert space $\mathbb{X}$ is called a complete system if every element $x \in \mathbb{X}$ can be approximated arbitrarily well by linear combinations of elements of this system. Now we consider overcomplete systems, i.e. when removal of one or more elements from the system results in a complete system. In other areas of research, for example in signal processing and function approximation, overcompleteness can help to achieve a more stable, more robust, or more compact decomposition than using a basis. Now we want to use overcomplete frames for finding a homogeneous theory of continuous and discrete wavelet transform using functional analysis. In German there is no special word for „frame". In science a frame is a structural system that supports other components of a scientific construction.

Definition 5.1 *Frame:* $a. := \{a_i \mid a_i \in \mathbb{X};\; i \in I\}$, *such that there is not any* $x \in \mathbb{X}$, $x \neq \mathbb{O}$ *with* $(x, a_i) = 0 \quad \forall i$.

$a.$ is redundant, because the elements a_i must be neither linear independent nor orthogonal.

5.1.1 Geometrical Interpretation - Introduction

First we consider a Hilbert space $\mathbb{X}$ of finite dimensions,
$dim\mathbb{X} = n; \quad a_1, ..., a_r \in \mathbb{X}; \quad r > n$
Let T be an operator which maps from $\mathbb{X}$ to $\mathbb{Y} = \mathbb{C}^r$ such that:
$(Tx)_j = (x, a_j)$ for $1 \leq j \leq r$.
Furthermore let $\{e_1, ...e_r\}$ be a basis of $\mathbb{C}^r \quad \Longrightarrow \quad Tx = \sum_{j=1}^{r}(x, a_j)e_j$.
$\mathbb{U} = im(T) = \{Tx \mid x \in \mathbb{X}\}; \quad \dim \mathbb{U} \leq n \quad \Longrightarrow \quad \mathbb{U} \subset \mathbb{Y}$.
Problems:

- Can we define an element $x \in \mathbb{X}$ uniquely by $y = Tx \in \mathbb{Y}$?
- If the answer is yes then the question is: How to calculate x from y?

$\mathbb{Y}$ is a Hilbert space with $(y, z) = \sum_{k=1}^{r} y_k \overline{z}_k$. We consider the adjoint operator T^* : $\mathbb{Y} \to \mathbb{X}$, defined by

$$(x, T^*y)_{\mathbb{X}} = (Tx, y)_{\mathbb{Y}} \qquad \forall x \in \mathbb{X},\ \forall y \in \mathbb{Y}$$

Then:

$$\begin{aligned} (x, T^*e_j) &= (Tx, e_j) \\ &= component\ number\ j\ of\ Tx \\ &= (x, a_j) \quad \forall x \in \mathbb{X} \end{aligned}$$

$\curvearrowright$

$$T^*e_j = a_j \qquad 1 \leq j \leq r \qquad (5.1.1)$$

$G = T^*T : \mathbb{X} \to \mathbb{X}$ is called Gram-Operator ($\neq$ Gram Matrix!)

$$Gx = T^*Tx = T^*\left(\sum_{j=1}^{r}(x, a_j)\, e_j\right) = \sum_{j=1}^{r}(x, a_j)\, T^*e_j = \sum_{j=1}^{r}(x, a_j)\, a_j \quad (5.1.2)$$

Furthermore: $\ker T = \ker G$, because

1. $Tx = 0 \quad \Longrightarrow \quad Gx = 0$
2. $\|Tx\|^2 = (Tx, Tx) = (T^*Tx, x) = (Gx, x) \quad \Longrightarrow \quad Gx = 0 \quad \Longrightarrow \quad Tx = 0.$

This implies: T is bijective $\quad \Longleftrightarrow \quad G$ is regular.
Now we have a look at G:

- $G = G^*$, because $(x, Gu) = (x, T^*Tu) = (Tx, Tu) = (T^*Tx, u) = (Gx, u)$
- $\curvearrowright$ The eigenvalues of G are real numbers and $\lambda(x,x) = (Gx,x) = \|Tx\|^2 \geq 0$
- $\curvearrowright$ $0 \leq A := \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n =: B$
- $\curvearrowright$ $\exists$ a basis $\{e_1, e_2, ..., e_n\}$, which diagonalizes G. In this basis we get:

$$Gx = \begin{pmatrix} \lambda_1 x_1 \\ \vdots \\ \lambda_n x_n \end{pmatrix} \qquad (5.1.3)$$

- $\curvearrowright$

$$\|Tx\|^2 = (Gx, x) = \sum_{k=1}^{n} \lambda_k |x_k|^2 \begin{cases} \geq A \|x\|^2 \\ \leq B \|x\|^2 \end{cases} .$$

Theorem 5.1 *$a. = \{a_1, ..., a_n\} \subset \mathbb{X}$ is a frame* $\iff$ $\exists\ B \geq A > 0 \mid$

$$A \|x\|^2 \leq \|Tx\|^2 \leq B \|x\|^2 \ \ \forall x \in \mathbb{X}.$$

A, B are called the bounds of the frame. If $A = B$, the frame is called „tight" and $\|Tx\|^2 = A \|x\|^2$. I this sense T is essentially an isometry.

Example 5.1 *Let $a. = \{a_1, ..., a_n\}$ be an ONS of $\mathbb{X}$. Then*

$$\|Tx\|^2 = \sum_{j=1}^{n} |(x, a_j)|^2 = \|x\|^2 \quad \forall x \in \mathbb{X}.$$

Thus the ONS is a tight frame with the bound $A = 1$. This gives information about the size of redundance: $A = 1$ means no redundance.

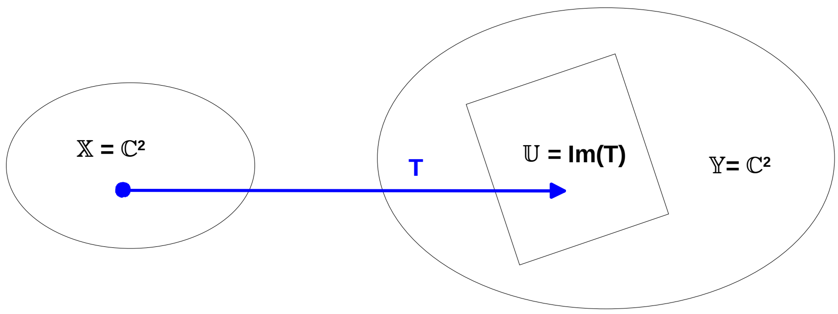
Example 5.2 $\mathbb{X} = \mathbb{C}^2$; $r \geq 2$; $\quad a_j = \frac{1}{\sqrt{2}} \begin{pmatrix} \exp(\frac{2\pi i j}{r}) \\ \exp(-\frac{2\pi i j}{r}) \end{pmatrix}$ $\quad j = 0, ..., r-1$. *Therefore we define the frame operator $T : \mathbb{X} \to \mathbb{C}^r$ for $x \in \mathbb{X} = \mathbb{C}^2$, $j = 0, ..., r-1$ by*

$$\begin{aligned} (Tx)_j &= (x, a_j) = \frac{1}{\sqrt{2}} \left(x_1 \exp(-\frac{2\pi i j}{r}) + x_2 \exp(\frac{2\pi i j}{r}) \right) \\ &= \frac{1}{\sqrt{2}} \left(x_1 \omega_r^{-j} + x_2 \omega_r^{j} \right) \quad \textit{with } \omega_r = \exp(\frac{2\pi i}{r}). \end{aligned}$$

Therefore ω_r is the r^{th} root of unity. $\curvearrowright$

$$\begin{aligned}
\|Tx\|^2 &= \frac{1}{2}\sum_{j=0}^{r-1}\left(x_1\omega_r^{-j}+x_2\omega_r^{j}\right)\left(\overline{x}_1\omega_r^{j}+\overline{x}_2\omega_r^{-j}\right)\\
&= \frac{1}{2}\sum_{j=0}^{r-1}\left(|x_1|^2+|x_2|^2+x_1\overline{x}_2\omega_r^{-2j}+\overline{x}_1x_2\omega_r^{2j}\right)\\
&= \frac{1}{2}\sum_{j=0}^{r-1}\|x\|^2+\frac{1}{2}x_1\overline{x}_2\sum_{j=0}^{r-1}\left(\omega_r^{-2}\right)^j+\frac{1}{2}\overline{x}_1x_2\sum_{j=0}^{r-1}\left(\omega_r^{2}\right)^j\\
&= \frac{r}{2}\|x\|^2+\frac{1}{2}x_1\overline{x}_2\frac{(\omega_r^{-2})^r-1}{\omega_r^{-2}-1}+\frac{1}{2}\overline{x}_1x_2\frac{(\omega_r^{2})^r-1}{\omega_r^{2}-1}\\
&= \frac{r}{2}\|x\|^2+\frac{1}{2}x_1\overline{x}_2\frac{(\omega_r^{r})^{-2}-1}{\omega_r^{-2}-1}+\frac{1}{2}\overline{x}_1x_2\frac{(\omega_r^{r})^{2}-1}{\omega_r^{2}-1}\\
&= \frac{r}{2}\|x\|^2.
\end{aligned}$$

Thus $a.$ is a tight frame with $A=\frac{r}{2}$, i.e. with $r=2$ we get $A=1$. That means in $\mathbb{C}^2$ two vectors of such kind are a basis in $\mathbb{X}$ and $\mathbb{U}$:



If $r>2$, then the number $\frac{r}{2}$ gives information about the redundancy of the set $a.$

Conclusion 5.1 *If $a.$ is a frame, then G is regular and T bijectiv, i.e. T is in principle invertible.*

Thus the second problem left over: Calculation of x from $Tx=y$.
Let $a.=\{a_1,...,a_n\}$ be a frame, G the corresponding Gram operator. $G^{-1}:\mathbb{X}\to\mathbb{X}$ exists because $G=T^*T$ is regular.
Now we consider the mapping $S=G^{-1}T^*:\mathbb{Y}\to\mathbb{X}$

$$ST=G^{-1}T^*T=G^{-1}G=E_{\mathbb{X}}.$$

Thus S is the left inverse of T. If $a.$ is tight then we get, with $\|Tx\|^2 = A\|x\|^2$,

$$\begin{aligned}(Gx, x) &= (T^*Tx, x) = (Tx, Tx) = \|Tx\|^2 = A(x, x)\\ &= A(G^{-1}Gx, x)\end{aligned}$$

$\curvearrowright$

$$\begin{aligned}G^{-1} &= \frac{1}{A}E_{\mathbb{X}}\\ S &= \frac{1}{A}E_{\mathbb{X}}T^* = \frac{1}{A}T^*.\end{aligned}$$

I.e. we get the left inverse without calculation! Furthermore S is the right inverse of T, too because

$$P = TS = TG^{-1}T^* = T(T^*T)^{-1}T^* = TT^{-1}\left(T^*\right)^{-1}T^* = E_{\mathbb{Y}}.$$

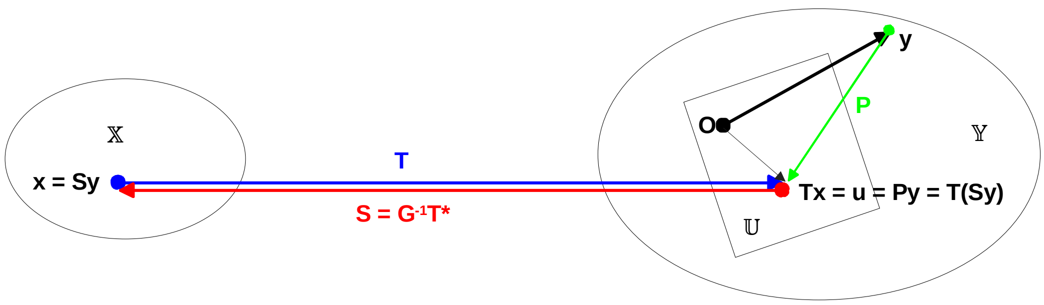
Therefore S is the inverse operator of T :

$$S = \frac{1}{A}T^*.$$

But we must have a look at the domains of S and T.
How can we invert the operator efficiently? What is the representation of y?

Theorem 5.2 *$P = TS$ is the orthogonal projection of $\mathbb{Y}$ to $\mathrm{Im}(T) = \mathbb{U}$. (proof: [1], p. 84)*

Interpretation:



I.e. for every $y \in \mathbb{Y}$ the preimage $x = Sy$ is the vector of $\mathbb{X}$, whose image Tx is mostly closed to y!! Therefore if $y \in \mathbb{U} \subset \mathbb{Y}$, then $x = Sy$ is the vector of $\mathbb{X}$ with $Tx = y$.

Definition 5.2 *$\widetilde{a}. = \{\widetilde{a}_1, ..., \widetilde{a}_r \mid \widetilde{a}_j = G^{-1}a_j \in X;\ 1 \leq j \leq r\}$ is called a dual frame corresponding to $a.$*

Notation 5.1 *If a. is tight, then:*

$$\begin{aligned}(\widetilde{T}x)_j &= (x, \widetilde{a_j}) = (x, G^{-1}a_j) = (x, \frac{1}{A}E_x a_j) = (x, \frac{1}{A}a_j) \quad \curvearrowright \\ \widetilde{a_j} &= \frac{1}{A}a_j\end{aligned}$$

Theorem 5.3 *Let a. be a frame with the bounds $0 < A \leq B$. $\widetilde{a}.$ is the corresponding dual frame. Then:*

1. $x = \sum_{j=1}^{r}(x, a_j)\widetilde{a_j} \quad \forall x \in \mathbb{X}$

2. $Sy = \sum_{j=1}^{r} y_j\widetilde{a_j} \quad \forall y \in \mathbb{Y}$

3. $\widetilde{a}.$ *is a frame with* $\frac{1}{A} \geq \frac{1}{B} > 0$

4. *a. is the dual frame to $\widetilde{a}.$, i.e.* $x = \sum_{j=1}^{r}(x, \widetilde{a_j})a_j \quad \forall x \in \mathbb{X}$

5. *Let $x = \sum_{j=1}^{r}\xi_j\widetilde{a_j}$ be an arbitrary representation of x as a linear combination of $\widetilde{a_j}$*

   $\implies \quad \sum_{j=1}^{r}|\xi_j|^2 \geq \sum_{j=1}^{r}|(x, a_j)|^2$

Proof:

1.

$$\begin{aligned}x &= G^{-1}Gx \overset{(5.1.2)}{=} G^{-1}\sum_{j=1}^{r}(x, a_j)a_j = \sum_{j=1}^{r}(x, a_j)G^{-1}a_j \\ &= \sum_{j=1}^{r}(x, a_j)\widetilde{a_j}\end{aligned}$$

2.

$$\begin{aligned}Sy &= G^{-1}T^*\sum_{j=1}^{r}y_je_j = G^{-1}\sum_{j=1}^{r}y_jT^*e_j \overset{(5.1.1)}{=} G^{-1}\sum_{j=1}^{r}y_ja_j \\ &= \sum_{j=1}^{r}y_jG^{-1}a_j = \sum_{j=1}^{r}y_j\widetilde{a_j}\end{aligned}$$

