Lecture

DIGITAL PROCESSING

OF

SPEECH AND IMAGE SIGNALS

RWTH Aachen, WS 2006/7

Prof. Dr.-Ing. H. Ney, Dr.rer.nat. R. Schlüter
Lehrstuhl für Informatik 6
RWTH Aachen

1. System Theory and Fourier Transform
2. Discrete Time Systems
3. Spectral Analysis
4. Fourier Transform and Image Processing
5. LPC Analysis
6. Wavelets
7. Coding
8. Image Segmentation and Contour-Finding

Completions: L. Welling, A. Eiden; April 1997
Completions: J. Dahmen, F. Hilger, S. Koepke; Mai 2000
Completions: F. Hilger, D. Keysers; Juli 2001
Translation: M. Popović, R. Schlüter; April 2003
Corrections: D. Stein; October 2006

Literature:

- A. V. Oppenheim, R. W. Schafer: Discrete Time Signal Processing, Prentice Hall, Englewood Cliffs, NJ, 1989.
- A. Papoulis: Signal Analysis, McGraw-Hill, New York, NY, 1977.
- A. Papoulis: The Fourier Integral and its Applications, McGraw-Hill Classic Textbook Reissue Series, McGraw-Hill, New York, NY, 1987.
- W. K. Pratt: Digital Image Processing, Wiley & Sons Inc, New York, NY, 1991.

Further reading:

- T. K. Moon, W. C. Stirling: *Mathematical Methods and Algorithms for Signal Processing.* Prentice Hall, Upper Saddle River, NJ, 2000.
- J. R. Deller, J. G. Proakis, J. H. L. Hansen: Discrete-Time Processing of Speech Signals, Macmillan Publishing Company, New York, NY, 1993.
- W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery: Numerical Recipes in C, Cambridge Univ. Press, Cambridge, 1992.
- L. Rabiner, B. H. Juang: Fundamentals of Speech Recognition, Prentice Hall, Englewood Cliffs, NJ, 1993.
- T. Lehmann, W. Oberschelp, E. Pelikan, R. Repges: Bildverarbeitung für die Medizin, Springer Verlag, Berlin, 1997.
- L. Berg: Lineare Gleichungssysteme mit Bandstruktur, VEB Deutscher Verlag der Wissenschaften, Berlin, 1986.

Contents

List of Figures

List of Tables

Chapter 1

System Theory and Fourier Transform

- Overview:

1.1 Introduction

What distinguishes the Fourier Transform (FT) from
other transformations?

1. Mathematical property of linear time-invariant systems:
   - FT decomposes the time signal into "eigenfunctions"
     "eigenfunctions" keep their form by passing the
     linear time-invariant system

     $$A\,x = \lambda\,x$$

   - Magnitude of FT: shift invariant
2. Physical observation:
   - Human ear produces sort of FT, essentially only magnitude of FT
     (strictly speaking: short-time FT)
   - Example:
     Time functions with different evolution can sound equally.
     The human ear either senses sense phase differences of partial tones of the complete sound of stationary processes very weakly, or does not sense them at all.

Fourier transform in speech processing:

- Calculation of the spectral components of speech
- Basic method for obtaining observations (features) for
  speech recognition

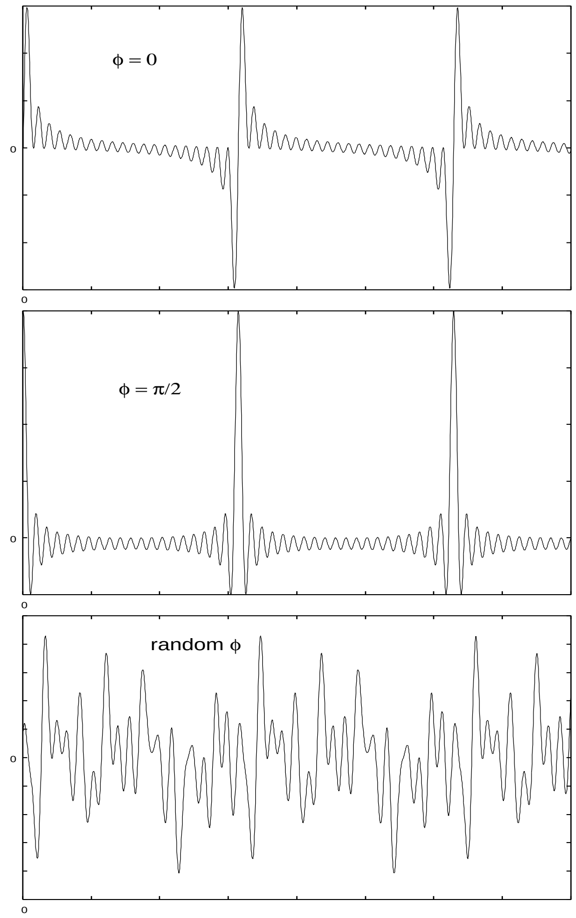


Figure 1.1: Oscillograms of three time functions composed as sum of 20 partial oscillations. a) $\phi_n = 0$, b) $\phi_n = \frac{\pi}{2}$, c) ϕ_n statistical.

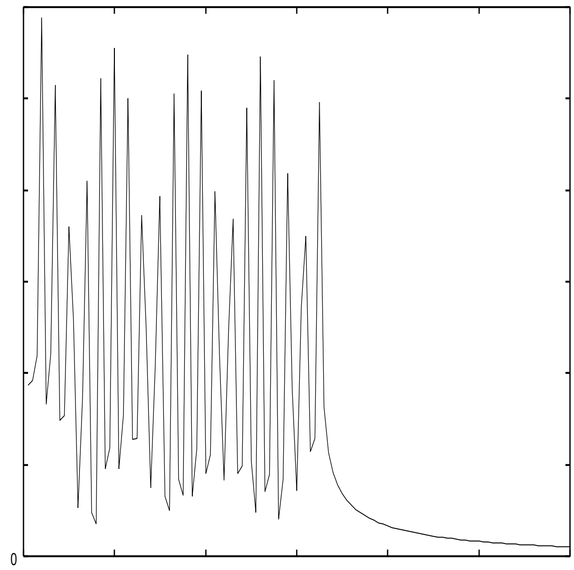


Figure 1.2: Amplitude spectrum of a time function composed as sum of 20 partial tones.

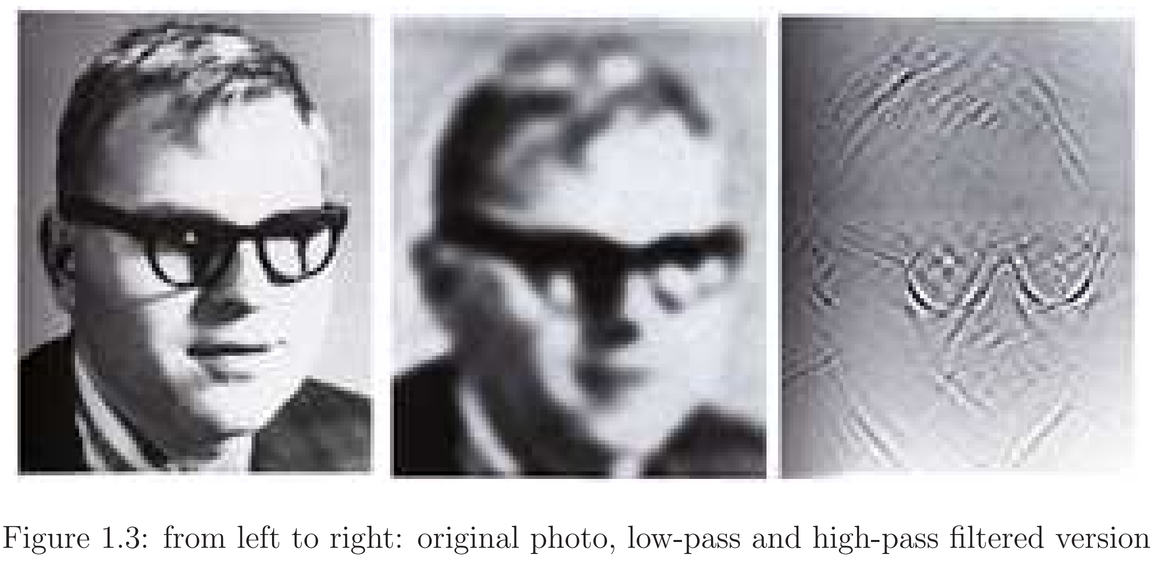


Figure 1.3: from left to right: original photo, low-pass and high-pass filtered version

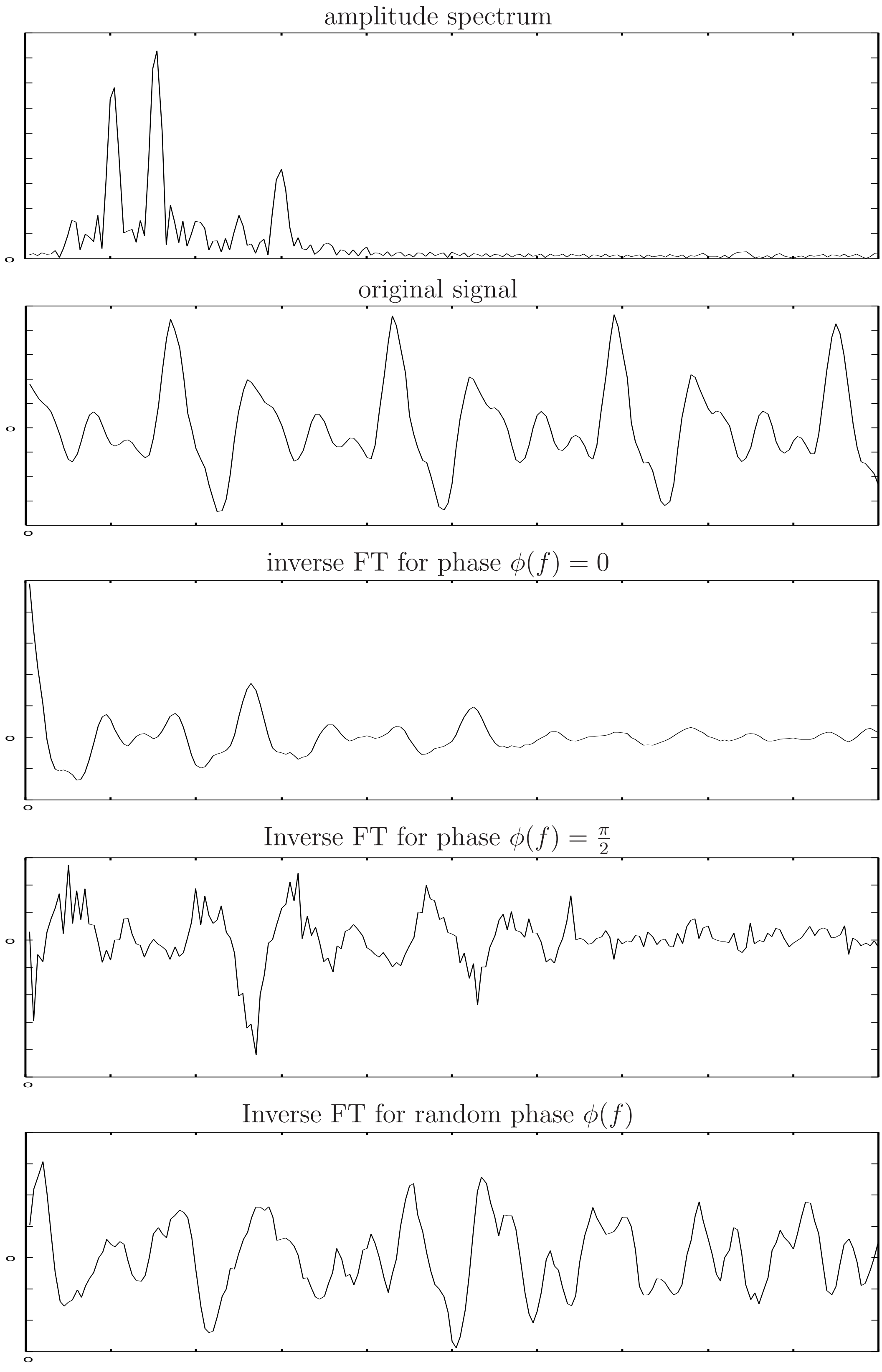


Figure 1.4: Phase manipulation for portion of a speech signal (vowel 'o') sampled at 8kHz, 25ms analysis window (200 samples), 512 point FFT

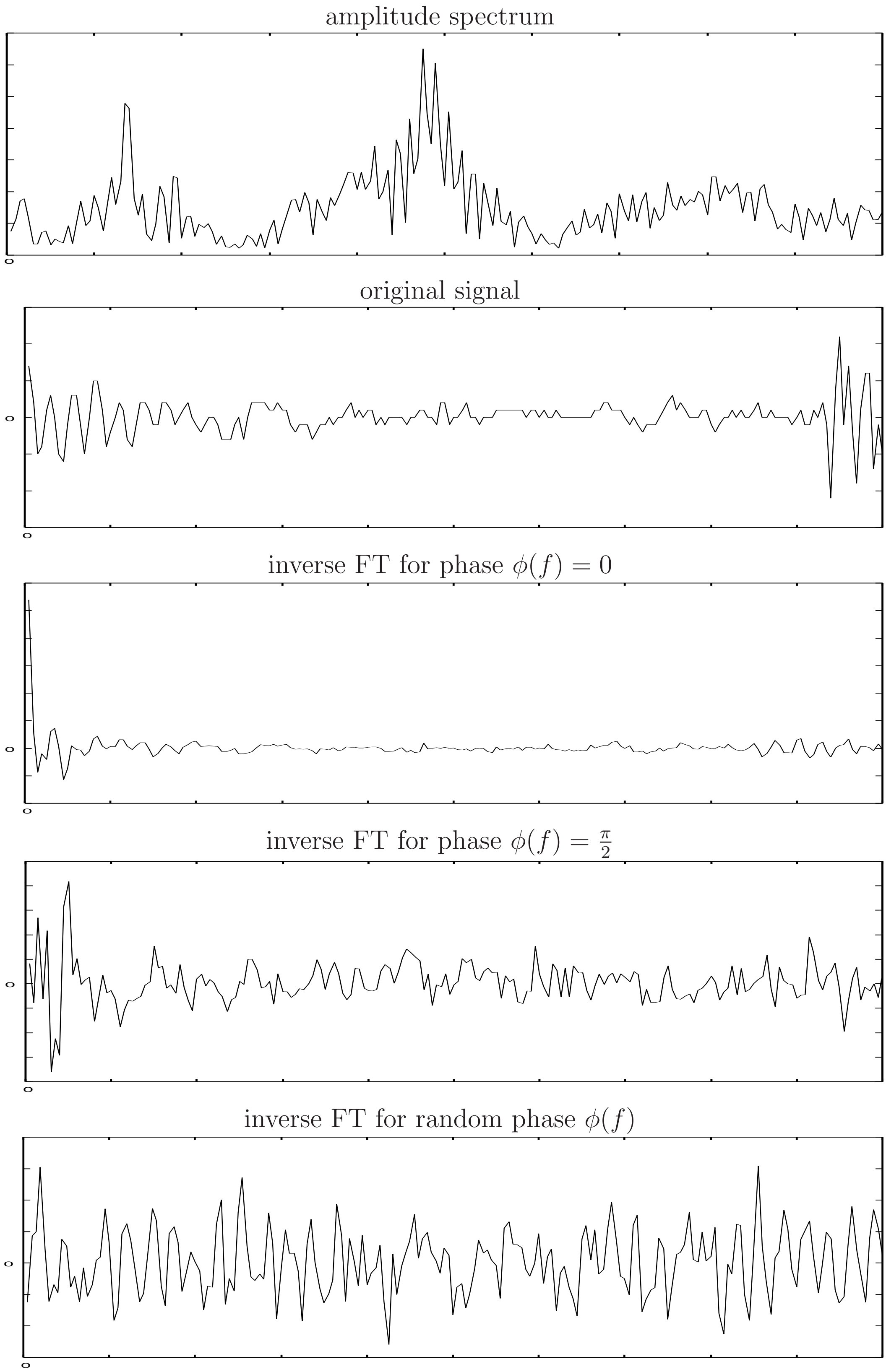


Figure 1.5: Phase manipulation for portion of a speech signal (consonant 'n') sampled at 8kHz, 25ms analysis window (200 samples), 512 point FFT

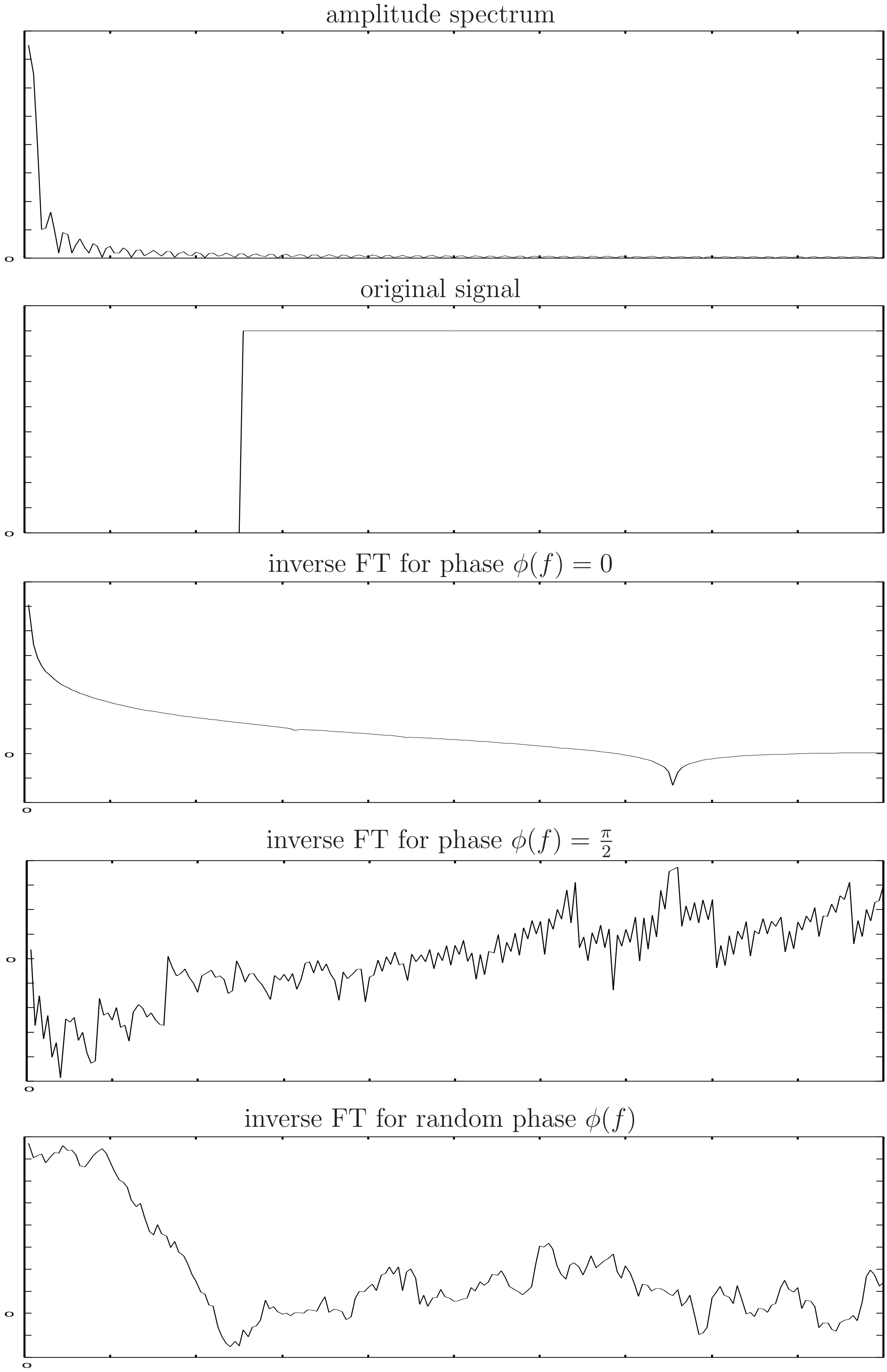


Figure 1.6: Phase manipulation for a Heaviside–function (step–function)

Why Fourier?

Roughly:

- Production, description and algorithmic operations on signals (functions or measurement curves over the time axis) can be described very well in Fourier domain (frequency domain).

Deeper reason:

- Production, description and algorithmic operations on signals are largely based on linear time-invariant (LTI) operations.
- Fourier Transform: simple representation of LTI-operations (later: convolution theorem)

Why continuous?

- Real world is continuous
- Computer (digital = time discrete = sampled)

  model of the real world

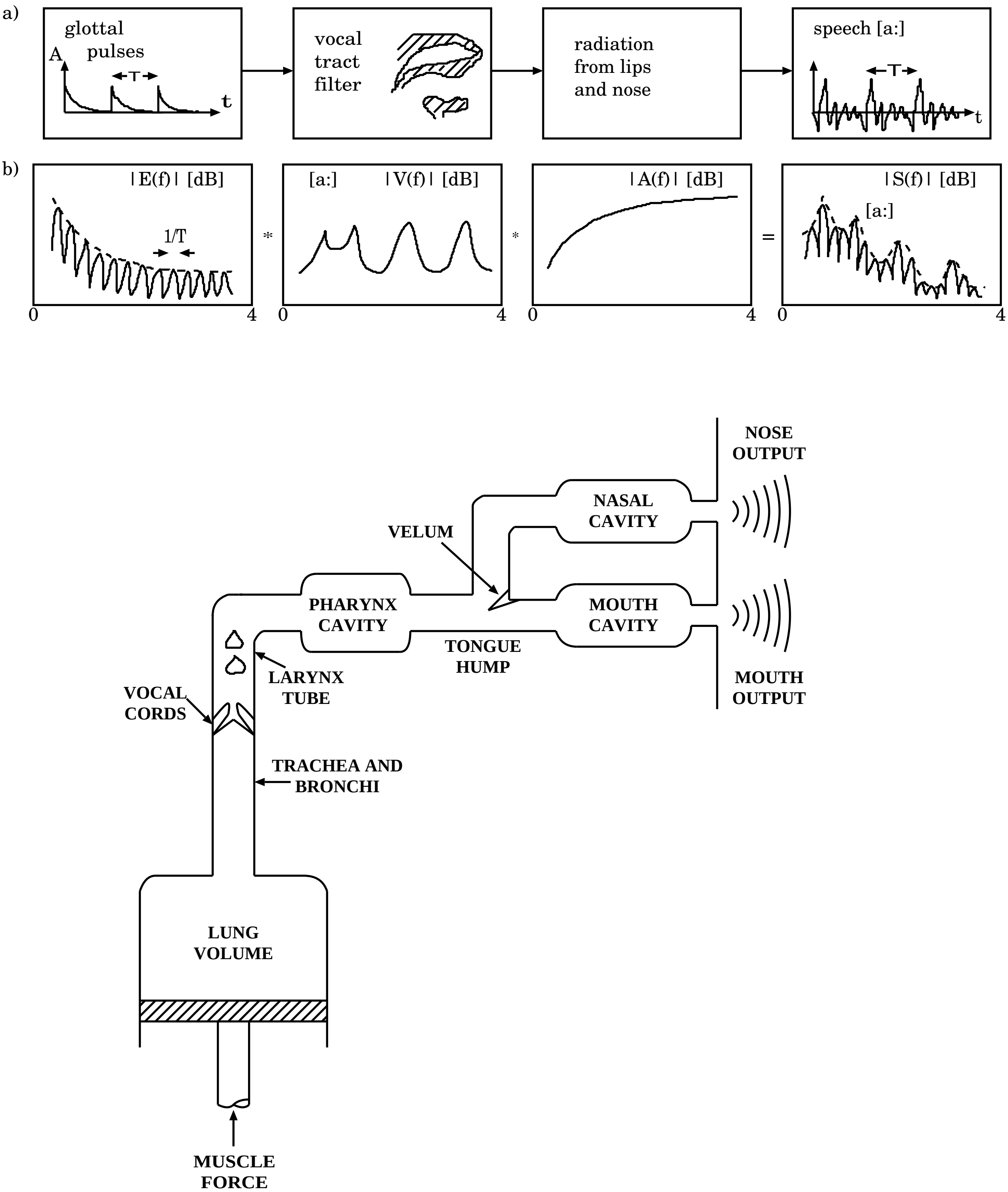
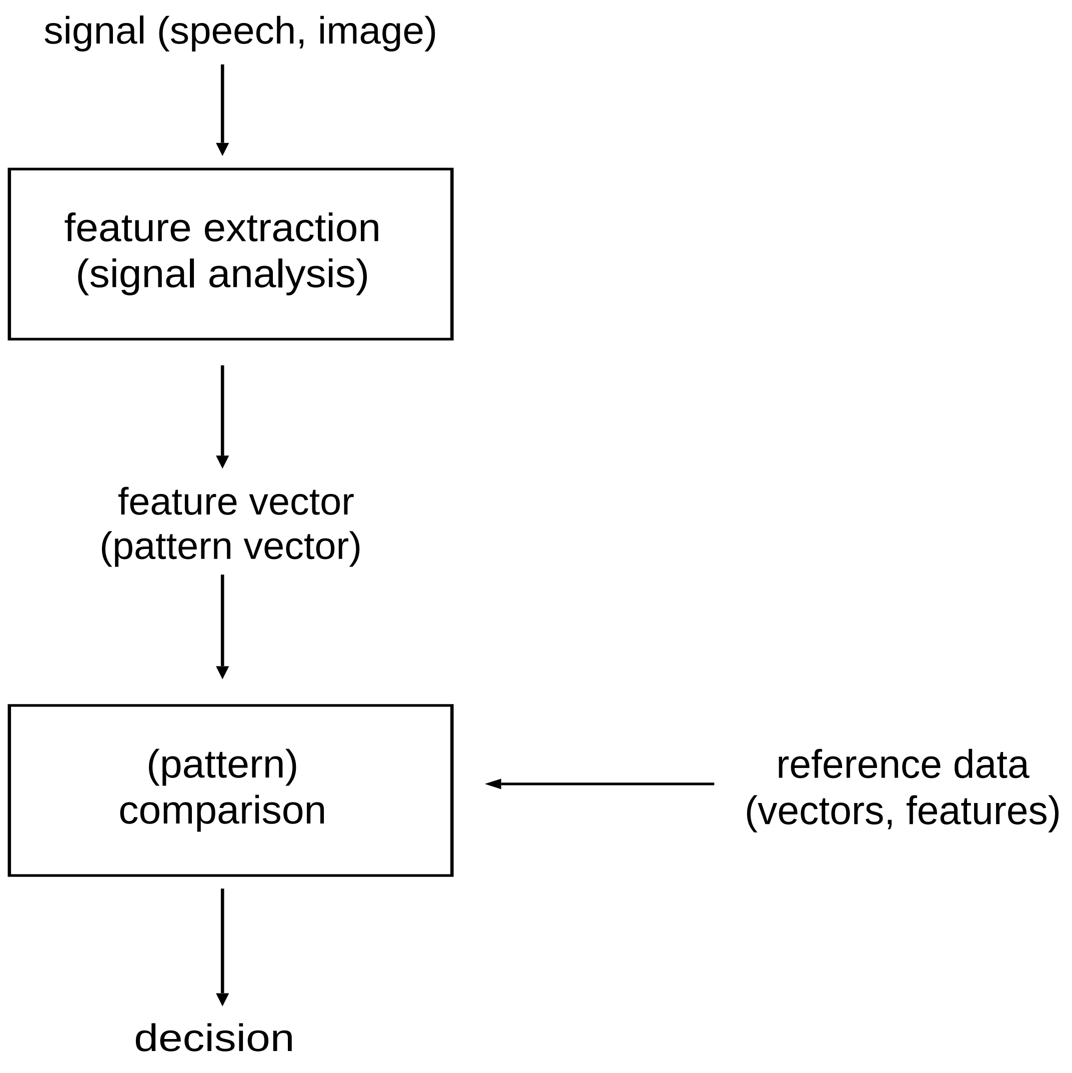


Figure 1.7: Schematic representation of the physiological mechanism of speech production



Examples:

- Spoken language
- Written numbers (letters)
- Cell recognition (red blood cells)

Examples of applications of Fourier Transform:

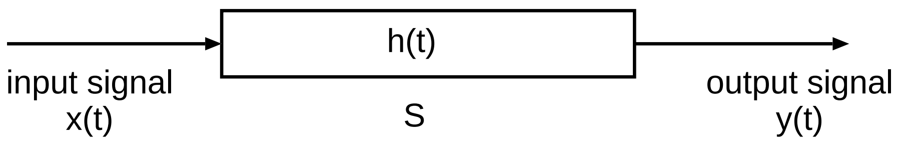
- Electrical switchgears
- Recognition and coding
  - Speech and general acoustic signals
  - Image signals
- Time series analysis:
  - Astronomical measurement curves
  - Stock-market course
  - ...
- Computer tomography
- Solving differential equations
- Description of image production in optical systems

1.2 Linear time-invariant Systems

Example:

– speech production

– electrical systems



symbolic:

$$\{t \to y(t)\} = S\,\{t \to x(t)\}$$

simplified:

$$y(t) = S\,\{x(t)\}$$

- Note: the complete time domain of the function is important, not individual positions in time t.

more exact:

$$y \;=\; S\,\{x\}$$

LTI–System: (LTI = Linear Time-Invariant)

- Linear:

  Additive:

$$S\,\{x_1 + x_2\} \;=\; S\,\{x_1\} + S\,\{x_2\}$$

  Homogeneous:

$$S\,\{\alpha\, x\} \;=\; \alpha S\,\{x\}\,, \qquad \alpha \in \mathbb{R}$$

- Time-invariant:

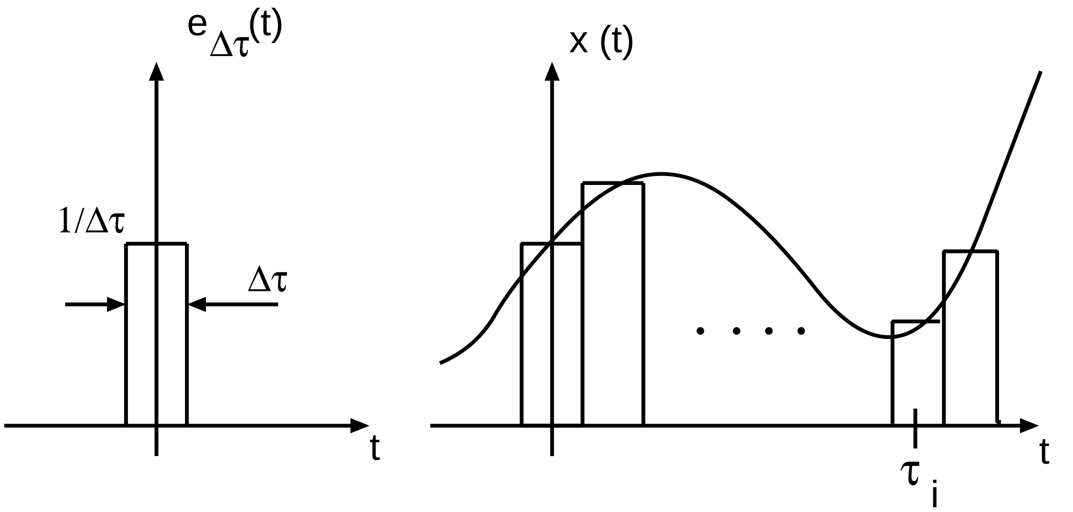
$$\{t \to y(t - t_0)\} = S\,\{t \to x(t - t_0)\}\,, \qquad t_0 \in \mathbb{R}$$

Mathematical theorem:

- Linearity and time invariance result in convolution representation
- Output signal $y(t)$ of LTI system S with input signal $x(t)$:

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} x(t-\tau)\, h(\tau)\, d\tau \\ &= \int_{-\infty}^{\infty} x(\tau)\, h(t-\tau)\, d\tau \\ &= x(t) * h(t) \end{aligned}$$

- h: impulse response of the system S



- system response $h_{\Delta\tau}(t)$ to excitation $e_{\Delta\tau}(t)$:

$$h_{\Delta\tau}(t) \quad = \quad S\ \{e_{\Delta\tau}(t)\}$$

- signal $x(t)$ is represented as sum of amplitude weighted and time shifted elementary functions $e_{\Delta\tau}(t)$:

$$x(t) \quad = \quad \lim_{\Delta\tau \to 0} \left[\sum_i x(\tau_i)\, e_{\Delta\tau}(t-\tau_i)\, \Delta\tau \right]$$

Hence the following holds for the output signal $y(t)$:

$$\begin{aligned} y(t) &= S\{x(t)\} \\ &= S\left\{ \lim_{\Delta\tau \to 0} \sum_i x(\tau_i)\, e_{\Delta\tau}(t-\tau_i)\, \Delta\tau \right\} \\ &= \lim_{\Delta\tau \to 0} \left[S\left\{ \sum_i x(\tau_i)\, e_{\Delta\tau}(t-\tau_i)\, \Delta\tau \right\} \right] \end{aligned}$$

additivity:

$$= \lim_{\Delta\tau \to 0} \left[\sum_i S\{\, x(\tau_i)\, e_{\Delta\tau}(t-\tau_i)\, \Delta\tau\} \right]$$

homogeneity (for $x(\tau_i)$ and $\Delta\tau$):

$$= \lim_{\Delta\tau \to 0} \left[\sum_i x(\tau_i)\, S\{\, e_{\Delta\tau}(t-\tau_i)\,\}\ \Delta\tau \right]$$

time invariance:

$$= \lim_{\Delta\tau \to 0} \left[\sum_i x(\tau_i)\, h_{\Delta\tau}(t-\tau_i)\, \Delta\tau \right]$$

limiting case $\Delta\tau \to 0$:

$$\begin{array}{rcl} \sum & \longrightarrow & \int \\ \Delta\tau & \longrightarrow & d\tau \\ \tau_i & \longrightarrow & \tau \\ h_{\Delta\tau}(t) & \longrightarrow & h(t) \end{array}$$

result:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)\, h(t-\tau)\, d\tau = x(t) * h(t)$$

$h(t)$: impulse response of the system

Examples of LTI-operations:

- Oscillatory systems (electrical or mechanical) with external excitation:

$$x(t) \longrightarrow \boxed{\quad h(\tau) \quad} \longrightarrow y(t)$$

$$y(t) \;=\; \int h(t-\tau)\; x(\tau)\; d\tau$$

$$y''(t) + 2\alpha y'(t) + \beta^2 y(t) = x(t)$$

α, β: parameters depending on the oscillatory system

- More general electrical engineering systems:
  high-pass, low-pass, band-pass

- Sliding average value:

$$x(t) \longrightarrow \boxed{\quad S \quad} \longrightarrow y(t) := \overline{x}(t)$$

$$\overline{x}(t) \;=\; \frac{1}{T} \int\limits_{-T/2}^{+T/2} x(t+\tau)\; d\tau$$

- Differentiator:

$$x(t) \longrightarrow \boxed{\quad S \quad} \longrightarrow y(t) := x'(t)$$

- Comb filter: "hypothesized" period T

$$x(t) \longrightarrow \boxed{\quad S \quad} \longrightarrow y(t) := x(t) - x(t-T)$$

- In general: linear differential equations with coefficients c_k and d_l

$$\sum_k c_k y^{(k)}(t) = \sum_l d_l x^{(l)}(t)$$

[+ further constraints]

Example of a non-linear system:

$$
\begin{aligned}
\text{system:} \quad y(t) &= x^2(t) \\
x(t) &= A\cos(\beta t) \\
\Longrightarrow y(t) &= A^2\cos^2(\beta t) = \frac{A^2}{2}(1+\cos(2\beta t))
\end{aligned}
$$

frequency doubling

1.3 Fourier Transform

Sinusoidal oscillation:

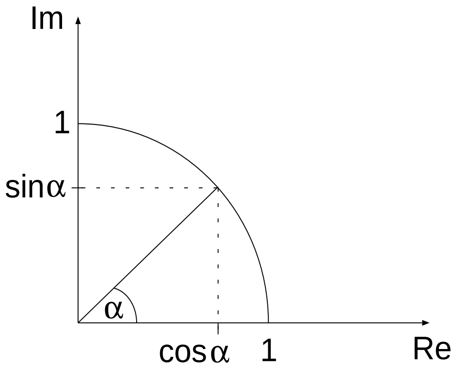
$$x(t) \;=\; A\,\sin\,(\,\omega\,t\;+\;\varphi\,)$$

amplitude A

phase / null phase φ

angular frequency $\omega \;=\; 2\,\pi\,f$

$$j^2 \;=\; -1, \qquad j \in \mathbb{C}$$



complex representation: $\quad e^{j\,\alpha} \;=\; \cos\,\alpha\;+\;j\,\sin\,\alpha, \qquad \alpha \in \mathbb{R}$

$$\cos\,\alpha \;=\; \frac{e^{j\alpha}\;+\;e^{-j\alpha}}{2} \qquad \text{and} \qquad \sin\,\alpha \;=\; \frac{e^{j\alpha} - e^{-j\alpha}}{2j}$$

dimension:

$$\begin{aligned} \mathrm{DIM}(\omega)\;\mathrm{DIM}(t) \;&=\; 1 \\ \mathrm{DIM}(\omega) \;&=\; \frac{1}{\mathrm{DIM}(t)} = \frac{1}{[\mathrm{sec}]} = [\mathrm{Hz}] \end{aligned}$$

LTI-System

$$y(t) \; = \; \int_{-\infty}^{\infty} x(t-\tau)h(\tau)d\tau \; = \; x(t) * h(t)$$

- Determine the following specific input signal:

$$x(t) \; = \; A\, e^{j(\omega t + \varphi)}$$

- For this input signal the output signal becomes:

$$\begin{aligned} y(t) \; &= \; \int_{-\infty}^{\infty} A\, e^{j(\omega(t-\tau)+\varphi)} h(\tau) d\tau \\ &= \; A\, e^{j(\omega t+\varphi)} \underbrace{\int_{-\infty}^{\infty} h(\tau) e^{-j\omega\tau} d\tau}_{H(\omega) \; = \; F\{h(\tau)\}} \\ &= \; x(t) \cdot H(\omega) \end{aligned}$$

- Definition of the Fourier transform:

$$H(\omega) \; = \; \int_{-\infty}^{\infty} h(\tau) e^{-j\omega\tau} d\tau \; = \; F\{h(\tau)\} \; = \; F\{\tau \rightarrow h(\tau)\}$$

($\rightarrow$ decomposition into $e^{-j\omega\tau}$)

- $H(\omega)$ is called transfer function of the system

Remark about $x(t) \; = \; A\, e^{j(\omega t+\varphi)}$:

- The shape of the input signal $x(t)$, i.e. its frequency ω ("eigenfunction") remains invariant
- Amplitude (intensity) and phase (time shift) are depending on $H(\omega)$ ("eigenvalue")

($\rightarrow$ analogy to the problem of eigenvalues in linear algebra)

Remarks

- FT is complex:

$$H(\omega) \;=\; Re\left\{H(\omega)\right\} \;+\; j\; Im\left\{H(\omega)\right\} \;=\; |H(\omega)|\; e^{j\Phi(\omega)}$$

- Amplitude (spectrum):

$$|H(\omega)| \;=\; \sqrt{Re\left\{H(\omega)\right\}^2 + Im\left\{H(\omega)\right\}^2}$$

- Phase (spectrum):

$$\Phi(\omega) \;=\; \begin{cases} \arctan\left(\dfrac{Im\left\{H(\omega)\right\}}{Re\left\{H(\omega)\right\}}\right) & Re\left\{H(\omega)\right\} \;>\; 0 \\[2ex] \arctan\left(\dfrac{Im\left\{H(\omega)\right\}}{Re\left\{H(\omega)\right\}}\right) \;+\; \pi & Re\left\{H(\omega)\right\} \;<\; 0 \\[2ex] \dfrac{\pi}{2} & \begin{array}{l} Re\left\{H(\omega)\right\} \;=\; 0, \\ Im\left\{H(\omega)\right\} \;>\; 0 \end{array} \\[2ex] -\dfrac{\pi}{2} & \begin{array}{l} Re\left\{H(\omega)\right\} \;=\; 0, \\ Im\left\{H(\omega)\right\} \;<\; 0 \end{array} \end{cases}$$

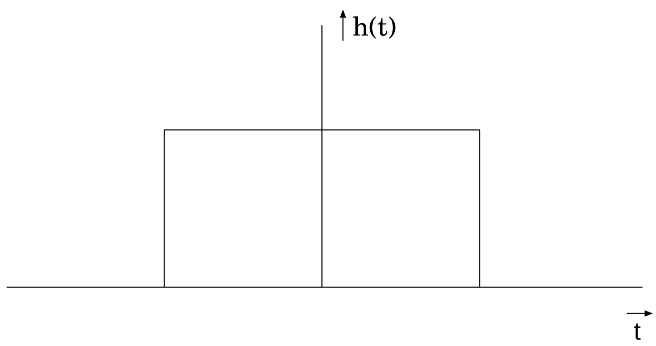
Examples of Fourier transforms:

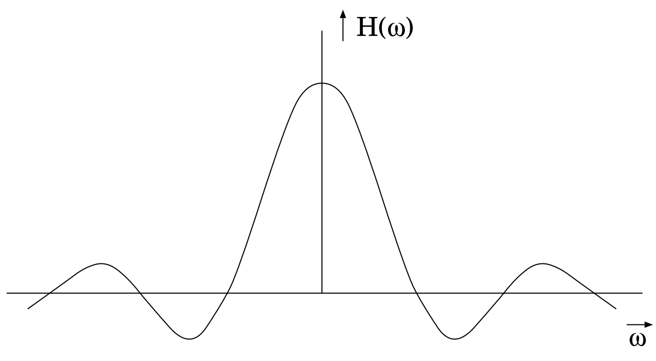
1. Rectangle function

$$h(t) \; = \; \mathrm{rect}(\frac{t}{T}) \; = \; \begin{cases} 1, & |t| \le T/2 \\ 0, & |t| > T/2 \end{cases}$$

$$\begin{aligned} H(\omega) \; &= \; \int\limits_{-\infty}^{\infty} h(t) e^{-j\omega t} dt = \int\limits_{-\frac{T}{2}}^{\frac{T}{2}} e^{-j\omega t} dt = \frac{1}{-j\omega}\left[e^{-j\omega\frac{T}{2}} - e^{j\omega\frac{T}{2}}\right] \\ &= \; \frac{2}{\omega}\sin(\frac{\omega T}{2}) \; = \; \frac{T\sin(\frac{\omega T}{2})}{\frac{\omega T}{2}} \end{aligned}$$

(here: $Im\{H(\omega)\} \; = \; 0$)



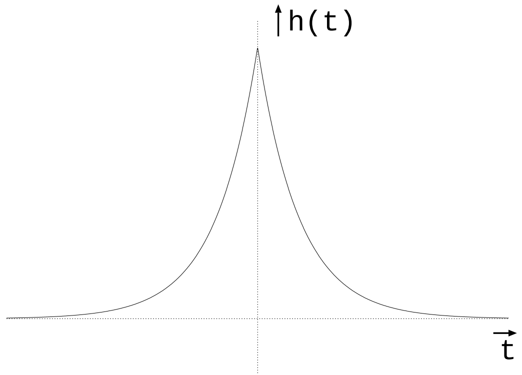


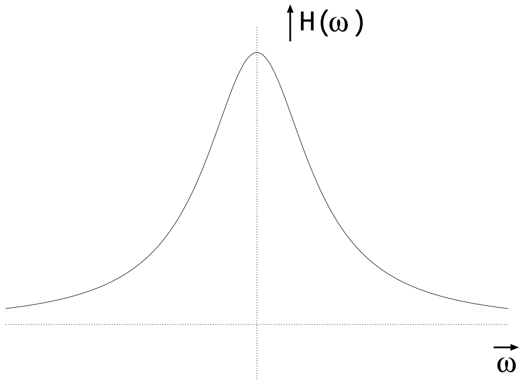
2. Double-sided exponential

$$h(t) \;=\; e^{-\alpha|t|} \quad \text{with} \quad \alpha > 0$$

$$\begin{aligned}
H(\omega) &= \int_{-\infty}^{\infty} h(t)e^{-j\omega t}dt \\
&= \int_{0}^{\infty} e^{-(\alpha+j\omega)t}dt + \int_{0}^{\infty} e^{-(\alpha-j\omega)t}dt \\
&= \left[\frac{e^{-(\alpha+j\omega)t}}{-(\alpha+j\omega)} + \frac{e^{-(\alpha-j\omega)t}}{-(\alpha-j\omega)}\right]_0^{\infty} \\
&= 0 + 0 - \frac{1}{-(\alpha+j\omega)} - \frac{1}{-(\alpha-j\omega)} \\
&= \frac{\alpha - j\omega + \alpha + j\omega}{\alpha^2+\omega^2} \\
&= \frac{2\alpha}{\alpha^2+\omega^2}
\end{aligned}$$

- Imaginary part equals 0
- Infinite spectrum
- No zeros





- If $h(t)$ is symmetric (i.e. $h(t) = h(-t)$), imaginary parts drop away and the real part is sufficient

3. Damped oscillations

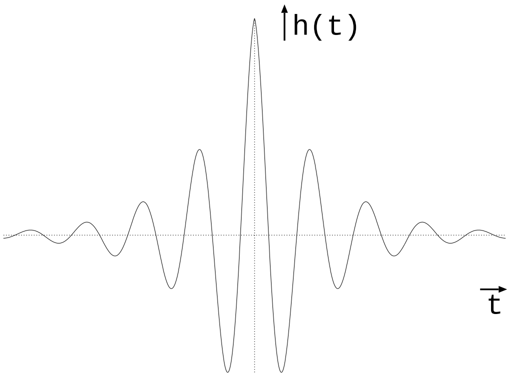
$$h(t) \;=\; e^{-\alpha|t|}\cos(\beta t) \quad \text{with} \quad \alpha > 0$$

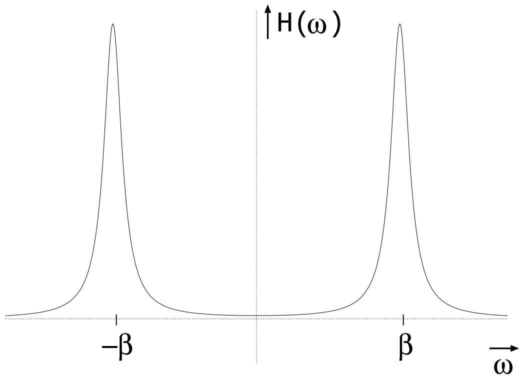
$$\begin{aligned}
H(\omega) &= \int\limits_{-\infty}^{\infty} h(t)e^{-j\omega t}dt \\
&= \int\limits_{0}^{\infty} e^{-(\alpha+j\omega)t}\cos(\beta t)dt + \int\limits_{0}^{\infty} e^{-(\alpha-j\omega)t}\cos(\beta t)dt \\
&= \int\limits_{0}^{\infty} e^{-(\alpha+j\omega)t}\,\frac{e^{j\beta t}+e^{-j\beta t}}{2}dt + \int\limits_{0}^{\infty} e^{-(\alpha-j\omega)t}\,\frac{e^{j\beta t}+e^{-j\beta t}}{2}dt \\
&= \quad \ldots \qquad \text{(elementary calculation)} \\
&= \frac{\alpha}{\alpha^2+(\omega-\beta)^2} + \frac{\alpha}{\alpha^2+(\omega+\beta)^2}
\end{aligned}$$

- Limiting case:

$$H(\omega)|_{\omega=\pm\beta} \;=\; \frac{1}{\alpha} \;+\; \frac{\alpha}{\alpha^2 \;+\; (2\beta)^2}$$

$\Longrightarrow$ tends towards ∞ or $-\infty$ if α tends towards 0

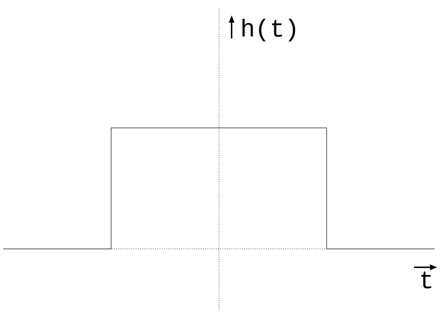


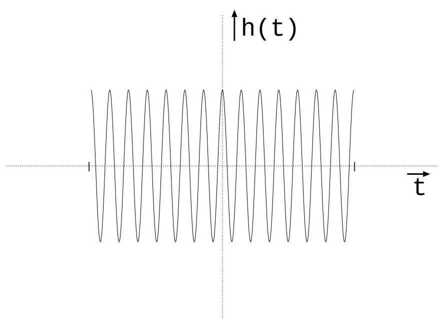


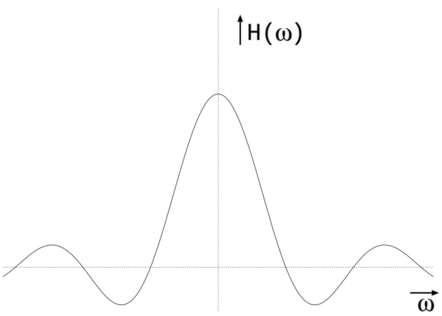
4. Modulated rectangle function ("truncated cosine")

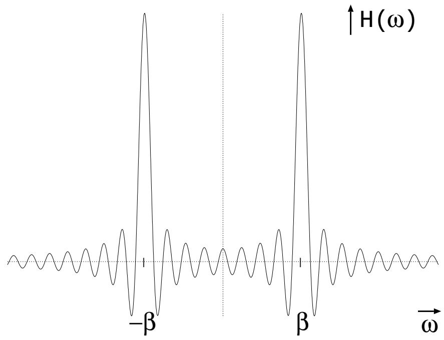
$$h(t) = \begin{cases} \cos(\beta\, t), & |t| \leq T/2 \\ 0, & |t| > T/2 \end{cases}$$

$$\begin{aligned} H(\omega) &= \int\limits_{-\infty}^{\infty} h(t) e^{-j\omega t} dt \\ &= \int\limits_{-\frac{T}{2}}^{\frac{T}{2}} \cos(\beta\, t) e^{-j\omega t} dt \\ &= \quad \ldots \qquad \text{(elementary calculation)} \\ &= \frac{T}{2} \left[\frac{\sin\left((\omega - \beta)\, \frac{T}{2} \right)}{(\omega - \beta)\, \frac{T}{2}} + \frac{\sin\left((\omega + \beta)\, \frac{T}{2} \right)}{(\omega + \beta)\, \frac{T}{2}} \right] \end{aligned}$$

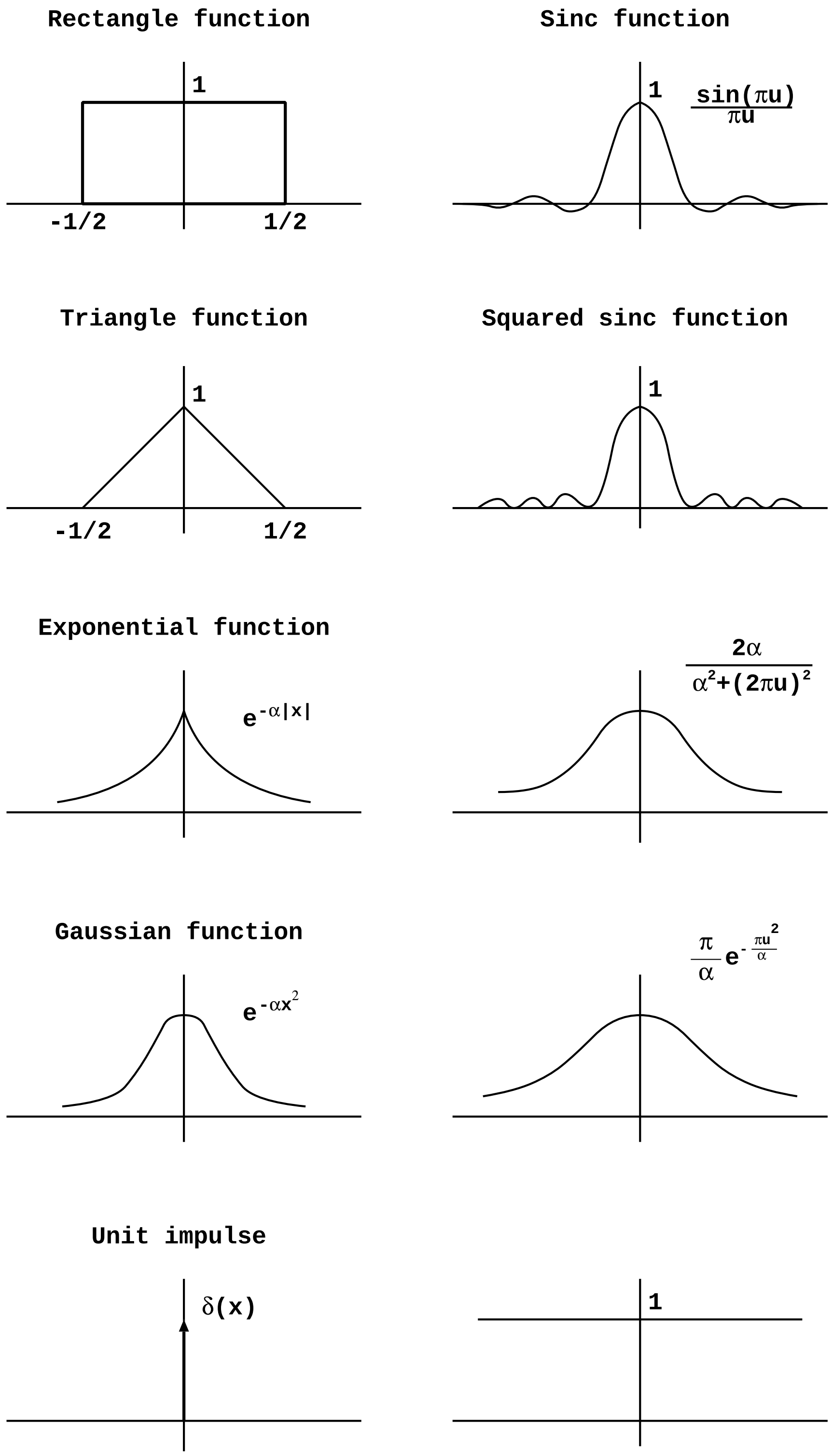








Fourier Transform pairs ($u = \omega/2\pi$)



Inverse Fourier–transform

$$H(\omega) \;=\; \int\limits_{-\infty}^{\infty} h(t) e^{-j\omega t} dt$$

assumption: $$\tilde{h}(t) \;=\; \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} H(\omega) e^{j\omega t} d\omega$$

with: $$H(\omega) = \int\limits_{-\infty}^{\infty} h(\tau) e^{-j\omega\tau} d\tau$$

inserting $H(\omega)$ in $\tilde{h}(t)$:

$$\begin{aligned}
\tilde{h}(t) \;&=\; \frac{1}{2\pi} \lim_{\Omega, T \to \infty} \int\limits_{-\Omega}^{\Omega} \left[\int\limits_{-T}^{T} h(\tau)\, e^{j\omega(t-\tau)}\, d\tau \right] d\omega \\
&=\; \frac{1}{2\pi} \lim_{\Omega \to \infty} \lim_{T \to \infty} \int\limits_{-T}^{T} \int\limits_{-\Omega}^{\Omega} e^{j\omega(t-\tau)}\, d\omega\, h(\tau)\, d\tau \\
&=\; \lim_{\Omega \to \infty} \lim_{T \to \infty} \frac{1}{\pi} \int\limits_{-T}^{T} \frac{\sin\left(\Omega(t-\tau)\right)}{t-\tau}\, h(\tau)\, d\tau \\
&=\; \lim_{\Omega \to \infty} \frac{1}{\pi} \int\limits_{-\infty}^{\infty} \frac{\sin\left(\Omega(t-\tau)\right)}{t-\tau}\, h(\tau)\, d\tau \\
&=\; h(t)
\end{aligned}$$

due to:

$$\lim_{\Omega \to \infty} \frac{1}{\pi} \int\limits_{-\infty}^{\infty} \frac{\sin(\Omega t)}{t}\, h(t)\, dt \;=\; h(0)$$

formal expression:

$$h(t) \;=\; \int\limits_{-\infty}^{\infty} \underbrace{\left[\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{j\omega(t-\tau)}\, d\omega \right]}_{=\; \delta(t-\tau)} h(\tau)\, d\tau$$

($\rightarrow$ distribution theory, see there for stronger proof)

1.4 Properties of the Fourier Transform

Symmetry

$$H(\omega) = \int_{-\infty}^{\infty} h(t)\, e^{-j\omega t}\, dt = F\left\{h(t)\right\}$$

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega)\, e^{j\omega t}\, d\omega = F^{-1}\left\{H(\omega)\right\}$$

$$F^2\{h(t)\} = F\{H(\omega)\} = 2\pi h(-t)$$

$$F^{-1}\, F\{h(t)\} = F^{-1}\{H(\omega)\} = h(t)$$

- Time domain and frequency domain are correlated symmetrically.
- Properties of FT are valid in both domains, especially the convolution theorem (see later).

