Modulation

Digital Modulations

Table of Content

3. Digital Modulation

Most message transmission systems today use digital signals for the transmission. If the input-signal of the modulator is digital (discrete value and time) one refers to it in simplified terms as Digital Modulation.

Analog source signals are converted into digital signals using an analog-digital-converter. These signals can then be further processed with the help of digital circuit technology and software algorithms (digital signal processing DSV) [9]. Programmable Digital Signal Processors (DSP) and Field-Programmable Gate Arrays (FPGA) allow comprehensive software implementation of modulators and demodulators, whose function as needed can be very easily modified without changing the hardware. For very high numbers of individual units, e.g. mobile telephones, all digital signal processing can be realized in an ASIC, whereby it is mostly the later modification of functions that are less comprehensive compared to solutions with DSP or FPGA.

Fig. 3-1 shows the block diagram for a transmission system with digital modulation. On the transmitter side, signals from analog sources are converted to digital signals by an analog-digital-converter and fed into the source coding. Signals from digital sources are directly fed to the source coding.
The source coding increases the channel capacity through a data compression in which all redundant and irrelevant data are removed from the source signal. Many compression procedures are standardized, e.g. MP3, MPEG, LPC, Linear Predictive Coding in GSM.
In the channel coding, the redundancy is increased again through the addition of more bits which leads to an increase of the bit rate. Several targets can be reached with channel coding:

- Long sequences of the same symbols (interfering DC-component) are avoided
- Missing symbol change is avoided in order to simplify timing recovery on the receiver's side (e.g. Manchester Line code)
- Error recognition and error correction on the receiver side

In digital modulation, the amplitude, frequency or phase of a sinusoidal carrier is influenced by the digital modulation signal. The modulated signal can only have discrete values.

In the transmission channel, interferences and noise are added.

And on the receiver side, the data clock and the carrier have to be recovered and synchronized for the demodulation. The demodulated signal gets into the channel-decoder so transmission errors can be detected and corrected. The signal is uncompressed in the source-decoder and fed into the digital sink or the digital-analog-converter.

Digital modulations have numerous advantages compared to analog transmission:

- Digital information processing
- Simple multiplex operation
- “Data encryption“, privacy communication
- High interference immunity
- Error correction is easily possible
- Low bandwidth requirements
- Smaller nonlinearities, practically constant S/N
- “simpler“ circuit technology (digital building blocks)

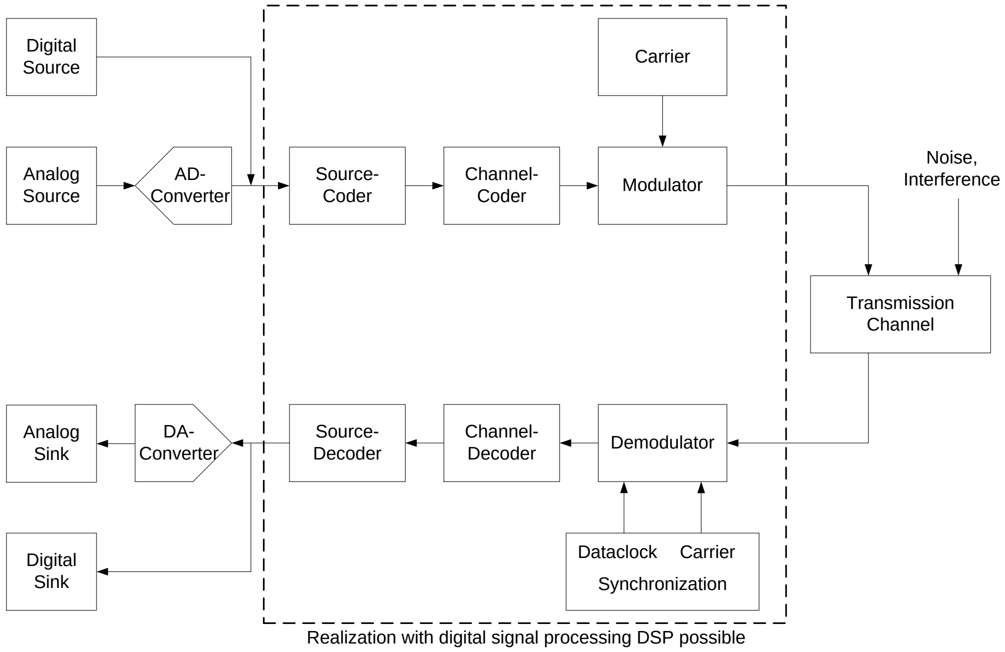


Fig. 3-1: Transmission system with digital modulation

Digital modulation of a sinusoidal carrier:

Sinusoidal carriers can basically be modulated using the same three options as with analog modulation signals:

- Amplitude Shift Keying (ASK)
- Frequency Shift Keying (FSK)
- Phase Shift Keying (PSK)

Different variants are possible for all three basic types.

3.1 Digital signals in baseband (line code) and their properties

Digital data is typically in the form of binary numbers (e.g. “10110001110”). For the transmission they must be implemented in chronological sequence of logical states or **coded**. A rectangular clock signal controls this coding and the transmission.

Below the basic line codes:

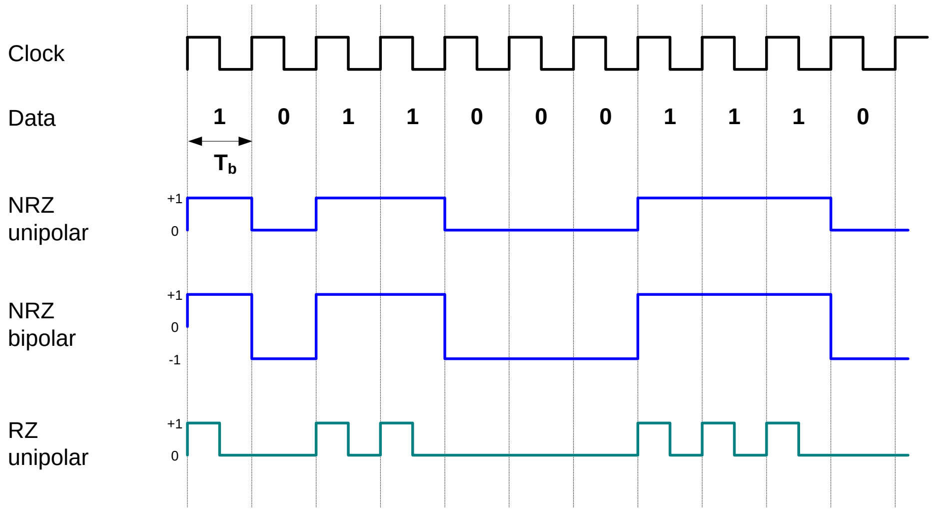


Fig. 3-2: Simple line code

Biphase Codes have the advantage compared to the already considered NRZ- and RZ-codes that they are free of DC and make it possible to facilitate easy clock recovery. They have at least one slope in each data bit.

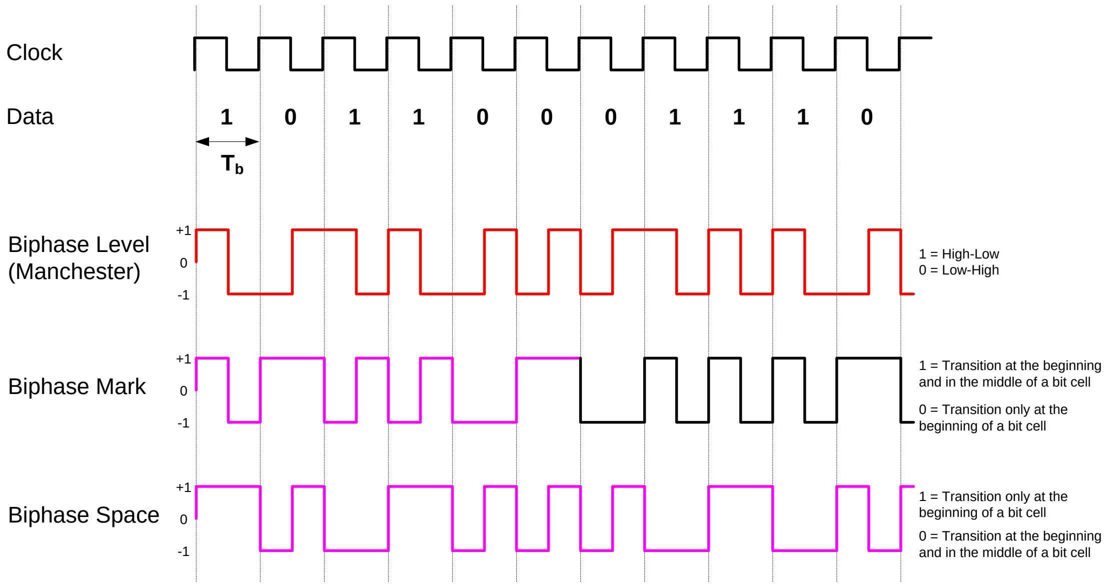


Fig. 3-3: Biphase line code

The baseband signals can be depicted by their Fourier series:

$$b_{unip}(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{i=1}^{\infty}\frac{1}{i}\sin\frac{i \cdot 2\pi}{T_s}t \qquad b_{bip}(t) = \frac{2}{\pi}\sum_{i=1}^{\infty}\frac{1}{i}\sin\frac{i \cdot 2\pi}{T_s}t \tag{3.1}$$

$i = 1,3,5,.....$

When these baseband signals are multiplied by the carrier $u_c(t) = \hat{u}_c\cos(\omega_c t)$, one gets:

$$u_{modunip}(t) = u_c(t) \cdot b_{unip}(t) = \frac{\cos(\omega_c t)}{2} + \frac{2}{\pi}\sum_{i=1}^{\infty}\left[\sin\left(\frac{i \cdot 2\pi}{T_s} + \omega_c\right)t + \sin\left(\frac{i \cdot 2\pi}{T_s} - \omega_c\right)t\right] \tag{3.2}$$

$$u_{modbip}(t) = u_c(t) \cdot b_{bip}(t) = \frac{2}{\pi}\sum_{i=1}^{\infty}\frac{1}{2i}\left[\sin\left(\frac{i \cdot 2\pi}{T_s} + \omega_c\right)t + \sin\left(\frac{i \cdot 2\pi}{T_s} - \omega_c\right)t\right] \tag{3.3}$$

With a **unipolar** baseband signal the result in the carrier with half amplitude and the upper and lower sideband.

With a **bipolar** baseband signal the carrier is not present, but only both sidebands.

Spectral viewing of the signal is a very important aspect in digital modulation techniques.

The multiplication of the baseband signals by a sinusoidal carrier corresponds to convolution of the signals in the frequency domain.

$$s(t) = s_1(t) \cdot s_2(t) \quad \circ\!\!-\!\!\bullet \quad S(f) = S_1(f) * S_2(f) \tag{3.4}$$

If $s_1(t)$ is a baseband signal and $s_2(t)$ is the carrier, the spectrum of the baseband signal will be mapped in the spectral domain on both sides of the carrier.

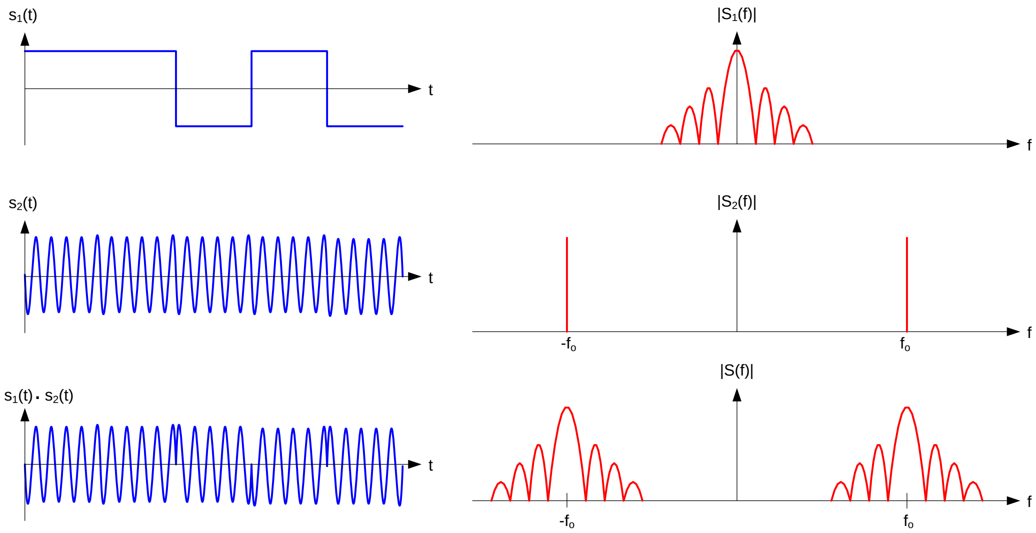


Fig. 3-4: Time and spectral domain, top: bipolar baseband signal, center: sinusoidal carrier, bottom: baseband signal multiplied by carrier signal

In practice, the baseband signals are used with a statistical distribution of the “zeros“ and “ones“, a pseudo-random bit sequence PRBS (see 3.1.4). This results in a line spectrum, whose envelop has a |si(x)|-function.

$$si(x) = sinc(x) = \frac{\sin(x)}{x} \tag{3.5}$$

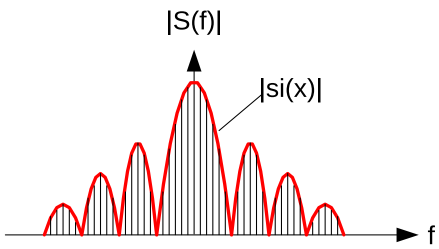


Fig. 3-5: Spectrum of a pseudo random bit sequence (PRBS)

In practical measurements with the spectrum analyzer, only the envelope is displayed. Additionally, not voltages but power in the form of **power spectral density G(f)** are displayed. This representation is consistently used in the digital modulation techniques.

$$G(f) = |S(f)|^2 \tag{3.6}$$

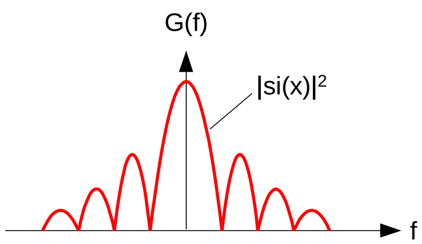


Fig. 3-6: Power spectral density of a pseudo random bit sequence (PRBS)

3.1.1 Important properties of line codes

Clock recovery

The clock content of a code should be as independent as possible from the content of the data being transmitted in order to facilitate clock recovery on the receiver side.

DC component

A DC-value cannot be transmitted without difficulty on transmission systems. For that reason DC-free code is to be striven for. Mostly this can only be fulfilled by statistical means.

Power Spectral Density G(f)

The power spectral density can be calculated using the Fourier transformation of the auto-correlation function.
Important characteristics are the amplitude distribution and the location of the zeros.

Error probability P_e

The error probability in a transmission channel disturbed by Gaussian noise (AWGN, Additive White Gaussian Noise) is represented as a function of the signal-to-noise ratio. The terms used here mean:

$$\frac{E_b}{N_0} = \frac{S}{N} B_{Noise} T_b = \frac{S}{N} \frac{B_{Noise}}{f_b} \tag{3.7}$$

E_b/N_0 = Value for S/N

E_b = Energy per Bit = $U^2 T$

$N_0/2$ = Power spectral density for AWGN

AWGN = Additive White Gaussian Noise

$$\mathrm{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^U e^{-U^2} du \tag{3.8}$$

$$\mathrm{erfc}(u) = 1 - \mathrm{erf}(u) = \frac{2}{\sqrt{\pi}} \int_U^\infty e^{-U^2} du \tag{3.9}$$

$$\mathrm{erf}(0) = 0 \qquad \mathrm{erf}(\infty) = 1 \tag{3.10}$$

erf = Error Function

erfc = Complementary Error Function

Nyquist bandwidth

The question of the minimum necessary bandwidth can be answered with help of the sampling theorem.

It is:

$$B_N = \frac{1}{2T_{min}} = \frac{\text{Baudrate}}{2} \qquad T_{min} = \text{shortest puls duration} \tag{3.11}$$

3.1.1.1 Unipolar NRZ (Non Return to Zero)

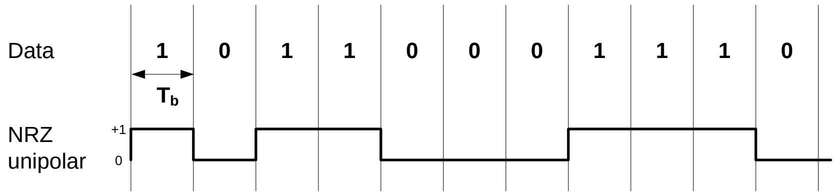


Fig. 3-7: Line code unipolar NRZ

Characteristics:

+ simple code
- DC-components
- Clock recovery in long sequences of "1" or "0" not possible

Power spectral density:

$$G_{U-NRZ}(f) = \frac{U^2 T_b}{4} si^2(\pi f T_b) = \frac{U^2 T_b}{4} \left[\frac{\sin(\pi f T_b)}{\pi f T_b} \right]^2 \qquad (3.12)$$

U = Voltage of the logical 1 - state

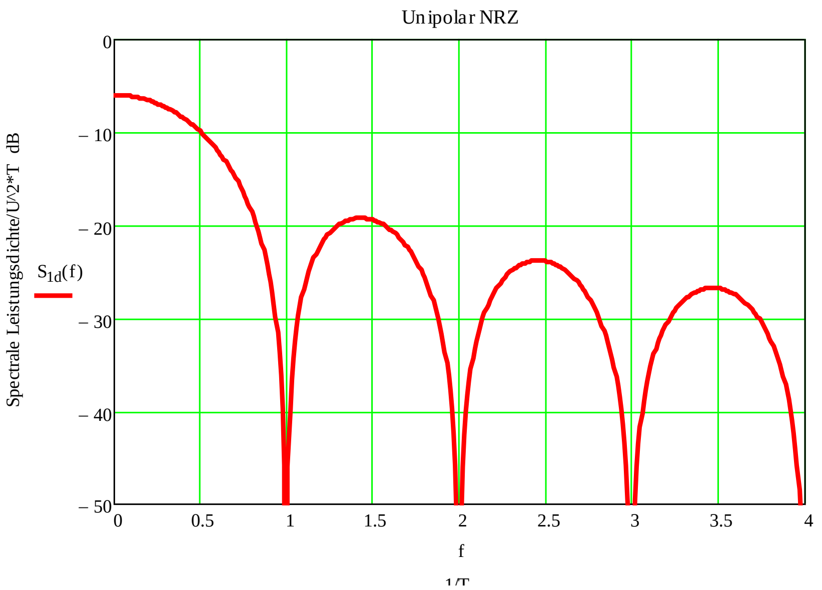


Fig. 3-8: Power spectral density of line code unipolar NRZ

Error probability:

$$P_{e_{U-NRZ}} = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{2N_o}} \right) \qquad (3.13)$$

Nyquist bandwidth:

$$B_N = \frac{1}{2T_b} \qquad T_b = \text{user data bit duration} \qquad (3.14)$$

3.1.1.2 Bipolar NRZ (Non Return to Zero)

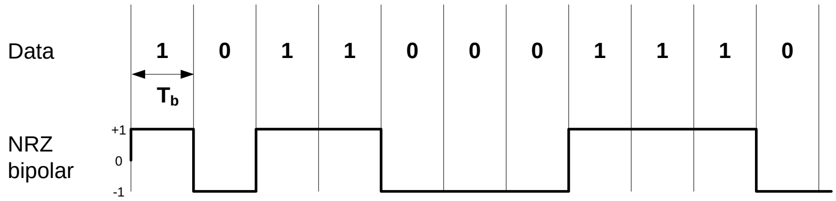


Fig. 3-9: Line code bipolar NRZ

Characteristics:

+ Simple code
+ No DC-components, to the extent that the distribution of “1“ and “0“ are equal
- Clock recovery for long sequences “1“ or “0“ not possible

Power spectral density:

$$G_{B-NRZ}(f) = U^2 T_b \ si^2(\pi f T_b) = U^2 T_b \left[\frac{\sin(\pi f T_b)}{\pi f T_b}\right]^2 \tag{3.15}$$

U = Voltage of the logical 1 - state

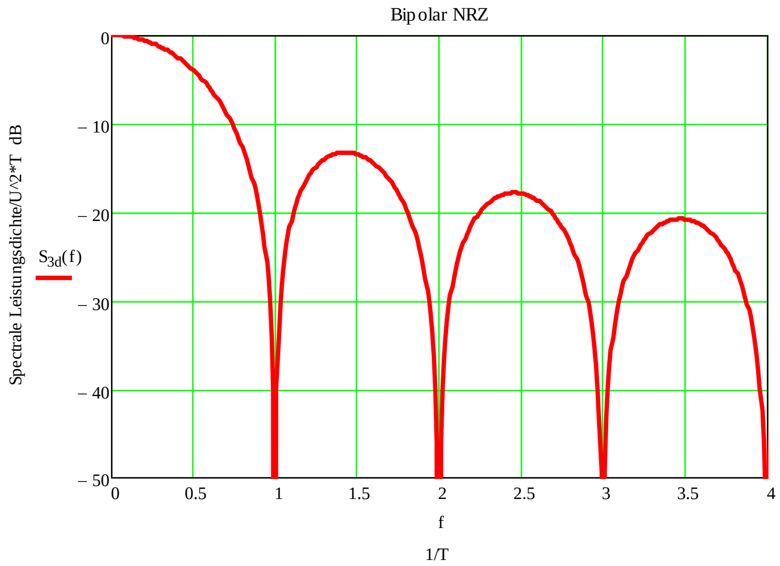


Fig. 3-10: Power spectral density of line code bipolar NRZ

Error probability:

$$P_{e_{B-NRZ}} = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_b}{N_o}}\right) \tag{3.16}$$

