ICTP-ITU-URSI School on Wireless Networking for Development
The Abdus Salam International Centre for Theoretical Physics ICTP, Trieste (Italy), 5 to 24 February 2007

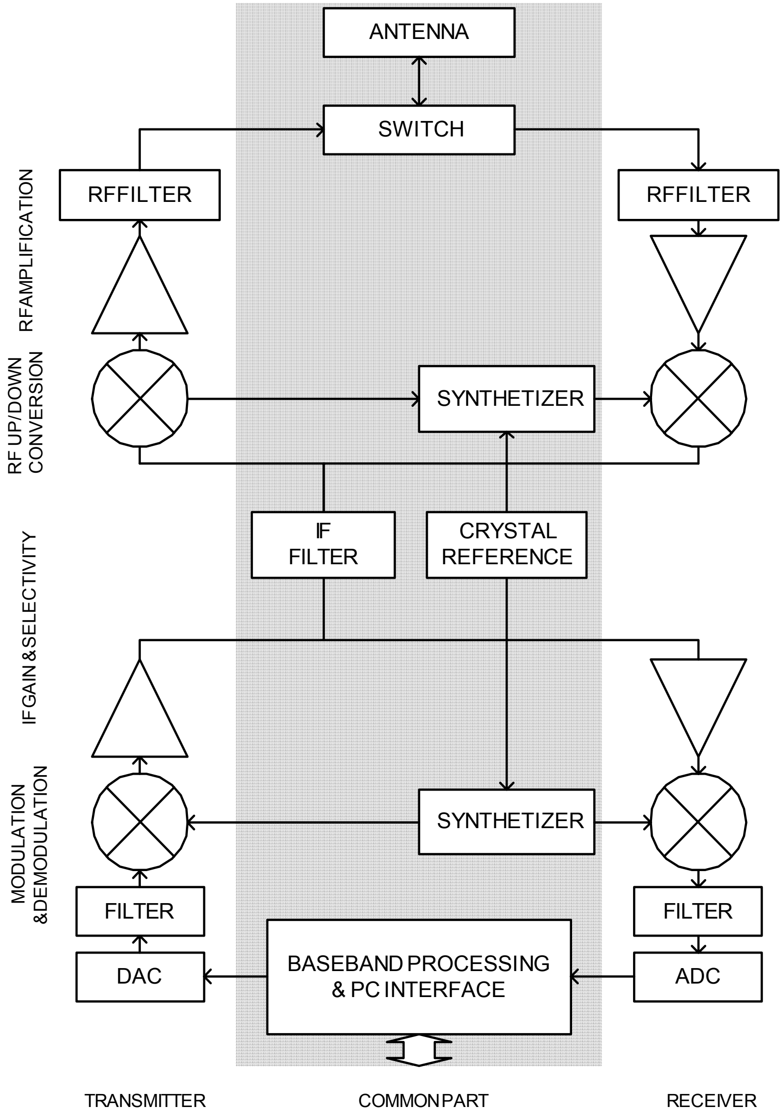
Channel & Modulation: Basics

Ryszard Struzak

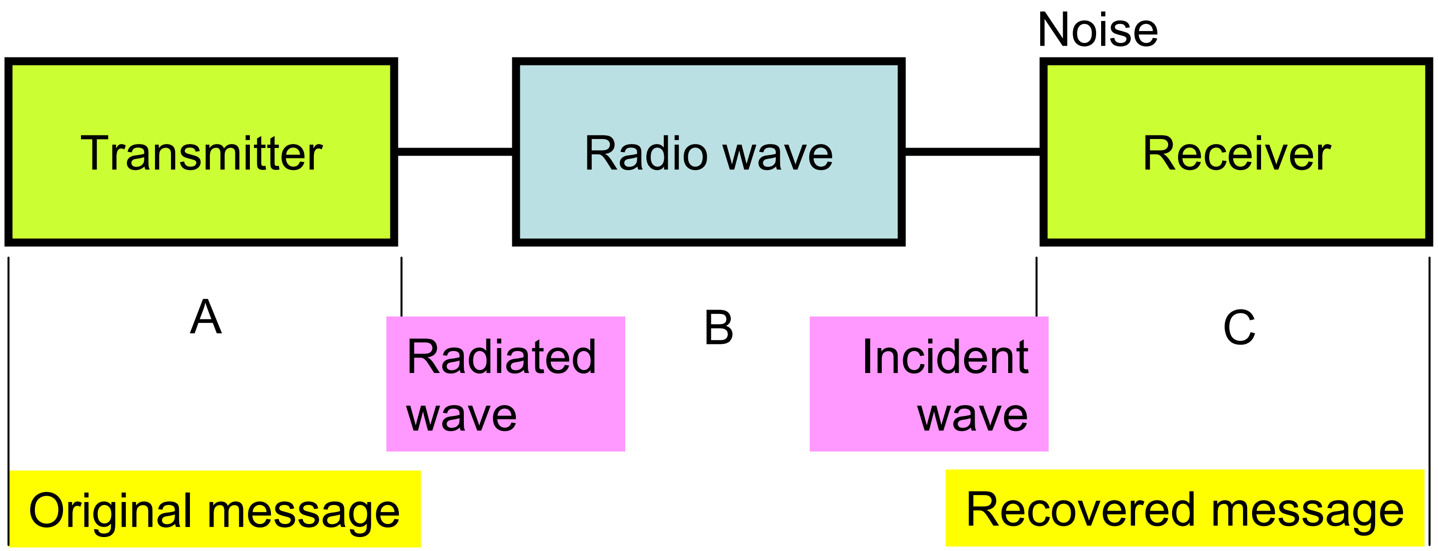
Outline

- Introduction to digital modulation
- Relevant modulation schemes
- Geometric representations
- Coherent & Non-Coherent Detection
- Modulation spectra

Modern radio



Isolated radio link

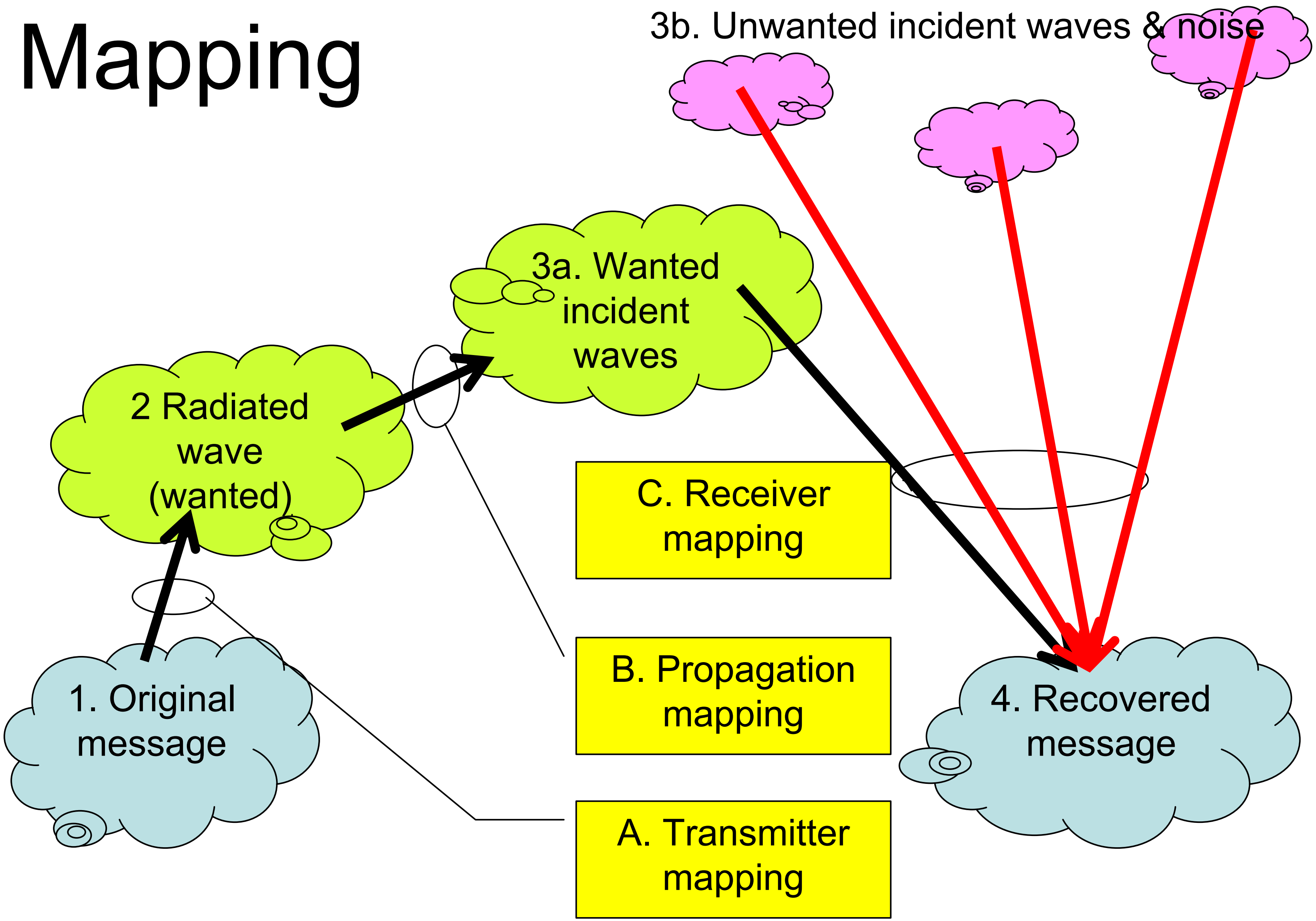


- Signal carrier & dimensionality vary along the channel
- Incident wave ≠ Radiated wave (noise & mapping errors)
- Recovered message ≠ original message

Transmission system

- A *message,* generated by a *source* of messages*,* to be delivered from the source to a distant *destination* via *telecommunication channel*
- The channel consists of a *transmitter node*, *propagation path* and *receiver node*.
    - » Message in its most general meaning is the object of communication. Depending on the context, the term may apply to both the information contents and its actual presentation, or signal.
    - » The baseband signal usually consist of a finite set of symbols. E.g. text message is composed of words that belong to a finite vocabulary of the language used. Each word in turn is composed by letters of a (finite) alphabet. (Analog-to-digital conversion)
- The transmitter and receiver process the signal using a common *modulation*, *communication protocol* and *communication policy*.

Mapping


3b. Unwanted incident waves & noise
3a. Wanted incident waves
2 Radiated wave (wanted)
C. Receiver mapping
B. Propagation mapping
A. Transmitter mapping
1. Original message
4. Recovered message

The transmitting station:

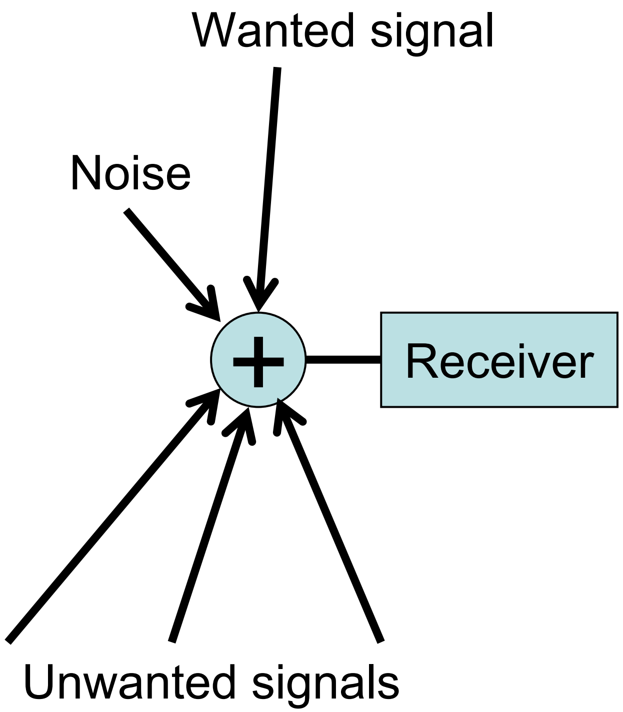
Maps the original message into the radio-wave signal launched at the transmitting antenna Generates a RF carrier

1. Combines it with the baseband signal into a RF signal through *modulation*
2. Performs additional operations
   - » E.g. *analog-to-digital conversion, formatting, coding, spreading,* adding additional messages/ characteristics such as coding, *error-control, authentication*, or *location information*
3. Radiates the resultant signal in the form of a modulated radio wave carrying the message

Propagation process:

- Transforms, or maps, the radio-wave signal launched by the transmitter into the incident radio wave at the receiver antenna
- The propagation mapping depends on extra variables (e.g. distance, latency), additional radio waves (e.g. reflected wave, waves originated in the environment), random uncertainty (e.g. noise, fading)
  and distortions

Receiver:



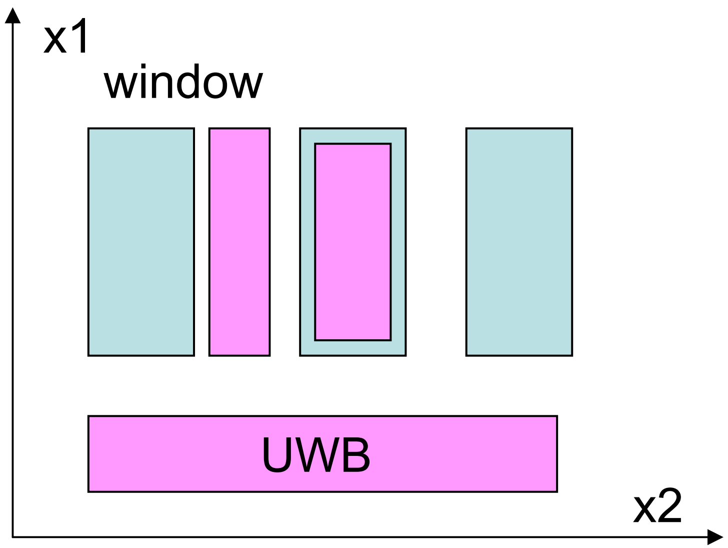
Maps the incident signals into the recovered message

1. Filters the incident signals : rejects unwanted signals and extract the wanted signal
   – The receiver's response defines a solid "window" in the signal hyperspace
2. Recovers the original message through
   - reversing the transmitter operations (demodulation, decoding, de-spreading, etc.),
   - compensating propagation transformations, and
   - correcting transmission distortions

- Multidimensional receiver’s reaction window
  - Accounts for various
    - modulations,
    - signal processing methods,
    - communication protocols,
    - etc.
- Impossible to show on a plane

