Basic Concepts of Digital Signal Processing

Prof. Dr. Michael Clausen, Dr. Meinard Müller
Institut für Informatik III
Römerstraße 164
Rheinische Friedrich-Wilhelms-Universität Bonn
{clausen, meinard}@cs.uni-bonn.de

Summer School 2003
International Program of Excellence (IPEC)
Bonn-Aachen International Center for Information Technology (B-IT)

Preface

These slides were used in the lecture "Basic Concepts of Digital Signal Processing" held within the Summer School 2003 of the International Programme of Excellence (IPEC) of the Bonn-Aachen International Center for Information Technology (B-IT). We hope that these slides give the students a rough guideline and a summary of the main concepts covered by the lecture. However, these slides are not meant to be self-contained or complete. Further details, illustrations, proofs, and explanations are given in the lecture.

The authors would be grateful for any comments and suggestions for improvement.

Bonn, August 2003

Michael Clausen
Meinard Müller

Contents

10 The Two-Dimensional (2D) Case

Literature

Chapter 1: Signals and Signal Spaces

1.1 Motivation and Examples

In their classical book "Digital Signal Processing." [Oppenheim/Schafer] the authors give the following characterization of a signal.

> A signal can be defined as a function that conveys information, generally about the state of behavior of a physical system. Although signals can be represented in many ways, in all cases the information is contained in a pattern of variations of some form. For example, the signal may take the form of a pattern of time variations or a spatially varying pattern. Signals are represented mathematically as functions of one or more independent variables. For example, a speech signal would be represented mathematically as a function of time and a picture would be represented as a brightness function of two spatial variables.

Signals include, for example,

- the sound of a piano as a signal represented by the amplitude of sound with respect to time,
- the distribution of light on a screen, or
- the light falling on a point on a surface or in space.

We first consider the case of a music signal. From a physical point of view, the musician, by means of his voice or instrument, generates an audio signal. This audio signal has the form of a sound wave emerging at its source and spreading through the air. Graphically, such an audio signal may be represented by its waveform which depicts the amplitude of the air pressure over the time:

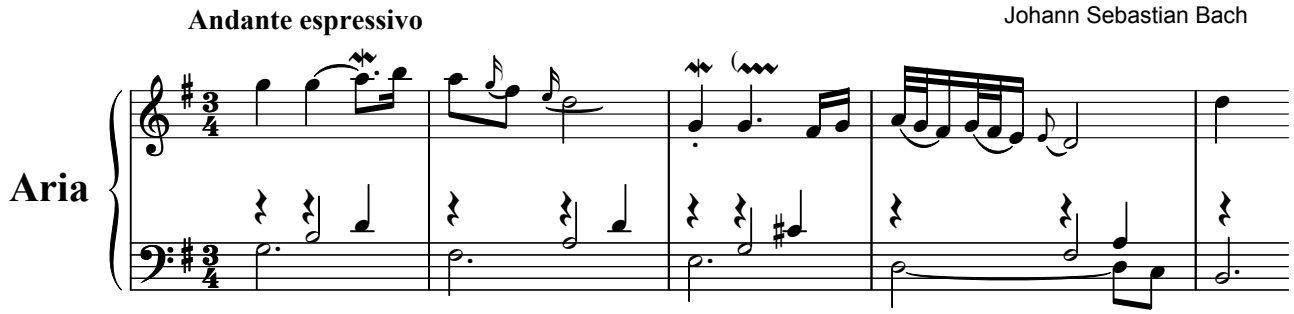


Figure 1: Beginning of the Aria con Variazioni by J. S. Bach, BWV 988.

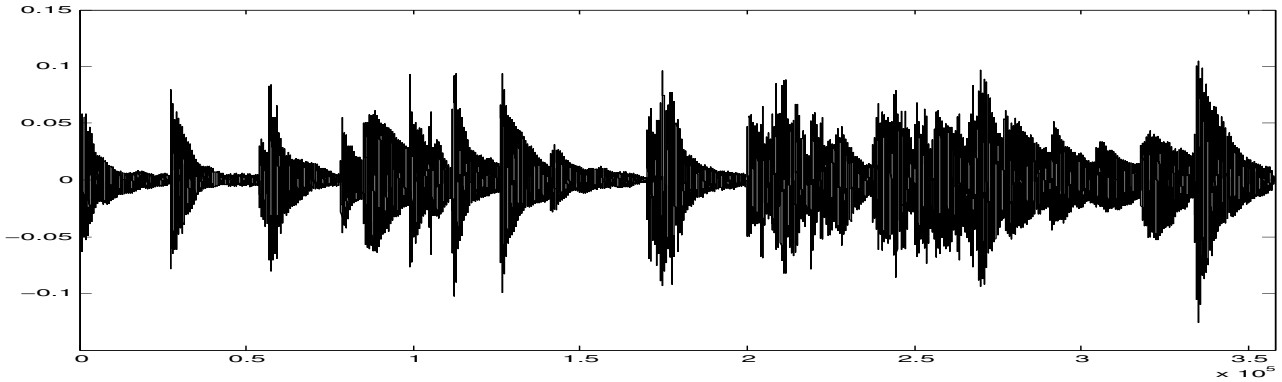


Figure 2: Waveform representation of an Aria interpretation of Fig. 1

In case of a digital image, the signal may be represented as a matrix with entries in the range $[0 : 255]$. Here each entry of the matrix represents a pixel and the value of the entry encodes the gray level of the pixel.

Figure 3: Signal represented as 256×256 matrix over $[0 : 255]$.

From a mathematical point of view a signal is just a function. Basically there are two different kinds of signals:

(1) continuous-time (CT) (or analog) signals

(2) discrete-time (DT) signals, also called discrete or sampled signals

For example, a physical sound wave can be thought of as a CT signal. However, a computer can only store a finite set of numbers and (usually) can only compute with finite precision. Therefore, in computer-based signal processing a continuous-time signal is transformed into some discrete signal where the signal is stored as a collection of samples. This transformation is also known as sampling. The discrete signal is only defined at particular, discrete locations.

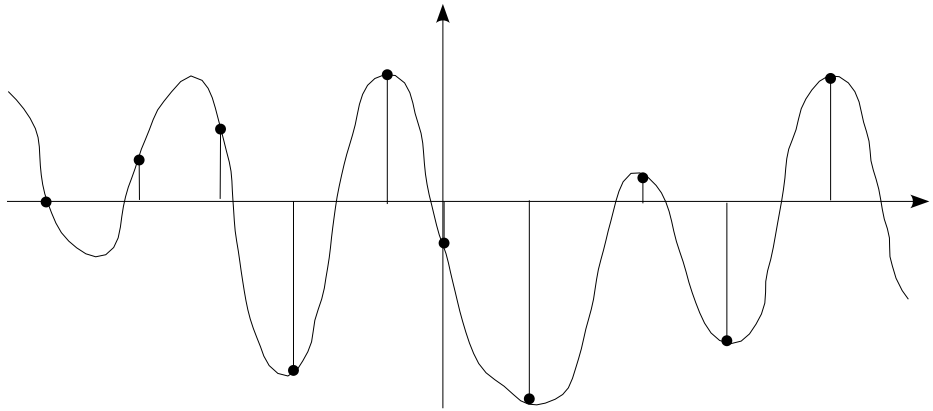


Figure 4: Example for sampling.

Digital signal processing involves the creation of digital signals from continuous-time signals. However, by sampling one looses in general information about the signal. The side effects caused by the discrete representation of some continuous signal are known as aliasing. In order to understand these side effects and the connection of the continuous and discrete world we need a thorough mathematical modeling of the concepts involved. In this chapter we start with the definition of signals and signal spaces.

1.2 Signals

1.2.1 Continuous-time (CT) signals

Depending on the nature of the signal, a continuous-time (CT) signal can be mathematically modeled as a

- one-dimensional real-valued function $f : \mathbb{R} \longrightarrow \mathbb{R}$
- one-dimensional complex-valued function $f : \mathbb{R} \longrightarrow \mathbb{C}$

For example, an audio signal may be represented by a function $f : \mathbb{R} \longrightarrow \mathbb{R}$, where the domain $\mathbb{R}$ represents the time axis and the range $\mathbb{R}$ the amplitude of the sound wave.

Complex-valued functions generalize the concept of real-valued functions.

Why complex-valued functions? $\rightsquigarrow$ Fourier transform

A signal, such as an image, can be modeled as a

- two-dimensional real-valued function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$
- two-dimensional complex-valued function $f : \mathbb{R}^2 \longrightarrow \mathbb{C}$

Here the domain $\mathbb{R}^2$ represents the two dimensional space (two spatial coordinates) and the range $\mathbb{R}$ the gray level or color value.

In general, a signal is a parametric function of any number of input dimension $n \in \mathbb{N}$ and output dimensions $m \in \mathbb{N}$:

- n-dimensional $\mathbb{R}$-vector-valued function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$
- n-dimensional $\mathbb{C}$-vector-valued function $f : \mathbb{R}^n \longrightarrow \mathbb{C}^m$

For example, a 3D-scene could be modeled by a 3-dimensional domain $\mathbb{R}^3$ with three spatial coordinates. Or a moving 3D-scene may by modeled by a 4-dimensional domain $\mathbb{R}^4$ with three spatial coordinates and one time coordinate.

Attention:

- A continuous-time signal need not to be continuous as a function!
- The term "continous-time" is unfortunate because it implies that the parameter of the function is a time value. In fact, the parameter (or argument) may represent anything, including time, space, or any "continuous parameter". Hence continuous-parameter would be a better expression. But the term "continous-time" is firmly established in literature, so we will use it here.

In the following we concentrate on signals of the form $f : \mathbb{R} \longrightarrow \mathbb{C}$ (audio signals) and later generalize to the case $f : \mathbb{R}^2 \longrightarrow \mathbb{C}$ (images).

Example 1.1. Let A, f be real numbers, then the function

$$y(t) = A \sin(2\pi f t)$$

defines a sine wave with amplitude A and frequency f. From a musical point of view, A represents the volume and f the pitch of the signal.

Example 1.2. The continuous-time box function has value 1 within some interval of width $W \in \mathbb{R}$ centered at $t = 0$, and is 0 outside:

$$b_W(t) := \begin{cases} 1 & \text{if } |t| \leq W/2, \\ 0 & \text{otherwise.} \end{cases}$$

Note that this is a continuous-time signal which is not continuous as a function.

Example 1.3. The continuous-time impulse signal, written δ, is zero everywhere except for $t = 0$ where it has an infinite value. It is an infinitely narrow spike of infinite height, which integrates to a value of 1. Strictly, we should call this a distribution. We define the impulse signal, also called the Dirac delta function, by

$$\delta(t) = 0 \text{ if } \quad t \neq 0, \quad \text{ and } \quad \int_{-\infty}^{+\infty} \delta(t)dt = 1.$$

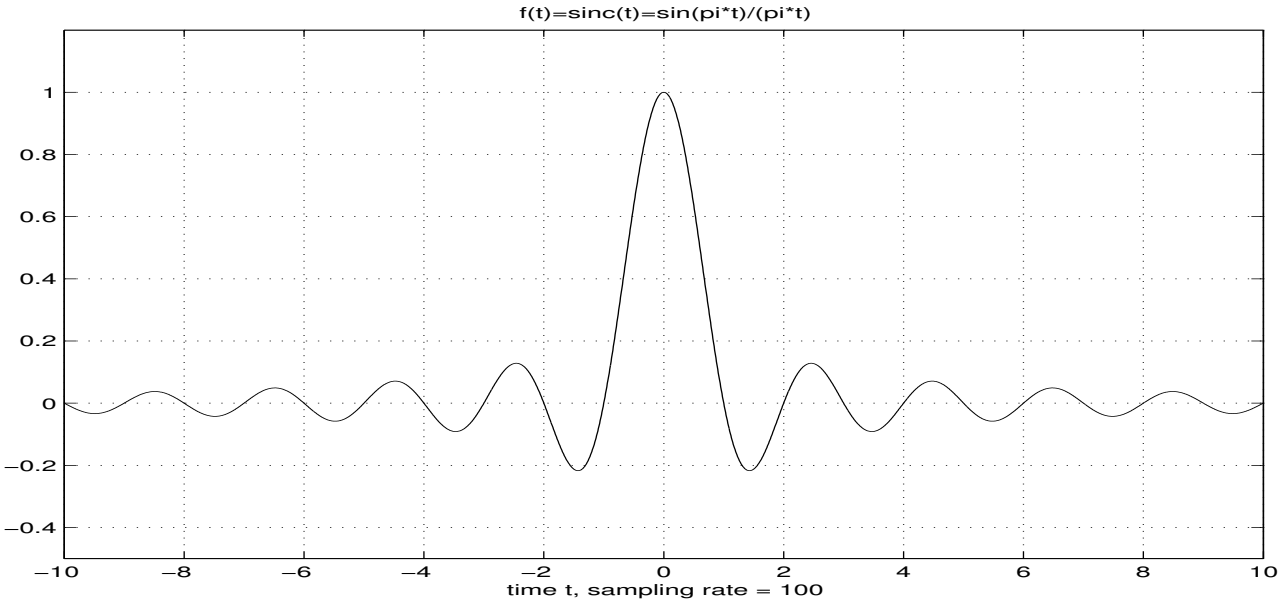
The impulse signal may be viewed as the derivative of the unit step function defined by

$$u(t) := \begin{cases} 0 & \text{if } t < 0, \\ 1 & \text{if } t \geq 0. \end{cases}$$

Example 1.4. A notational convenience is provided by the sinc (pronounced "sink") signal defined by:

$$\operatorname{sinc}(t) := \begin{cases} \frac{\sin \pi t}{\pi t} & \text{if } t \neq 0, \\ 1 & \text{if } t = 0. \end{cases}$$

Prove as an exercise that this function is continuous. The sinc-function has a number of remarkable properties which will play a crucial role in later chapters.



1.2.2 Discrete-time (DT) signals

In contrast to a continuous-time signal a discrete-time (DT) signal is defined only on a discrete set of (temporal or spatial) parameters. Mathematically one may assume that this discrete set of parameters is given by integer numbers. Hence, a DT-signal can be modeled as in the continuous case replacing the domain $\mathbb{R}$ by $\mathbb{Z}$:

- one-dimensional real-valued function $x : \mathbb{Z} \longrightarrow \mathbb{R}$
- one-dimensional complex-valued function $x : \mathbb{Z} \longrightarrow \mathbb{C}$
- two-dimensional real-valued function $x : \mathbb{Z}^2 \longrightarrow \mathbb{R}$
- two-dimensional complex-valued function $x : \mathbb{Z}^2 \longrightarrow \mathbb{C}$
- n-dimensional $\mathbb{R}$-vector-valued function $x : \mathbb{Z}^n \longrightarrow \mathbb{R}^m$
- n-dimensional $\mathbb{C}$-vector-valued function $x : \mathbb{Z}^n \longrightarrow \mathbb{C}^m$

We often use the symbols f,g to denote CT-signals and the symbols x,y to denote DT-signals. For the parameter we often use t in the CT- and n in the DT-case.

As mentioned before, DT-signals may be obtained from CT-signals by sampling. In this lecture we will just consider the case of equidistant sampling.

Definition 1.5. Let $f : \mathbb{R} \longrightarrow \mathbb{C}$ be a CT-signal. For a real number $T > 0$ we define T-sampling of f to be the discrete function $x\colon \mathbb{Z} \longrightarrow \mathbb{C}$ given by

$$x(n) := f(T \cdot n).$$

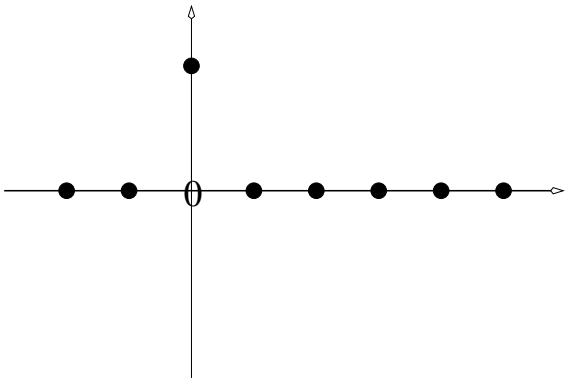
The sampling rate is the number of samples per second and is measured with the unit Hertz (Hz). Hence the sampling rate of x is $\frac{1}{T}$ Hz.

The sampling rate is crucial for the quality of the signal. Common sampling rates for speech and music signals are as follows (1 kHz $= 1000$ Hz):

- telephone: 8 kHz
- digital radio: 32 kHz
- CD: 44,1 kHz
- professional studio technology : 48 kHz

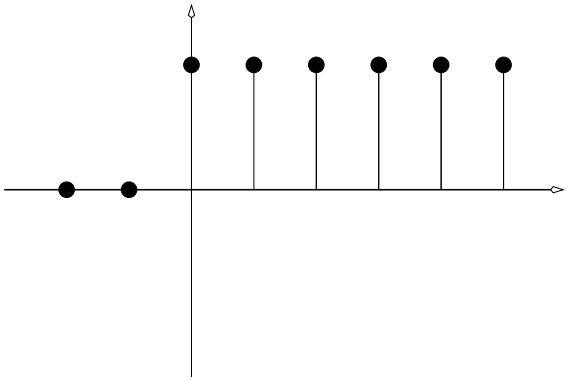
Example 1.6. The discrete-time unit impulse, written δ, is zero everywhere except for $n = 0$, where it has the value 1:

$$\delta(n) := \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n \neq 0. \end{cases}$$



Example 1.7. The discrete-time unit step function $u\colon \mathbb{Z} \longrightarrow \mathbb{C}$ is for all negative $n \in \mathbb{Z}$ zero and for all non-negative n equals 1:

$$u(n) := \begin{cases} 0 & \text{if } n < 0, \\ 1 & \text{if } n \geq 0. \end{cases}$$



Example 1.8. The discrete exponential signals are sequences of the form $(c \cdot a^n)_{n\in\mathbb{Z}}$, with arbitrary constants $a, c \in \mathbb{C}$. Sequences of the form $(c \cdot a^n u(n))_{n\in\mathbb{Z}}$ or $(c \cdot a^n u(-n))_{n\in\mathbb{Z}}$ are called truncated exponentials.

Example 1.9. The discrete frequency signals of frequency $\omega \in [0, 1)$ are special exponential signals of the form $(c \cdot e^{2\pi i \omega n})_{n \in \mathbb{Z}}$ with some complex $c \neq 0$. Frequency signals of frequency ω are periodic if and only if ω is a rational multiple of 1:

$$\omega = k/L, \quad \text{for suitable } k, L \in \mathbb{Z},\ L \neq 0.$$

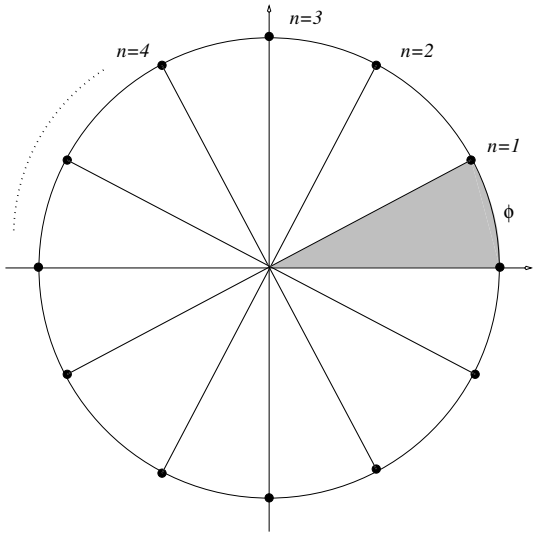
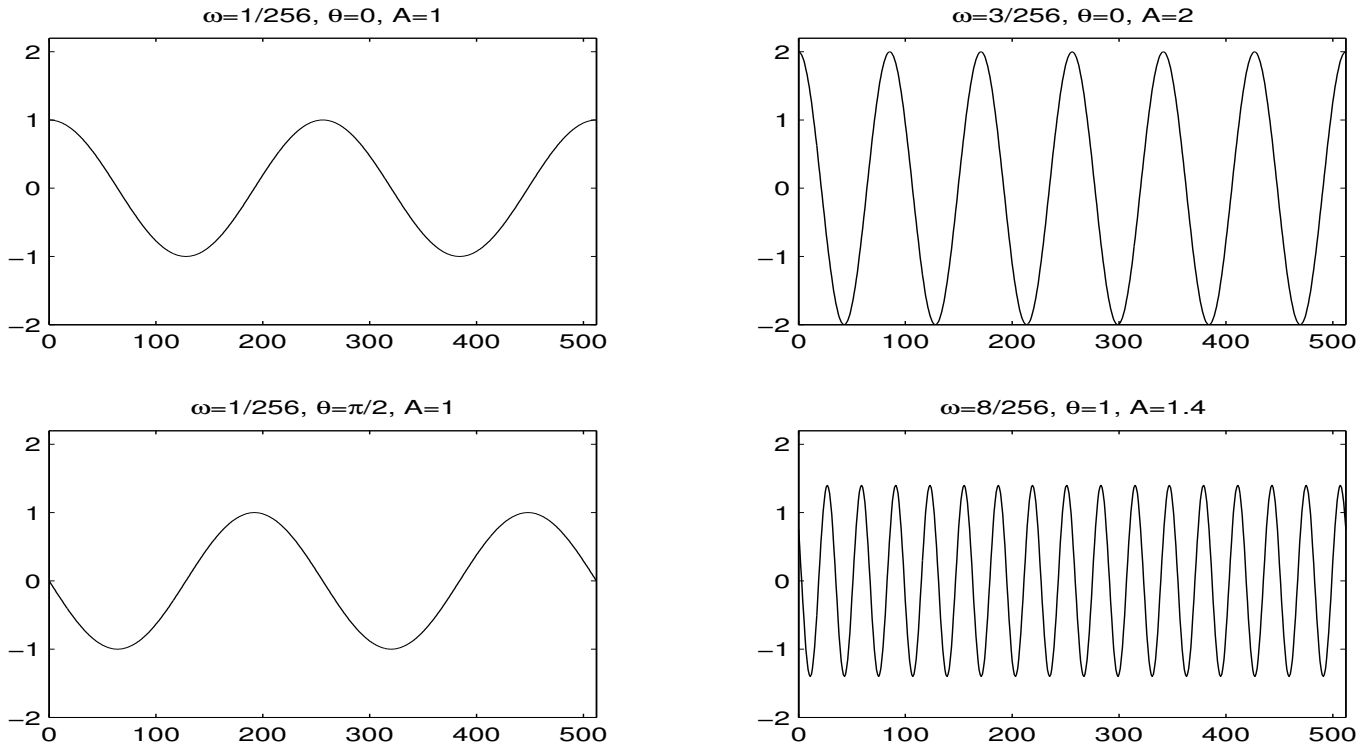


Figure 5: Geometric interpretation of the exponential signal $a^n = e^{2\pi i n \omega}$

Example 1.10. Sine-like signals are of the form $n \mapsto A\cos(2\pi\omega n + \theta)$ with real θ. This signal contains the frequencies ω as well as $-\omega$, since

$$\cos(2\pi\omega n + \theta) = \tfrac{1}{2}\left(e^{2\pi i(\omega n+\theta)} + e^{-2\pi i(\omega n+\theta)}\right).$$



1.3 Signal Spaces

Phenomenons such as superposition or amplification of signals — one may think of audio signals generated by the instruments of an orchestra — can be modeled by means of suitable operations which may by applied for signals from the same class or space. The following discussion holds for CT-signals as well as for DT-signals. Let D denote either the CT-parameter space $\mathbb{R}$ or the DT-parameter space $\mathbb{Z}$.

Definition 1.11. Let $x\colon D \to \mathbb{C}$, $y\colon D \to \mathbb{C}$ be signals and $\lambda \in \mathbb{C}$. Then the superposition of x and y is mathematically given by the sum $x + y$ defined by

$$(x + y)(t) := x(t) + y(t) \text{ for } t \in D.$$

The amplification of the signal x by the factor λ is mathematically given by the scalar-multiple λx defined by

$$(\lambda x)(t) := \lambda \cdot x(t) \text{ for } t \in D.$$

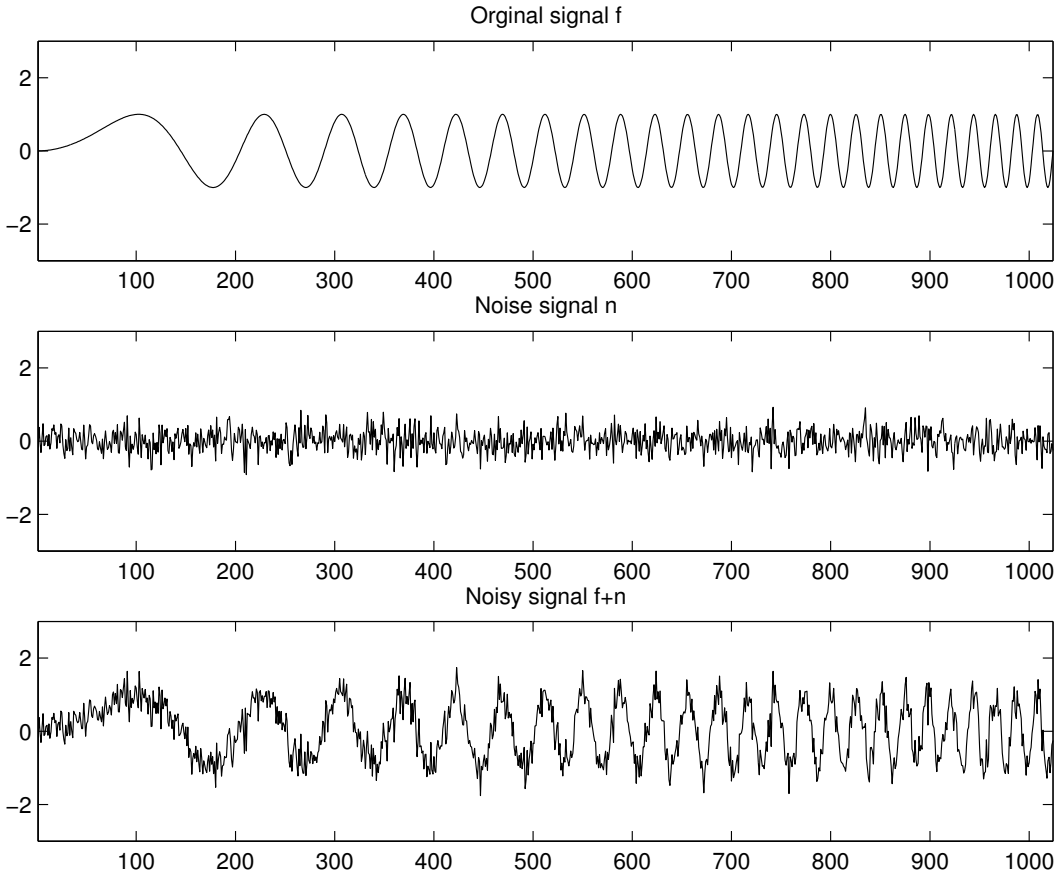
Theorem 1.12. *The set* $\mathbb{C}^D := \{x | x\colon D \to \mathbb{C}\}$ *with above addition and scalar multiplication defines a* $\mathbb{C}$*-vector space with* $\dim(\mathbb{C}^D) = \infty$ *in case* $|D| = \infty$.

Example 1.13. We consider the superposition of two sound waves of (nearly) equal frequency:

(1) Let $x(t) = Ae^{i\omega t}$ and $y(t) = Be^{i\omega t}$ with real A, B. Dependent on A and B we can observe the following:
$A > 0, B > 0 \Rightarrow$ amplification of the signal,
$A > 0, B < 0 \Rightarrow$ attenuation of the signal,
$A = -B \Rightarrow$ annihilation of the signals.

(2) Let $x(t) = Ae^{i\omega t}$ and $y(t) = Be^{i(\omega+\epsilon)t}$. The sum of x and y yields $(x + y)(t) = (A + Be^{i\epsilon t})e^{i\omega t}$, which indicates the so-called amplitude vibrato.

Example 1.14. The superposition of a signal with a noise signal is a typical example for the addition of two signals.



The spaces $\mathbb{C}^{\mathbb{Z}}$ or $\mathbb{C}^{\mathbb{R}}$ are far too large. From a physical perspective one is interested in signals of finite energy. From a mathematical point of view one is interested in spaces of signals which allow certain manipulations such as Fourier transforms or Wavelet transforms. Before we define suitable signal spaces (which will all be linear subspaces of $\mathbb{C}^{\mathbb{Z}}$ or $\mathbb{C}^{\mathbb{R}}$) we introduce some mathematical notations in the next subsection.

1.3.1 Banach Spaces and Hilbert Spaces

In signal processing one often has the problem to compare two signals with each other. For example, one is interested to determine "the distance" between two signals or one is interested to determine "the size" or "energy" of a signal. Next we remind on the general mathematical definition of such concepts.

Definition 1.15. A metric d on a set M is a map $d\colon M \times M \longrightarrow \mathbb{R}$ such that

- $d(x, y) \geq 0$,
- $d(x, y) = 0$ iff $x = y$,
- $d(x, y) = d(y, x)$,
- $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality) .

Definition 1.16. A **norm** on a $\mathbb{C}$-vector space V is a map $\|\cdot\|: V \longrightarrow \mathbb{R}_{\geq 0}$ such that

- $\|x\| = 0$ iff $x = 0$,
- $\|\lambda x\| = |\lambda| \cdot \|x\|$,
- $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality).

Note 1.17. Each norm induces a metric d on V via $d(x, y) := \|x - y\|$

Definition 1.18. A normed vector space V is called complete iff every Cauchy sequence in V converges in V.

Definition 1.19. A complete normed vector space is called Banach space.

Definition 1.20. An inner product or scalar product on a $\mathbb{C}$-vector space V is a map $\langle \cdot | \cdot \rangle \colon V \times V \to \mathbb{C}$ such that

- $\langle x|x\rangle \geq 0$, $= 0$ if and only if $x = 0$,
- $\langle x|y\rangle = \overline{\langle y|x\rangle}$,
- $\langle \cdot | \cdot \rangle$ is $\mathbb{C}$-linear in the first component.

Note 1.21. Each inner product induces a norm on V via $\|x\| := \sqrt{\langle x|x\rangle}$.

Definition 1.22. A Banach space, which norm is induced by an inner product, is called a Hilbert space.

These definitions are standard mathematical notions and can be found in any introductory textbook on (functional) analysis.

Example 1.23. The Hilbert space $\mathbb{C}^n$ with standard inner product is defined by the scalar product

$$\langle x|y\rangle := \sum_{i=1}^{n} x(i)\overline{y(i)}$$

for any $x = (x(1), x(2), \ldots, x(n)), y = (y(1), y(2), \ldots, y(n)) \in \mathbb{C}^n$.

The fact that there are orthonormal bases (ON-bases) in $\mathbb{C}^n$ with respect to the standard inner product generalizes to arbitrary Hilbert spaces. The following theorem characterizes ON-systems:

Theorem 1.24. *Let I be a countable set and $(x_i)_{i\in I}$ be an ON-system in the Hilbert space X, i.e., $\langle x_i|x_j\rangle = \delta_{ij}$ for $i, j \in I$. Then the following is equivalent:*

(1) *(Completeness) If $x \in X$ is orthogonal to all x_i, then $x = 0$.*

(2) *(Parseval-equality) For each $x \in X$ holds:*

$$\|x\|^2 = \sum_{i\in I} |\langle x|x_i\rangle|^2.$$

(3) *Each $x \in X$ has the following (generalized) Fourier expansion:*

$$x = \sum_{i\in I} \langle x|x_i\rangle x_i,$$

where on the right side at most a countable number of terms is nonzero and the series converges to x with respect to the norm regardless of the order of the summands.

Such a system is called a Hilbert basis of X.

Theorem 1.25. *Each Hilbert space has a Hilbert basis, and two Hilbert bases are of the same cardinality. This cardinality is called Hilbert dimension of X.*

Note 1.26. The norm and the inner product on a vector space are additional structures which allow to generalize many (geometric) constructions known from the two and three dimensional case to the higher dimensional or even infinite dimensional case:

(i) The norm $\|x\|$ allows to speak of the size of some vector or signal $x \in X$. In the case the norm is induced by an inner product one also calls $\|x\|^2$ the energy of x.

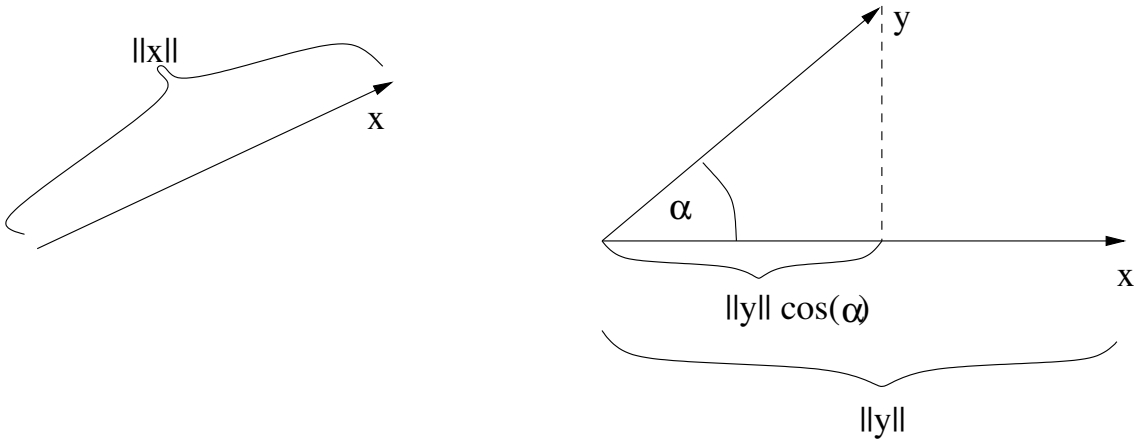
(ii) In a Hilbert space the norm $\|x\|^2 = \langle x|x\rangle$ is directly linked with the inner product. In this case the Cauchy-Schwarz inequality

$$|\langle x|y\rangle| \leq \|x\|\|y\|,$$

holds, which is for many estimations an indispensable mathematical tool.

(iii) By the inner product one can generalize the geometric concepts such as angles, orthogonality and orthonormality. This allows to define the concept of orthogonal subspaces and projection operators into these subspaces which will play a crucial role in Wavelet decompositions:

$$\langle x|y\rangle = \|x\| \cdot \|y\| \cdot \cos(\alpha)$$



(iv) Of fundamental importance in signal processing is the Fourier transform $x \mapsto \hat{x}$, to be defined below. On a certain Hilbert space this transform leaves the norm as well as the inner product invariant. This is the so-called Parseval equality:

$$\|x\| = \|\hat{x}\| \quad \text{and} \quad \langle x|y\rangle = \langle \hat{x}|\hat{y}\rangle .$$

This property is extremely useful for the frequency analysis of signals.

In view of the signal spaces introduced in the following sections a rough intuitive understanding of the following measure theoretic concepts is helpful. However, the foundations of measure theory and Lebesgue integration are of rather technical nature and a detailed introduction to this topic would require a lecture by itself.

- Riemann integral
- Borel measure on $\mathbb{R}^n$
- Lebesgue integral

Details can be found in [Folland].

1.3.2 Lebesgue Space $\ell^p(\mathbb{Z})$ for DT-Signals

We start with signal spaces of discrete-time signals and deal with the more difficult continuous-time case in the next subsection.

Definition 1.27. Let $1 \leq p < \infty$ be a real number. The (discrete-time) Lebesgue space $\ell^p(\mathbb{Z})$ consists of all sequences $x\colon \mathbb{Z} \to \mathbb{C}$ with $\sum_{n\in\mathbb{Z}} |x(n)|^p < \infty$:

$$\ell^p(\mathbb{Z}) := \{x\colon \mathbb{Z} \to \mathbb{C} \mid \sum_{n\in\mathbb{Z}} |x(n)|^p < \infty\}.$$

For $p = \infty$ let $\ell^\infty(\mathbb{Z})$ be the space of bounded signals with domain $\mathbb{Z}$:

$$\ell^\infty(\mathbb{Z}) := \{x\colon \mathbb{Z} \to \mathbb{C} \mid \exists B > 0\colon \forall n \in Z : \ |x(n)| \leq B\}.$$

The number p has the intuitive meaning to control the error sensitivity. A large p means that small errors are attenuated and large errors are amplified.

These classes of signals are closed under addition and scalar multiplication:

Theorem 1.28. *For each $1 \leq p \leq \infty$ the class $\ell^p(\mathbb{Z})$ is a linear subspace of $\mathbb{C}^{\mathbb{Z}}$.*

Proof: We have to show the following properties of $\ell^p(\mathbb{Z})$:

- $0 \in \ell^p(\mathbb{Z})$
- $x \in \ell^p(\mathbb{Z}), \lambda \in \mathbb{C} \Rightarrow \lambda x \in \ell^p(\mathbb{Z})$
- $x \in \ell^p(\mathbb{Z}), y \in \ell^p(\mathbb{Z}) \Rightarrow x + y \in \ell^p(\mathbb{Z})$

The first two properties are easy to see. In the case $p = \infty$ the third property is also easily seen. For $1 \leq p < \infty$ the third property follows from

$$
\begin{aligned}
\sum_{n\in\mathbb{Z}} |(x+y)(n)|^p &= \sum_{n\in\mathbb{Z}} |x(n)+y(n)|^p \\
&\leq \sum_{n\in\mathbb{Z}} (|x(n)|+|y(n)|)^p \\
&\leq \sum_{n\in\mathbb{Z}} (2\max\{|x(n)|,|y(n)|\})^p \\
&\leq 2^p \sum_{n\in\mathbb{Z}} (|x(n)|^p + |y(n)|^p) \\
&= \underbrace{2^p(\sum_{n\in\mathbb{Z}} |x(n)|^p)}_{<\infty} + \underbrace{2^p(\sum_{n\in\mathbb{Z}} |y(n)|^p)}_{<\infty} \\
&< \infty
\end{aligned}
$$

□

Theorem 1.29. *The maps*

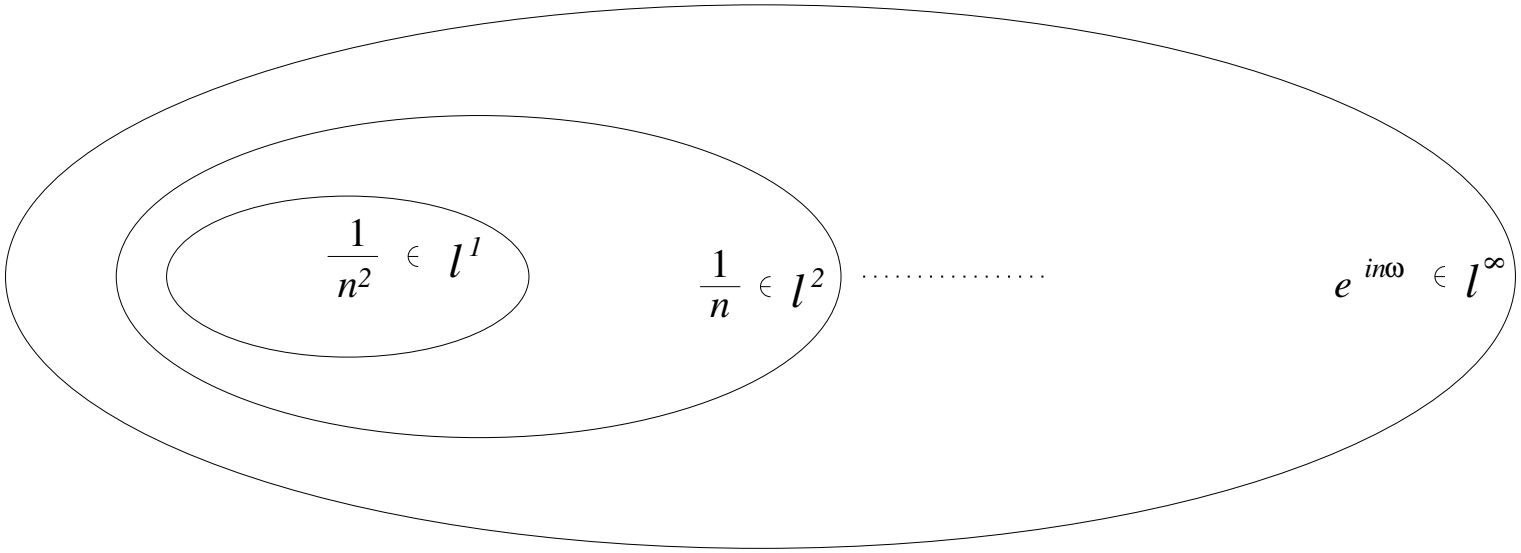
$$\begin{aligned} \|x\|_p &:= \left(\sum_{n\in\mathbb{Z}} |x(n)|^p \right)^{1/p} \quad \textit{for } 1 \le p < \infty \ \textit{and} \\ \|x\|_\infty &:= \sup\{|x(n)| : n \in \mathbb{Z}\} \end{aligned}$$

define a norm on $\ell^p(\mathbb{Z})$ and $\ell^\infty(\mathbb{Z})$, respectively. These spaces are complete with respect to the norms and, therefore, are Banach spaces.

A proof of this (not at all obvious) theorem can be found in [Folland].

Note 1.30. Let $1 \leq p < q \leq \infty$, then $\ell^p(\mathbb{Z}) \subseteq \ell^q(\mathbb{Z})$ and $\|x\|_q \leq \|x\|_p$ for all $x \in \ell^p(\mathbb{Z})$. The inclusion $\ell^p(\mathbb{Z}) \subseteq \ell^q(\mathbb{Z})$ is proper, i.e., there is some $x \in \ell^q(\mathbb{Z})$ with $x \notin \ell^p(\mathbb{Z})$. For example, the frequency sequences $(e^{i\omega n})_{n\in\mathbb{Z}}$ are in $\ell^\infty(\mathbb{Z})$ but not in $\ell^p(\mathbb{Z})$ for any $p < \infty$.

In the following figure some typical sequences are indicated.



For our considerations the three spaces $\ell^1(\mathbb{Z})$, $\ell^2(\mathbb{Z})$ and $\ell^\infty(\mathbb{Z})$ are of special interest.

$$\begin{aligned} \ell^1(\mathbb{Z}) &= \text{space of absolute-summable sequences} \\ \ell^2(\mathbb{Z}) &= \text{space of quadratic-summable sequences} \\ \ell^\infty(\mathbb{Z}) &= \text{space of bounded sequences.} \end{aligned}$$

Lemma 1.31. *The space $\ell^2(\mathbb{Z})$ is a Hilbert space with respect to the inner product defined by $\langle x|y\rangle := \sum_{n\in\mathbb{Z}} x(n)\overline{y(n)}$, $x, y \in \ell^2(\mathbb{Z})$. The indicator functions $(\delta_n)_{n\in\mathbb{Z}}$ of the elements of $\mathbb{Z}$ define a Hilbert basis (ON-basis) of $\ell^2(\mathbb{Z})$.*

Note that there are many other Hilbert bases of $\ell^2(\mathbb{Z})$.

1.3.3 Lebesgue Space $L^p(\mathbb{R})$ for CT-Signals

In this subsection we introduce the continuous-time signal spaces $L^p(\mathbb{R})$ which are the CT-counterpart of the DT-spaces $\ell^p(\mathbb{Z})$. From a formal point of view, one replaces $\mathbb{Z}$ by $\mathbb{R}$ and summation by integration. Some results carry over from the DT-case to the CT-case. However, there are also many phenomena in the CT-case which do not appear in the DT-case. One of the reasons is that the CT-parameter space $\mathbb{R}$ is much "bigger" than the DT-parameter space $\mathbb{Z}$ (for example, $\mathbb{R}$ is uncountable whereas $\mathbb{Z}$ is countable). Again we refer for the proofs to [Folland] and summarize the main definitions and properties.

Definition 1.32. Let $1 \leq p < \infty$ be a real number. The (continuous-time) Lebesgue space $L^p(\mathbb{R})$ consists of all functions $f\colon \mathbb{R} \to \mathbb{C}$ with $\int_{\mathbb{R}} |f(t)|^p dt < \infty$:

$$L^p(\mathbb{R}) := \{f\colon \mathbb{R} \to \mathbb{C} \mid \int_{\mathbb{R}} |f(t)|^p dt < \infty\}.$$

For $p = \infty$ let $L^\infty(\mathbb{R})$ be the space of essentially bounded signals with domain $\mathbb{R}$:

$$L^\infty(\mathbb{R}) := \{f\colon \mathbb{R} \to \mathbb{C} \mid \text{ess } \sup_{t\in\mathbb{R}} |f(t)| < \infty.$$

By definition ess $\sup_{t\in\mathbb{R}} |f(t)| := \inf\{a \geq 0 | \mu(\{x : |f(x)| > a\}) = 0\}$, where μ denotes the so-called Borel measure on $\mathbb{R}$.

Theorem 1.33. *For each $1 \leq p \leq \infty$ the class $L^p(\mathbb{R})$ is a linear subspace of $\mathbb{C}^{\mathbb{R}}$.*

Theorem 1.34. *The maps*

$$\begin{aligned} \|f\|_p &:= \sqrt[p]{\int_{\mathbb{R}} |f(t)|^p dt} \quad \text{für} \quad 1 \le p < \infty \\ \|f\|_\infty &:= \operatorname{ess\,sup}_{t\in\mathbb{R}} |f(t)| \end{aligned}$$

define a norm on $L^p(\mathbb{R})$ *and* $L^\infty(\mathbb{R})$ *respectively. These spaces are complete with respect to the norm and hence are Banach spaces.*

Note 1.35. Strictly speaking, the spaces $L^p(\mathbb{R})$ consists of equivalence classes of functions: two functions $f, g \in L^p(\mathbb{R})$ are considered as equal when $\|f - g\|_p = 0$. For further details concerning the L^p-spaces we refer to [Folland] or the *dtv-Atlas Mathematik II*.

Note 1.36. For the continuous-time Lebesgue-spaces one does not have an inclusion property as in the discrete-time case (see Note 1.30). For example, for the functions $f, g \in \mathbb{C}^{\mathbb{R}}$ defined by

$$f(t) := \begin{cases} \sqrt{\frac{1}{t}} & \text{if } t \in (0,1], \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad g(t) := \begin{cases} \frac{1}{t} & \text{if } t \in [1,+\infty), \\ 0 & \text{otherwise} \end{cases}$$

holds $f \in L^1(\mathbb{R}) \setminus L^2(\mathbb{R})$ and $g \in L^2(\mathbb{R}) \setminus L^1(\mathbb{R})$.

Lemma 1.37. *The space $L^2(\mathbb{R})$ is a Hilbert space with respect to the inner product defined by $\langle f|g\rangle := \int_{\mathbb{R}} f(t)\overline{g(t)}dt$, $f, g \in L^2(\mathbb{R})$.*

1.3.4 Lebesgue Space $L^p([0,1])$

In the last subsection we have defined the time-continuous signal spaces $L^p(\mathbb{R})$. In some sense, signals in $L^p(\mathbb{R})$ can be viewed as elements in $\mathbb{C}^{\mathbb{R}}$ satisfying some integrability condition.

In this subsection we introduce another class of time-continuous signals — the class of periodic signals — which is of fundamental importance.

Definition 1.38. A signal $f\colon \mathbb{R} \to \mathbb{C}$ is periodic of period $\lambda \in \mathbb{R}$ if for all $t \in \mathbb{R}$ holds $f(t) = f(t+\lambda)$.

Note 1.39. The following observations are more or less obvious.

(i) Any non-zero periodic function is not in $L^p(\mathbb{R})$ for $1 \leq p < \infty$.

(ii) Any periodic function f of period λ is already known when restricted to the interval $[0, \lambda]$.

(iii) Contrary any function $g\colon [0, \lambda] \longrightarrow \mathbb{C}$ can be extended in an obvious fashion to a periodic function $f\colon \mathbb{R} \longrightarrow \mathbb{C}$ of period λ.

(iv) For a λ-periodic function f the function defined by $t \mapsto f(\lambda \cdot)$ is 1-periodic, i.e., by applying the linear transformation $t \mapsto \lambda \cdot t$ one can switch from periodic functions with arbitrary period λ to the case where $\lambda = 1$. Hence, in the following we may assume $\lambda = 1$.

By the above note the space $\mathbb{C}^{[0,1]}$ coincides with the space of 1-periodic functions. Similar to the non-periodic one can now define linear subspaces $L^p([0, 1])$ for $1 \leq p < \infty$ which turn out to be Banach spaces. In the following we restrict to the case $p = 2$.

Theorem 1.40. *The space* $L^2([0,1]) := \{f\colon [0,1] \to \mathbb{C} \mid \int_0^1 |f(t)|^2 dt < \infty\}$ *of square-integrable* 1*-periodic functions is a Hilbert space with respect to the inner product*

$$\langle f|g\rangle := \int_0^1 f(t)\overline{g(t)}dt, \quad f,g \in L^2([0,1]).$$

Note 1.41. Similarly, for $a, b \in \mathbb{R}$, $a < b$, one can define the Hilbert space $L^2([a,b])$ of λ-periodic functions with $\lambda = b - a$.

Chapter 2: Fourier Transform

Barbara Burke Hubbard gives in her book "The world according to wavelets." [Hubbard] the following nice characterization of the Fourier transform:

> The Fourier transform is the mathematical procedure that breaks up a function into the frequencies that compose it, as a prism breaks up light into colors. It transforms a function f that depends on time into a new function, $\hat{f}$, which depends on frequency. This new function is called the Fourier transform of the original function (or, when the original function is periodic, its Fourier series).

A function and its Fourier transform are two faces of the same information:

- The function displays the time information and hides the information about frequencies. Intuitively, a signal corresponding to a musical recording shows when the notes are played (change of the air pressure) but not which notes are played.
- The Fourier transform displays information about frequencies and hides the time information. Intuitively, the Fourier transform of music tells what notes are played, but it is extremely difficult to figure out when they are played.

2.1 Fourier Series for Periodic CT-Signals

In the Hilbert space $H = L^2([0,1])$ there are two bases which are of special interest. The proof of the following theorem can be found in most books on Functional Analysis.

Theorem 2.1. *The Hilbert space $L^2([0,1])$ has (among others) the following two ON-bases:*

(1) $\{1, \sqrt{2}\cos(2\pi kt), \sqrt{2}\sin(2\pi kt) | k \in \mathbb{N}\}$

(2) $\{\mathbf{e}_k | k \in \mathbb{Z}\}$ *with* $\mathbf{e}_k(t) := e^{2\pi ikt}$ *for* $t \in [0,1]$.

Due to this theorem each $f \in L^2([0,1])$ can be expanded by a so-called Fourier series w.r.t (1)

$$f(t) = a_0 + \sqrt{2}\sum_{k=1}^{\infty} a_k \cos(2\pi kt) + \sqrt{2}\sum_{k=1}^{\infty} b_k \sin(2\pi kt).$$

The Fourier coefficients w.r.t. (1) are given by the inner products of the signal f with the basis functions of the ON-basis:

$$\begin{aligned} a_0 &= \langle f|1\rangle = \int_0^1 f(t)dt \\ a_k &= \langle f|\sqrt{2}\cos(2\pi kt)\rangle = \sqrt{2}\int_0^1 f(t)\cos(2\pi kt)dt \\ b_k &= \langle f|\sqrt{2}\sin(2\pi kt)\rangle f = \sqrt{2}\int_0^1 f(t)\sin(2\pi kt)dt \end{aligned}$$

The Fourier coefficient a_k expresses to which extend the functions $t \longrightarrow \cos(2\pi kt)$ (i.e., cosine function of frequency k Hertz) is "contained" in f. A similar interpretation holds for the coefficients b_k. A Fourier series takes only integer frequency $k \in \mathbb{N}$ into account. Note that the functions $t \mapsto \cos(2\pi kt)$ and $t \mapsto \sin(2\pi kt)$ represent the same frequency and differ only by some translations which is referred to as different phases.

Expansion of a signal $f \in L^2([0,1])$ with respect to the complex-valued ON-basis $\{\mathbf{e}_k : k \in \mathbb{Z}\}$ of $L^2([0,1])$ in (2) of Theorem 2.1 leads to the equality

$$f(t) = \sum_{k=-\infty}^{\infty} c_k \mathbf{e}_k = \sum_{k=-\infty}^{\infty} c_k e^{2\pi ikt}.$$

This expansion is also called Fourier series — this time w.r.t. (2). The coefficients

$$c_k = \langle f | e^{2\pi ikt} \rangle = \int_0^1 f(t) \overline{e^{2\pi ikt}} dt = \int_0^1 f(t) e^{-2\pi ikt} dt$$

are again called Fourier coefficients (w.r.t (2)).

The real and complex Fourier transform are closely related. Recall that $e^{2\pi ikt} = \cos(2\pi kt) + i\sin(2\pi kt)$. Then it is easy to see that

$$\begin{aligned} c_0 &= a_0 \\ c_k &= \frac{1}{\sqrt{2}}a_k - i\frac{1}{\sqrt{2}}b_k, \quad k > 0, \\ c_k &= \frac{1}{\sqrt{2}}a_{-k} + i\frac{1}{\sqrt{2}}b_{-k}, \quad k < 0, \end{aligned}$$

Similarly a_k and b_k can be recovered from the c_k. For notational reasons, the Fourier series w.r.t (2) is much easier to deal with. Therefore, we consider in the following only the Fourier series w.r.t. (2).

Note 2.2. The equality in the Fourier expansion is just an equality in the L^2-sense, i.e., equality up to a null set. Under additional conditions on f one also has pointwise equality. For example, in case f is continuously differentiable the Fourier series converges uniformly on $[0, 1]$ to f.

The following theorem says that $L^2([0,1])$ can be identified with $\ell^2(\mathbb{Z})$ via the Fourier coefficients. This is a special case of the general theory of Hilbert spaces and ON-systems (Parseval identity)

Theorem 2.3. *The function*

$$f \mapsto \hat{f} := (\langle f|\mathbf{e}_k\rangle)_{k\in\mathbb{Z}},$$

which assign to each signal $f \in L^2([0,1])$ *the sequence of Fourier coefficients, is a Hilbert space isomorphism:*

$$L^2([0,1]) \xrightarrow{\simeq} \ell^2(\mathbb{Z}).$$

In particular, for $f, g \in L^2([0,1])$ *holds*

$$\langle f|g\rangle_{L^2([0,1])} = \langle \hat{f}|\hat{g}\rangle_{l^2(\mathbb{Z})}.$$

The general case $L^2([a,b])$ for $a, b \in \mathbb{R}$, $a < b$, consisting of λ-periodic functions with $\lambda = b - a$ can be easily reduced to the above case $a = 0$ and $b = 1$. For example, the Fourier series transfers to the general case as follows.

Lemma 2.4. *Let $f \in L^2([a,b])$. Then one obtains the following representation as Fourier series of f:*

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{\frac{2\pi i k t}{b-a}}.$$

with coefficients

$$c_k = \frac{1}{b-a} \int_a^b f(t) e^{\frac{-2\pi i k t}{b-a}} dt.$$

2.2 Fourier Integral for non-periodic CT-Signals

For non-periodic continuous-time signals one can generalize the idea of the Fourier series. However, in this case the frequencies of integer values $k \in \mathbb{Z}$ do, in general, not suffice to "describe" a signal completely. Considering all frequencies $\omega \in \mathbb{R}$ and replacing summation by integration one gets the following "continuous" analog to the Fourier series:

Theorem 2.5. *For each signal $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ holds the equality*

$$f(t) = \int_{-\infty}^{\infty} c_\omega e^{2\pi i \omega t} d\omega \tag{1}$$

where c_ω is defined by

$$c_\omega = \int_{-\infty}^{\infty} f(t) e^{-2\pi i \omega t} dt. \tag{2}$$

Note 2.6. In the following let $\mathbf{e}_\omega \colon \mathbb{R} \longrightarrow \mathbb{C}$ denote the continuous exponential or frequency functions $t \mapsto e^{2\pi i \omega t}$ of frequency $\omega \in \mathbb{R}$.

(i) The assumption $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is a technical condition such that all integrals involved exist (i.e., are finite). Actually, there are even weaker conditions on f such that the integrals stills exist.

(ii) The equality (1) shows that any signal f (which satisfies a certain integrability condition) can be written as a (continuous) superposition of the frequency functions $\mathbf{e}_\omega$.

(iii) The number c_ω expresses the "intensity" with which the frequency function e_ω is "contained" in the signal f. Hence the numbers c_ω now play the role of the Fourier coefficients c_k in the Fourier series.

(iv) Note that the frequency functions $\mathbf{e}_\omega$ are $\frac{1}{\omega}$-periodic and are not contained in $L^p(\mathbb{R})$ for $1 \leq p < \infty$.

Definition 2.7. Let $f \in L^1(\mathbb{R})$ then the function $\hat{f}\colon \mathbb{R} \longrightarrow \mathbb{C}$ defined by

$$\hat{f}(\omega) := c_\omega = \int_{-\infty}^{\infty} f(t) e^{-2\pi i \omega t} dt, \quad \omega \in \mathbb{R},$$

is called Fourier integral or Fourier transform of f. Sometimes $\hat{f}$ is also denoted by $F(f)$.

It can be shown, that the definition of a Fourier transform of functions $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ can be extended to all signals $f \in L^2(\mathbb{R})$. (This is a non-trivial mathematical construction using the so-called Hahn-Banach Theorem.) The next theorem says that the Fourier transform is invariant under the inner product and hence preserves energy.

Theorem 2.8. (Plancherel) *The Fourier transform $f \mapsto \hat{f}$ defines a unitary transformation on $L^2(\mathbb{R})$. Hence, for $f \in L^2(\mathbb{R})$ holds $\hat{f} \in L^2(\mathbb{R})$ and $\|f\| = \|\hat{f}\|$. Furthermore, one has $\langle f | g \rangle = \langle \hat{f} | \hat{g} \rangle$ for any two functions $f, g \in L^2(\mathbb{R})$.*

Theorem 2.9. Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ (or more general $f \in L^2(\mathbb{R})$. Then the Fourier transform has the following properties:

(1) For $t_0 \in \mathbb{R}$, the translation of f by t_0 is defined by

$$f_{t_0}(t) := f(t - t_0).$$

Then

$$\widehat{f_{t_0}}(\omega) = e^{-2\pi i \omega t_0} \hat{f}(\omega).$$

(2) For $\omega_0 \in \mathbb{R}$, the modulation of f by ω_0 is defined by

$$f^{\omega_0}(t) := e^{-2\pi i \omega_0 t} f(t).$$

Then

$$\widehat{f^{\omega_0}}(\omega) = \hat{f}(\omega + \omega_0).$$

(3) Let f be differentiable with $f' \in L^2(\mathbb{R})$. Then

$$\widehat{f}'(\omega) = 2\pi i \omega \hat{f}(\omega).$$

(4) Let $\hat{f}$ be differentiable. Then

$$\hat{f}'(\omega) = -2\pi i \widehat{(t \mapsto tf(t))}(\omega).$$

(5) For $s \in \mathbb{R} \setminus \{0\}$ the scaled function $t \mapsto f(t/s)$ by s is also in $L^2(\mathbb{R})$ and

$$\widehat{f(\tfrac{\cdot}{s})}(\omega) = s\hat{f}(\omega s).$$

Proof: The proof, which amounts to a straightforward computation, is left as an exercise. □

Definition 2.10. Let $g \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$. The inverse Fourier transform of g is denoted by $\check{g}$ and defined by the integral

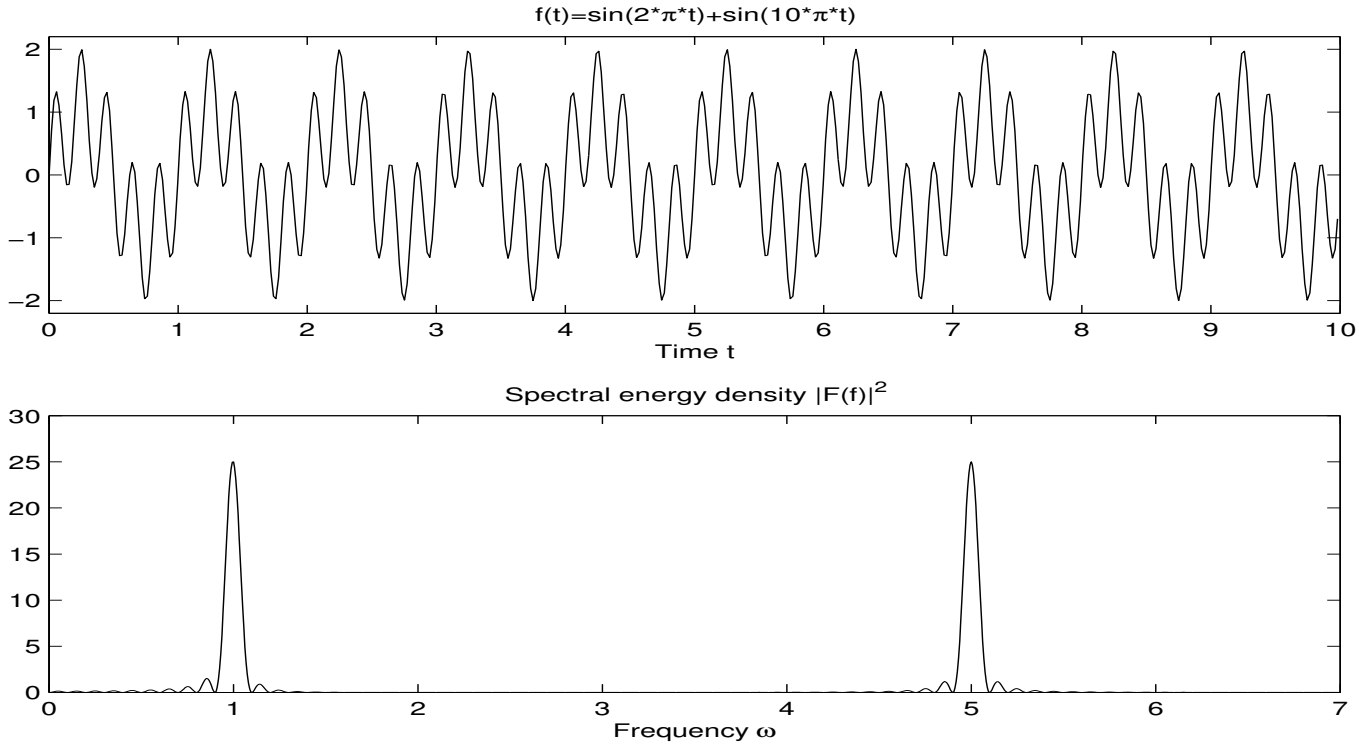
$$\check{g}(t) := \int_{-\infty}^{\infty} g(\omega) e^{2\pi i t\omega} d\omega.$$

It is easy to see that one has $\check{g}(t) = \hat{g}(-t)$. It is more difficult to show the next theorem whose proof can be found in [Folland].

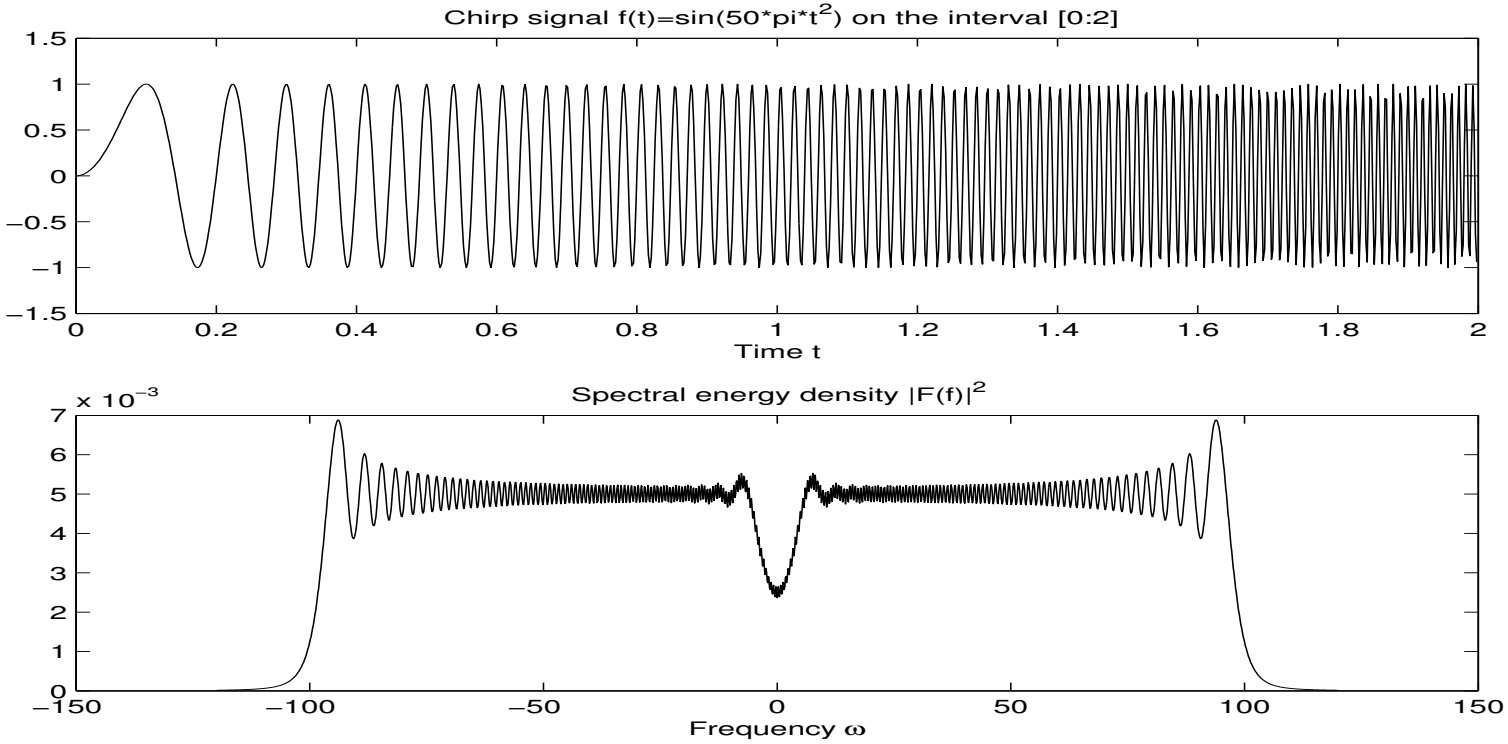
Theorem 2.11. *Let $g = \hat{f}$ be the Fourier transform of some signal $f \in L^2(\mathbb{R})$. Then $g \in L^2(\mathbb{R})$ and $\check{g} = f$. In other words, one has the identities*

$$(\hat{f})^\vee = f = (\check{f})^\wedge.$$

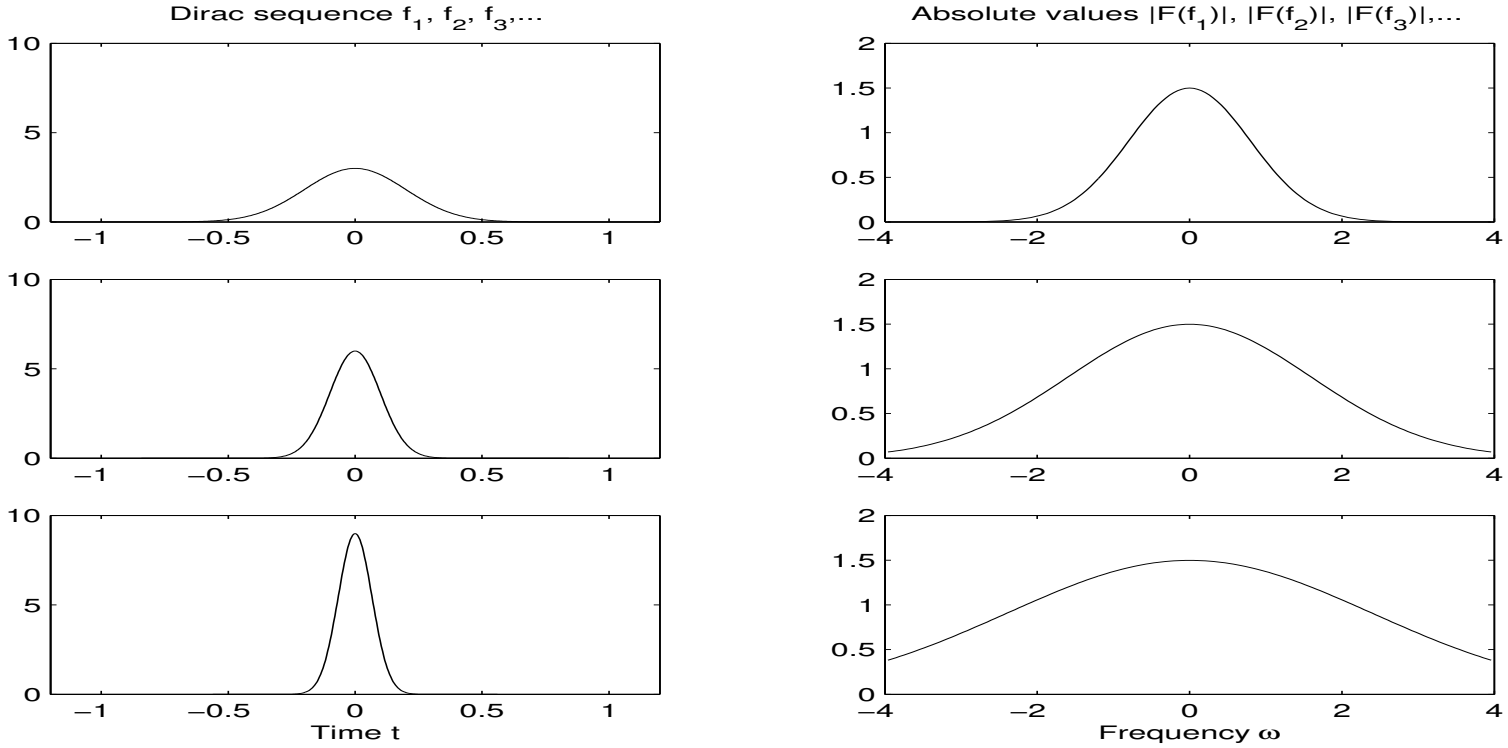
Example 2.12. In this example, the CT-function f is a superposition of two sines of frequency 1 Hz and 5 Hz on the interval $[0, 10]$ and zero outside this interval. The ripples in the spectrum come from the discontinuity of the signal at the boundaries $\rightsquigarrow$ "destructive interference".



Example 2.13. For the chirp signal f defined by $f(t) = \sin(50 \cdot \pi t^2)$, the frequency ω_0 at time $t = t_0$ is roughly given by derivative of the phase divided by 2π, i.e., $\omega_0 = 50 \cdot t_0$. Note that in the figure below, f is only defined on $[0 : 2]$ and zero outside this interval $\rightsquigarrow$ frequency band $[-100 : 100]$ and ripples.



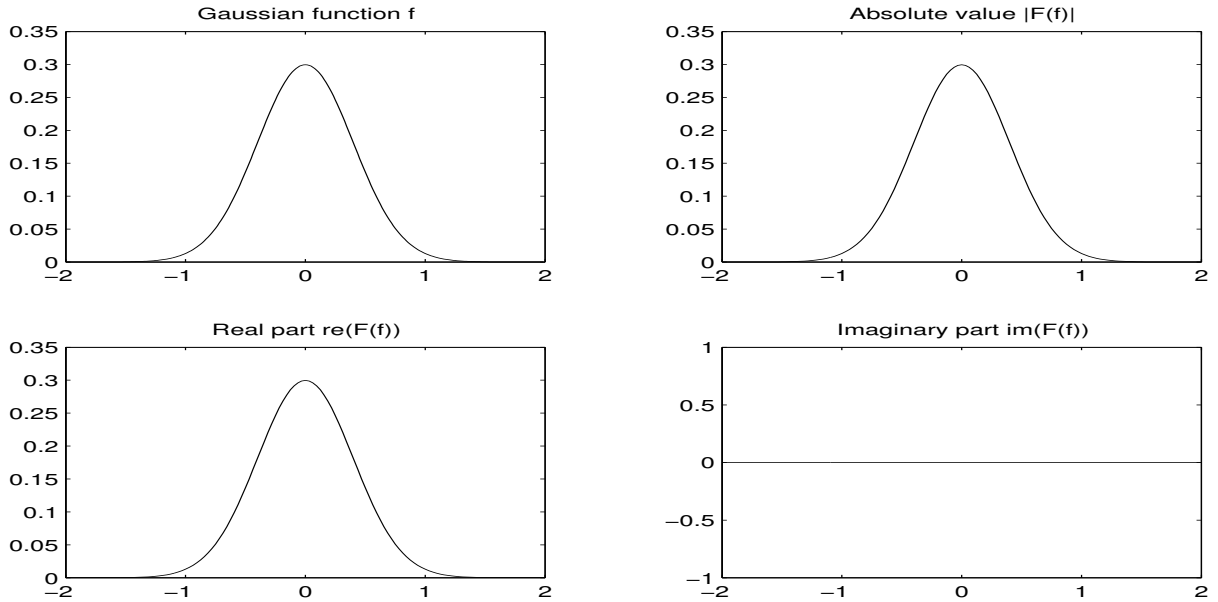
Example 2.14. A Dirac sequence is a sequence of functions $(f_n)_{n\in\mathbb{N}}$ of norm $\|f_n\| = 1$ such that for increasing n the functions f_n "concentrate more and more around the point $t = 0$." The limit of this sequence is the Dirac δ-function from Example 1.3. Intuitively, the δ-function is a superposition of all frequencies.



Example 2.15. The Gaussian function defined by the formula

$$f(t) = (2\pi)^{-\frac{1}{2}} \pi^{-\frac{1}{4}} e^{-\pi t^2}$$

has the remarkable property that it coincides with its Fourier transform. It has the minimal uncertainty in the sense of Heisenberg's uncertainty principle (see Chapter 6) and has good localizing properties in time as well as in frequency.