Nyquist bandwidth:

$$B_N = \frac{1}{2T_b} \qquad T_b = \text{user data bit duration} \tag{3.17}$$

3.1.1.3 Unipolar RZ (Return to Zero)

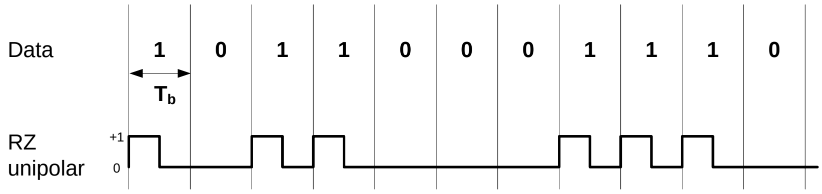


Fig. 3-11: Line code unipolar RZ

Characteristics:

+ Clock recovery more possible than with NRZ
- DC-components
- Clock recovery in long sequences of “0“ not possible
- Double bandwidth compared to NRZ

Power spectral density:

$$G_{U-RZ}(f) = \frac{U^2 T_b}{16} si^2(\pi f T_b / 2) = \frac{U^2 T_b}{16}\left[\frac{\sin(\pi f T_b / 2)}{\pi f T_b / 2}\right]^2 \tag{3.18}$$

U = Voltage of the logical 1 - state

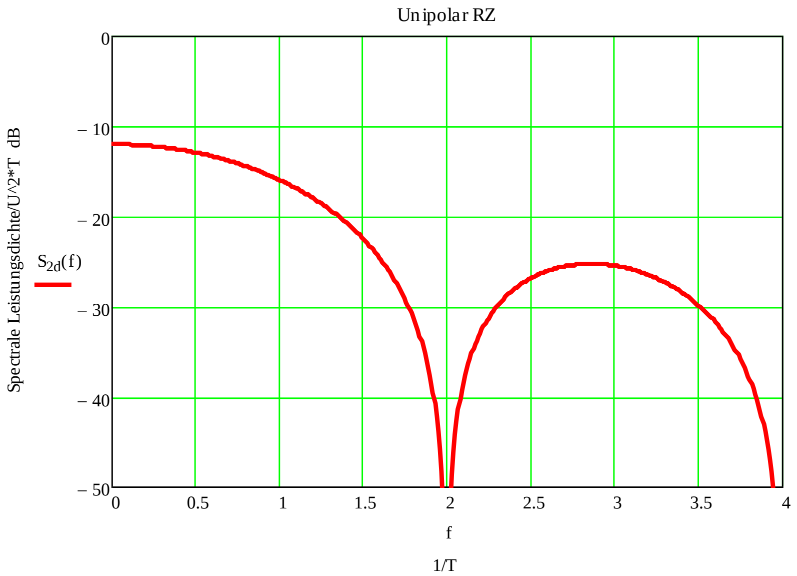


Fig. 3-12: Power spectral density of line code unipolar RZ

Error probability:

$$P_{e_{U-RZ}} = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{E_b}{8N_0}}\right) \tag{3.19}$$

Nyquist bandwidth:

$$B_N = \frac{1}{T_b} \qquad T_b = \text{user data bit duration} \tag{3.20}$$

3.1.1.4 Biphase Level (Manchester)

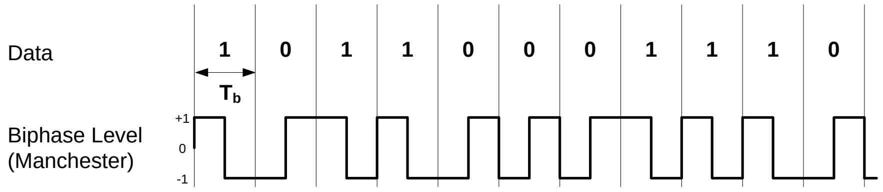


Fig. 3-13: Line code Biphase Level (Manchester)

Characteristics:

+ Simple clock recovery
+ No DC-components
- Double bandwidth compared to NRZ

Power spectral density:

$$G_{BL}(f) = U^2 T \left[\frac{\sin(\pi f T_b / 2)}{\pi f T_b / 2} \right]^2 \sin^2(\pi f T_b / 2) \tag{3.21}$$

U = Voltage of the logical 1 - state

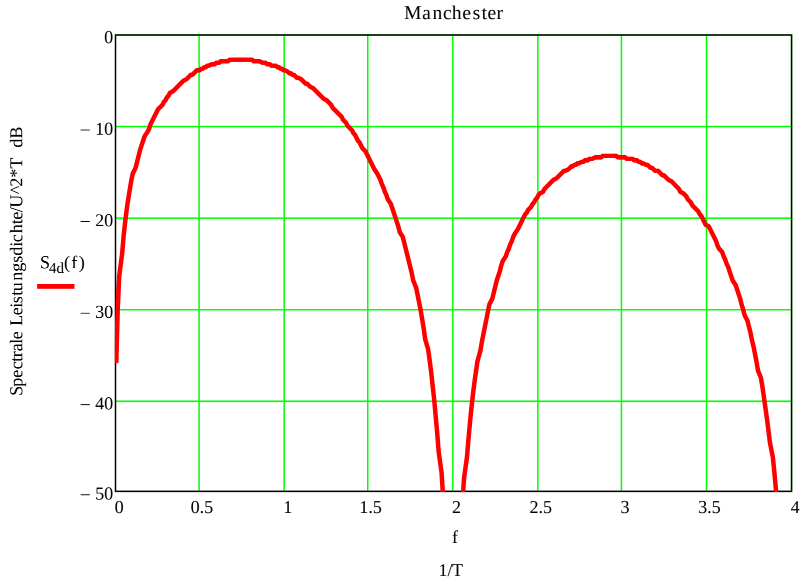


Fig. 3-14: Power spectral density of line code Biphase Level (Manchester)

Error probability:

$$P_{e_{U-NRZ}} = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_o}} \right) \quad \text{same as Bipolar NRZ} \tag{3.22}$$

Nyquist bandwidth:

$$B_N = \frac{1}{T_b} \qquad T_b = \text{user data bit duration} \tag{3.23}$$

3.1.1.5 Comparison of different codes

Code	*Nyquist-Bandwidth*	*Advantages*	*Disadvantages*
NRZ unipolar	$B_N = \frac{1}{2T_b}$	• Simple Code	• Clock recovery not possible in long “0“- or “1“- sequences • DC-component
NRZ bipolar	$B_N = \frac{1}{2T_b}$	• Simple Code • No DC-component as long as distribution of „1“ and „0“ are equal	• Clock recovery not possible for long “0“- or “1“- sequences • DC-component
RZ unipolar	$B_N = \frac{1}{T_b}$	• Clock recovery more possible than with NRZ	• Clock recovery not possible for long “0“- sequences • DC-component • Double bandwidth compared to NRZ
Biphase	$B_N = \frac{1}{T_b}$	• Simple clock recovery • No DC-component	• Double bandwidth compared to NRZ

Fig. 3-15: Comparison of different line codes

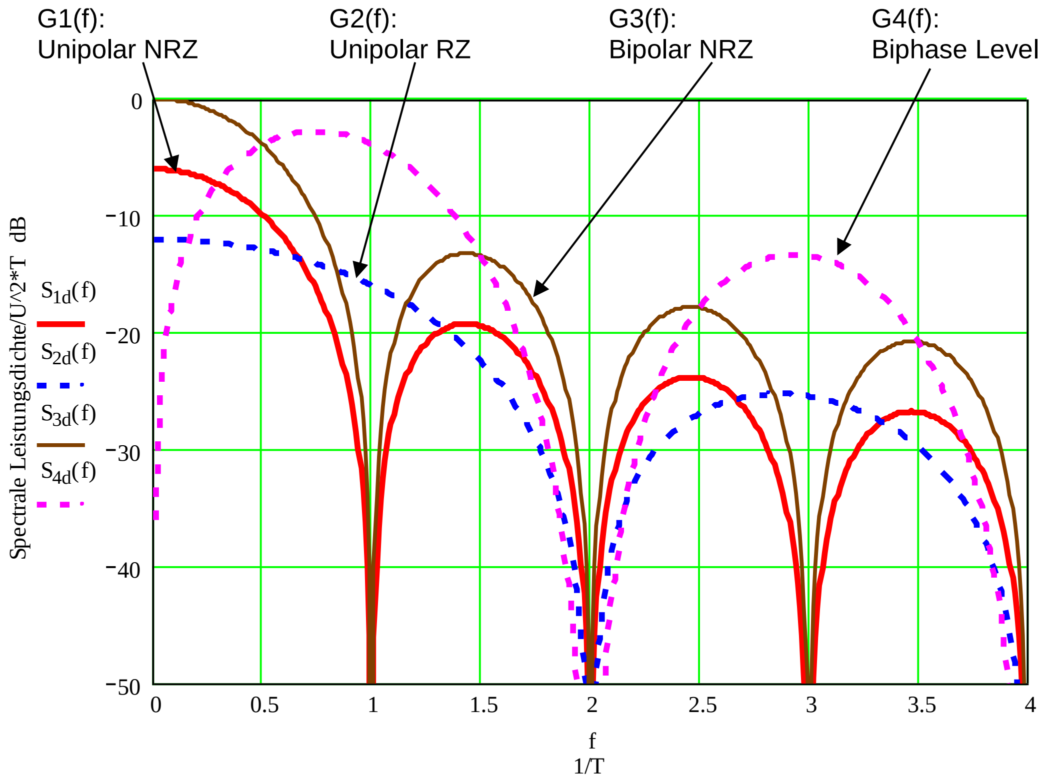


Fig. 3-16: Comparison of power spectral density of different line codes

3.1.2 Bit Error Rate (BER)

The BER is defined as

$$BER = \frac{n_e}{n_{tot}} \qquad n_e = \text{Number of error bits} \qquad n_{tot} = \text{Total number of transmitted bits} \tag{3.24}$$

Example:
If one error bit is detected in a transmission of 10'000 bits, this yields in BER = 10^{-4}

The length of a PRBS-sequence determines the minimum BER which can be measured:

$$BER_{min} = \frac{1}{n} \qquad n = \text{Length of PRBS in bits}$$

Example:

At a PRBS of $n = 2^{15} - 1$ the minimum BER that can be measured is:

$$BER_{min} = \frac{1}{2^{15} - 1} = 3.05 \cdot 10^{-5}$$

Error probability of various codes:

The stated bit error probability applies to interferences with "**Additive White Gaussian Noise**" (AWGN).

The other boundary conditions are ideal.

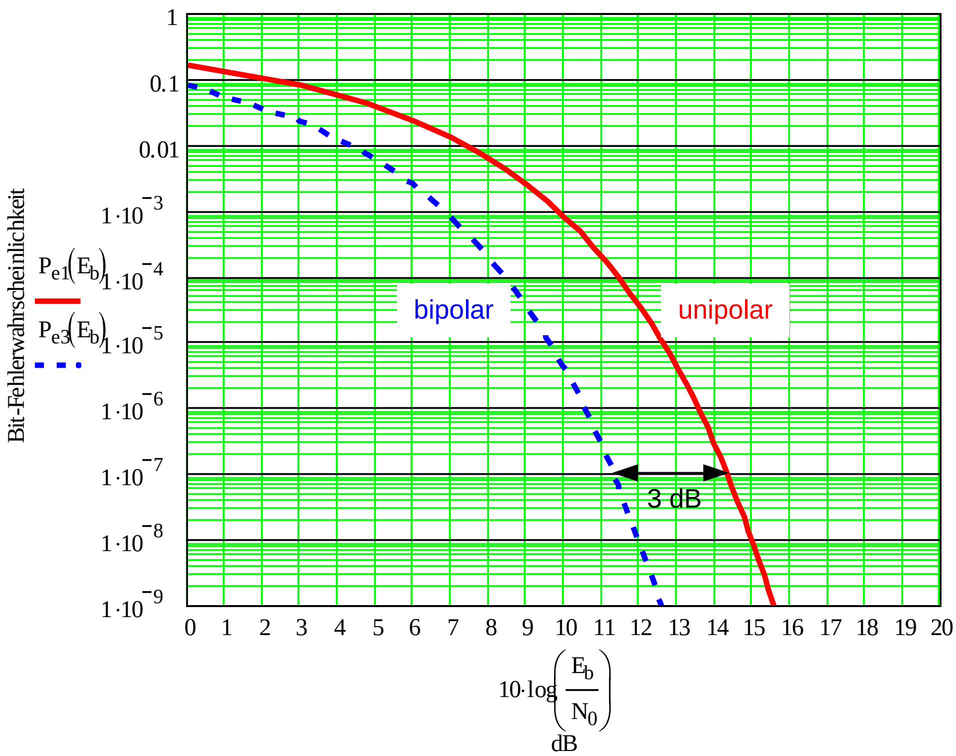


Fig. 3-17: Bit error probability as a function of signal-to-noise ratio

3.1.3 Inter Symbol Interference (ISI)

In the transmission of pulses through a band-limited system, the pulses are delayed and the transient responses and transient overshoots of a symbol can fall in the time slot for the next symbol.

This leads to inter-symbol interference (ISI) and makes it more difficult to identify the symbol on the receiver's side.

Correction: larger bandwidth $\rightarrow$ many disadvantages

Objective: Smallest possible bandwidth with low ISI

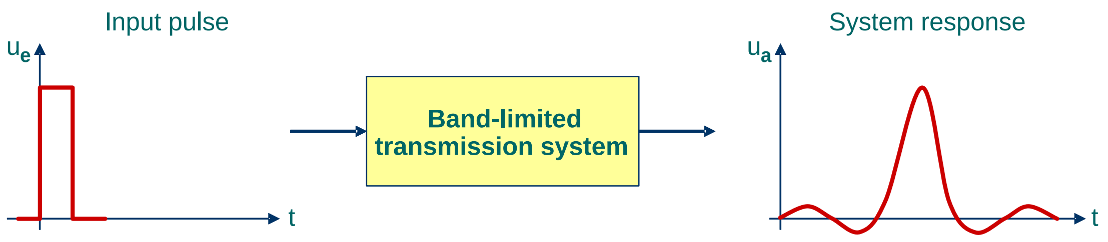


Fig. 3-18: System response of a band-limited transmission system

In a digital baseband transmission, pulses are transmitted (= **symbols)** continuously. If the transient responses and transient decay of the one symbol affect the neighboring symbols in a disadvantageous way, this makes it more difficult to identify the symbols on the receiver side.

Nyquist Criteria for ISI-extinction

Nyquist determined:

In an ISI-free channel (Transmitter-Receiver) all symbol responses must be zero at the sampling time with the exception of the momentary symbol.

$$h_{eff}(n \cdot T_s) = \begin{cases} K & n = 0 \\ 0 & n \neq 0 \end{cases} \tag{3.25}$$

h_{eff} = Pulse response
n = 0, 1, ..., n
T_s = Symbol period
K = Constant ($\neq 0$)

Transmission **with** intersymbol-interferences:

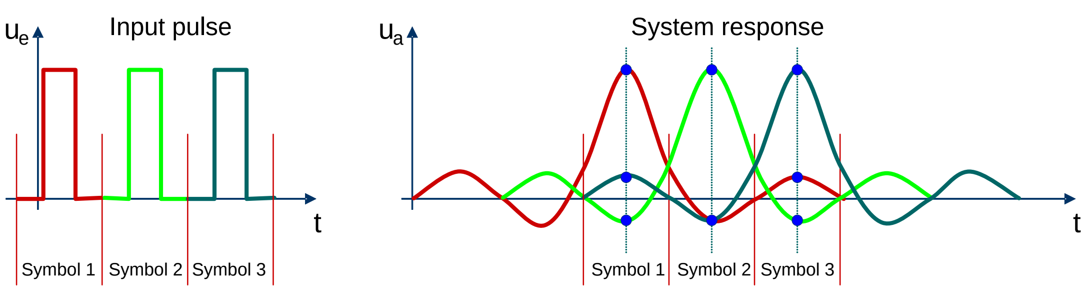


Fig. 3-19: Pulstransmission **with** ISI

Transmission **without** intersymbol-interferences:

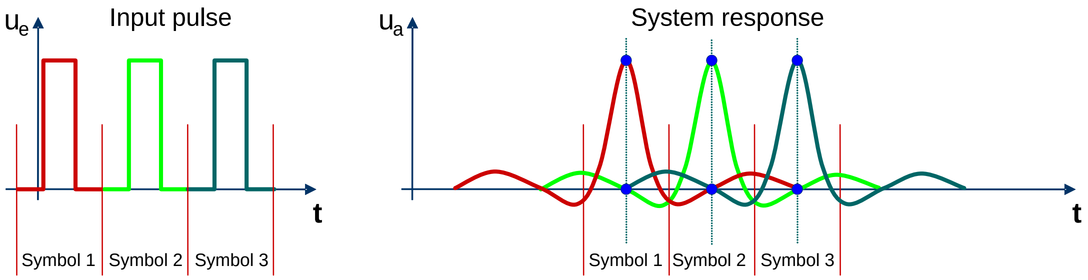


Fig. 3-20: Pulstransmission **without** ISI

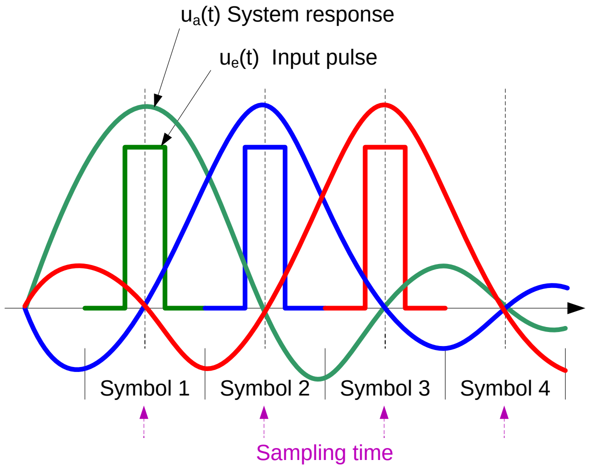


Fig. 3-21: Nyquist criteria for pulse transmission

ISI-free impulse response

Example: $$h_{eff}(t) = \frac{\sin\left(\frac{\pi t}{T_s}\right)}{\frac{\pi t}{T_s}} \tag{3.26}$$

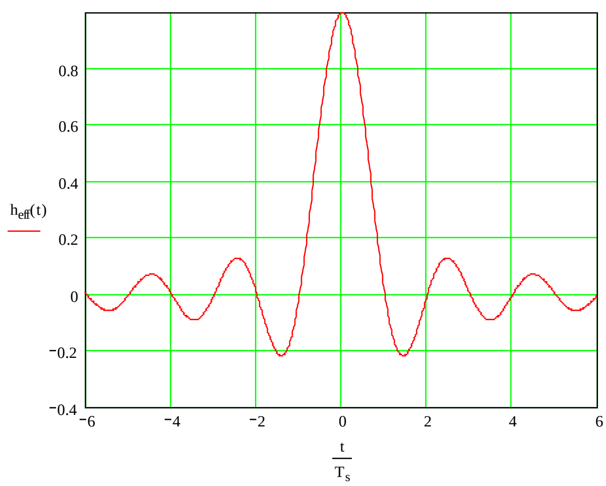


Fig. 3-22: ISI-free pulse response

This impulse response conforms to the Nyquist-Criterion.

The transmission function H(f) which an ISI-free channel must have can be gained from the Fourier transformation of h(t).

$$H_{eff}(f) = \frac{1}{f_s} \mathrm{rect} \frac{f}{f_s} \tag{3.27}$$

f_s = Symbol frequency

This transmission function has a “brick“-characteristic:

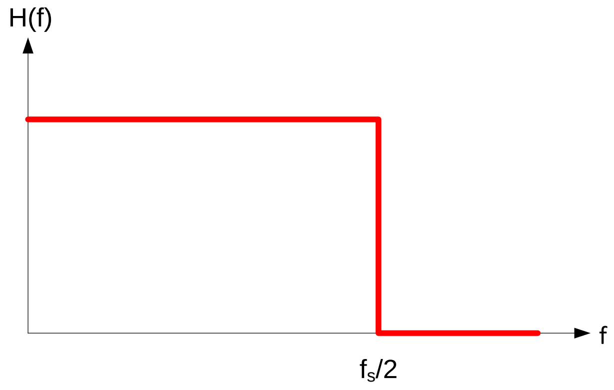


Fig. 3-23: Transfer function for an ideal ISI-free channel

Real low pass filter which evince their impulse response zero points at a distance of n·Ts, can, according to the so-called 1st Nyquist condition be realized with filters with point-symmetrical slopes (Nyquist slopes). The symmetry point lies at the Nyquist frequency $f_s/2$.