Receiver reaction window

e.g.
x1 = power, x2 = frequency

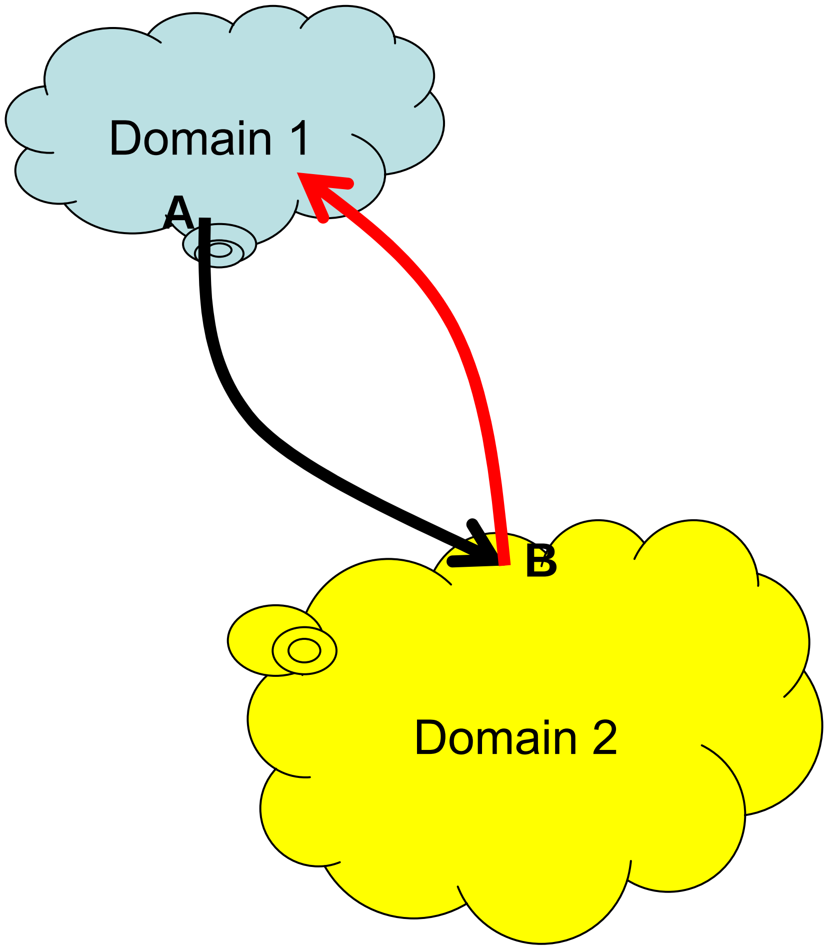


- Receiver ignores signals outside its reaction window
- E.g.: UWB signals with spectral power density below the sensitivity level of system "A" remain 'unnoticed' by "A"

EMC criteria

- The wanted incident signal must match the receiver reaction window
- Any unwanted signal must fall outside that window
  - It must be at a safe distance from it in at least one dimension
- Probability!

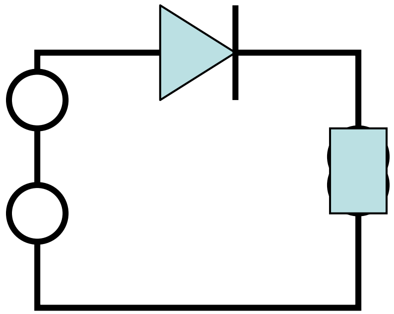
Direct & reverse mapping



- The direct mapping and the reverse-mapping do not result in a unique solution – they may be of type 'one-to-many'

Modulation

- Modulation - process of translation the baseband message signal to radio frequencies; demodulation is the reverse process
- Carrier:
  - continuous (sinusoidal),
  - pulsed (e.g. set of Walsh functions), or
  - random EM waves
- There are many modulation modes: http://en.wikipedia.org/wiki/Category:Radio_modulation_modes
      - There is no time to discuss here all of them!



The simplest mixer

$$i = \zeta(u) = a_0 + \frac{a_1}{1!}u + \frac{a_2}{2!}u^2 + \ldots + \frac{a_n}{n!}u^n + \ldots$$

$$u = A_1 \sin f_1 + A_2 \sin f_2$$

$$u^2 = A_1^2 \sin^2 f_1 + 2A_1 A_2 \sin f_1 \sin f_2 + A_2^2 \sin^2 f_2$$

$$= A_1^2(1-\cos 2f_1) + 2A_1A_2[\cos(f_1 - f_2) - \cos(f_1 + f_2)] + A_2^2(1-\cos 2f_2)$$

$$Components - const: a_0 + \frac{a_2}{2!}\left(A_1^2 + A_2^2\right)$$

$$f_1 : a_1 A_1$$

$$f_2 : a_1 A_2$$

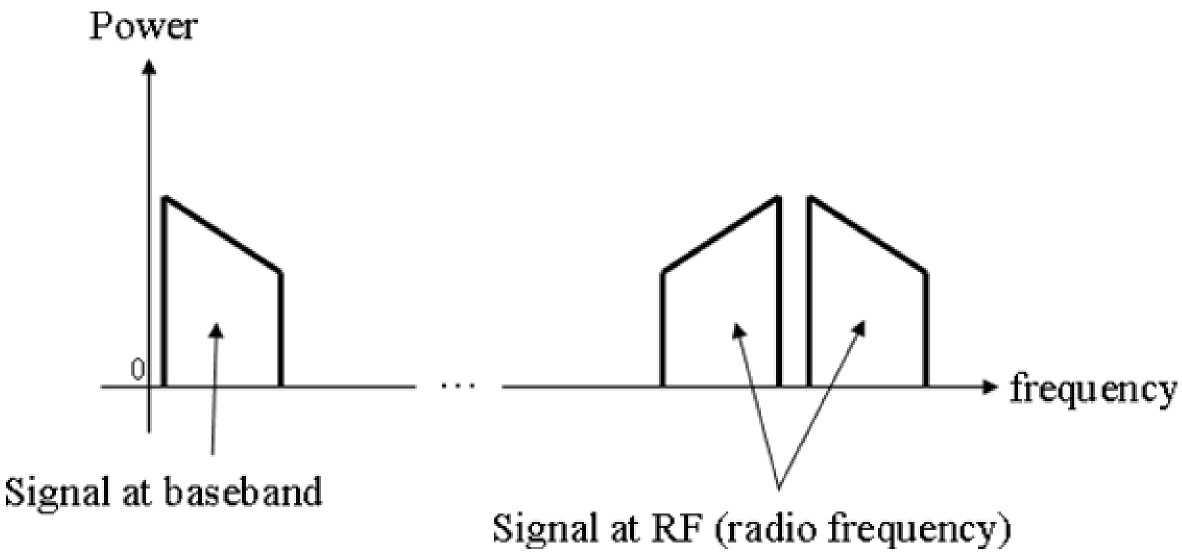
$$f_1 - f_2 : a_2 A_1 A_2$$

$$f_1 + f_2 : a_2 A_1 A_2$$

$$2f_1 : a_2 A_1^2$$

$$2f_2 : a_2 A_2^2$$

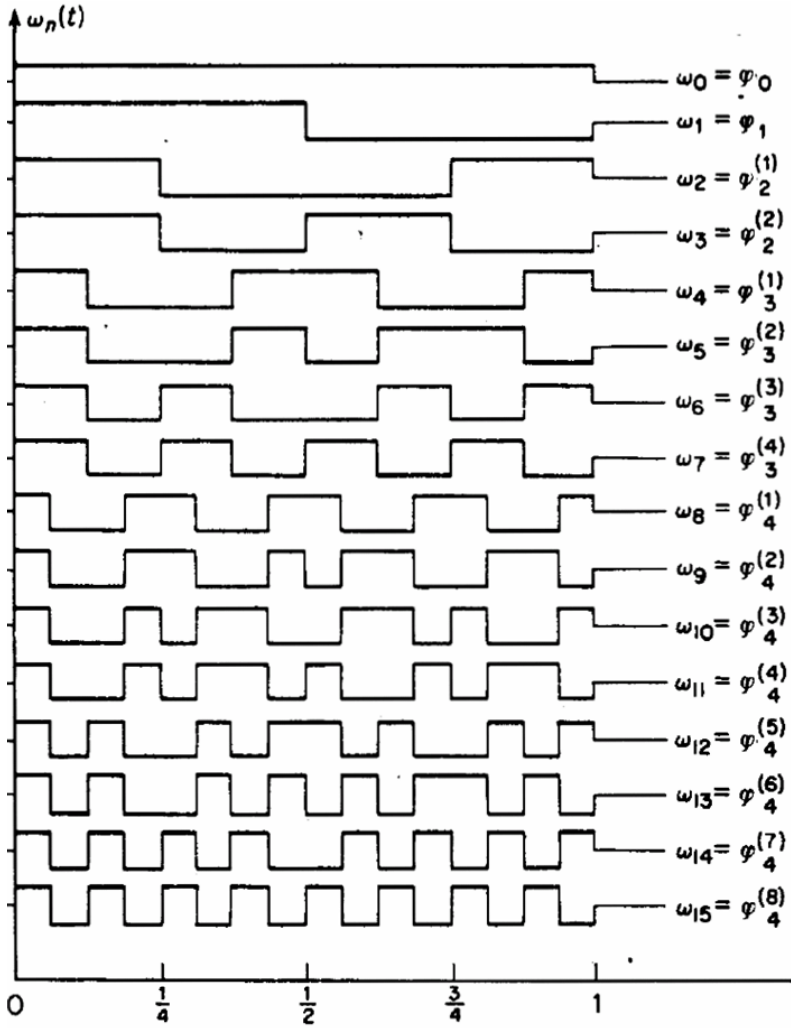
Frequency translation



Signal's baseband bandwidth is its bandwidth before modulation and multiplexing, or after demultiplexing and demodulation

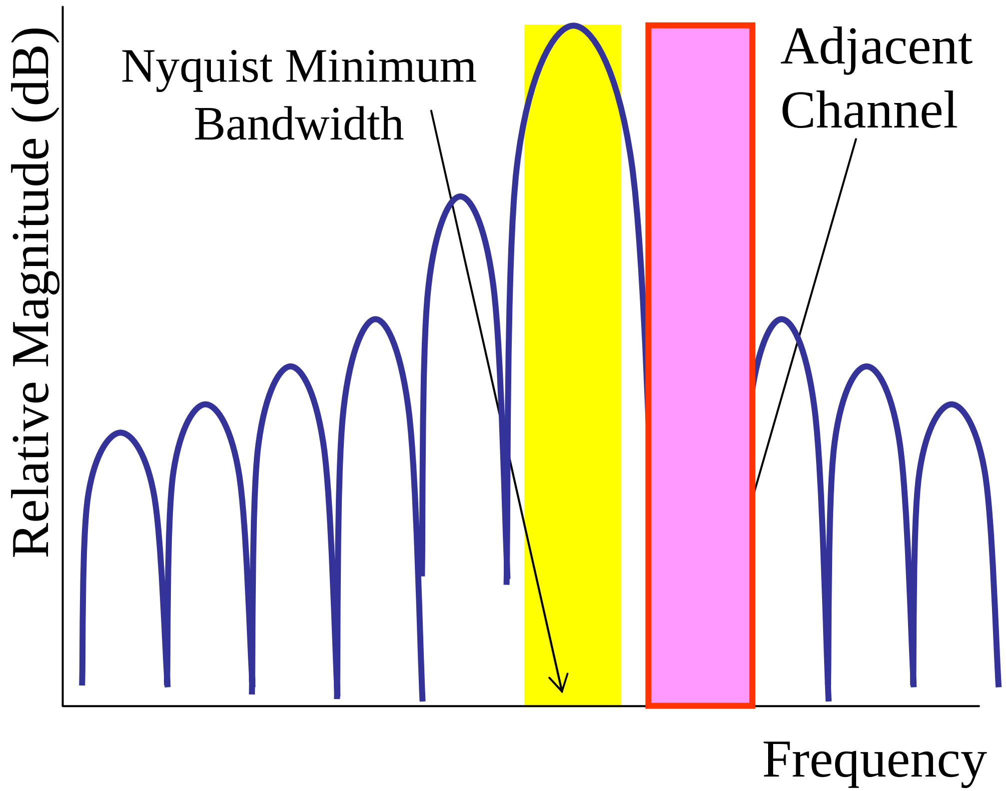
- Signal "at baseband" comprises all relevant frequency components carrying information. Modulation shifts the signal up to RF frequencies to allow for radio transmission. Usually, the process increases the signal bandwidth. Steps are often taken to reduce this effect, such as filtering the RF signal prior to transmission. http://en.wikipedia.org/wiki/Baseband

Rectangular carrier



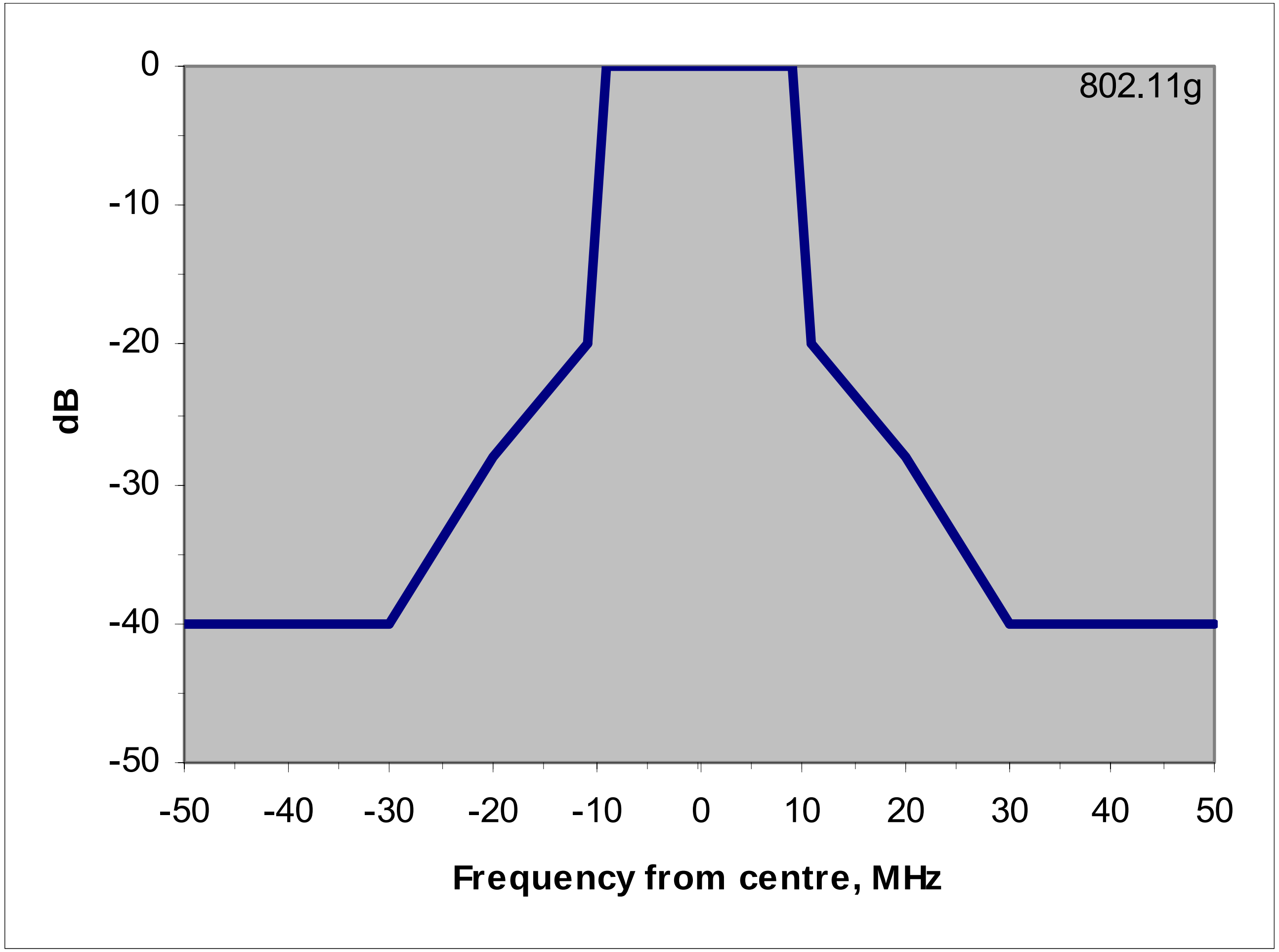
- Walsh functions
  - http://mathworld.wolfram.com/WalshFunction.html