Theorems for the Fourier transform

$$H(\omega) \;=\; \int\limits_{-\infty}^{\infty} e^{-j\omega t}\, h(t)\, dt$$

consider the equation:

$$H(\omega) \;=\; F\,\{h(t)\}$$

more exact:

$$\{\omega \rightarrow H(\omega)\} \;=\; F\,\{t \rightarrow h(t)\}$$

1. Linearity: integral operator is linear

2. Inverse scaling, similarity principle:

$$\int\limits_{-\infty}^{\infty} h(\alpha t)\, e^{-j\omega t}\, dt \;=\; \frac{1}{|\alpha|} \int\limits_{-\infty}^{\infty} h(\tau)\, e^{-j\frac{\omega}{\alpha}\tau}\, d\tau$$

$$F\{h(\alpha t)\} \;=\; \frac{1}{|\alpha|}\, H(\frac{\omega}{\alpha}), \qquad \alpha \in \mathbb{R}\backslash\{0\}$$

   Note:

   Absolute value, because integral boundaries are swapped for $\alpha < 0$.

3. Shift: $h(t - t_0)$

$$\begin{aligned}\int\limits_{-\infty}^{\infty} h(t-t_0)\, e^{-j\omega t}\, dt \;&=\; e^{-j\omega t_0} \int\limits_{-\infty}^{\infty} h(t-t_0)\, e^{-j\omega(t-t_0)}\, dt \\ &=\; e^{-j\omega t_0} \int\limits_{-\infty}^{\infty} h(\tau)\, e^{-j\omega\tau}\, d\tau\end{aligned}$$

$$\Longrightarrow F\{h(t-t_0)\} = e^{-j\omega t_0} H(\omega) \quad t_0 \in \mathbb{R}$$

$$\text{with} \quad H(\omega) = F\{h(t)\}$$

important:

$$| F\{h(t-t_0)\} | = | F\{h(t)\} | \qquad , \quad \text{because}$$

$$\begin{aligned} |e^{-j\omega t_0}| &= |e^{-ju}| = |\cos u - j \sin u| \\ &= \sqrt{\cos^2 u + sin^2 u} \\ &= 1 \end{aligned}$$

4. Symmetry and antisymmetry:

$$h(t) = h(-t) \qquad \text{results in} \qquad \text{Im}\{H(\omega)\} = 0$$

$$h(t) = -h(-t) \qquad \text{results in} \qquad \text{Re}\{H(\omega)\} = 0$$

5. Complex conjugation: suppose that $h(t)$ is a complex function

$$\begin{aligned} \int_{-\infty}^{\infty} \overline{h(t)}\, e^{-j\omega t}\, dt &= \overline{\int_{-\infty}^{\infty} \overline{h(t)\, e^{j\omega t}}\, dt} \\ &= \overline{\int_{-\infty}^{\infty} h(t)\, e^{j\omega t}\, dt} = \overline{H(-\omega)} \end{aligned}$$

$$F\{\overline{h(t)}\} = \overline{H(-\omega)} = \overline{F\{h(t)\}}$$

Special case: $h(t)$ is real, so $h(t) = \overline{h(t)}$

$$\Longrightarrow H(\omega) = \overline{H(-\omega)} \Longrightarrow | H(\omega) | = | \overline{H(-\omega)} | = | H(-\omega) |$$

6. Differentiation:

$$\begin{aligned}
\frac{dh}{dt} &= \frac{\partial}{\partial t} \left[\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} H(\omega)\, e^{j\omega t}\, d\omega \right] \\
&= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} H(\omega)\, j\omega\, e^{j\omega t}\, d\omega
\end{aligned}$$

$$F\{\frac{dh(t)}{dt}\} = j\omega\, F\{h(t)\}$$

Interpretation: differentiation = enhancement of high frequencies (due to the multiplication with ω)

7. Integration:

$$F\{\int\limits_{-\infty}^{t} h(\tau) d\tau\} = \frac{1}{j\omega}\, F\{h(t)\}$$

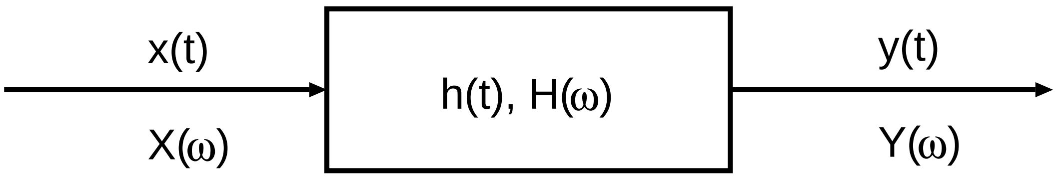
Proof: similar to differentiation or inversion

8. Modulation principle:

$$\begin{aligned}
F\{h(t)\cos(\omega_0 t)\} &= \int\limits_{-\infty}^{\infty} h(t)\, \cos(\omega_0 t)\, e^{-j\omega t}\, dt \\
&= \frac{1}{2} \left[\int\limits_{-\infty}^{\infty} h(t)\, e^{j\omega_0 t}\, e^{-j\omega t}\, dt + \int\limits_{-\infty}^{\infty} h(t)\, e^{-j\omega_0 t}\, e^{-j\omega t}\, dt \right] \\
&= \frac{1}{2} \left[\int\limits_{-\infty}^{\infty} h(t)\, e^{-j(\omega-\omega_0)t}\, dt + \int\limits_{-\infty}^{\infty} h(t)\, e^{-j(\omega+\omega_0)t}\, dt \right] \\
&= \frac{1}{2} \left[H(\omega-\omega_0) + H(\omega+\omega_0) \right]
\end{aligned}$$

and similarly

$$F\{ h(t)\, \sin(\omega_0 t) \} = \frac{1}{2j} \left[H(\omega-\omega_0) - H(\omega+\omega_0) \right]$$



Convolution theorem

- Convolution in time domain corresponds to multiplication in frequency domain

Time domain: $$y(t) \; = \; x(t) * h(t) \; = \int\limits_{-\infty}^{\infty} x(t-\tau)\; h(\tau)\; d\tau$$

Frequency domain:

$$\begin{aligned}
Y(\omega) &= \int\limits_{-\infty}^{\infty} e^{-j\omega t} \left[\int\limits_{-\infty}^{\infty} h(\tau)\; x(t-\tau)\; d\tau \right] dt \\
&= \int\limits_{-\infty}^{\infty} h(\tau) \left[\int\limits_{-\infty}^{\infty} x(t-\tau)\; e^{-j\omega t}\; dt \right] d\tau \\
&= \int\limits_{-\infty}^{\infty} h(\tau)\; X(\omega)\; e^{-j\omega\tau}\; d\tau \qquad \text{(shifting)} \\
&= X(\omega) \int\limits_{-\infty}^{\infty} h(\tau)\; e^{-j\omega\tau}\; d\tau \\
&= X(\omega)\; H(\omega)
\end{aligned}$$

- Likewise, multiplication in time domain corresponds to convolution in frequency domain (note the factor $\frac{1}{2\pi}$):

  Time domain: $\quad y(t) \; = \; a(t) \cdot b(t)$

  Frequency domain:

$$
\begin{aligned}
Y(\omega) &= \int_{-\infty}^{\infty} a(t) \cdot b(t) \; e^{-j\omega t} \; dt \\
&= \int_{-\infty}^{\infty} a(t) \frac{1}{2\pi} \int_{-\infty}^{\infty} B(\tilde{\omega}) e^{j\tilde{\omega}t} \; e^{-j\omega t} \; d\tilde{\omega} \; dt \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} B(\tilde{\omega}) \int_{-\infty}^{\infty} a(t) e^{-j(\omega-\tilde{\omega})t} \; dt \; d\tilde{\omega} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} A(\omega - \tilde{\omega}) \cdot B(\tilde{\omega}) d\tilde{\omega} \\
&= \frac{1}{2\pi} A(\omega) * B(\omega)
\end{aligned}
$$

- Motivation for the Fourier transform:
  FT gives the "simplest" representation of the system operation, because every LTI-System can be interpreted as convolution of the input signal $x(t)$ and the impulse response of the system $h(t)$. Convolution can be then efficiently calculated using FT and convolution theorem.

- Mathematical: eigenfunctions

Example: Oscillator with excitation

$$x(t) \longrightarrow \boxed{\quad Oscillator \quad} \longrightarrow y(t)$$

$$\begin{aligned}
y''(t) &+ 2\alpha\, y'(t) + \beta^2\, y(t) = x(t) \\
x(t) &= \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} X(\omega) e^{j\omega t} d\omega \\
y(t) &= \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} Y(\omega) e^{j\omega t} d\omega \\
y'(t) &= \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} Y(\omega) j\omega\; e^{j\omega t} d\omega \\
y''(t) &= \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} Y(\omega) [-\omega^2]\; e^{j\omega t} d\omega
\end{aligned}$$

$$\begin{aligned}
\int\limits_{-\infty}^{+\infty} [-\omega^2 + 2\alpha j\omega + \beta^2] Y(\omega) e^{j\omega t} d\omega &= \int\limits_{-\infty}^{+\infty} X(\omega) e^{j\omega t} d\omega \\
\int\limits_{-\infty}^{+\infty} \underbrace{\left\{[-\omega^2 + 2\alpha j\omega + \beta^2]\; Y(\omega) - X(\omega)\right\}}_{=0} e^{j\omega t} d\omega &= 0 \qquad \forall t
\end{aligned}$$

In this way we obtain the transfer function of an oscillator:

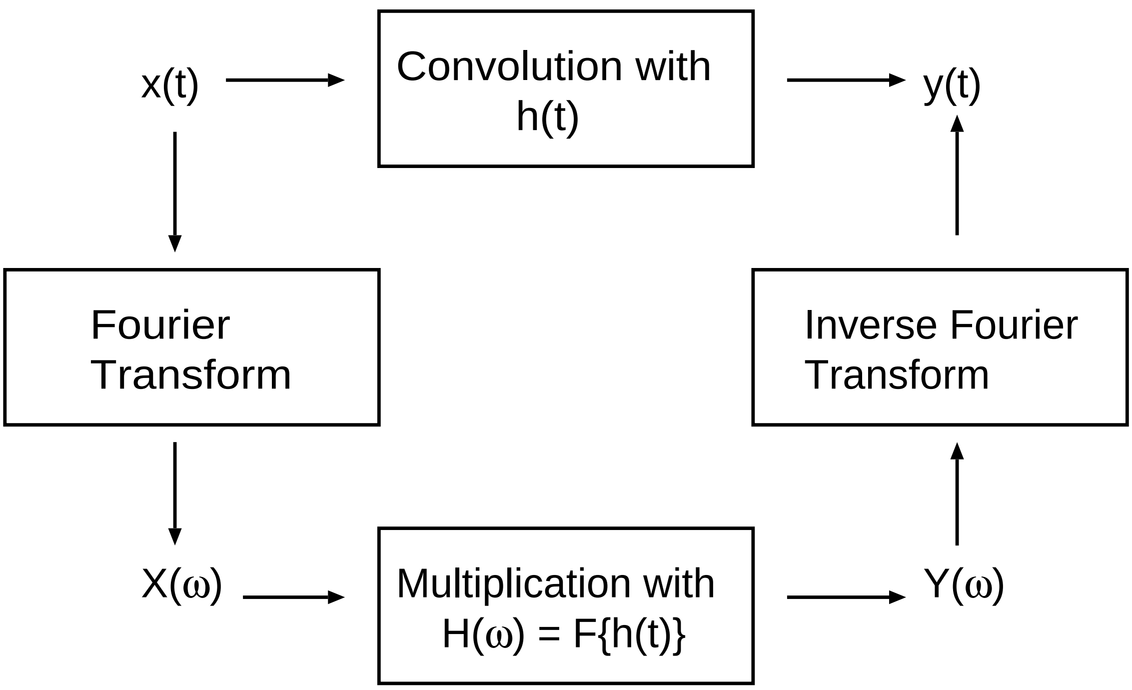
$$H(\omega) \;=\; \frac{Y(\omega)}{X(\omega)} = \frac{1}{-\omega^2 + 2\alpha j\omega + \beta^2}$$

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} H(\omega) e^{j\omega t} d\omega$$

(can be given explicitly)

$$y(t) = \int_{-\infty}^{+\infty} x(t)\ h(t-\tau) d\tau$$

Note:
$y(t)$ does not contain the component which corresponds to the homogeneous differential equation of the oscillator.



1.5 Parseval Theorem

Convolution theorem:

$$
\begin{aligned}
F^{-1}\{H(\omega)\,X(\omega)\} &= \int_{-\infty}^{\infty} h(t)\,x(\tau - t)\,dt \\
(\star) \qquad \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)\,X(\omega)\,e^{j\omega\tau}\,d\omega &= (h * x)(\tau)
\end{aligned}
$$

We make two special assumptions:

i) $x(-t) := \overline{h(t)}$, then: $X(\omega) \;=\; \overline{H(\omega)}$

ii) $\tau = 0$

Inserting in $(\star)$ results in:

$$
\begin{aligned}
\frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)\overline{H(\omega)}\,d\omega &= \int_{-\infty}^{\infty} h(t)\overline{h(t)}\,dt \\
\frac{1}{2\pi}\int_{-\infty}^{\infty} |H(\omega)|^2\,d\omega &= \int_{-\infty}^{\infty} |h(t)|^2\,dt \;=\; E
\end{aligned}
$$

- Energy E in time domain $\quad=\quad$ Energy E in frequency domain (up to the factor $\frac{1}{2\pi}$; aid: use normalization factor $\frac{1}{\sqrt{2\pi}}$ for both directions of Fourier Transform)
- Physical aspect: energy conservation
- Mathematical aspect: unitary (orthogonal) representation in vector space
- $|H(\omega)|^2$ is called power spectral density.

1.6 Autocorrelation Function

Autocorrelation function

- Autocorrelation function of time continuous signal or function $h(t)$ is defined as:

$$R(t) = \int_{-\infty}^{\infty} h(\tau)\ h(t+\tau) d\tau$$

- The following equation is valid:

$$R(t) = h(t) * h(-t) \qquad \rightarrow \quad \text{which results in} \qquad R(t) = R(-t)$$

- Fourier transform gives: (”Wiener-Khinchin Theorem”)

$$F\{R(t)\} = H(\omega)\ \overline{H(\omega)} = |H(\omega)|^2$$

- Thus: Fourier transform connects autocorrelation function $R(t)$ and power spectral density $|H(\omega)|^2$

$$|H(\omega)|^2 \ = \int_{-\infty}^{\infty} R(t)\ e^{-j\omega t}\ dt \ = \int_{-\infty}^{\infty} R(t)\ \cos(\omega t)\ dt$$

- Remark:
  autocorrelation is a special case of the cross correlation between signals $x(\tau)$ and $h(t)$

$$C_{h,x} = \int_{-\infty}^{\infty} h(\tau)\ x(t+\tau) d\tau$$

1.7 Existence of the Fourier Transform

Conditions for $h(t)$ for the existence of the Fourier transform

$$H(\omega) \;=\; \int_{-\infty}^{\infty} e^{-j\omega t}\, h(t)\, dt \quad , \qquad h(t) \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t}\, H(\omega)\, d\omega$$

When are those equations valid?

Sufficient conditions:

1. $h(t)$ is absolutely integrable:

$$\int_{-\infty}^{\infty} |h(t)| dt < \infty$$

2. $h(t)$ has finite number of jumps, minima and maxima in each interval of $\mathbb{R}$
3. $h(t)$ has no infinite jumps

More general conditions are possible (but rather complex set of conditions):

- Generalized functions, distributions,
  definition as functional

- Example: δ-function:

$$\int_{-\infty}^{\infty} \delta(t)\, h(t)\, dt \;=\; h(0) \quad \text{for all functions } h$$

Impulse response:

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(t-\tau)\delta(\tau)\, d\tau \\ &= h(t) * \delta(t) \\ &= h(t) \end{aligned}$$

Consequence:

$$h(t) \equiv 1 \;\Rightarrow\; \int_{-\infty}^{\infty} \delta(t)\, dt = 1$$

- A function like $\delta(t)$ does not "exist". But it is possible to define the functional for each function $t \to h(t)$:

$$[t \to h(t)] \;\longrightarrow\; \tilde{\delta}(h) \;:=\; h(0)$$

1.8 δ-Function

Starting point: definition of the δ-function as a boundary case of a function $\delta_\epsilon(t)$:

$$\lim_{\epsilon\to 0} \int_{-\infty}^{+\infty} f(t)\, \delta_\epsilon(t)\, dt \;= f(0) \tag{1.1}$$

- Possible realizations of $\delta_\epsilon(t)$

a) $\delta_\epsilon(t) = \begin{cases} \dfrac{1}{2\epsilon} & t \in [-\epsilon, +\epsilon] \\ 0 & \text{otherwise} \end{cases}$

b) $\delta_\epsilon(t) = \dfrac{1}{\pi}\dfrac{\epsilon}{\epsilon^2 + t^2}$

c) $\delta_\epsilon(t) = \dfrac{1}{\pi}\dfrac{\sin(t/\epsilon)}{t}$

d) $\delta_\epsilon(t) = \dfrac{1}{\sqrt{2\pi\epsilon^2}}\, e^{-\dfrac{t^2}{2\epsilon^2}}$

- During inversion of the Fourier transform we have "formally" obtained:

$$\delta(t) \;=\; \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} e^{j\omega t}\, d\omega \;=\; \lim_{\Omega\to\infty} \frac{1}{\pi}\,\frac{\sin(\Omega t)}{t} \tag{1.2}$$

Fourier transform $F\{\delta(t)\}$:

$$F\{\delta(t)\} \;=\; \int\limits_{-\infty}^{+\infty} e^{-j\omega t}\delta(t)\, dt$$

due to (1.1) the following holds:

$$F\{\delta(t)\} \;=\; e^{j\omega t}|_{t=0} \;=\; 1$$

- Another derivation using (1.2):

$$\begin{aligned} \delta(t) \;&=\; \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} e^{j\omega t}\; F\{\delta(t)\}\, d\omega \qquad \text{general} \\ &=\; \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} e^{j\omega t}\, d\omega \qquad \text{according to (1.2)} \end{aligned}$$

Comparison results in:

$$F\{\delta(t)\} \;=\; 1$$

From this we obtain the following equations:

From symmetry property:

$$F\{1\} \;=\; 2\,\pi\,\delta(\omega)$$

From shifting theorem:

$$F\{e^{j\omega_0 t}\,1\} \;=\; 2\,\pi\,\delta(\omega-\omega_0)$$

$$\begin{aligned}
\cos(\omega_0\,t) &= \frac{1}{2}\,\left[\,e^{j\omega_0 t} \;+\; e^{-j\omega_0 t}\,\right] \\
&= \frac{1}{2}\left[\int_{-\infty}^{+\infty} \delta(\omega-\omega_0)\,e^{j\omega t}\,d\omega \;+\; \int_{-\infty}^{+\infty} \delta(\omega+\omega_0)\,e^{j\omega t}\,d\omega\right] \\
&= \pi\,\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\,\delta(\omega-\omega_0)\;+\;\delta(\omega+\omega_0)\,\right]\,e^{j\omega t}\,d\omega \\
F\{\,\cos(\omega_0\,t)\,\} &= \pi\,\left[\,\delta(\omega-\omega_0)\;+\;\delta(\omega+\omega_0)\,\right]
\end{aligned}$$

Note: another derivation:

consider "damped oscillations"

$$\frac{1}{2\pi}\,e^{-\alpha|t|}\;\cos(\omega_0 t)$$

in the limit $\alpha \to 0$.

Comb function

- define "comb function" (pulse train, sequence of δ-impulses):

$$x(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT)$$

- Fourier transform of comb function:

$$\begin{aligned}
X(\omega) &= \int_{-\infty}^{+\infty} x(t)\, e^{-j\omega t}\, dt \\
&= \int_{-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \delta(t - nT)\, e^{-j\omega t}\, dt \\
&= \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(t - nT)\, e^{-j\omega t}\, dt \\
&= \sum_{n=-\infty}^{+\infty} e^{-j\omega nT} \\
&= \quad \ldots \qquad \text{(see Papoulis 1962, p. 44)} \\
&= \frac{2\pi}{T} \sum_{n=-\infty}^{+\infty} \delta(\omega - n\frac{2\pi}{T})
\end{aligned}$$

- in words:

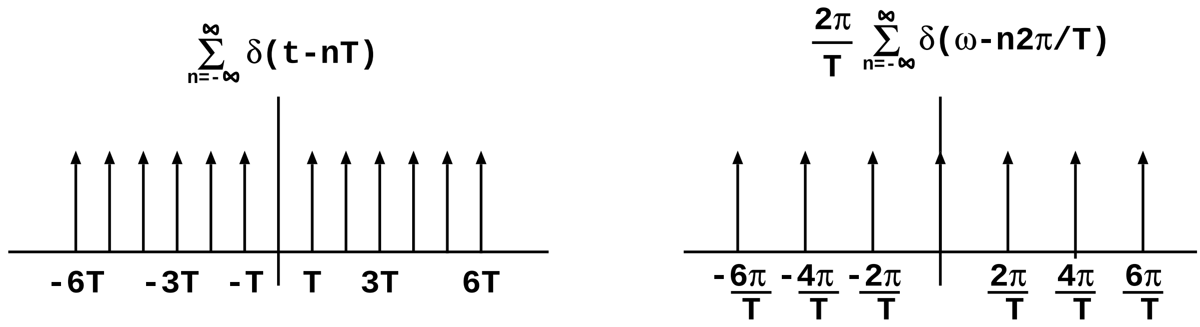
δ-impulse sequence with period T in time domain

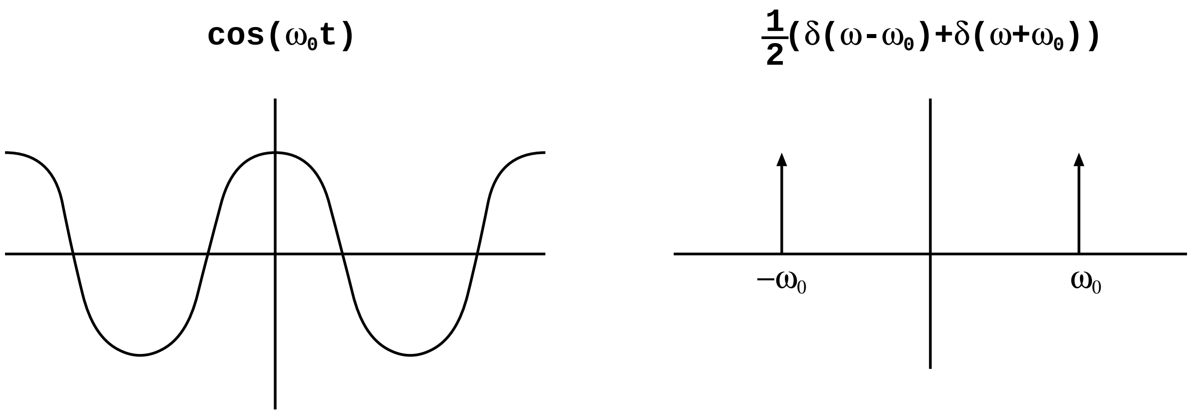
produces

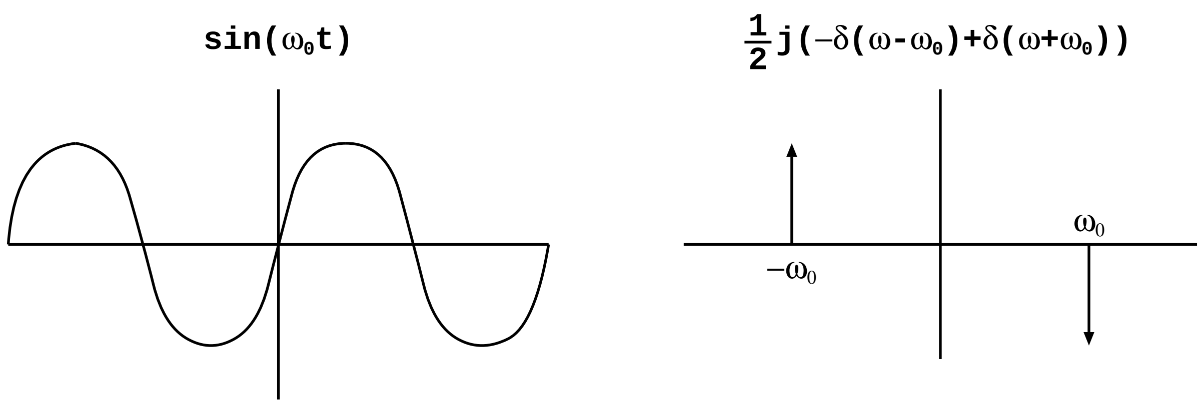
δ-impulse sequence with period $\frac{1}{T}$ in frequency domain
(i.e. $\frac{2\pi}{T}$ in ω-frequency domain)

comb function is transformed to comb function

Comb function







1.9 Motivation for Fourier Series

$$\begin{array}{rrcl} x: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & t & \longrightarrow & x(t) \end{array}$$

Consider a periodical function x with period T:

$$\begin{array}{rrcll} & x(t) & = & x(t+T) & \text{for each} \quad t \in \mathbb{R} \\ \text{then also} & x(t) & = & x(t+kT) & \text{for} \quad k \in \mathbb{Z} \end{array}$$

Examples:

- Constant function:

$$x_0(t) \quad = \quad A_0$$

- Harmonic oscillator:

$$x_1(t) \quad = \quad A_1 \, \cos\left(\frac{2\pi}{T} \, t \; + \; \varphi_1\right) , \qquad A_1 > 0$$

- All higher harmonic:

$$x_n(t) \quad = \quad A_n \, \cos\left(n\frac{2\pi}{T} \, t \; + \; \varphi_n\right) , \qquad A_n > 0$$

therefore

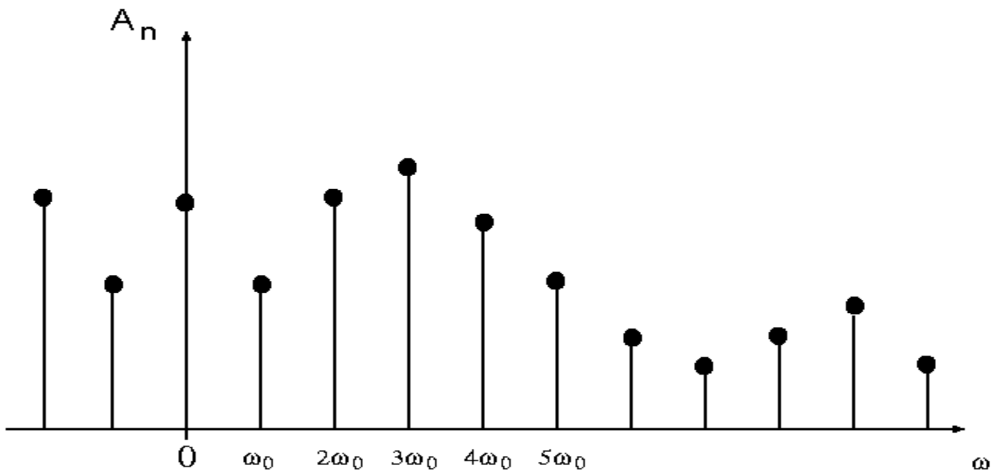
$$x(t) \; = \; \sum_{n=0}^{\infty} A_n \, \cos\left(n \, \omega_0 \, t \; + \; \varphi_n\right) \quad \text{with} \quad \omega_0 \; = \; \frac{2\,\pi}{T} , \qquad A_n \geq 0$$

is periodical with period $T = \frac{2\pi}{\omega_0}$

- Another notation:

$$x(t) \; = \; \sum_{n=-\infty}^{\infty} B_n \, e^{-j \, n \, \omega_0 \, t} \qquad \text{where} \quad B_n \quad \text{is a complex number}$$

Line spectrum representation



Real measured signal has always a "widespread" spectrum.

Reasons:

- Strictly periodical signal (almost) never exists
  - Period can fluctuate
  - "Wave form" within one period can fluctuate
  - Only a finite section of the signal is analyzed ("window function")
- Only a strictly periodical signal has a sharp line spectrum

Remarks:

- Fourier series are actually not strictly related to periodical functions: a finite interval of $\mathbb{R}$ is sufficient (the signal is then interpreted as infinitely prolonged).
- By transition from the finite interval to the complete real axis the Fourier series becomes Fourier integral.

Calculation of Fourier coefficients:

- Consider a periodical function $x(t)$ with period $T = \frac{2\pi}{\omega_0}$
- approach:

$$x(t) \;=\; \sum_{n=-\infty}^{+\infty} a_n \, e^{j\,n\,\omega_0\,t} \qquad a \in \mathbb{C}$$

- multiplication with $e^{-j\,m\,\omega_0\,t}$ where $m \in \mathbb{N}$ and integration over one period result in:

$$\int_{-T/2}^{+T/2} x(t)\, e^{-j\,m\,\omega_0\,t}\, dt \;=\; \sum_{n=-\infty}^{+\infty} a_n \int_{-T/2}^{+T/2} e^{j\,(n-m)\,\omega_0\,t}\, dt$$

- Due to "orthogonality" holds:

$$\int_{-T/2}^{+T/2} e^{j\,(n-m)\,\omega_0\,t}\, dt = \begin{cases} T & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

- Then:

$$\int_{-T/2}^{T/2} x(t)\, e^{-j\,m\,\omega_0\,t}\, dt \;=\; a_m\, T$$

- Result:

$$\begin{aligned} a_n \;&=\; \frac{1}{T} \int_{-T/2}^{+T/2} x(t)\, e^{-j\,n\,\omega_0\,t}\, dt \\ &=\; \frac{1}{T} \int_{-T/2}^{+T/2} x(t)\; \cos(n\,\omega_0\,t)\, dt \;-\; j\, \frac{1}{T} \int_{-T/2}^{+T/2} x(t)\; \sin(n\,\omega_0\,t)\, dt \end{aligned}$$

Spectrum of a periodical function

- If $x(t)$ is periodical with the period $T = \frac{2\pi}{\omega_0}$, then

$$x(t) \quad = \quad \sum_{n=-\infty}^{+\infty} a_n \, e^{j\,n\,\omega_0\,t}, \qquad a_n \in \mathbb{C}$$

- The Fourier transform $X(\omega)$ is:

$$\begin{aligned} X(\omega) &= F\{x(t)\} \\ &= \sum_{n=-\infty}^{+\infty} a_n \quad \underbrace{F\{e^{j\,n\,\omega_0\,t}\}}_{=2\pi\delta(\omega - n\omega_0)} \\ &= 2\,\pi \sum_{n=-\infty}^{+\infty} a_n \, \delta(\omega - n\omega_0) \end{aligned}$$

- Note:

  This derivation is formal, because the Fourier integral does not exist in the “usual sense”;
  strict derivation within the scope of distribution theory.

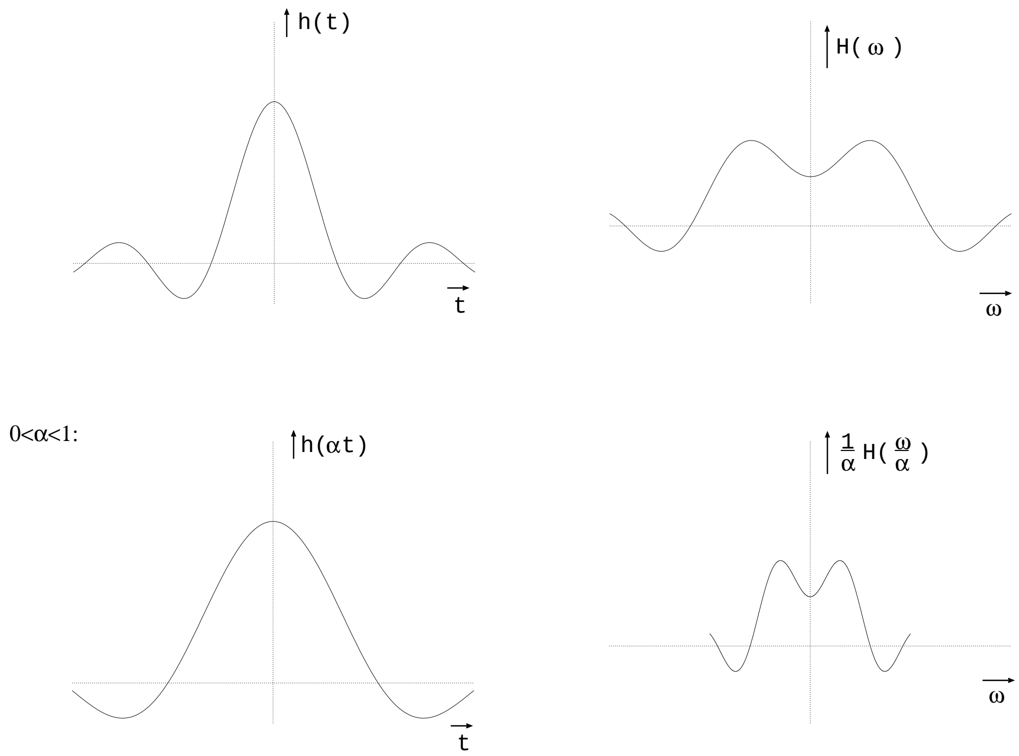
- In words:

  a periodic function with the period T has a Fourier transform in the form of a line spectrum with the distance $\omega_0 = \frac{2\pi}{T}$ between the components.

1.10 Time Duration and Band Width

1. Similarity principle:

$$F\{h(\alpha t)\} \;=\; \frac{1}{|\alpha|}\;H(\frac{\omega}{\alpha})$$



time duration T band width B

$$T \;\cdot\; B \quad = \quad const.$$

- High resolution in the time domain results in low resolution in the frequency domain and vice versa

2. Special case: $h(t)$ with

$$Im\{H(\omega)\} \quad = \quad 0 \quad (\ h(t) \text{ symmetrical }\ )$$

and

$$Re\{H(\omega)\} \quad \geq \quad 0$$

$h(t)$ has maximum for $t = 0$:

$$h(t) \ = \ \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} H(\omega)\ \cos(\omega t)\ d\omega \ \leq \ \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} H(\omega)\ d\omega \ = \ h(0)$$

define:

$$T \quad = \quad \frac{1}{h(0)} \int\limits_{-\infty}^{\infty} h(t)\ dt$$

$$B \quad = \quad \frac{1}{H(0)} \int\limits_{-\infty}^{\infty} H(\omega)\ d\omega$$

from

$$T \quad = \quad \frac{H(0)}{h(0)} \quad \text{and} \quad B \ = \ 2\,\pi\,\frac{h(0)}{H(0)}$$

follows

$$T \cdot B \quad = \quad 2\,\pi$$

3. In general: normalized impulse $h(t) \in \mathbb{R}$ with

$$\int_{-\infty}^{\infty} h^2(t)\, dt \;=\; 1, \qquad h(t) \in \mathbb{R}$$

$$T^2 \;:=\; \int_{-\infty}^{\infty} h^2(t)\, t^2\, dt$$

$$B^2 \;:=\; \int_{-\infty}^{\infty} |\, H(\omega)\,|^2\, \omega^2\, d\omega \qquad \left(=\; 2\,\pi \int_{-\infty}^{\infty} \left[h'(t)\right]^2\; dt \right)$$

- Results in "uncertainty relation":

$$T \;\cdot\; B \;\geq\; \sqrt{\frac{\pi}{2}}$$

- Proof: Cauchy-Schwarz inequality

$$|\; x^T\, y\;| \;\leq ||x||\; \cdot ||y||$$

$$\underbrace{\left| \int_{-\infty}^{\infty} [\; t\; h(t)\;]\;\; h'(t)\; dt \right|^2}_{=\frac{1}{4}} \;\leq\; \underbrace{\int_{-\infty}^{\infty} [\; t\; h(t)]^2\;\; dt}_{=T^2}\; \underbrace{\int_{-\infty}^{\infty} [\; h'(t)\;]^2\;\; dt}_{\frac{B^2}{2\pi}}$$

From: partial integration

$$\int u'(t)\; v(t)\; dt \;=\; u(t)\; v(t) - \int u(t)\; v'(t)\; dt$$

$$\int \underbrace{[\; h(t)\; h'(t)\;]}_{u'(t)} \underbrace{t}_{v(t)} dt \;=\; \frac{1}{2}\, h(t)^2\, t \;-\; \int \frac{1}{2}\, h^2(t)\; 1\; dt$$

$$\int_{-\infty}^{\infty} [\; h(t)\; h'(t)\;]\;\; t\; dt \;=\; 0 \;-\; \frac{1}{2}$$

Equality sign is valid for linear dependency:

$$\begin{aligned}
h'(t) &= -\lambda\, t\, h(t) \\
\frac{dh}{h} &= -\lambda\, t\, dt \\
\log(h) &= -\frac{1}{2}\,\lambda\, t^2 \;+\; \text{const.}, \qquad \lambda \;>\; 0
\end{aligned}$$

$$\Longrightarrow \qquad \text{Optimum} \quad T \cdot B \;=\; \sqrt{\frac{\pi}{2}} \quad \text{for Gauss' impulse}$$

$$h(t) \;=\; \frac{\lambda}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\lambda\, t^2}$$

$$\text{Variance:} \quad \sigma^2 = \frac{1}{\lambda}$$

Quantum Physics: similar statement about position and impulse of a particle

4. Finite positive signal

$$g(t) \quad \begin{cases} \geq 0 & 0 \leq t \leq T \\ = 0 & t < 0 \text{ or } T < t \end{cases}$$

The following is valid for the amplitude spectrum $|G(\omega)|$:

$$\begin{aligned} |G(\omega)| &= \left| \int_{-\infty}^{+\infty} g(t)\, e^{-j\omega t} dt \right| \\ &\leq \int_{-\infty}^{+\infty} |g(t)|\, |e^{-j\omega t}| dt \\ &= \int_{-\infty}^{+\infty} |g(t)|\, dt \\ &\quad \text{because } g(t) \geq 0 \\ &= G(0) \end{aligned}$$

Define the band width ω_B as:

$$|G(\omega_B)|^2 = \frac{G^2(0)}{2} \quad \text{and} \quad |G(\omega_B)|^2 \leq |G(\omega)|^2 \ \text{ for } \ |\omega| < \omega_B$$

Then:

$$T\, \omega_B \geq \frac{\pi}{2}$$

Proof:

The following inequalities are valid:

$$\begin{aligned} a^2 + b^2 &\geq \frac{(a-b)^2}{2} && \forall\, a, b \in \mathbb{R} \\ |\sin\alpha| + |\cos\alpha| &\geq 1 && \forall\, \alpha \in \mathbb{R} \end{aligned}$$

For the Fourier-Transform of $g(t)$ holds:

$$\begin{aligned} Re\{G(\omega)\} &= \int_0^T g(t)\ \cos\omega t\ dt \\ Im\{G(\omega)\} &= -\int_0^T g(t)\ \sin\omega t\ dt \end{aligned}$$

For $0 \leq \omega\, t \leq \frac{\pi}{2}$ holds: $\quad \cos\omega t \geq 0, \quad \sin\omega t \geq 0$
and therefore: $\quad \cos\omega t + \sin\omega t = |\cos\omega t| + |\sin\omega t| \geq 1$

$$\begin{aligned} Re\{G(\omega)\} - Im\{G(\omega)\} &= \int_0^T g(t)\ [\cos\omega t + \sin\omega t]\ dt \\ &\geq \int_0^T g(t)\ 1\ dt \\ &= G(0) \end{aligned}$$

$$\begin{aligned} |G(\omega)|^2 &= Re^2\{G(\omega)\} + Im^2\{G(\omega)\} \\ &\geq \frac{[Re\{G(\omega)\} - Im\{G(\omega)\}]^2}{2} \\ &\geq \frac{1}{2}\, G^2(0) \ \equiv\ |G(\omega_B)|^2 \end{aligned}$$

Chapter 2

Discrete Time Systems

- Overview:

2.1 Motivation and Goal

If we want to process a continuous time signal $x(t)$ with a computer, we have to sample it at discrete equidistant time points

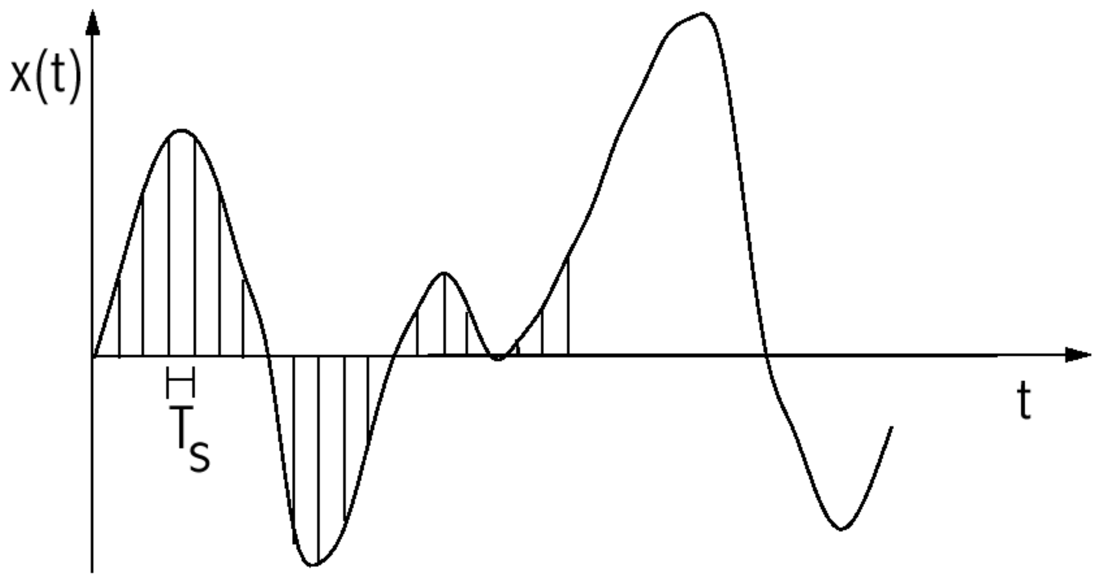
$$t_n = n \cdot T_S$$

where T_S is called sampling period.

Terminology:
"time discrete" is often called "digital", where this adjective often
(but not always) denotes the amplitude quantization,
i.e. the quantization of the value $x(n \cdot T_S)$.

Advantages of digital processing in comparison to analog components:

- independent of analog components and technical difficulties with respect to their realization;
- in principle arbitrary high accuracy;
- also non-linear methods are possible,
  in principle even every mathematical method.



2.2 Digital Simulation using Discrete Time Systems

Task definition:

- Given:
  Analog system with input signal $x(t)$ and output signal $y(t)$;
  Sampling with sampling period T_S
- Wanted:
  Discrete System with input signal $x[n]$ and output signal $y[n]$, such that

  $$x[n] \;=\; x(nT_S)$$

  results in

  $$y[n] \;=\; y(nT_S)$$

- For which signals is such a digital simulation possible?
- The sampling theorem gives (most of) the answer.

LTI System (analog to continuous time case):

- Linearity:

  Homogeneity:

$$S\{\alpha\, x[n]\} \;=\; \alpha\, S\{x[n]\}$$

  Additivity:

$$S\{x_1[n] \;+\; x_2[n]\} \;=\; S\{x_1[n]\} \;+\; S\{x_2[n]\}$$

- Shift invariance:

$$S\{x[n \;-\; n_0]\} \;=\; y[n - n_0], \qquad n_0 \quad \text{whole number}$$

Representation of an LTI System as discrete convolution:

Unit impulse:

$$\delta[n] \;=\; \begin{cases} 1, & n \;=\; 0 \\ 0, & n \;\neq\; 0 \end{cases}$$

The signal $x[n]$ is represented with amplitude weighted and time shifted unit impulses $\delta[n]$. The system reacts on $\delta[n]$ with $h[n]$:

$$h[n] \;=\; S\{\delta[n]\}$$

Input signal:

$$x[n] \;=\; \sum_{k=-\infty}^{\infty} x[k]\,\delta[n-k]$$

Output signal:

$$y[n] \;=\; S\left\{\sum_{k=-\infty}^{\infty} x[k]\,\delta[n-k]\right\}$$

Additivity

$$= \sum_{k=-\infty}^{\infty} S\{\,x[k]\,\delta[n-k]\,\}$$

Homogeneity

$$= \sum_{k=-\infty}^{\infty} x[k]\,S\{\,\delta[n-k]\,\}$$

Time invariance

$$= \sum_{k=-\infty}^{\infty} x[k]\,h[n-k]$$

Input signal $x[n]$ and output signal $y[n]$ of a discrete time LTI system are linked through discrete convolution.
$h[n]$ is called impulse response like in continuous time case.

2.3 Examples of Discrete Time Systems

- Difference calculation:

$$y[n] \; = \; x[n] \; - \; x[n-n_0]$$

- "1-2-1"-averaging:

$$y[n] \; = 0.5 \cdot x[n-1] + \; x[n] \; + 0.5 \cdot x[n+1]$$

- sliding window averaging ("smoothing")

$$y[n] \; = \; \frac{1}{2M+1} \sum_{k=-M}^{M} x[n-k]$$

- weighted averaging: instead of constant weight

$$h[n] \; = \; \frac{1}{2M+1}$$

  arbitrary weights can be used:

$$y[n] \; = \; \sum_{k=-M}^{M} h[k] \cdot x[n-k]$$

  Note: the only difference from general case is finite length of the convolution kernel $h[n]$.

