Introduction to Biomedical Engineering

Biosignal processing

Kung-Bin Sung

6/11/2007

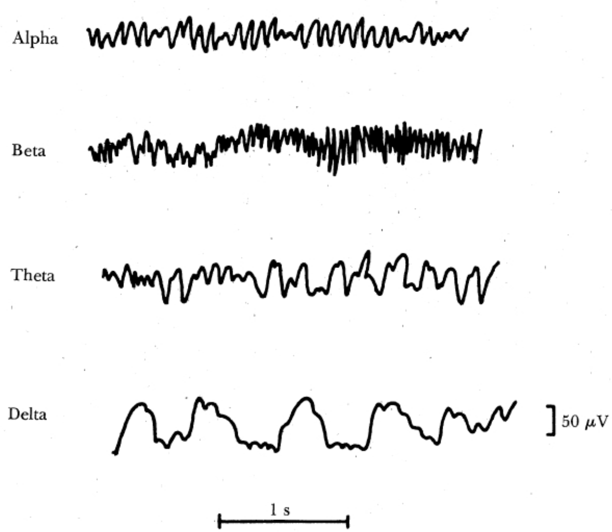
Outline

Chapter 10: Biosignal processing

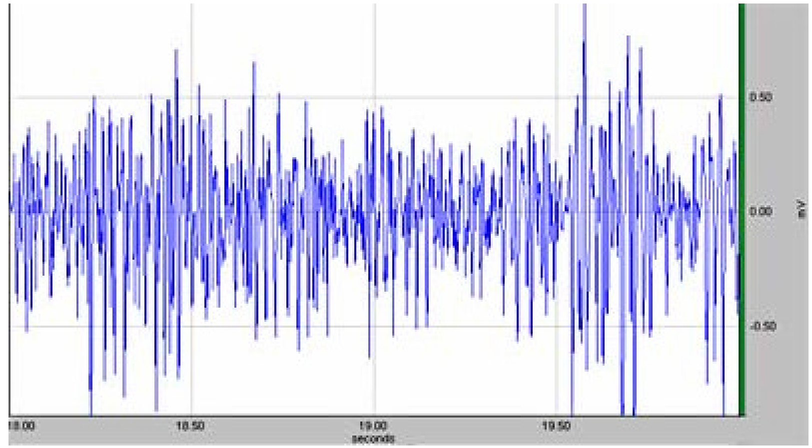
- Characteristics of biosignals
- Frequency domain representation and analysis
  - Fourier series, Fourier transform, discrete Fourier transform
  - Digital filters
- Signal averaging
- Time-frequency analysis
  - Short-time Fourier transform
  - Wavelet transform
- Artificial neural networks

Example biosignals

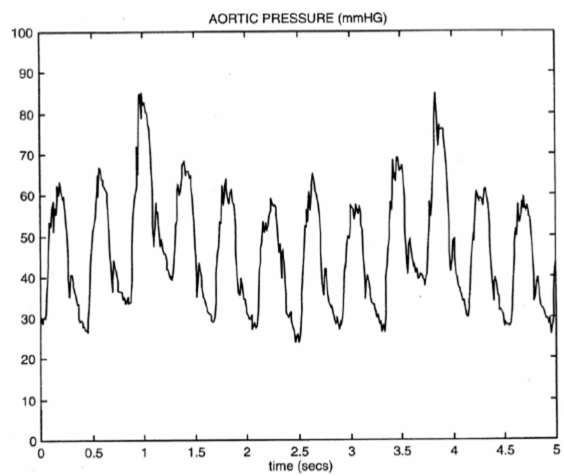
EEG



EMG



Blood pressure



ECG

Characteristics of (bio)signals

- Continuous vs. discrete

  $x(t)$ continuous variables such as time and space

  $x(n)$ n =0, 1, 2, 3… $\Rightarrow$ n is an integer

  sampled at a finite number of points

  we will deal with discrete signals in this module (a subset of digital signal processing)

- Deterministic vs. random
  - Deterministic: can be described by mathematical functions or rules

  Periodic signal is an example of deterministic signals

  $x(t) = x(t + nT)$ $\Rightarrow$ repeats itself every T units in time, T is the perio

  ECG and blood pressure both have dominant periodic components

Characteristics of biosignals

Random (stochastic) signals

- Contain uncertainty in the parameters that describe them, therefore, cannot be precisely described by mathematical functions
- Often analyzed using statistical techniques with probability distributions or simple statistical measures such as the mean and standard deviation
- Example: EMG (electromyogram)

Stationary random signals: the statistics or frequency spectra of the signal remain constant over time. It is the stationary properties of signals that we are interested in

Real biological signals always have some unpredictable noise or changes in parameters $\Rightarrow$ not completely deterministic

Signal processing

Ultimate goal of signal processing: to extract useful information from measured data

- Noise reduction and signal enhancement
- Signal conditioning
- Feature extraction
- Pattern recognition
- Classification such as diagnosis
- Data compression
- and more…

Time domain analysis

Some commonly used time-domain statistical measurements of biomedical signals

Root-mean-square $$RMS = \sqrt{\frac{\sum_{n=0}^{N-1} x^2(n)}{N}}$$

Average rectified value $$ARV = \frac{\sum_{n=0}^{N-1} \left| x(n) \right|}{N}$$

For example, the RMS value of an EMG signal is used to express the power of the signal, which can determine the fatigue, strength of force and ability of a muscle to handle mechanical resistance