Example 2.16. The box function $f = b_{1/2} = \chi_{[-1/2,1/2]}$ of length 1 centered at 0 (see Example 1.2) is given by

$$f(t) := \begin{cases} 1 & \text{if } -1/2 \le t \le 1/2, \\ 0 & \text{elsewhere.} \end{cases}$$

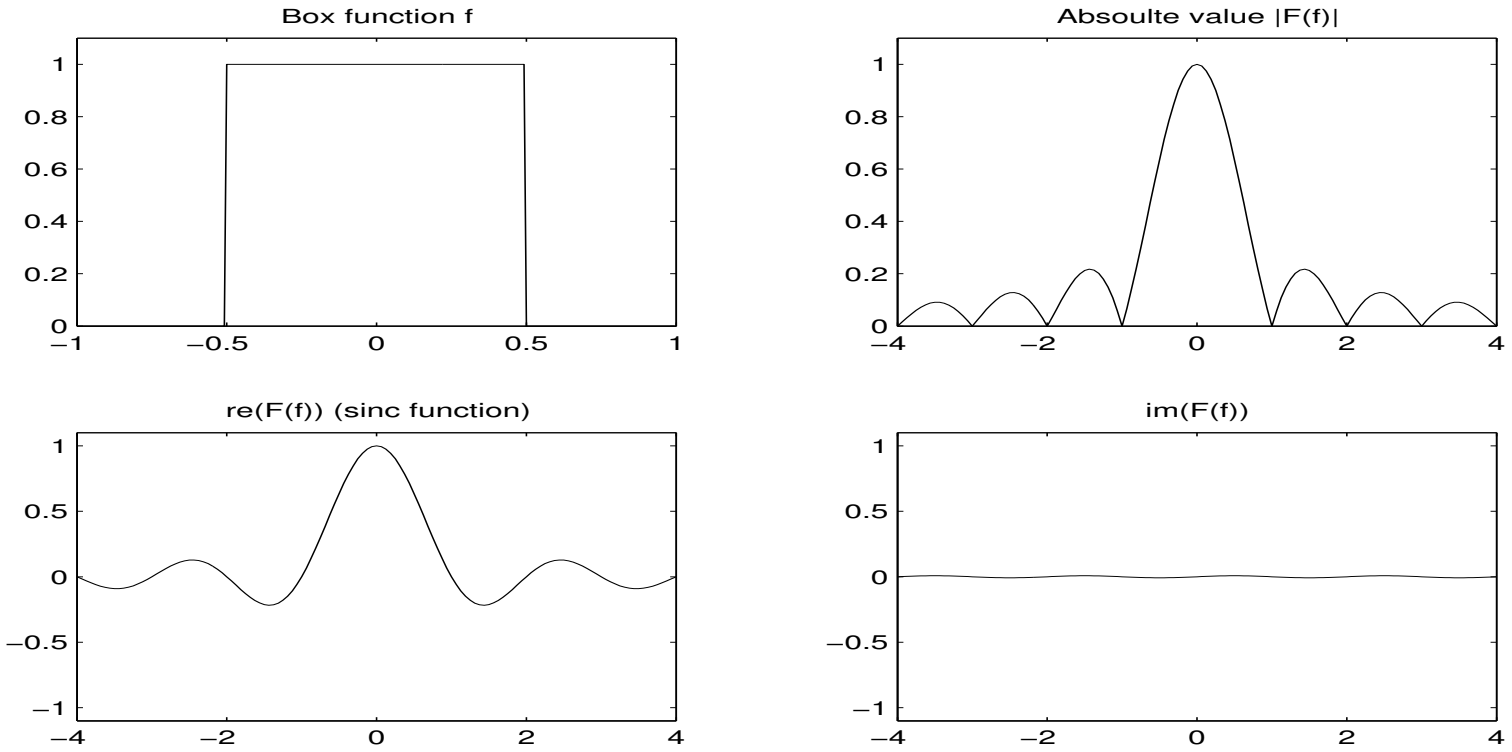
We compute the Fourier transform of f considering first the case for $\omega \neq 0$:

$$\begin{aligned} \hat{f}(\omega) &= \int_{-\infty}^{\infty} f(t) e^{-2\pi i \omega t} dt = \int_{-1/2}^{1/2} e^{-2\pi i \omega t} dt \\ &= \left[\frac{1}{-2\pi i \omega} e^{-2\pi i \omega t} \right]_{-1/2}^{1/2} \\ &= \frac{1}{-2\pi i \omega} \left(e^{-\pi i \omega} - e^{\pi i \omega} \right) = \frac{\sin(\pi \omega)}{\pi \omega}. \end{aligned}$$

For $\omega = 0$ we get

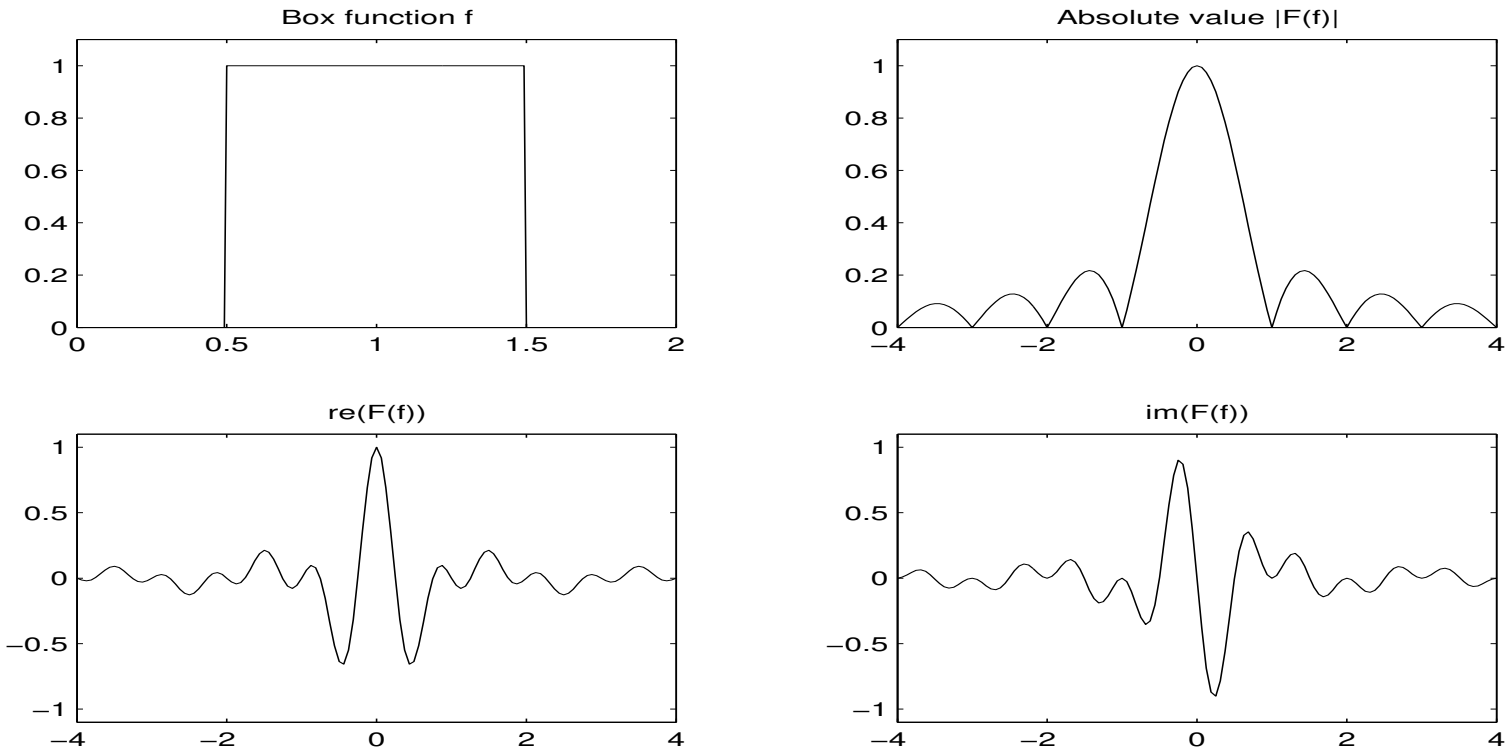
$$\hat{f}(0) = \int_{-\infty}^{\infty} f(t) dt = 1.$$

In other words, the Fourier transform of the box function is the sinc function from Example 1.4. The Fourier transform is in this case a real-valued function, i.e., the imaginary part of $\hat{f}$ is zero as is shown in the following figure.

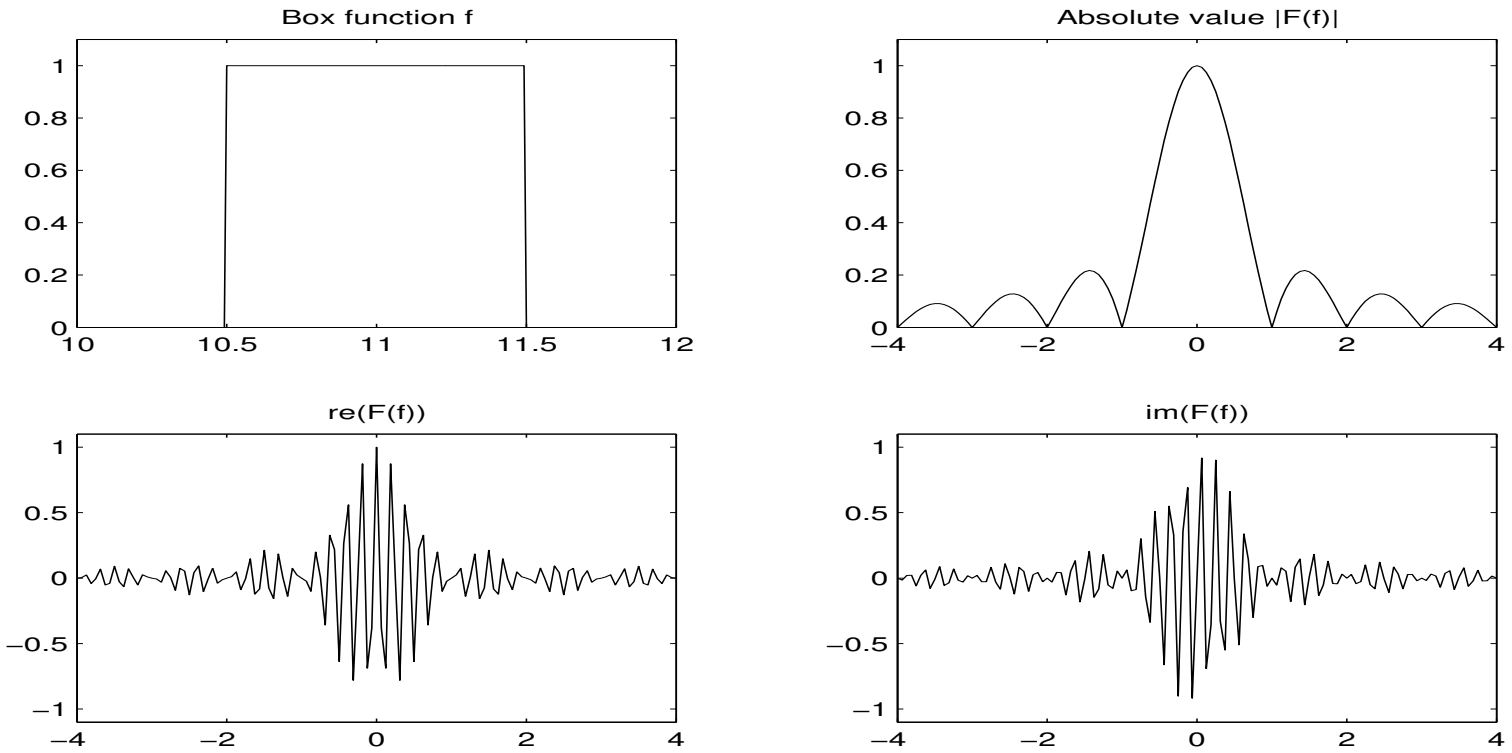


A translation of the signal in the time domain leads to a modulation in the Fourier domain. This is expressed in formula (1) of Theorem 2.9 and illustrated for the box function in the next two figures.

Translation of the box function by 1.



Translation of the box function by 11.



2.3 Fourier Transform for DT-Signals

In this section, we want to transfer the concept of a Fourier transform to the time-discrete case. The following definition is a kind of dual concept to the Fourier series:

Definition 2.17. The discrete-time (DT) Fourier transform $\hat{x}$ of a DT-signal $x \in \ell^2(\mathbb{Z})$ is defined by

$$\hat{x}(\omega) := \sum_{k=-\infty}^{\infty} x(k) e^{-2\pi i k \omega}, \quad \text{for } \omega \in [0,1].$$

Note 2.18. In Section 1 we have seen that for a periodic function $f \in L^2([0,1])$ the Fourier transform $\hat{f} := (\langle f | e^{2\pi i k t} \rangle)_{k \in \mathbb{Z}}$ is a DT-signal $\hat{f} \in \ell^2(\mathbb{Z})$ consisting of the Fourier coefficients. Now, for a DT-signal $x \in \ell^2(\mathbb{Z})$, the Fourier transform is a periodic function $\hat{x} \in L^2([0,1])$.

Note 2.19. Even for a signal $x \in \ell^1(\mathbb{Z})$ the Fourier transform $\hat{x}$ is defined as in Definition 2.17. However, in this case $\hat{x}$ is in general not any longer in $L^2([0,1])$ and the reconstruction of x from $\hat{x}$ becomes more complicate.

Similar to Theorem 2.9 DT-Fourier transform of DT-signals has the following properties.

Theorem 2.20. Let $x, y \in \ell^1(\mathbb{Z})$. Then

(1) Linearity: $\widehat{x+y} = \hat{x} + \hat{y}; \quad \widehat{\lambda x} = \lambda \hat{x}, \lambda \in \mathbb{C}$
(2) Time shift: $\widehat{x^k}(\omega) = e^{-2\pi i k\omega}\hat{x}(\omega)$
(3) Modulation: $\widehat{E_{\omega_0}[x]}(\omega) = \hat{x}(\omega + \omega_0)$
(4) Complex conjugation: $y = \overline{x} \;\Rightarrow\; \hat{y}(\omega) = \overline{\hat{x}(-\omega)}$
(5) Time reversal: $\forall n \in \mathbb{Z}$: $y(n) = x(-n) \Rightarrow \hat{y}(\omega) = \hat{x}(-\omega)$

Here, the modulation operator $E_{\omega_0}\colon \ell^1(\mathbb{Z}) \to \ell^1(\mathbb{Z})$ is defined by $E_{\omega_0}[x](n) := e^{-2\pi i\omega_0 n}x(n)$, $x \in \ell^1(\mathbb{Z})$, $n \in \mathbb{Z}$ (see also Example 3.6).

We compare the CT-Fourier transform and the DT-Fourier transform. Let $f \in L^2(\mathbb{R})$ be a continuous CT-signal and $x \in \ell^2(\mathbb{Z})$ the DT-signal defined by $x(k) := f(k), k \in \mathbb{Z}$. In other words, x is a sampled version of f. By definition we have

$$\hat{f}(\omega) := \int_{-\infty}^{\infty} f(t) e^{-2\pi i \omega t} dt$$

and

$$\hat{x}(\omega) := \sum_{k=-\infty}^{\infty} f(k) e^{-2\pi i \omega k}.$$

Hence, $\hat{x}(\omega)$ is a Riemann sum for $\hat{f}(\omega)$ for each $\omega \in \mathbb{R}$.

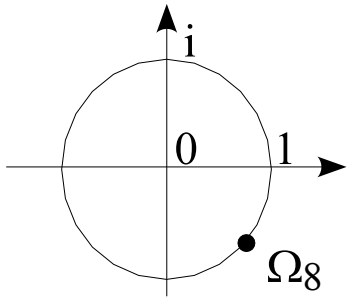
However, note that for increasing $|\omega|$ the functions $t \mapsto f(t)e^{-2\pi i\omega t}, t \in \mathbb{R}$ are increasingly oscillating. However, the sampled versions $k \mapsto f(k)e^{-2\pi i\omega k}, k \in \mathbb{Z}$, are not able to recognize oscillations of frequency greater one since such oscillations take place between two neighboring time points. This effect is known as aliasing and will be discussed in Chapter 4 in detail.

Note 2.21. Fourier series are ideal for analyzing periodic signals, since the harmonic modes $\mathbf{e}_k$, $k \in \mathbb{Z}$, unsed in the expansion are themselves periodic. By contrast, the Fourier intergral transform is a far less natural tool because it uses periodic functions to expand nonperiodic signals.

2.4 Discrete Fourier Transform

Computing the Fourier transform of a CT-signal or Fourier coefficients of a periodic CT-signal involves the evaluation of integrals which is computational infeasible. Also computing approximation of such integrals via Riemann sums can be very expensive.

Therefore one has to find fast algorithms for computing suitable approximations of Fourier coefficients for suitable frequencies — possibly at the expense of precision. In real world applications one usually has to deal with finite DT-signals. Let $N \in \mathbb{N}$ and $\Omega_N := e^{-2\pi i/N}$. ($\Omega_N$ is a so-called Nth primitive root of unity.) The next figure illustrates the case for $N = 8$:



Definition 2.22. The discrete Fourier transform (DFT) of size N is a linear map $\mathbb{C}^N \longrightarrow \mathbb{C}^N$ given by the $N \times N$-matrix

$$\mathrm{DFT}_N := \frac{1}{\sqrt{N}} \left(\Omega_N^{kj} \right)_{0 \le k,j < N}$$

$$= \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & \Omega_N & \cdots & \Omega_N^{(N-2)} & \Omega_N^{(N-1)} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & \Omega_N^{(N-2)} & \ddots & \Omega_N^{(N-2)(N-2)} & \Omega_N^{(N-2)(N-1)} \\ 1 & \Omega_N^{(N-1)} & \cdots & \Omega_N^{(N-1)(N-2)} & \Omega_N^{(N-1)(N-1)} \end{pmatrix}$$

Hence, for a vector (finite DT-signal) $v := (v_0, v_1, \ldots, v_{N-1})^T \in \mathbb{C}^N$ the DFT of v is again a vector $\hat{v} = \mathrm{DFT}_N(v) \in \mathbb{C}^N$ given by

$$\hat{v}_k := \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} v_j e^{-2\pi i j k/N}, \quad k = 0, 1, \ldots, N-1.$$

Note 2.23. The rows of the DFT_N-matrix given by

$$\mathbf{f}_k = \frac{1}{\sqrt{N}}(1, (\Omega_N)^k, \ldots, (\Omega_N)^{k(N-1)})^T \quad \in \mathbb{C}^N, \quad k = 0, 1, \ldots N-1,$$

are truncated versions of the discrete frequency signals from Example 1.9 and form an orthonormal basis of $\mathbb{C}^N$. Hence for the component $\hat{v}_k$ holds

$$\hat{v}_k = \langle v | \overline{\mathbf{f}_k} \rangle.$$

Note that the straightforward computation of the matrix-vector product $\mathrm{DFT}_N(v)$ requires $O(N^2)$ multiplications and additions. This is for most applications to slow — in many cases one has to deal with large $N >> 10^5$.

The important point is that there is an efficient algorithm, the so-called fast Fourier transform (FFT), to compute the DFT of an vector of length N in $O(N \log N)$. We refer for a detailed description of this algorithm to [Clausen/Baum].

The main idea of the FFT-algorithm — originally found by Gauss and rediscovered by Cooley and Tukey in 1965 — is based on a clever matrix factorization.

For $N = 2M$ holds:

$$\mathrm{DFT}_N \cdot \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \end{pmatrix} = \frac{1}{\sqrt{2}} \left(\begin{array}{c|c} \mathrm{id}_M & \Delta_M \\ \hline \mathrm{id}_M & -\Delta_M \end{array}\right) \left(\begin{array}{c|c} \mathrm{DFT}_M & 0 \\ \hline 0 & \mathrm{DFT}_M \end{array}\right) \begin{pmatrix} v_0 \\ v_2 \\ \vdots \\ v_{N-2} \\ \hline v_1 \\ v_3 \\ \vdots \\ v_{N-1} \end{pmatrix}$$

Here

$$\mathrm{id}_M = \mathrm{diag}\,(1, 1, \ldots, 1) \quad \text{and} \quad \Delta_M = \mathrm{diag}\,(1, \Omega_N, \ldots, \Omega_N^{M-1})$$

denote the $M \times M$-identity matrix and an $M \times M$-diagonal matrix, respectively. Furthermore, DFT_M corresponds to the DFT-matrix for $\Omega_M = \Omega_N^2$. If N is a power of two, this procedure can be performed recursively leading to an upper bound of $\frac{3}{2} N \log N$ additions and multiplications.

Next we describe, how the DFT can be used for an approximative computation of Fourier coefficients. For a periodic function $f \in L^2([0,1])$ we have the Fourier series

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi ikt} \quad \text{with} \quad c_k = \int_0^1 f(t) e^{-2\pi ikt} dt.$$

With respect to an equidistant partition of the interval $[0,1]$ into N segments, the integral for c_k is approximated by a Riemann sum. We denote this sum by γ_k for $k \in \mathbb{Z}$ which is given by

$$\gamma_k = \frac{1}{N} \sum_{j=0}^{N-1} f\left(\frac{j}{N}\right) e^{-2\pi ijk/N}.$$

By the identity $e^{-2\pi ijk/N} = e^{-2\pi ij(k+N)/N}$ the map

$$\mathbb{Z} \to \mathbb{C}, \quad k \mapsto \gamma_k$$

is N-periodic (this again is the so-called aliasing effect). Hence, the entire information of the sequence $(\gamma_k)_{k\in\mathbb{Z}}$ is contained in the vector

$$\Gamma := (\gamma_0, \gamma_1, \ldots, \gamma_{N-1})^T.$$

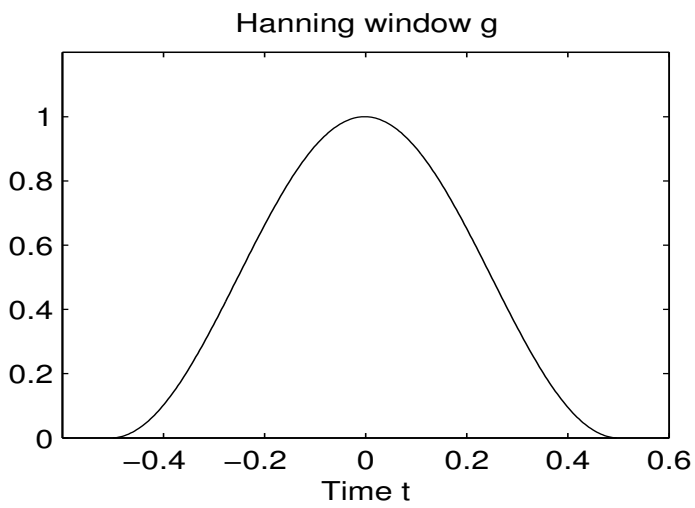
Note that Γ can be computed via the DFT of the vector $v := (v_0, v_1, \ldots, v_{N-1})^T \in \mathbb{C}^N$ were $v_j := \frac{1}{\sqrt{N}} f(\frac{j}{N})$.

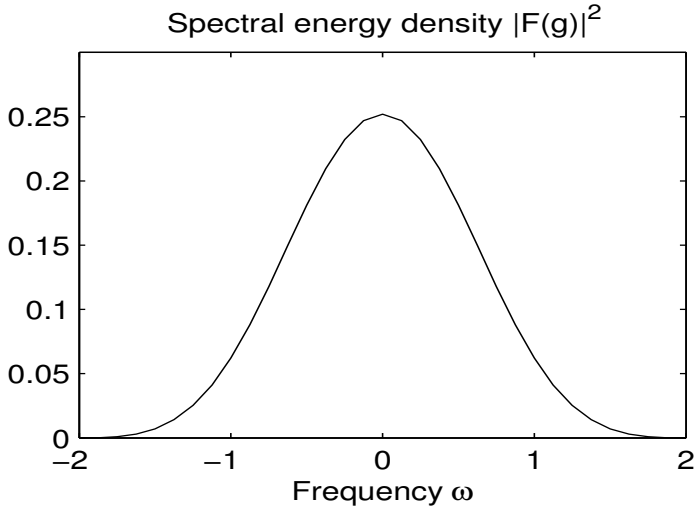
Note 2.24. The DFT computes Riemann approximations of N Fourier coefficients synchronously.

Note 2.25. In general, the quality of the approximation of c_k via γ_k decreases for increasing k. (Consider the number of oscillations of the function to be integrated!) In many cases, only half of the numbers γ_k for $0 \le k < \frac{N}{2}$ give acceptable approximations for c_k.

Example 2.26. In this example we need the Hanning-window g — a so-called window function — which is depicted below and defined by

$$g(u) := \begin{cases} 1 + \cos(\pi u) & \text{for} \quad -0.5 \leq u \leq 0.5 \\ 0 & \text{otherwise} \end{cases}$$





We multiply a suitable chirp function multiplied with a translated Hanning-window to get the CT-signal $f \in L^2([0,1])$ defined by

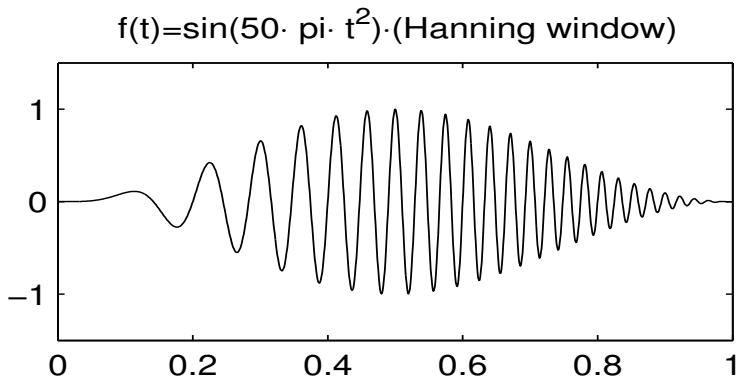
$$f(t) = \sin(50\pi t^2) \cdot g(t - 0.5), \quad t \in [0,1].$$

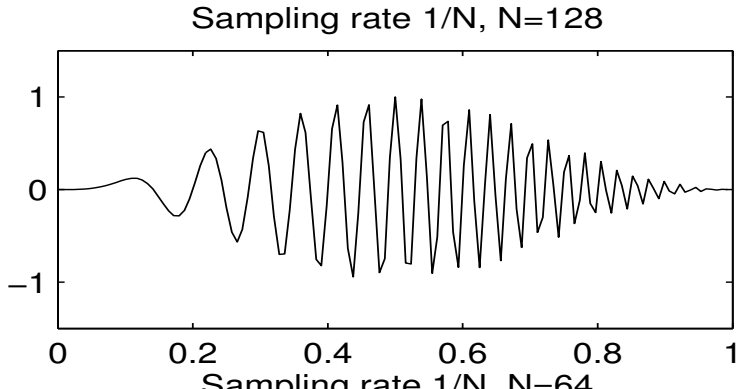
The rows of the next figure have the following meaning.

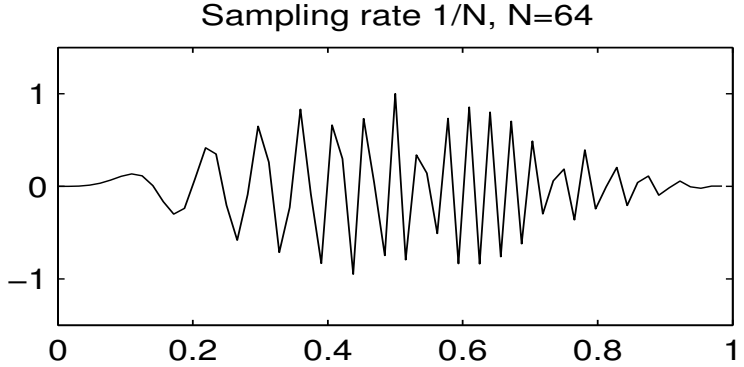
(1) The first row shows the function f and the absolute values $|c_k|$ for the Fourier coefficients c_k of the corresponding Fourier series expansion.

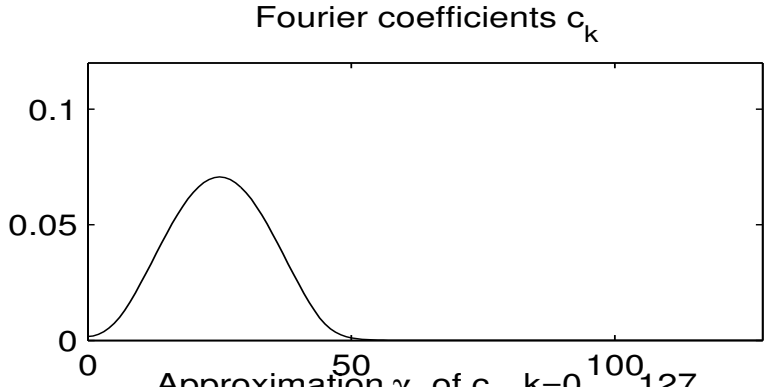
(2) A DFT of length $N = 128$ has been applied to the samples $f(k/128)$ for $k = 0, 1, \ldots, 127$. As explained before, the coefficients $\gamma_k, 0 \leq k \leq 127$, are an approximation of the corresponding c_k. The absolute values $|\gamma_k|$ are shown in the second row.

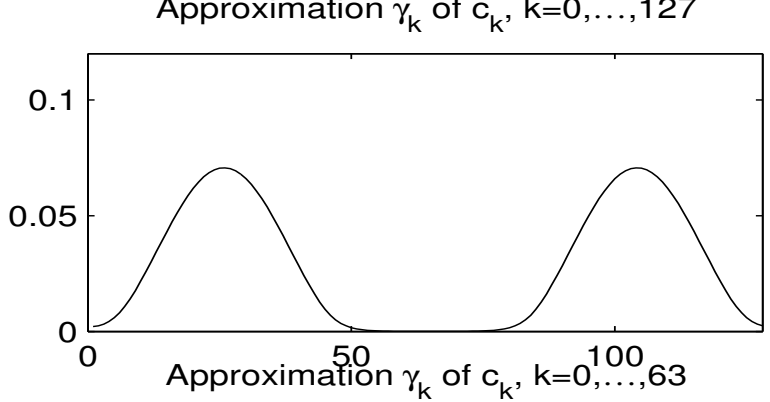
(3) Similar to (2), but now with $N = 64$.

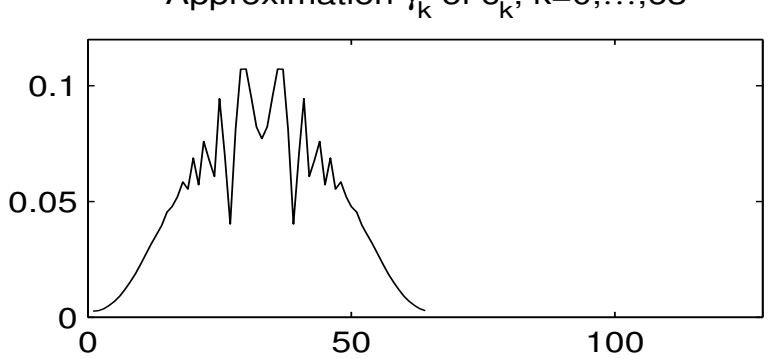

f(t)=sin(50· pi· t^2)·(Hanning window)
1
0
−1
0
0.2
0.4
0.6
0.8
1


Sampling rate 1/N, N=128
1
0
−1
0
0.2
0.4
0.6
0.8
1

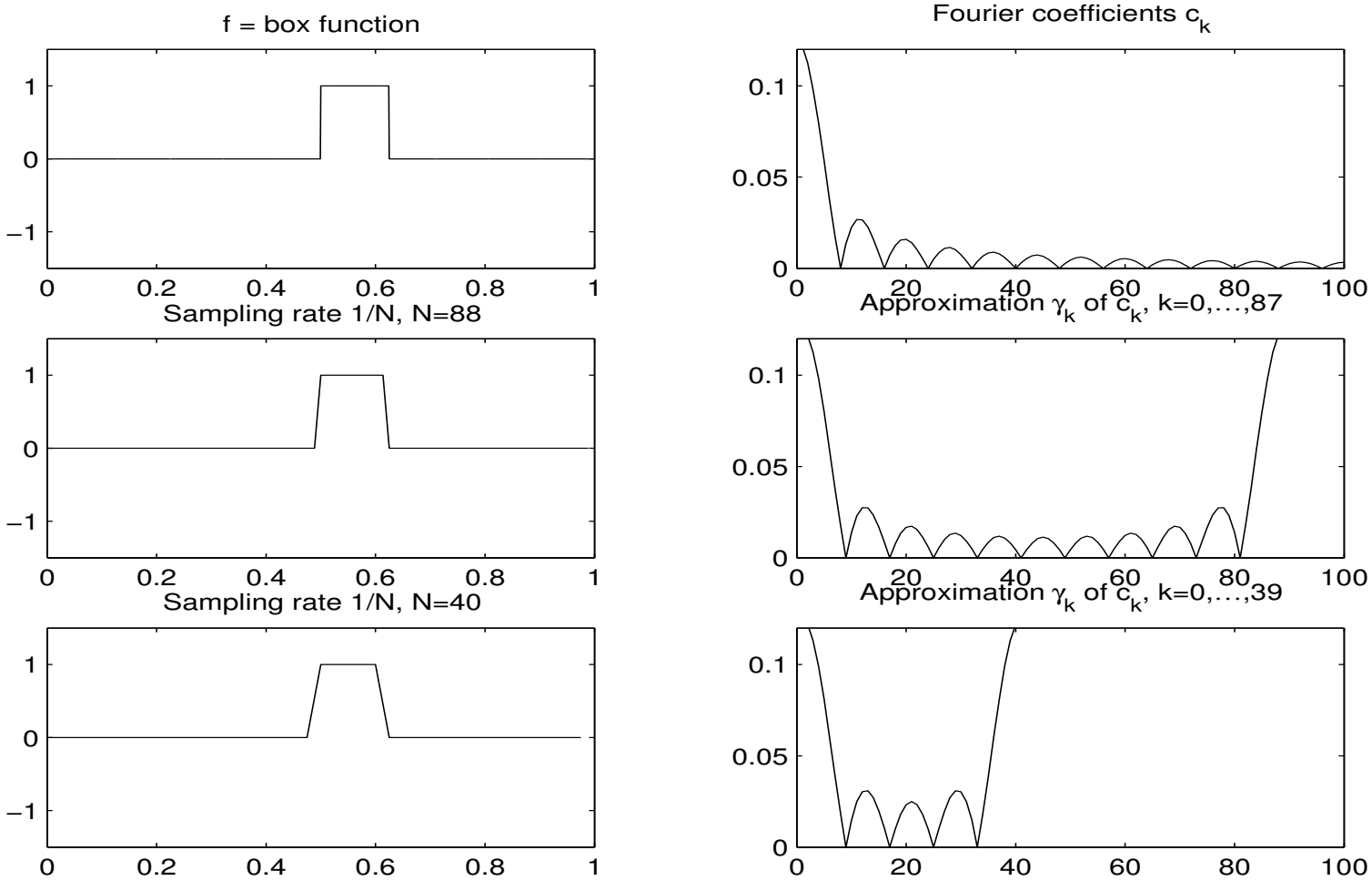

Sampling rate 1/N, N=64
1
0
−1
0
0.2
0.4
0.6
0.8
1


Fourier coefficients c_k
0.1
0.05
0
0
50
100


Approximation γ_k of c_k, k=0,...,127
0.1
0.05
0
0
50
100


Approximation γ_k of c_k, k=0,...,63
0.1
0.05
0
0
50
100

Example 2.27. The same as in Example 2.26 with f being a box function.

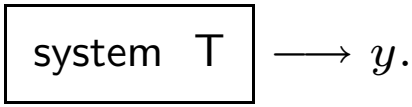


Chapter 3: Systems and Filters

Andrew S. Glassner writes in his book "Principles of Digital Image Synthesis." [Glassner]:

> Anything that alters a signal my be considered a system. For example, a concert hall may be considered a system. In this case, think of the sound of a violin as a signal represented by the amplitude of sound with respect to time. So a concert hall changes an input signal (a violin played on stage) to an output signal (the particular sound you hear at some particular seat).

Mathematically, a system $T : I \longrightarrow O$ transforms an input signal $x \in I$ into an output signal $y \in O$. Here I and O denote suitable signal spaces.



3.1 Linear Filter and LTI-Systems

We just consider the discrete-time case in detail and refer for a summary of the CT-case to Section 3.5. Mainly we are interested in the case $I = \ell^p(\mathbb{Z}) = O$, $1 \leq p \leq \infty$. The easiest class of systems between such spaces are linear systems.

Definition 3.1. Let I and O linear signal spaces. A linear map $T\colon I \to O$ is called a linear system. One has

$$T[x+y] = T[x] + T[y] \quad \text{and} \quad T[\lambda x] = \lambda T[x],$$

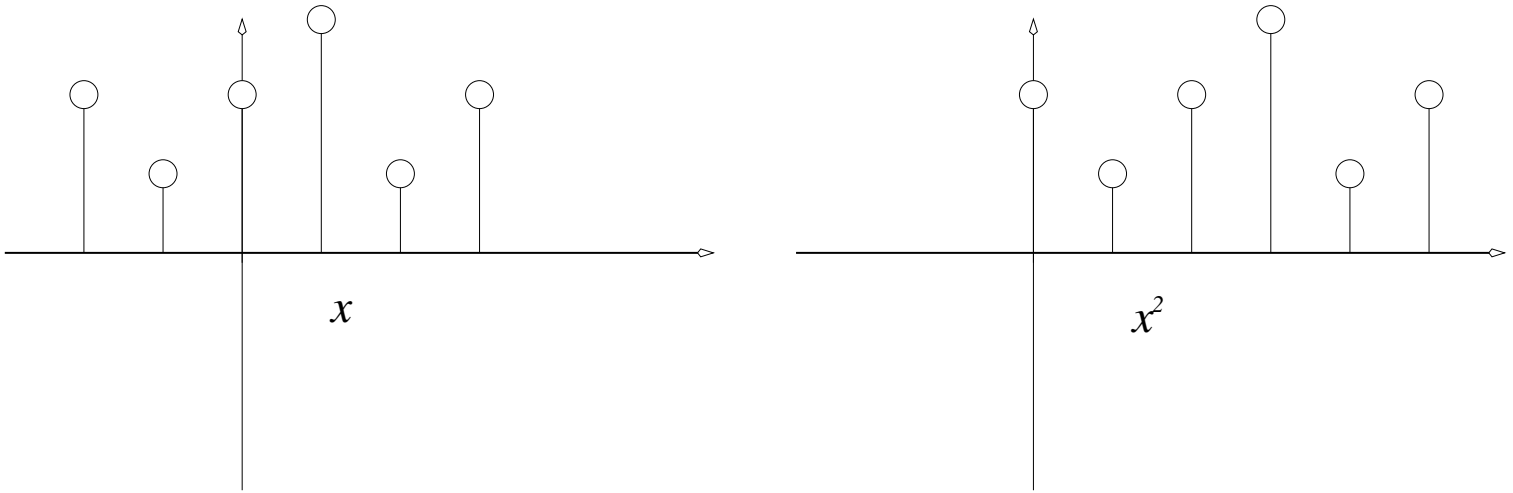
for all $x, y \in I$ and all $\lambda \in \mathbb{C}$.

Note 3.2. Note that T maps a signal to another signal, whereas a signal itself maps an element (the time point) into another element (the value of amplitude). In other words, T is a function between function spaces and is often referred to as an operator. This is also expressed in using different parenthesis: one often writes $T[x]$ instead of $T(x)$. Then $T[x](n)$ denotes the value of the output signal $T[x]$ at time n.

Example 3.3. The time shift by $k \in \mathbb{Z}$ is defined by

$$\tau_k[x](n) := x(n-k).$$

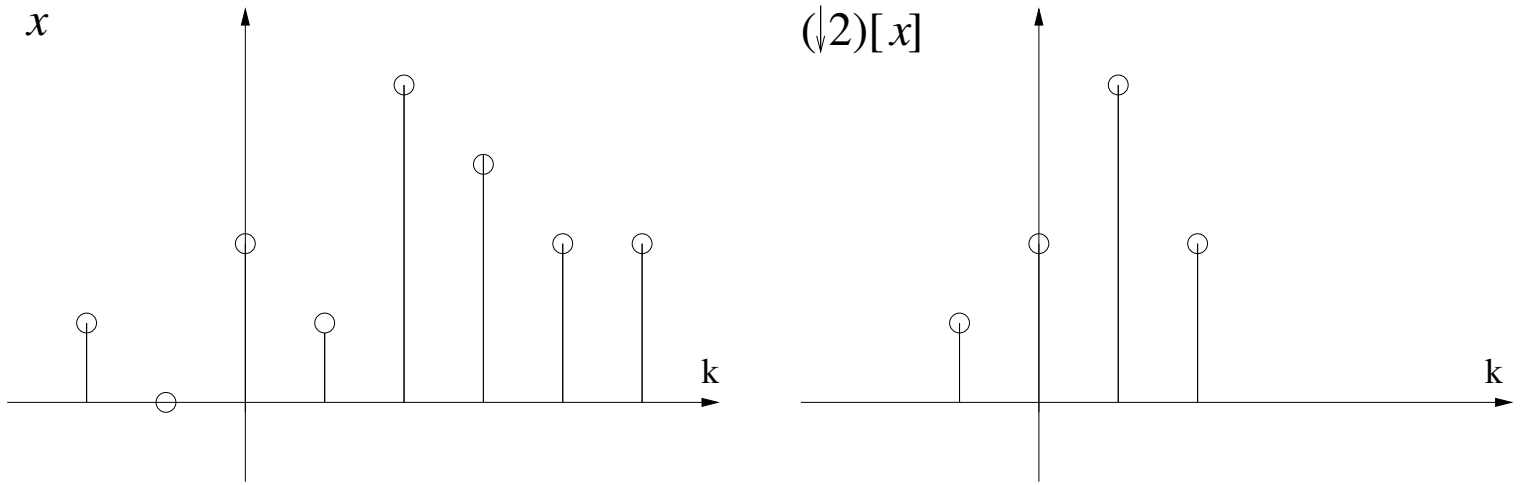
It is easy to see that τ_k is a linear operator from $\ell^p(\mathbb{Z})$ to $\ell^p(\mathbb{Z})$. Sometimes we also write write x^k for $\tau_k[x]$. In particular, $x^0 = x$ and δ^k is the indicator function for $k \in \mathbb{Z}$, i.e., $\delta^k(j) = \delta_{kj}$ for $j \in \mathbb{Z}$.



Example 3.4. The M-downsampler for some $M \in \mathbb{N}$ is defined by

$$(\downarrow M)[x](n) := x(M \cdot n).$$

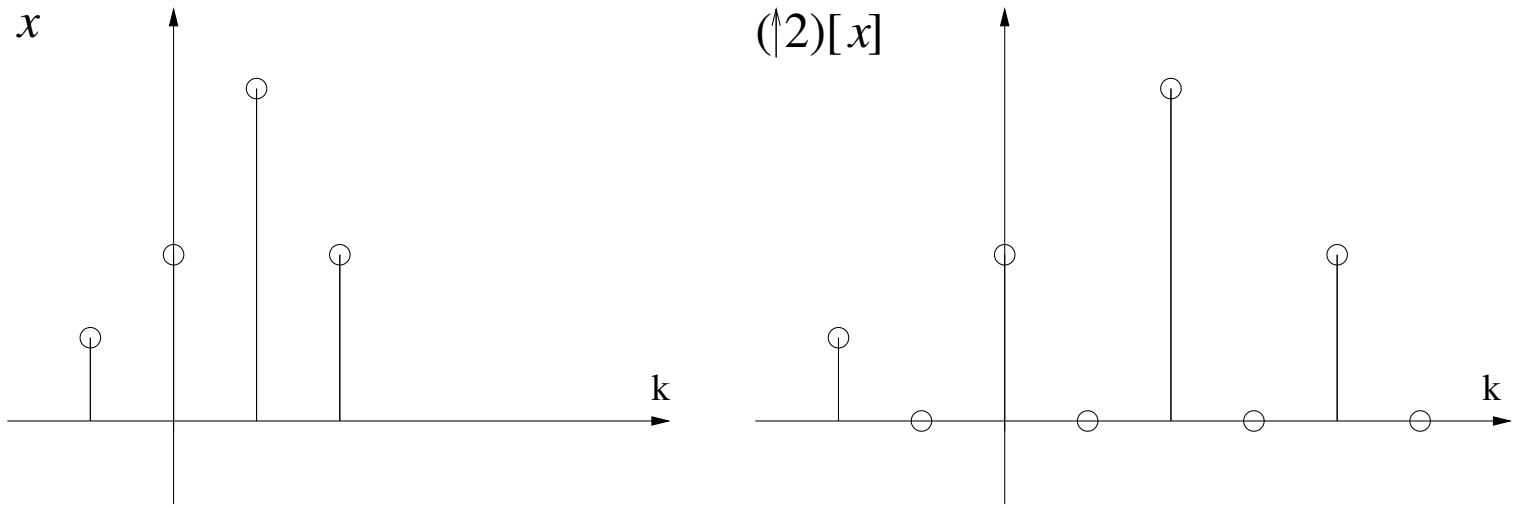
This linear operator from $\ell^p(\mathbb{Z})$ to $\ell^p(\mathbb{Z})$ takes only every Mth value of the signal x.



Example 3.5. Der M-upsampler for some $M \in \mathbb{N}$ is defined by

$$(\uparrow M)[x](n) = \begin{cases} x(n/M), & \text{if } M|n, \\ 0, & \text{otherwise.} \end{cases}$$

This linear operator from $\ell^p(\mathbb{Z})$ to $\ell^p(\mathbb{Z})$ widens the signal x and inserts $M-1$ additional time points with value 0 between any two neighboring time points of x.



Example 3.6. The amplitude modulation with a support signal $c \in \ell^\infty(\mathbb{Z})$ is defined by

$$m_c[x](n) := c(n) \cdot x(n).$$

An important special case is the frequency shift operator with respect to some $\omega \in (0, 1)$ defined by

$$E_\omega[x](n) := e^{-2\pi i \omega n} x(n).$$

In case of the signal $x = f_{\omega_0} = \left(e^{2\pi i \omega_0 n}\right)_{n\in\mathbb{Z}} \in \ell^\infty(\mathbb{Z})$ we get

$$E_\omega[f_{\omega_0}] = f_{\omega_0 - \omega}.$$

With respect to composition the operators E_ω und $E_{-\omega}$ are inverse to each other:

$$(E_{-\omega}) \circ E_\omega[x] = x.$$

The operator $E_{-\omega}$ is then called demodulation. By using modulation and subsequent demodulation, signals can be transmitted using high-frequency support signals.

We finally give two examples of systems which are not linear.

Example 3.7. The quantization operators given by

$$\text{rounding up:} \qquad \lceil x \rceil(n) := \lceil x(n) \rceil, \quad \text{for } x\colon \mathbb{Z} \to \mathbb{R}$$

$$\text{rounding down:} \qquad \lfloor x \rfloor(n) := \lfloor x(n) \rfloor, \quad \text{for } x\colon \mathbb{Z} \to \mathbb{R}$$

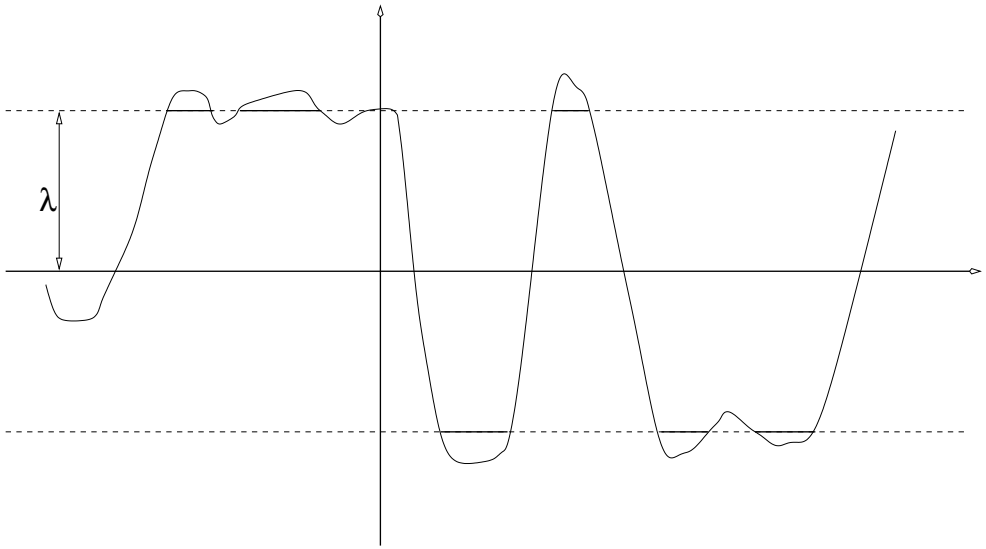
are not linear. However these operators are time invariant:

$$\lceil x^k \rceil = \lceil x \rceil^k, \quad \lfloor x^k \rfloor = \lfloor x \rfloor^k$$

Example 3.8. The cut-off at $\lambda > 0$ is defined by

$$\mathrm{Cut}_\lambda[x](n) := \begin{cases} x(n) & \text{if} \quad |x(n)| \leq \lambda \\ \frac{x(n)}{|x(n)|}\lambda & \text{else} \end{cases}$$

for $x : \mathbb{Z} \longrightarrow \mathbb{C}$. Cut_λ is not linear but time invariant. As an illustration the following figure shows the cut-off system for some CT-signal $x\colon \mathbb{R} \longrightarrow \mathbb{R}$:



We are interested in linear systems, which do not behave in a pathological way. One such important property is continuity.

Definition 3.9. A linear system $T\colon \ell^p(\mathbb{Z}) \to \ell^r(\mathbb{Z})$ is called continuous, if it maps all convergent sequences of input signals in $\ell^p(\mathbb{Z})$ to convergent sequences of output signals in $\ell^r(\mathbb{Z})$.

The following theorem describes a property of continuous operators which is often taken for granted:

Theorem 3.10. *A continuous linear system $T\colon \ell^p(\mathbb{Z}) \to \ell^r(\mathbb{Z})$ commutes with infinite sums. In particular, for $x \in \ell^p(\mathbb{Z})$ holds*

$$T[\sum_{k\in\mathbb{Z}} x(k)\delta^k] = \sum_{k\in\mathbb{Z}} x(k)T[\delta^k].$$

Note that the opposite statement also holds under certain additional conditions.

Proof: For $n \in \mathbb{N}$ let $\sigma_n := \sum_{|k|\leq n} x(k)\delta^k$. Then σ_n converges to x in the $\ell^p(\mathbb{Z})$-norm. Since T is continuous, $T[\sigma_n]$ converges to $T[x]$ in the $\ell^r(\mathbb{Z})$-norm, i.e.,

$$\left\| T[x] - \sum_{|k|\leq n} x(k)T[\delta^k] \right\|_r \to 0$$

for $n \to \infty$. Hence

$$T[x] = \lim_{n\to\infty} \sum_{|k|\leq n} x(k)T[\delta^k] = \sum_{k\in\mathbb{Z}} x(k)T[\delta^k].$$

□

Very important are systems which commute with time shifts.

Definition 3.11. A linear system $T: \ell^p(\mathbb{Z}) \to \ell^r(\mathbb{Z})$ is called a time invariant if $T \circ \tau_k = \tau_k \circ T$ for $k \in \mathbb{Z}$. In other words, for all $k \in \mathbb{Z}$ and all $x \in \ell^p(\mathbb{Z})$ holds

$$T[x^k] = T[x]^k.$$

A linear time invariant system is called linear time invariant (LTI) system.

LTI systems can be easily described and characterized by the so-called convolution operator. This will be the content of the next section.

3.2 Convolution Filter

The convolution of two signals is a kind of multiplication leading again to a signal. Convolution plays a crucial role in describing filters and hence is an indispensable mathematical tool in digital signal processing.

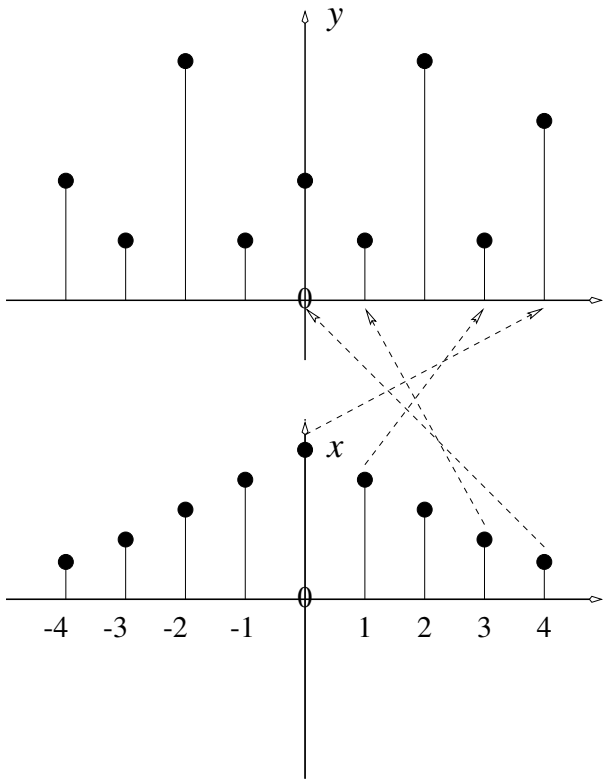
Definition 3.12. Let $x, y\colon \mathbb{Z} \longrightarrow \mathbb{C}$ be signals, then the convolution of x and y at position $n \in \mathbb{Z}$ is defined to be

$$(x * y)(n) := \sum_{k\in\mathbb{Z}} x(k)y(n-k).$$

Attention: $x * y$ exists only under suitable conditions on x and y, e.g., $x \in \ell^1(\mathbb{Z}), y \in \ell^\infty(\mathbb{Z})$ or $x, y \in \ell^2(\mathbb{Z})$. Further conditions will be summarized in Theorem 3.13.

Intuitively, if $y\colon \mathbb{Z} \to \mathbb{C}$ is a signal and x a probability distribution on $\mathbb{Z}$, i.e., $x\colon \mathbb{Z} \to [0, 1]$ with $\sum_{n\in\mathbb{Z}} x(n) = 1$, then $(x * y)(n) = \sum_{k\in\mathbb{Z}} x(k)y(n-k)$ can be thought of as weighted average of y around the neighborhood of n.

The following figure illustrates this for the position $n = 4$.



In the following theorem we summarize some important properties of the convolution and mulitplication operator. For a proof we refer to [Folland].

Theorem 3.13. *Let $1 \leq p, q, r \leq \infty$. Then the following holds.*

(1) *(Young Inequality) $\ell^1 * \ell^p \subseteq \ell^p$, i.e., for all $x \in \ell^1(\mathbb{Z})$ and $y \in \ell^p(\mathbb{Z})$ holds $x * y \in \ell^p(\mathbb{Z})$ and*

$$\|x * y\|_p \leq \|x\|_1 \cdot \|y\|_p.$$

(2) *Let p and q be conjugate exponents (i.e., $\frac{1}{p} + \frac{1}{q} = 1$ where one sets $\frac{1}{\infty} := 0$)). Then for all $x \in \ell^p$ and $y \in \ell^q$ holds*

$$x \cdot y \in \ell^1 \quad \text{and} \quad \|x \cdot y\|_1 \leq \|x\|_p \cdot \|y\|_q$$
$$x * y \in \ell^\infty \quad \text{and} \quad \|x * y\|_\infty \leq \|x\|_p \cdot \|y\|_q.$$

(3) *For all $x \in \ell^p(\mathbb{Z})$ and $y \in \ell^\infty(\mathbb{Z})$ one has $x \cdot y$ in $\ell^p(\mathbb{Z})$ and*

$$\|x \cdot y\|_p \leq \|x\|_p \cdot \|y\|_\infty$$

Theorem 3.14. *Let $x, y, z \in \mathbb{C}^{\mathbb{Z}}$ and suppose all of the following convolutions and products in question exist. The pointwise multiplication and convolution of signals are commutative and associative, i.e.,*

$$x \cdot y = y \cdot x, \quad x * y = y * x,$$

and

$$(x \cdot y) \cdot z = x \cdot (y \cdot z), \quad (x * y) * z = x * (y * z).$$

Furthermore, in combination with addition the respective laws of distributivity hold:

$$(x + y) \cdot z = x \cdot z + y \cdot z, \quad (x + y) * z = x * z + y * z.$$

From this theorem follows that for fixed $y \in \ell^q(\mathbb{Z})$ the convolution operator C_y defined by $C_y(x) := x * y$ is linear. From Theorem 3.13 follows that $C_y \colon \ell^p(\mathbb{Z}) \to \ell^\infty(\mathbb{Z})$ is continuous.

The convolution operator has a the following remarkable behaviour under Fourier transform.

Theorem 3.15. *Let $x, y \in \ell^2(\mathbb{Z})$ with $x * y \in \ell^2(\mathbb{Z})$. Then*

$$\widehat{x * y} = \hat{x} \cdot \hat{y}.$$

In other words, the Fourier transform of the convolution of two signals equals the pointwise multiplication of the Fourier transforms of the signals.

Note 3.16. *The equaltiy in 3.15 holds only in the $L^2([0,1])$-sense. This implies that $\widehat{x * y}(\omega) = \hat{x}(\omega)\hat{y}(\omega)$ for almost all $\omega \in [0,1]$.*

Example 3.17. For $x \in \ell^\infty(\mathbb{Z})$ holds

$$\tau_k[x] = x^k = x * \delta^k.$$

To show this, we first look at the left hand side of the equality:

$$\tau_k[x](n) = x(n-k).$$

For the right hand side one has

$$\begin{aligned}(x * \delta^k)(n) &= \sum_{\ell \in \mathbb{Z}} x(\ell)\delta^k(n-\ell) \\ &= \sum_{\ell \in \mathbb{Z}} x(\ell)\delta_{\ell=n-k} \\ &= x(n-k)\end{aligned}$$

In other words, the time shift operator τ_k coincides with the convolution operator C_{δ^k} on $\ell^\infty(\mathbb{Z})$.

Generalizing the previous example one gets the following theorem:

Theorem 3.18. *Let $1 \leq p, q \leq \infty$ and $T\colon \ell^p(\mathbb{Z}) \to \ell^q(\mathbb{Z})$ a continuous LTI system. Define $h := T[\delta]$, then $T = C_h$, i.e., for all $x \in \ell^p(\mathbb{Z})$ holds*

$$T[x] = h * x.$$

Proof: Since $\delta \in \ell^p(\mathbb{Z})$ the sequence $h := T[\delta]\ell^q(\mathbb{Z})$ is well defined. Furthermore, for all $n \in \mathbb{Z}$ and $x \in \ell^p(\mathbb{Z})$ holds

$$\begin{aligned}
T[x](n) &= T[\sum_k x(k)\delta^k](n) \\
&= \sum_k x(k)T[\delta^k](n) \quad \text{(linearity and continuity of } T\text{)} \\
&= \sum_k x(k)T[\delta]^k(n) \quad \text{(time invariance of } T\text{)} \\
&= \sum_k x(k)h(n-k) = (h * x)(n).
\end{aligned}$$

The previous theorem showed that continuous LTI systems can be expressed by a convolution operator. If we impose further conditions on the LTI systems they are even characterized by convolution.

Definition 3.19. A linear system $T\colon \ell^p(\mathbb{Z}) \longrightarrow \ell^p(\mathbb{Z})$ is called stable if

(1) T is continuous and
(2) $\forall k \in \mathbb{Z}\colon T[\delta^k] \in \ell^1(\mathbb{Z})$

In the case $p = \infty$ one also speaks from BIBO-stable (Bounded Input $\longrightarrow$ Bounded Output) linear systems.

Theorem 3.20. *For a linear system* $T\colon \ell^p(\mathbb{Z}) \to \ell^p(\mathbb{Z})$ *the following is equivalent:*

(1) T *is a stable LTI system.*

(2) *There is an* $h \in \ell^1(\mathbb{Z})$ *with* $T = C_h$, *i.e.,* $T[x] = h * x$, *for all* $x \in \ell^p(\mathbb{Z})$.

Proof: $(1) \Rightarrow (2)$: See the proof of Theorem 3.18.

$(2) \Rightarrow (1)$: Linearity is clear. Time invariance follows from the associativity of convolution:

$$T[x^k] = h * (x^k) = h * (x * \delta^k) = (h * x) * \delta^k = T[x]^k.$$

Continuity follows from Theorem 3.13 which implies

$$\|h * (x - x_m)\|_p \leq \|h\|_1 \cdot \|x - x_m\|_p \to 0 \quad \text{for } m \to \infty$$

for $x_m \to x$ in $\ell^p(\mathbb{Z})$. □

Definition 3.21. Let T be a continuous LTI-System and $h := T[\delta]$.

(1) The sequence h is called the impulse response of the system and $h(n)$ is called the nth filter coefficient.

(2) T is called FIR filter or FIR system (Finite Impulse Response) if only a finite number of filter coefficients are non zero. Otherwise T is called IIR filter or IIR system (Infinite Impulse Response).

Note 3.22. Very often one identifies the filter T with the impulse response h. Therefore, one often simply speaks of the filter h meaning the underlying convolution filter C_h.

Definition 3.23. The length $\ell(x)$ of non-zero DT-signal $x \in \mathbb{C}^{\mathbb{Z}}$ (i.e., $x_n = 0$ for all $n \in \mathbb{Z}$ but a finite number) is defined by

$$\ell(x) := 1 + \max\{n | x(n) \neq 0\} - \min\{n | x(n) \neq 0\}.$$

In other words, if $a \in \mathbb{Z}$ is the smallest index with $x(a) \neq 0$ and $b \in \mathbb{Z}$ the largest index with $x(b) \neq 0$, then $\ell(x) := b - a + 1$.

If $h \neq 0$ is the impulse response of some FIR filter, then $\ell(h)$ is also called the length of the FIR filter.

Lemma 3.24. *The length of the convolution of two finite sequences x and y is given by the formula*

$$\ell(x * y) = \ell(x) + \ell(y) - 1.$$

Proof: The proof is left as an exercise. □

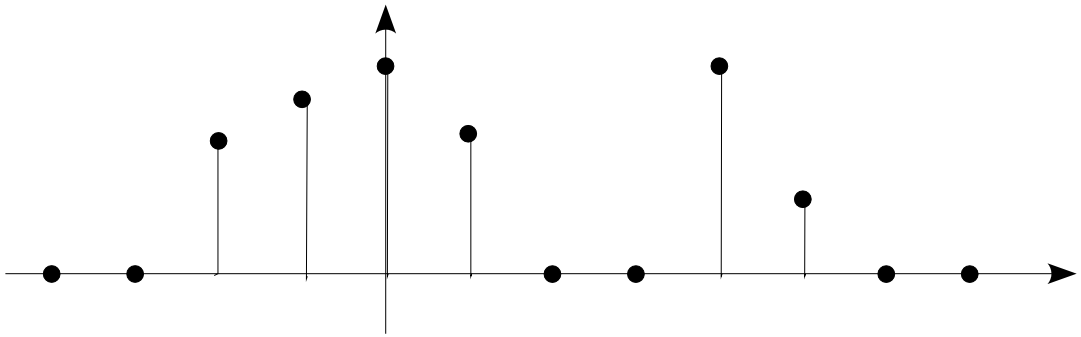


Figure 6: Example for the impulse response of an FIR filter of length 8

Definition 3.25. Let T be a continuous LTI-System and $h := T[\delta]$. T is called causal if $h(n) = 0$ for $n < 0$.

In [Proakis/Manolakis, p. 69] the importance of causality is explained as follows:

> It is apparent that in real-time signal processing applications we cannot observe the future values of the signal, an hence a noncausal system is physically unrealizable (i.e., it cannot be implemented). On the other hand, if the signal is recorded so that the processing is done off-line (nonreal time), it is possible to implement a noncausal system, since all values of the signal are available at the time of processing. This is often the case in the processing of geophysical signals and images.

Example 3.26. A causal FIR filter T of order N and length $N+1$ is of the form

$$T[x](n) = \sum_{\ell=0}^{N} h(\ell)x(n-\ell)$$

with filter coefficients $h(0), \ldots, h(N)$, $h(N) \neq 0$, and $h(0) \neq 0$. The output signal $T[x]$ depends at time point n only on the "past" $x(n-1), \ldots, x(n-N)$ and the "present" $x(n)$ of the input signal x. (Therefore one speaks of causality.) These values are weighted with the filter coefficients and added up.

Example 3.27. The time shifts satisfy $\tau_k \circ \tau_\ell = \tau_{k+\ell} = \tau_\ell \circ \tau_k$ and are therefore time invariant operators. Since $\tau_k[\delta] = \delta^k \in \ell^1(\mathbb{Z})$, the time shifts are stable LTI systems and coincide with the convolution operator: $\tau_k = C_{\delta^k}$. In particular, τ_k is an FIR system and for $k \geq 0$ it is causal.

Example 3.28. The downsampler $(\downarrow M)$ is linear and continuous. Note that $(\downarrow M)[\delta^k]$ is zero in the case that k is a multiple of M, and otherwise it is $\delta^{k/M}$. Hence the downsampler is a stable system. However, it is not time invariant since

$$((\downarrow M) \circ \tau_k)[x](n) = x(nM - k) \quad \text{but} \quad (\tau_k \circ (\downarrow M))[x](n) = x(n(M - k)).$$

Similarly, the upsampler $(\uparrow M)$ is also linear, continuous, and stable, but not time invariant.

Example 3.29. The frequency shift operator E_ω defined by $E_\omega[x](n) := e^{-2\pi i\omega n} \cdot x(n)$ is linear and continous. However, in general it is not time invariant. To be more specific, it holds

$$(E_\omega \circ \tau_k)[x](n) = E_\omega[x^k](n) = e^{-2\pi i\omega n} x(n-k)$$

which is in general not equal to

$$(\tau_k \circ E_\omega)[x](n) = E_\omega[x](n-k) = e^{-2\pi i\omega(n-k)} x(n-k).$$

E_ω commutes with all τ_k, $k \in \mathbb{Z}$, if and only if $e^{2\pi i\omega k} = 1$ for all $k \in \mathbb{Z}$. This only holds for the case $\omega = 0$.

Further examples will be given in Chapter 5.

3.3 Frequency Response

To characterize certain properties of an LTI system the Fourier transform of the corresponding impulse response plays — as we will see later — a crucial role.

Definition 3.30. Let T be a BIBO-stable LTI system with impulse response $h = T[\delta] \in \ell^1(\mathbb{Z})$. Then the Fourier transform (see Definition 2.17)

$$\hat{h}(\omega) := \sum_{n=-\infty}^{\infty} h(n) e^{-2\pi i n \omega}, \quad \omega \in [0, 1],$$

is called frequency response of T.

Note 3.31. In the literature one can find the following conventions which we will also use in the rest of this lecture.

(i) Identifying the system T with its impulse response h one often speaks from the frequence response $\hat{h}$ of the filter h. One then also writes H instead of $\hat{h}$.

(2) More general, one often uses small letters $f, g, h, \ldots, x, y \ldots$ to denote the discrete filters and DT-signals. The corresponding Fourier transforms are then denoted by the corresponding capital letters $F, G, H, \ldots, X, Y \ldots$.

Theorem 3.32. *Let $T = C_h$, i.e., $T[x] = h * x$, be a BIBO-stable LTI system. Then the frequency response $H(\omega) := \sum_{k\in\mathbb{Z}} h(k)e^{-2\pi i\omega k}$ at $\omega \in [0,1]$ is an eigenvalue of T. The frequency sequence $f_\omega := (e^{2\pi i\omega n})_{n\in\mathbb{Z}} \in l^\infty(\mathbb{Z})$ of frequency ω is an eigenvector to this eigenvalue, i.e.,*

$$T[f_\omega] = H(\omega) f_\omega.$$

The proof is left as an exercise which amounts to a straightforward computation. We have seen that a BIBO-stable LTI system maps a (normalized) frequency series of frequency ω into the $H(\omega)$-multiple of the same frequency series.

Definition 3.33. Let $\hat{h}(\omega) = H(\omega) = \sum_{n\in\mathbb{Z}} h(n)e^{-2\pi i\omega n}$ be the frequency response of some filter $h \in \ell(\mathbb{Z})$. Then the representation of H in polar coordinates is given by

$$H(\omega) = |H(\omega)| \cdot e^{2\pi i\Phi_h(\omega)}.$$

The function

$$\omega \mapsto |H(\omega)|$$

is called the magnitude response of h , whereas the function

$$\omega \mapsto \Phi_h(\omega)$$

is called phase response of h. Furthermore, the spectral energy density is defined to be the function

$$\omega \mapsto |H(\omega)|^2.$$

In digital signal processing it is usual to see the magnitude response displayed in decibels (dB). The decibel function $\mathrm{dB}\colon \mathbb{R}_{>0} \to \mathbb{R}$ is defined by

$$\mathrm{dB}(r) := 20 \log_{10}(r)$$

for positive real numbers r. Note that dB is monotonously decreasing, $\mathrm{dB}(1) = 0$ and

$$\lim_{r \to 0} \mathrm{dB}(r) = -\infty.$$

Hence the decibel function is suitable to study the behavior of positive functions close to 0. To get a feeling for decibels, observe that

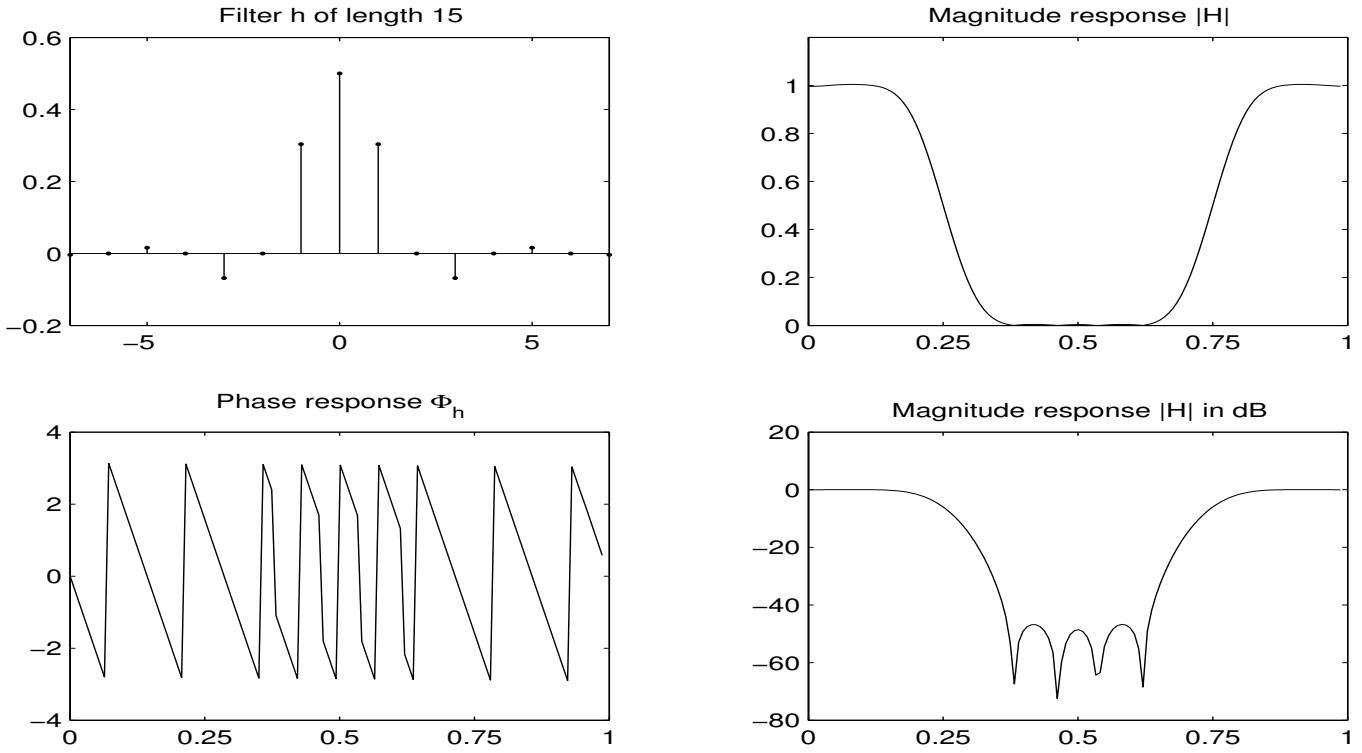
$$\mathrm{dB}(2r) \approx 6 + \mathrm{dB}(r),$$

i.e., doubling of the input corresponds to an increase of about 6 decibels.

The magnitude response measured in dB is given by

$$\omega \mapsto 20 \log_{10} |H(e^{i\omega})|.$$

Example 3.34. The following figure shows the magnitude response $|H|$ and the phase response Φ_h of a filter h of length 15. The non-zero filter coefficients of h are $h(-7) = -0.0036 = h(7), h(-5) = 0.0161 = h(5), h(-3) = -0.0682 = h(3), h(-1) = 0.3038 = h(1)$, and $h(0) = 0.5$. The characteristics of $|H|$ can be better seen in the decibel-scale.



Note 3.35. The frequency response H of a filter h exhibits the so-called spectral information of h. It has the following properties.

(1) For a real-valued filter h one has $\overline{H(\omega)} = H(-\omega)$. Therefore the magnitude response is an even function, i.e., $|H(\omega)| = |H(-\omega)|$, whereas the phase response is an odd function, i.e., $\Phi_h(-\omega) = -\Phi_h(\omega)$.

(2) The frequency response is a 1-periodic function. Since for real-valued filters h the frequency response satisfies $H = \overline{H}$, all information of H is already given by the interval $[0, \frac{1}{2}]$.

(3) Spectral information corresponding to low frequencies correspond to $\omega \in [0, \frac{1}{2}]$ around 0 whereas spectral information corresponding to high frequencies correspond to $\omega \in [0, \frac{1}{2}]$ around $\frac{1}{2}$.

From Theorem 3.15 follows that convolution of an input signal x with the filter h changes the frequency content $X = \hat{x}$ of the signal by pointwise multiplication with the frequency response H of the filter: depending on the properties of the magnitude response $|H|$ one distinguishes between lowpass, highpass, and bandpass filter.

Definition 3.36. Let $\omega_0, \omega_1 \in [0, \frac{1}{2}], \omega_1 < \omega_2$ denote a so-called cut-off frequency and h be a real-valued filter.

(1) h is called lowpass filter, if low frequencies ω with $0 \leq \omega \leq \omega_0$ are let through without attenuation whereas high frequencies ω with $\omega_0 \leq \omega \leq \frac{1}{2}$ are cut off. In other words,

$$|H(\omega)| \approx \begin{cases} 1 & \text{if } |\omega| \in [0, \omega_0] \\ 0 & \text{if } |\omega| \in [\omega_0, \frac{1}{2}] \end{cases}$$

(2) Similarly h is called highpass filter if

$$|H(\omega)| \approx \begin{cases} 1 & \text{if } |\omega| \in [\omega_0, \frac{1}{2}] \\ 0 & \text{if } |\omega| \in [0, \omega_0] \end{cases}$$

(3) Similarly h is called bandpass filter if

$$|H(\omega)| \approx \begin{cases} 1 & \text{if } |\omega| \in [\omega_0, \omega_1] \\ 0 & \text{if } |\omega| \in [0, \omega_0] \cup [\omega_1, \frac{1}{2}] \end{cases}$$

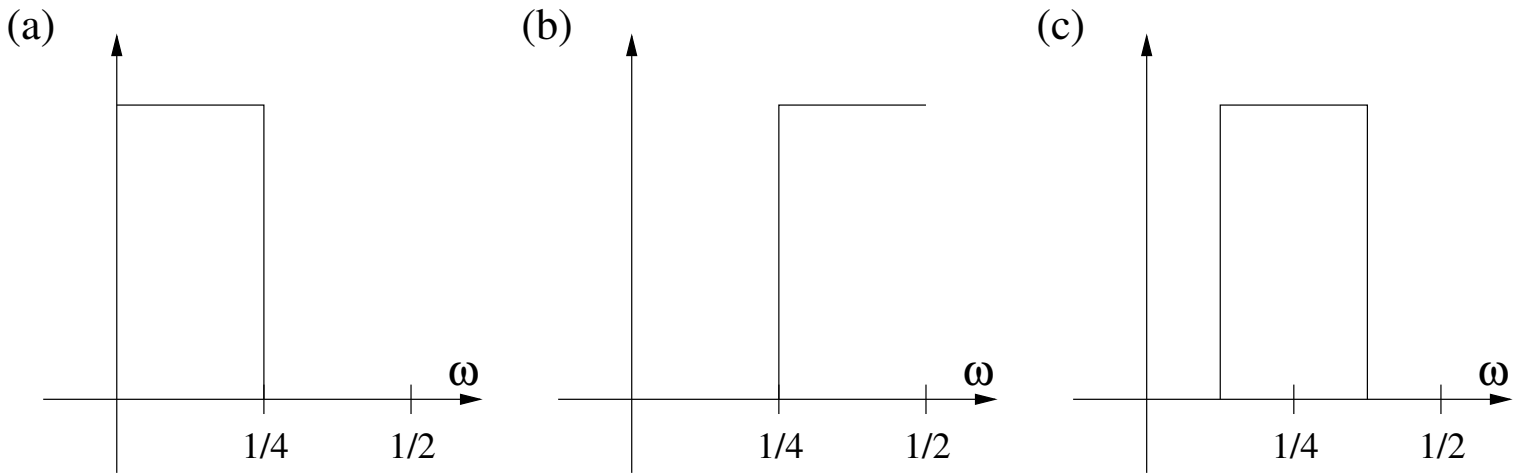


Figure 7: Magnitude responses of an ideal (a) lowpass, (b) highpass and (c) bandpass filter.

3.4 z-Transform and Transfer Function

The z-transform is a generalization of the discrete-time Fourier transform that allows many signals not having a Fourier transform to be described using related transform techniques.

Definition 3.37. Let $x\colon \mathbb{Z} \to \mathbb{C}$ be a signal. The z-transform of x, denoted by X, is defined to be the series

$$X(z) := \sum_{n\in\mathbb{Z}} x(n) z^{-n},$$

in the complex variable z.

Note 3.38. In complex analysis, series of the form $\sum_{n\in\mathbb{Z}} x(n) z^{-n}$ are also called Laurent series.

Note that when $z = e^{2\pi i\omega}$, the z-transform becomes the discrete-time Fourier transform

$$X(e^{2\pi i\omega}) = \sum_{n\in\mathbb{Z}} x(n) e^{-2\pi i n\omega}.$$

As with the DT-Fourier transform, the z-transfrom is only defined when the sum in Definition 3.37 converges. The sum generally does not converge for all values $z \in \mathbb{C}$.

Definition 3.39. Let X be the z-transform of the DT-signal $x\colon \mathbb{Z} \to \mathbb{C}$, then the region of convergence, denoted by D_x, is the largest open subset of $\mathbb{C}$ for which elements $z \in D_x$ the infinite series $\sum_{n\in\mathbb{Z}} x(n)z^{-n}$ converges.

