Lecture Notes

Digital Communications

Prof. Dr.-Ing. Johannes Huber
Prof. Dr.-Ing. Robert Schober

Friedrich-Alexander-Universität Erlangen-Nürnberg
Institute for Digital Communications

WS 2022/2023

Contents

Organizational Details

Classification: Compulsory in Master Program CME
Elective in EEI / IuK / CE / WING-IKS

Credit: 3 SWS lecture, 1 SWS tutorial, 5 ECTS

Location: Room 01.021, Cauerstr. 7

Time: Wednesday, $14^{15}-15^{45}$: Lecture
Thursday, $10^{15}-11^{45}$: Lecture / Tutorial (in a two week rotation)

Lecture: Prof. Dr. Laura Cottatellucci
Chair for Digital Communications
Cauerstraße 7, 5. floor, room 05.037
`laura.cottatellucci@fau.de`

Prof. Dr.-Ing. Robert Schober
Chair for Digital Communications
Cauerstraße 7, 5. floor, room 05.034
`robert.schober@fau.de`

Tutorial: Brikena Kaziu, M. Sc.
Chair for Digital Communications
Cauerstraße 7, 4. floor, room 04.035
`brikena.kaziu@fau.de`

Examination: Written exam
in first exam period (directly after the university's lecture session)

StudOn: `https://www.studon.fau.de/crs4622778.html`

Overview

- **Introduction**
- **Basics Concepts and Terminology**
- **Digital Pulse Amplitude Modulation (PAM)**
- **Variants of PAM-Transmission**
- **Representation of Signals in Signal Space**
- **Digital Frequency Modulation and Continuous Phase Modulation**
- **Digital Communication over Frequency Selective (Dispersive) Channels, Equalization of Data Signals**
- **ML Sequence Estimation**
- **Orthogonal Frequency Division Multiplexing (OFDM)**

Relevant Literature (Sample)

General Foundations:

[1] K.D. Kammeyer, A. Dekorsy. *Nachrichtenübertragung.* Springer Vieweg, Wiesbaden, 6-th edition, 2018.

[2] S. Haykin. *Communication Systems.* John Wiley & Sons, Inc., New York, 5-th edition, 2009.

Digital Communications:

[1] J.G. Proakis, M. Salehi. *Digital Communications.* McGraw–Hill, New York, 5-th edition, 2008.

[2] N. Benvenuto, G. Cherubini. *Algorithms for Communications Systems and their Application.* John Wiley & Sons, Inc., New York, 2-nd edition, 2021.

[3] R.E. Blahut. *Digital Transmission of Information.* Addison–Wesley, Reading, Mass., 1990.

[4] J.B. Anderson. *Digital Transmission Engineering.* Wiley–IEEE Press, Piscataway, NJ, 2-nd edition, 2005.

[5] J.R. Barry, E.A. Lee, D.G. Messerschmitt. *Digital Communication.* Kluwer Academic Publishers, Boston, 3-rd edition, 2003.

[6] S. Benedetto, E. Biglieri. *Principles of Digital Transmission—With Wireless Applications.* Kluwer Academic Press, New York, 1999.

[7] J.M. Wozencraft, I.M. Jacobs. *Principles of Communication Engineering.* John Wiley & Sons, Inc., New York, 1965.

Special Aspects:

[1] J.B. Anderson, T. Aulin, C.-E. Sundberg. *Digital Phase Modulation.* Plenum Press, New York, 1986.

[2] R. Fischer. *Precoding and Signal Shaping for Digital Transmission.* John Wiley & Sons, Inc., New York, 2002.

Additional Exercises:

[1] K.D. Kammeyer, A. Dekorsy, P. Klenner, M. Petermann. *Übungen zur Nachrichtenübertragung.* Springer Vieweg, Wiesbaden, 2-nd edition, 2020.

[2] G. Kramer. *Lerntutorial für NachrichtenTechnik im world wide web.* Online: https://www.lntwww.de/, 3-rd version, 2021.

1 Introduction

All modern information transmission systems are *digital*.

Reasons:

- Significantly higher *power efficiency* than analog transmission

  For a desired robustness of a signal against noise, much less transmission power is needed!

  For example, audio signal transmission (mono):
  Assuming same quality at receiver output (SNR $\hat{=}$60 dB), same signal attenuation, same signal bandwidth, and same noise power density

  FM: 100 kW transmit power

  PCM + digital transmission without source and channel coding: about 30 W

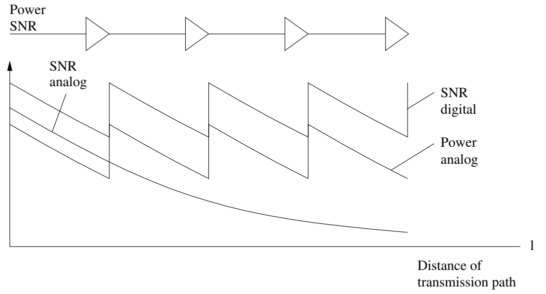
  with source and channel coding: < 1 W

- Continuous exchange between power and bandwidth efficiency possible

  ⇒ For a given power efficiency significantly higher bandwidth efficiency than analog transmission (above transmit power comparison for audio signals assumed equal bandwidth!)

- Signal regeneration

  Analog transmission:
  Intermediate amplifiers (relays) can greatly increase the signal's power but are unable to improve SNR.



Digital transmission:

Signal **regeneration/recovery** is possible: Detection of a binary information sequence and retransmission without noise.

- Arbitrarily long transmission paths are bridgeable without any loss in quality if the distance between the intermediate regenerators (relays) is small enough so that decision errors can be avoided, i.e., if transmission rate doesn't exceed the channel capacity.
- Copies of digitally stored information are exactly as good as the original (e.g. copies of compact discs (CDs) and analog audio tapes).

Examples of Digital Transmission:

Digital Baseband Transmission over Metallic Wire Pairs

- Inside all circuits for digital processing
- Symmetric wires

ISDN Base terminal	$144\,\frac{\mathrm{kbit}}{\mathrm{s}}$	$< 8\,\mathrm{km}$
ISDN Prime rate access	$2.048\,\frac{\mathrm{Mbit}}{\mathrm{s}}$	up to $5\,\mathrm{km}$

HDSL $\hat{=}$ ISDN Prime rate access

ADSL: up to $6\,\frac{\mathrm{Mbit}}{\mathrm{s}}$ in uplink; up to $500\,\frac{\mathrm{kbit}}{\mathrm{s}}$ in downlink and over analog or digital telephone (ISDN); up to a distance of 4 km without intermediate regenerators

ADSL 2+: up to $16\,\frac{\mathrm{Mbit}}{\mathrm{s}}$

VDSL: up to $100\,\frac{\mathrm{Mbit}}{\mathrm{s}}$

Coherent Optical Transmission using Light Waves

- Hero Experiment (April 2010): $64\,\frac{\mathrm{Tbit}}{\mathrm{s}}$ over a $320\,\mathrm{km}$ long optical wave guide ($640 \times 107\,\frac{\mathrm{Gbit}}{\mathrm{s}}$ in polarization and wavelength multiplexing with 36-QAM), spectral efficiency $\Gamma_{\mathrm{d}} = 8\,\frac{\mathrm{bit}}{\mathrm{scdotHz}}$

- Juli 2021: 319 $\frac{\text{Tbit}}{\text{s}}$ over a 3000 km long optical wave guide (16-QAM)

Digital Carrier Modulated Transmission

- Digital communication over switched telephone channel (300 Hz – 3.4 kHz):
  Voice Band Modem: 300 $\frac{\text{bit}}{\text{s}}$ (1960) up to 33.6 $\frac{\text{kbit}}{\text{s}}$ (1996)
  "PCM—Modems": $\leq$ 56 $\frac{\text{kbit}}{\text{s}}$ (1998)
- Digital mobile communication
  GSM, UMTS, LTE
  4G: 500 $\frac{\text{Mbit}}{\text{s}}$, 5G: 20 $\frac{\text{Gbit}}{\text{s}}$
- Digital satellite connections
  $n \times$ 64 $\frac{\text{kbit}}{\text{s}}$ with $n = 1(1)2000$
- Digital radio DAB: 1.5 $\frac{\text{Mbit}}{\text{s}}$
- Digital television DVB: 4 $\frac{\text{Mbit}}{\text{s}}$ up to 45 $\frac{\text{Mbit}}{\text{s}}$
- Digital directed radio: 140 $\frac{\text{Mbit}}{\text{s}}$ up to 565 $\frac{\text{Mbit}}{\text{s}}$
- Coherent optical transmission over air or wave guides
  (required received power: only ca. 4 $\frac{\text{photons}}{\text{bit}}$!!)
  Extremely high bandwith $\Rightarrow$ almost unlimited capacity available!

Digital Information Storage

- Magnetic recording: tape or disc
- Optical recording: CD, DVD, BlueRay

Driving Forces of the Rapid Development:

- Voice band modems: Due to the low data rates and carrier frequencies, highly complex procedures were realizeable (signal processors); highest power and bandwidth efficiency!
- Spacecraft: Highest power efficiency; money is not an issue!
- Mobile communications
- Internet
- "Ubiquitous Wireless Communications"

8

Key Points of the Historic Development:

1842	Morse telegraph
1928	Nyquist: Foundations and theoretical basis of digital communications (Karl Küpfmüller 1924!)
1948	Shannon: Information Theory
1965	Wozencraft/Jacobs: Principles of Communication Engineering
1971	Bingham: OFDM
1971	Viterbi: Convolutional codes
1972	Forney: Intersymbol Interference
1976	Ungerböck, Imai: Coded Modulation, Multi–level Codes
1993	Berrou, Glavieux, Thitimajshima: Turbo–Codes
1993	Wachsmann, Huber, Fischer: Optimal Coded Modulation
1996	McKay: LDPC–Codes
1996	Foschini: BLAST (MIMO)

2000 Ahlswede: Network Coding

2003 Laneman, Wornell: Wireless Relaying

2010 mmWave-Communication Systems (5G)

2015 THz-Communication (potenial in 6G), Drone based Systems (5G)

2020 Intelligent Reflecting Surfaces (IRS) (potential in 6G)

2 Basic Concepts and Terminology

Tasks of the **transmitter** (TX) in a digital communications system:

Mapping of an abstract *binary* sequence of source *symbols* (q_ν), $\nu \in \mathbb{Z}$, $q_\nu \in \{0,\ 1\}$, into a physical transmit *signal* $s(t)$ in such a way that

- The signal $s(t)$ does not exceed a predetermined bandwidth (**bandwidth efficiency**)
- The transmission – even in instances of low received power – is robust against noise (**power efficiency**)

Note: Without loss of generality (w.l.o.g), the digital information to be transmitted may be represented by means of a *binary* sequence (Information Theory: Source Coding Theorem)

Tasks of the **receiver** (RX) in a digital communication system:
Extraction of the source symbol sequence from the noisy and distorted received signal in such a way that:

- as much information as possible about the sequence of source symbols is extracted from the received signal
- the probability that an estimated binary symbol $\hat{q}_\nu$ is erroneously classified is as small as possible:

$$\min \overline{\mathrm{Pr}\left(\hat{q}_\nu \neq q_\nu\right)} = \min \text{BER}$$

- information about the reliability of an estimated symbol can be extracted (i.e. along with $\hat{q}_\nu$ an estimated value for $\mathrm{Pr}\left(\hat{q}_\nu = q_\nu\right)$ is obtained)

Without loss of generality, a memoryless binary source $Q \in \{0,1\}$ is assumed with a priori probabilities
$\Pr(q[l] = 0) = \Pr(q[l] = 1) = \frac{1}{2}, \quad \forall\, l \in \mathbb{Z}$, i.e., with entropy $H(Q) = 1$.

Quality Criteria of a Digital Communication Scheme:

- **Power Efficiency** (AWGN–Channel): Required signal to noise ratio

$$10\log(E_\mathrm{b}/N_0)$$

  to achieve a maximum tolerated Bit Error Probability

$$\text{BER} \quad (\underline{\text{B}}\text{it } \underline{\text{E}}\text{rror } \underline{\text{R}}\text{atio}) \quad \text{e.g.} \quad \text{BER} \leq 10^{-8}$$

  E_b: (equivalent) received signal energy per (information) bit, N_0: one–sided noise power spectral density

- **Bandwidth Efficiency:** $\Gamma_\mathrm{d} = \dfrac{R_\mathrm{T}}{B_\mathrm{RF}} \quad \left[\dfrac{\mathrm{bit/s}}{\mathrm{Hz}}\right]$, $R_\mathrm{T} = \frac{1}{T_\mathrm{b}}$: data rate to be transmitted in $\frac{\mathrm{bit}}{\mathrm{s}}$, T_b: bit duration, B_RF: (one–sided) bandwidth of the transmit signal

- **Complexity** of TX and RX, i.e., complexity of encoding, modulation, detection, receiver synchronization, channel coding etc.

- **Delay** (latency) of received data stream

16

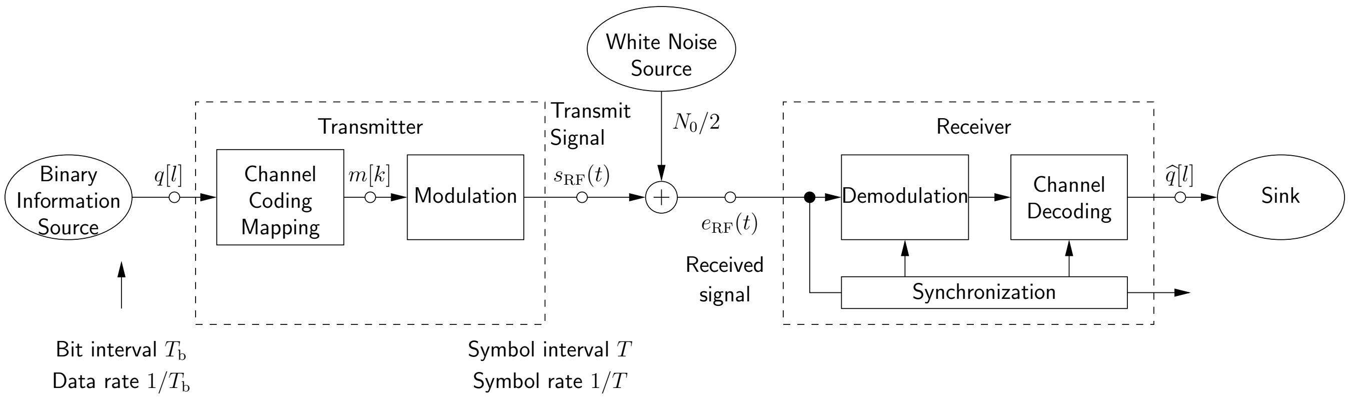
Parameters of a Digital Communication System:

– Transmitted data rate	R_T $\quad[\mathrm{bit/s}]$
– Bit interval	$T_\mathrm{b} = \frac{1}{R_\mathrm{T}}$ $\quad[\mathrm{s}]$
– Symbol interval	T $\quad[\mathrm{s}]$
– Symbol rate (also known as *baud rate*)	$1/T$ $\quad[\mathrm{symbol/s}]$
– Average information content per transmitted symbol, also known as the **rate of the communication scheme**	$R = \frac{T}{T_\mathrm{b}}$ $\quad[\mathrm{bit/symbol}]$
– Average power of the TX output signal	S_s
– Average power of the **desired** part of the RX input signal	S_e
– Average received energy per <u>s</u>ymbol	$E_\mathrm{s} = S_\mathrm{e} \cdot T$
– Average received energy per <u>b</u>it of information	$E_\mathrm{b} = S_\mathrm{e} \cdot T_\mathrm{b} = E_\mathrm{s}/R$
– (one–sided) bandwidth of the transmit signal	B_RF
– Bandwidth efficiency (also known as spectral efficiency)	$\Gamma_\mathrm{d} = \frac{R_\mathrm{T}}{B_\mathrm{RF}}$ $\quad\left[\frac{\mathrm{bit/s}}{\mathrm{Hz}}\right]$

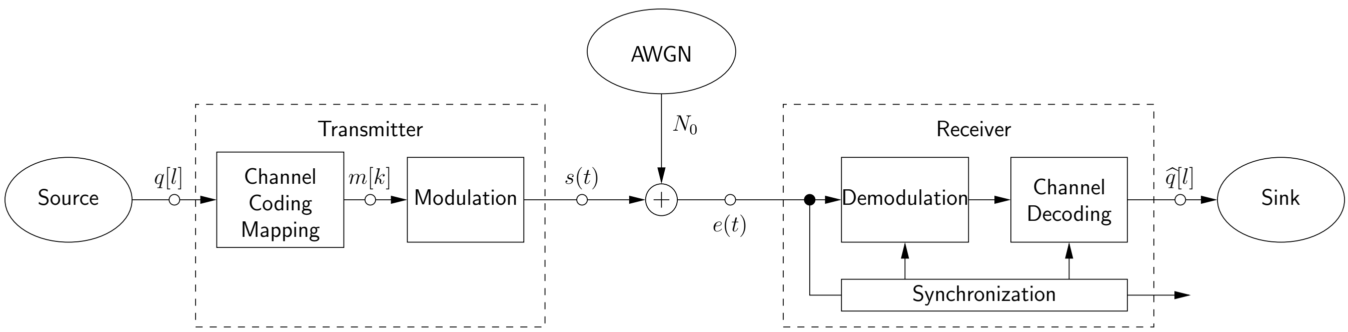
For AWGN Channel Model:

- Impairment by Additive White Gaussian Noise (AWGN) with
  (two–sided) noise power spectral density (physical signal) $N_0/2$
  (one–sided) noise power spectral density
  (Equivalent Complex Baseband <u>ECB</u> signal) N_0
- normalized signal to noise ratio (at RX input) $E_\mathrm{b}/N_0 = S_\mathrm{e} \cdot T_\mathrm{b}/N_0 = \frac{E_\mathrm{s}}{N_0}/R$

Block Diagram of a Digital Communication System (AWGN channel model):

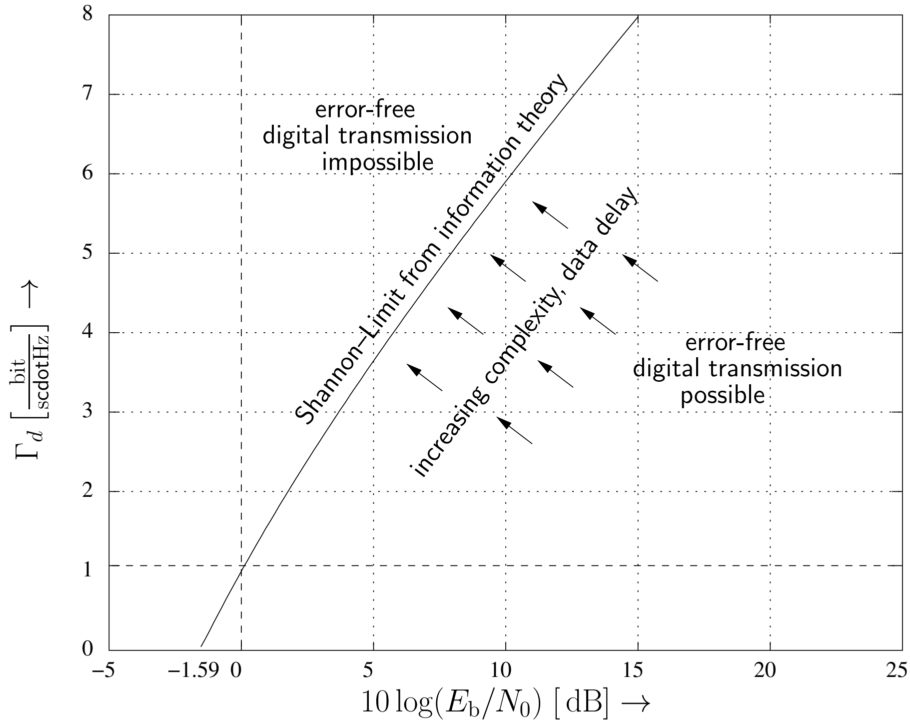


ECB–Signal (AWGN channel model):



Tradeoff between Power and Bandwidth Efficiency:

Power–bandwidth diagram for an AWGN channel



Basis of comparison for different communication schemes!

20

Shannon Limit for Digital Communication: Capacity of a Bandlimited AWGN Channel:

$$\boxed{C_\mathrm{T} \;=\; B_\mathrm{RF}\ \mathrm{ld}(1+\tfrac{S_\mathrm{e}}{N}) \quad \tfrac{\mathrm{bit}}{\mathrm{s}}}$$

B_RF	one–sided RF bandwidth
$S_\mathrm{e} = E_\mathrm{b}/T_\mathrm{b}$	signal power (at receiver)
$N = N_0 \cdot B_\mathrm{RF}$	noise power

For an ideal communication scheme, the data rate R_T is equal to the channel capacity

$$R_\mathrm{T} \;=\; B_\mathrm{RF}\ \mathrm{ld}\left(1+\frac{E_\mathrm{b}}{N_0}\frac{1}{B_\mathrm{RF}T_\mathrm{b}}\right)$$

with bandwidth efficiency $\Gamma_\mathrm{d} = R_\mathrm{T}/B_\mathrm{RF} = 1/(B_\mathrm{RF}\cdot T_\mathrm{b})$:

$$\Gamma_\mathrm{d} \;=\; \mathrm{ld}\left(1+\tfrac{E_\mathrm{b}}{N_0}\Gamma_\mathrm{d}\right) \qquad \boxed{\tfrac{E_\mathrm{b}}{N_0} \;=\; \tfrac{1}{\Gamma_\mathrm{d}}\left(2^{\Gamma_\mathrm{d}}-1\right)}$$

Shannon limit for the trade–off between power and bandwidth efficiency in digital communications

2.1 Introduction to Coding and Modulation

Mapping of a binary source symbol **sequence** to a transmit **signal** $s(t)$

$$q[l] \to s(t) \qquad l \in \mathbb{Z},\, t \in \mathbb{R}$$

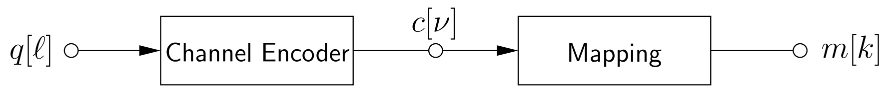
can be subdivided into

Channel Coding with Mapping

and

Modulation

2.1.1 Channel Coding with Mapping



$c[\nu], \quad \nu \in \mathbb{Z};$ code symbol sequence; $\mathcal{C}$: code symbol alphabet

$m[k], \quad k \in \mathbb{Z}, \quad m[k] \in \{0, \ldots, M-1\};$ sequence of signal numbers

2.1.1.1 Principles of Channel Coding

Channel coding: Reversible mapping of words of length k consisting of **binary** source symbols to code words of length n consisting of symbols from an M_c–ary *code symbol alphabet* $\mathcal{C}$

$$\vec{Q} = (Q_0 \quad Q_1 \quad Q_2 \quad \cdots \quad Q_{k-1}) \quad \longmapsto \quad (C_0 \quad C_1 \quad C_2 \quad \cdots \quad C_{n-1}) = \vec{C}$$

$$Q_i = q[i + \mu k] \in \{0,1\}, \quad \mu \in \mathbb{Z} \qquad\qquad C_j \in \mathcal{C}, \quad C_j = c[\,j + \mu n\,]$$

$$2^k \le M_\mathrm{c}^n$$

For $2^k < M_\mathrm{c}^n$ **redundancy** is introduced by the channel coding, primarily in the form of *check sums*.

Definition: *Parameters of a channel code*

- Code word length: n
- Code symbol alphabet: $\mathcal{C} = \{c_1, c_2, \ldots, c_{M_\mathrm{c}}\};\ |\mathcal{C}| = M_\mathrm{c}$
- Code rate (average information content per code symbol) $R_\mathrm{c} = \frac{1}{n}\,\mathrm{ld}\left(2^k\right) = \frac{k}{n} \quad \left[\frac{\text{bit}}{\text{code symbol}}\right]$
  (equally probable source and/or code words!)
- Code redundancy per code symbol $\rho_\mathrm{c} = \frac{1}{n}\left(\mathrm{ld}(M_\mathrm{c}^n) - \mathrm{ld}(2^k)\right) = \mathrm{ld}(M_\mathrm{c}) - R_\mathrm{c} \quad \left[\frac{\text{bit}}{\text{code symbol}}\right]$

If the encoding rule of source words into code words is time invariant and independent of preceding source words, then the encoding rule specifies a block code, otherwise the code is usually a **trellis code**.

Example: *Binary* $(n,\, k = n-1)$ *Parity Check Code (block code)*

$M_{\mathrm{c}} = 2$; n; $k = n-1$; $R_{\mathrm{c}} = \frac{n-1}{n}$, $\rho_{\mathrm{c}} = \frac{1}{n}$;

Extension to $\left\{ \begin{matrix} \text{even} \\ \text{odd} \end{matrix} \right\}$ number of symbols 1: $\left\{ \begin{matrix} \text{parity even} \\ \text{parity odd} \end{matrix} \right.$

Q_0	Q_1	$\cdots$	Q_{n-2}

$\longmapsto$

Q_0	Q_1	$\cdots$	Q_{n-2}	P

Parity even: $P = Q_0 \oplus Q_1 \oplus \ldots \oplus Q_{n-2} = \sum\!\!\!\!\!\bigoplus_{i=0}^{n-2} Q_i$ with

$\oplus$ and $\sum\!\!\!\!\!\bigoplus$: addition mod 2, i.e., $1 \oplus 1 = 0$!

Parity odd: $P = \sum\!\!\!\!\!\bigoplus_{i=0}^{n-2} Q_i \oplus 1$

Example: *Binary* $(15{,}11)$ *Hamming–Code (block code)*

$M_{\mathrm{c}} = 2$; $n = 15$; $k = 11$; $R_{\mathrm{c}} = \frac{11}{15}$, $\rho_{\mathrm{c}} = \frac{4}{15}$; $Q_i \in \{0{,}1\}$; $P_i \in \{0{,}1\}$

Q_0	Q_1	$\cdots$	Q_{10}

$\longmapsto$

Q_0	Q_1	$\cdots$	Q_{10}	P_1	P_2	P_3	P_4

with the checksums:

$$\begin{aligned}
Q_0 \oplus Q_1 \oplus Q_3 \oplus Q_4 \oplus Q_6 \oplus Q_8 \oplus Q_{10} \oplus P_1 &= 0 \\
Q_0 \oplus Q_2 \oplus Q_3 \oplus Q_5 \oplus Q_6 \oplus Q_9 \oplus Q_{10} \oplus P_2 &= 0 \\
Q_1 \oplus Q_2 \oplus Q_3 \oplus Q_7 \oplus Q_8 \oplus Q_9 \oplus Q_{10} \oplus P_3 &= 0 \\
Q_4 \oplus Q_5 \oplus Q_6 \oplus Q_7 \oplus Q_8 \oplus Q_9 \oplus Q_{10} \oplus P_4 &= 0
\end{aligned}$$

P_i: Parity symbols

Definition:

For a coding scheme for which k source symbols are directly mapped to k binary code symbols, the encoding rule is referred to as **systematic encoding**. (Note: Only for $M_{\mathrm{c}} = 2$ a systematic encoding exists)

Note:
For $(2^l-1, k=2^l-l-1)$ Hamming–Codes with $l=3,4,5,\ldots$, one error per code word can be corrected.

Example: *Quaternary code symbols*

$M_{\mathrm{c}}=4$; $n=4$; $k=8$; $C_i \in \{0{,}1{,}2{,}3\}$ quaternary code symbols

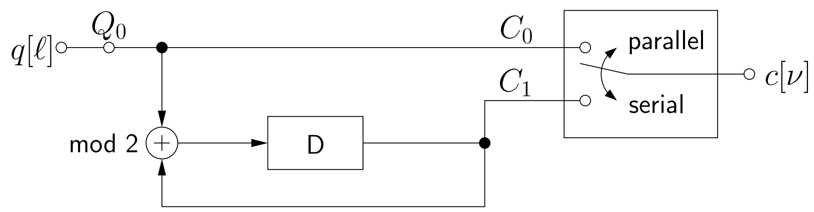
$$\boxed{Q_0 \mid Q_1 \mid \cdots \mid Q_7} \longmapsto \boxed{C_0 \mid C_1 \mid C_2 \mid C_3}$$

$$R_{\mathrm{c}} = 2\,\frac{\text{bit}}{\text{code symbol}}; \qquad \rho_{\mathrm{c}} = 2-2=0 \quad !$$

Redundancy–free recoding from binary to quaternary symbol alphabet.
(Such an encoding is actually a mapping, see next subsection)

Example: *Trellis Code*

$M_{\mathrm{c}}=2$; $n=2$; $k=1$; $R_{\mathrm{c}}=\frac{1}{2}$



$$\begin{array}{llllllll} q[l] & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ c[\nu] & 0\,\underline{0} & 1\,\underline{0} & 1\,\underline{1} & 0\,\underline{0} & 1\,\underline{0} & 0\,\underline{1} & 1\,\underline{1} & 1\,\underline{0} \\ & C_0\; C_1 & & & & & & & \end{array}$$

The mapping $Q_0 \to (C_0, C_1)$ is dependent on the **entire** sequence of past source symbols. The past values are represented by the current **state** of the delay element D.

2.1.1.2 Mapping

Definition: *Mapping*

> **Redundancy free encoding** of the code symbol sequence $c[\nu]$ with $\nu \in \mathbb{Z}$, $c[\nu] \in \mathcal{C}$, $|\mathcal{C}| = M_{\mathrm{c}}$ into a sequence of M–ary **signal numbers**, $\{0,1,\ldots, M-1\}$, (transform):
>
> $$c[\nu] \quad \rightarrow \quad m[k], \qquad \nu, k \in \mathbb{Z}, \qquad m[k] \in \{0,1,\ldots,\ (M-1)\}$$
>
> $m[k]$: Signal number in the k^{th} modulation interval
>
> M: Number of levels in the modulation scheme.

Collection of L code symbols (or L binary source symbols if no channel coding is used, i.e., $\rho_{\mathrm{c}} = 0$) into blocks which are mapped into a corresponding vector of V signal numbers:

$$(C_0, \ldots,\ C_{L-1}) \quad \mapsto \quad (\mathcal{M}_0, \ldots,\ \mathcal{M}_{V-1}) \quad \text{with} \quad M_{\mathrm{c}}^L \quad \stackrel{!}{=} \quad M^V$$

Note: If a particular pair $L,V \in \mathbb{N}$ does not exist for which this condition is fulfilled, an L and V are chosen such that $M^V > M_{\mathrm{c}}^L$, with the smallest possible difference, such that the *mapping redundancy* is minimized, $\rho_M = \frac{1}{V}\mathrm{ld}\left(\frac{M^V}{M_{\mathbf{c}}^L}\right) = \mathrm{ld}(M) - \frac{L}{V}\,\mathrm{ld}(M_{\mathrm{c}})$.

28

Example:

$M_{\mathrm{c}} = 2$; $M = 8$; $L = 3$; $V = 1$

$$\boxed{C_1}\boxed{C_2}\boxed{C_3} \quad \longmapsto \quad \boxed{M_1}$$

Representation of 3 binary symbols by one 8-ary symbol

Example:

$M_{\mathrm{c}} = 2$; $M = 3$; Representation of binary symbols by means of ternary symbols

Special case: $L = 11$; $V = 7$

$$2^{11} = 2048 \qquad 3^7 = 2187$$

$$\rho_{\mathrm{M}} = \mathrm{ld}(3) - \frac{11}{7} = 1.5850 - 1.5715 = 0.0135\,\frac{\mathrm{bit}}{\mathrm{ternary\ symbol}}$$

L	2^L	V	3^V	ρ_{M}
1	2	1	3	0.585
2	4	2	9	0.585
4	16	3	27	0.252

L	2^L	V	3^V	ρ_{M}
5	32	4	81	0.335
6	64	4	81	0.085
19	524288	12	531441	0.00163

2.1.2 Modulation

Definition: *Modulation*

Representation of a sequence of signal numbers $m[k]$, $k \in \mathbb{Z}$, by means of a continuous–time analog transmit signal $s_{\mathrm{RF}}(t)$, or alternatively ECB signal $s(t)$, for which

- each signal number $m[k]$ is mapped to one **signal pulse** per **modulation interval** T independent of preceding and following signal numbers

  $$s_{m[k]}(t-kT)$$

  taken from a **signal set** $\mathcal{S} = \{s_0(t), s_1(t), \ldots,\ s_{M-1}(t)\}$ with $|\mathcal{S}| = M$.

  The elements of the signal set (the signal pulses) are referred to as the **signal elements** of the digital communication scheme.

- The transmit signal consists of the **additive superposition** of these signal elements:

  $$s(t) = \sum_{k=-\infty}^{\infty} s_{m[k]}(t-kT) \qquad \text{ECB signal}$$

  $$s_{\mathrm{RF}}(t) = \sum_{k=-\infty}^{\infty} s_{\mathrm{RF},m[k]}(t-kT) \qquad \text{physical signal}$$

- The signal elements are **orthogonal** with respect to shifts in time which are integer multiples of the symbol interval T, see Section 2.1.2.1.

Parameters of a modulation scheme:

- Transmitted information flow ("data rate") $\quad R_{\mathrm{T}} = \frac{1}{T_{\mathrm{b}}}$
- Modulation interval, duration of modulation step $\quad T$
- Number of modulation levels $\quad M$
- Average signal energy per symbol (impulse) $\quad E_{\mathrm{s}} = \sum_{m=0}^{M-1} E_m \cdot \Pr(\mathcal{M} = m)$
  - for equally probable signal elements $\quad E_{\mathrm{s}} = \frac{1}{M}\sum_{m=0}^{M-1} E_m$
  - with the energy of the signal elements $s_m(t)$ $\quad E_m = \int_{-\infty}^{+\infty} |s_m(t)|^2 \,\mathrm{d}t$
- Average signal power (orthogonality!) $\quad S_{\mathrm{e}} = E_{\mathrm{s}}/T$
- Rate (of the modulation scheme):
  - Mapping of L code symbols, each with R_{c} bits of information, onto V signal elements $\quad R = R_{\mathrm{c}}\,\frac{L}{V}\ \left[\frac{\mathrm{bit}}{\mathrm{modulation\ step}}\right]$
  - Continuity condition: $\frac{1\,\mathrm{bit}}{T_{\mathrm{b}}} = \frac{R\ \mathrm{bit}}{T}$ $\quad R = \frac{T}{T_{\mathrm{b}}} \leq \mathrm{ld}(M)$
  - For redundancy–free transmission $\quad R = \mathrm{ld}(M)$
- Symbol speed (symbol rate or baud rate) $\quad \frac{1}{T} = \frac{1}{T_{\mathrm{b}}}\frac{1}{R} = \frac{1}{T_{\mathrm{b}}}\frac{1}{R_{\mathrm{c}}}\frac{V}{L}$
- Average signal energy per bit information (AWGN channel with compensation of attenuation!) $\quad E_{\mathrm{b}} = S_{\mathrm{e}} \cdot T_{\mathrm{b}} = E_{\mathrm{s}}/R$
- Random variable for signal number m $\quad \mathcal{M}$

Notes on terminology

Definition (Rate) Average information (with respect to information theory!) measured in bits per

- Symbol
- Modulation step, modulation interval
- Time

Examples:

1. Source: Rate $R_\mathrm{Q} \;=\; H(\mathrm{Q}) \quad \mathrm{bit/(source\,symbol)}$
2. Code: Rate $R_\mathrm{c} \;=\; k/n \quad \mathrm{bit/(code\,symbol)}$

Definition (Rate of a Modulation Scheme, R_m)

$$R_m \;=\; H(\mathcal{M}) \quad \text{with} \quad H(\mathcal{M}) \;=\; -\sum_{m=0}^{M-.1} \mathrm{Pr}(\mathcal{M}=m)\log_2\left(\mathrm{Pr}(\mathcal{M}=m)\right)$$

Average information per modulation step (i.e. per single signal element) in the case of redundancy–free mapping of redundancy–free source symbols onto the signal elements, i.e. , no channel coding, no mapping redundancy.

32

Note:

$R_m \leq \log_2(M)$ with

$R_m = \log_2(M)$ for equally probable signal elements (predominantly assumed here!)

Definition (Signal Shaping):

Procedure for creating a non–uniform distribution for the signal elements (e.g. Gaussian distributed amplitude coefficients in the case of ASK and QAM in order to come as close as possible to the Shannon limit)

Shaping redundancy $\rho_\mathrm{S} \;=\; \log_2(M) - H(\mathcal{M})$

Rate of a Digital Communication Scheme

$$\boxed{R = R_\mathrm{c}\cdot R_m \;=\; R_\mathrm{T}\cdot T \quad \frac{\mathrm{bit}}{\mathrm{modulation\ step}}}$$

Average information transmitted per modulation step, **including** channel coding, mapping, and signal shaping.

Remarks on E_b

Equivalent **received** energy per bit of transmitted information

$$E_\mathrm{b} = S_\mathrm{e} \cdot T_\mathrm{b}$$

E_b refers to the received signal $\tilde{e}(t)$ with average power S_e! ($S = S_\mathrm{s} = S_\mathrm{e}$ is valid only in the case of AWGN channels with compensation for signal attenuation by signal amplification at the receiver input)

$$E_\mathrm{b} = E_\mathrm{s}/R \quad \text{with} \quad E_\mathrm{s} = S_\mathrm{e} \cdot T$$

E_s : Signal energy per modulation step (per signal element)

Distortion–free channel with amplification by a factor of D (or alternatively, attenuation by $1/D$) and signal delay t_0:

$$\begin{aligned} e(t) &= \tilde{e}(t) + n(t) \quad \text{with} \quad \tilde{e}(t) = D \cdot s(t - t_0) \\ S_\mathrm{e} &= D^2 \cdot S_\mathrm{s} \\ E_\mathrm{b} &= S_\mathrm{s} \cdot D^2 \cdot T_\mathrm{b} \quad \text{and} \quad E_\mathrm{s} = S_\mathrm{s} \cdot D^2 \cdot T \end{aligned}$$

E_b depends on the data rate (transmitted information flow) R_T, the average transmit power S_s, and the signal attenuation. It is independent of the particular channel coding method as well as the modulation employed to map the source symbol sequence onto the transmit signal $s(t)$. This characteristic makes the

Average energy per bit to noise power spectral density ratio: E_b/N_0

especially well suited as a representation of the power efficiency of a digital communication scheme in the AWGN channel and for comparison of the power efficiency of different digital communication schemes; e.g. required min E_b/N_0 for BER$\leq 10^{-8}$.

2.1.2.1 Condition of Temporal Orthogonality

For the M elements $s_i(t)$ with $i \in \{0, \ldots, M-1\}$ of the signal set $\mathcal{S}$, i.e., the **signal elements**, a clear separation between coding and modulation is achieved by means of the following condition:

$$\varphi_{s_i,s_l}(kT) = \int_{-\infty}^{+\infty} s_i(t) \cdot s_l^*(t-kT)\,\mathrm{d}t \overset{!}{=} \left\{ \begin{array}{ccc} E_{il} & \text{for} & k=0 \\ 0 & \text{for} & k \neq 0 \end{array} \right\} = E_{il} \cdot \delta[k]$$

with E_{il}: cross energy between signal elements $s_i(t)$ and $s_l(t)$

specifically $E_{ii} \overset{\text{def}}{=} E_i$: energy of a signal element $s_i(t)$

This property is known as **temporal orthogonality** of the signal elements.

In particular, the temporal orthogonality of a particular signal element also applies to itself

$$\int_{-\infty}^{+\infty} s_i(t) \cdot s_i^*(t-kT)\,\mathrm{d}t = E_i \cdot \delta[k], \qquad i \in \{0, \ldots, M-1\}.$$

Motivation for temporal orthogonality:

We will later see that for impairment by additive white Gaussian noise, demodulation and detection for each symbol interval (i.e., for each pulse) can be carried out **independently**. This means that with respect to the represented symbol sequence (sequence of signal numbers) lossless discrete–time signal processing is possible using only **one sample** per symbol interval, even if (as is commonly the case) the sampling theorem is not fulfilled.

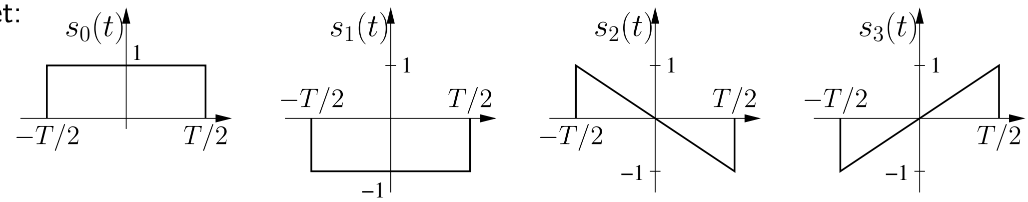
36

2.1.2.1.1 Time Limited Signal Elements

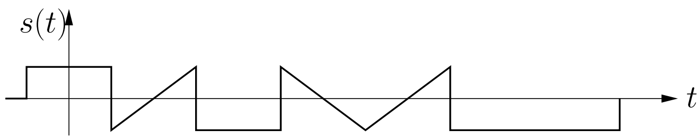
Signal elements which are time limited to one symbol interval fulfill the temporal orthogonality condition.
Reason: For the cross–correlation, $\varphi_{s_i,s_l}(\tau) = 0$ holds for $|\tau| \geq T$, i.e., $\varphi_{s_i,s_l}(kT) \equiv 0 \quad \forall k \in \mathbb{Z} \setminus \{0\}$

Example: *Quaternary Modulation Scheme*

Signal set:

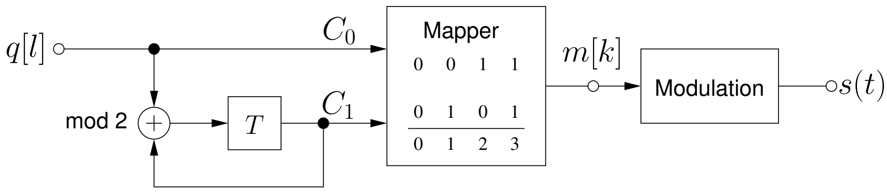


A. No channel coding (natural mapping)

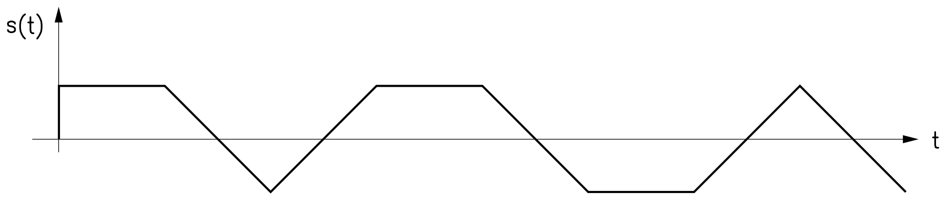


Rate: $R = 2\,\frac{\text{bit}}{\text{symbol}}$

B. With channel coding (compare with previous example for trellis code)



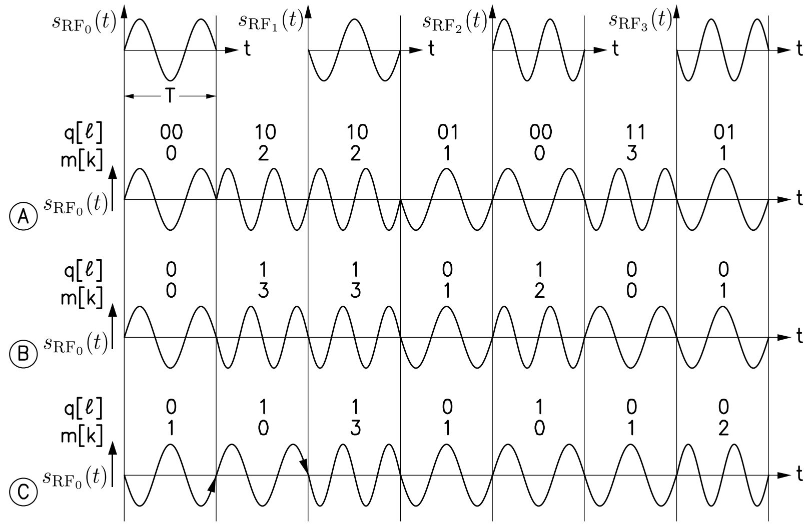
q[ℓ]:	0	1	1	0	1	0	1	1
c[ν]:	0 0	1 0	1 1	0 0	1 0	0 1	1 1	1 0
m[k]:	0	2	3	0	2	1	3	2



In this case, the channel coding has two purposes; increasing the bandwidth efficiency (i.e., discontinuities are avoided) and power efficiency. This comes at the cost of halving the rate:

$$R = 1\,\frac{\text{bit}}{\text{symbol}}$$

Example: *Digital Frequency– and Phase Modulation*



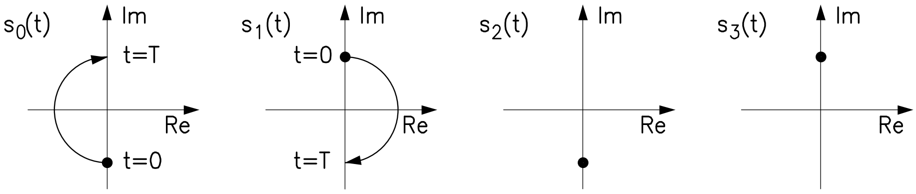
A No channel coding $R = 2\,\frac{\text{bit}}{\text{symbol}}$

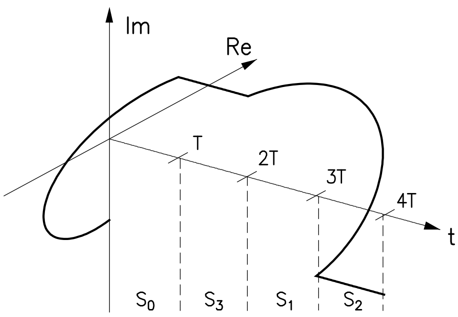
B (Recursive) channel coder $R = 1\,\frac{\text{bit}}{\text{symbol}}$
Continuous phase, the source symbols are represented by the **instantaneous frequency**

C (Nonrecursive) channel coder $R = 1\,\frac{\text{bit}}{\text{symbol}}$

Continuous phase, the source symbols are represented by the **phase** at the end of each symbol interval

Depiction of the above signal elements using ECB signals.
(Note: Here, the transformation frequency $f_0 = \frac{2}{T}$($\neq$ center frequency $\frac{1.75}{T}$ = carrier frequency) is used in order to have the minimum number of signal elements possible)



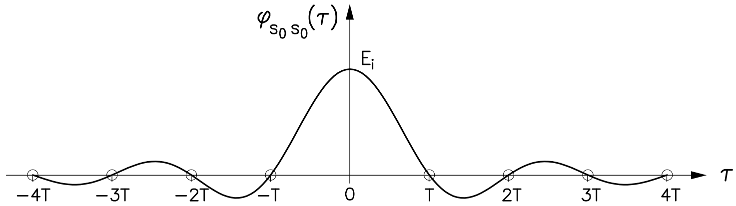


2.1.2.1.2 Bandlimited Signal Elements

Signal elements, whose autocorrelations, or alternatively crosscorrelations, contain zeros at multiples of the modulation interval fulfill the temporal orthogonality condition:

$$\varphi_{s_i,s_l}(kT) = \int_{-\infty}^{+\infty} s_i(t+kT)s_l^*(t)\,\mathrm{d}t = \begin{cases} 0, & \forall\, k \in \mathbb{Z} \setminus \{0\} \\ E_{il}, & k = 0 \end{cases}$$

e.g.

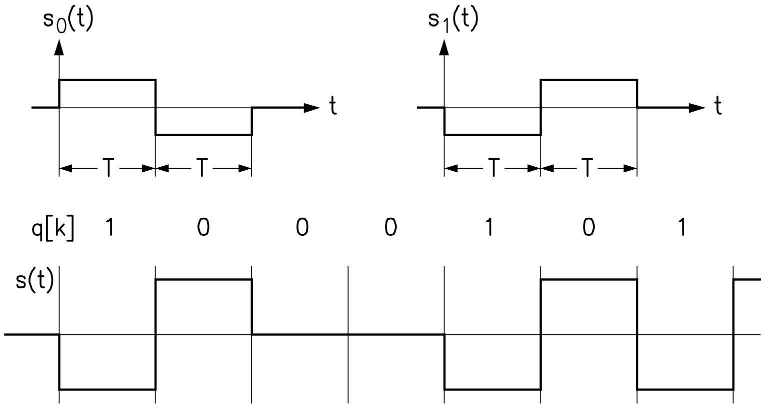


Bandlimited and therefore non–time–limited signal pulses can also fulfill the orthogonality condition!

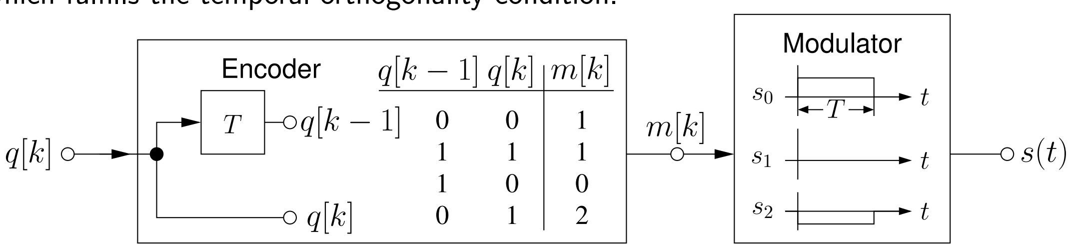
2.1.2.1.3 Equivalent Coding and Modulation

If the signal elements at the **receiver side** do **not** fulfill the temporal orthogonality condition due to signal distortions, caused by e.g. whitening filters in the case of colored noise or pulse shaping at the transmitter side, the receiver input signal can be represented in terms of an **equivalent encoding** and a **new signal set** for which the orthogonality condition is fulfilled. At the receiver side, demodulation and **decoding** of the equivalent encoding is carried out preferably with respect to this new signal set.

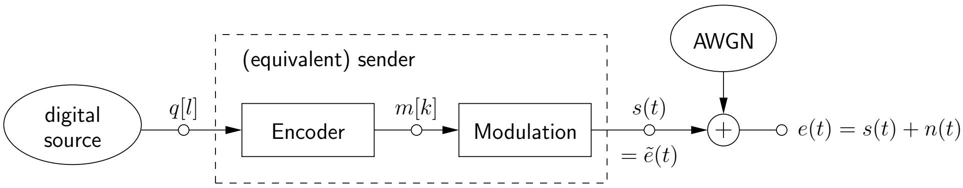
Example: *Equivalent Encoding*



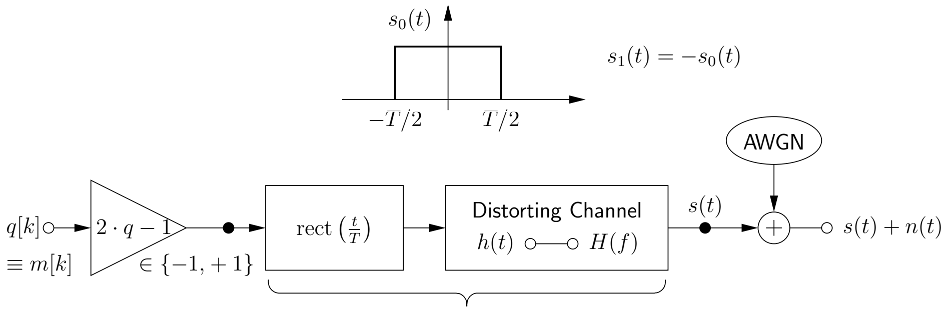
Division of the signal mapping in the above diagram into encoding and modulation with a new signal set, which fulfills the temporal orthogonality condition.



For the receiver input signal, it is always possible to find a model of signal representation in terms of an equivalent encoding, a signal set satisfying the temporal orthogonality condition, and AWGN!
Thus, the theory of optimum detection of encoded sequences of orthogonal signal elements in AWGN is sufficient in (almost) all cases!



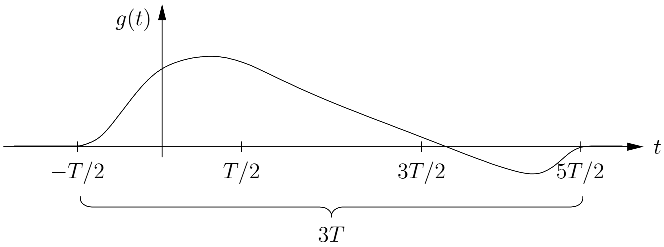
Example: *Distortion of a bipolar binary transmission with rectangular impulses*



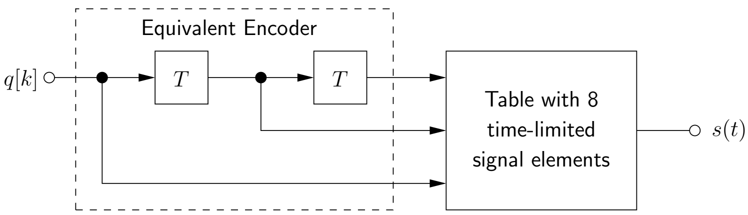
$$g(t) \;=\; \mathrm{rect}(\frac{t}{T}) \star h(t)$$

44

$$G(f) \;=\; T\mathrm{si}(\pi f T) \cdot H(f)$$



$\equiv$



3 Digital Pulse Amplitude Modulation

If all M signal elements of a digital modulation scheme are multiples of a single **fundamental pulse** $g(t)$, i.e., if

$$s_i(t) = a_i\, g(t), \qquad i = 0,1,\ldots,M-1,$$

holds, the communications scheme is referred to as **Digital Pulse Amplitude Modulation (PAM)**.

Characteristics of digital PAM:

1. The sequence of binary digital source symbols is mapped into a sequence of **amplitude coefficients** taken from a set $\boldsymbol{\mathcal{A}}$ with M elements.

$$\left.\begin{array}{l}\boldsymbol{\mathcal{A}}\text{: Signal constellation}\\ M = |\boldsymbol{\mathcal{A}}|\text{: Number of amplitude levels}\end{array}\right\}\text{ of the PAM–scheme}$$

2. The transmit signal consists of equidistant pulses $g(t-kT)$ which are weighted by amplitude coefficients $a_{m[k]}$:

$$s(t) = \sum_{k=-\infty}^{+\infty} a_{m[k]}\, g(t-kT), \qquad a_{m[k]} \in \boldsymbol{\mathcal{A}}, \qquad k \in \mathbb{Z}$$

Definition: $g(t) =$ *fundamental pulse of PAM*

Energy of the fundamental pulse: $E_\mathrm{g} = \int_{-\infty}^{+\infty} |g(t)|^2\,\mathrm{d}t = \int_{-\infty}^{+\infty} |G(f)|^2\,\mathrm{d}f$

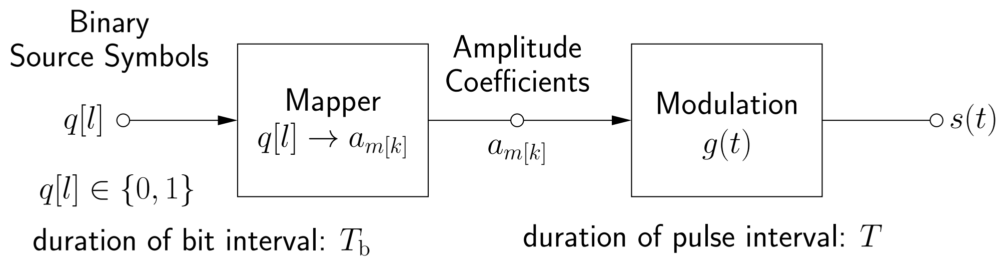
with fundamental pulse spectrum $G(f) = \int_{-\infty}^{+\infty} g(t)\mathrm{e}^{-\mathrm{j}2\pi f t}\,\mathrm{d}t$

T: Modulation interval or alternatively, pulse interval;

$m[k]$: Signal number at discrete time k

PAM with no channel coding:

Redundancy free mapping of a binary sequence of source symbol sequence $q[l]$; $l \in \mathbb{Z}$; to a sequence $a_{m[k]}$, $k \in \mathbb{Z}$, of amplitude coefficients at a data rate: $R_\mathrm{T}\ \left[\frac{\mathrm{bit}}{\mathrm{s}}\right]$; rate: $R = \mathrm{ld}(M) = R_\mathrm{T} \cdot T$



Example:

$M = 4$–ary PAM scheme with $\boldsymbol{\mathcal{A}} = \{-3, -1, +1, +3\}$

Each pair of source symbols is mapped to one amplitude coefficient:

$q[l]$:	0 1	1 0	1 1	0 0	0 0	0 1	1 0	0 1	1 1	0 1	0 ...
$a_{m[k]}$:	–1	+1	+3	–3	–3	–1	+1	–1	+3	–1	

symbol separation: $T = 2T_\mathrm{b}$

3.1 Average Power Spectral Density of a PAM Signal

Supplement 1: *Average Power Spectral Density (A-PSD)*

Given:

- Discrete time random process (wide sense stationary) $a[k]$ with autocorrelation sequence (ACS)

$$\phi_{aa}[\kappa] = \mathrm{E}\left\{a[k+\kappa]a^*[k]\right\}; \Phi_{aa}(\mathrm{e}^{j2\pi F}) = \sum_{\kappa=-\infty}^{+\infty} \phi_{aa}[\kappa] \cdot \mathrm{e}^{-j2\pi\kappa F} \quad : \text{PSD w.r.t. DTFT}$$

- Generation of a continuous time PAM signal by means of convolution of discrete time sequence $a[k]$ and continuous time pulse $g(t)$:

$$s(t) = \sum_{k=-\infty}^{+\infty} a[k]g(t-kT)$$

Results:

- PAM signal $s(t)$ is a cyclo–stationary (also termed "periodic stationary") random process, i.e., its stochastic parameters vary in time in a periodic manner with period T.

- Averaging of parameters over one period T:
  Average power, average ACF, and average power spectral density (A–PSD) i.e., forcing the cylco–stationary process to be wide sense stationary

- Model: Phase Randomization: Artificial extension of the set of sample function by all possible time shifts

$$s(\nu,t) = s(t-\nu T),$$

where ν is uniformally distributed in [0,1).

Average Autocorrelation Function (ACF) of a PAM signal $s(t)$

Averaging over one period T of the cylco–stationary random process

$$\begin{aligned}
\bar{\phi}_{ss}(\tau) &= \frac{1}{T}\int_0^T \mathrm{E}\left\{s(t+\tau)s^*(t)\right\}\mathrm{d}t \\
&= \frac{1}{T}\int_0^T \sum_k \sum_\ell \mathrm{E}\left\{a[k]a^*[\ell]g(t+\tau-kT)g^*(t-\ell T)\right\}\mathrm{d}t \\
&\overset{\kappa=k-\ell}{=} \frac{1}{T}\int_0^T \sum_k \sum_\kappa \mathrm{E}\left\{a[k]a^*[k-\kappa]\right\} g(t+\tau-kT)g^*(t-(k-\kappa)T)\mathrm{d}t
\end{aligned}$$

Using $\phi_{aa}[\kappa] = \mathrm{E}\{a[k+\kappa]a^*[k]\}$ leads to

$$\bar{\phi}_{ss}(\tau) \quad = \quad \frac{1}{T}\sum_{\kappa}\phi_{aa}[\kappa]\sum_{k}\int\limits_{0}^{T} g(t+\tau-kT)g^*(t-(k-\kappa)T)\mathrm{d}t$$

Using $\sum\limits_{k}\int_0^T f(t-kT)\mathrm{d}t = \int_{-\infty}^{+\infty} f(t)\mathrm{d}t$ yields

$$\bar{\phi}_{ss}(\tau) \quad = \quad \frac{1}{T}\sum_{\kappa}\phi_{aa}[\kappa]\int\limits_{-\infty}^{+\infty} g(t+\tau)g^*(t+\kappa T)\mathrm{d}t$$

Autocorrelation of the fundamental pulse $\varphi_{gg} = g(t) * g^*(-t) \stackrel{\text{def}}{=} \int\limits_{-\infty}^{+\infty} g(t+\tau)g^*(t)\mathrm{d}t$

$$\boxed{\bar{\phi}_{ss}(\tau) = \frac{1}{T}\sum_{\kappa}\phi_{aa}[\kappa]\varphi_{gg}(\tau-\kappa T)}$$

Average Power Spectral Density (A–PSD)

$$\begin{aligned}\bar{\Phi}_{ss}(f) &= \mathcal{F}\left\{\bar{\phi}_{ss}(\tau)\right\} \\ &= \mathcal{F}\left\{\frac{1}{T}\sum_{\kappa}\phi_{aa}[\kappa]\varphi_{gg}(\tau-\kappa T)\right\}\end{aligned}$$

52

$$\begin{aligned} &= \frac{1}{T}\sum_{\kappa}\phi_{aa}[\kappa]\mathcal{F}\left\{\varphi_{gg}(\tau-\kappa T)\right\} \\ &= \frac{1}{T}\sum_{\kappa}\phi_{aa}[\kappa]\mathcal{F}\left\{\varphi_{gg}(\tau)\right\}\cdot \mathrm{e}^{-j2\pi f\kappa T}\end{aligned}$$

Using $\varphi_{gg}(\tau) = \int\limits_{-\infty}^{\infty} g(t)g^*(t-\tau)\mathrm{d}t \circ\!\!-\!\!\bullet\; |G(f)|^2$ (Energy Spectral Density of $g(t)$) yields:

$$\bar{\Phi}_{ss}(f) \;=\; \frac{1}{T}\sum_{\kappa}\phi_{aa}[\kappa]|G(f)|^2\cdot \mathrm{e}^{-j2\pi f\kappa T}$$

With $\Phi_{aa}(z) = \sum\limits_{\kappa=-\infty}^{\infty}\phi_{aa}[\kappa]z^{-\kappa}$ we obtain for the A-PSD

$$\boxed{\bar{\Phi}_{ss}(f) = \frac{1}{T}\Phi_{aa}(\mathrm{e}^{j2\pi fT})|G(f)|^2}$$

Remark: Minimum bandwidth of pulse spectrum $G(f)$ is $\frac{1}{T}$. Otherwise less than one period of $\Phi_{aa}(\mathrm{e}^{j2\pi fT})$ would be contained in $\bar{\Phi}_{ss}(f)$, i.e., not all information present in $a[k]$ would be contained in $s(t)$!

End Supplement 1

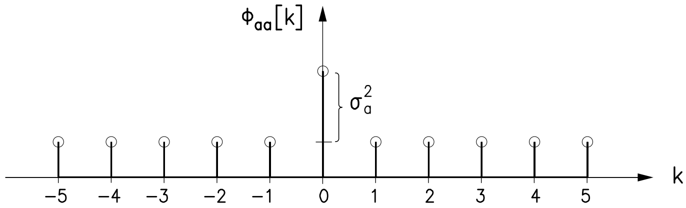
Special case: independent, equally probable source symbols $q[l]$, no channel coding: The amplitude coefficients are mutually independent and equally probable:

$$\Pr(a_m) = \frac{1}{M}, \quad \forall a_m \in \boldsymbol{\mathcal{A}}$$

$$\Rightarrow \text{ ACF } \phi_{aa}[k] = \begin{cases} \sigma_a^2 + |m_a|^2 & \text{for } k = 0 \\ \\ |m_a|^2 & \text{for } k \in \mathbb{Z}\backslash\{0\} \end{cases}$$

$m_a = \frac{1}{M} \sum\limits_{a_m \in \boldsymbol{\mathcal{A}}} a_m$ average taken over all amplitude coefficients

$\sigma_a^2 = \frac{1}{M} \sum\limits_{a_m \in \boldsymbol{\mathcal{A}}} |a_m|^2 - |m_a|^2$ variance of the amplitude coefficients



Power spectral density

$$\Phi_{aa}(\mathrm{e}^{j2\pi F}) = \sum_{k=-\infty}^{+\infty} \phi_{aa}[k] \cdot \mathrm{e}^{-\mathrm{j}2\pi kF} = \sigma_a^2 + \left\{ |m_a|^2 \cdot \sum_{k=-\infty}^{+\infty} \mathrm{e}^{-\mathrm{j}2\pi kF} \right\}$$

Fourier series of a train of δ–impulses

$$\sum_{k=-\infty}^{+\infty} \mathrm{e}^{\,-\mathrm{j}2\pi kF} \equiv \sum_{\ell=-\infty}^{+\infty} \delta(F - l)$$

Power spectral density of a discrete time sequence of amplitude coefficients

$$\Phi_{aa}(\mathrm{e}^{j2\pi F}) = \sigma_a^2 + |m_a|^2 \sum_{\ell=-\infty}^{+\infty} \delta(F - l)$$

Average power spectral density of a PAM–signal:

$$\bar{\Phi}_{ss}(f) = \sigma_a^2 \, \frac{\left|G(f)\right|^2}{T} + |m_a|^2 \, \frac{\left|G(f)\right|^2}{T^2} \sum_{\ell=-\infty}^{+\infty} \delta\left(f - \tfrac{\ell}{T}\right)$$

If the mean of the amplitude coefficients is not zero, discrete spectral lines appear at multiples of $1/T$ in the frequency domain (provided that $G(\frac{\ell}{T}) \neq 0$).
Average power of the transmit signal:

$$S_\mathrm{s} = (\sigma_a^2 + |m_a|^2) \cdot \frac{E_\mathrm{g}}{T}$$

Special case $m_a = 0$:

$$S_\mathrm{s} = \frac{\sigma_a^2}{T}\, E_\mathrm{g}\ ; \quad E_\mathrm{s} = \sigma_a^2\, E_\mathrm{g}; \quad E_\mathrm{b} = \frac{E_\mathrm{s}}{R} = \frac{\sigma_a^2}{R} E_\mathrm{g}$$

No channel coding: $\qquad R = \mathrm{ld}(M)$

3.2 Digital Baseband Transmission

If the (real) baseband signal $s(t)$ is not shifted to higher frequencies, the term **Digital Baseband Transmission** is used.

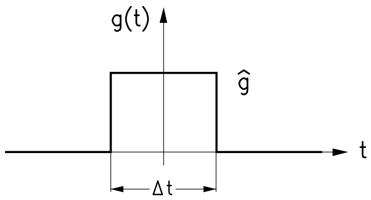
Application: In communication channels with lowpass filtering characteristics (e.g. cables) to avoid the large attenuation at higher frequencies.

$s(t)$ is a real physical signal:

$\Rightarrow$ Fundamental pulse $g(t)$, amplitude coefficients $a_i \in \mathbb{R}$ or alternatively $\boldsymbol{\mathcal{A}} \subset \mathbb{R}$

Example: *Digital binary baseband PAM with rectangular pulses*

$$M = 2; \quad T = T_\mathrm{b}; \quad R = 1 \frac{\mathrm{bit}}{\mathrm{symbol}}$$

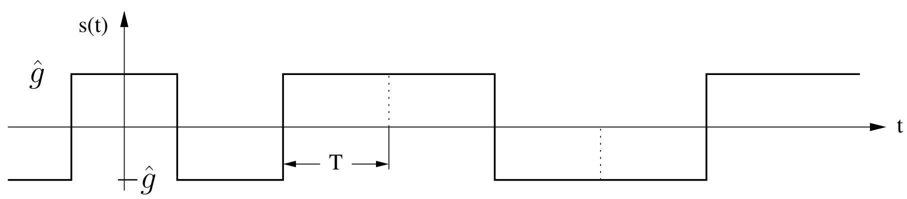


$$g(t) = \hat{g}\ \mathrm{rect}\left(\frac{t}{\Delta t}\right)$$

$$E_\mathrm{g} = \hat{g}^2 \Delta t$$

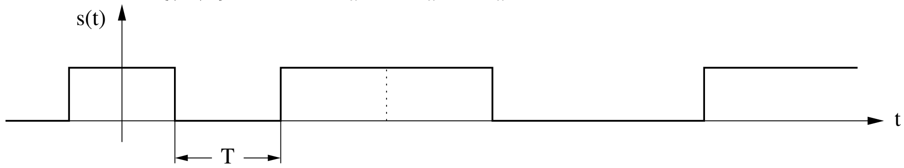
a) **Bipolar NRZ** (no return to zero) signal:

$\boldsymbol{\mathcal{A}} = \{-1; +1\}; \quad \sigma_a^2 = 1; \quad m_a = 0; \quad \Delta t = T; \quad S = \hat{g}^2$



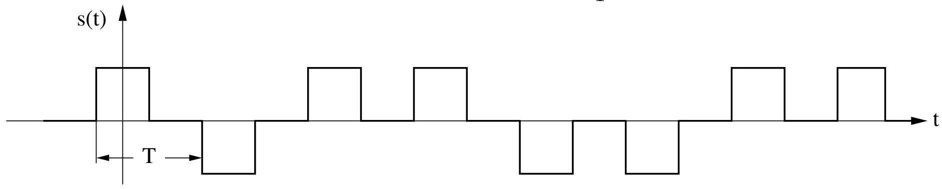
b) **Unipolar NRZ** signal:

$\boldsymbol{\mathcal{A}}= \{0; +2\}$; $\Delta t = T$; $E\{|a|^2\} = 2 = \sigma_a^2 + m_a^2$; $\sigma_a^2 = 1$; $m_a = 1$; $S = 2\hat{g}^2$



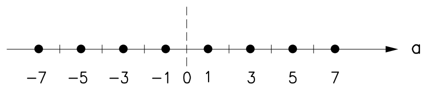
c) **Bipolar RZ** signal (return to zero): $\Delta t < T$

$\boldsymbol{\mathcal{A}}= \{-1; +1\}$; $\sigma_a^2 = 1$; $m_a = 0$; e.g. $\Delta t = T/2, S = \hat{g}^2 \frac{\Delta t}{T}$



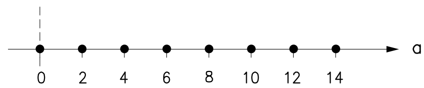
Advantage of RZ signals: Facilitates receiver synchronisation since the clock frequency can be recovered through full–wave rectification! Disadvantage of RZ signals: Extremely **bandwidth inefficient** and **higher crest factor** (Peak to Average Power Ratio (PAPR))

Conventional signal constellation for digital baseband transmission:

<table>
<tr><th>bipolar; M even</th><th>unipolar; $M \in \mathbb{N}$</th></tr>
<tr><td>


$\boldsymbol{\mathcal{A}}= \{\pm 1, \pm 3, \pm 5, \ldots, \pm (M-1)\}$

$\boxed{\sigma_a^2 = \frac{2}{M} \sum_{i=1}^{M/2} (2i-1)^2 = \frac{M^2-1}{3}}$

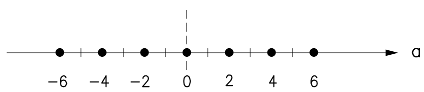
$m_a = 0$
</td><td rowspan="3">


$\boldsymbol{\mathcal{A}}= \{0, 2, 4, \ldots, (2M-2)\}$

$\bar{a}^2 = \frac{1}{M} \sum_{i=0}^{M-1} (2i)^2 = \frac{4M^2-6M+2}{3}$

$m_a = \frac{1}{M} \sum_{i=0}^{M-1} 2i = M - 1$

$\sigma_a^2 = \bar{a}^2 - m_a^2$

$\sigma_a^2 = \frac{M^2-1}{3}$
</td></tr>
<tr><th>bipolar; M odd</th></tr>
<tr><td>


$\boldsymbol{\mathcal{A}}= \{0, \pm 2, \pm 4, \ldots, \pm(M-1)\}$

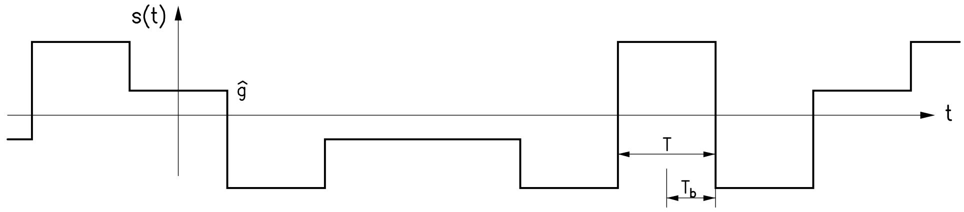
$\sigma_a^2 = \frac{M^2-1}{3}$

$m_a = 0$
</td></tr>
<tr><td colspan="2">$\sigma_a^2 = \frac{M^2-1}{3}$ in all cases!</td></tr>
</table>

Side note: Sequence formulas: $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$; $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$

Example: *4–ary (quaternary) bipolar NRZ baseband signal with square pulse*

$\boldsymbol{\mathcal{A}} = \{\pm 1, \pm 3\}$; $\sigma_a^2 = 5$; $S = 5\hat{g}^2$; $E_\mathrm{b} = 2.5 E_\mathrm{g}$; $E_\mathrm{g} = \hat{g}^2 \cdot T$; $E_\mathrm{s} = 5 E_\mathrm{g}$



Doubled bandwidth efficiency when compared to binary (2–ary) signal! $\Rightarrow$ particularly well–suited for transmission channels with lowpass behavior (cables), as high attenuation of significant signal components at high frequencies is avoided!

3.3 Carrier Modulated Digital PAM

The signal $s(t) = \sum\limits_{k=-\infty}^{+\infty} a_{m[k]}\ g(t-kT)$ is an ECB signal representing a high frequency, real, physical transmit signal:

$$s_{\mathrm{RF}}(t) = \sqrt{2}\mathrm{Re}\big\{s(t)\cdot \mathrm{e}^{\,\mathrm{j}2\pi f_{\mathrm{c}} t}\big\}$$

f_{c} : (Carrier frequency)

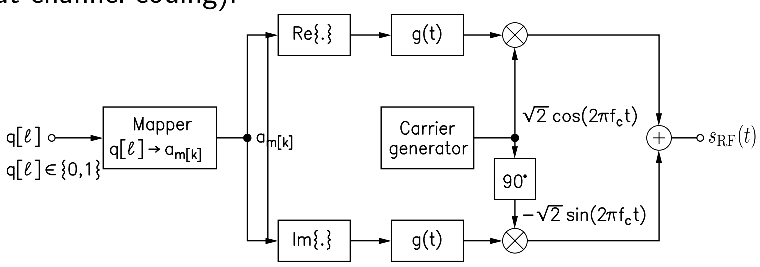
Normally **real–valued** fundamental pulses are used: $g(t) \in \mathbb{R}$;
In this case, we have

$$s_{\mathrm{RF}}(t) = \sqrt{2}\left(\sum_{k=-\infty}^{+\infty} \mathrm{Re}\{a_{m[k]}\}\ g(t-kT)\right)\cdot\cos(2\pi f_{\mathrm{c}} t) \qquad \text{In–phase component}$$

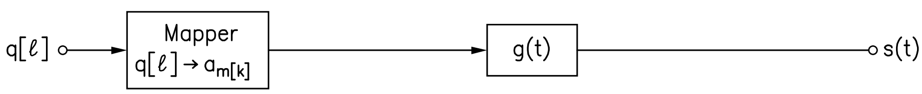
$$-\sqrt{2}\left(\sum_{k=-\infty}^{+\infty} \mathrm{Im}\{a_{m[k]}\}\ g(t-kT)\right)\cdot\sin(2\pi f_{\mathrm{c}} t) \qquad \text{Quadrature component}$$

Here, **complex** amplitude coefficients are also allowed: $\boldsymbol{\mathcal{A}} \subset \mathbb{C}$!

Block diagram (without channel coding):



ECB Model:



3.3.1 Amplitude Shift Keying (ASK)

One–dimensional amplitude modulation, i.e., restriction to real amplitude coefficients: $a_m \in \mathbb{R}$

Amplitude Shift–Keying (ASK)

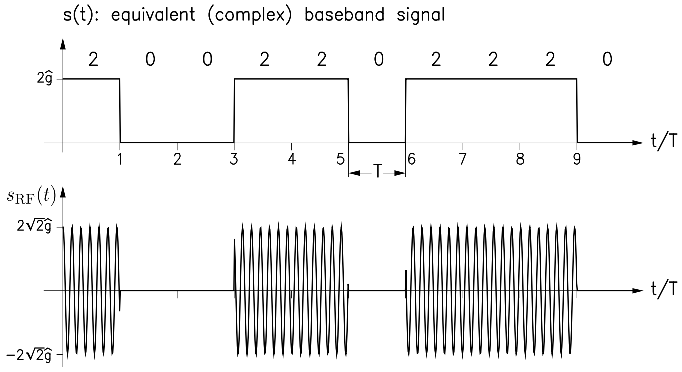
Information is transmitted using in–phase component only! Unipolar and bipolar constellations are possible as with baseband transmission, see Section 3.2.

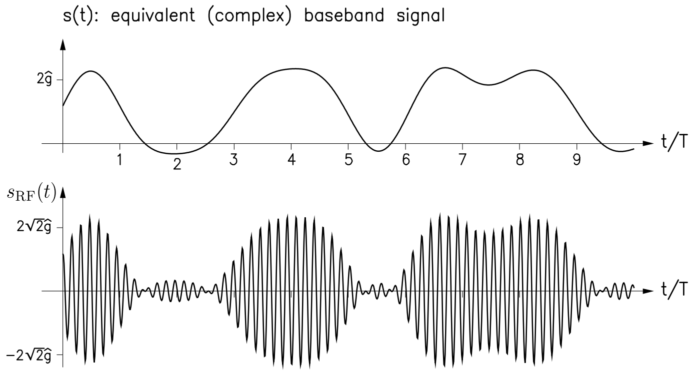
Example: *Unipolar, binary ASK*

$M = 2$; $\boldsymbol{\mathcal{A}} = \{0, 2\}$, also known as

<u>O</u>N/<u>O</u>FF–<u>K</u>eying (OOK)

- Hard keying: $g(t) = \hat{g}\,\mathrm{rect}(t/T)$
- Soft keying with so–called "Square Root–Nyquist" pulse ($\alpha = 0{,}5$), see Section 3.6



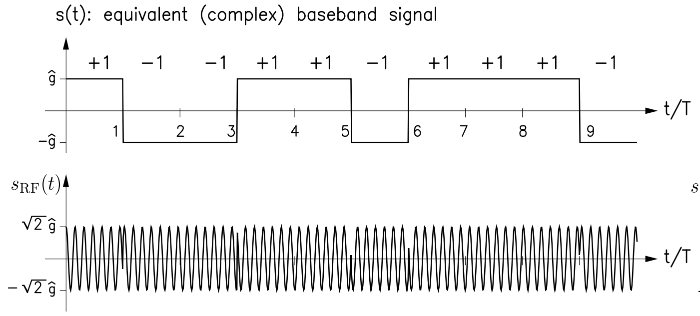


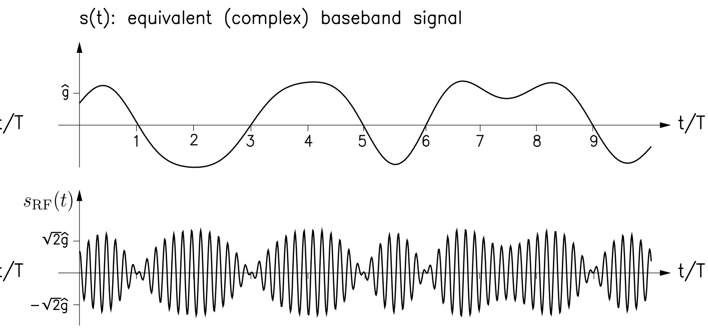
Example: *Bipolar, binary ASK*

$M = 2$; $\boldsymbol{\mathcal{A}} = \{-1, +1\}$, also known as

<u>B</u>inary <u>P</u>hase <u>S</u>hift <u>K</u>eying (BPSK)

- Hard keying: $g(t) = \hat{g}\,\mathrm{rect}(t/T)$
- Soft keying with so–called "Square Root–Nyquist" pulses ($\alpha = 0{,}5$)





discontinuities and peaks $\rightarrow$ bandwith inefficient

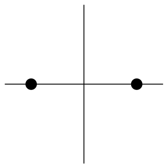
Bandwidth efficient, but oscillating envelope: high crest–factor; power inefficiency due to required "back–off" in high power amplifier (HPA) at transmitter output

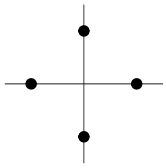
3.3.2 Digital Phase Modulation (PSK)

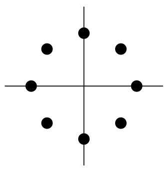
For digital phase modulation (phase–shift–keying, PSK), the signal constellation consists of equidistant points on the unit circle in the complex plane.

$$\boldsymbol{\mathcal{A}} = \left\{ \mathrm{e}^{\,\mathrm{j}2\pi\frac{m-1}{M}} \,\middle|\, m \in \{1, 2, \ldots, M\} \right\}$$

$$\Rightarrow \sigma_a^2 = 1, \quad m_a = 0, \quad \forall M \in \mathbb{N}$$

$M = 2$  2 ASK bipolar $\equiv 2$ PSK $=$ BPSK

$M = 4$  4 PSK $\equiv$ quadrature PSK (QPSK) most commonly used scheme

$M = 8$  8 PSK: often used together with $R_\mathrm{c} = 2/3$ channel coding: $R = 2$ like QPSK

68

Advantages of PSK:

- For hard keying $(g(t) = \hat{g}\,\mathrm{rect}(t/T))$ the transmit signal has a **constant envelope**. Signal segments inside of the modulations intervals are cosinus waves with M different phase offsets.

Disadvantages:

- For soft keying, strong variations in the signal envelope occur.
- For $M \geq 8$; (i.e., for bandwidth efficient schemes), signal points are very closely spaced $\rightarrow$ low power efficiency.

3.3.3 Digital Quadrature Amplitude Modulation (QAM)

Quadrature amplitude modulation (QAM): bipolar ASK modulation in in–phase and quadrature component: twice the bandwidth efficiency of ASK at equal power efficiency.

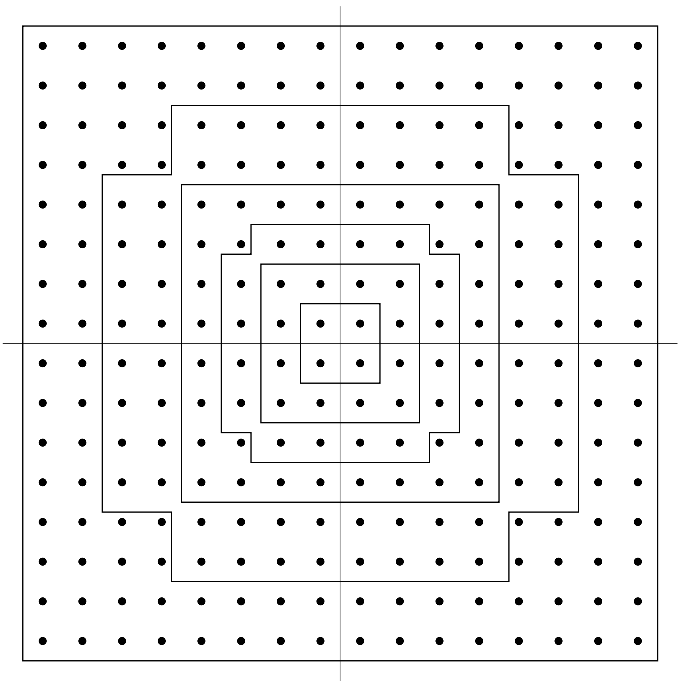
$$\boldsymbol{\mathcal{A}} \subset (2\mathbb{Z}+1) + \mathrm{j}(2\mathbb{Z}+1)$$ Lattice of complex numbers with odd–valued integer real and imaginary parts

For $\sqrt{M} \in \mathbb{N}$ **"quadratic constellations"** are common (rate for no channel coding):

4 QAM	$R = 2$ bit/symbol
16 QAM	$R = 4$ bit/symbol
64 QAM	$R = 6$ bit/symbol
256 QAM	$R = 8$ bit/symbol

For $\sqrt{M} \notin \mathbb{N}$ so–called **"cross–constellations"** are common:

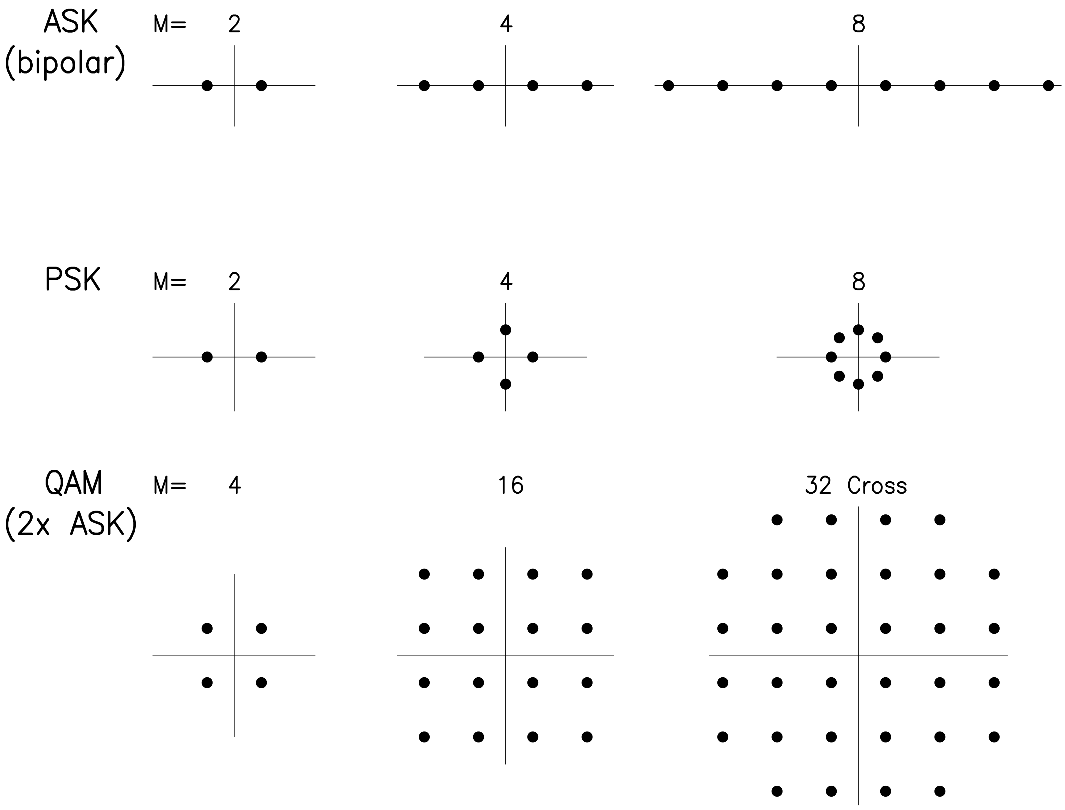
for example:	32 QAM	$R = 5$ bit/symbol
	128 QAM	$R = 7$ bit/symbol



QAM constellations with a high number of levels (signal points) are used, for example, in digital directional radio.

$M =$	4	16	32	64	128	256
			cross		cross	
$\sigma_a^2 =$	2	10	20	42	82	170
		↑		↑		↑
		digital directional radio				
$m_a =$	0	(DC–free, no spectral lines)				

Summary of commonly used signal constellations for digital PAM:



72

3.3.4 Crest–Factor of PAM Transmit Signals

Crest–factor: $$\zeta = \frac{\max\{|s(t)|\}}{\sqrt{S_\mathrm{s}}} = \frac{\max\limits_{\forall t \in \mathbb{R}; \langle a_{m[k]} \rangle} \left| \sum\limits_{k=-\infty}^{+\infty} a_{m[k]} g(t-kT) \right|}{\sqrt{\mathrm{E}\left\{|a_m|^2\right\} \cdot E_\mathrm{g}/T}} \quad ; \qquad \mathrm{PAR} = \zeta^2$$

with $|x \cdot y| = |x| \cdot |y|$ and $|x+y| \le |x| + |y|, x, y \in \mathbb{C}$, we have

$$\max_{\forall t \in \mathbb{R}; \langle a_{m[k]} \rangle} \left| \sum_{k=-\infty}^{+\infty} a_{m[k]} g(t-kT) \right| \le \max_{a_m \in \mathrm{A}} |a_m| \cdot \max_t \sum_{k=-\infty}^{+\infty} |g(t-kT)|$$

$$\Rightarrow \boxed{\zeta \le \zeta_a \cdot \zeta_g} \text{ with}$$

Crest–factor of signal constellation: $$\zeta_a = \frac{\max\limits_{a_m} |a_m|}{\sqrt{\mathrm{E}\{|a_m|^2\}}}$$

Crest–factor of fundamental pulse: $$\zeta_g = \frac{\max\limits_t \sum\limits_{k=-\infty}^{+\infty} |g(t-kT)|}{\sqrt{E_\mathrm{g}/T}}$$

Crest–Factors of Commonly–used PAM Constellations

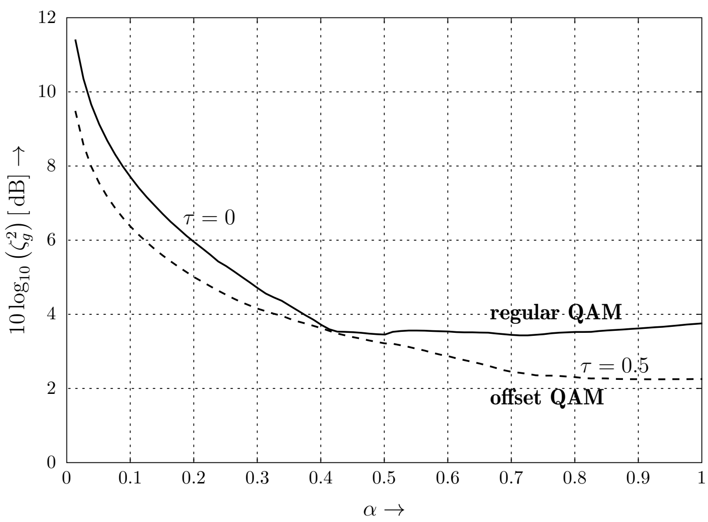
M	2	4	8	16	32	64	128	256
ASK bipolar	1	$\sqrt{\frac{9}{5}}$	$\sqrt{\frac{7}{3}}$	$\sqrt{\frac{45}{17}}$				
ASK unipolar	$\sqrt{2}$	$\sqrt{\frac{18}{7}}$	$\sqrt{\frac{14}{5}}$	$\sqrt{\frac{90}{31}}$				
PSK	1	1	1	1	1	1	1	1
QAM	–	1	–	$\sqrt{\frac{9}{5}}$	$\sqrt{\frac{17}{10}}$	$\sqrt{\frac{7}{3}}$	$\sqrt{\frac{85}{41}}$	$\sqrt{\frac{45}{17}}$

PAR in dB

M	2	4	8	16	32	64	128	256
ASK bipolar	0	2.55	3.68	4.23				
ASK unipolar	3	4.10	4.47	4.63				
PSK	0	0	0	0	0			
QAM	–	0	–	2.55	2.30	3.68	3.17	4.23

Crest–factor ζ_g for $\sqrt{\text{Nyquist}}$–pulse with cosine roll–off spectral slope as a function of the roll–off factor α, see Section 3.6

τ: Time offset between in–phase and quadrature components for so–called offset–QAM schemes.



3.4 Coherent Demodulation for Digital PAM and AWGN Channel

Receiver input signal (baseband transmission or ECB-representation of RF-signal)

$$e(t) = \mathrm{e}^{j\varphi} \cdot \left(\sum_{k=-\infty}^{+\infty} a_{m[k]} g(t - kT - \Delta T) \right) + n(t)$$

$n(t)$: complex Gaussian noise with noise power spectral density N_0 ($N_0/2$ per quadrature component)

ΔT: Symbol clock shift (with respect to the **receiver** clock)

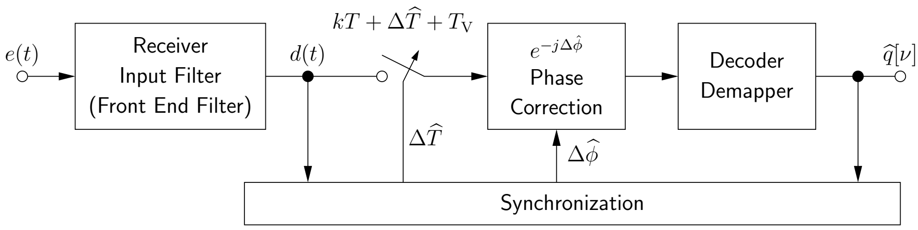
φ: Carrier phase offset (with respect to the **receiver** carrier oscillation)

(for digital baseband transmission we have always: $\varphi = 0$)

Receiver synchronisation: With the help of the highly attenuated and distorted receiver input signal, the symbol clock shift ΔT is estimated and eliminated

Coherent detection schemes for carrier–modulated transmission: Carrier phase offset φ is estimated and eliminated.

76



Synchronization methods: Training sequences, pilot symbols, exploitation of cyclo-stationary property of the signal.

(See lecture on "Synchronization Procedures for Digital Communications")

Note: If estimation and correction of the carrier phase are not performed (in order to save computing power or due to the rapid variations of φ), the resulting system is referred to as an **noncoherent communication system**. In this case, the same digital sequence will be represented by any phase rotated version of the ECB transmit signal. This will be discussed in greater detail later.

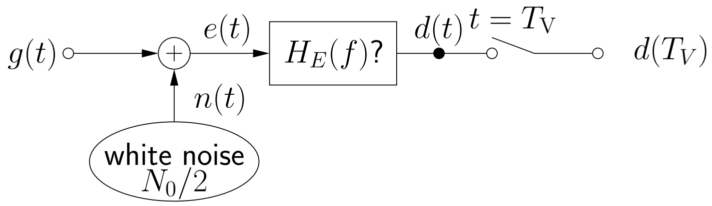
For the purpose of this lecture, we will assume the following:

- **Ideal symbol clock synchronization**: equivalent to: $\Delta T \overset{!}{=} 0$
- **Coherent demodulation with ideal carrier phase synchronization**: $\varphi \overset{!}{=} 0$

3.5 Optimal Detection of Pulses in AWGN

Problem: Detection of an energy–limited pulse $g(t)$, whose shape is known to the receiver, in the case of impairement by additive white Gaussian noise (AWGN).

Block diagram (real, physical signals and systems):



Question: What is the optimal receiver filter $H_\mathrm{E}(f)$ which will **bandlimit the noise** in such a way that the signal to noise ratio (SNR) is maximized at the **detection time** T_V.

Conflicting requirements for the receiver filter:

a) Narrowband so that the noise power in the detection signal $d(t)$ is minimized.

b) Wideband so that the signal will be minimally attenuated and distorted.

For which filter $H_\mathrm{E}(f)$ will the SNR of the detection signal SNR_d at detection time T_V

$$\mathrm{SNR}_d \stackrel{\mathrm{def}}{=} \frac{|\tilde{d}(T_\mathrm{V})|^2}{\sigma_{n_d}^2}$$

be maximized?

Signal:

$$\tilde{d}(t) = g(t) * h_\mathrm{E}(t)$$

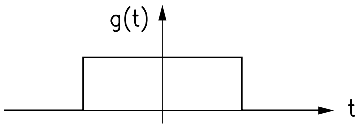
$$\tilde{d}(T_\mathrm{V}) = \big(g(t) * h_\mathrm{E}(t)\big)\Big|_{t = T_\mathrm{V}}$$

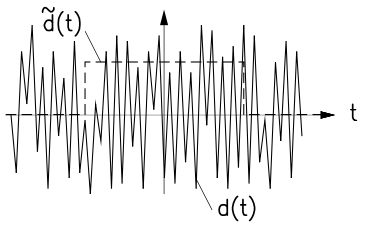
Noise variance (power): $\displaystyle \sigma_{n_d}^2 = \frac{N_0}{2} \int_{-\infty}^{+\infty} \left|H_\mathrm{E}(f)\right|^2 \mathrm{d}f$ with

$\sigma_{n_d}^2$: Variance of the **bandlimited** noise in the detection signal at the output of the receiver filter $h_\mathrm{E}(t)$

Two extremes as examples:

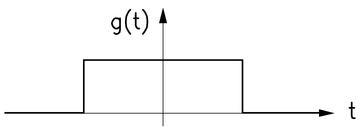
a) Receiver input filter with very high bandwidth, (bandwidth$\rightarrow \infty$; AWGN power $\rightarrow \infty$, theoretically infinite amount of noise!)

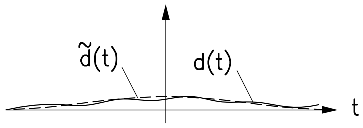




$\mathrm{SNR}_d \rightarrow 0$

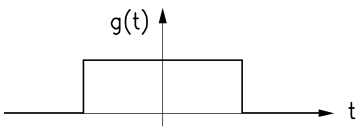
b) Receiver input filter with very low bandwidth, e.g.: serious distortion and attenuation of signal

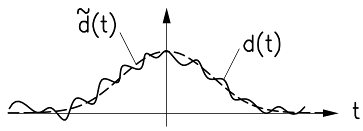




$\mathrm{SNR}_d \rightarrow 0$

Conclusion: A **compromise** between signal attenuation and noise bandlimiting is optimal.





$$\mathrm{SNR}_d = \frac{\left|\tilde{d}(T_\mathrm{V})\right|^2}{\sigma_{n_d}^2} = \frac{\left|\int\limits_{-\infty}^{+\infty} G(f)\; H_\mathrm{E}(f)\; \mathrm{e}^{\,\mathrm{j}2\pi f T_\mathrm{V}}\; \mathrm{d}f\right|^2}{\frac{N_0}{2}\int\limits_{-\infty}^{+\infty} \left|H_\mathrm{E}(f)\right|^2 \mathrm{d}f}$$

Derivation of an upper limit for the SNR of the detection signal SNR_d:

Nominator:

1. $$\left|\int\limits_{-\infty}^{+\infty} G(f)\; H_\mathrm{E}(f)\; \mathrm{e}^{\,\mathrm{j}2\pi f T_\mathrm{V}}\; \mathrm{d}f\right|^2 \leq \left[\int\limits_{-\infty}^{+\infty} \left|G(f)\right| \cdot \left|H_\mathrm{E}(f)\right| \cdot 1\; \mathrm{d}f\right]^2$$
   Absolute value of a sum is $\leq$ the sum of absolute values

2. $$\left(\int\limits_{-\infty}^{+\infty} \left|G(f)\right| \cdot \left|H_\mathrm{E}(f)\right|\; \mathrm{d}f\right)^2 \leq \left(\int\limits_{-\infty}^{+\infty} \left|G(f)\right|^2 \mathrm{d}f\right) \cdot \left(\int\limits_{-\infty}^{+\infty} \left|H_\mathrm{E}(f)\right|^2 \mathrm{d}f\right)$$
   Cauchy-Schwarz inequality

Note on Cauchy–Schwarz inequality: For $\vec{a}$, $\vec{b} \in \mathbb{R}^n$ it holds: $\vec{a} \cdot \vec{b}^{\mathrm{T}} = |\vec{a}| \cdot |\vec{b}| \cdot \cos(\phi_{\vec{a}\vec{b}}) \leq |\vec{a}| \cdot |\vec{b}|$; thus

$$\left(\sum_{\nu=1}^{n} a_\nu b_\nu\right)^2 \leq \left(\sum_{\nu=1}^{n} a_\nu^2\right) \cdot \left(\sum_{\nu=1}^{n} b_\nu^2\right)$$

here: $a_\nu \to \big|G(f)\big|$; $b_\nu \to \big|H_{\mathrm{E}}(f)\big|$; $\nu \to f$; $n \to \infty$

$$\Rightarrow \mathrm{SNR}_d \leq \frac{\left(\int\limits_{-\infty}^{\infty} |G(f)|^2 \mathrm{d}f\right) \cdot \left(\int\limits_{-\infty}^{\infty} |H_{\mathbf{E}}(f)|^2 \mathrm{d}f\right)}{\frac{N_0}{2} \int\limits_{-\infty}^{\infty} |H_{\mathbf{E}}(f)|^2 \mathrm{d}f} = \frac{\int\limits_{-\infty}^{+\infty} \big|G(f)\big|^2 \, \mathrm{d}f}{N_0/2} = \frac{E_{\mathrm{g}}}{N_0/2}$$

$$\text{with } E_{\mathrm{g}} = \int_{-\infty}^{+\infty} \big|G(f)\big|^2 \, \mathrm{d}f = \int_{-\infty}^{+\infty} \big|g(t)\big|^2 \, \mathrm{d}t, \quad \text{energy of the pulse } g(t)$$

> The SNR of the detection signal is less than or equal to the signal energy divided by the two–sided noise power spectral density (independent of pulse shape)!

A filter $H_{\mathrm{E}}(f)$, which provides an SNR of $\mathrm{SNR}_d = E_{\mathrm{g}}/(N_0/2)$ is an optimal filter!

Approach: Phase reversal on the real axis and adaption of amplitude response

$$H_{\mathrm{E}}(f) = \gamma \, G^*(f) \, \mathrm{e}^{-\mathrm{j}2\pi f T_{\mathrm{V}}}$$

where γ is an arbitrary real constant that can be used to correct the signal dimension

$$\mathrm{SNR}_d = \frac{\left|\int\limits_{-\infty}^{+\infty} G(f) \, \gamma \, G^*(f) \cdot \mathrm{e}^{-\mathrm{j}2\pi f T_{\mathrm{V}}} \cdot \mathrm{e}^{+\mathrm{j}2\pi f T_{\mathrm{V}}} \, \mathrm{d}f\right|^2}{\frac{N_0}{2} \, \gamma^2 \int\limits_{-\infty}^{+\infty} \big|G(f)\big|^2 \, \mathrm{d}f} = \frac{\gamma^2 E_{\mathrm{g}}^2}{\frac{N_0}{2} \, \gamma^2 E_{\mathrm{g}}} = \frac{E_{\mathrm{g}}}{N_0/2}$$

$$\mathrm{SNR}_d = \frac{E_{\mathrm{g}}}{N_0/2} = \frac{\text{pulse energy}}{\text{two–sided noise power spectral density}}$$

Therefore, $H_{\mathrm{E}}(f)$ is optimal!

Note:

From $g(t) \circ\!\!-\!\!\bullet \, G(f)$ follows: $g^*(t) \circ\!\!-\!\!\bullet \, G^*(-f)$ and $g^*(-t) \circ\!\!-\!\!\bullet \, G^*(f)$

Matched filter (MF) for an energy–limited signal pulse $g(t)$, in AWGN (also known as "optimal search filter"):

$$\text{Matched}\quad \boxed{\begin{aligned} H_{\mathrm{M}}(f) &= \gamma\, G^*(f) \cdot \mathrm{e}^{-\mathrm{j}2\pi f T_{\mathrm{V}}} \\ h_{\mathrm{M}}(t) &= \gamma\, g^*(T_{\mathrm{V}} - t) \end{aligned}}\quad \text{Filter}$$

Impulse response of a matched filter: Symmetrically reflected and possibly delayed version of the pulse $g(t)$ itself. Using the matched filter at the receiver will lead to the greatest possible SNR for a given detection time. The maximum is **independent** of the shape of the pulse! Only the **pulse energy** and the **two–sided noise power spectral density** specify the SNR: **Optimal filter for pulse detection (North, 1942)**.

Detection signal at the output of the matched filter excited by the pulse $g(t)$,

$$\tilde{d}(t) \;=\; g(t) * \gamma\, g^*(T_{\mathrm{V}} - t) = \gamma \int_{-\infty}^{+\infty} g(t')\, g^*(t' - t + T_{\mathrm{V}})\, \mathrm{d}t' = \gamma\, \varphi_{gg}(t - T_{\mathrm{V}}) \qquad \text{with}$$

$$\varphi_{gg}(t) := \int_{-\infty}^{+\infty} g(t'' + t)\, g^*(t'')\, \mathrm{d}t'' \quad \text{ACF of the energy limited pulse } g(t); \quad \varphi_{gg}(0) = E_{\mathrm{g}}$$

$\Rightarrow$The matched filter is also referred to as **correlation filter**

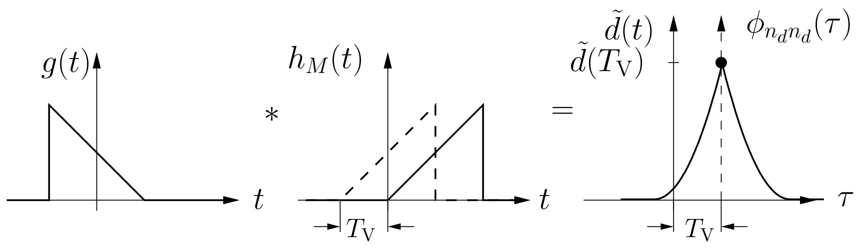
ACF of Gaussian (band limited) **noise process** at the output of a matched filter:

$$\begin{aligned} \phi_{n_d n_d}(\tau) &= \tfrac{N_0}{2}\, \delta(\tau) * h_{\mathrm{M}}(\tau) * h^*_{\mathrm{M}}(-\tau) = \tfrac{N_0}{2} \int_{-\infty}^{+\infty} h_{\mathrm{M}}(\tau - \tau')\, h^*_{\mathrm{M}}(-\tau')\, \mathrm{d}\tau'; \quad h_{\mathrm{M}}(t) = \gamma g^*(T_{\mathrm{V}} - t) \text{ inserted} \\ &= \tfrac{N_0}{2}\, \gamma^2 \int_{-\infty}^{+\infty} g^*(T_{\mathrm{V}} - \tau + \tau')\, g(T_{\mathrm{V}} + \tau')\, \mathrm{d}\tau'; \qquad \text{substitution: } \tau'' = T_{\mathrm{V}} - \tau + \tau' \\ &= \gamma^2\, \tfrac{N_0}{2} \int_{-\infty}^{+\infty} g(\tau'' + \tau)\, g^*(\tau'')\, \mathrm{d}\tau'' = \gamma^2\, \tfrac{N_0}{2}\, \varphi_{gg}(\tau) \end{aligned}$$

Theorem:

At the output of a matched filter, the desired signal $\tilde{d}(t)$ and the ACF of the noise $\phi_{n_d n_d}(\tau)$ are proportional to each other. Both are in turn proportional to the ACF $\varphi_{gg}(\tau)$ of the energy–limited pulse to which the filter is adapted.