3. *Let* $\widetilde{T}$ *be the frame operator to* $\widetilde{a}.$;
   G *is self adjoint, because* $G^* = (T^*T)^* = T^*T = G$.
   $\curvearrowright$ G^{-1} *is self adjoint, because* $(G^{-1})^* = (G^*)^{-1} = G^{-1}$.
   Thus
$$(\widetilde{T}x)_j = (x, \widetilde{a_j}) = (x, G^{-1}a_j) = (G^{-1}x, a_j) = (TG^{-1}x)_j$$
   $\curvearrowright$
$$\widetilde{T} = TG^{-1} \quad (5.1.4))$$
   $\curvearrowright$
$$\begin{aligned}\left\|\widetilde{T}x\right\|^2 &= \left\|TG^{-1}x\right\|^2 = (TG^{-1}x, TG^{-1}x) \\ &= (T^*TG^{-1}x, G^{-1}x) = (x, G^{-1}x)\end{aligned}$$
   We use an ONS $\{\widetilde{e_1}, ..., \widetilde{e_n}\}$ *which diagonalize* G *and* G^{-1}. $\curvearrowright$
$$\left\|\widetilde{T}x\right\|^2 = (x, G^{-1}x) \overset{(5.1.3)}{=} \sum_{i=1}^{n} \frac{1}{\lambda_i}|x_i|^2 \begin{cases} \geq \frac{1}{B}\|x\|^2 \\ \leq \frac{1}{A}\|x\|^2 \end{cases}$$

4. $$\widetilde{G} = \widetilde{T}^*\widetilde{T} \overset{(5.1.4)}{=} (TG^{-1})^*TG^{-1} = (G^{-1})^*T^*TG^{-1} = G^{-1}$$
   $\curvearrowright$
$$\widetilde{\widetilde{a_j}} = \widetilde{G}^{-1}\widetilde{a_j} = (G^{-1})^{-1}\widetilde{a_j} = (G^{-1})^{-1}G^{-1}a_j = a_j$$

5. $(\xi_1, ..., \xi_r)^T = y \in \mathbb{Y}$. *With 2. we get* $Sy = \sum_{j=1}^{r} y_j\widetilde{a_j} = x \in \mathbb{X}$. *Furthermore* $Tx = TSy = P_{\mathbb{U}}y$ *is the orthogonal projection from* $\mathbb{Y}$ *to* $\mathbb{U} = \mathrm{Im}(T)$ $\curvearrowright$
$$\sum_{j=1}^{r} |(x, a_j)|^2 = \|Tx\|^2 = \|P_{\mathbb{U}}y\|^2 \leq \|y\|^2 = \sum_{j=1}^{r} |\xi_j|^2 .$$
   If $y = P_{\mathbb{U}}y = Tx \in \mathbb{U}$ the inequality is an equation. ■

Interpretation:
By using frames $a.$ and the corresponding dual frames $\widetilde{a}.$ the element $x \in \mathbb{X}$ can be represented efficiently as the preimage of y. The „natural" representation 1. of the above theorem corresponds to the minimal „energy" of the coefficients.
In the next section we expand this theory to infinitely dimensional spaces.

5.1.2 Common Frames

Let $\mathbb{X}$ be a complex infinitely dimensional Hilbert space; M is a set of points.
For integral calculus we need a measure μ at $\mathbb{X}$ which allocate a volume to every measurable subset $E \subset M$.
Given the family $h. = \{h_m \mid m \in M,\ h_m \in \mathbb{X}\}$.

Definition 5.3 *Frame operator* $T : \mathbb{X} \longrightarrow \mathbb{C} : Tf(m) = (f, h_m); \quad f \in \mathbb{X}; \quad m \in M$ *with* $\|Tf\|^2 = \int_M |Tf(m)|^2 d\mu(m)$

Thus we get the dataset $\{Tf(m) \mid m \in M\}$, which saves information about $f \in \mathbb{X}$ by a „sensor" h_m.

Definition 5.4 *h. is a frame, if*

1. *Tf is $\mu-$mesasurable for every $f \in \mathbb{X}$ and if*
2. *there exists bounds A and B such that*
   $A\|f\|^2 \leq \|Tf\|^2 \leq B\|f\|^2 \qquad \forall f \in \mathbb{X}.$

Therefore the frame operator is linear, bounded (i.e. continuous) and invertable. But in the infinitely dimensional case an iterative algorithm is nessecary for calculating the inverse operator. It converges faster when the quotient $\frac{B}{A}$ is nearly equal to one :

Theorem 5.4 *Let h. be a frame with $B \geq A > 0$; $g \in \mathbb{X}$;*

$$\begin{aligned} f_0 &= 0 \\ f_{n+1} &= f_n + \frac{2}{A+B}(g - Gf_n); \quad n \geq 0 \end{aligned}$$

Then:

$$\lim_{n\to\infty} f_{n+1} = G^{-1}g.$$

Proof.

$$\begin{aligned} f_{n+1} &= \frac{2}{A+B}g + \left(E_{\mathbb{X}} - \frac{2}{A+B}G\right) f_n \\ &= \frac{2}{A+B}g + \qquad\qquad Rf_n \\ &= \widetilde{R}f_n \end{aligned}$$

Because h. is a frame we get:

$$A\|f\|^2 \leq \|Tf\|^2 = (Tf, Tf) = (T^*Tf, f) = (Gf, f) \leq B\|f\|^2$$

$\curvearrowright$

$$AE_{\mathbb{X}} \leq G \leq BE_{\mathbb{X}} \qquad (5.1.5)$$

$$\begin{aligned}\left\| G - \frac{A+B}{2} E_{\mathbb{X}} \right\| &\leq \left| \|G\| - \left\| \frac{A+B}{2} E_{\mathbb{X}} \right\| \right| \\ &\leq \left| B - \frac{A+B}{2} \right| = \frac{B-A}{2}\end{aligned}$$

$$\|R\| = \left\| \frac{2}{A+B} G - E_{\mathbb{X}} \right\| = \frac{2}{A+B} \left\| G - \frac{A+B}{2} E_{\mathbb{X}} \right\| \leq \frac{B-A}{B+A} < 1$$

Thus R is a contractive mapping $\curvearrowright$

$$\begin{aligned}\left\| \widetilde{R} z_n - \widetilde{R} f_n \right\| &= \left\| \frac{2}{A+B} g + R z_n - \frac{2}{A+B} g - R f_n \right\| \\ &\leq \|R\| \, \|z_{n-} f_n\|\end{aligned}$$

Therefore $\widetilde{R}$ is a contractive mapping, too and by the Banach fixpoint theorem there exists one and only one fixpoint:

$$\begin{aligned} f &= \lim_{n \to \infty} f_n \\ &= \frac{2}{A+B} g + \left(E_X - \frac{2}{A+B} G \right) f \\ &= f + \frac{2}{A+B} (g - Gf) \end{aligned}$$

$\curvearrowright$

$$g = Gf \quad \curvearrowright \quad f = G^{-1} g$$

■

Notation 5.2 *$A = B$ implies by (5.1.5)*

$$G = BE_{\mathbb{X}} \quad \curvearrowright \quad G^{-1} = \frac{1}{B} E_{\mathbb{X}},$$

i.e. an iterative algorithm is not necessary.

Application of the frame concept to the wavelet transform:
$\mathbb{X} := \mathbb{L}_2(\mathbb{R})$: space of time signals $f(t)$; $\dim \mathbb{X} = \infty$
$M := \mathbb{R}^2_- := \{(a, b) \mid \; a, b \in \mathbb{R}; \; a \neq 0\}$ with the measure $d\mu = \dfrac{dadb}{a^2}$
$\mathbb{Y} := \mathbb{L}_2(\mathbb{R}^2_-, d\mu) = \mathbb{H}$
Choose a mother wavelet ψ and generate the family

$\psi. := \{\psi_{ab} \mid \psi_{ab} = \frac{1}{\sqrt{c_\psi}} \frac{1}{|a|^{0.5}} \psi(\frac{t-b}{a}); \quad a, b \in \mathbb{R}; \ a \neq 0\}$

$Tf(a,b) = (f, \psi_{ab}) = Wf(a,b)$, i.e. the frame operator corresponding to the family ψ is the wavelet transform. The wavelet transform is an isometry and therefore we obtain the following formula and theorem

$$\|Wf(a,b)\|^2_{\mathbb{H}} = \|f\|^2_{\mathbb{L}_2}$$

Theorem 5.5 *$\psi.$ is a tight frame with bound 1 for every wavelet ψ.*

Thus $G^{-1} = E_{\mathbb{X}}; \quad \widetilde{\psi}. = \psi.$
Analogous to the considerations above we obtain

$$f = \int\limits_{\mathbb{R}^2_-} Wf(a,b)\psi_{ab} \frac{dadb}{a^2} \qquad \forall f \in \mathbb{L}_2(\mathbb{R}).$$

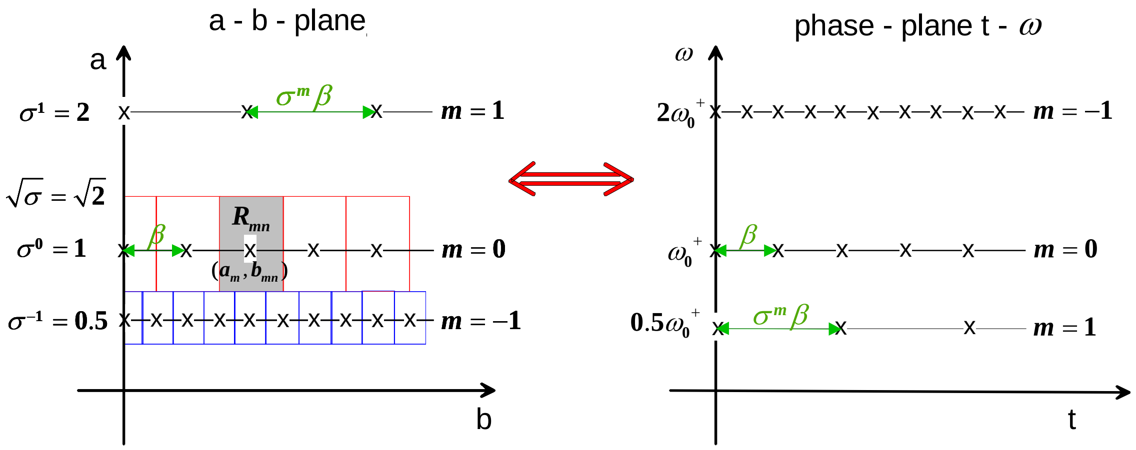
for the synthesis of $f \in \mathbb{L}_2(\mathbb{R})$ by the values $(Tf)_j = (f, a_j)$. This formula is the formula of the inverse wavelet transform.

5.2 Discrete Wavelet Transform

Problem: Is it necessary to know $Wf(a,b)$ at every point $(a,b) \in \mathbb{R}^2_-$ for synthesis of f?
We know from the Fourier transform theory that the sampling theorem from SHANNON gives the information that complete reconstruction of a bandlimited signal f is possible from a discrete dataset $f(kT), \ k \in \mathbb{Z}$.
Now we are looking for an analogon to the wavelet transform. But we don't discuss generally, for which subsets of $\mathbb{R}^2_-$ the reconstruction is possible. We only consider the set

$$M := \{(a_m, b_{mn}) \mid m, n \in \mathbb{Z}; \quad a_m = \sigma^m; \quad b_{mn} = n\sigma^m\beta; \quad \sigma > 1, \quad \beta > 0\}.$$

A common version is $\sigma = 2$. σ is called the zoom factor, β is called the basic step.
We only show set M for $a > 0$:



For the mother wavelet ψ we consider the family $\psi_.$ corresponding to this set:

$$\begin{aligned}
\psi_. &= \{\psi_{mn} \mid m, n \in \mathbb{Z}\} \quad \mathit{with} \\
\psi_{mn} &= \frac{1}{\sqrt{C_\psi}\sqrt{a_m}} \psi\left(\frac{t - b_{mn}}{a_m}\right) \\
&= \frac{1}{\sqrt{C_\psi}\sqrt{\sigma^m}} \psi\left(\frac{t - n\sigma^m \beta}{\sigma^m}\right) \\
&= \frac{1}{\sqrt{C_\psi}} \sigma^{-\frac{m}{2}} \psi\left(\sigma^{-m} t - n\beta\right)
\end{aligned}$$

In the phase space $t - \omega$ the functions ψ_{mn} are located around the point

$$\left(b_{mn}; \frac{\omega_0^\pm}{a_m}\right) = \left(n\sigma^m\beta; \sigma^{-m}\omega_0^\pm\right).$$

With increasing frequency ω the points lie closer with respect to the $t-$coordinate which implies the zoom property.
Now we are looking for properties of this family, i.e. conditions for ψ, σ, β, such that $\psi_.$ is a frame. To do this we need a measure on M analogous to the above. The point (a_m, b_{mn}) represents the rectangle R_{mn}. R_{mn} has the width $\sigma^m\beta$ and the height

$\sigma^m\sqrt{\sigma} - \sigma^m\frac{1}{\sqrt{\sigma}}$. By the HAAR measure we obtain

$$\begin{aligned}
\mu(R_{mn}) &= \sigma^m\beta\int_{\sigma^m/\sqrt{\sigma}}^{\sigma^m\sqrt{\sigma}} 1\frac{da}{a^2} \\
&= \sigma^m\beta\left(-\frac{1}{a}\right)_{\sigma^m/\sqrt{\sigma}}^{\sigma^m\sqrt{\sigma}} \\
&= \sigma^m\beta\left(-\sigma^{-m-0.5} + \sigma^{-m+0.5}\right) \\
&= \beta\left(-\sigma^{-0.5} + \sigma^{+0.5}\right) \\
&= \frac{\beta}{\sqrt{\sigma}}\left(\sigma - 1\right).
\end{aligned}$$

Notation 5.3 *With respect to the „normal" measure dadb these rectangles are not of equal size!*

$$M \sim \mathbb{Z}^2$$

Then the counting measure # can be assigned to M. This measure assigns the value $\mu(R_{mn}) = const.$ to every point of M. $\curvearrowright$ Then $\mathbb{Y} = \mathbb{L}_2(\mathbb{R}^2_-, d\mu)$ is equivalent to $\mathbb{Y} = \mathbf{l}^2(\mathbb{Z}^2)$.
If $\psi.$ is a frame, then the corresponding frame operator T

$$Tf(m,n) = (f, \psi_{mn}) = Wf(a_m, b_{mn})$$

is connected to the wavelet ψ by the wavelet transform.