The transitional range is determined by the roll-off-factor α.

$$\alpha = \frac{\Delta f}{B_N} \tag{3.28}$$

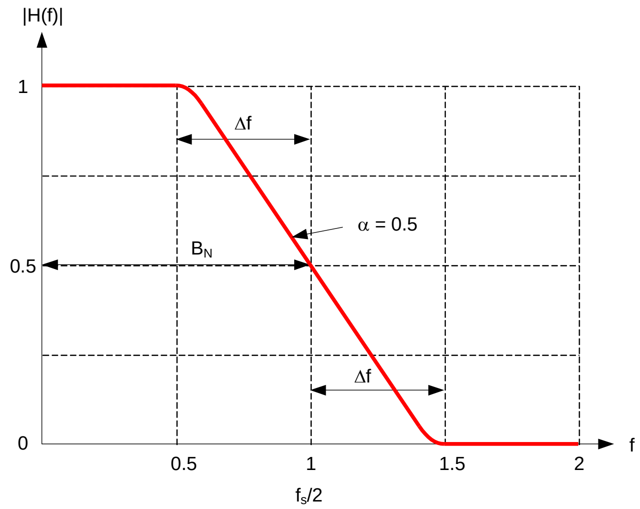


Fig. 3-24: Transfer function of a Nyquist-Filter

To keep the ISI as low as possible in a real system, h(t) must

- Fall off rapidly
- Evince low or no amplitude in the proximity of $n \cdot T_s (n \neq 0)$.

Filters which meet the Nyquist-Criteria will be referred to as Nyquist-filters.

An effective end-to-end transmission function $H_{eff}(f)$ can be easily realized in that the transmitter and receiver each have a transmission function of $\sqrt{H_{eff}(f)}$.

Filtering

To limit the bandwidth in a transmission system, the spectrum must be filtered in the transmission path. Basically there are two options available:

Baseband filtering (impulse formation):

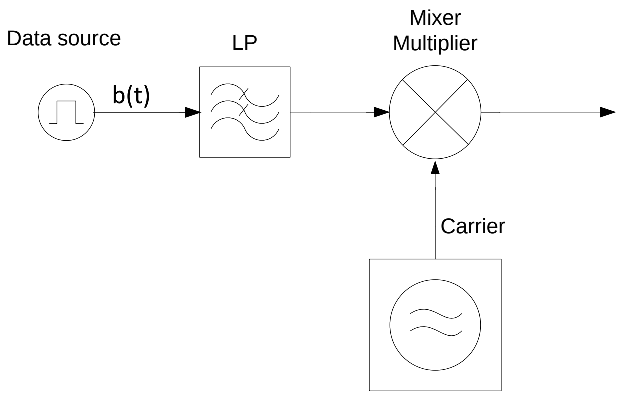


Fig. 3-25: Baseband filter

Band-pass filtering after modulation:

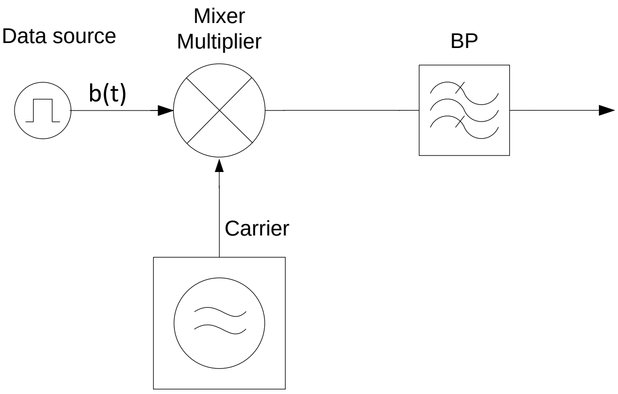


Fig. 3-26: Band-pass filter after modulation

Filtering after modulation requires very high filtering effort due to the relatively small bandwidth and therefore can only be realized with fixed intermediate frequencies at which high-grade quartz- or SAW-filters can be used. In practice both options are used in tandem, mostly.

Typical ISI-free or low-ISI filter types used here are:

- Gauss-filter
- Raised-cosine-filter

Both filters are described in [11].

Eye diagram

The quality of a digital transmission can easily be assessed with the **eye diagram**. In it, the interference from noise as well as jitter (phase fluctuations) along with the intersymbol-interferences are easily recognizable.

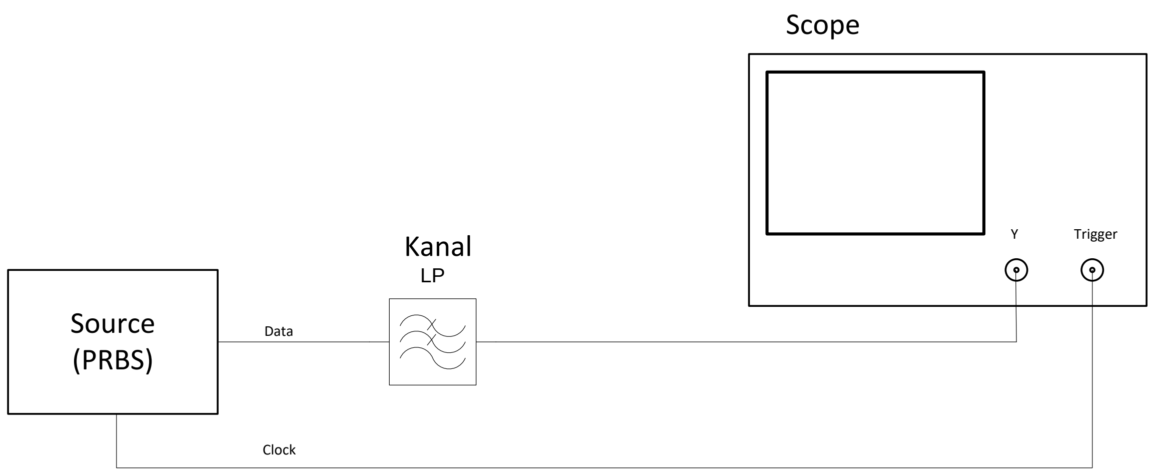


Fig. 3-27: Measurement setup for measuring the eye diagram

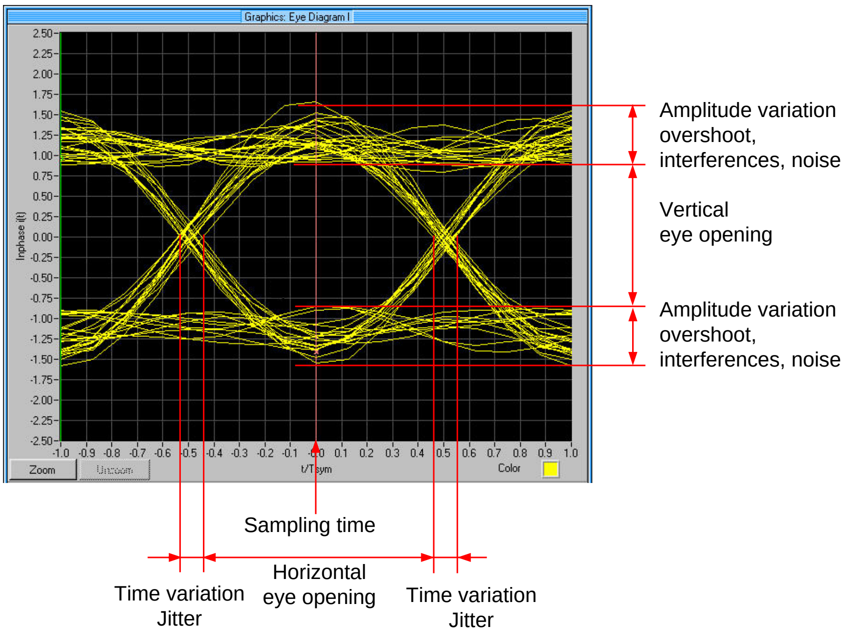


Fig. 3-28: Eye diagram of a real transmission channel

If the transmission quality is good, the eye diagram should have the biggest possible eye opening in the middle of the symbol (= sampling time).

The eye diagram shown in Fig. 3-28 has strong intersymbol-interferences and noise.

3.1.4 Pseudo Random Bit Sequence (PRBS)

A PRBS is used to investigate digital transmission systems in order to get a homogenous spectrum. In an analog system the PRBS corresponds to a pink noise at which all frequency components from 0 to a certain frequency are present.

In a PRBS, all bit sequences between ...00000... and ...010101... must be generated whereby the probability of a "1" is exactly the same as the probability of a "0".

PRBS are realized with the help of a feedback shift register. The pseudo-random sequence must have a length of n bits. Then the bit sequence repeats.

$$n = 2^m - 1 \qquad m = \text{number of shift registers} \qquad (3.29)$$

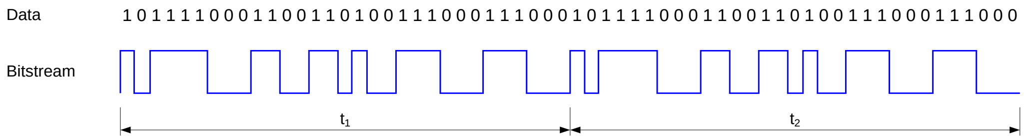


Fig. 3-29: Short Pseudo Random Bit Sequence (PRBS)

The circuit presented below generates a PRBS of 7 bits length. The bit sequence is ...0100111... and contains all combinations that are possible with 3 bits.

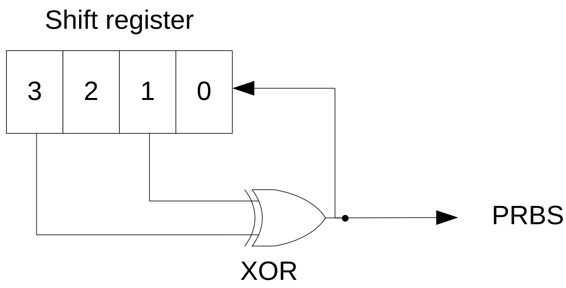


Fig. 3-30: Generation of a PRBS using a shift register

The spectral analysis shows individual discrete spectral lines at a distance of Δf from

$$\Delta f = \frac{1}{n} f_b \qquad \begin{aligned} n &= \text{Length of PRBS in bits} \\ f_b &= \text{Bit frequency} = \frac{1}{T_b} \\ T_b &= \text{Bit duration in s} \end{aligned} \qquad (3.30)$$

This shows that for n to infinity the spectrum merges into a continuous spectrum.

In practical applications, PRBS with a length of 511 bits and more are used.

PRBS-generator in PSPICE with unipolar and bipolar output:

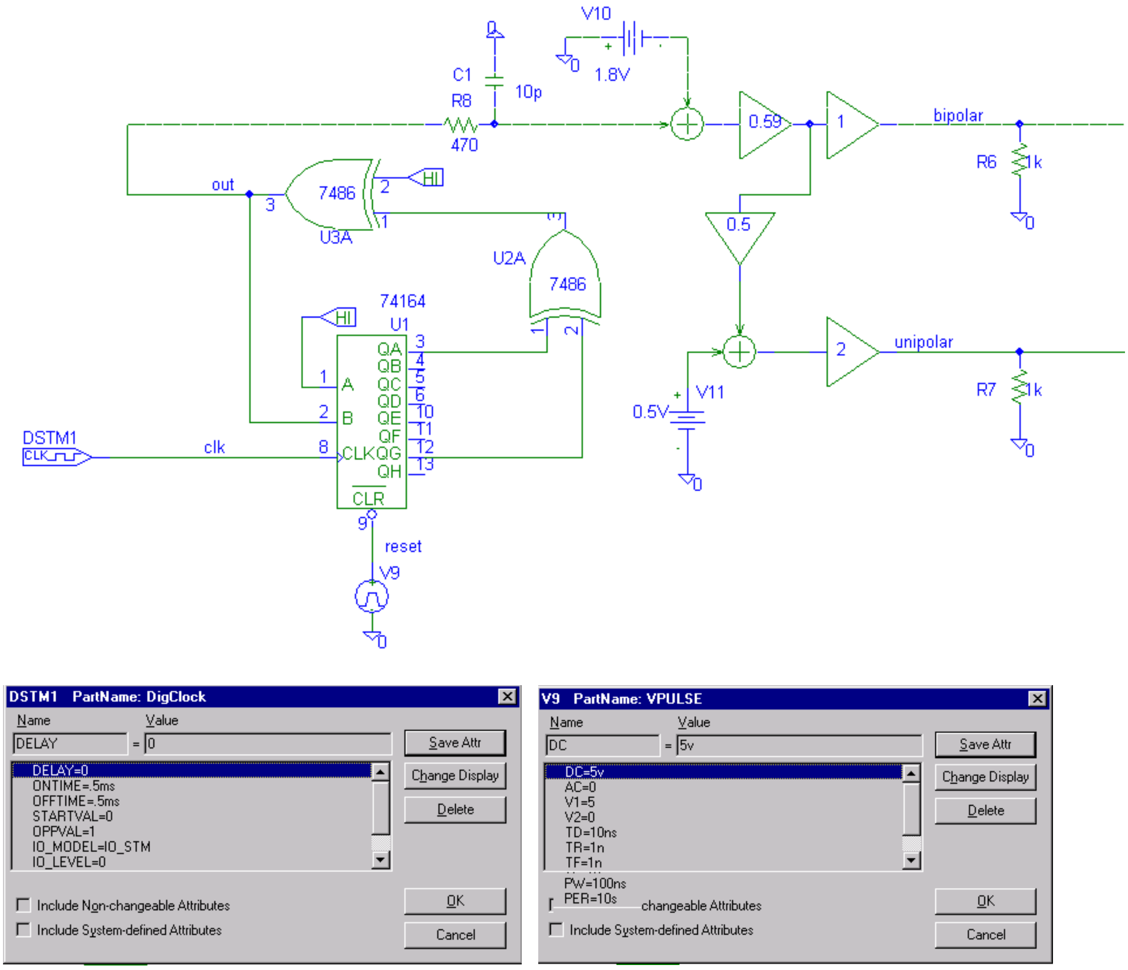


Fig. 3-31: PRBS simulation in SPICE

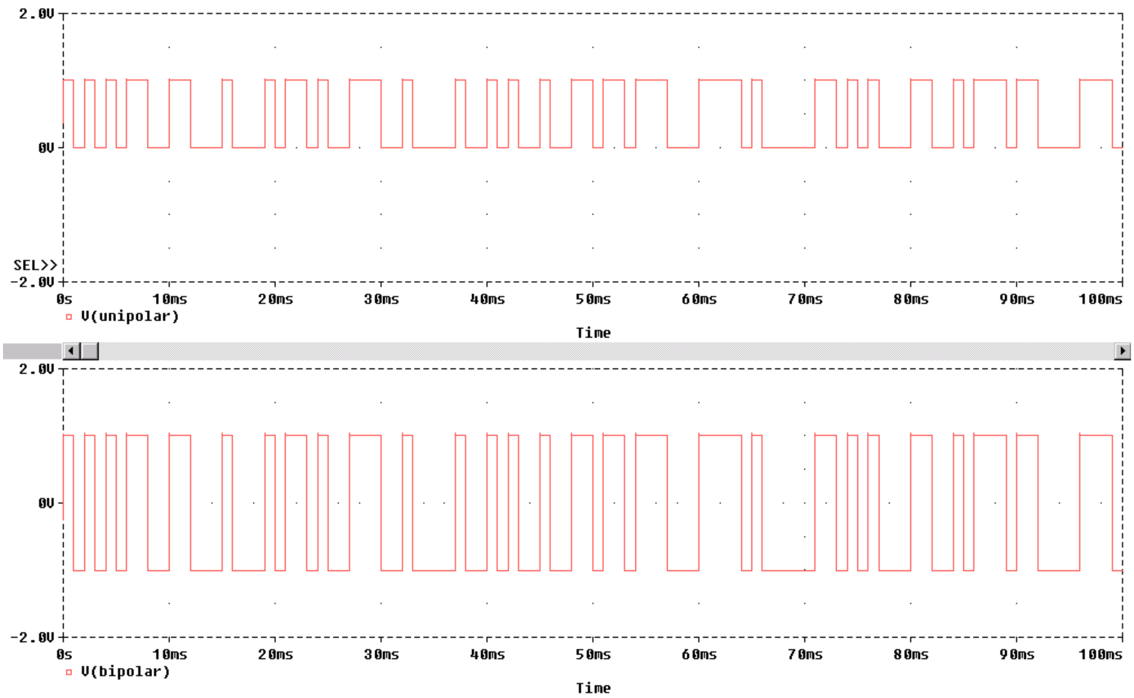
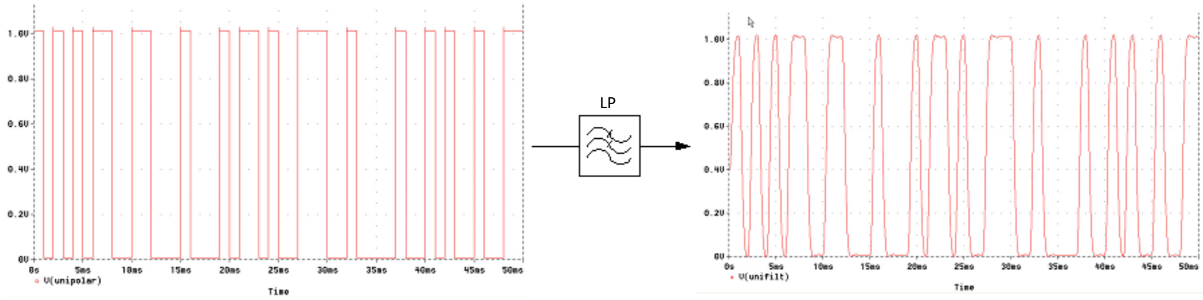


Fig. 3-32: Output signal unipolar and bipolar



Unfiltered baseband signal   Filtered baseband signal

Fig. 3-33: Baseband signal in the time domain, unfiltered (left) and filtered (right)

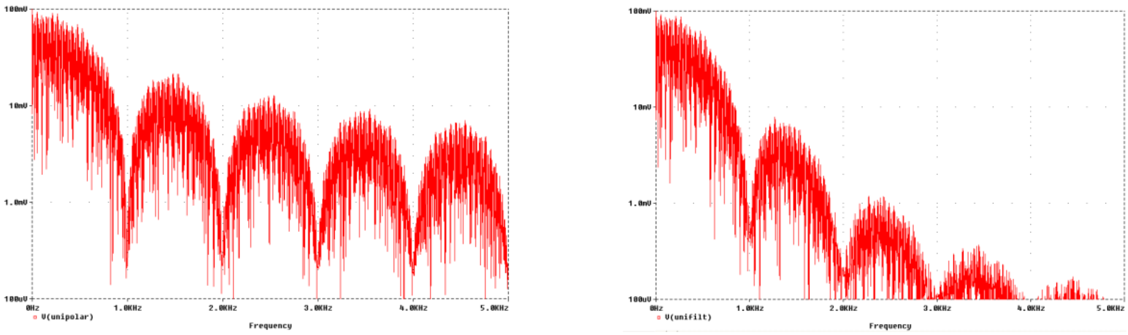


Fig. 3-34: Baseband signal in the frequency domain, unfiltered (left) and filtered (right)

3.2 Modulation possibilities of a sinusoidal carrier

We describe a sinusoidal high-frequency signal with

$$u_c(t) = A \cdot \cos(\omega_c t + \theta) \tag{3.31}$$

$A = \hat{u}_c$ = Amplitude (Peak value)

ω_c = Angular frequency of the carrier

θ = Phase shift (constant)

t = Time

Here we have three options for modulation: A, ω_c, θ

Representation of the sinusoidal signal in different domains:

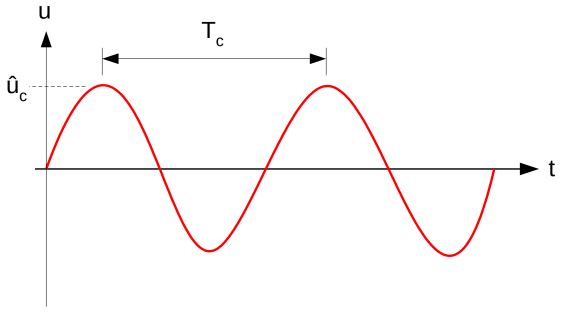


Fig. 3-35: Sinusoidal signal in the time domain

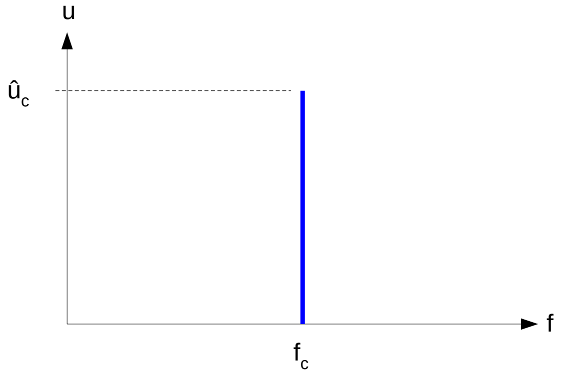


Fig. 3-36: Sinusoidal signal in the frequency domain

In the **phase state diagram**, this signal is displayed as a pointer (phasor).