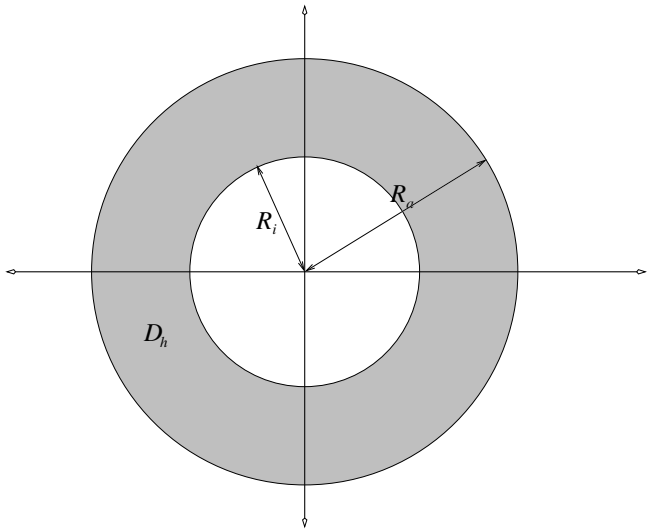
Note 3.40. The symbol X denotes the z-transform as well as the DT-Fourier transform (or frequency response) of some DT-signal x. This might be confusing since, for example, the z-transform $X(e^{2\pi i\omega})$ at $z = e^{2\pi i\omega}$ equals the DT-Fourier transfrom $X(\omega)$. However, from the context it should be clear which transform is meant.

The following theorem shows that the region of convergence D_x is an (possibly empty) annulus which is also illustrated by the figure below.

Theorem 3.41. *Let $x\colon \mathbb{Z} \to \mathbb{C}$ be a signal. Then the region of convergence D_x is an open annulus with inner radius $R_i(x)$ and outer radius $R_a(x)$, where*

$$R_i(x) = \overline{\lim}_{n\to\infty} |x(n)|^{1/n} \quad \textit{and} \quad R_a(x) = \frac{1}{\overline{\lim}_{n\to\infty} |x(-n)|^{1/n}}.$$

Hence, $D_x = \{z \in \mathbb{C} \mid R_i(x) < |z| < R_a(x)\}$.



Example 3.42. If $x : \mathbb{Z} \to \mathbb{C}$ is of finite length, i.e., $h(n) \neq 0$ only for a finite number of $n \in \mathbb{Z}$, then

$$R_i(h) = 0 \quad \text{and} \quad R_a(h) = \infty.$$

In other words, the region of concergence of the z-transform of a finite signal is $\mathbb{C} \setminus \{0\}$.

The following theorem shows, that two complex functions which are defined by Laurent series are equal if and only if all coefficients of the Laurent series coincide.

Theorem 3.43. [Identity theorem for Laurent series.] *Let $\sum_{n\in\mathbb{Z}} x_n z^n$ and $\sum_{n\in\mathbb{Z}} y_n z^n$ be two Laurent series which converge on the circle S_ρ of radius $\rho > 0$ around 0 uniformly to the same limit function f. Then for $n \in \mathbb{Z}$ holds*

$$x_n = y_n = \frac{1}{2\pi\rho^n} \int_0^{2\pi} f(\rho \cdot e^{i\varphi}) e^{-in\varphi} d\varphi.$$

The assumption are fulfilled if $S_\rho \subset D_x \cap D_y$.

Proof: See [Remmert], p. 284. □

What can be said about the order of magnitude of the signal values $x(n)$ and the region of convergence? The following theorem gives, in part, an answer.

Theorem 3.44. *Let $x\colon \mathbb{Z} \to \mathbb{R}$ and $\rho \in \mathbb{R}$ with $R_i(x) < \rho < R_a(x)$. Furthermore, let $\mu_\rho := \max_{|z|=\rho} |X(z)|$ denote the maximum of the z-transform X on the circle of radius ρ. Then the so called Gutzmersch formula holds:*

$$\sum_{n\in\mathbb{Z}} |x(n)|^2 \rho^{-2n} = \frac{1}{2\pi} \int_0^{2\pi} |X(\rho \cdot e^{i\varphi})|^2 d\varphi \leq \mu_\rho^2.$$

In addition, the Cauchy inequality holds:

$$|x(n)| \leq \mu_\rho \cdot \rho^n.$$

Proof: See [Remmert], p. 285. □

By the Gutzmersch formula the sequence $(x(n)\rho^{-n})_{n\in\mathbb{Z}}$ is in $\ell^2(\mathbb{Z})$. In the case $0 < \rho < 1$ the sequence x, if it is causal, is absolutely summable.

The z-transform has many properties generalizing the properties of the DT-Fourier transform (see Theorem 2.20).

Theorem 3.45. Let x, y and h DT-signals $\mathbb{Z} \longrightarrow \mathbb{C}$ with z-transforms X, Y and H, respectively. Then the following holds.

(1) Linearity: Let D be an open subset of $\mathbb{C}$ and $L(D)$ be the set all signals $x\colon \mathbb{Z} \longrightarrow \mathbb{C}$ such that $D \subseteq D_x$. Then the map

$$L(D) \ni x \mapsto X \downarrow D$$

is linear, where $X \downarrow D$ denotes the restriction of the z-transform to D.

(2) Time shift: Let $y = x^k$, i.e., $y(n) = x(n-k)$, then $D_y = D_x$ and $Y(z) = z^{-k}X(z)$.

(3) Modulation: Let $a \in \mathbb{C} \setminus \{0\}$ and $y(n) = a^n x(n)$, then $D_y = |a| \cdot D_x$ and $Y(z) = X(z/a)$.

(4) Complex conjugation: If $y(n) = \overline{x(n)}$ for all $n \in \mathbb{Z}$, then $D_y = D_x$ and $Y(z) = \overline{X(\overline{z})}$.

(5) Time reversal: If $y(n) = x(-n)$ for all $n \in \mathbb{Z}$, then $D_y = D_x^{-1}$ and $Y(z) = X(z^{-1})$.

(6) Convolution: If $y = h * x$, then for all $z \in D_h \cap D_x \subseteq D_y$ holds

$$Y(z) = H(z) \cdot X(z).$$

Proof: We give the proof for some of the properties and leave the rest as an exercise.

(2) Time shift:

$$\begin{aligned} Y(z) &= \sum_{n\in\mathbb{Z}} y(n)z^{-n} = \sum_{n\in\mathbb{Z}} x(n-k)z^{-n} \\ &= \sum_{n\in\mathbb{Z}} x(n-k)z^{-(n-k)}z^{-k} \\ &= z^{-k}\sum_{n\in\mathbb{Z}} x(n-k)z^{-(n-k)} = z^{-k}\sum_{n\in\mathbb{Z}} x(n)z^{-n} \\ &= z^{-k}X(z) \end{aligned}$$

(3) Modulation:

$$\begin{aligned} Y(\underbrace{z}_{\in D_y}) &= \sum_{n\in\mathbb{Z}} a^n x(n) z^{-n} \\ &= \sum_{n\in\mathbb{Z}} x(n) \left(\frac{z}{a}\right)^{-n} \\ &= \underbrace{X\left(\frac{z}{a}\right)}_{\in D_x} \end{aligned}$$

(4) Complex conjugation:

$$\begin{aligned} Y(z) &= \sum_{n\in\mathbb{Z}} y(n) z^{-n} = \sum_{n\in\mathbb{Z}} \overline{x(n)} z^{-n} \\ &= \overline{\sum_{n\in\mathbb{Z}} \overline{\overline{x(n)} z^{-n}}} = \overline{\sum_{n\in\mathbb{Z}} x(n) \overline{z}^{-n}} \\ &= \overline{X(\overline{z})} \end{aligned}$$

(6) Convolution:

$$
\begin{aligned}
H(z) \cdot X(z) &= \sum_{k\in\mathbb{Z}} h(k) z^{-k} \sum_{\ell\in\mathbb{Z}} x(\ell) z^{-\ell} \\
&= \sum_{k,\ell\in\mathbb{Z}} h(k) x(\ell) z^{-(k+\ell)} \\
&= \sum_{n\in\mathbb{Z}} \left(\sum_{k+\ell=n} h(k) x(\ell) \right) z^{-n} \\
&= \sum_{n\in\mathbb{Z}} \left(\sum_{k\in\mathbb{Z}} h(k) x(n-k) \right) z^{-n} \\
&= \sum_{n\in\mathbb{Z}} (h * x)(n) z^{-n} \quad = \quad \sum_{n\in\mathbb{Z}} y(n) z^{-n} \quad = \quad Y(z).
\end{aligned}
$$

We emphasize that the z-transform transforms the convolution of signals in pointwise multiplication. This fact is the basis for the specification of filters.

Similar to Definition 3.30 one has in the theory of filters another expression for the z-transform.

Definition 3.46. Let T be a BIBO-stable LTI system with impulse response $h = T[\delta] \in \ell^1(\mathbb{Z})$. Then the z-transform

$$H(z) := \sum_{n \in \mathbb{Z}} h(n) z^{-n}$$

is called transfer function of T (or h). One also speaks of the z-domain of T (or h).

3.5 Convolution for CT-Signals

Just as the analogy between the CT-signal spaces $L^p(\mathbb{R})$ and the DT-signal spaces $\ell^p(\mathbb{Z})$ there is also a similar analogy between filtering in the CT-case and in the DT-case. For example, the convolution of CT-signals can be defined in a similar fashion as for DT-signals (see Section 3.2). This is no coincidence, since convolution can be generally defined for so called locally compact groups G with a Haar measure. The convolution for CT-signals ($G = \mathbb{R}$) and for DT-signals ($G = \mathbb{Z}$) can then be considered as special cases of this more general concept.

In this section we give the main definitions and summarize some important properties of the continuous convolution.

Definition 3.47. For continuous-time signals $f, g\colon \mathbb{R} \to \mathbb{C}$ the convolution of f and g at $t \in \mathbb{R}$ is defined to be

$$(f * g)(t) := \int_{-\infty}^{\infty} f(s)g(t-s)ds.$$

Attention: $f * g$ does in general not exist! The next theorem gives some conditions on f and g which guarantees the the existence of $f * g$. (One compare this theorem with Theorem 3.13.)

Theorem 3.48. *Let $1 \leq p, q \leq \infty$, then the following properties hold.*

(1) *Let $f, g, h : \mathbb{R} \to \mathbb{C}$, such that all convolution integrals in question exist. Then $f * (g + h) = f * g + f * h$, $f * g = g * f$ and $(f * g) * h = f * (g * h)$. Furthermore holds $(\lambda f) * g = f * (\lambda g) = \lambda(x * g)$ for arbitrary $\lambda \in \mathbb{C}$.*

(2) *(Young Inequality) $L^1 * L^p \subseteq L^p$, i.e., for all $f \in L^1(\mathbb{R})$, $g \in L^p(\mathbb{R})$ holds $f * g \in L^p(\mathbb{R})$ and*

$$\|f * g\|_p \leq \|f\|_1 \cdot \|g\|_p.$$

(3) *Let p and q be conjugate exponents. Then for all $f \in L^p$ and $g \in L^q$ holds*

$$f \cdot g \in L^1 \quad \text{and} \quad \|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q$$
$$f * g \in L^\infty \quad \text{and} \quad \|f * g\|_\infty \leq \|f\|_p \cdot \|g\|_q.$$

(4) *For all $f \in L^p(\mathbb{R})$ and $g \in L^\infty(\mathbb{R})$ one has $f \cdot g$ in $L^p(\mathbb{R})$ and*

$$\|f \cdot g\|_p \leq \|f\|_p \cdot \|g\|_\infty.$$

For a proof we refer to [Folland]. Also in the CT-case the convolution of two functions transforms under the Fourier transform in pointwise multiplication (compare with Theorem 3.15).

Theorem 3.49. *Let $f, g \in L^2(\mathbb{R})$, then $f * g \in L^2(\mathbb{R})$ and*

$$\widehat{f * g} = \hat{f} \cdot \hat{g}.$$

3.6 Summary and Examples

In this chapter we have introduced a general system as a function between signal spaces. Imposing properties such as linearity, time-invariance, and stability lead us to stable LTI systems. One major result was formulated in Theorem 3.20: a stable LTI system T can be expressed as convolution operator C_h w.r.t. the impulse response $h = T[\delta] \in \ell^1(\mathbb{Z})$ and, vice versa, each sequence $h \in \ell^1(\mathbb{Z})$ defines a stable LTI system via convolution. Therefore, we often simply speak of a filter h and mean its associated convolution operator C_h.

In Definition 3.21 we distinguished the cases where the impulse response h is of finite length or infinite length. In the first case we called h an FIR filter and in the second case an IIR filter.

One important tool to describe the filter properties of some filter $h \in \ell^1(\mathbb{Z})$ is the frequency response H defined by the formula (see Definition 3.30):

$$H(\omega) := \sum_{n\in\mathbb{Z}} h(n)e^{-2\pi i n\omega}, \quad \omega \in [0,1].$$

One main property of the frequency response was that convolution of h with some signal x amounts to multiplication of H and X in the Fourier domain (see Theorem 3.15).

We cite [Glassner, p.218] to underscore the importance of the frequency response:

> In general, almost all filtering tasks can be usefully examined in frequency space, where we ask what happens to the spectrum of a signal as it passes through some system. Typically an important part of the analysis involves considering how the spectrum of the signal is scaled by the frequency response of the system, which for a linear, time invariant system is the Fourier representation of its impulse response. Thus such a filtering task may be viewed either as multiplication of two spectra or convolution of the signals.

Depending on the properties of the magnitude response we have in Definition 3.36 considered three types of filters: lowpass, highpass, and bandpass filters.

In 5 we will describe further filter characteristics and give some methods of how to construct FIR filters which approximate the ideal filters mentioned above. We conclude this chapter with two examples which illustrates the introduced notions.

3.6.1 Haar filter

We start with an easy example and consider a pair of FIR filters h and g defined by

$$h(n) = \begin{cases} \frac{1}{2} & \text{if } n = 0, 1, \\ 0 & \text{elsewhere} \end{cases} \quad \text{and} \quad g(n) = \begin{cases} \frac{1}{2} & \text{if } n = 0, \\ -\frac{1}{2} & \text{if } n = 1, \\ 0 & \text{elsewhere.} \end{cases}$$

These filters play a crucial role in the Haar wavelet transform as we will see in Chapter **??**.

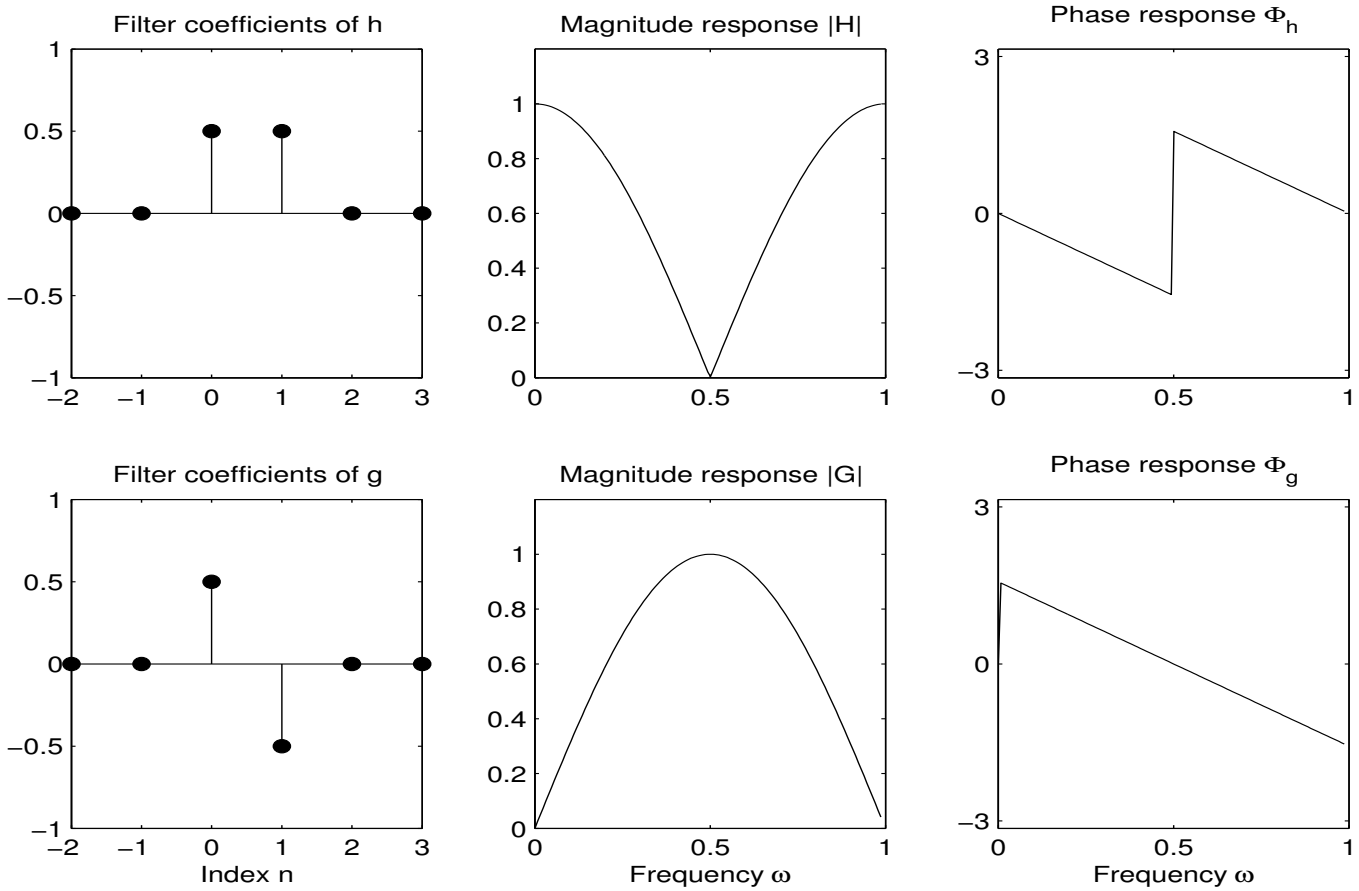
Their frequency response can easily be computed:

$$H(\omega) = \frac{1}{2} + \frac{1}{2}e^{-2\pi i\omega} = e^{-\pi i\omega} \cdot \frac{1}{2}(e^{\pi i\omega} + e^{-\pi i\omega}) = e^{-\pi i\omega}\cos(\pi\omega)$$

and

$$G(\omega) = \frac{1}{2} - \frac{1}{2}e^{-2\pi i\omega} = e^{-\pi i\omega} \cdot \frac{1}{2}(e^{\pi i\omega} - e^{-\pi i\omega}) = e^{-\pi i\omega} \cdot i \cdot \sin(\pi\omega).$$

The magnitude response and phase response of h and g are shown in the following figure.



Note 3.50. The filters h and g have the following properties:

(1) The frequency response H is a modulation of a cosine-function. Similar G is, up to the factor i, a modulation of a sine-function.

(2) $|H|$ is in the interval $[0, \frac{1}{2}]$ monotonously decreasing with $H(0) = 1$ and $H(\frac{1}{2}) = 0$. Hence, in some sense, H can be regarded as a lowpass filter. The quality, however, is very poor.

(3) Similarly $|G|$ is in the interval $[0, \frac{1}{2}]$ monotonously increasing with $G(0) = 0$ and $H(\frac{1}{2}) = 1$. Hence, G can be regarded as highpass filter.

(4) The filters h and g constitute a so-called pair of associated filters. For a signal x, the filtered signal $h * x$ contains the "low-frequency content" of x, whereas $g * x$ contains the "high-frequency content" of x.

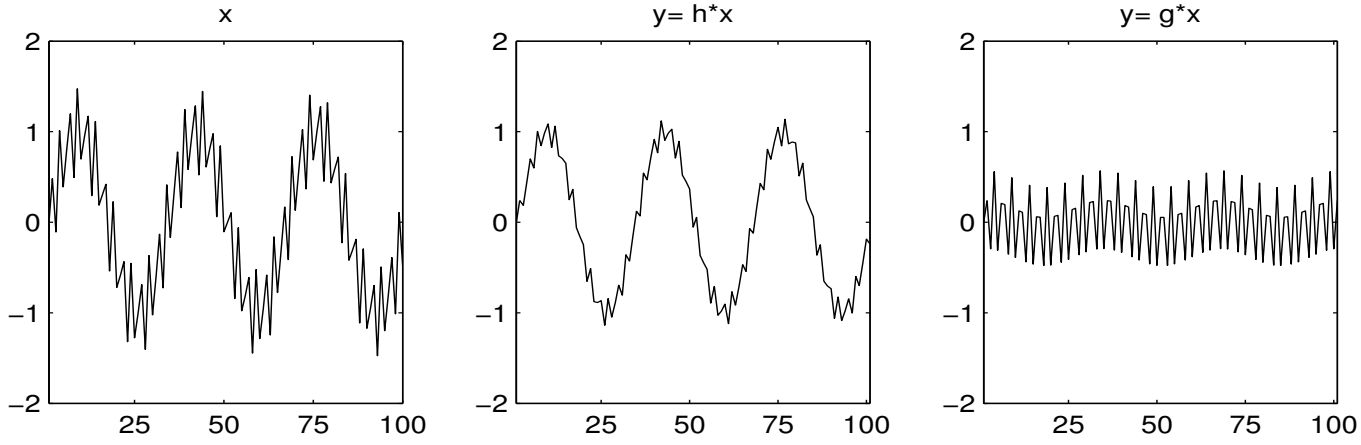
(5) The filters h an g considered as elements in $\ell^2(\mathbb{Z})$ are orthogonal.

(6) The filters h and g form the core of a so-called 2-band multirate filter bank.

The following figure shows as input signal x a superposition of two sines of frequency 3 Hz with amplitude 1 and of frequency 40 Hz with amplitude $1/2$ for the first second sampled with sampling rate $1/100$:

$$x(n) = \sin(3 \cdot 2\pi(n/100)) + \frac{1}{2} \cdot \sin(40 \cdot 2\pi(n/100)).$$

As explained before, the output signal $h * x$ captures the low frequency characteristics of x, whereas $g * x$ captures the high frequency characteristics of x.



3.6.2 Averaging Filter

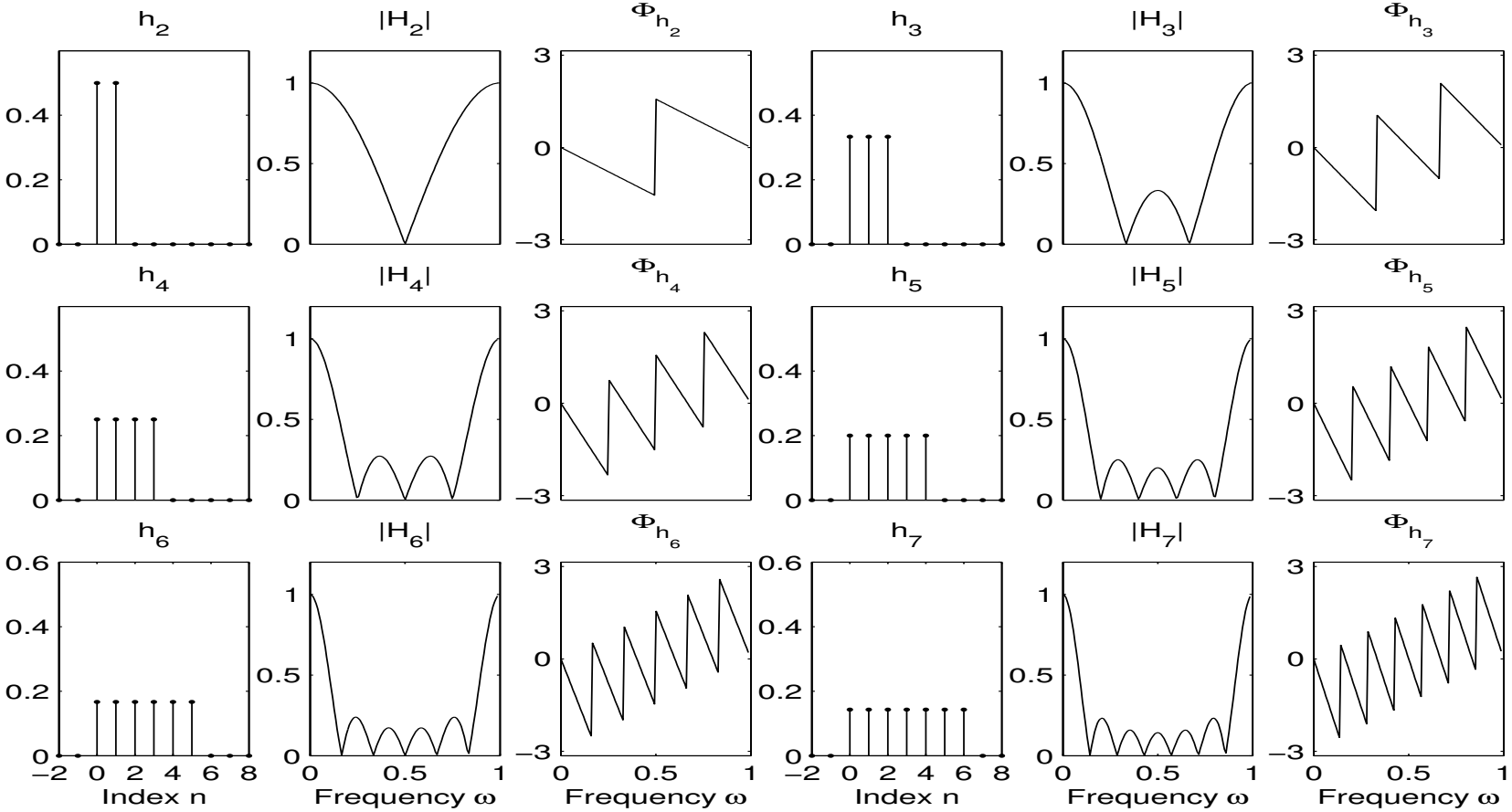
The averaging filters h_N of length N is defined by

$$h_N(n) = \begin{cases} \frac{1}{N} & \text{if } 0 \leq n \leq N-1 \\ 0 & \text{elsewhere.} \end{cases}$$

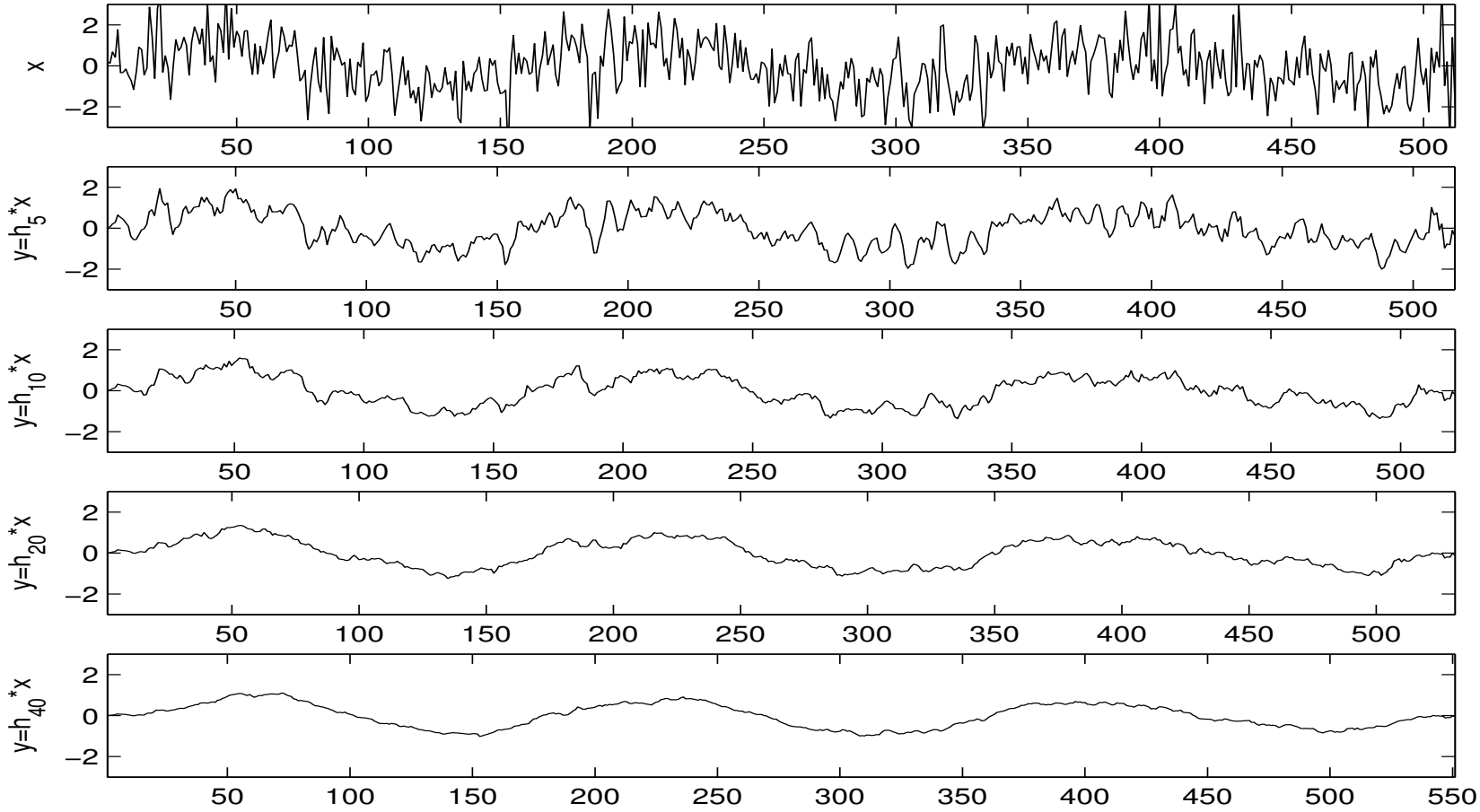
Hence the Haar lowpass filter of the last section is the averaging filter of length 2. The frequency response can easily be computed by using the geometric series:

$$\begin{aligned} H(\omega) &= \sum_{n=0}^{N-1} \frac{1}{N} e^{-2\pi i \omega n} = \frac{1}{N} \frac{1 - e^{-2\pi i \omega N}}{1 - e^{-2\pi i \omega}} \\ &= \frac{1}{N} \frac{e^{2\pi i \omega N/2} - e^{-2\pi i \omega N/2}}{e^{2\pi i \omega/2} - e^{-2\pi i \omega/2}} \cdot \frac{e^{-2\pi i \omega N/2}}{e^{-2\pi i \omega/2}} \\ &= \frac{1}{N} \frac{\sin(\pi \omega N)}{\sin(\pi \omega \omega)} e^{-2\pi i (N-1)\omega/2} \end{aligned}$$

The magnitude response and phase response of h_N are shown in the following figure for $N = 2$ to 7.



The following figure shows as an input signal x and for output signal $y = h * x$ filtered with $h = h_5, h_{10}, h_{20}$, and h_{40}, respectively.



Note 3.51. The following observation are illustrated by the last two figures.

(1) The cut-off frequency of h_N is given by $\omega_0 = 1/N$. However, the filter properties are not very good. For example, the pass band is not flat but drops right away and the stop band exhibits many ripples of large amplitude.

(2) The previous figure illustrates that with increasing N the filtered signal $y = h_N * x$ reflects better and better the low-frequency content of x.

(3) The length of the output signal $y = h * x$ does not only depend on the length of the input signal x but also on the length of the filter length N of the FIR filter h. For example, in our example the signal length of x is 512 and the length of the filtered signal $y = h_{40} * x$ is 551.

Chapter 4: Sampling and Aliasing

Doing signal processing with a computer one can, for example, store and process only a finite number of parameters. Therefore, a continuous-time signal has, in general, to be approximated in order to describe the approximation by a finite number of parameters. Mathematically, this corresponds to a projection of the signal onto a finite-dimensional vector space, which is the linear hull of a finite set of so-called synthesis functions. For example, the discrete set of parameters could be

- the Fourier coefficients (for periodic signals),
- the coefficients of polynomials (when representing a function by its Taylor series), or
- the values of a CT-signal on a finite number of points in time.

The third method, also referred to as sampling, is the easiest and most common way to discretize a CT-signal. However, when sampling a signal one looses, in general, information. This can lead to strong artefacts and distortions — also known as aliasing — when reconstruction a CT-signal from the sampled version. In this chapter we

- give a strict definition of sampling,
- discuss conditions on the CT-signals which are sufficient for a perfect reconstruction of the signal from the sampled version,
- desribe the aliasing effects, and
- discuss some methods to soften the resulting aretefacts and distortions.

4.1 Sampling

Definition 4.1. In the following we summarize the basic definitions and notions.

(1) The procedure to generate a DT-signal by taking the values of a CT-signal on a discrete set of points (in time) is referred to as sampling.

(2) The values of the DT-signal are then called samples, the points are called sample points.

(3) In case any two neighboring sample points have the same distance $T > 0$ (equidistant), on also speaks of T-sampling. Then a CT-signal $f\colon \mathbb{R} \longrightarrow \mathbb{C}$ transforms into the DT-signal $x\colon \mathbb{Z} \longrightarrow \mathbb{C}$ with

$$x(n) := f(T \cdot n).$$

(4) For T-sampling, the number of samples per time unit is $\frac{1}{T}$ which is referred to as sampling rate. The sampling rate is measured in unit Hertz or simply hz.

The transition from a CT-signal $f : \mathbb{R} \longrightarrow \mathbb{C}$ to a sampled DT-signal $x : \mathbb{Z} \longrightarrow \mathbb{C}$ results, in general, in a loss of information, i.e., f cannot be reconstructed from x. At least, one can try to reconstruct an approximation of f. This is done be using so-called synthesis functions. For example, in case of T-sampling one can choose the characteristic functions $\mathbf{1}_{[t_k,t_{k+1})}(t)$ of the intervals $[t_k, t_{k+1}), t_k := T \cdot k$, between two neighboring sample points. The one gets the function f_T defined by

$$f_T(t) := \sum_{k=-\infty}^{\infty} x(k)\mathbf{1}_{[t_k,t_{k+1})}(t) = \sum_{k=-\infty}^{\infty} f(T \cdot k)\mathbf{1}_{[t_k,t_{k+1})}(t).$$

In this case f is approximated by f_T, i.e., the value $f(t)$ is approximated by the constant $f(t_k)$ where t_k is the closest sampling point on the left of t.

Of course, synthesis functions are not uniquely determined. For example, instead of the characteristic functions one could also pick spline functions or trigonometric functions. Hence, the quality of the approximation depends not only on the number and quality of the sample values but also on the choice the synthesis function. However, there is no canonical choice. The "right" choice depends very much on the class of CT-signals under consideration and the application in mind.

For a certain class of CT-signals the choice of the so-called sinc-functions lead to a perfect reconstruction of the CT-signal from its samples. This amazing property is the content of the next section.

4.2 Shannon Sampling Theorem

The Fourier transform $\hat{f}$ of a continuous-time signal $f \in L^1(\mathbb{R})$ exhibits the spectral information on the signal. Intuitively, the value $|\hat{f}(\omega)|$ expresses the intensity in which the frequency $\omega \in \mathbb{R}$ is "contained" in the signal f. Sometimes it suffices to consider signals which only contain a certain range of frequencies.

For example, the human auditory system only recognizes frequencies which are below a certain threshold (around 20 kHz). Therefore, for a given audio signal wiping out all frequencies above this threshold results in a signal with no audible loss — at least for the human ear. (A bat, however, would probably protest against such an intervention.)

This is a good point for an excursion into the world of bats — creatures which make up one quarter of the world's mammal species. On Jim's Bat Page (`http://www.jimlev.warc.org.uk/bats.htm#bat1`) one finds the following nice account on echolocation.

Echolocation of bats

Bats find their way around in the dark and locate food by calling out and listening for echoes from nearby objects. In this way bats can detect such things as trees, buildings, the ground, telephone wires and flying insects. The echolocation systems of bats enable them to navigate, also to detect and home in on prey. In flight they are able to avoid obstacles and other bats. They are able to determine the type, location, direction and speed of their prey from the echoes they receive. An almost immediate echo will be received from a nearby object and it will be relatively loud. The echoes from further objects will be quieter and will take longer to return.

The sounds sent out by bats are very much higher in frequency than those that can be detected by humans. We can hear sounds in the frequency range 20 Hz to $20,000$ kHz whilst bats operate between 20 kHz and 150 kHz. This is known as ultrasonic sound.

Typically, Pipistrelles use $40 - 50$ kHz and Horseshoes use $80 - 100$ kHz. Bat calls are very complex. They may have several different frequencies or notes, and vary in loudness.

The speed of sound through air is fairly constant and is equal to the frequency of the sound multiplied by its wavelength. This means that if the frequency goes up, the wavelength goes down. If the speed of sound is about 0.3 kilometers per second, a 20 kHz sound has a wavelength of about 0.015 meters or 15 mm. A sound at 150 kHz has a wavelength of about 2 mm. This wavelength range between about 2 mm and 15 mm is also the approximate range of sizes of the insects they eat. The shorter the wavelength (or the higher frequency) used by the bats then the smaller the prey that can be detected. The shorter wavelengths give more accuracy in homing in on prey.

The loudness of the ultrasonic pulses also changes depending on whether the bat is searching for prey or capturing it. It is loudest during search becoming quieter during capture.

One would expect that the bats would fly around using the highest frequency so that they could navigate and avoid obstacles with greatest accuracy. Unfortunately the highest frequencies do not travel far compared with the lower frequencies (longer wavelengths). For this reason bats use different frequencies for different purposes. They use the lowest frequencies for travelling and searching for prey, changing to higher frequencies to home in as the distance to the prey shortens, then the highest frequencies at short range to contact and seize the prey. The sounds sent out by bats usually consist of a series of pulses of ultrasound which sweep down from a high frequency to a low frequency over a few thousands of a second. The time space between the pulses allows any echo to return and to be processed in the brain of the bat to give it information about its surroundings, and any prey that may be in range.
Some bats use pulses which do not change in frequency. It has been shown that these bats detect frequency changes in the echoes due to either their own movement, or the movement of their prey, or both. The frequency change is due to what is known as Doppler shift. This is the effect that occurs when a police car with its siren wailing

passes you. The frequency of the sound you hear drops as the car passes you. It is higher as it approaches you and lower as it speeds away from you. The amount of shift in frequency depends on the speed of the car - the higher the speed then the greater the Doppler shift. If the car was stationary and you ran past it fast enough you would also hear a Doppler shift due to your motion. The frequency shift that a bat detects is a combination of the relative speeds and directions if itself and its prey.
Humans see by building up a picture in the brain from light reflected from surrounding objects entering the eye. The bat brain builds up a similar picture from reflected sound. In the case of a bat, the sound is arriving in pulses and the sound fed to the brain must be a bit like the flashing light that we see at a disco. Just as our brains can make out something of the inside of the disco and see the dancers, the bats brain can build up a picture of its surroundings and nearby prey from the sound pulses it receives. In addition, from the nature of the echoes receive, the bat is able to determine what sort of prey it is seeing. For example, from the Doppler shift and the changes in loudness of the echoes due to pulses reflected from insect wings the bat can detect the speed that the insect is flapping its wings and deduce the size of the prey.

Small insects flap faster than large ones, for example, mosquitoes beat their wings more than 200 times a second whereas larger insects such as moths and large beetles may only beat about 50 times a second.
Larger bats generally use lower frequency ranges than smaller bats. This means that they can see further than smaller bats but do not see the smallest insects, - which they would find it hard to catch anyway.
The echolocation methods used by bats are good enough to allow capture of the prey in the bat's mouth most of the time. For near misses, the wing tips and tail are used to scoop up the prey.

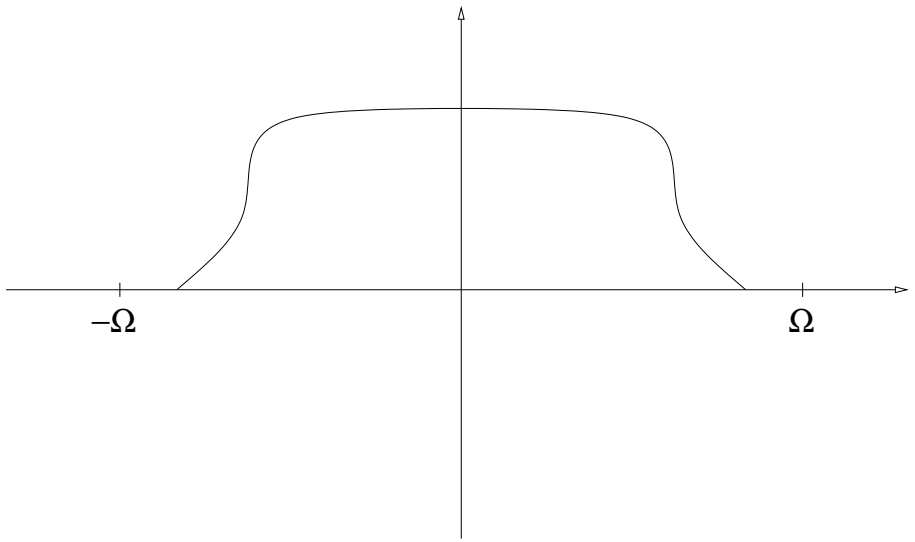
For further readings Jim refers to "Bats" by Phil Richardson published by Whittet Books (ISBN 0-905483-41-3) and "The Natural History of Hibernating Bats" by Roger Ransom published by Christopher Helm (ISBN 0-7470-2802-8).

Signals which contain only frequencies up to a certain threshold are called bandlimited. We give a more rigorous definition.

Definition 4.2. Let $\Omega > 0$. The CT-signal $f \in L^2(\mathbb{R})$ is called Ω-bandlimited if the Fourier transform $\hat{f}$ vanishes for $|\omega| > \Omega$:

$$\hat{f}(\omega) = 0 \quad \forall |\omega| > \Omega, \quad \text{i.e., } \operatorname{supp} \hat{f} \subset [-\Omega, \Omega].$$

The following figure shows a Fourier transform $\hat{f}$ of an Ω-bandlimited function f.



Bandlimited functions have surprising properties. For example, a continuous-time function $f \in L^2(\mathbb{R})$ may be, in general, discontinuous or even undefined at some points. If f is Ω-bandlimited, however, the following theorems hold:

Theorem 4.3. *If $f \in L^2(R)$ is bandlimited then f is smooth (i.e., infinitely often differentiable).*

Note 4.4. As mentioned in Note 1.35 $f \in L^2(\mathbb{R})$ denotes actually a whole class of functions. Therefore, to be more precise, Theorem 4.3 should read that in case f is bandlimited there is a smooth function equivalent to f. Only for this smooth (and hence continuous) representative it is possible to take samples.

The smaller Ω, i.e., the smaller the range of frequency contained in the signal, the smoother the signal. In other words, local bursts and discontinuity points in the signal result in high frequency components in the Fourier spectrum.

One other fact about bandlimited function, often referred to as uncertainty principle, is important in view of applications.

Theorem 4.5. *If $f \in L^2(R)$ is bandlimited then f cannot be compactly supported.*

Therefore, in real-world applications, when one has to deal with time-limited signals one cannot simply assume that f is bandlimited and has to struggle with the consequences which are known as aliasing.

The sampling theorem by Shannon says that a bandlimited function f can be reconstructed perfectly from a suitable equidistant set of samples.

Theorem 4.6. [Shannon.] [1] *Let $f \in L^2(\mathbb{R})$ be an Ω-bandlimited function and let x be the T-sampled version of f with $T := \frac{1}{2\Omega}$, i.e., $x(n) = f(nT)$, $n \in \mathbb{Z}$. Then f can be reconstructed from x:*

$$f(t) = \sum_{n=-\infty}^{\infty} x(n)\mathrm{sinc}\left(\frac{t-nT}{T}\right) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2\Omega}\right)\mathrm{sinc}\,(2\Omega t - n)\,,$$

where the scaled and translated versions of the sinc-function

$$\mathrm{sinc}\,(x) := \begin{cases} \frac{\sin \pi x}{\pi x} & x \neq 0 \\ 1 & x = 0. \end{cases}$$

are used as synthesis function.

[1] Shannon(1949), Kotel'nikov(1933), Whittaker(1915), de la Vallée Poissin(1908)

Sketch of proof: Since f is Ω-bandlimited, i.e., $\operatorname{supp}\hat{f} \subset [-\Omega, \Omega]$ by the inverse Fourier transform we get

$$f(t) = \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{2\pi i \omega t} d\omega \quad \text{and hence} \quad f\left(\frac{n}{2\Omega}\right) = \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{\frac{\pi i \omega n}{\Omega}} d\omega.$$

We extend $\hat{f}$ to a 2Ω-periodic function and denote this function by g. Then g can be represented by its Fourier series:

$$g(t) = \frac{1}{2\Omega} \sum_{n \in \mathbb{Z}} c_n e^{2\pi i t n}$$

with

$$c_n = \int_{\omega=0}^{2\Omega} g(\omega) e^{\frac{-2\pi i n \omega}{2\Omega}} dt = f\left(\frac{-n}{2\Omega}\right).$$

and hence

$$g(t) = \frac{1}{2\Omega} \sum_{n \in \mathbb{Z}} f\left(\frac{n}{2\Omega}\right) e^{-2\pi i n t}.$$

From this we get

$$\begin{aligned} f(t) &= \int_{-\Omega}^{\Omega} g(\omega) e^{2\pi i \omega t} d\omega \\ &= \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \sum_{n \in \mathbb{Z}} f\left(\frac{n}{2\Omega}\right) e^{-2\pi i n \omega} e^{2\pi i \omega t} d\omega \\ &= \sum_{n \in \mathbb{Z}} f\left(\frac{n}{2\Omega}\right) \underbrace{\frac{1}{2\Omega} \int_{-\Omega}^{\Omega} e^{2\pi i \omega (t-n)} d\omega}_{=\operatorname{sinc}(2\Omega t - n))} . \end{aligned}$$

□

Given an Ω-bandlimited function one needs a T-sampling with $T \geq \frac{1}{2\Omega}$. The sampling rate $\frac{1}{T} = 2\Omega$ Hz is sufficient for a perfect reconstruction of the signal.

Definition 4.7. For an Ω-bandlimited function, the sampling rate 2Ω is called Nyquist-Rate. Contrary, $\Omega = \frac{1}{2T}$ is called Nyquist frequence for the sampling interval of length T.

Example 4.8. Common sampling rates are:

Device	sampling rate	reproducible frequency
Telefone	8 kHz	≤ 4 kHz
DAT	32 kHz	≤ 16 kHz
	$44,1$ kHz	$\leq 22,05$ kHz
	48 kHz	≤ 24 kHz
CD	$44,1$ kHz	$\leq 22,05$ kHz

As mentioned before, the human auditory system only recognizes frequencies up to 20 kHz. Thus the reconstructed digitized CD-signals sounds like the original music piece — there is no audible loss of information.

4.3 Aliasing

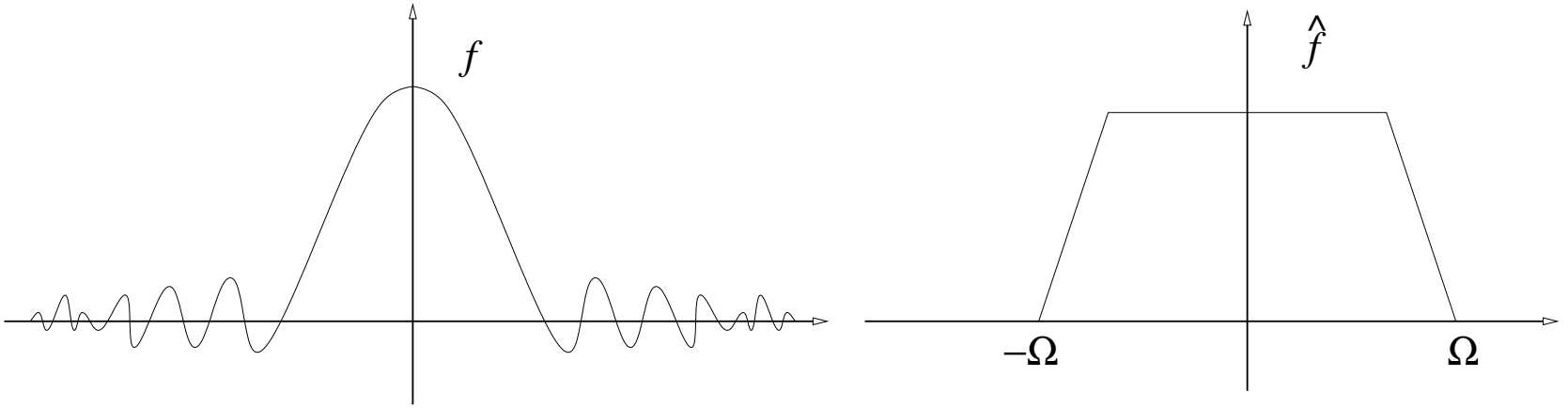
Ken C. Pohlmann gives in his book "Principles of Digital Audio" [Pohlmann] the following general description of the aliasing problem:

> One particular challenge to a digital audio system designer is that of aliasing, a kind of sampling confusion that can occur in the recording side of the signal chain. Just as a criminal can take many names and thus confuse identity, aliasing can create false signal components. These erroneous signals can appear within the audio bandwidth and are impossible to distinguish from legitimate signals. Obviously, it is the designer's obligation to prevent such distortion from ever occurring. In practice, aliasing is not a limitation. It merely underscores the importance of observing the criteria of sampling theory.

From an image processing point of view Andrew S. Glassner writes in [Glassner]:

> One way to make sure our pictures are as good as they can be is to make sure that they contain nothing extraneous or wrong. This sounds fine in principle, but it turns out that the mere process of representing an inherently continuous color picture on a device with a finite number of spatial locations usually introduces errors of its own. These errors are known collectively as aliasing, and they lead to phenomena like jagged edges, thin objects that seem to be broken into pieces, and, in animations, objects that suddenly appear and disappear.

In the following, we illustrate what happens with the spectrum when sampling a signal. Let $f \in L^2(\mathbb{R})$ be a Ω-bandlimited function as shown in the following figure.



Let $x : \mathbb{Z} \longrightarrow \mathbb{C}$ be the DT-signal by sampling f with a sampling rate $\frac{1}{T}$, i.e.,

$$x(n) = f(nT), \quad n \in \mathbb{Z}.$$

Then the spectrum $\hat{x}$ of x is the 1-periodization of a scaled version of the spectrum $\hat{f}$. The precise statement is formulated in the next theorem.

Theorem 4.9. *Let f be a continuous function with $f \in L^2(\mathbb{R})$ and x be the DT-signal obtained from f by T-sampling, i.e., $x(n) = f(nT)$ for $n \in \mathbb{N}$. Then*

$$\hat{x}(\omega) = \frac{1}{T} \sum_{k \in \mathbb{Z}} \hat{f}\left(\frac{\omega + k}{T}\right).$$

Proof: We first proof the result for the case $T = 1$. Let

$$g(\omega) := \sum_{k\in\mathbb{Z}} \hat{f}(\omega + k),$$

$\omega \in \mathbb{R}$. To show that $\hat{x} = g$ it suffices to show that the inverse Fourier transform of g and $\hat{x}$ coincide, i.e., $x = \check{g}$. To this means we compute the nth Fourier coefficient of g:

$$\begin{aligned}
\int_0^1 \hat{g}(\omega) e^{2\pi i\omega n} d\omega &= \int_0^1 \sum_{k\in\mathbb{Z}} \hat{f}(\omega + k) e^{2\pi i\omega n} d\omega \\
&= \sum_{k\in\mathbb{Z}} \int_k^{k+1} \hat{f}(\omega) e^{2\pi i(\omega-k)n} d\omega \\
&= \int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i\omega n} d\omega \\
&= f(n) \quad = \quad x(n).
\end{aligned}$$

For general $T > 0$ we consider the function h defined by $h(t) := f(t \cdot T)$ instead of f in the above proof. Then the general case follows by using rule (5) of Theorem 2.9. □

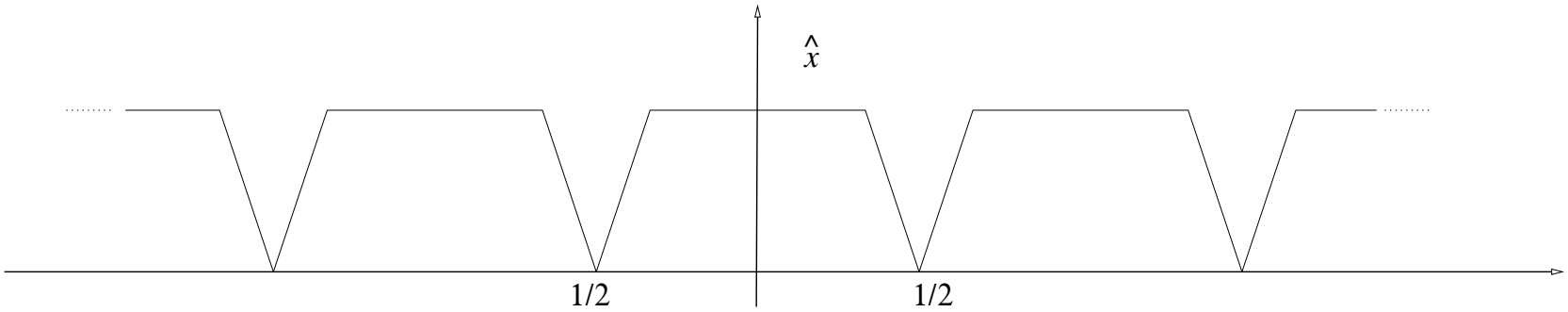
We now consider the three cases where the sampling rate $\frac{1}{T}$ is equal to, above and below the Nyquist rate 2Ω.

Case 1: $T = \frac{1}{2\Omega}$

When the sampling rate equals the Nyquist rate, the Fourier transform of $\hat{f}$ can be recovered from $\hat{x}$ since

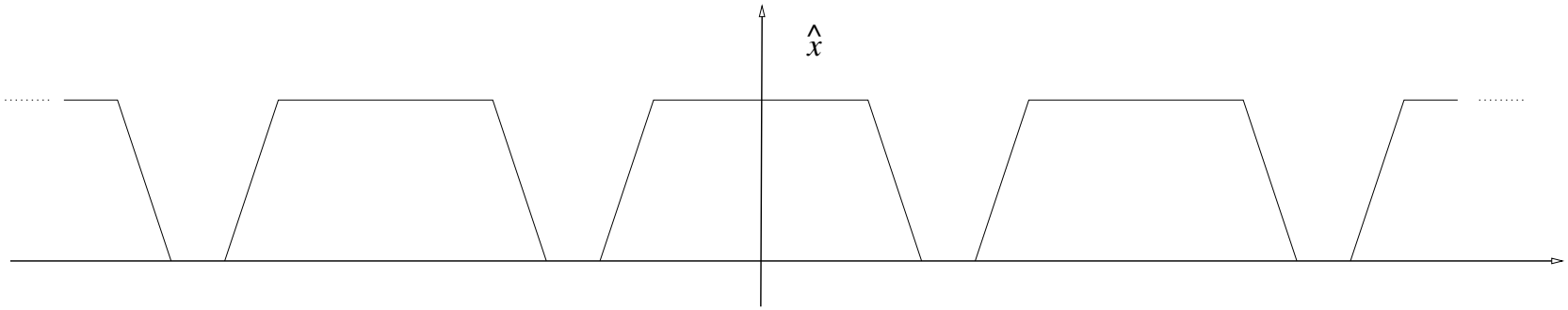
$$\hat{f}(\omega) = \chi_{[-1/2T,1/2T]} \cdot T \cdot \hat{x}(T\omega) = \chi_{[-\Omega,\Omega]} \cdot T \cdot \hat{x}(T\omega)$$

Therefore, f can be perfectly recovered from x by taking the inverse Fourier transform of $\chi_{[-\Omega,\Omega]} \cdot T \cdot \hat{x}(T\cdot)$. Actually, this is just a reformulation of Shannon's sampling theorem.



Case 2: $T < \frac{1}{2\Omega}$

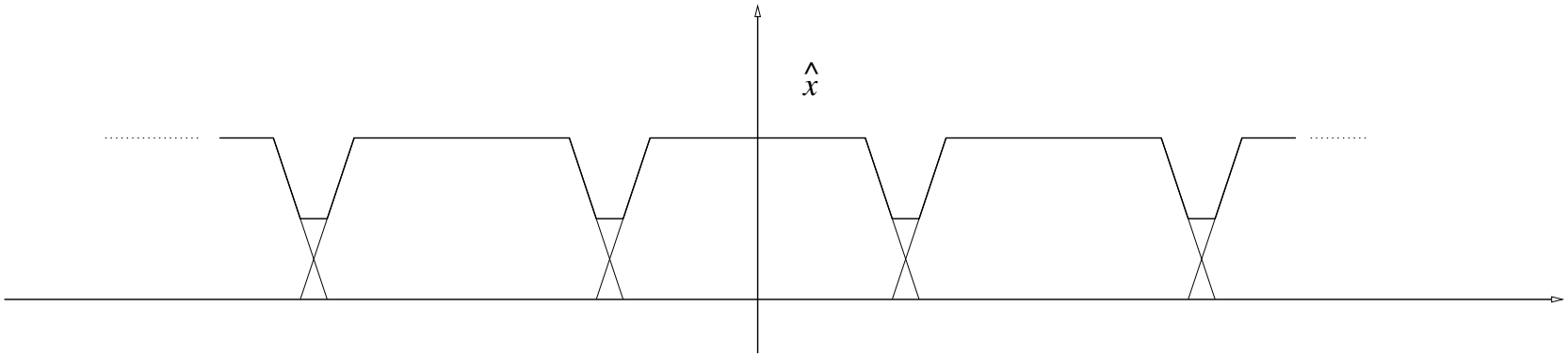
In this case the sampling rate $\frac{1}{T}$ is above the Nyquist rate 2Ω. One then also speaks from oversampling. As in Case 1, $\hat{f}$ can be recovered from $\hat{x}$ and f can be perfectly reconstructed from x.



One can show that in the case of oversampling, increasing the sampling rate leads to faster convergence properties of the reconstruction series $f(t) = \sum_{n=-\infty}^{\infty} x(n)\mathrm{sinc}\left(\frac{1}{T}(t - nT)\right)$.

Case 3: $T > \frac{1}{2\Omega}$

In this case the sampling rate $\frac{1}{T}$ is below the Nyquist rate 2Ω. One then also speaks from undersampling. In this case, reconstruction of f from x is not any longer possible. In the spectrum of $\hat{x}$ there is an "overlap" of the $\frac{1}{T}$-translated spectra of $\hat{f}$ and $\hat{f} = \mathbf{1}_{[-\Omega,\Omega]}\hat{x}$ does not hold any longer. This effect is known as aliasing.



We investigate the aliasing effect in a more rigorous way. Let $T = \frac{1}{\Omega}$ be a sampling rate for some $\Omega > 0$ and $f \in L^2(\mathbb{R})$ only be an Ω'-bandlimited function for some Ω' with

$$\Omega < \Omega' < 3\Omega.$$

Then T-sampling of f leads to the undersampled DT-signal x with $x(k) := f(kT)$. By the Fourier inversion holds

$$\begin{aligned} f(kT) &= \int_{-\Omega'}^{\Omega'} \hat{f}(\omega) e^{2\pi i \omega k T} d\omega \quad (\text{as } \hat{f} = 0 \text{ for } |\omega| > \Omega') \\ &= \int_{-3\Omega}^{3\Omega} \hat{f}(\omega) e^{2\pi i \omega k T} d\omega \\ &= \int_{-3\Omega}^{-\Omega} \hat{f}(\omega) e^{2\pi i \omega k T} d\omega + \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{2\pi i \omega k T} d\omega + \int_{\Omega}^{3\Omega} \hat{f}(\omega) e^{2\pi i \omega k T} d\omega. \end{aligned}$$

By a suitable substitution and using $e^{2\pi i\omega kT} = e^{2\pi i\omega' kT}$ for $\omega' - \omega \in \mathbb{Z}$ one obtains

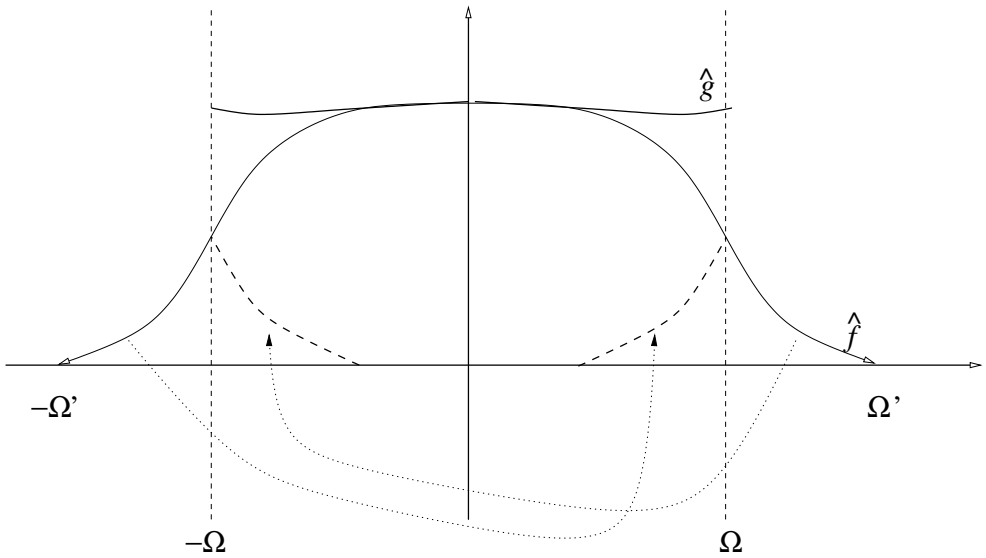
$$x(k) = f(kT) = \int_{-\Omega}^{\Omega} \underbrace{(\hat{f}(\omega) + \hat{f}(\omega + 2\Omega) + \hat{f}(\omega - 2\Omega))}_{=:(*)} e^{2\pi i\omega kT} d\omega \qquad (3)$$

We define a function g via its Fourier transform:

$$\hat{g}(\omega) := \begin{cases} (*) & \text{if } |\omega| \leq \Omega \\ 0 & \text{otherwise.} \end{cases}$$

Then, g is Ω-bandlimited and by Equation (3) holds $g(kT) = f(kT)$ for all $k \in \mathbb{Z}$.

Conclusion: The reconstruction of the CT-signal from the undersampled DT-signal x (by means of the sinc-synthesis functions of the Shannon Theorem 4.6) does not result in the original signal f but in the signal g.



As illustrated, in the spectrum of g high frequency above $|\Omega|$ are "fold" into the interval $[-\Omega, \Omega]$. In other words, in the undersampled DT-signal x high frequency components of f outside the interval $[-\Omega, \Omega]$ cannot be distinguished from certain low frequency components of f within $[-\Omega, \Omega]$, i.e., high-frequency components take on the identity of a lower frequency.

This effect is known as aliasing.

4.4 Down- und Upsampling

In Example 3.4 and 3.5 we have already introduced the downsampler $\downarrow M$ and upsampler $\uparrow M$ defined by

$$(\downarrow M)[x](n) := x(M \cdot n) \quad \text{and} \quad (\uparrow M)[x](n) = \begin{cases} x(n/M), & \text{if } M|n, \\ 0, & \text{otherwise.} \end{cases}$$

for some $M \in \mathbb{N}$, $n \in \mathbb{Z}$. Down- und upsampler play a central role in digital signal processing such as in the theory of multirate filterbanks — as we will see in a later chapter.

In this section we summarize some of the basic properties of the down- and upsampler and discuss the problem which arise in connection with the aliasing effect.

In the following discussion, we restrict to the case $M = 2$ which already exhibits the typical problems to cope with.

As mentioned in Example 3.28 the down- and upsampler are linear, continuous, and stable. The downsampler $\downarrow M$ destroys information. One does have $(\downarrow M)(\uparrow M) = I$, but $(\uparrow M)(\downarrow M) \neq I$.

A nice mathematical property is the fact that the sampling operators $(\downarrow 2)$ and $(\uparrow 2)$ are transposed maps with respect to the $\ell^2(\mathbb{Z})$-inner product. In other words, for all $x, y \in \ell^2(\mathbb{Z})$ holds

$$\langle (\downarrow 2)x | y \rangle_{\ell^2(\mathbb{Z})} = \langle x | (\uparrow 2)y \rangle_{\ell^2(\mathbb{Z})}.$$

4.4.1 Sampling operators in the z-domain

Next we investigate the effects of the sampling operators in the z-domain. Note that since $(\uparrow 2)$ and $(\downarrow 2)$ are not time invariant they are not representable as convolution filter. Therefore, there is no z-transforms for $(\uparrow 2)$ and $(\downarrow 2)$. However, the effects on the sampled signals can be studied in the z-domain.

Let $x \in \ell^1(\mathbb{Z})$ be the input signal, $u := (\uparrow 2)[x]$ be the upsampled, and $v := (\downarrow 2)[x]$ be the downsampled signal. Then the z-transforms X, V, and U, satisfy the following relations:

$$U(z) = \sum_{k\in\mathbb{Z}} x_k z^{-2k} = X(z^2) \tag{4}$$

$$\begin{aligned} V(z) &= \sum_{k\in\mathbb{Z}} x_{2k} z^{-k} \\ &= \frac{1}{2}\sum_{k\in\mathbb{Z}} x_k (z^{\frac{1}{2}})^{-k} + \frac{1}{2}\sum_{k\in\mathbb{Z}} x_k (-z^{\frac{1}{2}})^{-k} \qquad (5) \\ &= \frac{1}{2}(X(z^{\frac{1}{2}}) + X(-z^{\frac{1}{2}})). \end{aligned}$$

The operators $(\downarrow 2)$ und $(\uparrow 2)$, defined in the time domain, can also be defined in the z-domain via (4) and (5). We use for these operators the same symbols $(\downarrow 2)$ and $(\uparrow 2)$ as in the time domain. With this convention holds, for example, $(\downarrow 2)[x](n) = x(2 \cdot n)$ and $((\uparrow 2)X)(z) = X(z^2)$.

4.4.2 Sampling operators in the Fourier domain

By setting $z = e^{2\pi i\omega}$ in (4) and (5) we get the relations between x, u, and v in the Fourier (also called spectral domain or ω-domain):

$$U(\omega) = X(2\omega) \tag{6}$$

$$V(\omega) = \frac{1}{2}(X(\tfrac{\omega}{2}) + X(\tfrac{\omega}{2} + \tfrac{1}{2})) \tag{7}$$

By downsampling with $(\downarrow 2)$ one looses half of the data and, in general, also some information. In the Fourier domain this loss of data appears as mixing of different frequency components. As in the case of sampling CT-signals (see Section 3), the effect is known as aliasing.

The equation (7) says that the frequency component for $\frac{\omega}{2}$ and $\frac{\omega}{2} + \frac{1}{2}$ of the signal x are identified and summed up in the 2-downsampled signal v. This is illustrated in Figure 8.

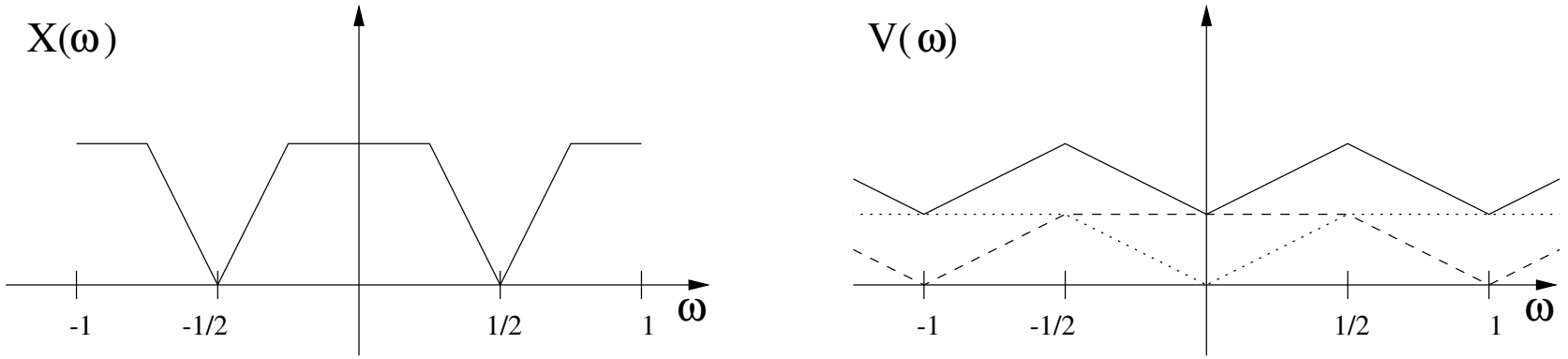


Figure 8: Downsampling in the frequency domain.

Upsampling with $(\uparrow 2)$ leads to the contrary effect, which is also known as imaging. Here the frequency component at ω of the original signal x is responsible for two frequency components at $\frac{\omega}{2}$ and $\frac{\omega}{2} + \frac{1}{2}$ of the 2-upsampled signal u. Equation (6) says that the frequency spectrum is compress by the factor of 2 (see Figure 9).

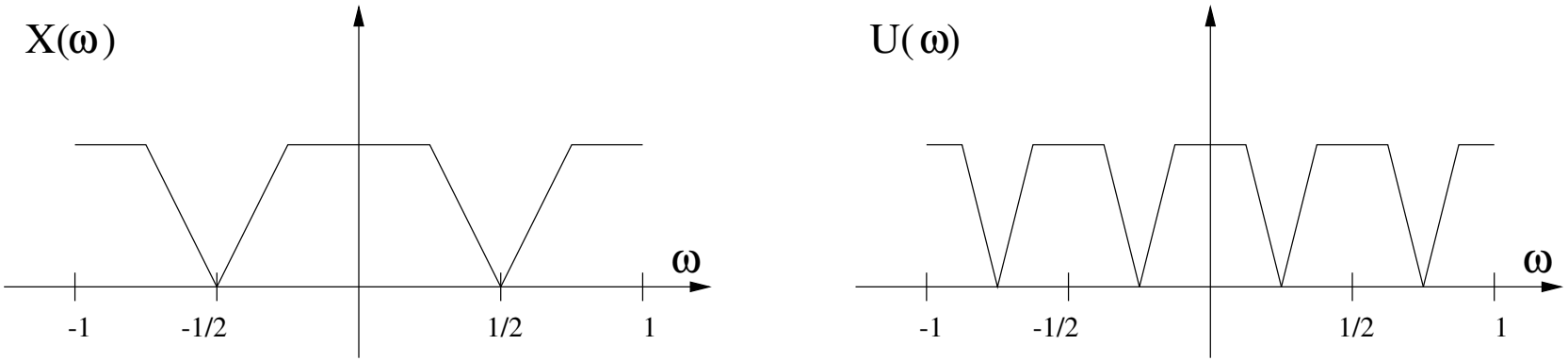


Figure 9: Upsampling in the frequency domain.

Due to the aliasing effects, the original signal x cannot in general be perfectly reconstructed from the downsampled signal $v = (\downarrow 2)[x]$. Similar to the CT-case some additional condition on the spectrum of x also guarantees perfect reconstruction.

Definition 4.10. A DT-signal x is called bandlimited by some $\Omega \in [0, \frac{1}{2}]$ if

$$X(\omega) = 0 \quad \text{for} \quad \Omega \leq |\omega| \leq \tfrac{1}{2}.$$

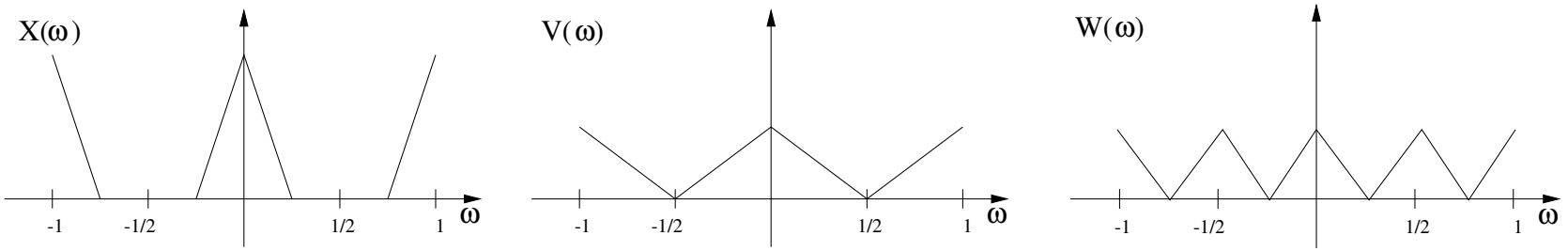


Figure 10: No aliasing if x is bandlimited by $\Omega \leq \frac{1}{4}$.

If the signal x is bandlimited by $\Omega = \frac{1}{4}$ then there is no aliasing and x can be reconstructed from v as follows (see also Figure 10):

(1) By $(\uparrow 2)$-upsampling one obtains the signal $w := (\uparrow 2)v$ with Fourier transform

$$W(\omega) = \frac{1}{2}(X(\omega) + X(\omega + \tfrac{1}{2})). \tag{8}$$

(2) Cut off the frequencies in W in the Fourier domain which resulted from the imaging-effect. This are all frequencies in

$$\left[-\frac{1}{2}, \frac{1}{2}\right] \setminus \left[-\frac{1}{4}, \frac{1}{4}\right]$$

and all translates by an integer number.

(3) Multiply the result by a factor of 2, then one recovers the Fourier transform X of the original signal x.

Note 4.11. In the time domain the reconstruction of the original signal from an downsampled version corresponds to an interpolation for the odd sample points of w (at these points w is zero). Therefore, in this context one also speaks form an interpolation filter which reconstructs the bandlimited signal x from w. Again the mathematical background is based on some version of Shannon's sampling theorem.

For a more detailed discussion we refer to Chapter 3 of [Strang/Nguyen].

Chapter 5: FIR Filters

In this chapter we follow in most part Chapter 8 of [Proakis/Manolakis].

In the design of frequency-selective filters, the desired filter characteristics are specified in the frequency domain in terms of the magnitude and the phase response. In the filter design process, one determines the coefficients of an FIR or IIR filter that closely approximates the desired frequency response specifications. The issue of which type of filter to design, FIR or IIR, depends on the nature of the problem and on the specifications of the desired frequency response.

There are many different theoretical methods for filter design which have been implemented and incorporated in numerous computer software programs such as MatLab. These programs allow the user to specify the desired filter characteristics and then computes the filter that best fits the desired design requirements.

This chapter allows us only to have a look at some aspects of filter design. We concentrate on the following kind of filters.

(1) We assume our filters to be causal, since this requirement is decisive for the physically realization of the filter.

(2) We only deal with FIR-filters, i.e., filters with a finite impulse response. Note that any FIR filter can be made causal by shifting the filter coefficients suitably.

(3) We want to restrict ourselves to the case of linear-phase filters which is an important property in view of applications. This property is equivalent to a certain symmetry or asymmetry condition on the filter coefficients as we will see later.

5.1 Causality and its Implications

We recall that a filter h is called causal if $h(n) = 0$ for $n < 0$. In this case the frequency response is given by

$$H(\omega) := \sum_{n=0}^{\infty} h(n) e^{-2\pi i n \omega}.$$

The Paley-Wiener Theorem gives necessary and sufficient conditions that a frequency response H must satisfy in order for the resulting filter to be causal. One version of this theorem can be stated as follows (see [Proakis/Manolakis, p. 616]):

Theorem 5.1 (Paley-Wiener). *If the filter $h \in \ell^1(\mathbb{Z})$ is causal then*

$$\int_0^1 |\log |H(\omega)|| d\omega < \infty.$$

Conversely, if $|H|$ is square integrable and if the above integral is finite, then one can associate with $|H|$ a phase response Φ such that the resulting filter with frequency response

$$H(\omega) = |H(\omega)| \cdot e^{i\Phi_h(\omega)}$$

is causal.

The Paley-Wiener Theorem has some immediate consequences:

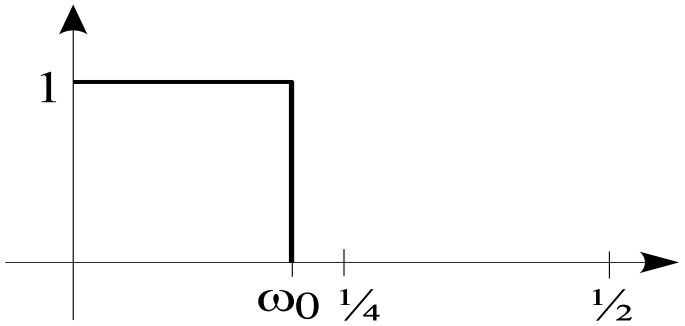
(1) For a causal filter h, the frequency response H cannot be zero except at a finite set of points in frequency, since the integral in the Paley-Wiener Theorem then becomes infinite. In particular, H cannot be zero over any finite band of frequencies.

(2) From (1) follows that there is no causal filter realizing an ideal lowpass, bandpass, or highpass filter.