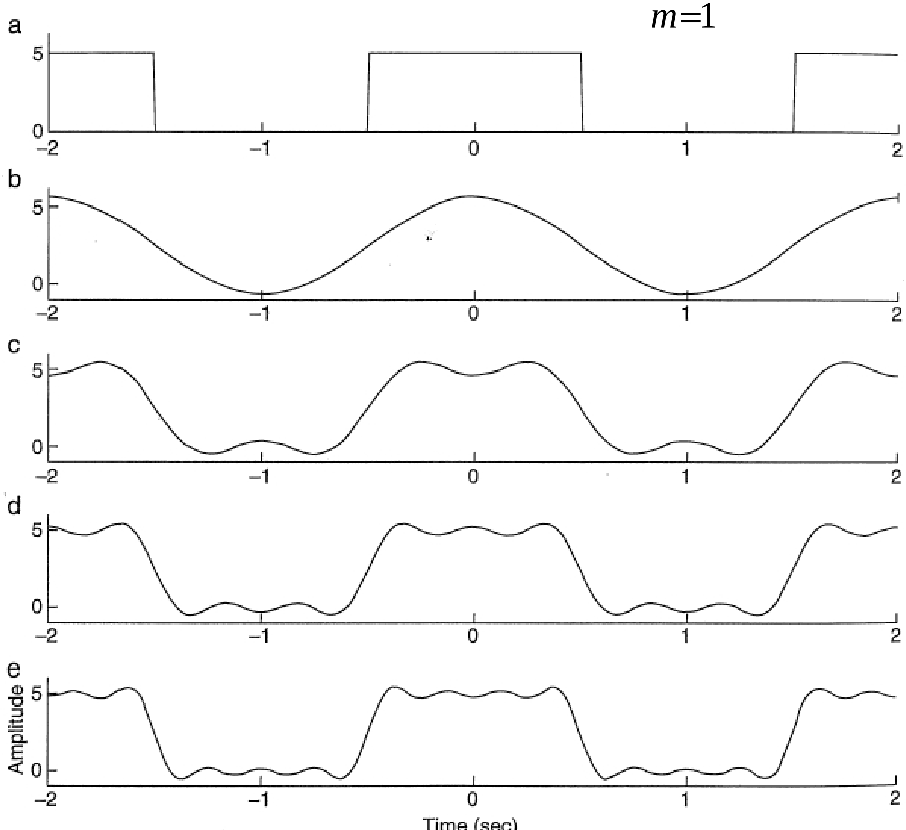
The ARV describes the smoothness of the EMG signal

Frequency domain representation of signals

Fourier's theory: a complex waveform can be approximated to any degree of accuracy with simpler functions

Example: a <u>periodic</u> square wave of period T can be represented by summing sinusoids with proper amplitudes and frequency

$$x(t) = a_0 + \sum_{m=1}^{\infty} (a_m \cos m w_0 t + b_m \sin m w_0 t) \qquad w_0 = \frac{2p}{T}$$



Number of harmonics

1

$$a_0 = \frac{1}{T}\int_T x(t)\,dt$$

2

$$a_m = \frac{2}{T}\int_T x(t)\cos(m w_0 t)\,dt$$

3

$$b_m = \frac{2}{T}\int_T x(t)\sin(m w_0 t)\,dt$$

4

Frequency domain representation of signals

But real biosignals are not periodic

As an expansion of the Fourier series in the previous slide, the Fourier integral or Fourier transform (FT) of a continuous signal is defined as

$$X(w) = \int_{-\infty}^{\infty} x(t) e^{-jwt} dt$$

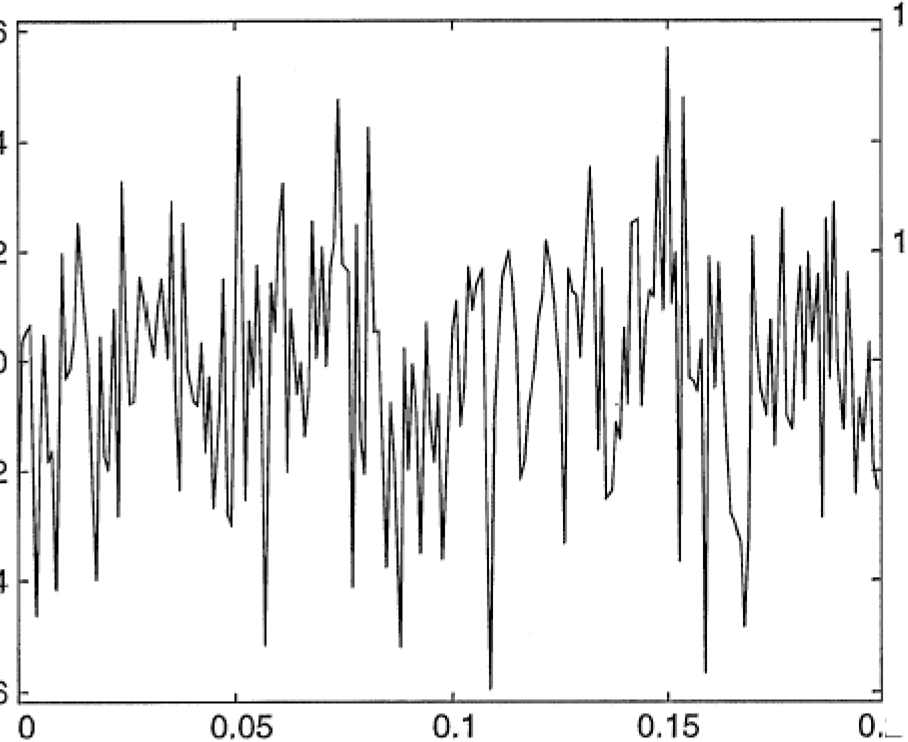
ω is continuous frequency, and X(ω) has complex values whose magnitude represents the amplitude of the frequency component at ω

The original (time-domain) signal can be completely recovered by the inverse Fourier transform (IFT), given sufficient sampling rate

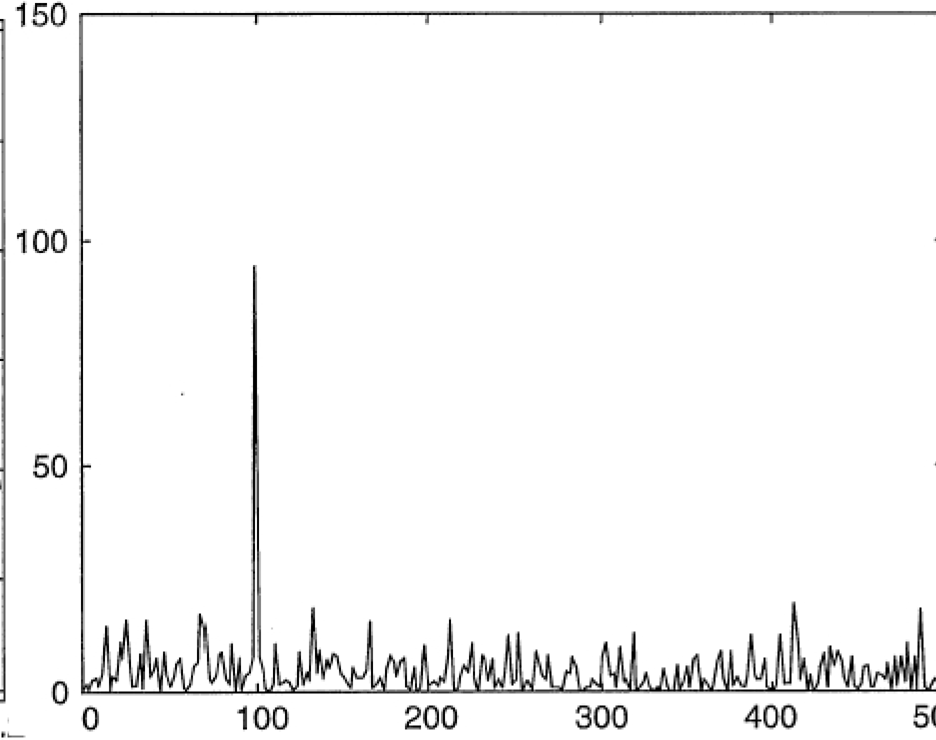
$$x(t) = \frac{1}{2p} \int_{-\infty}^{\infty} X(w) e^{jwt} dw$$