Definition 5.5 *Given $\sigma > 1$; ψ is called acceptable if*

1.
$$\exists\ \alpha > 0;\ \rho > 0;\ C \in \mathbb{R}\ |\ \left|\widehat{\psi}(\omega)\right| \leq \begin{cases} C|\omega|^\alpha & |\omega| \leq 1 \\ \frac{C}{|\omega|^{1+2\rho}} & |\omega| > 1 \end{cases}$$

2.
$$\exists\ A' > 0\ |\ \sum_{m=-\infty}^{\infty}\left|\widehat{\psi}(\sigma^m\omega)\right|^2 \geq A' \qquad 1 \leq |\omega| \leq \sigma$$

The constants $\alpha,\ \rho,\ A'$ and C are called the parameters of ψ.
For example Condition 1 is satisfied if $t\psi' \in \mathbb{L}_1$ and ψ' is of bounded variation with $\alpha = 1;\ \rho = 0.5$.
For example Condition 2 is satisfied if ψ is of finite order N.
Then there is $\widehat{\psi} = \gamma\omega^N + O\left(N+1\right)$
$\curvearrowright \quad \widehat{\psi}(\omega) \neq 0 \quad for \quad 0 < |\omega| < h$ ([1])

Theorem 5.6 *$\sigma > 1$, ψ is an acceptable wavelet with the parameters α, ρ, A', and C. Then there exists constants β_0, B', and C', such that: If $\beta < \beta_0$, then $\psi. = \{\psi_{mn}\}$ is a frame with constants*

$$A = \frac{2\pi}{\beta c_\psi}\left(A' - C'\beta^{1+\rho}\right); \quad B = \frac{2\pi}{\beta c_\psi}\left(B' - C'\beta^{1+\rho}\right).$$

Proof: [1]
Then it is possible to reconstruct $f \in \mathbb{L}_2(\mathbb{R})$ with respect to this frame.
$M \sim \mathbb{Z}^2 \subset \mathbb{R}^2$, i.e. values of the wavelet transform at these points of M are sufficient for reconstruction! (But it is still a countable infinite set.)
Corresponding to theorem 5.3 the dual frame $\widetilde{\psi}$ is needed to compute f :

$$\widetilde{\psi}_{mn} = G^{-1}(\psi_{mn}).$$

Because ψ is not tight in general the calculation of $\widetilde{\psi}_{mn}$ is troublesome. The functions $\widetilde{\psi}_{mn}$ are not dilated or translated copies of a single $\widetilde{\psi}$. That's why it is better to look for a tight frame. Then

$$\widetilde{\psi}_{mn} = \frac{1}{A}\psi_{mn}.$$

Theorem 5.7 *The Fourier transform $\widehat{\psi}$ of the wavelet ψ has a compact domain (band-limited signal!) in the interval $I = [\omega, \omega']$ with $\omega' > \omega > 0$. Given*

$$\sum_{m=-\infty}^{\infty} |\widehat{\psi}(\sigma^m \xi)|^2 \equiv A' > 0; \quad 1 \leq \xi \leq \sigma.$$

Then $\psi. = \{\psi_{mn}\}$, with the zoom factor σ and the basic step $\beta \leq \dfrac{2\pi}{\omega' - \omega}$, is a tight frame. (proof: [1])

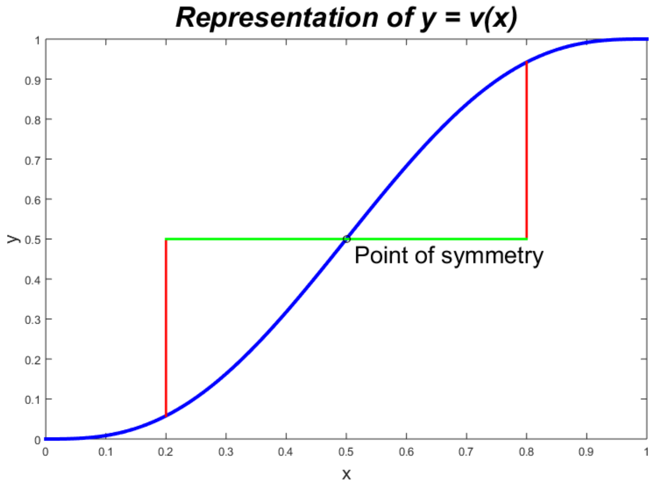
Therefore the reconstruction of f by elements of the set $Tf(m, n)$ can be done without problems because $G^{-1} = \frac{1}{A}E_x$. Furthermore

Theorem 5.8 *A tight frame $\psi. = \{\psi_{mn}\}$ to the wavelet ψ with the bounds $A = B = 1$ generates an ONS of $\mathbb{L}_2(\mathbb{R})$ if $\|\psi\|_{L_2} = 1$. (proof: [2])*

These two theorems are used for generating wavelets with the required properties concerning the connection to a frame, for example for generating the DAUBECHIES-GROSSMANN-MEYER-Wavelet or the MEYER-Wavelets. The following example shows this.

Example 5.3 [1] *Given an auxiliary function* $\nu = \nu(x) : \mathbb{R} \to \mathbb{R};\ \nu \in \mathbb{C}^k(\mathbb{R});\ k \geq 0$ *with* $\nu(x) = \begin{cases} 0 & for \quad x \leq 0 \\ 1 & for \quad x \geq 1 \end{cases}$ *, for example*

$$\nu(x) \quad = \quad \begin{cases} 0 & x \leq 0 \\ 10x^3 - 15x^4 + 6x^5 & 0 \leq x \leq 1 \\ 1 & x \geq 1 \end{cases}$$



Furthermore:

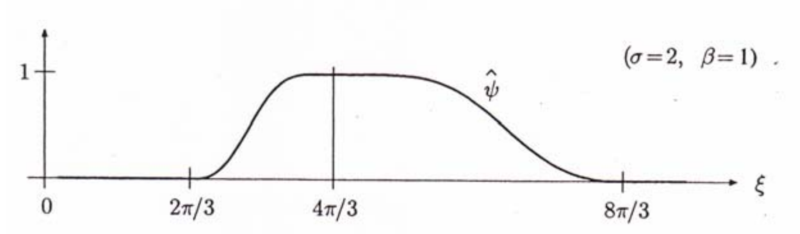
$$\begin{aligned} \nu(1-x) \quad &= \quad \begin{cases} 0 & for & 1-x \leq 0 & \curvearrowright & x \geq 1 \\ 1 & for & 1-x \geq 1 & \curvearrowright & x \leq 0 \\ 10(1-x)^3 - 15(1-x)^4 + 6(1-x)^5 & = & 1 - 10x^3 + 15x^4 - 6x^5 & for & 0 \leq x \leq 1 \end{cases} \\ &= \quad 1 - \nu(x) \qquad \curvearrowright \qquad \nu(x) = 1 - \nu(x) \end{aligned}$$

$k = 2$, *because* $\nu'(x) = 30x^2 - 60x^3 + 30x^4$ *and* $\nu''(x) = 60x - 180x^2 + 120x^3$ *imply* $\nu''(0) = 0;\ \nu''(1) = 0$.
$\nu(x)$ *is symmetric with respect to* $P = (0.5; 0.5)^T : \nu(0.5 - x) = 1 - \nu(0.5 + x)$.
Given $\sigma > 1;\ \beta > 0$. *We generate* $I = [\xi, \xi']$ *corresponding to the preconditions of the theorem 5.7 by*

$$\xi = \frac{2\pi}{(\sigma^2 - 1)\beta} > 0 \text{ and } \xi' = \sigma^2\xi > \xi > 0.$$

Then we construct ψ by $\widehat{\psi}$ with the domain I :

$$\widehat{\psi}(\omega) = \sqrt{A'} \begin{cases} \sin\left(\frac{\pi}{2}\nu(\frac{\omega-\xi}{\sigma\xi-\xi})\right) & for \quad \xi \le \xi \le \sigma\xi \\ \cos\left(\frac{\pi}{2}\nu(\frac{\omega-\sigma\xi}{\sigma^2\xi-\sigma\xi})\right) & for \quad \sigma\xi \le \omega \le \sigma^2\xi = \xi' \\ 0 & otherwise \end{cases}$$



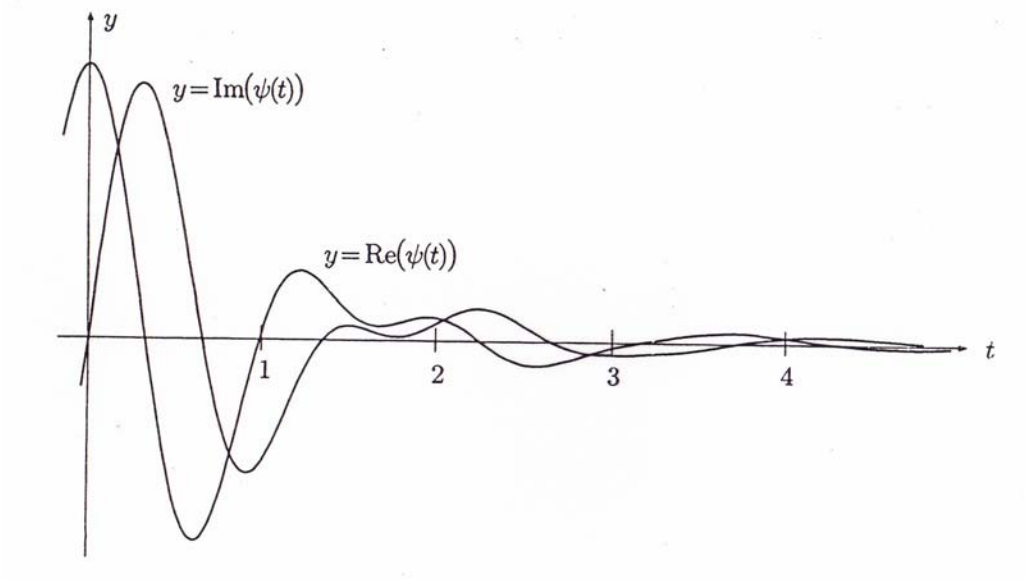
A' is calculated later by normalising ψ. Furthermore $\sigma = 2;\ \beta = 1$.
$\curvearrowright \quad \xi = \frac{2\pi}{3\beta} = \frac{2}{3}\pi; \quad \xi' = \frac{8}{3}\pi; \quad \xi' - \xi = 2\pi.$
We have to show

$$\sum_{k\in\mathbb{Z}} \left|\widehat{\psi}\left(\sigma^k\omega\right)\right|^2 = A'.$$

If $\xi < 0$ then $\sigma^k\xi < 0 \quad \curvearrowright \quad$ corresponding summands $\equiv 0$.
If $\xi \ge 0$ then $\exists\ k^* \in \mathbb{Z} \mid \xi \le \sigma^{k^*}\omega \le \sigma\xi \quad \wedge \quad \sigma\xi \le \sigma^{k^*+1}\omega \le \sigma^2\xi$. I.e., the sum goes from k^* to k^*+1 only.

$$\begin{aligned} \sum_{k\in\mathbb{Z}} \left|\widehat{\psi}\left(\sigma^k\omega\right)\right|^2 &= A'\left(\sin^2\left(\frac{\pi}{2}\nu(\frac{\sigma^{k^*}\omega-\xi}{\sigma\xi-\xi})\right) + \cos^2\left(\frac{\pi}{2}\nu(\frac{\sigma^{k^*+1}\omega-\sigma\xi}{\sigma^2\xi-\sigma\xi})\right)\right) \\ &= A'\left(\sin^2\left(\frac{\pi}{2}\nu(\frac{\sigma^{k^*}\omega-\xi}{\sigma\xi-\xi})\right) + \cos^2\left(\frac{\pi}{2}\nu(\frac{\sigma\left(\sigma^{k^*}\omega-\xi\right)}{\sigma\left(\sigma\xi-\xi\right)})\right)\right) \\ &= A' \end{aligned}$$

The inverse Fourier transform generates ψ, the DAUBECHIES-GROSSMANN-MEYER wavelet with $\sigma = 2;\ \beta = 1$.



furthermore [2]

$$A' = \frac{1}{\ln \sigma} = \frac{1}{\ln 2}; \quad A \geq \frac{\pi A'}{\beta} = \frac{\pi}{\ln 2}$$

Example 5.4 *The MEYER wavelet with*

$$\widehat{\psi}(\omega) = \tfrac{1}{\sqrt{2\pi}} e^{i\frac{\omega}{2}} (\mu(\omega) + \mu(-\omega))$$

and the defining function $\mu(\omega)$ *for* $\sigma = 2$ *and* $\beta = 1$

$$\mu(\omega) = \begin{cases} \sin(\frac{\pi}{2}\nu(\frac{3\omega}{2\pi} - 1)) & for \quad \frac{2\pi}{3} \leq \omega \leq \frac{4\pi}{3} \\ \cos(\frac{\pi}{2}\nu(\frac{3\omega}{4\pi} - 1)) & for \quad \frac{4\pi}{3} \leq \omega \leq \frac{8\pi}{3} \\ 0 & otherwise \end{cases}$$

generates a tight frame with the bound $A = 1$*. Furthermore* $\|\psi\|_{L_2}^2 = 1$*. Thus the frame* $\psi. = \{\psi_{mn}\}$ *is an ONS of* $\mathbb{L}_2(\mathbb{R})$.

Proof: Given $I=[\frac{2}{3}\pi;\frac{4}{3}\pi],\ J=[\frac{4}{3}\pi;\frac{8}{3}\pi]$. Because of $\|\psi\|^2_{\mathbb{L}_2}=\left\|\widehat{\psi}\right\|^2_{\mathbb{L}_2}$ we obtain:

$$\begin{aligned}
\|\psi\|^2_{\mathbb{L}_2} &= \frac{1}{2\pi}\left[\int_{|\omega|\in I}\sin^2\left(\frac{\pi}{2}\nu(\frac{3}{2\pi}|\omega|-1\right)d\omega+\int_{|\omega|\in J}\cos^2\left(\frac{\pi}{2}\nu(\frac{3}{4\pi}|\omega|-1)\right)d\omega\right]\\
&= \frac{1}{2\pi}\left[2\cdot\frac{2\pi}{3}\int_0^1\sin^2\left(\frac{\pi}{2}\nu(x)\right)dx+2\cdot\frac{4\pi}{3}\int_0^1\cos^2\left(\frac{\pi}{2}\nu(x)\right)dx\right]\\
&= \frac{2}{3}\int_0^1\left(1-\cos^2(\frac{\pi}{2}\nu(x))\right)dx+\frac{4}{3}\int_0^1\cos^2\left(\frac{\pi}{2}\nu(x)\right)dx\\
&= \frac{2}{3}+\frac{2}{3}\int_0^1\cos^2\left(\frac{\pi}{2}\nu(x)\right)dx
\end{aligned}$$

ν is symmetric with respect to $P=(0.5;0.5)^T$:
$\nu(y)=\nu(0.5+x)=1-v(\frac{1}{2}-x)$ with $y\in[0.5;1]$ and $x\in[0;0.5]$,
$\cos(\frac{\pi}{2}-z)=\sin(z)$ and a substitution $\alpha=\frac{1}{2}-x$. Therefore we get

$$\begin{aligned}
\int_0^1\cos^2\left(\frac{\pi}{2}\nu(x)\right)dx &= \int_0^{0.5}\cos^2\left(\frac{\pi}{2}\nu(x)\right)dx+\int_{0.5}^1\cos^2\left(\frac{\pi}{2}\nu(x))\right)dx\\
&= \int_0^{0.5}\cos^2\left(\frac{\pi}{2}\nu(x)\right)dx+\int_0^{0.5}\cos^2\left(\frac{\pi}{2}(1-\nu(\frac{1}{2}-x))\right)dx\\
&= \int_0^{0.5}\cos^2\left(\frac{\pi}{2}\nu(x)\right)dx+\int_0^{0.5}\sin^2\left(\frac{\pi}{2}\nu(\frac{1}{2}-x)\right)dx=\frac{1}{2}
\end{aligned}$$

In summery we obtain:

$$\|\psi\|^2_{\mathbb{L}_2}=\frac{2}{3}+\frac{2}{3}\int_0^1\cos^2\left(\frac{\pi}{2}\nu(x)\right)dx=\frac{2}{3}+\frac{2}{3}\cdot\frac{1}{2}=1 \qquad \blacksquare$$

Example 5.5 *The Mexican hat with*

$$\widehat{\psi}(\omega)=\frac{2}{\sqrt{3}}\sqrt[4]{\pi}\omega^2\exp(-\frac{\omega^2}{2})$$

doesn't have a compact domain. $\curvearrowright$ *The corresponding frame $\psi.=\psi_{mn}$ is not tight. We generate the family*

$$\psi.=\{\psi^j_{mn}\mid 0\le j\le N-1;\quad n,m\in\mathbb{Z}\}$$

with

$$\psi^j(t)=2^{-\frac{j}{N}}\psi(2^{-\frac{j}{N}}t);\quad j=0,...,N-1.$$

which gives a superposition of N grids in the phase space. This corresponds to „N voices per octave". We get the following estimations of A,B depending on $\beta,N,\sigma=2$

and a number b_0 *([2] p. 100):*

N	1			2		
b_0/π	A	B	B/A	A	B	B/A
0.25	13.091	14.183	1.083	27.273	27.278	1.000
0.50	6.546	7.092	1.083	13.637	13.639	1.000
0.75	4.364	4.728	1.083	9.091	9.093	1.000
1.00	3.223	3.596	1.116	6.768	6.870	1,015

N	3			4		
b_0/π	A	B	B/A	A	B	B/A
0.25	40.914	40.914	1.000	54.552	54.552	1.000
0.50	20.457	20.457	1.000	27.276	27.276	1.000
0.75	13.638	13.638	1.000	18.184	18.184	1.000
1.00	10.178	10.279	1.010	13.586	13.690	1.007

Now we know a possible partitioning for the $a-b-$plane to do the inverse wavelet transform and a possibility for construction of mother wavelets. But we don't know how to build a fast method for the inverse wavelet transform. The solution for this problem is the Multiscale Analysis.

