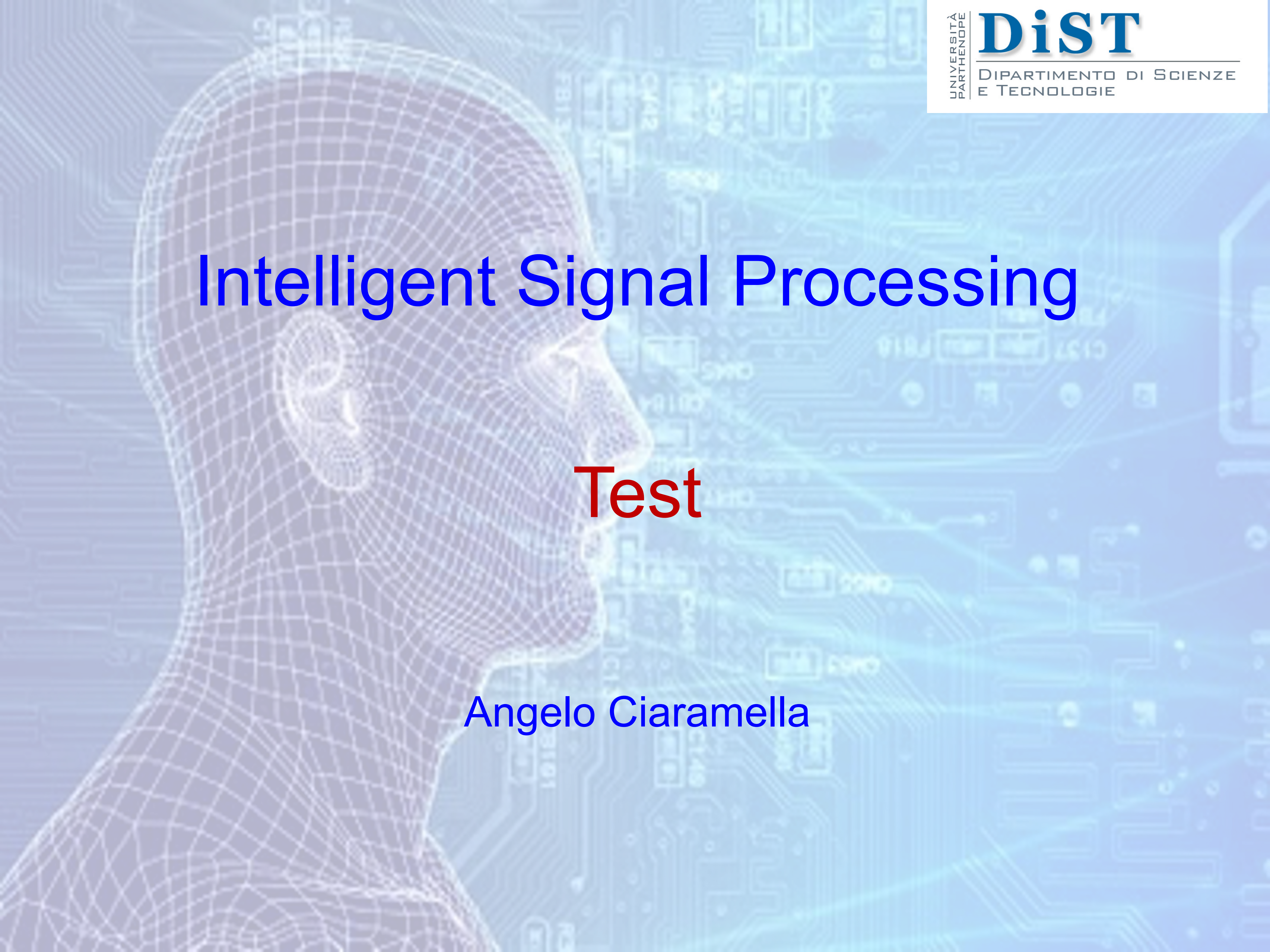


UNIVERSITÀ PARTHENOPE

DiST

DIPARTIMENTO DI SCIENZE E TECNOLOGIE

Intelligent Signal Processing

Test

Angelo Ciaramella

Question 1

- The waveform of the pure tone coincides with a sinusoidal trigonometric function

$$\boxed{y(t) = \mathbf{A}\sin(t)}$$

- Question
  - Explain its properties and the waveform

ISP – Verification tests

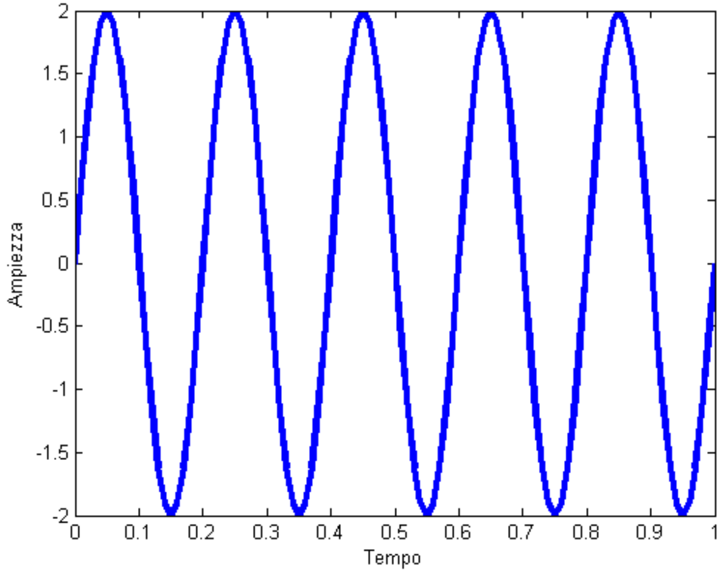
Pure tone properties

- The following are the properties of the pure tone
  - Frequency (f)
  - Angular frequency (ω)
  - Period (T)
  - Wavelength (λ)
  - Amplitude (A)
  - Phase (φ)
  - Initial phase (φ_0)
  - Speed (v)

ISP – Verification tests

Frequency

- Frequency
  - the number of cycles accomplished by the wave in a second
  - positive half-wave and a negative half-wave
  - measured in Hz [1 / sec]
  - equal to 1 Hz is a cycle every second



Frequency of # Hz

Angular frequency

- The angular frequency is defined as

$$\boxed{\omega = 2\pi f}$$

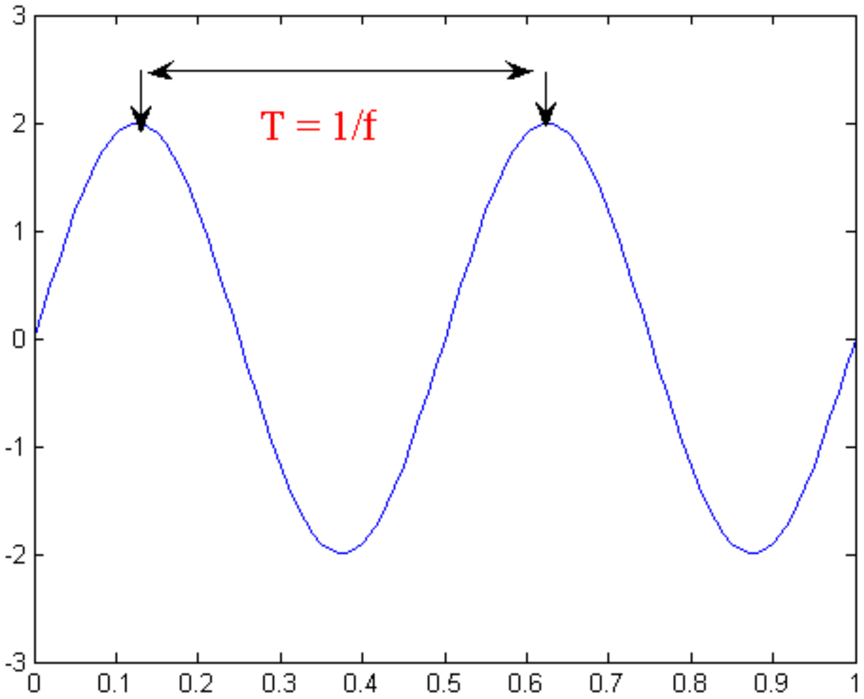
- It is expressed in radians

$$\boxed{2\pi \equiv 360^\circ}$$

Period

- The period is the time for achieving a complete cycle

$$\boxed{T = \frac{1}{f}}$$



Example of period

Wavelength

- The distance between two corresponding points along the waveform

$$\boxed{\lambda = \frac{c}{f}}$$

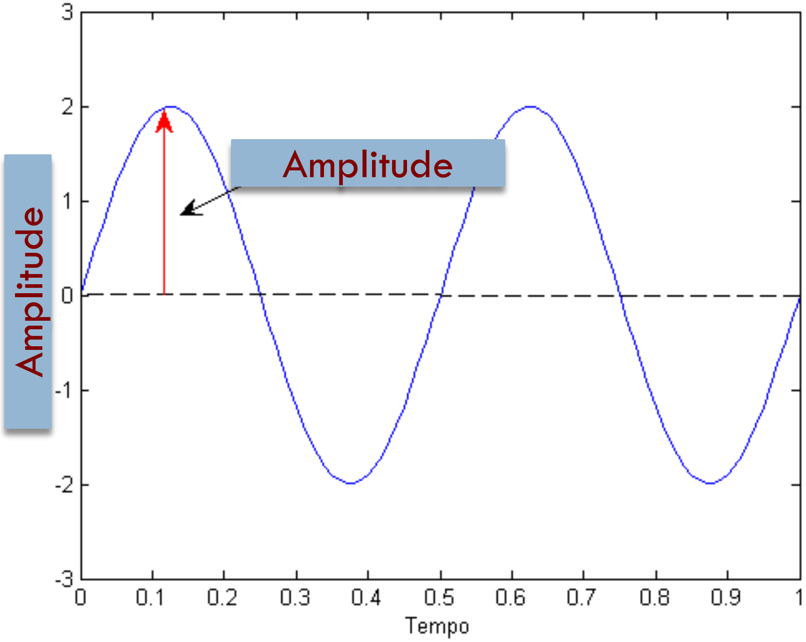
c is the speed of the sound in the considered medium (344 m/sec in air)

- Example of wavelength
    - 1 Hz frequency wave, travelling through the air

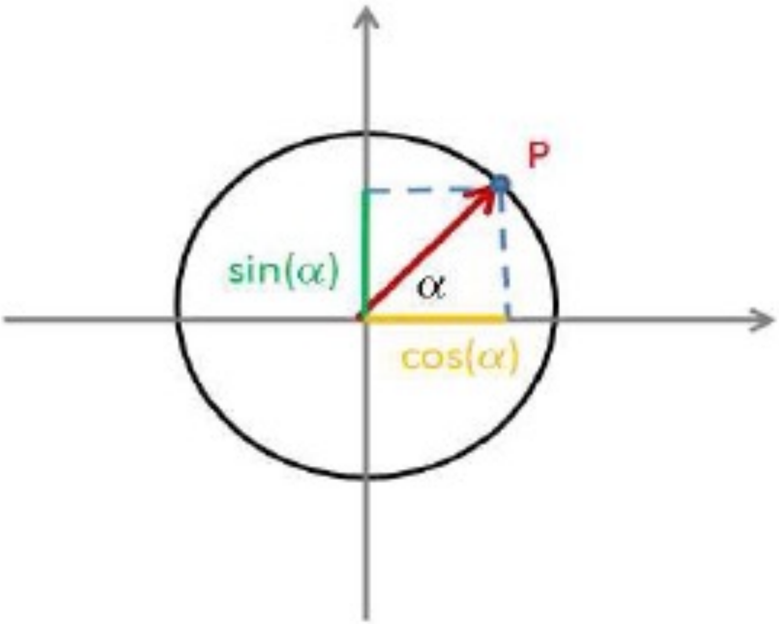
$$\lambda = \frac{c}{f} \Rightarrow \frac{344m/s}{1/\frac{1}{s}} = 344m$$

Amplitude

- It is the measure of the maximum deviation from the equilibrium position
- Larger amplitudes correspond to higher volumes

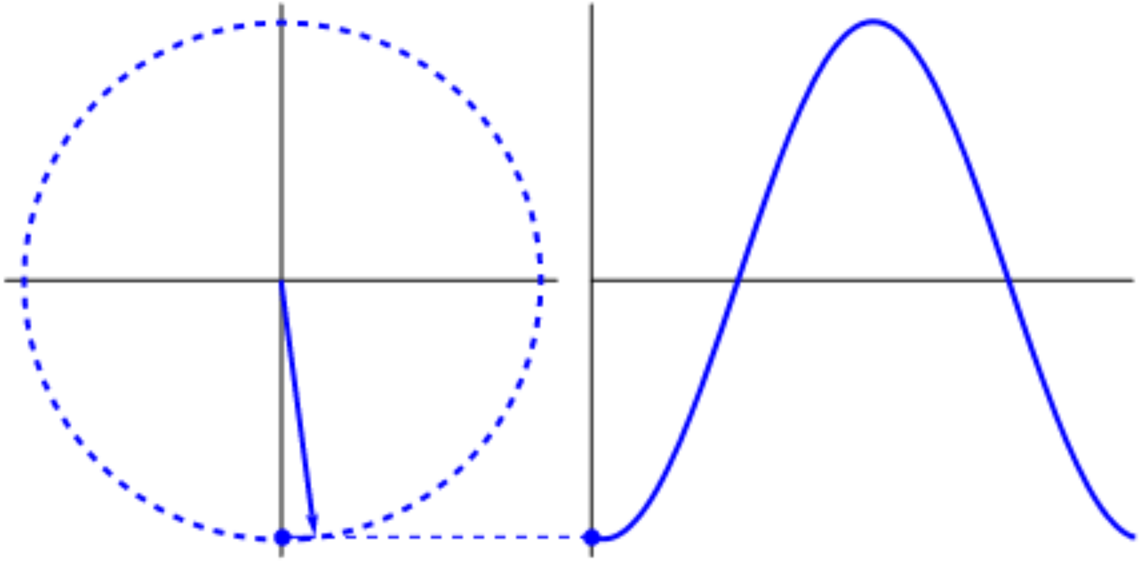


Phase



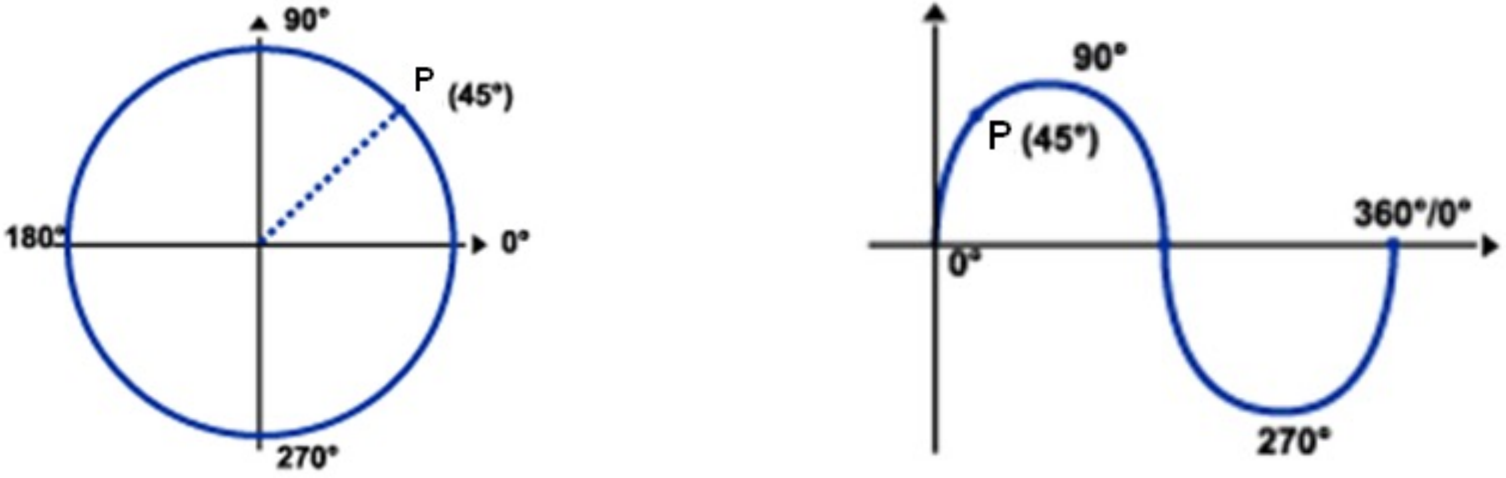
We consider a point moving on a circumference

Phase



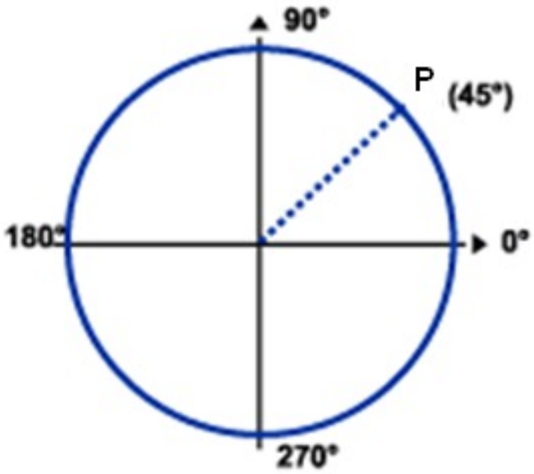
We consider a point moving on a circumference

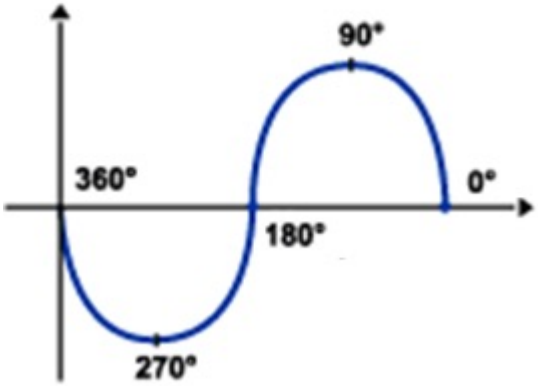
Phase



We imagine to rotate the point P counterclockwise and to observe its projection on the y axis

Phase





We imagine to rotate the point P clockwise and to observe its projection on the y axis

Frequency and time

- Alternative interpretation of frequency
  - the number of times that the point P makes a complete turn in a second
- The equation that correlates the phase with time is

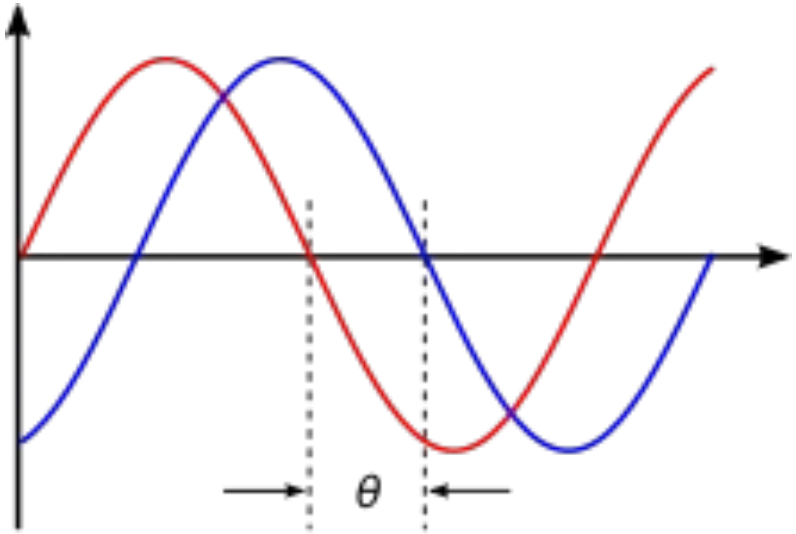
$$\boxed{\Delta t = t - t_0}$$

$$\boxed{\varphi = 2\pi f \Delta t}$$

$$\boxed{t_0 = 0 \Rightarrow \Delta t = t}$$

Initial phase

- The initial phase φ_0 is the offset from where you start to look at the pure tone



Pure tones with different phases

Pure tone equation

- By introduced parameters, the waveform of the pure tone is

$$\boxed{y(t) = A\sin(\varphi + \varphi_0) = A\sin(2\pi f t + \varphi_0)}$$

References

- Material
  - Slides
  - Video Lessons
- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 2

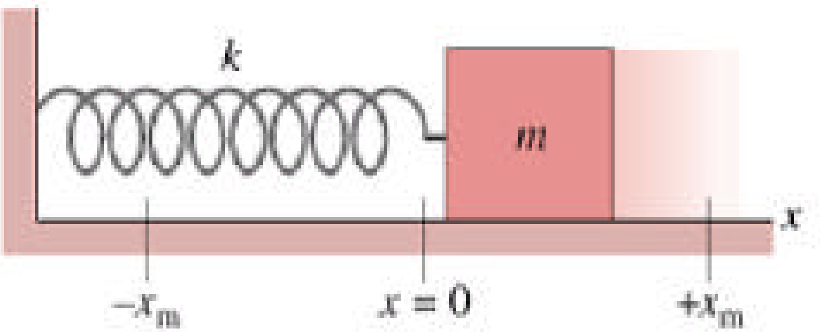
- The waveform of the pure tone coincides with a trigonometric function

$$\boxed{x(t) = A\cos(\omega t + \phi_0)}$$

- Question
  - Which is the correlation between the pure tone and the oscillating systems?

Oscillating systems

- A body of mass m moves along the x-axis under the action of a ideal spring with elastic constant *k* and in the absence of dissipative forces



- From the second Newton's law

$$\boxed{-kx = ma}$$

Oscillating systems

- In particular we obtain

$$\boxed{\frac{d^2x}{dt^2}+\frac{k}{m}x=0}$$

- The solution of the differential equation is

A = x$_m$

$$\boxed{x(t)=A\cos(\omega t+\phi_0)}$$

$$\boxed{\omega=\sqrt{\frac{k}{m}}=2\pi f}$$

Initial phase

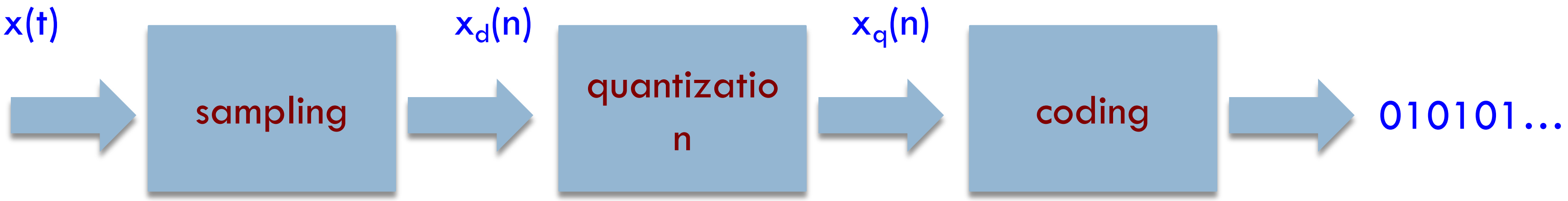
References

- Material
  - Slides
  - Video Lessons
- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 3

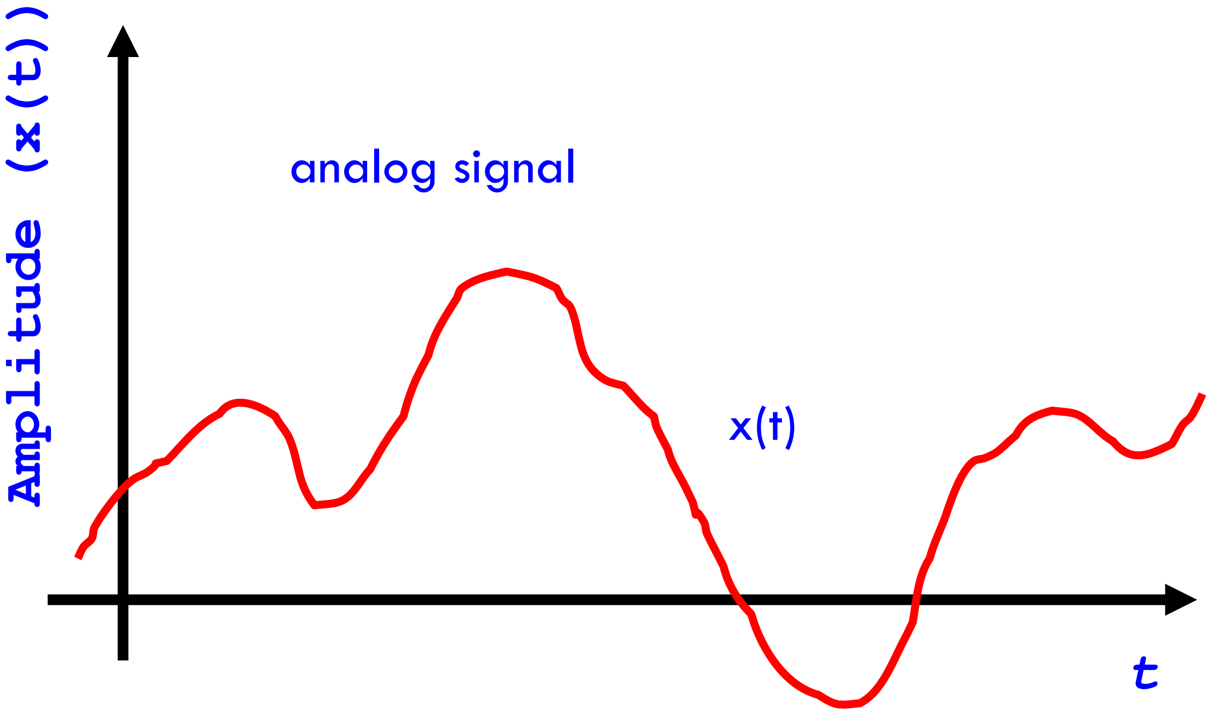
- Question
    - Describe the process for sampling a signal and the Nyquist theorem

Analog to Digital Converter



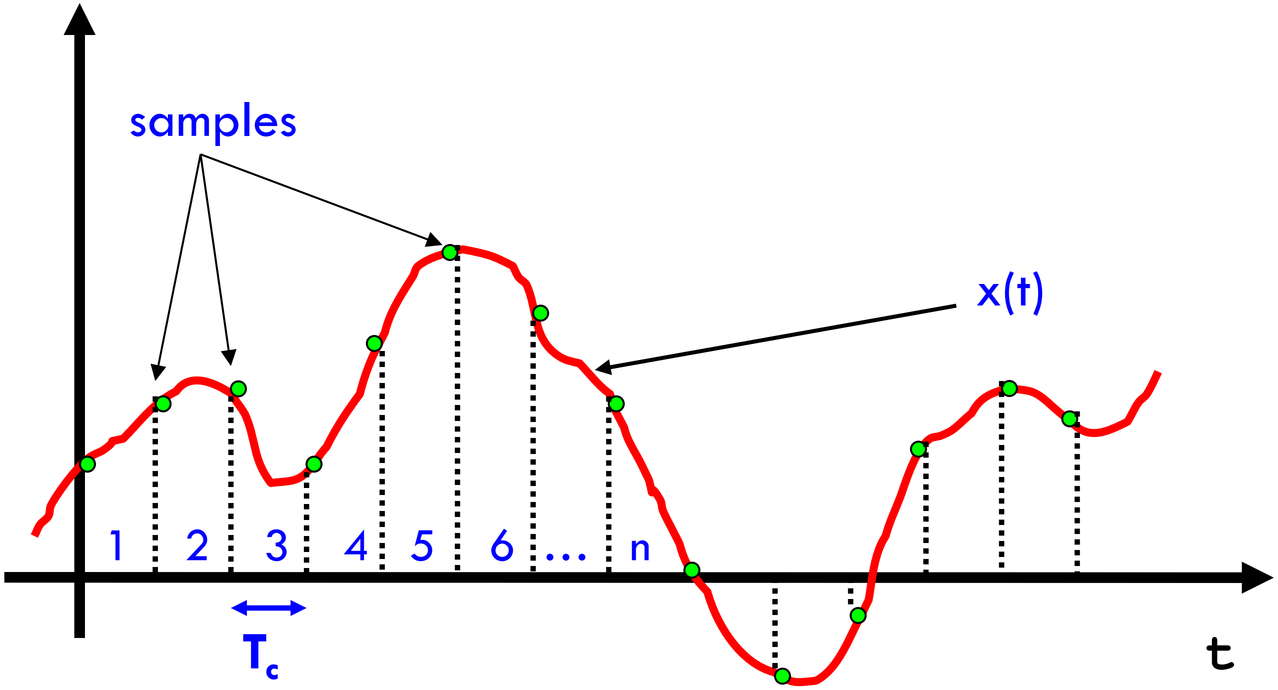
Main phases of the ADC process: sampling, quantization, coding.

Analog signal



We start from a temporal representation of analog signal

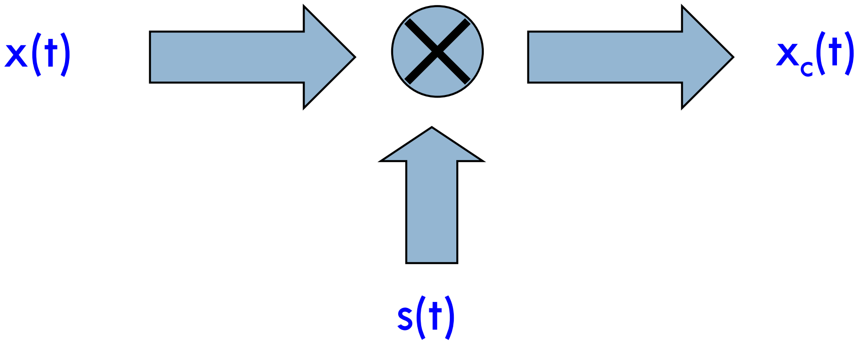
Analog signal sampling



We define a sampling period T_c obtaining a sampled signal

$$x_d(n) = x(nT_c)$$

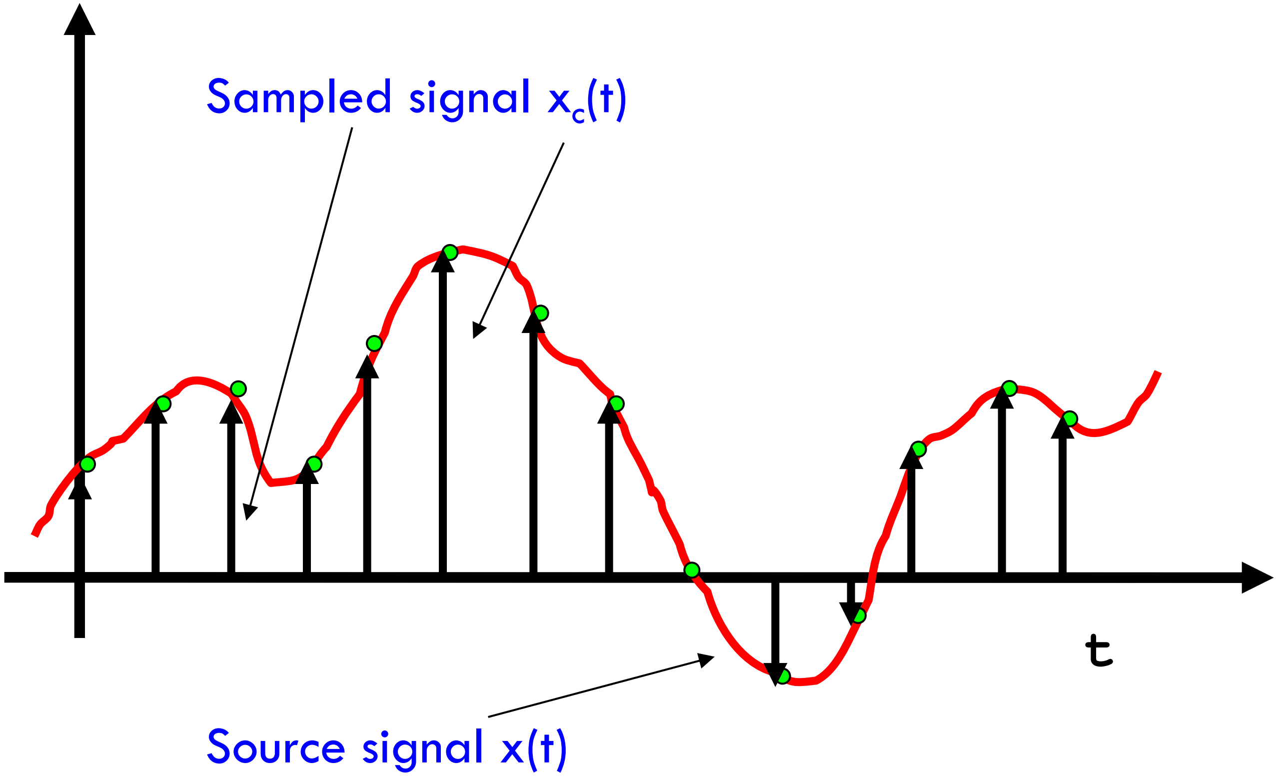
Pulse Code Modulation (PCM)



Pulse Code Modulation (PCM) is a mathamtical method used to digitally represent analog signals.

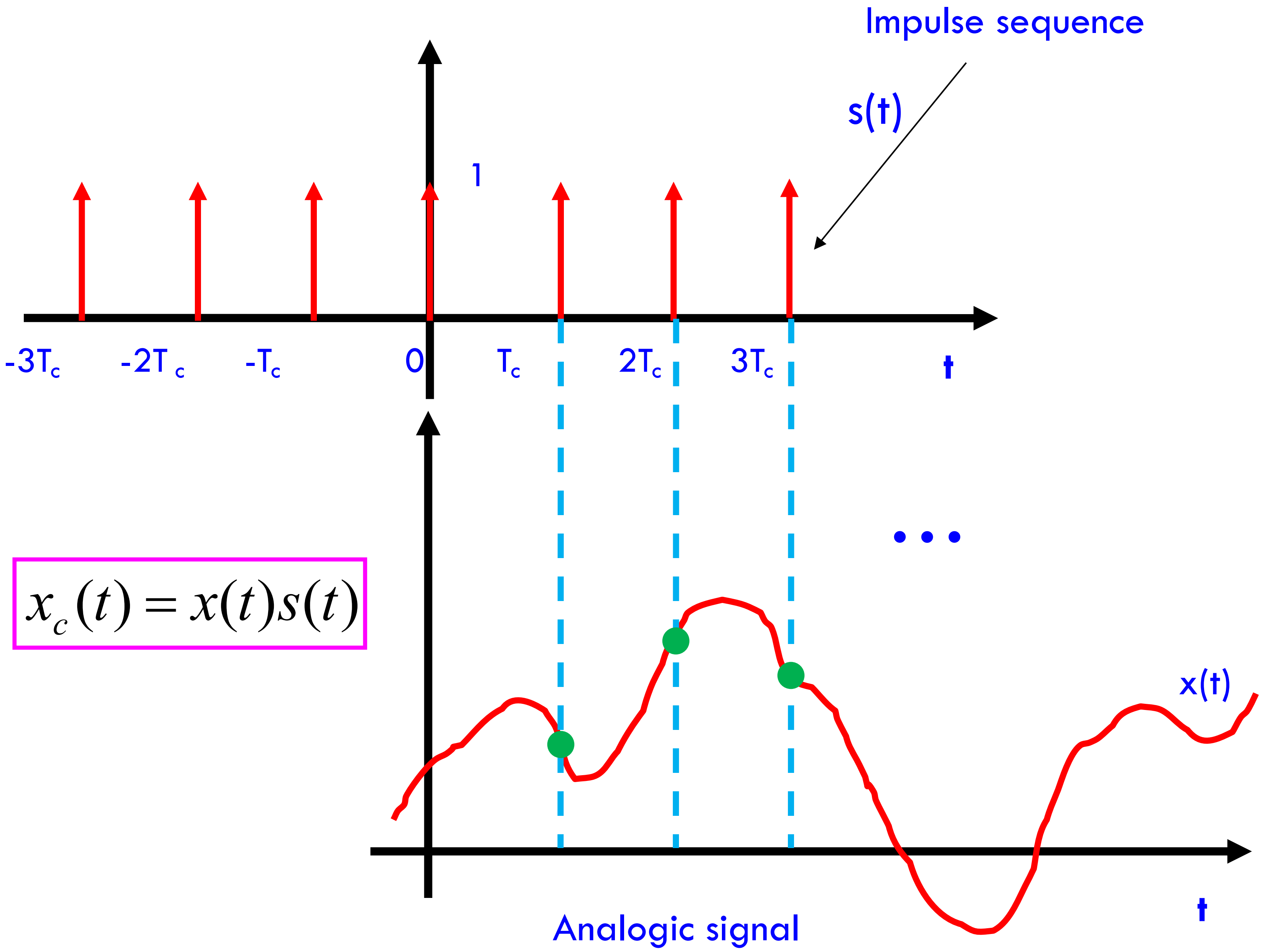
We consider s(t) as a periodic impulse sequence obtained by $\delta(t)$ (delta di Dirac)

Pulse Code Modulation (PCM)



Result of the PCM

Pulse Code Modulation (PCM)



$$x_c(t) = x(t)s(t)$$

Sampling theorem

- The Sampling theorem defines the minimum sampling frequency which is necessary to avoid distortions in the signal reconstruction

- Introduced by Harold Nyquist, and appeared in 1949 in an article authored by E. C. Shannon

- Result
  - Given a signal with a limited and known bandwidth, the minimum sampling frequency of this signal must be at least twice its highest frequency

Sampling theorem

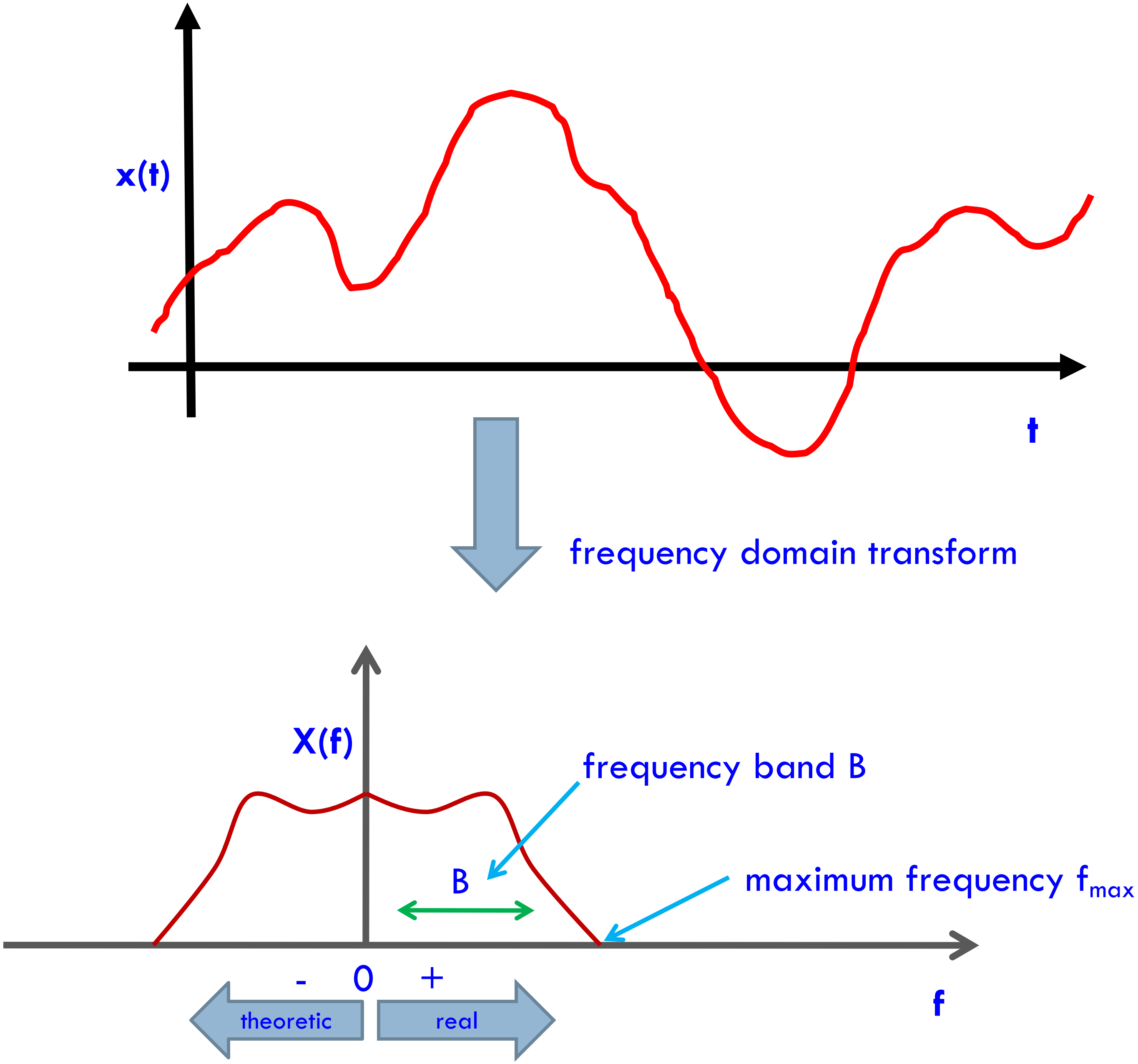
- **Theorem**

A continuous-time signal x(t) with a spectral band B strictly limited (X(f) = 0 for | f | > ± B) can be uniquely reconstructed from its sampled version

x (n) (n = 0, ± 1, ± 2, ± 3, ...) if the sampling frequency $f_c = 1/T_c$ satisfies the following relation

$$f_c = \frac{1}{T_c} \geq 2\mathbf{B}$$

Sampling theorem

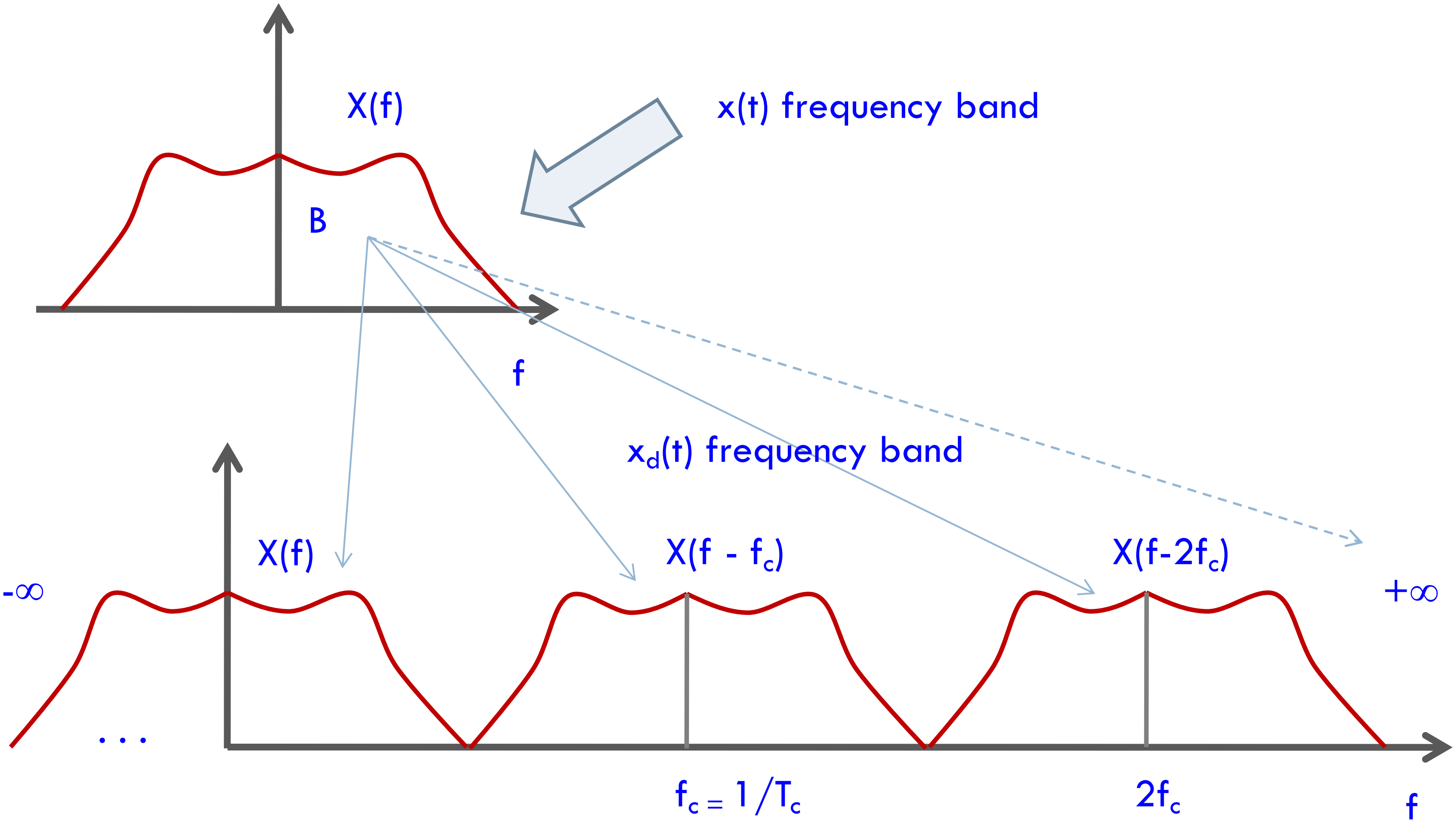


frequency domain transform

frequency band B

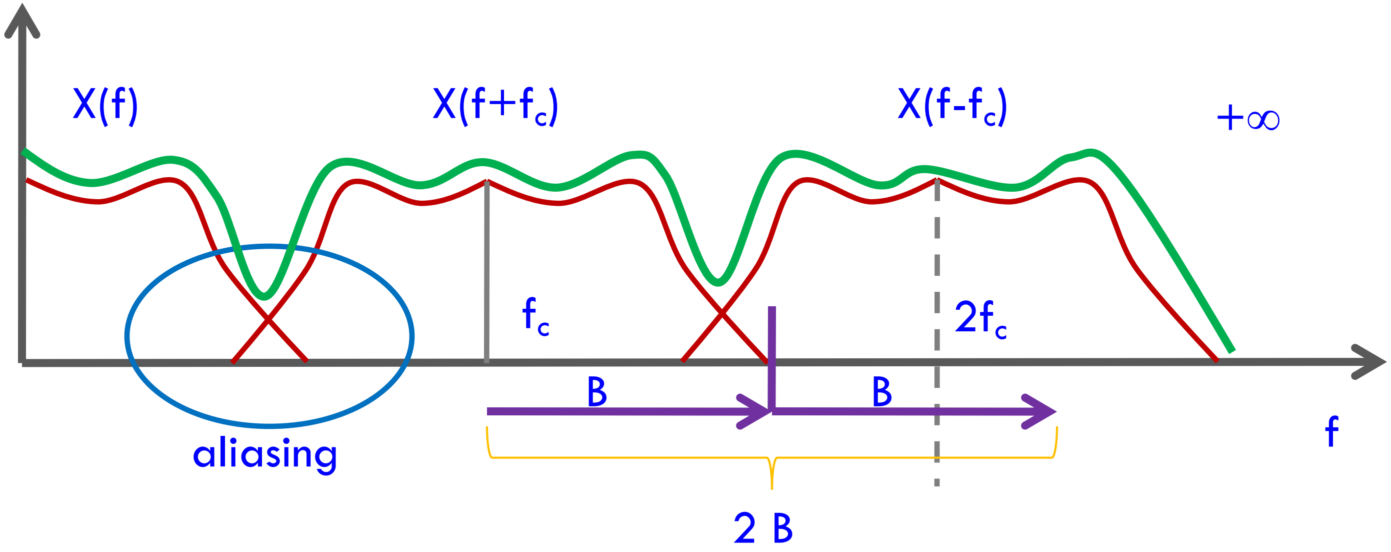
maximum frequency f_{max}

Sampling theorem



Sampling theorem

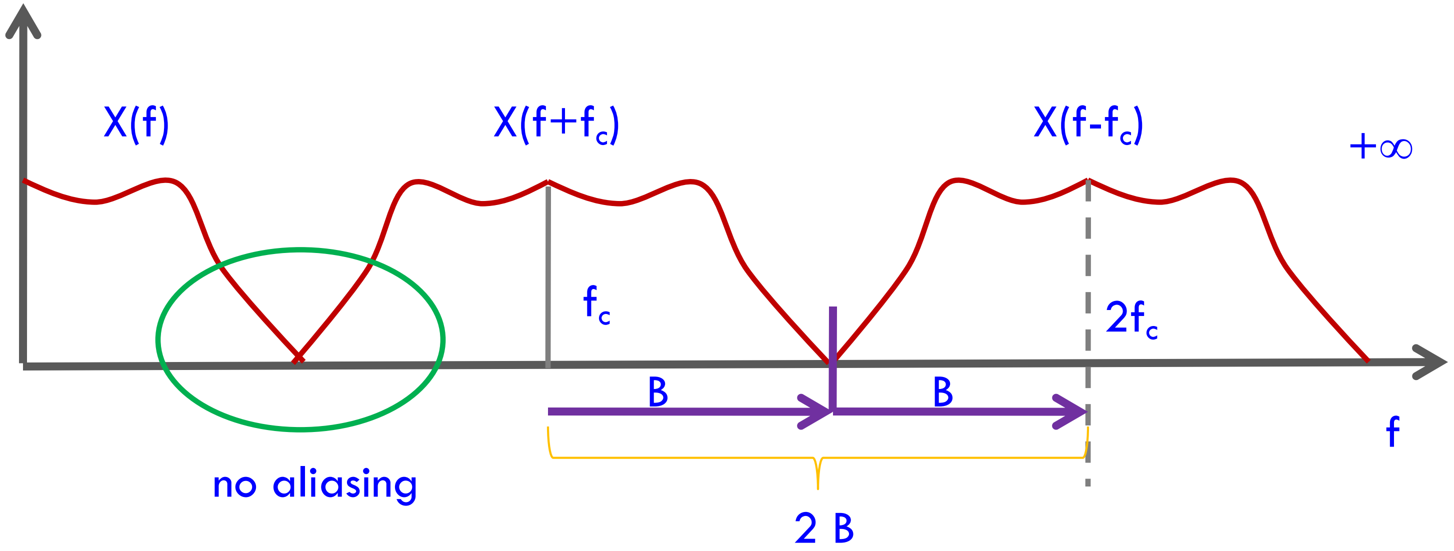
$$\boxed{f_c < 2B}$$



Subsampling of the signal generates aliasing

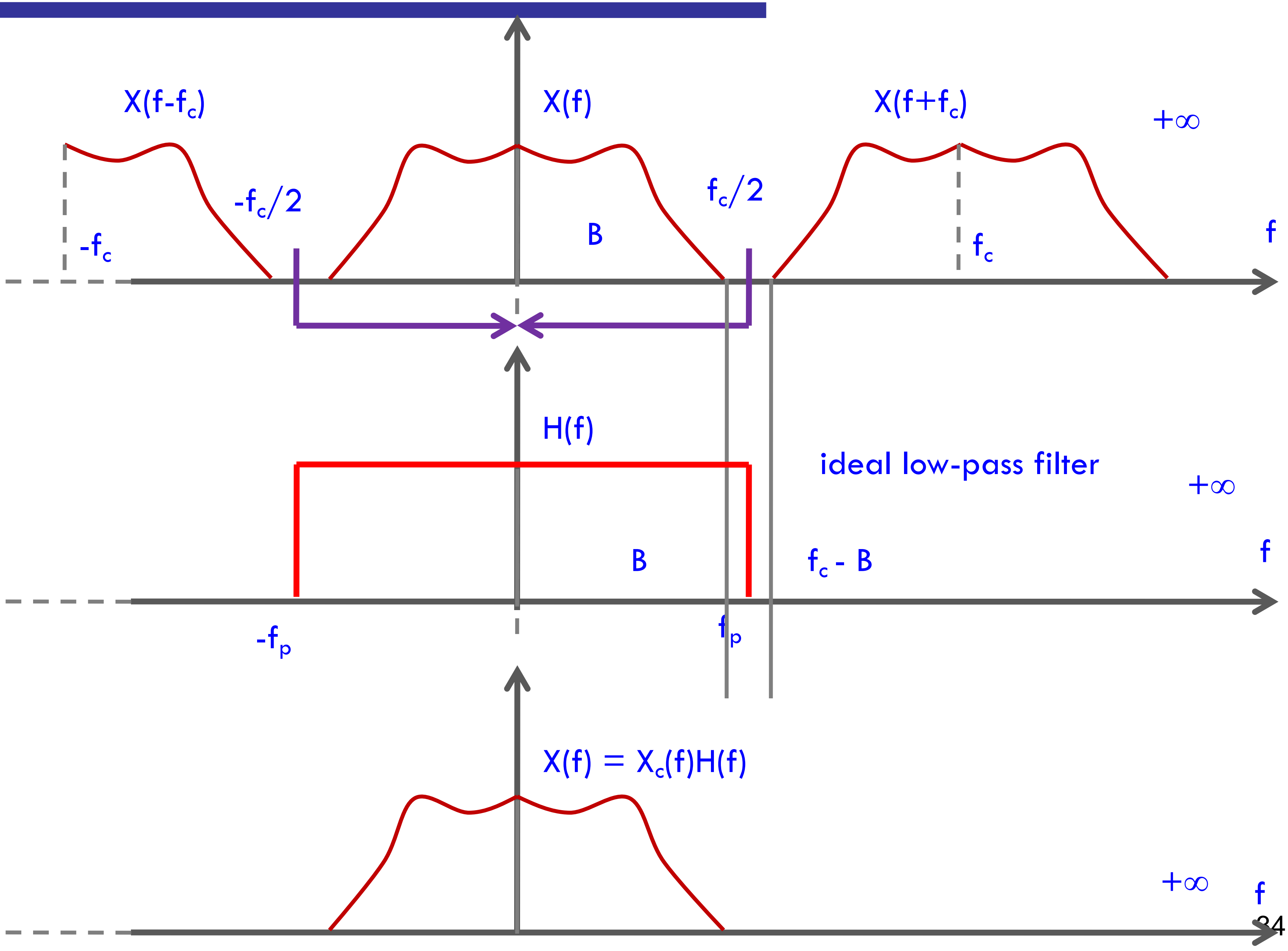
Sampling theorem

$$\boxed{f_c \geq 2B}$$

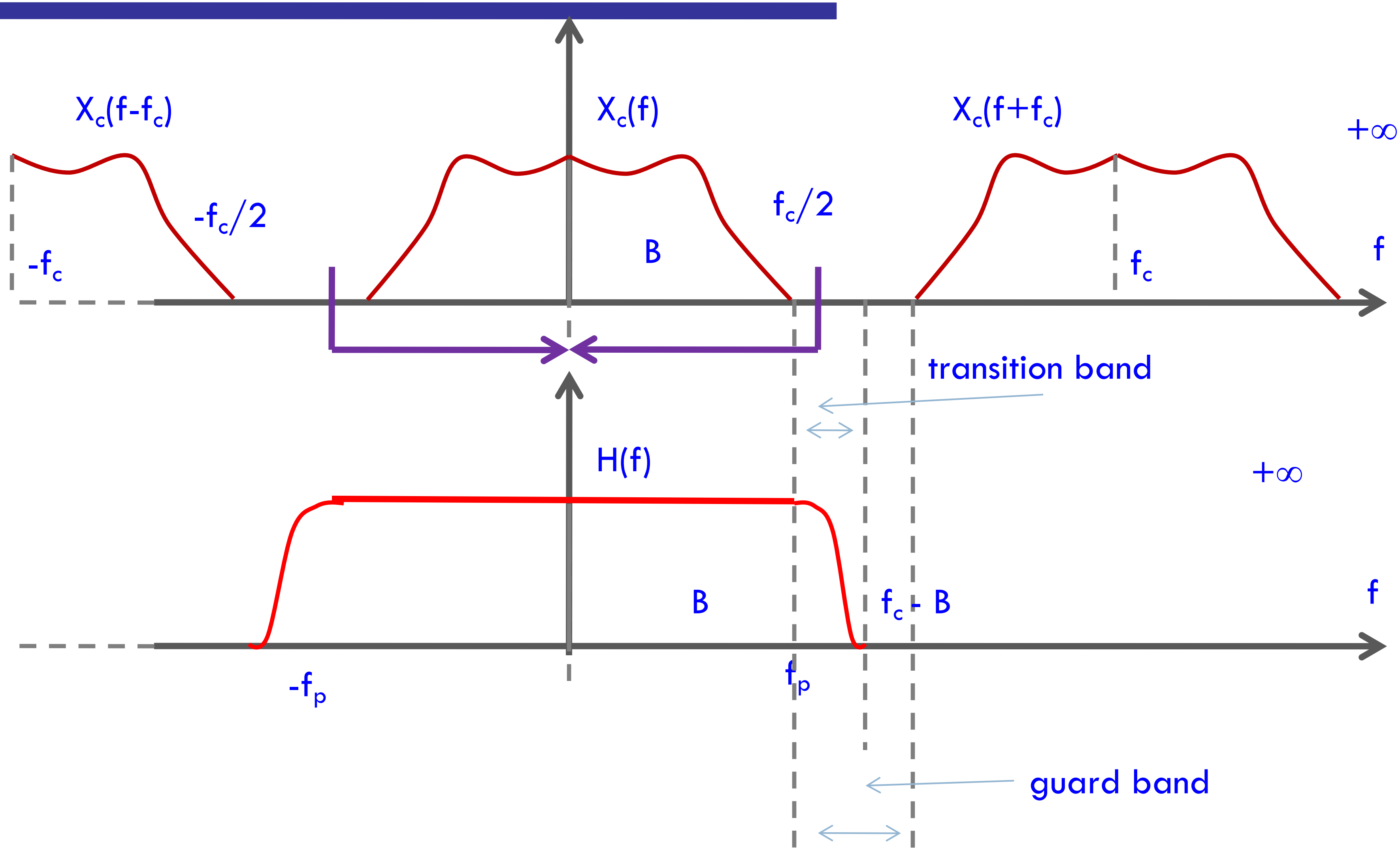


Oversampling of the signal

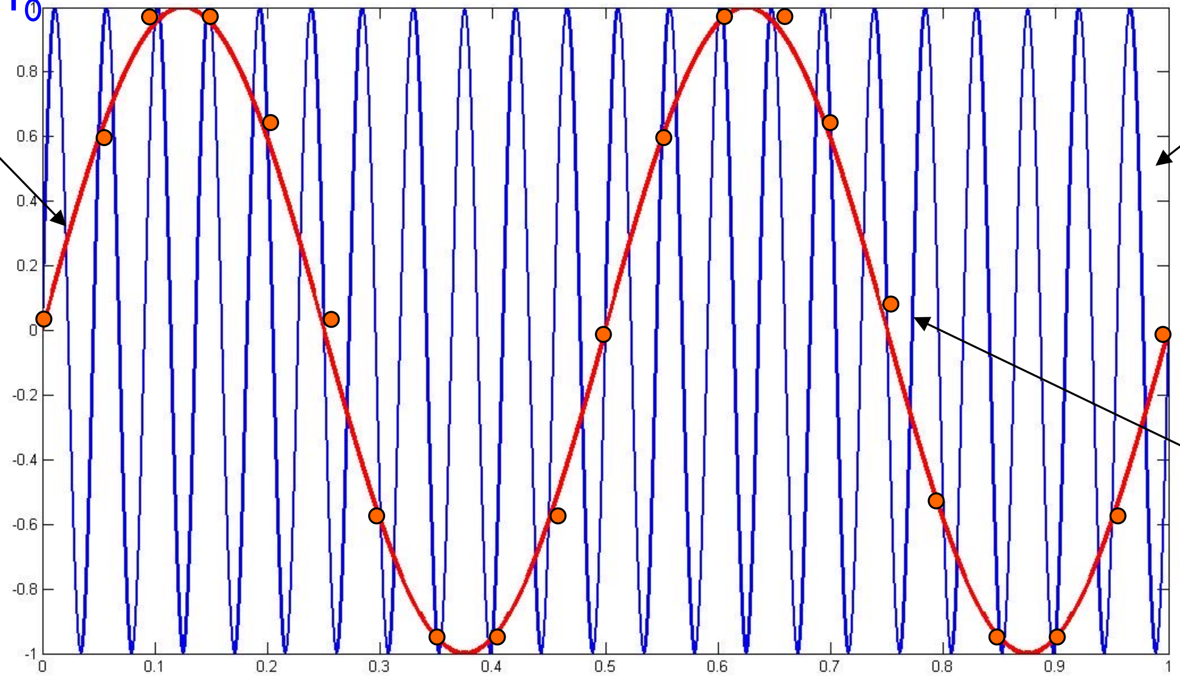
Signal reconstruction



Real filters



Example of reconstructed signal



Reconstruct signal

f_{alias} =

f_c - $f_0 < f_0$

Source signal with frequency f_0

Samples with $f_c < 2\ f_0$

$$f_{alias} = f_c - f_{reale} < f_{reale}$$

References

- Material
  - Slides
  - Video Lessons
- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 4

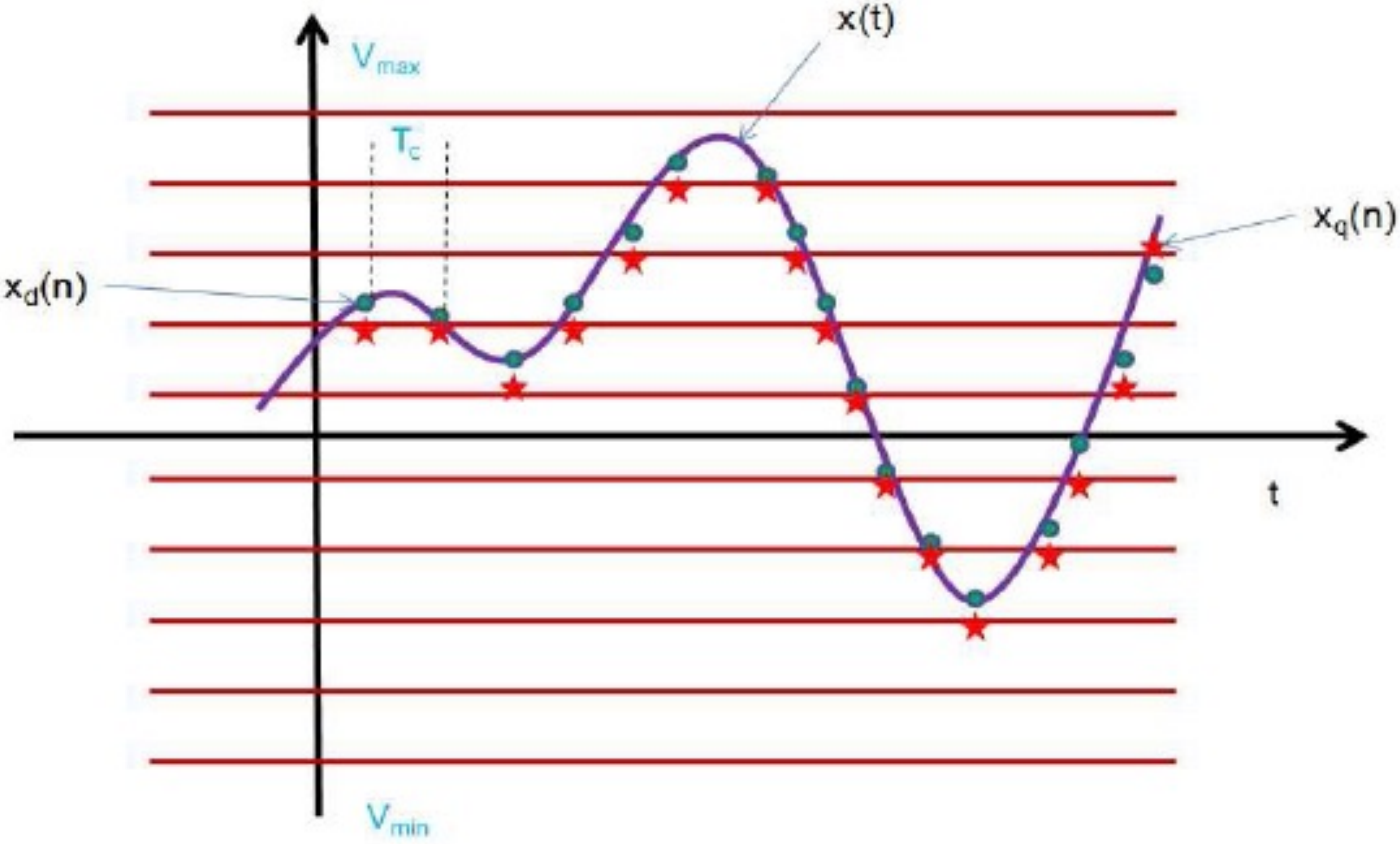
- Question
  - Describe the process of quantization

Quantization

- Quantization is the procedure of constraining something from a continuous set of values (such as the real numbers) to a relatively small discrete set (such as the integers)

- Quantization replaces each real number with an approximation from a finite set of discrete values (levels)
  - values are represented as fixed-point words or floating-point words
  - common word-lengths are 8-bit (256 levels), 16-bit (65,536 levels), 32-bit (4.3 billion levels

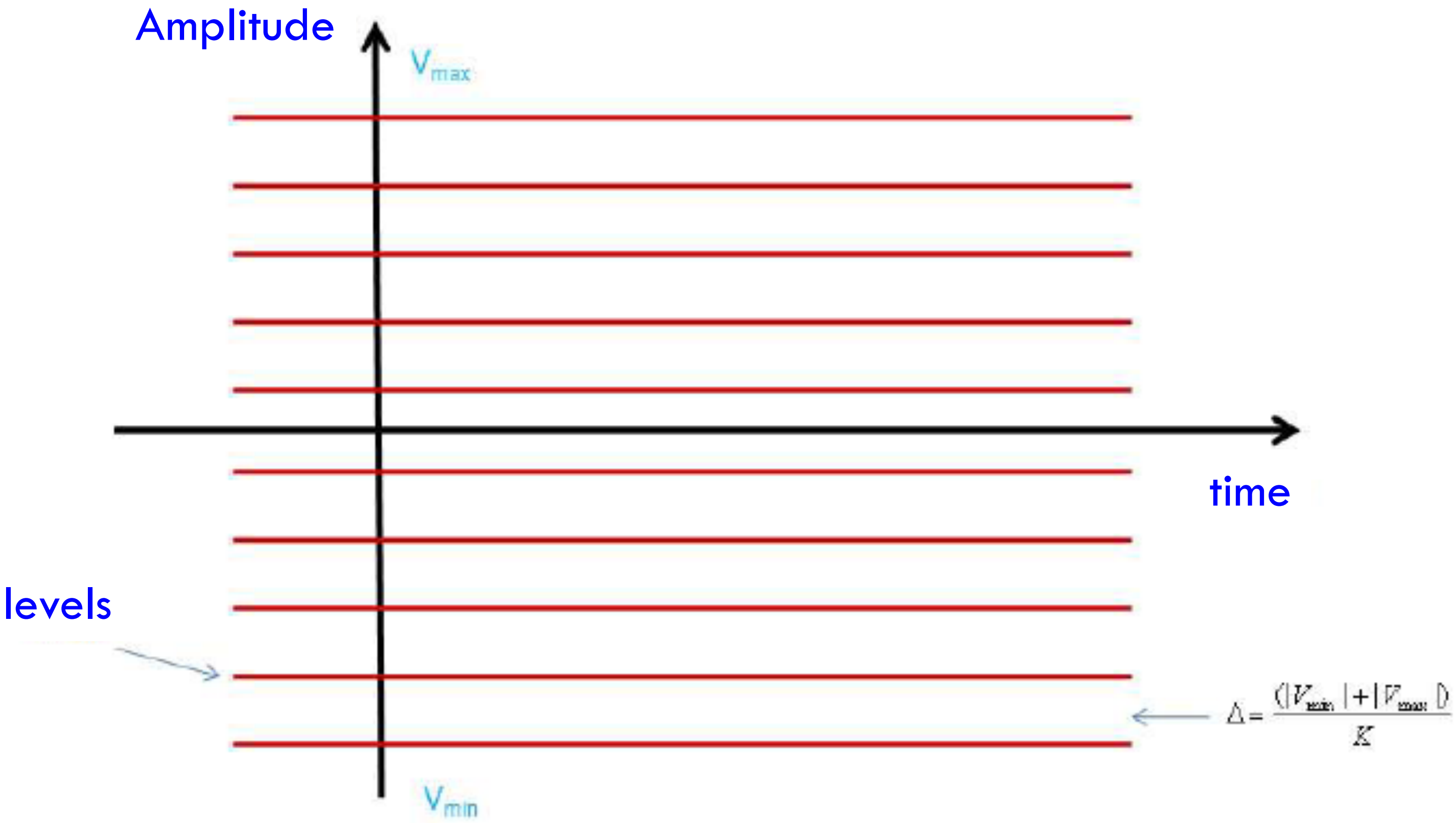
Example of quantization



$$e(n) = x_q(n) - x_d(n)$$

quantization error

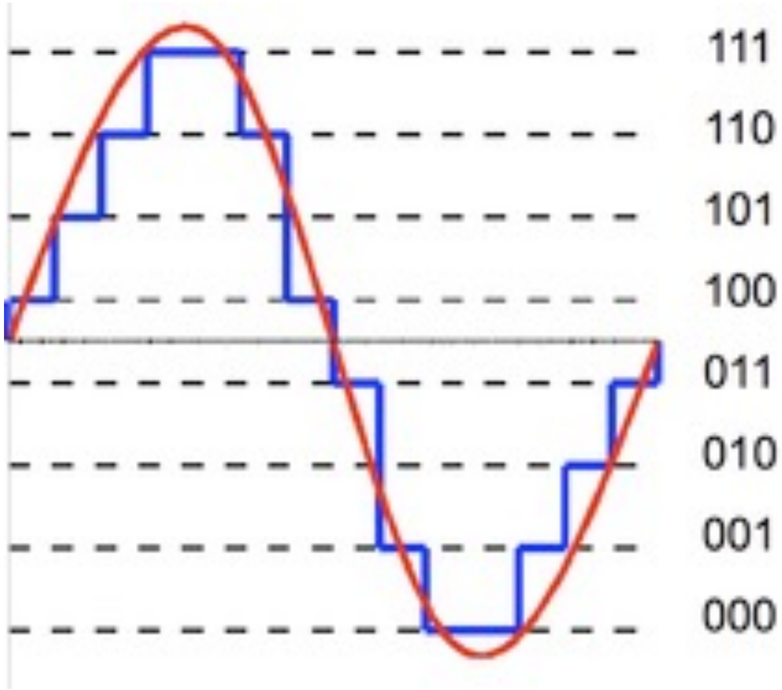
Coding



We start with an uniform quantization

Coding

- With N bits K = 2^N quantization levels are obtained
  - at each level a code of N bits can be associated



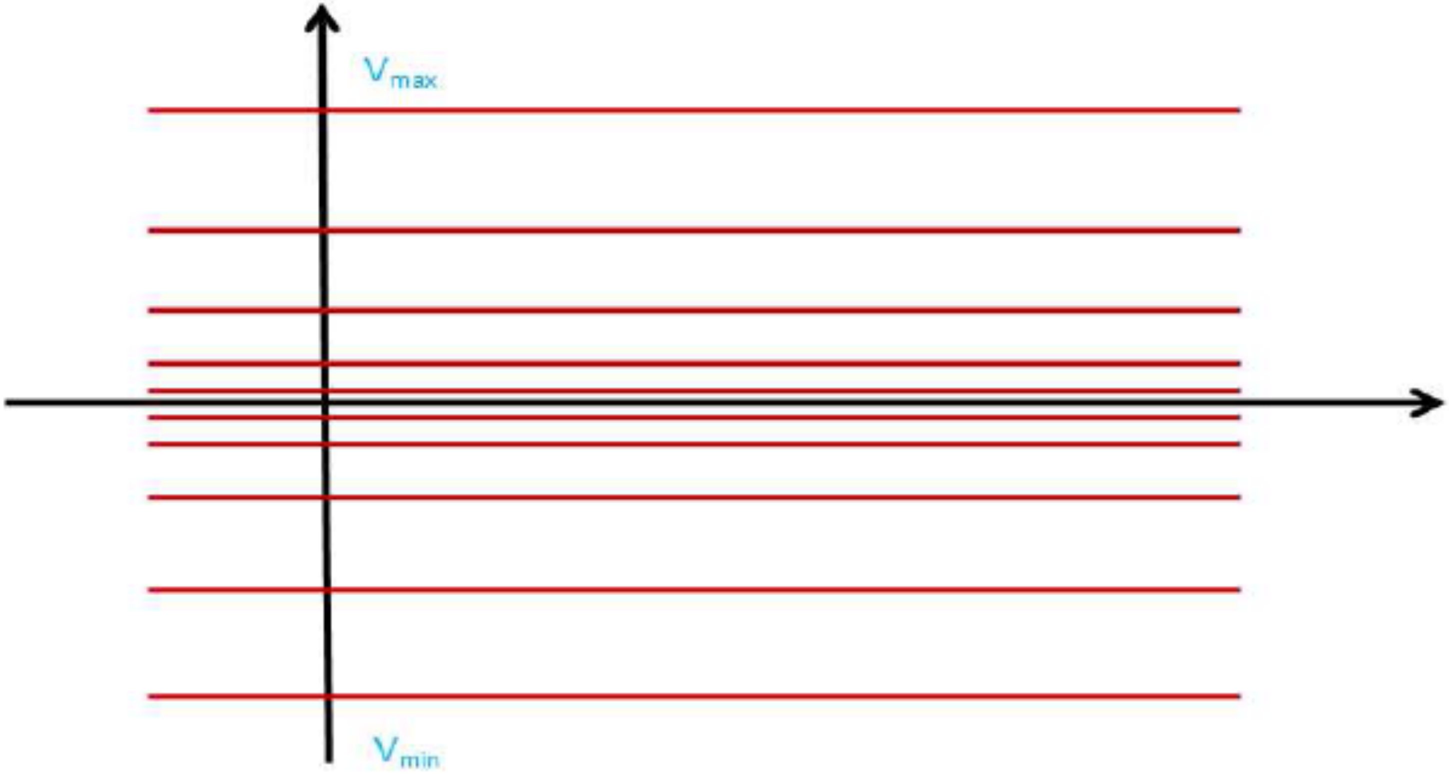
3-bit resolution with eight levels

Bit rate

- Bit rate
  - number of bits per second
  - product between the sampling frequency and the number of quantization bits

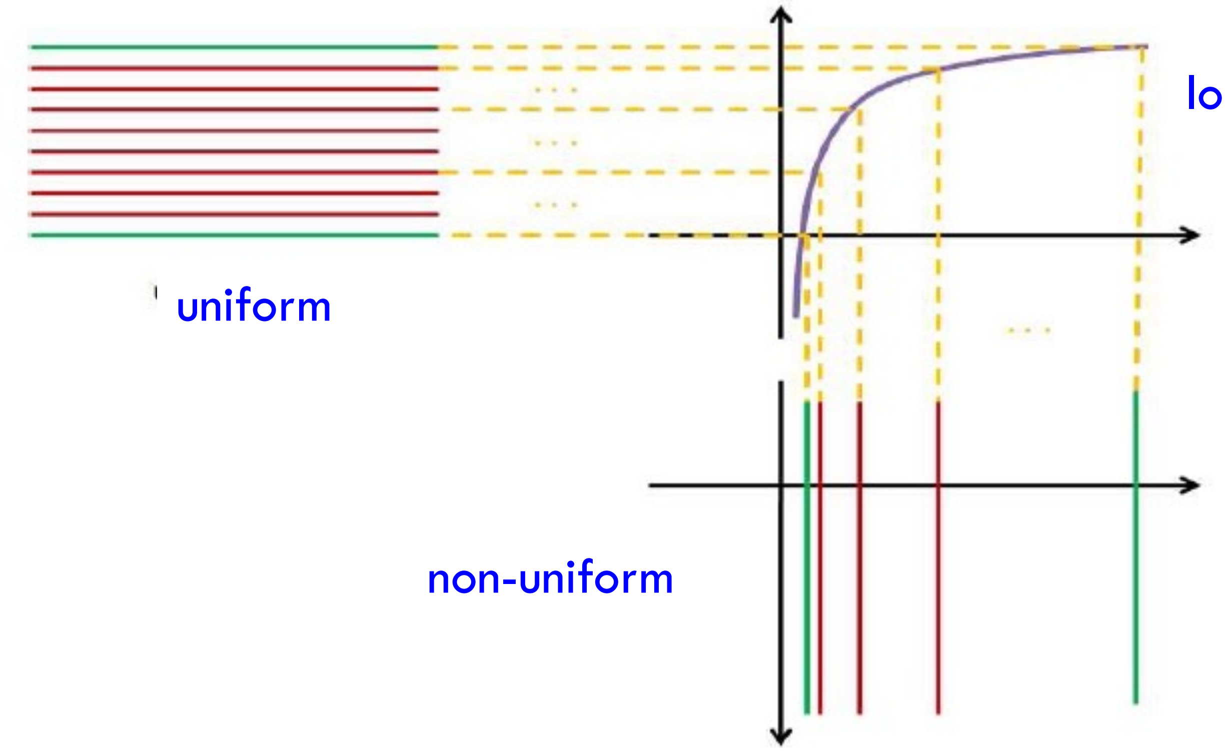
$$\text{bit rate } = f_c \cdot N$$

Non-linear quantization



non-uniform quantization

Logarithmic quantization

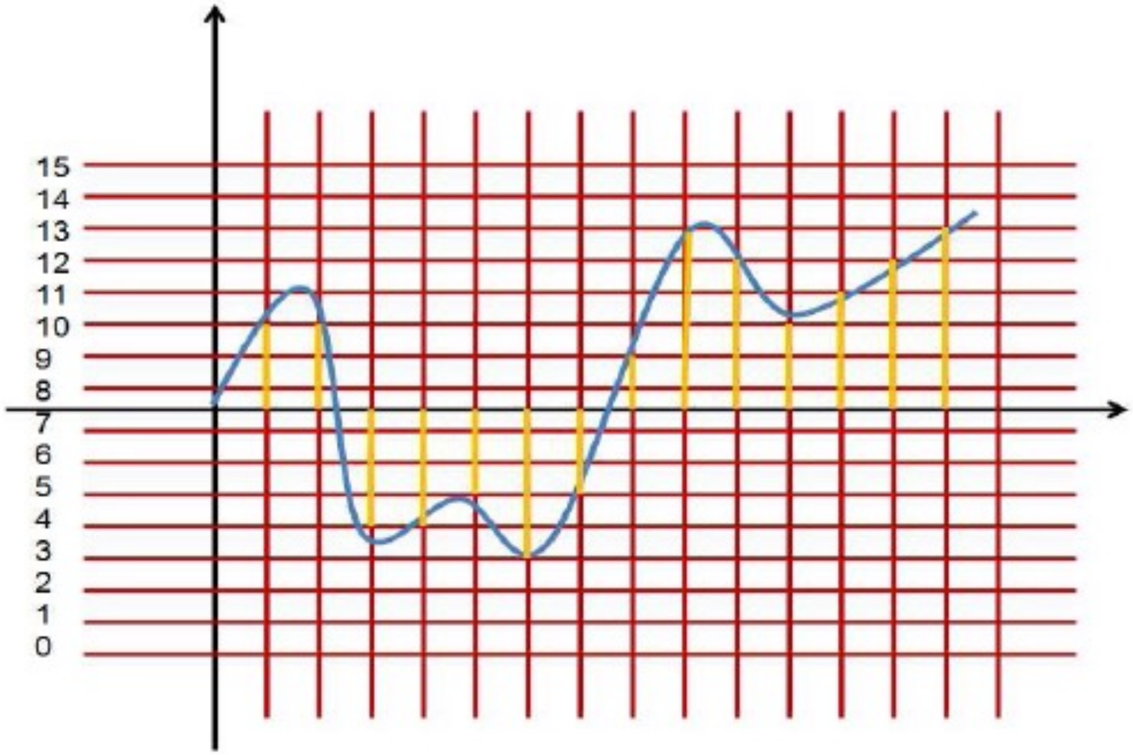


logarithm

uniform

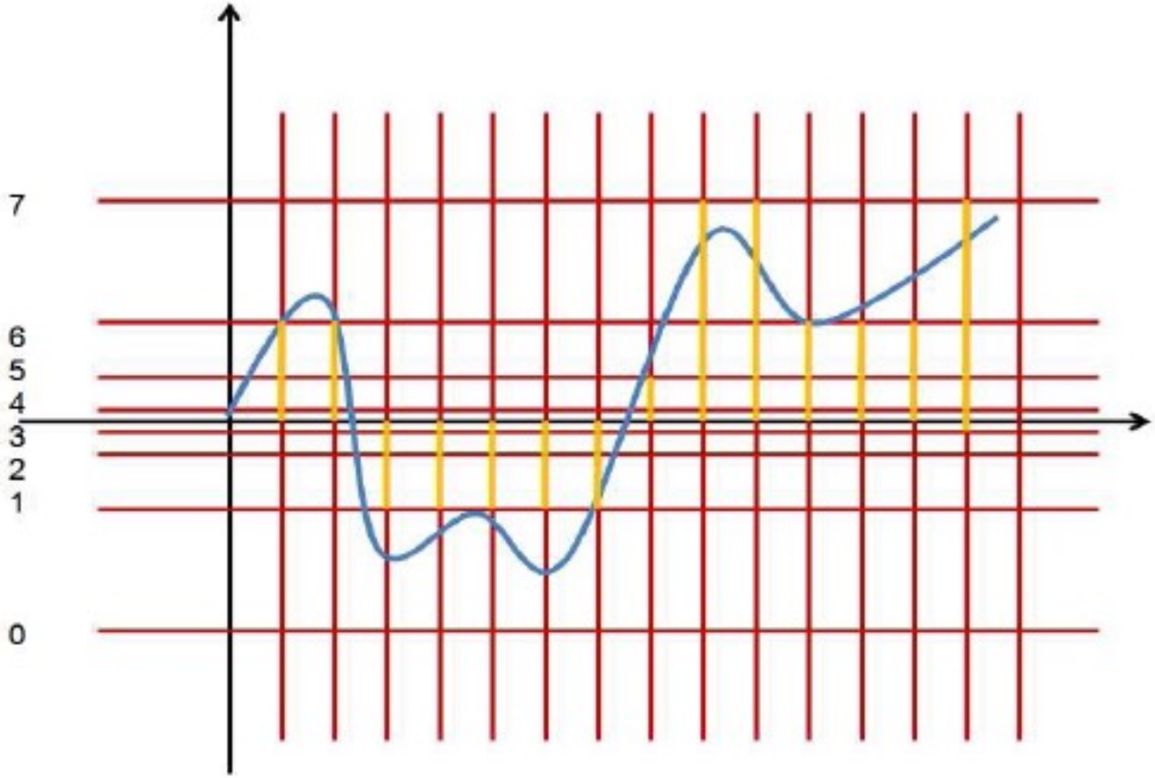
non-uniform

Example



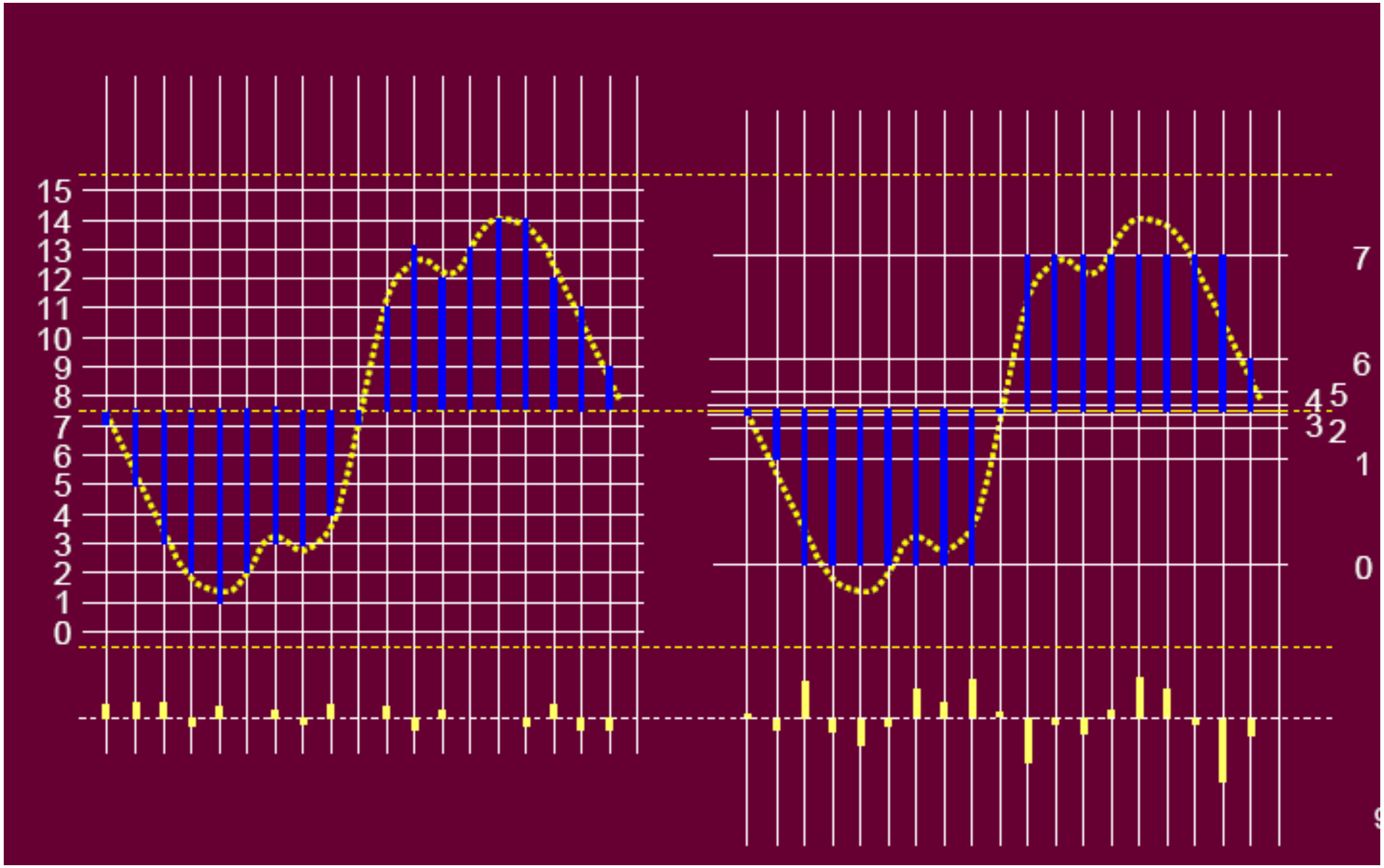
Uniform quantization

Example



non-uniform quantization

Comparison



4 bit linear

3 bit logarithmic

Comparison

- Quality
  - the dynamic range of the 8-bit logarithmic quantization corresponds to 13-14 bits linear quantizer
- Signal to Noise Rate (SNR)
  - a 8-bit logarithmic converter is better than a 8-bit linear at low amplitudes, but worse at high amplitudes

References

- Material
  - Slides
  - Video Lessons
- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 5

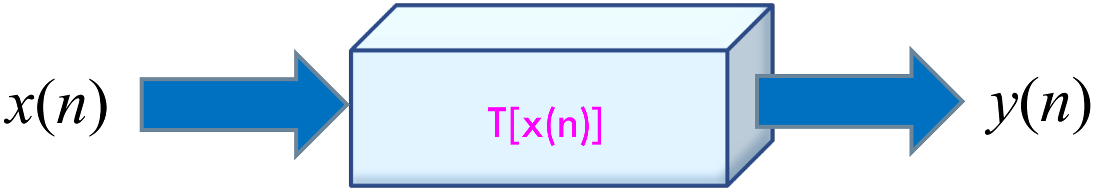
- Automated controls play an essential role in the technological progress of human civilization
  - e.g., washing machines, refrigerators, ovens, automatic pilots of airplanes, robots, etc.
  - a real world problem can be described by a system

- Question
  - Describe the Linear Time-Invariant systems and the Impulse Response

Systems

- Mathematically a system is an
  - unique transformation mapping an input sequence x(*n*) into an output y(*n*)

$$\boxed{y(n) = T\left[x(n)\right]}$$

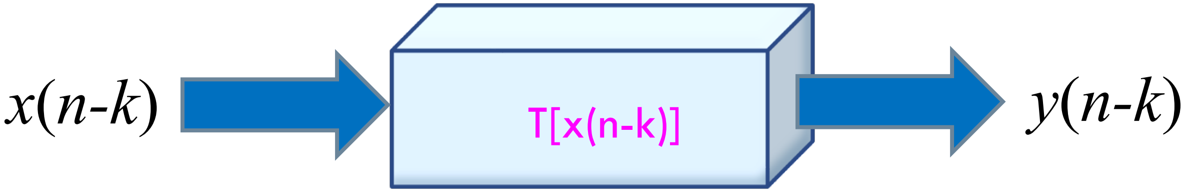


LTI systems

- Linear Time-Invariant (LTI) theory
  - comes from applied mathematics
  - has direct applications in
    - NMR spectroscopy, seismology, circuits, signal processing, control theory, and other technical areas
- It investigates the response of a linear and time-invariant system to an arbitrary input signal

Time-invariant systems

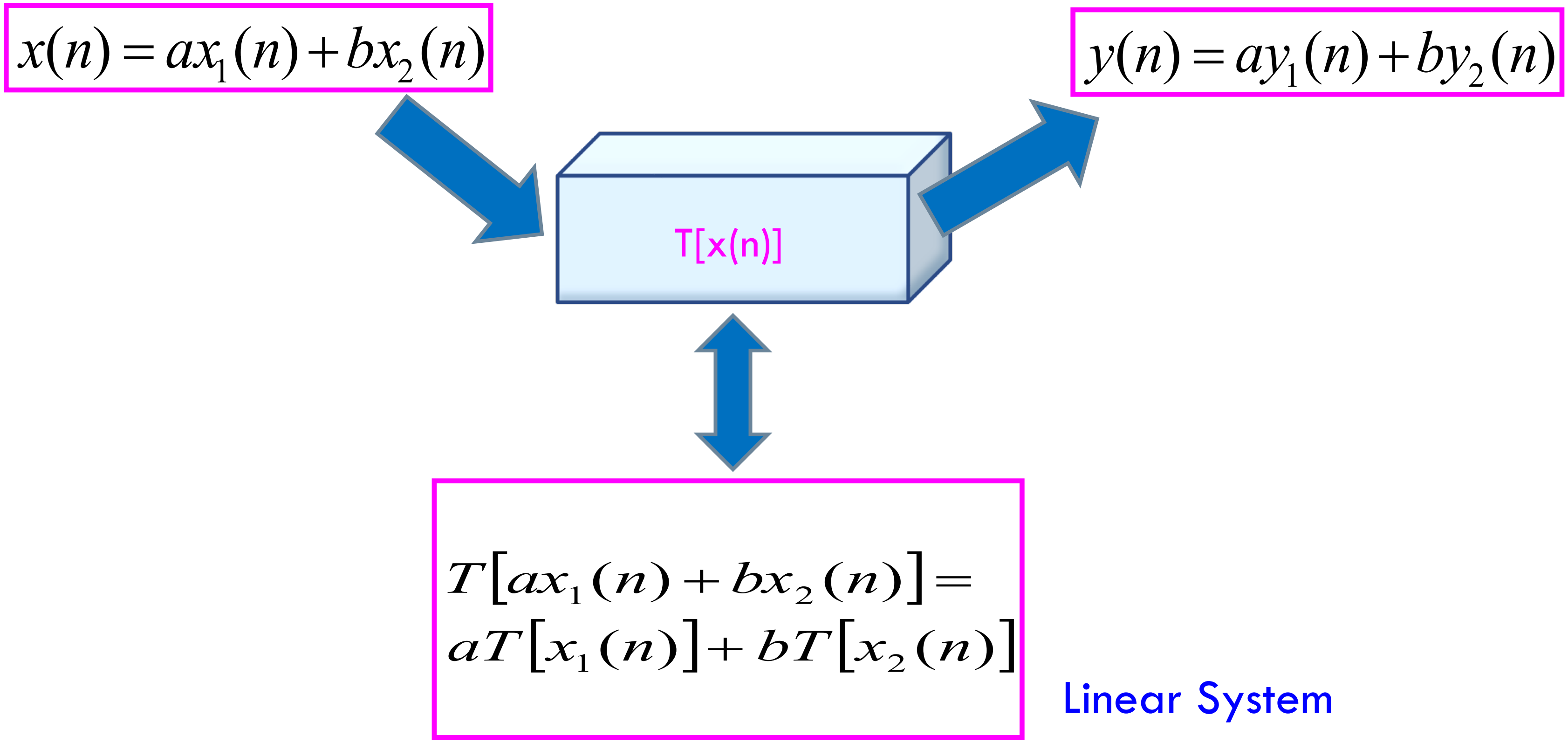
- Time-Invariant condition
  - If y(n) is the response to x(n) then y(n-k) is the response to x (n-k)
  - k is a positive or negative integer



Time-Invariant System

Linear Systems

- The class of Linear Systems is defined by the principle of superposition



$x(n) = ax_1(n) + bx_2(n)$

$y(n) = ay_1(n) + by_2(n)$

$$T[ax_1(n) + bx_2(n)] = \\ aT[x_1(n)] + bT[x_2(n)]$$

Linear System

Impulse Response

- A linear system can be completely characterized by its Impulse Response

Convolution

$$y(n) = T\left[x(n)\right] = T\left[\sum_{k=-\infty}^{\infty} x(k)\delta(n-k)\right]$$

$$y(n) = \sum_{k=-\infty}^{\infty} T\left[x(k)\delta(n-k)\right] = \sum_{k=-\infty}^{\infty} x(k)T\left[\delta(n-k)\right] = \sum_{k=-\infty}^{\infty} x(k)h_k(n)$$

Impulse Response, $h_k(n) = h(k-n)$

Impulse Response

- The convolution operation is denoted as

$$\boxed{y(n) = x(n) * h(n)}$$

- Equivalently we can write

$$\boxed{y(n) = \sum_{k=-\infty}^{\infty} h(k)x(n-k) = h(n) * x(n)}$$

References

- Material
  - Slides
  - Video Lessons

- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 6

- Question
  - Describe the Fourier Theorem and the Discrete Time Fourier Transform

Fourier theorem

- Theorem

any continuous periodic signal can be obtained by the superposition of simple sine waves, each with its amplitude and phase, and whose frequencies are harmonics of the fundamental frequency of the signal

Continuous periodic signal with period T_0

$$x(t) = A_0 + \sum_{k=1}^{\infty} A_k \cos(2\pi k f_0 t + \theta_k) = \boxed{A_0 + \sum_{k=1}^{\infty} A_k \cos(\omega_0 k t + \theta_k)}$$

$$f_0 = 1/T_0 \qquad \omega_0 = 2\pi f_0$$

Fourier series

Fourier transform

- Real environments
  - non-periodic signals

Inverse Fourier Transform
(Analysis)

$$x(t) = \int_{-\infty}^{\infty} X(\omega) e^{j\omega t} d\omega = \int_{-\infty}^{\infty} X(f) e^{j2\pi f t} df.$$

$$X(f)df = A_k$$

$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt \qquad X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi f t} dt.$$

Direct Fourier Transform
(Syntesis)

ISP – Verification tests

Discrete Time Fourier Transform

- For a discrete sequence x(*n*)

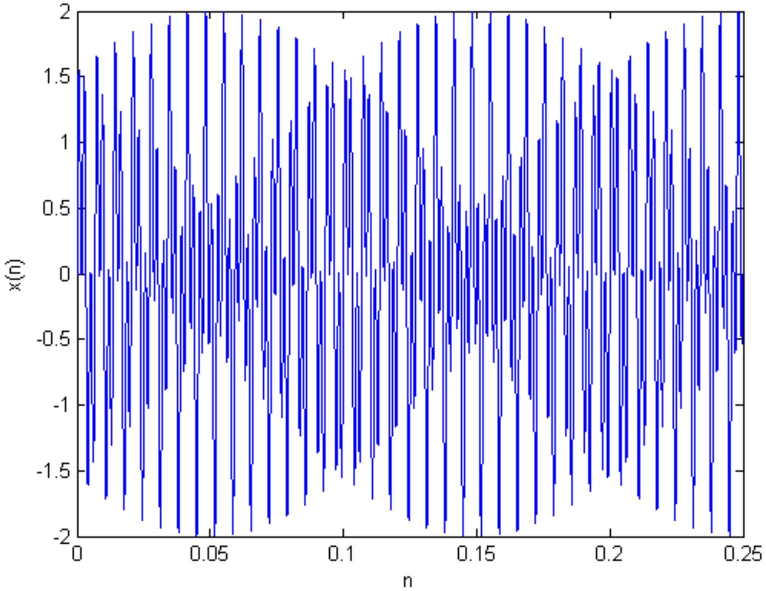
$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}$$

DTFT

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

Trasformata di Fourier
a tempo discreto inversa

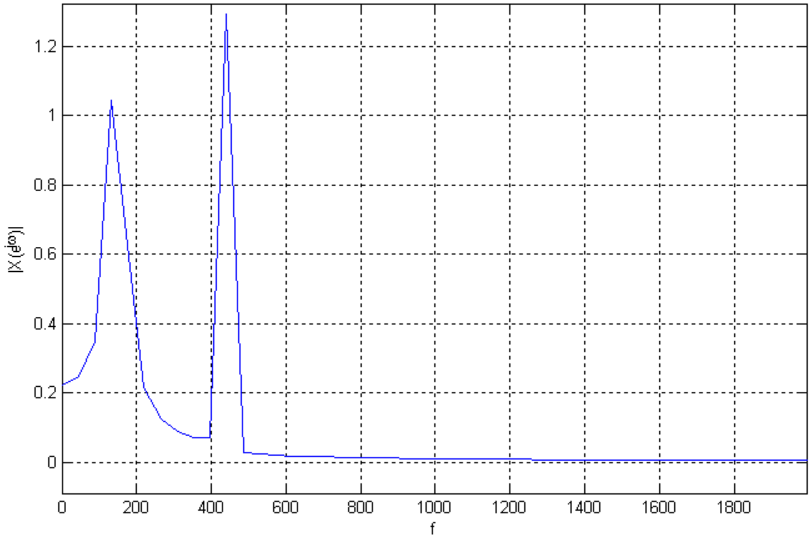
DTFT example



Sequence with two pure tones with 150 and 440 Hz, respectively. Sampling frquence of 44100 Hz.