Modulation Spectra



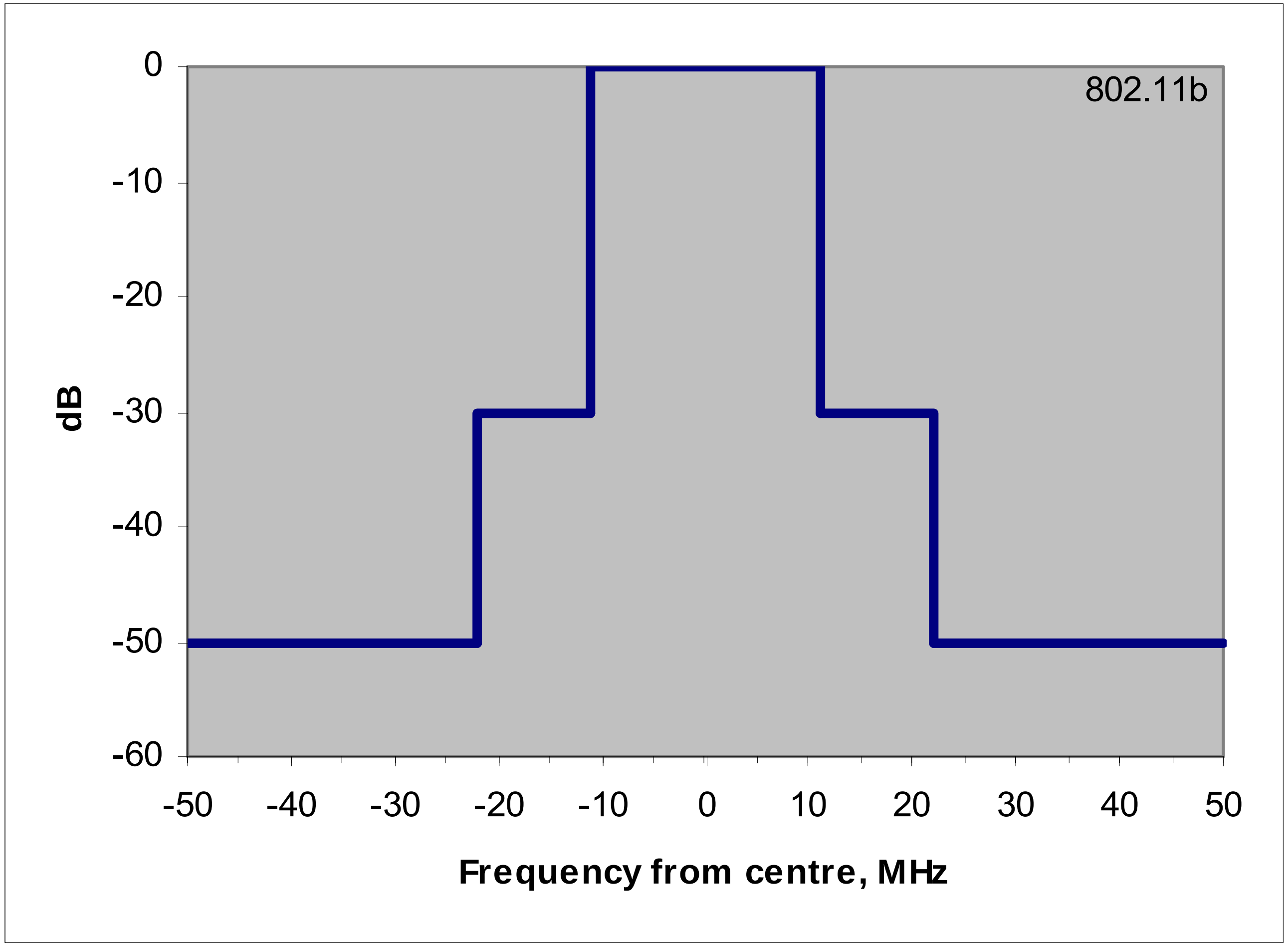
- The Nyquist bandwidth is the minimum bandwidth that can represent a signal (within an acceptable error)
- The spectrum occupied by a signal should be as close as practicable to that minimum, otherwise adjacent channel interference occur
- The spectrum occupied by a signal can be reduced by application of filters

802.11g spectrum mask



Source: R Morrow: Wireless network coexistence, p. 201

802.11b spectrum mask



Source: R Morrow: Wireless network coexistence, p. 201

802.11b/g channels

(Center frequencies in GHz)

1.	***2,412***
2.	2,417
3.	2,422
4.	2,427
5.	2,432
6.	***2,437***
7.	2,442
8.	2,447
9.	2,452
10.	2,457
11.	***2,462***
12.	2,467
13.	2,472
14.	2,484

- Different subsets of these channels are made available in various countries
- Spacing: ~5 (12) MHz
- Occupied bandwidth: ~22 MHz (802.11b), ~16.6 MHz (802.11g)

Source: R Morrow: Wireless network coexistence, p. 198

Why Carrier?

- Effective radiation of EM waves requires antenna dimensions to be comparable with the wavelength:
  - Antenna for 3 kHz would be ~100 km long
  - Antenna for 3 GHz is 10 cm long
- Sharing the access to the telecommunication channel resources

Modulation Process

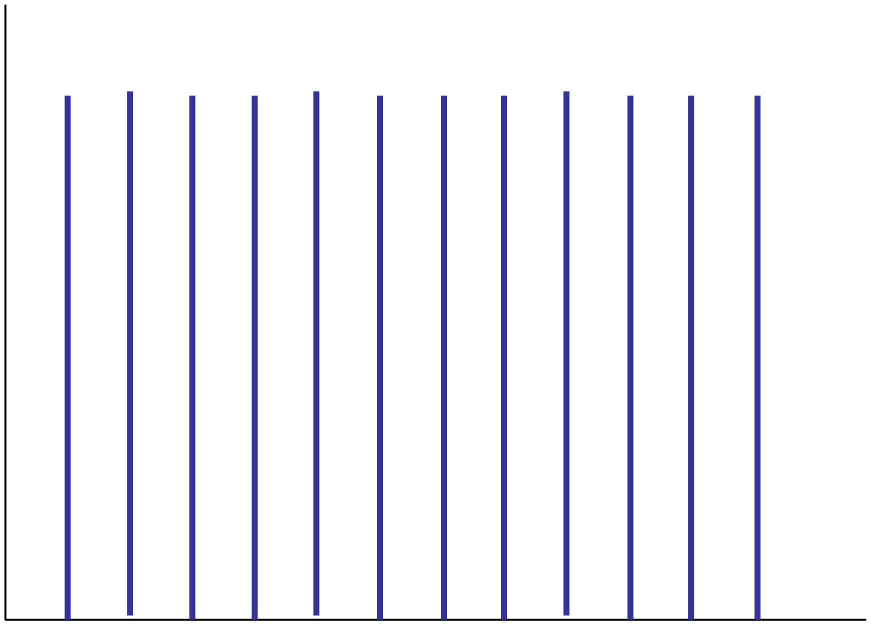
$$f = f\left(a_1, a_2, a_3, \ldots a_n, t\right) \text{ (= carrier)}$$

$$a_1, a_2, a_3, \ldots a_n \text{ (= modulation parameters)}$$

$$t \text{ (= time)}$$

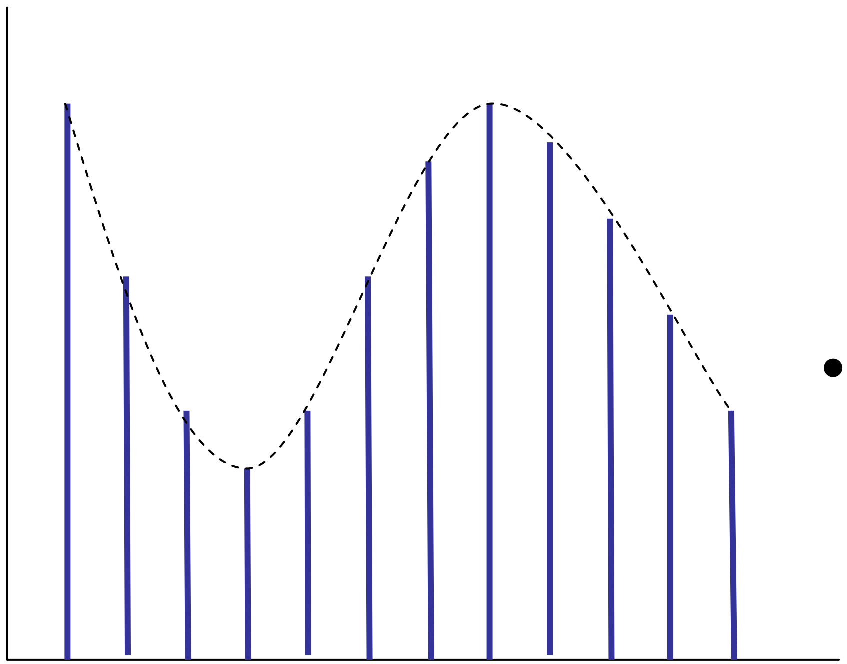
- Modulation implies varying one or more characteristics (modulation parameters a_1, a_2, … a_n) of a carrier *f* in accordance with the information-bearing (modulating) baseband signal.

Pulse Carrier



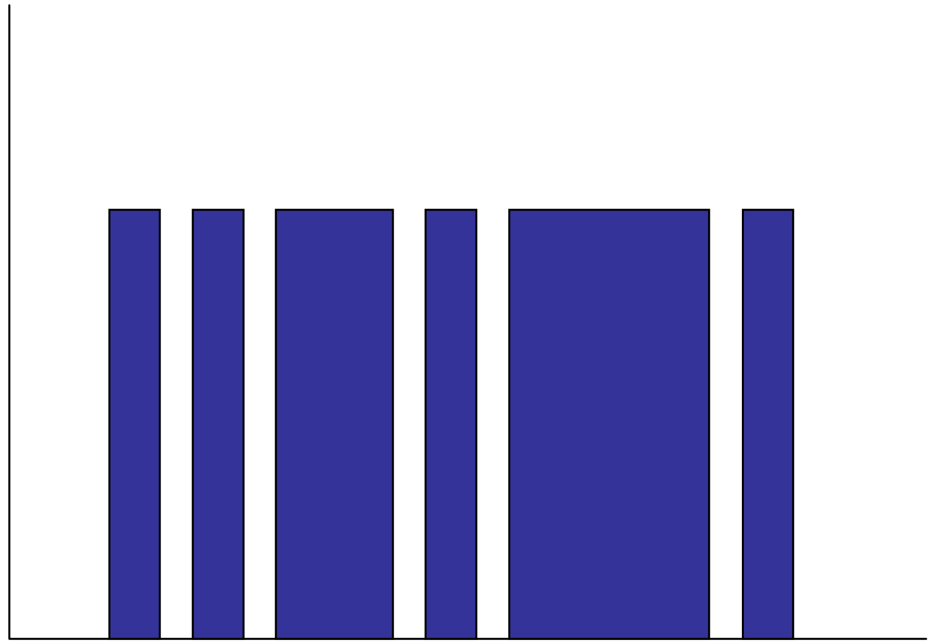
- Carrier:
  A train of identical pulses regularly spaced in time

Pulse-Amplitude Modulation (PAM)



- Modulation in which the amplitude of pulses is varied in accordance with the modulating signal.
- Used e.g. in telephone switching equipment such as a private branch exchange (PBX)

Pulse-Duration Modulation (PDM)

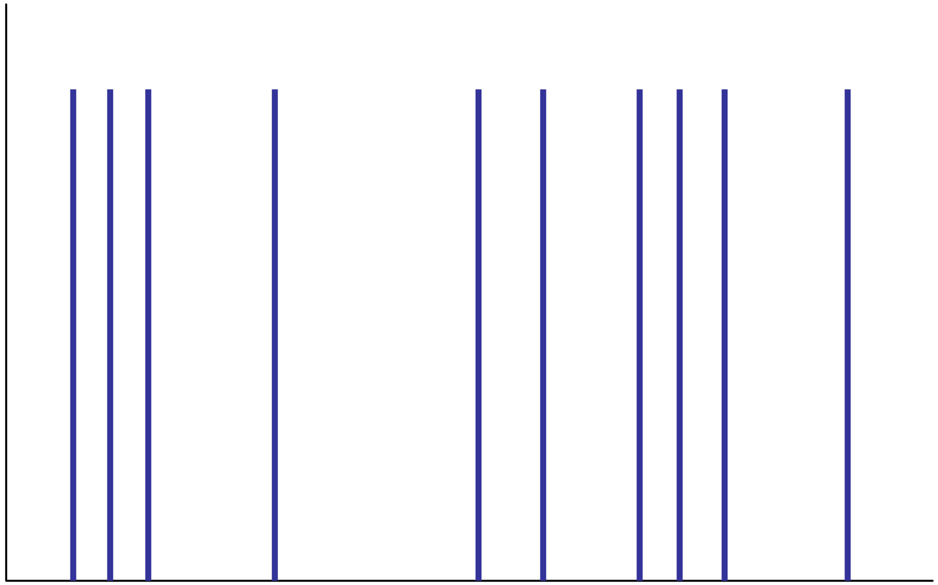


Used e.g. in telephone switching equipment such as a private branch exchange (PBX)

Modulation in which the duration of pulses is varied in accordance with the modulating signal.

Deprecated synonyms: pulse-length modulation, pulse-width modulation.

Pulse-Position Modulation (PPM)



- Modulation in which the temporal positions of the pulses are varied in accordance with some characteristic of the modulating signal.