5.3 Multiscale Analysis - MSA or Multiresolution Analysis MRA

MSA was founded by MEYER and MALLAT. It represents an independent way to obtain the discrete wavelet transform. The goal is the construction of very fast algorithms for the discrete wavelet transform, so that it has a chance in comparison with the FFT. It is the design method of most of the practically relevant discrete wavelet transforms (DWT). To this we need wavelets ψ, for which the frame $(\psi, 2, 1)$ is an ONS of $\mathbb{L}_2(\mathbb{R})$:

$$f = \sum_{m,n} (f, \psi_{mn}^{2,1})_{\mathbb{L}_2} \psi_{mn}^{2,1}$$

MSA provides the ability to construct such wavelet basis.

5.3.1 Introduction

Motivation:
Splitting the signal $f \in \mathbb{V}_{-1} \subset \mathbb{L}_2(\mathbb{R})$ in its

low-frequency	and	high-frequency component
$\downarrow$		$\downarrow$
1. by the orthogonal projection $\mathbf{P}_0 f$ to $\mathbb{V}_0 \subset \mathbb{V}_{-1}$, which contains the „smooth" functions of $\mathbb{V}_{-1}$.	$\Longrightarrow$	**2**. $\mathbb{W}_0 \oplus \mathbb{V}_0 = \mathbb{V}_{-1}$ $\mathbb{W}_0$ is the orthogonal complement to $\mathbb{V}_0$ with respect to $\mathbb{V}_{-1}$. It contains the „rough" functions.

Let $Q_0 f$ be the projection of f to $\mathbb{W}_0$. $\curvearrowright$

$$\begin{aligned} \mathbb{V}_{-1} &= \mathbb{V}_0 \oplus \mathbb{W}_0 \\ f &= P_0 f + Q_0 f \end{aligned}$$

In an analogous way we split $P_0 f$ by projectors P_1, Q_1, such that:

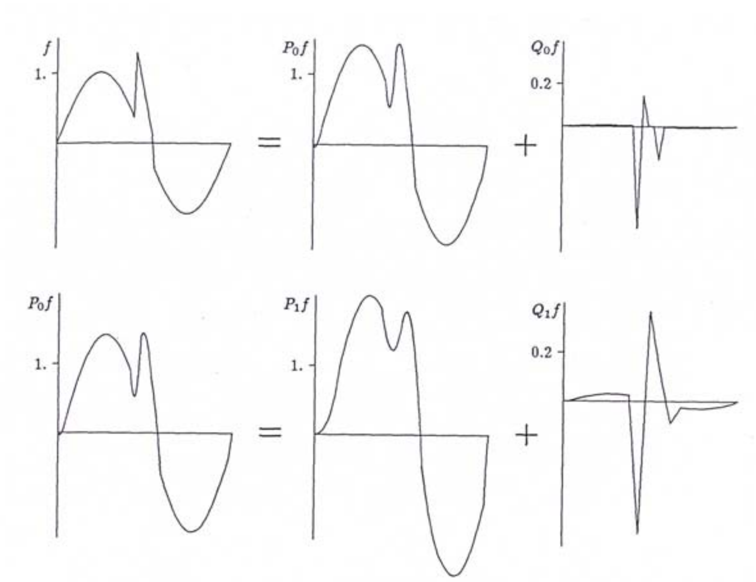
$$P_1(P_0 f) = P_1 f; \quad Q_1(P_0 f) = Q_1 f.$$

Therefore we get:

$$\begin{aligned} P_0 f &= P_1 f + Q_1 f \qquad \widehat{=} \qquad \mathbb{V}_0 = \mathbb{V}_1 \oplus \mathbb{W}_1 \\ f &= P_1 f + Q_1 f + Q_0 f \\ &\vdots \\ f &= \underbrace{P_n f} + \underbrace{Q_n f + Q_{n-1} f + \ldots + Q_0 f} \end{aligned}$$

This corresponds to splitting the signal into a mixture of low frequencies ($P_n f$) and bands of high frequencies ($Q_i f$). Thus the $Q_i f$ contain f shares of certain details size according to the overswept frequency band. $Q_0 f$ is the band with the highest frequencies, i.e. it corresponds to the smallest details. This decomposition process is called multi-scale analysis (MSA) and has a certain similarity to the decomposition process in multigrid algorithms by numerically solving partial differential equations.

Example 5.6 *Function splitting ([2], p. 104)*



The question is now: How to find the spaces $\mathbb{V}_j$?

Definition 5.6 *A multiresolution analysis (MSA) of the space $\mathbb{L}_2(\mathbb{R})$ consists of a sequence of closed nested subspaces $\mathbb{V}_m \subset \mathbb{L}_2(\mathbb{R})$:*

$$\{\mathbf{0}\} \subset ... \subset \mathbb{V}_1 \subset \mathbb{V}_0 \subset \mathbb{V}_{-1} \subset ... \subset \mathbb{L}_2(\mathbb{R})$$

such that:

$$a) \underbrace{\overline{\bigcup_{m\in\mathbb{Z}} \mathbb{V}_m} = \mathbb{L}_2(\mathbb{R})}_{\textbf{completeness}} \quad and \quad \underbrace{\bigcap_{m\in\mathbb{Z}} \mathbb{V}_m = \{\mathbf{0}\}}_{\textbf{separation axiom}}$$

$$b) \mathbb{V}_{j+1} = D_2(\mathbb{V}_j) \quad \forall j \in \mathbb{Z} \quad (\textit{self-similarity})$$

$$c) \exists \phi \in \mathbb{L}_2 \cap \mathbb{L}_1 \quad | \quad \{\phi(\cdot - k) \mid k \in \mathbb{Z}\} \textit{ is an ONS in } \mathbb{V}_0.$$

*The generating function ϕ is called the **scaling function or the father wavelet**.*

Notation 5.4 *$f \in \mathbb{V}_j$ contains only details of the extent of $\geq 2^j$ on the time axis. The more "negative" j, the finer the details which are included; in the limit, any $f \in \mathbb{L}_2(\mathbb{R})$ is reached ($\widehat{=}$ low-pass filter).*

Notation 5.5 *b) implies*

$$f(\cdot) \in \mathbb{V}_{j+1} \Longleftrightarrow f(2\cdot) \in \mathbb{V}_j \Longleftrightarrow f(2^{j+1}\cdot) \in \mathbb{V}_0.$$

MSA is distinguished by property b) and

Notation 5.6 *c) implies*

$$\mathbb{V}_0 = \{f \in \mathbb{L}_2(\mathbb{R}) \mid f(t) = \sum_k c_k \phi(t-k);\ \sum_k |c_k|^2 < \infty\}.$$

Analogous to the wavelet functions ψ_{mn}, functions $\phi_{jk}(t)$ are formed:

$$\phi_{jk}(t) = 2^{-\frac{j}{2}}\phi(\frac{t-k2^j}{2^j}) = 2^{-\frac{j}{2}}\phi(\frac{t}{2^j}-k) \qquad (5.3.1).$$

Notation 5.7 *Then b) implies that $\{\phi_{jk} \mid k \in \mathbb{Z}\}$ is an ONS of $\mathbb{V}_j$. ϕ_{jk} and ϕ_{jk+1} are shifted by a step size 2^j against each other. Thus the spaces $\mathbb{V}_j$ are scaled versions of the basic space $\mathbb{V}_0$ which is spanned by translation of the scaling function ϕ.*

Example 5.7 *Choose*

$$\phi_H = \begin{cases} 1 & for \quad 0 \le t < 1 \\ 0 & otherwise \end{cases}$$

and choose $\mathbb{V}_0$ such that it is the space of all functions which are constant at the intervals $[k; k+1)$:

$$\begin{aligned} \mathbb{V}_0 &= \overline{span\{\phi_{0k} \mid k \in \mathbb{Z}\}} \quad \curvearrowright \\ \mathbb{V}_j &= D_{2^j}(\mathbb{V}_0); \quad for \quad j \neq 0. \end{aligned}$$

By this definition the inclusion of the spaces, the separation axiom and the completeness axiom are satisfied (because of the completeness of step functions). Thus $\{\mathbb{V}_m\}_{m\in\mathbb{Z}}$ is a MSA. This choice of ϕ leads to the HAAR wavelet.

Because of the proper subset relationship of the spaces $\mathbb{V}_j$ pairwise orthogonal subspaces $\mathbb{W}_j \subset \mathbb{L}_2(\mathbb{R})$ are designed such that $\mathbb{L}_2(\mathbb{R})$ can be completed:

$$\mathbb{V}_{j-1} = \mathbb{V}_j \oplus \mathbb{W}_j; \quad \mathbb{W}_j \perp \mathbb{V}_j \quad \forall j \in \mathbb{Z}. \quad (5.3.2)$$

Furthermore:

$$\mathbb{W}_{j+1} = D_2(\mathbb{W}_j) \quad i.e. \quad f \in \mathbb{W}_j \Longleftrightarrow f\left(2^j \cdot\right) \in \mathbb{W}_0.$$

Theorem 5.9 *If $\{\mathbb{V}_j\}_{j\in\mathbb{Z}}$ has the properties up to and including a) of a MSA, then the subspaces $\mathbb{W}_j$ are pairwise orthogonal, and we obtain*

$$\mathbb{L}_2(\mathbb{R}) = \overline{\bigoplus_{j\in\mathbb{Z}} \mathbb{W}_j}.$$

proof: [1], p. 108

Let P_j be the orthoprojector from $\mathbb{L}_2(\mathbb{R})$ to $\mathbb{V}_j$, then the image $P_j f$ of the signal $f \in \mathbb{L}_2(\mathbb{R})$ still contains all details $\geq 2^j$ on the time axis. The following applies:

$$P_j f = \sum_{k=-\infty}^{\infty} \left(f, \phi_{jk}\right) \phi_{jk}.$$

Thus P_j corresponds to a low-pass filter.
Let Q_j be the orthoprojector from $\mathbb{L}_2(\mathbb{R})$ to $\mathbb{W}_j$. Because of (5.3.2) we get

$$P_{j-1} = P_j + Q_j \quad i.e. \quad Q_j = P_{j-1} - P_j.$$

Therefore $P_{j-1} f$ contains all details $\geq 2^{j-1}$ on the time axis. $-P_j f$ removes the details $\geq 2^j$. Thus Q_j corresponds to a (band) filter, which extract details of the length $2^{j-0.5} = 2^j \frac{1}{\sqrt{2}}$.

Notation 5.8

$$\begin{tikzcd}
\mathbb{L}_2(\mathbb{R}) \arrow[r, dashed] & \mathbb{V}_{-1} \arrow[r, "P_0"] \arrow[ddr, "Q_0"'] & \mathbb{V}_0 \arrow[r, "P_1"] \arrow[ddr, "Q_1"'] & \mathbb{V}_1 \arrow[r, "P_2"] \arrow[ddr, "Q_2"'] & \mathbb{V}_2 \arrow[r, dashed] & \{0\} \\
 & & \oplus & \oplus & \oplus & \\
 & & \mathbb{W}_0 & \mathbb{W}_1 & \mathbb{W}_2 \arrow[r, dashed] & \{0\}
\end{tikzcd}$$

Obviously:

$$\mathbb{V}_m = \bigoplus_{j \geq m+1} \mathbb{W}_j \quad \stackrel{m \to -\infty}{\longrightarrow} \quad \bigoplus_{j \in \mathbb{Z}} \mathbb{W}_j = \mathbb{L}_2(\mathbb{R}).$$

Therefore $f \in \mathbb{L}_2(\mathbb{R})$ can be decomposed into

$$f = \sum_{j \in \mathbb{Z}} Q_j f = \sum_{j \leq m} Q_j f + \sum_{j \geq m+1} Q_j f = P_m f + \sum_{j \geq m+1}^{\infty} Q_j f,$$

i.e. in a lowpass and a sum of band filters.

Outlook:
For every MSA there exists a wavelet ψ, which translated and dilated versions $\psi_{mk}(t) = 2^{-\frac{m}{2}} \psi(2^{-m} t - k)$ for fixed $m \in \mathbb{Z}$ are an ONS of $\mathbb{W}_m$.
The mother wavelet can be constructed by the scaling function explicitly.
$\curvearrowright Q_m f = \sum_{-\infty}^{\infty} (f, \psi_{mk}) \psi_{mk}$. The collection of all ψ_{mk} i.e. the family ψ. is then an ONS of $\mathbb{L}_2(\mathbb{R})$.

5.3.2 Scaling function

It is the most important part of the MSA, because corresponding to remark 5.6 from 5.3.1 we define first the space $\mathbb{V}_0$ by ϕ and then the other spaces $\mathbb{V}_j$ by definition b). $\phi = \phi_{0,0}$ must be defined in such a way that the functions $\phi_{0k} = \phi(\cdot - k)$ are orthonormal (see definition c). If necessary it must be aided by the Schmidt orthonormalisation process. In addition the definition of ϕ must guarantee the inclusions of the spaces $\mathbb{V}_j$ and the completeness axiom (properties a) and b) of the MSA).