- First order difference equation:
  (recursive averaging, averaging with memory)

$$y[n] - \alpha \, y[n-1] \; = x[n]$$

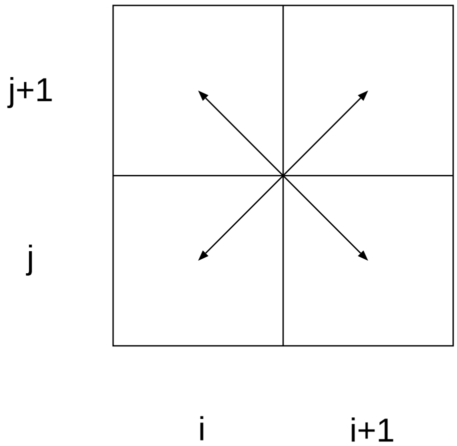
- (Digital) resonator (second order difference equation)

$$y[n] - \alpha \, y[n-1] - \beta \, y[n-2] = x[n]$$

- Image processing:
  Gradient calculation and image enhancement
  (Roberts Operator, Laplace Operator)

- Roberts Cross Operator

  gray values $x[i,j]$



2

$$|\nabla x[i,j]|^2 \;=\; (x[i,j]-x[i+1,j+1])^2+(x[i,j+1]-x[i+1,j])^2$$

Note: <u>non-linear operation</u>

simplified:

$$|\nabla x[i,j]| \;=\; |x[i,j]-x[i+1,j+1]|+|x[i,j+1]-x[i+1,j]|$$

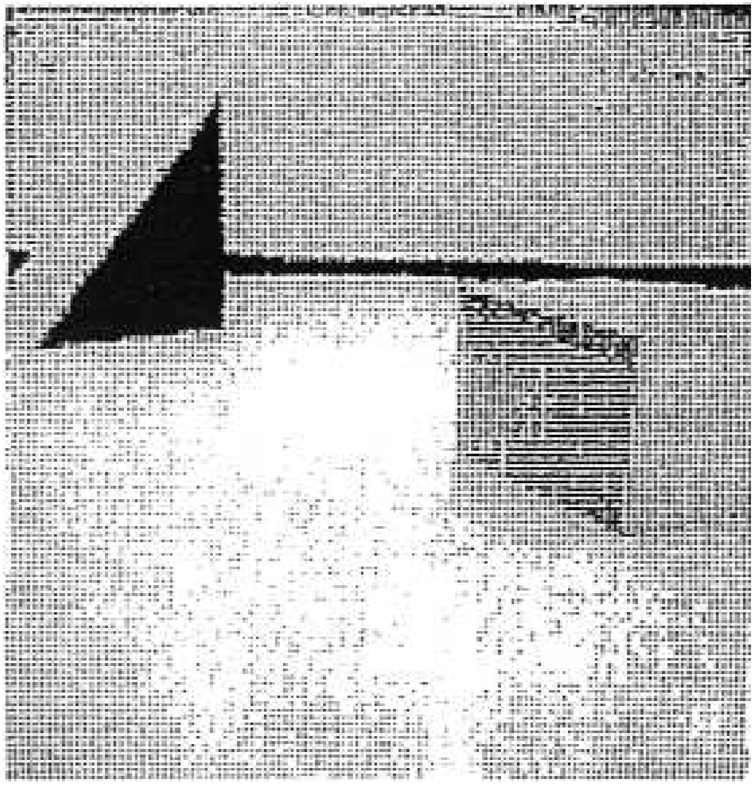


Figure 2.1: Digital photo

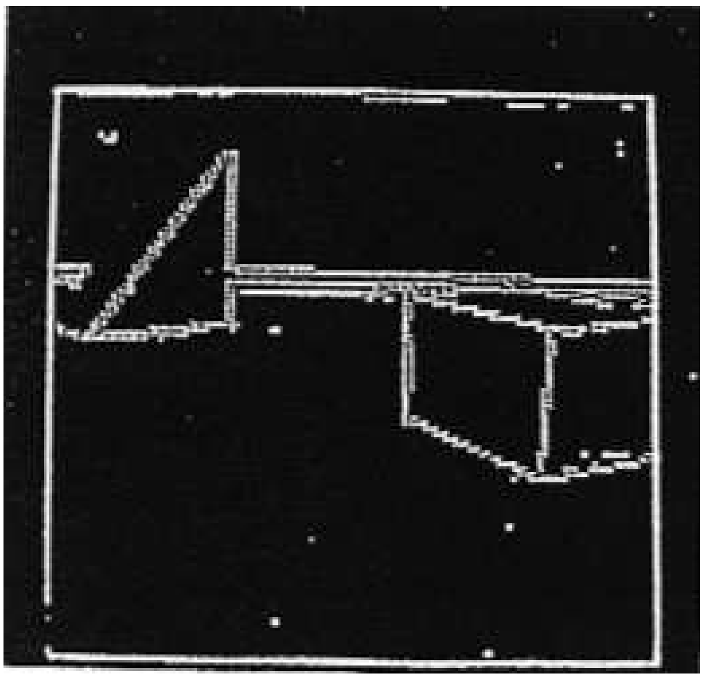


Figure 2.2: Gradient image

- Laplace Operator $\equiv$ discrete approximation of the second derivation

$$
\begin{aligned}
\nabla^2 x[i,j] &= \triangle_i^2 x[i,j] + \triangle_j^2 x[i,j] \\
&= x[i+1,j] - 2x[i,j] + x[i-1,j] + \\
&\quad x[i,j+1] - 2x[i,j] + x[i,j-1] \\
&= x[i+1,j] + x[i-1,j] + x[i,j+1] + x[i,j-1] - 4x[i,j]
\end{aligned}
$$

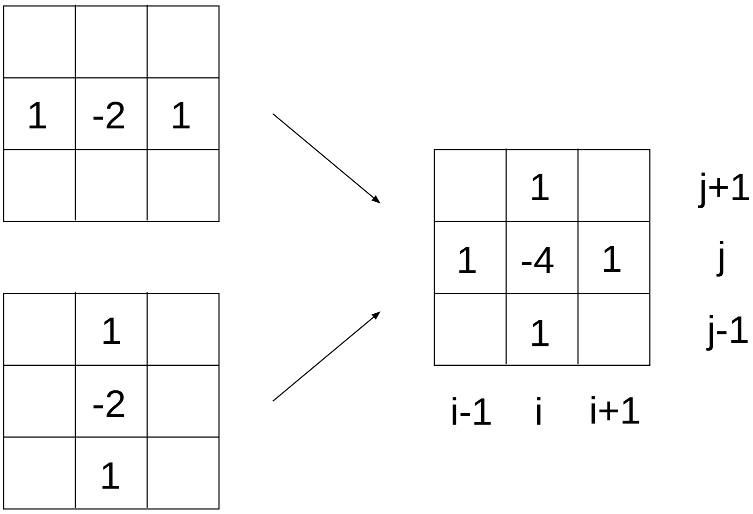


Image enhancement:

$$
\begin{aligned}
y[i,j] &= x[i,j] - \nabla^2 x[i,j] \\
&= h[i,j] * x[i,j]
\end{aligned}
$$

$$
\begin{array}{ccc|c}
 & -1 & & j+1 \\
-1 & +5 & -1 & j \\
 & -1 & & j-1 \\
\hline
i-1 & i & i+1 &
\end{array}
$$

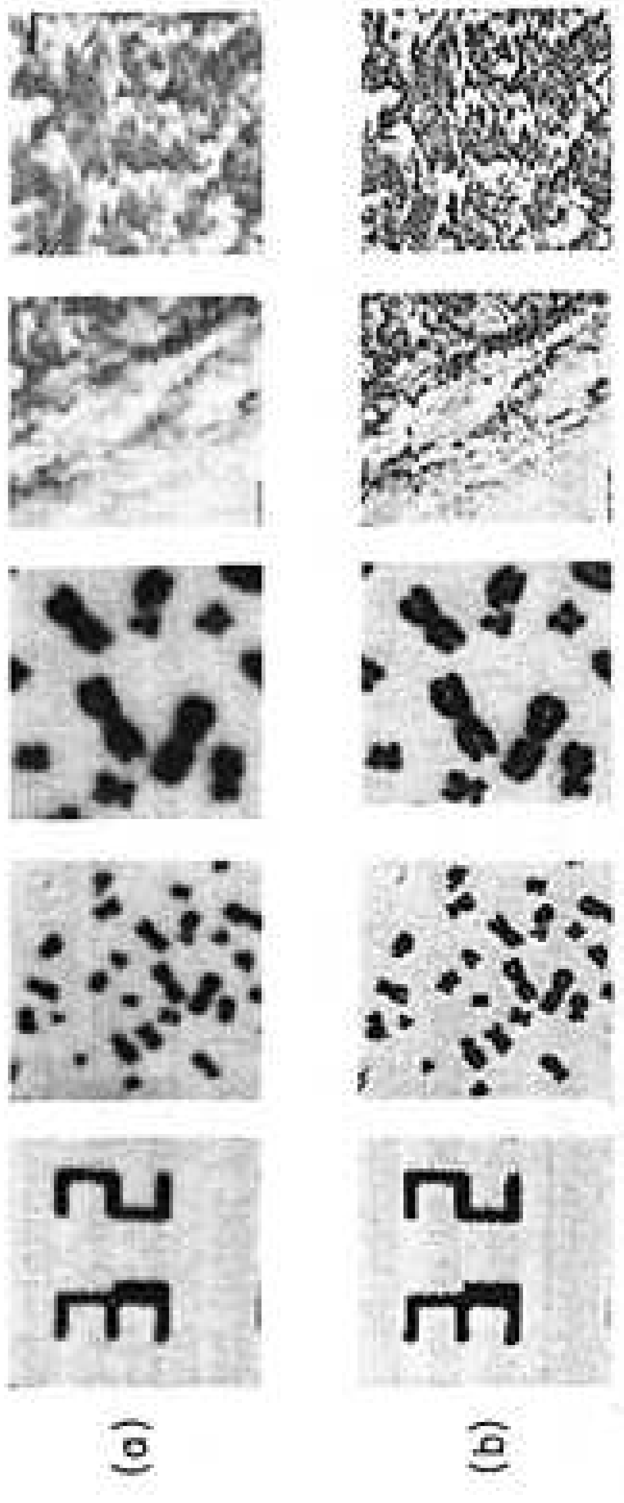


Figure 2.3: Several real cases of Laplace Operator subtraction from original image. a) Original image b) Original image minus Laplace Operator (negative values are set to 0 and values above the grey scale are set to the highest grade of grey)

2.4 Sampling Theorem (Nyquist Theorem) and Reconstruction

The following will be analyzed and derived respectively:

How should we choose the sampling period T_S, if we want to represent a continuous signal $x(t)$ with its sample values $x(nT_S)$ so that the signal $x(t)$ can be exactly reconstructed from its sample values?

- Fourier transform of the continuous time signal $x(t)$:

$$\begin{aligned} X(\omega) &= F\{\, x(t) \,\} = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt \\ x(t) &= F^{-1}\{\, X(\omega) \,\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t}\, d\omega \end{aligned} \tag{2.1}$$

- Signal $x(t)$ has limited bandwidth with upper limit ω_B, which means:
  $X(\omega) \;=\; 0 \qquad \text{for all} \quad |\omega| \;\geq \omega_B$
  Note: $X(\omega_B) = 0$

- $X(\omega)$ in domain $-\omega_B < \omega < \omega_B$ can be represented as *Fourier Series*:

$$X(\omega) = \sum_{n=-\infty}^{\infty} a_n \exp(-jn\pi \frac{\omega}{\omega_B}) \tag{2.2}$$

- The coefficients a_n are given by:

$$a_n = \frac{1}{2\omega_B} \int_{-\omega_B}^{\omega_B} X(\omega) \exp(jn\pi \frac{\omega}{\omega_B})\, d\omega \tag{2.3}$$

- Comparison of the equations (2.1) and (2.3) shows that the coefficients a_n are given by the values of the inverse Fourier transform of $x(t)$ at points

$$t_n = \frac{n\pi}{\omega_B} \tag{2.4}$$

  The band limitation of $X(\omega)$ has to be considered for the integration limits in (2.1). Result:

$$a_n = x(\frac{n\pi}{\omega_B}) \cdot \frac{\pi}{\omega_B} \tag{2.5}$$

- Inserting Eq. (2.5) into Eq. (2.2) and then in Eq. (2.1) results in:

$$x(t) \;=\; \frac{1}{2\pi} \int\limits_{-\omega_B}^{\omega_B} \frac{\pi}{\omega_B} \sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\omega_B}) \exp(-jn\pi\frac{\omega}{\omega_B}) \exp(j\omega t)\, d\omega$$

- After swapping summation and integration and subsequent integration:

$$x(t) \;=\; \sum_{n=-\infty}^{\infty} x(\frac{n\pi}{\omega_B}) \frac{\sin(\omega_B\,(t-\frac{n\pi}{\omega_B}))}{\omega_B\,(t-\frac{n\pi}{\omega_B})}$$

- Reconstruction of the signal $x(t)$ from sample values is possible if equidistant sample values $x(\frac{n\pi}{\omega_B}) \;=\; x(n\cdot T_s)$ have the distance T_S

$$T_S \;=\; \frac{\pi}{\omega_B} \tag{2.6}$$

- The sampling period T_S corresponds to the sampling frequency Ω_S:

$$\Omega_S \;=\; \frac{2\pi}{T_S}$$

- Equation (2.6) shows that if the sampling frequency is

$$\Omega_S \;:=\; 2\,\omega_B$$

  the original signal $x(t)$ can be reconstructed exactly.

- In the Fourier series representation of $X(\omega)$ in equation (2.2), the period $2\cdot\omega_B$ has been supposed.
  ω_B is the highest frequency component of the signal $x(t)$.

- Since $X(\omega)$ is equal to zero for $|\omega| \geq \omega_B$, the period $2 \cdot \omega_B$ can be substituted with every period $2 \cdot \widetilde{\omega}_B$ where $\widetilde{\omega}_B \geq \omega_B$. The previous derivation is also valid for this $\widetilde{\omega}_B$.

- When

$$\widetilde{\omega}_B = \frac{\pi}{T_S}$$

then:

$$x(t) = \sum_{n=-\infty}^{\infty} x(n\,T_S)\,\frac{\sin(\pi\,(t - n\,T_S)/T_S)}{\pi\,(t - n\,T_S)/T_S}$$

(reconstruction formula)

Note: $\lim_{t \to 0} \frac{\sin(t)}{t} = 1$ (l'Hôpital's rule)

- The condition $\widetilde{\omega}_B \geq \omega_B$ results in:

$$T_S \leq \frac{\pi}{\omega_B} \tag{2.7}$$

for the sampling period T_S and in:

$$\Omega_S \geq 2 \cdot \omega_B \tag{2.8}$$

for the sampling frequency Ω_S.

- The equations (2.7) and (2.8) are denoted as sampling theorem. The sampling frequency has to be at least twice as high as the upper limit frequency of the signal ω_B where $X(\omega) = 0$ for $|\omega| \geq \omega_B$. If and only if this condition is satisfied, an exact reconstruction (without approximation) of a continuous signal $x(t)$ from its sample values $x(nT_S)$ is possible.

- Note: The sampling frequency $\Omega_S = 2 \cdot \omega_B$ is also called *Nyquist frequency.*

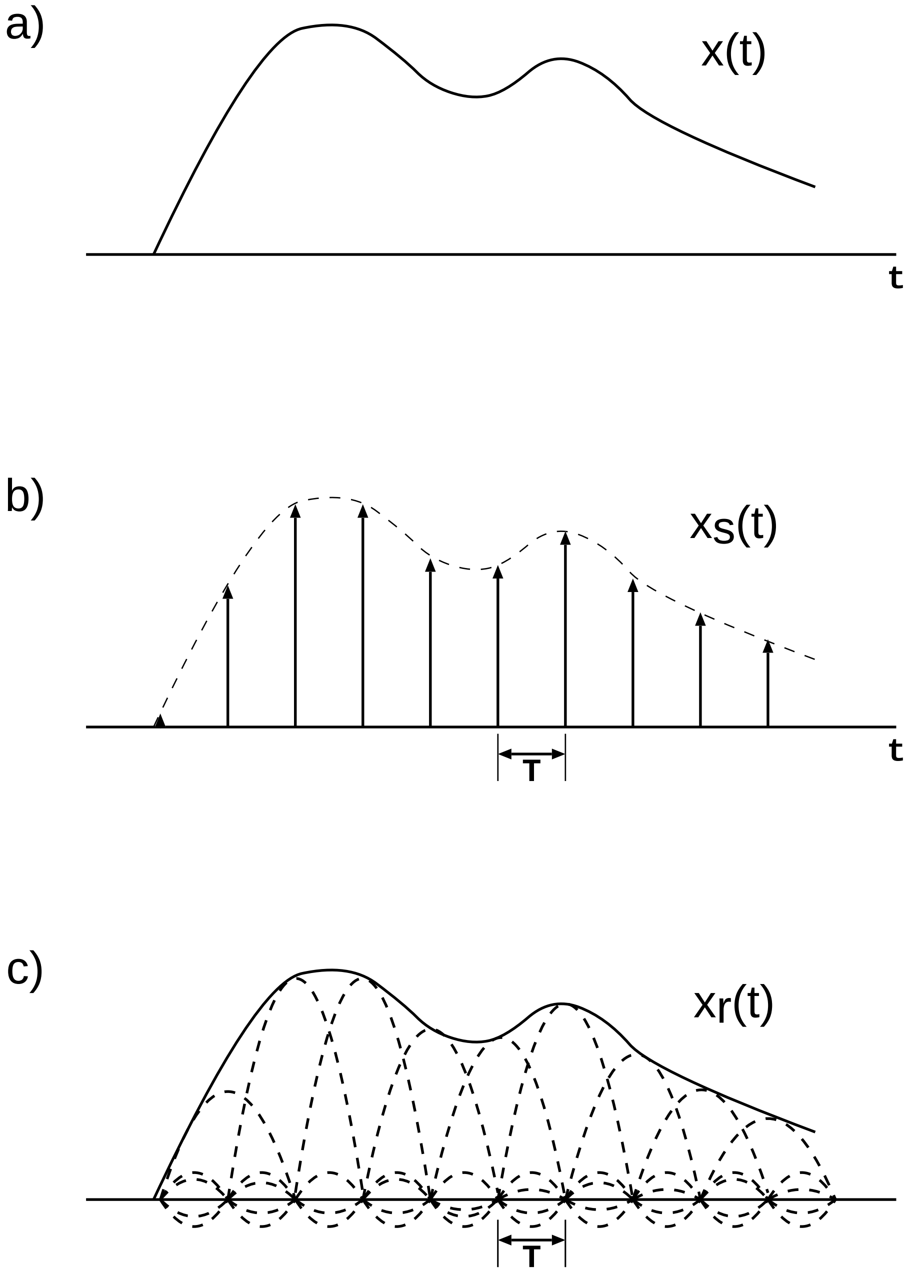


Figure 2.4: Ideal reconstruction of a band-limited signal (from Oppenheim, Schafer)
a) original signal b) sampled signal c) reconstructed signal

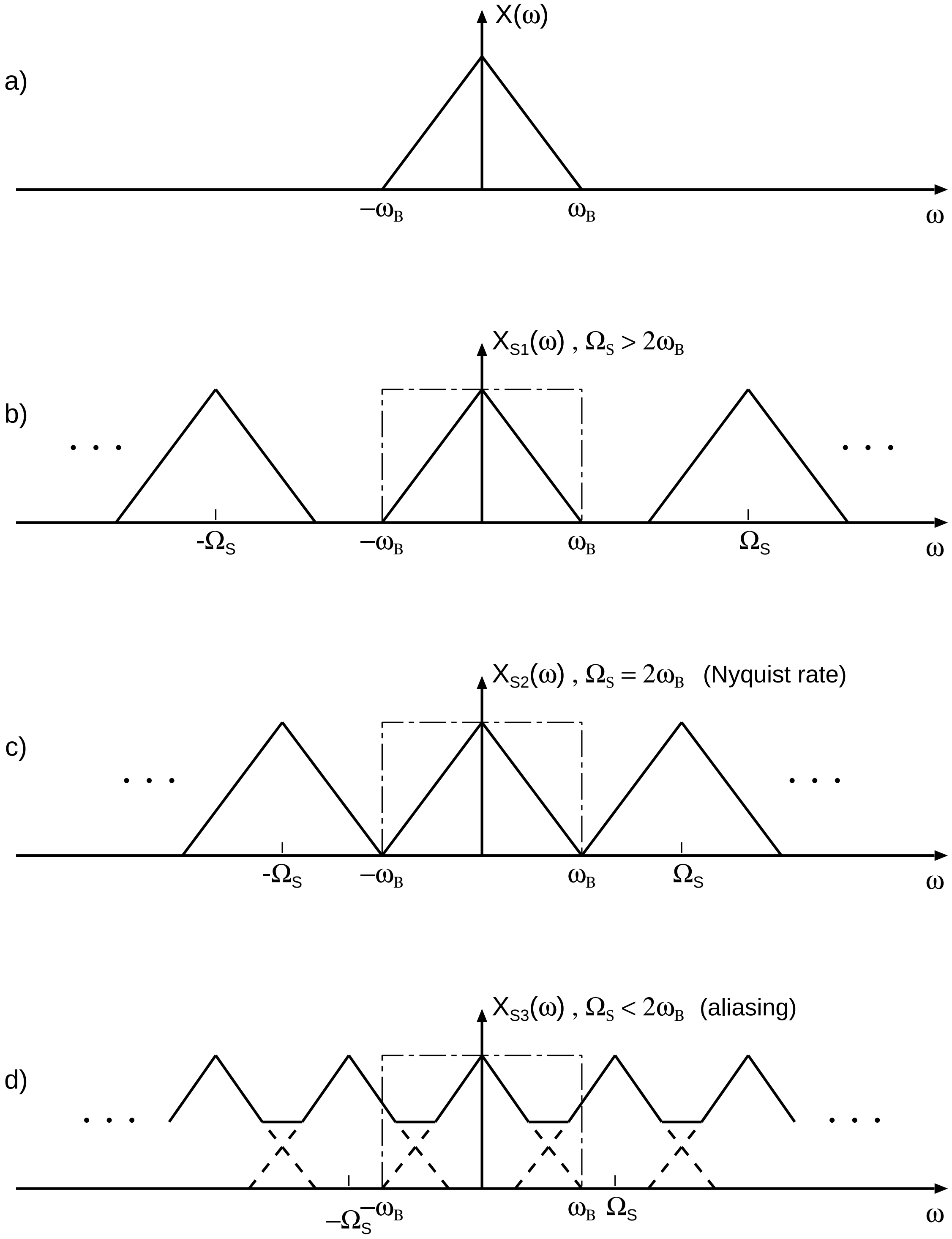


Figure 2.5: Sampling of band-limited signal with different sampling rates:
b) sampling rate higher than Nyquist rate - exact reconstruction possible
c) sampling rate equal to Nyquist rate - exact reconstruction possible
d) sampling rate smaller than Nyquist rate - aliasing - exact reconstruction not possible

Another proof using delta- and comb-function:

Sampling of the continuous signal $x(t)$ with $\Omega_S = \frac{2\pi}{T_S}$

- Band limitation: $X(\omega) = 0$ for $|\omega| \geq \omega_B$

  (always possible: analog to low-pass with $T(\omega) = 0$ for $|\omega| \geq \omega_B$)

- Sampling procedure

  Multiplication of a function with a comb-function in time domain

$$x_s(t) = T_s\, x(t) \cdot \sum_{n=-\infty}^{+\infty} \delta(t - nT_s)$$

  results in a convolution with a comb-function in frequency domain:

$$\begin{aligned} X_s(\omega) &= T_s \cdot \frac{1}{2\pi} X(\omega) * \frac{2\pi}{T_s} \sum_{n=-\infty}^{+\infty} \delta\left(\omega - \tilde{\omega}\frac{2\pi n}{T_s}\right) \\ &= \int_{-\infty}^{+\infty} X(\tilde{\omega}) \sum_{n=-\infty}^{+\infty} \delta\left(\omega - \tilde{\omega}\frac{2\pi n}{T_s}\right) d\tilde{\omega} \\ &= \sum_{n=-\infty}^{+\infty} X\left(\omega - n\frac{2\pi}{T_s}\right) \end{aligned}$$

  $\Longrightarrow$ sampled signal has periodical Fourier spectrum

  (Analogy to Fourier series: periodical signal has line spectrum, i.e. discrete spectrum)

  No overlap if:

$$\begin{aligned} \omega_B &\leq \Omega_S - \omega_B \\ 2\omega_B &\leq \Omega_S \end{aligned}$$

- In so-called digital simulation, the signal $x(t)$ is represented by its sampled values $x(n \cdot T_S)$ measured at equidistant time points with distance T_S. With a proper sampling period T_S an exact reconstruction of the signal $x(t)$ from the sampled values $x(n \cdot T_S)$ is possible.

- If it is possible to exactly reconstruct the signal $x(t)$ from the sampled values $x(n \cdot T_S)$, then it is possible to perform a discrete time processing of the sampled values $x(n \cdot T_S)$ on a computer, which is equivalent to the continuous time processing of the signal $x(t)$ (digital simulation).

- Continuous time processing:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)\, h(t-\tau)\, d\tau$$

- Discrete time processing:
  - Sampling period T_S
  - $x[n] := x(nT_S)$

$$\begin{aligned} y(nT_S) &= \sum_{k=-\infty}^{\infty} x(kT_S)\, h(nT_S - kT_S)\, T_S, \qquad \tilde{h}[n] = h(nT_S) \\ y[n] &= \sum_{k=-\infty}^{\infty} x[k]\, \tilde{h}[n-k] \end{aligned}$$

- As a result of the convolution theorem (convolution in time domain corresponds to multiplication in frequency domain), the band limited input signal gives an also band limited output signal which is exactly determined by its sampled values.

Important:

- In the domain $|\omega| < \Omega_S/2$ the Fourier transform of a continuous time signal $x(t)$ is identical with the Fourier–transform of the corresponding sampled discrete time signal $x(nT_S)$:

$$
\begin{aligned}
X(\omega) &= \int_{-\infty}^{\infty} x(t)\,\exp(-j\omega t)\,dt \\
&\quad \text{for} \quad |\omega| \leq \Omega_S/2 \quad \text{is identical to} \\
T_S \cdot X_S(\omega) &= T_S \cdot \sum_{n=-\infty}^{\infty} x(nT_S)\,\exp(-j\omega T_S n) \\
&= T_S \cdot \sum_{n=-\infty}^{\infty} x(nT_S)\,\exp(-j\frac{2\pi\omega}{\Omega_S}n)
\end{aligned}
$$

- Inverse Fourier transform of discrete time signal:

$$
x(nT_S) = \frac{1}{\Omega_S} \int_{-\Omega_S/2}^{\Omega_S/2} X_S(\omega)\,\exp(j\omega T_S n)\,d\omega
$$

- One period:

$$
\begin{aligned}
-\frac{\Omega_S}{2} \;\leq\; \omega \;\leq\; \frac{\Omega_S}{2} \\
-\pi \;\leq\; \frac{2\pi\omega}{\Omega_S} \;\leq\; \pi
\end{aligned}
$$

- The Fourier transform of a discrete time signal is periodic in ω with the period $2\,\pi/T_S = \Omega_S$.

- The Fourier transform of a discrete time signal is continuous in ω.

Frequency normalization

- Define the normalized frequency ω_N:

$$\omega_N \ := \ 2\pi \frac{\omega}{\Omega_S}$$

- Definition: (ω now denotes a normalized frequency)

  – Fourier transform of discrete time signal $x[n]$:

$$X(e^{j\omega}) \ = \ \sum_{n=-\infty}^{+\infty} x[n] \exp(-j\omega n)$$

  Note the notation $X(e^{j\omega})$.

  – Inverse Fourier transform of discrete time signal $x[n]$:

$$x[n] \ = \ \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) \exp(j\omega n) \, d\omega$$

2.5 Logarithmic Scale and dB

Why?

– large dynamic range for the amplitude values of a signal

$$x(t) \;=\; A \cos \beta t$$

$A :=$ amplitude
(pressure, velocity, inclination, current, voltage, ...)
"linear" variable

$A_0 :=$ reference amplitude
predefined value for calibration

$dB :=$ "decibel"

$$\begin{aligned} A[dB] \;&\equiv\; 20 \cdot \lg \frac{A}{A_0}, \qquad \lg \equiv \log_{10} \\ &=\; 10 \cdot \lg \frac{A^2}{A_0^2}, \qquad A^2 = \text{ quadratic variable} = \text{energy, intensity} \end{aligned}$$

because of $2^{10} = 1024 \cong 10^3$:

$$1 \text{ bit more} \;\hat{=}\; \begin{cases} \text{factor 2 for amplitude} \;\hat{=}\; 6 \text{ dB} \\ = \text{factor 4 for intensity} \end{cases}$$

3 dB $\hat{=}$ factor 2 for intensity

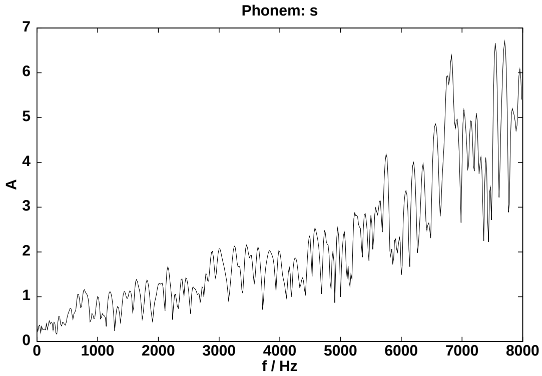


Figure 2.6: Amplitude spectrum of the voiceless phoneme "s" from the word "ist"

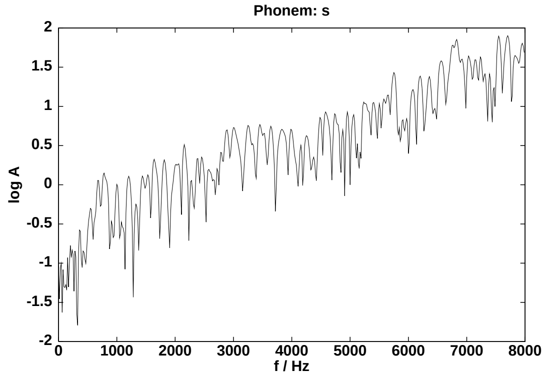


Figure 2.7: Logarithmic amplitude spectrum of the phoneme "s"

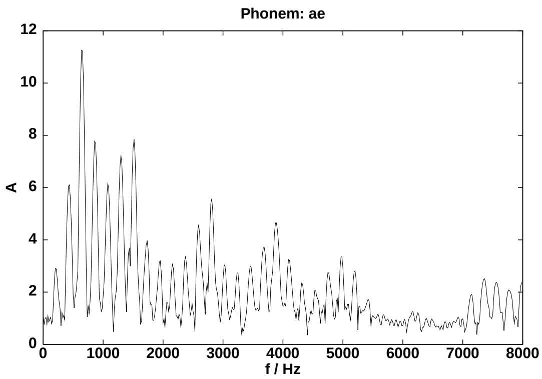


Figure 2.8: Amplitude spectrum of the voiced phoneme "ae" from the word "Äh"

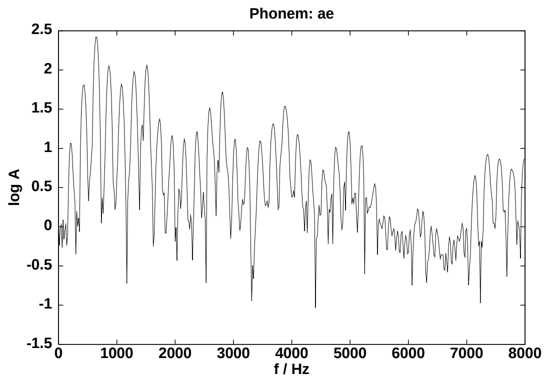


Figure 2.9: Logarithmic amplitude spectrum of the phoneme "ae"

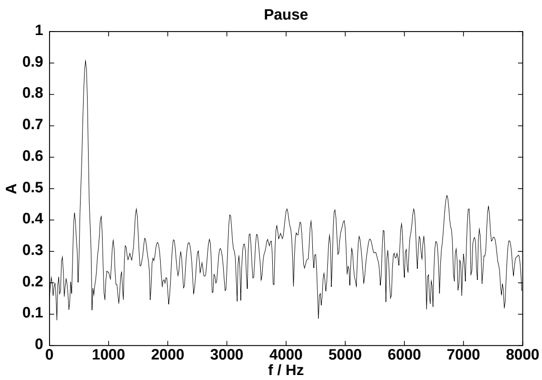


Figure 2.10: Amplitude spectrum of a speech pause

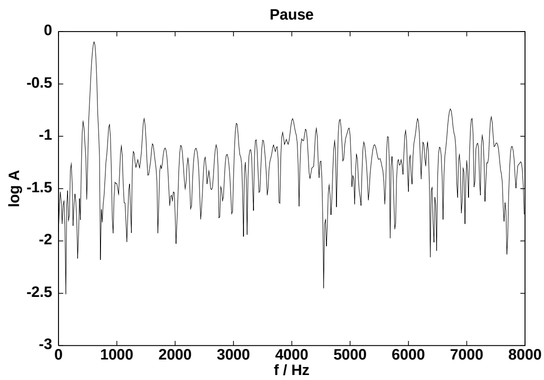
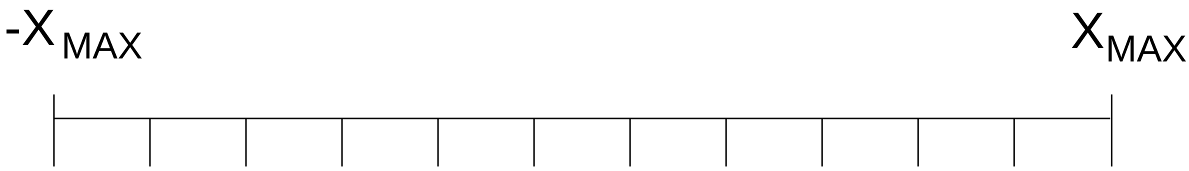


Figure 2.11: Logarithmic amplitude spectrum of a speech pause

2.6 Quantization

- Uniform quantization



- Quantisation: $\hat{x} = Q(x)$
- B bits correspond to 2^B quantisation levels
- Boundaries: $x_0, x_1, \ldots, x_k, \ldots, x_K$ where $K = 2^B$
- Width $\triangle$ of one quantisation level using uniform quantisation:

$$\triangle = \frac{2 \cdot X_{MAX}}{2^B}$$

- Quantisation error:

$$\sigma_e^2 = \int_{-\infty}^{+\infty} (x - \hat{x})^2 \, p(x) \, dx = \sum_{k=1}^{K} \int_{x_{k-1}}^{x_k} (x - \hat{x}_k)^2 \, p(x) \, dx$$

– for uniform quantisation:

a) $x_k - x_{k-1} = \triangle = const(k)$

b) $\hat{x}_k = \frac{1}{2}(x_{k-1} + x_k)$

– uniform distribution with $p(x) = const(x)$ results in:

$$\sigma_e^2 = \sum_k \frac{\triangle^2}{12} \cdot \frac{1}{K} = \frac{\triangle^2}{12} = \frac{X_{MAX}^2}{3 \cdot 2^{2B}}$$

- signal-to-noise ratio in dB (general definition):

$$SNR[dB] \ := \ 10 \lg \frac{\sigma_x^2}{\sigma_n^2}$$

σ_x^2 = power of the signal x
σ_n^2 = power of the noise n

SNR = signal-to-noise ratio

- signal-to-quantisation noise ratio (special case):

$$SNR[dB] \ := \ 10 \lg \frac{\sigma_x^2}{\sigma_e^2}$$

σ_e^2 = power of the noise caused by quantisation errors

- uniform quantisation using B bits:

$$SNR[dB] \ = \ 6.02\ B + 4.77 - 20 \lg \frac{X_{MAX}}{\sigma_x}$$

- if signal amplitude has Gaussian distribution, only 0.064% of samples have amplitude greater than $4\sigma_x$:

$$SNR[dB] \ = \ 6.02\ B - 7.2 \qquad \text{for } X_{MAX} = 4\sigma_x$$

2.7 Fourier Transform and z–Transform

Transfer function and Fourier transform

Eigenfunctions of discrete linear time invariant systems (analog to time continuous case):

$$x[n] \;=\; e^{j\,\omega\,n} \qquad -\infty \;<\; n \;<\; \infty$$

(ω is dimensionless here)

Proof:

$$\begin{aligned} x[n] &= e^{j\,\omega\,n} \\ y[n] &= \sum_{k=-\infty}^{\infty} h[k]\; e^{j\,\omega\,(n-k)} \\ &= e^{j\,\omega\,n} \sum_{k=-\infty}^{\infty} h[k]\; e^{-j\,\omega\,k} \end{aligned}$$

Define: $\quad H(e^{j\,\omega}) \;=\; \sum_{k=-\infty}^{\infty} h[k]\; e^{-j\,\omega\,k}$

Remark:
The Fourier transform of a discrete time signal is already introduced as Fourier series during the derivation of sampling theorem and reconstruction formula (equation (2.2)).

Result: $\quad y[n] \;=\; e^{j\,\omega\,n}\; H(e^{j\,\omega})$

z–transform:

- Fourier transform of a discrete time signal: $x[n]$

$$X(e^{j\omega}) \;=\; \sum_{n=-\infty}^{+\infty} x[n]\; e^{-j\omega n}$$

  – periodic in ω

  – ω is normalized frequency, thence:

$$-\pi \;<\; \omega \;\leq\; \pi$$

  – X is evaluated on the unit circle ($e^{j\omega}$)

- Generalization: X is evaluated for any complex values z.

- That results in z–transform:

$$X(z) = \sum_{n=-\infty}^{+\infty} x[n]\; z^{-n}$$

- Reasons for z–transform

  1. analytically simpler, function theory methods are applicable
  2. better handling of convergence problem:
     – convergence of finite signal, i.e. $x[n] = 0$ for each $n > N_0$
     – convergence of infinite signal depends on z

- Inverse z–transform:

$$x[n] = \frac{1}{2\pi j} \oint X(z)\; z^{n-1}\; dz$$

formally: $z = e^{j\omega}$ $\qquad dz = jzd\omega$

$$x[n] = \frac{1}{2\pi} \int_{0}^{2\pi} X(e^{j\omega})\; e^{j\omega n}\; d\omega$$

Example of Fourier transform and z–transform:

- "Truncated geometric series"

$$x[n] = \begin{cases} a^n & 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases}$$

- z–transform

$$\begin{aligned} X(z) &= \sum_{n=0}^{N-1} a^n \, z^{-n} = \sum_{n=0}^{N-1} (a \, z^{-1})^n = \frac{1-(a \, z^{-1})^N}{1 - a \, z^{-1}} \\ &= \frac{1}{z^{N-1}} \frac{z^N - a^N}{z - a} \end{aligned}$$

- Fourier transform

  z–transform results in Fourier transformation using substitution

$$z = e^{j\omega}$$

$$X(e^{j\omega}) = \frac{1 - a^N \, e^{-j\omega N}}{1 - a \, e^{-j\omega}}$$

special case for $a = 1$ (discrete time rectangle):

$$= \exp\left(-j\frac{\omega(N-1)}{2}\right) \frac{\sin\left(\dfrac{\omega N}{2}\right)}{\sin\left(\dfrac{\omega}{2}\right)}$$

Proof for the z–transform inversion

- Statement:

$$x[k] \;=\; \frac{1}{2\pi j} \oint X(z)\, z^{k-1}\, dz$$

- Cauchy integration rule

$$\begin{aligned}
\frac{1}{2\pi j} \oint z^{-k} dz \;&=\; \begin{cases} 1 & k = 1 \\ 0 & k \neq 1 \end{cases} \\
\frac{1}{2\pi j} \oint X(z)\, z^{k-1} dz \;&=\; \frac{1}{2\pi j} \oint \sum_n x[n]\, z^{-n+k-1} dz \\
&=\; \sum_n x[n] \underbrace{\frac{1}{2\pi j} \oint z^{-n+k-1} dz}_{\neq 0 \text{ only for } n = k} \\
&=\; x[k]
\end{aligned}$$

- Fourier:

$$z \;=\; e^{j\omega} \qquad \Longrightarrow \qquad dz \;=\; j\, e^{j\omega}\, d\omega$$

Then:

$$x[n] \;=\; \frac{1}{2\pi j} \int_{-\pi}^{+\pi} X(e^{j\omega})\, (e^{j\omega})^{n-1}\, j\, e^{j\omega} d\omega$$

Integration path is unit circle because of $e^{j\omega}$

$$=\; \frac{1}{2\pi} \int_{-\pi}^{+\pi} X(e^{j\omega})\, e^{j\omega n}\, d\omega$$

2.8 System Representation and Examples

Example 1: Difference calculation

- Difference equation

$$y[n] \;=\; x[n] \;-\; x[n-n_0], \qquad n_0 \text{ integral number}$$

- Fourier transform gives:

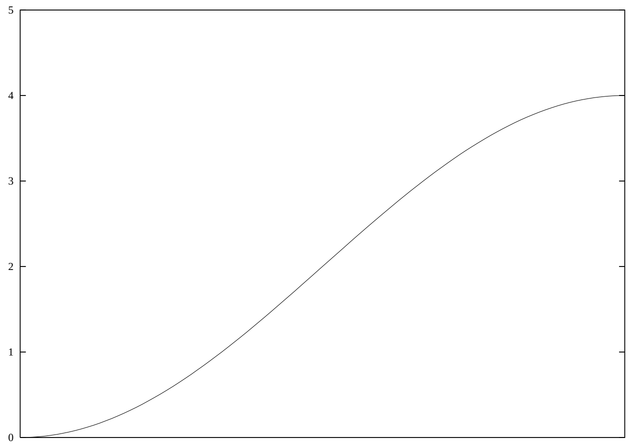
$$\begin{aligned}
\sum_{n=-\infty}^{\infty} y[n]\, e^{-j\omega n} \;&=\; \sum_{n=-\infty}^{\infty} x[n]\, e^{-j\omega n} \;-\; \sum_{n=-\infty}^{\infty} x[n-n_0]\, e^{-j\omega n} \\
Y(e^{j\omega}) \;&=\; X(e^{j\omega}) \;-\; \sum_{n=-\infty}^{\infty} \left(x[n]\, e^{-j\omega n}\, e^{-j\omega n_0} \right) \\
&=\; X(e^{j\omega}) \;-\; e^{-j\omega n_0}\, X(e^{j\omega})
\end{aligned}$$

- Then follows:

$$\begin{aligned}
H(e^{j\omega}) \;&=\; \frac{Y(e^{j\omega})}{X(e^{j\omega})} \\
&=\; 1 \;-\; e^{-j\omega n_0}
\end{aligned}$$

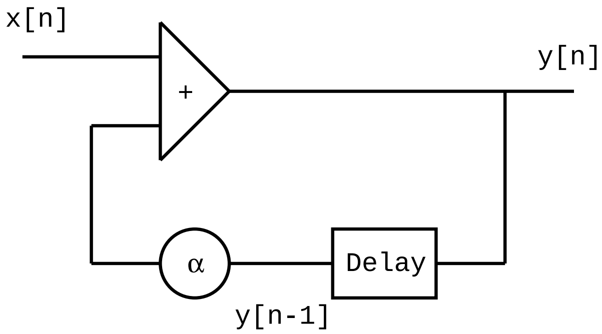
$$\begin{aligned}
|H(e^{j\omega})|^2 \;&=\; (1-\cos(\omega n_0))^2 \;+\; \sin^2(\omega n_0) \\
&=\; 1 \;-\; 2\cos(\omega n_0) \;+\; \cos^2(\omega n_0) \;+\; \sin^2(\omega n_0) \\
&=\; 2\,(1 \;-\; \cos(\omega n_0))
\end{aligned}$$

$|H(e^{i\omega})|^2$



ω $\qquad \frac{\pi}{n_0}$

Example 2: First order difference equation



$$
\begin{array}{rrcl}
 & x[n] \;+\; \alpha\, y[n-1] & = & y[n] \\
\Longleftrightarrow & y[n] \;-\; \alpha\, y[n-1] & = & x[n]
\end{array}
$$

Method 1: Estimation of transfer function $H(e^{j\omega})$
from impulse response $h[n]$:

- From the Eq. above with $y[n] = h[n]$ and $x[n] = \delta[n]$ follows:

$$
\begin{aligned}
h[n] &= \delta[n] \;+\; \alpha\, h[n-1] \\
&= \delta[n] \;+\; \alpha\, \delta[n-1] \;+\; \alpha^2\, \delta[n-2] \;+\; \cdots \\
&= \begin{cases} \alpha^n, & n \geq 0 \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
$$

- Fourier spectrum/transfer function $H(e^{j\omega})$

$$
\begin{aligned}
H(e^{j\omega}) &= \sum_{k=-\infty}^{+\infty} h[k]\, e^{-j\omega k} \\
&= \sum_{k=0}^{+\infty} \alpha^k\, e^{-j\omega k} \\
&= \sum_{k=0}^{+\infty} \left(\alpha\, e^{-j\omega}\right)^k \\
&= \frac{1}{1 \;-\; \alpha\, e^{-j\omega}} \qquad \text{for} \quad |\alpha| < 1
\end{aligned}
$$

Method 2: Estimation of transfer function $H(e^{j\omega})$ using Fourier transform of difference equation:

- Difference equation:

$$y[n] \; - \; \alpha \, y[n-1] \quad = \quad x[n]$$

- Fourier–transform:

$$Y(e^{j\omega}) \; - \; \alpha \, e^{-j\omega} \, Y(e^{j\omega}) \quad = \quad X(e^{j\omega})$$

- Result:

$$\begin{aligned} H(e^{j\omega}) \quad &= \quad \frac{Y(e^{j\omega})}{X(e^{j\omega})} \\ &= \quad \frac{1}{1 \; - \; \alpha \, e^{-j\omega}} \end{aligned}$$

Example 3: Linear difference equations (with constant coefficients)

- Difference equation:

$$y[n] = \sum_{i=0}^{I} b[i]\ x[n-i] - \sum_{j=1}^{J} a[j]\ y[n-j]$$

- z-transform:

$$Y(z) = X(z) \sum_{i=0}^{I} b[i] z^{-i} - Y(z) \sum_{j=1}^{J} a[j] z^{-j}$$

- Result:

$$\begin{aligned} H(z) &= \frac{Y(z)}{X(z)} \\ &= \frac{\sum\limits_{i=0}^{I} b[i]\ z^{-i}}{1 + \sum\limits_{j=1}^{J} a[j]\ z^{-j}} \\ &= \sum_{n=-\infty}^{+\infty} h[n]\ z^{-n} \end{aligned}$$

Using the definition of $H(z)$ we can optain the impulse response as a function of the coefficients of the difference equation in the above term.

Remark:

If we factorise denominator and numerator polynoms into linear factors, we can obtain a zero-pole-representation of a discrete time LTI system:

$$H(i) = \frac{\Pi_{i=1}^{I}(z - v_i)}{\Pi_{j=1}^{J}(z - w_j)}$$

with zeros $v_i \in \mathbb{C}$ and poles $w_j \in \mathbb{C}$.

- in general:

  $h[n]$ has infinite number of non-zero values

$$\Longrightarrow \text{IIR–filter: Infinite Impulse Response}$$

- but if: $a[j] \equiv 0 \qquad \forall j$
  $h[n]$ identical to zero outside of a finite interval

$$h[n] = \begin{cases} b[n] & n = 0, \ldots, I \\ 0 & \text{otherwise} \end{cases}$$

$$\Longrightarrow \text{FIR–filter: Finite Impulse Response}$$

Example 4:
Impulse response as "truncated geometric series"

$$h[n] = \begin{cases} a^n & 0 \leq n \leq M \qquad a \in \mathbb{R} \\ 0 & \text{otherwise} \end{cases}$$

$$H(z) = \sum_{n=0}^{M} a^n \, z^{-n} = \frac{1 - a^{M+1} \, z^{-(M+1)}}{1 - a \, z^{-1}}$$

system operation:

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[k] \, x[n-k] \\ &= \sum_{k=0}^{M} a^k x[n-k] \end{aligned}$$

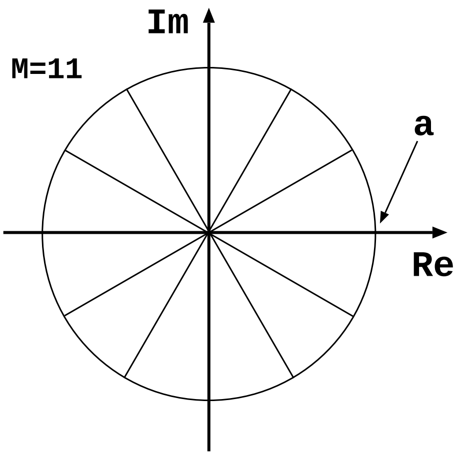
or as difference equation ("recursively")

$$y[n] - a \, y[n-1] = x[n] - a^{M+1} x[n-M-1]$$

For this example we consider the zero-pole-representation:

$$H(z) \;=\; \frac{1-\left(\frac{z}{a}\right)^{-(M+1)}}{1-\left(\frac{z}{a}\right)^{-1}} \qquad \alpha > 0$$

$$\begin{aligned} \text{Zeros:} \quad z_k &= a \cdot e^{j\frac{2\pi k}{M+1}} \qquad k \;=\; 0, 1, \ldots, M \\ \text{Pole:} \quad z_0 &= a \\ &\text{(cancelled by zero } z_0 \;=\; a\text{)} \end{aligned}$$



Example 5:
Fibonacci numbers

Difference equation:

$$
\begin{array}{llcl}
n \geq 0 & h[n+2] & = & h[n+1] \; + \; h[n] \\
 & h[0] & = & h[1] \; = \; 1 \\
n < 0 & h[n] & = & 0
\end{array}
$$

$$
\begin{aligned}
H(z) &= \sum_{n=-\infty}^{\infty} h[n] z^{-n} \\
&= 1 \; + \; z^{-1} \; + \; \sum_{n=0}^{\infty} h[n+2] z^{-(n+2)} \\
&= 1 \; + \; z^{-1} \; + \; \sum_{n=0}^{\infty} h[n+1] z^{-(n+2)} \; + \; \sum_{n=0}^{\infty} h[n] z^{-(n+2)} \\
&= 1 \; + \; z^{-1} \; + \; z^{-1} \sum_{n=0}^{\infty} h[n+1] z^{-(n+1)} \; + \; z^{-2} \sum_{n=0}^{\infty} h[n] z^{-n} \\
&= 1 \; + \; z^{-1} \; \underbrace{\left(1 \; + \; \sum_{n=1}^{\infty} h[n] z^{-n}\right)}_{H(z)} + z^{-2} \underbrace{\sum_{n=0}^{\infty} h[n] z^{-n}}_{H(z)} \\
&= 1 \; + \; z^{-1} H(z) \; + \; z^{-2} H(z)
\end{aligned}
$$

$$
\begin{aligned}
H(z)(1 - z^{-1} - z^{-2}) &= 1 \\
H(z) &= \frac{1}{1 - z^{-1} - z^{-2}} \\
&= \frac{1}{\sqrt{5}} \left(\frac{a}{1 - a z^{-1}} - \frac{b}{1 - b z^{-1}} \right)
\end{aligned}
$$

where $a = \dfrac{1 + \sqrt{5}}{2}$ and $b = \dfrac{1 - \sqrt{5}}{2}$

For a and b the following holds:

$$\frac{1}{1-\left(\frac{a}{z}\right)} = \sum_{n=0}^{\infty} \left(\frac{a}{z}\right)^n = \sum_{n=0}^{\infty} a^n z^{-n}$$

That results in:

$$\begin{aligned} H(z) &= \sum_{n=0}^{\infty} \frac{1}{\sqrt{5}} \left(a^{n+1} - b^{n+1} \right) z^{-n} \\ &\stackrel{!}{=} \sum_{n=0}^{\infty} h[n]\, z^{-n} \end{aligned}$$

$$h[n] = \frac{1}{\sqrt{5}} \left(a^{n+1} - b^{n+1} \right)$$

Table 2.1: Fourier transform pairs

	signal	Fourier–transform
1.	$\delta[n]$	1
2.	$\delta[n-n_0]$	$e^{-j\omega n_0}$
3.	$1 \qquad (-\infty < n < \infty)$	$\sum_{k=-\infty}^{\infty} 2\pi\delta(\omega + 2\pi k)$
4.	$a^n u[n] \qquad (\|a\| < 1)$	$\frac{1}{1-ae^{-j\omega}}$
5.	$u[n]$	$\frac{1}{1-e^{-j\omega}} + \sum_{k=-\infty}^{\infty} \pi\delta(\omega + 2\pi k)$
6.	$(n+1)a^n u[n] \qquad (\|a\| < 1)$	$\frac{1}{(1-ae^{-j\omega})^2}$
7.	$\frac{r^n \sin\omega_p(n+1)}{\sin\omega_p} u[n] \qquad (\|r\| < 1)$	$\frac{1}{1-2r\cos\omega_p\, e^{-j\omega} \ + \ r^2 e^{-j2\omega}}$
8.	$\frac{\sin\omega_c n}{\pi n}$	$X(e^{j\omega}) = \begin{cases} 1, & \|\omega\| < \omega_c, \\ 0, & \omega_c < \|\omega\| \leq \pi \end{cases}$
9.	$x[n] = \begin{cases} 1, & 0 \leq n \leq M \\ 0, & otherwise \end{cases}$	$\frac{\sin[\omega(M+1)/2]}{\sin(\omega/2)} \, e^{-j\omega M/2}$
10.	$e^{j\omega_0 n}$	$\sum_{k=-\infty}^{\infty} 2\pi\delta(\omega - \omega_0 + 2\pi k)$
11.	$\cos(\omega_0 n + \phi)$	$\pi \sum_{k=-\infty}^{\infty} [e^{j\phi}\delta(\omega - \omega_0 + 2\pi k) + e^{-j\phi}\delta(\omega - \omega_0 + 2\pi k)]$

$$u[n] = \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases}$$

2.9 Discrete Time Signal Fourier Transform Theorem

Basically there is no difference between FT theorem for the continuous time and the discrete time case because summation has the same properties as integration. Only differentiation and difference calculation are not completely analog, because it is not possible to form a derivative in the discrete time case.