(3) Since any FIR filter can be made causal by a shift (which amounts to a modulation in the frequency response and hence leaves the magnitude response unchanged), it follows from (2) that there is no FIR filter realizing an ideal filter.

5.2 The Ideal Lowpass Filter

We have just seen that an ideal lowpass filter with cut-off frequency ω_0 as illustrated below cannot be realized by a causal filter or FIR filter.



In the following, we investigate if there is at least a noncausal filter realizing the ideal lowpass filter. This might give us some ideas of how to approximate the ideal filter by FIR filters.

The frequency response of an ideal lowpass filter with cut-off frequency ω_0, $0 < \omega_0 \leq 1/2$, and with real coefficients is symmetric and 1-periodic. Hence it is specified on $[0, 1/2]$ by the formula

$$H_{\omega_0}(\omega) := \begin{cases} 1 & 0 \leq \omega \leq \omega_0 \\ 0 & \omega_0 < \omega \leq 1/2 \end{cases}$$

We now want to know, if there is a filter $h = (h(n))_{n\in\mathbb{Z}}$ with real coefficients whose frequency response equals H_{ω_0}, i.e.,

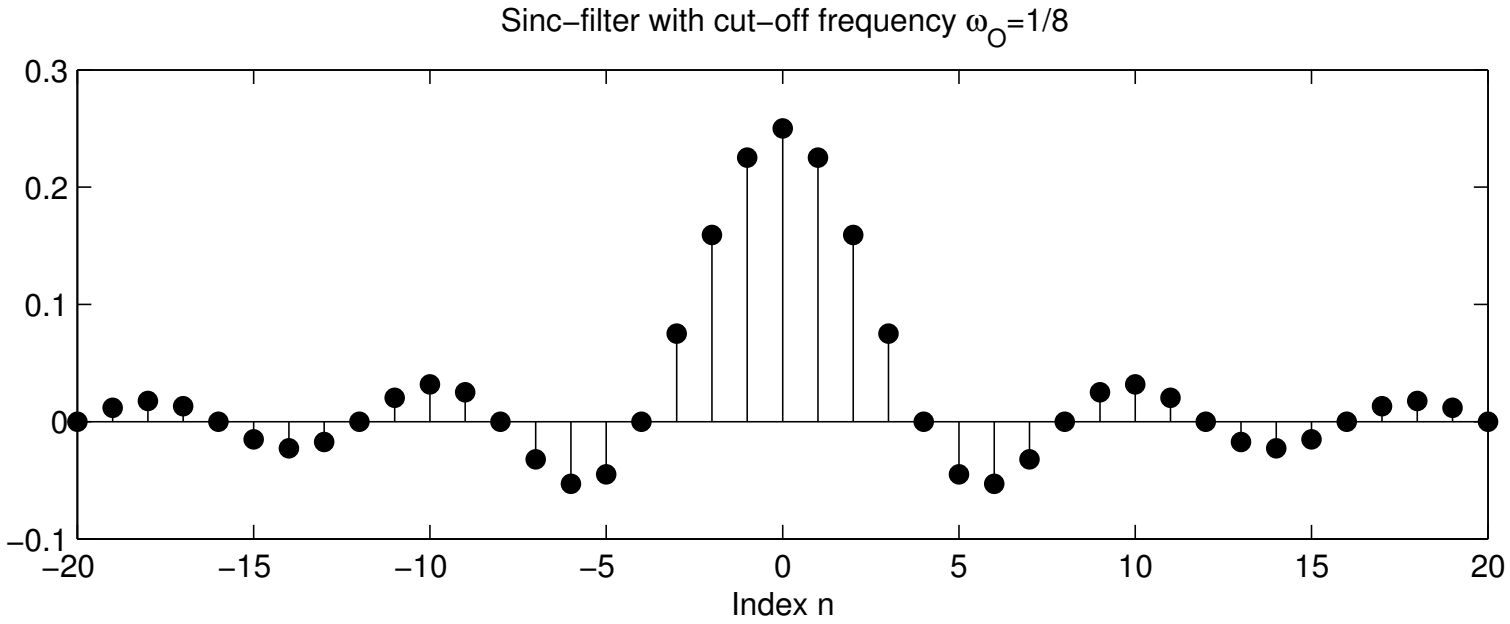
$$H(\omega) = \sum_{n\in\mathbb{Z}} h(n)e^{-2\pi i\omega n} = H_{\omega_0}(\omega) = H_{\omega_0}(-\omega)$$

for $\omega \in [0,1]$. From this we see that the filter coefficients $h(n)$ must be the Fourier coefficients of the Fourier series of H_{ω_0}, i.e.,

$$\begin{aligned} h(n) &= \int_0^1 H_{\omega_0}(\omega)e^{2\pi i\omega n}d\omega \\ &= \int_{-\omega_0}^{\omega_0} e^{2\pi i\omega n}d\omega \\ &= 2\omega_0\mathrm{sinc}\,(2\omega_0 n), \end{aligned}$$

where the computation follows as in Example 2.16. In other words, the filter corresponding to the ideal lowpass filter with cut-off frequency ω_0 is the sinc-function $t \mapsto 2\omega_0\mathrm{sinc}\,(2\omega_0 t)$ sampled at the integers $n \in \mathbb{Z}$.

The following figure shows the sinc-filter coefficients for the case $\omega_0 = 1/8$.



However, as mentioned before, there are several problems with the sinc-filter coefficients:

(1) The sinc-filter is not causal.

(2) The sinc-filter is not an FIR filter.

(3) Even worse, the filter is not even stable. For example, for $\omega_0 = 1/4$ one can show that the sequence $(h(2n))_{n\in\mathbb{Z}}$ behaves, up to a constant, like the sequence $(1/n)_{n\in\mathbb{Z}}$ which is clearly not in $\ell^1(\mathbb{Z})$.

5.3 Characteristics of Practical Frequency-Selective Filters

From our previous discussion follows that ideal filters are noncausal and hence physically not realizable for real-time signal processing applications. When an ideal filter is converted into a realizable FIR filter, the perfect behavior is degraded. For example, one can observe the following phenomena:

- Causality implies that the frequency response H of the filter cannot be zero, except at a finite set of points in the frequency domain. This leads to ripples in the passband and stopband,
- In addition, H cannot have an infinitely sharp cutoff from passband to stopband, i.e., H cannot drop from unity to zero abruptly and each of the sharp discontinuities is smeared into a range of frequencies.

In applications some degradations in the fequency response may be tolerable. For example, a small amount of ripples in the passband or in the stopband as shown in Figure 11 may be acceptable.

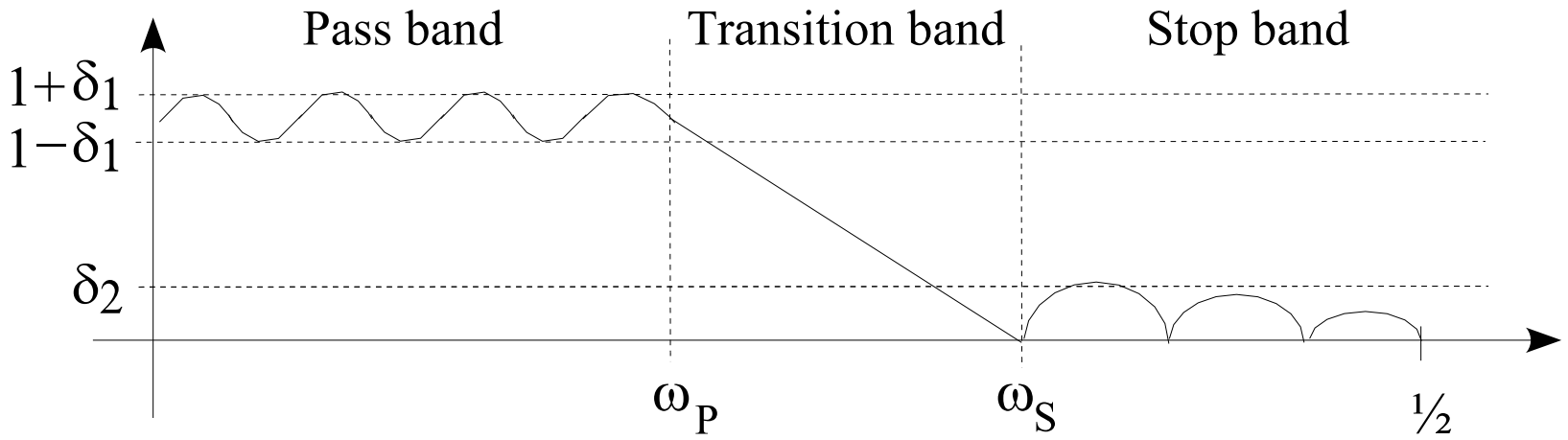


Figure 11: Magnitude characteristics of physically realizable lowpass filters.

The following filter characteristics are illustrated by Figure 11.

- The transition of the frequency response from passband to stopband defines the transition band of the filter.
- The band-edge frequency ω_p define the edge of the passband, while the frequency ω_s denotes the beginning of the stopband.
- The width of the transition band is $\omega_s - \omega_p$.
- The width of the passband is called bandwidth of the filter. For example, if the filter is lowpass with a passband edge frequency ω_p, its bandwidth is ω_p.
- If there are ripples in the passband of the filter, its value is denoted as δ_1, and the magnitude $|H|$ varies between the limits $1 \pm \delta_1$. The ripples in the stopband of the filter is denoted as δ_2.

To accommodate a large dynamic range in the graph of the magnitude response of a filter, it is common practice to use the decibel-scale as introduced in Section 3. Consequently, the ripples in the passband is $20 \log_{10} \delta_1$ decibels, and that in the stopband is $20 \log_{10} \delta_2$.

To summarize, in any filter design problem we can specify

(1) the maximum tolerable passband ripples,

(2) the maximum tolerable stopband ripples,

(3) the passband edge frequency ω_p, and

(4) the stopband edge frequency ω_s.

Based on these specifications, one is now interested in the construction of a filter h whose frequency response H best approximates these specifications. Of course, the degree to which H approximates the specifications depends in part on the number of non-zero filter coefficients.

5.4 Linear-Phase FIR Filters

Let h be a causal FIR filter of length M. Then the transfer function of h is given by

$$H(z) = \sum_{k=0}^{M-1} h(k) z^{-k}$$

which can be viewed as a polynomial of degree $M-1$ in the variable z^{-1}. By definition, the zeros of the filter h are the roots of this polynomial.

An FIR filter h has linear phase if the filter coefficients satisfy the following symmetry or antisymmetry conditions:

$$h(n) = \pm h(M-1-n), \quad n = 0, 1, \ldots, M-1.$$

From this follows that the transfer function of a causal, linear-phase FIR filter of length M is of the form:

$$H(z) = z^{-(M-1)/2} \left(h\left(\frac{M-1}{2}\right) + \sum_{k=0}^{(M-3)/2} h(k) \left[z^{(M-1-2k)/2} \pm z^{-(M-1-2k)/2} \right] \right)$$

if M is odd and

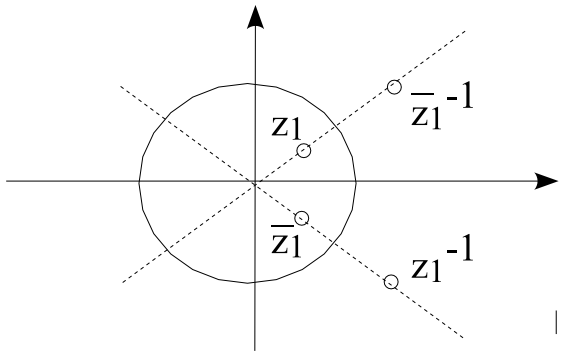
$$H(z) = z^{-(M-1)/2} \sum_{k=0}^{(M/2)-1} h(k) \left[z^{(M-1-2k)/2} \pm z^{-(M-1-2k)/2} \right]$$

if M is even.

Now, if we substitute z^{-1} for z in the transfer function of h and multiply both sides of the resulting equation by $z^{-(M-1)}$, we obtain

$$z^{-(M-1)}H(z^{-1}) = \pm H(z).$$

This result implies that the roots of the polynomial $H(z)$ are identical to the roots of the polynomial $H(z^{-1})$. In other words, if z_1 is a root of $H(z)$, then $1/z_1$ is also a root. Furthermore, since the filter coefficients $h(k)$ are assumed to be real, complex-valued roots must occur in complex-conjugate pairs. Hence, if z_1 is a complex-valued root, $\overline{z_1}$ is also a root. Therefore, $1/\overline{z_1}$ must also be a root. The next figure illustrates the symmetry that exists in the location of the zeros of a linear-phase FIR filter.



The frequency response of a linear-phase FIR filters are obtained by evaluating the formulas of the transfer function on the unit circle, i.e., by setting $z = e^{2\pi i\omega}$. From the formulas for $H(z)$ one then derives the following formulas for $H(\omega)$. If h is symmetric then $H(\omega)$ can be expressed as

$$H(\omega) = H_r(\omega)e^{-2\pi i\omega(M-1)/2}$$

where $H_r(\omega)$ is a real function of ω and can be expressed as

$$H_r(\omega) = h\left(\frac{M-1}{2}\right) + 2\sum_{k=0}^{(M-3)/2} h(k)\cos 2\pi\omega\left(\frac{M-1}{2} - k\right)$$

if M is odd and

$$H_r(\omega) = 2\sum_{k=0}^{(M/2)-1} h(k)\cos 2\pi\omega\left(\frac{M-1}{2} - k\right)$$

if M is even. The antisymmetric case is similar and left as an exercise.

The phase response of the filter for both M odd and M even is

$$\Phi_h(\omega) = \begin{cases} -\omega\left(\dfrac{M-1}{2}\right), & \text{if } H_r(\omega) > 0, \\ -\omega\left(\dfrac{M-1}{2}\right) + \dfrac{1}{2}, & \text{if } H_r(\omega) < 0. \end{cases}$$

Note that, for a symmetric filter h, the number of filter coefficients that specify the frequency response is $(M+1)/2$ when M is odd or $M/2$ when M is even.

5.5 Design of Linear-Phase FIR Filters Using Windows

In this section we assume that our filter specifications are given in form of some desired frequency response, say H_d. H_d could be, for example, the ideal lowpass, highpass or bandpass filter. Suppose $H_d \in L^2([0,1])$, then H_d has a Fourier series expansion (see Section 1). From this follows that the filter coefficients $h_d(n)$ of our desired filter h_d satisfy

$$h_d(n) = \int_0^1 H_d(\omega) e^{2\pi i \omega n} d\omega$$

and

$$H_d(\omega) = \sum_{n \in \mathbb{Z}} h_d(n) e^{-2\pi i \omega n}.$$

In general, the filter h_d is infinite in length and must be truncated at some point to get an FIR filter h of length $M \in \mathbb{N}$, say. By a time shift, we may assume that h_d was truncated at point $n = 0$ and $n = M - 1$, so that h is causal. Then truncation of h_d to a length M is equivalent to multiplying h_d by a discrete box-function w defined as

$$w(n) := \begin{cases} 1 & \text{for } n = 0, 1, \ldots M - 1, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Thus the filter coefficients of the FIR filter h becomes

$$h(n) = h_d(n)w(n), \quad n \in \mathbb{N}.$$

We now consider the effect of the window function on the desired frequency response H_d. Recall that multiplication of the window function w with h_d is equivalent to convolution H_d with W, where W denotes the frequency response of w. The convolution of H_d and W yields the frequency response of the FIR filter h. That is

$$H(\omega) = \int_0^1 H_d(\nu) W(\omega - \nu) d\nu.$$

The window function w has a great impact on the approximation quality of the resulting FIR filter $h = h_d \cdot w$. The convolution of H_d with W has the effect of smoothing H_d. When the window length M is increased then

(1) W becomes narrower,

(2) the smoothing provided by W is reduced, and

(3) the transition band in the frequency response of h becomes smaller.

The window technique is best described in terms of a specific example which we will consider now.

5.5.1 Windowing with the Box-Function

Suppose that we want to design a symmetric causal lowpass linear-phase FIR filter having a desired symmetric frequency response

$$H_d(\omega) := \begin{cases} e^{-2\pi i\omega(M-1)/2} & 0 \leq \omega \leq \omega_0 \\ 0 & \omega_0 < \omega \leq 1/2 \end{cases}$$

for some cut-off frequency ω_0 with $0 < \omega_0 \leq 1/4$. A delay of $(M-1)/2$ units is incorporated into H_d in anticipation of forcing the filter to be of length M. As in Section 2 one can prove that the filter coefficients of h_d are given by

$$h_d(n) = 2\omega_0 \mathrm{sinc}\left(2\omega_0\left(n - \frac{M-1}{2}\right)\right).$$

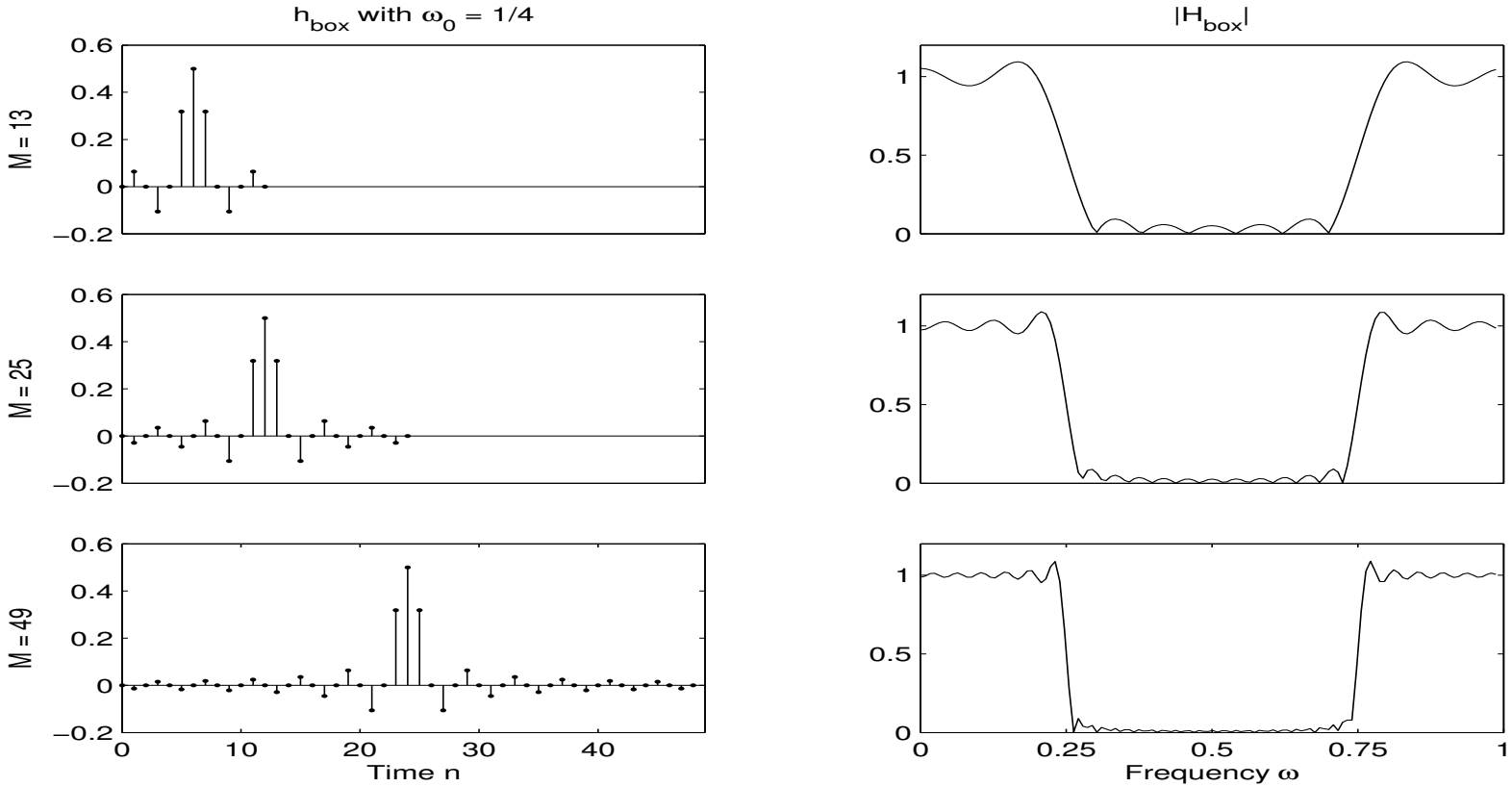
Clearly, h_d is noncausal and infinite in length. If we multiply h_d by the discrete box-function w of length M as given in Equation (9), we obtain an FIR filter h_{box} given by

$$h_{\mathrm{box}}(n) = 2\omega_0 \mathrm{sinc}\left(2\omega_0\left(n - \frac{M-1}{2}\right)\right), \quad 0 \leq n \leq M-1,$$

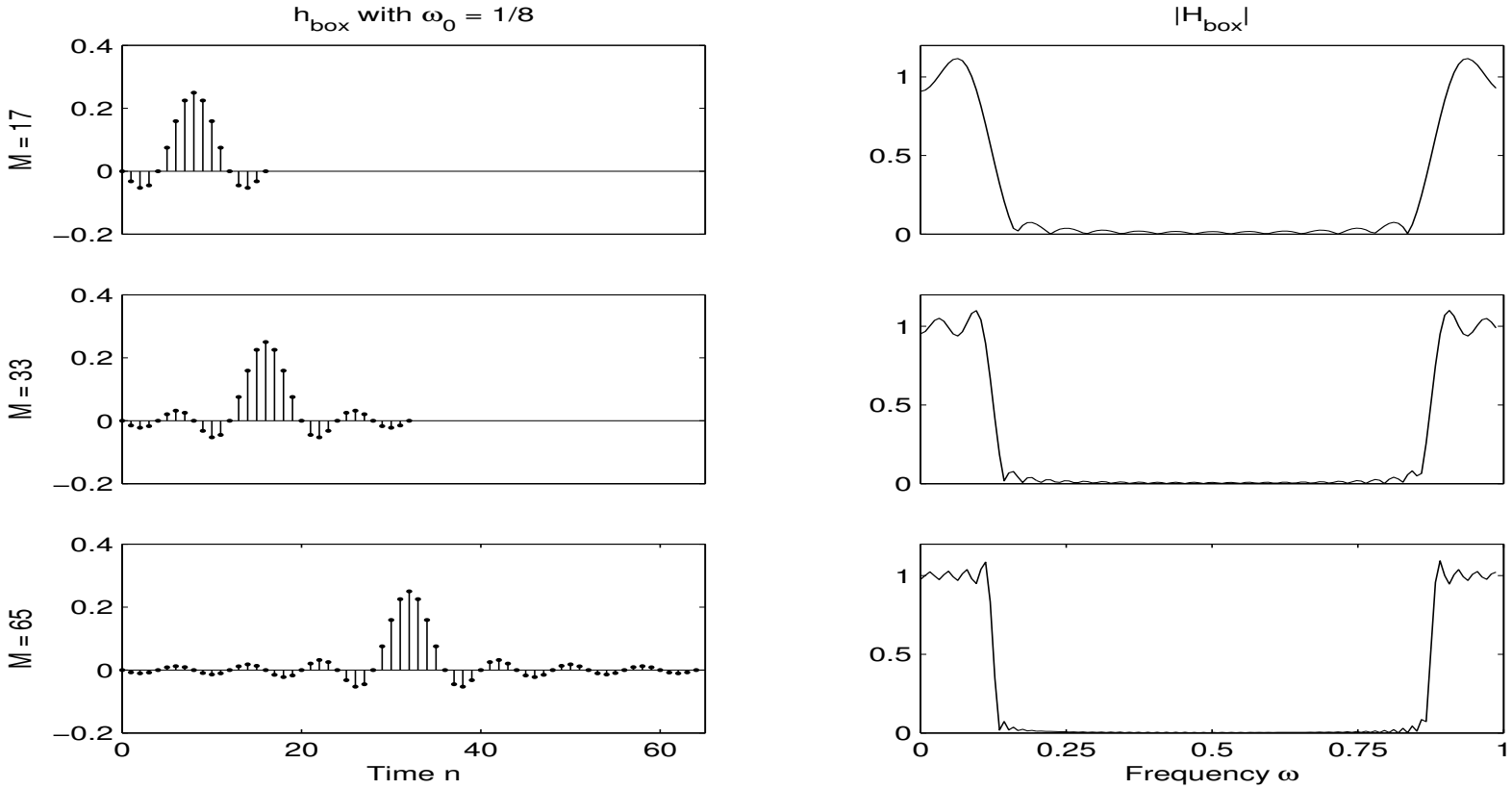
such that

- h is of length M,
- it is symmetric and has therefore linear phase, and
- it approximates the ideal lowpass filter with cut-off frequency ω_0.

The following figure illustrate the approximation behavior of h_{box} in the frequency domain for $\omega_0 = \frac{1}{4}$ and $M = 13$, $M = 25$, and $M = 49$.



The following figure illustrate the approximation behavior of h_{box} in the frequency domain for $\omega_0 = \frac{1}{8}$ and $M = 17$, $M = 33$, and $M = 65$. Note that the pass band of this lowpass filter is just half the size of the one before.



Note that there is a significant oscillatory overshoot of H_{box} at $\omega = \omega_0$, independent of the value M. As M increases, the oscillations become more rapid, but the size of the ripples remains the same. One can show that for $M \longrightarrow \infty$, the oscillations converge to the point of discontinuity $\omega = \omega_0$, but their amplitude does not go to zero. This oscillatory behavior of the approximation h_{box} to the ideal frequency response H_d is known as Gibbs phenomenon. This nonuniform convergence phenomenon is identical to the study of the convergence of Fourier series and manifests itself in the design of FIR filters.

Indeed, from our discussion above it is not hard to see that the demodulated frequency response G defined by $G(\omega) := e^{2\pi i\omega(M-1)/2}H(\omega)$ is given by the formula

$$G(\omega) = \sum_{n=-N}^{N} 2\omega_0 \cdot \operatorname{sinc}(2\omega_0 n)e^{-2\pi in\omega}$$

when $M = 2N + 1$. This is nothing else then the N-truncated Fourier series of the ideal frequency response H_d with cut-off frequency ω_0. It is a well known fact from the theory of Fourier series that the N-truncated Fourier series converge in the $L^2([0, 1])$-norm — also referred to as mean-square convergence — to the periodic limit function (in this case H_d). However, mean-square convergence does not imply pointwise convergence or even uniform convergence. One phenomenon arising from this fact is the Gibbs phenomenon.

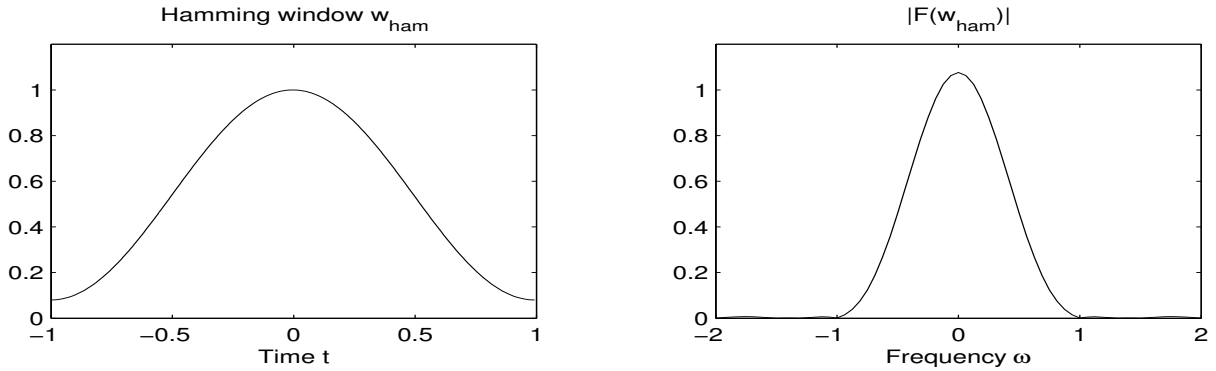
For details on this phenomenon we refer, for example, to [Folland] (mathematical point of view) or to the books [Oppenheim/Schafer] and [Proakis/Manolakis] (filter design point of view).

5.5.2 Windowing with the Hamming Window

The ringing and ripples in the frequency response, especially the Gibbs effects near the band edges, can be soften by using a window function that contains a taper and decay toward zero gradually, instead of abruptly, as it occurs in a rectangular window. For our next example, the underlying CT-window function is the Hamming window w_{ham} defined by the formula

$$w_{\mathrm{ham}}(t) = 0.54 - 0.46\cos(2\pi t).$$

Together with its CT-Fourier transform it is shown in the following figure.

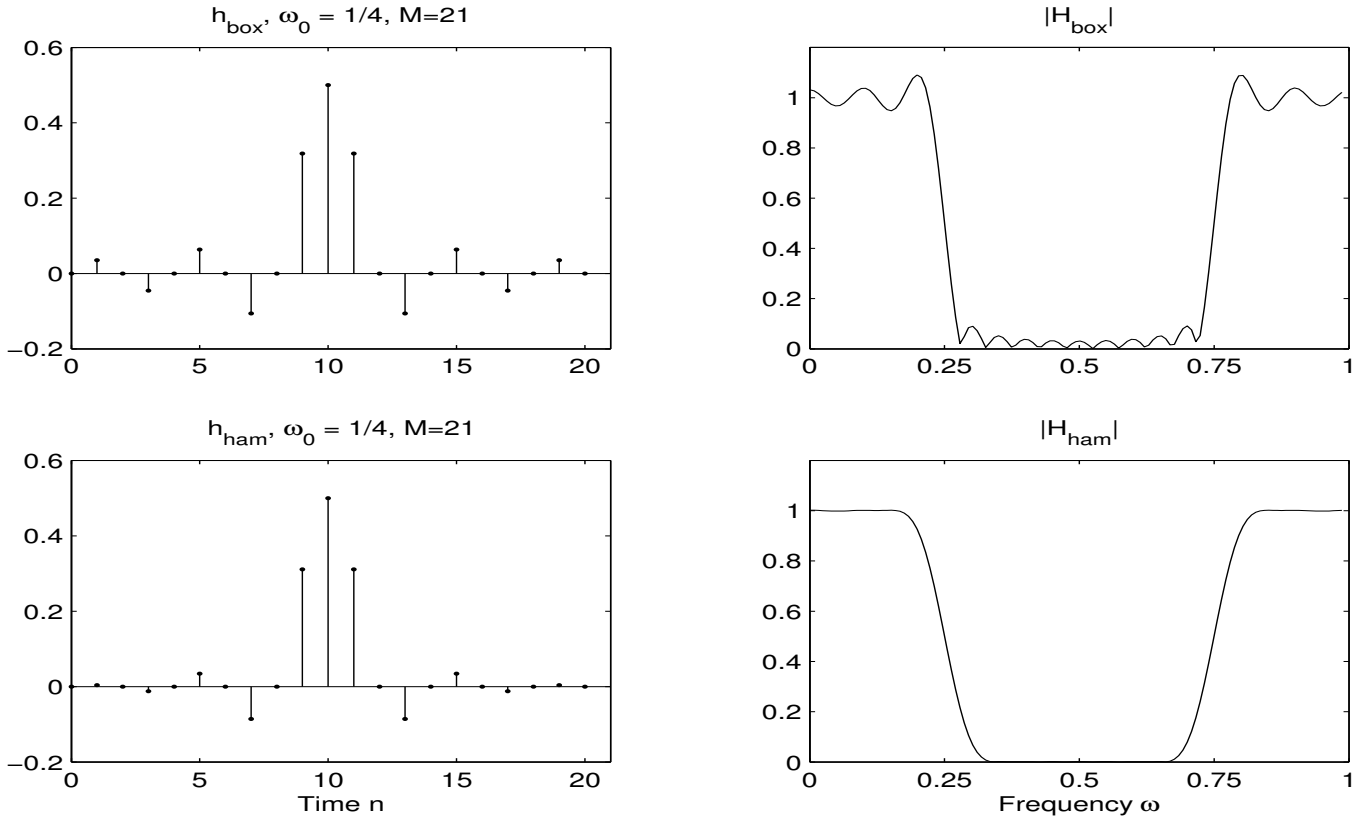


If one multiplies the filter coefficients of our desired filter h_d with suitable samples of the Hamming window, ones gets a smoothed FIR filter denoted by h_{ham}:

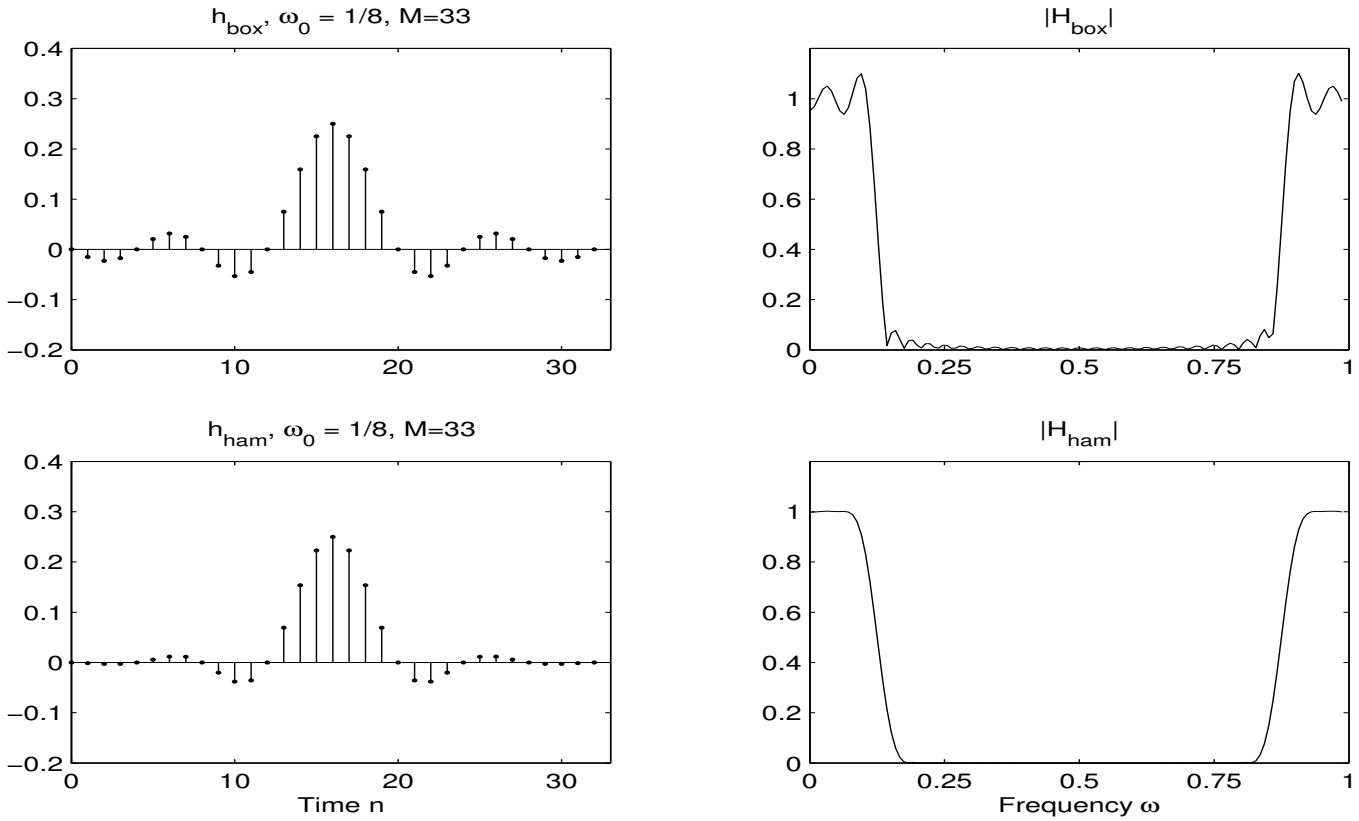
$$h_{\mathrm{ham}}(n) = h_d(n) \cdot w_{\mathrm{ham}}\left(n - \frac{M-1}{2}\right), \quad 0 \leq n \leq M-1.$$

The smoothing effect of windowing the filter coefficients with the Hamming window is shown in the following figures for various M and ω_0.

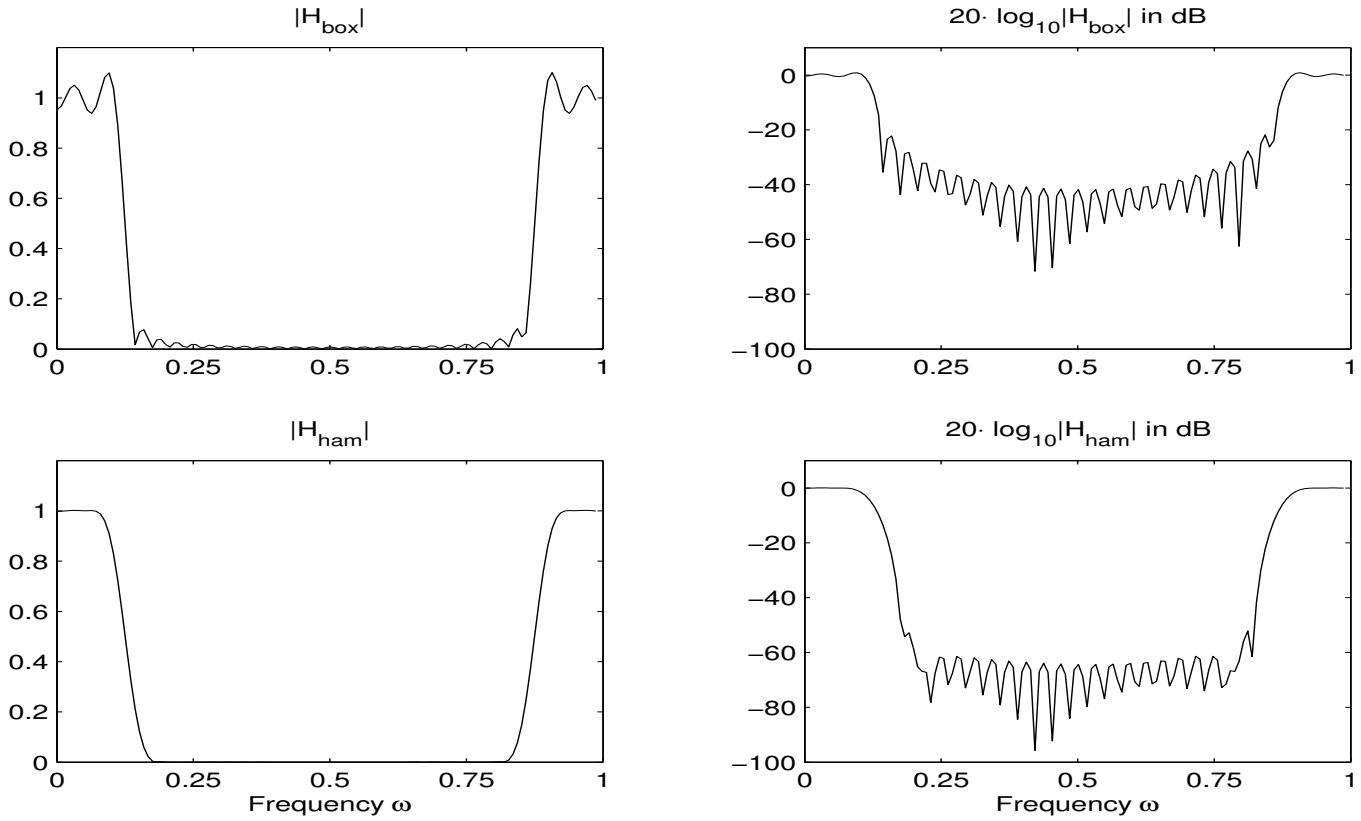
In this example we consider the desired filter h_d with $\omega_0 = 1/4$. The figure shows two approximations of h_d by FIR filters of length $M = 21$, the first FIR filter is h_{box}, the second h_{ham}.



Here, the desired filter h_d with $\omega_0 = 1/8$ is approximated by FIR filters h_{box} and h_{ham}. This time the filter length is $M = 33$.



The ripple effects in the passband or in the stopband can be better seen in the decibel-scale. The next figure shows the magnitude response of the previous example ($\omega_0 = 1/8$, $M = 33$) in the normal and in the decibel-scale.



As illustrated by the figures, windowing by the Hamming window does indeed decrease the ripple effects in the passband and stopband — however at the expense of an increase in the width of the transition band of the filter.

In the literature beside the Hamming window many other window-functions have been suggested which further decrease the ringing effects or lead to other improvements. We mention some of the most commom window functions:

- Barlett (triangle) window
- Blackman window
- Kaiser window
- Hanning window

Historically, the design method on the use of windows to truncate the desired filter h_d and obtaining an approximation of the desired spectral shaping, was the first method proposed for designing linear-phase FIR filters.

The major disadvantage of the window design method is the lack of precise control of the critical frequencies, such as ω_p and ω_s, in the design of a lowpass FIR filter. The values ω_p and ω_s, in general, depend of the type of window and the filter length. Also, the size of the ripples, δ_1 and δ_2, cannot very well be controlled.

Many other filter design methods have been suggested such as th frequency sampling method or the Chebyshev approximation method. The latter one provides total control or the filter specifications in terms of ω_p, ω_s, δ_1, and δ_2.

For further details we refer to [Proakis/Manolakis].

5.6 Bandpass Filter from Lowpass Filter

In the last sections, we have only discussed lowpass filters. However, the same approximation properties and design principles also apply to arbitrary bandpass filters. Actually, there is a straightforward method to construct bandpass filters from lowpass filters having the same quality characteristics.

The underlying idea is very simple and based on the fact that translation in the Fourier domain corresponds to modulation in the time domain. This is expressed by (3) of Theorem 2.20:

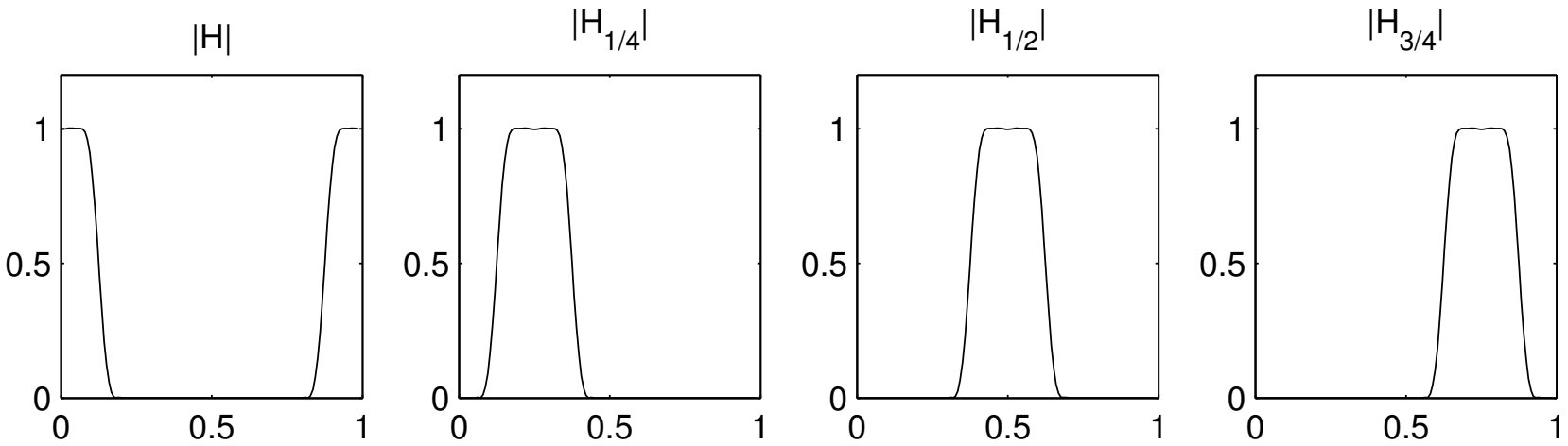
$$\widehat{E_\lambda x}(\omega) = \hat{x}(\omega + \lambda),$$

with $x \in \ell^1(\mathbb{Z})$. Recall that E_λ denotes the modulation operator (see Example 3.6):

$$E_\lambda[x](n) := e^{-2\pi i \lambda n} x(n), \quad x \in \ell^1(\mathbb{Z}),\ n \in \mathbb{Z}.$$

for some $\lambda \in \mathbb{R}$.

Therefore, applying the modulation operator E_λ on the filter coefficients of some filter h leads to a shift by λ in the frequency response H. This is illustrated by the next figure, where the frequency response of the modulated filter $E_\lambda[h]$ is denoted by H_λ.



The problem, however, is that $E_\lambda[h]$ has in general complex filter coefficients even if h has real coefficients. One possible trick is to consider the real part (or imaginary part) of the modulated filter. For example, let

$$g_\lambda := \mathrm{re}(E_\lambda[h]),$$

then

$$g_\lambda(n) = \cos(2\pi\lambda n)h(n) = \frac{1}{2}(e^{2\pi i\lambda n} + e^{-2\pi i\lambda n})h(n).$$

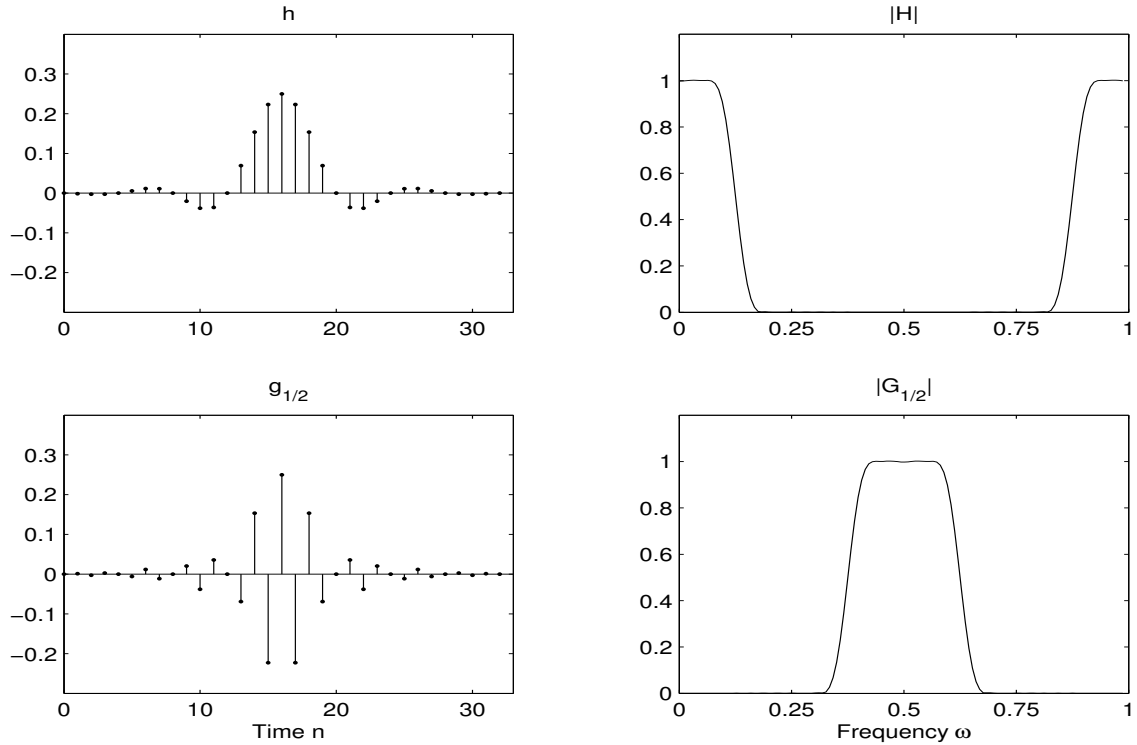
One also says that g_λ arises from cosine-modulation from h.

We now investigate which effect cosine-modulation has in the frequency domain. Let G_λ denote the frequency response of g_λ. Then

$$\begin{aligned} G_\lambda(\omega) &= \sum_{n\in\mathbb{Z}} g(n)e^{-2\pi i\omega n} \\ &= \sum_{n\in\mathbb{Z}} h(n)\cdot\frac{1}{2}(e^{2\pi i\lambda n}-e^{-2\pi i\lambda n})e^{-2\pi i\omega n} \\ &= \frac{1}{2}\sum_{n\in\mathbb{Z}} h(n)e^{-2\pi i(\omega-\lambda)n}+\frac{1}{2}\sum_{n\in\mathbb{Z}} h(n)e^{-2\pi i(\omega+\lambda)n} \\ &= \frac{1}{2}(H(\omega+\lambda)+H(\omega-\lambda)) \end{aligned}$$

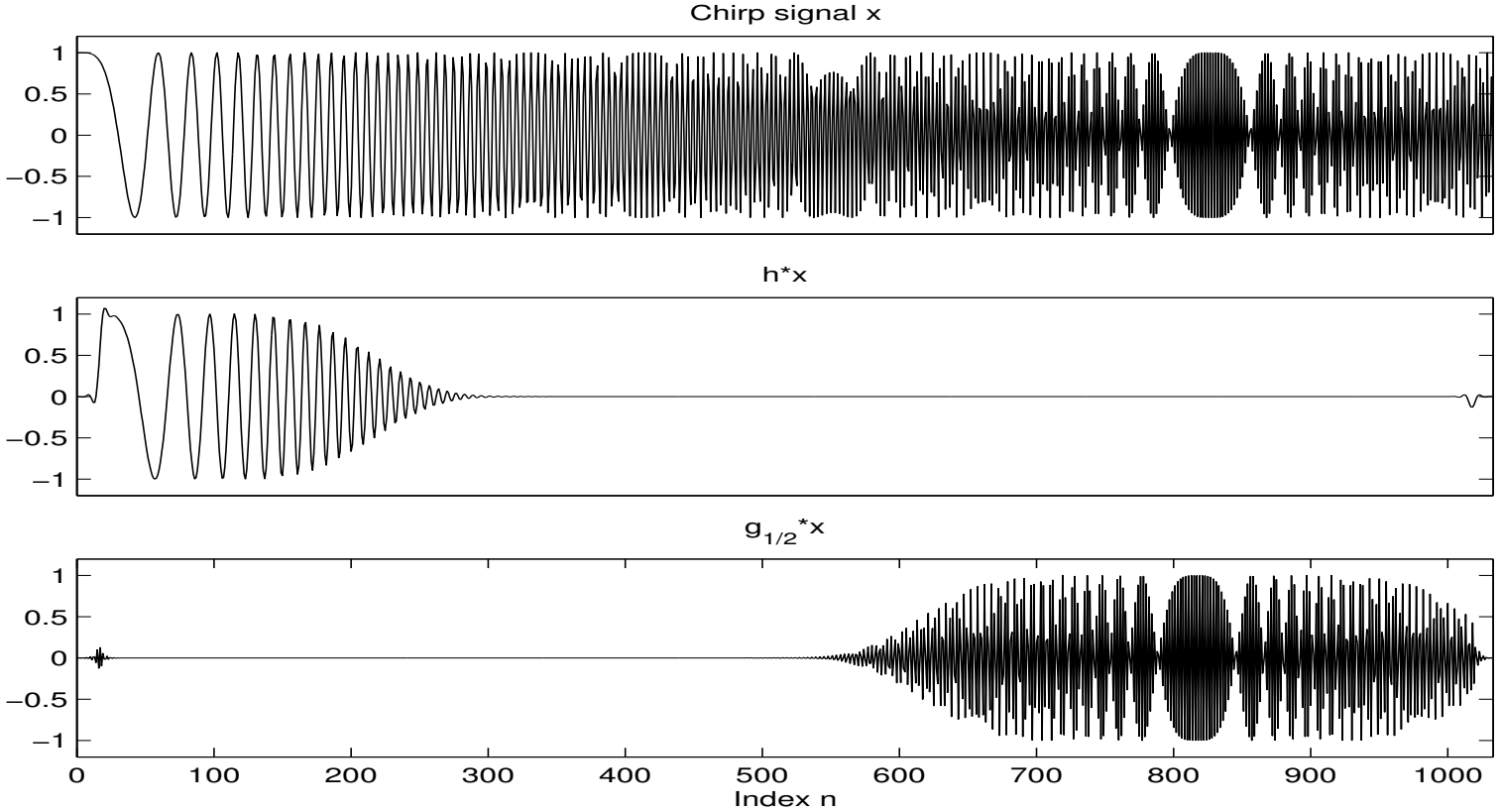
We give some examples to illustrate this effect. To distinguish the cut-off frequencies of different filters, we write the filter as argument of the cut-off frequency.

Example 5.2. Let h be an ideal lowpass filter with cut-off frequency $\omega_0(h) \leq 1/4$ and $\lambda = 1/2$. Then $g_{1/2}$ is the ideal highpass with cut-off frequency $\omega_0(g_{1/2}) = 1/2 - \omega_0(h)$.

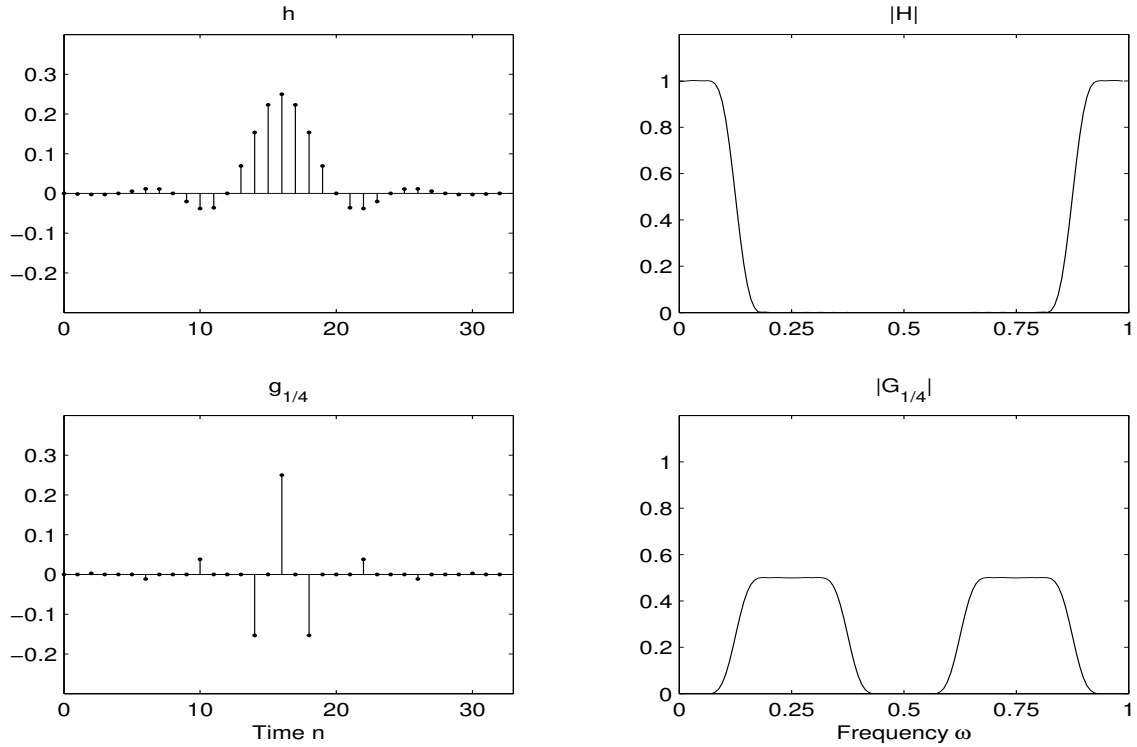


Note that in this case the cosine-modulation amounts to $g_{1/2}(n) = (-1)^n h(n)$.

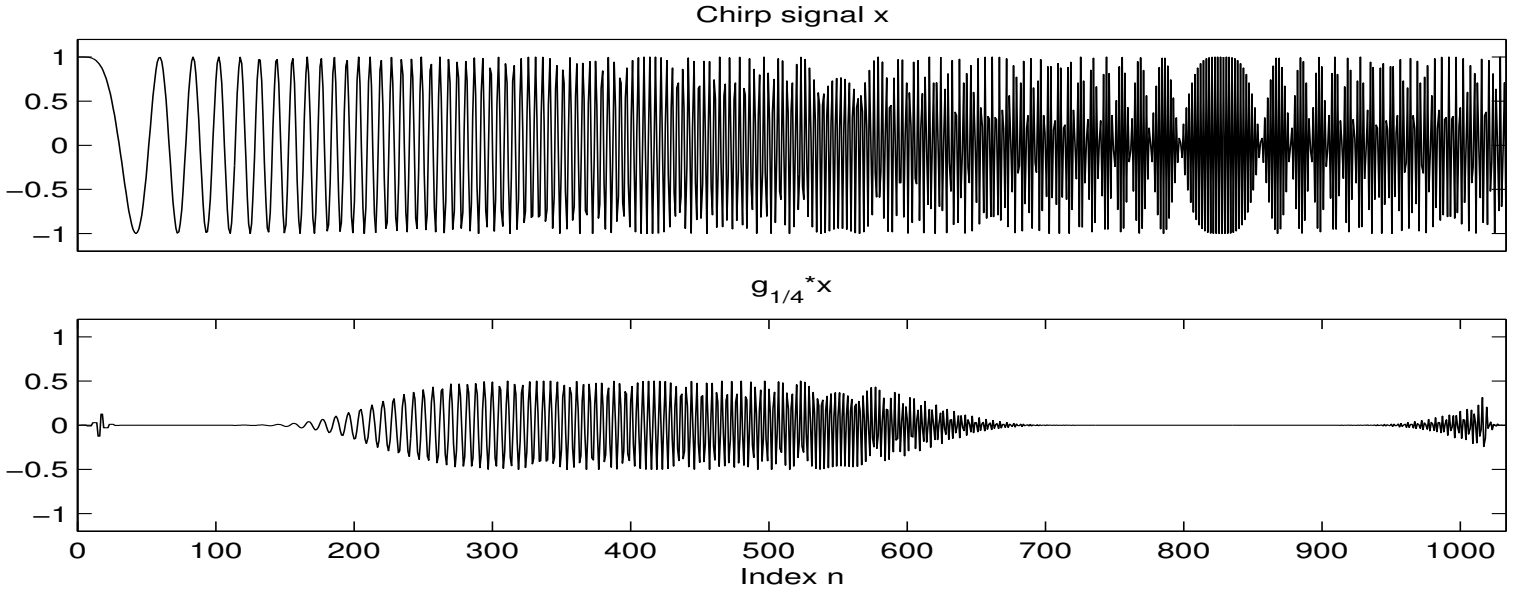
The lowpass filter h and the highpass filter $g_{1/2}$ have been applied to some chirp signal x. Note that due to aliasing the frequency decreases at the end of the signal.



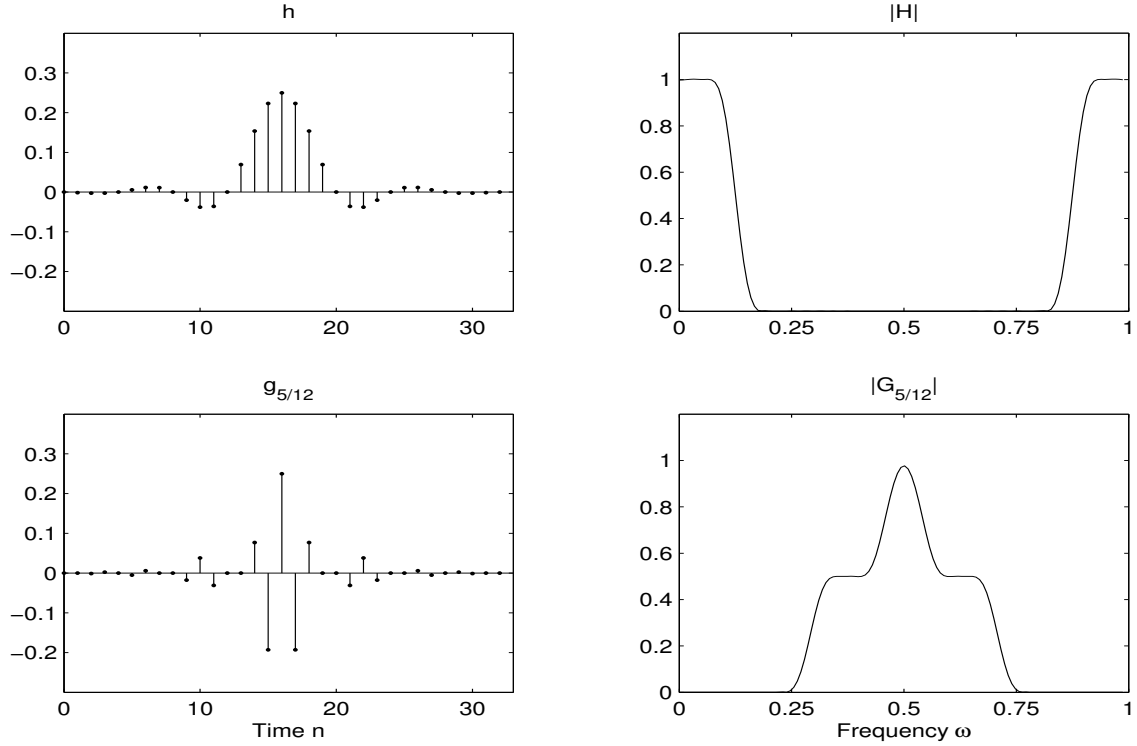
Example 5.3. Let h be an ideal lowpass filter with cut-off frequency $\omega_0(h) \leq 1/8$ and $\lambda = 1/4$ Then $g_{1/4}$ is (up to a factor $1/2$) an ideal bandpass filter with $\omega_0(g_{1/4}) = 1/4 - \omega_0(h)$ and $\omega_1(g_{1/4}) = 1/4 + \omega_0(h)$.



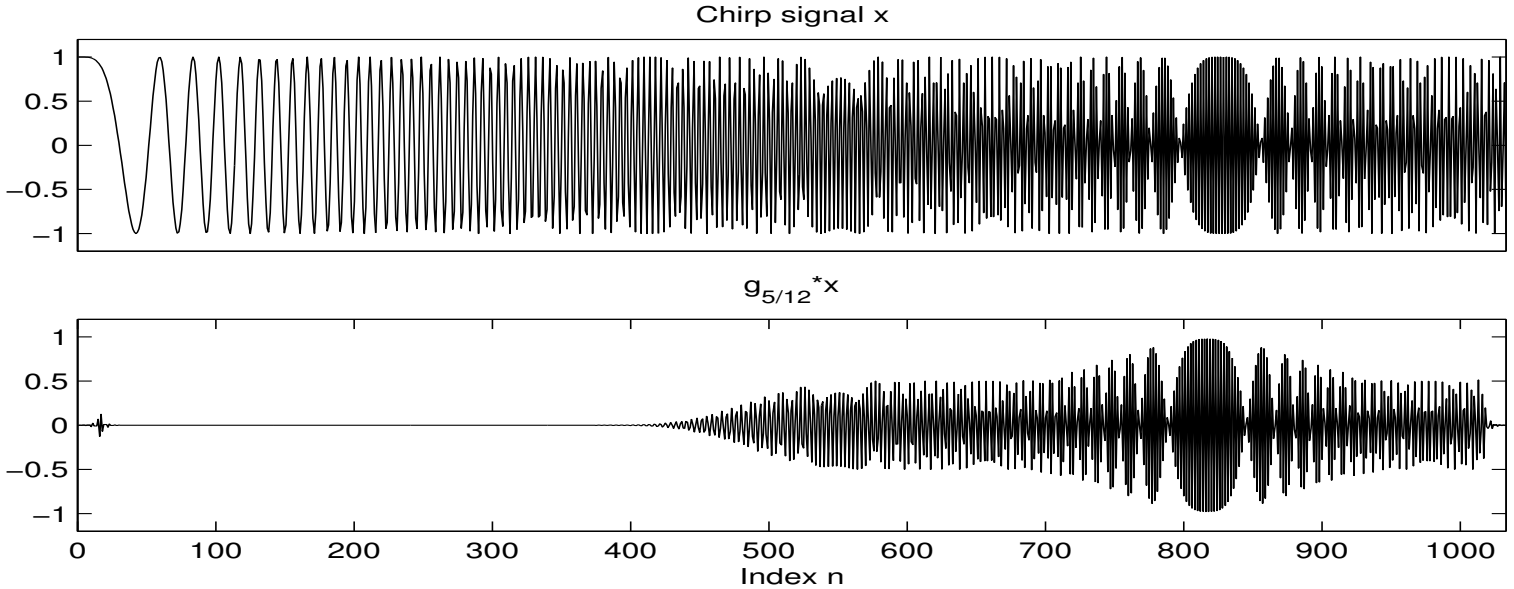
The bandpass filter $g_{1/4}$ has been applied to the same chirp signal x as in the last example.



Example 5.4. The following figure shows some highpass filter which let high frequencies pass through, attenuates the frequencies in the middle and cuts off low frequencies.



Again we see the effect of the filter $g_{5/12}$ on the chirp signal x.



Note 5.5. In case h ia a causal, linear-phase FIR filter of length M a slight modification of above cosine-moulation

$$g_\lambda(n) = \cos\left(2\pi\lambda\left(n - \frac{M-1}{2}\right)\right) h(n)$$

again yields a causal, linear-phase FIR filter g_λ of the same length.

Modulation of filters play an important role in the design of filter banks. Here one starts with a single "prototype" filter and generates a whole set of other filters by suitable modulations of this prototype filter. Ideally

- each of these filters covers a band in the spectral domain,
- the bands of different filters do not overlap, and
- the union of the bands cover the whole spectral domain.

In view of the perfect reconstruction condition of a filter bank and an efficient simultaneous evaluation of all filters, the cosine-modulation has to be carefully designed (amounting in choosing suitable phases in the cosines).

For further details on this topic we refer to [Burrus/Gopinath/Guo].

Chapter 6: Windowed Fourier Transform (WFT)

The Fourier transform $\hat{f}$ of a signal $f \in L^2(\mathbb{R})$ describes the frequency content of the signal. The time dependent f is transformed into a frequency dependent signal

$$\hat{f}(\omega) := \int_{-\infty}^{\infty} f(t) e^{-2\pi i \omega t} dt.$$

Intuitively, the signal f is analyzed by means of the exponential functions

$$\mathbb{R} \to \mathbb{C}, \quad t \mapsto e^{2\pi i \omega t}$$

of different frequencies $\omega \in \mathbb{R}$. These analysis functions are periodic and not localized w.r.t. time. In this sense they are well suited to analyze the in general non-periodic signal $f \in L^2(\mathbb{R})$.

Summarizing, the Fourier transform can be interpreted as follows:

- The Fourier transform hides the information about time (in the phase). It tells which "notes" (frequencies) are played, but it does not tell when these notes are played.
- Sudden changes und local variations of the signal as well as the beginning and the end of events cannot be detected by the Fourier transform.
- The frequency information is always averaged over the entire time interval.
- Local phenomena of the signal become global phenomena in the Fourier transform.
- Contrary, small changes in the phase of the Fourier transform can have considerable effects in the time domain.

To remedy the drawbacks of the Fourier transform Dennis Gabor introduced in the year 1946 the modified Fourier transform, now known as windowed Fourier transform or simply WFT. This transform is a compromise between a time- and a frequency-based representation of the signal. The WFT does not only tell which frequencies are "contained" in the signal but also at which points of times or, to be more precise, in which time intervals these frequencies appear.

6.1 Defintion of the WFT

For a given signal $f \in L^2(\mathbb{R})$ we want to find a transform, $\tilde{f}(\omega, t)$ which exhibits the frequency distribution at the point t. The basic idea for the design of such a transform comes from the way the human auditory system performs a realtime analysis of audio signals:

- For the analysis only a small section of the signal which lies directly behind the present point of time is used for the analysis.
- The further the points of time in this section lie in the past, the less they will be considered.
- Similar, the audio signal at "current" points of time is not yet percieved very well and therefore also less considered.

Mathematically, this weighting of the signal is modeled by multiplying the signal with a window function. The window function can be thought of as a bell-shaped function which localizes around $t = 0$.

If $f \in L^2(\mathbb{R})$ is a signal and $g : \mathbb{R} \to \mathbb{C}$ is a window function, then the function $f_{g,t}$ localized at point t is defined by

$$f_{g,t}(u) := f(t,u) := \bar{g}(u-t)f(u).$$

If g is known from the context, then one also writes just f_t instead of $f_{g,t}$. For the moment, we just assume $g \in L^2(\mathbb{R})$ and $\|g\|_2 \neq 0$.

Definition 6.1. Let $g \in L^2(\mathbb{R})$, $\|g\|_2 \neq 0$, be a window function. Then for a signal $f \in L^2(\mathbb{R})$ the transform

$$\tilde{f}(\omega, t) := \int_{-\infty}^{\infty} f_t(u) e^{-2\pi i \omega u} du = \int_{-\infty}^{\infty} \bar{g}(u-t) f(u) e^{-2\pi i \omega u} du$$

is called the windowed Fourier transform (WFT) of f (with respect to g).

If we define the function $g_{\omega,t} : \mathbb{R} \longrightarrow \mathbb{C}$ by

$$g_{\omega,t}(u) := e^{2\pi i \omega u} g(u-t), \quad u \in \mathbb{R},$$

then $\|g_{\omega,t}\| = \|g\|$, $g_{\omega,t} \in L^2(\mathbb{R})$, and the WFT of f can be written as

$$\tilde{f}(\omega, t) = \langle f | g_{\omega,t} \rangle.$$

Intuitively, the WFT can be thought of as follows:

- The function $g_{\omega,t}$ represents a "musical note" of frequency ω which oscillates within the translated window given by $u \mapsto |g(u-t)|$.
- The inner product $\langle f | g_{\omega,t} \rangle$ measures the correlation between the signal f and the musical note $g_{\omega,t}$. If f and $g_{\omega,t}$ have a similar course in time within the window, the inner product $\langle f | g_{\omega,t} \rangle$ has a large absolute value and vice versa.
- The signal

$$u \mapsto \langle f | g_{\omega,t} \rangle g_{\omega,t}(u)$$

can be considered as the "projection" of the signal f in direction of the musical note $g_{\omega,t}$.

6.2 Examples

6.2.1 Window Functions

As an example, Figure 12 shows three different window functions with their respective spectral energy density.

Using the box-window g, the signal f is preserved on the support $\mathrm{supp}(g) = [-0.5, 0.5]$ and is set to zero outside this support. The box-window has a major drawbacks. The localized functions $f_{g,t}$ defined by

$$f_{g,t}(u) := f(t,u) := \bar{g}(u-t)f(u)$$

have in general considerable discontinuities at the cuts which cause artefacts and interferencies seen in the frequency domain.

We remind that short-time events and abrupt changes in the signal lead to high-frequency phenomena in the Fourier domain.

The Hanning window g defined by

$$g(u) := \begin{cases} 1 + \cos(\pi u) & \text{falls} \quad -1 \leq u \leq 1 \\ 0 & \text{sonst} \end{cases}$$

was already introduced in Example 2.26. The Hanning window has also compact support, however, it abates smoothly when reaching the support boundaries. Therefore, the above mentioned artefacts in the Fourier transform of the windowed signal are softened. This is also illustrated by comparing Figure 13 with Figure 14.

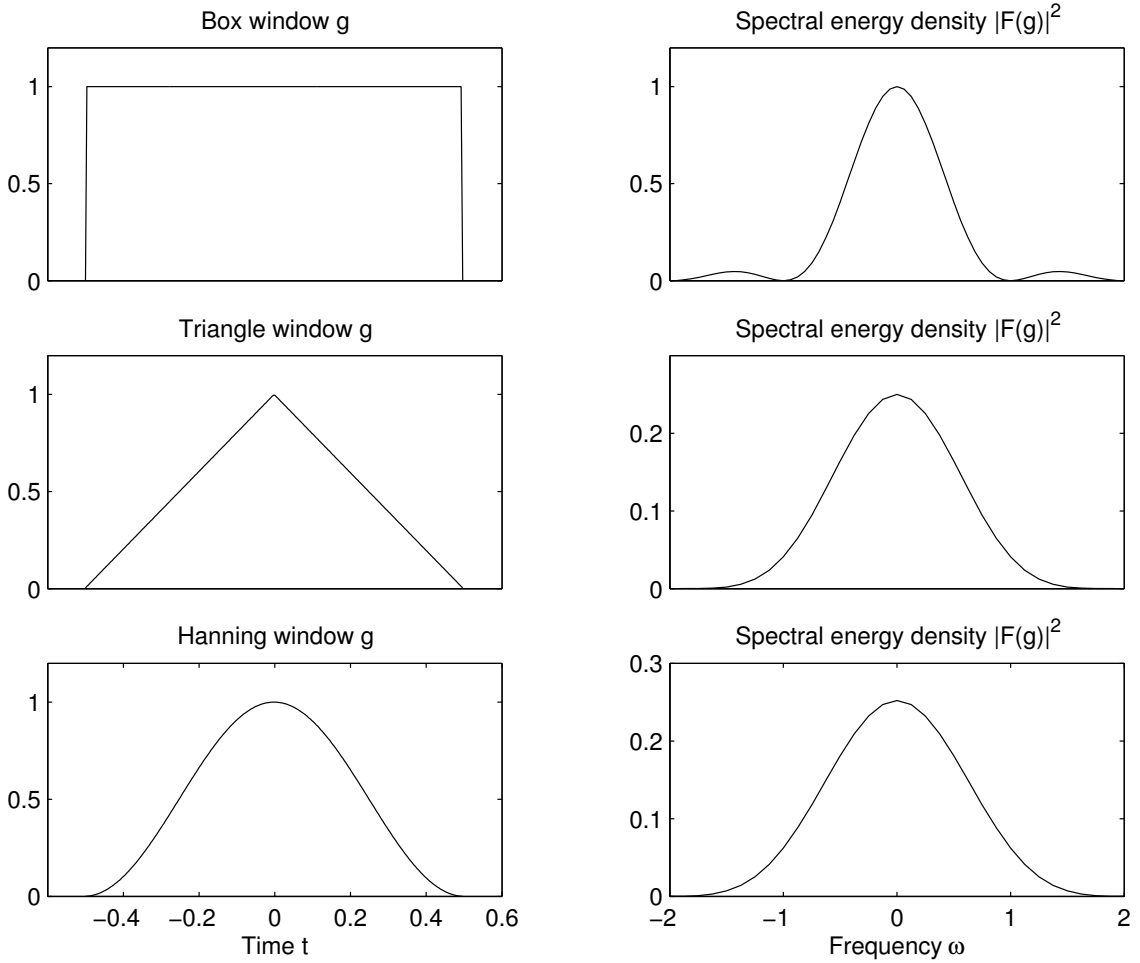


Figure 12: Window functions.

There are many other window functions used in the WFT. For further information we refer to the MatLab handbook.

Finally, we want to mention, that the window function does not necessarily have to have a compact support. One example is the Gauss window shown in Example 2.15. The choice of the "right" window function constitutes often a difficult problem and depends on the respective application.

6.2.2 WFT of a Chirp Signal

Figure 13 shows the time-frequency representation of the chirp signal f defined by

$$f(t) = \sin(400\pi t^2), \quad t \in \mathbb{R}$$

using the WFT w.r.t. the Hanning window. The values

$$|\tilde{f}(\omega, t)| = |\langle f | g_{\omega,t} \rangle|$$

at (t, ω) is represented by different gray levels, which are lighter for small values $|\tilde{f}(\omega, t)|$ and darker for large values $|\tilde{f}(\omega, t)|$.