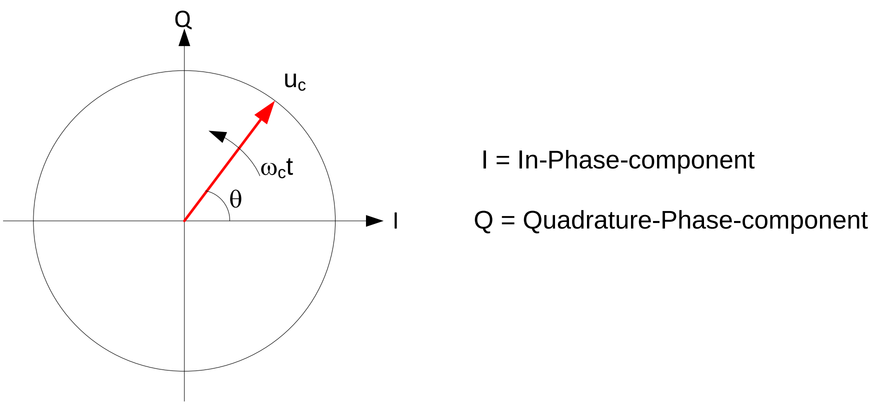


Fig. 3-37: Sinusoidal signal in the phase domain

3.2.1 Amplitude modulation with analog modulation signal

The amplitude A is influenced by the modulation content.

$$A(t) = f(u_m(t)) = \hat{u}_c + u_m(t) = \hat{u}_c + \hat{u}_m \cos(\omega_m t) \tag{3.32}$$

$$u_{AM}(t) = A(t)\cos(\omega_c t + \theta) \tag{3.33}$$

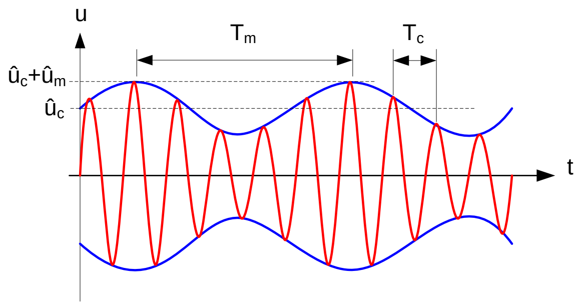


Fig. 3-38: Amplitude modulated signal in the time domain

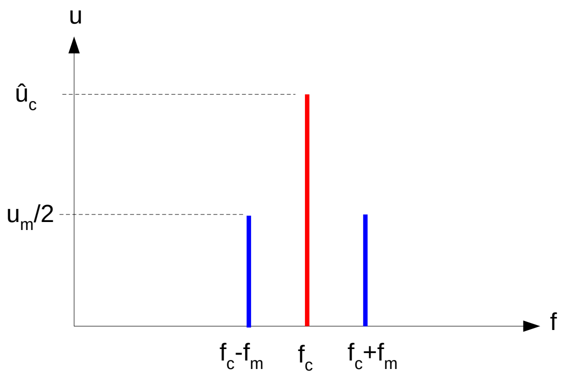


Fig. 3-39: Amplitude modulated signal in the frequency domain

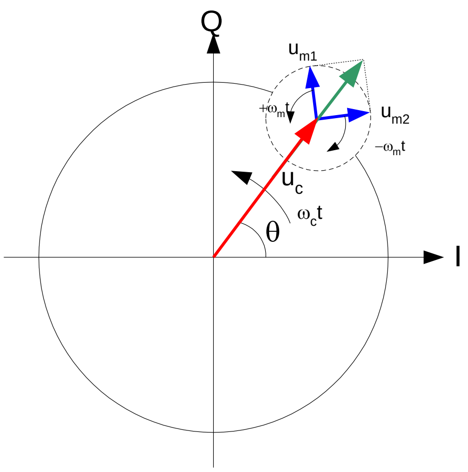


Fig. 3-40: Amplitude modulated signal in the phase domain

3.2.2 Angle modulation (Frequency and phase mod.) with analog modulation signal

Frequency modulation:

The carrier frequency ω_c is influenced by the modulation content:

$$\omega_c(t) = f(u_m(t)) \tag{3.34}$$

Phase modulation:

The carrier phase θ_c is influenced by the modulation content:

$$\theta_c(t) = f(u_m(t)) \tag{3.35}$$

The instantaneous phase angle of the carrier is

$$\phi_c(t) = \omega_c(t) \cdot t + \theta_c(t) \tag{3.36}$$

The instantaneous frequency of the carrier is described by

$$\omega_c(t) = \frac{d\phi_c(t)}{dt} \tag{3.37}$$

With that we can describe the angular modulation as follows

$$u_c(t) = A \cdot \cos(\omega_c t + \eta \cdot \sin(\omega_m t)) \tag{3.38}$$

$\eta =$ Peak value of $\theta(t) =$ Modulation index

From this we get the instantaneous frequency

$$f(t) = \frac{\omega_c}{2\pi} + \frac{\eta\omega_m}{2\pi}\cos(\omega_m t) \tag{3.39}$$

and with $\Delta f = \eta f_m \; \eta = \dfrac{\Delta f}{f_m}$ (3.40)

$$u_c(t) = A \cdot \cos\left(\omega_c t + \frac{\Delta f}{f_m}\sin(\omega_m t)\right) \tag{3.41}$$

For spectral analysis, the Fourier coefficients have to be determined.

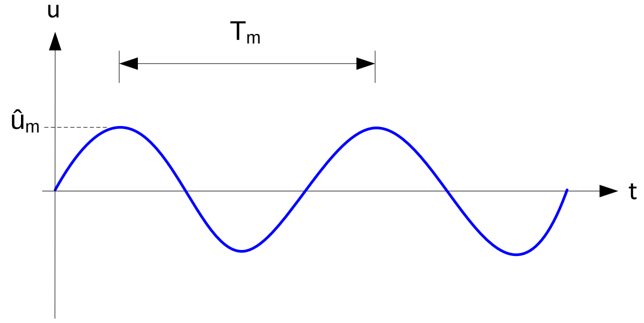


Fig. 3-41: Analog modulation signal in the time domain

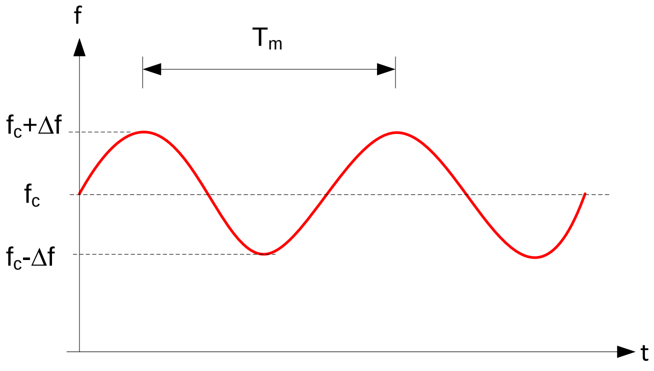


Fig. 3-42: Angular modulated signal: instantaneous frequency in the time domain

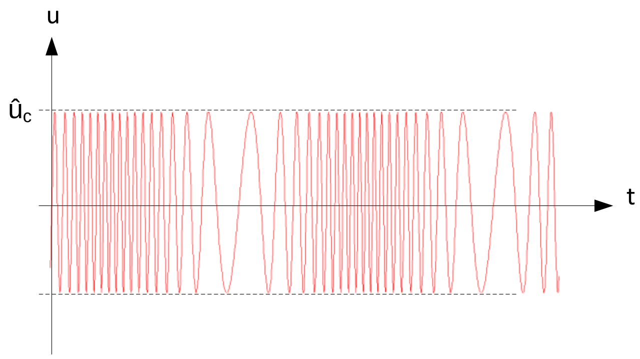


Fig. 3-43: Angular modulated signal in the time domain

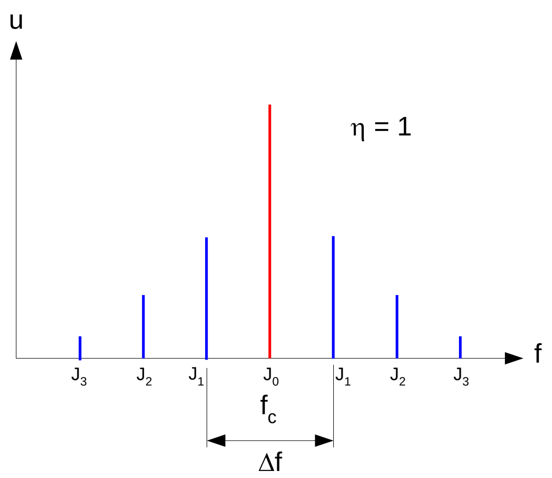


Fig. 3-44: Angular modulated signal in the frequency domain

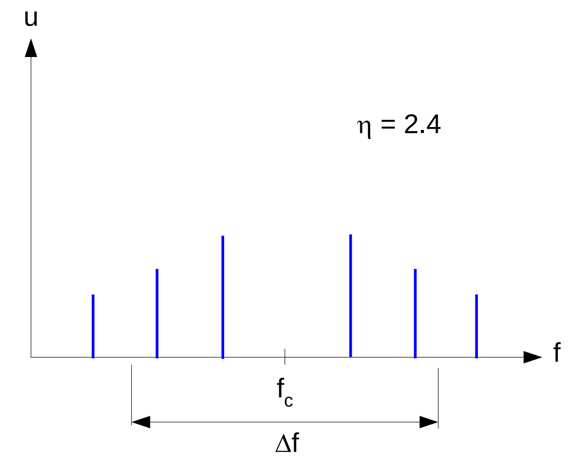


Fig. 3-45: Angular modulated signal in the frequency domain

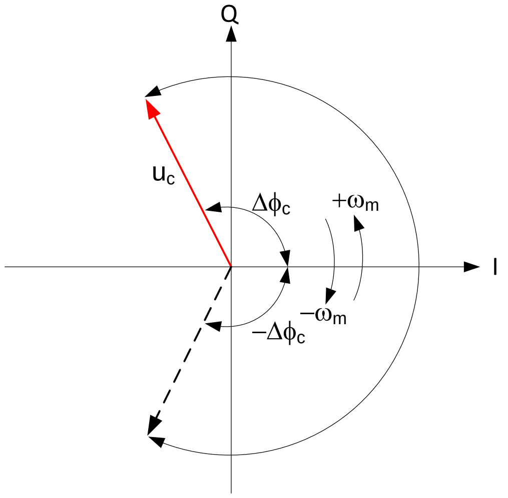


Fig. 3-46: Angular modulated signal in the phase domain

3.3 Digital modulation of a sinusoidal carrier

Sinusoidal carriers can be modulated with digital modulation signals basically with the same three options as with analog modulation signals:

- Amplitude Shift Keying (ASK)
- Frequency Shift Keying (FSK)
- Phase Shift Keying (PSK)

Different variants are possible for all three basic types.

3.3.1 Amplitude Shift Keying ASK

With the binary modulation signal b(t), the amplitude of the carriers must be keyed between two discrete amplitude values. In so-called “On-Off-Keying“(OOK) the carrier is turned on and off.

Amplitude Shift Keying is only used in very simple systems, e.g. Keyless-Entry-Systems, and is of little significance for the transmission of digital baseband signals in more complex systems.

3.3.1.1 On-Off-Keying OOK

We can describe the binary signal consisting of a serial bit stream of “0” and “1“as follows:

$$b(t) = \left| \begin{matrix} +1 & \text{"1"} \\ 0 & \text{"0"} \end{matrix} \right. \tag{3.42}$$

The amplitude of b(t) is normalized to the maximum voltage (e.g. +5V).

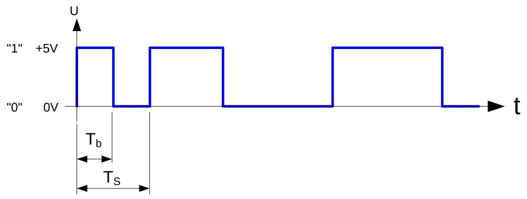


Fig. 3-47: Modulation signal for amplitude shift keying

The bitrate is stated in bit/s and equals $r_b = \frac{1\,\text{bit}}{T_b}$

The bitrate is in numerical value identical to the bit sequence frequency or the bit clock frequency

$$f_b = \frac{1}{T_b} \tag{3.43}$$

The Nyquist bandwidth (minimum necessary bandwidth to transmit a 0101 bit sequence) is

$$B_N = \frac{1}{2T_b} = \frac{1}{2} f_b \tag{3.44}$$

The carrier is described with

$$s_C(t) = \hat{u}_C \cos(\omega_c t) \tag{3.45}$$

The modulated signal results from multiplication of carrier and modulation signal

$$u_{ASK}(t) = b(t)s_C(t) = b(t)\hat{u}_C \cos(\omega_c t) \tag{3.46}$$

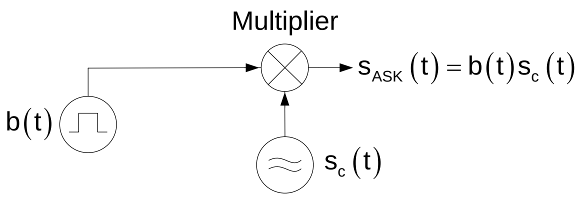


Fig. 3-48: ASK-Modulation with multiplier

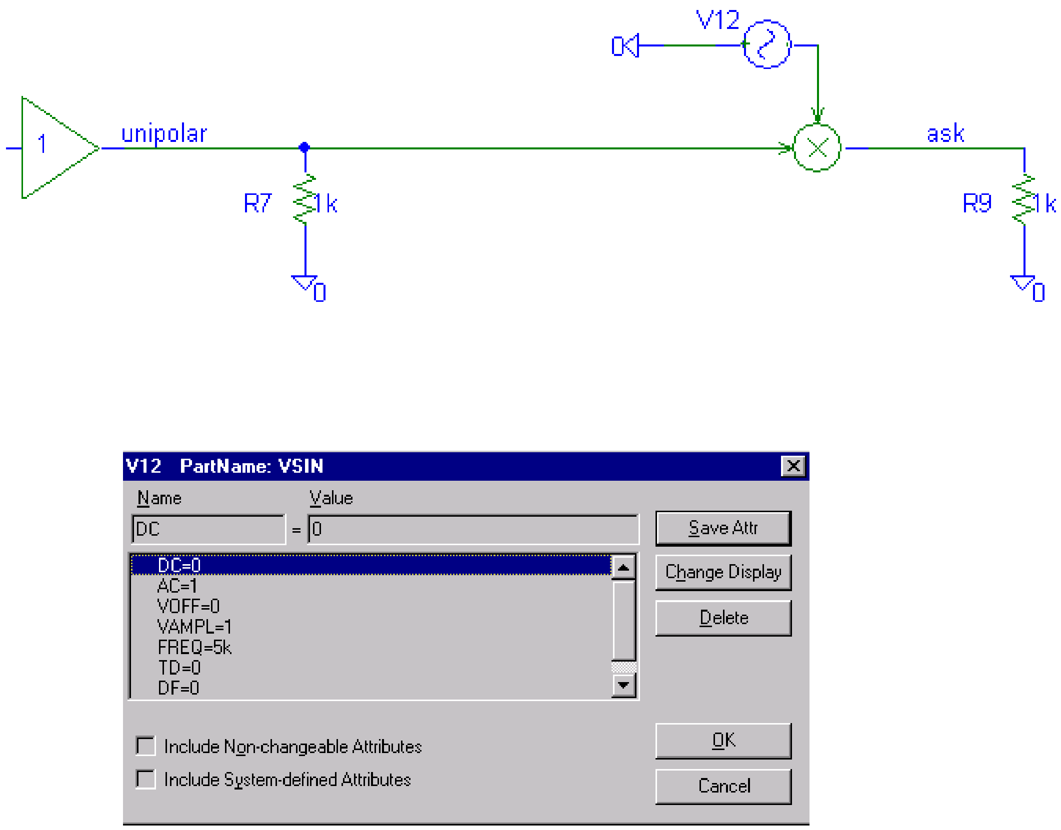


Fig. 3-49: ASK-Modulation with multiplier (PSPICE-Schematic)

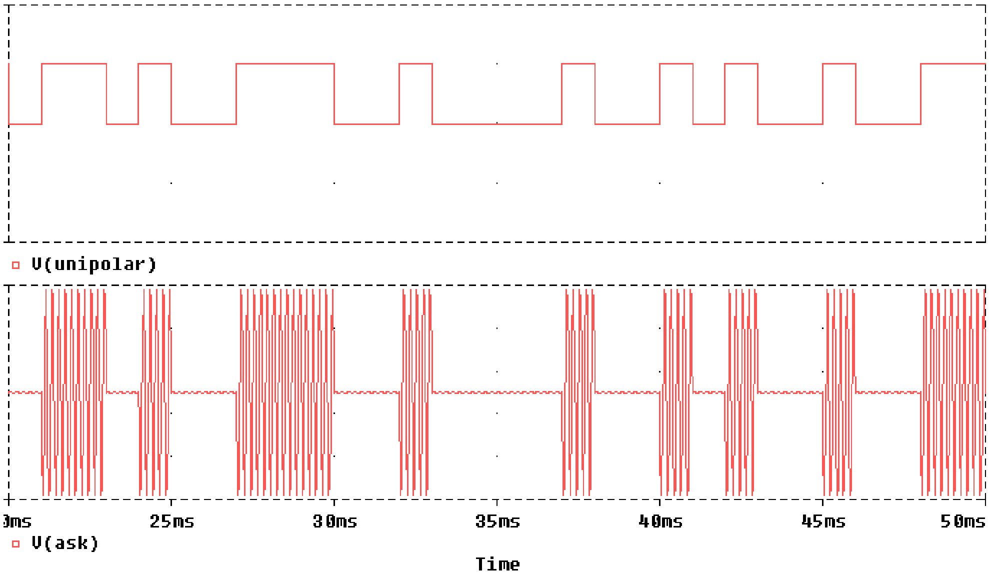


Fig. 3-50: ASK- (OOK) modulated signal in the time domain $(T_b = 1\,\text{ms})$

One gets the spectrum of the modulated signal by convolution (multiplication of the Fourier-series with binary signal in time domain).

$$s(t) = \hat{u}_c \cos(\omega_c t)\left\{\frac{1}{2} + \frac{2}{\pi}\left[\cos\left(\frac{2\pi t}{T_S}\right) - \frac{1}{3}\cos\left(\frac{3\cdot 2\pi t}{T_S}\right) + ...\right]\right\} \tag{3.47}$$

$$s(t)\cdot b(t) \quad \circ\!\!-\!\!\bullet \quad S(f) * B(f)$$

For simplicity's sake the modulation signal here was assumed as a 1-0 sequence with a period duration of

$T_S = 2\,T_b$

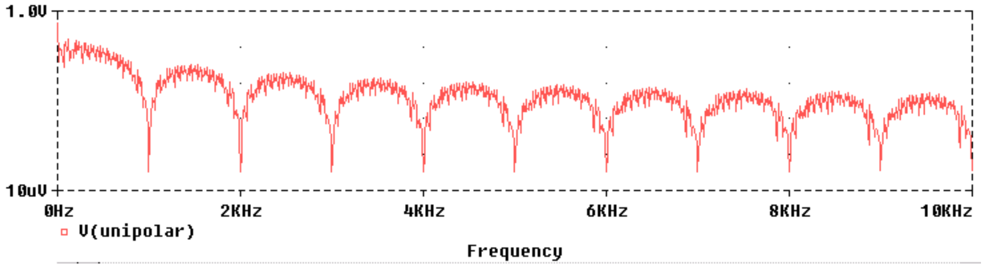


Fig. 3-51: Spectrum of a binary modulation signal (PRBS $T_b = 1\,\text{ms}$)

The spectrum of the modulated signal is symmetrical to the carrier:

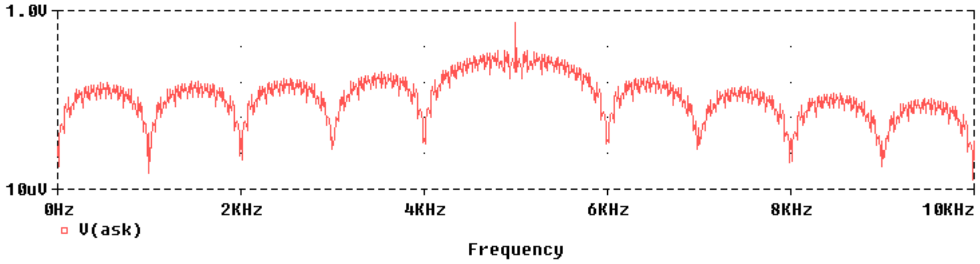


Fig. 3-52: Spectrum of a 5kHz carrier modulated with a PRBS ($T_b = 1\,\text{ms}$)

3.3.1.2 Multi-level amplitude shift keying (mASK)

Amplitude Key Shifting can also be done with multiple levels with m-signal states (m-ary ASK, mASK) ($m = 2,4,8,......,2^n$, $n = 1,2,3,......$). In this the carrier will be modulated through an m-step baseband signal with the symbol rate $r_s = \frac{1}{T_s}$.

Bits are combined into m symbols. In the modulation interval $kT_s \le t \le (k+1)T_s$ each symbol of the baseband will be assigned a discrete amplitude step of the carrier.

Example:

With n=2 we get $m = 2^n = 4$ symbols.

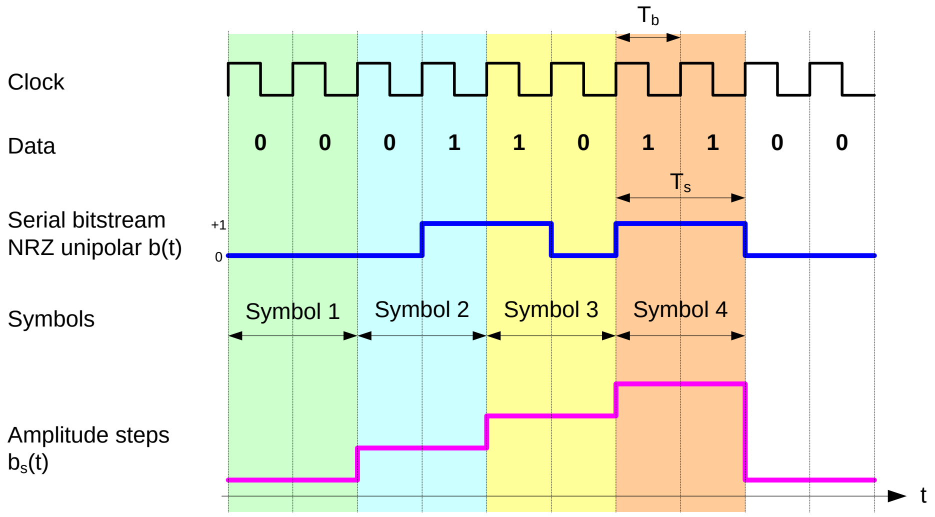


Fig. 3-53: Symbol generation

Bit	Symbol
00	Symbol 1
01	Symbol 2
10	Symbol 3
11	Symbol 4

Fig. 3-54: Symbols

In the German literature, the term “Dibit“ is also used in place of a symbol if a symbol combines two bits.