Example: *Pulse, Impulse Response of the Matched Filter, Detection Pulse:*



T_{V}: delay for causal implementation of the matched filter

At time $t = T_{\mathrm{V}}$, (an optimal a decision on a discrete–valued amplitude coefficient a can be made based on the output signal of the matched filter as long as no interference caused by overlapping of previous and subsequent pulses occurs (i.e., in the case where only a single, isolated pulse is transmitted)!

3.6 Inter Symbol Interference–free Communication Over an AWGN Channel

AWGN channel:

– Neither linear nor nonlinear distortion of the signal, signal attenuation only.

– Interference by AWGN

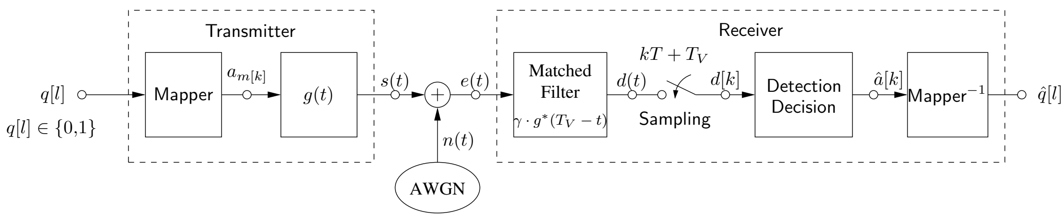
$$\mathrm{PSD} = \begin{cases} N_0 & \text{for ECB signals } \big(n(t) \text{ complex }\big) \\ N_0/2 & \text{for digital baseband transmission } \big(n(t) \text{ real }\big) \end{cases}$$

Optimal extraction of single pulses from the noisy receiver input signal by means of

matched filters

PAM Transmission Model (without channel coding)

Digital baseband transmission or alternatively ECB model for carrier modulated transmission:



3.6.1 Detection Signal

$$\begin{aligned}\tilde{d}(t) &= d(t) * h_{\mathrm{E}}(t) = d(t) * \gamma\, g^*(T_{\mathrm{V}} - t) = \left[\sum_{k=-\infty}^{+\infty} a_{m[k]}\, g(t-kT)\right] * \gamma\, g^*(T_{\mathrm{V}} - t) = \\ &= \gamma \sum_{k=-\infty}^{+\infty} a_{m[k]} \big[g(t-kT) * g^*(T_{\mathrm{V}} - t)\big]\end{aligned}$$

$$\begin{aligned}g(t-kT) * g^*(T_{\mathrm{V}} - t) &= \int_{-\infty}^{+\infty} g^*(T_{\mathrm{V}} - t + \tau)\, g(\tau - kT)\, \mathrm{d}\tau; \quad \text{substitution: } \tau' = T_{\mathrm{V}} - t + \tau;\ \mathrm{d}\tau = \mathrm{d}\tau' \\ &= \int_{-\infty}^{+\infty} g(\tau' + t - T_{\mathrm{V}} - kT)\, g^*(\tau')\, \mathrm{d}\tau' = \varphi_{gg}(t - kT - T_{\mathrm{V}}) \qquad \text{with}\end{aligned}$$

$$\varphi_{gg}(t) = \int_{-\infty}^{+\infty} g(\tau' + t)\, g^*(\tau')\, \mathrm{d}\tau'$$

ACF of the fundamental pulse!

$$\boxed{\tilde{d}(t) = \gamma \sum_{k=-\infty}^{+\infty} a_{m[k]}\, \varphi_{gg}(t - kT - T_{\mathrm{V}})}$$

PAM signal with ACF of the fundamental pulse as pulse shape

Conditions for Inter-Symbol Interference–Free Detection:

Sample at time $\ell T + T_{\mathrm{V}}$ is dependent **only** on **corresponding amplitude coefficient** $a_{m[\ell]}$ and **not** on **preceeding** samples $a_{m[\ell-i]}$ or **subsequent** $a_{m[\ell+i]}$ ones; $i \in \mathbb{N}$; i.e., *pulses* do not interfere with each other at the optimal detection time instants:

Intersymbol interference–free transmission: No ISI

$$\tilde{d}[l] \stackrel{\mathrm{def}}{=} \tilde{d}(\ell T + T_{\mathrm{V}}) \stackrel{!}{=} a_{m[\ell]}$$

$$\tilde{d}(\ell T + T_{\mathrm{V}}) = \gamma \sum_{k=-\infty}^{+\infty} a_{m[k]}\, \varphi_{gg}(\ell T + T_{\mathrm{V}} - kT - T_{\mathrm{V}}) = \gamma \sum_{k=-\infty}^{+\infty} a_{m[k]}\, \varphi_{gg}\big((\ell - k)T\big) \stackrel{!}{=} a_{m[\ell]}$$

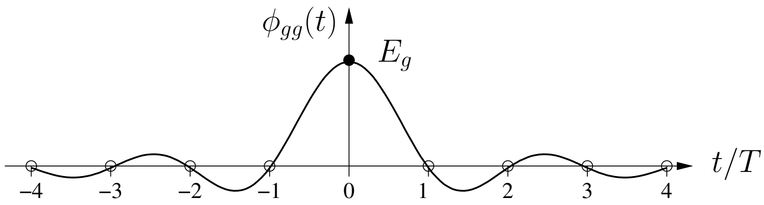
The **autocorrelation** function $\varphi_{gg}(t)$ of the pulse $g(t)$ must have **equidistant zero crossings** for all λT with $\lambda \in \mathbb{Z}$, $\lambda \neq 0$. Theredore, it must be so called **Nyquist–function**:

The autocorrelation $\varphi_{gg}(t)$ must fulfill the **first Nyquist criterion** (Nyquist 1927):

$$\boxed{\varphi_{gg}(\lambda T) = \int_{-\infty}^{+\infty} g(t' + \lambda T) g^*(t') dt' \stackrel{!}{=} \left\{\begin{array}{ll} E_{\mathrm{g}} & \text{for } \lambda = 0 \\ 0 & \text{for } \lambda \in \mathbb{Z}\backslash\{0\} \end{array}\right\} = \delta[\lambda]\, E_{\mathrm{g}}} \tag{3.1}$$

In other words, for optimal, ISI–free detection of digital PAM signals transmitted over AWGN channels, the fundamental pulse, $g(t)$, must be chosen so that the pulse $g(t)$ itself and delayed pulses $g(t-\lambda T)$, delayed by multiples, λ, of the modulation interval T are mutually

orthogonal.



$\Rightarrow$ **Orthogonality condition for PAM**

This condition for PAM exactly corresponds to the requirement of **temporal** orthogonality of signal elements $s_i(t)$ stated in Section 2.1.2.1:
Temporal orthogonality allows individual pulses (signal elements) in a continuous signal to be detected in an optimal manner exactly as if neither preceeding pulses nor subsequent pulses existed.

If orthogonality holds, the proportionality constant of the matched filter γ is obtained as

$$\gamma \; a_{m[\ell]} \; E_{\mathrm{g}} \stackrel{!}{=} a_{m[\ell]} \Rightarrow \boxed{\gamma = \frac{1}{E_g}}$$

Note: This choice of γ produces a **dimensionless** output signal of the matched filter, i.e., a normalized, discrete–time detection signal $d[k]$ which corresponds to the sequence $a_{m[k]}$ of (dimensionless) amplitude coefficients at the transmitter side.

3.6.2 Detection Noise

Power spectral density (PSD) of noise at the output of the matched filter:

$$\text{PSD:} \quad \Phi_{n_d n_d}(f) = \frac{N_0}{x} \left|H_{\mathrm{E}}(f)\right|^2 = \frac{N_0}{x} \gamma^2 \left|G(f)\right|^2 \quad \text{with}$$

$$x = \begin{cases} 1 & \text{carrier modulated PAM in ECB domain} \\ 2 & \text{digital baseband transmission} \end{cases}$$

$$\text{ACF:} \quad \phi_{n_d n_d}(\tau) = \frac{N_0}{x} \gamma^2 \varphi_{gg}(\tau) \quad , \qquad \text{where } \left|G(f)\right|^2 \bullet\!\!-\!\!\circ \; \varphi_{gg}(\tau)$$

Noise samples from a discrete–time sequence $n_d[k] \stackrel{\text{def}}{=} n_d(kT)$

$$\text{Variance (noise power): } \sigma_n^2 = \phi_{n_d n_d}(0) = \frac{N_0}{x} \cdot \frac{E_\mathrm{g}}{E_\mathrm{g}^2} = \frac{N_0}{x E_\mathrm{g}} \qquad (3.2)$$

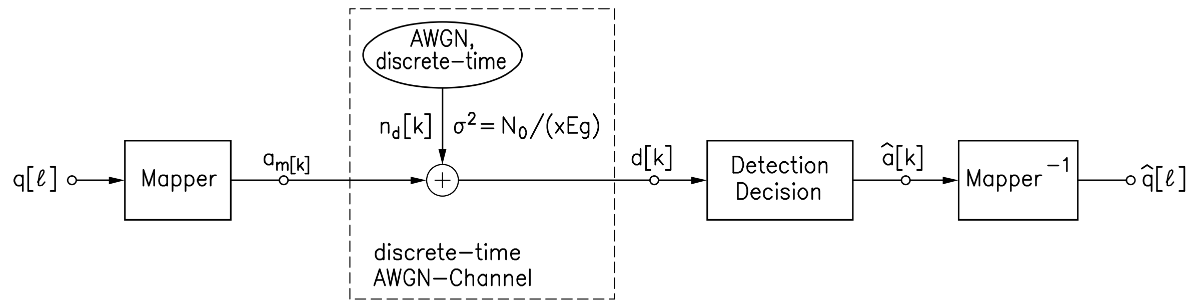
Since the signal pulse and the autocorrelation function of the noise at the output of a matched filter are proportional to each other, the following holds:

Theorem:

If the signal pulse at the output of a matched filter fulfills the first Nyquist criterion and the PAM fundamental pulse (or alternatively, the signal elements in general) is orthogonal with respect to shifted versions (shifted by multiples of the modulation interval), the noise samples $n_d[k] = n_d(kT + T_\mathrm{V}), k \in \mathbb{Z}$, are mutually uncorrelated, and therefore for Gaussian noise, all noise samples $n_d[k]$ are therefore also statistically independent!

$\Rightarrow$ The sequence of noise values $n_d[kT]$ (discrete–time noise signal) constitutes **discrete–time complex, white, Gaussian noise**. As long as the fundamental pulse $g(t)$ fulfills the temporal orthogonality condition $\int_{-\infty}^{+\infty} g(t + \lambda T)\ g^*(t)\ \mathrm{d}t = \delta[\lambda]\ E_\mathrm{g}$, the real and imaginary components of the noise are statistically independent and have variance $\frac{N_0}{2E_\mathrm{g}}$ (alternatively, zero variance for the imaginary component for digital baseband transmission). In this case, **each individual pulse** $a_{m[k]}\ g(t - kT)$ can be received **independently with respect to signal and noise from all other pulses.**

Discrete–Time Equivalent Block Diagram for ISI–Free PAM Transmission (without channel coding):



3.6.3 The Nyquist Criterion in the Frequency Domain

Interpretation of the temporal orthogonality condition

$$\int_{-\infty}^{+\infty} g(t+\lambda T)\, g^*(t)\, \mathrm{d}t = \delta[\lambda]\, E_\mathrm{g}, \quad \lambda \in \mathbb{Z}$$

as a convolution:

$$g(t) * g^*(-t)\Big|_{t=\lambda T} = \delta[\lambda]\, E_\mathrm{g}$$

combined with **ideal** sampling, which is equivalent to a multiplication with an impulse train $T \cdot \sum\limits_{\lambda=-\infty}^{+\infty} \delta(t-\lambda T)$:

$$T\left(\sum_{\lambda=-\infty}^{+\infty} \delta(t-\lambda T)\right) \cdot \big(g(t) * g^*(-t)\big) = T\delta(t)\, E_\mathrm{g}$$

Taking the Fourier transform, we obtain

$$\mathcal{F}\left\{T \cdot \sum_{\lambda=-\infty}^{+\infty} \delta(t-\lambda T)\right\} * \mathcal{F}\big\{g(t) * g^*(-t)\big\} = T\, E_\mathrm{g} \tag{3.3}$$

With $\mathcal{F}\left\{T \sum\limits_{\lambda=-\infty}^{+\infty} \delta(t-\lambda T)\right\} = \sum\limits_{\ell=-\infty}^{+\infty} \delta(f-\ell/T)$

and $\mathcal{F}\big\{g(t) * g^*(-t)\big\} = \mathcal{F}\big\{g(t)\big\} \cdot \mathcal{F}\big\{g^*(-t)\big\} = G(f) \cdot G^*(f) = \big|G(f)\big|^2$

we get $\left(\sum\limits_{\ell=-\infty}^{+\infty} \delta(f-\ell/T)\right) * \big|G(f)\big|^2 = T\, E_\mathrm{g}$

For the convolution, the distributive property holds:

$$\sum_{\ell=-\infty}^{+\infty} \Big(\big|G(f)\big|^2 * \delta(f-\ell/T)\Big) = T\, E_\mathrm{g}$$

The convolution with a δ–function corresponds to a delay.

$$\boxed{\sum_{\ell=-\infty}^{+\infty} \big|G(f-\ell/T)\big|^2 = T\, E_\mathrm{g}}$$

Theorem:

If a function $\varphi_{gg}(t) = g(t) * g^*(t)$ meets the first Nyquist Criterion, the sum of all possible ℓ/T-shifted versions of its spectrum $|G(f)|^2$ is a constant

Corollary:

A function, for which the superposition of all versions of its spectrum delayed by $\frac{\ell}{T}$, $\forall \ell \in \mathbb{Z}$, sums up to a constant, possesses equidistant zero crossings in time domain, which are located at $t = kT, \quad \forall k \in \mathbb{Z}\backslash\{0\}$.

Pulses $g(t)$ whose **autocorrelation** $\varphi_{gg}(\tau)$ exhibit this property are termed **"Root Nyquist–Pulses"** or, alternatively, square root Nyquist, or short $\sqrt{\text{Nyquist}}$–pulses. This characteristic allows the phase response of a $\sqrt{\text{Nyquist}}$–pulse to be freely chosen! (The phase response is equalized by the matched filter, which ultimately transforms the $\sqrt{\text{Nyquist}}$–pulse into a Nyquist pulse.)

For an optimal, ISI–free digital PAM transmission over an AWGN–channel, $\sqrt{\textbf{Nyquist}}$–pulses are temporally orthogonal.

Interpretation: Before detection, the detection signal $d(t)$ is sampled at the symbol clock rate $kT + T_{\text{V}}$. This sampling creates a periodic repetition in the frequency domain. Only if a constant value arises from this repetition in the frequency domain, the discrete–time sequence of amplitude coefficients is transmitted without distortion (i.e. without any dispersive effect). This means that $\tilde{d}[k] = a_{m[k]}$ holds. If the receiver input filter is matched to the fundamental $\sqrt{\text{Nyquist}}$–PAM pulse, then the maximum SNR is achieved for the discrete–time detection values $d[k]$.

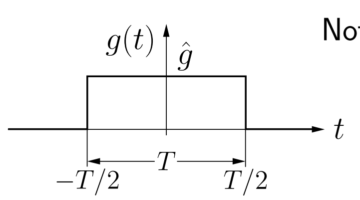
96

3.6.4 Examples of Nyquist–Functions

Example: *Examples of $\sqrt{\textit{Nyquist}}$ pulses and Nyquist Pulses:*

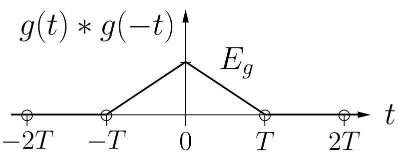
A. Rectangle in time domain (hard keying)

Pulse ($\sqrt{\text{Nyquist}}$): $g(t) = \hat{g}\ \mathrm{rect}\left(\dfrac{t}{T}\right)$; $\quad G(f) = \hat{g}\ T\ \mathrm{sinc}(\pi f T)\,;\quad E_{\text{g}} = \hat{g}^2\ T$

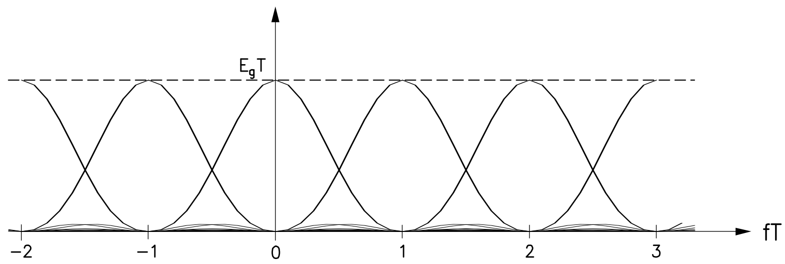


Note: $\mathrm{sinc}(x) = \begin{cases} \frac{\sin(x)}{x} & \text{for } x \neq 0 \\ 1 & \text{for } x = 0 \end{cases}$

autocorrelation (Nyquist–pulse) = detection pulse $g_d(t)$: Triangle



We have: $\hat{g}^2\, T^2 \sum_{\ell=-\infty}^{+\infty} \mathrm{sinc}^2\big(\pi(fT-\ell)\big) = \hat{g}^2\, T^2 = E_{\mathrm{g}}T, \quad \forall f$



Theorem:

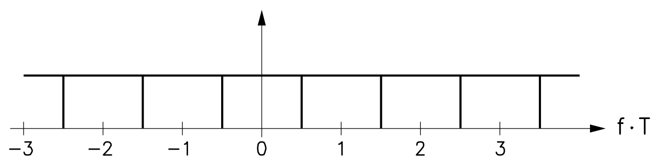
Generally: If a pulse $g(t)$ is time limited within $t \in \left[-\frac{T}{2}, \frac{T}{2}\right]$, i.e., $g(t) \neq 0$ only for $t \in \left[-\frac{T}{2}, \frac{T}{2}\right]$, the orthogonality condition is fulfilled and its autocorrelation (matched filter output signal) is limited to an interval of size $2T$. Therefore, $g(t)$ is a $\sqrt{\text{Nyquist}}$–pulse.

B. Rectangle in the frequency domain

Pulse ($\sqrt{\text{Nyquist}}$): $g(t) = \hat{g}\,\mathrm{sinc}(\pi t/T)\,, \qquad G(f) = \hat{g}\; T\;\mathrm{rect}(fT), \qquad E_{\mathrm{g}} = \hat{g}^2\; T$

Autocorrelation (Nyquist pulse): $g(t) * g(-t) = E_{\mathrm{g}}\,\mathrm{sinc}(\pi t/T) \sim g(t)$

we have $\hat{g}^2\; T^2 \sum_{\ell=-\infty}^{+\infty} \mathrm{rect}(fT-\ell) = E_{\mathrm{g}}\; T, \quad \forall f$



Theorem:

This pulse is strictly limited in the frequency domain. Therefore, the $\sqrt{\text{Nyquist}}$–pulse in time domain has infinite duration! The rectangle–pulse $\mathrm{rect}(fT)$ in the frequency domain, is the pulse with the smallest bandwidth for which inter-symbol interference–free transmission is possible! For smaller bandwidth, gaps between the periodic spectral repetitions are unavoidable, and therefor ISI is unavoidable!

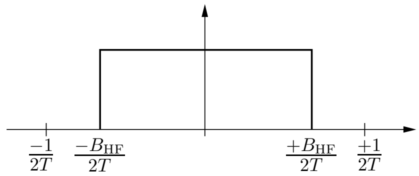
Corollary:

The smallest possible bandwidth of a digital PAM signal for which inter-symbol interference–free detection for transparent transmission is possible is

$$B_{\mathrm{RF}} \;=\; \frac{1}{T}$$

For smaller bandwidths, different source symbol sequences may result in the same transmit signals.

Example:

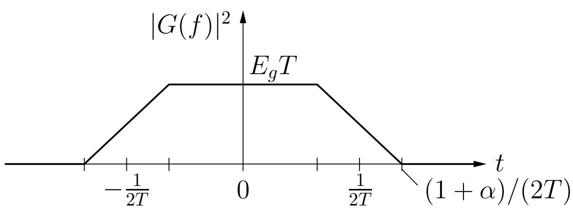


The two sequences

$$\begin{array}{ccccccc} \cdots & +1 & -1 & +1 & -1 & +1 & \cdots \\ \cdots & -1 & +1 & -1 & +1 & -1 & \cdots \end{array}$$

create periodic signals with frequency $\frac{1}{2T}$. However, since neither their fundamental frequency component nor their harmonics are allowed in the spectrum, the **same** transmit signal is generated for both sequence $s(t) \equiv 0$. The signals are therefore indistinguishable!

C. Nyquist pulse with trapezoidal shape in the frequency domain:



α: **roll–off–factor** or bandwidth excess factor $\alpha \in [0;1]$

$\alpha = 0$ Rectangle in frequency domain

$\alpha = 1$ Triangle in frequency domain

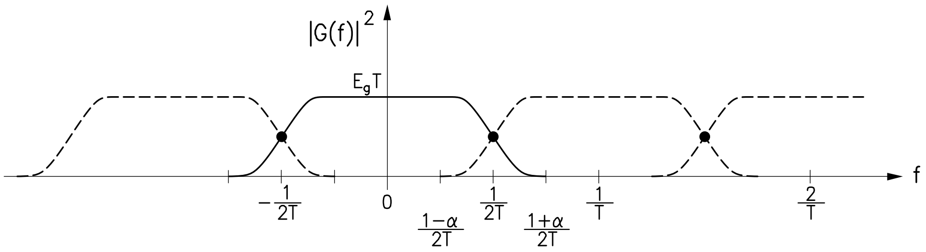
α: The relative extension of bandwith

$$B_{HF} = (1+\alpha)\frac{1}{T}$$

over its strict minimum $\frac{1}{T}$

In general: Functions with symmetric slope with respect to the points $\left(\pm\frac{1}{2T}; \frac{E_g T}{2}\right)$ are Nyquist functions.

D. Root-raised cosine pulses in the frequency domain:



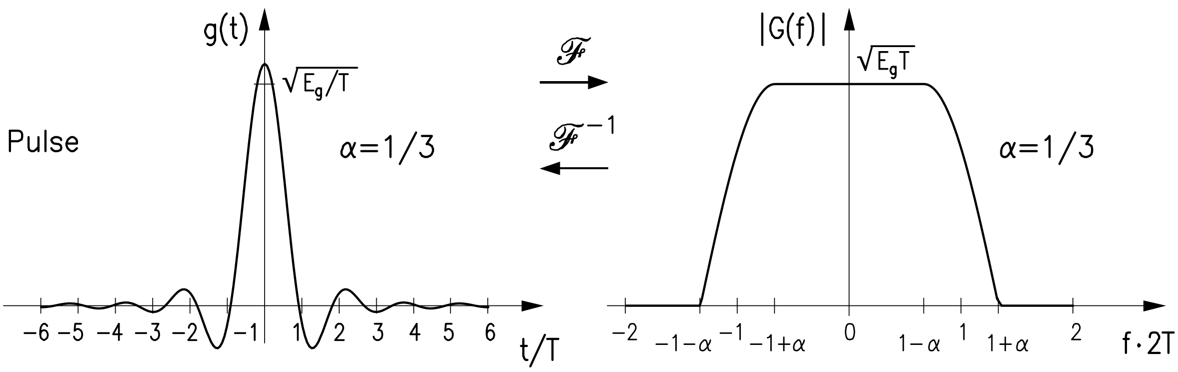
$$\left|G(f)\right|^2 = E_\mathrm{g}T \begin{cases} 1 & \text{for } |f| < \frac{1-\alpha}{2T} \\ \frac{1}{2}\Big[1-\sin\left(\frac{\pi T}{\alpha}\left(|f|-\frac{1}{2T}\right)\right)\Big] & \text{for } \frac{1-\alpha}{2T} \le |f| \le \frac{1+\alpha}{2T} \\ 0 & \text{for } |f| > \frac{1+\alpha}{2T} \end{cases}$$

Relative extension of the Nyquist interval $\left[-\frac{1}{2T};\frac{1}{2T}\right]$:
roll–off factor or bandwidth excess–factor $\alpha \in [0;1]$

In practice, this is a very commonly used Nyquist–/$\sqrt{\text{Nyquist}}$–function (approximated by digital signal processing)!

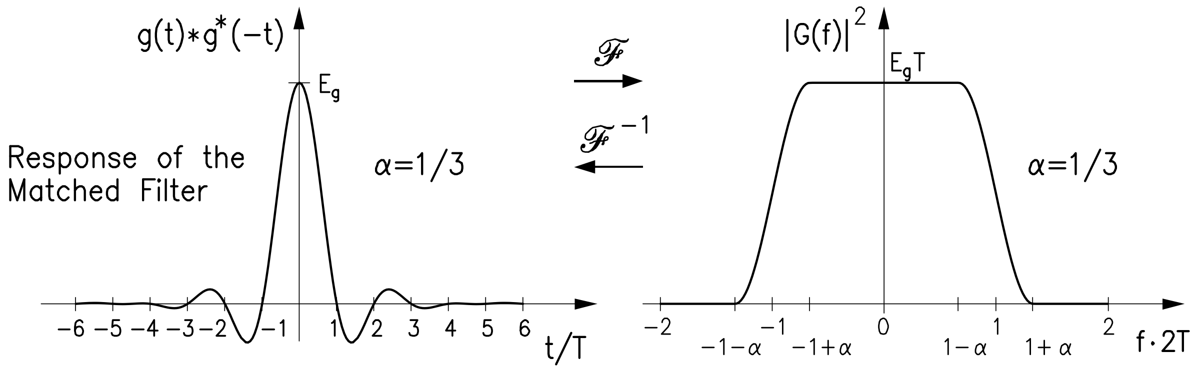
For root-raised cosine pulses with linear phase responses (symmetric $\sqrt{\mathrm{Nyquist}}$–pulse), the following holds:

$$\sqrt{\mathrm{Nyquist}}\text{–pulse:} \quad g(t) = \sqrt{\frac{E_\mathrm{g}}{T}}\;\frac{4\alpha t\cos\big(\pi(1+\alpha)t/T\big) + T\sin\big(\pi(1-\alpha)t/T\big)}{\pi t\big(1-(4\alpha t/T)^2\big)}$$



$\sqrt{\mathrm{Nyquist}}$–pulses have, in contrast to Nyquist–pulses in general, **no** zero crossings at kT, $k \in \mathbb{Z}\backslash\{0\}$ for $\alpha \neq 0$.

Nyquist pulse:



Bandwidth of a PAM–Signal using $\sqrt{\text{Nyquist}}$–functions with spectral roll–off slope:

– For carrier modulated transmission:

$$B_{\text{RF}} = \frac{1}{T}(1+\alpha)\ ; \quad \Gamma_{\text{d}} = \frac{\text{ld}(M)}{1+\alpha} = \frac{R}{1+\alpha}$$

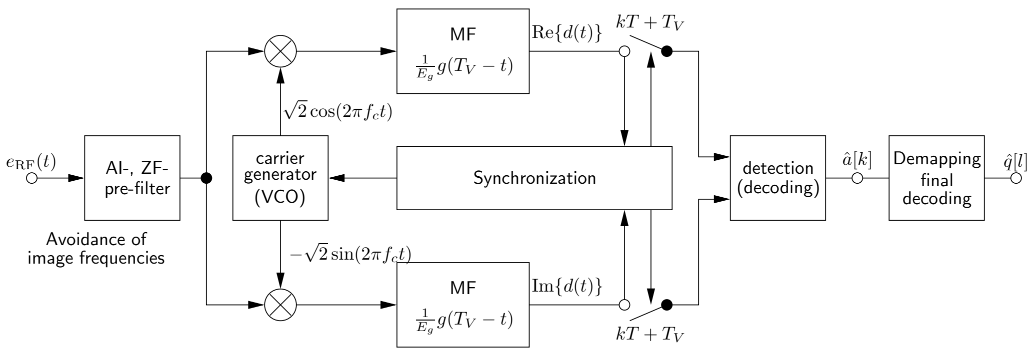
– For digital baseband transmission:

$$B_{\text{RF}} = \frac{1}{2\,T}(1+\alpha)\ ; \quad \Gamma_{\text{d}} = \frac{2\,R}{1+\alpha} = \frac{2\,\text{ld}(M)}{1+\alpha}$$

104

3.6.5 Eye Pattern

Block diagram for a carrier–modulated RF–PAM–transmission system for real signals with a real fundamental pulse $g(t)$:



MF: Matched filter for fundamental pulse $g(t)$

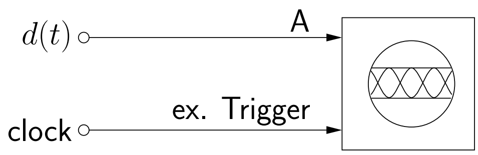
Visualization of the detection signal $d(t)$, i.e., $\mathrm{Re}\{d(t)\}$ and $\mathrm{Im}\{d(t)\}$:

The depiction of **all** sample functions of a cyclo–stationary stochastic process in one **single** diagram creates a periodic pattern with period T, the so–called

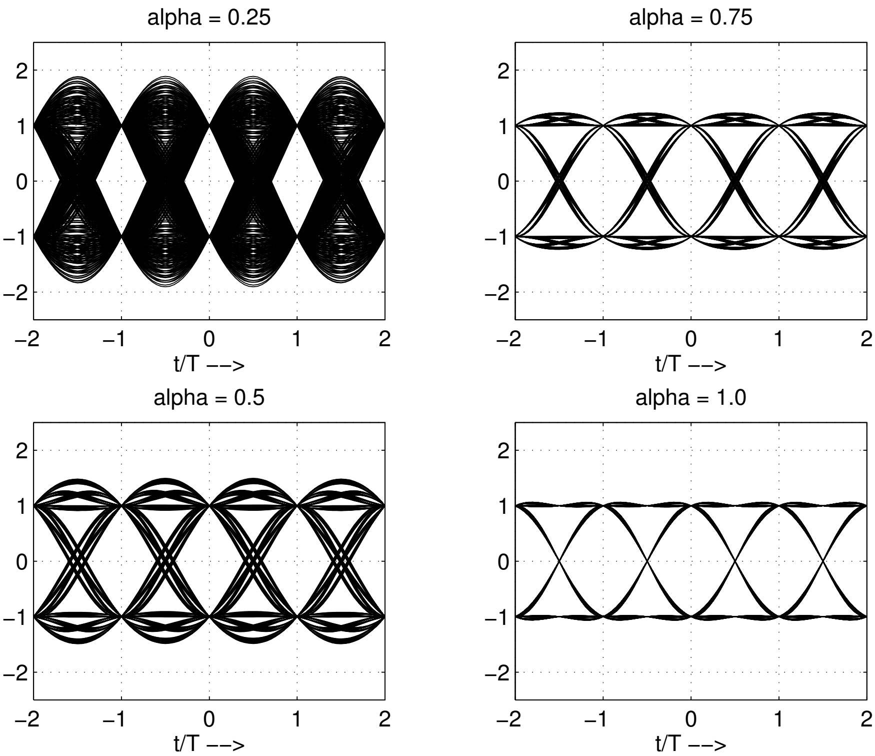
Eye Pattern.

Here: Simultaneous depiction of the detection signals of all possible sequences of amplitude coefficients.

Display of the eye pattern on an oscilloscope: Trigger with the symbol clock, memory oscilloscope or oscilloscope with slow intensity decay



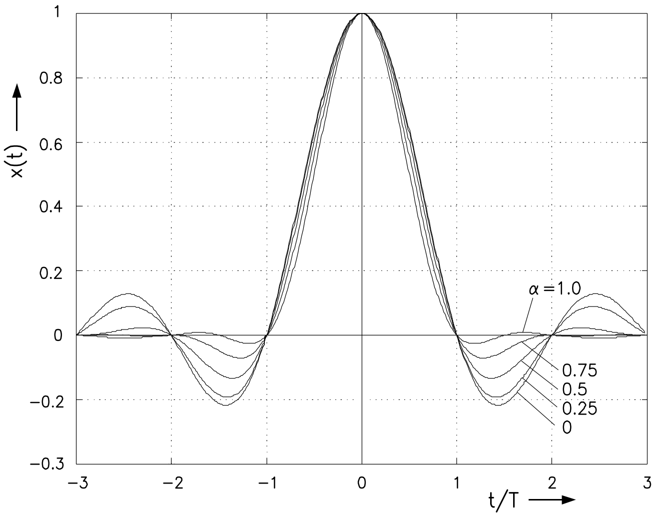
Eye pattern for binary ASK transmission with root-raised cosine pulses with roll–off factor α



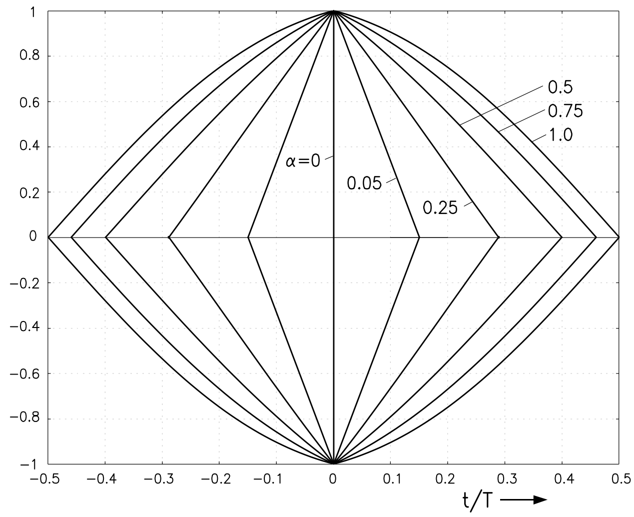
Horizontal eye–opening:

Resistance to "Jitter" in the phase of the receiver's symbol clock

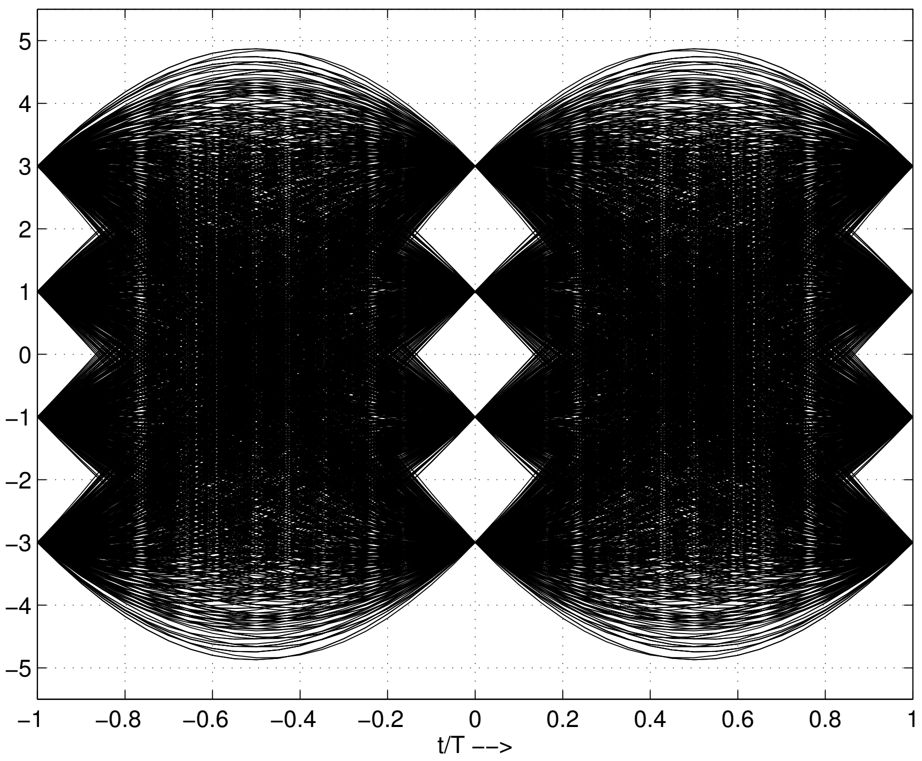
Raised cosine pulses for various roll–off factors α



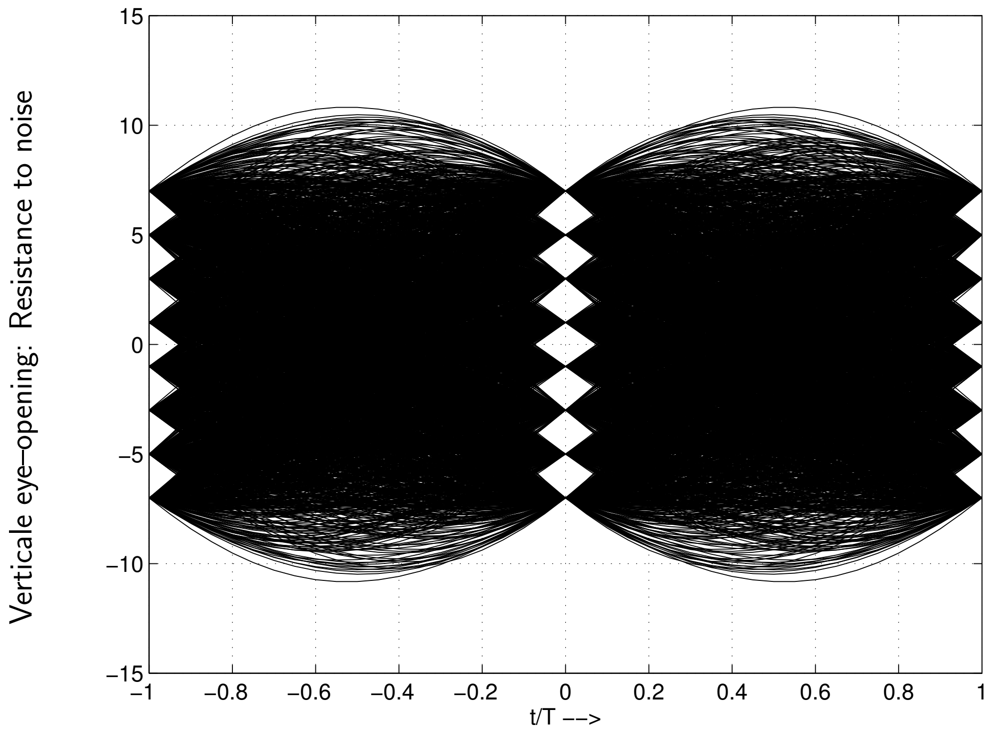
Inner eye rim for binary ASK transmission with raised cosine pulses



Eye pattern for quaternary ASK transmission in the baseband (or 16QAM) using raised cosine pulses with roll–off factor; $\alpha = 0.33$



Eye pattern for 8–ASK transmission (or 64QAM) using raised cosine pulse with roll–off factor; $\alpha = 0.33$



3.7 Detection and Decoding for digital PAM–Transmission

3.7.1 Complete Signal Representation

Theorem:

If the symbol timing synchronization is ideal, i.e., $\Delta\hat{T} = \Delta T$ for PAM, the continuous time received signal $r(t)$ is represented **without information loss** with respect to the data sequence $a_{m[k]}$ by the sequence of samples $d[k] = d(kT + T_{\mathrm{V}})$ at the output of the matched filter after the receiver front end. In this case, one (real or complex valued) sample per modulation step is sufficient to extract all information contained about the transmitted data sequence in the receiver input signal . This holds although the **sampling theorem** is **not fulfilled** in general (bandwidth excess factor $\alpha > 0$ for horizontal eye–opening!).

Note:
A perfect reconstruction of the continuous time receiver input signal is therefore generally not possible. For imperfect symbol clock synchronization, also information with regard to $a_{m[k]}$ is lost (i.e., a loss in power efficiency results).

Proof:

Description of the continuous–time receiver input signal $r(t)$ using a sum of D **orthogonal basis functions** $g_i(t-kT)$ (signal expansion by means of orthogonal functions as usual):
The number D of dimension per modulation interval has to be chosen sufficiently large.

$$r(t) = \sum_{k=-\infty}^{+\infty} \sum_{i=0}^{D-1} d_i[k] g_i(t-kT) \tag{3.4}$$

with **double orthogonality** with respect to k and i:

$$\int_{-\infty}^{+\infty} g_i(t+kT)\, g_l^*(t)\,\mathrm{d}t = \left\{ \begin{array}{ll} E_\mathrm{g} & \text{for } k=0 \text{ and } i=l \\ 0 & \text{for } k\neq 0 \text{ or } i\neq l \end{array} \right\} = E_\mathrm{g}\cdot\delta[k]\cdot\delta[i-l]. \tag{3.5}$$

By choosing an adequate number of basis functions, the continuous–time receiver input signal $r(t)$ (desired signal + noise) can be represented with arbitrarily small error.

Computation of the components $d_l[\nu]$ of the signal representation:

$$d_l[\nu] = \frac{1}{E_\mathrm{g}} \int_{-\infty}^{+\infty} e(t)\cdot g_l^*(t-\nu T)\,\mathrm{d}t \quad ; \qquad \nu\in\mathbb{Z}$$

Proof:
Substitution of (3.4) for $e(t)$: All terms for which $i\neq l$ and $k\neq\nu$ vanish due to the double orthogonality, while for $i=l$ and $k=\nu$, the integral yields E_g.

Special choice for PAM:

$$g_0(t) \stackrel{!}{=} g(t)$$

$$\boxed{d_0[k] = d[k] = \text{sequence of samples at the matched filter output!}}$$

(Desired) Signal:

$$\tilde{r}(t) = \sum_{k=-\infty}^{+\infty} a_{m[k]}\, g(t-kT)$$

$$\tilde{d}_0[k] = a_{m[k]} = \tilde{d}[k] \quad \text{desired part of the samples at the matched filter output}$$

$$\tilde{d}_i[k] = 0 \quad \forall k\in\mathbb{Z}, i\in\{1,2,\ldots,D-1\} \quad \text{(orthogonality)}$$

Additive <u>White</u> Gaussian Noise (AWGN) $n(t)$

$n_i[k]$: Mutually independent Gaussian random variables with variance

$$\sigma_{n,\ell}^2 = \sigma_n^2 = \frac{N_0}{E_\mathrm{g}} = \sigma_I^2+\sigma_Q^2; \qquad \sigma_I^2=\sigma_Q^2=\frac{N_0}{2E_\mathrm{g}} \quad \forall k\in\mathbb{Z}; \qquad \ell\in\{0,1,2,\ldots,D-1\}$$

Derivation analogous to Section 3.6.1, see (3.2).

Specifically $g_0(t) = g(t)$: $n_0[k] = n_d[k]$ noise components of the matched filter output

Complete representation of desired signal + noise:

i	Components
0	$d_0[k] = a_{m[k]} + n_d[k]$
1	$0 + n_1[k]$
2	$0 + n_2[k]$
$\vdots$	$\vdots$

Note: For a complete representation of white noise, an infinite number of dimensions, (i.e., $D \to \infty$), would be required (*infinite bandwidth* implies an *infinite number of dimensions*!)

Theorem of irrelevant data (Information Theory):

All parameters of a signal that have no causal relationship with the information which it represents, are useless for the extraction of this information from the signal.

$\Rightarrow$ Noise components $n_1[k]$, $n_2[k]$, ... are irrelevant. This means that they cannot help for detection or decoding.

$\Rightarrow$ The sequence $d[k]$ of samples at the output of the matched filter constitute a **sufficient statistics** of the receiver input signal with respect to data detection!

■

3.7.2 Signal Space Representation of PAM

Since the PAM fundamental pulse $g(t)$ meets the orthogonality condition (see Section 2.1.2.1) or Eq. (3.1), a signal space with one **dimension** for each modulation step is spanned: A specific **sequence** of amplitude coefficients $a_{m[k]}, k \in \{0, 1, \dots, N-1\}$, of length N can be equivalently represented by one **vector** $\vec{a} = (a_{m[0]}, a_{m[1]}, \dots, a_{m[N-1]})$ in an N dimensional space $\mathbb{R}^N$ (digital baseband transmission) or $\mathbb{C}^N$ (ECB–domain, i.e. $2N$ real dimensions). To the transmitted vector, an N–dimensional ($2N$-dimensional) Gaussian noise vector $\vec{n}$ with an isotropic probability density function (spherical N-dimensional pdf) is added. For $\vec{n}$ we have:

Baseband transmission: $$f_{\vec{n}}(n_0, n_2, \dots, n_{N-1}) = \left(\frac{1}{\sqrt{2\pi}\sigma_n}\right)^N \mathrm{e}^{-\frac{1}{2\sigma_n^2}\sum_{k=0}^{N-1} n_k^2}; \quad n_k \in \mathbb{R}; \quad \sigma_n^2 = N_0/(2E_\mathrm{g})$$

ECB–domain: $$f_{\vec{n}}(n_0, n_2, \dots, n_{N-1}) = \left(\frac{1}{\pi\sigma_n^2}\right)^N \mathrm{e}^{-\frac{1}{\sigma_n^2}\sum_{k=0}^{N-1} |n_k|^2}; \quad n_k \in \mathbb{C}; \quad \sigma_n^2 = N_0/E_\mathrm{g};$$
$$\sigma_I^2 = \sigma_Q^2 = N_0/(2E_\mathrm{g})$$

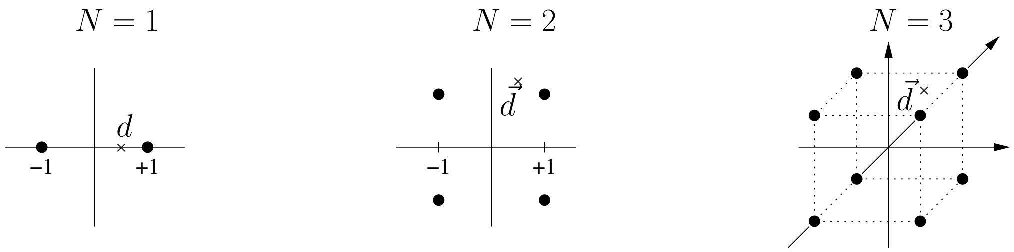
The sequence $d[k] = d(kT + T_\mathrm{V})$, $k = 0, 1, \dots, N-1$, of samples at the matched filter output is thus equivalently represented by vector $\vec{d} = (d[0], d[1], \dots, d[N-1])$ in the signal space $\mathbb{R}^N$ or $\mathbb{C}^N$.

Signal Space Representation Without Application of Channel Coding

In N real or N complex (i.e., $2N$ real) dimensions there are M^N signal points which represent all possible sequences of N amplitude coefficients.

Example:

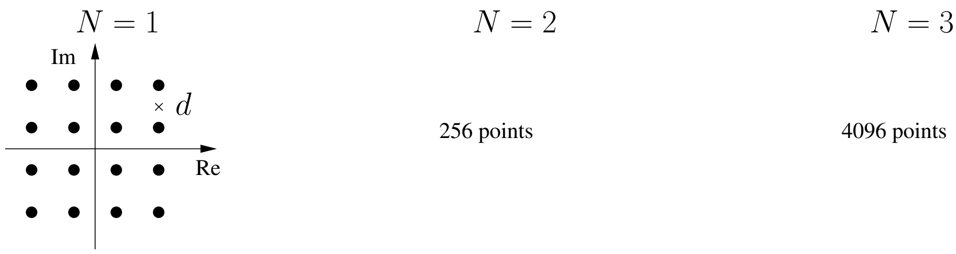
a) 2ASK = 2PSK = BPSK (without channel coding) $\mathcal{A} = \{-1, +1\}$



ect.

116

b) 16QAM (without channel coding)



ect.

Signal Space Representation if Channel Coding is Applied

Block code of length n with code symbols from an M_c–ary alphabet: one codeword is mapped to N amplitude coefficients or, alternatively, onto N modulation steps with:

$$M^N = M_\mathrm{c}^n; N = n \log_M(M_\mathrm{c}) = n\frac{\mathrm{ld}M_\mathrm{c}}{\mathrm{ld}M} \tag{3.6}$$

(Without loss of generality, we assume $n\,\frac{\mathrm{ld}(M_\mathrm{c})}{\mathrm{ld}(M)}$ to be an integer)

From the M^N possible signal points, only

$$2^k = 2^{nR_\mathrm{c}} = 2^{NR} \qquad \text{with} \qquad R = \frac{\mathrm{ld}(M)}{\mathrm{ld}(M_\mathrm{c})} \cdot R_\mathrm{c} < \mathrm{ld}(M)$$

codewords are used for transmission. This means that for each (residual) signal point in the N–dimensional signal space,

$$2^{N(\mathrm{ld}(M)-R)} \qquad \text{signal points}$$

are discarded: Thinning out of the signal space

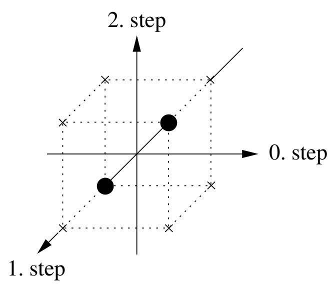
$\Rightarrow$ Probability of a erroneous detection of the transmitted codeword is reduced!

Example: *Simplest possible case*

Binary repetition code of length 3, $n=3$, $k=1$, $M_\mathrm{c}=2$, $R_\mathrm{c}=\frac{1}{3}$

$$\mathbf{C} = \{(000),(111)\}$$

2ASK with mapping $a = 2c-1$, $(0 \to -1\,,\, 1 \to +1)$, $N=3$, $R=\frac{1}{3}$



only two signal vectors remain out of eight possible single vectors in the three–dimensional signal space.

3.7.3 Maximum A–Posteriori Decision Rule

The decision over a sequence of amplitude coefficients $\hat{a}[k]$, $k = 0,1,\ldots,N-1$, is considered, e.g. a codeword of a block code over N modulation steps.

Notation as row vectors: $(a_{m[0]}, a_{m[1]}, \ldots, a_{m[N-1]}) = \vec{a}$; $(d[0], d[1], \ldots, d[N-1]) = \vec{d}$; etc.

Optimal decision rule:

Decide in favor of that sequence $\vec{a}$ for which the probability **after** (= a–posteriori (latin)) observation of the received signal $\vec{d}$ is the maximum:

Maximum A–Posteriori (MAP) Rule $$\hat{\vec{a}} = \underset{\mathcal{M}^{-1}\{\vec{a}\}\in\mathbf{C}}{\arg\max} \quad \Pr\left(\vec{a}\,|\,\vec{d}\right)$$

Here, all 2^{NR} possible vectors $\vec{a}$ contained in the code book after the mapping $\mathcal{M}\{\cdot\}$ (i.e. $\vec{a} = \mathcal{M}\{\vec{c}\}$ with $\vec{c} \in \mathbf{C}$) have to be considered.

Bayes Rule: $$\Pr\left(\vec{a}\,|\,\vec{d}\right) = f_{\vec{d}}\left(\vec{d}\,|\,\vec{a}\right) \cdot \frac{\Pr\left(\vec{a}\right)}{f_{\vec{d}}(\,\vec{d}\,)}$$

with $f_{\vec{d}}(\vec{d})$: N– (respectively $2N$–) dimensional probability density function (pdf) for the observed sequence $d[k]$.

For the AWGN channel, from $d[k] = a_{m[k]} + n_d[k]$ it follows that:

$$f_{\vec{d}}\left(\vec{d}\,|\,\vec{a}\right) = \left(\frac{1}{\pi\sigma_n^2}\right)^N \mathrm{e}^{-\frac{|\vec{d}-\vec{a}|^2}{\sigma_n^2}} \qquad \text{(ECB–domain)}$$

N– ($2N$–) dimensional Gaussian pdf for noise vector $\vec{n}_d[k]$

1. Constant denominator $f_{\vec{d}}(\vec{d})$ is irrelevant for max.–search $\Rightarrow$ can be omitted
2. Constant factor $(1/\pi\sigma_n^2)^N$ with constant σ_n^2 is irrelevant for max.–search $\Rightarrow$ can be omitted
3. $-\ln(\cdot)$ is a strictly monotonically decreasing function. For $a > b$, $-\ln(a) < -\ln(b)$ holds

Equivalent MAP rule:

$$\boxed{\hat{\vec{a}} = \underset{\mathcal{M}^{-1}\{\vec{a}\}\in\mathbf{C}}{\arg\min} \quad \frac{1}{\sigma_n^2}\left|\vec{d}-\vec{a}\right|^2 - \ln\left(\Pr\left(\vec{a}\right)\right)}$$

3.7.4 Maximum–Likelihood Decision Rule

If all sequences $\vec{a}$ with $\mathcal{M}^{-1}\{\vec{a}\} \in \mathbf{C}$ are a–priori equally probable, we have

$$\Pr(\vec{a}) = \frac{1}{2^{NR}} \quad .$$

Then, the

Maximum–Likelihood (ML–) Rule: $\hat{\vec{a}} = \underset{\mathcal{M}^{-1}\{\vec{a}\}\in\mathbf{C}}{\arg\min} \left|\vec{d} - \vec{a}\right|^2$

is equivalent to the MAP rule.

Due to the temporal orthogonality of the fundamental pulse form, the following holds:

$$\left|\vec{d} - \vec{a}\right|^2 = \sum_{k=0}^{N-1} \left|d[k] - a_{m[k]}\right|^2$$

ML–Rule: Decision in favor of the sequence of amplitude coefficients $\hat{\vec{a}}$, which have the shortest Euclidean distance to the received signal point $\vec{d}$ in the N– ($2N$–) dimensional signal. The squared Euclidean distance between the detected point $\vec{d}$ and a signal point $\vec{a}$ is known as the

metric for the signal point $\vec{a}$

or the metric of the corresponding codeword.

Definition: *Metric with respect to a codeword* $\vec{c}$

$$\Lambda(\vec{c}) \overset{\text{def}}{=} c_1 \ln\Big(f_{\vec{d}}(\vec{d}\,|\,\vec{a}\,)\Big) + c_2 \qquad \text{with } \vec{a} = \mathcal{M}\{\vec{c}\}, \quad \vec{c} \in \mathbf{C}$$

with appropriate constants c_1 and c_2

here: $\Lambda(\vec{c}) = \left|\vec{d} - \vec{a}\right|^2$ with $\vec{a} = \mathcal{M}(\vec{c})$

It is also important to note that the metric for a **sequence** of **symbols** (here: sequence of amplitude coefficients) and thus for a codeword can be defined as the **sum of squared Euclidean distances** for the individual modulation steps.

Definition: *Metric with respect to a signal element in step* k

$$\lambda_i[k] = c_1 \ln f_d(d[k]|a_i) + c_2, \qquad i \in \{0,1,\ldots,\ M-1\},$$

$$\textbf{here}: \quad \lambda_i[k] = |d[k] - a_i|^2$$

We have

$$\Lambda(\vec{c}) = \sum_{k=0}^{N-1} \lambda_{m[k]}[k] \qquad \text{with } \mathcal{M}^{-1}(a_{m[0]}, a_{m[1]}, \ldots,\ a_{m[N-1]}) = \vec{c}$$

This **additivity of the metric** facilitates the implementation of low-complexity decoding algorithms **greatly**.

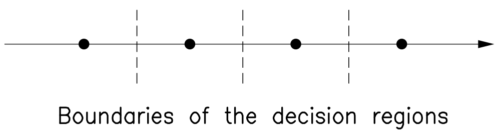
$\Rightarrow$ Subdivision of the N– $(2N$–) dimensional signal space into 2^{NR} **decision regions** with respect to the smallest Euclidean distance.

For decisions according to the ML rule, the decision regions are demarcated using the **perpendicular bisecting planes** between each pair of signal points.

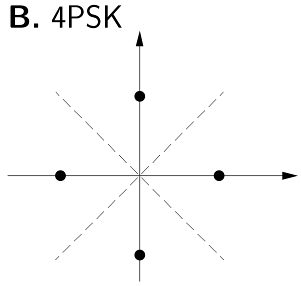
Example:

$N = 1$, symbol by symbol ML decision (ML detection) is possible

A. Baseband transmission or ASK

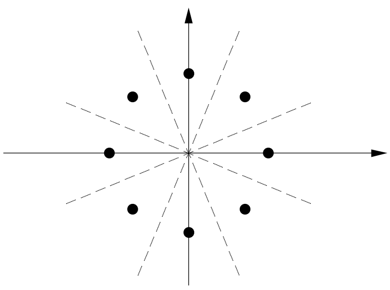


B. 4PSK

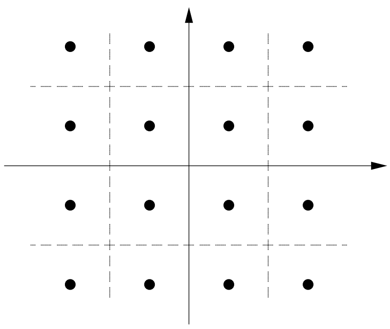


124

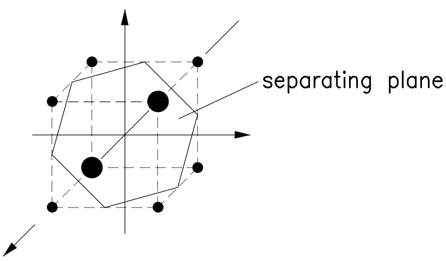
C. 8PSK



D. 16QAM



E. $N = 3$, 2PSK, ML decision for the $(3{,}1)$ repetition code (ML decoding)



Distance between points: $2\sqrt{3}$

Distance to separating plane: $\sqrt{3}$

For PAM transmission **without channel coding** ($R = \mathrm{ld}(M)$), **symbol by symbol** detection is possible due to the independence between the desired signal and noise in a 1–(2–) dimensional subspace and the orthogonality of signal subspaces for different modulation steps:

Maximum–Likelihood Detection

$$\hat{a}[k] = \arg\min_{a \in \mathcal{A}} |d[k] - a|^2$$

For PAM transmission **with channel coding** ($R < \mathrm{ld}(M)$), the ML decision has to be performed in N– (or $2N$–) dimensions by comparing the distances between all 2^{NR} codewords representating the vectors $\vec{a} \in \mathbf{C}$ and the received signal point.

Maximum–Likelihood Decoding

$$\hat{\vec{a}}[k] = \arg\min_{\mathcal{M}^{-1}\{\vec{a}\} \in \mathcal{C}} |\vec{d} - \vec{a}|^2$$

Due to the extremely large number 2^{NR} of codewords for technically interesting codeword lengths N, ML decoding through a comparison of the metrics of all codewords is not implementable:

$\Rightarrow$ 1. Efficient decoding algorithms

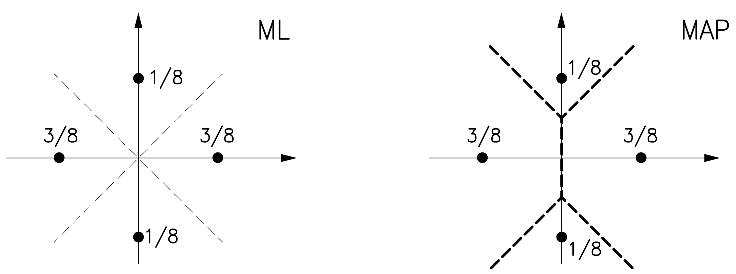
$\Rightarrow$ 2. Suboptimal coding schemes

Connection between MAP– and ML–Detection

When not all signal points have equal a–priori probabilities, the **ML** decision rule may be converted to the **MAP** decision rule by shifting the decision boundaries ($N = 1$) or, alternatively, the decision planes ($N > 1$) between two signal points $\vec{b}$ and $\vec{c}$ by $\frac{\sigma_n^2}{2} \ln\left(\frac{\Pr(\vec{b})}{\Pr(\vec{c})}\right)$ in direction of the signal point with lower a–priori probability.

Example:

4PSK, $N = 1$



3.8 Euclidean Distance

Given: PAM–signals for two different sequences of signal elements $a_{m[k]}$ and $a_{\ell[k]}$

Energy of the difference of corresponding PAM signals:

$$\begin{aligned}\int_{-\infty}^{+\infty}\left|\left(\sum_{k=-\infty}^{+\infty} a_{m[k]}g(t-kT)\right)-\left(\sum_{k=-\infty}^{+\infty} a_{\ell[k]}g(t-kT)\right)\right|^2 \mathrm{d}t &= \int_{-\infty}^{+\infty}\left|\left(\sum_{k=-\infty}^{+\infty}\left(a_{m[k]}-a_{\ell[k]}\right) g(t-kT)\right)\right|^2 \mathrm{d}t \\ &= \sum_{k=-\infty}^{+\infty}\left|a_{\ell[k]}-a_{m[k]}\right|^2 \int_{-\infty}^{+\infty}|g(t-kT)|^2\,\mathrm{d}t \\ &= \sum_{k=-\infty}^{+\infty}\left|a_{\ell[k]}-a_{m[k]}\right|^2 \cdot E_{\mathrm{g}}, \qquad (3.7)\end{aligned}$$

where we explored in (3.7) the temporal orthogonality of the signal elements to discard all mixed terms.

The squared Euclidean distance between two signal points in the signal space is proportional to the energy of the difference of the two signals (Parseval's theorem!).

Note: This is also true for the Euclidean distance between receiver input signal and hypotheses $\vec{a}$ of the transmitted sequences. The ML receiver: Decide in favour of the transmit signal for which the difference to the receiver input signal has minimum energy.

- $\Rightarrow$ **The larger the energy of the difference of two transmit signals is, the easier it is to differentiate between them.**
- $\Rightarrow$ The larger the **minimum squared Euclidean distance** between two signal points is, the more power efficient the corresponding digital communications scheme is.

Definition: *Normalized minimum squared Euclidean distance of a digital PAM communication scheme (NMQED*

$$\boxed{d_{\min}^2 \overset{\text{def}}{=} \frac{E_{\mathrm{g}}}{2E_{\mathrm{b}}} \min_{\vec{m}\neq\vec{l}} \sum_{k=-\infty}^{+\infty} \left|a_{m[k]}-a_{l[k]}\right|^2}$$

Note: For historical reasons, the minimum squared Euclidean distance is normalized with respect to twice the average signal energy per bit of information.

Without channel coding:

$$\boxed{d_{\min}^2 \stackrel{\text{def}}{=} \frac{E_\text{g}}{2E_\text{b}} \min_{m \neq l} \; |a_m - a_l|^2 \quad ; \qquad \forall\, m,l \in \{0,1,\ldots,M-1\}}$$

With channel coding:

$$\boxed{d_{\min}^2 \stackrel{\text{def}}{=} \frac{E_\text{g}}{2E_\text{b}} \min_{\vec{m} \neq \vec{l}} \sum_{k=0}^{N-1} \left|a_{m[k]} - a_{l[k]}\right|^2 \quad ; \quad \mathcal{M}^{-1}\{\vec{m}\},\, \mathcal{M}^{-1}\{\vec{l}\} \,\in\, \mathbf{C}}$$

Example: *(without channel coding)*

- **ASK bipolar:** $E_\text{b} = E_\text{g} \cdot \dfrac{M^2-1}{3} / \text{ld}(M) \quad \min\limits_{m \neq l} |a_m - a_l|^2 = 4$

$$\boxed{d_{\min}^2 = \frac{6\,\text{ld}(M)}{M^2-1}}$$

M	2	4	8	16
$d_{\min}^2$	2	0.80	0.286	0.0941

- **PSK:** $E_\text{b} = E_\text{g} / \text{ld}(M)$

$$\begin{aligned}\min_{m \neq l} |a_m - a_l|^2 \;&=\; |\text{e}^{\text{j}2\pi/M} - 1|^2 = 1 - 2\,\text{Re}\left\{\text{e}^{\text{j}2\pi/M}\right\} + 1 \\ &=\; 2 - 2\cos\left(\frac{2\pi}{M}\right) = 4 \cdot \sin^2\left(\frac{\pi}{M}\right)\end{aligned}$$

$$\boxed{d_{\min}^2 = 2\,\text{ld}(M) \cdot \sin^2\left(\frac{\pi}{M}\right)}$$

M	2	4	8	16
$d_{\min}^2$	2	2	0.878	0.304

- **QAM: Quadratic constellations**

$$E_\text{b} = 2E_\text{g} \cdot \frac{M-1}{3} / \text{ld}(M)$$

$$\boxed{d_{\min}^2 = \frac{3\,\text{ld}(M)}{M-1}}$$

M	4	16		64		256
$d_{\min}^2$	2	0.8		0.286		0.0941
			$\uparrow$		$\uparrow$	
▪ **Cross-QAM**			32		128	
			0.5		0.171	

3.9 Error Probability for Coherent Demodulation

3.9.1 ML Detection (without Channel Coding)

Decision errors occur symbolwise when using ML–detection (no channel coding!) if the received signal $d[k]$ falls outside the decision region $\mathcal{G}_m$ to which the transmitted signal point a_m belongs.

$$\Pr(\hat{m}[k] \neq m[k]) = \int\limits_{d \notin \mathcal{G}_{m[k]}} f_{\mathbf{d}}(d|a_{m[k]})\,\mathrm{d}d$$

with $f_n(n) = \dfrac{1}{\pi\sigma_n^2} \exp\left(-|n|^2 \Big/ \sigma_n^2\right)$ for ECB–signals

$$= \left(\frac{1}{\sqrt{2\pi}\sigma_I}\right)^2 \exp\left(-\left((\mathrm{Re}\{n\})^2 + (\mathrm{Im}\{n\})^2\right) \Big/ (2\sigma_I^2)\right)$$

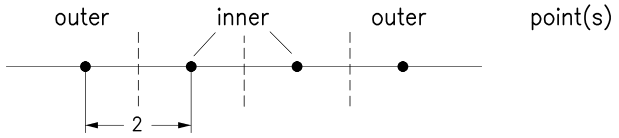
$f_n(n) = \dfrac{1}{\sqrt{2\pi}\sigma_I} \exp\left(-n^2 \Big/ (2\sigma_I^2)\right)$ for real signals

$$\sigma_n^2 = \frac{N_0}{E_\mathrm{g}} \quad ; \quad \sigma_I^2 = \frac{N_0}{2E_\mathrm{g}} = \sigma_n^2/2$$

For **equally probable** signal elements the **average symbol error probability (Symbol Error Ratio) SER** can be defined as

$$\mathsf{SER} = \sum_{m=0}^{M-1} \frac{1}{M} \, \Pr(\hat{m} \neq m)$$

3.9.1.1 Bipolar ASK



$D = 1$, only real part

"Inner points": $\Pr(\hat{m} \neq m) = \int_{-\infty}^{a_m - 1} f_n(d - a_m)\,\mathrm{d}d + \int_{a_m+1}^{+\infty} f_n(d - a_m)\,\mathrm{d}d$

$$= 2 \int_1^{\infty} \frac{1}{\sqrt{2\pi}\sigma_I} \mathrm{e}^{-x^2/(2\sigma_I^2)}\,\mathrm{d}x = 2\,Q\left(\frac{1}{\sigma_I}\right) = 2\,Q\left(\sqrt{\frac{2E_\mathrm{g}}{N_0}}\right)$$

with $\boxed{Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-z^2/2}\,\mathrm{d}z}$ Complementary Gaussian Error Integral

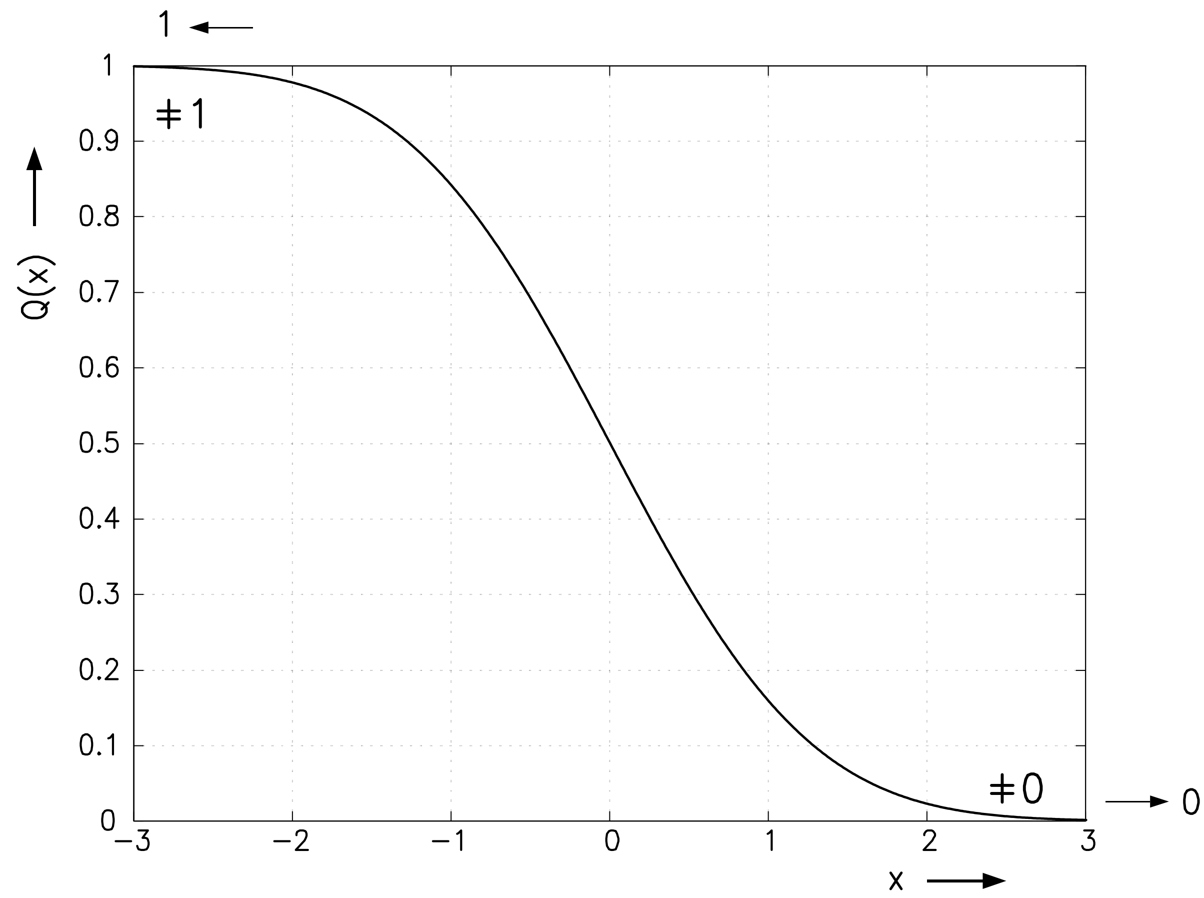
$$Q(x) = 1 - \Phi(x) \quad \text{with} \quad \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{z^2}{2}}\,\mathrm{d}z \quad \text{Gaussian Error Integral}$$

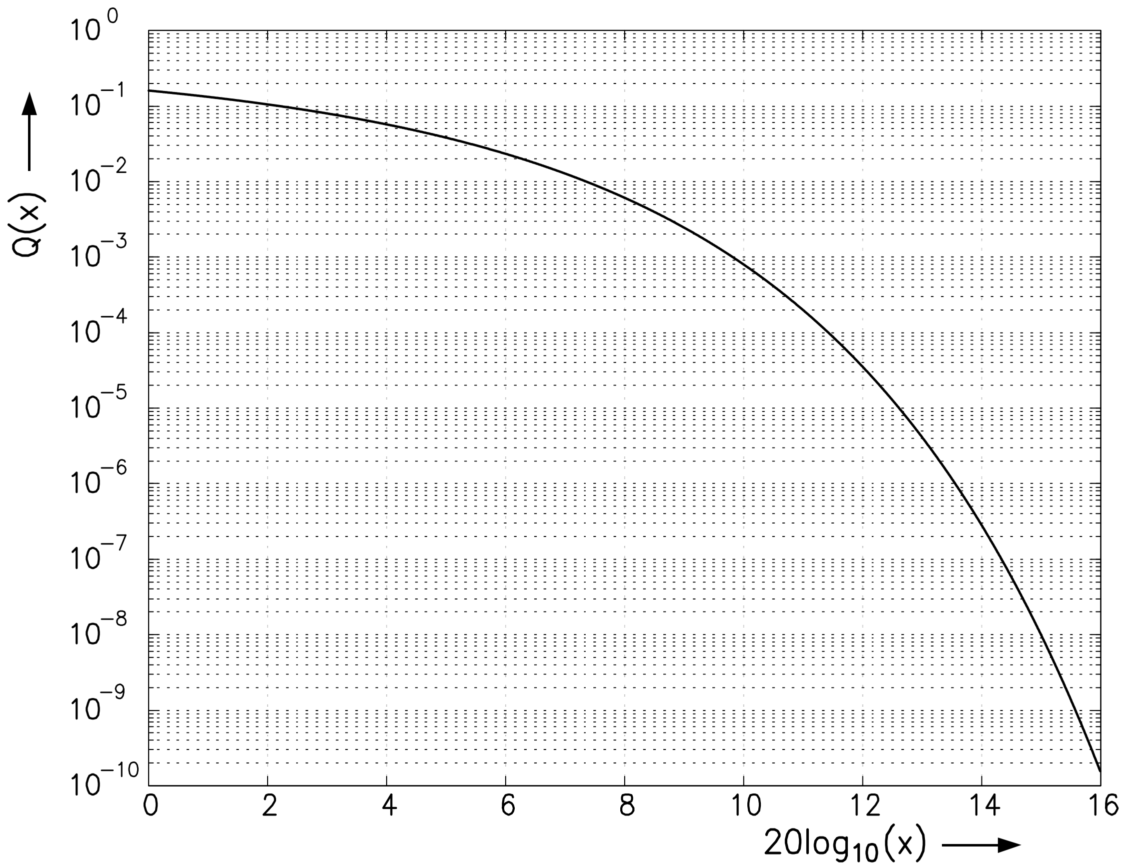
Note for $x \geq 0$: $\mathrm{erf}(x) \mathrel{\widehat{=}} \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,\mathrm{d}t$ $\quad \boxed{\mathrm{erf}(x) = 2\left(\Phi(\sqrt{2}x)\right) - 1}$

$$\boxed{\Phi(x) = \frac{1}{2}\,\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right) + \frac{1}{2}}$$

$$\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$$

$$\boxed{\mathrm{erfc}(x) = 2\,Q\left(\sqrt{2}x\right); \quad Q(x) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right)}$$


1
≑1
Q(x)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
≑0
0
−3
−2
−1
0
1
2
3
x


Q(x)
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-10}
0
2
4
6
8
10
12
14
16
$20\log_{10}(x)$

"outer points": $\Pr(\hat{m} \neq m) = 1 \cdot Q\left(\sqrt{2\frac{E_\mathrm{g}}{N_0}}\right)$

$$E_\mathrm{g} = E_\mathrm{b}\,\frac{3 \cdot \mathrm{ld}(M)}{M^2 - 1} = E_\mathrm{b}\,\frac{d_\mathrm{min}^2}{2}$$

Symbol error probability (2 outer, $M-2$ inner points):

$$\boxed{\mathsf{SER} = \frac{2M-2}{M}\,Q\left(\sqrt{d_\mathrm{min}^2\frac{E_\mathrm{b}}{N_0}}\right)}$$

3.9.1.2 PSK

$M = 2$: 2PSK $\equiv$ 2ASK BER $=$ SER $= Q\left(\sqrt{2\frac{E_\mathrm{b}}{N_0}}\right)$

$M = 4$: 4PSK $\equiv$ 4QAM corresponds to 2ASK in I– and Q–component

$$\mathsf{SER} = 1 - \left(1 - Q\left(\sqrt{2\frac{E_\mathrm{b}}{N_0}}\right)\right)^2 \leq 2\,Q\left(\sqrt{2\frac{E_\mathrm{b}}{N_0}}\right) \quad \text{(two nearest neigbours)}$$

$$\mathsf{BER} = Q\left(\sqrt{2\,\frac{E_\mathrm{b}}{N_0}}\right) \quad \text{for Gray–mapping}$$

$M > 4$: SER $\leq 2\,Q\left(\sqrt{d_\mathrm{min}^2\frac{E_\mathrm{b}}{N_0}}\right)$

3.9.1.3 General Bound for the Signal Error Probability for Digital PAM

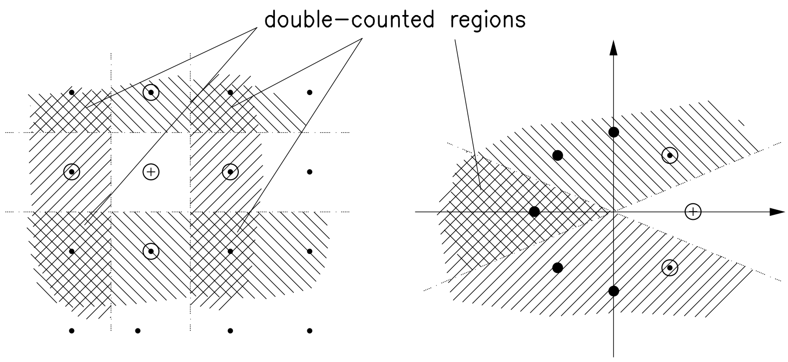
$$\boxed{\mathrm{SER} \leq N_\mathrm{min}\mathrm{Q}\left(\sqrt{d_\mathrm{min}^2\,\frac{E_\mathrm{b}}{N_0}}\right)}$$

with N_min : average number of nearest neighbor signal points

Illustration of the bound for

QAM and PSK

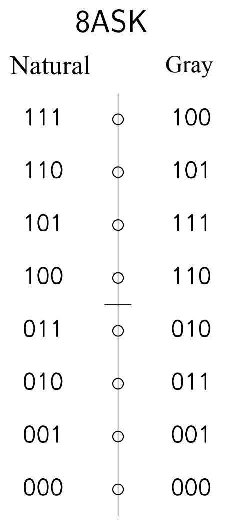
Number of neighbor signal points
for $\oplus$ signal point is $= 4$ $\qquad N_\mathrm{min} = 2$

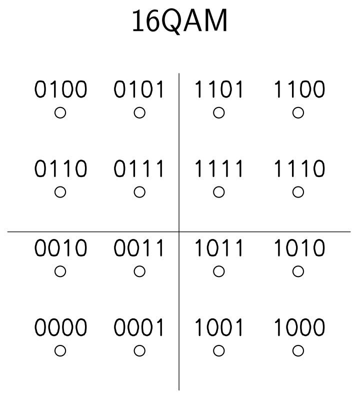


3.9.1.4 Bit Error Probability for Digital PAM

For a given symbol error probability, the bit error probability can be minimized with Gray mapping. Neighboring signal points are mapped to words of $\mathrm{ld}(M)$ binary symbols in such a way that they only differ in **one** binary symbol.

Examples for Gray–mapping:





Note: A Gray mapping exists for all regular (standard) signal constellations in one and two (real) dimensions. However, it is easy to find constellations for which a Gray mapping does not exist.

140

Bit Error Probability for **Gray Mapping** (Approximation):

$$\boxed{\mathrm{BER} \approx \frac{N_{\min}}{\mathrm{ld}(M)}\, \mathrm{Q}\left(\sqrt{d_{\min}^2 \frac{E_\mathrm{b}}{N_0}}\right)}$$

Assumption: The SNR is high enough to make a misclassification to the nearest neighbor points significantly higher than other points.

ASK bipolar:

M:	2	4	8	16	320
$d_{\min}^2$:	2	0.8	0.286	0.094	0.0293
$N_{\min}$:	1	1.5	1.75	1.875	1.9375
σ_a^2:	1	5	21	85	341

ASK unipolar:

M:	2	4	8	16	32
$d_{\min}^2$:	1	0.286	0.0857	0.0258	0.0077
$N_{\min}$:	1	1.5	1.75	1.875	1.9375
σ_a^2:	1	5	21	85	341
m_a:	1	3	7	15	31

PSK:

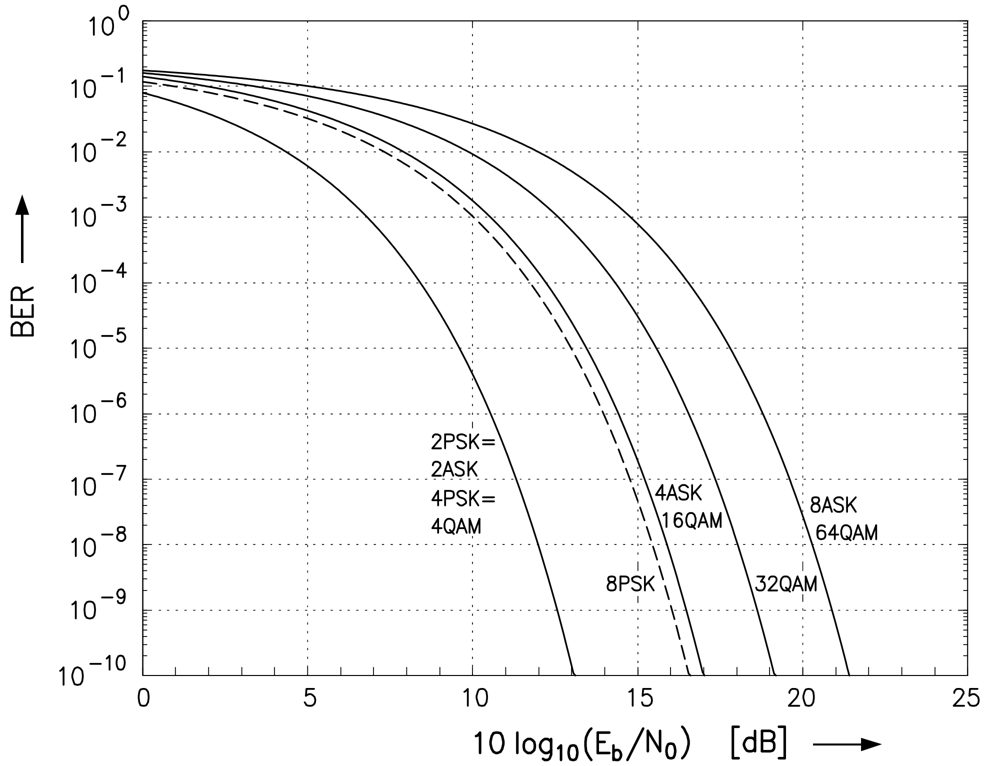
M:	2	4	8	16
$d^2_{\min}$:	2	2	0.878	0.304
$N_{\min}$:	1	2	2	2
σ^2_a:	1	1	1	1

QAM:

M:	2	4	16	32 (cross)
$d^2_{\min}$:	2	2	0.8	0.5
$N_{\min}$:	1	2	3	3.25
σ^2_a:	1	2	10	20

M:	64	128 (cross)	256
$d^2_{\min}$:	0.286	0.171	0.094
$N_{\min}$:	3.5	3.625	3.75
σ^2_a:	42	82	170

Bit error probabilities of digital PAM transmission schemes
as a function of the average energy per bit to noise PSD: E_{b}/N_0:



3.9.1.5 Digital PAM Scheme in the Power–Rate–Diagram (AWGN Channel)

Power efficiency: $\mathrm{BER} = f\left(\frac{E_\mathrm{b}}{N_0}\right) \Rightarrow E_\mathrm{b}/N_0 = f^{-1}(\mathrm{BER_T})$ for a max. tolerated $\mathrm{BER_T}$

E.g.: $\mathrm{BER_T} = 10^{-5}$; $\mathrm{BER_T} = 10^{-8}$

Rate of a communications scheme: R [bit/symbol]

Shannon Limit:

Capacity of the discrete–time AWGN channel (ECB–signals!):

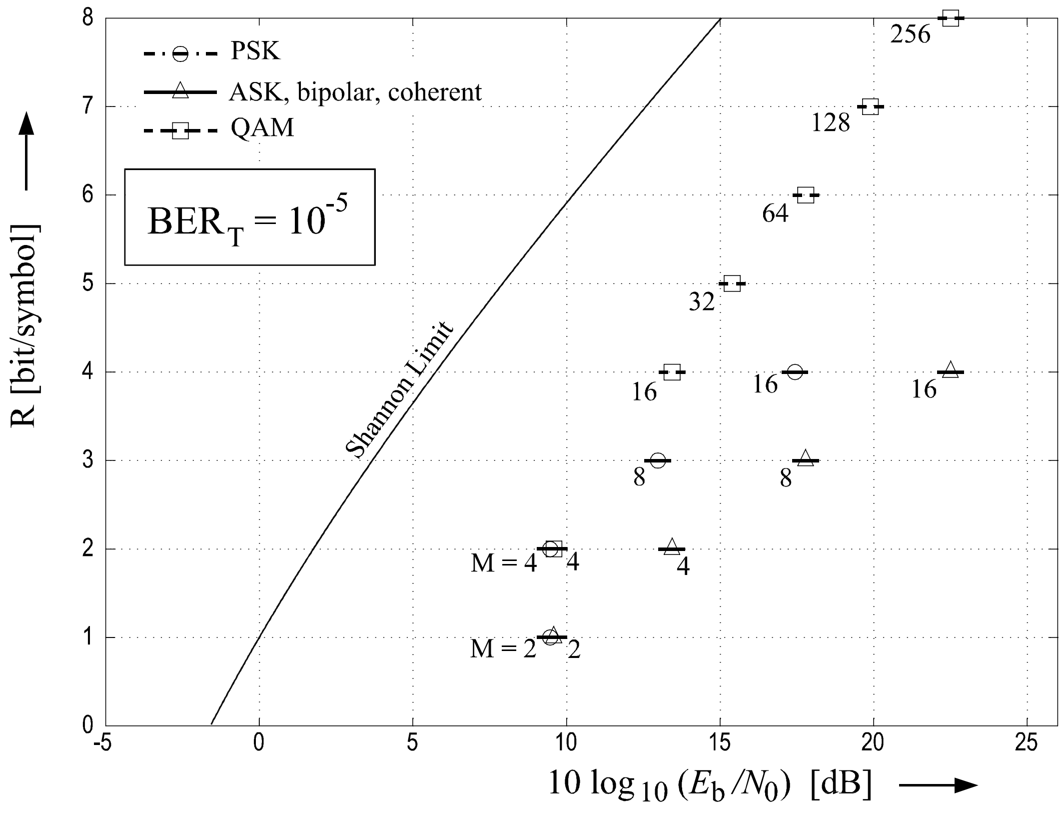
$$C = \log_2\left(1 + \frac{S}{N}\right) \qquad \left[\frac{\text{bit}}{\text{channel use}}\right] \qquad N = \frac{N_0}{T} \ \text{ for } \sqrt{\text{Nyquist}}\text{–pulses}$$

Optimal communication scheme, $R \overset{!}{=} C$:

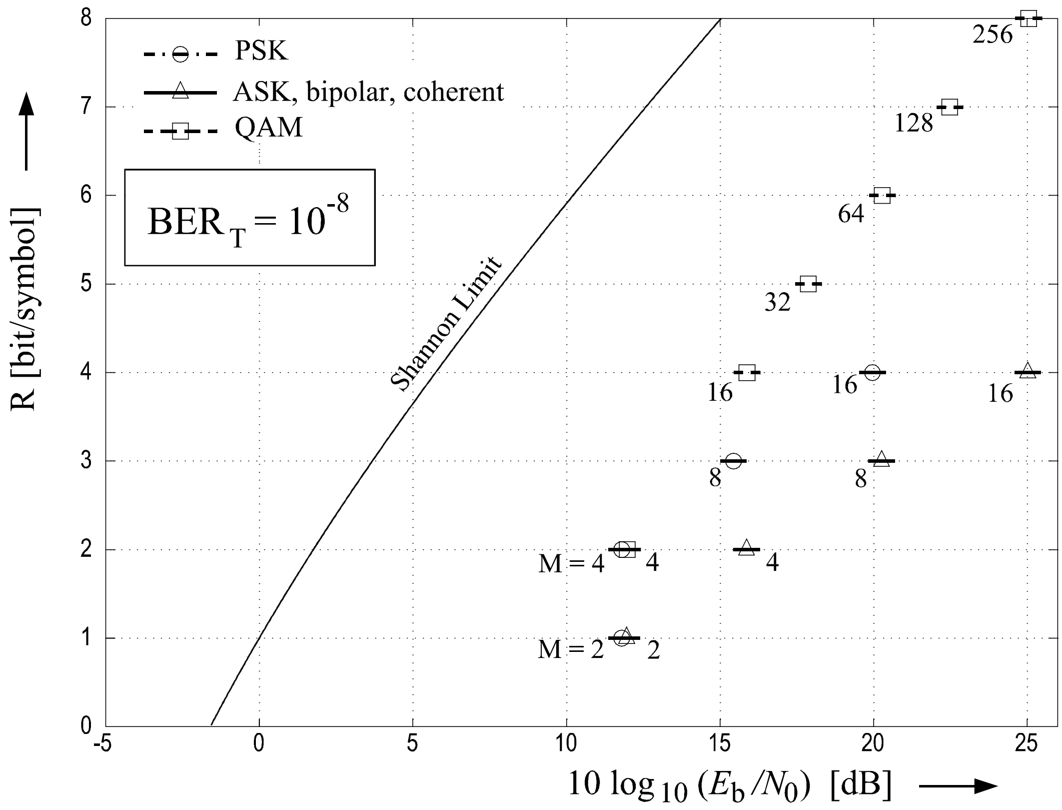
$$R = \log_2\left(1 + \frac{S}{N_0/T} \cdot \frac{T_\mathrm{b}}{T_\mathrm{b}}\right) \qquad \text{with} \quad R = \frac{T}{T_\mathrm{b}}$$

$$\boxed{R = \log_2\left(1 + \frac{E_\mathrm{b}}{N_0} \cdot R\right) \qquad \text{alternatively} \qquad \frac{E_\mathrm{b}}{N_0} = \frac{1}{R}\left(2^R - 1\right)}$$

Power–Rate–Diagram for PAM with $\mathrm{BER_T} = 10^{-5}$:



Power–Rate–Diagram for PAM with $\mathrm{BER_T} = 10^{-8}$:



3.9.1.6 Digital PAM in the Power–Bandwidth–Diagram

Bandwidth efficiency: $\sqrt{\text{Nyquist}}$–pulse with roll–off factor $\alpha \in [0; 1]$

Carrier–modulated transmission:

$$B_{\mathrm{HF}} = \frac{1+\alpha}{T} = \frac{R_{\mathrm{T}}}{R}(1+\alpha)$$

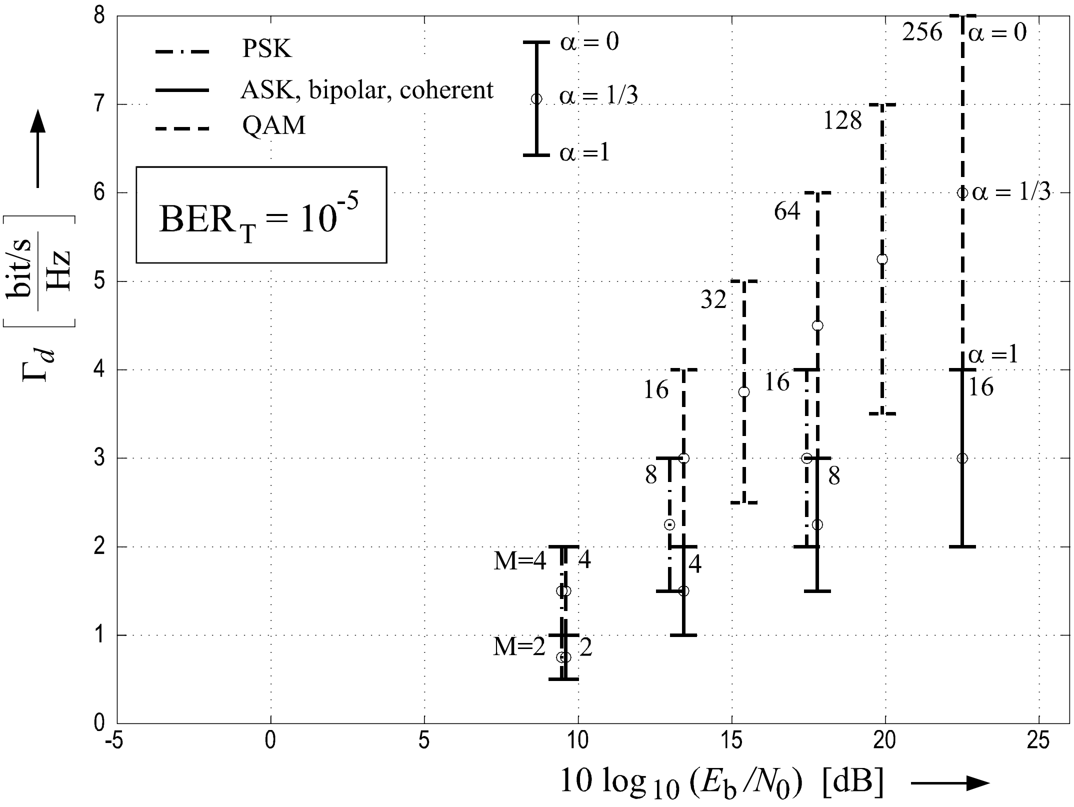
$$\boxed{\Gamma_d = \frac{R_{\mathrm{T}}}{B_{\mathrm{HF}}} = \frac{R}{1+\alpha} = \frac{\mathrm{ld}(M)}{1+\alpha}}$$

Digital baseband transmission:

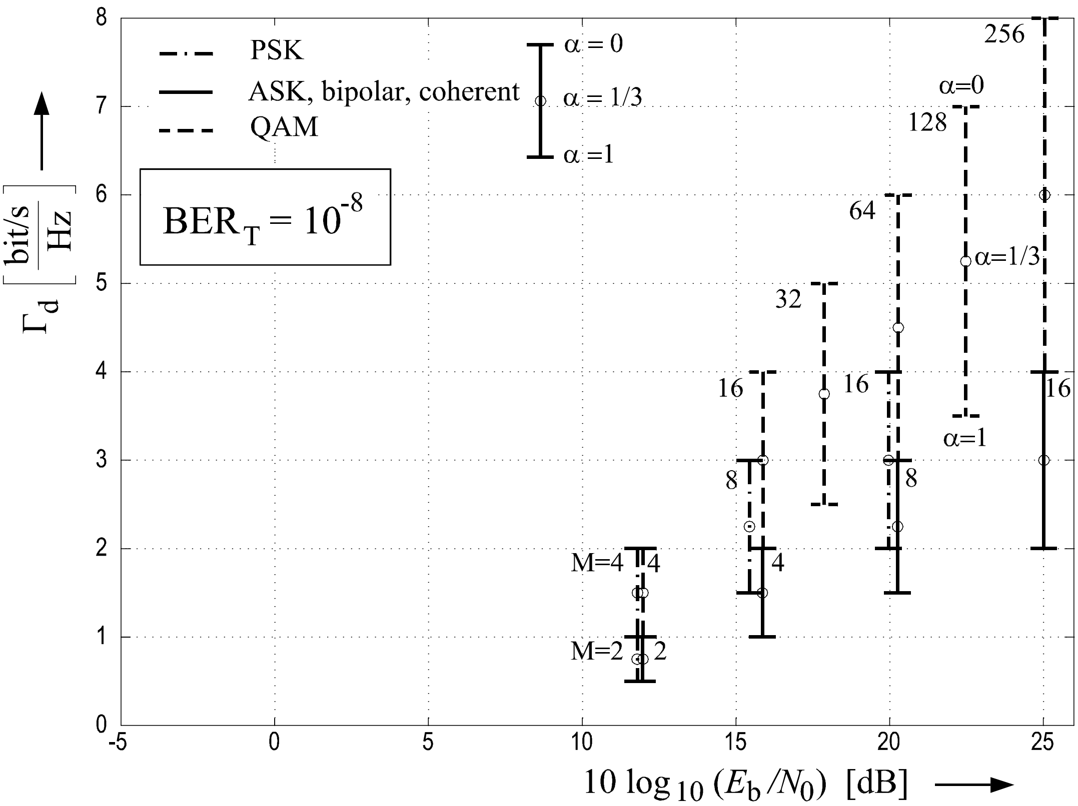
$$B_{\mathrm{s}} = \frac{1+\alpha}{2T}$$

$$\boxed{\Gamma_d = 2\,\frac{R}{1+\alpha} = 2\,\frac{\mathrm{ld}(M)}{1+\alpha}}$$

Power–Bandwidth–Diagramm for PAM with $\mathrm{BER_T} = 10^{-5}$:



Power–Bandwidth–Diagramm for PAM with $\mathrm{BER_T} = 10^{-8}$:

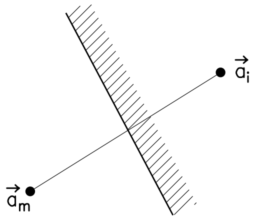


3.9.1.7 General Upper Bound of Symbol Error Probability (Union Bound)

For any modulation scheme (not only PAM!) with or without channel coding (word length N), the following holds:

$$\Pr(\hat{m} \neq m) = \Pr\left(\bigcup_{\substack{i=0 \\ i\neq m}}^{M-1} (\hat{m} = i)\right) \leq \sum_{\substack{i=0 \\ i\neq m}}^{M-1} \Pr(\hat{m} = i|m)$$

Probability of detecting signal point a_m as point a_i: $\Pr(\hat{m} = i|m)$



$2(N)$–dimensional signal space

For the misclassification of point a_m as point a_i, the only relevant component of the isotropic noise is the component in the direction of the straight line connecting the two points. All other components (which are orthogonal to this direction) are irrelevant. $\Rightarrow$ A one–dimensional (real) Gaussian Error Integral is sufficient for the calculation of the probability of misclassification, independent of the dimensionality of the signal space for each modulation step!

$$\Pr(\hat{m} = i|m) = Q\left(\frac{\parallel \vec{a}_m - \vec{a}_i \parallel /2}{\sigma_I}\right) = Q\left(\sqrt{\frac{\parallel \vec{a}_m - \vec{a}_i \parallel^2}{2}\frac{E_{\mathrm{g}}}{N_0}}\right)$$

$|\vec{a}_m - \vec{a}_i|^2$: Squared Euclidean distance between signal vectors $\vec{a}_m$ and $\vec{a}_i$ (energy of the difference of the signal elements)

Definition: *Normalized squared Euclidean distance*

$$d^2(m,i) = \frac{|\vec{a}_m - \vec{a}_i|^2 \cdot E_{\mathrm{g}}}{2\,E_{\mathrm{b}}}$$

$$\boxed{\Pr(\hat{m} = i|m) = Q\left(\sqrt{d^2(m,i) \cdot \frac{E_{\mathrm{b}}}{N_0}}\right)}$$

Expectation over all transmitted signal elements:

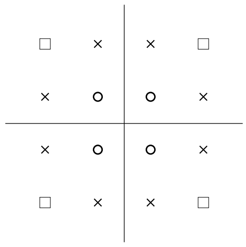
$$\Rightarrow \boxed{\mathsf{SER} \leq \sum_{\forall\, d} N(d)\, Q\left(\sqrt{d^2 \cdot \frac{E_{\mathrm{b}}}{N_0}}\right)}$$

where $N(d)$ denotes the average number of signal points with normalized distance d. $N(d)$ is also referred to as the distance profile or distance spectrum.

Definition: *Average number of nearest neighbor error events*

$$\boxed{N_{\min} = N(d_{\min})}$$

Example: *Distance profile for 16QAM*



4 inner points ○

distance	2	$2\sqrt{2}$	4	$\sqrt{20}$	$4\sqrt{2}$
number	4	4	2	4	1

8 points ×

distance	2	$2\sqrt{2}$	4	$\sqrt{20}$	$4\sqrt{2}$	6	$\sqrt{40}$	$\sqrt{52}$
number	3	2	2	3	1	1	2	1

4 points □

distance	2	$2\sqrt{2}$	4	$\sqrt{20}$	$4\sqrt{2}$	6	$\sqrt{40}$	$\sqrt{52}$	$6\sqrt{2}$
number	2	1	2	2	1	2	2	2	1

152

Averages

distance	2	$2\sqrt{2}$	4	$\sqrt{20}$	$4\sqrt{2}$	6	$\sqrt{40}$	$\sqrt{52}$	$6\sqrt{2}$
average number	3	$\frac{36}{16}$	2	3	1	1	$\frac{24}{16}$	1	$\frac{4}{16}$

$$E_\mathrm{b} = 2\,\frac{M-1}{3}\,E_\mathrm{g}/\mathrm{ld}(16) = 2.5E_\mathrm{g}$$

Average distance profile $N(d)$

d^2	0.8	1.6	3.2	4	6.4	7.2	80	88	96
$N(d)$	3	$\frac{9}{4}$	2	3	1	1	$\frac{6}{4}$	1	$\frac{1}{4}$
	$\uparrow$ $N_{\min}$								

where d is the normalized squared Euclidean distance.

3.9.2 ML–Decoding (with Channel Coding)

Channel code: Block code over N symbols

Decision in favor of the codeword $\vec{\hat{c}} = \mathcal{M}^{-1}\{\vec{\hat{m}}\}$, for which the Euclidean distance between the transmitted signal point

$$\vec{a} = (a_{m[0]}, a_{m[1]}, \ldots, a_{m[N-1]})$$

and the received signal point

$$\vec{d} = (d[0], d[1], \ldots, d[N-1])$$

is minimized.

True Maximum–Likelihood–Decoding for long block codes is not possible (exponential complexity in N!). It can, however, be approximated in certain cases using **iterative decoding methods**. A calculation of the error probability is possible only in special cases. Otherwise the union bound has to be employed to arrive at an upper bound.

Word Error Rate (WER)

$$\boxed{\text{WER} < \sum_d A(d)\, Q\left(\sqrt{d^2\,\frac{E_\text{b}}{N_0}}\right)}$$

with average distance profile $A(d)$ with respect to the normalized squared Euclidean distance of the code.

Bit error probability (Gray–mapping)

$$\boxed{\text{BER} \approx \sum_d A(d) \cdot \frac{d}{N\ \text{ld}(M)} Q\left(\sqrt{d^2\frac{E_\text{b}}{N_0}}\right)}$$

Note: Due to the energy consumed by the transmission of redundant code bits in general, a so–called

rate loss

occurs. The normalized squared Euclidean distance between two codewords is defined as

$$d^2 \overset{\text{def}}{=} \frac{|\vec{a}_m - \vec{a}_i|^2 \cdot E_\text{g}}{2\,E_\text{b}}, \qquad m \neq i$$

with energy per bit information $E_\text{b} = E_\text{s}\frac{T_\text{b}}{T} = \frac{E_\text{s}}{R}$ and energy per symbol $E_\text{s} = \text{E}\{E_m\}$, specifically for **binary** channel codes with the rate R_c, we have

$$E_\text{b} = E_\text{s}/(R_\text{c}\,\text{ld}(M)) \Rightarrow \quad d^2 = \frac{1}{2}R_\text{c} \cdot \text{ld}(M) \cdot |\vec{a}_m - \vec{a}_i|^2 \cdot \frac{E_\text{g}}{E_\text{s}}$$

with growing code redundancy, i.e., a decreasing code rate R_c, the normalized squared Euclidean distance becomes smaller due to the rate loss! On the other hand for decreasing code rate R_c, less words are valid codewords and the distances between codewords increases: Optimization of code rate for maximum power efficiency.

3.9.2.1 Channel Coding for Binary Bipolar Transmission Schemes

We consider the mapping of binary, bipolar amplitude coefficients $\in \{-1, +1\}$ to binary code symbols $\in \{0{,}1\}$

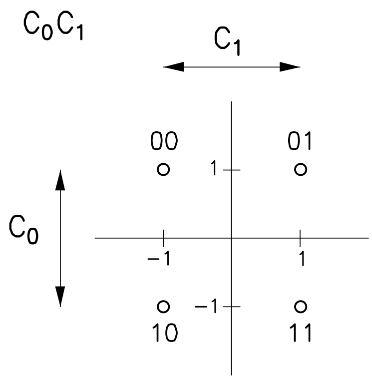
e.g. $\quad 0 \rightarrow +1 \quad ; \quad 1 \rightarrow -1$

in each dimension of a signal space.

This applies to e.g. 2PSK = 2ASK ($M = 2$, $E_\text{s} = E_\text{g}$)

and to

4PSK $\equiv$ 4QAM with Gray–mapping ($M = 4$, $E_\text{s} = 2E_\text{g}$)



If two binary codewords $\vec{c}_m$ and $\vec{c}_i$ of length n differ at δ positions,

$$\text{e.g.}\qquad \begin{array}{cccccc} 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ & & \uparrow & & \uparrow & \uparrow \end{array} \qquad \delta = 3$$

then the corresponding squared Euclidean distance in the N–dimensional signal space is equal to

$$|\vec{a}_m - \vec{a}_i|^2 = 4\,\delta$$

Definition:

The **number** δ of symbols in which two codewords $\vec{c}_n$ and $\vec{c}_i$ differ is referred to as the **Hamming distance** of the two codewords.

Note : Hamming distance induces a **metric** on the set of codewords (the code), the code becomes a **metric space**

Def. of Metric Space : Given an arbitrary set $\mathcal{A} = \{e_1,e_2,e_3,\cdots\}$. Define a mapping $d : \mathcal{A} \times \mathcal{A} \to \mathbb{R}_0^+$, i.e., non negative real number d is mapped to any tupel (e_i,e_j)

d is a metric if:

- $d \geq 0$
- $d(e_i,e_j) = d(e_j,e_i)$
- $d(e_i,e_i) = 0$
- $d(e_i,e_k) \leq d(e_i,e_j) + d(e_j,e_k)$

Example: *Average Distance Profile of Random Binary Codes*

Random selection of 2^k codewords out of the set of 2^n possible words

A. With respect to the Hamming distance of the 2^n possible codewords $\vec{c}_i$ of length n, exactly $\binom{n}{\delta}$ codewords differ from a random codeword $\vec{c}_m$ in δ positions (binominal distribution). With the probability $\frac{2^k}{2^n}$ a word of length n is chosen as a codeword in a code with rate $R_\mathrm{c} = \frac{k}{n}$.