Absolute value of the DTFT

$0 \leq \omega \leq \pi \rightarrow 0 \leq f \leq 22050$



Convolution theorem

$$x(n) \xleftrightarrow[\text{Transform}]{\mathcal{F}} X\left(e^{j\omega}\right)$$

$$h(n) \xleftrightarrow{\mathcal{F}} H\left(e^{j\omega}\right)$$

$$\boxed{y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) = x(n) * h(n)}$$

$$\Downarrow$$

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

- Highlighting
  - A convolution in the time domain corresponds to a product in the frequency domain

Modulation theorem (windowing)

$$x(n) \xleftrightarrow{\mathcal{F}} X\left(e^{j\omega}\right)$$

$$w(n) \xleftrightarrow{\mathcal{F}} W\left(e^{j\omega}\right)$$

$$\boxed{y(n) = x(n)w(n)}$$

$$\Downarrow$$

$$\boxed{Y(e^{j\omega}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\theta})W(e^{j(\omega-\theta)})d\theta}$$

- Highlighting
  - The DTFT of the product of sequence corresopnds to a periodic convolution of the single DTFTs
  - e.g., FIR con windowing

Teorema di Parseval

$$x(n) \overset{\mathcal{F}}{\Longleftrightarrow} X\left(e^{j\omega}\right)$$

- Energy

$$\boxed{E = \sum_{n=-\infty}^{\infty} \left|x(n)\right|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|X(e^{j\omega})\right|^2 d\omega}$$

- $|X(e^{i\omega})|^2$ Energy Spectral Density

References

- Material
  - Slides
  - Video Lessons
- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 7

- Question
  - Describe the z Transform and its properties

z Transform

- For continuous time systems
  - The Laplace Transform can be considered a generalization of the Fourier Transform

- For Discrete time systems
  - The z Transform can be considered a generalization of the Discrete Time Fourier Transform

z Transform

- z Transform
  - use a generic complex number
  - when $z = e^{j\omega}$ a DTFT is obtained
  - contains further details on the nature of the signal
- The z Transform (bilateral) of a sequence *x(n)*

$$\boxed{X(z) = \sum_{n=-\infty}^{\infty} x(n) z^{-n}}$$

$$\boxed{X(z) = \sum_{n=0}^{\infty} x(n) z^{-n}}$$

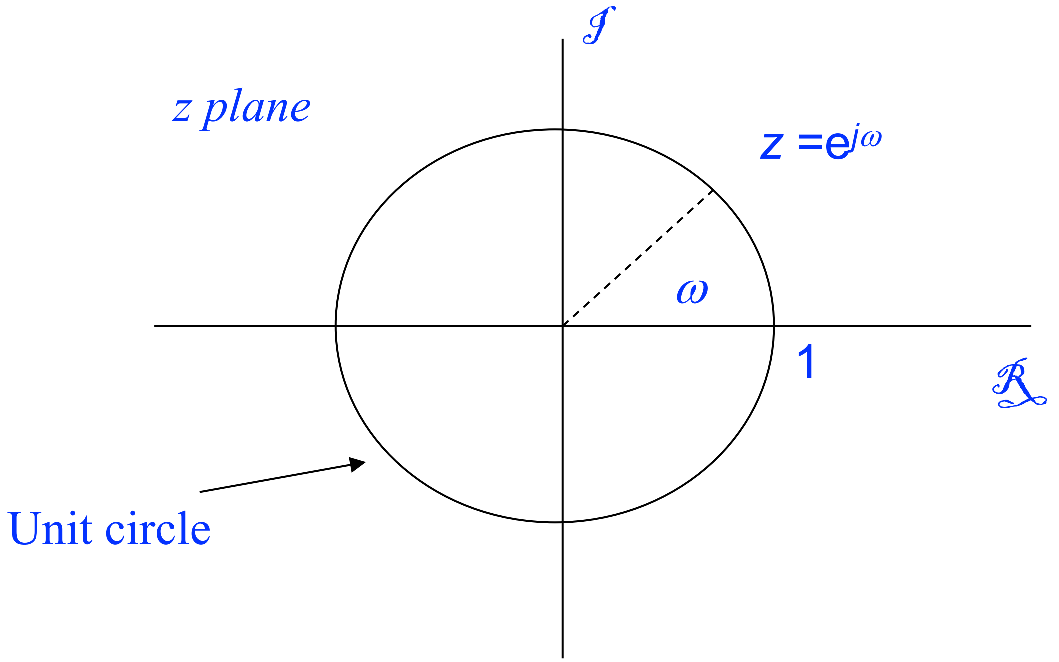
right unilateral

z Transform and DTFT

- Setting $z = re^{j\omega}$ we obtain

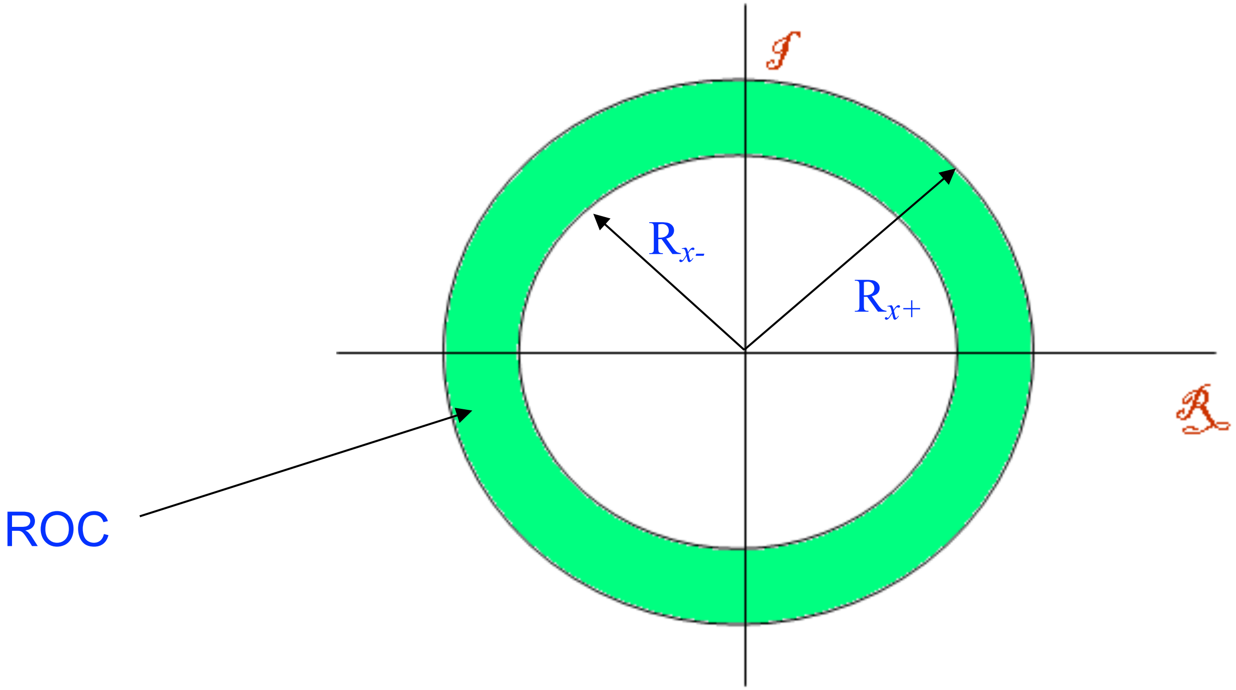
$$X(re^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)\left(re^{j\omega}\right)^{-n} = \sum_{n=-\infty}^{\infty} x(n)r^{-n}e^{-j\omega n}$$

- For $r = 1$ $(|z| = 1)$ the z Transform becomes the DTFT



Region of Convergence

- Given a sequence x(n) the set of z values for which the z Transform converges is named
  - Region of Convergence



Region of Convergence

- Properties
  - The outer boundary is a circle or can be extended to infinity
  - The inner border is a circle and can be extended to become the origin

- If the ROC
  - includes the unit circle, this implies convergence of the z-transform also the Fourier transform converges
  - does not include the unit circle the Fourier transform it is not absolutely convergent

Zeros and poles

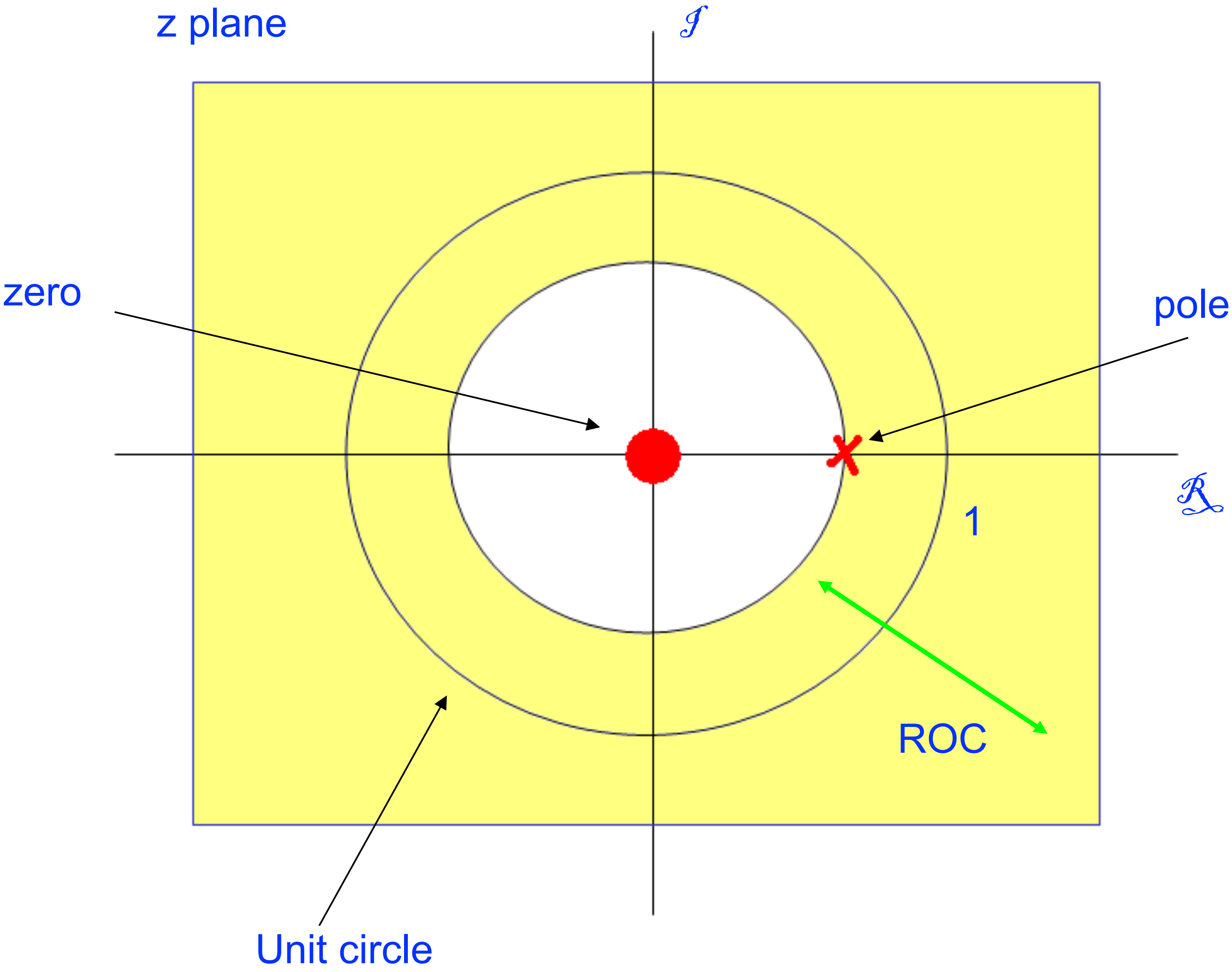
- An important class is the rational function (polynomials ratio in z)

$$\boxed{X(z)=\frac{P(z)}{Q(z)}}$$

- the zeros of the system are roots of the numerator polynomial

- the poles of the system are roots of the denominator polynomial

Example of ROC



Convergence properties

- The ROC must be a connected region
- The ROC is a ring or a disc in the z-plane centered at the origin
- The Fourier Transform of x(n) converges absolutely if and only if the ROC of the z Transform of x(n) comprises the unit circle
- The ROC does not contain any pole and is bounded by poles or zeros or infinite
- If x(n) is a sequence of finite duration the ROC is the entire z-plane except for possible z = 0 and z = +∞
- If x (n) is the monolateral right the ROC is the outside of a circle (pole amplitude increased up to (possibly) +∞)
- If x (n) is monolateral left the ROC is the inside of a circle (pole different from zero with lower amplitude up (possibly) to 0)
- If x (n) is the two-sided ROC consists of a ring in the z plane limited from the inside and from the outside by a pole and in accordance with the property 3 contains no pole

References

- Material
  - Slides
  - Video Lessons

- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 8

- Question
  - Describe the Discrete Fourier Transform and the Discrete Cosine Transform

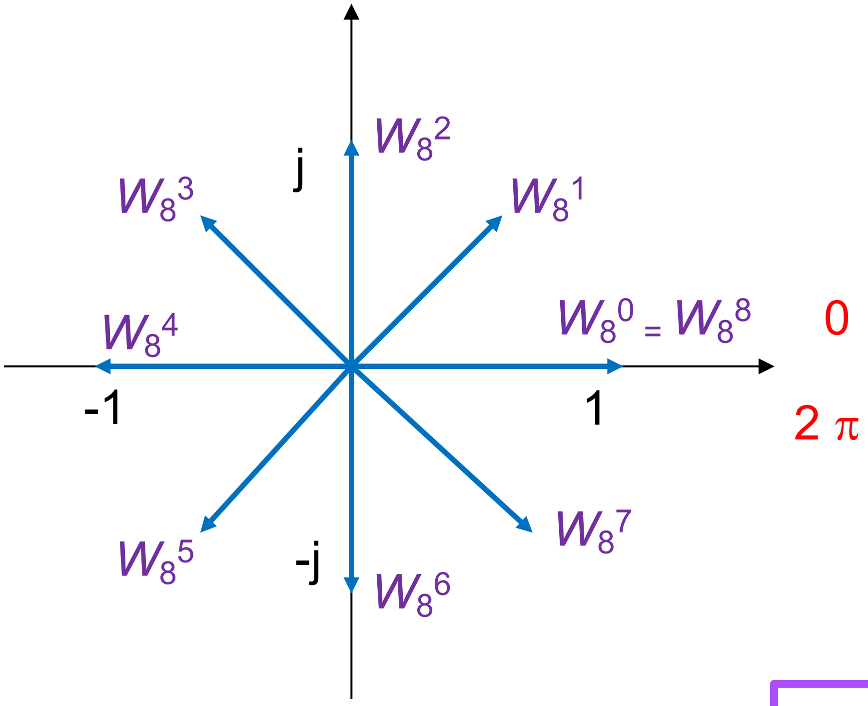
Discrete Fourier Transform

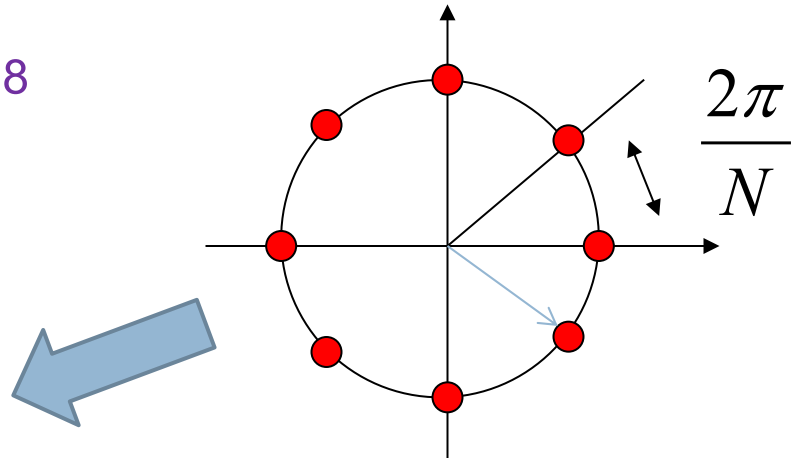
- Continuous time transforms
  - Fourier Theorem
  - Continuous Fourier Transform
  - Discrete Time Fourier Transform (DTFT)
  - z-transform

- Transformation for finite duration sequences
  - Discrete Fourier Transform (DFT)
  - Discrete Cosine Transform (DCT)

Roots of the unit circle

N=8





$$\boxed{W_N = e^{-j\frac{2\pi}{N}}}$$

The N-th root of the unity

$$\boxed{W_N^k = e^{-j\frac{2\pi}{N}k}}$$

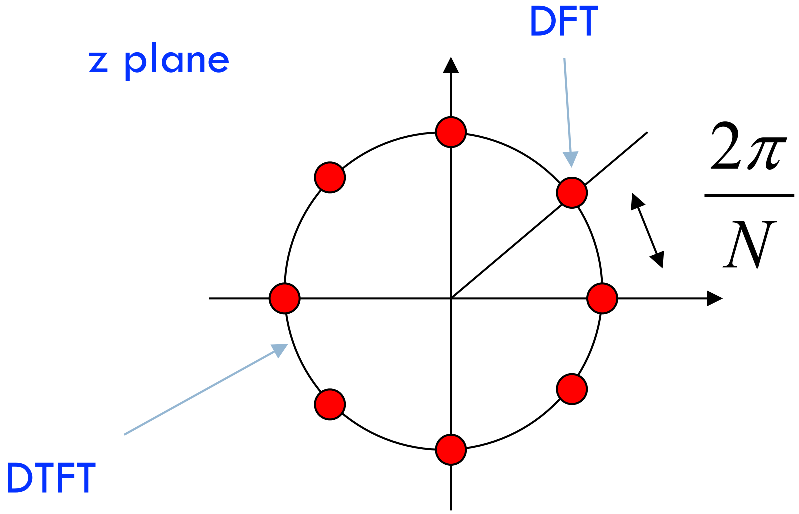
N roots of the unity

$$0 \leq k \ \leq N - 1$$

ISP – Verification tests

DFT and z-Transform

- It corresponds to sample the z-transform, X (z), in N points equally spaced on the unit circle



DFT

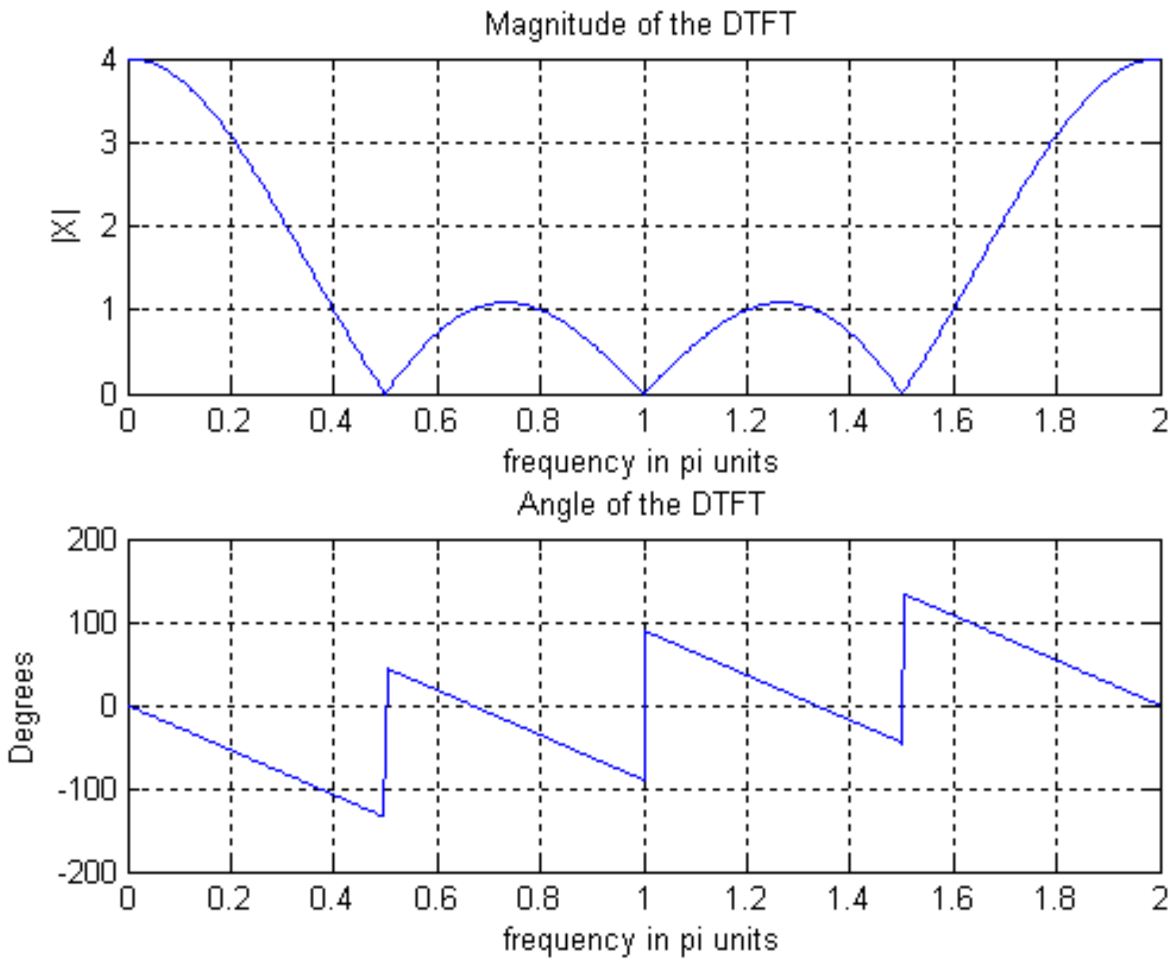
$$X(k) = \begin{cases} \sum_{n=0}^{N-1} x(n) W_N^{kn} & 0 \le k \le N-1 \\ 0 & \text{altrove} \end{cases}$$

Analysis

$$x(n) = \begin{cases} \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{-kn} & 0 \le n \le N-1 \\ 0 & \text{altrove} \end{cases}$$

Synthesis

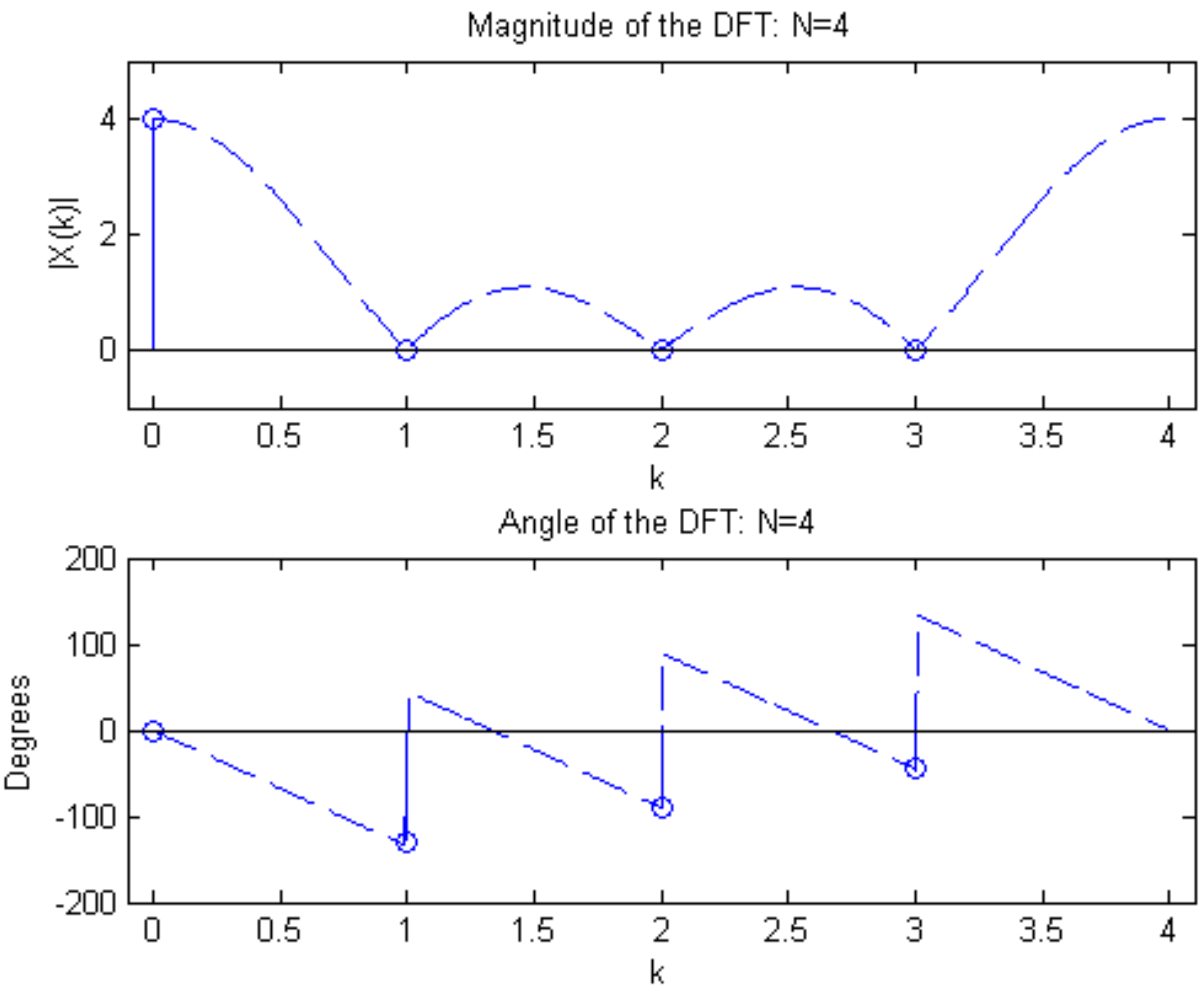
Example of DFT



DTFT of a 4 points sequence

$$x(n) = \begin{cases} 1 & 0 \le n \le 3 \\ 0 & \text{altrove} \end{cases}$$

Example of DFT



DFT of a 4 points sequence. The DFT is a sampling of the DTFT

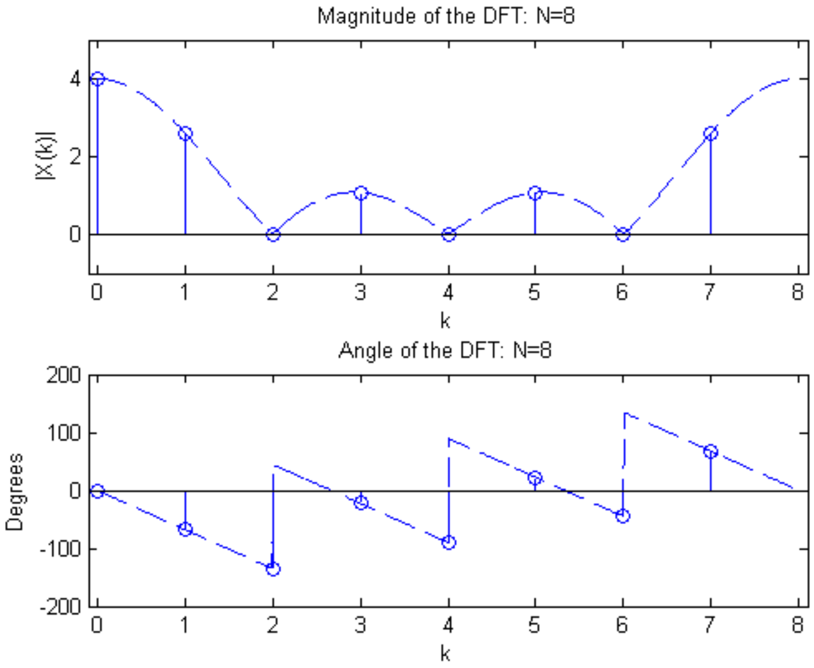
Padding

- We add some zeros to the previous sequence

$$x(n)=[1\ \ 1\ \ 1\ \ 1\ \ 0\ 0\ 0\ 0]$$

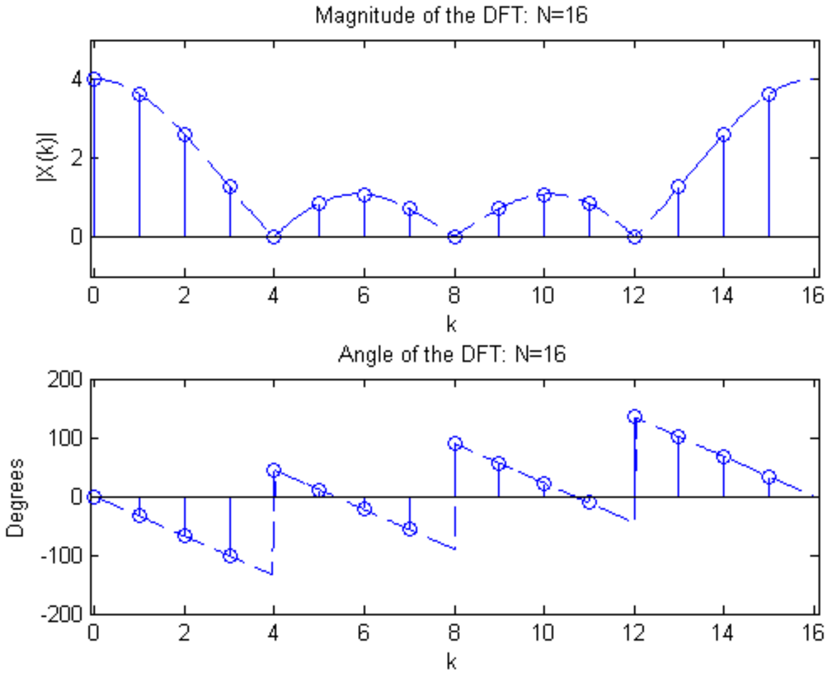
- This operation is named zero-padding
- It is needed to obtain a dense spectrum

DFT


Magnitude of the DFT: N=8

Angle of the DFT: N=8

DFT after zero-padding

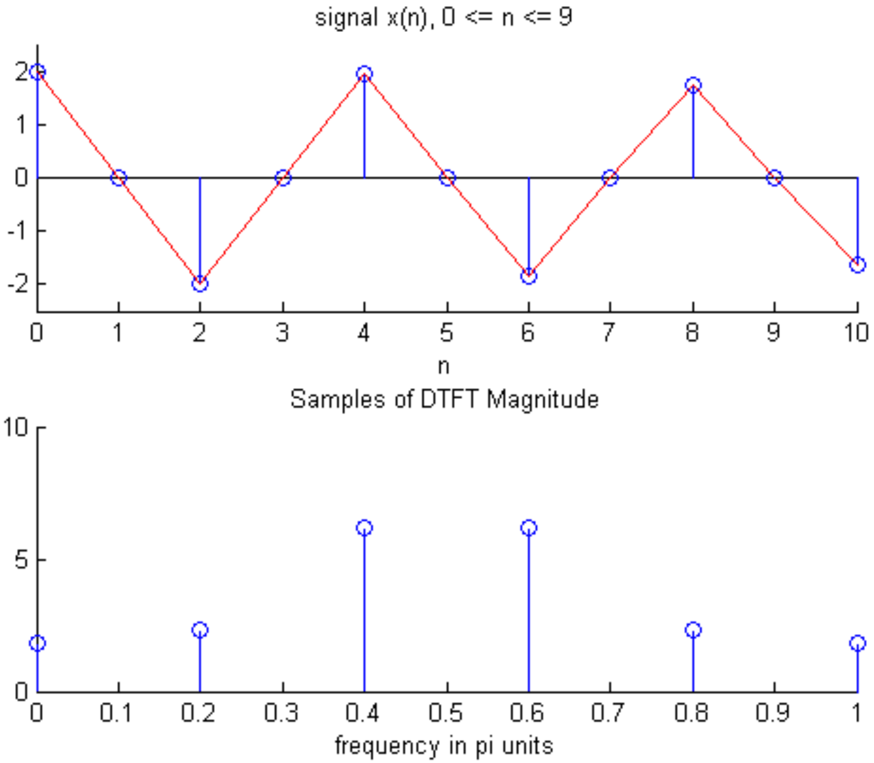

Magnitude of the DFT: N=16

Angle of the DFT: N=16

High density spectrum

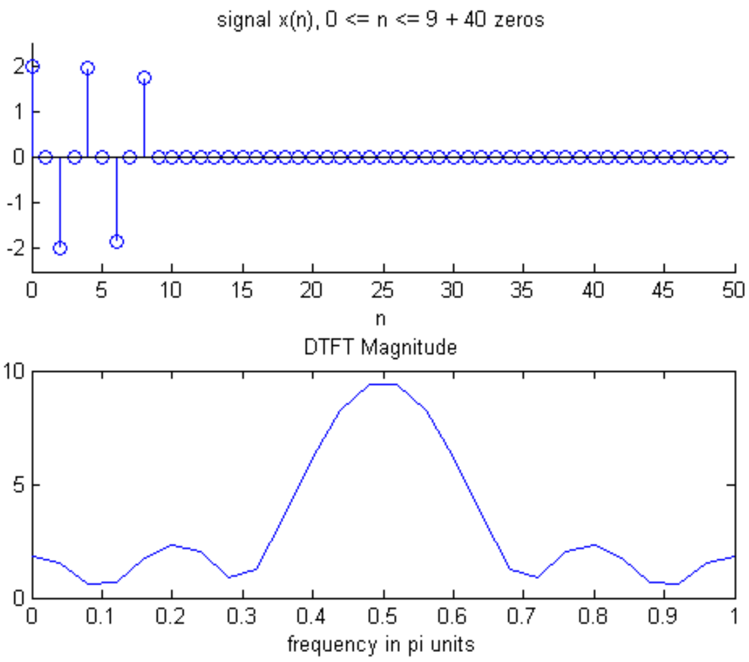
- We consider the following signal

$$x(n) = \cos(0.48\pi n) + \cos(0.52\pi n)$$

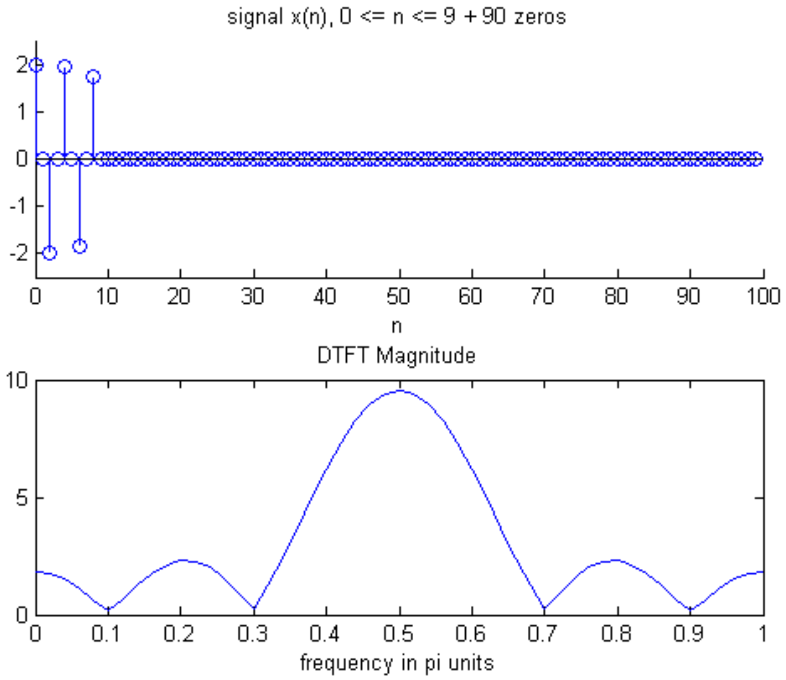


DFT on 10 points

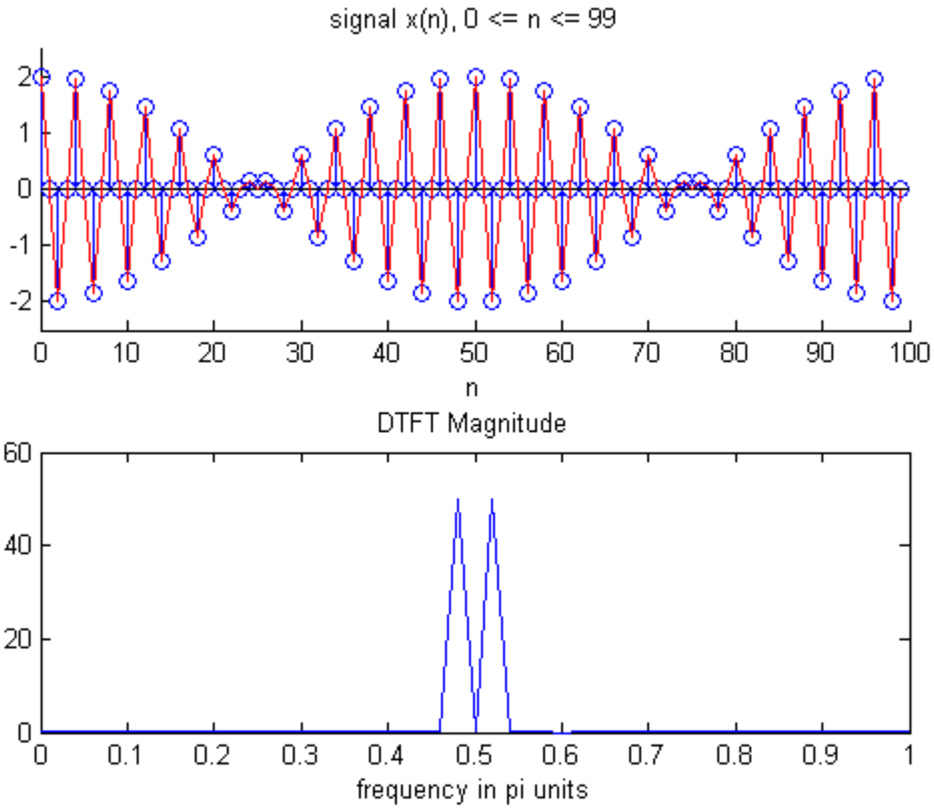
High density spectrum



Increasing the points ...



High resolution spectrum



To obtain an High Resolution Spectrum we increse the sampling points of the source sequence. The estimated frequencies corrispond to the frequencies of the analyzed signal.

Discrete Cosine Transform

- The Discrete Cosine Transform (DCT) is very similar to DFT but in real domain

- DCT is used for feature extraction
  - data decorrelation
  - low loss compression
  - the basis functions are orthogonal
  - it is symmetric

1D DCT

$0 \leq k \leq$ N-1

monodimensional signal

$$X(k) = \alpha(k) \sum_{n=0}^{N-1} x(n) \cos\left(\frac{\pi(2n+1)k}{2N} \right)$$

Analysis

$$x(n) = \sum_{k=0}^{N-1} \alpha(k) X(k) \cos\left(\frac{\pi(2n+1)k}{2N} \right)$$

Synthesis

$$\alpha(k) = \begin{cases} \sqrt{\dfrac{1}{N}} & \text{se} \quad k = 0 \\ \sqrt{\dfrac{2}{N}} & \text{se} \quad k \neq 0 \end{cases}$$

1D DCT

$0 \leq n \leq$ N-1
$0 \leq m \leq$ N-1

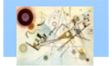
bidimensional signal

Analysis

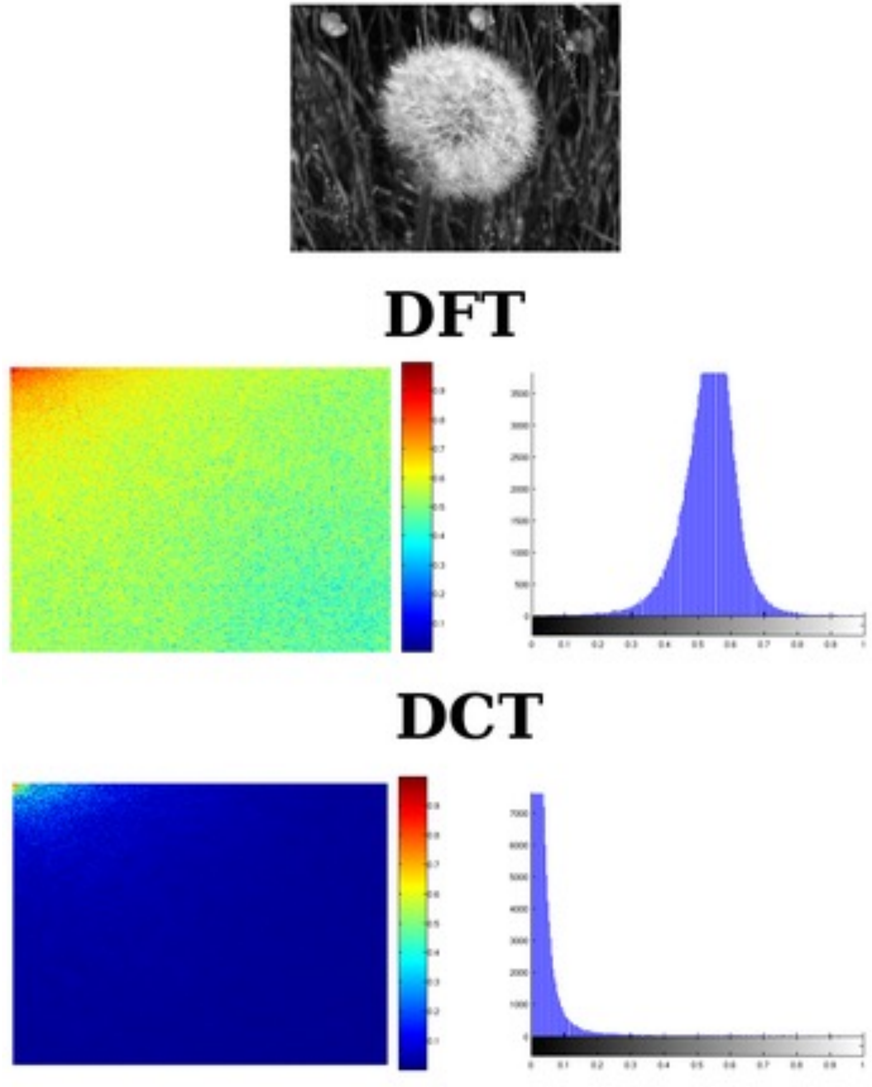
$$X(n,m) = \alpha(n)\alpha(m)\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} x(i,j)\cos\left(\frac{\pi(2i+1)n}{2N}\right)\cos\left(\frac{\pi(2j+1)m}{2N}\right)$$

Synthesis

$$x(i,j) = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} \alpha(n)\alpha(m)X(n,m)\cos\left(\frac{\pi(2i+1)n}{2N}\right)\cos\left(\frac{\pi(2j+1)m}{2N}\right)$$



DFT vs DCT



DCT provides the spatial compression, able to detect changes of information between contiguos area avoiding the repetitions

References

- Material
  - Slides
  - Video Lessons
- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

ISP – Verification tests

Question 9

- Question
  - Describe the Fast Fourier Transform algorithm

ISP – Verification tests

Introduction

- The Discrete Fourier Transform (DFT) has an important role for signal analysis

- In the sixties of the last century a fast approach for DFT was introduced by Cooley and Tukey
    - Fast Fourier Transform

Classis

- Decimation in time
  - The source signal x(n) is divided in shorter sequences

- Decimation in frequency
  - The DFT coefficients X(k) are divided in shorter sequences

DFT

$$X(k)=\begin{cases}\sum_{n=0}^{N-1} x(n)W_N^{kn} & 0\le k\le N-1\\ 0 & \text{altrove}\end{cases}$$

Analysis

$$x(n)=\begin{cases}\frac{1}{N}\sum_{k=0}^{N-1} X\left(k\right)W_N^{-kn} & 0\le n\le N-1\\ 0 & \text{altrove}\end{cases}$$

Synthesis

$$X(k)=\sum_{n=0}^{N-1}\left\{\left(\mathrm{Re}[x(n)]\mathrm{Re}\left[W_N^{kn}\right]-\mathrm{Im}[x(n)]\mathrm{Im}\left[W_N^{kn}\right]\right)+j\left(\mathrm{Re}[x(n)]\mathrm{Re}\left[W_N^{kn}\right]+\mathrm{Im}[x(n)]\mathrm{Im}\left[W_N^{kn}\right]\right)\right\}$$

$$k=0,1,\ldots,N-1$$

DFT

$$X(k)=\begin{cases}\sum_{n=0}^{N-1} x(n)W_N^{kn} & 0\le k\le N-1\\ 0 & \text{altrove}\end{cases}$$

Analysis

$$x(n)=\begin{cases}\frac{1}{N}\sum_{k=0}^{N-1} X(k)W_N^{-kn} & 0\le n\le N-1\\ 0 & \text{altrove}\end{cases}$$

Synthesis

$$X(k)=\sum_{n=0}^{N-1}\left\{\left(\mathrm{Re}[x(n)]\mathrm{Re}\left[W_N^{kn}\right]-\mathrm{Im}[x(n)]\mathrm{Im}\left[W_N^{kn}\right]\right)+j\left(\mathrm{Re}[x(n)]\mathrm{Re}\left[W_N^{kn}\right]+\mathrm{Im}[x(n)]\mathrm{Im}\left[W_N^{kn}\right]\right)\right\}$$

$$k=0,1,\ldots,N-1$$

X(*k*) needs of 4*N* real products and (4*N*-1) real sums for ech *k*. Totally, we have $4N^2$ real products e *N*(4*N*-1) real sums.

Time decimation

- We use the symmetry and periodicity of the complex exponential

$$\boxed{W_N^{kn} = e^{-j\left(\frac{2\pi}{N}\right)kn}}$$

- The sequence is a power of two

$$N = 2^v$$

Time decimation

- X(k) is calculated dividing x(n) in two subsequences

$$\boxed{X(k)=\sum_{n=0}^{N-1} x(n)e^{-j\left(\frac{2\pi}{N}\right)kn}}$$

$$\Downarrow$$

$$\boxed{X(k)=\sum_{\text{n even}}^{N-1} x(n)e^{-j\left(\frac{2\pi}{N}\right)kn}+\sum_{\text{n odd}}^{N-1} x(n)e^{-j\left(\frac{2\pi}{N}\right)kn}}$$

Time decimation

`n = 2r` `n = 2r+1`

$$X(k) = \sum_{\mathrm{r=0}}^{N/2-1} x(2r) W_N^{2rk} + \sum_{\mathrm{r=0}}^{N/2-1} x(2r+1) W_N^{(2r+1)k}$$

$$= \sum_{\mathrm{r=0}}^{N/2-1} x(2r)(W_N^2)^{rk} + W_N^k \sum_{\mathrm{r=0}}^{N/2-1} x(2r+1)(W_N^2)^{rk}$$

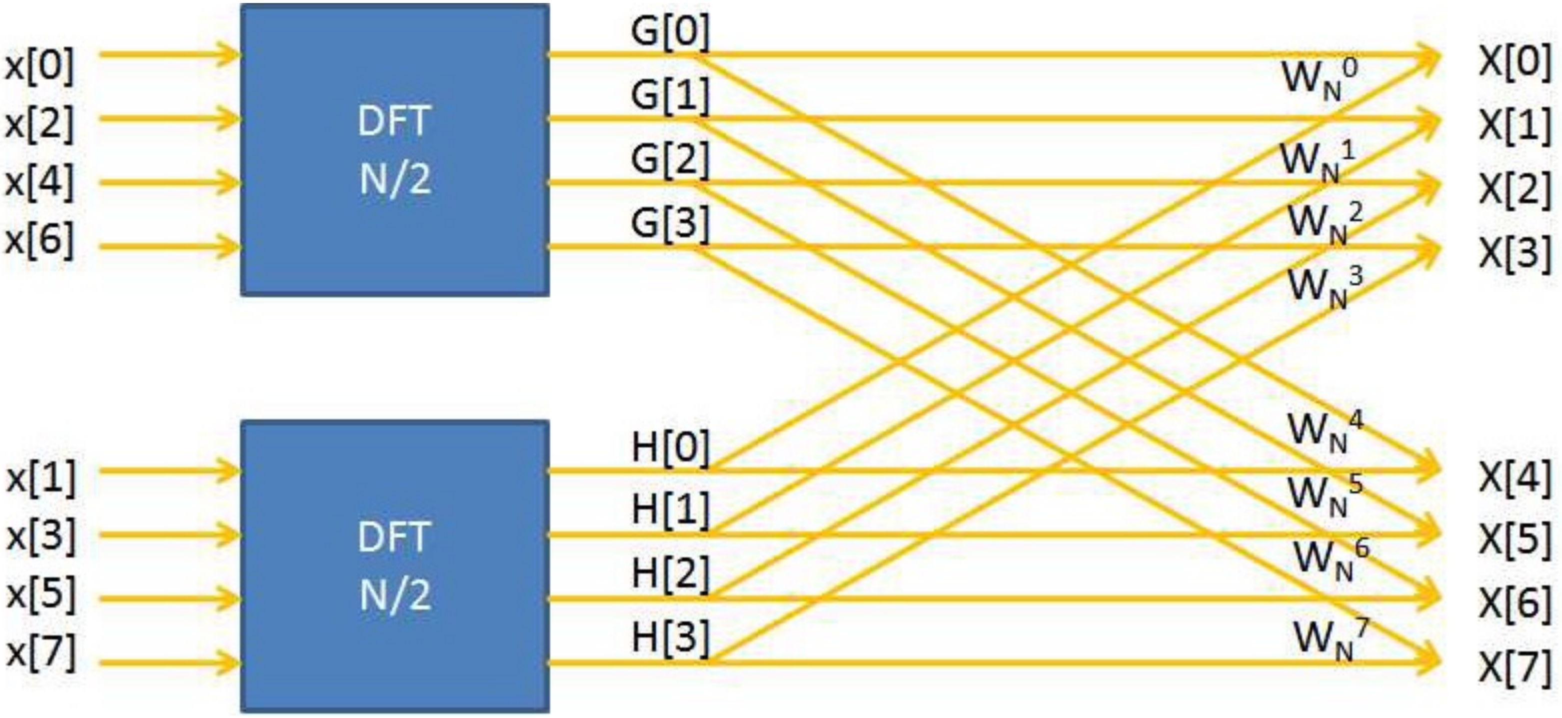
$$W_N^2 = e^{-2j(2\pi/N)} =$$

$$e^{-j2\pi/(N/2)} = W_{N/2}$$

⇩

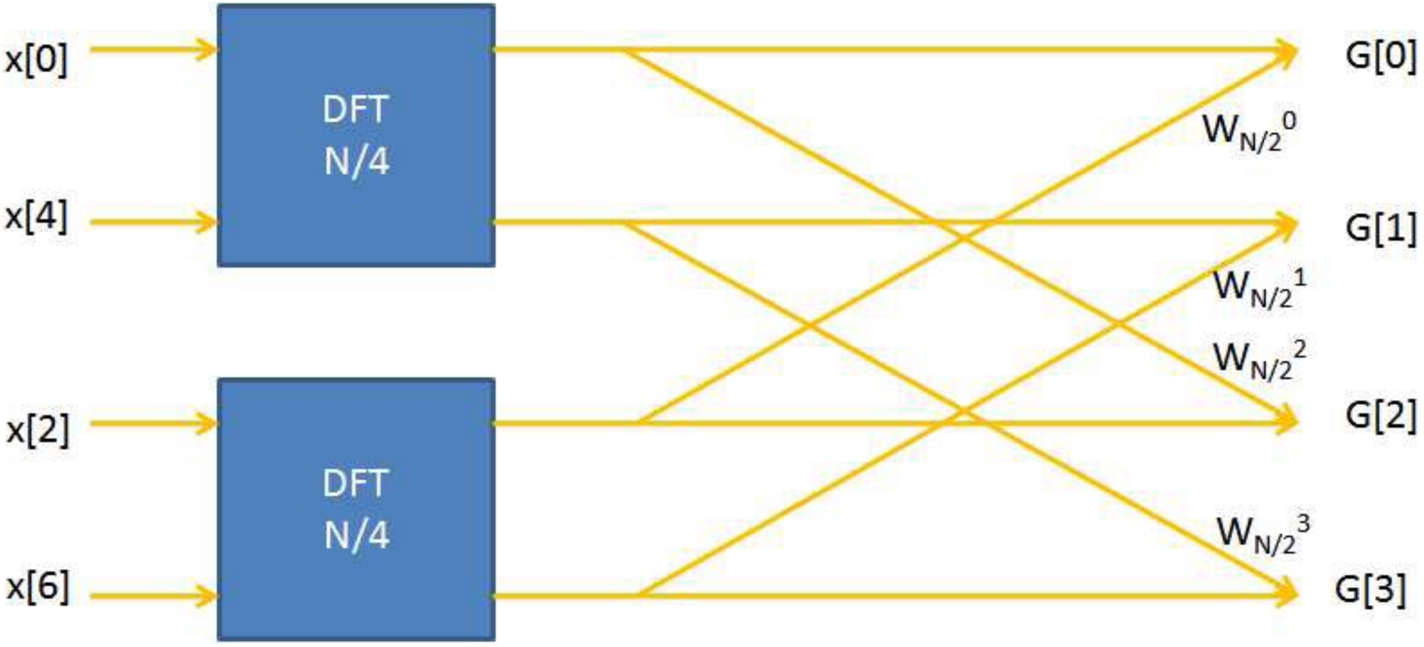
$$\boxed{\begin{aligned} X(k) &= \sum_{\mathrm{r=0}}^{N/2-1} x(2r) W_{N/2}^{rk} + W_N^k \sum_{\mathrm{r=0}}^{N/2-1} x(2r+1) W_{N/2}^{rk} \\ &= G(k) + W_N^k H(k) \end{aligned}}$$

Flow graph



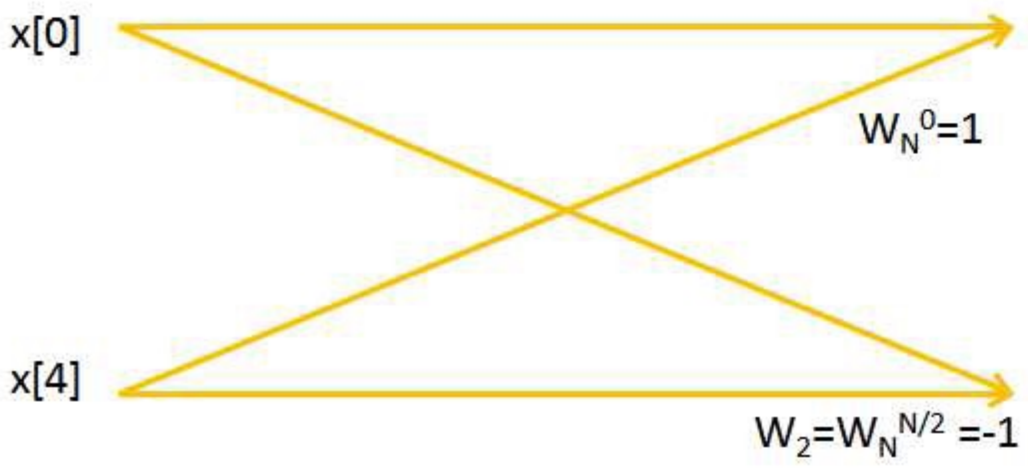
Flow Graph for a DFT with *N*=8

Flow graph



Flow Graph for a DFT with $N=4$

Flow graph



Flow Graph for a DFT with $N=2$

FFT algorithm

- The FFT is obtained by a recursive algorithm based on a divide-et-impera strategy

- Fourier coefficients

$$X[k] = \sum_{j=0}^{N/2-1} x[j] W_N^{kj}$$

$$\uparrow$$

$$x = (x[0], x[1], \ldots, x[N-1])$$

FFT algorithm

```
FFT-Ricorsiva(x)
   N = length(x);
   if N==1
      then return x[0]
   WN = exp(j  2 PI/N)
   W = 1
   Xp = FFT-Ricorsiva([x[0], x[2], x[4], . . . , x[N-2]])
   Xd = FFT-Ricorsiva([x[1], x[3], x[5], . . . , x[N-1]])
   for k = 0 to N/2 - 1
         do
            X[k] = Xp + W Xd
            X[k+N/2] = Xp - W Xd
            W = W W_N
     return X
```

Time complexity

- The asymptotic time complexity is

$$\boxed{T(N) = 2T(N/2) + \Theta(N) = \Theta(N \log N)}$$

- It is the same also for the inverse transform

Convolution theorem

- A faster convolution can be obtained

$$\boxed{a*b=\mathbf{DFT}_{2N}^{-1}\left(\mathbf{DFT}_{2N}(a)\mathbf{DFT}_{2N}(b)\right)}$$

zero padding

References

- Material
  - Slides
  - Video Lessons
- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 10

- Filter
  - device that increases or reduces the energy connected to certain regions of the spectrum sound
- Question
  - Describe the Finite Impulse Response (FIR) filters

Introduction

- Filter
  - device that increases or reduces the energy connected to certain regions of the spectrum sound
  - these operations are typically performed by equalizers
    - bank of bandpass filters
  - the critical bands of the auditory membrane are bandpass filters

Convolution

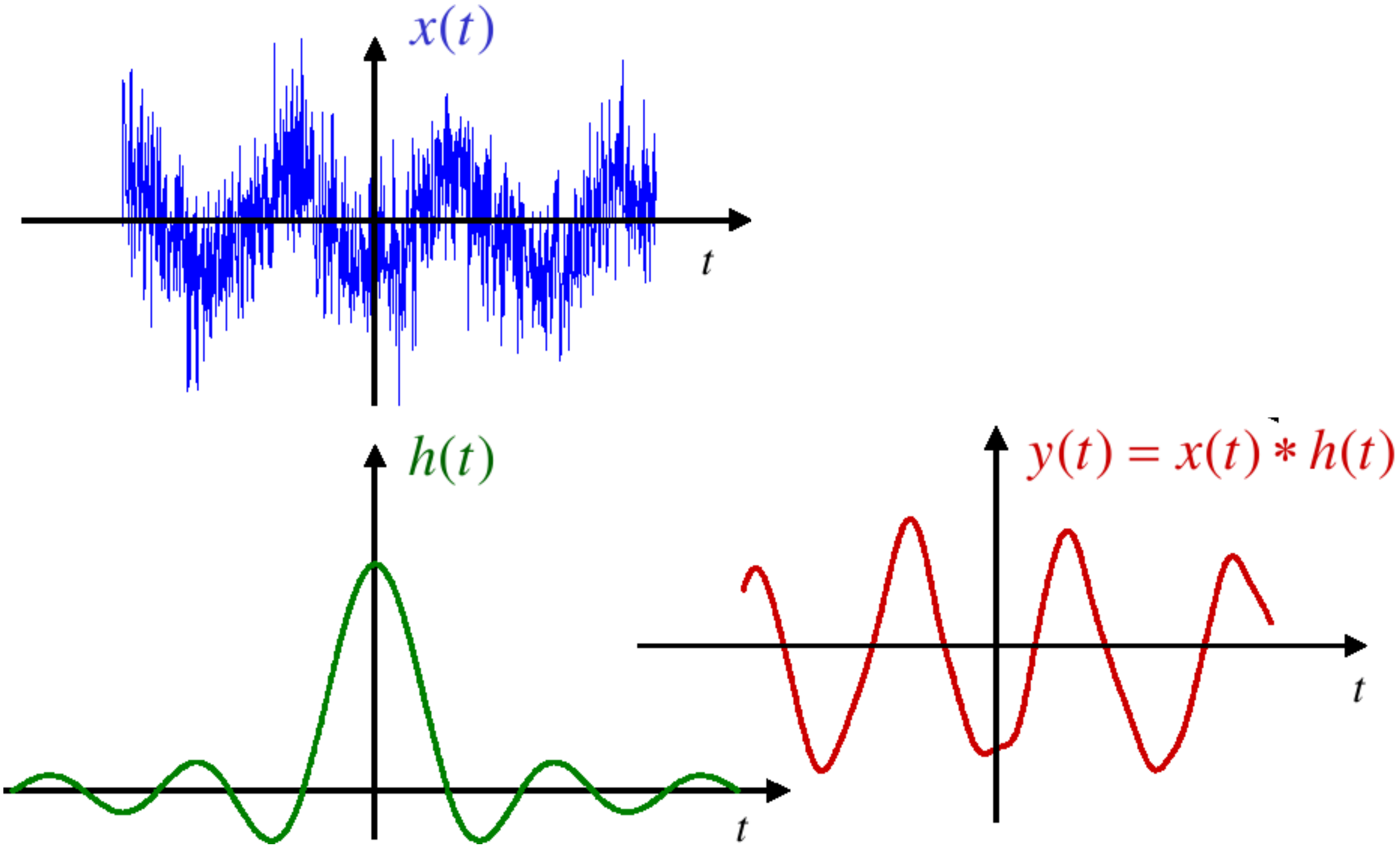
- LTI are characterized by the impulse response

$$\boxed{y(n)=\sum_{k=-\infty}^{\infty} x(k)h(n-k)}$$

convolution

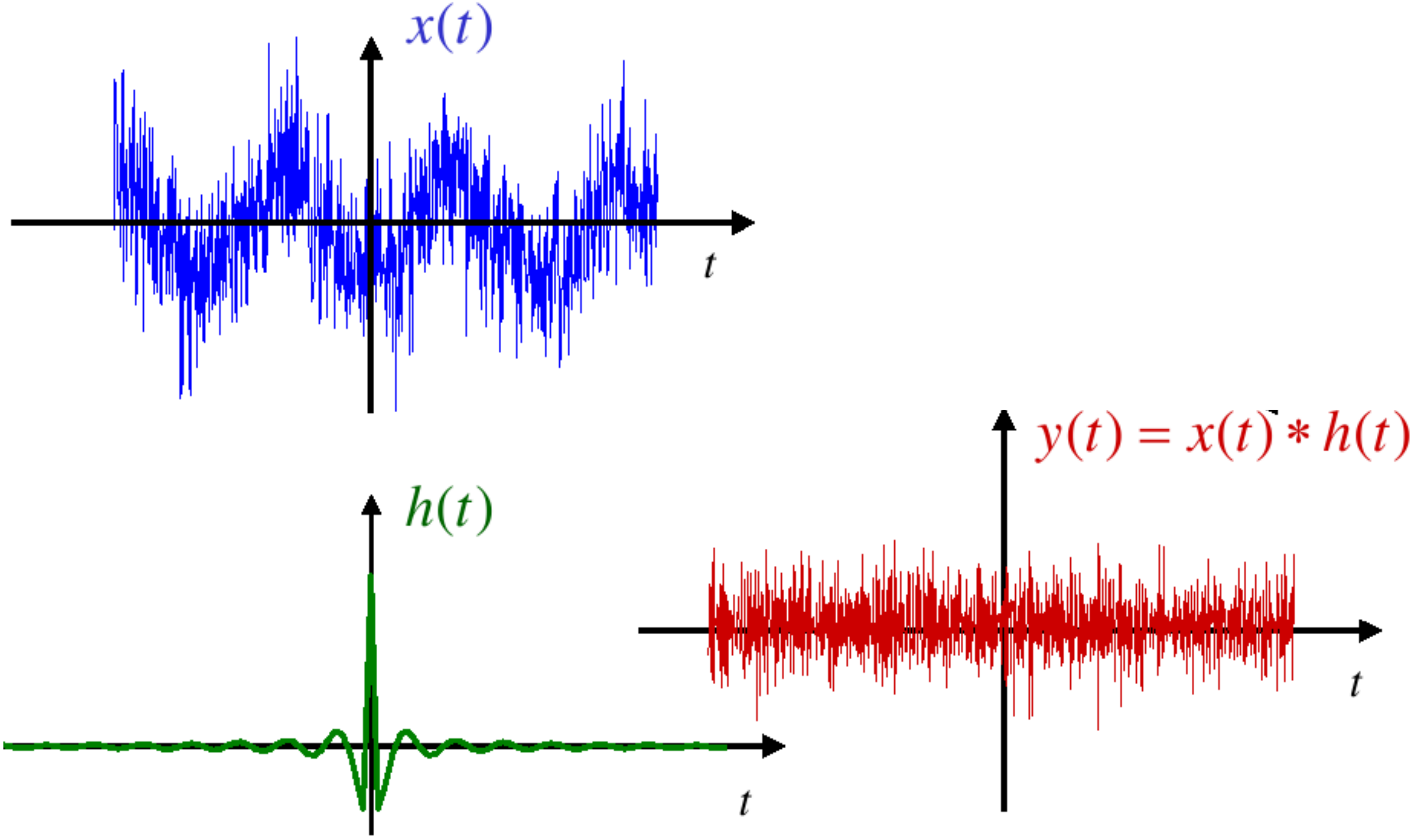
$$\boxed{y(n)=x(n)*h(n)}$$

Low-pass filter



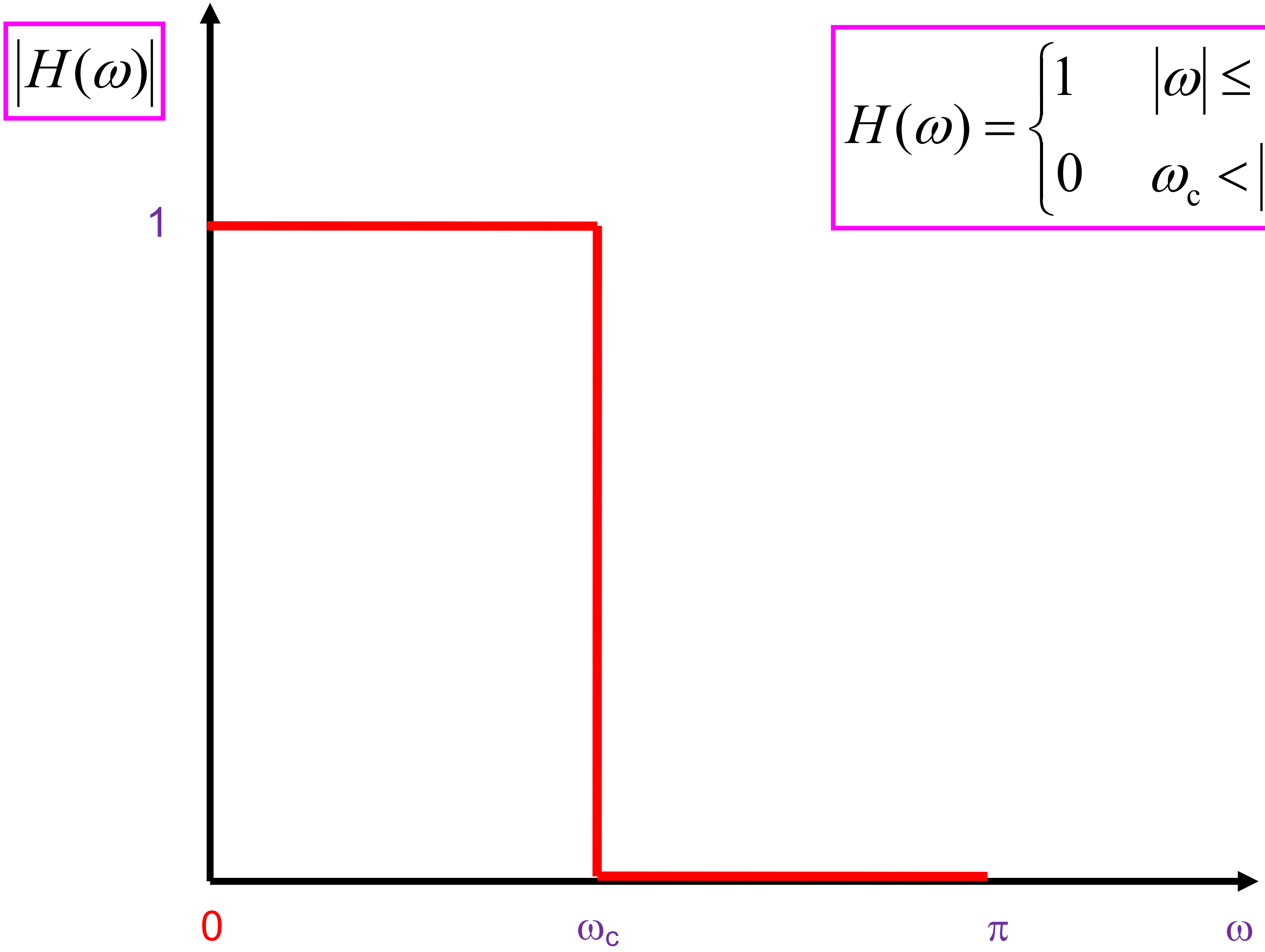
Filtering properties of the convolution

High-pass filter



Filtering properties of the convolution

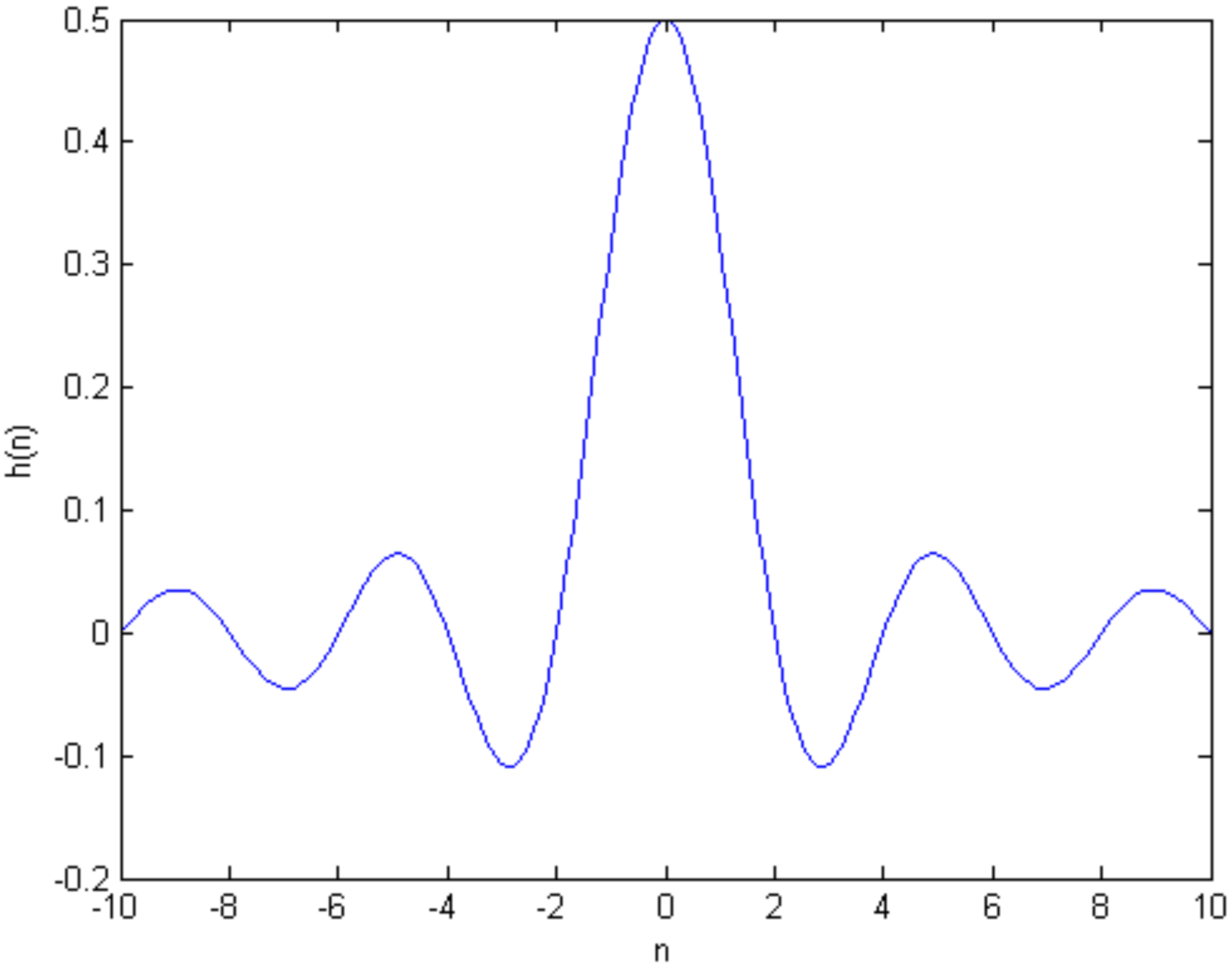
Ideal low-pass filter



$$H(\omega) = \begin{cases} 1 & |\omega| \le \omega_c \\ 0 & \omega_c < |\omega| < \pi \end{cases}$$

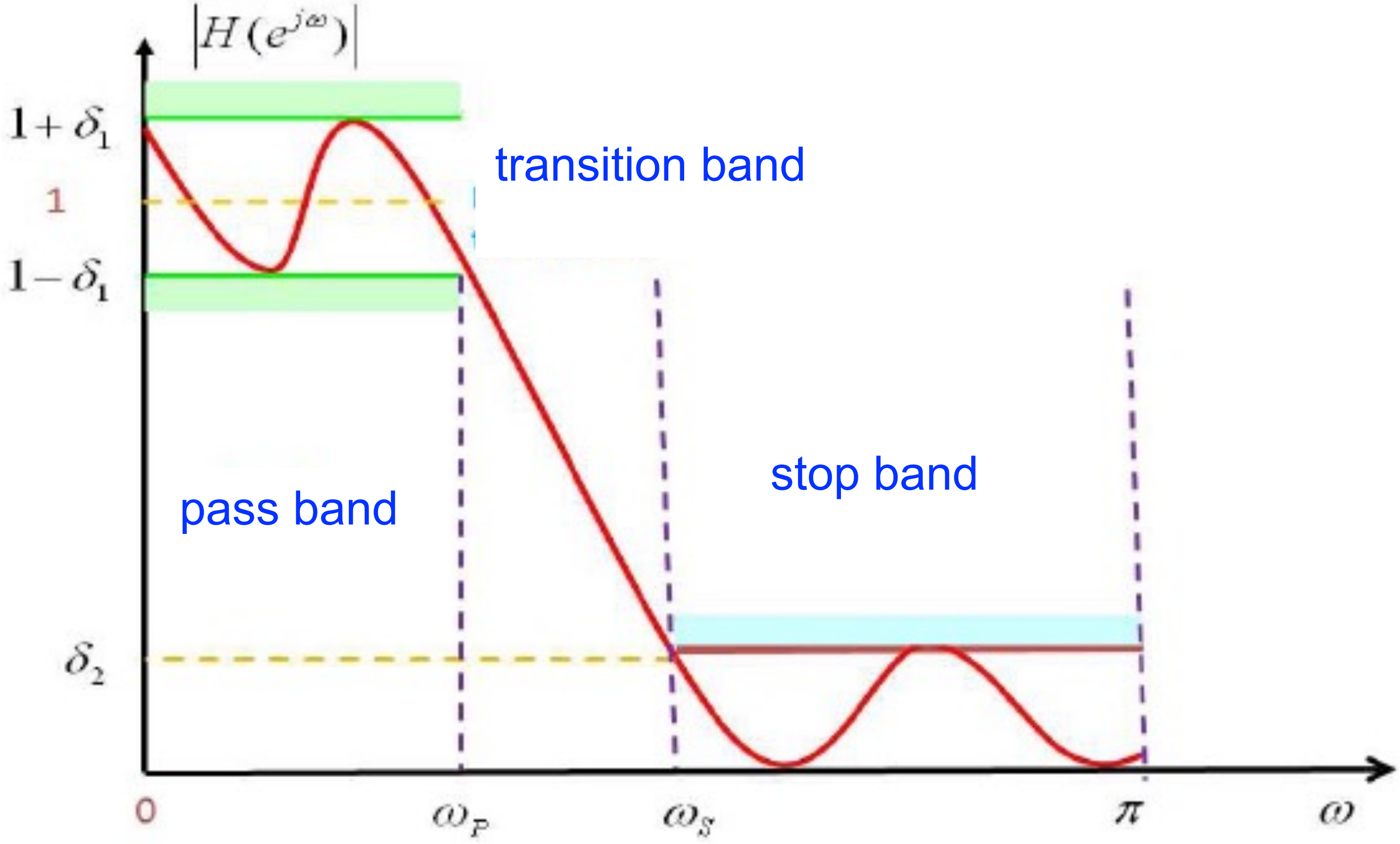
Ideal low-pass filter

h(n)



Impulse response corrisponding to low-pass filter

Real low-pass filter



transition band

stop band

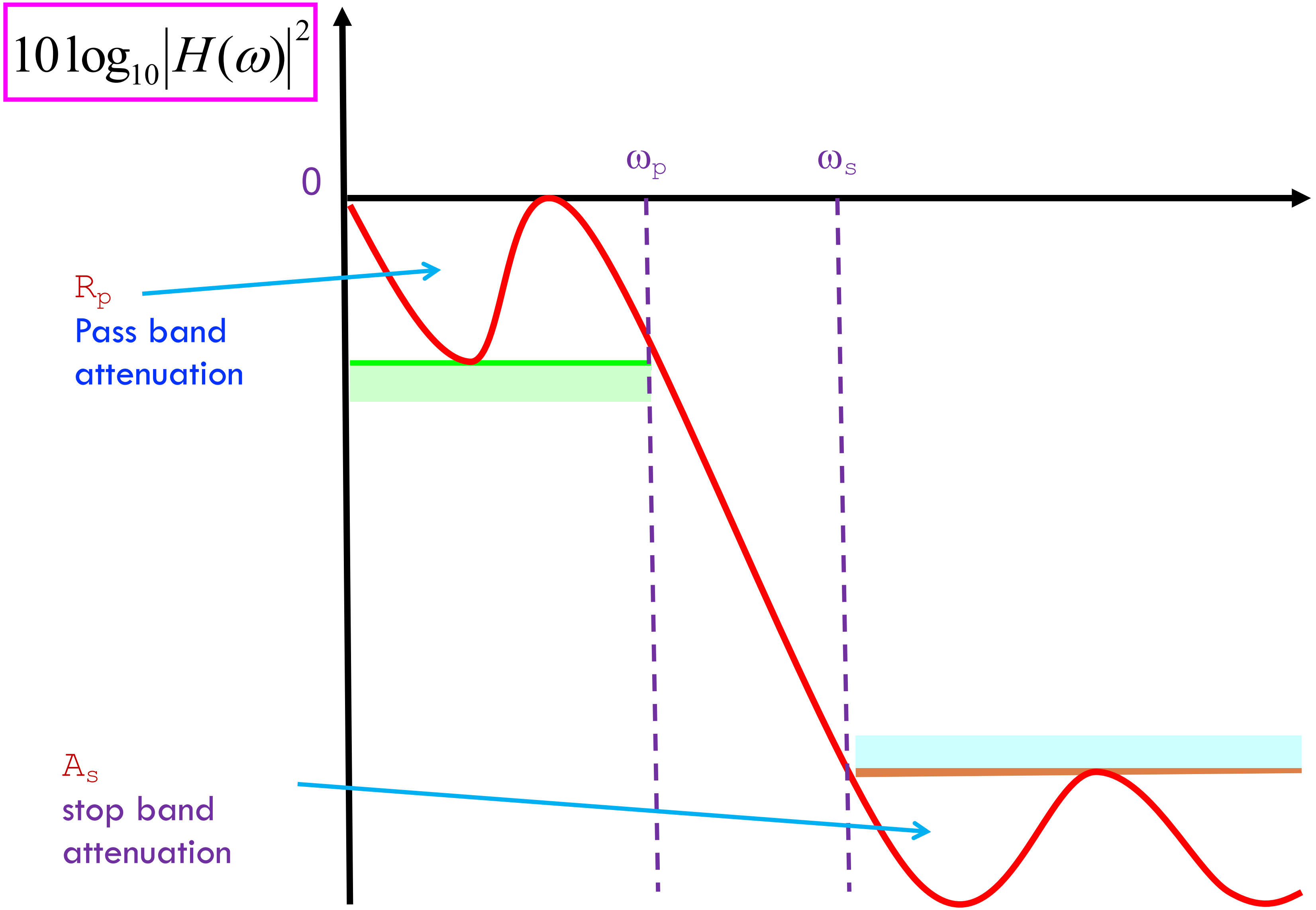
pass band

$$\boxed{1-\delta_1 \leq |H(\omega)| \leq 1+\delta_1 \qquad |\omega| \leq \omega_p}$$

$$\boxed{|H(\omega)| \leq \delta_2 \qquad \omega_s \leq |\omega| \leq \pi}$$

tollerance limits

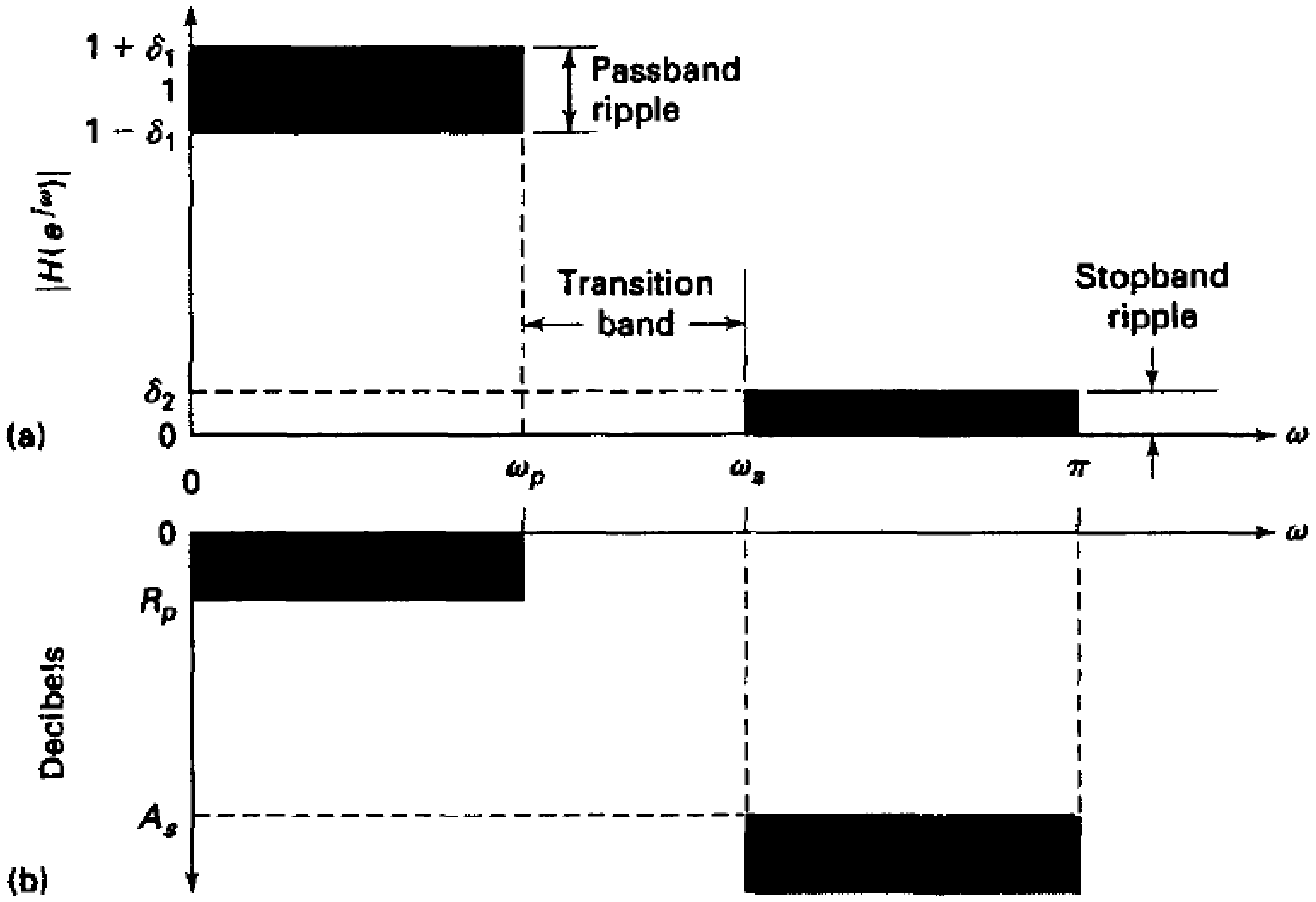
Decibel parameters



R_p
Pass band
attenuation

A_s
stop band
attenuation

Comparison



Comparison between paramaters with or without decibels

IIR and FIR

- Finite Impulse Response (FIR)
  - Polinomial Transfer function
  - Stable and linear phase

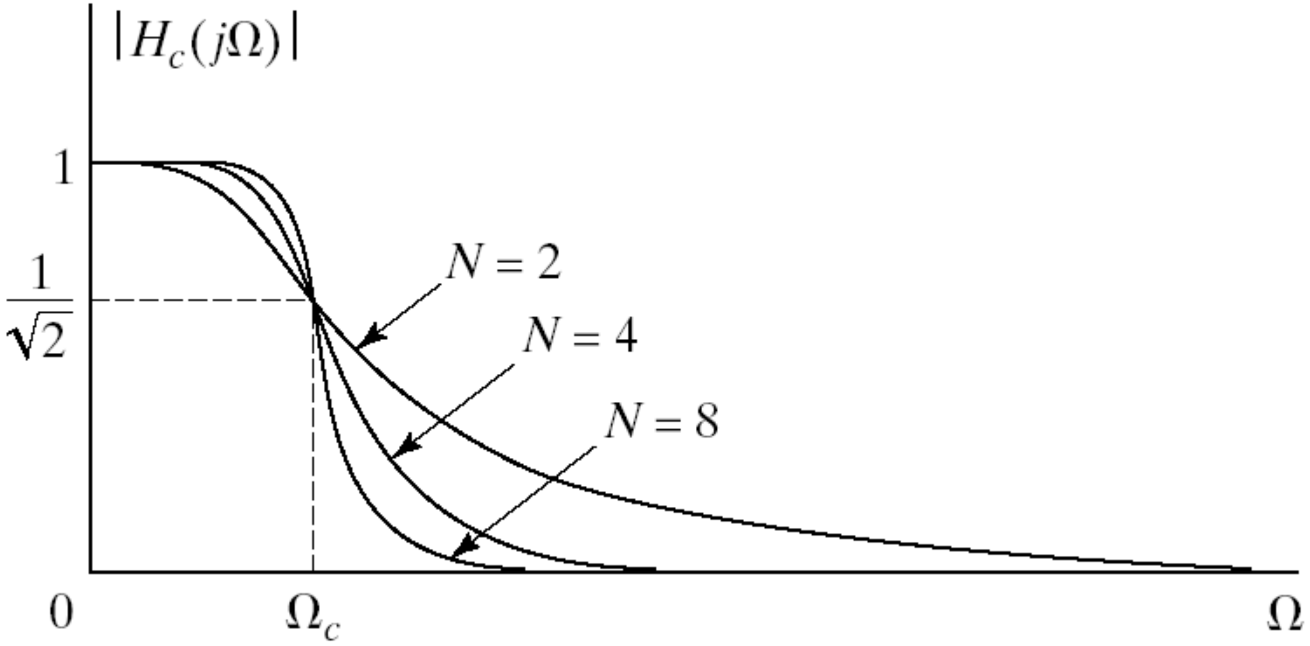
- Infinite Impulse Response (IIR)
  - Rational function
  - Non-linear phase and no stable
  - Better frequency cut

IIR

- To develop a muneric IIR filter
  - Transformation of an analogic filter in a numeric filter

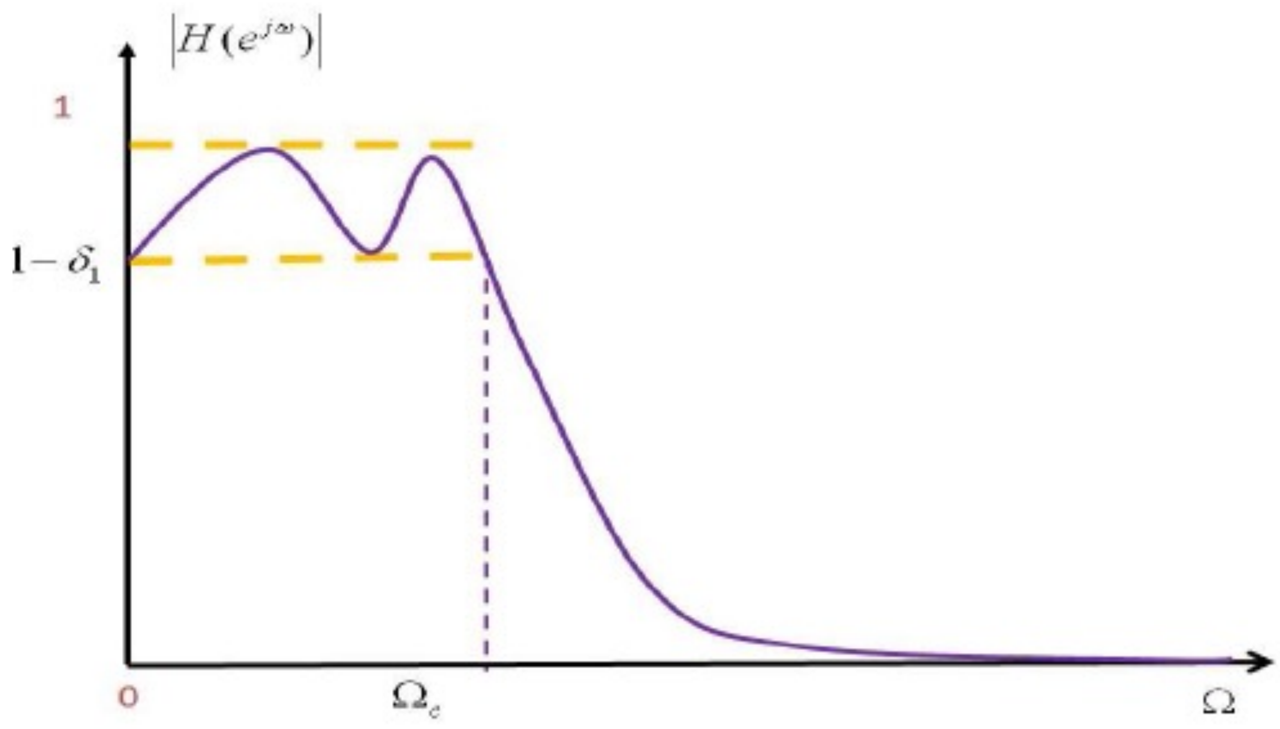
- Known analogic filters
  - Butterworth
  - Chebyshev
  - Elliptic

Butterworth



Butterworth analogic filter

Chebychev



Chebychev analogic filter

Ideal low-pass filter

$$H_d(\omega) = \begin{cases} 1 & |\omega| \le \omega_c \\ 0 & \omega_c < |\omega| < \pi \end{cases}$$

frequencies

$$h_d(n) = \frac{\sin \omega_c n}{\pi n}$$

time

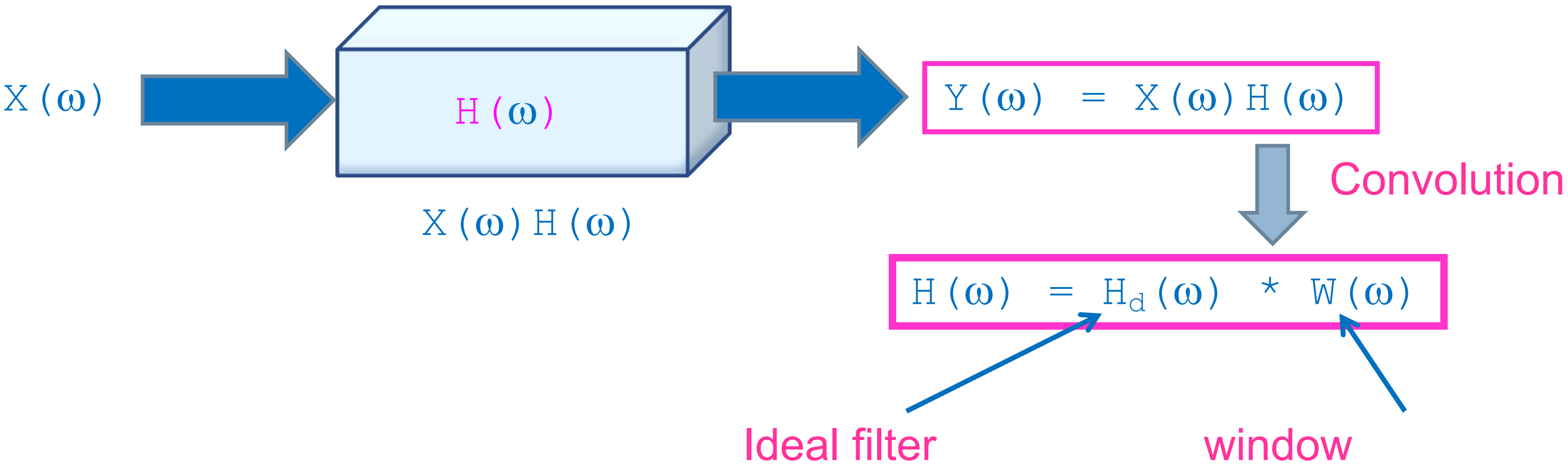
$$h(n) = h_d(n)w(n) \quad \text{dove} \quad w(n) = \begin{cases} 1 & 0 \le n \le M-1 \\ 0 & altrove \end{cases}$$

finite duaration

window

Filtering

Frequencies domain



$X(\omega)$

$H(\omega)$

$X(\omega)H(\omega)$

$Y(\omega) = X(\omega)H(\omega)$

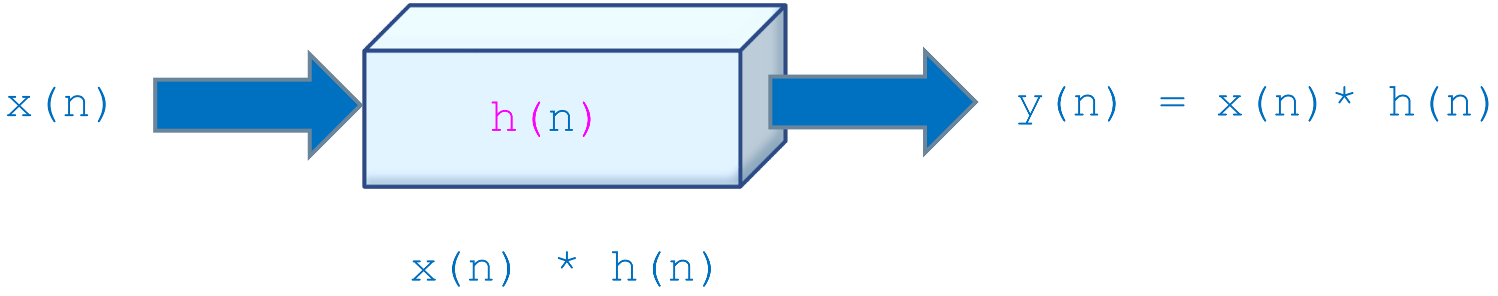
Convolution

$H(\omega) = H_d(\omega) * W(\omega)$

Ideal filter

window

Time domain



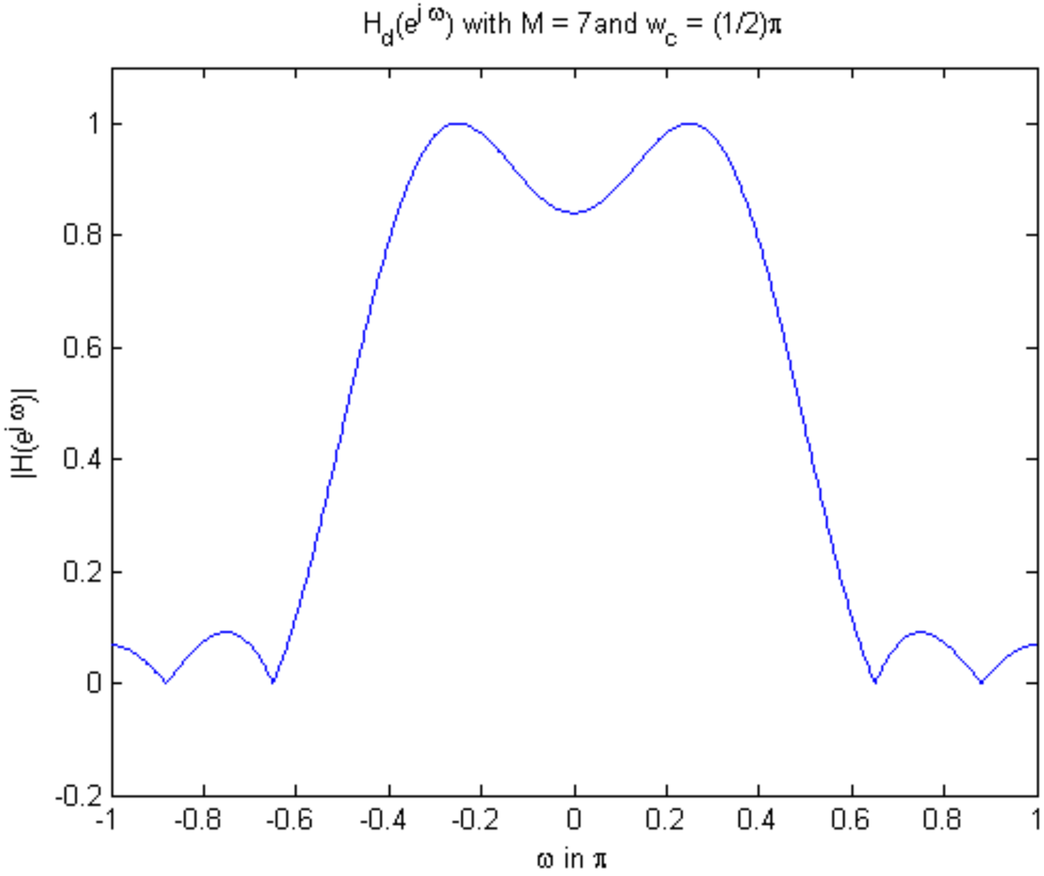
$x(n)$

$h(n)$

$x(n) * h(n)$

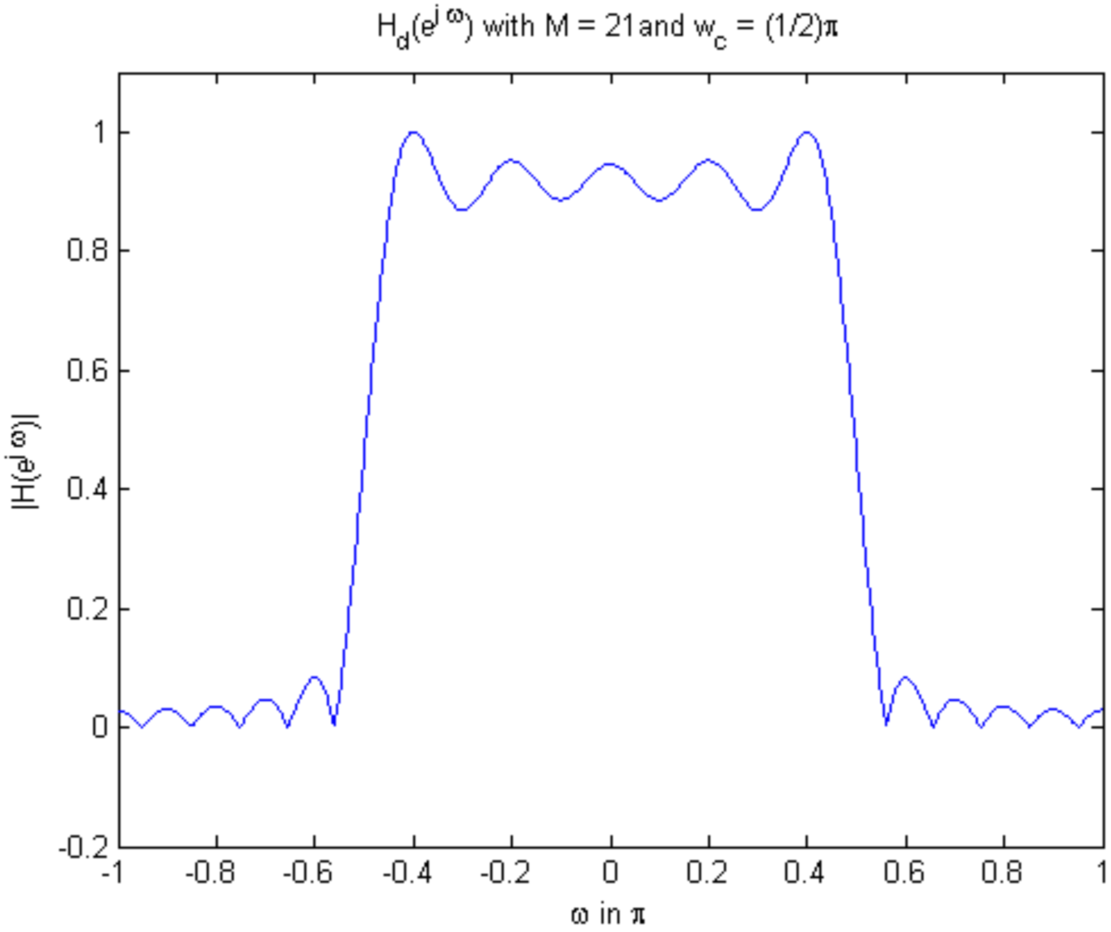
$y(n) = x(n) * h(n)$

Gibbs phenomenon



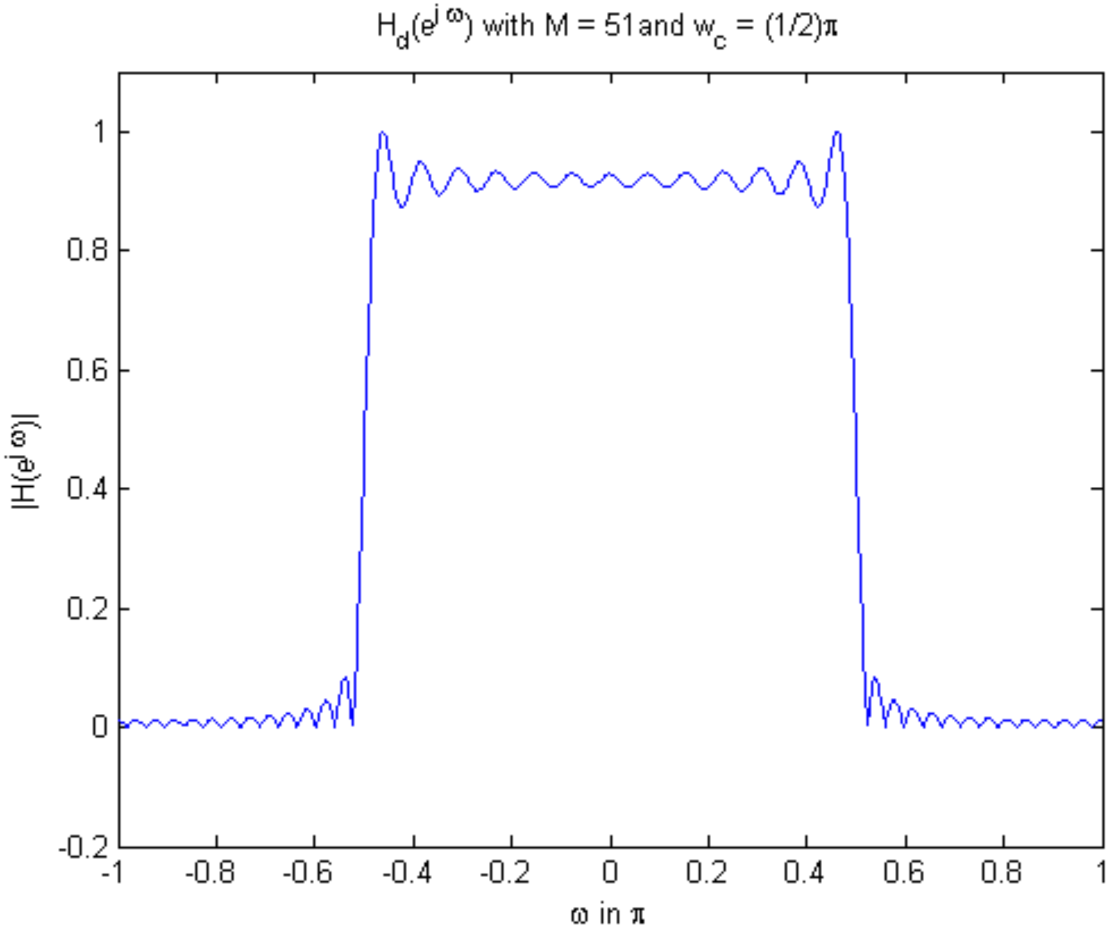
Rettangular window with M = 7

Gibbs phenomenon



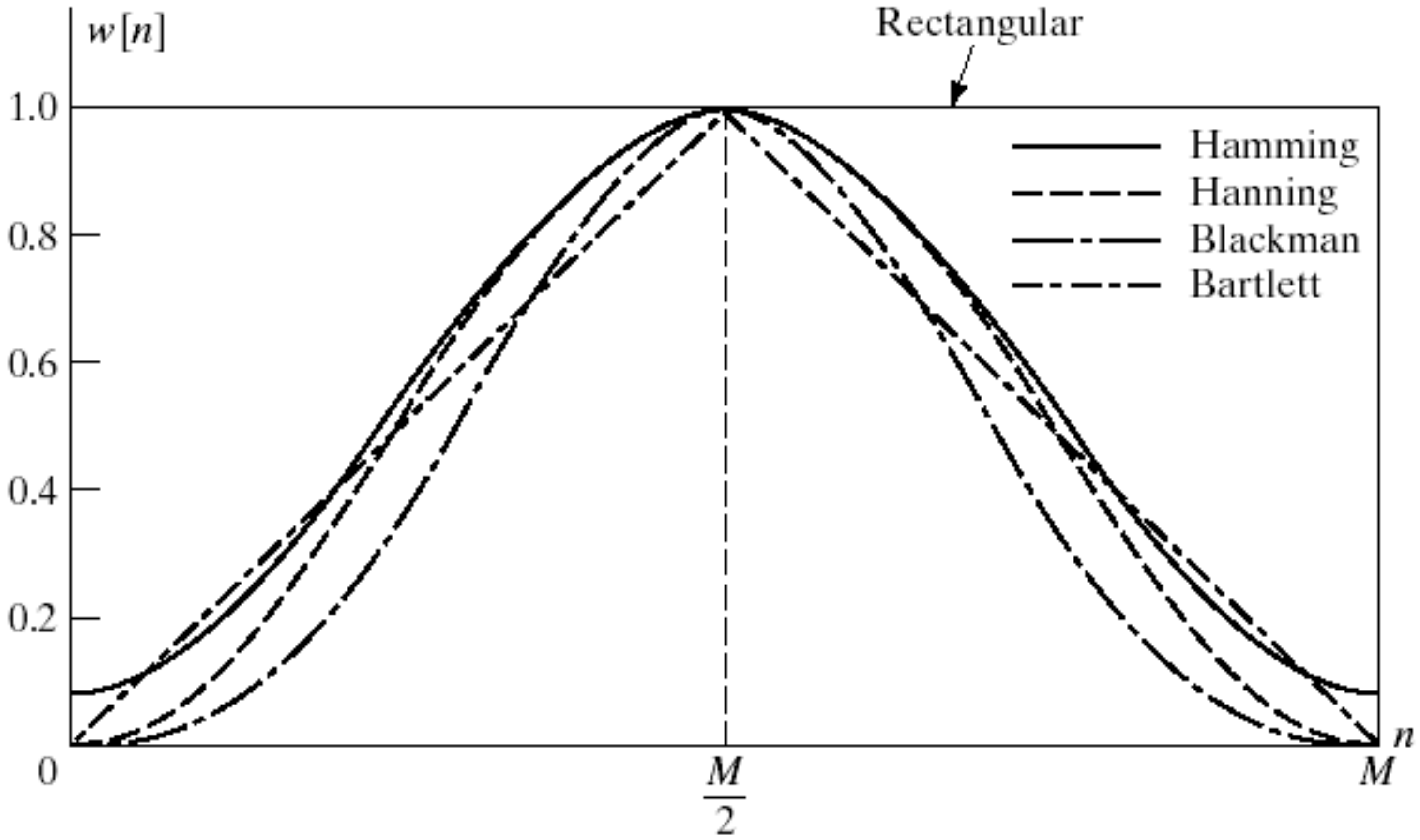
Rettangular window with M = 21

Gibbs phenomenon



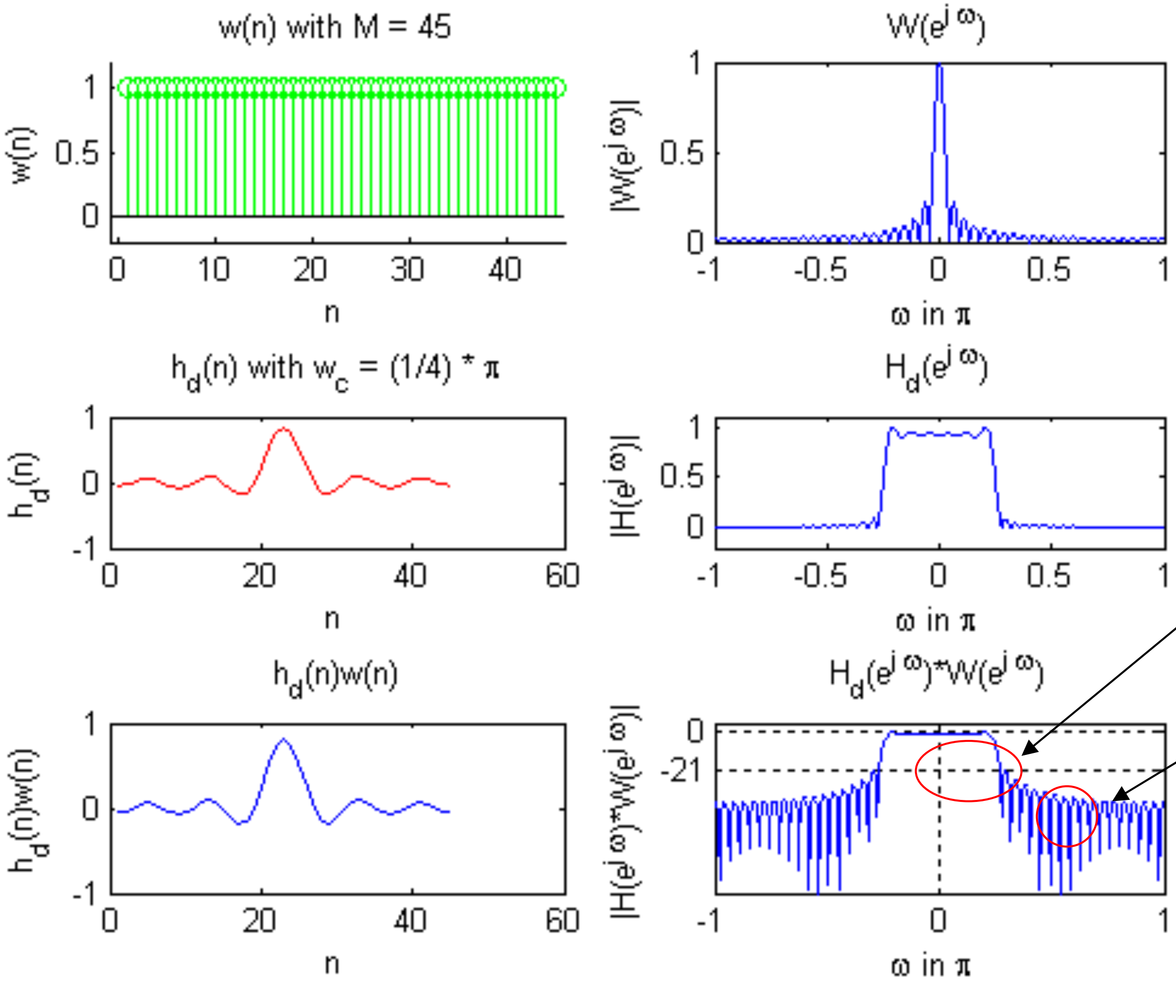
Rettangular window with M = 51

Gibbs phenomenon



To decrease the height of the side lobes different windows are used

Rectangular window



bandwidth

$$\omega_s - \omega_p = \frac{1.8\pi}{M}$$

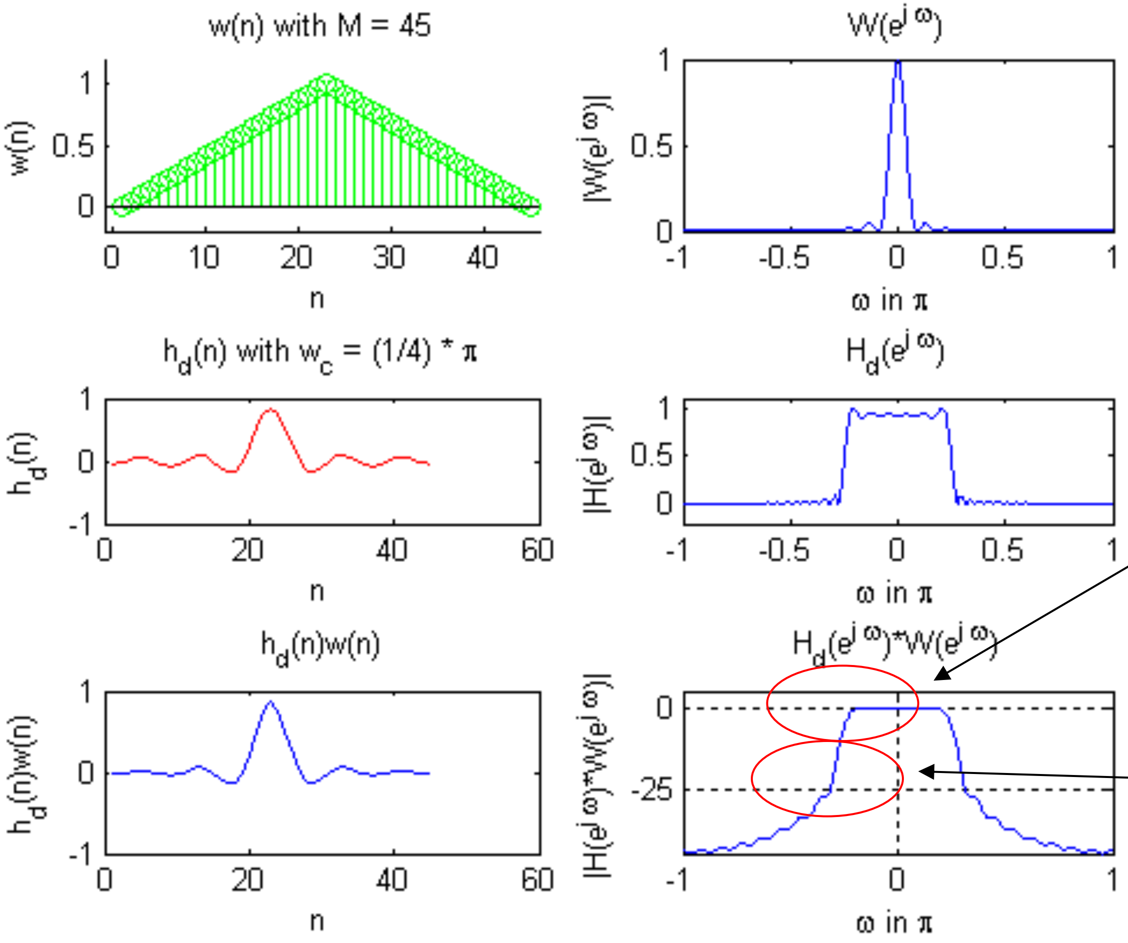
attenuation of 21 dB

M = 45

Bartlett window

$$\boxed{w(n)=\begin{cases} 2n/M & 0\le n\le M/2 \\ 2-2n/M & M/2\le n\le M \\ 0 & else \end{cases}}$$

Bartlett window



bandwidth

$$\omega_s - \omega_p = \frac{6.1\pi}{M}$$

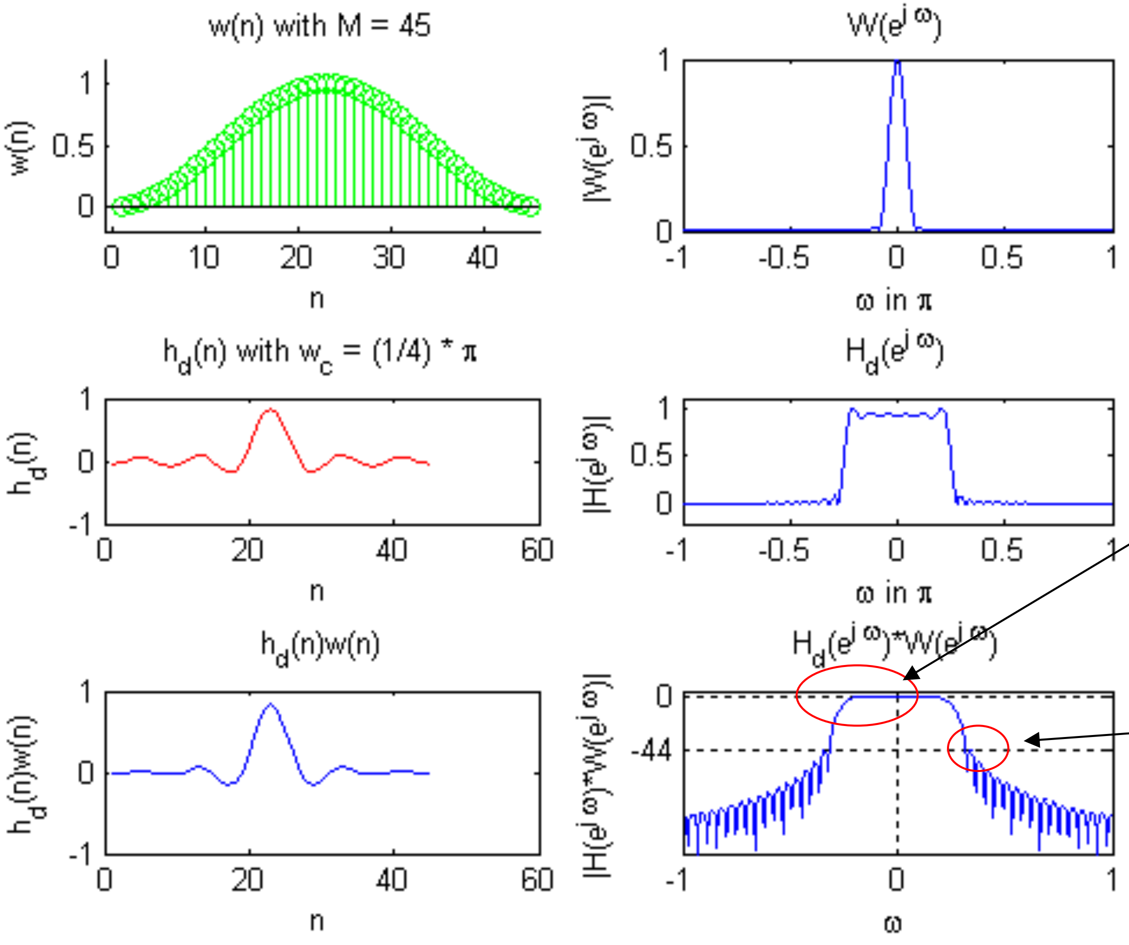
attenuation of 26 dB

M = 45

Hanning window

$$w(n)=\begin{cases}\frac{1}{2}\left[1-\cos\left(\frac{2\pi n}{M}\right)\right] & 0\le n\le M\\ 0 & else\end{cases}$$

Hanning window



bandwidth

$$\omega_s - \omega_p = \frac{6.2\pi}{M}$$

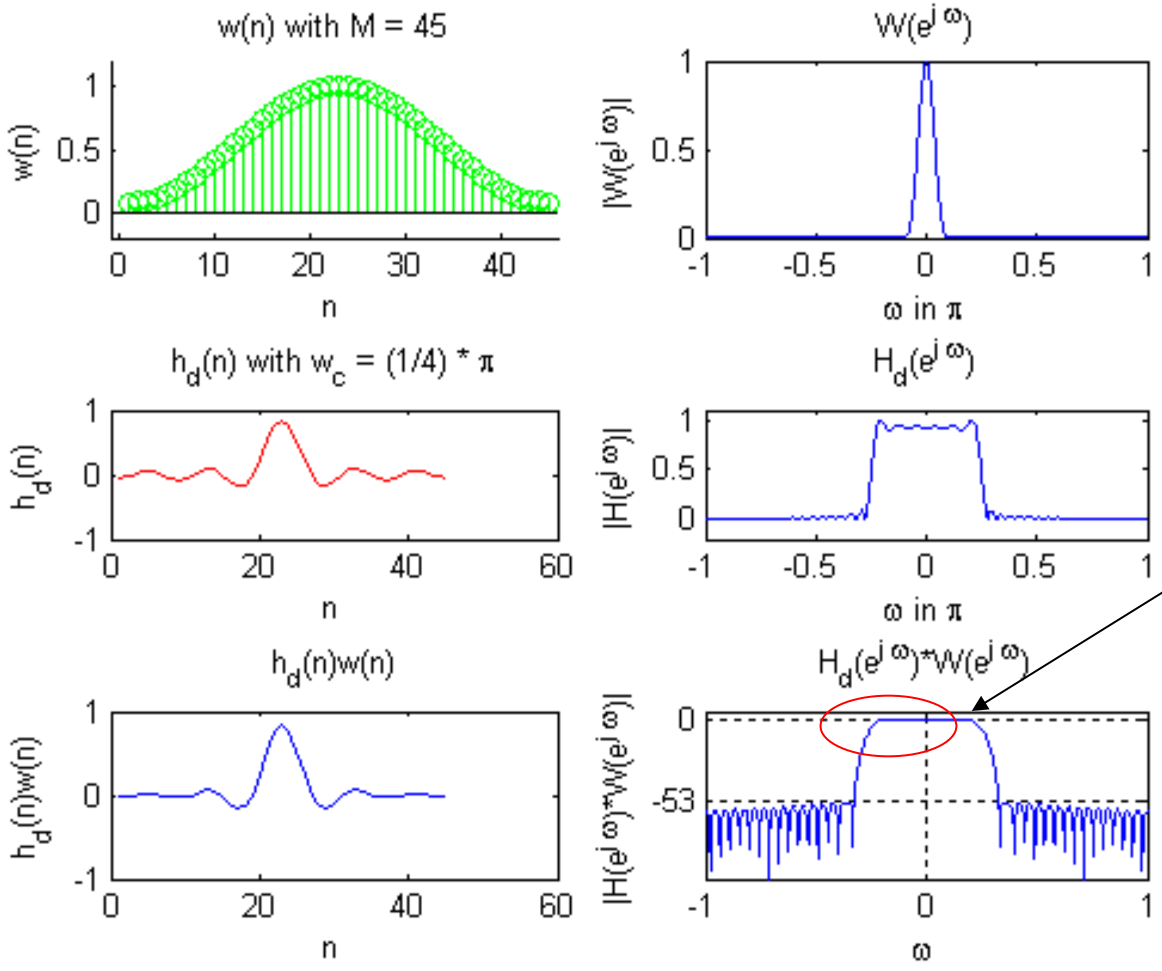
attenuation of 44 dB

M = 45

Hamming window

$$\boxed{w(n) = \begin{cases} 0.54 - 0.46\cos\left(\dfrac{2\pi n}{M}\right) & 0 \le n \le M \\ 0 & else \end{cases}}$$

Hamming window



bandwidth

$$\omega_s - \omega_p = \frac{6.6\pi}{M}$$

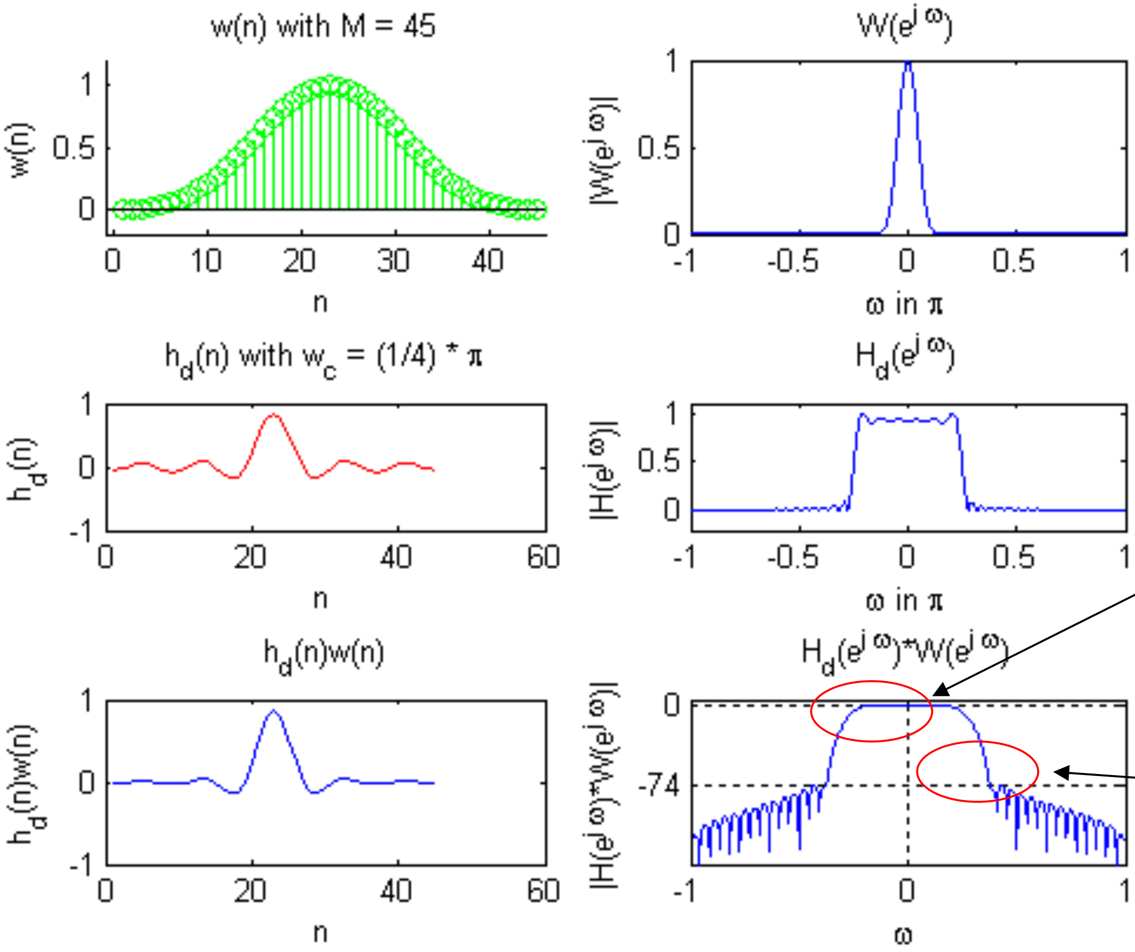
attenuation of
53 dB

M = 45

Blackman window

$$w(n) = \begin{cases} 0.42 - 0.5\cos\left(\dfrac{2\pi n}{M}\right) + 0.08\cos\left(\dfrac{4\pi n}{M}\right) & 0 \le n \le M \\ 0 & else \end{cases}$$

Blackman window



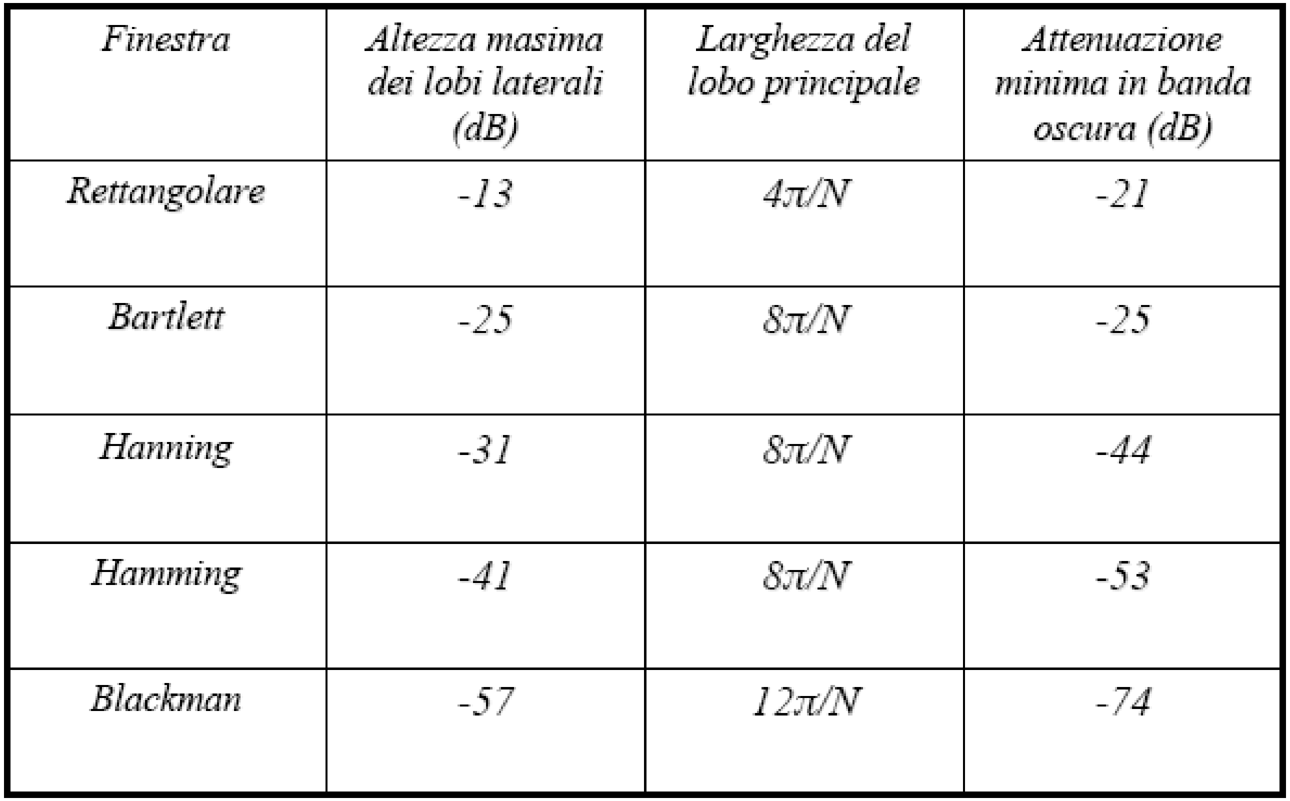
bandwidth

$$\omega_s - \omega_p = \frac{11\pi}{M}$$

attenuation of 74 dB

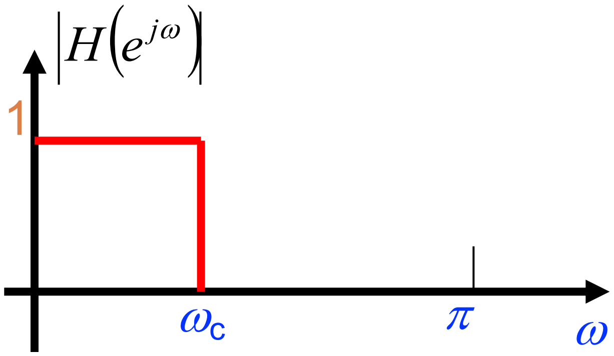
M = 45

Summarazing



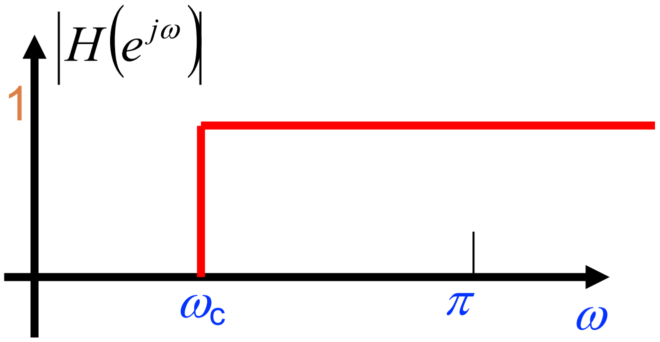
Finestra	*Altezza masima dei lobi laterali (dB)*	*Larghezza del lobo principale*	*Attenuazione minima in banda oscura (dB)*
Rettangolare	-13	$4\pi/N$	-21
Bartlett	-25	$8\pi/N$	-25
Hanning	-31	$8\pi/N$	-44
Hamming	-41	$8\pi/N$	-53
Blackman	-57	$12\pi/N$	-74