The serial bit-stream is converted into the desired amplitude steps through serial-parallel conversion and D/A-conversion. The assigning of amplitude steps to the symbols is arbitrary and matches the desired system specifications and is not standardized. The following assignment is just one example:

$$b_s(t) = \left| \begin{array}{ll} 1.00\text{V} & \text{"11"} \\ 0.66\text{V} & \text{"10"} \\ 0.33\text{V} & \text{"01"} \\ 0\text{V} & \text{"00"} \end{array} \right. \qquad (3.48)$$

The symbol rate r_s is n-times smaller than the bitrate and thus the Nyquist bandwidth is also n-times smaller than the binary ASK (OOK).

$$r_s = \frac{1}{T_s} = \frac{r_b}{n} = \frac{1\,\text{Bit}}{nT_b} = \frac{r_b}{\log_2 m} \tag{3.49}$$

$$B_N = \frac{1}{2T_s} = \frac{1}{2nT_b} = \frac{1}{2}\frac{f_b}{\log_2 m} \tag{3.50}$$

The required RF-bandwidth is twice as large as the bandwidth of the baseband signal due to the formation of two sidebands:

$$B_{mASK} = \frac{1}{T_s} \tag{3.51}$$

When using a Raised Cosine Filter, the necessary RF-bandwidth is:

$$B_{mASK} = \frac{1}{T_s}(1+\alpha) \qquad \alpha = \text{Roll-off-Faktor des Filters} \tag{3.52}$$

The **modulation** results from multiplying $u_c(t)$ by $b_s(t)$

$$s_{mASK}(t) = u_c(t)\cdot b_s(t) \tag{3.53}$$

The envelope curve of the mASK-modulated signal is determined by the impulse form of the baseband signal. When filtered with a Nyquist filter the signal transfers are "soft", which is why one refers to "soft keying", in contrast to "hard keying" with rectangular impulses.

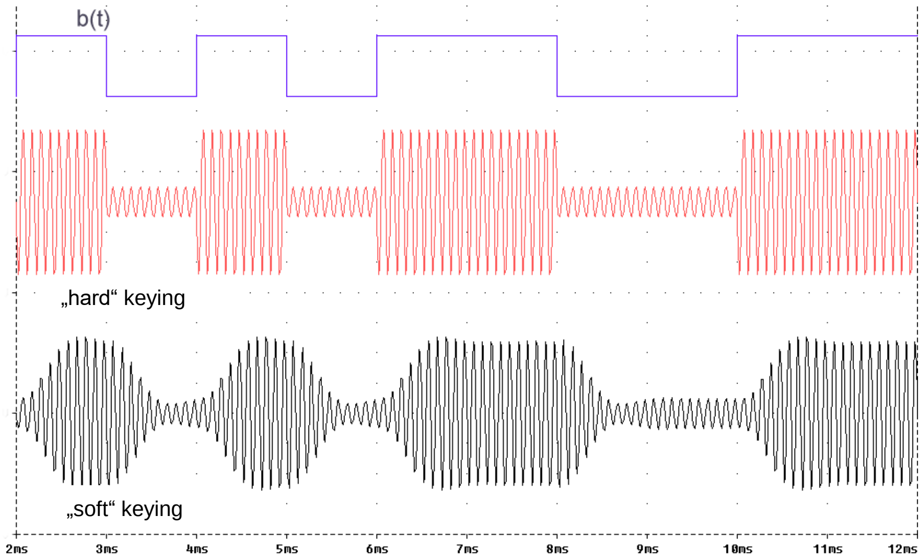


Fig. 3-55: Baseband filtering

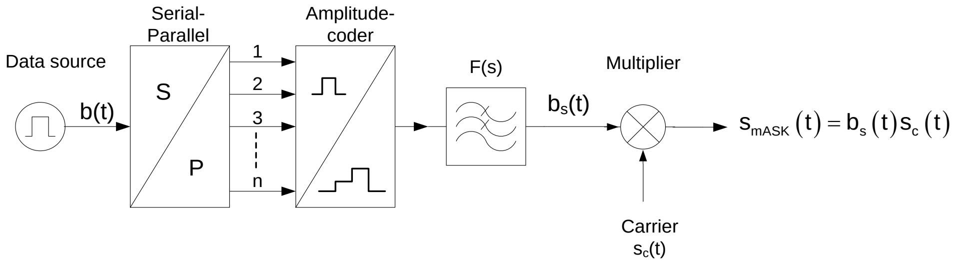


Fig. 3-56: mASK-Modulator

3.3.1.3 Considerations for the Multiplier:

In nearly all systems of digital modulation, multipliers (mixers) are used for modulation and demodulation. In the modulator, the baseband signal to be transferred is multiplied by a sinusoidal carrier after pulse shaping. Since the multiplication cannot be ideally realized, a bandpass filter follows at the multiplier output in order to suppress unwanted mixing products. If the baseband signal has a DC component (unipolar), then a spectral line at the carrier frequency will be generated in the spectrum of the modulated signal. If the baseband signal does not have any DC component (Manchester), the carrier will be suppressed.

When one observes the multiplication of two baseband signals, one unipolar with DC component and one bipolar without a DC component, by the carrier, the following description is possible:

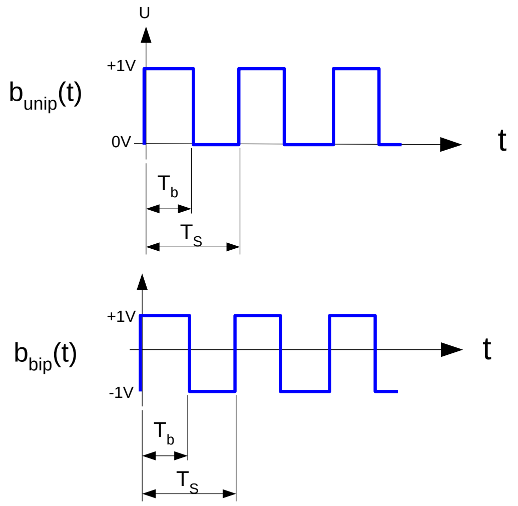


Fig. 3-57: Unipolar and bipolar modulation signal

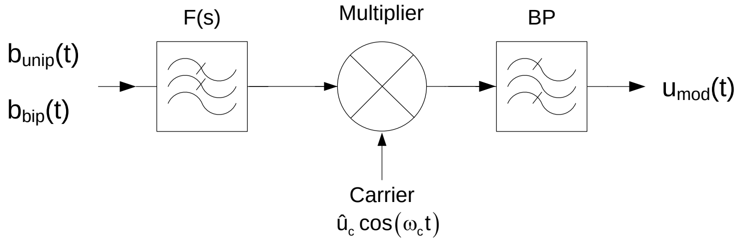


Fig. 3-58: Filtering and multiplication with a sinusoidal carrier

Both baseband signals can be described by their Fourier series:

$$b_{unip}(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{i=1}^{\infty}\frac{1}{i}\sin\frac{i\cdot 2\pi}{T_s}t$$

$$b_{bip}(t) = \frac{2}{\pi}\sum_{i=1}^{\infty}\frac{1}{i}\sin\frac{i\cdot 2\pi}{T_s}t \tag{3.54}$$

If these baseband signals are multiplied by the carrier

$$u_c(t) = \hat{u}_c\cos(\omega_c t) \tag{3.55}$$

the result is:

$$u_{modunip}(t) = u_c(t)\cdot b_{unip}(t) = \frac{\cos(\omega_c t)}{2} + \frac{2}{\pi}\sum_{i=1}^{\infty}\left[\sin\left(\frac{i\cdot 2\pi}{T_s}+\omega_c\right)t + \sin\left(\frac{i\cdot 2\pi}{T_s}-\omega_c\right)t\right] \tag{3.56}$$

and

$$u_{modbip}(t) = u_c(t) \cdot b_{bip}(t) = \frac{2}{\pi}\sum_{i=1}^{\infty}\frac{1}{2i}\left[\sin\left(\frac{i \cdot 2\pi}{T_s} + \omega_c\right)t + \sin\left(\frac{i \cdot 2\pi}{T_s} - \omega_c\right)t\right] \tag{3.57}$$

With an unipolar baseband signal the carrier with half amplitude appears plus the upper and lower sideband.

With a bipolar baseband signal, the carrier is not present, but only both sidebands.

Circuits for Mixer, Multiplier

As already mentioned, **double balanced modulators** (DBM) are mostly what is used in digital modulation systems. A very common type is the diode ring modulator. Its simple circuit is weighted against the disadvantage of high LO-power requirement, which is typically +7 dBm.

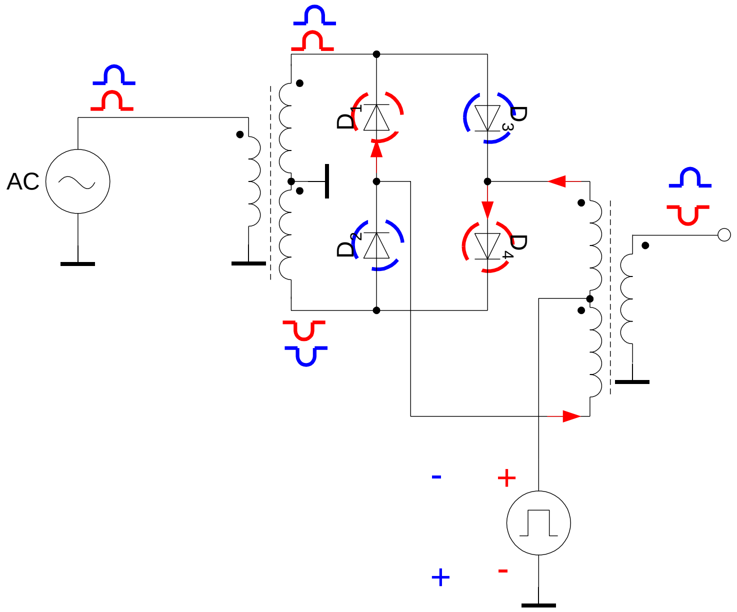


Fig. 3-59: Double-Balanced Diode Modulator

The more elaborate **Gilbert-Cell** (named after its inventor Barry Gilbert) has various advantages:

- Simple integration in an IC
- Only one transformer is necessary (if any)
- Low LO-power requirement of approx. –10dBm

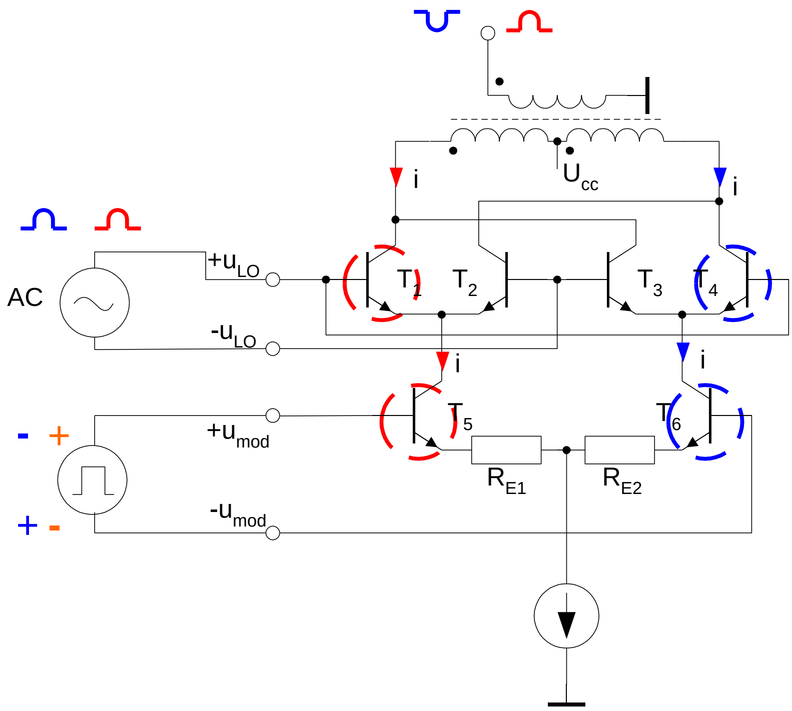


Fig. 3-60: Gilbert-Cell Modulator

3.3.1.4 Demodulation of ASK

Demodulation can be coherent or incoherent. Demodulation is understood as coherent, in which the demodulation carrier is phase locked to the transmitter carrier (synchronous demodulator). Incoherent demodulation is e.g. the envelope detector.

For coherent demodulation, the carrier has to be recovered from the received signal.

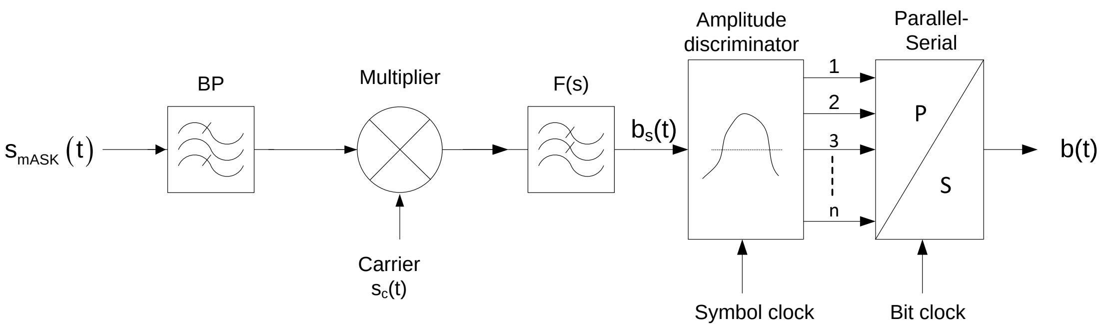


Fig. 3-61: Coherent mASK-Demodulator

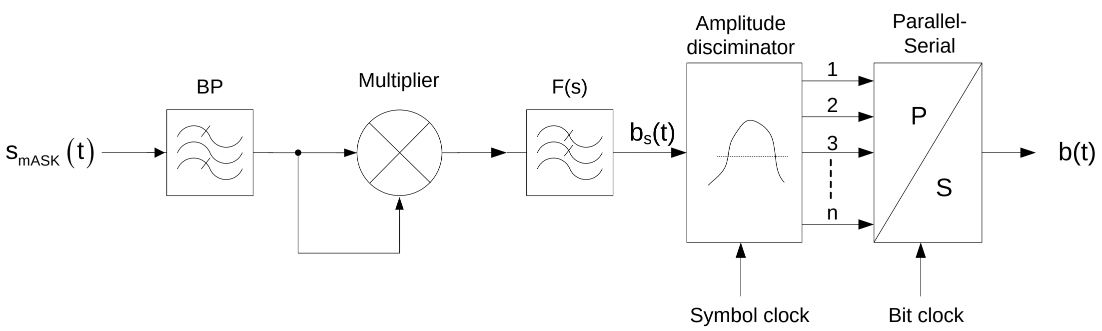


Fig. 3-62: Coherent mASK-Demodulator with squaring

The principle of demodulation with squaring will be analyzed in later discussion of PSK.

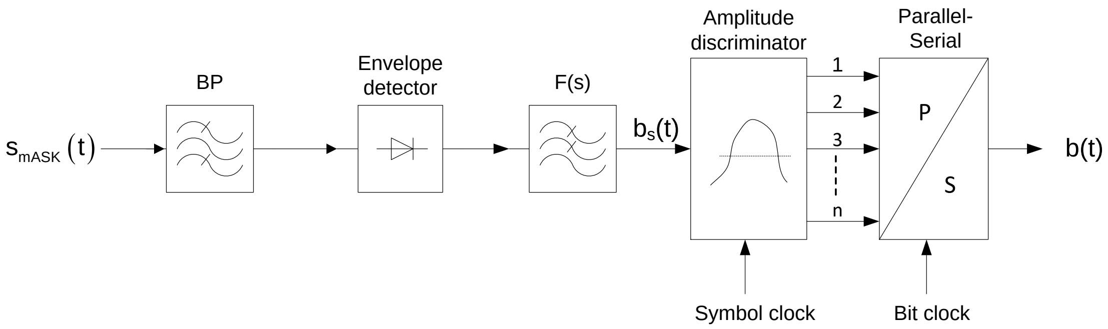


Fig. 3-63: Incoherent mASK-Demodulator

The discrete signal states are visually represented in the **state diagram**. Each point characterizes a mASK symbol in the modulation interval. The highest amplitude in the state diagram is usually normalized to 1. This representation allows to visually estimate the immunity to interference and noise. As the distance of the signal points decreases, the susceptibility to interference decreases because the distance to the decision limit decreases (Fig. 3-64 e).

With coherent demodulation it is more advantageous to use mASK-systems with state diagrams, in which the signal state “zero“(no carrier, Fig. 3-64 a and c) is also used.
With incoherent demodulation however it is recommended not to use the signal point “zero“ (no carrier) (Fig. 3-64 b and d).

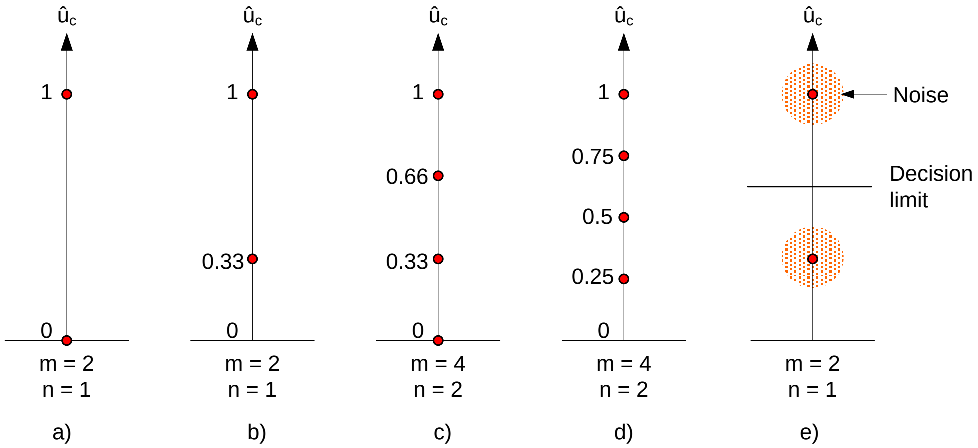


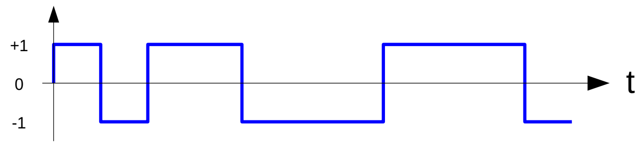
Fig. 3-64: State diagram of some ASK-systems

3.3.2 Phase Shift Keying PSK

3.3.2.1 Binary Phase Shift Keying, BPSK

With the binary modulation signal b(t), the phase of the carrier is keyed between discrete phase values. If the phase is keyed only between two discrete values ($0^{o}, 180^{o}$) the procedure is called Binary Phase Shift Keying (BPSK). Multi-value Phase Shift Keying is also possible (m-PSK). Systems with 4 phase states (Quadrature Phase Shift Keying QPSK) are very common. Similarly common is Differential Phase Shift Keying Modulation (DPSK).