   Average distance profile: $A_H(\delta) = \frac{2^k}{2^n}\binom{n}{\delta} = 2^{-n(1-R_\mathrm{c})}\binom{n}{\delta}$

B. Translation to

   the normalized squared Euclidean distance

$$\boxed{d^2 = 4 \cdot \tfrac{1}{2} R_\mathrm{c} \delta}$$

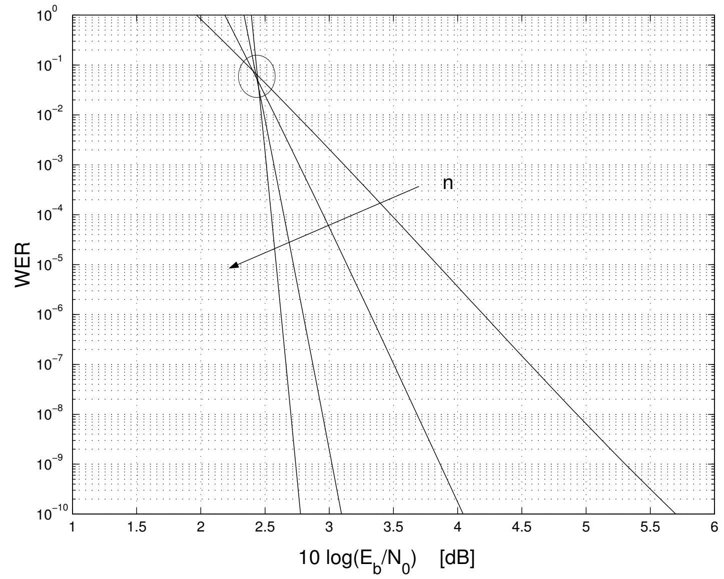
   This leads to :

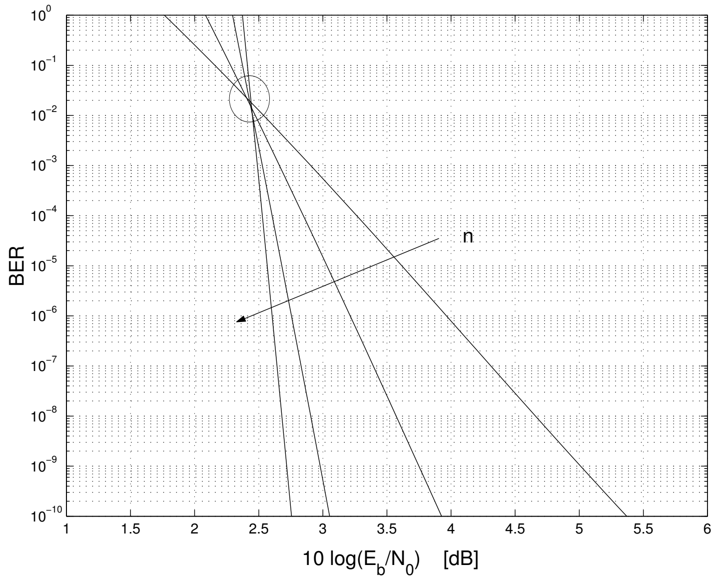
$$\text{Word error probability:}\qquad \mathrm{WER} \leq 2^{-n(1-R_\mathrm{c})} \sum_{\delta=0}^{n} \binom{n}{\delta}\, Q\left(\sqrt{2R_\mathrm{c}\delta\frac{E_\mathrm{b}}{N_0}}\right)$$

$$\text{Bit error probability:}\qquad \mathrm{BER} \leq 2^{-n(1-R_\mathrm{c})} \frac{1}{n} \sum_{\delta=0}^{n} \binom{n}{\delta} \cdot \delta \cdot Q\left(\sqrt{2R_\mathrm{c}\delta\frac{E_\mathrm{b}}{N_0}}\right)$$

   (Systematic coding assumed)

For $R_c = 1/2$ and $n = 100$, 200, 500, 1000 the following upper limits on word– and bit–error probabilities are observed for random channel codes and (soft decision) ML decoding.





<u>Intersection</u>: Cut-off rate $R_0 = \frac{1}{2}$

160

Estimation of word error probability according to Bhattacharyya (from Information Theory)

$$\mathrm{WER} < 2^{-n(R_0 - R_c)} \quad \text{with } R_0\text{: cut–off–Rate}$$

Binary–Input–AWGN channel:

$$R_0 = 1 - \log_2\left(1 + \mathrm{e}^{-R_c E_b/N_0}\right)$$

For $R_c = 1/2 \quad \mathrm{WER} \leq \left(2^{-\left(1/2 - \log_2\left(1+\mathrm{e}^{-E_b/(2N_0)}\right)\right)}\right)^n = \left(\frac{1+\mathrm{e}^{-E_b/(2N_0)}}{\sqrt{2}}\right)^n$

With Chernoff–bound for the Q–function (corresponds to Bhattacharyya–bound)

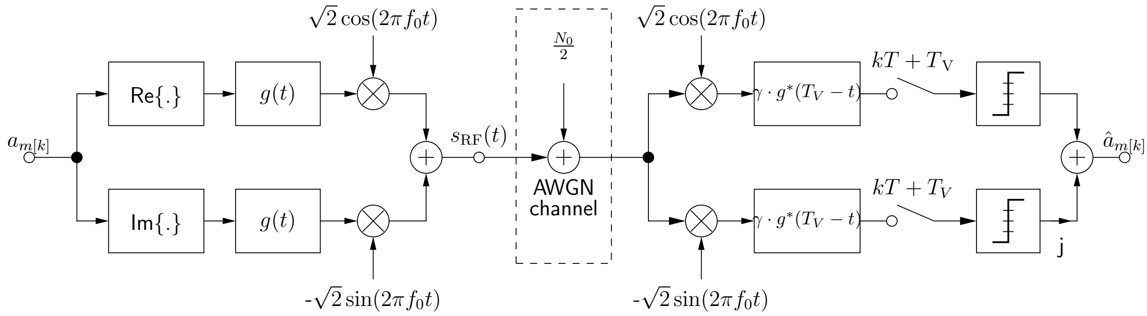
$$Q(x) \leq \mathrm{e}^{-x^2/2}$$

the same result as with the random code arises.

4 Variants of PAM–Transmission Schemes

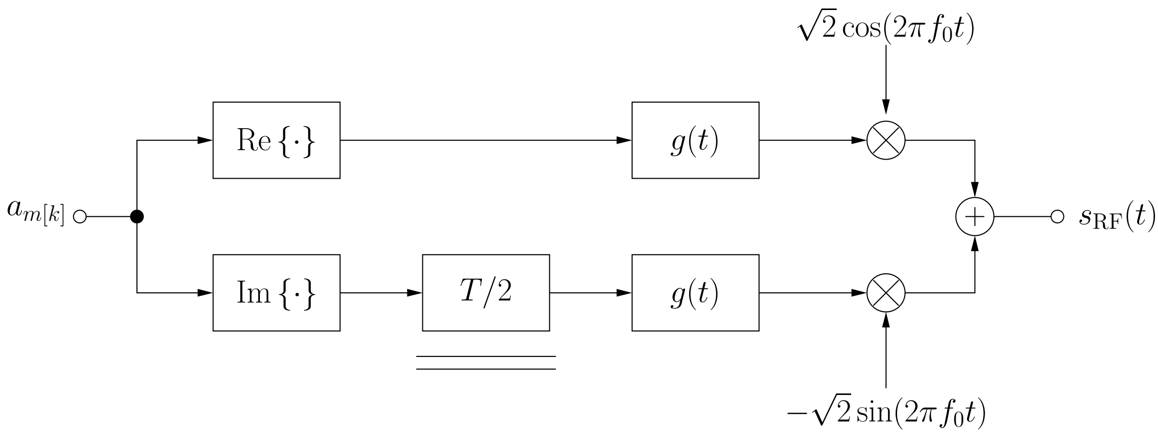
4.1 Offset–QAM (Staggered QAM)

QAM: Fundamental transmit pulse $g(t)$, $g(t) \in \mathbb{R}$



Offset–QAM :

Temporal offset of modulation in in–phase and quadrature component by one half of a symbol duration, i.e., $T/2$



Purpose:

Minimize variations in the envelope of the transmit signal, i.e., minimization of the Crest factor ζ_g

$\Rightarrow$ Less amplitude modulation (AM) (avoid zeros!)

$\Rightarrow$ Less AM/AM and AM/phase modulation (PM) conversion

$\Rightarrow$ Less back–off for the power amplifier at the transmitter necessary

$\Rightarrow$ Higher power efficiency

Crest factor of the pulse for Offset–QAM:

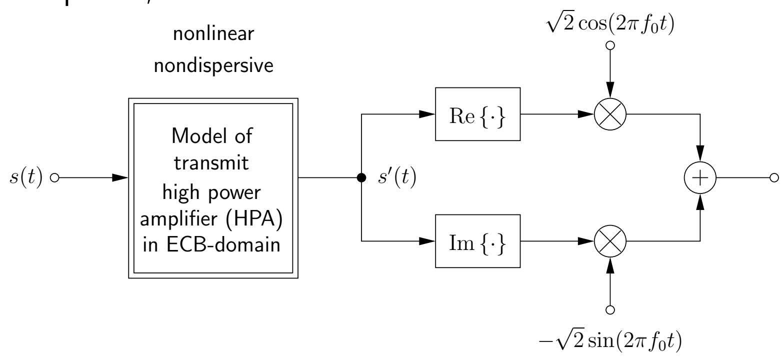
$$\zeta_g = \max_t \sqrt{\left(\left(\sum_{\mu=-\infty}^{+\infty} |g(t-\mu T)|\right)^2 + \left(\sum_{\mu=-\infty}^{+\infty} |g(t-\mu T - T/2)|\right)^2\right) \Big/ (2E_\mathrm{g}/T)}$$

The delay in the quadrature branch does not change the average power spectral density.

164

Transmitter amplification stages exhibit *nonlinear, dispersive* characteristics:
Amplitude distortion, amplitude dependent phase distortion.

Most stages are well approximated using nonlinear, *non*dispersive models for the ECB signal, e.g. traveling–wave tubes, GAsFET amplifiers, etc.



Description of the model using AM/AM and AM/PM conversion

$$s(t) = |s(t)| \cdot \mathrm{e}^{\mathrm{j}\arg(s(t))} \qquad s'(t) = A\left(\left|\frac{s(t)}{s_0}\right|\right) \cdot \mathrm{e}^{\mathrm{j}(P(|s(t)/s_0|)+\arg(s(t)))}$$

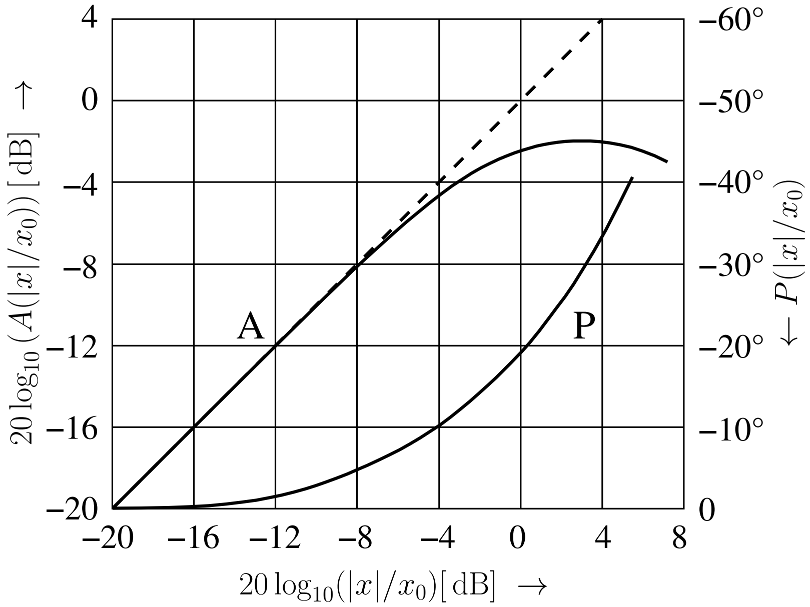
with

$A(\cdot)$: AM/AM characteristic

$P(\cdot)$: AM/PM characteristic

Example:

Travelling wave tube

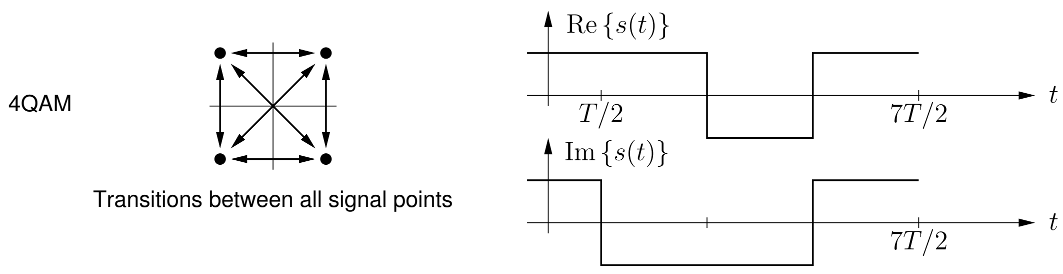


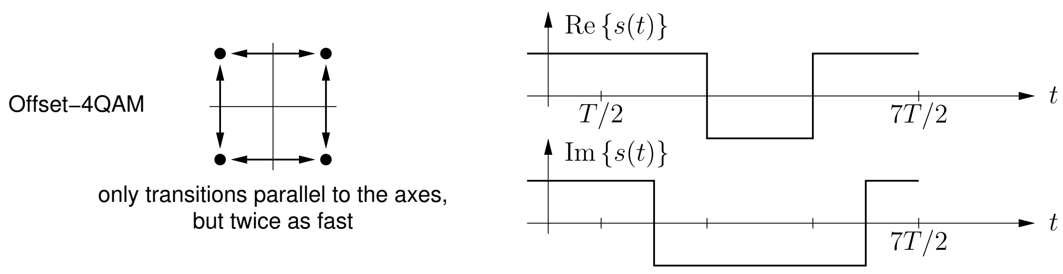
Examples:

A. Hard keying: $g(t) = \sqrt{\frac{E_\mathrm{g}}{T}}\,\mathrm{rect}(t/T)$

Phasor of the equivalent lowpass signal

(Nyquist plot of the signal in the complex plane)

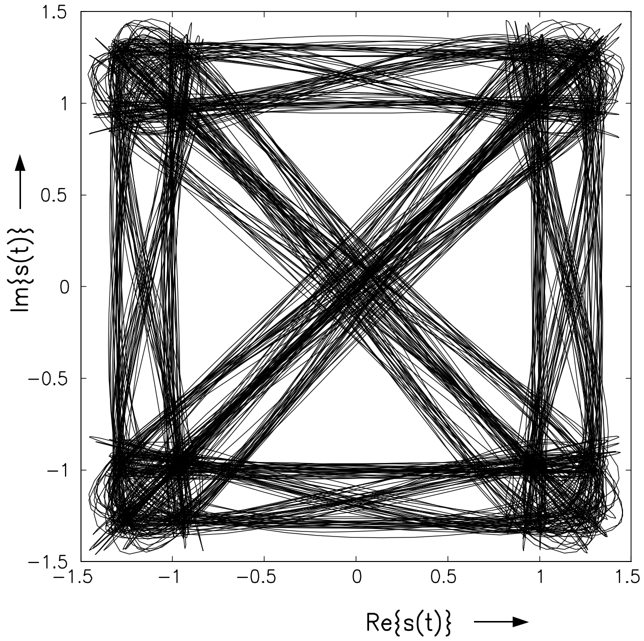




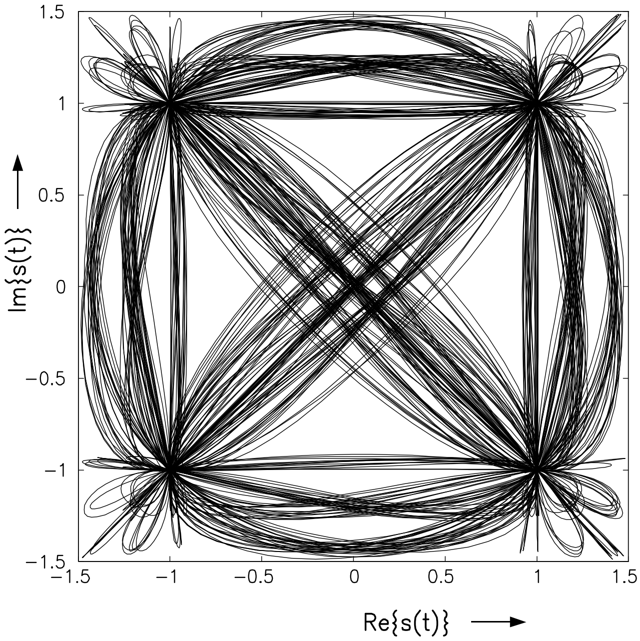
B. Soft keying: $\sqrt{\cos}$–roll–off $\alpha = 0.5$

4QAM:

transmit signal

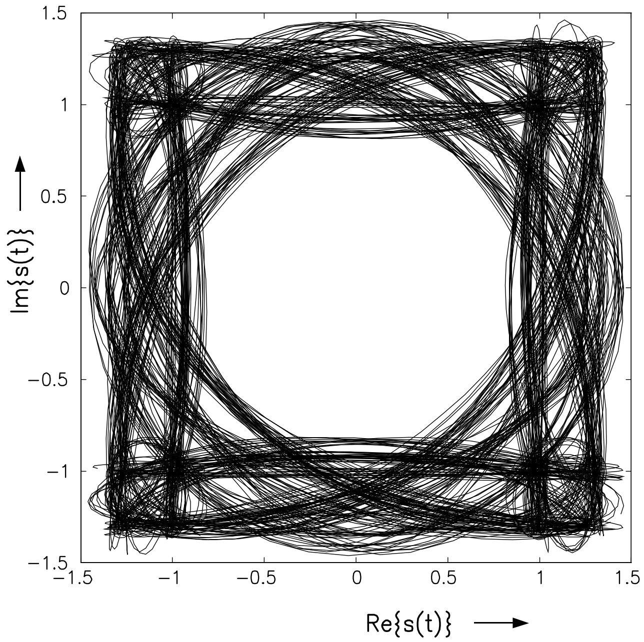


received signal (after matched–filtering)

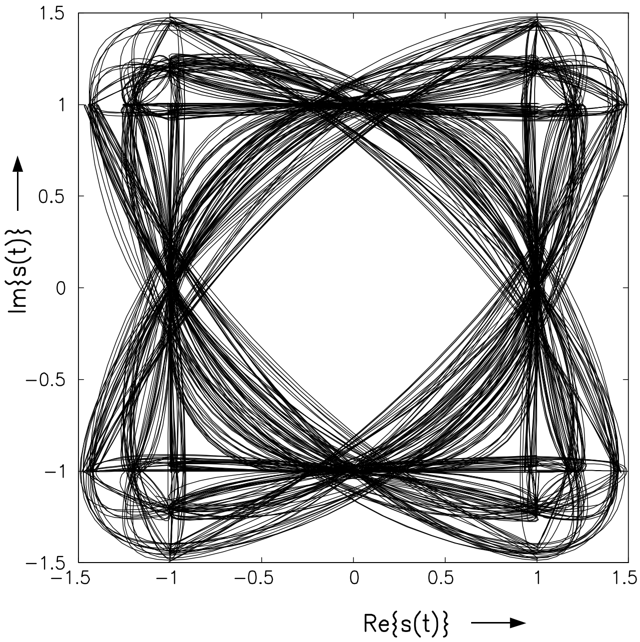


Offset–QAM:

transmit signal



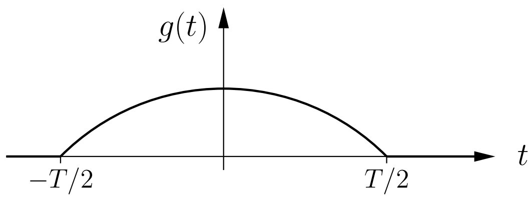
received signal (after matched–filtering)



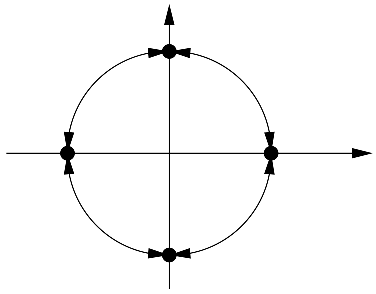
4.2 Minimum Shift–Keying (MSK)

MSK: Offset–4QAM with the pulse

$$g(t) = \sqrt{\frac{2E_\mathrm{g}}{T}} \cos(\pi t/T) \cdot \mathsf{rect}(t/T)$$



Phasor of the transmit signal



$\cos^2\varphi + \sin^2\varphi = 1$

Constant envelope, pure phase modulation

- ⇒ Crest factor $\zeta = 1$
- ⇒ very efficient power amplification, as no AM is used
  - ⇒ No AM/AM and no AM/PM conversion! No back–off necessary (high power amplifier at limit!)

Power efficiency of MSK: $d^2_\mathrm{min} = 2$

Average power spectral density of MSK:

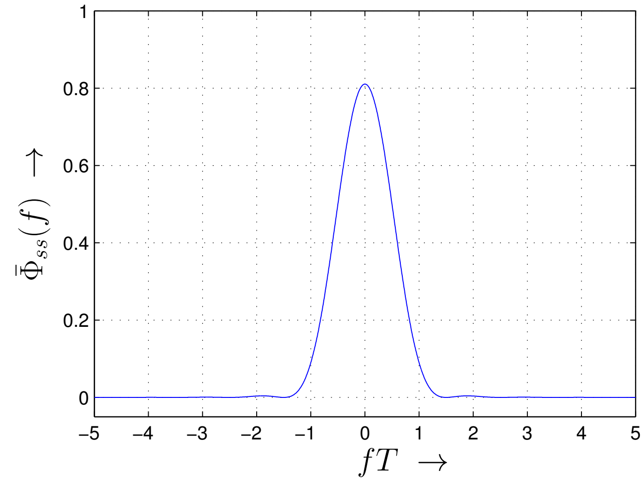
$$\begin{aligned}
\bar{\Phi}_{ss}(f) &= \frac{|G(f)|^2}{T} \cdot \mathrm{E}\{|a|^2\} \\
g(t) &= \sqrt{\frac{2E_\mathrm{g}}{T}} \cos(\pi t/T) \cdot \mathrm{rect}(t/T) \\
&\ \ \circ\!\!-\!\!\bullet \\
G(f) &= \sqrt{\frac{2E_\mathrm{g}}{T}} \frac{2T}{\pi} \cdot \frac{\cos(\pi f T)}{1-(2fT)^2} \\
\mathrm{E}\{|a|^2\} &= 2 \\
\Rightarrow \quad \bar{\Phi}_{ss}(f) &= \frac{16E_\mathrm{g}}{\pi^2} \cdot \frac{\cos^2(\pi f T)}{(1-(2fT)^2)^2}
\end{aligned}$$

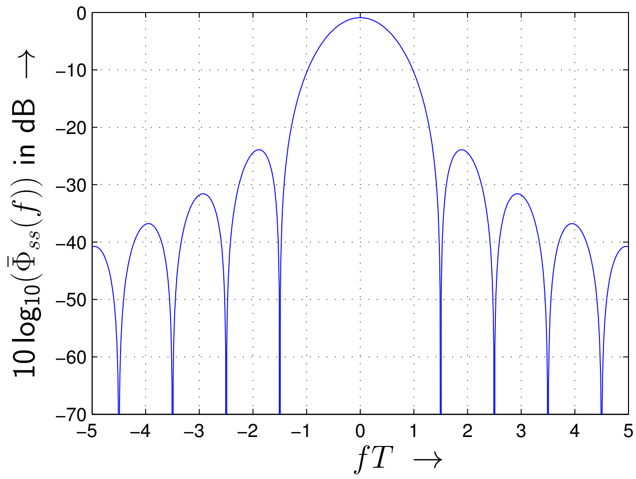
$$B_{99\%} = 1{,}11 R_\mathrm{T}\,; \qquad \Gamma_{\mathrm{d}\,99\%} = \frac{R_\mathrm{T}}{B_{99\%}} = 0{,}9\ \frac{\mathsf{bit/s}}{\mathsf{Hz}}$$

with $B_{X\%}$ defined as the bandwidth which contains $X\%$ of the transmit signal power (ECB–domain).

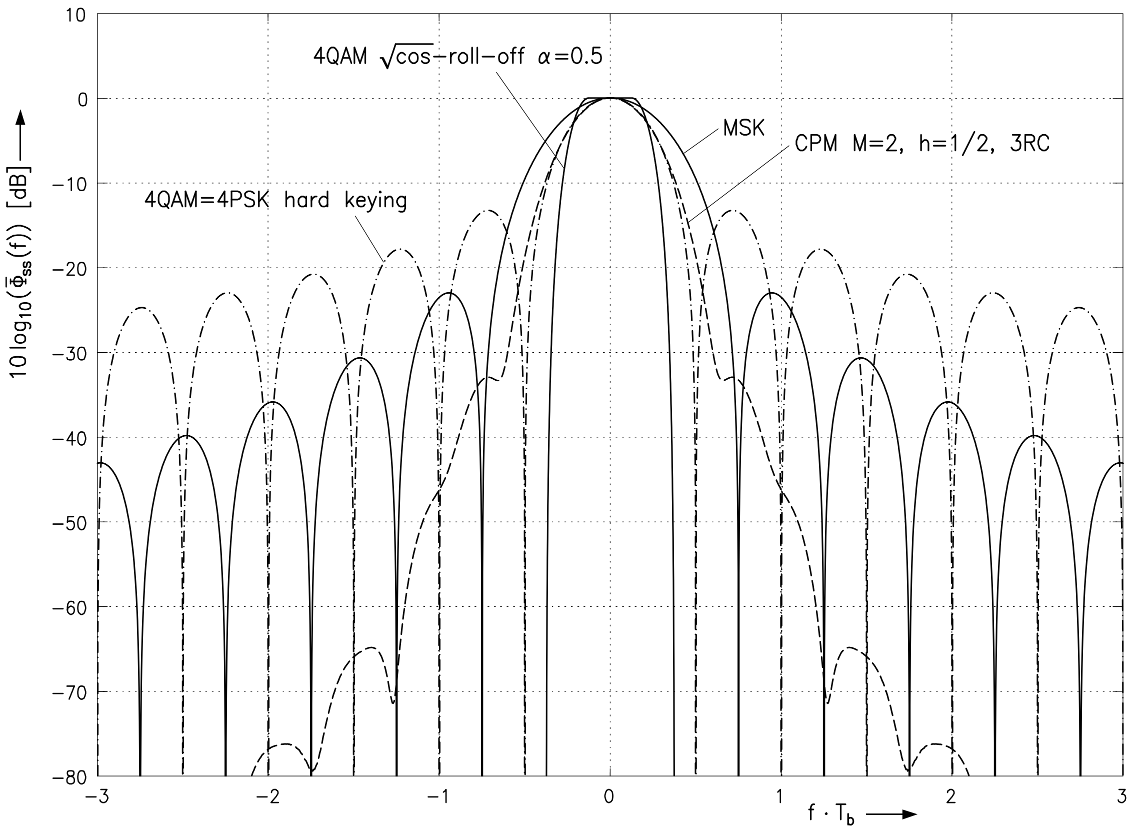
$$\int\limits_{-B_{X\%}/2}^{+B_{X\%}/2} \bar{\Phi}_{ss}(f)\,\mathrm{d}f = \frac{X}{100} \int\limits_{-\infty}^{+\infty} \bar{\Phi}_{ss}(f)\,\mathrm{d}f$$

Average Power Spectral Density of MSK: (linear and logarithmic scale for ordinate)

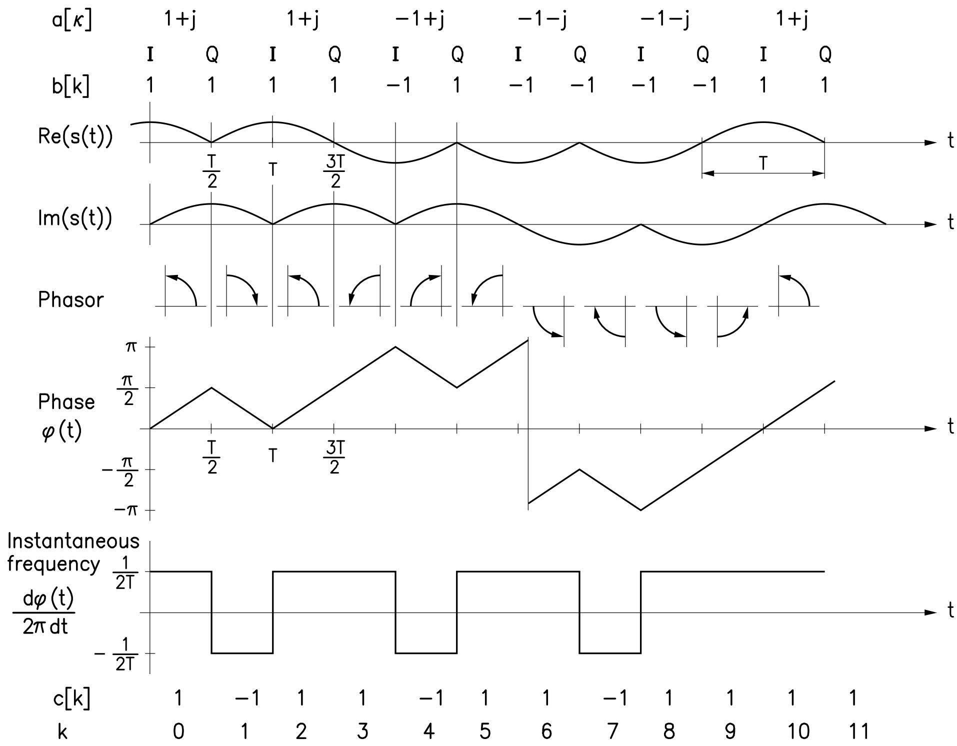




Comparison between average Power Spectral Densities



Equivalence between Minimum Shift Keying and Binary Frequency Modulation



The amplitude coefficients $a_{m[\kappa]} \in \{\pm 1 \pm \mathrm{j}\}$ of the QAM interpretation of MSK are given by

$$a_{m[\kappa]} = b[2\kappa - 1] + \mathrm{j}\, b[2\kappa]\ , \quad \text{mit} \quad b[k] \in \{\pm 1\}\ , \quad k, \kappa \in \mathbb{Z}.$$

The pair $(b[k-1], b[k])$ defines the *instantaneous frequency* within the time interval

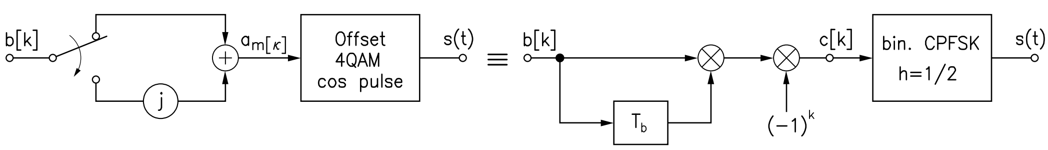
$[k\,T_\mathrm{b}, (k+1)T_\mathrm{b})$ with $T_\mathrm{b} = T/2$:

$$\frac{1}{2\pi}\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t} = \frac{1}{4T_\mathrm{b}}\, c[k] \quad \text{with} \quad c[k] \in \{\pm 1\}$$

$c[k]$ is the input symbol of an equivalent *frequency modulator*

Equivalent block diagram:

$$c[k] = (-1)^k \cdot b[k] \cdot b[k-1]$$

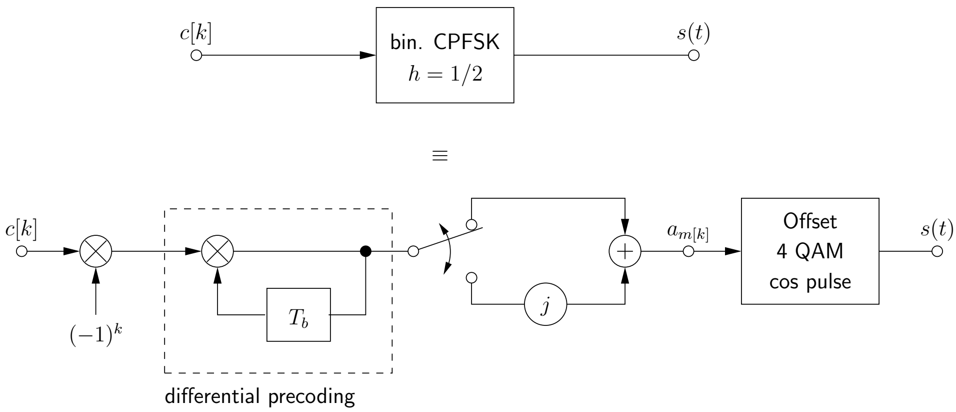


CPFSK: Continuous Phase Frequency Shift Keying

h: Modulation index

(separation between the instantaneous frequencies, related to the baudrate of CPFSK)

Alternatively, interpretation of binary CPFSK with $h = 1/2$ as Offset–QAM:



4.3 Gaussian Minimum Shift–Keying (GMSK)

MSK :

- constant envelope
  $\Longrightarrow$ efficient power amplification
- very wide power spectral density
  $\Longrightarrow$ low spectral efficiency

Desired : "Compact Spectrum"

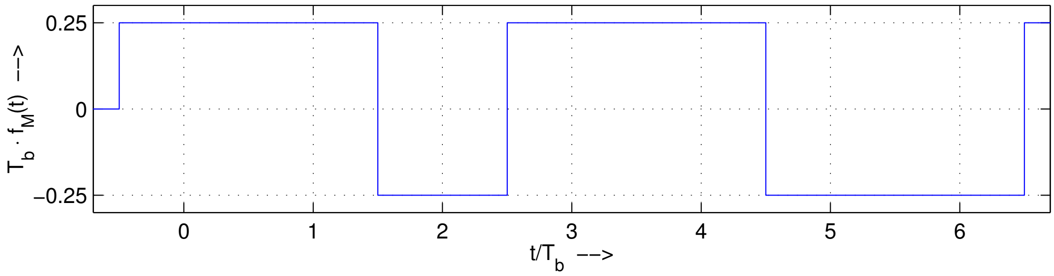
$\Rightarrow$ Avoidance of the frequency jumps in MSK via filtering of the instantaneous frequency signal

Generalization : Continous Phase Modulation (CPM) see Chapter 6

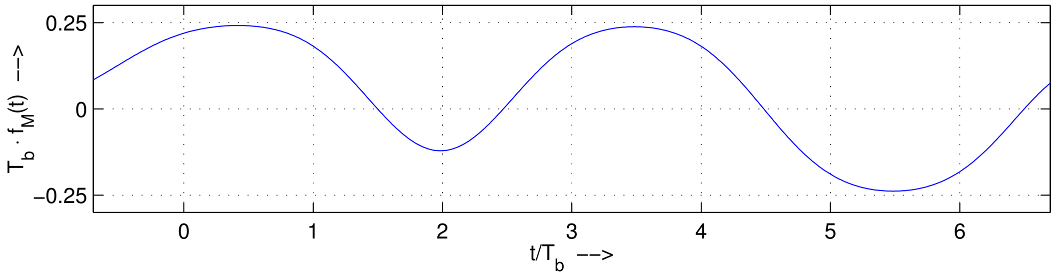
Example:

Instantaneous frequency over time

- MSK:

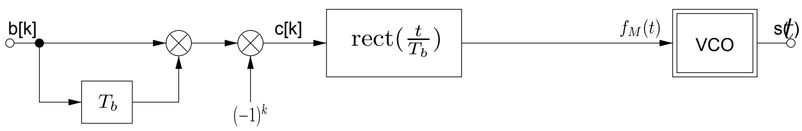


- GMSK:

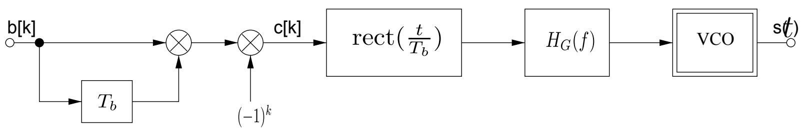


Modulation:

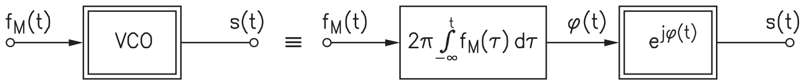
MSK:



GMSK:



Voltage Controlled Oscillator (VCO):



with

$$\boxed{H_G(f) = e^{-(f/B_b)^2 \cdot \ln 2/2}}$$

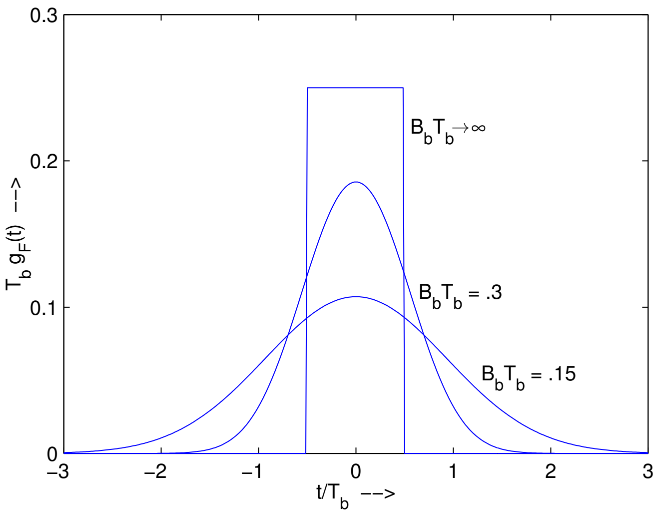
GSM: $B_b \cdot T_b = 0.3$

Parameter: $B_b \cdot T_b$

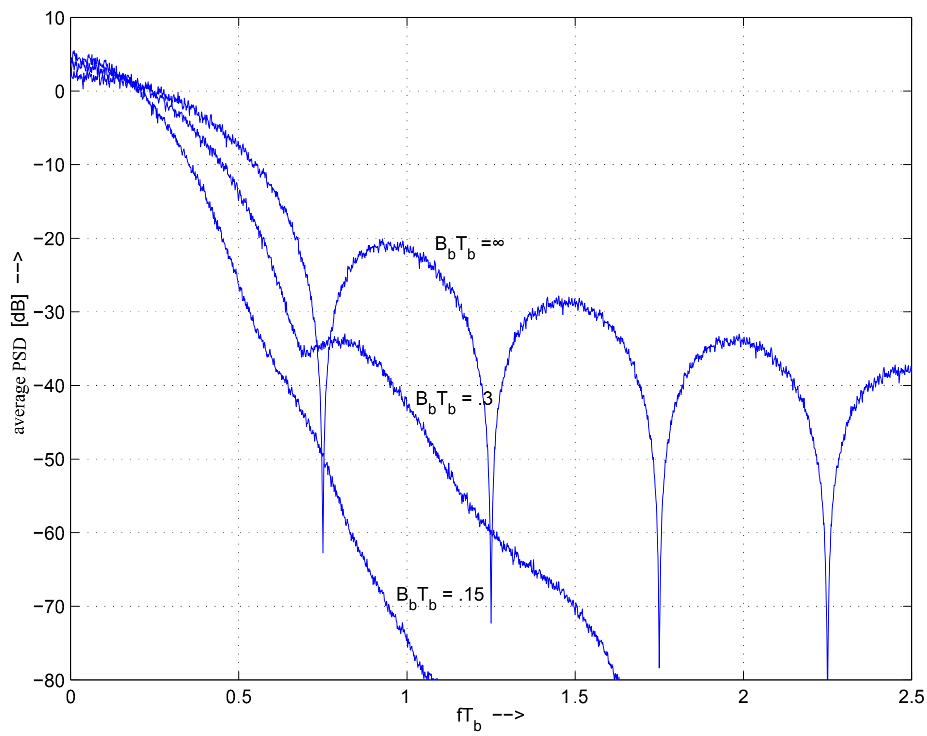
Frequency pulse:

$$
\begin{aligned}
H_{\mathrm{G}}(f) &= \mathrm{e}^{-(f/B_{\mathrm{b}})^2 \cdot \ln 2/2} \\
g_{\mathrm{F}}(t) &= \frac{1}{4T_{\mathrm{b}}} \cdot \mathrm{rect}\left(\frac{t}{T_{\mathrm{b}}}\right) * h_{\mathrm{G}}(t) \quad \text{with } h_{\mathrm{G}}(t) = \mathscr{F}^{-1}\left\{H_{\mathrm{G}}(f)\right\} \\
&= \frac{1}{4T_{\mathrm{b}}} \left(\mathrm{Q}\left(\frac{t + T_{\mathrm{b}}/2}{\sigma T_{\mathrm{b}}}\right) - \mathrm{Q}\left(\frac{t - T_{\mathrm{b}}/2}{\sigma T_{\mathrm{b}}}\right) \right)
\end{aligned}
$$

with $\sigma = \frac{\sqrt{\ln 2}}{2\pi B_{\mathrm{b}} T_{\mathrm{b}}}$, $Q(x)$: Gaussian Q–function

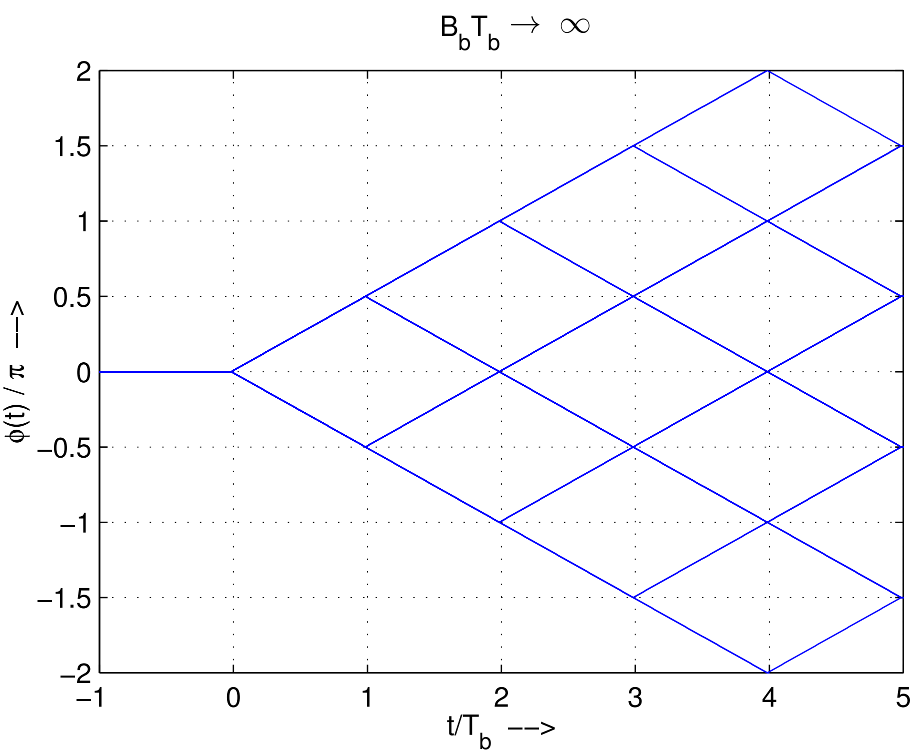


Average power spectral density of GMSK:

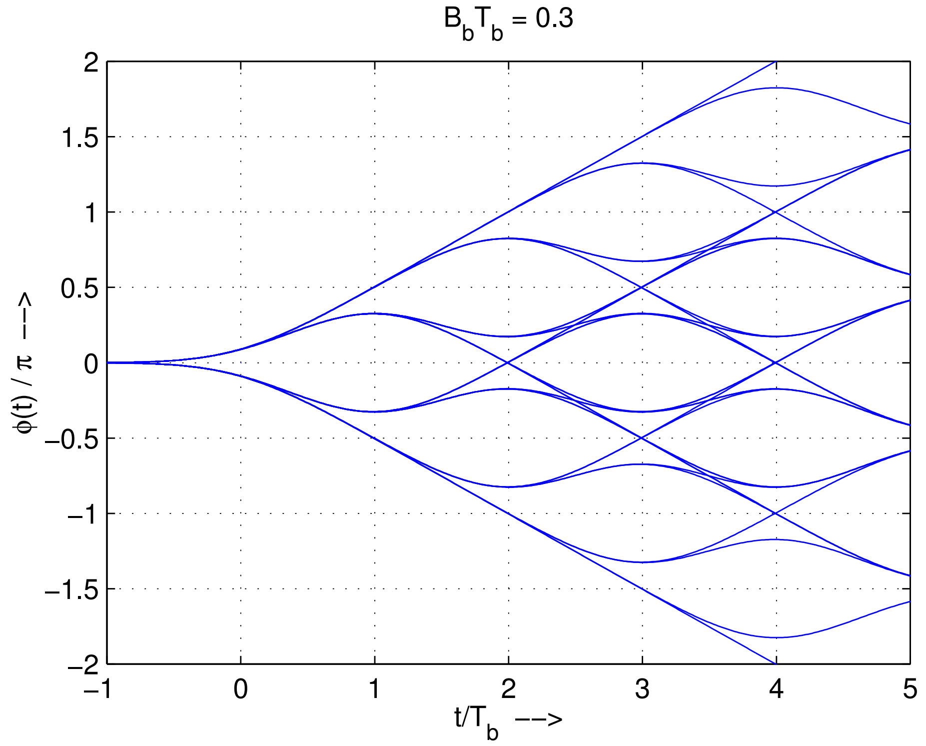


Phase tree (for different source symbols):

MSK:

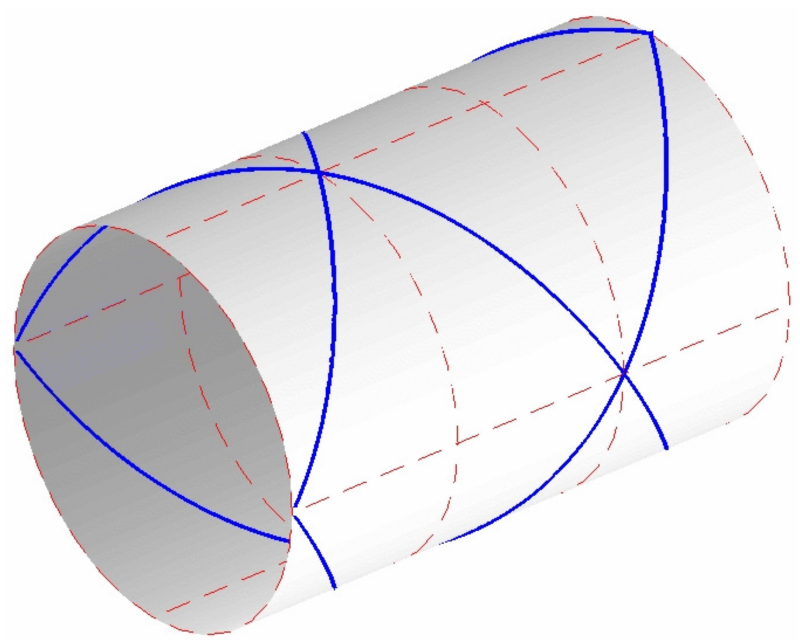


GMSK:

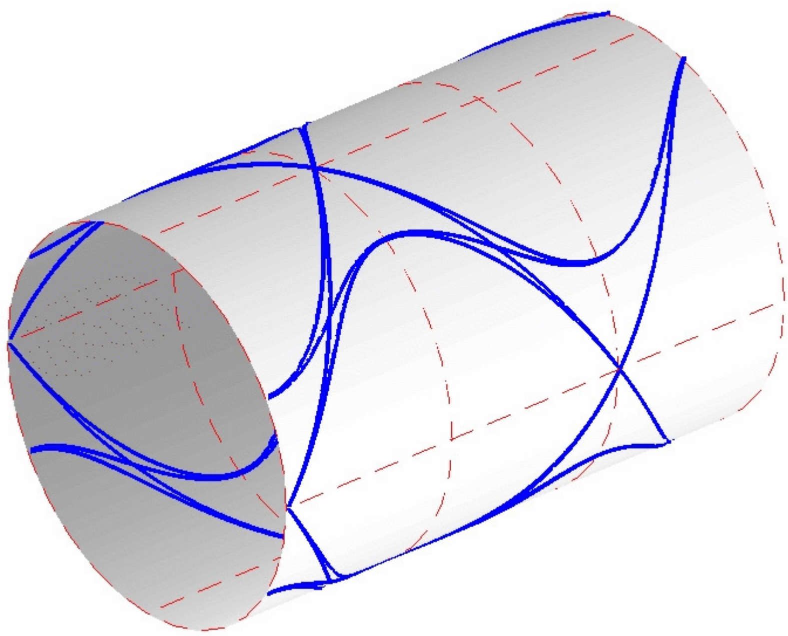


Phase cylinder:

MSK:



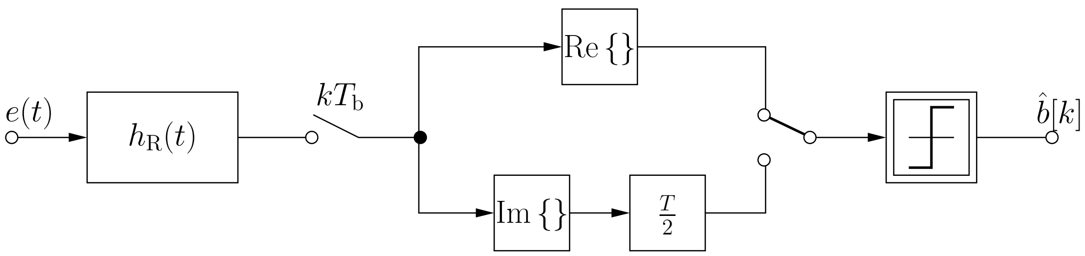
GMSK:



Demodulation:

- GMSK is a *nonlinear* modulation scheme
- However: In practice, GMSK is often approximated as offset–QAM for demodulation

ECB Representation:

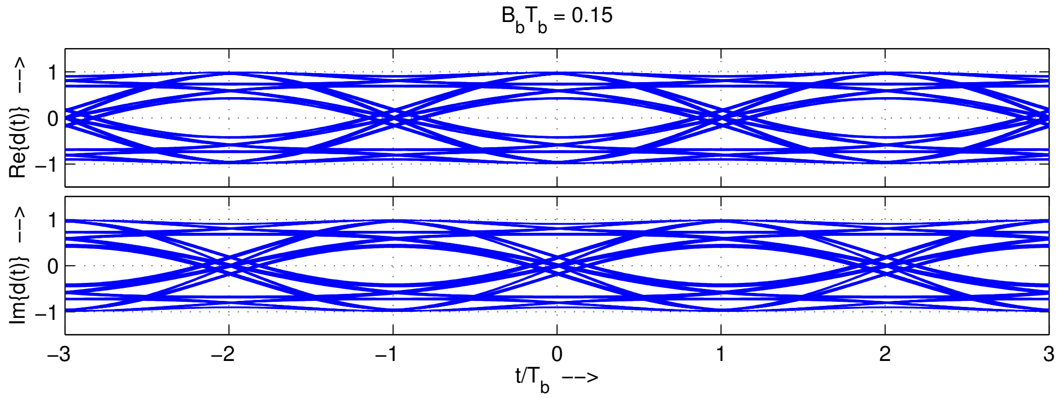


Receiver (input) filter $h_{\mathrm{R}}(t)$ is optimized (depending on SNR)

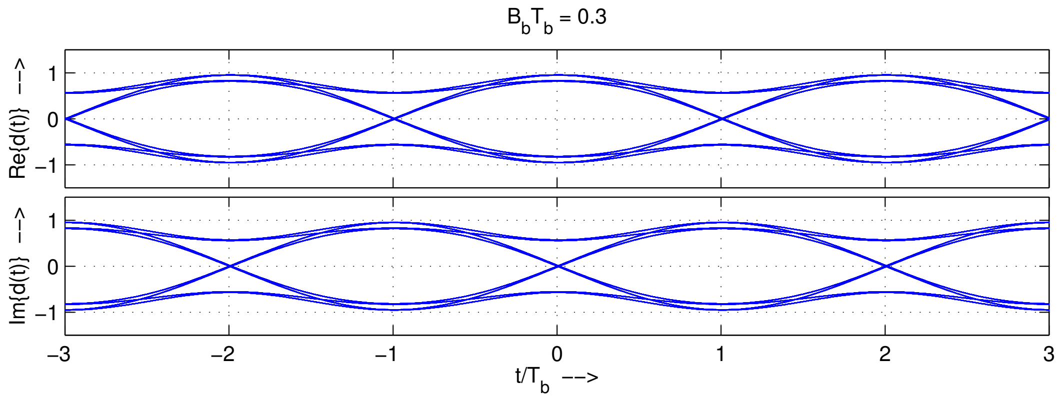
Compromise:

- Gaussian lowpass $H_{\mathrm{R}}(f) = \mathrm{e}^{-(f/B_{\mathrm{R}})^2 \cdot \frac{\ln 2}{2}}$
- Optimization of the filter bandwidth B_{R}
  Typical choice: $B_{\mathrm{R}} \cdot T_{\mathrm{b}} \approx 0.63$

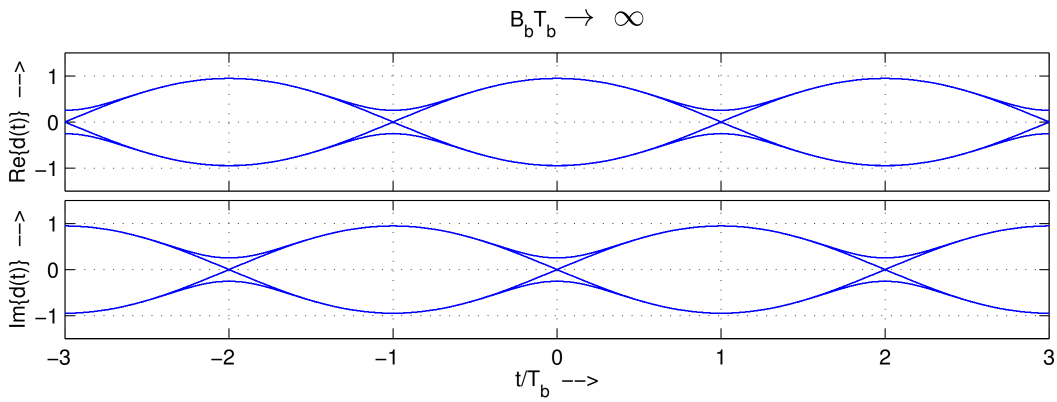
Eye pattern:



$B_bT_b = 0.15$
Re{d(t)} -->
Im{d(t)} -->
1
0
−1
−3
−2
−1
0
1
2
3
t/T_b -->



$B_bT_b = 0.3$
Re{d(t)} -->
Im{d(t)} -->
1
0
−1
−3
−2
−1
0
1
2
3
t/T_b -->

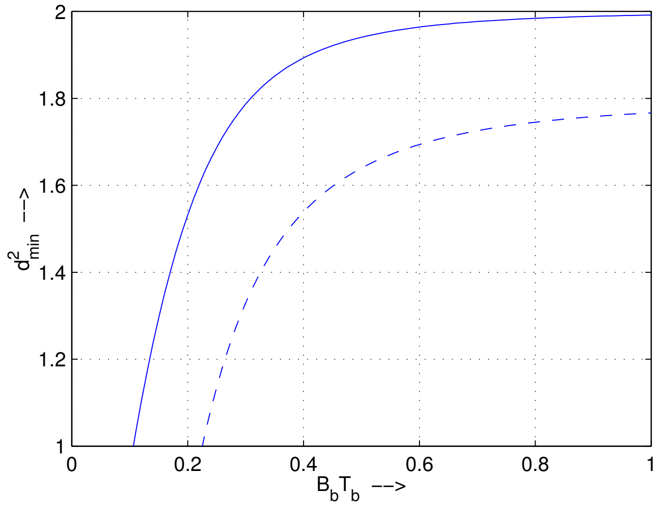


$B_bT_b \rightarrow \infty$
Re{d(t)} -->
Im{d(t)} -->
1
0
−1
−3
−2
−1
0
1
2
3
t/T_b -->

Power efficiency of GMSK:

$$\mathrm{BER} \approx \mathrm{Q}\left(\sqrt{d^2_{\min} \cdot \frac{E_\mathrm{b}}{N_0}}\right) \quad \text{with} \quad d^2_{\min} = \frac{1}{2E_\mathrm{b}} \cdot \min_{\substack{s^{(i)}(t), s^{(j)}(t) \\ s^{(i)}(t) \neq s^{(j)}(t)}} \int_{-\infty}^{\infty} |s^{(i)}(t) - s^{(j)}(t)|^2 \, \mathrm{d}t$$

minimum normalized squared Euclidean distance $d^2_{\min}$ as a function of $B_\mathrm{b}T_\mathrm{b}$:



— $d^2_{\min}$

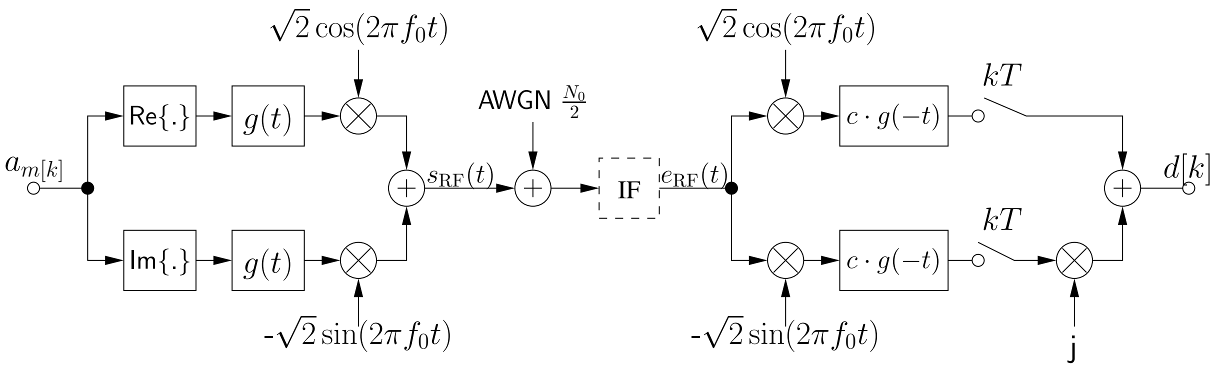
– – Degradation because of simplified demodulation

4.4 "Carrierless" Amplitude and Phase Modulation

Carrierless AM/PM (CAP)

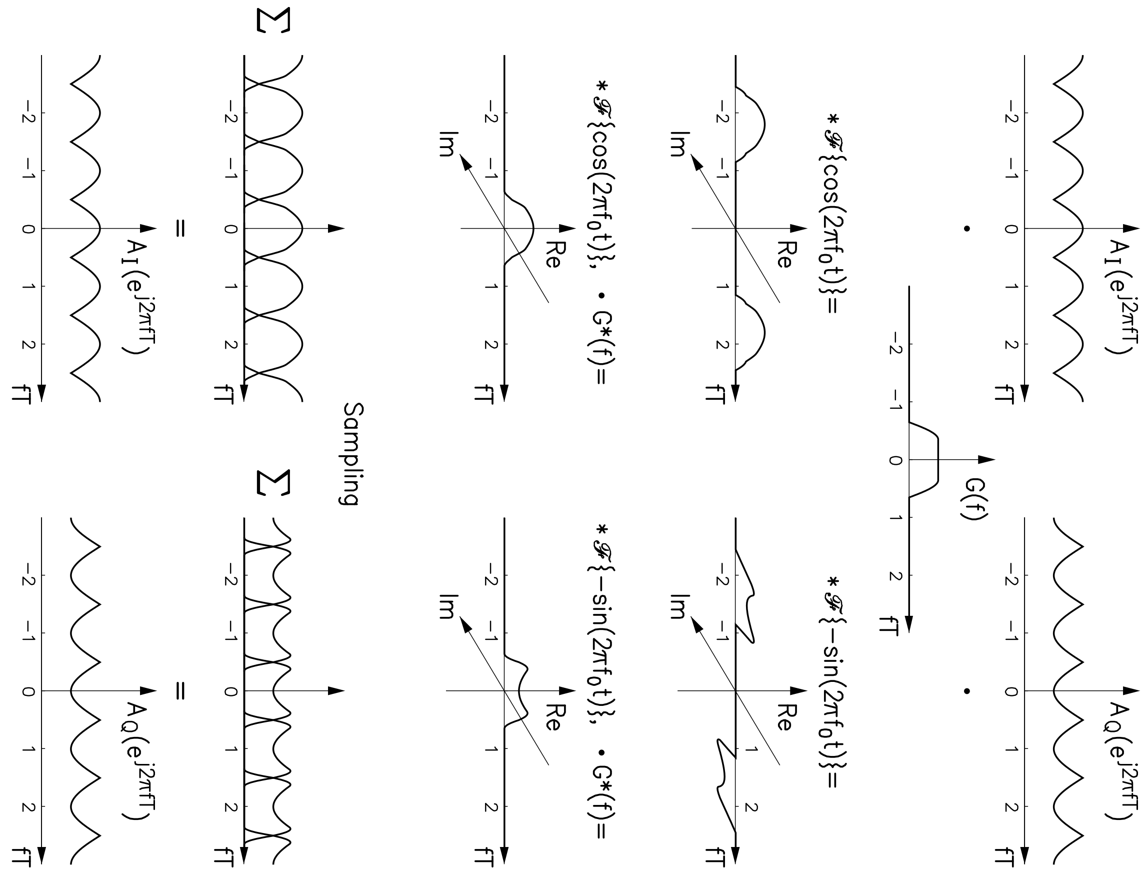
Application field: High rate data transmission over paired cables (Asymmetric Digital Subscriber Lines Systems (ADSL))

Review: QAM, fundamental transmit pulse $g(t)$



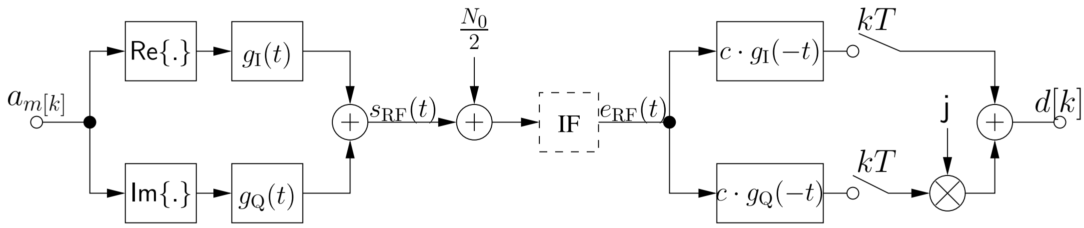
Example:

- $g(t)$ is a $\sqrt{\cos}$–roll–off pulse (here: $\alpha = 0.3$)
- $f_0 = 1.8\ /T$
- Spectra of the real sub-sequence of:
  $A_{\mathrm{I}}\left(\mathrm{e}^{\mathrm{j}2\pi fT}\right) = \mathscr{F}\left\{\mathrm{Re}\{a_{m[k]}\}\right\}$ and $A_{\mathrm{Q}}\left(\mathrm{e}^{\mathrm{j}2\pi fT}\right) = \mathscr{F}\left\{\mathrm{Im}\{a_{m[k]}\}\right\}$
  simplified representation for the case where $A_{\mathrm{I}}\left(\mathrm{e}^{\mathrm{j}2\pi fT}\right)$ and $A_{\mathrm{Q}}\left(\mathrm{e}^{\mathrm{j}2\pi fT}\right)$ begin real



Alternative: CAP

$g_{\mathrm{I}}(t)$ and $g_{\mathrm{Q}}(t)$ are real mutually orthogonal in–phase and quadrature transmit pulses



Conditions:

- For intersymbol interference–free signalling in in–phase and quadrature components: $g_{\mathrm{I}}(t)$ and $g_{\mathrm{Q}}(t)$ have to be $\sqrt{\text{Nyquist}}$–pulses (orthogonality w.r.t. to time shifts kT):

$$\int_{-\infty}^{\infty} g_{\mathrm{I}}(t+kT) \cdot g_{\mathrm{I}}(t) \, \mathrm{d}t = E_{g_{\mathrm{I}}} \cdot \delta_{0k} \; , \qquad k \in \mathbb{Z}$$

$$\int_{-\infty}^{\infty} g_{\mathrm{Q}}(t+kT) \cdot g_{\mathrm{Q}}(t) \, \mathrm{d}t = E_{g_{\mathrm{Q}}} \cdot \delta_{0k} \; , \qquad k \in \mathbb{Z}$$

- No crosstalk between in–phase and quadrature components:

$$\int_{-\infty}^{\infty} g_{\mathrm{I}}(t+\mu T) \cdot g_{\mathrm{Q}}(t) \, \mathrm{d}t = 0 \; , \qquad \mu \in \mathbb{Z}$$

  $\Rightarrow$ $g_{\mathrm{I}}(t)$ and $g_{\mathrm{Q}}(t)$ must be mutually orthogonal $\sqrt{\text{Nyquist}}$–pulses.

  (see twofold orthogonality condition for basis functions in chapter on signal space representation)

- In order to generate an RF signal, $g_{\mathrm{I}}(t)$ and $g_{\mathrm{Q}}(t)$ must have bandpass spectra centered at the carrier frequency f_0
  $\Rightarrow$ no mixing in the RF–band is necessary ($\Rightarrow$ "carrierless" AM/PM)

Possible choice of $g_{\mathrm{I}}(t)$ and $g_{\mathrm{Q}}(t)$:

a) Choose real, bandlimited (with respect to $f_{\max}$) $\sqrt{\text{Nyquist}}$–pulse $g(t)$ in the baseband, i.e.,

$$\int_{-\infty}^{\infty} g(t+kT) \cdot g(t) \, \mathrm{d}t = E_g \cdot \delta_{0k} \; , \qquad k \in \mathbb{Z}$$

and

$$G(f) = \mathscr{F}\{g(t)\} = 0,\ |f| > f_{\max}$$

b) With $f_0 > f_{\max}$:

$$\begin{aligned} g_{\mathrm{I}}(t) &= \sqrt{2}g(t) \cdot \cos(2\pi f_0 t) \\ g_{\mathrm{Q}}(t) &= -\sqrt{2}g(t) \cdot \sin(2\pi f_0 t) \end{aligned}$$

Note:

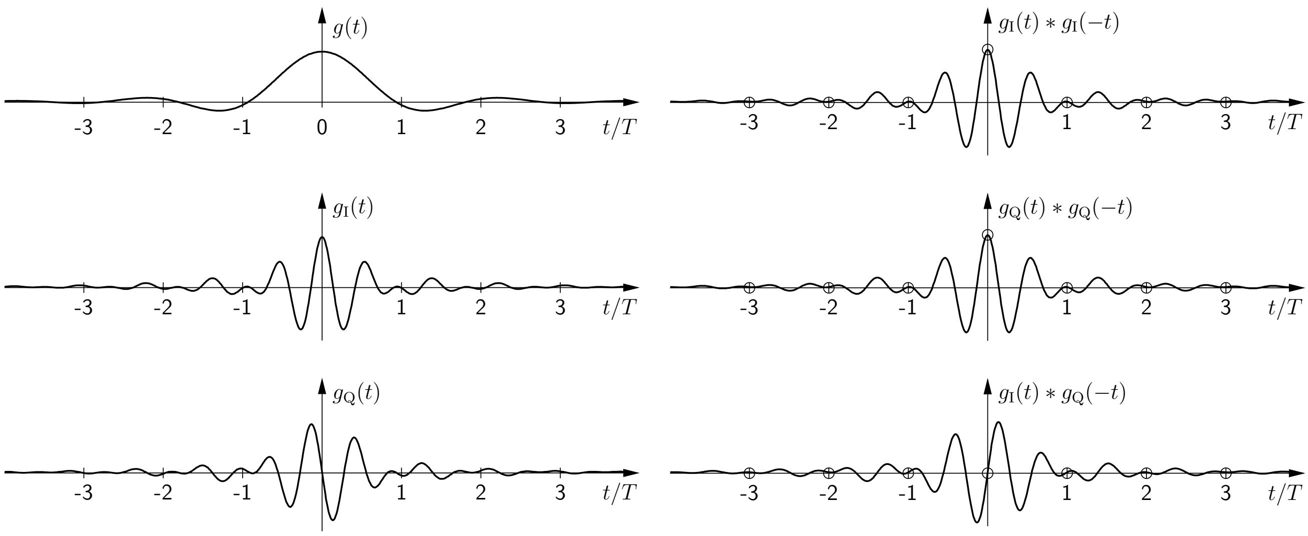
$g_{\mathrm{I}}(t)$ and $g_{\mathrm{Q}}(t)$ constitute a Hilbert transform pair

$$\begin{array}{ccc} \mathscr{H}\{g(t)\cos(2\pi f_0 t)\} & = & g(t)\sin(2\pi f_0 t) \\ \bullet\!\!-\!\!\circ & & \bullet\!\!-\!\!\circ \\ \frac{1}{2}\left(G(f-f_0)+G(f+f_0)\right)\cdot(-\mathrm{j}\cdot\mathrm{sign}(f)) & = & \frac{1}{2\mathrm{j}}\left(G(f-f_0)-G(f+f_0)\right) \end{array}$$

Note: $\mathscr{H}\{x(t)\} \,\widehat{=}\, \dfrac{1}{\pi}\displaystyle\int_{-\infty}^{+\infty}\frac{x(\tau)}{t-\tau}\mathrm{d}\tau$

Example:

- $g(t)$ is a $\sqrt{\cos}$–roll–off pulse (here: $\alpha = 0.3$)
- $f_0 = 1.8/T$



○ : Zero Crossings inforced

Proof: a) $g_\mathrm{I}(t+kT)$ orthogonal to $g_\mathrm{Q}(t)$, $k \in \mathbb{Z}$

b) $g_\mathrm{I}(t)$ and $g_\mathrm{Q}(t)$ are $\sqrt{\text{Nyquist}}$–pulses

$$\begin{aligned}
I &= \int_{-\infty}^{\infty} g_{\mathrm{I/Q}}(t+kT)\cdot g_{\mathrm{I/Q}}(t)\,\mathrm{d}t \qquad k\in\mathbb{Z}\\
&= \int_{-\infty}^{\infty} \sqrt{2}\,g(t+kT)\left\{\begin{matrix}\cos\\ -\sin\end{matrix}\right\}(2\pi f_0(t+kT))\cdot\sqrt{2}\,g(t)\left\{\begin{matrix}\cos\\ -\sin\end{matrix}\right\}(2\pi f_0 t)\,\mathrm{d}t\\
&= 2\Big(g(t)\left\{\begin{matrix}\cos\\ -\sin\end{matrix}\right\}(2\pi f_0 t) * g(-t)\left\{\begin{matrix}\cos\\ -\sin\end{matrix}\right\}(-2\pi f_0 t)\Big)\Big|_{t=kT}
\end{aligned}$$

$$\bullet\!\!-\!\!\circ$$

$$\begin{aligned}
&2\sum_i\Big(\left\{\begin{matrix}1/2\\ \mathrm{j}/2\end{matrix}\right\}\left(G(f-f_0-\tfrac{i}{T})\left\{\begin{matrix}+\\ -\end{matrix}\right\}G(f+f_0-\tfrac{i}{T})\right)\\
&\qquad\cdot\left\{\begin{matrix}1/2\\ -\mathrm{j}/2\end{matrix}\right\}\left(G^*(f-f_0-\tfrac{i}{T})\left\{\begin{matrix}+\\ -\end{matrix}\right\}G^*(f+f_0-\tfrac{i}{T})\right)\Big)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2}\left\{\begin{matrix}1\\ -\mathrm{j}\\ \mathrm{j}\\ 1\end{matrix}\right\}\sum_i\Big(|G(f-f_0-\tfrac{i}{T})|^2\left\{\begin{matrix}+\\ -\\ -\\ +\end{matrix}\right\}|G(f+f_0-\tfrac{i}{T})|^2\\
&\qquad\left\{\begin{matrix}+\\ +\\ -\\ -\end{matrix}\right\}G(f-f_0-\tfrac{i}{T})\cdot G^*(f+f_0-\tfrac{i}{T})\left\{\begin{matrix}+\\ -\\ +\\ -\end{matrix}\right\}G(f+f_0-\tfrac{i}{T})\cdot G^*(f-f_0-\tfrac{i}{T})\Big)
\end{aligned}$$

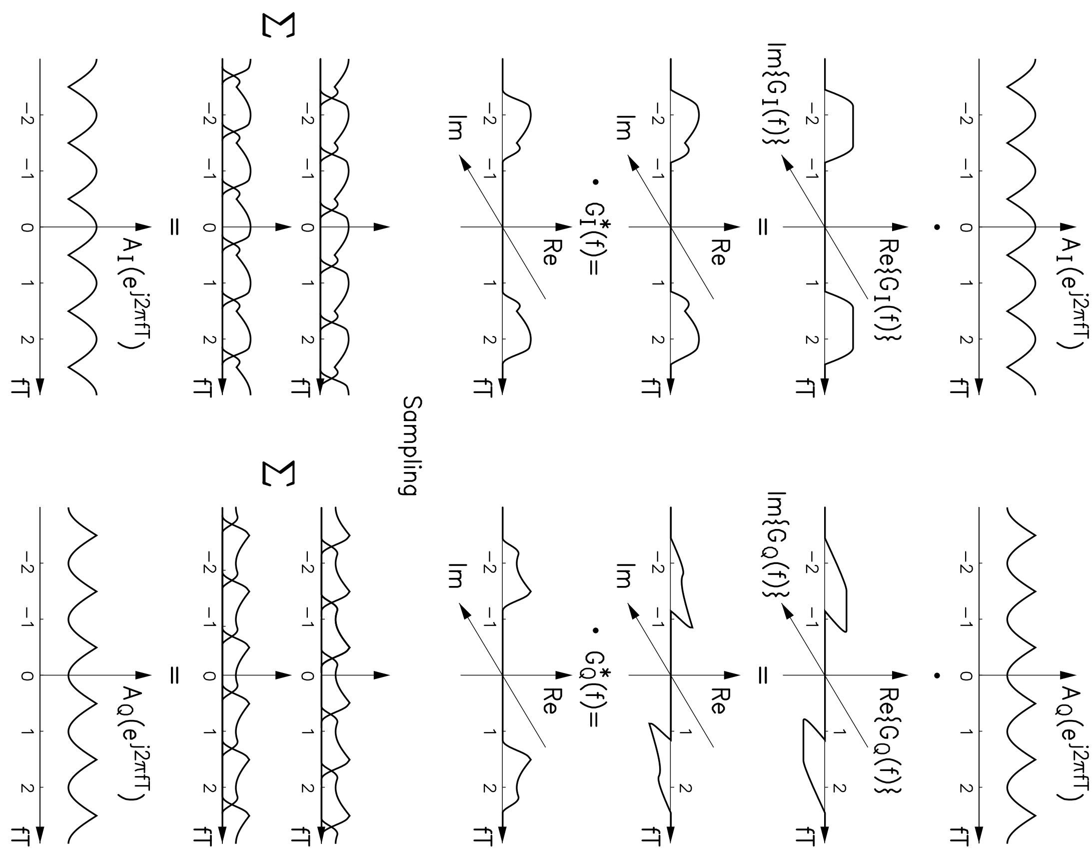
$$= \frac{1}{2}\left\{\begin{matrix}1\\ -\mathrm{j}\\ \mathrm{j}\\ 1\end{matrix}\right\}\left(E_\mathrm{g}\left\{\begin{matrix}+\\ -\\ -\\ +\end{matrix}\right\}E_\mathrm{g}\right) = \begin{cases} E_\mathrm{g} & \text{“I}\cdot\text{I”}\\ 0 & \text{“I}\cdot\text{Q”}\\ 0 & \text{“Q}\cdot\text{I”}\\ E_\mathrm{g} & \text{“Q}\cdot\text{Q”}\end{cases}$$

$$\circ\!\!-\!\!\bullet$$

$$I = \begin{cases} E_\mathrm{g}\cdot\delta_{0k} & \text{“I}\cdot\text{I”, “Q}\cdot\text{Q”}\\ 0 & \text{“I}\cdot\text{Q”, “Q}\cdot\text{I”}\end{cases}, \qquad k\in\mathbb{Z}$$

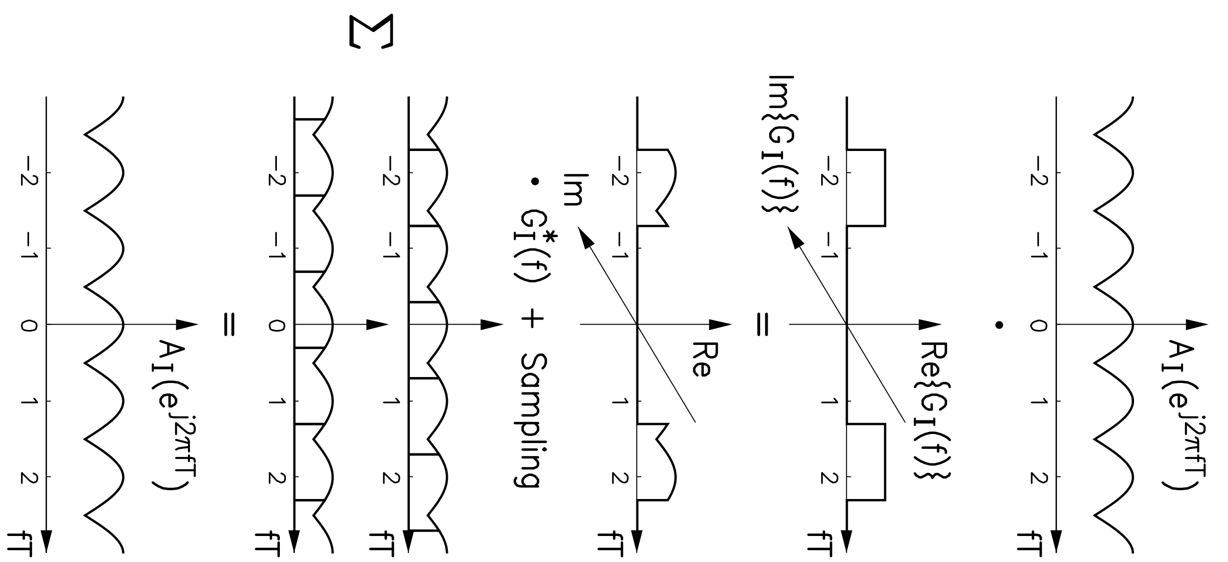
Interpretation: (corresponding spectra)

- $g(t)$ is a $\sqrt{\cos}$–roll–off pulse (here $\alpha = 0.3$)
- $f_0 = 1.8\ /T$
- Spectra of the real subsequences:
  $A_\mathrm{I}\left(\mathrm{e}^{\mathrm{j}2\pi fT}\right) = \mathscr{F}\left\{\mathrm{Re}\{a_{m[k]}\}\right\}$ and $A_\mathrm{Q}\left(\mathrm{e}^{\mathrm{j}2\pi fT}\right) = \mathscr{F}\left\{\mathrm{Im}\{a_{m[k]}\}\right\}$
  Simplified representation for real–valued $A_\mathrm{I}\left(\mathrm{e}^{\mathrm{j}2\pi fT}\right)$ and $A_\mathrm{Q}\left(\mathrm{e}^{\mathrm{j}2\pi fT}\right)$



200

Further simplified representation (only for in–phase branch)
for $\quad g(t) = \mathrm{si}(\pi t/T), \quad g_{\mathrm{I}}(t) = \sqrt{2}\,\mathrm{si}(\pi t/T)\cdot\cos(2\pi f_0 t)$



Discussion:

- In contrast to QAM, CAP requires no explicit mixing into the RF band. Typically, however, very low "carrier" frequencies f_0 are used (e.g. $f_0 = \frac{1}{T}$).

  The complete (RF) signal generation can be realized using digital signal processing.

- In the case of CAP, a local oscillator is not required at the receiver side for shifting the signal into the baseband.

  $\Rightarrow$ *no carrier synchronization necessary!*

  However: Highly accurate symbol clock synchronization is required.
  Symbol clock phase errors result in crosstalk between in–phase and quadrature components!

- CAP can be precisely demodulated using bandpass filtering and sampling if the following holds:

$$f_0 = \frac{k}{T} \qquad k \in \{1, 2, 3, \ldots\}$$

- Interpretation for transmitter pulses with rectangular spectra of width $\frac{1}{T}$:
  - average power spectral density of the sequence of amplitude coefficients $a_{m[k]}$ is periodic in the frequency with period $\frac{1}{T}$
  - QAM selects the baseband part of width $\frac{1}{T}$ and shifts it to frequency f_0.
  - CAP selects a frequency band of width $\frac{1}{T}$ around an arbitrary frequency f_0.

204

5 Signal Space Representation, General Digital Communications Schemes, Coherent and Incoherent Demodulation

General digital transmission scheme, see Chapter 2:

- Transmit signal (ECB–domain):

$$s(t) = \sum_{k=-\infty}^{+\infty} s_{m[k]}(t - kT)$$

- Modulation interval: T
- Number of modulation levels: M
- Signal set $\mathcal{S} = \{s_0(t), s_1(t), \ldots, s_{M-1}(t)\}$
  ($\widehat{=}$ M–ary modulation scheme)

- **Condition** that has to be fulfilled so that detection can be carried out independently from modulation intervals:
  Signal elements $s_i(t)$ must fulfill the temporal orthogonality condition

$$\int_{-\infty}^{+\infty} s_i(t + kT) \cdot s_\ell^*(t) \, \mathrm{d}t = \begin{cases} E_{i\ell}, & k = 0 \\ 0, & k \neq 0 \end{cases} \quad , \quad \forall i,\ell \in \{0, 1, \ldots, M-1\}.$$

  See Chapter 2, separation of coding and modulation

5.1 Signal Representation over Orthonormal Basis Functions

For arbitrary modulation schemes used in digital communications, the elements $s_m(t)$ of the signal set $\mathcal{S}$ – the signal elements – can be represented as linear combinations of D **doubly orthogonal basis functions**

$$g_l(t)\ , \qquad l \in \{0, 1, \ldots,\ D-1\},$$

which fulfill the condition

$$\frac{1}{E_\mathrm{g}} \int\limits_{-\infty}^{+\infty} g_l(t+kT) g_\ell^*(t)\,\mathrm{d}t = \delta_{l\ell} \cdot \delta_{0k}\ , \quad l,\ell \in \{0,1,\cdots, D-1\}\ , \quad k \in \mathbb{Z}$$

with $\delta_{l\ell} = \begin{cases} 1 & \text{for } l = \ell \\ 0 & \text{for } l \neq \ell \end{cases}$ (Kronecker symbol)

$\Rightarrow$ *double orthogonality with respect to l and kT!*

E_g: Energy of the basis function (a normalization factor)

Representation of the signal elements of an arbitrary modulation interval as:

linear combination of basis functions:
$$s_i(t) = \sum_{l=0}^{D-1} s_{l,i}\, g_l(t) \qquad \longleftrightarrow \qquad \boldsymbol{s}_i = [s_{0,i}, \ldots,\ s_{D-1,i}]$$
$$i \in \{0, 1, \ldots,\ M-1\}$$

with the *linear weights* $s_{l,i}$ of the i^{th} signal element w.r.t. the l^{th} dimension. The choice of the basis function is almost arbitrary, but it is generally advantageous to chose the basis functions such that the *dimensionality* D is as small as possible. D should not be larger than the dimensionality of the signal space spanned by the M functions $s_i(t)$. This condition is assumed to be fulfilled in the following analysis, and therefore D is known as the *dimensionality of the modulation scheme per modulation step*.

equivalent representation $\boxed{\textbf{time function} \longleftrightarrow \textbf{signal points}}$

D: Dimensions per modulation step

ND: Dimensions for N modulation steps

Due to the temporal orthogonality, the subspaces for each individual modulation step in the ND dimensional signal space are mutually orthogonal for sequences of signal elements.

Note: Digital PAM is a special case with **one** (complex) dimension per modulation step, i.e., $D = 1$. The PAM pulse $g(t)$ is equal to $g_0(t)$.

Calculation of the *linear weights* $s_{l,i}$:

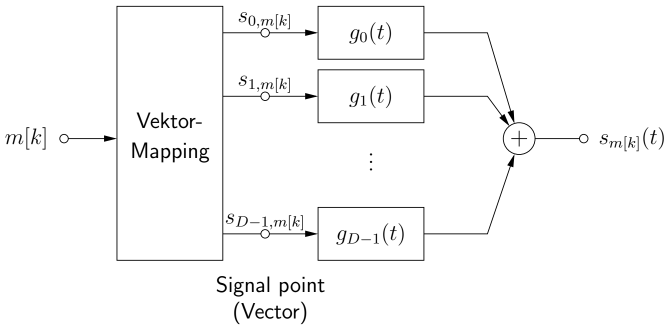
$$s_{l,i} = \frac{1}{E_\mathrm{g}} \int\limits_{-\infty}^{+\infty} s_i(t)\, g_l^*(t)\,\mathrm{d}t = \frac{1}{E_\mathrm{g}} \int\limits_{-\infty}^{+\infty} \sum_{\ell=0}^{D-1} s_{\ell,i}\, g_\ell(t)\, g_l^*(t)\,\mathrm{d}t = \sum_{\ell=0}^{D-1} s_{\ell,i}\, \frac{1}{E_\mathrm{g}} \int\limits_{-\infty}^{+\infty} g_\ell(t)\, g_l^*(t)\,\mathrm{d}t = \sum_{\ell=0}^{D-1} s_{\ell,i} \cdot \delta_{\ell l} = s_{l,i}$$

The mapping of a signal number m to a signal element $s_m(t)$

is equivalent to

the mapping of a signal number m to a **signal vector** $\boldsymbol{s}_m = [s_{m,0}, s_{m,1}, \ldots, s_{m,(D-1)}]$

Generalization of PAM to general digital transmission schemes – interpretation as D–fold parallel PAM transmission:



Representation with vectors and matrices:

- Vector of the basis functions:

$$\boldsymbol{g}(t) \mathrel{\widehat{=}} [g_0(t),\, g_1(t),\ \ldots,\ g_{D-1}(t)] \qquad \text{(row vector)}$$

- Double orthogonality condition:

$$\frac{1}{E_\mathrm{g}} \int\limits_{-\infty}^{+\infty} \left[\boldsymbol{g}^\mathsf{T}(t+kT) \cdot \boldsymbol{g}^*(t)\right]\,\mathrm{d}t = \boldsymbol{I}_{D\times D} \cdot \delta_{0k}$$

- Vector of signal elements:

$$\boldsymbol{s}(t) \mathrel{\widehat{=}} [s_0(t),\, s_1(t), \ldots,\ s_{M-1}(t)] \qquad \text{(row vector)}$$

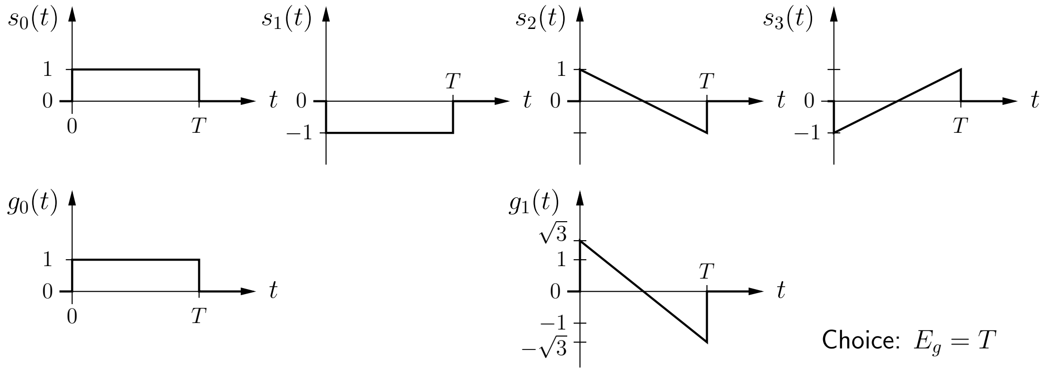
- Matrix of linear weights ($D \times M$ matrix):

$$\boldsymbol{S} = \left[\, s_{l,i} \,\right]$$

- Relationship:

$$\boldsymbol{s}(t) = \boldsymbol{g}(t) \cdot \boldsymbol{S}$$

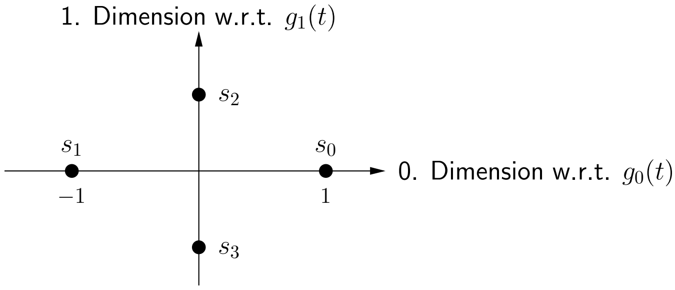
Example: Signal elements, basis functions, transmit signal



Choice: $E_g = T$

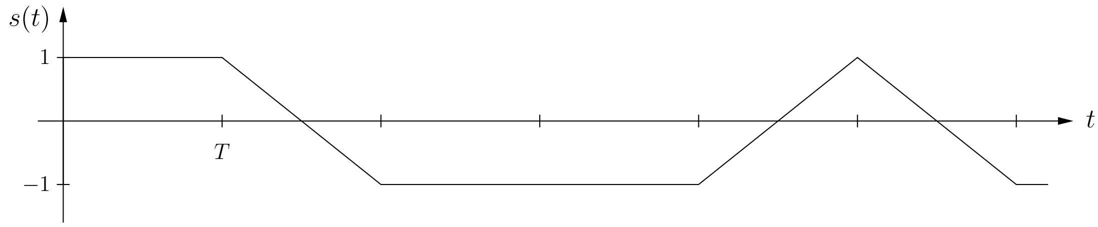
$$s_0(t) = g_0(t), \qquad s_1(t) = -g_0(t), \qquad s_2(t) = \frac{1}{\sqrt{3}} g_1(t), \qquad s_3(t) = -\frac{1}{\sqrt{3}} g_1(t)$$

Constellation diagram



212

Transmit signal



$m[k]$	0	2	1	1	3	2	1
$s_{0,m[k]}$	1	0	-1	-1	0	0	-1
$s_{1,m[k]}$	0	$\frac{1}{\sqrt{3}}$	0	0	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	0

5.2 Gram–Schmidt Procedure

Problem statement: How can we obtain an orthonormal set of basis functions $\boldsymbol{g}_l(t)$ from a given set of signal elements $\boldsymbol{s}_i(t)$?

Observation: There are many ways to solve this problem, each with different characteristics.

Example:

For signal elements limited to the modulation interval T, their corresponding

Fourier series expansion

$$s_i(t) \;=\; \sum_{l=-\infty}^{+\infty} s_{l,i} \mathrm{e}^{\mathrm{j}2\pi i t/T} \text{ for } t \in [0,T)$$

with the (complex) basis functions

$$g_l(t) \;=\; \begin{cases} \mathrm{e}^{\mathrm{j}2\pi i t/T} & \text{for } t \in [0,T) \\ 0 & \text{for } t \in \mathbb{R}\backslash[0,T) \end{cases}$$

is a representation in terms of orthonormal basis functions, but usually with a very high number of dimensions

One solution (which in most cases is not unique) consists of generating orthonormal basis functions using *linear combinations* of the signal elements themselves:

$$g_l(t) = \sum_{i=0}^{M-1} g_{i,l}\, s_i(t)$$

with *orthogonalization factors* $g_{i,l}$.

This can be accomplished with the help of the Gram–Schmidt process, which was originally developed to obtain an orthonormal (Cartesian) basis with minimal dimension for a given set of vectors.

Vector notation:

- Matrix of orthogonalization factors ($M \times D$ matrix):

$$\boldsymbol{G} = \left[\, g_{i,l} \,\right]$$

- Relationship:

$$\boldsymbol{g}(t) = \boldsymbol{s}(t) \cdot \boldsymbol{G}$$

- As $\boldsymbol{s}(t) = \boldsymbol{g}(t) \cdot \boldsymbol{S}$ is valid, we have:

$$\boldsymbol{g}(t) = \boldsymbol{g}(t) \cdot \boldsymbol{S} \cdot \boldsymbol{G}$$

or alternatively $\boldsymbol{S} \cdot \boldsymbol{G} = \boldsymbol{I}_{D \times D}$

- *Cross–energy matrix of signal elements:*

$$\boldsymbol{E} = \left[\, E_{il} \,\right] = \left[\int\limits_{-\infty}^{+\infty} s_i(t) s_l^*(t)\, \mathrm{d}t \right] = \int\limits_{-\infty}^{+\infty} \left[\boldsymbol{s}^{\mathsf{T}}(t) \boldsymbol{s}^*(t) \right] \, \mathrm{d}t$$

Note:

1. The elements E_{ii} on the main diagonal of the cross–energy matrix are equal to the energy E_i of the M signal elements $E_i = \int s_i(t) s_i^*(t)\, \mathrm{d}t = E_{ii}$
2. The cross–energy matrix is Hermitian, i.e., $\boldsymbol{E} = \boldsymbol{E}^{\mathsf{H}} = (\boldsymbol{E}^*)^{\mathsf{T}}$ holds. Therefore, $\boldsymbol{E}$ possesses real, non–negative eigenvalues and the modal matrices (the matrices of the left or right eigenvectors) are unitary.

Using $\boldsymbol{s}(t) = \boldsymbol{g}(t)\boldsymbol{s}$, we obtain

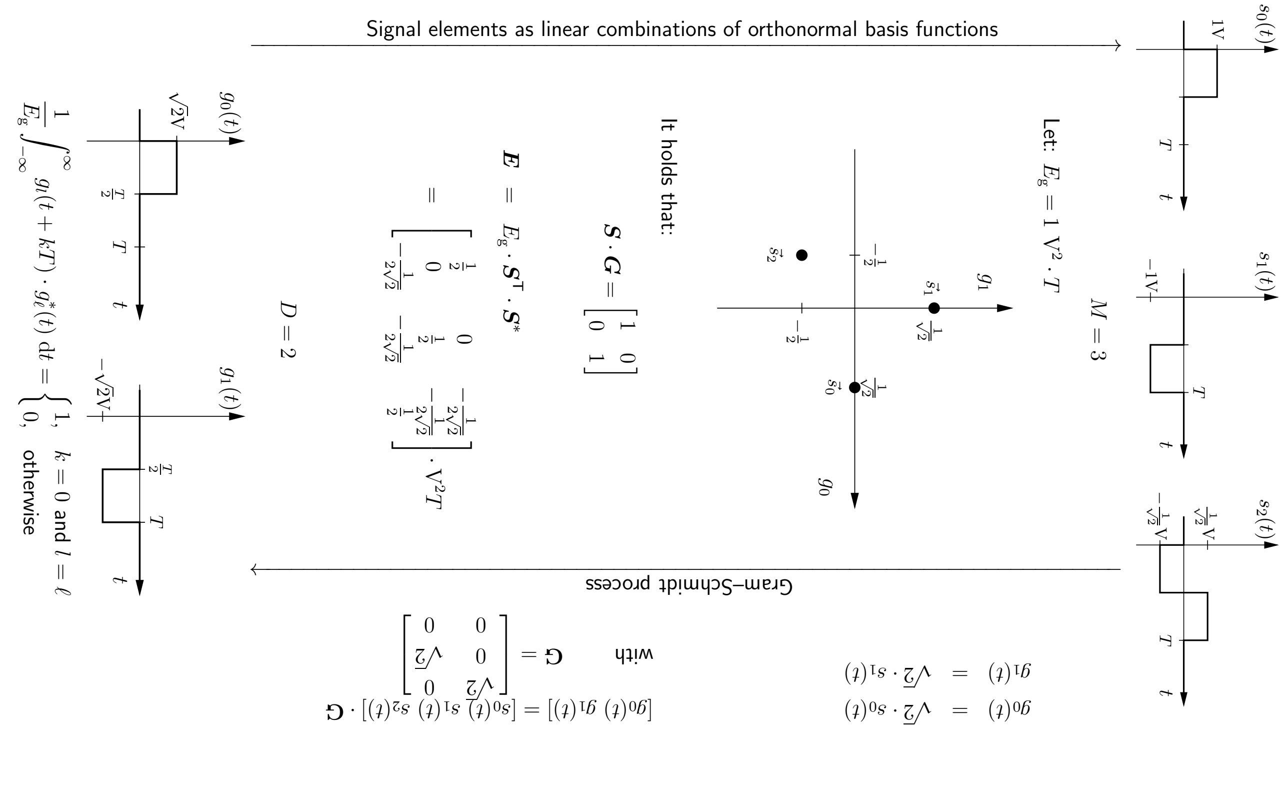
$$\boldsymbol{E} = E_{\mathrm{g}} \cdot \boldsymbol{S}^{\mathsf{T}} \cdot \boldsymbol{S}^*$$

Example:

$$\begin{aligned} s_0(t) &= \tfrac{1}{\sqrt{2}} \cdot g_0(t) \\ s_1(t) &= \tfrac{1}{\sqrt{2}} \cdot g_1(t) \\ s_2(t) &= -\tfrac{1}{2} \cdot (g_0(t) + g_1(t)) \end{aligned}$$

$[s_0(t)\ s_1(t)\ s_2(t)] = [g_0(t)\ g_1(t)] \cdot \boldsymbol{S}$

mit $\boldsymbol{S} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} \end{bmatrix}$



Discussion of Units (see related literature):

Example: $s_i(t)$ and thus $s(t)$ and $n(t)$ have units of signals, e.g. V (volt); t has the unit s (second);
Thus, N_0 and E_g have unit of (normalized) energy e.g. V^2/Hz=V^{2}s

In the literature, the unit of the corresponding basis functions is typically $\frac{1}{\sqrt{\mathrm{s}}}$!

$$\frac{1}{E_\mathrm{g}}\int\limits_{-\infty}^{\infty} g_l(t+kT)\cdot g_\ell^*(t)\,\mathrm{d}t \quad \Bigg| \quad \int\limits_{-\infty}^{\infty} g_l(t+kT)\cdot g_\ell^*(t)\,\mathrm{d}t$$

$$= \begin{cases} 1, & k=0 \text{ and } l=\ell \\ 0, & \text{otherwise} \end{cases}, \quad k\in\mathbb{Z}, \quad l,\ell = 0\,(1)\,D-1$$

Units of the corresponding components in the literature is typically V$\sqrt{\mathrm{s}}$!

$$d_l[k] = \frac{1}{E_\mathrm{g}}\int\limits_{-\infty}^{\infty} e(t)\cdot g_l^*(t-kT)\,\mathrm{d}t \quad \Bigg| \quad d_l[k] = \int\limits_{-\infty}^{\infty} e(t)\cdot g_l^*(t-kT)\,\mathrm{d}t$$

Unit	Parameter	mostly in literature
V	$g_l(t)$	$\frac{1}{\sqrt{\mathrm{s}}}$
–	$d_l[k]$, $s_{l,m[k]}$, $n_l[k]$	$\mathrm{V}\sqrt{\mathrm{s}}$
–	σ^2	$\mathrm{V}^2\mathrm{s}$
$\sigma^2 = \frac{N_0}{E_\mathrm{g}}$		$\sigma^2 = N_0$

Note: Often in the literature, a variance, specifying a signal power has unit of energy (very confusing and source of a numerous errors in literature!)

Gram–Schmidt Procedure:

Recursive calculation of the

- *Linear weights* (linear weight matrix $\boldsymbol{S}$) and the
- *Orthogonalization factors* (orthogonalization matrix $\boldsymbol{G}$)

1. **First basis function:**

$$g_0(t) = \sqrt{\frac{E_\mathrm{g}}{E_0}}\, s_0(t)$$

i.e.,

$$g_{00} = \frac{1}{s_{00}} = \sqrt{\frac{E_\mathrm{g}}{E_0}}\,, \qquad \begin{array}{ll} g_{i,0} = 0\,, & i = 1,2,\ldots,M-1\,, \\ s_{l,0} = 0\,, & l = 1,2,\ldots \end{array}$$

The energy of the first signal element is

$$E_0 = \int\limits_{-\infty}^{+\infty} |s_0(t)|^2 \;\mathrm{d}t$$

2. **Recursion:** Let the l basis functions for the representation of the first $i \geq l$ signal elements be $s_0(t), s_1(t), \ldots,\; s_{i-1}(t)$.

Representation of the next signal element $s_i(t)$:

a) Determination of the linear factors $s_{\ell,i}$ with respect to the the already established first l basis functions $g_\ell(t)$, $\ell = 0{,}1, \ldots,\; l-1$,

$$s_{\ell,i} = \frac{1}{E_\mathrm{g}} \int\limits_{-\infty}^{+\infty} s_i(t)\, g_\ell^*(t)\;\mathrm{d}t.$$

With the cross energies $E_{in} = \int\limits_{-\infty}^{+\infty} s_i(t)\, s_n^*(t)\;\mathrm{d}t$, the following representation holds

$$\begin{aligned} s_{\ell,i} \;&=\; \frac{1}{E_\mathrm{g}} \int\limits_{-\infty}^{+\infty} s_i(t)\, g_\ell^*(t)\;\mathrm{d}t \;=\; \frac{1}{E_\mathrm{g}} \int\limits_{-\infty}^{+\infty} s_i(t) \sum_{n=0}^{i-1} g_{n,\ell}^*\, s_n^*(t)\;\mathrm{d}t \\ &=\; \frac{1}{E_\mathrm{g}} \sum_{n=0}^{i-1} g_{n,\ell}^* \int\limits_{-\infty}^{+\infty} s_i(t)\, s_n^*(t)\;\mathrm{d}t = \frac{1}{E_\mathrm{g}} \sum_{n=0}^{i-1} g_{n,\ell}^*\, E_{in} \end{aligned}$$

Auxiliary calculation:

Determination of the energy of a signal element, assuming all basis functions are known:

$$
\begin{aligned}
E_i &= \int_{-\infty}^{+\infty} |s_i(t)|^2 \, \mathrm{d}t \\
&= \int_{-\infty}^{+\infty} \left(\sum_{l=0}^{D-1} s_{l,i}\, g_l(t) \right) \left(\sum_{\ell=0}^{D-1} s^*_{\ell,i}\, g^*_\ell(t) \right) \mathrm{d}t \\
&= E_\mathrm{g} \sum_{l=0}^{D-1} \sum_{\ell=0}^{D-1} s_{l,i}\, s^*_{\ell,i} \; \underbrace{\frac{1}{E_\mathrm{g}} \int_{-\infty}^{+\infty} g_l(t)\, g^*_\ell(t)\, \mathrm{d}t}_{\delta_{l,\ell}} \\
&= E_\mathrm{g} \cdot \sum_{l=0}^{D-1} |s_{l,i}|^2 \\
&= E_\mathrm{g} \cdot ||\boldsymbol{s}_i||^2 \;=\; E_\mathrm{g} \cdot \boldsymbol{s}_i \cdot \boldsymbol{s_i}^\mathrm{H} \qquad \text{(Parseval's Theorem!)}
\end{aligned}
$$

The energy of a signal is proportional to squared norm (metric) of the vector, representing the signal in a signal space, spanned by orthogonal basis functions.

b) **Test**: Can the signal element $s_i(t)$ be completely represented using the first l basis functions $g_\ell(t)$?

To test this we calculate: $\Delta E_i \overset{\text{def}}{=} E_i - E_\mathrm{g} \sum\limits_{\ell=0}^{l-1} |s_{\ell,i}|^2$

A: $\Delta E_i = 0$

The $(i+1)$–th signal element can be completely represented by the already established l signal elements.

$$s_{\ell,i} = 0 \,, \qquad \ell = l, l+1, \ldots$$

Go to **3.**

B: $\Delta E_i > 0$

For the complete representation of $s_i(t)$, a new additional basis function $g_l(t)$ is needed:

$$g_l(t) = \sqrt{\frac{E_\mathrm{g}}{\Delta E_i}} \; \underbrace{\left(s_i(t) - \sum_{\ell=0}^{l-1} s_{\ell,i}\, g_\ell(t) \right)}_{\text{orthogonal residual with energy } \Delta E_i}$$

i.e.,

$$s_{l,i} = \sqrt{\Delta E_i / E_\mathrm{g}} \,, \qquad s_{\ell,i} = 0 \,, \qquad \ell = l+1, l+2, \ldots$$

By this, the entire i^{th} column of the matrix $\boldsymbol{S}$ is known, and the corresponding orthogonalization factors are immediately seen to be

$$g_{i,l} = \sqrt{\frac{E_\mathrm{g}}{\Delta E_i}}$$

The remaining orthogonalization factors $g_{n,l}$, $n = 0,\ 1, \ldots,\ i-1$, are still missing and are needed to represent the $g_l(t)$ as a linear combination of the first i signal elements. To this end, we consider $g_l(t) = \sum\limits_{n=0}^{i} g_{n,l}\, s_n(t)$ alternatively,

$$\begin{aligned} g_l(t) - \sqrt{\tfrac{E_\mathrm{g}}{\Delta E_i}} s_i(t) &= \sum_{n=0}^{i-1} g_{n,l}\, s_n(t) \\ &= -\sqrt{\tfrac{E_\mathrm{g}}{\Delta E_i}} \sum_{\ell=0}^{l-1} s_{\ell,i}\, g_\ell(t) = -\sqrt{\tfrac{E_\mathrm{g}}{\Delta E_i}} \sum_{\ell=0}^{l-1} s_{\ell,i} \sum_{n=0}^{i-1} g_{n,\ell}\, s_n(t) \\ &= \sum_{n=0}^{i-1} \left(-\sqrt{\tfrac{E_\mathrm{g}}{\Delta E_i}} \sum_{\ell=0}^{l-1} s_{\ell,i}\, g_{n,\ell} \right) s_n(t) \end{aligned}$$

Consequently, the orthogonalization factors are obtained as

$$g_{n,l} = -\sqrt{\tfrac{E_\mathrm{g}}{\Delta E_i}} \sum_{\ell=0}^{l-1} s_{\ell,i}\, g_{n,\ell}$$

alternatively,

$$g_{n,l} = -\sqrt{\tfrac{1}{E_\mathrm{g}\, \Delta E_i}} \sum_{m=0}^{i-1} E_{im} \sum_{\ell=0}^{l-1} g_{n,\ell}\, g^*_{m,\ell} \qquad n = 0, 1, \ldots, i-1$$

Through this, the entire column l of $\mathbf{G}$ is defined!