Types of filters



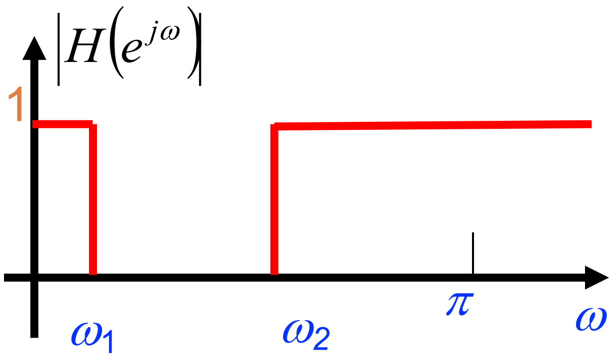
low-passa filter

a)



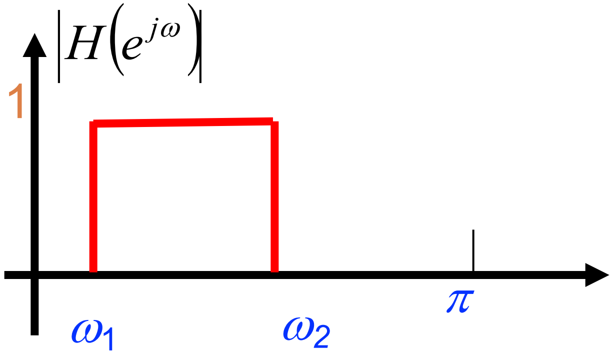
b)

high-passa filter



stop-band filter

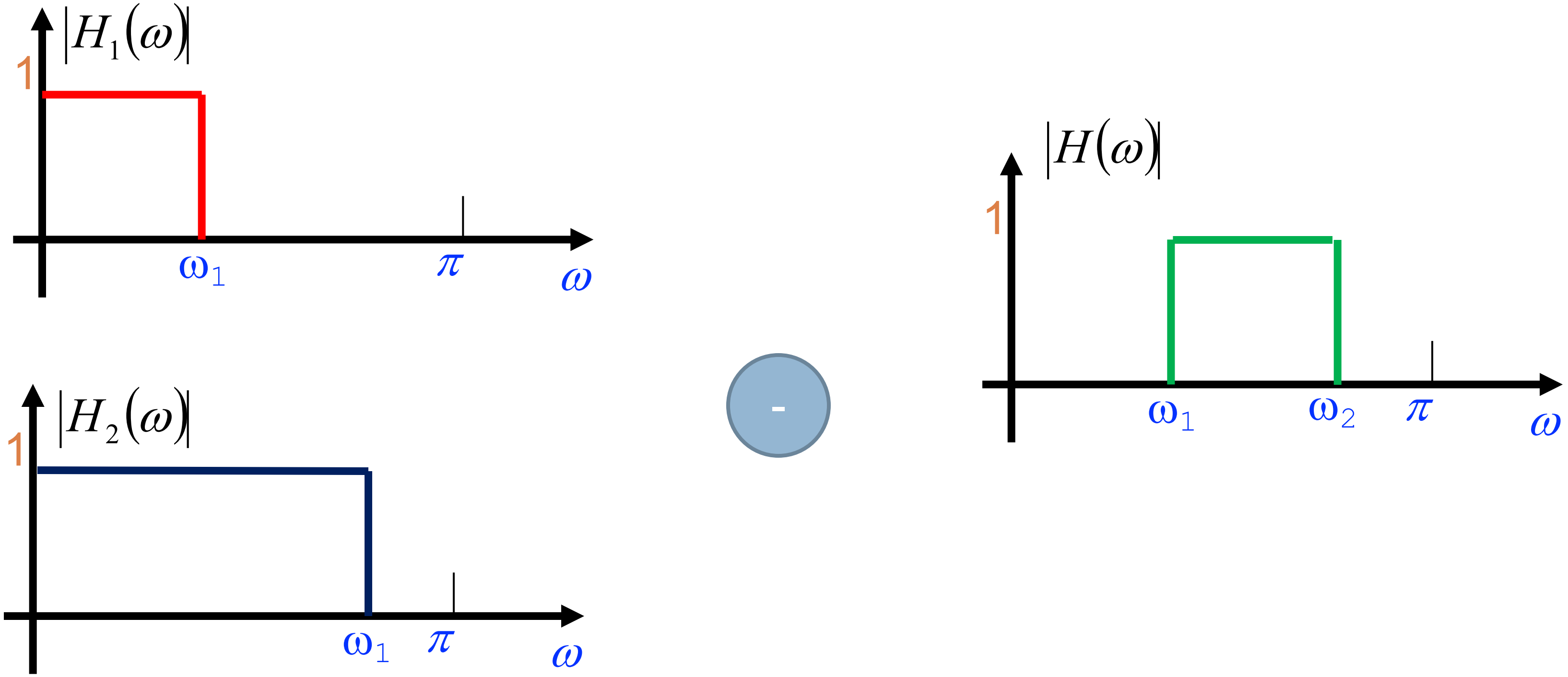
c)



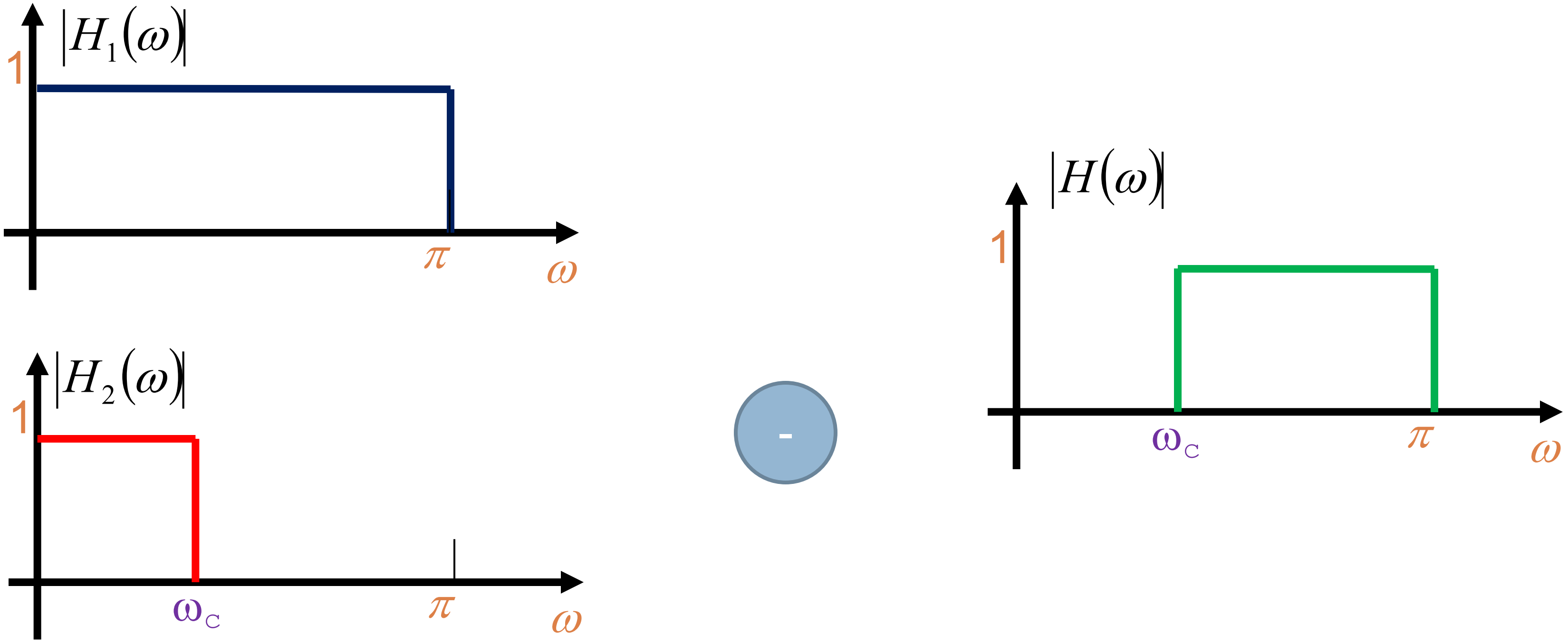
d)

band-pass filter

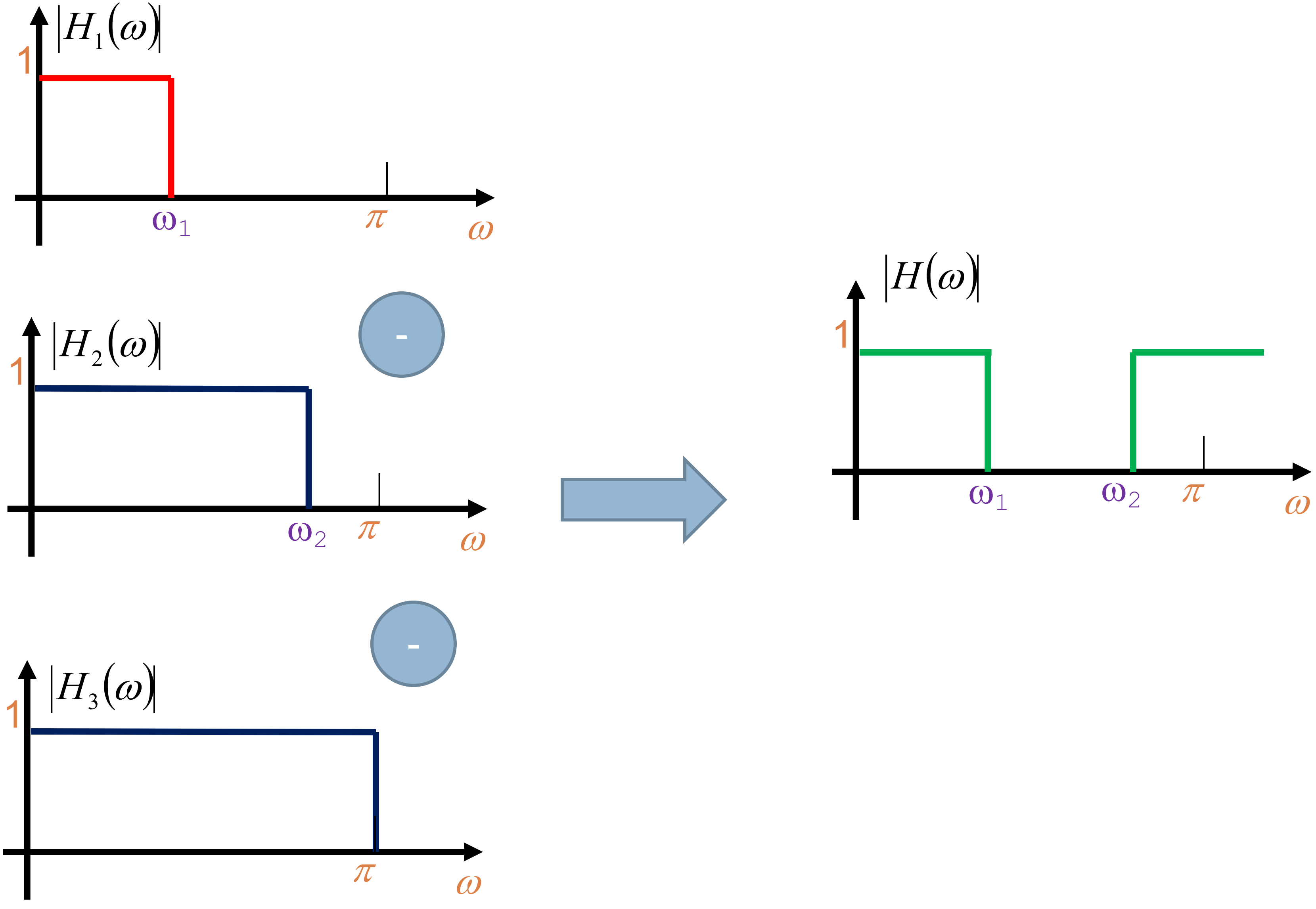
band-pass filter



high-pass filter



band-stop filter



References

- Material
  - Slides
  - Video Lessons
- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

ISP – Verification tests

Question 11

- Spectrum estimation
  - the frequency content of the signals can be estimated
- Question
  - Describe the Periodogram

Spectrum estimation techniques

Periodogram

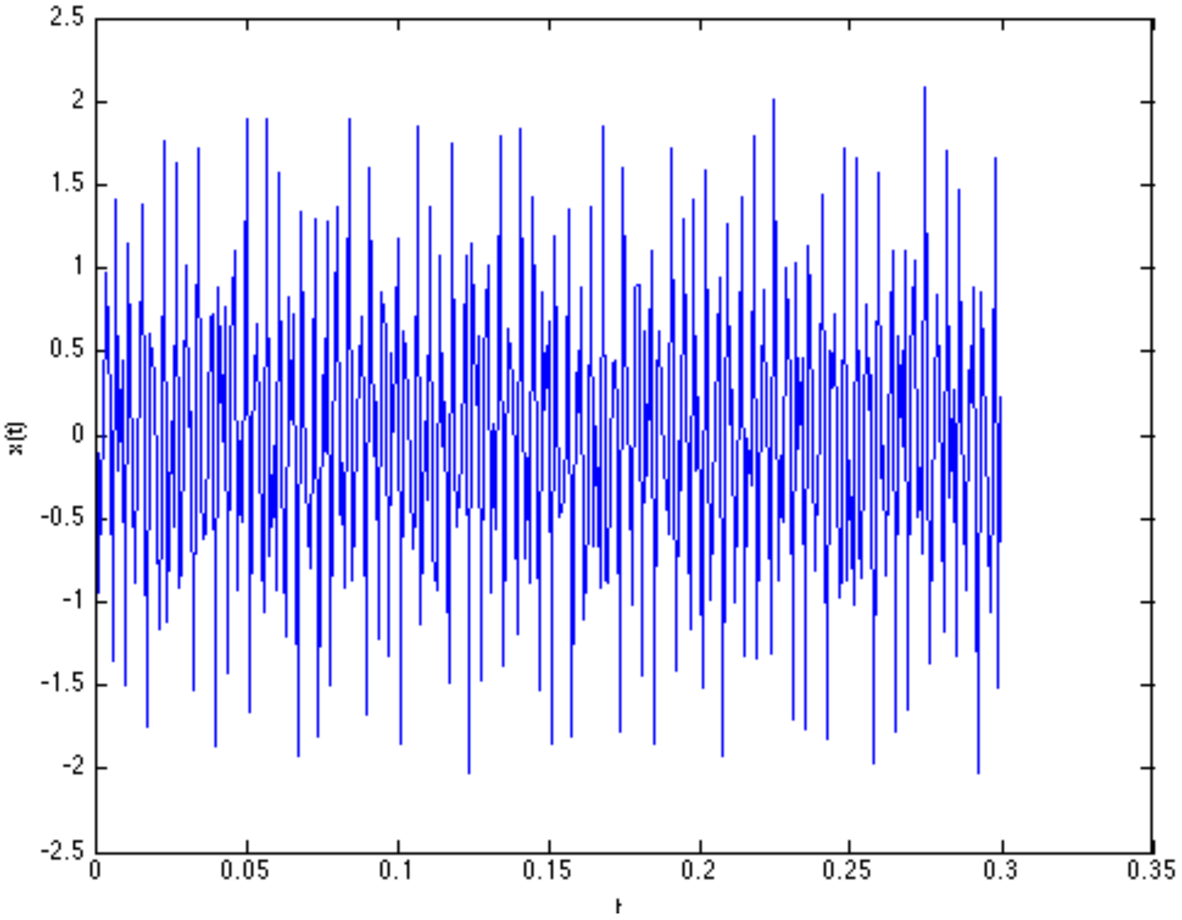
- DTFT of a sequence x(n)

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

- Periodogram

$$I_N(\omega) = \frac{1}{N} \left| X(e^{j\omega}) \right|^2 = \frac{1}{N} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} x(l) x(m) e^{j\omega m} e^{-j\omega l}$$

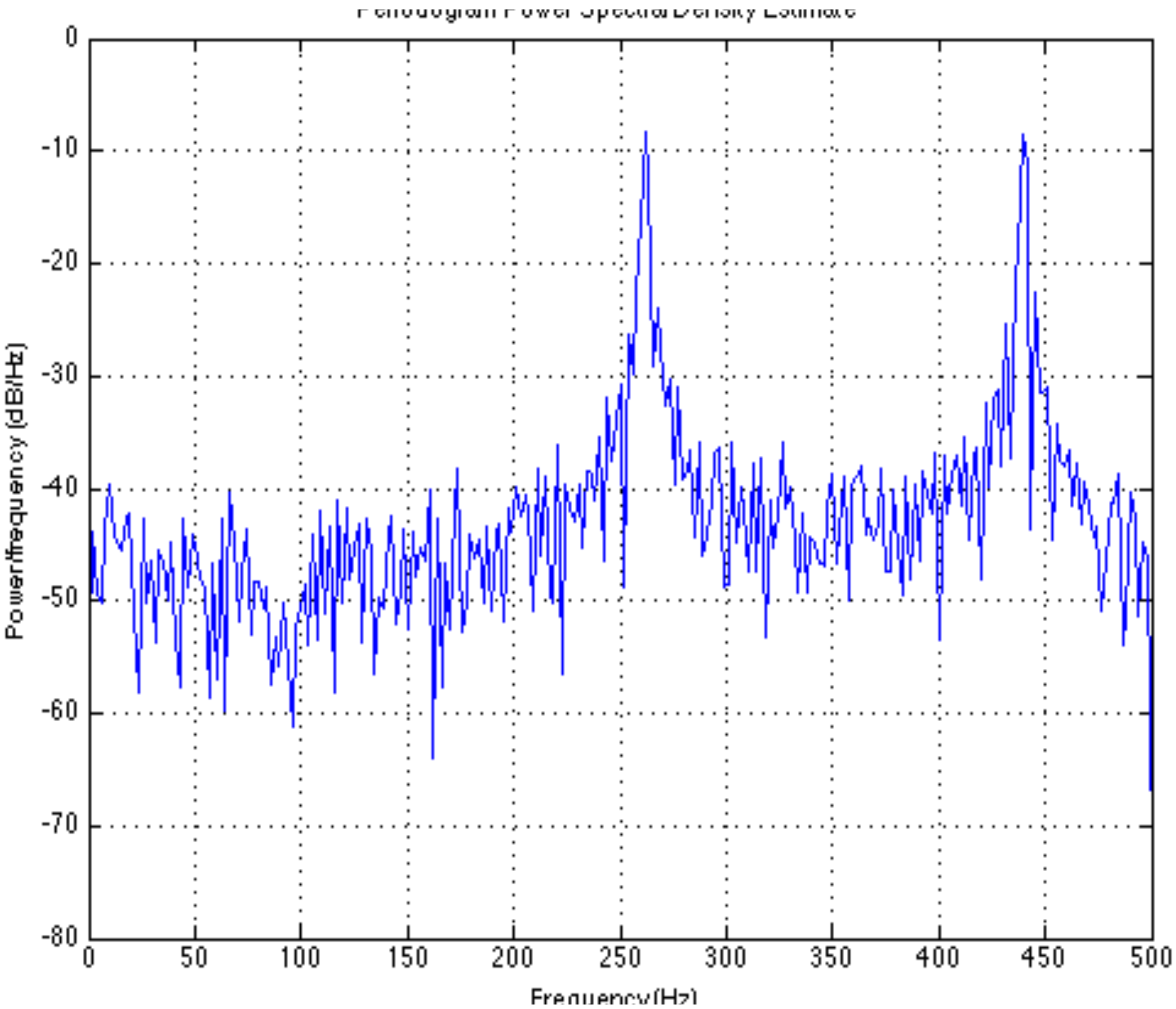
Periodogram



Source signal

$$x(n) = \sin(2\pi \cdot 262 \cdot n) + \sin(2\pi \cdot 440 \cdot n) + 0.1 \cdot randn(n)$$

Periodogram



Estimated Periodogram

Considerations

- The Periodogram
  - We have problems with increasing the N
  - The variance does not approach zero as the data length N increases
  - The periodogram is not a consistent estimator (i.e. converges in some sense to the true value)
  - Why does the variance not decrease with increasing N?
    - Increasing N means increasing the number of individual frequencies (instead of increasing the accuracy of each frequency)
- “smoothed” Periodograms are defined

Bartlett's method

- A sequence x(n) is divided in *K* segments of *M* samples (*N* = *KM*)

$$x^{(i)}(n) = x(n + iM - M) \qquad 0 \le n \le M - 1; 1 \le i \le K$$

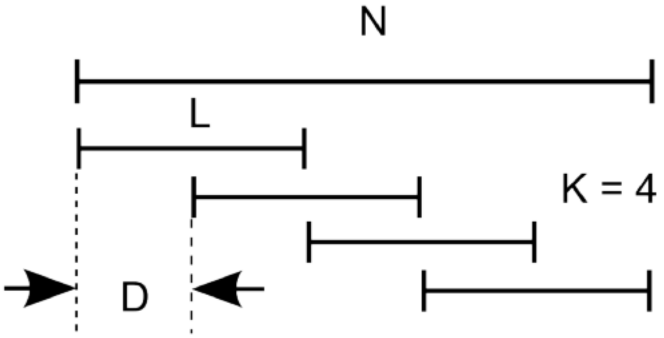
- *K* Periodograms are calculated

$$\boxed{I_M^{(i)}(\omega) = \frac{1}{M}\left|\sum_{n=0}^{M-1} x^{(i)}(n) e^{-j\omega n}\right|^2}$$

Bartlett's method

- Estimation of the spectrum

$$\boxed{B_{xx}(\omega) = \frac{1}{K}\sum_{i=1}^{K} I_M^{(i)}(\omega)}$$



Welch's method

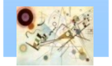
- A window w(n) is applied

$$J_M^{(i)}(\omega) = \frac{1}{MU}\left|\sum_{n=0}^{M-1} x^{(i)}(n)w(n)e^{-j\omega n}\right|^2$$

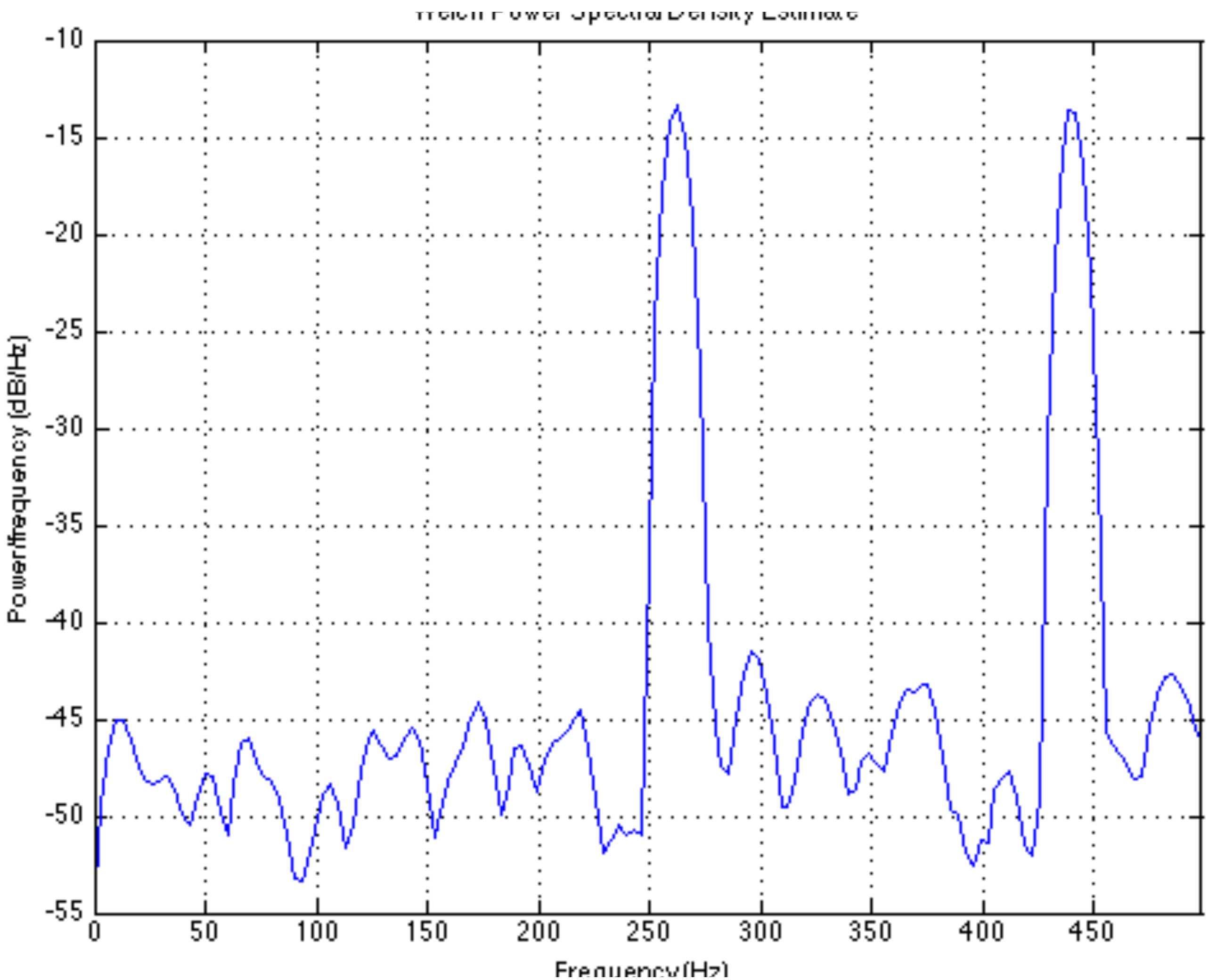
$$U = \frac{1}{M}\sum_{n=0}^{M-1} w^2(n)$$

$$\Rightarrow$$

$$B_{xx}^{\omega}(\omega) = \frac{1}{K}\sum_{i=1}^{K} J_M^{(i)}(\omega)$$



Welch



Estimated Periodogram

MUSIC

- Very important are the methods based on the decomposition by eigenvectors and eigenvalues

- Methods
  - Pisarenko
  - MUltiple Signal Classification (MUSIC)
  - Estimation of Signal Parameters via Rational Invariance Technique (ESPRIT)

- Applications
  - Spectral estimation
  - Direction Of Arrival (DOA)

MUSIC

- The sequence x(n) is

$$\boxed{x(n) = \sum_{i=1}^{p} A_i e^{jn\omega_i} + w(n)}$$

- Autocorrelation matrix

$$\boxed{r_x(k) = \sum_{i=1}^{p} P_i e^{jk\omega_i} + \sigma_\omega^2 \delta(k)}$$

$$\boxed{P_i = \left|A_i\right|^2}$$

MUSIC

- We write the autocorrelation matrix as

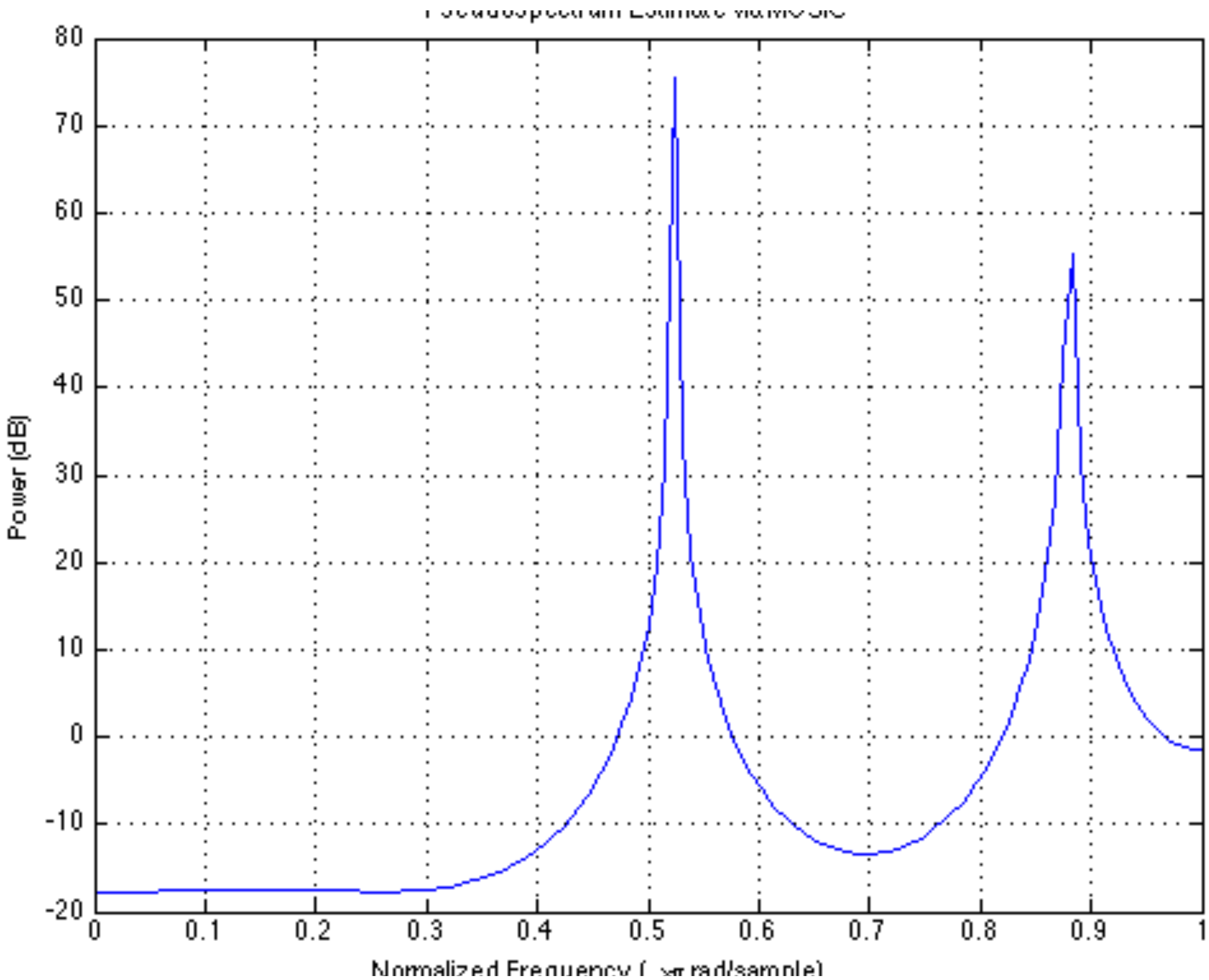
$$\mathbf{R}_x = \mathbf{R}_s + \mathbf{R}_n = \sum_{i=1}^{p} P_i \mathbf{e}_i \mathbf{e}_i^H + \sigma_w^2 \mathbf{I}$$

$$\mathbf{e}_i = \left[1, e^{j\omega_i}, e^{j2\omega_i}, \ldots, e^{j(M-1)\omega_i}\right]$$

- The p principal components span the signal space

$$P_{MUSIC}(e^{j\omega}) = \frac{1}{\sum_{i=p+1}^{M} \left|\mathbf{e}^H \mathbf{v}_i\right|^2}$$

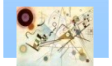
Example



Estimated Periodogram

References

- Material
  - Slides
  - Video Lessons
- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Discrete-time signal processing, A. V. Oppenheim, R. W. Schafer, J.R. Buck, Upper Saddle River, N.J., Prentice Hall, 1999, ISBN 0-13-754920-2
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

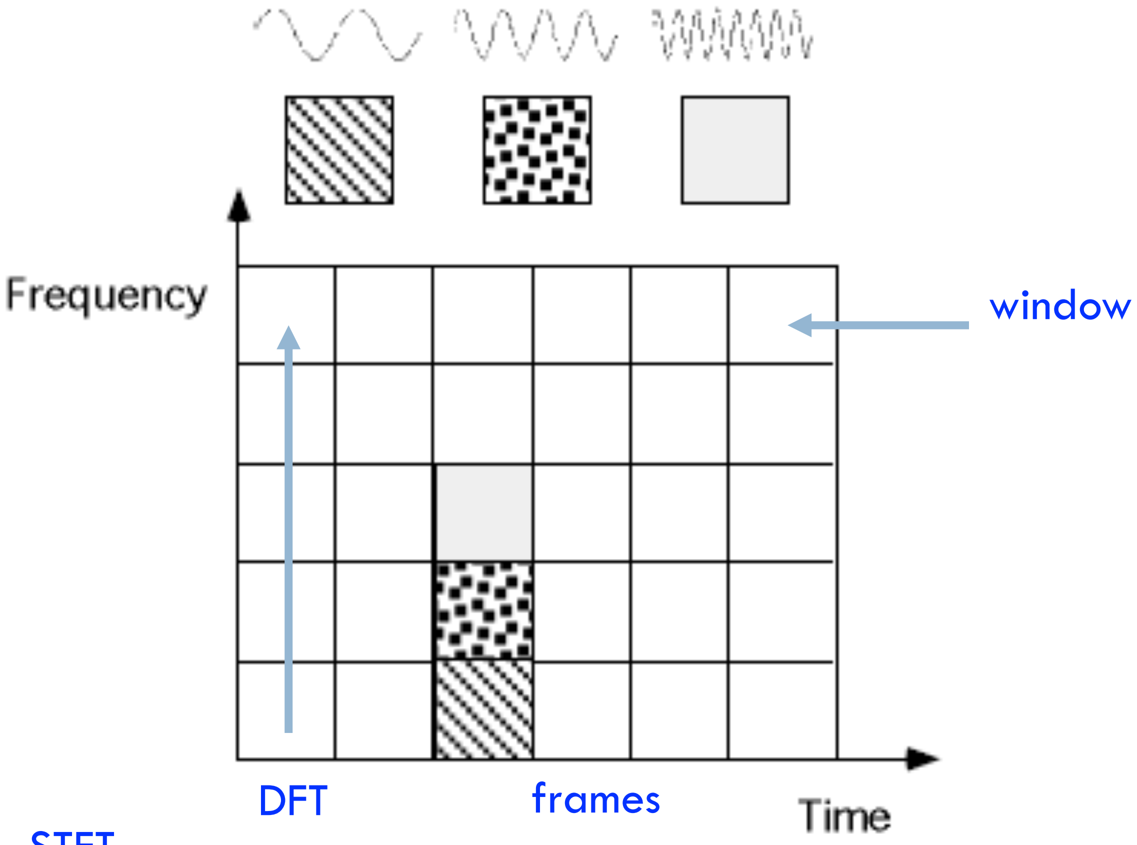


Question 12

- Spectrum estimation
  - the content of the signals can be estimated both in time and frequency
- Question
  - Describe the Short Time Fourier and the Wavelet Transforms

Short Time Fourier Transform

- Short Time Fourier Transform (STFT)
  - resolution in both time and frequency domains



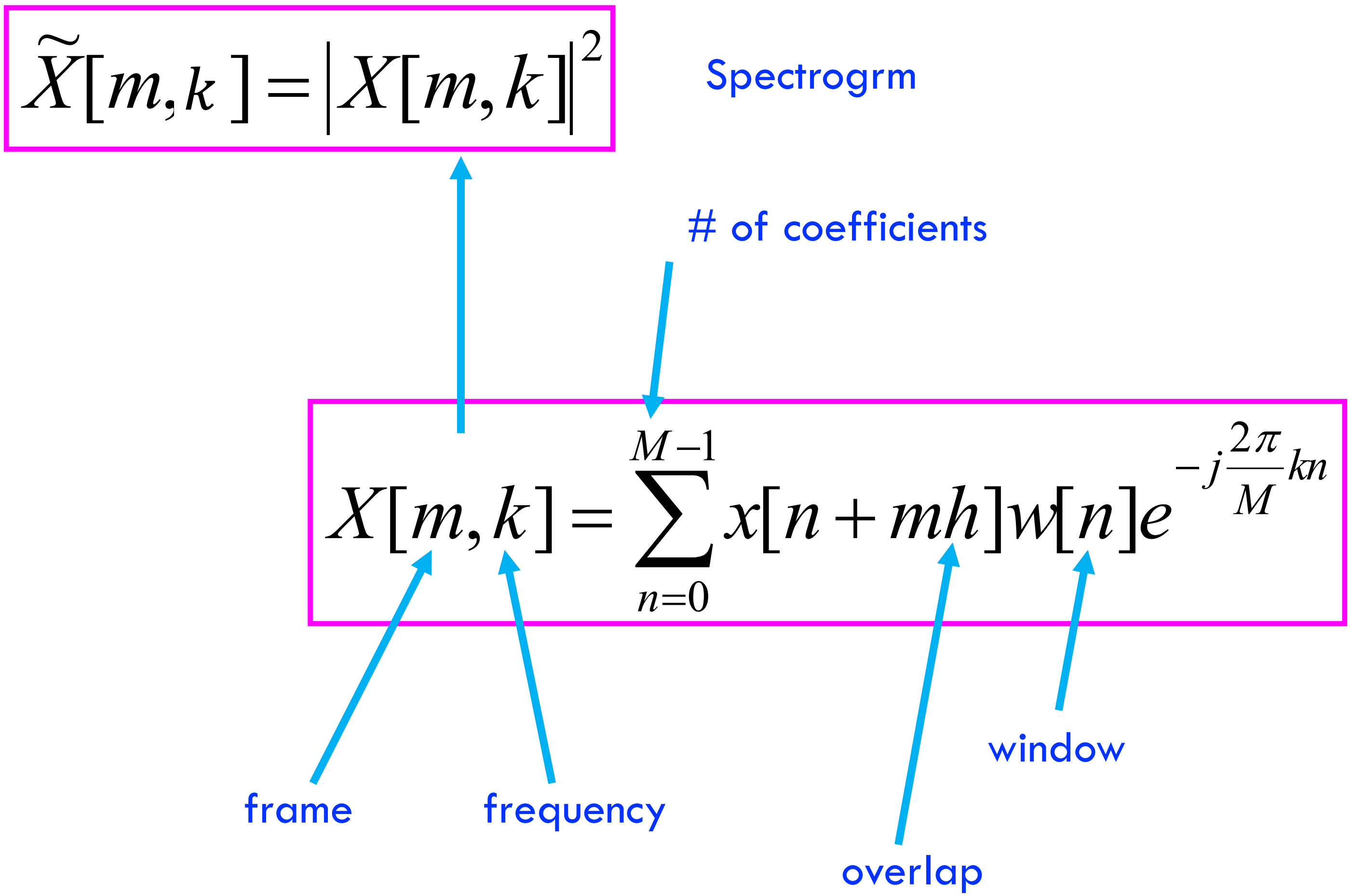
window

DFT

frames

STFT

STFT



$$\tilde{X}[m,k] = \left|X[m,k]\right|^2$$

Spectrogrm

# of coefficients

$$X[m,k] = \sum_{n=0}^{M-1} x[n+mh]w[n]e^{-j\frac{2\pi}{M}kn}$$

frame

frequency

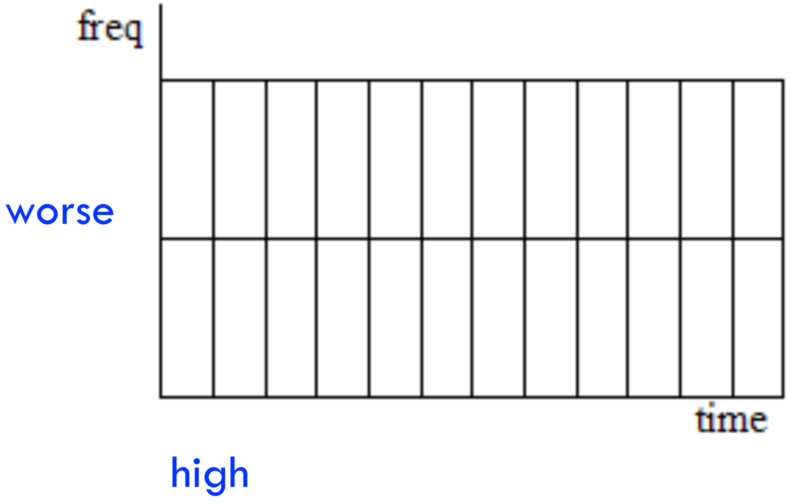
window

overlap

STFT Uncertainty

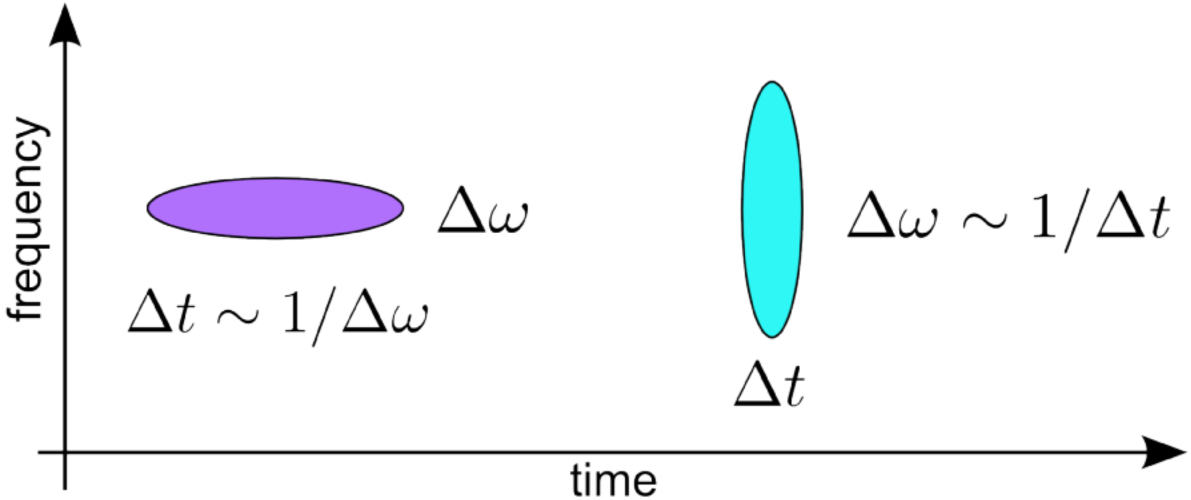
- spectrogram behaviour
  - similar to the *Heisenberg uncertainty principle*
  - higher time resolution imply worse frequencies resolution and vice versa

$$\Delta t \Delta \omega \geq \frac{1}{2}$$

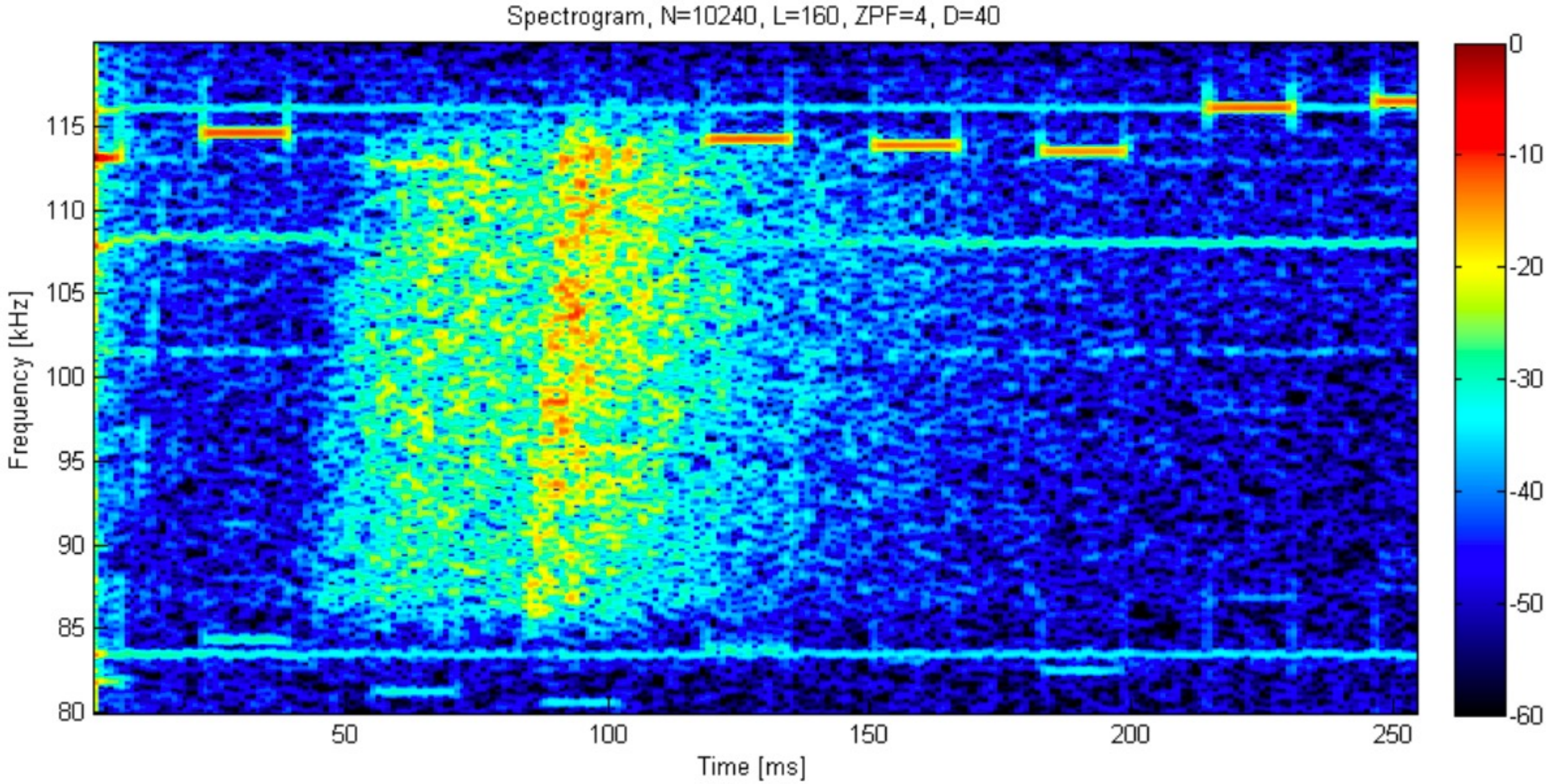


STFT Uncertainty

$$\Delta t \Delta\omega \geq \frac{1}{2}$$



STFT of Sonar data



Single ping of sonar data

Wavelet

- ## Wavelet Transform
  - solves the resolution problem
  - the signal is analyzed at different frequencies ans resolutions
    - High frequencies
      - High time resolution, low frequencies resolution
    - Low frequencies
      - High frequencies resolution, low time resolution

ISP – Verification tests

Wavelet Analysis

$$G_X(t, f) = \int_{-\infty}^{+\infty} w(t-\tau)\, e^{-j2\pi f\tau} x(\tau) d\tau$$

Mother Wavelet

shift

$$w(t-\tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t)\, \psi\left(\frac{t-\tau}{s}\right) dt$$

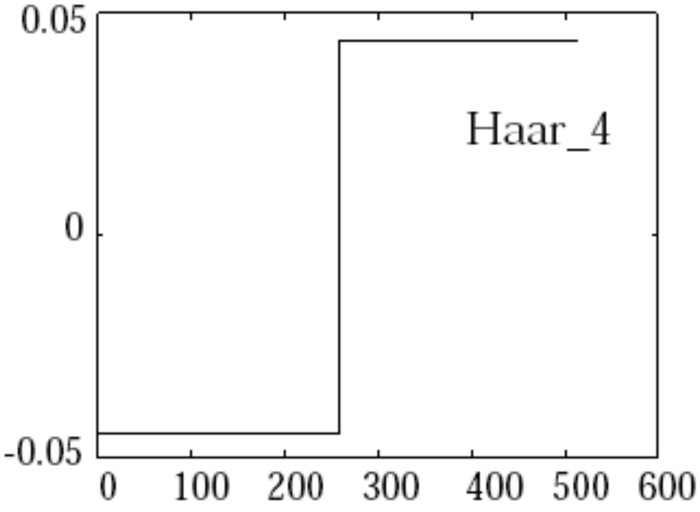
scale

The wavelet transform is simply a kind of correlation function between the mother wavelet scaled and shifted, and the input signal

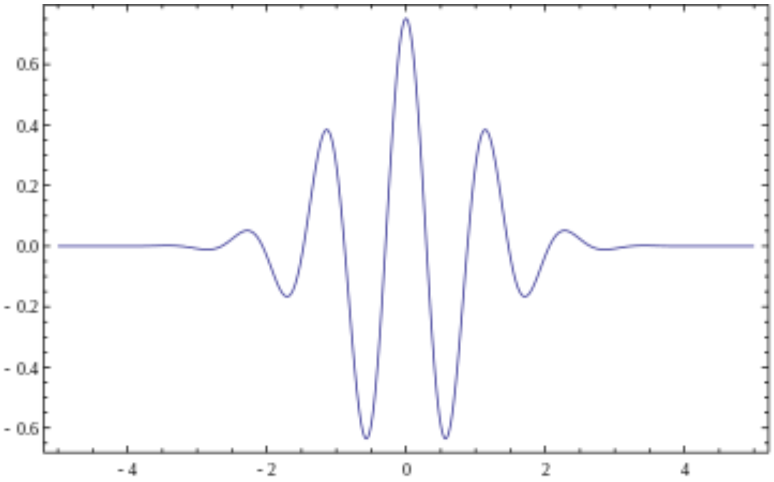
Scale factor

- s>1: dilated
- s<1: compressed

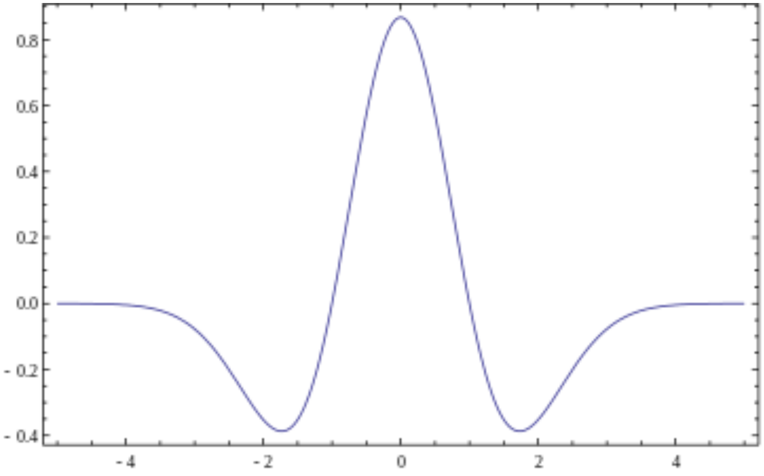
Mother Wavelet



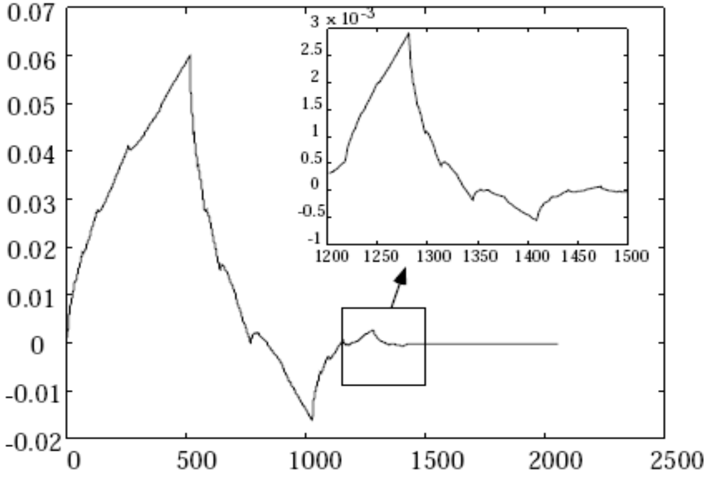
Haar



Morlet

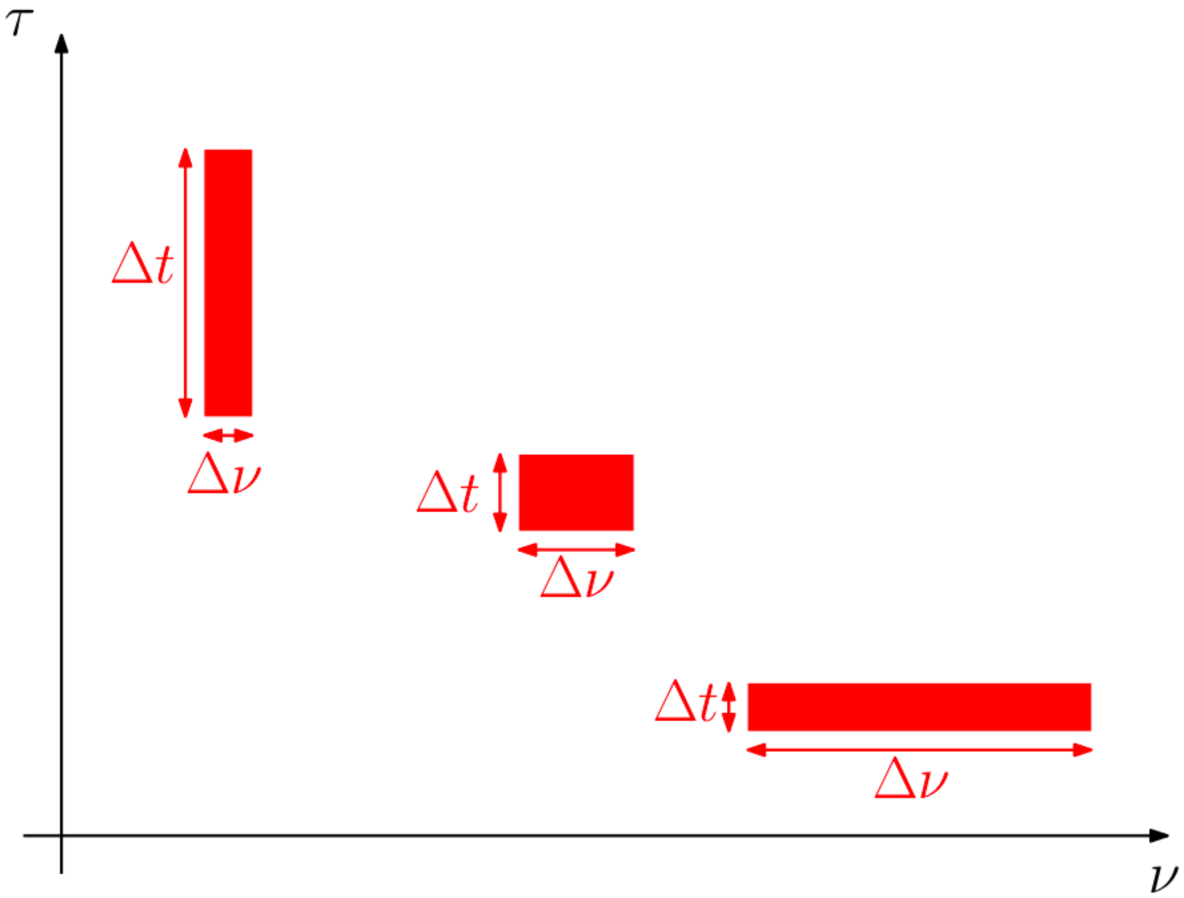


Mexican hat



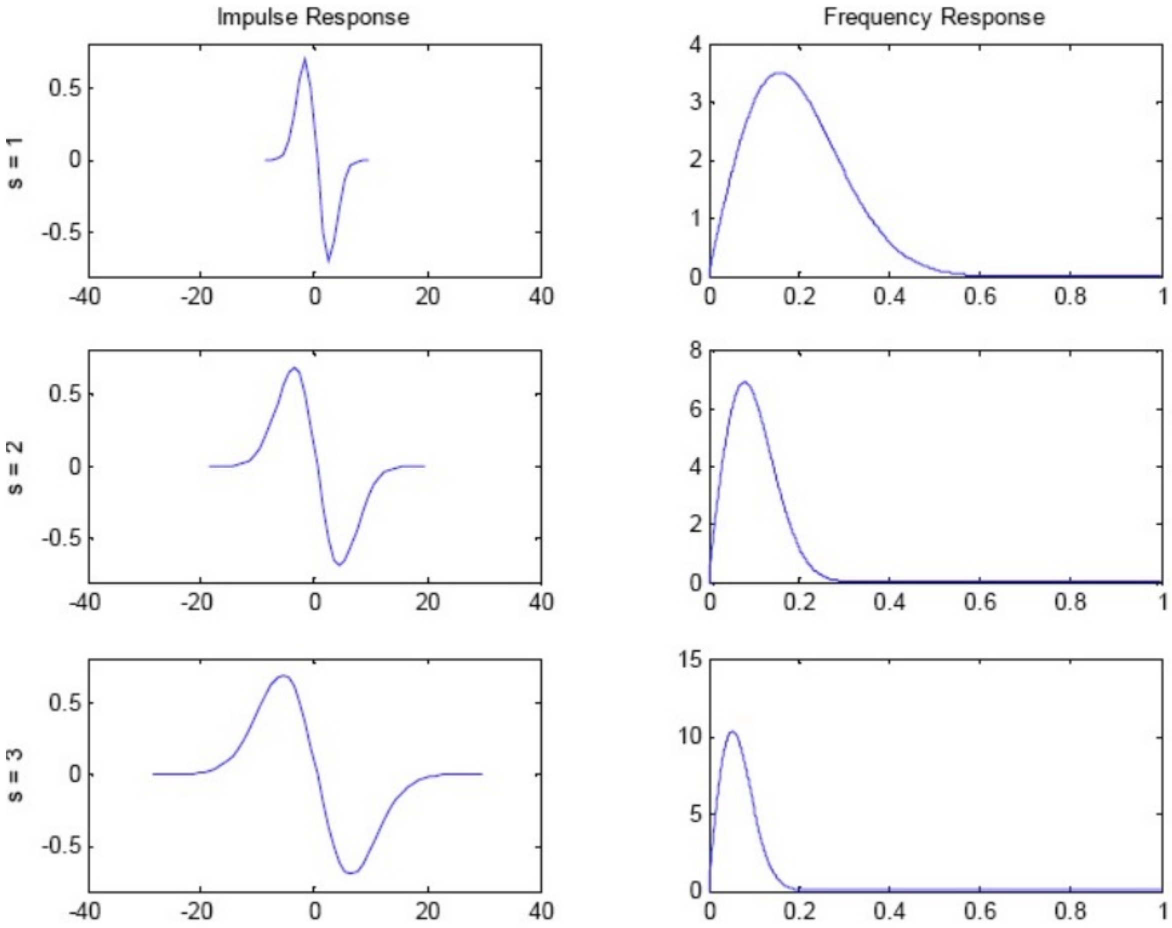
Daubechies (fractal)

Resolutions



Fast changes: low frequency resolution, high time resolution
Slow changes: high frequency resolution, low time resolution

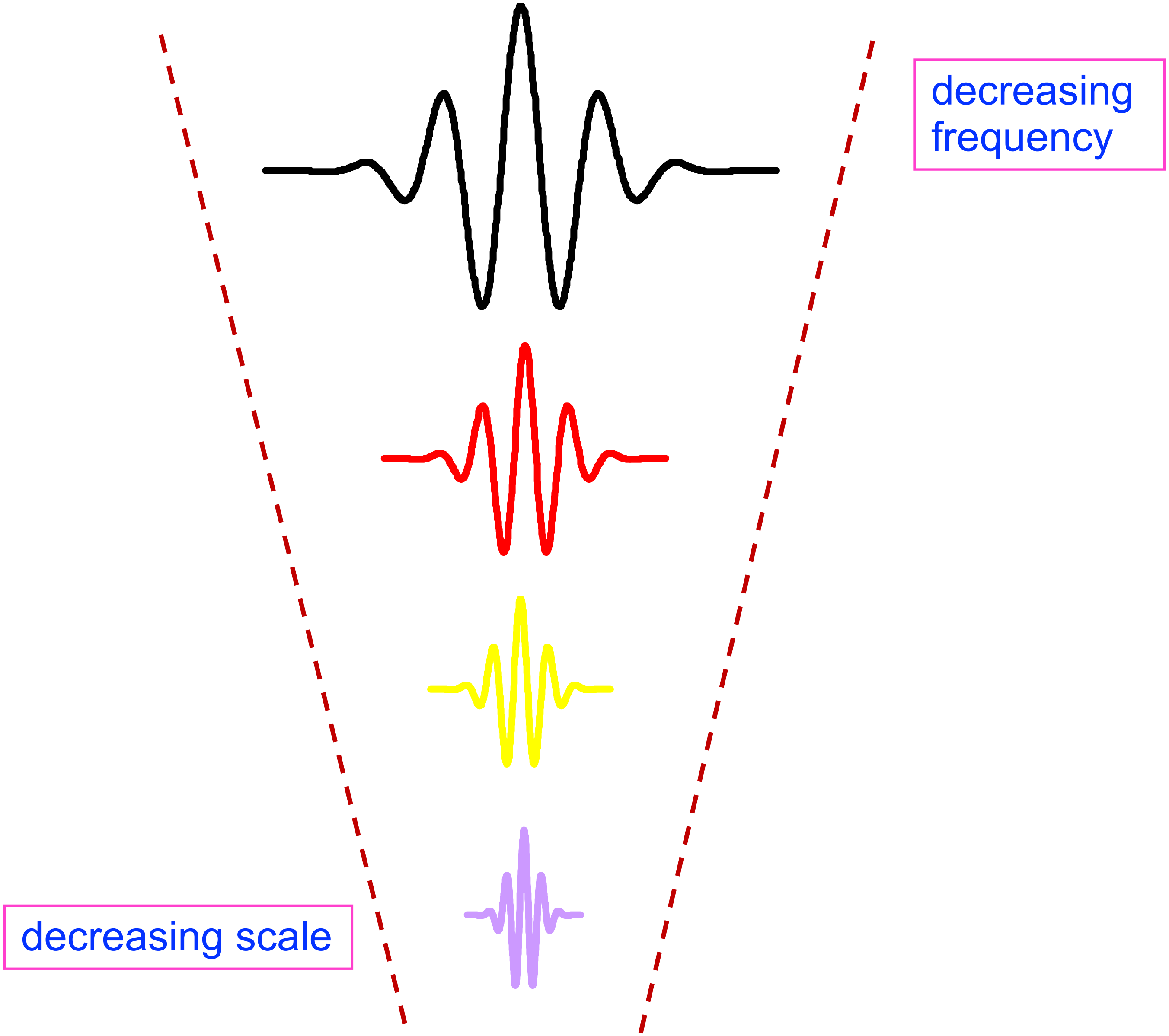
Resolutions



Low (time) scales is equivalent to study low frequency components, i.e. the rough features of the signal

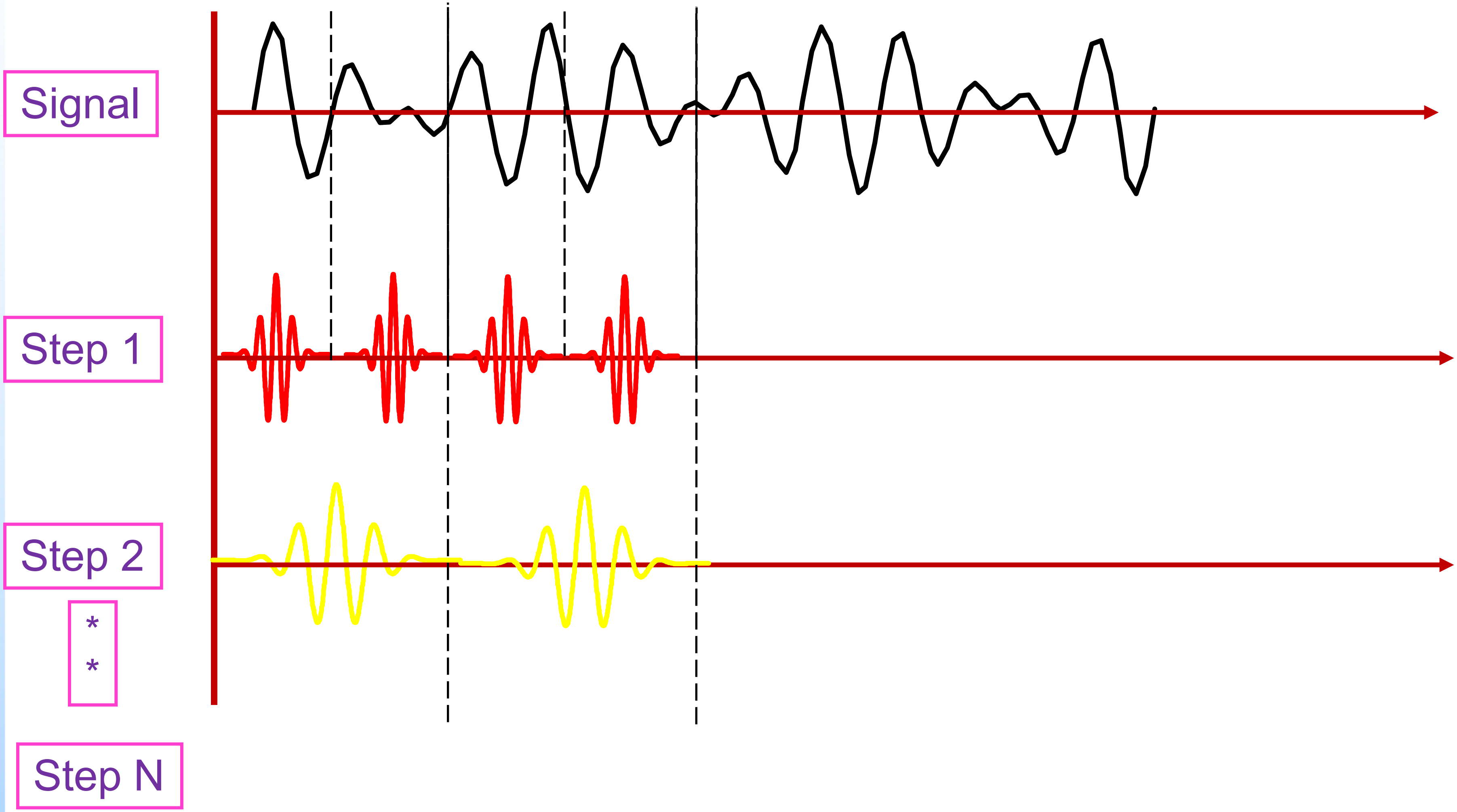
High (time) scales is equivalent to study high frequency components, i.e. the details in the signal

Mother Wavelet


decreasing frequency

decreasing scale

Wavelet Transform



Signal

Step 1

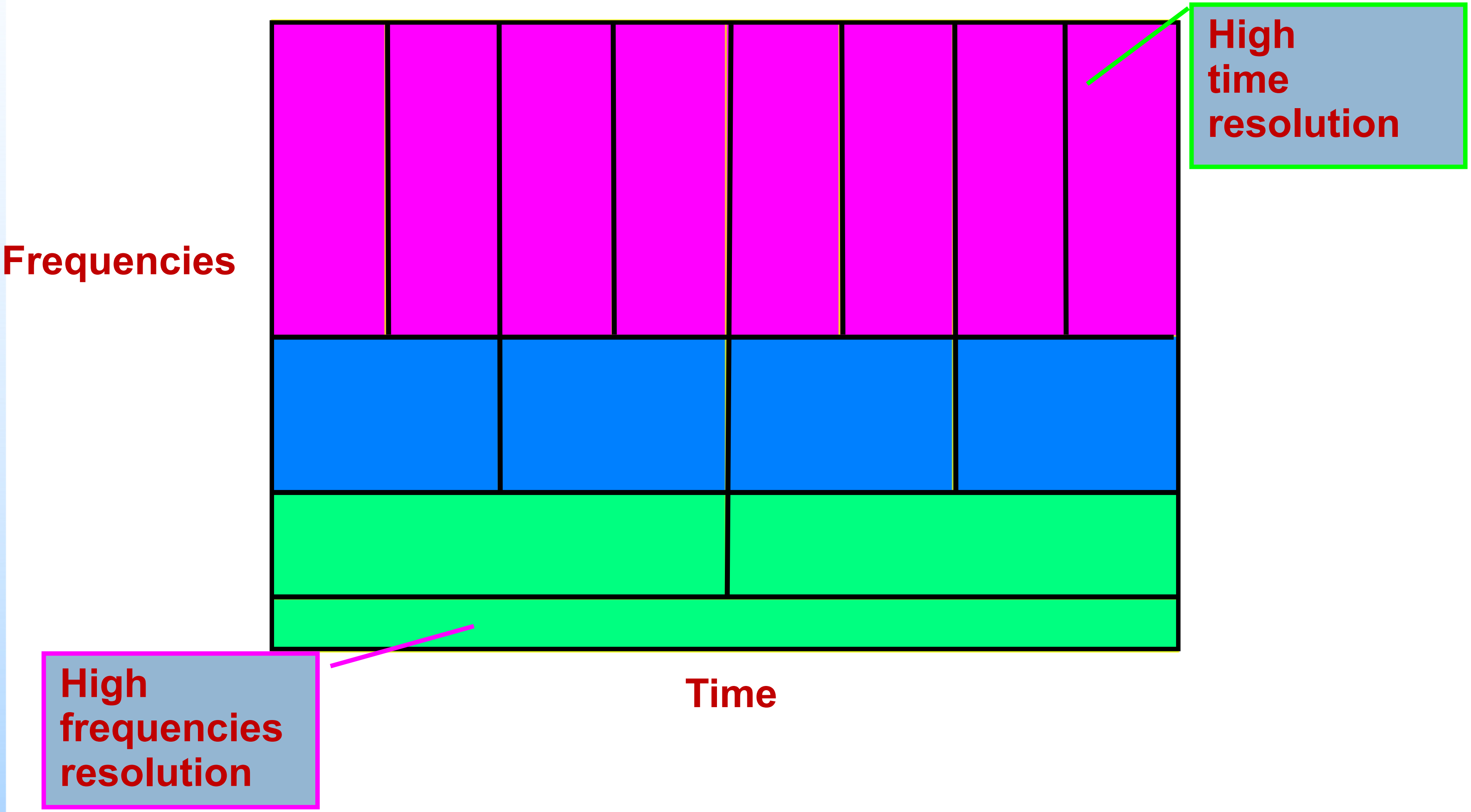
Step 2

*

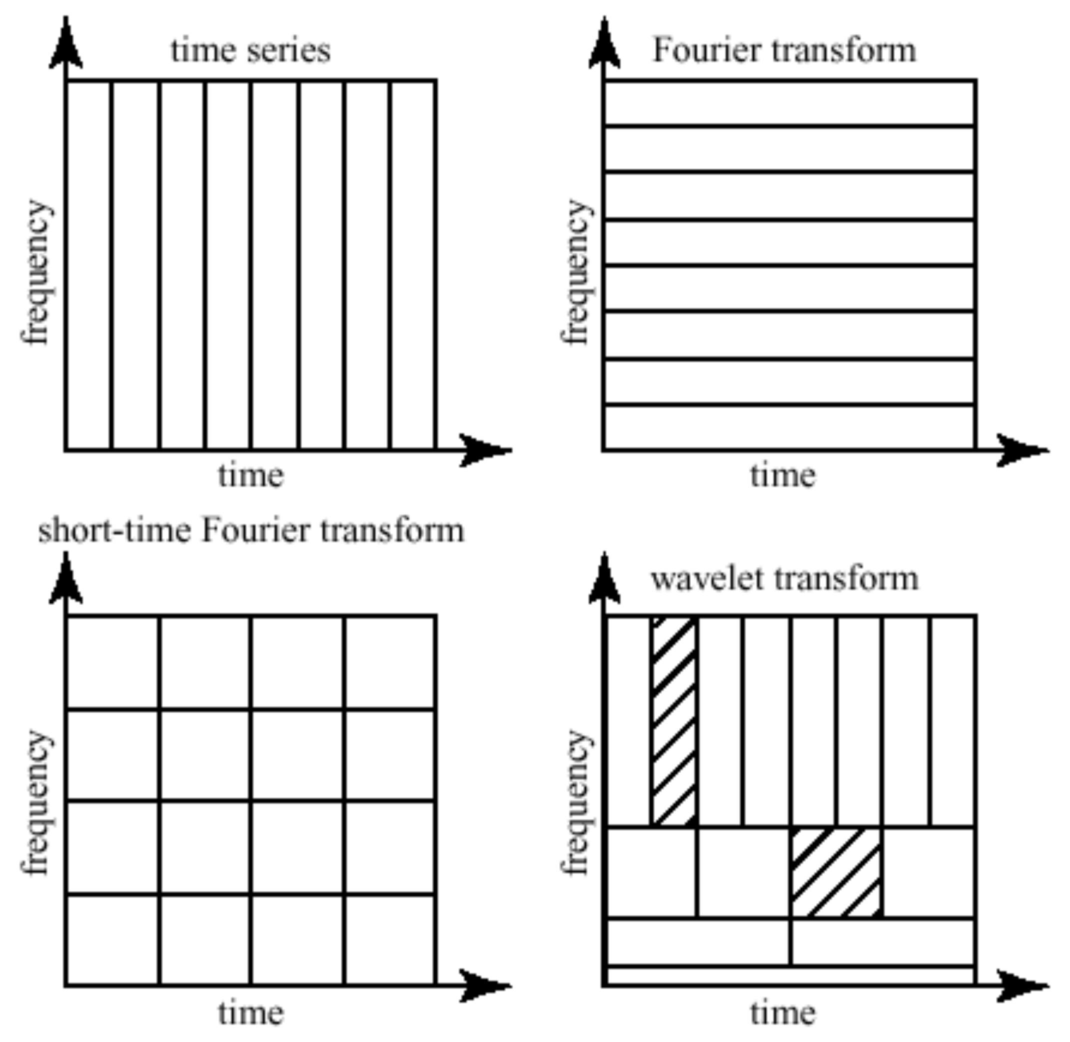
*

Step N

Resolution


Frequencies
Time
High time resolution
High frequencies resolution

STFT vs CWT



From http://www.cerm.unifi.it/EUcourse2001/Gunther_lecturenotes.pdf, p.10

Discrete Wavelet Transform (DWT)

- Sub-bands encoding
  - High pass filters
    - impulse response g[n]
  - Low pass filters
    - impulse response h[n]

DWT

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

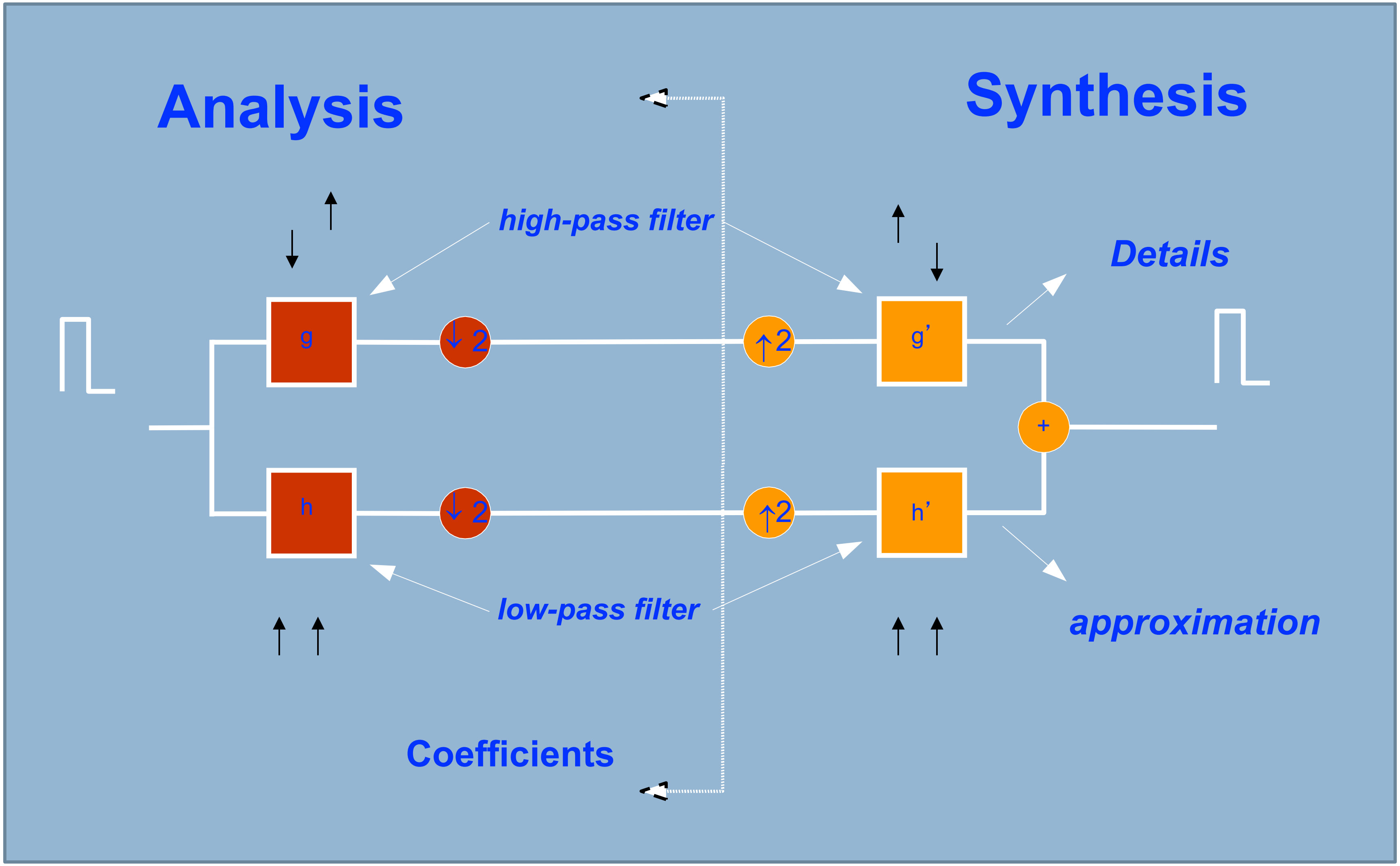
filtering

$$y_H[k]=\sum_{n=-\infty}^{\infty}x[n]g[2k-n]$$

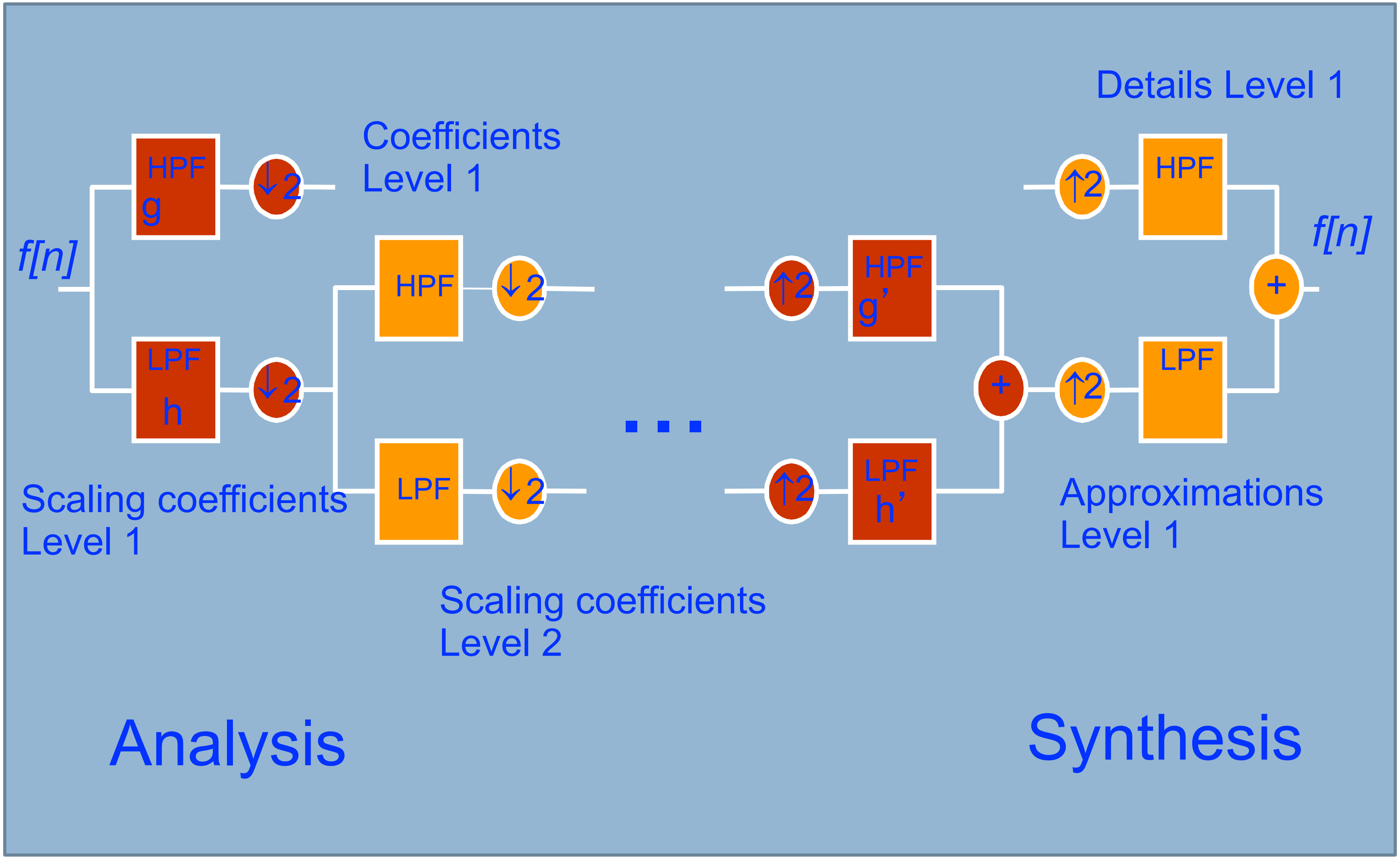
$$y_L[k]=\sum_{n=-\infty}^{\infty}x[n]h[2k-n]$$

Filtering and downsampling

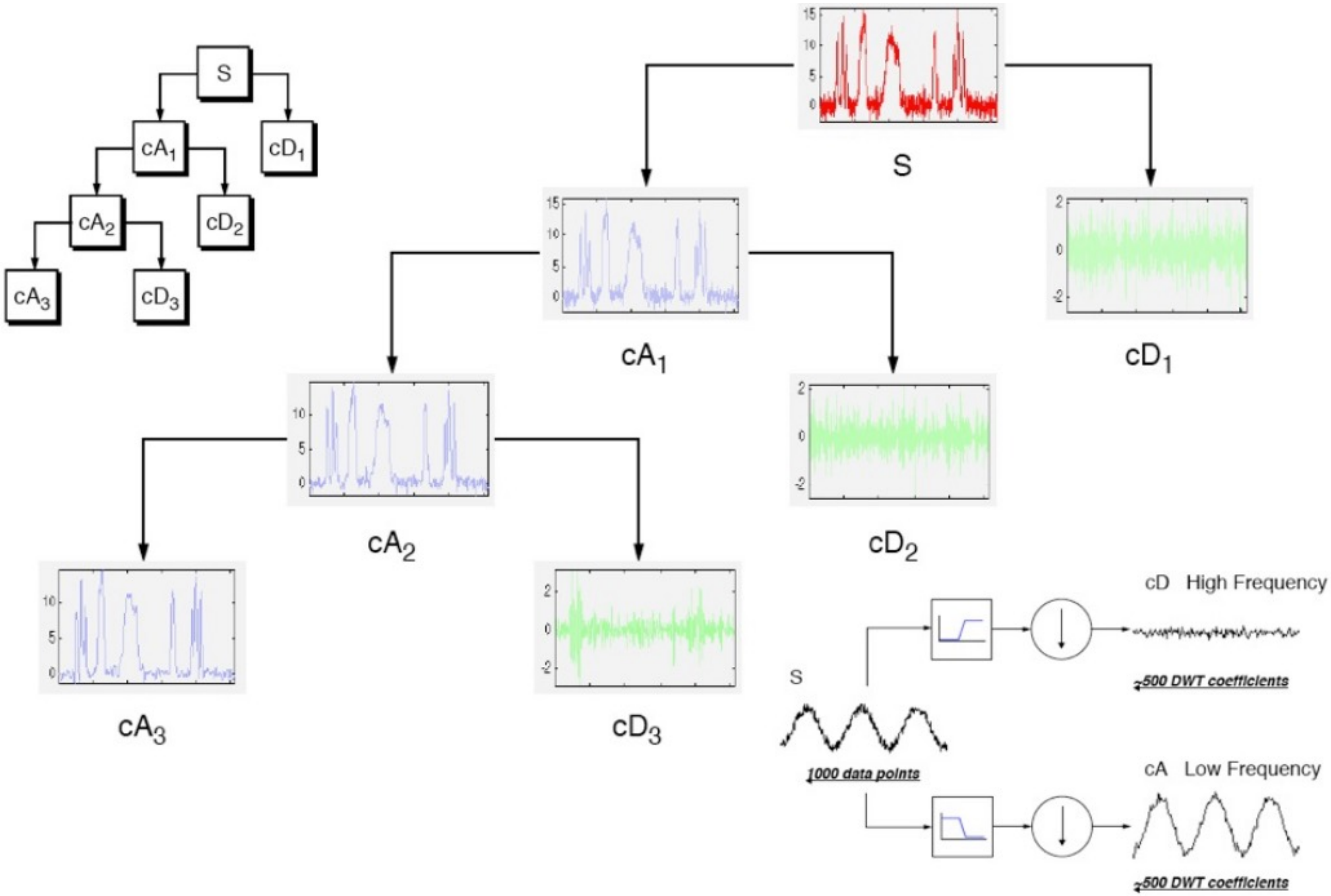
DWT


Analysis
Synthesis
high-pass filter
g
↓2
↑2
g’
Details
+
h
↓2
↑2
h’
low-pass filter
approximation
Coefficients

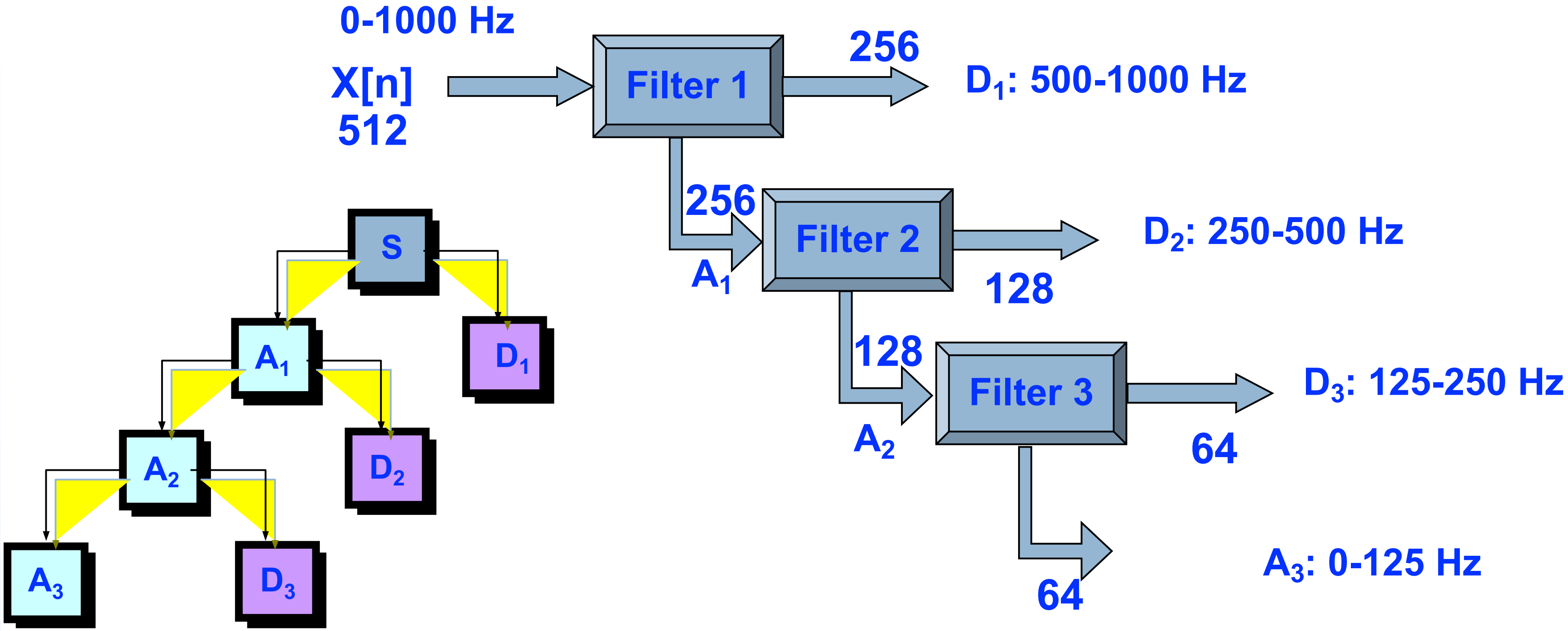
Mallat's algorithm



Mallat's algorithm



Mallat's algorithm



0-1000 Hz

X[n]

512

Filter 1

256

D_1: 500-1000 Hz

256

A_1

Filter 2

D_2: 250-500 Hz

128

128

A_2

Filter 3

D_3: 125-250 Hz

64

64

A_3: 0-125 Hz

S

A_1

D_1

A_2

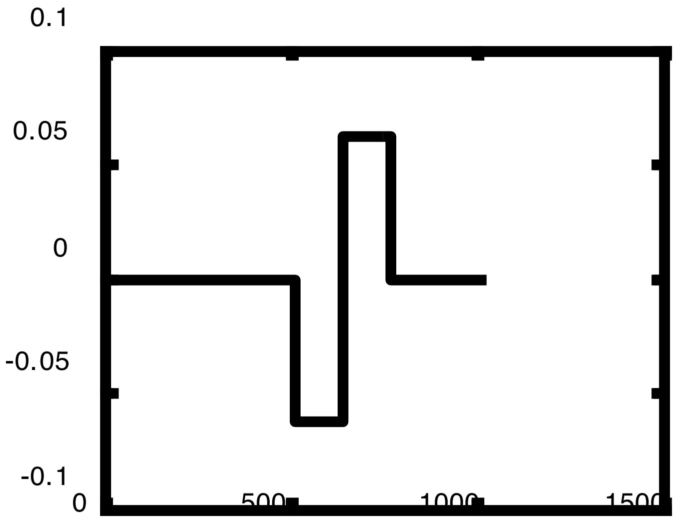
D_2

A_3

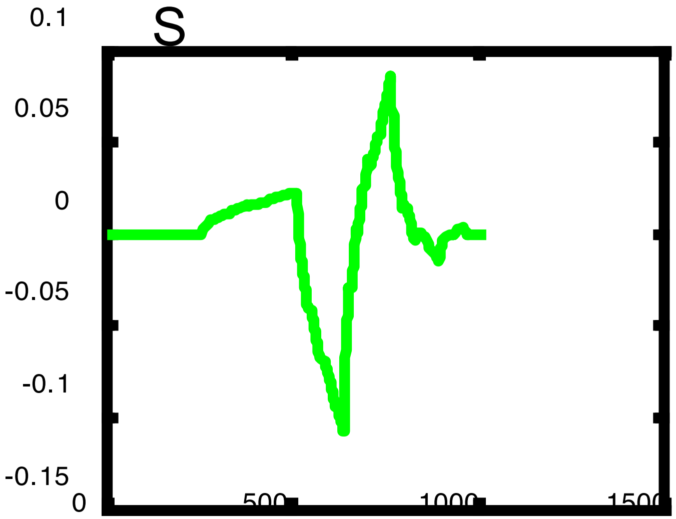
D_3

Wavelets

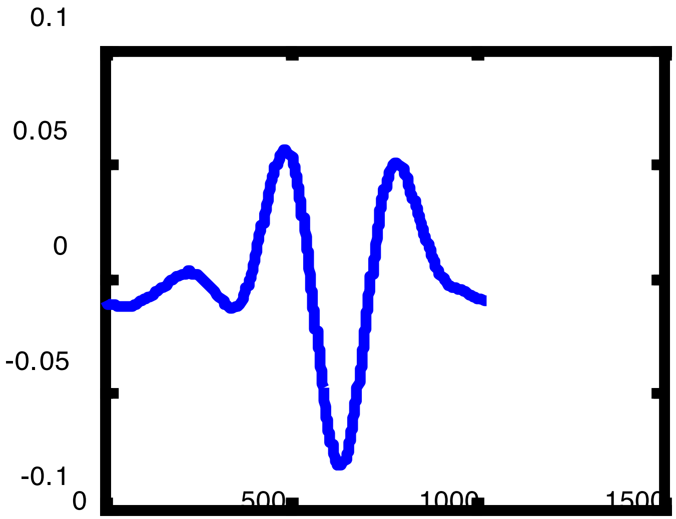
HAAR



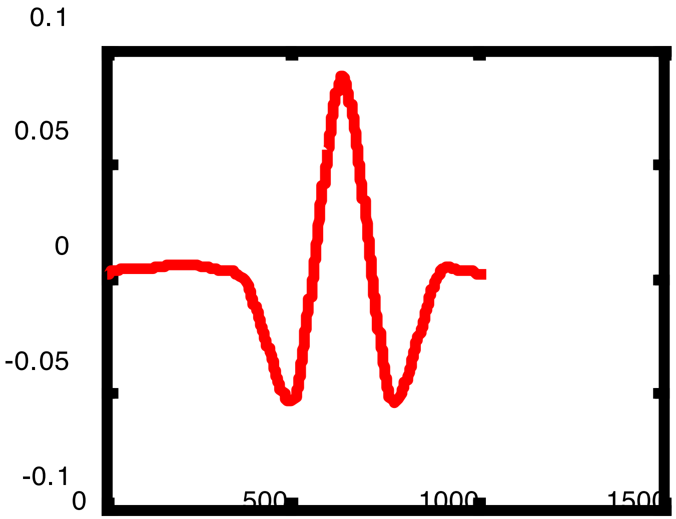
DAUBECHIES



SYMMLET



COIFLET



Example – 1D signal

Data (Size): xn1 (1001)

Wavelet: db 4

Level: 5

Display mode : Full Decomposition

at level: 5

Signal:

0.0-0.4: 2 Hz

0.4-0.7: 10 Hz

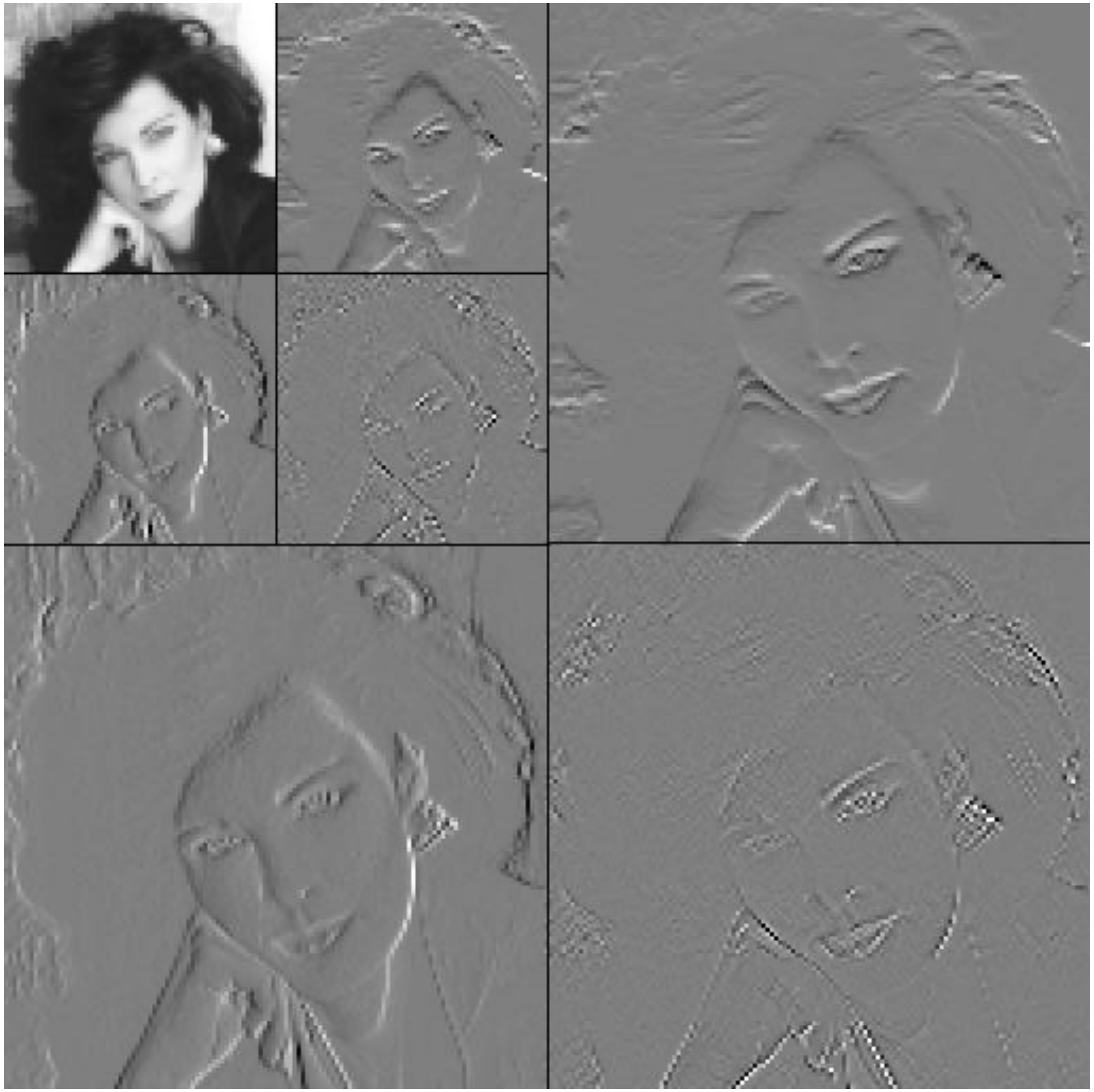
0.7-1.0: 20Hz

Wavelet:

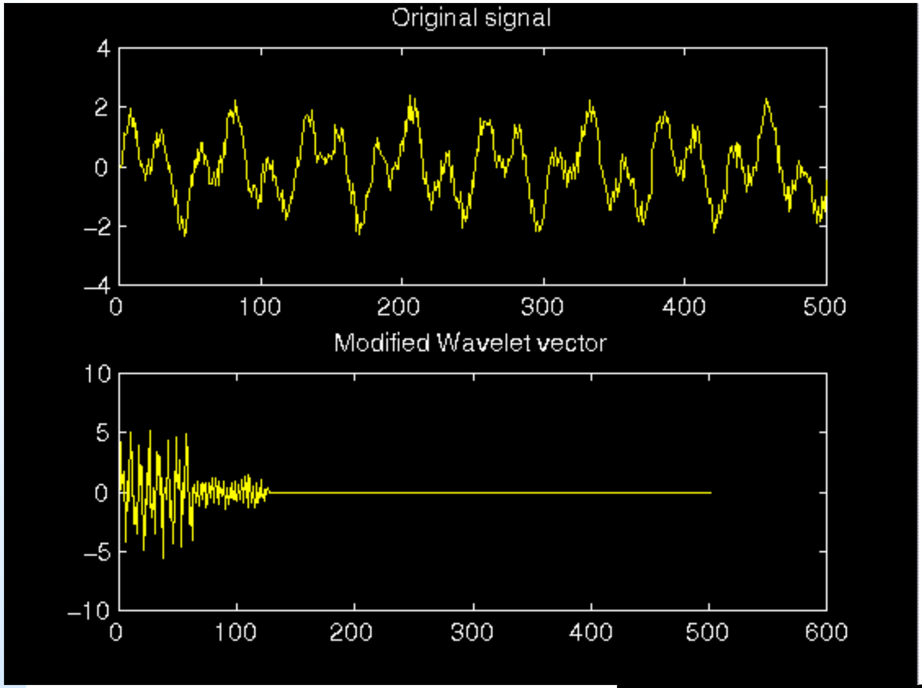
db4

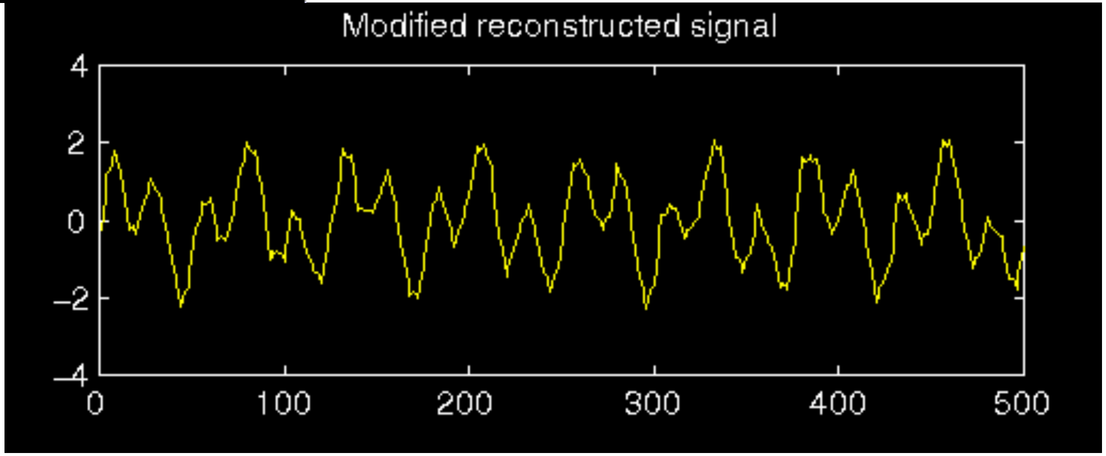
Level: 6

Example – 2D signal

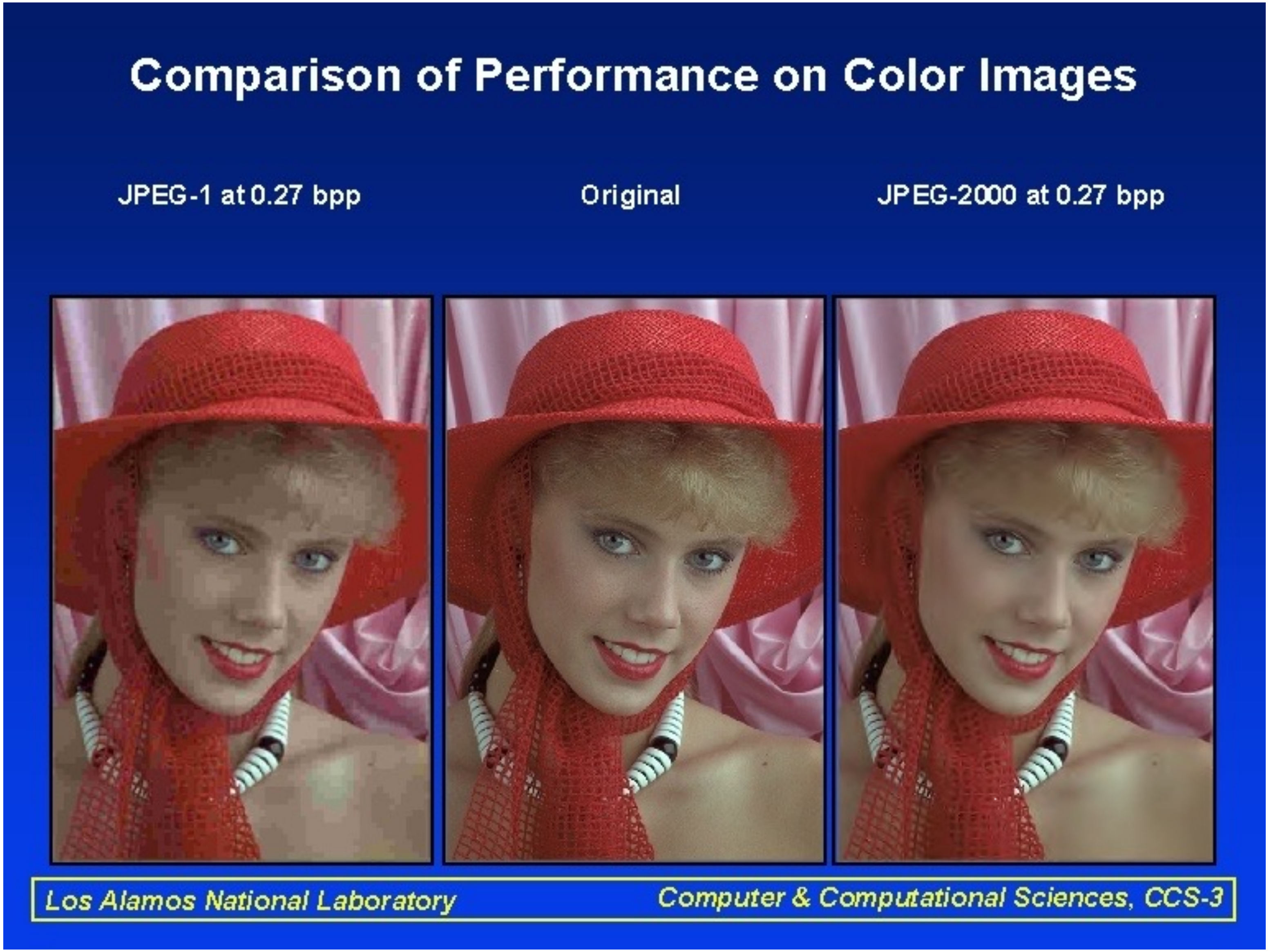


Example – 1D signal denoising





Example – compression



Comparison of Performance on Color Images

JPEG-1 at 0.27 bpp

Original

JPEG-2000 at 0.27 bpp

Los Alamos National Laboratory

Computer & Computational Sciences, CCS-3

References

- Material
  - Slides
  - Video Lessons

- Books
  - Digital Signal Processing, System Analysis and Design, P. S. R. Diniz, E. A. B. da Silva, S. L. Netto, Cambridge University Press, 2012

Question 13

- Content estimation
  - the content of the signals can be estimated by cepstral coefficients

- Question
  - Describe the MFCC and LPC methodologies

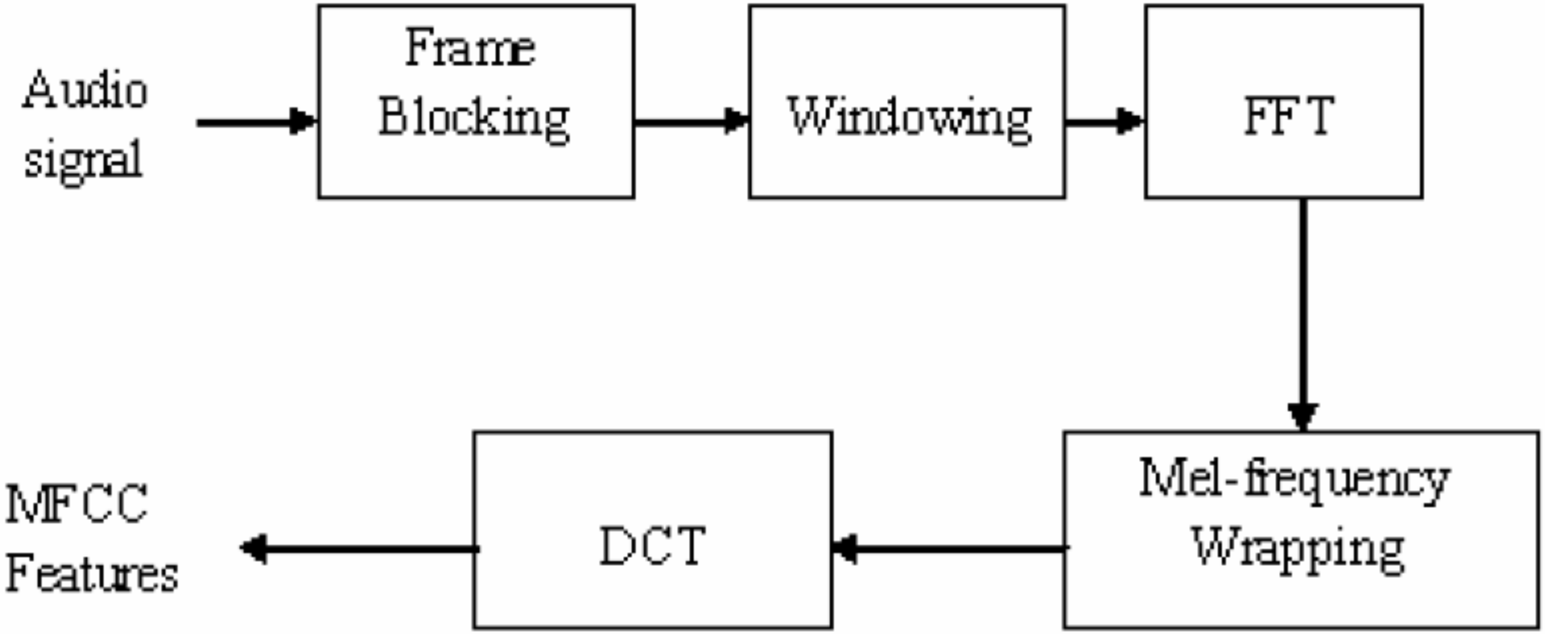
Mel Frequency Cepstral Coefficients

- Mel Frequency Cepstral Coefficients (MFCC)
  - based on perceptual techniques

- Main applications
  - Speech recognition
  - Music information retrieval
  - Musical genre classification

MFCC

Bartlett or Hamming



Block diagram

MFCC wrapping

- Melody scale (mel)
  - proposed by Stevens, Volkman and Newman in 1937
  - based on the non-linear human auditory perception
  - the human hearing system cannot differentiate very close frequencies
  - A 1000 Hz tone at 40 dB corresponds to 1000 mels

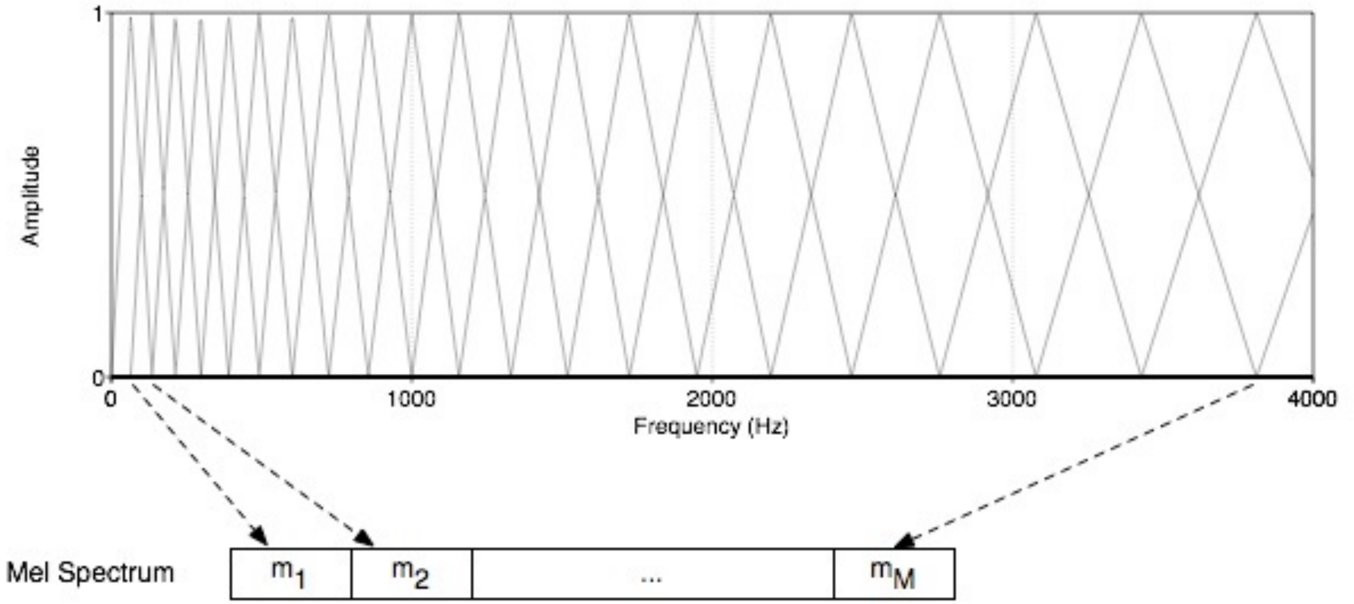
$$\boxed{Mel(f) = 2595\log_{10}\left(1+\frac{f}{700}\right)}$$