We can describe the binary signal consisting of a serial bit stream of “0” and “1“ as follows:



$$b(t) = \begin{vmatrix} +1 & \text{"1"} \\ -1 & \text{"0"} \end{vmatrix} \qquad (3.58)$$

Fig. 3-65: Bipolar modulation signal

The BPSK-signal is again generated with multiplying the baseband signals by the carrier signal:

$$u_{BPSK}(t) = b(t) \cdot u_c(t) = b(t) \cdot \hat{u}_c \cos(\omega_c t) \qquad (3.59)$$

This yields:

$$s_{BPSK}(t) = \hat{u}_c \begin{vmatrix} \cos(\omega_c t) & \text{"0"} \\ \cos(\omega_c t + \pi) & \text{"1"} \end{vmatrix} \qquad (3.60)$$

$$\cos(\omega_c t + \pi) = -\cos(\omega_c t) \qquad (3.61)$$

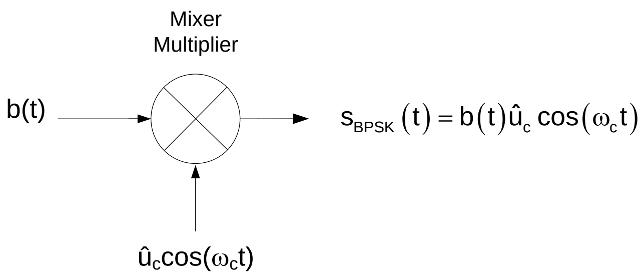


Fig. 3-66: BPSK-modulator without filter

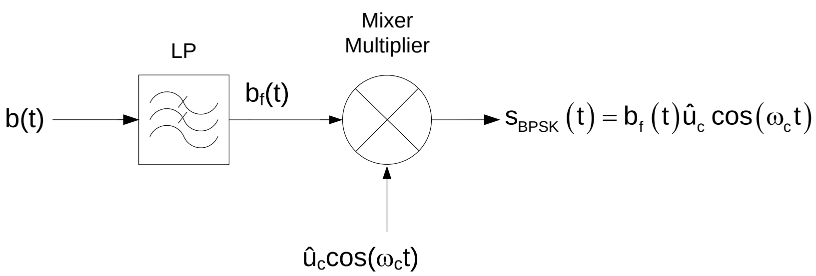


Fig. 3-67: BPSK-modulator with filter

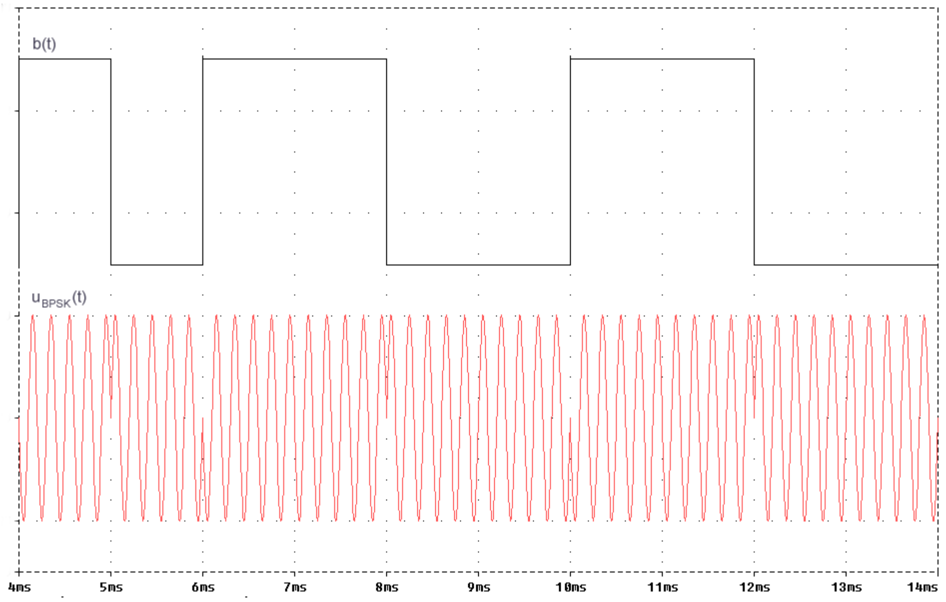


Fig. 3-68: BPSK-signal (unfiltered) in the time domain

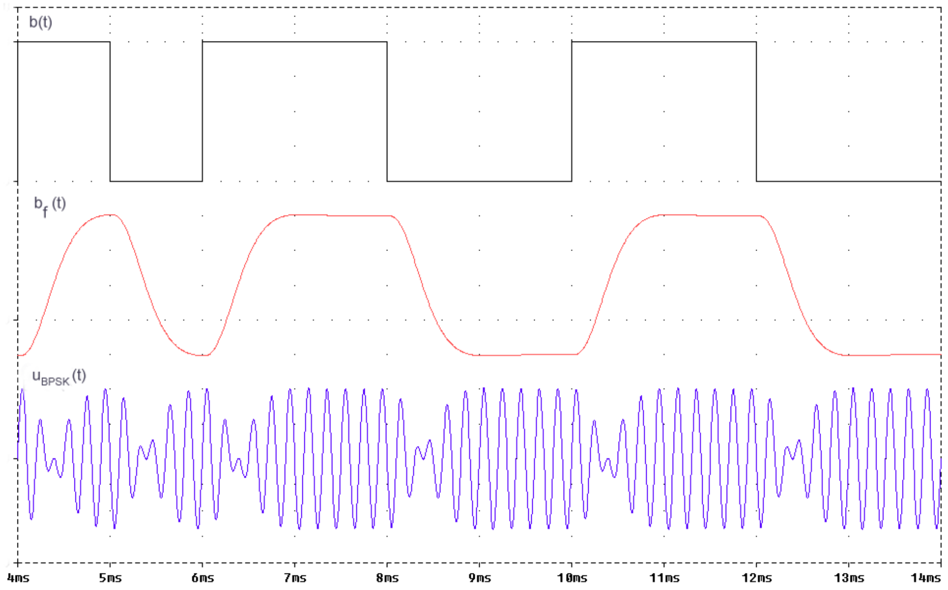


Fig. 3-69: BPSK-signal (filtered) in the time domain

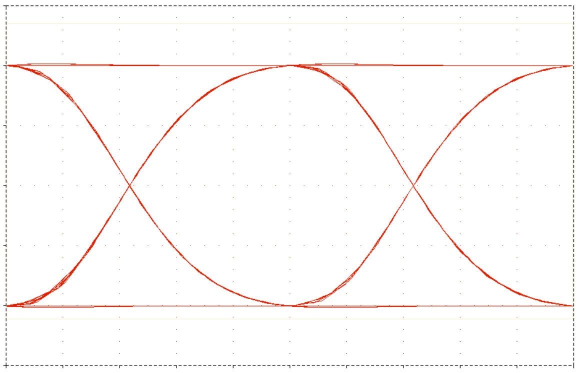


Fig. 3-70: Eye diagram of $b_f(t)$

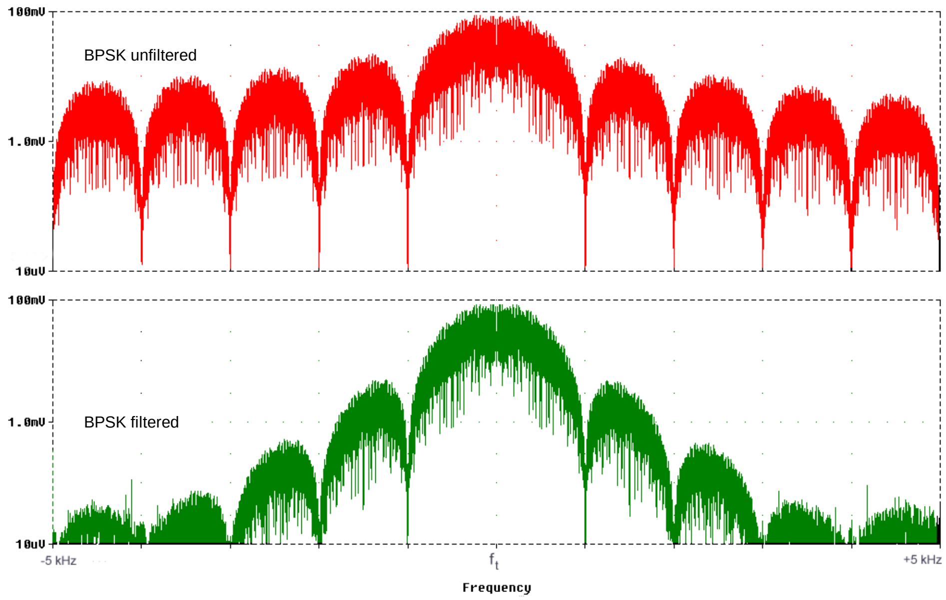


Fig. 3-71: BPSK-signal (unfiltered and filtered) in the frequency domain

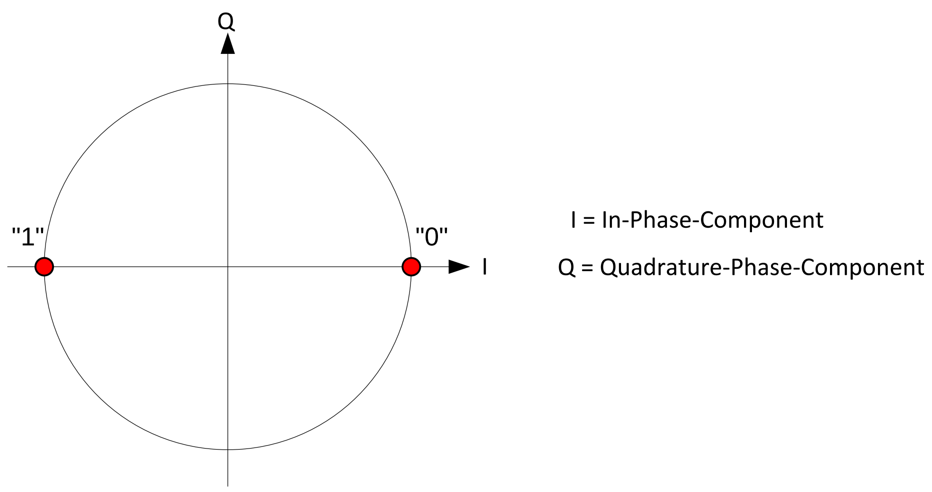


Fig. 3-72: Phase state diagram of a BPSK-signal

3.3.2.2 Quadrature-PSK (QPSK)

A multi-level PSK can be used exactly like multi-level amplitude keying. 4-PSK (Quadrature Phase Shift Keying QPSK) is particularly common. QPSK is generated through over-layering two BPSK-systems, whereby the carrier of the one system has a phase shift of 90 degree compared to the carrier of the second system.

Bit	Symbol
00	Symbol 1
01	Symbol 2
10	Symbol 3
11	Symbol 4

Fig. 3-73: Symbols

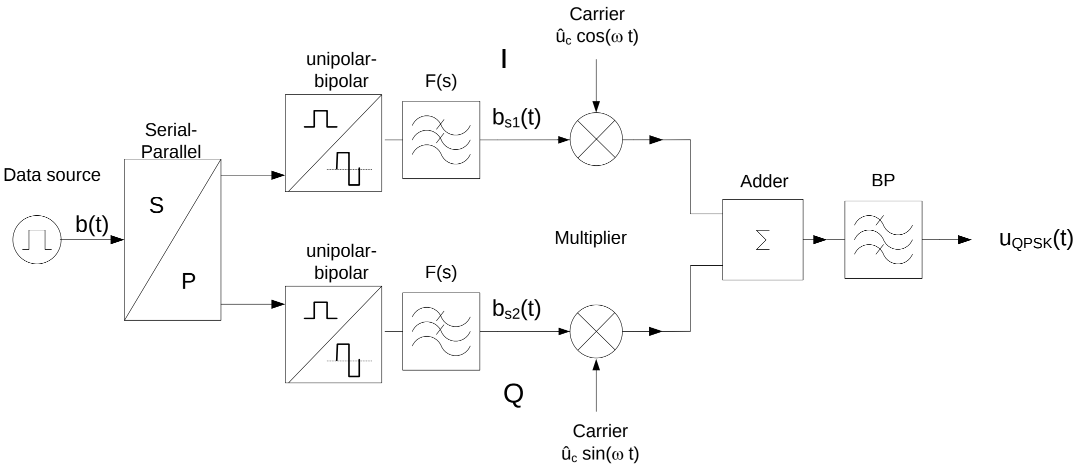


Fig. 3-74: QPSK-modulator

A serial-parallel-conversion and data filtering is used generate I- and Q-signals (I = In phase, Q = Quadrature phase).

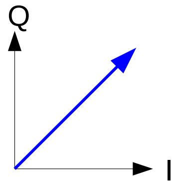
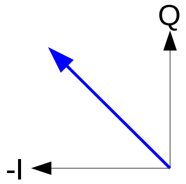
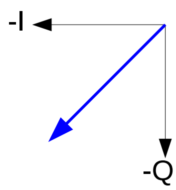
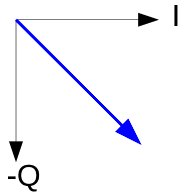
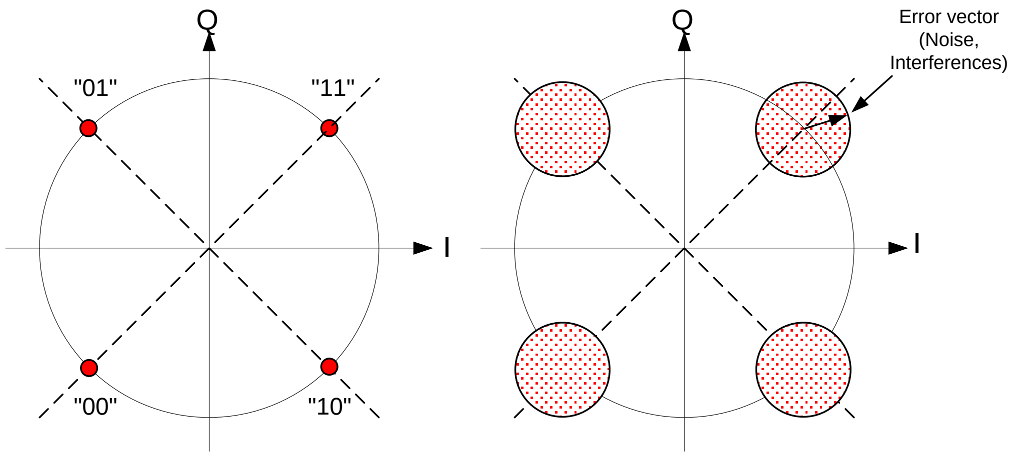
I $\left[b_{s1}(t)\right]$	Q $\left[b_{s2}(t)\right]$	$u_{QPSK}(t)$	Vector diagram
+1	+1	$+\cos(\omega_c t)+\sin(\omega_c t)$	
-1	+1	$-\cos(\omega_c t)+\sin(\omega_c t)$	
-1	-1	$-\cos(\omega_c t)-\sin(\omega_c t)$	
+1	-1	$+\cos(\omega_c t)-\sin(\omega_c t)$	

Fig. 3-75: Vector diagram for 4 symbols



Gray-coded

I = In-Phase-Component
Q = Quadrature-Phase-Component

Fig. 3-76: IQ-diagram

The symbol assignment as shown here is called Gray-Coding. The advantage of this type of coding lies in that only one bit is incorrected identified if a symbol error appears in an adjacent quadrant.

3.3.2.3 Circuits for carrier quadrature and generation of symbols

Carrier quadrature

For low carrier frequencies up to a few 100 kHz, an Allpass-Phase Shifter can be used in active filtering technology with OpAmp’s. With two 2nd order filters as shown in the illustration below, it is possible to achieve a phase precision of a few degrees over a frequency range of a decade.

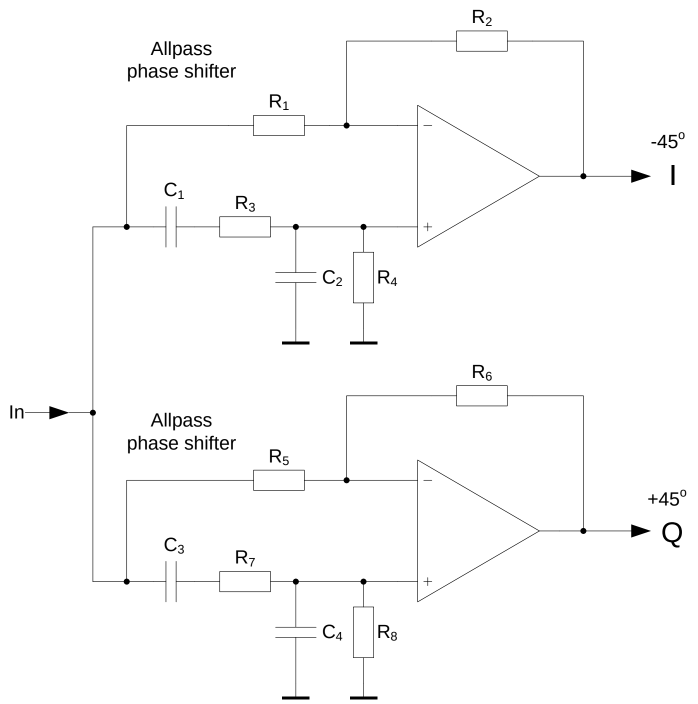


Fig. 3-77: All-pass phase shifter

In a frequency range of up to a few 100 MHz, a high-pass low-pass circuit with lumped L and C is an easily realizable solution. The bandwidth is however highly restricted.

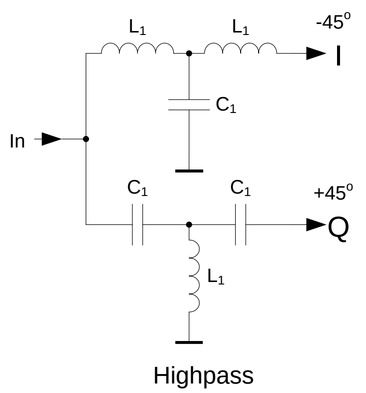


Fig. 3-78: Low-pass-high-pass phase shifter

In the GHz-range, solutions with directional couplers can be easily realized (here is an example of a branch line-coupler with micro-strip line). When using substrates with a high ε_r small dimensions result. The bandwidths lie in the magnitude of an octave.

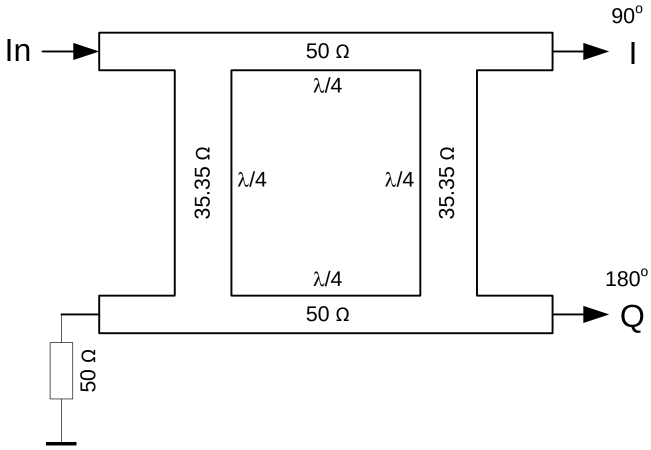


Fig. 3-79: Branchline-coupler using mikrostrip lines

Digital solutions with flip-flops are only limited in their bandwidth by the maximum clock frequency of the logic technology in use (up to a few GHz). The main disadvantage is that there must be double or quadruple carrier input frequencies.

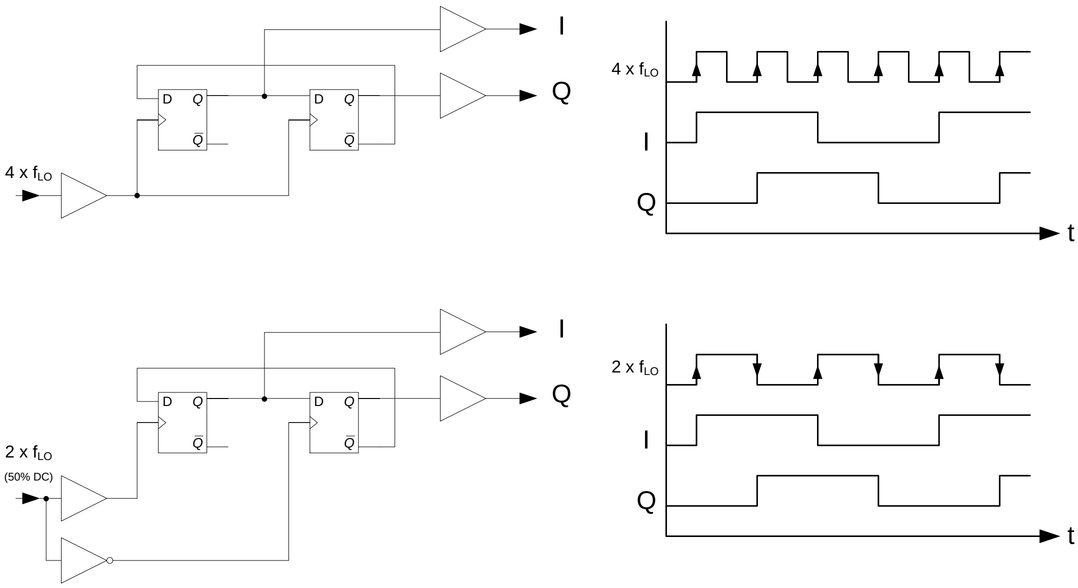


Fig. 3-80: Phase shifter using Flip-Flops

Generation of Symbols

A simple circuit to generate symbols with Gray-Coding is shown in the following figure:

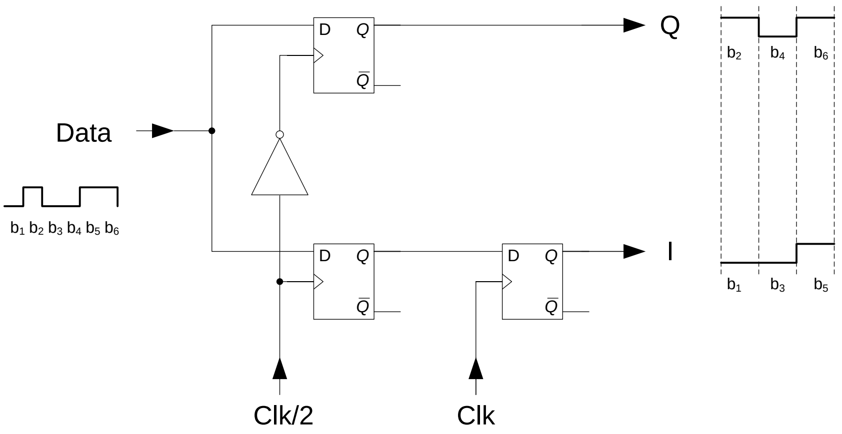


Fig. 3-81: Generation of symbols using Gray-Coding

3.3.2.4 Spectrum Efficiency

When two bits are combined into a symbol, the symbol frequency is only half as great as the bit frequency. This means that the bandwidth required for QPSK is only half as large as for BPSK, or that twice the bit rate can be transmitted in the same bandwidth. The spectrum efficiency is sometimes stated as Bit/s/Hz for an ideal Nyquist-System. The practically available spectrum efficiency is about 70% of the theoretical spectrum efficiency.