Example: extracting frequency-domain information in the signal

Time-domain signal (100Hz sine wave with random noise)

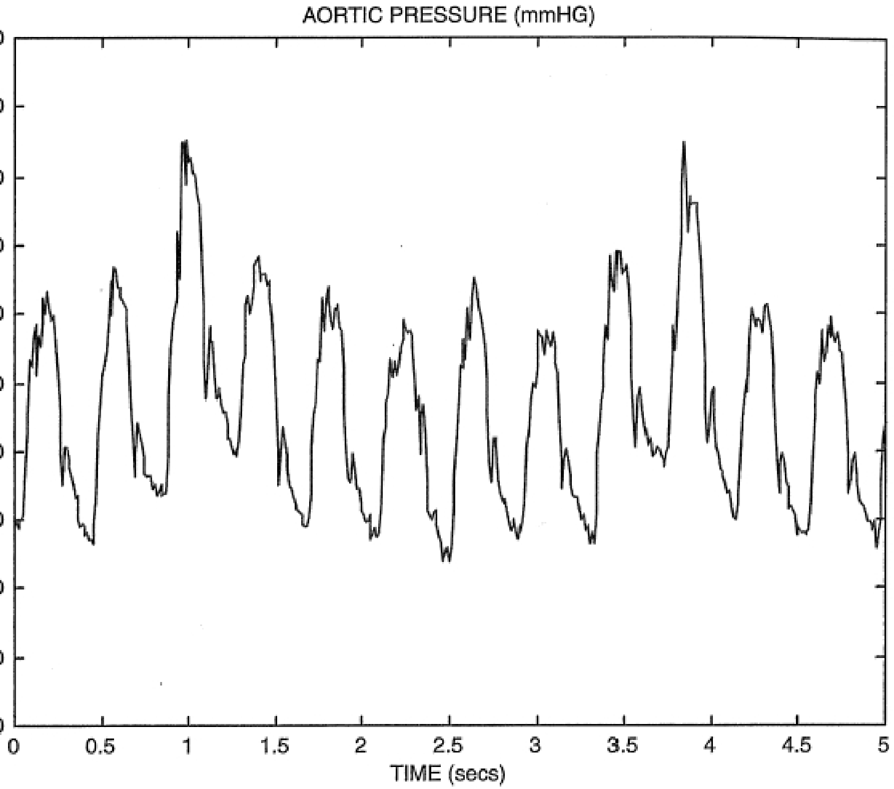


Frequency-domain (magnitude)

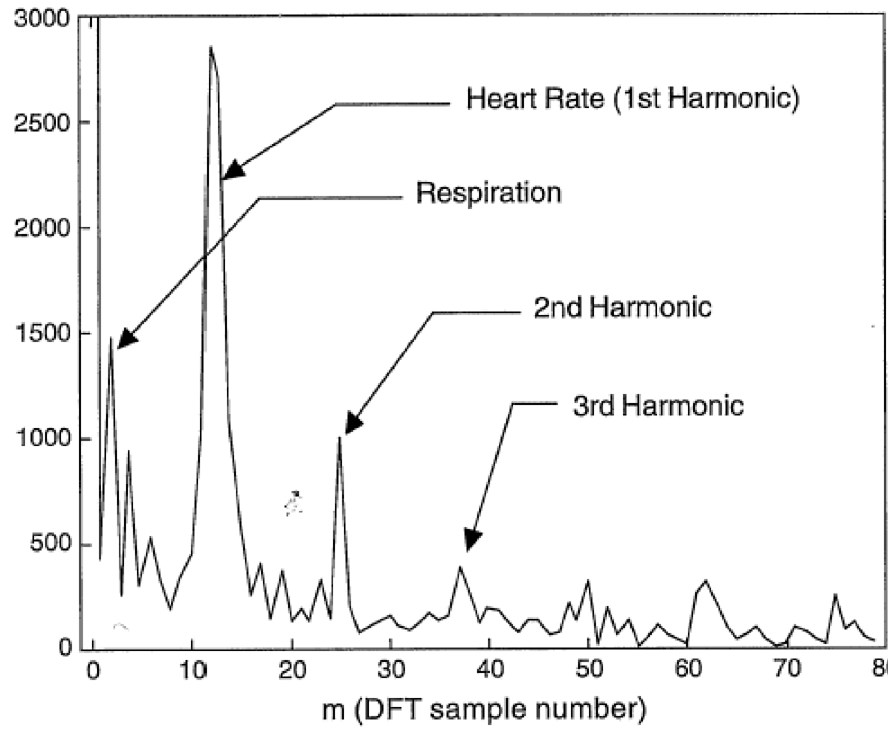


Frequency domain representation of signals

Example: blood pressure waveform (sampled at 200 points/s)



Fourier domain (magnitude)



In practice, the signal is discrete both in time and magnitude, and a discrete version of Fourier transform is carried out to get the Fourier (frequency) domain representation

Discrete Fourier transform (DFT)

Recall that the input is a discrete signal, which is basically a series of numbers

$x(n)$ $\quad n$ =0, 1, 2, …, N-1

The discrete Fourier transform (DFT) of the discrete signal is

$$X(m) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2pmn}{N}} \qquad n = 0, 1, \ldots, N-1$$