Theorem 5.10 *Choose* $\phi \in \mathbb{L}_2(\mathbb{R}),\ \phi \neq 0$

$$\mathbb{V}_0 = \{f \in \mathbb{L}_2 \mid f\,(t) = \sum_k c_k \phi(t-k); \quad \sum_k |c_k|^2 < \infty\}; \quad \mathbb{V}_{j+1} = D_2 \mathbb{V}_j; \quad j \in \mathbb{Z}$$

If $\mathbb{V}_0 \subset \mathbb{V}_{-1}$, *then* $\{\mathbf{0}\} \subset ... \subset \mathbb{V}_1 \subset \mathbb{V}_0 \subset \mathbb{V}_{-1} \subset ... \subset \mathbb{L}_2(\mathbb{R})$

Theorem 5.11

$$\mathbb{V}_0 \subset \mathbb{V}_{-1} \quad \Longleftrightarrow \quad \exists\ h. \in \mathbf{l}^2(\mathbb{Z}) \mid \phi\,(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} h_k \phi(2t-k) \quad \textit{almost everywhere} \tag{5.3.3}$$

This representation of $\phi\,(t)$ is called the **Scaling equation**.
Proof. I) $\Longrightarrow$
$\mathbb{V}_0 = \{f \in \mathbb{L}_2(\mathbb{R}) \mid f(t) = \sum_k c_k \phi(t-k),\ \sum_k |c_k|^2 < \infty\}$

$$\begin{aligned} f &\in \mathbb{V}_{-1} \quad \Longleftrightarrow \quad f(2^{-1}\cdot) \in \mathbb{V}_0 \quad \curvearrowright \\ \phi_{-1,k} &= 2^{\frac{1}{2}} \phi\left(\frac{t-k2^{-1}}{2^{-1}}\right) \text{ is an ONS in } \mathbb{V}_{-1}. \end{aligned}$$

$\curvearrowright$

$$\mathbb{V}_{-1} = \{f \in \mathbb{L}_2(\mathbb{R}) \mid f(t) = \sum_k h_k \phi_{-1,k};\ h. \in \mathbf{l}^2(\mathbb{Z})\}$$

$\curvearrowright$

$$\phi(t) = \sum_k h_k \sqrt{2} \phi(2t-k) \quad \textit{almost everywhere with } h. \in \mathbf{l}^2(\mathbb{Z})$$

II) $\Longleftarrow$
Assume

$$\begin{aligned} \phi(t-l) &= \sqrt{2} \sum_k \widetilde{h}_k \phi(2(t-l)-k) \\ &= \sqrt{2} \sum_k \widetilde{h}_k \phi(2t-(k+2l)) \end{aligned}$$

$\curvearrowright$

$$\phi_{0,l} = \phi(t-l) = \sum_k \widetilde{h}_k \phi_{-1,k+2l} \in \mathbb{V}_{-1}$$

$\curvearrowright \quad \mathbb{V}_0 \subset \mathbb{V}_{-1}$ ∎

Notation 5.9 *The scaling equation rules the MSA, because h. defines the scaling function uniquely.*

Notation 5.10 *For the fast algorithm we need h., not ϕ and ψ.*

Notation 5.11 *The scaling equation describes the self-similarity of ϕ (analogous to the theory of fractal sets) which results in restrictions on the selection of ϕ.*

Properties of the coefficients h_k :
$\{\phi_{0k}\}_{k\in\mathbb{Z}}$ is an ONS in $\mathbb{V}_0$ which implies:

1. $\delta_{0n} = \sum_k h_k \overline{h_{2n+k}} \quad \forall n \in \mathbb{Z}$ (5.3.4, consistency condition)

$$\begin{aligned}
\delta_{0n} &= (\phi_{0n}, \phi) = \int \phi(t-n)\overline{\phi(t)}dt \\
&= 2\sum_{k,l} h_k \overline{h_l} \int \phi(2t-2n-k)\overline{\phi(2t-l)}dt \\
&= \sum_{k,l} h_k \overline{h_l} \int \phi(\widetilde{t}-2n-k)\overline{\phi(\widetilde{t}-l)}d\widetilde{t} \\
&= \sum_{k,l} h_k \overline{h_l}\delta_{l,2n+k} = \sum_k h_k \overline{h_{2n+k}} \qquad \forall n \in \mathbb{Z}
\end{aligned}$$

2. $\sum_k |h_k|^2 = 1.$

Theorem 5.12 $h \in \mathbf{l}^1(\mathbb{Z})$; $\int_{\mathbb{R}} \phi(t)\,dt = q \neq 0 \implies \sum_{k=-\infty}^{\infty} h_k = \sqrt{2}.$

Theorem 5.13 *If the scaling function ϕ is a function with compact support, then there are only finitely many $h_k \neq 0$.*

Theorem 5.14 *Let ϕ be a function with compact support,*

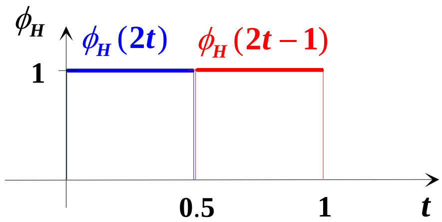
$$\begin{aligned}
a &= a(\phi) = \inf\{t \mid \phi(t) \neq 0\} > -\infty \\
b &= b(\phi) = \sup\{t \mid \phi(t) \neq 0\} < +\infty,
\end{aligned}$$

then: $a, b \in \mathbb{Z}$ and there are at most these h_k with $a \leq k \leq b$ not equal zero.

(proofs: [1])

Example 5.8

$$\begin{aligned}\phi &= \phi_H = \begin{cases} 1 & for \quad 0 \le t < 1 \\ 0 & otherwise \end{cases} \\ &= \phi_H(2t) + \phi_H(2t-1) \\ &= \frac{1}{\sqrt{2}}\phi_{-1,0} + \frac{1}{\sqrt{2}}\phi_{-1,1} \\ &= h_0\phi_{-1,0} + h_1\phi_{-1,1}\end{aligned}$$



$h_k = 0$ for all other h_k. Moreover, the function ϕ satisfies the scaling equation and $\sum_k |h_k|^2 = 1$.

As in the fast algorithm only the coefficients h_k are significant, furthermore we use only functions ϕ with compact support.
Under which conditions on ϕ in the MSA, are the completeness and the separation axioms true?

Theorem 5.15 *Choose $\phi \in \mathbb{L}_2(\mathbb{R})$ such that $|\phi(t)| \le \frac{c}{1+t^2}$; $\quad t \in \mathbb{R}$; $\quad$ and let $\{\phi_{0k}\}_{k\in\mathbb{Z}}$ be an ONS of $\mathbb{V}_0$. Then*

$$\bigcap_j \mathbb{V}_j = \{\mathbf{0}\} \quad \wedge \quad \overline{\bigcup_j \mathbb{V}_j} = \mathbb{L}_2(\mathbb{R}) \quad \Longleftrightarrow \quad \left|\int \phi dt\right| = 1.$$

(Proof: [1] p. 114-117)

When does the set $\{\phi_k = \phi(\cdot - k)\}_{k\in\mathbb{Z}}$ form an ONS of $\mathbb{V}_0$?

Theorem 5.16 *$\phi \in \mathbb{L}_2(\mathbb{R})$. $\{\phi_k = \phi(\cdot - k)\}_{k\in\mathbb{Z}}$ forms an ONS*
$\Longleftrightarrow \quad \Phi(\omega) = \sum_l |\widehat{\phi}(\omega + 2\pi l)|^2 = \frac{1}{2\pi}$ *almost everywhere*

Proof. Given $\phi \in \mathbb{L}_2(\mathbb{R})$

$$\begin{aligned}
(\phi_0, \phi_k) &= (\widehat{\phi}_0, \widehat{\phi}_k) = \int_{\mathbb{R}} \widehat{\phi}(\omega)\overline{\exp(-ik\omega)\widehat{\phi}(\omega)}d\omega \\
&= \int_{\mathbb{R}} \left|\widehat{\phi}(\omega)\right|^2 \exp(ik\omega)d\omega \\
&= \sum_l \int_0^{2\pi} \left|\widehat{\phi}(\omega + 2\pi l)\right|^2 \exp(ik\omega)d\omega \\
&= \int_0^{2\pi} \sum_l \left|\widehat{\phi}(\omega + 2\pi l)\right|^2 \exp(ik\omega)d\omega \qquad (FUBINI\ theorem) \\
&= \int_0^{2\pi} \Phi(\omega)\exp(ik\omega)d\omega \\
&= 2\pi\widehat{\Phi}(-k) \stackrel{!}{=} \delta_{0k} \qquad (\Phi\ is\ \ 2\pi - periodic) \\
\Longleftrightarrow\ & \widehat{\Phi}(-k) = \frac{1}{2\pi}\delta_{0k} \quad \curvearrowright \quad \Phi(\omega) = \frac{1}{2\pi} \quad almost\ everywhere
\end{aligned}$$

■

5.3.3 Construction of the Scaling Function and the Mother Wavelet

We transfer the construction into the codomain of the Fourier transform because there the calculation is easier. First we apply the Fourier transform to the scaling equation:

$$\begin{aligned}
\phi(t) &= \sqrt{2}\sum_{k=-\infty}^{\infty} h_k\phi(2t-k) = \sqrt{2}\sum_{-\infty}^{\infty} h_k\phi(2(t-\frac{k}{2})) \quad \curvearrowright \\
\widehat{\phi}(\omega) &= \sqrt{2}\sum_{k=-\infty}^{\infty} h_k e^{-i\frac{\omega}{2}k}\frac{1}{2}\widehat{\phi}(\frac{\omega}{2}) \qquad (R1, R3) \\
\widehat{\phi}(\omega) &= H\left(\frac{\omega}{2}\right)\widehat{\phi}(\frac{\omega}{2}) \qquad (5.3.5) \\
\text{with } H(\frac{\omega}{2}) &= \frac{1}{\sqrt{2}}\sum_{k=-\infty}^{\infty} h_k e^{-i\frac{\omega}{2}k}. \qquad (5.3.5.a)
\end{aligned}$$

If finitely many $h_k \neq 0$ then the function $H(\omega)$ is almost everywhere convergent because $\|h.\| = 1$, is $2\pi-$periodic and a trigonometric polynomial.

Theorem 5.16 implies

$$\begin{aligned}
\frac{1}{2\pi} &= \sum_l \left|\widehat{\phi}(\omega+4\pi l)\right|^2 + \sum_l \left|\widehat{\phi}(\omega+2\pi+4\pi l)\right|^2 \\
&= \sum_l \left|H(\frac{\omega}{2}+2\pi l)\right|^2 \left|\widehat{\phi}(\frac{\omega}{2}+2\pi l)\right|^2 + \sum_l \left|H(\frac{\omega}{2}+\pi+2\pi l)\right|^2 \left|\widehat{\phi}(\frac{\omega}{2}+\pi+2\pi l)\right|^2 \\
&= \left|H(\frac{\omega}{2})\right|^2 \sum_l \left|\widehat{\phi}(\frac{\omega}{2}+2\pi l)\right|^2 + \left|H(\frac{\omega}{2}+\pi)\right|^2 \sum_l \left|\widehat{\phi}(\frac{\omega}{2}+\pi+2\pi l)\right|^2 \\
&= \left|H(\frac{\omega}{2})\right|^2 \Phi(\frac{\omega}{2}) + \left|H(\frac{\omega}{2}+\pi)\right|^2 \Phi(\frac{\omega}{2}+\pi) \\
&= \left(\left|H(\frac{\omega}{2})\right|^2 + \left|H(\frac{\omega}{2}+\pi)\right|^2\right)\frac{1}{2\pi}.
\end{aligned}$$

Therefore we get the „FOURIER version” of the consistency condition:

Theorem 5.17 *The generating function H of a MSA satisfies the equation*

$$|H(\omega)|^2 + |H(\omega+\pi)|^2 = 1 \quad \textit{almost everywhere}$$

Thus:

- $|H(\omega)| \le 1; \quad \omega \in \mathbb{R}$
- $\widehat{\phi}(0) \neq 0$ implies: $\widehat{\phi}(0) = H(0)\,\widehat{\phi}(0) \quad \curvearrowright \quad H(0) = 1$
- $|H(0+\pi)|^2 = 1 - |H(0)|^2 = 0 \quad \curvearrowright \quad H(\pi) = 0$
- $H(0) = \frac{1}{\sqrt{2}}\sum_k h_k = 1 \quad \curvearrowright \quad \sum_k h_k = \sqrt{2}$
- $H(\pi) = \frac{1}{\sqrt{2}}\sum_k h_k e^{-ik\pi} = \frac{1}{\sqrt{2}}\sum_k h_k (-1)^k = 0 \quad \curvearrowright \quad \sum_k h_k (-1)^k = 0.$

Theorem 5.18 $f \in \mathbb{L}_2(\mathbb{R}),\ f \in \mathbb{W}_0$
$\Longleftrightarrow \quad \exists \nu = \nu(.) \in \mathbb{L}_2(\mathbb{R}/2\pi) \quad | \quad \widehat{f}(\omega) = \exp(i\frac{\omega}{2})\nu(\omega)\overline{H(\frac{\omega}{2}+\pi)}\widehat{\phi}(\frac{\omega}{2})$. *(Proof se* [1]*)*

The wavelet basis must fill the space $\mathbb{L}_2(\mathbb{R})$. By elimination of the spaces $\mathbb{W}_j$ band filters are defined. The wavelet basis is a basis in the spaces $\mathbb{W}_j$ and the mother wavelet ψ has to belong to $\mathbb{W}_0$. $\curvearrowright$ Ansatz for $\widehat{\psi}$:

$$\widehat{\psi}(\omega) = e^{i\frac{\omega}{2}}\overline{H\left(\frac{\omega}{2}+\pi\right)}\widehat{\phi}(\frac{\omega}{2}) \qquad (5.3.9)$$

Theorem 5.19 *If ψ is defined by (5.3.9) then $\{\psi_{0k}\}_{k\in\mathbb{Z}}$ forms an ONS of $\mathbb{W}_0$.*

Proof. $1 \in \mathbb{L}_2(\mathbb{R}/2\pi) \quad \curvearrowright \quad \psi \in \mathbb{W}_0$ because of the previous theorem
$\curvearrowright \ \psi_{0k} \in \mathbb{W}_0$
Using theorems 5.16 and 5.18 we prove the orthonormality:

$$\begin{aligned}
\Psi(\omega) &= \sum_n \left|\widehat{\psi}(\omega+2\pi n)\right|^2 \\
&= \sum_l \left|\widehat{\psi}(\omega+4\pi l)\right|^2 + \sum_l \left|\widehat{\psi}(\omega+2\pi+4\pi l)\right|^2 \\
&= \sum_l \left|\overline{H(\frac{\omega}{2}+2\pi l+\pi)}\right|^2 \left|\widehat{\phi}(\frac{\omega}{2}+2\pi l)\right|^2 + \sum_l \left|\overline{H(\frac{\omega}{2}+\pi+2\pi l+\pi)}\right|^2 \left|\widehat{\phi}(\frac{\omega}{2}+\pi+2\pi l)\right|^2 \\
&= \sum_l \left|\overline{H(\frac{\omega}{2}+\pi)}\right|^2 \left|\widehat{\phi}(\frac{\omega}{2}+2\pi l)\right|^2 + \sum_l \left|\overline{H(\frac{\omega}{2})}\right|^2 \left|\widehat{\phi}(\frac{\omega}{2}+\pi+2\pi l)\right|^2 \\
&= \left|\overline{H(\frac{\omega}{2}+\pi)}\right|^2 \sum_l \left|\widehat{\phi}(\frac{\omega}{2}+2\pi l)\right|^2 + \left|\overline{H(\frac{\omega}{2})}\right|^2 \sum_l \left|\widehat{\phi}(\frac{\omega}{2}+\pi+2\pi l)\right|^2 \\
&= \frac{1}{2\pi}\left(\left|\overline{H(\frac{\omega}{2}+\pi)}\right|^2 + \left|\overline{H(\frac{\omega}{2})}\right|^2\right) = \frac{1}{2\pi}
\end{aligned}$$

$f \in \mathbb{W}_0$ implies $\exists \nu = \nu(.) \in \mathbb{L}_2(\mathbb{R}/2\pi)$ such that:

$$\begin{aligned}
\widehat{f}(\omega) &= \nu(\omega)\exp(i\frac{\omega}{2})\overline{H(\frac{\omega}{2}+\pi)}\widehat{\phi}(\frac{\omega}{2}) \quad (\textit{theorem } 5.18) \\
&= \nu(\omega)\widehat{\psi}(\omega) \quad (\textit{see } 5.3.9).
\end{aligned}$$

Because of $\nu(.) \in \mathbb{L}_2(\mathbb{R}/2\pi)$ we get $\nu(\omega) = \sum_k \nu_k \exp(-ik\omega)$
with $\sum_k |\nu_k|^2 = \|\nu\| < \infty$ which implies

$$\begin{aligned}
\widehat{f}(\omega) &= \sum_k \nu_k \exp(-ik\omega)\widehat{\psi}(\omega) \\
\underset{almost\ everywhere}{\overset{R1}{\Longleftrightarrow}} f(t) &= \sum_k \nu_k \psi(t-k) \\
&= \sum_k \nu_k \psi_{0k} \quad \text{is convergent.}
\end{aligned}$$

$\curvearrowright \quad \{\psi_{0k}\}_{k\in\mathbb{Z}}$ form a basis in $\mathbb{W}_{0.}$ ■

Notation 5.12 *The approach for $\widehat{\psi}(\omega)$ is not unique. Factors like*

$$(-1)\, e^{i\alpha} e^{-iN\omega}$$

with $\alpha \in \mathbb{R}$; $N \in \mathbb{N}$ are allowed. The multiplication by $e^{-iN\omega}$ corresponds to a translation of the support of ψ by N units to the right.