Table 2.2: Fourier transform Theorems

	signal $x[n], y[n]$	Fourier–transform $X(e^{j\omega}), Y(e^{j\omega})$
1.	$ax[n] + by[n]$	$aX(e^{j\omega}) + bY(e^{j\omega})$
2.	$x[n - n_d]$, n_d is integral number	$e^{-j\omega n_d} X(e^{j\omega})$
3.	$e^{j\omega_0 n} x[n]$	$X(e^{j(\omega-\omega_0)})$
4.	$x[-n]$	$X(e^{-j\omega})$ $\overline{X}(e^{j\omega})$ if $x[n]$ is real
5.	$nx[n]$	$j\dfrac{dX(e^{j\omega})}{d\omega}$
6.	$x[n] * y[n]$	$X(e^{j\omega})Y(e^{j\omega})$
7.	$x[n]y[n]$	$\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} X(e^{j\Theta})Y(e^{j(\omega-\Theta)})d\Theta$
8.	$x[n] - x[n-1]$	$(1 - e^{-j\omega})X(e^{j\omega})$ $\lvert 1 - e^{-j\omega}\rvert^2 = 2(1 - \cos\omega)$
	Parseval theorem	
9.	$\displaystyle\sum_{n=-\infty}^{\infty} \lvert x[n]\rvert^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \lvert X(e^{j\omega})\rvert^2 d\omega$	
10.	$\displaystyle\sum_{n=-\infty}^{\infty} x[n]\overline{y}[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})\overline{Y}(e^{j\omega}) d\omega$	

Example 1 corresponding to Theorem 5:

$$
\begin{aligned}
X(e^{j\omega}) &= \sum_{k=-\infty}^{+\infty} x[k]\; e^{-j\omega k} \\
\frac{d}{d\omega}\, X(e^{j\omega}) &= \frac{d}{d\omega} \left(\sum_{k=-\infty}^{+\infty} x[k]\; e^{-j\omega k} \right) \\
&= \sum_{k=-\infty}^{+\infty} \frac{d}{d\omega} \;\big(\, x[k]\; e^{-j\omega k} \,\big) \\
&= \sum_{k=-\infty}^{+\infty} x[k]\; (-jk)\; e^{-j\omega k} \\
\Longleftrightarrow \qquad j \frac{d}{d\omega} X(e^{j\omega}) &= \sum_{k=-\infty}^{+\infty} k\; x[k]\; e^{-j\omega k}
\end{aligned}
$$

$$
F\left\{n \cdot x[n]\right\} \;=\; j\, \frac{d}{d\omega}\; F\left\{x[n]\right\}
$$

Example 2 corresponding to Theorem 8:

$$
\begin{aligned}
F\left\{x[n] - x[n-1]\right\} &= \sum_{k=-\infty}^{+\infty} x[k]\; e^{-j\omega k} \;-\; \sum_{k=-\infty}^{+\infty} x[k-1]\; e^{-j\omega k} \\
&= \sum_{k=-\infty}^{+\infty} x[k]\; e^{-j\omega k} \;-\; \sum_{k=-\infty}^{+\infty} x[k]\; e^{-j\omega k}\; e^{-j\omega} \\
&= X(e^{j\omega})\; \left(1 - e^{-j\omega}\right)
\end{aligned}
$$

$$
\begin{aligned}
\Longrightarrow \qquad \left|F\left\{x[n] - x[n-1]\right\}\right|^2 &= \left|F\left\{x[n]\right\}\right|^2 \,|1 \;-\; e^{-j\omega}|^2 \\
&= \left|F\left\{x[n]\right\}\right|^2\; 2 \cdot (1 - \cos(\omega))
\end{aligned}
$$

2.10 Discrete Fourier Transform: DFT

The Fourier transform for discrete time signals and systems has been explained on the previous pages. For discrete time signals with finite length there is also another Fourier representation called "Discrete Fourier Transform" (DFT).

The DFT plays a central role in digital signal processing.

Decisive reasons:

- fast algorithms exist for DFT calculation
  (Fast Fourier Transform, FFT).
- discrete frequencies ω_k can be better represented in the computer than continuous frequencies ω.

Assume a discrete time signal $x[n]$ with finite length (see also chapter 3.2 on page 135):

$$x[n] = \begin{cases} x[n] & 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases}$$

Note: For a continuous time signal it is *impossible* in the strict sense to be band-limited *and* time-limited (truncation effect $\hat{=}$ Windowing).

- The discrete time signal Fourier transform for $x[n]$ is:

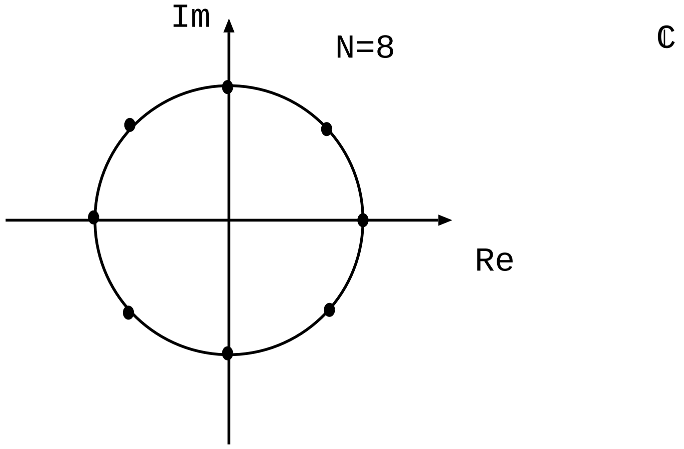
$$X(e^{j\omega}) = \sum_{n=0}^{N-1} x[n] \exp(-j\omega n)$$

- ω is a continuous variable. The period is 2π. Frequency discretisation is made by sampling along the frequency axis.
- The Fourier transform $X(e^{j\omega})$ is evaluated at

$$\omega_k = \frac{2\pi}{N} k \qquad \text{where} \quad k = 0, 1, \ldots, N-1$$

- Define:

$$X[k] := X(e^{j\omega})|\omega = \omega_k$$



- Discrete Fourier Transform (DFT):

$$X[k] \;=\; \sum_{n=0}^{N-1} x[n] \exp(-j\,\frac{2\pi}{N}\,k\,n), \qquad k \;=\; 0, 1, \ldots, N-1$$

- Inverse DFT:

$$x[n] \;=\; \frac{1}{N} \sum_{k=0}^{N-1} X[k] \exp(j\,\frac{2\pi}{N}\,k\,n), \qquad n \;=\; 0, 1, \ldots, N-1$$

- Remark:

  This equation can be proven by inserting the equation for $X[k]$ in the equation for $x[n]$ and using the orthogonality:

$$\frac{1}{N} \sum_{n=0}^{N-1} \exp\left(j \frac{2\pi}{N} kn\right) \;=\; \begin{cases} 1 & k = m\,N, \quad m \quad \text{is integral number} \\ 0 & \text{otherwise} \end{cases}$$

- Note:

  Consider the "analogy" between inverse DFT (above) and inverse Fourier transform of discrete time signal:

$$x[n] \;=\; \frac{1}{2\pi} \int_0^{2\pi} X(e^{j\omega})\, e^{j\omega n}\, d\omega$$

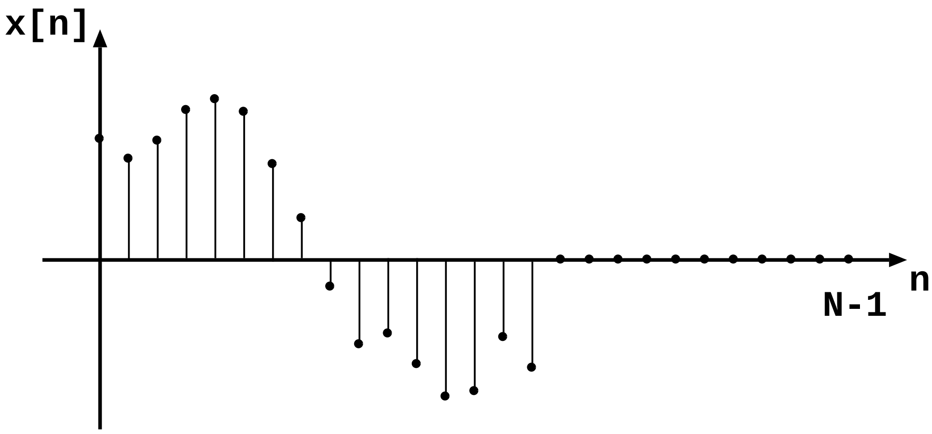
  Under the given conditions the integral is equal to the sum (*without approximation!*).

Remarks:

- DFT coefficients $X[k]$ are not an approximation of the discrete time signal Fourier transform $X(e^{j\omega})$. On the contrary:

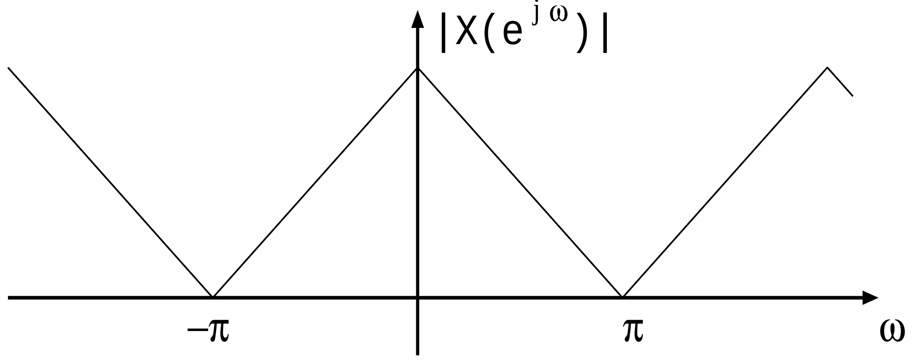
$$X[k] \quad = \quad X(e^{j\omega})|\omega = \omega_k$$

- Number of the coefficients $X[k]$ depends on the signal length N. A finer sampling of the discrete time signal Fourier transform is possible by appending zeros to the signal $x[n]$ (Zero–Padding).



Interpretation of Fourier coefficients

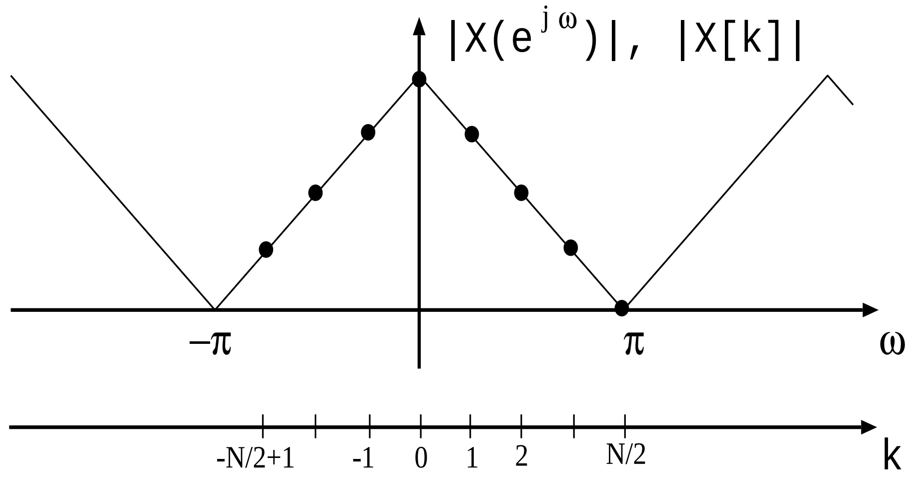
- Fourier transform $X(e^{j\omega})$ of the time discrete signal $x[n]$



- Evaluation at N discrete sampling points

$$\omega_k \;=\; \frac{2\pi}{N}\,k$$

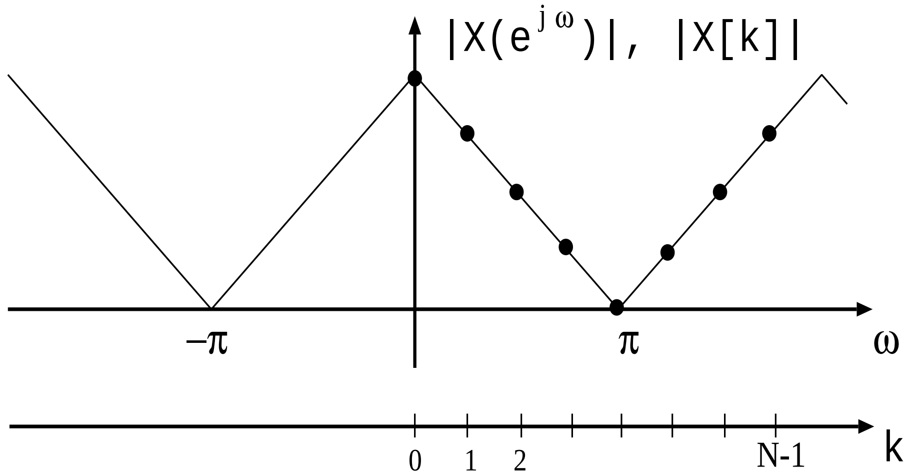
yields the DFT coefficients $X[k]$.
At first k lies in the domain $k \;=\; -\frac{N}{2}+1,\ldots,0,\ldots,\frac{N}{2}$.



- Because of the periodicity of $X(e^{j\omega})$ the coefficients $X[k]$ can also be obtained by shifting the sampling points with negative frequency into the positive frequency domain (by one period).

  Then $k = 0, \ldots, N/2, \ldots, N-1$.

$$X[k] \;=\; \sum_{n=0}^{N-1} x[n] \exp(-j\,\frac{2\pi}{N}\,k\,n)$$



- Interpretation of coefficients for general signal $x[n]$:

$$\begin{aligned}
k = 0 \qquad &\longleftrightarrow \qquad f = 0 \\
1 \le k \le \frac{N}{2} - 1 \qquad &\longleftrightarrow \qquad 0 < f < \frac{f_S}{2} \\
k = \frac{N}{2} \qquad &\longleftrightarrow \qquad \pm\frac{f_S}{2} \\
\frac{N}{2} + 1 \le k \le N - 1 \qquad &\longleftrightarrow \qquad -\frac{f_S}{2} < f < 0
\end{aligned}$$

Symmetric relations by real signals:

- For DFT coefficients $X[k]$ of a real signal $x[n]$ the following holds:

$$X[k] \quad = \quad \overline{X[N-k]}$$

$$\begin{aligned} \mathrm{Re}(X[k]) &= \mathrm{Re}(X[N-k]) \\ \mathrm{Im}(X[k]) &= -\mathrm{Im}(X[N-k]) \end{aligned}$$

- For the amplitude spectrum $|\,X[k]\,|$ the following holds:

$$\begin{aligned} |\,X[k]\,|^2 &= \mathrm{Re}^2\{X[k]\} + \mathrm{Im}^2\{X[k]\} \\ &= |\,X[N-k]\,|^2 \end{aligned}$$

Realization of DFT:

```
/*      PI = 3.14159265358979                                        */
/*      x:        input signal                                       */
/*      N: length of input signal                                    */
/*      Xre, Xim: real and imaginary part of DFT coefficients        */

void    dft (int N, float x[], float Xre[], float Xim[]) {
        int     n, k;
        float   SumRe, SumIm;

        for (k=0; k<=N-1; k++)
        {
             SumRe = 0.0;
             SumIm = 0.0;

             for (n=0; n<=N-1; n++)
             {
                  SumRe += x[n]*cos(2*PI*k*n/N);
                  SumIm -= x[n]*sin(2*PI*k*n/N);
             }

             Xre[k] = SumRe;
             Xim[k] = SumIm;
        }
}
```

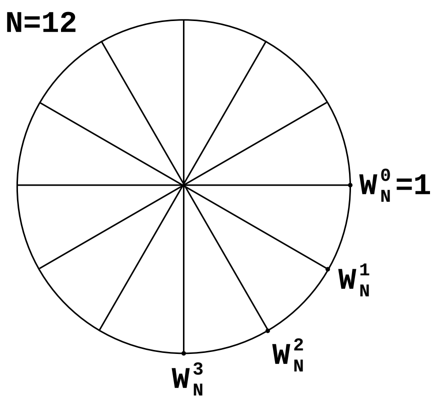
Remark:

- discrete realization
- Reduction of "Fourier powers" $e^{-\frac{2\pi j}{N} \cdot kn}$ to $e^{-\frac{2\pi j}{N} \cdot l}$ $(l = 0, 1, \ldots, N-1)$ is possible, because they are periodical (on the unit circle).

2.11 DFT as Matrix Operation

- Notation with unit roots

$$
\begin{aligned}
X[k] &= \sum_{n=0}^{N-1} x[n] \exp(-\frac{2\pi j}{N} k\, n) \\
&= \sum_{n=0}^{N-1} x[n]\, W_N^{kn} \\
\text{where} \quad & W_N := \exp(-\frac{2\pi j}{N})
\end{aligned}
$$



- Periodicity of W_N
  unit root:

$$
\begin{aligned}
\exp(-j\,\omega_k) &= \exp(-j\frac{2\pi}{N}k) = (W_N)^k \\
W_N &:= \exp(-j\frac{2\pi}{N})
\end{aligned}
$$

Note:

$$
\begin{array}{lll}
1. & W_N^r & = W_N^{r \bmod N} \\[1em]
2. & W_N^{kN} & = (W_N^N)^k = 1^k = 1 \qquad k \in \mathbb{Z} \\[1em]
3. & W_N^2 & = [\exp(-\frac{2\pi j}{N})]^2 = \exp(-\frac{2\pi j}{N}\, 2) \\[1em]
 & & = \exp(-\frac{2\pi j}{N/2}) = W_{N/2} \qquad N \text{ even} \\[1em]
4. & W_N^{N/2} & = \exp(-\frac{2\pi j}{N}\,\frac{N}{2}) = \exp(-\pi j) = -1 \\[1em]
5. & W_N^{r+N/2} & = W_N^{N/2}\, W_N^r = -W_N^r
\end{array}
$$

DFT as matrix multiplication

$$
\begin{aligned}
X[k] &= \sum_{n=0}^{N-1} x[n] \exp\left(-\frac{2\pi j}{N}\, k\, n\right) \\
&= \sum_{n=0}^{N-1} W_N^{kn}\, x[n] \\
&= \sum_{n=0}^{N-1} \{W_N\}_{kn}\, x[n]
\end{aligned}
$$

with the matrix $\{W_N\}$ and the matrix elements:

$$
\{W_N\}_{kn} := W_N^{kn}
$$

Inversion:

$$
\begin{aligned}
x[n] &= \frac{1}{N} \sum_{k=0}^{N-1} X[k] \exp\left(\frac{2\pi j}{N}\, k\, n\right) \\
&= \frac{1}{N} \sum_{k=0}^{N-1} (W_N^{-1})^{kn}\, X[k] \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \{W_N^{-1}\}_{kn}\, X[k]
\end{aligned}
$$

Therefore for the matrix $\{W_N\}^{-1}$ holds:

$$
\{W_N\}^{-1} := \frac{1}{N} \{W_N^{-1}\}
$$

DFT matrix operation: properties

- DFT: invertible linear mapping

$$N \text{ complex signal values} \leftrightarrow N \text{ complex Fourier components}$$

$$N \text{ real signal values} \leftrightarrow \frac{N}{2} \text{ complex Fourier components}$$

(due to symmetry)

in words:
DFT causes no "information loss" in the signal.

- Parseval theorem for DFT
general Fourier:

$$\sum_{n=0}^{N-1} |x[n]|^2 \; = \; \frac{1}{2\pi} \int_{-\pi}^{+\pi} |X(e^{j\omega})|^2 d\omega$$

special DFT: (recalculate for yourself!)

$$\sum_{n=0}^{N-1} |x[n]|^2 \; = \; \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2$$

in words:
Disregarding the factor $\frac{1}{N}$, the DFT is a norm conserving ($\hat{=}$ energy conserving) transformation (mathematical terminology: "unitary").

2.12 From Continuous Fourier Transform to Matrix Representation of Discrete Fourier Transform

Assumption: band-limited signal $x(t)$

Fourier transform of the continuous time signal $x(t)$:

$$X(\omega) \;=\; F\{x(t)\} \;=\; \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t} dt \tag{2.9}$$

For the *exact reconstruction* (without approximation) of the continuous time signal from sampled values, the samples $x[n] \;=\; x(n \cdot T_s)$ must have the distance of at most

$$T_s \;=\; \frac{\pi}{\omega_B}$$

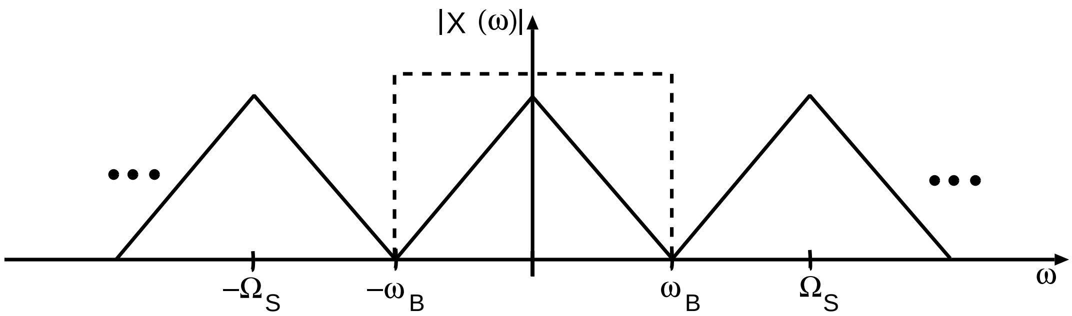
(sampling theorem).

This results in the Fourier transform of the discrete time signal $x[n]$:

$$X(e^{j\omega}) \;=\; \sum_{-\infty}^{\infty} x[n]\, e^{-j\omega n} \tag{2.10}$$

where ω is frequency "normalised on T_s"

Functions (2.9) and (2.10) agree in interval $[-\Omega_S/2, +\Omega_S/2] = [-\omega_B, +\omega_B]$.



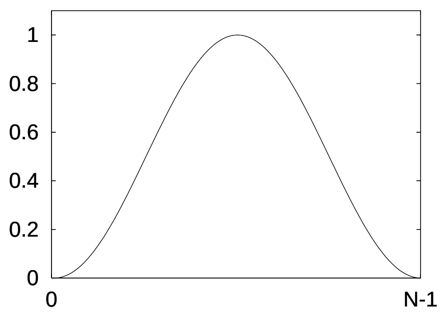


Figure 2.12: Hanning window

The signal $x[n]$ is further decomposed by applying a window function $w[n]$ (windowing):

$$w[n] \;=\; \begin{cases} \dots & n = 0, \dots, N-1 \\ 0 & otherwise \end{cases}$$

Windowed signal $y[n]$:

$$y[n] \;=\; w[n] \cdot x[n]$$

can be analyzed using Fourier transform or DFT.

$$Y(e^{j\omega_k}) \;=\; \sum_{n=0}^{N-1} y[n]\, e^{-j\omega_k}$$

DFT:

$$\begin{aligned} \omega_k &= \frac{2\pi k}{N} \qquad \text{where } k \;=\; 0, \dots, N-1 \\ Y[k] &= \sum_{n=0}^{N-1} y[n]\, e^{-\frac{2\pi}{N}kn} \end{aligned}$$

Matrix representation:

$$K = N \quad \begin{bmatrix} Y[0] \\ \vdots \\ Y[k] \\ \vdots \\ Y[K-1] \end{bmatrix} = \begin{bmatrix} & \vdots & \\ & & \\ & e^{-\frac{2\pi j}{N} \cdot n \cdot k} & \\ & & \\ & \vdots & \end{bmatrix} \begin{bmatrix} y[0] \\ \vdots \\ y[n] \\ \vdots \\ y[N-1] \end{bmatrix}$$

2.13 Frequency Resolution and Zero Padding

Task: signal $x[n]$ with finite length N is given.

Wanted: Fourier transform $X(e^{j\omega_k})$ at

$$\omega_k = \frac{2\pi}{K}k, \quad \text{where } k = 0, 1, \ldots, K-1 \text{ and } K > N$$

Inserting the definitions:

$$\begin{aligned}
X(e^{j\omega_k}) &= \sum_{n=0}^{N-1} x[n] \exp\left(-\frac{2\pi j}{K} k\, n\right) \\
&= \sum_{n=0}^{K-1} \tilde{x}[n] \exp\left(-\frac{2\pi j}{K} k\, n\right) \\
\text{where} \quad \tilde{x}[n] &= \begin{cases} x[n] & n = 0, \ldots, N-1 \\ 0 & n = N, \ldots, K-1 \end{cases}
\end{aligned}$$

i.e. Zero Padding (appending zeros).

Matrix representation:

$$\begin{bmatrix} X[0] \\ \vdots \\ \\ \\ \vdots \\ X[K-1] \end{bmatrix} = \begin{bmatrix} & & \\ & & \\ & W_K^{nk} & \\ & & \\ & & \end{bmatrix} \begin{bmatrix} x[0] \\ \vdots \\ \vdots \\ x[N-1] \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} n = 0 \\ \\ \\ n = N-1 \\ n = N \\ \\ n = K-1 \end{array}$$

Note:
"Zero Padding" does not introduce any additional information into the signal. This is only a trick so that DFT and particularly FFT (Fast Fourier Transform) can be performed with a

higher frequency resolution .

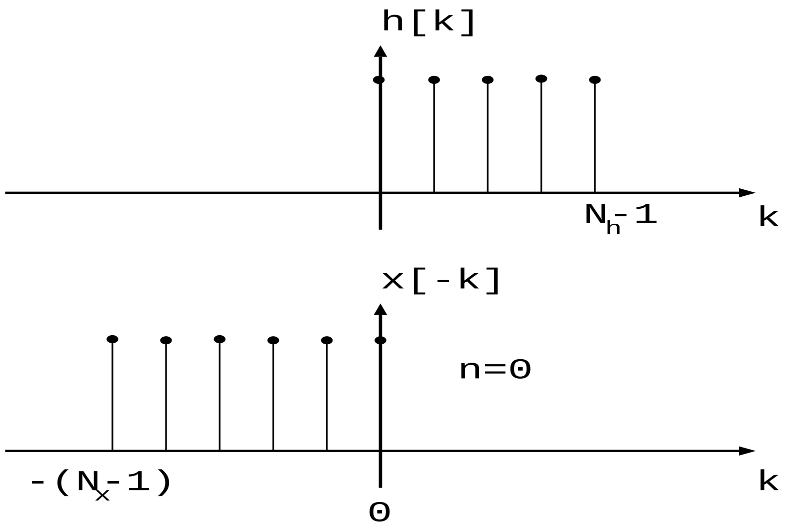
2.14 Finite Convolution

Input signal and convolution kernel have finite duration

Consider "finite" convolution:

- Impulse response: $h[n] \equiv 0$ for $n \notin \{0, 1, 2, \ldots, N_h - 1\}$
- Input signal: $x[n] \equiv 0$ for $n \notin \{0, 1, 2, \ldots, N_x - 1\}$
- Output signal:

$$
\begin{aligned}
y[n] &= \sum_{k=-\infty}^{\infty} h[k]\, x[n-k] \\
&= \sum_{k=0}^{N_h-1} h[k]\, x[n-k]
\end{aligned}
$$



- Altogether: $N_x + N_h - 1$ positions with "overlap"
- Therefore only $N_x + N_h - 1$ values of output signal can be different from zero:

$$
y[n] = \begin{cases} 0 & n > N_x + N_h - 2 \\ \ldots & n = 0, 1, \ldots, N_x + N_h - 2 \\ 0 & n < 0 \end{cases}
$$

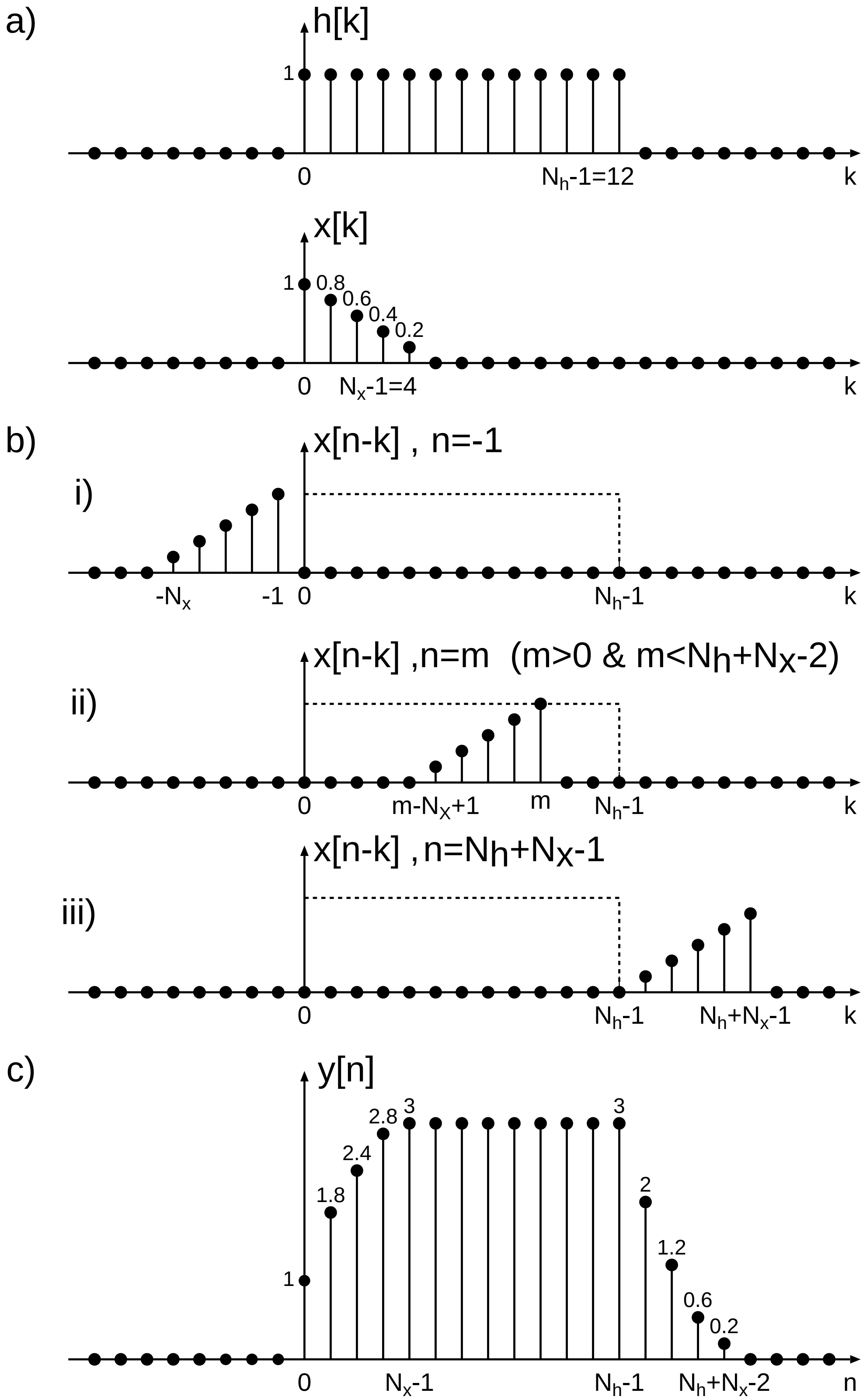


Figure 2.13: Example of a linear convolution of two finite length signals: a) two signals; b) signal x[n-k] for different values of n:
i) $n < 0$, no overlap with h[k], therefore convolution y[n] = 0
ii) n between 0 and $N_h + N_x - 2$, convolution $\neq 0$
iii) $n > N_h + N_x - 2$, no overlap with h[k], convolution y[n] = 0
c) resulting convolution y[n].

Finite convolution using DFT

Convolution theorem:

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]\, x[n-k]$$

Fourier:

$$Y(e^{j\omega}) = H(e^{j\omega})\, X(e^{j\omega}), \qquad 0 \leq \omega \leq 2\pi$$

Also valid for sample frequencies:

$$\omega_k := \frac{2\pi}{N}\, k, \qquad k = 0, \ldots, N-1 \text{ for any } N$$

Notation: $Y[k] = H[k]\, X[k]$

- Question: How to choose the length N of the DFT ?
  - Reminder: different "lengths"
  - $x[n]$: N_x non-zero values
  - $h[n]$: N_h non-zero values
  - $y[n]$: $N_y = N_x + N_h - 1$ non-zero values
- Answer:
  - The convolution theorem is certainly correct for any $N > 0$.
  - If we want to calculate the output signal completely from $Y[k]$, we have to know $Y[k]$ for

    at least $N = N_x + N_h - 1$

    frequency values $k = 0, 1, \ldots, N-1$.
  - In words: for the DFT length N must be valid:

    $$N \geq N_x + N_h - 1$$

    Method: Zero Padding, i.e. appending zeros.

Note: The FFT will be introduced on the next pages. A comparison of costs for realization of the finite convolution by DFT and FFT can be found at the end of paragraph 2.15.

2.15 Fast Fourier Transform (FFT)

Principle of FFT:
Calculation of the DFT can be done by successive decomposition into smaller DFT calculations. In this way, the number of elementary operations (multiplications and additions) is dramatically reduced:

FFT:

$$N^2 \;\rightarrow\; \frac{N}{2}\,\mathrm{ld}\,N \quad \text{operations}$$

$$N = 1024: \;\frac{2 \cdot N}{\mathrm{ld}\,N} \;=\; \frac{2 \cdot 1024}{10} \;=\; 200 \qquad \text{factor of velocity gain}$$

The matrix is decomposed into a product of sparse matrices, therefore N with lot of prime factors is convenient (not necessarily only powers of two).

Terminology for different variants of FFT:

- in time $\leftrightarrow$ in frequency
- in place: yes/no
- radix 2 $\leftrightarrow$ radix 4
- decomposition to prime factors instead of $N = 2^n$

History

1965	Cooley and Tukey
1942	Danielson and Lanczos
1905	Runge
1805	Gauss

Algorithms which are based on a decomposition of the signal $x[n]$ are called "decimation–in–time algorithms".

The case $N = 2^\nu$ is considered in the following.

$$\begin{aligned} X[k] &= \sum_{n=0}^{N-1} x[n] \exp(-j \frac{2\pi}{N} k\, n) \qquad \text{where} \quad k = 0, 1, \ldots, N-1 \\ &= \sum_{n=0}^{N-1} x[n]\, W_N^{nk} \qquad \text{where} \quad W_N^{nk} \;=\; \exp(-j \frac{2\pi}{N} k\, n) \end{aligned}$$

- Decomposition of the sum over n into the sums over even and odd n:

$$\begin{aligned} X[k] &= \sum_{r=0}^{N/2-1} x[2r]\, W_N^{2rk} \;+\; \sum_{r=0}^{N/2-1} x[2r+1]\, W_N^{(2r+1)k} \\ &= \sum_{r=0}^{N/2-1} x[2r]\, (W_N^2)^{rk} \;+\; W_N^k \sum_{r=0}^{N/2-1} x[2r+1]\, (W_N^2)^{rk} \end{aligned}$$

- Because of

$$W_N^2 \;=\; \exp(-2j \frac{2\pi}{N}) \;=\; \exp(-j \frac{2\pi}{N/2}) \;=\; W_{N/2}$$

  for $k = 0, \ldots, N-1$ holds:

$$\begin{aligned} X[k] &= \sum_{r=0}^{N/2-1} x[2r]\, W_{N/2}^{rk} \;+\; W_N^k \sum_{r=0}^{N/2-1} x[2r+1]\, (W_{N/2})^{rk} \\ &= G[k] \;+\; W_N^k\, H[k] \end{aligned}$$

- Each of the two sums corresponds to the DFT with the length $N/2$. The first sum is a $N/2$–DFT of the even indexed signal values $x[n]$, the second sum is a $N/2$–DFT of the odd indexed values.

- The DFT of the length N can be obtained by getting the two $N/2$–DFT's together, with the factor W_N^k.

Complexity:
The complexity $O(N^2)$ of one-dimensional FT can be reduced by adequate resorting values from two FTs with length $\frac{N}{2}$ and complexity $O(2 \cdot (\frac{N}{2})^2) = \frac{N^2}{2}$. By successive application of this resorting the complexity can be reduced to $O(N \log N)$.
The case $N = 2^3 = 8$ is considered in the following.

- $X[4]$ can be obtained from $H[4]$ and $G[4]$ according to previous equation.
- Because of the DFT–length $\frac{N}{2} = 4$:

$$H[4] \;=\; H[0] \quad \text{and} \quad G[4] \;=\; G[0]$$

And then:

$$X[4] \;=\; G[0] \;+\; W_N^4 \, H[0]$$

The values $X[5]$, $X[6]$ and $X[7]$ can be obtained analogously.

Flow diagram for decomposition of one N-DFT into two $N/2$–DFTs:

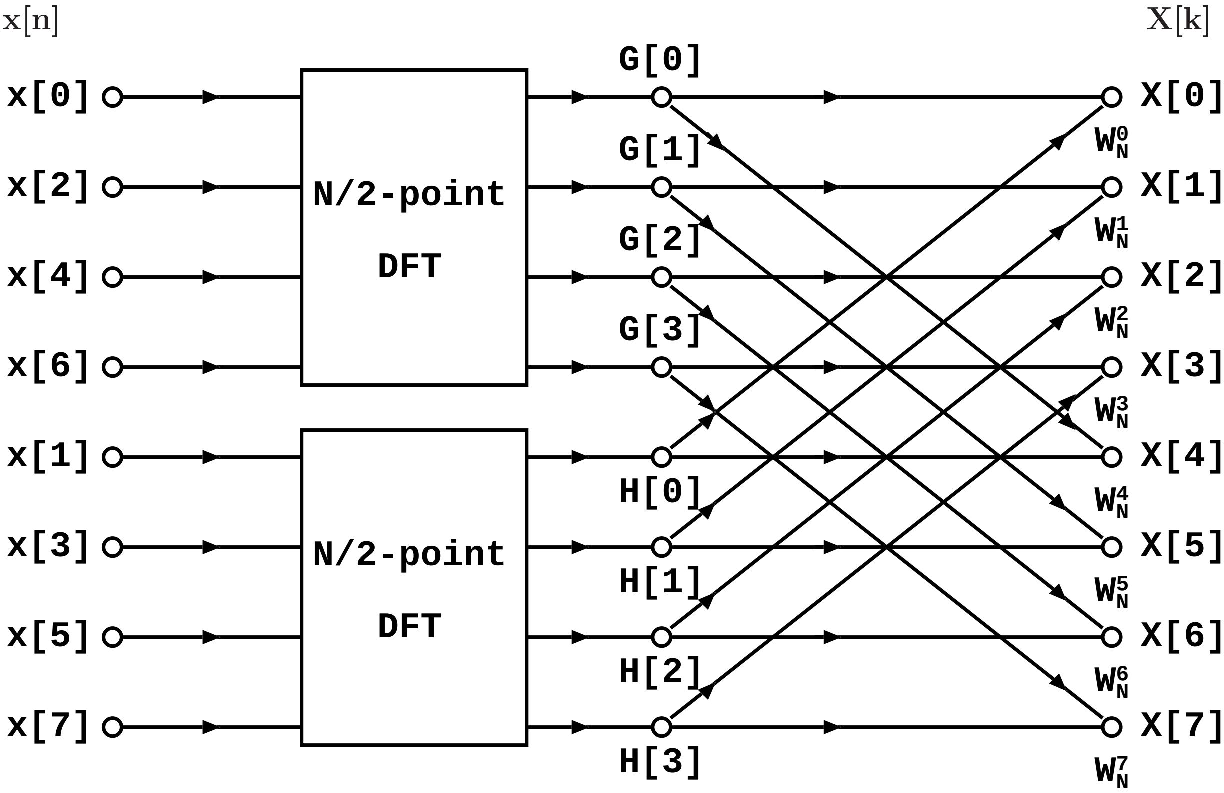


Figure 2.14: Flow diagram for decomposition of one N-DFT to two $N/2$–DFTs with $N = 8$

- Further analogous decomposition, until only DFT's with the length $N = 2$ remain (so called Butterfly Operation)
- Resulting flow diagram of the FFT:

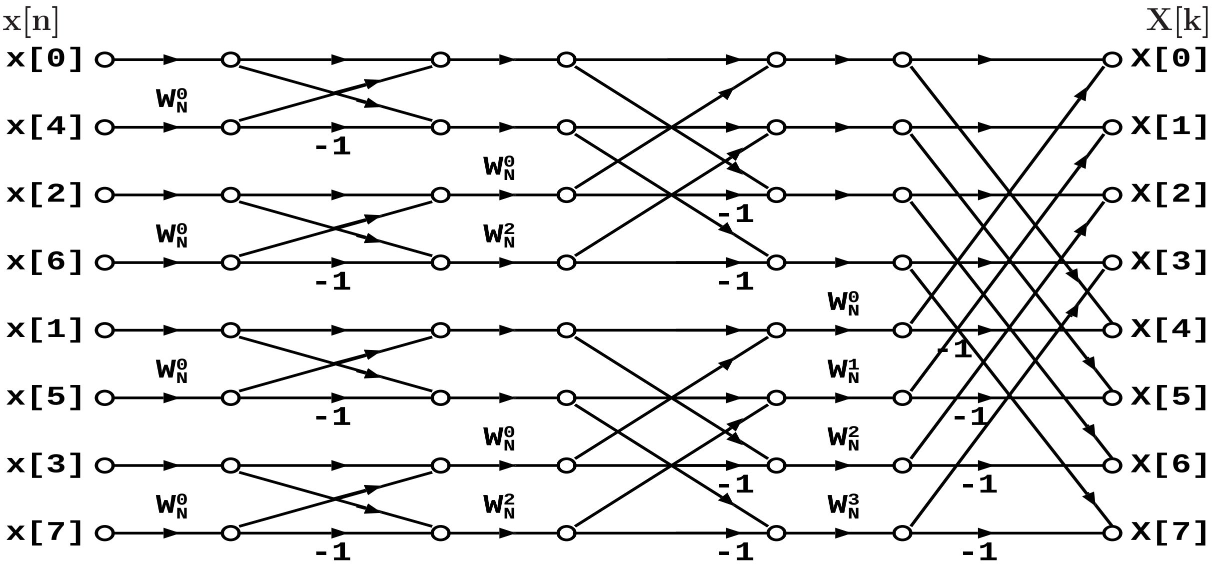


Figure 2.15: Flow diagram of an 8–point–FFT using Butterfly operations.

Complexity reduction

- Number of complex multiplications in FFT
  is $N/2 \cdot \mathrm{ld}\, N$.
- Comparison:
  Direct application of the DFT definition needs
  N^2 complex multiplications.
- Example: $N = 1024 = 2^{10}$

  $$\frac{N^2}{N/2 \cdot \mathrm{ld}\, N} \approx 200$$

  Complexity reduction by factor 200

  FFT with the base 2 is not minimal according to number of additions, FFT with the base 4 can be better.

Matrix representation of the FFT principle

- The complex Fourier matrix can be decomposed into the product of $r = \operatorname{ld} N$ matrices, each of them having only two non-zero elements in each column.
- The following graph shows the decomposition of the Fourier matrix in the case of inverse transformation.
- w corresponds to W_N^{-1}

$$X = |w^{nk}|X = T_3 \cdot T_2 \cdot T_1 \cdot T_S \cdot x$$

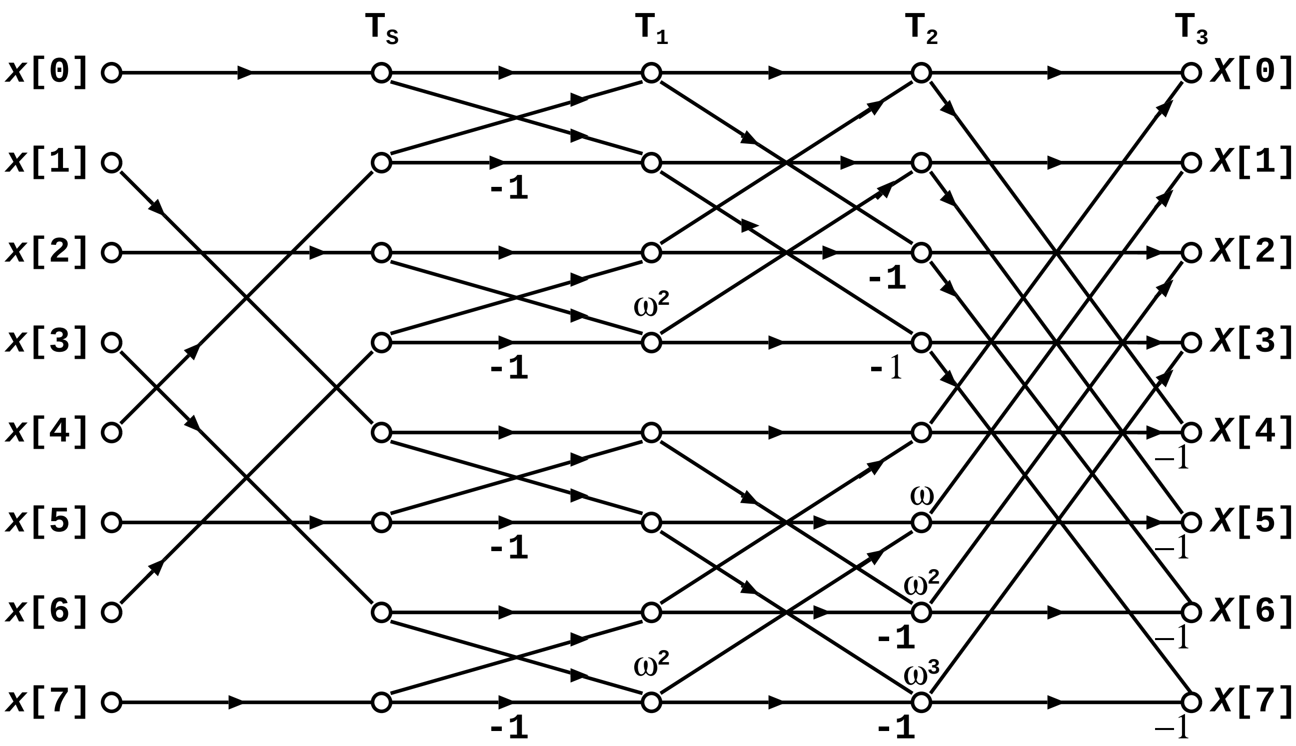
This is how the decomposition into $r + 1 = 4$ matrices looks like ($w^4 = -1, w^8 = 1$):

$$|w^{nk}| = \begin{array}{|cccccccc|}\hline 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & w & w^2 & w^3 & w^4 & w^5 & w^6 & w^7 \\ 1 & w^2 & w^4 & w^6 & w^8 & w^{10} & w^{12} & w^{14} \\ 1 & w^3 & w^6 & w^9 & w^{12} & w^{15} & w^{18} & w^{21} \\ 1 & w^4 & w^8 & w^{12} & w^{16} & w^{20} & w^{24} & w^{28} \\ 1 & w^5 & w^{10} & w^{15} & w^{20} & w^{25} & w^{30} & w^{35} \\ 1 & w^6 & w^{12} & w^{18} & w^{24} & w^{30} & w^{36} & w^{42} \\ 1 & w^7 & w^{14} & w^{21} & w^{28} & w^{35} & w^{42} & w^{49} \\ \hline \end{array} = \begin{array}{|cccccccc|}\hline 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & w & w^2 & w^3 & w^4 & w^5 & w^6 & w^7 \\ 1 & w^2 & w^4 & w^6 & 1 & w^2 & w^4 & w^6 \\ 1 & w^3 & w^6 & w & w^4 & w^7 & w^2 & w^5 \\ 1 & w^4 & 1 & w^4 & 1 & w^4 & 1 & w^4 \\ 1 & w^5 & w^2 & w^7 & w^4 & w & w^6 & w^3 \\ 1 & w^6 & w^4 & w^2 & 1 & w^6 & w^4 & w^2 \\ 1 & w^7 & w^6 & w^5 & w^4 & w^3 & w^2 & w \\ \hline \end{array}$$

Signal flow diagram

The calculation operations which correspond to the matrix representation of FFT can be showed in a signal flow diagram.

T_3		T_2		T_1		T_S	
1 0 0 0	1 0 0 0	1 0 1 0	0 0 0 0	1 1 0 0	0 0 0 0	1 0 0 0	0 0 0 0
0 1 0 0	0 w 0 0	0 1 0 w^2	0 0 0 0	1 -1 0 0	0 0 0 0	0 0 0 0	1 0 0 0
0 0 1 0	0 0 w^2 0	1 0 -1 0	0 0 0 0	0 0 1 1	0 0 0 0	0 0 1 0	0 0 0 0
0 0 0 1	0 0 0 w^3	0 1 0 $-w^2$	0 0 0 0	0 0 1 -1	0 0 0 0	0 0 0 0	0 0 1 0
1 0 0 0	-1 0 0 0	0 0 0 0	1 0 1 0	0 0 0 0	1 1 0 0	0 1 0 0	0 0 0 0
0 1 0 0	0 $-w$ 0 0	0 0 0 0	0 1 0 w^2	0 0 0 0	1 -1 0 0	0 0 0 0	0 1 0 0
0 0 1 0	0 0 $-w^2$ 0	0 0 0 0	1 0 -1 0	0 0 0 0	0 0 1 1	0 0 0 1	0 0 0 0
0 0 0 1	0 0 0 $-w^3$	0 0 0 0	0 1 0 $-w^2$	0 0 0 0	0 0 1 -1	0 0 0 0	0 0 0 1



- Matrices T_1, T_2 and T_3 contain exactly two non-zero elements in each row.

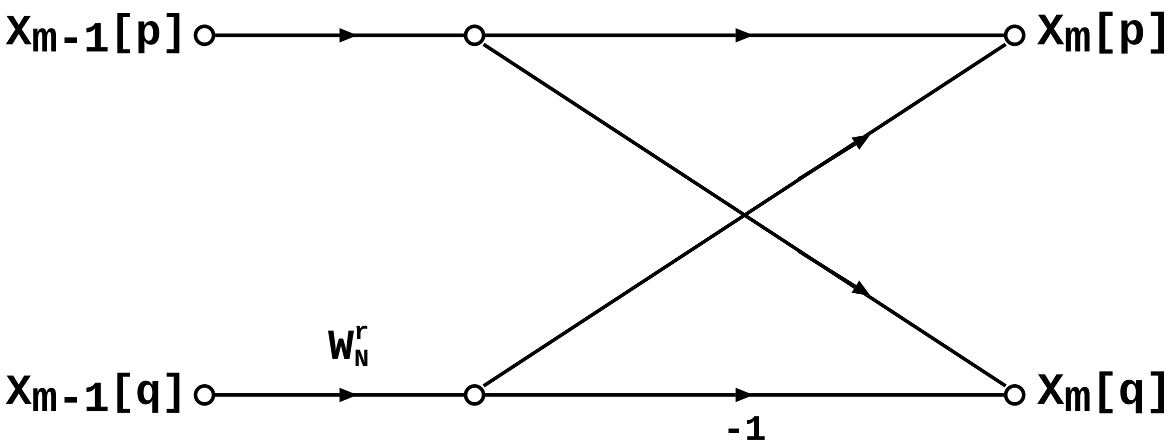
- Non-zero elements are realizing the Butterfly Operation.

- Matrix T_1: step width of the Butterfly Operation is 1
  Matrix T_2: step width of the Butterfly Operation is 2
  Matrix T_3: step width of the Butterfly Operation is 4

- step widths can be found:
  - in signal flow diagram
  - distance between the non-zero elements in T_1, T_2 and T_3

Butterfly Operation

- Signal flow diagram and matrix representation of the FFT are based on the following basic operation:



- For two input values $X_{m-1}[p]$ and $X_{m-1}[q]$ this operation produces two output values $X_m[p]$ and $X_m[q]$. The output values are thereby a linear combination of the input values.

- Because of the flow graph, the operation is called "Butterfly Operation".

$$
\begin{aligned}
X_m[p] &= X_{m-1}[p] + W_N^r \, X_{m-1}[q] \\
X_m[q] &= X_{m-1}[p] - W_N^r \, X_{m-1}[q]
\end{aligned}
$$

$$
\begin{bmatrix} X_m[p] \\ X_m[q] \end{bmatrix} = \begin{bmatrix} 1 & W_N^r \\ 1 & -W_N^r \end{bmatrix} \begin{bmatrix} X_{m-1}[p] \\ X_{m-1}[q] \end{bmatrix}
$$

Bit Reversal

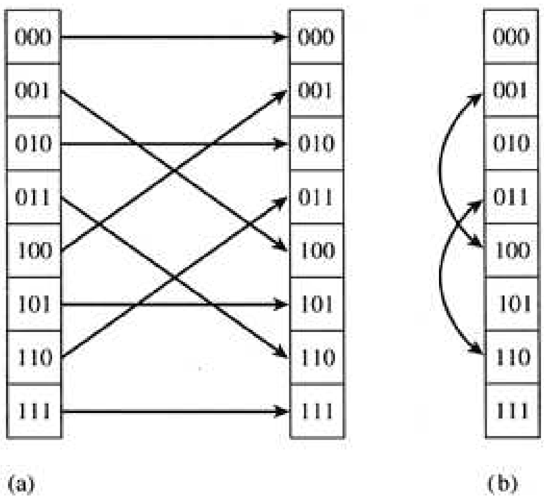
- The matrix representation of the FFT uses a sorting matrix, i.e. the signal which is to be transformed is at first resorted.