Increment the recursion: $l + 1 \to l$

3. **Test:**
   $i = M - 1$? Yes: Terminate recursion
   No: $i + 1 \to i$, next step; Go to **2.**
4. **Termination of recursion:** Dimensionality $D = l$

Discussion:

- Using Gram–Schmidt, the matrices $\boldsymbol{S}$ and $\boldsymbol{G}$ take the form of *upper triangular matrices*.
- Interpretation of Gram–Schmidt process with respect to the cross–energy matrix:

$$\begin{aligned}
\boldsymbol{E} &= \left[E_{in}\right] = \int_{-\infty}^{+\infty} \left[\boldsymbol{s}^{\mathsf{T}}(t)\, \boldsymbol{s}^{*}(t)\right]\, \mathrm{d}t \\
&= \int_{-\infty}^{+\infty} \left[(\boldsymbol{g}(t)\boldsymbol{S})^{\mathsf{T}} \cdot (\boldsymbol{g}^{*}(t)\boldsymbol{S}^{*})\right]\, \mathrm{d}t \\
&= E_{\mathrm{g}} \cdot \boldsymbol{S}^{\mathsf{T}} \cdot \underbrace{\frac{1}{E_{\mathrm{g}}} \int_{-\infty}^{+\infty} \left[\boldsymbol{g}^{\mathsf{T}}(t) \cdot \boldsymbol{g}^{*}(t)\right]\, \mathrm{d}t}_{=\boldsymbol{I}_{D\times D}} \cdot \boldsymbol{S}^{*} \\
&= E_{\mathrm{g}} \cdot \boldsymbol{S}^{\mathsf{T}} \cdot \boldsymbol{S}^{*}
\end{aligned}$$

  Decomposition of the Hermitian matrix $\boldsymbol{E}$ into triangular matrices: *Choleski factorization*

$$\boldsymbol{E} = E_{\mathrm{g}} \cdot \boldsymbol{S}^{\mathsf{T}} \cdot \boldsymbol{S}^{*}$$

  Dimensionality: $D =$ Rank of $\boldsymbol{E}$

Note: The basis function system $\boldsymbol{g}(t)$ provided by the Gram–Schmidt process is dependent on the order in which the signal elements $s_i(t)$ are considered by the algorithm. A different ordering of signal elements will yield different basis functions.

Note: The representation of a signal in an incomplete set of basis functions is advantageous in some applications. With an incomplete set, the computational load can be reduced by discarding dimensions which only contain small amounts of energy, thus creating an (incomplete) *rank reduced signal representation*.
In this case, for example, a *sorted* Gram–Schmidt process may be used, for which the signal elements are considered by the algorithm in order of decreasing energy such that the signal elements with less energy are sufficiently represented by the previously calculated basis functions. This yields linear weights which become very small for higher dimensions, allowing these dimensions to be discard without significant loss (e.g. with very little increase in error probability during detection/decoding).

5.3 Representation of Received Signals

Without loss of generality, the **noise–free** part of the receiver input signal $\tilde{e}(t)$ can be represented as a sum of signal elements $s_i(t)$ which fulfill the double orthogonality condition, see Section 2.1.2.1. In the case of distortions (linear or nonlinear), an equivalent coding and modulation (i.e. a modified set of signal elements) must be defined to describe the noise–free part of the receiver input signal (see Section 2.1.2.1.3). Thus, all definitions discussed in the previous sections also apply to any noise–free part of the receiver input signal.

For a **complete** representation of the receiver input signal

$$e(t) = \tilde{e}(t) + n(t)$$

many more than D basis functions $g_l(t)$, $l \in \{0, 1, \ldots,\ D-1, D, D+1, \ldots,\ \Delta - 1\}$, seem to be required for a complete representation of signal and noise $n(t)$, i.e., $\Delta \gg D$.

$$e(t) \stackrel{!}{=} \sum_{k=-\infty}^{+\infty} \sum_{l=0}^{\Delta} d_l[k]\, g_l(t - kT)$$

where $d_l[k]$ denotes the l^{th} component of the received vector.

Def.: *Received vector* with respect to the k^{th} modulation step

$$\Big[\, d_0[k], d_1[k], \ldots,\ d_{D-1}[k], \ldots,\ d_{\Delta-1}[k] \,\Big]$$

Note: In the case of true "white noise" (AWGN), $\Delta \to \infty$ would approach infinity! This happens because the number of dimensions in the signal space is proportional to the time–bandwidth product of the signal to be represented. [1]

With $\tilde{e}(t) = \sum\limits_{k=-\infty}^{+\infty} \sum\limits_{l=0}^{\Delta} s_{l,m[k]}\, g_l(t - kT)$ and the double orthogonality condition (S. 207) the components $d_l[k]$ of the detection vector can be calculated as

$$d_l[k] = \frac{1}{E_g} \int\limits_{-\infty}^{+\infty} e(t) g_l^*(t - kT) \mathrm{d}t = s_{l,m[k]} + n_l[k],$$

where

$$\begin{array}{rcccll} d_l[k] & = & s_{l,m[k]} & + & n_l[k] & \text{for } l = 0, 1, \ldots,\ D-1 \\ d_l[k] & = & 0 & + & n_l[k] & \text{for } l = D, D+1, \ldots,\ \Delta - 1 \end{array}$$

Theorem of irrelevant data from information theory:

Parameters Y_i, which do not have statistical dependencies with the transmitted information X and, for which therefore, mutual information $I(X; Y_i|Z) \equiv 0$ for any condition Z is zero, are useless for detection of X and therefore irrelevant for the receiver.

[1]Note: The theorem of Dollard (1964) states: For all elements $s_i(t)$ of a signal set $\mathbf{S}$, having 90% of their energy concentrated within a frequency band $[-W, +W]$, $D \leq \lceil 2{,}22WT \rceil$ basis functions provide a sufficiently accurate representation.

Proof: Chain rule of mutual information

$$\begin{aligned} I\left(X;(Y_1,Y_2,\ldots,Y_i,\ldots,Y_L)\right) &= I\left(X;(Y_1,Y_2,\ldots,Y_{i-1},Y_{i+1},\ldots,Y_L)\right) \\ &\quad + \underbrace{I\left(X;Y_i|\,(Y_1,Y_2,\ldots,Y_{i-1},Y_{i+1},\ldots,Y_L)\right)}_{=0} \\ &= I\left(X;(Y_1,Y_2,\ldots,Y_{i-1},Y_{i+1},\ldots,Y_L)\right) \end{aligned}$$

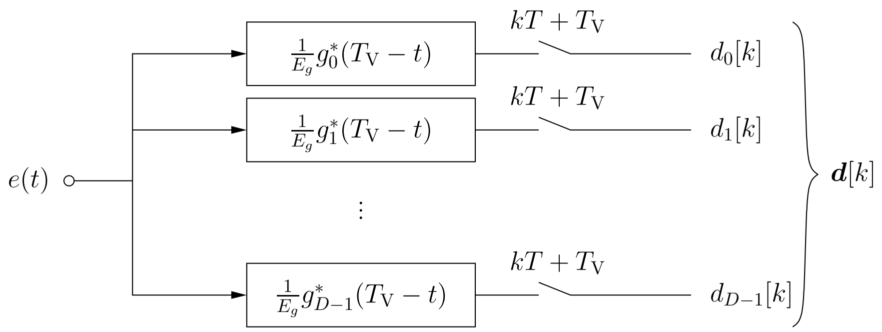
$\Rightarrow$ Only those components $d_l[k]$ of the received vector for which a noiseless part exists (i.e., for $l = 0, 1, \ldots,\ D-1$), are **necessary and sufficient** for an **optimal** detection or decoding. Additional components are irrelevant and therefore need not be calculated at the receiver.

Definition: Received Vector

$$\boldsymbol{d}[k] = \Big[\, d_0[k], d_1[k], \ldots,\ d_{D-1}[k] \,\Big]$$

The sequence of received vectors $d[k], k \in \mathbb{Z}$, provide sufficient statistics for the receiver input signal w.r.t data estimation (not for complete reconstruction of the receiver input signal including noise).

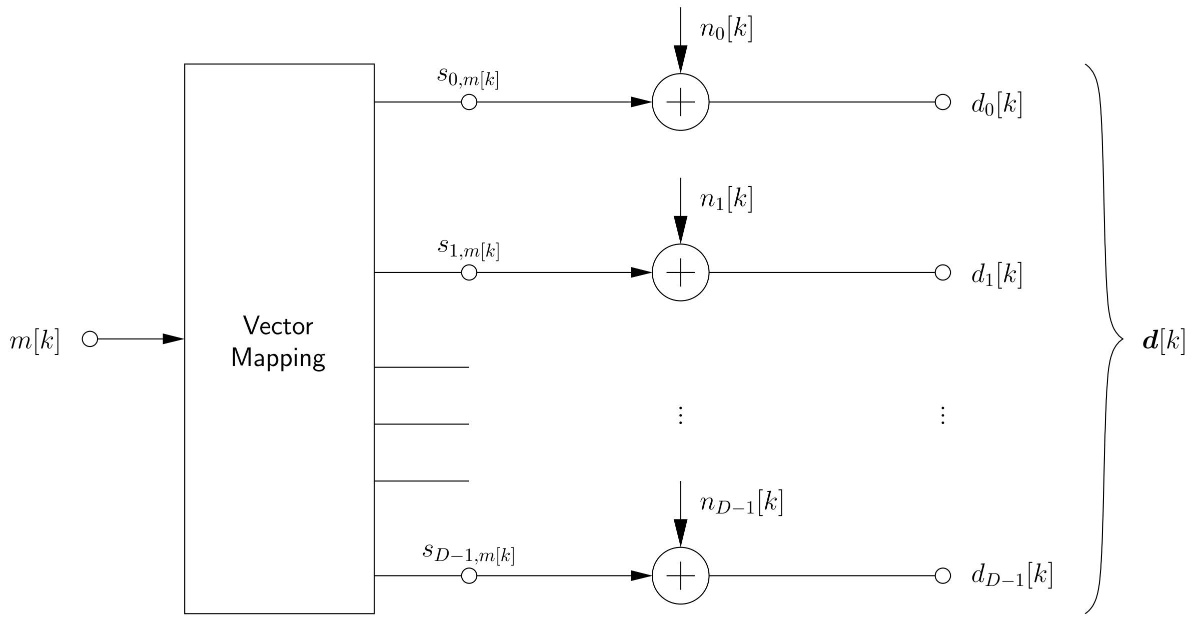
General Vector Receiver:



Vector Demodulator

The bank of D parallel matched filters with respect to the D basis functions together with sampling at the symbol rate (baud rate; T_V: delay to enable causal implementation of the matched filter).

Equivalent Vector Channel Model (ECB domain):



Note: For $D = 1$, (i.e., digital PAM), a filter with respect to the fundamental pulse $g(t) = g_0(t)$ is sufficient (i.e., a matched filter for the fundamental pulse). Alternative derivation of the matched filter.

232

Variance of the noise components $n_l[k]$ in the received vector in case of white Gaussian noise with (one–sided) noise power spectral density N_0 (two sided in the case of ECB signal representation: AWGN ECB):

$$\sigma^2 = N_0 \cdot \delta(\tau) * \frac{1}{E_\mathrm{g}} g_\ell(\tau) * \frac{1}{E_\mathrm{g}} g_\ell^*(-\tau)\bigg|_{\tau=0} = N_0 \underbrace{g_\ell(\tau) * g_\ell^*(-\tau)|_{\tau=0}}_{E_\mathrm{g}} / E_\mathrm{g}^2 = N_0/E_\mathrm{g}, \qquad \ell \in \{0, \dots,\ D-1\}$$

$\Rightarrow$ Due to the double orthogonality of the basis function $g(t)$, all detector noise values $n_l[k]$ at any discrete time k and dimension l are mutually uncorrelated. Therefore, for the AWGN channel, they are also statistically independent! and have the same variance N_0/E_g.

$$\begin{aligned} \mathrm{E}\{n_l[k] \cdot n_\ell^*[k+\kappa]\} &= N_0/E_g \cdot \delta[\kappa] \cdot \delta[l-\ell] \\ &= \begin{cases} N_0/E_g & \text{for } \kappa = 0, l = \ell, \\ 0 & \text{for } \kappa \neq 0, l \neq \ell, \end{cases} \quad \forall\, k \in \mathbb{Z};\ l, \ell \in \{0, 1, \dots,\ D-1\} \end{aligned}$$

Definition: *Relevant Part of the Receiver Input Signal*

$$\breve{e}(t) \;=\; \sum_{k=-\infty}^{+\infty} \sum_{l=0}^{D-1} d_l[k]\, g_l(t-kT) = \tilde{e}(t) + \breve{n}(t)$$

with **relevant part of noise**

$$\breve{n}(t) \;=\; \sum_{k=-\infty}^{+\infty} \sum_{l=0}^{D-1} n_{l,k}\, g_l(t-kT)$$

"Part of the noise which actually interfers with the signal"

Interpretation: $\breve{n}(t)$ can be thought of as a good approximation for the part of the noise which overlaps in the frequency domain with the band occupied by the noise–free part of the receiver input signal. The bank of matched filters limits the noise bandwidth and thus the noise variance. Only the relevant dimensions have to be further processed.

5.4 Optimal Coherent Demodulation and Decoding

5.4.1 Maximum a Posteriori and Maximum Likelihood Sequence Estimation

General case: Transmission using channel coding

Optimal decision rule:

Decide in favor of the sequence of signal numbers $\langle \hat{m}[k] \rangle$ which, out of all possible sequences allowed by the code, has the maximum probability **after** observation of the *entire receiver input signal*:
Maximum a–Posteriori Sequence Estimation (MAPSE) (a–posteriori = after observation of receiver input signal (latin))

Notation: $\langle m[k] \rangle$ denotes entire sequence

$$\langle \hat{m}[k] \rangle \;=\; \underset{\forall \mathcal{M}^{-1}(\langle m[k] \rangle) \in \mathcal{C}}{\operatorname{argmax}} \; \Pr\left\{ \langle m[k] \rangle | e(t) \right\}$$

with (Bayes' Rule)

$$\Pr\left\{ \langle m[k] \rangle | e(t) \right\} \;=\; \frac{\Pr\left\{ e(t) | \langle m[k] \rangle \right\} \cdot \Pr\left\{ \langle m[k] \rangle \right\}}{\sum\limits_{\forall \mathcal{M}^{-1}(\langle i[k] \rangle) \in \mathcal{C}} \Pr\left\{ e(t) | \langle i[k] \rangle \right\} \cdot \Pr\left\{ \langle i[k] \rangle \right\}}$$

Note: Denominator is constant for all hypothesis $i[k]$ and must therefore not be evaluated.

For equally probably code words: $\Pr\{\langle m[k] \rangle\} = \text{const.}$

Maximum Likelihood Sequence Estimation (MLSE):

$$\langle \hat{m}[k] \rangle = \underset{\forall \mathcal{M}^{-1}(\langle m[k] \rangle) \in \mathcal{C}}{\operatorname{argmax}} \; \Pr\left\{ e(t) | \langle m[k] \rangle \right\}$$

Decide in favor of the sequence of signal numbers $\langle \hat{m}[k] \rangle$ allowed by the code which maximize the probability of the *already observed* receiver input signal.

Representation of the relevant receiver input signal $\breve{e}(t)$ with vector components $d_l[k]$:

$$\Pr\left\{ \breve{e}(t) | \langle m[k] \rangle \right\} \longrightarrow f_{\boldsymbol{d}}\left(\langle \boldsymbol{d}[k] \rangle | \langle m[k] \rangle \right)$$

$d_l[k]$: independent, complex Gaussian random variables with mean $s_{l,m[k]}$ and variance $\sigma^2 = N_0/E_{\mathrm{g}}$ (ECB–domain)

$$f_{\boldsymbol{d}}\left(\langle \boldsymbol{d}[k] \rangle | \langle m[k] \rangle \right) = \prod_{k=-\infty}^{+\infty} \prod_{l=0}^{D-1} \left(\frac{1}{\pi \sigma^2} \, \mathrm{e}^{-|d_l[k] - s_{l,m[k]}|^2/\sigma^2} \right)$$

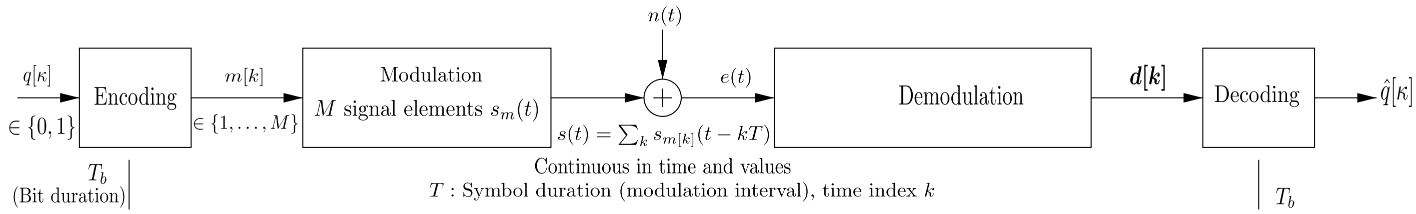
$\ln(x)$ for $x > 0$ a strongly monotonically increasing function:

$$\langle \hat{m}[k] \rangle = \underset{\forall \mathcal{M}^{-1}(\langle m[k] \rangle) \in \mathcal{C}}{\operatorname{argmax}} \sum_{k=-\infty}^{+\infty} \lambda_{m[k]} \quad \text{with} \quad \lambda_{m[k]} = -\sum_{l=0}^{D-1} |d_l[k] - s_{l,m[k]}|^2$$

Metric with respect to the signal element $s_m(t-kT)$ in the k-th step = negative Euclidean distance in the signal space

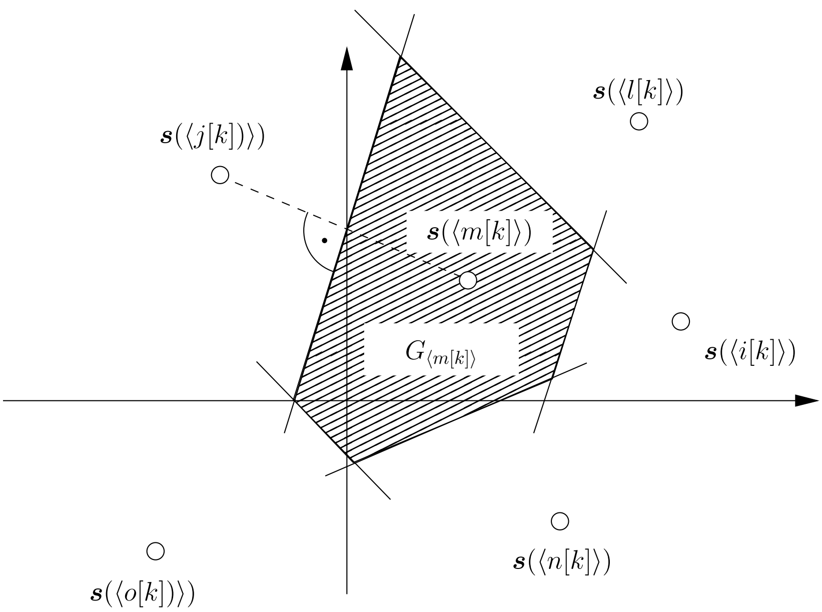
MLSE: Decide in favor of the sequence of signal numbers $\langle \hat{m}[k] \rangle$ allowed by the code which exhibits the shortest Euclidean distance $\sqrt{\sum_{k=-\infty}^{+\infty} -\lambda_{m[k]}}$, between the received signal point $(\ldots \boldsymbol{d}[k], \boldsymbol{d}[k+1], \ldots$ and hypothetical signal point $(\ldots, \boldsymbol{s}_{\hat{m}[k]}, \boldsymbol{s}_{\hat{m+1}[k]}, \ldots)$ in the multi-dimensional signal space.

$$\text{MLSE:} \quad \hat{m}[k] = \underset{\mathcal{M}^{-1}(\langle m[k] \rangle) \in \mathbf{C}}{\operatorname{argmin}} \sum_{k=-\infty}^{+\infty} \sum_{l=0}^{D-1} \left| d_l[k] - s_{l,m[k]} \right|^2$$



Maximum Likelihood Sequence Estimation:

Division of the multi-dimensional signal space into *decision regions* $G_{\langle m[k] \rangle}$ with perpendicularly bisecting separating hyperplanes.



5.4.2 Maximum-Likelihood-Detection

Transmission system without coding: All sequences $\langle m[k] \rangle$ of signal elements are allowed and equally probable; independent decision from modulation step to modulation step possible.

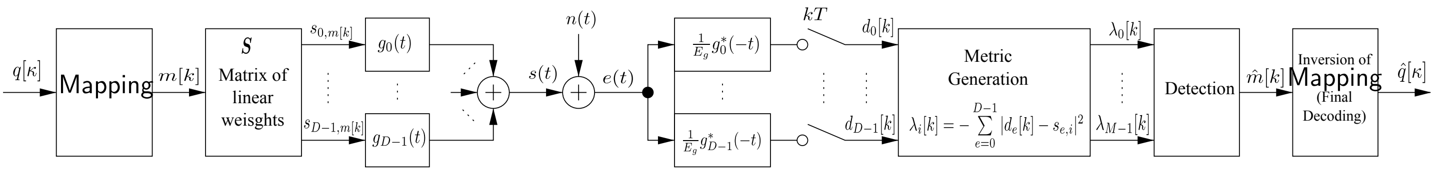
$\Rightarrow$ Maximum Likelihood Detection (MLD):

Decide at step k in favor of the signal element with possesses the smallest Euclidean distance

$$\sqrt{\sum_{d=0}^{D-1} |d_l[k] - s_{l,m[k]}|^2}$$

from the received signal $\boldsymbol{d}[k]$.

Uncoded transmission – ML *detection*:



5.4.3 Minimal Squared Euclidean Distance

Primary indicator for the power efficiency of a digital communications scheme:
Smallest (squared) distance between two signal points in the multi–dimensional signal space

Minimal (squared) Euclidean Distance in the signal space (MSED):

$$d_{E,\min}^2 \stackrel{\text{def}}{=} \min_{\substack{\forall \langle m[k] \rangle, \langle i[k] \rangle \in \mathcal{C} \\ m[k] \neq i[k]}} \sum_{k=-\infty}^{\infty} \sum_{l=0}^{D-1} \left| s_{l,m[k]} - s_{l,i[k]} \right|^2$$

Theorem:

The squared Euclidean distance in the signal space between two signals $s_m(t)$ and $s_i(t)$ corresponds to the energy of their difference $(s_m(t) - s_i(t))$.

Proof:

Exploiting $\left|\sum_l a_l - b_l\right|^2 = \left(\sum_l a_l - b_l\right)\left(\sum_\lambda a_\lambda^* - b_\lambda^*\right) = \sum_l \sum_\lambda a_l a_\lambda^* - a_l b_\lambda^* - b_l a_\lambda^* + b_l b_\lambda^*$

We obtain

$$\begin{aligned}
\int_{-\infty}^{\infty} |s_m(t) - s_i(t)|^2 \,\mathrm{d}t &= \int_{-\infty}^{\infty} \left| \sum_{k=-\infty}^{\infty} \sum_{l=0}^{D-1} \left(s_{l,m[k]} - s_{l,i[k]}\right) g_l(t-kT) \right|^2 \,\mathrm{d}t \\
&= \sum_{k=-\infty}^{\infty} \sum_{\kappa=-\infty}^{\infty} \sum_{l=0}^{D-1} \sum_{\lambda=0}^{D-1} \quad s_{l,m[k]} s^*_{\lambda,m[\kappa]} \varphi_{l,\lambda}[k,\kappa] - s_{l,m[k]} s^*_{\lambda,i[\kappa]} \varphi_{l,\lambda}[k,\kappa] \\
&\qquad\qquad - s_{l,i[\kappa]} s^*_{\lambda,m[k]} \varphi_{l,\lambda}[k,\kappa] + s_{l,i[\kappa]} s^*_{\lambda,i[k]} \varphi_{l,\lambda}[k,\kappa] \\
&= E_\mathrm{g} \sum_{k=-\infty}^{\infty} \sum_{l=0}^{D-1} |s_{l,m[k]}|^2 + |s_{l,i[k]}|^2 \underbrace{- s_{l,m[k]} s^*_{l,i,[k]} - s^*_{l,m[k]} s_{l,i[k]}}_{-2\mathrm{Re}\left\{s_{l,m[k]} s^*_{l,i[k]}\right\}} \\
&= \underbrace{E_\mathrm{g} \sum_{k=-\infty}^{\infty} \sum_{l=0}^{D-1} \left| s_{l,m[k]} - s_{l,i[k]} \right|^2}_{d_E^2} \qquad \text{q.e.d.}
\end{aligned}$$

where we used $\varphi_{l,\lambda}[k,\kappa] = \int_{-\infty}^{\infty} g_l(t-kT) g_\lambda(t-kT)\,\mathrm{d}t = \begin{cases} E_\mathrm{g} & \text{for } l=\lambda \text{ and } k=\kappa \\ 0 & \text{for } l\neq\lambda \text{ or } k\neq\kappa \end{cases}$ double orthogonality of the basis functions

and $|a-b|^2 = |a|^2 + |b|^2 - 2\mathrm{Re}\{ab^*\}$

$\Rightarrow$ The first order decisive measure for the differentiability between two signals is the energy of their difference, i.e. its squared Euclidean distance in the signal space

Analog: **MLSE and MLD**

Decide in favor of the sequence of signal elements $\langle \hat{m}[k] \rangle$ (in favor of the signal element $\hat{m}[k]$), for which the difference between the relevant portion of the receiver input signal $\breve{e}(t)$ and the corresponding hypothetical signal exhibits the smallest energy.

MLSE:

$$\langle \hat{m}[k] \rangle = \underset{\forall \langle \hat{m}[k] \rangle \in \mathcal{C}}{\operatorname{argmin}} \int_{-\infty}^{+\infty} \left| \breve{e}(t) - \sum_{k=-\infty}^{+\infty} s_{\hat{m}[k]}(t-kT) \right|^2 \mathrm{d}t$$

MLD:

$$\hat{m}[k] = \underset{\forall \hat{m}[k] \in \mathcal{C}}{\operatorname{argmin}} \int_{-\infty}^{+\infty} \left| \breve{e}(t) - s_{\hat{m}[k]}(t-kT) \right|^2 \mathrm{d}t$$

Definition: *Normalized, Minimum, Squared Euclidean Distance (NMSE)*

$$d^2_{\min} = \frac{d^2_{E,\min} \cdot E_{\mathrm{g}}}{\underset{\uparrow}{2}\; E_{\mathrm{b}}}$$

for historical reasons!

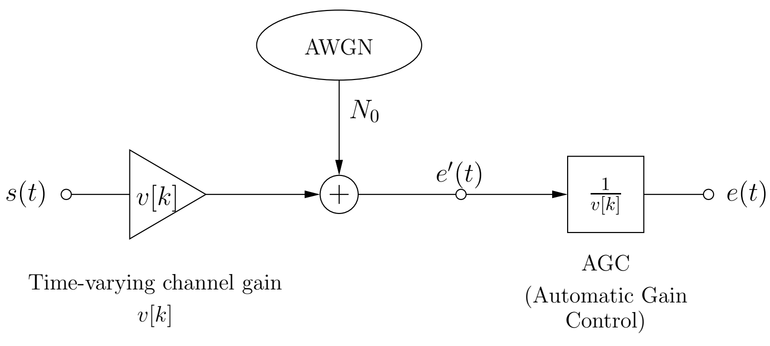
Application of channel coding:

For differing (at least by one symbol) sequences of source symbols, sequences of signal elements are generated which differ in *several (as many as possible) steps*. In this way, the minimal Euclidean distance $d_{E,\min}$ between two possible desired signal points in the multi–dimensional signal space is increased!

$\Rightarrow$ Increase in power efficiency!

MLSE for time–varying channels:

The model also includes *time–varying*, memoryless AWGN channels, e.g. a channel with non–frequency selective fading



I.e. N_0 has to be replaced by $N_0/|v[k]|^2$:

$$\langle \hat{m}[k] \rangle = \underset{\forall \langle m[k] \rangle \in \mathcal{C}}{\operatorname{argmin}} \sum_{k=-\infty}^{+\infty} |v[k]|^2 \cdot \sum_{l=0}^{D-1} |d_l[k] - s_{l,m[k]}|^2$$

5.4.4 Demodulation by Correlation

Maximum Likelihood Estimation: (AWGN assumed)

$$\text{MLSE:} \quad \hat{s}(t) = \underset{\forall s(t)}{\operatorname{argmin}} \int_{-\infty}^{+\infty} |e(t) - s(t)|^2 \, \mathrm{d}t$$

Following holds:

$$\begin{aligned}
\int_{-\infty}^{+\infty} |e(t) - s(t)|^2 \, \mathrm{d}t &= \int_{-\infty}^{+\infty} (e(t) - s(t))\,(e(t) - s(t))^* \, \mathrm{d}t \\
&= \int_{-\infty}^{+\infty} |e(t)|^2 \, \mathrm{d}t - \int_{-\infty}^{+\infty} (e(t)s^*(t) + e^*(t)s(t)) \, \mathrm{d}t + \int_{-\infty}^{+\infty} |s(t)|^2 \, \mathrm{d}t \\
&= E_e - 2\,\mathrm{Re}\left\{ \int_{-\infty}^{+\infty} e(t)s^*(t) \, \mathrm{d}t \right\} + E_s
\end{aligned}$$

with
- Energy of the receiver input signal: $E_e = \int_{-\infty}^{+\infty} |e(t)|^2 \, \mathrm{d}t$
- *Correlation* of the receiver input signal with the hypothetical transmit signal

$$\int_{-\infty}^{+\infty} e(t)s^*(t) \, \mathrm{d}t = \int_{-\infty}^{+\infty} e(t) \sum_{k=-\infty}^{+\infty} s^*_{m[k]}(t - kT) \, \mathrm{d}t \overset{\text{def}}{=} E_g \cdot \sum_{k=-\infty}^{+\infty} K_{m[k]}[k]$$

with the components

$$K_{m[k]}[k] \overset{\text{def}}{=} \frac{1}{E_g} \int_{-\infty}^{+\infty} e(t)s^*_m(t - kT) \, \mathrm{d}t = \left(\frac{1}{E_g} e(t) * s^*_m(T_v - t) \right)\Bigg|_{t = kT + T_v}$$

a *correlation vector* $\boldsymbol{K}[k] = (K_0[k], K_1[k], \ldots, K_{M-1}[k])$ at step k, which represent the sample values at modulation step k at the output of the Matched Filter adapted to the signal elements $s_m(t)$.

- Energy of the hypothetical desired signal:

$$\begin{aligned}
E_s &= \int_{-\infty}^{+\infty} |s(t)|^2 \, \mathrm{d}t = \int_{-\infty}^{+\infty} \left| \sum_{k=-\infty}^{+\infty} s_{m[k]}(t - kT) \right|^2 \mathrm{d}t \overset{\text{temporal orthogon.}}{=} \sum_{k=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left| s_{m[k]}(t) \right|^2 \, \mathrm{d}t \\
&\overset{\text{def}}{=} \sum_{k=-\infty}^{+\infty} E_{m[k]}
\end{aligned}$$

Therefore:

$$\begin{aligned}
\int_{-\infty}^{+\infty} |e(t) - s(t)|^2 \, \mathrm{d}t &= E_e - 2\,\mathrm{Re}\left\{ \int_{-\infty}^{+\infty} e(t)s^*(t) \, \mathrm{d}t \right\} + E_s \\
&= E_e - 2\,\mathrm{Re}\left\{ E_g \cdot \sum_{k=-\infty}^{+\infty} K_{m[k]}[k] \right\} + \sum_{k=-\infty}^{+\infty} E_{m[k]} \\
&= E_e - E_g \sum_{k=-\infty}^{+\infty} \left(2\,\mathrm{Re}\left\{ K_{m[k]}[k] \right\} - E_{m[k]}/E_g \right)
\end{aligned}$$

As E_e is irrelevant (independent of the hypothesis of a transmit signal) with respect to the calculation of the minimum squared Euclidean distance (i.e. the difference energy in the case of ML detection), and the energies E_m of all signal elements at the receiver are known, the *correlation* of the receiver input signal with all hypothetical transmit signal elements allows an ML detection, i.e. optimum decoding is possible (generalization of the matched filter receiver).

$\Rightarrow$ The sequence of the

correlation vectors $\quad \boldsymbol{K}[k] = \Big[K_0[k],\, K_1[k]\, , \ldots, K_{M-1}[k] \Big]\, , \quad k \in \mathbb{Z}$

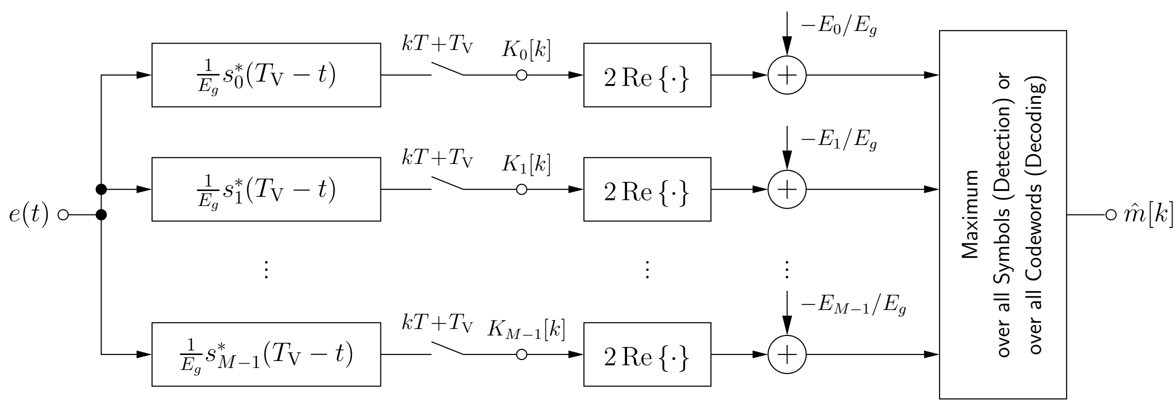
constitutes a **sufficient statistics** on the receiver input signal with respect to the transmitted sequence of signal numbers $m[k]$.

The value

$$\lambda'_m[k] \stackrel{\text{def}}{=} 2\,\mathrm{Re}\{K_m[k]\} - E_m/E_g$$

is therefore, along with the squared Euclidean distance $\lambda_m[k] = |\boldsymbol{d}[k] - \boldsymbol{s}_m|^2$, an optimal metric with respect to the hypothesis that in the k-th step the signal element $s_m(t)$ had been transmitted.
(NB: The correlation needs to be maximized instead of minimization of Euclidean distance!)

General Correlation Receiver (Matched Filter Receiver):



As $M \geq D$ holds, the correlation demodulator is in general more computationally expensive than the vector demodulator. For $M = D$, when all signal elements are linearly independent (e.g. as in the case of orthogonal modulation schemes like pulse position modulation), these demodulation procedures are equivalent.

5.5 Non–coherent Demodulation

Receiver input signal for an AWGN channel (ECB domain). Carrier phase φ is usually unkown at the receiver:

$$e(t) = s(t) \cdot \mathrm{e}^{\mathrm{j}\varphi} + n(t)$$

with complex white Gaussian noise $n(t)$ having noise power spectral density

$$N_0 \quad \text{or} \quad N_0/2 \text{ per quadrature component}$$

Strategies:

- *Coherent demodulation*:
  φ is estimated through carrier phase synchronization (estimate: $\hat{\varphi}$)
  For analysis, an ideal carrier phase synchronization is assumed, i.e.

  $$\hat{\varphi} = \varphi$$

  Without loss of generality, we set

  $$\varphi = 0 \;, \qquad \hat{\varphi} = 0$$

  An estimation error $\hat{\varphi} \neq \varphi$ usually leads to a degradation in power efficiency (BER increases!)

- *Non–coherent demodulation*:
  The carrier phase is not estimated, i.e., all phases are considered equally likely by the demodulator.
  Probability density function (pdf) $f_\varphi(\varphi)$ of the carrier phase:

  $$f_\varphi(\varphi) = \begin{cases} \frac{1}{2\pi} & \text{for } -\pi \leq \varphi < +\pi \\ 0 & \text{otherwise} \end{cases}$$

  However, in the case of non–coherent demodulation schemes it is assumed that the carrier phase φ remains constant over one or even several modulation intervals T. This implies that we assume for the difference between the carrier *frequencies* at the transmitter and receiver

  $$\Delta f_c \cdot T \ll 1$$

  Thus, a carrier *frequency* synchronization may still be necessary for non–coherent demodulation. (Especially for mobile communication applications – Doppler Effect!)

Optimal Non–coherent Demodulation:

- Without loss of generality, we limit the analysis to a single modulation step at $k = 0$. Due to the temporal orthogonality, these results are also valid for sequences of signal elements.
- The information contained in the receiver input signal $e(t)$ with respect to the signal element $s_{m[0]}(t)$ at step $k = 0$ is completely represented by vector

$$\boldsymbol{d}[0] = \left[d_0[0], d_1[0], \ldots,\; d_{D-1}[0]\right]$$

at the output of the *vector demodulator*

Combined pdf of vector $\boldsymbol{d}[0]$ and carrier phase (difference) φ, assuming the signal element $s_m(t)$ was transmitted:

$$f_{\boldsymbol{d}[0],\varphi}(\boldsymbol{d}[0], \varphi \mid m) = \left(\frac{E_g}{\pi N_0}\right)^D \mathrm{e}^{-\frac{\sum\limits_{l=0}^{D-1}\left|d_l[0]-s_{l,m}\cdot \mathrm{e}^{\mathrm{j}\varphi}\right|^2}{N_0/E_g}} \cdot \frac{1}{2\pi}\mathrm{rect}\left(\frac{\varphi}{2\pi}\right)$$

252

- Interpretation of the squared Euclidean distance as the energy of the difference between the relevant portion of the receiver input signal $\breve{e}(t)$ and the assumed signal element:

$$\begin{aligned} E_g \sum_{l=0}^{D-1} \left|d_l[0] - s_{l,m} \cdot \mathrm{e}^{\mathrm{j}\varphi}\right|^2 &= \int\limits_{-\infty}^{+\infty} \left|\breve{e}(t) - \mathrm{e}^{\mathrm{j}\varphi} s_m(t)\right|^2 \,\mathrm{d}t \\ &= E_{\breve{e}} + E_m - 2E_g \,\mathrm{Re}\left\{K_m[0]\mathrm{e}^{-\mathrm{j}\varphi}\right\} \end{aligned}$$

with $K_m[0]$ denoting the complex correlation, see Section 5.4.4:

$$K_m[k] = \frac{1}{E_g} \int\limits_{-\infty}^{+\infty} e(t)\, s_m^*(t - kT)\,\mathrm{d}t$$

The combined pdf can therefore be defined as

$$f_{\boldsymbol{d}[0],\varphi}(\boldsymbol{d}[0],\varphi \mid m) = \frac{1}{2\pi} \cdot C_{\mathrm{irr}} \cdot \mathrm{e}^{-E_m/N_0} \cdot \mathrm{e}^{+2\,\mathrm{Re}\{K_m[0]\cdot \mathrm{e}^{-j\varphi}\}\cdot E_g/N_0} \cdot \mathrm{rect}\left(\frac{\varphi}{2\pi}\right)$$

with an *irrelevant factor* $C_{\mathrm{irr}} = \left(\frac{E_g}{\pi N_0}\right)^D \cdot \mathrm{e}^{-E_{\breve{e}}/N_0}$

- ML detection of the signal number m:

$$\hat{m} = \underset{m}{\operatorname{argmax}} \; f_{\boldsymbol{d}[0]}(\boldsymbol{d}[0] \mid m)$$

- The pdf

$$f_{\boldsymbol{d}[0]}(\boldsymbol{d}[0] \mid m)$$

is obtained by means of marginalization of the joint pdf with respect to φ:

$$f_{\boldsymbol{d}[0]}(\boldsymbol{d}[0] \mid m) = \int\limits_{-\pi}^{+\pi} f_{\boldsymbol{d}[0],\varphi}(\boldsymbol{d}[0],\varphi \mid m) \, \mathrm{d}\varphi$$

- ML metric

$$\lambda'_m \sim \ln(f_{\boldsymbol{d}}(\boldsymbol{d} \mid m))$$

for non–coherent demodulation (after discarding all irrelevant terms):

$$\lambda'_m[0] = \ln\left(\frac{1}{2\pi}\int\limits_{-\pi}^{+\pi} \mathrm{e}^{2\,\mathrm{Re}\{K_m[0]\cdot \mathrm{e}^{-j\varphi}\}\cdot E_g/N_0} \, \mathrm{d}\varphi\right) - E_m/N_0$$

With $\qquad K_m[0] = |K_m[0]| \cdot \mathrm{e}^{\mathrm{j}\varphi_{Km}} \qquad$ and $\quad \varphi_{Km} = \arg\{K_m[0]\}$

we obtain $\qquad \lambda'_m[0] = \ln\left(\frac{1}{2\pi}\int\limits_{-\pi}^{+\pi} \mathrm{e}^{2\frac{E_g}{N_0}\cdot|K_m[0]|\cdot\cos(\varphi_{Km}-\varphi)} \, \mathrm{d}\varphi\right) - E_m/N_0$

- *Modified Bessel function of order* 0:

$$I_0(x) = \frac{1}{2\pi}\int\limits_{-\pi}^{+\pi} \mathrm{e}^{x\cos(\varphi-\varphi_0)} \, \mathrm{d}\varphi$$

$I_0(x)$ is independent from φ_0, as the integration is carried out over one period of a periodic function. (Note: The following Bessel functions exist $J_n(x)$, $\quad I_n(x)$, $\quad K_n(x)$, $\quad H_n(x)$).

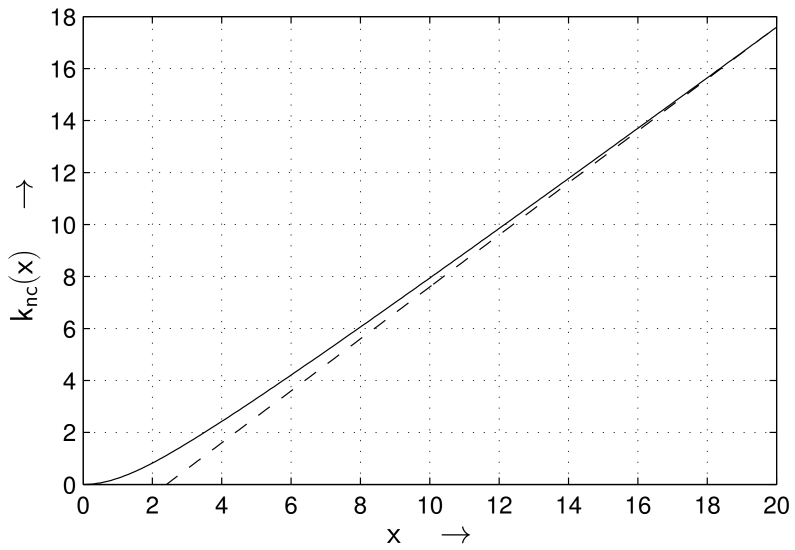
- Optimal metric (or optimal detection variable) for non–coherent demodulation:

$$\lambda'_m[k] = \mathrm{k}_{\mathrm{nc}}\left(2\frac{E_g}{N_0}\,|K_m[k]|\right) - E_m/N_0$$

with $\qquad \mathrm{k}_{\mathrm{nc}}(x) = \ln(I_0(x)) \quad$ for $x > 0$

(*optimal weighting function for n̲on-c̲oherent demodulation*)

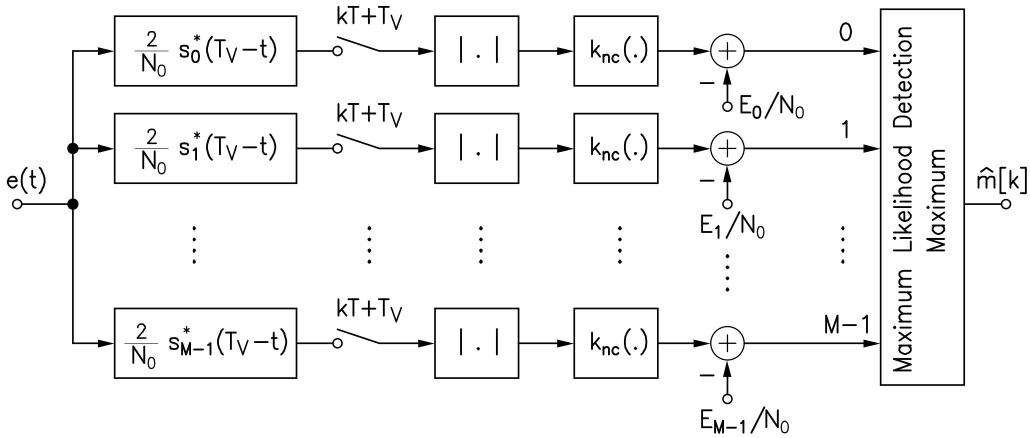
- Graph of function $\mathrm{k_{nc}}(x)$:



Function $\mathrm{k_{nc}}(x)$ is monotonically increasing and almost linear for $x > 6$:

$\Rightarrow$ Approximation: $\quad \mathrm{k_{nc}}(x) \approx x - 2.4 \quad$ for $\quad x > 6$

Matched Filter Demodulation (Non–coherent):



Notes:

- If all signal elements have the same energy (e.g., for digital frequency modulation), i.e., for

$$E_m = \text{const.} \quad \forall\, m = 0(1)M-1,$$

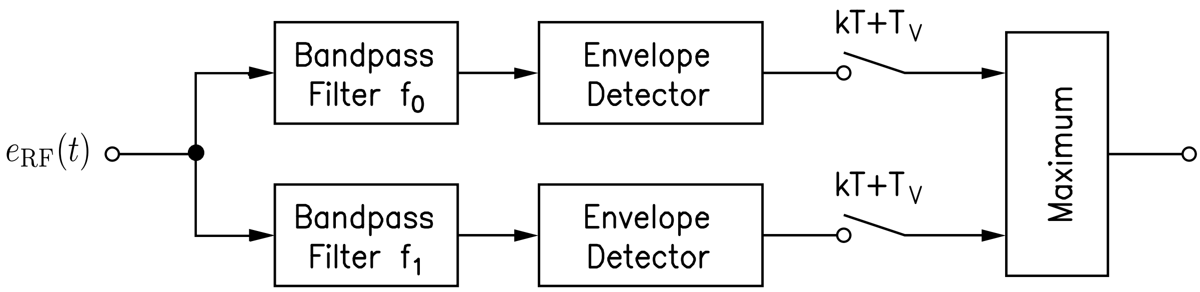
  the subtraction of E_m/N_0 can be discarded (irrelevant with respect to maximization).

- For ML *detection*, i.e., for an uncoded transmission, and $E_m = \text{const.}$ the nonlinear weighting can also be disregarded, as $\mathrm{k}_{\mathrm{nc}}(x)$ is monotonic (a maximum remains a maximum). $\Longrightarrow$ Envelope demodulation:

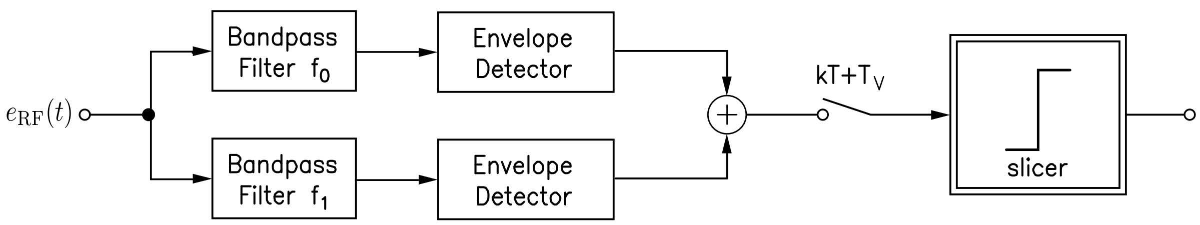
$$\hat{m}[k] = \underset{m}{\operatorname{argmax}}\, |K_m[k]|$$

  M parallel filters for the signal elements; decision for the signal element whose corresponding filter has the *greatest envelope value* at sampling time instant kT.

Example: Binary frequency modulation

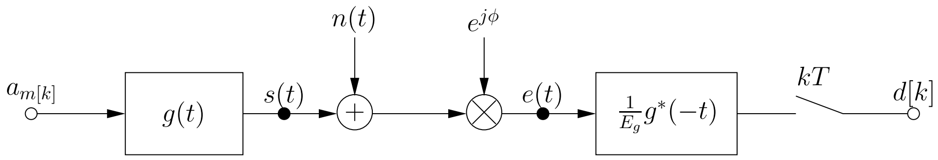


or



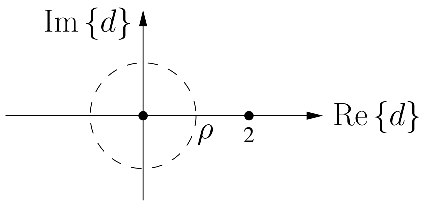
5.5.1 Non–coherent Demodulation of Unipolar ASK

Block diagram:



Non–coherent demodulation is possible for *unipolar* ASK, but not for bipolar ASK, since in bipolar ASK the phase is used for representation of information.

For simplicity, only binary transmission, $M = 2$, is considered at first: on-off keying (OOK)



Decision boundary: Circle centered at origin with radius ρ (for $M > 2$-step ASK: Decision boundaries are concentric circles)

260

Error probability:

A. $s_0 = 0$ transmitted

A bit error arises if the zero mean complex Gaussian variable $d = n$ with variance N_0/E_g has a magnitude greater than ρ. The pdf of d is given by

$$f_e(d_I, d_Q \mid 0) = \frac{1}{2\pi\sigma_I^2}\,\mathrm{e}^{-(d_I^2+d_Q^2)/(2\sigma_I^2)}$$

with

$$\begin{aligned} d &= d_I + jd_Q \quad ; \\ \sigma_I^2 &= N_0/(2E_g) \end{aligned}$$

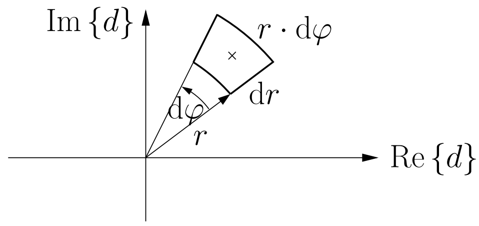
Cumulative distribution function (cdf) for the magnitude ρ of a complex, zero–mean Gaussian random variable:

$$F_\rho(\rho) = \Pr\{|d| \le \rho\} = \iint\limits_{d_I^2+d_Q^2\le\rho^2} f_e(d_I, d_Q \mid 0)\;\mathrm{d}d_I\mathrm{d}d_Q$$

Transformation from Cartesian to polar coordinates

$$\mathrm{d}_I = r \cdot \cos\varphi \quad ; \quad d_Q = r \cdot \sin\varphi$$

with $$r = |d| = \sqrt{d_I^2 + d_Q^2}\,, \quad \varphi = \arg\{d\}\,, \quad \mathrm{d}d_I \mathrm{d}d_Q = \mathrm{d}r \cdot r \cdot \mathrm{d}\varphi$$



$$\Longrightarrow \qquad F_\rho(\rho) = \frac{1}{2\pi\sigma_I^2} \int_0^\rho \int_0^{2\pi} \mathrm{e}^{-r^2/(2\sigma_I^2)} \cdot r \; \mathrm{d}\varphi \mathrm{d}r$$

With $\int x\,\mathrm{e}^{-x^2}\,\mathrm{d}x = -\frac{1}{2}\,\mathrm{e}^{-x^2} + C$ we obtain $F_\rho(\rho) = 1 - \mathrm{e}^{-\rho^2/(2\sigma_I^2)}$

Therefore, we get for the error probability if $s_0 = 0$ is transmitted $\mathrm{P_e}(s_0 = 0) = 1 - F_\rho(\rho) = \mathrm{e}^{-\rho^2/(2\sigma_I^2)}$

Rayleigh Distribution:

pdf $f_\rho(\rho)$ for the magnitude of a complex, zero–mean Gaussian random variable with variance $\sigma^2 = 2\sigma_I^2$

$$f_\rho(\rho) = \frac{\mathrm{d}}{\mathrm{d}\rho} F_\rho(\rho) = \begin{cases} \dfrac{2\rho}{\sigma^2}\,\mathrm{e}^{-\rho^2/\sigma^2}, & \text{for } \rho \geq 0 \\ 0, & \text{for } \rho < 0 \end{cases}$$

B. $s_1 = 2$ transmitted

A bit error arises when the complex Gaussian random variable $d = 2 + n$ with mean 2 and variance N_0/E_g has a magnitude smaller than ρ. The conditional pdf of d is now given by

$$f_e(d_I, d_Q \mid a) = \frac{1}{2\pi\sigma_I^2}\,\mathrm{e}^{-((d_I - a)^2 + d_Q^2)/(2\sigma_I^2)}$$

with $a = 2$

Cumulative distribution function (cdf) for the magnitude ρ of a complex non–zero–mean Gaussian random variable

$$F_\rho(\rho) = \Pr\{|d| \leq \rho\} = \iint_{d_I^2 + d_Q^2 \leq \rho^2} f_e(d_I, d_Q \mid a)\; \mathrm{d}d_I \mathrm{d}d_Q$$

in polar coordinates:

$$F_\rho(\rho) = \frac{1}{\sigma_I^2}\int_0^\rho r\,\mathrm{e}^{-(a^2+r^2)/(2\sigma_I^2)} \cdot \underbrace{\frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{(r\,a\cos\varphi)/\sigma_I^2}\,\mathrm{d}\varphi}_{=I_0\left(\frac{ar}{\sigma_I^2}\right)}\;\mathrm{d}r$$

where $I_0(x) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{x\cos\varphi}\,\mathrm{d}\varphi$

is a zeroth order modified Bessel function of the first kind.

Therefore we obtain $\mathrm{P_e}(s_1 = 2) = F_\rho(\rho) =: 1 - Q_M\left(\frac{2}{\sigma_I}, \frac{\rho}{\sigma_I}\right)$

with the "**Marcum Q–function**"

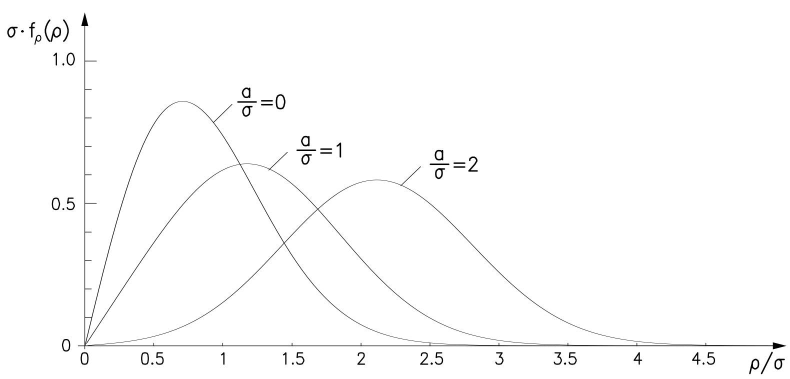
$$Q_M(y,x) := \int_x^\infty r\,\mathrm{e}^{-(r^2+y^2)/2} \cdot I_0(yr)\,\mathrm{d}r,$$

which is the complementary cumulative distribution function $(F_\rho(\rho) = 1 - Q_M(y,x))$ for Rice–distributed random variables with variance 2.

Rice Distribution:

The probability density function $f_\rho(\rho)$ of the magnitude of a complex Gaussian random variable with variance $\sigma^2 = 2\sigma_I^2$ and mean value magnitude a (Note: Phase of the complex mean value is irrelevant!)

$$f_\rho(\rho) = \frac{\mathrm{d}}{\mathrm{d}\rho}F_\rho(\rho) = \begin{cases} \frac{2}{\sigma^2}\cdot\rho\,\mathrm{e}^{-(\rho^2+a^2)/\sigma^2}\cdot I_0\left(\frac{2a\rho}{\sigma^2}\right) & \text{for } \rho \geq 0 \\ 0 & \text{for } \rho < 0 \end{cases}$$



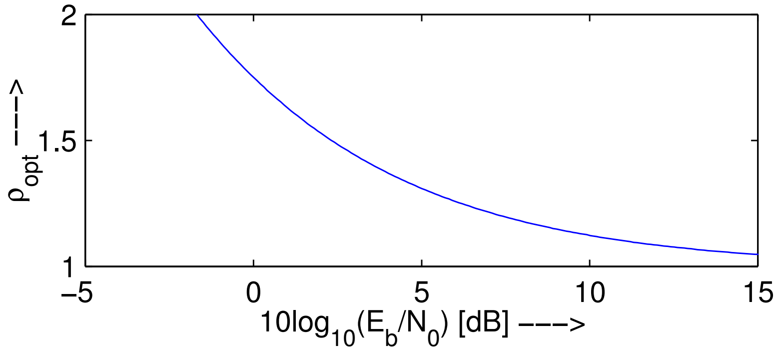
Special case $a = 0$: Rayleigh Distribution

A + B: Average bit error rate for non–coherent demodulation of OOK:

With $E_\mathrm{b} = 2E_g$ the following holds:

$$\mathrm{BER} = \frac{1}{2}\left(\mathrm{e}^{-\rho^2 E_\mathrm{b}/(2N_0)} + 1 - Q_M\left(2\sqrt{\frac{E_\mathrm{b}}{N_0}}, \sqrt{\frac{\rho^2 E_\mathrm{b}}{N_0}}\right)\right)$$

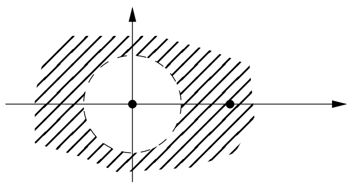
Important: The optimal radius ρ for the decision boundary circle depends on the actual SNR.



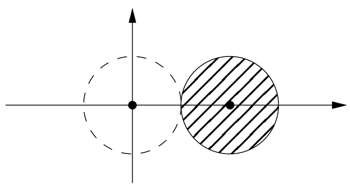
In practice: Radius $\rho = 1$ is usually chosen.

Upper bound for error probabilities:

Integration over the region for the correct reception of $s_1 = 2$



is approximated by integrating over a circle centered on the sent signal point.



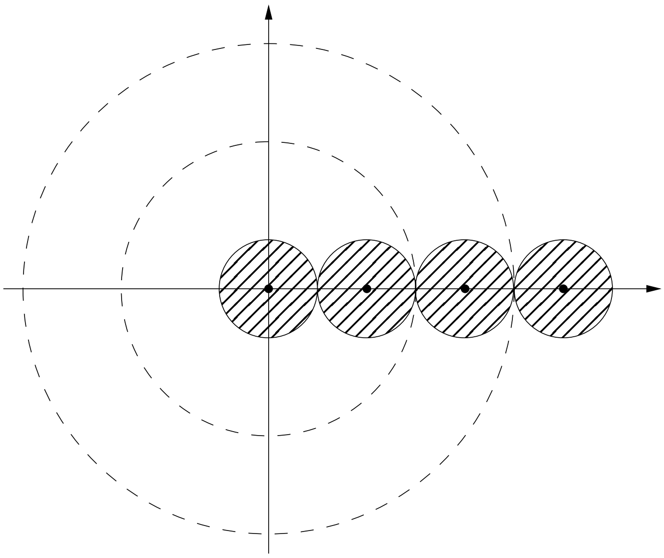
Upper bound is relatively tight for high SNR

$$\begin{aligned} \mathrm{P_e}(s_1 = 2) &< \mathrm{e}^{-\rho^2/(2\sigma_I^2)} \\ &= \mathrm{P_e}(s_0 = 0) \\ \text{choice for high SNR: } \rho &= 1 \text{ (see figure)} \end{aligned}$$

Thus, we obtain $\quad \mathrm{BER} < \mathrm{e}^{-\rho^2/(2\sigma_I^2)} = \mathrm{e}^{-1/(2N_0/(2E_g))}$

or $\quad \mathrm{BER} < \mathrm{e}^{-E_\mathrm{b}/(2N_0)}$

Generalization of this approximation to multi–level, unipolar ASK with non–coherent demodulation:



$$\mathrm{SER} < \mathrm{e}^{-E_g/N_0} = \mathrm{e}^{-d_{\min}^2 \cdot E_\mathrm{b}/(2N_0)}$$

Bit error rate for Gray mapping

$$\mathrm{BER} \approx \frac{1}{\mathrm{ld}(M)} \cdot \mathrm{SER}$$

268

Comparison with coherent demodulation:

Symbol error probability for coherent demodulation:

$$\mathrm{SER} = \frac{2M-2}{M} Q\left(\sqrt{d_{\min}^2 \frac{E_\mathrm{b}}{N_0}}\right)$$

Chernoff bound for Q function:

$$Q(x) \leq \frac{1}{2}\,\mathrm{e}^{-x^2/2} \quad \text{for} \quad x \geq 0$$

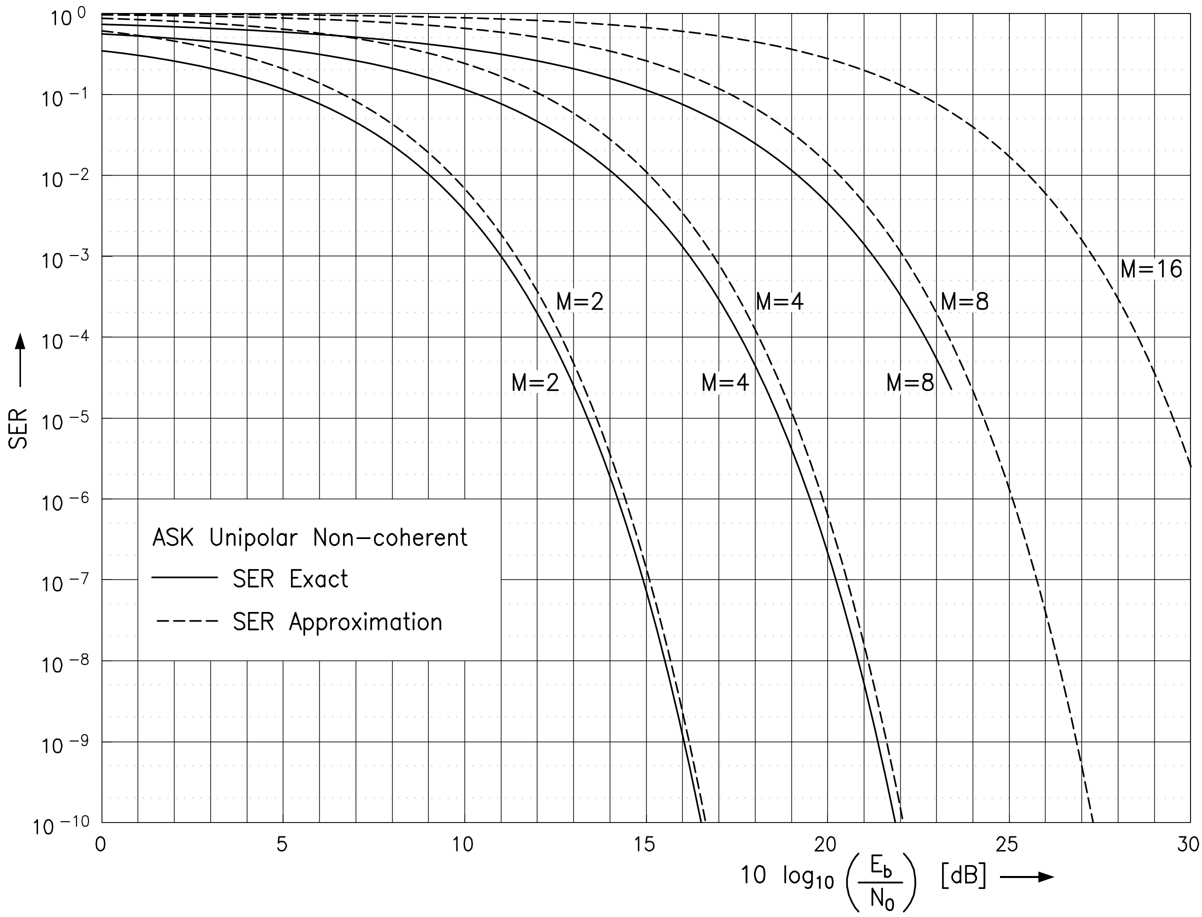
$$\mathrm{SER} \leq \frac{M-1}{M}\,\mathrm{e}^{-d_{\min}^2 E_\mathrm{b}/(2N_0)}$$

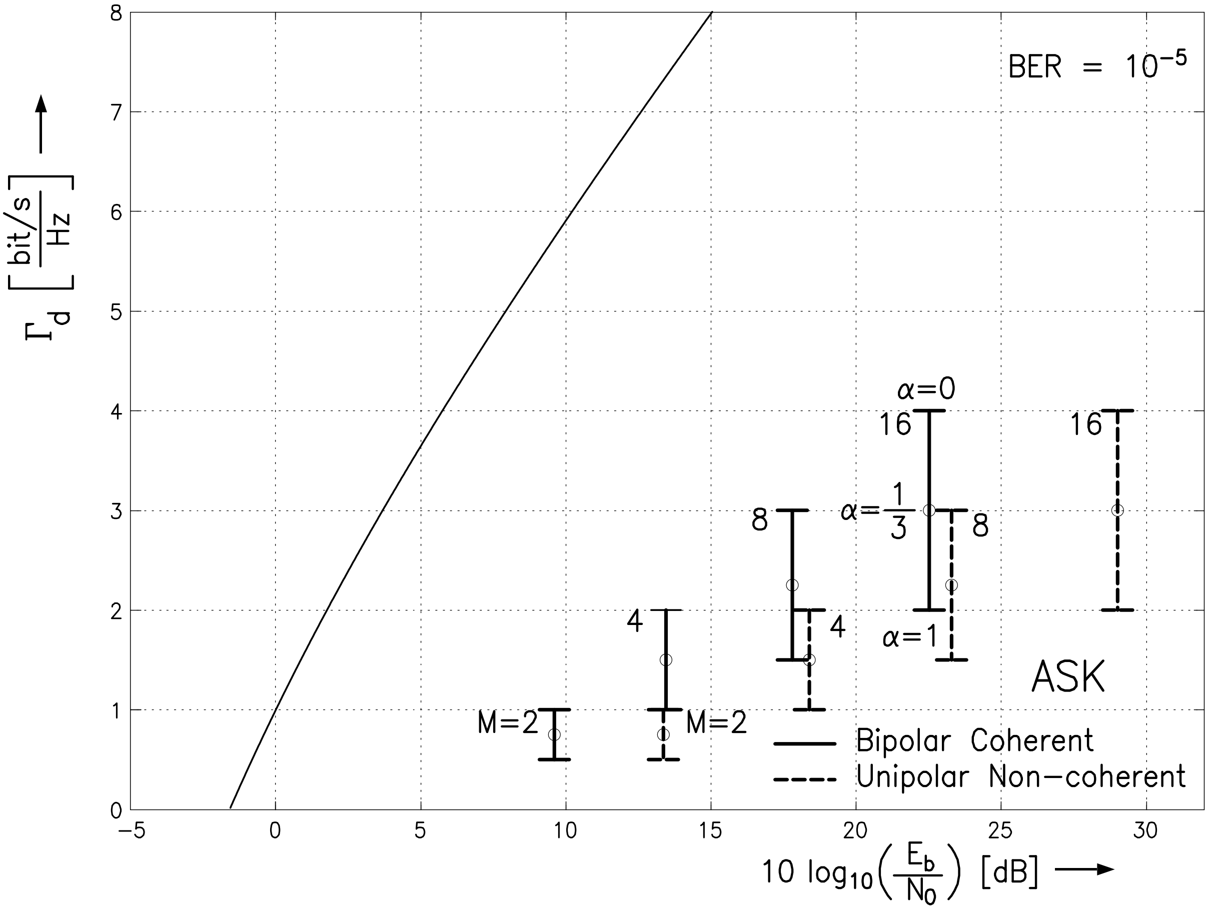
$\Longrightarrow$ The result for non–coherent demodulation (upper bound) is equivalent to the Chernoff bound for coherent demodulation

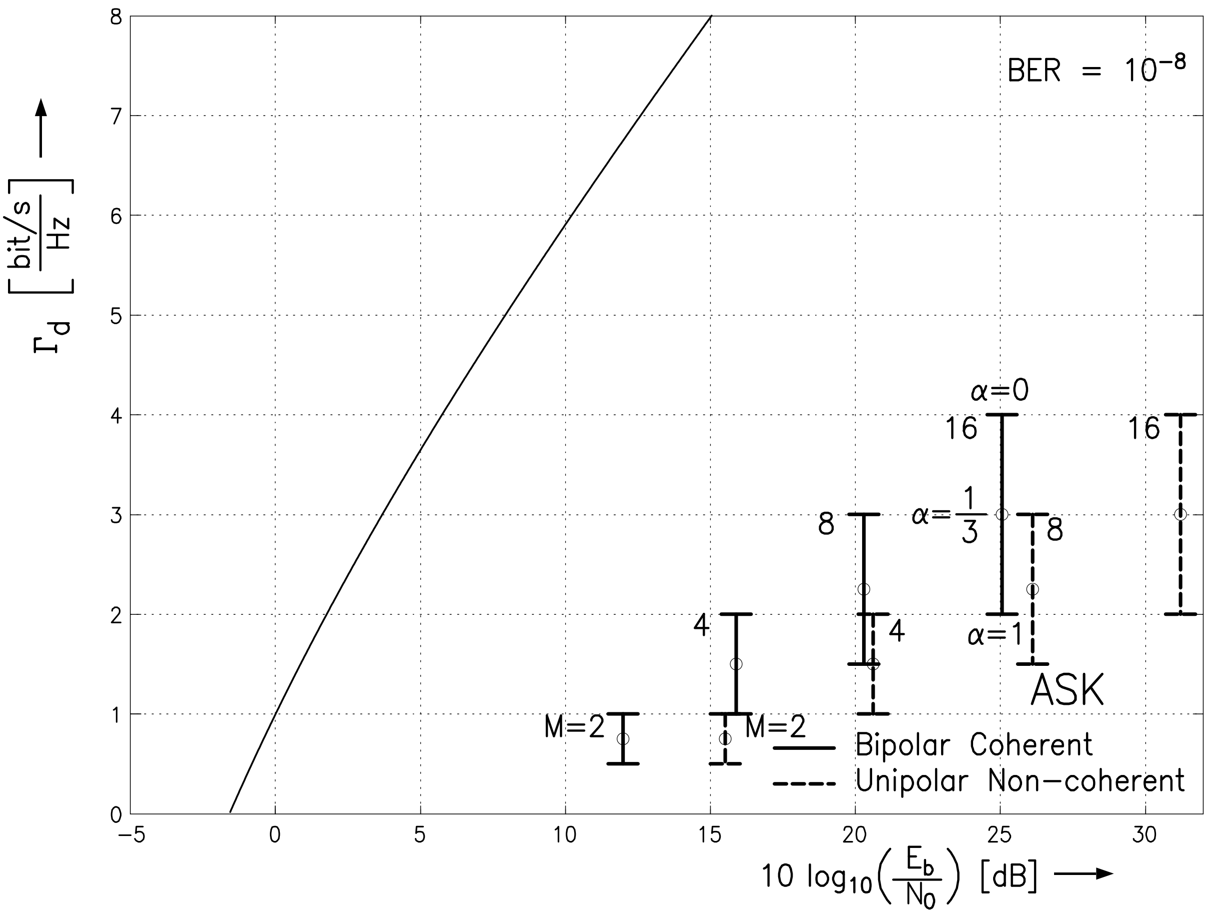
Interpretation:

Non–coherent demodulation itself causes only a small loss in power efficiency compared to coherent demodulation. However, a noticeable performance loss (for $M \gg 2$ up to 6 dB) is caused by the fact that non–coherent demodulation schemes are restricted to unipolar ASK instead of significantly more power efficient bipolar ASK methods.

$$\mathrm{E}\{|a|^2\} = \frac{4M^2 - 6M + 2}{3} = \frac{M^2-1}{3} + (M-1)^2 = \sigma_a^2 + m_a^2$$


ASK Unipolar Non−coherent
SER Exact
SER Approximation
M=2
M=4
M=8
M=16
SER
10 log10 (Eb/N0) [dB]

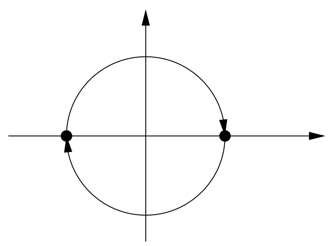

BER = 10^−5
Γd [bit/s / Hz]
α=0
α=1/3
α=1
M=2
4
8
16
ASK
Bipolar Coherent
Unipolar Non−coherent
10 log10 (Eb/N0) [dB]



5.5.2 Differential PSK

Problem at the receiver: Where is $0°$?

Example: 2PSK



=> Incorrect locking of the carrier phase synchronization leads to an inversion of all symbols!

In general: M–fold ambiguity for M–PSK

(also for ASK, bipolar: 2–fold ambiguity

QAM 4–fold ambiguity)

Possible solutions:

- Synchronizing sequence periodically inserted into data stream
- Differential precoding:

  Representation of the message using the *difference* of the phases between two adjacent signal points, but not the absolute phase values.
  (Differential phase modulation $\rightarrow$ a step towards frequency modulation!)

M–ary information symbol (for the k–th modulation step)

$$b[k] \in \{0, 1, 2, \dots, (M-1)\}$$

Differential PSK:

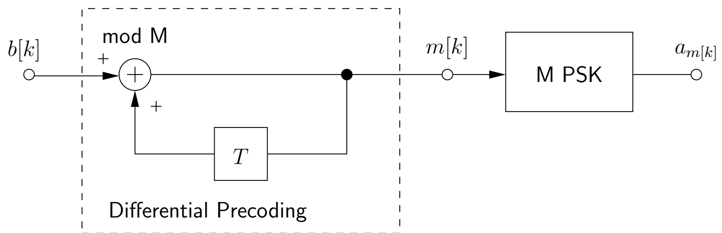
$$a_{m[k]} = a_{m[k-1]} \cdot \mathrm{e}^{\mathrm{j}2\pi b[k]/M}$$

Alternatively, with $m[k] := \dfrac{M}{2\pi} \arg\{a_{m[k]}\} \in \{0, 1, 2, \dots, (M-1)\}$, i.e., $a_{m[k]} = \mathrm{e}^{\mathrm{j}2\pi \frac{m[k]}{M}}$ the following holds

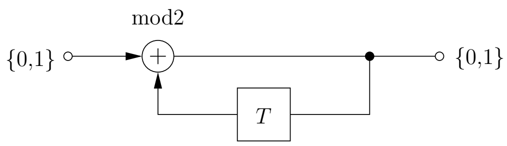
$$m[k] = (m[k-1] + b[k]) \bmod M$$

=> *Cyclic integration* $\{0, 1, \dots, (M-1)\}$ with respect to modulo M arithmetic

Block diagram of the transmitter:



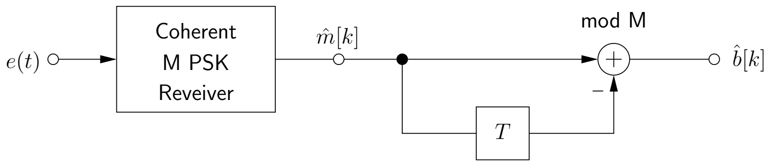
Special case: $M = 2$



5.5.2.1 Differential Coherent Demodulation: DCPSK

- Coherent PSK-receiver detects symbols $\hat{m}[k]$

=> Inversion of integration using *cyclic differentiation* $\{0, 1, \ldots, (M-1)\}$ according to the mod M arithmetic



At the receiver side, differentiation suppresses the constant components (DC components)
=> A constant phase offset created by false locking of the carrier phase synchronization is suppressed!

Error Probability for DCPSK:

$$
\begin{aligned}
M = 2: \qquad & \mathrm{BER}_{\mathrm{DCPSK}} &=& \ \mathrm{BER}_{\mathrm{PSK}} \cdot (1 - \mathrm{BER}_{\mathrm{PSK}}) + (1 - \mathrm{BER}_{\mathrm{PSK}}) \cdot \mathrm{BER}_{\mathrm{PSK}} \\
& &=& \ 2\,\mathrm{BER}_{\mathrm{PSK}} \cdot (1 - \mathrm{BER}_{\mathrm{PSK}}) \\
M > 2: \qquad & \mathrm{SER}_{\mathrm{DCPSK}} &\approx& \ 2\,\mathrm{SER}_{\mathrm{PSK}} - 1{,}5\,\mathrm{SER}_{\mathrm{PSK}}^2
\end{aligned}
$$

High probability of double errors!

276

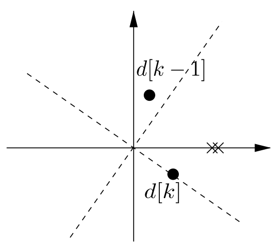
5.5.2.2 Differential Demodulation: DPSK

- Phase of the preceding signal point is used as phase reference.
- Assumption: Carrier phase φ remains approximately constant between two subsequent symbol intervals.

Interpretation as coherent scheme:

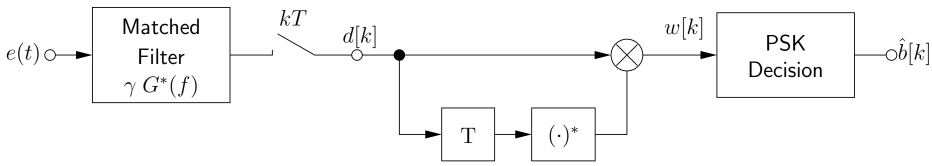
Simple carrier phase synchronization by using one signal point for carrier phase estimation.
The signal used for carrier phase estimation is affected by noise in the same way as the information carrying signal point!

Example: $M = 2$; two $+1$'s are sent



differential demodulation: error
coherent demodulation: no error

Practical implementation: (only valid for ECB signals not for RF signals!)



Decision variable $w[k] \stackrel{\text{def}}{=} d[k] \cdot d^*[k-1] = |d[k]| \cdot |d[k-1]| \cdot \mathrm{e}^{\mathrm{j}\,(\arg\{d[k]\} - \arg\{d[k-1]\})}$

Error probability:

Here, only binary 2DPSK ($M = 2$) is considered for simplicity; for a general derivation, see e.g. Benedetto, Biglieri, Castellani: *Digital Transmission Theory*

Without loss of generality, assume $b[k] = 0$ is transmitted and normalization $\gamma = \frac{1}{E_g}$:

$$\mathrm{BER} = \Pr\left\{\pi/2 < \arg\{w[k]\} < 3\pi/2\right\} = \Pr\left\{\mathrm{Re}\{w[k]\} < 0\right\}$$

with

$$w[k] = (\mathrm{e}^{\mathrm{j}\psi} + n[k]) \cdot (\mathrm{e}^{\mathrm{j}\psi} + n[k-1])^*$$

ψ: arbitrary (but constant over at least two symbol intervals) carrier phase and symbol phase

With

$$\mathrm{Re}\{z_1 \cdot z_2^*\} = \left|\frac{z_1 + z_2}{2}\right|^2 - \left|\frac{z_1 - z_2}{2}\right|^2$$

we obtain

$$\mathrm{BER} = \Pr\left\{\frac{|Y_1|^2}{4} < \frac{|Y_2|^2}{4}\right\} = \Pr\{|Y_1| < |Y_2|\}$$

we have

- $Y_1 = 2\,\mathrm{e}^{\mathrm{j}\psi} + n[k] + n[k-1]$:
  complex Gaussian random variable with mean $2 \cdot \mathrm{e}^{\mathrm{j}\psi}$ and variance
  $2\sigma^2 = 4\sigma_I^2$, since $\mathrm{E}\{n[k] \cdot n[k-1]^*\} = 0$ (temporal orthogonality)

  => $|Y_1|$ is Rice–distributed with the parameters
  $2\sigma^2 = 2N_0/E_g$ and mean value $a = |2\,\mathrm{e}^{\mathrm{j}\psi}| = 2$

- $Y_2 = n[k] - n[k-1]$:
  complex, zero–mean Gaussian random variable with variance $2\sigma^2 = 4\sigma_I^2$

  => $|Y_2|$ is Rayleigh–distributed with the parameter $2\sigma^2$

- Random variables Y_1 and Y_2 are uncorrelated and thus statistically independent (Gaussian variables!):

$$\begin{aligned}\mathrm{E}\{Y_1 \cdot Y_2^*\} &= 2\,\mathrm{e}^{\mathrm{j}\psi} \cdot \mathrm{E}\{n[k]^* - n[k-1]^*\} + \mathrm{E}\{|n[k]|^2\} \\ &\quad - \mathrm{E}\{n[k] \cdot n[k-1]^*\} + \mathrm{E}\{n[k]^* n[k-1]\} - \mathrm{E}\{|n[k-1]|^2\} \\ &= 0 + \sigma_n^2 - 0 + 0 - \sigma_n^2 \;=\; 0\end{aligned}$$

Bit Error Probability:

$$\mathrm{BER} = \Pr(|Y_1| < |Y_2|) = \int\limits_0^\infty f_{\mathrm{Rice}}(y_1) \int\limits_{y_1}^\infty f_{\mathrm{Rayleigh}}(y_2)\,\mathrm{d}y_2\,\mathrm{d}y_1$$

Probability that a Rayleigh–distributed variable Y_2 (magnitude of a complex Gaussian variable with zero mean!) exceeds a Ricean–distributed variable Y_1 (magnitude of a complex Gaussian variable with non-zero mean)

Auxiliary calculation:

$$\int\limits_{y_1}^\infty f_{\mathrm{Rayleigh}}(y_2)\,\mathrm{d}y_2 = \int\limits_{y_1}^\infty \frac{y_2}{\sigma^2} \cdot \mathrm{e}^{-y_2^2/(2\sigma^2)}\,\mathrm{d}y_2 = \mathrm{e}^{-y_1^2/(2\sigma^2)}$$

Therefore, we obtain

$$\mathrm{BER} \;=\; \int\limits_0^{+\infty} \frac{y_1}{\sigma^2} \cdot \mathrm{e}^{-(a^2+y_1^2)/(2\sigma^2)} \cdot I_0(a y_1/\sigma^2) \cdot \mathrm{e}^{-y_1^2/(2\sigma^2)}\,\mathrm{d}y_1$$

Subst.: $z^2 = 2\,y_1^2$, i.e. $\mathrm{d}y_1 = \frac{1}{\sqrt{2}}\,\mathrm{d}z$ and $b^2 = a^2/2 = 2$

$$\begin{aligned}&= \int\limits_0^\infty \frac{z}{\sqrt{2}\sigma^2} \cdot \mathrm{e}^{-(2b^2+z^2)/(2\sigma^2)}\, I_0(bz/\sigma^2)\,\frac{\mathrm{d}z}{\sqrt{2}} \\ &= \mathrm{e}^{-b^2/(2\sigma^2)} \cdot \frac{1}{2}\int\limits_0^\infty f_{\mathrm{Rice}}(z)\,\mathrm{d}z \\ &= \frac{1}{2}\,\mathrm{e}^{-\frac{2}{2N_0/E_g}}\end{aligned}$$

$$\mathrm{BER} \;=\; \frac{1}{2}\,\mathrm{e}^{-E_b/N_0} \qquad \text{DPSK},\; M = 2$$

Comparison: Coherent 2PSK: $\mathrm{BER} = Q(\sqrt{2E_b/N_0})$

Chernoff bound for Q-function: $Q(x) \le \frac{1}{2}\,\mathrm{e}^{-x^2/2}$

=> The result for DPSK is equivalent to the Chernoff bound for coherent 2PSK

=> For very high SNR, *binary* DPSK is as power efficient as 2PSK!

General solution: $M > 2$

$$\mathrm{SER} < 1 + Q_M(X_1,X_2) - Q_M(X_2,X_1)$$

with

$$X_1 = \sqrt{\mathrm{ld}(M)(1-\sin(\tfrac{\pi}{M}))\tfrac{E_b}{N_0}} \quad ; \quad X_2 = \sqrt{\mathrm{ld}(M)(1+\sin(\tfrac{\pi}{M}))\tfrac{E_b}{N_0}}$$

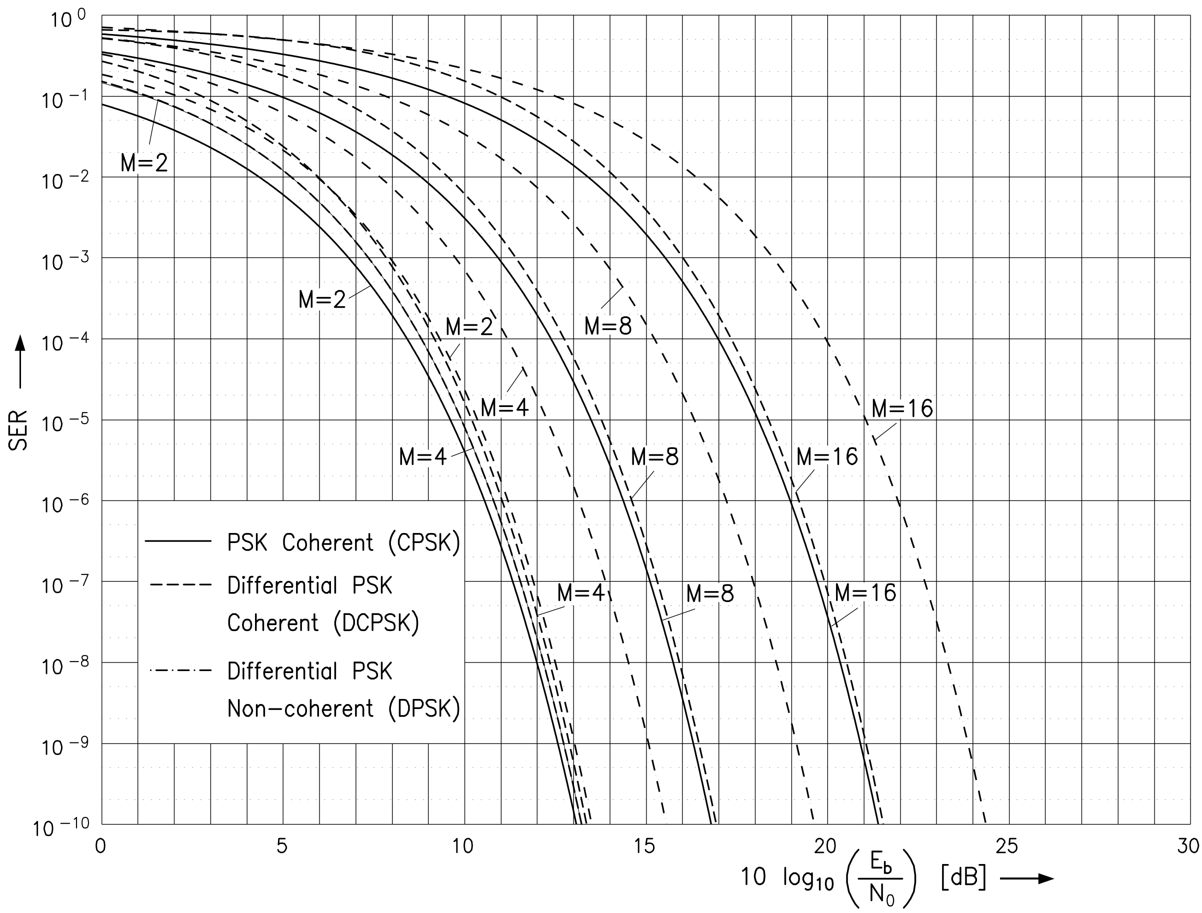
and “Marcum Q–Function”:

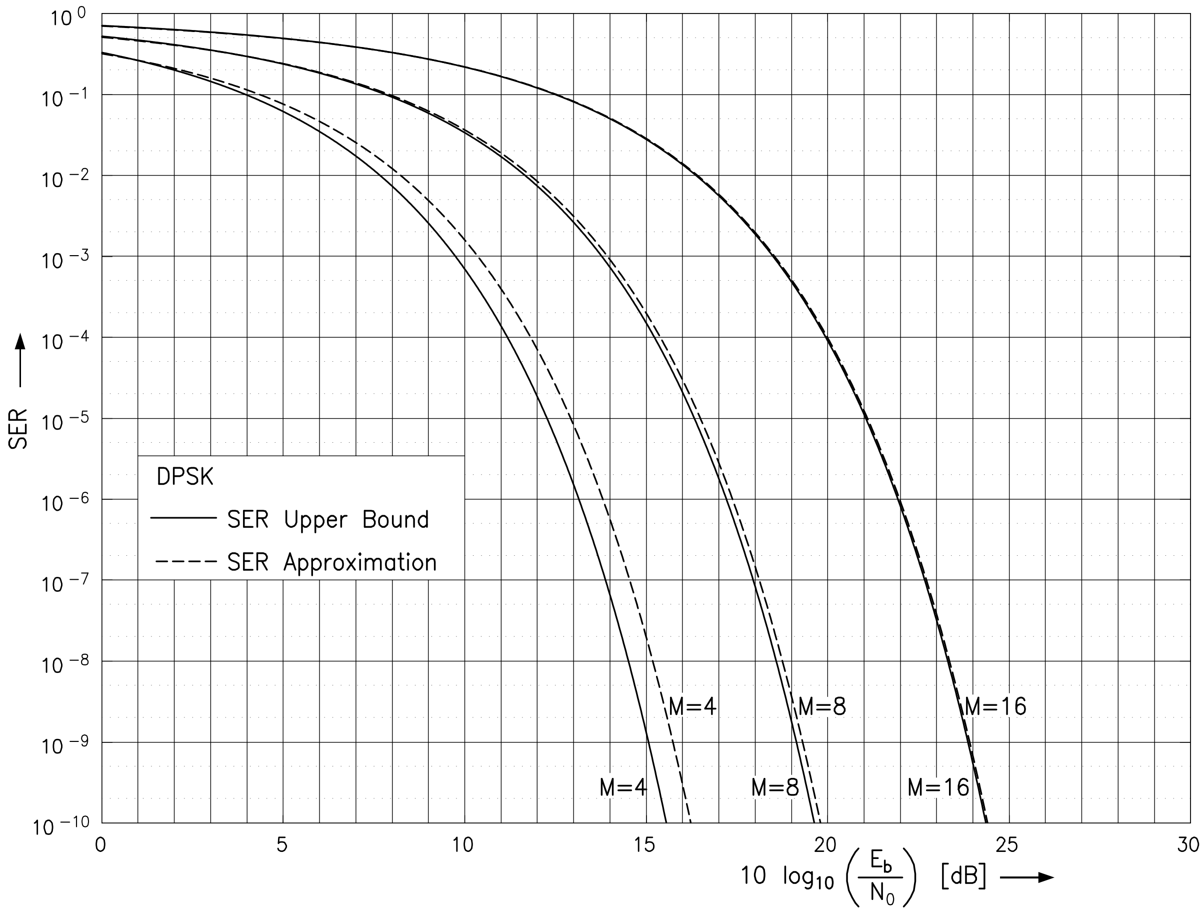
$$Q_M(y,x) := \int\limits_x^\infty r\,\mathrm{e}^{-(r^2+y^2)/2} \cdot I_0(yr)\,\mathrm{d}r$$

For $M \geq 4$ and $E_b/N_0 \to \infty$, the following approximation is useful

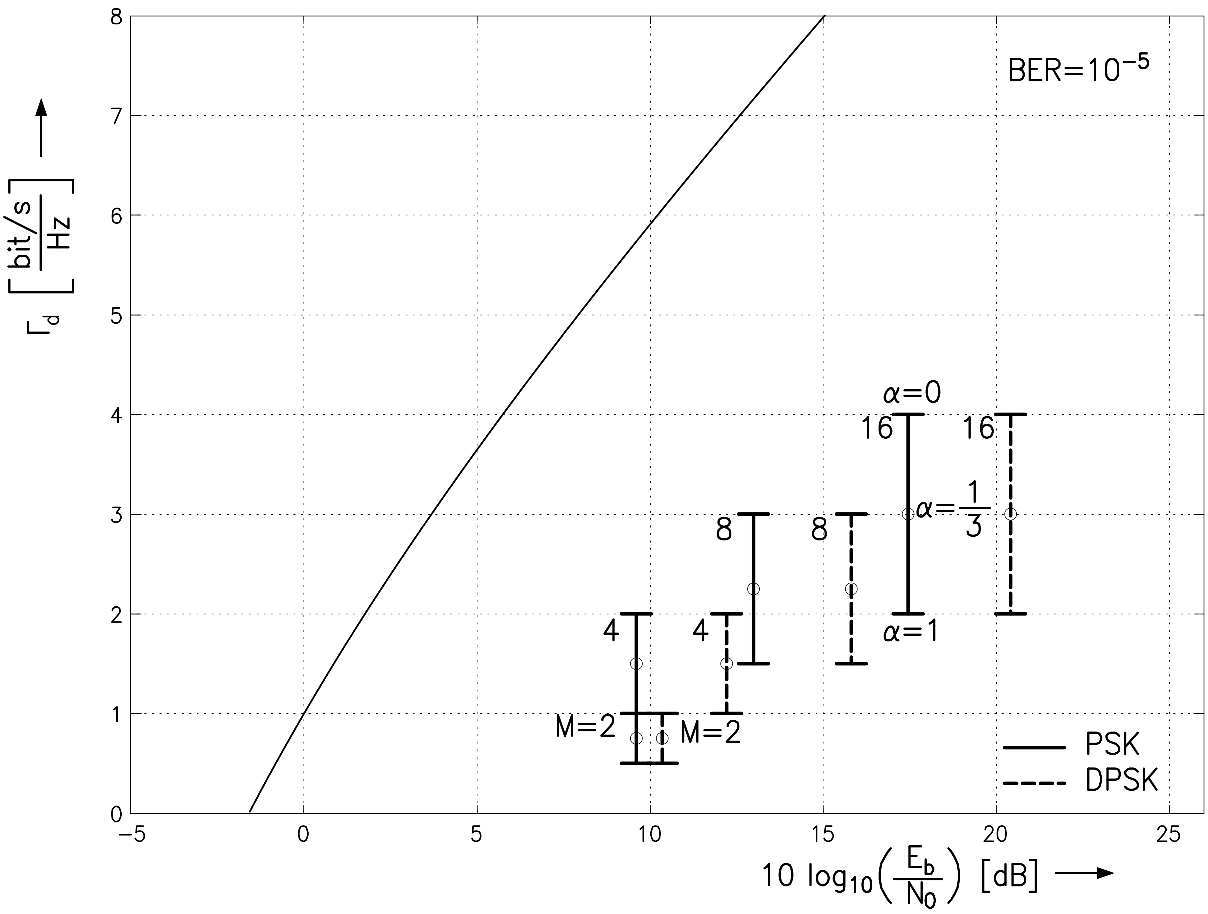
$$\mathrm{SER} \approx 2\,Q\left(\sqrt{d^2_{\min}\frac{E_b}{N_0}\Big/2}\right)$$

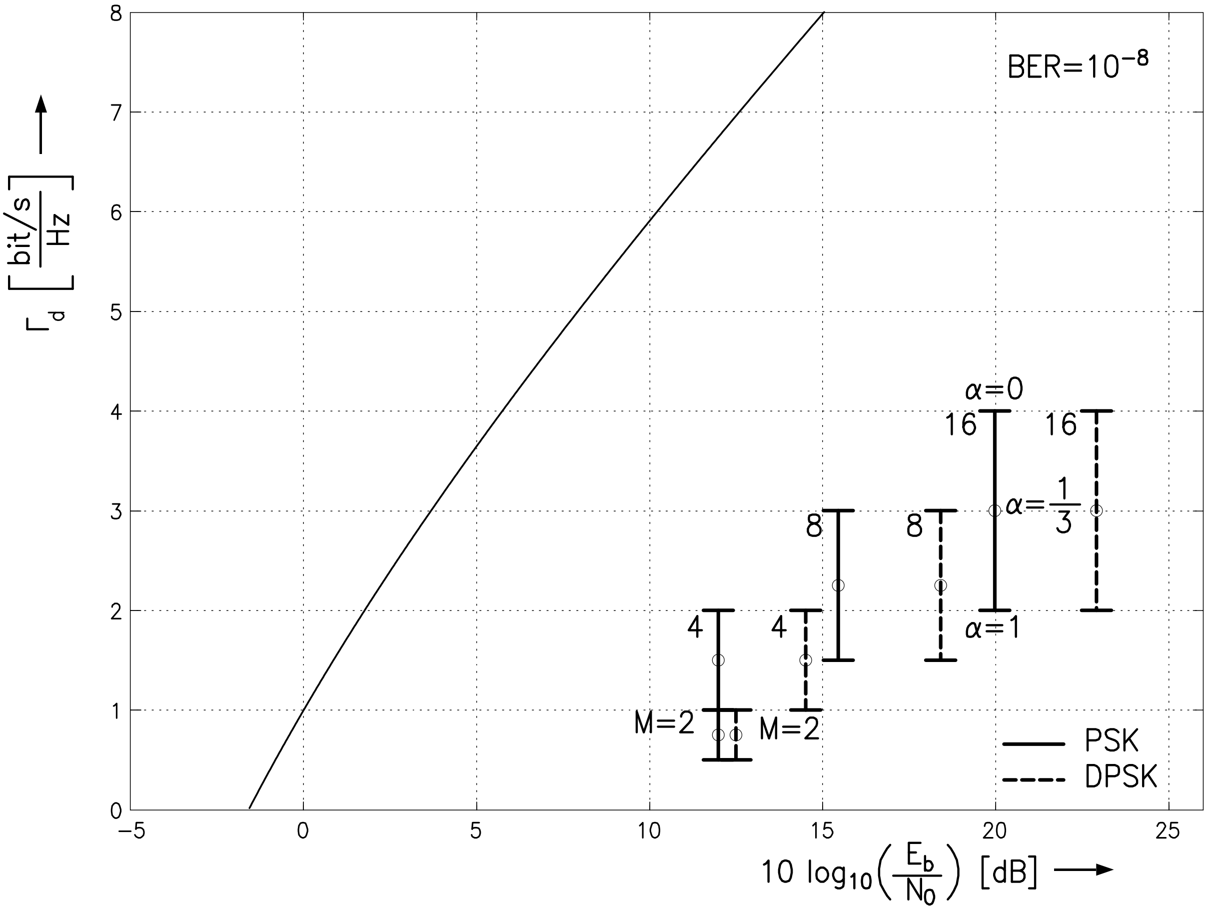
$\Longrightarrow$ 3 dB loss compared to coherent PSK




DPSK
SER Upper Bound
SER Approximation
M=4
M=8
M=16
M=4
M=8
M=16
SER
10 log10 (Eb/N0) [dB]

284

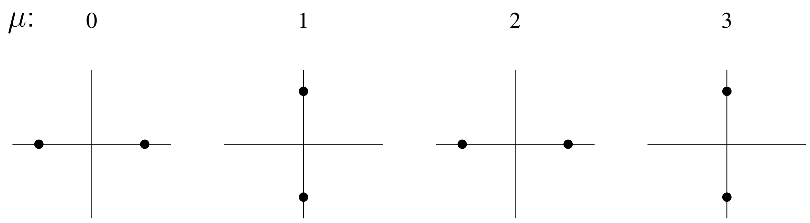

BER=10−5
α=0
α=1/3
α=1
16
16
8
8
4
4
M=2
M=2
PSK
DPSK
Γd [bit/s / Hz]
10 log10 (Eb/N0) [dB]



5.5.2.3 Rotating DPSK Schemes

- Phase–shifts of constellations in each second step
- **Purpose:** Minimization of the oscillations in the envelope of the transmit signal

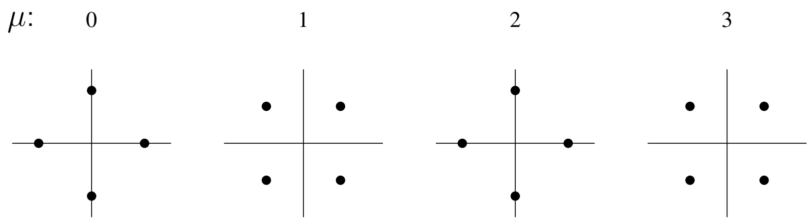
$M = 2 : \pi/2$–shift DPSK



Symbol 0: Phase jump $-\frac{\pi}{2}$; symbol 1: Phase jump $+\frac{\pi}{2}$

(very closely related to frequency modulation, especially with pulses
$g(t) = \sqrt{2E_g/(2T_\mathrm{b})}\,\cos(\pi t/(2T_\mathrm{b}))\,\mathrm{rect}(t/(2T_\mathrm{b}))$: MSK with frequency mapping!)

$M = 4$: $\pi/4$-shift QDPSK



Applications: – Japanese digital mobile radio system PDC (Personal Digital Cellular)

– American digital mobile radio system IS-54

5.5.2.4 Multiple Symbol Differential Detection

Loss of DPSK over coherent PSK:
Detection of $a_{m[k]}$ is based on a noisy reference point, i.e., the previously received symbol $d[k-1]$.

Idea (Wilson et. al. 1989, Divsalar and Simon 1990): More stable reference by basing decision on multiple symbols **simultaneously**

Multiple Symbol Detection

Multiple Symbol Differential Detection, MSDD

Method: Non–coherent correlation (matched filter) demodulation for **blocks of length** N of amplitude coefficients $a_{m[k]}$

Non–coherent ML **Sequence** Detection

Without loss of generality, the detection of the block begins at $k = 0$

$$\boldsymbol{a} = \left[a_{m[0]}, a_{m[1]}, \dots a_{m[N-1]}\right],$$

Inversion of the differential coding follows subsequently as for D**C**PSK

There are M^N possible transmit signals $s_i(t) = \sum\limits_{l=0}^{N-1} \tilde{a}_l^{(i)} g(t-lT)$, for which the correlation values K_i must be defined:

$$K_i = \frac{1}{E_g}\int\limits_{-\infty}^{+\infty} e(t)\cdot s_i^*(t)\mathrm{d}t = \sum_{l=0}^{N-1}\tilde{a}_l^{(i)*}\frac{1}{E_g}\int\limits_{-\infty}^{+\infty} e(t)\cdot g^*(t-lT)\mathrm{d}t$$

$$\boxed{K_i = \sum_{l=0}^{N-1} d[l]\tilde{a}_l^{(i)*}} \quad \text{with } d[l] = \frac{1}{E_g}\int\limits_{-\infty}^{+\infty} e(t)\cdot g^*(t-lT)\mathrm{d}t$$

=> All correlation values K_i can be obtained from the vector of N samples at the output of the matched filter (for the fundamental pulse $g(t)$)

$$\boldsymbol{d} = \left[\, d[0], d[1], \ldots, d[N-1] \,\right]$$

for all M^N hypothetical vectors

$$\tilde{\boldsymbol{a}}_i = \left[\, \tilde{a}_0^{(i)}, \tilde{a}_1^{(i)}, \ldots, \tilde{a}_{N-1}^{(i)} \,\right]$$

with

$$K_i = \boldsymbol{d}\,\tilde{\boldsymbol{a}}_i^{\mathsf{H}} = \sum_{l=0}^{N-1} d[l]\tilde{a}_l^{(i)*}$$

Some facts for DPSK:

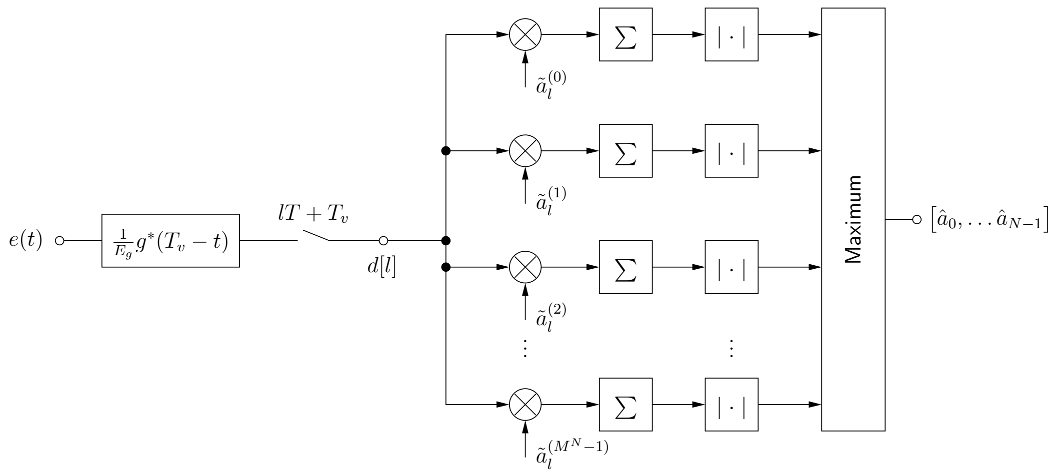
- All M^N hypothetical signals $s_i(t)$ have the same energy $N\cdot E_g$, since $\left|\tilde{a}_l^{(i)*}\right| = 1$ and $g(t)$ is a $\sqrt{\text{Nyquist}}$ pulse.
  Subtraction of the energy term E_i/E_g can be discarded, as it is irrelevant to the maximum (see page 248).
- The weighting function $k_{\mathrm{nc}}(\cdot)$ is monotonically increasing and also irrelevant for the maximum

$$\boxed{\text{Decision Rule } \hat{\boldsymbol{a}} = \left[\, \hat{a}_0^{(i)}, \hat{a}_1^{(i)}, \ldots, \hat{a}_{N-1}^{(i)} \,\right] = \underset{\forall \tilde{\boldsymbol{a}}_i}{\mathrm{argmax}} \left|\boldsymbol{d}\tilde{\boldsymbol{a}}_i^{\mathsf{H}}\right| \text{ for MSDD}}$$

- Without loss of generality, $\tilde{a}_0^{(i)} = 1\ \forall i$ may be chosen: only M^{N-1} hypotheses necessary

$$|K_i| = \left|\sum_{l=0}^{N-1} d[l]\tilde{a}_l^{(i)*}\right| = \underbrace{\left|\tilde{a}_0^{(i)*}\right|}_{=1}\cdot\left|d[0] + \sum_{l=1}^{N-1} d[l]\tilde{\tilde{a}}_l^{(i)*}\right|$$

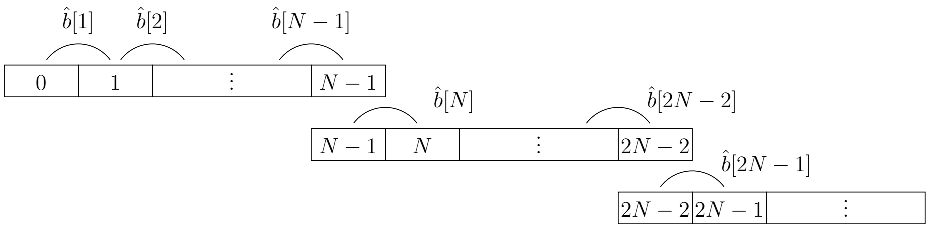
with $\tilde{\tilde{a}}_l^{(i)*} = \tilde{a}_l^{(i)*}/\tilde{a}_0^{(i)*} \in \mathcal{A}$; i.e., auxiliary symbols $\tilde{\tilde{a}}_l^{(i)*}$ are elements of the M PSK constellation.