Similarly, an inverse discrete Fourier transform is of this form:

$$x(n) = \frac{1}{N} \sum_{m=0}^{N-1} X(m) e^{-j\frac{2pnm}{N}}$$

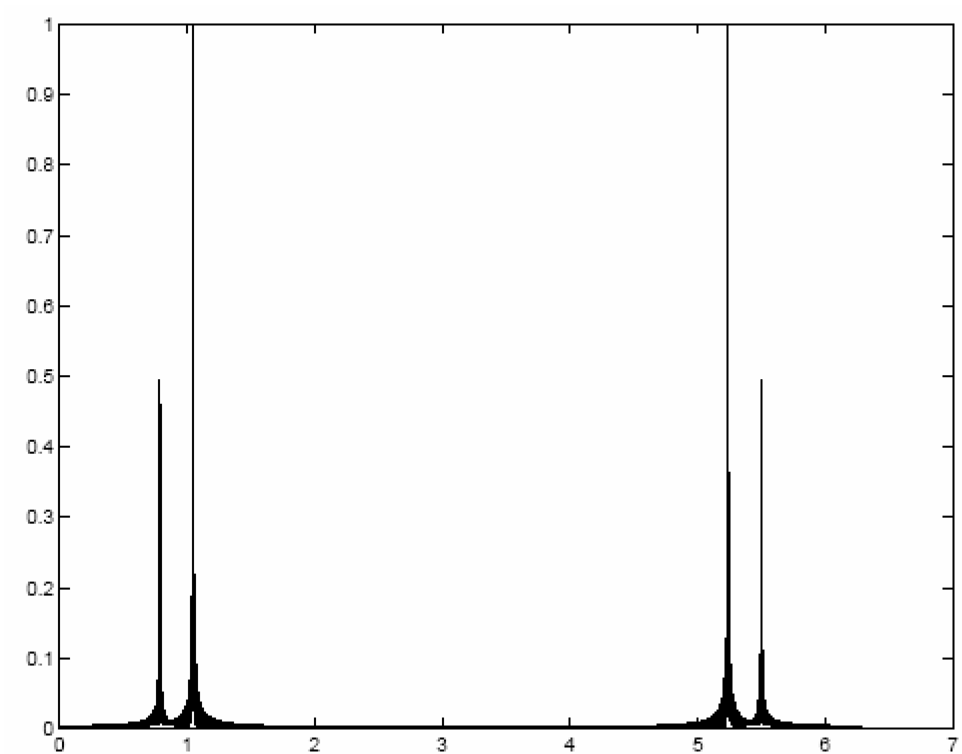
Note that the number of data points in $x(n)$ and $X(m)$ are always <u>the same</u>

The frequency in the Fourier domain is related to the sampling frequency f_s

Discrete Fourier transform

Example problem 10.13 $x(n) = \sin(\frac{p}{4} n) + 2\cos(\frac{p}{3} n)$

The magnitude of its DFT:



$m = 0$ $m = N\text{-}1$

Frequency (rad)