Inverse FOURIER Transform into the time domain of $\widehat{\psi}(\omega)$:

$$
\begin{aligned}
\widehat{\psi}(\omega) &= e^{i\frac{\omega}{2}}\overline{H\left(\frac{\omega}{2}+\pi\right)}\widehat{\phi}(\frac{\omega}{2}) \\
&= \frac{1}{\sqrt{2}}\sum_k \overline{h_k}e^{i(\frac{\omega}{2}+\pi)k}e^{i\frac{\omega}{2}}\widehat{\phi}(\frac{\omega}{2}) \quad \text{(because of (5.3.5.}a\text{))} \\
&= \frac{1}{\sqrt{2}}\sum_k (-1)^k\,\overline{h_k}e^{i\frac{\omega}{2}(k+1)}\widehat{\phi}(\frac{\omega}{2}) \\
&= \frac{1}{\sqrt{2}}\sum_{\widetilde{k}} (-1)^{(\widetilde{k}-1)}\,\overline{h}_{-\widetilde{k}-1}e^{-i\frac{\omega}{2}\widetilde{k}}\widehat{\phi}(\frac{\omega}{2}) \quad \text{with } k=-\widetilde{k}-1
\end{aligned}
$$

Application of $(R1)$ and $(R2)$ implies:

$$
\begin{aligned}
\psi(t) &= \sqrt{2}\sum_k (-1)^{k-1}\,\overline{h}_{-k-1}\phi(2t-k) \\
&= \sqrt{2}\sum_k g_k\phi(2t-k) \quad \text{with } g_k=(-1)^{k-1}\,\overline{h}_{-k-1}
\end{aligned}
$$

Another possible definition of g_k is the following:

$$g_k = (-1)^k\,\overline{h}_{2N-1-k}.$$

$h_k \neq 0$ for $0 \le k \le 2N-1$ implies $g_k \neq 0$ for $0 \le k \le 2N-1$,
$\curvearrowright$ Summation in the algorithms: $0 \le k \le 2N-1$

Theorem 5.20 *Let $\{\mathbb{V}_j\}_{j\in\mathbb{Z}}$ be a MSA with the scaling function ϕ and the generating function H, $\psi(t)=\sqrt{2}\sum_k g_k\phi(2t-k)$ with $g_k=(-1)^{k-1}\,\overline{h}_{-k-1}$. This implies:*

$$\{\psi_{jk} \mid j\in\mathbb{Z},\quad k\in\mathbb{Z}\}$$

is an ONS of $\mathbb{L}_2(\mathbb{R})$, the wavelet basis. (Proof [1], p. 124)

Example 5.9 *HAAR Wavelet:*

$$\phi=\phi_H=\left\{\begin{array}{ll} 1 & t\in[0;1) \\ 0 & otherwise \end{array}\right\} \quad \curvearrowright \quad \widehat{\phi}_H(\omega)=\frac{1}{\sqrt{2\pi}}si\left(\frac{\omega}{2}\right)e^{-i\frac{\omega}{2}}$$

According to the scaling equation we got:

$\phi = \frac{1}{\sqrt{2}}\phi_{-1,0} + \frac{1}{\sqrt{2}}\phi_{-1,1} \quad \curvearrowright \quad h_0 = h_1 = \frac{1}{\sqrt{2}} \quad \curvearrowright$

$$\begin{aligned} H(\omega) &= \frac{1}{\sqrt{2}}\sum_k h_k e^{-ik\omega} \\ &= \frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}e^{-i\omega}) \\ &= \frac{1}{2}(1+e^{-i\omega}) \\ &= e^{-i\frac{\omega}{2}}\frac{1}{2}(e^{i\frac{\omega}{2}} + e^{-i\frac{\omega}{2}}) \\ &= e^{-i\frac{\omega}{2}}\cos\frac{\omega}{2} \end{aligned}$$

Thus

$$\begin{aligned} H\left(\frac{\omega}{2}\right)\widehat{\phi}\left(\frac{\omega}{2}\right) &= e^{-i\frac{\omega}{4}}\cos\frac{\omega}{4}\frac{1}{\sqrt{2\pi}}si\left(\frac{\omega}{4}\right)e^{-i\frac{\omega}{4}} \\ &= e^{-i\frac{\omega}{2}}\frac{1}{\sqrt{2\pi}}\frac{\cos\frac{\omega}{4}\sin\left(\frac{\omega}{4}\right)}{\frac{\omega}{4}}\frac{2}{2} \\ &= e^{-i\frac{\omega}{2}}\frac{1}{\sqrt{2\pi}}\frac{\sin\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}} \\ &= \widehat{\phi}_H(\omega) \end{aligned}$$

After an inverse Fourier transform we get, by the approach $\widehat{\psi}(\omega) = e^{i\frac{\omega}{2}}\overline{H\left(\frac{\omega}{2}+\pi\right)}\widehat{\phi}\left(\frac{\omega}{2}\right)$, *a function* ψ *which is translated one unit to the left and multiplied by* (-1) *with respect to* ψ_H*. Therefore we use :*

$$\begin{aligned} \widehat{\psi}(\omega) &= -e^{-i\omega}e^{i\frac{\omega}{2}}\overline{H\left(\frac{\omega}{2}+\pi\right)}\widehat{\phi}\left(\frac{\omega}{2}\right) \\ &= -e^{-i\frac{\omega}{2}}\cos(\frac{1}{2}(\frac{\omega}{2}+\pi))e^{i\frac{1}{2}(\frac{\omega}{2}+\pi)}\frac{1}{\sqrt{2\pi}}si\left(\frac{\omega}{4}\right)e^{-i\frac{\omega}{4}} \\ &= -e^{-i\frac{\omega}{2}}(-\sin\frac{\omega}{4})\cdot i\cdot\frac{1}{\sqrt{2\pi}}si\left(\frac{\omega}{4}\right) \\ &= \frac{i}{\sqrt{2\pi}}\frac{\sin^2\left(\frac{\omega}{4}\right)}{\frac{\omega}{4}}e^{-i\frac{\omega}{2}} \\ &= \widehat{\psi}_H(\omega) \end{aligned}$$

We calculate the coefficients g_k *by the inverse Fourier transform of* $\widehat{\psi}_H(\omega)$ *into the* *time domain:*

$$\begin{aligned}
\widehat{\psi}(\omega) &= -e^{-i\omega}e^{i\frac{\omega}{2}}\overline{H\left(\frac{\omega}{2}+\pi\right)}\widehat{\phi}\left(\frac{\omega}{2}\right) \\
&= \frac{1}{\sqrt{2}}\sum_k \overline{h}_k e^{ik(\frac{\omega}{2}+\pi)}(-e^{-i\frac{\omega}{2}})\widehat{\phi}\left(\frac{\omega}{2}\right) \quad \textit{(replacement of H)} \\
&= \frac{1}{\sqrt{2}}\sum_k (-1)^k\,\overline{h}_k e^{i(k-1)\frac{\omega}{2}}\widehat{\phi}\left(\frac{\omega}{2}\right)(-1) \\
&= \frac{1}{\sqrt{2}}\sum_k (-1)^{k+1}\,\overline{h}_k e^{-i(1-k)\frac{\omega}{2}}\widehat{\phi}\left(\frac{\omega}{2}\right) \\
&= \frac{1}{\sqrt{2}}\sum_{\widetilde{k}} (-1)^{2-\widetilde{k}}\,\overline{h}_{1-\widetilde{k}} e^{-i\widetilde{k}\frac{\omega}{2}}\widehat{\phi}\left(\frac{\omega}{2}\right) \quad (\widetilde{k}=1-k).
\end{aligned}$$

and get ψ :

$$\begin{aligned}
\psi(t) &= \sqrt{2}\sum_k (-1)^{2-k}\,\overline{h}_{1-k}\phi(2t-k) \\
&= \sqrt{2}\sum_k g_k\phi(2t-k) \quad \textit{with} \quad g_k = (-1)^{2-k}\,\overline{h}_{1-k}
\end{aligned}$$

$$\begin{aligned}
g_k &= (-1)^{2-k}\,\overline{h}_{1-k} = (-1)^k\,\overline{h}_{1-k} \\
g_0 &= \overline{h}_1 = \frac{1}{\sqrt{2}} \\
g_1 &= -\overline{h}_0 = -\frac{1}{\sqrt{2}}
\end{aligned}$$

Thus:

$$\begin{aligned}
\psi(t) &= \sqrt{2}(\frac{1}{\sqrt{2}}\phi(2t) - \frac{1}{\sqrt{2}}\phi(2t-1)) \\
&= \phi(2t) - \phi(2t-1) \\
&= \frac{1}{\sqrt{2}}\phi_{-1,0} - \frac{1}{\sqrt{2}}\phi_{-1,1} \\
&= \psi_H(t)
\end{aligned}$$

As we are able to construct a MSA by $\phi = \phi_H$ (see previous example) there exists two paths to the HAAR wavelet:

1. construction by the spaces $\mathbb{V}_m,\ \mathbb{W}_m$ and the corresponding projection operators $P_m,\ Q_m$ and

2. a general construction by the induced wavelet.

Summary:

Theorem 5.21 *$\phi \in \mathbb{L}_1 \cap \mathbb{L}_2$ and satisfies the scaling equation:*
$\phi(t) = \sqrt{2} \sum_{-\infty}^{\infty} h_k \phi(2t - k);\quad \int_{\mathbb{R}} \phi(t)\, dt \neq 0;$
$f \in \mathbb{V}_j \Longleftrightarrow f(2^j \cdot) \in \mathbb{V}_0;$
$\mathbb{V}_0 = \{f \in \mathbb{L}_2(\mathbb{R}) \mid f(t) = \sum_k c_k \phi(t-k);\quad \sum_k |c_k|^2 < \infty\}.$
If there exists constants A, B such that
$0 < A \leq B$ *and* $A \leq \Phi(\omega) = \sum_l |\widehat{\phi}(\omega + 2\pi l)|^2 \leq B$ *almost everywhere, then:*

1. *$\{\phi_{0k} = \phi(\cdot - k)\}_{k\in\mathbb{Z}}$ is a frame with the bounds $2\pi A$ and $2\pi B$.*

2. *$\widehat{\widetilde{\phi}}(\omega) = \dfrac{\widehat{\phi}(\omega)}{\sqrt{2\pi\Phi(\omega)}}$ defines a MSA with the same spaces $\mathbb{V}_j$;*
$\{\phi_{0k} = \phi(\cdot - k)\}_{k\in\mathbb{Z}}$ is an ONS of $\mathbb{V}_0$.

Proof: [1], p. 128ff

Notation 5.13 *Every MSA induces orthogonal wavelet families. The converse is not true. (counterexample see [2], p. 118).*

Notation 5.14 *A method of construction for a wavelet basis:*

1. *Definition of a scaling function ϕ with a nonvanishing average*

2. *Construction of the spaces $\mathbb{V}_j$*

3. *Check whether $\{\phi_{0k}\}_{k\in\mathbb{Z}}$ is an ONS of $\mathbb{V}_0$:*

$$\sum_l |\widehat{\phi}(\omega + 2\pi l)|^2 \stackrel{?}{=} \frac{1}{2\pi} \quad \textit{almost everywhere}$$

Yes	*No*
$\{\phi_{0k}\}_{k\in\mathbb{Z}}$ is a tight frame	*Search constants A, B for the theorem above* *$\{\phi_{0k}\}_{k\in\mathbb{Z}}$ is a frame*
$\widehat{\widetilde{\phi}}(\omega) = \widehat{\phi}(\omega)$	$\widehat{\widetilde{\phi}}(\omega) = \dfrac{\widehat{\phi}(\omega)}{\sqrt{2\pi\Phi(\omega)}}$

4. *Construction of the MSA corresponding to* $\widehat{\widetilde{\phi}}(\omega)$
5. *calculation of the scaling coefficients* h_k
6. *Construction of the wavelet basis by the wavelet functions*

Notation 5.15 *Problems during this process are*

- *the orthogonalisation and*
- *the calculation of the coefficients* h_k *if* $\widetilde{\phi}$ *has not a compact support.*

5.4 Fast Algorithms

The starting point is the scaling equation together with the corresponding equation for $\psi(t)$:

$$\begin{aligned}\phi(t) &= \sqrt{2}\sum_{l=-\infty}^{\infty} h_l\phi(2t-l)\\ \psi(t) &= \sqrt{2}\sum_{l=-\infty}^{\infty} g_l\phi(2t-l)\\ \textit{with}\quad g_l &= (-1)^{l-1}\,\overline{h}_{-l-1}\quad \textit{or}\quad g_l=(-1)^l\,\overline{h}_{2N-l-1}\end{aligned}$$

This results in:

$$\begin{aligned}\phi_{jk}(t) &= 2^{-\frac{j}{2}}\phi\left(\frac{t}{2^j}-k\right)\\ &= 2^{-\frac{j}{2}+\frac{1}{2}}\sum_{l=-\infty}^{\infty} h_l\phi\left(2\left(\frac{t}{2^j}-k\right)-l\right)\\ &= 2^{-(\frac{j-1}{2})}\sum_{l=-\infty}^{\infty} h_l\phi\left(\frac{t}{2^{j-1}}-2k-l\right)\\ &= \sum_{l=-\infty}^{\infty} h_l\phi_{j-1,2k+l}(t) \qquad (5.4.1.a)\end{aligned}$$

and similarly

$$\psi_{jk}(t)=\sum_{l=-\infty}^{\infty} g_l\phi_{j-1,2k+l}(t)\,. \qquad (5.4.1.b)$$

This recursion forms the basis for a fast algorithm.

5.4.1 Analysis of $f \in \mathbb{L}_2(\mathbb{R})$

The finest scale to be considered belongs to $j = 0$.