Determination of the $N-1$ difference symbols from the N estimated PSK symbols $\hat{a}_l$

$$\hat{b}[l] = \frac{M}{2\pi} \arg\{\hat{a}_l \hat{a}_{l-1}^*\} \quad l = 1, 2, \dots, \; N-1$$

For continuous transmission, **overlapping** detection blocks are formed at the receiver side



292

Special case $N = 2$: Conventional detection

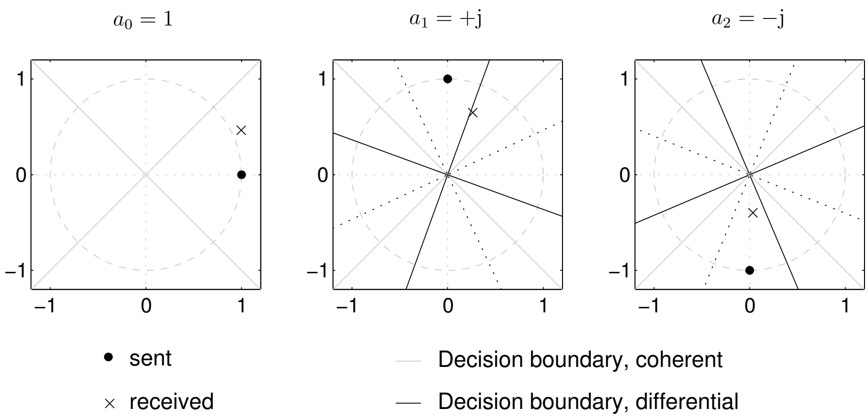
$$\hat{b}[k] = \underset{b \in \{0, \dots M-1\}}{\operatorname{argmax}} \left| 1 + \frac{d[k]}{d[k-1]} \cdot \mathrm{e}^{-\mathrm{j}2\pi b/M} \right| \quad \text{(with } a_{k-1} = 1 \text{ we get } b_k = \tfrac{M}{2\pi} \arg\{a_k\}\text{)}$$

Advantages: The generation of an average, reference phase inherently defined over $N-1$ symbols increase the power efficiency **(large gains are possible especially in the case of (slowly) time–varying channels!)**

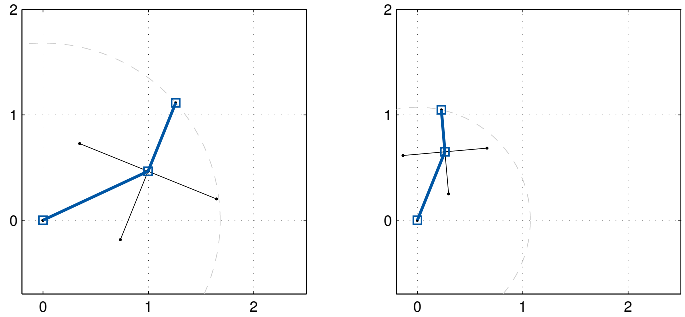
Disadvantages:

- High complexity due to the testing of M^{N-1} hypotheses instead of just M.
- Sufficient stability of the carrier phase over NT instead of just $2T$ required (only relatively low carrier frequency errors are supported. This may be probelmatic in mobile communications because of the Doppler effect.).

Example: $M = 4$: *4DPSK Comparison of Conventional Detection ($N = 2$) with MSDD with $N = 3$*

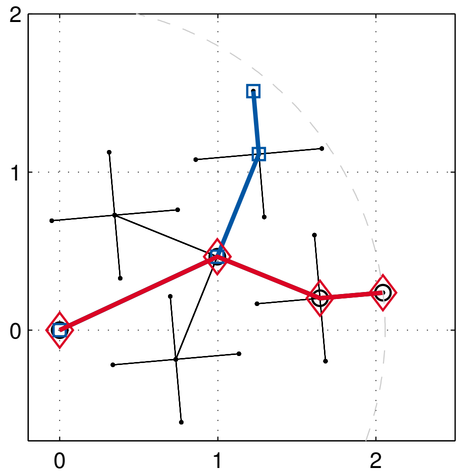


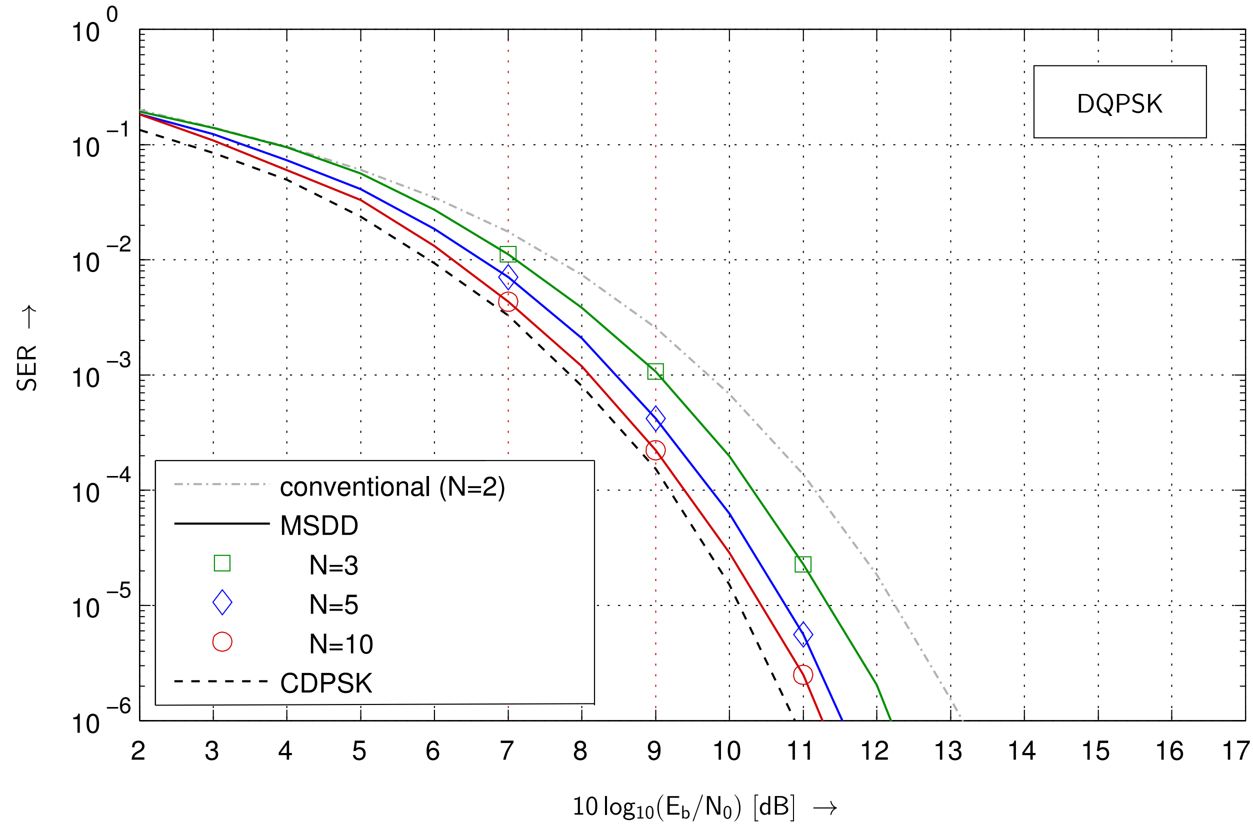
Conventional detection:



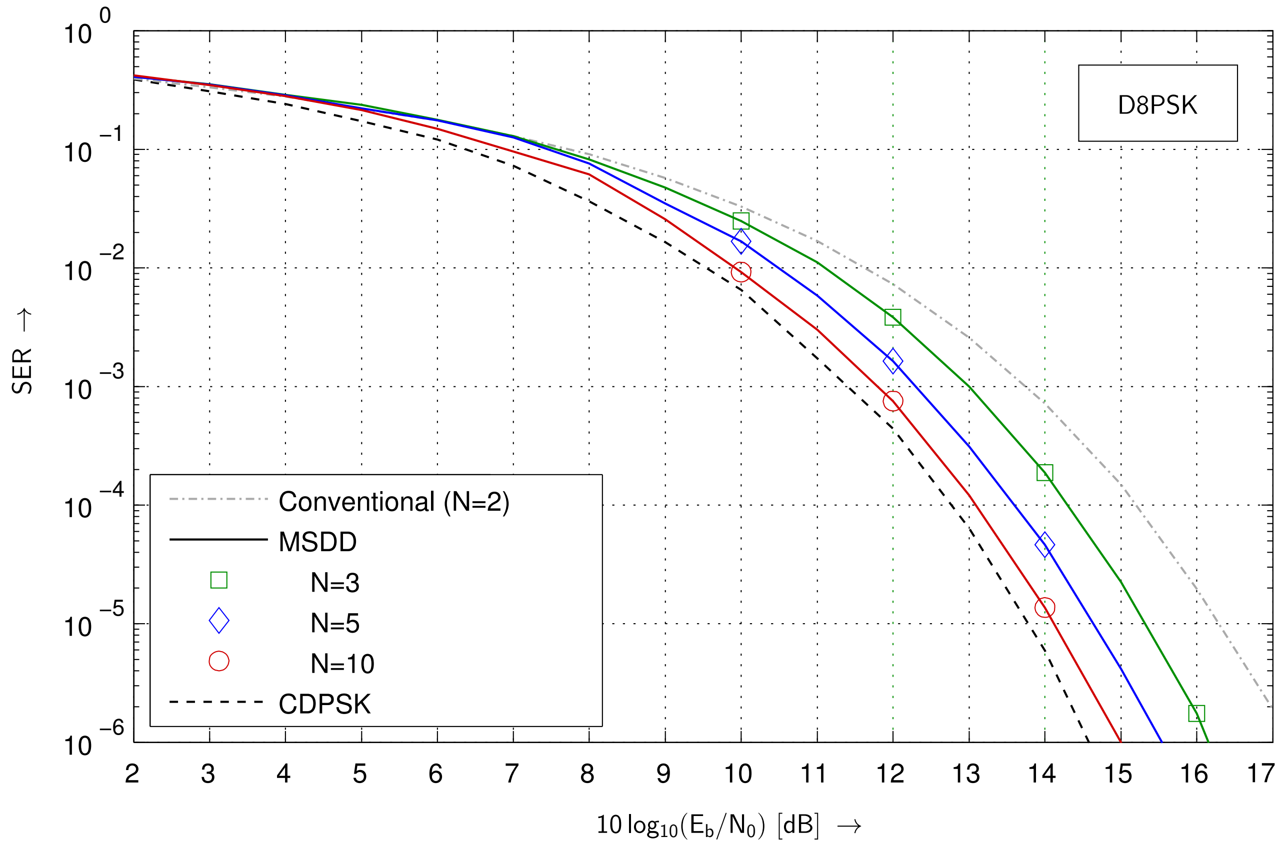
294

MSDD $N = 3$:




DQPSK
conventional (N=2)
MSDD
N=3
N=5
N=10
CDPSK
SER →
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
$10\log_{10}(E_b/N_0)$ [dB] →

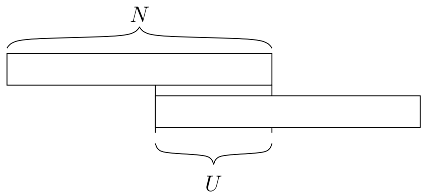
296


D8PSK
Conventional (N=2)
MSDD
N=3
N=5
N=10
CDPSK
SER →
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
$10\log_{10}(E_b/N_0)$ [dB] →

Modifications for MSDD:

1. Intelligent search algorithms for simplification of the search of $\underset{\forall \tilde{\boldsymbol{a}}_i}{\operatorname{argmax}} \left| \boldsymbol{d}\tilde{\boldsymbol{a}}_i^{\mathsf{H}} \right|$ using (on average) much fewer hypotheses (sphere decoder; Mackenthun Algorithm, 1994)

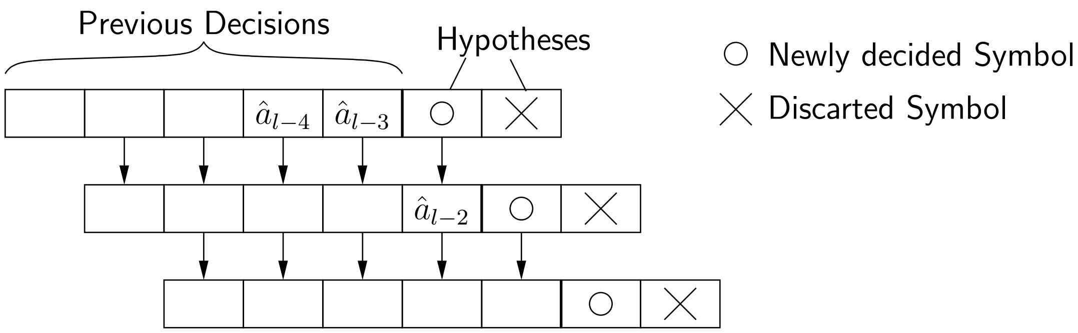
2. Decision–Feedback Multiple Symbol Differential Detection (DF-MSDD)
   Construction of blocks of length N, which overlap each other by $U > 1$ symbols

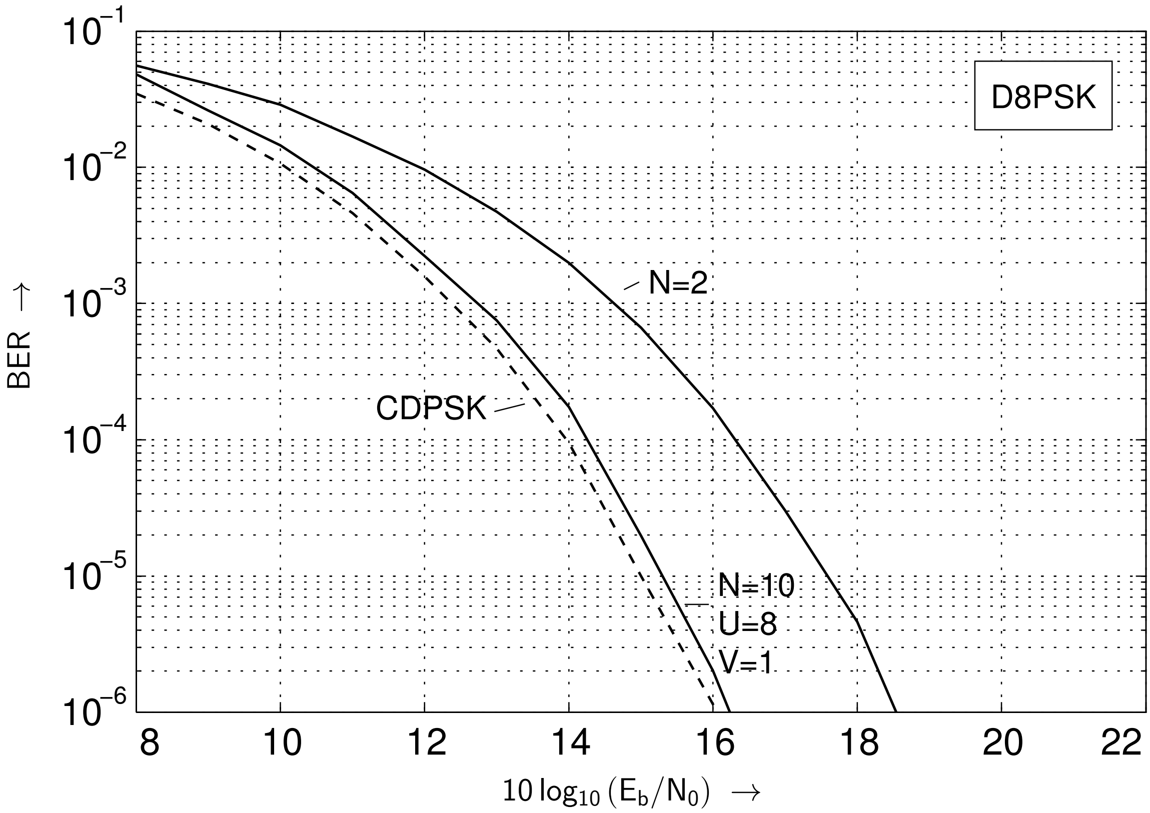


   Substitution of U with the already decided $\hat{a}_l$ from the current block:
   - Only M^{N-U} hypotheses are necessary, but the reference phase is still based on N symbol intervals.
   - Error propagation effects (especially at low SNR)

3. Subset Multiple Symbol Differential Detection (Subset-MSDD)
   Reliability of the estimated symbols $\hat{a}_l$ is lower at the block edges than at the middle:
   - Discarding of V symbols at block edges (especially the last symbols $b[N-1]$). Overlapping of the blocks by $V+1$ symbols.

4. Combination of 2 and 3: Subset DF-MSDD
   Particulary advantageous variant with respect to achievable efficiency and reduction of computational complexity. Choose N to be as large as possible (i.e., as large as the phase stationarity allows).
   $U = N-2$, $V = 1$, detection of M^2 hypotheses for each symbol, discarding of the last symbol, block progression by one.




D8PSK
N=2
CDPSK
N=10
U=8
V=1
BER →
10 log10 (Eb/N0) →
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
8
10
12
14
16
18
20
22

6 Digital Frequency Modulation (FSK) and Continuous Phase Modulation (CPM)

6.1 Digital Frequency Modulation (FSK)

Acronym: FSK = Frequency Shift Keying

Signal elements for M–level FSK (ECB domain):

$$s_m(t) = \begin{cases} \sqrt{E_g/T}\ \mathrm{e}^{\mathrm{j}(2\pi m\,h\,t/T+\varphi_0)} & \text{for } t \in [0,\,T) \\ \\ 0 & \text{for } t \notin [0,\,T) \end{cases}$$

with $m = 0, 1, \ldots, M-1$

Modulation index: h

= Smallest frequency difference between signal elements, expressed in oscillation periods per modulation interval T

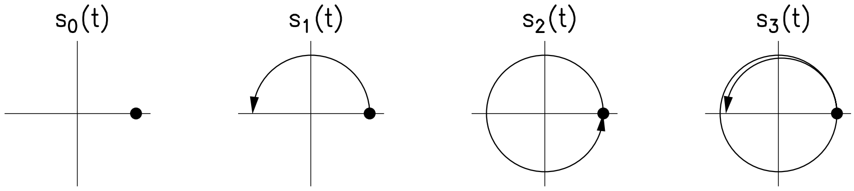
Transmit signal:

$$s(t) = \sum_{k=-\infty}^{+\infty} s_{m[k]}(t-kT)$$

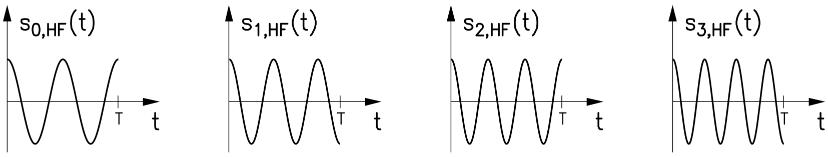
Since the time–limited signal elements do not overlap, the temporal orthogonality condition is obviously satisfied.

Example: $M = 4$, $h = 1/2$

Equivalent lowpass signal elements:



Example for corresponding real signal elements $(f_0 = 2/T)$:



Remark: The transformation frequency f_0 is in this case *not* equal to the carrier frequency f_c. Specifically, the following holds:

$$\text{Carrier frequency:} \quad f_c = f_0 + h\frac{(M-1)}{2T}$$

Discussion:

- For $h \notin \mathbb{Z}$ and $\varphi_0 = \text{const.}$, phase jumps appear at the boundaries of the modulation intervals:

  discontinuous FSK

- Within the interval k, to avoid this behavior we may set the phase to

$$\varphi_0[k] = \lim_{\Delta t \to 0} \arg(s_{m[k-1]}(T - \Delta t))$$

  In this way, FSK is implemented with continuous phase at the interval boundaries:

  Continuous Phase Frequency Shift Keying (CPFSK)

  Continuous phase FSK implies inherent *coding* necessary to ensure the phase continuity, which can increase both power– and bandwidth–efficiency.

- **Special case:**
  For $h = 1$ (generally $h \in \mathbb{Z}$), a CPFSK signal always results, but without any inherent coding as each signal element is interchangeable without altering the phase boundaries (Sunde's FSK).

Remark:

In Section 6.1, only FSK schemes are considered for which detection is carried out in a symbol–by–symbol manner. The extension to sequence estimation is presented in Section 6.2.

To analyze the increased robustness against noise created by the inherent coding in continuous phase FSK, techniques from channel coding have to be employed (maximum–likelihood–sequence estimation, Viterbi algorithm, see Supplement 2).

Average Power Spectral Density of FSK:

(without inherent coding, i.e., for statistically independent, equally probable signal elements)

$$\bar{\Phi}_{ss}(f) \;=\; \frac{1}{TM}\sum_{m=0}^{M-1}\left(|S_m(f)|^2 - |\bar{S}(f)|^2\right) + \frac{1}{T^2}\left|\bar{S}(f)\right|^2 \sum_{i=-\infty}^{+\infty}\delta(f-i/T)$$

with

$$S_m(f) \;=\; \sqrt{E_g T}\;\mathrm{si}(\pi T(f - h\cdot \tfrac{m}{T}))\cdot \mathrm{e}^{\mathrm{j}(\varphi_0-\pi f T)}\cdot \mathrm{e}^{\mathrm{j}\pi h m}$$

$$\bar{S}(f) \;=\; \frac{1}{M}\sum_{m=0}^{M-1} S_m(f)$$

For $\bar{S}(f) \not\equiv 0$, i.e., for

$$\sum_{m=0}^{M-1} s_m(t) \not\equiv 0\;,$$

discrete spectral lines appear which are a separation of $\frac{1}{T}$. These lines consume power but do not carry any information!

6.1.1 Binary FSK

Squared Euclidean distance:

$$d_E^2 = \frac{1}{E_g}\int_0^T |s_0(t) - s_1(t)|^2 \, \mathrm{d}t = 2(1 - \mathrm{Re}\{\rho\})$$

with

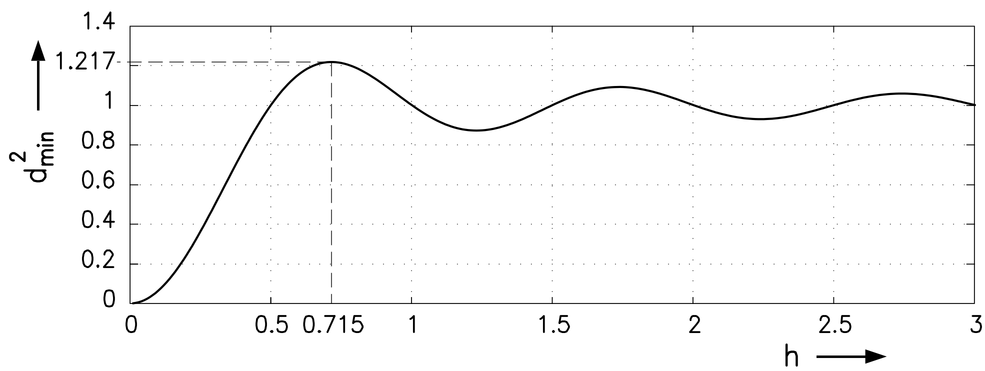
$$E_g = \int_0^T |s_m(t)|^2 \, \mathrm{d}t \,, \qquad m = 0, 1$$

Correlation coefficient:

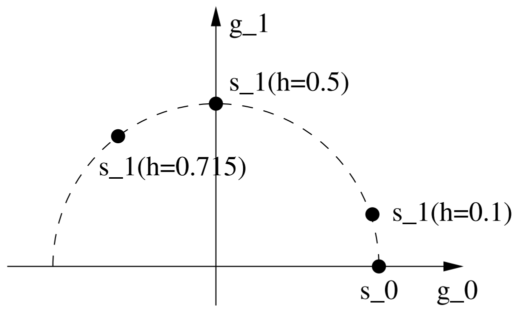
$$\begin{aligned}\rho &= \frac{1}{E_g}\int_0^T s_0(t)\, s_1^*(t) \, \mathrm{d}t \\ &= \mathrm{j}\,\frac{\mathrm{e}^{-\mathrm{j}2\pi h} - 1}{2\pi h} = \frac{\sin(2\pi h)}{2\pi h} + \mathrm{j}\,\frac{\cos(2\pi h) - 1}{2\pi h}\end{aligned}$$

Since in this case, $E_b = E_g$ holds, we have

$$d_{\min}^2 = \frac{E_g}{2E_b}\, d_E^2 = (1 - \mathrm{si}(2\pi h))$$



Signal space illustration:



Signal elements are *orthogonal* (with respect to coherent demodulation) for $\mathrm{Re}\{\rho\} = 0$, which is equivalent to

$$h = \frac{i}{2} \qquad i \in \mathbb{Z} \quad ; \quad d^2_{\min} = 1$$

Minimal frequency difference between orthogonal signal elements:

$$h = 1/2 \qquad \text{Minimum Shift Keying (MSK)}$$

Note: There is a problem:

Here, $d^2_{\min} = 1$ for MSK,
but $d^2_{\min} = 2$ for MSK when interpreted as Offset–QPSK?

Solution:

Take advantage of the coding inherent to the continuous phase!

Maximum Euclidean distance for

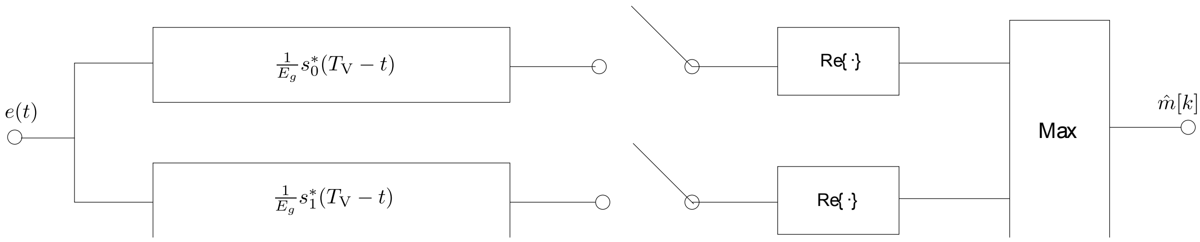
$$h = 0.715\ , \qquad d^2_{\min} = 1.217$$

A modulation index ($\widehat{=}$ frequency deviation) of $h > 0.715$ is not beneficial, since bandwidth consumption is increased without a gain in power efficiency!

$\Longrightarrow$ (apparently) very different from analog FM!

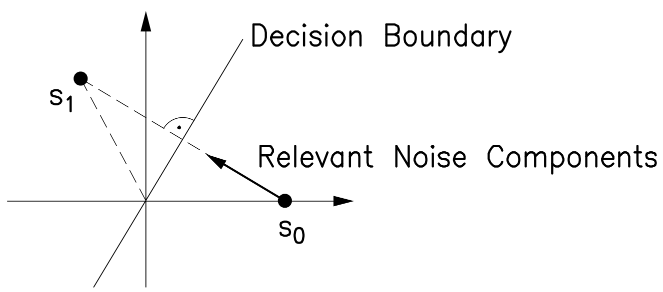
6.1.1.1 Coherent Demodulation of Binary FSK

Optimal receiver: (without decoding!)



=> Bank of matched filters

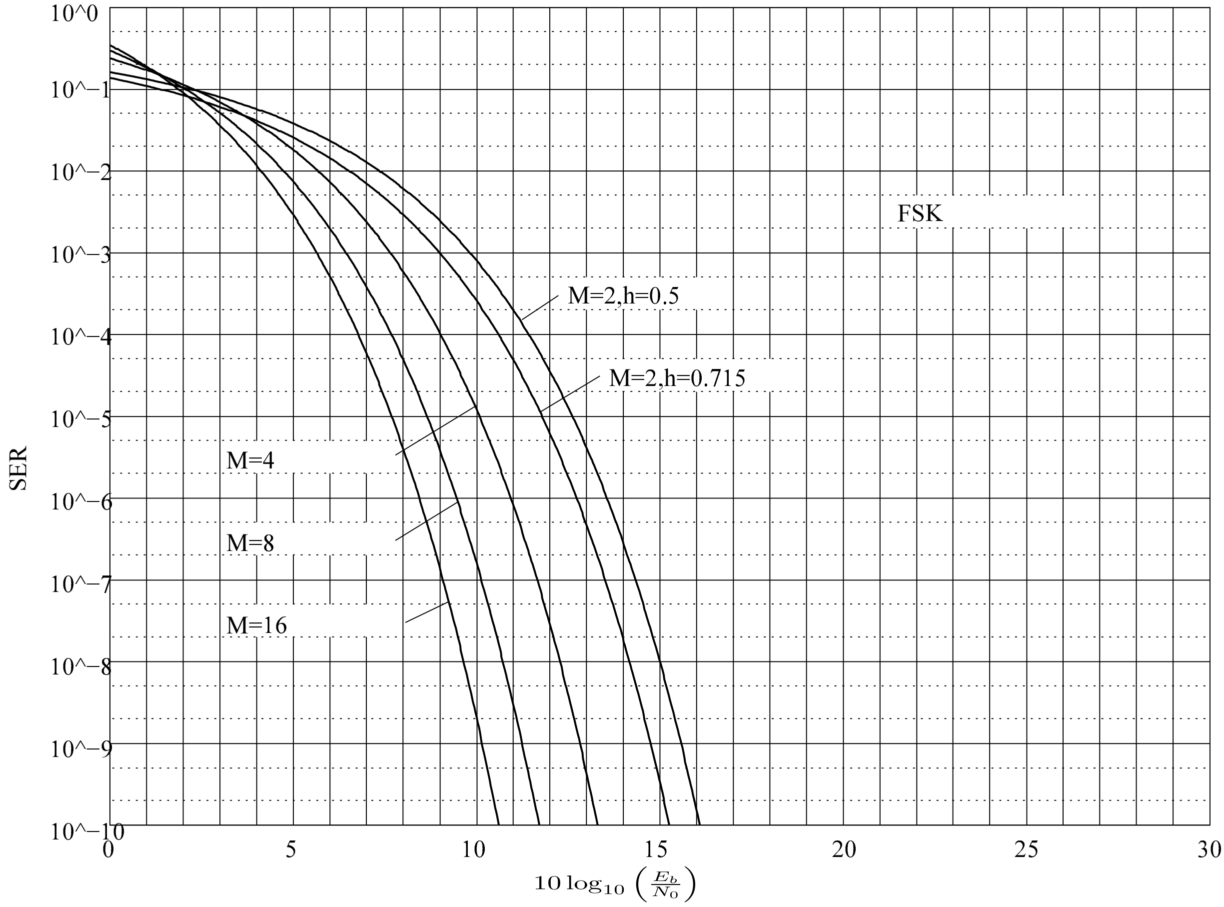
Bit error probability:



Since the variance of the Gaussian noise in the direction of the other signal points is equal to $\sigma_I^2 = \frac{N_0}{2E_g}$, we obtain

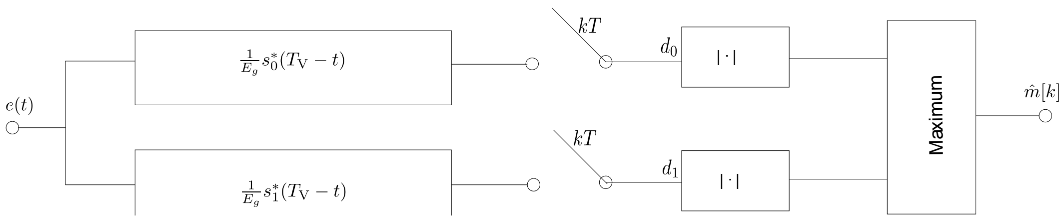
$$\begin{aligned} \mathrm{BER} &= \mathrm{Q}\left(\frac{d_E/2}{\sigma_I}\right) = \mathrm{Q}\left(\sqrt{d_{\min}^2 \frac{E_b}{N_0}}\right) \\ &= \mathrm{Q}\left(\sqrt{(1 - \mathrm{si}(2\pi h))\frac{E_b}{N_0}}\right) \end{aligned}$$

Examples:	$h = 0.5$:	3 dB loss with respect to 2PSK
	$h = 0.715$:	2.16 dB loss with respect to 2PSK

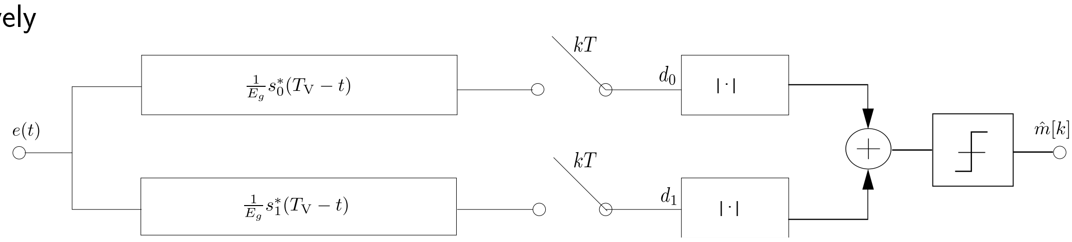


6.1.1.2 Non–coherent Demodulation of Binary FSK

Optimal receiver (without decoding!):



Alternatively



Discriminator as widely used for analog FM!

Bit Error Probability:

Without loss of generality, we assume: $s_0(t)$ was transmitted. Sampling values at the output of the matched filter:

$$\begin{aligned} d_0 &= 1 + n_0 \\ d_1 &= \rho + n_1 \end{aligned}$$

with

$$\begin{aligned} \mathrm{E}\{|n_0|^2\} &= \mathrm{E}\{|n_1|^2\} = N_0/E_g \\ \mathrm{E}\{n_0 n_1^*\} &= \rho^* N_0/E_g \end{aligned}$$

$|d_0|$ and $|d_1|$ are **correlated** Ricean–distributed random variables.

The bit error rate may be calculated as the probability that one Ricean-distributed variable exceeds another:

$$\mathrm{BER} = \Pr\left(|d_1| > |d_0|\right)$$

The analysis analogous to that of DPSK. After some calculations, we obtain

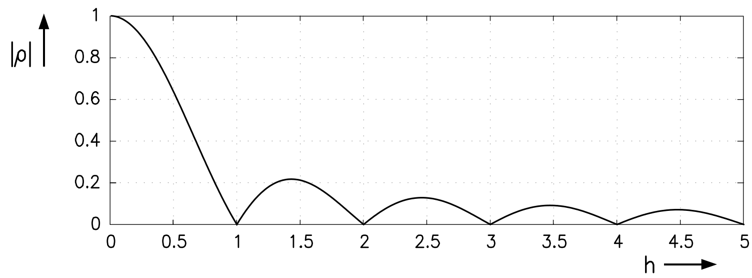
$$\mathrm{BER} = Q_M(x,y) - \frac{1}{2}\,\mathrm{e}^{-(x^2+y^2)/2}\mathrm{I}_0(x \cdot y)$$

with

$$\begin{aligned} x &= \sqrt{\frac{E_b}{2N_0}(1 - \sqrt{1-|\rho|^2})} \\ y &= \sqrt{\frac{E_b}{2N_0}(1 + \sqrt{1-|\rho|^2})} \end{aligned}$$

316

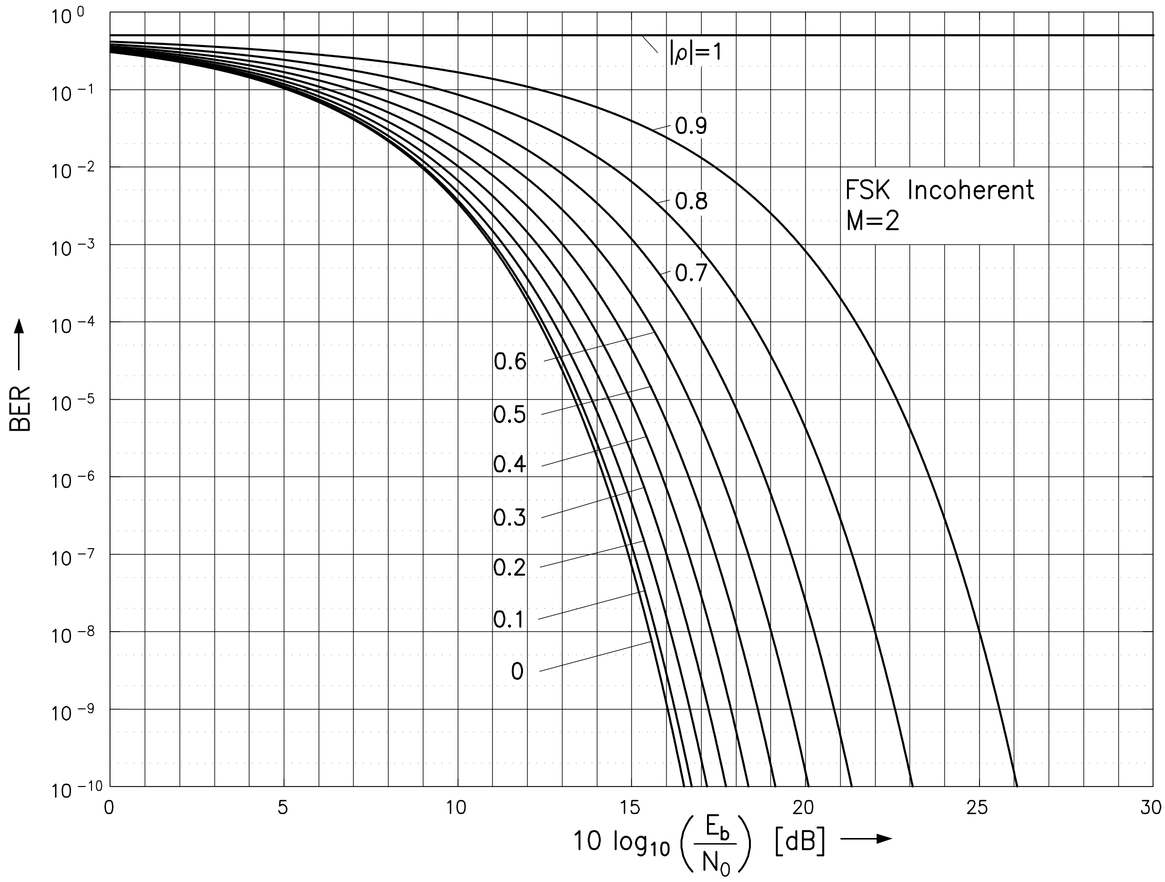
Correlative coefficient (*with respect to non–coherent demodulation*):



For orthogonality $|\rho| = 0$ => $h \in \mathbb{N}$

In this case, we have $x = 0$, $y = \sqrt{E_b/N_0}$, and

$$\mathrm{BER} = \tfrac{1}{2}\,\mathrm{e}^{-E_b/(2N_0)}$$



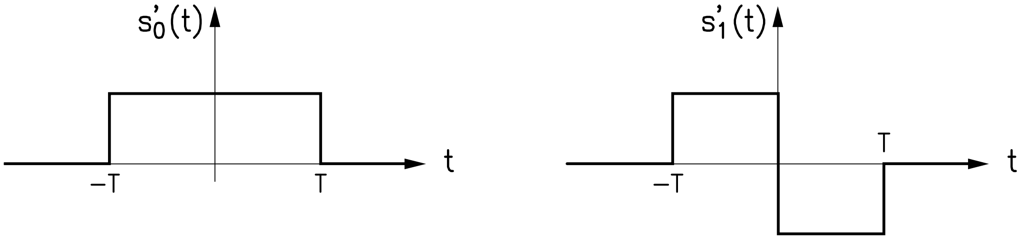
Observation: For $h \in \mathbb{N}$, 2FSK is exactly 3 dB less power efficient than 2DPSK

Interpretation:

2DPSK interpreted as two virtual orthogonal signal elements of length $2T$

$$\begin{aligned} s_0'(t) &= g(t+T) + g(t) \qquad \text{no phase jump} \\ s_1'(t) &= g(t+T) - g(t) \qquad \text{phase jump} \end{aligned}$$

for example, hard keying, rectangular pluses:



Both virtual signal elements $s'_0(t)$ and $s'_1(t)$ are obviously orthogonal. For $\sqrt{\text{Nyquist}}$–Impulses $g(t)$, they are also temporally orthogonal.
Since

$$\int\limits_{-\infty}^{+\infty} |s'_0(t)|^2 \,\mathrm{d}t = \int\limits_{-\infty}^{+\infty} |s'_1(t)|^2 \,\mathrm{d}t = 2E_g$$

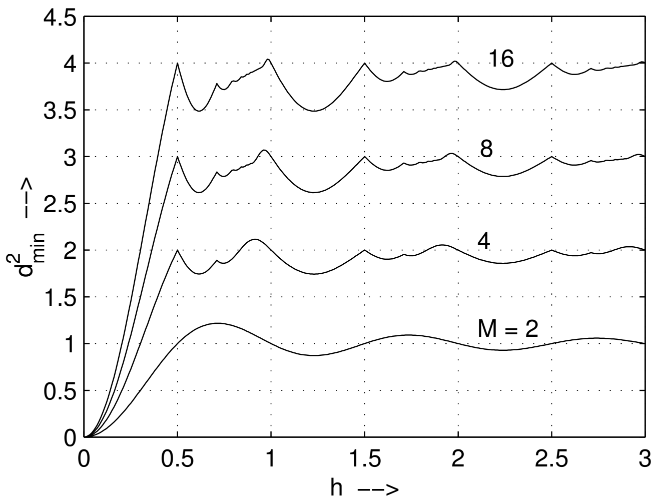
<u>twice</u> the impulse energy is exploited! Each signal element is used *twice* during the detection process: $3\ \mathrm{dB}$ better than 2 FSK with $|\rho| = 0$

6.1.2 Multi–Level FSK

Minimum squared Euclidean distance:

With $E_b = E_g/\mathrm{ld}(M)$ for M-ary FSK:

$$d^2_{\min} = \frac{E_g}{2E_b} \min_{\substack{\forall s_i(t), s_j(t) \\ s_i(t) \neq s_j(t)}} \frac{1}{E_g} \int\limits_0^T |s_i(t) - s_j(t)|^2 \,\mathrm{d}t = \mathrm{ld}(M) \min_{m=1,2,\ldots,M-1} (1 - \mathrm{si}(2\pi m h))$$



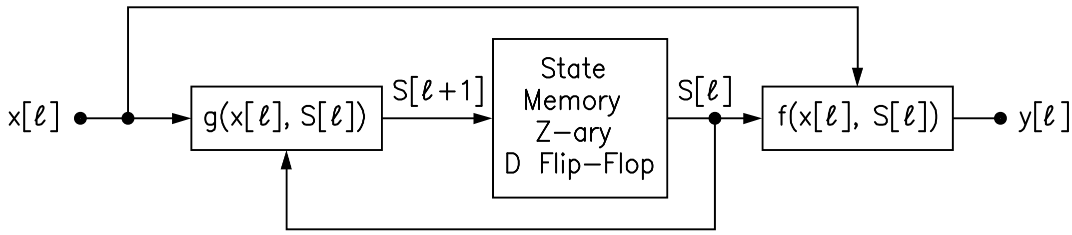
Supplement 2: *Trellis Coding*

In trellis coding, a sequence of M_y–ary output symbols $y[l]$ is created from a sequence of M_x–ary input symbols $x[l]$ using a **Mealy machine**, where previously observed input symbols influence the coding process. We restrict our consideration to the following case:

$$M_x = 2^k\,; \quad M_c = 2\,; \quad M_y = 2^n\,; \quad R_c = \frac{k}{n}$$

Normally, k and n are chosen to be very small. The Mealy machine possesses Z different internal states $S[l] \in \{0, 1, \ldots, Z-1\}$, $l \in \mathbb{Z}$.

Trellis encoder represented as a Mealy machine:



Output function: Calculation of output symbols $y[l]$

$$y[l] = f\big(x[l], S[l]\big)$$

State transition function: Calculation of subsequent state

$$S[l+1] = g\big(x[l], S[l]\big)$$

Transmission of the trellis encoder output symbol: Mapping onto $V, V \leq 1$, signal elements $s_{m[k]}(t)$

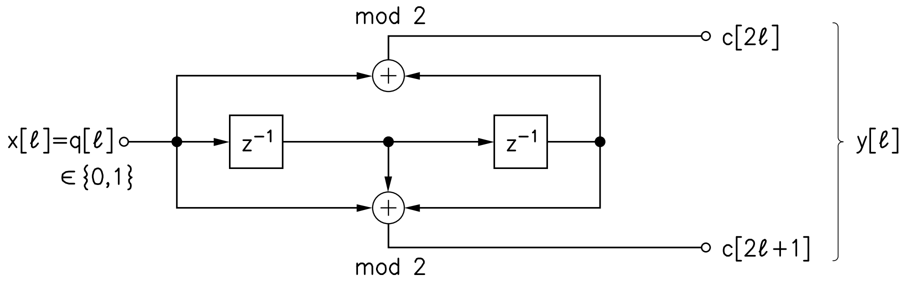
$$y[l] \longrightarrow m[l]$$

or alternatively,

$$y[l] \longrightarrow (m\,[\,l \cdot V\,], m\,[\,l \cdot V + 1\,], \ldots,\ m\,[\,l \cdot V + V - 1\,])$$

Example: *Trellis Encoder*

- Convolutional encoder
  e.g. code of rate $R_c = \frac{1}{2}$ with 4 states; $k = 1$; $n = 2$

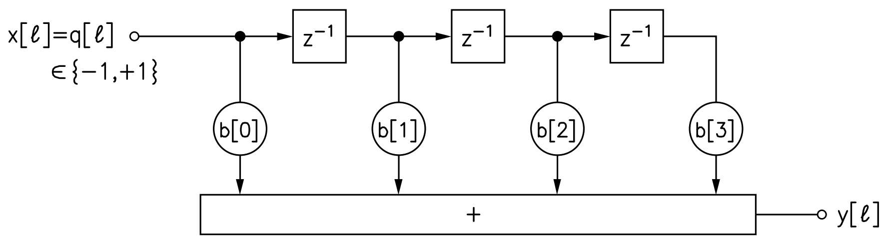


States:
All possible combinations for the contents of the shift register elements

State	Memory
0	00
1	01
2	10
3	11

Example: Transmission of both binary code symbols $c[2l]$ and $c[2l+1]$ using 4PSK

- Linear distortion of a binary, bipolar data signal: equivalent discrete–time block diagram for the samples at the output of a well-suited receiver input filter (so–called **Whitened**–Matched Filter (**not** Matched Filter)):



Filtering of the discrete time binary, bipolar data signals using a linear, time–invariant system $B(z) \longleftrightarrow b[l]$

$$y[l] \; = \; q[l] * b[l]$$

$$Y(z) \; = \; Q(z) \cdot B(z)$$

$y[l]$: multi–level ASK signal (e.g. for a 3rd order system with up to 16 levels; $M_y \leq 16$), the signal points are not uniformly arranged.

Trellis diagram of a trellis encoder:

Representation of all possible inputs and outputs of a Mealy machine as well as the state transitions over discrete time l.

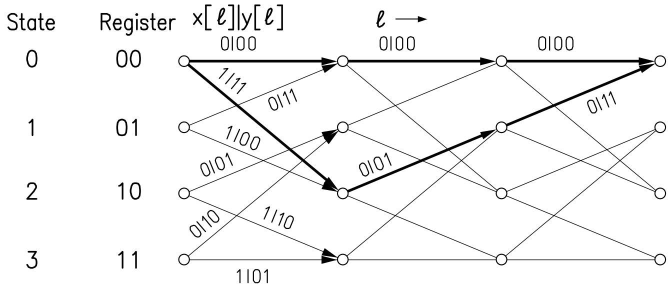
Possible state transitions are represented using **labelled branches**:

$$S[l] \circ \xrightarrow{x[l]|y[l]} \circ \; S[l+1]$$

Branch label: Input symbol x, which evokes a state transition, and output symbol y, which is then created by this state transmission.

Example: *Trellis diagram for the above encoder*

(Convolutional encoder from page 323)



A sequence of branches constitutes a **path** through the trellis diagram. Each message (i.e., a sequence of binary source symbols) creates one path. Not all paths are allowed: Code redundancy.

Task of **trellis decoding:**

Find the a–posteriori most probable path through the trellis diagram. With the help of the sequence of received signal samples $d[l]$, $l \in \mathbb{Z}$, at the output of the matched filter.

Optimal trellis decoding for MPSK transmission over an AWGN channel: Define the path through the trellis diagram for which the aggregate metric

$$\Lambda = \sum_{l=-\infty}^{+\infty} \lambda_{m[l]}[l] \quad \text{with}$$

$\lambda_{m[l]}[l]$: Metric of the (hypothetical) signal element $s_{m[l]}(t - lT)$ in the l–th time step of the received signal $d[l]$

is maximized.

Metric:	AWGN channel	negative squared Euclidean distance
	BSC	Hamming distance

Solution: **Viterbi Algorithm (1958/1967)**

1. At each discrete time l and for each of the Z possible states, the path to this state is selected for which the hitherto aggregate metric

$$\Lambda_i[l] = \sum_{\mu=-\infty}^{l} \lambda_{m[\mu]}[\mu] \quad \text{with } i = 0,1,\ldots,Z-1$$

is maximal.

Recursive solution: for $i = 0,\ldots,Z-1$

A. Define

$$\Lambda_i[l] = \max\left(\left(\Lambda_{i_1}[l-1] + \lambda_{m_1}[l]\right),\ldots,\left(\Lambda_{i_{M_x}}[l-1] + \lambda_{m_{M_x}}[l]\right)\right)$$

with

i_j : Previous state j, from which a branch to state i is generated by means of a signal element m_j. There are M_x different previous states $j, 1 \leq j \leq M_x$, which end in state i.

B. Take as the new path register for state i the **path register** of the selected previous state augmented by the source symbol which was generated by the current branch under consideration.

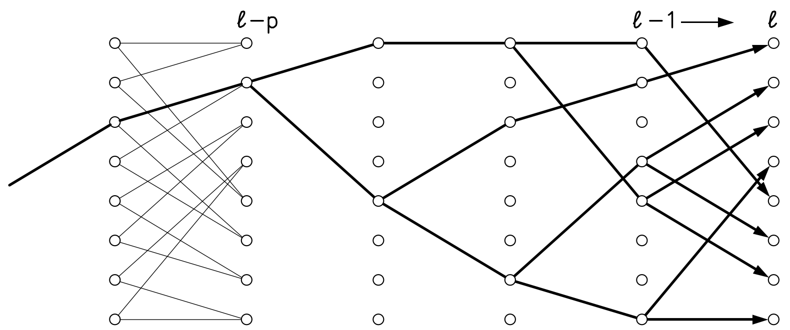
328

Explanation: Out of all paths which lead through state i at discrete time l only the one with the highest metric up to time l can also reach the highest total aggregate metric at the end of reception.

2. Once a state $i_{\max}$ has been reached, read from the path register the input symbol from time $l-p$ (i.e., the value recorded p time intervals previously, $x_{i_{\max}}[l-p]$) which has the larges aggregate value $\Lambda_i[l]$: $i_{\max} = \arg\max_i \Lambda_i[l]$

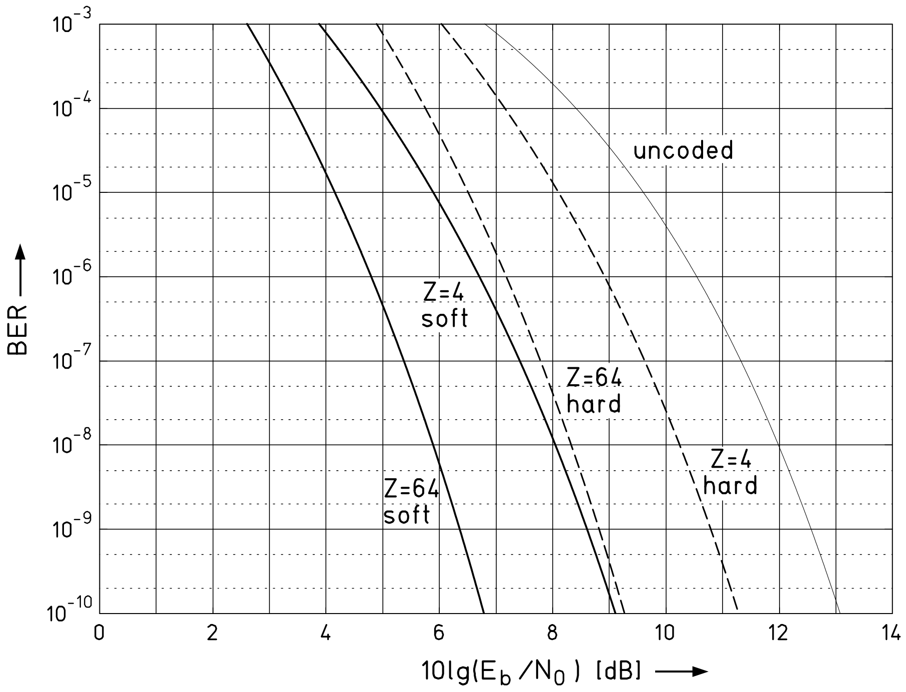
p: Path register length, decoding delay

Rule of thumb: $p \geq 5\,\mathrm{ld}(Z)$



Example: Bit error rates for convolutional codes of rate $\frac{1}{2}$ with Z states

Binary, bipolar transmission in each dimension (2ASK = 2PSK, 4PSK = 4QAM) of the signal space



End Supplement 2

6.2 Continuous Phase Modulation (CPM)

Frequency modulation with Continuous Phase boundaries

Data: $a[k] \in \{\pm 1, \pm 3, \ldots, \pm(M-1)\}$ $\quad$ (M even, see ASK)

Transmit signal:

$$s_{\mathrm{RF}}(t) = \sqrt{\frac{2E_s}{T}} \cos\left(2\pi f_c t + \phi(t)\right)$$

with phase $\phi(t)$ as "PAM signal" with fundamental pulse $q(t)$

$$\phi(t) = 2\pi h \sum_{k=-\infty}^{+\infty} a[k] q\left(t - kT\right) = 2\pi h \sum_{k=-\infty}^{+\infty} a[k] \int\limits_{-\infty}^{t} g\left(t' - kT\right) \mathrm{d}t'$$

$h = \dfrac{p}{q}$: modulation index ; $\qquad p, q \in \mathbb{N} \qquad \text{rel. prim}$

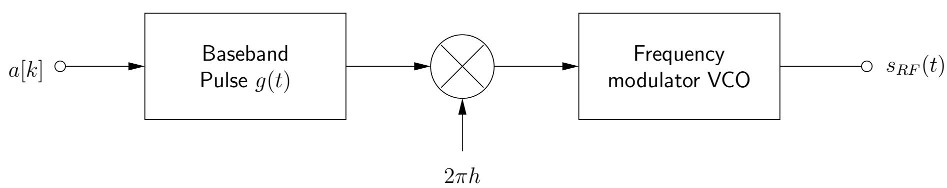
$g(t)$: frequency pulse, extended over L symbol durations T

$q(t) = \int\limits_{0}^{t} g\left(t'\right) \mathrm{d}t'$ $\quad$ phase impulse, normalization $\quad q\left(t \to \infty\right) = 1/2$

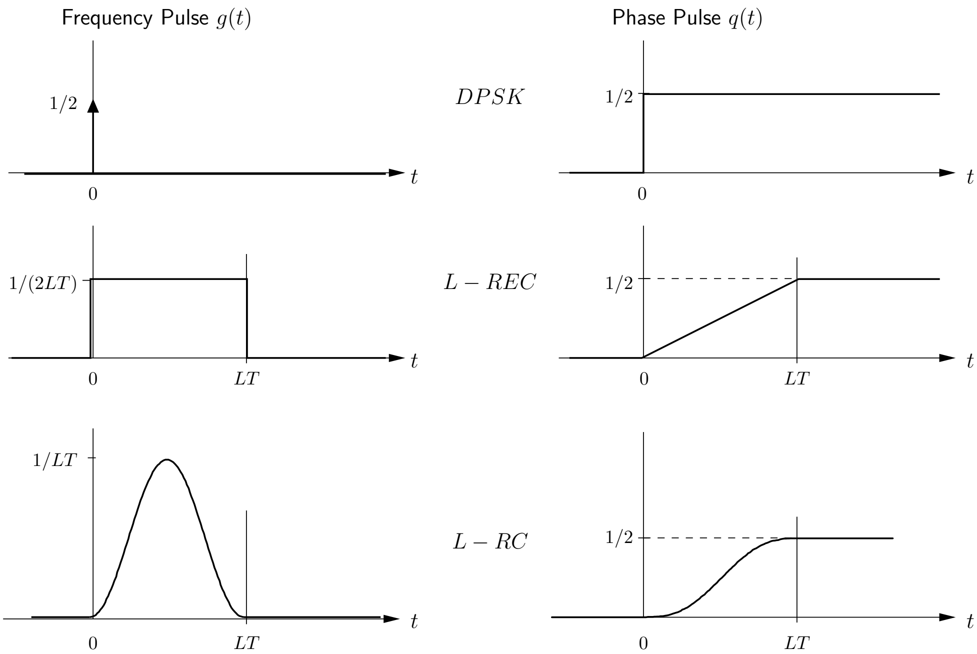
Features of CPM:

- Soft, continuous phase transitions result in the attenuation of spectral side lobes (bandwidth efficiency)
- Constant signal envelope: Efficient power amplification (power efficiency)
- Intersymbol interference in the phase signal!
- Inherent trellis coding (power efficiency)
- Efficient demodulation and synchronization using a "signal representation of spectral samples"
- Decoding using a (reduced-state) Viterbi algorithm

Block diagram for the generation of a CPM signal (historic interpretation!)



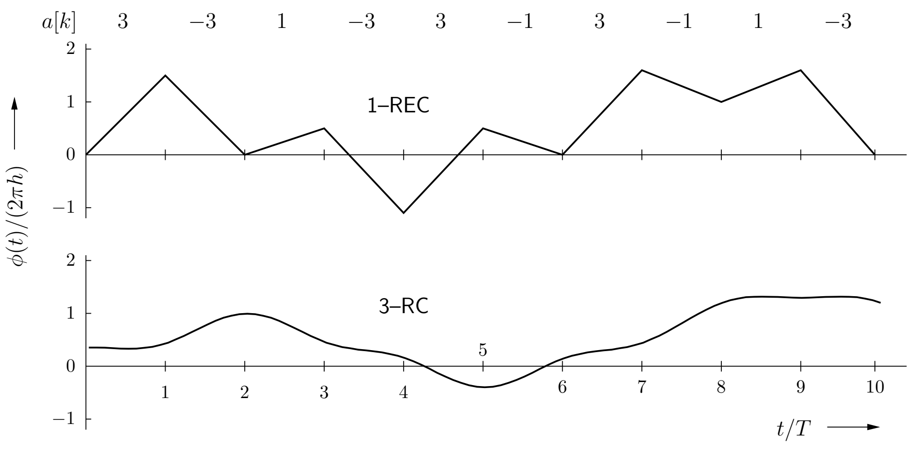
Frequency and phase pulses for various CPM schemes



REC: rectangular pulse

RC: raised cosine pulse

Example: *Phase trajectories for 1–REC– and 3–RC–CPM; M = 4*



Frequency pulses ($L \in \mathbb{N}$) usually applied:

L–REC $\rightarrow$ CPFSK (for $L > 1$: partial response)

$$g(t) = \begin{cases} \frac{1}{2LT} & \text{for } 0 \leq t < LT \\ 0 & \text{otherwise} \end{cases} = \frac{1}{2LT}\,\mathrm{rect}\left(\frac{t - L \cdot \frac{T}{2}}{LT}\right)$$

L–RC $\rightarrow$ Raised Cosine CPM

$$g(t) = \begin{cases} \frac{1}{LT}\left(\sin\left(\pi \frac{t}{LT}\right)\right)^2 & \text{for } 0 \leq t \leq LT \\ 0 & \text{otherwise} \end{cases}$$

6.2.1 Interpretation of CPM as a Trellis–Coded Signal

ECB signal with respect to a reference frequency ($\neq$ carrier frequency !)

$$f_0 = f_c - h\,\frac{M-1}{2T} \quad \Rightarrow \text{ frequency for symbol } a[k] \equiv -(M-1)$$

$$s(t) = \sqrt{\frac{E_s}{T}}\mathrm{e}^{+\mathrm{j}\psi(t)} \quad \text{with} \quad \psi(t) = \phi(t) + 2\pi h\,\frac{M-1}{2T}t$$

In the intervall $\nu T \leq t < (\nu+1)T$, we have

$$\psi(t) = \frac{2\pi}{q}\Theta_{\nu-L} + 2\pi h \sum_{k=\nu-L+1}^{\nu} b(a[k], t-kT) + \psi_0$$

with the **unipolar** phase signal

$$b(a,t) \overset{\text{def}}{=} \begin{cases} 0 & \text{for } \; t < 0 \\ aq(t) + \dfrac{M-1}{2LT}t & \text{for } \; 0 \leq t \leq LT \\ \dfrac{a}{2} + \dfrac{M-1}{2} & \text{for } \; t > LT \end{cases}$$

and the "phase state" contributions of all previous unipolar phase impulses which have reached their final

value $(a[k] + M - 1)/2)$

$$\Theta_{\nu-L} = \left(p \sum_{k=-\infty}^{\nu-L} \frac{(a[k]+M-1)}{2}\right) \bmod q \in \{0,1,2,\ldots,q-1\}$$

Definition: Unipolar phase increment $\quad b[k] \overset{\text{def}}{=} \frac{a[k]+M-1}{2} \quad \in \{0,1,2,\ldots,M-1\}$

Representation of the information $a[k]$ as a phase growth $2\pi h b[k]$ with respect to the reference phase $2\pi(f_\mathrm{c} - h\frac{M-1}{2T})t$.

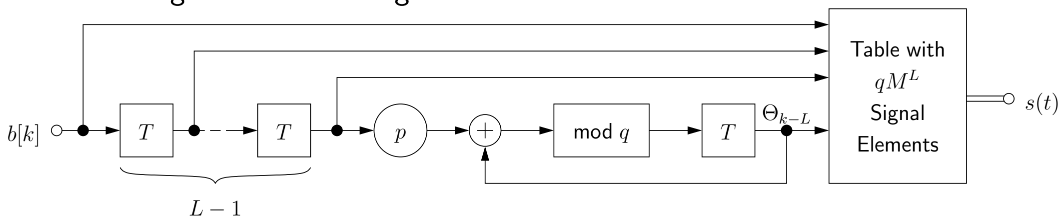
The following recursion holds: $\boxed{\Theta_{\nu-L+1} = (\Theta_{\nu-L} + p \cdot b[\nu-L+1]) \bmod q}$

Note: Information $b[k]$ can be interpreted as the derivative of the phase, i.e., as an "instantaneous frequency" like in analog FM.

Memory state of a trellis encoder: $S[\nu] = \Big(\Theta_{\nu-L},\, b[\nu-L+1], \ldots,\, b[\nu-1]\Big)$

Representation of CPM as a trellis–coded modulation scheme for M even and p odd (q even)

Generation of a CPM signal as an ECB signal:



Number of signal elements = all possible signal trajectories within the symbol interval $0 \le t < T$:
$q \cdot M^L$ signal elements

$$s_i(t) = \mathrm{e}^{\mathrm{j}\psi_i(t)} \quad i \in \{0, 1, \dots, qM^L - 1\}$$

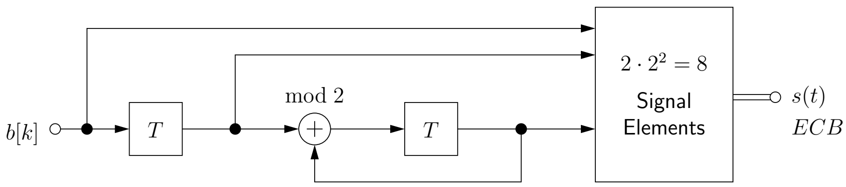
with

$$\psi_i(t) = 2\pi \left(h \sum_{\mu=-L+1}^{0} b\left(a^{(i)}, t - \mu T\right) + \frac{\Theta_i}{q} \right)$$

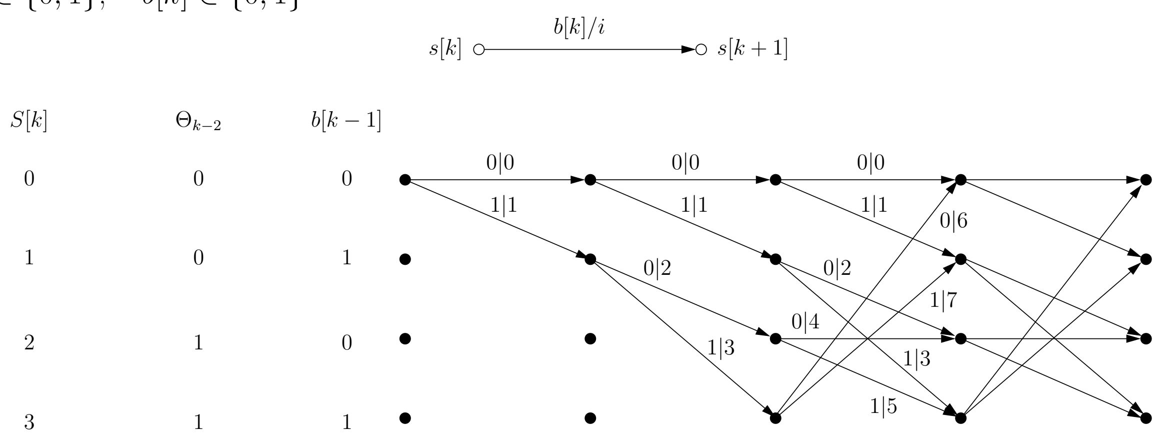
for all M^L sequences $\langle a[-L+1], \dots, a[0] \rangle$
q phase states Θ_{-L}

Number of states: $Z = M^{L-1} \cdot q$

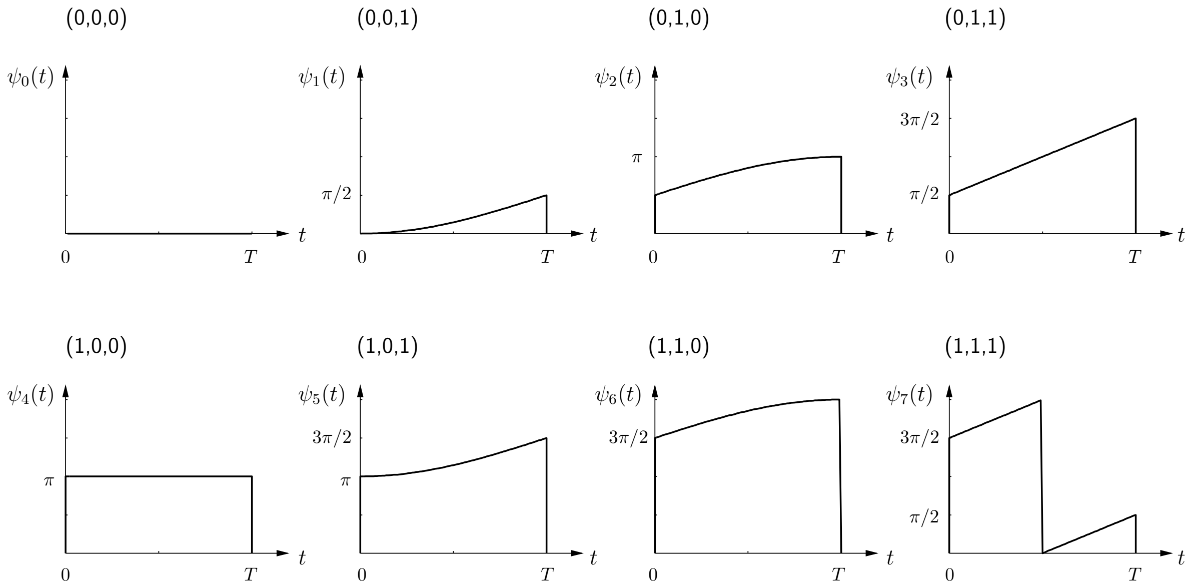
Example: *Trellis diagram for CPM $M = 2$, 2–RC, $h = 1/2$, $p = 1$, $q = 2$*



$\Theta_k \in \{0, 1\}; \quad b[k] \in \{0, 1\}$

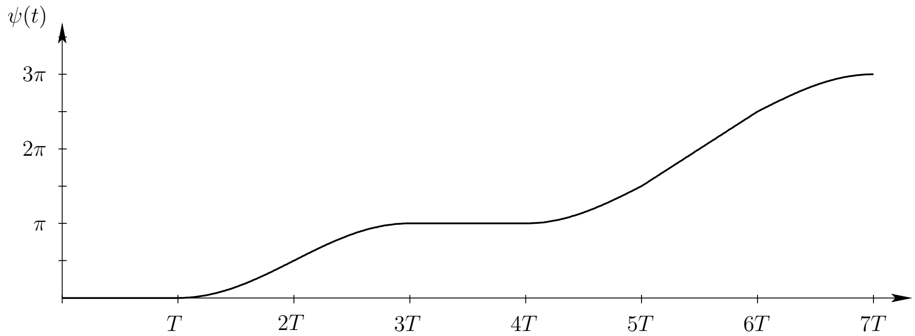


Phase functions of the 8 signal elements $(\Theta_{k-2}, b[k-1], b[k])$



Example

Data sequence $b[k]$	0	1	0	0	1	1	0
State $S[k]$	0	0	1	2	2	3	1
Signal number i	0	1	2	4	5	7	2
Phase $\psi(t)$							



6.2.2 Minimum Euclidean Distance of CPM Signals

$$d_{\min}^2 = \frac{E_s/T}{2E_b} \min_{\substack{\langle a[k]^{(1)}\rangle, \langle a[k]^{(2)}\rangle \\ a[0]^{(1)} \neq a[0]^{(2)}}} \int_{-\infty}^{+\infty} \left| \mathrm{e}^{+\mathrm{j}\psi(\langle a[k]^{(1)}\rangle, t)} - \mathrm{e}^{+\mathrm{j}\psi(\langle a[k]^{(2)}\rangle, t)} \right|^2 \mathrm{d}t =$$

$$= \frac{\mathrm{ld}(M)}{T} \min_{\substack{\langle \alpha[k]\rangle \\ \alpha[0]\neq 0}} \int_{-\infty}^{+\infty} \left(1 - \cos\left(2\pi h \sum_{\mu=-\infty}^{+\infty} \alpha[\mu] q(t-\mu T) \right) \right) \mathrm{d}t$$

with $\alpha[k] := a[k]^{(1)} - a[k]^{(2)}$ difference coefficients
$\alpha[k] \in \{0, \pm 2, \ldots, \pm(2M-2)\}$

Definition of the minimum over all difference sequences of sufficient length F, i.e., $\alpha[\mu] = 0$ for $\mu < 1$ and $\mu > F$.

Weak modulation index:

$$\exists \langle \alpha[k] \rangle \text{ with } \left(\pi h \sum_{\mu=1}^{F} \alpha[\mu] \right) \bmod 2\pi = 0$$

Within only a few steps, the difference symbols have no effect on subsequent signal components. Thus, there is no aggregation of energy of the difference of signals => low minimum squared Euclidean distance.

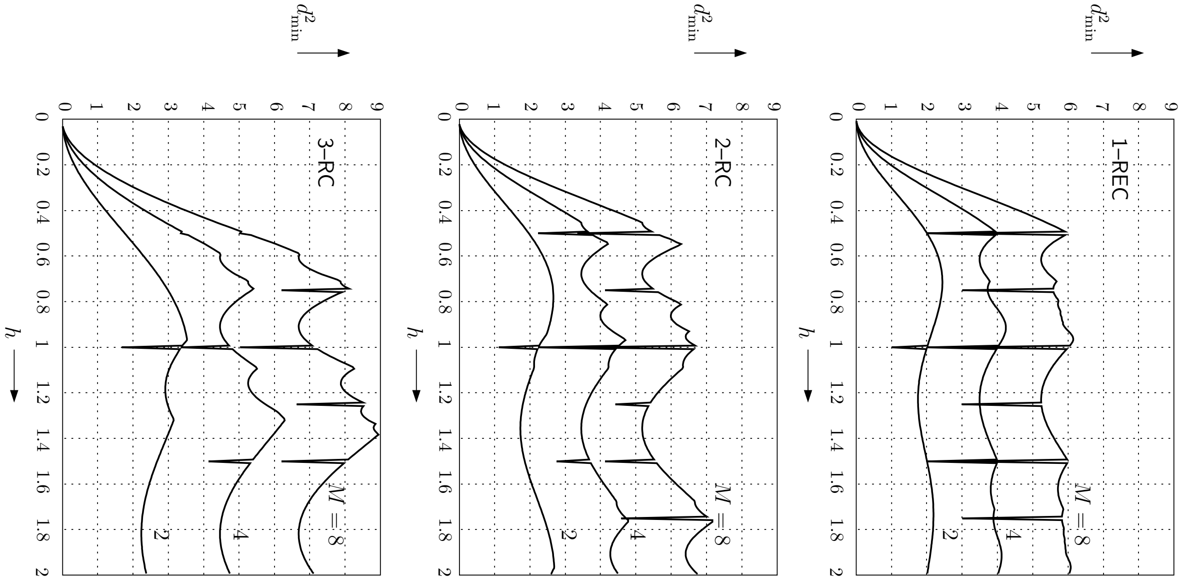
Examples

$M = 2,\ h \in \mathbb{N}$ (1–REC, $h = 1$: Sunde's FSK $d_{\min}^2 = 1$!)

$M = 4,\ h \in i/2,\ i \in \mathbb{N}$

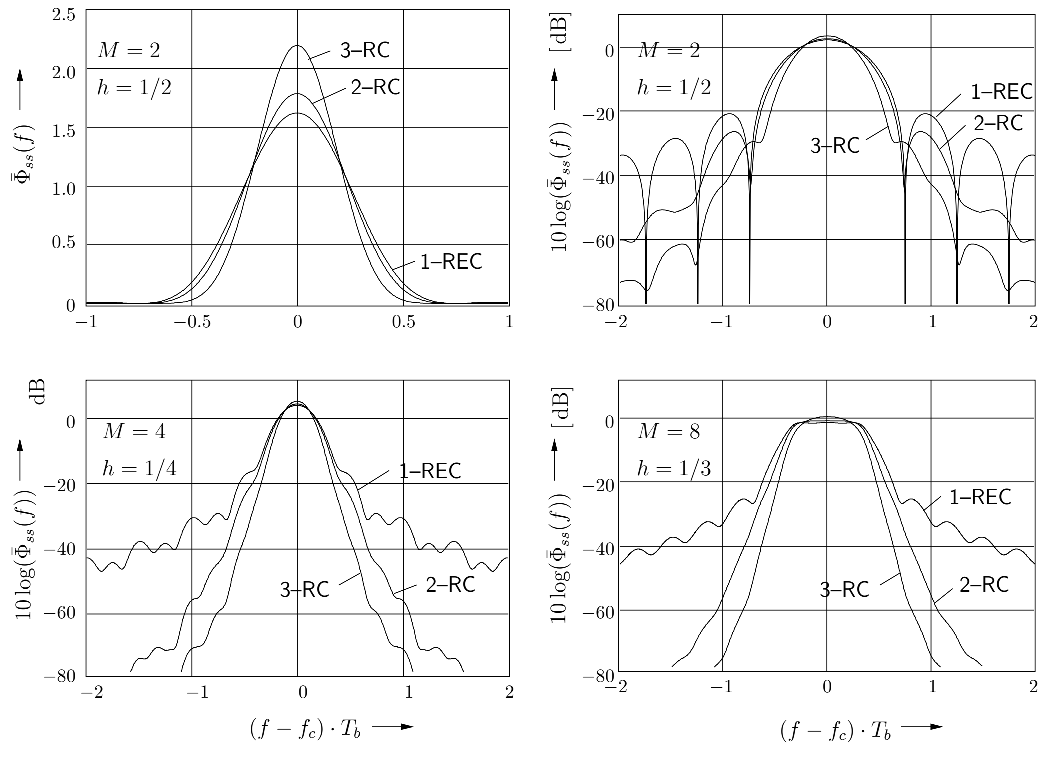
$F = 1$ is sufficient!

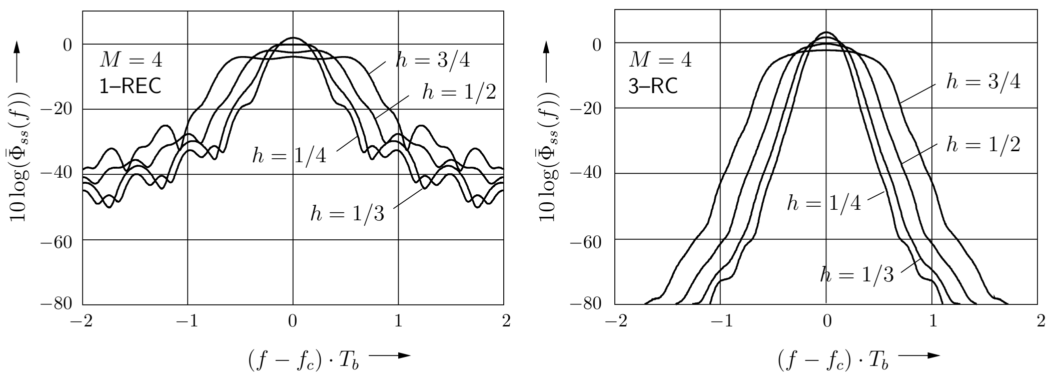
The Minimal Normalized Squared Euclidean Distance versus the modulation index h for $h = i/256, i \in \mathbb{N}$, frequency pulse 1–REC, 2–REC, 3–REC, 2–, 4–, 8–ary CPM



344

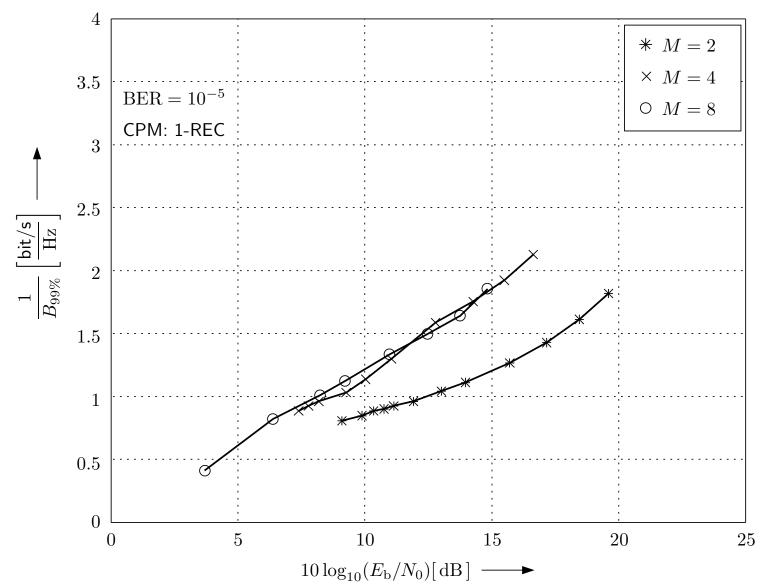
Average Power Spectral Densities of various CMP–Schemes



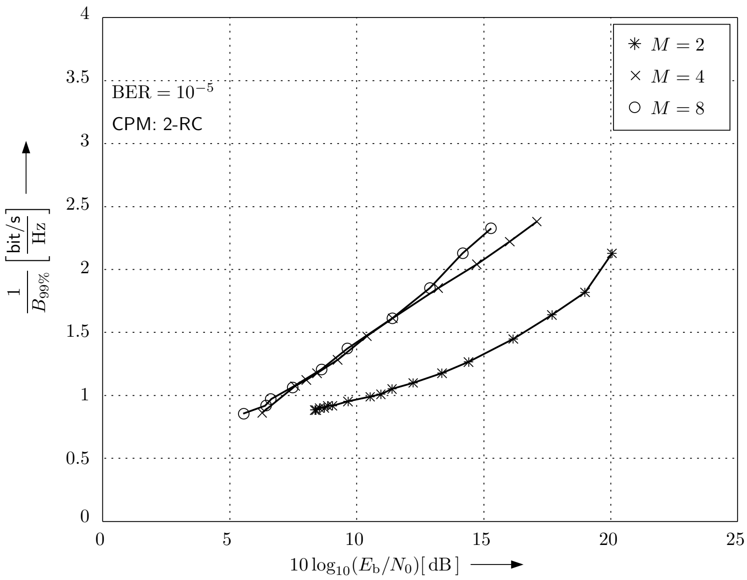


6.2.3 Power–Bandwidth Diagrams for CPM Schemes

Power–bandwidth diagram for 1–REC frequency pulse

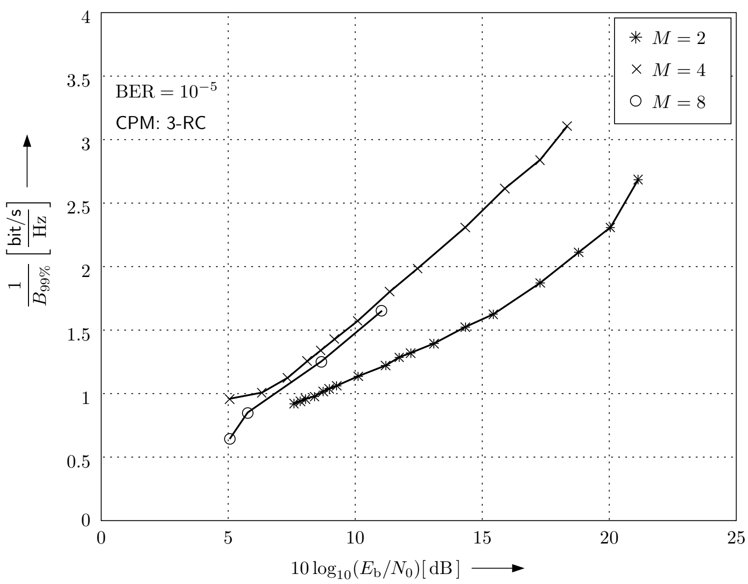


Power–bandwidth diagram for 2–RC frequency pulse



348

Power–bandwidth diagram for 3–RC frequency pulse



6.2.4 Coherent Receiver for CPM

- Interpretation of signal segment of duration T as signal element. There exist $M^L \cdot q$ signal elements.
  Example: $M = 4 \quad h = 1/2 \quad L = 3 \quad : 128$ signal elements

- Matched filter bank for all signal elements

- Receiver with reduced dimensionality (vector demodulator with incomplete basis)

  A number D of filters (in the ECB region) where

$$\lceil h(M-1)+1 \rceil \leq D \leq \lceil 1{,}11h(M-1)+2{,}22 \rceil$$

  is sufficient to ensure an accurate representation of the signal elements.

  Example: $M = 4; \quad L = 3; \quad h = 1/2: \quad 3 \leq D \leq 4$

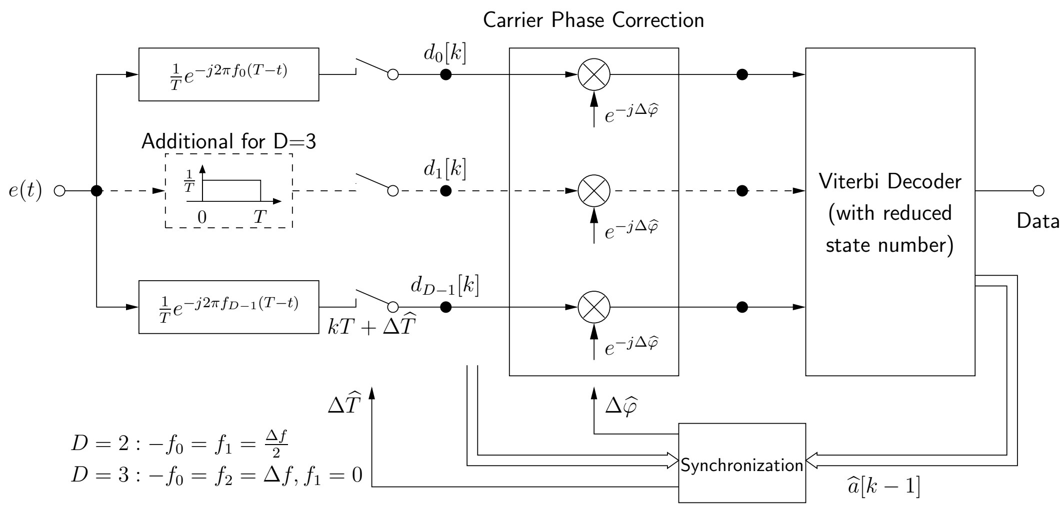
  If the following basis for the signal space is chosen:

$$\phi_{\mathrm{d}}(t) = \begin{cases} \mathrm{e}^{+\mathrm{j}2\pi f_{\mathrm{d}} t} & \text{for } t \in [0,\, T) \\ 0 & \text{for } t \notin [0,\, T) \end{cases}$$

$$\text{with} \qquad f_{\mathrm{d}} = \triangle f \, \frac{2d-(D-1)}{2} \quad \text{for } \; d = 0, \ldots, D-1,$$

  then a basis in the frequency domain at frequencies f_{d} is obtained through *spectral sampling* (special case $\triangle f = 1/T$: finite Fourier series). In general, however, frequency increments $\triangle f$ in the range

  of $0{,}5 \leq \triangle f \cdot T \leq 0{,}7$ are favorable. (An orthogonal basis will only result for $\triangle f = 1/T$ but this is normally not a good choice!). For receivers utilizing spectral sampling, carrier phase and symbol clock synchronization are particularly simple!

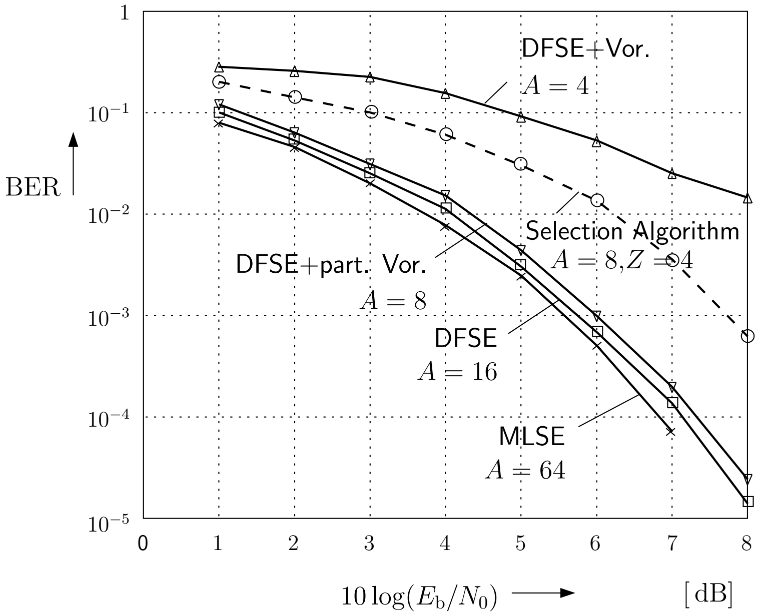


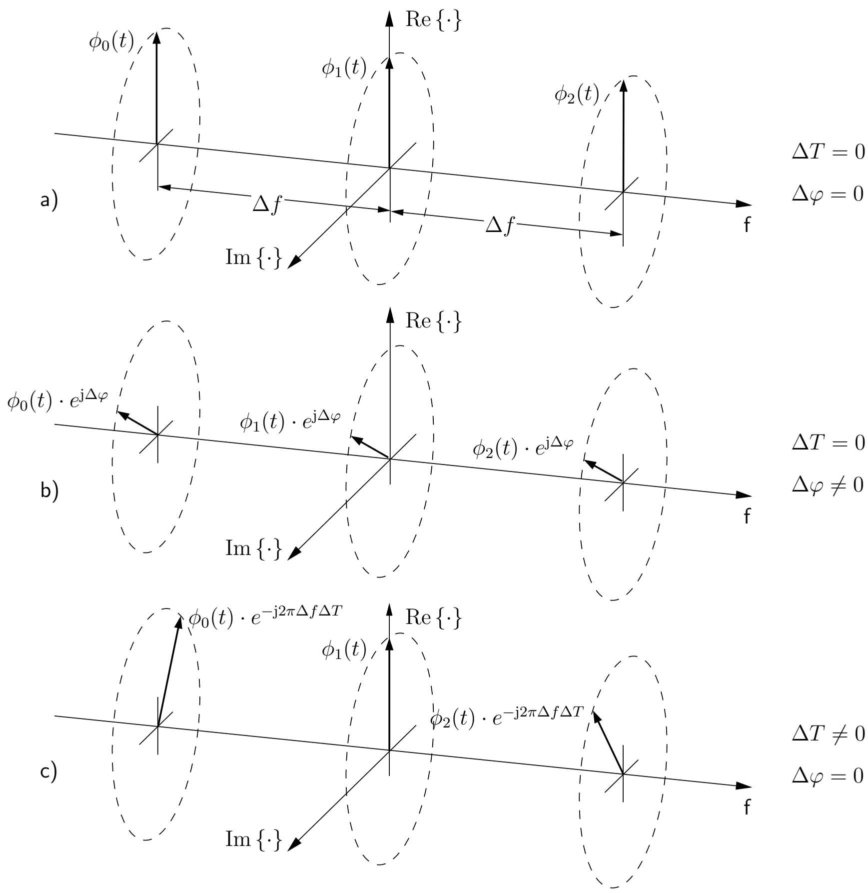
- Viterbi algorithm for $Z = M^{L-1} \cdot q$ states of the equivalent trellis encoder
  Complexity number of MLSE: $A = M^L \cdot q / \operatorname{ld}(M)$
  Example: $M = 4; \quad L = 3; \quad h = 1/2: \quad Z = 32; \quad A = 64$

- Application of Reduced–State Sequence–Estimation (RSSE) algorithms

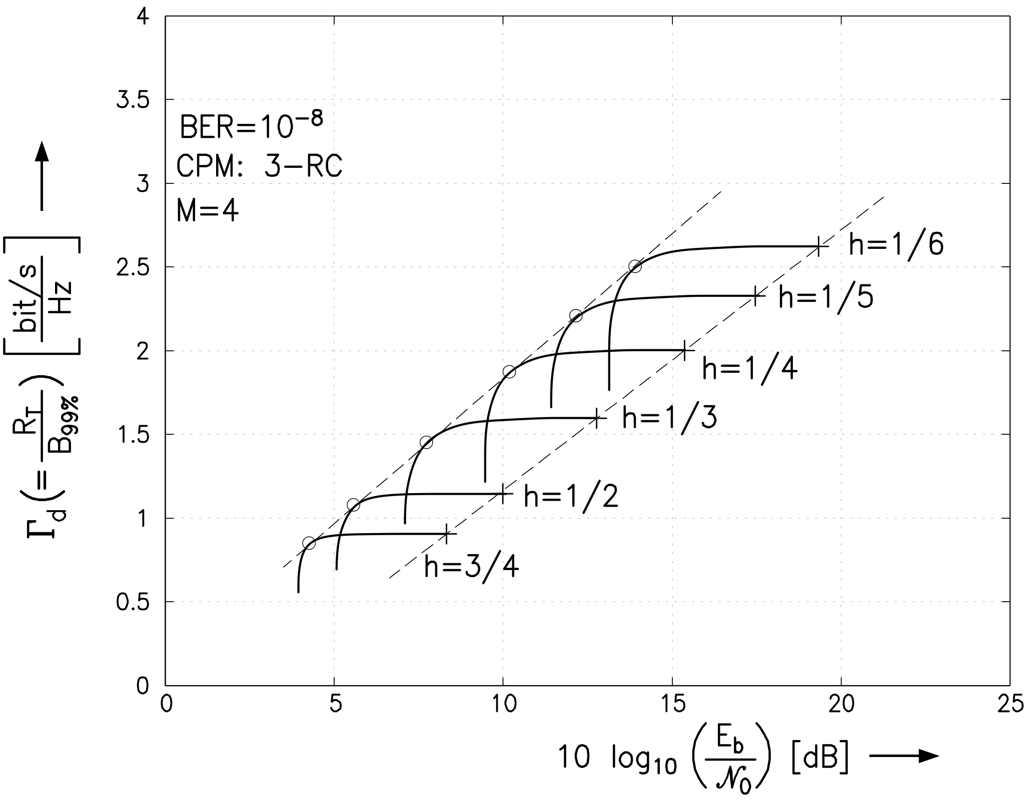
  Example: $M = 4; \quad L = 3; \quad h = 1/2:$

  Reduction up to $Z = 4$ states, $(A = 8)$ is possible with only 0.3 dB loss in performance.





Concatenation of CPM with $(255, k, t = \frac{255-k}{2})$ Reed Solomon codes over $\mathbb{F}_{256}$



7 Orthogonal Constellations

Orthogonal constellations and their variants are very interesting from both a theoretical (for $M \to \infty$ optimal scheme with respect to power efficiency) and a practical point of view. These constellations have found widespread use in mobile communications (e.g. for Code Division Multiple Access (CDMA) in UMTS) and in ultra-wideband communications (UWB).

Orthogonal constellation:
All M signal elements are mutually orthogonal and have equal energy E_g:

$$E_{il} = \int\limits_{-\infty}^{+\infty} s_i(t) s_l^*(t)\,\mathrm{d}t = \begin{cases} E_g & \text{for } i = l \\ 0 & \text{for } i \neq l \end{cases}$$

$$\forall i, l \in \{0, 1, \ldots,\ M-1\}$$

The signal elements $s_m(t)$ form are the basis functions $g_m(t)$ of an M–dimensional signal space.

Examples:

a) M–FSK with $h = i/2 \quad ; \quad i \in \mathbb{N}$
Relevant in practice $h = 1/2 \quad ; \quad$ ($h = 1$ for non–coherent demodulation)

b) Pulse Position Modulation (PPM)

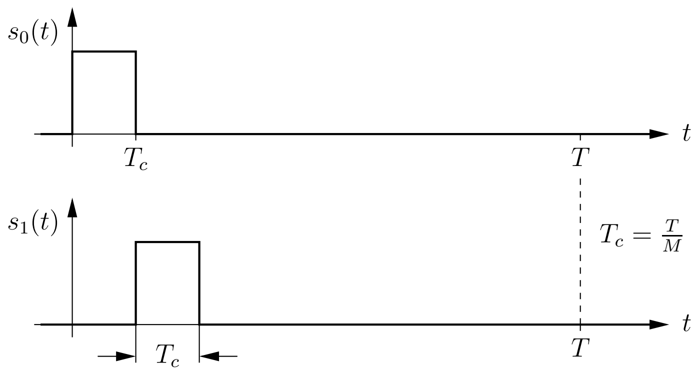
$$s_m(t) = g_c\left(t - (m-1)\frac{T}{M}\right) \quad ; \quad m \in \{0, 1, \ldots,\ M-1\}$$

Chip pulse $g_c(t)$ fulfills the temporal orthogonality condition with respect to the chip interval $T_c = T/M$:

$$\int\limits_{-\infty}^{+\infty} g_c\left(t + k \cdot \frac{T}{M}\right) g_c^*(t)\,\mathrm{d}t = \begin{cases} E_c & \text{for} \quad k = 0 \\ 0 & \text{for} \quad k \neq 0 \end{cases}$$

Here, we have $E_c = E_g$.

Application: ultra–wideband communications (UWB) with impulse radio and free space optical communication.



c) Orthogonal sequences (OSM)

$$s_m(t) = \sum_{i=0}^{N-1} b_{im} g_c \left(t - i\frac{T}{N} \right)$$

Chipsymbols: $b_{im} \in \{-1, +1\}$

N : Sequence length

Orthogonality:

$$\sum_{i=0}^{N-1} b_{im} b_{il} = \begin{cases} N & \text{for} \quad m = l \\ 0 & \text{for} \quad m \neq l \end{cases} \qquad m,l \in \{0, 1, \ldots, M-1\}$$

M orthogonal sequences only exist for $N \geq M$!

Application: CDMA in mobile communications, spread spectrum techniques (especially for military applications)

Example:

Walsh sequences

$n \times n$ Hadamard matrices $\mathbf{H}_n$ are defined recursively by

$$\mathbf{H}_{2n} = \begin{bmatrix} \mathbf{H}_n & \mathbf{H}_n \\ \mathbf{H}_n & -\mathbf{H}_n \end{bmatrix}$$

and

$$\mathbf{H}_1 = [1]$$

$$\Rightarrow \mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \text{orthogonality of rows:} \qquad 1 \cdot 1 + 1(-1) = 0$$

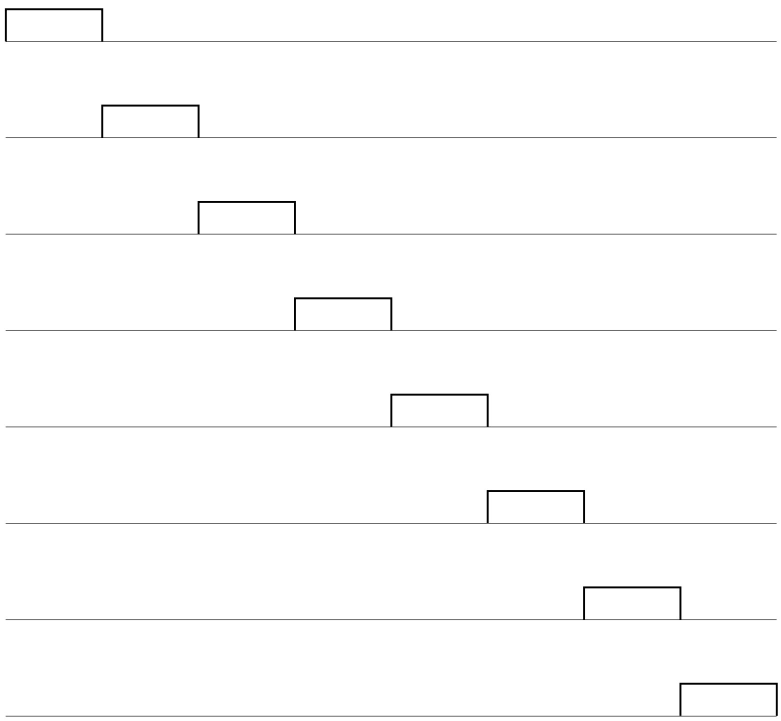
$$\mathbf{H}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

If the rows of $\mathbf{H}_n$ are mutually orthogonal, then the first n rows and the second set of n rows of $\mathbf{H}_{2n}$ are also mutually orthogonal. Likewise, the rows $i \in \{1, \ldots, n\}$ and $j \in \{n+1, \ldots, 2n\} \backslash \{n+i\}$ are mutually orthogonal, as the respective half of the summands yields zero. For rows i and $i+n$, the pairs of respective negative summands are obtained, guaranteeing orthogonality in this case as well.
Energy of signal elements: $E_g = N \cdot E_c$

Advantage of orthogonal sequences over pulse position modulation: Higher energy of signal element at smaller peak to average power ratio (lower crest factor of the signal).

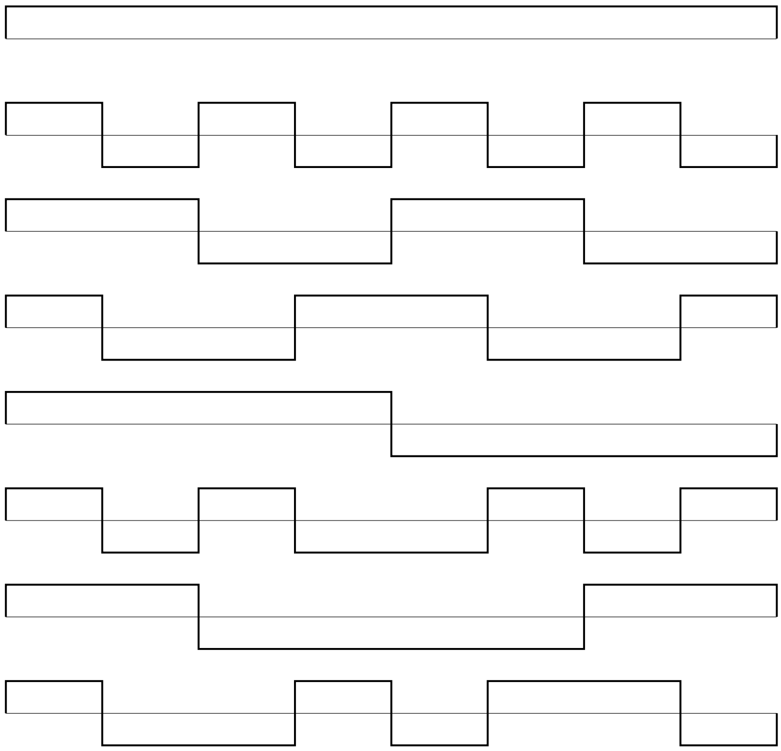
Application: CDMA (e.g. UMTS mobile communications)

Orthogonal functions for time multiplex transmission, pulse position modulation



364

8th Order Walsh function



Spectral efficiency

RF bandwidth for a set of orthogonal signals when both in–phase and quadrature component are used:

$$B_{\mathrm{RF}} = \frac{M}{2T}(1+\alpha) \quad \text{with}$$

α: roll–off factor of the chip impulse or, alternatively, to achieve a "soft envelope" of the signal elements, for e.g. for FSK with $h = 1/2$ or PPM / OSM ($\sqrt{\text{Nyquist}}$–pulse with roll–off factor α)

Spectral efficiency

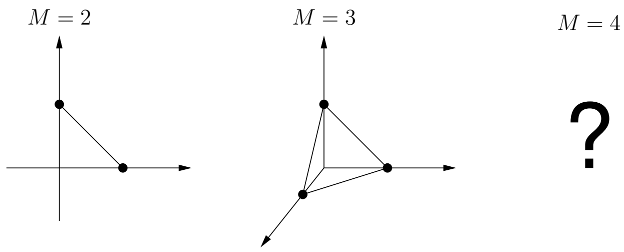
$$\Gamma_{\mathrm{d}} = \frac{R_{\mathrm{T}}}{B_{\mathrm{RF}}} = \frac{\log_2(M)/T}{M(1+\alpha)/(2T)} = \frac{\log_2(M)}{M} \cdot \frac{2}{1+\alpha} = \frac{R}{2^R} \cdot \frac{2}{1+\alpha}$$

Note: Decreasing with increasing number of levels M (this is in contrast to digital PAM!)

The rate of the modulation is $R = \log_2(M)$ in $\left[\frac{\text{bit}}{\text{modulation interval}}\right]$

Power efficiency

Euclidean distance



$$d_E^2 = 2 \qquad \forall M$$

for each pair of signal elements.

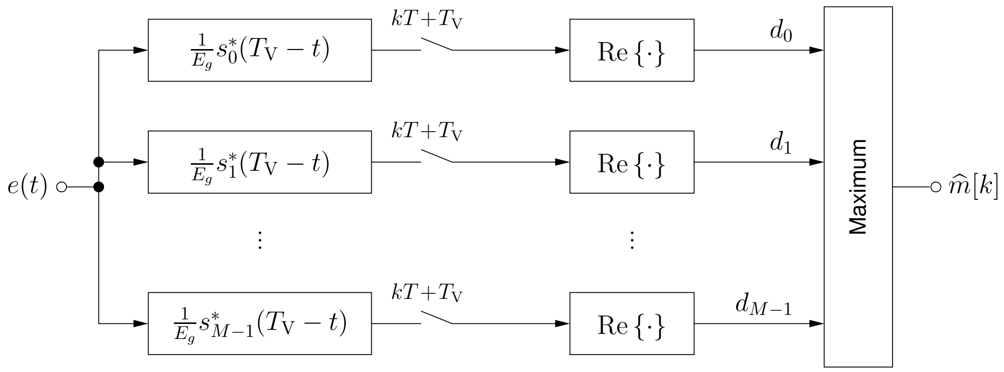
Normalized squared Euclidean distance

$$d_{\min}^2 = \frac{d_E^2 E_g}{2E_b} = \frac{2E_g}{2(E_g/\log_2(M))} = \log_2(M)$$

increases with step size M (in contrast to digital PAM!)