2 frequency components in $x(n)$

Total number of data points N

$\Rightarrow$ The step size in frequency is

$$\frac{2p}{N} \quad \text{(rad)}$$

The corresponding frequency ω of the $X(m)$

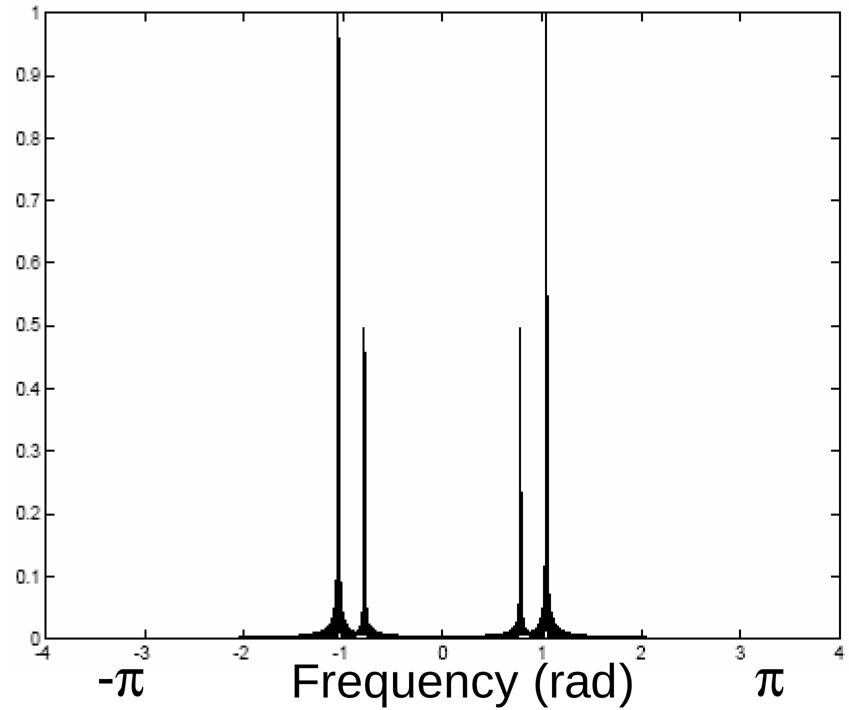
$$\frac{m}{N} \cdot 2p$$

Discrete Fourier transform

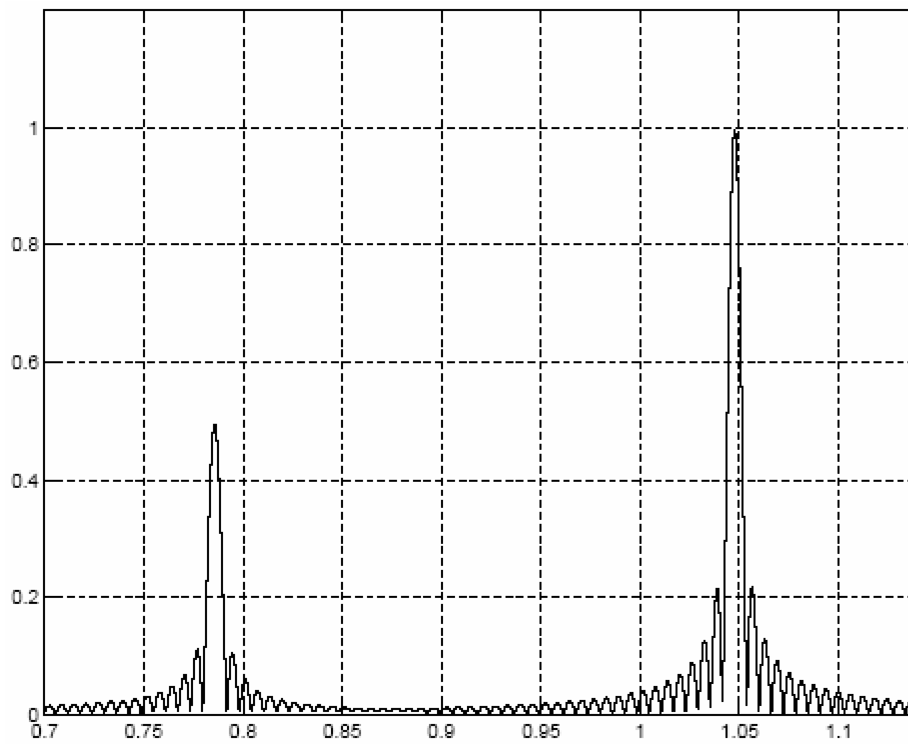
Example problem 10.13

For any real-valued signal, its Fourier transform has symmetric values with respect to ω=0. Conventionally, the DC (ω=0) component is plotted in the middle $\Rightarrow$ switch the left and right halves of DFT ("*fftshift*" function in Matlab)

DFT magnitude after "*fftshift*"



zoom in around the peaks



Discrete Fourier transform

- Another important feature of DFT is that it is periodic with a period of 2π
- Moreover, it is implied that the time-domain signal is also periodic
- Due to the symmetric values of DFT, it is sufficient to show only the frequency range $0 \sim \pi$

Discrete Fourier transform

Relationship between frequency-domain sequence and the time domain signal

For a signal that is sampled at a sampling frequency f_s (Hz)

- Its DFT is a sampled version of the continuous FT of the signal (sampling interval = $2\pi/N$ or f_s/N)

- Its DFT has a frequency range of -π~π (rad) which corresponds to

$$-\frac{1}{2} f_s \sim \frac{1}{2} f_s \quad \text{(Hz)}$$

Frequency domain analysis – comments

- Fourier transform describes the global frequency content of the signal
  - At each frequency ω, the magnitude of FT represents the amount of that frequency contained in the signal
  - At each frequency ω, the phase of FT measures the location (relative shift) of that frequency component. However, the phase information is more difficult to interpret and less often used than the magnitude
- Methods that provide time-frequency information of the signal
  - Short-time Fourier transform
  - Wavelet transform

Digital filters

As in the analog case, digital filters can be implemented in the frequency domain

time-domain $x(t)$ → Fourier transform → frequency-domain $X(w)$

$X(w)$ ↓ $X(w) \cdot H(w)$

$H(w)$ is the transfer function of the filter

$X(w) \cdot H(w)$ → inverse Fourier transform → filtered signal

<u>For discrete signals</u>

time-domain $x(n)$ → DFT → frequency-domain $X(m)$

$X(m)$ ↓ $X(m) \cdot H(m)$

$H(m)$ is a discrete transfer function sampled from $H(w)$

$X(m) \cdot H(m)$ → inverse DFT → filtered signal

Digital filters

Alternatively, digital filters can be implemented in the time-domain

The general form of a real-time digital filter (difference equation)

$$y(n) = \sum_{m=0}^{M} b_m x(n-m) - \sum_{m=1}^{M} a_m y(n-m)$$

$y(n)$ output of the current time n

$x(n)$ input of the current time n

$y(n-1)$ and $x(n-1)$ are output and input of the previous data point

- It is real-time because it does not need the value of any "future" samples
- Can be calculated easily

Digital filters

Finite impulse response (FIR) filter: impulse response has a finite number of nonzero points

Example: $$y(n) = \frac{1}{3}x(n) + \frac{1}{3}x(n-1) + \frac{1}{3}x(n-2)$$

Infinite impulse response (IIR) filter: impulse response has an infinite number of nonzero points

Example: $$y(n) = \frac{1}{2}x(n) + \frac{1}{2}y(n-1)$$

(depends on value of previous output)

Digital filters

The frequency-domain characteristics of digital filters can be analyzed by using the <u>z transform</u>

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

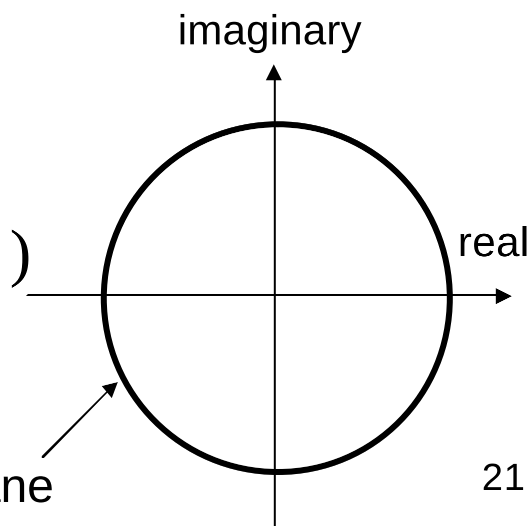
The z transform is similar to the Laplace transform which converts a continuous time-domain signal into frequency domain

We can describe the frequency response of a digital filter by using its transform function

$$H(z) = \frac{Y(z)}{X(z)}$$

$$z = e^{j2pf/f_s} = \cos(2pf / f_s) + j\sin(2pf / f_s)$$

for the frequency range $0 \leq f < 0.5 f_s$



Note that z is on the unit circle in the complex plane

Digital filters

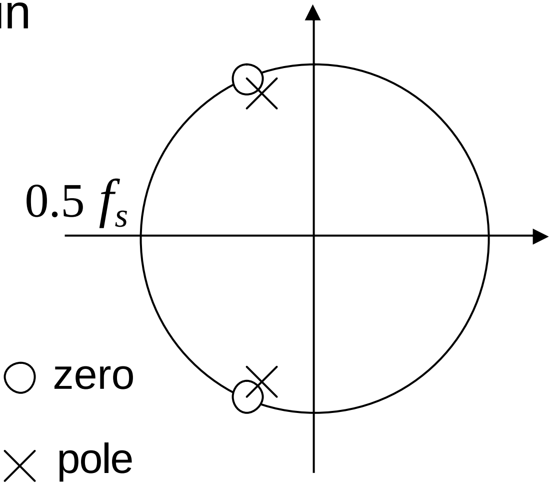
In a simple method of designing digital filters, the transfer function can be expressed as

$$H(z) = \prod_i \frac{(z - z_i)}{(z - p_i)}$$

where z_i is zero and p_i is pole

- We can set zeros on the unit circle to obtain low gain near the zero

- The poles are located near the zeros to obtain sharp transitions



Digital filters – example 1

60 Hz notch filter, sampling frequency f_s = 244.14Hz, we can set zeros and poles as

$$z_1 = e^{j2pf/f_s} = \cos q + j\sin q \qquad p_1 = az_1$$

$$z_2 = e^{-j2pf/f_s} = \cos q - j\sin q \qquad p_2 = az_2$$

$0 < \alpha < 1$

So the transfer function is

$$H(z) = \frac{Y(z)}{X(z)} = \frac{(z - z_1)(z - z_2)}{(z - p_1)(z - p_2)} = \frac{1 - 2\cos q \cdot z^{-1} + z^{-2}}{1 - 2a \cdot \cos q \cdot z^{-1} + a^2 \cdot z^{-2}}$$