MFCC wrapping

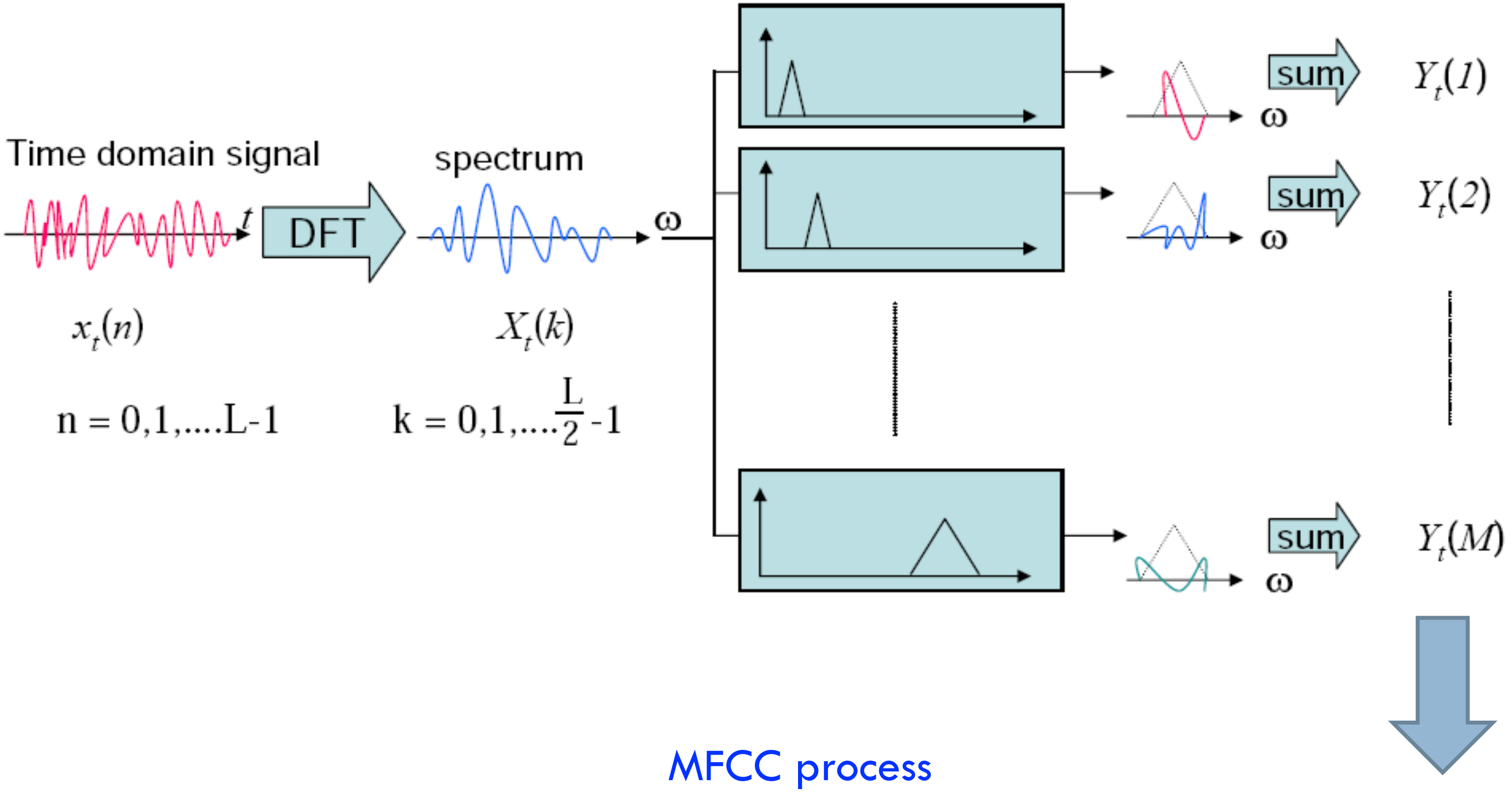
$$BW = 25 + 75\left[1 + 1.4(f_{mel}/1000)^2\right]^{0.69}$$

Band-pass filter

Filters Bank



MFCC process



MFCC process

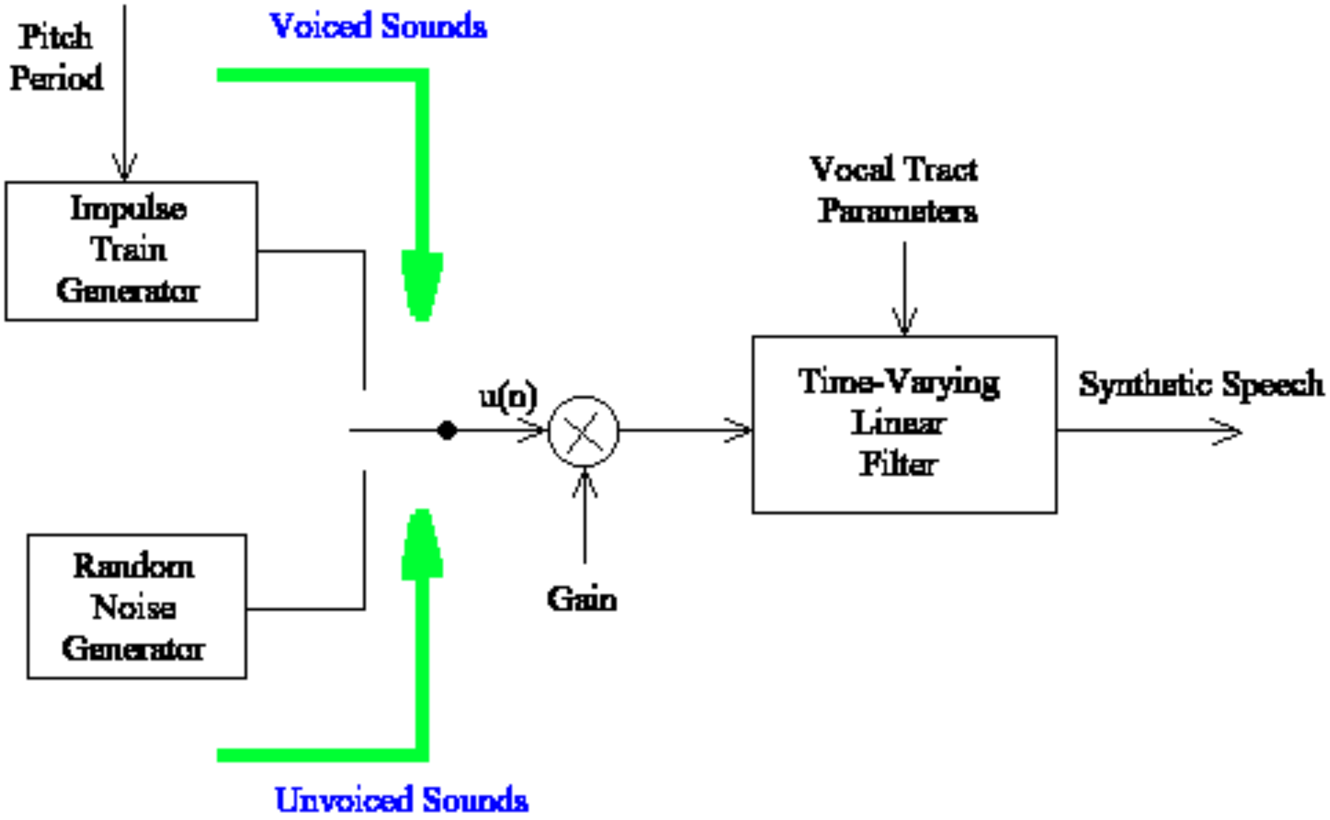
Time domain coefficients by PCA or DCT

Linear Predictive Coding

- Linear Predictive Coding (LPC)
  - Analysis and synthesis of signals
  - Feature extraction
  - Compression
  - Synthesis of the vocal tract

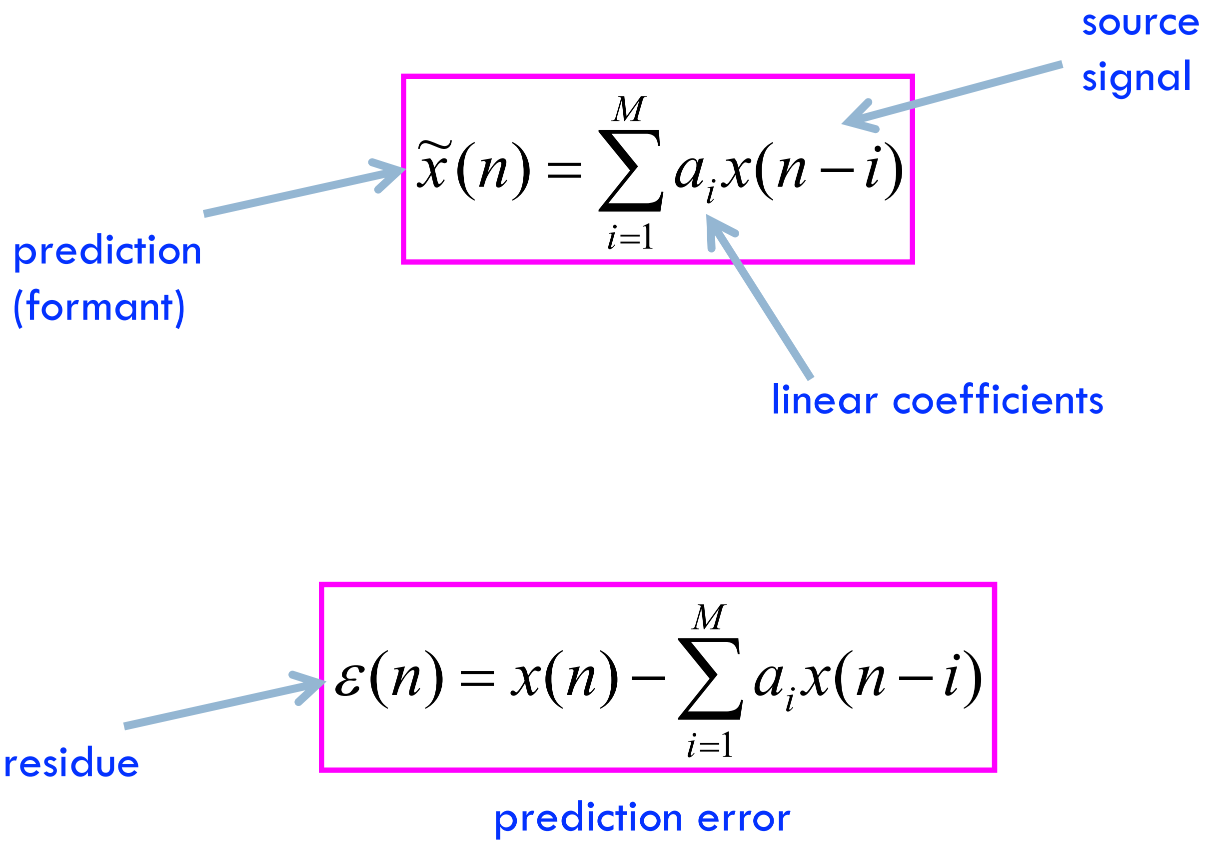
- Voice
  - modulation result caused by the throat and mouth (formant) on the sound emitted by the vocal cords (residue)

LPC



Speech Synthesis model based on LPC model

LPC



source signal

$$\tilde{x}(n) = \sum_{i=1}^{M} a_i x(n-i)$$

prediction (formant)

linear coefficients

$$\varepsilon(n) = x(n) - \sum_{i=1}^{M} a_i x(n-i)$$

residue

prediction error

LPC

- Mean Squared Error (MSE)

$$E = \sum_n \varepsilon(n)^2 = \sum_n \left(x(n) - \sum_{i=1}^{M} a_i x(n-i) \right)^2$$

coefficients to estimate

- Optimization
  - Autocorrelation method (N^3)
    - QR decomposition
    - Gauss elimination
  - Levison-Durbin Algorithm (N^2)

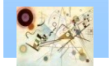
References

- Material
  - Slides
  - Video Lessons

- Books
  - Digital Signal Processing, System Analysis and Design, P. S. R. Diniz, E. A. B. da Silva, S. L. Netto, Cambridge University Press, 2012

Question 14

- Filtering
  - the frequency of the signal can be filtered by adaptive methodologies
- Question
  - Describe the Adaptive Filters

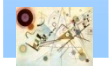


Adaptive filters

- Adaptive filter
  - The parameters are estimated
    - learning algorithm
  - An error function is used
    - e.g., Linear Artificial Neural Network (Adaline)

Adaptive filters

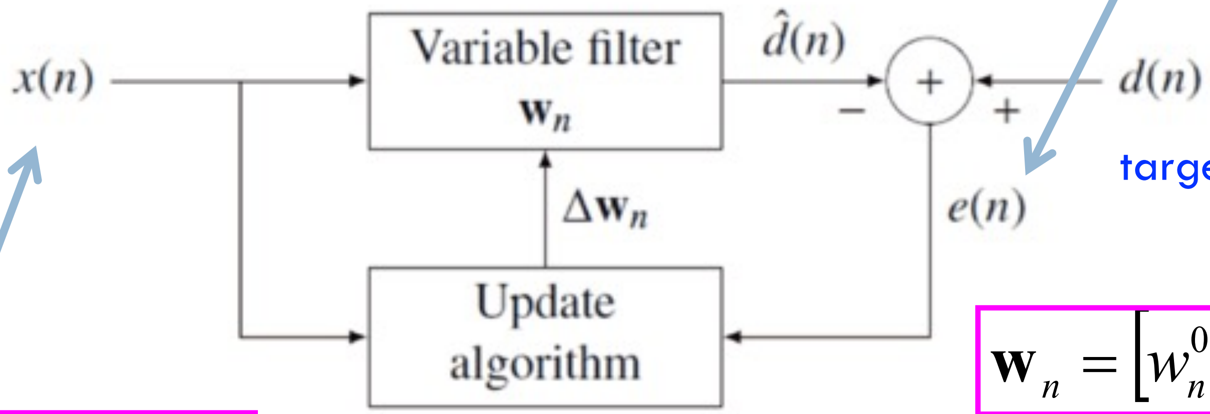
- Hospital
  - ECG (electrocardiogram) corrupted by noise at 50 Hz (electricity)
  - The current can vary between 47 Hz and 53 Hz
  - A filter for the elimination of static noise at 50 Hz could give errors
  - An adaptive filter can learn from the current shape of noise
- Helicopter
  - Pilot speaking with noise from rotating propeller
  - The noise has not a spectrum well defined
  - An adaptive filter learns the shape of the noise
  - The noise can be subtracted from the signal for only the pilot's voice



Adaptive filters

$$e(n) = d(n) - \hat{d}(n)$$

input



target signal

$$\mathbf{w}_n = \left[w_n^0, w_n^1, \ldots, w_n^p\right]$$

$$x(n) = d(n) + v(n)$$

FIR structure

$$\widetilde{d}(n) = \mathbf{w}_n \cdot \mathbf{x}(n)$$

noise

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \Delta\mathbf{w}_n$$

$$\mathbf{x}(n) = \left[x(n), x(n-1), \ldots, x(n-p)\right]$$

Learning algorithm (e.g., LMS and descent gradient)

source signal

References

- Material
  - Slides
  - Video Lessons

- Books
  - Digital Signal Processing, System Analysis and Design, P. S. R. Diniz, E. A. B. da Silva, S. L. Netto, Cambridge University Press, 2012

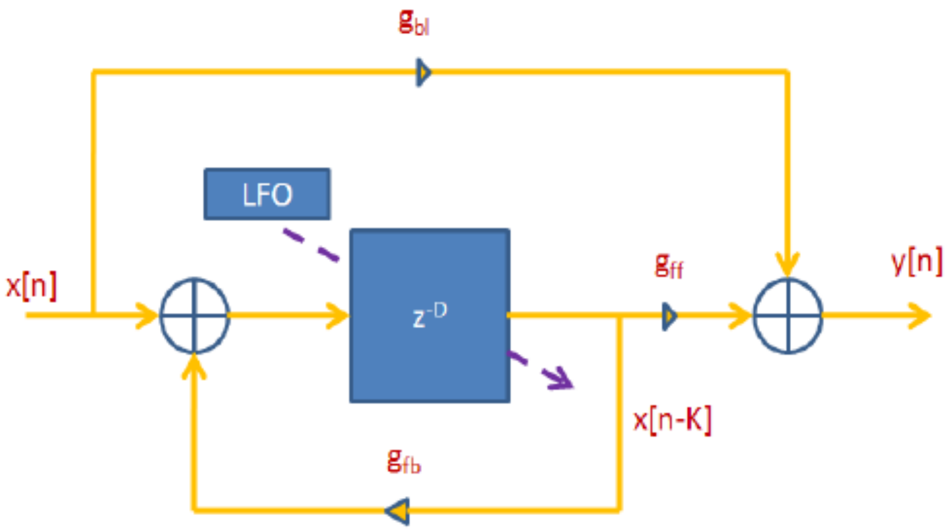
Question 15

- Effects
  - are correlated with the modifications of an acoustic signal

- Question
  - Describe the general model for effects

Effects

- General model for effects



$$
H(z)=\frac{g_{bl}+g_{ff}z^{-D[n]}}{1+g_{fb}z^{-K}}
$$

Vibrato

- From the general model eliminating feedback and blending

- Delay less than 5 ms

- Size of delay line

$$\boxed{D_1 = \frac{f_c}{2f_0} m}$$

Flanging

- From the general model eliminating feedback and blending

- Delay in the range 1 ms – 10 ms

- Equations

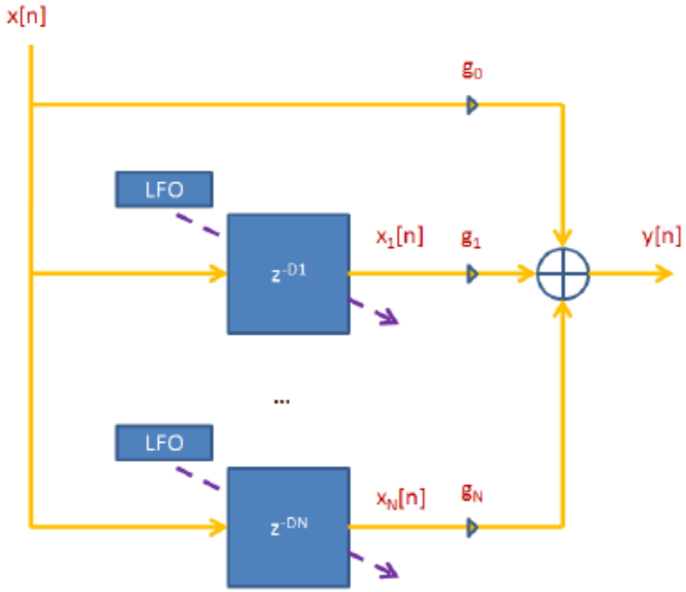
$$\boxed{y[n] = x[n] + g_{ff} x[n - D[n]]}$$

$$\boxed{D[n] = D_0 + D_1 \sin(2\pi f_{FL} n)}$$

ISP – Verification tests

Chorus

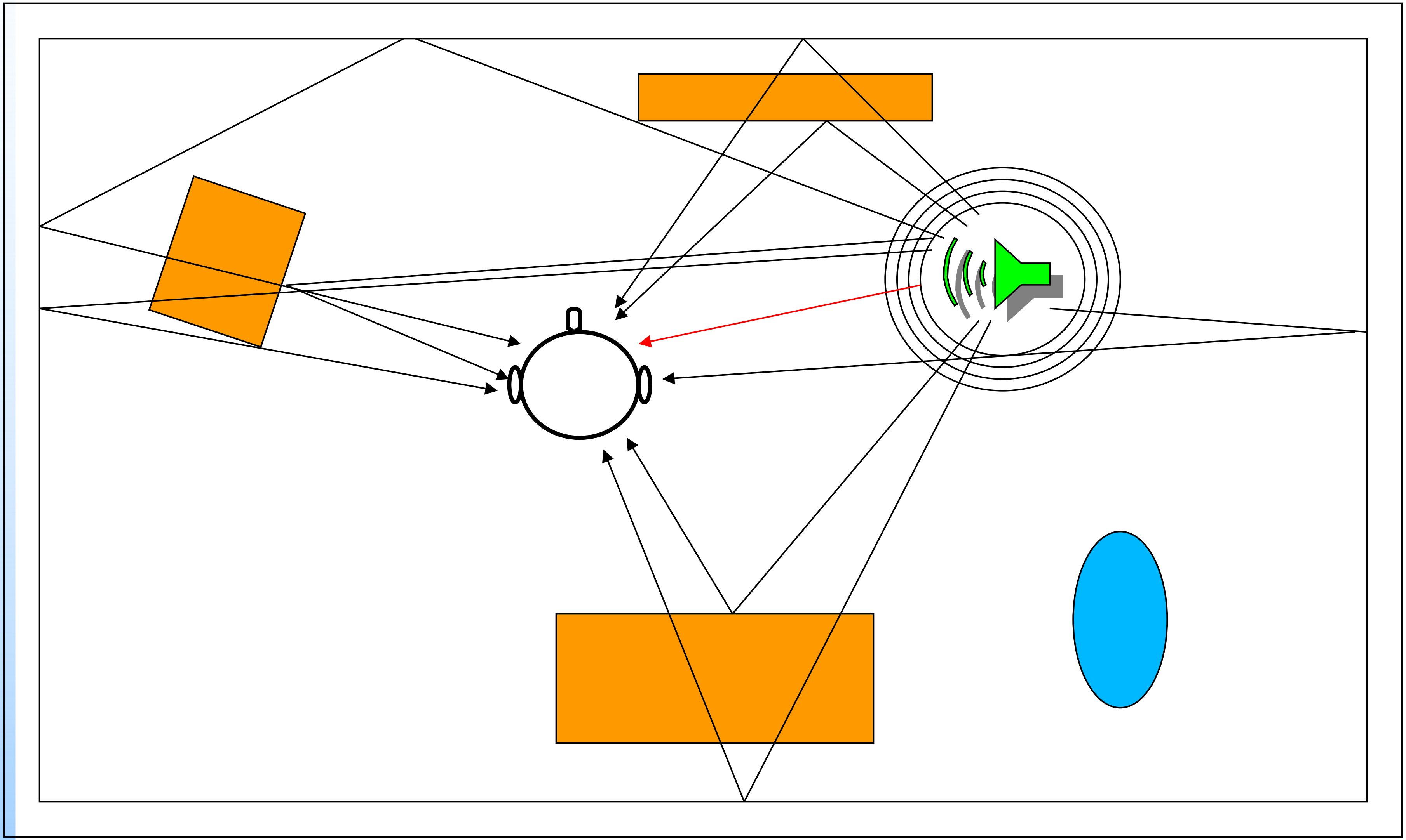
- At least two voices



Effects

	g_{bl}	g_{ff}	g_{fb}	Onset	Depth	Modulazione
Vibrato	0.0	1.0	0.0	0 ms	0-5ms	0.1-5 Hz sinusoidale
Flanger	0.707	0.707	-0.707	0ms	1-10 ms	0.1- 1 Hz sinusoidale
Chorus	1.0	0.707	0.0	1-30 ms	5-30 ms	Lowpass noise
White chorus	0.707	1.0	0.707	1-30 ms	5-30 ms	Lowpass noise
Doubling	0.707	0.707	0.0	10-100 ms	1-100 ms	Lowpass noise
Eco	1.0	≤ 1.0	< 0	50-∞	80-∞	-

Reverberation



Reverberation

- A simple approach is based on the convolution of the room impulse response

$$\boxed{y(n) = h_i \bigotimes x(n)}$$

- A more sophisticated methodology is based on a perspective approach

References

- Material
  - Slides
  - Video Lessons

- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Audio digitale, A. Uncini, McGraw-Hill Education, 2006
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 16

- A waveform of the sound to be generated is computed by using models

- Question
  - Describe the general model for audio synthesis

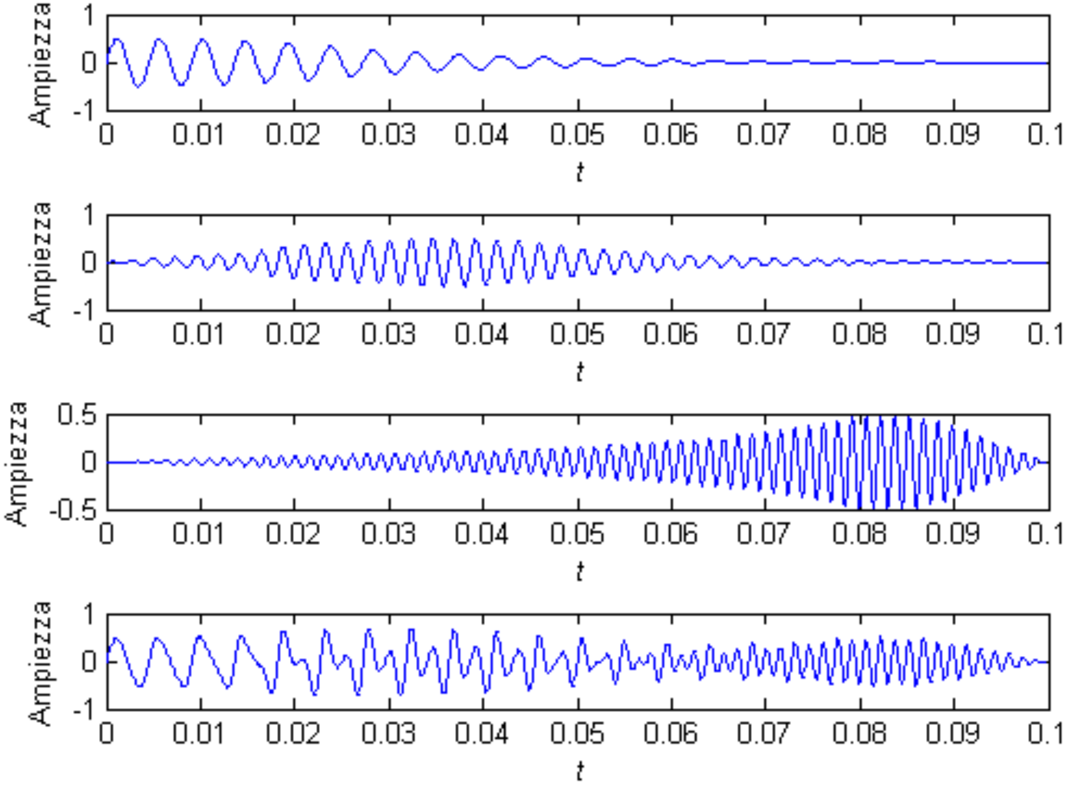
Sound Synthesis

- A waveform of the sound to be generated is computed by using models

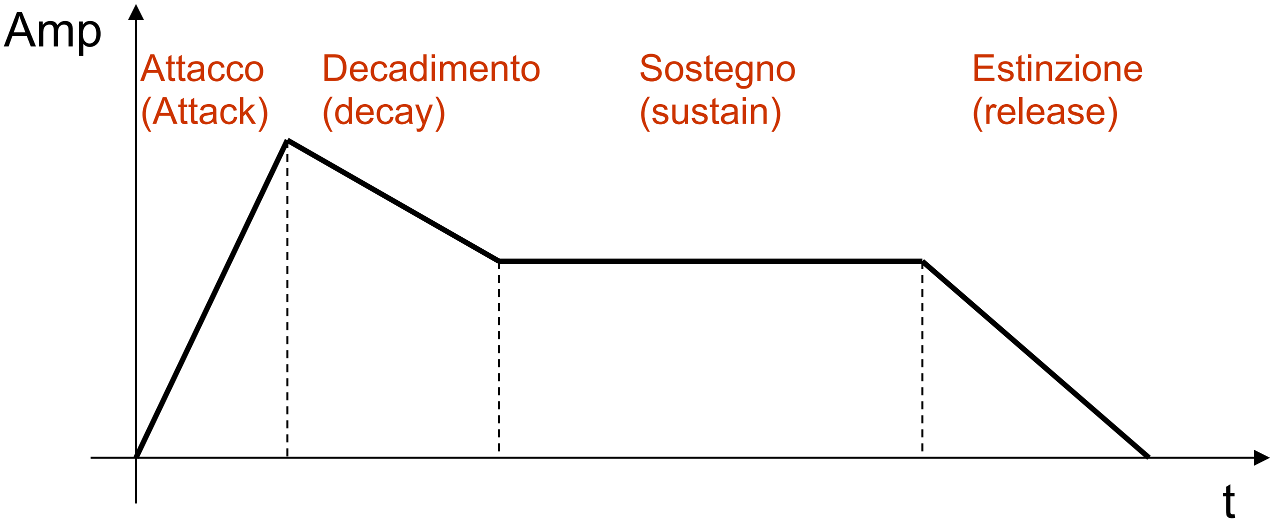
- Some approaches
  - Additive synthesis
  - Physical modelling synthesis

Additive syntheis

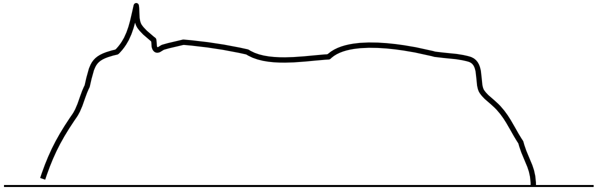
- Directly from the Fourier Theorem



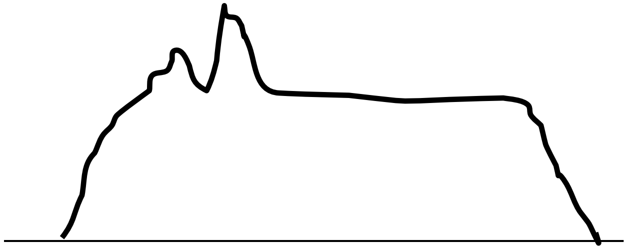
Envelope


Amp
Attacco
(Attack)
Decadimento
(decay)
Sostegno
(sustain)
Estinzione
(release)
t

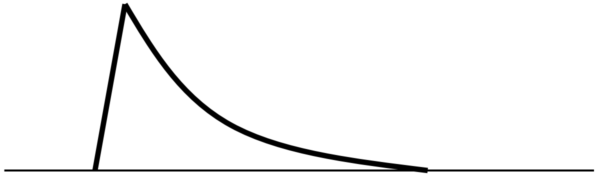
Envelope



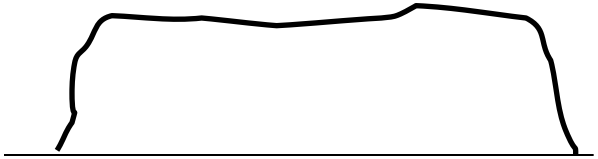
Flute



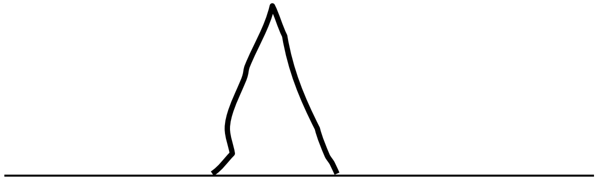
Trumpet



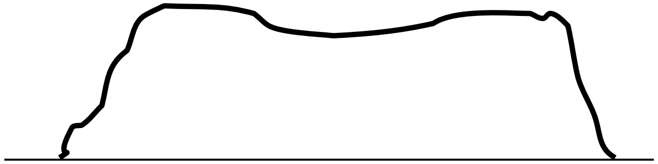
Piano



Violin

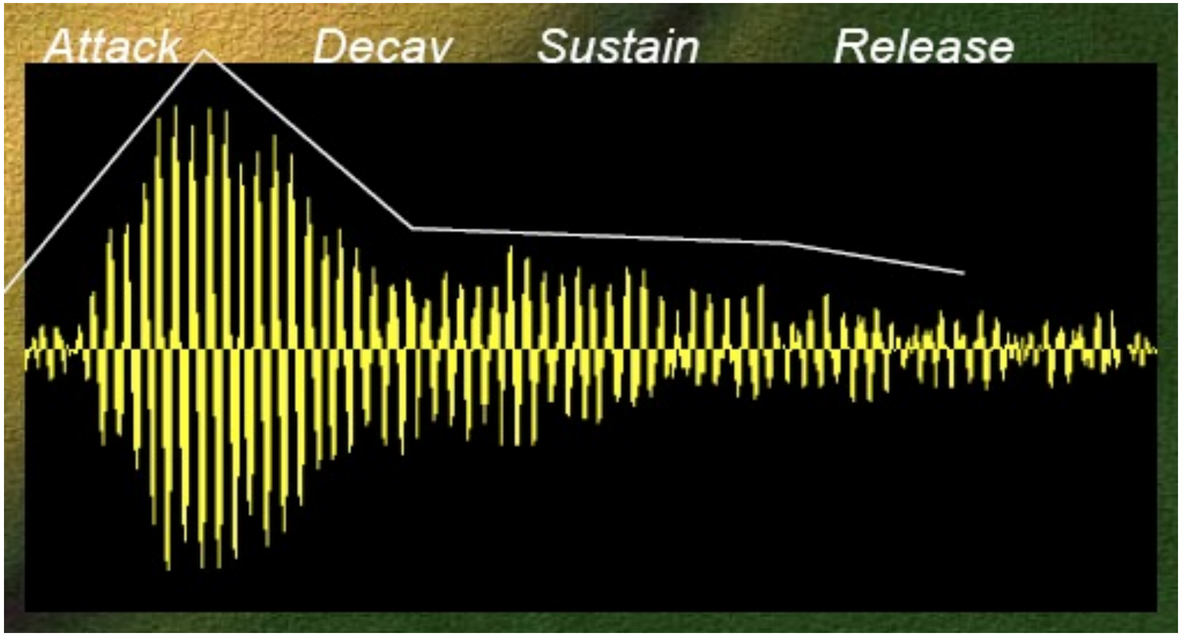


Wood blocks

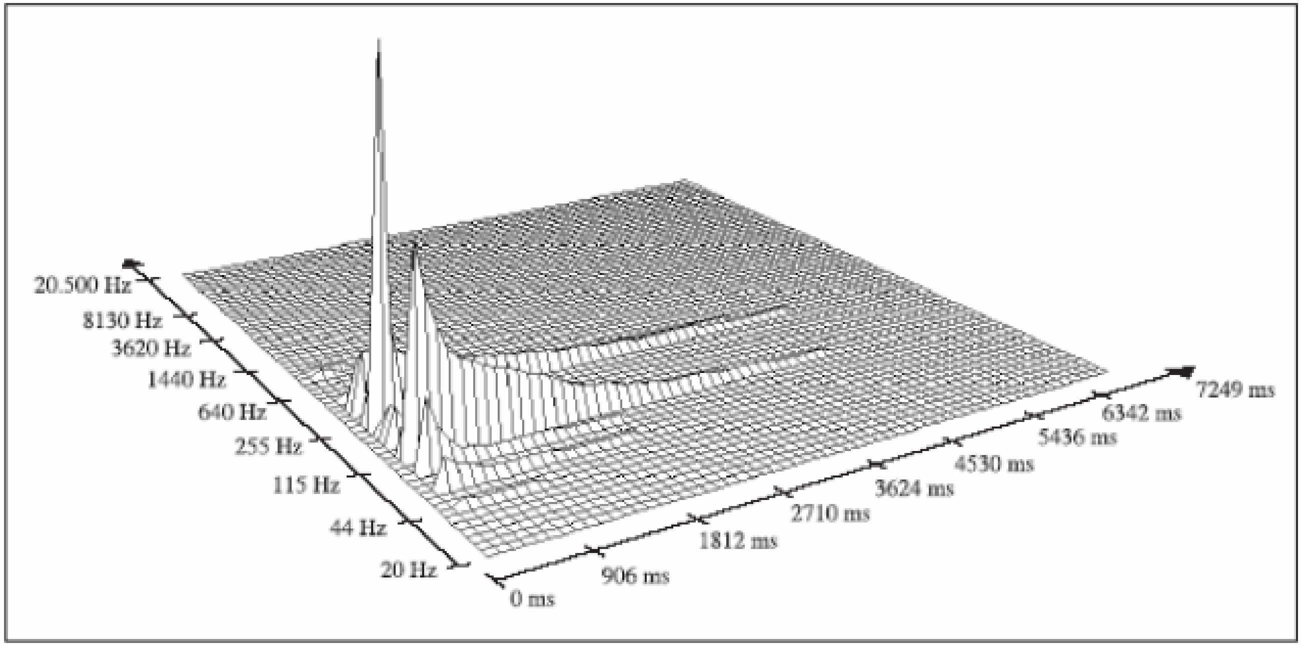


Contrabass

Envelope



Spectogram


20.500 Hz
8130 Hz
3620 Hz
1440 Hz
640 Hz
255 Hz
115 Hz
44 Hz
20 Hz
0 ms
906 ms
1812 ms
2710 ms
3624 ms
4530 ms
5436 ms
6342 ms
7249 ms

Physical modelling synthesis

- A mathematical model is used
  - equations or algorithms to simulate a physical source of sound
    - usually a musical instrument

- Methodology
  - Karplus-Strong algorithm

Karplus-Strong algorithm

- Karplus-Strong method
  - use the delay line

$$\boxed{y[n] = x[n] + R^L y[n-L]}$$

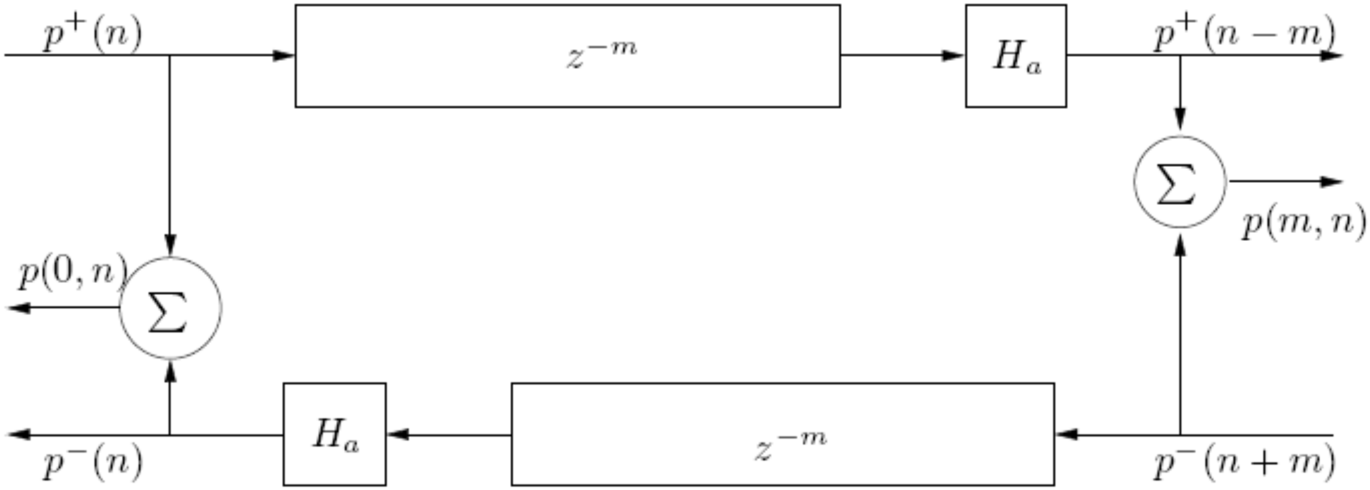
Difference equation

Comb filter

$$\boxed{H(z) = \frac{1}{1-R^L z^{-L}} = \frac{1}{z^L - R^L}}$$

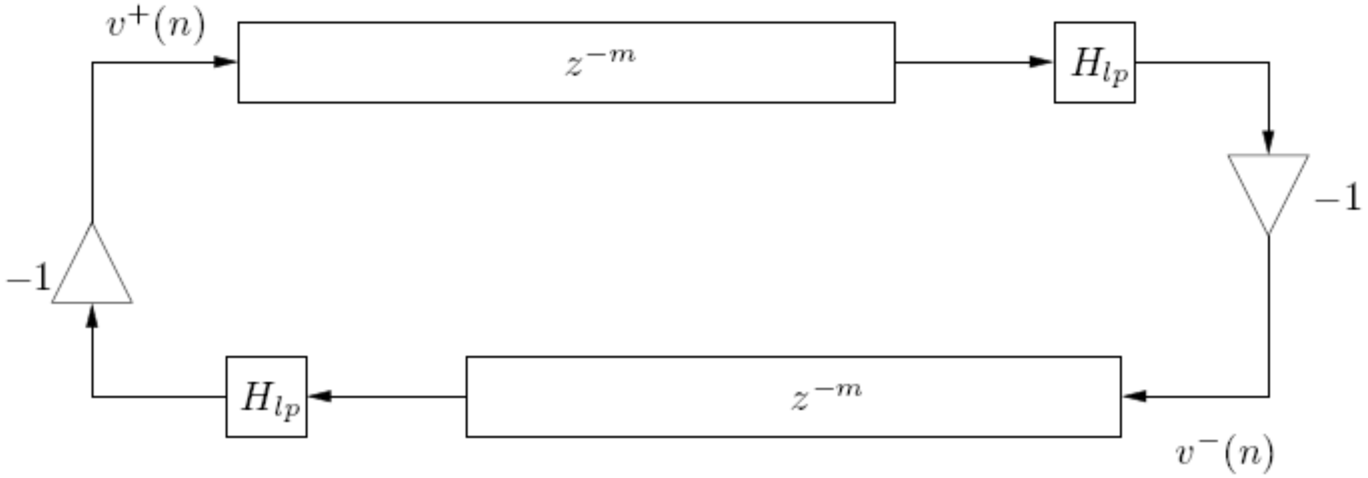
Transfer fuction

Chord of an acoustic guitar (A - 440Hz)



Using waveguides

Chord of an acoustic guitar (A - 440Hz)



Ideal chord with dissipation

$$\boxed{L = \frac{lF_s}{\sqrt{\frac{T}{m/l}}}}$$

Lenght of the delay line considering physical parameters

T – Tension of the chord

M – mass

l – length

F_s – sampling frequency

References

- Material
  - Slides
  - Video Lessons

- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Audio digitale, A. Uncini, McGraw-Hill Education, 2006
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 17

- To be transmitted over the network multimedia content must be digitized and compressed

- Question
    - Describe the PCM based compression approaches

Introduction

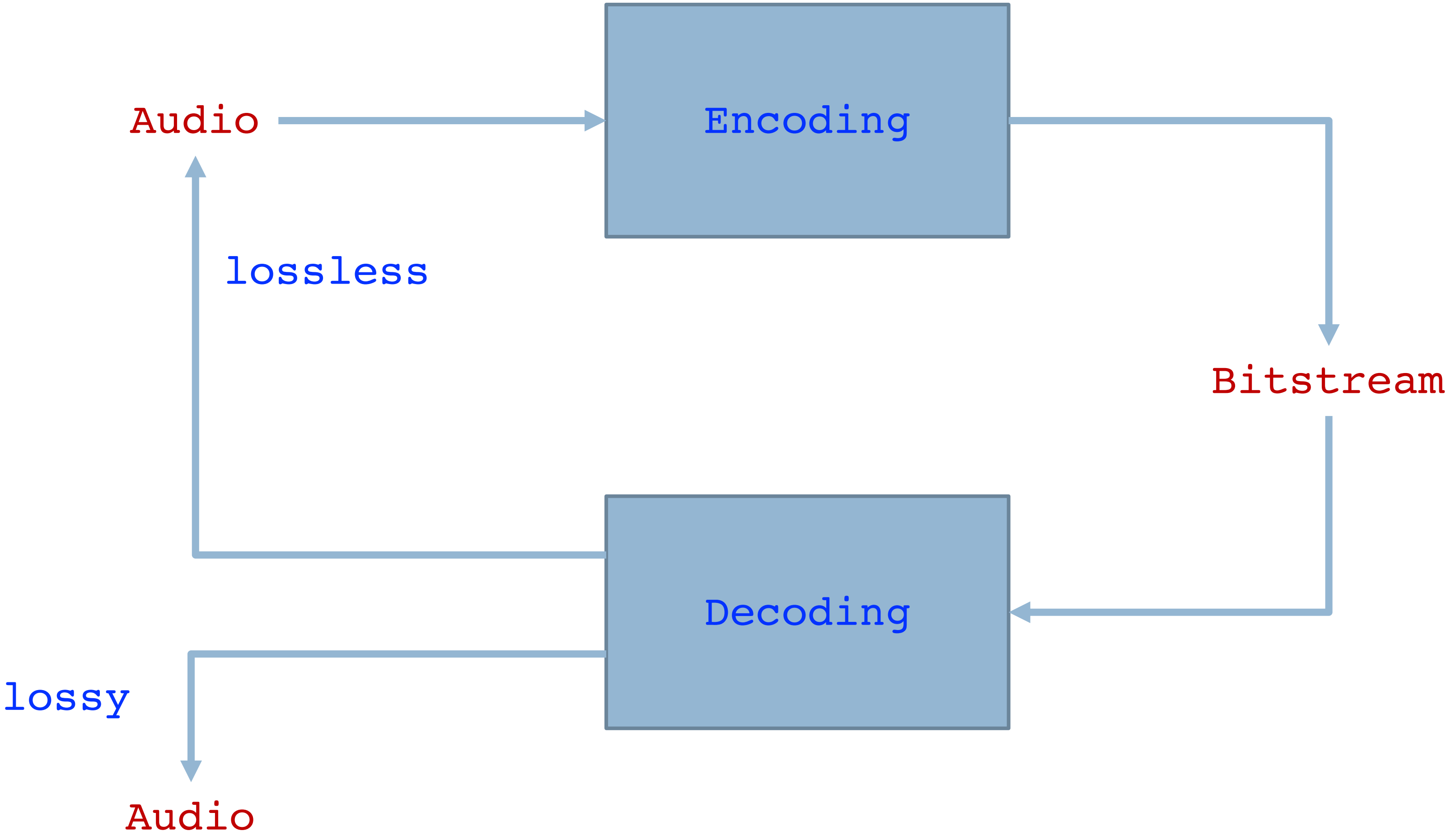
- To be transmitted over the network multimedia content
  - must be digitized and compressed
- Image
  - uncompressed 1024 x 1024 image
    - 8 bits for each color (RGB)
  - 3 Mbyte of memory
  - the transmission on a 64 Kbps channel needs of 7 minutes

Pulse Code Modulation

- Pulse Code Modulation (PCM)
  - sampling frequency
  - quantization bits

- Examples
  - voice
    - 8000 Hz
    - 8 bits
  - CD audio
    - 44100 Hz
    - 16 bits

Encoding-Decoding



Audio

Encoding

Bitstream

Decoding

lossless

lossy

Audio

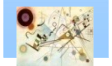
μ-law and A-law compressions

- μ-law coding
  - North America and Japan
  - digital phone on ISDN

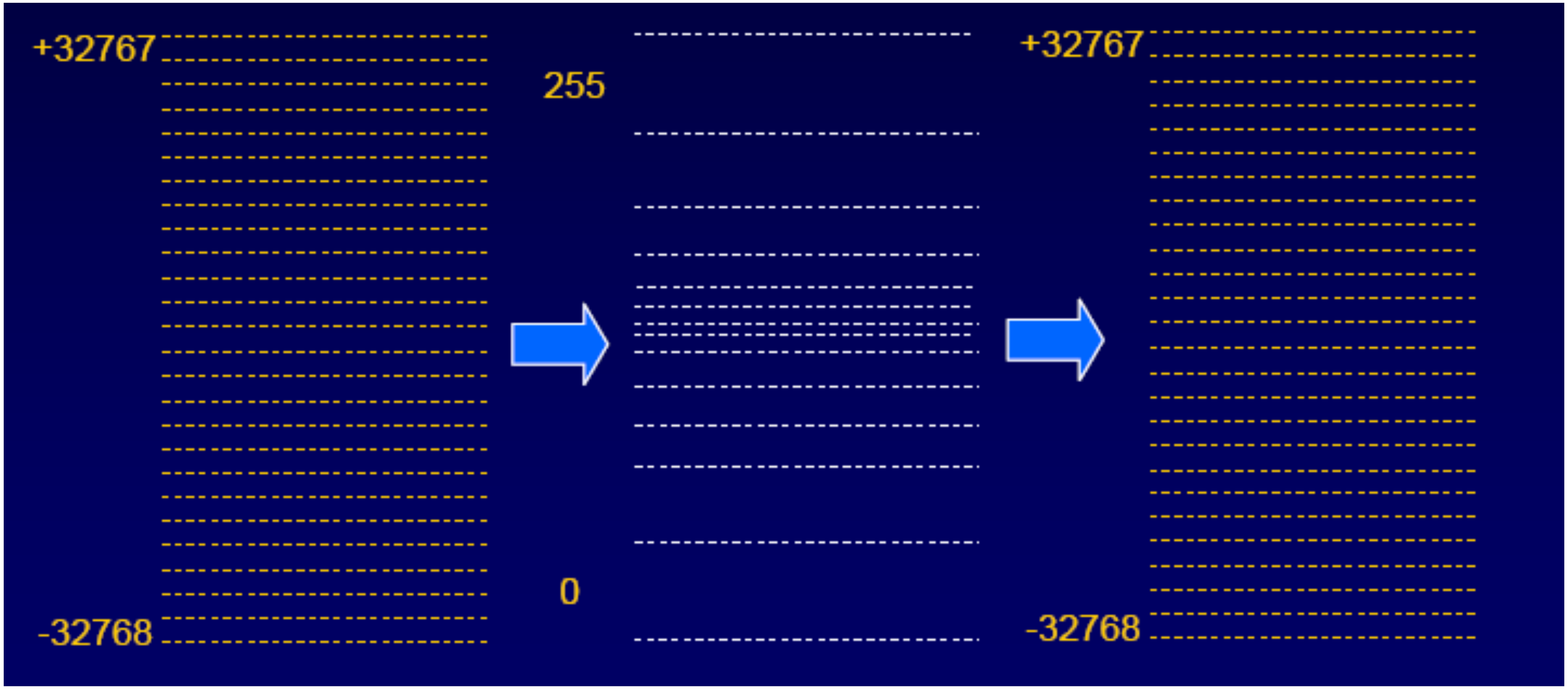
- A-law
  - Europe
  - International traffic on ISDN

- Both uses a 8 bits for quantization

- It is a lossy compression



μ-law and A-law compressions



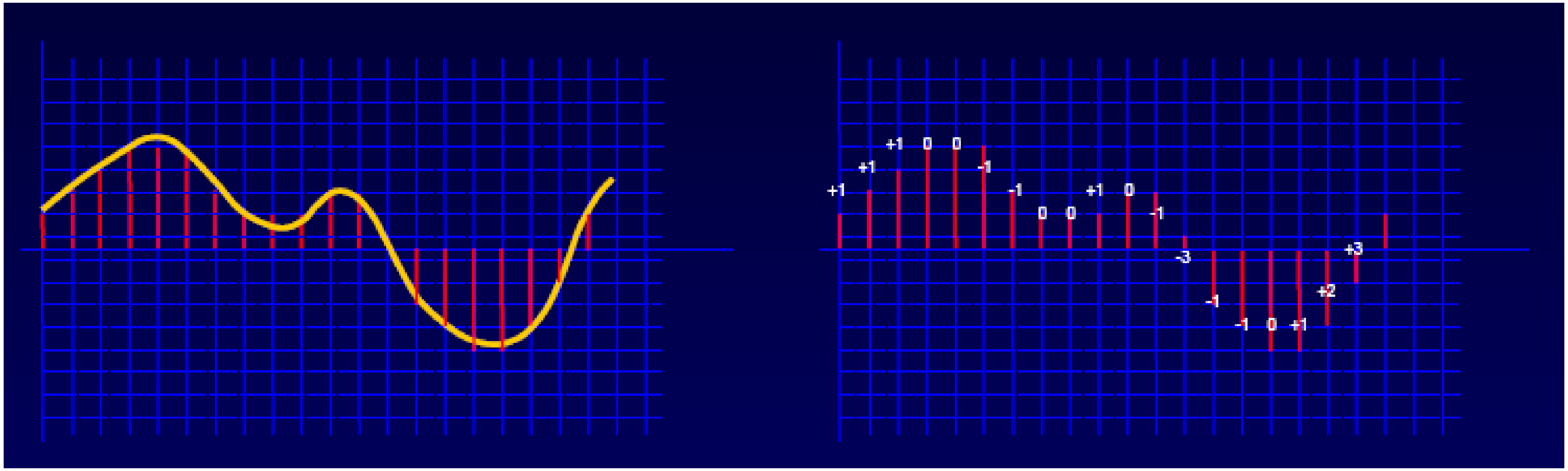
16 bits 8 bits 16 bits

ln exp

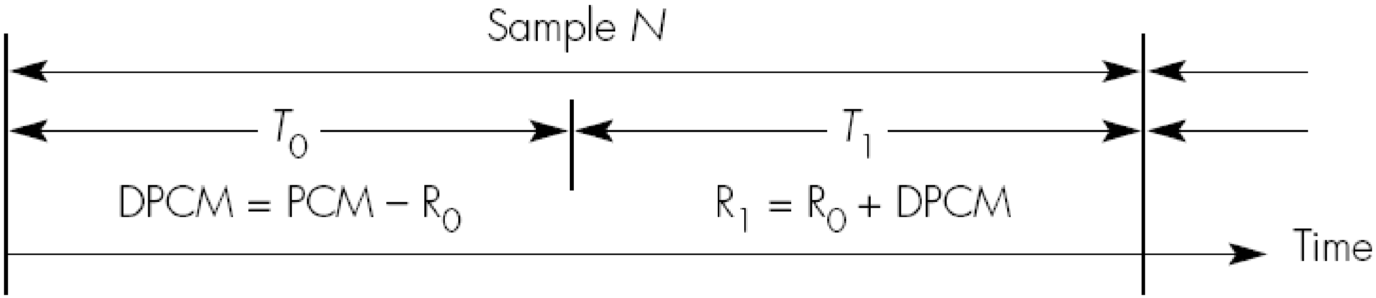
Aim – to use 8 bits instead of 8 bits

Differential Pulse Code Modulation

- Differential Pulse Code Modulation (DPCM)
  - derives from PCM
  - difference between two consectuive samples



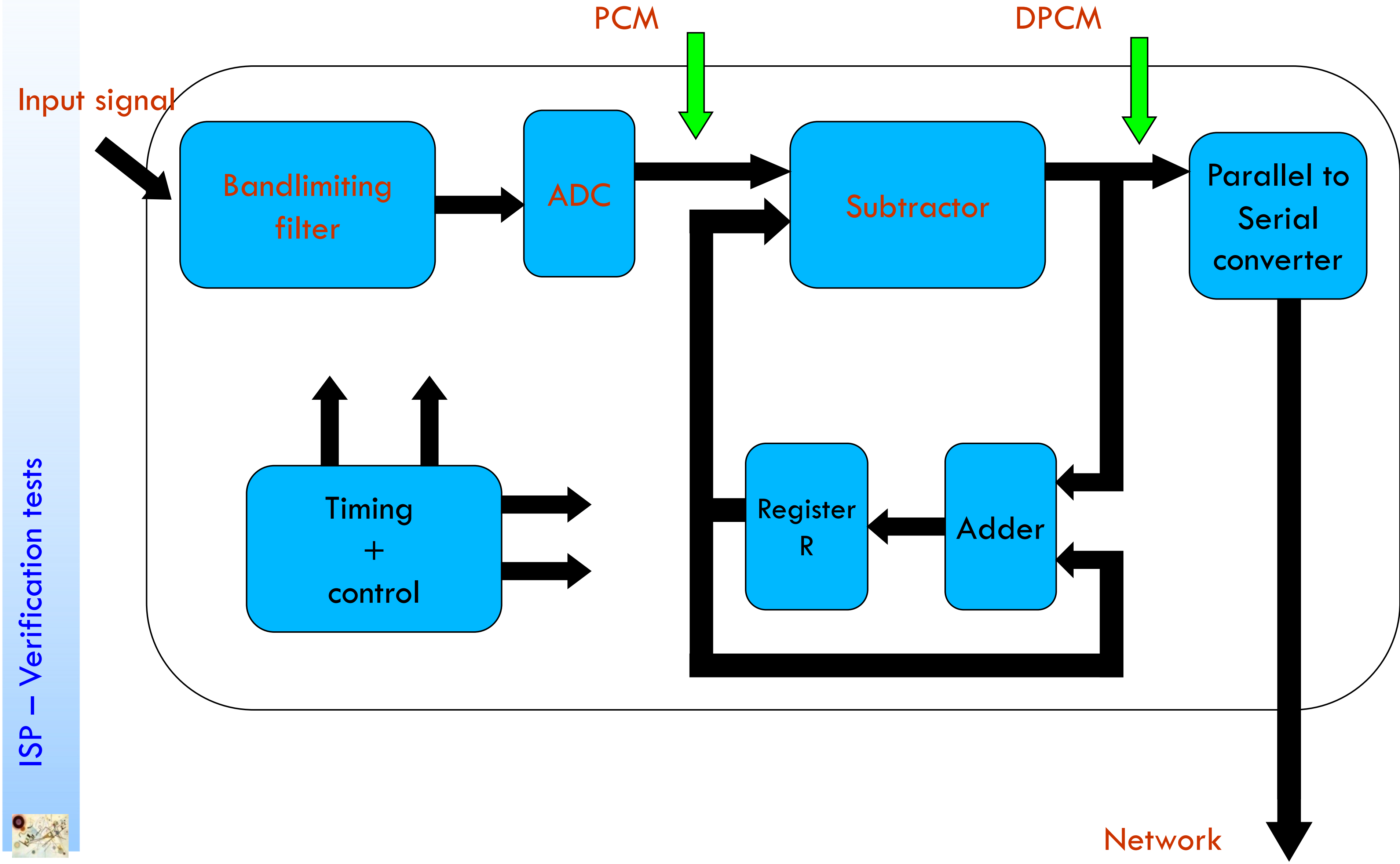
DPCM



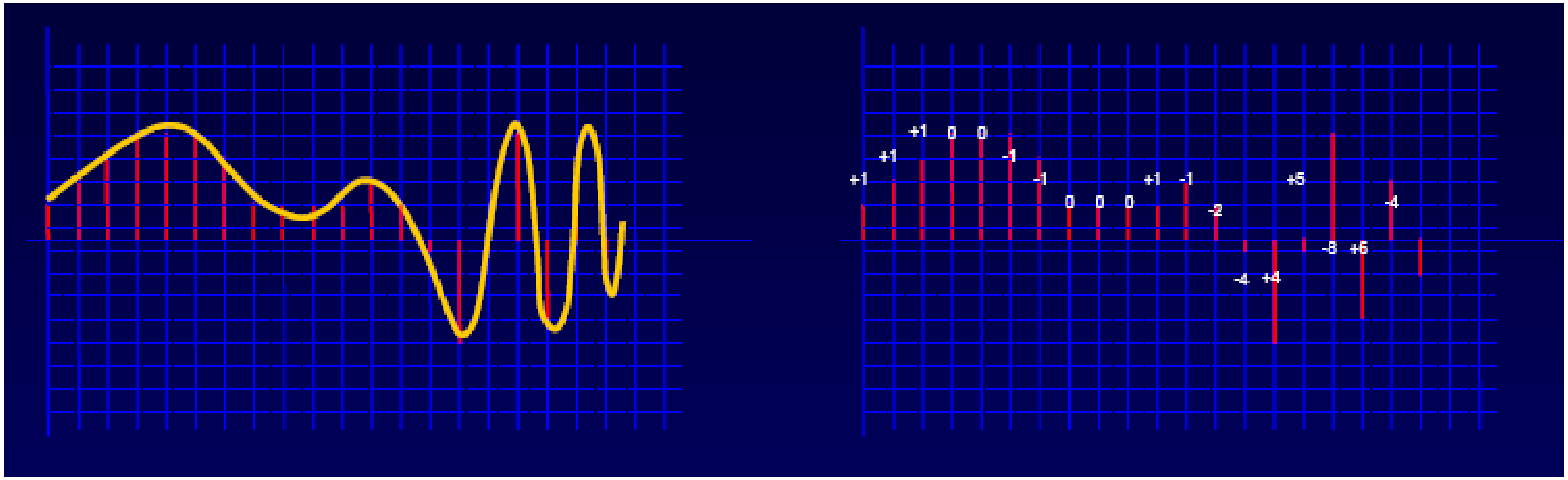
R_0 = current contents of register R and R_1 = new/updated contents

Timing phase. Two registers are used.

DPCM encoder

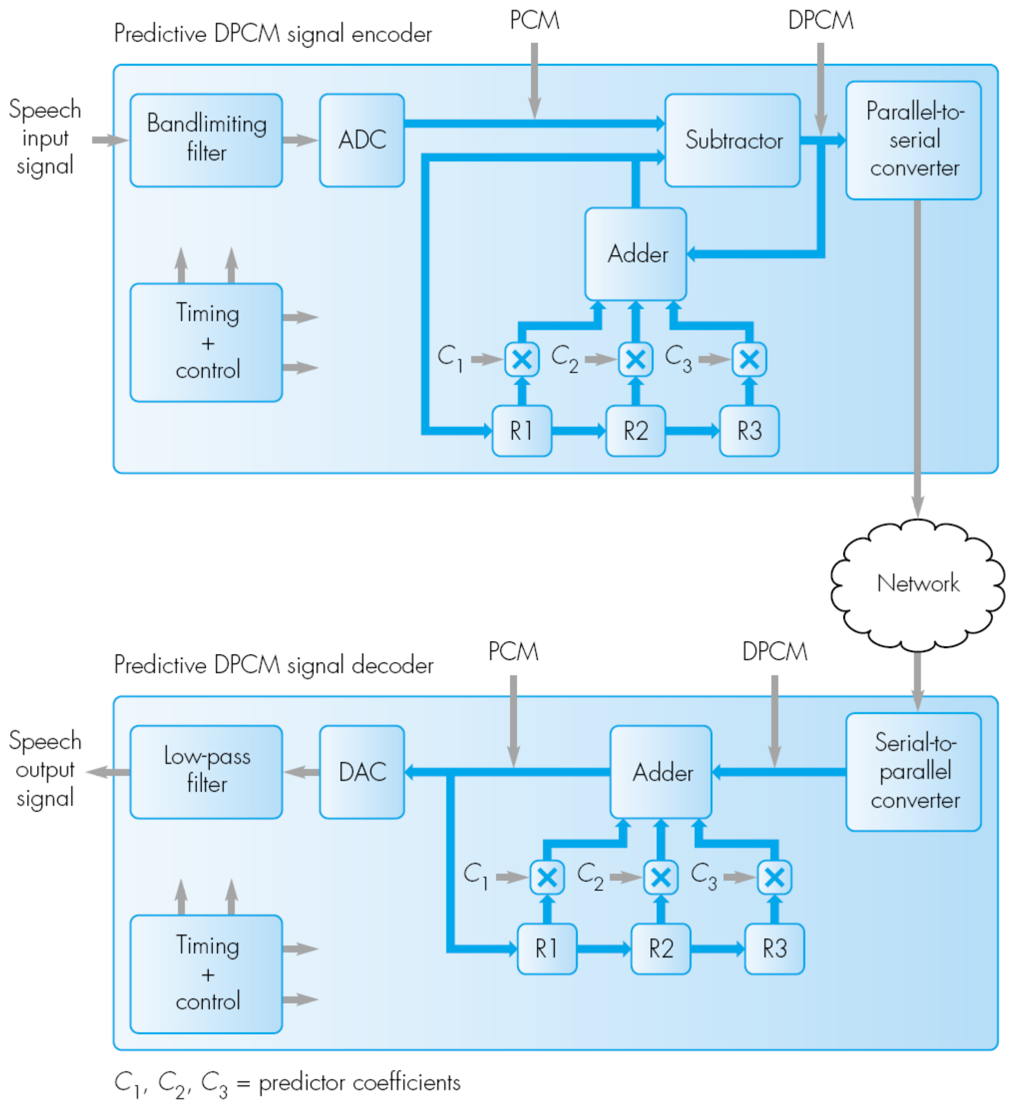


DPCM slope overload



High frequencies differences needs a higher number of bits

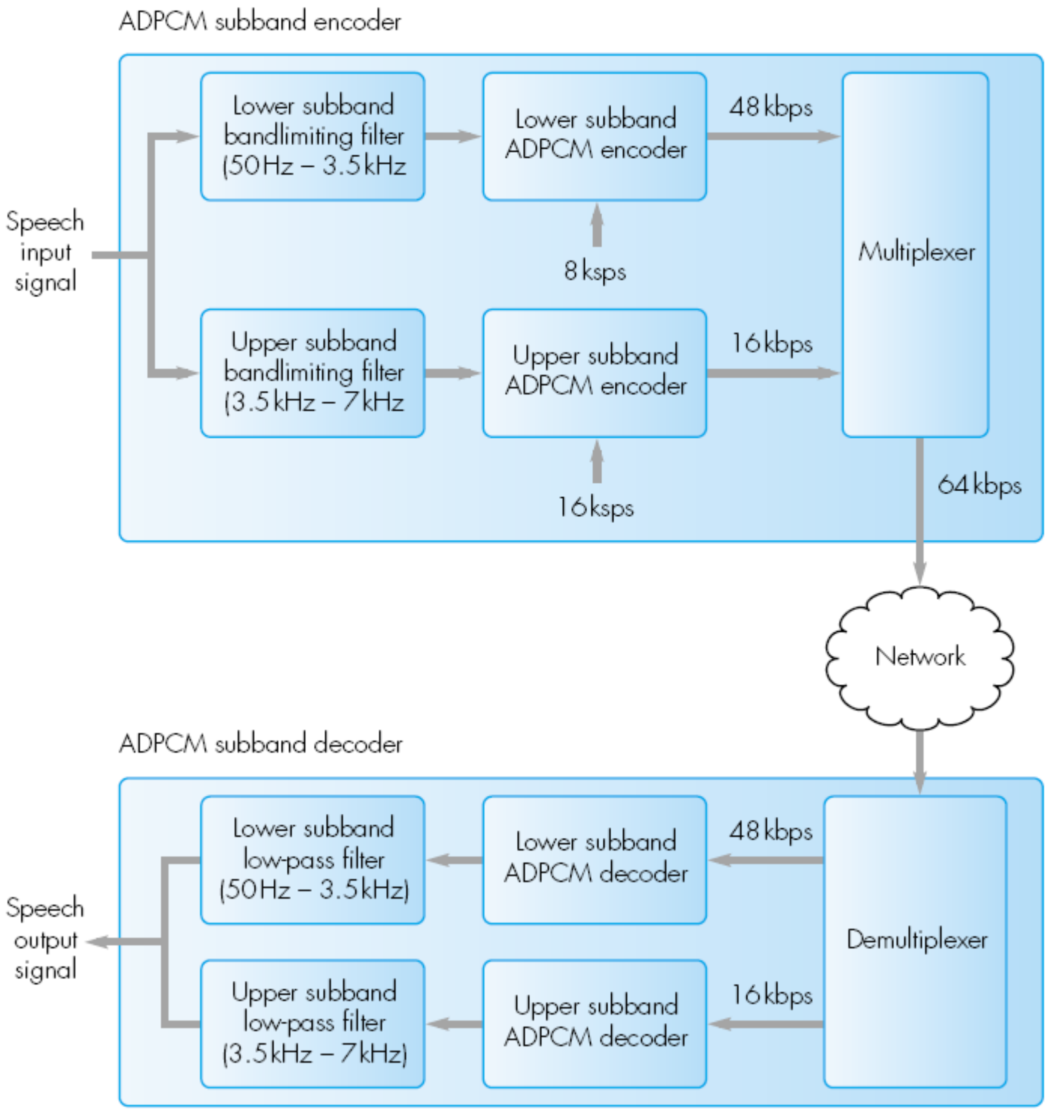
Predictive DPCM



C_1, C_2, C_3 = predictor coefficients

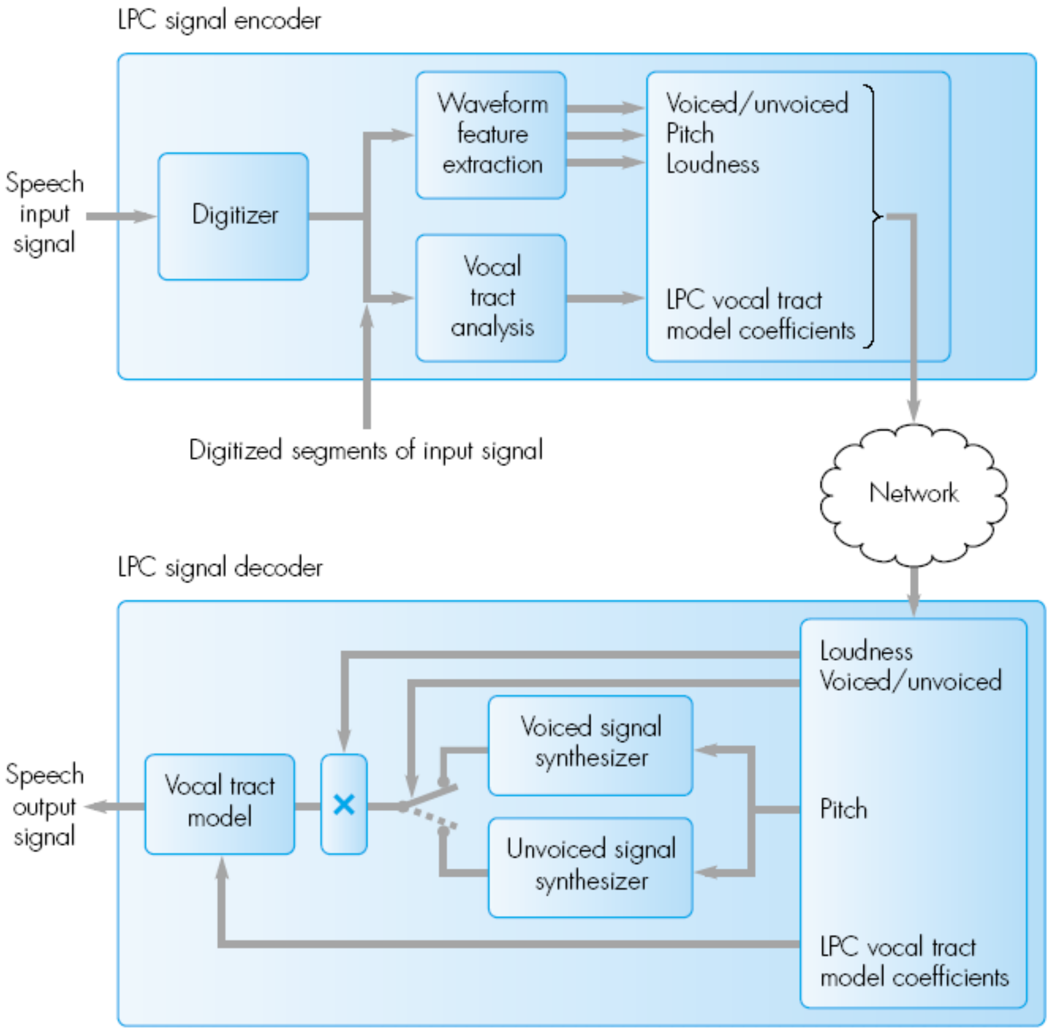
Prediction by using 3 registers and 3 coefficients

Adaptive DPCM



A filtering scheme is used

Linear Preditive Coding



LPC scheme

Code-excited LPC

- CELP
  - Set of segments (templates)
    - named codebook
  - transmitted codeword
    - template with best matching with an input segment

References

- Material
  - Slides
  - Video Lessons

- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Audio digitale, A. Uncini, McGraw-Hill Education, 2006
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 17

- To be transmitted over the network multimedia content must be digitized and compressed

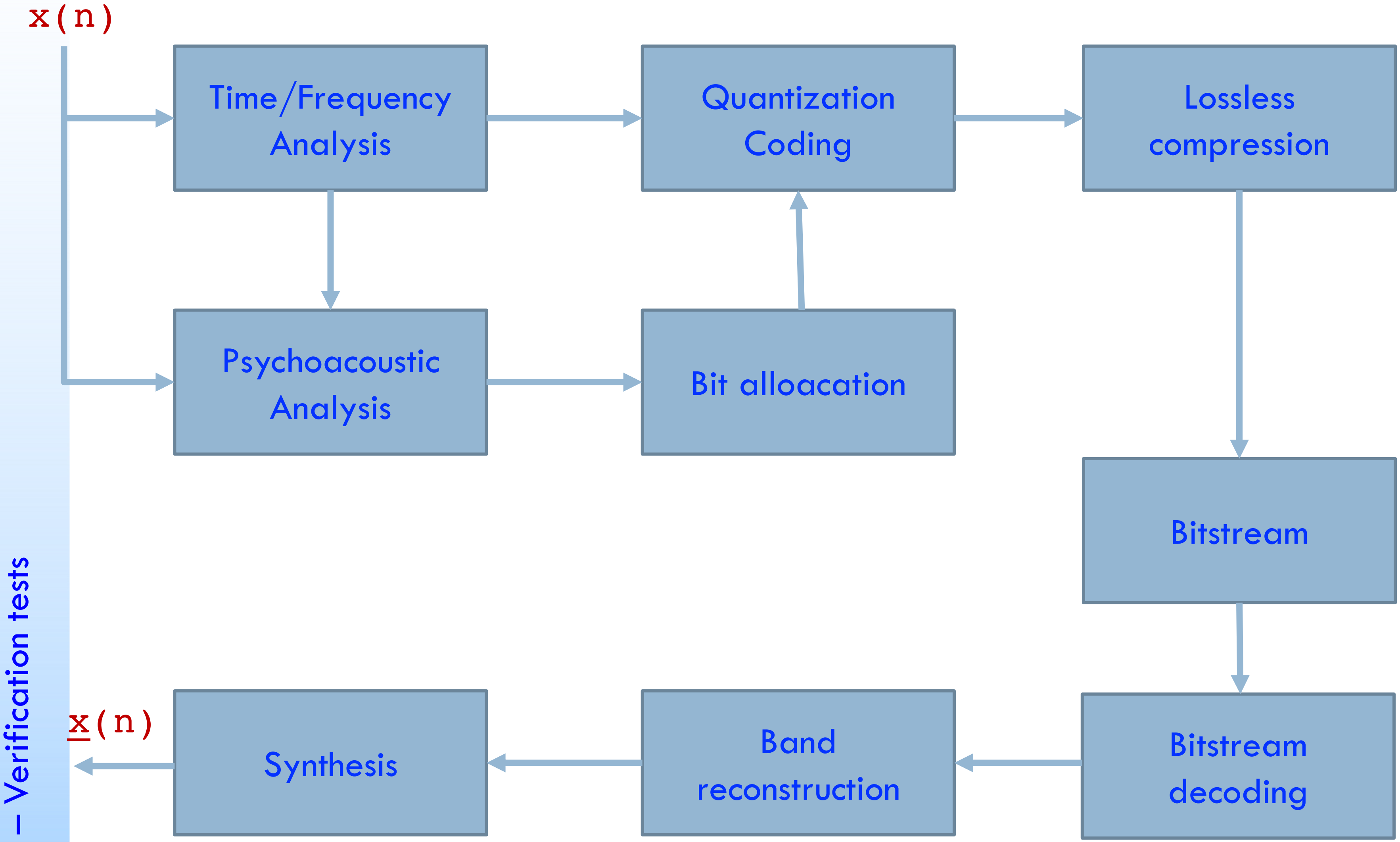
- Question
  - Describe the perceptual compression

Perceptual coding

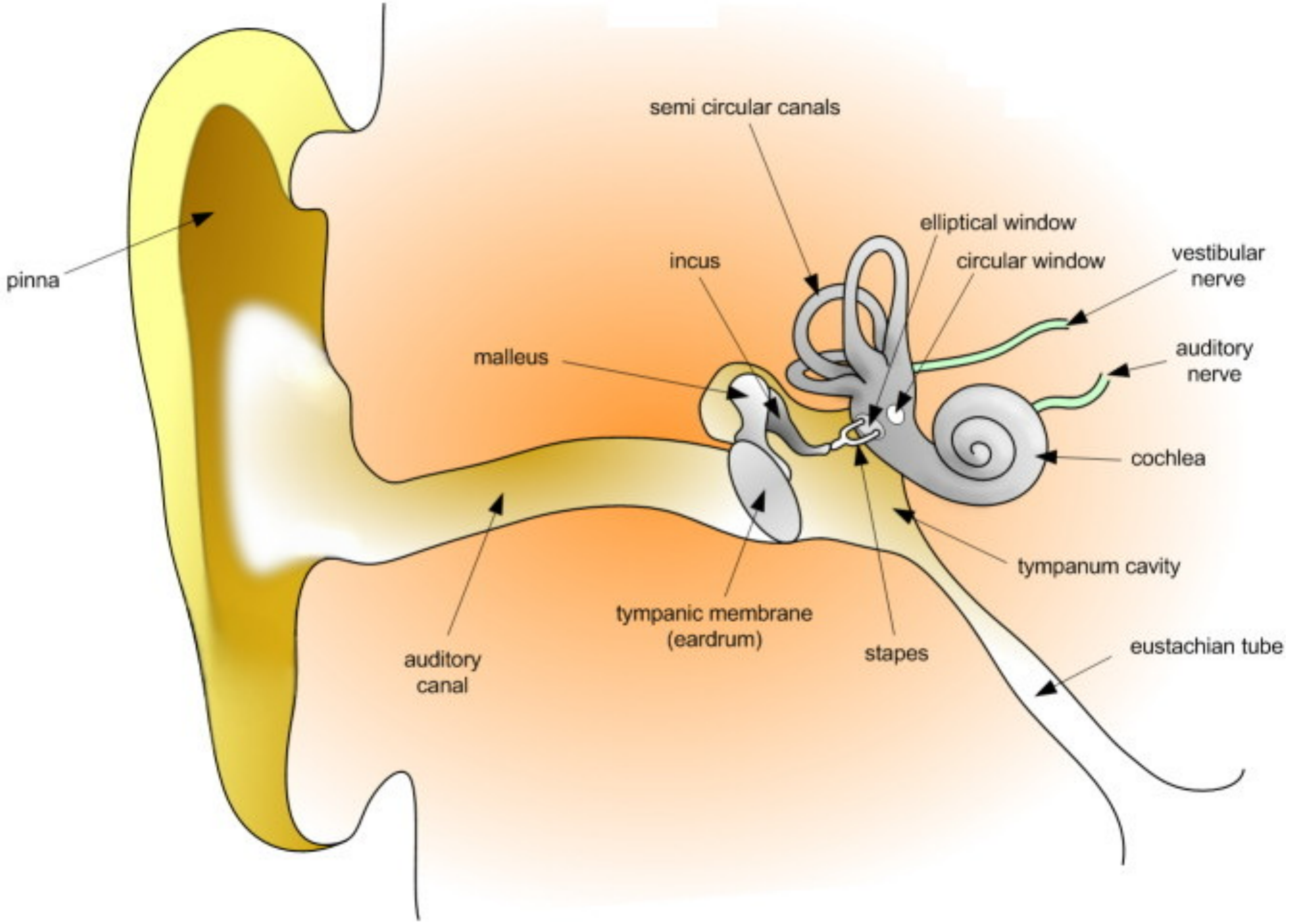
- psychoacoustic models
  - exploit the characteristics of the human ear
  - only the perceptual characteristics are transmitted

- Main aspects
  - frequency masking
  - temporal masking

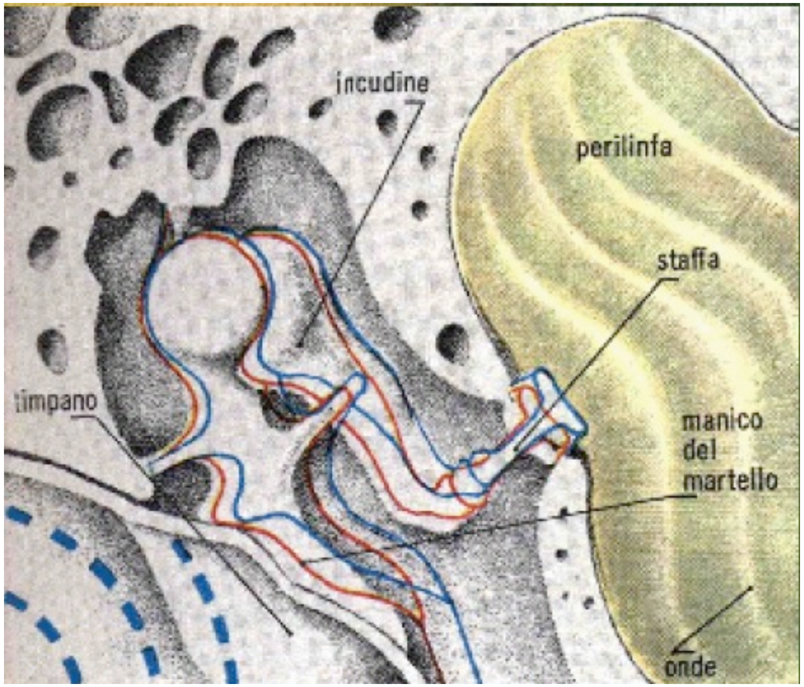
General scheme



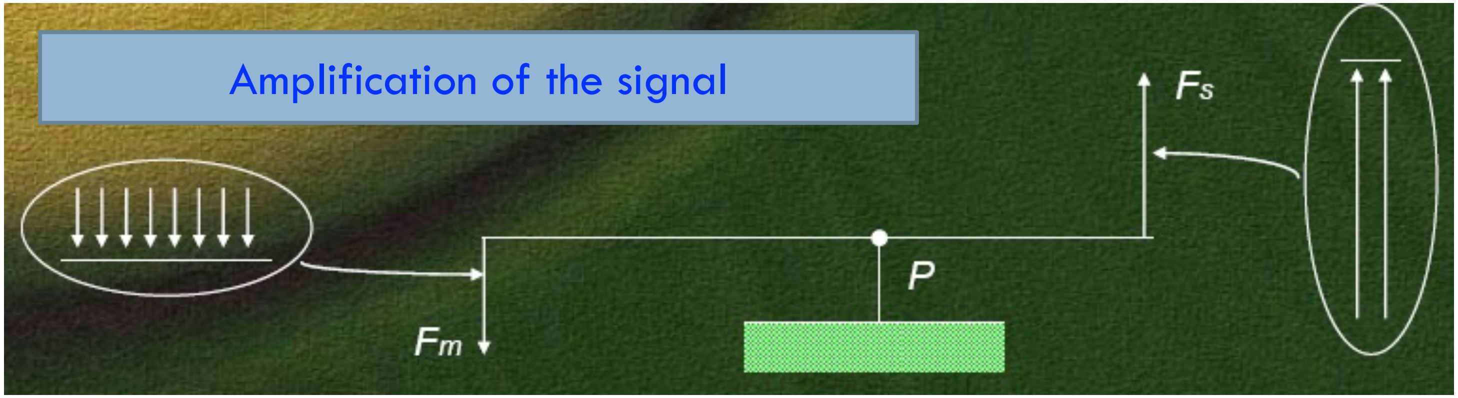
Auditory perception



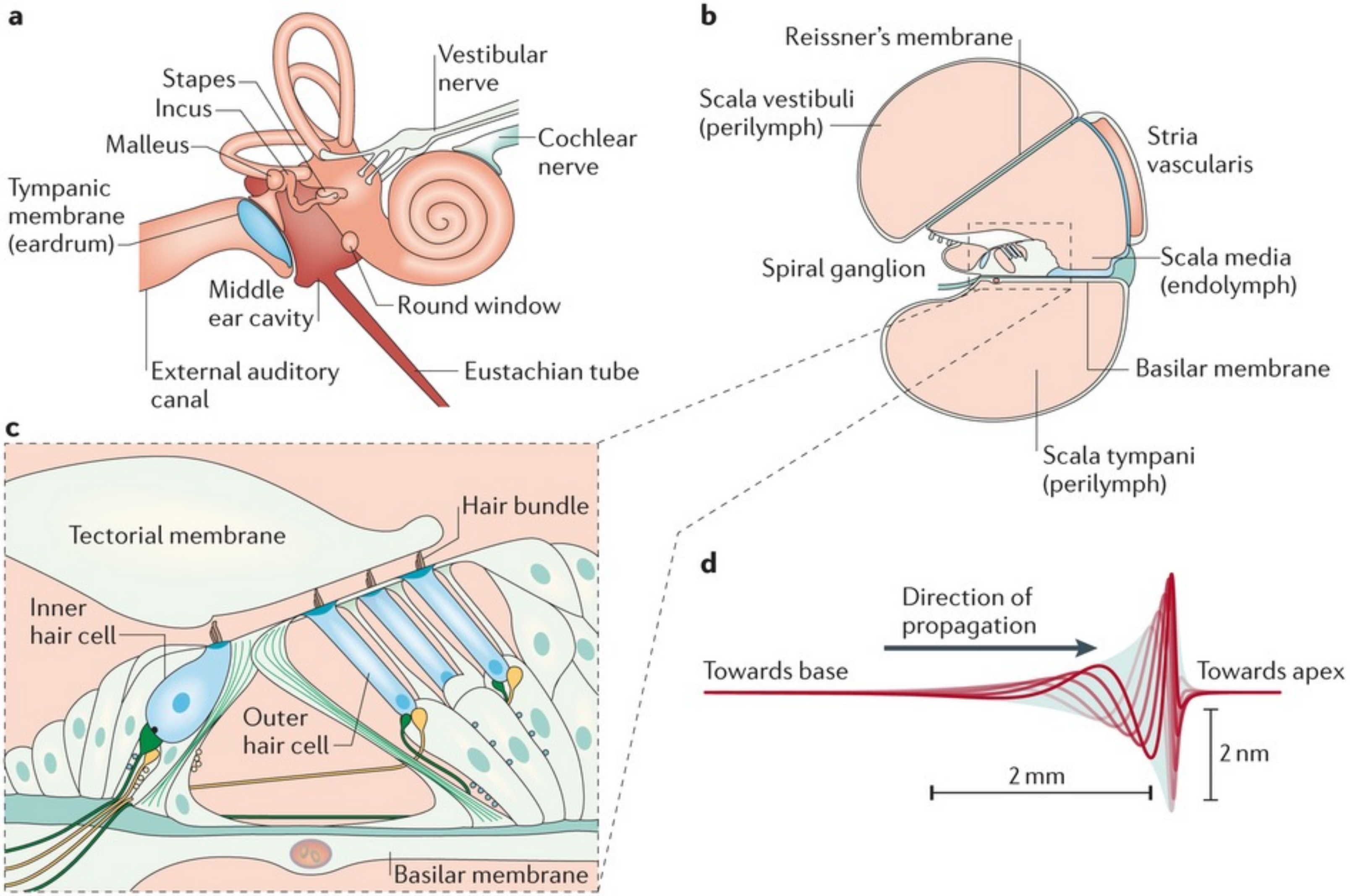
Middle ear



Vibration propagation

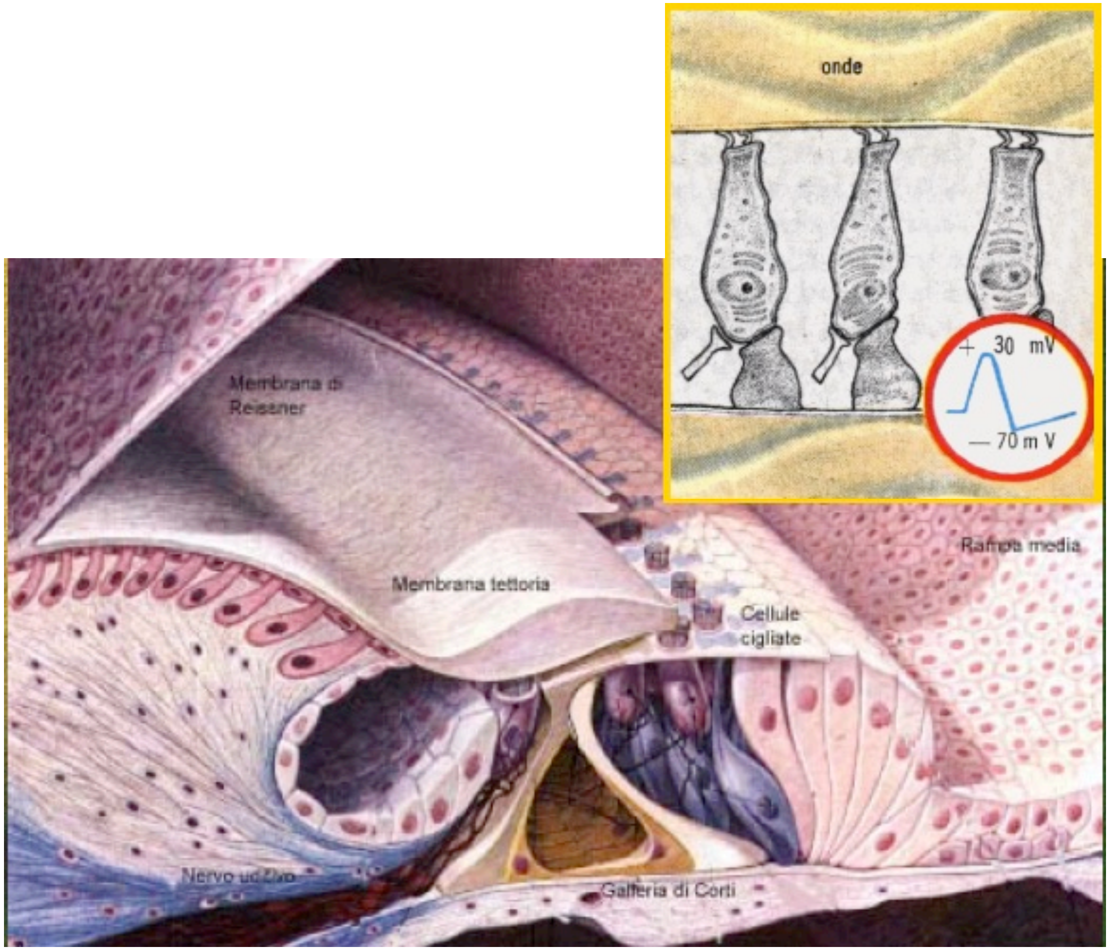


Cochlea

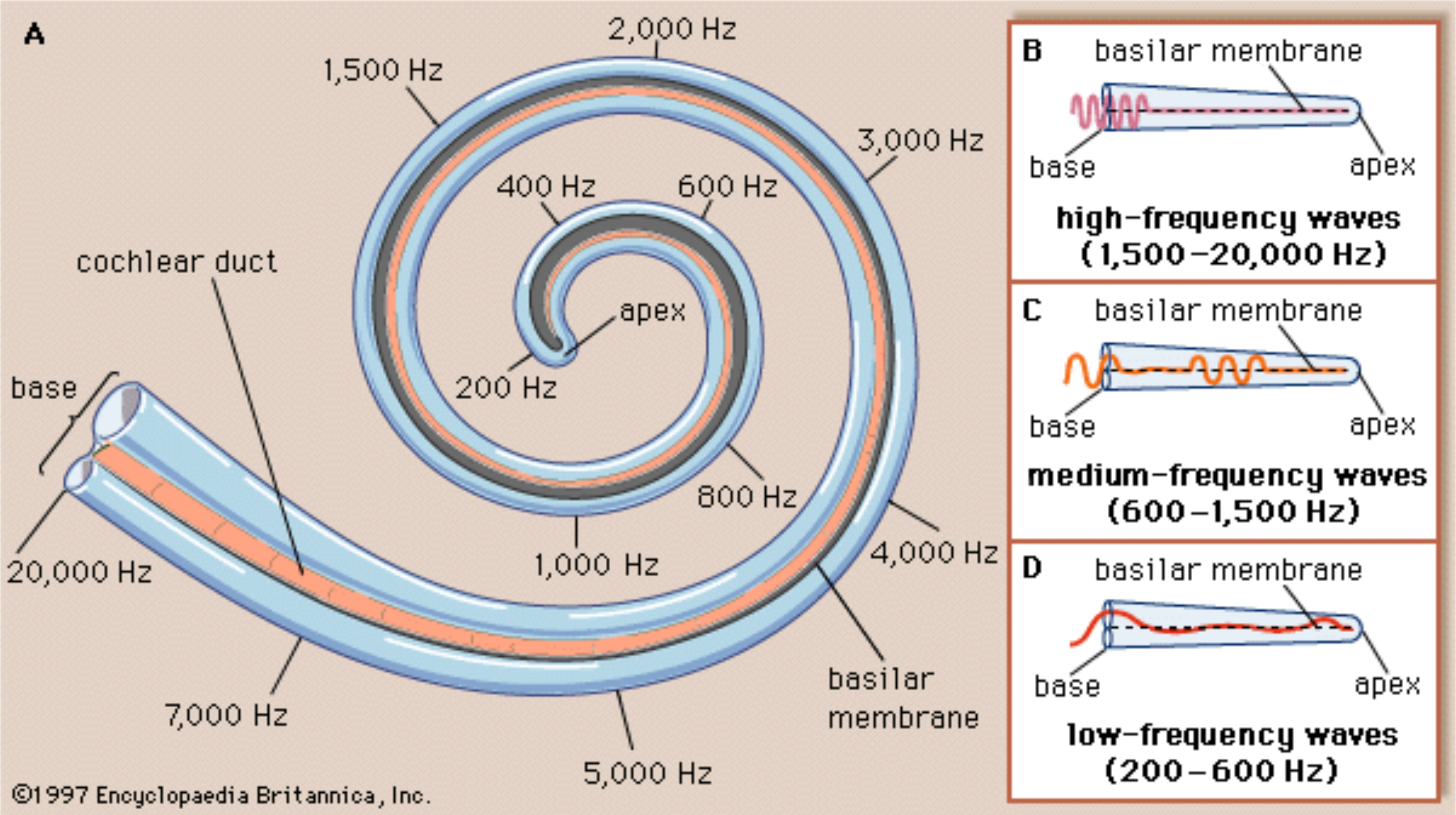


Nature Reviews | **Neuroscience**

Corti organ

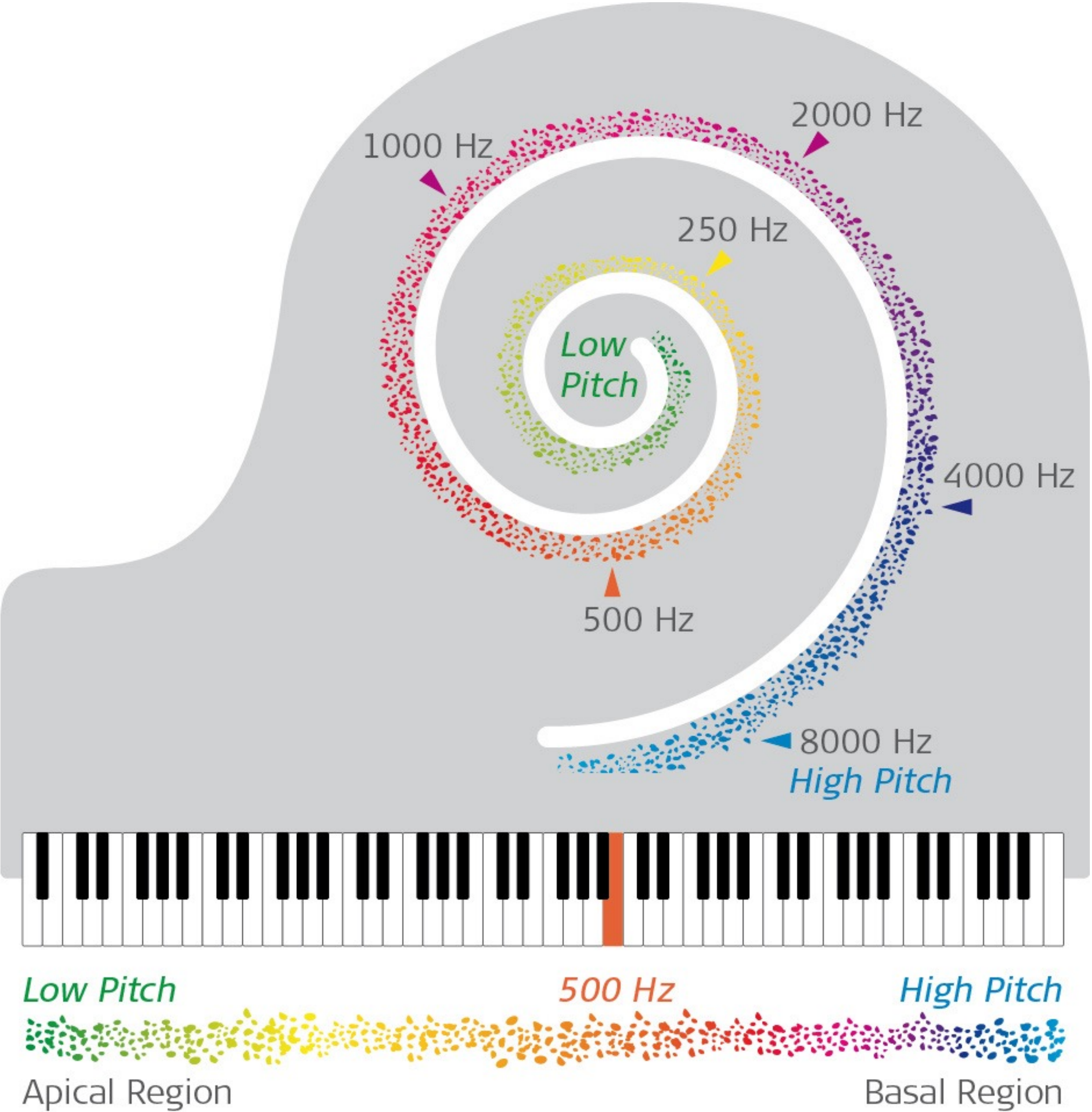

onde
+ 30 mV
— 70 m V
Membrana di Reissner
Membrana tettoria
Cellule cigliate
Rampa media
Nervo uditivo
Galleria di Corti

Basilar membrane



Bank of filters

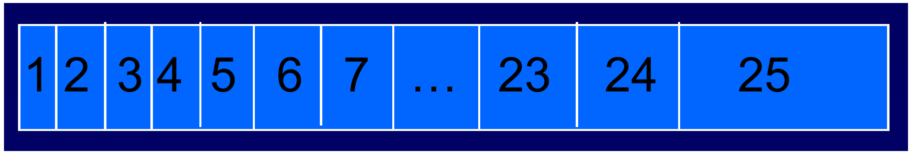
Basilar membrane



Herman
von Helmholtz

Basilar membrane

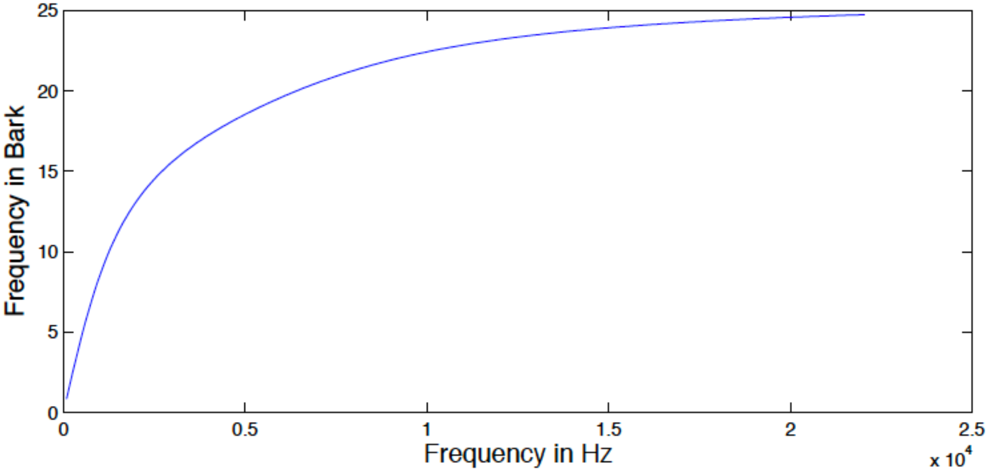
- Basilar membrane
  - 25 critical bands
  - non-linear behaviour (logartitmic)
  - pass-band bank filters



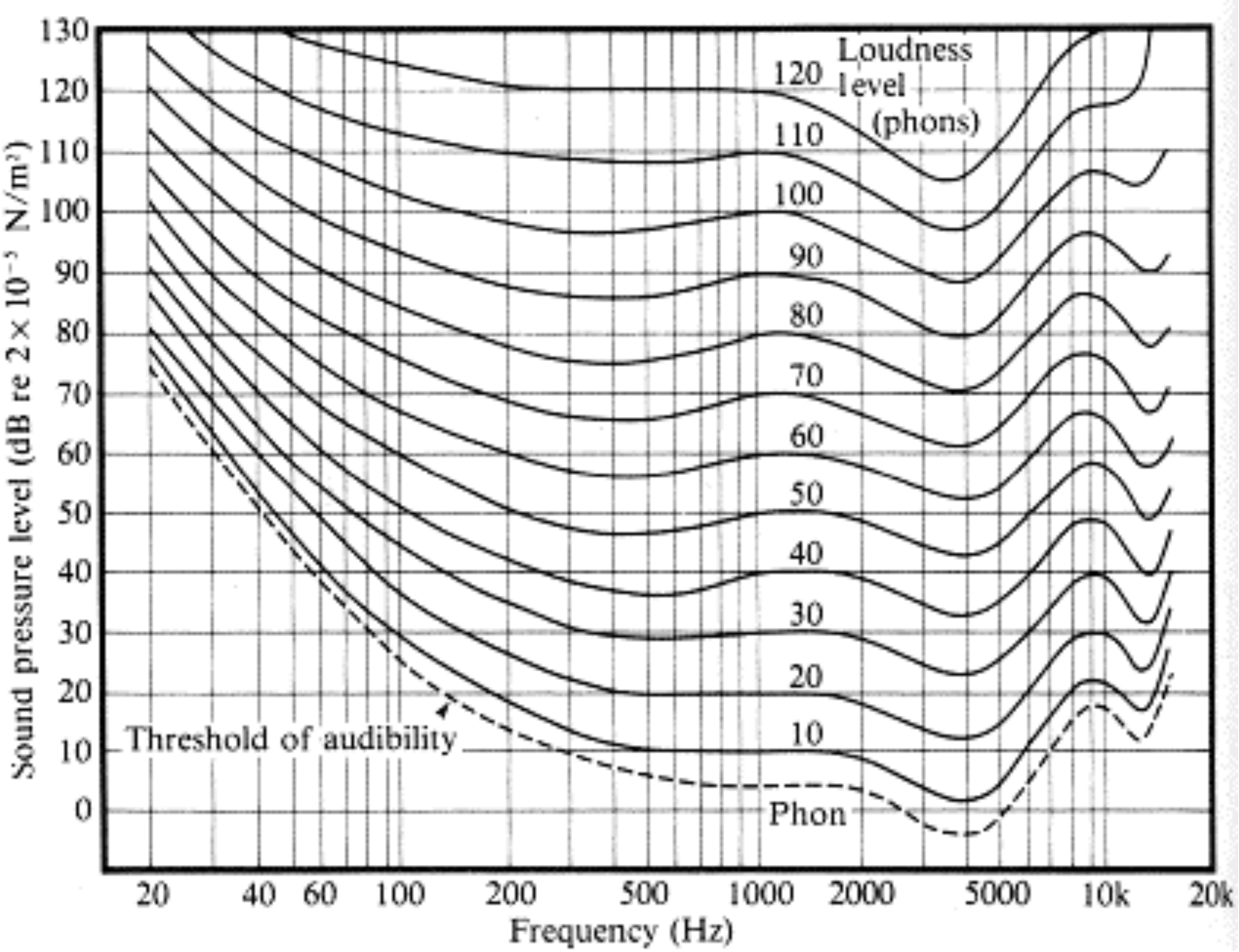
Non-linear basilar membrane critical bands

Bark scale

- Bark scale
  - $f < 500\ Hz$ $$f_{bark} = \frac{f}{100}$$
  - $f \geq 500\ Hz$ $$f_{bark} = 9 + 4 \log \frac{f}{100}$$

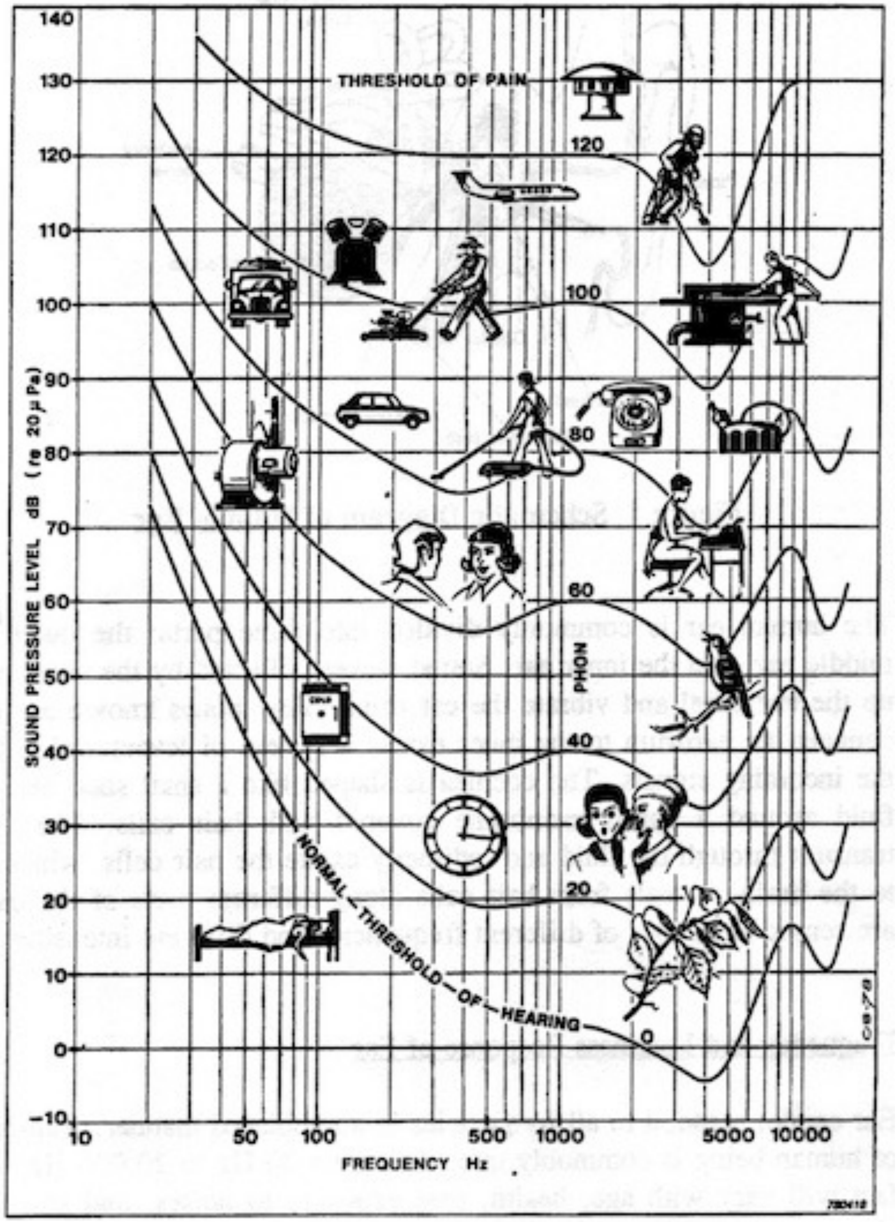


Fletcher – Munson Diagram



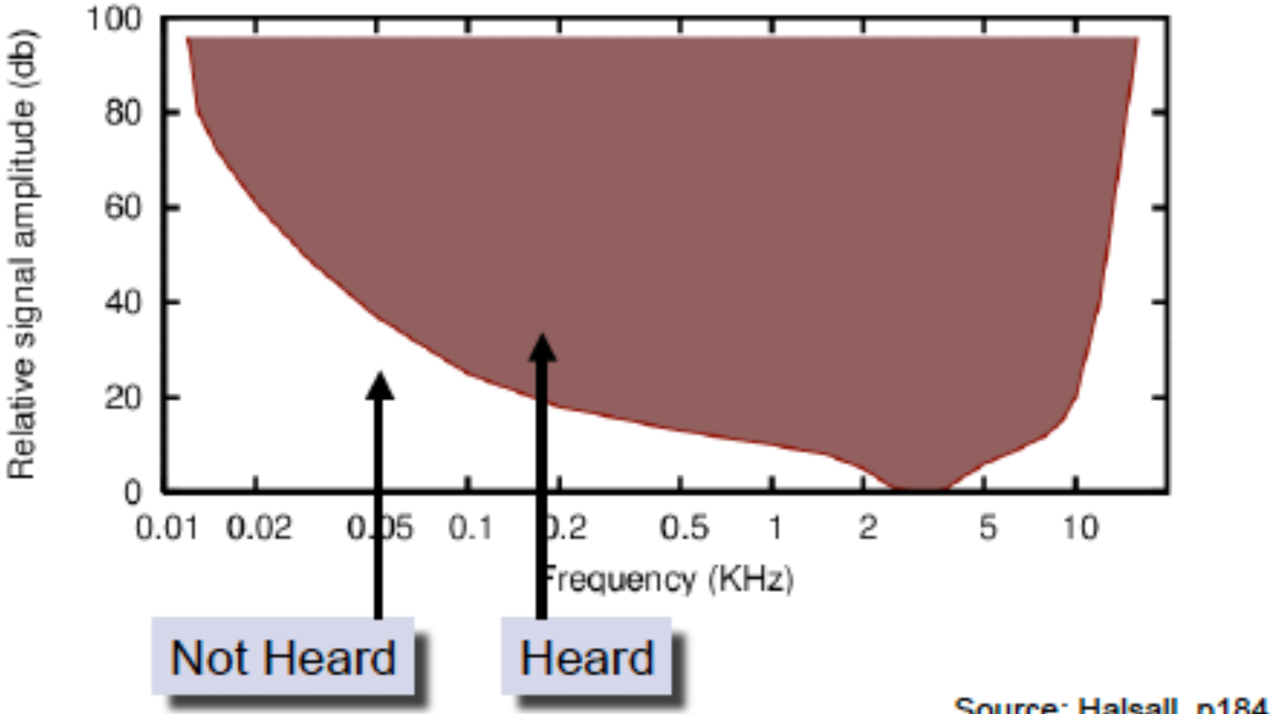
Equal-loudness contours for the human ear determined experimentally

Fletcher – Munson Diagram



Example of phones

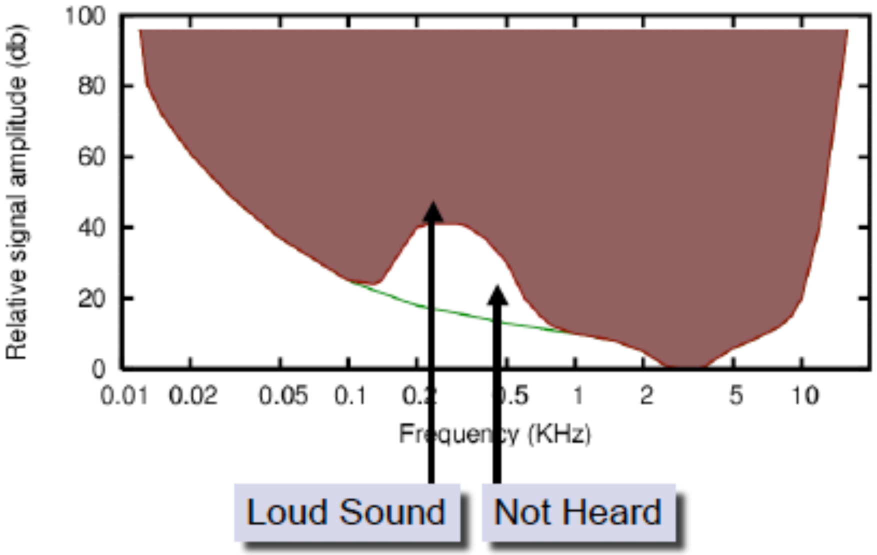
Hearing threshold



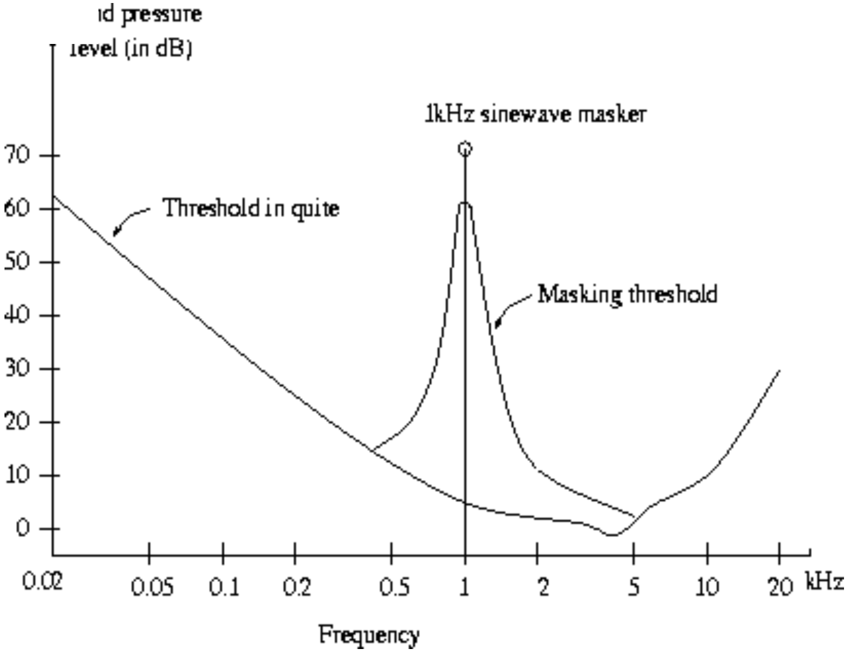
Source: Halsall, p184

Normal Treshold of Hearing

Simultaneous masking

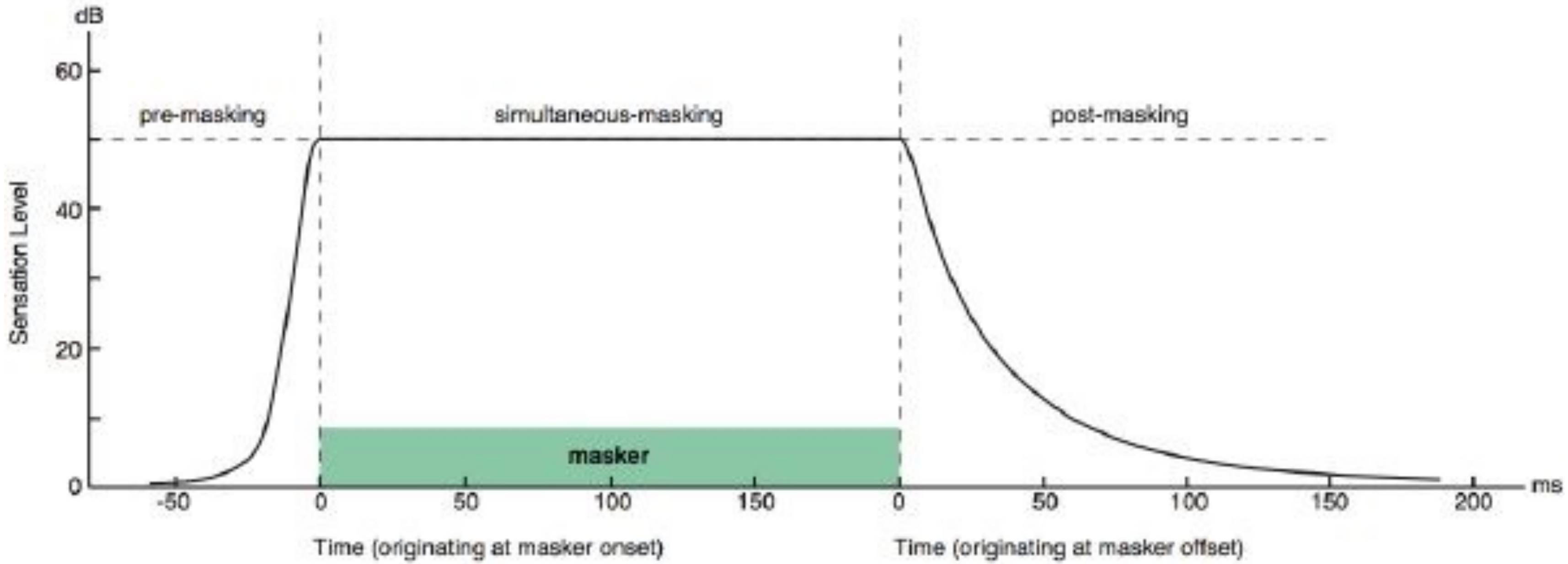


Masking threshold



Example of simultaneous masking

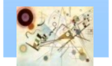
Temporal masking



Example of temporal masking

MPEG Audio Encoders

- Perceptual coding
  - unnecessary information eliminated
  - psychoacoustic model
  - masking mechanism
    - number of quantization bits