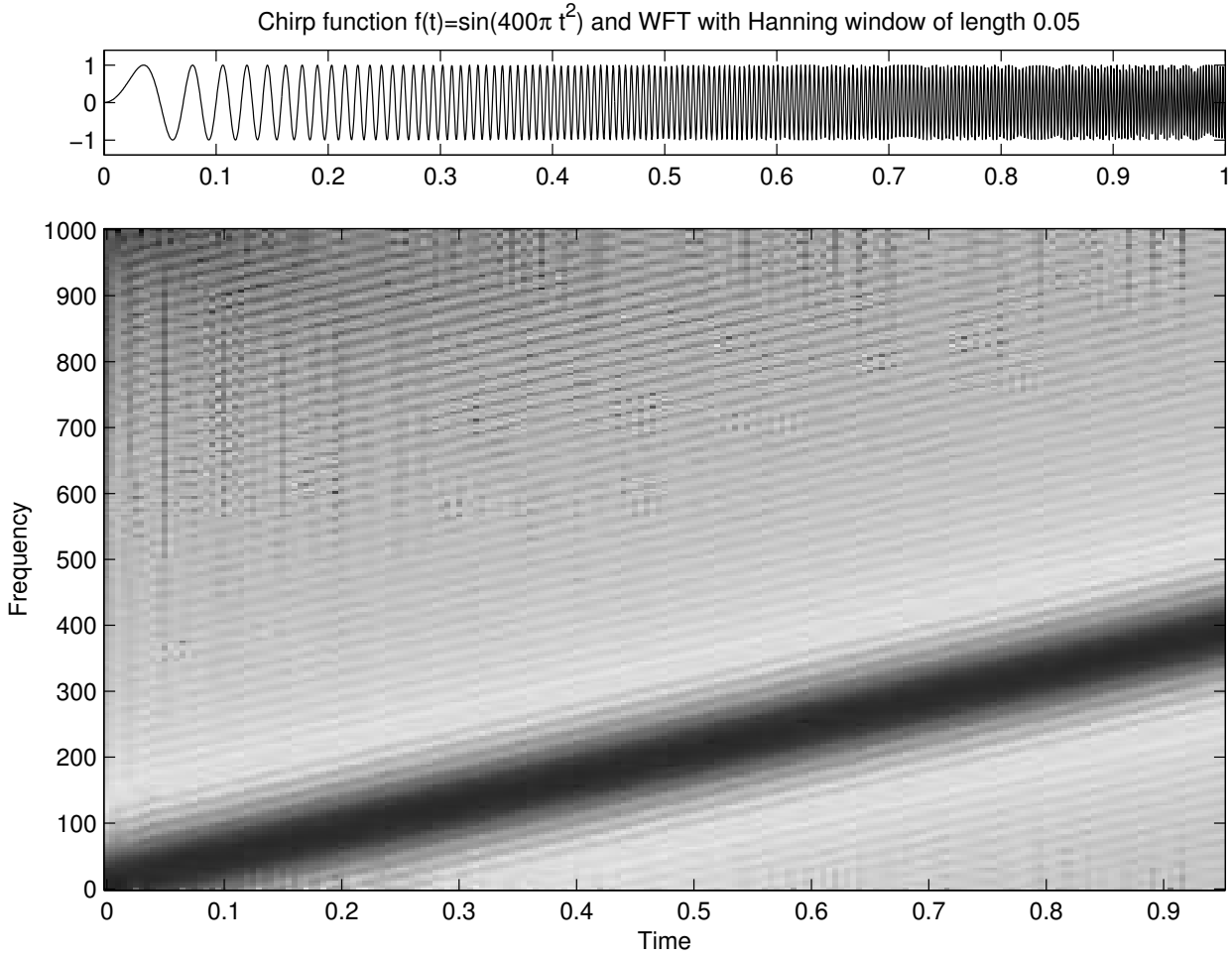


Figure 13: WFT with Hanning window.

We want to make some comments on Figure 13.

(1) As one might expect from a time-frequency representation of a chirp signal, the large values $|\tilde{f}(\omega, t)|$ lie on the diagonal $\omega = 800 \cdot t$ (which corresponds up to a factor 2π to the derivative of the phase).

(2) Furthermore, the figure shows some diagonals below and above the main diagonal which reflect small parasitic frequency arising from the destructive intereference mentioned in Example 2.12.

(3) The dark areas in the upper left and lower right corner in the time-frequency representation are not based on the properties of f or g but are caused by approximation errors. In the actual computations of the WFT by computer, the integrals $\langle f | g_{\omega,t} \rangle$ have been approximated by Riemann sums. The more the function $f \cdot g_{\omega,t}$ oscillates, the worse get the approximations of the integral $\langle f | g_{\omega,t} \rangle$ by these sums. Therefore, we have large approximation errors in the mentioned areas: in the upper area left since $g_{\omega,t}$ has large oscillations for large ω and in the lower right area since f has large oscillations for large t.

Figure 14 shows the time-frequency representation where for the WFT the box window was used instead of the Hanning window.

The abrupt cuts at the support boundaries introduced by box window cause interferences all over the spectrum. Figure 14 shows that the dark areas and hence the large values of $|\tilde{f}(\omega, t)|$ are spread all over the time-frequency domain. The large values of $|\tilde{f}(\omega, t)|$ are not as well concentrated around the diagonal $\omega = 800 \cdot t$ as it was in case of the Hanning window. Hence, using the box functions, leads to an "inferior" time-frequency respresentation of the chirp signal.

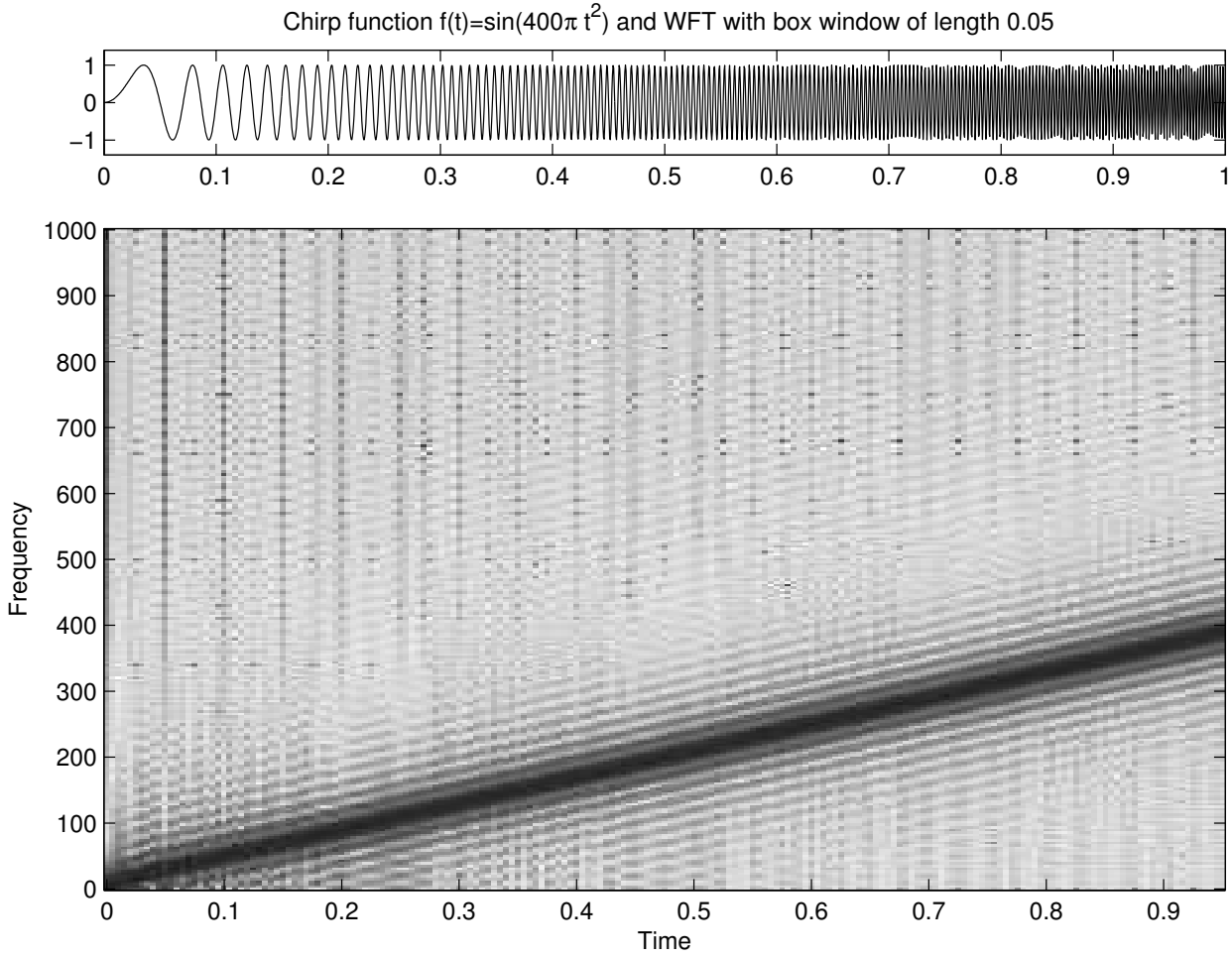


Figure 14: WFT with box window.

6.2.3 WFT in Dependence of the Window Size

In this example we consider the signal f defined on $[0, 1]$ by

$$f(t) = \sin(800\pi t) + \sin(900\pi t) + c[\delta(t - 0.45) + \delta(t - 0.5)]$$

with a small constant $c \in \mathbb{R}$ and zero on $\mathbb{R} \setminus [0, 1]$. f is the superposition of two pure sines of frequency $\omega_1 = 400$ and $\omega_2 = 450$, respectively, and two additional impulses at $t_1 = 0.45$ and $t_2 = 0.5$.

The Figures 15 and 16 show the time-frequency representation of f using a WFT w.r.t. a Hanning window — however, with different window sizes. In the first case the window size was chosen to be 0.02 and in the second case to be 0.1.

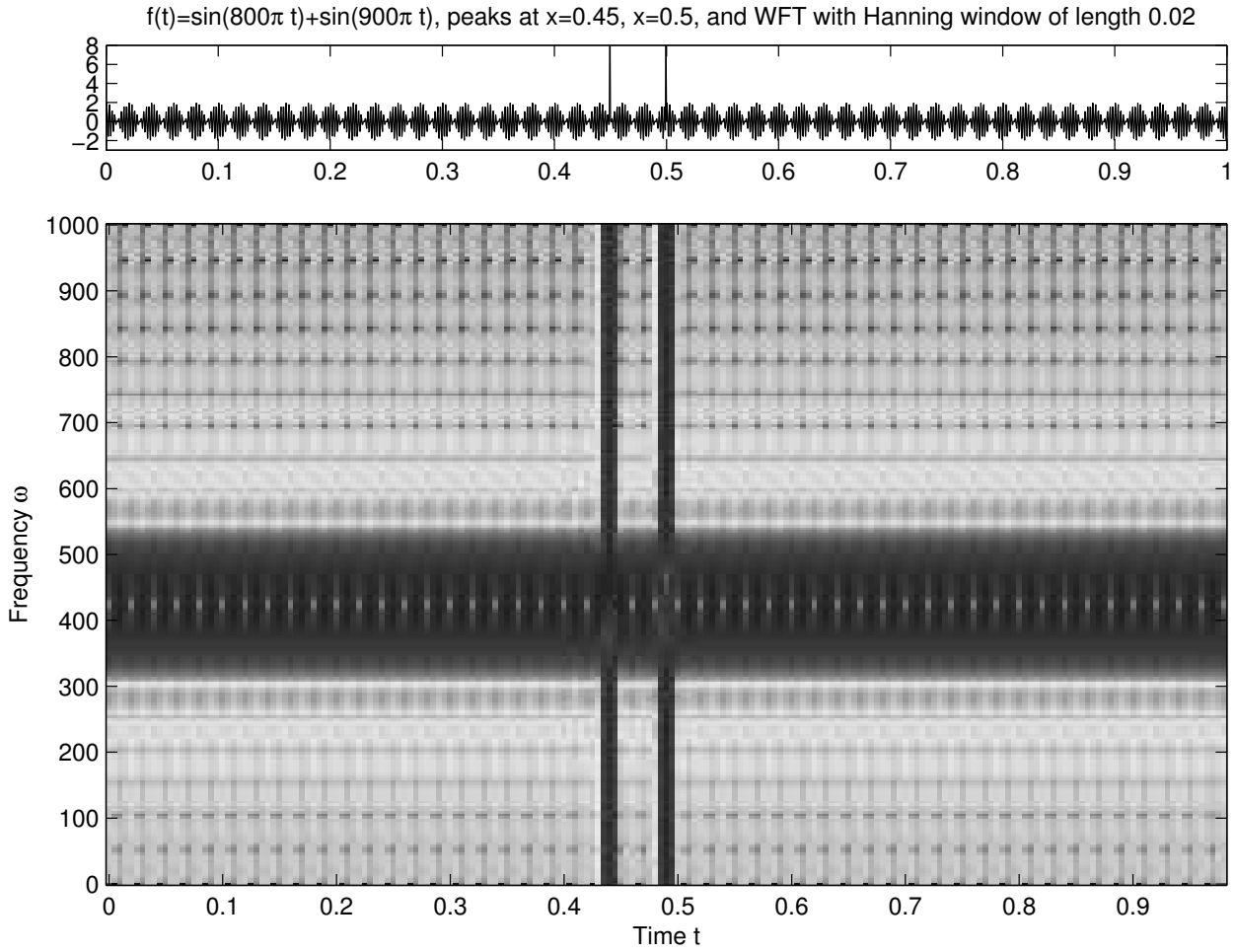


Figure 15: WFT using a small window size.

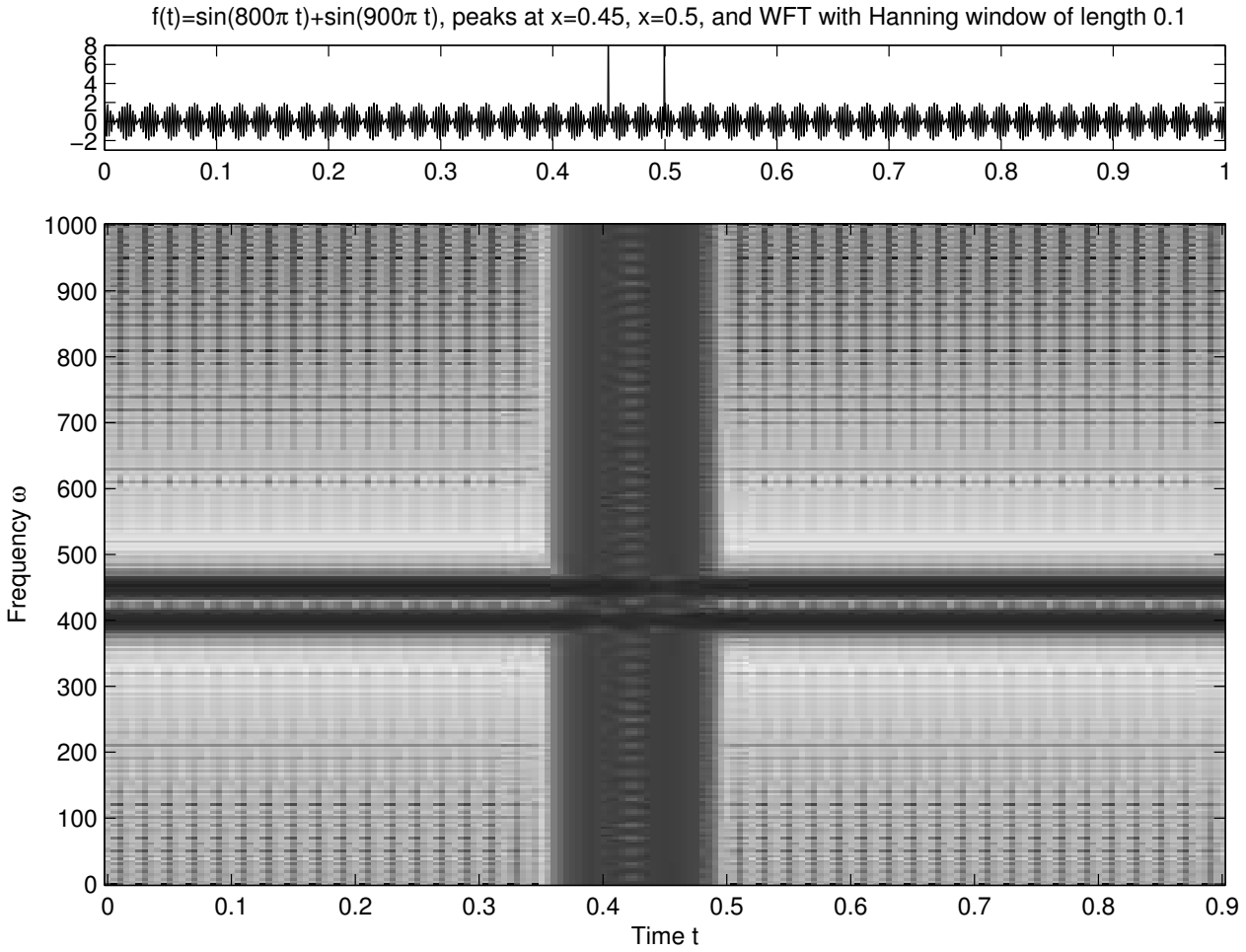


Figure 16: WFT using a large window size.

Comparing the two figures one can make the following observations:

(1) For the larger window size of 0.1 (Figure 16) one has a good resolution in the time-frequency representation of the two pure sine oscillations. However, the tow peaks at t_1 and t_2 are poorly separated.

(2) Using the smaller window size of 0.02 (Figure 15) one has a good resolution of the two peaks but the two frequencies ω_1 and ω_2 of the sines are "smeared".

This phenomena are based on a general principle:

(1) Increasing the window size of the window function g leads to a WFT (w.r.t. g) which "averages" the frequencies of the signal f over a greater time segment. Therefore, time information is lost. (In the limit case of an "infinite window size" one derives at the usual Fourier transform which averages the frequencies over all of $\mathbb{R}$.)

(2) Decreasing the window size leads to a better time resolution of the corresponding WFT. However, in this case the lower frequencies of the signal f cannot any longer be seized by the WFT since they are dominated by the window size. This leads to the above mentions smearing effects for low-frequency phenomena. (In the limit case, g is an impulse and the WFT gives back the signal f: perfect time information, no frequency information.)

6.3 Time-Frequency Localization of the WFT

We want to look at the coefficients $\tilde{f}(\omega, t)$ from two different perspectives: a time- and a frequency-based one. Let $f \in L^2(\mathbb{R})$ be a signal and $g \in L^2(\mathbb{R})$, $\|g\|_2 \neq 0$ be a window. Then, by Parseval's Identity,

$$\tilde{f}(\omega, t) = \langle f | g_{\omega,t} \rangle = \langle \hat{f} | \widehat{g_{\omega,t}} \rangle.$$

In other words, one has the following two representations for $\tilde{f}(\omega, t)$:

(1) $\tilde{f}(\omega, t) = \langle f | g_{\omega,t} \rangle = \int_{-\infty}^{\infty} f(u) \bar{g}(u - t) e^{-2\pi i \omega u} du$

(2) $\tilde{f}(\omega, t) = \langle \hat{f} | \widehat{g_{\omega,t}} \rangle = e^{-2\pi i \omega t} \int_{-\infty}^{\infty} \hat{f}(v) \bar{\hat{g}}(v - \omega) e^{2\pi i v t} dv$

These two representations allow the following interpretations:

(1) Let t_0 be a point of time. The signal f is windowed by $u \mapsto \bar{g}(u - t_0)$ und Fourier transformed. Then $|\tilde{f}(\omega, t_0)|$ tells which frequencies "appear" in f in the neighborhood of t_0.

(2) Let ω_0 be a frequency. The Fourier transform $\hat{f}$ is windowed by $v \mapsto \bar{\hat{g}}(v - \omega_0)$ and inverse Fourier transformed. Then $|\tilde{f}(\omega_0, t)|$ tells in which points of time the signal f contains the frequencies in a neighborhood of ω_0.

The neighborhood of t_0 in (1) is determined by the localization property of the window function g. Similarly, the neighborhood of ω_0 in (2) is determined by the localization property of the Fourier transform $\hat{g}$ of the window function.

If one wants to have in the time-frequency representation of a signal f a good resolution in time as well as in frequency, the window function g should have a good localization in time (property on g) as well as in frequency (property on $\hat{g}$).

However, the Heisenberg uncertainty principle says that this simultaneous localization of g is only possible up to a certain degree.

6.3.1 Heisenberg Uncertainty Principle

Formally, the localization property of a window function $g \in L^2(\mathbb{R})$ with $\|g\| = 1$ can be defined by using the notions of the center $t_0(g)$ and the width $T(g)$ of g, where

$$t_0 = t_0(g) := \int_{-\infty}^{\infty} t|g(t)|^2 dt \quad \text{and} \quad T(g) := \left(\int_{-\infty}^{\infty} (t - t_0)^2 |g(t)|^2 dt \right)^{\frac{1}{2}}.$$

Analogously, one defines the center $\omega_0(g)$ and the width $\Omega(g)$ for the Fourier transform $\hat{g}$ by

$$\omega_0 = \omega_0(g) := \int_{-\infty}^{\infty} \omega|\hat{g}(\omega)|^2 d\omega \quad \text{and} \quad \Omega(g) := \left(\int_{-\infty}^{\infty} (\omega - \omega_0)^2 |\hat{g}(\omega)|^2 d\omega \right)^{\frac{1}{2}}.$$

If the window function g is known from the context, we also simply write t_0, T, ω_0 and Ω without the arguement g. Mathematically, the center t_0 is the expectation and T the standard deviation of the random variable $t \mapsto |g(t)|^2$. In general, the center and width for arbitrary functions $g \in L^2(\mathbb{R})$ are not defined. (These values could be infinite.) In case both values are finite, we say that g localizes at t_0 with window width T. Analogously, we say that $\hat{g}$ localizes at frequency ω_0 with bandwidth Ω.

As motivated before, we are interested in a window function which localizes in time as well as in frequency. The Heisenberg uncertainty principle says that this is not possible with arbitrary precision.

Theorem 6.2. [Heisenberg Uncertainty Principle] *Let $g \in L^2(\mathbb{R})$ with $\|g\| = 1$, center $t_0(g)$ and width $T(g)$. Furthermore, let $\omega_0(g)$ and $\Omega(g)$ be the center and width of $\hat{g}$, respectively. Then*

$$T(g) \cdot \Omega(g) \geq \frac{1}{4\pi}.$$

Squaring both sides yields

$$\left(\int_{-\infty}^{\infty} (t - t_0)^2 |g(t)|^2 dt \right) \left(\int_{-\infty}^{\infty} (\omega - \omega_0)^2 |\hat{g}(\omega)|^2 d\omega \right) \geq \frac{1}{16\pi^2}.$$

Proof:

We prove this assertion only for $g \in L^2(\mathbb{R})$, which are continuously differentiable such that the derivative g' is also in $L^2(\mathbb{R})$. One can show that such functions form a dense subset of $L^2(\mathbb{R})$. For arbitrary functions in $L^2(\mathbb{R})$ the assertion follows then by some approximation argument. We refer to the literature for such a proof.

We simplify the problem stepwise.

(1) In case $T(g)$ or $\Omega(g)$ is infinite, the assertion of the theorem becomes obvious. Therefore, we now may restrict ourselves to functions $g \in L^2(\mathbb{R})$ for which both of these values are finite.

(2) For some $g \in L^2(\mathbb{R})$ define $h \in L^2(\mathbb{R})$ by

$$h(t) := e^{-2\pi i \omega_0 t} g(t + t_0).$$

By Theorem 2.9,

$$\hat{h}(\omega) = e^{2\pi i t_0(\omega+\omega_0)} \hat{g}(\omega + \omega_0).$$

From this it is easy to see that h is centered, i.e., $t_0(h) = 0$ and $\omega_0(h) = 0$, and

$$\begin{aligned} & \left(\int_{-\infty}^{\infty} t^2 |h(t)|^2 dt \right) \left(\int_{-\infty}^{\infty} \omega^2 |\hat{h}(\omega)|^2 d\omega \right) \\ = & \left(\int_{-\infty}^{\infty} (t - t_0)^2 |g(t)|^2 dt \right) \left(\int_{-\infty}^{\infty} (\omega - \omega_0)^2 |\hat{g}(\omega)|^2 d\omega \right). \end{aligned}$$

In other words, it suffices to show the Heisenberg uncertainty principle for centered functions.

In the following let $g \in L^2(\mathbb{R})$ with $\|g\| = 1$, $t_0(g) = 0$, and $\omega_0(g) = 0$. Furthermore, let $T(g)$ and $\Omega(g)$ be finite, g differentiable, and $g' \in L^2(\mathbb{R})$. Then using $\omega\hat{g}(\omega) = \frac{1}{2\pi i}\widehat{g'}(\omega)$ and Parseval's equation $\|\widehat{g'}\| = \|g'\|$, we obtain

$$\begin{aligned} T(g)^2 \cdot \Omega(g)^2 &= \left(\int_{-\infty}^{\infty} t^2 |g(t)|^2 dt\right)\left(\int_{-\infty}^{\infty} \omega^2 |\hat{g}(\omega)|^2 d\omega\right) \\ &= \frac{1}{4\pi^2}\left(\int_{-\infty}^{\infty} t^2 |g(t)|^2 dt\right)\left(\int_{-\infty}^{\infty} |g'(t)|^2 dt\right) \end{aligned}$$

From the Cauchy-Schwarz inequality $\|f_1\|^2\|f_2\|^2 \geq |\langle f_1|f_2\rangle|^2$ with $f_1(t) = |tg(t)|$ and $f_2(t) = |g'(t)|$ follows

$$\begin{aligned} & \frac{1}{4\pi^2}\left(\int_{-\infty}^{\infty} t^2|g(t)|^2dt\right)\left(\int_{-\infty}^{\infty}|g'(t)|^2dt\right) \\ \geq \quad & \frac{1}{4\pi^2}\left(\int_{-\infty}^{\infty}|tg(t)g'(t)|dt\right)^2 \quad \geq \ldots \end{aligned}$$

For arbitrary complex numbers $a, b \in \mathbb{C}$ holds $|ab| = |a\bar{b}| \geq \mathrm{Re}(a\bar{b}) = \frac{1}{2}(a\bar{b} + \bar{a}b)$. Using this for $a = tg(t)$ and $b = g'(t)$ one gets

$$\ldots \geq \frac{1}{4\pi^2}\left(\frac{1}{2}\int_{-\infty}^{\infty}(tg(t)\overline{g'(t)} + t\overline{g(t)}g'(t))dt\right)^2$$

Using $\frac{d}{dt}|g(t)|^2 = g(t)\overline{g'(t)} + g'(t)\overline{g(t)}$, $\int_{-\infty}^{\infty} t|g(t)|^2 dt = t_0(g) = 0$ and hence $\lim_{t\to\infty} t|g(t)|^2 = 0$, it follows by partial integration that

$$\begin{aligned} & \frac{1}{4\pi^2}\left(\frac{1}{2}\int_{-\infty}^{\infty}(tg(t)\overline{g'(t)} + t\overline{g(t)}g'(t))dt\right)^2 \\ \geq \quad & \frac{1}{16\pi^2}\left(-\int_{-\infty}^{\infty}|g(t)|^2 dt\right)^2 = \frac{1}{16\pi^2}\|g\|^4. \end{aligned}$$

Since $\|g\| = 1$, the assertion follows. □

Note 6.3. The boundary in the Heisenberg uncertainty principle is sharp. Indeed, one can show that for the Gauss function (see also Example 2.15) g_{ω_0,t_0} defined by

$$g_{\omega_0,t_0}(t) := \pi^{-\frac{1}{4}} \frac{1}{\sqrt{2\pi}} e^{-2\pi i \omega_0 t} e^{-\pi(t-t_0)^2}$$

holds $t_0(g_{\omega_0,t_0}) = t_0$, $\omega_0(g_{\omega_0,t_0}) = \omega_0$ and

$$T(g_{\omega_0,t_0}) \cdot \Omega(g_{\omega_0,t_0}) = \frac{1}{4\pi}.$$

Furthermore, it holds that this function is the only function with minimal uncertainty at (t_0, ω_0).

Note 6.4. We close this subsection with some more intuitive notes on the principle which are partly taken from [Hubbard].

- The Heisenberg uncertainty principle has its origin in quantum physics. Intuitively it says that an elementary particle does not simultaneously have a precise position and a precise momentum. In our case this principle says that a signal does not simultaneously have a precise location in time and a precise frequency.
- A very brief signal, well localized in time, necessarily has a Fourier transform that is spread out: a broad range of frequencies. Conversely, a signal with a very narrow range of frequencies is necessarily spread in time; it's not possible to convince just a few sines and cosines to cancel out so that the vlaues of the signal is small outside a narrow time interval.
- The time parameter t in $\tilde{f}(\omega, t)$ of the WFT is not sharp but represents a time interval, which depends on the window width of g. Similar holds for the frequency parameter ω. The choice of the window function g determines the "resolution proportion" between time and frequency.

6.3.2 Information Cells

In all methods which analyze a signal simultaneously in time and frequency the Heisenberg uncertainty principle comes into play. If one wants a good resolution in time, one has to put up with a poor frequency resolution and vice versa. These compromises can be illustrated in the so-called time-frequency domain where the horizontal axis represents time and the vertical axis represents frequency.

Let $g \in L^2(\mathbb{R})$, $\|g\| = 1$, be window with center t_0 and width T, such that the Fourier transform $\hat{g}$ has center ω_0 and bandwidth Ω. Then g can be represented in the time-frequency domain by a rectangle which is parallel to the axes having width T, height Ω, and center of gravity (t_0, ω_0). Such a rectangle is called information cell and will be denoted by $\mathrm{IC}(g)$ (see Figure 17).

Then the Heisenberg uncertainty principle sasy the the area of each such information cell is at least $1/4\pi$:

$$\mathrm{Area}(\mathrm{IC}(g)) \geq \frac{1}{4\pi}.$$

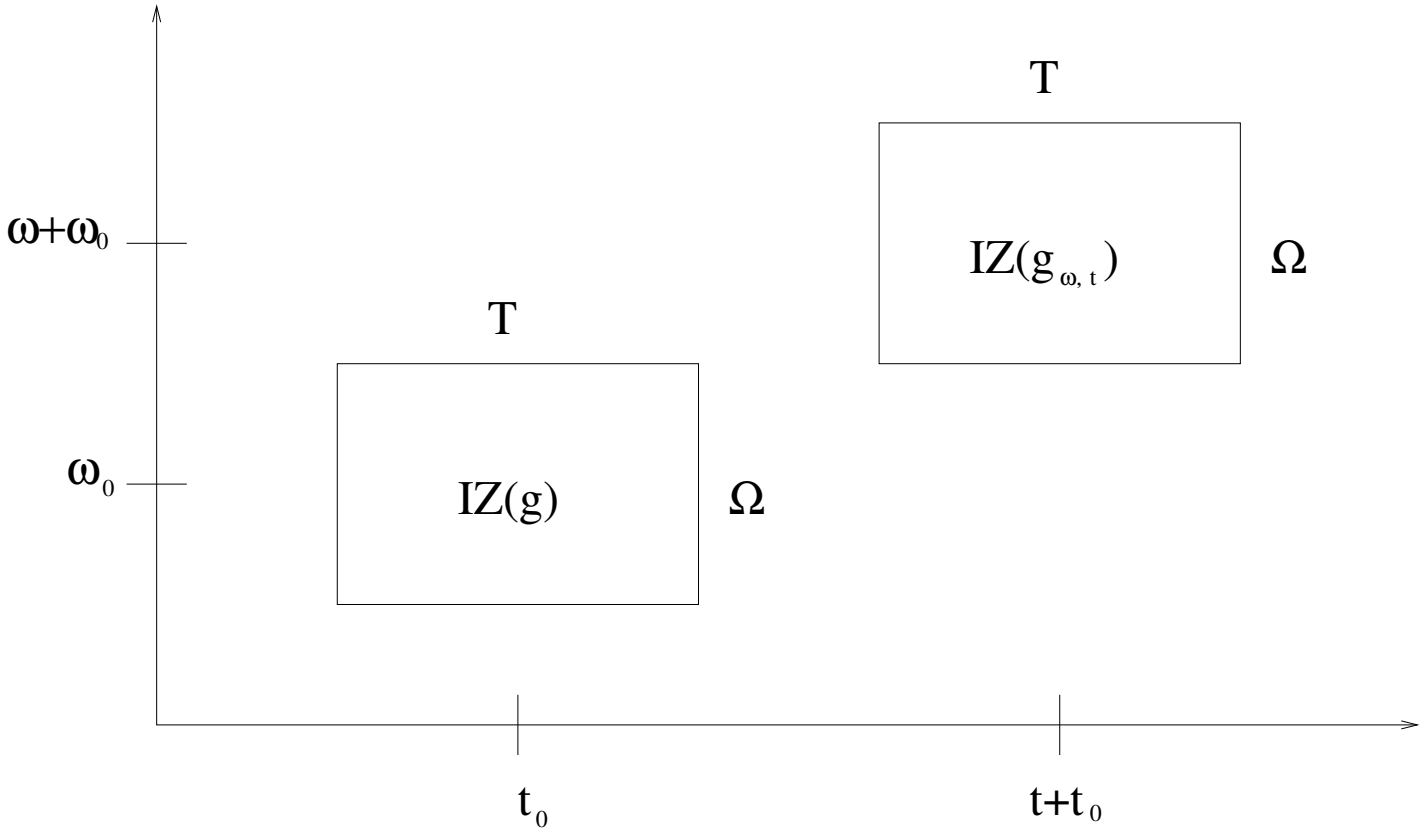


Figure 17: Information cells for g and $g_{\omega,t}$.

One easily computes that the information cell of the "musical note"

$$g_{\omega,t} : \mathbb{R} \to \mathbb{C}, \quad g_{\omega,t}(u) := e^{2\pi i\omega u} g(u-t)$$

for a window g as above has also width T and height Ω with center of gravity $(t_0 + t, \omega_0 + \omega)$. The WFT

$$\tilde{f}(\omega, t) = \langle f | g_{\omega,t} \rangle$$

of a signal $f \in L^2(\mathbb{R})$ w.r.t. the window function g gives an analysis of f in the area of the information cell of $g_{\omega,t}$. In the time-frequency domain the WFT is depicted by representing the value $\langle f | g_{\omega,t} \rangle$ over the information cell of $g_{\omega,t}$ with a suitable grey color.

If one considers the WFT only on a discrete grid of points in the time-frequency domain, one gets a tiling of the time-frequency domain by (possibly overlapping) congruent information cells, where the grey levels reflect the time-frequency information of the signal f.

6.3.3 Reconstruction of the Signal from its WFT

Let $f \in L^2(\mathbb{R})$ be a signal with WFT $\tilde{f}(\omega, t)$ w.r.t the window function $g \in L^2(\mathbb{R})$, $\|g\| \neq 0$. As explained in Section 1, the value $\tilde{f}(\omega, t)$ expresses in which intensity the note $g_{\omega,t}$ is "contained" in the signal f. Therfore, intuitively, one should be able to reconstruct the signal f as superposition of the notes $g_{\omega,t}$ weighted by $\langle f | g_{\omega,t} \rangle$:

$$f(u) \sim \int_{\mathbb{R}} \int_{\mathbb{R}} \langle f | g_{\omega,t} \rangle g_{\omega,t}(u) d\omega dt.$$

We will now show that this is indeed correct up to a normalizing factor.

By definition holds $\tilde{f}(\omega, t) = \hat{f}_t(\omega)$. Applying the inverse Fourier transform to $\hat{f}_t(\omega)$ w.r.t. the variable ω one gets

$$\bar{g}(u-t)f(u) = f_t(u) = \int_{\mathbb{R}} \tilde{f}(\omega, t) e^{2\pi i \omega u} d\omega.$$

Note that $f(u)$ cannot simply recovered through division by $\bar{g}(u-t)$ since this function coud be zero. Instead, we multiply both sides by $g(u-t)$ and integrate over t:

$$\int_{\mathbb{R}} |g(u-t)|^2 f(u) dt = \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{f}(\omega, t) g(u-t) e^{2\pi i \omega u} d\omega dt.$$

Since by assumption $\|g\| \neq 0$, the reconstruction formula

$$f(u) = \frac{1}{\|g\|^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{f}(\omega, t) g_{\omega,t}(u) d\omega dt$$

follows. In case $\|g\| = 1$, this coincides exactly with our intuitive considerations above.

6.4 Discrete Version of the WFT

The synthesis formula

$$f(u) = \frac{1}{\|g\|^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{f}(\omega, t) g_{\omega,t}(u) d\omega dt$$

for the reconstruction of a signal f from its WFT is in general a rather redundant representation: a one-dimensional parameter space (the time denoted by the variable u) is represented by an integral over a two-dimensional parameter space (the time-frequency domain represented by the variables t and ω). In this section, we want to investigate when a discrete (or even finite) set of values $\tilde{f}(\omega, t)$ is sufficient for the reconstruction of the signal f. This is also important in view of practical computations.

We assume in this section that the window function g has compact support contained in the interval $[a, b] \subset \mathbb{R}$. Then the localized signal $f_t(u) := \bar{g}(u-t)f(u)$ has support in $[a+t, b+t]$ and can therefore on this interval be represented by the Fourier series

$$f_t(u) = \sum_{m\in\mathbb{Z}} \langle f_t | \frac{1}{\sqrt{T}} e_{T,m} \rangle \frac{1}{\sqrt{T}} e_{T,m}(u), \tag{10}$$

where $T := b - a$ and the functions $e_{T,m}$ are defined by

$$e_{T,m}(u) := e^{2\pi i u m/T}, \quad m \in \mathbb{Z}.$$

Note that $(\frac{1}{\sqrt{N}} e_{T,m})_{m\in\mathbb{Z}}$ form an ONB of the Hilbert space $L^2([a+t, b+t])$ with inner product $\langle f|g\rangle = \int_{a+t}^{b+t} f(u)\overline{g(u)}du$. We emphasize that this equality only hold for $u \in [a+t, b+t]$.

The value $\nu := \frac{1}{b-a} = \frac{1}{T}$ is called frequency step width and represents the distance between two neighboring frequencies to be analyzed. Then

$$\begin{aligned}\langle f_t | e_{T,m} \rangle &= \int_{a+t}^{b+t} \bar{g}(u-t) f(u) e^{-2\pi i m \nu u} du \\ &= \int_{-\infty}^{\infty} \bar{g}(u-t) f(u) e^{-2\pi i m \nu u} du \\ &= \tilde{f}(m\nu, t).\end{aligned}$$

Multiplying both sides of the Fourier series representation (10) by $g(u-t)$ one gets

$$|g(u-t)|^2 f(u) = \nu \sum_{m \in \mathbb{Z}} g(u-t) \tilde{f}(m\nu, t) e^{2\pi i m \nu u}.$$

The goal is to isolate $f(u)$ on the left hand side. To this means we define the function

$$H_\tau(u) := \tau \sum_{n\in\mathbb{Z}} |g(u-n\tau)|^2$$

w.r.t. the time step width $\tau > 0$. H_τ is well-defined since the support of g is assumed to be compact. Furthermore, if g is Riemann integrable, H_τ defines a Riemann sum for $\|g\|^2$ and therefore

$$H_\tau(u) \to \|g\|^2 \quad \text{for} \quad \tau \to 0.$$

Summation over n with time step width τ results in

$$H_\tau(u) f(u) = \tau\nu \sum_{n\in\mathbb{Z}} \sum_{m\in\mathbb{Z}} g(u-n\tau)\tilde{f}(m\nu, n\tau) e^{2\pi i m\nu u}.$$

We want to divide by H_τ which is only possible if $H_\tau(u) \neq 0$ for almost all u. Let

$$A_\tau := \inf_{u\in\mathbb{R}} H_\tau(u) \quad \text{and} \quad B_\tau := \sup_{u\in\mathbb{R}} H_\tau(u).$$

Then one can show that for "nice" window functions g and for sufficiently small τ and all $u \in \mathbb{R}$ holds

$$0 < A_\tau \leq H_\tau(u) \leq B_\tau < \infty.$$

In this case we obtain a discrete WFT-reconstruction

$$f(u) = \tau\nu \sum_{n\in\mathbb{Z}} \sum_{m\in\mathbb{Z}} \frac{g(u-n\tau)}{H_\tau(u)} \tilde{f}(m\nu, n\tau) e^{2\pi i m\nu u}. \tag{11}$$

The finiteness of B_τ assures the numerical stability of the summation.

Note 6.5. For $A_\tau > 0$ to be satisfied one needs $0 < \tau \leq b - a$, since otherwise $H_\tau(u) = 0$ for $b < u < a + \tau$. With $0 < \tau \leq b - a$ and by definiton of ν holds

$$0 < \tau \cdot \nu \leq 1.$$

Then the sample density defined by

$$\rho(\nu, \tau) := \frac{1}{\nu \cdot \tau} = \frac{b - a}{\tau}$$

is a measure for the redundancy of the representation (11) of f. For example, if $b - a = 1024$ and $\tau = 256$, then the sample density is 4, which is in real-time applications a common value. The extreme case $\tau \cdot \nu = 1$ contains no redundancy and, as it turns out, is in practical computations numerically not stable.

6.5 Drawback of the WFT

Even though the introduction of window functions leads to localization in time as well as in frequency the analysis of a signal by the WFT has several drawbacks which we discuss next.

The window function g introduces a kind of scaling in the time-frequency domain. In other words, for a given g the width T and the height Ω of the corresponding information cell $\mathrm{IC}(g)$ is determined once and for all. The signal f is analyzed by the notes $g_{\omega,t}$ on the translated information cell $\mathrm{IC}(g_{\omega,t}) = \mathrm{IC}(g) + (t, \omega)$.

As we explain next, the WFT is not efficient when the signal f contains phenomena which are either of short duration or of long duration as compared to T.

(1) Suppose the signal f contains a local oscillation or a peak of width $\Delta u \ll T$ at $u = u_0$. Such a peak is in the WFT-synthesis represented as a superposition of notes $g_{\omega,t}$ each of it having length T. This is only possible, if many of the notes $g_{\omega,t}$ with $t \approx u_0$ of many different frequencies are used, due to the principle of constructive and destructive interference. Therefore, $\tilde{f}(\omega, u_0)$ is spread out in frequency at point of time u_0.

(2) We now suppose that f contains some large, global variations with $\Delta u \gg T$ such as a low-frequency sine oscillation. In this case many of the low-frequency notes $g_{\omega,t}$ are needed to synthesize f over Δu. Therefore, $\tilde{f}(\omega, t)$ is spread out in time.

The following principles

Principle 1: Properties of f, which are much shorter than T, are synthesized in the frequency domain.

Principle 2: Properties of f, which are much longer than T, are synthesized in the time domain.

In both cases many of the notes $g_{\omega,t}$ are needed to synthesize phenomena of the signal f which are relatively simple in nature such as a peak or a low-frequency oscillation. In other words, the WFT is not capable to capture such features in few, but large coefficients $\tilde{f}(\omega, t)$ and is therefore not efficient.

In the following chapter we introduce the time-frequency analysis based on wavelets which overcome some of the drawbacks of the WFT.

Chapter 7: Continuous Wavelet Transform (CWT)

Just like the WFT, the wavelet transform also gives a time-frequency representation of a signal. Analogously to the WFT, the continuous wavelet transform (CWT) of a signal $f \in L^2(\mathbb{R})$ is defined by

$$\tilde{f}(s,t) := |s|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(u) \bar{\psi} \left(\frac{u-t}{s} \right) du,$$

where $\psi \in L^2(\mathbb{R})$ is a suitable function called mother wavelet.

Hence the WFT and CWT are very similar in their definition: both transforms compute inner products of the signal f with a two parameter family of functions, namely

$$g_{\omega,t}(u) := e^{2\pi i\omega u} g(u-t)$$

for the WFT and

$$\psi^{s,t}(u) := |s|^{-\frac{1}{2}} \psi\left(\frac{u-t}{s}\right)$$

for the CWT. In the latter case, the functions $\psi^{s,t}, t \in \mathbb{R}, s \in \mathbb{R} \setminus \{0\}$, are called wavelets. This are scaled and translated versions of the mother wavelet $\psi = \psi^{1,0}$ (therefore, the notation s and t for the parameters).

However, there is a fundamental difference between the WFT and CWT. For the WFT the analyzing functions $g_{\omega,t}$ have constant window size where ω specifies the frequency, i.e., the number of oscillations. For the CWT, the number of oscillations of the analyzing window $\psi^{s,t}$ is constant where s specifies the window size. Increasing the scale parameters s leads to a large window size, which induces for a fixed number of oscillations a decrease in frequency. In other words,

- a WFT analyzes a signal by notes $g_{\omega,t}$ of different frequencies all having the same duration, whereas
- a CWT analyzes a signal by notes $\psi^{s,t}$ where notes of low frequencies are long and notes of high frequencies are short.

Therefore, a CWT can recognize — by means of the short and high-pitched notes — short, high-frequency phenomena of the signal (such as peaks), but can also detect — by means of long and low-pitched notes — smooth, low-frequency characteristics of the signal.

In this chapter we introduce the CWT in a similar fashion as the WFT. For the CWT the mathematical background is more complicate than for the WFT, so that we are not able to give proofs for many of the facts we need. However, we also attach importance to mathematically exact definitions and rigorous argumentations. Our goal is to present the main ideas of the wavelet transform and to give at some points insight into the underlying mathematics. In Chapter 8 we will be awarded by a beautiful algorithm known as fast discrete wavelet transform. Actually, this algorithm computes the continuous wavelet transform on a discrete time grid when the input signal is given in form of a linear combination of certain basis functions.

7.1 Definition of the CWT

For the windowed Fourier transform or WFT the starting point of the analysis was a window function $g \in L^2(\mathbb{R})$, which can be thought of as a bell-shaped real function localized at zero. From a mathematical point of view, however, all functions $g \in L^2(\mathbb{R})$ with $\|g\| \neq 0$ could be used to enable the reconstruction of a signal from its WFT.

Similarly, for the continuous wavelet transform or CWT the starting point of the analysis is a Wavelet ψ, which can be thought of a "small wave" or "ripple" localized at zero. From a mathematical point of view not all $\psi \in L^2(\mathbb{R})$ with $\|\psi\| \neq 0$ can be used as wavelet. Ones needs some technical assumption on ψ whose meaning will become clear when one tries to reconstruct the signal form its CWT. The definition of a wavelet is as follows.

Definition 7.1. A function $\psi \in L^2(\mathbb{R})$, which satisfies the admissibility condition

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

is called a wavelet.

We only give some intuitive comments on the admissibility condition.

- One can show that the set of wavelets $\{\psi \in L^2(\mathbb{R}) | \psi \text{ is admissible}\}$ together with the zero function forms a dense, linear subspace of $L^2(\mathbb{R})$. In this sense, one has a "large assortment" of wavelets.
- The admissibility condition implies for wavelets $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ that the mean value of ψ vanishes, i.e., $\int_{\mathbb{R}} \psi(t) dt = 0$ and hence $\hat{\psi}(0) = 0$. Intuitively, this means the a wavelet oscillates (low frequencies do not appear!) around the x-axis (mean value is zero!).

For further details we refer to [Louis/Maaß/Rieder].

For a wavelet ψ und $s, t \in \mathbb{R}, s \neq 0$, we define the functions

$$\psi^{s,t} : \mathbb{R} \to \mathbb{C}, \quad \psi^{s,t}(u) := |s|^{-\frac{1}{2}} \psi \left(\frac{u-t}{s} \right) .$$

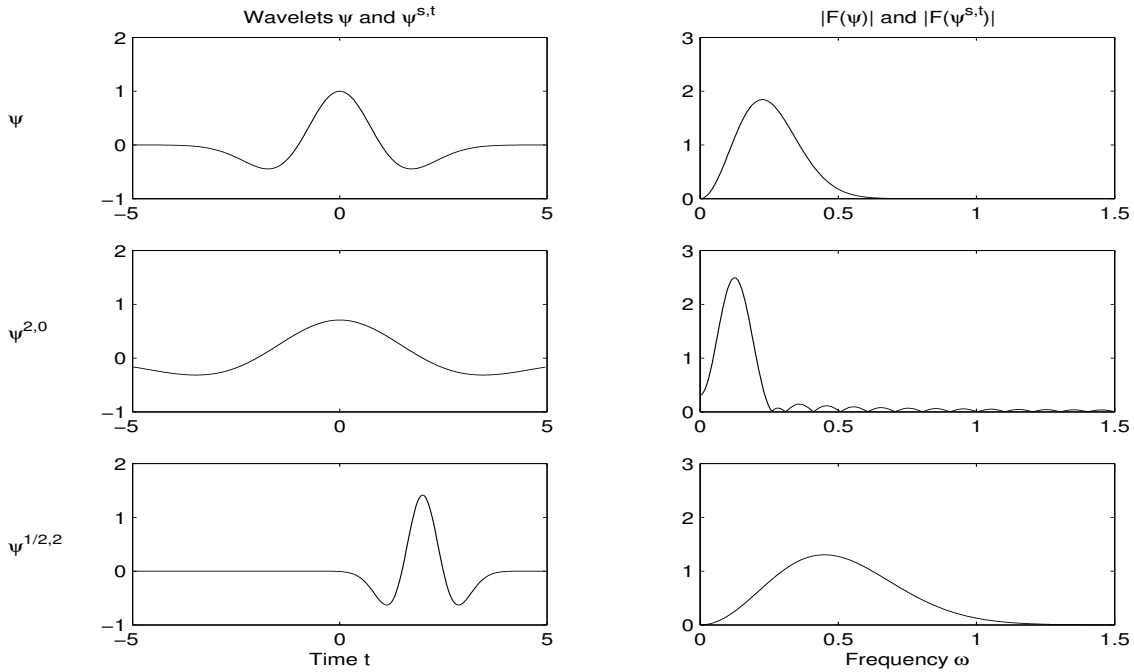


Figure 18: Scaled and translated versions $\psi^{s,t}$ of the wavelet ψ.

Theorem 7.2. *If ψ is a wavelet, then all $\psi^{s,t}$ are wavelets as well, for $s, t \in \mathbb{R}$, $s \neq 0$. In particular,*

$$\begin{aligned} \|\psi^{s,t}\| &= \|\psi\|, && (12)\\ \widehat{\psi^{s,t}}(\omega) &= |s|^{1/2} e^{-2\pi i \omega t} \hat{\psi}(\omega s), && (13)\\ c_{\psi^{s,t}} &= |s| c_\psi. && (14) \end{aligned}$$

Proof: The assertion that $\psi^{s,t}$ is a wavelet in case ψ is one, follows immediately from (12) and (14). In the following, we only prove (13). Assertion (12) and (14) follow by a similar computation which are left as an exercise.

$$\begin{aligned} \widehat{\psi^{s,t}}(\omega) &= \int_{\mathbb{R}} \psi^{s,t}(u) e^{-2\pi i \omega u} du = |s|^{-\frac{1}{2}} \int_{\mathbb{R}} \psi\left(\frac{u-t}{s}\right) e^{-2\pi i \omega u} du \\ &= |s|^{-\frac{1}{2}} \int_{\mathbb{R}} |s| \psi(x) e^{-2\pi i \omega (xs+t)} dx = |s|^{1/2} e^{-2\pi i \omega t} \hat{\psi}(\omega s). \end{aligned}$$

□

Definition 7.3. The continuous wavelet transform or CWT of a signal $f \in L^2(\mathbb{R})$ w.r.t. the wavelet ψ is defined by

$$\tilde{f}(s,t) := \langle f|\psi^{s,t}\rangle = |s|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(u)\bar{\psi}\left(\frac{u-t}{s}\right) du,$$

where $s \in \mathbb{R} \setminus \{0\}$ and $t \in \mathbb{R}$.

We use the same symbol $\tilde{f}$ for the WFT and CWT. However, from the context and from the parameters (ω, t) and (s, t), respectively, it should be clear which transform is meant. Analogously to the WFT, the inner products $\langle f|\psi^{s,t}\rangle$ measure the correlation of the signal f with the "musical note" $\psi^{s,t}$. The signal

$$u \mapsto \langle f|\psi^{s,t}\rangle \psi^{s,t}(u)$$

is the "projection" of the signal f onto the subspace spanned by the musical note $\psi^{s,t}$ and expresses the share in which $\psi^{s,t}$ is "contained" in f.

7.2 Examples of Wavelets

In this section, we will present several examples of wavelets given explicitly by some formula or defined implicitly by some recursive construction. In order to show that these functions indeed define wavelets, we apply some theorems which give sufficient conditions. For the proofs of these theorems we refer to [Louis/Maaß/Rieder].

Theorem 7.4. *Suppose $\psi \in L^2(\mathbb{R}) \setminus \{0\}$ has compact support. Then*

$$\psi \quad \text{wavelet} \qquad \Longleftrightarrow \int_{\mathbb{R}} \psi(t)dt = 0.$$

This theorem shows that the function $\psi \in L^2(\mathbb{R}) \setminus \{0\}$ shown in Figure 19 and defined by

$$\psi(t) := \begin{cases} 1 & \text{if} \quad 0 \le t < \frac{1}{2} \\ -1 & \text{if} \quad \frac{1}{2} \le t < 1 \\ 0 & \text{otherwise.} \end{cases}$$

is a wavelet, since obviously $\int_{\mathbb{R}} \psi(t) = 0$. This wavelet is also known as Haar wavelet and constitutes an easy example, which will be used later as illustration.

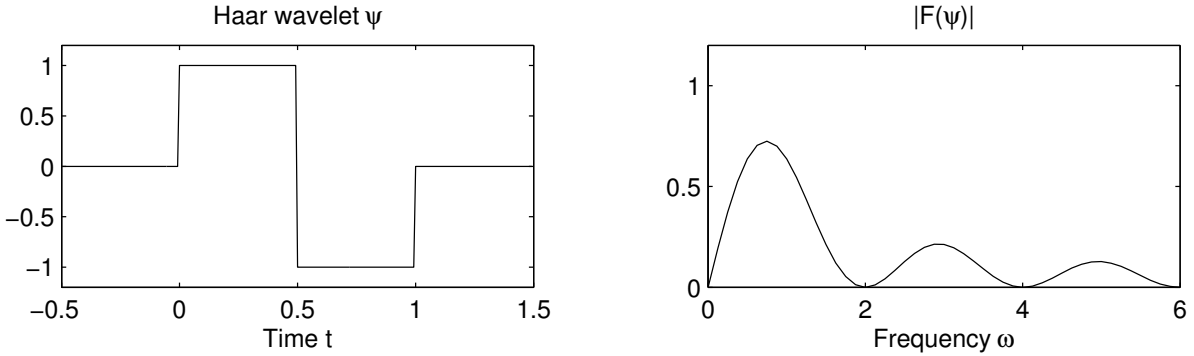


Figure 19: Haar wavelet and its Fourier transform.

Theorem 7.5. *Suppose $\phi \in L^2(\mathbb{R})$ is continuously differentiable with*

$$\psi := \phi' \in L^2(\mathbb{R}) \setminus \{0\}.$$

Then ψ is a wavelet.

As application of this theorem, we consider the so-called Mexican hat $\psi : \mathbb{R} \longrightarrow \mathbb{R}$ shown in Figure 20 and defined by

$$\psi(t) := -\frac{d^2}{dt^2} e^{\frac{-t^2}{2}} = (1 - t^2) e^{\frac{-t^2}{2}}.$$

Obviously, the function $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$\phi(t) := -\frac{d}{dt} e^{\frac{-t^2}{2}}$$

satisfies the conditions in Theorem 19. The function ϕ is, up to a sign, the derivative of the Gauss function.

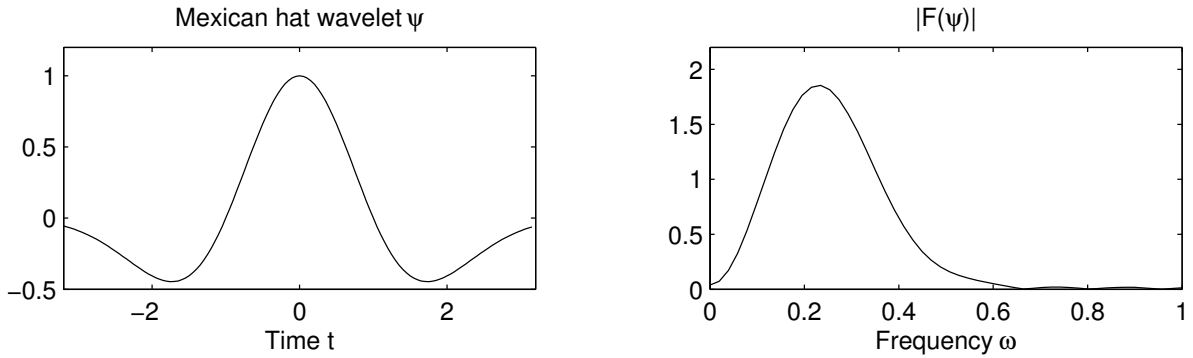


Figure 20: Mexican hat wavelet and its Fourier transform.

The illustration of ψ (up to a normalizing factor) in Figure 20 tells, how this wavelet came to its name. In contrast to the Haar wavelet, the Mexican hat wavelet is a C^∞-function and has a much better localization property in the frequency domain.

As another example, we introduce the so-called Meyer wavelet which is shown in Figure 21.

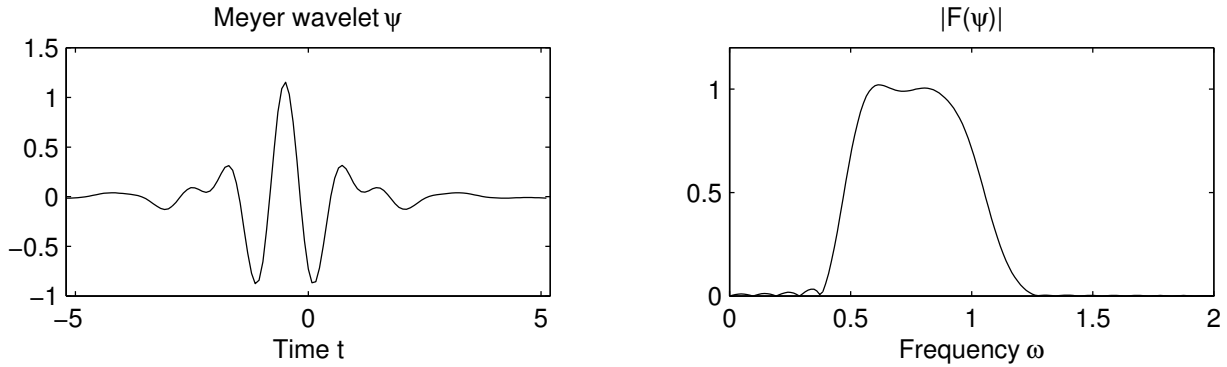


Figure 21: Meyer wavelet and its Fourier transform.

The Meyer wavelet ψ is defined by means of its Fourier transform $\hat{\psi}$. Let

$$\hat{\psi}(y) := \frac{1}{\sqrt{2\pi}} e^{\frac{iy}{2}} (w(y) + w(-y))$$

with

$$w(y) := \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(\frac{3y}{2\pi} - 1\right)\right) & \text{if} \quad \frac{2\pi}{3} \leq y < \frac{4\pi}{3}, \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{3y}{2\pi} - 1\right)\right) & \text{if} \quad \frac{4\pi}{3} \leq y < 2\pi, \\ 0 & \text{otherwise}, \end{cases}$$

where $\nu : \mathbb{R} \longrightarrow [0, 1]$ is a smooth function such that $\nu(y) = 0$ for $y \leq 0$, $\nu(y) = 1$ for $y \geq 1$, and $\nu(y) + \nu(1 - y) = 1$. The admissibility condition for ψ follows from the explicit formula of $\hat{\psi}$.

Note 7.6. Up to now, we gave examples for wavelets which were all given by explicit formulas. However, for many of the common wavelets such explicit formulas do not exist. In most cases, wavelets are implicitly defined by means of the so-called associated filter coefficients which will be introduced in Chapter 8. The wavelets can then be constructed from these filter coefficients by means of some recursive algorithm.

One important class of such implicitly defined wavelets are the Daubechies wavelets. Due to their central importance in the theory of wavelets, we have illustrated the first five wavelets of this family in Figure 22.

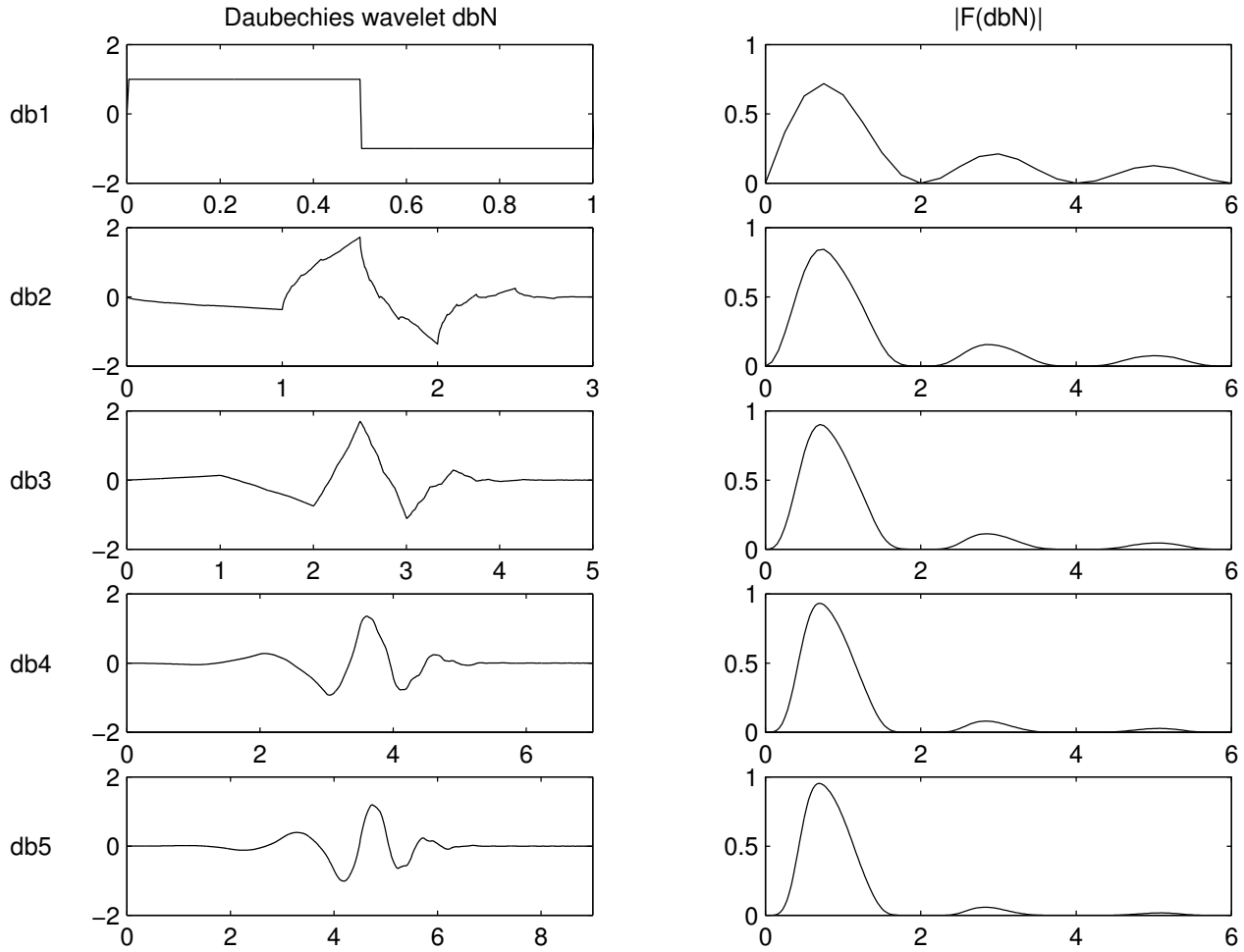


Figure 22: Daubechies wavelets and their corresponding Fourier transform.

The family of Daubechies wavelets is parameterized by $N \in \mathbb{N}$, where the Nth wavelet is denoted by db$N$. The wavelet db1 is just the Haar wavelet from Figure 19. All other Daubechies wavelets do not have explicit formulas. As is illustrated by Figure 22, the regularity of the wavelets db$N$ (e.g., in the sense of differentiability or in the sense of localization property in the frequency domain) increases with increasing N. The wavelet db$N$ has compact support $[0, 2N - 1]$.

For a definition of the wavelets and further properties wie refer to [Strang/Nguyen].

7.3 Time-frequency localization of the CWT

In a similar fashion as in Section 3 for the WFT we now want to investigate the time-frequency representation $\tilde{f}(s,t)$ of the CWT. In the following, let $f \in L^2(\mathbb{R})$ be a signal and $\psi \in L^2(\mathbb{R})$ a wavelet with $\|\psi\| = 1$. As in Subsection 1, $t_0(\psi)$ denotes the center and $T(\psi)$ the width of ψ. Similarly, $\omega_0(\psi)$ denotes the center and $\Omega(\psi)$ the width of $\hat{\psi}$. We assume that all values are finite. Then, as in Subsection 2, the inner products

$$\langle f | \psi \rangle$$

can be interpreted as time-frequency analysis of f over the information cell $\mathrm{IC}(\psi)$ given by the parameters $t_0(\psi)$, $\omega_0(\psi)$, $T(\psi)$ and $\Omega(\psi)$.

The following theorem tells, how the information cell $\mathrm{IC}(\psi)$ behaves under translation and scaling of the wavelet ψ.

Theorem 7.7. *Let ψ be a wavelet with $\|\psi\| = 1$ and finite information cell $\mathrm{IC}(\psi)$. Then for all $t, s \in \mathbb{R}$, $s \neq 0$, holds*

$$t_0(\psi^{s,t}) = s \cdot t_0(\psi) + t \quad \text{and} \quad T(\psi^{s,t}) = |s| \cdot T(\psi), \tag{15}$$

as well as

$$\omega_0(\psi^{s,t}) = \frac{1}{s}\omega_0(\psi) \quad \text{and} \quad \Omega(\psi^{s,t}) = \frac{1}{|s|}\Omega(\psi). \tag{16}$$

Proof: Using $\|\psi\| = \|\psi^{s,t}\|$, the assertions follow by some straightforward substitution. We just prove the assertion for $t_0(\psi^{s,t})$ and leave the other cases as exercise.

$$\begin{aligned} t_0(\psi^{s,t}) &= \int_{\mathbb{R}} u|\psi^{s,t}(u)|^2 du \\ &= \frac{1}{|s|}\int_{\mathbb{R}} u\left|\psi\left(\frac{u-t}{s}\right)\right|^2 du \\ &= \frac{1}{|s|}\int_{\mathbb{R}} (vs+t)|\psi(v)|^2 \cdot |s| dv \\ &= s\int_{\mathbb{R}} v|\psi(v)|^2 dv + t\|\psi\|^2 \\ &= s \cdot t_0(\psi) + t \end{aligned}$$

□

The width of the information cell $\mathrm{IC}(\psi)$ w.r.t. time is proportional to the scaling parameter s, whereas w.r.t. frequency it is reciprocal to s. Hence, the shape of the information cells for a CWT depends on s whereas the area is invariant under scaling and translation. This is also illustrated by Figure 23.

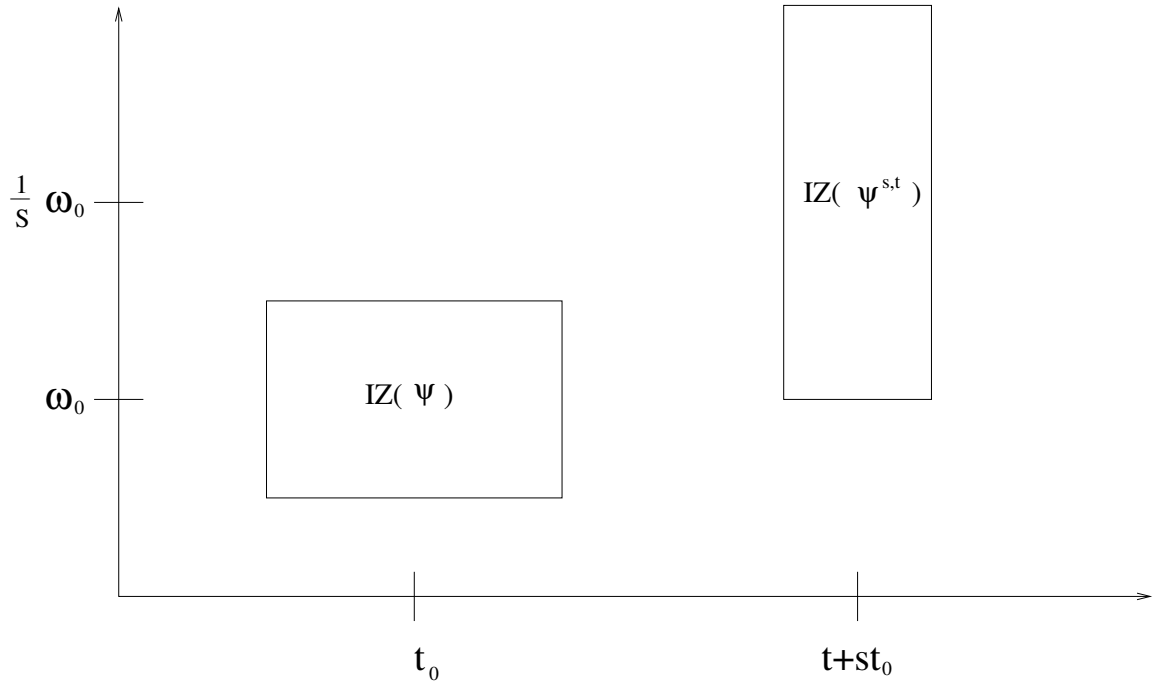


Figure 23: Information cells for ψ and $\psi^{s,t}$ with $s = \frac{1}{2}$.

We want to get an intuitive idea of this result.

- The scaling of a wavelet by a factor $s > 0$ is nothing else then compressing ($s < 1$) or expanding ($s > 1$) the wavelet.
- For increasing s the wavelet $\psi^{s,t}$ is expanded and the corresponding frequency spectrum is compressed. Intuitively, the number of oscillations per time unit decreases when stretching the wavelet.
- Contrary, for decreasing s the wavelet $\psi^{s,t}$ is compressed and the corresponding frequency spectrum is extended. Intuitively, the number of oscillations per time unit increases when compressing the wavelet.

These phenomena are reflected by the information cells. For a large parameter s, the information cell $\mathrm{IC}(\psi^{s,t})$ is wide-stretched in time having width sT and center $t + st_0$ and pressed in frequency having width $\frac{1}{s}\Omega$ and center $\frac{1}{s}\omega_0$.

In other words,

- for large s the inner products $\langle f|\psi^{s,t}\rangle$ analyze f w.r.t. long time segments and low frequencies picking up global and low-frequency changes in f.
- For small s the inner product $\langle f|\psi^{s,t}\rangle$ analyze f w.r.t. short time segments and high frequencies picking up sudden events of f such as peaks and high-frequency oscillations.

Principle: The scaling factor s is reciprocal to the frequency parameter ω. If the note ψ has pitch (i.e., frequency) ω_0, the note $\psi^{s,t}$ has frequency $\frac{\omega_0}{s}$.

Note 7.8. The above definition of localization in the frequency domain at ω_0 is not suitable for many wavelet transforms since for many common wavelets, $\hat{\psi}$ is an even (symmetric) function which have a peak at some positive and at the corresponding negative frequency. Therefore, let

$$\omega_0^+ := \int_0^\infty \omega |\hat{\psi}(\omega)|^2 d\omega \quad \text{und} \quad \omega_0^- := \int_{-\infty}^0 \omega |\hat{\psi}(\omega)|^2 d\omega.$$

Then, we say that ψ localizes at $(t_0, \omega_0^\pm)$. The parameter $\omega_0^\pm$ behaves under translation and scaling of the wavelet just as ω_0 and the interpretation remains the same. For further details we refer to [Louis/Maaß/Rieder, p.31].

7.4 Examples of some CWTs

7.4.1 CWT of some Chirp Signal

Figure 24 shows the time-scale representation of the CWT of a chirp signal f using a db4-wavelet. The representation is analogous to the time-frequency representation of the WFT in Subsection 6.2.2. The point at (t, s) has a gray color which is proportional to the values

$$|\tilde{f}(s,t)| = |\langle f|\psi^{s,t}\rangle|,$$

i.e., the larger the value the darker the color.

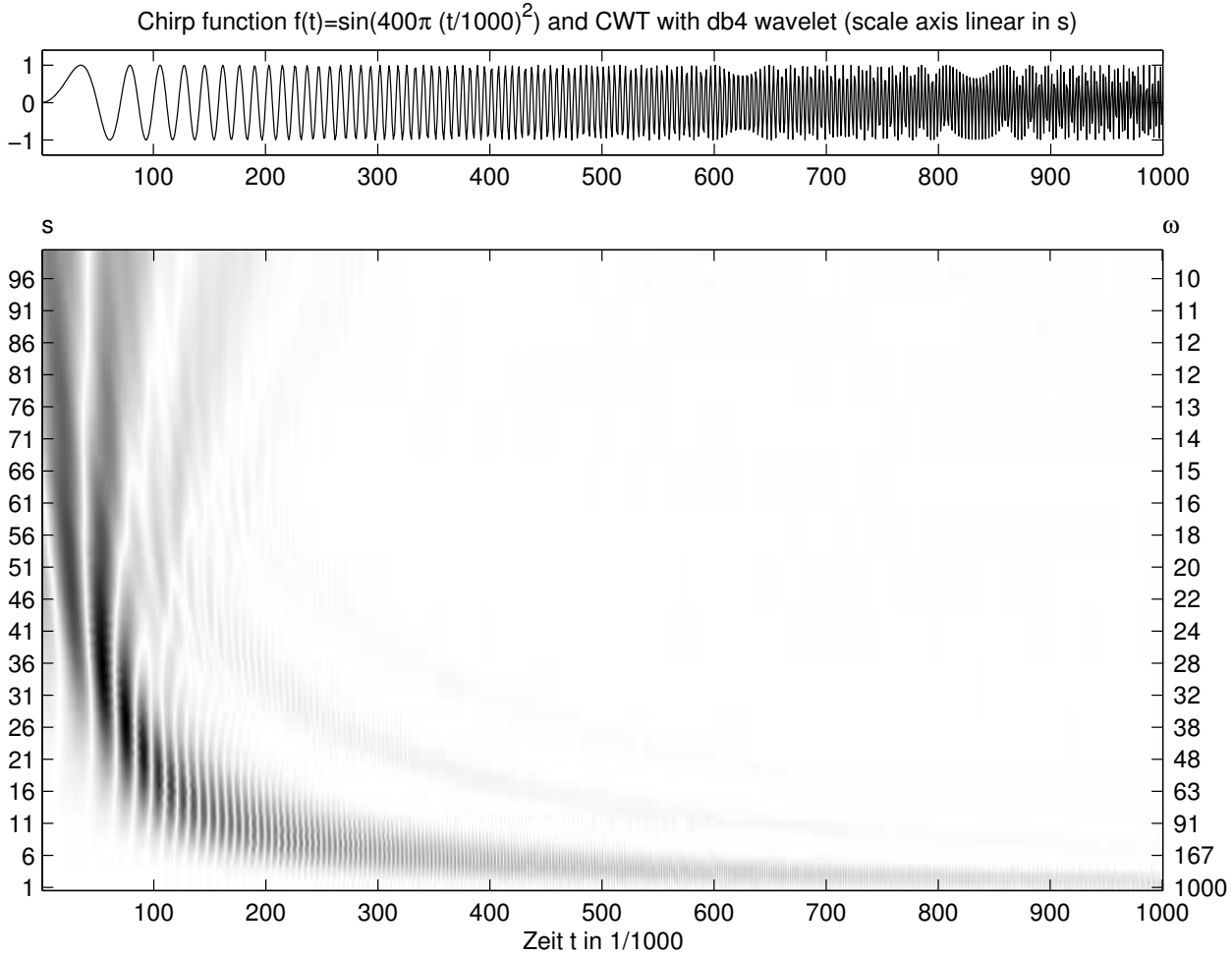


Figure 24: Time-scale representation of the CWT of some chirp function.

As was shown in the last section, the scaling parameter s is antiproportional to the analyzed frequency ω. Since the chirp function has low frequencies for small points of time t, there is a strong correlation between the wavelets $\psi^{s,t}$ with large scaling parameter s. Therefore, one has in Figure 24 dark points for small t and large s. With increasing time t the frequency ω of the chirp signal also increases. So, one expects large values $|\langle f|\psi^{s,t}\rangle|$ for decreasing s. Since the frequency of the chirp signal f increases linearly, there is a hyperbola like shaped dark area in the time-scale representation of f.

In Figure 24, the frequencies ω corresponding to s are shown at the right hand side of the time-scale representation. For a better comparison to the time-frequency representation (such as in the figures 13 Figure 16 of the WFT) we have reparameterized the scale axis such that the corresponding frequencies are linearly arranged (amounting to a hyperbola-like arrangement of the scale axis). This is shown in Figure 25. In this figure one again recognizes the diagonal analogously to the time-frequency representation of the WFT of Figure 13.