A very useful property of the z transform (time-shifting)

$$Z[x(n-k)] = z^{-k} Z[x(n)]$$

Then we can get the digital filter

$$y(n) = x(n) - 2\cos q \cdot x(n-1) + x(n-2) + 2a\cos q \cdot y(n-1) - a^2 y(n-2)$$

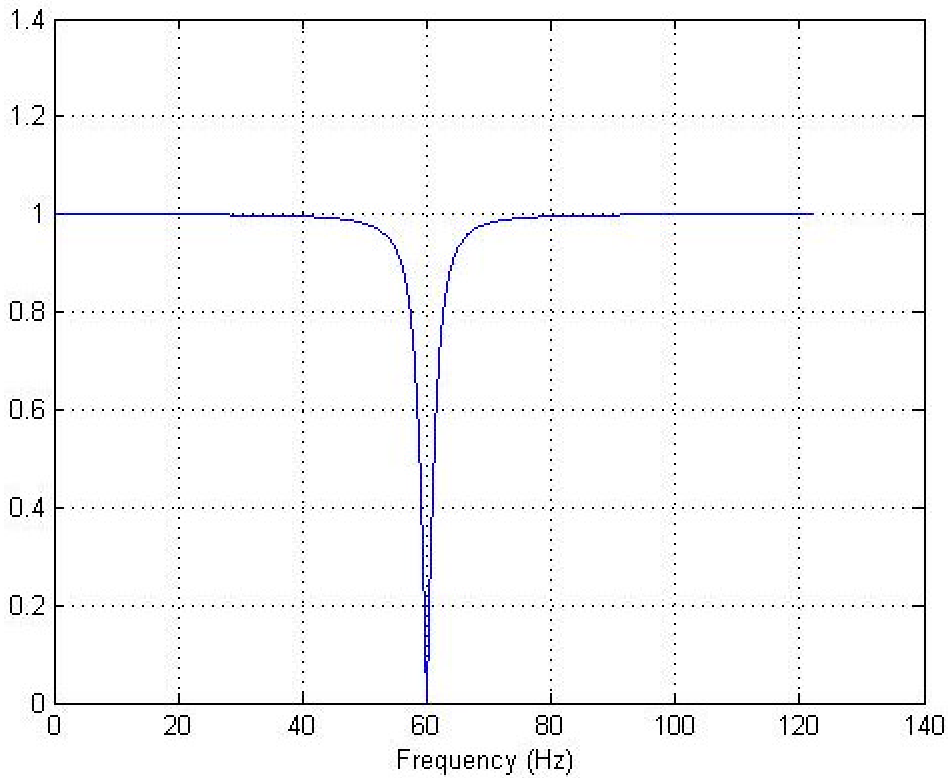
Digital filters – example 1

Substitute the numbers into the transfer function

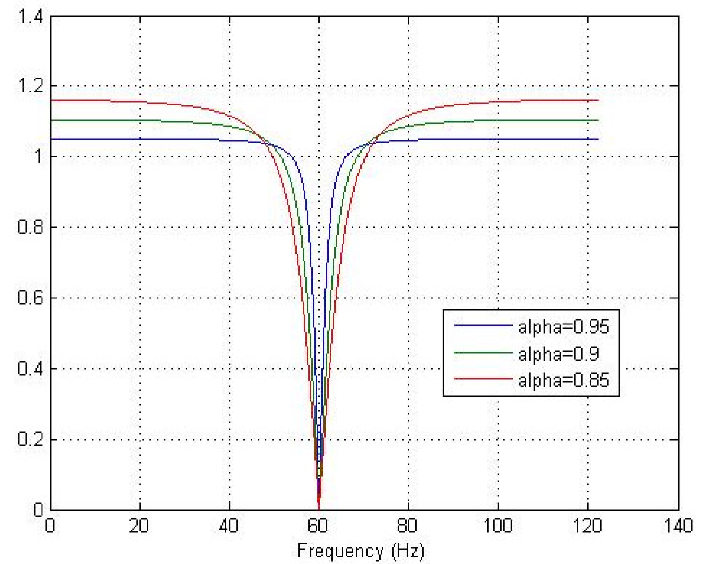
$$q = 2p(60/244.14) = 1.544$$

$\alpha = 0.95$

Frequency response of the 60Hz notch filter
(DC gain adjusted to 1)

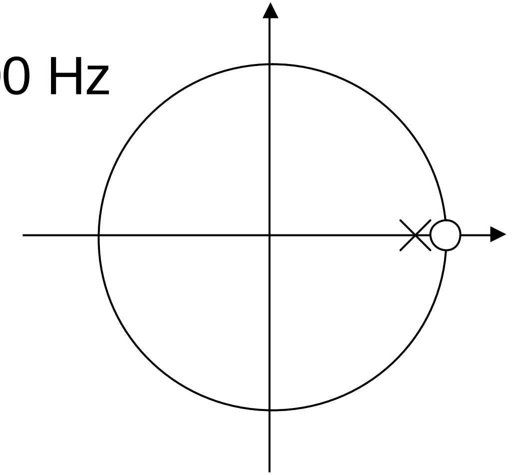


Effects of different α



Digital filters – example 2

A high-Q high pass filter, f_s = 100 Hz



Select:

double zeros at z=1

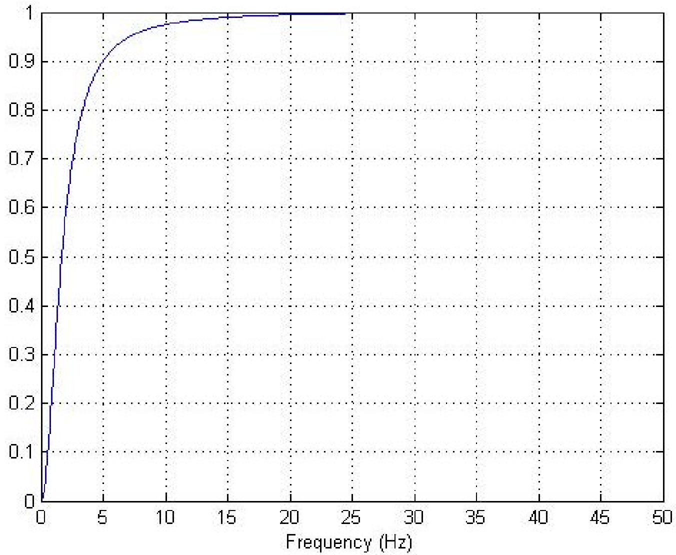
double poles at z=0.9

Transfer function:

$$H(z) = \frac{(z-1)^2}{(z-0.9)^2}$$

$$z = e^{j2pf/f_s}$$

Frequency response of the high pass filter



Digital filter:

$$y(n) = 0.9025x(n) - 1.805x(n-1)$$
$$+ 0.9025x(n-2) + 1.8y(n-1) - 0.81y(n-2)$$

Signal averaging

Averaging can be considered as a low-pass filter since high frequency components will be attenuated by averaging

For most biological signals there is a random noise superimposed on the quantity of interest

$$y_i(t) = x(t) + n(t)$$

$y_i(t)$ is the measured signal; subscript i indicates multiple measurements are obtained

$x(t)$ is the deterministic component, assuming it exists

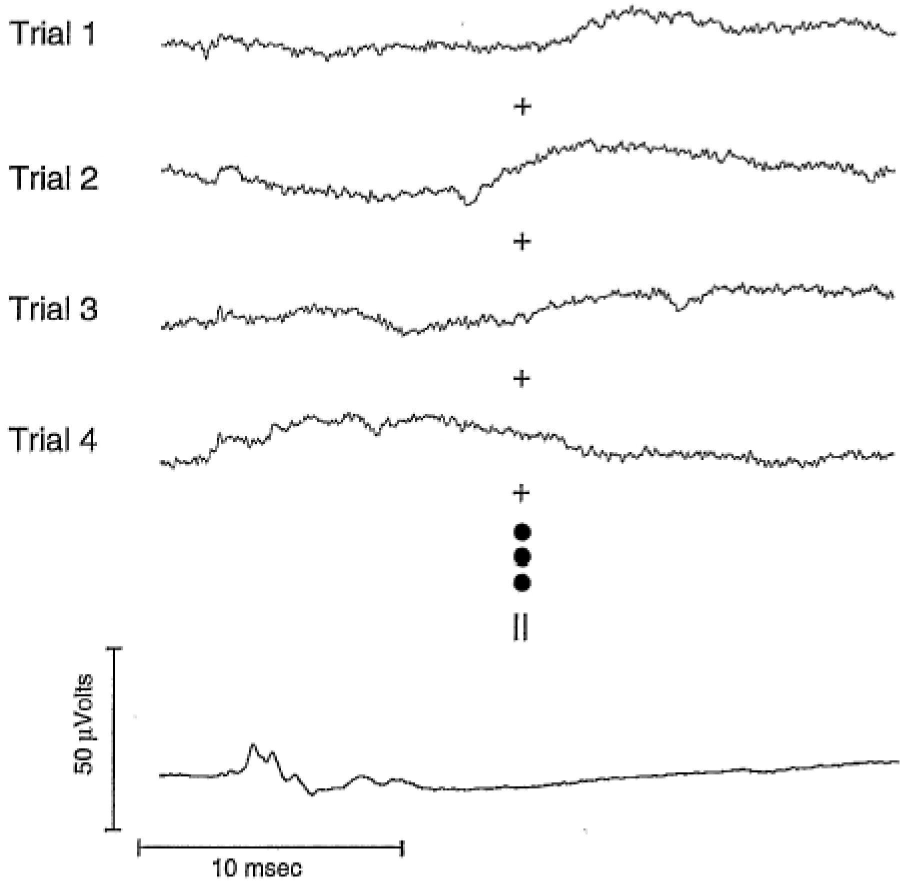
$n(t)$ is random noise

If we take the average of measurements from N separate trials

$$\bar{y}(k) = \frac{1}{N}\sum_{i=1}^{N} y_i(k) \quad \Rightarrow \quad \bar{y}(k) = x(k) + \frac{1}{N}\sum_{i=1}^{N} n(k)$$

If the noise is purely random the error of measurement will decrease as the number of trials N gets larger

Signal averaging – example 1



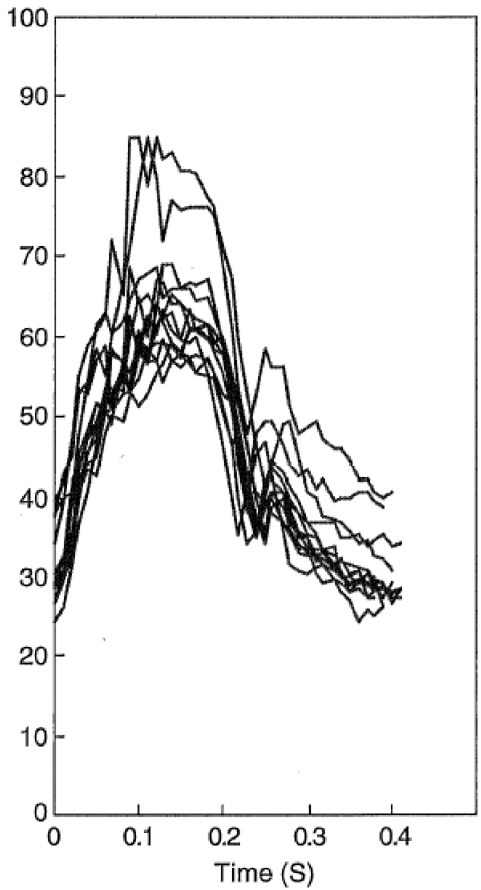
Auditory response averaged from 1000 trials (measurements) to reduce the effects of noise