MPEG

- Moving Picture Experts Group (MPEG)
  - working group of authorities that was formed by ISO and IEC
  - standards for audio and video compression and transmission
  - established in 1988 by the initiative of
    - Hiroshi Yasuda (Nippon Telegraph and Telephone)
    - Leonardo Chiariglione
  - The first meeting was in May 1988 in Ottawa, Canada

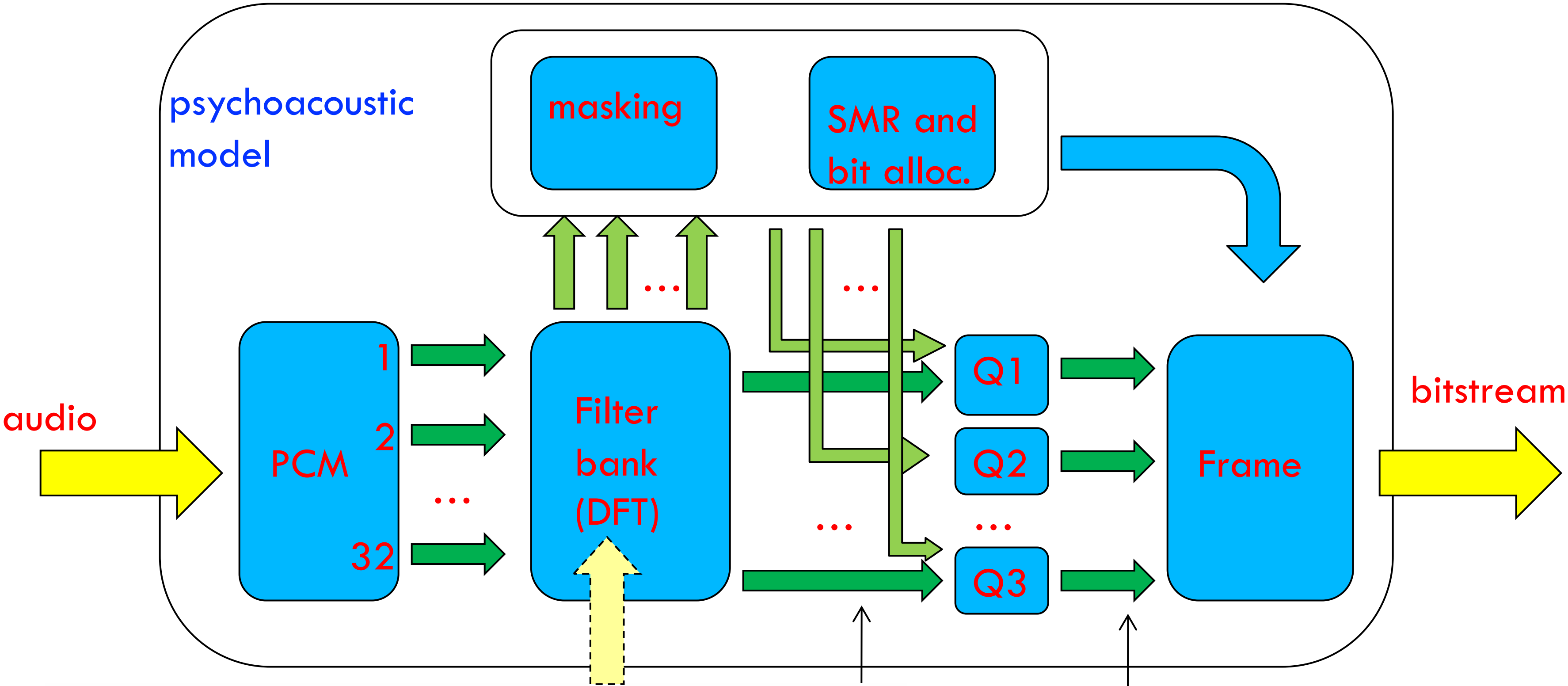
MPEG-1

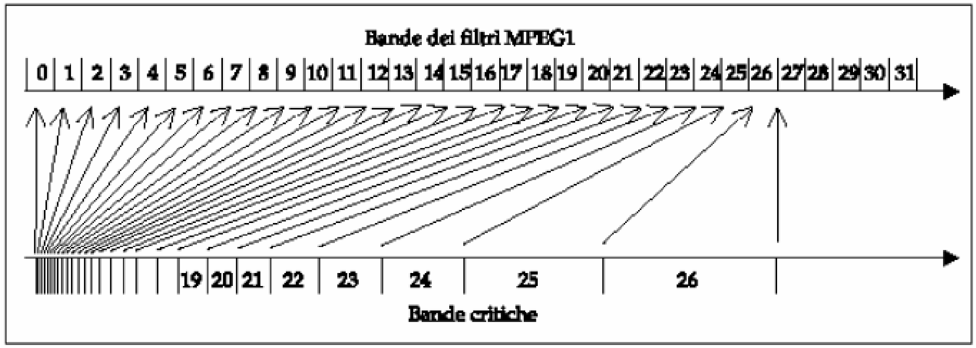
- MPEG-1
  - standard for lossy compression of video and audio
  - designed to compress VHS-quality raw digital video and CD audio down to 1.5 Mbit/s (26:1 and 6:1 compression ratios respectively)
  - without excessive quality loss
    - video CDs
    - digital cable/satellite TV
    - digital audio broadcasting (DAB)

MPEG 1

- MPEG 1
  - Layer 1
    - Compressed bit rate: 32-448 Kbps
  - Layer II
    - Compressed bit rate: 32-192 Kbps
  - Layer III
    - Compressed bit rate: 64 Kbps
    - mp3

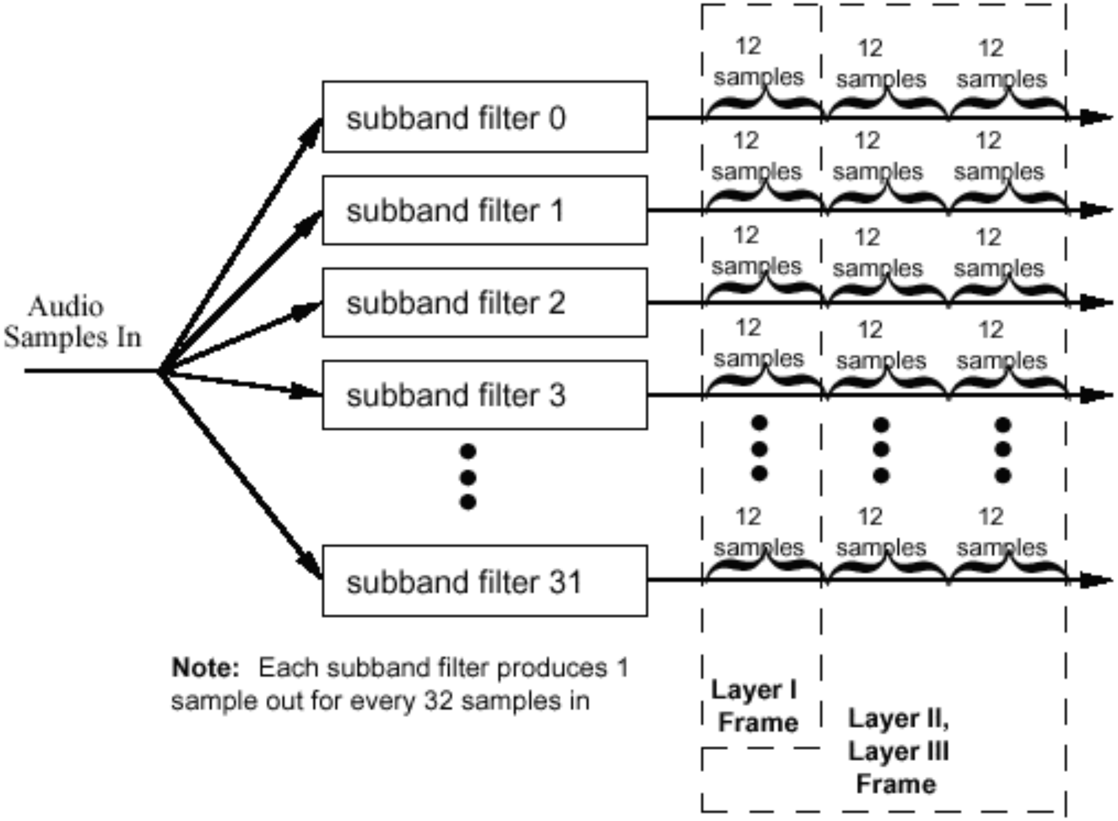
MPEG 1 – Layer 1





Quantized subbands samples

MPEG 1 – Layer 1



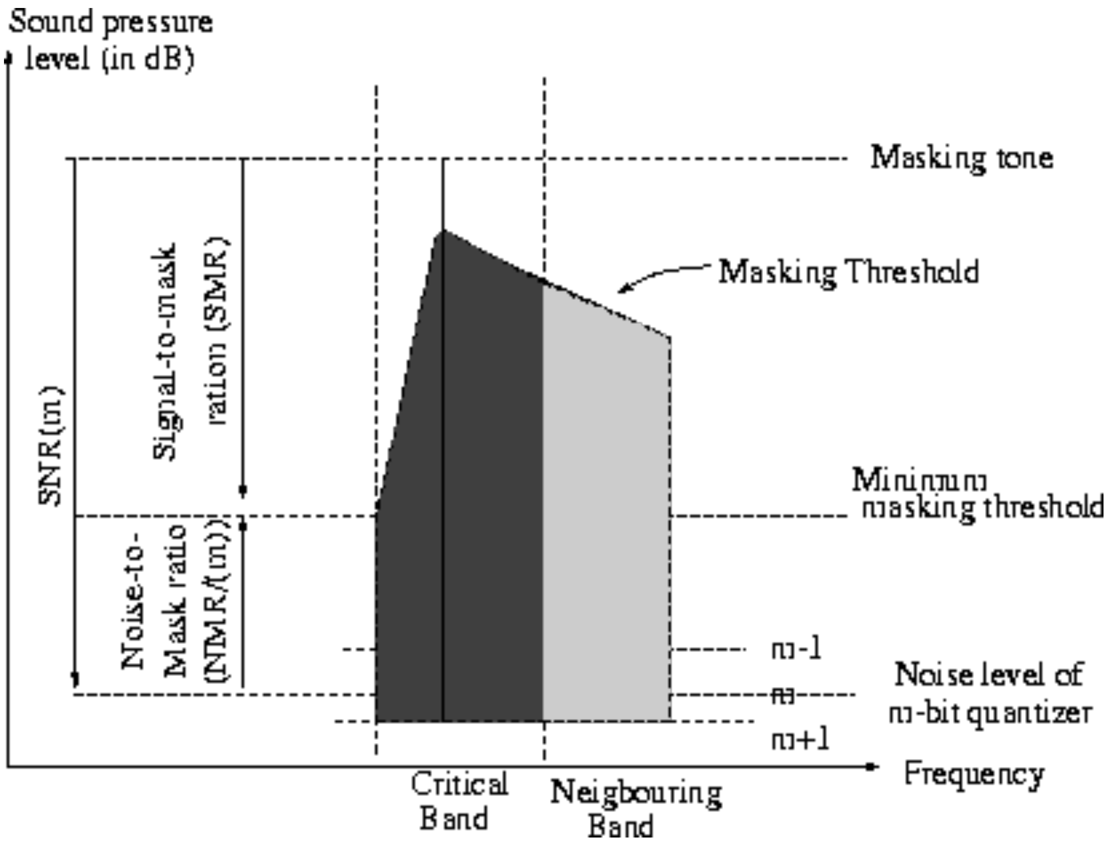
Note: Each subband filter produces 1 sample out for every 32 samples in

Definition of segments

Psychoacoustic model

- A 1024 points FFT is used
- Global masking thresholds

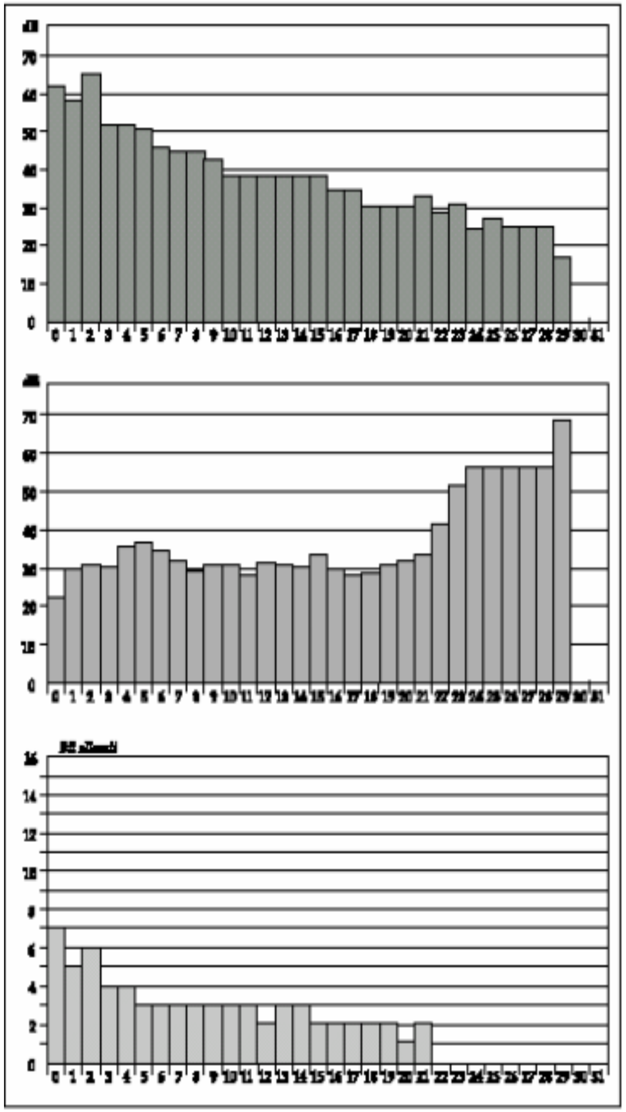
Global bit allocation



Bit allocation for each subband (NMR = SNR - SMR)

Then the subbands are placed in order of lowest to highest mask-to-noise ratio, and the lowest subband is allocated the smallest number of code bits and this process continues until no more code bits can be allocated

Global bit allocation



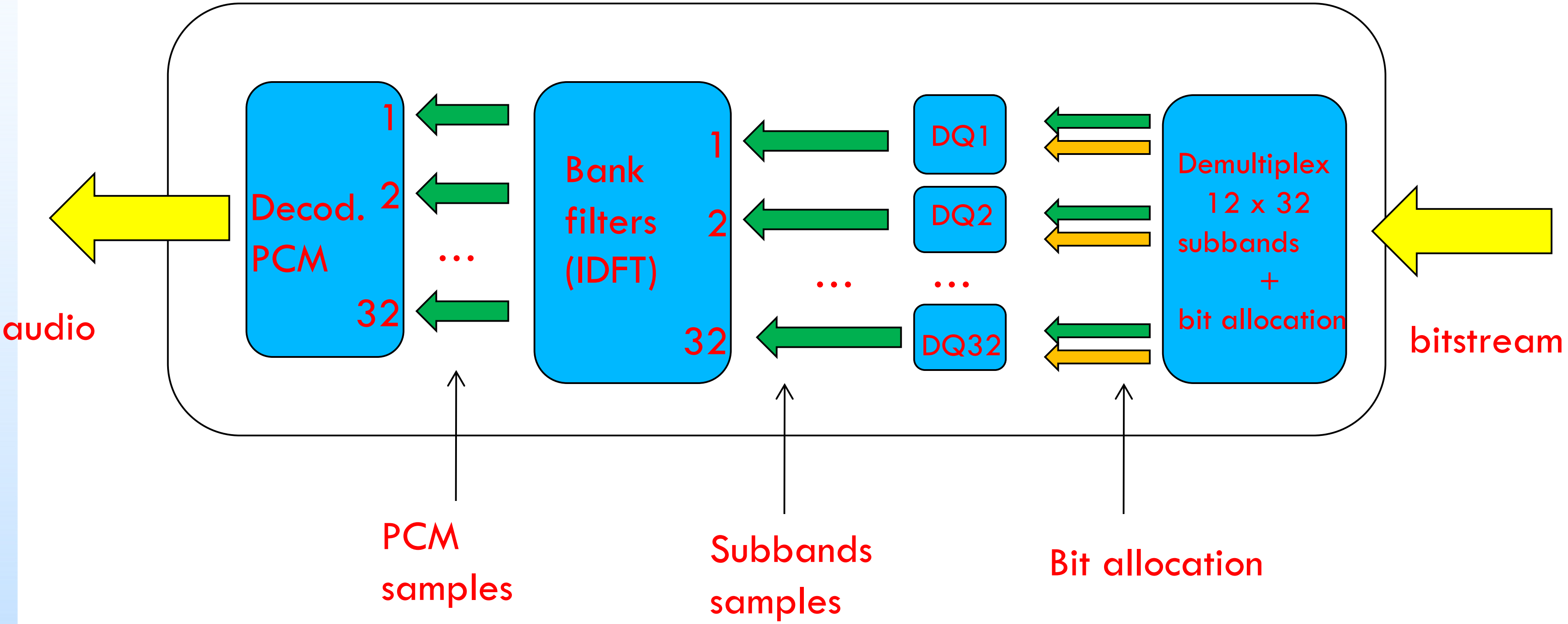
Amplitude spectrum

Global masking threshold

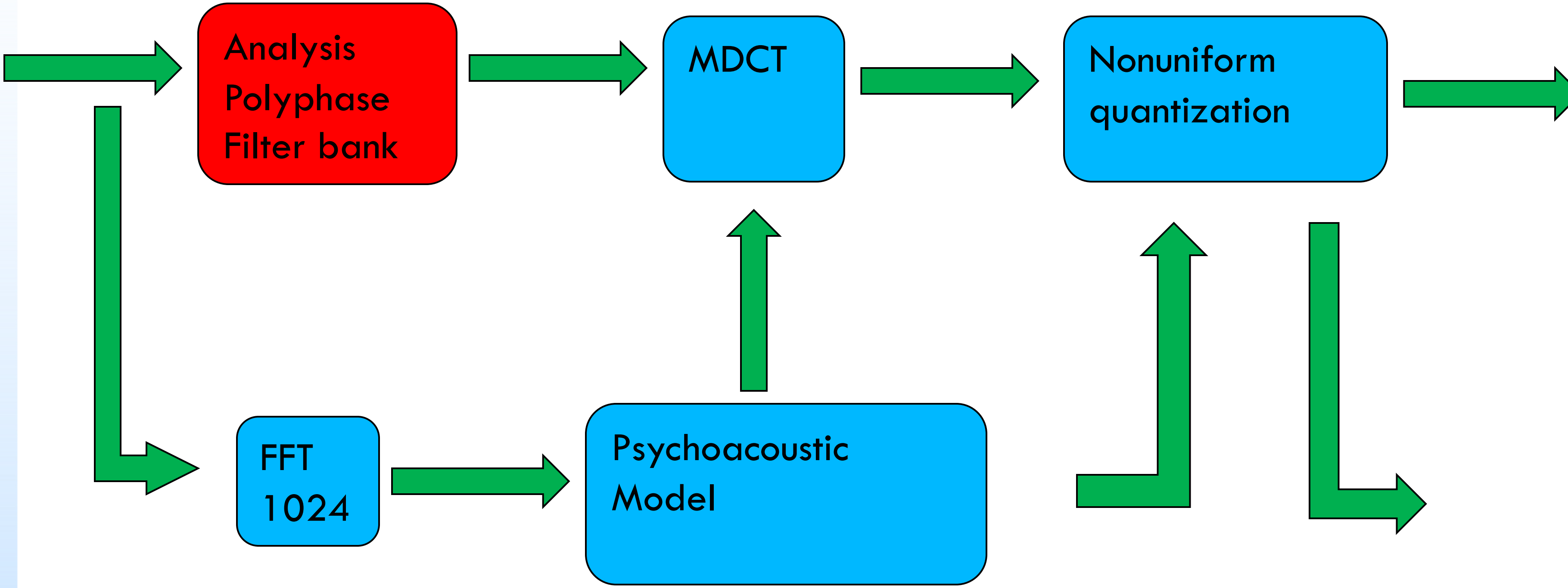
Subbands bit allocation

ISP – Verification tests

Decoder


Decod.
PCM
1
2
...
32
Bank
filters
(IDFT)
1
2
...
32
DQ1
DQ2
...
DQ32
Demultiplex
12 x 32
subbands
+
bit allocation
audio
bitstream
PCM
samples
Subbands
samples
Bit allocation

MPEG – Layer 3



Analysis Polyphase Filter bank

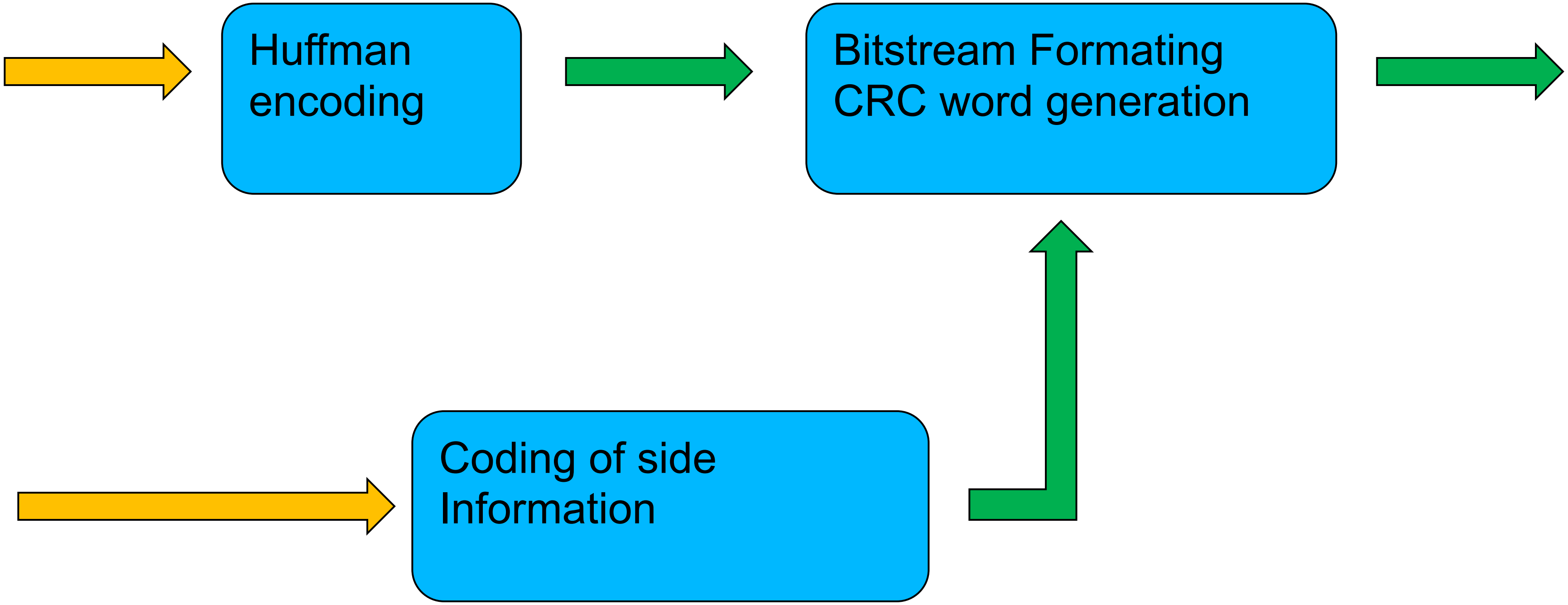
MDCT

Nonuniform quantization

FFT 1024

Psychoacoustic Model

MPEG – Layer 3



Huffman encoding

Bitstream Formating CRC word generation

Coding of side Information

References

- Material
  - Slides
  - Video Lessons

- Books
  - Signal Processing Book (Ciaramella)
    - free download on the e-learning platform
  - Fundamentals of Multimedia, Z.-N. Li, M. S. Drew, J. Liu, Springer, 2021
  - Digital Signal Processing, J. Proakis, D. Manolakis, Prentice Hall, 4 edition, 2006

Question 18

- Lossy compression for digital images

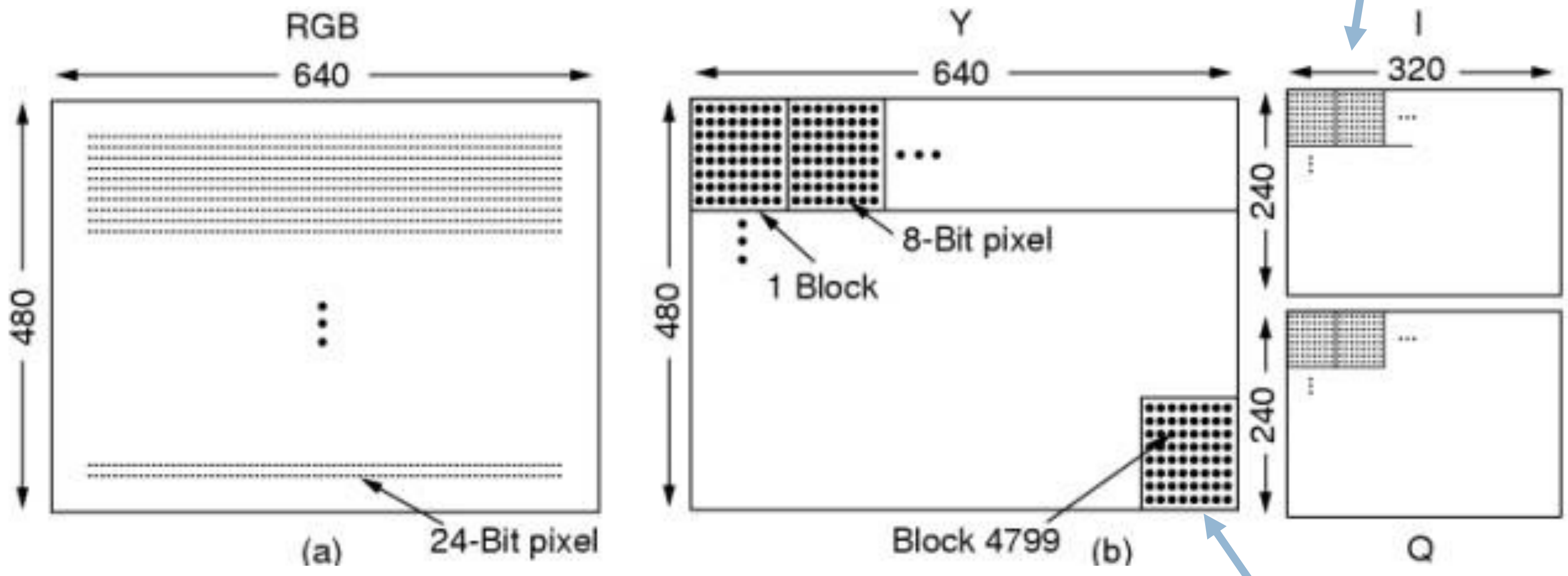
- Question
  - Describe the JPEG compression algorithm

JPEG standard

- JPEG (Joint Photographic Expert Group)
  - developed by experts on behalf of the ISO-IEC
  - International Standard 10918
  - lossy compression for digital images
    - images produced by digital photography
  - degree of compression can be adjusted
    - tradeoff between storage size and image quality.
    - typically achieves 10:1 compression with little perceptible loss in image quality

JPEG encoding - step 1

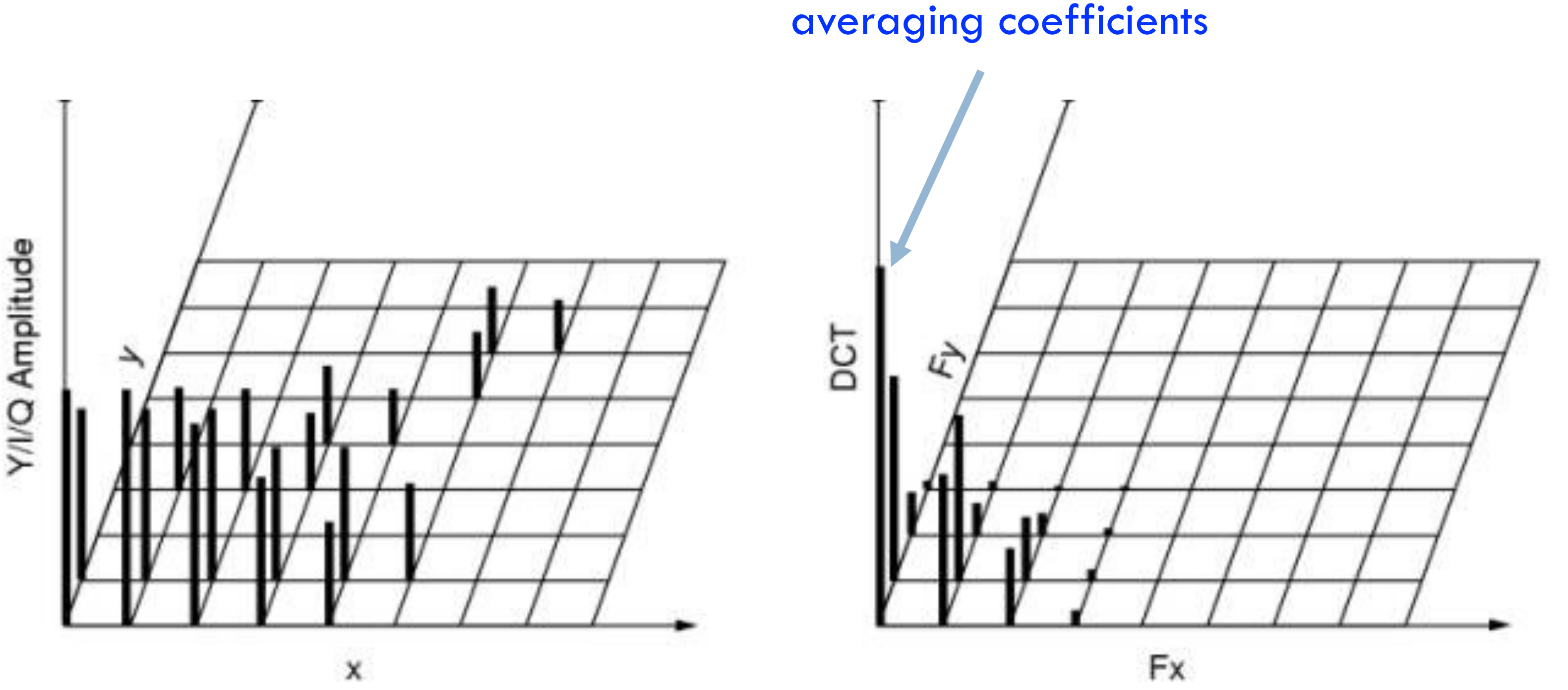
averaging 2x2
block pixels



128 is subtract

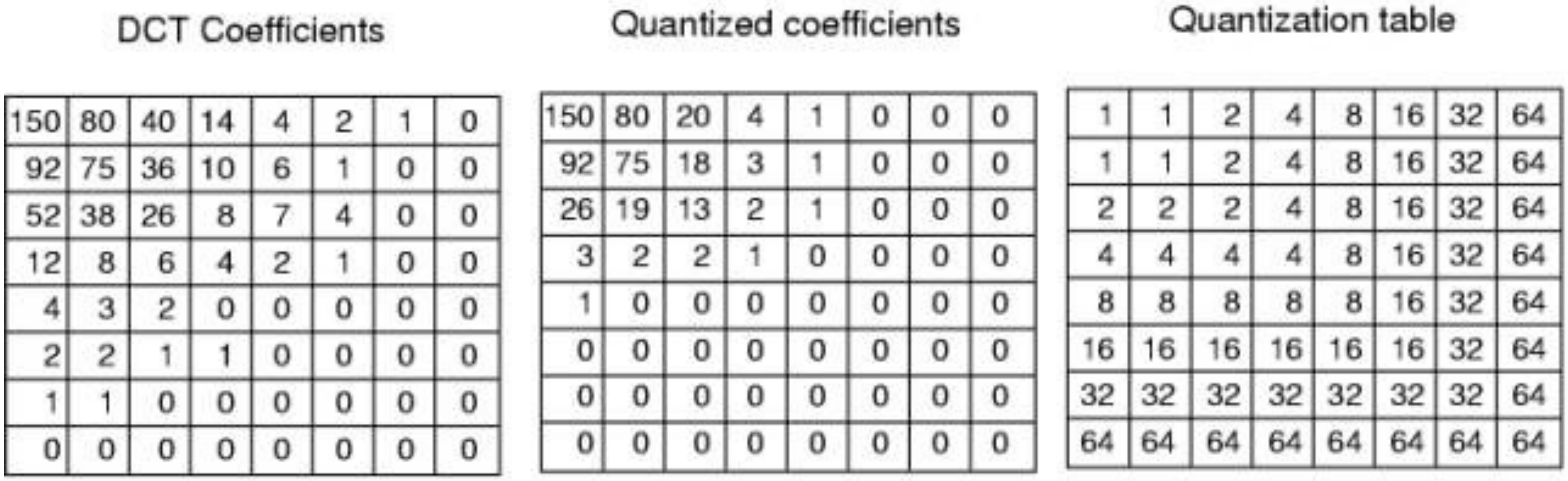
JPEG encoding – YIQ Block preparation

JPEG encoding - step 2



JPEG encoding – DCT coefficients

JPEG encoding – step 3



DCT Coefficients

150	80	40	14	4	2	1	0
92	75	36	10	6	1	0	0
52	38	26	8	7	4	0	0
12	8	6	4	2	1	0	0
4	3	2	0	0	0	0	0
2	2	1	1	0	0	0	0
1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantized coefficients

150	80	20	4	1	0	0	0
92	75	18	3	1	0	0	0
26	19	13	2	1	0	0	0
3	2	2	1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantization table

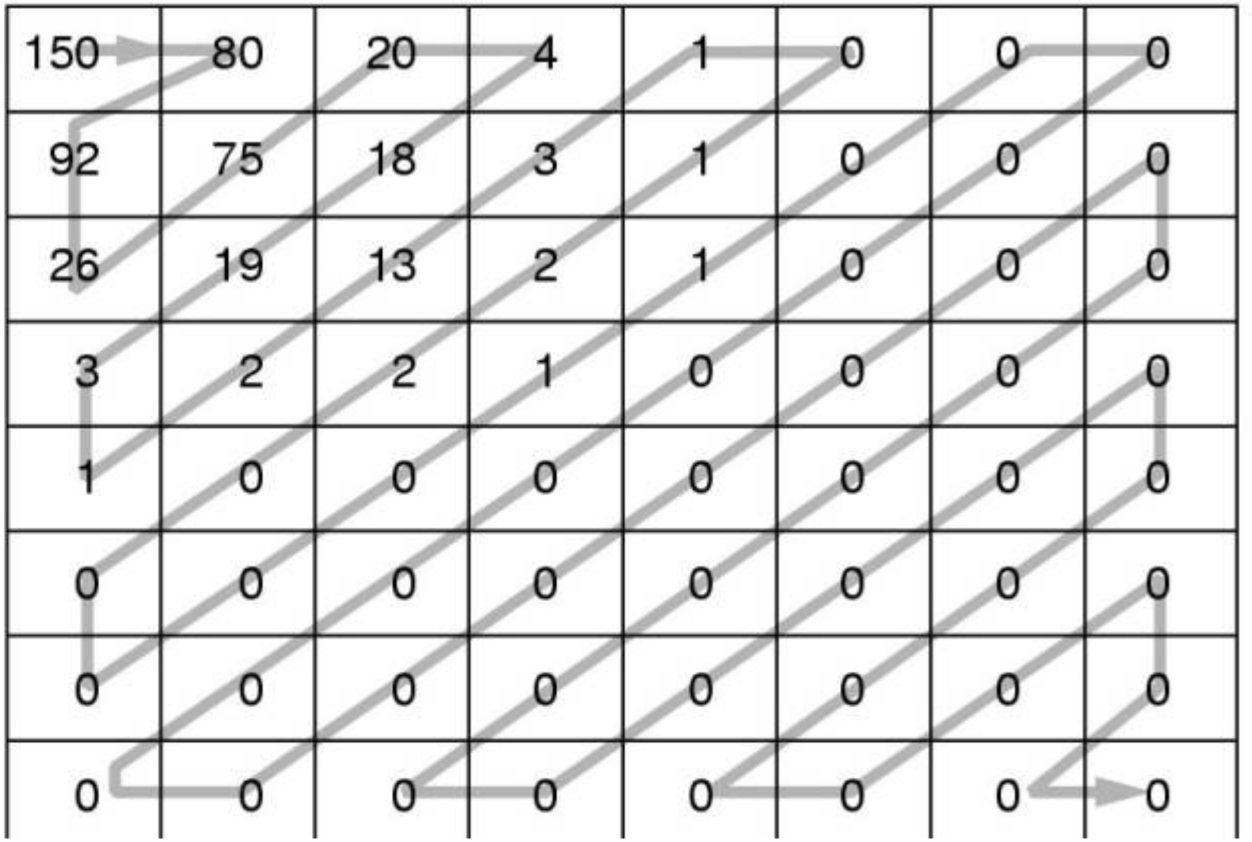
1	1	2	4	8	16	32	64
1	1	2	4	8	16	32	64
2	2	2	4	8	16	32	64
4	4	4	4	8	16	32	64
8	8	8	8	8	16	32	64
16	16	16	16	16	16	32	64
32	32	32	32	32	32	32	64
64	64	64	64	64	64	64	64

JPEG encoding – quantization

JPEG encoding - step 4

- The coefficient (0,0) is substituted by the difference with the same coefficient of the adjacency matrix
  - a low value since the coefficients are similar

JPEG encoding - step 5



150	80	20	4	1	0	0	0
92	75	18	3	1	0	0	0
26	19	13	2	1	0	0	0
3	2	2	1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

JPEG encoding – matrix linearization

RLE is used

JPEG encoding - step 6

- A Huffman encoding scheme is used

- Decoding is obtained by inverting the steps

References

- Material
  - Slides
  - Video Lessons

- Books
  - Fundamentals of Multimedia, Z.-N. Li, M. S. Drew, J. Liu, Springer, 2021

Question 19

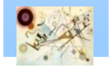
- Lossy compression for digital images

- Question
  - Describe the MPEG 1 standard

DV standard

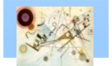
- DV standard
  - each frame is encoded with JPEG
  - *high compression rate*

Source	Mbps	GB/ora
MPEG-2 (640x480)	4	1.76
DV (720x480)	25	11



MPEG

- Moving Picture Experts Group (MPEG)
  - working group of authorities that was formed by ISO and IEC
  - standards for audio and video compression and transmission
  - established in 1988 by the initiative of
    - Hiroshi Yasuda (Nippon Telegraph and Telephone)
    - Leonardo Chiariglione
  - The first meeting was in May 1988 in Ottawa, Canada



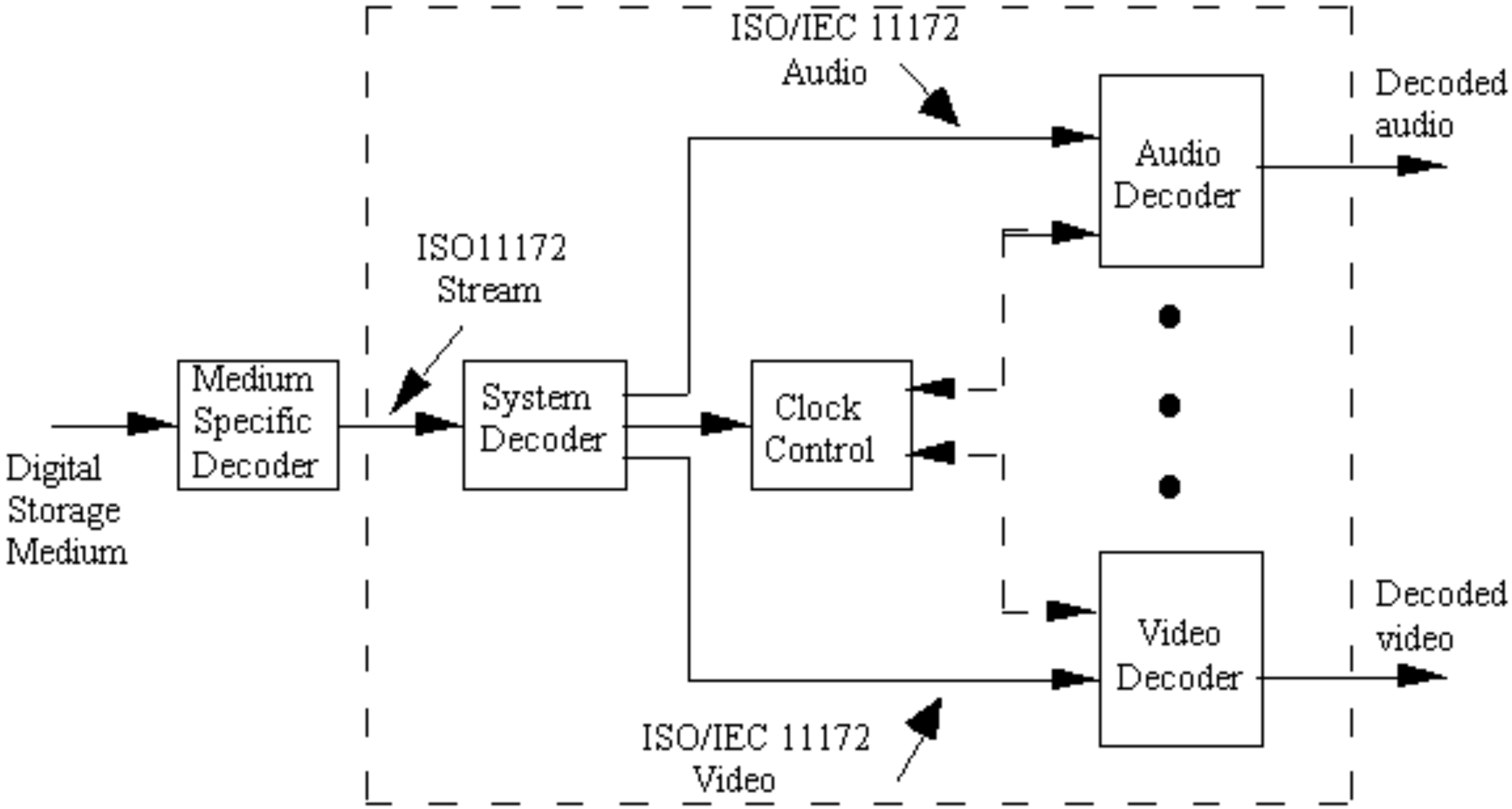
MPEG-1

- MPEG-1
  - standard for lossy compression of video and audio
  - designed to compress VHS-quality raw digital video and CD audio down to 1.5 Mbit/s (26:1 and 6:1 compression ratios respectively)
  - without excessive quality loss
    - video CDs
    - digital cable/satellite TV
    - digital audio broadcasting (DAB)

MPEG-1

- The standard consists of five Parts
  - ISO/IEC 11172-1 (1993)
    - System
  - ISO/IEC 11172-2 (1993)
    - Video
  - ISO/IEC 11172-3 (1993)
    - Audio
  - ISO/IEC 11172-4 (1995)
    - Compliance Testing
  - ISO/IEC TR 11172-5 (1998)
    - Software simulation

MPEG-1 - System

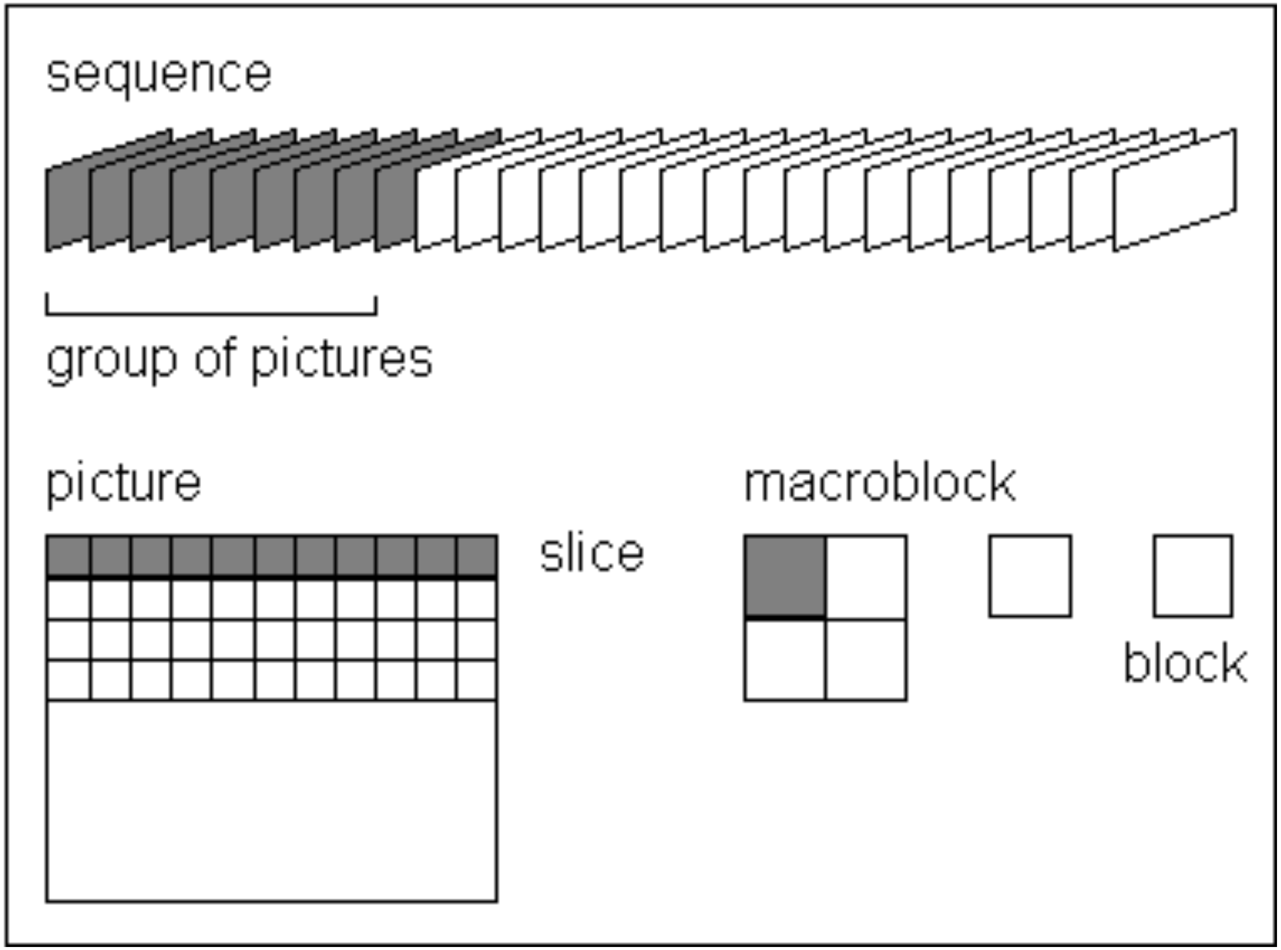


ISO/IEC 11172-1: System

MPEG-1 - Video

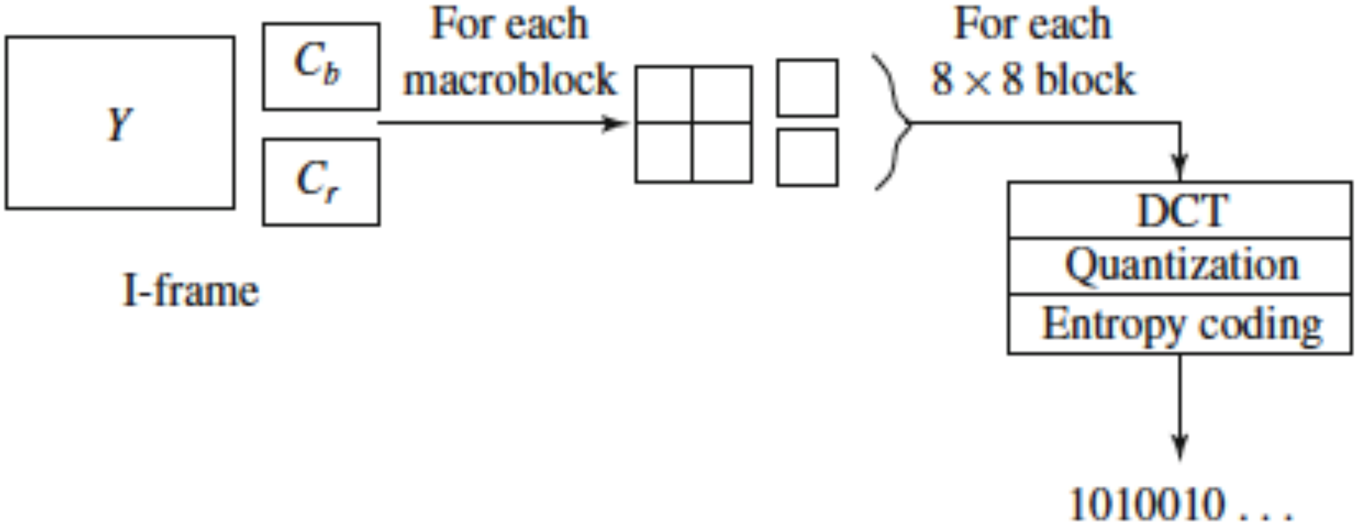
- MPEG-1 has several frame/picture types
  - I-frame (Intra-frame)
    - decoded independently of any other frames
    - can be considered effectively identical to baseline JPEG images
    - also in H.261 encoding standard
  - P-frame (Predicted-frame)
    - also be called forward-predicted frames
    - improve compression by exploiting the temporal redundancy in a video
    - store only the *difference* in image from the frame (either an I-frame or P-frame) immediately preceding it (*anchor frame*)
    - the difference between a P-frame and its anchor frame is calculated using *motion vectors* on each *macroblock* of the frame
    - Motion vector data will be embedded in the P-frame for use by the decoder
    - also in H.261 encoding standard

MPEG-1 - Video



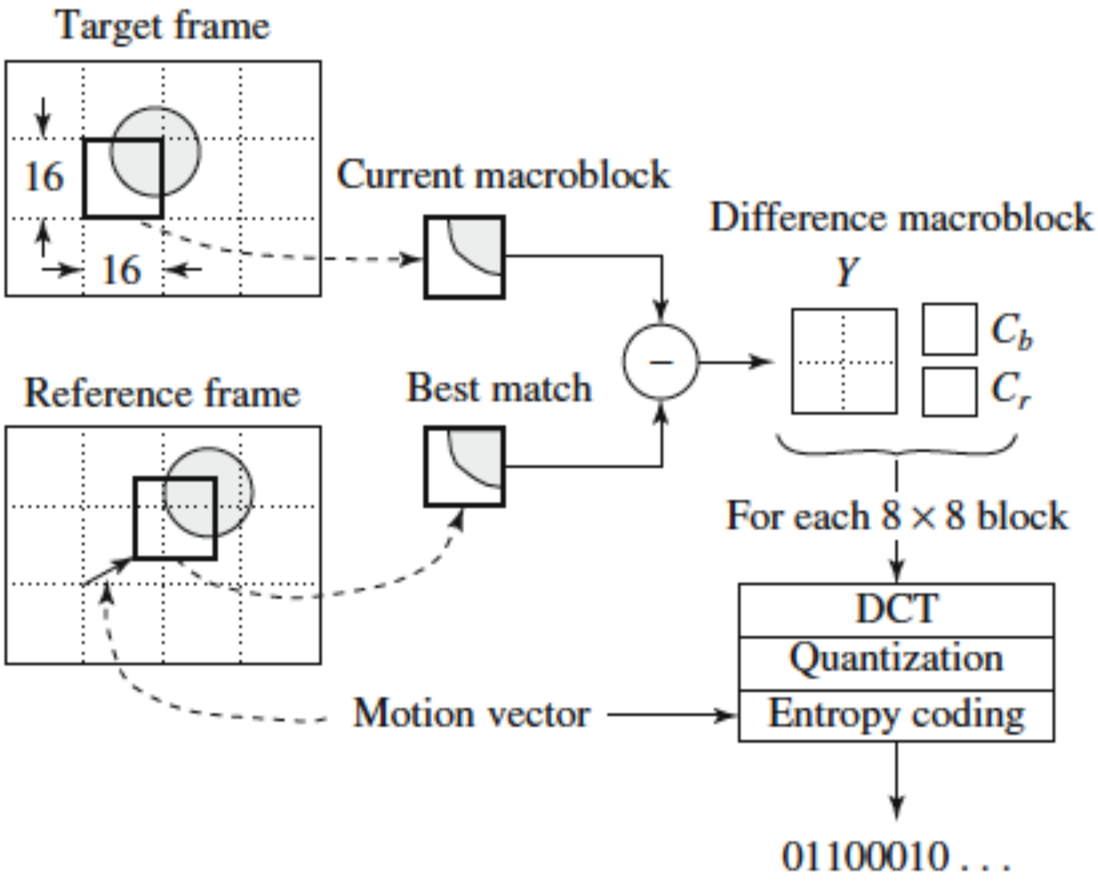
Sequence of pictures and macroblocks

MPEG-1 - Video



I-frame coding

MPEG-1 - Video

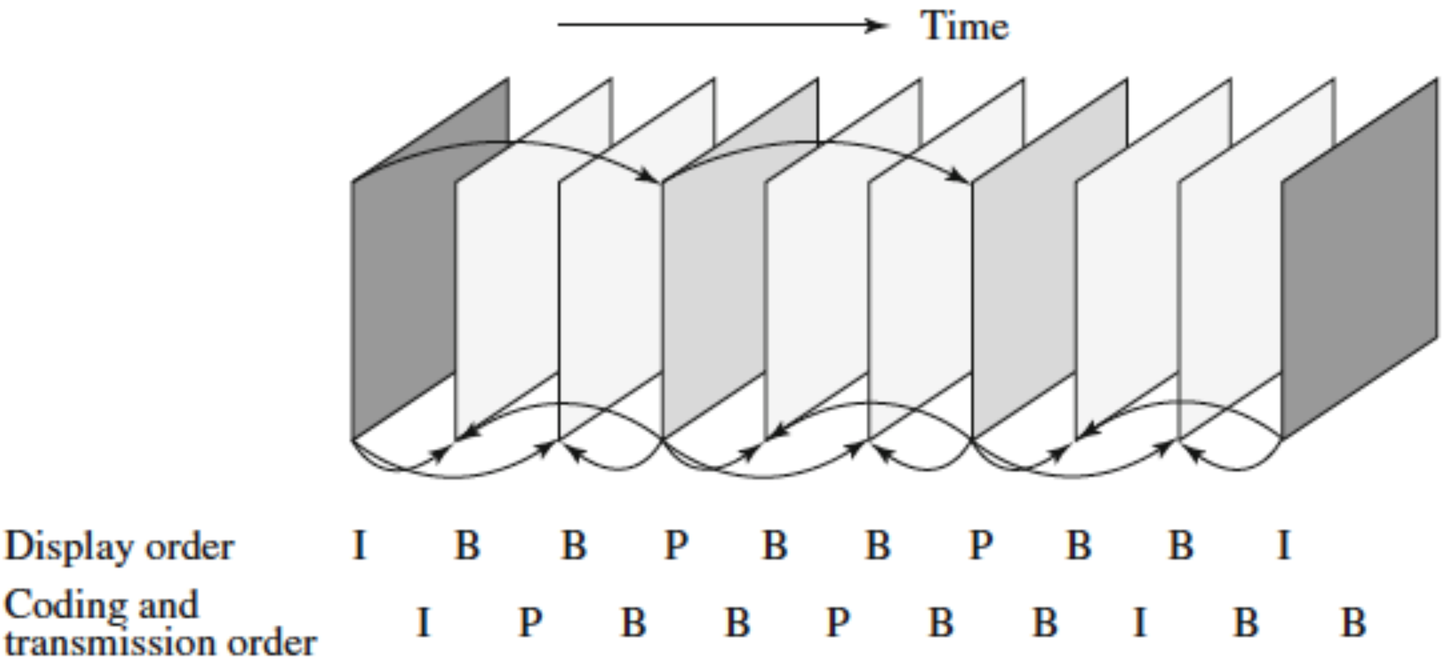


P-frame coding based on motion compensation

MPEG-1 - Video

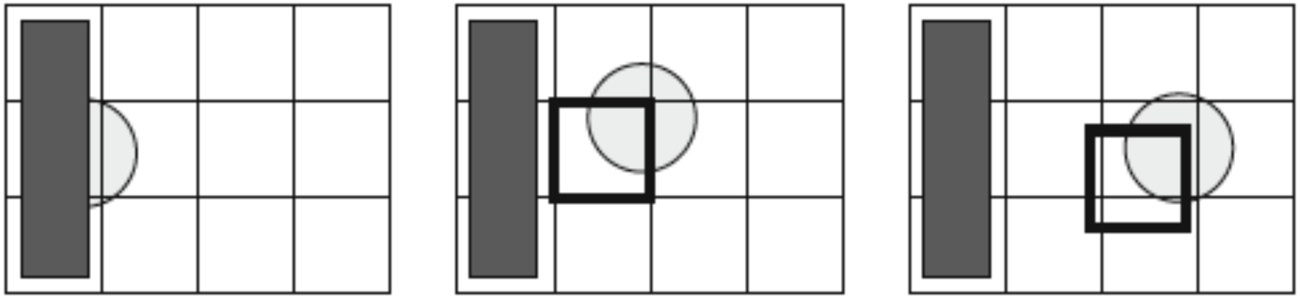
- MPEG-1 has several frame/picture types
  - B-frame (Bidirectional-frame)
    - make predictions using both the previous and future frames (i.e. two anchor frames)
    - requires larger data buffers and causes an increased delay on both decoding and during encoding
  - D-frame
    - independent images (intra-frames) that have been encoded using DC transform coefficients only
    - very low quality
    - are only used for fast previews of video, for instance when seeking through a video at high speed

MPEG-1 - Video



MPEG frame sequence

MPEG-1 - Video



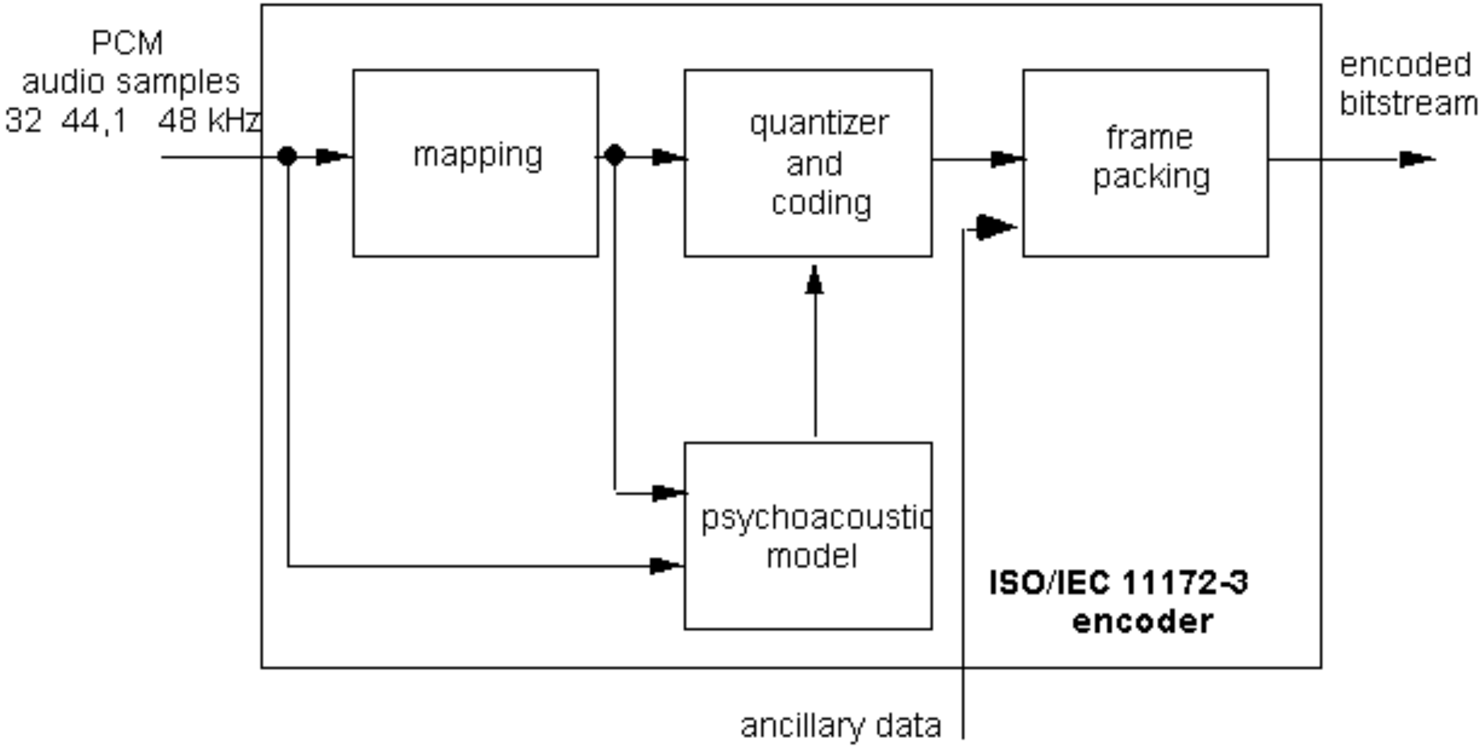
Previous frame

Target frame

Next frame

The need for bidirectional frame

MPEG-1 - Audio



Audio encoding

MPEG-1

- Part 4
  - procedures for testing conformance
  - provides two sets of guidelines and reference bitstreams for testing the conformance of MPEG-1 audio and video decoders, as well as the bitstreams produced by an encoder
- Part 5
  - Reference software
  - C reference code for encoding and decoding of audio and video, as well as multiplexing and demultiplexing

References

- Material
  - Slides
  - Video Lessons
- Books
  - Fundamentals of Multimedia, Z.-N. Li, M. S. Drew, J. Liu, Springer, 2021

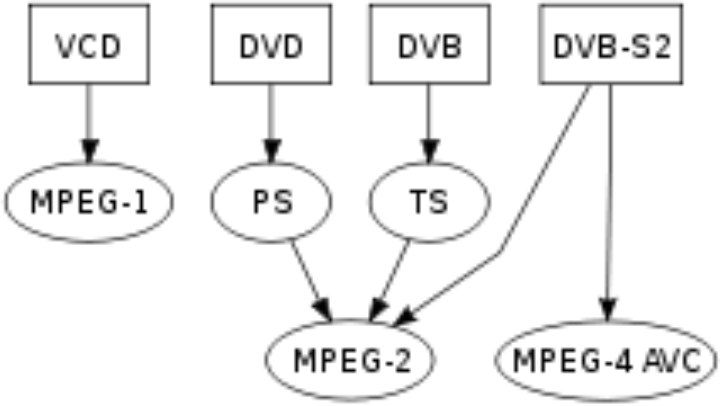
Question 20

- Lossy compression for digital images

- Question
  - Describe the MPEG 2 standard

MPEG-2

- MPEG-2
  - generic coding of moving pictures and associated audio information
  - combination of lossy video compression and lossy audio data compression methods
  - storage and transmission of movies using currently available storage media and transmission bandwidth



MPEG-2 is used in Digital Video Broadcast and DVDs.
The MPEG transport stream, TS, and MPEG program stream, PS, are container formats

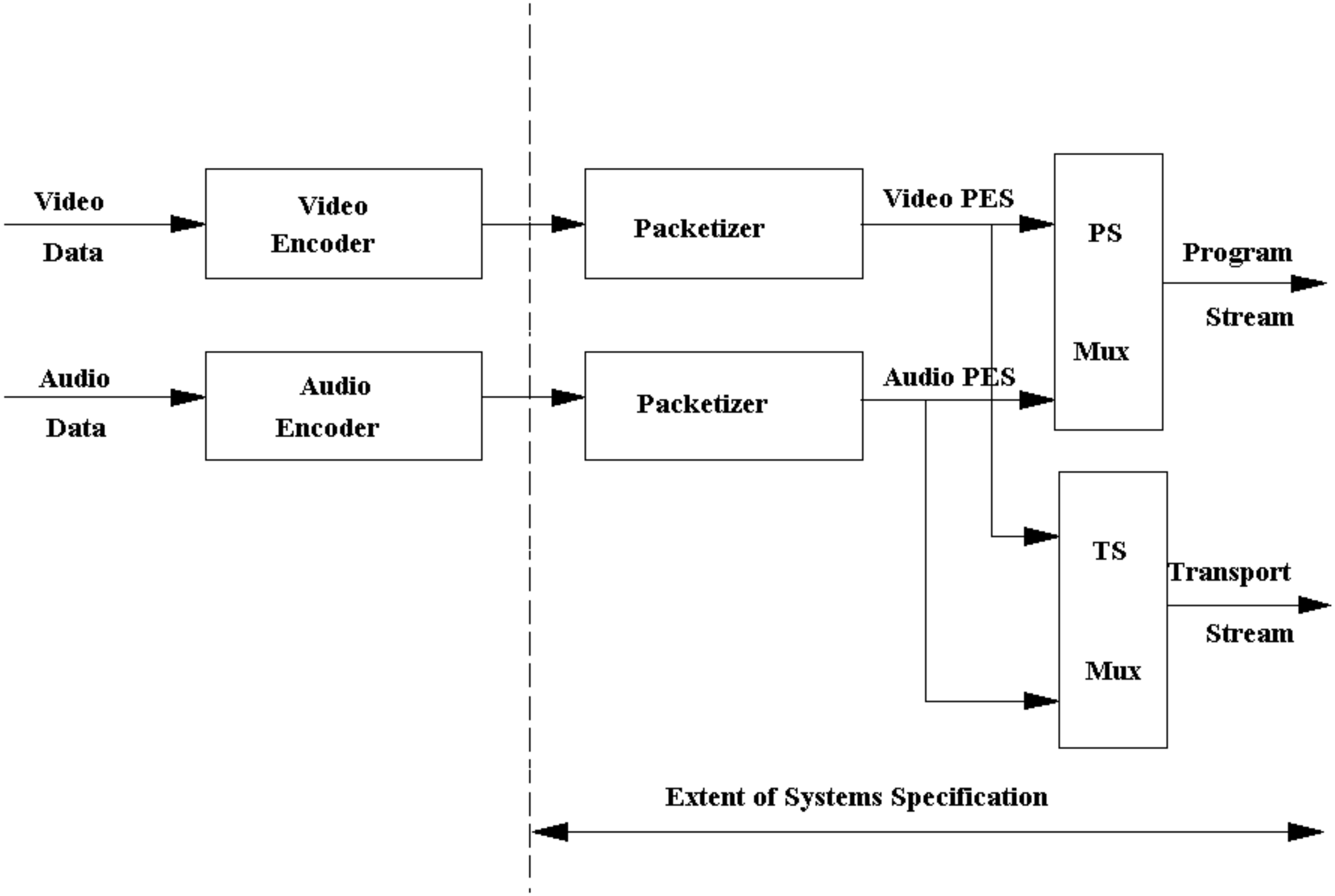
MPEG-2

- The standard consists of 9 Parts
  - ISO/IEC 13818-1 (2000)
    - Systems
  - ISO/IEC 13818-2 (2000)
    - Video
  - ISO/IEC 13818-3 (1998)
    - Audio
  - ISO/IEC 13818-4 (1998)
    - Conformance Testing
  - ISO/IEC 13818-1 (1997)
    - Software simulation

MPEG-2

- The standard consists of 9 Parts
  - ISO/IEC 13818-6 (1998)
    - Extensions for DSM-CC
  - ISO/IEC 13818-7 (1997)
    - Advanced Audio Coding (AAC
  - ISO/IEC 13818-8 (1996)
    - Extension for real time interface for systems decoders
  - ISO/IEC 13818-9 (1999)
    - Conformance extensions for Digital Storage Media Command and Control (DSM-CC)

MPEG-2 - System



ISO/IEC 13818 - System

MPEG-2 - Video

- Video encoding
  - similar to the previous MPEG-1 standard
  - provides support for interlaced video, the format used by analog broadcast TV systems
  - MPEG-2 Video and Systems are also used in some HDTV transmission systems

MPEG-2 - Audio

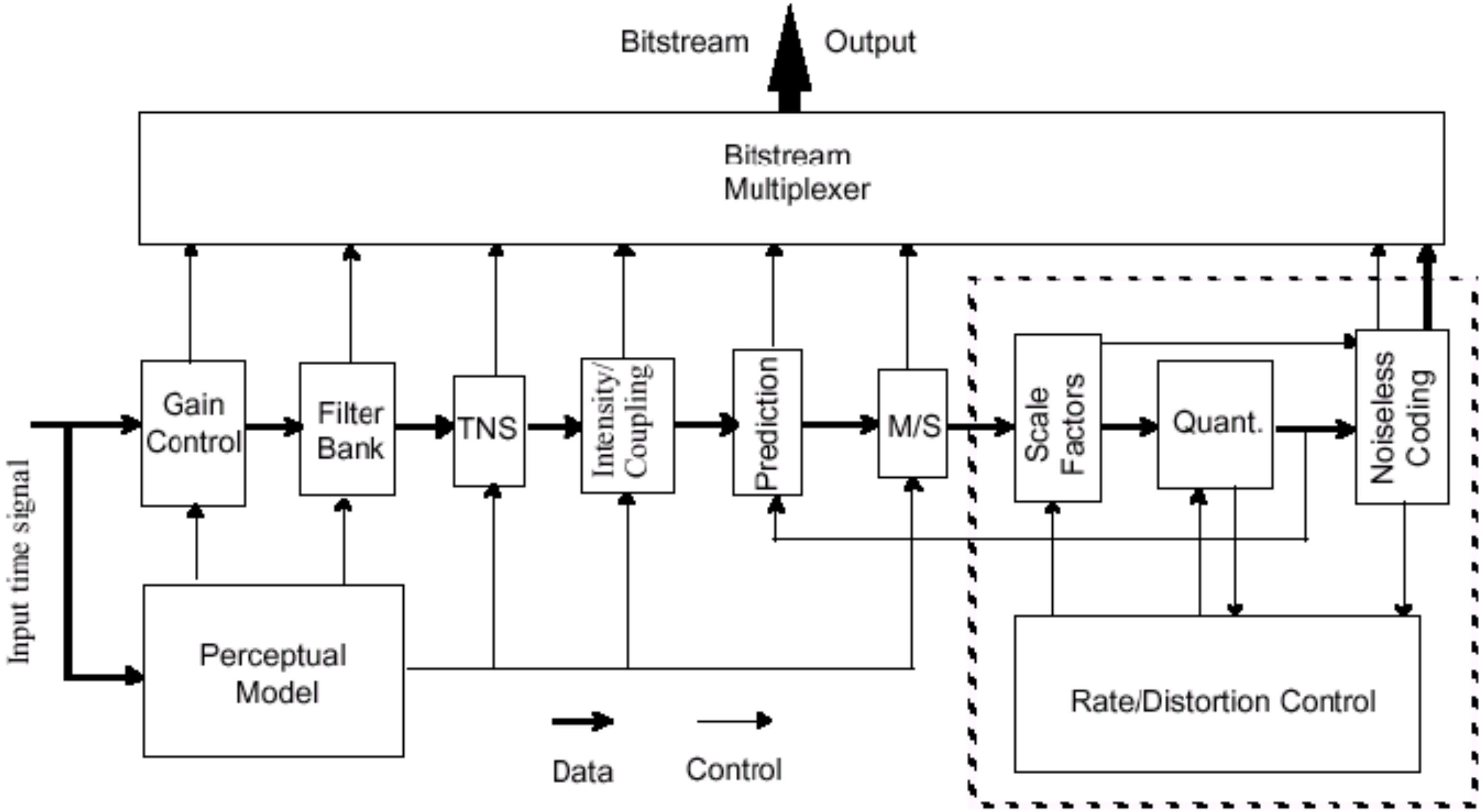
- Audio encoding
  - MPEG-2 introduces new audio encoding methods compared to MPEG-1
    - MPEG-2 Part 3
      - enhances MPEG-1's audio by allowing the coding of audio programs with more than two channels, up to 5.1 multichannel
    - MPEG-2 Part 7
      - specifies a rather different, non-backwards-compatible audio format
      - is referred to as MPEG-2 AAC (Advanced Audio Coding)
      - AAC is more efficient

ISP – Verification tests

MPEG 2 - AAC

- Adanced Audio Coding (AAC)
  - improvement for multichannel encoding
  - 48 channels
  - samplig frequency from 8 to 96 KHz for each channel

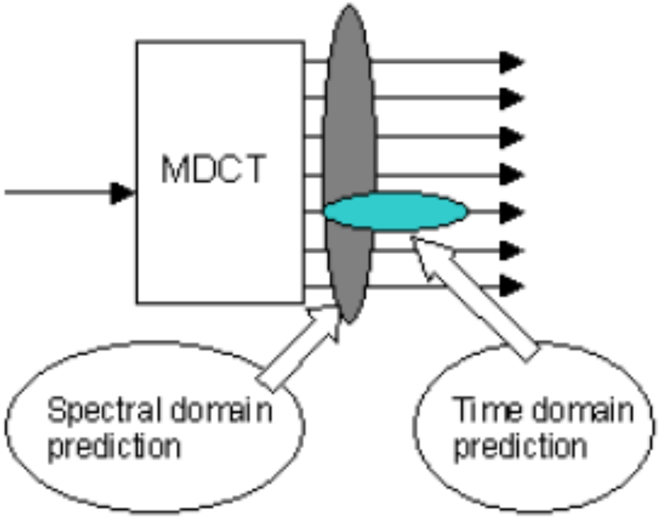
MPEG 2 - AAC



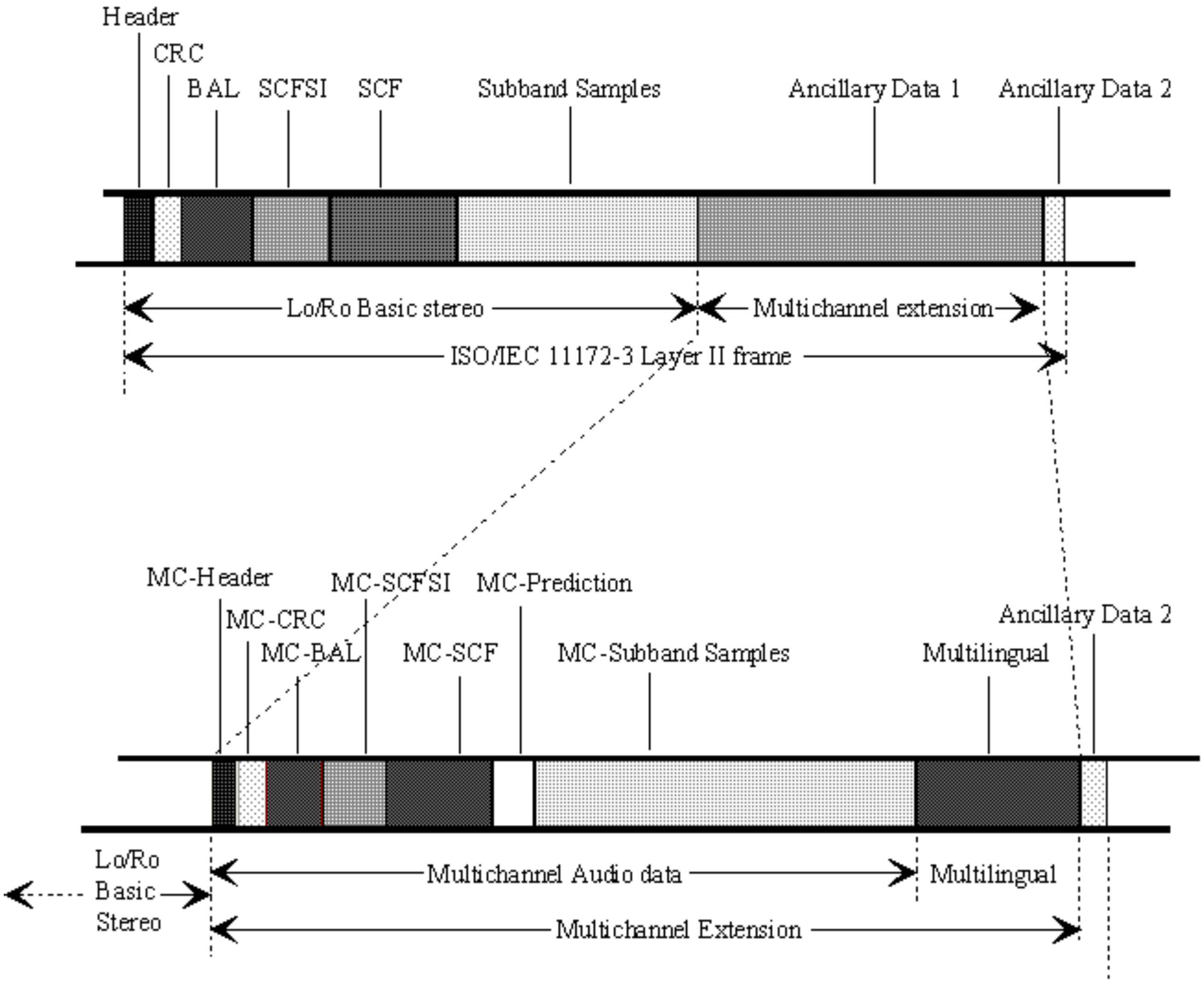
AAC encoding scheme

MPEG 2 - AAC

- Main concept – prediction
  - Prediction
  - Temporal Noise Shaping (TNS)



MPEG-2 - Audio



Multichannel Audio information

References

- Material
  - Slides
  - Video Lessons

- Books
  - Fundamentals of Multimedia, Z.-N. Li, M. S. Drew, J. Liu, Springer, 2021

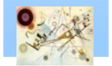
Question 21

- Lossy compression for digital images

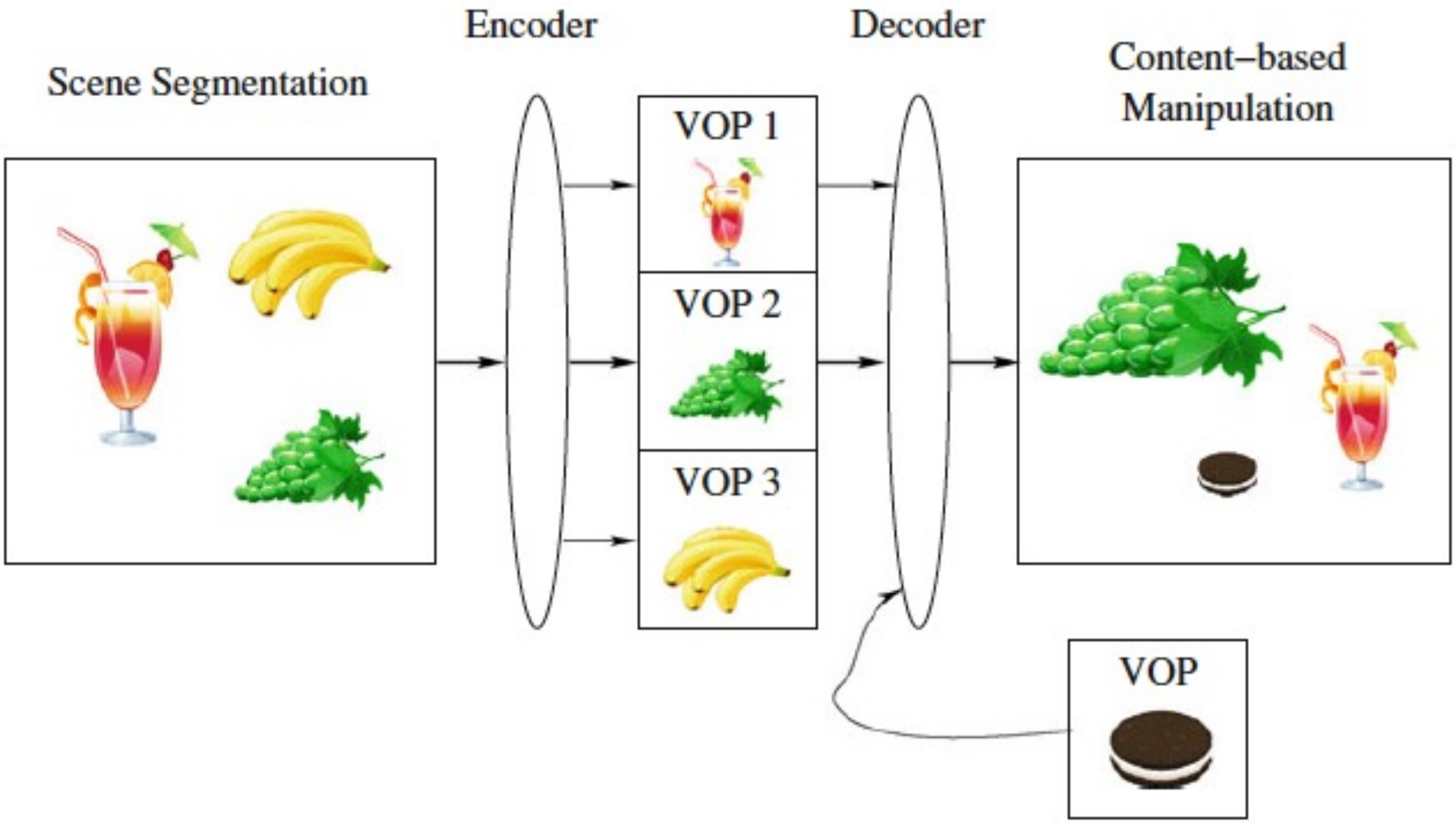
- Question
  - Describe the MPEG 4 standard

MPEG-4

- Besides compression, it pays great attention to user interactivity
  - allows a larger number of users to create and communicate their multimedia presentations and applications on new infrastructures
    - Internet, mobile/wireless networks, …
  - adopt a new object-based coding approach
    - *media objects* are entities
    - media objects (audio and visual objects) can be either natural or synthetic
- bitrate covers a large range, between 5 kbps and 10 Mbps

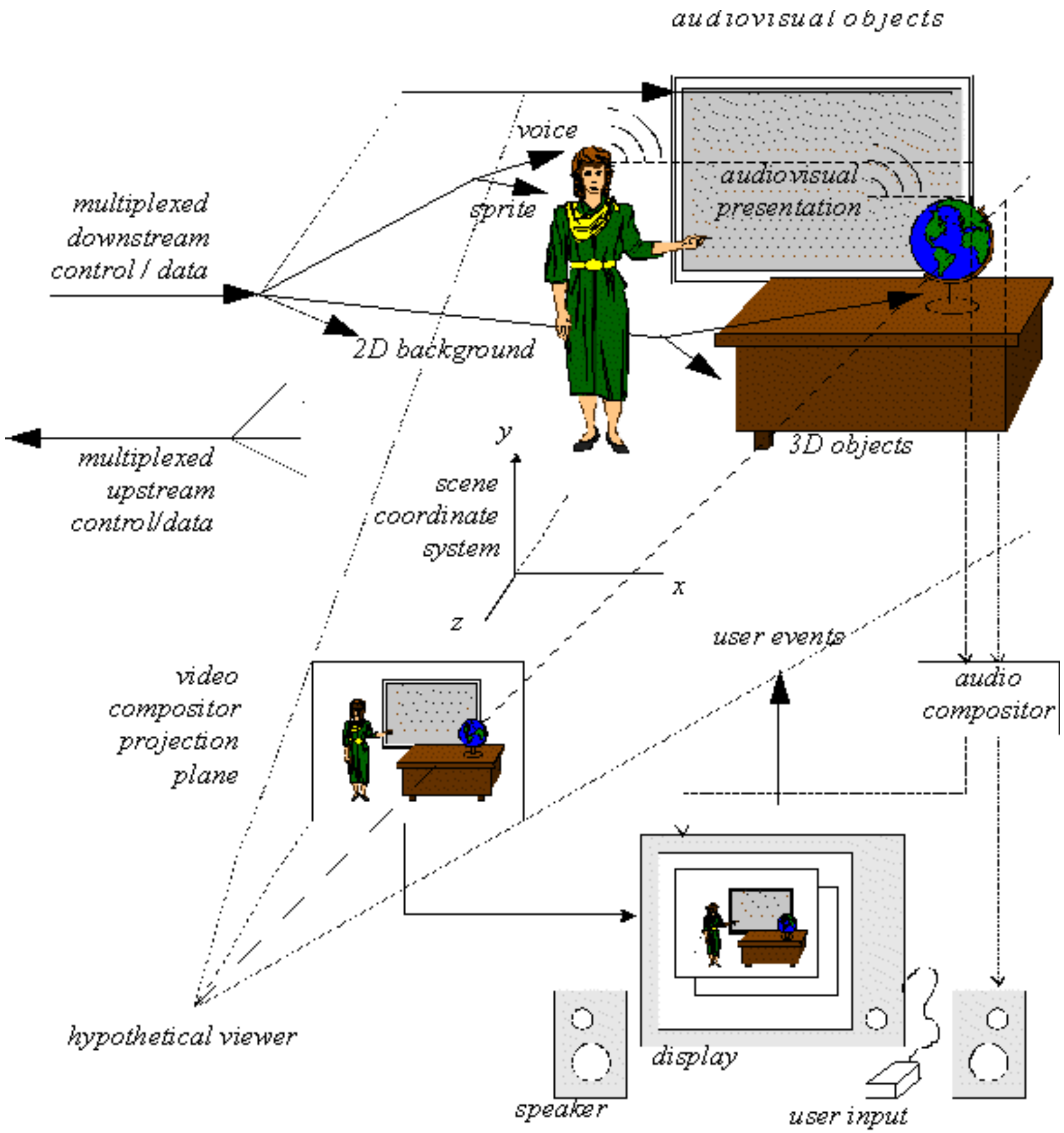


MPEG-4



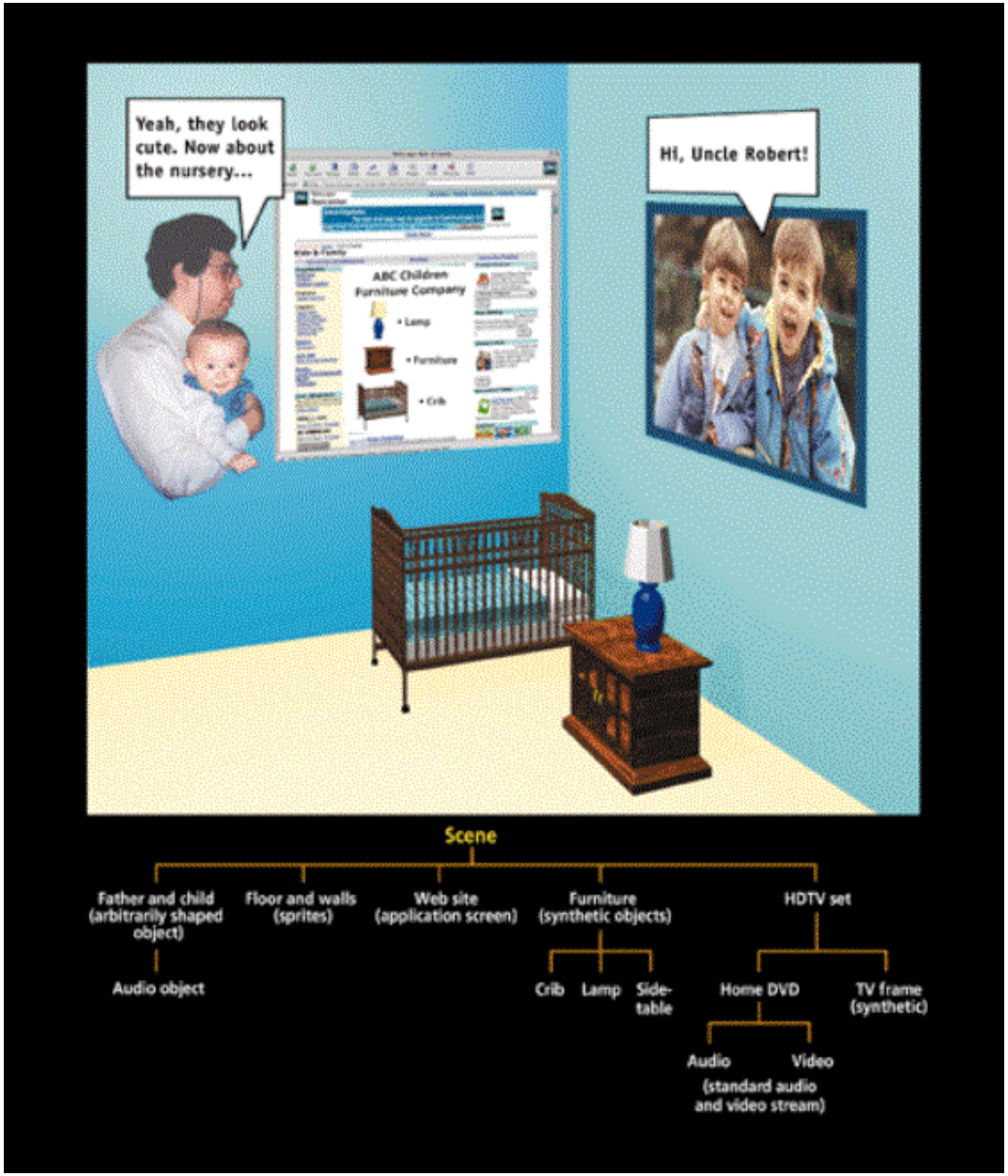
Composition and manipulation of MPEG-4 videos (VOP = Video object plane)

MPEG-4



Example of a MPEG-4 scene

MPEG-4

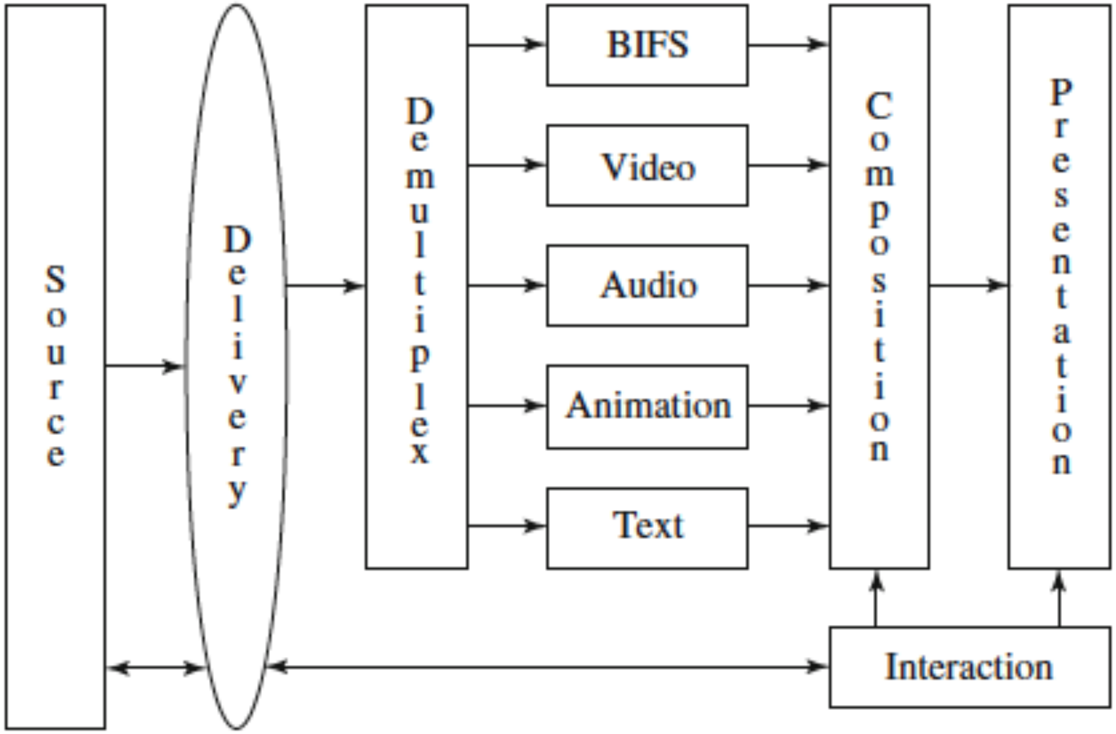


Example of a MPEG-4 scene

Hierachical scene composition

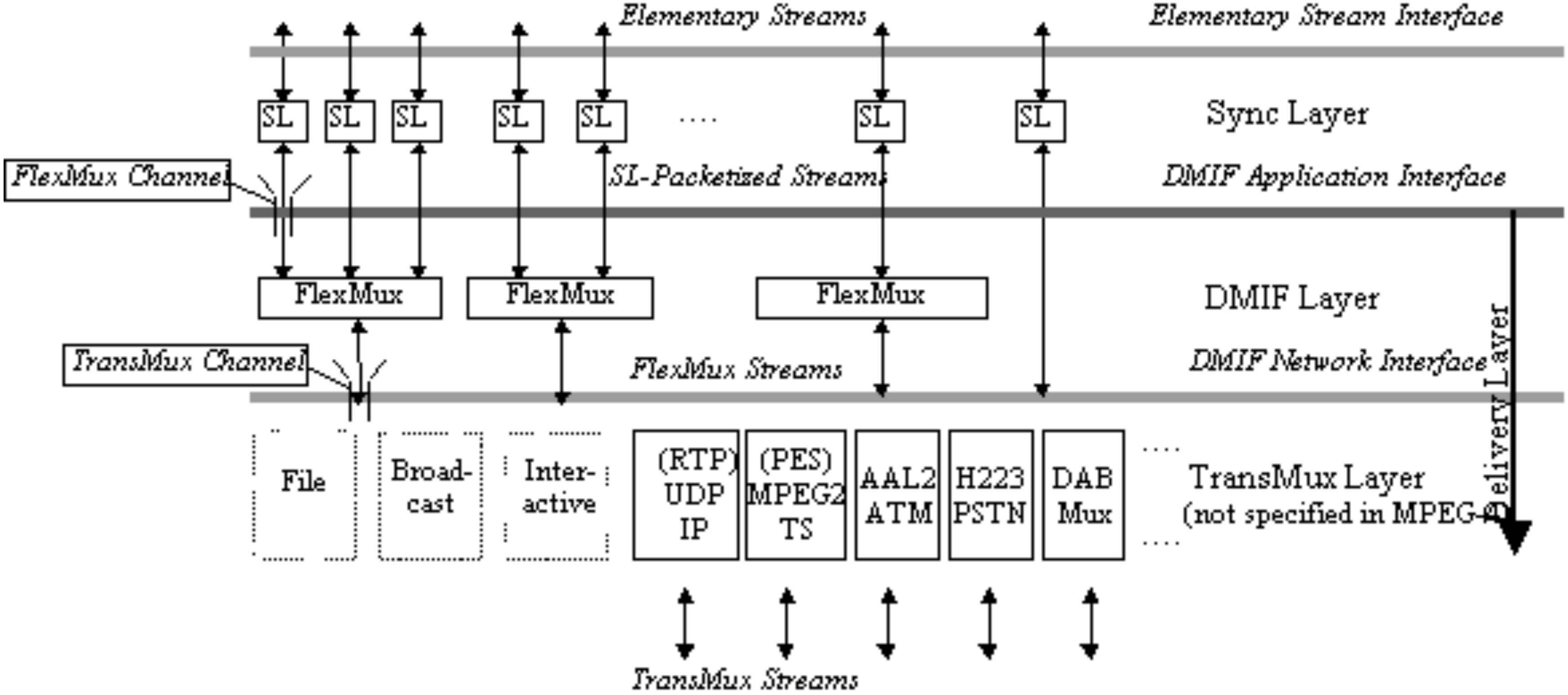
ISP – Verification tests

MPEG-4



MPEG-4 reference model

MPEG-4

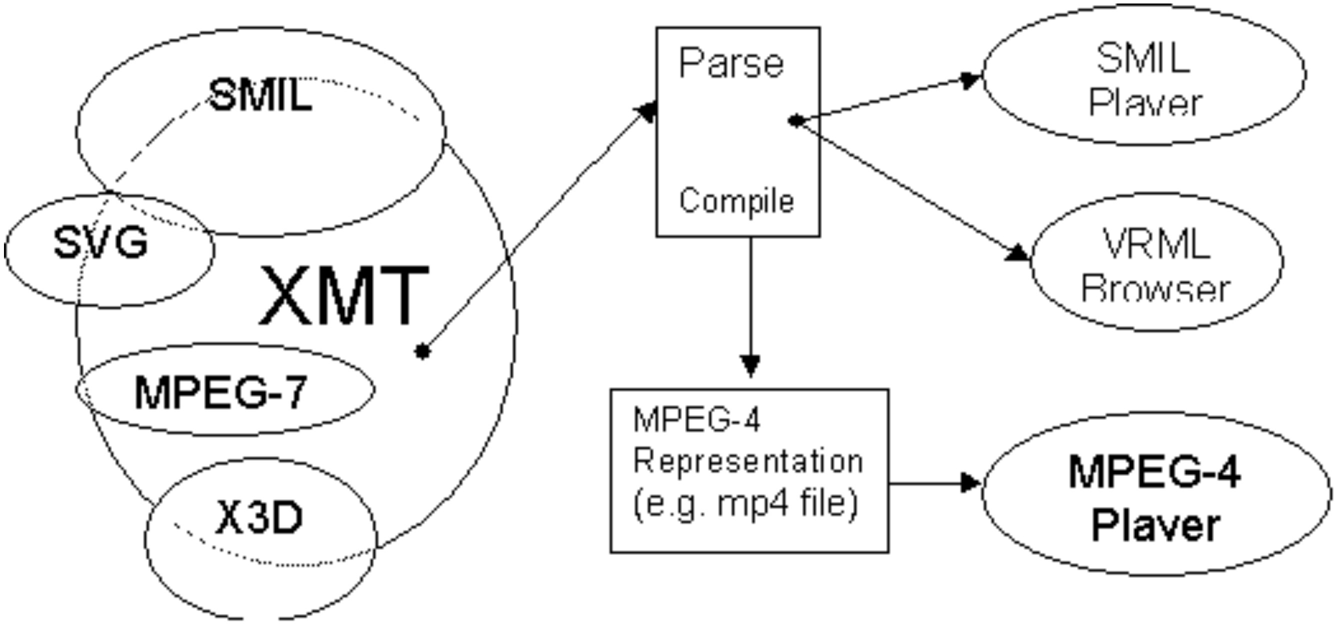


Layers of the system

BIFS

- BIFS - Binary Format for Scenes
  - facilitates the composition of media object in the scene
  - scene graph
    - nodes – audiovisual primitives and attributes
    - graph structure – spatial and temporal relationship of objects in the scene
  - enhancement of Virtual Reality Modeling Language (VRML)

BIFS

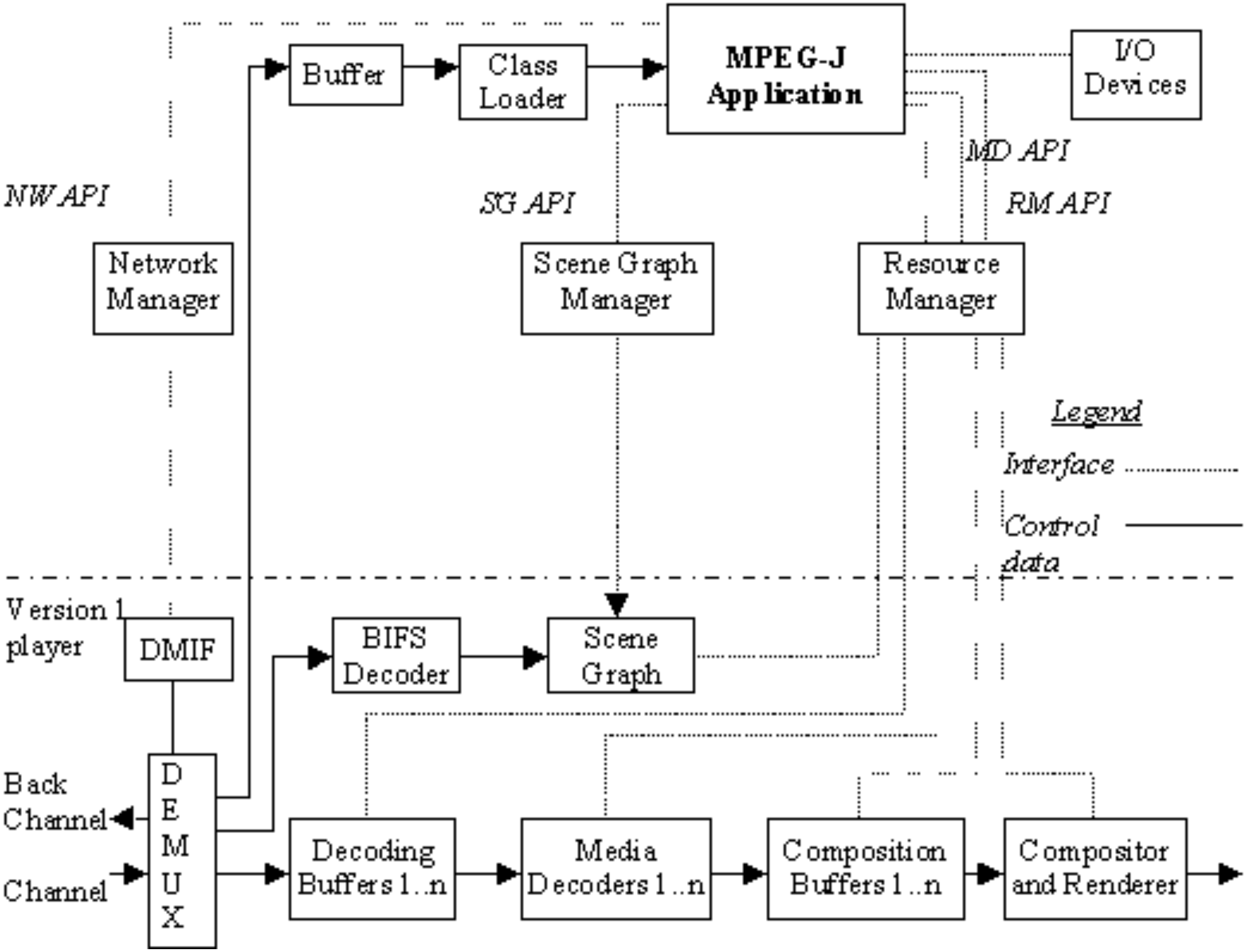


BIFS interfaces

MPEG-J

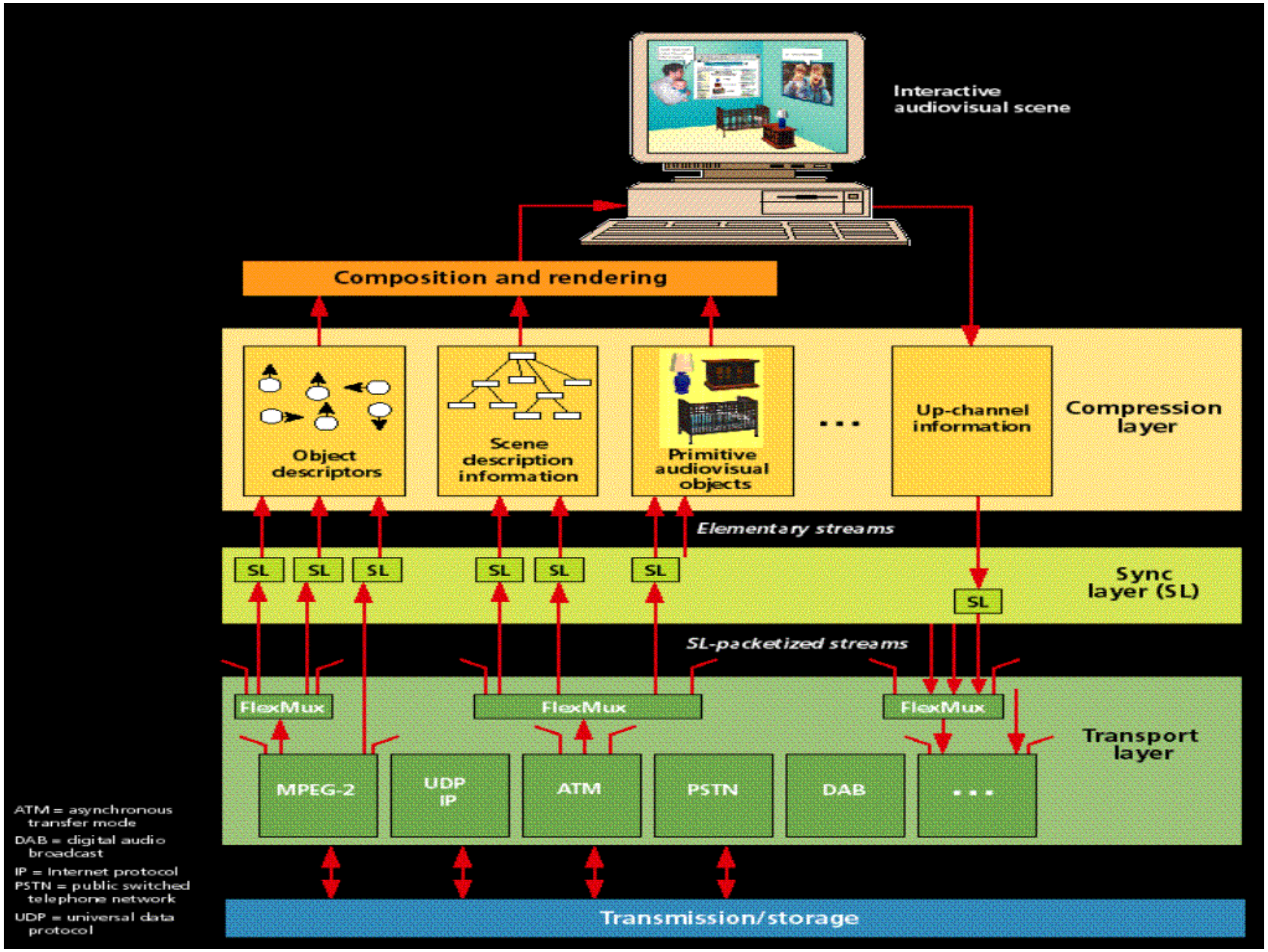
- MPEG-J
  - programming environment
  - Java applications can access Java packages and APIs and enhance users' interactivity

MPEG-J



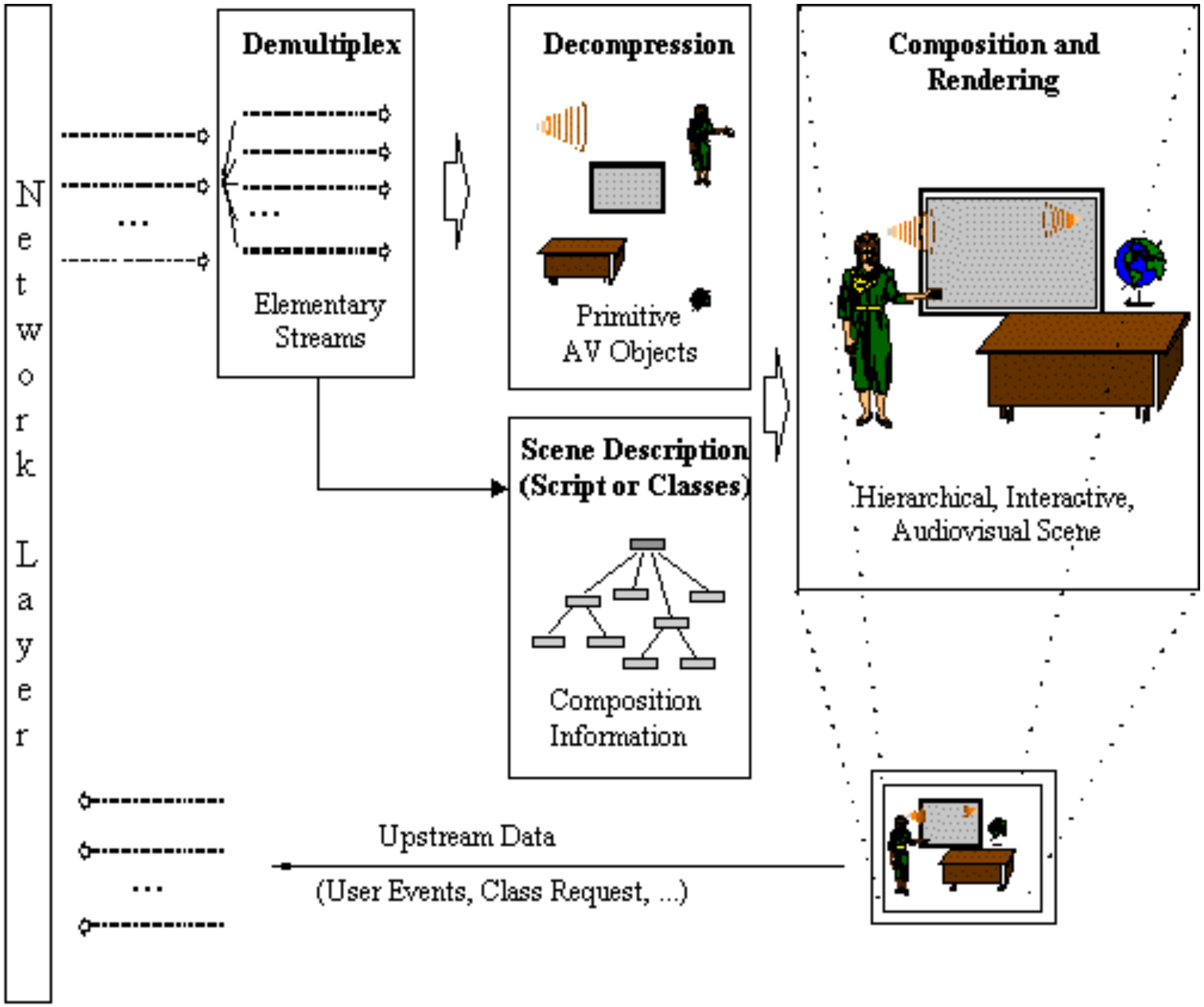
MPEG-J interfaces

MPEG-4



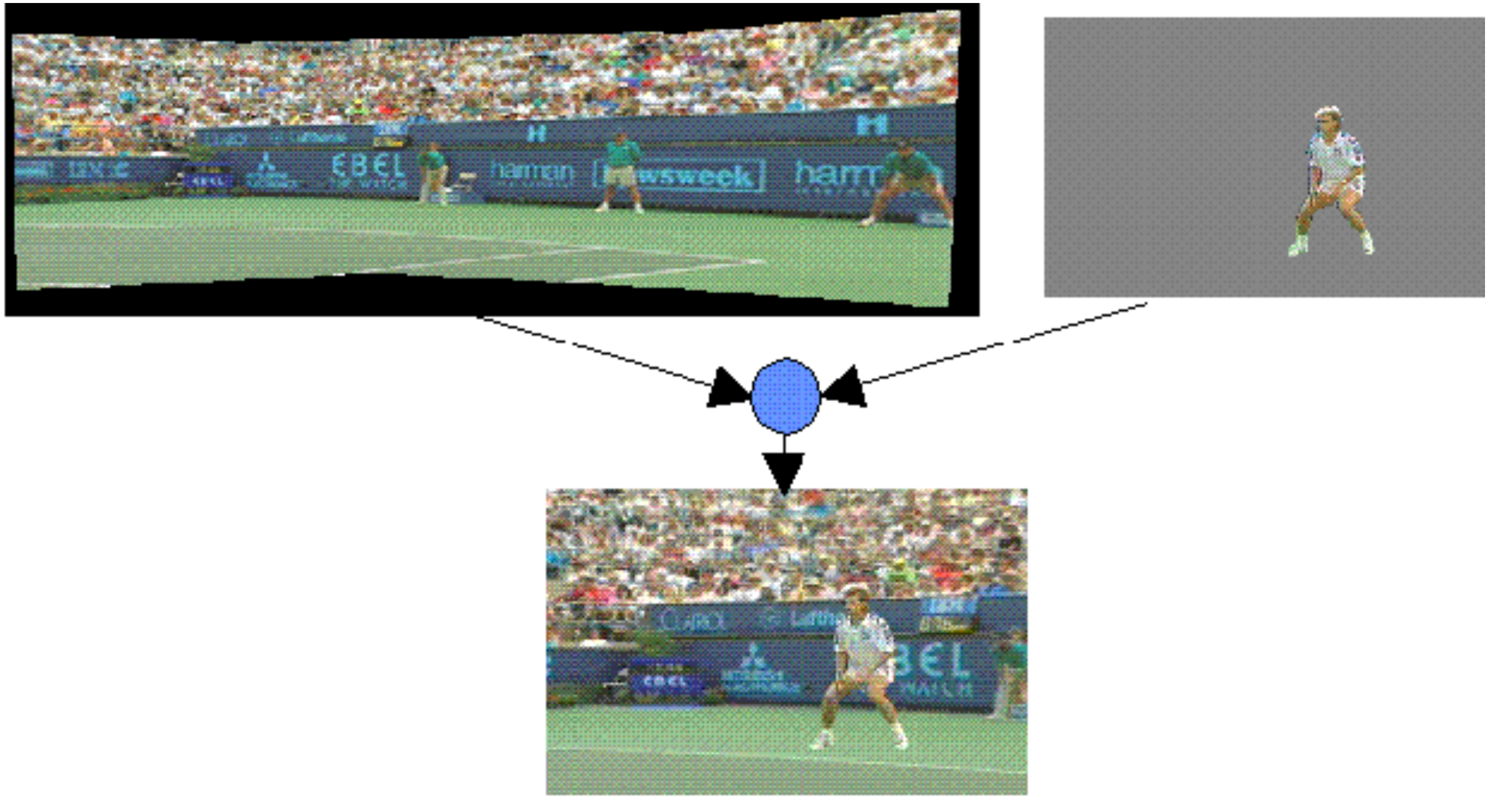
MPEG-4 components and layers

MPEG-4



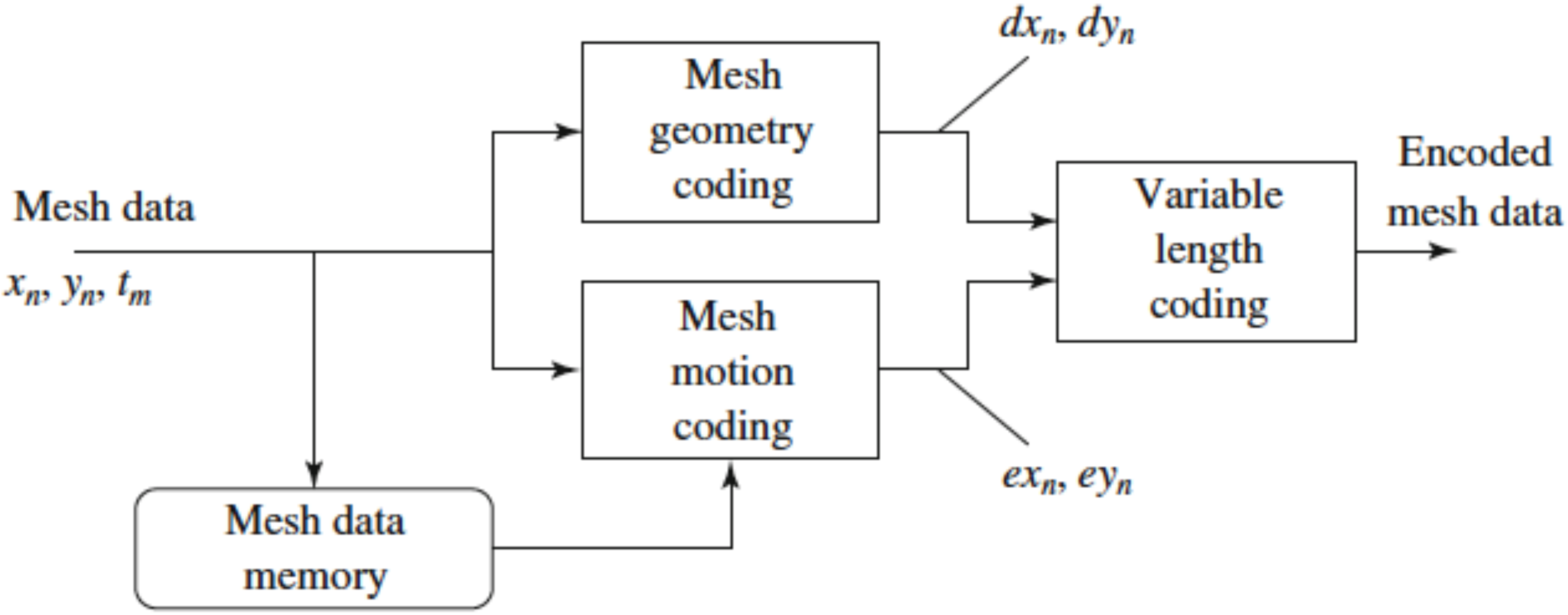
Decoding, composition and rendering

Sprite coding



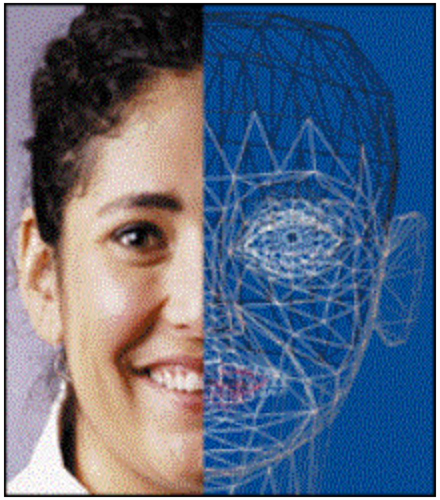
Example of sprite coding to compose an image