7.1 Coherent Demodulation

Optimal receiver: Bank of M matched filters $\equiv$ vector receiver



Symbol Error Probability:

Without loss of generality, assume $s_0(t)$ was transmitted:

$$
\begin{aligned}
d_0 &= 1 + n_0 \\
d_1 &= n_1 \\
d_2 &= n_2 \\
&\vdots \\
d_{M-1} &= n_{M-1}
\end{aligned}
$$

Noise n_m: real, mutually independent Gaussian random variables with variance $\sigma_I^2 = N_0/(2E_g)$

$$1 - \mathrm{SER} = \Pr(d_0 > d_1, d_0 > d_2, d_0 > d_3 \ldots)$$

Because of the mutual statistical independence of the n_m, we have

$$\Pr(d_0 > d_i) = \Pr(d_0 > d_j) \quad \forall \quad i,j \in \{1, 2, 3, \dots,\ M-1\}$$

$$\Rightarrow \mathrm{SER} = 1 - \int\limits_{-\infty}^{\infty} \Pr{}^{M-1}(d < d_0) f_{d_0}(d_0)\,\mathrm{d}d_0 \qquad \text{with}$$

$$\Pr(d < d_0) = 1 - Q\left(\frac{d_0}{\sigma_I}\right)$$

$$f_{d_0}(d_0) = \frac{1}{\sqrt{2\pi}\sigma_I} \mathrm{e}^{-\frac{(d_0-1)^2}{2\sigma_I^2}}$$

Substitution:

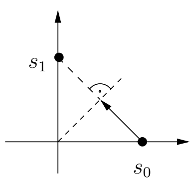
$$y = d_0/\sigma_I \quad ; \quad a = \frac{1}{\sigma_I} = \sqrt{E_g/(N_0/2)} = \sqrt{2\mathrm{ld}(M)\frac{E_b}{N_0}}$$

$$\boxed{\mathrm{SER} = 1 - \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} (1 - Q(y))^{M-1}\, \mathrm{e}^{-(y-a)^2/2} \mathrm{d}y}$$

Upper Bounds for the Symbol Error Probability:

a) Union Bound

$$\mathsf{SER} = \Pr\left(\bigcup_{i=1}^{M-1} (d_i > d_0)\right) \leq \sum_{i=1}^{M-1} \Pr(d_i > d_0) = (M-1) \cdot \Pr(d_1 > d_0)$$



$$\Pr(d_1 > d_0) = Q\left(\frac{\sqrt{2}/2}{\sigma_I}\right) = Q\left(\sqrt{\frac{E_g}{N_0}}\right) = Q\left(\sqrt{\mathrm{ld}(M)\frac{E_b}{N_0}}\right)$$

$$\mathrm{SER} \leq (M-1) Q\left(\sqrt{\mathrm{ld}(M)\frac{E_b}{N_0}}\right)$$

Chernoff–Bound for the Gaussian $Q-$function

$$Q(x) \leq \frac{1}{2}\,\mathrm{e}^{-x^2/2} \qquad \text{for } x \geq 0$$

$$\mathrm{SER} < \frac{M}{2}\mathrm{e}^{-\mathrm{ld}(M)E_b/(2N_0)} = \frac{1}{2}\mathrm{e}^{\mathrm{ld}(M)-\mathrm{ld}(M)E_b/(2N_0)}$$

$$\boxed{\mathrm{SER} < \tfrac{1}{2}\;\mathrm{e}^{-\mathrm{ld}(M)(E_b/N_0-2\ln(2))/2}}$$

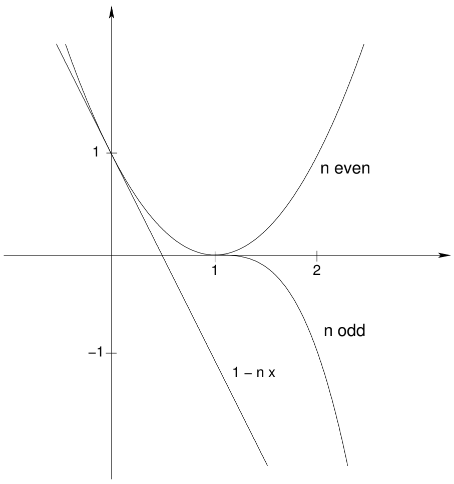
b) Tighter Approximation for Low SNR

Exact equation:

$$\mathrm{SER} = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}(1-(1-Q(y))^{M-1})\mathrm{e}^{-(y-a)^2/2}\mathrm{d}y$$

372

Bounding of the expression in parentheses
$(1-x)^n \geq 1-nx$ for $0 \leq x \leq 1$; $n \in \mathbb{N}$,
see sketch below:



$$\alpha)\;\; 1-(1-Q(y))^{M-1} \;\leq\; 1-(1-(M-1)Q(y)) \;<\; MQ(y) \;<$$

$$< M\,\mathrm{e}^{-y^2/2} \;\leftarrow \qquad \text{for } y \geq y_0 > 0 \text{ with } M\cdot\mathrm{e}^{-y_0^2/2} \overset{!}{=} 1 \text{ alternatively } y_0 = \sqrt{2\ln(M)}$$

$$\beta)\;\; 1-(1-Q(y))^{M-1} \;\leq\; 1 \quad \text{(probability)}$$

Summary:

$$1-(1-Q(y))^{M-1} \leq \begin{cases} 1 & \text{for} \quad y \leq y_0 \\ M\mathrm{e}^{-y^2/2} & \text{for} \quad y > y_0 > 0 \end{cases}$$

Expansion of the integral into the output equation

$$\mathrm{SER} = \int\limits_{-\infty}^{+\infty} \cdots \mathrm{d}y = \int\limits_{-\infty}^{y_0} \cdots \mathrm{d}y + \int\limits_{y_0}^{\infty} \cdots \mathrm{d}y =: I_1 + I_2$$

with

1. Integral:

$$I_1 \leq \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{y_0} 1 \cdot \mathrm{e}^{-(y-a)^2/2} \mathrm{d}y = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{y_0-a} \mathrm{e}^{-z^2/2} \mathrm{d}z = 1 - Q(y_0 - a)$$

Substitution: $z = y - a \quad ; \quad Q(x) = 1 - Q(-x) \leq \frac{1}{2}\,\mathrm{e}^{-x^2/2}$ for $x \geq 0$

$$I_1 \leq Q(a - y_0) \leq \frac{1}{2}\,\mathrm{e}^{-(a-y_0)^2/2} \text{ for } y_0 \leq a$$

2. Integral:

$$I_2 \leq M \cdot \frac{1}{\sqrt{2\pi}} \int\limits_{y_0}^{\infty} \mathrm{e}^{-\frac{y^2+(y-a)^2}{2}} \mathrm{d}y$$

quadratic expansion of the exponent

$$\begin{aligned} y^2 + (y-a)^2 &= 2\left(y^2 - ay + \frac{a^2}{4} + \frac{a^2}{4}\right) \\ &= 2\left(\left(y - \frac{a}{2}\right)^2 + \frac{a^2}{4}\right) \end{aligned}$$

Thus,

$$I_2 \leq M \cdot \mathrm{e}^{-a^2/4} \cdot \frac{1}{\sqrt{2\pi}} \int\limits_{y_0}^{\infty} \mathrm{e}^{-(y-a/2)^2} \mathrm{d}y$$

Subst.: $z = \sqrt{2}(y - a/2), \quad \mathrm{d}y = \mathrm{d}z/\sqrt{2}$

$$I_2 \leq \frac{M}{\sqrt{2}}\,\mathrm{e}^{-a^2/4} \frac{1}{\sqrt{2\pi}} \int\limits_{\sqrt{2}(y_0-a/2)}^{\infty} \mathrm{e}^{-z^2/2} \mathrm{d}z = \frac{M}{\sqrt{2}} \mathrm{e}^{-a^2/4} Q\left(\sqrt{2}\left(y_0 - \frac{a}{2}\right)\right)$$

$$I_2 \leq \frac{M}{2\sqrt{2}}\,\mathrm{e}^{-a^2/4} \cdot \mathrm{e}^{-(y_0-\frac{a}{2})^2} \text{ for } y_0 \geq \frac{a}{2}$$

Using $M = \mathrm{e}^{+y_0^2/2}$ in the previous equation:

$$I_2 \leq \frac{1}{2\sqrt{2}}\,\mathrm{e}^{-\frac{a^2-2y_0^2+4y_0^2-4ay_0+a^2}{4}} = \frac{1}{2\sqrt{2}}\,\mathrm{e}^{-(a-y_0)^2/2}$$

Altogether, we obtain

$$\mathrm{SER} < \mathrm{e}^{-(a-y_0)^2/2} \text{ for } \frac{a}{2} \leq y_0 \leq a \text{ with}$$

$$a = \sqrt{2\mathrm{ld}(M)E_b/N_0} \quad ; \quad y_0 = \sqrt{2\mathrm{ld}(M)\cdot\ln(2)}$$

$$\frac{2\mathrm{ld}(M)E_b/N_0}{4} \leq 2\mathrm{ld}(M)\cdot\ln(2) \leq 2\mathrm{ld}(M)E_b/N_0$$

For $\ln(2) \leq E_b/N_0 \leq 4\ln(2)$, we have

$$\mathrm{SER} < \mathrm{e}^{-\mathrm{ld}(M)\left(\sqrt{E_b/N_0}-\sqrt{\ln(2)}\right)^2}$$

376

Rate of modulation $R = \mathrm{ld}(M)$

Of the bounds a) and b), the tighter one is chosen, yielding

$$\boxed{\mathrm{SER} < 2^{-RE_{\mathrm{ortho}}(E_b/N_0)}}$$

with the

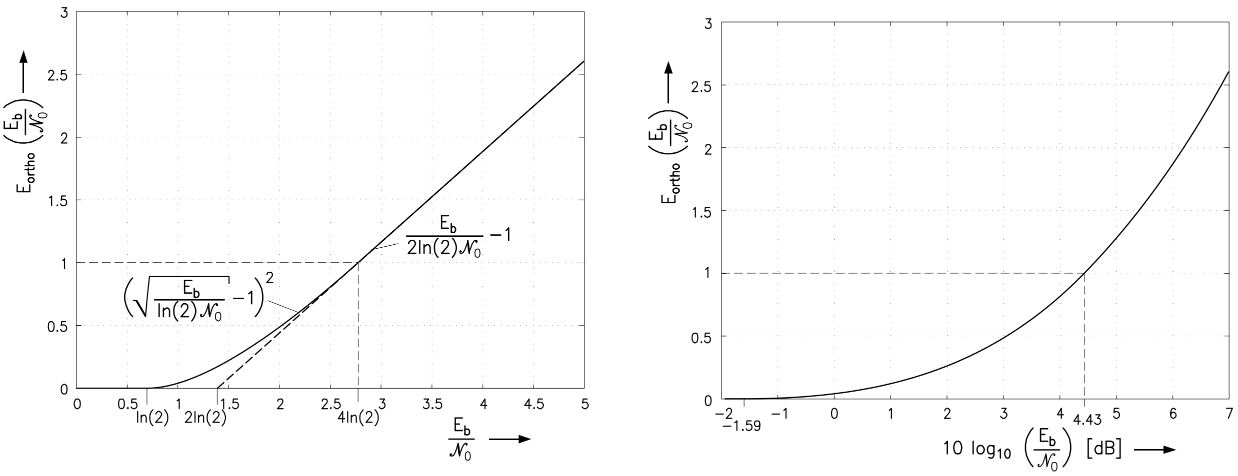
Error exponent for orthogonal constellations:

$$E_{\mathrm{ortho}}(E_b/N_0) = \begin{cases} 0 & \text{for } E_b/N_0 < \ln(2) \\ \left(\sqrt{E_b/(\ln(2)N_0)}-1\right)^2 & \text{for } \ln(2) \leq E_b/N_0 \leq 4\ln(2) \\ E_b/(2\ln(2)N_0)-1 & \text{for } E_b/N_0 \geq 4\ln(2) \end{cases}$$

For $E_b/N_0 > \ln(2)$ the error exponent $E_{\text{ortho}}(E_b/N_0)$ is positive: Therefore, we can conclude the following:

By means of construction of sufficiently long blocks of R binary source symbols and subsequent transmission of each block by one signal element of an orthogonal constellation with 2^R elements, an arbitrarily small symbol error probability may be achieved by increasing R as long as the following holds for the SNR:

$$\boxed{E_b/N_0 > \ln(2) \quad \text{or alternatively} \quad 10\lg(E_b/N_0) > -1.59 \text{ dB}}$$



7.2 Comparisons with Bounds from Information Theory

The capacity of an AWGN channel according to Shannon's famous formula:

$$C_{\text{T}} = B_{\text{RF}} \,\text{ld}\left(1 + \frac{S}{N}\right) \left[\frac{\text{bit}}{s}\right] \qquad \text{with}$$

B_{RF} : Signal bandwidth (one–sided)

$S = E_b/T_b$: Signal power

$N = N_0 \cdot B_{\text{RF}}$: Noise power

R_{T} : Transmission data rate (information stream)
$T_{\text{b}} = 1/R_{\text{T}}$: (Equivalent) transmission time per bit of information

For an ideal digital transmission scheme, we have

$$R_{\mathrm{T}} = C_{\mathrm{T}}$$

with bandwidth efficiency $\Gamma_{\mathrm{d}} = R_{\mathrm{T}}/B_{\mathrm{RF}}\ [\mathrm{bit/s/Hz}]$

The following compromise between power and bandwidth efficiency holds for the above bound from information theory:

$$R_{\mathrm{T}} = B_{\mathrm{RF}}\,\mathrm{ld}\left(1 + \frac{E_b}{N_0}\frac{1}{B_{\mathrm{RF}}T_{\mathrm{b}}}\right)$$

$$\boxed{\begin{aligned} \Gamma_{\mathrm{d}} &= \mathrm{ld}\left(1 + \frac{E_b}{N_0}\cdot\Gamma_{\mathrm{d}}\right) \\ \frac{E_{\mathrm{b}}}{N_0} &= \tfrac{1}{\Gamma_{\mathrm{d}}}\left(2^{\Gamma_{\mathrm{d}}} - 1\right) \end{aligned}}$$

This relationship marks a performance bound for all digital communications schemes in the *power bandwidth diagram*.

380

Even for infinite bandwidth $\Gamma_{\mathrm{d}} \to 0$, an SNR of

$$\lim_{x\to 0}\frac{1}{x}(2^x - 1) = \lim_{x\to 0}\frac{\mathrm{e}^{\ln(2)x}\cdot\ln(2)}{1} = \ln(2)$$

is required. In other words,

$$\boxed{10\mathrm{lg}(E_b/N_0) > -1.59\ \mathrm{dB}}$$

⇒ **The channel capacity can be achieved in principle, with orthogonal constellations for $M \to \infty$ or $R \to \infty$.**

However, this means of reaching the channel capacity is prohibitively bandwidth *inefficient*, as for coherent demodulation we obtain for orthogonal constellations (c.f. p.365 for $\alpha = 0$):

$$\Gamma_{\mathrm{d}} \leq \frac{2\mathrm{ld}(M)}{M} = \frac{R}{2^{R-1}}.$$

For the power–bandwidth diagram of orthogonal constellation for a certain tolerated SER, we require:

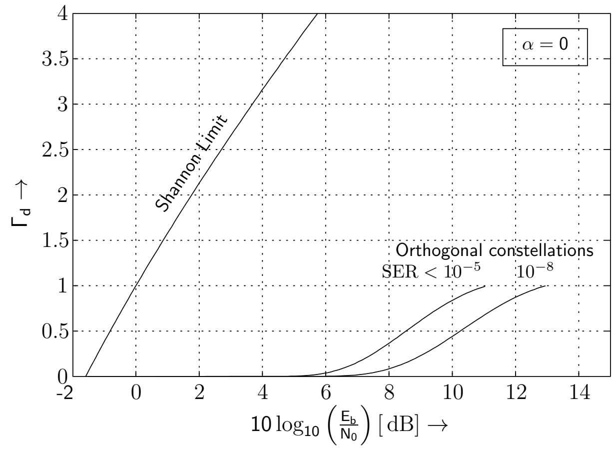
a) $\Gamma_{\mathrm{d}} \leq \dfrac{R}{2^{R-1}}$

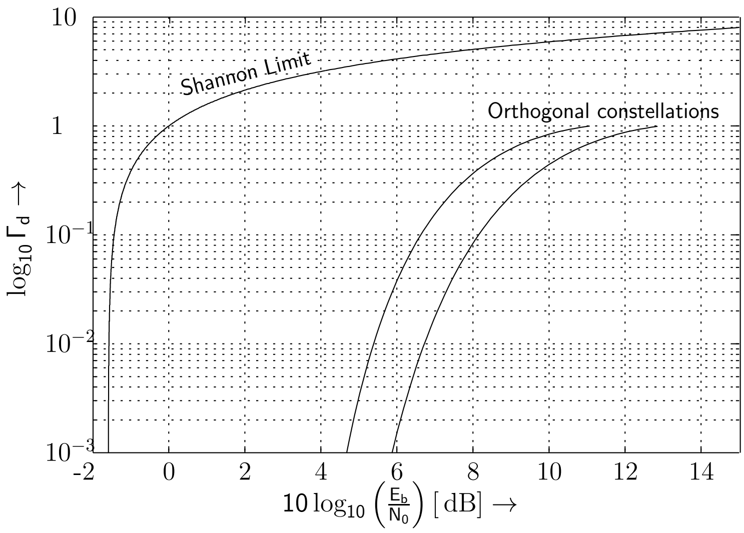
b) $E_{\mathrm{ortho}}\left(\dfrac{E_b}{N_0}\right) = -\mathrm{ld}(\mathrm{SER})/R$

For $R < -\mathrm{ld}(\mathrm{SER}) : \dfrac{E_b}{N_0} = 2\ln(2)(1 - \mathrm{ld}(\mathrm{SER})/R)$

For $R > -\mathrm{ld}(\mathrm{SER}) : \dfrac{E_b}{N_0} = \ln(2)\left(1 + \sqrt{-\mathrm{ld}(\mathrm{SER})/R}\right)^2$

Note: For $R \to \infty : \dfrac{E_b}{N_0} = \ln(2)$





Observation:
Orthogonal constellations reach the optimal power efficiency in a **very bandwidth inefficient way.** In this respect, transmission schemes using orthogonal modulation are extremely inefficient for high number of levels M !

7.3 Non–coherent Demodulation

In contrast to coherent transmission, no separate use of the in–phase and quadrature components of a signal is possible for non–coherent demodulation. For this reason, the following holds for bandwidth non–coherent demodulation of orthogonal constellations:

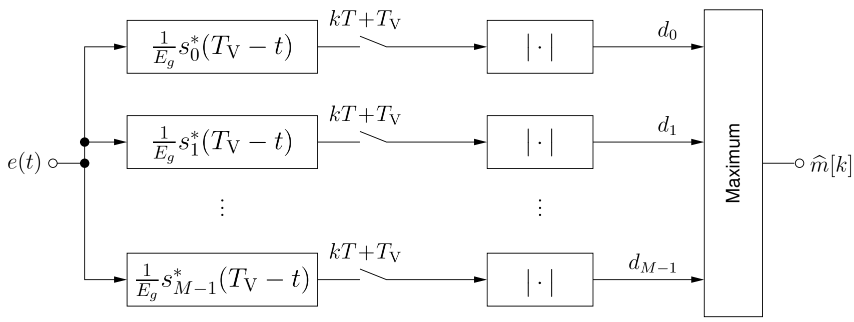
$$B_{\mathrm{RF}} = \frac{M}{T}\,(1+\alpha) \qquad \alpha: \text{ roll–off factor}$$

e.g. M-FSK with modulation index $h = 1$

$$\text{bandwidth efficiency } \Gamma_{\mathrm{d}} \leq \frac{R}{2^R}$$

$R = \mathrm{ld}(M)$

Optimal receiver



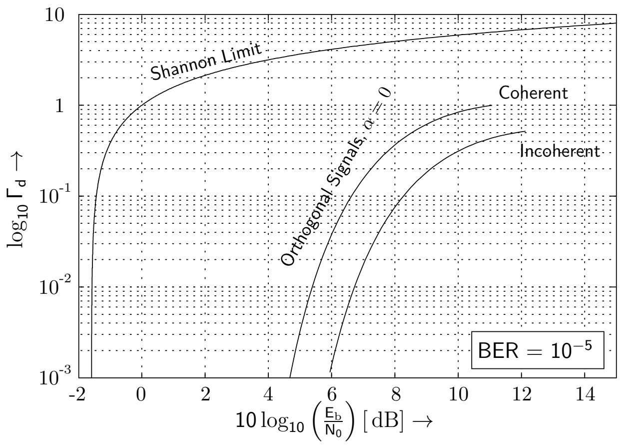
For non–coherent demodulation, an error exponent

$$E'_{\mathrm{ortho}}(E_b/N_0)$$

may be defined which is not significantly smaller than in the case of coherent demodulation. Especially for $R \gg 1$, the error exponents of for non–coherently and coherently demodulated orthogonal constellations are almost equal:

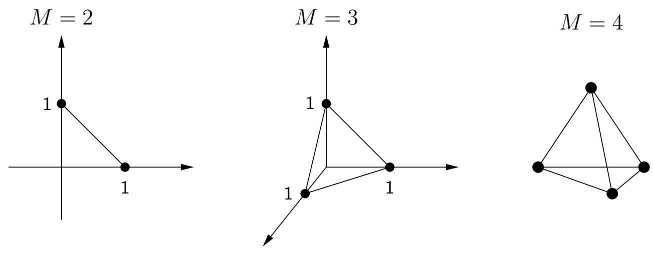
$$E'_{\mathrm{ortho}}(E_b/N_0) > 0 \qquad \text{for } E_b/N_0 > \ln(2)$$

Power–Bandwidth Diagram for Orthogonal Constellations (BER $= 10^{-5}$)



7.4 Simplex Constellation

If all of the end points M unit vectors of an M–dimensional space in Cartesian coordinates are connected, an $(M-1)$–dimensional figure is created which is known as an $(M-1)$–dimensional *simplex*.



Line

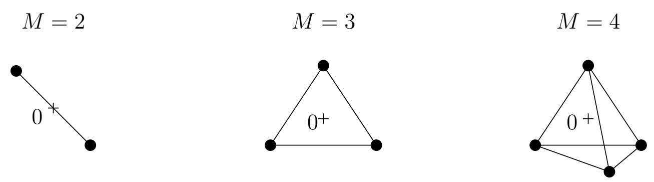
Equilateral Triangle

Tetrahedon

All these points possess a mutual distance of $\sqrt{2}$ from one another.
The M signal points $s_{om}(t)$ of an orthogonal constellation may thus also be described as those of an $(M-1)$–dimensional simplex in the M–dimensional signal space. Through a translation with respect to the negative centroid

$$-\frac{1}{M}(1, 1, \ldots, 1)$$

one obtains the *regular* simplex constellation (i.e. with the centroid located at the origin):



If the arithmetic mean

$$\bar{s}(t) = \frac{1}{M} \sum_{m=0}^{M-1} s_{om}(t)$$

of the M signal elements $s_{om}(t)$, $m \in \{0, 1, \ldots, M-1\}$ of an orthogonal constellation is subtracted, the result is a (regular) *simplex constellation* in an $(M-1)$–dimensional signal space:

$$s_m(t) = s_{om}(t) - \bar{s}(t)$$

Defining E_{go} as the energy of the signal elements in the corresponding orthogonal constellation, we obtain for the simplex constellation:

$$E_g = \int_{-\infty}^{+\infty} |s_m(t)|^2 \mathrm{d}t = \int_{-\infty}^{+\infty} |s_{om}(t)|^2 \mathrm{d}t - \int_{-\infty}^{+\infty} |\bar{s}(t)|^2 \mathrm{d}t \quad \text{(Steiner's Theorem)}$$

The orthogonality of the signal elements $s_{om}(t)$ yields the following integral

$$\int_{-\infty}^{+\infty} |\bar{s}(t)|^2 \mathrm{d}t = \frac{1}{M^2} \sum_{m=1}^{M} \int_{-\infty}^{+\infty} |s_{om}(t)|^2 \mathrm{d}t = E_{go}/M$$

$$\Rightarrow \qquad E_s = E_{go}(1 - 1/M) = E_{go}\frac{M-1}{M} \quad \forall m \in \{0, 1, \ldots,\ M-1\}$$

For orthogonal as well as simplex constellations, we have,

$$\mathrm{d}_E^2 = 2$$

Distances are translation invariant!

Simplex constellation: $\quad d_{\min}^2 = \dfrac{2E_{go}}{2\frac{M-1}{M}E_{go}\frac{1}{\mathrm{ld}(M)}} = \dfrac{M}{M-1}\,\mathrm{ld}(M)$

Comparison

M	Orthogonal:	$d_{\min}^2$	Simplex:	$d_{\min}^2$	Gain
2	e.g. 2FSK	1	antipodal (2PSK, 2ASK)	2	3 dB
3	e.g. 3FSK	$\mathrm{ld}(3) = 1.58$	3PSK	$\frac{3}{2} \cdot \mathrm{ld}(3) = 2.38$	1.77 dB
4	e.g. 4FSK	2	e.g. PPM, regular	2.66	1.25 dB

For $M > 4$, the difference diminishes.

In contrast to orthogonal constellation, for simplex constellations, the energy required for the transmission of an information–less average signal is saved: No discrete lines in power spectral density. Simplex constellations have (probably) the highest power efficiency among all constellations with the same number of signal points.

The expression for the error probability of orthogonal constellations also hold for simplex constellation if E_b is substituted by $\frac{M}{M-1}E_b$.

7.5 Biorthogonal Constellations

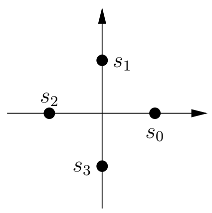
For each signal element $s_m(t)$ of an orthogonal constellation, the negative function

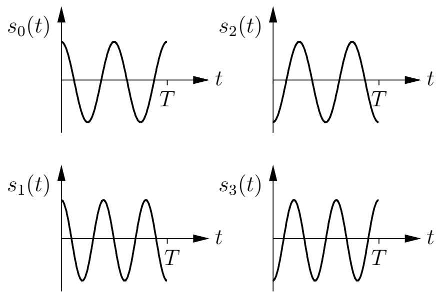
$$-s_m(t) =: s_{m+M/2}(t) \quad m \in \{0, 1, \ldots, M/2\}$$

is also included in the signal set (i.e., M is even). Biorthogonal constellations have therefore zero mean. All signal elements possess equal signal energy E_g.

Example:

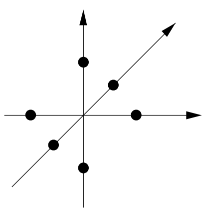
$M = 4$, $D = 2$





Example:

$M = 6$, $D = 3$



Min. Euclidean distance:

Only the distances $\sqrt{2}$ and 2 occur.

For $M \geq 4$, we have

$$\left.\begin{aligned} d^2_{E,\min} &= 2 \\ d^2_{\min} &= \mathrm{ld}(M) \end{aligned}\right\} \text{as with orthogonal constellations}$$

Special case: $M = 2 \quad d^2_{E,\min} = 4 \quad ; \quad d^2_{\min} = 2$

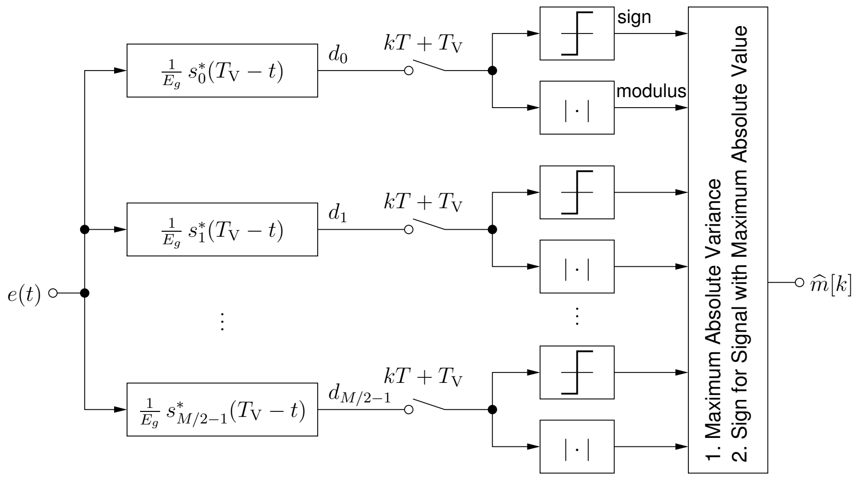
Bandwidth efficiency:

Example: Pulse position modulation $\quad B_{\mathrm{RF}} \approx \dfrac{M}{2}\dfrac{1}{2T}\,(1+\alpha)$

$$\Gamma_{\mathrm{d}} \approx \frac{4\mathrm{ld}(M)}{M(1+\alpha)} = \frac{4R}{2^R} \cdot \frac{1}{1+\alpha} = \frac{R}{2^{R-2}} \cdot \frac{1}{1+\alpha}$$

For $M \gg 1$, the biorthogonal constellation is by a factor of 2 more bandwidth efficient than the standard orthogonal constellation.

Optimal coherent receiver (only coherent reception is possible!): Bank of $M/2$ matched filters



Symbol error probability:

Without loss of generality, assume: $s_0(t)$ is sent

$$\mathrm{SER} = 1 - \Pr(d_0 > 0, |d_1| < |d_0|, |d_2| < |d_0|, \ldots, |d_{M/2-1}| < |d_0|)$$

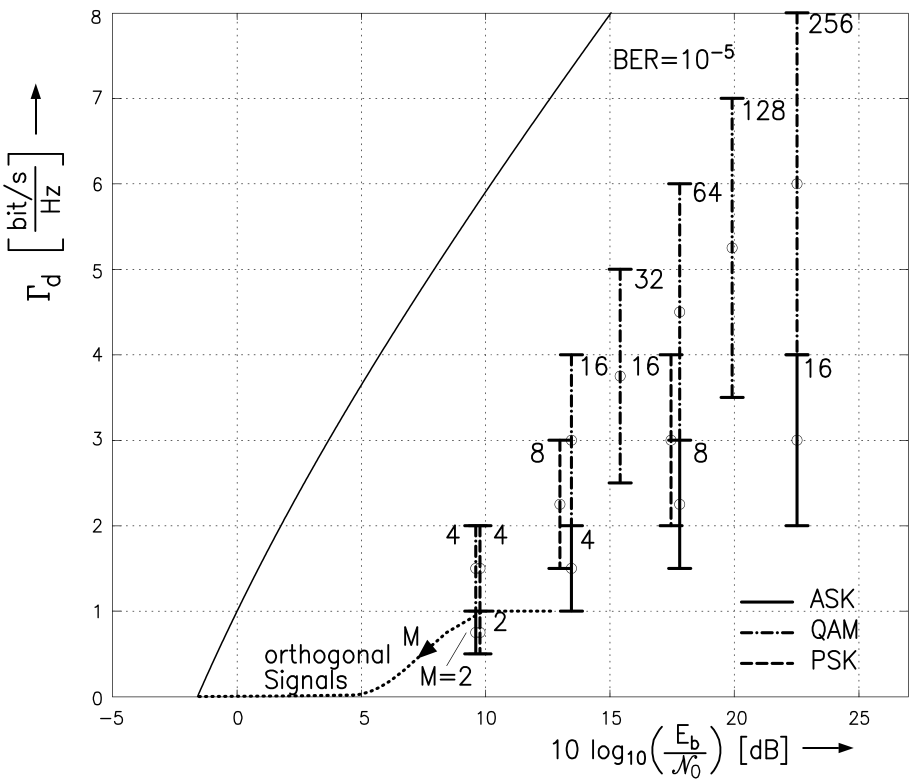
$$= 1 - \frac{1}{\sqrt{2\pi}} \int\limits_0^\infty \mathrm{e}^{\frac{-(y-a)^2}{2}} \left(1 - 2Q(y)\right)^{\frac{M}{2}-1} \mathrm{d}y$$

with

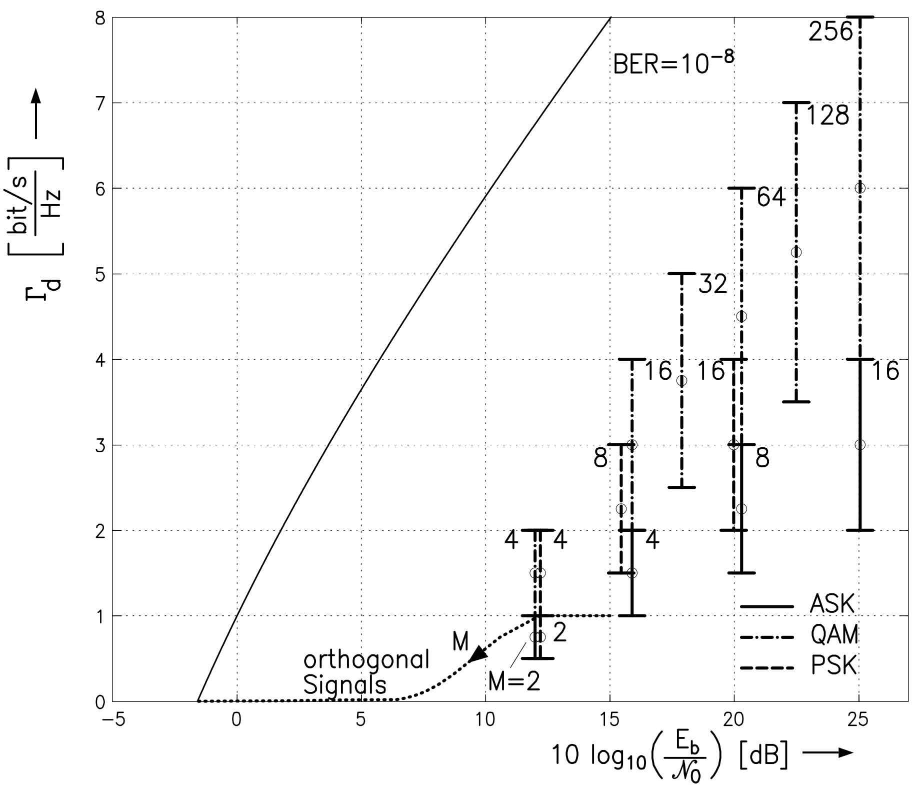
$$y = d_0/\sqrt{N_0/(2T)} \quad ; \quad a = \sqrt{2E_g/N_0}$$

The error exponent may be derived with the same approach as in the case of orthogonal constellations.

Power–bandwidth diagram for uncoded transmission schemes (BER $= 10^{-5}$)



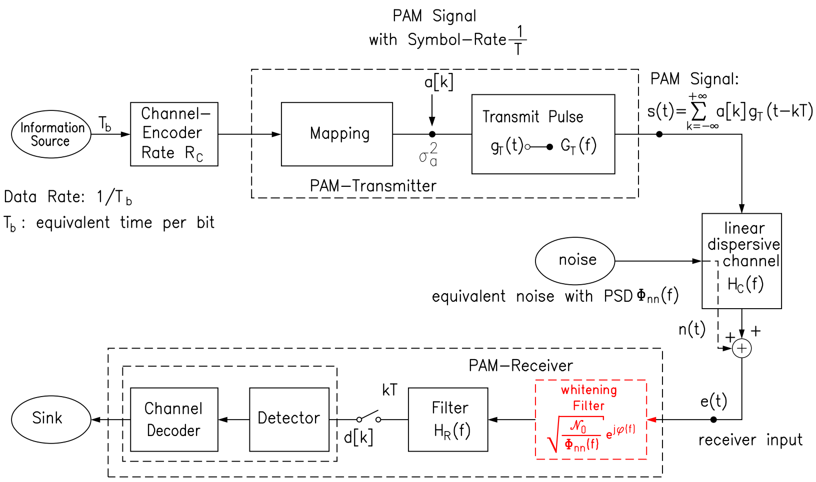
Power–bandwidth diagram for uncoded transmission schemes (BER $= 10^{-8}$)



8 Equalization of Communication Signals

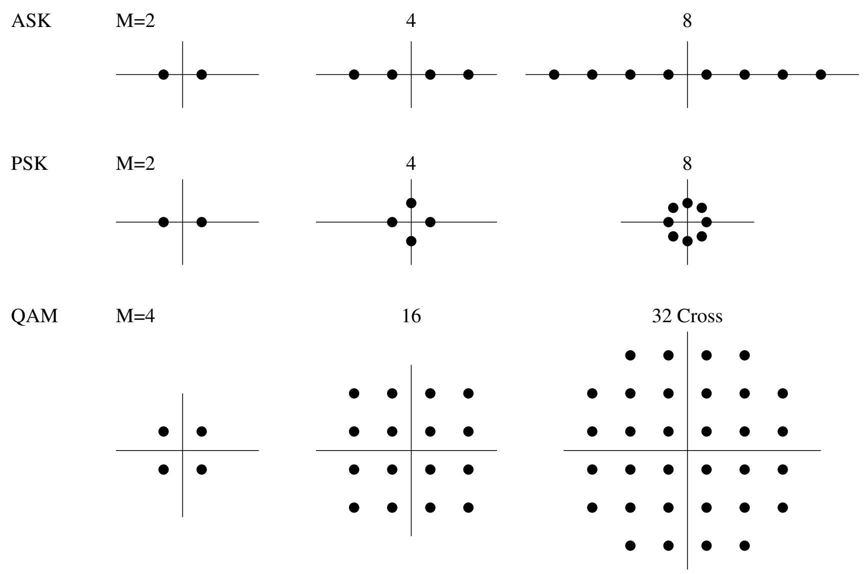
8.1 Basics of Digital PAM Transmission and Coding

Block diagram: (equivalent complex baseband signals (ECB signals))



spectral Signal–to–Noise Ratio: $$\mathrm{SNR}(f) = \frac{\sigma_a^2 |G_T(f)|^2 |H_C(f)|^2}{T \cdot \Phi_{nn}(f)}$$

Typical signal constellations of digital PAM for M–ary signalling



Modulation rate $$R = R_C \cdot \mathrm{ld}\,(M) = \frac{T}{T_b} \qquad \left[\frac{\text{bit}}{\text{symbol}}\right]$$

Examples for linearly distorting channels

– wire of length l:

$$H_{\mathrm{C}}(f) = 10^{-a(f)\cdot l/20} \cdot \mathrm{e}^{-\mathrm{j}\beta(f)l}$$

$$a(f)\text{: Attenuation in dB/km}$$

Approximation $a(f) \approx \sum_{i=1}^{N} a_i \left(\frac{f}{f_0}\right)^{b_i}$ with normalization frequency f_0

At high frequencies: Skin–effect

$$a(f) \approx a_1 \left(\frac{|f|}{f_0}\right)^{0,5}$$

$$\left|H_{\mathrm{C}}(f)\right|^2 = 10^{-a_{\mathrm{c}}\sqrt{2\,|f|\,T}/10}$$

a_{c} [dB]: Characteristic attenuation at **Nyquist frequency** = half symbol rate = $\frac{1}{2T}$

$$a_{\mathrm{c}} = a(f = \tfrac{1}{2T})l \quad \text{[dB]}$$

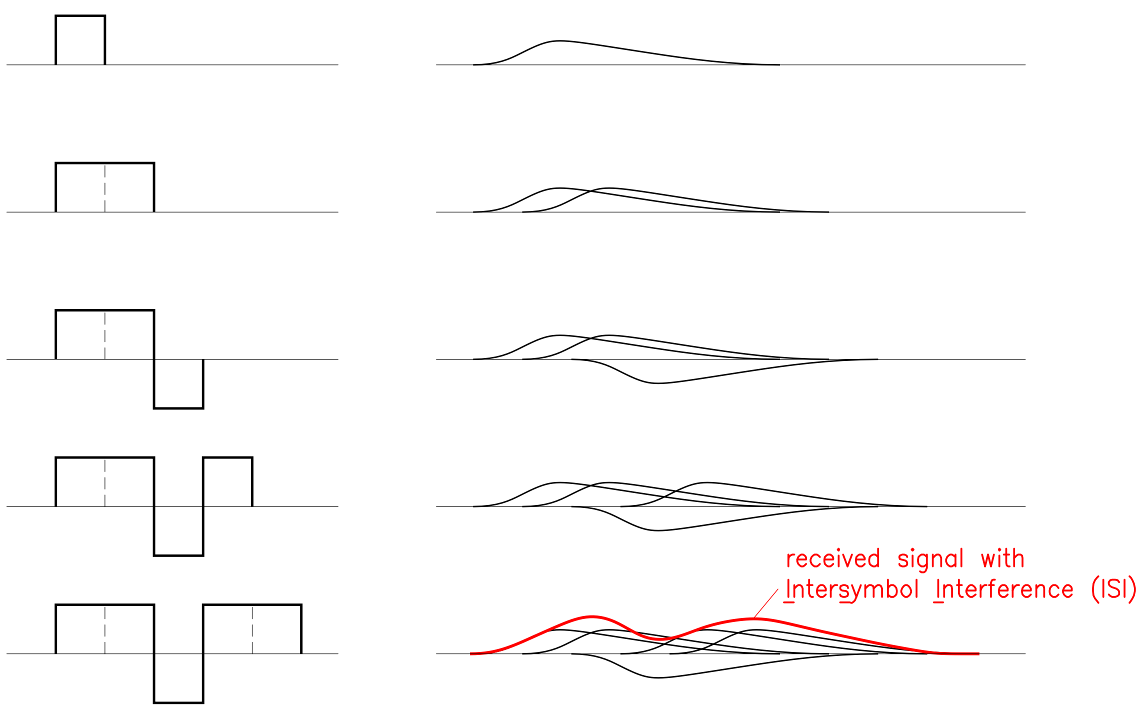
– Multipath propagation in wireless communications:

$$H_{\mathrm{C}}(f) = \sum_{i=1}^{N} a_i \, \mathrm{e}^{-\mathrm{j}2\pi f T_i}$$

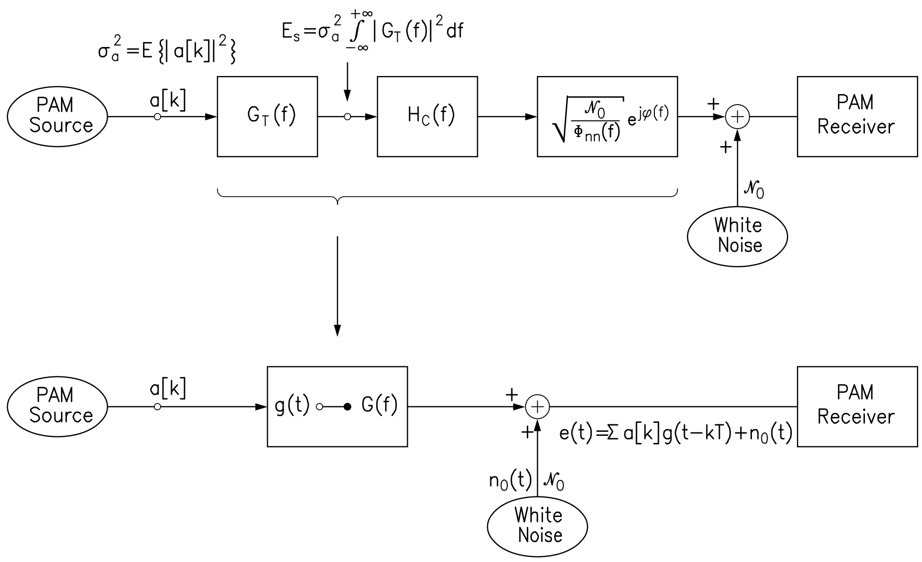
Mobile radio: a_i and T_i are time–variant!

400

Intersymbol Interference (ISI)



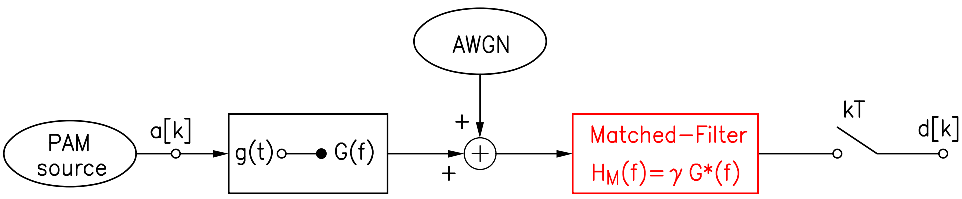
Recapitulation: PAM tansmission, equivalent block diagram



$$\text{with } |G(f)|^2 = |G_T(f)|^2 \cdot |H_C(f)|^2 \cdot \frac{N_0}{\Phi_{nn}(f)}$$

Spectral signal–to–noise ratio at the receiver input: $\mathrm{SNR}(f) = \dfrac{\sigma_a^2 |G(f)|^2}{T N_0}$

PAM transmission without distortion: AWGN channel



Received signal: $$d[k] = \underbrace{\gamma \sum_{l=-\infty}^{+\infty} a[l]\,(g(t) * g^*(-t))|_{t=(k-l)T}}_{\text{signal } \tilde{d}[k]} + \underbrace{n_d[k]}_{\text{noise}}$$

Condition for ISI–free detection: $$\tilde{d}[k] = \gamma \sum_{l=-\infty}^{+\infty} a[l]\,(g(t) * g^*(-t))|_{t=(k-l)T} \stackrel{!}{=} a[k]$$

$$\Rightarrow \boxed{\int_{-\infty}^{+\infty} g(t+kT)g^*(t)\,\mathrm{d}t = \begin{cases} E_g & \text{for } \; k=0 \\ 0 & \text{for } \; k \in \mathbb{Z}\setminus\{0\} \end{cases}}$$

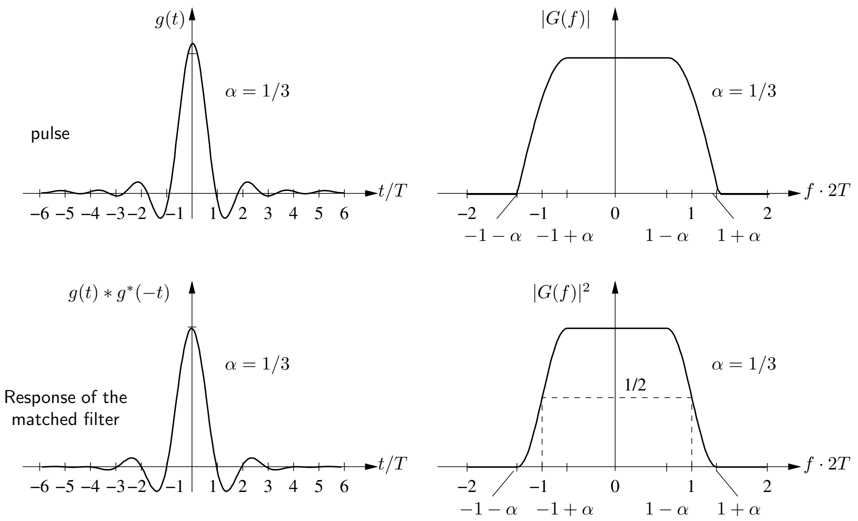
Orthogonality of transmit pulse $g(t)$ with respect to time shifts $k \cdot T$

Normalization: $\gamma = \dfrac{1}{E_g}$ with E_g = pulse energy $= \displaystyle\int_{-\infty}^{+\infty} |G(f)|^2\,\mathrm{d}f = \int_{-\infty}^{+\infty} |g(t)|^2\,\mathrm{d}t$

$$\left.\begin{array}{l} g(t) * g(-t) \quad : \text{Nyquist pulse} \\ \displaystyle\sum_{i=-\infty}^{+\infty} |G(f - i/T)|^2 = E_g\, T = \text{ const.} \end{array}\right\} \text{Nyquist's first criterion (1928)}$$

$$\Rightarrow g(t) : \sqrt{\text{Nyquist}}\text{–pulse}$$

Example: Raised–cosine with roll–off factor $\alpha = \frac{1}{3}$



$\sqrt{\text{Nyquist}}$

Nyquist

Autocorrelation function (ACF) of noise at the filter output:

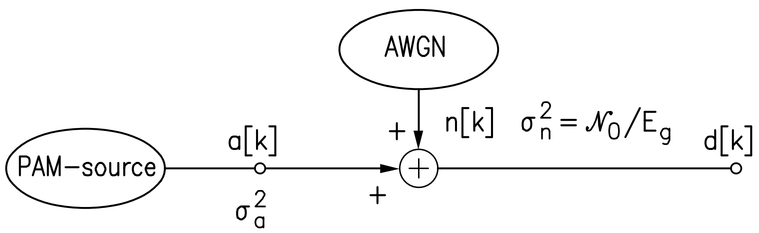
$$\phi_{nn}(\tau) = N_0\, \gamma^2\, g(\tau) * g^*(-\tau)$$

ACF of discrete–time noise:

$$\phi_{nn}[k] = \phi_{nn}(\tau = kT) = \frac{N_0}{E_g}\, \delta[k] = \begin{cases} \dfrac{N_0}{E_g} & \text{for } \; k = 0 \\ 0 & \text{for } \; k \in \mathbb{Z} \setminus \{0\} \end{cases}$$

Discrete white Gaussian noise

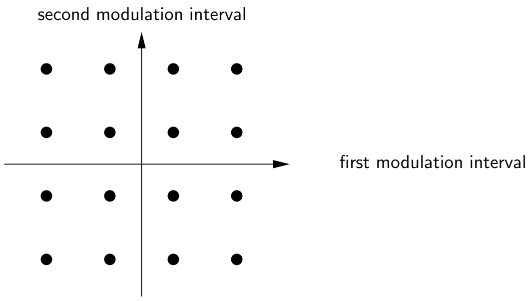
Discrete–time model



PAM transmission

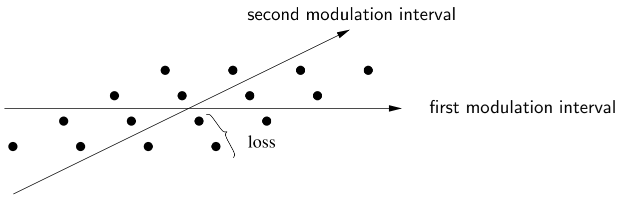
PAM $\stackrel{\text{def}}{=}$ Use of one dimension of a signal space per modulation interval T

Without distortion, ISI–free signalling: Mutual orthogonality of all modulation intervals



With dispersive distortion: ISI generating pulses $g(t)$ that are no longer orthogonal

$\exists\, k \in \mathbb{Z} \setminus \{0\}$ with $\displaystyle\int_{-\infty}^{+\infty} g(t+kT) g^*(t)\,\mathrm{d}t \neq 0$



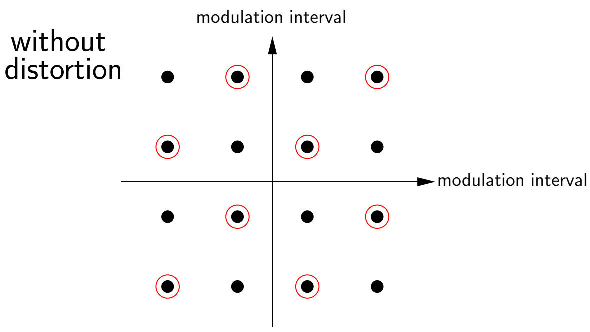
Channel coding, coded modulation

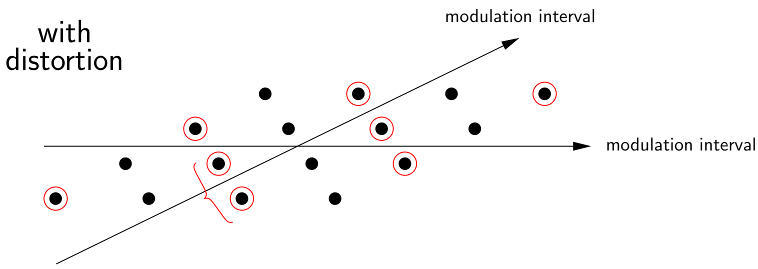
Thinning of all signal points of the resulting constellation in N dimensions by usage of only 2^{RN} instead of $M^N = 2^{\mathrm{ld}(M)\cdot N}$ points in the N–dimensional constellation

Example: 8-ASK, i.e., $M = 8$ signal points per real dimension; choice: $R =$ 2.5 bit/symbol, codeword length $N = 400$

Usage of $2^{2.5\,\cdot\,400} \approx 10^{300}$ codewords out of $2^{3\,\cdot\,400} \approx 10^{360}$ possible words (points in 400 dimensions)

Expurgation of about 10^{60} points per one codeword!





Distance increase

Loss in Minimum distance:
coding inefficient!

PAM transmission with distortion

⇒ Digital PAM transmission with pulses $g(t)$ which are not orthogonal:

$$\int_{-\infty}^{+\infty} g(t+kT)\, g^*(t)\, \mathrm{d}t \neq 0 \quad \text{for } k \in \mathbb{Z}\backslash\{0\}$$

⇒ At the output of the matched filter $\gamma g^*(T_\nu - t)$ ISI occurs

⇒ Symbol by symbol detection is no longer optimum!

digital PAM with linear distortion

$$\equiv$$

digital PAM with non–orthogonal pulses

$$\equiv$$

digital PAM with intersymbol interference (ISI)

ISI or (and) coloured noise are equivalent as far as detection is concerned

8.2 Optimum Receiver

8.2.1 Matched Filter Bound

Brute force solution: Receiver for data **sequences** of length N

Correlation of the receiver input signal $e(t) = \tilde{e}(t) + n(t)$ with all M^N possible noise–free signals:

$$\tilde{e}_i(t) = \sum_{k=0}^{N-1} a_i[k] \cdot g(t - kT)$$

Decide for sequence $a_i[k]$ if

$$i = \underset{\forall \tilde{i}}{\operatorname{argmax}} \left\{ \mathrm{Re}\left\{ \int_{-\infty}^{+\infty} e(t) \cdot \tilde{e}_{\tilde{i}}^*(t)\, \mathrm{d}t \right\} - \frac{1}{2} \int_{-\infty}^{+\infty} |\tilde{e}_{\tilde{i}}(t)|^2 \, \mathrm{d}t \right\}$$

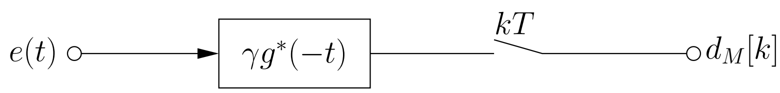
Complexity increases exponentionally with N, i.e., with the duration of data transmission!

but $$\gamma \cdot \int_{-\infty}^{+\infty} e(t) \cdot \tilde{e}_i^*(t)\, \mathrm{d}t = \sum_{k=0}^{N-1} a_i^*[k] \gamma \int_{-\infty}^{+\infty} e(t) g^*(t - kT)\, \mathrm{d}t \overset{\text{def}}{=} \sum_{k=0}^{N-1} a_i^*[k] d_M[k]$$

$\longrightarrow$ Matched Filter provides sufficient statistics for optimum receiver

T–spaced sampling is sufficient and does not cause an information loss

Matched filter $\gamma g^*(-t)$ and T–spaced sampling generates a sequence $d_M[k]$, representing the data sequence of the receiver input signal without loss of information, although the sampling theorem usually is not satisfied!



Sequence $d_M[k]$: Sufficient statistics, but not a practicable solution! ($N \to \infty$)

Theorem: For any equalization method, the transfer function of the optimal receiver input filter includes the continuous–time matched filter, matched to the overall continuous–time pulse $g(t)$, as a factor.

Matched filter bound

Signal to Noise Ratio (SNR) if ISI is ignored; i.e., for transmission of a single, isolated pulse:

Matched–filter–bound:

$$\mathrm{SNR}_{\mathrm{MFB}} = \frac{\sigma_a^2}{N_0/E_g} = \frac{\sigma_a^2}{N_0}\int_{-\infty}^{+\infty} |G(f)|^2 \,\mathrm{d}f =$$

$$= T\int_{-1/(2T)}^{+1/(2T)} \sum_{i=-\infty}^{+\infty} \frac{\sigma_a^2 |G(f-i/T)|^2}{N_0 T}\,\mathrm{d}f = T\int_{-1/(2T)}^{+1/(2T)} \sum_{i=-\infty}^{+\infty} \mathrm{SNR}(f-i/T)\,\mathrm{d}f$$

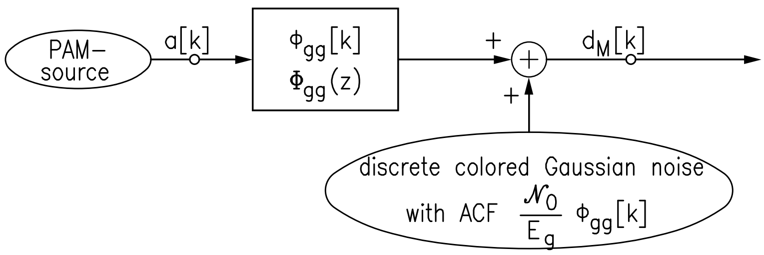
with $\widetilde{\mathrm{SNR}}(f) := \sum_{i=-\infty}^{+\infty} \mathrm{SNR}(f-i/T)$: **folded** spectral SNR at the receiver input after prefilter and T–spaced sampling

$$\boxed{\mathrm{SNR}_{\mathrm{MFB}} = \underbrace{T\int_{-1/(2T)}^{+1/(2T)} \widetilde{\mathrm{SNR}}(f)\,\mathrm{d}f}_{\textbf{Arithmetic mean}\text{ over the folded spectral SNR}}}$$

412

Equivalent discrete–time model

Discrete–time model for the matched filter output signal:



with $\phi_{gg}[k] = \gamma \cdot g(t) * g^*(-t)|_{t=kT} = \frac{1}{E_g}\int_{-\infty}^{+\infty} g(t) g^*(t-kT)\,\mathrm{d}t$ normalized discrete autocorrelation of the overall impulse

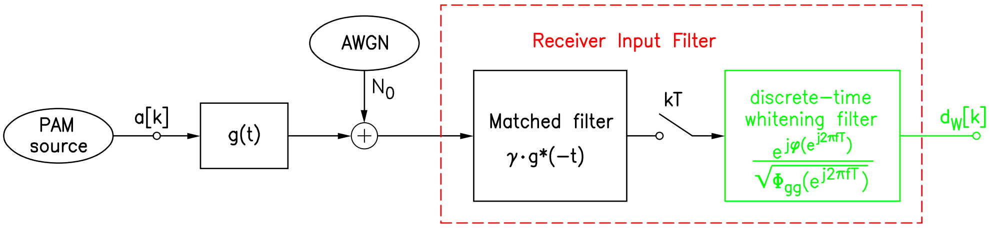
Normalized PSD: $\Phi_{gg}(z) \overset{\text{def}}{=} \sum_{k=-\infty}^{+\infty} \phi_{gg}[k] \cdot z^{-k}$ notice: $\phi_{gg}[0] = 1$

Special case: $z = \mathrm{e}^{\mathrm{j}2\pi fT}$ (Fourier Transform) $\rightarrow \Phi_{gg}(\mathrm{e}^{\mathrm{j}2\pi fT}) = \frac{1}{E_g T}\sum_{i=-\infty}^{+\infty} |G(f-i/T)|^2$

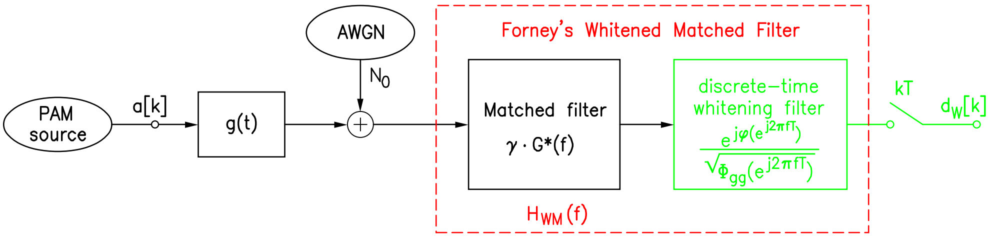
normalized folded energy spectrum of the overall pulse $g(t)$!

But: ISI over L modulation intervals in the receiver input signal is now extended to ISI over $2L$ samples and **the noise is coloured !!**

8.2.2 Forney's Whitened Matched Filter



Sampling at the output of the receiver input filter:



Continuous–time receiver input filter: Whitened matched filter

$$\boxed{H_{WM}(f) = \sqrt{\frac{T}{E_g}} \frac{G^*(f) \cdot e^{j\varphi(f)}}{\sqrt{\sum\limits_{i=-\infty}^{+\infty} |G(f-i/T)|^2}}}$$

Power transfer function: $|H_{WM}(f)|^2 = \frac{T}{E_g} \frac{|G(f)|^2}{\sum_{i=-\infty}^{+\infty} |G(f-i/T)|^2}$

is a Nyquist–function because $\sum\limits_{l=-\infty}^{+\infty} |H_{WM}(f-l/T)|^2 \equiv \frac{T}{E_g}$

$\Rightarrow H_{WM}(f)$ is a Square–Root Nyquist–Function:
$h_{WM}(t) = \mathscr{F}^{-1}\{H_{WM}(f)\}$ produces a set of

orthogonal functions: $\{h_{WM}(t-kT), k \in \mathbb{Z}\}$

Variance of noise in the filter output samples $d_W[k]$:

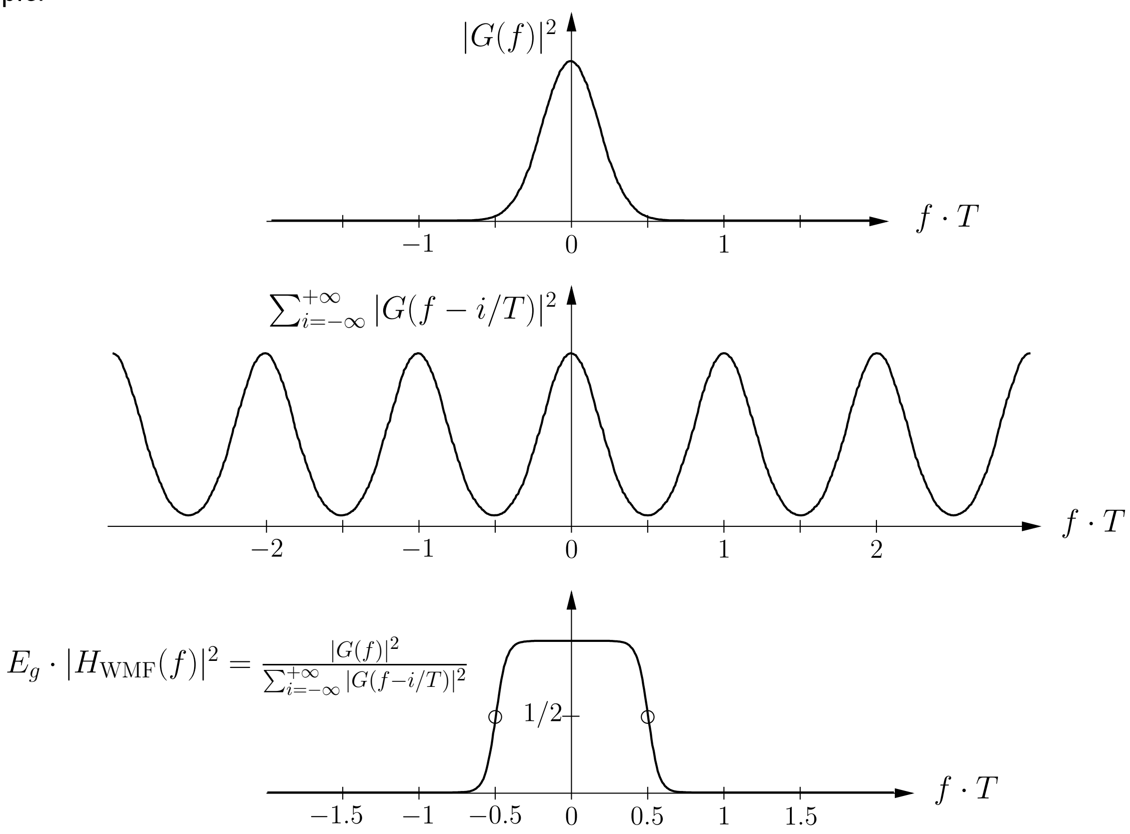
$$\sigma_n^2 = N_0 \int_{-\infty}^{+\infty} |H_{WM}(f)|^2 \, \mathrm{d}f = N_0 \sum_{l=-\infty}^{+\infty} \int_{-\frac{1}{2T}}^{+\frac{1}{2T}} |H_{WM}(f-l/T)|^2 \, \mathrm{d}f = \frac{N_0}{E_g} \cdot \frac{T}{T}$$

Spectral Factorization of: $\Phi_{gg}(z) \stackrel{\text{def}}{=} B(z) \cdot B^*(z^{*-1})$

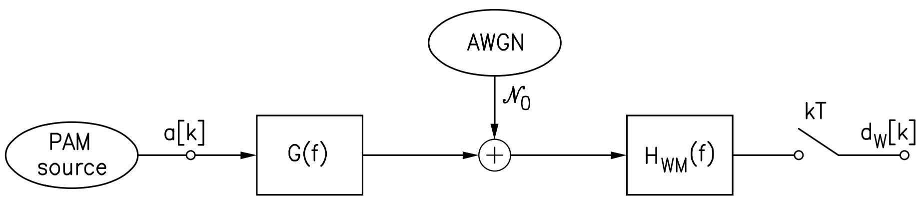
$$\boxed{H_{WM}(f) = \frac{1}{E_g} \frac{G^*(f)}{B^*(e^{j2\pi fT})}}$$

ISI only over L symbol intervals, again!

Example:



416

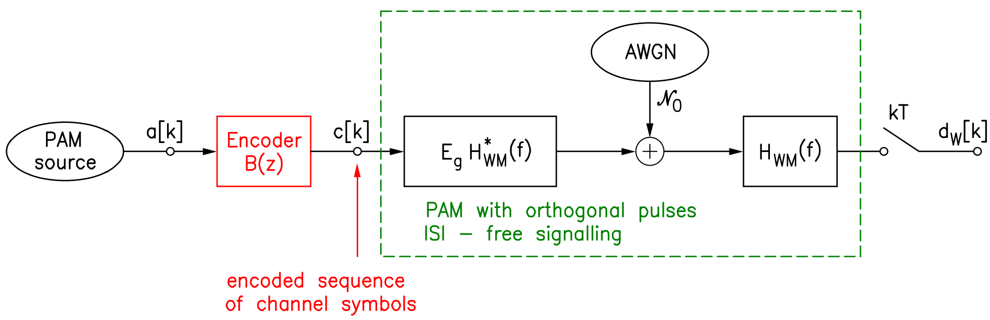


$$G(f) \cdot H_{WM}(f) = G(f) \cdot \frac{1}{E_g} \frac{G^*(f)}{B^*(\mathrm{e}^{\mathrm{j}2\pi fT})} \cdot \frac{B(\mathrm{e}^{\mathrm{j}2\pi fT})}{B(\mathrm{e}^{\mathrm{j}2\pi fT})} = B(\mathrm{e}^{\mathrm{j}2\pi fT}) \cdot E_g \cdot H^*_{WM}(f) \cdot H_{WM}(f)$$

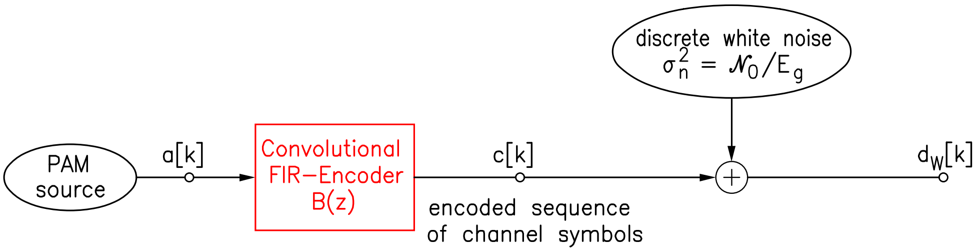
Expansion: $G(f) = B(\mathrm{e}^{\mathrm{j}2\pi fT}) \cdot E_g \cdot H^*_{WM}(f)$ or $g(t) = E_g \cdot \sum_{i=0}^{L} b[i]\, h^*_{WM}(-(t - iT))$

Expansion of $g(t)$ into a sequence of $\sqrt{\text{Nyquist}}$-pulses with coefficients: $b[k] = \mathcal{Z}^{-1}\{B(z)\}$

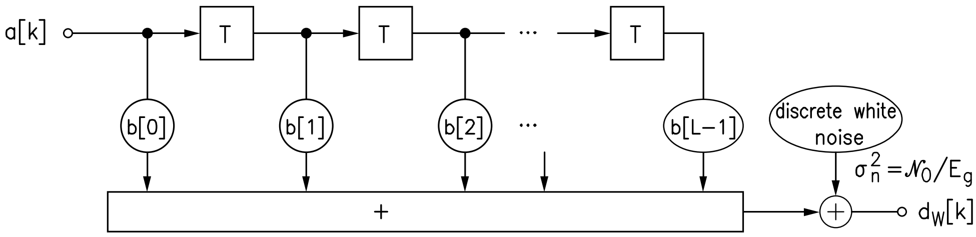
Equivalent block diagram:



Discrete–time equivalent block diagram



FIR encoder:



Trellis encoder with M^{L-1} states notice: $\sum_{k=0}^{L} |b[k]|^2 = \phi_{gg}[0] = 1\,!$

SNR at the output of the WMF is identical to the SNR at the output of the MF!

Properties of $H_{WM}(f)$:

- No equalization, but restriction of noise bandwidth to $\frac{1}{T}$ ($\frac{1}{2T}$ for baseband PAM). Noise whitening balances matched filter around DC: Flat transfer function around DC, high attenuation beyond Nyquist–frequency $\frac{1}{2T}$.
- Almost no "matching" to the equivalent pulse $g(t)$; any $\sqrt{\text{Nyquist}}$–Filter with proper designed slope around the Nyquist–frequency can serve as an approximate whitened matched filter without noticable losses.

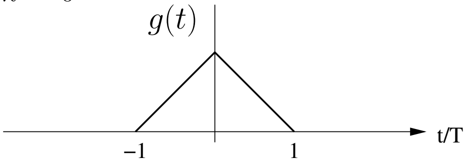
Optimum Data Estimation

- Interpretation of ISI as a trellis code via whitened matched filter
- Application of an efficient trellis decoder:
  - Maximum–Likelihood (or A–Posteriori) Sequence Estimation (MLSE): Viterbi algorithm
  - Maximum–Likelihood (or A–Posteriori) Symbol–by–Symbol Decoder (MLSSD): BCJR algorithm

Example:

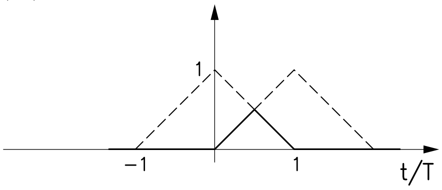
1. **Calculation of $\phi_{gg}[k]$:** $\quad k \in \mathbb{Z}$

$k=0$



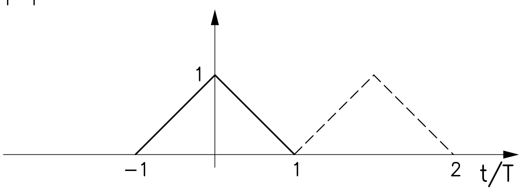
$$2\int_0^1 x^2 \mathrm{d}x = \frac{2}{3} = E_g$$

$|k|=1$



$$\int_0^1 (1-x)x\mathrm{d}x = \frac{1}{6}$$

$|k|>1$



$$\int_0^1 (k-x)x\mathrm{d}x = 0 \quad \forall |k|>1$$

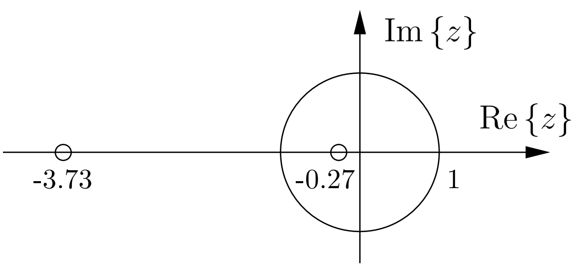
$$\phi_{\mathrm{gg}}[k] = \frac{1}{E_g} g(t) * g(kT-t) = \begin{cases} 1 & \text{for } k=0 \\ \frac{1/6}{2/3} & \text{for } |k|=1 \\ 0 & \text{for } |k|>1 \end{cases}$$

2. **Spectral Factorization of $\Phi_{gg}(z) = B(z)\cdot B^*(z^{*^{-1}})$:**

$$\Phi_{gg}(z) = \frac{1}{4}z + 1 + \frac{1}{4}z^{-1} = \frac{z^{-1}}{4}(z^2+4z+1)$$

Zeros of $\quad z^2+4z+1$: $\quad z_{1/2} = \frac{-4\pm\sqrt{16-4}}{2}$

$z_1 = -2+\sqrt{3} \quad z_2 = -2-\sqrt{3}$ $\quad$ Note: $z_2 = \frac{1}{z_1}$



$$\Phi_{\mathrm{gg}}(z) = \frac{z^{-1}}{4}(z-z_1)(z-z_2) = \frac{(1-z_1z^{-1})}{2\sqrt{-z_1}}\frac{(-z_1z+1)}{2\sqrt{-z_1}}$$
$$= B(z)\cdot B^*(z^{*^{-1}}) \qquad \text{Def.: } c = 2\sqrt{-z_1} = 1.035$$

$$B(z) = \frac{1}{c}(1-z_1z) \quad \Rightarrow \quad b[-1] = 0.259 \quad b[0] = 0.966$$

Note: $b^2[0] + b^2[1] = 1!$

$$B^*(z^{*^{-1}}) = \frac{1}{c}(1-z_1z^{-1}) \quad \Rightarrow \quad b'[0] = \frac{1}{c} = b[0], b'[1] = b[-1]$$

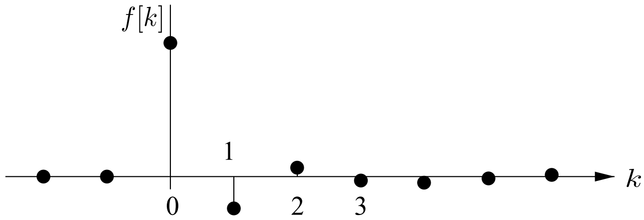
$B^*(z^{*^{-1}})$ has zeros only inside the unit circle in the complex plane: Minimum phase property!
$\Rightarrow$ The inverse $F(z) = 1/B^*(z^{*^{-1}})$ exists and is stable!

3. **Equalization of $B^*(z^{*^{-1}})$:**

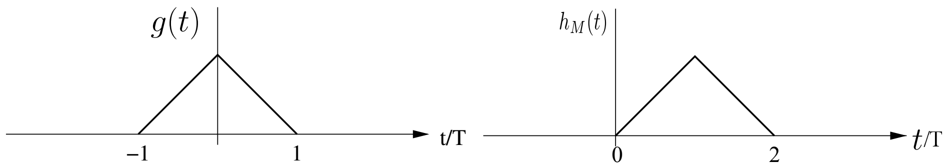
$$F(z) = \frac{1}{B^*(z^{*^{-1}})} = \frac{c}{1-z_1z^{-1}}$$

Geometric Series: $\frac{1}{1-x} = \sum\limits_{i=0}^{\infty} x^i \qquad \forall |x| < 1$

$$= c(1 + z_1z^{-1} + z_1^2z^{-2} + z_1^3z^{-3} + z_1^4z^{-4} + z_1^5z^{-5}\ldots)$$
$$f[k] = (c, cz_1, cz_1^2, cz_1^3, cz_1^4, cz_1^5, \ldots)$$
$$= (+1.035, -0.277, +0.074, -0.02, \ldots)$$

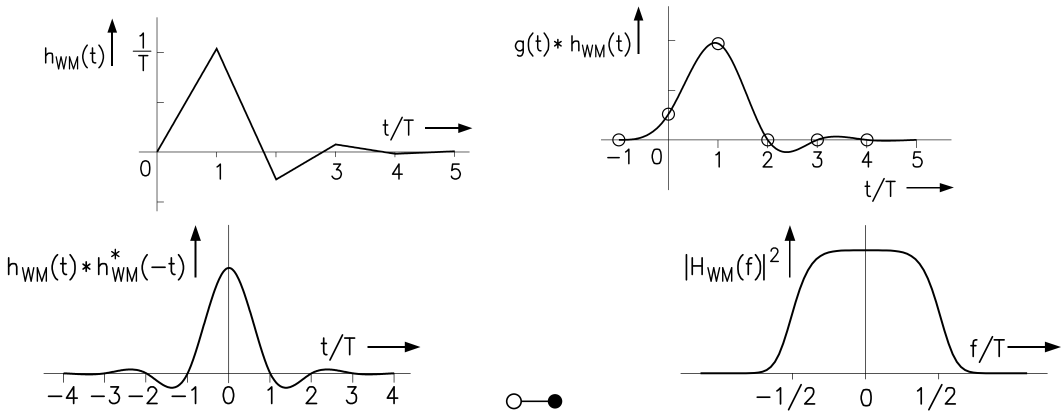


4. **Impulse response of the whitened matched filter:**



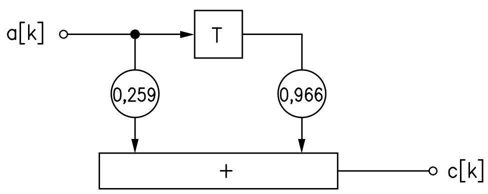
$$h_{\mathrm{WM}}(t+T) \;=\; \sum_{k=0}^{\infty} f[k] g\left(-(t-kT)\right) \qquad (+T \text{ to enforce causality!})$$

$$\text{i.e.,}\quad h_{\mathrm{WM}}(t) \;=\; \sum_{k=0}^{\infty} f[k] h_{\mathrm{M}}(t-kT)$$



5. **Equivalent encoder:** Enforcing causality by one step delay

$$B_c(z) \stackrel{\text{def}}{=} \frac{1}{c}(-z_1 + z^{-1}) \qquad b_c[0] = 0.259 \quad b_c[1] = 0.966$$



Error Probability for MLSE or MLSSD

$\Pr(\text{symbol error}) \approx K \cdot Q\left(\sqrt{\frac{\Delta E_{\min}}{2\,N_0}}\right) = K \cdot Q\left(\sqrt{d^2_{\min}\frac{E_b}{N_0}}\right)$ with:

K: small factor (typical: 0.5 . . . 4)

$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}\,\mathrm{d}t$ Gaussian $Q-$function

$\Delta E_{\min}$: minimum energy of difference sequences: $\Delta E_{\min} = \min\limits_{i\,j} \Delta E_{i\,j}$ with

$$\begin{aligned}\Delta E_{ij} &= \int_{-\infty}^{+\infty} \left| \sum_{k=-\infty}^{+\infty} a_i[k] g(t-kT) - \sum_{k=-\infty}^{+\infty} a_j[k] g(t-kT) \right|^2 \,\mathrm{d}t \\ &= \int_{-\infty}^{+\infty} \left| \sum_{k=-\infty}^{+\infty} \alpha_{ij}[k] g(t-kT) \right|^2 \,\mathrm{d}t\end{aligned}$$

with $\alpha_{ij}[k] = a_i[k] - a_j[k]$ = sequence of difference symbols

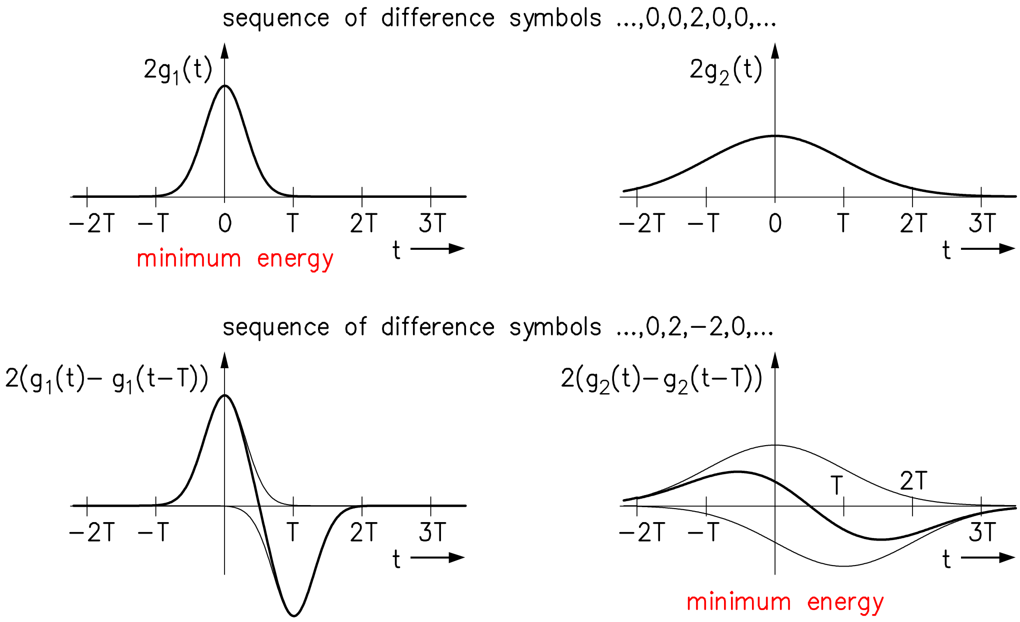
$$\Delta E_{\min} = \min_{\alpha[k]} \Delta E_{ij}$$

$$d^2_{\min} = \Delta E_{\min}/(2E_b) = \text{normalized squared Euclidean distance}$$

$$d^2_{\min} = \frac{1}{2}\min_{\alpha[k]} \left((b[k] * b^*[-k]) * (\alpha[k] * \alpha^*[-k])\right)\big|_{k=0}$$

Example:

Slim and extended pulse $g(t)$



Example: Metallic wire, dispersion by skin effect, white noise

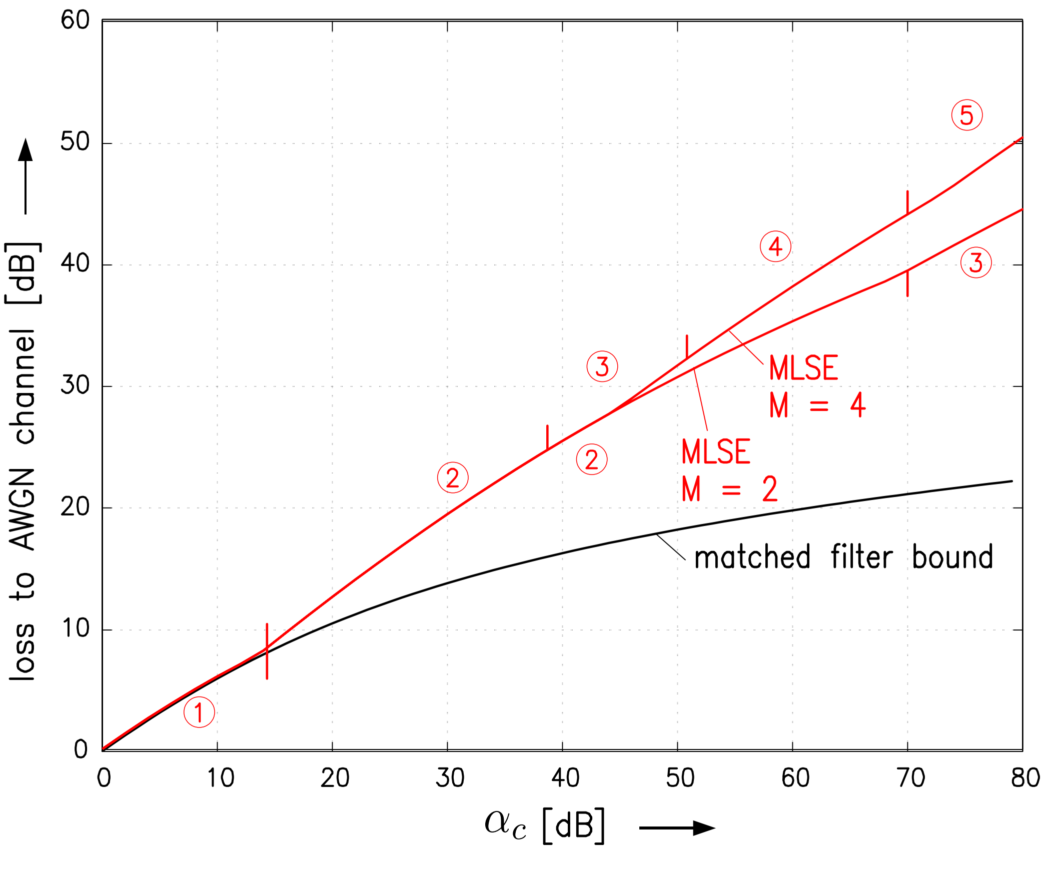
$$|H_C(f)|^2 = 10^{-\alpha_c \sqrt{2|f|T}/10}$$

α_c: Characteristic attenuation at the Nyquist frequency

minimum distance generating difference sequences

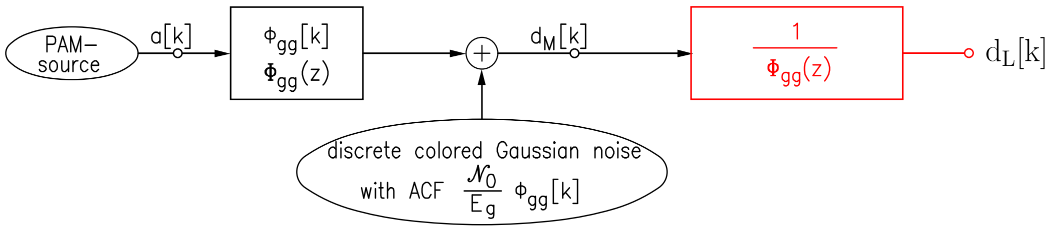
Region	Binary	Quaternary
1	2	2
2	2, -2	2, -2
3	2, -2, -2,2	2, -4, 4, -2
4	–	2, -4, 2
5	–	2, -6, 6, -2

Cable with skin effect: Binary and quaternary signalling



8.3 Linear Equalization

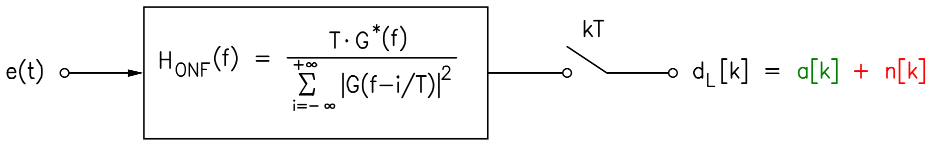
Equivalent block diagram for the output of the matched filter



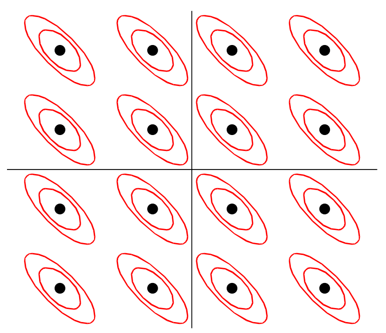
Linear zero–forcing equalization (LZFE) with a discrete–time filter $\dfrac{1}{\Phi_{gg}(z)}$,

if $\Phi_{gg}(e^{j2\pi fT}) = \dfrac{1}{E_g T}\displaystyle\sum_{i=-\infty}^{+\infty} |G(f-i/T)|^2 \neq 0$ for $f \in [0, \dfrac{1}{T})$

Receiver input filter: Optimum Nyquist filter (ONF)



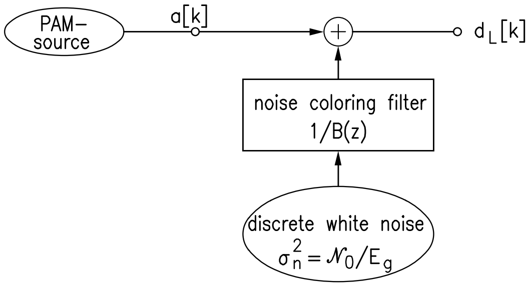
PSD of discrete noise $n[k]$: $\dfrac{N_0}{E_g} \cdot \Phi_{gg}(z) \cdot \dfrac{1}{\Phi_{gg}(z) \cdot \Phi_{gg}^*(z^{*-1})} = \dfrac{N_0/E_g}{\Phi_{gg}^*(z^{*-1})}$



orthogonality for signal

enhanced correlated noise

equivalent
discrete–time
block diagram



spectral factorization:
$\Phi_{gg}(z) = B(z)\cdot B^*(z^{*-1})$

Signal–to–noise ratio, $\mathrm{SNR}_{\mathrm{LZFE}}$, at filter output

Noise variance: $\dfrac{N_0}{E_g} T \displaystyle\int_{-\frac{1}{2T}}^{+\frac{1}{2T}} \frac{1}{\Phi_{gg}(e^{j2\pi fT})}\,\mathrm{d}f = \frac{N_0}{E_g} E_g T^2 \int_{-\frac{1}{2T}}^{+\frac{1}{2T}} \frac{\mathrm{d}f}{\sum_{i=-\infty}^{+\infty} |G(f-i/T)|^2}$

Signal power: $E\left\{|a[k]|^2\right\} = \sigma_a^2 \qquad \Longrightarrow \qquad \mathrm{SNR}_{\mathrm{LZFE}} = \dfrac{\sigma_a^2}{N_0 T^2} \Big/ \displaystyle\int_{-\frac{1}{2T}}^{+\frac{1}{2T}} \frac{\mathrm{d}f}{\sum_{i=-\infty}^{+\infty} |G(f-i/T)|^2}$

Remember: spectral signal–to–noise ratio: $\mathrm{SNR}(f) = \frac{\sigma_a^2}{N_0 T} \cdot |G(f)|^2$

Folded spectral signal–to–noise ratio $\widetilde{\mathrm{SNR}}(f) = \sum_{i=\infty}^{+\infty} \mathrm{SNR}(f - i/T)$

$\mathrm{SNR}_{\mathrm{LZFE}} = 1 \Big/ \left(T \cdot \displaystyle\int_{-\frac{1}{2T}}^{+\frac{1}{2T}} \frac{\mathrm{d}f}{\widetilde{\mathrm{SNR}}(f)} \right) \quad \hat{=}$	Harmonic mean of the folded spectral SNR at the receiver input

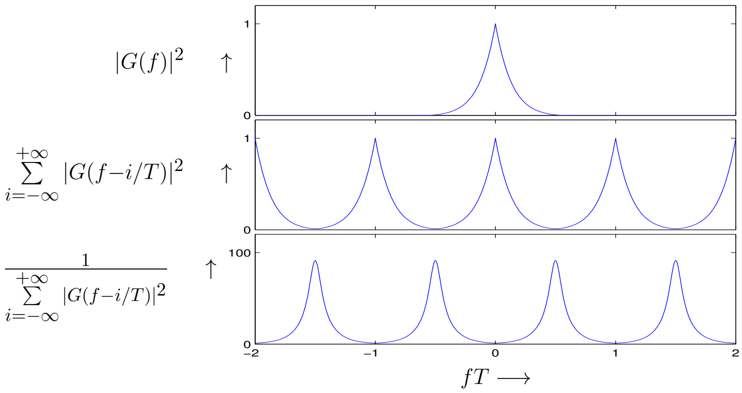
Remember: Harmonic Mean $= 1/(\frac{1}{N}\sum_{i=1}^{N} \frac{1}{a_i})$

432

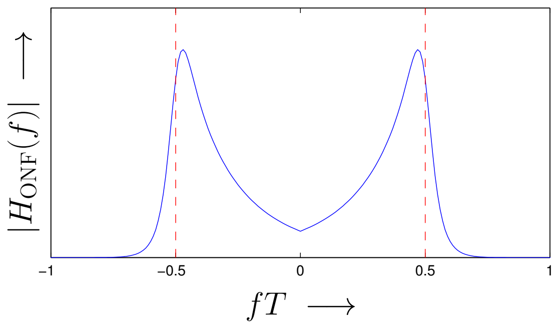
Example: Optimum linear equalization

- simplified DSL baseband transmission (DSL: Digital Subscriber Lines)
- Downstream (Office $\rightarrow$ Subscriber)
- AWGN
- Symmetric pair of length $d = 3\,\mathrm{km}$ (26AWG cable)

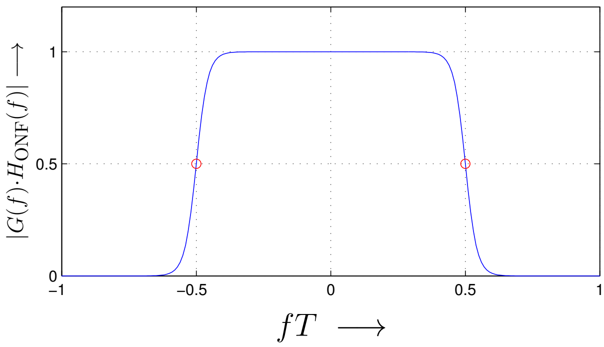
Impulse spectrum $G(f) = H_{\mathrm{T}}(f) \cdot H_{\mathrm{C}}(f)$



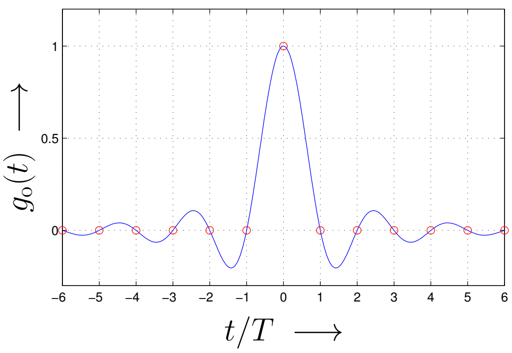
Transfer function of the ONF:



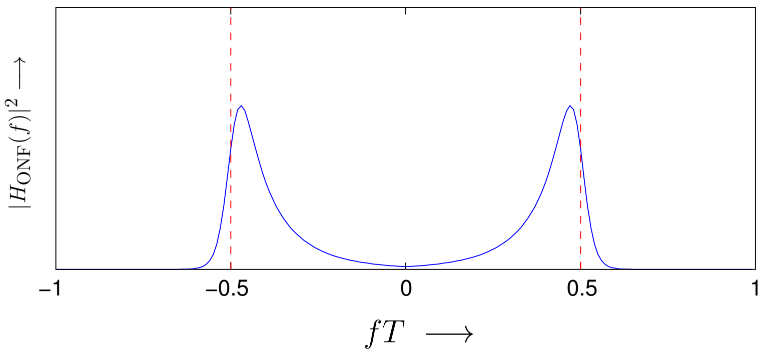
Transfer function $G(f) \cdot H_{\mathrm{ONF}}(f)$



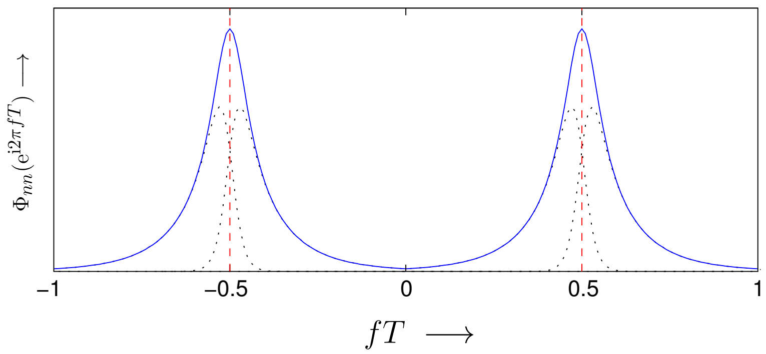
Impulses at the output of the ONF: Nyquist function



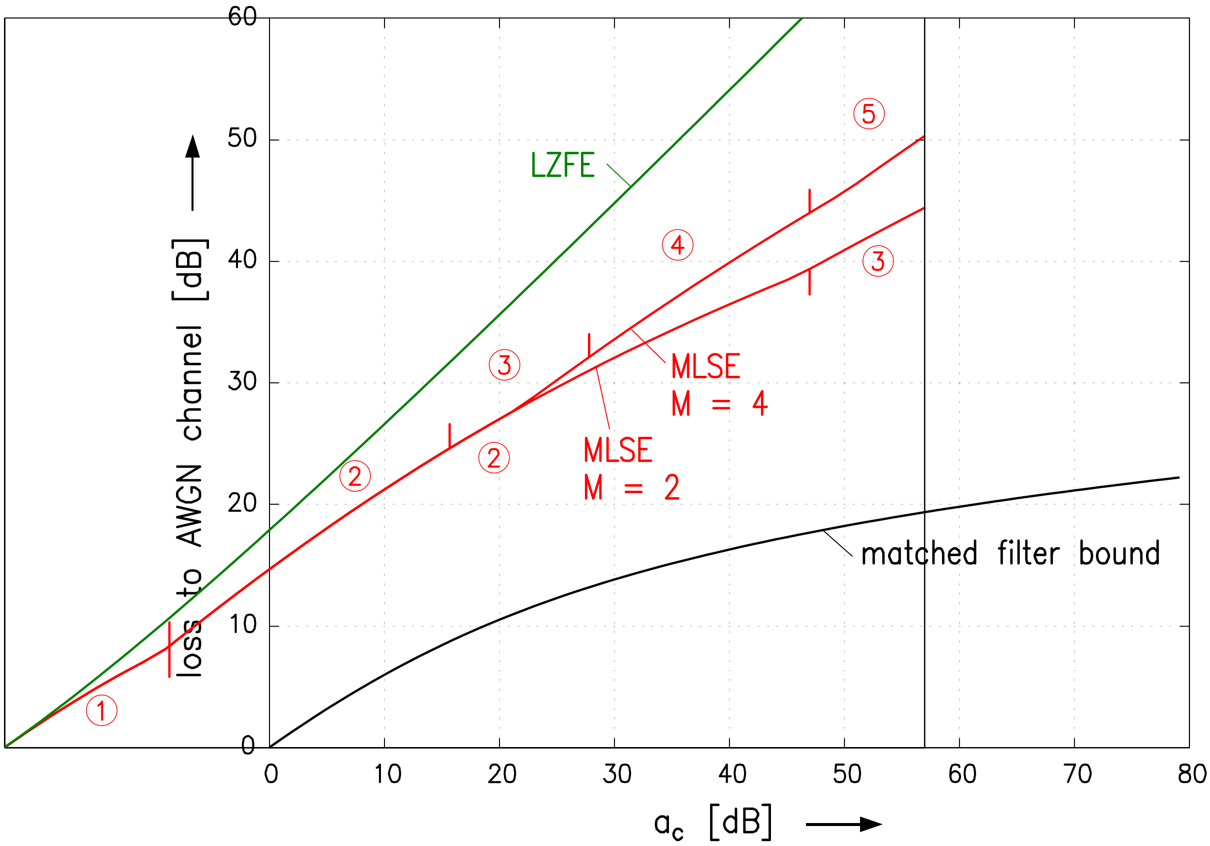
Square of magnitude of the transfer function of the ONF:



Spectral power density of discrete time noise at the output of the ONF.



Cable with skin effect: Binary and quaternary signalling

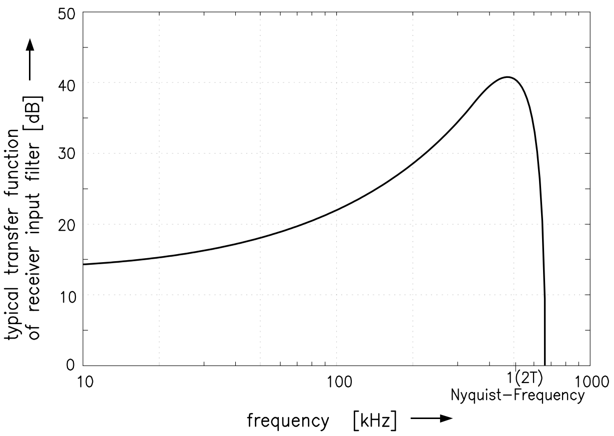


Example: HDSL1: Quaternary signalling over twisted pairs

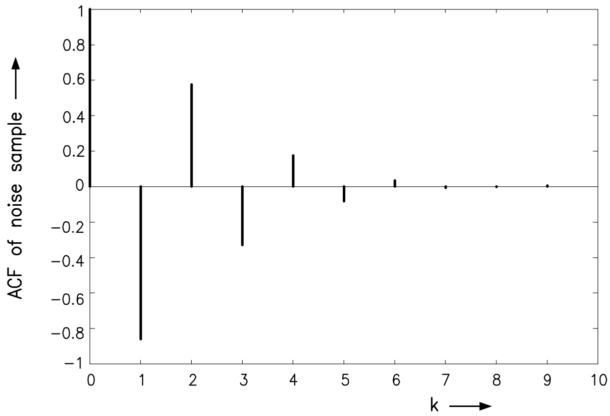
$R_T = 2.048\,\frac{\text{MBit}}{\text{s}}$; $M = 4$ ASK baseband transmission $\frac{1}{T} = 1.024\,\frac{\text{Msymbols}}{\text{s}}$

Nyquist frequency: 512 kHz

Optimum LZF equalization



Autocorrelation function (ACF) of noise samples



8.4 Noise Prediction by Decision Feedback

Supplement 3: *Linear Prediction*

Given: Discrete–time wide sense stationary real-valued stochastic process $q[k]$ with

Autocorrelation sequence: $\phi_{qq}[\lambda] = \mathrm{E}\left\{q[k+\lambda]\,q[k]\right\} \quad \forall k$

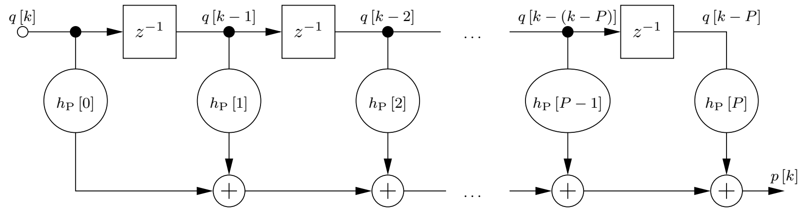
Power spectral density: $\Phi_{qq}\left(\mathrm{e}^{\mathrm{j}2\pi f T_\mathrm{A}}\right) = \sum_{\lambda=-\infty}^{+\infty} \phi_{qq}[\lambda]\cdot \mathrm{e}^{-\mathrm{j}2\pi f T_\mathrm{A}\lambda}$

Linear prediction:

- Calculation of a predicted value $p[k]$ for the next value $q[k+1]$ using the values observed so far $q[\kappa]$, $\kappa = k, k-1, k-2, \ldots$,
by means of a linear system: Prediction filter

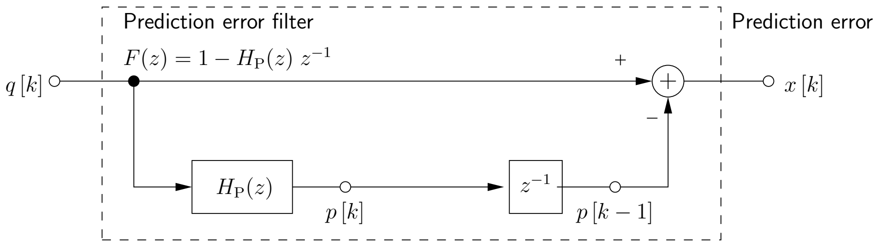
FIR filter $H_\mathrm{P}(z) \bullet\!\!-\!\!\circ\, h_\mathrm{P}[k]$ with degree P:

$$H_\mathrm{P}(z) = h_\mathrm{P}[0] + h_\mathrm{P}[1]\,z^{-1} + h_\mathrm{P}[2]\,z^{-2} + \ldots + h_\mathrm{P}[P]\,z^{-P}$$



- Predicted value: $p[k] \;=\; q[k] * h_\mathrm{P}[k] \;=\; \sum_{\kappa=0}^{P} h_\mathrm{P}[\kappa]\, q[k-\kappa]$
- Prediction error: $x[k] \;=\; q[k] - p[k-1]$
- Prediction error filter: $F(z) \;=\; 1 - H_\mathrm{P}(z) \cdot z^{-1} \;=\; \sum_{k=0}^{P+1} f[k]\, z^{-k}$,

  where $f[0] = 1$, $f[k] = -h_\mathrm{P}[k-1]$, $k = 1, 2, \ldots, P+1$



Optimization of the predictor:

Goal: Minimize variance $\sigma_x^2 = \mathrm{E}\{x^2[k]\}$ of the prediction error signal

$$x[k] \;=\; q[k] - \sum\nolimits_{\kappa=0}^{P} h_\mathrm{P}[\kappa]\, q[k-(\kappa+1)]$$

For the minimum, we must have

$$\frac{\partial}{\partial h_\mathrm{P}[l]}\, \mathrm{E}\left\{ \left(q[k] - \sum_{\kappa=0}^{P} h_\mathrm{P}[\kappa]\, q[k-(\kappa+1)] \right)^2 \right\} \;\overset{!}{=}\; 0 \qquad \forall l = 0, 1, \ldots, P$$

$$-2\,\mathrm{E}\left\{ \left(q[k] - \sum_{\kappa=0}^{P} h_\mathrm{P}[\kappa]\, q[k-(\kappa+1)] \right) \cdot q[k-(l+1)] \right\} \;\overset{!}{=}\; 0$$

$$\Rightarrow \underbrace{\mathrm{E}\{q[k]\, q[k-(l+1)]\}}_{\phi_{qq}[-(l+1)]} = \sum_{\kappa=0}^{P} h_\mathrm{P}[\kappa] \cdot \underbrace{\mathrm{E}\{q[k-(\kappa+1)]\, q[k-(l+1)]\}}_{\phi_{qq}[l-\kappa]}$$

The coefficients $h_\mathrm{P}[k]$ of the optimum prediction filter can be calculated from the ACF $\phi_{qq}[\kappa]$ of the input process $q[k]$ by solving a system of linear equations!

Yule–Walker equations with $\phi_{qq}[-\kappa] = \phi_{qq}^*[\kappa]$:

$$\begin{bmatrix} \phi_{qq}[0] & \phi_{qq}[1] & \cdots & \phi_{qq}[P] \\ \phi_{qq}[-1] & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \phi_{qq}[1] \\ \phi_{qq}[-P] & \cdots & \phi_{qq}[-1] & \phi_{qq}[0] \end{bmatrix} \cdot \begin{bmatrix} h_{\mathrm{P}}[0] \\ h_{\mathrm{P}}[1] \\ \vdots \\ h_{\mathrm{P}}[P] \end{bmatrix} = \begin{bmatrix} \phi_{qq}[-1] \\ \phi_{qq}[-2] \\ \vdots \\ \phi_{qq}[-P-1] \end{bmatrix}$$

Minimum variance of the prediction error $x[k]$:

$$\sigma_{x,\min}^2 = \underbrace{\phi_{qq}[0]}_{\text{Variance before the prediction filter}} - \underbrace{\sum_{\kappa=0}^{P} h_{\mathrm{P}}[\kappa]\,\phi_{qq}[\kappa+1]}_{\text{Variance reduction using optimal prediction}}$$

Gain in average signal power (reduction of average signal power) by means of linear prediction:

$$\begin{aligned} G_{\mathrm{P,opt}} &= 10\ \lg\left(\frac{\sigma_q^2}{\sigma_{x,\min}^2}\right) \\ &= -10\ \lg\left(1 - \sum_{\kappa=0}^{P} h_{\mathrm{P}}[\kappa]\,\phi_{qq}[\kappa+1]/\phi_{qq}[0]\right) \end{aligned}$$

Remarks:

- The solution Yule–Walker equations always corresponds to a causal minimum phase prediction error filter $F(z)$, i.e., all zeros of $1/F(z)$ are located inside the unit circle of the complex plane. Therefore, $1/F(z)$ can be implemented as a causal and stable filter.
- For a sufficiently high degree P of the prediction filter, the prediction error signal will be a white stochastic process, i.e., its PSD is constant. When all correlations are exploited, i.e., the error signal is a white process, no further gain in variance is possible with linear prediction.

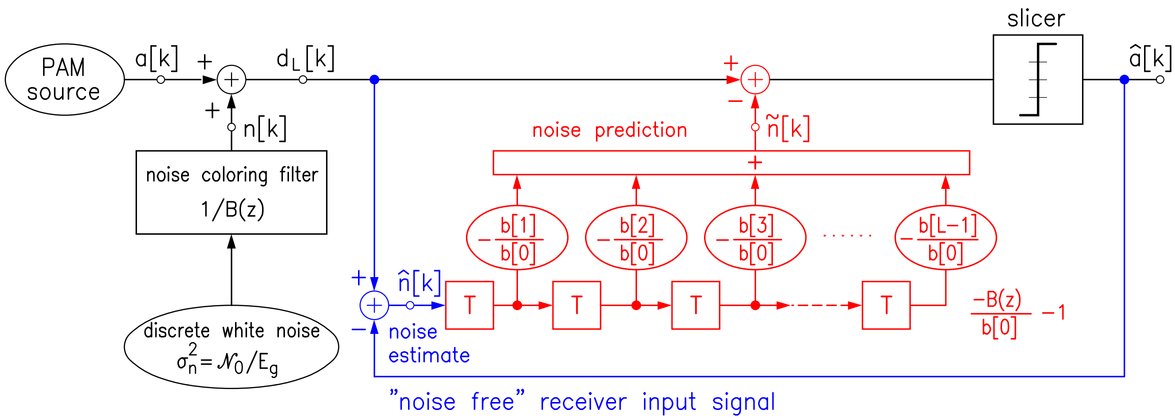
  This means: Prediction error filter $\equiv$ Whitening filter

  resulting in a white process with minimum variance.
- For sufficiently long prediction filters, the pdf of the error signal $x[k]$ is approaching a Gaussian distribution due to the central limit theorem.