- Example for $N = 8$:

n	Binary representation	Reversed	n'
0	000	000	0
1	001	100	4
2	010	010	2
3	011	110	6
4	100	001	1
5	101	101	5
6	110	011	3
7	111	111	7

- Bit Reversal is a necessary part of the FFT–Algorithm.

Bit Reversal for $N = 2^3$



2.16 FFT Implementation

- Fortran version

```
C      adapted from:  Oppenheim, Schafer p.  608
C SUBROUTINE FFT_DecimationInTime (X, ld_n) *********************************
C ***************************************************************************
       PARAMETER PI = 3.14159265358979
       PARAMETER N_max = 2048

       COMPLEX X(N_max)                     !  array for input AND output
       COMPLEX Temp                         !  temporary storage
       COMPLEX W_uni                        !  root of unity
       COMPLEX W_pow                        !  powers of W_uni
       INTEGER N, ld_N, ip, iq, iqbeq, j, k, i_exp, istp

       N = 2**ld_n
       IF (N.GT.N_max) STOP
C BIT Reversed Sorting *********************************************************
       j = 1
       DO i = 1, N-1
            IF (i.LT.j) THEN                !  swap X(j) and X(i)
                 Temp = X(j)
                 X(j) = X(i)
                 X(i) = Temp
            ENDIF
            k = N/2
            DO WHILE (k.LT.j)
                 j = j - k
                 k = k / 2
            ENDDO
            j = j + k
       ENDDO
C End of Bit Reversed Sorting **************************************************

C FFT Butterfly Operations *****************************************************
       DO i=1, ld_N
            i_exp = 2**i                    !  exponent
            istp = i_exp/2                  !  stepsize
            W_pow = (1.0,0.0)
            W_uni = CMPLX (COS (PI/FLOAT(istp)), -SIN(PI/FLOAT(istp)))
            DO ipbeg = 1, istp
                 DO ip = ipbeg, N, i_exp
                      iq = ip + istp
                      Temp = X(iq) * W_pow
                      X(iq) = X(iq) - Temp
                      X(ip) = X(iq) + Temp
                 ENDDO
                 W_pow = W_pow * W_uni
            ENDDO
       ENDDO
C End of FFT Butterfly Operations **********************************************
       RETURN
       END
```

Explanations about Fortran Program

Two program parts:

1. Bit Reversal
2. Butterfly Operations
   - 3 loops with variables i, ipbeg, ip are controlling the execution of the Butterfly operations
   - outer loop i:
     i specifies the level of the FFT
   - With exception of the first level, Butterfly operations are "nested". Therefore two loops are used for the Butterfly operations within one level.
   - middle loop, ipbeg:
     ipbeg: goes over the "nested" Butterfly operations
     i=1: ipbeg=1
     i=2: ipbeg=1,2
     i=3: ipbeg=1,2,3,4
     iqbeg: specifies the sequence of starting points for inner loop
   - inner loop, ip:
     ip: specifies the first element of the Butterfly operation
     istp: step width of the Butterfly operation
     iq=ip+istp: specifies the second element for Butterfly operation
     inner loop is "started" once per "nesting"

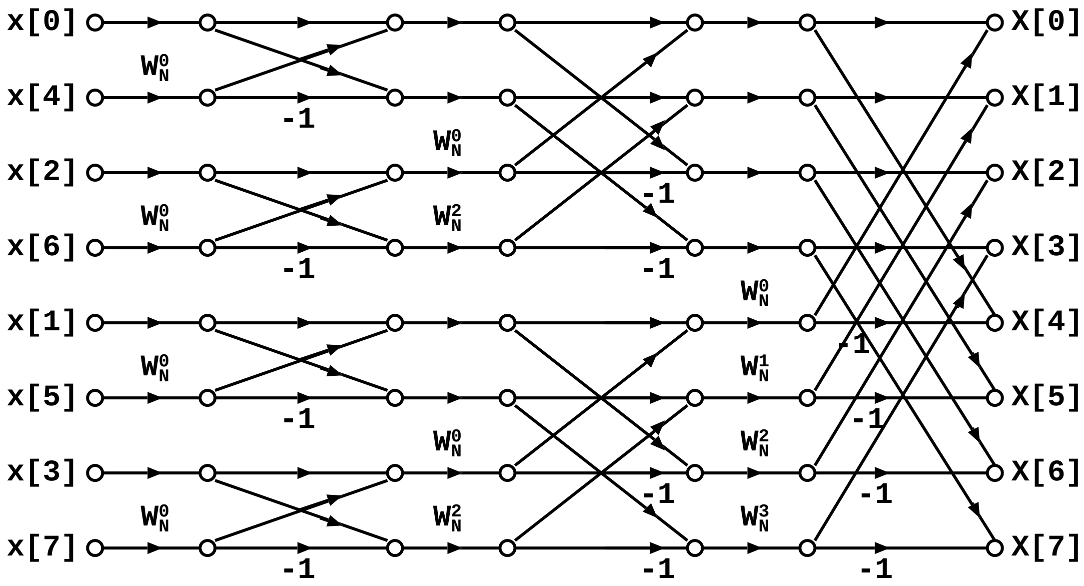


Figure 2.16: Flow diagram of an 8–point–FFT using Butterfly operations.

• C version (from Numerical Recipes in C)

```
#include <math.h>
#define SWAP(a,b) tempr=(a);(a)=(b);(b)=tempr

void four(float data[], unsigned long nn, int isign)
 Replaces data[1..2*nn] by its discrete Fourier transform, if isign is input as 1; or replaces
 data[1..2*nn] by nn times its inverse discrete Fourier transform if insign is input as -1.
 data is a complex array of lenght nn or, equivalently, a real array of lenght 2*nn. nn MUST
 be an whole number power of 2 (this is not checked for!).
{
   unsigned long n, mmax, m, j, istep, i;
   double wtemp, wr, wpr, wpi, wi, theta;    Double prec. for the trigonometric recurrences.
   float tempr, tempi;
   n=nn << 1;
   j=1;
   for (i=1; i<n; i+=2) {                     This is the bit-reversal section of the routine.
       if (j > i) {
          SWAP (data[j], data[i]);             Exchange the two complex numbers.
          SWAP (data[j+1], data[i+1]);
       }
       m=n >> 1;
       while (m >= 2 && j > m) {
          j -= m;
          m >>= 1;
      }
      j += m;
   }
   Here begins the Danielson-Lanczos section of the routine.
   mmax=2;
   while (n > mmax) {                         Outer loop executed log2 nn times
       istep=mmax << 1;
       theta=isign*(6.28318530717959/mmax);   Initialise the trigonometric recurrence.
       wtemp=sin(0.5*theta):
       wpr = -2.0*wtemp*wtemp;
       wpi=sin(theta);
       wr=1.0;
       wi=0.0;
       for (m=1;m <mmax;m+=2) {               Here are the two nested loops.
          for (i=m;i<=n;i+=istep) {
             j=i+mmax;                          This is the Danielson-Lanczos formula:
             tempr=wr*data[j]-wi*data[j+1];
             tempi=wr*data[j+1]+wi*data[j];
             data[j]=data[i]-tempr;
             data[j+1]=data[i+1]-tempi;
             data[i] += tempr;
             data[i+1] += tempi;
          }
          wr=(wtemp=wr)*wpr-wi*wpi+wr;         Trigonometric recurrence
          wi=wi*wpr+wtemp*wpi+wi;
       }
       mmax=istep
   }
}
```

- Input and output arrays (C version)

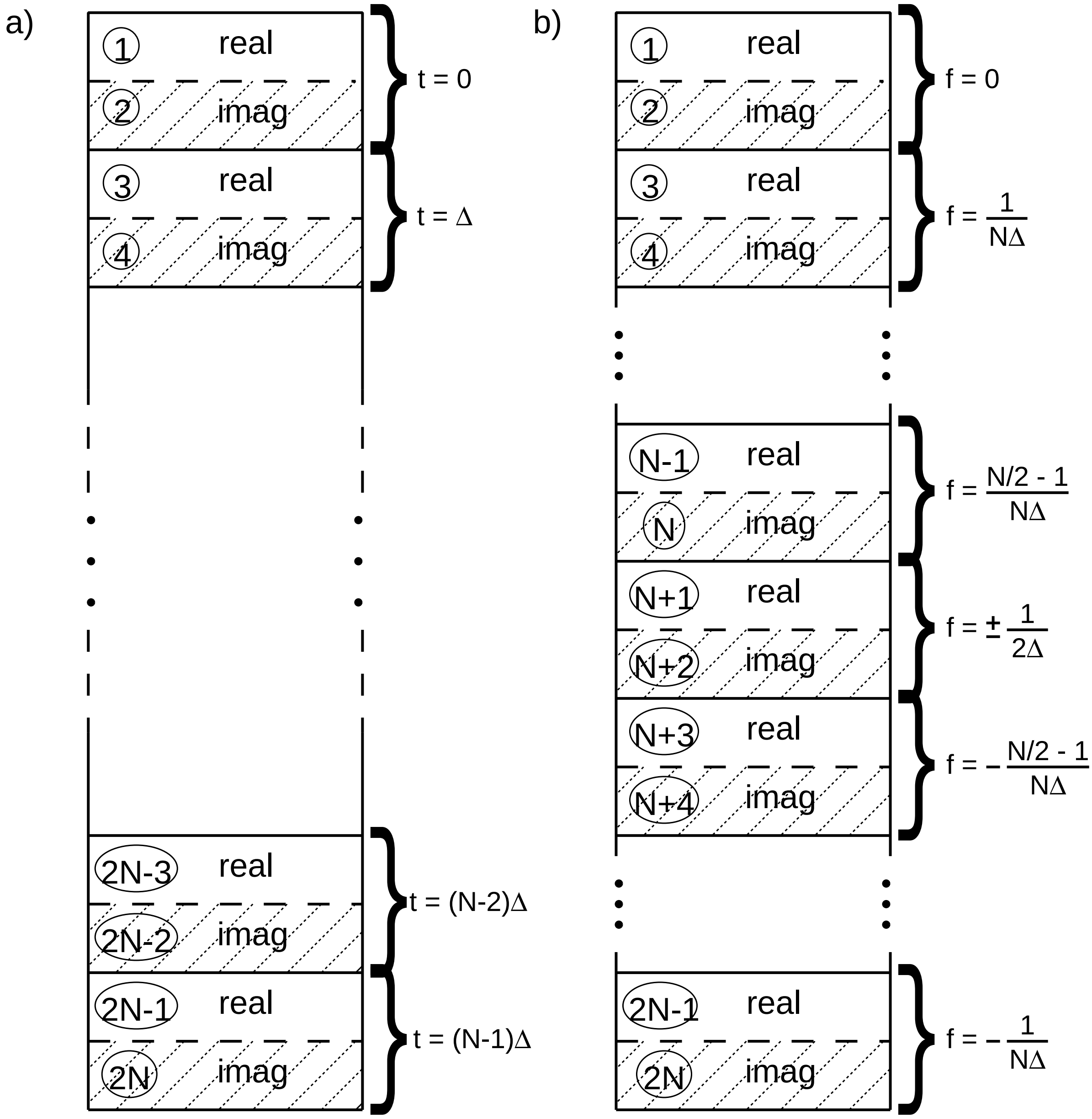


Figure 2.17: Input and output arrays of an FFT. a) The input array contains N (N is power of 2) complex input values in one `real` array of the length $2N$. with alternating real and imaginary parts. b) The output array contains complex Fourier spectrum at N frequency values. Again alternating real and imaginary parts. The array begins with the zero-frequency and then goes up to the highest frequency followed with values for the negative frequencies.

Finite convolution: complexity by the application of FFT

Estimation of number of necessary multiplications for a convolution of $x[n]$ and $h[n]$

$x[n]$: N_x non-zero values

$h[n]$: N_h non-zero values

<table>
<tr><td colspan="3">Realisation</td></tr>
<tr><td>direct implementation</td><td>DFT</td><td>FFT</td></tr>
<tr><td rowspan="6">$N_x \cdot N_h$</td><td colspan="2">transformation</td></tr>
<tr><td>$(N_x + N_h)^2$</td><td>$\frac{N_x+N_h}{2} \log_2(N_x + N_h)$</td></tr>
<tr><td colspan="2">multiplication in frequency domain</td></tr>
<tr><td>$N_x + N_h$</td><td>$N_x + N_h$</td></tr>
<tr><td colspan="2">inverse transformation</td></tr>
<tr><td>$(N_x + N_h)^2$</td><td>$\frac{N_x+N_h}{2} \log_2(N_x + N_h)$</td></tr>
</table>

2.17 Cyclic Matrices and Fourier Transform

The Fourier transform plays a significant role for so-called cyclic matrices (cf. chapter 3.8):

$$
\begin{array}{ccc}
 & \begin{array}{ccccccc} 0 & & & n & & & N-1 \end{array} & \\
H \;=\; & \left[\begin{array}{ccccccc}
h_0 & h_1 & h_2 & & \cdots & & h_{N-1} \\
h_{N-1} & \ddots & \ddots & \ddots & & & h_{N-2} \\
h_{N-2} & \ddots & \ddots & \ddots & \ddots & & \vdots \\
 & & \ddots & \ddots & \ddots & \ddots & \\
\vdots & & & \ddots & \ddots & \ddots & h_2 \\
 & & & & \ddots & \ddots & h_1 \\
h_1 & & & & & h_{N-1} & h_0
\end{array}\right] & \begin{array}{c} 0 \\ \\ \\ m \\ \\ \\ N-1 \end{array}
\end{array}
$$

so: $$H_{mn} \;=\; h_{(n-m)\mathrm{mod}N}$$

with so-called kernel vector $(h_0, h_1 \ldots, h_{N-1})^T$ and $h_n \in \mathbb{C}$, mostly $h_n \in \mathbb{R}$.

Remark: using cyclic matrices it is possible to define cyclic i.e. periodic convolutions and to build a cyclic variant of the system theory.

The eigenvectors of a cyclic matrix can be obtained from the columns of DFT matrix:

$$
\begin{array}{cc}
 & \begin{array}{ccccc} 0 & & n & & N-1 \end{array} \\
\begin{array}{c} 0 \\ \\ \\ m \\ \\ \\ N-1 \end{array} & \left[\begin{array}{ccccc}
w^{0\cdot 0} & \cdots & w^{n\cdot 0} & \cdots & w^{(N-1)\cdot 0} \\
\vdots & & w^{n\cdot 1} & & \vdots \\
\vdots & & \vdots & & \\
 & \cdots & \vdots & \cdots & \vdots \\
\vdots & & \vdots & & \\
 & \cdots & \vdots & \cdots & \vdots \\
\vdots & & w^{n\cdot(N-1)} & & \vdots
\end{array}\right]
\end{array}
\quad \text{where } w \;=\; e^{-\frac{2\pi j}{N}}
$$

The eigenvalues λ_k of a cyclic matrix with the kernel vector $(h_0, h_1 \ldots, h_{N-1})^T$ are:

$$\lambda_k \;=\; \sum_{n=0}^{N-1} h_n \, e^{-\frac{2\pi j}{N} \cdot k \cdot n} \qquad \text{where} \quad k \;=\; 0, 1 \ldots, N-1$$

The representation is oriented on L. Berg: Linear equation systems with band structure (page 52 ff).

The special case of a cyclic matrix when the kernel vector h is symmetric and real is especially interesting for many applications:

$$h_n \;=\; h_{N-n} \text{ and } h_n \in \mathbb{R}$$

That means that the matrix is *real, symmetric and cyclic:*

$$\begin{array}{cc} & \begin{array}{ccccccc} 0 & & & & & & N-1 \end{array} \\ \begin{array}{c} 0 \\ \\ \\ \\ \\ \\ N-1 \end{array} & \begin{bmatrix} h_0 & h_1 & h_2 & & \ldots & h_2 & h_1 \\ h_1 & \ddots & \ddots & \ddots & & & h_2 \\ & \ddots & \ddots & \ddots & \ddots & & h_3 \\ & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \\ h_2 & h_3 & & & \ddots & \ddots & h_1 \\ h_1 & h_2 & & & & h_1 & h_0 \end{bmatrix} \end{array}$$

For such a matrix is valid:

a) the eigenvalues λ_k are:

$$\lambda_k \;=\; \sum_{n=0}^{N-1} h_n \,\cos\left(\frac{2\pi kn}{N}\right)$$

b) The eigenvectors can be obtained from the columns of the “discrete cos–matrix”:

$$\begin{array}{cc} & n \\ m & \left[\begin{array}{ccc} \quad & \begin{pmatrix} \vdots \\ \vdots \\ \vdots \\ \cos\left(\frac{2\pi nm}{N}\right) \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} & \quad \end{array} \right] \end{array}$$

Application: diagonalisation of covariance matrices, e.g. for coding of image and speech signals.

Proof:

We will prove that

$$v_k = \begin{pmatrix} \cos\frac{2\pi k\cdot 0}{N} \\ \cos\frac{2\pi k\cdot 1}{N} \\ \vdots \\ \vdots \\ \vdots \\ \cos\frac{2\pi k\cdot (N-1)}{N} \end{pmatrix}$$

are eigenvectors of the given matrix with eigenvalues λ_k.

For a symmetric cyclic matrix is valid:

$$H_{mn} := h_{(n-m)\mathrm{mod}N} \quad \text{where} \quad h_n = h_{N-n}$$

One row results in (for odd N):

$$\sum_n H_{mn}\cdot v_{kn} = h_0\cdot\cos\frac{2\pi km}{N} + \sum_{l=1}^{\frac{N-1}{2}} h_l\cdot\left(\cos\frac{2\pi k(m-l)}{N} + \cos\frac{2\pi k(m+l)}{N}\right)$$

For even N only *one* term for $l = (N-1)/2$ can go into the sum.

According to addition theorem:

$$\cos(x+y)+\cos(x-y) = 2\cdot\cos x\cdot\cos y$$

With $x = 2\pi km/N$ and $y = 2\pi kl/N$ follows:

$$\begin{aligned}
\sum_n H_{mn}\cdot v_{kn} &= h_0\cdot\cos\frac{2\pi km}{N} + 2\sum_{l=1}^{\frac{N-1}{2}} h_l\cdot\left(\cos\frac{2\pi kl}{N}\cdot\cos\frac{2\pi km}{N}\right) \\
&= \underbrace{\left[h_0 + 2\sum_{l=1}^{\frac{N-1}{2}} h_l\cos\frac{2\pi kl}{N}\right]}_{\lambda_k}\cdot\underbrace{\cos\frac{2\pi km}{N}}_{v_{km}} \\
&= \lambda_k\cdot v_{km}
\end{aligned}$$

Excursion: Toeplitz matrices

We consider only quadratic real matrices:

$$H \in \mathbb{R}^{N \times N} \qquad \text{with } H_{mn} \in \mathbb{R}$$

a) H is a (general) *Toeplitz matrix*, if:

$$H_{mn} = h_{n-m}$$

i.e.

$$H = \begin{bmatrix} h_0 & h_1 & h_2 & & \dots & h_{N-2} & h_{N-1} \\ h_{-1} & \ddots & \ddots & \ddots & & & h_{N-2} \\ h_{-2} & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots & \ddots & h_2 \\ h_{2-N} & & & \ddots & \ddots & \ddots & h_1 \\ h_{1-N} & h_{2-N} & & & h_{-2} & h_{-1} & h_0 \end{bmatrix}$$

b) For a *symmetric Toeplitz matrix* then holds:

$$H_{mn} = h_{|n-m|}$$

i.e.

$$H = \begin{bmatrix} h_0 & h_1 & h_2 & & \dots & h_{N-2} & h_{N-1} \\ h_1 & \ddots & \ddots & \ddots & & & h_{N-2} \\ h_2 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots & \ddots & h_2 \\ h_{N-2} & & & \ddots & \ddots & \ddots & h_1 \\ h_{N-1} & h_{N-2} & & & h_2 & h_1 & h_0 \end{bmatrix}$$

c) A *cyclic matrix* can be obtained by special choice:

$$H_{mn} = h_{(n-m)\mathrm{mod}N}$$

$$H = \begin{bmatrix} h_0 & h_1 & h_2 & & \dots & h_{N-2} & h_{N-1} \\ h_{N-1} & \ddots & \ddots & \ddots & & & h_{N-2} \\ h_{N-2} & \ddots & \ddots & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & \ddots & h_2 \\ h_2 & & & & \ddots & \ddots & h_1 \\ h_1 & h_2 & & & & h_{N-1} & h_0 \end{bmatrix}$$

Also valid: each cyclic matrix is a Toeplitz matrix.

d) A *symmetric cyclic matrix* can be obtained by the following choice of the kernel vectors h:

$$h_n = h_{N-n} \quad \text{for } n = 0, \dots, N$$

For example, we obtain for $N = 8$:

$$H = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 & h_4 & h_3 & h_2 & h_1 \\ h_1 & h_0 & h_1 & h_2 & h_3 & h_4 & h_3 & h_2 \\ h_2 & h_1 & h_0 & h_1 & h_2 & h_3 & h_4 & h_3 \\ h_3 & h_2 & h_1 & h_0 & h_1 & h_2 & h_3 & h_4 \\ h_4 & h_3 & h_2 & h_1 & h_0 & h_1 & h_2 & h_3 \\ h_3 & h_4 & h_3 & h_2 & h_1 & h_0 & h_1 & h_2 \\ h_2 & h_3 & h_4 & h_3 & h_2 & h_1 & h_0 & h_1 \\ h_1 & h_2 & h_3 & h_4 & h_3 & h_2 & h_1 & h_0 \end{bmatrix}$$

Chapter 3

Spectral analysis

- Overview:

  3.1 Features for Speech Recognition

  3.2 Short Time Analysis and Windowing

  3.3 Autocorrelation function and Power Spectral Density

  3.4 Spectrograms

  3.5 Filter Bank Analysis

  3.6 Mel–Scale

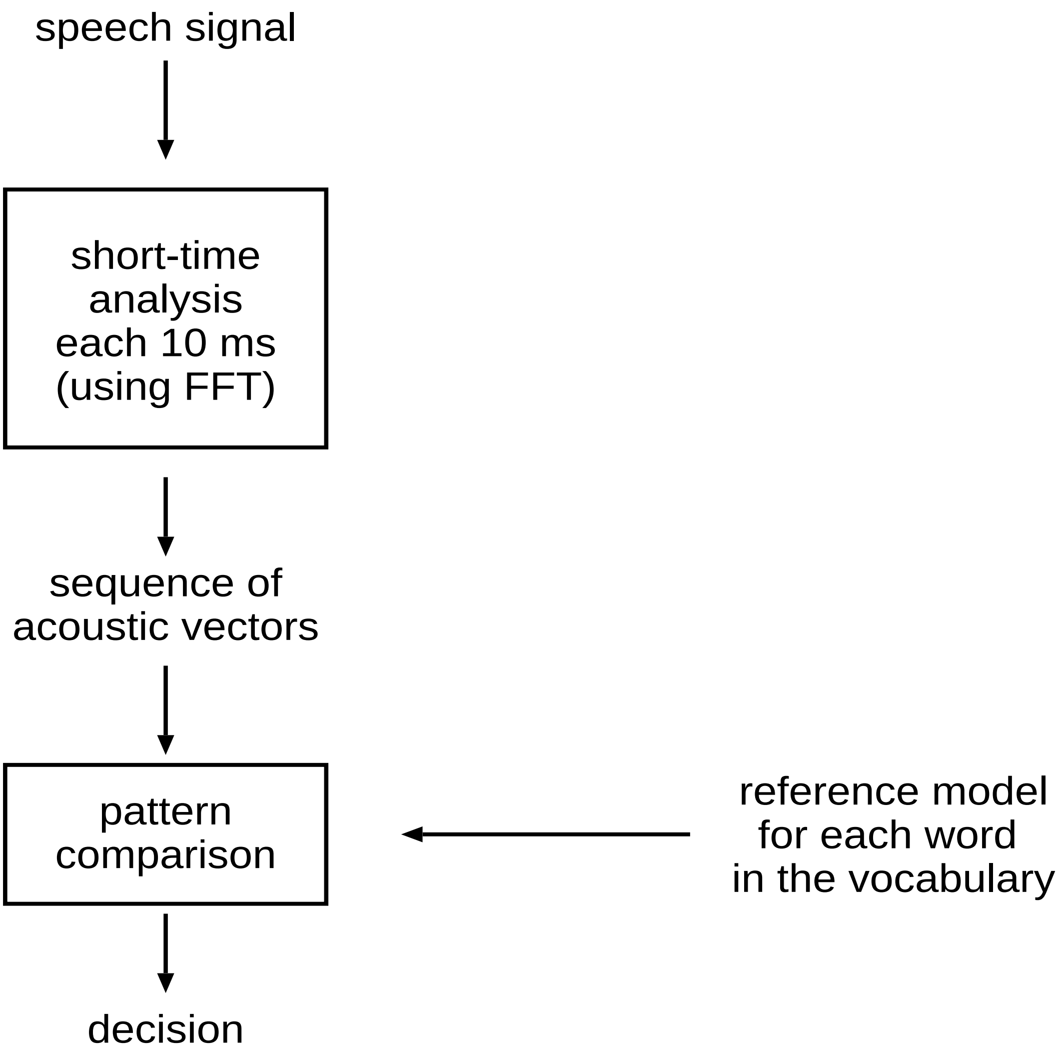
  3.7 Cepstrum
  - Cepstrum Calculation from Filter Bank Output
  - Mel–Cepstrum according to Davis and Mermelstein

  3.8 Statistical Interpretation of Cepstrum Transformation

  3.9 Energy in acoustic Vector

3.1 Features for Speech Recognition

Architecture of an automatic speech recognition system



Short time analysis:

- window length 10–40ms
- sampling period 10–20ms
- in case of sampling rate of 10kHz:
  - Window: 100–400 samples
  - sampling period (frame shift): 100–200 samples

Recommended windows:

- Hamming
- Kaiser
- Blackman

Model parameters:

- Energy, intensity (“loudness”)
- Fundamental frequency (“height”)
- Spectral parameters (“colour”, “smoother” amplitude spectrum)

Goal:

- Ideally: Real features for the recognition
- In practice: Data reduction, i.e. compact description of the speech signal (amplitude spectrum)

Side effect:

- Method also enables coding of speech signals using lowest possible number of bits

Key words:

- Fourier transform: wide band/narrow band, autocorrelation function
- Filter bank
- Cepstrum
- Linear Predictive Coding (LPC) analysis
- Fundamental frequency analysis

3.2 Short Time Analysis and Windowing

- The DFT is defined for signals with finite duration.
- Speech signal $s[n]$:
  quasi stationary, i.e. properties do not change within 20-50 ms.
- Window function $w[n]$:
  Decomposition of the original signal $s[n]$ into (overlapping) segments using a window function $w[n]$:

$$x[n] \quad = \quad s[n] \cdot w[n]$$

where for example

$$w[n] \quad = \quad \begin{cases} 1, & |n| \leq N/2 \\ 0, & \text{otherwise} \end{cases}$$

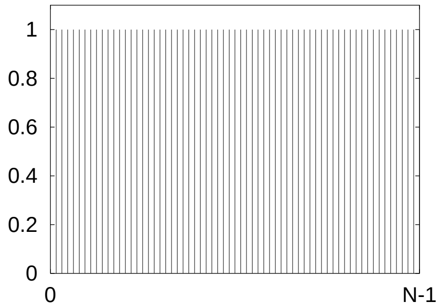
- The windowed signal $x[n]$ is analyzed with a Fourier Transform or DFT.
- The multiplication of the original signal $s[n]$ with the window function $w[n]$ in the time domain corresponds to the convolution of the spectra of two signals $S(e^{j\omega})$ and window function $W(e^{j\omega})$ in the frequency domain:

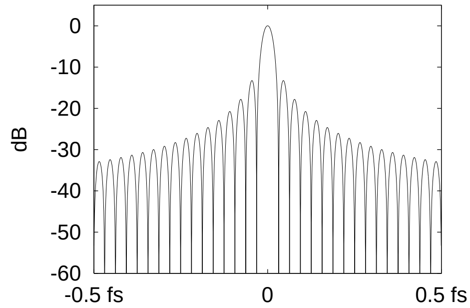
$$X(e^{j\omega}) \;=\; \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} S(e^{j\theta})\, W(e^{j(\omega-\theta)})\, d\theta$$

- This convolution performs a (spectral) smearing in the frequency domain (*leakage*).

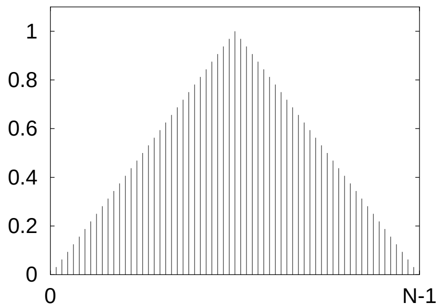
Window function:

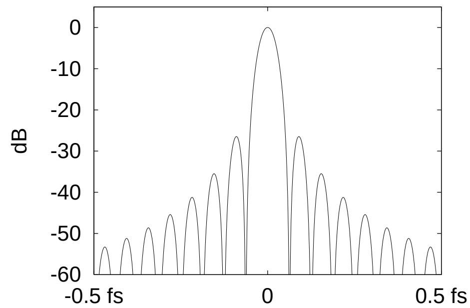
Impulse response:



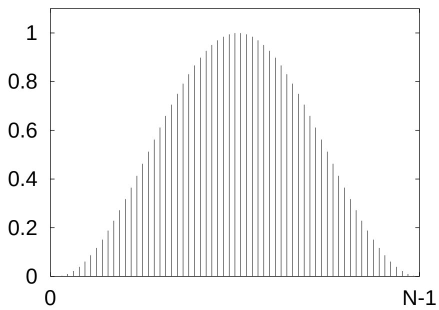


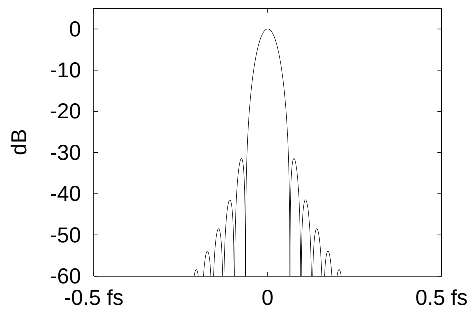
Rectangle



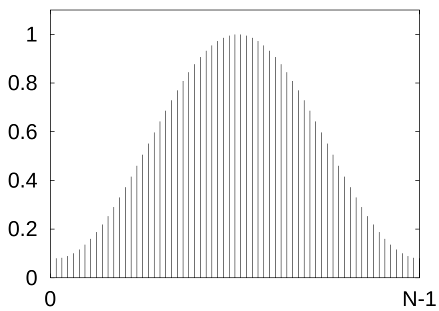


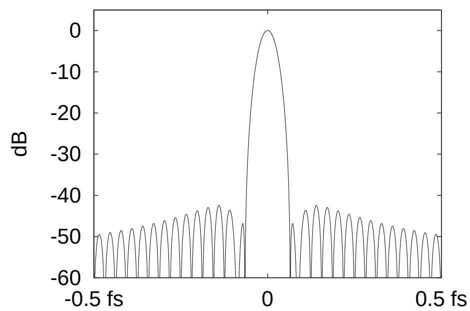
Triangle





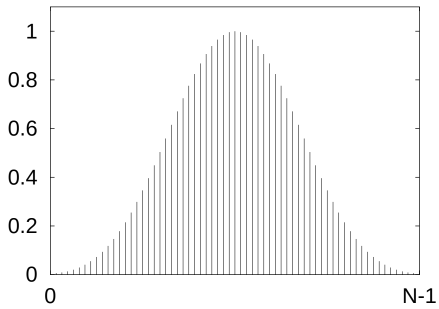
Hanning

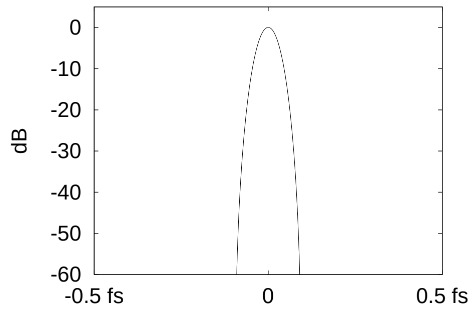




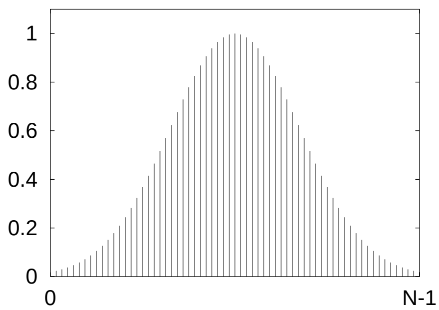
Hamming

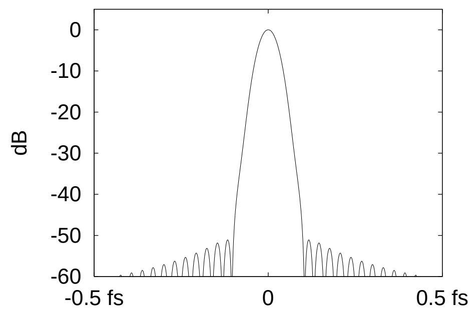
Window function: Impulse response:



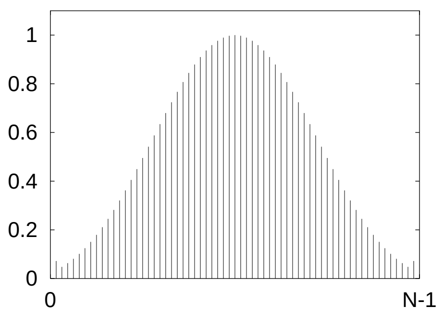


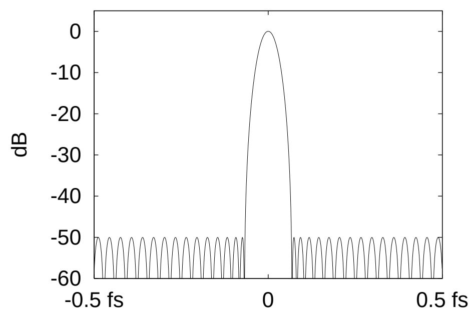
Nuttall





Gauss





Chebyshev

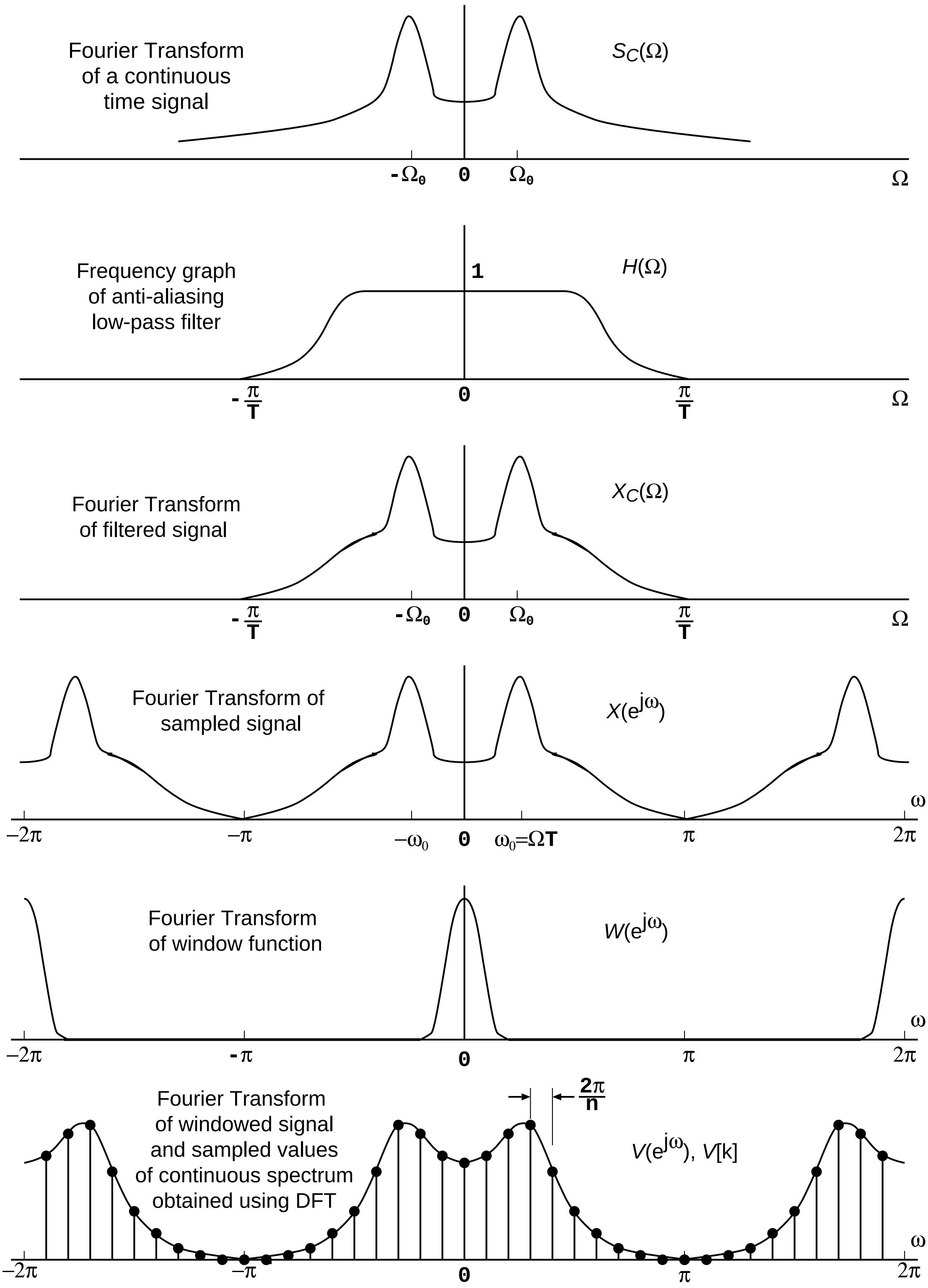


Figure 3.1: Example for the application of the Discrete Fourier Transform (DFT).

Properties of short-time DFT–analysis

Important effects:

- *Picket Fence*
  If not enough sampled values of continuous spectrum are available, spectral sampling can yield delusive results. This problem can be reduced using Zero Padding (inter-space between the coefficients $S[k]$ becomes smaller, i.e. frequency resolution becomes better)

- *Leakage*: Spreading of the line spectrum
  Because the window function is limited in time, a spreaded spectrum is measured instead of the spectrum of the original signal unlimited in time. That means, the line spectrum even becomes spreaded for pure sinusoidal signals.

3 examples of DFT analysis

- we observe a continuous time signal $x(t)$ composed of two sinusoids:

$$x(t) \;=\; A_0\;\cos(\Omega_0\,t)\;+\;A_1\;\cos(\Omega_1\,t) \qquad -\infty < t < \infty$$

- sampling according to sampling theorem (with negligible quantization errors)
- discrete time signal $x[n]$:

$$x[n] \;=\; A_0\;\cos(\omega_0\,n)\;+\;A_1\;\cos(\omega_1\,n) \qquad -\infty < n < \infty$$

where $\omega_0 = \Omega_0 T_S$ and $\omega_1 = \Omega_1 T_S$

- with the window function $w[n]$:

$$v[n] \;=\; A_0\,w[n]\;\cos(\omega_0\,n)\;+\;A_1\,w[n]\;\cos(\omega_1\,n)$$

Intermediate calculations:

$$\begin{aligned} v[n] &= \frac{A_0}{2}\,w[n]\exp(j\,\omega_0\,n)\;+\;\frac{A_0}{2}\,w[n]\exp(-j\,\omega_0\,n) \\ &+ \frac{A_1}{2}\,w[n]\exp(j\,\omega_1\,n)\;+\;\frac{A_1}{2}\,w[n]\exp(-j\,\omega_1\,n) \end{aligned}$$

also modulation principle

- Fourier Transform of the windowed signal:

$$\begin{aligned} V(e^{j\omega}) &= \frac{A_0}{2}\,W(e^{j(\omega-\omega_0)})\;+\;\frac{A_0}{2}\,W(e^{j(\omega+\omega_0)}) \\ &+ \frac{A_1}{2}\,W(e^{j(\omega-\omega_1)})\;+\;\frac{A_1}{2}\,W(e^{j(\omega+\omega_1)}) \end{aligned}$$

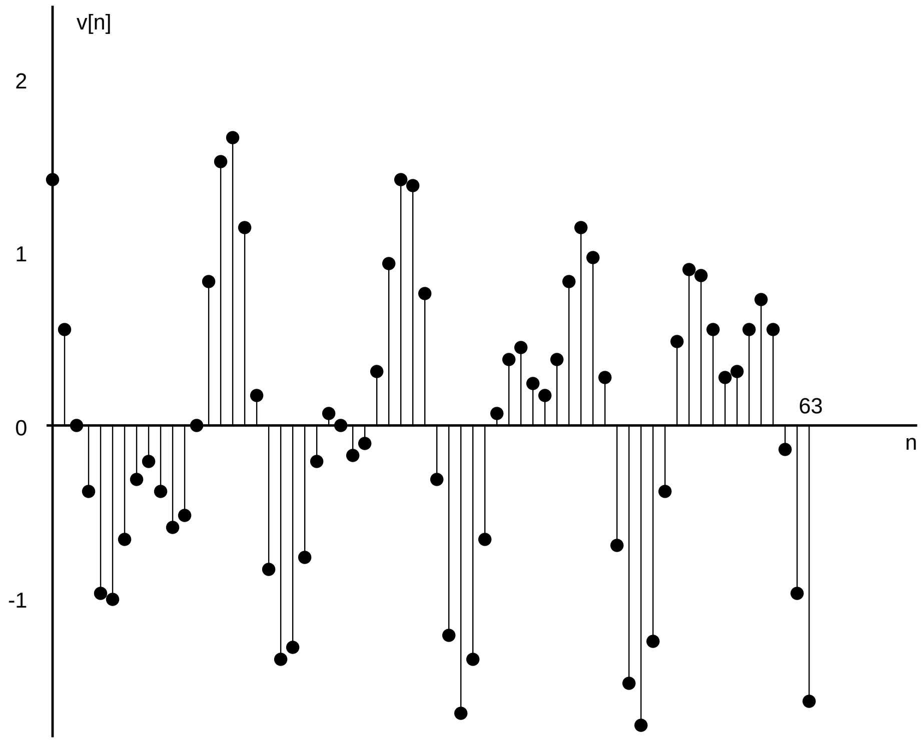
- Assume:

$$\Omega_0 = \frac{2\pi}{14} \cdot 10\text{kHz}, \Omega_1 = \frac{4\pi}{15} \cdot 10\text{kHz}$$

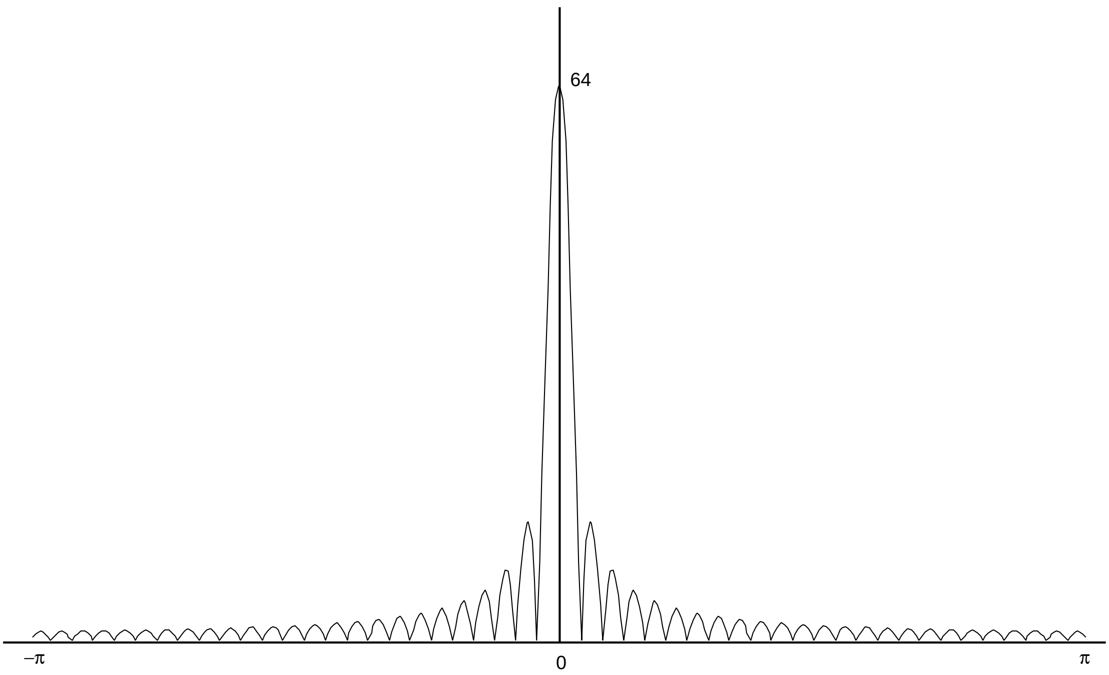
$1/T_S = 10\text{kHz}$, rectangle window with $N = 64$, $A_0 = 1$, $A_1 = 0.75$

- The windowed signal $v[n]$ for the discrete time signal $x(n)$ is therefore:

$$v[n] = \begin{cases} \cos(\frac{2\pi}{14}n) + 0.75\ \cos(\frac{4\pi}{15}n) & : \quad 0 \le n \le 63 \\ & : \\ 0 & : \quad \text{otherwise} \end{cases}$$



- Fourier Transform $W(e^{j\omega})$ of the rectangle window function



Example 1: **Leakage Effect**

Variation of ω_0 and ω_1 resp. Ω_0 and Ω_1
Difference between frequencies ω_0 and ω_1 is reduced gradually

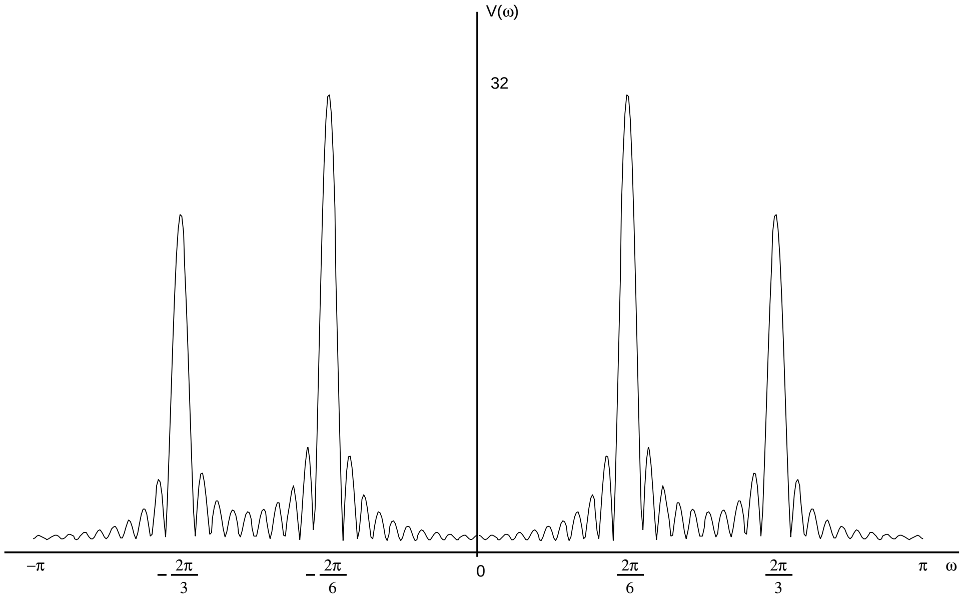
Case 1a:

$$\Omega_0 \;=\; \frac{2\pi}{6}\,10^4\ \mathrm{Hz}, \qquad \Omega_1 \;=\; \frac{2\pi}{3}\,10^4\ \mathrm{Hz}$$

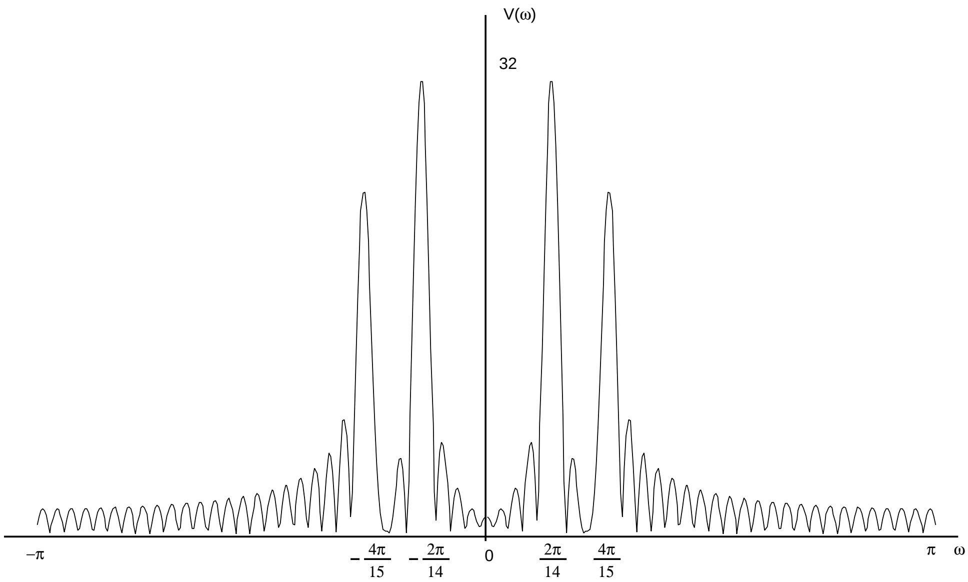
$$\omega_0 \;=\; \Omega_0\,T_S \;=\; \frac{2\pi}{6}\,10^4\ \mathrm{Hz}\ 10^{-4}\ \mathrm{s} \;=\; \frac{2\pi}{6}$$

$$\omega_1 \;=\; \Omega_1\,T_S \;=\; \frac{2\pi}{3}\,10^4\ \mathrm{Hz}\ 10^{-4}\ \mathrm{s} \;=\; \frac{2\pi}{3}$$

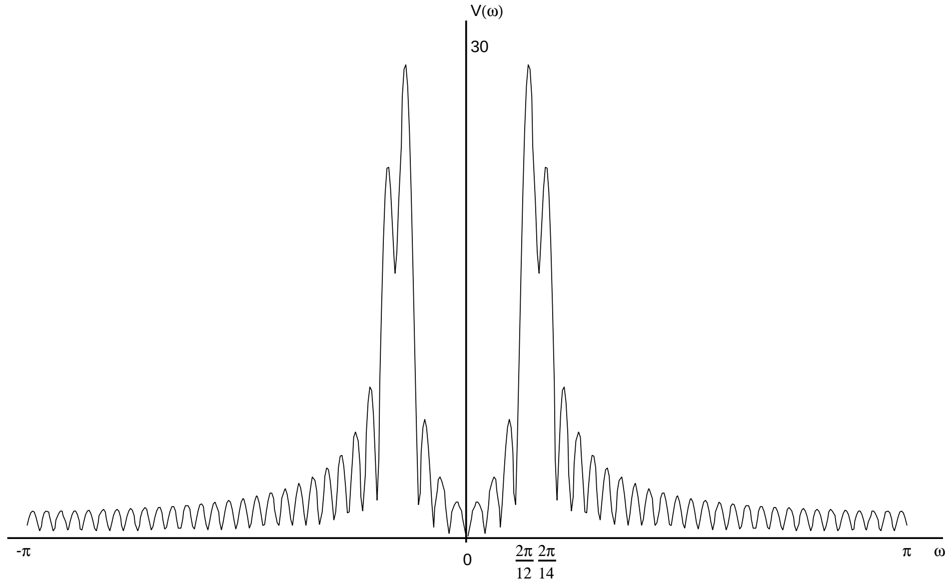
Case 1a (continued): $\omega_0 = \dfrac{2\pi}{6} \qquad \omega_1 = \dfrac{2\pi}{3}$



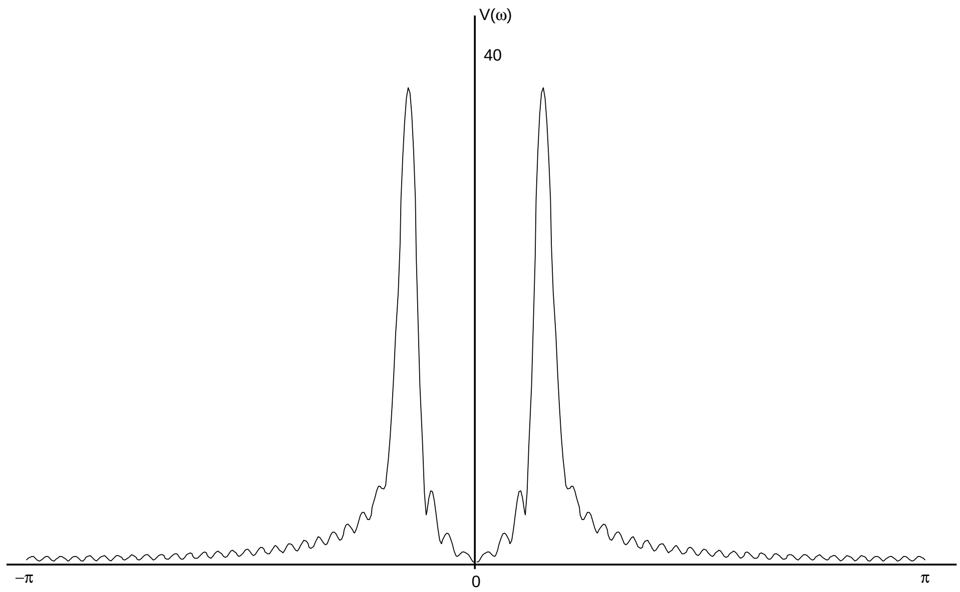
Case 1b: $\omega_0 = \dfrac{2\pi}{14} \qquad \omega_1 = \dfrac{4\pi}{15}$



Case 1c: $\omega_0 = \dfrac{2\pi}{14} \qquad \omega_1 = \dfrac{2\pi}{12}$



Case 1d: $\omega_0 = \dfrac{2\pi}{14} \qquad \omega_1 = \dfrac{4\pi}{25}$



Example 2: **Picket Fence Effect**

DFT gives sampled values of the spectrum of the windowed signal.
Spectral sampling can yield delusive results.