Ultra-Wideband (UWB) Systems

- Radio or wireless devices where the occupied bandwidth is greater than 25% of the center frequency or greater than 1.5 GHz.
- Radio or wireless systems that use narrow pulses (on the order of 1 to 10 nanoseconds), also called carrierless or impulse systems, for communications and sensing (short-range radar).
- Radio or wireless systems that use time-domain modulation methods (*e.g.,* pulse-position modulation) for communications applications, or time-domain processing for sensing applications.

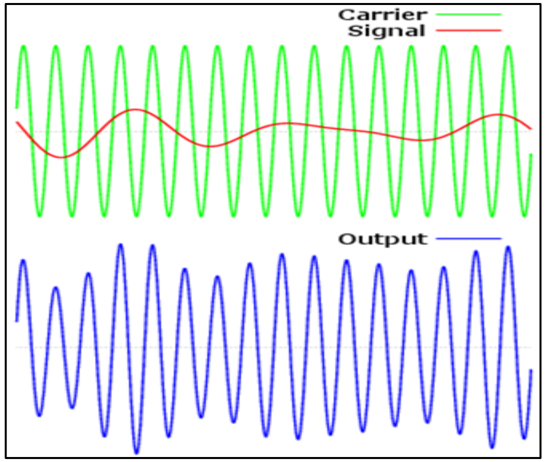
Continuous (sinusoidal) carrier

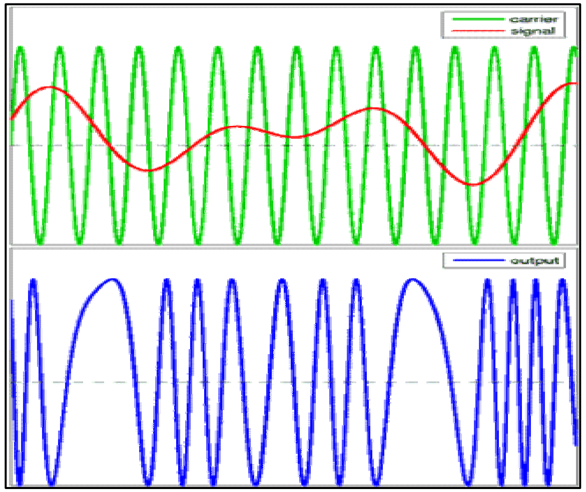
Carrier: $A \sin[\omega t + \varphi]$

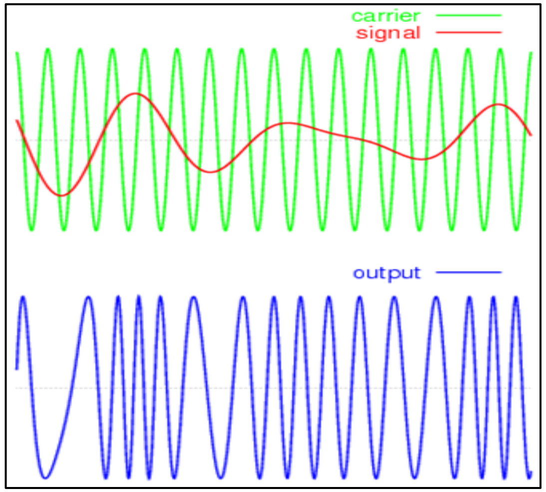
- A = const
- ω = const
- φ = const

- Amplitude modulation (AM)
  - A = A(t) – carries information
  - ω = const
  - φ = const
- Frequency modulation (FM)
  - A = const
  - ω = ω(t)– carries information
  - φ = const
- Phase modulation (PM)
  - A = const
  - ω = const
  - φ = φ(t) – carries information

AM, FM, PM





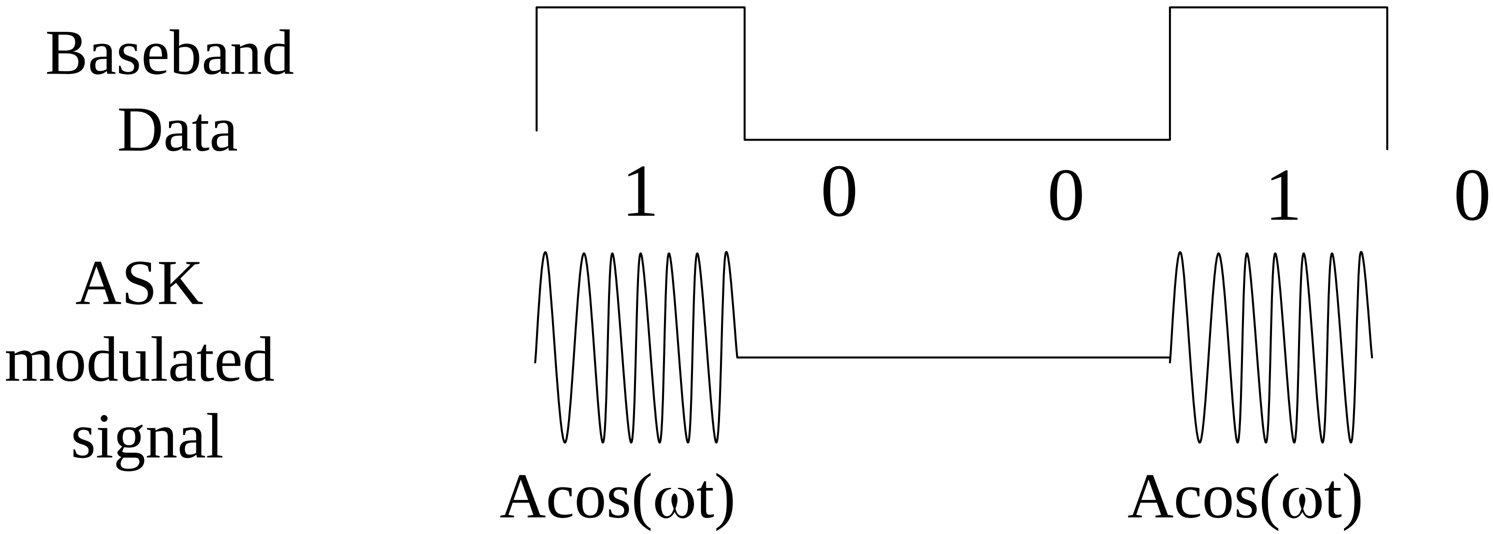


The modulating signal superimposed on the carrier wave & the resulting modulated signal.
The spectrum
[wikipedia]

Digital modulation

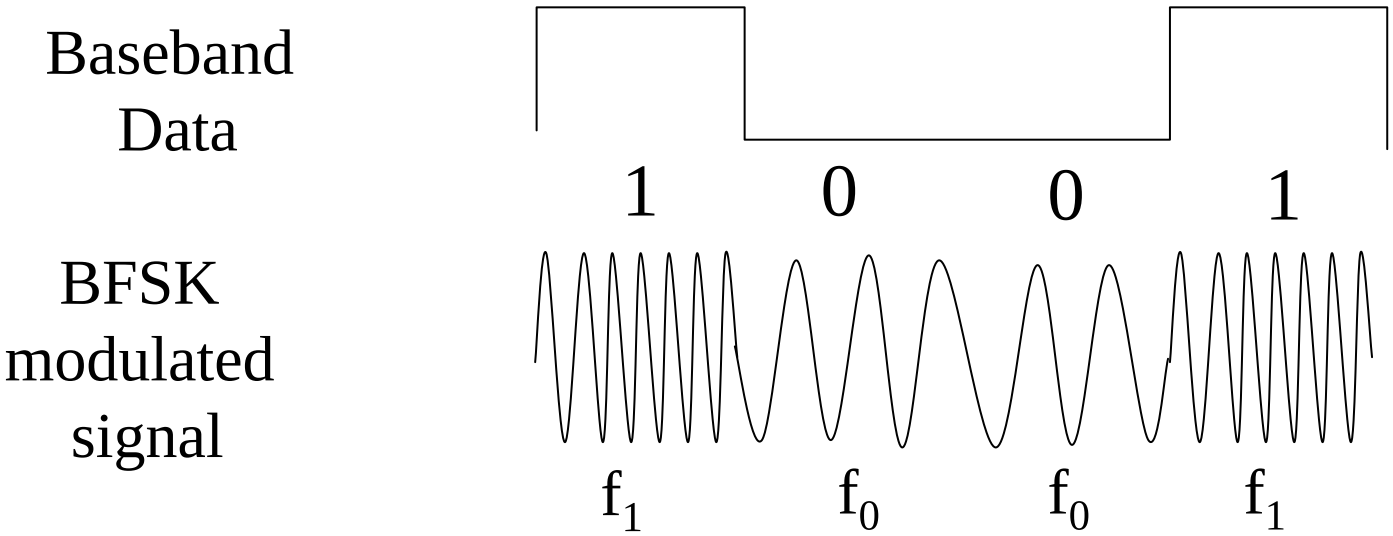
- Any form of digital modulation necessarily uses a finite number of distinct signals to represent digital data.
- In the case of PSK, a finite number of phases are used.
- In the case of FSK, a finite number of frequencies are used.
- In the case of ASK, a finite number of amplitudes are used.

Amplitude Shift Keying (ASK)



- Pulse shaping can be employed to remove spectral spreading
- ASK demonstrates poor performance, as it is heavily affected by noise, fading, and interference

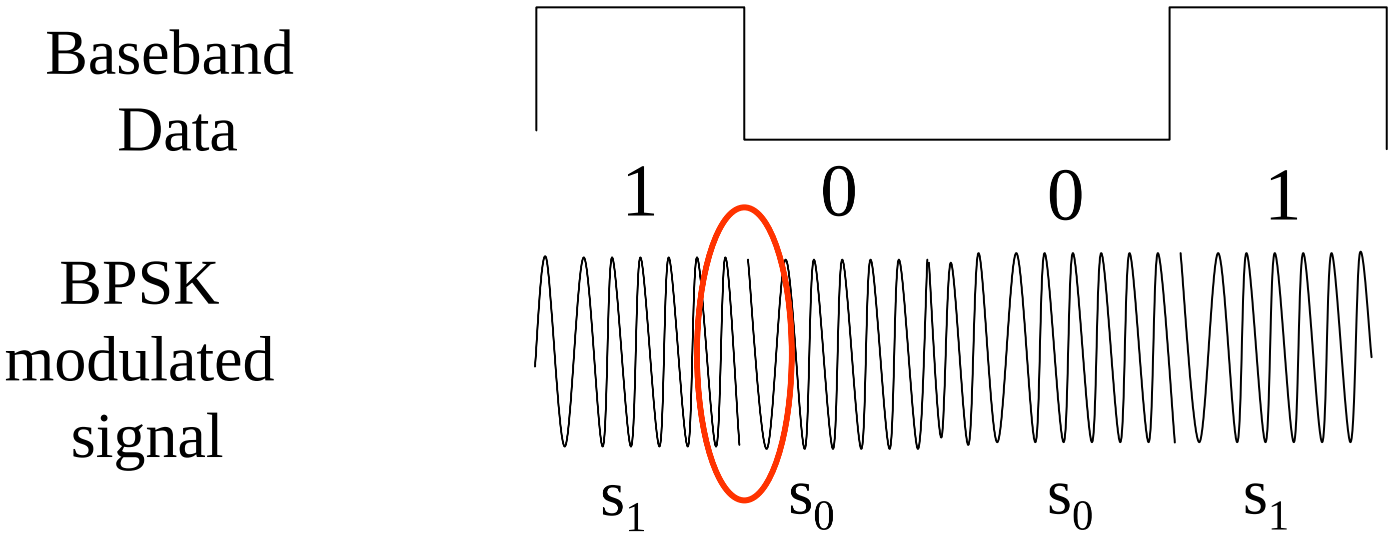
Frequency Shift Keying (FSK)



where $f_0 = A\cos(\omega_c - \Delta\omega)t$ and $f_1 = A\cos(\omega_c + \Delta\omega)t$

- Example: The ITU-T V.21 modem standard uses FSK
- FSK can be expanded to a M-ary scheme, employing multiple frequencies as different states

Phase Shift Keying (PSK)



where $s_0 = -A\cos(\omega_c t)$ and $s_1 = A\cos(\omega_c t)$

- Major drawback – rapid amplitude change between symbols due to phase discontinuity, which requires infinite bandwidth. Binary Phase Shift Keying (BPSK) demonstrates better performance than ASK and BFSK
- BPSK can be expanded to a M-ary scheme, employing multiple phases and amplitudes as different states

Differential Modulation

- In the transmitter, each symbol is modulated relative to the previous symbol and modulating signal, for instance in BPSK 0 = no change, 1 = $+180^0$
- In the receiver, the current symbol is demodulated using the previous symbol as a reference. The previous symbol serves as an estimate of the channel. A no-change condition causes the modulated signal to remain at the same 0 or 1 state of the previous symbol.

DPSK

- Differential modulation is theoretically 3dB poorer than coherent. This is because the differential system has 2 sources of error: a corrupted symbol, and a corrupted reference (the previous symbol)
- DPSK = Differential phase-shift keying: In the transmitter, each symbol is modulated relative to (a) the phase of the immediately preceding signal element and (b) the data being transmitted.

Demodulation & Detection

- Demodulation
  - Is process of removing the carrier signal to obtain the original signal waveform
- Detection – extracts the symbols from the waveform
  - Coherent detection
  - Non-coherent detection

Coherent Detection

- An estimate of the channel phase and attenuation is recovered. It is then possible to reproduce the transmitted signal and demodulate.
- Requires a replica carrier wave of the same frequency and phase at the receiver.
- The received signal and replica carrier are cross-correlated using information contained in their amplitudes and phases.
- Also known as synchronous detection

Coherent Detection 2

- Carrier recovery methods include
  - Pilot Tone (such as Transparent Tone in Band)
    - Less power in the information bearing signal, High peak-to-mean power ratio
  - Carrier recovery from the information signal
    - E.g. Costas loop
- Applicable to
  - Phase Shift Keying (PSK)
  - Frequency Shift Keying (FSK)
  - Amplitude Shift Keying (ASK)

Non-Coherent Detection

- Requires no reference wave; does not exploit phase reference information (envelope detection)
  - Differential Phase Shift Keying (DPSK)
  - Frequency Shift Keying (FSK)
  - Amplitude Shift Keying (ASK)
  - Non coherent detection is less complex than coherent detection (easier to implement), but has worse performance.

Geometric Representation

- Digital modulation involves choosing a particular signal $s_i(t)$ form a finite set S of possible signals.
- For binary modulation schemes a binary information bit is mapped directly to a signal and S contains only 2 signals, representing 0 and 1.
- For M-ary keying S contains more than 2 signals and each represents more than a single bit of information. With a signal set of size M, it is possible to transmit up to log_2M bits per signal.