End Supplement 3

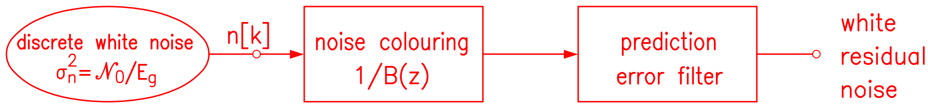
8.4.1 Noise Prediction by Decision Feedback

LZFE: Equivalent discrete–time block diagram

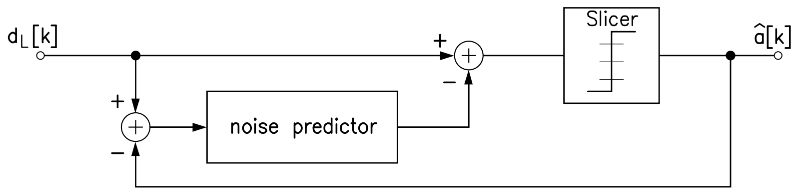


Error free decision $\Rightarrow$ Cancellation of noise

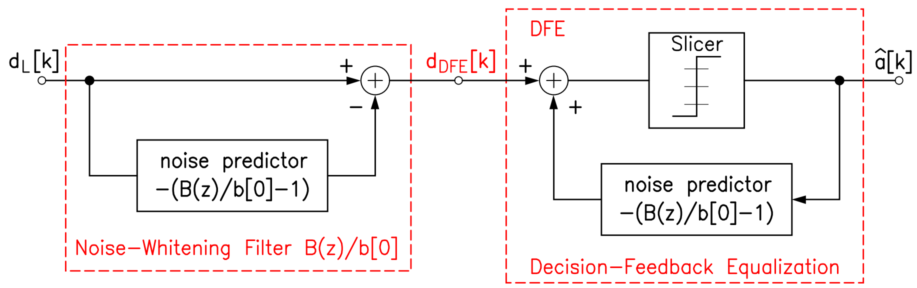
Equivalent block diagram for (discrete–time) noise



Decision–feedback equalization (DFE)



$$\equiv$$



Introduces ISI again

Equalization $b[0]/B(z)$ for signal only!

No noise enhancement!

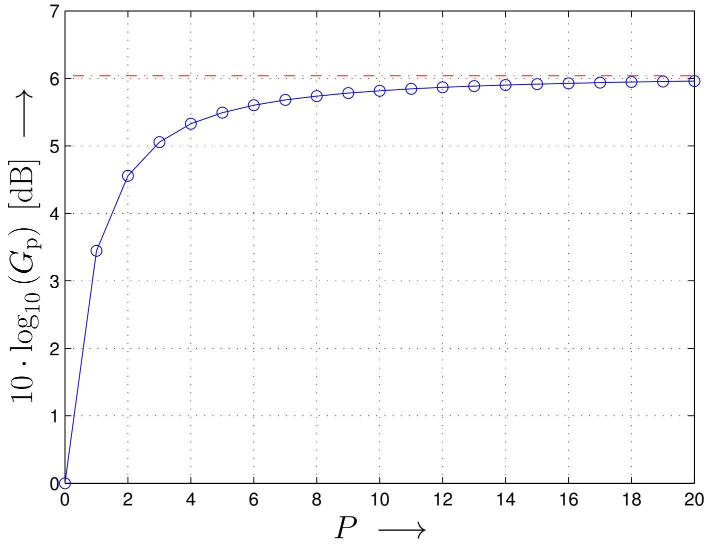
Split of signal and noise by the slicer.

Noise prediction and DFE are equivalent!

Example: Noise prediction and prediction gain

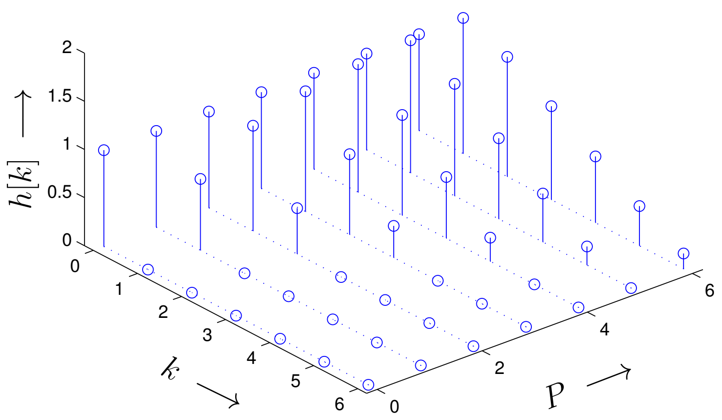
Simplified DSL upstream example; length of cable $\ell = 3\,\text{km}$ (26AWG cable)

Prediction gain versus degree of prediction filter P:



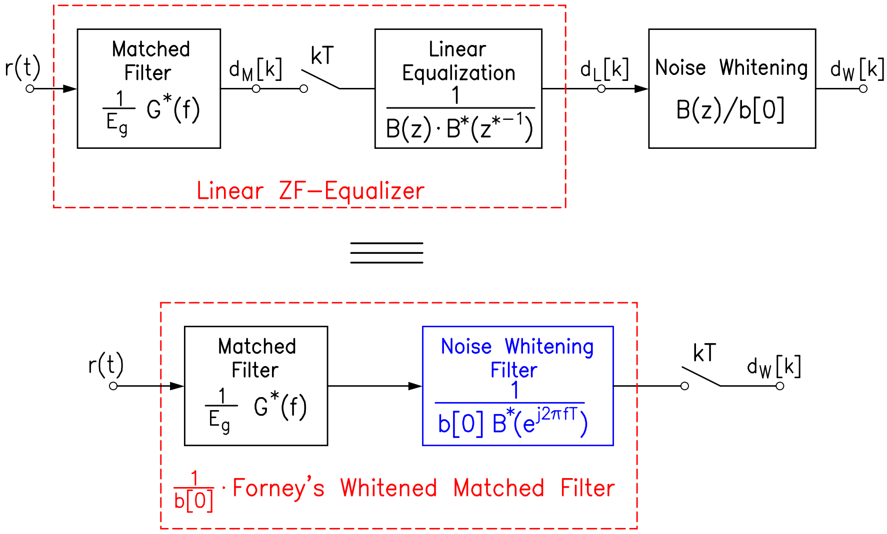
Here: Asymptotic Gain 6.04 dB

Impulse response of the prediction error filter $B(z)$ versus degree P:



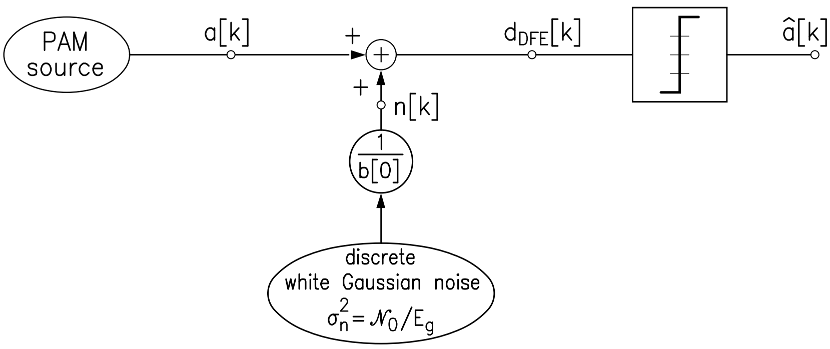
8.4.2 Zero Forcing Decision–Feedback Equalization

Receiver Input–Filter



Optimum receiver input filter for zero–forcing decision–feedback equalization (ZF-DFE):
Forney's whitened matched filter

Equivalent discrete–time block diagram for correct decisions (i.e. without error propagation)



$$\boxed{\mathrm{SNR}_{\mathrm{ZF-DFE}} = \frac{\sigma_a^2 E_g}{N_0} \cdot |b[0]|^2}$$

Maximum SNR $\Rightarrow$ maximum $b[0]$ with $B(z) \cdot B^*(z^{*-1}) = \Phi_{gg}(z)$

Decomposition of $\Phi_{gg}(z)$ into causal **minimum phase** filter $B(z)$
and (anti–) causal **maximum phase** filter $B^*(z^{*-1})$

Optimum receiver input filter: Whitened matched filter, generating a **minimum phase** discrete–time pulse $b[k] \longleftrightarrow B(z)$ after sampling from the continuous–time pulse $g(t)$.

Spectral Factorization of the discrete–time PSD of the Impulse

(equivalent) PAM pulse: $g(t) \circ\!\!-\!\!\bullet\, G(f)$

ACF: $g(t) * g^*(-t) \circ\!\!-\!\!\bullet\, |G(f)|^2$

discrete–time ACF: $g(t) * g^*(-t)|_{t=kT} = E_g\, \phi_{gg}[k]$

PSD: $\Phi_{gg}(z) = \sum\limits_{k=-\infty}^{+\infty} \phi_{gg}[k]\, z^{-k}$; $\quad \Phi_{gg}(\mathrm{e}^{\mathrm{j}2\pi fT}) = \frac{1}{E_g T} \sum\limits_{i=-\infty}^{+\infty} |G(f - i/T)|^2$

Spectral factorization: $\Phi_{gg}(z) = B(z) \cdot B^*(z^{*-1})$

Cepstrum method: $\ln\left(\Phi_{gg}(z)\right) = \ln\left(B(z)\right) + \ln\left(B^*(z^{*-1})\right)$

$B(z) \longleftrightarrow b[k]$ minimum phase part

$$\boxed{b[0] = \exp\left[\frac{1}{2} T \int_{-1/(2T)}^{1/(2T)} \ln\left(\Phi_{gg}(\mathrm{e}^{\mathrm{j}2\pi fT})\right)\, \mathrm{d}f\right]}$$

Remember: $\sum\limits_{i=-\infty}^{+\infty} |G(f - i/T)|^2 = \widetilde{\mathrm{SNR}}(f) \cdot \frac{T N_0}{\sigma_a^2}$

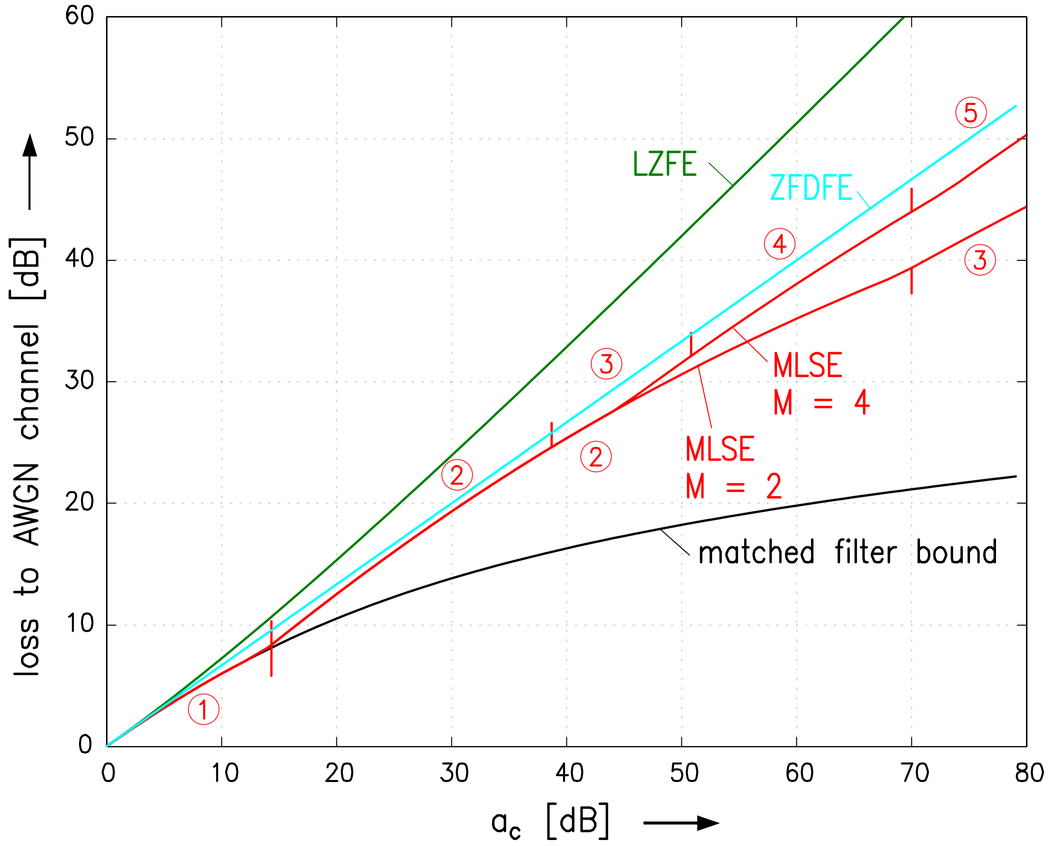
$$\boxed{\mathrm{SNR}_{\mathrm{ZF-DFE}} = \exp\left[T \int_{-1/(2T)}^{1/(2T)} \ln\left(\widetilde{\mathrm{SNR}}(f)\right)\, \mathrm{d}f\right]}$$

SNR for ZF-DFE:

Geometric Mean of the folded SNR at the receiver input

Remember: Geometric mean $\left(\prod\limits_{i=1}^{D} a_i\right)^{1/D} = \exp\left[\frac{1}{D} \sum\limits_{i=1}^{D} \ln(a_i)\right]$

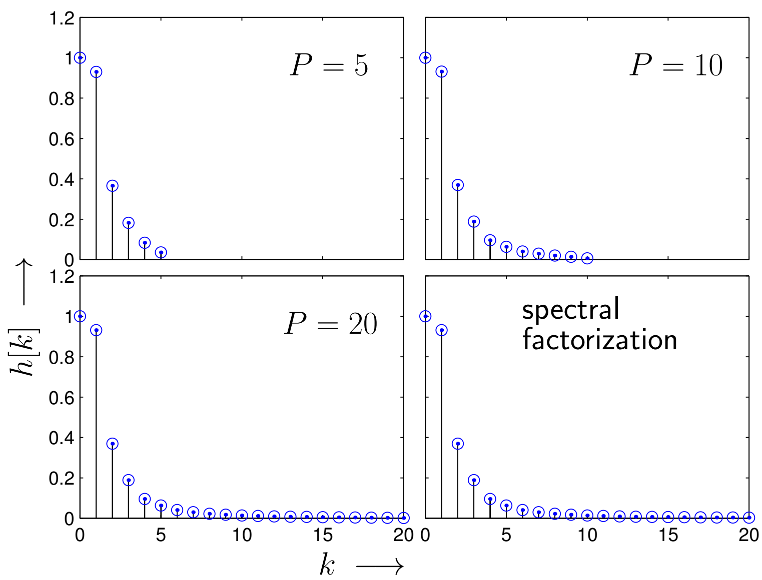
Cable with skin effect: Binary and quaternary signalling



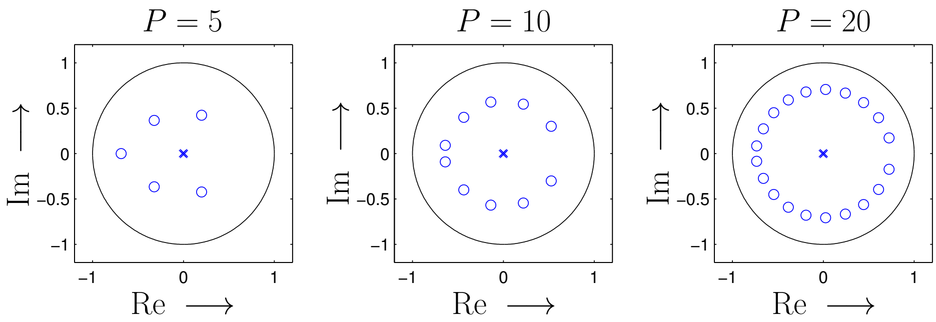
Example: Whitened matched filter

Simplified DSL downstream modell

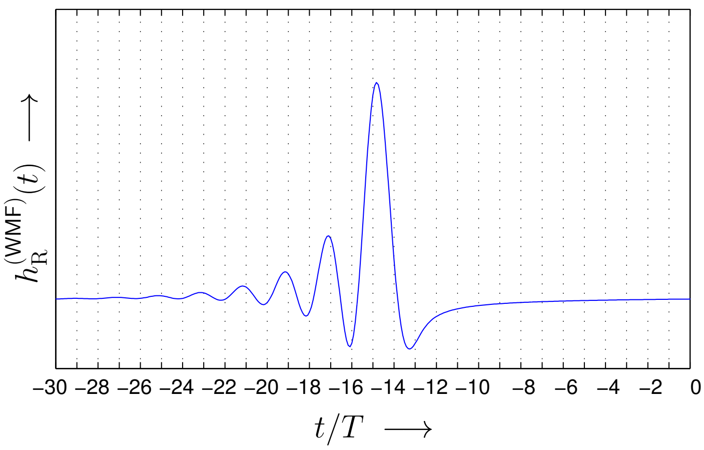
Discrete time end–to–end impulse response



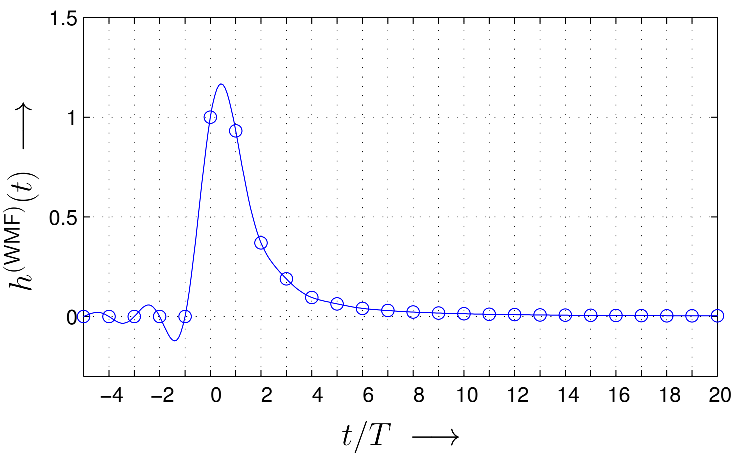
Poles–Zeros diagram:



Impulse response of the WMF:



End–to–end impulse response



8.4.3 Error Probability for Uncoded ZF-DFE

Error propagation is ignored (Genie-aided)

$$\Pr(\text{symbol error}) \approx N_{\min}\, Q\left(\sqrt{\mathrm{SNR}_0}\right)$$

$Q(x)$: Gaussian Q-function $\quad Q(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_x^{\infty} e^{-t^2/2}\, \mathrm{d}t$

$N_{\min}$: number of nearest–neighbour signal points

ASK, baseband PAM: $N_{\min} = 1 \ldots 2$

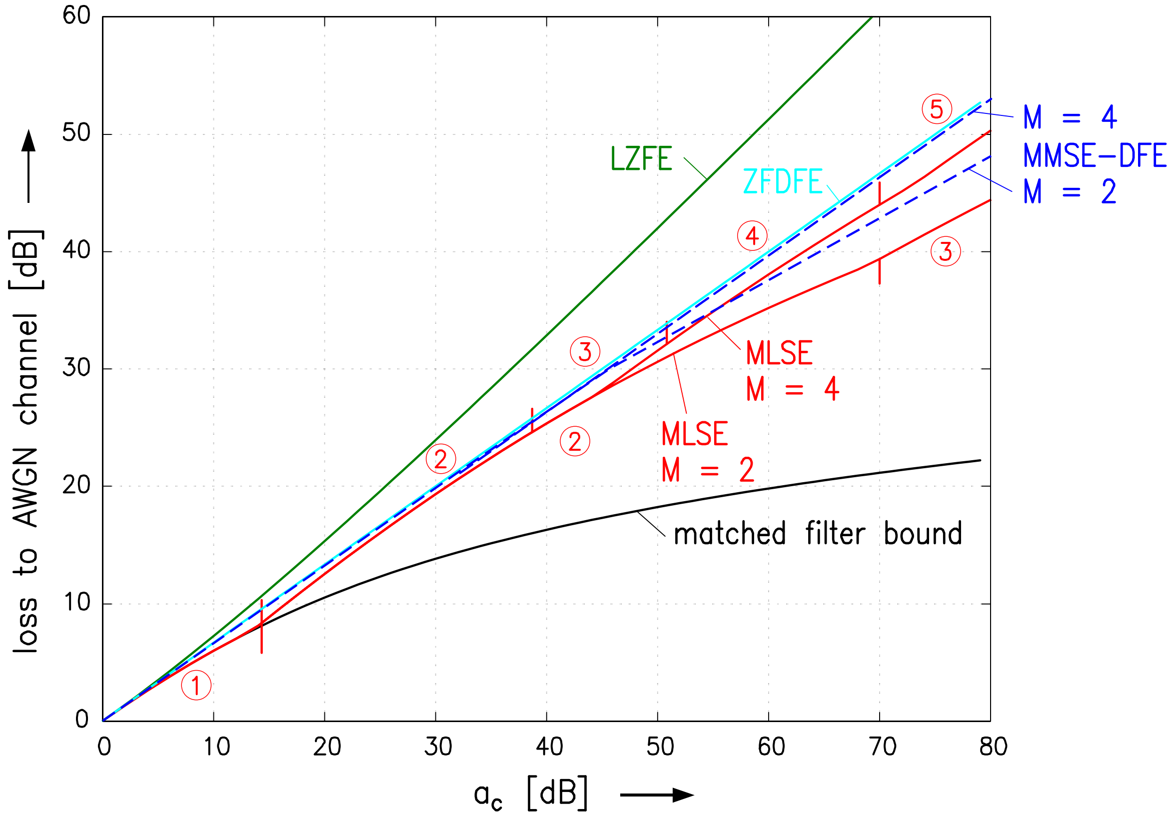
QAM : $N_{\min} = 2 \ldots 4$

$$\mathrm{SNR}_0 = \frac{(\text{min. Euclidean distance}/2)^2}{\sigma_n^2} = \mathrm{SNR}_{\mathrm{ZF-DFE}}/\sigma_a^2$$

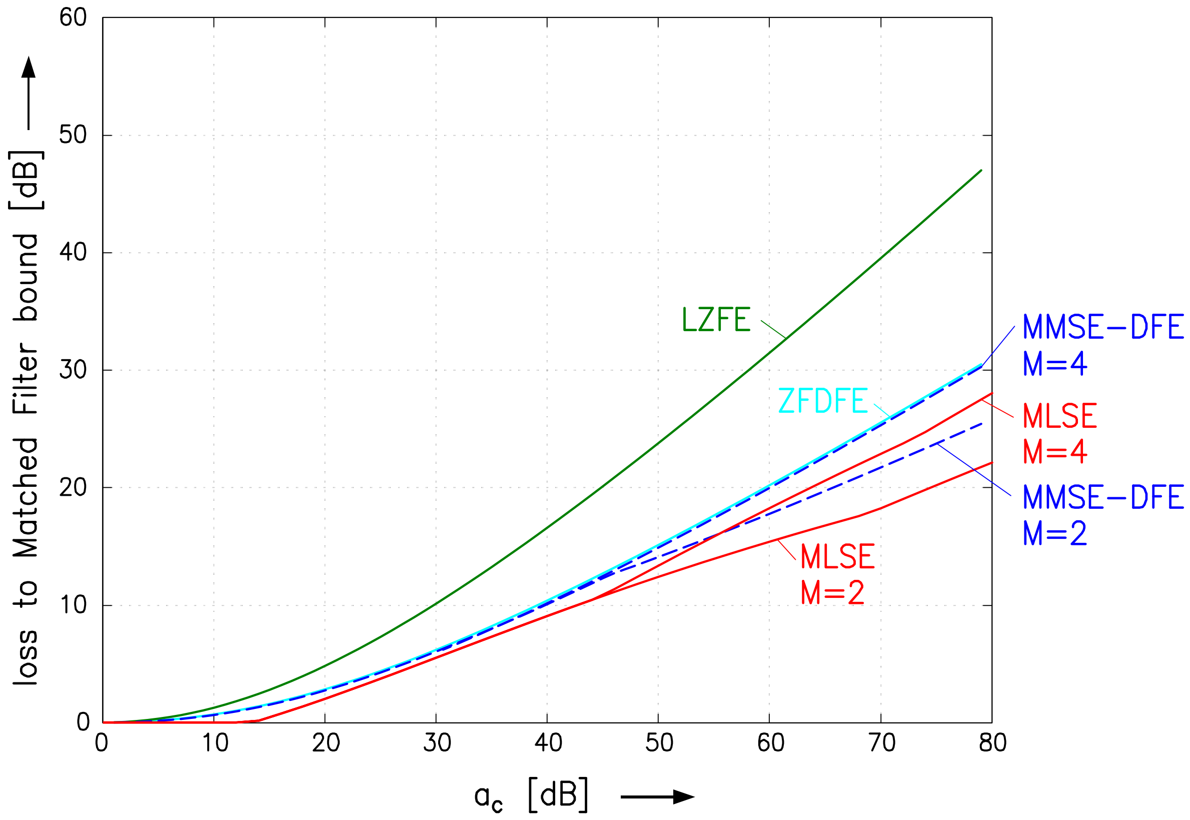
QAM: $\sigma_a^2 = 2\dfrac{(\sqrt{M})^2 - 1}{3} \approx \dfrac{2}{3} 2^R$ with rate R in $\left[\dfrac{\mathrm{bit}}{\mathrm{symbol}}\right]$

$$\mathrm{SNR}_0 \approx \mathrm{SNR}_{\mathrm{ZF-DFE}} \cdot \frac{3}{2} 2^{-R}$$

Cable with skin effect: Binary and quaternary signalling – loss to AWGN channel



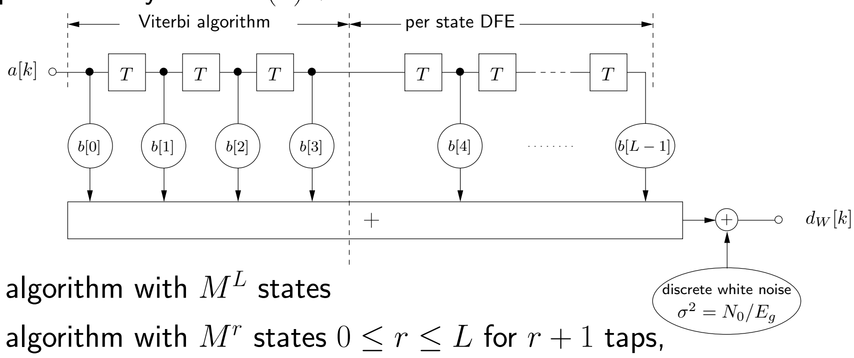
Cable with skin effect: Binary and quaternary signalling – loss to matched filter bound



8.4.4 Reduced–State Sequence Estimation

Decision feedback–sequence estimation (DFSE)

- equivalent structure for discrete–time signal at the output of the whitened matched filter
- minimum phase FIR system: $B(z)+$ white noise



MLSE: Viterbi algorithm with M^L states

DFSE: Viterbi algorithm with M^r states $0 \leq r \leq L$ for $r+1$ taps, ISI of taps $b[r+1]$ to $b[L]$ cancelled by DFE's, using the symbols from the path registers of the different states

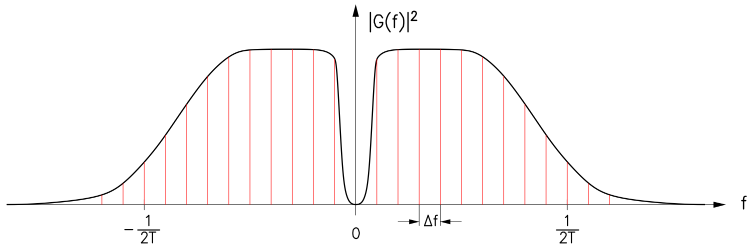
DFE: $r=0$: State reduction to only one state

$\Longrightarrow$ almost continuous tradeoff between performance and complexity

Good news: Often a few states (e.g. $r=1,2$) are sufficient to come very close to MLSE

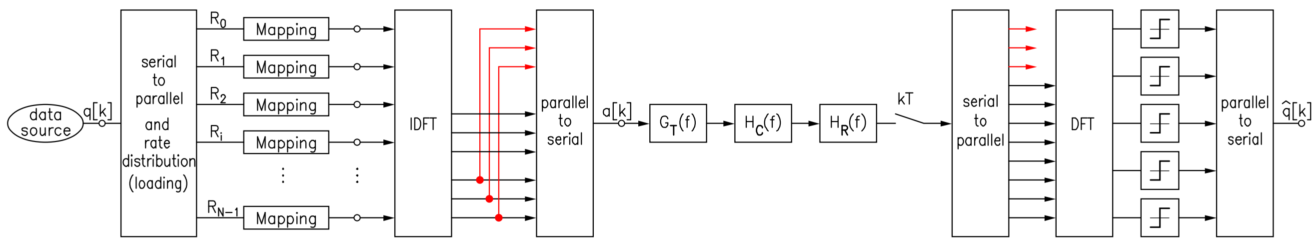
Bad news: Minimum phase filter output pulse indispensible for state reduction. Complex all–pass filter may be necessary.

8.5 Discrete Multitone (DMT or OFDM)



Dividing the spectrum into N subchannels of bandwidth $\Delta f = \dfrac{1}{NT}$

favorable implementation via Discrete Fourier Transform (DFT)



Cyclic prefix of N_{pre} samples (guard interval): Transform of linear convolution into cyclic convolution

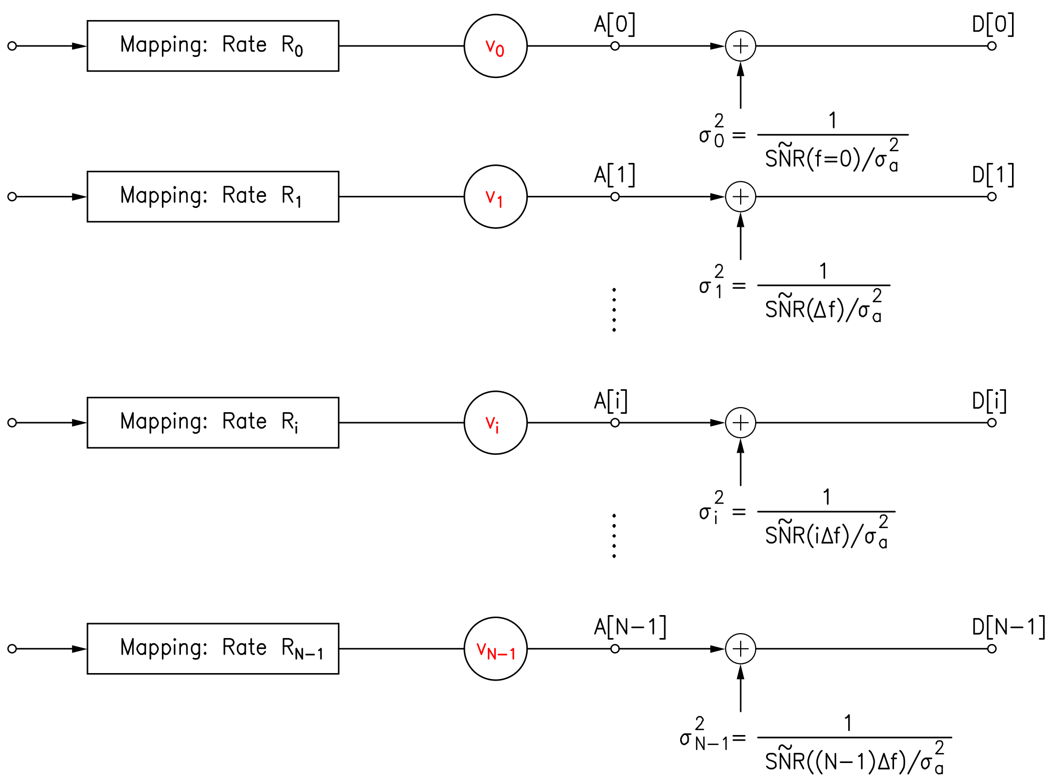
$\Longrightarrow$ Orthogonal nondispersive subchannels along discrete frequency axis

Transmission of frames of $N \cdot R$ bit in time interval $(N + N_{\mathrm{pre}})T$

Energy per frame: $(N + N_{\mathrm{pre}}) \cdot E_g \cdot \sigma_A^2$

Equivalent discrete time block diagram

Including inversion of different subchannel attenuations



Loading of uncoded DMT

$$\mathrm{SNR}_{0,i} = \frac{v_i^2}{\sigma_i^2}$$

Condition $E\left\{|A[i]|^2\right\} =$ constant $= \sigma_A^2$ (Spectral shaping included in $G_T(f)$, i.e., $\widetilde{\mathrm{SNR}}(f)$!)

QAM: $E\left\{|A[i]|^2\right\} \approx \frac{2}{3} \cdot 2^{R_i} \cdot v_i^2 \stackrel{!}{=} \sigma_A^2$ with $R_i =$ Rate per subchannel in $\left[\frac{\text{bit}}{\text{subcarrier}}\right]$

Loading rule for balanced error probabilities:

$\mathrm{SNR}_{0,i} =$ constant $=: \mathrm{SNR}_{0,\mathrm{DMT}}$

$$\mathrm{SNR}_{0,\mathrm{DMT}} = \frac{3}{2}\frac{\sigma_A^2}{\sigma_i^2}\,2^{-R_i} = \frac{3}{2}\frac{\sigma_A^2}{\sigma_a^2}\,\widetilde{\mathrm{SNR}}(i\,\Delta f)\cdot 2^{-R_i}$$

$$\boxed{R_i \stackrel{!}{=} \mathrm{ld}\left(\frac{\widetilde{\mathrm{SNR}}(i\,\Delta f)}{\mathrm{SNR}_{0,\mathrm{DMT}}}\right) + \mathrm{ld}\left(\frac{3}{2}\frac{\sigma_A^2}{\sigma_a^2}\right)}$$

Comparison of loaded DMT and PAM

Simplified Analysis

1. Cyclic prefix (guard interval) ignored
2. Equal overall rate: $\displaystyle\sum_{i=0}^{N-1} R_i \stackrel{!}{=} N \cdot R$
3. Equal signal power: $\sigma_A^2 = \sigma_a^2$, optimum loading

$$\sum_{i=0}^{N-1} R_i = \sum_{i=0}^{N-1} \mathrm{ld}(\widetilde{\mathrm{SNR}}(i\,\Delta f)) - N\,\mathrm{ld}(\mathrm{SNR}_{0,\mathrm{DMT}}) + N\,\mathrm{ld}\left(\frac{3}{2}\right) \stackrel{!}{=} N\,R$$

$$\mathrm{SNR}_{0,\mathrm{DMT}} = 2^{\left[\frac{1}{N}\sum_{i=0}^{N-1}\mathrm{ld}(\widetilde{\mathrm{SNR}}(i\,\Delta f))\right]} \cdot \frac{3}{2}\,2^{-R}$$

geometric mean

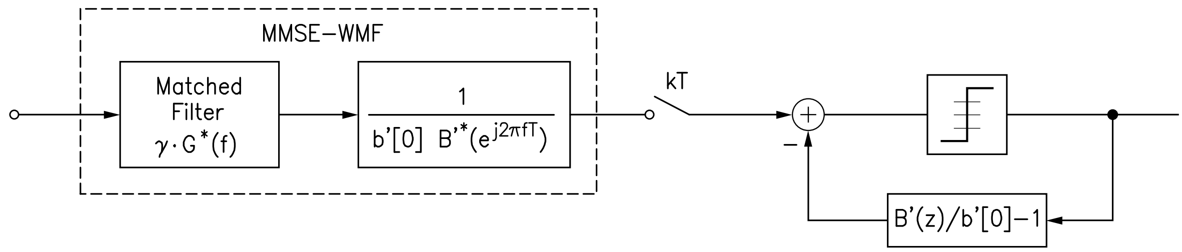
$$N \longrightarrow \infty\ :\ \mathrm{SNR}_{0,DMT} = \exp\left\{T\int_{-1/(2T)}^{+1/(2T)} \ln(\widetilde{\mathrm{SNR}}(f))\,\mathrm{d}f\right\}\cdot\frac{3}{2}\,2^{-R} = \mathrm{SNR}_{0,\mathrm{ZF-DFE}}$$

Identical error performance !

Optimum loaded DMT and ZF-DFE are equivalent !!

8.6 Aspects from Information Theory

Optimum DFE with minimum mean square error whitened matched filter



Spectral factorization of modified PSD $\Phi'_{gg}(z) = B'(z) \cdot B'^*(z^{*-1})$
with $\Phi'_{gg}(f) = \dfrac{1}{E_g T} \displaystyle\sum_{i=-\infty}^{+\infty} |G(f - i/T)|^2 + \dfrac{N_0}{\sigma_a^2 E_g}$

Minimum mean–squared error consisting of noise **and residual ISI**, i.e., toleration of small amounts of ISI for improved noise suppression:

$$\mathrm{SNR}_{\mathrm{MMSE-DFE,}\ \text{unbiased}} = \exp\left[T \int_{-1/(2T)}^{+1/(2T)} \ln(\widetilde{\mathrm{SNR}}(f)\ +\ 1)\,\mathrm{d}f\right] - 1$$

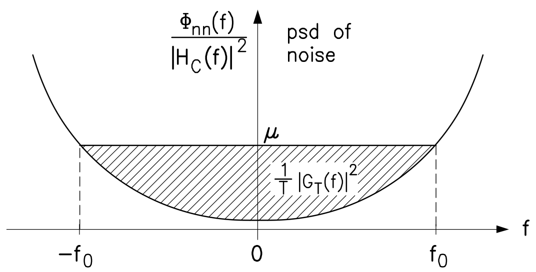
Optimum PAM transmit signal

- Concentration of signal energy on the first Nyquist–interval: $G_T(f) \equiv 0$ for $|f| > \frac{1}{2T} \Longrightarrow$ $\widetilde{\mathrm{SNR}}(f) = \mathrm{SNR}(f)$ for $|f| \le \frac{1}{2T}$

$\Longrightarrow$
$$\boxed{\begin{aligned} \mathrm{SNR}_{\mathrm{MMSE-DFE}} &= 2^{T \cdot I_M} - 1 \\ I_M &:= \int_{-\infty}^{+\infty} \log_2(\mathrm{SNR}(f) + 1)\,\mathrm{d}f \end{aligned}}$$

I_M: Mutual information of the **continuous**–time channel

- Maximum SNR $\longrightarrow$ Maximum mutual information $\Longrightarrow$ Channel capacity
  $\Longrightarrow$ Choose transmit spectrum according to the **water–filling rule**



Condition:
$$\int_{-\infty}^{+\infty} |G_T(f)|^2\,\mathrm{d}f \overset{!}{=} E_S/\sigma_a^2$$
Transmit energy per symbol

Cut–off–frequency from water–filling rule: f_0

$\Rightarrow$ Optimum PAM symbol interval: $\boxed{T_{\mathrm{opt}} = \dfrac{1}{2f_0}}$

$T > \dfrac{1}{2f_0}$: Nyquist interval can not be completely filled: Loss in Euclidean distance
i.e., some sequences are hardly distinguishable!

$T < \dfrac{1}{2f_0}$: Waste of signal power outside the first Nyquist–interval

$\Longrightarrow$ Optimum PAM rate for transmission of data rate $\dfrac{1}{T_b}$

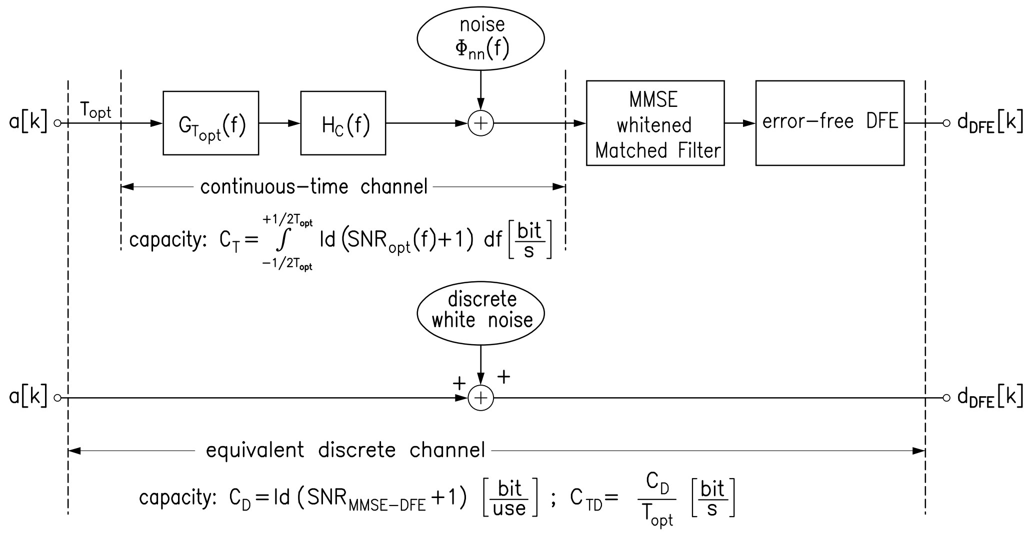
$$\boxed{R = \frac{T_{\mathrm{opt}}}{T_b} \quad \left[\frac{\text{bit}}{\text{symbol}}\right]}$$

$$\boxed{\mathrm{SNR}_{\mathrm{MMSE-DFE}} = 2^{T_{\mathrm{opt}} \cdot C_T} - 1} \qquad \text{with } C_T = \int_{-1/(2T_{\mathrm{opt}})}^{+1/(2T_{\mathrm{opt}})} \mathrm{ld}(\mathrm{SNR}_{\mathrm{opt}}(f) + 1)\,\mathrm{d}f \quad \left[\frac{\text{bit}}{\text{s}}\right]$$

C_T : Continuous–time capacity

Continuous and equivalent discrete–time block diagram

for MMSE–DFE with genius feedback



$$\mathrm{SNR}_{\mathrm{MMSE-DFE}} = 2^{T_{\mathrm{opt}}\, C_T} - 1 \qquad \Longrightarrow \qquad \boxed{C_{TD} = C_T}$$

DFE is sufficient to achieve capacity

MMSE–DFE: Canonical equalization

- Suboptimum MMSE decision–feedback equalization is sufficient to achieve channel capacity!
- In principle, intersymbol interference (ISI) and channel code can be decoded separately.
  A super–decoder is not necessary for optimality!
- Channel codes that well suited for the non–dispersive AWGN–channel, are also well suited for linear dispersive channels. Special codes designed for the dispersive channel are not necessary.
- Information Theory provides optimum rate R_{opt} and symbol interval T_{opt}. $\Longrightarrow$ Favourable transmission of code's redundancy by PAM constellation expansion. About 0.5 bit of redundancy per (real) dimension is sufficient as long as signal shaping is not applied.

  M–ary QAM: Choose $\quad 2^{R_{\text{opt}}+1} \leq M < 2^{R_{\text{opt}}+2}$

  $\Longrightarrow$ Optimum system design from Information Theory

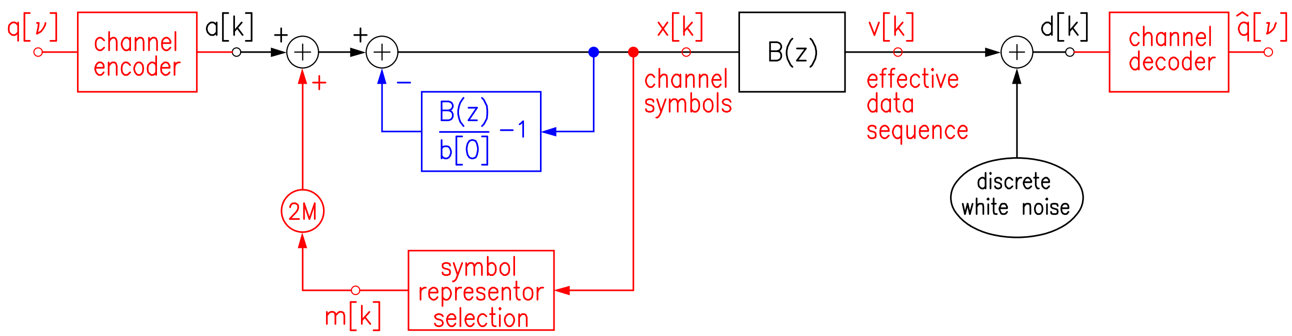
8.7 Precoding

Problems with DFE in practice:

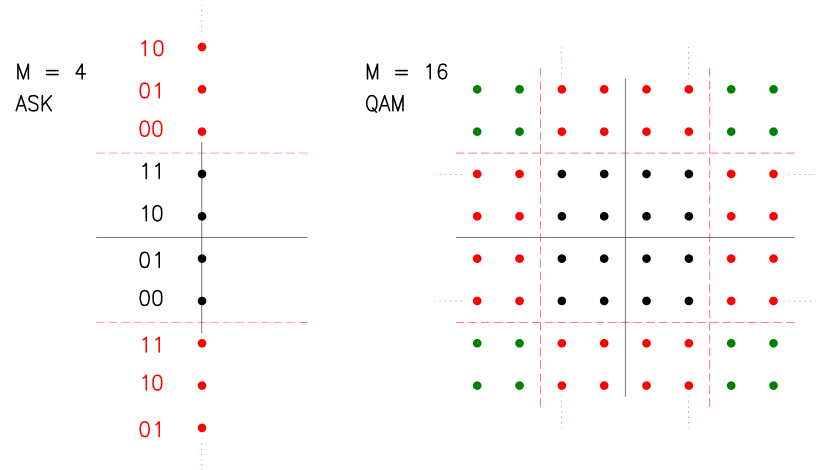
- Error propagation by feedback
- No decision delay tolerable because immediate feedback is essential
  $\Longrightarrow$ DFE incompatible with channel coding!!

Solution: Equivalent structure with Tomlinson–Harashima precoding

Discrete–time model for ZF whitened–matched filter output



Multiple representation of symbols



Representation of symbols a_m by **sets of symbols**

ASK: $\mathcal{A} = \{a_m + m[k] \cdot 2M \,|\, m[k] \in \mathbb{Z}\}$; $m[k]$: actual selection

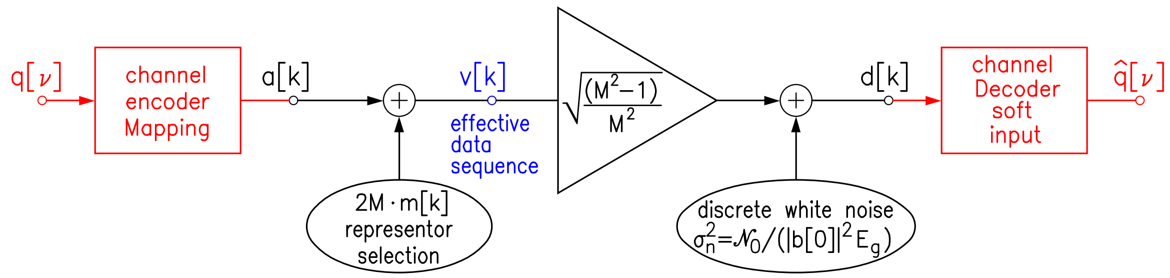
QAM: $\mathcal{A} = \{a_m + m_I[k] \cdot 2M + jm_Q[k] \cdot 2M \,|\, m_I[k], m_Q[k] \in \mathbb{Z}\}$

Select $m[k]$, $(m_I[k],\, m_Q[k])$ such that channel symbol $x[k]$ has minimum magnitude

ASK: Channel symbols $x[k]$ are continuous and equally distributed in interval $[-M, +M)$

Tomlinson–Harashima Precoding (THP)

Equivalent discrete–time model



AWGN channel with periodic continuation of signal constellation

Equivalent $\mathrm{SNR}_{\mathrm{TOM}} = \dfrac{\sigma_a^2 E_g}{N_0} |b[0]|^2 \cdot \dfrac{M^2-1}{M^2} = \dfrac{M^2-1}{M^2} \mathrm{SNR}_{\mathrm{ZF-DFE}}$

$(M^2-1)/M^2$: THP loss: Discrete vs. continuous distribution of signal points

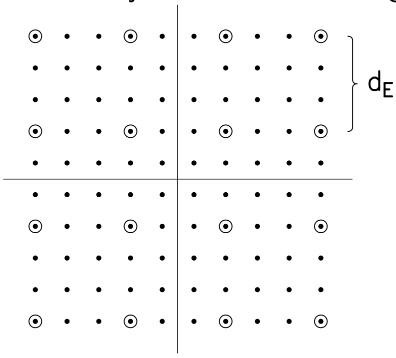
ASK $M = 2$: 1.23 dB; $M = 4$: 0.28 dB; $M = 8$: 0.07 dB

THP is equivalent to genie-aided ZF-DFE

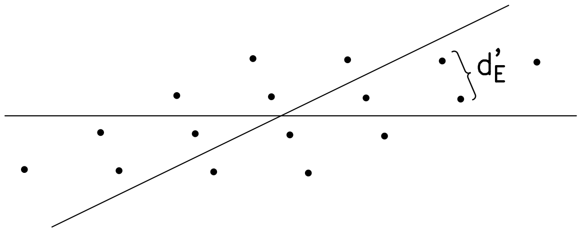
Geometric interpretation of THP or trellis precoding— signal space (1)

1. Without Precoding

Transmitter side: Orthogonal signalling per symbol interval, thinning of constellation by channel coding



Receiver side: Distortion by the dispersive system: Loss of orthogonality, i.e.,
$\exists i \in \mathbb{Z} \backslash \{0\}$
with $\displaystyle\int_{-\infty}^{+\infty} g(t) \cdot g^*(t - iT)\, \mathrm{d}t \neq 0$



Loss in minimum Euclidean distance
"decoding" indispensible due to intersymbol interdependencies

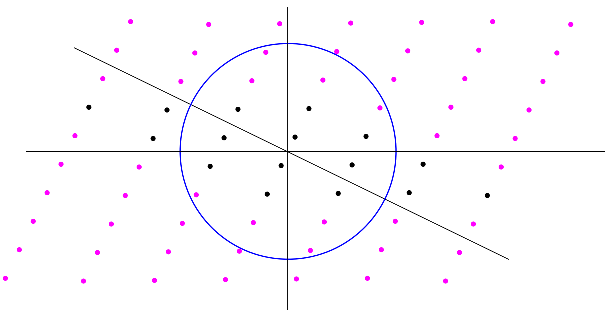
Geometric interpretation of THP or trellis precoding— signal space (2)

2. With Precoding

Transmitter side: Pre–equalization

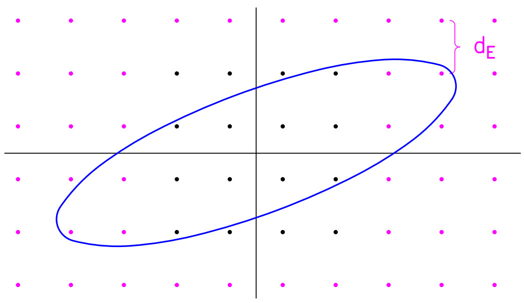
precoding with shaping

multiple representation of signal points



Enhancement of signal power

Receiver side: Re–orthogonalization by the dispersive channel



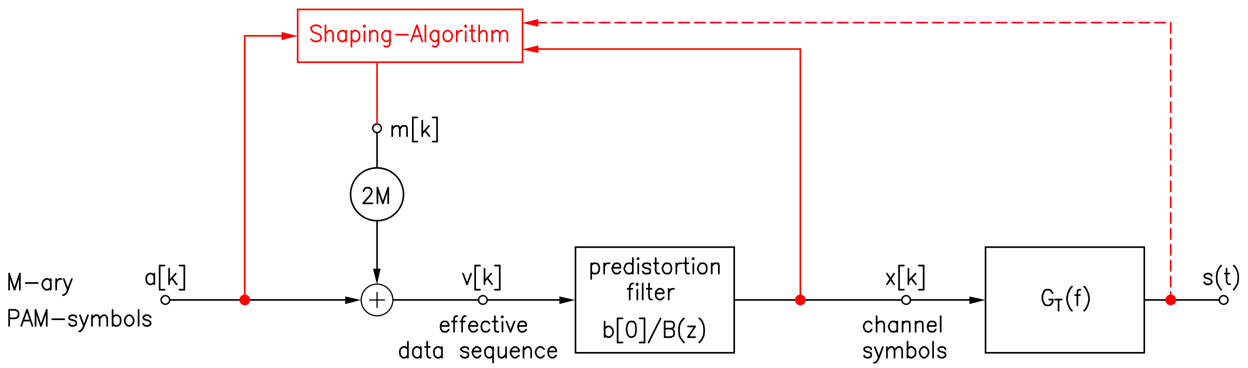
Euclidean distances reconstructed
boundary of received constellation

Advantages of Tomlinson–Harashima precoding and shaping

- Equivalent to genie-aided ZF-DFE
- Original distances between sequences of PAM–symbols are reconstructed
- Channel coding possible in the same way as for a non-distorting channel
- Max. signal amplitude (THP) or average signal power (shaping) minimized

Precoding + signal shaping

Equivalent structure of Tomlinson–Harashima precoding:



- **THP:** Independent choice of $m[k]$ for min. $|x[k]|$ in each modulation interval (dimension)
- **Precoding and shaping:** Choose **sequence** $m[k]$ such that some **shaping criteria** for the transmitter's output signal are optimized.

Signal–Shaping Criteria

- Minimum average transmitter output power, i.e., $\min E\left\{|x[k]|^2\right\}$
- Forcing a zero at DC, i.e., $\Phi_{xx}(z=1)=0$
- Reduction of the Peak–to–Average Power Ratio (PAR), $\max_t |s(t)|^2 \Big/ E\left\{|s(t)|^2\right\}$, of the continuous–time transmitter output signal
- Mixed criteria
- Extension of the PAM constellation from M to $M \cdot 2^{R_s}$ points
  $\Longrightarrow$ On average R_s free shaping bits $s[k]$ available per step for signal shaping
- THP for extended constellation
- Calculation of the optimum **sequence** $s[k]$ of shaping bits via Viterbi algorithm on a trellis of an (imaginary) scrambler for the binary sequences $s[k]$
  - Metric definition with regard to shaping criteria
  - ISI due to predistortion filter taken into account by DFSE at the receiver

Use of predistortion filter $1/B(z)$ for mixing and dispersion of shaping and data bits
$\Longrightarrow$ No extra scrambling of shaping bits with data bits (cf. trellis precoding)
$\Longrightarrow$ No inversion of shaping necessary at the receiver side (cf. trellis precoding)
$\Longrightarrow$ No error propagation due to a de–shaping (cf. trellis precoding)

Shaping for minimum average signal power $\min E\left\{|x[k]|^2\right\}$:

- Inclusion of sequence of channel symbols $x[k]$ within a multi–dimensional hypersphere (instead of inclusion into a hypercube for THP)
  $\Longrightarrow$ Gaussian distributed channel symbols per dimension:
  Capacity achieving distribution for the AWGN channel

- Shaping gains of up to 1,53 dB achievable in principle, up to 1 dB in practice
- Shaping gain achievable with much fewer states of a Viterbi decoder than what would be required for an additional coding gain of 1 dB via trellis coded modulation

Disadvantage of Tomlinson–Harashima Precoding and Shaping without Scrambling

- Extreme high dynamics of the "effective data sequence".
- The signal observeable after the receiver input filter and T–spaced sampling at the **receiver side** is given by

$$v[k] = a[k] + m[k] \cdot 2M$$

$$V_{\max} = \underbrace{\max_k v[k] \approx 2 \cdot M \cdot}_{\text{for additional shaping}} \sum_{i=0}^{L} \left| b[i] \big/ b[0] \right| + 1 \quad \gg \quad M - 1$$

- Extreme precision in equalization necessary (very large pulses interfere with very small pulses!)
- High sensitivity to symbol timing errors and jitter
- Approximately discrete Gaussian distribution of the effective data sequence

  $\Longrightarrow$ Blind methods for linear fine tuning equalization at the receiver side are not applicable!

  Solution: Dynamics Limited Precoding and Dynamics Shaping (Fischer, Huber 1995)

Dynamics Limited Precoding

$=$ THP, but with restricted sets of PAM symbols representations, limited to V_L

$$|a_m + 2M \cdot m[k]| \leq V_L$$

Tradeoff between dynamics V_L at the receiver side and average signal power at the transmitter side. Extreme choices:

- $V_L = M - 1$: Linear predistortion (equivalent to LZFE)
- $V_L \geq V_{\max}$: original THP

Dynamics Shaping

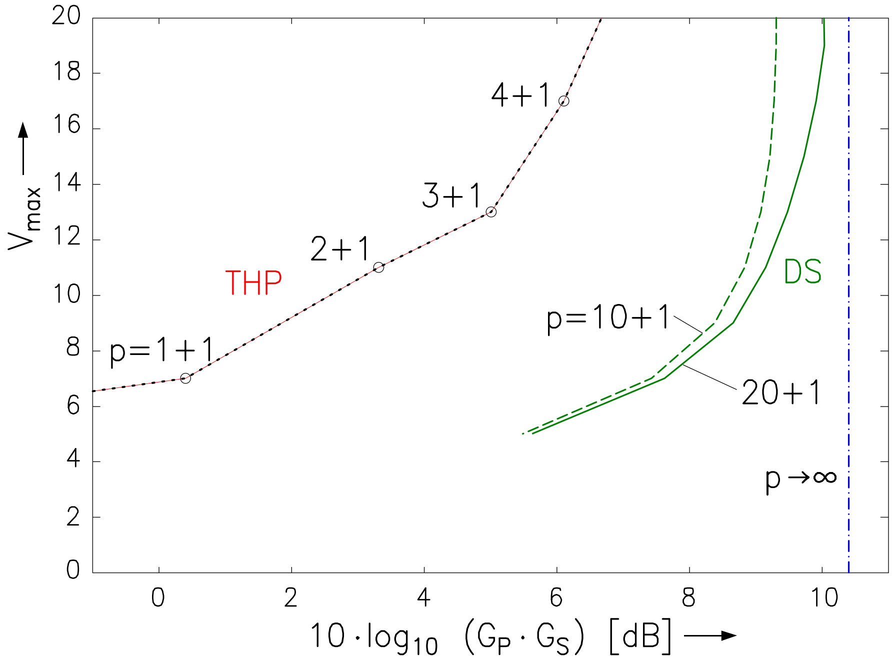
$=$ Shaping without scrambling, but with restricted sets of of PAM symbols representations, limited to V_L
Continuous tradeoff between shaping gain and reduction of dynamics at the receiver side possible

- Various shaping criteria applicable by proper choice of metric: Average transmit power, PAR at transmitter and receiver side
- Blind methods for linear fine tuning equalization at the receiver side now are applicable!
- Blind equalization methods for precoded signalling (Gerstacker, Huber 1997)

Example: HDSL I

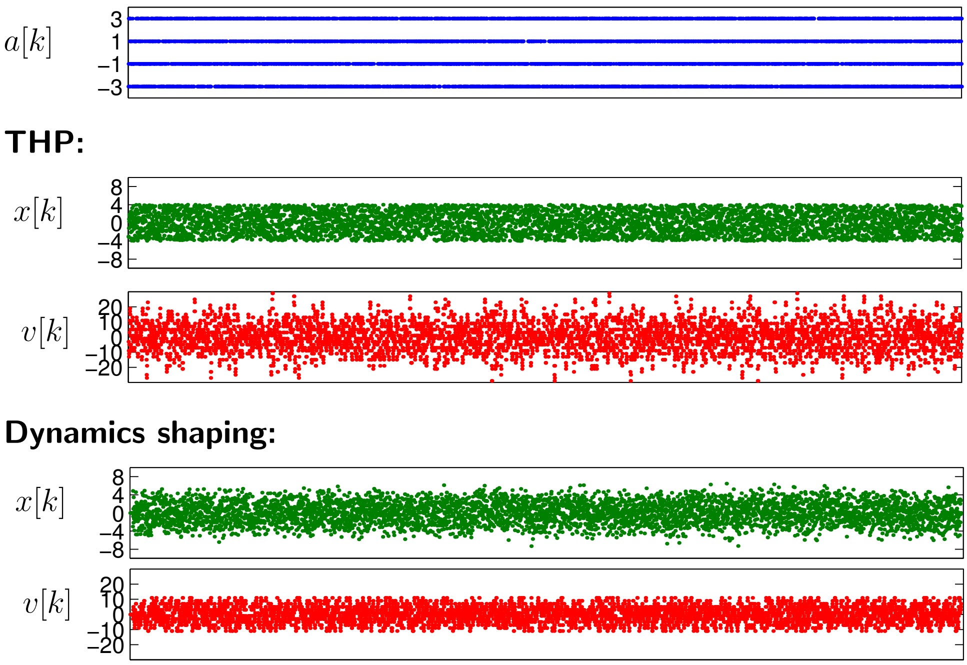
4-ary signalling over twisted pair lines

Gain due to noise prediction, G_p, times gain due to shaping, G_S. Both gains are with respect to LZFE and are shown as functions of the dynamics limitation $V_{\max}$ of the effective data sequence



Simulation of symbol sequences

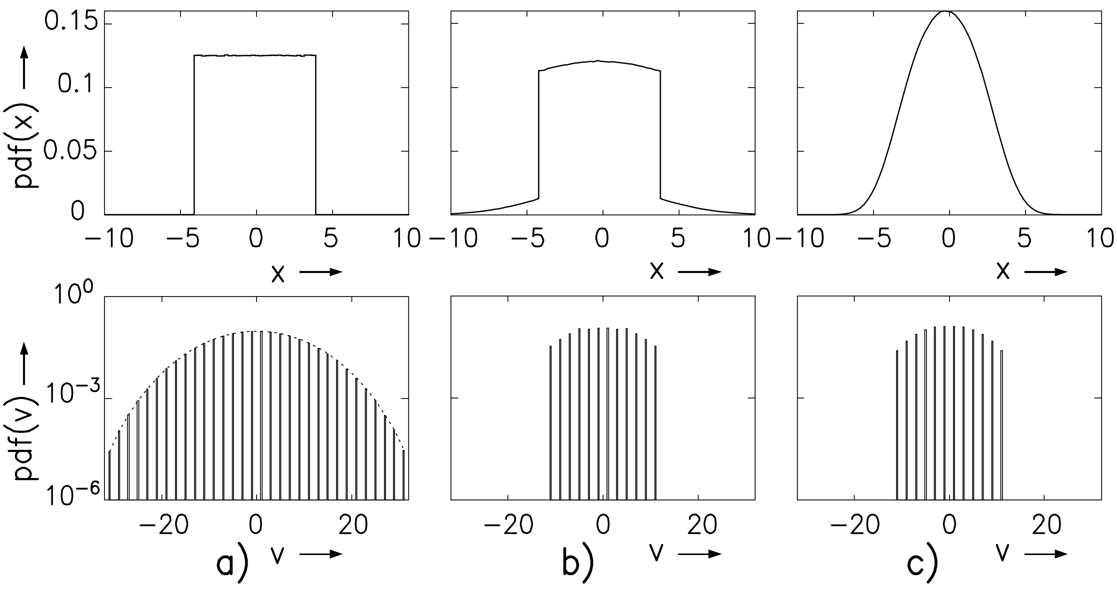
Symbol sequence:



Pdfs of transmitter output signal and pdfs of effective data sequence

transmitter output

receiver input (effective data sequence)



THP | Dyn. Lim. Precoding $V_{\max} = 12$ | Dynamics Shaping $V_{\max} = 12$

Conclusions

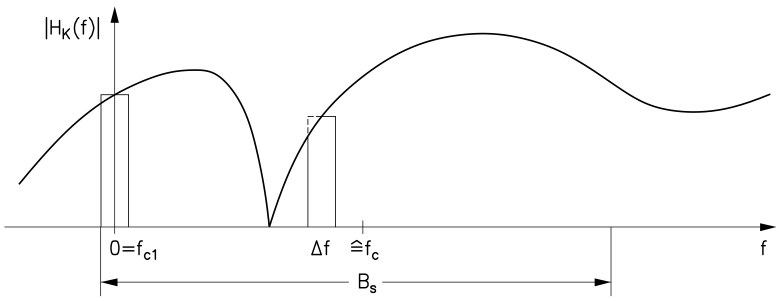
- Theory of ISI and equalization is very rich and includes many symmetries and equivalences

  - Relationships between optimum receiver input filters
  - Arithmetic, harmonic, geometric mean over folded $\widetilde{\mathrm{SNR}}(f)$
  - Equivalence of noise prediction and DFE
  - Duality of DFE and loaded DMT
  - DFE and channel capacity, system design based on information theoretical principles
  - Optimum system design based on information theory
  - Duality of DFE and THP, transformation into equivalent AWGN channel
  - Capacity achievable with codes designed for AWGN channel

9 Multi–Carrier Schemes

Idea: Digital transmission over frequency selective, i.e., dispersive and noisy channels using the division of the frequency spectrum into multiple bands.
Individual digital PAM–transmission schemes within each partial frequency band:

Illustration:



Terms:

- Orthogonal Frequency Division Multiplexing (OFDM)
- Discrete Multitone (DMT)
- Multicarrier Modulation (MCM)

Fields of application: (Examples)

- Digital Audio Broadcasting (DAB): digital radio
- Digital Video Broadcasting (DVB): digital television
- Digital Subscriber Lines (ADSL): fast digital transmission over metallic twisted pairs
- Local Area Network (WLAN): IEEE 802.11
- Local Wide Area Network (WiMAX): IEEE 802.16
- Mobile Communication: LTE, LTE–advanced

Principle:

Decomposition of the available frequency band with bandwidth B_s into D subchannels (so–called *subcarriers*) with

bandwidth	: $\Delta f = B_\mathrm{s}/D$
(min.) symbol duration	: $T_\mathrm{s} = 1/\Delta f$
carrier frequencies	: $f_{ci} = (i-1)\,\Delta f\;,\; i = 1, \ldots,\; D$

For $D \gg 1$, there are D independent *non-frequency–selective* AWGN channels with bandwidth Δf and weighting with $H_K(f = f_{ci})$, i.e., attenuation

$$\alpha_i = -20\log_{10}(|H_K((i-1)\,\Delta f)|)$$

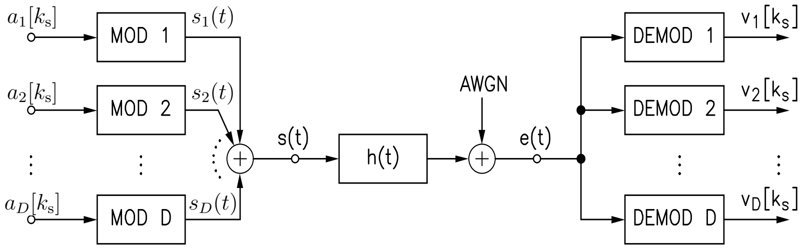
over which a *distortion–free* transmission is possible.

Note:

To maintain a consistent illustration, the lower boundary, rather than the middle, of the frequency band is used as the transformation frequency for the ECB transformation.

9.1 OFDM/DMT

Basic principle:



Transmit signal (ECB representation):

$$s(t) = \sum_{i=1}^{D} s_i(t)$$

with D identical PAM transmission schemes (pulse $g(t)$):

$$\begin{aligned} s_i(t) &= \left(\sum_{k_\mathrm{s}} a_i[k_\mathrm{s}] \cdot g(t - k_\mathrm{s} T_\mathrm{s}) \right) \cdot \mathrm{e}^{\mathrm{j}2\pi(i-1)\Delta f t} \\ &= \sum_{k_\mathrm{s}} a_i[k_\mathrm{s}] \cdot g(t - k_\mathrm{s} T_\mathrm{s}) \cdot \mathrm{e}^{\mathrm{j}2\pi((i-1)t - k_\mathrm{s} T_\mathrm{s})/T_\mathrm{s}} \end{aligned}$$

$$= \sum_{k_\mathrm{s}} a_i[k_\mathrm{s}] \cdot g_i(t - k_\mathrm{s} T_\mathrm{s})$$

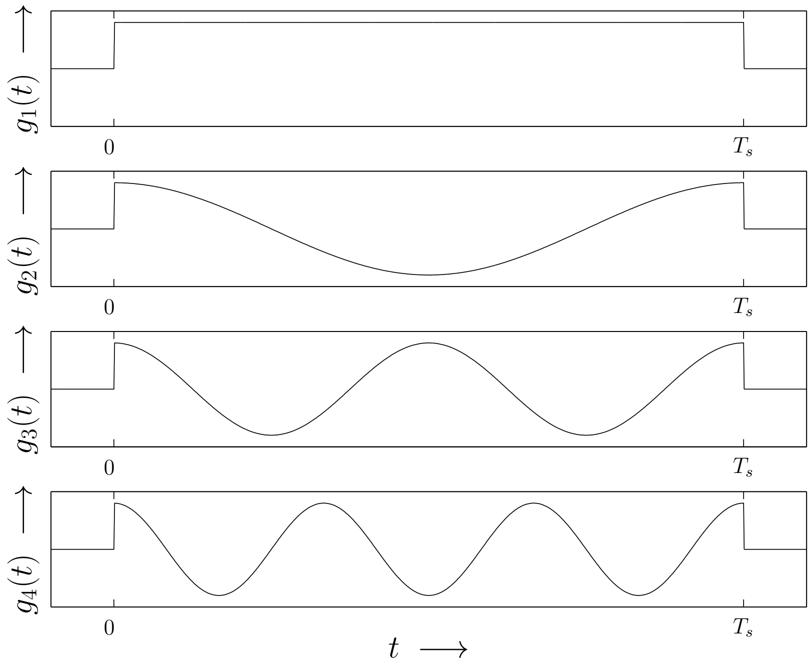
Usually, the fundamental pulses $g_i(t)$ are be thought of as *chips* $g_\mathrm{T}(t)$, weighted with the coefficients $b_i[\kappa]$ ($T = T_\mathrm{s}/D$):

$$g_i(t) = \sum_{\kappa=0}^{D-1} b_i[\kappa] \cdot g_\mathrm{T}(t - \kappa T)$$

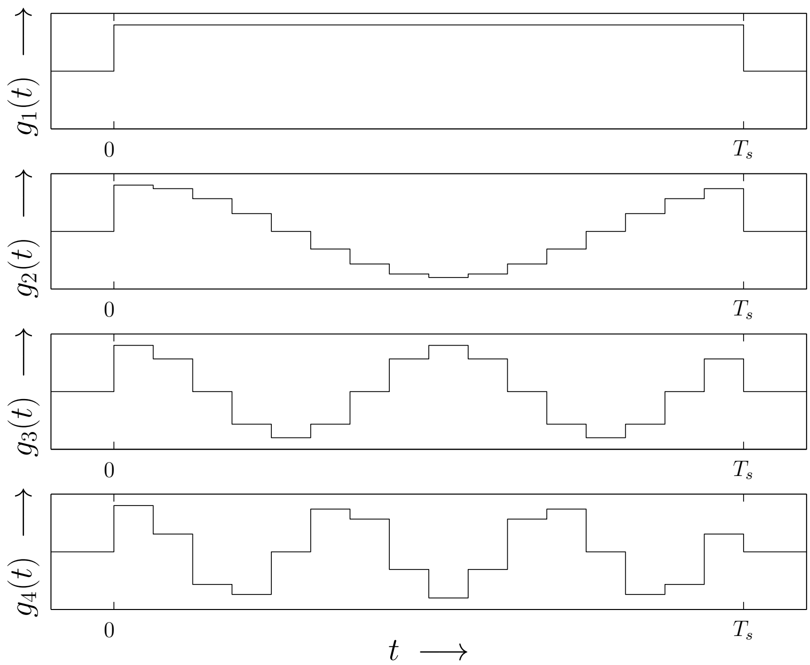
Example:

(real signals)

Theory:



Practice: $D = 16$ (here: $g_{\mathrm{T}}(t)$ is rectangular in the time domain)



Approach:

- The chip pulse shape $g_{\mathrm{T}}(t)$, channel transfer function, receiver filter $H_{\mathrm{R}}(f)$, and sampling at rate $1/T$ may be combined into an equivalent *discrete–time channel model*.

  Desired transfer function: $H(z)$ of order p
  PSD of the additive noise: $\Phi_{nn}(\mathrm{e}^{\mathrm{j}2\pi fT})$

- From a *block* of D data symbols, a block for D channel symbols is computed at the transmitter

$$
\begin{aligned}
s(t) &= \sum_{i=1}^{D} \sum_{k_{\mathrm{s}}} a_i[k_{\mathrm{s}}] \cdot \sum_{\kappa=0}^{D-1} b_i[\kappa] \cdot g_{\mathrm{T}}(t - \kappa T - k_{\mathrm{s}} T_{\mathrm{s}}) \\
&= \sum_{k_{\mathrm{s}}} \sum_{\kappa=0}^{D-1} \underbrace{\sum_{i=1}^{D} a_i[k_{\mathrm{s}}] \cdot b_i[\kappa]}_{\stackrel{\mathrm{def}}{=} x_{\kappa+1}[k_{\mathrm{s}}]} \cdot g_{\mathrm{T}}(t - \kappa T - k_{\mathrm{s}} T_{\mathrm{s}})
\end{aligned}
$$

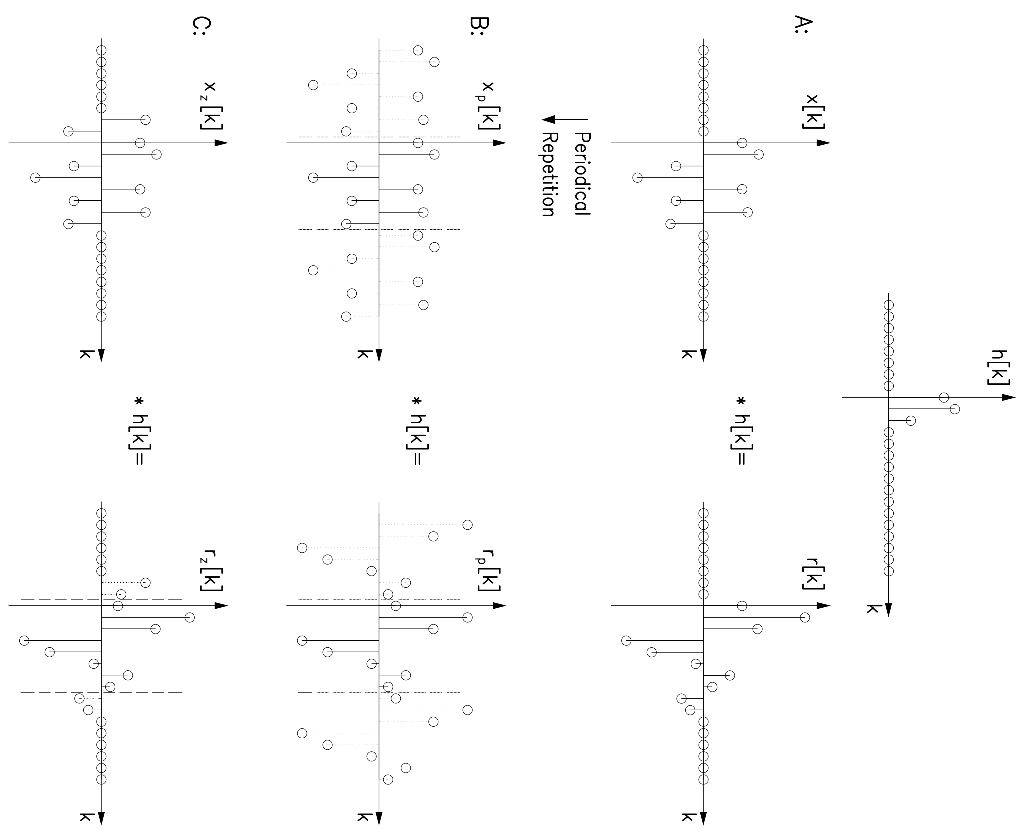
  $\Longrightarrow$ Block transmission over an ISI–Channel

  At the receiver side, values of the amplitude coefficients $a_i[k_{\mathrm{s}}]$ are estimated from the received blocks.

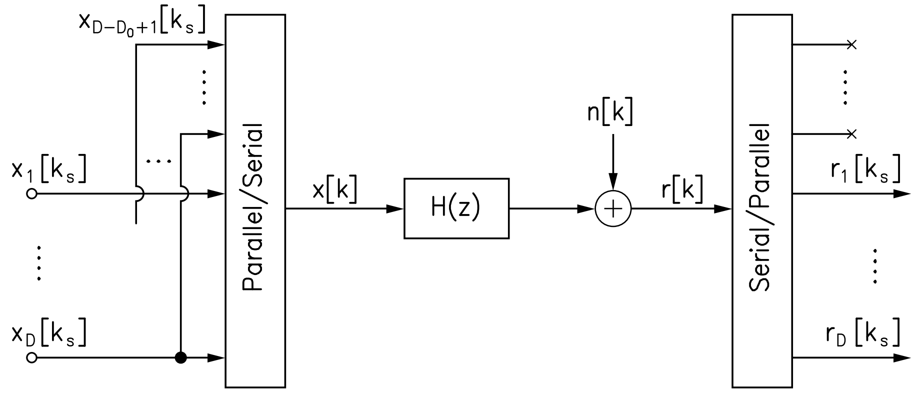
- Desired: Neighboring blocks should be independent from each other at the detector— no inter–*block* interference

- Possible approaches:
  - Insertion of p zeros between blocks
    – No waste of signal power
    – Pre– and post–processing is channel–dependent
  - Partial cyclic repetition of the blocks
    – Small increase in signal power
    – Pre– and post–processing is not affected by the actual channel
  - Insertion of a pre–determined word (unique word OFDM)

Linear Convolution — Cyclic Convolution:



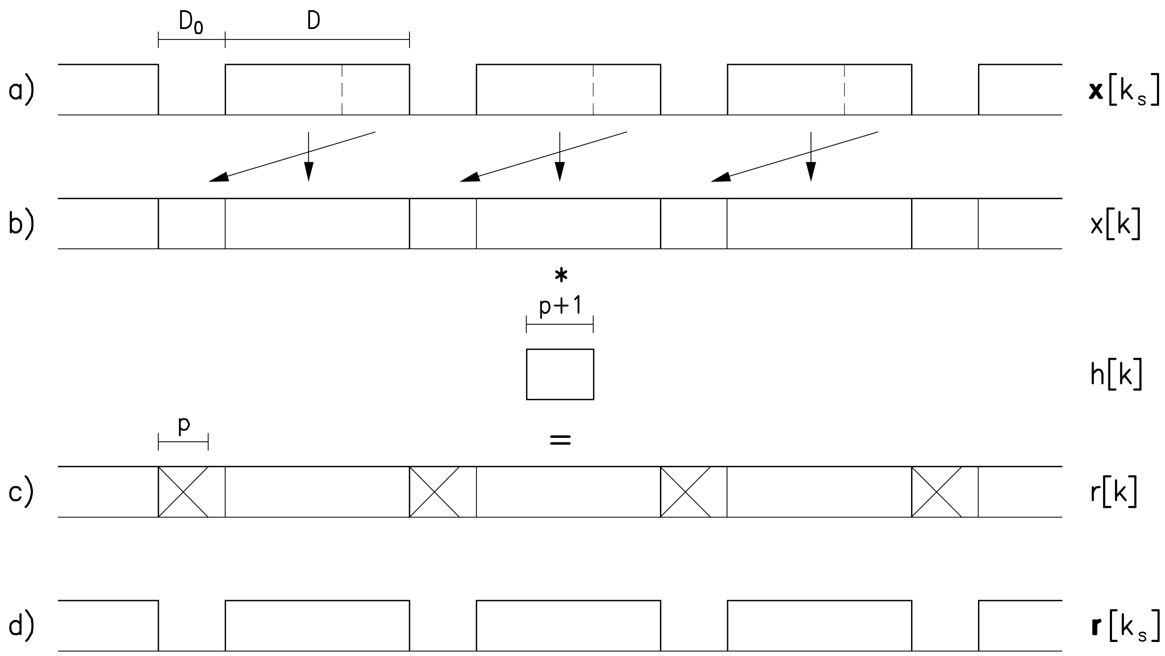
Approach: Cyclic extension by D_0 ($D_0 \geq p$) symbols
$\Longrightarrow$ *Guard intervall*



Symbol clock per OFDM frame (timing of the DMT/OFDM symbols) T_s (time index k_s);

Timing of the channel symbols (chip timing) T (time index k) with $T_s = (D + D_0) \cdot T$

Ongoing transmission:



For a block of length D

$$\mathbf{x}[k_s] = [\, x_1[k_s], \, \ldots \, , x_D[k_s] \,]$$

and a cyclically extended block of length $D + D_0$

$$\mathbf{x}_z[k_s] = [\, x_{D-D_0+1}[k_s], \, \ldots \, , x_D[k_s], \, x_1[k_s], \, \ldots \, , x_D[k_s] \,]$$

the following holds

$$\begin{aligned}
\mathbf{x}_z[k_s] &= \mathbf{x}[k_s] \cdot \left[\begin{array}{c|c} \mathbf{0}_{(D-D_0)\times D_0} & \mathbf{I}_D \\ \hline \mathbf{I}_{D_0} & \end{array} \right] \\
&= \mathbf{x}[k_s] \cdot \left[\begin{array}{ccc|cccc} & \mathbf{0} & & 1 & & & \\ \hline 1 & & & & \ddots & & \\ & \ddots & & & & \ddots & \\ & & 1 & & & & 1 \end{array} \right] .
\end{aligned}$$

Transmission over the channel:

$$\mathbf{r}_z[k_s] = \mathbf{x}_z[k_s] \cdot \mathbf{H}$$

with channel matrix

$$\mathbf{H} := \begin{bmatrix} h[0] & h[1] & \cdots & h[p] & & \mathbf{0} \\ & \ddots & \ddots & & \ddots & \\ & & \ddots & \ddots & & h[p] \\ & & & \ddots & \ddots & \vdots \\ & \mathbf{0} & & & \ddots & h[1] \\ & & & & & h[0] \end{bmatrix}$$

The removal of the D symbols $r_i[k_s]$, $i = 1, 2, \ldots, D$, and allocation to the vector

$$\mathbf{r}[k_s] = [\, r_1[k_s], \, \ldots \, , r_D[k_s] \,]$$

yields

$$\mathbf{r}[k_s] = \mathbf{r}_z[k_s] \cdot \left[\begin{array}{c} \mathbf{0}_{D_0 \times D} \\ \hline \mathbf{I}_D \end{array} \right] = \mathbf{r}_z[k_s] \cdot \left[\begin{array}{ccc} & \mathbf{0} & \\ \hline 1 & & \\ & \ddots & \\ & & 1 \end{array} \right]$$

Combining everything, we obtain

$$\mathbf{r}[k_s] = \mathbf{x}[k_s] \cdot \mathbf{H}_z$$

with *cyclic matrix*

$$\mathbf{H}_z \overset{\text{def}}{=} \left[\begin{array}{ccc|ccc} \multicolumn{3}{c|}{\mathbf{0}} & 1 & & \\ \cline{1-3} 1 & & & & \ddots & \\ & \ddots & & & & \ddots \\ & & 1 & & & 1 \end{array} \right] \cdot \begin{bmatrix} h[0] & h[1] & \cdots & h[p] & & \mathbf{0} \\ & \ddots & \ddots & & \ddots & \\ & & \ddots & \ddots & & h[p] \\ & & & \ddots & \ddots & \vdots \\ & \mathbf{0} & & & \ddots & h[1] \\ & & & & & h[0] \end{bmatrix} \cdot \left[\begin{array}{ccc} & \mathbf{0} & \\ \hline 1 & & \\ & \ddots & \\ & & 1 \end{array} \right]$$

$$= \begin{bmatrix} h[0] & h[1] & \cdots & h[p] & & \mathbf{0} \\ & \ddots & \ddots & & \ddots & \\ & & \ddots & \ddots & & h[p] \\ h[p] & & \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & & & \ddots & h[1] \\ h[1] & \cdots & h[p] & & & h[0] \end{bmatrix}$$

Independent of $\langle h[k] \rangle$, the *DFT matrix*

$$\mathbf{W}_D \overset{\text{def}}{=} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}} & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}2} & \cdots & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}(D-1)} \\ 1 & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}2} & & & \vdots \\ \vdots & \vdots & & & \vdots \\ 1 & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}(D-1)} & \cdots & \cdots & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}(D-1)^2} \end{bmatrix}$$

is the modal matrix of $\mathbf{H}_z$. Thus, the following holds

$$\sqrt{D}\mathbf{W}_D^{-1} \cdot \mathbf{H}_z \cdot 1/\sqrt{D}\mathbf{W}_D = \mathbf{diag}\left(\lambda_1, \ldots, \lambda_D\right) = 1/\sqrt{D}\mathbf{W}_D \cdot \mathbf{H}_z \cdot \sqrt{D}\mathbf{W}_D^{-1}$$

Note: $\dfrac{1}{\sqrt{D}} \cdot \mathbf{W}_D$ corresponds to DFT
$\sqrt{D} \cdot \mathbf{W}_D^{-1}$ corresponds to IDFT

Transmission scheme:

- Transmitter: Matrix $\sqrt{D} \cdot \mathbf{W}_D^{-1}$
  (Due to the factor $\sqrt{D}$ the transformation is power preserving)
  From the data symbols $a_i[k_\mathrm{s}]$, $i = 1, \ldots, D$, the channel symbols

  $$x_k[k_\mathrm{s}] = \sqrt{D} \cdot \sum_{i=0}^{D-1} a_{i+1}[k_\mathrm{s}] \cdot \mathrm{e}^{+\mathrm{j}\frac{2\pi}{D}i(k-1)}$$

  $k = 1, \ldots, D$, are calculated.

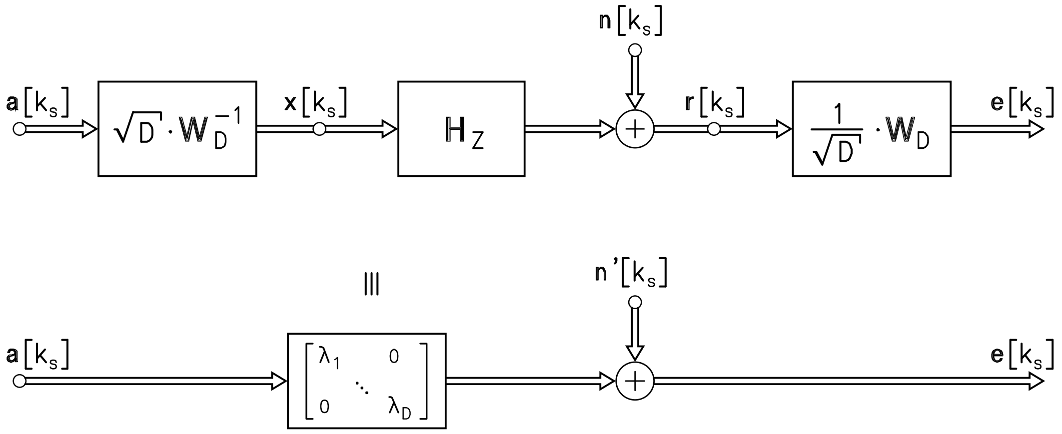
- Receiver: Matrix $1/\sqrt{D} \cdot \mathbf{W}_D$
  From the received symbols $r_i[k_\mathrm{s}]$, $i = 1, \ldots, D$, the symbols for detection,

  $$e_i[k_\mathrm{s}] = \frac{1}{\sqrt{D}} \cdot \sum_{k=0}^{D-1} r_{k+1}[k_\mathrm{s}] \cdot \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}(i-1)k}$$

  $i = 1, \ldots, D$, are calculated.

- Transmission: D parallel, independent subchannels (AWGN) with gain factors λ_i, $i = 1, \ldots, D$ (eigenvalues of the matrix $\mathbf{H}_z$) with

$$\lambda_i = \sum_{k=0}^{D-1} h[k] \cdot \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}(i-1)k} = H(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})$$



Discussion:

- The relationship between signals $\mathbf{a}[k_s]$ and $\mathbf{x}[k_s]$ corresponds to an IDFT — $\mathbf{a}[k_s]$ can be interpreted as the spectrum of $\mathbf{x}[k_s]$

- Specification of desired symbols in the frequency domain — formed using the transmit signals and their inverse Fourier transform

- The subchannels for DMT/OFDM are therefore known as *carriers*

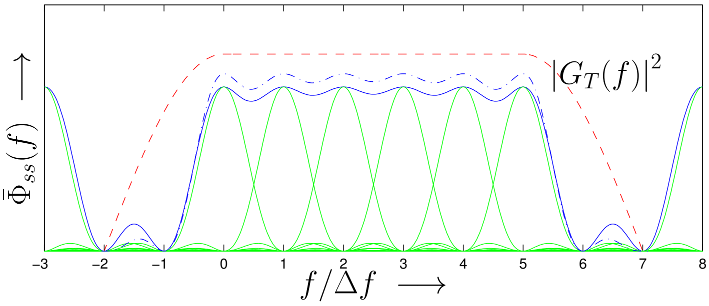
- Average PSD of the continuous–time DMT signal $s(t)$, as introduced by the DFT of length D and the guard interval of length D_0 ($T = T_s/(D + D_0)$)

$$\bar{\Phi}_{ss}(f) = \frac{T_s}{D} \sum_{\mu=-\infty}^{+\infty} \sigma^2_{a_{(\mu \bmod D)+1}} \cdot \left(\frac{\sin \pi(fT_s - \mu(D+D_0)/D)}{\pi(fT_s - \mu(D+D_0)/D)}\right)^2 \cdot |G_T(f)|^2$$

  - The individual carriers are weighted by $\sin(x)/x$ – function due to rectangular block structure ($D + D_0$ values) in the time domain
  - Power per sub–carrier corresponds to the variance of the data symbols in the subchannel under consideration
  - Weightening of the entire signal with $|G_T(f)|^2$

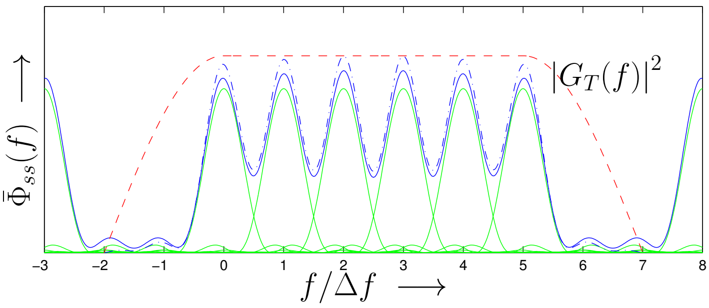
Example: $D = 8$; $D_0 = 0$;

$\sigma^2_{a_i} = \mathrm{const.}$, $i = 1, \ldots, 6$; $\sigma^2_{a_i} = 0$, $i = 7, 8$



Example: $D = 8$; $D_0 = 2$;

$\sigma^2_{a_i} = \mathrm{const.}$, $i = 1, \ldots, 6$; $\sigma^2_{a_i} = 0$, $i = 7, 8$



- As long as

$$\varphi_{nn}[\kappa] = 0\,, \qquad |\kappa| \geq D$$

is valid for the ACF of the noise, the noise power in the i-th subchannel is given by

$$\begin{aligned}\sigma^2_{n'_i} &= \sum_{\kappa=-D+1}^{D-1} \left(1 - \frac{|\kappa|}{D}\right) \varphi_{nn}[\kappa] \cdot \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}\kappa(i-1)} \\ &\approx \sum_{\kappa=-\infty}^{+\infty} \varphi_{nn}[\kappa] \cdot \mathrm{e}^{-\mathrm{j}\frac{2\pi}{D}\kappa(i-1)} \\ &= \Phi_{nn}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})\end{aligned}$$

- If the DFT yields a real–valued signal (baseband transmission), then the Hermitian symmetry conditions must be satisfied

$$a_i[k_s] \;=\; a^*_{D-i+2}[k_s]\,, \qquad i = 2, \ldots, D/2$$

and

$$a_1[k_s] \qquad \text{and} \qquad a_{D/2+1}[k_s] \qquad \text{real.}$$

=> $D/2 - 1$ complex subchannels ($i = 2, \ldots, D/2$) and two real channels ($i = 1$ and $i = D/2 + 1$)

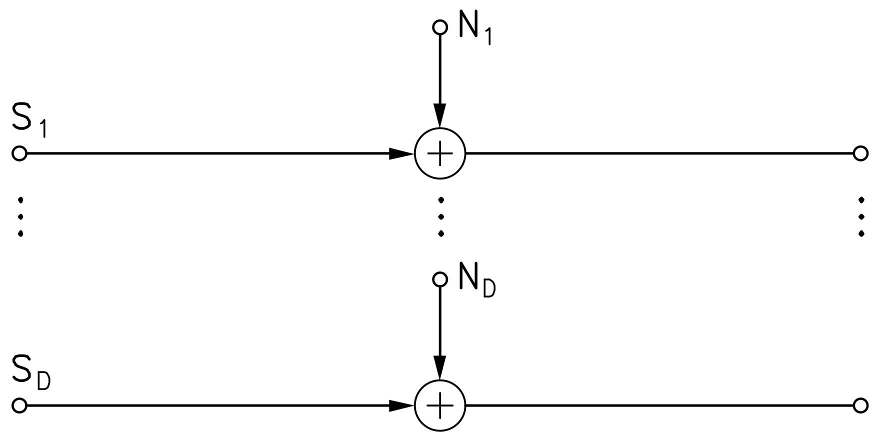
9.2 Optimization of DMT/OFDM

Scenarios:

- Radio Applications for Broadcasting (Radio):
  Transmission channels to all participants are different: Optimization of OFDM system is difficult.
  Typical: All subcarriers use identical modulation scheme and power

- Point–to–point transmission (e.g. ADSL):
  The use of the subcarriers (parallel subchannels) can be optimized.
  Typical: Optimization of the rate and transmit power for the individual sub–carriers.
  => *Loading–Algorithms*

starting point:



Rate allocation according to channel capacity:

(Loading algorithm by Chow/Cioffi/Bingham. 1995)

Capacity of a subchannel i: $C_i = \mathrm{ld}(1 + S_i/N_i)$

Choose: $S_i = \text{const.}$ — rate in i-th subchannel $R_i = \mathrm{ld}(1 + 1/(N_i \cdot \gamma))$

with γ such that the desired total rate R is met $R = \sum_{i=1}^{D} R_i = \sum_{i=1}^{D} \mathrm{ld}(1 + 1/(N_i \cdot \gamma))$

Summary of algorithm:

- Rounding to integer rates
- The rate quantization can be equalized using an adapted power distribution
- Optimal with respect to channel capacity
- Not the optimal criterion for the division of a given total rate

Rate allocation according to the best power allocation:

(Loading algorithm by Hughes–Hartogs, 1987/89)

Given: Desired rate R; desired error rate

Algorithm:

1. Definition of the transmission power $S_{i,R}$, such that information may be transmitted in the i-th subchannel at a rate of $R = 1, 2, \ldots$ bits for a given error rate
2. Definition of the *"power increments"* $\Delta S_{i,R} = S_{i,R} - S_{i,R-1}$
3. The transmitted bits are successively mapped to the particular subchannel for which the transmission of an additional binary symbol will require the least amount of additional power ($\Delta S_{i,R}$ minimal)
4. Necessary total transmit power is thus automatically obtained

- Computationally complex algorithm
- Since the error rate is not known a priori, the optimal rate and power distribution is not obtained in general.

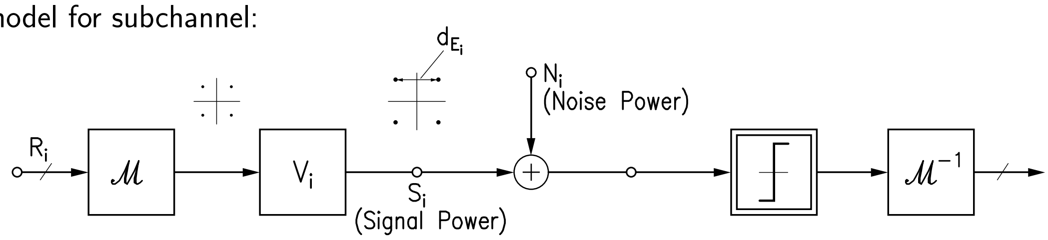
Rate allocation for the minimization of the symbol error rate:

(Loading algorithm by Fischer/Huber, 1995)

Given: Desired rate R; total transmit power S

Goal: Constant and minimal error rate in all subchannels

Equivalent model for subchannel:



Symbol error rate:

$$\mathrm{SER}_i = K_i \cdot \mathrm{Q}\left(\sqrt{\mathrm{SNR}_{0,i}}\right), \qquad \text{with} \quad \mathrm{SNR}_{0,i} = \frac{d_{Ei}^2/4}{N_i/2}$$

with d_{Ei}^2: minimum squared Euclidean distance in i–the subchannel

$\mathrm{Q}(x)$: Gaussian Q-function

512

Optimization problem:

$$\mathrm{SNR}_{0,i} = \mathrm{SNR}_0 \longrightarrow \max$$

with a) $R = \sum_{i=1}^{D} R_i = \text{const.}$ and

b) $S = \sum_{i=1}^{D} S_i = \text{const.}$

Simplification for the optimization:

Consider hypothetical signal constellations with $2^{R_i} \notin \mathbb{N}$ signal points and point separation of 2

Average transmission power (continuous approximation and $d_{Ei} = 2V_i$):

$$S_i = V_i^2 \cdot \frac{2}{3} 2^{R_i} = \mathrm{SNR}_0 \frac{N_i}{2} \frac{2}{3} 2^{R_i}$$

Alternatively, with condition b)

$$S = \sum\nolimits_{i=1}^{D} S_i = \frac{\mathrm{SNR}_0}{3} \sum\nolimits_{i=1}^{D} N_i \cdot 2^{R_i}$$

and thus,

$$\mathrm{SNR}_0 = \frac{3S}{\sum_{i=1}^{D} N_i \cdot 2^{R_i}}$$

Lagrangian function for optimization:

$$L = \sum\nolimits_{i=1}^{D} N_i \cdot 2^{R_i} - \mu \left(\sum\nolimits_{i=1}^{D} R_i - R \right)$$

We have

$$\frac{\partial L}{\partial R_i} = N_i \cdot \ln(2) \cdot 2^{R_i} - \mu \overset{!}{=} 0$$

Or alternatively,

$$N_i \cdot 2^{R_i} = \mathrm{const.}, \qquad i = 1, 2, \ldots, D$$

Multiplying this product over all terms yields

$$(\mathrm{const.})^D = \left(N_i \cdot 2^{R_i} \right)^D = \prod_{l=1}^{D} N_l \cdot 2^{R_l} = 2^{\sum_{l=1}^{D} R_l} \cdot \prod_{l=1}^{D} N_l = 2^R \cdot \prod_{l=1}^{D} N_l$$

with

$$\left(N_i \cdot 2^{R_i} \right)^D = 2^R \cdot \prod_{l=1}^{D} N_l,$$

the rate distribution can be directly obtained as

$$R_i = \frac{R}{D} + \frac{1}{D} \cdot \mathrm{ld} \left(\frac{\prod_{l=1}^{D} N_l}{N_i^D} \right)$$

If $N_i \cdot 2^{R_i} = \mathrm{const.}$, one obtains (compare with above formulation)

$$S_i = S/D$$

Ultimately, with the choice of

$$S = D \cdot \frac{2}{3} 2^{R/D}$$

(D identical systems with rate R/D as comparison)

We obtain for the SNR

$$\begin{aligned}
\mathrm{SNR}_0 &= \frac{3S}{\sum_{i=1}^{D} N_i \cdot 2^{R_i}} \\
&= \frac{3S}{\sum_{i=1}^{D} \sqrt[D]{2^R \cdot \prod_{l=1}^{D} N_l}} \\
&= \frac{3D \cdot \frac{2}{3} 2^{R/D}}{D 2^{R/D} \sqrt[D]{\prod_{l=1}^{D} N_l}} \\
&= \frac{1}{\sqrt[D]{\prod_{l=1}^{D} N_l / 2}}
\end{aligned}$$

516

Summary: *Optimal Rate and Power Distribution*

$$\begin{aligned}
R_i &= \frac{R}{D} + \frac{1}{D} \cdot \mathrm{ld}\left(\frac{\prod_{l=1}^{D} N_l}{N_i^D}\right) \\
S_i &= S/D \\
\mathrm{SNR}_0 &= \frac{1}{\sqrt[D]{\prod_{i=1}^{D} N_i / 2}} \qquad \text{(with } S = D \cdot \tfrac{2}{3} 2^{R/D}\text{)}
\end{aligned}$$

Procedure in practice:

- Quantization of the rates; restriction of the maximal rate $R_{\max}$
- Compensation of the quantization using power adaptation
- For encoding: Substitute N_i for N_i/G_i (G_i: coding gain)

Algorithm:

Allocation of the rate according to

$$R_i' = \left(R + \sum_{l=1}^{D} \mathrm{ld}\,(N_l) \right) \Big/ D - \mathrm{ld}\,(N_i)$$

Exclude carriers with $R_i' \leq 0$ and proceed with new (iterative) distribution

Round R_i' to $R_i \in \{0, 1, 2, \ \ldots, R_{\max}\}$, with $\sum_{i=1}^{D} R_i = R$

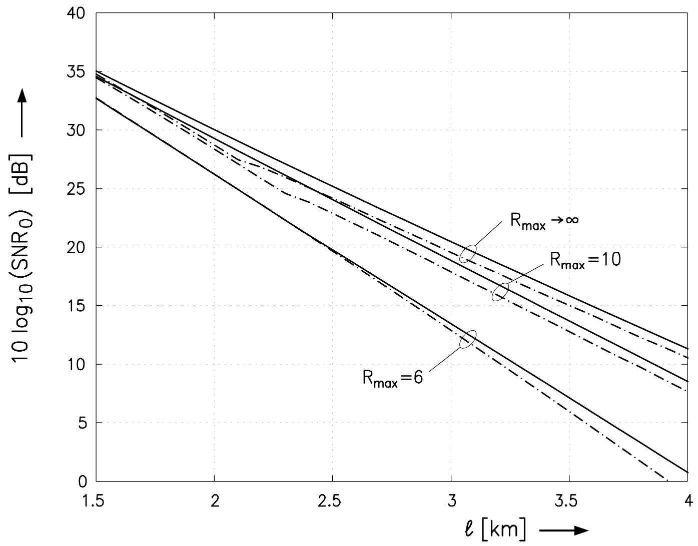
Adapt partial transmit power such that

– all used subchannels have the same error rate and

– $\sum_{i=1}^{D} S_i = S$

Example:

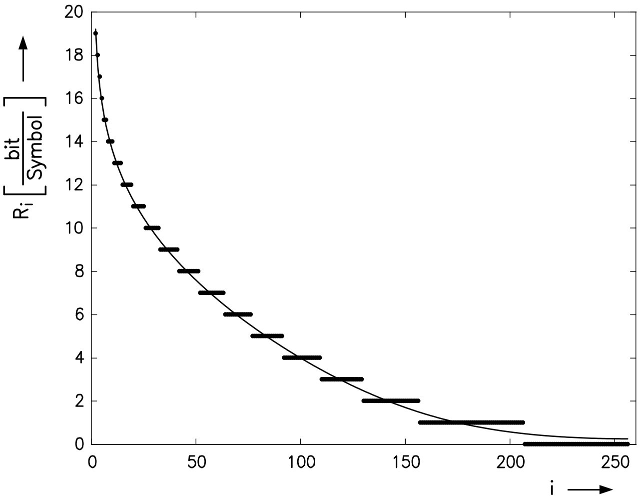
- 2.048 $\frac{\mathrm{Mbit}}{\mathrm{s}}$–transmission over a symmetric cable with a conductor diameter of 0.4 mm (26AWG cable)
- Typical near–end crosstalk situation
- Channel symbol clock $1/T = 1.024$ MHz; DFT of length 512
  => 1024 bit are accommodated in each DMT frame



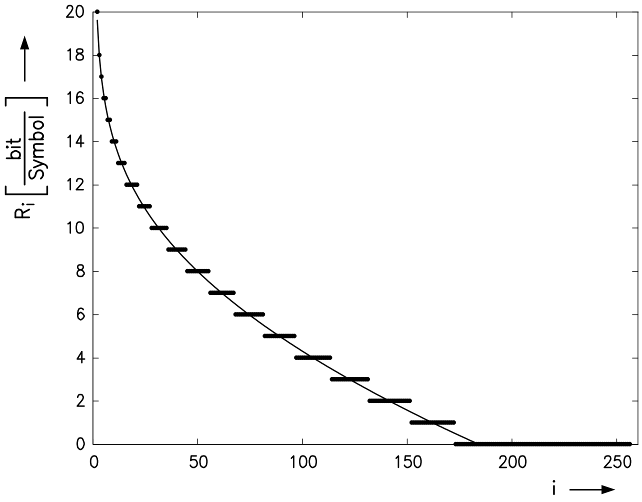
Solid line: Algorithm from Fischer/Huber;
Dotted line: Algorithm from Chow et al.

Rate Allocation versus number of carriers: $\ell = 2.745\,\text{km}$, $R_{\max} \to \infty$

Algorithm from Chow et al.