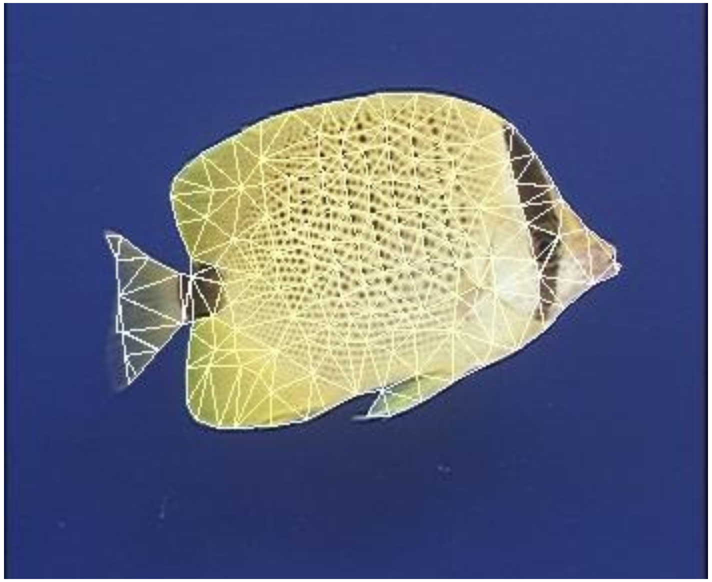
Synthetic objects



2D Mesh object plane encoding process

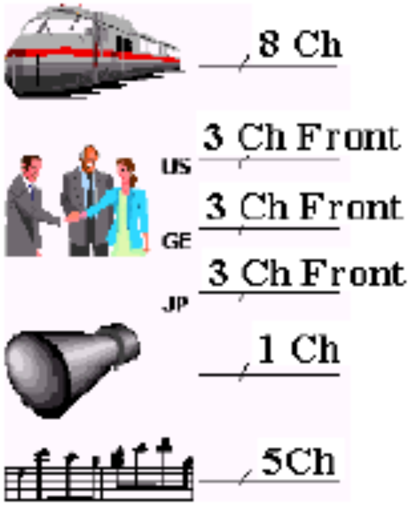
Synthetic objects





Examples of mashes applications in MPEG-4 (avatar)

MPEG 4



MPEG 4 considers each audio as an independent object

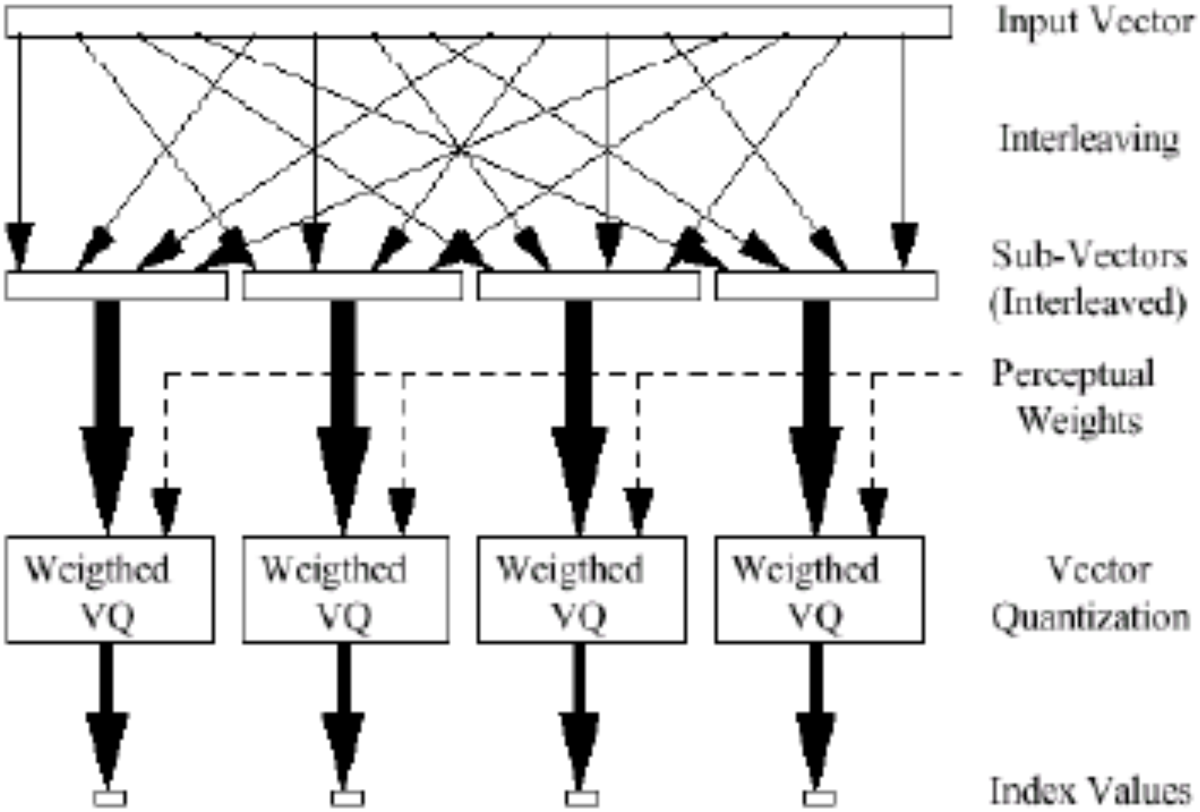
MPEG 4 – Speech Signal

- Speech signal
  - Synthesis Decoding Code Excited Linear Predictive (CELP)
    - Bitrate from 4 to 24 Kbit/s
  - Harmonic vector eXcitation Coding (HVXC)
    - Bitrate from 2 to 4 Kbit/s

General Audio

- General Audio
  - Transform-domain Weighted Interleaved Vector Quanrization (TwinVQ)
    - less than 16 Kbit/s
  - AAC for greater bitrates

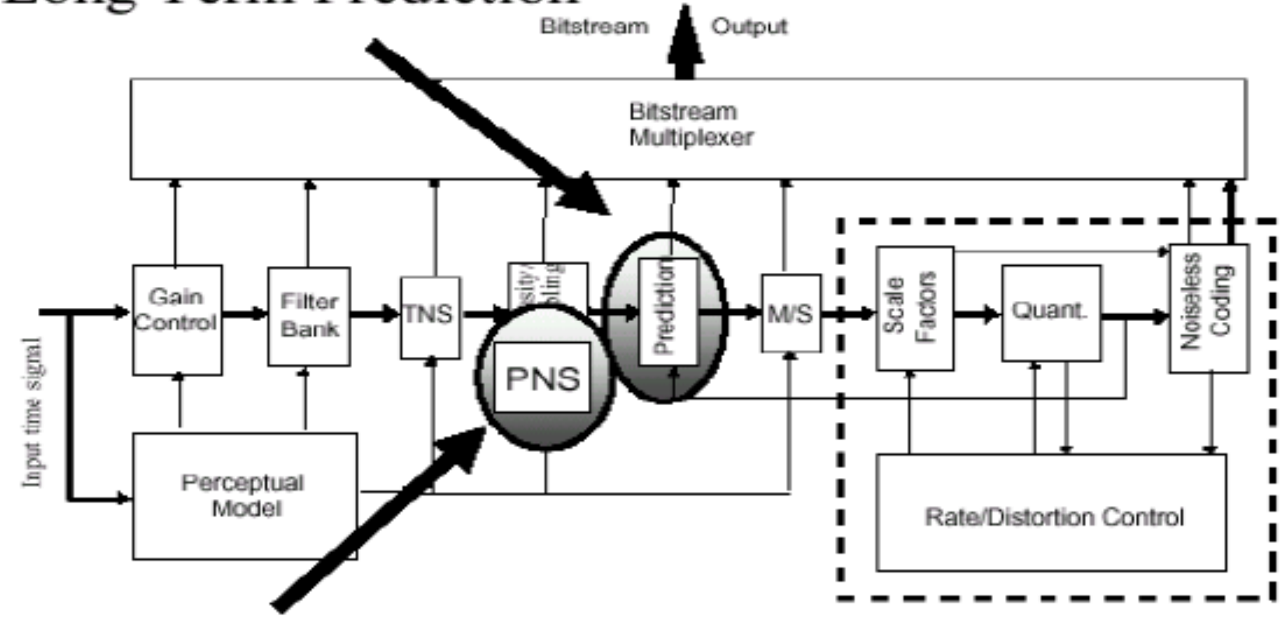
General Audio



TwinVQ scheme

AAC

Long Term Prediction



Perceptual Noise Substitution

TwinVQ scheme

Synthesized Speech

- Text to Speech
  - production of a sound voice from a text
  - Interface for compressed data

Synthesized Audio

- Structured Audio Orchestra Language (SAOL)
  - Set of musical instruments for reproducing
- Structured Audio Score Language (SASL)
  - what to produce

References

- Material
  - Slides
  - Video Lessons

- Books
  - Fundamentals of Multimedia, Z.-N. Li, M. S. Drew, J. Liu, Springer, 2021

Question 22

- Buffering strategy multimedia content

- Question
  - Describe the Organ-Pipe Algorithm

Multiple files

- On a Video Server
  - time will be wasted moving the disk head from movie to movie when multiple movies are being viewed simultaneously by different customers
- Observation
  - some movies are more popular than others
  - taking popularity into account when placing movies on the disk

Zipf's law

- Zipf's law
  - George Zipf, Harvard professor of linguistics
  - if the movies, books, Web pages, or words are ranked on their popularity, the probability that the next customer will choose the item ranked k-th in the list is C/k
  - If there are N movies, C is computed such that

$$\boxed{C + C/2 + C/3 + \ldots + C/N = 1}$$

Zipf's law

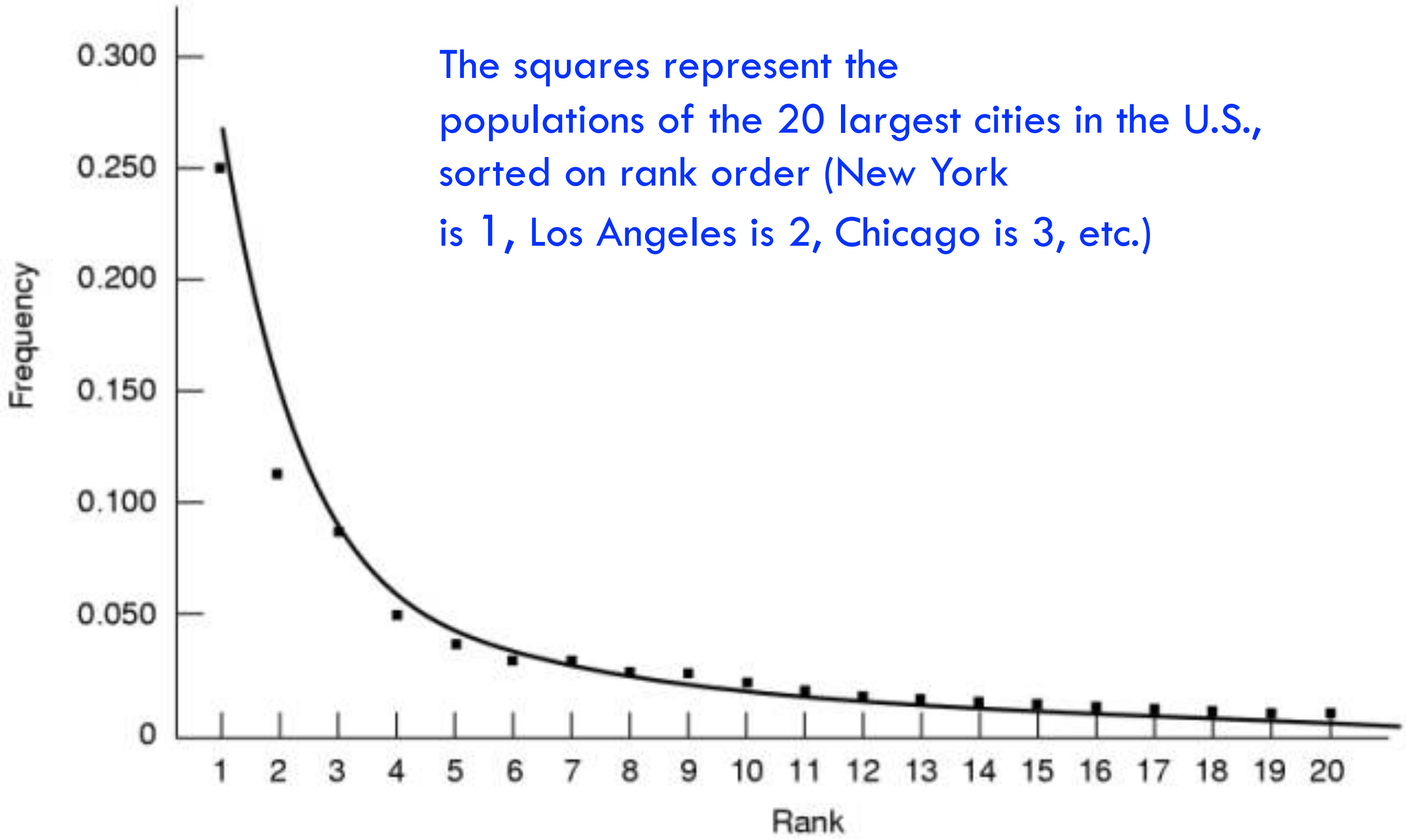
N. population	*C*
10	0.341
100	0.193
1000	0.134
10000	0.102

C values varying *N*

N. movies	*Probabilities*
1	0.134
2	0.067
3	0.045
4	0.034
5	0.027

Probabilities
for the top five movies
with *N*=1000

Zipf's law



The squares represent the populations of the 20 largest cities in the U.S., sorted on rank order (New York is 1, Los Angeles is 2, Chicago is 3, etc.)

Zipf's law predicts that the second largest city should have a population half of the largest city and the third largest city should be one third of the largest city, and so on.

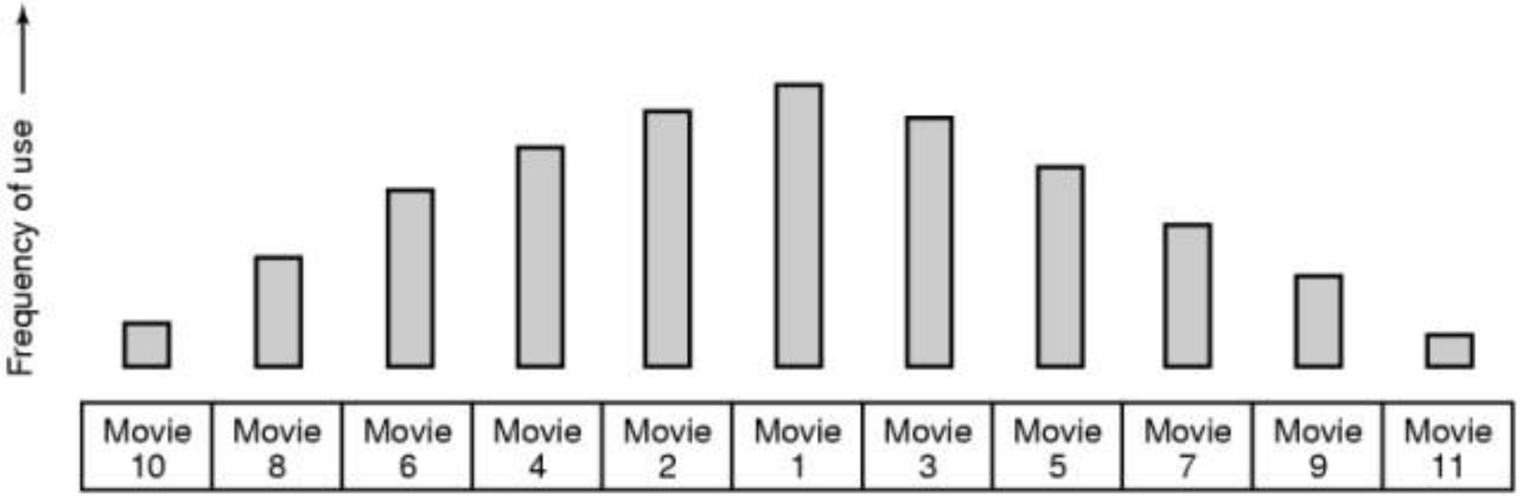
Zipf's law

- For movies on a video server
  - Zipf's law states that the most popular movie is chosen twice as often as the second most popular movie, three times as often as the third most popular movie, and so on
  - e.g., movie 50 has a popularity of C/50 and movie 51 has a popularity of C/51 , so movie 51 is 50/5 1as popular as movie 50, only about a 2% difference

Organ-Pipe algorithm

- Organ-Pipe algorithm
  - Grossman and Silverman (1973) and Wong (1983)
  - Studies have shown that the best strategy is surprisingly simple and distribution independent
  - placing the most popular movie in the middle of the disk, with the second and third most popular movies on either side of it

Organ-Pipe algorithm

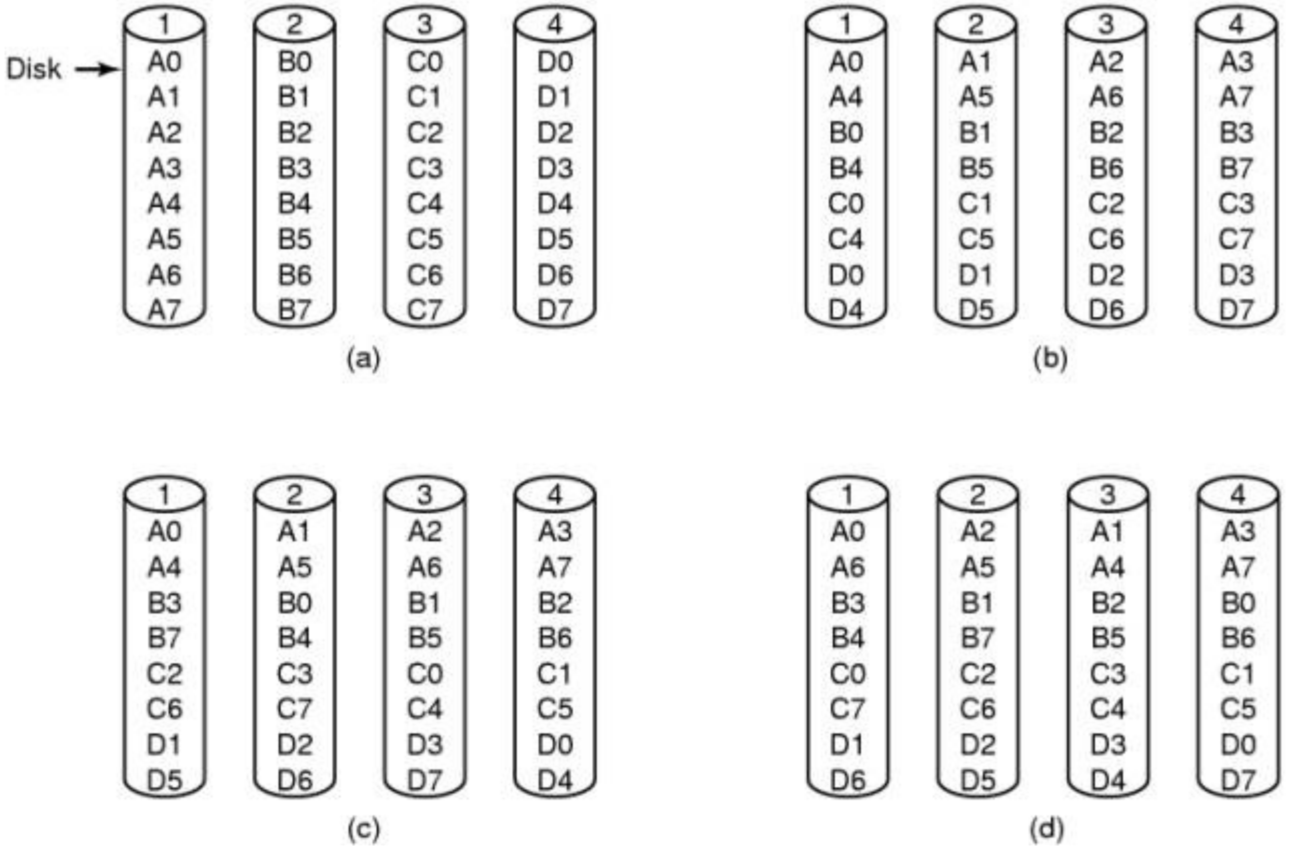


The organ-pipe distribution of files on a video server

Organ-Pipe algorithm

- Organ-Pipe algorithm
  - With 1000 movies and a Zipf’s law distribution
    - the top five movies represent a total probability of 0.307
    - the disk head will stay in the cylinders allocated to the top five movies about 30% of the time

Multiple Disks – Disk Farm



1	2	3	4
A0	B0	C0	D0
A1	B1	C1	D1
A2	B2	C2	D2
A3	B3	C3	D3
A4	B4	C4	D4
A5	B5	C5	D5
A6	B6	C6	D6
A7	B7	C7	D7

(a)

1	2	3	4
A0	A1	A2	A3
A4	A5	A6	A7
B0	B1	B2	B3
B4	B5	B6	B7
C0	C1	C2	C3
C4	C5	C6	C7
D0	D1	D2	D3
D4	D5	D6	D7

(b)

1	2	3	4
A0	A1	A2	A3
A4	A5	A6	A7
B3	B0	B1	B2
B7	B4	B5	B6
C2	C3	C0	C1
C6	C7	C4	C5
D1	D2	D3	D0
D5	D6	D7	D4

(c)

1	2	3	4
A0	A2	A1	A3
A6	A5	A4	A7
B3	B1	B2	B0
B4	B7	B5	B6
C0	C2	C3	C1
C7	C6	C4	C5
D1	D2	D3	D0
D6	D5	D4	D7

(d)

Four ways of organizing multimedia files over multiple disks. (a) No striping. (b) Same striping pattern for all files. (c) Staggered striping. (d) Random striping.

References

- Material
  - Slides
  - Video Lessons
- Books
  - Modern Operating Systems, A. S. Tanenbaum, Pearson, 4th edition , 2015,

Question 23

- Scheduling strategy for multimedia processes

- Question
    - Describe the scheduling of multimedia processes

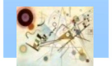
Homogeneous processes

- The simplest kind of video server
  - support the display of a fixed number of movies
    - same frame rate, video resolution, data rate, and other parameters
  - For each movie, there is a single process (or thread)
- NTSC 30 times per second
  - number of processes is small enough that all the work can be done in one frame time
    - round-robin scheduling
    - this model is rarely applicable in reality

Real-time scheduling

- Real applications
  - the number of users changes as viewers come and go
  - frame sizes vary wildly due to the nature of video compression
  - different movies may have different resolutions
  - multiple processes competing for the CPU

- Real time scheduling
  - the system knows the frequency at which each process must run
  - how much work it has to do
  - what its next deadline is

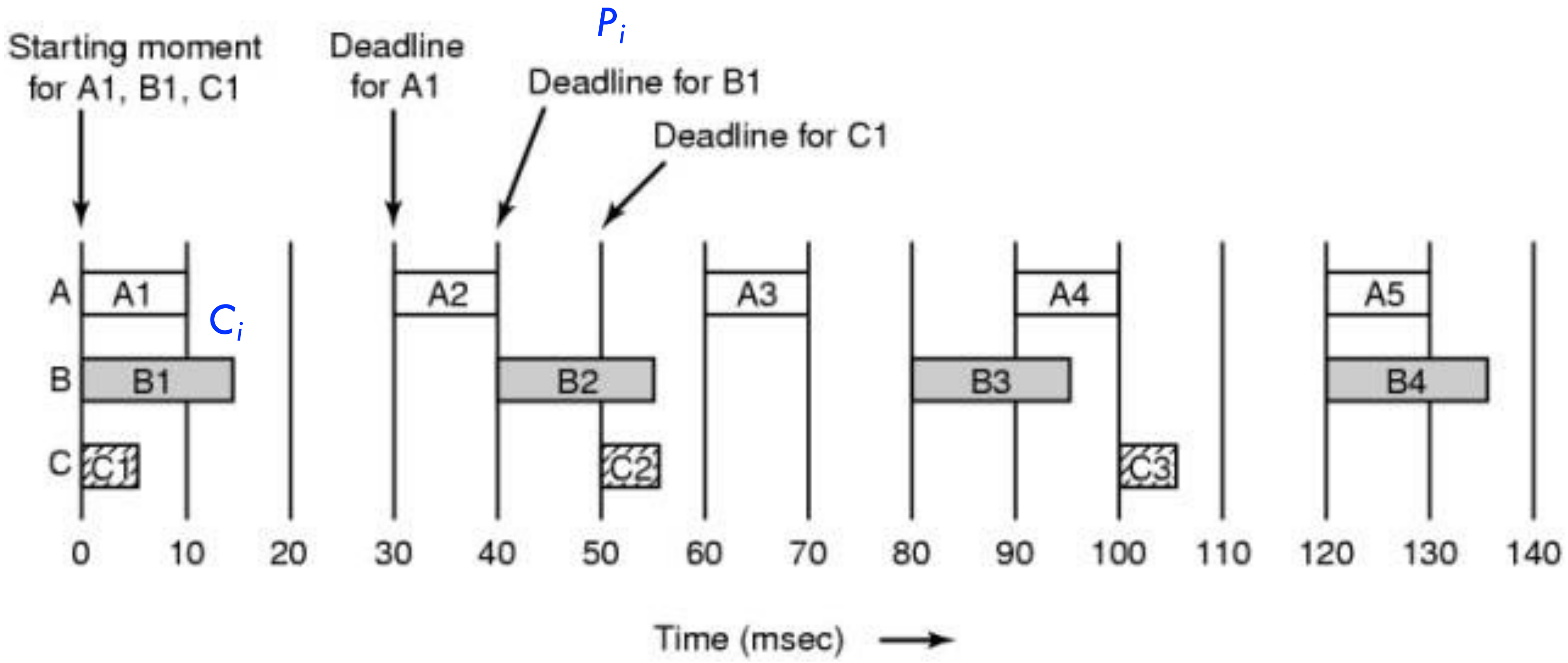


Real time-scheduling

Process	***Priodicity***	***CPU time***
A	33 Hz (NTSC)	10 ms
B	25 Hz (PAL)	15 ms
C	20 Hz (PAL slow connection)	5 ms

Example of processes

Real time-scheduling



Three periodic processes, each displaying a movie. The frame rates and processing requirements per frame are different for each movie.

Schedulable processes

- Schedulable condition
  - if process i has period P_i msec and requires C_i msec of CPU time per frame, the system is schedulable if and only if

$$\sum_{i=1}^{m} \frac{C_i}{P_i} \leq 1$$

Schedulable processes

Process	*Ci/Pi*
A	10/30
B	15/40
C	5/50

The system of processes is schedulable since the total is 0.808 of the CPU

Real-time algorithms

- Real-time algorithms can be
  - static
    - assign each process a fixed priority in advance and then do prioritized preemptive scheduling using those priorities
  - dynamic
    - does not have fixed priorities

Rate Monotonic Scheduling

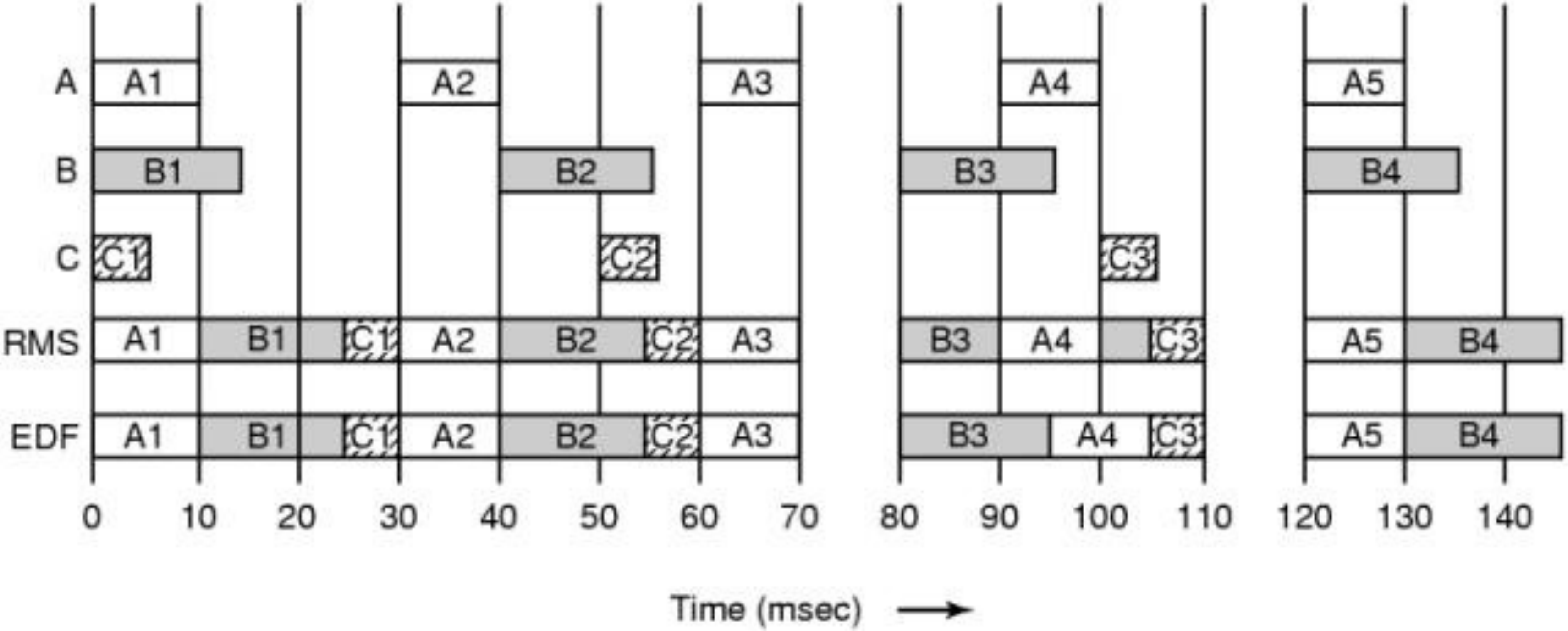
- Rate Monotonic Scheduling (RMS)
  - Liu and Layland, 1973
  - real-time scheduling algorithm for preemptable, periodic processes
- Conditions
  - each periodic process must complete within its period
  - no process is dependent on any other process
  - each process needs the same amount of CPU time on each burst
  - any nonperiodic processes have no deadlines
  - process preemption occurs instantaneously and with no overhead

Rate Monotonic Scheduling

- Rate Monotonic Scheduling
  - works by assigning each process a fixed priority equal to the frequency of occurrence of its triggering event
  - Liu and Layland proved that RMS is optimal among the class of static scheduling algorithms

Process	*Priority*
A	33
B	25
C	20

RMS



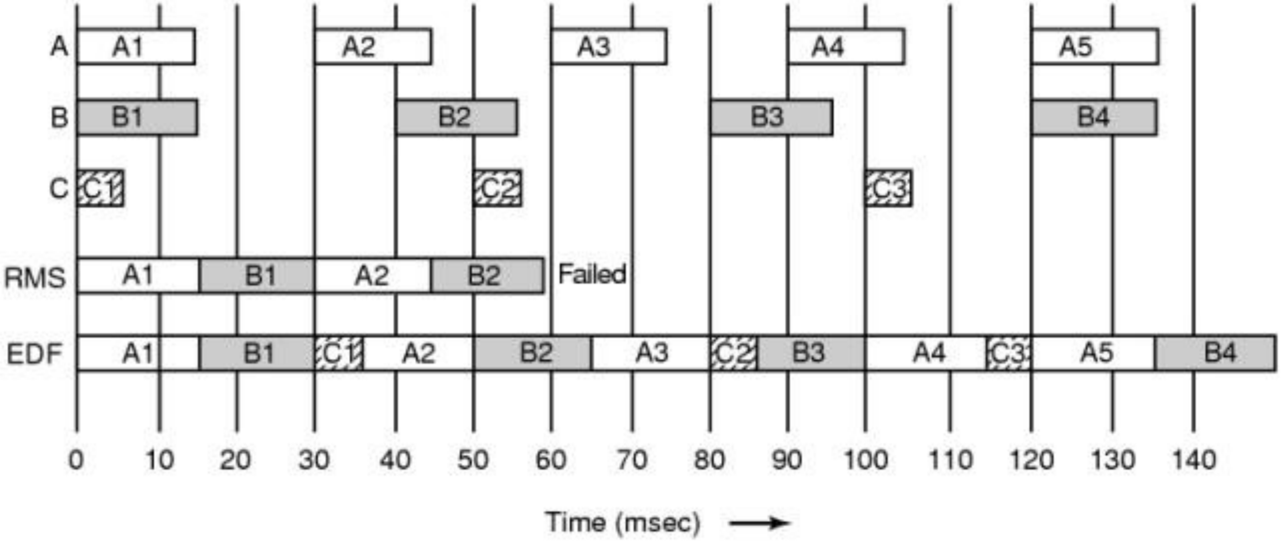
An example of RMS and EDF real-time scheduling

Earliest Deadline First Scheduling

- Earliest Deadline First Scheduling (EDF)
  - dynamic algorithm that does not require processes to be periodic

- Algorithm
  - a process needs CPU time, it announces its presence and its deadline
  - the scheduler keeps a list of runnable processes, *sorted on deadline*
  - the algorithm runs the first process on the list, the one with the closest deadline
  - whenever a new process becomes ready, the system checks to see if its deadline occurs before that of the currently running process
    - If so, the new process preempts the current one

RMS vs EDF



Example of RMS and EDF real-time scheduling (schedulable processes)

RMS

- Any system of periodic processes, if

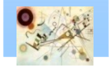
$$\sum_{i=1}^{m} \frac{C_i}{P_i} \leq m\left(2^{1/m} - 1\right)$$

then RMS is guaranteed to work

RMS

# of processes	**CPU utilization**
3	0.780
4	0.757
5	0.743
10	0.718
20	0.705
100	0.696
infinity	*ln 2*

CPU utilization by using RMS



EDF

- EDF
  - always works for any schedulable set of processes
  - it can achieve 100% CPU utilization
  - the price paid is a more complex algorithm

References

- Material
  - Slides
  - Video Lessons

- Books
  - Modern Operating Systems, A. S. Tanenbaum, Pearson, 4th edition , 2015,

Question 24

- Disk scheduling strategy for multimedia content

- Question
  - Describe the disk scheduling of multimedia processes

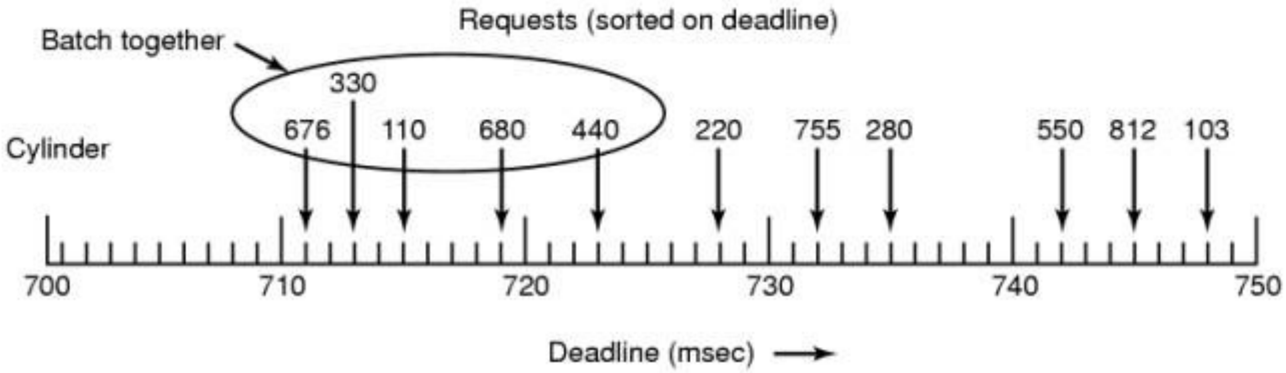
Dynamic disk scheduling

- different data rates
  - different movies may now have different data rates
  - it is not possible to have one round every 33.3 msec and fetch one frame for each stream

- selecting the next disk request
  - deadlines and cylinders
  - scan-EDF algorithm (Reddy and Wyllie, 1994)

Scan-EDF algorithm

- Basic idea
  - *collect requests whose deadlines are relatively close together into batches and process these in cylinder order*



The scan-EDF algorithm uses deadlines and cylinder numbers for scheduling

Scan-EDF algorithm

- When different streams have different data rates
  - *should the customer be admitted?*

- *If there is enough of each left (disk bandwidth, memory buffers, CPU time) for an average customer, the new one is admitted*

- *If the server has enough capacity for the specific film the new customer wants, then admission is granted; otherwise it is denied*

References

- Material
  - Slides
  - Video Lessons

- Books
  - Modern Operating Systems, A. S. Tanenbaum, Pearson, 4th edition , 2015,

Question 25

- Multimedia Networking

- Question
  - Describe the packet loss in VoIP

Introduction

- Internet telephony
  - commonly called Voice-over-IP (VoIP)

- Limitations
  - IP provides best-effort service
    - delay bound
    - percentage of packets lost

- Scenario
  - sender generates bytes at a rate of 8,000 bytes per second
  - every 20 msecs the sender gathers these bytes into a chunk
  - a chunk and a special header are encapsulated in a UDP segment, via a call to the socket interface
  - the number of bytes in a chunk is 20 msecs
  - a *UDP segment is sent every 20 msecs*
  - the receiver can simply play back each chunk as soon as it arrives

Packet loss

- UDP segment
  - is encapsulated in an IP datagram
  - datagram wanders through the network
  - it passes through router buffers
  - It is possible that one or more of the buffers in the path from sender to receiver is full
    - IP datagram may be discarded
- Loss elimination
  - sending the packets over TCP
  - unacceptable for conversational real-time audio applications such as VoIP
- UDP is used by Skype unless a user is behind a NAT or firewall that blocks UDP segments (in which case TCP is used)

Packet loss

- Packet loss ratios
  - between 1 and 20 percent can be tolerated
    - packet loss concealment
  - packet loss exceeds 10 to 20 percent (for example, on a wireless link
    - no acceptable audio quality

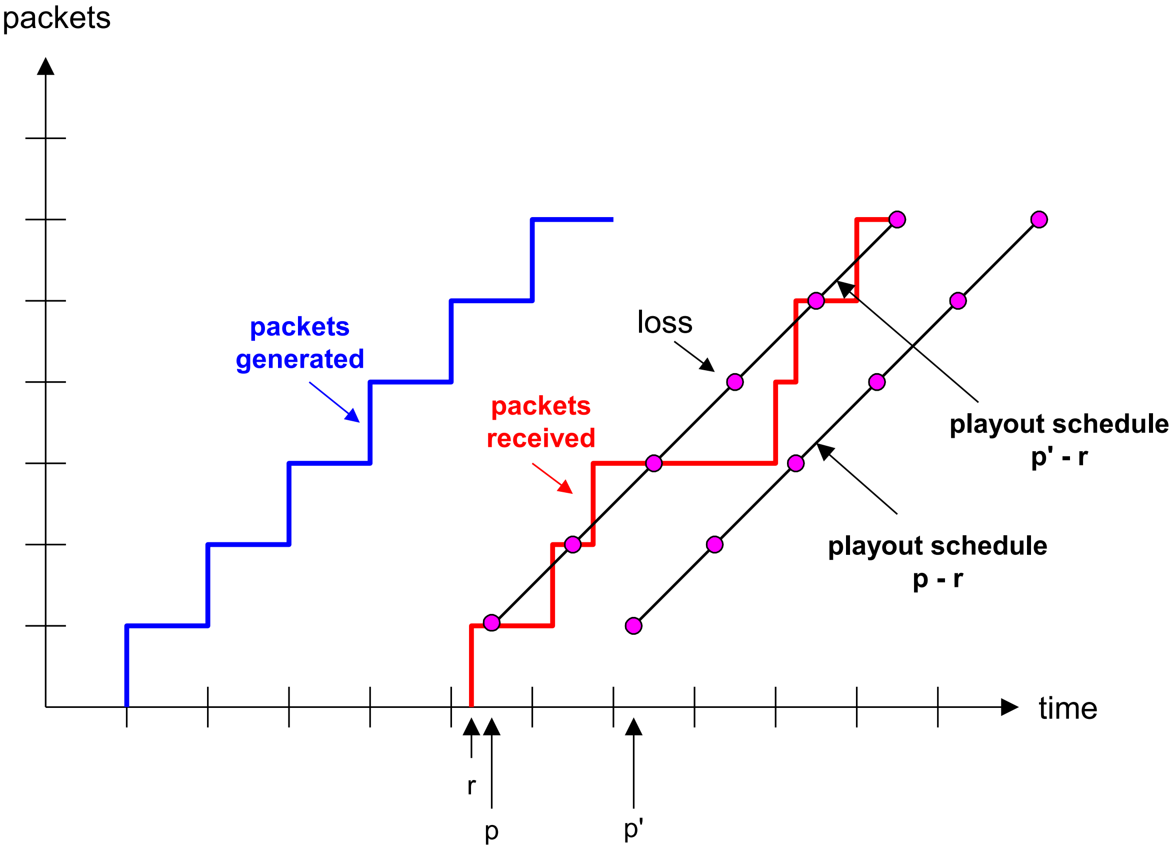
End to end delay

- End-to-end delay
  - accumulation of transmission, processing, and queuing delays
  - VoIP
    - end-to-end delays smaller than 150 msecs are not perceived by a human listener
    - delays between 150 and 400 msecs can be acceptable but are not ideal
    - delays exceeding 400 msecs can seriously hinder the interactivity in voice conversations

Packet jitter

- Packet jitter
  - varying delays
    - the time from when a packet is generated at the source until it is received at the receiver can fluctuate from packet to packet
  - e.g., different queues for different rooters

Removing jitter



Fixed playout delay

Adaptive Playout Delay

t_i = the timestamp of the ith packet = the time the packet was generated by the sender

r_i = the time packet i is received by receiver

p_i = the time packet i is played at receiver

$$d_i = (1 - u)\, d_{i-1} + u\, (r_i - t_i)$$

Estimate of the average network delay

$$v_i = (1 - u)\, v_{i-1} + u \mid r_i - t_i - d_i \mid$$

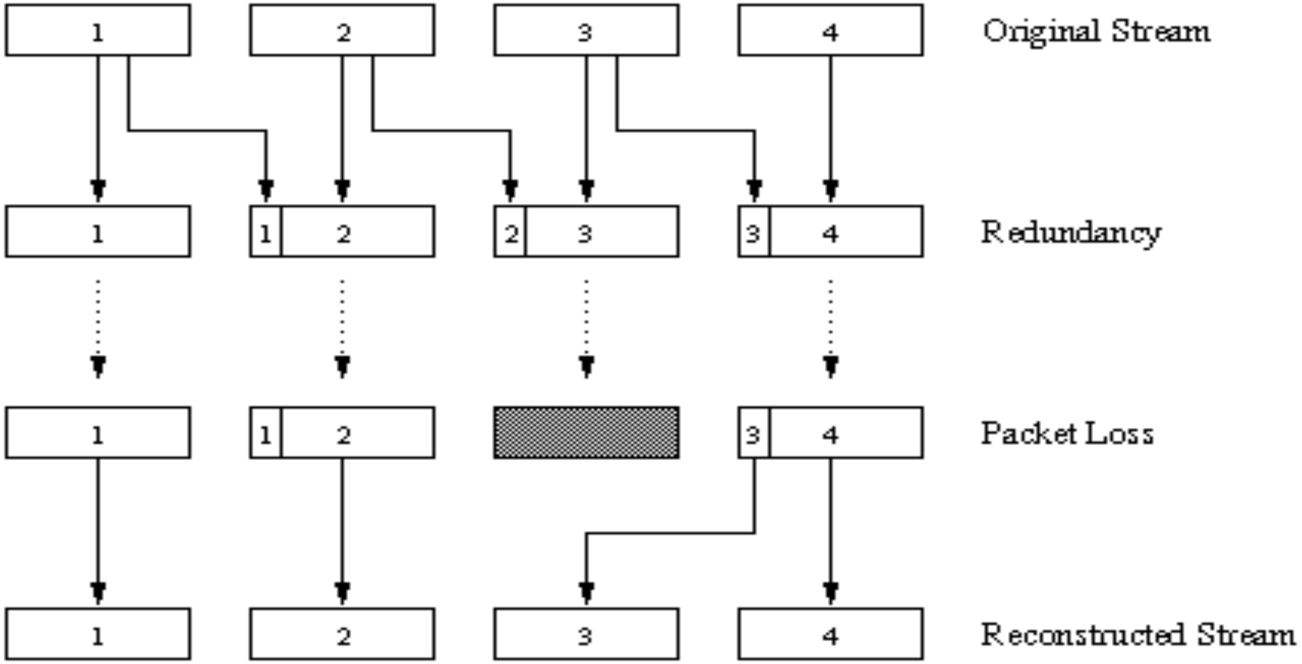
Estimate of the average deviation of the network delay

costant 4

$$p_i = t_i + d_i + Kv_i$$

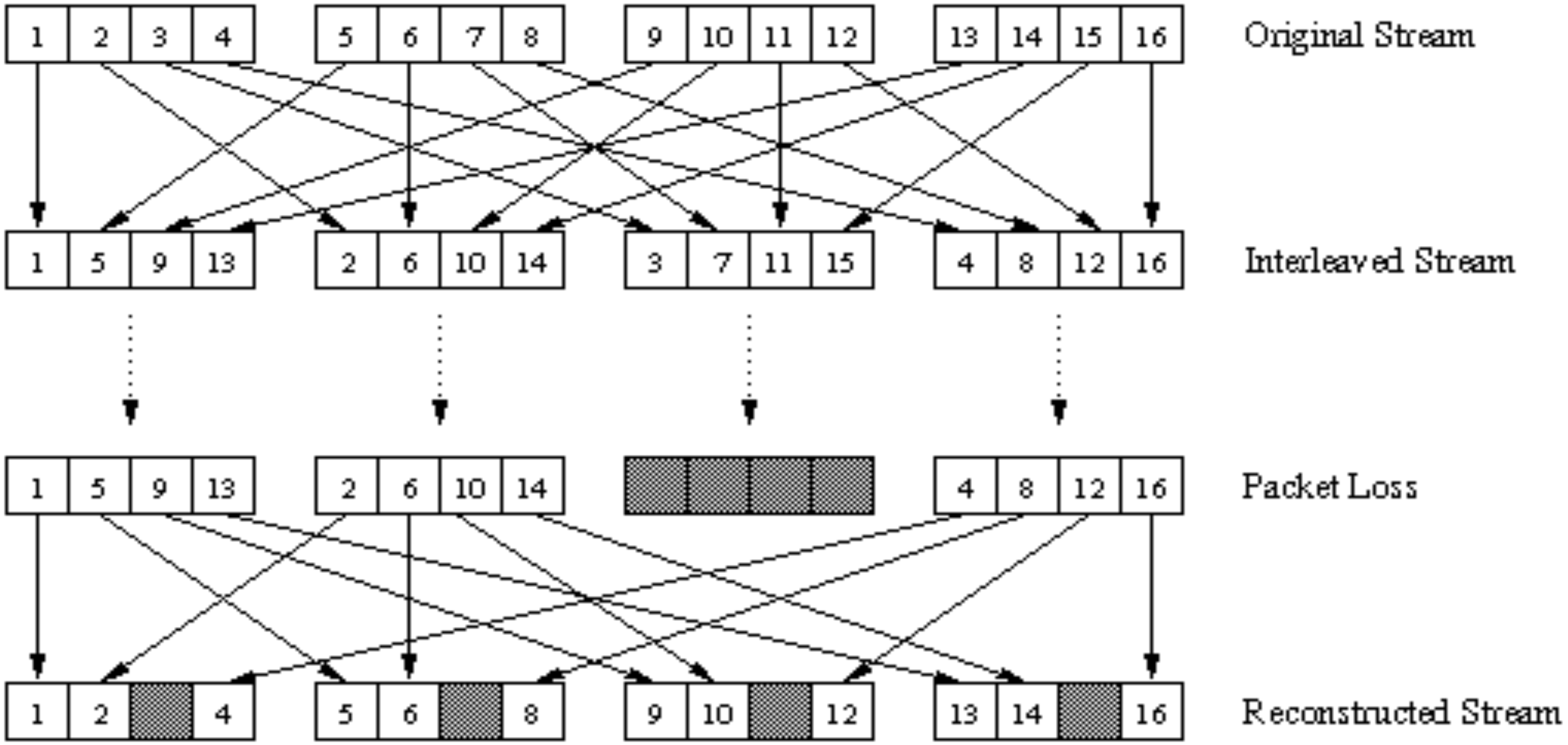
Playout of the packets

Forward Error Correction



FEC - FEC mechanism is to send a lower-resolution audio stream as the redundant information

Interleaving



Sending interleaved audio

Error Concealment

- Error concealment schemes
  - attempt to produce a replacement for a lost packet that is similar to the original
  - the simplest form of receiver-based recovery is packet repetition
  - methodology based on compressive sensing

References

- Material
  - Slides
  - Video Lessons

- Books
  - Computer Networking: A Top-Down Approach, J. F. Kurose, K. W. Ross, Pearson, 6 edition, 2013

Question 26

- Multimedia Networking

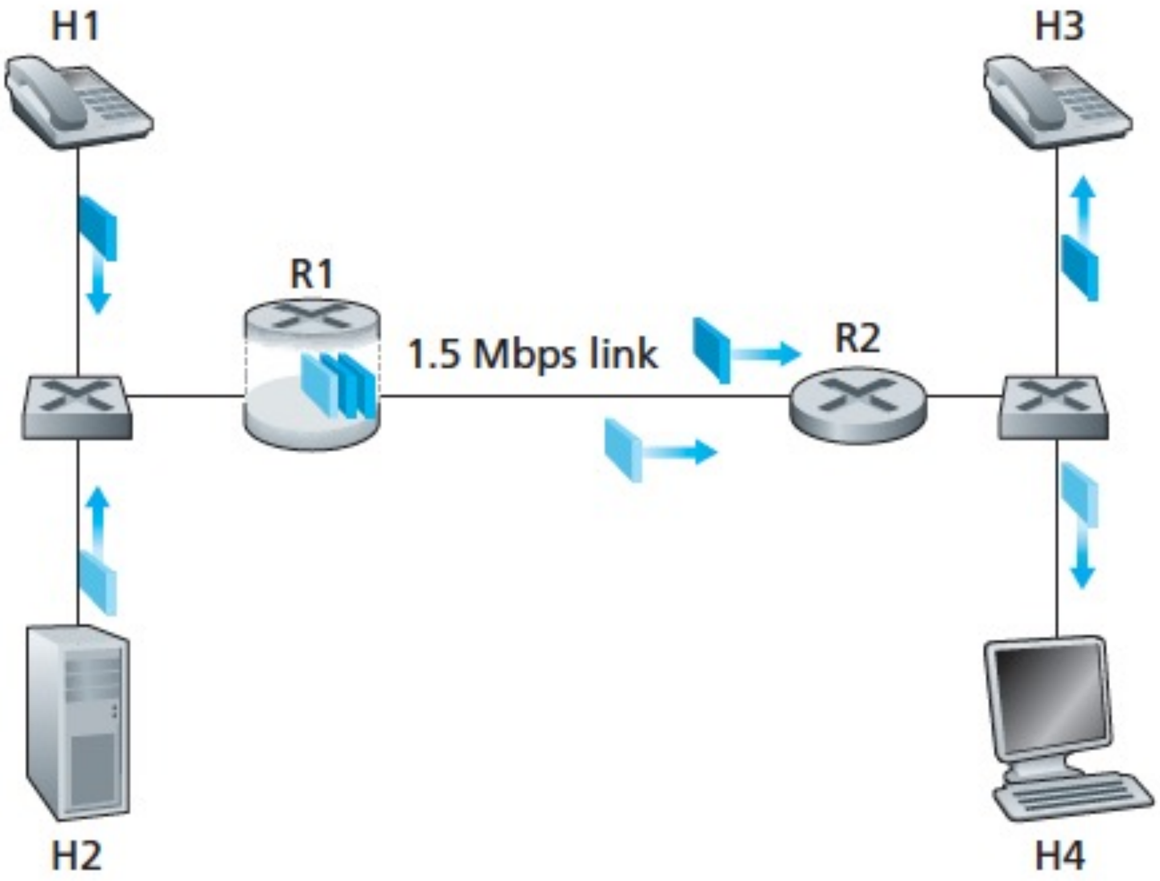
- Question
    - Describe the QoS

Introduction

- simplest enhancement to the one-size-fits-all best-effort service
    - divide traffic into classes

- parameters are to be used to guide the selection of the actual service parameters when transmitting a datagram through a particular network
    - Type-of-Service (ToS) field in the IPv4 header

Networking scenario



Competing audio and http applications

Multiple Classes of Service

Insight 1

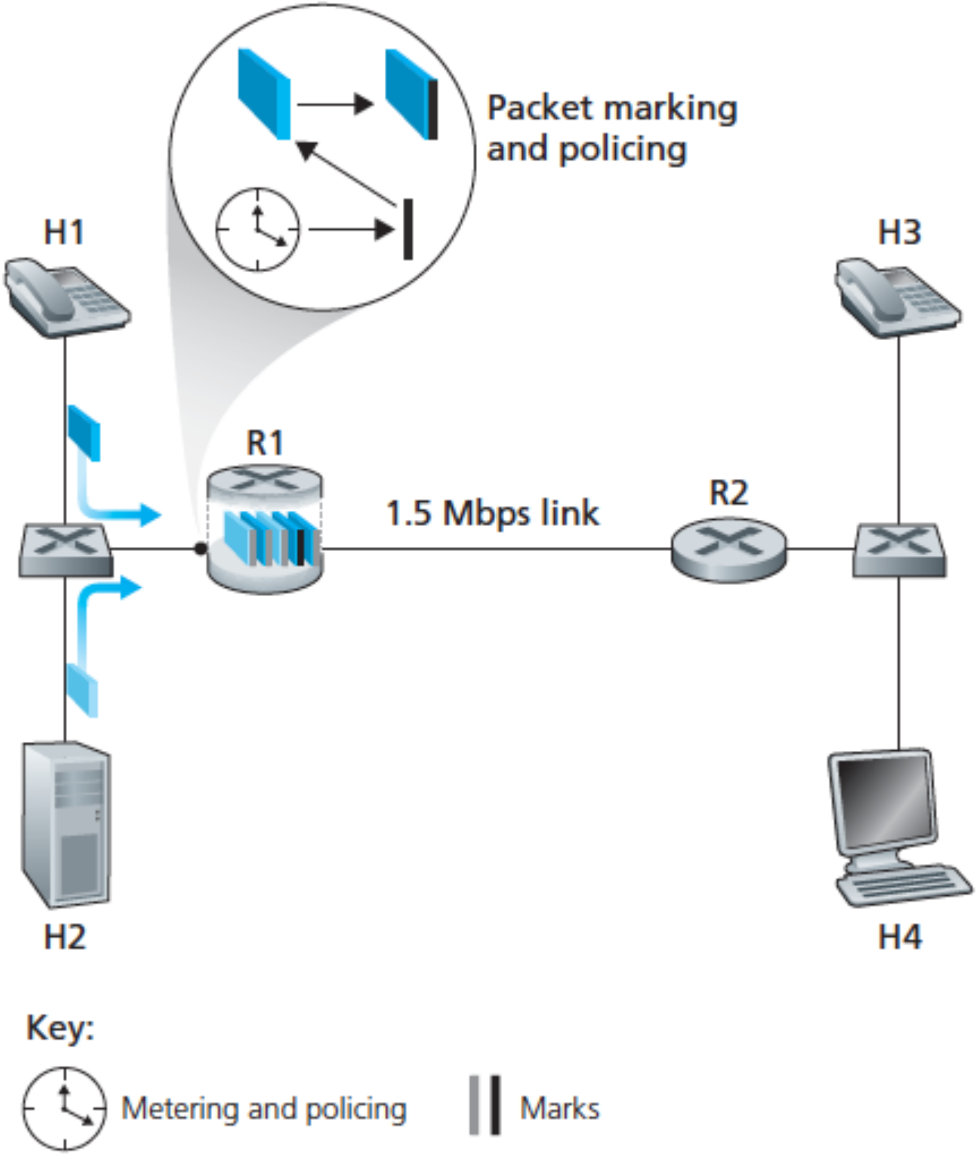
Packet marking allows a router to distinguish among packets belonging to different classes of traffic

Insight 2

It is desirable to provide a degree of traffic isolation among classes so that one class is not adversely affected by another class of traffic that misbehaves

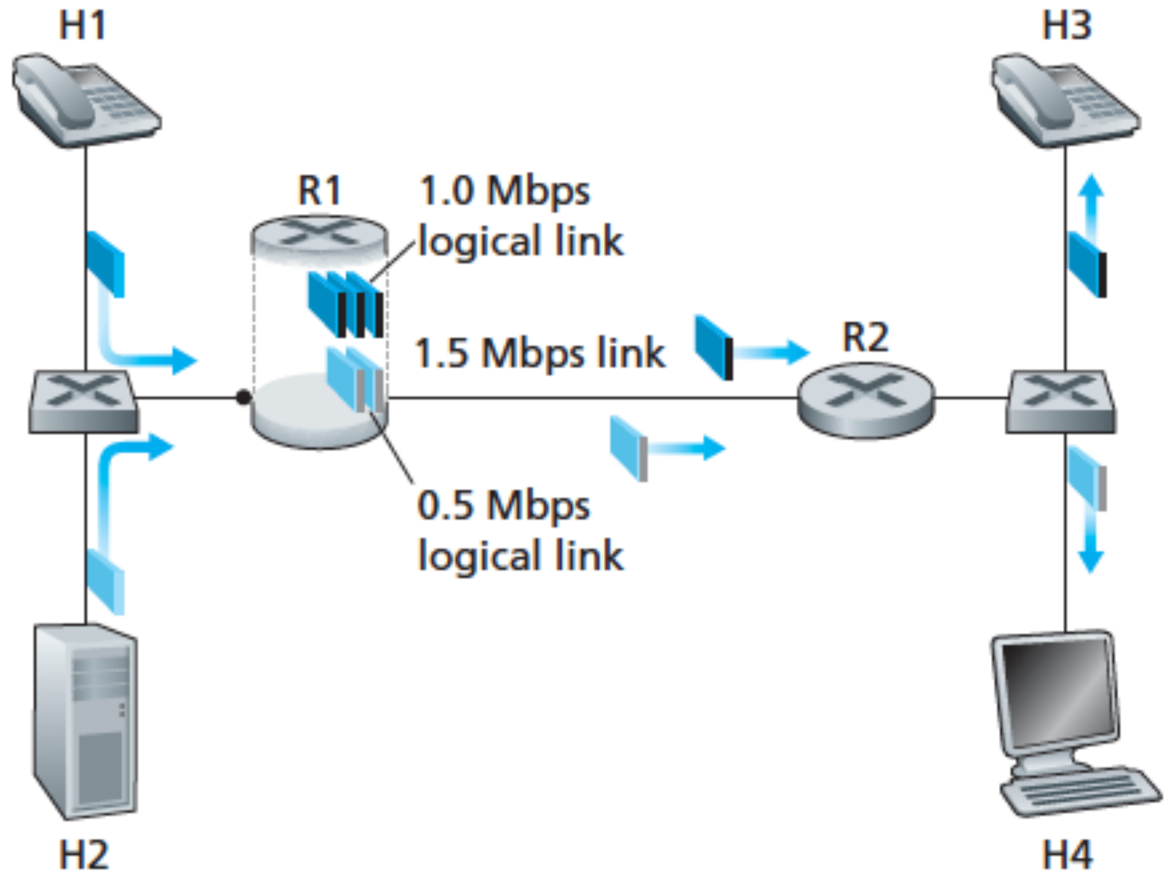
Audio 1.5 Mbps or higher (policing)

Networking scenario



Marking and Policing the audio and http traffic classes

Networking scenario



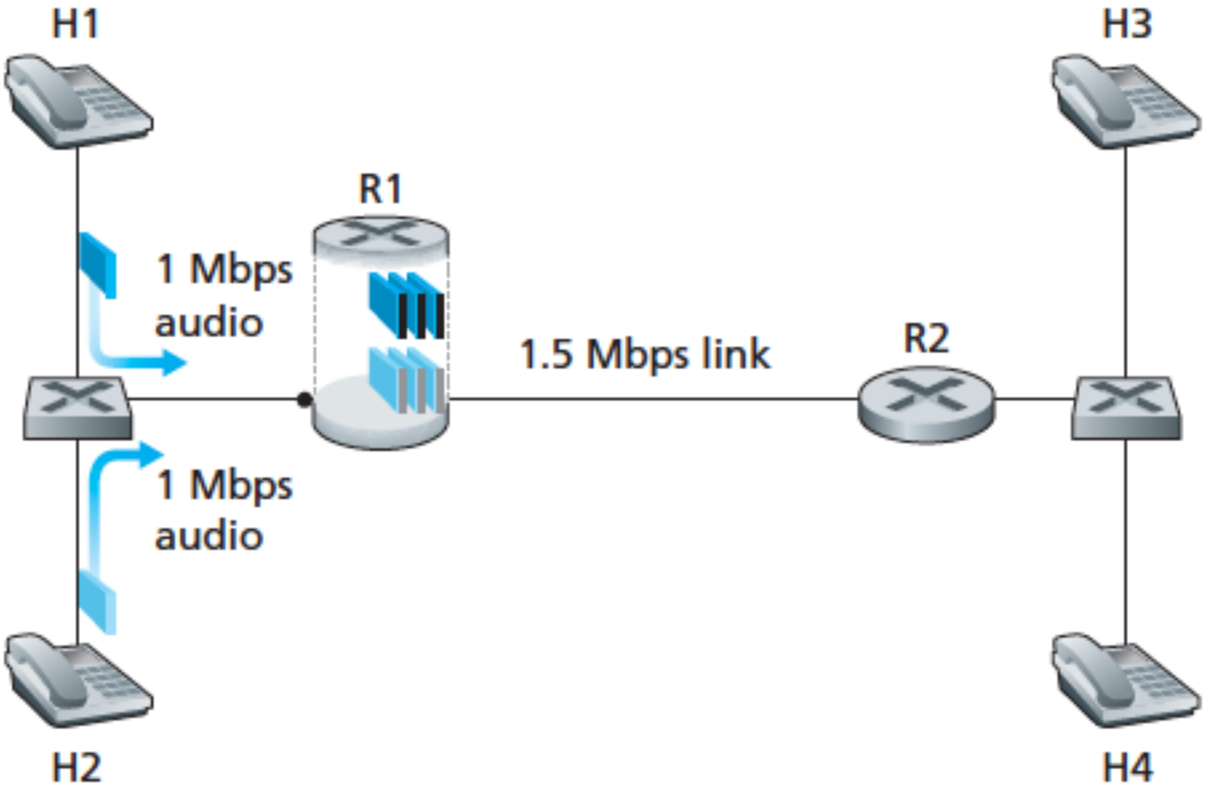
Logical isolation of audio and http traffic classes

Multiple Classes of Service

Insight 3

While providing isolation among classes or flows, it is desirable to use resources (for example, link bandwidth and buffers) as efficiently as possible

Multiple Classes of Service



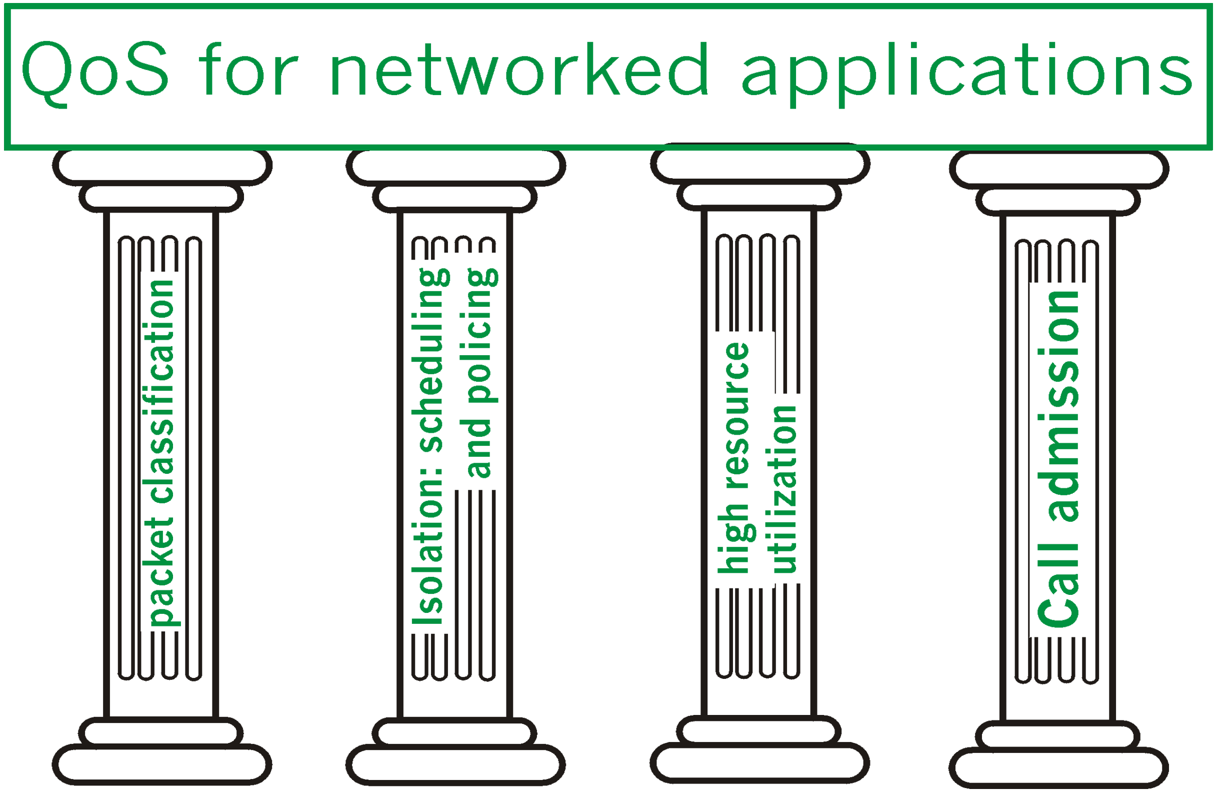
Two competing audio applications overloading the R1-to-R2 link

Multiple Classes of Service

Insight 4

If sufficient resources will not always be available, and QoS is to be guaranteed , a call admission process is needed in which flows declare their QoS requirements and are then either admitted to the network (at the required QoS) or blocked from the network (if the required QoS cannot be provided by the network)

Quality of Service



QoS for networked applications

packet classification

Isolation: scheduling and policing

high resource utilization

Call admission

4 pillars of QoS

ISP – Verification tests

References

- Material
  - Slides
  - Video Lessons

- Books
  - Computer Networking: A Top-Down Approach, J. F. Kurose, K. W. Ross, Pearson, 6 edition, 2013

Question 27

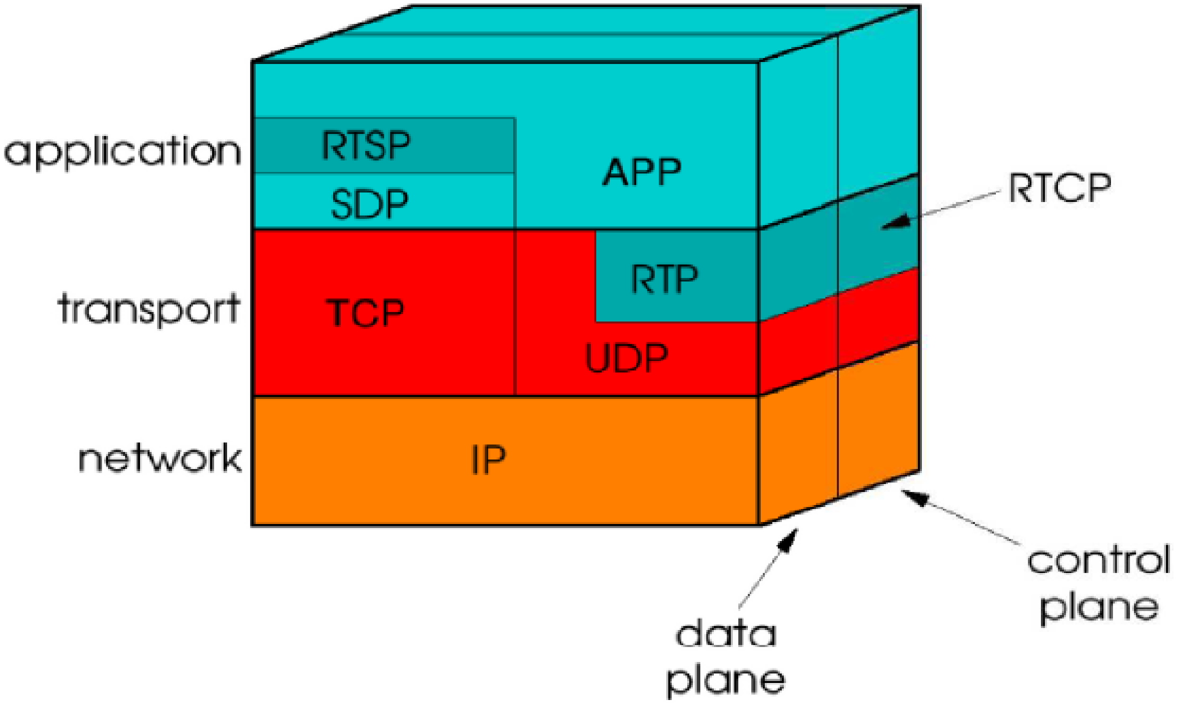
- Multimedia Networking

- Question
  - Describe the RTP

Real Time Protocol

- Real Time Protocol
  - defined in RFC 3550
  - used for transporting common formats
    - such as PCM, ACC, and MP3 for sound and MPEG and H.263 for video

Real Time Protocol



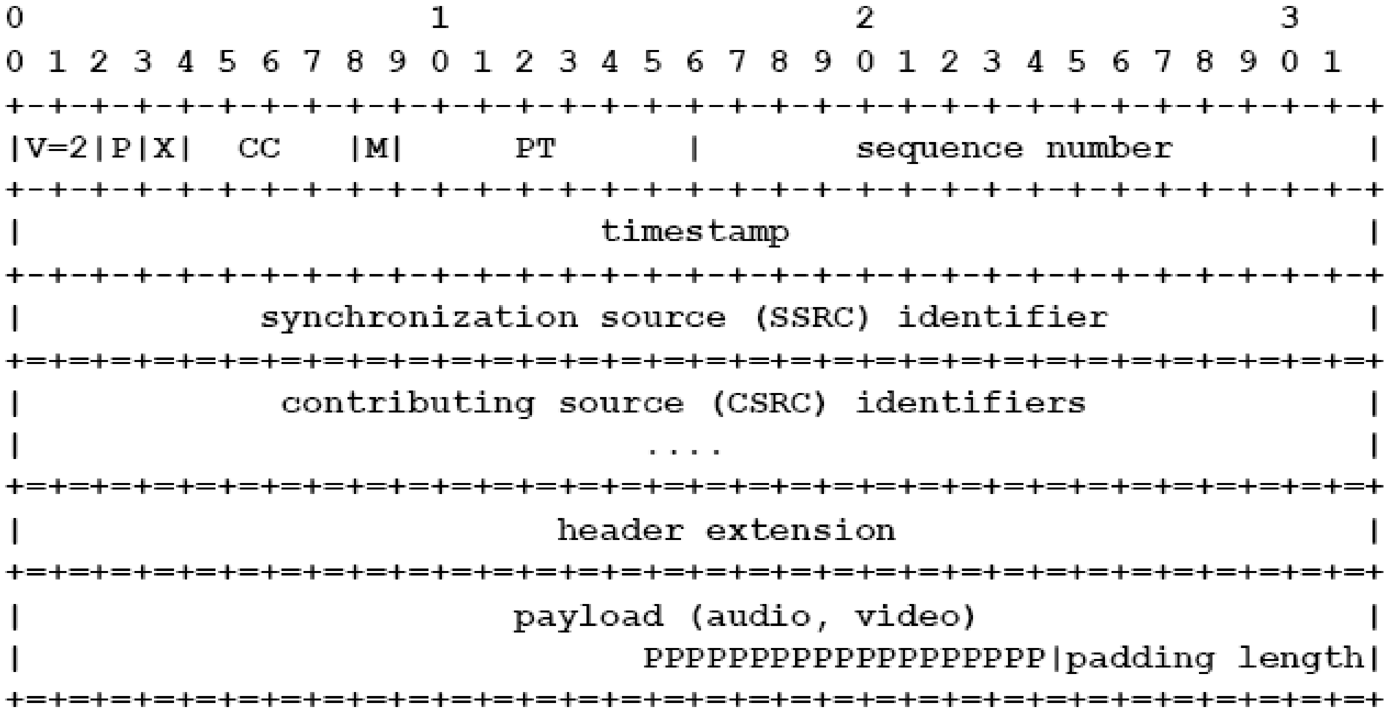
Levels of the communication protocol

Real Time Protocol

RTP Header

Code	Format	Sampling frequency	Frequency
0	PCM legge μ	8 KHz	64 Kbps
1	1016	8 KHz	4,8 Kbps
3	GSM	8 KHz	13 Kbps
7	LPC	8 KHz	2,4 Kbps
9	G.722	16 KHz	48-64 Kbps
14	Audio MPEG	90 KHz	-
15	G.728	8 KHz	16 Kbps

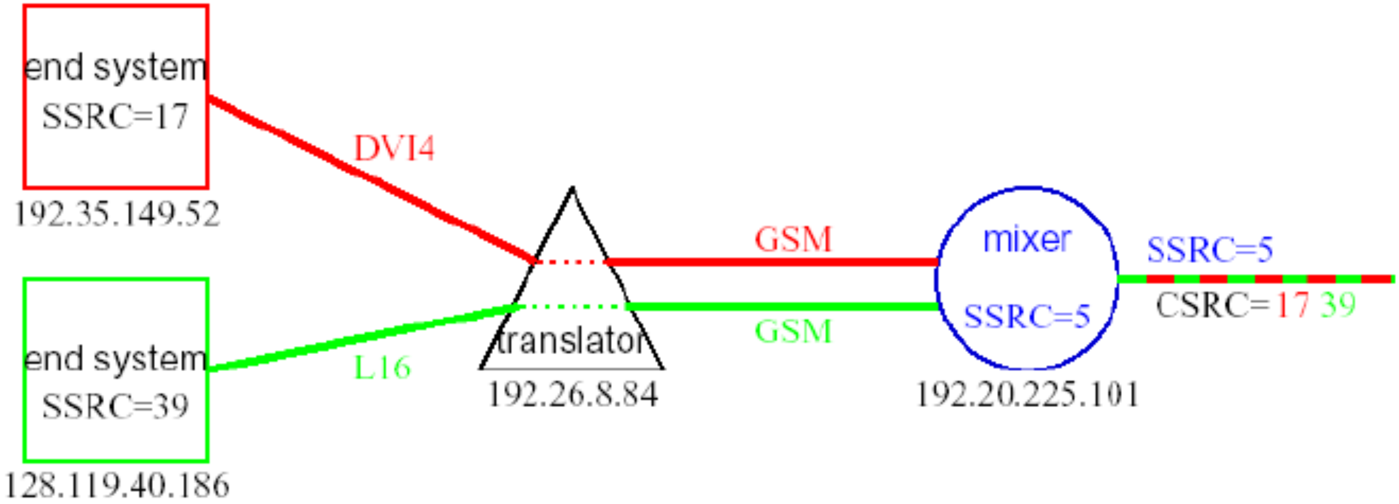
Real Time Protocol



```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|X|  CC   |M|     PT      |       sequence number         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           timestamp                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           synchronization source (SSRC) identifier            |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|            contributing source (CSRC) identifiers             |
|                             ....                              |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|                         header extension                      |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|                       payload (audio, video)                  |
|                             PPPPPPPPPPPPPPPPPPP|padding length|
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
```

RFC packet definition

Real Time Protocol



CSRC = Contributing Source

Mixer and traslator

Real Time Protocol

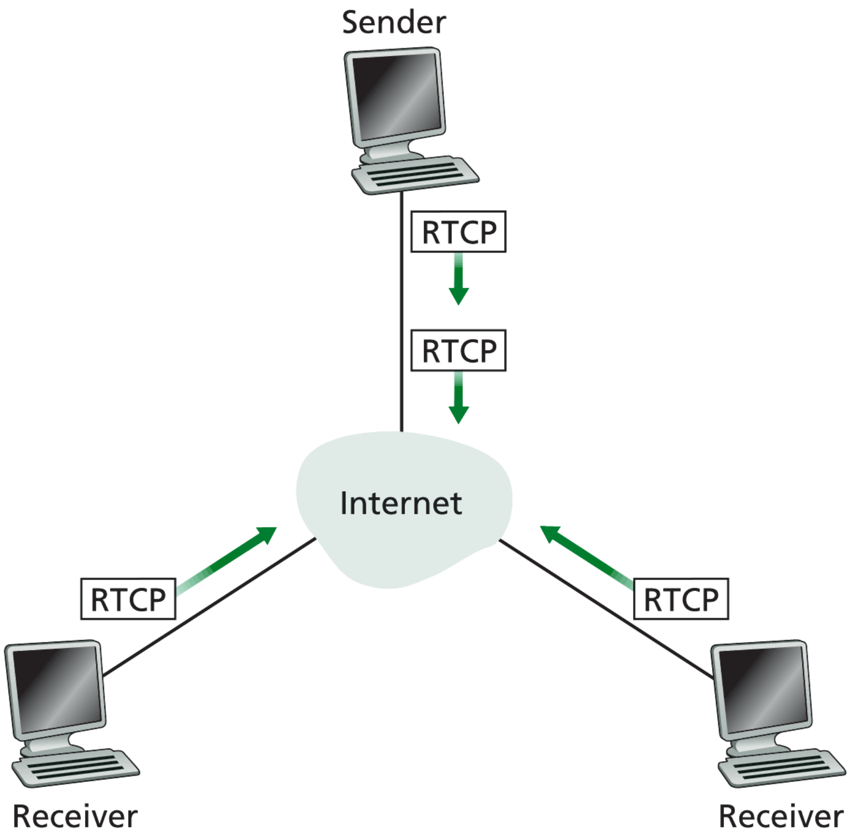


Figure 7.12 ◆ Both senders and receivers send RTCP messages.

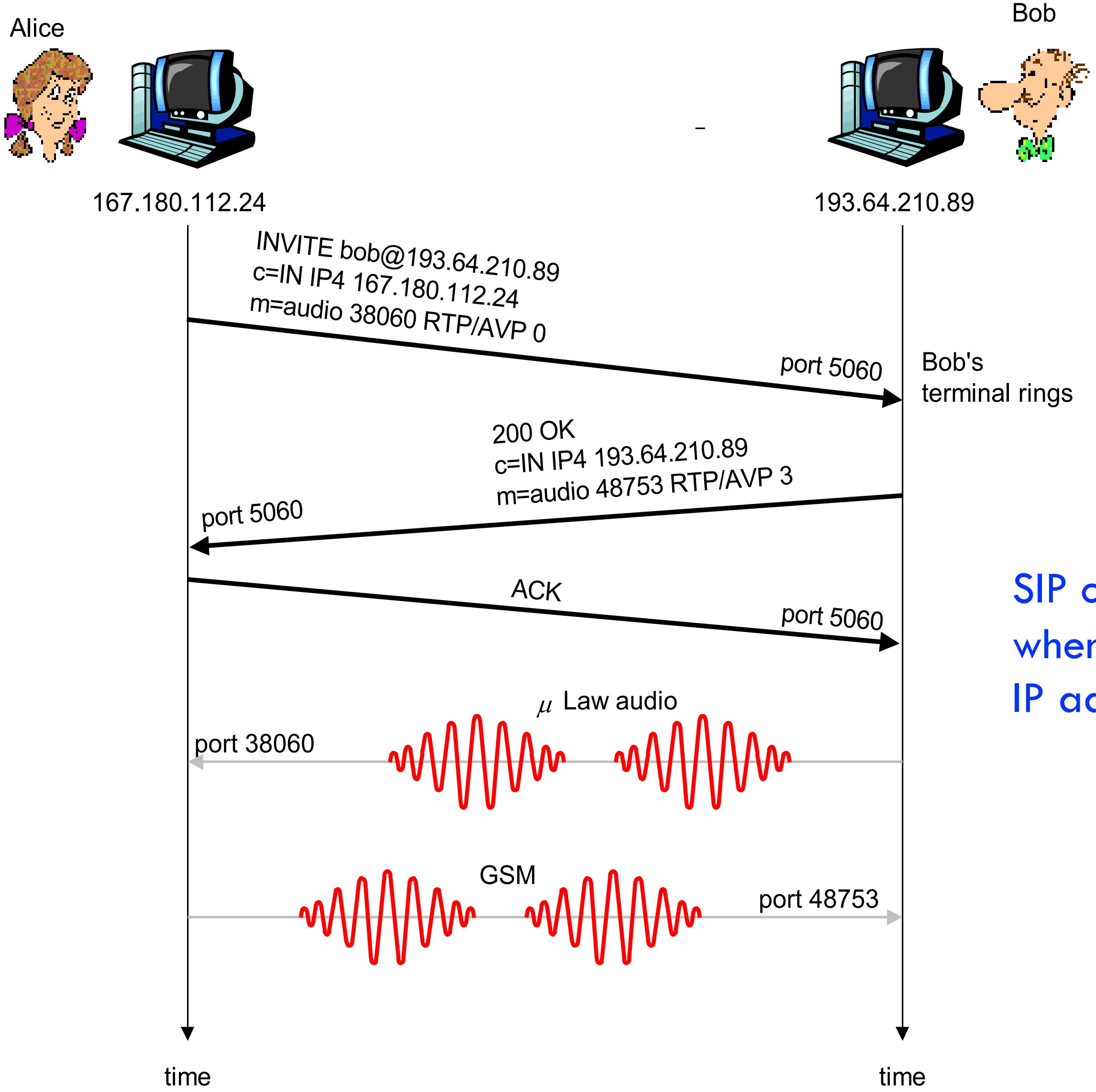
Real Time Control Protocol

- Streaming information
  - SSRC identifier
  - temporal stamp
  - number of sent packets
  - stream bytes

SIP

- Session Initiation Protocol
  - RFC 3261
  - provides mechanisms for establishing calls between a caller and a callee over an IP network
  - It provides mechanisms for the caller to determine the current IP address of the callee
  - It provides mechanisms for call management, such as adding new media streams
  - during the call, changing the encoding during the call, inviting new participants during the call, call transfer, and call holding

SIP



SIP call establishment when Alice knows Bob's IP address

References

- Material
  - Slides
  - Video Lessons

- Books
  - Computer Networking: A Top-Down Approach, J. F. Kurose, K. W. Ross, Pearson, 6 edition, 2013

ISP – Verification tests

Question 28

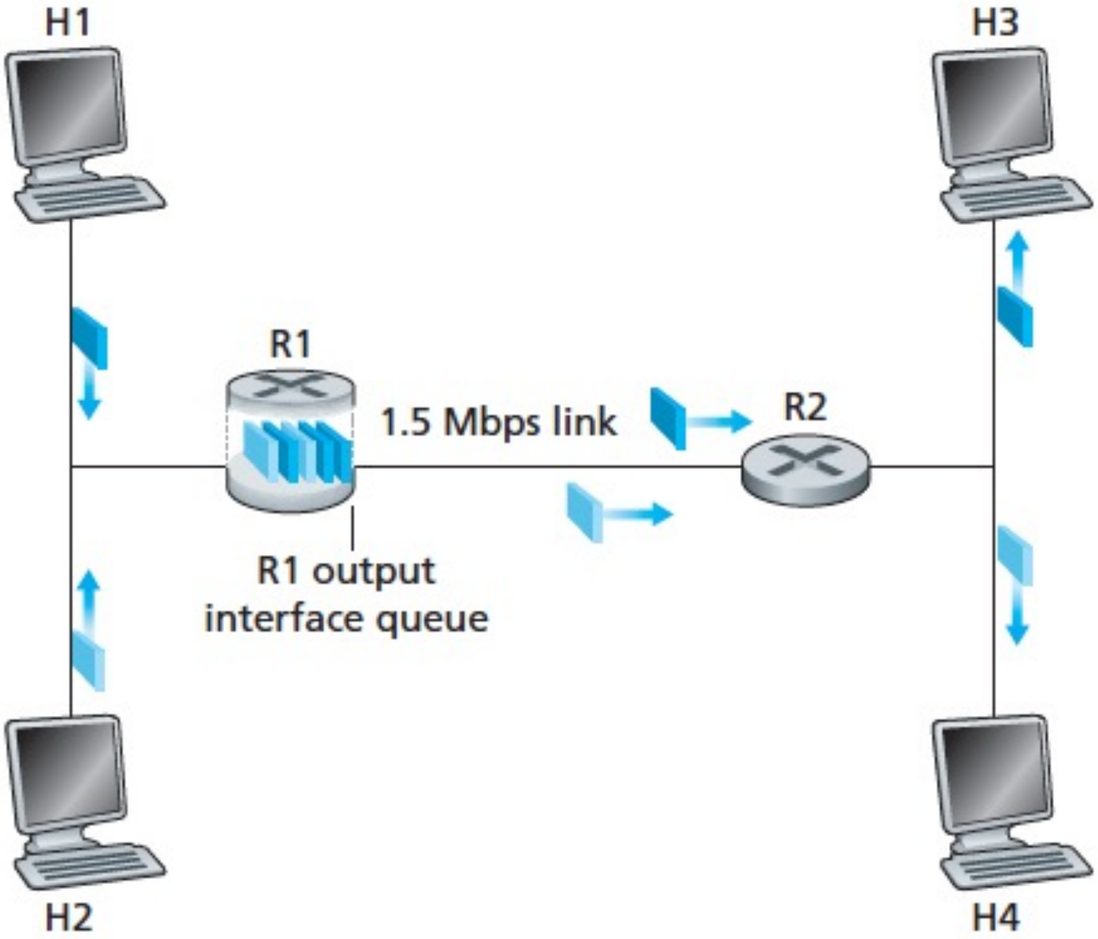
- Multimedia Networking

- Question
  - Describe the scheduling and policing mechanisms

Scheduling mechanisms

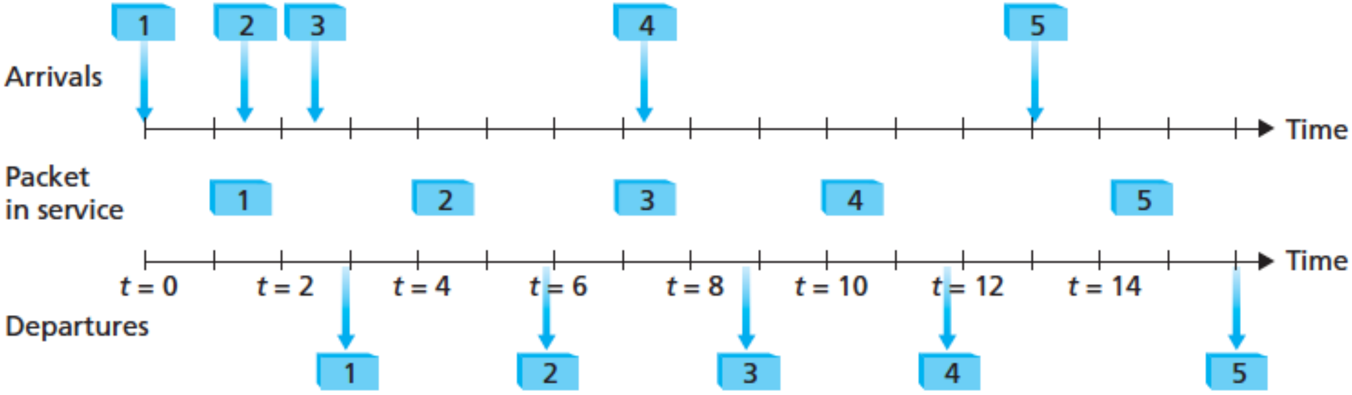
- packets belonging to various network flows are multiplexed and queued for transmission
  - *link-scheduling discipline*
  - *packet-discarding policy*
    - *determines whether the packet will be dropped (lost) or whether other packets will be removed from the queue to make space for the arriving packet*

FIFO



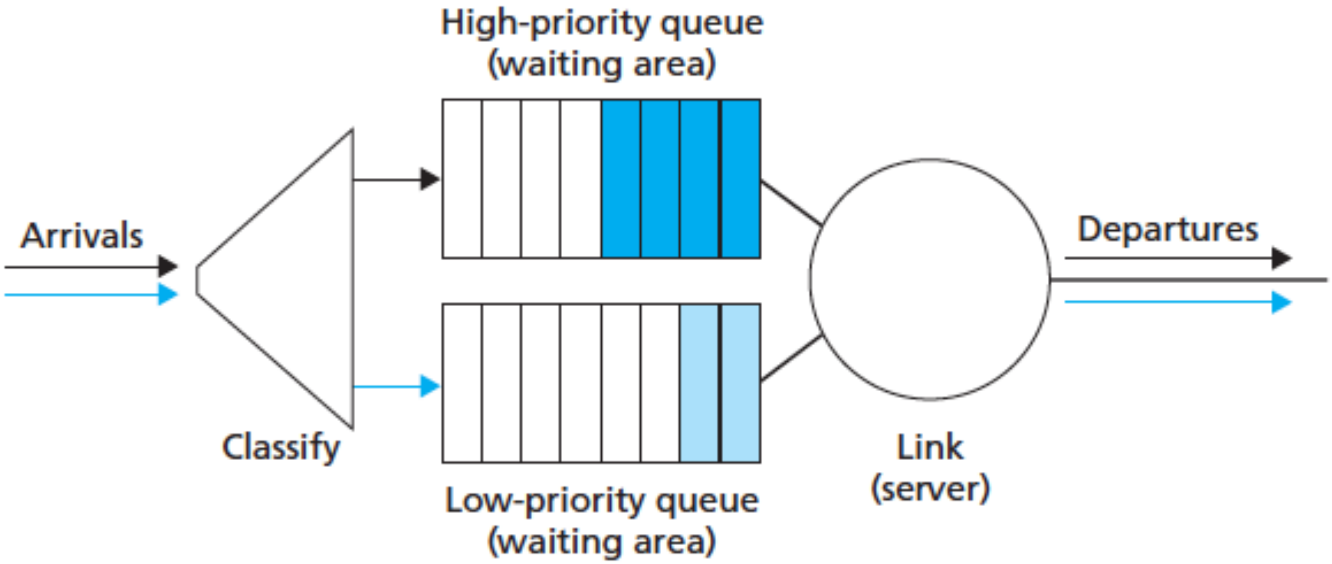
FIFO queuing abstraction

FIFO



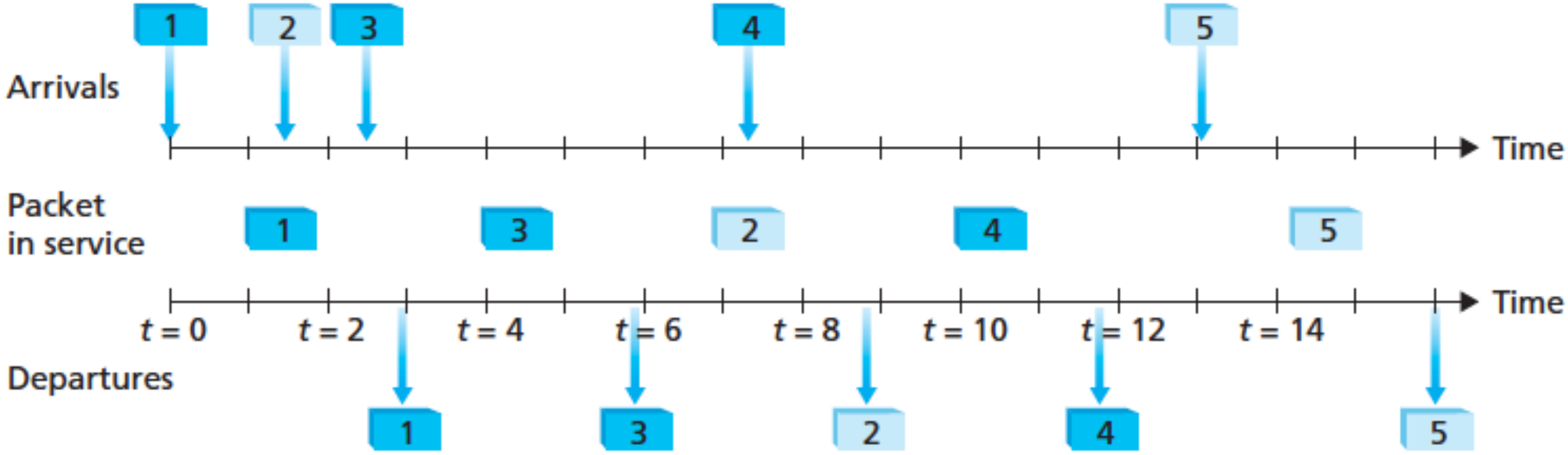
The FIFO queue in operation

Priority queuing model



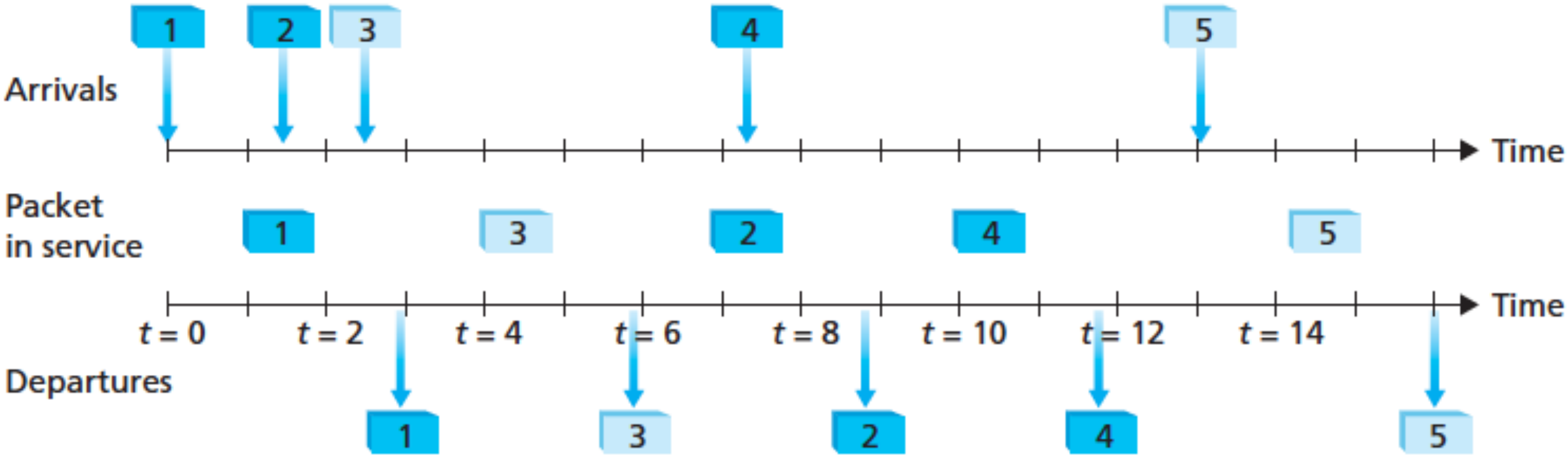
Priority queuing model

Priority queuing model



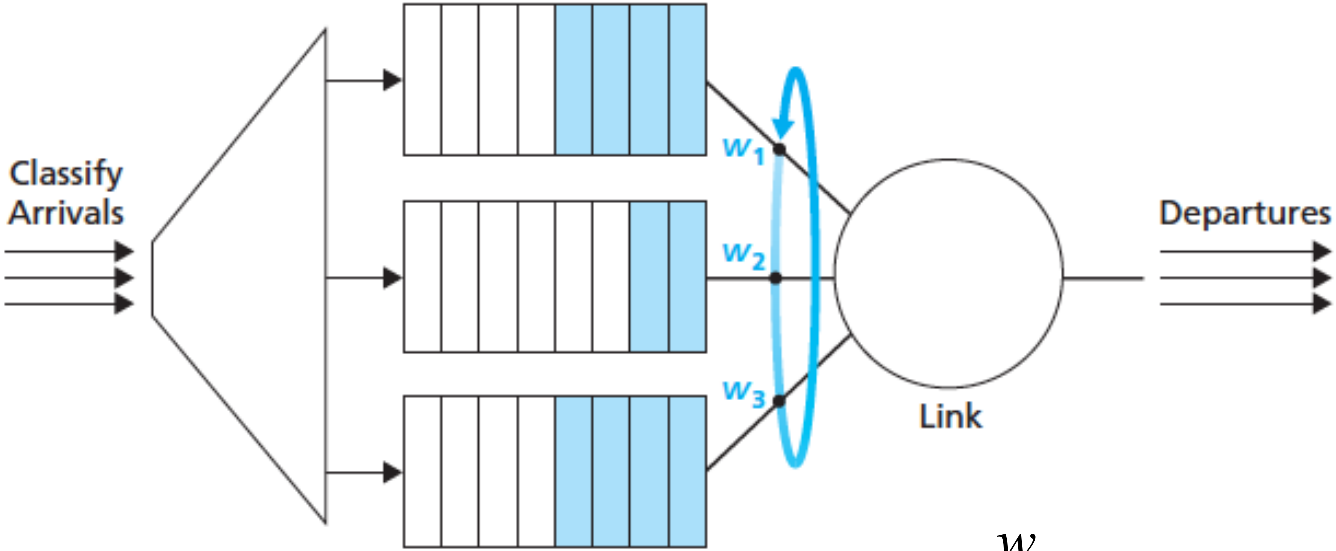
Operation of the priority queue

Round-Robin model



Operation of two-class round robin queue

Weighted Fair Queuing model

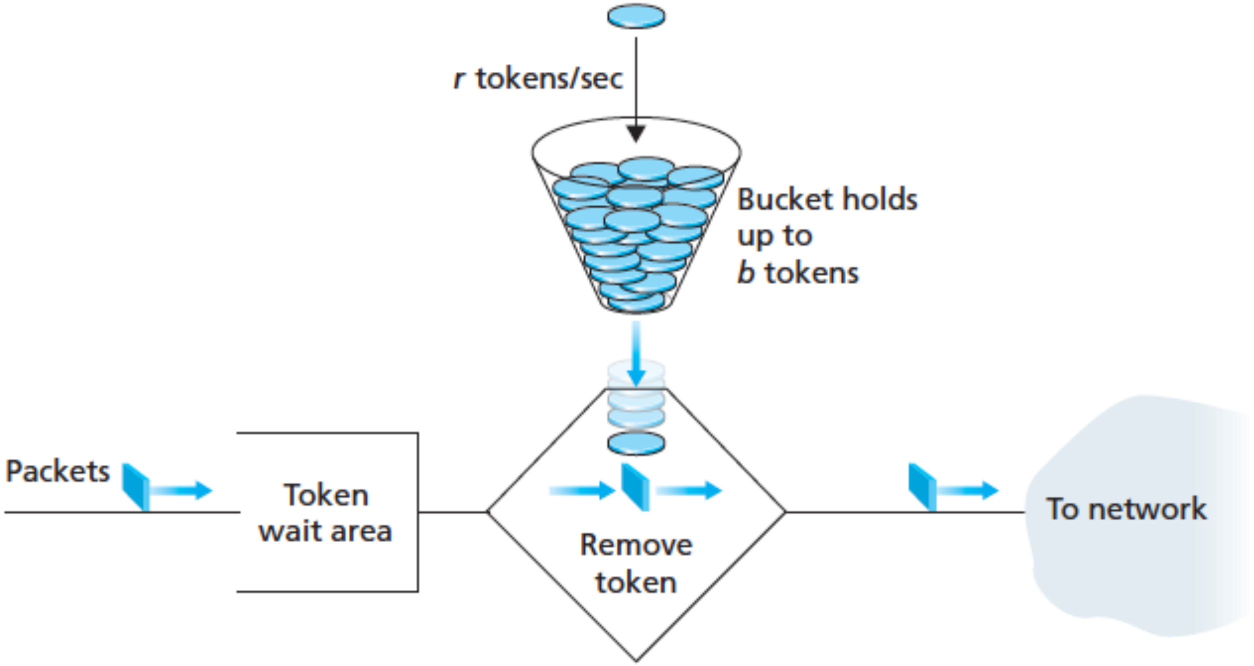


$$R\frac{w_i}{\sum w_j}$$

Weighted fair queuing (WFQ)

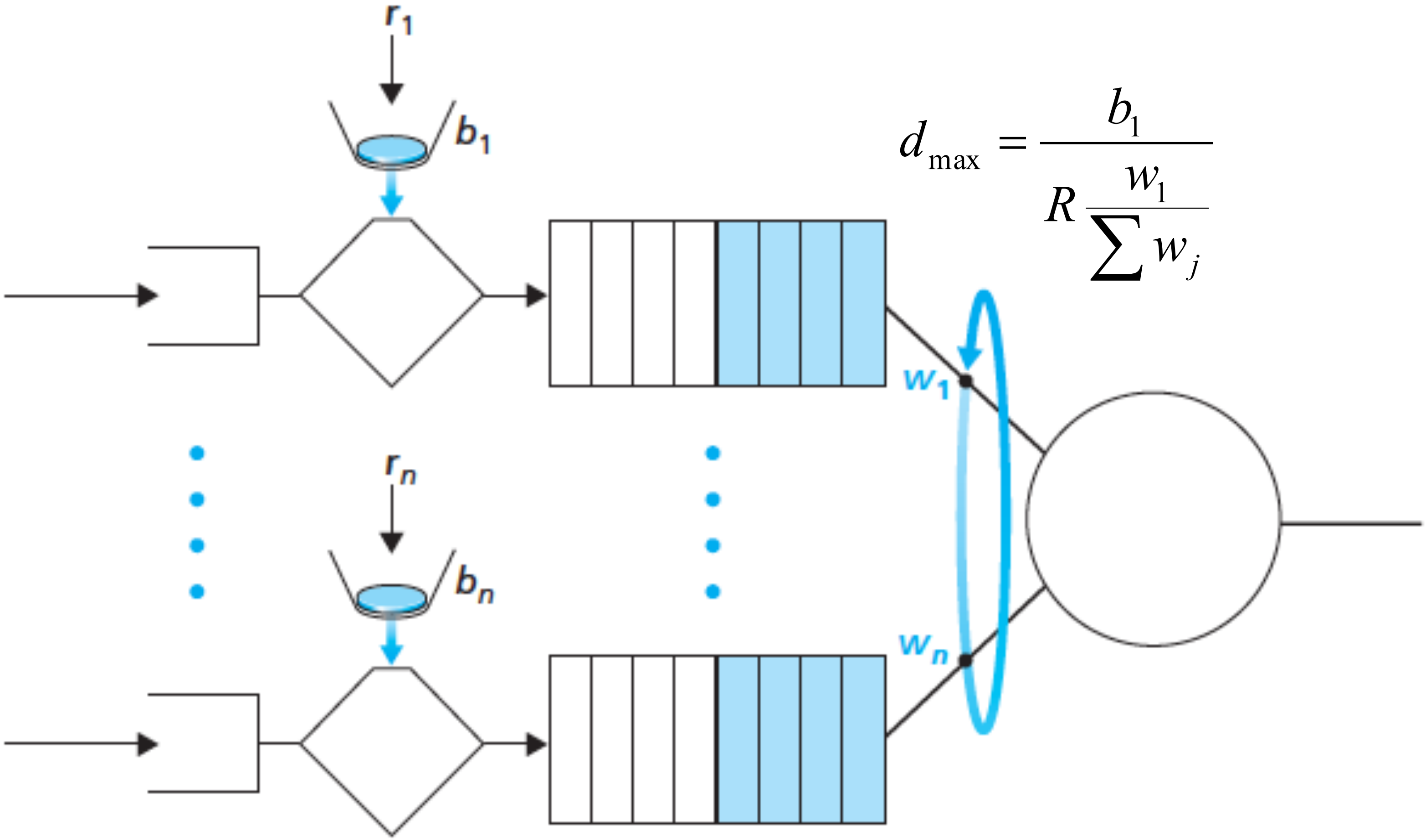
Policing

- Policing
  - regulation of the rate at which a class or flow is allowed to inject packets into the network



The leaky bucket policer

Policy



$$d_{\max} = \frac{b_1}{R\frac{w_1}{\sum w_j}}$$

n multiplexed leakybucket flows with WFQ scheduling

References

- Material
  - Slides
  - Video Lessons

- Books
  - Computer Networking: A Top-Down Approach, J. F. Kurose, K. W. Ross, Pearson, 6 edition, 2013

ISP – Verification tests

Question 29

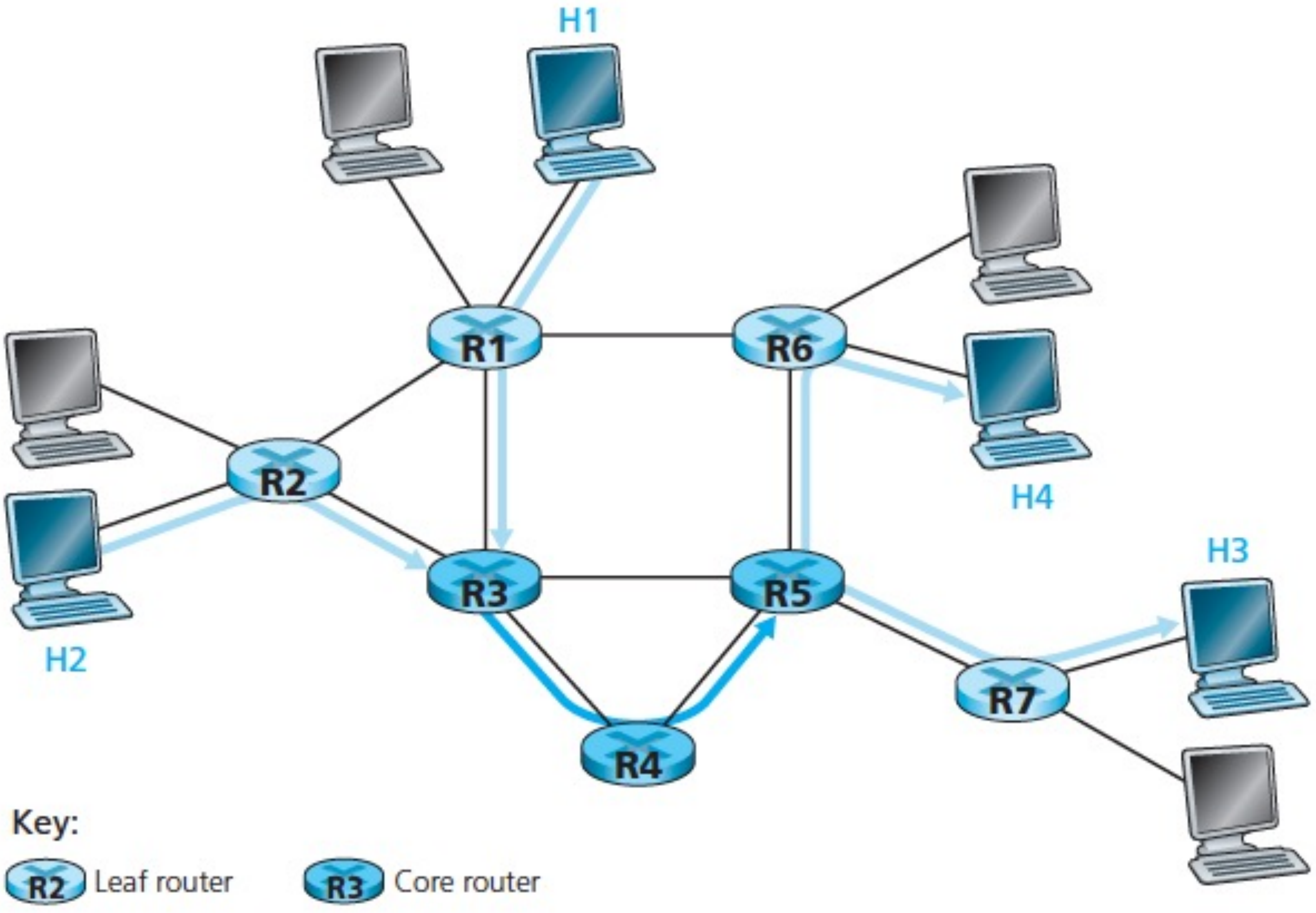
- Multimedia Networking

- Question
  - Describe the Diffserv and Intserv mechanisms

Diffserv

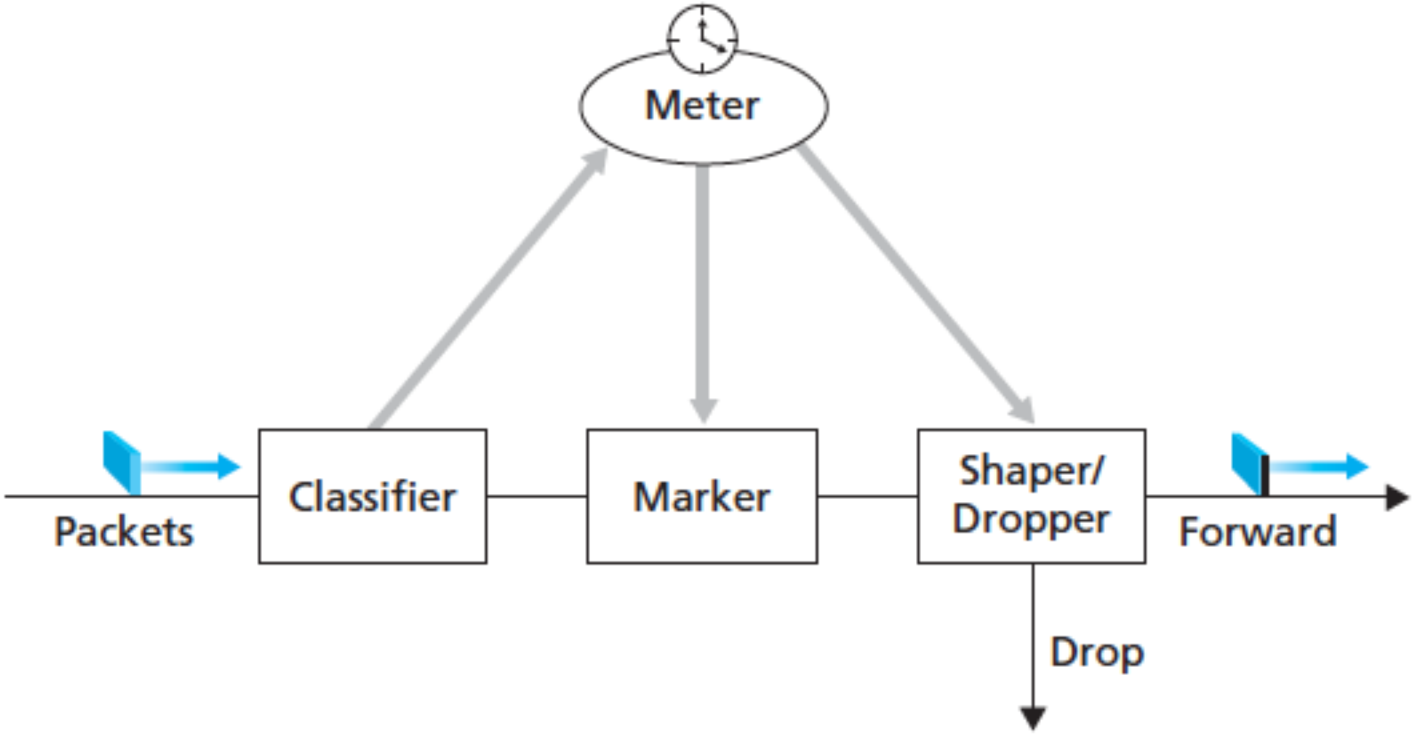
- Diffserv
  - provides service differentiation
    - *the ability to handle different classes of traffic in different ways within the Internet in a scalable manner*
  - Edge functions
    - packet classification and traffic conditioning
  - Core function
    - forwarding