Geometric Representation 2

- Any element of set S can be represented as a point in a vector space whose coordinates are basis signals $\phi_j(t)$ such that

$$\int_{-\infty}^{\infty} \phi_i(t)\phi_j(t)\,dt = 0, i \neq j; \text{ (= are orthogonal)}$$

$$E = \int_{-\infty}^{\infty} \left[\phi_i(t)\right]^2 dt = 1; \text{ (= normalization)}$$

Then $\quad s_i(t) = \sum_{j=1}^{N} s_{ij}\phi_j(t)$

Example: BPSK (Binary Phase Shift Keying) Constellation Diagram

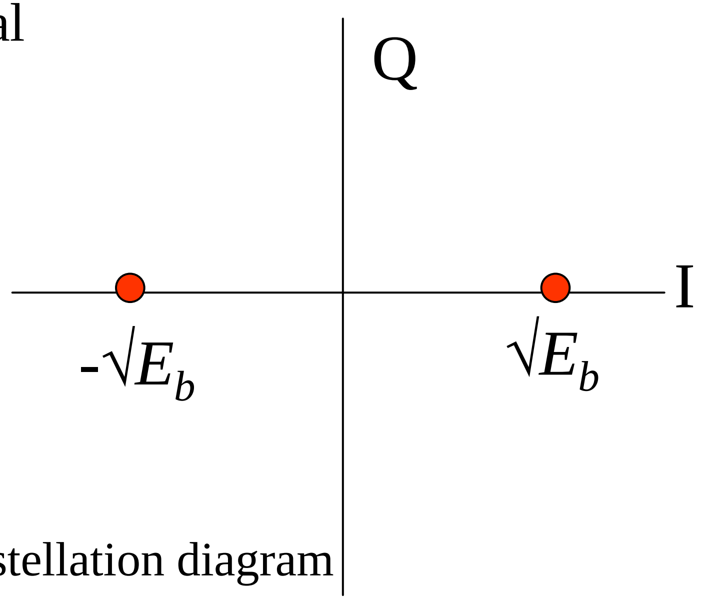
$$S_{BPSK} = \left\{ \left[s_1(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t) \right], \left[s_2(t) = -\sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t); \right] \right\}; \quad 0 \le t \le T_b$$

$E_b =$ energy per bit; $T_b =$ bit period

For this signal set, there is a single basic signal

$$\phi_1(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_c t); \quad 0 \le t \le T_b$$

$$S_{BPSK} = \left\{ \left[\sqrt{E_b}\,\phi_1(t) \right], \left[-\sqrt{E_b}\,\phi_1(t) \right] \right\}$$



Constellation diagram

Constellation diagram

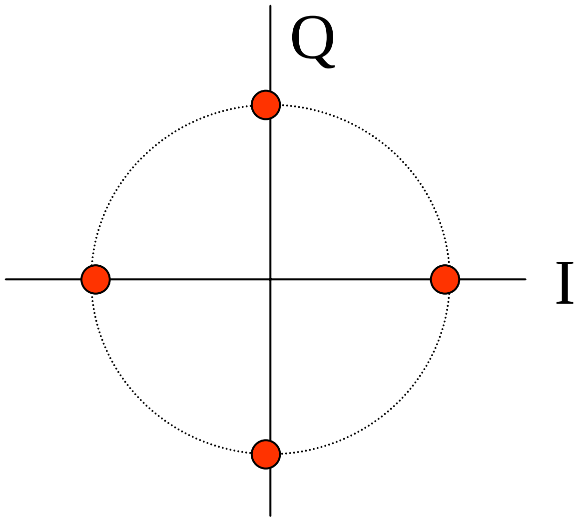
= graphical representation of the complex envelope of each possible symbol state

- The x-axis represents the in-phase component and the y-axis the quadrature component of the complex envelope
- The distance between signals on a constellation diagram relates to how different the modulation waveforms are and how easily a receiver can differentiate between them.

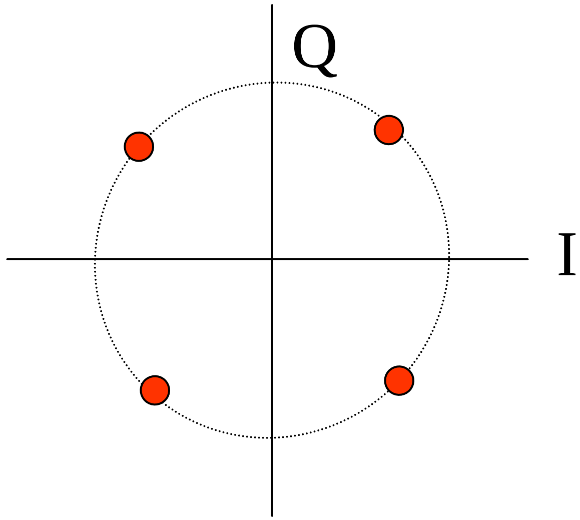
QPSK

- Quadrature Phase Shift Keying (QPSK) can be interpreted as two independent BPSK systems [one on the I-channel (in-phase) and one on Q (quadraure phase)], and thus the same performance but twice the bandwidth efficiency
- Large envelope variations occur due to abrupt phase transitions, thus requiring linear amplification

QPSK Constellation Diagram



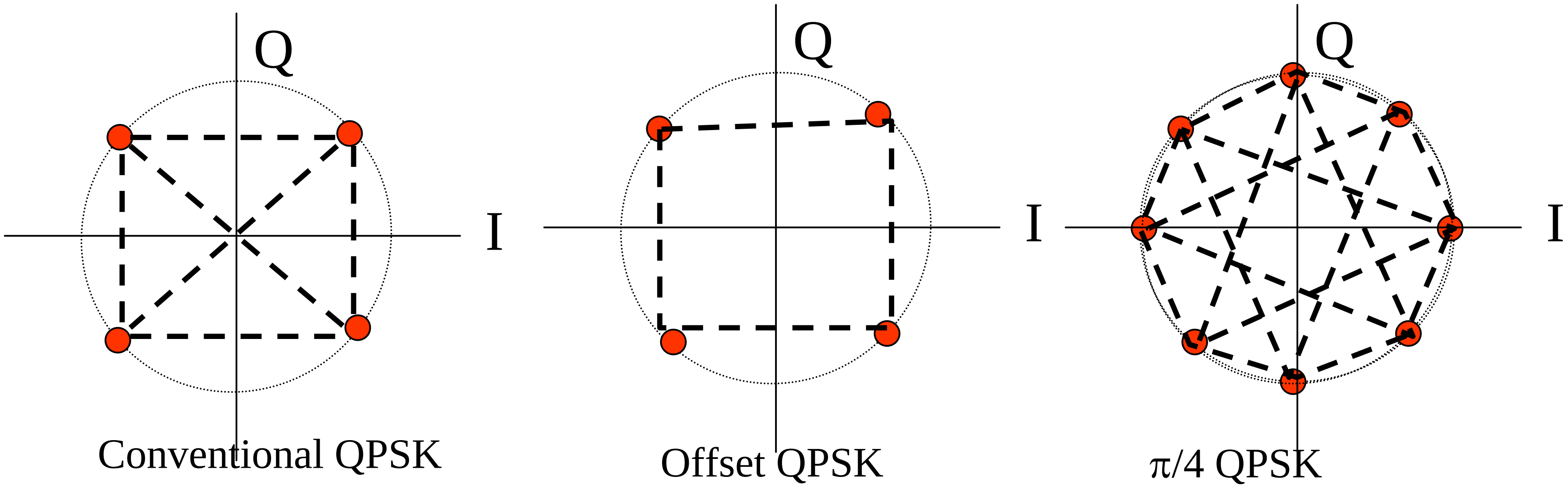
Carrier phases
$\{0, \pi/2, \pi, 3\pi/2\}$



Carrier phases
$\{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}$

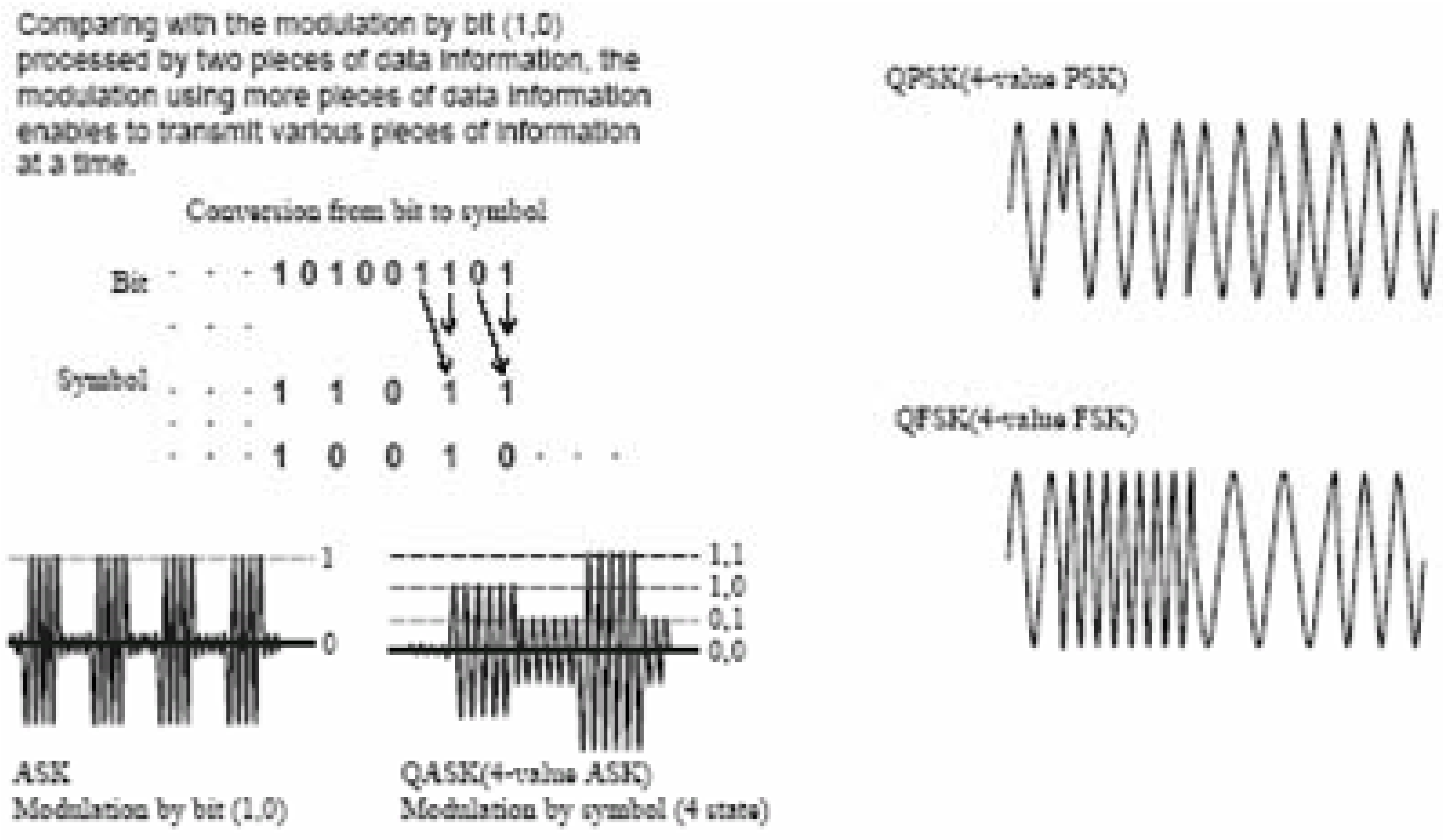
- Quadrature Phase Shift Keying has twice the bandwidth efficiency of BPSK since 2 bits are transmitted in a single modulation symbol

Types of QPSK

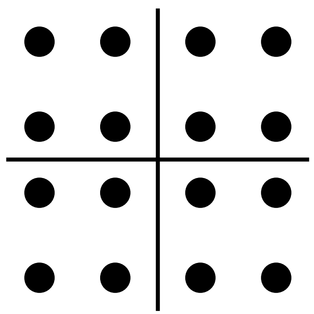


- Conventional QPSK has transitions through zero (i.e. 180^0 phase transition). Highly linear amplifiers required.
- In Offset QPSK, the phase transitions are limited to 90^0, the transitions on the I and Q channels are staggered.
- In π/4 QPSK the set of constellation points are toggled each symbol, so transitions through zero cannot occur. This scheme produces the lowest envelope variations.
- All QPSK schemes require linear power amplifiers

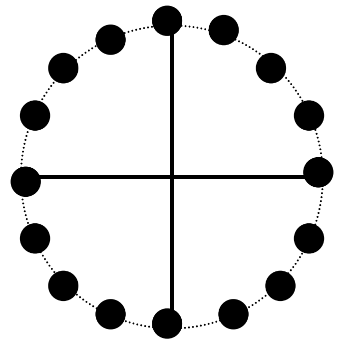
Multi-level digital modulation



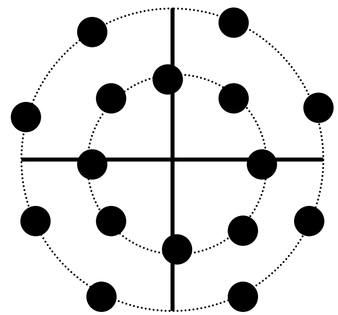
Multi-level (M-ary) Phase and Amplitude Modulation



16 QAM



16 PSK

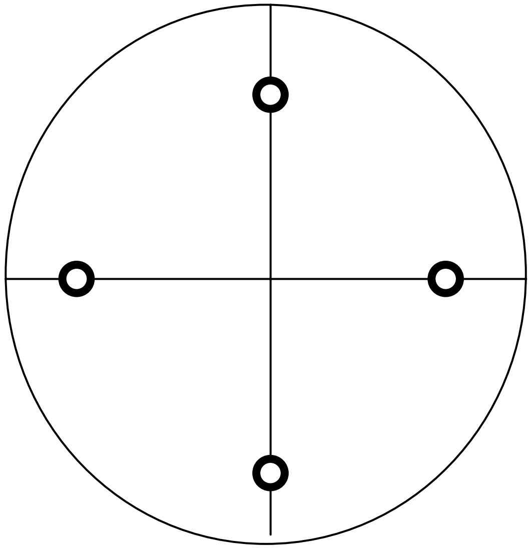


16 APSK

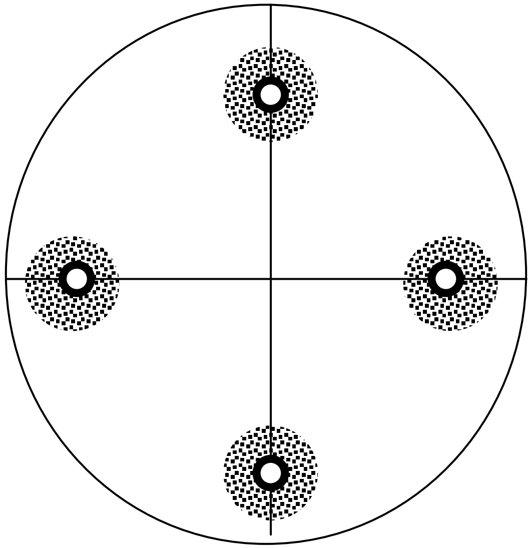
- Amplitude and phase shift keying can be combined to transmit several bits per symbol.
  - Often referred to as *linear* as they require linear amplification.
  - More bandwidth-efficient, but more susceptible to noise.
- For M=4, 16QAM has the largest distance between points, but requires very linear amplification. 16PSK has less stringent linearity requirements, but has less spacing between constellation points, and is therefore more affected by noise.