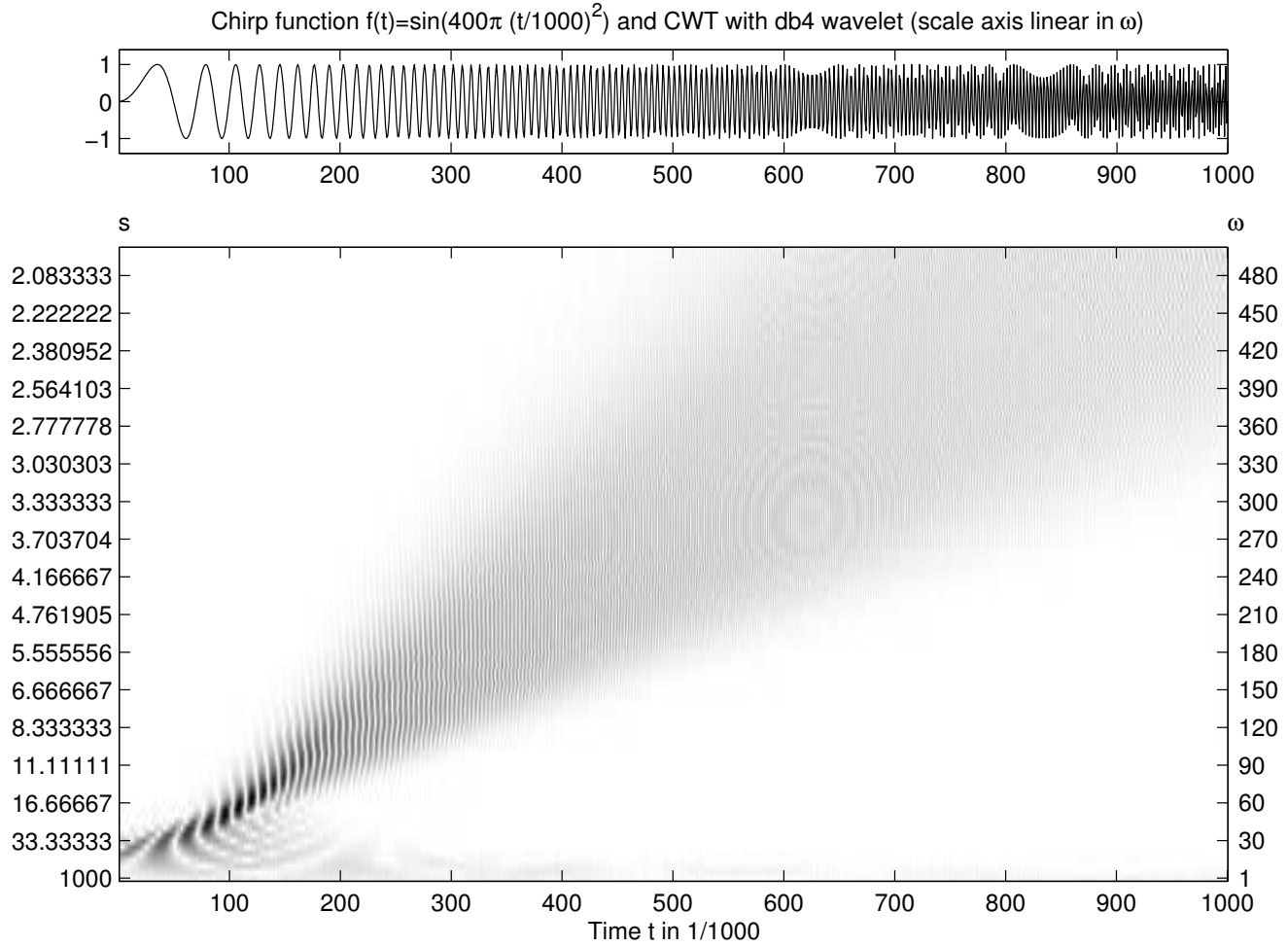


Figure 25: Reparameterized time-scale representation of Figure 24.

However, Figure 25 shows that, in particular for high frequencies contained in the chirp signal, the CWT does not localize as good as the WFT. In other words, in the CWT high frequencies are "smeared".

This effect becomes even more evident when one uses the discontinuous Haar wavelet instead of the db4 wavelets in the CWT analysis. This is illustrated by Figure 26. The discontinuities introduced by the Haar wavelet lead to parasitic frequencies in all frequency bands.

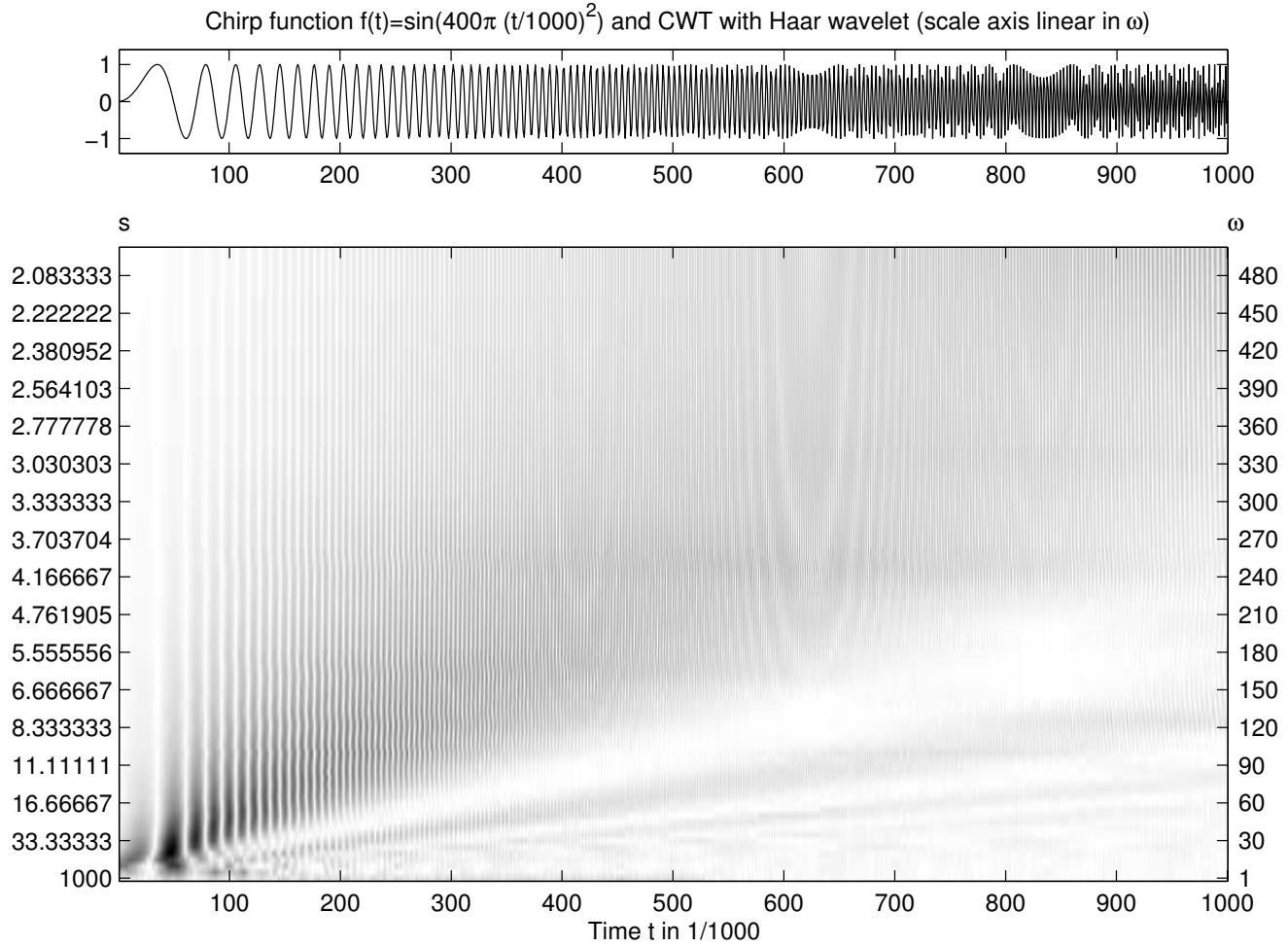


Figure 26: Smearing effects in the CWT using a Haar wavelet.

7.4.2 CWT of Sines with Impulses.

Analogous to the last subsection, Figure 27 shows the time-scale representation of the CWT w.r.t. the db4. This time the input signal is the superposition of two sines of frequency $\omega_1 = 10$ and $\omega_2 = 50$, respectively, and two additional impulses at $t_1 = 0.4$ and $t_2 = 0.5$. Figure 28 shows the reparameterized time-scale representation with a linear arrangement of the frequencies.

As in the time-frequency representation in Figure 16 one would expect two horizontal stripes corresponding to the two frequencies ω_1 and ω_2 of the sines. In Figure 27 this two stripes can be recognized — at least with some imagination — at the scale parameters around $s = 100$ and $s = 20$ which corespond to the frequencies $\omega_1 = 10$ and $\omega_2 = 50$. Also, the impulses can be recognized for small scale parameters s corresponding to high frequencies.

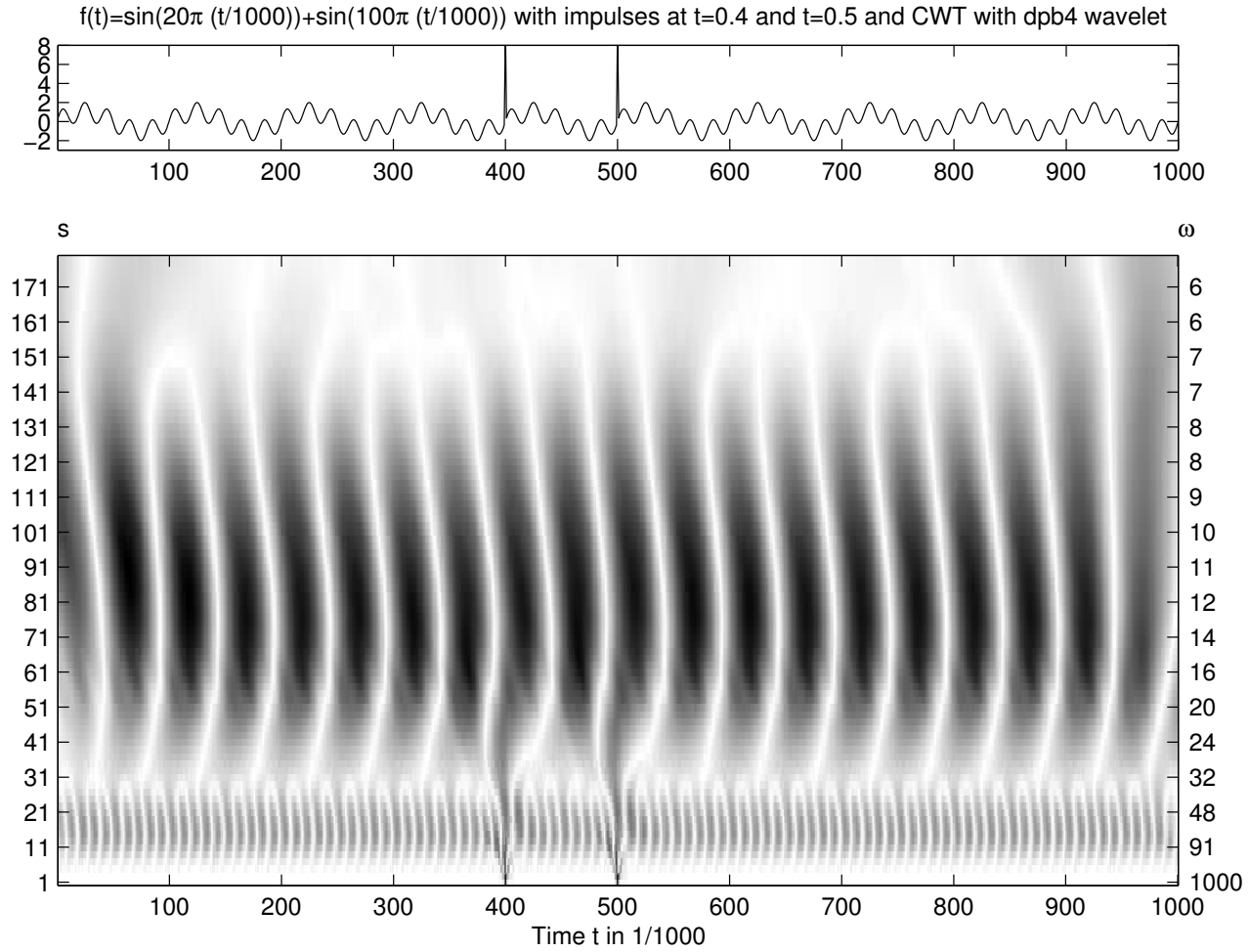


Figure 27: Time-scale representation of the CWT w.r.t. the db4 wavelet.

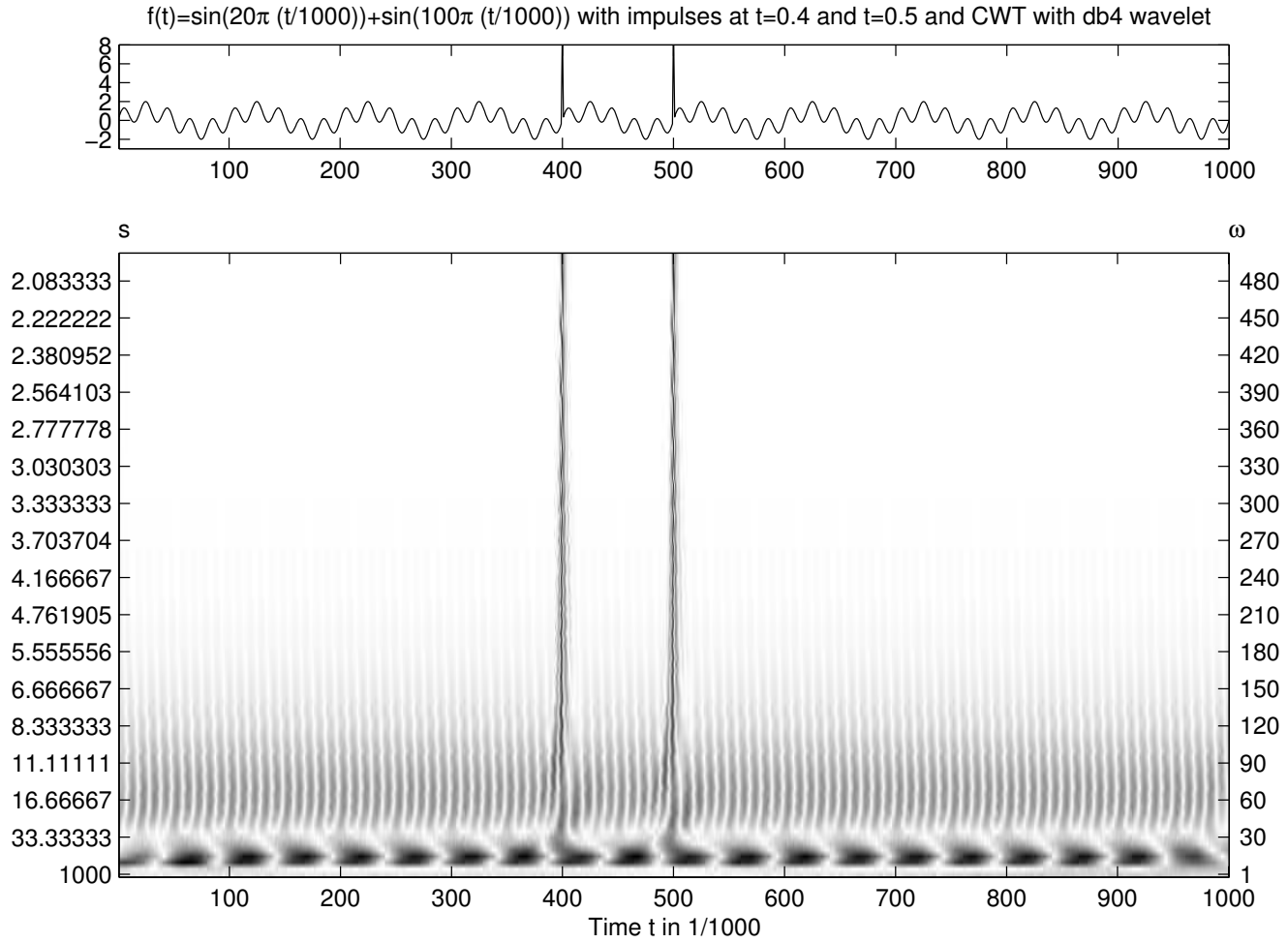


Figure 28: Reparameterized time-scale representation of Figure 27.

Note that the two horizontal stripes of Figure 27 and Figure 28, which correspond to the frequencies ω_1 and ω_2 of the sines, are smeared in frequency. Furthermore, these stripes are periodically interrupted by "vertical white stripes". In the following we give some explanation of such phenomena.

To this means, we fix a scale parameter $s = s_0$ and a point of time $t = t_1$ so that the absolute value of the wavelet coefficient $\langle f|\psi^{s_0,t_1}\rangle$ is large as, e.g., for $s_0 = 100$ and $t_1 = 0.575$. The wavelet ψ^{s_0,t_1} and the function f have then a high correlation at $t_0(\psi^{s_0,t_1}) = t_1$. In other words, f and ψ^{s_0,t_1} have the same sign at most points, i.e., ψ^{s_0,t_1} "nestles" against the function f (in our case against the low-frequency osciallation $\omega_1 = 10$ of f). This leads to a large wavelet coefficient $\langle f|\psi^{s_0,t_1}\rangle$. Figure 29 (a) illustrates this situation.

Next, we look at the point of time $t_2 = 0.625$ which corresponds to a translation of the wavelet ψ^{s_0,t_1} of half a period of the sine of frequency $\omega_1 = 10$. Then, ψ^{s_0,t_2} and f have opposite sign at most points as is illustrated in Figure 29 (b). Again the absolute value of the wavelet coefficient $\langle f|\psi^{s_0,t_2}\rangle$ is large. However, $\langle f|\psi^{s_0,t_2}\rangle$ and $\langle f|\psi^{s_0,t_1}\rangle$ have opposite sign.

Now, it is not difficult to see that the function $t \mapsto \langle f|\psi^{s_0,t}\rangle$ is continuous, which assumes positive values (such as for $t = t_1$) as well as negative values (such as for $t = t_2$). Therefore, this function has a zero at some point of time t_{zero} between t_1 and t_2. This is illustrated by Figure 29 (c).

This explains the periodic interruption by "vertical white stripes" (corresponding to neighborhoods of the zeros). Note that this periodic interruption takes place two times the period $\omega = 10$ of the low-frequency sine component of f.

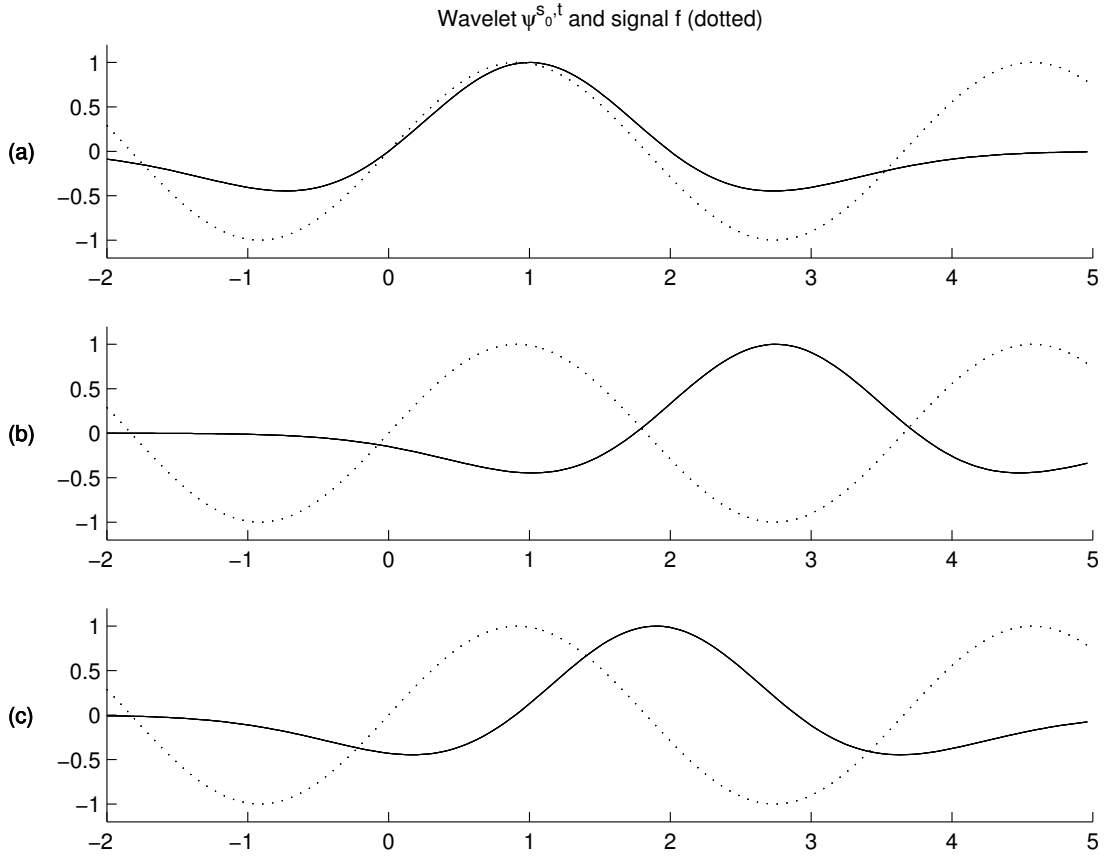


Figure 29: Strong positive (a), negative (b), and weak (c) correlation of the wavelets $\psi^{s_0,t}$ with the signal f.

7.4.3 Reconstruction of the Signal from its CWT

Let $f \in L^2(\mathbb{R})$ be a signal with CWT $\tilde{f}(s,t)$ w.r.t. a wavelet $\psi \in L^2(\mathbb{R})$ having admissability constant

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

In this section, we want to reconstruct the signal f from its CWT. Recall that, as explained in Section 7.3, the wavelet coefficient $\tilde{f}(s,t)$ can be interpreted as the intensity in which the note $\psi^{s,t}$ is contained in the signal f. If the note ψ is of frequency ω_0, the note $\psi^{s,t}$ has frequency $\frac{\omega_0}{s}$. Intuitively, the signal f should be representable as a superposition of the notes $\psi^{\frac{\omega_0}{\omega},t}$ of frequency ω weighted by $\tilde{f}(\frac{\omega_0}{\omega},t)$:

$$f(u) \sim \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{f}\left(\frac{\omega_0}{\omega}, t\right) \psi^{\frac{\omega_0}{\omega},t}(u) d\omega dt = \omega_0 \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{f}(s,t) \psi^{s,t}(u) \frac{1}{s^2} ds dt,$$

where the last equation follows by the substitution $s = \frac{\omega_0}{\omega}$. We will see that this reconstruction formula holds up to a constant normalizing factor.

In the following computations we need the Fourier transform $\widehat{\psi^{s,t}}$ of $\psi^{s,t}$. By (13) of Theorem 7.2 holds

$$\widehat{\psi^{s,t}}(\omega) \quad = \quad e^{-2\pi i\omega t}|s|^{\frac{1}{2}}\hat{\psi}(s\omega). \tag{17}$$

Using this formula we obtain by Parseval's equality

$$\begin{aligned} \tilde{f}(s,t) &= \langle f|\psi^{s,t}\rangle = \langle \hat{f}|\widehat{\psi^{s,t}}\rangle \\ &= \int_{\mathbb{R}} |s|^{\frac{1}{2}}\overline{e^{-2\pi i\omega t}}\overline{\hat{\psi}}(s\omega)\hat{f}(\omega)d\omega \\ &= |s|^{\frac{1}{2}}\left(\overline{\hat{\psi}}(s\cdot)\hat{f}(\cdot)\right)^{\vee}(t). \end{aligned} \tag{18}$$

Applying the Fourier transform on both sides one gets

$$\int_{\mathbb{R}} e^{-2\pi i\omega t}\tilde{f}(s,t)dt \quad = \quad |s|^{\frac{1}{2}}\overline{\hat{\psi}}(s\omega)\hat{f}(\omega) \tag{19}$$

Now, if we could isolate $\hat{f}(\omega)$ on the right hand side then f could be recovered by an inverse Fourier transform. However, it is not possible to simply divide by $|s|^{\frac{1}{2}}\overline{\hat{\psi}}(s\omega)$ since this value could vanish. Instead, we multiply both sides of (19) with $\hat{\psi}(s\omega)|s|^{-\frac{3}{2}}$ and integrate over s:

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \hat{\psi}(s\omega)|s|^{-\frac{3}{2}} e^{-2\pi i\omega t}\tilde{f}(s,t)dsdt = \int_{\mathbb{R}} |s|^{-1}|\hat{\psi}(s\omega)|^2\hat{f}(\omega)ds \tag{20}$$

Now the admissibility condition $0 < c_\psi < \infty$ comes into play, where

$$c_\psi = \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi)|^2}{|\xi|}d\xi = \int_{\mathbb{R}} \frac{|\hat{\psi}(s\omega)|^2}{|s|}ds$$

using the substitution $\xi = s\omega$.

Dividing both sides of (20) by c_ψ and using (17) one gets the following reconstruction formula of $\hat{f}$:

$$\begin{aligned}\hat{f}(\omega) &= \frac{1}{c_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{f}(s,t)\hat{\psi}(s\omega)|s|^{-\frac{3}{2}}e^{-2\pi i\omega t}dsdt \\ &= \frac{1}{c_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{f}(s,t)\widehat{\psi^{s,t}}(\omega)|s|^{-2}dsdt. \end{aligned} \tag{21}$$

Applying the inverse Fourier transform one obtains the reconstruction formula of the signal f:

$$\begin{aligned} f(u) &= \frac{1}{c_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{f}(s,t)\widehat{\psi^{s,t}}(\omega)e^{2\pi i\omega u}\frac{1}{s^2}d\omega dsdt \\ &= \frac{1}{c_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{f}(s,t)\psi^{s,t}(u)\frac{1}{s^2}dsdt. \end{aligned} \tag{22}$$

The result is summarized in the following theorem.

Theorem 7.9. *Let ψ be a wavelet, i.e., $\psi \in L^2(\mathbb{R})$ satisfying the admissability condition*

$$0 < c_\psi := \int_{\mathbb{R}} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty.$$

Then any signal $f \in L^2(\mathbb{R})$ can be reconstructed from its CWT $\tilde{f}(s,t)$ by

$$f(u) = \frac{1}{c_\psi} \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{f}(s,t) \psi^{s,t}(u) \frac{1}{s^2} ds dt.$$

7.5 Discrete Version of the CWT

In the previous sections we have introduced the continuous version of the wavelet transform. This case is important in order to get the "right" intuition of how to interpret wavelet coefficients. In the following we discuss some of the problems which arise when one actually wants to compute the wavelet transform. Thus we are faced with the questions of

(1) how to discretize the continuous wavelet transform,

(2) how to compute this discretized wavelet transform efficiently, and

(3) how to compute the reconstruction of a signal from its discretized wavelet transform (inversion) efficiently.

7.5.1 CWT-Adapted Grid

By Theorem 7.9, a signal $f \in L^2(\mathbb{R})$ can by reconstructed by the synthesis formula

$$f(u) = \frac{1}{c_\psi} \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{f}(s,t) \psi^{s,t}(u) \frac{1}{s^2} ds dt,$$

$\tilde{f}$ is the CWT w.r.t. some wavelet ψ. As discussed in Section 6.4 for the WFT, this representation is in general rather redundant: a one-dimensional parameter space (the time denoted by the variable u) is represented by an integral over a two-dimensional parameter space (the time-scale domain represented by the variables t and s). In the following, we discuss a discrete version of the CWT where a discrete (or even finite) set of values $\tilde{f}(s,t)$ is sufficient for the reconstruction of the signal f.

Warning: The word "discrete" does not refer to the wavelet itself, which always will be a time-continuous function in $L^2(\mathbb{R})$. It refers to the discrete number of points in the time-scale parameter space on which the CWT is evaluated.

Of course, the discrete subset in $\mathbb{R} \setminus \{0\} \times \mathbb{R}$ on which the CWT is to be evaluated depends on the wavelet. For example, in the case of the discrete version of the WFT in Section 6.4, where we only considered window functions with compact support, the mesh of the discrete grid in the time-frequency plane depended on the respective window function.

A similar assertion holds for the CWT. However, in this case we do not consider an equidistant grid, but a grid of the form

$$\{(a^m, nba^m)|m, n \in \mathbb{Z}\} \subset \mathbb{R} \setminus \{0\} \times \mathbb{R} \tag{23}$$

with $a > 1, b > 0$ and the corresponding set of functions

$$\left\{ \psi_{m,n}^{(a,b)} := \psi^{a^m, nba^m} = a^{-m/2} \psi(a^{-m} \cdot -nb) | m, n \in \mathbb{Z} \right\}.$$

The form of this grid is motivated by the time-frequency interpretation of the CWT in Section 7.3. Let ψ be a wavelet with $\|\psi\| = 1$ and let $\mathrm{IC}(\psi)$ be the corresponding information cell with center $(t_0(\psi), \omega_0(\psi))$, width $T(\psi)$, and height $\Omega(\psi)$. We consider the case $t_0 = 0$ und $\omega_0 > 0$. Then, by Equation (15), the center of the information cell $\mathrm{IC}(\psi_{m,n}^{a,b})$ of the note $\psi_{m,n}^{a,b}$ is given by

$$(a^{-m}\omega_0, nba^m)$$

and, by Equation (16), the width and height are given by

$$T(\psi_{m,n}^{(a,b)}) = a^m T(\psi) \quad \text{bzw.} \quad \Omega(\psi_{m,n}^{(a,b)}) = a^{-m}\Omega(\psi).$$

With increasing frequency ($m \to -\infty$) the time component of the centers of the functions $\{\psi_{m,n}^{(a,b)} | n \in \mathbb{Z}\}$ are closer together. This is also illustrated by Figure 30). Therefore, the grids as defined by Equation (23) model the zoom effect of the CWT.

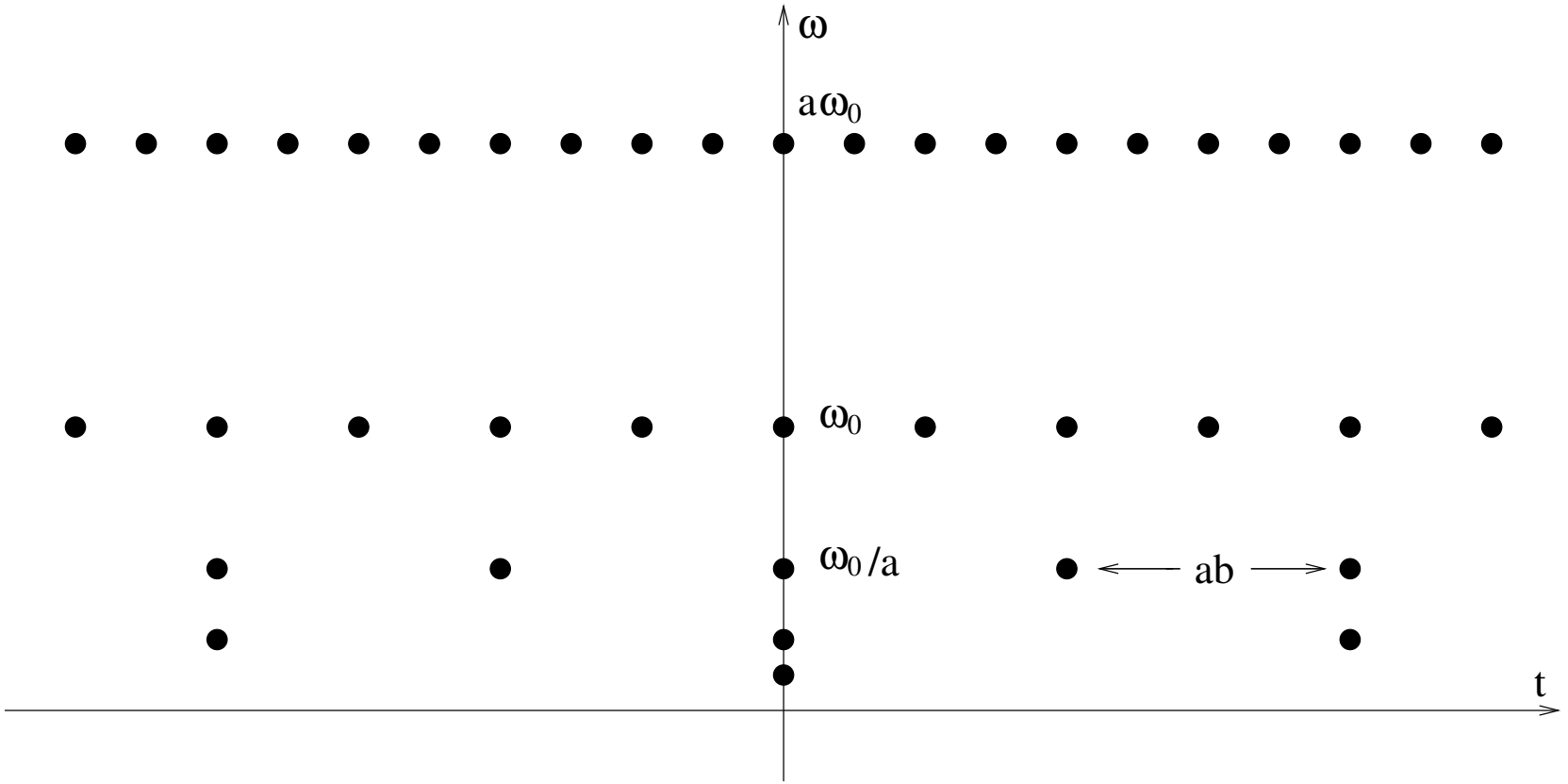


Figure 30: Distribution of the points $(nba^m, a^{-m}\omega_0)$ for $a = 2$ and $b = 1$ in the time-frequency plane.

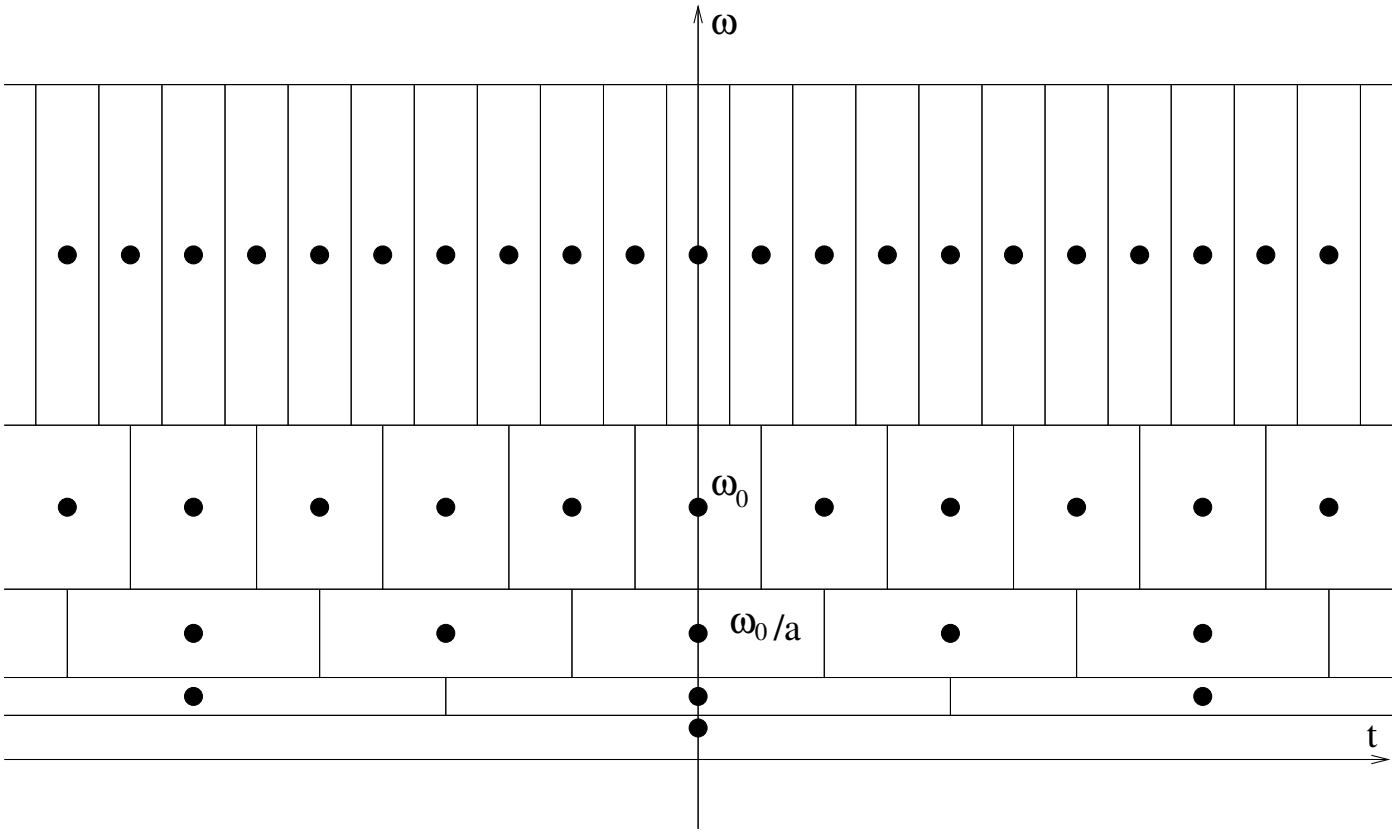


Figure 31: Ideal tiling of the time-frequency plane by the information cells of a CWT corresponding to the points in the grid of Figure 30.

7.5.2 Wavelet Frames

The investigation of the wavelet transform on grids as defined by Equation (23) leads to so-called wavelet frames which are the content of this subsection.

Definition 7.10. Let $a > 1, b > 0$ and $\psi \in L^2(\mathbb{R})$. The family of functions

$$\left(\psi_{m,n}^{(a,b)}\right)_{m,n\in\mathbb{Z}}$$

forms a so-called wavelet frame for $L^2(\mathbb{R})$, if there are constants $A, B > 0$ such that for all $f \in L^2(\mathbb{R})$ one has

$$A\|f\|^2 \leq \sum_{m,n\in\mathbb{Z}} |\langle f|\psi_{m,n}^{(a,b)}\rangle|^2 \leq B\|f\|^2. \tag{24}$$

One says that (ψ, a, b) generates the frame. The constants A and B are called frame boundaries. In the case $A = B$ the frame is called tight.

The next theorem shows that the concept of wavelet frames generalizes the concept of ON bases.

Theorem 7.11. *A tight wavelet frame (ψ, a, b) with frame boundaries $A = B = 1$ and $\|\psi\| = 1$ forms an ON basis $L^2(\mathbb{R})$.*

Proof: Let (ψ, a, b) be a tight wavelet frame. Then

$$\|\psi_{m,n}^{(a,b)}\| = \|\psi\| = 1$$

for all $m, n \in \mathbb{Z}$. Since $A = B = 1$, one has equality in (24) and setting $f = \psi_{\mu,\nu}^{(a,b)}$, $\mu, \nu \in \mathbb{Z}$, one gets

$$\begin{aligned} \|\psi_{\mu,\nu}^{(a,b)}\|^2 &= \sum_{m,n\in\mathbb{Z}} |\langle \psi_{\mu,\nu}^{(a,b)} | \psi_{m,n}^{(a,b)} \rangle|^2 \\ &= \|\psi_{\mu,\nu}^{(a,b)}\|^2 + \sum_{m,n\in\mathbb{Z},(m,n)\neq(\mu,\nu)} |\langle \psi_{\mu,\nu}^{(a,b)} | \psi_{m,n}^{(a,b)} \rangle|^2 \end{aligned}$$

Therefore, it follows that $\langle \psi_{\mu,\nu}^{(a,b)} | \psi_{m,n}^{(a,b)} \rangle = 0$ for $(m,n) \neq (\mu,\nu)$. This proves the orthonormality. Completeness of the family follows from Parseval's equation and (24) with $A = B = 1$. □

In general, it can be rather difficult to prove if the frame conditions are fulfilled or not. For sufficient conditions we refer to [Louis/Maaß/Rieder]. In this book it is also shown that the Meyer wavelet ψ (see Figure 21) forms a tight wavelet frame with $a = 2$ and $b = 1$. Then, Theorem 7.11 implies the following result.

Corollary 7.12. *The Meyer wavelet ψ generates a frame $(\psi, 2, 1)$ which forms an ON basis of $L^2(\mathbb{R})$.*

Each wavelet frame (ψ, a, b) defines a linear operator

$$T : L^2(\mathbb{R}) \quad \rightarrow \quad \ell(\mathbb{Z} \times \mathbb{Z})$$

$$f \quad \mapsto \quad (\langle f | \psi_{m,n}^{(a,b)} \rangle)_{m,n \in \mathbb{Z}}.$$

This operator fulfills

$$A^{\frac{1}{2}} \|f\|_{L^2} \leq \|Tf\|_{\ell^2} \leq B^{\frac{1}{2}} \|f\|_{L^2}. \tag{25}$$

Therefore, T is continuous since its norm is bounded by

$$\|T\| := \sup_{\|f\|=1} \|Tf\| \leq B^{\frac{1}{2}}.$$

Furthermore, from $A > 0$ follows that T is injective and continuously invertible on its image since

$$\|T^{-1}|_{\mathrm{Image}\, T}\| \leq A^{-\frac{1}{2}}.$$

In this case f can be reconstructed from the discrete set of values $Tf = (\langle f|\psi_{m,n}^{(a,b)}\rangle)_{m,n\in\mathbb{Z}}$.

Next, we explain how this reconstruction can be achieved in a more general framework. let H be a (separable) Hilbert space and $\Phi = (\varphi_j)_{j\in\mathbb{Z}}$ a frame, i.e.,

$$\exists A, B > 0 \forall v \in H: \quad A\|v\|^2 \leq \sum_{j\in\mathbb{Z}} |\langle v|\varphi_j\rangle|^2 \leq B\|v\|^2. \tag{26}$$

For given Φ we define the frame operator $S: H \longrightarrow H$ by

$$Sv := \frac{2}{A+B} \sum_{j\in\mathbb{Z}} \langle v|\varphi_j\rangle \varphi_j.$$

Lemma 7.13. *The frame operator S is a positive, bounded, continuously invertible operator with*

$$\frac{2A}{A+B}\|v\|^2 \leq \langle Sv|v\rangle \leq \frac{2B}{A+B}\|v\|^2 \tag{27}$$

for all $v \in H$. Furthermore,

$$\|I - S\| \leq \rho := \frac{B-A}{A+B} < 1. \tag{28}$$

Sketch of proof: First, we show that S is well defined. For $N \in \mathbb{N}$ define the operator S_N by

$$S_N(v) := \sum_{|j|\leq N} \langle v|\varphi_j\rangle\varphi_j, \quad v \in H.$$

Using the Cauchy-Schwarz inequality and the frame boundary B one can show that the sequence $(S_N)_{N\in\mathbb{N}}$ is a Cauchy sequence which converges to a continuous operator which is just S. The inequality in (27) follows from the definition of S and the frame condition (26).

Now, we show (28). From (27) follows

$$\left(1 - \frac{2B}{A+B}\right) \|v\|^2 \leq \langle (I-S)v|v\rangle \leq \left(1 - \frac{2A}{A+B}\right) \|v\|^2$$

and therefore

$$|\langle (I-S)v|v\rangle| \leq \frac{B-A}{A+B}\|v\|^2 = \rho\|v\|^2.$$

From this follows (28) using the well-known fact from functional analysis that

$$\|I-S\| = \sup_{\|v\|=1} |\langle (I-S)v|v\rangle|.$$

For further details we refer to the literature. □

Important for the construction of the inverse S^{-1} of S is property (28). Since $\rho < 1$, the so-called Neumann series

$$\sum_{k=0}^{\infty} (I - S)^k$$

converges towards S^{-1}. Therefore, the sequence

$$v_n := \sum_{k=0}^{n} (I - S)^k Sv$$

converges to v. Furthermore, $(v_n)_{n\in\mathbb{N}}$ satisfies the recursion

$$v_{n+1} = Sv + (I - S) \sum_{k=1}^{n+1} (I - S)^{(k-1)} Sv = Sv + (I - S)v_n \tag{29}$$

and the error estimate

$$\|v - v_n\| = \|(I - S)^{n+1} v\| \leq \rho^{n+1} \|v\|. \tag{30}$$

The recursion (29) is a common algorithm for the approximation of S^{-1}. If the frame boundaries A and B are close together, i.e., $B/A \approx 1$, then $\rho \approx 0$ and only a few steps in the recursion are sufficient to guarantee a good approximation of v.

We come back to our specific situation. We have discussed that there exists wavelet frames (ψ, a, b) and that in this case a signal $f \in L^2(\mathbb{R})$ can be perfectly recovered from ths wavelet coefficients

$$\tilde{f}(a^{-m}, nba^{-m}), m, n \in \mathbb{Z}.$$

If one just uses a finite number of grid points (the grid points which correspond to the "large" wavelet coefficients $\tilde{f}(a_0^{-m}, nb_0a_0^m)$), then at least an approximation of the original signal f can be reconstructed from this finite set of wavelet coefficients. Then the following primitive "algorithm" can be used to compute a discrete wavelet transform:

Pick a suitable wavelet frame (ψ, a, b) and compute some approximation of the wavelet coefficients

$$\tilde{f}(a^m, nba^m) = \langle f | \psi_{m,n}^{(a,b)} \rangle$$

for all "essential" grid points (a finite number). The approximation of the wavelet coefficients could, for example, be done by computing suitable Riemann sums.

This "algorithm" is in general neither efficient nor practicable.

- The approximation of the relevant integrals $\langle f | \psi_{n,m}^{(a_0,b_0)} \rangle$ is rather expensive.
- Furthermore, the choice of the "essential" grid points depends on the signal to be analyzed.
- It is not clear how to determine, a priori, which grid points turn out to be essential.

In the year 1986 the connection between the so-called multiresolution analysis and wavelet transforms was discovered which led to completely new insights into the theory of wavelets. Based on these ideas fast algorithms for the discrete wavelet transform as well as its reconstruction were found. These recursive algorithms, which will be the content of the next chapter, are well suited in view of efficient implementations.

Chapter 8: Multiresolution Analysis (MRA) and Wavelet Transform

Again we cite as an introduction Barbara Burke Hubbard who gives in [Hubbard] the following nice account of multiresolution.

> Coming from different directions — pure mathematics and computer vision — Stéphane Mallat und Yves Meyer created multiresolution theory. One result of their work was the fast wavelet transform. Another result was a mathematical theory of orthogonal wavelets. This theory defines four mathematical conditions which, if met, make it possible to view a signal at resolutions differing by a factor of two, and to encode the difference of information from one resolution to the next as orthogonal wavelet coefficients.

The theory also gives a recipe for creating new orthogonal wavelet bases. Both scaling functions and wavelets can be created using a function that is the Fourier transform of a low-pass filter. (In geometrical terms, this function parameterizes a curve on a sphere in four-dimensional space.) Curiously, the bottom line of multiresolution theory is that to compute the wavelet transform of a signal we need neither scaling functions nor wavelets: just very simple digital filters.

In this chapter we first introduce the concept of multiresolution, describe their links to wavelet theory, and then discuss the resulting fast wavelet algorithms (FDWT and FIDWT).

8.1 Multiresolution Analysis (MRA)

8.1.1 Motivating Analogy

We first want to illustrate the main idea of the multiresolution analysis (MRA) by means of some MRA-like situation. We want to emphasize that this analogy only holds w.r.t. some aspects of MRA and exclusively serves as motivation.

For this purpose, we consider the binary representation of a real number $f \in \mathbb{R}$ (the symbol f should remind on a signal) in the following way. Let

$$V_m := 2^m \mathbb{Z} \quad \text{for} \quad m \in \mathbb{Z}.$$

Then the following properties hold:

$$\{0\} \subset \ldots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \ldots \subset \mathbb{R},$$

$$\overline{\bigcup_{m\in\mathbb{Z}} V_m} = \mathbb{R},$$

$$\bigcap_{m\in\mathbb{Z}} V_m = \{0\},$$

$$f \in V_m \iff f \cdot 2^{-m} \in V_0.$$

If the real number f has the binary representation

$$f = \sum_{n \in \mathbb{Z}} w_n 2^n, \quad w_n \in \{0, 1\},$$

then the "projection" $P_i f$ of f onto V_i is given by

$$P_i f := \sum_{n \geq i} w_n 2^n.$$

In other words, setting all digits of f up to position $i - 1$ to zero results in the number $P_i f$. Increasing i results in a "coarser resolution" $P_i f$ of f and, contrary, decreasing i results in a "finer resolution". In practice, the computation of the binary representation (at least some approximation with some prescribed precision) is the goal. One starts with some coarse approximation given by some $P_i f$ (often $i = 0$). The one computes successively the numbers $w_n, n = i, i + 1, i + 2, \ldots$. Doing so, in each step from j to $j + 1$ one records the "detail" $w_j 2^j$, $w_j \in \{0, 1\}$.

8.1.2 Definition of the MRA

Before we give the mathematical definition of a multiresolution analysis (MRA) we present the main underlying idea. Suppose we want to split up a signal f, which lies in some linear subspace $V_{-1} \subset L^2(\mathbb{R})$, into some high-frequency and some low-frequency component. The smooth and low-frequency component is described by some orthogonal projection $P_0 f$ into some smaller subspace $V_0 \subset V_{-1}$ which contains the "smooth" functions of V_{-1}. We denote the orthogonal complement of V_0 in V_{-1} by W_0. This linear subspace contains by construction the "jagged" (high-frequency) signals. Denoting the projection of f onto W_0 by $Q_0 f$, then

$$\begin{aligned} f &= P_0 f + Q_0 f, \\ V_{-1} &= V_0 \oplus W_0. \end{aligned}$$

Analogously, we proceed with $P_0 f$ by splitting up V_0 into some orthogonal sum of subspaces V_1 ("smooth" elements) and W_1 ("jagged" elements). The respective projections are denoted by P_1 and Q_1. Since $P_1 P_0 f = P_1 f$ as well as $Q_1 Q_0 f = Q_1 f$ we get

$$P_0 f = P_1 f + Q_1 f$$

and therefore

$$f = P_1 f + Q_1 f + Q_0 f. \tag{31}$$

Proceeding recursively, P_1 is split up into $P_2 f$ and Q_2, and so on

$$\begin{array}{ccccccccccc} L^2(\mathbb{R}) & \ldots & \longrightarrow & V_{-1} & \xrightarrow{P_0} & V_0 & \xrightarrow{P_1} & V_1 & \ldots & \longrightarrow & \{0\} \\ & & & & \overset{Q_0}{\searrow} & & \overset{Q_1}{\searrow} & & & & \\ & & & \ldots & & W_0 & & W_1 & \ldots & & \end{array}$$

The projections can be interpreted as follows:

- $P_1 f$ represents the low-frequency components of f representing the details of f below some specific size.
- $Q_0 f$ and $Q_1 f$ "contain" components of f corresponding to certain a frequency band where $Q_0 f$ represents higher-frequency details of f than $Q_1 f$.
- Equation (31) can be thought of as a decomposition of the signal corresponding to frequency bands consisting of high frequencies components and a mixture of lower frequencies.

For an illustration we refer to Figure 34 and Figure 35.

This decomposition process can mathematically modeled by the multiresolution analysis.

Definition 8.1. A multiresolution analysis (MRA) of $L^2(\mathbb{R})$ is a sequence of closed subspaces $V_m \subset L^2(\mathbb{R})$

$$\{0\} \subset \ldots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \ldots \subset L^2(\mathbb{R}), \tag{32}$$

such that the following holds:

$$\overline{\bigcup_{m\in\mathbb{Z}} V_m} = L^2(\mathbb{R}), \tag{33}$$

$$\bigcap_{m\in\mathbb{Z}} V_m = \{0\}, \tag{34}$$

$$f(\cdot) \in V_m \iff f(2^m \cdot) \in V_0. \tag{35}$$

Furthermore, one postulates the existence of some function $\varphi \in V_0$, the so-called scaling function, whose integral translates form an ON-basis of V_0, i.e.,

$$V_0 = \overline{\operatorname{span}\{\varphi(\cdot - k) | k \in \mathbb{Z}\}}. \tag{36}$$

The easiest example for an MRA is the so-called Haar-MRA which will be presented in Section 8.3 In the following we discuss Definition 8.1 and summarize some immediate consequences.

(1) Conditions (33) and (34) are fulfilled by many families of subspaces $(V_m)_{m\in\mathbb{Z}}$. An MRA is distinguished by Condition (35): the spaces V_m are scaled versions of the base space V_0, spanned by the translations of the scaling function φ (36).

(2) For increasing m ($m \to \infty$) the functions in V_m become broad, details are bloat. Contrary, when decreasing m ($m \to -\infty$), the functions in V_m become jagged containing more and more details. Let P_m denote the orthogonal projector onto V_m. Then the limits

$$\lim_{m\to+\infty} \|P_m f\| = 0 \quad \text{and} \qquad \lim_{m\to-\infty} \|P_m f - f\| = 0$$

state this interpretation more precisely. One says that $P_m f$ is the representation of f on "scale" V_m and contains all details of f up to size 2^m.

(3) Because of Condition (36), the space V_0 is translation invariant, i.e.,

$$f \in V_0 \Longleftrightarrow f(\cdot - k) \in V_0 \quad \text{for} \quad k \in \mathbb{Z}.$$

From (35) follows

$$f \in V_m \Longleftrightarrow f(\cdot - 2^m k) \in V_m \quad \text{for} \quad k \in \mathbb{Z}.$$

(4) The space V_m is spanned by the functions

$$\varphi_{m,k}(x) := 2^{-m/2}\varphi(2^{-m}x - k),$$

in fact, $(\varphi_{m,k})_{m\in\mathbb{Z}}$ is an ON-basis of V_m:

$$V_m = \overline{\operatorname{span}\{\varphi_{m,k} | k \in \mathbb{Z}\}}.$$

This follows from (35) and (36). Furthermore holds $\|\varphi_{m,k}\| = \|\varphi\|$.

(5) In the definition of an MRA one often encounters the condition, that the translates of the scaling function φ form a so-called Riesz basis. This seems to be a weaker condition then the condition of forming an ON-basis as stated in Definition 8.1. However, one can show that, starting with a Riesz basis, one can construct an ON-basis with the same properties. Hence, our definition does not mean any restriction in generality.

As in the introductory explanation of the main idea, we define the space W_m as the orthogonal complement of V_m in V_{m-1},

$$V_{m-1} = W_m \oplus V_m, \quad V_m \perp W_m, \tag{37}$$

and the operators Q_m as orthogonal projection from $L^2(\mathbb{R})$ onto W_m,

$$P_{m-1} = Q_m + P_m. \tag{38}$$

Then

$$V_m = \bigoplus_{j \geq m+1} W_j \quad \text{and therefore} \quad L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j. \tag{39}$$

The spaces W_m inherit the scaling property of the spaces V_m (see (35)):

$$f(\cdot) \in W_m \Longleftrightarrow f(2^m \cdot) \in W_0. \tag{40}$$

Using Equation (39), one sees that a function $f \in L^2(\mathbb{R})$ can be decomposed as

$$f = \sum_{j \in \mathbb{Z}} Q_j f = \sum_{j \geq m+1} Q_j f + \sum_{j \leq m} Q_j f = P_m f + \sum_{j \leq m} Q_j f. \tag{41}$$

This decomposition of f explains the denotation of multiresolution analysis. $P_m f$ represents f on scale m, which corresponds to applying a low-pass filter on f having a decreasing cut-off frequency for increasing m. The remaining high-frequency component is split into various frequency bands $Q_j f, -\infty < j \leq m$. In doing so, Q_j contains the details by which P_{j-1} differs from P_j:

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

8.2 MRA und Wavelets

At first sight, there does not seem to be any connection between the definition of the MRA and the wavelet theory introduced in the last chapter. We present the main result of this section, which appears like a miracle:

Main Result: To each MRA there is a wavelet ψ whose translated and dilated versions

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

for a fixed $m \in \mathbb{Z}$ form an orthonormal basis (ONB) of the space W_m. Furthermore, the wavelet can be constructed explicitly from the scaling function φ.

Surprisingly, digital filters

$$(h_k)_{k\in\mathbb{Z}} \quad \text{and} \quad (g_k)_{k\in\mathbb{Z}}$$

come into play when constructing such a wavelet ψ from the scaling function φ of the MRA. These filters do not only establish the connection between MRAs and wavelet theory but also between time-continuous and time-discrete signal processing.

The following diagram give an overview:

MRA

$((V_m), \varphi)$

$\varphi \in V_0 \subset V_{-1}$

$\Downarrow$

Scaling equation

$\varphi = \sum_{k\in\mathbb{Z}} h_k \varphi_{-1,k}$

$\Downarrow$

Scaling coefficients

$(h_k)_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$

$\Longrightarrow$

Wavelets

ψ defines

a wavelet

$\Uparrow$

Wavelet equation

$\psi := \sum_{k\in\mathbb{Z}} g_k \varphi_{-1,k} \in V_{-1}$

$\Uparrow$

Wavelet coefficients

$(g_k)_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$

$g_k := (-1)^k \bar{h}_{1-k}$

Digital filters

8.2.1 Filter Coefficients and Scaling Equation

By Definition 8.1 an MRA has a scaling function φ whose translates form on ON-basis of the base space V_0. In combination with the other axioms we derive in the following further properties of φ.

Lemma 8.2. *The scaling function φ satisfies the so-called scaling equation, i.e., there is a sequence $(h_k)_{k\in\mathbb{Z}}$ of complex numbers with*

$$\varphi(x) = \sqrt{2}\sum_{k\in\mathbb{Z}} h_k \varphi(2x-k). \tag{42}$$

Proof: Noting the inclusion $\varphi \in V_0 \subset V_{-1} = \overline{\operatorname{span}\{\sqrt{2}\varphi(2\cdot -k)|k\in\mathbb{Z}\}}$ the lemma follows immediately. □

Note 8.3. In the scaling equation

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

lies the key for the construction of orthonormal wavelet bases as well as for fast wavelet algorithms. The coefficients

$$(h_k)_{k \in \mathbb{Z}}$$

will come to the fore whereas the scaling function φ will only stay in the background.

Lemma 8.4. *The delated and translated versions $\varphi_{m,k}$ of the scaling function φ satisfies the scaling equation*

$$\varphi_{m,k} = \sum_{j\in\mathbb{Z}} h_j \varphi_{m-1,2k+j}. \tag{43}$$

Proof: This follows directly from the scaling equation (42) and the definition of $\varphi_{m,k}$:

$$\begin{aligned}
\varphi_{m,k}(x) &= 2^{-m/2}\varphi(2^{-m}x - k) \\
&\overset{(42)}{=} 2^{-m/2}2^{1/2}\sum_{j\in\mathbb{Z}} h_j \varphi(2(2^{-m}x - k) - j) \\
&= 2^{-(m-1)/2}\sum_{j\in\mathbb{Z}} h_j \varphi(2^{-(m-1)}x - (2k + j)) \\
&= \sum_{j\in\mathbb{Z}} h_j \varphi_{m-1,2k+j}(x).
\end{aligned}$$

□

Lemma 8.5. *The coefficients $(h_k)_{k\in\mathbb{Z}}$ of the scaling equation (42) satisfy the following orthogonality relations:*

$$\sum_{k\in\mathbb{Z}} h_{k+2j}\bar{h}_k = \delta_{0,j}. \tag{44}$$

Proof: This follows directly from the orthogonality relations satisfied by the translates of the scaling functions:

$$\begin{aligned}
\delta_{0,j} &= \langle\varphi(\cdot)|\varphi(\cdot+j)\rangle \\
&\overset{(43)}{=} \langle\sum_{l\in\mathbb{Z}} h_l\varphi_{-1,l}|\sum_{k\in\mathbb{Z}} h_k\varphi_{-1,2j+k}\rangle \\
&= \sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} h_l\bar{h}_k\langle\varphi_{-1,l}|\varphi_{-1,2j+k}\rangle \\
&= \sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} h_l\bar{h}_k\delta_{l,k+2j} \quad = \quad \sum_{k\in\mathbb{Z}} h_{k+2j}\bar{h}_k
\end{aligned}$$

□

One can show that $(h_k)_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$. In the following we want to assume the stronger condition $(h_k)_{k\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$. (Since in practice one often deals with finite sequences this is not an essential restriction.) Then the sequence of scaling filter coefficients $h := (h_k)_{k\in\mathbb{Z}}$ can be regarded as a convolution filter defining a stable LTI system:

$$T : \ell^\infty(\mathbb{Z}) \to \ell^\infty(\mathbb{Z}), \quad T[x] := h * x.$$

We recall that $h = T[\delta]$ is also called the impulse response of the system T, where δ denotes the unit impulse. The frequency response of the system T is defined as Fourier transform $H = \hat{h}$ of h:

$$H(\omega) := \sum_{k\in\mathbb{Z}} h_k e^{-2\pi i k\omega}.$$

Lemma 8.6. *The Fourier transform translates the scaling equation (42) into the product*

$$\hat{\varphi}(\omega) = \frac{\sqrt{2}}{2} H\left(\frac{\omega}{2}\right) \hat{\varphi}\left(\frac{\omega}{2}\right). \tag{45}$$

Proof: We apply the Fourier transform to both sides of the scaling equation $\varphi(x) = \sqrt{2}\sum_{k\in\mathbb{Z}} h_k \varphi(2x-k)$:

$$\begin{aligned}
\hat{\varphi}(\omega) &= \int_{\mathbb{R}} \sqrt{2} \sum_{k\in\mathbb{Z}} h_k \varphi(2x-k) e^{-2\pi i x\omega} dx \\
&= \sqrt{2} \sum_{k\in\mathbb{Z}} h_k \int_{\mathbb{R}} \frac{1}{2} \varphi(x) e^{-2\pi i (\frac{x+k}{2})\omega} dx \\
&= \frac{\sqrt{2}}{2} \sum_{k\in\mathbb{Z}} h_k e^{-2\pi i k \frac{\omega}{2}} \int_{\mathbb{R}} \varphi(x) e^{-2\pi i x \frac{\omega}{2}} dx \quad = \quad \frac{\sqrt{2}}{2} H(\tfrac{\omega}{2}) \hat{\varphi}(\tfrac{\omega}{2}).
\end{aligned}$$

□

The product (45) is easier to deal with when using analytic methods. The frequency response H exhibits information about the filter properties of the filter $h = (h_k)_{k\in\mathbb{Z}}$. The following theorem gives the mathematical fundament for the interpretation of h.

Theorem 8.7. *Let $h = (h_k)_{k\in\mathbb{Z}}$ be the scaling filter of an MRA. Then the frequency response H satisfies the so-called orthogonality relation*

$$|H(\omega)|^2 + |H\left(\omega + \tfrac{1}{2}\right)|^2 = 2. \tag{46}$$

Furthermore, the frequency response assumes the following values:

$$H(0) = \sqrt{2} \quad \textit{and} \quad H\left(\tfrac{1}{2}\right) = 0, \tag{47}$$

which is equivalent to

$$\sum_{k\in\mathbb{Z}} h_k = H(0) = \sqrt{2} \quad \textit{and} \quad \sum_{k\in\mathbb{Z}} (-1)^k h_k = H\left(\tfrac{1}{2}\right) = 0. \tag{48}$$

Proof: We refer to p. 114 of [Louis/Maaß/Rieder] or to [Burrus/Gopinath/Guo]. □

Note 8.8. Using the language of filter design, Theorem 8.7 can be interpreted as follows:

- In the case that the scaling function φ is a real-valued function, the scaling filter $h = (h_k)_{k\in\mathbb{Z}}$ are real-valued as well.

- Furthermore, assuming $h \in \ell^1(\mathbb{Z})$, H defines a (1-periodic) continuous function. Then, the values of Equation (47) imply that $|H|$ is small in a neighborhood of $\omega = \frac{1}{2}$ and $|H| \approx 1$ in a neighborhood of $\omega = 0$.

- In other words, the scaling filter $h = (h_k)_{k\in\mathbb{Z}}$ can be regarded as a low-pass filter.

To the scaling filter $h = (h_k)_{k\in\mathbb{Z}}$ we associate the filter $g = (g_k)_{k\in\mathbb{Z}}$ defined by

$$g_k := (-1)^k \bar{h}_{1-k}. \tag{49}$$

Then for the frequency response G follows

$$\begin{aligned}
G(\omega) &= \sum_{k\in\mathbb{Z}} g_k e^{-2\pi ik\omega} = \sum_{k\in\mathbb{Z}} (-1)^k \bar{h}_{1-k} e^{-2\pi ik\omega} \\
&= \sum_{k\in\mathbb{Z}} \bar{h}_{1-k} e^{-2\pi ik(\omega+\frac{1}{2})} \\
&= \sum_{k\in\mathbb{Z}} \bar{h}_k e^{-2\pi i(1-k)(\omega+\frac{1}{2})} \\
&= e^{-2\pi i(\omega+\frac{1}{2})} \sum_{k\in\mathbb{Z}} \overline{h_k e^{-2\pi ik(\omega+\frac{1}{2})}} \\
&= -e^{-2\pi i\omega} \overline{H\left(\omega+\tfrac{1}{2}\right)}.
\end{aligned}$$

From this and Theorem 8.7 we obtain the following result.

Corollary 8.9. *The frequency responses G and H bear the following relation:*

$$|H(\omega)|^2 + |G(\omega)|^2 = 2. \tag{50}$$

Furthermore, the frequency response G assumes the values

$$G(0) = 0, \quad \text{and} \quad G\left(\tfrac{1}{2}\right) = \sqrt{2}. \tag{51}$$

In other words, the convolution filter defined by g is a high-pass filter.

8.2.2 Filter Coefficients and Wavelets

The scaling filter coefficients allow us to construct the wavelets as announced in the main result at the beginning of this section.

Theorem 8.10. *Let $\{V_m|m \in \mathbb{Z}\}$ be an MRA with scaling function $\varphi \in V_0$. The function $\psi \in V_{-1}$ defined by*

$$\psi(x) = \sqrt{2}\sum_{k\in\mathbb{Z}} g_k\varphi(2x-k) := \sum_{k\in\mathbb{Z}} g_k\varphi_{-1,k}(x), \tag{52}$$

$$g_k := (-1)^k\bar{h}_{1-k}, \tag{53}$$

where $\{h_k|k \in \mathbb{Z}\}$ are the coefficients of the scaling equation (42) has the following properties:

(i) $\{\psi_{m,k}(\cdot) := 2^{-m/2}\psi(2^{-m}\cdot -k)|k \in \mathbb{Z}\}$ *is an ONB for* W_m.

(ii) $\{\psi_{m,k}|m,k \in \mathbb{Z}\}$ *is an ONB for* $L^2(\mathbb{R})$.

(iii) ψ *is a wavelet with* $c_\psi = \int_{\mathbb{R}} |\omega|^{-1}|\hat{\psi}(\omega)|^2 d\omega = 2\ln 2$.

Proof: In the first step we show $\psi \in W_0 \subset V_{-1}$:

$$\begin{aligned}
\langle \psi(\cdot) | \varphi(\cdot - n) \rangle &\overset{(42)}{=} 2 \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} g_k \bar{h}_l \langle \varphi(2 \cdot -k) | \varphi(2 \cdot -2n - l) \rangle \\
&= \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} g_k \bar{h}_l \delta_{k,2n+l} = \sum_{l \in \mathbb{Z}} g_{2n+l} \bar{h}_l \\
&\overset{(53)}{=} \sum_{l \in \mathbb{Z}} (-1)^{2n+l} \bar{h}_{1-(2n+l)} \bar{h}_l \\
&= \sum_{l \in \mathbb{Z}} \bar{h}_{1-2(n+l)} \bar{h}_{2l} - \sum_{l \in \mathbb{Z}} \bar{h}_{-2(n+l)} \bar{h}_{2l+1} \\
&= \sum_{\lambda \in \mathbb{Z}} \bar{h}_{1+2\lambda} \bar{h}_{-2(\lambda+n)} - \sum_{l \in \mathbb{Z}} \bar{h}_{-2(n+l)} \bar{h}_{2l+1} \\
&= 0.
\end{aligned}$$

Using Equation (44), a similar computation verifies the orthonormality of

$$\{\psi(\cdot - k)|k \in \mathbb{Z}\}.$$

To complete the proof of (i) and (ii) we have to show the completeness of the system

$$\{\psi(\cdot - k)|k \in \mathbb{Z}\}$$

in W_0. To this means it is sufficient to show the completeness of the orthonormal system

$$\{\varphi(\cdot - k), \psi(\cdot - k)|k \in \mathbb{Z}\}$$

in V_{-1} since $V_0 \oplus W_0 = V_{-1}$. For this part, it is sufficient to show that

$$\varphi_{-1,0} = \sqrt{2}\varphi(2\cdot)$$

is representable by

$$\{\varphi(\cdot - k), \psi(\cdot - k)|k \in \mathbb{Z}\}.$$

We use Equation (42) and Equation (52) and show that Parseval's identity holds:

$$
\begin{aligned}
& 2\sum_{k\in\mathbb{Z}}\Big(|\langle\varphi(2\cdot)|\varphi(\cdot-k)\rangle|^2+|\langle\varphi(2\cdot)|\psi(\cdot-k)\rangle|^2\Big) \\
= \quad & 4\sum_{k\in\mathbb{Z}}\Big(\Big|\sum_{l\in\mathbb{Z}}h_l\langle\varphi(2\cdot)|\varphi(2\cdot-2k-l)\rangle\Big|^2+\Big|\sum_{l\in\mathbb{Z}}g_l\langle\varphi(2\cdot)|\psi(2\cdot-2k-l)\rangle\Big|^2\Big) \\
= \quad & \sum_{k\in\mathbb{Z}}\Big(\Big|\sum_{l\in\mathbb{Z}}h_l\delta_{0,2k+l}\Big|^2+\Big|\sum_{l\in\mathbb{Z}}(-1)^l h_{1-l}\delta_{0,2k+l}\Big|^2\Big) \\
= \quad & \sum_{k\in\mathbb{Z}}h_{2k}^2+\sum_{k\in\mathbb{Z}}h_{2k+1}^2 \\
= \quad & \sum_{k\in\mathbb{Z}}h_k^2 \quad \overset{Lemma 8.5}{=} \quad 1 \quad = \quad \|\varphi_{-1,0}\|^2.
\end{aligned}
$$

The proof of (iii) is more extensive and can be found in [Louis/Maaß/Rieder]. □

Lemma 8.11. *The translated and dilated versions $\psi_{m,k}$ of the wavelet ψ satisfy the scaling equation:*

$$\psi_{m,k} = \sum_{j \in \mathbb{Z}} g_j \varphi_{m-1,2k+j}. \tag{54}$$

Proof: This follows from the from the scaling equation (52) in the same way as in Lemma 8.4. □

The coefficients $(h_k)_{k\in\mathbb{Z}}$ are denoted as scaling filter coefficients and the coefficients $(g_k)_{k\in\mathbb{Z}}$ as wavelet filter coefficients associated to the MRA. We also speak from the scaling filter $h = (h_k)_{k\in\mathbb{Z}}$ and wavelet filter $g = (g_k)_{k\in\mathbb{Z}}$. For an example of the theory introduced so far, we refer to Section 8.3 where the Haar-MRA is discussed in detail.

Note 8.12. The wavelet associated to the MRA is not uniquely determined. However, if the wavelet ψ is defined by equations (52) and (53), the wavelet is real valued in case the scaling function φ is real valued. This is an important property in view of an implementation of the wavelet algorithms discussed in the next section. For further details we refer to p. 117 of [Louis/Maaß/Rieder].

8.2.3 Fast Wavelet Algorithms

From the interplay of MRA and wavelet theory one can derive the fast discret wavelet transform, which is the content of this section. It seems to be an odd consequence that a wavelet transform based on a wavelet associated to some MRA can be performed without knowing the wavelet explicitly. All we need to know are the scaling filter coefficients $(h_k)_{k\in\mathbb{Z}}$ defined by the scaling equation (42).

Instead of computing wavelet coefficients $\langle f|\psi^{s,t}\rangle$ by computing the integrals of the inner products, one does smoothing much easier: the signal (or, to be more precise, a discretized form of the signal) is convolved with the low-pass filter $h = (h_k)_{k\in\mathbb{Z}}$ (scaling coefficients) and the associated high-pass filter $g = (g_k)_{k\in\mathbb{Z}}$ (wavelet filter coefficients). The details of this procedure will be explained in the next subsections.

8.2.4 Fast Discrete Wavelet Transform (FDWT)

In this subsection we introduce the algorithm for the fast computation of the discrete wavelet transform. Based on the underlying MRA the algorithm can be derived from the scaling equation (42) in an elegant fashion.

We fix an MRA with scaling function φ and start with a signal f which lies in the base space V_0 of the MRA. As it turns out, this assumption on f will be the actual discretization step (see subsection 8.2.7). By Definition 8.1, the signal f can be written as

$$f(x) = \sum_{k \in \mathbb{Z}} v_k^0 \varphi(x - k) \tag{55}$$

with a suitable family of coefficients

$$v^0 = (v_k^0)_{k \in \mathbb{Z}}.$$

As before, let ψ denote the orthonormal wavelet associated to φ (see (52)). Then

$$\{\psi_{m,k} = 2^{-m/2}\psi(2^{-m} \cdot -k) | m, k \in \mathbb{Z}\}$$

defines an ON basis of $L^2(\mathbb{R})$. The goal is to compute the wavelet coefficients

$$\tilde{f}(2^m, 2^m k) = \langle f | \psi_{m,k} \rangle, m \in \mathbb{Z}, k \in \mathbb{Z}. \tag{56}$$

Note that this amounts to evaluating the CWT $\tilde{f}$ on a CWT-adapted discrete-time grid with $a = 2$ and $b = 1$ (see Subsection 7.5.1). The corresponding wavelet frame is an ON-basis (see Subsection 7.5.2).

Since $f \in V_0$ and $L^2(\mathbb{R}) = \overline{V_0 \oplus \bigoplus_{j \leq 0} W_j}$ by equation (39), it follows immediately that

$$\langle f | \psi_{m,k} \rangle = 0 \quad \text{for} \quad m \leq 0, k \in \mathbb{Z}$$

Therefore, it suffices to compute $\tilde{f}(2^m, 2^m k)$ for $m \in \mathbb{N}$, $k \in \mathbb{Z}$. We introduce the following notation:

$$\begin{aligned} w_k^m &:= \langle f | \psi_{m,k} \rangle, \quad w^m = (w_k^m | k \in \mathbb{Z}) \in \ell^2(\mathbb{Z}), \\ v_k^m &:= \langle f | \varphi_{m,k} \rangle, \quad v^m = (v_k^m | k \in \mathbb{Z}) \in \ell^2(\mathbb{Z}). \end{aligned}$$

Using the wavelet equation (54) and scaling equation (43) we obtain

$$w_k^m = \langle f|\psi_{m,k}\rangle = \sum_{l\in\mathbb{Z}} \bar{g}_l \langle f|\varphi_{m-1,2k+l}\rangle = \sum_{l\in\mathbb{Z}} \bar{g}_{l-2k} v_l^{m-1}, \tag{57}$$

$$v_k^m = \langle f|\varphi_{m,k}\rangle = \sum_{l\in\mathbb{Z}} \bar{h}_l \langle f|\varphi_{m-1,2k+l}\rangle = \sum_{l\in\mathbb{Z}} \bar{h}_{l-2k} v_l^{m-1}. \tag{58}$$

These recursive formulas constitute already the wavelet decomposition algorithm: starting with the sequence v^0 we can recursively compute the wavelet coefficients w^m, $m \in \mathbb{N}$, from v^{m-1} which in turn can be recursively computed. Once v^0 is known, all operations are performed on the v^m and w^m and the signal f is not used any longer. As mentioned before, the assumption $f \in V_0$ is the actual discretization step of the time-continuous signal f. In Subsection 8.2.7 we will discuss the problems which arise from this assumption in practice.

We further investigate the computations in (57) and (58) which will be expressed by the decomposition operators $\mathcal{H}$ and $\mathcal{G}$, respectively. We define

$$\begin{aligned} \mathcal{H} : \ell^2(\mathbb{Z}) &\rightarrow \ell^2(\mathbb{Z}) \\ v &\mapsto \left((\mathcal{H}v)_k = \sum_{l\in\mathbb{Z}} \bar{h}_{l-2k} v_l \right)_{k\in\mathbb{Z}} \end{aligned} \tag{59}$$

and

$$\begin{aligned} \mathcal{G} : \ell^2(\mathbb{Z}) &\rightarrow \ell^2(\mathbb{Z}) \\ v &\mapsto \left((\mathcal{G}v)_k = \sum_{l\in\mathbb{Z}} \bar{g}_{l-2k} v_l \right)_{k\in\mathbb{Z}}. \end{aligned} \tag{60}$$

The operators $\mathcal{H}$ and $\mathcal{G}$ have a nice interpretation in terms of filter theory: define the filters $\tilde{h} = (\tilde{h}(k))_{k\in\mathbb{Z}}$ and $\tilde{g} = (\tilde{g}(k))_{k\in\mathbb{Z}}$ by

$$\tilde{h}(k) = \bar{h}_{-k} \quad \text{and} \quad \tilde{g}(k) = \bar{g}_{-k}, \quad k \in \mathbb{Z},$$

then $\mathcal{H}$ and $\mathcal{G}$ are nothing else than convolution with $\tilde{h}$ and $\tilde{g}$, respectively, and downsampling by a factor of 2. In other words,

$$\begin{aligned} w^m &= (\downarrow 2)[\tilde{g} * v^{m-1}], \\ v^m &= (\downarrow 2)[\tilde{h} * v^{m-1}]. \end{aligned}$$

The algorithm for the discrete wavelet transform performed on the first M scales can be summarized as follows.