1. Start at $j = 0$:

$$\begin{aligned} a_{0k} &= (f, \phi_{0k})_{\mathbb{L}_2} = \int_{\mathbb{R}} f(t)\overline{\phi(t-k)}dt \quad \text{(numerical integration)} \\ P_0 f &= \sum_k a_{0k}\phi_{0k} \end{aligned}$$

2. Proceed in the direction of increasing $j = j+1$, i.e. in the direction of longer and longer wavelengths.
   Let $j \geq 1$; $a_{j-1,k}$ are known.
   P_{j-1} describes details of f having a spread $\geq 2^{j-1}$ on time axis.

$$P_{j-1}f = \sum_{k=-\infty}^{\infty} a_{j-1,k}\phi_{j-1,k.}$$

$$\begin{aligned} a_{jk} &= (f, \phi_{jk})_{\mathbb{L}_2} \\ &= \sum_{l=-\infty}^{\infty} \overline{h_l}(f, \phi_{j-1,2k+l}) \qquad (with\ 5.4.1a) \\ &= \sum_{l=-\infty}^{\infty} \overline{h_l} a_{j-1,2k+l} \qquad (5.4.2) \\ P_j f &= \sum_k a_{jk}\phi_{jk} \end{aligned}$$

   This corresponds to the next more coarse approximation of f. Further

$$P_{j-1}f = P_j f + Q_j f,$$

   such that $Q_j f \in \mathbb{W}_j$ where $\{\psi_{jk}\}$ is an ONS. $\curvearrowright$

$$\begin{aligned} Q_j f &= \sum_{k=-\infty}^{\infty} d_{jk}\psi_{jk} \\ d_{jk} &= (f, \psi_{jk}) \\ &= \sum_{l=-\infty}^{\infty} \overline{g_l}(f, \phi_{j-1,2k+l}) \qquad (with\ 5.4.1.b) \\ &= \sum_{l=-\infty}^{\infty} \overline{g_l} a_{j-1,2k+l} \qquad (5.4.3) \end{aligned}$$

Therefore the information about the signal f which is extracted from $a_{j-1,k}$ is saved in the vector $\underline{d}_j$. It contains information about details of f which have a spread of size $\approx 2^{j-\frac{1}{2}}$ on the time axis. During this process f is made „coarser by factor two".
Calculation scheme:

$$\begin{tikzcd}
\underline{\boldsymbol{a}}_0 \arrow[r, "(5.4.2)"] \arrow[dr, "(5.4.3)"'] & \underline{\boldsymbol{a}}_1 \arrow[r, "(5.4.2)"] \arrow[dr, "(5.4.3)"'] & \underline{\boldsymbol{a}}_2 \cdots\cdots\cdots & \arrow[r, "(5.4.2)"] \arrow[dr, "(5.4.3)"'] & \underline{\boldsymbol{a}}_J \\
& \underline{\boldsymbol{d}}_1 & \underline{\boldsymbol{d}}_2 & & \underline{\boldsymbol{d}}_J
\end{tikzcd}$$

Algorithm:
Input of $\underline{a}_0$ and of the depth of the decomposition J
for $j = 1$ to J

calculate $\underline{a}_j$ by (5.4.2)

calculate $\underline{d}_j$ by (5.4.3)

end
Output of $\underline{a}_J,\ \underline{d}_1, ...\underline{d}_J$

Notation 5.16 *Vector $\underline{a}_0$ should be calculated numerically. But usually f is only represented by a discrete data set $\{f(k)\}$. If suppϕ is small, the values $f(k)$ are nearly constant and if $\int_{\mathbb{R}} \phi dt = 1$ then we can use $a_{0k} = f(k)\,, \quad k \in \mathbb{Z}$.*

Notation 5.17 *When computing (5.4.2) and (5.4.3) we only need the coefficients h_k and g_k, not the functions ϕ and ψ!!! h_k and g_k are computed and saved once and used for all subsequent calculations (see tables below).*

5.4.2 Synthesis

1. Start: Given $\underline{a}_J,\ \underline{d}_1, ...\underline{d}_J$. We are looking for $\underline{a}_0 \mid P_0 f = \sum_k a_{0k}\phi_{0k}$.

2.

$$P_{j-1}f = P_j f + Q_j f = \sum_k a_{j,k}\phi_{j,k} + \sum_k d_{jk}\psi_{jk}$$

On the other hand we have

$$\begin{aligned}
P_{j-1}f &= \sum_k a_{j-1,k}\phi_{j-1,k} \\
a_{j-1,n} &= \left(P_{j-1}f, \phi_{j-1,n}\right) \\
&= \sum_k a_{jk}\left(\phi_{jk}, \phi_{j-1,n}\right) + \sum_k d_{jk}\left(\psi_{jk}, \phi_{j-1,n}\right) \\
&= \sum_k a_{jk}\left(\sum_{l=-\infty}^{\infty} h_l\phi_{j-1,2k+l}, \phi_{j-1,n}\right) + \sum_k d_{jk}\left(\sum_{l=-\infty}^{\infty} g_l\phi_{j-1,2k+l}, \phi_{j-1,n}\right) \\
&= \sum_k a_{jk}h_{n-2k} + \sum_k d_{jk}g_{n-2k} \qquad (5.4.4)
\end{aligned}$$

because $n = 2k + l$ implies $l = n - 2k$.

Calculation scheme:

$$\begin{tikzcd}
\underline{\boldsymbol{a}}_J \arrow[r, "\underline{\boldsymbol{h}}"] & \underline{\boldsymbol{a}}_{J-1} \arrow[r, "\underline{\boldsymbol{h}}"] & \underline{\boldsymbol{a}}_{J-2} & \cdots\cdots & \underline{\boldsymbol{a}}_1 \arrow[r, "\underline{\boldsymbol{h}}"] & \underline{\boldsymbol{a}}_0 \\
\underline{\boldsymbol{d}}_J \arrow[ur, "\underline{\boldsymbol{g}}"] & \underline{\boldsymbol{d}}_{J-1} \arrow[ur, "\underline{\boldsymbol{g}}"] & & \cdots\cdots & \underline{\boldsymbol{d}}_1 \arrow[ur, "\underline{\boldsymbol{g}}"] &
\end{tikzcd}$$

Complexity:

Assumption: ϕ has a compact support.

$\curvearrowright \quad a(\phi) = \inf\{t \mid \phi(t) \neq 0\} \overset{!}{=} 0$

$\curvearrowright \quad b(\phi) = \sup\{t \mid \phi(t) \neq 0\} \overset{!}{=} 2N - 1; \quad N \geq 1$

$\curvearrowright \quad h_k \neq 0; \quad g_k = (-1)^k\,\overline{h}_{2N-1-k} \neq 0 \quad \text{for} \quad 0 \leq k \leq 2N-1$

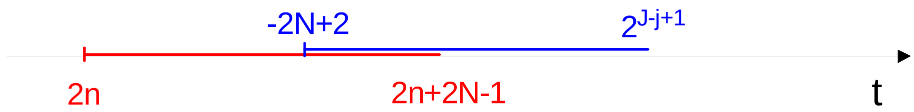
The vector $\underline{a}_0$ contains all information we want to use about the signal $f \in \mathbb{L}_2(\mathbb{R})$.

Theorem 5.22 *supp* $(\underline{a}_0) \subset [0\ ;\ 2^J)$; *length* $(\underline{a}_0) = 2^J$
implies *supp* $\left(\underline{a}_j\right) \subset [-2N+2\ ;\ 2^{J-j})$; *for* $j \geq 0$

Proof. Base clause: $j = 0$, *supp* $(\underline{a}_0) \subset [-2N+2\ ;\ 2^J)$: satisfied by assumption
Induction hypothesis: *supp* $\left(\underline{a}_{j-1}\right) \subset [-2N+2\ ;\ 2^{J-j+1})$

Induction step: $j \geq 1$: transition from $j-1$ to j :

$$\begin{aligned} a_{jn} &= \sum_{l=0}^{2N-1} \overline{h_l} a_{j-1,2n+l} \neq 0 \\ &\Longleftrightarrow \emptyset \neq \{2n, 2n+1, ..., 2n+2N-1\} \cap [-2N+2; 2^{J-j+1}) \\ &\Longleftrightarrow 2n < 2^{J-j+1} \quad \wedge \quad 2n+2N-1 \geq -2N+2 \\ &\Longleftrightarrow n < 2^{J-j} \quad \wedge \quad n \geq -2N+\frac{3}{2} \end{aligned}$$



$\curvearrowright n \in [-2N+2; 2^{J-j})$ ■
The process terminates after $j = J$ steps because *supp* $(\underline{a}_j)$ stagnates at $[-2N+2\ ;\ 0)$.

Now we are looking for the number of multiplications μ carried out up to this point:

$length\ (\underline{a}_j) = length\ (\underline{d}_j) \leq 2^{J-j} + 2N - 2$

Computation of a_{jn} requires at most $length(\underline{h}) = 2N$ multiplications.

$$\begin{aligned} \mu &\leq 2 \cdot 2N \cdot \sum_{j=1}^{J} (2^{J-j} + 2N - 2) \\ &= 2 \cdot 2N \cdot (2^J - 1 + J(2N-2)) \quad \text{(geometric series)} \\ &= 2 \cdot 2N \cdot 2^J \left(1 + \frac{-1 + J(2N-2)}{2^J}\right) \\ &= 2 \cdot length(\underline{h}) \cdot length\ (\underline{a}_0)\ (1 + o(1)) \\ &\sim O(length\ (\underline{a}_0)) \end{aligned}$$

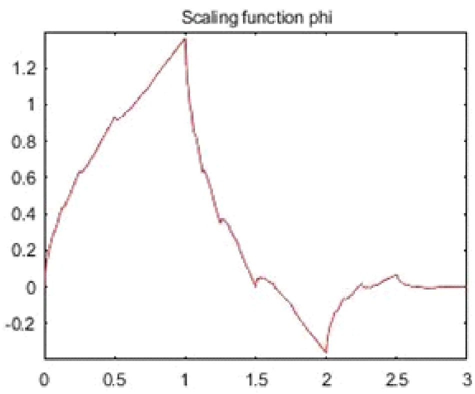
5.4.3 Tables

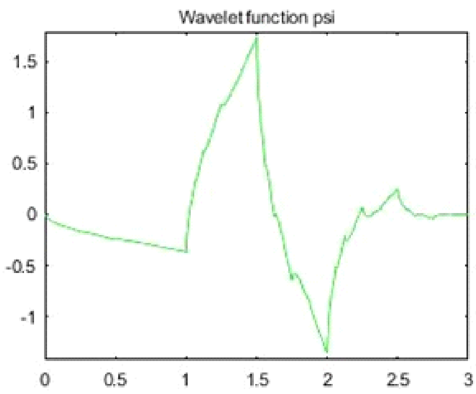
Tables of the coefficients h_k and g_k and figures of the mother wavelet and the Scaling functions are shown in [1], [2], [8]. For example the following representations were taken from there:

Example 5.10 *DAUBECHIES-Wavelets: Table of the coefficients h_k*
$N=1$ corresponds to the HAAR wavelet.

k	$N=1$	$N=2$	$N=3$	$N=4$	$N=5$
0	$\frac{1}{\sqrt{2}}$	$\frac{1-\sqrt{3}}{4\sqrt{2}}$	0.332671	0.230378	0.160102
1	$\frac{1}{\sqrt{2}}$	$\frac{3-\sqrt{3}}{4\sqrt{2}}$	0.806892	0.714847	0.603829
2		$\frac{3+\sqrt{3}}{4\sqrt{2}}$	0.459878	0.630881	0.724309
3		$\frac{1+\sqrt{3}}{4\sqrt{2}}$	− 0.135011	− 0.027984	0.138428
4			− 0.085441	− 0.187035	− 0.242295
5			0.035226	0.030841	− 0.032245
6				0.032883	0.077571
7				− 0.010597	− 0.006241
8					− 0.012581
9					0.003336

DAUBECHIES Wavelet for $N=2$:

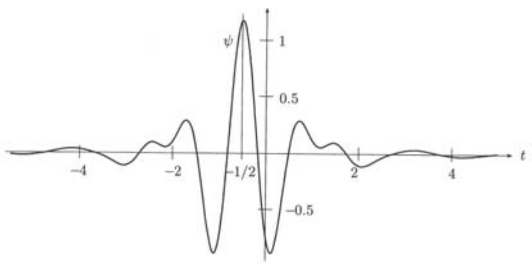




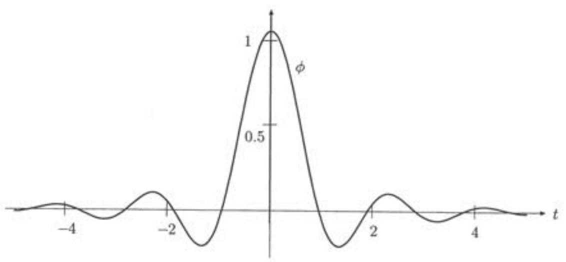
Example 5.11 *MEYER-Wavelet, table of the coefficients:*

k	$h_k = h_{-k}$
0	.748791
1	.442347
2	−.039431
3	−.127928
4	.033278
5	.057120
6	−.024807
7	−.025310
8	.016000
9	.009538
10	−.008556
11	−.002451
12	.003416
13	.000058
14	−.000647
15	.000225

k	$h_k = h_{-k}$
16	−.000329
17	.000061
18	.000333
19	−.000231
20	−.000059
21	.000174
22	−.000115
23	−.000027
24	.000115
25	−.000067
26	−.000028
27	.000066
28	−.000040
29	−.000015
30	.000046
31	−.000027



Meyer Wavelet with $n = 3$

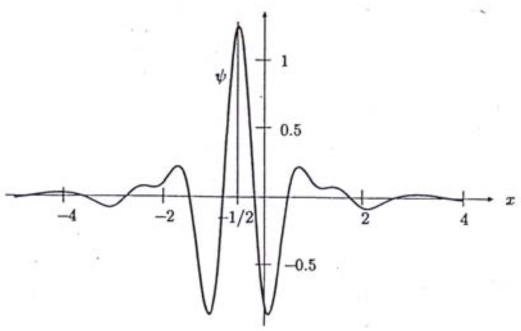


Meyer Scaling function with $n = 3$

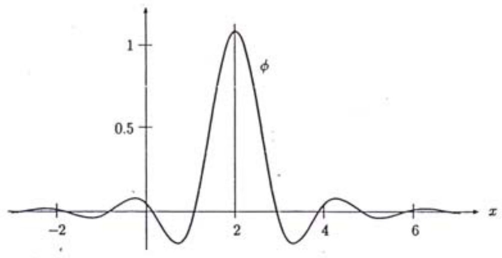
Example 5.12 *BATTLE-LEMARIE-Wavelet for* $n = 3$:
Table of the coefficients:

r	$h^\#_r = h^\#_{4-r}$
2	.766130
3	.433923
4	−.050202
5	−.110037
6	.032081
7	.042068
8	−.017176
9	−.017982
10	.008685
11	.008201
12	−.004354
13	−.003882
14	.002187
15	.001882
16	−.001104

r	$h^\#_r = h^\#_{4-r}$
17	−.000927
18	.000560
19	.000462
20	−.000285
21	−.000232
22	.000146
23	.000118
24	−.000075
25	−.000060
26	.000039
27	.000031
28	−.000020
29	−.000016
30	.000010
31	.000008



BATTLE-LEMARIE-
Wavelet with $n = 3$



BATTLE-LEMARIE-
Scaling function
with $n = 3$

6 Applications

The goal of signal processing is to extract information from the signal $s \in \mathbb{L}_2(\mathbb{R})$, for example the occurrence of

- predefined patterns,
- periodic components
- jumps and
- irregularities.

The wavelet transform is useful if the desired phenomena have multi-scale structure, such as edges, jumps, locally varying order of differentiability etc., which can be recognized by the asymptotic behavior at the discontinuity. This is a big advantage compared with the FOURIER transform, which „smeared" these phenomena over $\mathbb{R}$.

6.1 Preliminaries

In practice we know discrete values of a measured signal:

$$s_k = s(kh); \quad k \in \mathbb{Z}; \quad h > 0 : \text{ sampling rate.}$$

It is convenient to choose $h = 2^j$.