Modulation	Spectrum Efficiency Bit/s/Hz	Applications
MSK	1	GSM
BPSK	1	Telemetry, cable modems
OQPSK	1	Satellite communications
QPSK	2	Satellite communications, TETRA, CDMA, NADC, PHS, DVB-S, Modems
DQPSK, π/4-QPSK	3	NDAC, TACS
8PSK	3	Satellite communications, Telemetry, aeronautical radios
16QAM	4	Microwave radios, Modems, DVB-C, DVB-T
32QAM	5	Microwave radios, DVB-T
64QAM	6	DVB-C, Modems, Microwave radios
256QAM	8	DVB-C, Modems, Microwave radios

Fig. 3-82: Spectrum efficiency and applications of digital modulations

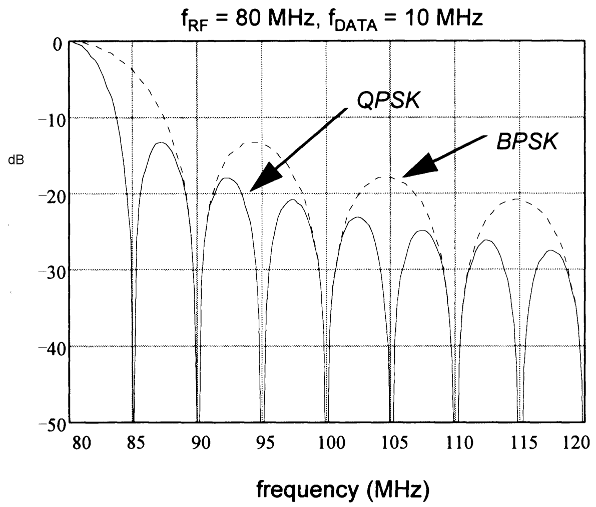


Fig. 3-83: Spectrum comparison BPSK-QPSK

The required RF-bandwidths with baseband filtering are calculated according to:

$$B_{BPSK} = \frac{1}{T_s}(1+\alpha) = \frac{1}{T_b}(1+\alpha)$$

$$B_{OQPSK} = \frac{1}{T_s}(1+\alpha) = \frac{1}{T_b}(1+\alpha)$$

$$B_{QPSK} = \frac{1}{T_s}(1+\alpha) = \frac{1}{2T_b}(1+\alpha) \qquad \begin{aligned} T_s &= \text{Symbol length} \\ T_b &= \text{Bit length} \\ \alpha &= \text{Roll-off-factor of the filter} \end{aligned} \tag{3.62}$$

$$B_{8PSK} = \frac{1}{T_s}(1+\alpha) = \frac{1}{3T_b}(1+\alpha)$$

$$B_{16QAM} = \frac{1}{T_s}(1+\alpha) = \frac{1}{4T_b}(1+\alpha)$$

The spectrum of the modulated signal results in multiplication of the baseband signal (symbols with multi-value modulation) by the carrier and is a two-sided spectrum around the carrier.

Example: Power spectral density of the BPSK modulation

A bipolar NRZ-baseband signal has the power spectral density of

$$G_{B-NRZ}(f) = U^2 T_b \left[\frac{\sin(\pi f T_b)}{\pi f T_b}\right]^2 \tag{3.63}$$

For BPSK-modulation, the baseband b(t) is multiplied by the carrier $\hat{u}_c \cos(\omega_c t)$

$$G_{BPSK}(f) = U^2 T_b \left[\frac{\sin(\pi f T_b)}{\pi f T_b}\right]^2 \hat{u}_C \cos(\omega_c t) \tag{3.64}$$

After a trigonometric transformation, the two-sided power spectral density is:

$$G_{BPSK}(f) = U^2 \hat{u}_C T_b \left\{\left[\frac{\sin(\pi T_b [f - f_c])}{\pi T_b [f - f_c]}\right]^2 + \left[\frac{\sin(\pi T_b [f + f_c])}{\pi T_b [f + f_c]}\right]^2\right\} \tag{3.65}$$

3.3.2.5 Offset QPSK (OQPSK)

In QPSK the amplitude has the value of zero for a short time when there is a phase shift of 180°. This means that large amplitude variation has to be processed in the system and the entire system chain must have linear behavior (no amplitude limiting). In order to avoid a phase shift of 180°, and thus reduce the amplitude variation, the Q-data are delayed by a bit period (1/2-symbol period).

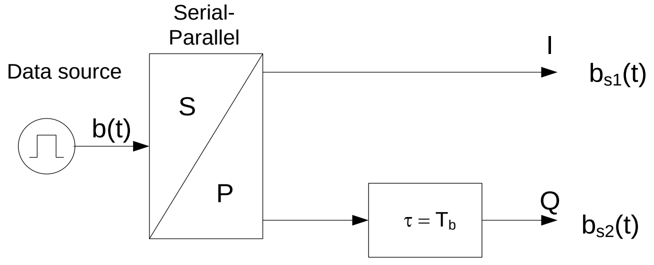


Fig. 3-84: Offset QPSK

This means, as shown in the impulse diagram below, that only phase jumps of ±90° occur and that the amplitude of the carrier reduce to the minimal value of 0.707 of the maximum value.

The advantage of spectrum efficiency of QPSK however is lost again in OQPSK and is equal to BPSK.

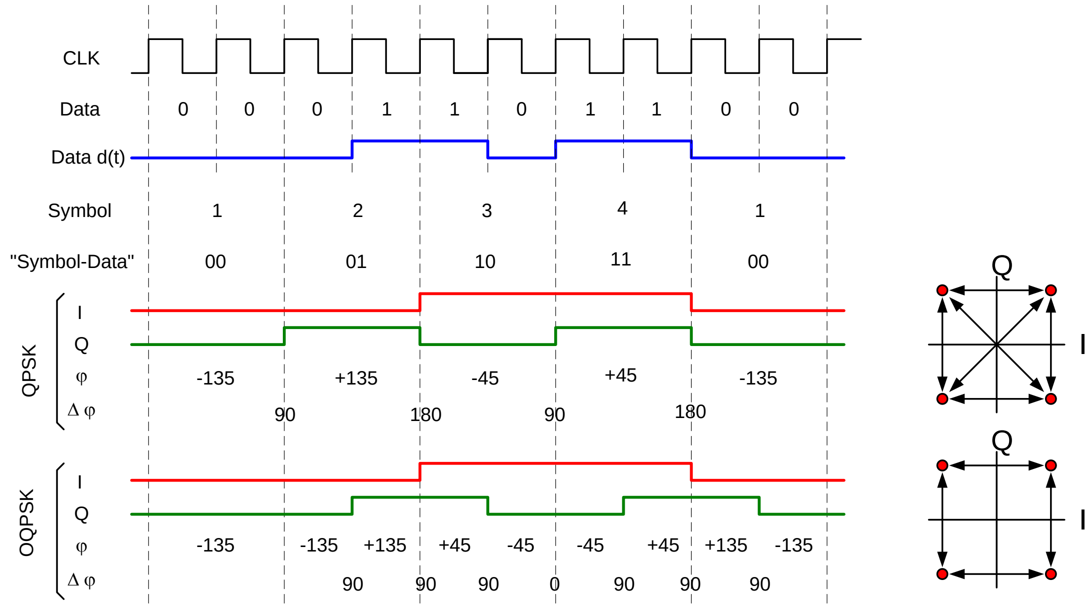


Fig. 3-85: QPSK and OQPSK

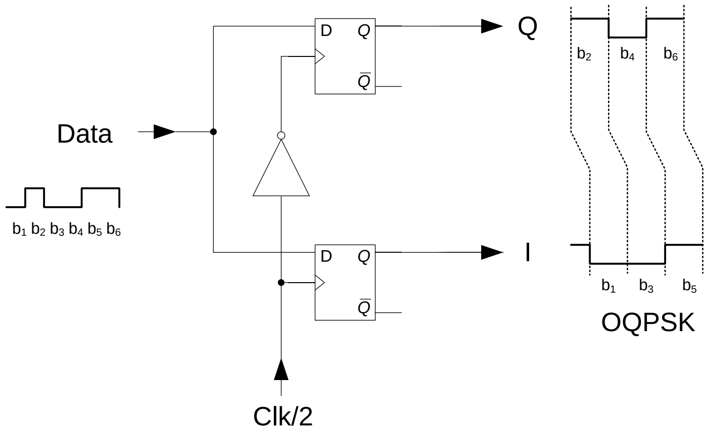


Fig. 3-86: Simple circuit to generate OQPSK-Symbols

3.3.2.6 Differential QPSK (DQPSK), π/4-QPSK

Another modulation procedure consists in only permitting phase jumps of $\pm\pi/4$ and $\pm3\pi/4$. The information is differential encoded: Symbols are transmitted as phase changes and not as absolute phase positions.

Symbol	Phase Change
00	$\pi/4$
01	$3\pi/4$
10	$-\pi/4$
11	$-3\pi/4$

Fig. 3-87: Phase change for DQPSK

This yields 8 possible phase states with a spectrum efficiency of 3 Bit/s/Hz. The amplitude variation is greater than with OQPSK but smaller than with QPSK.

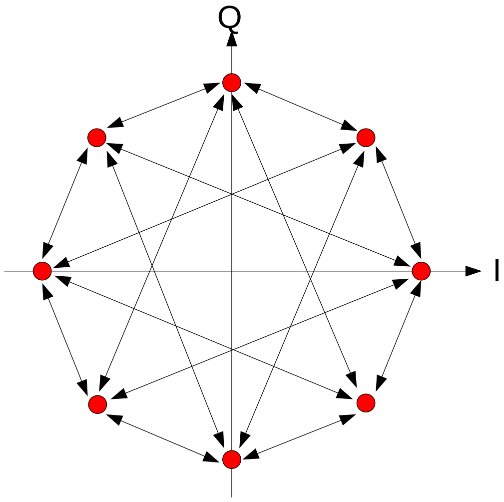


Fig. 3-88: Phase transitions for π/4-QPSK

Since there is a phase change for every symbol, the clock recovery on the receiver side is especially easy. Likewise the demodulation can be incoherent (not synchronous), which means that circuit can be simplified.

3.3.2.7 Quadrature Amplitude Modulation QAM

If two QPSK-systems are combined as shown below, one gets a 16-QAM-system, i.e. there are 16 states of phase and amplitude. The generation of these 16 states can easily be derived from vector addition of the individual modulation components. 16-QAM is used as a standard in microwave radio systems with bitrates of 140 Mbit/s.

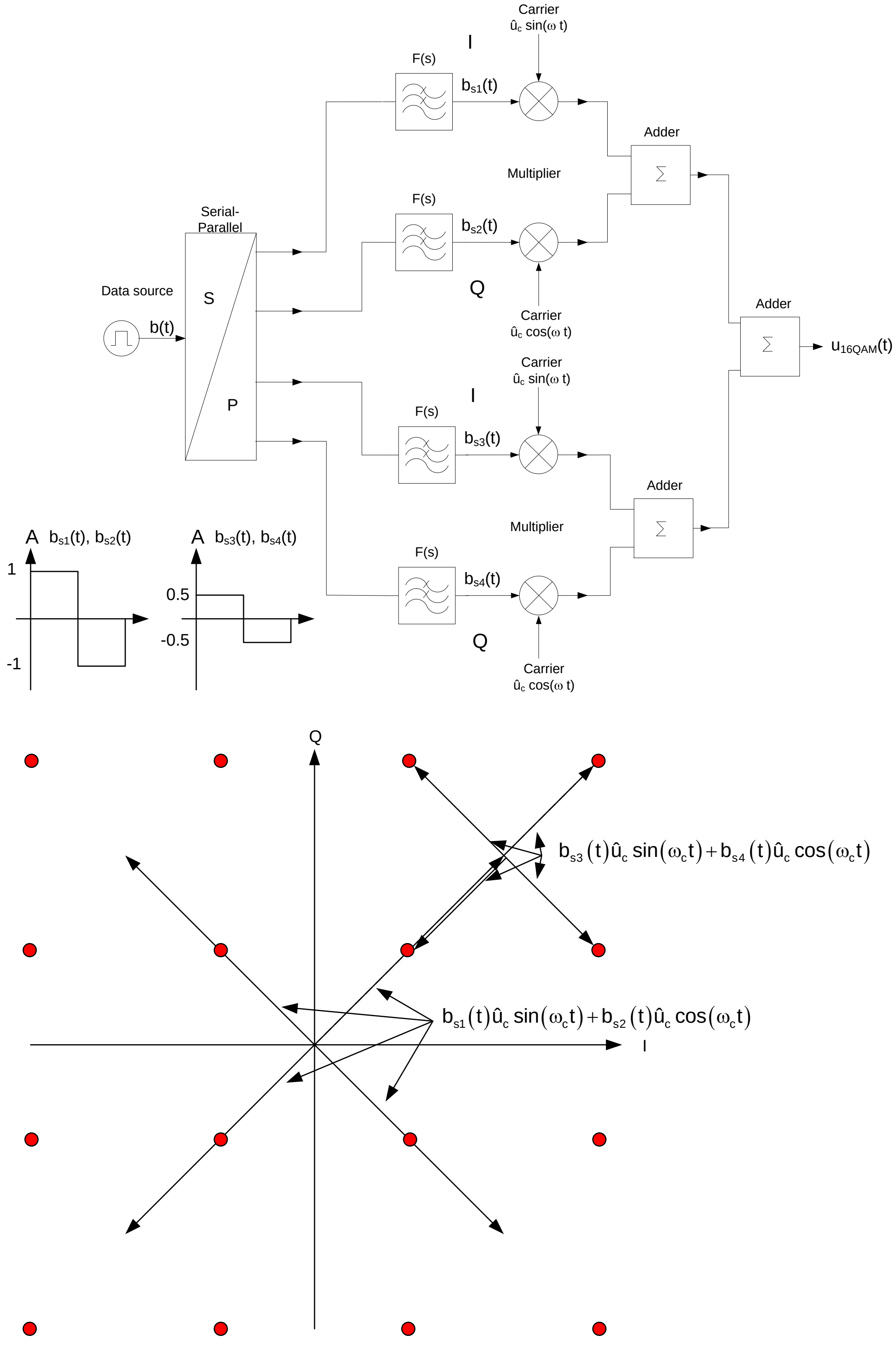


Fig. 3-89: Circuit and IQ-Diagram for 16-QAM

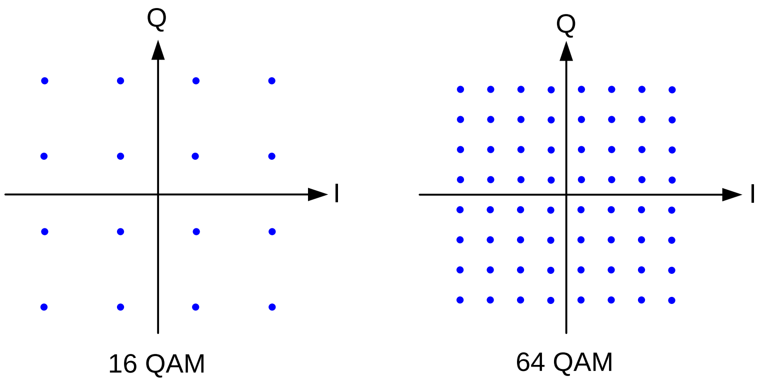


Fig. 3-90: IQ-Diagramm für 16QAM, 64QAM

3.3.2.8 Demodulation of PSK

Basically the demodulation is the reversal of the modulation circuit. For synchronous demodulation (coherent), the carrier must be derived and phase-locked from the received signal on the receiver side. Likewise, the data- or symbol clock must be recovered.

Synchronous demodulation of BPSK

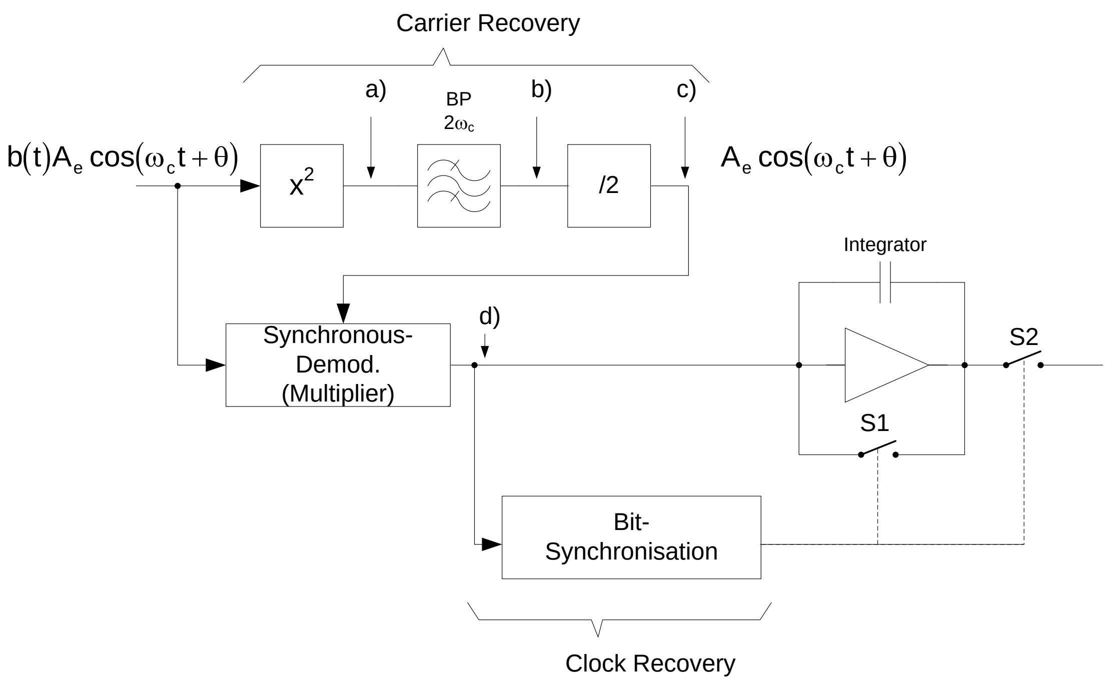


Fig. 3-91: Block diagram for BPSK-Demodulation

The squaring of the input signal yields (point a in the block diagram):

$$\left[b(t)A_e\cos(\omega_c t+\theta)\right]^2 = A_e^{\,2}\left[\frac{1}{2}+\frac{1}{2}\cos 2(\omega_c t+\theta)\right] = \frac{1}{2}A_e^{\,2}\left[1+\cos(2\omega_c t+2\theta)\right] \qquad (3.66)$$

After the bandpass the following remains (DC removed) (point b):

$$\frac{1}{2}A_e^{\,2}\cos(2\omega_c t+2\theta) \qquad \text{b)} \qquad (3.67)$$

After dividing by 2, the carrier is recovered (point c):

$$A_e\cos(\omega_c t+\theta) \qquad \text{c)} \qquad (3.68)$$

On multiplying the input signal by the recovered carrier, one gets:

$$b(t)A_e\cos(\omega_c t+\theta)\hat{u}_c\cos(\omega_c t+\theta)$$
$$=b(t)A\cos^2(\omega_c t+\theta)=b(t)A\frac{1}{2}\left[1+\cos(2\omega_c t+2\theta)\right] \quad \text{d)} \qquad (3.69)$$

The values of the amplitudes are not important in these discussions.

The bit-synchronization has the following function:

- The end of a bit is recognized

At the end of a bit switch S1 is briefly closed in order to discharge the integrator-C. Shortly <u>before</u> S1 is closed, S2 is briefly closed in order to get a sample of the output of the integrator. This sample is the desired output signal.

For the analysis, we will, for simplicity's sake, assume that the bit length T_b is equal to an even number n cycles of the carrier frequency f_c:

$$T_b = nf_c$$

In this case the output voltage from the integrators at the end of a bit-intervals of $(k\text{-}1)T_b$ to kT_b according the equation (3.69) is:

$$u_o(kT_b)=\int_{(k-1)T_b}^{kT_b} b(kT_b)A\frac{1}{2}\left[1+\cos(2\omega_c t+2\theta)\right]dt$$
$$=b(kT_b)A\int_{(k-1)T_b}^{kT_b}\frac{1}{2}dt+\underbrace{b(kT_b)A\int_{(k-1)T_b}^{kT_b}\frac{1}{2}\cos 2(\omega_c t+\theta)dt}_{=0,\text{ Integral over a whole cycle}} \qquad (3.70)$$

$$=b(kT_b)T_b\frac{A}{2}$$

This proves that this demodulator provides an image of the transmitted bit sequence b(t).

Synchronous demodulation of QPSK

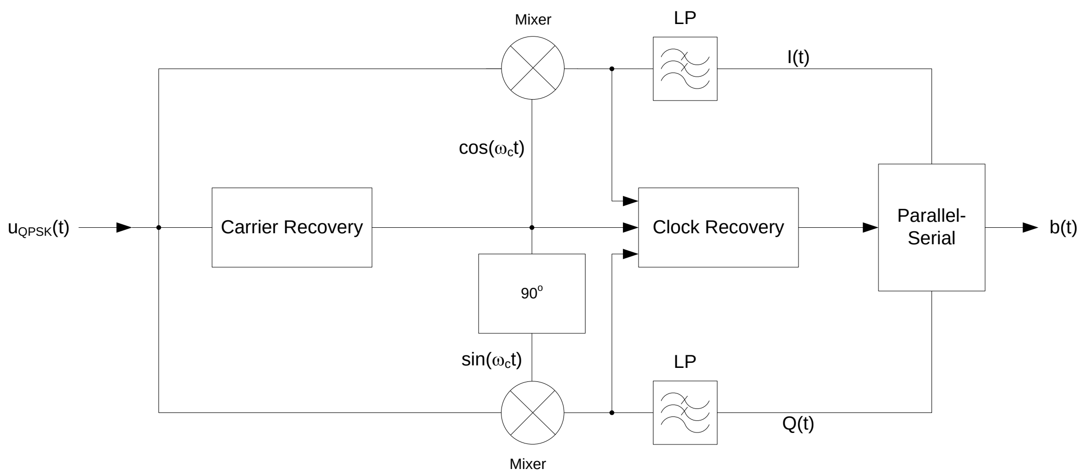


Fig. 3-92: Block diagram for QPSK-Demodulation

More Circuit Details

Costas-Loop:

A Costas-Loop consists of a PLL-control loop for recovery of the carrier frequency. When frequency or phase deviations arise between the input signal and the VCO a control voltage is generated u3(t), which readjusts the VCO.