- Java animations: http://www.educatorscorner.com/index.cgi?CONTENT_ID=2478

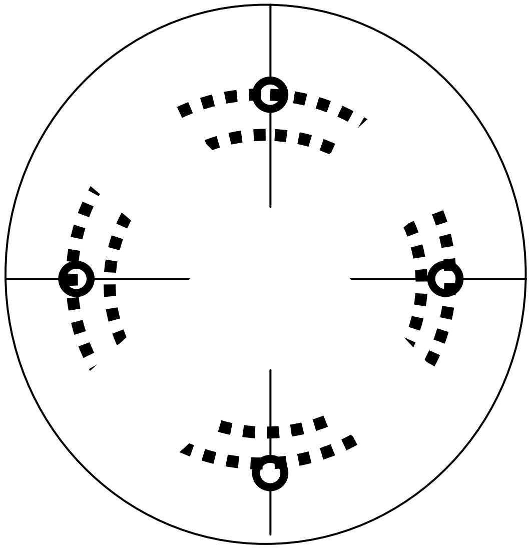
Distortions



Perfect channel

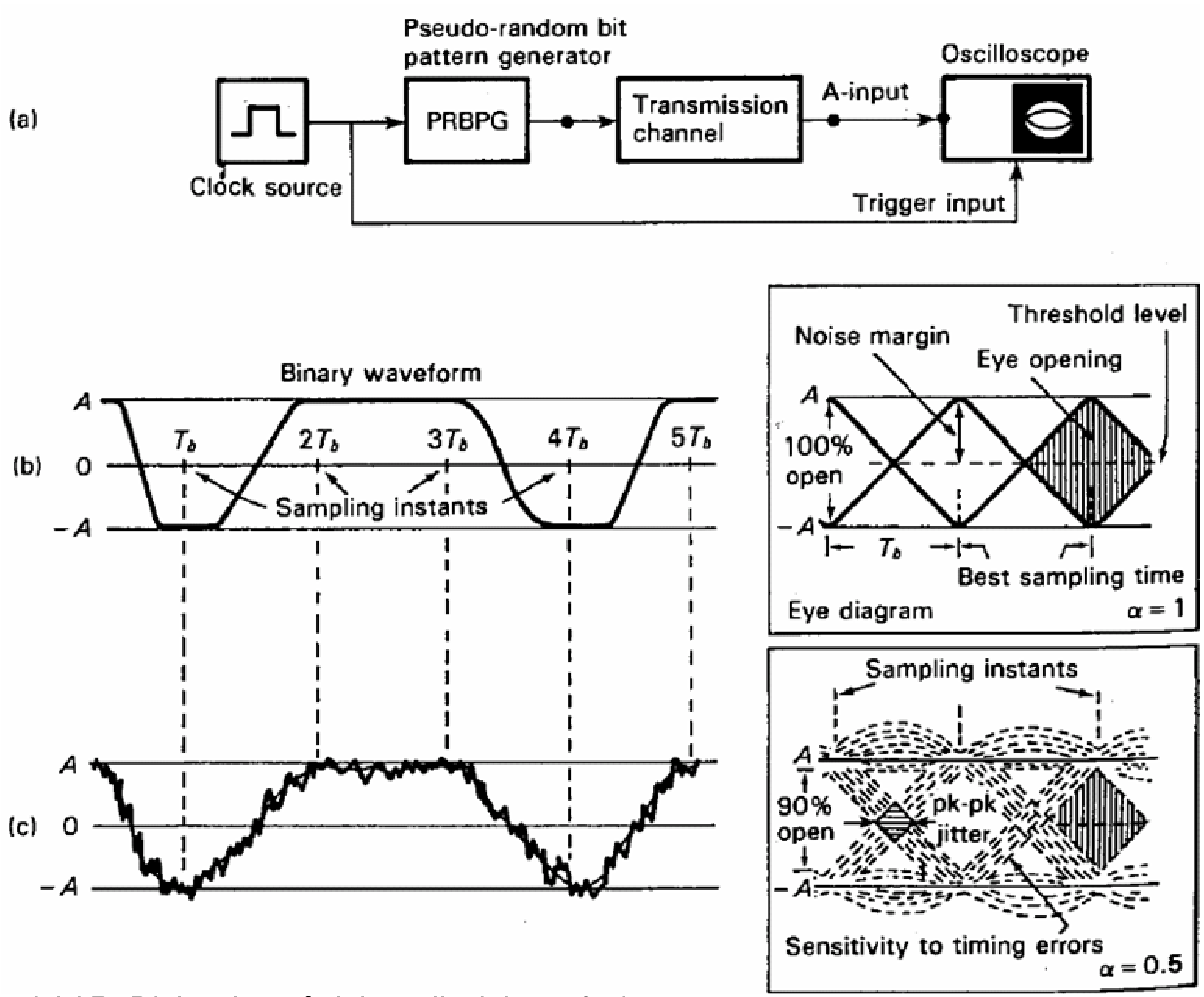


White noise



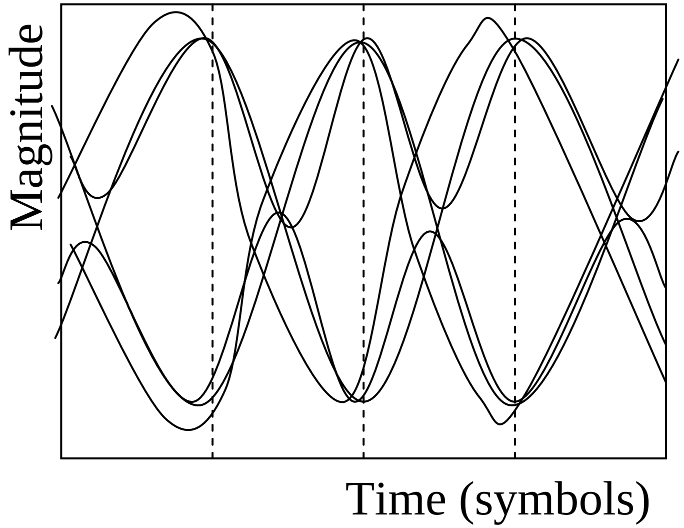
Phase jitter

‘Eye’ diagram



Townsend AAR: Digital line-of-sight radio links p. 274

Eye Diagram



- Eye pattern is an oscilloscope display in which digital data signal from a receiver is repetitively superimposed on itself many times (sampled and applied to the vertical input, while the data rate is used to trigger the horizontal sweep).
- It is so called because the pattern looks like a series of eyes between a pair of rails.
- If the “eye” is not open at the sample point, errors will occur due to signal corruption.

GMSK

- Gaussian Minimum Shift Keying (GMSK) is a form of continuous-phase FSK in which the phase change is changed between symbols to provide a constant envelope. Consequently it is a popular alternative to QPSK
- The RF bandwidth is controlled by the Gaussian low-pass filter bandwidth. The degree of filtering is expressed by multiplying the filter 3dB bandwidth (B) by the bit period of the transmission (T), i.e. by BT
- GMSK allows efficient class C non-linear amplifiers to be used

Bandwidth Efficiency

$$\frac{f_b}{W} = \log_2\left(1 + \frac{E_b f_b}{\eta W}\right)$$

$f_b =$ capacity (bits per second)

$W =$ bandwidth of the modulating baseband signal (Hz)

$E_b =$ energy per bit

$\eta =$ noise power density (watts/Hz)

Thus

$E_b f_b =$ total signal power

$\eta W =$ total noise power

$\frac{f_b}{W} =$ bandwidth use efficiency

= bits per second per Hz

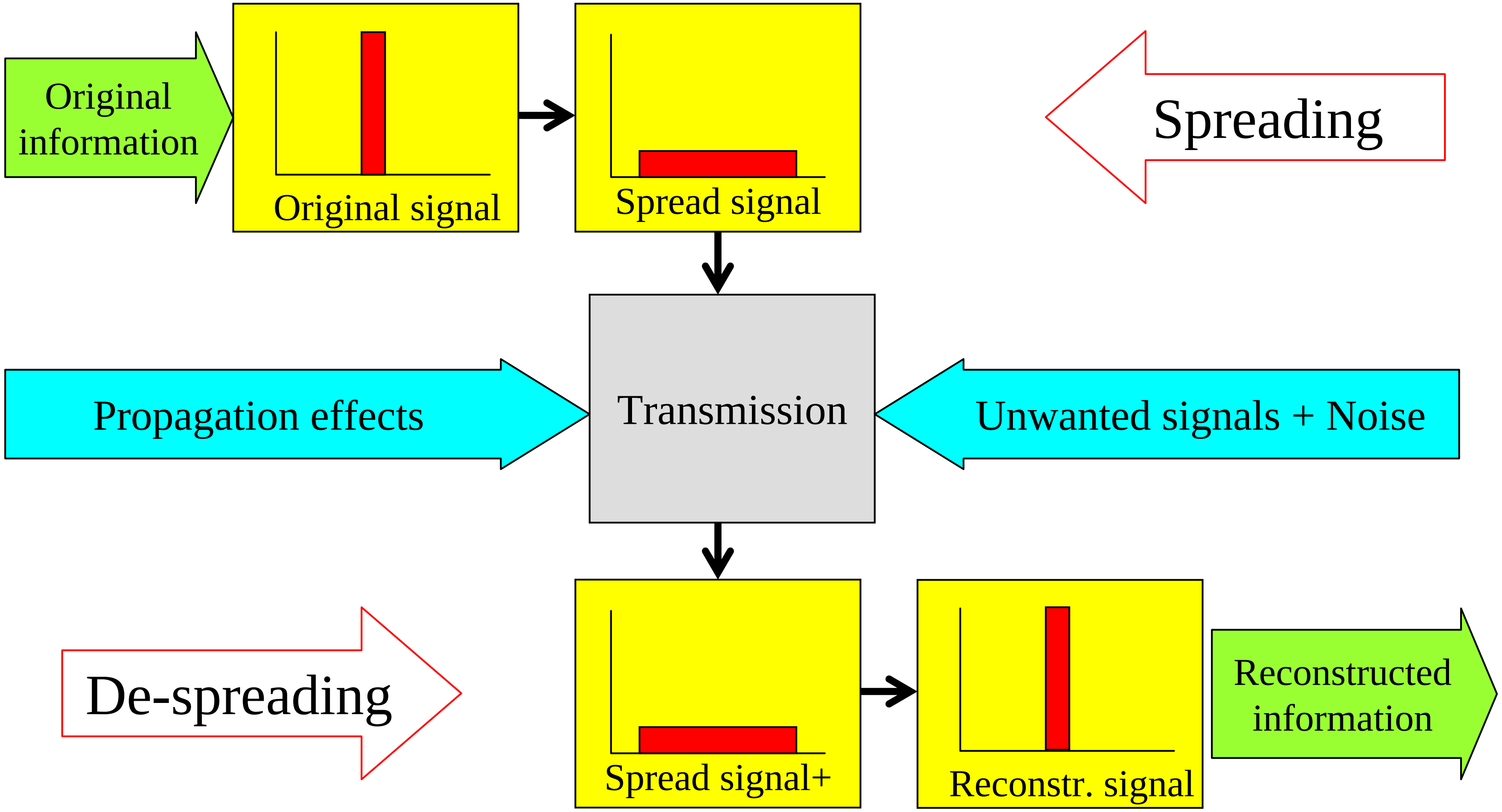
Comparison of Modulation Types

Modulation Format	Bandwidth efficiency C/B	Log2(C/B)	Error-free Eb/No
16 PSK	4	2	18dB
16 QAM	4	2	15dB
8 PSK	3	1.6	14.5dB
4 PSK	2	1	10dB
4 QAM	2	1	10dB
BFSK	1	0	13dB
BPSK	1	0	10.5dB

Modulation Summary

- Phase Shift Keying (PSK) is often used as it provides efficient use of RF spectrum. $\pi/4$ QPSK (Quadrature PSK) reduces the envelope variation of the signal.
- High level M-array schemes (such as 64-QAM) are very bandwidth-efficient but more susceptible to noise and require linear amplification
- Constant envelope schemes (such as GMSK) allow for non-linear power-efficient amplifiers
- Coherent reception provides better performance but requires a more complex receiver

SS communications basics



Capacity of communication system

$$C = B*\log_2\{1 + [S/(N_o*B)]\}$$

Noise density, W/Hz

Received signal power, W

Bandwidth, Hz

Capacity, bit/s

The capacity to transfer error-free information is enhanced with increased bandwidth B, even though the signal-to-noise ratio is decreased because of the increased bandwidth.