1. **Adaptation to the wavelet transformation:**
   It is possible to interpret the values s_k as expansion coefficients of a function $\widetilde{f}$ in accordance with the scaling function because $\left\{ \phi_{jk} = 2^{-j} \phi \left(\frac{t}{2^j} - k \right) \right\}_{k \in \mathbb{Z}}$ is an ONS in $\mathbb{V}_j$:

$$\widetilde{f}(t) = \sum_{k \in \mathbb{Z}} s_k \phi(h^{-1}t - k) = \sum_{k \in \mathbb{Z}} s(kh) \phi \left(\frac{t}{h} - k \right) \qquad \textit{with } h = 2^j. \tag{6.1.1}$$

   But we need the coefficients of an expansion of $\widetilde{f}$ in $\mathbb{V}_0$, calculated by s_k.
   1. Possibility: Use of special quadrature formulas for $a_{0k} = (f, \phi_{0k})_{\mathbb{L}_2(\mathbb{R})}$.
   2. Possibility: It can be shown:

$$\widetilde{f}(kh) = \frac{1}{h} \int_{\mathbb{R}} s(t) \phi \left(\frac{t}{h} - k \right) dt = s_k + O(h),$$

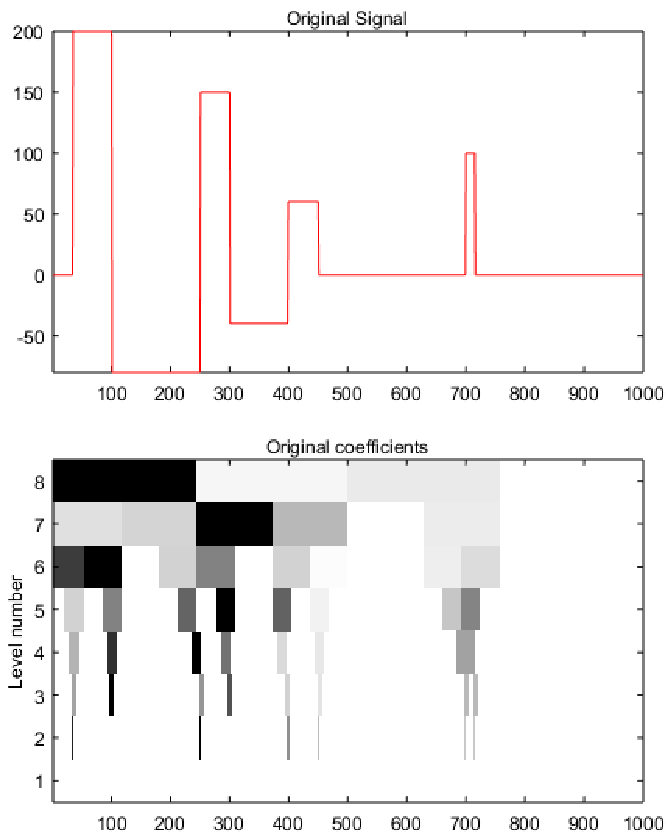
assumed that $s \in \mathbb{C}^1(\mathbb{R})$, $\int_{\mathbb{R}} \phi(t)\,dt = 1$, $\widetilde{f} = \widetilde{f}(t)$ ([2], p. 203). With the substitution $x = \frac{t}{h}$ respectively $t = hx$ we get by (6.1.1):

$$f(x) = \widetilde{f}(hx) = \sum_{k \in \mathbb{Z}} s_k \phi(\frac{hx}{h} - k) \in \mathbb{V}_0.$$

That's why the sampled signal $\{s_k\}$ can be used as input data!

2. **Displaying of the results of the wavelet transform:**
   Consider for example scale diagrams.
   Colour the rectangles of the $a-b-$plane corresponding to the to the size of the associated coefficient. If the mother wavelet is concentrated about μ then a large wavelet coefficient d_k^m means that a significant detail of size $h2^m$ occurs in f, in the neighborhood of the point $(\mu + hk)2^m$ in the scale 2^m. Thus the coefficient d_k^m and the rectangle of width $h2^m$ around the point $(\mu + hk)2^m$ are coupled.

   Example: scale diagram with 8 steps:



3. **Selection: Continuous Wavelet Transform (CWT) or Fast Discrete Wavelet Transform (FDWT)**
   This question is relevant, because the translation invariance of the CWT is lost by discretisation onto the plane using a logarithmic grid. Then the points $(2^m, n2^m)$ of the grid do not go into other grid points (see transparency no. 23).
   The numerical calculation of the CWT requires a large computational effort. Integrals of inner products must be computed by discretisation. If $f(nh), \quad n \in \mathbb{Z}$ is a sampled sequence, then it is useful to require the translation invariance only to multiples of h.
   $\curvearrowright$ discretisation $(a_m, b_n) = (s_m, nh)\,; \quad n \in \mathbb{Z}; \quad$ for every s_m
   $\curvearrowright$ application of quadrature formulae
   $\curvearrowright$ fine raster image
   The direct comparison of CWT and FDWT results:

CWT		FDWT
high	computational effort	less
yes	translational invariance	no
easy	interpretability	difficult
easy, but very high effort	extension to several dimensions	difficult, but already done

   To avoid this selection problem between CWT and FDWT a mixed form was designed: the „Algorithme à trous", which combines the advantages of both wavelet transforms. It is relatively fast: Assume the length of the signal is N.Then the FDWT calculates $2N$ coefficients and the „Algorithme a trous" $2NJ$ coefficients during J steps of the algorithm. The scale diagrams are „continuous" with respect to t and they are easier to interpret than diagrams of the FDWT [3] (see transparency no. 24).

6.2 Data Compression

Data compression is the most successful application of the wavelet transform, but it also works with the Fourier transform or the cosine transform. Data compression is necessary for the real-time transmission and saving of images. This makes high demands on computing time and memory capacity. The assessment is based on the compression rate k :

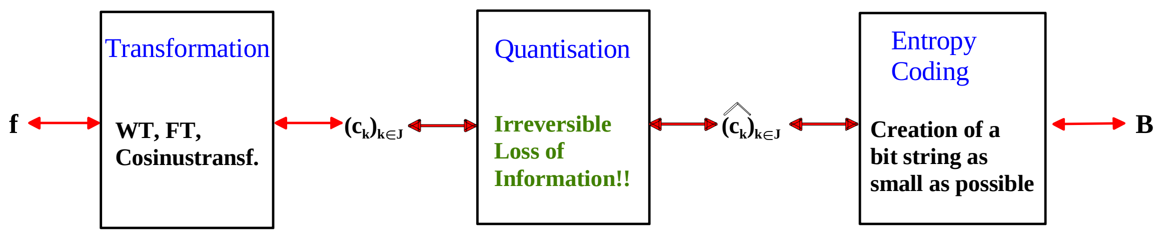
$$k = \frac{\text{Memory requirements of the original}}{\text{Memory requirements of the compressed file}}$$

Technical details:

- A digitalized black and white image corresponds to a (n, n) −matrix of gray values of the respective image pixels.

- Coloured images correspond to 3 pictures for the colours red, blue and green. This RGB representation is transferred into a another representation which uses one brightness and two colour values because brightness variations are registered by the eye very well, however, less color differences.

- Principle:



Transform:
For the Fourier and cosine transforms the image must first be broken down into sub-images for resolving local details because sine and cosine functions have unrestricted support. The boundaries of the sub-images are usually visible after the inverse transform with large k.
In the case of the wavelet transform it is possible to use the two-dimensional wavelet transform. In this case no decomposition into sub-images is required. But a more common case is the use of tensor wavelets. In this case rows and columns are transformed by one-dimensional algorithms.Then we get sequences of matrices with the decomposition coefficients which can be saved in place. The selection of the mother wavelet is a problem because the choice depends on the structure of the signals.
Quantisation:
The simplest type of quantisation is the uniform scalar quantisation. This means rounding to integer multiples of the quantisation step $\triangle > 0$. If „quantisation tables" are used, then they have a significant influence on the quality of compressed images. The error arises here. After the inverse transform it manifests itself in the approximated signal values $\widetilde{f}$. $f - \widetilde{f}$ is called quantisation noise. Measurements are, for example, $\left\| f - \widetilde{f} \right\|$ or $\left\| f - \widetilde{f} \right\|^2 = \sum\limits_{k \in J} |c_k - \widetilde{c}_k|^2$ or the MSE (mean square error) $= \frac{1}{N} \sum\limits_{k=0}^{N-1} \left| f_k - \widetilde{f}_k \right|^2$. If M is the allowed extension of the range of f, then we obtain for the ratio $\frac{M^2}{MSE}$ measured

in decibels

$$PSNR = 10\log_{10}\frac{M^2}{MSE} \quad or \quad PSNR = 10\log_{10}\frac{NM^2}{\sum\limits_{k=0}^{N-1}\left|f_k-\widetilde{f}_k\right|^2}$$

(PSNR: Peak Signal to Noise Ratio) by sampling of N values, $h = 1$.

Entropy-Coding:

Techniques like Huffman coding or run-length coding are used to get vectors as close as possible to the optimal bit length.

For examples of original and compressed images see [2] and transparencies no. 26-27.

6.3 Denoising - Noise Suppression

Denoising by wavelet transform gives the best results with the least effort. As above, thresholding is performed instead of quantisation (small wavelet coefficients are zeroized). By the inverse transform we then get a less noisy signal. To obtain good results, the following requirements must be satisfied

- The noise level $\|f^{noise}\|$ is small in relation to $\left\|f^{original}\right\|$.
- The original signal can be well represented by few coefficients in the new basis. Therefore it is very suitable for compression.
- The noise is not well compressed in the new basis . (Random noise cannot be compressed by any ONS.)

There are essentially two options for thresholding:

1. „Hard Thresholding":

$$c_k = \begin{cases} 0 & \text{for } |c_k| \leq \tau \\ c_k & \text{otherwise} \end{cases}$$

2. „Soft Thresholding":

$$c_k = \begin{cases} 0 & \text{for } |c_k| \leq \tau \\ sgn(c_k)(|c_k|-\tau) & \text{otherwise} \end{cases}$$

For the definition of τ, a statistical model for the noise is required, for example white noise with the standard deviation σ is used

$$f_k = f_k^{original} + \sigma g_k, \quad g_k : \text{values of } N(0,1), \quad k = 0, \ldots N-1.$$

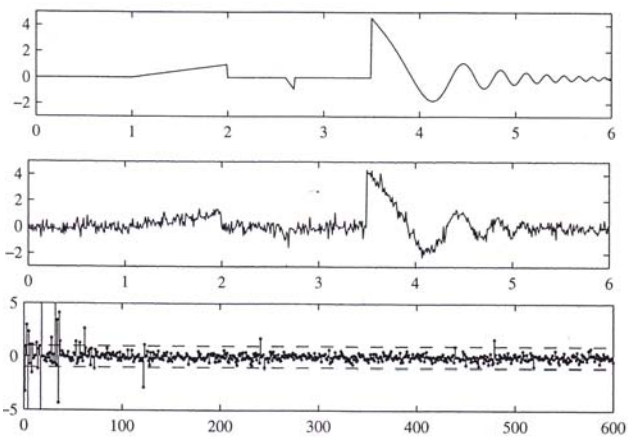
This results in an estimation for τ and σ :

$$\begin{aligned}\tau &= \kappa\sqrt{2\ln(N)}\sigma; \quad \kappa = O(1)\\ \sigma &\approx \frac{\text{Median of the modulus of the wavelet coefficients}}{0.6745}\end{aligned}$$

Application of denoising:

- Clarification of audio signals
- speech recognition
- Preparing data for ill-conditioned inversion problems, such as in computer tomography or economic data ....

Example 6.1 *Clarification of a noisy signal* [3]*:*

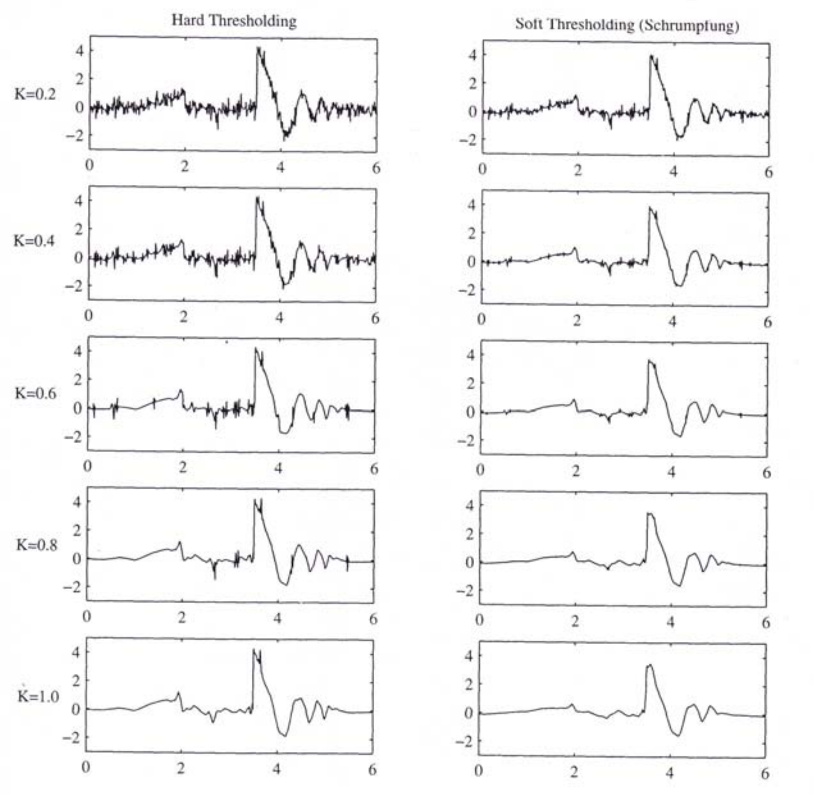


original signal

noisy signal, $\sigma = 0.3$

wavelet coefficients, $K = 1$

$K = \frac{1}{k}$ *with the compression rate* k.



6.4 Feature Detection

A feature detection algorithm finds many applications: such as in medical examination for the evaluation of ECG, EEG, ultrasonography, phonocardiogram etc. Comparisons between the wavelet transform, the FOURIER transform, the Short-Time FOURIER transform, the WIGNER distribution and other transforms always bring the same result: The wavelet transform gives the best results. The reason for this lies in the transient character of biological signals. Basically, these signals are periodic functions with short time disorders and/or varying period and/or with a high noise level. Usually it is desirable to detect short peaks in the signal which can be of high energy and which can recur at irregular intervals. Problems lie in the delimitation of noise, in the delineation of similar signal components and the definition of artifacts.

Example 6.2 *ECG* [2]
For example the following questions are investigated: Is the rhythm of the heart valves synchronous to the rhythm of the main heart muscle? Is the main muscle relaxed bet-

ween contractions? If not, then that is an acute alarm signal of sudden cardiac death and other serious heart diseases (see transparency no. 29).

Example 6.3 *EEG* [6]
In the evaluation of EEG epileptic spikes are sought. These are short energetic spikes, followed by long flat waves. These epileptic spikes are also found between attacks and they are used for the diagnosis.
The signal is a sampled electrical voltage. The energy of the signals is proportional to $\|f\|$ *, and by PLANCHEREL's theorem we get :*

$$\begin{aligned} \|f\|^2 &= \int |f(t)|^2 dt = \frac{1}{C} \iint \left| \frac{Wf(a,b)}{a} \right|^2 dadb \\ &\sim \sum_l \frac{1}{l^2} \sum_k |c_{lk}|^2. \end{aligned}$$

(see transparencies no. 30-31)

Example 6.4 *Reliability of gear systems* [5]
The problem is to optimize the maintenance strategy, to ensure the reliability of the gear drive. The gear wheels are the greatest source of error, and thus they are the focus of the diagnosis. Therefore the vibration signal which is caused by the gearbox will be analysed. This is a long and difficult process. It goes from a rotor synchronized sampling, over denoising and elimination of other periodic but non rotor synchronous signal components, to the analysis of the true vibration signal of the gearbox by CWT with the Morlet wavelet and to a user-friendly angle-order representation (see transparencies no. 32-36).

There are many other uses for the wavelet transform, for example edge detection in images, troubleshooting in woven and knitted fabrics, image analysis in mammography, effective memory algorithms, broadband communications .... In mathematics, the wavelet basis will also be investigated for use in the Galerkin method for boundary value problems for partial differential equations.