Diffserv



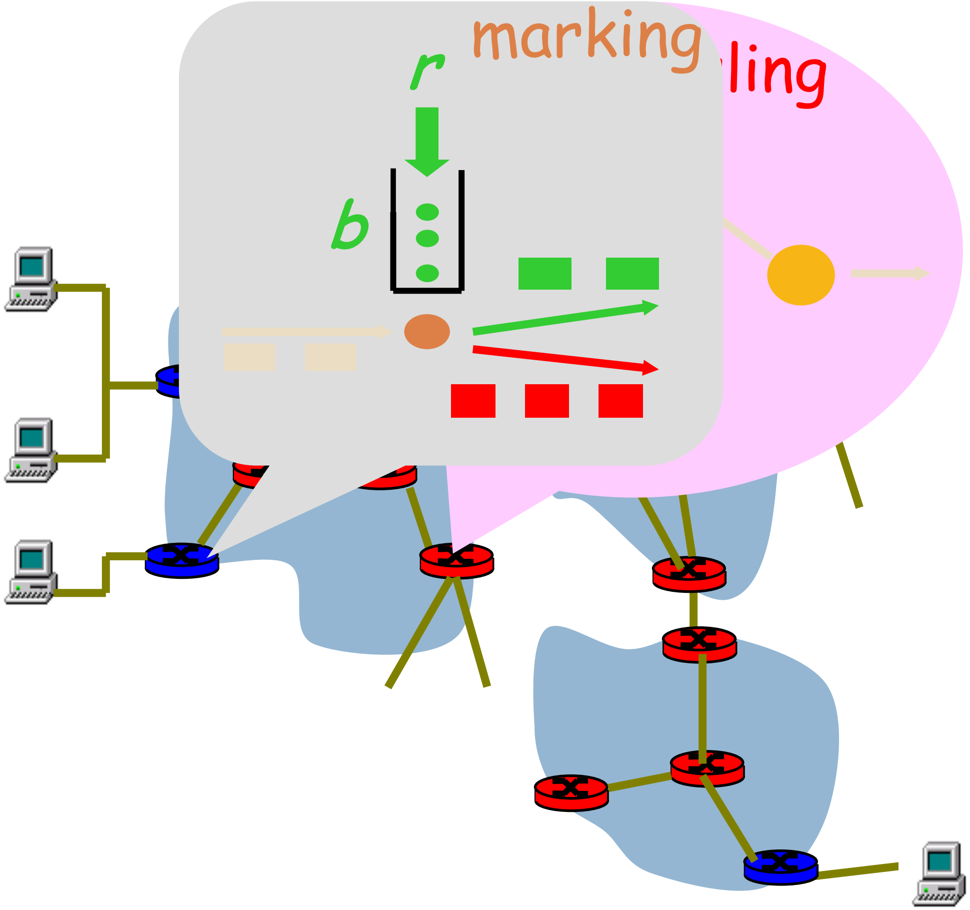
A simple Diffserv network example

Diffserv



A simple Diffserv network example

Diffserv



Diffserv network example

Intserv

- to be guaranteed a given QoS
  - Resource reservation
    - *resources are reserved*
  - Call admission
    - *network must have a mechanism for calls to request and reserve resources*
  - Call setup signaling
    - protocol is needed to coordinate these various activities (call setup protocol)
    - ReSerVation Protocol (RSVP)

Intserv

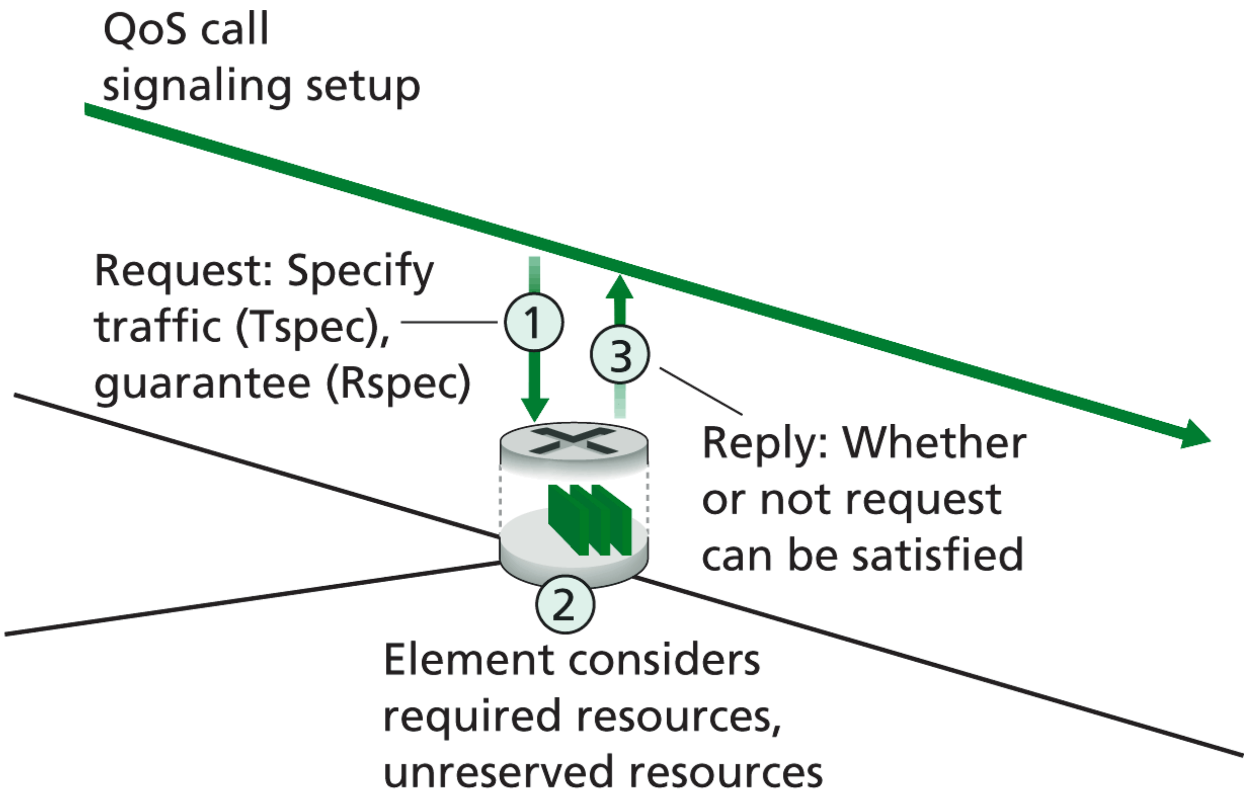


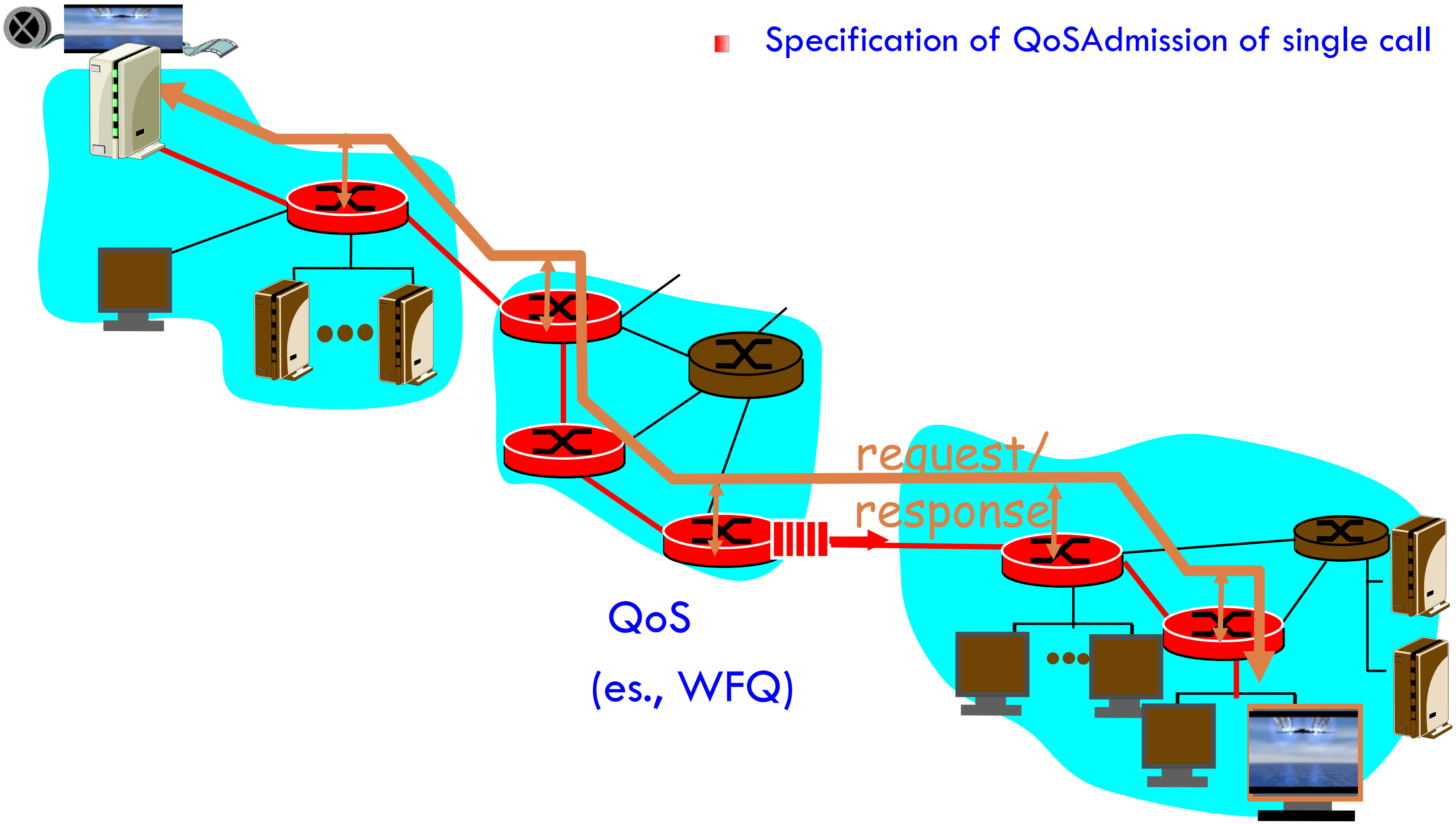
Figure 7.32 ♦ Per-element call behavior

The call setup process

Intserv

Resources reservation

- Call Setting Message (RSVP)
- Traffic Characterization and
- Specification of QoSAdmission of single call



References

- Material
  - Slides
  - Video Lessons
- Books
  - Computer Networking: A Top-Down Approach, J. F. Kurose, K. W. Ross, Pearson, 6 edition, 2013

Question 30

- *Information theory is a branch of applied mathematics and electrical engineering involving the quantication of information*

- Question
  - Describe the basis of Information Theory

Introduction

- *Information theory is a branch of applied mathematics and electrical engineering involving the quantication of information*

- *Claude E. Shannon (1948)*
  - Finds fundamental limits on signal processing operations, such as compressing data and reliably storing and communicating data

Information Theory

- What's information?
  - Information is the reduction of uncertainty
  - Some (informal) axioms
    - if something is certain its uncertainty = 0
    - uncertainty should be maximum if all choices are equally probable
    - uncertainty (information) should add for independent sources

Information Theory

- How to measure information content?
  - Let X be a random variable whose outcome x takes values in $\{a_1, \dots, a_L\}$ with probabilities $\{p_1, \dots, p_L\}$
  - Shannon's information content for the outcome x = a_i

$$H(x = a_i) = \log_2\left(\frac{1}{P(x = a_i)}\right) = \log_2\left(\frac{1}{p_i}\right)$$

  - Entropy

$$H(X) = \sum_i p_i \log_2\left(\frac{1}{p_i}\right) = -\sum_i p_i \log_2(p_i)$$

sensible measure of expected (average) information content

Information Theory

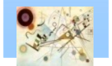
- Information content
  - How many bits needed to compress your data?
  - Example
    - Observe a sequence «…00000100» with p_1 = 0.1 (or p_0 = 0.9)

$$H(x = 1) = \log_2\left(\frac{1}{0.1}\right) = 3.3 bits$$

$$H(x = 0) = \log_2\left(\frac{1}{0.9}\right) = 0.15 bits$$

Information Theory

- Intuition
  - The «1» has less information
    - you don't get too much surprised with a 0
  - You don’t learn too much with a 0
  - The «1» is
    - more improbable
    - more surprising
    - more informative

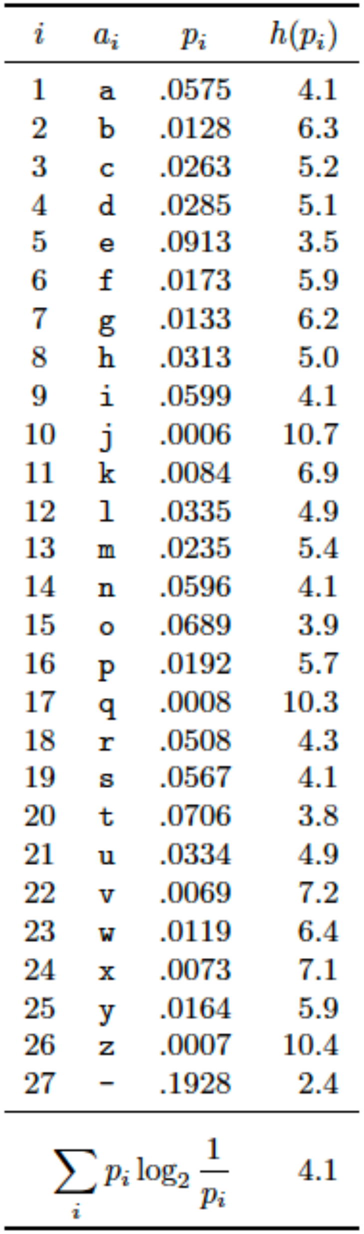


Information Theory

The entropy of an ensemble

$$H(X) \equiv \sum_{x \in \mathcal{A}_X} P(x) \log \frac{1}{P(x)},$$

$P(x) = 0$ that $0 \times \log 1/0 \equiv 0$ $\lim_{\theta \to 0^+} \theta \log 1/\theta = 0$

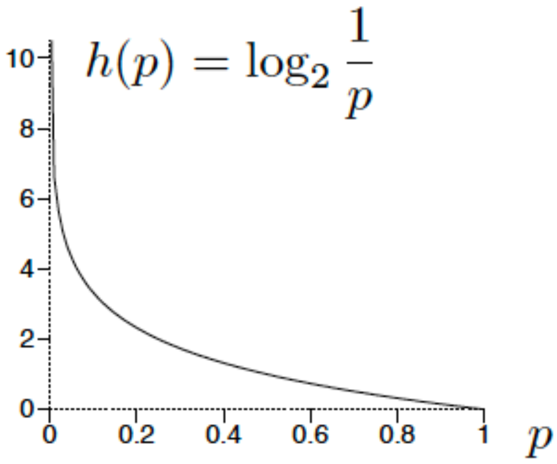


i	a_i	p_i	$h(p_i)$
1	`a`	.0575	4.1
2	`b`	.0128	6.3
3	`c`	.0263	5.2
4	`d`	.0285	5.1
5	`e`	.0913	3.5
6	`f`	.0173	5.9
7	`g`	.0133	6.2
8	`h`	.0313	5.0
9	`i`	.0599	4.1
10	`j`	.0006	10.7
11	`k`	.0084	6.9
12	`l`	.0335	4.9
13	`m`	.0235	5.4
14	`n`	.0596	4.1
15	`o`	.0689	3.9
16	`p`	.0192	5.7
17	`q`	.0008	10.3
18	`r`	.0508	4.3
19	`s`	.0567	4.1
20	`t`	.0706	3.8
21	`u`	.0334	4.9
22	`v`	.0069	7.2
23	`w`	.0119	6.4
24	`x`	.0073	7.1
25	`y`	.0164	5.9
26	`z`	.0007	10.4
27	`-`	.1928	2.4
$\sum_i p_i \log_2 \frac{1}{p_i}$			4.1

Table 2.9. Shannon information contents of the outcomes `a`–`z`.

Information and uncertainty

- Consider a binary random variable that can take two values with probabilities *p* and *1 – p*

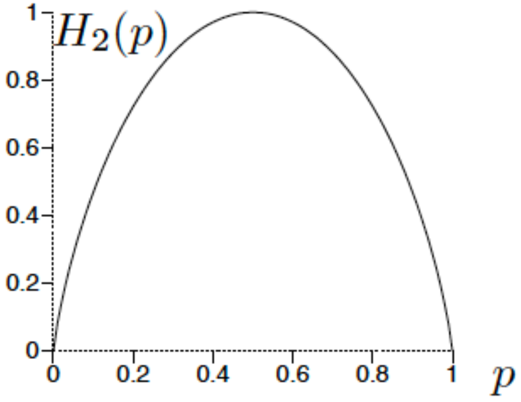


p	$h(p)$	$H_2(p)$
0.001	10.0	0.011
0.01	6.6	0.081
0.1	3.3	0.47
0.2	2.3	0.72
0.5	1.0	1.0

Shannon information content of an outcome with probability *p*, as a function of *p*. The less probable an outcome is, the greater its Shannon information content.

Information and uncertainty

- Consider a binary random variable that can take two values with probabilities *p* and *1 – p*



p	$h(p)$	$H_2(p)$
0.001	10.0	0.011
0.01	6.6	0.081
0.1	3.3	0.47
0.2	2.3	0.72
0.5	1.0	1.0

$$H_2(p) = H(p, 1-p) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{(1-p)}$$

Information and uncertainty

- Improbable events are more informative, but less frequent on average

- The entropy satisfies the two first axioms
  - observation of a certain event carries no information
  - maximum information is carried by uniformly probable events

Information under independence

- Variables x and y that are independent

$$P(x,y) = P(x)P(y)$$

$$h(x,y) = \log \frac{1}{P(x,y)} = \log \frac{1}{P(x)P(y)} = \log \frac{1}{P(x)} + \log \frac{1}{P(y)}$$

$$h(x,y) = h(x) + h(y)$$

- Shannon's information content

$$H(X,Y) = H(X) + H(Y)$$

Differential Entropy

- Vecrtor **a** with PDF *P*(**a**)

$$H(\mathbf{a}) = \int P(\mathbf{a}) \log_2\left(\frac{1}{P(\mathbf{a})}\right) d\mathbf{a} = - \int P(\mathbf{a}) \log_2(P(\mathbf{a})) d\mathbf{a}$$

entropy is related to the PDF volume

$$H(\mathbf{a}) = \frac{1}{2} \ln(2\pi e\sigma^2)$$ Unidimensional Gaussian

$$H(\mathbf{a}) = \frac{1}{\log(2)} \ln((2\pi e\sigma)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}})$$ Multidimensional Gaussian

More about Entropy

- Joint Entropy

$$H(X,Y) = \sum_{xy \in \mathcal{A}_X \mathcal{A}_Y} P(x,y) \log \frac{1}{P(x,y)}$$

$$H(X,Y) = H(X) + H(Y) \;\text{ iff }\; P(x,y) = P(x)P(y)$$

- Conditional Entropy

$$\begin{aligned} H(X \,|\, Y) &\equiv \sum_{y \in \mathcal{A}_Y} P(y) \left[\sum_{x \in \mathcal{A}_X} P(x \,|\, y) \log \frac{1}{P(x \,|\, y)} \right] \\ &= \sum_{xy \in \mathcal{A}_X \mathcal{A}_Y} P(x,y) \log \frac{1}{P(x \,|\, y)}. \end{aligned}$$

More about Entropy

- Chain rule for information content

$$\log \frac{1}{P(x,y)} = \log \frac{1}{P(x)} + \log \frac{1}{P(y\,|\,x)} \qquad h(x,y) = h(x) + h(y\,|\,x)$$

- Chain rule for entropy

$$H(X,Y) = H(X) + H(Y\,|\,X) = H(Y) + H(X\,|\,Y)$$

More about Entropy

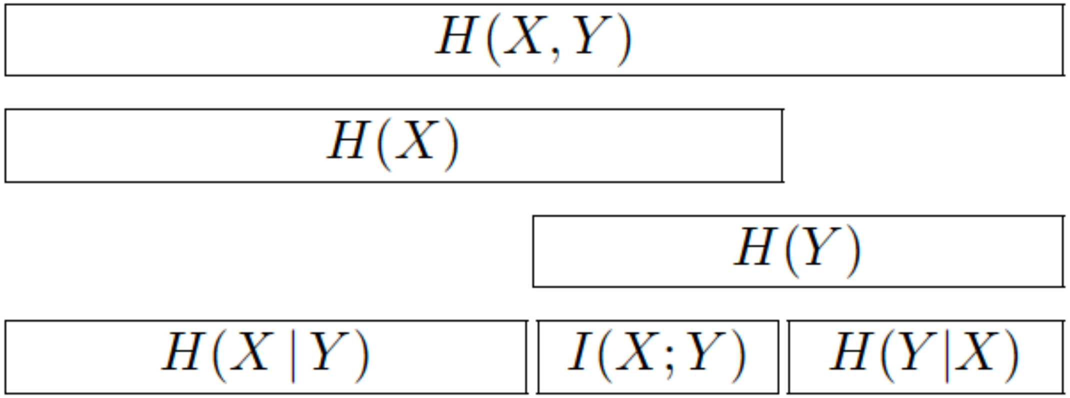
- Mutual Information

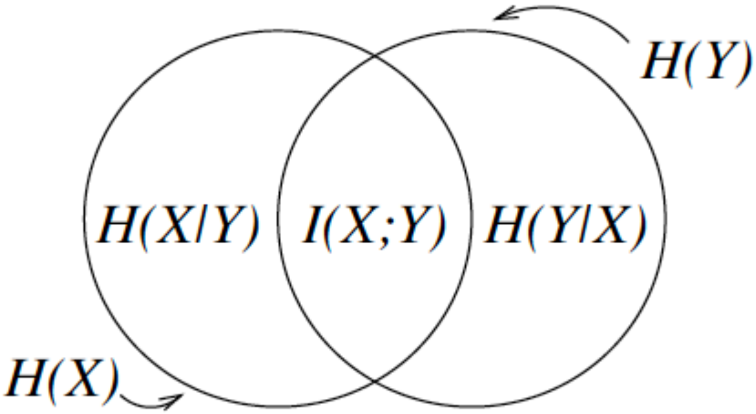
$$I(X;Y) \equiv H(X) - H(X|Y)$$

$$I(X;Y) = I(Y;X)$$

$$I(X;Y) \geq 0$$

More about Entropy





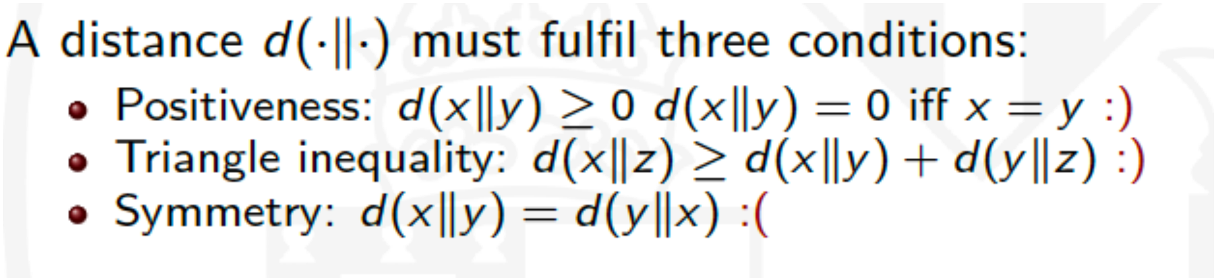
Kullback-Leibler Divergence

- Kind of distance

$$D_{KL}\left(P(\mathbf{a}),Q(\mathbf{a})\right) = \int P(\mathbf{a})\log_2\left(\frac{P(\mathbf{a})}{Q(\mathbf{a})}\right) d\mathbf{a}$$

$$D_{KL} \geq 0$$

$$D_{KL} = 0 \quad iff\ P(\mathbf{a}) = Q(\mathbf{a})$$



A distance $d(\cdot\|\cdot)$ must fulfil three conditions:

- Positiveness: $d(x\|y) \geq 0$ $d(x\|y) = 0$ iff $x = y$:)
- Triangle inequality: $d(x\|z) \geq d(x\|y) + d(y\|z)$:)
- Symmetry: $d(x\|y) = d(y\|x)$:(

Cross-Entropy

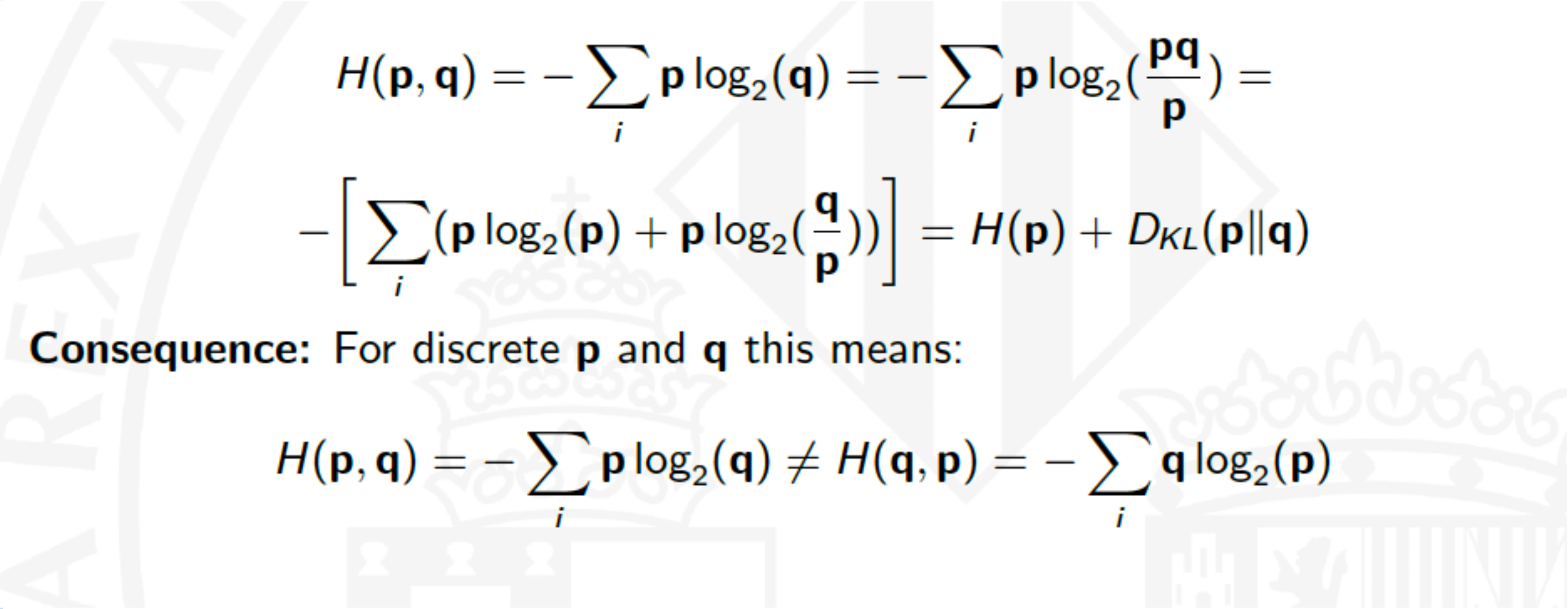
- Two distributions **p** and **q**

$$H(\mathbf{p},\mathbf{q}) = H(\mathbf{p}) + D_{KL}(\mathbf{p}||\mathbf{q})$$

$$D_{KL}(\mathbf{p}||\mathbf{q}) = H(\mathbf{p},\mathbf{q}) - H(\mathbf{p})$$

Cross-Entropy

- Two distributions **p** and **q**



$$H(\mathbf{p},\mathbf{q}) = -\sum_i \mathbf{p}\log_2(\mathbf{q}) = -\sum_i \mathbf{p}\log_2(\frac{\mathbf{pq}}{\mathbf{p}}) =$$

$$-\left[\sum_i (\mathbf{p}\log_2(\mathbf{p}) + \mathbf{p}\log_2(\frac{\mathbf{q}}{\mathbf{p}}))\right] = H(\mathbf{p}) + D_{KL}(\mathbf{p}\|\mathbf{q})$$

Consequence: For discrete **p** and **q** this means:

$$H(\mathbf{p},\mathbf{q}) = -\sum_i \mathbf{p}\log_2(\mathbf{q}) \neq H(\mathbf{q},\mathbf{p}) = -\sum_i \mathbf{q}\log_2(\mathbf{p})$$

More on MI

- Mutual Information

$$I(x,y) = \sum_x \sum_y p(x,y) \log\left(\frac{p(x,y)}{p_1(x)p_2(x)}\right)$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

More on MI

- Mutual Information

$$I(x,y) = \sum_{x} \sum_{y} p(x,y) \log\left(\frac{p(x,y)}{p_1(x)p_2(x)}\right)$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

References

- Material
  - Slides
  - Video Lessons

- Books
  - Information Theory, Inference and Learning Algorithms, D. J. C. MacKay, Cambridge: Cambridge University Press., 2003

Question 31

- Compressive Sensing (CS) is a new sensing modality, which compresses the signal being acquired at the time of sensing

- Question
  - Describe the basis of Compressive Sensing

Question 31

- Compressive Sensing (CS) is a new sensing modality, which compresses the signal being acquired at the time of sensing

- Question
  - Describe the basis of Compressive Sensing

Introduction

- Compressive Sensing (o Compressed Sensing) technique
- client-server architecture
- Compressice Sensing for
  - compression
  - packet loss reconstruction

Compressive Sensing

- Compressive Sensing (CS)
  - is a new sensing modality, which compresses the signal being acquired at the time of sensing
  - Signals can have sparse or compressible representation either in original domain or in some transform domain
  - Relying on the sparsity of the signals, CS allows us to sample the signal at a rate much below the Nyquist sampling rate
  - the varied reconstruction algorithms of CS can faithfully reconstruct the original signal back from fewer compressive measurements

Compressive Sensing

- CS was introduced by Donoho, Candès, Romberg, and Tao in 2004

Compressive Sensing

- Emerging technique for signal processing
  - acquisition/reconstruction that violates the Nyquist-Shannon limit
    - less samples

- A signal can have sparse/compressible representation either in original domain or in some transform domains
  - Fourier transform, cosine transform, wavelet transform, etc. A few examples of signals having sparse

- Domains
  - natural images which have sparse representation in wavelet domain
  - speech signal can be represented by fewer components using Fourier transform
  - better model for medical images can be obtained using Radon transform
  - etc.

Linear inverse problems

- Many classic problems in computer can be posed as linear inverse problems

- Notation
  - Signal of interest $x \in \mathbb{R}^N$
  - Observations $y \in \mathbb{R}^M$
  - Measurement model $y = \Phi x + e$

    measurement matrix

    measurement noise
  - Problem definition: given y, recover x

Linear inverse problems

- Scenario 1

$$M \geq N$$

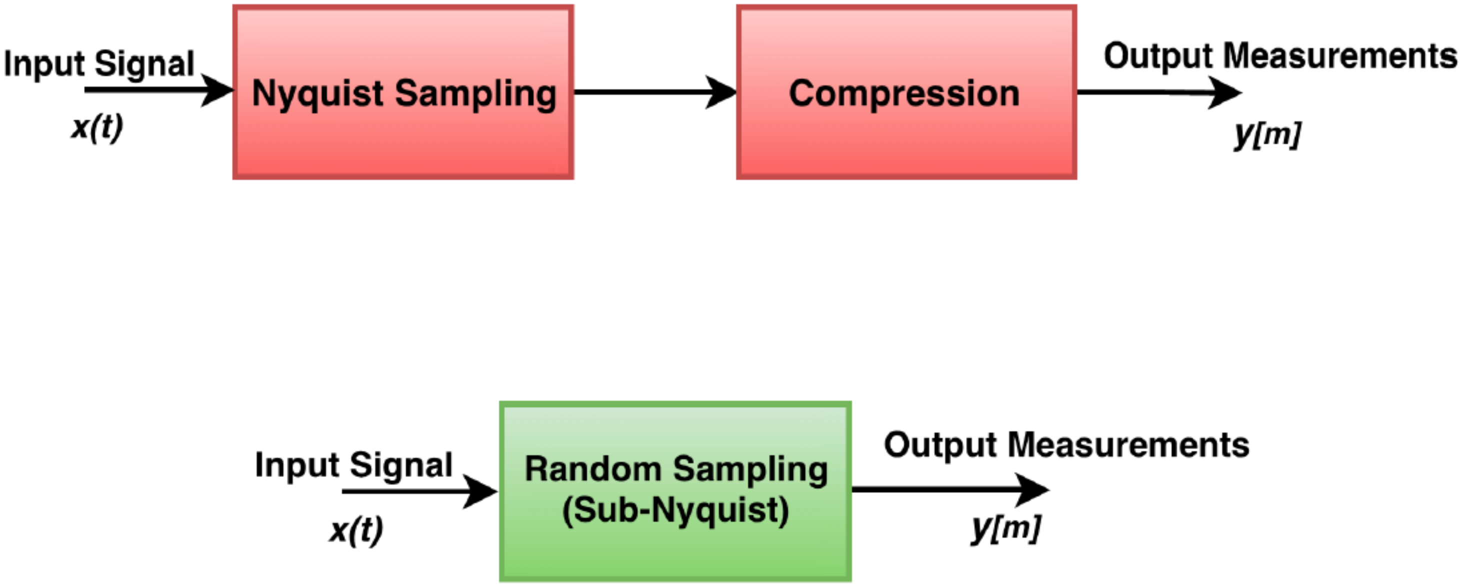
$$\hat{x} = \Phi^{-1} y$$

- Scenario 2

$$M < N$$

  - Measurement matrix has a (N-M) dimensional null-space
  - Solution is no longer unique
  - Under-sampling ratio M/N

Compressive Sensing



Input Signal

x(t)

Nyquist Sampling

Compression

Output Measurements

y[m]

Input Signal

x(t)

Random Sampling (Sub-Nyquist)

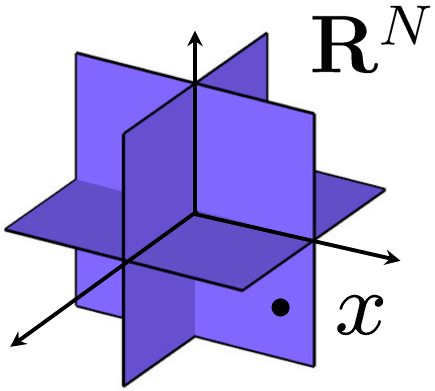
Output Measurements

y[m]

A comparision of sampling techniques: (a) traditional sampling, (b) compressive sensing.

Sparsity

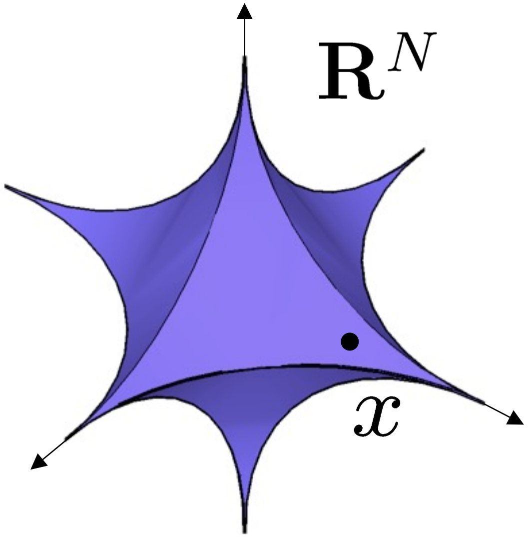
- Sparse signal
  - only K out of N coordinates nonzero
  - Model – union of *k*-dimensional subspaces

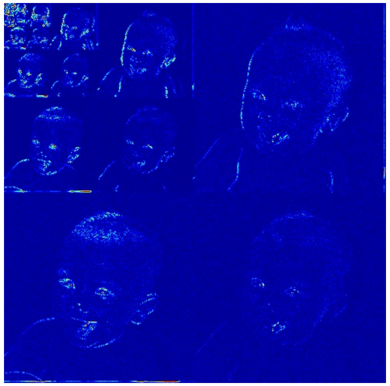


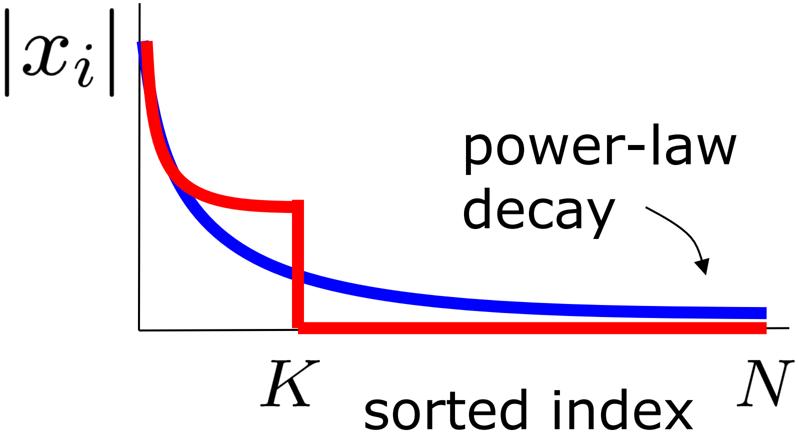
Sparsity

- Compressible signal
  - sorted coordinates decay rapidly with power-law
  - Model based on ℓ_p

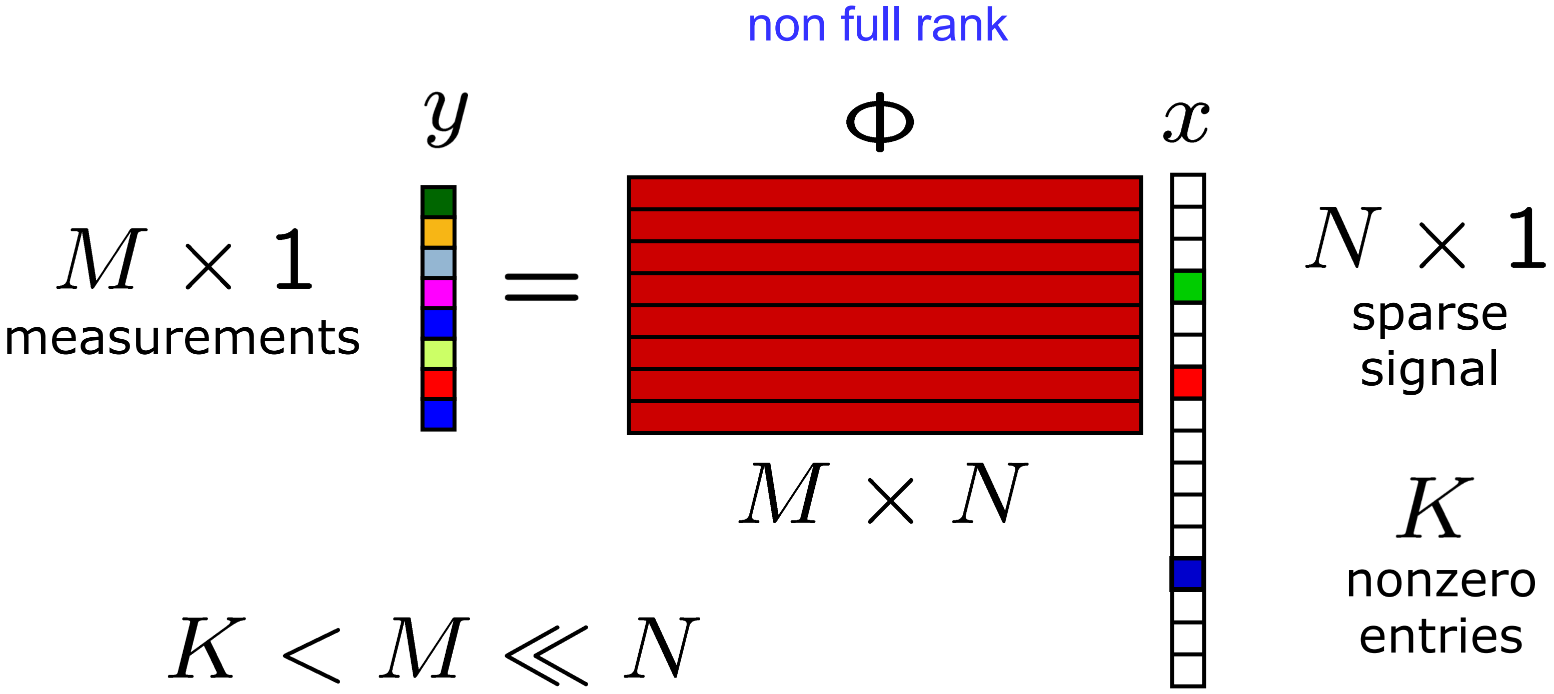
$$\|x\|_p^p = \sum_i |x_i|^p \leq 1, \;\; p \leq 1$$







Compressive Sampling



non full rank

y Φ x

$M \times 1$
measurements

$=$

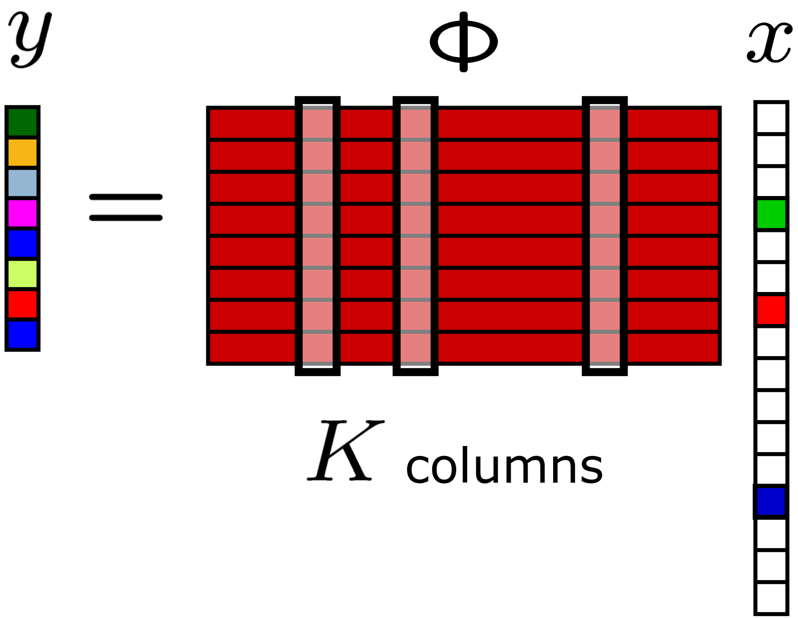
$M \times N$

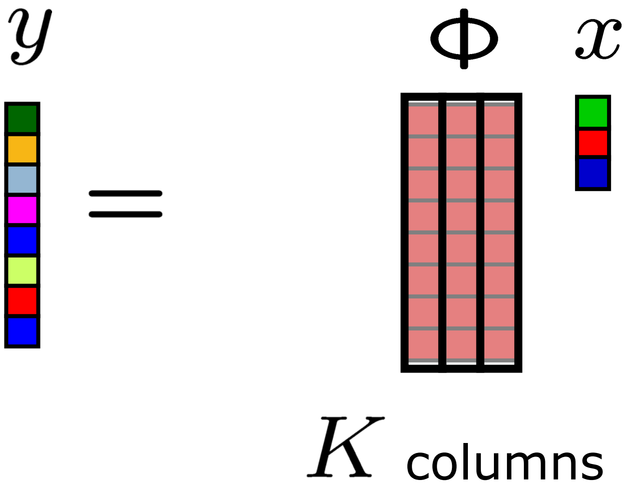
$N \times 1$
sparse
signal

K
nonzero
entries

$$K < M \ll N$$

How can it work?

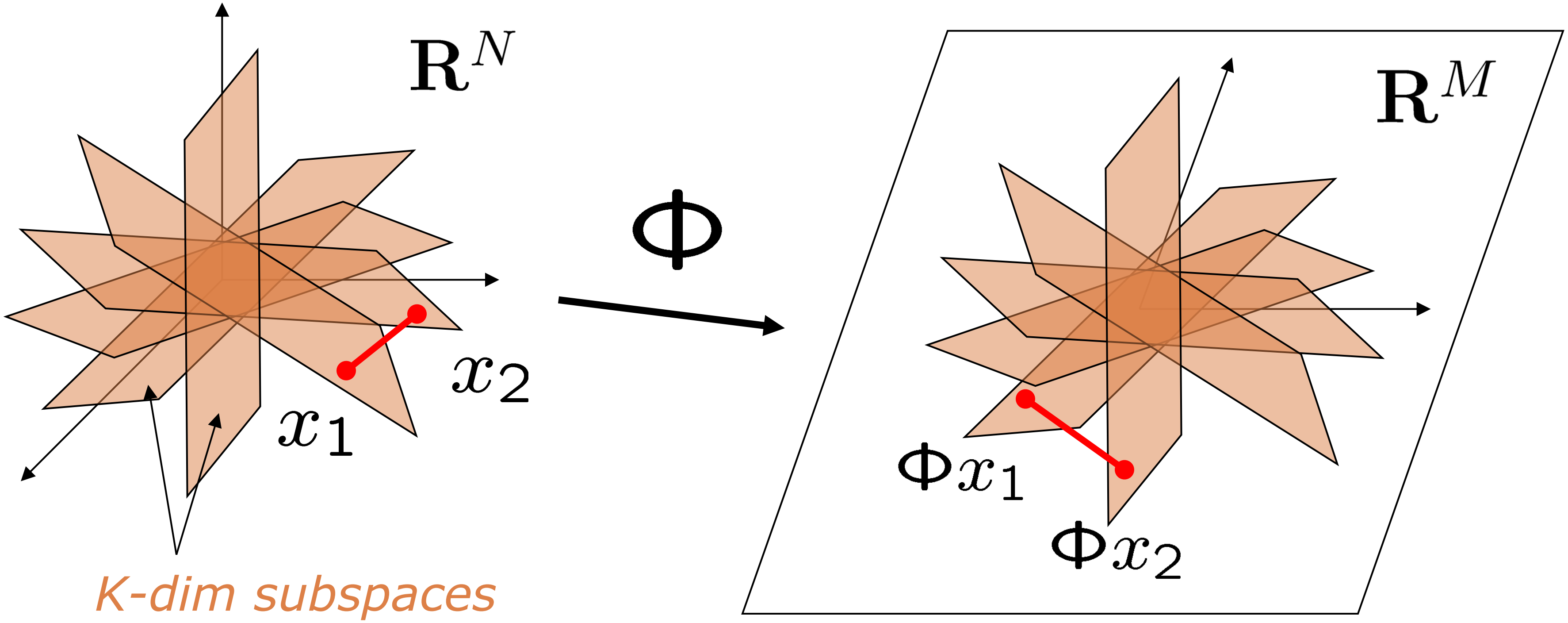




Restricted Isometry Property

- Design Φ so that each of its MxK submatrices are full rank (ideally close to orthobasis)
  - Restricted Isometry Property (RIP)

Preserve the structure of sparse/compressible signals



K-dim subspaces

Restricted Isometry Property

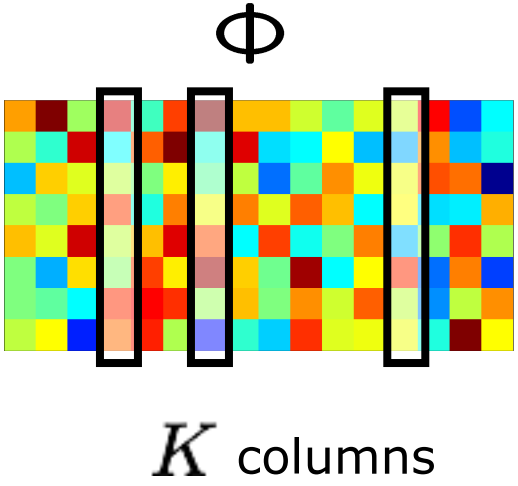
- RIP of order 2K implies
  - for all K-sparse x_1 and x_2

$$(1 - \delta_{2K}) \leq \frac{\|\Phi x_1 - \Phi x_2\|_2^2}{\|x_1 - x_2\|_2^2} \leq (1 + \delta_{2K})$$

  - Ensure that

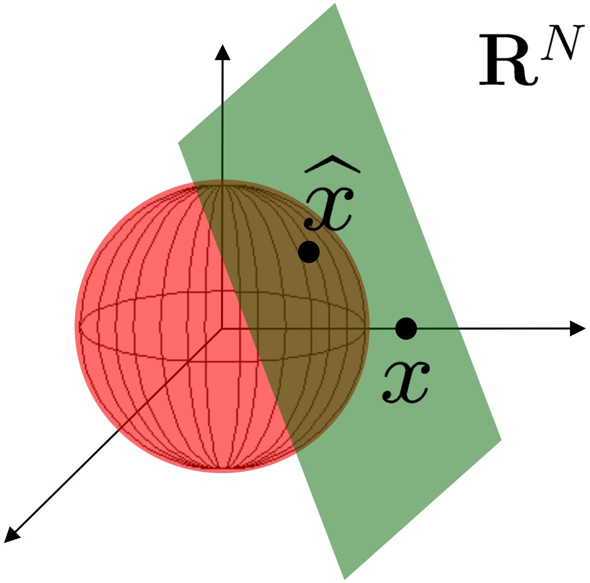
$$\|x_1 - x_2\|_2 \approx \|\Phi x_1 - \Phi x_2\|_2$$

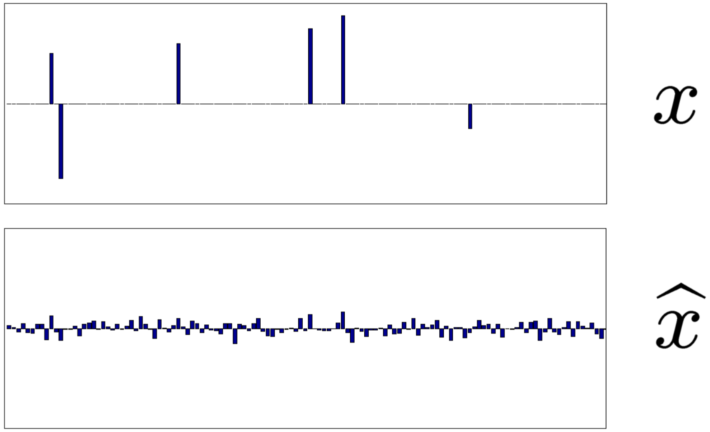
  - Draw Φ at random
    - iid Gaussian
    - iid Bernoulli



L_2 signal recovery

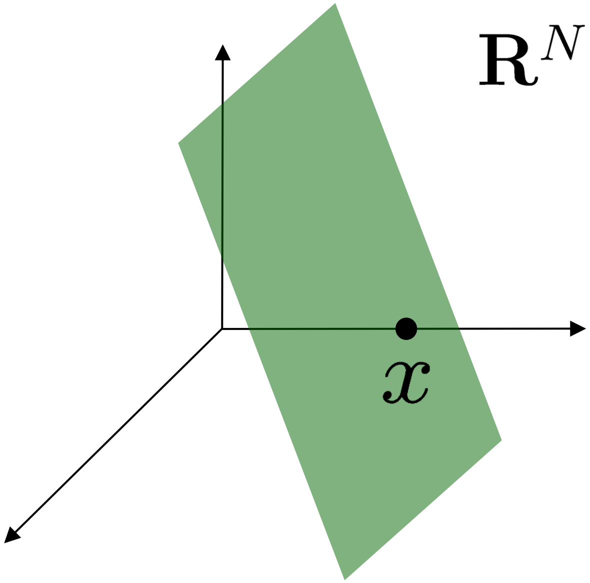
$$\widehat{x} = \arg\min_{y=\Phi x} \|x\|_2$$

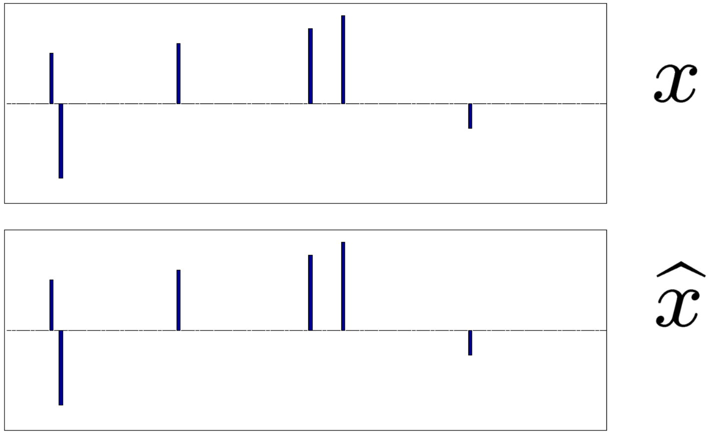




L_0 signal recovery

$$\widehat{x} = \arg\min_{y=\Phi x} \|x\|_0$$

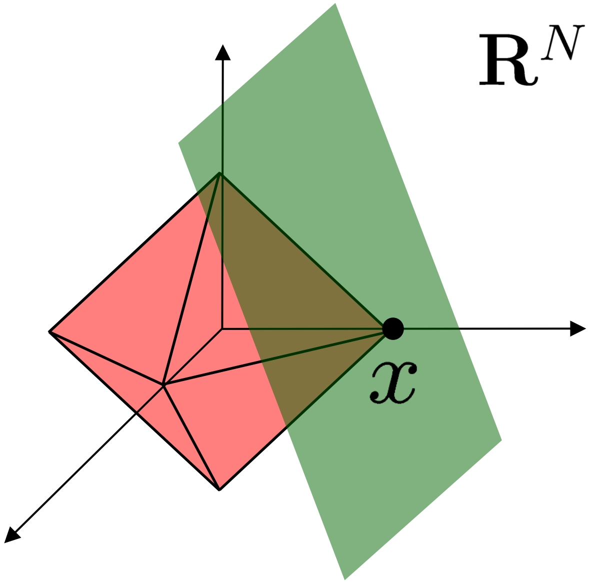




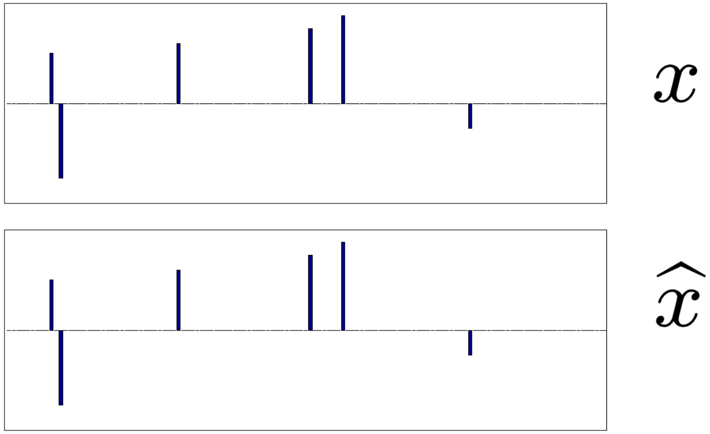
NP-Complete

L_0 signal recovery

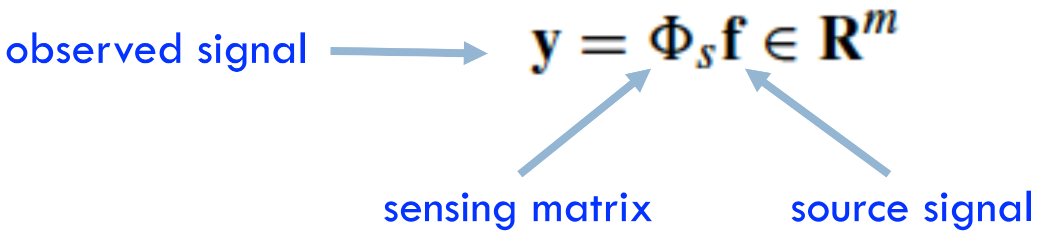
$$\widehat{x} = \text{arg}\min_{y=\Phi x} \|x\|_1$$



Polynomial time alg
(linear programming)



Compressive Sensing

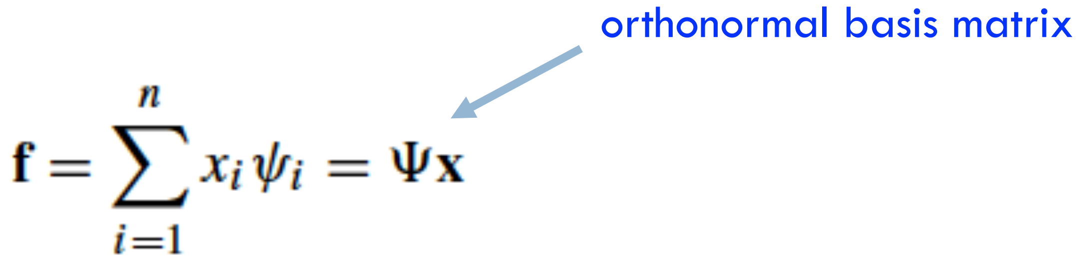


observed signal

$$\mathbf{y} = \Phi_s \mathbf{f} \in \mathbf{R}^m$$

sensing matrix

source signal



orthonormal basis matrix

$$\mathbf{f} = \sum_{i=1}^{n} x_i \psi_i = \Psi \mathbf{x}$$

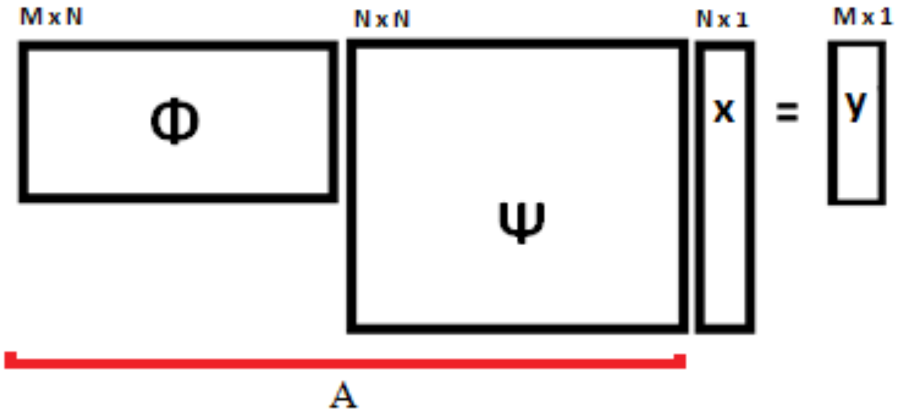
sparse representation

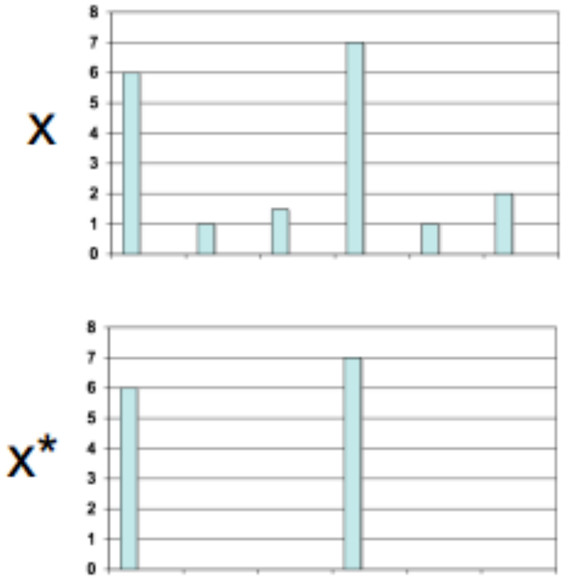
$$\mu(\Phi, \Psi)$$

coherence misure

In our case ϕ is the indentity matrix and ψ is a dictionay (*learned* or obtained by DCT)

Compressive Sensing





k=2

Optimization algorithm

l_0 -minimization
problem is NP-hard

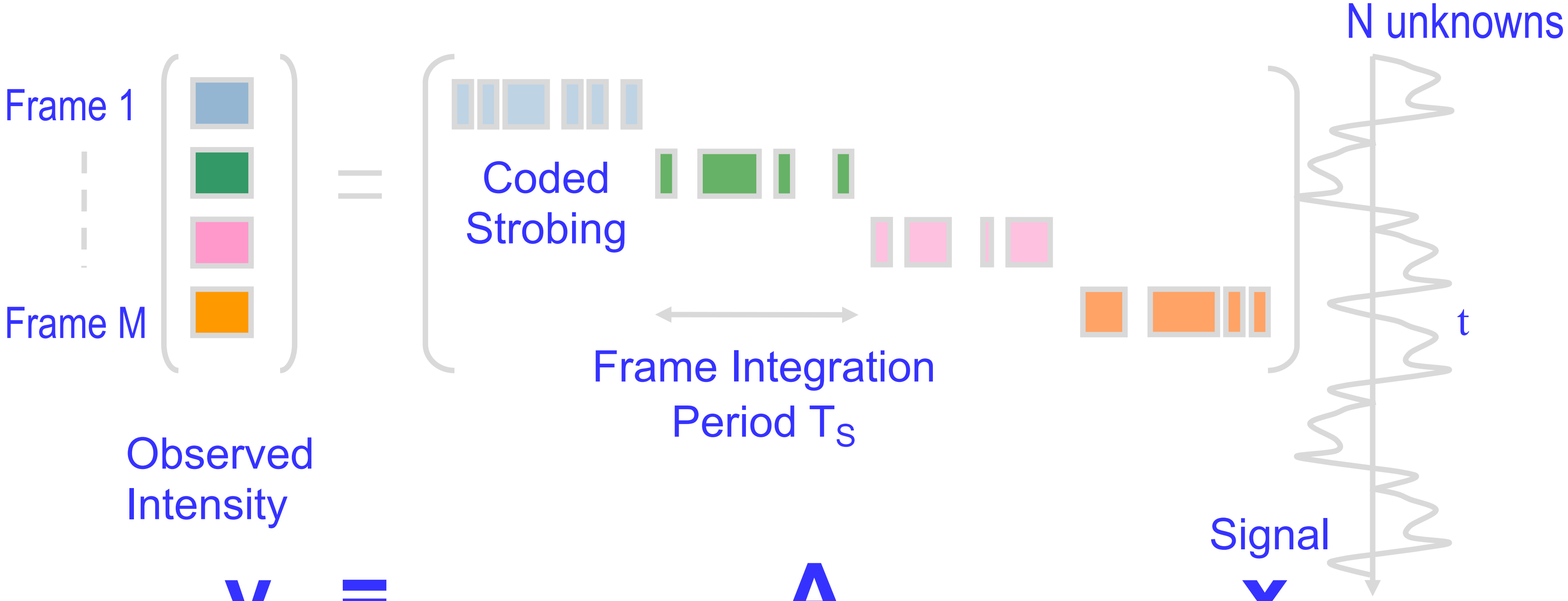
$$\mathbf{f}^* = \Psi \mathbf{x}^*$$

reconstruction

$$\min \|\mathbf{x}\|_{L_1} \quad \text{subject to} \quad y_k = \langle \phi_k, \Psi \mathbf{x} \rangle \quad \forall k \in M$$
$$\mathbf{x} \in R^n$$

Convex optimization algorithm

https://statweb.stanford.edu/~candes/l1magic/#code



N unknowns

Frame 1

Frame M

Coded Strobing

Frame Integration Period T_S

t

Observed Intensity

Signal

$$\mathbf{y} = \mathbf{A} \mathbf{x}$$

Introduction

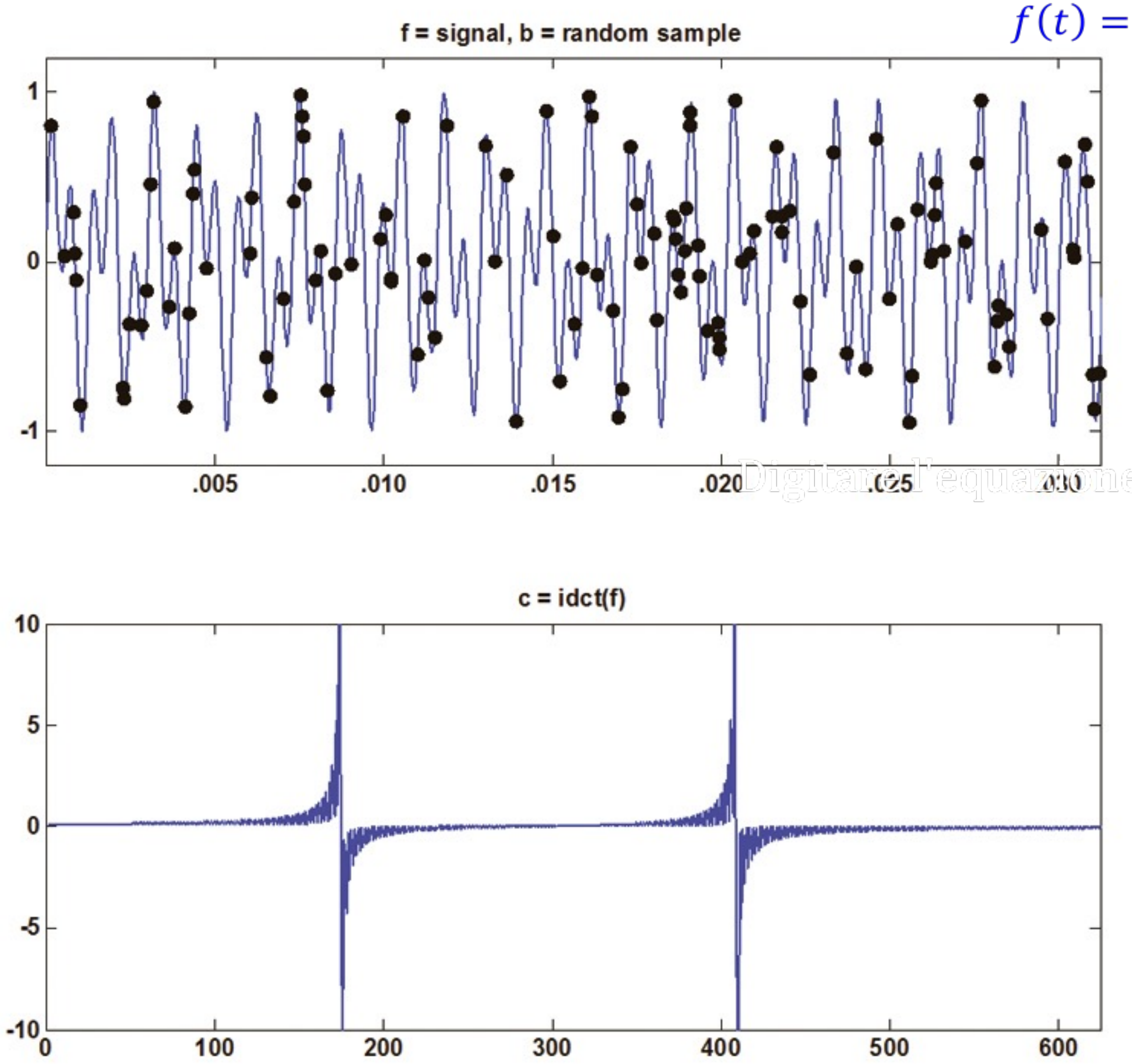
$$f(t) = \sin(1394\,\pi\,t) + \sin(3266\,\pi\,t)$$

1/8 of esecond at 4 kHz

$$n = 5000$$

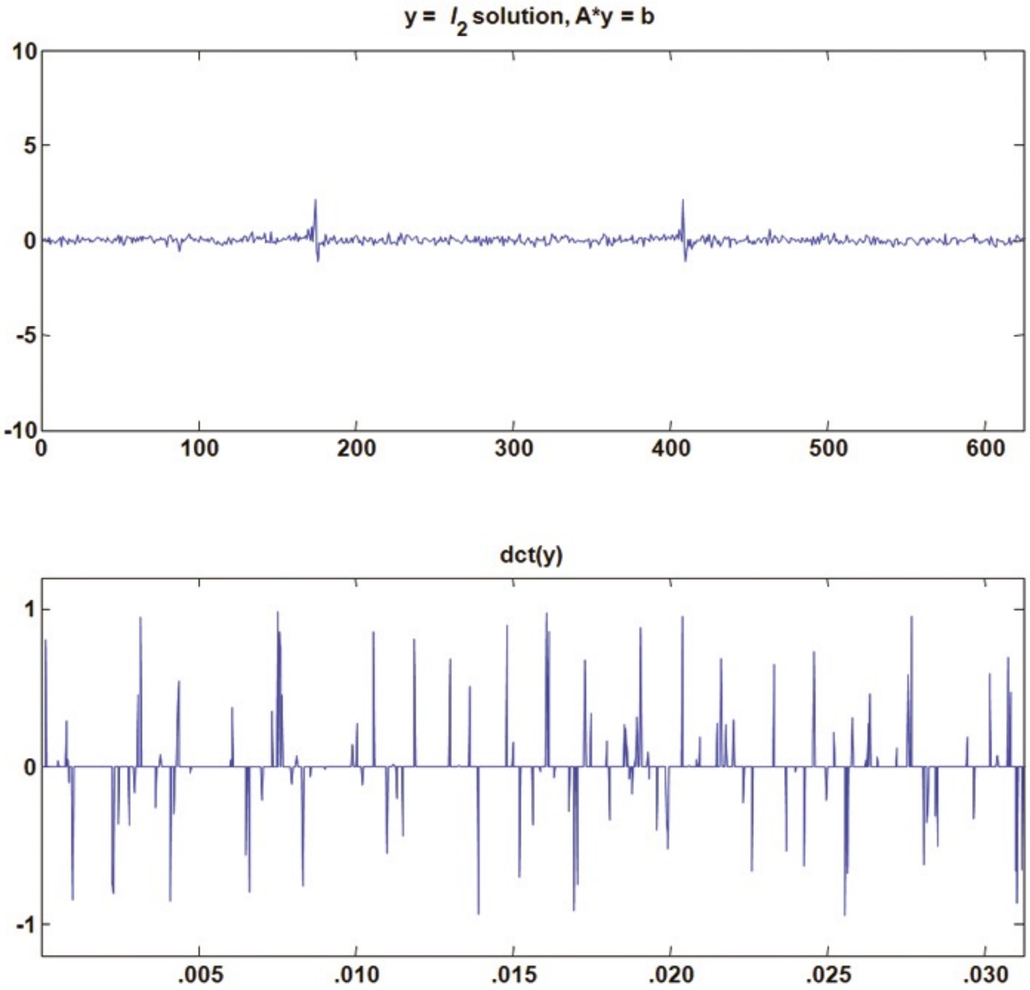
$$m = 500\ random\ samples$$

A DCT based dictionary is used



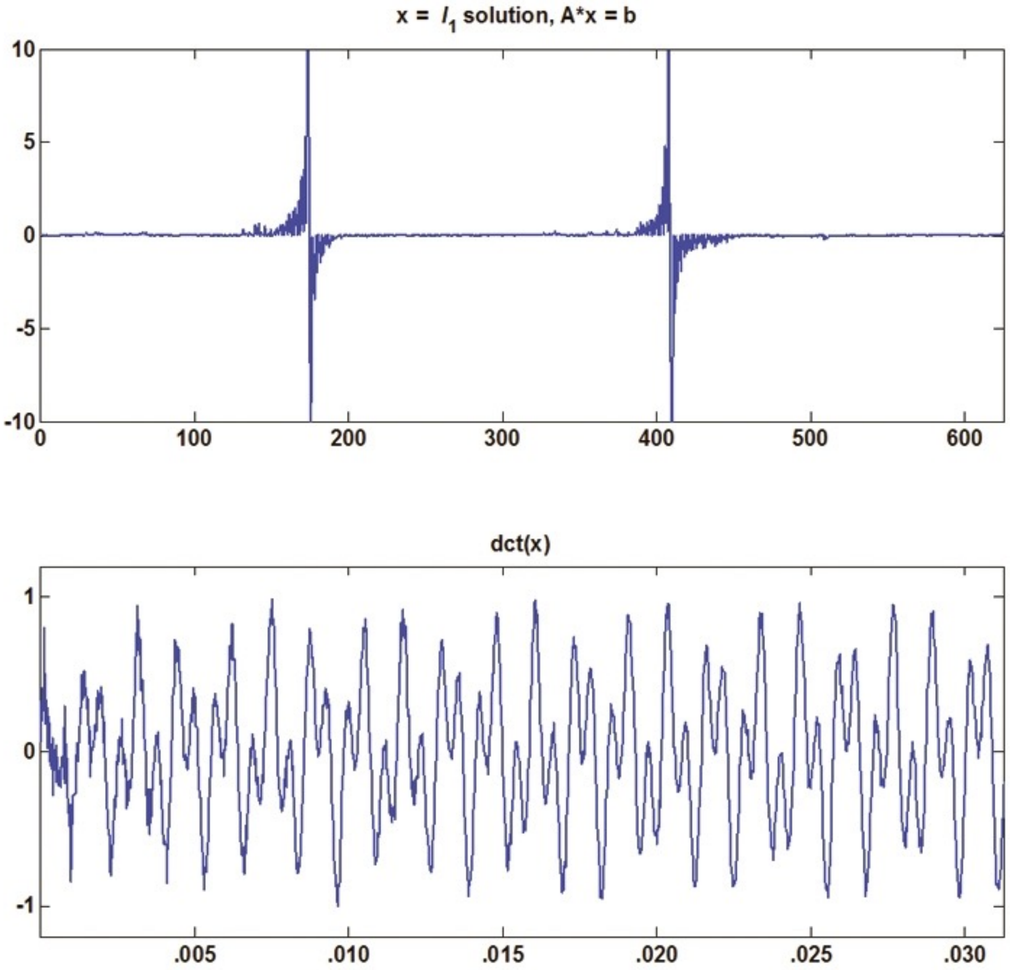
Top: Random samples of the original signal generated by the “A” key on a touch-tone phone. Bottom: The inverse discrete cosine transform of the signal.

Introduction



Results by uisng L_2 norm

Introduction



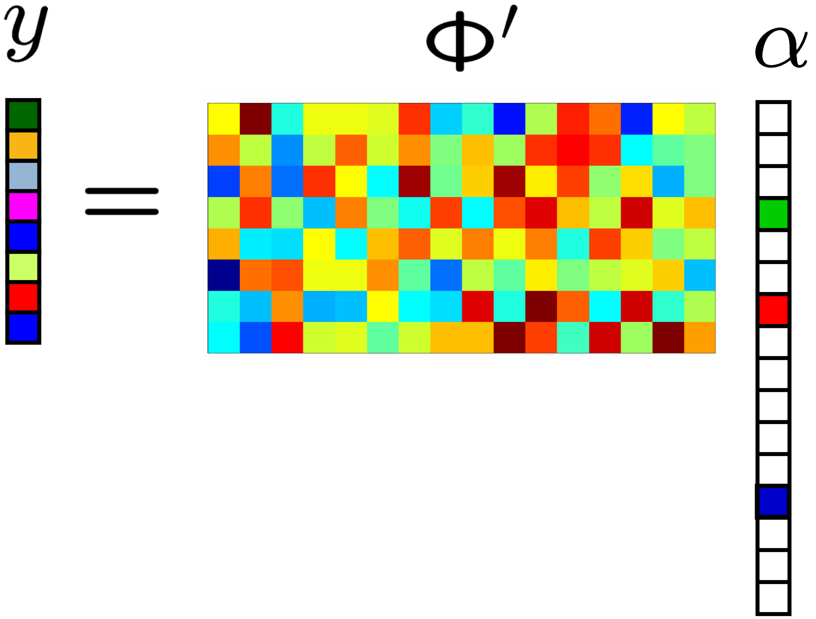
Results by uisng L_1 norm

Optimization algorithms

- Signal recovery via iterative greedy algorithm
  - (orthogonal) matching pursuit
  - iterated thresholding
  - CoSaMP

Universality

$$y = \Phi x = \Phi \Psi \alpha = \Phi' \alpha$$



$N \times 1$
sparse
coefficient
vector

K
nonzero
entries

Random measurements can be used for signals sparse in *any* basis: DCT/FFT/Wavelet/Learned Dictionary

References

- Material
  - Slides
  - Video Lessons

- Books
  - A Mathematical Introduction to Compressive Sensing, S. Foucart , H. Rauhut, 2013

Question 32

- Compressive Sensing (CS) is a new sensing modality, which compresses the signal being acquired at the time of sensing

- Question
  - Describe the basis of Dictionary Learning

Dictionay learning

- Goal
  - Given training data

$$x_1, x_2, \ldots, x_T \qquad x_i \in \mathbb{R}^N$$

  - learn a dictionary D

$$x_i = Ds_i \qquad D \in \mathbb{R}^{N \times Q}$$

$$s_i \in \mathbb{R}^Q$$

where s_i are sparse

Dictionay learning

- Optimization approach

$$\min_{D,S} \|X - DS\|_F$$

$$\text{s.t}$$

$$\forall i,\ \|s_i\|_0 \leq K$$

Non-convex constraint

Non-convex constraint

Bilinear in *D* and *S*

Bilinear in *D* and *S*

Dictionay learning

- Optimization approach

$$\min_{D,S} \|X - DS\|_F$$

$$\text{s.t}$$

$$\forall i,\ \|s_i\|_0 \leq K$$

Non-convex constraint

Bilinear in *D* and *S*

- Biconvex in D and S

$$\min_{D,S} \|X - DS\|_F + \lambda \sum_k \|s_k\|_1$$

Given *D,* the optimization problem is convex in s_k
Given *S,* the optimization problem is a least squares problem

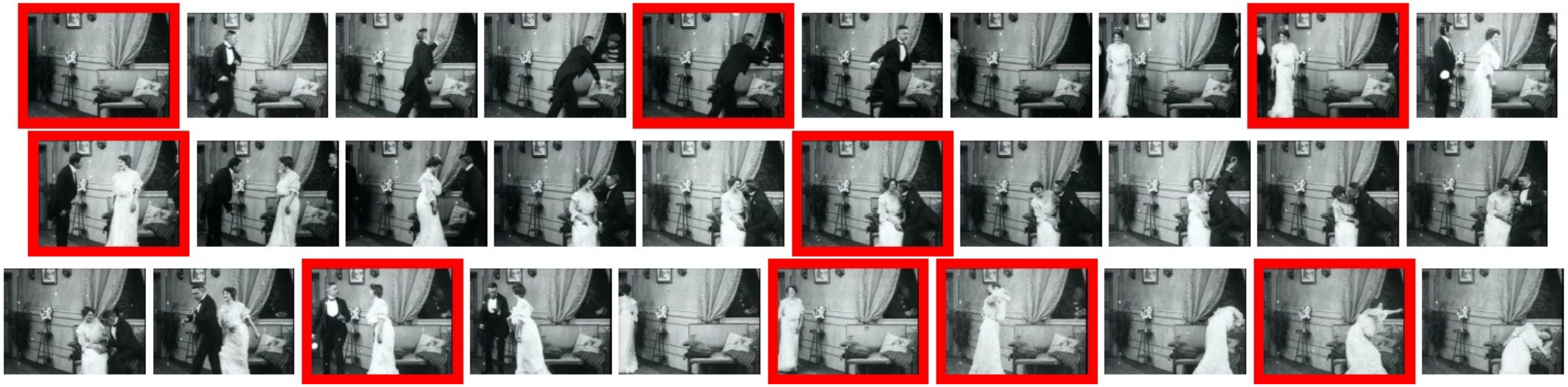
Dictionay learning

- K-SVD
  - Solve using alternate minimization techniques
    - Start with D = wavelet or DCT bases
    - Additional pruning steps to control size of the dictionary

- Sparse Modeling for Finding Representative Objects

$$\min \|\boldsymbol{Y} - \boldsymbol{Y}\boldsymbol{C}\|_F^2 \quad \text{s.t.} \quad \|\boldsymbol{C}\|_{1,q} \le \tau, \ \mathbf{1}^\top \boldsymbol{C} = \mathbf{1}^\top$$

Finding Representative Objects

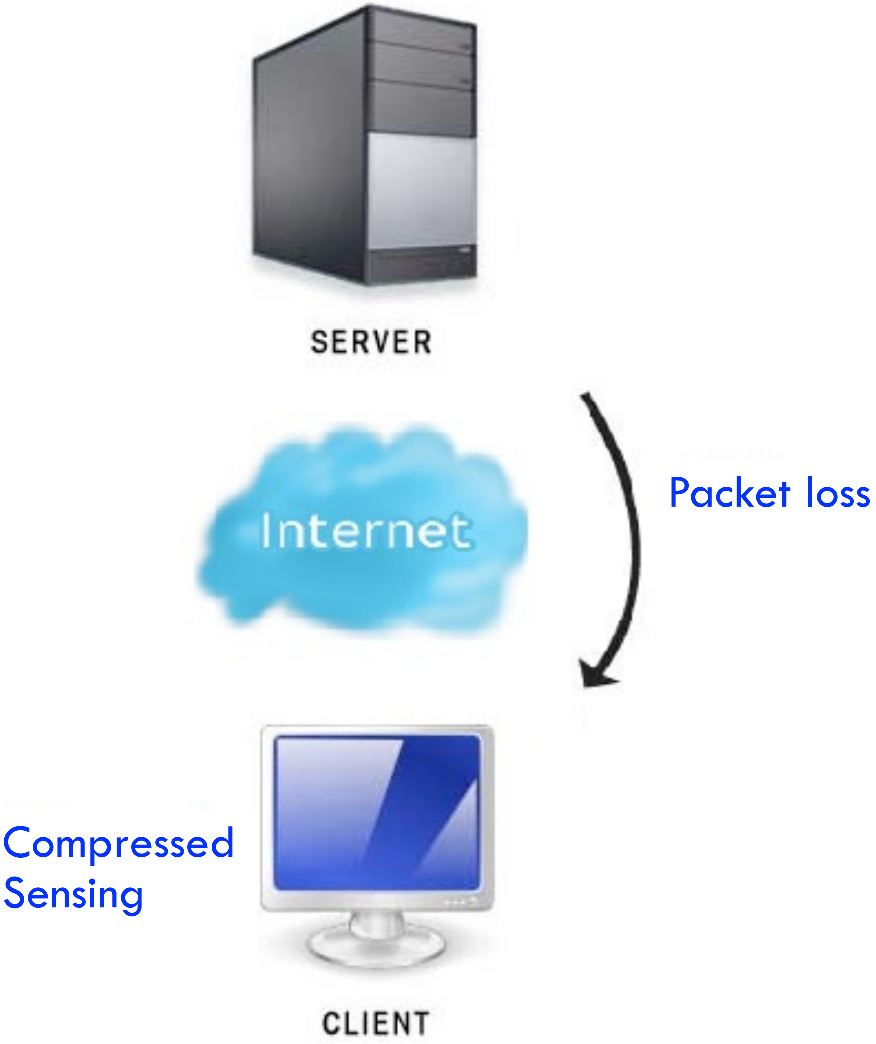


Deblurreing

Blurred Photos

Deblurred Result

First scenario



Second Scenario

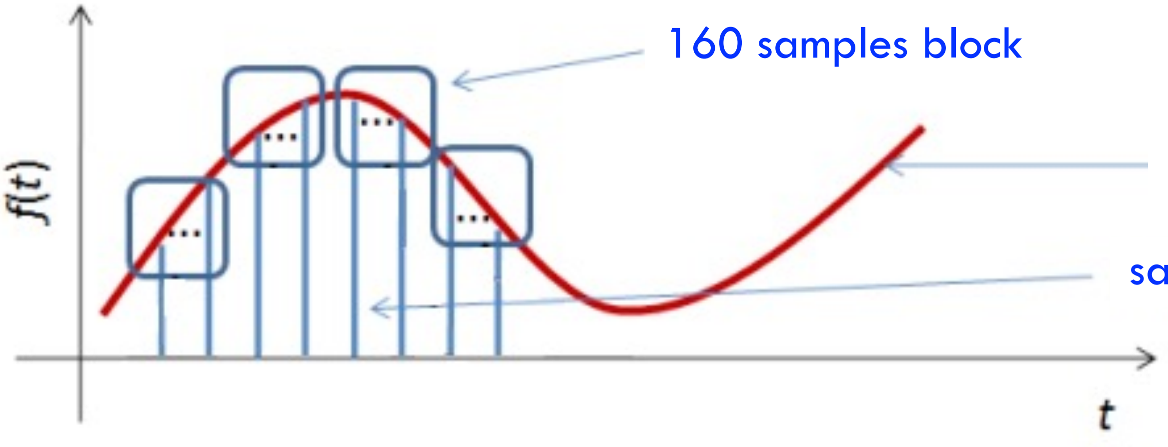
Coding

Compressed Sensing

Decoding

Packet loss

1 sec = 64kbit/sec ---> 1 msec = 64bit ---> 20msec = 1280bit ---> 160 byte

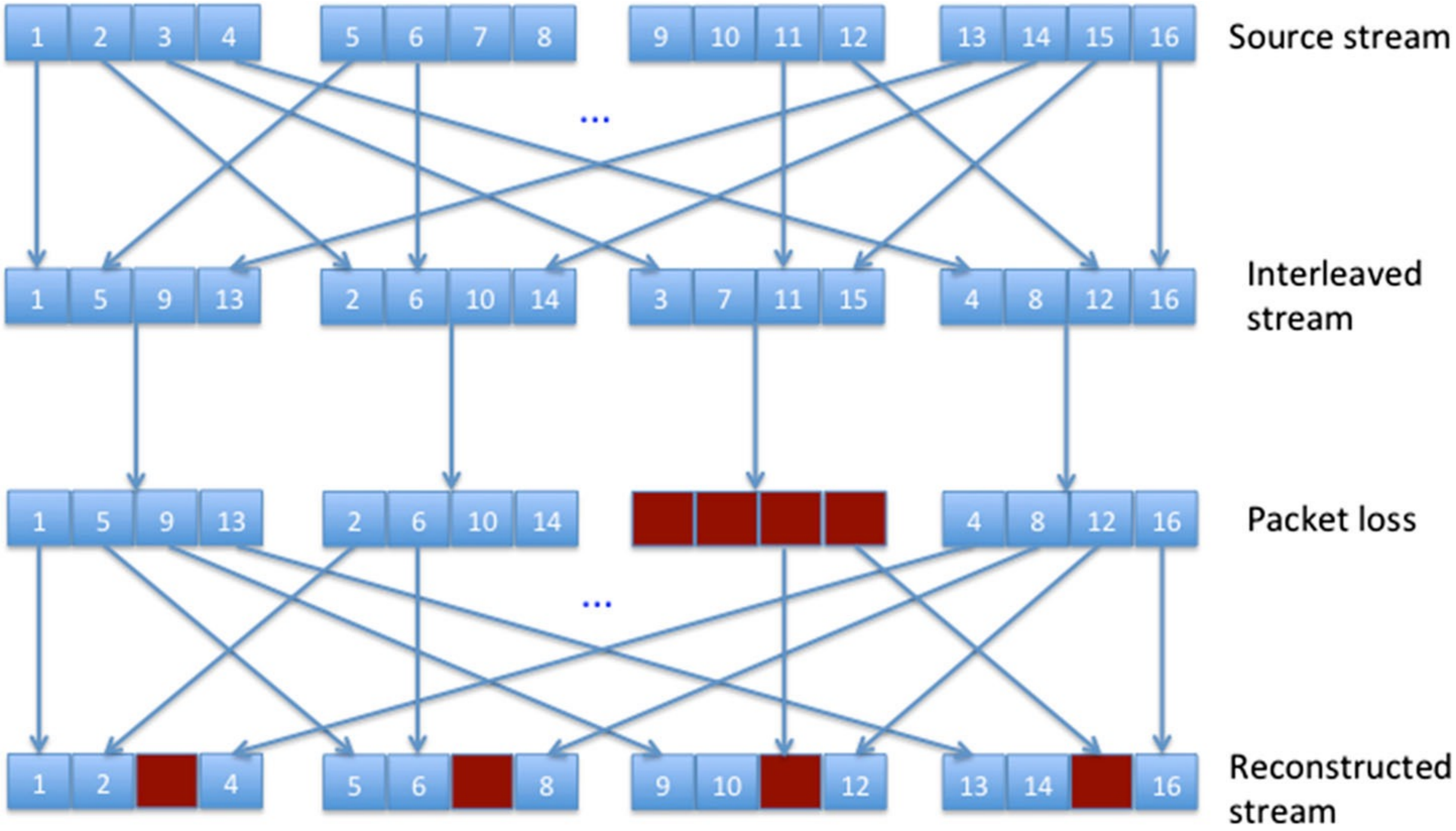


160 samples block

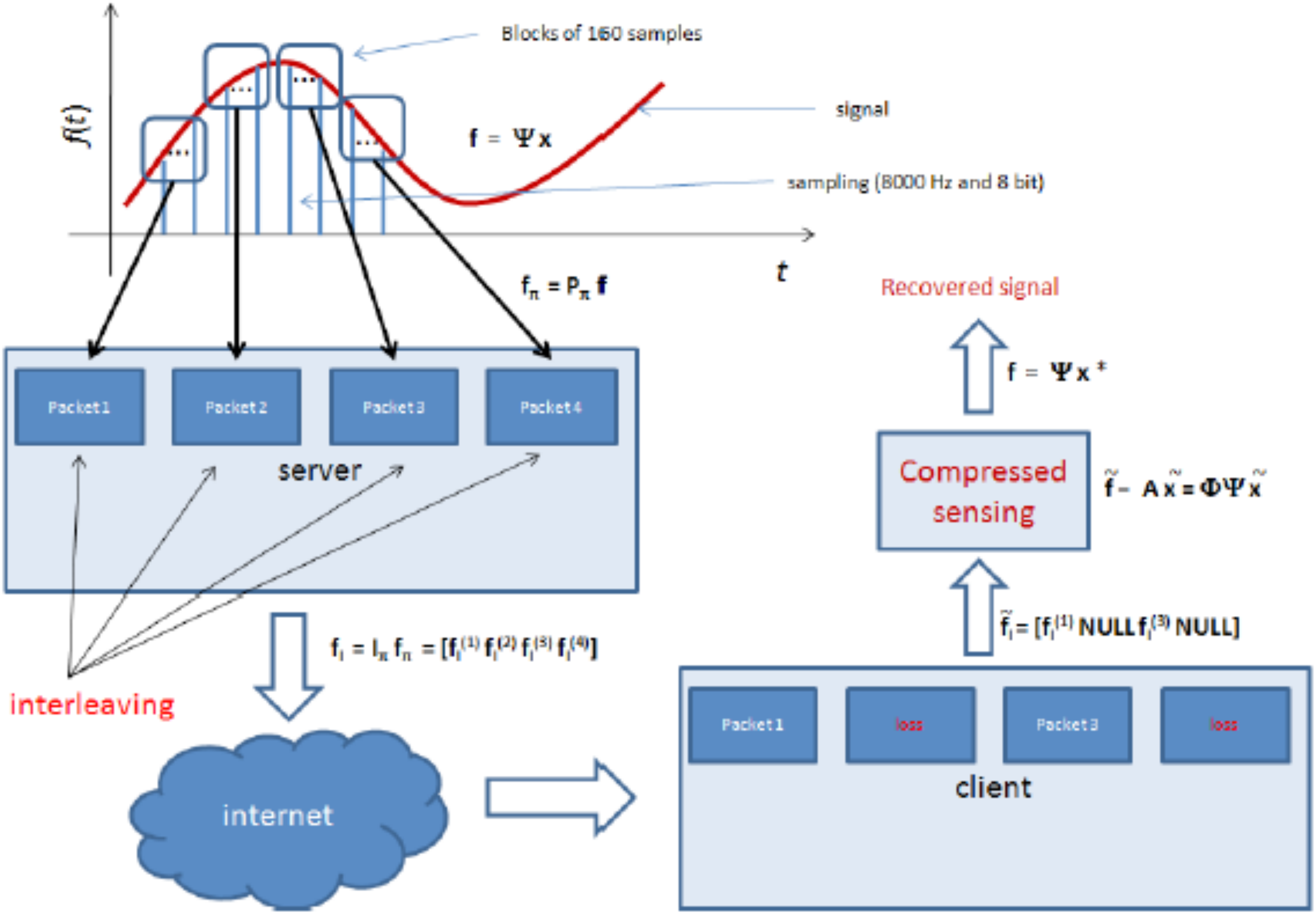
signal

sampling (8 KHz, 8 bit)

Interleaving



Reconstruction scheme



$f(t)$

Blocks of 160 samples

signal

$f = \Psi x$

sampling (8000 Hz and 8 bit)

t

$f_\pi = P_\pi f$

Packet 1

Packet 2

Packet 3

Packet 4

server

interleaving

$f_i = I_\pi f_\pi = [f_i^{(1)} f_i^{(2)} f_i^{(3)} f_i^{(4)}]$

internet

Packet 1

loss

Packet 3

loss

client

$\tilde{f}_i = [f_i^{(1)} \text{ NULL } f_i^{(3)} \text{ NULL}]$

Compressed sensing

$\tilde{f} - A\tilde{x} = \Phi\Psi\tilde{x}$

$f = \Psi x^*$

Recovered signal

Experimental results

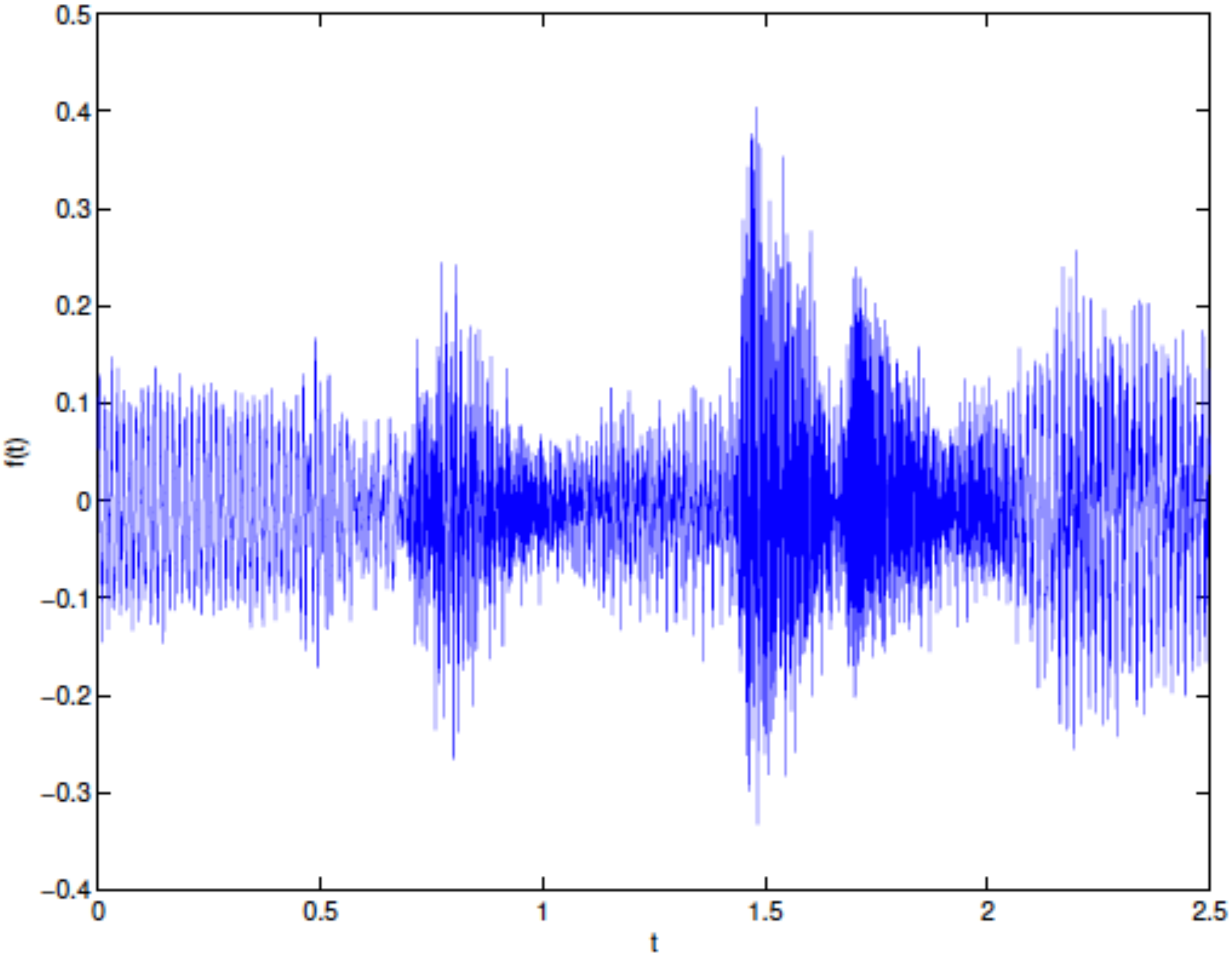


Fig. 4. Audio signal of a female speaker.

Experimental results

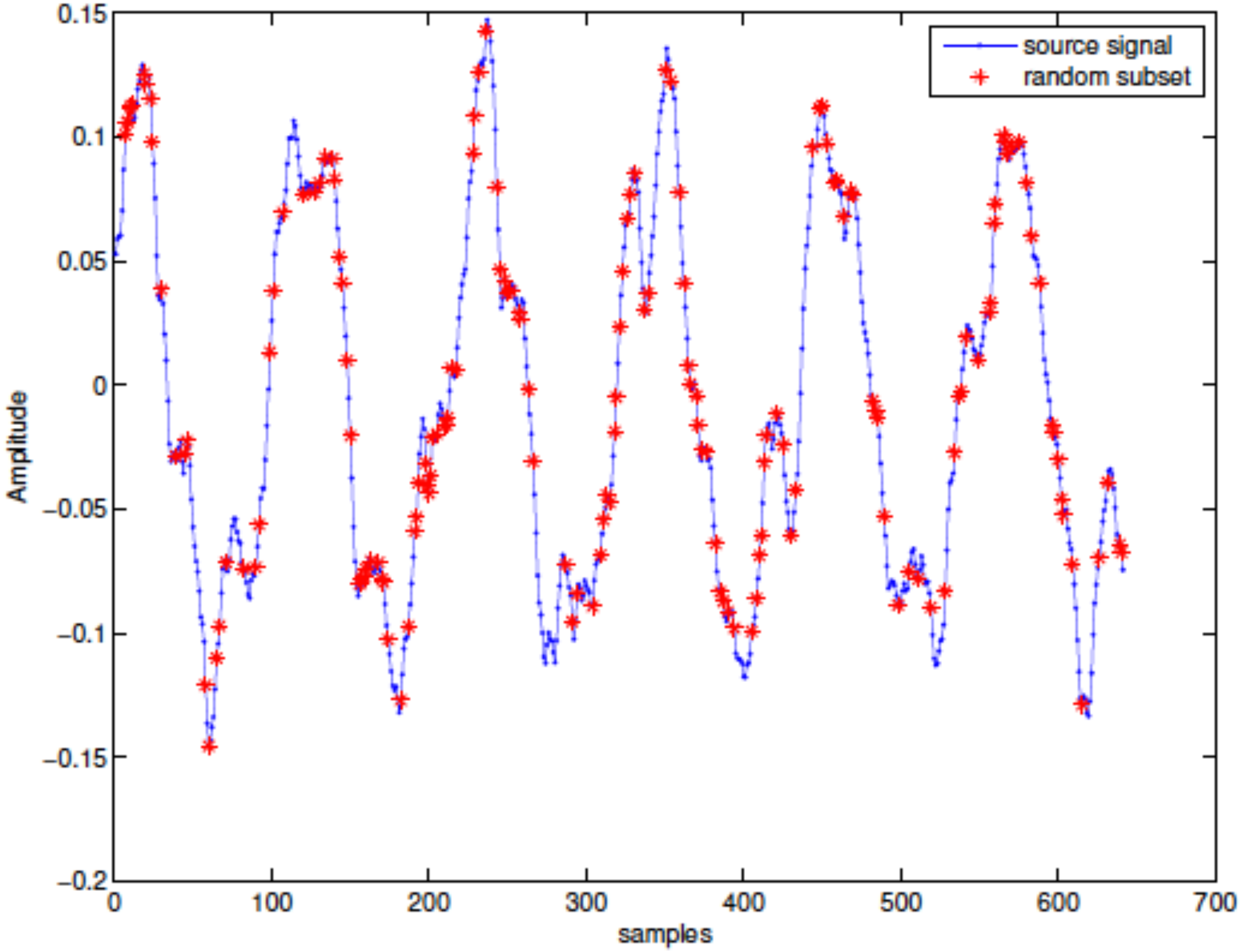


Fig. 5. Frame information after 3 packets lost.

Experimental results

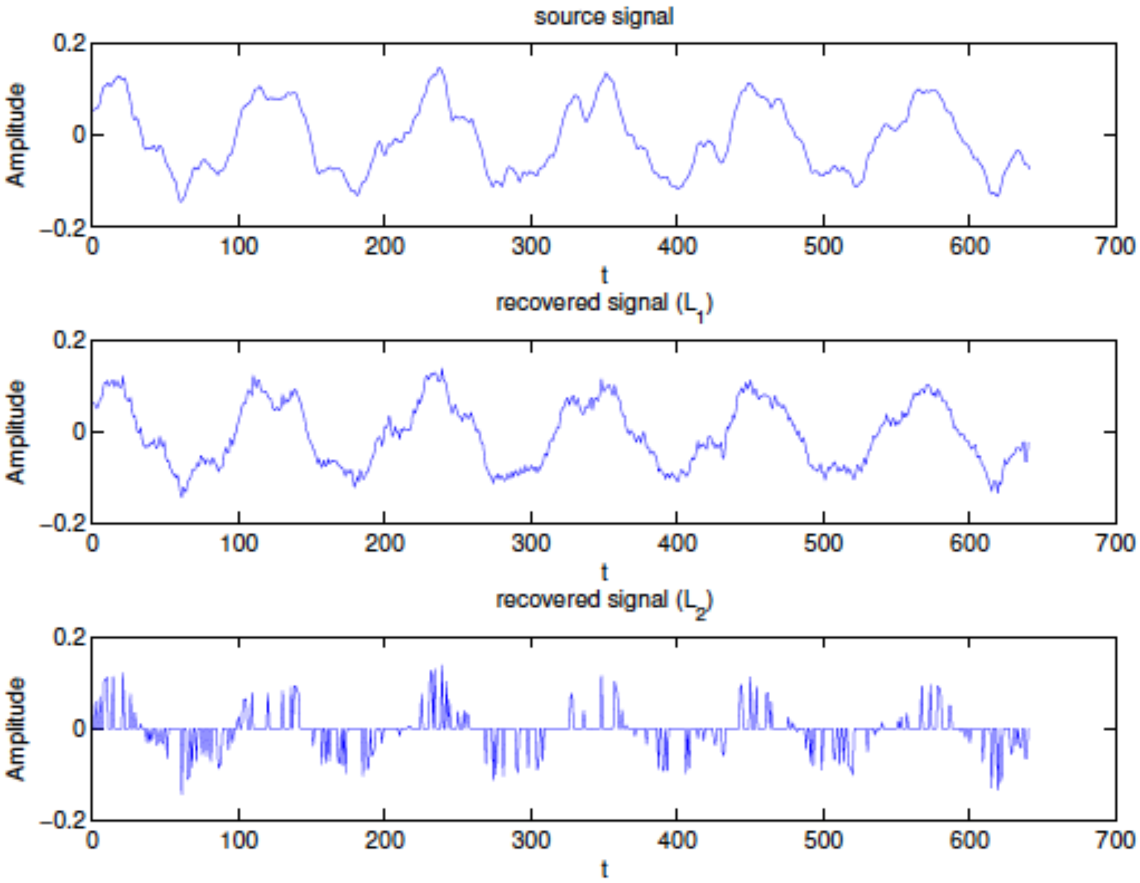


Fig. 6. Comparison between a frame of the source signal and those of the recovered signals by using L_1 and L_2 norms.

Experimental results

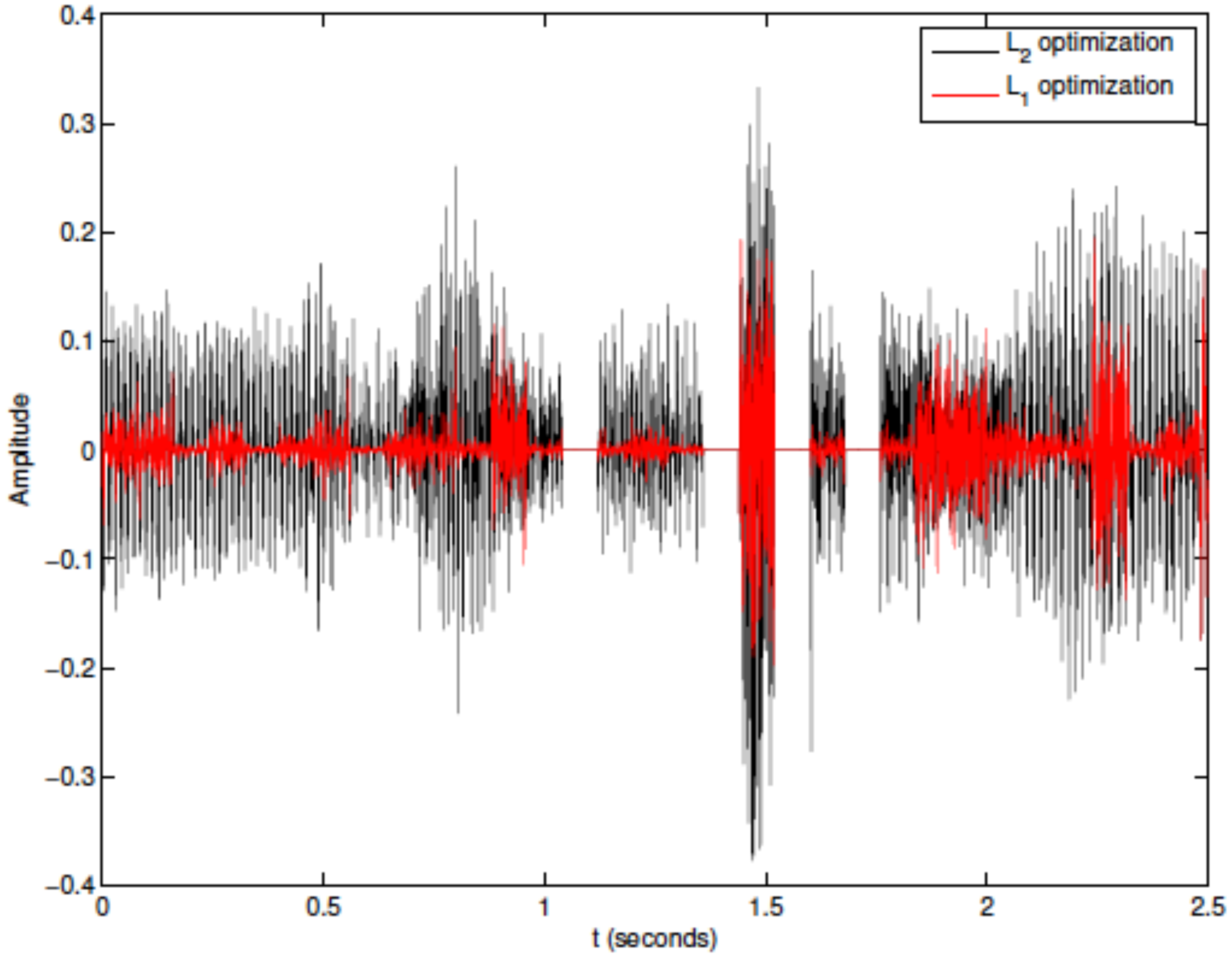


Fig. 7. Resisdua between the source signal and the recovered signals by using L_1 and L_2 norms.

Experimental results

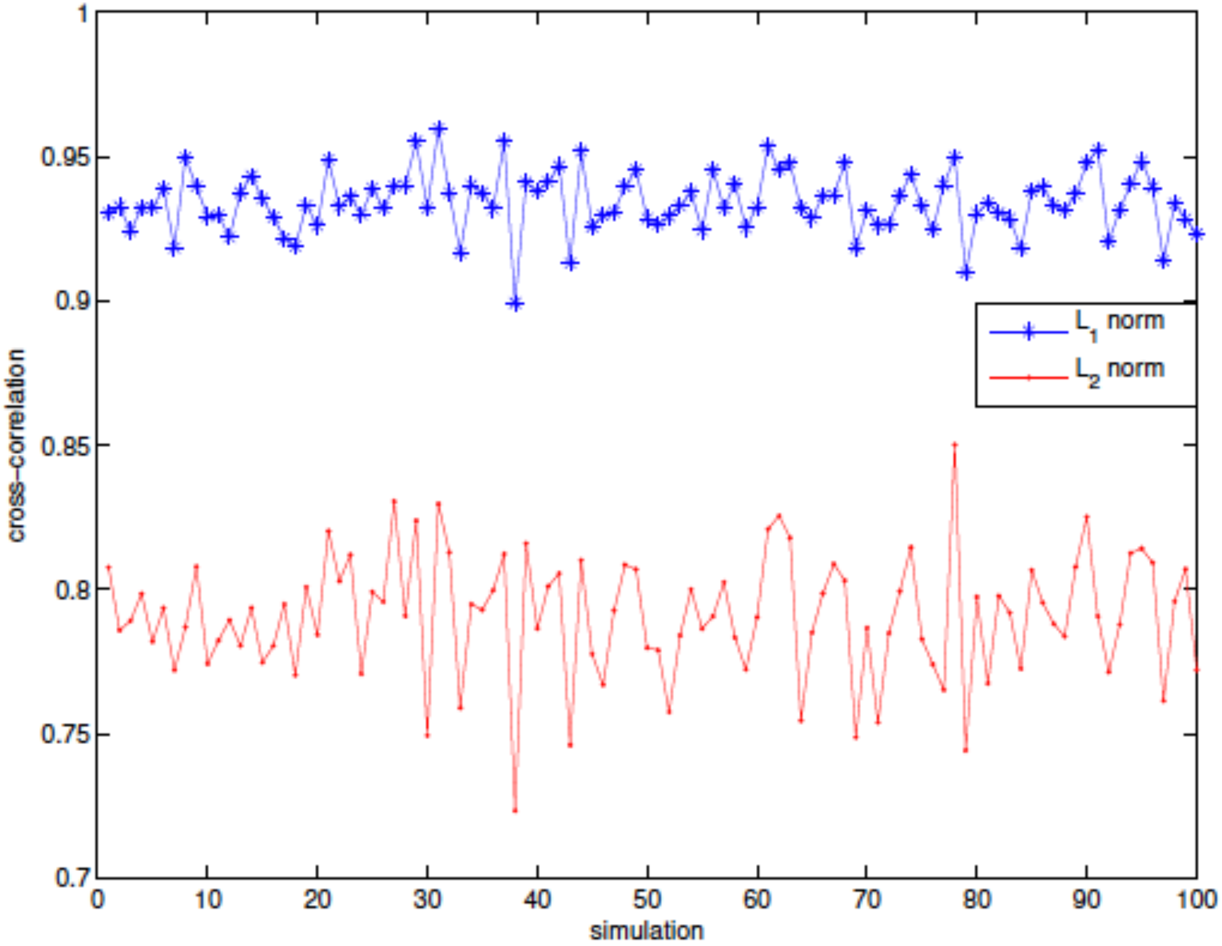


Fig. 8. Cross-correlation coefficients after 100 simulations: audio female speaker.