Fast Discrete Wavelet Transform (FDWT)

Input: $v^0 = (v_k^0)_{k\in\mathbb{Z}}$
$M =$ number of scales

Compute: $w^m = \mathcal{G}v^{m-1}$ for $m = 1, \ldots, M$
$v^m = \mathcal{H}v^{m-1}$

Output: v^M
$w^m, m = 1, \ldots, M$

$$\begin{array}{ccccccccc} v^0 & \xrightarrow[\searrow]{\mathcal{H}\ \mathcal{G}} & v^1 & \xrightarrow[\searrow]{\mathcal{H}\ \mathcal{G}} & v^2 & \ldots & v^{M-1} & \xrightarrow[\searrow]{\mathcal{H}\ \mathcal{G}} & v^M \\ & & w^1 & & w^2 & & \ldots & & w^M \end{array}$$

Note 8.13. The algorithm starts with a signal f corresponding to some fine scale (here $f \in V_0$ corresponding to scale 0) and then proceeds in direction of coarser scales. The choice of V_0 as start of the algorithm is arbitrary. One could, as well, start with the space V_{-1254} and proceed in the same way.

8.2.5 Fast Discrete Wavelet Reconstruction (FIDWT)

We now investigate how to reconstruct the sequence v^0 and hence the original signal $f \in V_0$,

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

from the sequences of coefficients

$$\{v^M, w^m | m = 1, \ldots, M\}.$$

This amounts to the inverse of the discrete wavelet transfrom discussed in the previous subsection. The corresponding fast algorithm is called fast discrete wavelet reconstruction or fast inverse discrete wavelet transform (FIDWT).

We first investigate how to reconstruct the sequence v^0 from v^1 and w^1. Recall that the space V_0 is the orthogonal sum of the subspaces V_1 and W_1. Therefore

$$\begin{aligned}\sum_{k\in\mathbb{Z}} v_k^0 \varphi_{0,k} &= \sum_{j\in\mathbb{Z}} v_j^1 \varphi_{1,j} + \sum_{j\in\mathbb{Z}} w_j^1 \psi_{1,j} \\ &= \sum_{j\in\mathbb{Z}} v_j^1 \sum_{l\in\mathbb{Z}} h_l \varphi_{0,2j+l} + \sum_{j\in\mathbb{Z}} w_j^1 \sum_{l\in\mathbb{Z}} g_l \varphi_{0,2j+l},\end{aligned}$$

where we have used the scaling equation (43) and wavelet equation (54). Comparing coefficients results in

$$v_k^0 = \sum_{j\in\mathbb{Z}} v_j^1 h_{k-2j} + \sum_{j\in\mathbb{Z}} w_j^1 g_{k-2j}. \tag{61}$$

In the same way, one can reconstruct the sequence v^{M-1} from w^M and v^M, $M \in \mathbb{N}$. Proceeding recursively, one can reconstruct v^0 from v^M and the sequences of wavelet coefficients w^m of the scales $m = M, M-1, \ldots, 1$.

This reconstruction algorithm can be expressed by the operators $\mathcal{H}^*$ and $\mathcal{G}^*$ defined by

$$\begin{aligned} \mathcal{H}^* : \ell^2(\mathbb{Z}) &\rightarrow \ell^2(\mathbb{Z}) \\ v &\mapsto ((H^*v)_k = \sum_{j\in\mathbb{Z}} h_{k-2j}v_j | k \in \mathbb{Z}), \qquad (62) \\ \mathcal{G}^* : \ell^2(\mathbb{Z}) &\rightarrow \ell^2(\mathbb{Z}) \\ w &\mapsto ((G^*w)_k = \sum_{j\in\mathbb{Z}} g_{k-2j}w_j | k \in \mathbb{Z}). \qquad (63) \end{aligned}$$

It is not difficult to show that $\mathcal{H}^*$ and $\mathcal{G}^*$ are the adjoint operators to the decomposition operators $\mathcal{H}$ and $\mathcal{G}$ of the last subsection, i.e.,

$$\langle \mathcal{H}v|w\rangle = \langle v|\mathcal{H}^*w\rangle \quad \text{and} \quad \langle \mathcal{G}v|w\rangle = \langle v|\mathcal{G}^*w\rangle, \quad v, w \in \ell^2(\mathbb{Z}).$$

A single reconstruction step from scale m to $m-1$ is described by

$$v^{m-1} = \mathcal{H}^* v^m + \mathcal{G}^* w^m. \tag{64}$$

As for $\mathcal{H}$ and $\mathcal{G}$, the operators $\mathcal{H}^*$ and $\mathcal{G}^*$ also have a nice interpretation in terms of filter theory: $\mathcal{H}^*$ and $\mathcal{G}^*$ are nothing else then upsampling by a factor of 2 and subsequent convolution with h and g, respectively. In other words,

$$v^{m-1} = h * ((\uparrow 2)[v^m]) + g * ((\uparrow 2)[w^m]). \tag{65}$$

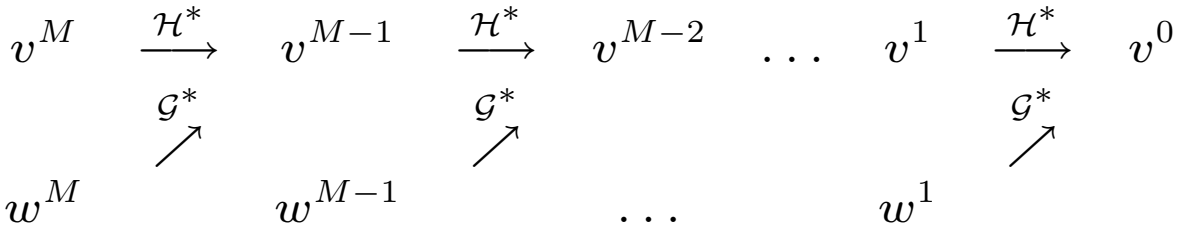
The algorithm for the discrete inverse wavelet transform (FIDWT) starting at scale M and proceeding up to scale 0 can be summarized as follows.

Fast Discrete Wavelet Reconstruction (FIDWT)

Input: M = number of scales
v^M
w^m, $m = 1, \ldots, M$

Compute: $v^{m-1} = \mathcal{H}^* v^m + \mathcal{G}^* w^m$ for $m = M, \ldots, 1$

Output: v^0



Note 8.14. The reconstruction algorithm starts with some sequence v^M which corresponds to some coarse representation of the original signal f. Using successively the wavelet coefficients w^m of the scales $m = M, M-1, \ldots 1$, the sequences v^m are reconstructed which correspond to finer versions of f for decreasing m. Reaching the scale $m = 0$ one has finally reconstructed the original signal $f \in V_0$.

8.2.6 Complexity of the FDWT and FIDWT

We now want to determine the complexity of the algorithms for discrete wavelet transform and its inverse.

Recall that $\ell(x)$ denotes the length of a finite DT-signal $x \in \mathbb{C}^{\mathbb{Z}}$ (see Definition 3.23). We suppose that the length of the input sequence $v^0 = (v_k^0)_{k\in\mathbb{Z}}$ has finite length

$$N := \ell(v^0).$$

Furthermore, we only consider the case of finite filters $h = (h_k)_{k\in\mathbb{Z}}$ and $g = (g_k)_{k\in\mathbb{Z}}$. One can show that this holds for any wavelets with compact support such as for the Haar wavelet or the Daubechies wavelets.

By definition of g (see (53) of Theorem 8.10) we have $\ell(g) = \ell(h)$. In the discrete wavelet decomposition the coefficient sequences

$$w^m = \mathcal{G}v^{m-1} \quad \text{und} \quad v^m = \mathcal{H}v^{m-1}$$

are successively computed for $m = 1, \ldots, M$. The operators H and G can be expressed as convolution with suitable filters of length $\ell(h)$ and subsequent 2-downsampling.

By Lemma 3.24 the length of the convolution of two finite sequences x and y is given by the formula

$$\ell(x * y) = \ell(x) + \ell(y) - 1.$$

By 2-downsampling, the length of the resulting sequence is about half the length.

Therefore, we get for $m = 1, \ldots, M$

$$\begin{aligned} \ell(v^m) &\leq \lceil \tfrac{1}{2}(\ell(v^{m-1}) + \ell(h) - 1) \rceil \leq \tfrac{1}{2}(\ell(v^{m-1}) + \ell(h)) \\ \ell(w^m) &\leq \lceil \tfrac{1}{2}(\ell(v^{m-1}) + \ell(g) - 1) \rceil \leq \tfrac{1}{2}(\ell(v^{m-1}) + \ell(h)) \end{aligned}$$

From this, we inductively get the estimates

$$\ell(v^m) \leq 2^{-m}\ell(v^0) + \ell(h), \tag{66}$$

$$\ell(w^m) \leq 2^{-m}\ell(v^0) + \ell(h). \tag{67}$$

Furthermore, note that the algorithm terminates if $M \geq \log_2 N$. In other words, the number M of scales, which have to be computed, is bounded by $\log_2 N$.

Using a straightforward algorithm to compute the convolution $x * y$ of two finite signals x and y (note that there are in general more efficient algorithms) takes at most $2\ell(x)\ell(y)$ operations (additions and multiplications). Therefore, the overall cost OC for computing the sequences v^m and w^m for the scales $m = 1, \ldots, M$ is bounded by

$$\begin{aligned} \mathrm{OC} &\leq \sum_{m=1}^{M} (2\ell(v^m)\ell(h) + 2\ell(w^m)\ell(g)) \\ &\leq \sum_{m=1}^{M} 4(2^{-m}\ell(v^0)\ell(h) + \ell(h)^2) \\ &\leq 4\ell(v^0)\ell(h) + 4M\ell(h)^2. \end{aligned}$$

Since M is bounded by $\log_2 N$ we get the result as summarized in the next theorem.

Theorem 8.15. *Let V_0 be the base space of an MRA with scaling function φ, associated wavelet ψ and filter coefficients $h = (h_k)_{k\in\mathbb{Z}}$ of finite length $\ell(h)$. Furthermore, let $f \in V_0$ be a signal with $f(x) = \sum_{k\in\mathbb{Z}} v_k^0 \varphi(x-k)$ such that the coefficient sequence $v^0 = (v_k^0)_{k\in\mathbb{Z}}$ is of finite length $N = \ell(v^0)$. Then the overall complexity OC to compute the discrete wavelet transform f with respect to the mother wavelet ψ is*

$$\mathsf{OC} = O(N).$$

To be more explicit, one has the bound $\mathsf{OC} \le c_{\ell(h)} N$, where $c_{\ell(h)}$ is a constant which depends linearly on the filter length $\ell(h)$.

The overall complexity of the inverse transform (FIDWT) can be determined in a similar fashion and results in a complexity of the same order as in Theorem 8.15.

8.2.7 Discretization step in the DWT

Again, we want to emphasize that in the fast discrete wavelet transform neither the analyzing wavelets $\psi^{s,t}$ nor the analyzed signal f are discrete. Discrete is only the grid on which the wavelet coefficients $w_k^m = \langle f|\psi_{m,k}\rangle$ are computed given by the integrals

$$\langle f|\psi_{m,k}\rangle = \int_{\mathbb{R}} f(u)\bar{\psi}_{m,k}(u)du.$$

Also these integrals are computed precisely by the FDWT and not approximately, as one might guess. However, the integrals are not computed directly but recursively using the values of integrals corresponding to lower scales.

These facts are surprising at first sight and the following kind of questions come into mind:

- How is it possible to compute integrals of time-continuous functions by some discrete algorithm?
- Where lies the actual discretization step which connects the CT-world (the signal f and the wavelets $\psi_{m,k}$) with the DT-world?

The answer to these questions lies in the assumption (55) that the signal f lies in the base space V_0 of the MRA. In other words, in order for the FDWT algorithm to get started ones needs the assumption that f can be developed as

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

with a family of coefficients

$$v^0 = (v_k^0 = \langle f | \varphi(\cdot - k) \rangle)_{k \in \mathbb{Z}}.$$

The coefficient sequence v^0 is the starting point for the recursive FDWT. In theory, this algorithm is very elegant when having the "right" assumptions. In practice, however, there are quite some open problems:

- How does one obtain from the original CT-signal $f \in L^2(\mathbb{R})$ a signal in the base space V_0, which is needed as starting point for the recursive FDWT?
- The assumption $f \in V_0$ guarantees the existence of a coefficient sequence $v^0 = (v_k^0)_{k\in\mathbb{Z}}$. However, how can the coefficients $v_k^0 = \langle f|\varphi(\cdot - k)\rangle$, which are given by integrals, be computed in practice? Even computing the integrals numerically (for example, approximation by Riemann sums) is in general very expensive and inefficient.
- In digital signal processing the signals are in general not given as CT-signals $f \in L^2(\mathbb{R})$ but as sampled DT-signals $(f(k))_{k\in\mathbb{Z}}$. How can the FDWT used for such DT-signals?

In practice there is an easy "solution" to all of these problems: just use the samples $(f(k))_{k\in\mathbb{Z}}$ of the CT-signal f as coefficient sequence $(v_k^0)_{k\in\mathbb{Z}}$. Strictly speaking, this identification does in general not make sense. In the following we cite Strang und Nguyen who write in their book [Strang/Nguyen]:

> ... Is this legal [to use the samples $f(k)$ instead of the coefficients (v_k^0)]? **No. It is a wavelet crime.** Some can't imagine doing it, others can't imagine not doing it. Is this crime convenient? **Yes.** We may not know the whole function $t \mapsto f(t)$, it may not be a combination of $(\varphi(\cdot - k)_{k\in\mathbb{Z}}$, and computing the true coefficients $(v_k^0)_{k\in\mathbb{Z}}$ may take too long. But the crime cannot go unnoticed - we have to discuss it.

For a detailed discussion we refer to p. 232 of [Strang/Nguyen]. We restrict ourselves to some remarks. The basic assumption when identifying samples with coefficients is

$$v_k^0 \approx f(k), \quad \text{for} \quad k \in \mathbb{Z}.$$

Actually, this is true for many scaling functions φ which behave like the Dirac function δ. In other words, if φ is concentrated around the point of time $t = 0$ like a peak and its integral is 1 then $\langle f|\varphi\rangle \approx \langle f|\delta\rangle$. Then, the same holds for the corresponding translates of φ and δ and we get

$$v_k^0 = \langle f|\varphi(\cdot - k)\rangle \approx \langle f|\delta(\cdot - k)\rangle = f(k).$$

Therefore, in practice it is important to check if this approximation property is actually satisfied — depending on the underlying MRA and its associated wavelet and the signal f.

If the approximation property does not hold, the samples values have to be suitably preprocessed using so-called prefilters. These modified samples can then be used as starting coefficients in the FDWT instead of the $(v_k^0)_{k\in\mathbb{Z}}$.

Conclusion: If for a given MRA and a signal f one has $v_k^0 \approx f(k)$, the samples $(f_k)_{k\in\mathbb{Z}}$ (possibly preprocessed) can be used instead of the coefficients $(v_k^0)_{k\in\mathbb{Z}}$ as starting values for the FDWT. In this case, however, the wavelet coefficients w_k^m are faulty, even though the FDWT computes exactly. The quality of the w_k^m depends solely on the approximation quality of the input data.

8.2.8 P_m, Q_m and Wavelet Coefficients

The FDWT computes from the input sequence of scaling coefficients $v^0 = (v_k^0)_{k\in\mathbb{Z}}$ of some signal $f \in V_0$ the sequence of wavelet coefficients $w^m = (w_k^m)_{k\in\mathbb{Z}}$ for the scales $m = 1, \ldots, M$ and the sequence of scaling coefficients $v^M = (v_k^M)_{k\in\mathbb{Z}}$ on scale M. With these data the projections Q_m of f on W_m are given by

$$Q_m(f) = \sum_{k\in\mathbb{Z}} \langle f|\psi_{m,k}\rangle \psi_{m,k} = \sum_{k\in\mathbb{Z}} w_k^m \psi_{m,k},$$

$m = 1, \ldots, M$, and the projection P_M of f on V_M is given by

$$P_M(f) = \sum_{k\in\mathbb{Z}} \langle f|\varphi_{M,k}\rangle \varphi_{M,k} = \sum_{k\in\mathbb{Z}} v_k^M \varphi_{M,k}.$$

Similarly, the projections $P_m(f)$ can be computed by using the scaling coefficients v^m for the scales $m = M, \ldots, 1$, wich in turn can be computed by the FIDWT.

Since in practice neither the scaling function φ nor the mother wavelet ψ are known explicitly, the projections can not very well be computed even though the wavelet and scaling coefficients are known. Many "wavelet specialists" read the properties of some signal f directly from the wavelet coefficients. For a graphical representation of the wavelet coefficients (if there are only a finite number, i.e., if the sequence v^0 has finite length N) one constructs one sequence

$$w(f) := (v^M, w^M, w^{M-1}, \ldots, w^1) \tag{68}$$

from the $M+1$ sequences $w^1, \ldots, w^M$ and v^M. This representation is, for example, used in Figure 40 of Chapter 9.

In concrete applications the length $N = \ell(v^0)$ of the input sequence, which also determines the complexity of the FDWT and FIDWT, is much larger then the filter lengths $\ell(h) = \ell(g)$. For example, one has often $N > 10^4$, whereas $\ell(h) \leq 8$. Therefore, from the estimations (66) and (67) follows

$$\ell(w(f)) \approx N.$$

If one considers signals $f \in V_0$ with $\ell(v^0) = N \in \mathbb{N}$, where N is a power of two, one can achieve by a small modification of the FDWT that the output sequence $w(f)$ as given in Equation (68) has also length N. The so modified FDWT mapping the N coefficients of v^0 to the N coefficients of $w(f)$ can be realized by an orthonormal $N \times N$-matrix. For details we refer to [Strang/Nguyen].

Warning: Finally, we repeat one of our warnings concerning the L^2 norm. We have seen that any signal $f \in L^2(\mathbb{R})$ can be represented as

$$f = \sum_{m\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \langle f|\psi_{m,k}\rangle \psi_{m,k}.$$

This equation holds again only in the L^2-sense! It is important to be conscious of this fact since otherwise paradoxical implications arise. In the following we give an example. For each $n \in \mathbb{N}$ we define

$$f_n := \sum_{-n\le m\le n} \sum_{-n\le k\le n} \langle f|\psi_{m,k}\rangle \psi_{m,k}.$$

Since for all wavelets $\psi_{m,k}$, $m, k \in \mathbb{Z}$, the mean value

$$\int_{\mathbb{R}} \psi_{m,k}(t)dt$$

is zero, the same holds for each function f_n. Since f is the limit function of f_n for $n \longrightarrow \infty$, one could think that the mean value of f also has to be zero even though this does not hold for a general function $f \in L^2(\mathbb{R})$.

Here the fact is decisive that the sequence $(f_n)_{n\in\mathbb{N}}$ does not converge in the L^1-norm (in this case the mean value of f would have to vanish) but converges in the L^2-norm. In other words, it is possible to represent an arbitrary function in $f \in L^2(\mathbb{R})$ as an L^2-limit of functions whose mean value is zero.

8.3 Example: Haar-MRA

In this section, we illustrate the theory introduced before by means of the Haar-MRA as an example. Already in 1910, Haar showed that certain translated and scaled version of the box function form a basis of $L^2(\mathbb{R})$. In much later years, it was discovered that this system — also referred to as Haar system — constitutes the easiest example for an MRA.

Let V_0 be the space of functions in $L^2(\mathbb{R})$, which are constant on all intervals

$$[k, k+1[, \quad k \in \mathbb{N}.$$

Any function $f \in V_0$ is determined by its values $f(k)$ at the points of time $t = k$. $k \in \mathbb{N}$. Note that the function $t \mapsto f(2t)$ ist constant at all intervals of the form

$$[k/2, (k+1)/2[\quad \text{for} \quad k \in \mathbb{N}.$$

The space of all such functions is denoted by V_{-1}. In a similar fashion, we define the space $V_m, m \in \mathbb{Z}$, as space of functions being constant on the so-called dyadic intervals of length 2^m. Here, a dyadic interval of length 2^m is defined by

$$I_{m,k} := [k \cdot 2^m, (k+1) \cdot 2^m[\quad \text{for} \quad k \in \mathbb{N}.$$

Obviously, one has the inclusions

$$\{0\} \subset \ldots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \ldots \subset L^2(\mathbb{R}),$$

since each function being constant at the intervals of length 2^{-m} are, of course, also constant on the intervals of half the length. The properties

$$\bigcap_{m\in\mathbb{Z}} V_m = \{0\} \quad \text{and} \quad f(\cdot) \in V_m \Longleftrightarrow f(2^m\cdot) \in V_0$$

are also easy to see. The actual achievement of Haar was to show that

$$\overline{\bigcup_{m\in\mathbb{Z}} V_m} = L^2(\mathbb{R}).$$

In other words, the L^2-step functions on the dyadic intervals are dense in $L^2(\mathbb{R})$, i.e., each function in $L^2(\mathbb{R})$ can be approximated by such step functions (in the L^2 sense). For a proof of this fact, we have to refer to the literature.

Up to the existence of a scaling function, we have seen that the spaces V_m, $m \in \mathbb{Z}$, satisfy the axioms of an MRA (see Definition 8.1). Define

$$\varphi(t) := \begin{cases} 1 & \text{if} \quad 0 \leq t < 1 \\ 0 & \text{otherwise,} \end{cases}$$

which is also illustrated in Figure 32. The box function φ is obviously orthogonal to its integral translates $t \mapsto \varphi(t-k), k \in \mathbb{Z}$, and these translates form an ONB of V_0. Therefore, φ defines a scaling function as postulated and makes V_m, $m \in \mathbb{Z}$, to an MRA which is also denoted as Haar-MRA.

For the Haar-MRA we summarize the results which were derived for general MRAs in the last section. The existence of a scaling equation ist guaranteed by Lemma 8.2. In our example, one can easily guess the scaling coefficients. One has

$$h_0 = \frac{1}{\sqrt{2}}, h_1 = \frac{1}{\sqrt{2}}, \tag{69}$$

and all other coefficients are zero. This filter was, up to a factor, already investigated in Subsection 3.6.1. Recall that the frequency response H of $h := (h_k)_{k\in\mathbb{Z}}$ is given by

$$H(\omega) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} e^{-2\pi i\omega} = e^{-\pi i\omega}\sqrt{2}\cos(\pi\omega).$$

As was shown in Theorem 8.7, one indeed has $H(0) = \sqrt{2}$ and $H(\frac{1}{2}) = 0$.

The scaling filter coefficients $h = (h_k)_{k\in\mathbb{Z}}$ define a lowpass filter. The associated wavelet filter coefficients $g = (g_k)_{k\in\mathbb{Z}}$ are given by

$$g_0 = \bar{h}_1 = \frac{1}{\sqrt{2}}, g_1 = (-1)\bar{h}_0 = -\frac{1}{\sqrt{2}}, \tag{70}$$

and all other coefficients are zero. The mother wavelet of the MRA is by Theorem 8.10 defined by

$$\begin{aligned}
\psi(x) &= \sqrt{2}\sum_{k\in\mathbb{Z}} g_k \varphi(2x-k) \\
&= \begin{cases} 1 & \text{if} \quad 0 \le t < \frac{1}{2} \\ -1 & \text{if} \quad \frac{1}{2} \le t < 1 \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

This is the so-called Haar wavelet, which we already encountered in Figure 19. We already know from Theorem 7.4 that the Haar wavelet is indeed a wavelet in the sense of Definition 7.1. This follows now again from Theorem 8.10. Furthermore, the

$$\{\psi_{m,k} | k \in \mathbb{Z}\}, \quad k \in \mathbb{Z},$$

form an ONB of W_m and

$$\{\psi_{m,k} | m, k \in \mathbb{Z}\},$$

from an ONB of $L^2(\mathbb{R})$. Among others, it follows that the mother wavelet $\psi \in W_0$ is orthogonal to the scaling function $\varphi \in V_0$ and its translates.

Figure 32 shows the Haar scaling function φ, the corresponding Haar wavelet ψ, and the frequency responses of the scaling and wavelet filter coefficients.

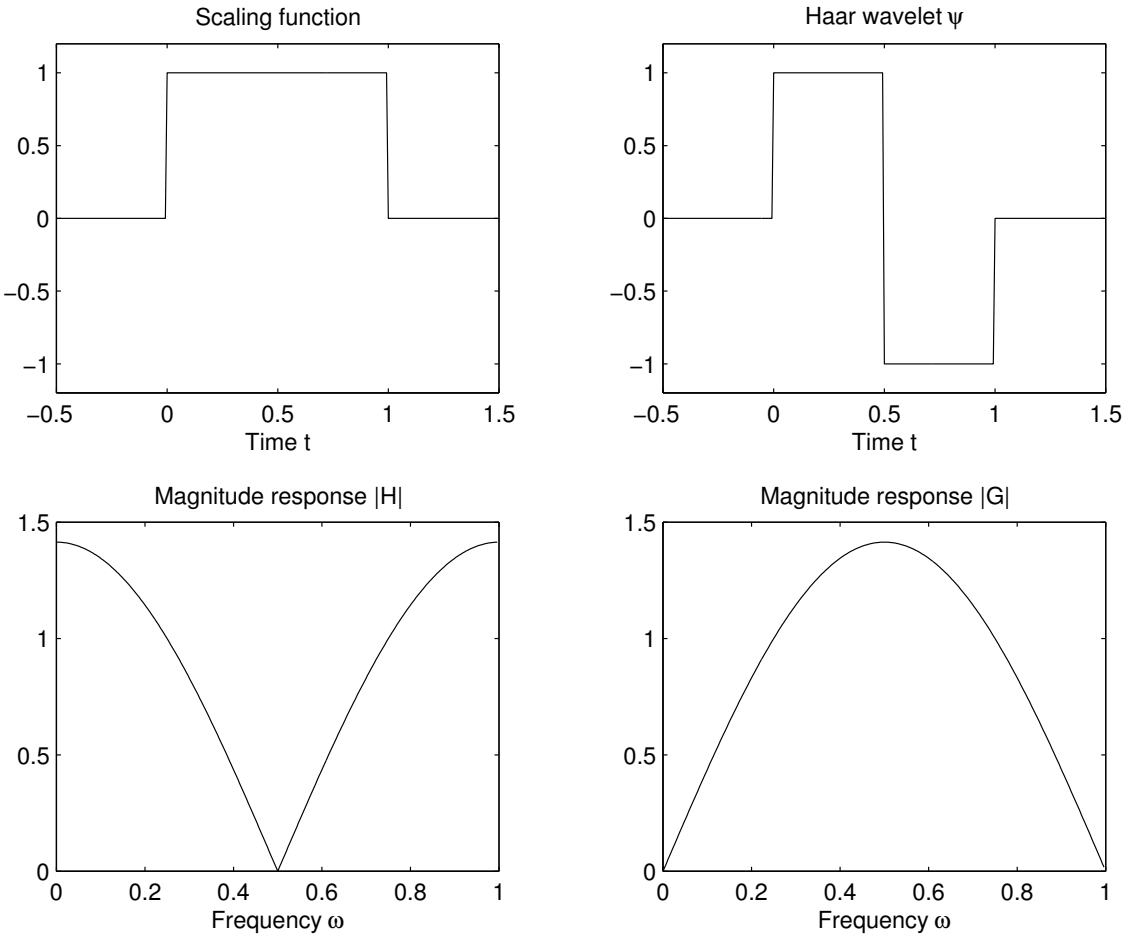


Figure 32: Haar-MRA.

We now want to perform the FDWT from Subsection 8.2.4 w.r.t. the Haar wavelet ψ on some concrete input signal. The signal $f \in V_0$ shown in Figure 33 is given by

$$\begin{aligned} f(t) = \quad & 1 \cdot \varphi(t) + 1 \cdot \varphi(t-1) + (-1) \cdot \varphi(t-2) + (-1) \cdot \varphi(t-3) \\ & +(-1) \cdot \varphi(t-4) + (-1) \cdot \varphi(t-5) + (-2) \cdot \varphi(t-6) + 2 \cdot \varphi(t-7). \end{aligned}$$

Using the notation of Subsection 8.2.4, the input coefficient sequence $v^0 = (v_k^0)_{k \in \mathbb{Z}}$ for the FDWT is given by is given by

$$v^0 = (1, 1, -1, -1, -1, -1, -2, 2).$$

(All other coefficients are zero and we only write in the following the above coefficients.)

Using the filters h and g (see (69) and (70)) we recursively compute the sequences v^m of scaling coefficients for $m = 1, \ldots, M$ by the formula

$$v_k^m = \sum_{l \in \mathbb{Z}} \bar{h}_{l-2k} v_l^{m-1} = \frac{1}{\sqrt{2}} v_{2k}^{m-1} + \frac{1}{\sqrt{2}} v_{2k+1}^{m-1}$$

and the sequence w^m of wavelet coefficients by

$$w_k^m = \sum_{l \in \mathbb{Z}} \bar{g}_{l-2k} v_l^{m-1} = \frac{1}{\sqrt{2}} v_{2k}^{m-1} - \frac{1}{\sqrt{2}} v_{2k+1}^{m-1}.$$

Using these formulas, the FDWT proceeds as follows:

v^0	1	1	-1	-1	-1	-1	-2	2
v^1	$\sqrt{2}$		$-\sqrt{2}$		$-\sqrt{2}$		0	
w^1	0		0		0		$-2\sqrt{2}$	
v^2		0				-1		
w^2		2				-1		
v^3				$\frac{-1}{\sqrt{2}}$				
w^3				$\frac{1}{\sqrt{2}}$				

Figure 33 shows the signal f and its projections $P_i(f)$ and $Q_i(f)$ onto the subspaces V_i and W_i, respectively, for $i = 1, 2, 3$. The projection $Q_2(f)$, for example, is computed from w^2 by

$$Q_2(f) = \sum_{k\in\mathbb{Z}} w_k^2 \psi_{2,k} = 2\psi_{2,0} - 1\psi_{2,1} = \psi\left(\frac{\cdot}{4}\right) - \frac{1}{2}\psi\left(\frac{\cdot - 4}{4}\right).$$

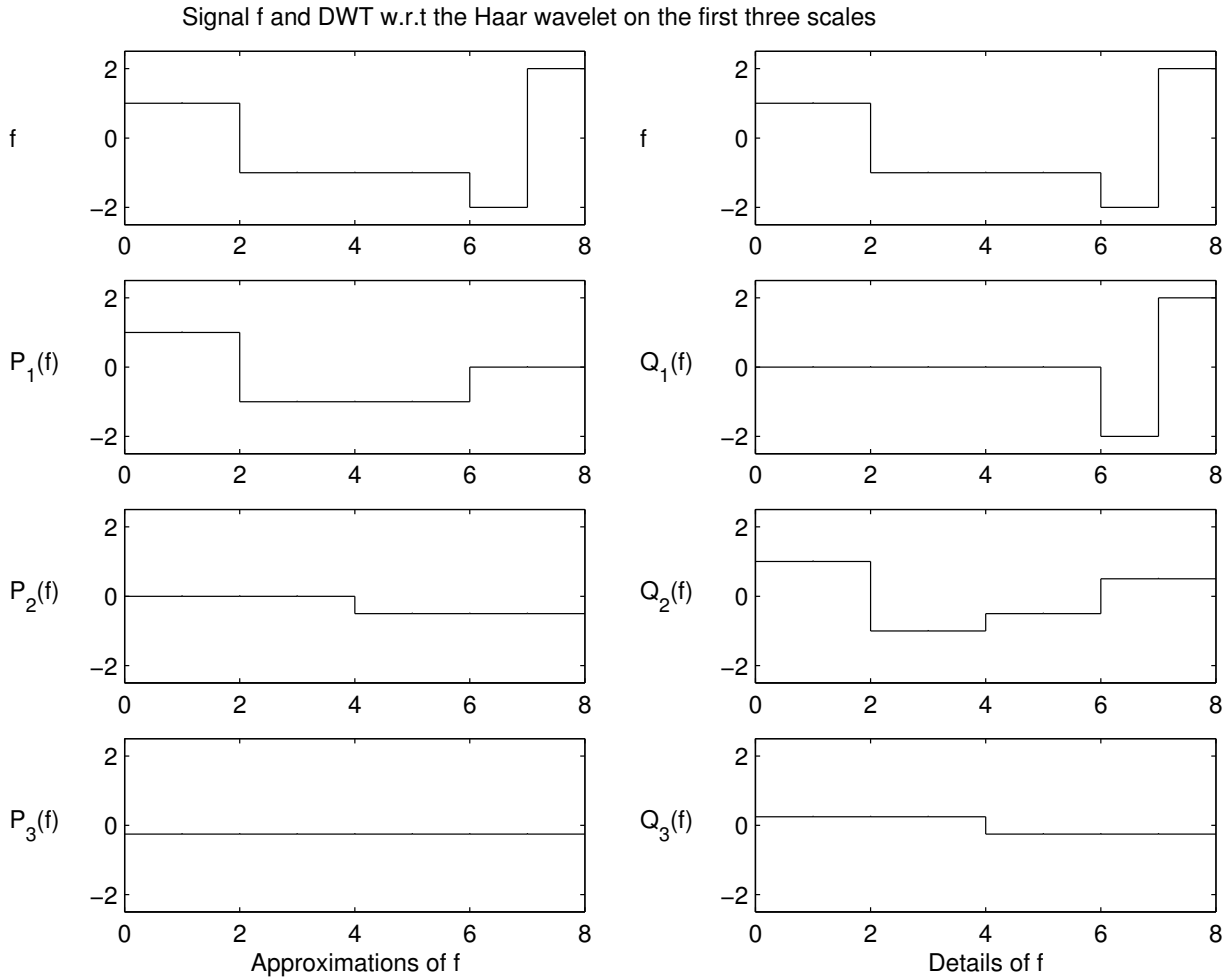


Figure 33: Example of a DWT w.r.t. the Haar wavelet of the signal f

Finally, we want to look at the DWT w.r.t. the Haar wavelet of two signals we have already encountered in several examples.

Figure 34 shows the projections of the chirp signal

$$f(t) = \sin(20\pi(t/N)^2)$$

onto the spaces V_i and W_i for the scales $i = 1, 2 \ldots, M$, $M = 7$. To this means we have sampled f at the points $t = 1, 2, \ldots, N$, $N = 128$, and the sequence of samples was taken as sequence v^0 of scaling coefficients based on the assumption

$$v^0_{k-1} \approx f(k) \quad \text{für} \quad k = 1, 2, \ldots, N.$$

We recall that this identification can lead to faulty and meaningless wavelet coefficients in case the above approximation assumption does not hold. For a discussion we refer to Section 8.2.7.

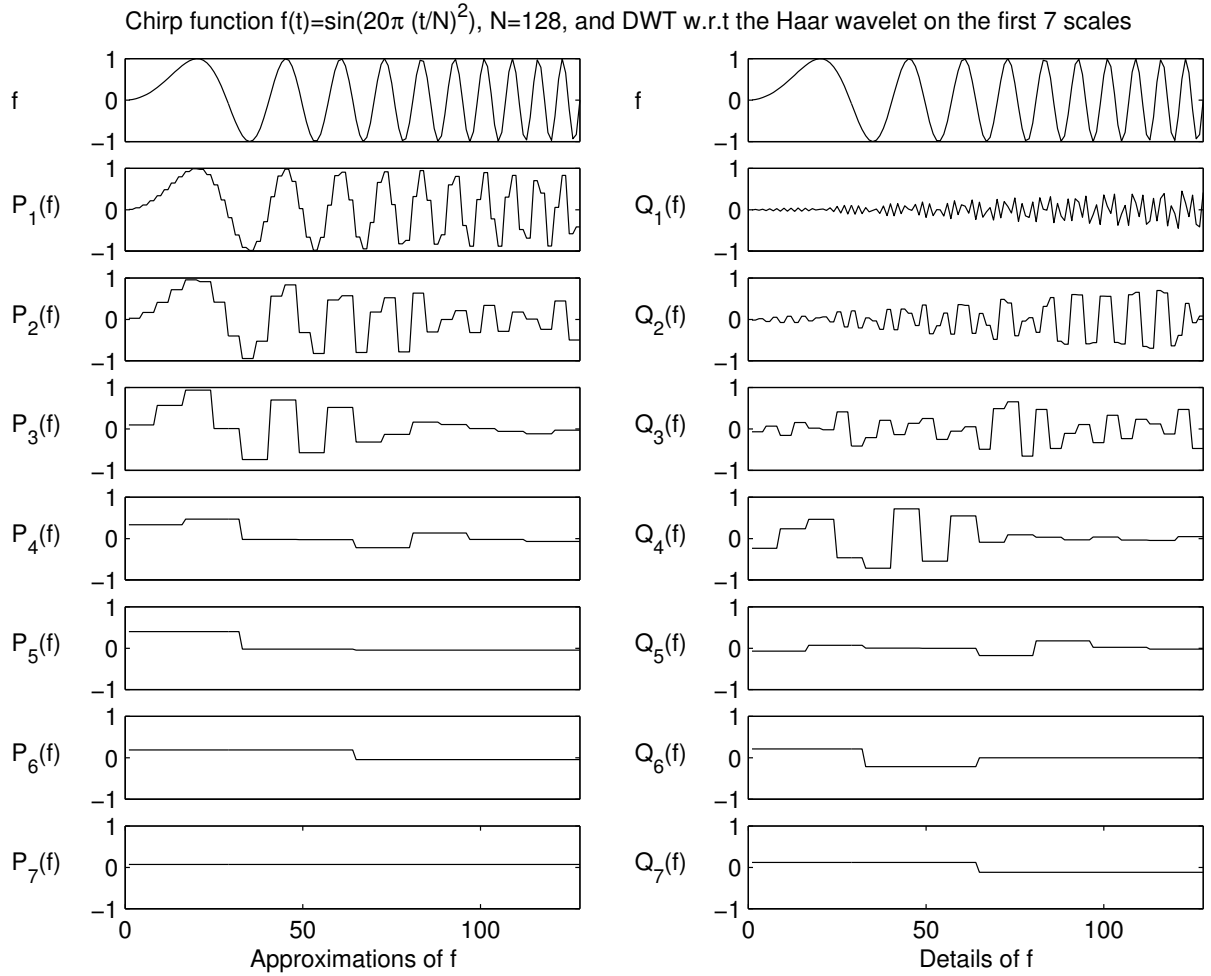


Figure 34: Haar-DWT of some chirp signals

Figure 34 illustrated very well the following principles.

- The wavelet coefficients of scales with small index encode the high-frequency components of the signal f.
- On the first scale the "details" $Q_1(f)$ increase in time which corresponds to the increasing frequency of the chirp signal f.
- The projection $P_1(f)$ is the difference of the signals f and the detail $Q_1(f)$.
- Computing the second scale, the projection $P_1(f)$ is further decomposed into a high-frequency component, the detail $Q_2(f)$, and a low-frequency component $P_2(f)$. $P_2(f)$ is a smoothed version of $P_1(f)$.
- This procedure is now iterated up to projections $P_7(f)$ which is a constant function on $[1 : N]$. This is the case, when in the DFT there is only one non-trivial coefficient left in the sequence v^M (here $M = 7$).

Similarly, Figure 35 shows a signal f which is the superposition of two sines of frequencies $\omega = 50$ and $\omega = 5$, respectively, with two impulses at $t = N/4$ and $t = N/2$, $N = 128$. The interpretation of the projections $P_i(f)$ and $Q_i(f)$ for $i = 1, 2 \ldots, M$, $M = 7$, is as in the last example. The impulses can be very well identified in the high-frequency details $Q_1(f)$ und $Q_2(f)$. The sine of frequency $\omega = 50$ is represented mainly in detail $Q_2(f)$, whereas the sine of frequency $\omega = 5$ is mainly reflected by the projections $P_2(f)$ and $P_3(f)$.

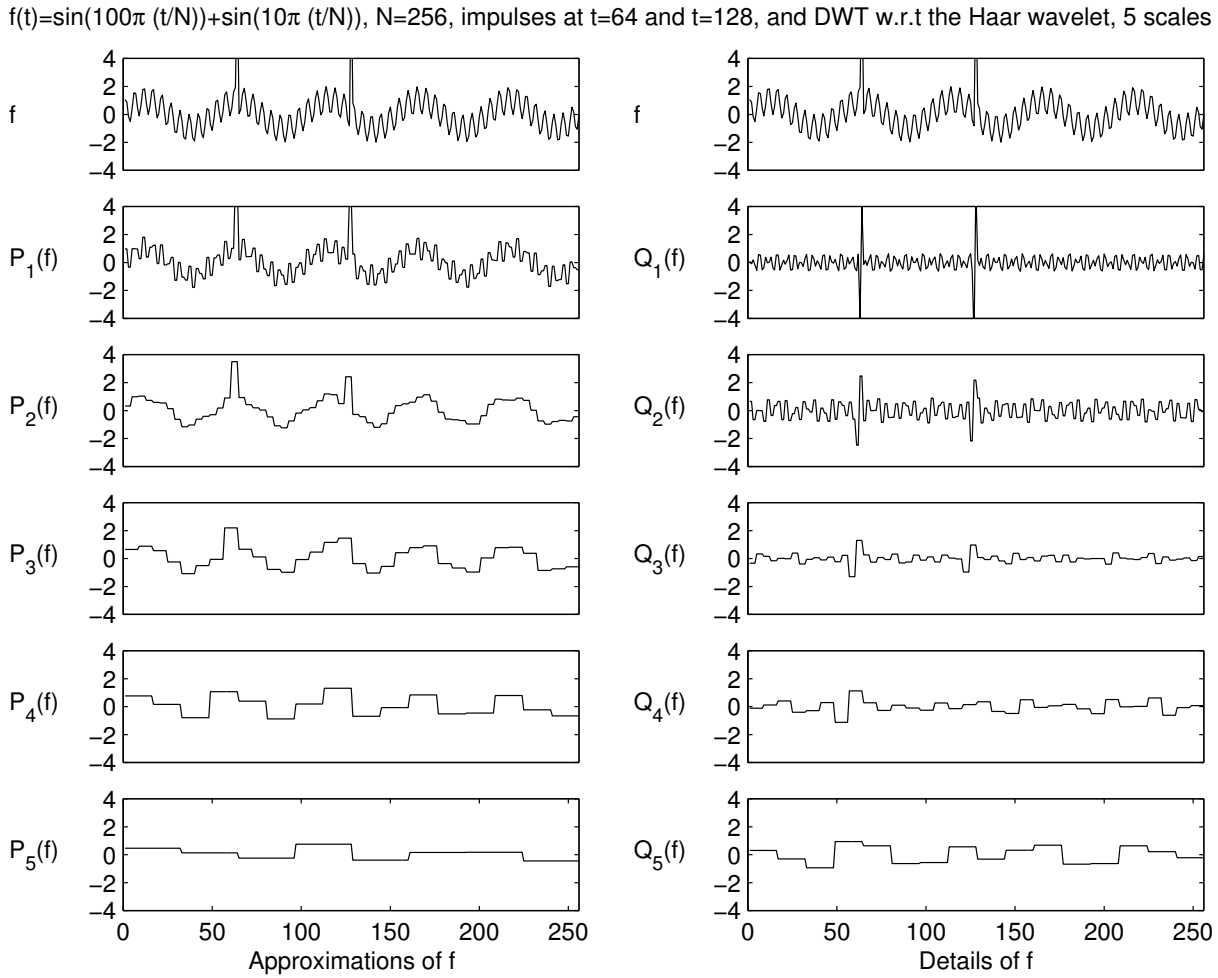


Figure 35: Haar-DWT of the superposition of two sines with two impulses

Chapter 9: DWT-Based Applications

In the previous chapters we have studied various transforms of a signal f such as the Fourier transform, the WFT and the DWT. The signal f gives, for example, information about the amplitude of some waveform at a given point in time or the pixel value of some picture at a given point in space. The goal of the transforms is to exhibit other information about the signal such as the frequency content of the signal or the occurrences of certain singularities (e.g., in higher derivatives of f which cannot be seen in the time-representation). The transforms provide a different representation of the signal, one also often speaks of transformation domain, in which properties of f can be read off which can not be seen in the time domain.

For example, from the Fourier transform of f or from the Fourier domain one can read off the spectral content of f. Or in the DWT domain, i.e., from the wavelet coefficients, one can determine which details (scale parameter s) occur in f at a given time (time parameter t) in which intensity (size of wavelet coefficient).

If no information is lost by applying the transform, i.e., if the transform is invertible, the signal f can be reconstructed from the data of the transformation domain.

Of course, it does not make sense to transform a signal and to immediately reconstruct it. Between these two steps the actual signal processing takes places. We transform the signal signal f by a suitable transform T and obtain a transparent representation $T(f)$ in view of the application in mind. The signal processor then processes the data $T(f)$ resulting in some modified data $T(f)^*$. The inverse transform T^{-1} reconstructs from this modified data a signal in the time domain denoted by $f^* := T^{-1}(T(f)^*)$.

All transforms T we have considered so far are continuous operators with continuous inverse operator T^{-1}. This property is important in signal processing since small modifications of the data $T(f)$ also result only in small deviations of the original signal when transformed back into the time domain.

In this chapter we discuss some DWT-based applications. Important is the property of the wavelet transform that for a large class of relevant signals the energy of a signal is concentrated in few wavelet coefficients when transformed to the DWT domain. This property can be exploited to develop methods to represent a signal in a compressed form. For example, the FBI uses DWT-based compression methods to store finger prints in a digital library.

Furthermore, the property of the DWT that the essential properties of a signal are reflected by only few, but large wavelet coefficients can be used to detect noise in a signal and to possibly separate it from the signal. DWT-based denoising methods have been successively applied to examine a recording played by Brahms in 1889 of one of his own compositions (a version of the Hungarian Dance No. 1). In 1935 an LP disc was directly cut from the original wax cylinder recording. However, the quality of this record was so poor — the actual interpretation was submerged into noise — that the recording was more or less useless for musicological examinations. Using DWT-based methods it was possible for the first time to extract musicologically relevant information, which exhibited interesting information about the way how Brahms interpreted one of his own pieces. For example, one found out that Brahms took the liberty to considerably deviate from the score and extemporize freely (see [Hubbard] for more details).

9.1 DWT-Based Denoising

The main of idea of DWT-based denoising can be summarized as follows:

(1) The noisy signal f is transformed by a suitable DWT.

(2) In the transformation domain, $T(f)$ is processed by some thresholding method, i.e., wavelet coefficients are removed which lie beyond some suitably chosen thresholds. This results in the data $T(f)^*$.

(3) The reconstruction by means of the inverse DWT from the modified wavelet coefficients $T(f)^*$ gives the denoised signal f^*.

This method is based on the assumption that the energy of the signal f in the DWT domain is concentrated in few, but large wavelet coefficients. In other words, the essential properties of f are captured in a small number of wavelet coefficients and these coefficients are large in absolute value. In contrast, the energy of noise-like components is spread over the whole range of wavelet coefficients which then are small in absolute value.

Therefore, by means of thresholding as explained above the small wavelet coefficients corresponding to the noisy components are removed in the DWT domain and the desired signal is reconstruced by the inverse DWT — in general, at the cost of some acceptable loss of details.

In classical denoising methods, the Fourier transform was used to separate the actual signal from the noise components in spectral domain, e.g., by means of linear filtering. However, this method could not applied in the case that the spectra of the actual signal and the noise components overlapped. The new DWT-based thresholding method is a non-linear process and is based on some different principle: not a frequency-based but an amplitude-based separation principle.

Concerning the denoising method described above, a lot of questions arise.

- What actually is noise? What are the noisy components? (For example, for a researcher, who is looking for oil in the sea, a submarine causes annoying and disruptive noise. However, for military the same submarine could be the actual signal.)

- In practice one has to deal with noisy signals where the signal is submerged in noise and is hardly perceptiable. How can one, in such a case, determine which components belong the actual signal and which belong to the noise component? How can one determine the noise level?

- Which criteria can be used to valuate the quality of the denoised and thus improved signal?

- How can the threshold used in the thresholding method be determined?

- Which wavelets should be chosen in the DWT?

There is no solution to these probleme and the choice of the suitable parameters depends very much on the class of signals under consideration. In this area of digital signal processing there are still many questions left open and it constitutes a current field of research. In the following, we go into some more detail concerning some of the questions above.

9.1.1 White Noise

One often uses the following fomula to model a noisy and destorted digital signal $(s(n))_{n\in\mathbb{Z}}$:

$$s(n) = f(n) + \sigma e(n),$$

where

$$f = (f(n))_{n\in\mathbb{Z}}$$

is the actual digital signal and

$$e = (e(n))_{n\in\mathbb{Z}}$$

is a $\mathcal{N}(0,1)$-distributed Gaussian white noise. The constant

$$\sigma \in \mathbb{R}^+$$

denotes the noise level. In this subsection we explain how white noise can be modeled. To this means we recall some basic notions from probability theory.

Let $(\Omega, \mathcal{A}, P)$ be a probability space (P-space). A random variable X (RV X) is a measurable map

$$X : \Omega \to \mathbb{R},$$

i.e., all preimages under X of some interval in $\mathbb{R}$ are in $\mathcal{A}$. An easy example is the so-called Bernoulli-RV

$$X : \Omega \to \{-1, 1\},$$

which only assumes the values 1 and -1. One has

$$P(X = 1) = p \quad \text{and} \quad P(X = -1) = 1 - p$$

for some $p \in [0, 1]$. Here, the notions $P(X = 1)$ is an abbreviation for $P(\{\omega \in \Omega : X(\omega) = 1\})$. The Bernoulli-RV models, for example, the experiment of throwing a coin, where the number 1 means heads and the number -1 means tails. In case of a "fair" coin one has $p = \frac{1}{2}$.

The distribution function of an RV X is defined by

$$F_X(\alpha) := P(X \leq \alpha).$$

In case F_X is differentiable, the probability density of X is defined by

$$f_X(\alpha) := \left(\frac{d}{d\alpha} F_X\right)(\alpha).$$

The mean value or expectation μ_X is a kind of average value, or point of gravity, of the RV X and is defined by

$$\mu_X := \mathrm{E}(X) := \int_{-\infty}^{\infty} \alpha f_X(\alpha) d\alpha.$$

The variance σ_X^2 and the standard deviation σ_X form a measure for the variation around this average. These values are defined by

$$\sigma_X^2 := \mathrm{Var}(X) := \mathrm{E}([X - \mathrm{E}(X)]^2) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$$

If there exist the values $\mathrm{E}(X^2)$ und $\mathrm{E}(Y^2)$ for two random variables X and Y, then the covariance $\mathrm{Cov}(X, Y)$ defined by

$$\mathrm{Cov}(X, Y) = \mathrm{E}((X - \mathrm{E}(X))(Y - \mathrm{E}(Y))) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$$

and the correlation coefficient ρ_{XY} defined by

$$\rho_{XY} := \mathrm{Cov}(X, Y)/(\sigma_X \sigma_Y)$$

exist. The random variables X and Y are called uncorrelated, if

$$\mathrm{Cov}(X, Y) = 0.$$

The Gaussian RV or normal distribution is a very important in probability theory. This RV is defined via its probability density

$$f_{\mu,\sigma}(\alpha) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\alpha-\mu)^2}{2\sigma^2}\right).$$

This is the Gaussian function with parameters σ and μ. Recall that we have seen this function already in connection with the Heisenberg uncertainty principle (see Subsection 6.3.1) One can compute that for the Gaussian RV X with parameters σ and μ holds

$$\mu_X = \mu \quad \text{and} \quad \sigma_X = \sigma.$$

One also says that X is a $\mathcal{N}(\mu, \sigma)$-distribution.

Definition 9.1. Under white noise one understands a sequence $(X_n)_{n\in\mathbb{Z}}$ of uncorrelated random variables X_n having the same mean values $\mu_{X_n} = \mu$ and the same variance $\sigma^2_{X_n} = \sigma^2$ for all $n \in \mathbb{Z}$. In other words,

$$\mathrm{Cov}(X_i, X_j) = 0, \quad \text{if} \quad i \neq j, i, j \in \mathbb{N}$$

and

$$\mathrm{Cov}(X_i, X_i) = \mathrm{Var}(X_i) = \sigma^2 \quad \text{for} \quad i \in \mathbb{N}.$$

We discuss this definition by means of two examples.

For the Bernoulli white noise all random variables X_n, $n \in \mathbb{N}$, are Bernoulli-RVs. In the case $p = \frac{1}{2}$ holds $\mu = 0$ and $\sigma = 1$. The RVs X_n are pairwise uncorrelated.

Again if we think of the RV as throwing a fair coin, each X_n corresponds to throwing the coin one time. The property that the X_n are pairwise uncorrelated means that the result of the ith throw does not depend on the result of the jth throw for $i \neq j$. Figure 36 shows a typical realization of Bernoulli-distributed white noise with $p = \frac{1}{2}$.

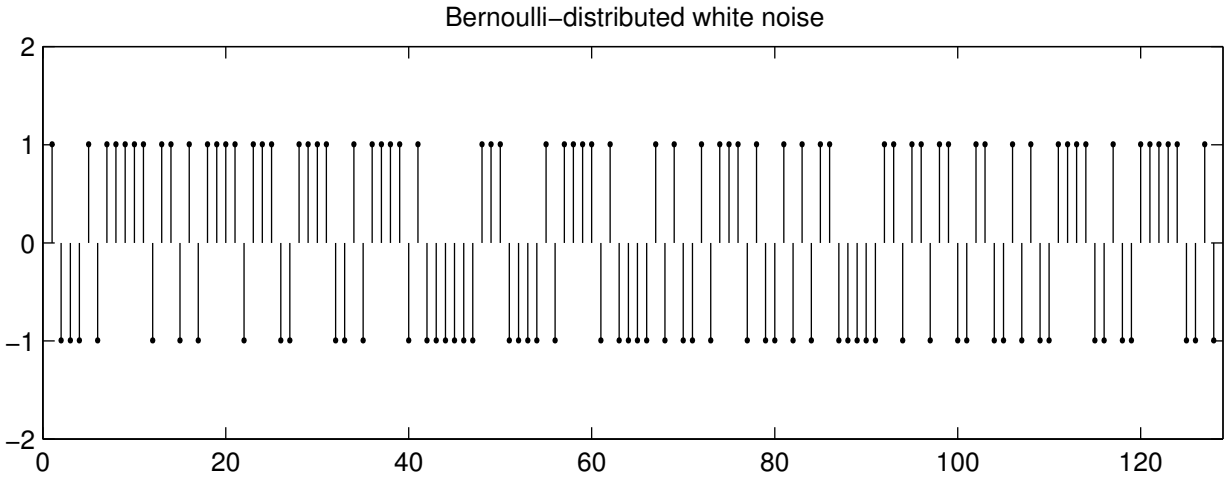


Figure 36: Bernoulli-distributed white noise with $p = \frac{1}{2}$.

For the Gaussian white noise all random variables X_n, $n \in \mathbb{N}$, are $\mathcal{N}(\mu, \sigma)$-RVs. Figure 37 shows a typical realization of $\mathcal{N}(0, 1)$-distributed Gaussian white noise with parameters $\mu = 0$ and $\sigma = 1$.

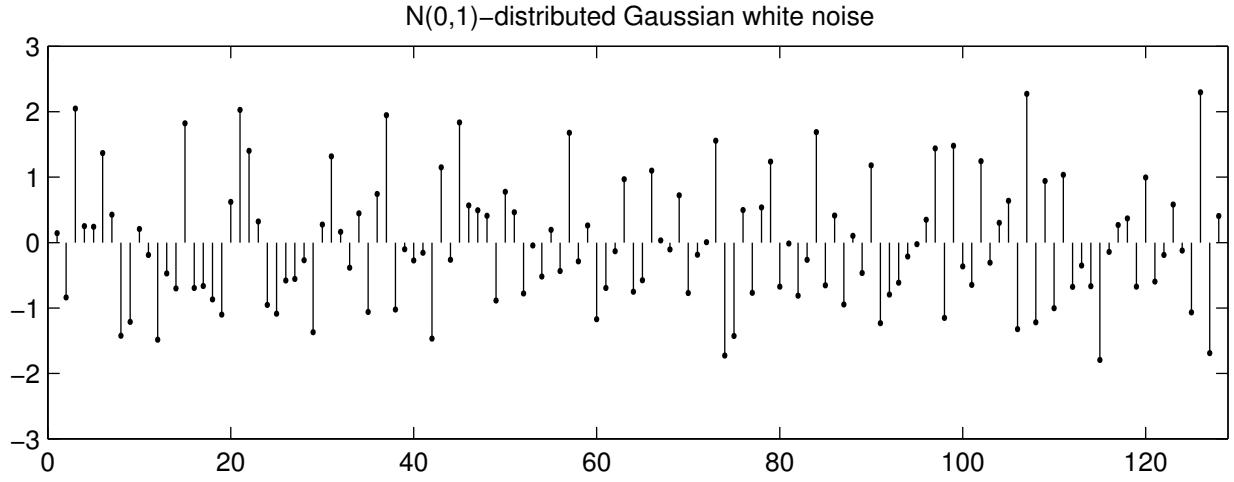


Figure 37: $\mathcal{N}(\mu, \sigma)$-distributed Gaussian white noise with $\mu = 0$ and $\sigma = 1$.

In audio signal processing one often assumes that the noise component $(e(n))_{n\in\mathbb{Z}}$ is $\mathcal{N}(0,1)$-distributed Gaussian white noise, i.e., $e(n)$ is a $\mathcal{N}(0,1)$-distributed RV for each $n \in \mathbb{Z}$. The noise level σ is modeled by a scaling factor which expresses the noise intensity contained in the signal. One can show that

$$\mathrm{Var}(\sigma e(n)) = \sigma^2 \mathrm{Var}(e(n)) = \sigma^2,$$

i.e., the standard deviation of the RV $\sigma e(n)$ equals the noise level σ. This explains why σ is also referred to as noise level.

The first line of Figure 38 shows a typical realization of $\mathcal{N}(0,1)$-distributed Gaussian white noise which was generated by a random processors. The signal is sampled with a sampling rate of $N = 1024$.

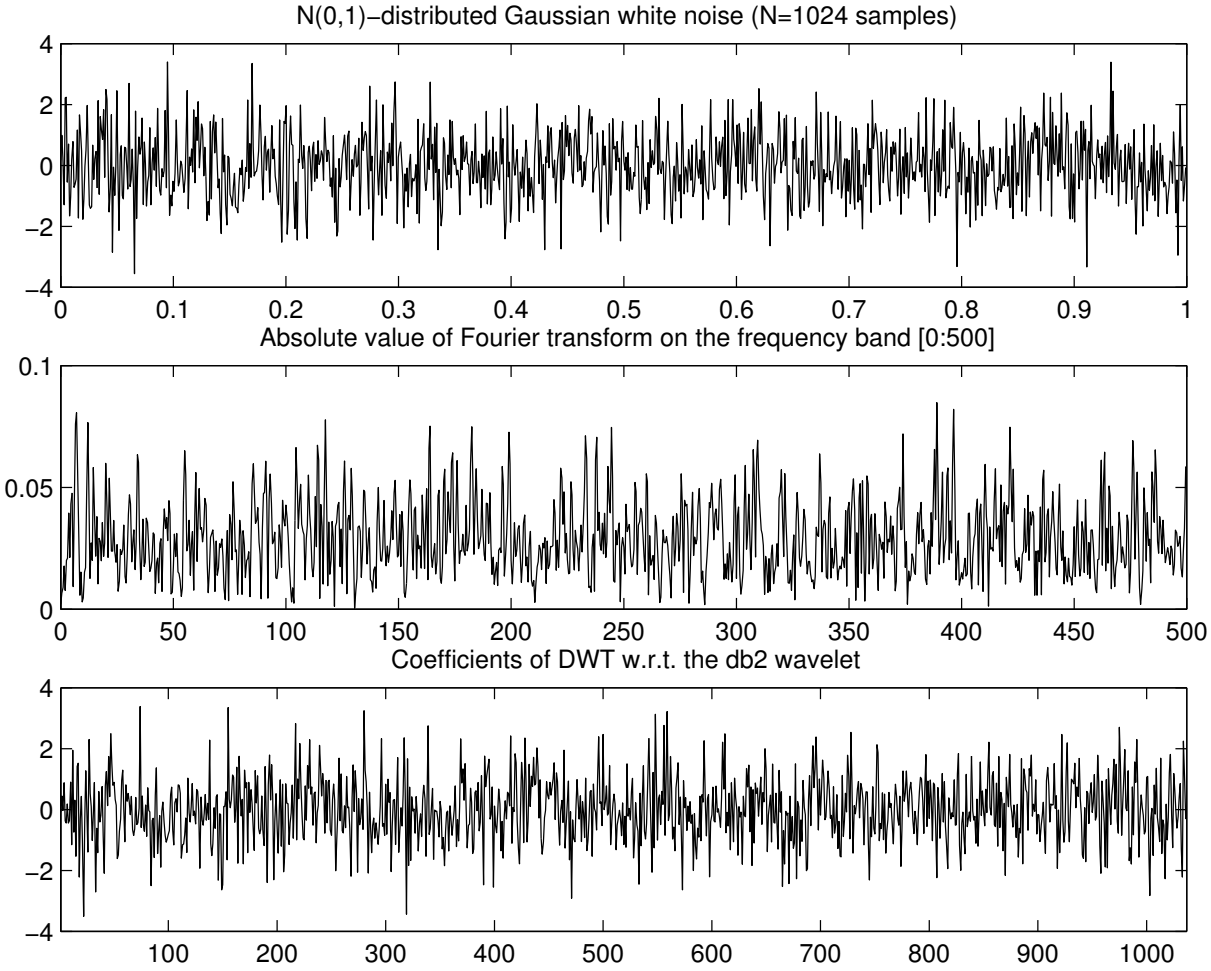


Figure 38: White noise in time domain, Fourier domain, and DWT domian.

White noise is characterized by the fact that all frequencies (im the statistical sense) appear with the same energy. This is also illustrated by the second line of Figure 38 which shows the absolute values of the Fourier transform: no frequency band is distinguished from the others. Also in the frequency domain the noise looks like noise.

This observation can also be made when transforming the noise signal by some orthogonal DWT. The last line of Figure 38 shows the wavelet coefficients of the noise signal computed on the first $M = 10$ scales w.r.t. the db2 wavelet. Here the M sequences $w^1, w^2, \ldots, w^M$ and the sequence v^M are — as in Equation (68) of Chapter 8 — written as one sequence

$$w := (v^M, w^M, w^{M-1}, \ldots, w^1),$$

where the length of this sequence is about the same as for the noise signal ($N = 1024$). Note that again the sequence w of wavelet coefficients looks just as noisy as the noise signal itself. The energy of the white noise is uniformly distributed over all wavelet coefficients; there is no concentration of the energy in a small number of wavelet coefficients.

9.1.2 Thresholding

Donoho describes in [Donoho] two methods to remove "disturbing" wavelet coefficients.

(1) In the so-called hard thresholding method all coefficients are set to zero whose absolute value lie below a given threshold δ. Hard thresholding can be defined as a function T_{hard} on input coefficients c:

$$T^{\delta}_{hard}(c) := \begin{cases} c & \text{if} \quad |c| \geq \delta \\ 0 & \text{if} \quad |c| < \delta. \end{cases}$$

(2) The so-called soft thresholding is an extension of hard thresholding. Again all coefficients are set to zero whose absolut value lie below a given threshold δ. However, this time all other coefficients are also modified such that their absolute values reduces by δ. This procedure is also known as wavelet shrinkage. Soft thresholding is defined

by the function

$$T^{\delta}_{soft}(c) := \begin{cases} \operatorname{sign}(c)(|c| - \delta) & \text{if} \quad |c| \geq \delta \\ 0 & \text{if} \quad |c| < \delta. \end{cases}$$

The two functions T^{δ}_{hard} and $T^{\delta}_{soft}(c)$ are also shown in Figure 39.

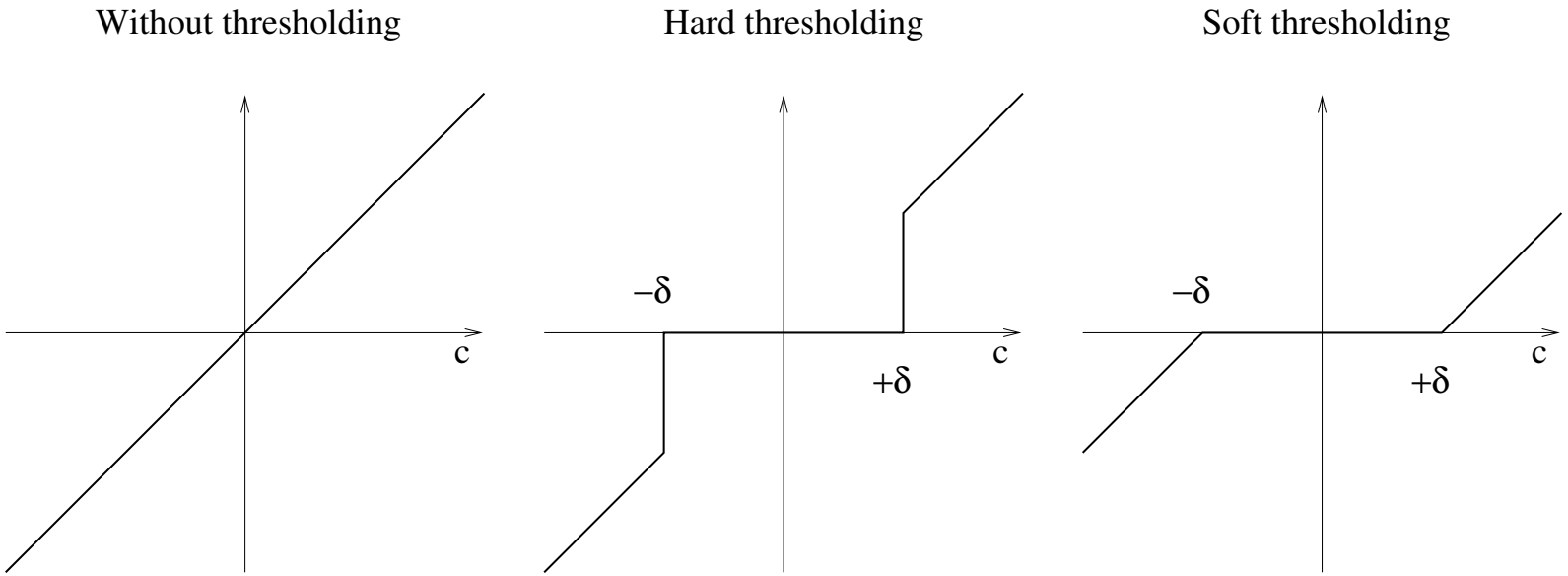


Figure 39: Hard and Soft-Thresholding.

9.1.3 Choice of the Threshold

In the following let $T = DWT$ denote the discrete wavelet transform w.r.t. to some fixed orthogonal wavelet such as a Daubechies wavelet. Transforming an input signal $x = (x_1, x_2, \ldots, x_N)$ of length $N = 2^M, M \in \mathbb{N}$, by T on the first M scales, the output sequence is given by

$$w := (v^M, w^M, w^{M-1}, \ldots, w^1).$$

By a small modification of the DWT algorithm one can achieve that the output length of w is also N and T is an orthogonal transform. Then, fixing bases, T can be represented by an orthogonal $N \times N$-matrix denoted by T_N. For simplicity denote the sequence of wavelet coefficients by

$$w^x := T_N(x) = (w_n^x)_{0 \le n < N}.$$

As in subsection 9.1.1 we use again the model assumption

$$s(n) = f(n) + \sigma e(n).$$

We recall that a $\mathcal{N}(0,1)$-distributed noise signal $(e(n))_{n\in\mathbb{Z}}$ transforms again into a $\mathcal{N}(0,1)$-distributed noise signal under some orthogonal transform. Therefore, the wavelet coefficients $w^e(n)$ are also uncorrelated $\mathcal{N}(0,1)$-distributed random variables.

Comparing the wavelet coefficients $w^s(n)$ of the noisy signal s with the wavelet coefficients $w^f(n)$ of the original signal, the expected deviation is σ in a statistical sense, namely the standard deviation of the RV

$$w^s(n) - w^f(n) = \sigma w^e(n).$$

Therefore, the $\delta := \sigma$ would be a good choice for the threshold δ. Actually, one can show that this choice in the hard thresholding procedure leads in a statistical sense (w.r.t. to the so-called risk measure) to an optimally denoised signal f^*.

In practice, however, we have to deal with a completely different situation: neither the original signal f, nor the kind of noise and the noise level σ are known. The only information we have on hand is the noisy signal s.

In the following, we assume that the noise signal is an $\mathcal{N}(0,1)$-distributed white noise. In many applications this is a good model, even though an idealistic assumption. (Note that there are also many applications where this assumption does not reflect the kind of noise one has to cope with!)

Then the noise level σ has to be estimated and the threshold δ has to be suitably chosen. Estimation of parameters such as σ is itself a complicated problem with no standard solution. The concepts of estimators is part of probability theory. For details on this topic we have to refer to the literature.

Some procedures to estimate the noise level δ have been implemented in MATLAB (see `thselect`). Among others there are

- Steins unbiased estimate of risk (quadratic loss function). This leads to the so-called SURE thresholding.
- Minimax estimators.
- Combination of methods.

Under the assumption that such a method is at hand, we now want to describe the actual thresholding algorithm.

9.1.4 Algorithm for Denoising

Given: Digital noisy signal s with N sampling values.

Possibly information on the properties of the original signal (the original signal is in general not known, but one often can assume that the original signal lies in some signal class, e.g., possessing some smoothness properties.

Assumption: The noise component is modelled by an $\mathcal{N}(0,1)$-distributed Gaussian white noise.

Parameters: Choice of a suitable wavelet on which bases the DWT is performed.

Choice of a suitable threshold δ by means of some estimator described above.

Choice of thresholding method (hard or soft).

In view of the anticipated properties of the original signal one chooses suitable parameters. Note that this choice depends on the respective application. The denoising algorithm proceeds in three steps:

(1) Transform the noisy signal s by means of the DWT. In the wavelet domain this signal is represented by the wavelet coefficients $(w^s(n))$.

(2) Apply hard or soft thresholding with threshold δ on the wavelet coefficients (signal processing in the transform domain).

(3) Reconstruct the signal from the so modified wavelet coefficients (inverse DWT) and obtain the denoised signal f^*.

This algorithm can be implemented in MATLAB as follows. The Examples in Subsection 5 were generated by such a program.

```
%DWT-based Denoising by Thresholding implemented in MATLAB
clear;
N=1024;                          %number of samples
t=1:N;
wavelet='db8';                   %wavelet
M=floor(log2(N));                %number of scales
e = randn(1,N);                  %white N(0,1)-distributed noise
sigma = 0.3;                     %noise level sigma
f=sin(50*pi*(t/N)^ 2)            %original signal (chirp signal)
s=f+sigma*e;                     %noisy signal
thr=thselect(s,'rigrsure');      %threshold
```

```
%DWT of s, wavelet coefficients are w
[w,l]=wavedec(s,M,wavelet);

%hard thresholding, modified wavelet coefficients are whard
whard=w;
for k=(l(1)+1):length(whard),
   if abs(whard(k))<=thr whard(k)=0;   end
end

%soft thresholding, modified wavelet coefficients are wsoft
wsoft=whard;
for k=(l(1)+1):length(wsoft),
   if wsoft(k)<>0 wsoft(k)=sign(wsoft(k))*(abs(wsoft(k))-thr);end
end
```

```
%reconstruction, denoised signal fhard and fsoft
fhard = wrcoef('a',whard,l,wavelet,0);
fsoft = wrcoef('a',wsoft,l,wavelet,0);

%plot original signal f, noisy signal s and
%denoised signals fhard, fsoft
subplot(4,1,1); plot(t,f); title('original signal f')
subplot(4,1,2); plot(t,s); title('noisy signal s')
subplot(4,1,3); plot(t,fhard); title('denoised signal fhard')
subplot(4,1,4); plot(t,fsoft); title('denoised signal fsoft')
```

The result of this program is shown in the left column of Figure 40.

9.1.5 Examples for Denoising

In this example, the original signal f is a chirp signal as shown in Figure 40. A DWT was performed w.r.t. to the db8 wavelet. As described before, the output data is written as one sequence

$$w := (v^M, w^M, w^{M-1}, \ldots, w^1),$$

where w^m, $m = 1, 2, \ldots, M$, denote the sequences of wavelet coefficients of scale m, v^M denotes the sequence of scaling coefficients of scale M We recall that the sequence w^m has roughly the length $N \cdot 2^{-m}$.

The right column of Figure 40 shows the sequence w. Note there are only few wavelet coefficients of f which have large absolute value. In other words, the energy of f is concentrated in few wavelet coefficients.

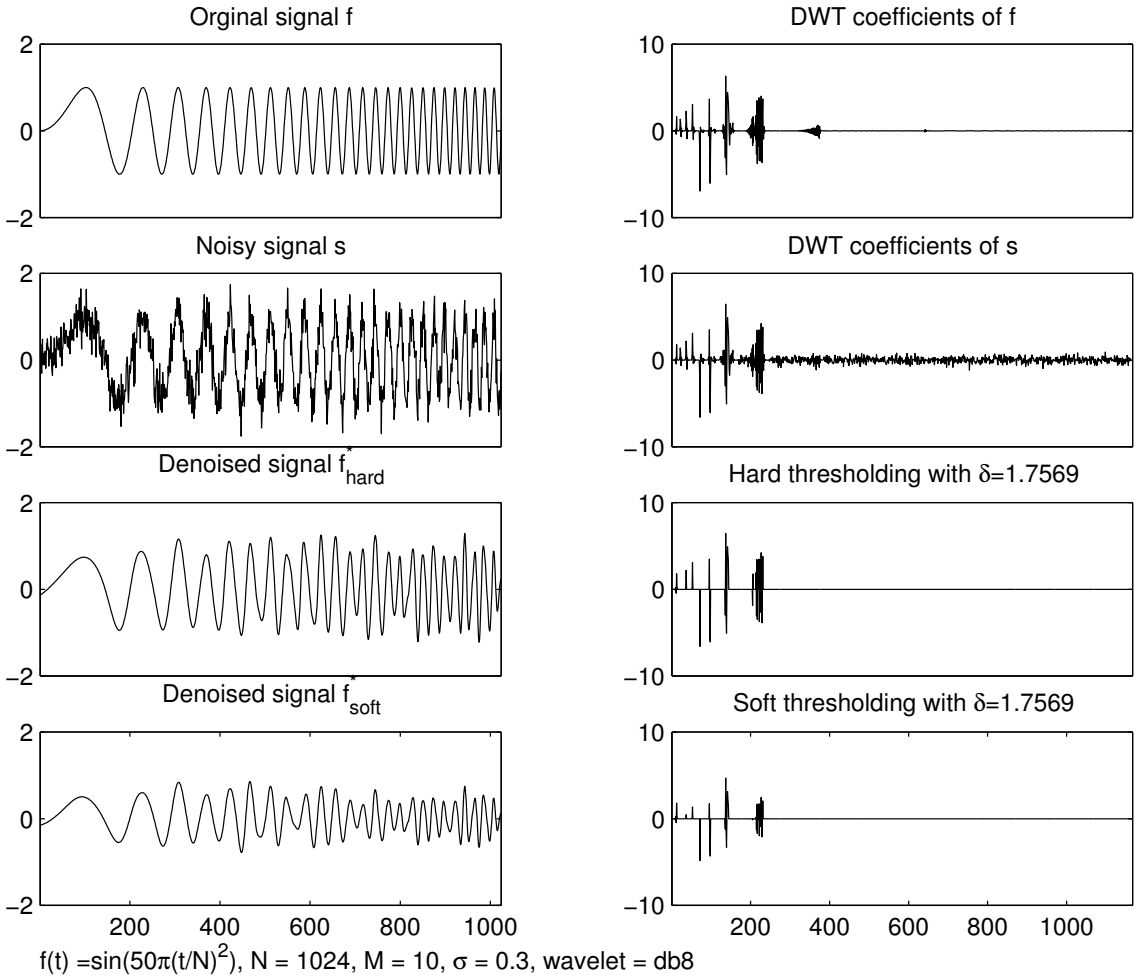


Figure 40: Denoising of a noisy chirp signal.