Case 2a:

- Windowed signal $v[n]$:

$$v[n] \ = \ \begin{cases} \cos(\frac{2\pi}{14}n) \ + \ 0.75 \ \cos(\frac{4\pi}{15}n) & : \quad 0 \leq n \leq 63 \\ 0 & : \quad \text{otherwise} \end{cases}$$

- DFT of the length $N = 64$ without Zero Padding a)

b)

c)

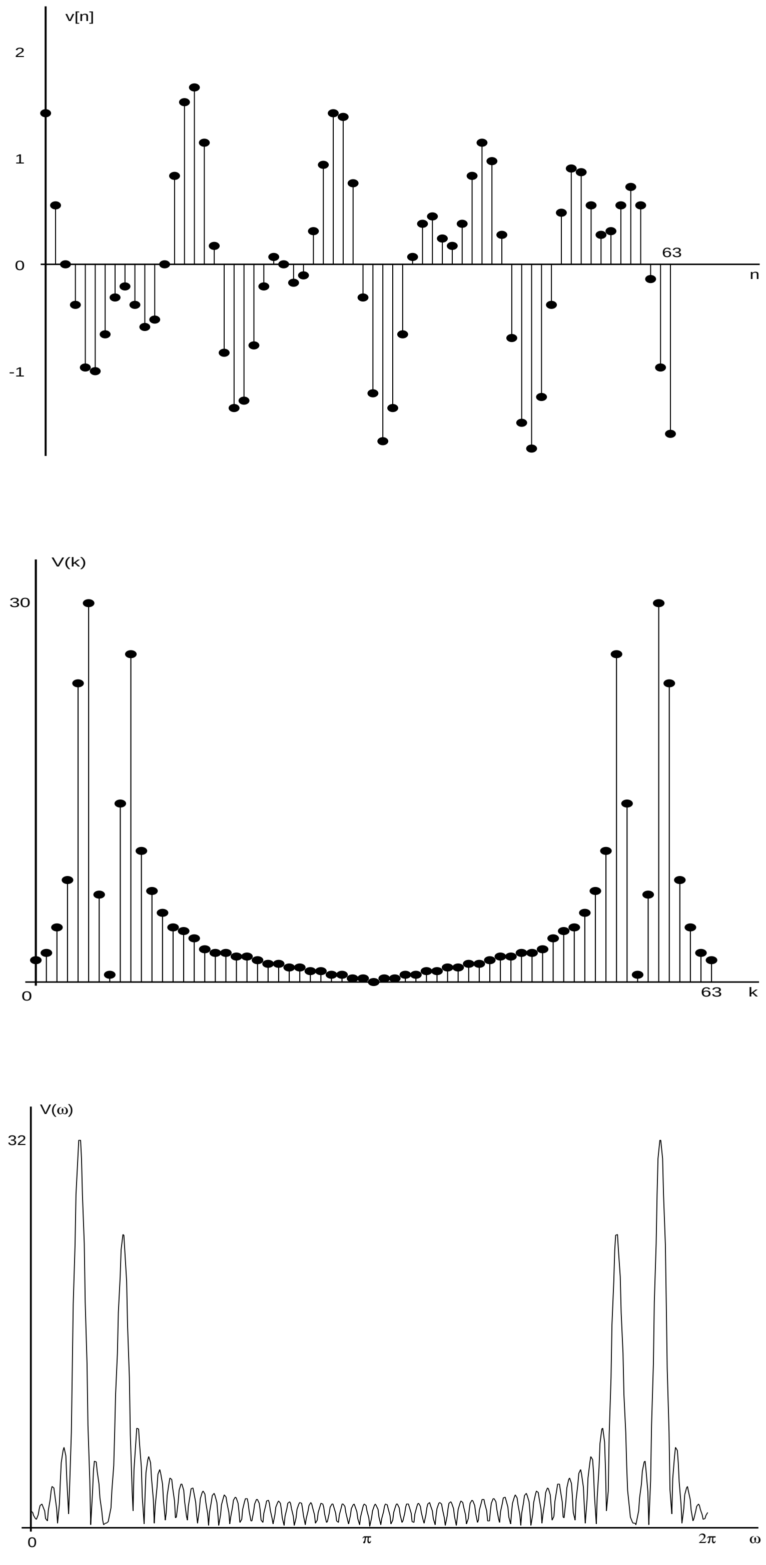


Figure 3.2: a) signal $v[n]$; b) DFT-spectrum $V[k]$; c) Fourier spectrum $V(e^{j\omega})$.

Case 2b:

- In contrast to case 2a, the frequencies of sinusoids are changed only slightly.

- Windowed signal $v[n]$:

$$v[n] = \begin{cases} \cos(\frac{2\pi}{16}n) + 0.75 \cos(\frac{2\pi}{8}n) & : \quad 0 \le n \le 63 \\ 0 & : \quad \text{otherwise} \end{cases}$$

- DFT of the length $N = 64$ without Zero Padding

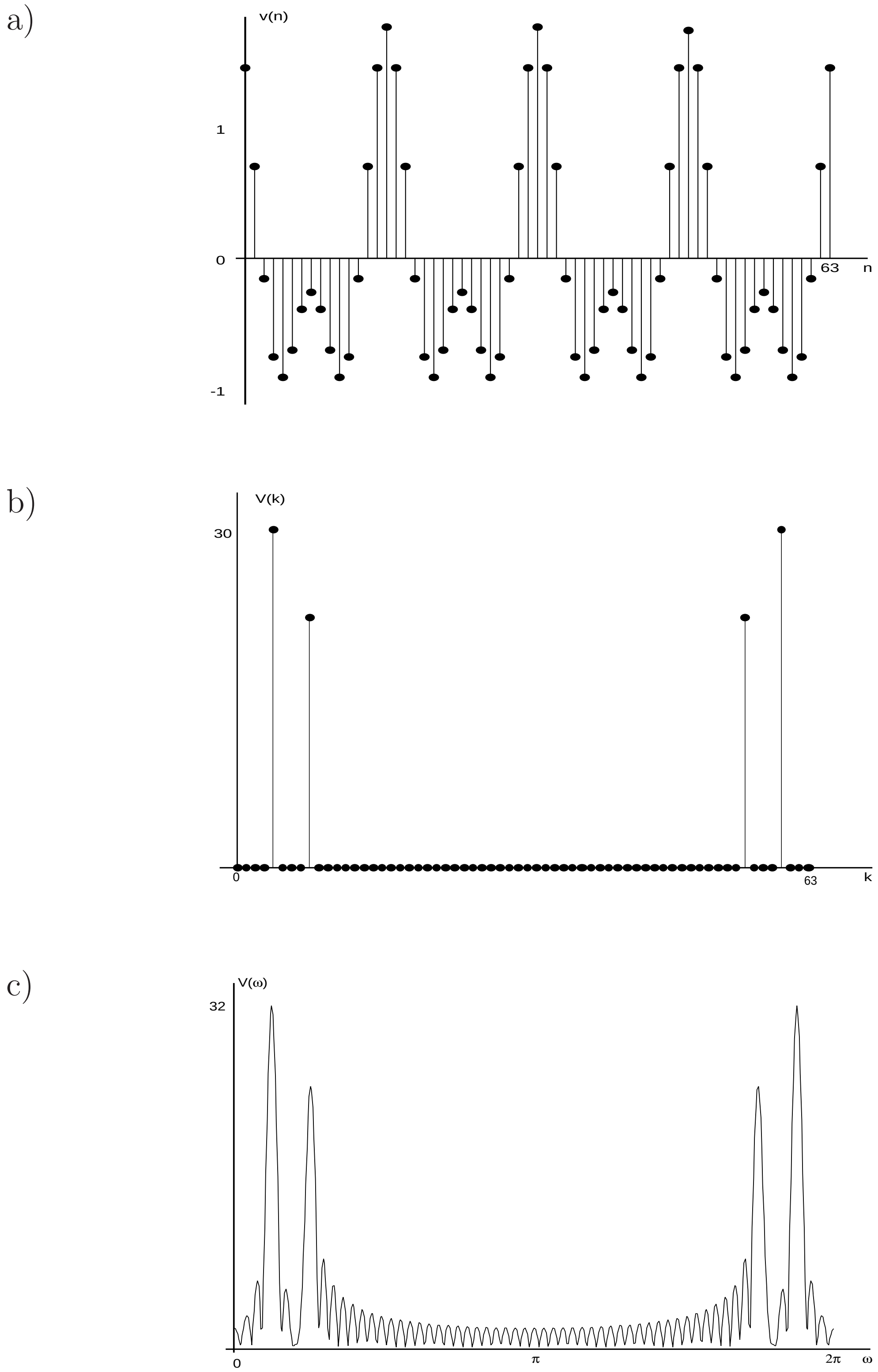


Figure 3.3: a) signal $v[n]$; b) DFT-spectrum $V[k]$; c) Fourier spectrum $V(e^{j\omega})$.

Analysis of Example 2:

- The manifestation of the DFT can be put down to the spectral sampling. Although in Case 2b the windowed signal $v[n]$ contains a significant number of frequencies beyond ω_0 and ω_1, they do not show in the DFT spectrum of length $N = 64$.
- Using a rectangle window, the DFT of the sinusoidal signal gives sharp spectral lines, if the period N of the transformation is a whole multiple of the signal period and no Zero Padding is applied.
- Explanation for the case of a complex exponential function:
  - Assume the signal $x[n]$:

$$x[n] \;=\; \frac{1}{N} \exp(j\, \frac{2\pi}{n_0}\, n)$$

  - Then:

$$X[k] \;=\; \delta(k - \frac{N}{n_0})$$

  - For the DFT of rectangle window holds:

$$W[k] \;=\; \frac{\sin(\pi k)}{\sin(\pi k/N)}$$

  - Convolution theorem for windowed signal $v[n]$ gives:

$$V[k] \;=\; X[k] \,*\, W[k] \;=\; \frac{\sin\left(\pi(k - \frac{N}{n_0})\right)}{\sin\left(\pi(k - \frac{N}{n_0})/N\right)}$$

  - In case of

$$\frac{N}{n_0} \in \mathbb{N}$$

    only the DFT coefficient $k = \dfrac{N}{n_0}$ is non-zero.

Example 2 (continued)

- Assume signal $v[n]$ of Case 2b:

$$v[n] \;=\; \begin{cases} \cos(\frac{2\pi}{16}n) \;+\; 0.75\;\cos(\frac{2\pi}{8}n) & : \quad 0 \le n \le 63 \\ 0 & : \quad \text{otherwise} \end{cases}$$

- In contrast to Case 2b, a DFT with length $N = 128$ is applied (*Zero Padding*).

- Result:

  Using finer sampling, existing additional frequency components emerge.

a)

b)

c)

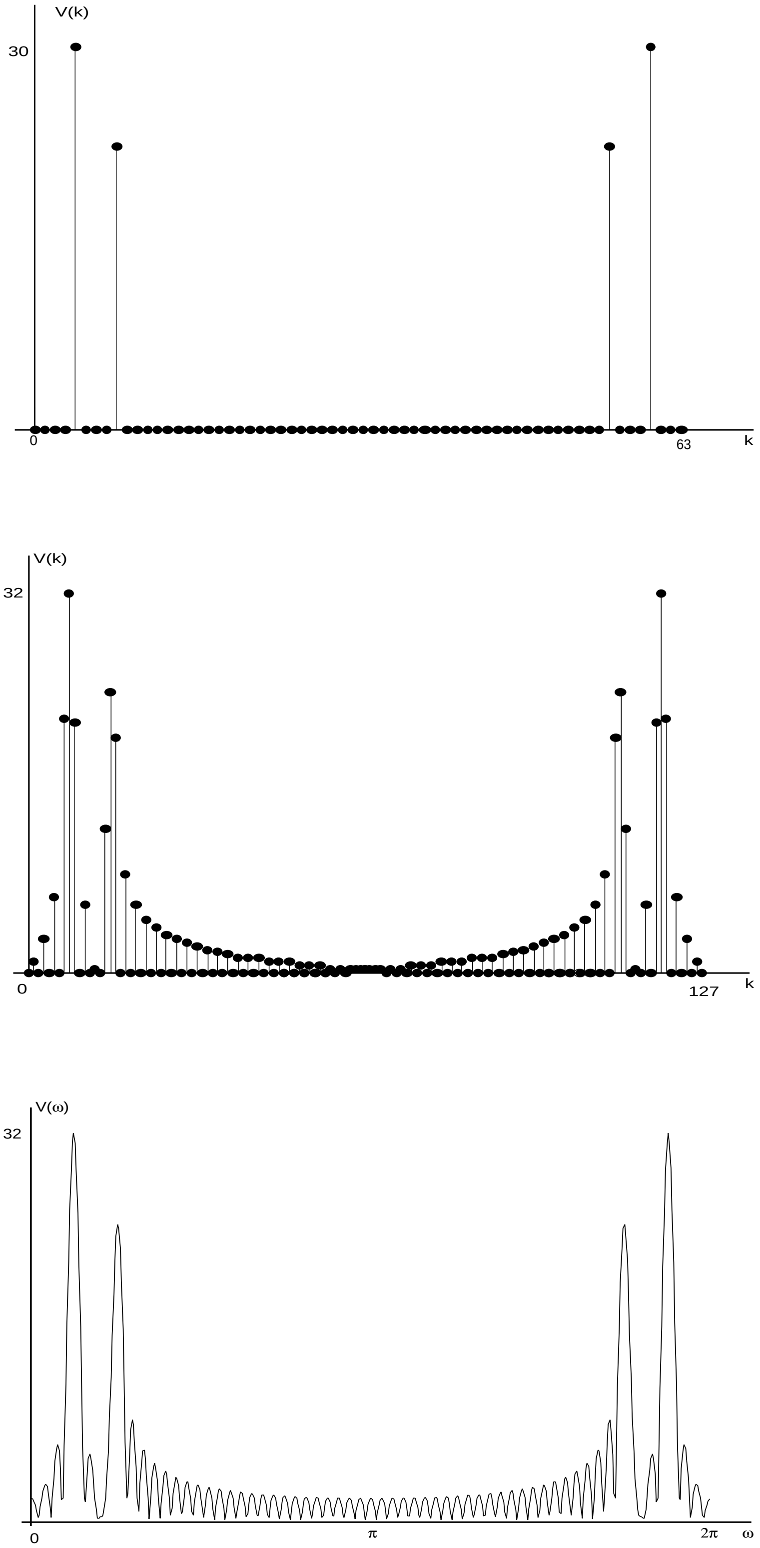


Figure 3.4: a) DFT of length $N = 64$; b) DFT of length $N = 128$; c) Fourier spectrum $V(e^{j\omega})$.

Example 3:

Explanation of following illustrations:

- Assume: signal of Example 2, Case 2a.
- Window: *Kaiser window* is applied instead of *rectangle window*.
- First: window length $L = 64$ and DFT length $N = 64$.
- Then: window length L and DFT length N are halved.
- Afterwards: for the case $L = 32$, the DFT length N is gradually increased up to $N = 1024$ (Zero Padding).
- Finally: DFT spectrum for the case $N = 1024$ and $L = 64$.

The Kaiser window is defined as:

$$w_K[n] = \begin{cases} \dfrac{I_0\left[\beta\left(1 - [(n-\alpha)/\alpha]^2\right)^{1/2}\right]}{I_0(\beta)} & : \quad 0 \le n \le L-1 \\ 0 & : \quad \text{otherwise} \end{cases}$$

In this example:

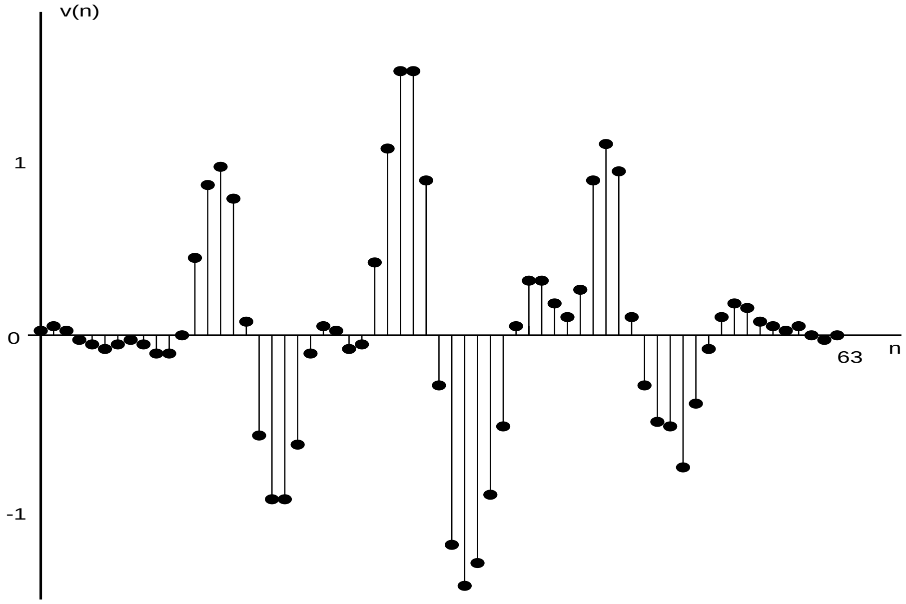
$$\beta = 0.8 \qquad \text{and} \qquad \alpha = \frac{L-1}{2}$$

The windowed signal $v[n]$:

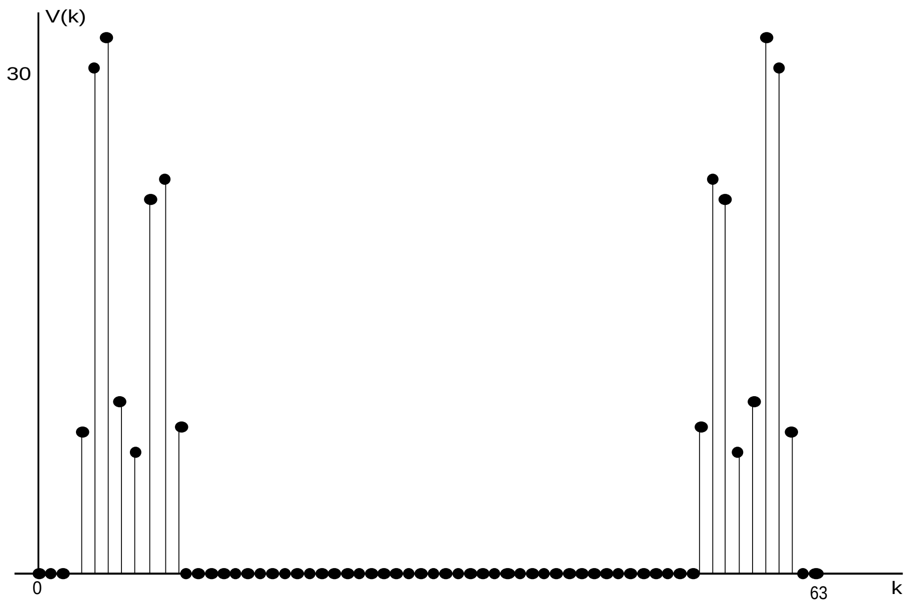
$$v[n] = w_K[n] \cos(\frac{2\pi}{14}n) + 0.75\, w_K[n] \cos(\frac{4\pi}{15}n)$$

Example 3: (continued)
DFT length $N = 64$, window length $L = 64$

- Windowed signal



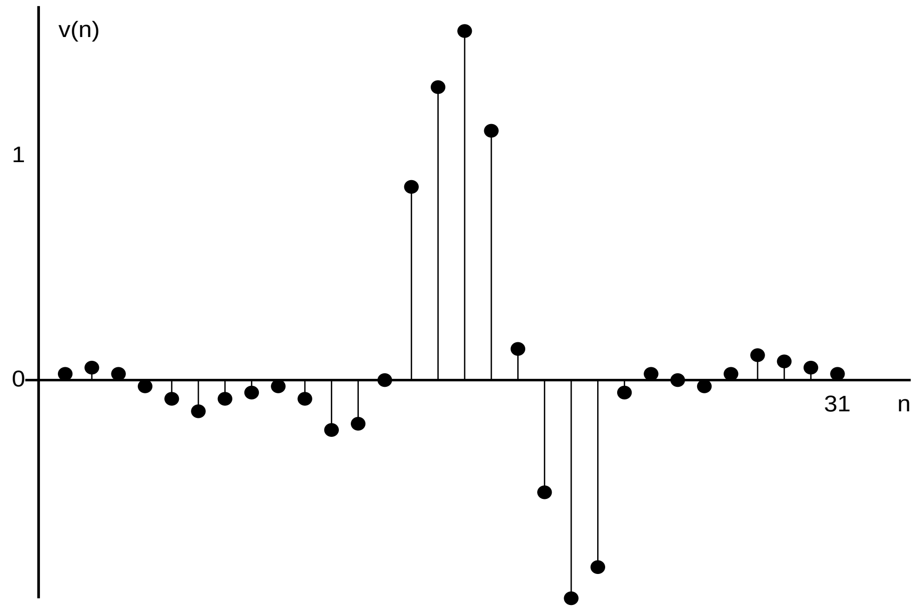
- DFT spectrum



Example 3: (continued)
DFT length $N = 32$, window length $L = 32$

(N and L halved)

- Windowed signal



- DFT spectrum

Example 3: (continued)

Effect of changing DFT length N at constant window length $L = 32$ (Zero Padding)

- DFT length $N = 32$, window length $L = 32$

- DFT length $N = 64$, window length $L = 32$

Example 3: (continued)

- DFT length $N = 128$, window length $L = 32$

- DFT length $N = 1024$, window length $L = 32$

Example 3: (continued)

Increasing the window length (L)

- DFT length $N = 1024$, window length $L = 64$

Example 4: Window function influence on the spectrum

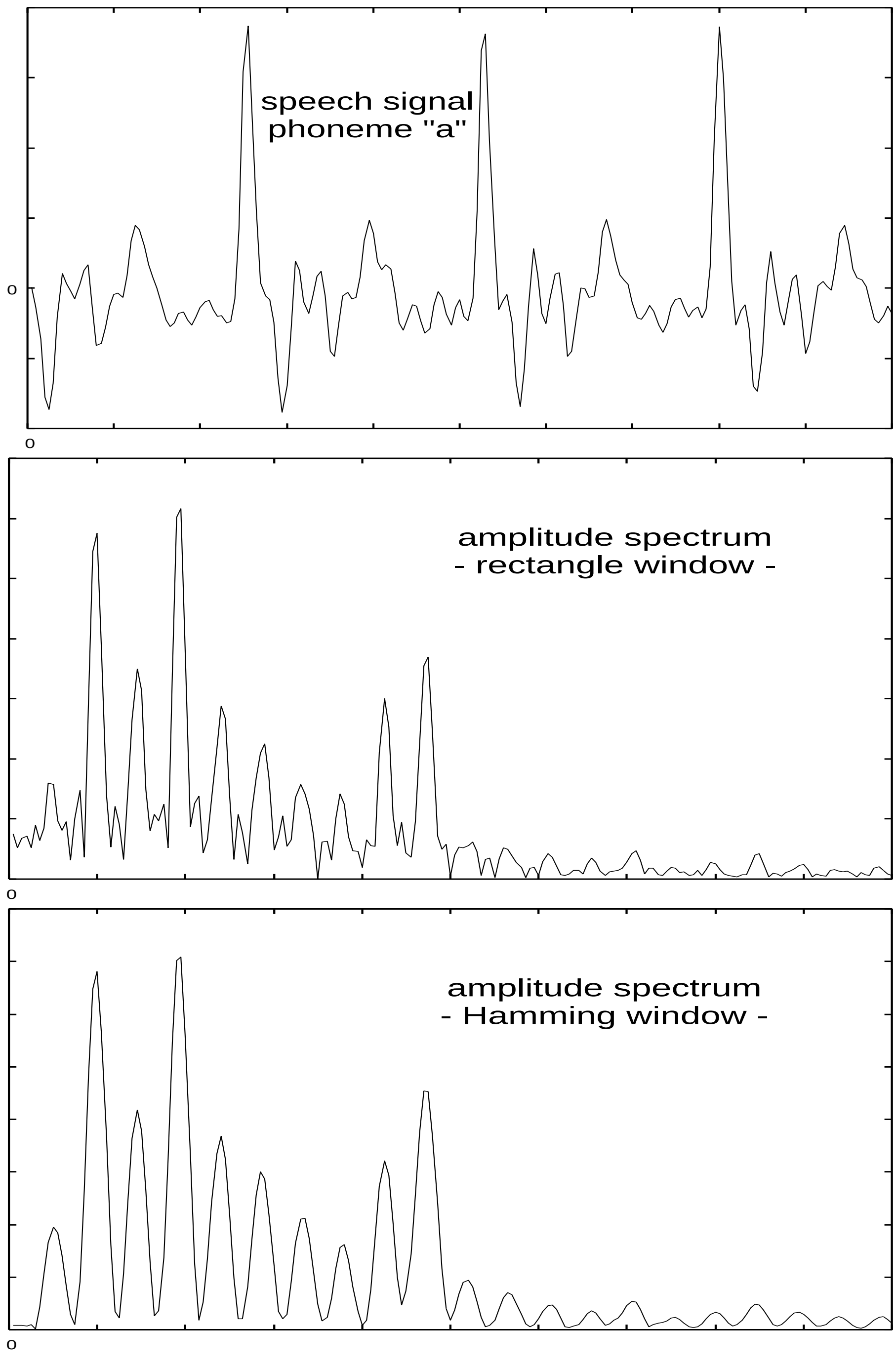


Figure 3.5: Influence of the window function:
above: speech signal (vowel “a”); central: 512 point FFT using rectangle window; below: 512 point FFT using Hamming window

3.3 Autocorrelation Function and Power Spectral Density

Definition of Autocorrelation Function (ACF) analog to the continuous time case:

$$R[k] \; := \; \sum_{n=-\infty}^{\infty} x[n] \, x[n+k]$$

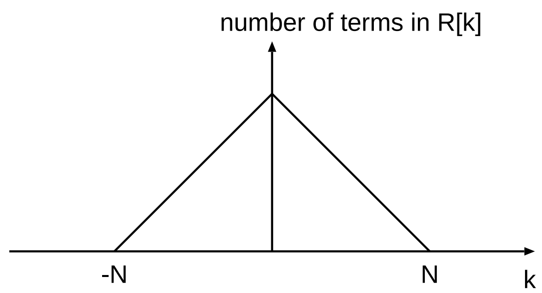
For a signal $x[n]$ assume (e.g. after some suitable windowing):

$$x[n] \; = \; \begin{cases} x[n] & 0 \leq n \leq N-1 \\ 0 & \text{otherwise} \end{cases}$$

In this case the ACF gives:

$$R[k] \; = \; \sum_{n=0}^{N-1-k} x[n] \, x[n+k] \qquad \text{because } x[n] = 0 \text{ for } n < 0 \text{ and } n \geq N$$

"triangular effect"



Cross correlation:

$$R_{xy}[k] \; = \; \sum_{n=-\infty}^{\infty} x[n] \cdot y[n-k]$$

In contrast to convolution:

$$O_{xy}[k] \; = \; \sum_{n=-\infty}^{\infty} x[n] \cdot y[k-n]$$

Properties of ACF:

1. $R[k] \;=\; R[-k]$

2. $R[k] \;\leq R[0]$ for each $k \in \mathbb{N}$ ($R[0]$: energy, intensity)

3. If $x[n] \;\longrightarrow\; R[k]$, then $\alpha\, x[n] \;\longrightarrow\; \alpha^2\, R[k]$

4. Intensity spectrum is the Fourier Transform of the ACF:

$$\begin{aligned}
\mid X(e^{j\omega}) \mid^2 &= X(e^{j\omega}) \cdot \overline{X(e^{j\omega})} \\
&= \sum_{k=-\infty}^{\infty} x[k] \exp(-j\omega k) \cdot \sum_{l=-\infty}^{\infty} x[l] \exp(j\omega l) \\
&= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} x[k]\, x[l] \exp(-j\omega k) \exp(j\omega l) \\
&= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} x[k+l]\, x[l] \exp(-j\omega k) \exp(-j\omega l) \exp(j\omega l) \\
&= \sum_{k=-\infty}^{\infty} \left(\sum_{l=-\infty}^{\infty} x[k+l]\, x[l] \right) \exp(-j\omega k) \\
&= \sum_{k=-\infty}^{\infty} R[k] \exp(-j\omega k)
\end{aligned}$$

Note: The phase spectrum is removed.

5. Because of the symmetry $R[k] = R[-k]$ the DFT becomes the cosine transform:

$$
\begin{aligned}
\mid X(e^{j\omega}) \mid^2 &= \sum_{k=-\infty}^{\infty} R[k] \exp(-j\omega k) \\
&= \sum_{k=-(N-1)}^{N-1} R[k] \exp(-j\omega k) \\
&= R[0] \; + \; \sum_{k=1}^{N-1} R[k] \; (\exp(-j\omega k) \; + \; \exp(j\omega k)) \\
&= R[0] \; + \; 2 \cdot \sum_{k=1}^{N-1} R[k] \; \cos(\omega k) \qquad \text{because} \qquad R[k] = R[-k]
\end{aligned}
$$

6. The intensity spectrum $\mid X(e^{j\omega}) \mid^2$ is a polynom of $\cos(\omega)$ with grade $N-1$.

   Reason: Moivre formula:

$$
\cos(\omega k) \; = \; \cos^k(\omega) \; - \; \binom{k}{2} \cos^{k-2}(\omega) \sin^2(\omega) \; + \; \binom{k}{4} \cos^{k-4}(\omega) \sin^4(\omega) \; - \; \ldots
$$

Example 1: Spectral analysis using ACF

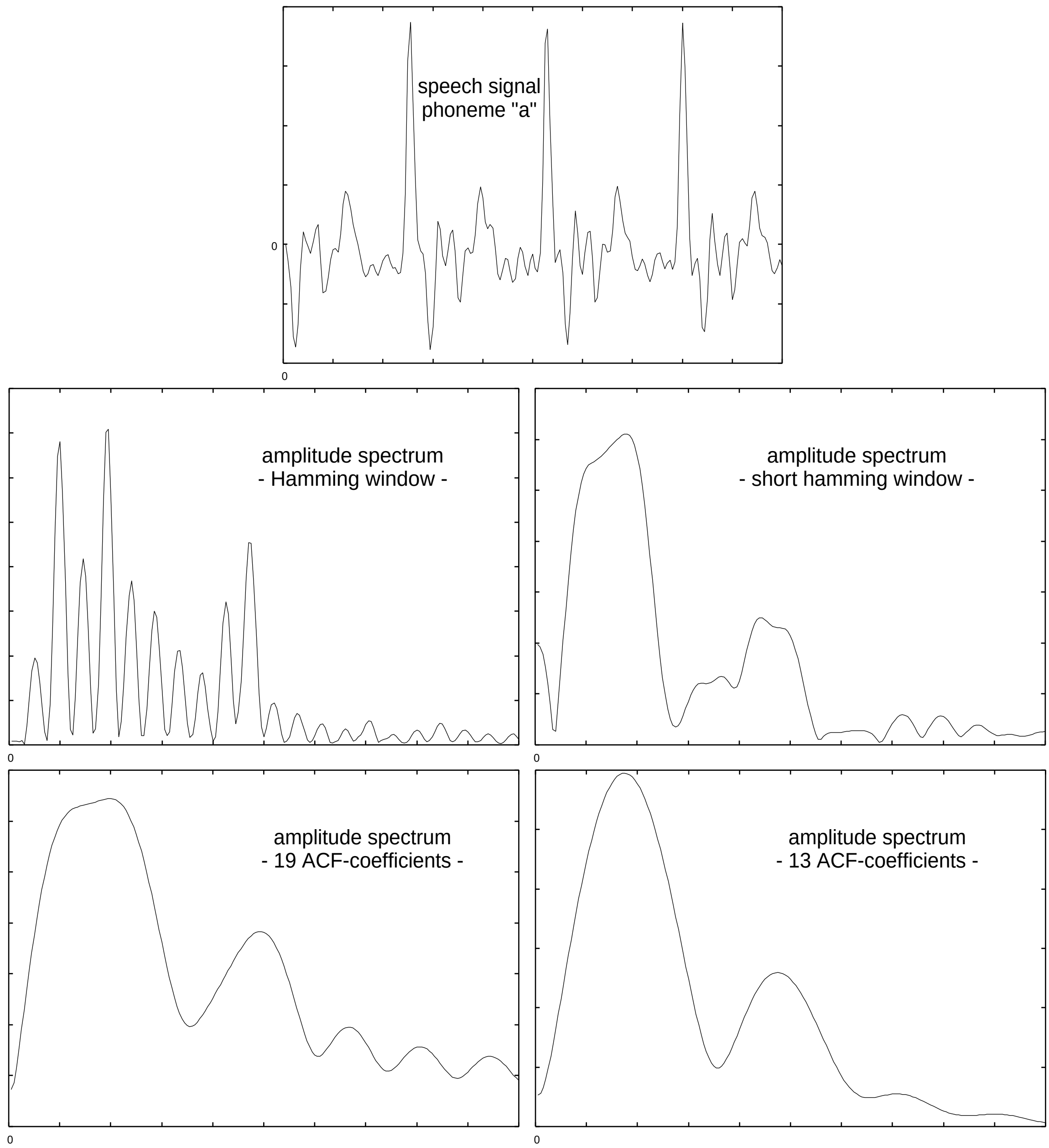


Figure 3.6: Fourier Transform of a voiced speech segment:
a) signal progression, b) high resolution Fourier Transform, c) low resolution Fourier Transform with short Hamming window (50 sampled values), d) low resolution Fourier Transform using autocorrelation function (19 coefficients), e) low resolution Fourier Transform using autocorrelation function (13 coefficients)

Example 2: ACF of voiced and unvoiced speech segments

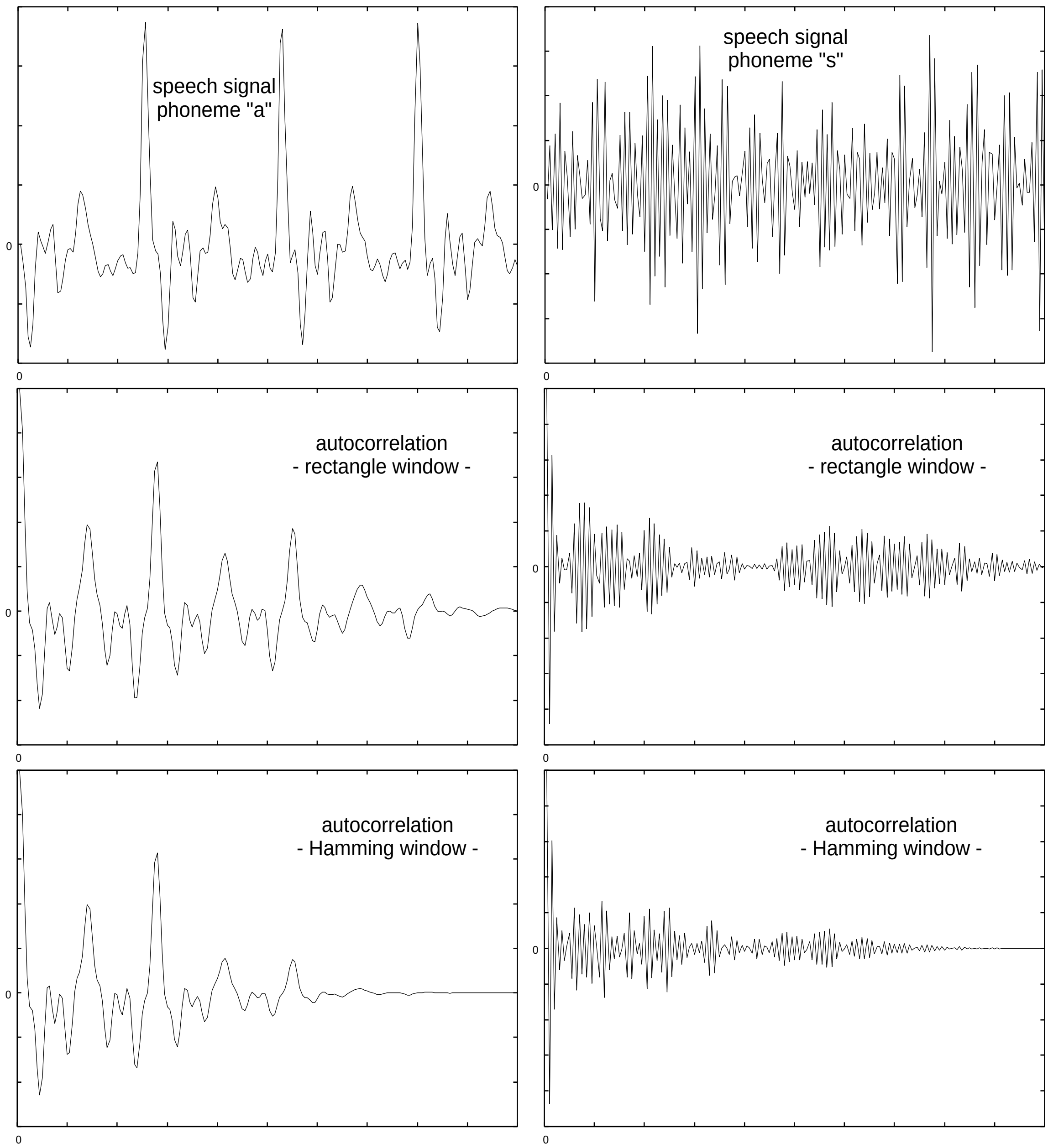


Figure 3.7: Signal progression and autocorrelation function of voiced (left) and unvoiced (right) speech segment

Example 3: Temporal progression of autocorrelation coefficients

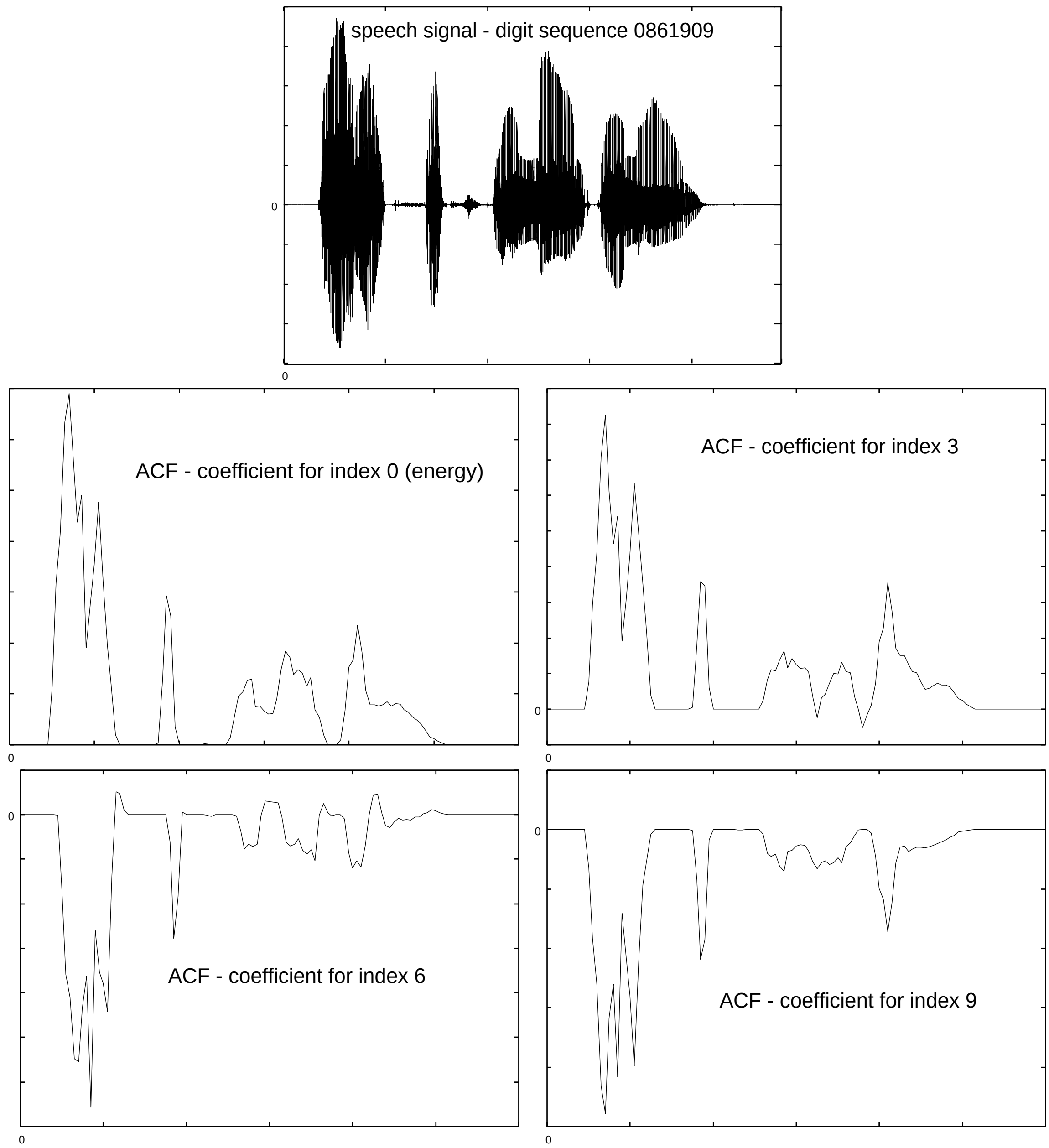


Figure 3.8: Temporal progression of speech signal and four autocorrelation coefficients

3.4 Spectrograms

Using DFT

- Wide-band: in frequency domain:
  - short time window
  - “interaction” in the “synchronization” between time window and “pitch impulses”
  - vertical lines
  - no resolution of spectral fine structure
- Narrow-band: in frequency domain:
  - long time window
  - good resolution of the spectral fine structure

Example 1: speech spectrograms

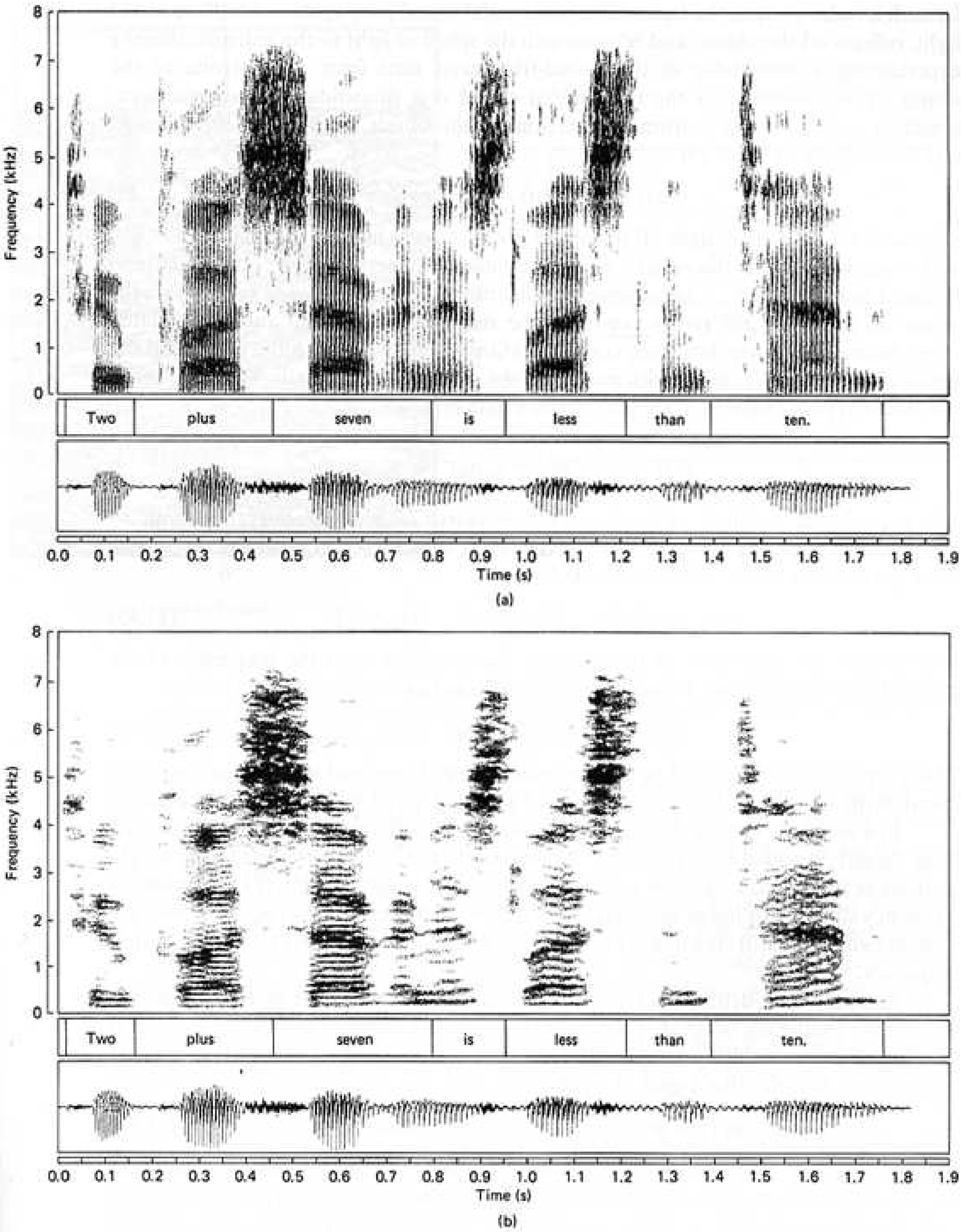


Figure 3.9: a) wide-band spectrogram: short time window, high time resolution (vertical lines), no frequency resolution; for voiced signals provides information on formant structure b) narrow-band spectrogram: long time window, no time resolution, high frequency resolution (horizontal lines); for voiced signals provides information on fundamental frequency (pitch)

Example 2: speech spectrograms

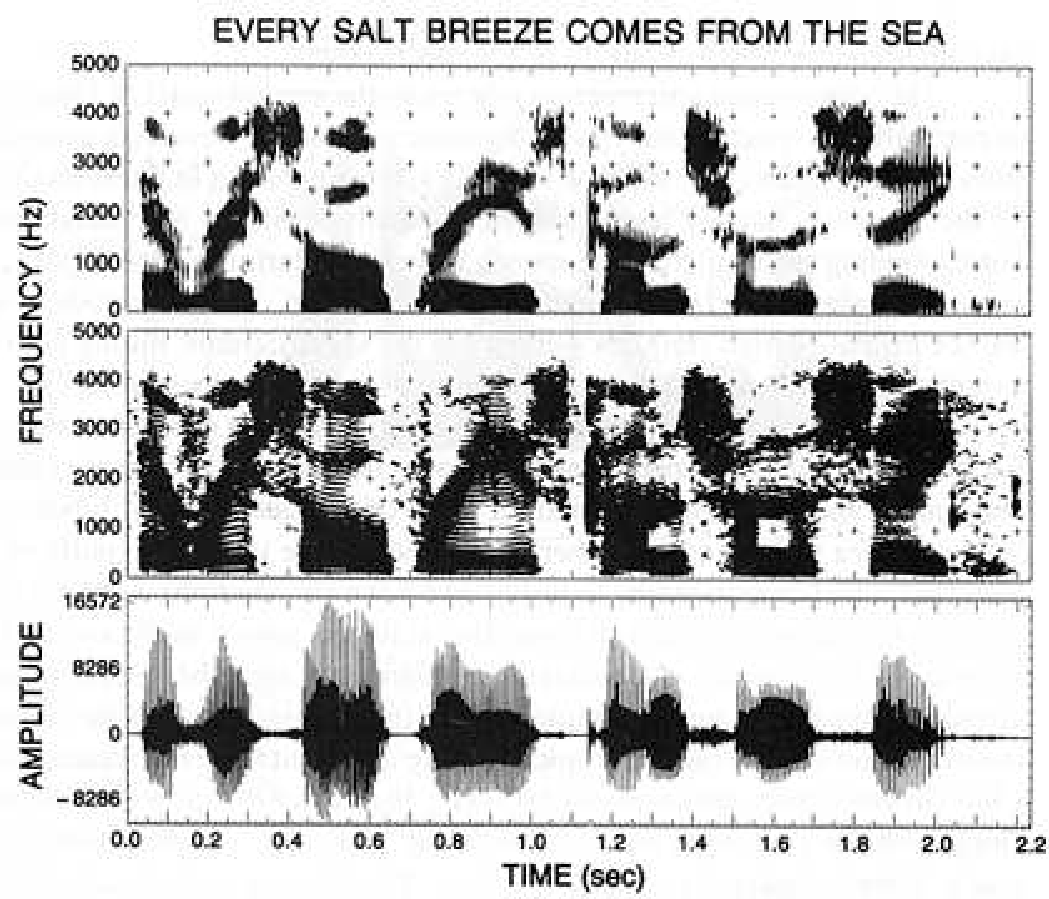
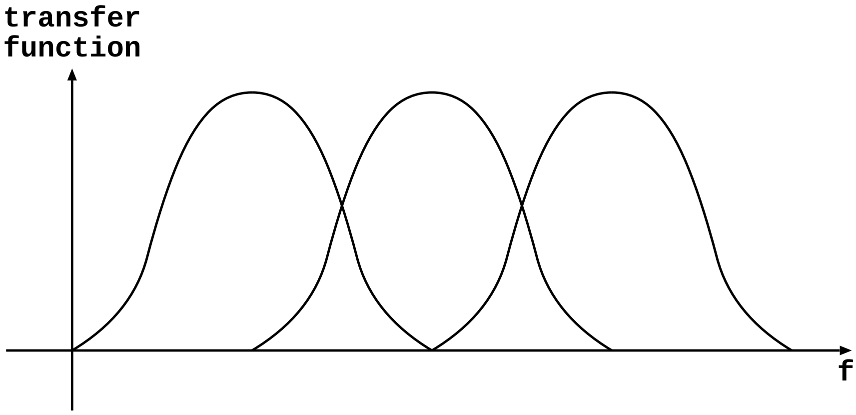


Figure 3.10: Wide-band and narrow-band spectrogram and speech amplitude for the sentence “Every salt breeze comes from the sea”.