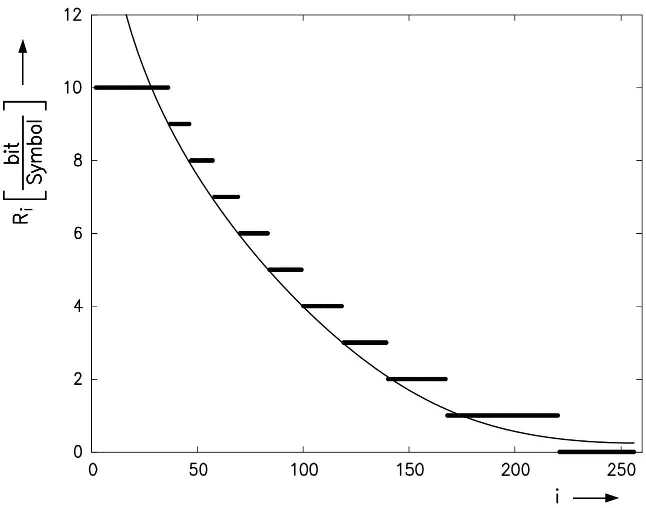


Algorithm from Fischer/Huber

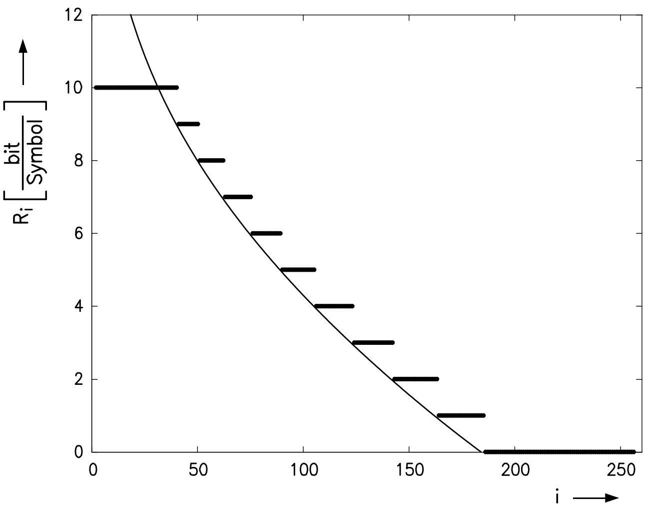


Rate Allocation versus Number of Carriers: $\ell = 2.745\,\text{km}$, $R_{\max} = 10$

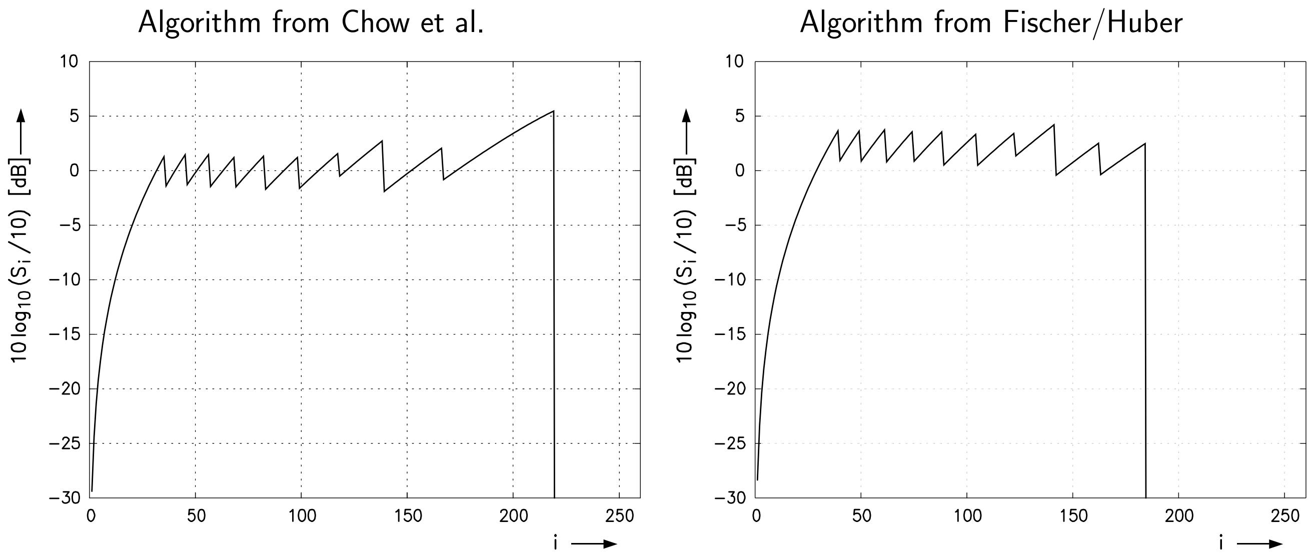
Algorithm from Chow et al.



Algorithm from Fischer/Huber



Corresponding power allocation for $R_{\max} = 10$



9.3 Comparison of DMT and DFE

DFE: *non*linear, time–*in*variant procedure

DMT/OFDM: linear, time–variant procedure

Comparison criterion: SNR_0

Minimal squared distance between the signal point and the decision region boundary relative to the noise variance per dimension

1. ONF: (ZF–LE)

$$\mathrm{SNR}_0^{(\mathsf{ZF-LE})} = \frac{2}{T \displaystyle\int_{-1/(2T)}^{1/(2T)} \Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi fT})\,\mathrm{d}f}$$

2. DFE:

$$\mathrm{SNR}_0^{(\mathsf{ZF-DFE})} = 2 \cdot \exp\left\{ -T \int_{-1/(2T)}^{1/(2T)} \ln\left(\Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi fT})\right)\,\mathrm{d}f \right\}$$

3. DMT/OFDM:

Assumption: Given overall transfer function $H^{(\mathsf{DMT})}(\mathrm{e}^{\mathrm{j}2\pi fT})$ and PSD of the noise:

$$\Phi_{nn}^{(\mathsf{DMT})}(\mathrm{e}^{\mathrm{j}2\pi fT}) = \Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi fT}) \cdot \left| H^{(\mathsf{DMT})}(\mathrm{e}^{\mathrm{j}2\pi fT}) \right|^2$$

Transmission gains in the parallel sub–channels:

$$\lambda_i = H^{(\mathsf{DMT})}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})$$

Noise power in the parallel sub–channels:

$$\sigma_{n_i'}^2 \approx \Phi_{nn}^{(\mathsf{DMT})}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})$$

After compensation of gain factors, we have AWGN channels where the variance of the complex–valued noise is given by

$$\begin{aligned} \frac{\sigma_{n_i'}^2}{|\lambda_i|^2} &\approx \frac{\Phi_{nn}^{(\mathsf{DMT})}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})}{|H^{(\mathsf{DMT})}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})|^2} \\ &= \frac{\Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})|H^{(\mathsf{DMT})}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})|^2}{|H^{(\mathsf{DMT})}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})|^2} \\ &= \Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D}) \end{aligned}$$

524

$\Longrightarrow$ (Approximately) independent from the actual channel impulse response!
(as long as the impulse response is shorter than the guard interval)

Optimal loading:

$$\mathrm{SNR}_0^{(\mathsf{DMT})} = \frac{2}{\sqrt[D]{\prod_{i=1}^{D} \Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi(i-1)/D})}}.$$

Asymptotic behavior ($D \to \infty$)

$$
\begin{aligned}
\sqrt[D]{\prod_{i=0}^{D-1} \Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi i/D})} &= \exp\left\{\frac{1}{D}\ln\left(\prod_{i=0}^{D-1} \Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi i/D})\right)\right\} \\
&= \exp\left\{\frac{1}{D}\sum_{i=0}^{D-1} \ln\left(\Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi i/D})\right)\right\} \\
&\xrightarrow{D\to\infty} \exp\left\{\int_{-1/2}^{1/2} \ln\left(\Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi\nu})\right)\,\mathrm{d}\nu\right\} \\
&\overset{\nu=fT}{=} \exp\left\{T\int_{-1/(2T)}^{1/(2T)} \ln\left(\Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi fT})\right)\,\mathrm{d}f\right\}
\end{aligned}
$$

alternatively

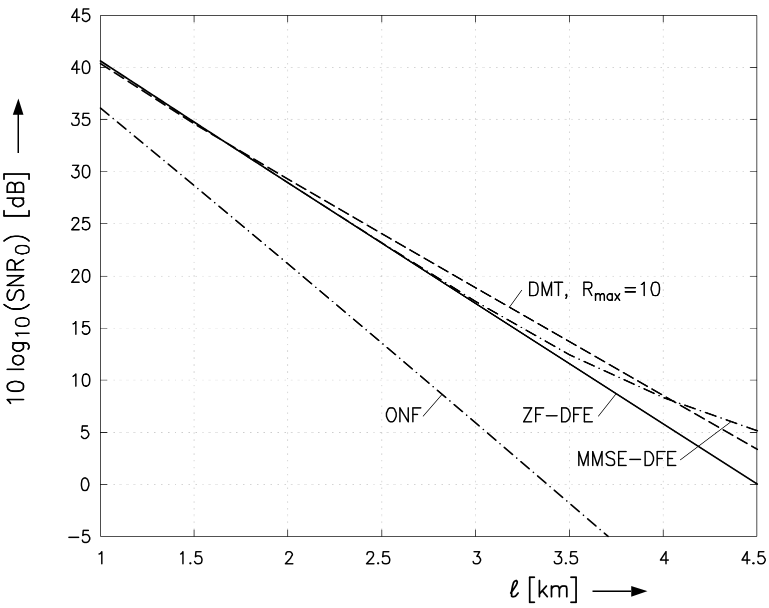
$$
\mathrm{SNR}_0^{(\mathsf{DMT})} = 2\cdot\exp\left\{-T\int_{-1/(2T)}^{1/(2T)} \ln\left(\Phi_{nn}^{(\mathsf{ZF-LE})}(\mathrm{e}^{\mathrm{j}2\pi fT})\right)\,\mathrm{d}f\right\}
$$

therefore: $\mathrm{SNR}_0^{(\mathsf{DMT})} = \mathrm{SNR}_0^{(\mathsf{ZF-DFE})}$

=> DFE and DMT have (asymptotically) exactly the same performance!

Example:

SNR_0 over a cable of length ℓ



- ONF, DFE: 4–ASK
- DMT: DFT of length 512, optimal loading for rate 4 (16 QAM) per sub–channel on average

Discussion of DMT/OFDM:

Advantages:

- Transformation of dispersive linear distortion into non–dispersive variable gain factors by means of DFT at the receiver–side, as long as the guard interval is sufficiently long.

  $\Rightarrow$ Simplest and most efficient transmission scheme for frequency selective fading channels (i.e. mobile communications), since no equalization of ISI is necessary!

- Robust transmission over dispersive linearly distorting channels when used in conjunction with channel coding (parallel AWGN channels)

- By constructing long DMT symbols (i.e. by using a large number of sub–carriers), the loss by insertion of a guard interval is kept small.

  Dispersion duration $T_D \ll DT \; : \; D_0 \ll D$ is sufficient

- For a large number of sub–carriers D, an average power spectral density of near–rectangular shape is achievable by setting sub-carriers to zero.

- Not sensitive to phase shifts in the symbol clock; sampling clock errors results in increasing phase shift for the sub–carriers which is:
  - Reversible through phase synchronization
  - Insignificant for DPSK along the subcarrier

Disadvantages:

- Very high signal delay (initial run–time)

$$> 2T_\mathrm{s} = 2RT_b \ ,$$

  as a large amount of binary data is represented in a single DMT symbol (typically $R > 1000$)

- Guard interval results in loss of power and bandwidth efficiency — DMT/OFDM is only efficient for $D \gg 1$.

- DMT is extremely susceptible to problems caused by nonlinear distortion (intermodulation between sub–carriers) and quickly time–varying channels.

- Transmit signal has noise–like characteristics (almost Gaussian distributed; Central Limit Theorem!)

  $\Rightarrow$ extremely high peak value (crest–) factor $\quad \zeta \leq \sqrt{D}\, \zeta_a \, \zeta_g \ ,$

  Assumption: Each of the D PAM schemes use a signal constellation $\mathcal{A}$ with same Crest–factor ζ_a.

  This problem should be considered from a statistical point of view. Many algorithms exist for peak–value reduction.

- Very rigid conditions on the linearity of the high power amplifier (HPA) at the transmitter because the peak value occurs rather rarely.

  $\Rightarrow$ Large back–off required

  $\Rightarrow$ High losses in power efficiency

  Example: Digital radio DAB:

  - 1536 modulated carriers; 4 DPSK modulation;
  - 5 kW transmitter needed for 100 W of average transmit power!

532

A Basics of Signal and System Theory

A-1 Fourier Transform

Time-domain function: $x(t)$

Frequency-domain function: $X(f)$

$$
\begin{aligned}
X(f) &= \int_{-\infty}^{+\infty} x(t)\, \mathrm{e}^{-\mathrm{j}2\pi f t} \mathrm{d}t = \mathcal{F}\{x(t)\} \\
x(t) &= \int_{-\infty}^{+\infty} X(f)\, \mathrm{e}^{+\mathrm{j}2\pi f t} \mathrm{d}f = \mathcal{F}^{-1}\{X(f)\}
\end{aligned}
$$

Examples of Fourier pairs:

$x(t)$	$\circ\!\!-\!\!\bullet$	$X(f)$
1		$\delta(f)$
$\delta(t)$		1
$\sin(2\pi f_0 t)$		$\frac{\mathrm{j}}{2}\left(\delta(f+f_0) - \delta(f-f_0)\right)$
$\cos(2\pi f_0 t)$		$\frac{1}{2}\left(\delta(f+f_0) + \delta(f-f_0)\right)$
$\mathrm{rect}(t/T)$		$T\mathrm{si}(\pi f T)$
$\mathrm{si}(\pi t/T)$		$T\mathrm{rect}(fT)$
$\mathrm{e}^{-\pi(t/T)^2}$		$T\mathrm{e}^{-\pi(fT)^2}$

$$\mathrm{e}^{-t/T}\epsilon(t) \qquad \frac{T}{1+\mathrm{j}2\pi fT}$$

$$\mathrm{e}^{-|t|/T} \qquad \frac{2T}{1+(2\pi fT)^2}$$

$$\frac{1}{1+(t/T)^2} \qquad \pi T\mathrm{e}^{-2\pi|f|T}$$

$$\epsilon(t) \qquad \frac{1}{\mathrm{j}2\pi f}+\frac{1}{2}\delta(f)$$

$$\mathrm{sign}(t) \qquad \frac{1}{\mathrm{j}\pi f}$$

$$\frac{1}{\pi t} \qquad -\mathrm{j}\ \mathrm{sign}(f)$$

Definitions:

$$\mathrm{rect}(x) \overset{\mathrm{def}}{=} \begin{cases} 1 & \text{for } |x| < 0{,}5 \\ 0.5 & \text{for } |x| = 0{,}5 \\ 0 & \text{for } |x| > 0{,}5 \end{cases}$$

$$\mathrm{si}(x) \overset{\mathrm{def}}{=} \begin{cases} 1 & \text{for } x = 0 \\ \frac{\sin(x)}{x} & \text{for } x \neq 0 \end{cases}$$

$$\epsilon(x) \overset{\mathrm{def}}{=} \begin{cases} 1 & \text{for } x > 0 \\ 0.5 & \text{for } x = 0 \\ 0 & \text{for } x < 0 \end{cases}$$

$$\mathrm{sign}(x) \overset{\mathrm{def}}{=} \begin{cases} 1 & \text{for } x > 0 \\ 0 & \text{for } x = 0 \\ \text{-1} & \text{for } x < 0 \end{cases}$$

A-2 Transfer of Signals through Systems

$\mathcal{T}$: Mapping of a set of functions to a set of functions: Transformation

$$\mathcal{T}\{x(t)\} = y(t)$$

$$\text{Shorthand notation}: \quad x(t) \rightarrow y(t)$$

Dispersive system:

All function values $x(t)$, $\forall t \in \mathbb{R}$, influence the value the output signal $y(t_0)$ at distinct time instant t_0.

Causal dispersive system:

All function values $x(t)$, $t \leq t_0$, influence the value of the output signal $y(t_0)$ at distinct time instant t_0.

Definition: Signal distortion

The transform $\mathcal{T}$ causes a signal distorition if

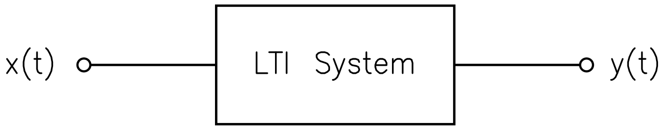
$$y(t) \neq cx(t - t_0) \quad \forall c \in \mathbb{C} \text{ and } t_0 \in \mathbb{R}$$

I.e., if the transformed signal $y(t)$ does not have the same shape as the original signal $x(t)$ or if $y(t)$ cannot be generated by means of scaling and shifting of $x(t)$.

Definition: Eigen functions of a transform $\mathcal{T}$

Functions $e(t)$ which are not distorted by transform $\mathcal{T}$ are called eigen functions of T.

A-3 Linear, Time-invariant, Dispersive Systems (LTI Systems)



Linearity: From $x_1(t) \to y_1(t)$ and $x_2(t) \to y_2(t)$,
$(\lambda \cdot x_1(t) + \beta \cdot x_2(t)) \to (\lambda \cdot y_1(t) + \beta \cdot y_2(t))$ follows

Time-invariance: From $x(t) \to y(t)$, $\forall t_0 \in \mathbb{R}$:
$x(t - t_0) \to y(t - t_0)$ follows
Shifting principle

Impulse response of LTI systems: Response signal $h(t)$ to input signal $x(t) = \delta(t)$. If $h(t)$ is expanded in time, i.e., for

$$h(t) \neq c\delta(t - t_0) \qquad \forall c \in \mathbb{C} \; ; \; t_0 \in \mathbb{R}$$

then the LTI system is said to be *dispersive*.

– Remark 1:
  In a linear system, the output signal is in general **not** proportional to the input signal. This only applies for dispersionless LTI systems; dispersive LTI systems usually generate **linear** distortions.

– Remark 2:
  Distinguish: Systems *with-memory* and *dispersive* systems.
  E.g.: Time delay: $h(t) = \delta(t - t_0)$
  dispersionless but for $t_0 \neq 0$, the system has memory!

Representation of a signal as a sum (integral) of (differentially small) δ–functions:
Definition of δ–function as a distribution

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

From linearlity and time-invariance, as well as the transformation $\delta(t-\tau) \rightarrow h(t-\tau)$, we get

$$\boxed{y(t) = \int\limits_{-\infty}^{+\infty} x(\tau) h(t-\tau)\,\mathrm{d}\tau}$$

Convolution integral

Shorthand notation:

$$y(t) = x(t) * h(t) = h(t) * x(t) = \int\limits_{-\infty}^{+\infty} h(\tau) x(t-\tau)\,\mathrm{d}\tau$$

For the "convolution product", commutativity, associativity and distributivity apply.

Exponential functions are eigenfunctions of LTI systems.

Special case: $x(t) = \mathrm{e}^{\mathrm{j}2\pi f t}$

$$y(t) = \int\limits_{-\infty}^{+\infty} h(\tau) \mathrm{e}^{\mathrm{j}2\pi f(t-\tau)}\,\mathrm{d}\tau = \mathrm{e}^{\mathrm{j}2\pi f t}\, H(f)$$

$$\text{with} \quad H(f) = \int\limits_{-\infty}^{+\infty} h(\tau)\,\mathrm{e}^{-\mathrm{j}2\pi f \tau} \mathrm{d}\tau = \mathcal{F}\{h(t)\}$$

From $x(t) = \int\limits_{-\infty}^{+\infty} X(f) \mathrm{e}^{\mathrm{j}2\pi f t} \mathrm{d}f$ and the superposition principle, we obtain

$$y(t) = \int\limits_{-\infty}^{+\infty} X(f) \cdot H(f) \cdot \mathrm{e}^{\mathrm{j}2\pi f t} \mathrm{d}f = \mathcal{F}^{-1}\{Y(f)\}$$

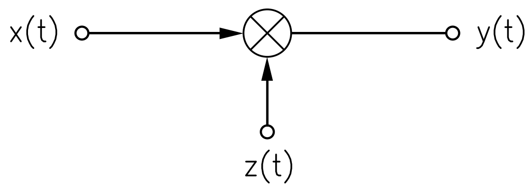
$$\Rightarrow \boxed{Y(f) = X(f) \cdot H(f)}$$

$H(f)$: Transfer function

Convolution in time domain corresponds to multiplication in frequency domain:

$$* \circ\!\!-\!\!\bullet \cdot$$

A-4 Multiplication of Signals



y(t)=x(t). z(t)

Considering $x(t) \to y(t)$ with a given function $z(t)$ presents a **linear**, time-**variant**, dispersion**less** signal transformation.

$$\boxed{Y(f) = Z(f) * X(f) = X(f) * Z(f)}$$

Multiplication in time domain corresponds to convolution in frequency domain:

$$\cdot \circ\!\!-\!\!\bullet *$$

Examples:

$z(t)$	$Z(f)$	$Y(f)$
$\mathrm{e}^{\mathrm{j}2\pi f_0 t}$	$\delta(f - f_0)$	$X(f - f_0)$
$\cos(2\pi f_0 t)$	$\frac{1}{2}(\delta(f + f_0) + \delta(f - f_0))$	$\frac{1}{2}(X(f + f_0) + X(f - f_0))$
$\sin(2\pi f_0 t)$	$\frac{\mathrm{j}}{2}(\delta(f + f_0) - \delta(f - f_0))$	$\frac{\mathrm{j}}{2}(X(f + f_0) - X(f - f_0))$
$\sum_{i=-\infty}^{+\infty} \delta(t - iT)$	$\frac{1}{T} \sum_{k=-\infty}^{+\infty} \delta\left(f - \frac{k}{T}\right)$	$\frac{1}{T} \sum_{k=-\infty}^{+\infty} X\left(f - \frac{k}{T}\right)$

A-5 Signal Energy and Energy Spectral Density

If

$$\int_{-\infty}^{+\infty} |x(t)|^2 \,\mathrm{d}t = E_x < \infty$$

holds, then the signal $x(t)$ is *energy-limited* with *energy* E_x.

Remark:

Energy-limited signals are usually of impulsive shape.

For energy-limited signals, the Fourier integral converges.

The following equivalence holds:

$$\begin{aligned} E_x &= \int_{-\infty}^{+\infty} x(t)x^*(t)\,\mathrm{d}t = \mathcal{F}\{x(t)\cdot x^*(t)\}\big|_{f=0} \\ &= X(f) * X^*(-f)\big|_{f=0} = \int_{-\infty}^{+\infty} X(\varphi)X^*(\varphi)\,\mathrm{d}\varphi \end{aligned}$$

$$\boxed{\underbrace{E_x = \int_{-\infty}^{+\infty} |x(t)|^2\,\mathrm{d}t = \int_{-\infty}^{+\infty} |X(f)|^2\,\mathrm{d}f}_{\text{Parseval's theorem}}}$$

Definition: *Energy spectral density* of signal $x(t)$:

$$|X(f)|^2$$

$$|X(f)|^2 = X(f)\cdot X^*(f) = \mathcal{F}\{x(t) * x^*(-t)\}$$

A-6 Autocorrelation

Definition:

Autocorrelation of signal $x(t)$:

$$\varphi_{xx}(\tau) = x(\tau) * x^*(-\tau) = \int_{-\infty}^{+\infty} x(t+\tau)x^*(t)\,\mathrm{d}t$$

$$\varphi_{xx}(\tau) \circ\!\!-\!\!\bullet\ |X(f)|^2$$

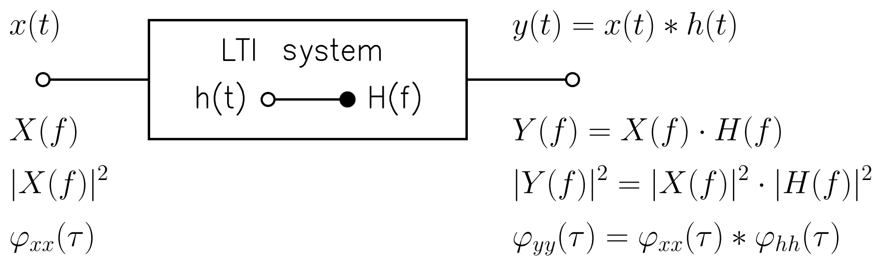
Analogous definition:

Crosscorrelation and *cross energy spectral density* of the two signals $x(t)$ and $y(t)$:

$$\varphi_{xy}(\tau) = x(\tau) * y^*(-\tau) = \int_{-\infty}^{+\infty} x(t+\tau)y^*(t)\,\mathrm{d}t$$

$$\varphi_{xy}(\tau) \circ\!\!-\!\!\bullet\ X(f)\cdot Y^*(f)$$

Transmission over an LTI System



Definition:

System autocorrelation $\varphi_{hh}(\tau) = h(\tau) * h^*(-\tau)$

Definition:

Power transfer function: $|H(f)|^2$

$$\varphi_{hh}(\tau) \circ\!\!-\!\!\bullet\ |H(f)|^2$$

A-7 Stochastic Process

A-7.1 Probability

Let **H** be a set of (usually many) possible elementary outcomes η_i of a random experiment. The one-time realization of a random experiment results in an elementary outcome. Let the random experiment be repeatable under the same conditions.

Definition: An **event** A is a subset of **H**

$$A \subseteq \mathbf{H}$$

If the observed elementary outcome η belongs to the subset A, $\eta \in A$, then event A has occurred. The outcome η complies with the proporties which are characteristic for all elementary outcomes in the set A.

Definition: Probability

An event is assigned a real number in the interval [0;1] as probability

$$\Pr(A) \in [0;1] \subset \mathbb{R}$$

in such a way that: (Axioms of probability)

1. $\Pr(\mathbf{H}) = 1$ One elementary outcome surely occurs
2. $\Pr(\emptyset) = 0$
3. If $A \cap B = \emptyset$ (disjoint, incompatible events)
   is satisfied, then
   $\Pr(A \cup B) = \Pr(A) + \Pr(B)$ is valid

Corollaries:

1. $\Pr(A \cup B) \;=\; \Pr(A) + \Pr(B) - \Pr(A \cap B)$
   Proof: Disassemble $A \cup B$ to disjoint subsets $A \setminus (A \cap B), \quad B \setminus (A \cap B), \quad (A \cap B)$
2. $\Pr(\mathbf{H} \setminus A \stackrel{\text{def}}{=} \bar{A}) \;=\; 1 - \Pr(A)$
   Proof: Axiom 3 und 1 with $B = \bar{A}, \quad A \cup \bar{A} = \mathbf{H}, \quad A \cap \bar{A} = 0$
   $1 \;=\; \Pr(A) + \Pr(\bar{A})$

Definition: Conditional Probability

$$\Pr(A|B) \;=\; \frac{\Pr(A\cap B)}{\Pr(B)}$$

For the conditional probability, it is assumed that an outcome of the subset B has occured, therefore the condition $\eta \in B$ is met. This condition is surely satisfied for the intersection $A \cap B$. Thus, B takes the role of the set of all possible outcomes **H** as well as the axioms 1-3, and the corollaries are valid for conditional probabilities in the same way as for normal probabilities.

1. $\Pr(B|B) \;=\; 1$
2. $\Pr(\emptyset|B) \;=\; 0$
3. For disjoint events A und C, thus for $A \cap C = \emptyset$, we have;
   $\Pr(A\cup C|B) \;=\; \Pr(A|B) + \Pr(C|B)$

Bayes' rule: Change of event and condition in conditional probabilities:

$$\Pr(A|B) \;=\; \frac{\Pr(A\cap B)}{\Pr(B)} \;=\; \Pr(B|A)\cdot\frac{\Pr(A)}{\Pr(B)}$$

Definition: Independent events

$$\Pr(A\cap B) \;=\; \Pr(A)\cdot\Pr(B)$$

552

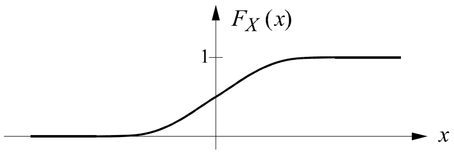
Corollary: $\quad \Pr(A|B) \;=\; \dfrac{\Pr(A\cap B)}{\Pr(B)} \;=\; \Pr(A)$

A-7.2 Continuous Real Random Variable

Let **H** be the set of outcomes η. Let each outcome η be assigned a real number $\boldsymbol{x}(\eta)$, which will be referred to as a real random variable. Thus, the result of the realization of a random experiment is a random variable.
Cumulative distribution function (CDF):

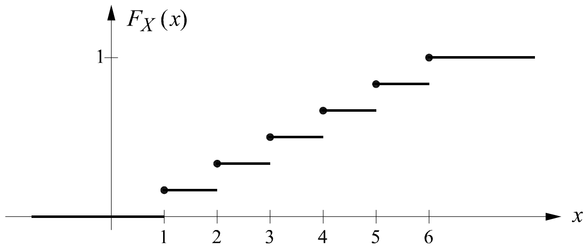
$$F_{\boldsymbol{x}}(x) \;=\; \Pr(\boldsymbol{x}(\eta) \le x) \quad , \quad x \in \mathbb{R}$$

The CDF takes values from 0 to 1 and is a monotonically non-decreasing function



If X takes only discrete values, then the CDF shows steps at discrete values (right-hand-side continuity)

Example: Dice , $X \in \{1,2,3,4,5,6\}$



Let $b \geq a$; $a,b \in \mathbb{R}$, then $(-\infty,a] \cup [a,b] = (-\infty,b]$ holds. The events $\boldsymbol{x}(\eta) \leq a$ and $a < \boldsymbol{x}(\eta) \leq b$ are disjoint, therefore we get

$$\Pr(\boldsymbol{x}(\eta) \leq a) + \Pr(a < \boldsymbol{x}(\eta) \leq b) = \Pr(\boldsymbol{x}(\eta) \leq b)$$

$$\boxed{\Pr(a < \boldsymbol{x}(\eta) \leq b) = F_{\boldsymbol{x}}(b) - F_{\boldsymbol{x}}(a)}$$

The probability that a real random variable falls in the interval between a and b is equal to the difference between the distribution function at the upper interval limit and the distribution function at the lower interval limit!

554

Corollary: If $F_{\boldsymbol{x}}(a)$ is continuous at point a, then

$$\boxed{\Pr(\boldsymbol{x}(\eta) = a) = 0}$$

Probability Density Function

Typically, the probability that a truly continuously distributed random variable equals a fixed value is zero. For example, a random noise voltage N will not be equal to the exact value of $\pi = 3.14159....$ At some point after the decimal place there is certainly a deviation.

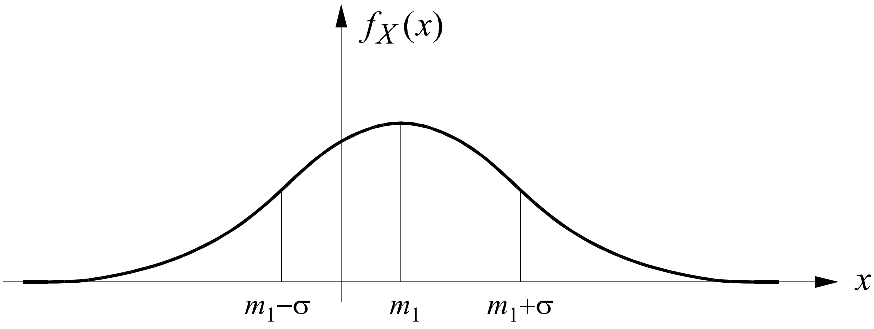
Definition: Probability Density Function (PDF)

$$\begin{aligned} f_{\boldsymbol{x}}(x) &= \frac{\mathrm{d}F_{\boldsymbol{x}}(x)}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{F_{\boldsymbol{x}}(x+\Delta x) - F_{\boldsymbol{x}}(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{\Pr(\boldsymbol{x}(\eta) \leq x + \Delta x) - \Pr(\boldsymbol{x}(\eta) \leq x)}{\Delta x} \end{aligned}$$

$$\int_{-\infty}^{x} f_{\boldsymbol{x}}(\alpha)\mathrm{d}\alpha = F_{\boldsymbol{x}}(x) \quad ; \quad \int_{-\infty}^{+\infty} f_{\boldsymbol{x}}(x)\mathrm{d}x = 1 \quad ; \quad f_{\boldsymbol{x}}(x) \geq 0$$

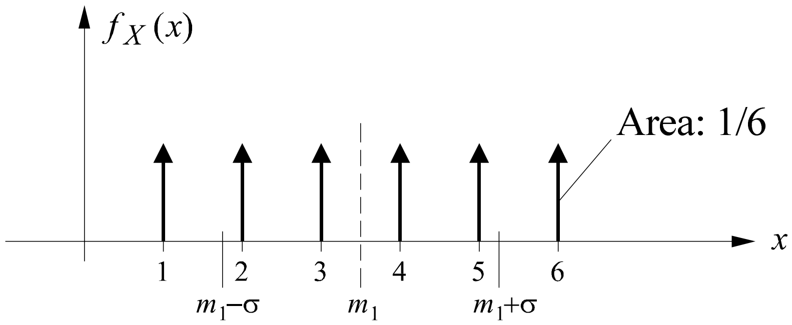
$$\Pr(a < \boldsymbol{x}(\eta) \leq b) = \int_{a}^{b} f_{\boldsymbol{x}}(x)\mathrm{d}x$$

Examples:



For discrete random variables, δ–functions arise after differentiation.

Example: Dice



556

Expected values, mean values

Definition: Expected value with respect to function $g(x)$

$$E\{g(X)\} = \int_{-\infty}^{+\infty} g(x) f_{\boldsymbol{x}}(x) \mathrm{d}x$$

Special cases:

- Moments $\quad m_i : g(x) = x^i$

1. (Linear) mean value, first moment
   $g(x) = x$
   $$m_1 = E\{X\} = \int_{-\infty}^{+\infty} x f_{\boldsymbol{x}}(x) \mathrm{d}x = \overline{x}$$
2. Quadratic mean value, second moment, power
   $g(x) = x^2$
   $$m_2 = E\{X^2\} = \int_{-\infty}^{+\infty} x^2 f_{\boldsymbol{x}}(x) \mathrm{d}x = m_2 = \overline{x^2} \qquad \text{"power"}$$

- Central moment $\qquad g(x) = (x - m_1)^i$

Variance

$$\mu_2 \;=\; \int\limits_{-\infty}^{+\infty} (x - \overline{x})^2 f_{\boldsymbol{x}}(x)\mathrm{d}x \;=\; m_2 - m_1^2 \;=\; \sigma^2$$

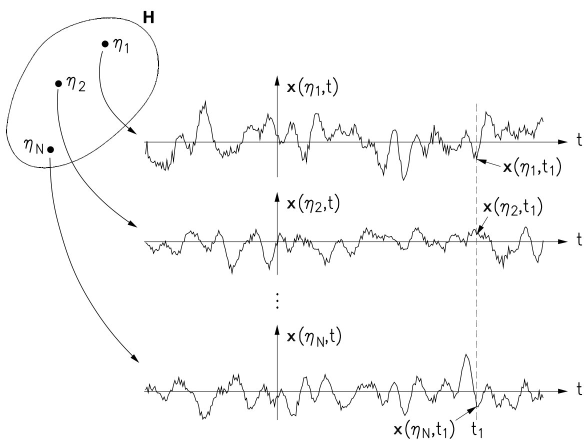
$$\sigma \;=\; \sqrt{\mu_2} \quad : \quad \text{root-mean-square deviation, standard deviation, effective value}$$

A-7.3 Definition of a Stochastic Process

Let **H** be the set of events of disjoint outcomes η.
Each outcome η of a random experiment will be assigned to a (real or complex-valued) *function of time*

$$\boldsymbol{x}(\eta, t).$$

The outcome of the random experiment is also one of many possible functions of time. These functions are referred to as *sample functions* of the stochastic process.



A-7.4 Joint Stochastic Processes

Each outcome is assigned two (or more) time-functions

$$\boldsymbol{x}(\eta, t)\ ,\ \boldsymbol{y}(\eta, t).$$

Complex stochastic processes

$$\boldsymbol{z}(\eta, t) = \boldsymbol{x}(\eta, t) + \mathrm{j}\boldsymbol{y}(\eta, t)$$

will be treated in the following as two real joint stochastic processes.

A-7.5 The Probability Density Function (PDF)

PDF at time instant t_1: Variable $\boldsymbol{x}(\eta,t_1)$ at fixed time instant t_1 is a real random variable with PDF

$$f_{\boldsymbol{x}}(x,t_1) = \lim_{\Delta x \to 0} \frac{\Pr(x < \boldsymbol{x}(\eta,t_1) \le x + \Delta x)}{\Delta x}$$

Joint PDF for multiple time instants $(t_1,t_2,\ldots,t_n) = \vec{t}$

$$f_{\vec{x}}(\vec{x},\vec{t}\,) = \lim_{\Delta x \to 0} \frac{\Pr\left(\bigcap_{i=1}^{n}(x_i < \boldsymbol{x}(\eta,t_i) \le x_i + \Delta x)\right)}{(\Delta x)^n}$$

A-7.6 Expected Value

– Mean value at the observed time instant

$$m_x(t_1) = E\left\{\boldsymbol{x}(\eta, t_1)\right\} = \underbrace{\int_{-\infty}^{+\infty} x f_{\boldsymbol{x}}(x,t_1)\,\mathrm{d}x}_{\text{for real-valued signals}}$$

– Normalized (instantaneous) power

$$S_x(t_1) = E\left\{|\boldsymbol{x}(\eta, t_1)|^2\right\} = \int_{-\infty}^{+\infty} |x|^2 f_{\boldsymbol{x}}(x,t_1)\,\mathrm{d}x$$

– Autocorrelation function (ACF)

$$\phi_{xx}(t_1,t_2) = E\left\{\boldsymbol{x}(\eta,t_1)\cdot\boldsymbol{x}^*(\eta,t_2)\right\}$$

$$\Rightarrow S_x(t_1) = \phi_{xx}(t_1,t_1)$$

Real-valued signals:

$$\phi_{xx}(t_1,t_2) = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} x_1 x_2 f_{\boldsymbol{x_1 x_2}}(x_1,x_2,t_1,t_2)\,\mathrm{d}x_1\mathrm{d}x_2$$

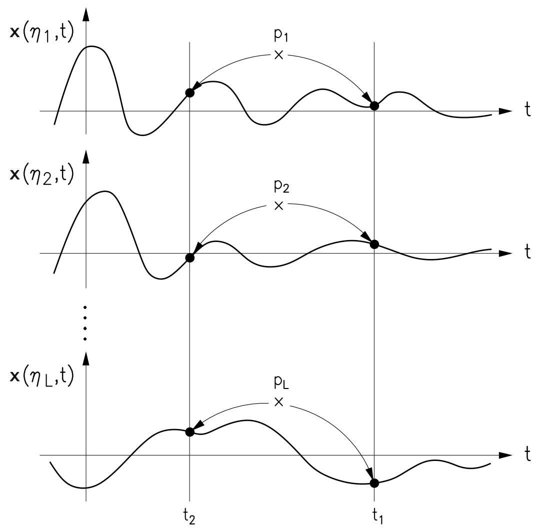
Complex-valued signals:

$$\phi_{xx}(t_1,t_2) =$$
$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}\left[(ac+bd)+\mathrm{j}(bc-ad)\right] f_{\boldsymbol{abcd}}(a,b,c,d)\,\mathrm{d}a\,\mathrm{d}b\,\mathrm{d}c\,\mathrm{d}d$$

with

$\boldsymbol{a} = \mathrm{Re}\{\boldsymbol{x}(\eta,t_1)\}$; $\boldsymbol{b} = \mathrm{Im}\{\boldsymbol{x}(\eta,t_1)\}$

$\boldsymbol{c} = \mathrm{Re}\{\boldsymbol{x}(\eta,t_2)\}$; $\boldsymbol{d} = \mathrm{Im}\{\boldsymbol{x}(\eta,t_2)\}$



$$\phi_{xx}(t_1,t_2) = \lim_{L\to\infty}\frac{1}{L}\sum_{i=1}^{L} p_i \quad \text{with} \quad p_i = x(\eta_i,t_1)\cdot x(\eta_i,t_2)$$

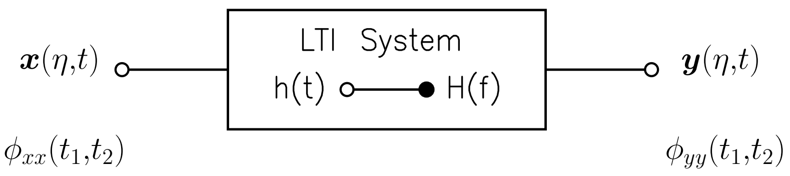
(real-valued signals)

– Crosscorrelation function (CCF) of joint random processes:

$$\phi_{xy}(t_1,t_2) = E\left\{\boldsymbol{x}(\eta,t_1)\cdot\boldsymbol{y}^*(\eta,t_2)\right\}$$

$$\Rightarrow S_{xy}(t_1) = \phi_{xy}(t_1,t_1) \quad \text{cross power}$$

A-7.7 Transfer of Stochastic Processes through LTI Systems



– Mean value function of the output signal:

$$\begin{aligned} m_y(t_1) &= E\left\{\boldsymbol{y}(\eta,t_1)\right\} = E\left\{\int\limits_{-\infty}^{+\infty} h(t_1-t')\,\boldsymbol{x}(\eta,t')\,\mathrm{d}t'\right\} \\ &= \int\limits_{-\infty}^{+\infty} h(t_1-t')\cdot E\left\{\boldsymbol{x}(\eta,t')\right\}\mathrm{d}t' = \int\limits_{-\infty}^{+\infty} h(t_1-t')\,m_x(t')\,\mathrm{d}t' \end{aligned}$$

$$\boxed{m_y(t_1) = m_x(t_1) * h(t_1)}$$

– Crosscorrelation function between input and output signals:

$$\phi_{xy}(t_1,t_2) = E\left\{\boldsymbol{x}(\eta,t_1)\cdot\int\limits_{-\infty}^{+\infty} h^*(t_2-t')\,\boldsymbol{x}^*(\eta,t')\,\mathrm{d}t'\right\} = \int\limits_{-\infty}^{+\infty} h^*(t_2-t')\,\phi_{xx}(t_1,t')\,\mathrm{d}t'$$

$$\begin{aligned} \phi_{xy}(t_1,t_2) &= \phi_{xx}(t_1,t_2) * h^*(t_2) \\ \phi_{yx}(t_1,t_2) &= \phi_{xx}(t_1,t_2) * h(t_1) \end{aligned}$$

– Autocorrelation function of the output signal:

$$\phi_{yy}(t_1,t_2) = E\left\{\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} h(t_1-t')h^*(t_2-t'')\,\boldsymbol{x}(\eta,t')\boldsymbol{x}^*(\eta,t'')\,\mathrm{d}t'\mathrm{d}t''\right\}$$

$$= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} h(t_1-t')h^*(t_2-t'')\phi_{xx}(t',t'')\,\mathrm{d}t'\mathrm{d}t''$$

$$\phi_{yy}(t_1,t_2) = \phi_{xx}(t_1,t_2) * h(t_1) * h^*(t_2)$$

A-7.8 Weakly Stationary Random Process

– Conditions for weakly stationarity of **real-valued** random processes $\boldsymbol{x}(\eta,t)$ and $\boldsymbol{y}(\eta,t)$:

1. $m_x(t) = m_x, \quad \forall t \in \mathbb{R}$, constant mean
2. $\phi_{xx}(t+\tau, t) =: \phi_{xx}(\tau), \quad \forall t \in \mathbb{R}$ Definition: $\phi_{xx}(\tau) = E\{\boldsymbol{x}(\eta, t+\tau) \cdot \boldsymbol{x}(\eta, t)\}$
   (Def., because $\phi_{xx}(\tau)$ was not introduced so far!)
   ACF and/or CCF depend only on the **time difference** $\tau = t_1 - t_2$,
   not on the absolute time t.
   Special case: $S_x = \phi_{xx}(0)$ (normalized) power of the random process
3. $\phi_{xy}(t+\tau, t) =: \phi_{xy}(\tau), \quad \forall t \in \mathbb{R}$
   Special case: $S_{xy} = \phi_{xy}(0)$ (normalized) cross power

Additional conditions for **complex-valued** random processes:

Notation: $\mathrm{Re}\{\boldsymbol{x}(\eta, t)\} = \boldsymbol{x}_{\mathrm{I}}(\eta, t)$ inphase component
$\mathrm{Im}\{\boldsymbol{x}(\eta, t)\} = \boldsymbol{x}_{\mathrm{Q}}(\eta, t)$ quadrature component

4. $m_x(t) \equiv 0, \quad \forall t$, zero mean value
5. $\phi_{x_\mathrm{I} x_\mathrm{I}}(\tau) = \phi_{x_\mathrm{Q} x_\mathrm{Q}}(\tau)$
   Real– and imaginary parts have the same ACF.

6. $\phi_{x_\mathrm{I} x_\mathrm{Q}}(\tau) = -\phi_{x_\mathrm{I} x_\mathrm{Q}}(-\tau)$
   Odd-symmetric CCF between the real and imaginary parts.

→ Weakly stationary complex-valued processes are therefore *rotational invariant*, i.e., $\forall \varphi \in \mathbb{R}$, $\boldsymbol{x}(\eta, t)$ and $\boldsymbol{x}(\eta, t)\, \mathrm{e}^{\mathrm{j}\varphi}$ have the same ACF.

Example: $\varphi = \frac{\pi}{2}: \quad \mathrm{j}\, \boldsymbol{x}(\eta, t) = -\boldsymbol{x}_\mathrm{Q}(\eta, t) + \mathrm{j}\, \boldsymbol{x}_\mathrm{I}(\eta, t)$
ACFs of the real and imaginary parts remain the same.
CCF: $E\{-\boldsymbol{x}_\mathrm{Q}(\eta, t+\tau) \cdot \boldsymbol{x}_\mathrm{I}(\eta, t)\} =$
$= -E\{\boldsymbol{x}_\mathrm{I}(\eta, t-\tau) \cdot \boldsymbol{x}_\mathrm{Q}(\eta, t)\} =$
$= -\phi_{x_\mathrm{I} x_\mathrm{Q}}(-\tau) = \phi_{x_\mathrm{I} x_\mathrm{Q}}(+\tau)$ because of odd symmetry.
CCF remains the same!

Remark:
A complex-valued random process, whose imaginary part is zero, is not stationary!

– Symmetry property: $\phi_{xx}(\tau) = \phi^*_{xx}(-\tau)$,
Special case: real-valued random processes:
$\phi_{xx}(\tau) = \phi_{xx}(-\tau), \phi_{xy}(\tau) = \phi_{yx}(-\tau)$

– *Power spectral density (PSD)* of weakly stationary random processes

Definition:

$$\Phi_{xx}(f) := \mathcal{F}\{\phi_{xx}(\tau)\} = \int\limits_{-\infty}^{+\infty} \phi_{xx}(\tau) \cdot \mathrm{e}^{-\mathrm{j}2\pi f\tau} \mathrm{d}\tau$$

Properties: $\Phi_{xx}(f) \in \mathbb{R} \quad ; \quad \Phi_{xx}(f) \geq 0 \quad \forall f \in \mathbb{R}$

Power of the random process:

$$S_x = \int\limits_{-\infty}^{+\infty} \Phi_{xx}(f)\,\mathrm{d}f$$

special case of real-valued random process:

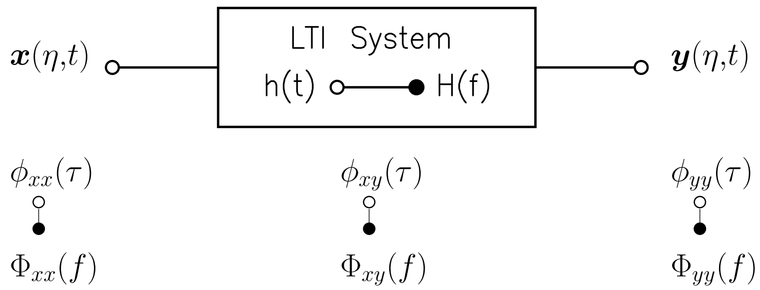
$\Phi_{xx}(f) = \Phi_{xx}(-f)$

– *Cross power spectral density (CPSD)* joint weakly stationary random processes

Definition:

$$\Phi_{xy}(f) := \mathcal{F}\{\phi_{xy}(\tau)\} = \int\limits_{-\infty}^{+\infty} \phi_{xy}(\tau) \cdot \mathrm{e}^{-\mathrm{j}2\pi f\tau} \mathrm{d}\tau$$

– Transfer of weakly stationary random processes through LTI system



With $t_1 = t + \tau$ and $t_2 = t$, we obtain

$$m_y = m_x \int\limits_{-\infty}^{+\infty} h(t)\,\mathrm{d}t = m_x \cdot H(0)$$

$$\begin{aligned}
\phi_{xy}(\tau) &= \phi_{xx}(\tau) * h^*(-\tau) \\
\phi_{yx}(\tau) &= \phi_{xx}(\tau) * h(\tau) \\
\phi_{yy}(\tau) &= \phi_{xx}(\tau) * h(\tau) * h^*(-\tau) = \phi_{xx}(\tau) * \varphi_{hh}(\tau)
\end{aligned}$$

The corresponding PSDs are given by

$$\Phi_{xy}(f) \;=\; \Phi_{xx}(f)\cdot H^*(f)$$
$$\Phi_{yx}(f) \;=\; \Phi_{xx}(f)\cdot H(f)$$
$$\Phi_{yy}(f) \;=\; \Phi_{xx}(f)\cdot |H(f)|^2$$

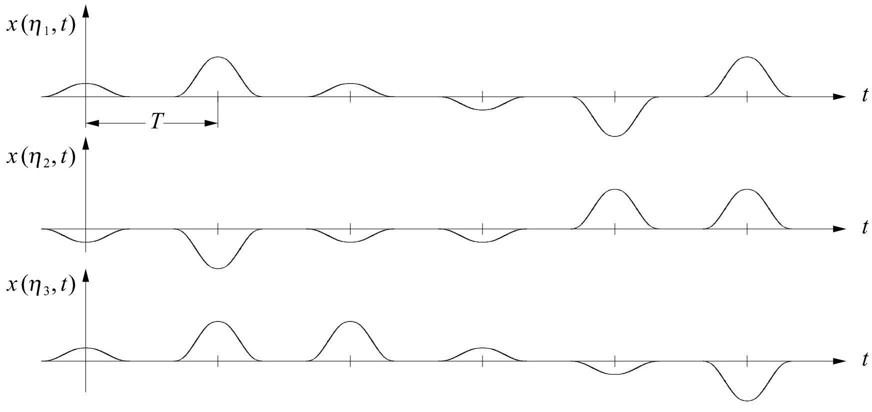
A-7.9 Weakly Cyclostationary Random Process (also known as periodic weakly stationary random process)

ACF: $$\phi_{xx}(t+\tau,t) \;=\; E\{\boldsymbol{x}(\eta,t+\tau)\boldsymbol{x}^*(\eta,t)\}$$
$$= E\{\boldsymbol{x}(\eta,t+kT+\tau)\boldsymbol{x}^*(\eta,t+kT)\} \;=\; \phi_{xx}(t+kT+\tau,t+kT), \quad \forall k\in\mathbb{Z}$$

T: Period, cycle of the weakly cyclostationary (periodic stationary) random process

Example: *PAM–signals with weakly stationary discrete-time sequence of amplitude coefficients*

$$\boldsymbol{x}(\eta,t) \;=\; \sum_{k=-\infty}^{+\infty} a[k]g(t-kT) \quad \text{with} \quad E\{a[k+\nu]a^*[k]\} \;=\; \phi_{aa}[\nu], \quad \forall k\in\mathbb{Z}$$



Average ACF: Averaging over one period $\Rightarrow$ Time independent!

$$\bar{\phi}_{xx}(\tau) := \frac{1}{T} \int_{-T/2}^{+T/2} \phi_{xx}(t+\tau,t)\,\mathrm{d}t$$

Average power spectral density

$$\bar{\Phi}_{xx}(f) = \mathcal{F}\{\bar{\phi}_{xx}(\tau)\} = \frac{1}{T} \int_{-T/2}^{+T/2} \mathcal{F}_\tau\{\phi_{xx}(t+\tau,t)\}\,\mathrm{d}t$$

Interpretation:

For each sample funtion $\boldsymbol{x}(\eta,t)$, all time-shifted functions $\boldsymbol{x}(\eta,t-\lambda T)$, with λ being uniformally distributed in the interval $[0,1)$, i.e., PDF $f_\lambda(\lambda) = \mathrm{rect}\left(\lambda - \frac{1}{2}\right)$, are included in the set of sample functions to enforce weakly stationarity $\Rightarrow$ **Phase randomization**.

The average ACF $\bar{\phi}_{xx}(\tau)$ and average PSD $\bar{\Phi}_{xx}(f)$ are the ACF and PSD of these weakly stationary auxiliary random processes.

572

A-8 Gaussian Processes

According to the central limit theorem of statistics, a stochastic process that arises from the sum of many identical, but statistically independent subprocesses, is a Gaussian process.

$\Rightarrow$ Gaussian processes are very suitable for modelling of noise and interference.

Gaussian processes are, for a given power spectral density, the "most random" of all random processes.

$\Rightarrow$ Interference by a Gaussian process is the "worst" interference.

$\Rightarrow$ Gaussian processes are also suitable for modelling (useful) signals.

A-8.1 Definition of a Gaussian Process

A real-valued random process $\boldsymbol{x}(\eta,t)$ is a Gaussian process, if for all natural numbers N, the N-dimensional joint PDF of random variables $\vec{x} = (x_1,x_2,\ldots,x_N)$ at time instants $\vec{t} = (t_1,t_2,\ldots,t_N)$ are given by

$$f_{\boldsymbol{x}}(\vec{x},\vec{t}) = \frac{1}{\sqrt{(2\pi)^N \det\{\mathbf{M}\}}} \cdot \exp\left(-\frac{1}{2}(\vec{x}-\vec{m})\mathbf{M}^{-1}(\vec{x}-\vec{m})^T\right)$$

with

$$\vec{m} = (m_x(t_1),m_x(t_2),\ldots,\, m_x(t_N))$$
$$\mathbf{M} = (\mu_{ik}) \qquad \forall i,k \in \{1,2,\ldots,\, N\}$$

and

$$\begin{aligned}\mu_{ik} &= E\{(\boldsymbol{x}(\eta,t_i) - m_x(t_i))(\boldsymbol{x}(\eta,t_k) - m_x(t_k))\} \\ &= \phi_{xx}(t_i,t_k) - m_x(t_i)m_x(t_k)\end{aligned}$$

$\mathbf{M}$ is called the *covariance matrix* for the time instants $t_1, t_2, \ldots, t_N$.

Remark:
All joint densities for Gaussian processes are completely described by the **first** and **second** order moments (expected values). Thus, a Gaussian process is completely specified by the mean function $m_x(t)$ and the ACF $\phi_{xx}(t_1,t_2)$!

Special case:
Weakly stationary Gaussian processes:

$$m_x(t) = m_x; \qquad \phi_{xx}(t+\tau,t) = \phi_{xx}(\tau)\ \forall t$$

$\mathbf{M}$ does **not** depend on the absolute time, but ratner only on the time difference.

$\Rightarrow$ A weakly stationary Gaussian process is thus strictly stationary as well!

Example:

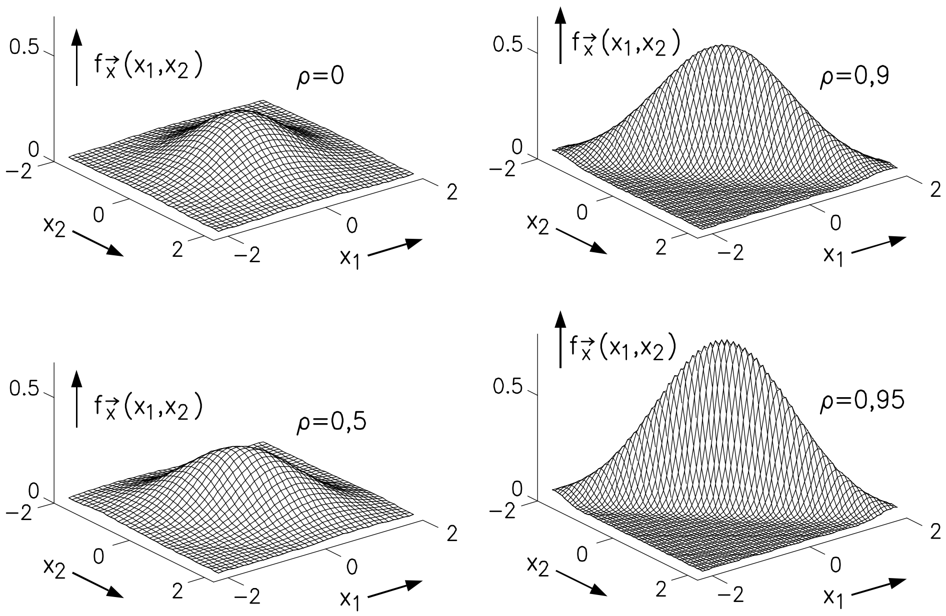
$N = 1$

$$f_{\boldsymbol{x}}(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \cdot \exp\left(-\frac{(x-m)^2}{2\sigma_x^2}\right)$$

$$N = 2: \quad \mathbf{M} = \sigma_x^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \vec{m} = (0{,}0)$$

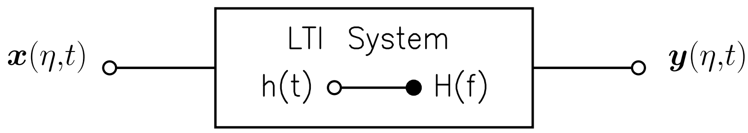
ρ : Correlation coefficient

$$f_{\boldsymbol{x}}(x_1, x_2) = \frac{1}{2\pi\sigma_x^2\sqrt{1-\rho^2}} \cdot \exp\left(-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2\sigma_x^2(1-\rho^2)}\right)$$

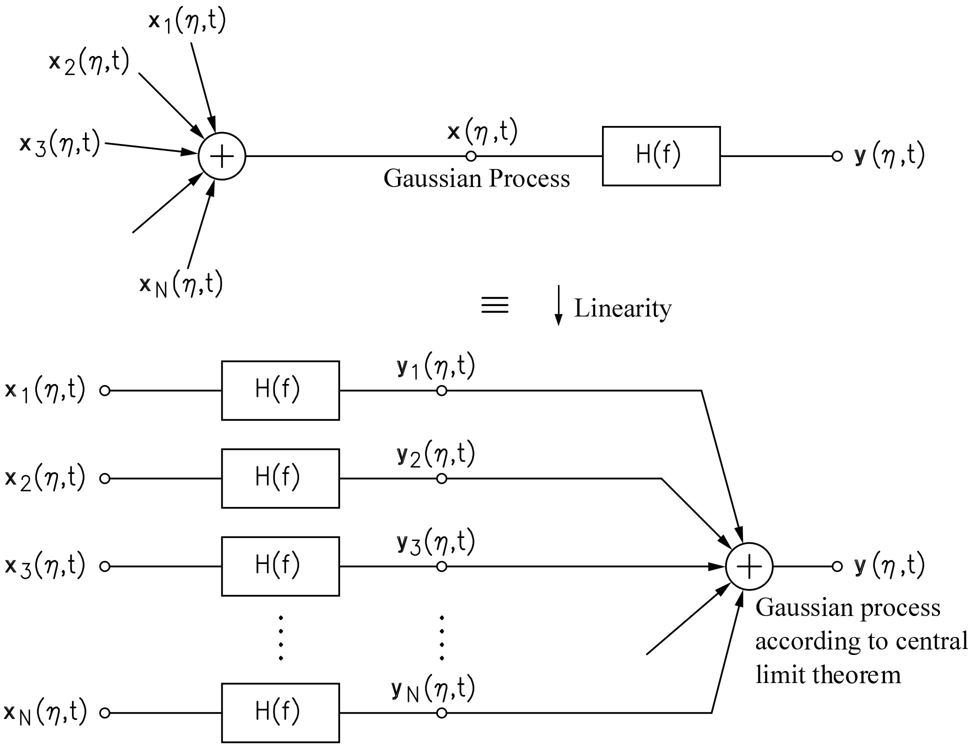


2–dim. Gaussian joint densities ($\sigma_x^2 = 1$)

A-8.2 Transfer of a Random Process through a LTI System



Model for Gaussian process using the central limit theorem for vector random variables:



A Gaussian process remains Gaussian after passing through a LTI system, and also the output process can be modelled as a sum of independent sub-processes. The mean function

$$m_y(t) = m_x(t) * h(t)$$

and the ACF

$$\phi_{yy}(t_1, t_2) = \phi_{xx}(t_1, t_2) * h(t_1) * h^*, (t_2)$$

fully specify the output process.

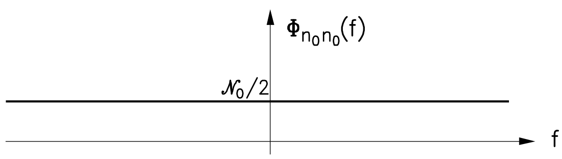
Note: This only holds for Gaussian random processes, not in general.

A-8.3 White Gaussian Noise

A stationary Gaussian process $\boldsymbol{n}_0(\eta, t)$ with constant PSD

$$\Phi_{n_0 n_0}(f) \;=\; N_0/2 \quad \forall f \in \mathbb{R}$$

is referred to as white (Gaussian) noise.



N_0 : <u>one</u>-sided noise power spectral density (historical reasons)

$N_0/2$: <u>two</u>-sided noise power spectral density

$\mathrm{ACF} : \phi_{n_0 n_0}(\tau) = \frac{N_0}{2}\,\delta(\tau)$

Power: $\phi_{n_0 n_0}(0) \rightarrow \infty \Rightarrow$ White noise does not exist!!!

But white noise is an appropriate model for noise and/or interference.

White Gaussian noise is the most "random" of all random processes! (cf. Information Theory)

Example: Thermal resistance noise

$$N_0/2 = 2\,\mathrm{k}\,T_K \cdot R \quad \text{with}$$

$$\begin{array}{ll} \mathrm{k} = 1{,}38 \cdot 10^{-23}\frac{\mathrm{Ws}}{\mathrm{K}} & \text{Boltzmann constant} \\ T_K & \text{absolute temperature in K} \\ R & \text{resistance} \end{array}$$

Note: Constant PSD $\frac{N_0}{2}$ holds for up to about 300 GHz only!!

Remark:

Distinguish **continuous-time** white noise from **discrete-time** white noise:

$$\boldsymbol{n}_0[\eta,k] \quad \text{with} \quad \begin{array}{l} \phi_{n_0n_0}[\kappa] = \sigma_n^2 \cdot \delta[\kappa] \\ \Phi_{n_0n_0}(\mathrm{e}^{\mathrm{j}2\pi fT}) = \sigma_n^2 \end{array}$$

Discrete-time white noise does exist (finite power).

A-8.4 Complex Stationary Gaussian Processes

For a complex-valued Gaussian random variable $\boldsymbol{z} = \boldsymbol{x} + \mathrm{j}\boldsymbol{y}$ with mutually uncorrelated real– and imaginary parts, i.e., $\sigma_{xy}^2 = 0$, which have identical variances

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2/2,$$

the joint PDF is given by

$$f_{\boldsymbol{xy}}(x,y) = \frac{1}{2\pi\sigma_x^2}\mathrm{e}^{-((x-m_x)^2+(y-m_y)^2)/(2\sigma_x^2)}.$$

This joint PDF can be formally expressed with respect to complex random variable $\boldsymbol{z}$:

$$f_{\boldsymbol{z}}(z) = \frac{1}{\pi\sigma_z^2}\mathrm{e}^{-|z-m_z|^2/\sigma_z^2}$$

with $m_z = m_x + \mathrm{j}m_y$.

Definition:

A complex weakly stationary random process $\boldsymbol{z}(\eta, t)$ as defined in Chapter A-7.8, is a *complex-valued stationary Gaussian process*, if the joint PDF of the real and imaginary parts of n random variables

$$\boldsymbol{z}_k = \boldsymbol{z}(\eta, t_k),$$

observed at n arbitrary points in time, is a $2n$–dimensional Gaussian PDF.

– $2n$–dim. joint PDF

$$f_{\vec{\boldsymbol{x}},\vec{\boldsymbol{y}}}(\vec{x},\vec{y}) = \frac{1}{(2\pi)^n\sqrt{\det\{\mathbf{N}\}}}\exp\left(-\frac{1}{2}(\vec{x},\vec{y})\mathbf{N}^{-1}(\vec{x},\vec{y})^T\right)$$

$\mathbf{N}$: $2n\times 2n$ covariance matrix, decomposed in $n\times n$ matrices:

$$\mathbf{N} := \begin{pmatrix} \mathbf{N}_{xx} & \mathbf{N}_{xy} \\ \\ \mathbf{N}_{yx} & \mathbf{N}_{yy} \end{pmatrix}$$

From rotational invariance results:

From $\phi_{xx}(\tau) = \phi_{yy}(\tau) = \phi_{xx}(-\tau)$, we obtain

$$\mathbf{N}_{xx} = \mathbf{N}_{yy} = \mathbf{N}_{xx}^T = \mathbf{N}_{yy}^T \quad \text{symmetric}$$

From $\phi_{xy}(\tau) = \phi_{yx}(-\tau) = -\phi_{xy}(-\tau) = -\phi_{yx}(\tau)$, we obtain

$$\mathbf{N}_{xy} = -\mathbf{N}_{yx} = -\mathbf{N}_{xy}^T = \mathbf{N}_{yx}^T \quad \text{skew-symmetric}$$

– **Complex-valued** $n\times n$ covariance matrix

$$\mathbf{M} = (E\{\boldsymbol{z}_\ell \boldsymbol{z}_k^H\}) = \mathrm{Re}\{\mathbf{M}\} + \mathrm{j}\,\mathrm{Im}\{\mathbf{M}\} \quad \text{for} \quad k,\ell = 1(1)n$$

$$\begin{aligned} E\{\boldsymbol{z}_\ell \boldsymbol{z}_k^H\} &= E\{(\boldsymbol{x}_\ell + \mathrm{j}\,\boldsymbol{y}_\ell)(\boldsymbol{x}_k^T - \mathrm{j}\,\boldsymbol{y}_k^T)\} = \\ &= E\{\boldsymbol{x}_\ell \boldsymbol{x}_k^T\} + E\{\boldsymbol{y}_\ell \boldsymbol{y}_k^T\} + \mathrm{j}\,(E\{\boldsymbol{y}_\ell \boldsymbol{x}_k^T\} - E\{\boldsymbol{x}_\ell \boldsymbol{y}_k^T\}) = \\ &= 2\,\underbrace{E\{\boldsymbol{x}_\ell \boldsymbol{x}_k^T\}}_{\mathbf{N}_{xx}} - \mathrm{j}\,2\,\underbrace{E\{\boldsymbol{x}_\ell \boldsymbol{y}_k^T\}}_{\mathbf{N}_{xy}} \end{aligned}$$

$$\Rightarrow \mathbf{M}^T = \mathbf{M}^* \text{ Hermitian symmetric}$$

Compare with $2n\times 2n$ covariance matrix

$$\mathbf{N} = \frac{1}{2}\begin{pmatrix} \mathrm{Re}\{\mathbf{M}\} & -\mathrm{Im}\{\mathbf{M}\} \\ \\ \mathrm{Im}\{\mathbf{M}\} & \mathrm{Re}\{\mathbf{M}\} \end{pmatrix}$$

Thus, with the help of the $n\times n$ complex covariance matrices, the following **formal** notation holds for the $2n$–dim. joint PDF of n complex Gaussian random variables

$$f_{\vec{\boldsymbol{z}}}(\vec{z}) = \frac{1}{\pi^n|\det\{\mathbf{M}\}|}\exp(-\vec{z}\,\mathbf{M}^{-1}\vec{z}^{*T})$$

A-9 Equivalent Complex Baseband Signals and Systems

Reasons for transformation of signals to equivalent complex baseband signals (ECB signals):

– Handling modulation and RF transmission schemes independent of the carrier frequency

– Unified treatment of (digital) baseband and carrier-modulated transmission schemes

– Simplified mathematical representation and calculation methods for carrier-modulated transmission schemes

A-9.1 Transformation of Signals

Physical (radio frequency) signals $x_{\mathrm{RF}}(t)$ are real.

Thereby, the following applies for the spectrum $X_{\mathrm{RF}}(f) = \mathcal{F}\{x_{\mathrm{RF}}(t)\}$ of physical signals:

$$X_{\mathrm{RF}}(f) = X^*_{\mathrm{RF}}(-f)$$

$\Rightarrow$ One side of the two-sided spectrum of real-valued signals is redundant.

Analytical Signal
By suppressing the spectrum for $f < 0$ and doubling the spectrum for $f > 0$, the *analytical signal* $x^+_{\mathrm{RF}}(t)$ is formed from the real physical signal $x_{\mathrm{RF}}(t)$.

$$X^+_{\mathrm{RF}}(f) = (1 + \mathrm{sign}(f)) \cdot X_{\mathrm{RF}}(f) \text{ with}$$

$$\mathrm{sign}(f) = \begin{cases} 1 & \text{for} \quad f > 0 \\ 0 & \text{for} \quad f = 0 \\ -1 & \text{for} \quad f < 0 \end{cases}$$

Fourier transform

$$\mathrm{sign}(f) \bullet\!\!-\!\!\circ \frac{\mathrm{j}}{\pi t} \quad ; \quad \mathrm{sign}(t) \circ\!\!-\!\!\bullet \frac{1}{\mathrm{j}\pi f}$$

$$\Rightarrow x^+_{\mathrm{RF}}(t) = x_{\mathrm{RF}}(t) + \mathrm{j}x_{\mathrm{RF}}(t) * \frac{1}{\pi t} = x_{\mathrm{RF}}(t) + \mathrm{j}\,\mathcal{H}\{x_{\mathrm{RF}}(t)\}$$

with

$$\mathcal{H}\{y(t)\} := y(t) * \frac{1}{\pi t} = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{y(t')}{t - t'}\mathrm{d}t' \quad \text{Hilbert Transform}$$

The analytical signal has twice the energy of the physical signal, since

$$\int_0^{\infty} |2X_{\mathrm{RF}}(f)|^2 \,\mathrm{d}f = 2\int_{-\infty}^{+\infty} |X_{\mathrm{RF}}(f)|^2 \,\mathrm{d}f$$

holds.

Equivalent Complex Baseband Signal (ECB Signal)

Definition:

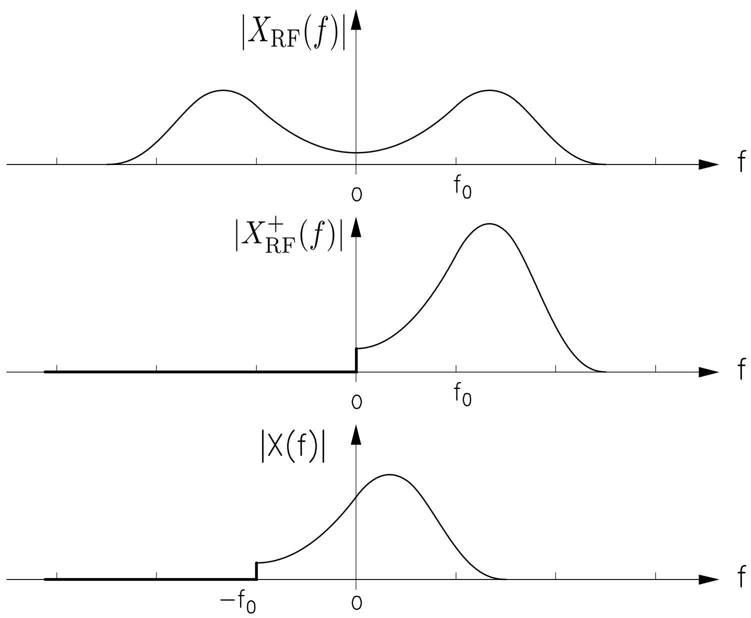
ECB signal $x(t)$ of physical signal $x_{\mathrm{RF}}(t)$:

$$x(t) = \frac{1}{\sqrt{2}}\,x_{\mathrm{RF}}^{+}(t)\cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 t} = \frac{1}{\sqrt{2}}\left(x_{\mathrm{RF}}(t) + \mathrm{j}\,\mathcal{H}\left\{x_{\mathrm{RF}}(t)\right\}\right)\cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 t}$$

$$\circ\!\!-\!\!\!-\!\!\bullet$$

$$X(f) = \frac{1}{\sqrt{2}}\,X_{\mathrm{RF}}^{+}(f+f_0) = \frac{1}{\sqrt{2}}\left(1 + \mathsf{sign}(f+f_0)\right)\cdot X_{\mathrm{RF}}(f+f_0)$$

588



The signal transformation $x_{\mathrm{RF}}(t) \rightarrow x(t)$ is linear.

f_0: Transformation frequency

Inverse transform:

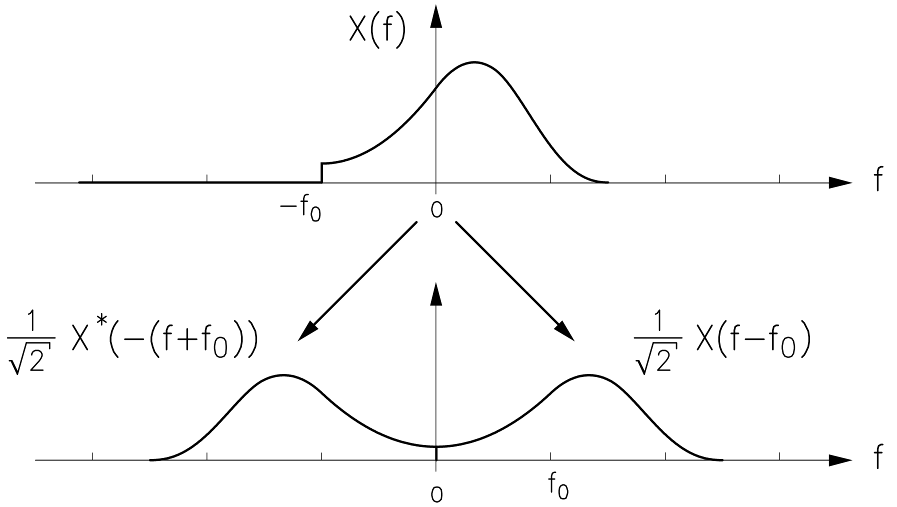
$$x_{\mathrm{RF}}(t) = \sqrt{2}\,\mathrm{Re}\left\{x(t)\mathrm{e}^{+\mathrm{j}2\pi f_0 t}\right\}$$

With $\mathrm{Re}\{z\} = \frac{1}{2}(z + z^*)$, one gets:

$$x_{\mathrm{RF}}(t) = \frac{\sqrt{2}}{2}\left(x(t)\cdot \mathrm{e}^{\mathrm{j}2\pi f_0 t} + x^*(t)\cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 t}\right)$$

$$\circ\!\!-\!\!\!-\!\!\bullet$$

$$X_{\mathrm{RF}}(f) = \frac{\sqrt{2}}{2}\left(X(f - f_0) + X^*(-(f + f_0))\right)$$



– Remark 1:

By introducing the factor $\frac{1}{\sqrt{2}}$ for the transformation to the ECB and $\sqrt{2}$ for the inverse transformation, the physical signal $x_{\mathrm{RF}}(t)$ and the ECB signal have the **same energy**.

$$\int_{-\infty}^{+\infty} |X_{\mathrm{RF}}(f)|^2 \,\mathrm{d}f = \int_{-\infty}^{+\infty} |X(f)|^2 \,\mathrm{d}f = \int_{-f_0}^{+\infty} \left|\frac{2}{\sqrt{2}} X_{\mathrm{RF}}(f + f_0)\right|^2 \mathrm{d}f$$

In this way, many confusion that occur frequently in the literature are avoided!

– Remark 2:

The most important application of system theory for *complex-valued* signals in the field of telecommunications is the use of ECB signals.

– Remark 3:

The spectrum of ECB signals for $f < -f_0$ is zero.

Quadrature Components of ECB Signals

$$x(t) = x_{\mathrm{I}}(t) + \mathrm{j}\, x_{\mathrm{Q}}(t) \quad \text{with}$$

$$\left.\begin{aligned} x_{\mathrm{I}}(t) &= \mathrm{Re}\{x(t)\} && \underline{\text{i}}\text{nphase component} \\ x_{\mathrm{Q}}(t) &= \mathrm{Im}\{x(t)\} && \underline{\text{q}}\text{uadrature component} \end{aligned}\right\} \text{quadrature components}$$

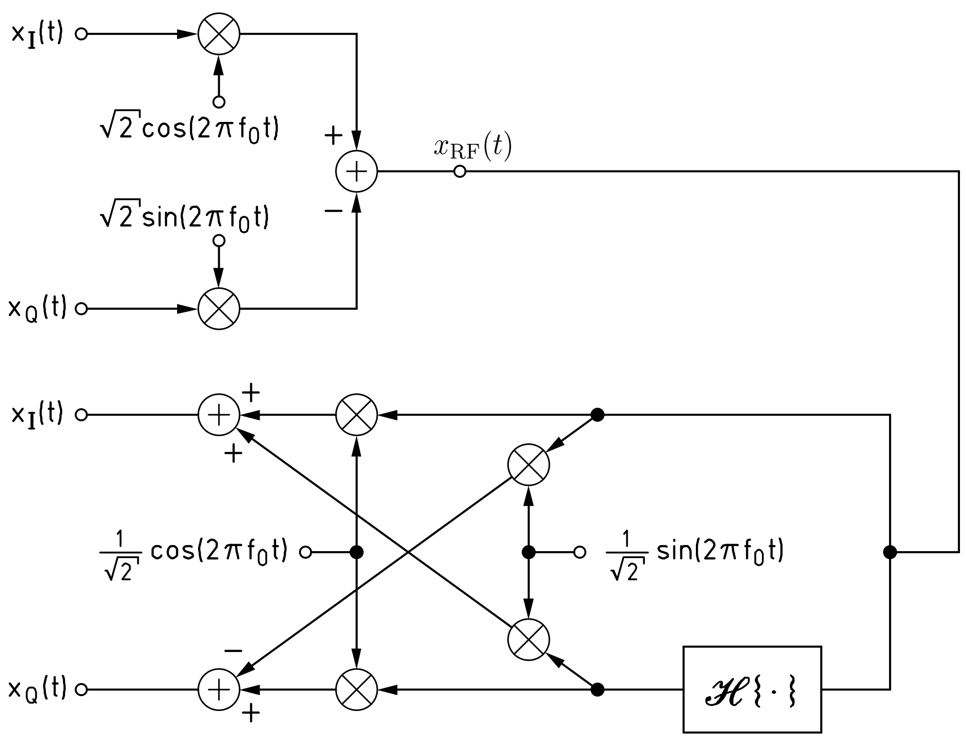
$$x_{\mathrm{I}}(t) = \frac{1}{\sqrt{2}} \left(x_{\mathrm{RF}}(t) \cdot \cos(2\pi f_0 t) + \mathcal{H}\{x_{\mathrm{RF}}(t)\} \cdot \sin(2\pi f_0 t) \right)$$

$$x_{\mathrm{Q}}(t) = \frac{1}{\sqrt{2}} \left(\mathcal{H}\{x_{\mathrm{RF}}(t)\} \cdot \cos(2\pi f_0 t) - x_{\mathrm{RF}}(t) \cdot \sin(2\pi f_0 t) \right)$$

Generation of the physical signal from the quadrature components

$$\begin{aligned} x_{\mathrm{RF}}(t) &= \sqrt{2}\, \mathrm{Re}\{(x_{\mathrm{I}}(t) + \mathrm{j} x_{\mathrm{Q}}(t)) \cdot \mathrm{e}^{\mathrm{j}2\pi f_0 t}\} \\ &= \sqrt{2} \left(x_{\mathrm{I}}(t) \cdot \cos(2\pi f_0 t) - x_{\mathrm{Q}}(t) \sin(2\pi f_0 t) \right) \end{aligned}$$

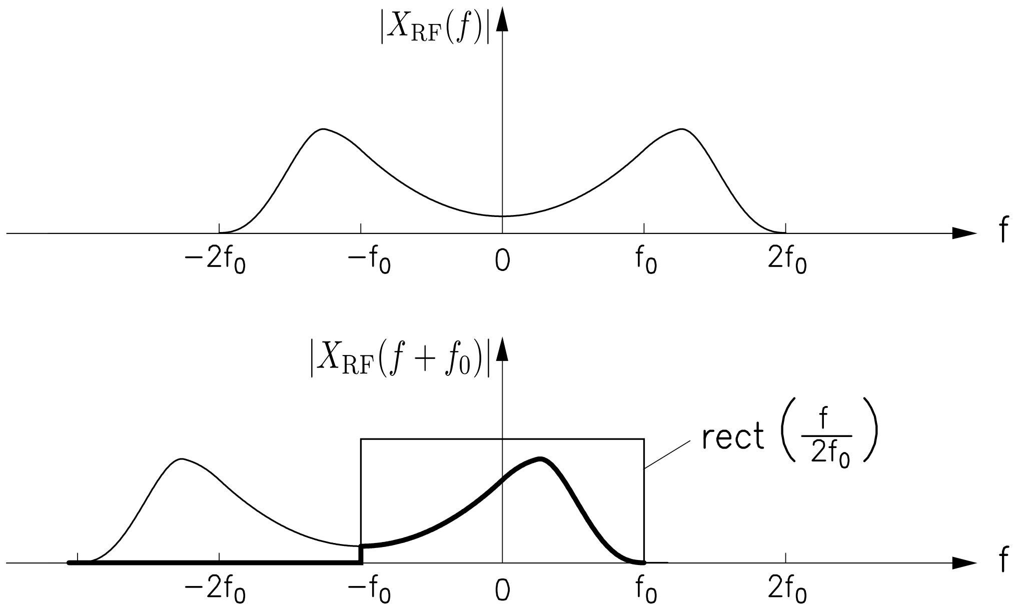
592



Simplified extraction of the quadrature components if

$$X_{\mathrm{RF}}(f) = 0 \text{ for} |f| > \underline{\underline{2}} \cdot f_0$$

holds.



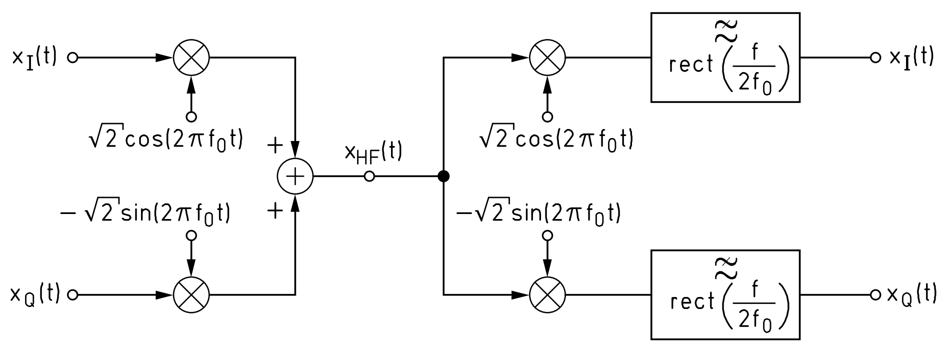
Instead of the sign–function, a lowpass filter with cut-off frequency $f_g = f_0$ can be used to suppress the redundant portion of the previously shifted spectrum,

$$X(f) = \sqrt{2} X_{\mathrm{RF}}(f + f_0) \cdot \mathrm{rect}\left(\frac{f}{2f_0}\right)$$

$$\bullet\!\!-\!\!\circ$$

$$x(t) = \sqrt{2}\,(x_{\mathrm{RF}}(t) \cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 t}) * (2f_0 \cdot \mathrm{si}(2\pi f_0 t))$$

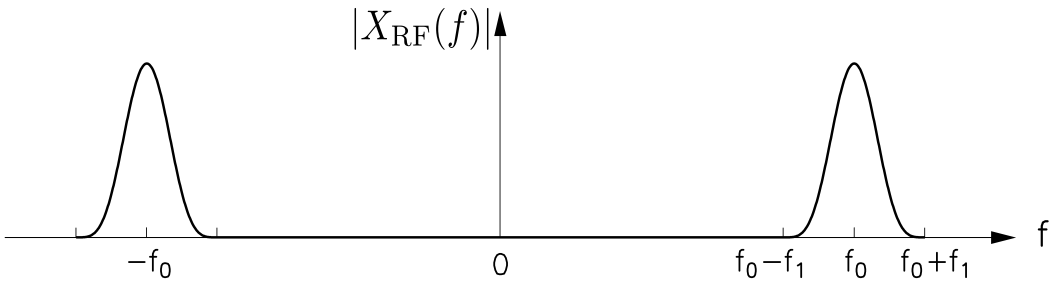
$$\Rightarrow \begin{aligned} x_{\mathrm{I}}(t) &= 2\sqrt{2} f_0\,(x_{\mathrm{RF}}(t) \cdot \cos(2\pi f_0 t)) * \mathrm{si}(2\pi f_0 t) \\ x_{\mathrm{Q}}(t) &= -2\sqrt{2} f_0 (x_{\mathrm{RF}}(t) \cdot \sin(2\pi f_0 t)) * \mathrm{si}(2\pi f_0 t) \end{aligned}$$



If

$$X(f) \approx 0 \quad \text{for} \quad ||f| - f_0| > f_1 \text{ und } f_1 < f_0$$

holds, then a lowpass filter with a flat transition between the passband $(|f| \leq f_1)$ and the stopband $(|f| > 2f_0 - f_1)$ can be used to extract the quadrature components:
Bandpass signal



Representation of ECB Signals in Polar Coordinates:

Let:

$$x(t) = |x(t)| \cdot \mathrm{e}^{\mathrm{j}\varphi_x(t)}$$

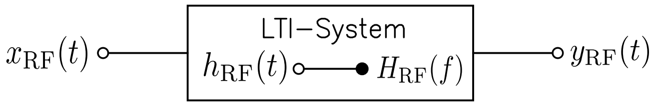
with

$$\begin{aligned} |x(t)| &= \sqrt{x_{\mathrm{I}}^2(t) + x_{\mathrm{Q}}^2(t)} \\ \varphi_x(t) &= \arg(x(t)) \\ &= \arctan\left(\frac{x_{\mathrm{Q}}(t)}{x_{\mathrm{I}}(t)}\right) + \frac{\pi}{2}(1 - \mathrm{sign}(x_{\mathrm{I}}(t))) \end{aligned}$$

Definition:
The signal $|x(t)|$ is referred to as the *envelope* of the signal $x(t)$.
The envelope is independent of the transformation frequency f_0!

Remark:
The envelope of the physical signal $x_{\mathrm{RF}}(t)$, which represents a modulated $\cos$–wave of a relatively narrow-band bandpass signal, is greater than the envelope of the equivalent ECB signal by a factor of $\sqrt{2}$!

A-9.2 Transformation of LTI Systems



Physical LTI system: Impulse response $h_{\mathrm{RF}}(t) \in \mathbb{R}$

Transfer function $H_{\mathrm{RF}}(f)$

Definition of ECB systems

$$h(t) = \frac{1}{\underline{\underline{\underline{2}}}}\left(h_{\mathrm{RF}}(t) + \mathrm{j}\mathcal{H}\{h_{\mathrm{RF}}(t)\}\right) \cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 t}$$

$$H(f) = \frac{1}{\underline{\underline{\underline{2}}}}(1 + \mathrm{sign}(f + f_0))H_{\mathrm{RF}}(f + f_0)$$

Inverse transform

$$\begin{aligned} h_{\mathrm{RF}}(t) &= \underline{\underline{2}}\,\mathrm{Re}\left\{h(t) \cdot \mathrm{e}^{\mathrm{j}2\pi f_0 t}\right\} \\ H_{\mathrm{RF}}(f) &= H(f - f_0) + H^*(-(f + f_0)) \end{aligned}$$

Remark:

Note here the prefactors $1/2$ and 2 in contrast to the transformation of signals ($1/\sqrt{2}$ and $\sqrt{2}$)!!

Transmission of signals over a LTI system:

a) Physical signals:

$$y_{\mathrm{RF}}(t) = x_{\mathrm{RF}}(t) * h_{RF}(t)$$

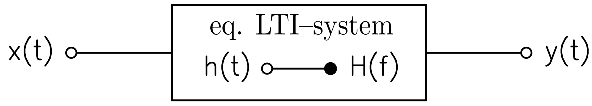
$$Y_{\mathrm{RF}}(f) = X_{\mathrm{RF}}(f) \cdot H_{\mathrm{RF}}(f)$$

are equivalent to:

b) ECB signals:

$$y(t) = x(t) * h(t)$$

$$Y(f) = X(f) \cdot H(f)$$

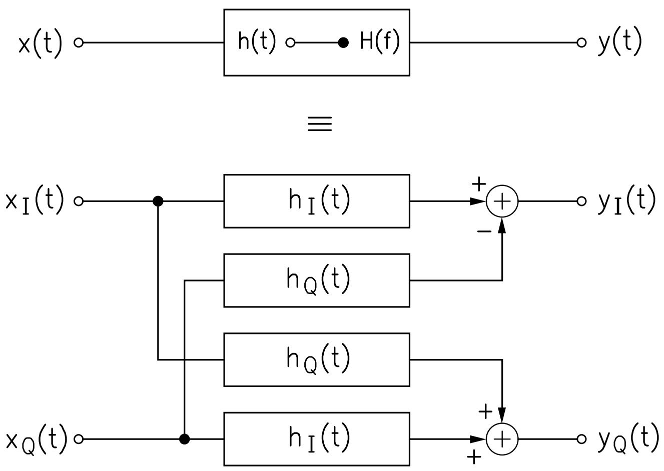


Proof:

$$
\begin{aligned}
Y_{\mathrm{RF}}(f) &= X_{\mathrm{RF}}(f) \cdot H_{\mathrm{RF}}(f) \\
&= \frac{1}{\sqrt{2}}\left(X(f-f_0) + X^*(-(f+f_0)\right) \cdot \left(H(f-f_0) + H^*(-(f+f_0))\right) \\
&= \frac{1}{\sqrt{2}}\left(X(f-f_0) \cdot H(f-f_0) + X^*(-(f+f_0))H^*(-(f+f_0))\right) + \\
&\quad + \frac{1}{\sqrt{2}}(\underbrace{X(f-f_0)}_{f>0} \cdot \underbrace{H^*(-(f+f_0))}_{f<0} + \underbrace{X^*(-(f+f_0))}_{f<0} \cdot \underbrace{H(f-f_0))}_{f>0} \qquad \neq 0 \text{ for:} \\
&= \frac{1}{\sqrt{2}}\left(Y(f-f_0) + Y^*(-(f+f_0))\right) = Y_{\mathrm{RF}}(f) \qquad \text{q.e.d.}
\end{aligned}
$$

An ECB system $h(t) = h_{\mathrm{I}}(t) + \mathrm{j}h_{\mathrm{Q}}(t)$ comprises four real LTI systems with respect to the quadrature components:

$$
y(t) = x_{\mathrm{I}}(t) * h_{\mathrm{I}}(t) - x_{\mathrm{Q}}(t) * h_{\mathrm{Q}}(t) + \mathrm{j}(x_{\mathrm{I}}(t) * h_{\mathrm{Q}}(t) + x_{\mathrm{Q}}(t) * h_{\mathrm{I}}(t))
$$



– Remark 1:

  For other signal transforms such as linear filtering and frequency conversion, the use of ECB signals and systems should be considered with extreme caution. For example, an ECB signal of the product of two signals does not correspond to the product of the associated ECB signals! On the other hand, some forms of **dispersive** nonlinear systems can be conveniently modelled by equivalent **disperionless** nonlinear systems in terms of ECB signals. E.g., description of non-linear dispersive RF amplifiers by means of AM/AM and AM/PM characteristics for ECB signals!

–

Remark 2:

Since the spectrum $X(f)$ of the ECB input signal disappears by definition for $f < -f_0$, limiting the ECB transfer function to $H(f) = 0$ for $f < -f_0$ is formally not necessary.

For an *alternative system*

$$H'(f) = H_{\mathrm{RF}}(f + f_0)$$

one gets in principle the same ECB output signal as for the actual system $H(f) = \frac{1}{2}(1+\mathrm{sign}(f + f_0))H(f + f_0)$. Using such an alternative system, one often gets a simplified ECB representation.

Example: Delay element

$$\begin{aligned} h_{\mathrm{RF}}(t) &= \delta(t-t_0)\,;\; H_{\mathrm{RF}}(f) = \mathrm{e}^{-\mathrm{j}2\pi f t_0} \\ H(f) &= \tfrac{1}{2}(1+\mathrm{sign}(f+f_0)) \cdot \mathrm{e}^{-\mathrm{j}2\pi(f+f_0)t_0} \\ h(t) &= \tfrac{1}{2}\left(\delta(t-t_0) + \frac{\mathrm{j}}{\pi(t-t_0)}\right) \cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 t_0} \end{aligned}$$

Alternative system

$$\begin{aligned} H'(f) &= \mathrm{e}^{-\mathrm{j}2\pi(f+f_0)t_0} \\ h'(t) &= \delta(t-t_0) \cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 t_0} \end{aligned}$$

Note: If the input signal is not suppressed at frequencies $f < -f_0$, the alternative system can not be applied.

A-9.3 Transformation of Weakly Stationary Random Processes

A-9.3.1 Transformation of the ACF and the PSD

Let $\boldsymbol{x}_{\mathrm{RF}}(\eta, t)$ be a real, weakly stationary random process with ACF $\phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(\tau)$.

ECB transformation of the sample functions:

$$\boldsymbol{x}(\eta, t) := \frac{1}{\sqrt{2}}\left(\boldsymbol{\alpha}(t) + \mathrm{j}\,\boldsymbol{\beta}(t)\right)\mathrm{e}^{-\mathrm{j}2\pi f_0 t}$$

with abbreviation $\boldsymbol{\alpha}(t) = \boldsymbol{x}_{\mathrm{RF}}(\eta, t),\ \boldsymbol{\beta}(t) = \mathcal{H}\left\{\boldsymbol{x}_{\mathrm{RF}}(\eta, t)\right\}$

Mean value: $m_x(t) = \dfrac{1}{\sqrt{2}} m_{x_{\mathrm{RF}}} \cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 t}$

⇒ The ECB process $\boldsymbol{x}(\eta,t)$ is weakly stationary only if the real-valued process $\boldsymbol{x}_{\mathrm{RF}}(\eta, t)$ has **zero mean value** $(m_{x_{\mathrm{RF}}} = m_x = 0)$.

ACF of the transformed random process:

$$\phi_{xx}(\tau) = E\{\boldsymbol{x}(\eta, t+\tau) \cdot \boldsymbol{x}^*(\eta, t)\} =$$

$$= \frac{1}{2}E\{(\boldsymbol{\alpha}(t+\tau) + \mathrm{j}\,\boldsymbol{\beta}(t+\tau)) \cdot (\boldsymbol{\alpha}(t) - \mathrm{j}\,\boldsymbol{\beta}(t)) \cdot \mathrm{e}^{-\mathrm{j}2\pi f_0(t+\tau)} \cdot \mathrm{e}^{+\mathrm{j}2\pi f_0 t}\} =$$

$$= \frac{1}{2}\, \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau} (\phi_{\alpha\alpha}(\tau) + \phi_{\beta\beta}(\tau) + \mathrm{j}\,(\phi_{\beta\alpha}(\tau) - \phi_{\alpha\beta}(\tau)))$$

The Hilbert transform in the time domain corresponds to the filtering of the signal $\boldsymbol{\alpha}(t)$ with a system whose transfer function is $-\mathrm{j} \cdot \mathrm{sign}(f)$, i.e.,

$$\boldsymbol{\beta}(t) \circ\!\!-\!\!\!-\!\!\bullet\; \mathcal{F}\{\boldsymbol{\alpha}(t)\} \cdot (-\mathrm{j} \cdot \mathrm{sign}(f)).$$

Hence, we have

A)

$$\Phi_{\beta\beta}(f) = \Phi_{\alpha\alpha}(f) \cdot |-\mathrm{j} \cdot \mathrm{sign}(f)|^2 = \Phi_{\alpha\alpha}(f)$$

$$\Rightarrow \phi_{\alpha\alpha}(\tau) = \phi_{\beta\beta}(\tau)$$

B)

$$\Phi_{\alpha\beta}(f) = \Phi_{\alpha\alpha}(f) \cdot (-\mathrm{j} \cdot \mathrm{sign}(f))^*$$

$$\Rightarrow \Phi_{\alpha\beta}(-f) = -\Phi_{\alpha\beta}(f) \;\; \text{odd symmetry}$$

$$\Rightarrow \phi_{\alpha\beta}(\tau) = -\phi_{\alpha\beta}(-\tau) \quad \text{odd symmetry}$$

In addition, the following holds for the real-valued random processes $\boldsymbol{\alpha}(t), \boldsymbol{\beta}(t)$

$$\phi_{\alpha\beta}(\tau) = \phi_{\beta\alpha}(-\tau)$$

$$\Rightarrow \phi_{\alpha\beta}(\tau) = -\phi_{\beta\alpha}(\tau)$$

C)

$$\text{From } \mathcal{H}\{\phi_{\alpha\alpha}(\tau)\} \circ\!\!-\!\!\!-\!\!\bullet\; \Phi_{\alpha\alpha}(f) \cdot (-\mathrm{j} \cdot \mathrm{sign}(f)) = \Phi_{\beta\alpha}(f), \text{ we get}$$

$$\mathcal{H}\{\phi_{\alpha\alpha}(\tau)\} = \phi_{\beta\alpha}(\tau)$$

With A–C, the ACF of the transformed random process becomes:

$$\begin{aligned} \phi_{xx}(\tau) &= \underline{\underline{1}} \cdot (\phi_{\alpha\alpha}(\tau) + \mathrm{j}\,\mathcal{H}\{\phi_{\alpha\alpha}(\tau)\}) \cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau} \\ \phi_{xx}(\tau) &= \underline{\underline{1}} \cdot (\phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(\tau) + \mathrm{j}\,\mathcal{H}\left\{\phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(\tau)\right\}) \cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau} \end{aligned}$$

Accordingly, the following holds for the PSD:

$$\Phi_{xx}(f) = \underline{\underline{1}} \cdot \Phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(f + f_0) \cdot (1 + \mathrm{sign}(f + f_0))$$

The transformation rule for signals applies also for the ACF and the PSD of weakly stationary random processes up to a pre-factor of 1 instead of $\frac{1}{\sqrt{2}}$!! The ECB processes of zero-mean processes are also weakly stationary and rotational invariant.

Note: The reason for introducing the pre-factor $\frac{1}{\sqrt{2}}$ in the ECB transformation is to conserve energy and power. This conservation is now obvious by means of the pre-factor 1 for the transformation of the ACF. Thus, both pre-factors are consequences of each other and cannot be chosen indicidually free!

A-9.3.2 Summary

Sample functions, ACF und PSD:

$$
\begin{aligned}
\boldsymbol{x}_{\mathrm{RF}}(\eta,t) \quad &\leftrightarrow \quad \boldsymbol{x}(\eta,t) &=& \ \tfrac{1}{\sqrt{2}}\left(\boldsymbol{x}_{\mathrm{RF}}(\eta,t) + \mathrm{j}\,\mathcal{H}\{\boldsymbol{x}_{\mathrm{RF}}(\eta,t)\}\right)\cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 t} \\
\phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(\tau) \quad &\leftrightarrow \quad \phi_{xx}(\tau) &=& \ \underline{\underline{1}}\cdot\left(\phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(\tau) + \mathrm{j}\mathcal{H}\{\phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(\tau)\}\right)\cdot \mathrm{e}^{-\mathrm{j}2\pi f_0 \tau} \\
\Phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(f) \quad &\leftrightarrow \quad \Phi_{xx}(f) &=& \ \underline{\underline{1}}\cdot \Phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(f+f_0)\cdot(1+\mathrm{sign}(f+f_0))
\end{aligned}
$$

Inverse transform:

$$
\begin{aligned}
\boldsymbol{x}_{\mathrm{RF}}(\eta,t) &= \sqrt{2}\,\mathrm{Re}\{\boldsymbol{x}(\eta,t)\cdot \mathrm{e}^{\mathrm{j}2\pi f_0 t}\} \\
\phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(\tau) &= \underline{\underline{1}}\,\mathrm{Re}\{\phi_{xx}(\tau)\cdot \mathrm{e}^{\mathrm{j}2\pi f_0 \tau}\} \\
\Phi_{x_{\mathrm{RF}}x_{\mathrm{RF}}}(f) &= \underline{\underline{\tfrac{1}{2}}}(\Phi_{xx}(f-f_0) + \Phi_{xx}(-(f+f_0)))
\end{aligned}
$$

Conditions for weakly stationarity

1) $m_x = 0$
2) $\phi_{x_{\mathrm{I}}x_{\mathrm{I}}}(\tau) = \phi_{x_{\mathrm{Q}}x_{\mathrm{Q}}}(\tau)$
3) $\phi_{x_{\mathrm{I}}x_{\mathrm{Q}}}(\tau) = -\phi_{x_{\mathrm{Q}}x_{\mathrm{I}}}(\tau) = -\phi_{x_{\mathrm{I}}x_{\mathrm{Q}}}(-\tau)$

These conditions are summarized in the theorem of Grettenberg:

A weakly stationary complex random process is an ECB process of a real physical weakly stationary random process if and only if it has zero mean, its real and imaginary parts have identical autocorrelation functions, and its crosscorrelation function is odd symmetric.

This implies the necessity of **rotational symmetry** of a stationary complex random process.

The rotation $\boldsymbol{x}(\eta,t)\cdot \mathrm{e}^{\mathrm{j}2\pi f_0 t}$ must not cause the first and second order moments to be time-variant (i.e. dependent on t)!

Condition 3) states that in a stationary, rotational symmetric complex random process, the real and the imaginary parts at the same time instant are always uncorrelated:

$$E\{\boldsymbol{x}_{\mathrm{I}}(\eta,t)\boldsymbol{x}_{\mathrm{Q}}(\eta,t)\} = 0 \quad \forall t$$

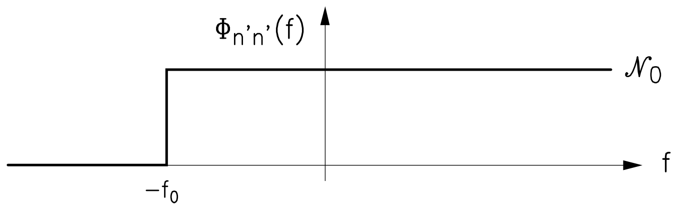
A-9.3.3 ECB Transformation of White Noise

For a real physical white noise $\boldsymbol{n}_0(\eta, t)$ with PSD $\Phi_{n_0 n_0}(f) = N_0/2$, the corresponding ECB noise

$$\boldsymbol{n}'(\eta, t) = \boldsymbol{n}'_{\mathrm{I}}(\eta, t) + \mathrm{j}\, \boldsymbol{n}'_{\mathrm{Q}}(\eta, t) \quad \text{with}$$

$$\mathrm{PSD} \quad \Phi_{n'n'}(f) = \frac{N_0}{2}\,(1 + \mathrm{sign}(f + f_0)) = \begin{cases} N_0 & \text{for } f \geq -f_0 \\ 0 & \text{for } f < -f_0 \end{cases}$$

is not white!

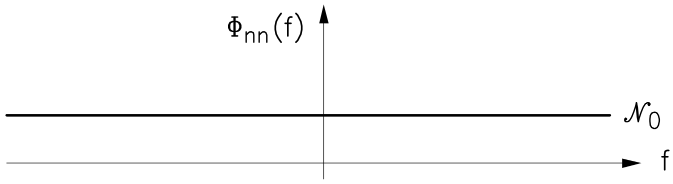


White noise is merely a simple alternative model at the input of band-limited filters $H_{\mathrm{RF}}(f)$, which limits the noise power any ways. However, for the ECB filter $H(f)$, the following applies in any case:

$$H(f) = 0 \text{ for } f < -f_0.$$

(Therefore, for a band-limited system, the ECB equivalent system without the suppression of the transfer function for $f < -f_0$ here is not applicable!)

Hence, a non-zero noise power spectral density for $f < -f_0$ is irrelevant. Without loss of generality, an *alternative complex-valued white process* $\boldsymbol{n}(\eta, t)$ with PSD $\Phi_{nn}(f) = N_0 \quad \forall f$ can be used instead of the noise $\boldsymbol{n}'(\eta,t)$ with PSD $\Phi_{n'n'}(f)$.



We also **define** $\boldsymbol{n}(\eta, t)$ as an ECB process of a real-valued white noise $\boldsymbol{n}_0(\eta, t)$.

ACF $\phi_{nn}(\tau) = N_0\,\delta(\tau)$

Quadrature components: $\boldsymbol{n}(\eta,t) = \boldsymbol{n}_{\mathrm{I}}(\eta,t) + \mathrm{j}\boldsymbol{n}_{\mathrm{Q}}(\eta,t)$

$\phi_{n_{\mathrm{I}}n_{\mathrm{I}}}(\tau) = \phi_{n_{\mathrm{Q}}n_{\mathrm{Q}}}(\tau) = \frac{N_0}{2}\cdot\delta(\tau)$

$\Phi_{n_{\mathrm{I}}n_{\mathrm{I}}}(f) = \Phi_{n_{\mathrm{Q}}n_{\mathrm{Q}}}(f) = \dfrac{N_0}{2}$

$\phi_{n_{\mathrm{I}}n_{\mathrm{Q}}}(\tau) \equiv 0$ (Rotational symmetry!)

Transmission of a complex-valued white noise over an ECB system with impulse response $h(t) = h_{\mathrm{I}}(t) + \mathrm{j}\,h_{\mathrm{Q}}(t)$:

$$\boldsymbol{y}(\eta,t) = h(t) * \boldsymbol{n}(\eta,t)$$

$$\Phi_{yy}(f) = N_0\,|H(f)|^2$$

$$\phi_{y_{\mathrm{I}}y_{\mathrm{I}}}(\tau) = \phi_{y_{\mathrm{Q}}y_{\mathrm{Q}}}(\tau) = \frac{N_0}{2}\left(\varphi_{h_{\mathrm{I}}h_{\mathrm{I}}}(\tau) + \varphi_{h_{\mathrm{Q}}h_{\mathrm{Q}}}(\tau)\right)$$

$$\phi_{y_{\mathrm{I}}y_{\mathrm{Q}}}(\tau) = \frac{N_0}{2}\left(\varphi_{h_{\mathrm{I}}h_{\mathrm{Q}}}(\tau) - \varphi_{h_{\mathrm{I}}h_{\mathrm{Q}}}(-\tau)\right)$$

The output process is stationary, rotationally symmetric and remains a complex-valued Gaussian random process.