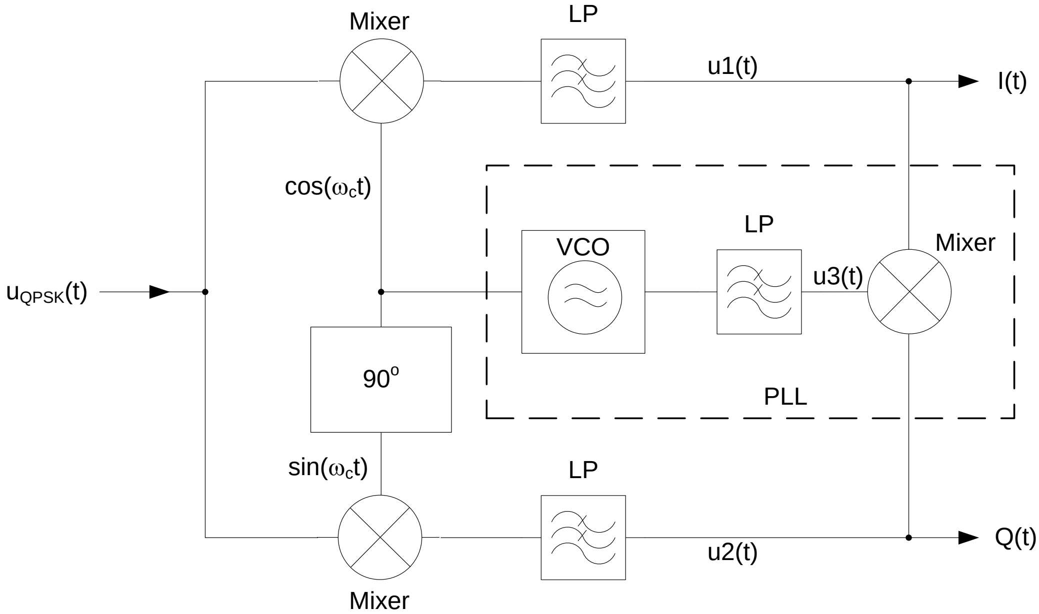


Fig. 3-93: Block diagram of Costas-Loop

Clock Recovery:

The bit- or symbol clock can either be derived directly from the input signal (clock information must be included in the amplitude) or derived from the demodulated signal.

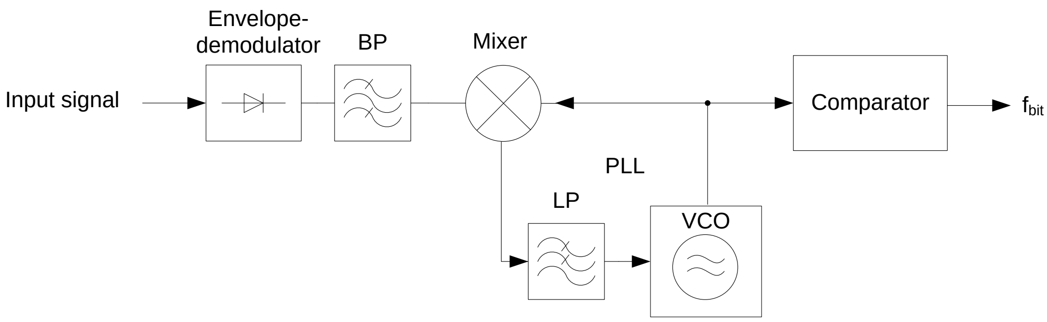


Fig. 3-94: Block diagram of clock recovery

Data Recovery and Sampler:

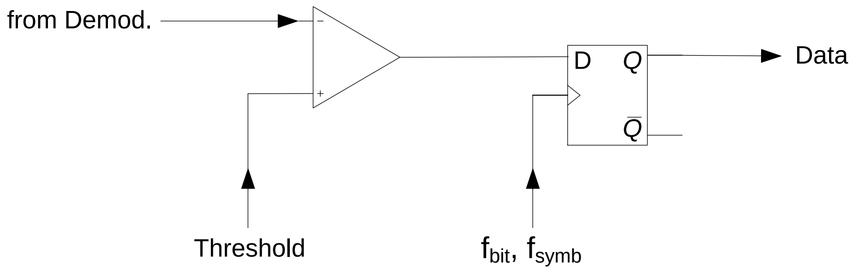


Fig. 3-95: Threshold detector and sampler

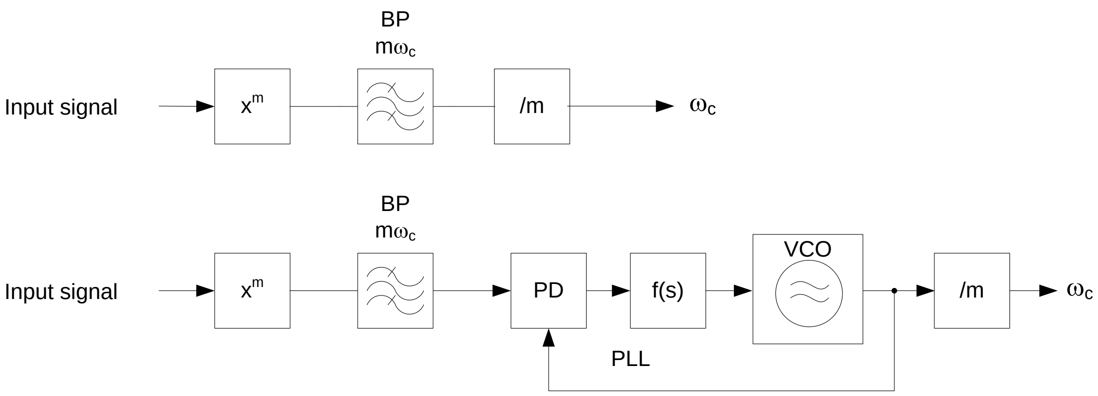


Fig. 3-96: Carrier recovery for m-ary-PSK

3.3.3 Frequency Shift Keying FSK

In Frequency Shift Keying one distinguishes between FSK with **discontinuous** phase change and FSK with **continuous** phase change.

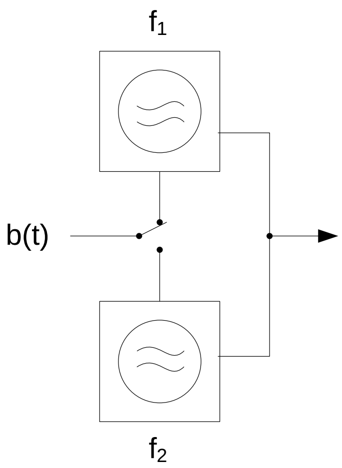


Fig. 3-97: Discontinuously frequency shift keying

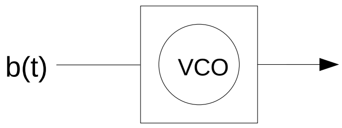


Fig. 3-98: Continuously frequency shift keying

Phase modulation can also be regarded as frequency modulation. The relationships between phase and frequency are known:

$$\omega(t) = \frac{d\varphi(t)}{dt} \qquad \varphi(t) = \int \omega(t)dt \tag{3.71}$$

As in Phase Shift Keying there are also various types of Frequency Shift Keying.

- Tamed Frequency Modulation, **TFM**: similar to MSK, but with an even steeper spectrum amplitude drop
- Four Frequency Keying, **4-FSK**: This is used in ERMES.
- **MSK** (Minimum Shift Keying)
- **GMSK** (Gauss-filtered Minimum Shift Keying)

GMSK (GSM, DECT) has by far the largest group of users. For that reason we will devote most of our attention here to this variant.

3.3.3.1 Minimum Shift Keying MSK and Gauss-filtered Minimum Shift Keying GMSK

MSK has the following characteristics:

- MSK can be regarded as a phase- or frequency modulation
- The phase rotates during the time of a bit length by $\pm 90^o$
- MSK is a frequency modulation with the modulation index of 0.5

$$\beta = \frac{\Delta f}{f_m} = \frac{H}{f_m} = 0.5 \tag{3.72}$$

$$H = \frac{f_m}{2} \qquad \begin{array}{l} H = \text{ Deviation} \\ f_m = \text{ Modulation frequency = Bit frequency} \end{array} \tag{3.73}$$

MSK = FFSK (Fast Frequency Shift Keying)

Advantages:

Constant amplitude (non-linear amplifiers can be used)

Sideband amplitudes drop off more rapidly than in BPSK, QPSK

Lower ISI

Disadvantages:

Main lobe of the spectrum is 1.5 times broader than in QPSK

Higher BER then QPSK at same S/N

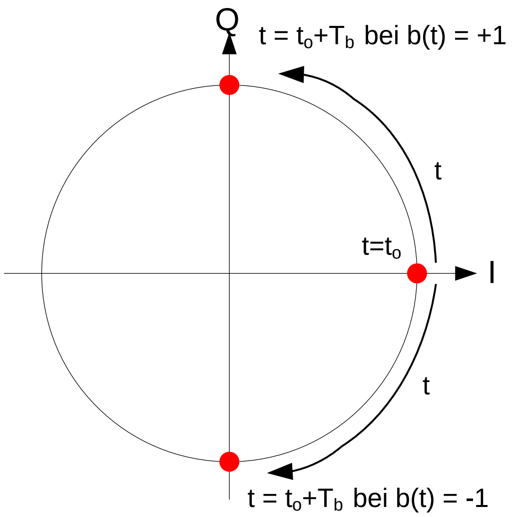


Fig. 3-99: I-Q-Diagram for MSK

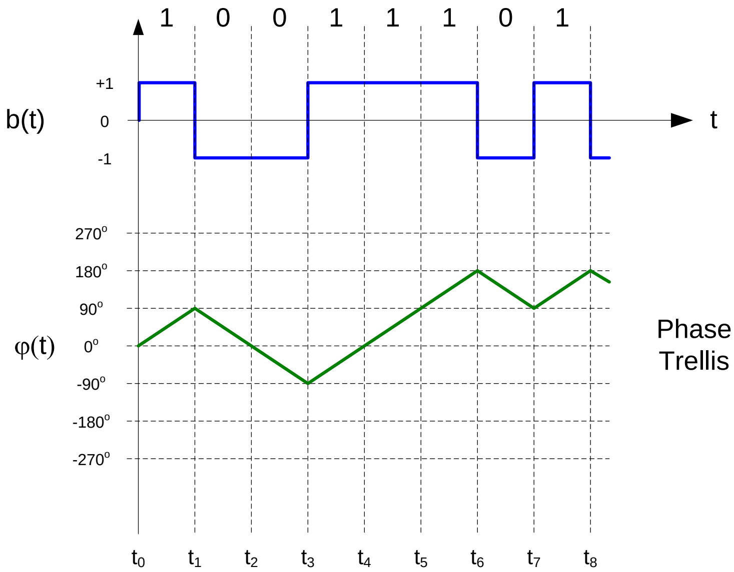


Fig. 3-100: Trellis-Diagram for MSK

The modulated voltage is calculated from:

$$u_{MSK}(t) = A \cdot \cos\left[\omega_c t + \int_{-\infty}^{t} d(t) \cdot \frac{\pi}{2} dt\right] = A \cdot \cos\left[\omega_c t + \frac{\pi}{2} \cdot (\pm 1) \cdot t + \varphi_o\right] \tag{3.74}$$

$$d(t) = \begin{vmatrix} +1 & "1" \\ -1 & "0" \end{vmatrix} \tag{3.75}$$

To determine the deviation, we assume that the phase changes during a bit length T_b is $b(t) = +1$ by $+90^o = \frac{\pi}{2}$ and for $b(t) = -1$ by $-90^o = \frac{-\pi}{2}$:

$$\Delta f^{+} = \frac{\Delta\omega}{2\pi} = \frac{\Delta\varphi(t)}{\Delta t \cdot 2\pi} = \frac{\pi/2}{T_b \cdot 2\pi} = \frac{1}{4T_b} = \frac{f_b}{4} \tag{3.76}$$

$$\Delta f^{-} = \frac{\Delta\omega}{2\pi} = \frac{\Delta\varphi(t)}{\Delta t \cdot 2\pi} = \frac{-\pi/2}{T_b \cdot 2\pi} = \frac{-1}{4T_b} = -\frac{f_b}{4} \tag{3.77}$$

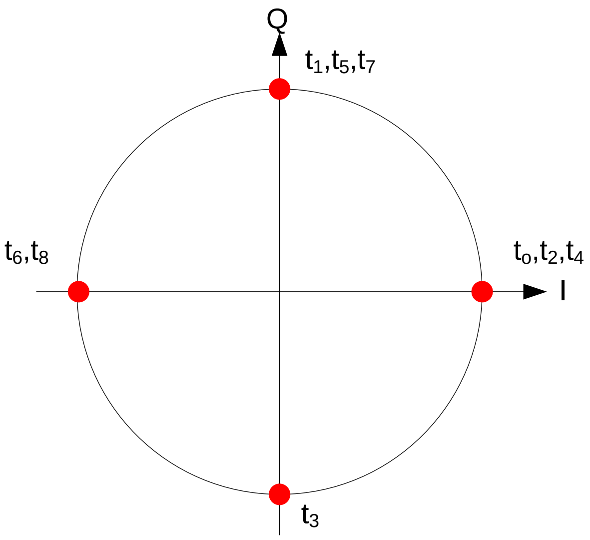
The frequency deviation H is then:

$$H = \Delta f^{+} - \Delta f^{-} = \frac{f_b}{4} - \left(-\frac{f_b}{4}\right) = \frac{f_b}{2} \tag{3.78}$$

The designation **Minimum** Shift Keying stands for:

Minimum frequency difference between „1“ and +0+ for synchronous demodulation.
A multiple of $\Delta\varphi = 90^{\circ}$ results in „1“- and „0“-frequencies that are always equal.

If an I-Q-modulator is used for MSK, then the voltages $u_I(t)$ and $u_Q(t)$ can be determined from the I-Q-diagram:



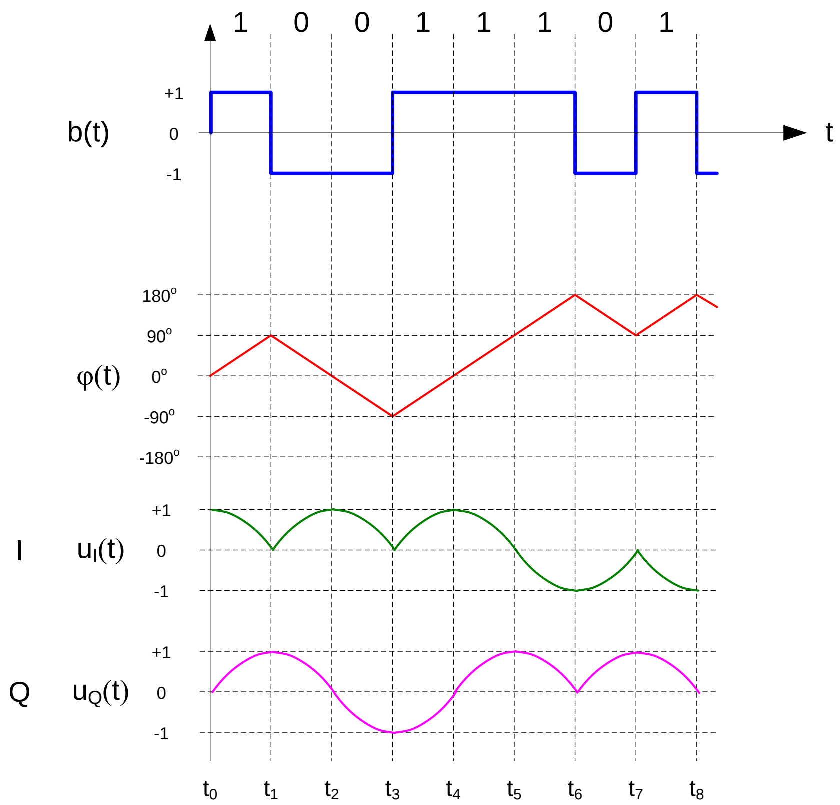


Fig. 3-101: Generation of I-Q-Voltages for MSK

Power spectral density of MSK:

$$S_{MSK}(f) = \frac{8}{\pi^2}\hat{u}\left[\left\{\frac{\cos\frac{2\pi(f-f_0)}{f_b}}{1-\left(\frac{4(f-f_0)}{f_b}\right)^2}\right\}^2 + \left\{\frac{\cos\frac{2\pi(f+f_0)}{f_b}}{1-\left(\frac{4(f+f_0)}{f_b}\right)^2}\right\}^2\right] \tag{3.79}$$

$$u_{MSK}(t) = A\cdot\left[b_e(t)\sin\frac{2\pi t}{4T_b}\right]\cos(\omega_c t) + A\cdot\left[b_o(t)\cos\frac{2\pi t}{4T_b}\right]\sin(\omega_c t) \tag{3.80}$$

$b_e(t) =$ even data bits (2,4,6,...), (I)

$b_o(t) =$ odd data bits (1,3,5,...), (Q)

In contrast to BPSK and QPSK the carrier is not abruptly switched with b(t), but “softly“ with $b_e(t)\sin(xt),\ b_o(t)\cos(xt)$.

Although the spectrum beside the carrier drops off very quickly, the side lobes in an adjacent channel still interfere. The side lobes can be further limited if the hard peaks of the phase change are “rounded off”. Sudden phase changes can be prevented by impulse shaping in the baseband. If a Gauss-Filter is used for impulse shaping, one gets “Gaussian Minimum Shift Keying“ **GMSK**. This procedure is used in GSM (Groupe Spécial Mobile, Global System Mobile).

Depending on the bandwidth-time-product BT of the Gauss filter, the side lobes are further reduced though admittedly with the problem of a larger BER.

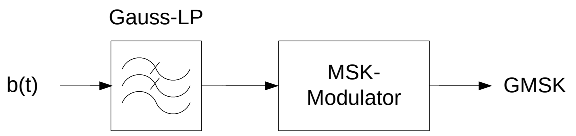


Fig. 3-102: Baseband Filter for GMSK

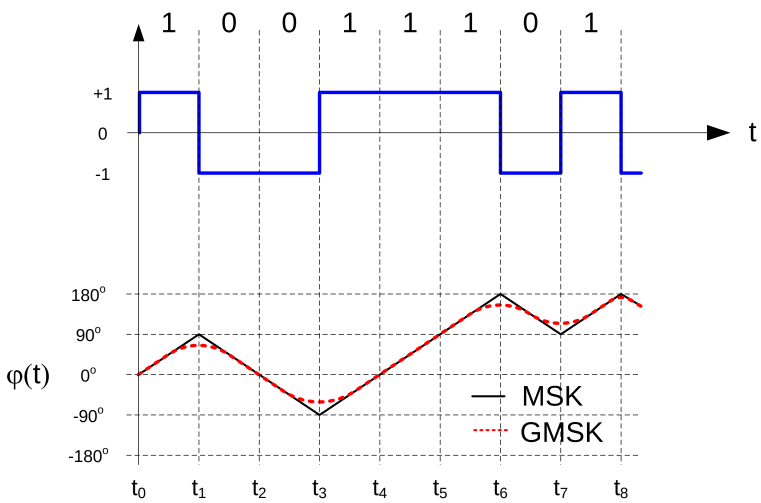


Fig. 3-103: Trellis-Diagram for MSK and GMSK

3.4 References

[1] Taub, H., Schilling, D.L.: *Principles of communication systems*. McGraw-Hill, 2nd Edition 1986

[2] Kammeyer, K.D. : *Nachrichtenübertragung*. Vieweg+Teubner, 4. Auflage 2008

[3] Roppel, C.: *Grundlagen der digitalen Kommunikationstechnik*. Carl Hanser Verlag, 2006

[4] Ohm, J-R., Lüke, H.D.:*Signalübertragung*. Springer Verlag Berlin, 10. Auflage 2007

[5] Schwartz, M.: Information, Transmission, Modulation, and Noise. McGraw-Hill, 1980

[6] Zinke, O., Brunswig, H.: *Hochfrequenztechnik 2*, Springer Verlag Berlin, 5. Auflage 1999

[7] Stumpers, F.L.M.H.: *Theory of frequency modulation noise*. Proc. Inst. Radio Engrs. 36, 1948, 1081-1092

[8] Rice, S.O. : *Statistical properties of a sine wave plus random noise*. Bell Syst. Techn.J. 27, 1948, 109-157

[9] von Grünigen, D.Ch.: *Digitale Signalerarbeitung mit einer Einführung in die kontinuierlichen Signale und Systeme.* Carl Hanser Verlag, 5. Auflage 2014

[10] von Grünigen, D.Ch.: *Digitale Signalerarbeitung: Bausteine, Systeme, Anwendungen.* Fotorotar Print und Media, 2008

[11] Dellsperger, F.: *Passive Filter der Hochfrequenz- und Nachrichtentechnik.* Lecture Script, 2012