3.5 Filter Bank Analysis

- History:
  Decomposition of the signal using a "bank" of band-pass filters and energy calculation in each frequency band



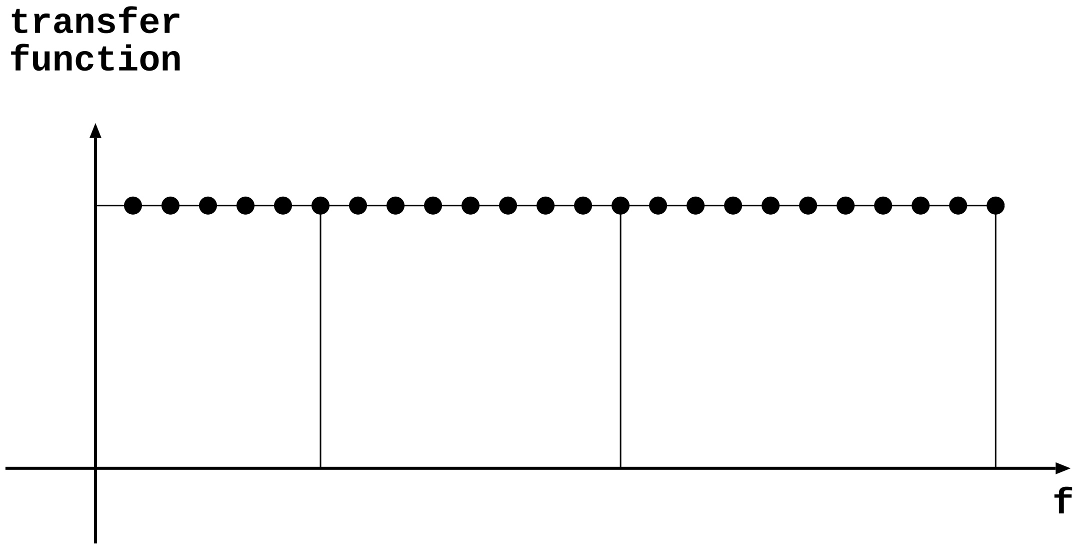
- Today digitally:
  - Digital filters:

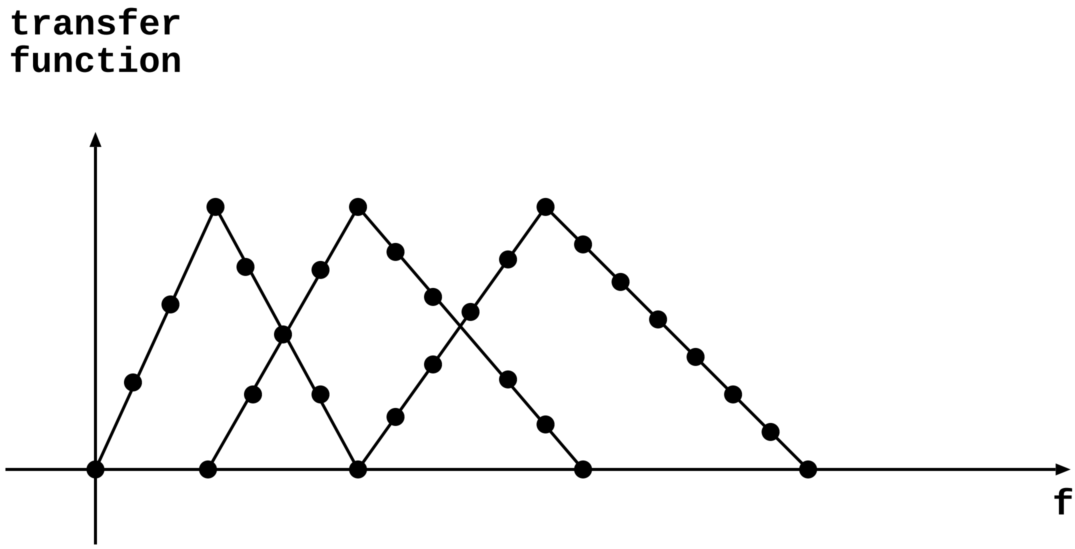
$$y_k[n] \;=\; \sum_{m=-\infty}^{\infty} h_k[n-m]\, x[m]\;, \quad k = 1, \ldots, K$$

  - FIR: Finite Impulse Response
    IIR: Infinite Impulse Response (recursive filters)
  - DFT (FFT) + further processing

- DFT/FFT Method:
  - Window function
  - Appending zeros for desired "resolution" (*zero padding*)
  - FFT
  - "Energy" calculation: $|X(e^{j\omega})|$, $|X(e^{j\omega})|^2$, $\log|X(e^{j\omega})|$
  - Weighted averaging for each channel and frequency band respectively

DFT/FFT filter bank:



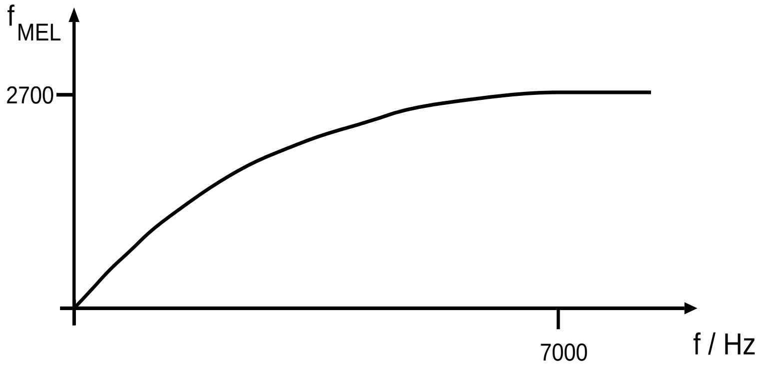


- Averaging:
  summation should be as smooth as possible over all channels
  Form: rectangle, triangle, trapeze, etc.
- Choosing the central frequencies f_k:
  - constant:
    $\Delta f_k = \text{const. for all } k$
    e.g. 20 channels with $\Delta f = 200\text{Hz}$ for $0 - 4$ kHz
  - constant relative band width:
    $$\frac{\Delta f_k}{f_k} = \text{const. for all } k$$
  - frequency groups of the ear (total number 24):
    $$\begin{aligned} f &< 500\text{Hz}: & \Delta f &= 100 \\ f &\geq 500\text{Hz}: & \frac{\Delta f}{f} &= 20\% \end{aligned}$$
  - adjusted to vowels or sounds

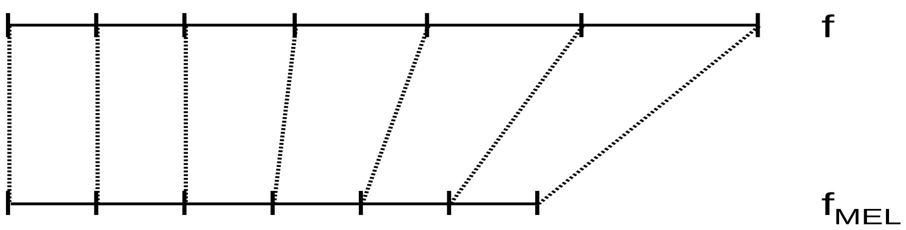
3.6 Mel-frequency scale

The frequency resolution of the human ear is decreasing on the higher frequencies. This empirical dependency results in the definition of the Mel scale, which is approximately calculated as (from: Hidden Markov Toolkit, Cambridge University Engineering Departement, S.J.Young):

$$f_{\mathrm{MEL}} \;=\; 2595 \log_{10}\left(1 + \frac{f}{700\mathrm{Hz}}\right)$$



Compression of the high frequencies



A filter bank with constant band-widths can be used on the Mel scale:

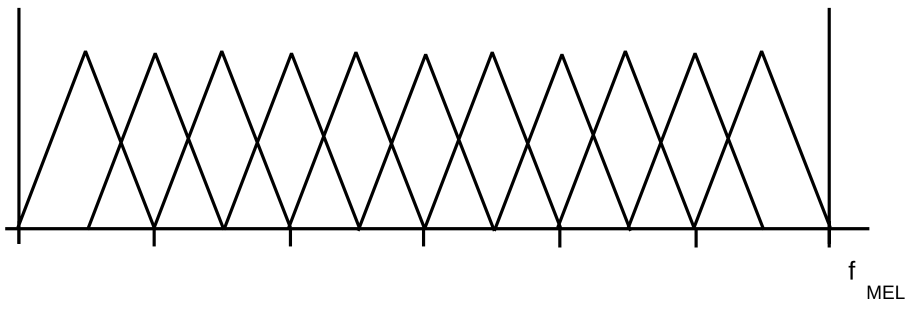


Table: MEL Scale:

f/Hz	f_{MEL}
65	100
136	200
213	300
298	400
391	500
492	600
603	700
724	800
856	900
1000	1000
1158	1100
1330	1200
1519	1300
1724	1400
1949	1500
2195	1600
2464	1700
2757	1800
3078	1900
3429	2000
3812	2100
4230	2200
4688	2300
5187	2400
5734	2500
6331	2600
6984	2700

3.7 Cepstrum

The Cepstrum is the Fourier series expansion of the logarithm of the spectrum.
Comparison: autocorrelation function is a Fourier series of the normal spectrum.

We consider:

$$y[n] \; = \; \sum_{k=-\infty}^{\infty} h[n-k] \; x[k]$$

Goal:
Separating the kernel $h[n]$ from the input signal $x[n]$.
This problem is also called inversion or deconvolution.

- Convolution theorem:

$$Y(e^{j\omega}) \; = \; H(e^{j\omega}) \; X(e^{j\omega})$$

- Logarithm (complex):

$$\log Y(e^{j\omega}) \; = \; \log H(e^{j\omega}) \; + \; \log X(e^{j\omega})$$

- Inverse Fourier Transform:

$$F^{-1}\left\{\log Y(e^{j\omega})\right\} \; = \; F^{-1}\left\{\log H(e^{j\omega})\right\} \; + \; F^{-1}\left\{\log X(e^{j\omega})\right\}$$

- Another notation:

$$\hat{y}[n] \;=\; \hat{x}[n] \;+\; \hat{h}[n]$$

using the definition of the cepstrum for $x[n]$
(analogous for $\hat{y}[n]$ and $\hat{h}[n]$)

$$\begin{aligned}
\hat{x}[n] &= F^{-1}\left\{\log X(e^{j\omega})\right\} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(j\omega n)\,\log X(e^{j\omega})\,d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(j\omega n)\,\log\left[\sum_m x[m]\,\exp(-j\omega m)\right] d\omega \\
&= C\left\{x[n]\right\}
\end{aligned}$$

- Note:
  - Cepstrum = artificial word derived from “spectrum”
  - Cepstrum is located in time domain

Through the cepstrum transformation

$$x[n] \quad \longrightarrow \quad \hat{x}[n] \; = \; C\left\{x[n]\right\}$$

the convolution comes down to a simple addition.

In the cepstrum domain, a linear operation L (time invariance is not necessary) on $\hat{y}[n]$ is performed separately on $\hat{h}[n]$ and $\hat{x}[n]$:

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[n-k]\, x[k] \\ \hat{y}[n] &= \hat{h}[n] \; + \; \hat{x}[n] \\ L\left\{\hat{y}[n]\right\} &= L\left\{\hat{h}[n]\right\} \; + \; L\left\{\hat{x}[n]\right\} \end{aligned}$$

With the definition G_L for the concatenation of the cepstrum, the operation L, and the inverse cepstrum

$$G_L \; := \; C^{-1} \circ L \circ C$$

we obtain

$$G_L\left\{h[n] * x[n]\right\} \; = \; G_L\left\{h[n]\right\} \; * \; G_L\left\{x[n]\right\}.$$

Such a transformation G_L acts on $h[n]$ and $x[n]$ separately, and is called:

homomorph (structure preserving)

Complex cepstrum:

$$\hat{x}[n] \;=\; \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} \exp(j\omega n)\, \log X(e^{j\omega})\; d\omega$$

Note: complex logarithm

Simple cepstrum (real cepstrum):

$$\tilde{x}[n] \;=\; \frac{1}{2\pi}\int\limits_{0}^{2\pi} \exp(j\omega n)\, \log|X(e^{j\omega})|\; d\omega$$

- Cepstrum: Fourier coefficients of the logarithmized power spectral density
- ACF: Fourier coefficients of Fourier series of the power spectral density

Setting cepstral coefficients $\hat{x}[n]$ to zero for high n results in smoothing of the power spectral density.

Implementation:

Fourier Transform via N–FFT ($N = 512$, 1024, 2048)
(But: discretisation error):

$$\hat{x}[n] \;:=\; \frac{1}{N}\sum_{k=0}^{N-1} e^{j\dfrac{2\pi}{N}kn} \log|X(e^{j\dfrac{2\pi}{N}k})|$$

Example 1: Real cepstrum

Fine structure of power spectral density with the period $1/T$ results in a single peak in the cepstrum at time T.

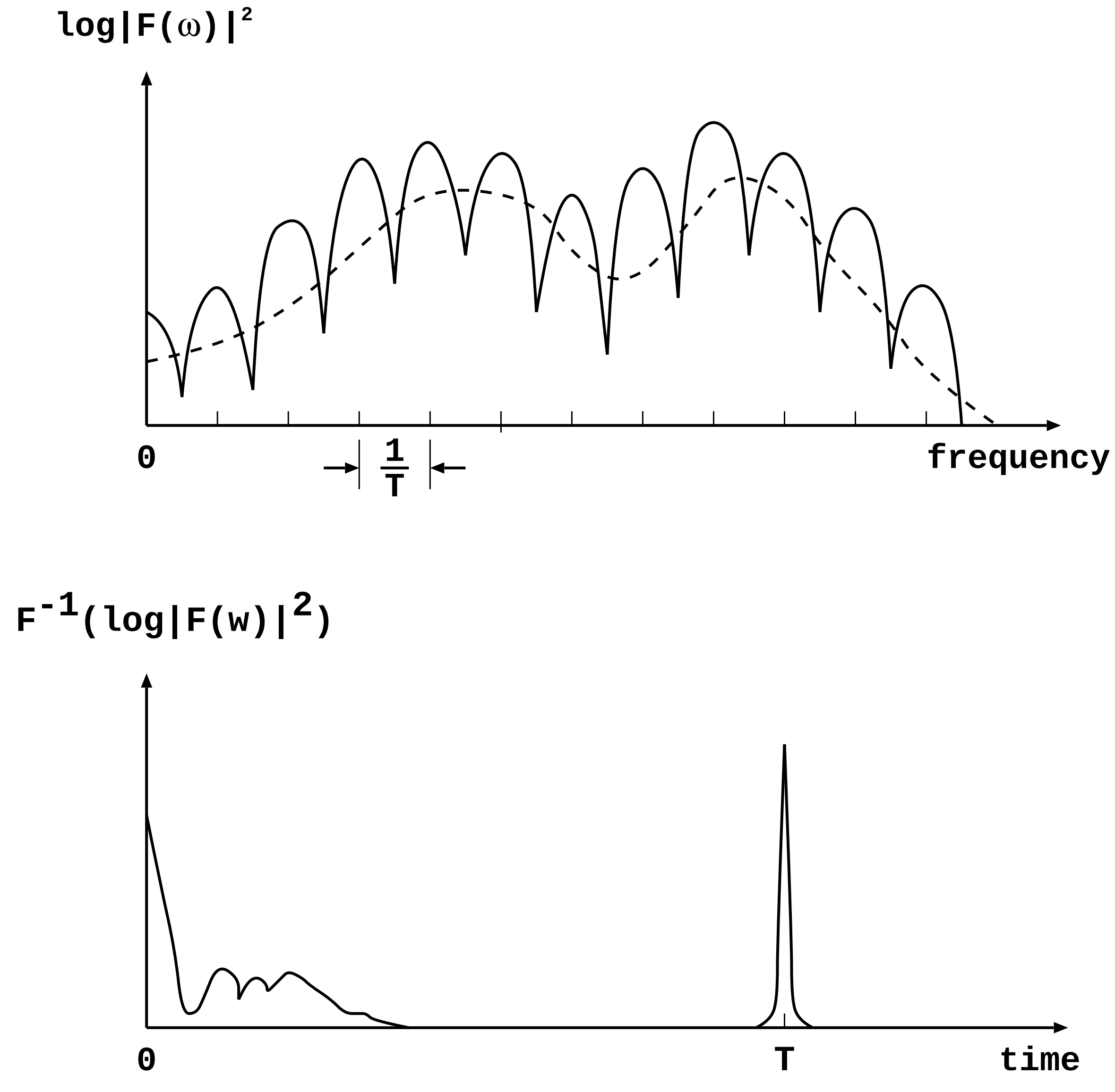


Figure 3.11: Above: logarithmized power spectrum of a spoken vowel (schematic). Below: corresponding cepstrum (inverse Fourier–transform of the logarithmized power spectrum).

Example 2: Smoothing

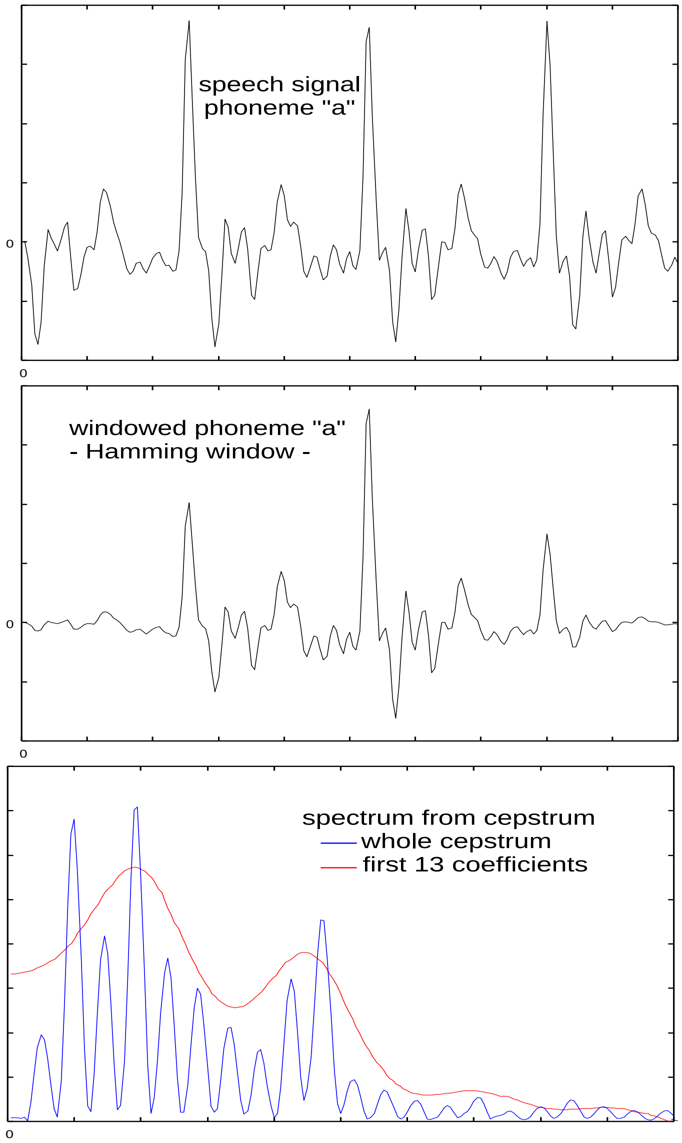


Figure 3.12: Cepstral smoothing: speech signal (vowel “a”), windowed speech signal (Hamming window), spectrum obtained from the whole cepstrum (blue) and smoothed spectrum obtained from the first 13 cepstral coefficients (red).

Example 3: Smoothing with different numbers of cepstral coefficients

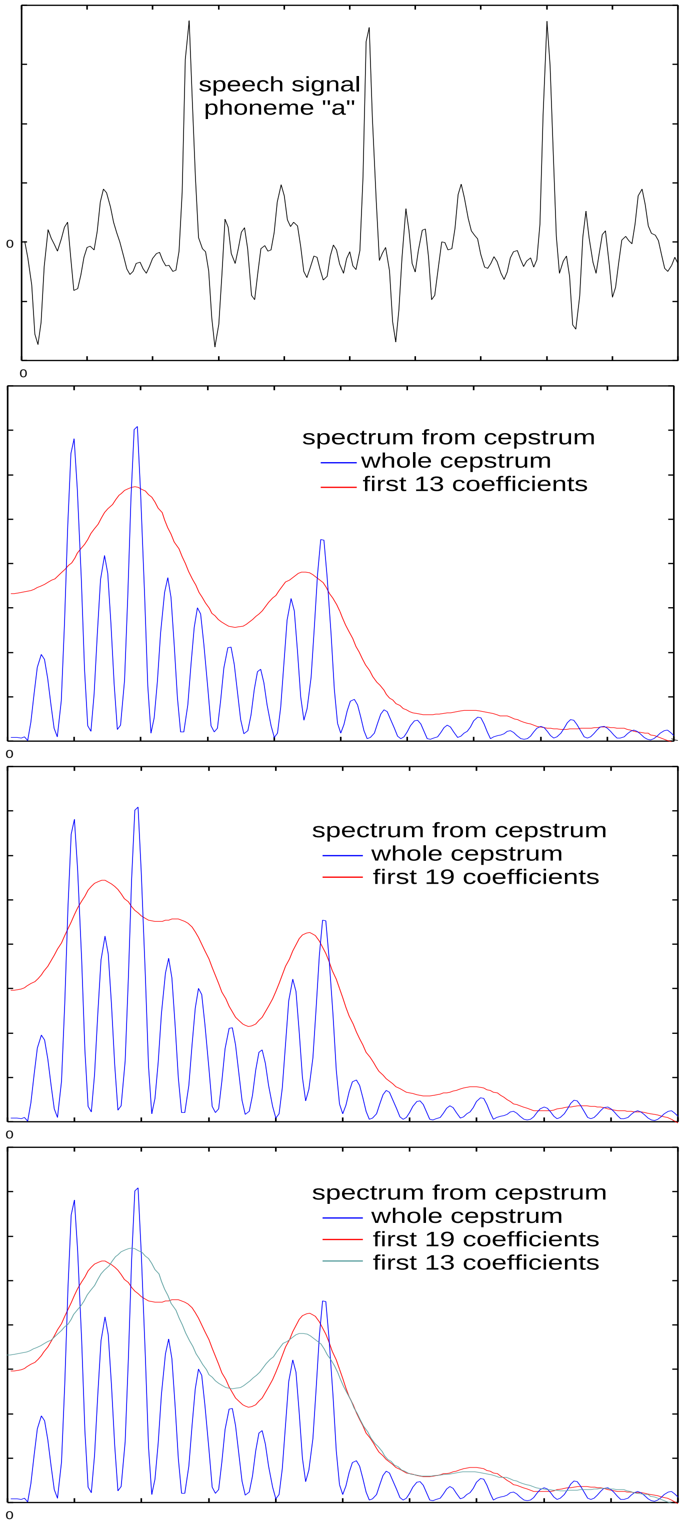
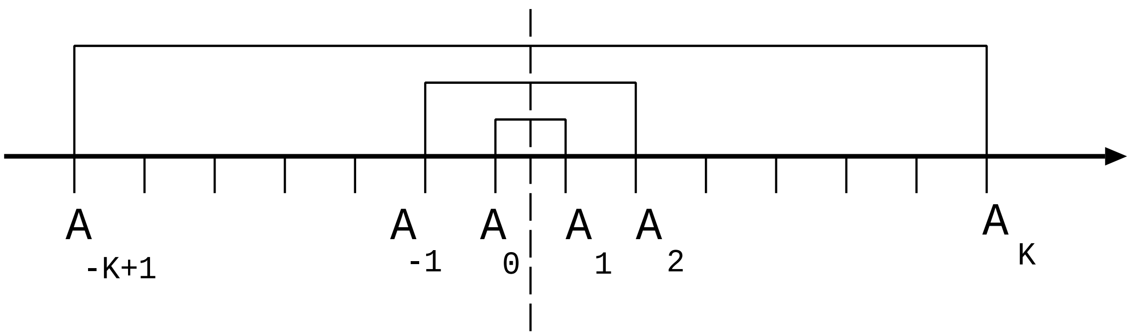


Figure 3.13: Homomorph analysis of a speech segment: signal progression, homomorph smoothed spectrum using 13 and 19 cepstral coefficients

Cepstrum calculation using Filter Bank Output

- Filter bank outputs $A[k]$ for $k = 1, \ldots, K$
  Note: $k = 0$ is missing.
- We complete the outputs symmetrically:



- Symmetry $A_{-k+1} \; = \; A_k$ for all $k = 1, \ldots, K$.

Inverse DFT $a[n]$ of the symmetric sequence $A_{-K+1}, \ldots, A_K$:

$$a[n] = \frac{1}{2K} \sum_{k=-K+1}^{K} A_k \exp\left(\frac{2\pi j}{2K} nk\right)$$

$$= \frac{1}{2K} \sum_{k=1}^{K} A_k \left[\exp\left(\frac{2\pi j}{2K} nk\right) + \exp\left(\frac{2\pi j}{2K} n(-k+1)\right)\right]$$

$$= \exp\left(\frac{2\pi j}{2K} 0.5\right) \frac{1}{2K} \sum_{k=1}^{K} A_k \left[\exp\left(\frac{2\pi j}{2K} n(k-0.5)\right) + \exp\left(-\frac{2\pi j}{2K} n(k-0.5)\right)\right]$$

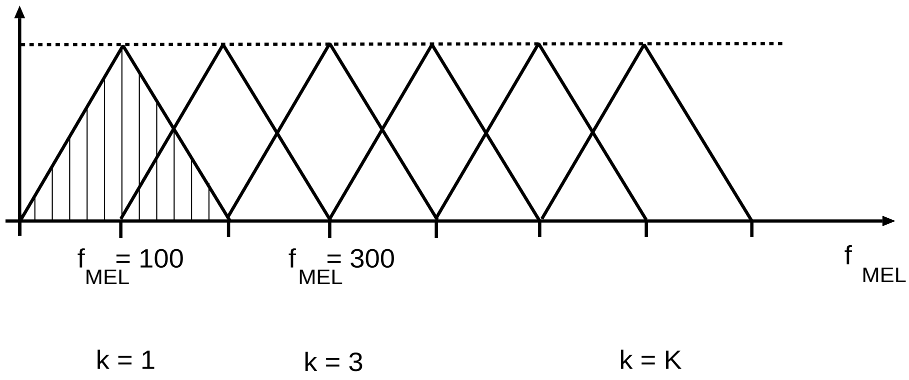
$$= \exp\left(\frac{2\pi j}{2K}\, 0.5\right) \frac{1}{K} \sum_{k=1}^{K} A_k \cos\left(\frac{\pi n}{K}(k-0.5)\right)$$

The phase term $\exp\left(\frac{2\pi j}{2K}\, 0.5\right)$ depends on the position of the symmetry axis around $k = 0.5$.

Cepstrum is defined as:

$$a[n] = \frac{1}{K} \sum_{k=1}^{K} A_k \cos\left(\frac{\pi n}{K}(k-0.5)\right)$$

Mel Cepstrum according to Davis and Mermelstein



Filter bank:

- overlapping band-pass filters triangular shape,
- all channels have equal band width, and filter positioning is equidistant on a Mel scale.

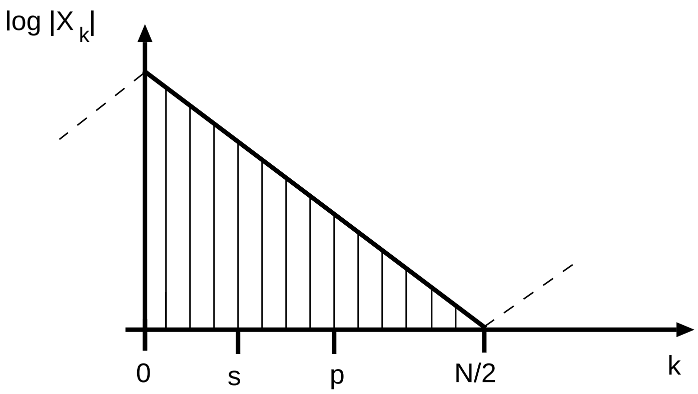
Calculation of the filter bank outputs:

- magnitude of DFT coefficients,
- for each channel summation of the magnitudes according to triangular weight function,
- for each channel logarithm of the sum.

Thus the filter outputs $A[k]$ with $k = 1, \ldots, K$ are obtained. Using the filter bank outputs, the cepstrum is calculated using a cosine transform. (see previous description)

3.8 Statistical Interpretation of the Cepstrum Transformation

We consider the filter bank outputs $\log|X_k|$.



Assumption: The correlation between the outputs s and p, i.e. the element C_{sp} of the covariance matrix does not depend directly on s or p, but only on their difference. Because the spectrum is periodical there is no distance greater than N:

$$C_{sp} \;=\; c_{(s-p)\mathrm{mod}N}$$

It is further assumed that the correlation is locally symmetric:

$$C_{s,s+n} = C_{s,s-n}$$

Then:

$$c_{(s-s-n)\mathrm{mod}N} = c_{(s-s+n)\mathrm{mod}N}$$

$$c_{(-n)\mathrm{mod}N} = c_{(+n)\mathrm{mod}N}$$

With $0 \leq n \leq N$ follows:

$$c_n = c_{N-n}$$

i.e. we have a symmetric cyclic matrix with the kernel vector c.

Example: the covariance matrix for $N = 8$:

$$C = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 & c_4 & c_3 & c_2 & c_1 \\ c_1 & c_0 & c_1 & c_2 & c_3 & c_4 & c_3 & c_2 \\ c_2 & c_1 & c_0 & c_1 & c_2 & c_3 & c_4 & c_3 \\ c_3 & c_2 & c_1 & c_0 & c_1 & c_2 & c_3 & c_4 \\ c_4 & c_3 & c_2 & c_1 & c_0 & c_1 & c_2 & c_3 \\ c_3 & c_4 & c_3 & c_2 & c_1 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_4 & c_3 & c_2 & c_1 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_4 & c_3 & c_2 & c_1 & c_0 \end{bmatrix}$$

Such a covariance matrix will be diagonalised using the cosine transform (or Fourier Transform, which results in the cosine transform due to the symmetry) (see excursion in chapter 2.17).

3.9 Energy in acoustic Vector

The energy is usually added as zeroth (or first) component to the acoustic vector.

For the logarithmic energy we have:

$$\log E \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} \log|X(e^{j\omega})|^2 \, d\omega$$

For the (short time) spectrum or cepstrum it approximately holds:

$$\log E \;\approx\; \frac{1}{K}\sum_{k=1}^{K} \log|X_k|^2$$

Spectra are usually normalized with log E:

$$\log Y_k^2 \;=\; \log|X_k|^2 - \log E$$

such that:

$$\sum_{k=1}^{K} \log Y_k^2 \equiv 0$$

The cepstral coefficient $\hat{x}[0]$ is the logarithmized energy.

Chapter 4

Fourier Transform and Image Processing

- Overview:

  4.1 Spatial Frequencies and Fourier Transform for Images

  4.2 Discrete Fourier Transform for Images

  4.3 Fourier Transform in Computer Tomography

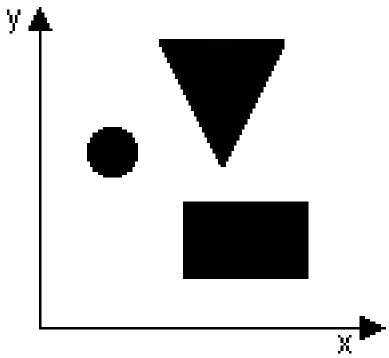
  4.4 Fourier Transform and RST Invariance

4.1 Spatial Frequencies and Fourier Transform for Images

A grey-valued image $g(x, y)$ can be interpreted as:

$$\begin{aligned} g : \mathbb{R}^2 &\rightarrow [0, \infty[\\ (x, y) &\rightarrow g(x, y) \end{aligned}$$

Space coordinates (x, y) are at first considered as continuous. Discretization and DFT will be analyzed later.



Convention: $g(x, y) \equiv 0$ *outside* of the image.

The Fourier–transform $G(f_x, f_y)$ of the image $g(x, y)$ is defined as:

$$\begin{aligned} G(f_x, f_y) &= F\{(x, y) \rightarrow g(x, y)\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)\; e^{-2\pi j (f_x x + f_y y)}\; dx dy \end{aligned}$$

The arguments f_x and f_y are called *spatial frequencies*.

The two-dimensional Fourier–transform can be obtained by using *two one-dimensional* Fourier–transforms.
We consider one "image row" with a constant value of y:

$$x \;\rightarrow\; g(x,y)$$

Corresponding Fourier–transform $G_y(f_x)$:

$$\begin{aligned} G_y(f_x) &= F\{x \;\rightarrow\; g(x,y)\} \\ &= \int\limits_{-\infty}^{\infty} g(x,y)\; e^{-2\pi j f_x x}\; dx \end{aligned}$$

Then we compute the Fourier–transform of the function:

$$y \;\rightarrow\; G_y(f_x)$$

and obtain:

$$\begin{aligned} F\{y \;\rightarrow\; G_y(f_x)\} &= \int\limits_{-\infty}^{+\infty} G_y(f_x)\; e^{-2\pi j f_y y}\; dy \\ &= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} g(x,y)\; e^{-2\pi j (f_x x + f_y y)}\; dx dy \end{aligned}$$

In this way we get the result:

$$G(f_x, f_y) \;=\; F\{y \;\rightarrow F\{x \;\rightarrow\; g(x,y)\}\}$$

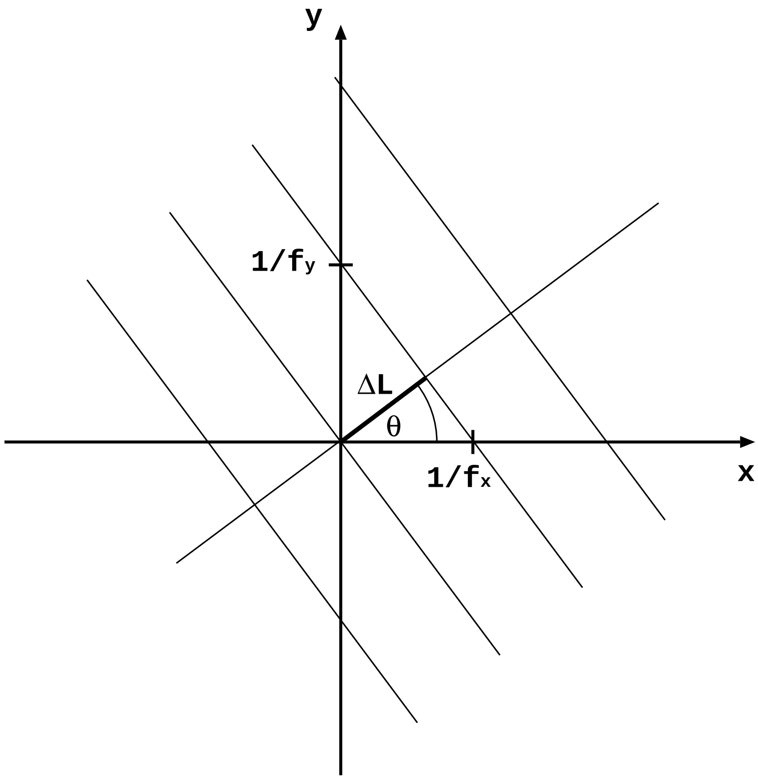
For the *inverse* FT we have (as can be expected):

$$\begin{aligned} g(x,y) &= F^{-1}\{(f_x, f_y) \;\rightarrow\; G(f_x, f_y)\} \\ &= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} G(f_x, f_y)\; e^{2\pi j (f_x x + f_y y)}\; df_x df_y \end{aligned}$$

We would like to interpret the two-dimensional FT visually. For this purpose, we consider the exponential factor in the FT and require the following condition:

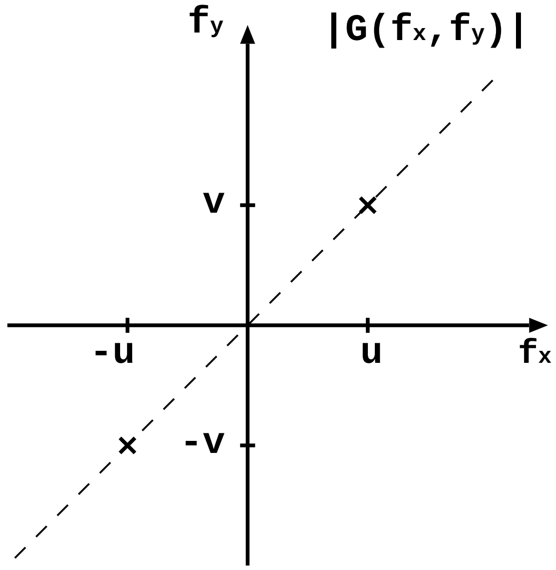
$$\begin{aligned} e^{2\pi j(f_x x+f_y y)} &\stackrel{!}{=} 1 \\ \Rightarrow\; 2\pi j(f_x x+f_y y) &= 2\pi n \qquad \text{for } n \in \mathbb{N} \end{aligned}$$

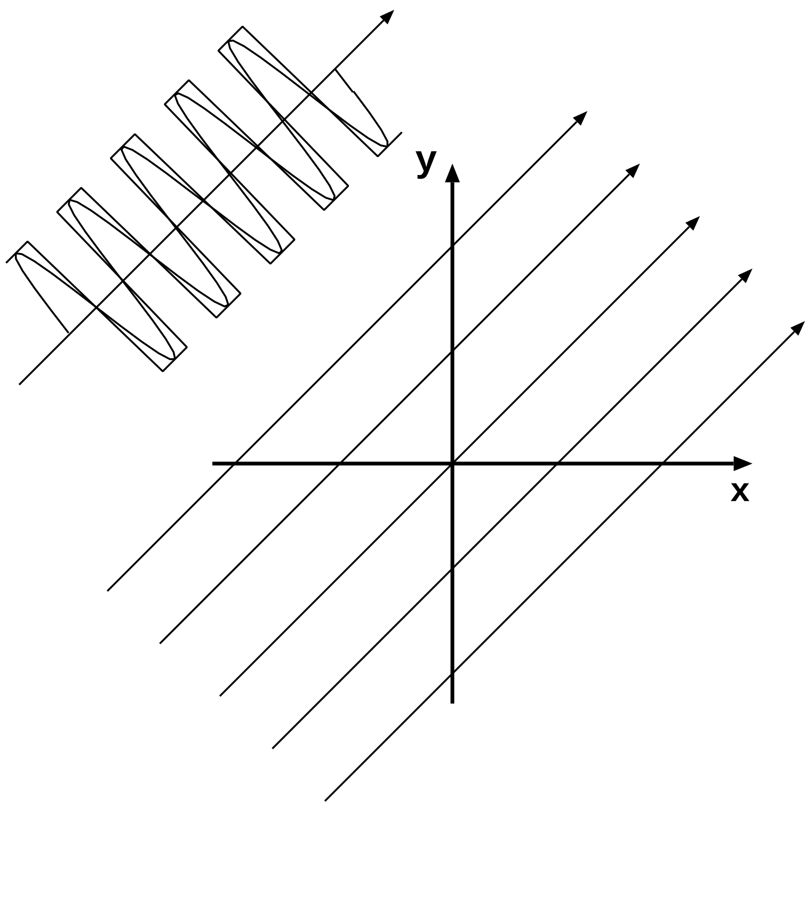
$$\Rightarrow \quad y = -\frac{f_x}{f_y}x + \frac{n}{f_y}$$



$$\text{spatial period} \qquad \Delta L = \frac{1}{\sqrt{f_x^2+f_y^2}}$$

Special case:
$|G(f_x, f_y)|$ has a large value *only* at "one" point $(u, v) = (f_x, f_y)$ in the spatial frequency plane





Since $G(f_x, f_y) = \overline{G}(-f_x, -f_y)$ for a *real* image $g(x, y)$, we have *two dominant frequency pairs* in the Fourier–transform integral:

$$|G(u,v)| \cdot [e^{2\pi j(ux+vy)} + e^{-2\pi j(ux+vy)}] = 2|G(u,v)| \cdot \cos 2\pi(ux+vy)$$

This function describes a black-white cosine wave pattern with

$$(f_x, f_y) = (u, v)$$

<u>Where is the value of $G(f_x, f_y)$ large ?</u>

ideally: points (u, v) and $(-u, -v)$ represent *cosine–variant* of the grey values

really: straight line through (u, v) and $(-u, -v)$ represents *abrupt changes* of the grey values

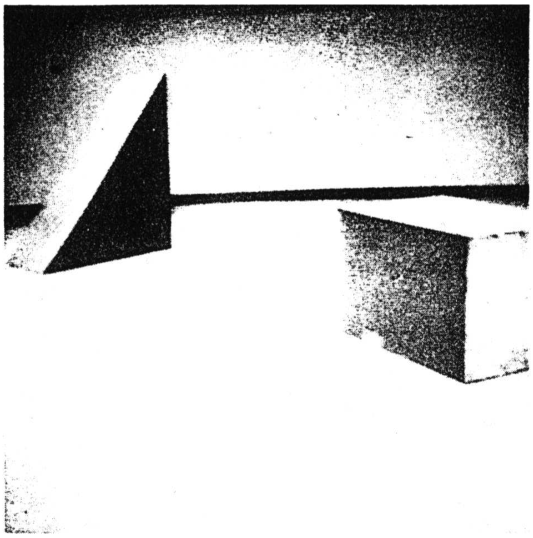


Figure 4.1: TV–image (analog)

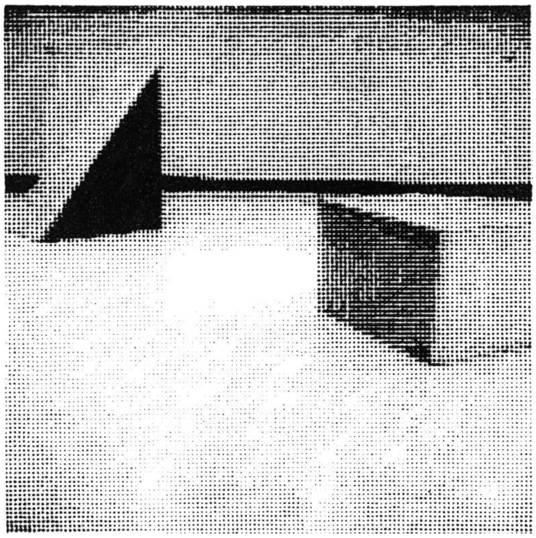


Figure 4.2: Digitized TV–image

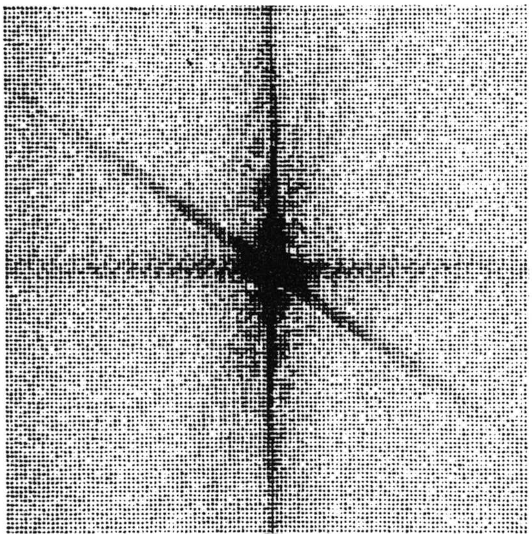


Figure 4.3: Amplitude spectrum of Figure 4.2

Figure 4.4: Low-pass filtered

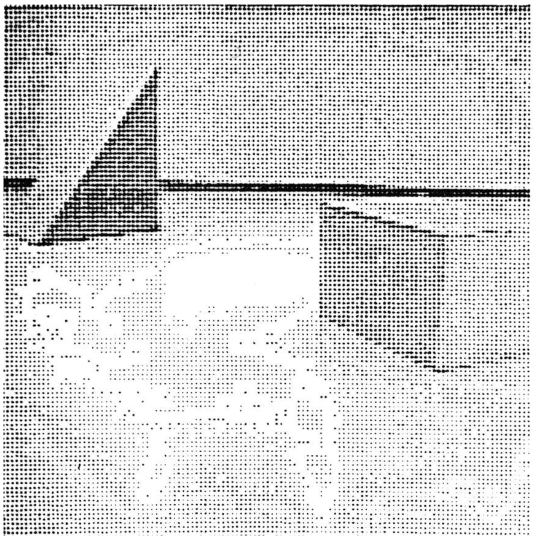


Figure 4.5: High-pass filtered

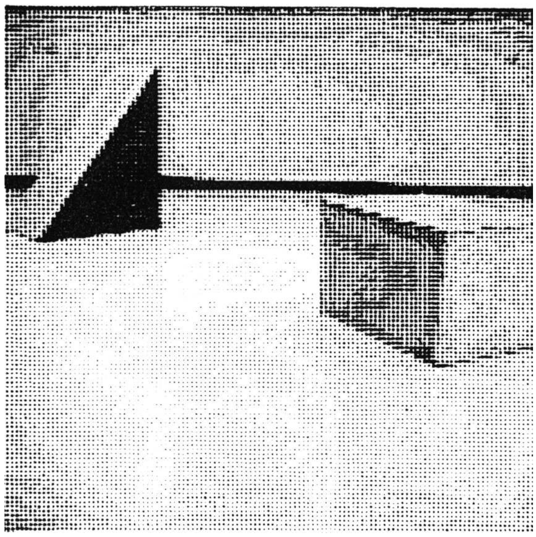


Figure 4.6: High-pass enhancement

Explanation for figures 4.1–4.6 (from Duda & Hart 1973, pp. 310–312):

Figure 4.1:	TV–image (analog)
Figure 4.2:	digitized TV–image - 120×120 pixels - grey values from 0 (black) to 15 (white)
Figure 4.3:	Fourier–transform of the image from Figure 4.2 (amplitude spectrum) $\log\lvert G(f_x, f_y)\rvert$: black $\hat{=}$ high amplitude note: 1. strong components along the axes $\hat{=}$ vertical and horizontal image edges 2. concentration around $(f_x, f_y) = (0, 0)$ $\hat{=}$ regions with constant grey values
Figure 4.4:	Low-pass filter: $H(f_x, f_y) \ = \ [cos(\pi f_x) \cdot cos(\pi f_y)]^{16}$ $0 \leqq H \leqq 1$
Figure 4.5:	High pass filter: $H(f_x, f_y) \ = \ 1.5 - [cos(\pi f_x) \cdot cos(\pi f_y)]^{4}$ $0.5 \leqq H \leqq 1.5$
Figure 4.6:	High pass enhancement: $H(f_x, f_y) \ = \ 2.0 - [cos(\pi f_x) \cdot cos(\pi f_y)]^{4}$ $1.0 \leqq H \leqq 2.0$

Following general rules for $\mathbf{G(f_x, f_y)}$ ensue:

Edges in the image $g(x, y)$:

- An image edge produces "strong" spatial frequency components along one straight line in the spatial frequency plane which is *orthogonal* to the edge.
- The "sharper" the edge is, the "longer" is the corresponding line in the spatial frequency domain.

Regions with constant grey values:

- Regions with constant grey values increase the values of $|G(f_x, f_y)|$ around the origin $(f_x, f_y) = (0, 0)$. $(f_x, f_y) = (0, 0)$ is called "DC component" (average grey value, DC=direct current).

4.2 Discrete Fourier Transform for Images

The analog image $g(x, y)$ is discretized ("sampled") along both axes. We obtain the discrete image:

$$g[j, k] \;:=\; g(j \cdot \Delta x, k \cdot \Delta y) \quad \text{where} \quad j, k \;=\; 0, 1, \ldots, N-1$$

Change in notation: $i \;=\; \sqrt{-1}$ instead of j

$$G(e^{\frac{2\pi i}{N}\cdot u}, e^{\frac{2\pi i}{N}\cdot v}) \;=\; \sum_{j=0}^{N-1}\sum_{k=0}^{N-1} g[j, k] e^{-\frac{2\pi i}{N}(uj\,+\,vk)} \quad \text{where} \quad u, v \;=\; 0, 1, \ldots, N-1$$

Discretization is written as:

$$\begin{aligned} G[u, v] \;&=\; \sum_{j=0}^{N-1}\sum_{k=0}^{N-1} g[j, k]\, e^{-\frac{2\pi i}{N}(uj\,+\,vk)} \\ &=\; \sum_{j=0}^{N-1} e^{-\frac{2\pi i}{N}uj}. \left(\sum_{k=0}^{N-1} g[j, k]\, e^{-\frac{2\pi i}{N}vk} \right) \end{aligned}$$

Interpretation:
Fourier–transform of the image is first performed row by row, then column by column.
Using usual definition of the "Fourier matrix" W (i.e. $W_{vk} \;=\; (e^{-\frac{2\pi i}{N}})^{vk}$), we obtain the matrix representation of Fourier–transform.

Using the notation:

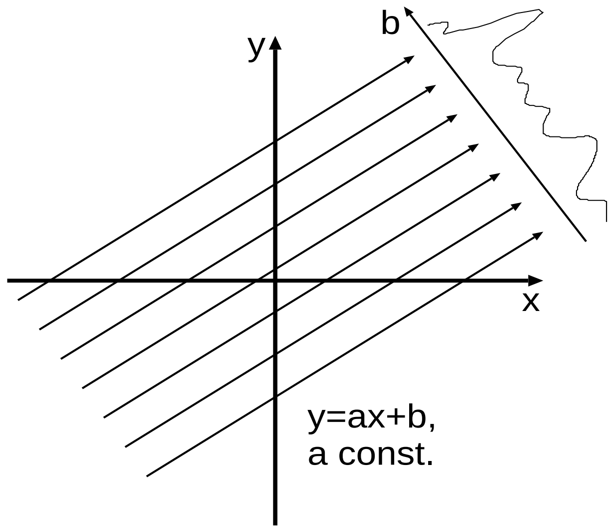
$$\begin{aligned} g \;&\in\; \mathbb{R}^{N\mathrm{x}N} \\ W \;&\in\; \mathbb{C}^{N\mathrm{x}N} \\ G \;&\in\; \mathbb{C}^{N\mathrm{x}N} \end{aligned}$$

we obtain

$$\begin{aligned} G \;&=\; [W\, g\, W] \\ g \;&=\; \frac{1}{N^2} \cdot [W^{-1}\, G\, W^{-1}] \end{aligned}$$

Note: In the corresponding definition of the Fourier–transform, instead of the factors 1 and $1/N^2$ we have the factors $1/N$ and $1/N$ or $1/N^2$ and 1.

4.3 Fourier Transform in Computer Tomography



We consider a projection of the image $g(x, y)$ along the straight line:

$$y \;=\; ax \;+\; b$$

We produce a set of straight lines by keeping a constant and varying b.

Projection:

$$g_a(b) \;=\; \int g(x, ax+b)\, dx$$

Fourier–transform:

$$\begin{aligned} G_a(f_b) \;&=\; \int g_a(b)\, e^{-2\pi j f_b b}\, db \\ &=\; \int\int g(x, ax+b)\, e^{-2\pi j f_b b}\, db\, dx \end{aligned}$$

We substitute $b \;=\; y \;-\; ax$ and obtain:

$$\begin{aligned} G_a(f_b) \;&=\; \int\int g(x, y)\, e^{-2\pi j (y f_b - x a f_b)}\, dy dx \\ &=\; G(-a f_b, f_b) \\ &=\; \text{Fourier–transform } G(f_x, f_y) \text{ of } g(x, y) \text{ along} \\ & \text{the spatial frequency straight line } (f_x, f_y) \text{ with } f_y \;=\; -\frac{1}{a} f_x \end{aligned}$$

Remarks:

a. Straight line in spatial frequency domain: $(f_x, f_y) = (-af_b, f_b)$ is orthogonal to $y = ax + b$:

$$y = ax + b \Rightarrow \text{ in Fourier–transform } f_y = -\frac{1}{a} f_x$$

The angle between these straight lines is a right angle because $a \cdot \left(-\frac{1}{a}\right) = -1$.