From the illustration of the DWT of the noisy signal s one recognizes that the noisy component σe results in many wavelet coefficients on all scales of small absolute value. Using the SURE procedure, the threshold δ was chosen (note that only s is known, but not f, e or σ). Using hard thresholding and soft thresholding the wavelet coefficients of the noisy signal were modified and from these the signals f^*_{hard} and f^*_{soft}, respectively, reconstructed. The result of this procedure is shown in Figure 40.

The choice of the wavelet db8 was motivated as follows: assuming that the original signal f is of high regularity (smooth signal), we also picked a wavelet of high regularity, since smooth properties of the signal can be better synthesized by smooth synthesis functions.

The last remark is illustrated by the next example. Here, the original signal consists of some discontinuous part with a peak (left part) and a smooth part (right part). The thresholding procedure has been applied with two different wavelets: the db16-Wavelet of high regularity and the discontinuous Haar wavelet of low regularity. The denoising algorithm w.r.t. to both wavelets is illustrated in Figures 41 and 42, respectively.

As is seen from the right column of Figure 42, the skyline-like part of f is very efficiently encoded by the Haar wavelet which leads to few, but large wavelet coefficients in the big scales. However, the smooth sine-like left part of f is only poorly approximated by the discontinuous Haar wavelets leading to many small wavelet coefficients spread over all scales. This last fact leads to problems in the denoising procedure: For the noisy signal s the small wavelet coefficients coming from the original signal f cannot be distinguished (by means of the amplitude) from the small wavelet coefficients coming from the noise component σe.

In the thresholding procedure, all wavelet coefficients are set to zero whose absolute value lies below the threshold. Therefore, also many wavelet coefficients which actually encode the sine-like part of f are set to zero. As a result, in the denoised signals f^*_{hard} and f^*_{soft} the skyline-like part is very well reconstruced. The sine-like left part, however, is strongly distorted or nearly erased.

This example illustrates what can happen, if the original signal and the wavelet used for the DWT do not have corresponding properties w.r.t. regularity.

Instead of using the Haar wavelet we next use the db16 wavelet — a wavelet of high regularity. In this case, as is shown in Figure 41, the sine-like right part of f is encoded very efficiently but not the skyline-like left part of f. For the same reason as above, now the sine-like part is very well reconstructed in the denoised signals f^*_{hard} and f^*_{soft}, whereas the skyline-like part is considerably distorted. In f^*_{soft}, even the peak is erased. In such cases one also speaks from "over-smoothing".

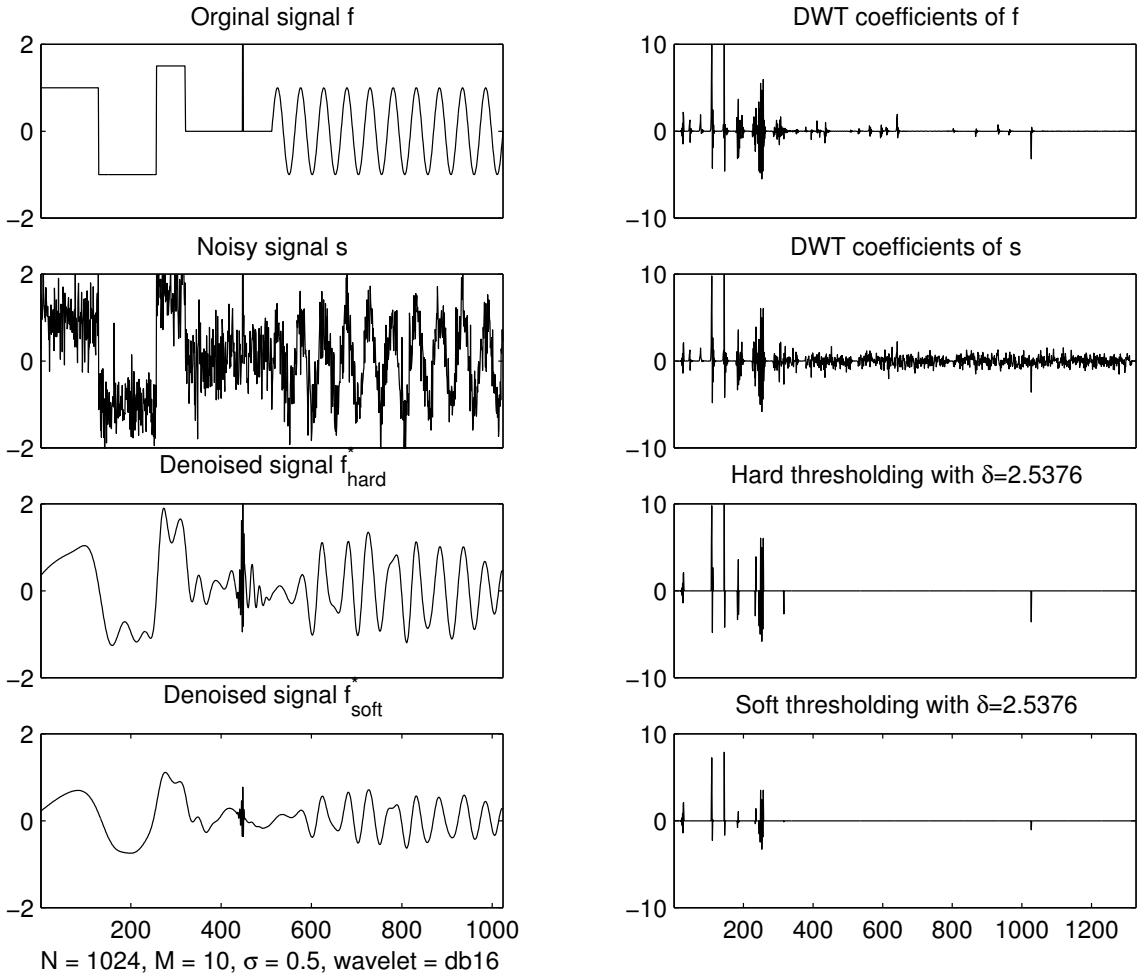


Figure 41: Denoising with wavelet of high regularity (db16 wavelet).

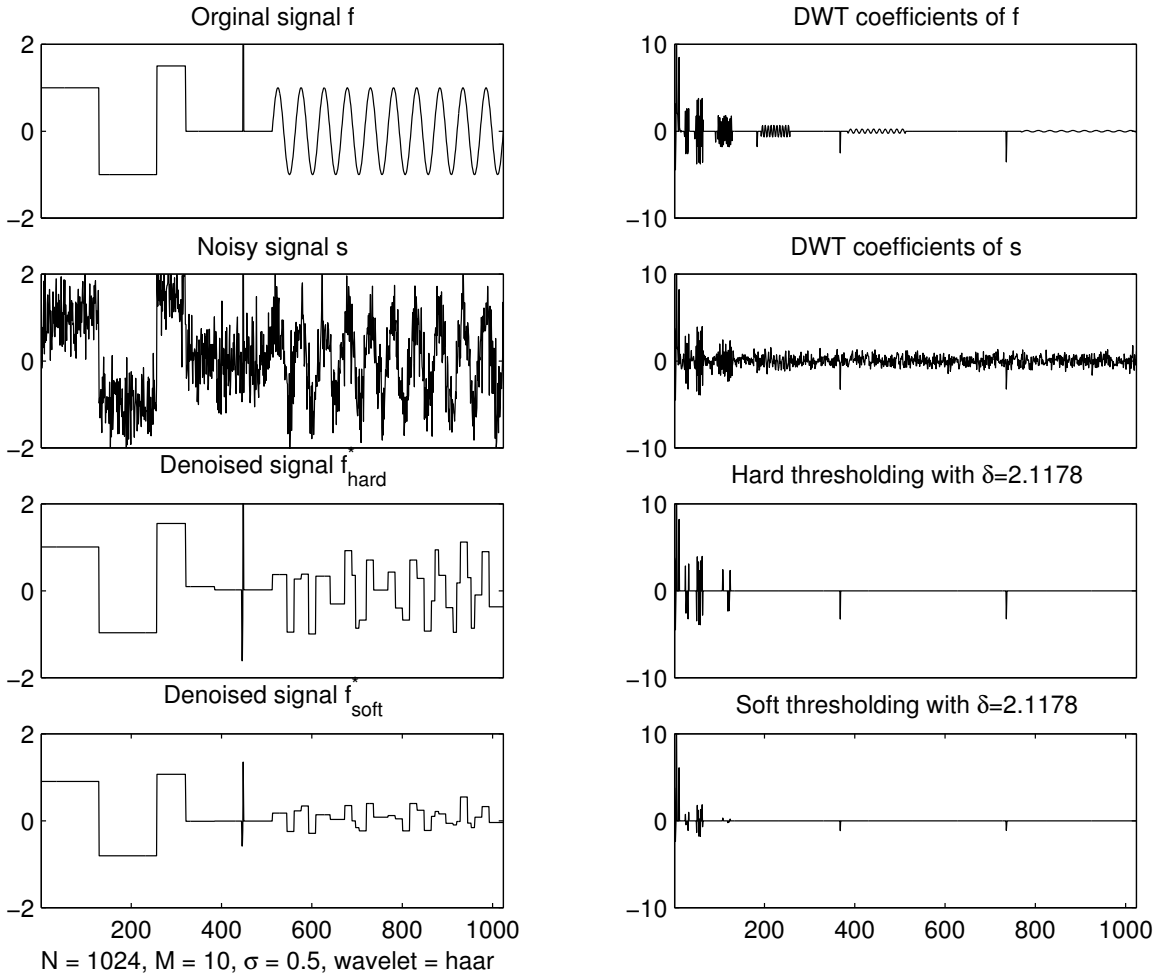


Figure 42: Denoising with Haar wavelet.

9.1.6 Problems and Remarks

First we want to comment on the differences between hard and soft thresholding. In hard thresholding, the noise component is in general quite well removed; however, in this procedure many artefacts such as psudo-oscillations and small abrupt distortions may arise in the reconstructed signal which substantially degrade the quality. These artefacts are small when measured in the ℓ^2-norm (i.e., the ℓ^2-distance between f and f^*_{hard} is small), but they may be responsible that the resulting signal f^*_{hard} does not have the same regularity properties as the original signal f.

For that reason, one often uses soft thresholding in the denoising procedure. In this case, one has to pay with an additional shrinkage of the wavelet coefficients leading to a loss of the signal's energy (and therefore a larger ℓ^2-distance between f and f^*_{soft}). However, one can show that the denoised signal f^*_{soft} lies with high probability in the signal class concerning the regularity as the original signal f.

In other words, with high probability f^*_{soft} is at least as smooth as the original signal f, where this assertions holds for a large class of regularity measures (e.g., in the sense of n-fold differentiability).

The proof of this result and further details can be found in [Donoho].

As in traditional denoising procedures, e.g., by means of lowpass filtering, one has in the DWT-based approach a trade-off between denoising and "over-smoothing" of the signal. The larger the threshold δ, the better the denoising. However, at the same time more and more details of the original signal f are lost.

For example, in Figure 41 the impulse of the original signal f in the denoised signal f^*_{soft} is completely erased.

Finally, we want to indicate some possible refinements of the denoising algorithm. So far, we have used in the thresholding procedure one global threshold δ for all wavelet coefficients — independent of the scale m.

For many applications it has turned out to be advantageous, if a scale-dependent thresholding in the denoising process. In other words, for each scale m, $1 \leq m \leq M$, a separate threshold δ_m is chosen. This idea is based on the following observations:

(1) In practice, the noise signal e may not be a $\mathcal{N}(0,1)$-distributed Gaussian white noise where the energy is uniformly spread over all frequencies, but a so-called colored noise where the energy is only spread over certain frequency bands. Often one deals with high-frequency noise so that by this noise the wavelet coefficients of low scales (small details) are affected.

(2) The perception of the human auditory system depends very much on the frequency band components of the signal in question. For example, for some relatively smooth original signal one may not be as uncomfortable with some low-frequency noise as with some high-frequency noise.

In particular, using results from psychoaccoustics one has achieved considerable progress in denoising procedures.

One should not lose sight of the actual goal of denoising: the goal is not to extract a signal from a noisy signal which has nice mathematical properties such as smoothness (mathematical quality measure). The goal is to extract a signal which, for the human listener, is sensed as low-noise signal (subjective quality measure).

9.2 DWT-Based Compression

The DWT-based compression is based on the same principles as the DWT-based denoising discussed in the last section. Fundamental is the property of the DWT, that the energy and therefore the essential information of a large class of signals, which are relevant in practice (such as audio or video signals), is concentrated in a relatively small number of large wavelet coefficients.

Let f be the signal to be compressed, $w^1, w^2, \ldots, w^M$ be the sequences of wavelet coefficients of the first M scales, and v^M the sequence of scaling coefficients of scale M.

The compression algorithm is based on the following principle: If the signal f is sufficiently smooth, then for a suitable chosen M only a small number of scaling and wavelet coefficients are sufficient for a good approximation of f.

Similar to the denoising procedure, the compression algorithm consists of three steps:

(1) Transform the signal f by the DWT w.r.t. to some suitably chosen wavelet.

(2) Apply thresholding on the wavelet and scaling coefficients (signal processing in the DWT domain).

(3) By the inverse DWT an approximation of the original signal can be reconstructed.

- Suppose the original signal f is given by a vector of length N. Applying the DWT one obtains about the same number of scaling and wavelet coefficients.
- By thresholding the number of non-zero coefficients can be reduced dramatically depending on the signal class.
- For the reconstructed signal f^c it suffices to store these non-zero coefficients together with some additional information such as the wavelet used in the DFT and the scale and translation parameters of the non-zero coefficients.
- In other words, storing the approximation f^c in the transform domain is much cheaper than storing the original signal f which amounts in a (lossy) compression of f.

For the compression algorithm there are two strategies. In the first strategy, one uses a single threshold δ for all scaling and wavelet coefficients. In the second strategy, one uses different thresholds depending on the respective scale of the coefficients.

In the first strategy, one chooses the largest wavelet and scaling w.r.t. their absolut value and all other coefficients are set to zero. The number of the non-zero coefficients can be determined in different ways such as

- Chose a global threshold δ in the thresholding procedure.
- A compression rate is given which is to be achieved (e.g., $1:10$). Then choose, according to this rate, the largest coefficients in the DWT domain. For example, for the compression rate $1:10$ one sorts the coefficients according to their absolute value in decreasing order and then pick the first ten percent. The other 90% of the coefficients are set to zero.
- A bound for the approximation error is given. To be more precise, the ℓ^2-performance of f^c relative to f w.r.t. the ℓ^2-norm is measured in percent and defined by

$$\frac{\|f^c\|}{\|f\|} \cdot 100\%.$$

  Then the number of coefficients to be stored depends on the given bound of the ℓ^2-performance which has to be exceed by the reconstruction f_c.

In all three cases, the choice of the scaling and wavelet coefficients set to zero depend on only one parameter (namely, the threshold δ, the compression rate, or the ℓ^2-performance).

In the second strategy the threshold is chosen subject to the respective scale. In other words, one has several thresholds

$$\delta_1, \delta_2, \ldots, \delta_M.$$

Such a strategy is, for example, the so-called Birgé-Massart strategy, which is based on the "adaptive functional estimation in regression or density contexts". For details we refer to [Birgé/Massart].

Although this estimation procedure has a complicated mathematical background, it is easy to implement. Let $M \geq 2$ be the number of scales to be computed in the DWT. As usual $\ell(w^M)$ denotes the length of the sequence w^M consisting of the wavelet coefficients of the highest scale. Furthermore, let α be a real number greater than one and m_0 be an integer with $1 \leq m_0 \leq M$. The numbers m_0, $\ell(w^M)$, and α determine the strategy:

- On the scales $m_0 + 1, \ldots, M$ the coefficients are not modified, i.e., $\delta_{m_0+1} = \ldots = \delta_M = 1$.

- For the scale m with $1 \leq m \leq m_0$ one picks the k_m largest (w.r.t. to their absolut value) wavelet coefficients in w^m, where k_m is defined by

$$k_m := \frac{\ell(w^M)}{(m_0 + 1 - m)^\alpha}.$$

In the compression algorithm one often sets the parameter $\alpha = 1.5$. This strategy is also used for denoising where one often sets $\alpha = 3$.

Finally, the compression algorithm is illustrated by two examples. Figure 43 shows again a chirp signal f, our favorite example, and the reconstructed signal f^c for different thresholding procedures. Here, we used the smooth db8 wavelet for the DWT.

Using hard thresholding with threshold $\delta = 1$, 95.2% of the scaling and wavelet coefficients are set to zero which corresponds to a compression rate of about $1 : 20$. The ℓ^2-performance is about 95.7%. This shows that most of the signal's energy is concentrated in a relative small number of coefficients in the DWT domain.

As one expect, increasing δ leads to a better compression rate, however at the expense of the ℓ^2-performance.

The interpretation of the example shown Figure 44 is similar.

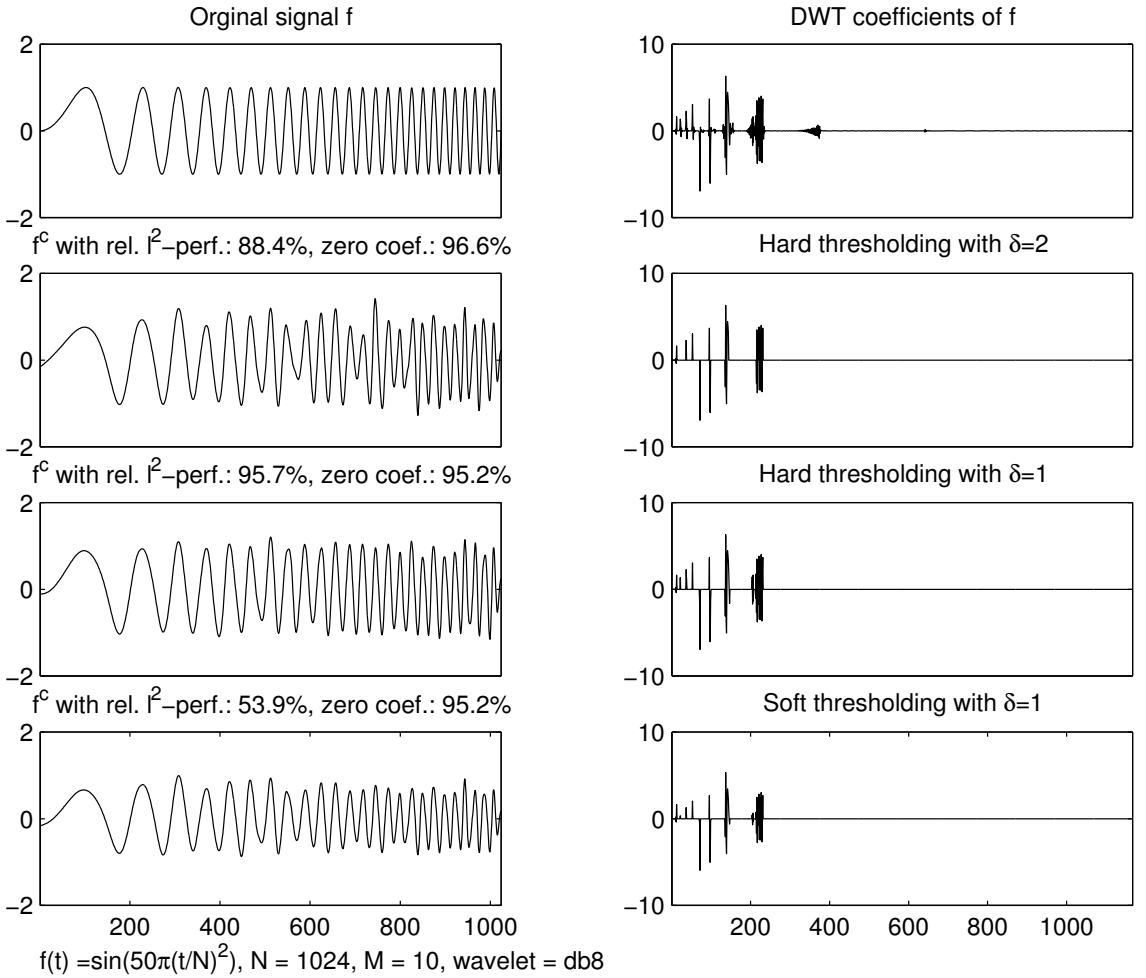


Figure 43: DWT-based compression of a chirp signal.

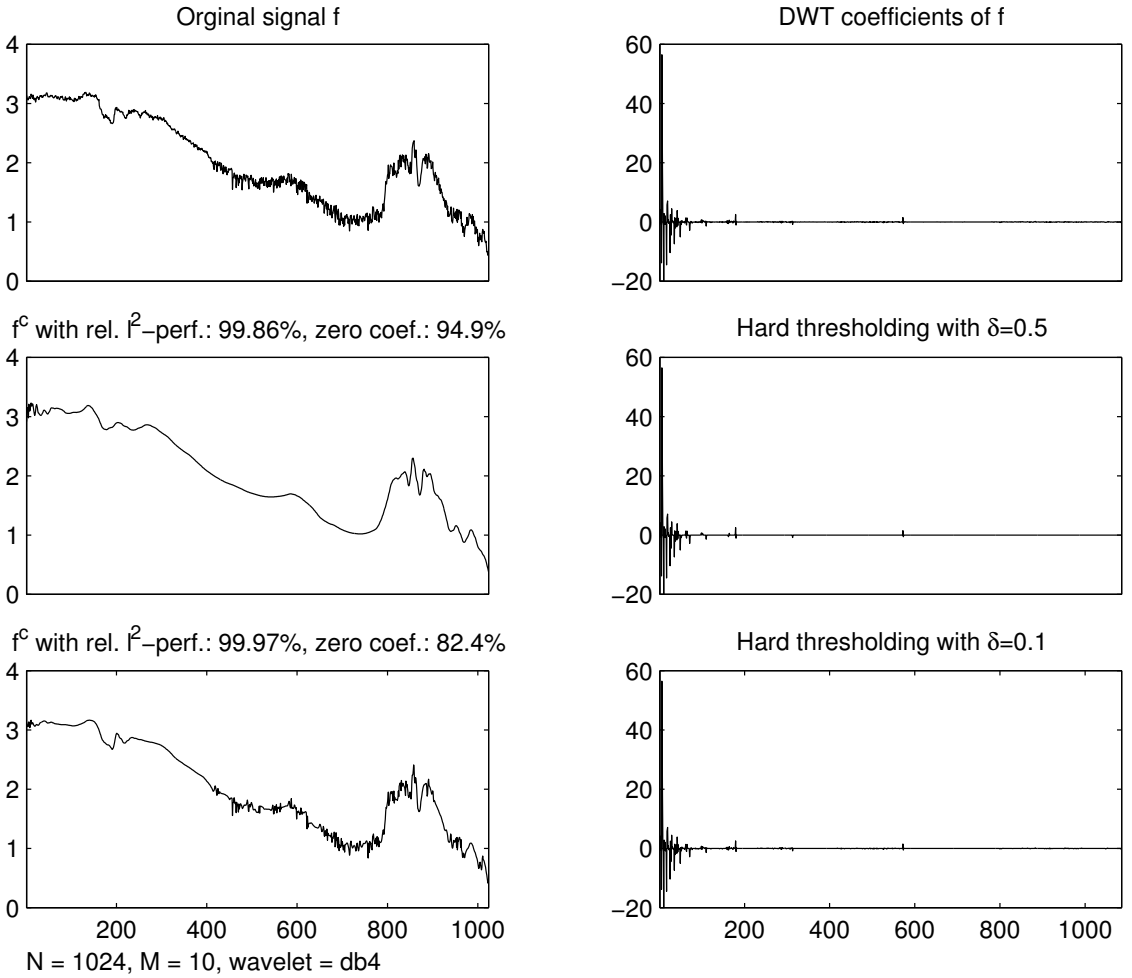


Figure 44: DWT-based compression of a noisy signal.

Warning: The excellent ℓ^2-performance in our examples should not be overvalued. Especially in the first example we are dealing with a synthetic signal of high regulartiy. The explicit knowledge of the signal was used to pick a suitable wavelet as well as to determine the threshold δ. To put it in an overstated fashion: if one wants to compress a pure sine wave, the knowledge of the period results in a compression using only one coefficient — one can just store the period itself using the sine function as analyizing function. This would result in a fantastic compression rate with perfect ℓ^2-performance.

However, typically one wants to compress signals such as natural images or audio signals. The choice of suitable parameters constitutes a serious problem. Picking the "wrong" parameters can lead to more or less useless algorithms even though the underlying theory has not changed.

When designing algorithms, the set of test signals on which these algorithms are tested and suitable parameters are investigated play a crucial role. In this testing and optimization process it often happens that the properties of the (seemingly representative) test data itself steal in the designing process. In other words, the improvements of the algorithms may be based on the specific properties of the test data. Applying the same algorithms to "real data" may lead to an unexpected poor performance.

In this sense, the examples in this script are solely an illustration of the underlying theory and do not describe the behaviour of the presented algortihms on real data.

Chapter 10: The Two Dimensional (2D) Case

The techniques presented in the earlier chapters can be generalized to higher dimensions. If one understands how to generalize the one-dimensional case to the two-dimensional case, then it is easy to figure out how to generalize to the n-dimensional case $n \in \mathbb{N}$. Therefore, we only discuss the 2D signals as a useful special case.

As was already mentioned in Chapter 1, two-dimensional signals are important in image rendering. In this case the two parameters express spatial information. However, a signal's two dimension need not be spatial. For example, some filtering techniques work in time and one spatial dimension.

10.1 2D-Signals and Systems

To make the analogy between the 1D case and the 2D case as explicit as possible, we use a straightforward generalization of the notation.

A 2D complex-valued continuous-time (CT) signal is given by a function

$$f : \mathbb{R}^2 \to \mathbb{C}.$$

For the two arguments we use the symbols t_1 and t_2.

A 2D complex-valued discrete-time (DT) signal is given by a function

$$x : \mathbb{Z}^2 \to \mathbb{C}.$$

For the two arguments we use the symbols n_1 and n_2, k_1 and k_2, etc.

As for the 1D case, the symbols $f, g, \ldots$ are typically used to denote CT-signals whereas the symbols $x, y, \ldots$ are typically used to denote DT-signals.

Example 10.1. The discrete time unit impulse δ in the 1D case (see Example 1.6) generalizes to the 2D unit impulse, also denoted by δ, as follows:

$$\delta(n_1, n_2) := \begin{cases} 1 & \text{if } n_1 = n_2 = 0, \\ 0 & \text{if } n_1, n_2 \neq 0. \end{cases}$$

As with the one-dimensional case, an arbitrary time-discrete 2D signal $x : \mathbb{Z}^2 \to \mathbb{C}$ can be expressed as a linear combination of shifted unit impulses:

$$x(n_1, n_2) = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} x(k_1, k_2)\delta(n_1 - k_1, n_2 - k_2).$$

Example 10.2. The discrete-time unit step function u in the 1D case (see Example 1.7) generalizes to the 2D unit step function, also denoted by u, as follows:

$$u(n_1, n_2) := \begin{cases} 1 & \text{if } n_1 \geq 0, n_2 \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

Example 10.3. A 2D exponential signal is of the form

$$a^{n_1} b^{n_2}$$

for $a, b \in \mathbb{R}$ and the 2D sinusoidal signal is of the form

$$A \cos(\omega_1 n_1 + \phi_1) \cos(\omega_2 n_2 + \phi_2)$$

with amplitude $A \in \mathbb{R}$, frequencies $\omega_1, \omega_2 \in \mathbb{R}$, and phases $\phi_1, \phi_2 \in \mathbb{R}$.

Definition 10.4. A two-dimensional signal is called separable if it can be expressed as a product of one-dimensional signals. To be more precise, $x : \mathbb{Z}^2 \longrightarrow \mathbb{C}$ is separable if it can be expressed in the form

$$x(n_1, n_2) = x_1(n_1) x_2(n_2), \quad n_1, n_2 \in \mathbb{Z},$$

for suitable 1D signals x_1 and x_2. An analog definition applies for CT-time signals $f : \mathbb{R}^2 \longrightarrow \mathbb{C}$.

It is easy to see that the signals in the Examples 10.1 to 10.3 are separable.

Completely analog to the 1D case one can define the Lebesgue spaces $\ell^p(\mathbb{Z}^2)$ for DT signals and $L^p(\mathbb{R}^2)$ for CT signals, $1 \leq p \leq \infty$. For example,

$$\ell^p(\mathbb{Z}) := \left\{ x \colon \mathbb{Z}^2 \to \mathbb{C} \;\middle|\; \sum_{n_1 \in \mathbb{Z}} \sum_{n_2 \in \mathbb{Z}} |x(n_1, n_2)|^p < \infty \right\}.$$

or

$$L^p(\mathbb{R}^2) := \left\{ f \colon \mathbb{R}^2 \to \mathbb{C} \;\middle|\; f \text{ measurable and } \int_{\mathbb{R}} \int_{\mathbb{R}} |\, f(t_1, t_2)|^p dt_1 dt_2 < \infty \right\}$$

for $1 \leq p < \infty$. From this it should be clear how to define the corresponding norms (and the inner products in the case $p = 2$).

Again the spaces $\ell^p(\mathbb{Z}^2)$ and $L^p(\mathbb{R}^2)$ are Banach spaces for $1 \leq p \leq \infty$. In the case $p = 2$ one obtains Hilbert spaces.

In Chapter 3, a system

$$T : I \to O$$

was defined for a general space I of input signals and space O of output signals. For linear spaces I and O the system was called linear if

$$T[x + y] = T[x] + T[y] \quad \text{and} \quad T[\lambda x] = \lambda T[x],$$

for all $x, y \in I$ and all $\lambda \in \mathbb{C}$ (see Definition 3.1). In case I and O are normed spaces, T is called continuous if it maps all convergent sequences of input signals in I to convergent sequences of output signals in O.

Since the spaces $\ell^p(\mathbb{Z}^2)$ are normed vector spaces, all these concepts can be applied for systems

$$T\colon \ell^p(\mathbb{Z}^2) \to \ell^r(\mathbb{Z}^2).$$

As in Definition 3.11 such a system is called time invariant or shift invariant if

$$T \circ \tau_{k_1,k_2} = \tau_{k_1,k_2} \circ T$$

for $k_1, k_2 \in \mathbb{Z}$. Here, τ_{k_1,k_2} denotes the 2D shift operator defined by

$$\tau_{k_1,k_2}[x](n_1, n_2) := x(n_1 - k_1, n_2 - k_2).$$

For short, a linear time invariant system $T\colon \ell^p(\mathbb{Z}^2) \to \ell^r(\mathbb{Z}^2)$ is again simply called an LTI-system.

As for the 1D case, a continuous LTI-systems can be easily described and characterized by the convolution operator. In the 2D case, the convolution $x * y$ of two signals $x, y\colon \mathbb{Z}^2 \to \mathbb{C}$ is defined by

$$(x * y)(n_1, n_2) := \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} x(k_1, k_2) y(n_1 - k_1, n_2 - k_2).$$

The convolution product only exists with additional assumptions on the signals just as in the 1D case (see Theorem 3.13). For example, let p and q be conjugated exponents, then for a fixed $y \in \ell^q(\mathbb{Z}^2)$ the convolution operator

$$C_y\colon \ell^p(\mathbb{Z}^2) \to \ell^\infty(\mathbb{Z}^2)$$

defined by

$$C_y(x) := x * y, \quad x \in \ell^p(\mathbb{Z}^2)$$

defines a linear and continuous operator.

Then, also in the 2D case, a continuous LTI-System $T\colon \ell^p(\mathbb{Z}^2) \to \ell^q(\mathbb{Z}^2)$ is characterized by its impulse response

$$h := T[\delta].$$

In particular, one has $T[x] = h * x$ for all $x \in \ell^p(\mathbb{Z}^2)$, i.e., $T = C_h$. For a stable LTI-system, i.e., $T[\delta] \in \ell^1(\mathbb{Z}^2)$, one obtains the following analog result to Theorem 3.20.

Theorem 10.5. *For a linear system $T\colon \ell^p(\mathbb{Z}^2) \to \ell^p(\mathbb{Z}^2)$ the following is equivalent:*

(1) *T is a stable LTI-system.*
(2) *There is an $h \in \ell^1(\mathbb{Z}^2)$ with $T = C_h$.*

It is straightforward to transfer notions such as FIR filter, IIR filter, causality, etc. from the 1D to the 2D case.

10.2 2D Fourier Transform

The Fourier transform, in both continuous-time and discrete-time cases, can be directly generalized to two dimensions. The only conceptual change is that we now project a 2D signal onto a set of 2D basis functions. The practical result is that the summations and integrals become double summations and double integrals.

10.2.1 2D Fourier Transform for CT-Signals

In the 1D case, a signal can be represented as weighted superposition of the frequency functions

$$t \mapsto e^{2\pi i\omega t}, \quad t \in \mathbb{R},$$

of frequency $\omega \in \mathbb{R}$ where the weights are given by the Fourier coefficients $\hat{f}(\omega)$. In the 2D case, one uses for the analysis of the signal the following 2D frequency functions:

$$(t_1, t_2) \mapsto e^{2\pi i\omega_1 t_1} e^{2\pi i\omega_2 t_2} = e^{2\pi i(\omega_1 t_1 + \omega_2 t_2)}, \quad t_1, t_2 \in \mathbb{R},$$

where $\omega_1 \in \mathbb{R}$ is the frequency in the first and $\omega_2 \in \mathbb{R}$ the frequency in the second dimension.

Theorem 2.5 and Definition 2.7 for 1D signals generalizes as follows:

Theorem 10.6. *For each signal* $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ *holds the equality*

$$f(t_1, t_2) = \int_{\mathbb{R}} \int_{\mathbb{R}} c_{\omega_1,\omega_2} e^{2\pi i(\omega_1 t_1 + \omega_2 t_2)} d\omega_1 d\omega_2 \tag{71}$$

where c_{ω_1,ω_2} *is defined by*

$$c_{\omega_1,\omega_2} = \int_{\mathbb{R}} \int_{\mathbb{R}} f(t_1, t_2) e^{-2\pi i(\omega_1 t_1 + \omega_2 t_2)} dt_1 dt_2. \tag{72}$$

Definition 10.7. Let $f \in L^1(\mathbb{R}^2)$ then the function $\hat{f}\colon \mathbb{R}^2 \to \mathbb{C}$ defined by

$$\hat{f}(\omega_1, \omega_2) := c_{\omega_1,\omega_2} = \int_{\mathbb{R}} \int_{\mathbb{R}} f(t_1, t_2) e^{-2\pi i(\omega_1 t_1 + \omega_2 t_2)} dt_1 dt_2, \quad \omega_1, \omega_2 \in \mathbb{R},$$

is called Fourier integral or Fourier transform of f.

Note that the basis frequency functions are separable. This observation can be a great help in practice. It means that we may write the double integral of the Fourier transform as a cascade of two 1D transforms. To be more precise, let $f \in L^1(\mathbb{R}^2)$. For each fixed $t_2 \in \mathbb{R}$ we define the function

$$f_{t_2}(t_1) := f(t_1, t_2).$$

Then f_{t_2} is a 1D signal which is in $L^1(\mathbb{R})$ (at least for almost all t_2). The Fourier transform $\widehat{f_{t_2}}$ of f_{t_2} is given by

$$\widehat{f_{t_2}}(\omega_1) = \int_{\mathbb{R}} f_{t_2}(t_1) e^{-2\pi i \omega_1 t_1} dt_1.$$

Now, fixing ω_1 the function $t_2 \mapsto \widehat{f_{t_2}}(\omega_1)$ is again a 1D signal which is in $L^1(\mathbb{R})$. The Fourier transform of this function with frequency parameter ω_2 gives

$$\int_{\mathbb{R}} \widehat{f_{t_2}}(\omega_1) e^{-2\pi i \omega_2 t_2} dt_2 = \int_{\mathbb{R}} \int_{\mathbb{R}} f_{t_2}(t_1) e^{-2\pi i \omega_1 t_1} dt_1 e^{-2\pi i \omega_2 t_2} dt_2 = \hat{f}(\omega_1, \omega_2).$$

The last equation has important practical ramifications.

- It tells us that we can take a 2D Fourier transform of a signal by taking two successive 1D passes: first transforming all the "rows" (fixing the second variable t_2) and then all the "columns" (fixing the first variable t_1).

- Interchanging the role of t_1 and t_2, the 2D Fourier transform may also be computed by first transforming all the "columns" and then all the "rows".

- The synthesis formula shares this property.

- Because of this property, the 2D Fourier transform inherits all the properties of the 1D transform, such as scaling, convolution, modulation, and so forth.

Example 10.8. The box function b_W from Example 1.2 has a 2D version denoted by b_{W_1,W_2} and defined by

$$b_{W_1,W_2}(t_1,t_2) := \begin{cases} 1 & \text{if } |t_1| \leq W_1/2 \text{ and } |t_2| \leq W_2/2 \\ 0 & \text{otherwise.} \end{cases}$$

Before we take the Fourier transform we will pause for a moment to consider what to expect. The box function b_{W_1,W_2} is a separable function, built from the product of two 1D boxes. We know that the Fourier transform of a box is a sinc-function (see Example 2.16), and we saw above that a 2D transform may be taken as a sequence of two 1D transforms. Intuitively, we would expect the horizontal pass to spread out the horizontal box into a sinc, and the vertical pass to do the same. Thus we would expect the Fourier transform of the 2D box to be the product of two 1D sinc functions, one along each axis.

To confirm this analysis, we find the transform directly, starting with Definition 10.7:

$$\begin{aligned}
\widehat{b_{W_1,W_2}}(\omega_1,\omega_2) &= \int_{\mathbb{R}}\int_{\mathbb{R}} b_{W_1,W_2}(t_1,t_2)e^{-2\pi i(\omega_1 t_1+\omega_2 t_2)}dt_1dt_2 \\
&= \int_{-W_1/2}^{W_1/2}\int_{-W_2/2}^{W_2/2} e^{-2\pi i\omega_1 t_1}e^{-2\pi i\omega_2 t_2}dt_1dt_2 \\
&= \int_{-W_1/2}^{W_1/2} e^{-2\pi i\omega_1 t_1}dt_1 \int_{-W_2/2}^{W_2/2} e^{-2\pi i\omega_2 t_2}dt_2 \\
&= W_1W_2\ \mathrm{sinc}\,(W_1\omega_1)\mathrm{sinc}\,(W_2\omega_2).
\end{aligned}$$

Figure 45 shows the 2D box function and its Fourier transform. The first line shows a 3D plot of f and $\hat{f}$. In the second line f and $\hat{f}$ are represented by an image, where the value of f at (t_1,t_2) and $\hat{f}$ at (ω_1,ω_2) is indicated by the gray level, where smaller values correspond to dark pixels and larger values to bright pixels.

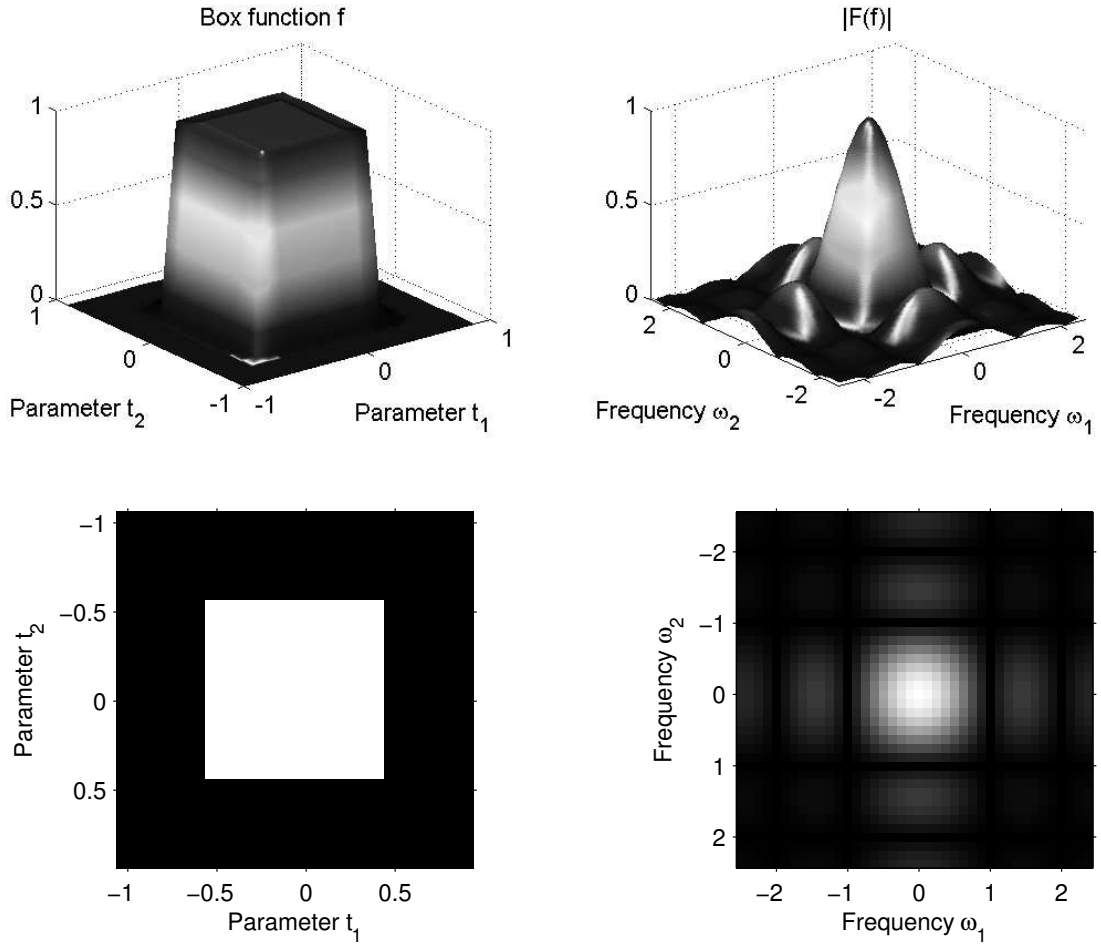


Figure 45: 2D box function and its Fourier spectrum.

10.2.2 2D Fourier Transform for DT-Signals

In the discrete-time case, the generalization of the Fourier transform from 1D to 2D signals is just as in the continuous-time case. We only give a short summary of the main definitions and some properties. The reader should compare the following definition with Definition 2.17.

Definition 10.9. The discrete-time 2D Fourier transform $\hat{x}$ of a DT-signal $x \in \ell^2(\mathbb{Z}^2)$ is defined by

$$\hat{x}(\omega_1, \omega_2) := \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} x(k_1, k_2) e^{-2\pi i (k_1 \omega_1 + k_2 \omega_2)}$$

for $\omega_1, \omega_2 \in [0, 1]$.

Note 10.10. In the 2D case, the Fourier transform $\hat{x}$ is a $(1,1)$-periodic function meaning that it is 1-periodic w.r.t. to both parameters t_1 and t_2:

$$\hat{x}(\omega_1 + n_1, \omega_2 + n_2) = \hat{x}(\omega_1, \omega_2)$$

for all $n_1, n_2 \in \mathbb{Z}$ and all $\omega_1, \omega_2 \in \mathbb{R}$. The space of $(1,1)$-periodic functions square-integrable over $[0,1]^2$ can be identified with the Hilbert space $L^2([0,1]^2)$ just as in the 1D case. One can show that for a DT signal $x \in \ell^2(\mathbb{Z}^2)$, the Fourier transform is square-integrable over $[0,1]^2$, i.e., $\hat{x} \in L^2([0,1]^2)$.

Lemma 10.11. *The signal $x \in \ell^2(\mathbb{Z}^2)$ can be reconstructed from its Fourier transform $\hat{x}$ via*

$$x(k_1, k_2) = \int_{[0,1]} \int_{[0,1]} \hat{x}(\omega_1, \omega_2) e^{-2\pi i(k_1\omega_1 + k_2\omega_2)} d\omega_1 d\omega_2.$$

As for the 1D case, the frequency response of a 2D convolution filter is defined as 2D Fourier transform of the filter coefficients. To be more precise, let $h := T[\delta] \in \ell^1(\mathbb{Z}^2)$ be the impulse response of a stable continuous LTI system T, then the frequency response H of T is defined by

$$H(\omega_1, \omega_2) := \hat{h}(\omega_1, \omega_2) = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} h(k_1, k_2) e^{-2\pi i (k_1 \omega_1 + k_2 \omega_2)}.$$

As in the 1D case, discrete filters and DT-signals are often denoted by small letters $f, g, h, \ldots, x, y \ldots$ and the corresponding Fourier transforms are then denoted by the corresponding capital letters $F, G, H, \ldots, X, Y \ldots$.

- It can be shown that if h is separable, then H is separable as well, i.e., it can be expressed in the form
$$H(\omega_1, \omega_2) = H_1(\omega_1) H_2(\omega_2).$$
- The filter coefficients $h(k_1, k_2)$ can be recovered from the frequency response as indicated in Lemma 10.11.
- By applying the Fourier transform on the convolution $y = T[x] = h * x$, it follows that the Fourier transforms of the input signal x and the output signal y of a two-dimensional LTI system are related by
$$Y(\omega_1, \omega_2) = H(\omega_1, \omega_2) X(\omega_1, \omega_2),$$
for $\omega_1, \omega_2 \in \mathbb{R}$.

It is the last property which allows to characterize some LTI system T by means of its frequency response H: depending on the properties of the magnitude response $|H|$ one distinguishes, as for the 1D case, between lowpass, highpass, and bandpass filters.

Example 10.12. As an example, we consider the 2D averaging filter. The reader should compare the 2D case with the 1D case discussed in Subsection 3.6.2. The 2D averaging filters h_{N_1,N_2} of length (N_1, N_2) is defined by

$$h_{N_1,N_2}(n_1, n_2) = \begin{cases} \frac{1}{N_1 N_2} & \text{if } 0 \leq n_1 \leq N_1 - 1 \text{ and } 0 \leq n_2 \leq N_2 - 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Clearly, h_{N_1,N_2} is a separable filter:

$$h_{N_1,N_2}(n_1, n_2) = h_{N_1}(n_1) h_{N_2}(n_2),$$

where h_{N_1} denotes the 1D averaging filter of length N_1 and h_{N_2} denotes the 1D averaging filter of length N_2. Therefore, the frequency response H_{N_1,N_2} of h_{N_1,N_2} is given by

$$\begin{aligned} H_{N_1,N_2}(\omega_1, \omega_2) &= H_{N_1}(\omega_1) H_{N_2}(\omega_2) \\ &= \frac{1}{N_1 N_2} \frac{\sin(\pi\omega_1 N_1)}{\sin(\pi\omega_1)} \frac{\sin(\pi\omega_2 N_2)}{\sin(\pi\omega_2)} e^{-\pi i((N_1-1)\omega_1 + (N_2-1)\omega_2)}. \end{aligned}$$

Figure 46 and Figure 47 show averaging filters h_{N_1,N_2} and their corresponding frequency responses H_{N_1,N_2} for various N_1 and N_2.

The left hand side of Figure 48 shows a discrete-time 2D signal f which we already encountered in Figure 3. f is a 1-sampled version of an image of size 256×256. The Fourier transform of f was approximated by using a DFT of size 256 is described in Section 2.4. Here, in each dimension, the upper 128 coefficients of the output vector are interpreted as Fourier coefficients of the corresponding negative frequencies. The result is shown on the right hand side of Figure 48, where a logarithmic and normalized scale is used for the gray levels of the pixels.

Finally, Figure 49 shows the result in time and frequency domain when filtering the signal f by various averaging filters $h_{N,N}$.

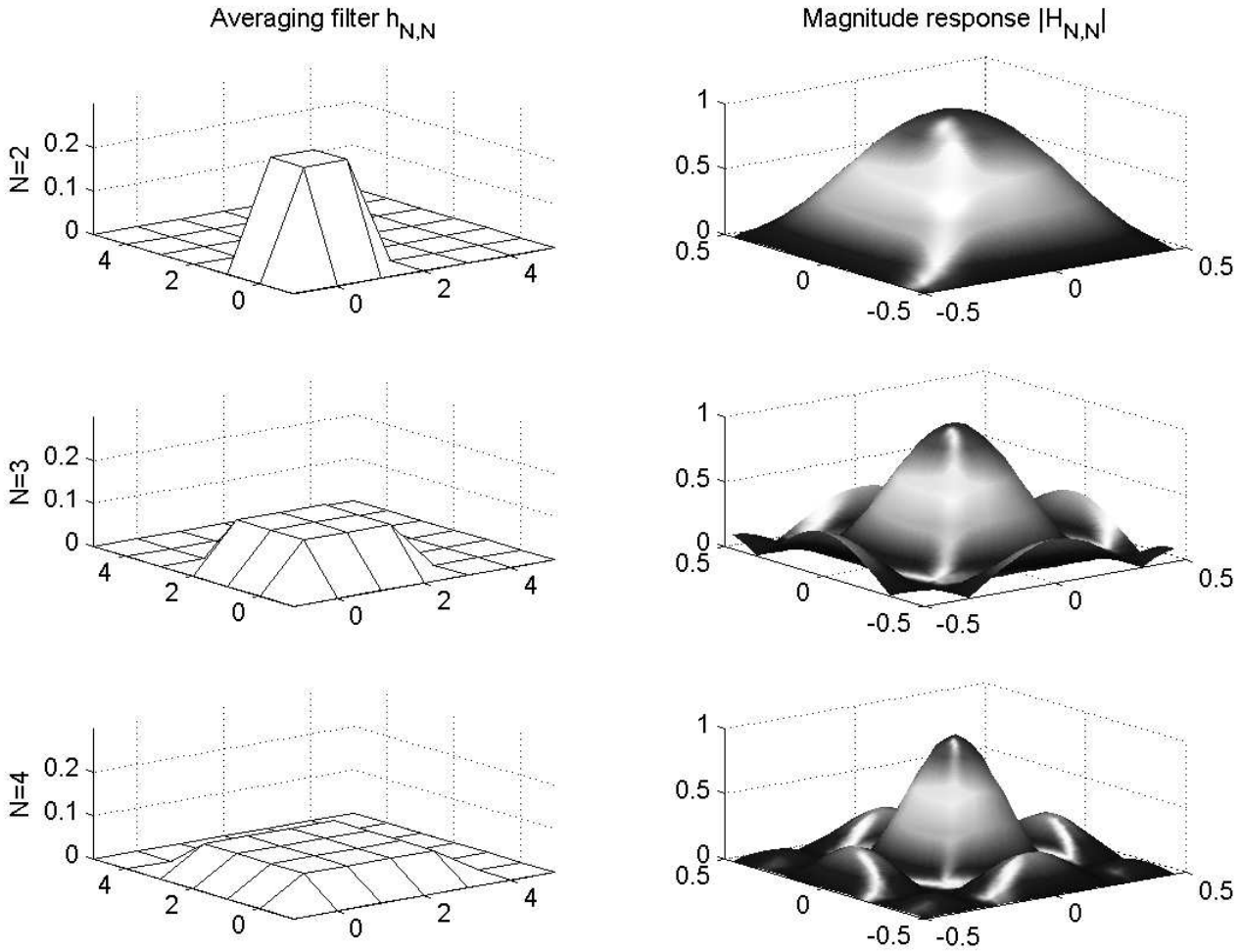


Figure 46: Averaging filter $h_{N,N}$ and magnitude response $|H_{N,N}|$ for $N = 2, 3, 4$.

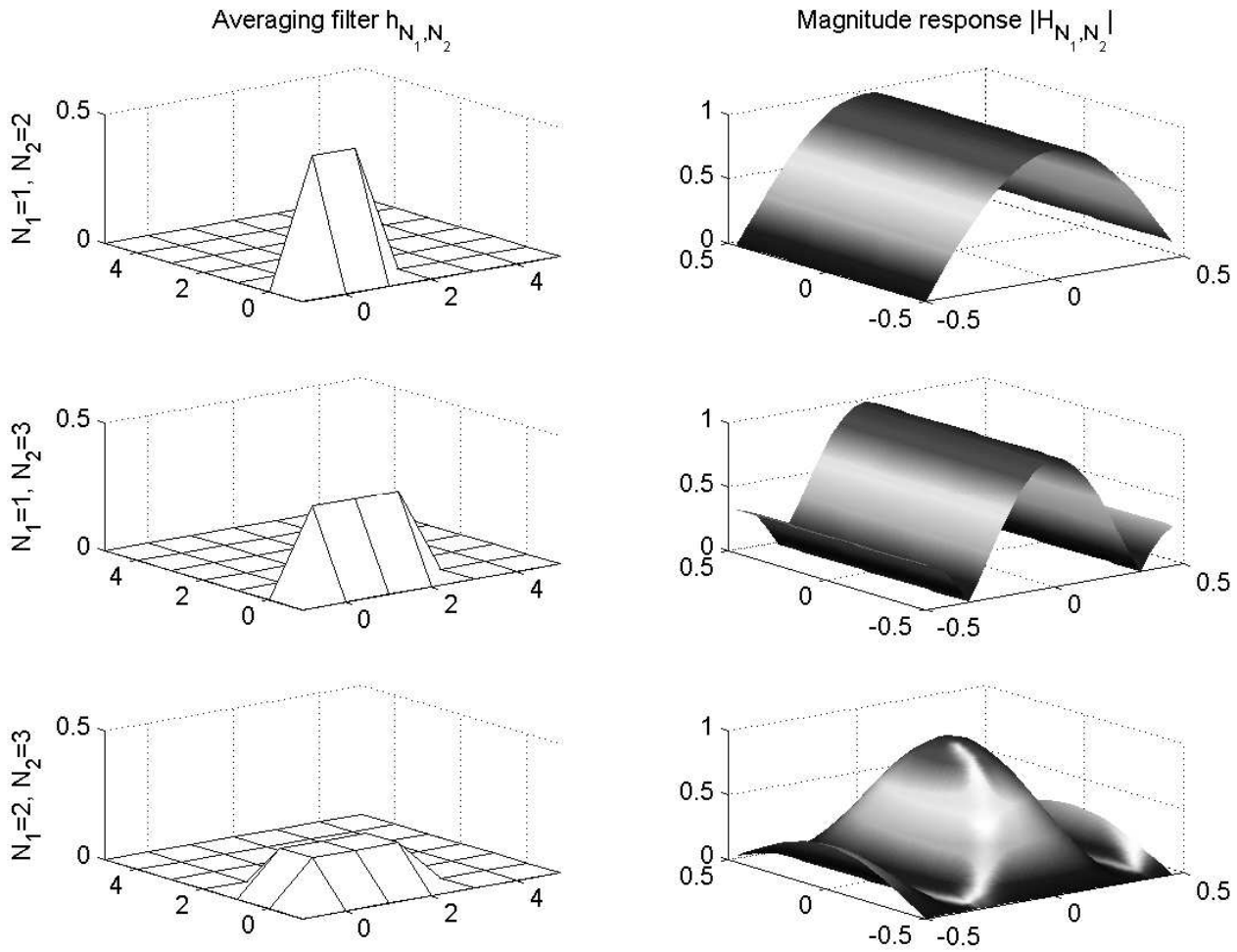


Figure 47: Averaging filter h_{N_1,N_2} and magnitude response $|H_{N_1,N_2}|$.

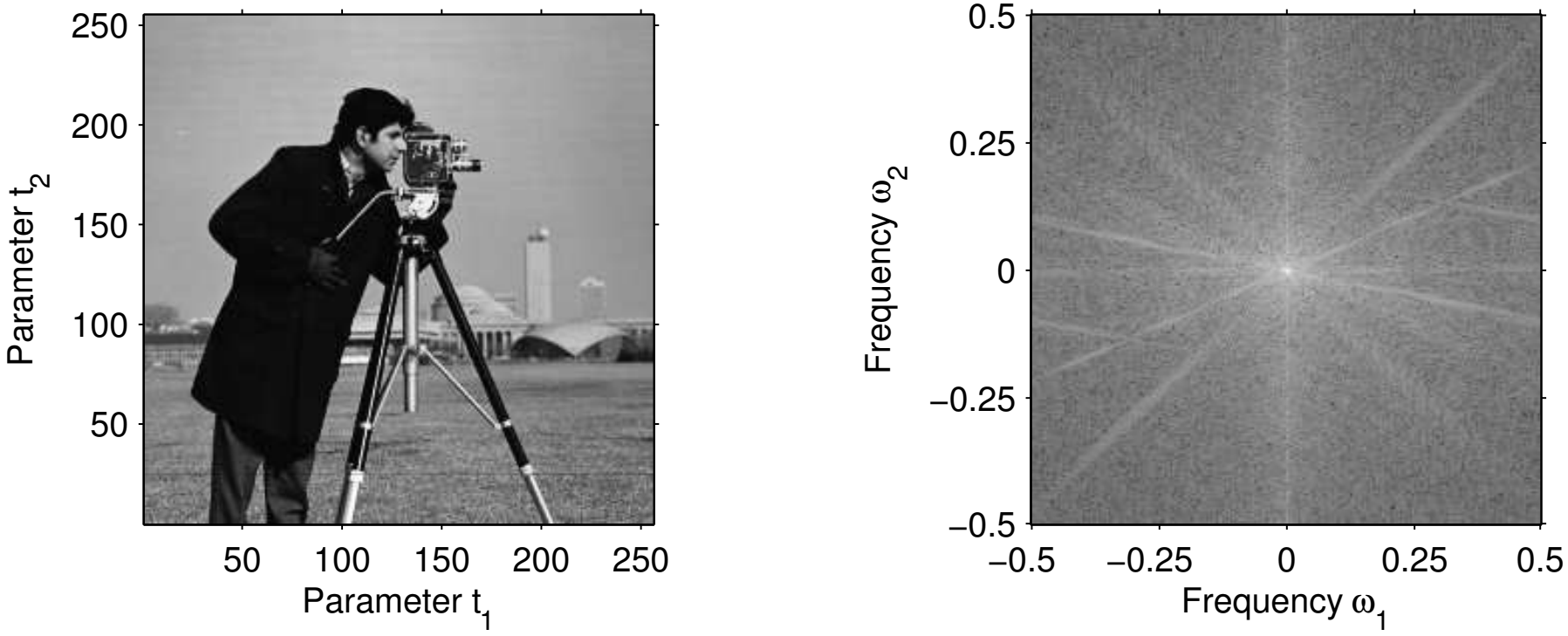


Figure 48: Discrete-time 2D signal f and its approximated Fourier spectrum in decibel computed by the DFT.

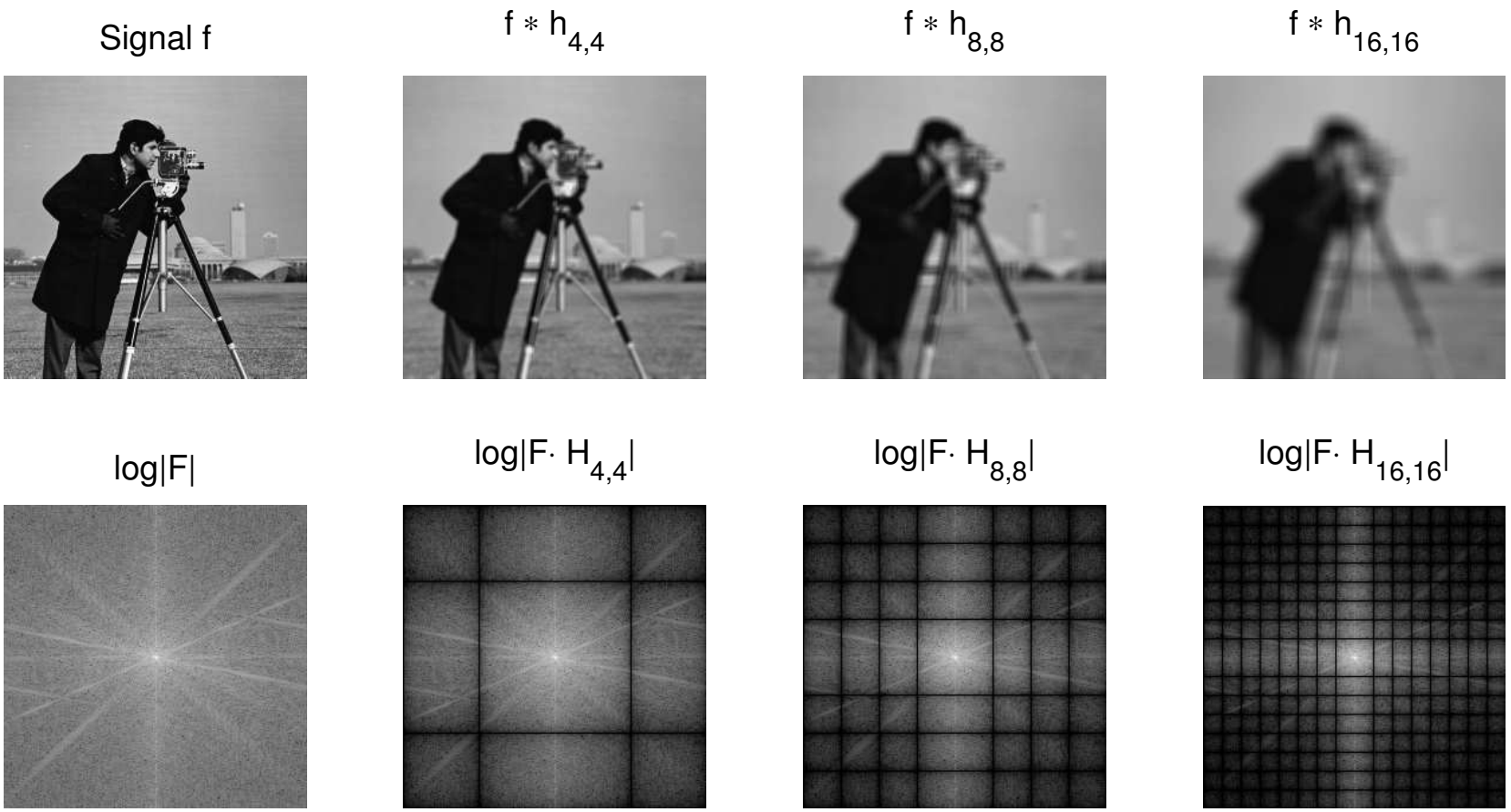


Figure 49: Filtered 2D signal f with various averaging filters $h_{N,N}$ in time and spectral domain.

10.2.3 2D z-Transform

The z-transform for 1D signals as introduced in Section 3.4 has a straightforward generalization to the 2D case. For a 2D signal $x : \mathbb{Z}^2 \to \mathbb{C}$, the z-transform X is defined by

$$X(z_1, z_2) = \sum_{n_1 \in \mathbb{Z}} \sum_{n_2 \in \mathbb{Z}} x(n_1, n_2) z_1^{-n_1} z_2^{-n_2},$$

where z_1 and z_2 are complex variables. As for the 1D case, we use the same symbol X to denote the z-transform as well as the Fourier transform. With z_1 and z_2 expressed in polar coordinates

$$z_1 = r_1 e^{2\pi i \omega_1} \quad \text{and} \quad z_2 = r_2 e^{2\pi i \omega_2}$$

the z-transform can be written as

$$X(r_1 e^{2\pi i \omega_1}, r_2 e^{2\pi i \omega_2}) = \sum_{n_1 \in \mathbb{Z}} \sum_{n_2 \in \mathbb{Z}} x(n_1, n_2) r_1^{-n_1} r_2^{-n_2} e^{2\pi i \omega_1 n_1} e^{2\pi i \omega_2 n_2}.$$

Thus, as in the 1D case, the 2D z-transform can be interpreted as the 2D Fourier transform of x multiplied by the 2D exponential sequence

$$(r_1^{-n_1} r_2^{-n_2})_{n_1 \in \mathbb{Z}, n_2 \in \mathbb{Z}}.$$

For $|z_1| = |z_2| = 1$, i.e., for $r_1 = r_2 = 1$, the z-transform is equal to the Fourier transform. For convergence of the 2D z-transform we require that the sequence $(x(n_1, n_2) z_1^{-n_1} z_2^{-n_2})_{n_1 \in \mathbb{Z}, n_2 \in \mathbb{Z}}$ be absolutely summable, i.e., that

$$\sum_{n_1 \in \mathbb{Z}} \sum_{n_2 \in \mathbb{Z}} |x(n_1, n_2) z_1^{-n_1} z_2^{-n_2}| < \infty.$$

The set of values z_1 and z_2 for which the 2D z-transform converges defines the region of convergence.

A 2D z-transform X is said to be separable if it can be expressed in the form

$$X(z_1, z_2) = X_1(z_1)X_2(z_2).$$

It is easy to see that X is separable only if x is separable, i.e., $x(n_1, n_2) = x_1(n_1)x_2(n_2)$ for suitable 1D signals x_1 and x_2. In that case X_1 and X_2 are the 1D z-transforms of x_1 and x_2, respectively.

The 2D z-transform of a convolution of two 2D signals is the product of their z-transforms:

$$y = h * x \quad \Rightarrow \quad Y(z) = H(z)X(z).$$

As in one dimension, the z-transform of the 2D impulse response is referred to as the transfer function of the system.

10.3 Sampling and Aliasing in 2D

We now turn our attention to sampling in two dimensions. Let $f : \mathbb{R}^2 \to \mathbb{C}$ be a continuous CT-signal. We only consider the case of sampling on a rectangular grid with periods (T_1, T_2), $T_1 > 0$, $T_1 > 0$. In this case one obtains a DT-signal $x\colon \mathbb{Z}^2 \to \mathbb{C}$ via

$$x(n_1, n_2) := f(n_1 \cdot T_1, n_2 \cdot T_2), \quad n_1, n_2 \in \mathbb{Z}.$$

One then also speaks from (T_1, T_2)-sampling and sampling rate $(1/T_1, 1/T_2)$ measured in each of the two dimensions separately.

Similar to the 1D case, the spectrum $\hat{x}$ of x is the $(1,1)$-periodized spectrum of a scaled version of the spectrum $\hat{f}$. We summarize the result in the following theorem (compare to Theorem 4.9).

Theorem 10.13. *Let f be a continuous function with $f \in L^2(\mathbb{R}^2)$ and x be the DT-signal obtained from f by (T_1, T_2)-sampling. Then*

$$\hat{x}(\omega_1, \omega_2) = \frac{1}{T_1 T_2} \sum_{k_1 \in \mathbb{Z}} \sum_{k_1 \in \mathbb{Z}} \hat{f}\left(\frac{\omega_1 + k_1}{T_1}, \frac{\omega_1 + k_1}{T_1}\right) \quad \text{for} \quad \omega_1, \omega_2 \in \mathbb{R}.$$

From this it is not difficult to see how the sampling theorem by Shannon extends to the 2D case. First, we need the 2D version of Definition 4.2.

Definition 10.14. Let $\Omega_1, \Omega_2 > 0$. The CT-signal $f \in L^2(\mathbb{R}^2)$ is called (Ω_1, Ω_2)-bandlimited if the Fourier transform $\hat{f}$ vanishes for $|\omega_1| > \Omega_1$ and $|\omega_2| > \Omega_2$:

$$\hat{f}(\omega_1, \omega_2) = 0 \quad \forall |\omega_1| > \Omega_1, |\omega_2| > \Omega_2, \quad \text{i.e., } \operatorname{supp} \hat{f} \subset [-\Omega_1, \Omega_1] \times [-\Omega_2, \Omega_2].$$

Then the 2D version of Shannon's theorem is as follows (compare with Theorem 4.6):

Theorem 10.15. *Let $f \in L^2(\mathbb{R}^2)$ be an (Ω_1, Ω_2)-bandlimited function and let x be the (T_1, T_2)-sampled version of f with $T_1 := \frac{1}{2\Omega_1}$ and $T_2 := \frac{1}{2\Omega_2}$, i.e.,*

$$x(n_1, n_2) := f(n_1 \cdot T_1, n_2 \cdot T_2), \quad n_1, n_2 \in \mathbb{Z}.$$

Then f can be reconstructed from x as follows:

$$f(t_1, t_2) = \sum_{n_1 \in \mathbb{Z}} \sum_{n_2 \in \mathbb{Z}} x(n_1, n_2) \operatorname{sinc}\left(\frac{t_1 - n_1 T_1}{T_1}\right) \operatorname{sinc}\left(\frac{t_2 - n_2 T_2}{T_2}\right).$$

Note 10.16. Theorem 4.3 and Note 4.4 also apply for the 2D case.

10.4 2D MRA and Wavelets

In this section, we indicate how to generalize the 1D multiresolution analysis presented in Chapter 8 following in most part Section 2.2.2 of [Louis/Maaß/Rieder]. We also refer for further details to this book. The properties of 2D MRAs are the basis for usage of wavelets in applications such as image analysis and image compression.

A multiresolution analysis of $L^2(\mathbb{R}^2)$ consists — just as for the 1D case — of an increasing sequence of closed subspaces $V_m \subset L^2(\mathbb{R}^2)$, $m \in \mathbb{Z}$, such that

$$\overline{\bigcup_{m\in\mathbb{Z}} V_m} = L^2(\mathbb{R}^2) \quad \text{and} \quad \bigcap_{m\in\mathbb{Z}} V_m = \{0\}.$$

In contrast to the 1D MRA, we now admit translates $k \in \mathbb{Z}^2$ and the transition from V_m to V_{m-1} is described by a regular matrix A, also denoted as dilation matrix:

$$f(\cdot) \in V_m \Longleftrightarrow f(A\cdot) \in V_{m-1}.$$

The base space V_0 is again generated by an ONB consisting of all integer translates of a scaling function $\varphi \in V_0$:

$$V_0 = \overline{\mathrm{span}\{\varphi(\cdot - k)|k \in \mathbb{Z}^2\}}.$$

Then it follows that

$$\{\varphi_{m,k} := |\det A|^{-m/2}\varphi(A^m \cdot -k)|k \in \mathbb{Z}^2\}$$

is an ONB of V_m. According to the philosophy of the MRA decomposition the dilation matrix should stretch in all directions. In other words, the absolute values of the eigenvalues of A are greater then one. Furthermore, A should have integer entries. This is equivalent to

$$A\mathbb{Z}^2 \subset \mathbb{Z}^2.$$

2D wavelets are those functions which span the orthogonal complement of V_0 in V_{-1}. This is, in general, not any longer possible for one single function with all its integer translates. The following heuristic note should convince us of this fact and should motivate the following theorem of Y. Meyer.

Note 10.17. We chose as scaling function the characteristic function of some set Ω. The basis functions of V_{-1} is then the characteristic function of the set $A^{-1}\Omega$ with measure

$$\mu(A^{-1}\Omega) = \mu(\Omega)/|\det A|.$$

A basis of V_{-1} needs therefore "$|\det A|$-times" the number of elements as a basis of V_0. The orthogonal complement of V_0 is in this sense "$(|\det A| - 1)$-times" larger then V_0.

Theorem 10.18. [Meyer.] *Let $(V_m)_{m\in\mathbb{Z}}$ be an MRA with dilation matrix A. Then there exist $|\det A| - 1$ wavelets*

$$\psi_1, \psi_2, \ldots, \psi_{|\det A|-1} \in V_{-1}$$

which generate an ONB of the orthogonal complement of V_0 in V_{-1}, i.e.,

$$\left\{ \psi_{j,m,k} = |\det A|^{-m/2} \psi_j(A^{-m} \cdot -k) \middle| j = 1, 2, \ldots |\det A| - 1, m \in \mathbb{Z}, k \in \mathbb{Z}^2 \right\}$$

is an ONB of $L^2(\mathbb{R}^2)$.

By this theorem we obtain an orthogonal decomposition of V_{-1} in $|\det A|$ subspaces

$$V_{-1} = \bigoplus_{j=1}^{|\det A|-1} W_{0,j} \oplus V_0,$$

where the spaces $W_{0,j}$ are given by

$$W_{0,j} = \overline{\operatorname{span}\{\psi_j(\cdot - k) | k \in \mathbb{Z}^2\}}.$$

Note 10.19. In case of the definition of 1D MRA one may replace the dilation parameter 2 (see condition (35) in Definition 8.1) by some arbitrary integer parameter $p \geq 2$. Then, however, one also needs $p-1$ one-dimensional wavelets to generate W_0.

As example, we consider the so-called tensor product wavelets. In general, the tensor product $f \otimes g : \mathbb{R}^2 \longrightarrow \mathbb{C}$ of two functions $f, g : \mathbb{R} \longrightarrow \mathbb{C}$ is defined by

$$f \otimes g(t_1, t_2) := f(t_1)g(t_2).$$

We start with a 1D orthogonal wavelet ψ and some orthogonal scaling function φ coming from an 1D MRA. Note that the set of functions

$$\{\psi_{m,k} \otimes \psi_{\mu,\kappa} \mid m, \mu, k, \kappa \in \mathbb{Z}\}$$

defines an ONB of $L^2(\mathbb{R}^2)$. However, it does not constitute a wavelet basis, since for $m \neq \mu$ the basis functions of different scales are mixed up.

One has to chose a different approach. The tensor product $\varphi \otimes \varphi$ of the scaling function φ generates a 2D scaling function for a 2D MRA w.r.t. the dilation matrix

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

By Theorem 10.18 there are $|\det A| - 1 = 3$ wavelets whose integer translates form a basis of W_0. One can show that these wavelets are given by

$$\varphi \otimes \psi, \ \psi \otimes \varphi, \ \text{and} \ \psi \otimes \psi.$$

In this basis, there is no mixing up of different scales. A decomposition w.r.t. this basis has a typical multiresolution property.

- The 2D wavelets and the scaling function defined by tensor products are obviously separable functions. Therefore, their properties can easily by deduced from the 1D case.
- However, there is also a disadvantage concerning these wavelets: one needs three different wavelets to generate W_0.
- The wavelets constructed as above have the property that certain directions (the t_1- and t_2-direction as well as the diagonal) are "privileged". Such wavelets are also called anisotrop. This property can be an advantage when used for edge detection in images. However, tensor product wavelets are not suitable in view of applications in data compression. This application demands isotropy and a number of wavelets as small as possible.

In view of the last application, one may study 2D MRAs having a dilation matrix A with $|\det A| = 2$. For further details on this topic we have to refer to the literature such as [Louis/Maaß/Rieder].

Literature

[Clausen/Müller] Michael Clausen, Meinard Müller: Zeit-Frequenz-Analyse und Wavelettransformation. Script, summer term 2001, University of Bonn.

[Blatter] Christian Blatter: *Wavelets - Eine Einführung.* Vieweg 1998

[Burrus/Gopinath/Guo] C. Sidney Burrus, Ramesh A. Gopinath, Haitao Guo: *Wavelets and Wavelet Transforms.* Prentice Hall, New Jersey 1998

[Clausen/Baum] Michael Clausen, Ulrich Baum: *Fast Fourier Transforms.* BI Wissenschaftsverlag 1993

[Daubechies] Ingrid Daubechies: *Ten lectures on wavelets.* CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM 1992

[Donoho] David L. Donoho: *Denoising by Soft-Thresholding.* IEEE Transaction on Information Theory, Vol. 41, Nr. 3, S. 613 - 627, Mai 1995

[Folland] Gerald B. Folland: *Real Analysis.* John Wiley & Sons, 1984

[Glassner] Andrew S. Glassner: *Principles of Digital Image Synthesis.* Morgan Kaufmann Publishers, Inc. 1999

[Hubbard] Barbara Burke Hubbard: *The world according to wavelets.* AK Peters, Wellesley, Massachusetts, 1996

[Kaiser] Gerald Kaiser: *A friendly guide to wavelets.* Birkhäuser 1994

[Krengel] Ulrich Krengel: *Einführung in die Wahrscheinlichkeitstheorie und Statistik.* Vieweg, 1991

[Lang] Serge Lang: *Algebra.* Addison-Wesley Publishing Company, 1993

[Louis/Maaß/Rieder] Alfred K. Louis, Peter Maaß, Andreas Rieder: *Waveltes.* B. G. Teubner, Stuttgart 1994

[Matlab] M. Misiti, Y. Misiti, G. Oppenheim J.-M. Poggi: *Wavelet Toolbox for Use with* MATLAB. The MathWorks, Inc. 1996

[Pohlman] Ken C. Pohlmann: *Principles of Digital Audio.* McGraw-Hill, Inc. 1995

[Proakis/Manolakis] John G. Proakis, Dimitris G. Manolakis: *Digital Signal Processing.* Prentice Hall, 1996

[Oppenheim/Schafer] Alan V. Oppenheim, Ronald W. Schafer: *Digital Signal Processing.* VCH, 1975

[Oppenheim/Willsky] Alan V. Oppenheim, Alan S. Willsky: *Signale und Systeme.* VCH, 1992

[Remmert] Remmert: *Funktionentheorie 1.* 3. Auflage, Springer-Verlag

[Strang/Nguyen] Gilbert Strang, Truong Nguyen: *Wavelets and Filter Banks.* Wellesley-Cambridge Press 1996

[Wickerhauser] Mladen V. Wickerhauser: *Adaptive Wavelet-Analysis.* Vieweg 1993