SS: basic characteristics

- Signal spread over a wide bandwidth >> minimum bandwidth necessary to transmit information
- Spreading by means of a code independent of the data
- Data recovered by de-spreading the signal with a synchronous replica of the reference code
  - TR: transmitted reference (separate data-channel and reference-channel, correlation detector)
  - SR: stored reference (independent generation at T & R pseudo-random identical waveforms, synchronization by signal received, correlation detector)
  - Other (MT: T-signal generated by pulsing a matched filter having long, pseudo-randomly controlled impulse response. Signal detection at R by identical filter & correlation computation)

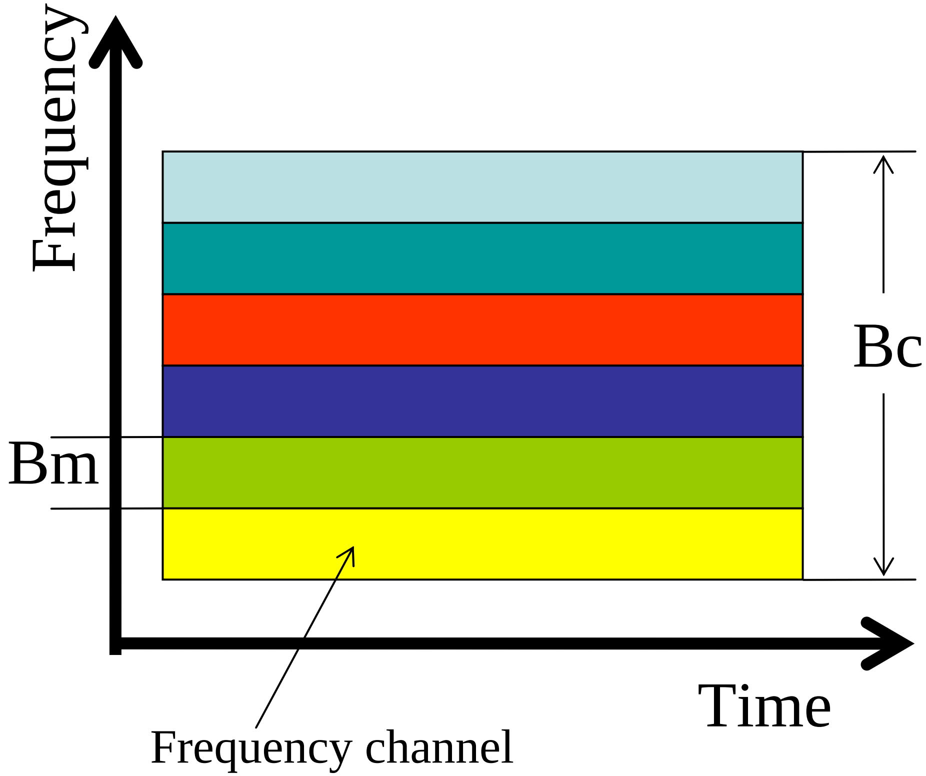
SS communication techniques

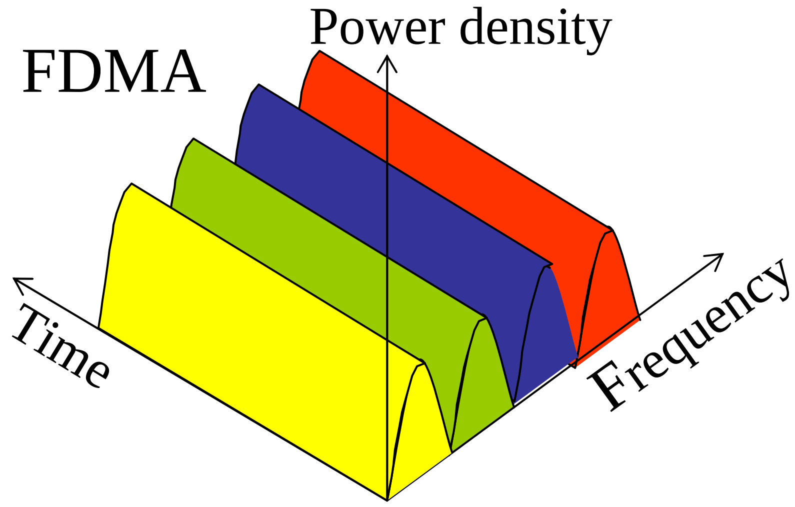
- FH: frequency hoping (frequency synthesizer controlled by pseudo-random sequence of numbers)
- DS: direct sequence (pseudo-random sequence of pulses used for spreading)
- TH: time hoping (spreading achieved by randomly spacing transmitted pulses)
- Random noise as carrier
- Combination of the above
- Other techniques (radar and other applications)

Multiple-access techniques

- TDMA: time-division multiple access
- FDMA: frequency-division multiple access
- CDMA: code-division multiple access
- OFDM: orthogonal frequency multiple access

FDMA

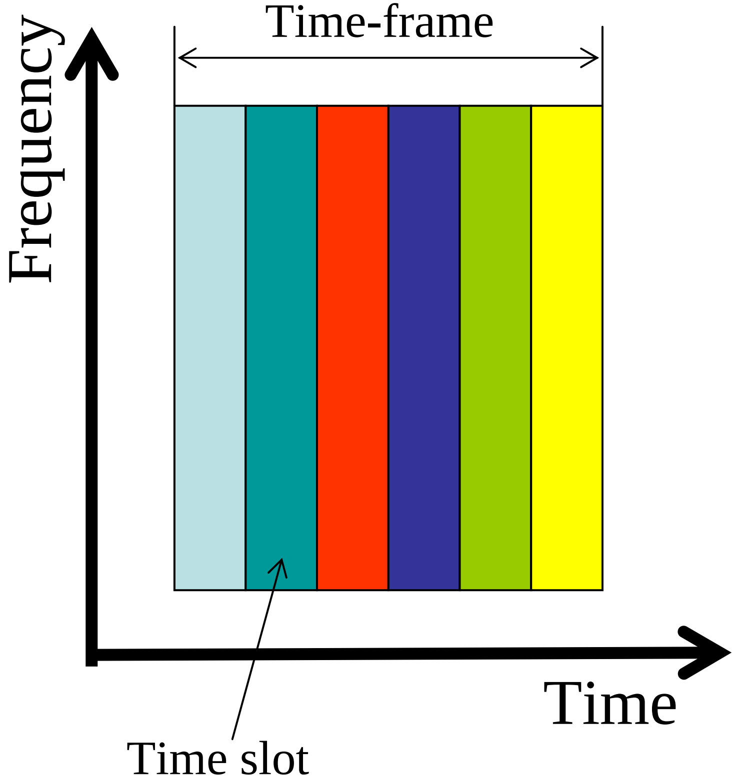


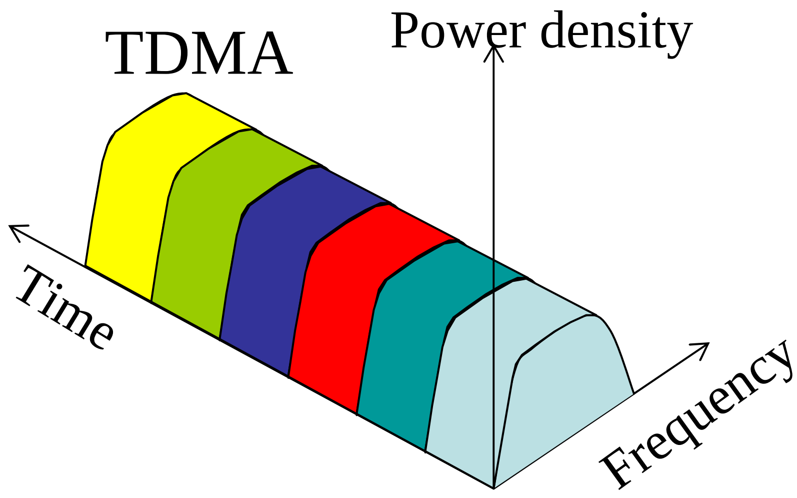


Transmission is organized in frequency channels. Each link is assigned a separate channel.

Example: Telephony Bm = 3-9 kHz

TDMA

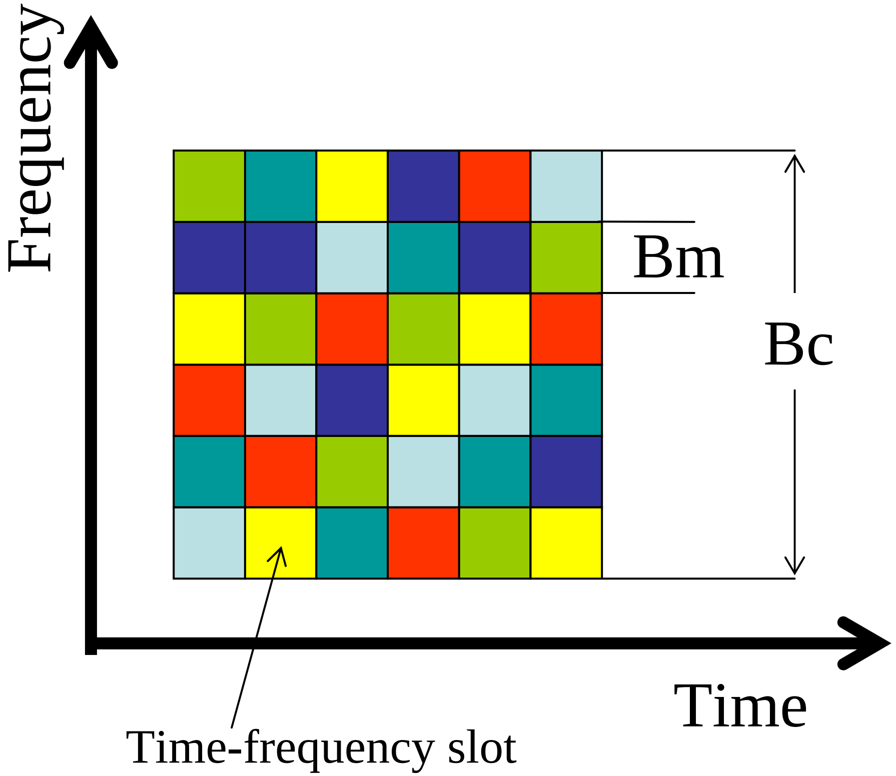


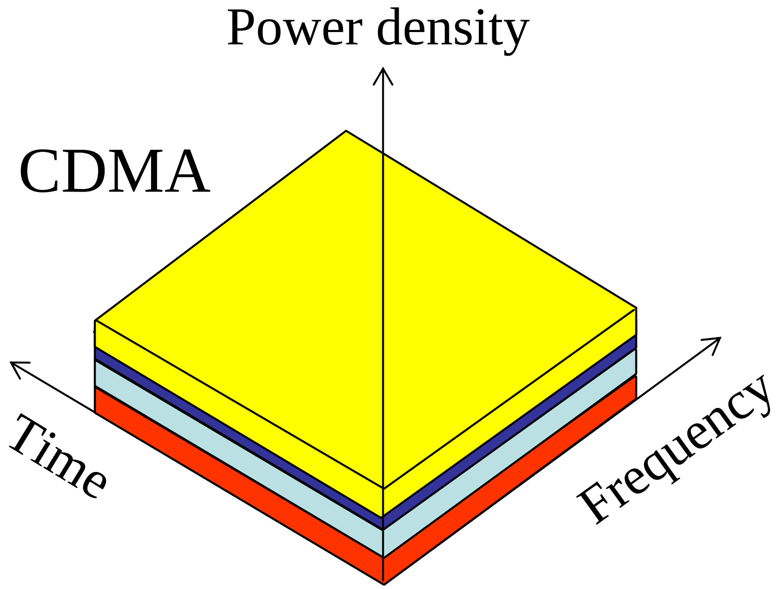


Transmission is organized in repetitive "time-frames". Each frame consists of groups of pulses - time slots. Each user/ link is assigned a separate time-slot.

Example: DECT (Digital enhanced cordless phone) Frame lasts 10 ms, consists of 24 time slots (each 417μs)

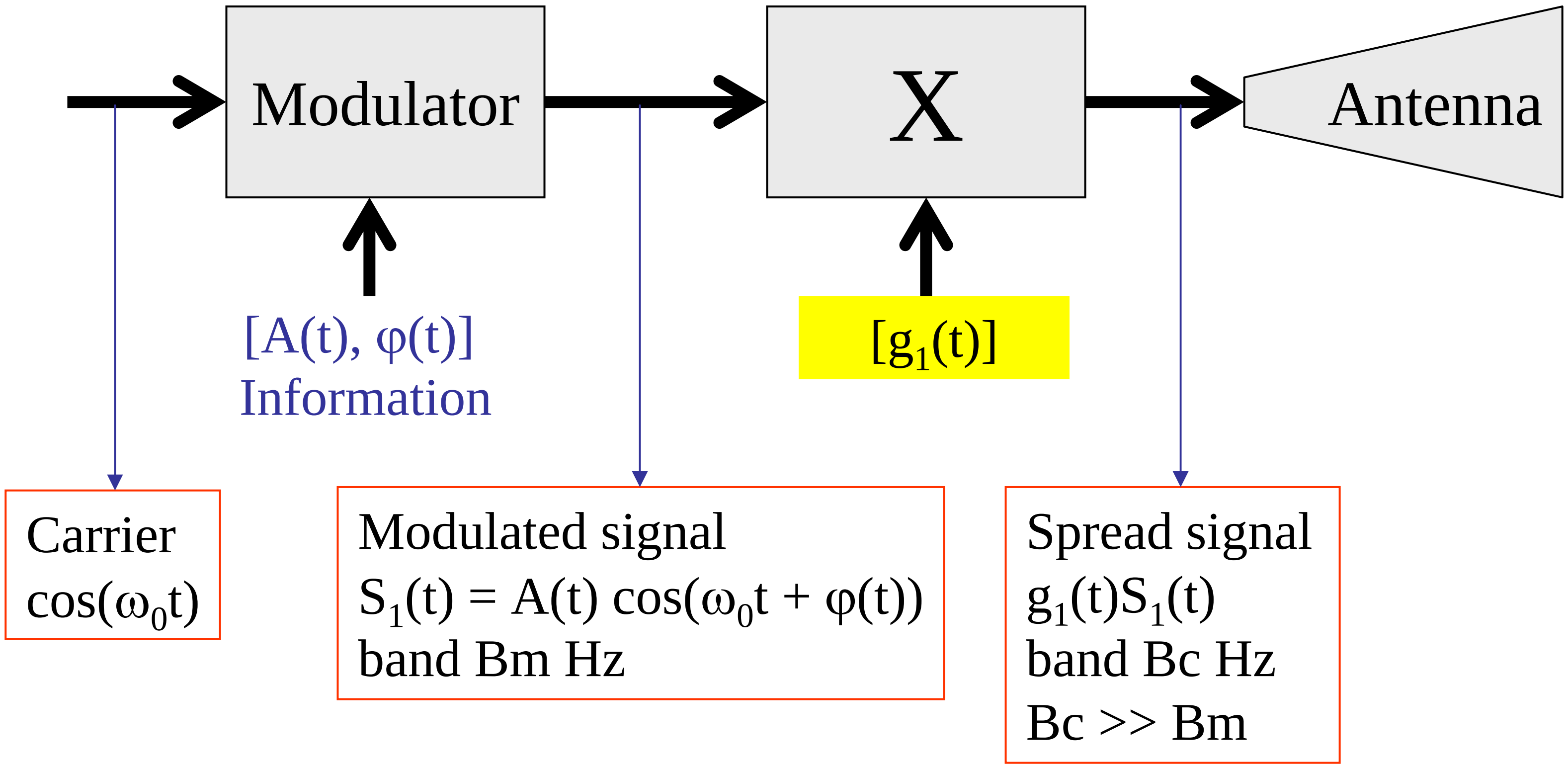
FH SS (CDMA)