In general:

$$\begin{aligned} y_1(x) &= m_1\, x + b_1 \\ y_2(x) &= m_2\, x + b_2 \\ y_1(x) \perp y_2(x) &\Leftrightarrow m_1 \cdot m_2 = -1 \end{aligned}$$

b. The value $G_a(f_b)$ is independent of the "offset" b and depends only on the orientation a of the straight line. Therefore, if we calculate the projection for *many different inclinations* a and apply the one-dimensional FT, we obtain the two-dimensional FT of the image $g(x, y)$.

4.4 Fourier Transform and RST Invariance

We will investigate invariance of the Fourier–transform to

$$\begin{aligned} R &: \text{Rotation} \\ S &: \text{Scaling} \\ T &: \text{Translation} \end{aligned}$$

We will use vector notation for the two-dimensional Fourier–transform:

$$\begin{array}{llcll} \text{coordinates:} & z & = & \begin{pmatrix} x \\ y \end{pmatrix} & \in \mathbb{R}^2 \\ \text{image grey values:} & g(z) & = & g(x,y) & \in \mathbb{R}^+ \\ \text{spatial frequency:} & f & = & \begin{pmatrix} f_x \\ f_y \end{pmatrix} & \in \mathbb{R}^2 \end{array}$$

We ignore the discretization.

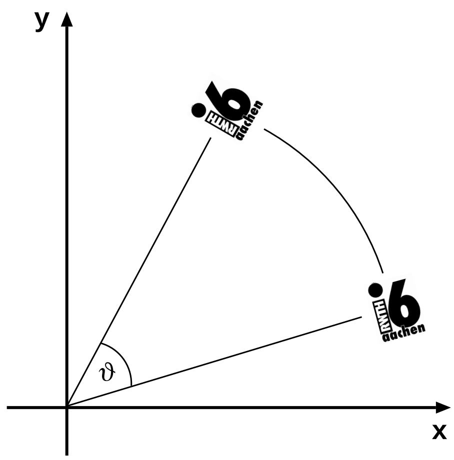
Translation

$$
\begin{aligned}
z &\rightarrow z + z_0 \\
& \quad \text{with translation vector } z_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \in \mathbb{R}^2 \\
\text{Image}: \; g(z) &\rightarrow \tilde{g}(z) := g(z + z_0)
\end{aligned}
$$

$$
\text{FT}: \; \tilde{G}(f) = \exp\left(i[f_x x_0 + f_y y_0]\right) \cdot G(f)
$$

Rotation

$$
\begin{aligned}
z &\rightarrow D_\vartheta z \\
& \quad \text{with rotation matrix } D_\vartheta = \begin{bmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{bmatrix}
\end{aligned}
$$



Scaling

$$z \;\rightarrow\; \alpha \cdot z$$

with scaling factor $\alpha \; > \; 0$

$$\text{Image}: \quad g(z) \;\rightarrow\; \tilde{g}(z) \;=\; g(\alpha \cdot z)$$

$$\text{FT}: \quad \tilde{G}(f) \;=\; \frac{1}{\alpha^2} \cdot G\left(\frac{f}{\alpha}\right)$$

basically: similarity principle (S.**??**) for one-dimensional FT transferred to two dimensions

Invertible linear mapping

$$z \;\rightarrow\; A \cdot z$$

where $A \in \mathbb{R}^{2\times 2}$ invertable

$$\text{Image}: \quad g(z) \;\rightarrow\; \tilde{g}(z) \;=\; g(Az)$$

$$\begin{aligned} \text{FT}: \quad \tilde{G}(f) \;&=\; \ldots \\ &=\; \frac{1}{\det(A)} \cdot \tilde{G}((A^{-1})^T f) \end{aligned}$$

Proof: transformation of two-dimensional integration variables

We apply two basic rules to obtain the RST-invariance:

1. Invariance to translation (=T) can be obtained by using the square of the absolute value
$$g(z) \;\rightarrow\; G(f) \;\rightarrow\; |G(f)|^2$$
2. To obtain RS-invariance we transfer to polar coordinates in the spatial frequency domain. We write in *complex* notation:
$$f_z \;\equiv\; f_x \,+\, i\, f_y \;=\; r\, e^{i\phi} = \exp\left(\ln r \,+\, i\phi\right) \qquad \in \mathbb{C}^2$$

   Complex logarithm:
$$f'_z \;:=\; \ln f_z \;=\; \ln r \,+\, i\phi$$

   We already know:

   a) rotation by angle ϕ_0 in spatial domain
   $\hat{=}$ rotation by angle ϕ_0 in spatial frequency domain

   b) scaling with factor α in spatial domain
   $\hat{=}$ scaling with factor $\left(\dfrac{1}{\alpha}\right)$ and $\left(\dfrac{1}{\alpha}\right)^2$ respectively in spatial frequency domain

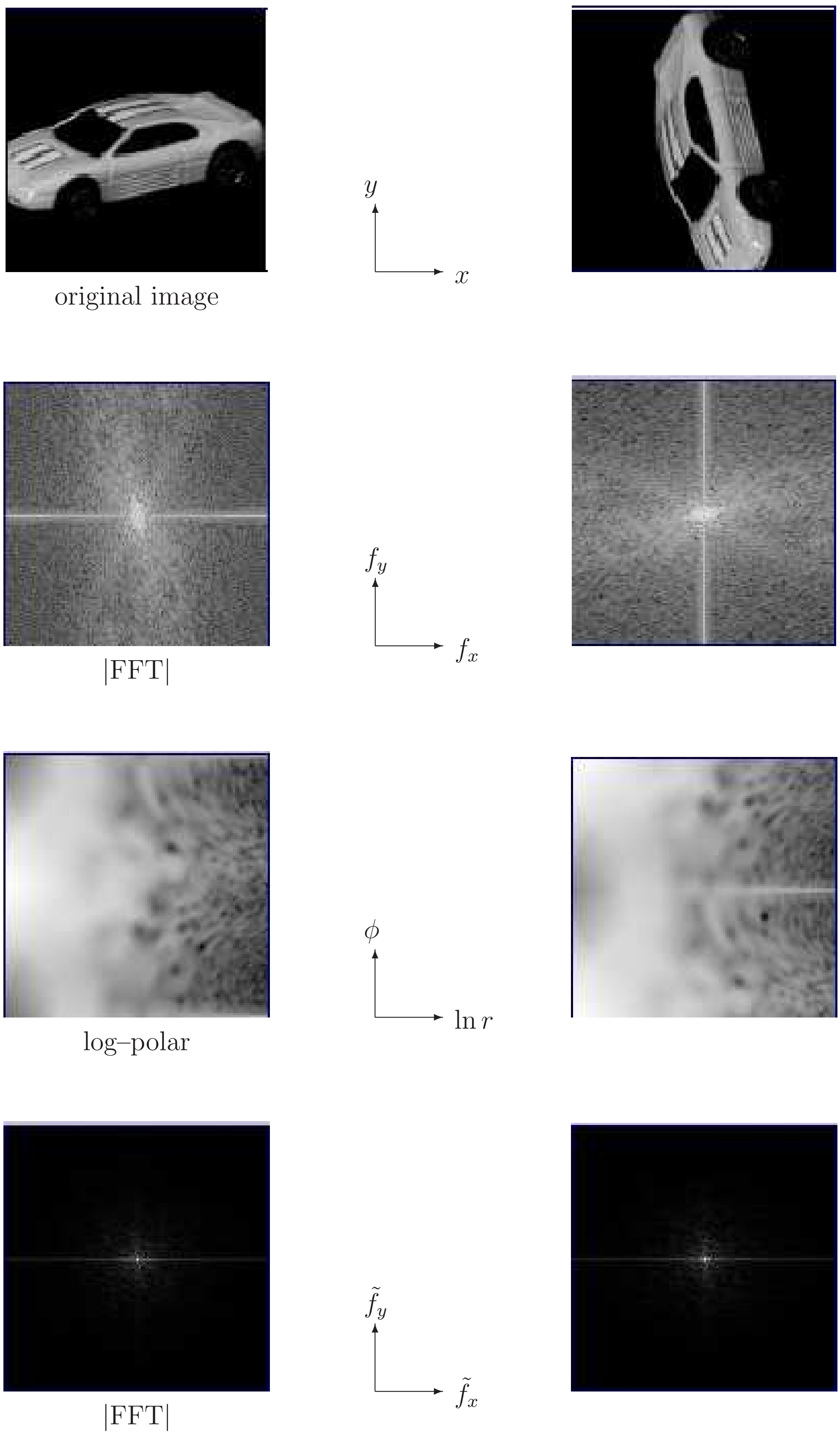
$$f'_z \xrightarrow[\text{rotation}]{\text{scaling and}} \tilde{f}'_z$$

$$
\begin{aligned}
\tilde{f}'_z &= \ln \tilde{z} \\
&= \ln \tilde{r} \,+\, i\tilde{\phi} \\
&= \ln \left(\frac{r}{\alpha} \right) \,+ i(\phi \,-\, \phi_0) \\
&= \ln r \,+\, i\,\phi \,-\, \ln \alpha \,-\, i\,\phi_0 \\
&= f'_z \,-\, \ln \alpha \,-\, i\,\phi_0 \\
&= \text{translation with the shift vector} \quad (-\ln \alpha \,-\, i\,\phi) \in \mathbb{C}^2 \\
& \text{ in \textit{logarithmic polar coordinates} of the spatial frequency plane}
\end{aligned}
$$

RST-invariant features can therefor be obtained as follows:

$$
\begin{array}{ll}
g(x,y) & \text{image: } (x,y) \in \mathbb{R}^2 \\
\downarrow & \text{first Fourier–transform} \\
G(f_x,f_y) & = F\{g(x,y)\} \quad \text{mit } (f_x,f_y) \in \mathbb{R}^2 \\
\downarrow & \boxed{\text{squared absolute value}} \\
|G(f_x,f_y)|^2 & \\
\downarrow & \text{logarithmic polar coordinates:} \quad (\ln r, \phi) \\
|G(\ln r,\phi)|^2 & \\
\downarrow & \text{second Fourier–transform} \\
F(|G(\ln r,\phi)|^2) & \\
\downarrow & \boxed{\text{squared absolute value}} \\
|F(|G(\ln r,\phi)|^2)|^2 &
\end{array}
$$

Next page:
Analysis of the RST-invariant features.
Original grey valued images are identical up to a 90°–rotation.


original image
y
x
|FFT|
f_y
f_x
log–polar
ϕ
$\ln r$
|FFT|
$\tilde{f}_y$
$\tilde{f}_x$

Warning:

a) Invariant observations are not necessarily good for classification.

b) Observations that are calculated using the two-dimensional Fourier–transform are not complete, i.e. the original image *cannot* be reconstructed completely.

Chapter 5

LPC Analysis

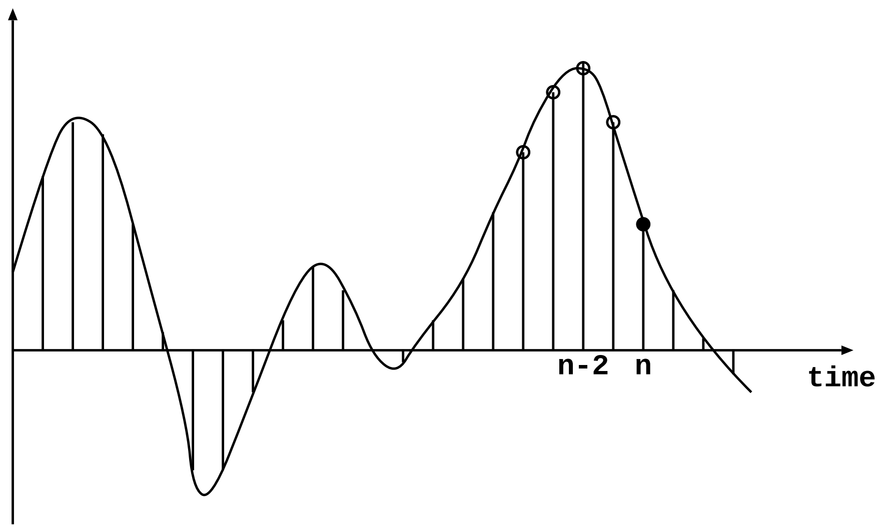
- Overview:

The acronym *LPC* stands for

Linear Predictive Coefficients / Coding

and is utilized in signal processing and frequency analysis, as well as in signal coding.

5.1 Principle of LPC Analysis



We consider a discrete time signal $x[n]$, possibly multiplied with a window function. The goal of an LPC analysis is to predict each signal value $x[n]$ by its preceding values $x[n-1], x[n-2], ..., x[n-K]$. We distinguish:

$$\begin{aligned} x[n]: &\quad \text{signal value} \\ \hat{x}[n]: &\quad \text{predicted value} \end{aligned}$$

We assume the predicted value $\hat{x}[n]$ to be a linear combination of the preceding values of $x[n]$:

$$\hat{x}[n] \; := \; \sum_{k=1}^{K} \alpha_k \, x[n-k]$$

with at first unknown coefficients α_k, $k = 1, ..., K$, which are called

LPC–coefficients or prediction coefficients.

The value K is called prediction order, e.g. $K = 8, \ldots, 10$ at a sampling frequency of 4 kHz (about 2 coefficients per kHz).

Outlook

Starting point: “coding” in time domain (goal: bit reduction)

$\downarrow$ Parseval Theorem

parametric model for power spectrum of Fourier–transform
(more exact: rough structure of power spectrum for speech signal)

LPC analysis applications:

- speech coding
  (ADPCM = adaptive differential pulse code modulation)
- signal processing:
  parametric modelling with autoregressive or all-pole models (order K)
- time curves:
  resonance and oscillator curves, sun spots, stock-market course, ...
- image coding
- also: interpretation as Maximum Entropy Approach

The coefficients α_k are unknown at first. To estimate these, we define the prediction error for each point n in time:

$$\begin{aligned} e[n] &:= x[n] - \hat{x}[n] \\ &= x[n] - \sum_{k=1}^{K} \alpha_k \, x[n-k] \end{aligned}$$

For a reliable set of LPC–coefficients we calculate the squared error criterion E as sum of the squared prediction errors $e[n]$:

$$\begin{aligned} E &= \sum_n e^2(n) \\ &= \sum_n \left[x[n] - \sum_{k=1}^{K} \alpha_k \, x[n-k] \right]^2 \\ &\stackrel{!}{=} \quad \text{minimum with respect to } \alpha_1, \ldots, \alpha_k, \ldots, \alpha_K \end{aligned}$$

Taking the derivative $\frac{\partial}{\partial \alpha_l}$ for $l = 1, \ldots, K$ results in:

$$\sum_n \left(x[n] - \sum_k \alpha_k \, x[n-k] \right) x[n-l] \stackrel{!}{=} 0$$

$$\sum_k \alpha_k \sum_n x[n-k] x[n-l] = \sum_n x[n-l] x[n]$$

Here, the summation limits are not specified on purpose.
If the squared error criterion E is considered as a function of LPC–coefficients, the following properties ensue:

- E is quadratic in $\alpha_1, \ldots, \alpha_k, \ldots, \alpha_K$; it is guaranteed to be non-negative and it has a single well-defined minimum.
- The optimal LPC–coefficients are invariant to linear scaling of the signal values $x[n]$.

Minimization of the squared error criterion with respect to the LPC–coefficients results either from taking the derivative or from the "quadratic complement" (recalculate for yourself!). The linear equation system for the LPC–coefficients α_k ensues:

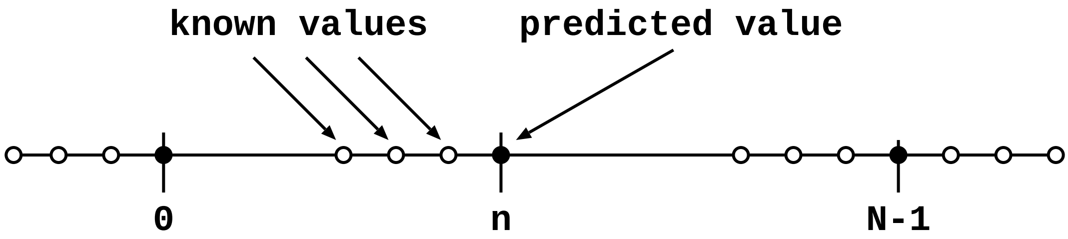
$$l = 1, \ldots, K: \qquad \sum_{k=1}^{K} \alpha_k \cdot \sum_n x[n-k]\, x[n-l] = \sum_n x[n-l]\, x[n]$$

with still unspecified summation limits over n. We consider two methods for the choice of summation limits:

1. covariance method
2. autocorrelation method

Warning: terminology is *not* consistent.

5.2 LPC: Covariance Method



Covariance Method

- *No* window function is applied, such that we obtain the following summation limits:

$$\sum_{n} e^2(n) \;=\; \sum_{n=0}^{N-1} e^2(n)$$

  i.e. we also use signal values $x[n]$ with $n < 0$ for prediction.

- The resulting equation system for LPC–coefficients:

$$l = 1, \ldots, K : \quad \sum_{k=1}^{K} \alpha_k \, \Phi(l,k) = \Phi(l,0)$$

  with the definition:

$$\Phi(l,k) \;:=\; \sum_{n=0}^{N-1} x[n-l] \, x[n-k]$$

  For the above terms hold:

  – they describe a kind of cross correlation between two “signals”
  – they are similar to a *covariance* matrix

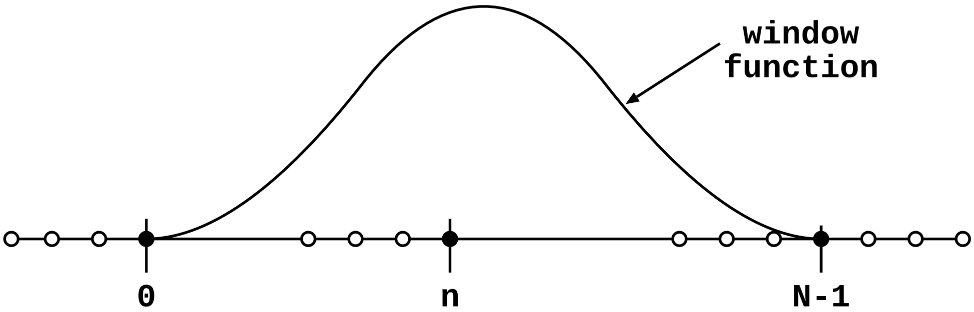
  Computational complexity for solving the equation system:

$$O(K^3) \,+\, O(NK)$$

  – autocorrelation method has more favorable complexity: $O(K^2)$
  – but: calculation of auto/cross-correlation function dominates

In contrast to covariance method, autocorrelation method offers an interpretation in the frequency domain and therefore is often preferred.

5.3 LPC: Autocorrelation Method



We consider the signal after multiplication with a convenient window function, usually Hamming window:

In principle, the summation limits now are

$$\sum_{n} e^2[n] \;=\; \sum_{n=-\infty}^{n=+\infty} e^2[n] \; .$$

Since, due to windowing the signal $x[n]$ is identical to zero outside the window function, i.e.

$$x[n] \equiv 0 \qquad \text{for } n < 0 \text{ or } N-1 < n$$

we obtain the following for the prediction error $e[n]$:

$$e[n] \equiv 0 \qquad \text{for } n < 0 \text{ or } N-1+K < n$$

Therefore, the total error E becomes:

$$E \;=\; \sum_{n=0}^{N+K-1} \; e^2[n]$$

The prediction error $e[n]$ can become "large" on the window function boundaries:

- Beginning:	prediction from "zeros"
- End:	prediction of "zeros"

Inserting the summation limits:

$$\begin{aligned}
\sum_n x[n-k]\, x[n-l] &= R(|l-k|), \quad \sum_n x[n]\, x[n-l] = R(|l|) \\
\text{where} \quad R(|l-k|) &= \sum_{n=0}^{N-1-l} x[n-k]\, x[n-l] \\
R(|l|) &= \sum_n x[n]\, x[n-l] = \sum_{n=0}^{N-1-l} x[n]\, x[n-l]
\end{aligned}$$

In this way we obtain the following equation system for the LPC–coefficients α_k:

$$l = 1, ..., K: \quad \sum_{k=1}^{K} \alpha_k R(|l-k|) = R(l)$$

or in matrix form:

$$\begin{bmatrix}
R(0) & R(1) & \dots & & R(K-1) \\
R(1) & R(0) & \dots & & R(K-2) \\
\vdots & & & & \vdots \\
& & \ddots & & \vdots \\
\vdots & & & & R(1) \\
R(K-1) & R(K-2) & \dots & R(1) & R(0)
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\ \alpha_2 \\ \\ \vdots \\ \\ \alpha_K
\end{bmatrix}
=
\begin{bmatrix}
R(1) \\ R(2) \\ \\ \vdots \\ \\ R(K)
\end{bmatrix}$$

Note that this equation system is completely determined by the autocorrelation coefficients

$$R(0), ..., R(k), ..., R(K).$$

Hence, the autocorrelation coefficients will "only" be converted to obtain the LPC–coefficients

$$\alpha_1, ..., \alpha_k, ..., \alpha_K.$$

The matrix of this equation system has the following properties:
- Toeplitz structure (follows from time invariance)
- solution: Durbin–algorithm with complexity $O(K^2)$

5.4 LPC: Interpretation in Frequency Domain

The LPC autocorrelation method allows prediction error conversion from time domain into frequency domain using Parseval theorem so that LPC analysis can be interpreted as adaptation of parametric model spectrum to the observed signal spectrum.
We start with the prediction error $e[n]$:

$$e[n] \;=\; x[n] - \sum_{k=1}^{K} \alpha_k \, x[n-k]$$

and apply the z–transform to this equation. The z–transform is restricted to the unit circle.

$$z \;=\; e^{j\omega} \qquad \in \mathbb{C}$$

For the z-transforms $E(z)$ and $X(z)$ we obtain:

$$E(z) \;=\; X(z) \cdot \left[1 - \sum_{k=1}^{K} \alpha_k z^{-k} \right]$$

The total error E_{tot} for the squared error criterion becomes:

$$\begin{aligned}
E_{tot} &= \sum_{n=0}^{N+K-1} e^2[n] \\
&= \frac{1}{2\pi} \int_{-\pi}^{+\pi} |E(e^{j\omega})|^2 \, d\omega \qquad \text{(Parseval Theorem)} \\
&= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left| 1 - \sum_{k=1}^{K} \alpha_k \, e^{-j\omega k} \right|^2 \cdot |X(e^{j\omega})|^2 \, d\omega \\
&= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left| P(e^{j\omega}) \right|^2 \cdot |X(e^{j\omega})|^2 \, d\omega
\end{aligned}$$

with the so-called predictor polynom:

$$P(e^{j\omega}) \;:=\; 1 - \sum_{k=1}^{K} \alpha_k \, e^{-j\omega k}$$

Squared absolute value of the predictor polynom

$$\begin{aligned} \left|P(e^{j\omega})\right|^2 \;&=\; \left|1 - \sum_{k=1}^{K} \alpha_k \, e^{-j\omega k}\right|^2 \\ &=\; ... \\ &=\; \sum_{k=1}^{K} B_k \cdot \cos(\omega k) \end{aligned}$$

(with suitable coefficients B_k resulting from the predictor coefficients) is a polynom with respect to $\cos(\omega)$, which can be obtained via application of trigonometric transformations.

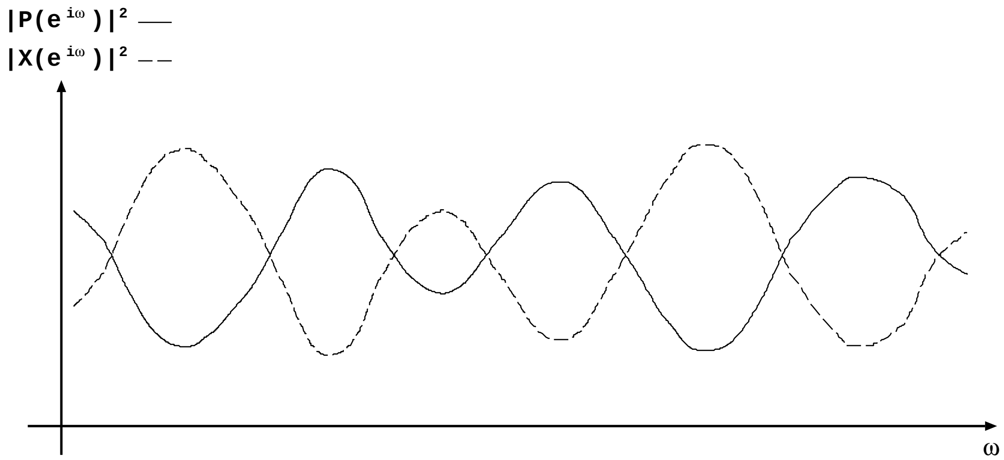
The predictor polynom tries to "compensate" for $|X(e^{j\omega})|^2$ – especially at maxima – and to generate a "white" spectrum for the prediction error $e[n]$.

The *complex* predictor polynom $P(z)$ with $z \in \mathbb{C}$ has exactly K zeros in the complex plane and therefore can be factorised into linear factors:

$$P(z) \;=\; \prod_{k=1}^{K} (z - z_k)$$

Observations:

- These zeros are complex conjugated pairs because $\alpha_k \in \mathbb{R}$.
- The zeros *can* cause "minima" of $\left|P(e^{j\omega})\right|^2$. The minima of $|P(e^{j\omega})|^2$ approximately correspond to the maxima of the *smoothed* spectrum $|X(e^{j\omega})|^2$, because for minimization of the error integral it is first of all necessary to "compensate" for the maxima of the signal spectrum. The LPC analysis could therefore be used to describe of the speech signal formant structure.



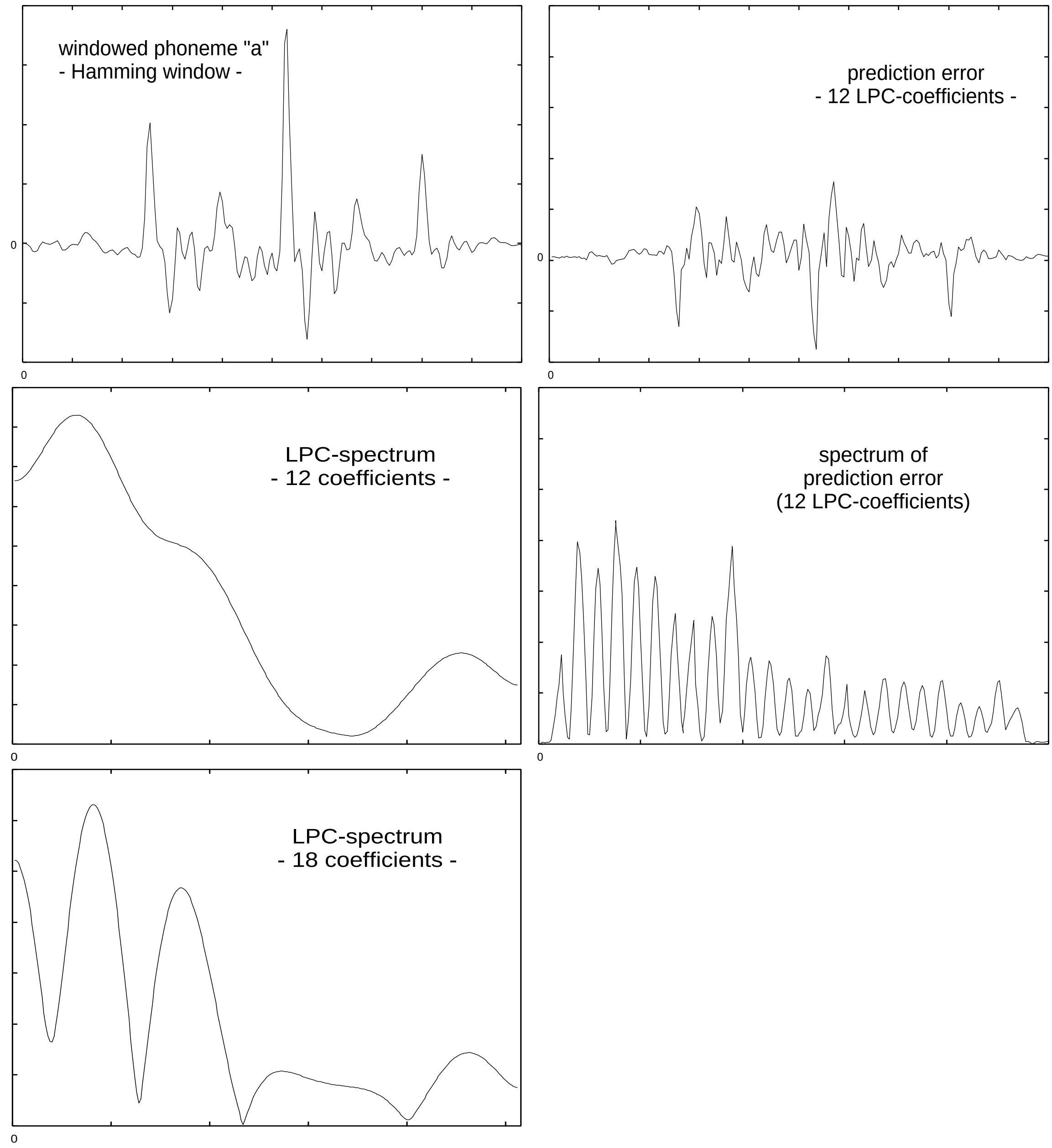


Figure 5.1: LPC–analysis of one speech segment
a) signal progression, b) prediction error (K=12), c) LPC–spectrum with K=12 coefficients, d) spectrum of the prediction error (K=12), e) LPC–spectrum with K=18 coefficients

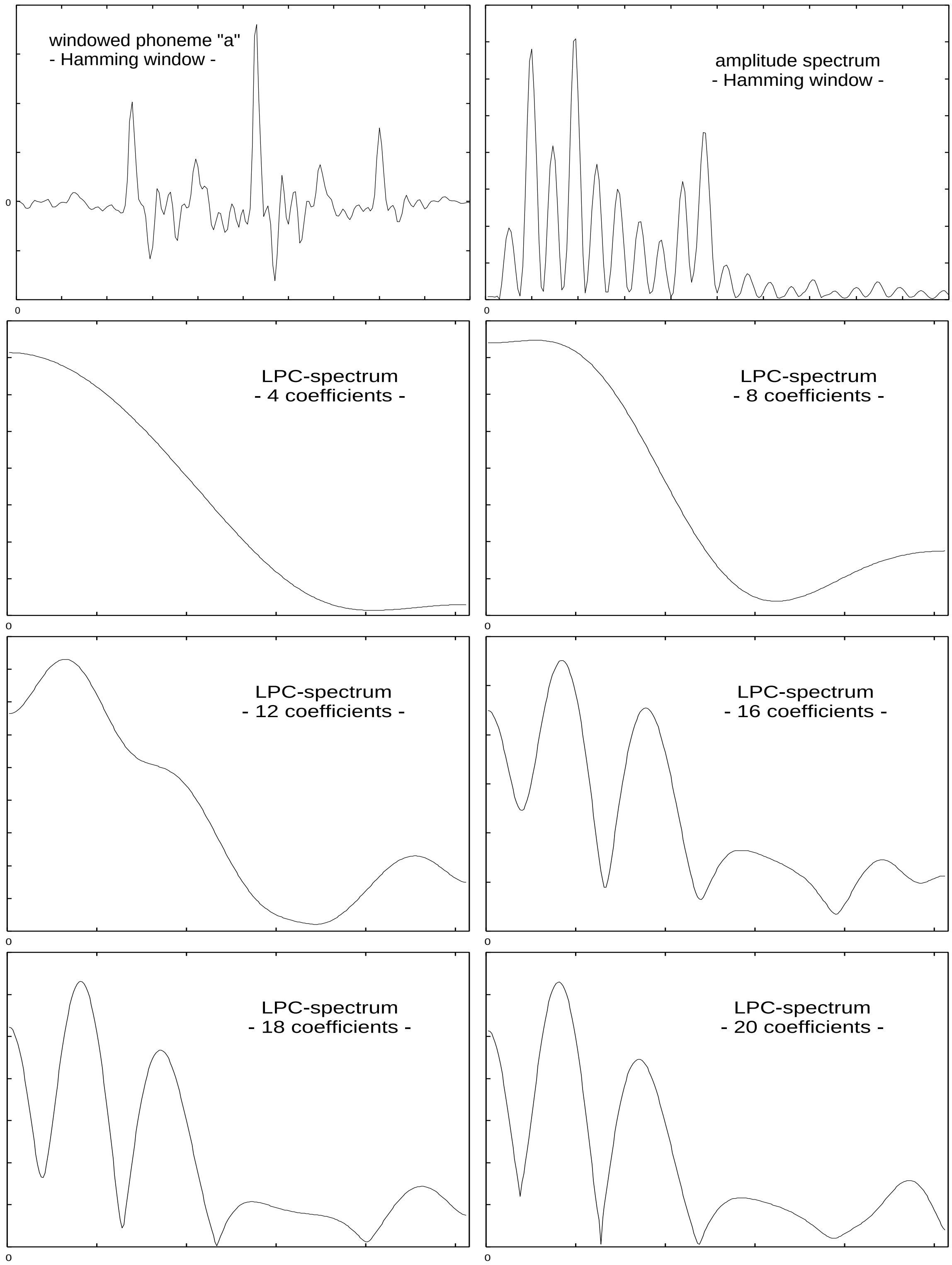
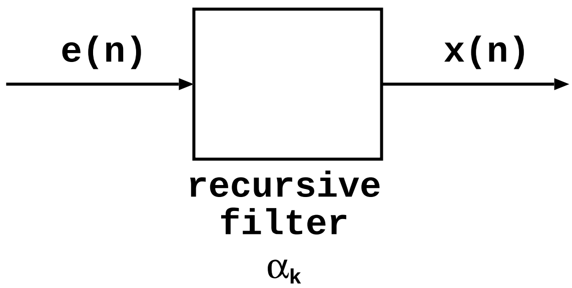


Figure 5.2: LPC–Spectra for different prediction orders K

5.5 LPC: Generative Model



For the prediction error $e[n]$ and its z–transform holds:

$$\begin{aligned} e[n] &= x[n] - \sum_{k=1}^{K} \alpha_k \, x[n-k] \\ E(z) &= X(z) - \sum_{k=1}^{K} \alpha_k \, X(z) \, z^{-k} \\ &= X(z) \cdot [1 - \sum_{k=1}^{K} \alpha_k \, z^{-k}] \end{aligned}$$

If we consider prediction error as input signal, we can also interpret the LPC–theorem as generative model which generates an output signal $x[n]$ from an adequate "input signal" $e[n]$:

$$x[n] = e[n] + \sum_{k=1}^{K} \alpha_k \, x[n-k] \; .$$

For the signal spectrum $X(z)$ holds:

$$X(z) = \frac{E(z)}{1 - \sum_{k=1}^{K} \alpha_k \, z^{-k}}$$

This model is called *autoregressive* model. The excitation has to be chosen such that $E(z)$ is "white", i.e. it does not have fine structure due to the fundamental frequency ("pitch–frequency").
In other words:

$$E(z) = G = \text{const. ("gain")}$$

Special case:

$$E[n] \;=\; G \cdot \delta[n]$$

Then for LPC model spectrum $X(z)$ holds:

$$X(z) \;=\; \frac{G}{1 - \sum\limits_{k=1}^{K} \alpha_k \, z^{-k}}$$

This spectrum is often interpreted as LPC model spectrum $X(z)$ of observed signal. It is reasonable to set (without explanation):

$$G^2 \;=\; R(0) - \sum_{k=1}^{K} \alpha_k \, R(k) \;=\; R(0) \left[1 - \sum_{k=1}^{K} \alpha_k \, \frac{R(k)}{R(0)} \right]$$

This LPC model spectrum *does not have any* zeros, it has *only* poles, and therefore is also called all–pole model.

Remarks:

- stability problems by solving the equation system
  ($\longleftarrow$ truncation error in autocorrelation)
- way out: preemphasis through difference calculation
- absolute rule for choice of order K:

  1 formant needs 2 LPC–coefficients

  1 formant per kHz

  + excitation pulse shape + radiation: 2 LPC–coefficients

  $\Longrightarrow$ rule of thumb:

bandwidth	
$4\,kHz$	$K = 10$
$5\,kHz$	$K = 12$
$6\,kHz$	$K = 14$

5.6 LPC: Alternative Representations

- so far:
  G gain
  α_k LPC–coefficients

- impulse response of generative model

- impulse response of squared absolute value of ”predictor polynom”

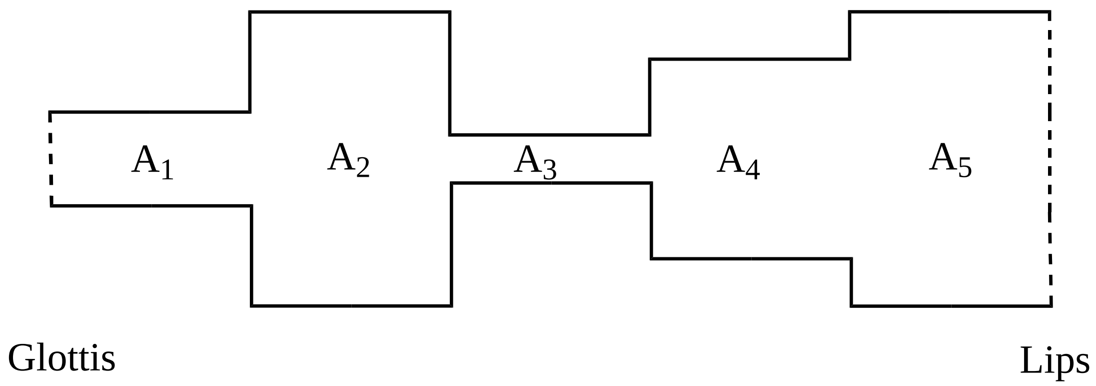
- cepstrum

- poles / zeros of synthesis model / ”predictor polynom”
  $\Longrightarrow$ formants / bandwidths
  problem: noise susceptible

- PARCOR–coefficients: partial correlation

- Area–coefficients: cross-section surfaces A_k

- reflexion coefficients $\sim$ PARCOR; tube model



Chapter 6

Outlook: Wavelet Transform

- Overview:

  6.1 Motivation: from Fourier to Wavelet Transform

  6.2 Definition

  6.3 Discrete Wavelet Transform

6.1 Motivation: from Fourier to Wavelet Transform

Fourier transform uses infinitely extended basis functions

$$e^{-j\omega t}$$

and therefore does *not* have any time resolution, i.e. there is no information about the localization along the time axis.

Therefore, a window function often is used

$$t \;\rightarrow\; w(t), \qquad \text{complex in general}$$

This function has finite support such that it is possible to investigate a segment of a function of interest.

We define a short–time Fourier transform $F_b(w)$ of time signal $t \;\rightarrow\; f(t)$ at position $b \in \mathbb{R}$ in time:

$$F_b(w) \;:=\; \int\limits_{-\infty}^{+\infty} f(t)\overline{w}(t-b)e^{-j\omega t}\,dt$$

where $\overline{w}(t)$ denotes the complex conjugated value of $w(t)$.

The wavelet–transform can be derived from this equation in two steps:

a) we ignore the basis function $e^{-j\omega t}$ and we define the window function as the new basis function.

b) in addition to the *localization parameter* b, we also introduce a *scaling parameter* $a > 0$.

We consider the family of window functions $W_{ab}(t)$:

$$w_{ab}(t) \;:=\; w\left(\frac{t-b}{a}\right)$$

6.2 Definition

The following notation is usually used for the Wavelet–transform:

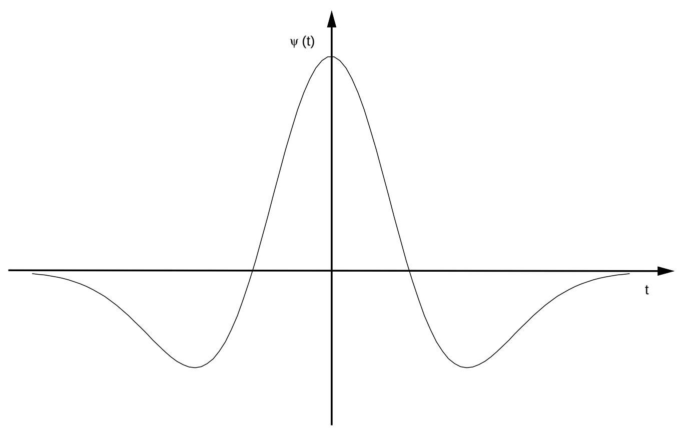
$$\psi_{ab}(t) \ := \ \frac{1}{\sqrt{a}} \cdot \psi\left(\frac{t-b}{a}\right)$$

which is the so-called *Mother-Wavelet* $t \to \psi(t)$.

Like the window function for the short–time Fourier transform the Mother-Wavelet should be "localized" as much as possible.

<u>Example:</u>

Mexican-Hat Function: $\quad \psi(t) \ = \ (1-t^2)e^{-\frac{1}{2}t^2}$



The *wavelet–transform* of $f(t)$ with respect to $\psi(t)$ is defined as:

$$F_\psi(a,b) \;=\; \frac{1}{a}\int\limits_{-\infty}^{+\infty} f(t)\;\cdot\;\overline{\psi}\left(\;\frac{t-b}{a}\;\right)\;dt$$

with scaling parameter $a > 0$ and localization parameter $b \in \mathbb{R}$.

For the inverse transformation holds:

$$f(t) \;=\; \frac{1}{C_\psi}\cdot\frac{1}{a^2}\cdot\int\limits_{0}^{+\infty} da \int\limits_{-\infty}^{+\infty} db\, F_\psi(a,b)\;\cdot\;\psi\left(\;\frac{t-b}{a}\;\right)$$

with

$$C_\psi \;:=\; \int\limits_{0}^{\infty}\frac{|\psi(\omega)|^2}{\omega}d\omega \;<\; \infty$$

Proof (principle only):

The proof uses the (generalized) Parseval Theorem:

$$\begin{aligned}
F_\psi(a,b) \;&=\; \frac{1}{\sqrt{a}}\int\limits_{-\infty}^{+\infty} f(t)\;\cdot\;\overline{\psi}_{ab}(t)\,dt \\
&=\; \frac{1}{\sqrt{a}}\int\limits_{-\infty}^{+\infty} F(\omega)\;\cdot\;\psi_{ab}(\omega)\,d\omega \\
\text{with}\qquad \psi_{ab}(t) \;&=\; \frac{1}{\sqrt{a}}\cdot\psi\left(\frac{t-b}{a}\right)
\end{aligned}$$

using further conversions.

6.3 Discrete Wavelet Transform

For the scaling parameters $a > 0$ we choose:

$$a \;=\; \alpha_0^m \quad \text{where} \quad \alpha_0 > 1 \quad \text{and} \quad m \in \mathbb{Z}.$$

The values of m determine the width of wavelet $\psi_{ab}(t)$.

In order to adjust the localization parameter properly, we define:

$$b \;=\; n\, b_0\, a_0^m \quad \text{where} \quad b_0 > 0 \quad \text{and} \quad n \in \mathbb{Z}.$$

Thus we constrain the Wavelet transform to *discrete values*:

$$F_\psi(a,b) \qquad \rightarrow \qquad F_\psi(n,m)$$

The choice of the function $\psi(t)$ is still open.

It is useful to choose $\psi(t)$ such that the function system $\{\psi_{mn}|m,n \in \mathbb{Z},\ m > 0\}$

$$\text{with} \quad \psi_{mn}(t) \;:=\; a_0^{-\frac{m}{2}} \cdot \psi(a_0^{-m} t \;-\; n b_0)$$

represents an orthonormal basis for functions $t \rightarrow f(t) \quad \in L^2(\mathbb{R})$.

Note: The scalar product $< f(t), g(t) >$ of two functions $f(t)$ and $g(t)$ is defined as:

$$< f(t), g(t) > = \int f(t) \cdot \overline{g}(t)\, dt$$

In this way we obtain the following representation for the *discrete Wavelet–transform*:

$$\begin{aligned} F_\psi(m,n) &= \int\limits_{-\infty}^{+\infty} f(t) a_0^{-\frac{m}{2}} \psi(a_0^{-m} t \,-\, n b_0)\, dt \\ &= < f(t), \psi_{mn}(t) > \end{aligned}$$

Due to the orthogonality it is possible to convert the integral of the inverse Wavelet–transform into an infinite series:

$$\begin{aligned} f(t) &= \frac{1}{C_\psi} \sum_m \sum_n F_\psi(m,n) \psi_{mn}(t) \\ &= \frac{1}{C_\psi} \sum_m \sum_n F_\psi(m,n)\, a_0^{-\frac{m}{2}} \psi(a_0^{-m} t \,-\, n b_0) \end{aligned}$$

Example: *Haar*–function and *Haar*–basis

special choice: $a_0 = 2$
$b_0 = 1$

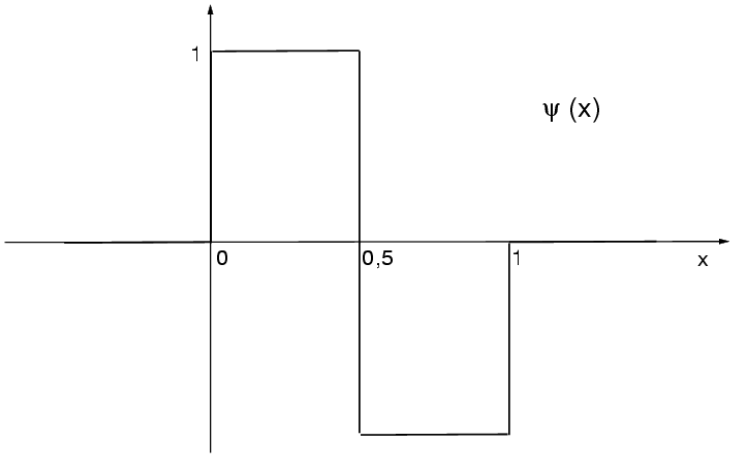
The *Haar*–function is defined as:

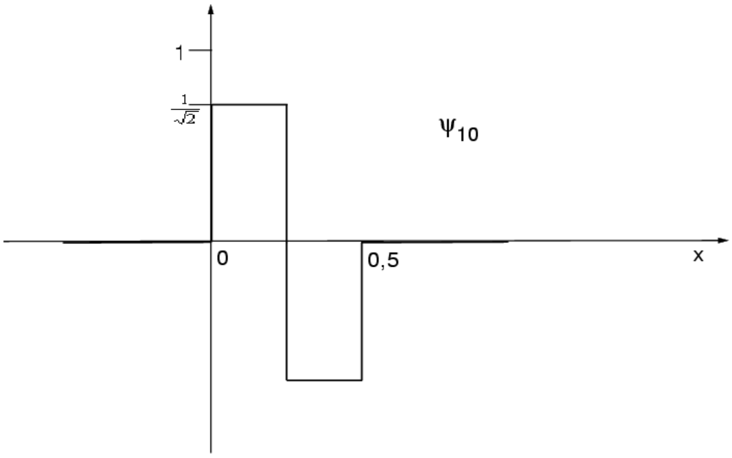
$$\psi(t) = \begin{cases} 1 & 0 \leqq t \leqq \frac{1}{2} \\ -1 & \frac{1}{2} \leqq t < 1 \\ 0 & otherwise \end{cases}$$

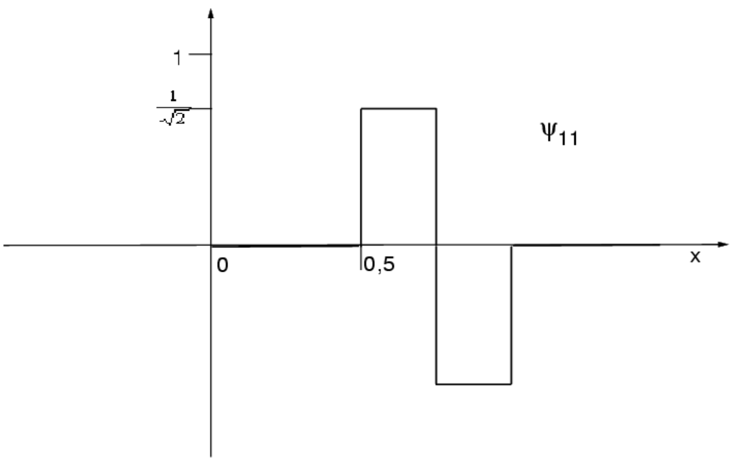
This defines the *Haar*–basis $\left\{ \psi(t) \mid m, n \in \mathbb{Z}, m > 0 \right\}$:

$$\psi_{mn}(t) = \frac{1}{\sqrt{2^m}} \cdot \psi \left(2^m t - n \right)$$

It is easy to see that for increasing m a increasingly finer resolution is obtained and that n determines localization in time.


1
ψ (x)
0
0,5
1
x


1
$\frac{1}{\sqrt{2}}$
ψ_{10}
0
0,5
x


1
$\frac{1}{\sqrt{2}}$
ψ_{11}
0
0,5
x

Chapter 7

Coding

The following types of coding are distinguished:

- source coding (data compression)
  goal: transmission (storage) using *as few bits as possible without or with few* errors
- channel coding
  goal: preferably faultless data transmission (storage)
  e.g. error-recognizing and error-correcting codes
- simultaneous source and channel coding
  goal: simultaneous optimization

The following data types are distinguished:

- discrete alphabet
- continuous signal (audio, video, ...)

Source coding

- lossless coding (compression)
  usually discrete sources, e.g. text compression
- lossy coding
  usually continuous signals
  notation:
  rate - distortion theory

  distortion, error

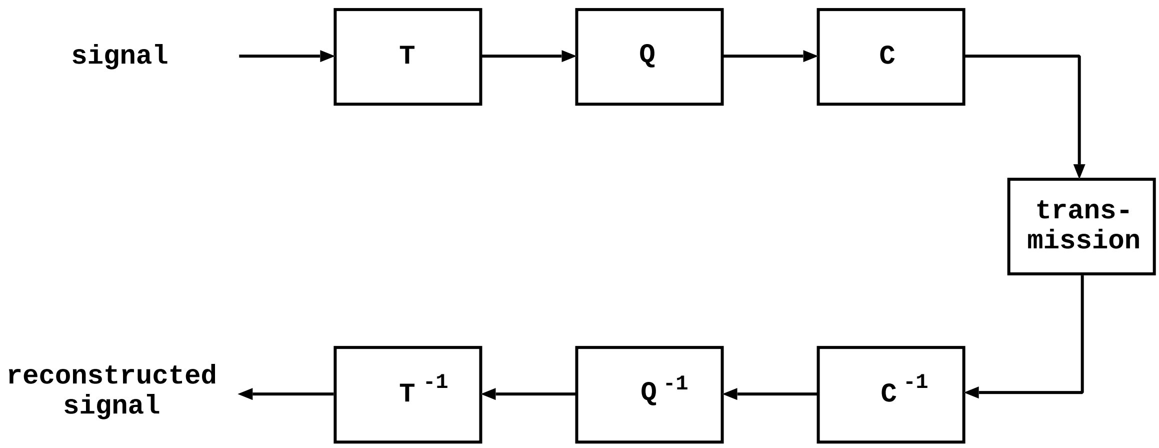
  bit rate

Three effects can be utilized for signal coding:

a) statistical redundancy and correlation:
   samples are not independent.

b) perceptive properties of the receiver (ear and eye):
   some fine structures in the signal are irrelevant to the receiver

c) signal distortion:
   coded signal differs from the original signal without significant quality deterioration.



T: transformation, e.g. DCT

Q: quantization, e.g. vector quantization

C: mapping of bit representation

References:

- **Ze-Nian Li: CMPT 365 Multimedia Systems. Simon Fraser University, British Columbia, Canada, fall 1999, Version Jan.2000;**

  `http://www.cs.sfu.ca/CourseCentral/365/li/index_prev.html`.

- Peter Noll: MPEG Digital Audio Coding. IEEE Signal Processing Magazine, pp.59-81, Sep. 1997.
- Thomas Sikora: MPEG Digital Video-Coding Standards. IEEE Signal Processing Magazine, pp.82-100, Sep. 1997.
- A. Ortega, K. Ramchandran: Rate-Distortion Methods for Image and Video Compression. IEEE Signal Processing Magazine, pp.23-50, Nov. 1998.
- G. J. Sullivan, Th. Wiegand: Rate-Distortion Optimization for Video Compression. IEEE Signal Processing Magazine, pp.74-90, Nov. 1998.

Chapter 8

Image Segmentation and Contour-Finding

The lecture notes for this chapter are available as a separate document.