Experimental results

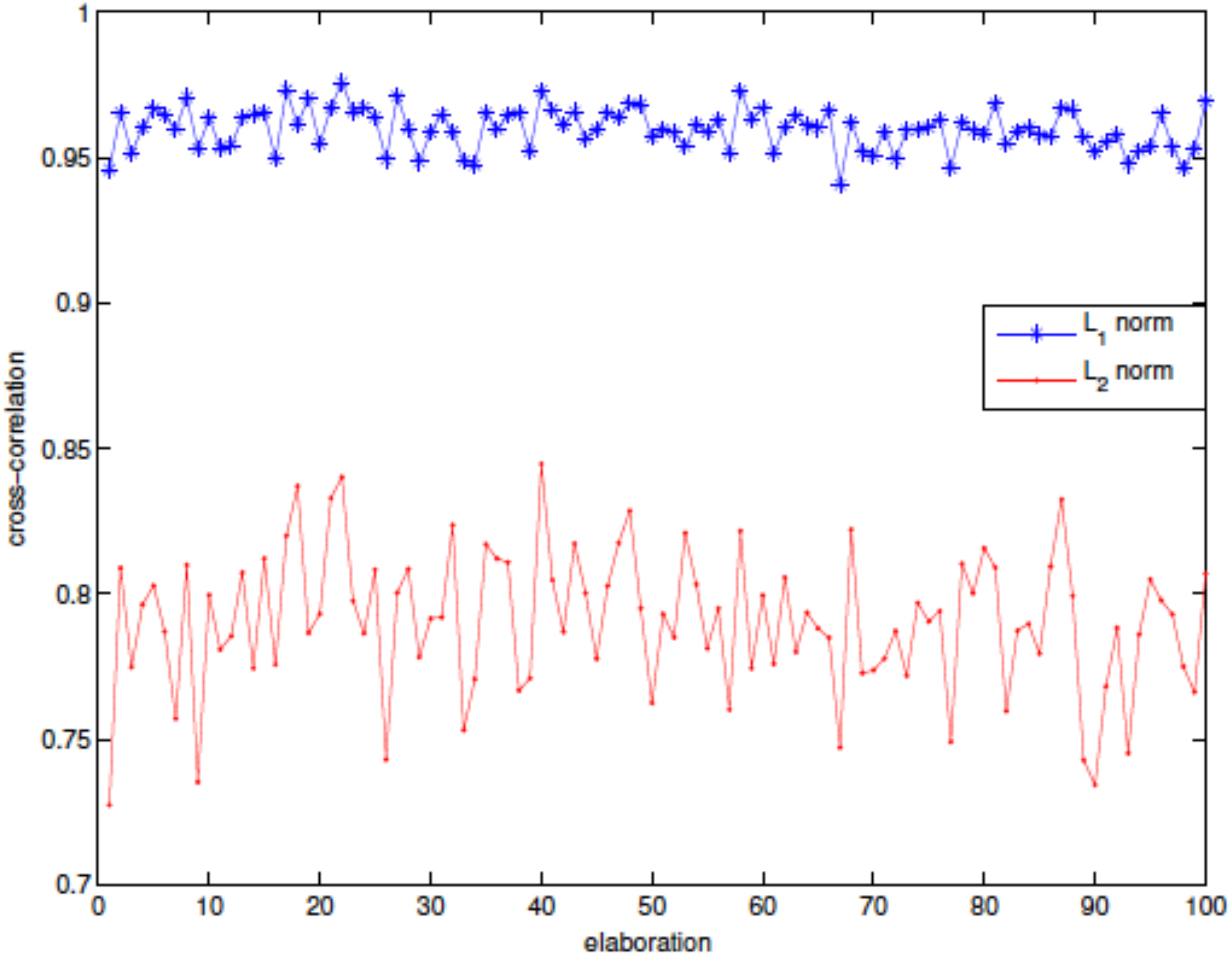


Fig. 9. Cross-correlation coefficients after 100 simulations: audio male speaker.

Experimental results

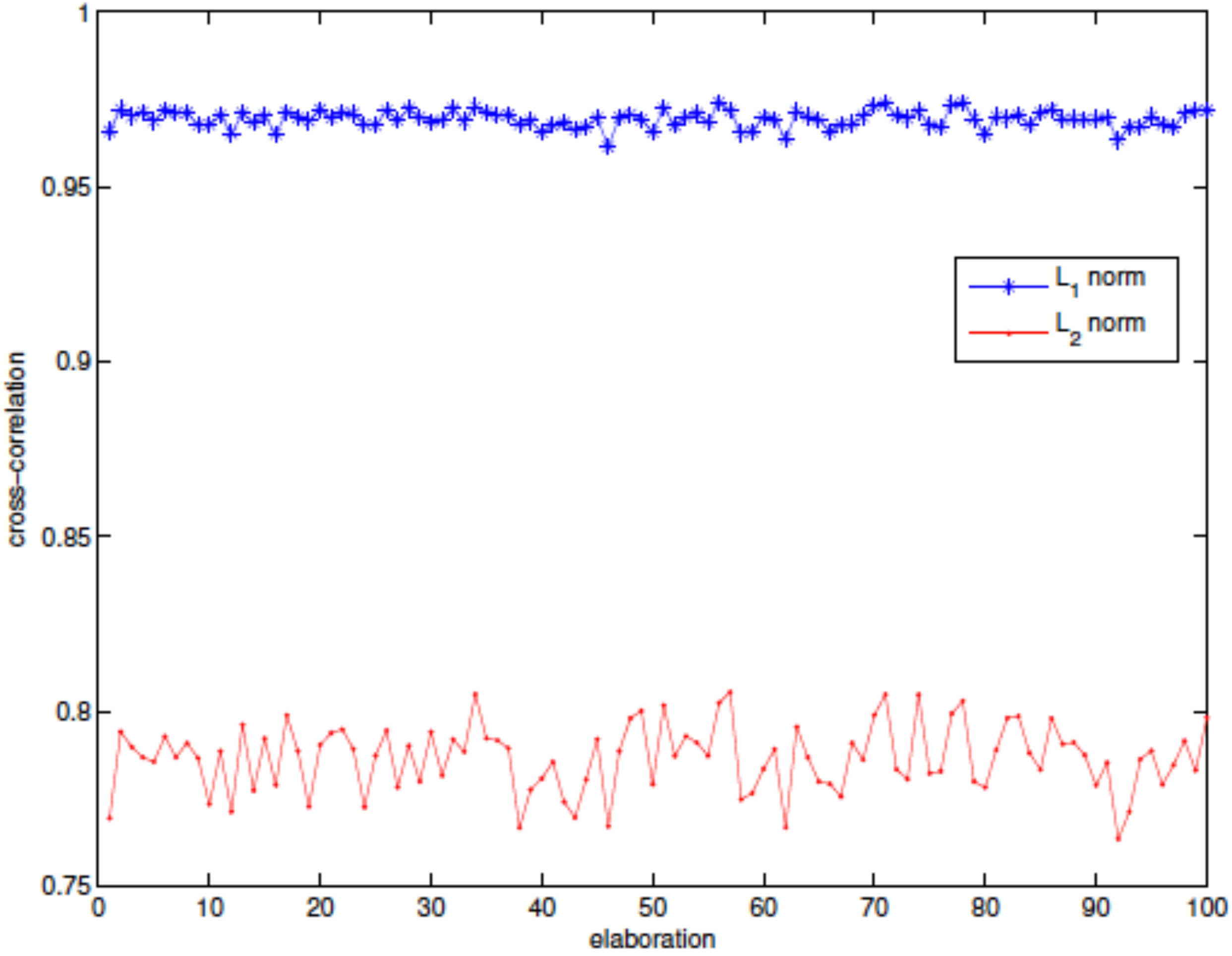
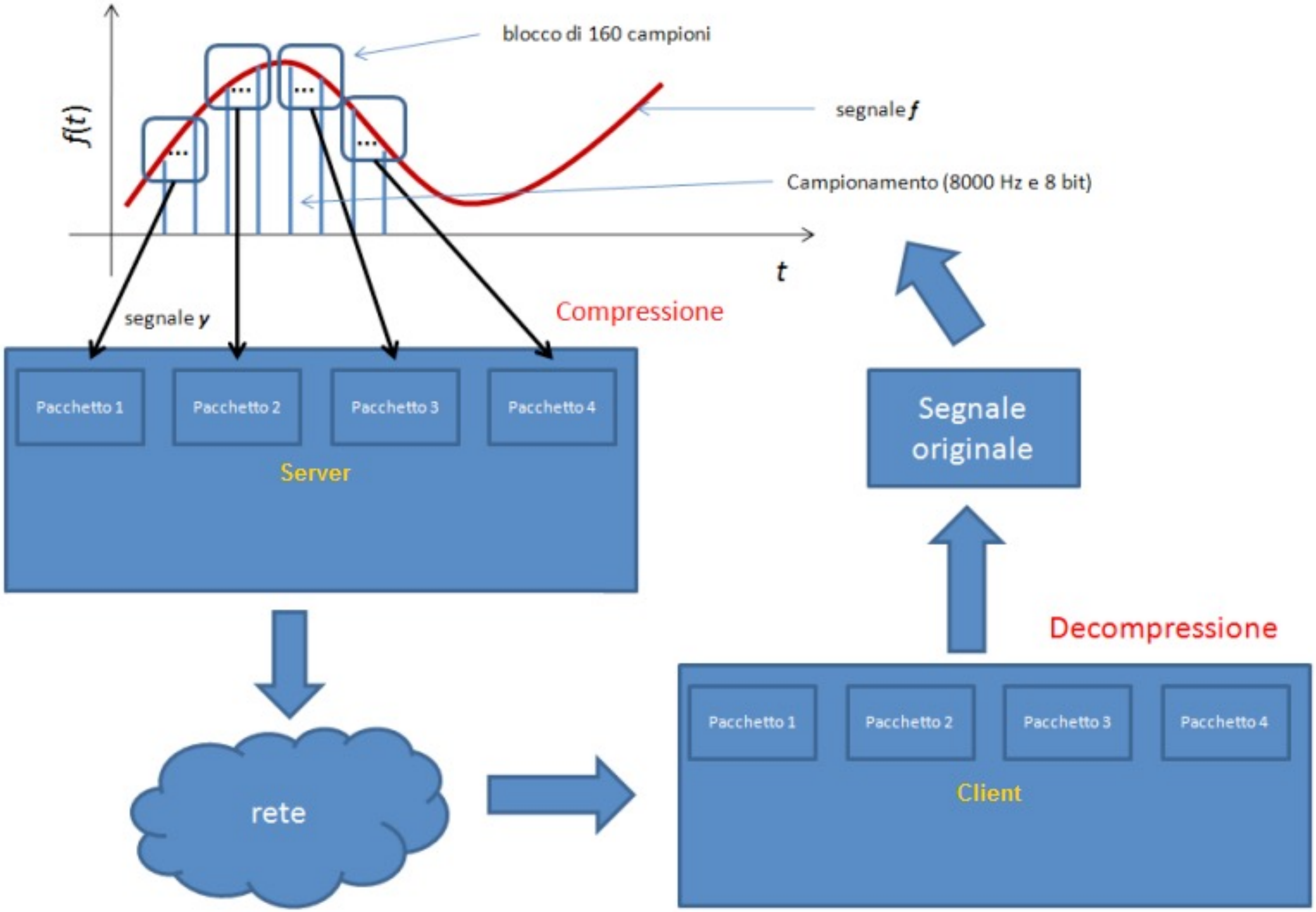


Fig. 10. Cross-correlation coefficients after 100 simulations: audio song.

Compression scheme



References

- Material
  - Slides
  - Video Lessons

- Books
  - A Mathematical Introduction to Compressive Sensing, S. Foucart , H. Rauhut, 2013

Question 33

- PCA can be defined as the principal subspace such that the variance of the projected data is maximized

- Question
  - Describe the basis of PCA

Principal Component Analysis

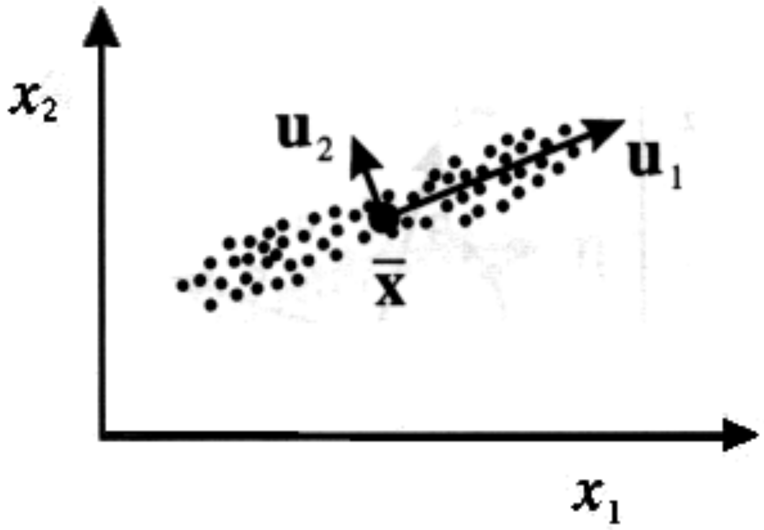
- Principal Component Analysis (PCA) is a statistical technique
  - Dimensionality reduction
  - Lossy data compression
  - Feature extraction
  - Data visualization

- It is also known as the *Karhunen-Loeve* transform

Second-Order methods

- The second-order methods are the most popular methods to find a linear transformation
- This methods find the representation using only the information contained in the covariance matrix of the data vector **x**
- PCA is widely used in signal processing, statistics, and neural computing

Principal Components



In a linear projection down to one dimension, the optimum choice of projection, in the sense of minimizing the sum-of-squares error, is obtained first subtracting off the mean of the data set, and then projecting onto the first eigenvector $\mathbf{u}_1$ of the covariance matrix.

Projection error minimization

- We introduce a complete orthonormal set of *D*-dimensional basis vectors (i=1,…,*D*)

$$\boxed{\mathbf{u}_i^T \mathbf{u}_j = \delta_{ij}}$$

- Because this basis is complete, each data point can be represented by a linear combination of the basis vectors

$$\boxed{\mathbf{x}_n = \sum_{i=1}^{D} \alpha_{ni} \mathbf{u}_i}$$

Projection error minimization

- We can write also that

$$\boxed{\mathbf{x}_n = \sum_{i=1}^{D} \left(\mathbf{x}_n^T \mathbf{u}_i\right) \mathbf{u}_i} \longleftarrow \boxed{\alpha_{nj} = \mathbf{x}_n^T \mathbf{u}_j}$$

- Our goal is to approximate this data point using a representation involving a restricted number *M* < *D* of variables corresponding to a projection onto a lower-dimensional subspace

$$\boxed{\widetilde{\mathbf{x}}_n = \sum_{i=1}^{M} z_{ni} \mathbf{u}_i + \sum_{i=M+1}^{D} b_i \mathbf{u}_i}$$

Projection error minimization

- As our distortion measure we shall use the squared distance between the original point and its approximation averaged over the data set so that our goal is to minimize

$$\boxed{J = \frac{1}{N}\sum_{n=1}^{N}\left\|\mathbf{x}_n - \tilde{\mathbf{x}}_n\right\|^2}$$

- The general solution is obtained by choosing the basis to be eigenvectors of the covariance matrix given by

$$\boxed{\mathbf{S}\mathbf{u}_i = \lambda_i \mathbf{u}_i}$$

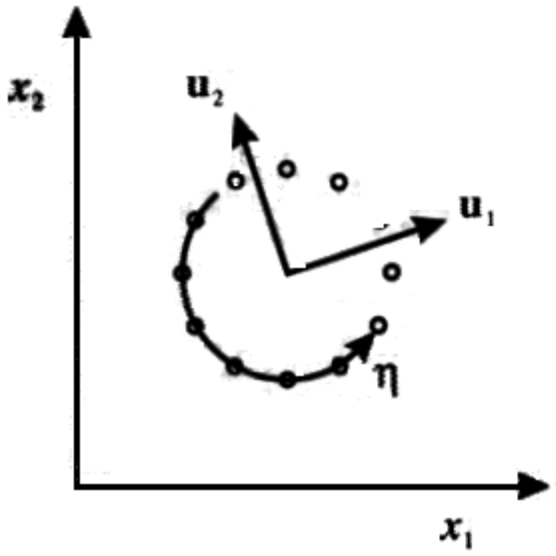
Projection error minimization

- The corresponding value of the distortion measure is then given by

$$\boxed{J = \sum_{i=M+1}^{D} \lambda_i}$$

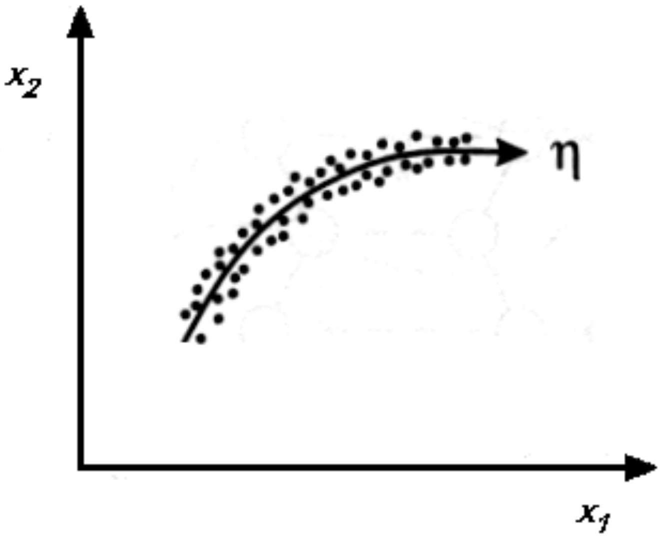
- We minimize this error selecting the eigenvectors defining the principal subspace are those corresponding to the *M* largest eigenvalues

Complex distributions



A linear dimensionality reduce technique, such as PCA, is unable to detect the lower dimensionality. In this case PCA gives two eigenvectors with equal eigenvalues. The data can described by a single eigenvalue

Addition of a small level of noise to data having an intrinsic. Dimensionality to 1 can increase its intrinsic dimensionality to 2. The data can be represented to a good approximation by a single variable η and can be regarded as having an intrinsic dimensionality of 1.



References

- Material
  - Slides
  - Video Lessons

- Books
  - Independent Component Analysis, A. Hyvärinen, J. Karhunen, E. Oja, John Wiley & Sons, 2001

Question 34

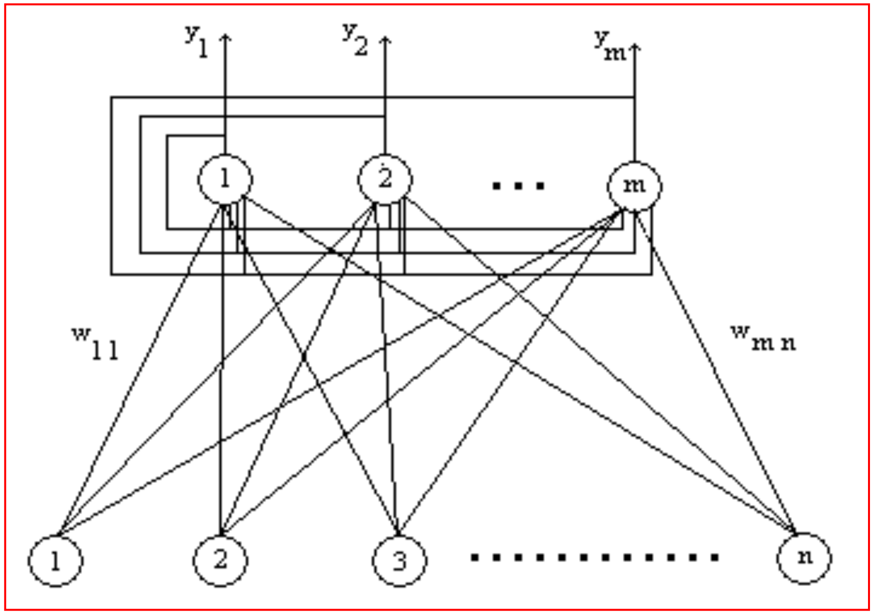
- PCA can be defined as the principal subspace such that the variance of the projected data is maximized

- Question
  - Describe the basis of non-linear PCA Neural Network

Unsupervised Neural Networks

- Typically Hebbian type learning rules are used
- There are two type of NN able to extract the Principal Components:
  - Symmetric (Oja, 1989)
  - Hierarchical (Sanger, 1989)

PCA and Unsupervised Neural Network

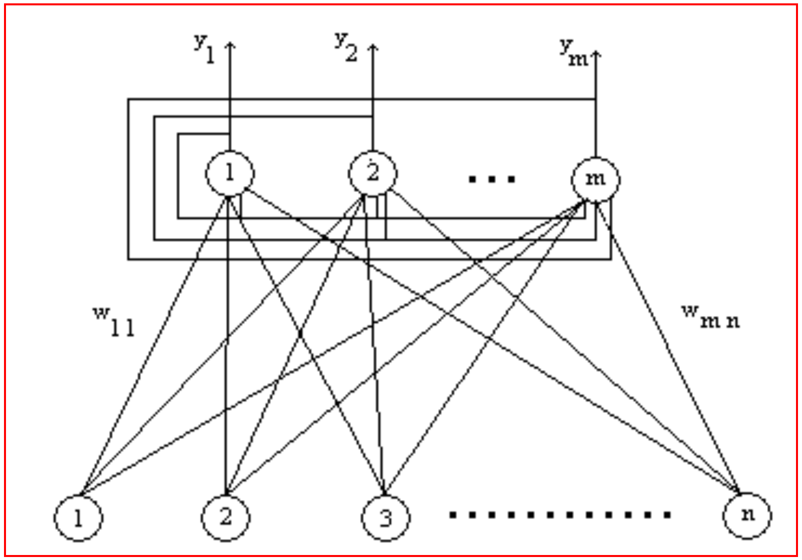


Symmetric PCA NN

Single layer Neural Network

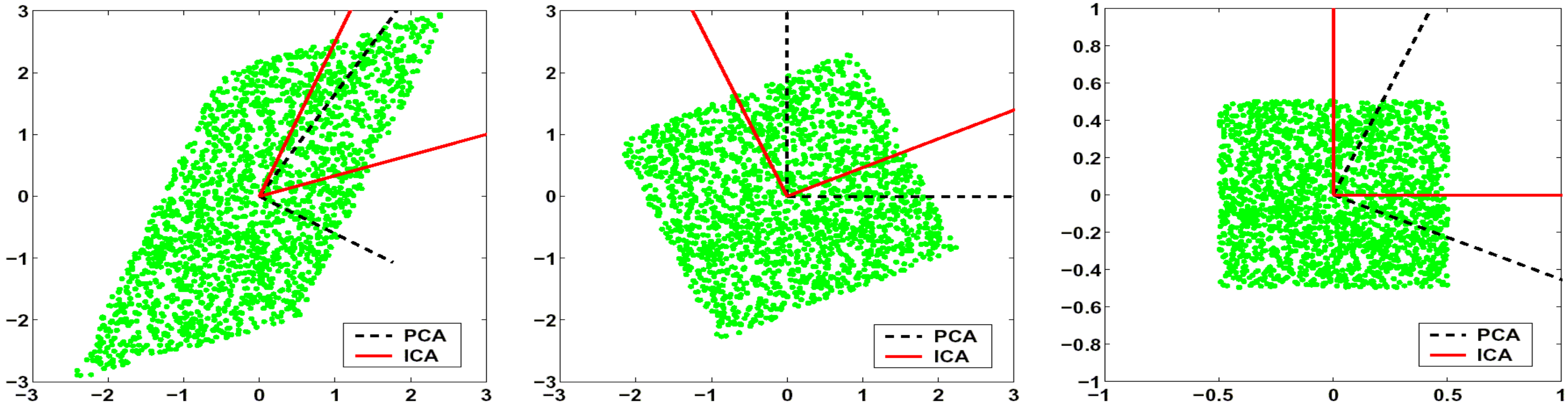
$$E\left[\mathbf{y}^2\right] = E\left[\left(\mathbf{w}^T\mathbf{x}\right)^2\right]$$

Objective function



Hierarchical PCA NN

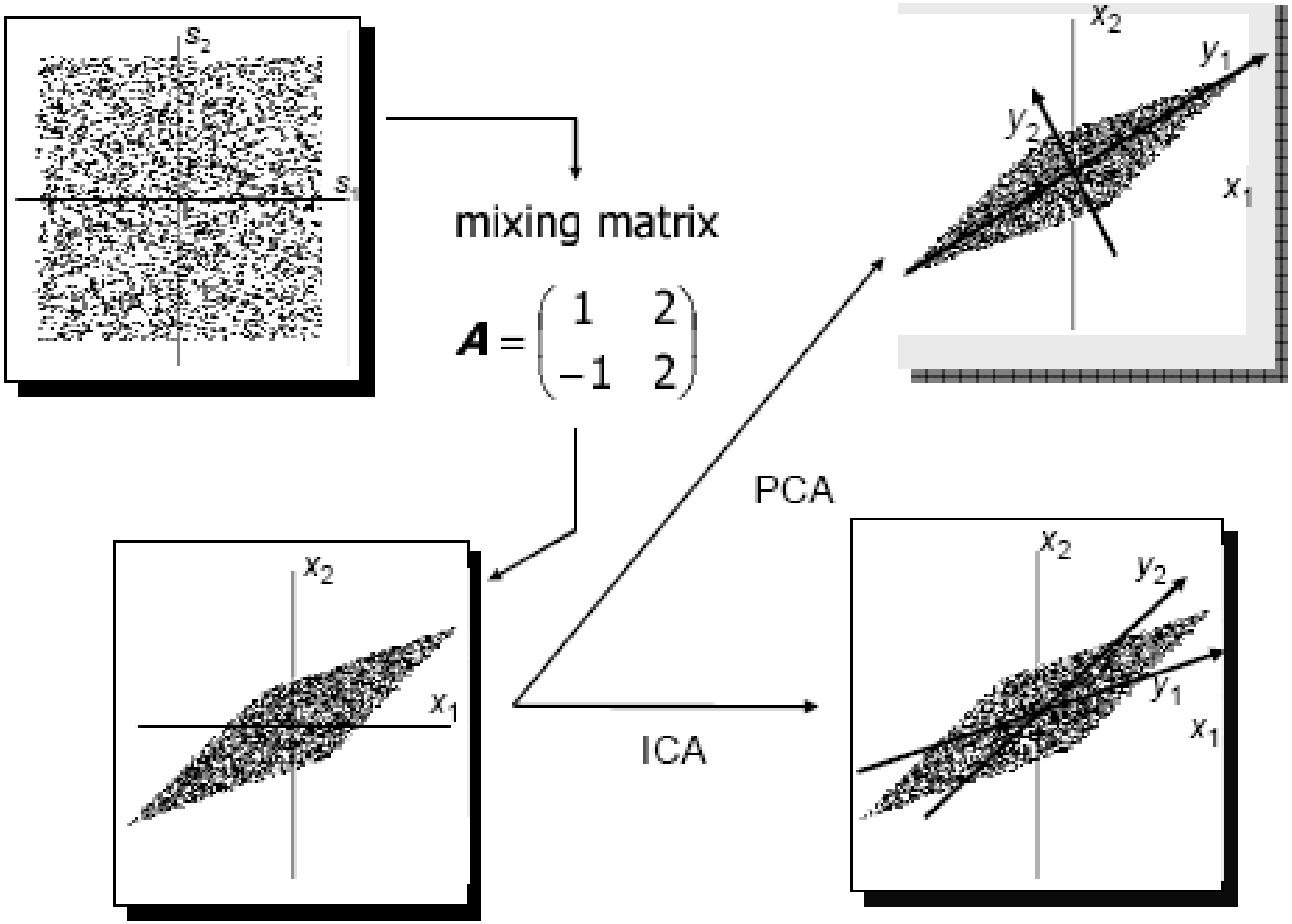
ICA versus PCA



(a) Original (b) After PCA pre-whitening (c) After ICA projection

PCA maximises the variance and projections onto the basis vectors are mixtures. ICA correctly finds the two vectors onto which the *projections are independent*.

Mixing matrix



mixing matrix

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix}$$

PCA

ICA

Non-linear objective function

Maximization

$$\mathbf{x} \xrightarrow{\mathbf{w}\text{ (weights)}} \boxed{E\left\{f\left(\mathbf{w}^T\mathbf{x}\right)\right\}}$$

L-dimensional vector

where E is the expectation with respect to the (unknown) density of **x** and $f(.)$ is a continue function (e.g. ln cosh(.))

Taylor series

$$\boxed{\ln\cosh(y) = \frac{1}{2}y^2 - \frac{1}{12}y^4 + \frac{1}{45}y^6 + O(y^8)}$$

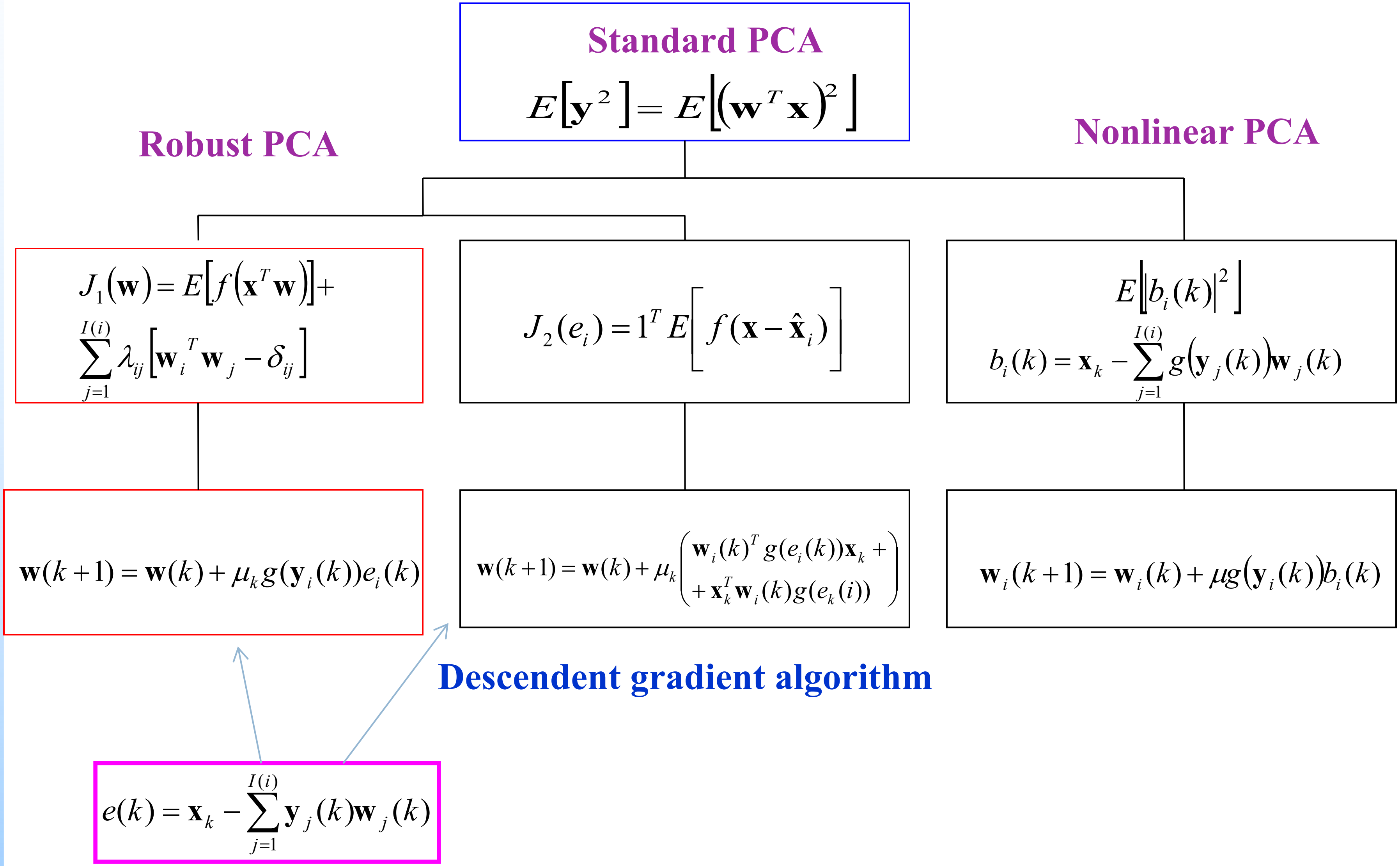
$$E\{\ln\cosh(y)\} = {}^1\!/_2\, E\left\{(w^T x)^2\right\} - {}^1\!/_{12}\, E\left\{(w^T x)^4\right\} + {}^1\!/_{45}\, E\left\{(w^T x)^6\right\} + E\left\{O((w^T x)^2)\right\}$$

$$\boxed{C = I \quad \text{and} \quad {}^1\!/_2\, E\left\{\left(w^T x\right)^2\right\} = {}^1\!/_2}$$

$$\boxed{-\frac{1}{12}E\left\{\left(w^T x\right)^4\right\}}$$

That is dominating, and the kurtosis is optimized

Unsupervised Neural Network



Standard PCA

$$E\left[\mathbf{y}^2\right] = E\left[\left(\mathbf{w}^T\mathbf{x}\right)^2\right]$$

Robust PCA

Nonlinear PCA

$$J_1(\mathbf{w}) = E\left[f\left(\mathbf{x}^T\mathbf{w}\right)\right] + \sum_{j=1}^{I(i)} \lambda_{ij}\left[\mathbf{w}_i{}^T\mathbf{w}_j - \delta_{ij}\right]$$

$$J_2(e_i) = 1^T E\left[f(\mathbf{x} - \hat{\mathbf{x}}_i)\right]$$

$$E\left[\left|b_i(k)\right|^2\right]$$

$$b_i(k) = \mathbf{x}_k - \sum_{j=1}^{I(i)} g\left(\mathbf{y}_j(k)\right)\mathbf{w}_j(k)$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu_k g(\mathbf{y}_i(k))e_i(k)$$

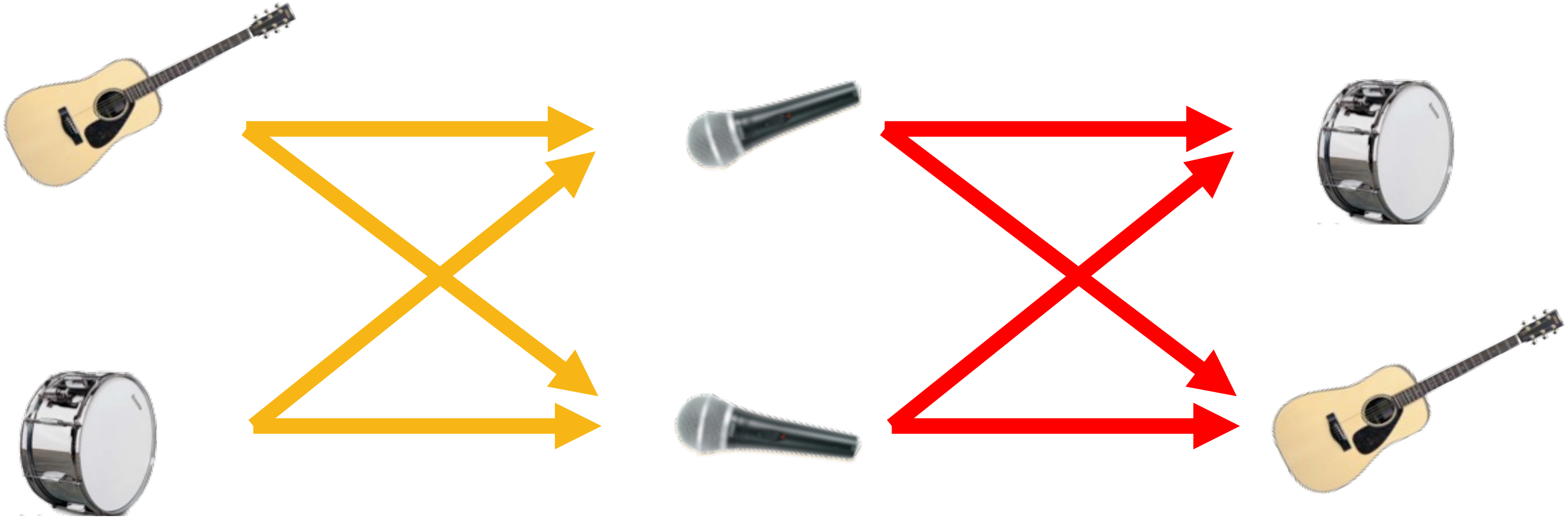
$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu_k \begin{pmatrix} \mathbf{w}_i(k)^T g(e_i(k))\mathbf{x}_k + \\ + \mathbf{x}_k^T\mathbf{w}_i(k)g(e_k(i)) \end{pmatrix}$$

$$\mathbf{w}_i(k+1) = \mathbf{w}_i(k) + \mu g\left(\mathbf{y}_i(k)\right)b_i(k)$$

Descendent gradient algorithm

$$e(k) = \mathbf{x}_k - \sum_{j=1}^{I(i)} \mathbf{y}_j(k)\mathbf{w}_j(k)$$

Cocktail party



Sources

Mixtures

Estimated-Sources

s

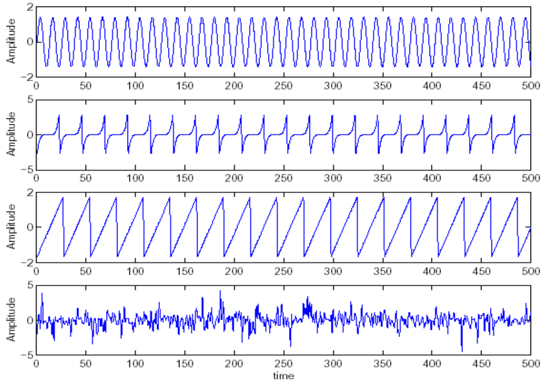
A

x

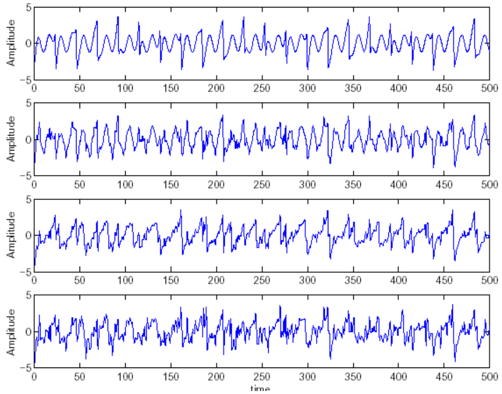
W

y

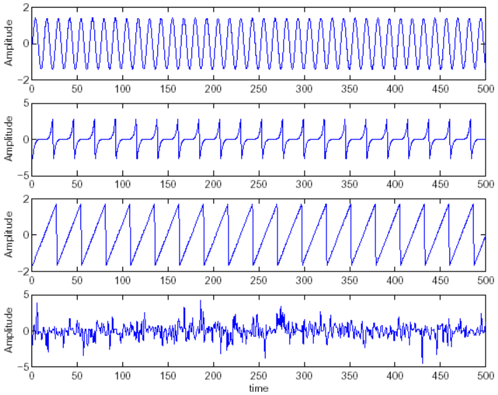
Source estimation



Source signals



Mixed signals



Estimated signals

$$\begin{array}{lllllll} x_1(t) & = & a_{11}s_1(t) & + & a_{12}s_2(t) & + & a_{13}s_3(t) \\ x_2(t) & = & a_{21}s_1(t) & + & a_{22}s_2(t) & + & a_{23}s_3(t) \\ x_3(t) & = & a_{31}s_1(t) & + & a_{32}s_2(t) & + & a_{33}s_3(t) \end{array}$$

$$\begin{array}{lllllll} y_1(t) & = & w_{11}x_1(t) & + & w_{12}x_2(t) & + & w_{13}x_3(t) \\ y_2(t) & = & w_{21}x_1(t) & + & w_{22}x_2(t) & + & w_{23}x_3(t) \\ y_3(t) & = & w_{31}x_1(t) & + & w_{32}x_2(t) & + & w_{33}x_3(t) \end{array}$$

x_1(t), $x_2(t)$,$x_3(t)$ are the observed signals, $s_1(t)$, $s_2(t)$, $s_3(t)$ the source signals

y_1(t), $y_2(t)$,$y_3(t)$ are the separated signals

References

- Material
  - Slides
  - Video Lessons
- Books
  - Independent Component Analysis, A. Hyvärinen, J. Karhunen, E. Oja, John Wiley & Sons, 2001

Question 35

- PCA can be defined as the principal subspace such that the variance of the projected data is maximized

- Question
  - Describe the basis of Independent Component Analysis

Independent Component Analysis

- Independent Component Analysis (ICA)
  - statistical and computational technique for revealing hidden factors that underlie sets of random variables, measurements, or signals

- ICA can be seen an extension of Principal Component Analysis (PCA) and Factor Analysis (FA)

- The technique of ICA was firstly introduced in early 1980s in the context of the Neural Networks (NNs) modeling

- ICA is becoming one of the exciting new topics, both in the field of NNs, mainly unsupervised learning, and in advanced statistics and signal processing

Probability distributions and densities

- random variable (rv) or stochastic variable is a variable whose value results from a measurement on some type of random process

- The cumulative distribution function (cdf) F_x of a random variable x at point $x = x_0$ is defined as the probability

$$\boxed{F_x(x_0) = P(x \leq x_0)}$$

- For continuous rv the cdf is a nonnegative, nondecreasing continuous function

$$\boxed{0 \leq F_x(x_0) \leq 1}$$

Probability distributions and densities

- The probability density function (pdf) $p_x(x)$ is obtained as the derivative of its cumulative distribution function

$$\boxed{p_x(x_0) = \left.\frac{dF_x(x)}{dx}\right|_{x=x_0}}$$

- The cdf is computed by using

$$\boxed{F_x(x_0) = \int_{-\infty}^{x_0} p_x(\xi)d\xi}$$

Distribution of a random vector

- Assume now that **x** is a n-dimensional random vector of continuous random variables

$$\boxed{\mathbf{x} = \left(x_1, x_2, \ldots, x_n\right)^T}$$

- The cdf is computed by using

$$\boxed{F_{\mathbf{x}}(\mathbf{x_0}) = P(\mathbf{x} \le \mathbf{x_0})}$$

$$\boxed{p_{\mathbf{x}}(\mathbf{x_0}) = \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\cdots\frac{\partial}{\partial x_n}F_{\mathbf{x}}(\mathbf{x})\bigg|_{\mathbf{x}=\mathbf{x}_0}}$$

Joint and marginal distributions

- The cdf called the joint distribution function of **x** and **y** is

$$\boxed{F_{\mathbf{x},\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0) = P(\mathbf{x} \leq \mathbf{x}_0, \mathbf{y} \leq \mathbf{y}_0)}$$

- The joint density function $p_{\mathbf{x},\mathbf{y}}(\mathbf{x},\mathbf{y})$ is defined by differentiating the joint distribution function

- The marginal densities are (e.g. on **x**)

$$\boxed{p_{\mathbf{x}}(\mathbf{x}) = \int_{-\infty}^{\infty} p_{\mathbf{x},\mathbf{y}}(\mathbf{x},\eta)d\eta}$$

Expectation and moments

- Let **g**(**x**) denote any quantity derived from the random vector **x** the expectation of **g**(**x**) is

$$E\{\mathbf{g}(\mathbf{x})\} = \int_{-\infty}^{\infty} \mathbf{g}(\mathbf{x}) p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}$$

- Moments are expectations used to characterize a random vector. The mean vector is

$$\mathbf{m}_{\mathbf{x}} = E\{\mathbf{x}\} = \int_{-\infty}^{\infty} \mathbf{x} p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}$$

- The n x n correlation matrix is

$$\mathbf{R}_{\mathbf{x}} = E\{\mathbf{x}\mathbf{x}^T\} = \mathbf{C}_{\mathbf{x}} + \mathbf{m}_{\mathbf{x}}\mathbf{m}_{\mathbf{x}}^T$$

$$\mathbf{C}_{\mathbf{x}} = E\{(\mathbf{x}-\mathbf{m}_{\mathbf{x}})(\mathbf{x}-\mathbf{m}_{\mathbf{x}})^T\}$$

Covariance matrix

Uncorrelatedness and independence

- Two random vectors **x** and **y** are uncorrelated if their cross-covariance matrix is a zero matrix

$$\boxed{\mathbf{C}_{\mathbf{xy}} = E\left\{(\mathbf{x}-\mathbf{m}_{\mathbf{x}})(\mathbf{y}-\mathbf{m}_{\mathbf{y}})^T\right\} = 0}$$

- The rvs x and y are said independent if and only if

$$\boxed{p_{x,y}(x,y) = p_x(x)\,p_y(y)}$$

- For random vectors is

$$\boxed{p_{\mathbf{x},\mathbf{y},\mathbf{z},\ldots}(\mathbf{x},\mathbf{y},\mathbf{z},\ldots) = p_{\mathbf{x}}(\mathbf{x})\,p_{\mathbf{y}}(\mathbf{y})\,p_{\mathbf{z}}(\mathbf{z})\ldots}$$

- Uncorrelated Gaussian rvs are also independent. This property is not shared by other distributions in general

Higher-order statistics

- Consider a scalar rv x, the *j*-th moment is defined as (j=1,2,…)

$$\alpha_j = E\{x^j\} = \int_{-\infty}^{\infty} \xi^j p_x(\xi) d\xi$$

- The j-th central moment

$$\mu_j = E\{(x - \alpha_1)^j\} = \int_{-\infty}^{\infty} (\xi - m_x)^j p_x(\xi) d\xi$$

Skewness

- The third central moment is called the skewness (asymmetricity of the pdf)

$$\boxed{\mu_3 = E\left\{(x - m_x)^3\right\}}$$

- The 4-th moment and central moment are applied in ICA

Kurtosis

- Usually the 4-order statistic (i.e. cumulants) is employed and it is called Kurtosis

$$\boxed{\mathbf{kurt}(x) = E\left\{x^4\right\} - 3\left[E\left\{x^2\right\}\right]^2}$$

- A distribution having kurtosis
  - Zero is called mesocurtic
  - Negative platykurtic (subgaussian)
  - Positive leptokurtic (supergaussian)

Differential entropy

- The differential entropy of a rv is defined as

$$\boxed{H(x) = -\int p_x(\xi)\log p_x(\xi)d\xi = -E\{\log p_x(x)\}}$$

- Can be interpreted as a measure of randomness. If the rv is concentrated on certain small intervals, its differential entropy is small

Mutual Information

- Mutual information is a measure of the information that members of a set of random variables have on other random variables in the set

$$\boxed{I(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} H(x_i) - H(\mathbf{x})}$$

where **x** is the vector containing all the x_i

- If x_i are independent they give no information on each other

Kullback-Leibler divergence

- Mutual information can be considered a distance using the Kullback-Leibler divergence

$$\boxed{\delta(p^1, p^2) = \int p^1(\xi) \log \frac{p^1(\xi)}{p^2(\xi)} d\xi}$$

- Can be considered as a distance between pdfs
  - Is always nonnegative
  - Is zero if and only if the two distributions are equal
  - Can be symmetrized

Negentropy

- The Negentropy is a measure that is zero for a Gaussian variable and always nonnegative

$$\boxed{J(\mathbf{x}) = H(\mathbf{x}_{Gauss}) - H(\mathbf{x})}$$

- A simple approximation is (standardized rv)

$$\boxed{J(x) \approx \frac{1}{12} E\left\{x^3\right\}^2 + \frac{1}{48}\,\mathbf{kurt}(x)^2}$$

Negentropy

- A more robust approximation is

$$J(x) \approx k_1\left(E\left\{G^1(x)\right\}\right)^2 + k_2\left(E\left\{G^2(x)\right\} - E\left\{G^2(\nu)\right\}\right)^2$$

where k_1 and k_2 are positive constants, G^1 and G^2 are odd and even function, respectively (e.g. $G^1(x) = x^3$ and $G^2(x) = x^4$)

Newton's method

- Newton's method is one of the most efficient ways for function minimization $F(\mathbf{w})$

- The updating rule is (by using the gradient and the *Hessian*)

$$\boxed{\Delta\mathbf{w} = -\left[\frac{\partial^2 F(\mathbf{w})}{\partial \mathbf{w}^2}\right]^{-1} \frac{\partial F(\mathbf{w})}{\partial \mathbf{w}}}$$

- The convergence of the Newton's method is quadratic

The Lagrange method

- In many cases we have constrained optimizations

$$\boxed{\begin{array}{l} \min F(\mathbf{w}) \\ \text{subject to } H_i(\mathbf{w}) = 0, \quad i = 1, \ldots, k \end{array}}$$

- The most used way to take the constraints into account is the method of Lagrange multipliers $(\lambda_1, \ldots, \lambda_k)$

- We form the Lagrange function

$$\boxed{L(\mathbf{w}, \lambda_1, \ldots \lambda_k) = F(\mathbf{w}) + \sum_{i=1}^{k} \lambda_i H_i(\mathbf{w})} \Longrightarrow \boxed{\frac{\partial L(\mathbf{w}, \lambda_1, \ldots \lambda_k)}{\partial \mathbf{w}} = 0}$$

ICA

- We want to maximize the negentropy using this approximation

$$J_G(\mathbf{w}) = \left[E\left\{G(\mathbf{w}^T\mathbf{x})\right\} - E\left\{G(\nu)\right\}\right]^2$$

- The multi-unit problem is

$$\max \sum_{i=1}^{n} J_G(\mathbf{w}_i) \quad \mathbf{w}_i, i = 1,..,n$$

$$\text{such that } E\left\{(\mathbf{w}_k^T\mathbf{x})(\mathbf{w}_j^T\mathbf{x})\right\} = \delta_{jk}$$

- A fixed point algorithm is obtained by applying the Newton’s method to the Lagrangian of this optimization problem (FastICA)

ICA

- In many cases we have constrained optimizations

$$G(y) = \frac{1}{a_1} \log \cosh a_1 y$$

$$G(y) = -\exp(-y^2 / 2)$$

$$G(y) = y^4$$

$$g(y) = \tanh(a_1 y)$$

$$g(y) = y \exp(-y^2 / 2)$$

$$g(y) = y^3$$

Question 36

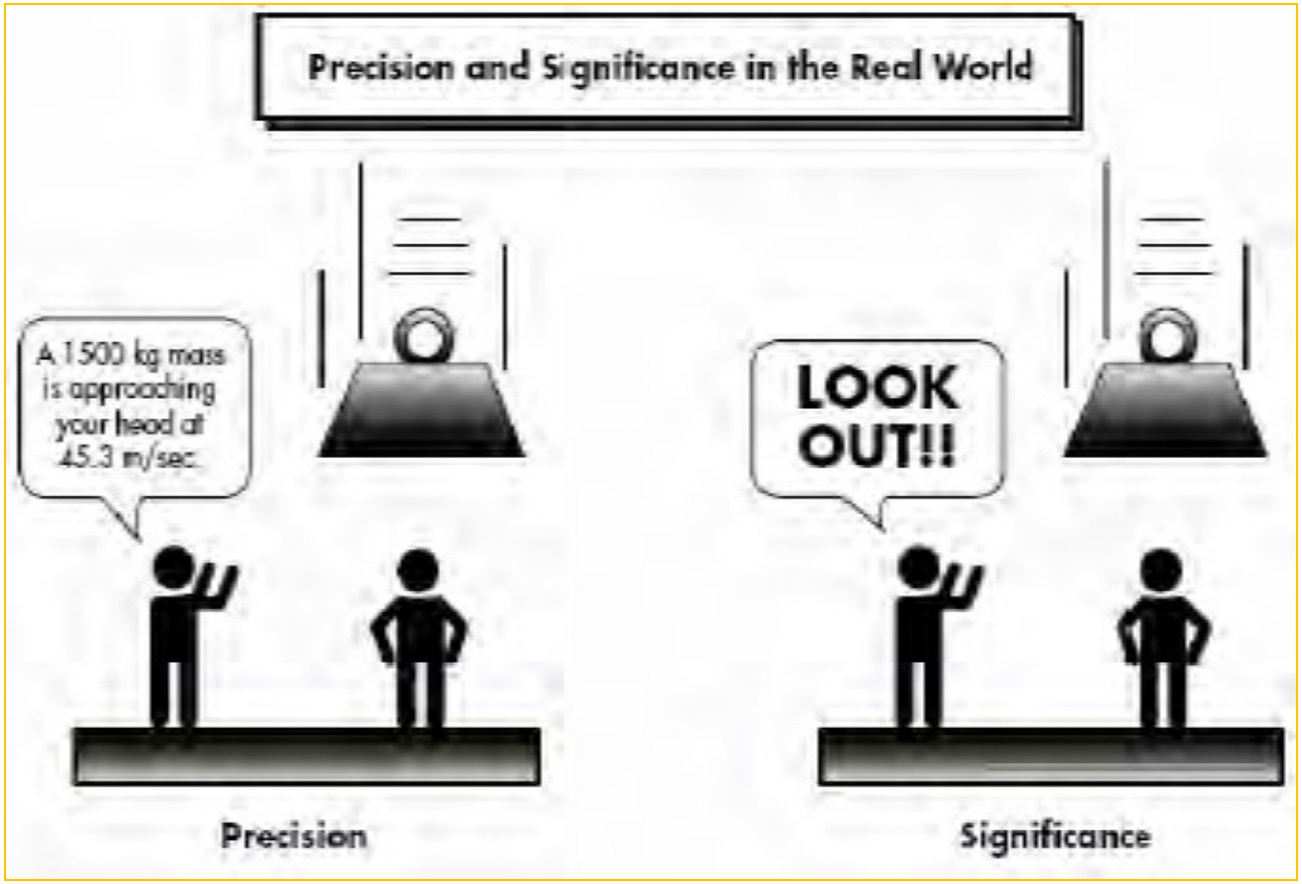
- PCA can be defined as the principal subspace such that the variance of the projected data is maximized

- Question
  - Describe the basis of Fuzzy Logic

Fuzzy Logic

- Fuzzy Logic is used to describe and operate with vague definitions
    - Example (control of a cement plant)
        - if the temperature is high add a little cement and increase the water a lot
- Fuzzy logic is a form of many-valued logic
    - the truth values of variables may be any real number between 0 and 1 inclusive

Meaning vs precision



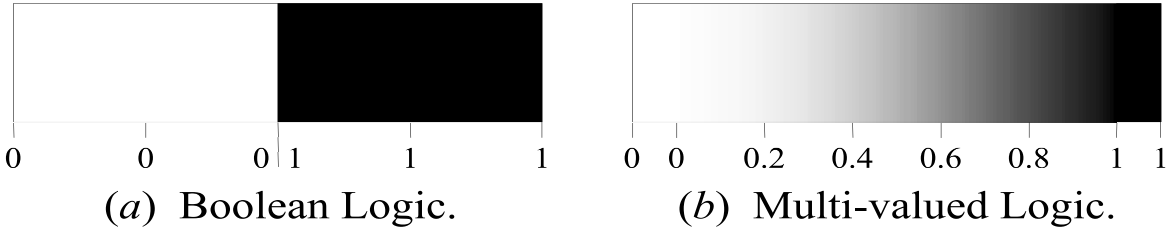
Difference between meaning and precision

In brief …

- Boolean logic
  - Boole (1854)
- Classical set theory (1900)
  - traditional sets (boolean belonging) and set operations
- Multivariate logic
  - Russell (1920)
  - Lukasiewicz (1930)
- Fuzzy Logic theory
  - Zadeh (1965)
  - extension of traditional sets (non boolean belonging) and operations on the elements
- Neutrosophic logic
  - Smarandache (1998)

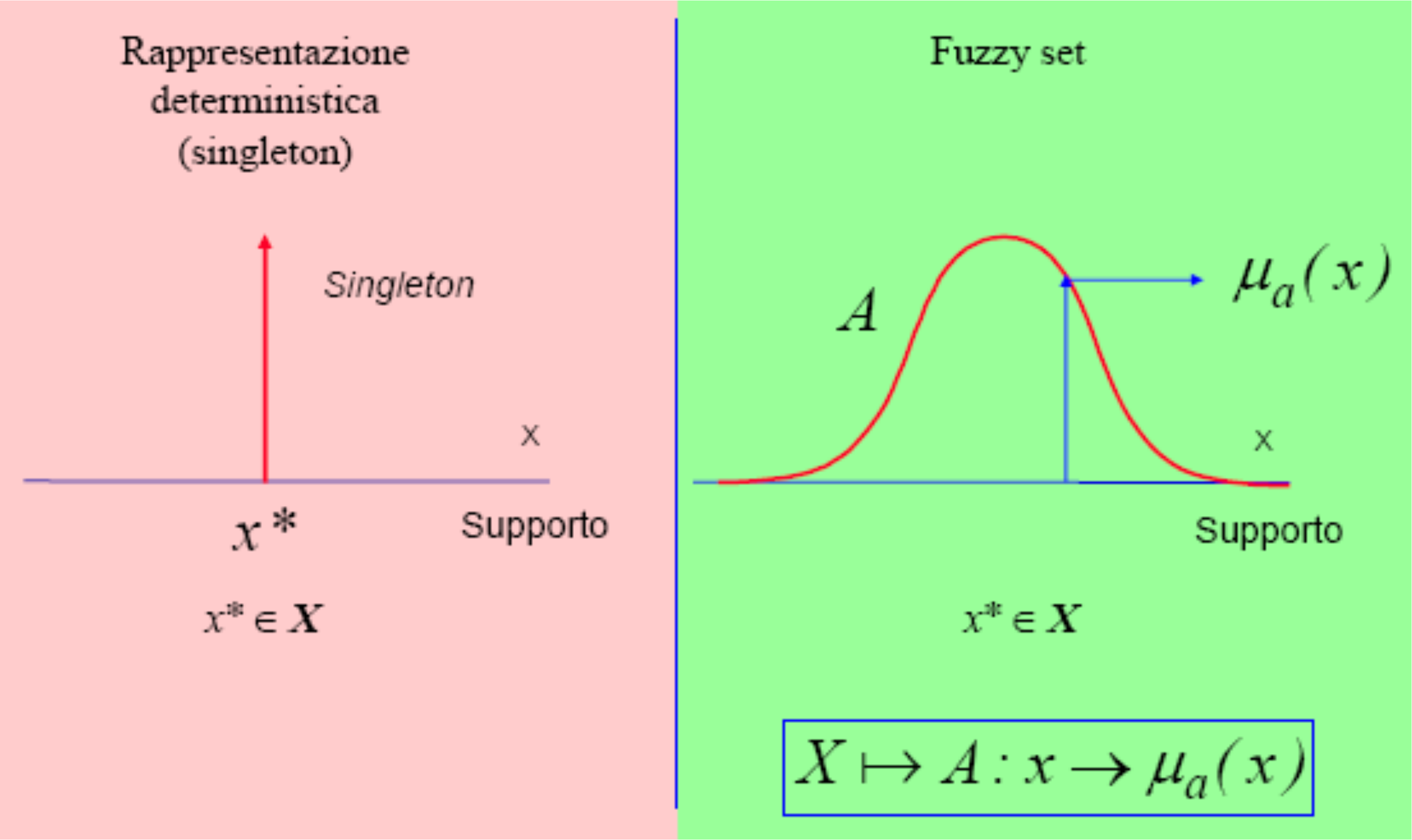
Crisp vs Fuzzy sets

- Fuzzy logic is a set of mathematical principles for representing knowledge based on the degree of belonging to a set



(*a*) Boolean Logic. (*b*) Multi-valued Logic.

Crisp vs Fuzzy



Rappresentazione deterministica (singleton)

$x^* \in \boldsymbol{X}$

Fuzzy set

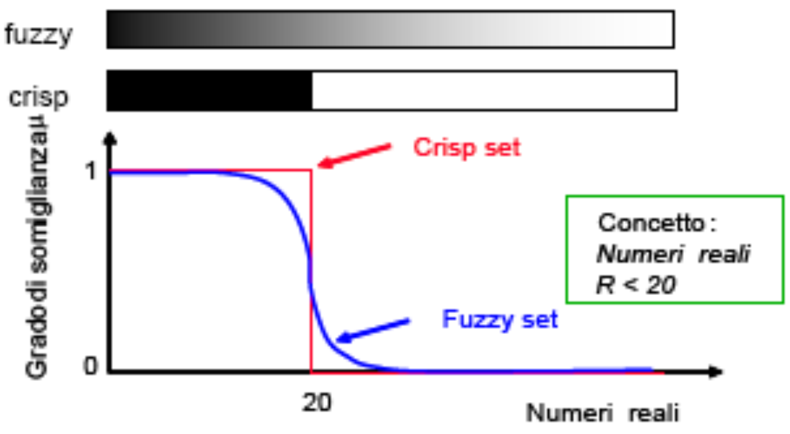
$x^* \in \boldsymbol{X}$

$$\boxed{X \mapsto A : x \to \mu_a(x)}$$

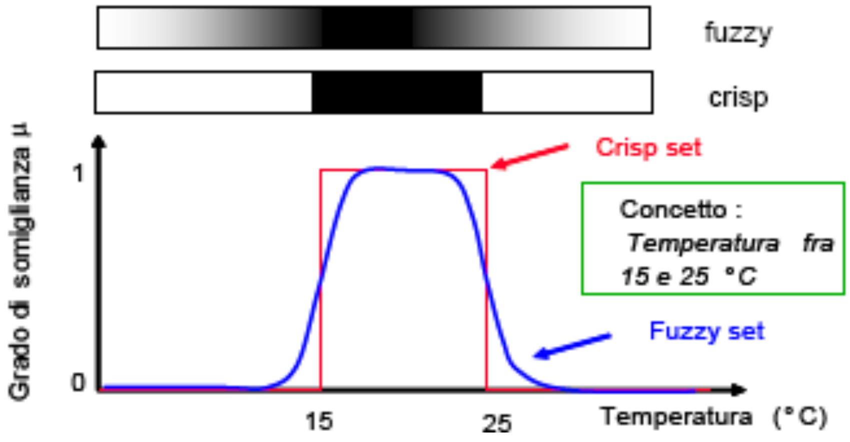
Linguistic variables

- A linguistic variable is a label that defines a concept
- This corresponds to a membership function (qualifier)
- It determines the degree of truth μ of any support value

Linguistic variables

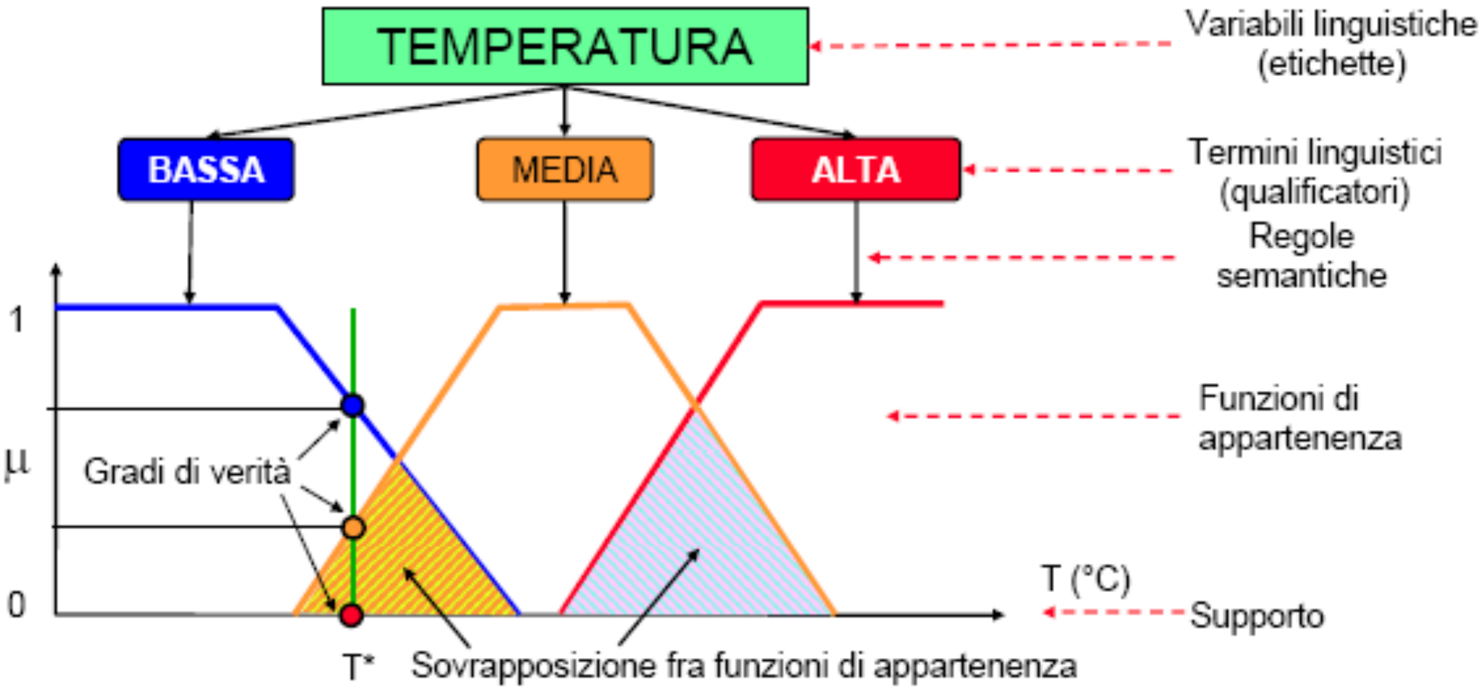


Numeri reali



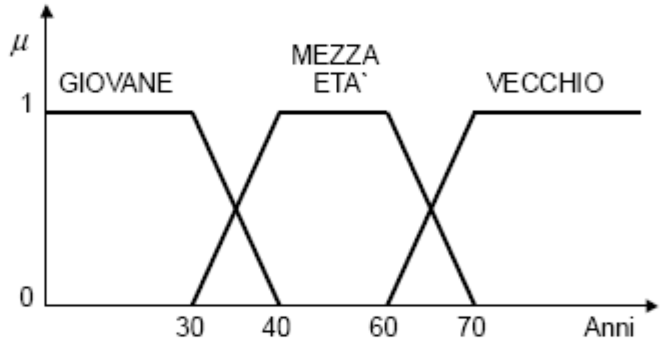
Temperatura

Linguistic variables

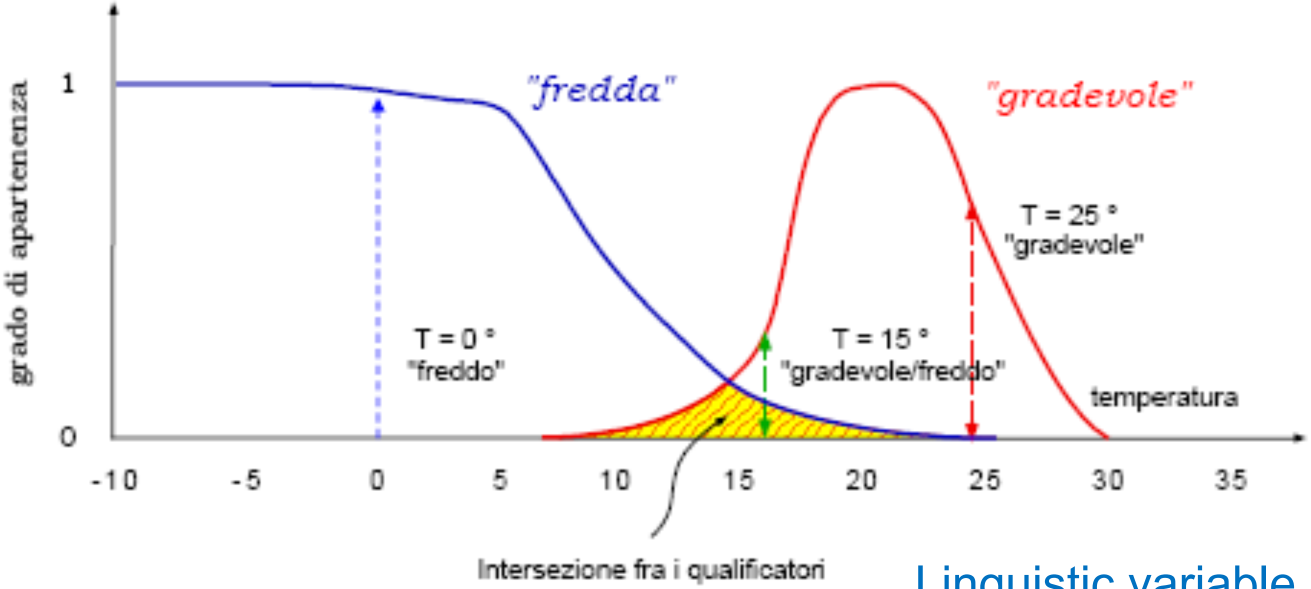


Esempio di fuzzificazione

Linguistic variables examples

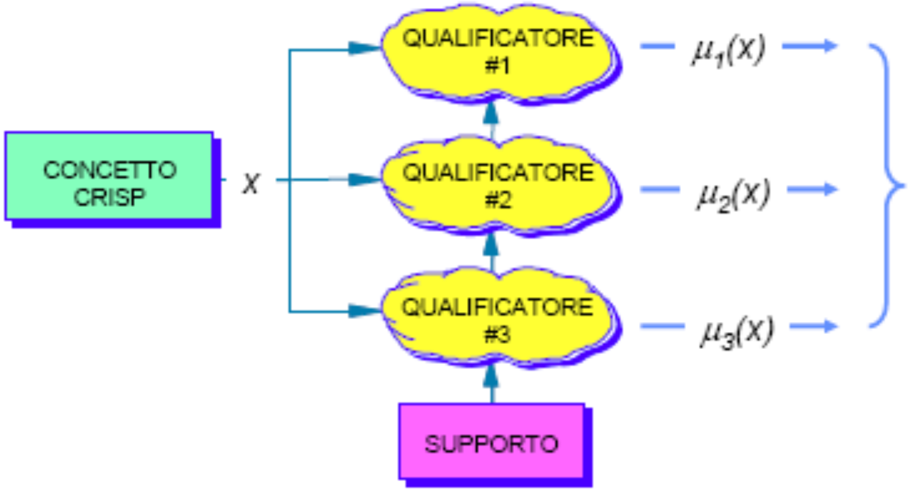


Linguistic variable “anni”



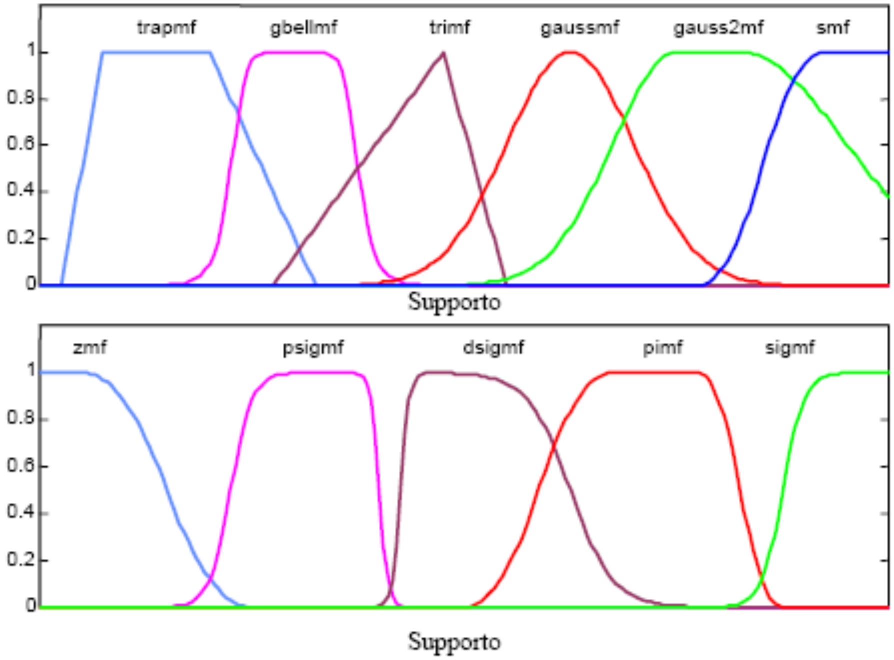
Linguistic variable “temperatura”

Fuzzification



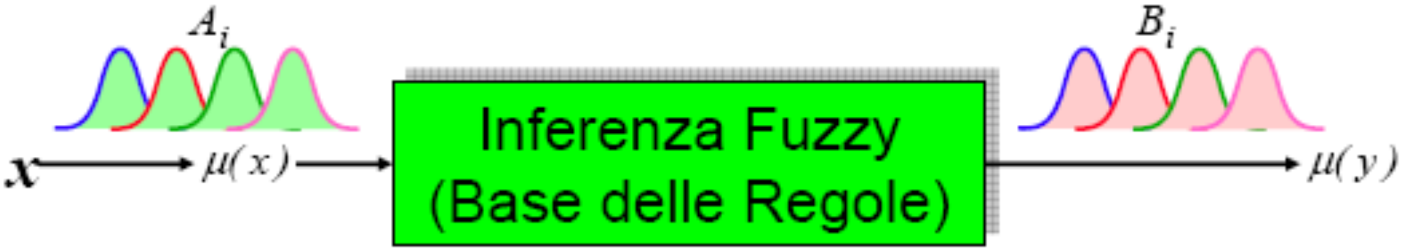
Fuzzification phase

Gradi di verità



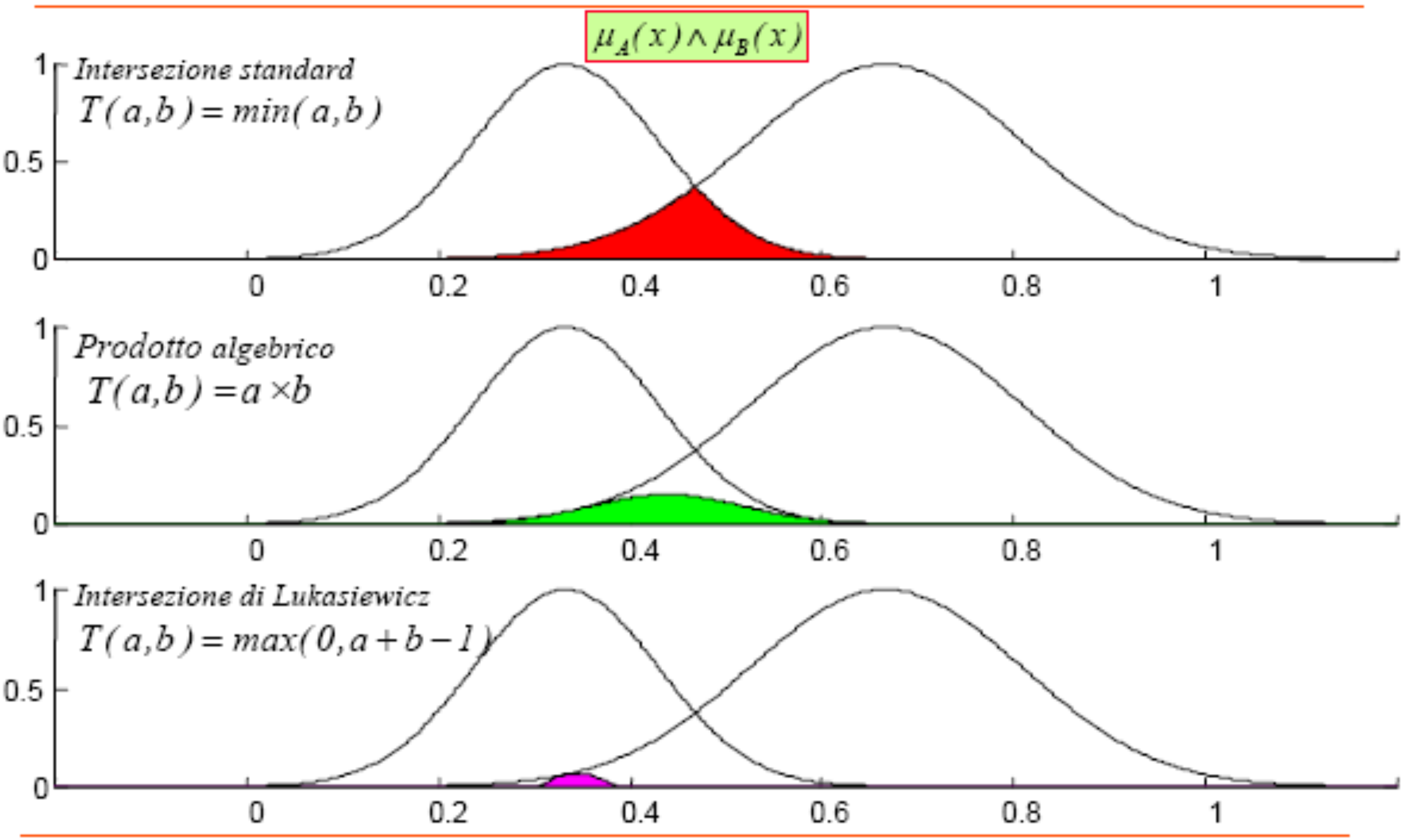
Kinds of memberships

Inference system



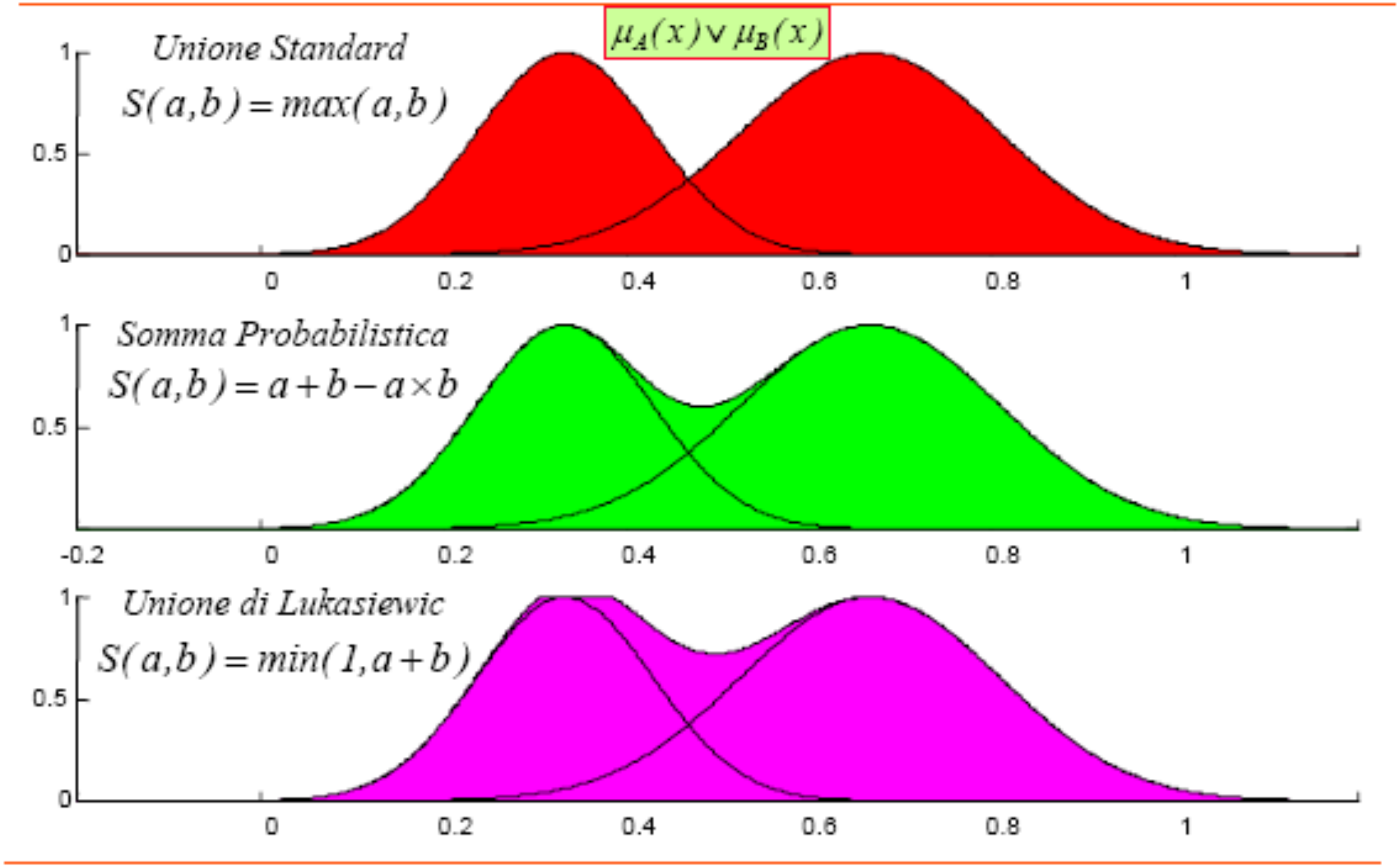
$$R_i : \ IF \underbrace{x_1 \ is \ A_1 \ AND \ x_2 \ is \ A_2}_{antecedente} THEN \underbrace{y \ is \ B}_{conseguente}$$

Operators



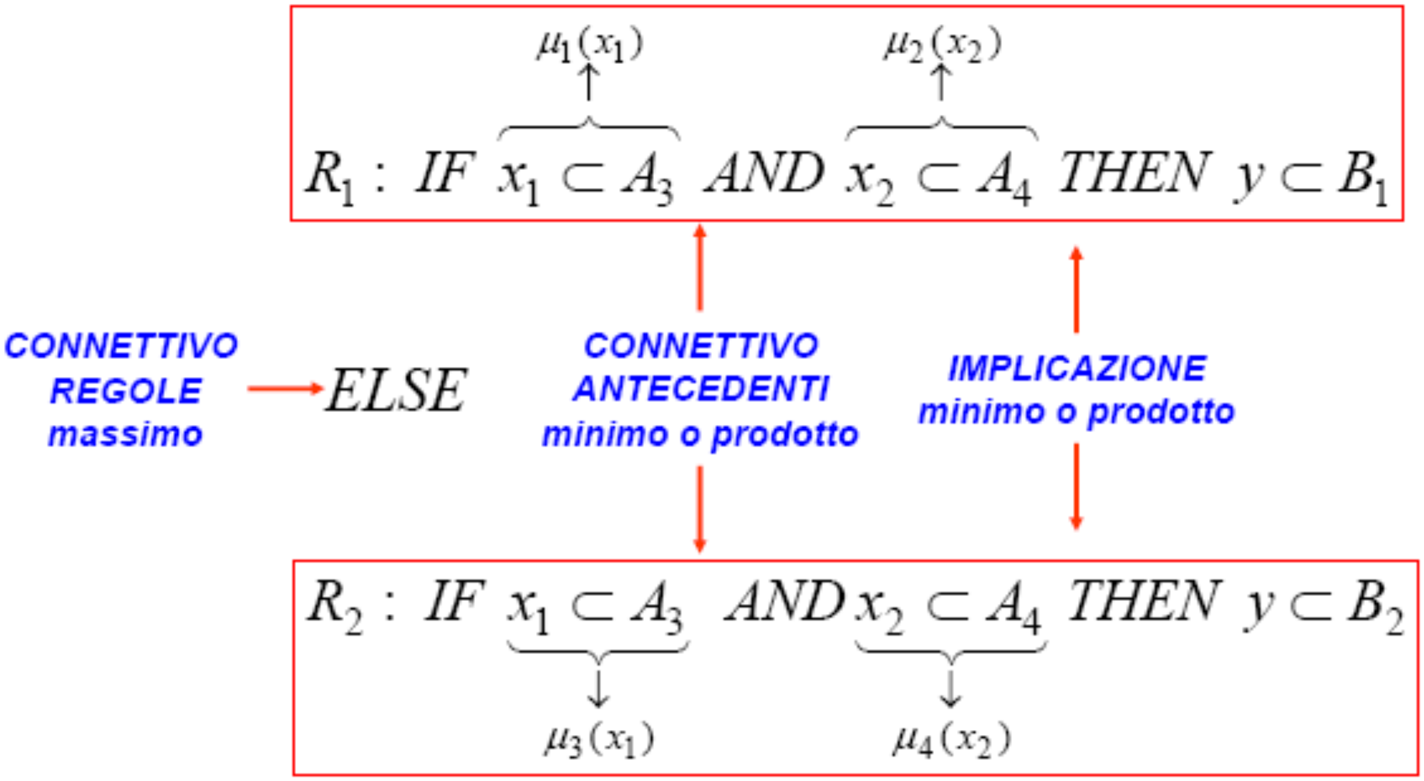
Intersection operators

Operators



Union operators

Inference rules



$$R_1 : IF\ \overbrace{x_1 \subset A_3}^{\overset{\mu_1(x_1)}{\uparrow}}\ AND\ \overbrace{x_2 \subset A_4}^{\overset{\mu_2(x_2)}{\uparrow}}\ THEN\ y \subset B_1$$

CONNETTIVO
REGOLE
massimo → $ELSE$

CONNETTIVO
ANTECEDENTI
minimo o prodotto

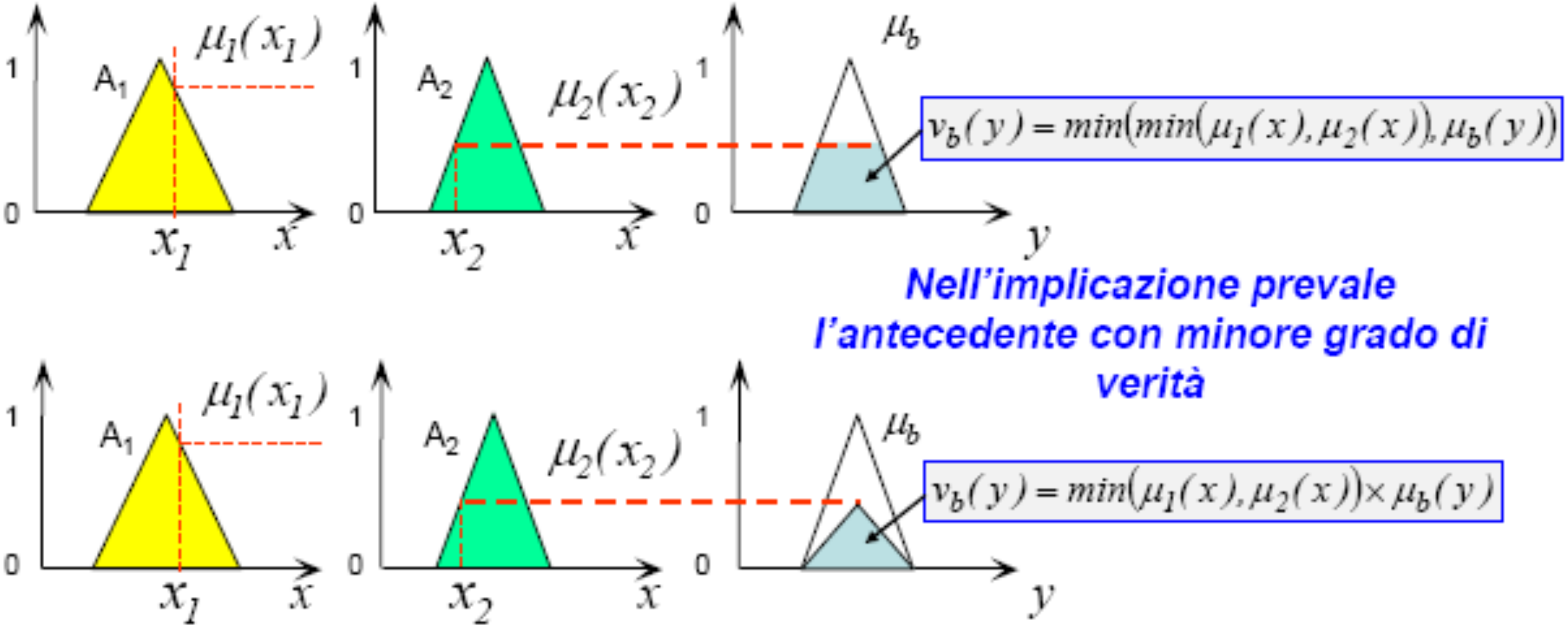
IMPLICAZIONE
minimo o prodotto

$$R_2 : IF\ \underbrace{x_1 \subset A_3}_{\underset{\mu_3(x_1)}{\downarrow}}\ AND\ \underbrace{x_2 \subset A_4}_{\underset{\mu_4(x_2)}{\downarrow}}\ THEN\ y \subset B_2$$

Inference (Mamdani)

$$IF\ (x_1\ is\ A_1)\ AND\ (x_2\ is\ A_2)\ THEN\ y\ is\ B$$

$$\mu_b(y) = (\mu_1(x^*) \wedge \mu_2(x^*)) \wedge \mu_b(y)$$



$$v_b(y) = min(min(\mu_1(x), \mu_2(x)), \mu_b(y))$$

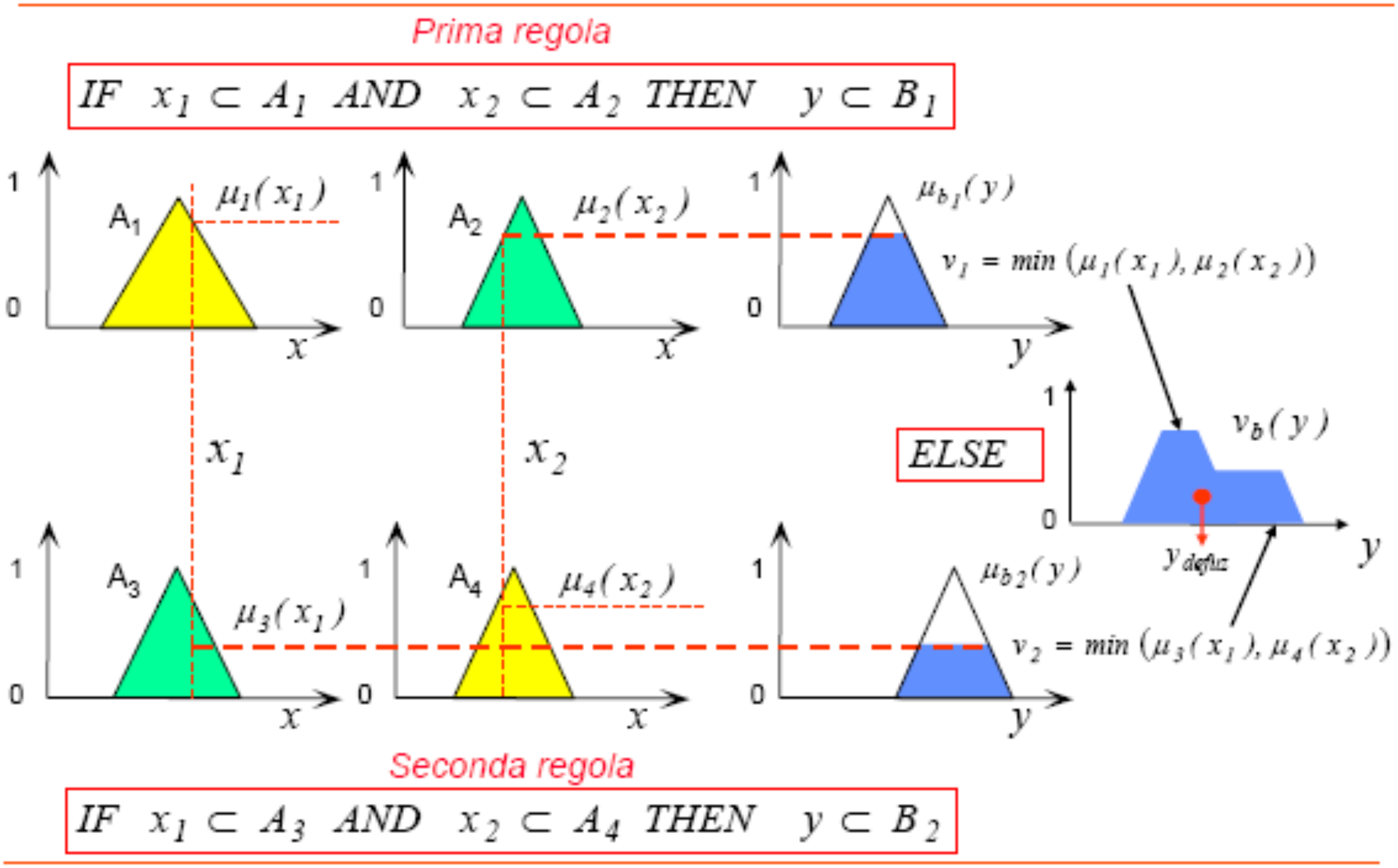
Nell'implicazione prevale l'antecedente con minore grado di verità

$$v_b(y) = min(\mu_1(x), \mu_2(x)) \times \mu_b(y)$$

Mamdani based inference

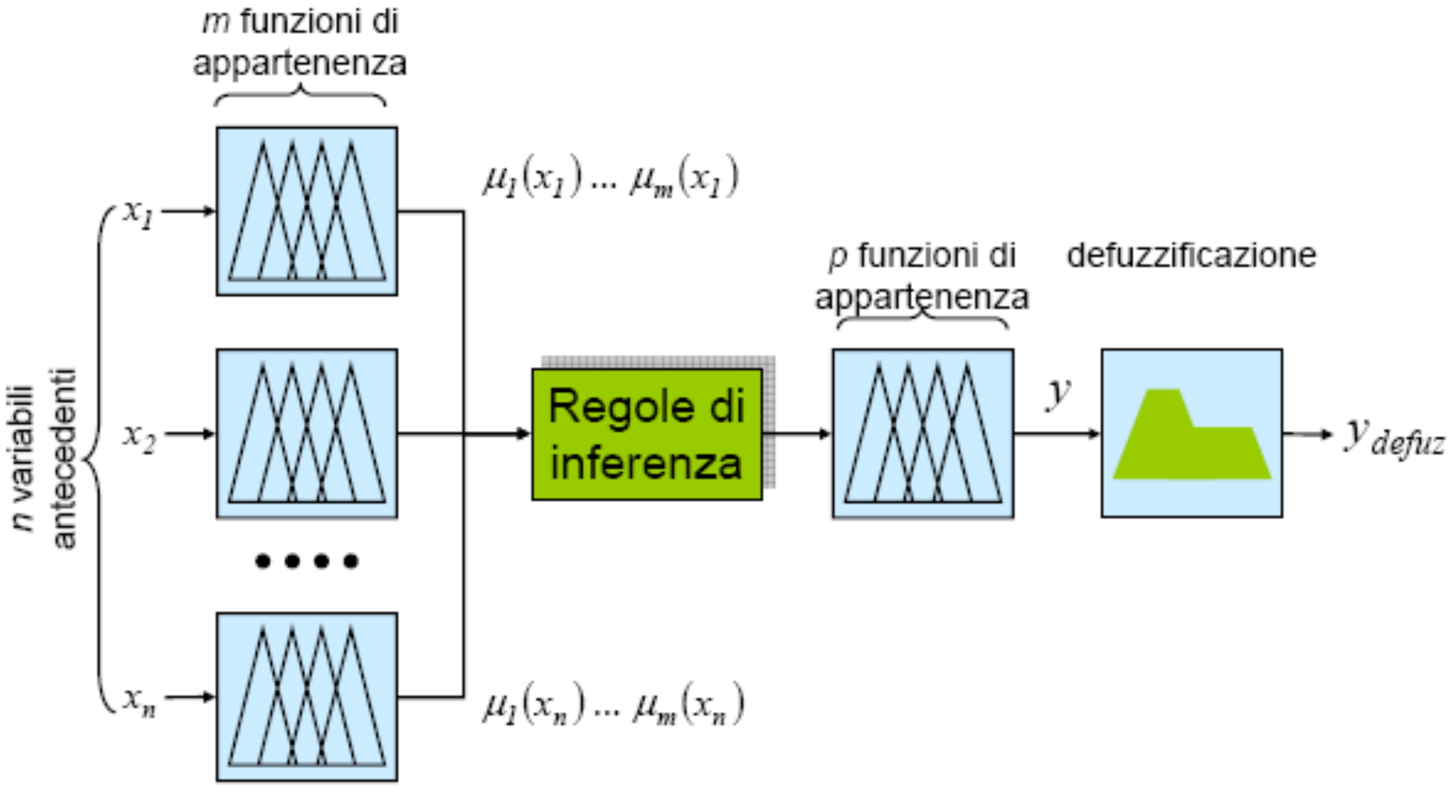
ISP – Verification tests

Defuzzification



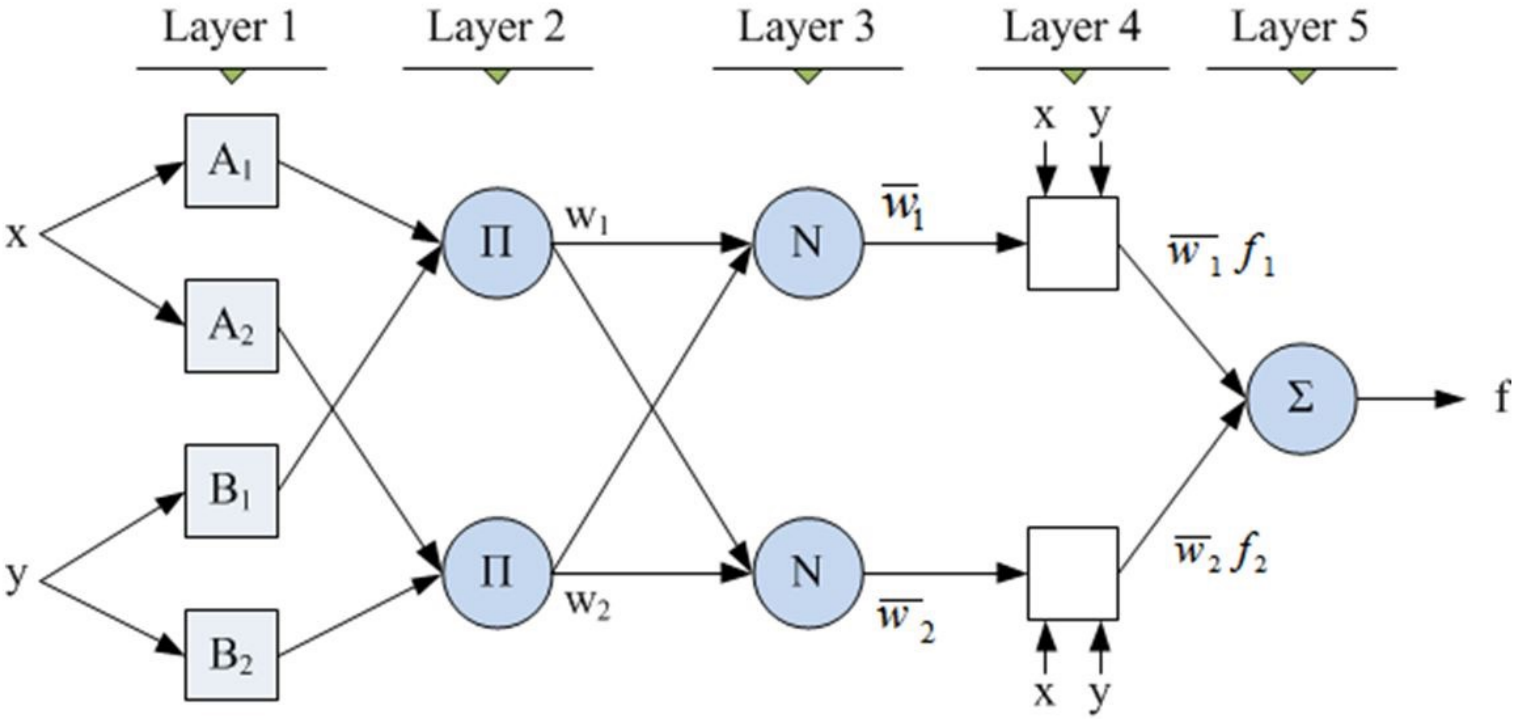
Inference and defuzzification

Fuzzy systems



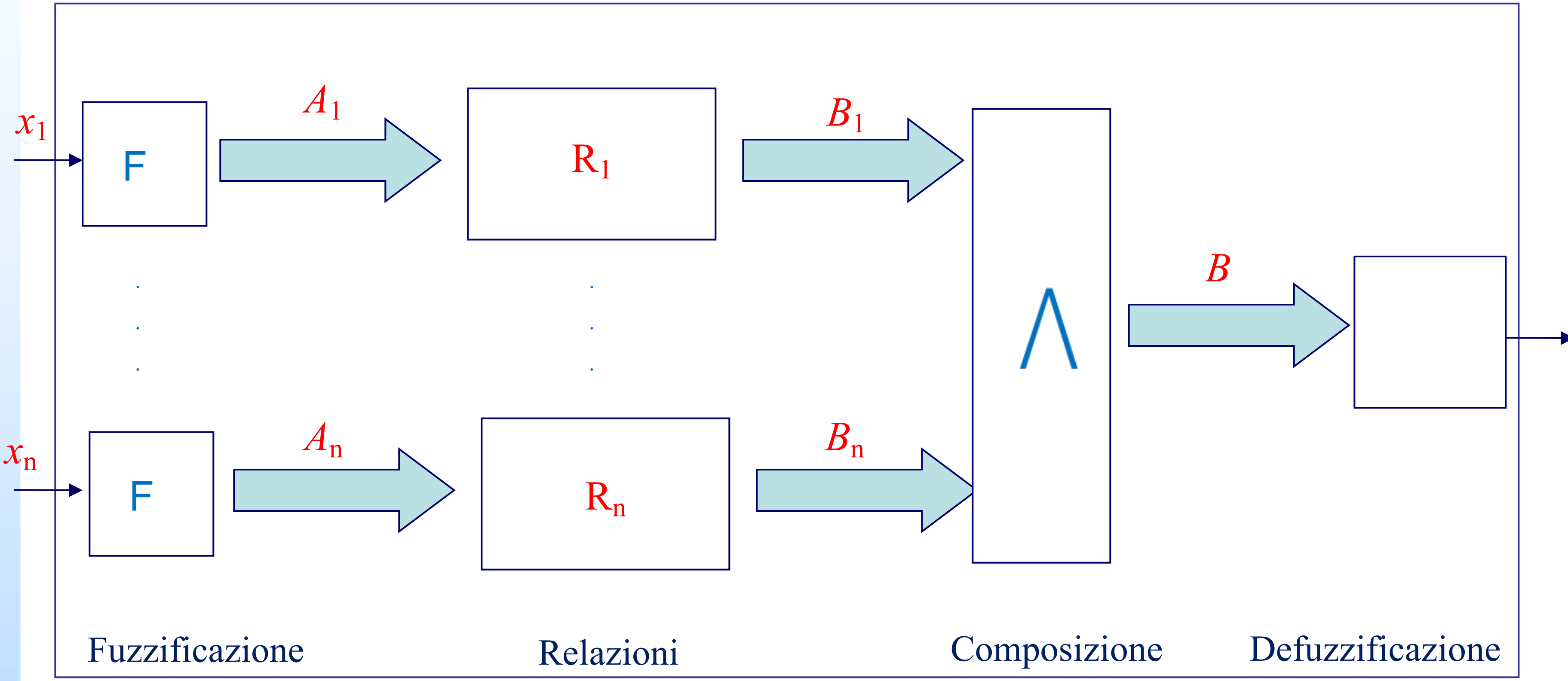
… neuro-fuzzy systems

ANFIS



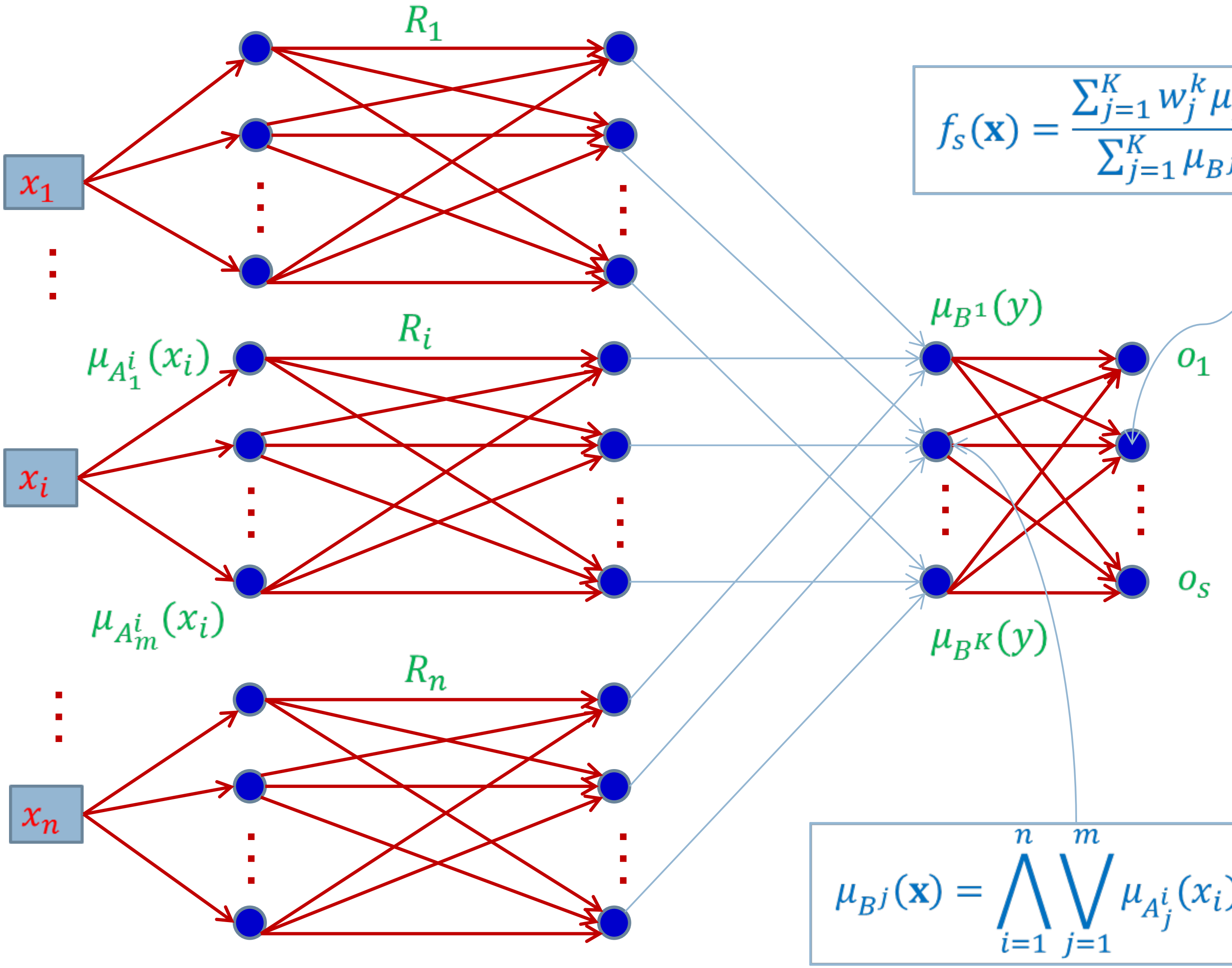
ANFIS model

FRNN



Fuzzy Relation Neural Network Model

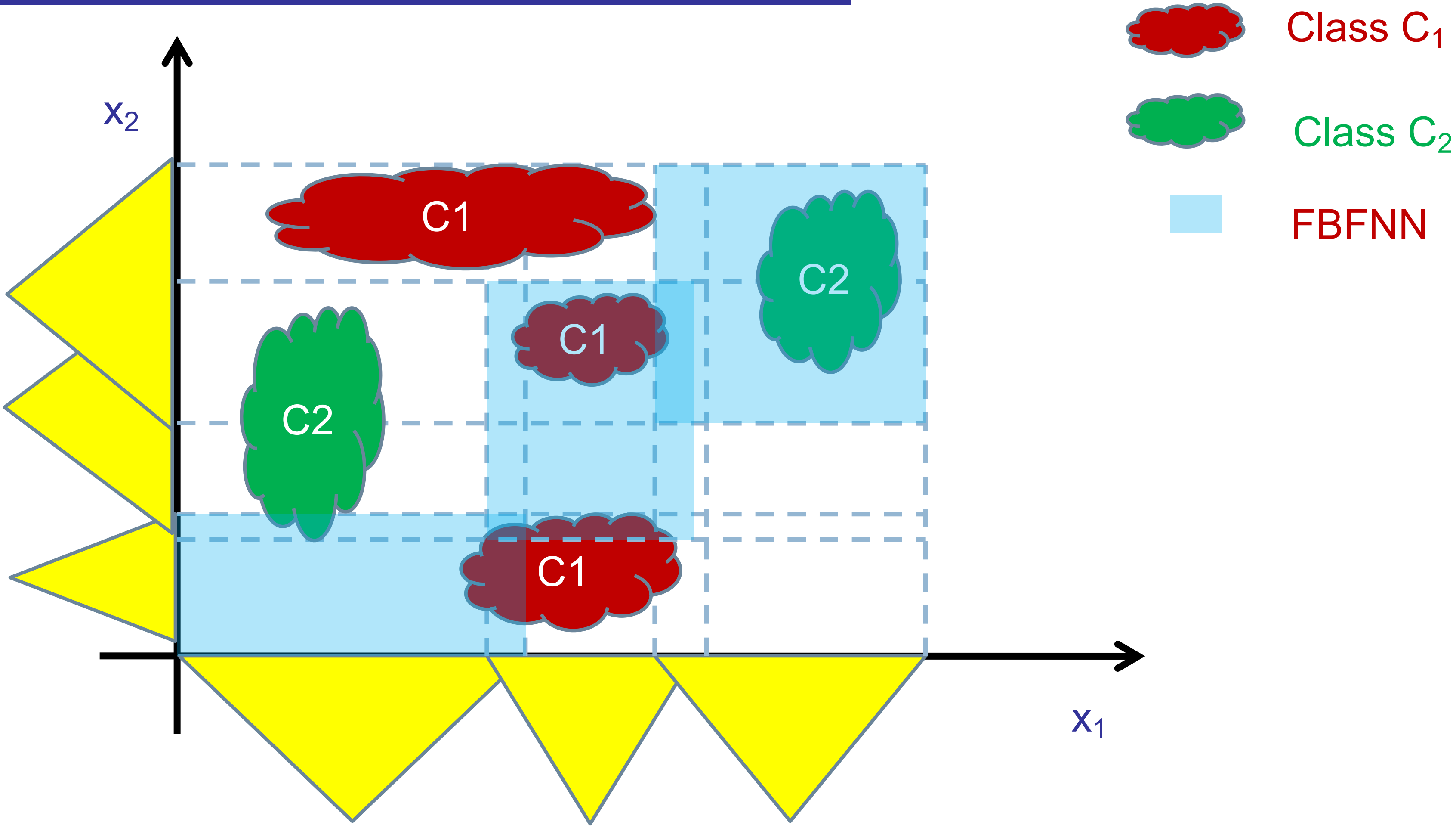
FRNN



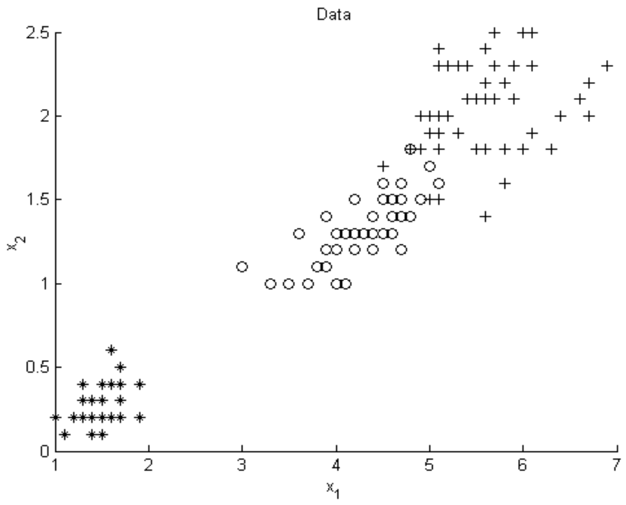
$$f_s(\mathbf{x}) = \frac{\sum_{j=1}^{K} w_j^k \mu_{B^j}(\mathbf{x})}{\sum_{j=1}^{K} \mu_{B^j}(\mathbf{x})}$$

$$\mu_{B^j}(\mathbf{x}) = \bigwedge_{i=1}^{n} \bigvee_{j=1}^{m} \mu_{A_j^i}(x_i) \circ R_i^j$$

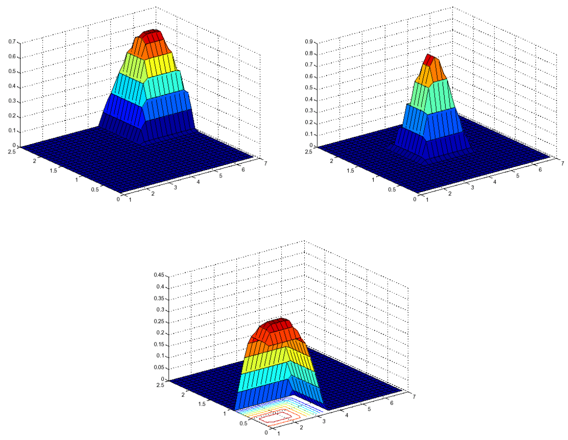
Granulation



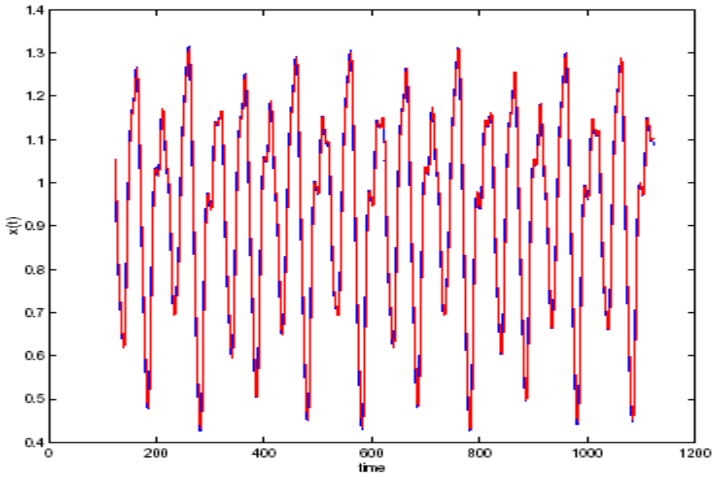
Some results



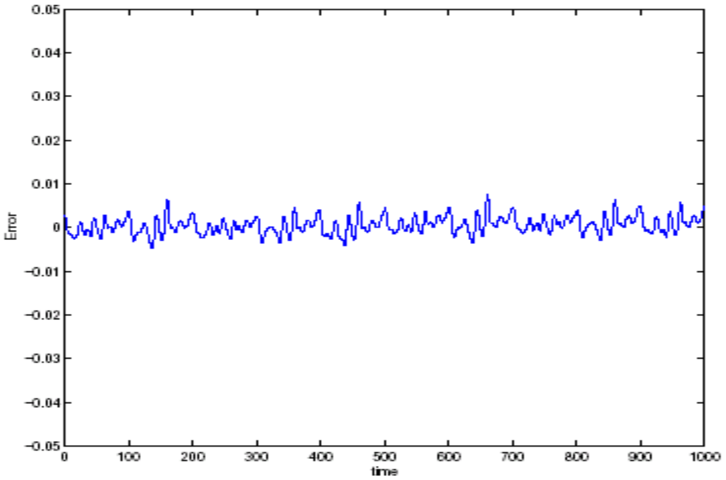
IRIS data set



Memberships



Mackey-Glass chaotic time series



Residum

References

- Material
  - Slides
  - Video Lessons

- Books
  - Fuzzy Logic with Engineering Applications, T. J. Ross, 4th Edition, 2016