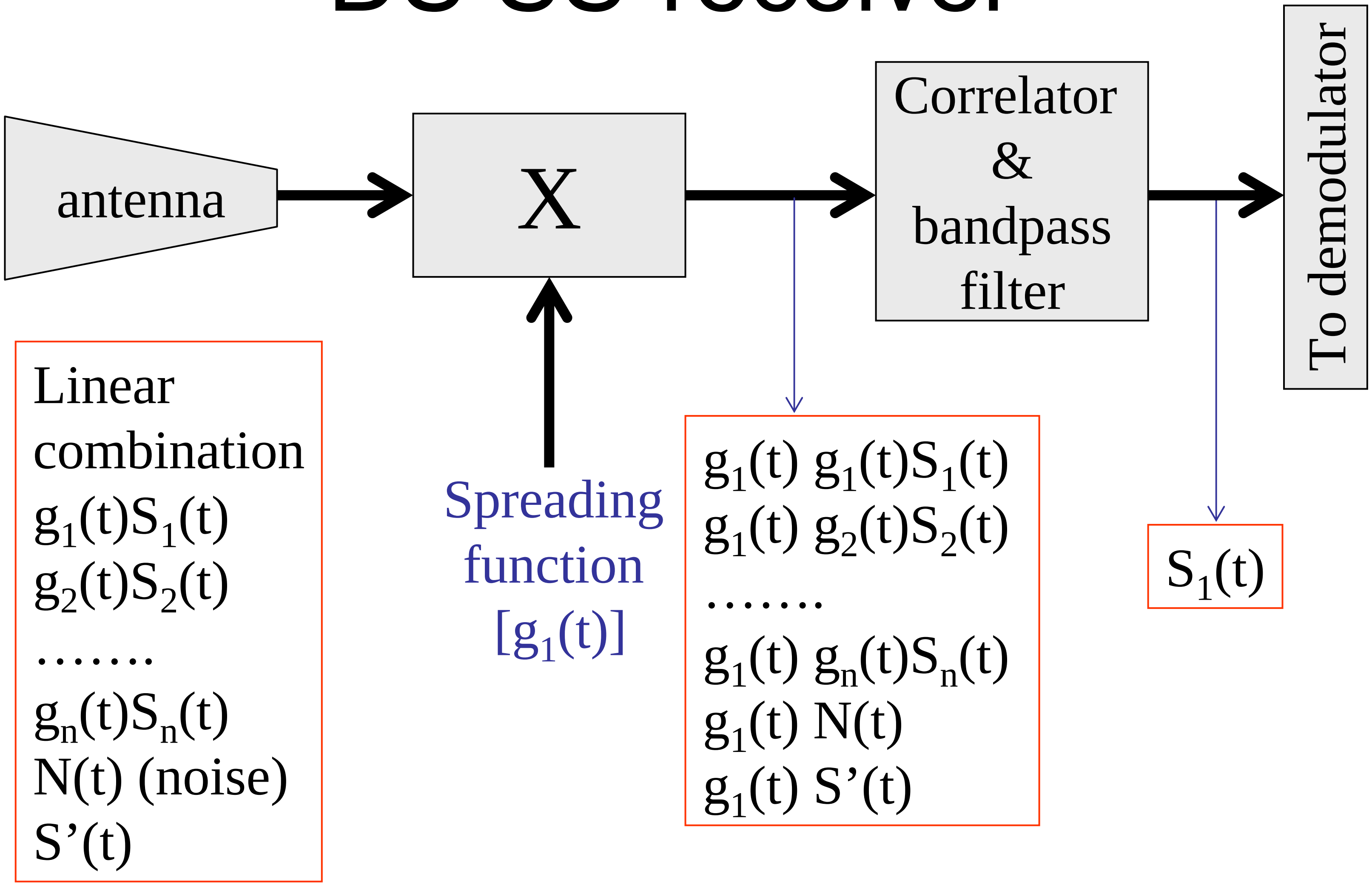
Transmission is organized in time-frequency “slots”. Each link is assigned a sequence of the slots, according to a specific code.

DS SS: transmitter

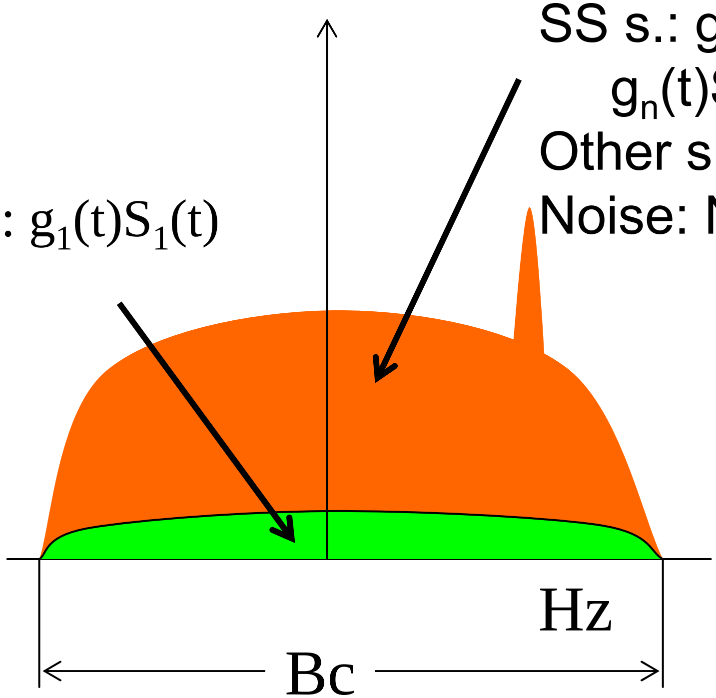


$g_i(t)$: pseudo-random noise (PN) spreading functions that spreads the energy of $S_1(t)$ over a bandwidth considerably wider than that of $S_1(t)$: ideally $g_i(t)\, g_j(t) = 1$ if $i = j$ and $g_i(t)\, g_j(t) = 0$ if $i \neq j$

DS SS-receiver



SS-receiver's Input



Signal-to-interference ratio $(S/ I)_{in} = S/ [I(\omega)*Bc]$

Bc = Input correlator bandwidth

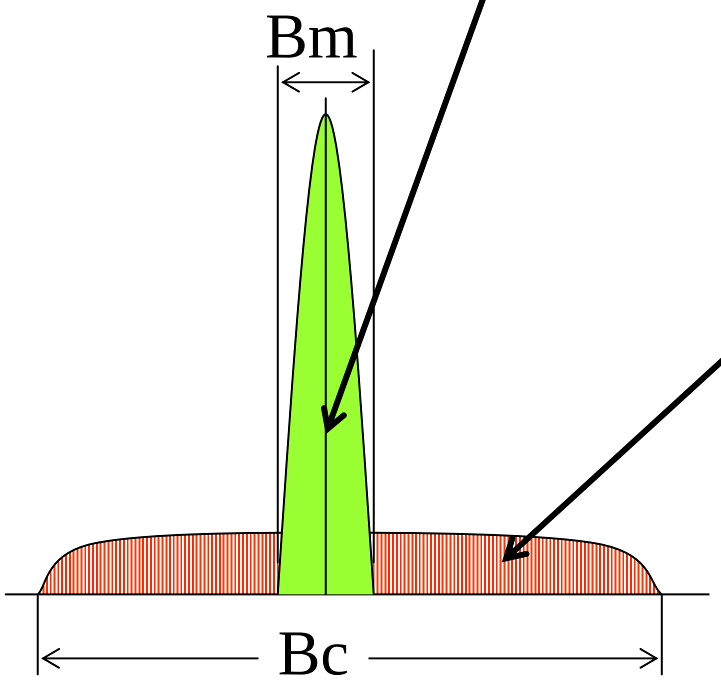
$I(\omega)$ = Average spectral power density of unwanted signals in Bc

S = Power of the wanted signal

SS-correlator/ filter output

Wanted (correlated) signal: de-spread to its original bandwidth

as $g_1(t)\, g_1(t) S_1(t) = S_1(t)$ with $g_1(t)\, g_1(t) = 1$



Uncorrelated (unwanted) signals

spread & rejected by correlator + noise

$g_1(t)\, S'(t);\ g_1(t)\, N(t);\ g_1(t)\, g_j(t) S_j(t) = 0$

as $g_i(t)\, g_j(t) = 0$ for $i \neq j$

Signal-to-interference ratio

$$(S/\ I)_{out} = S/\ [I(\omega)*Bm]$$

Bc = Input correlator bandwidth

Bm = Output filter bandwidth

I(ω) = Average spectral power density of unwanted signals & noise in Bm

S = power of the wanted signal at the correlator output

Spreading = reducing spectral power density

SS Processing Gain =

$$= [(S/I)_{in} / (S/I)_{out}] = \sim Bc/Bm$$

Example: GPS signal

RF bandwidth Bc ~ 2MHz   Filter bandwidth Bm ~ 100 Hz

Processing gain ~20’000 (+43 dB)

Input S/N = -20 dB   (signal power = 1% of noise power)

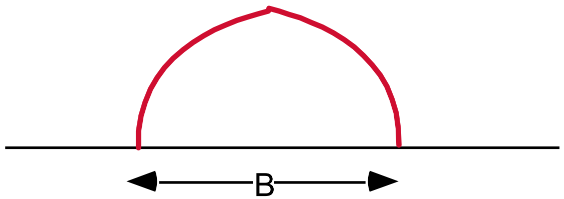
Output S/N = +23 dB   (signal power = 200 x noise power)

(GPS = Global Positioning System)

OFDM

- Basic idea:
  - Using a large number of parallel narrow-band sub-carriers instead of a single wide-band carrier to transport information
- Advantages
  - Efficient in dealing with multi-path and selective fading
  - Robust again narrow-band interference
- Disadvantages
  - Sensitive to frequency offset and phase noise
  - Peak-to-average problem reduces the power efficiency of RF amplifier at the transmitter
- Adopted for various standards
    - DSL, 802.11a, DAB, DVB

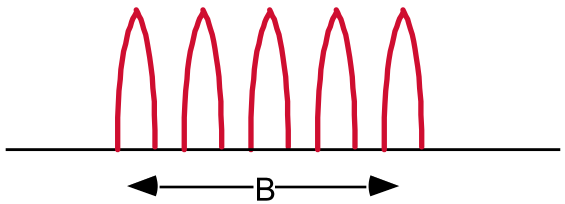
OFDM



Pulse length ~1/B

– Data are transmited over **only one carrier**

N carriers



Pulse length ~ N/B

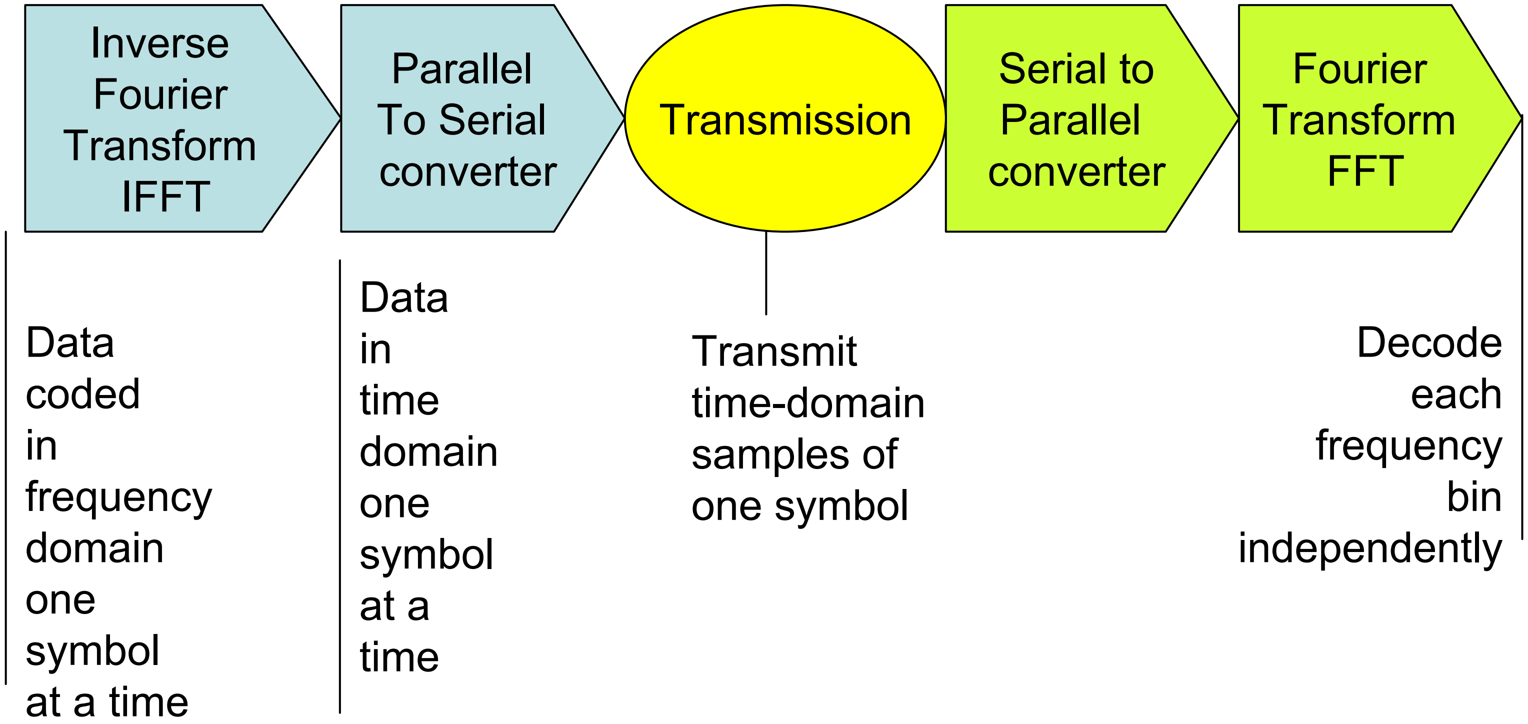
– Data are shared among **several carriers** and simultaneously transmitted

Similar to FDM technique

– Selective Fading
– Very short pulses

– Flat Fading per carrier
– N long pulses

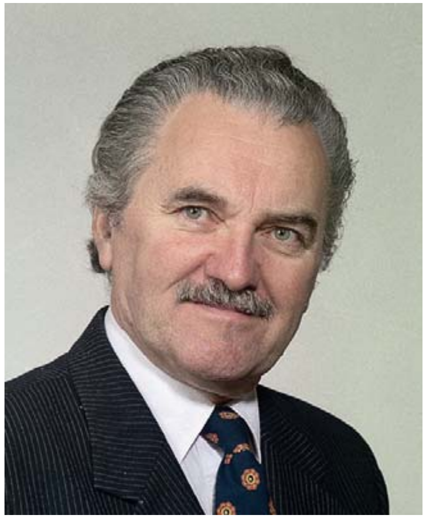
OFDM modulation & demodulation



http://www.iss.rwth-aachen.de/Projekte/Theo/OFDM/node2.html

- http://en.wikipedia.org/wiki/QPSK
- http://en.wikipedia.org/wiki/Orthogonal_frequency-division_multiplexing
- "The how and why of COFDM" Jonathan Stott. EBU: *EBU Technical Review* 278 (winter 1998)

- Thank yoy for your attention



Ryszard STRUZAK *PhD., DSc.*
Co-Director, ICTP-ITUD School on Wireless Networking, IT
Academician, International Telecommunication Academy
Life Fellow IEEE
ryszard@struzak.com
www.ryszard.struzak.com

Important notes

Beware of misprints!!! These materials are preliminary notes for my lectures and may contain misprints. If you notice some, or if you have comments, please send these to r.struzak@ieee.org.