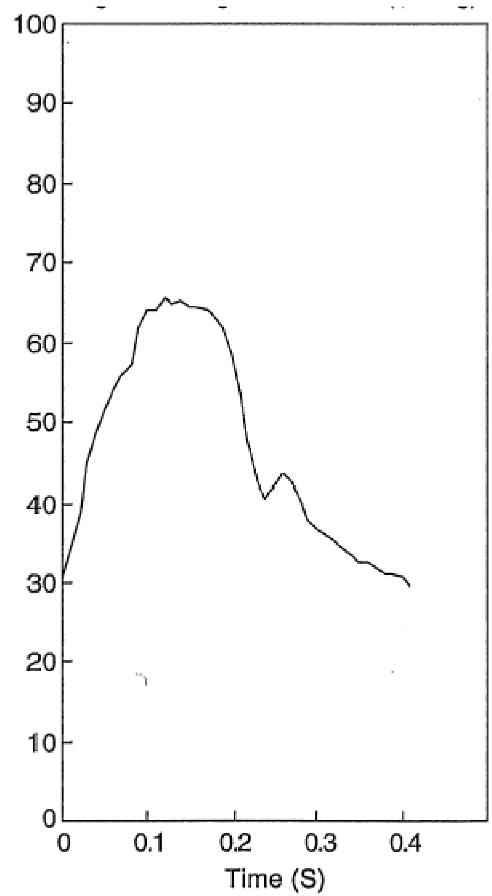
Signal averaging – example 2

Use the blood pressure data (slide 9) as an example

Note that the blood pressure waveform is approximately periodic

Overlay of many periods of the pressure waveform





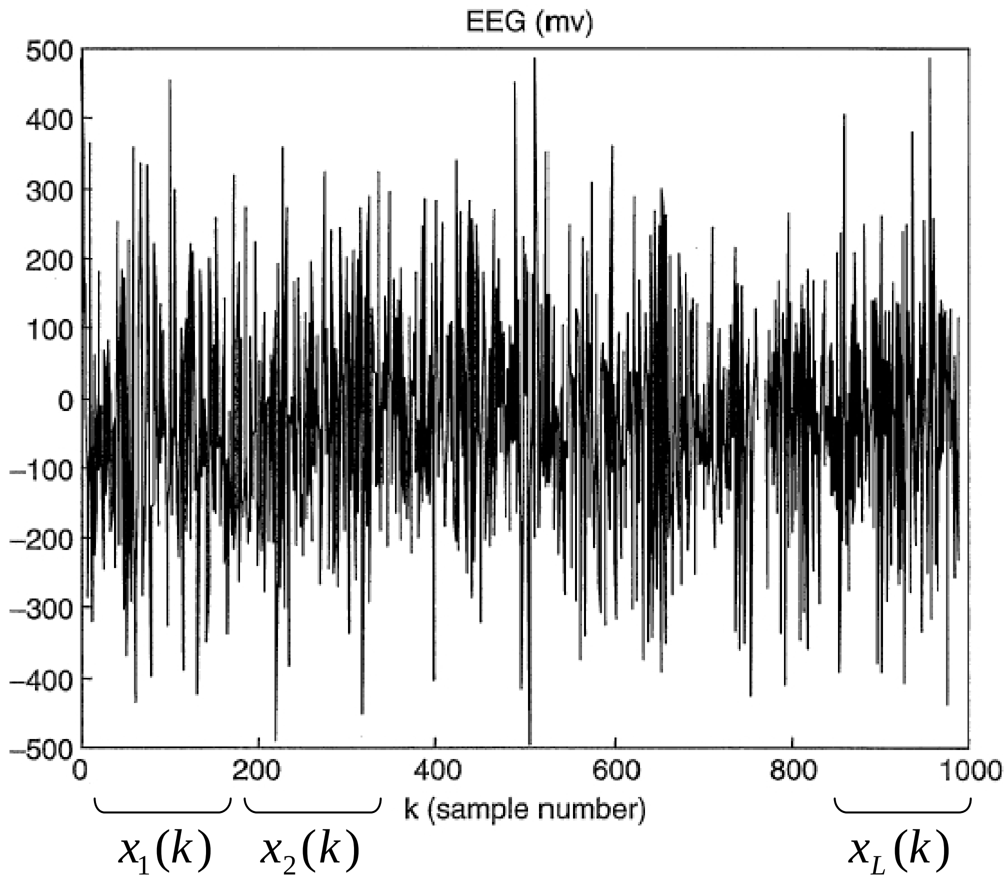
Average over all the periods

From the average waveform, we can get many useful parameters such as the maximum and minimum pressures, derivative of pressure rise during systole, and rate of decay during diastole

Signal averaging – example 3

For signals that are more random (aperiodic), signal averaging in the frequency domain may be useful

Example: EEG signal is aperiodic and the frequency of the signal is of interest because it indicates the activity level of the brain



multiple segments of data

$x_i(n)$ $i=1,2,\ldots,L$

↓ DFT

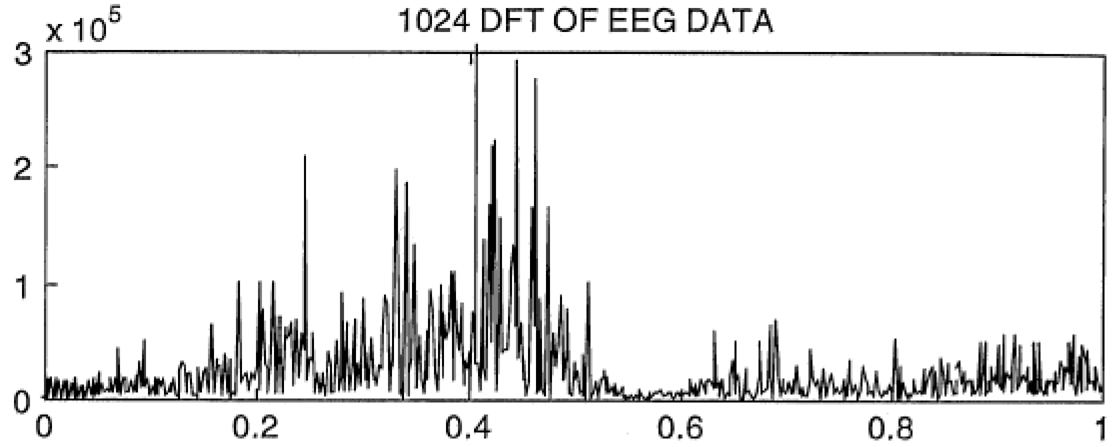
$X_i(m)$

↓ $|\bullet|^2$

$P_i(m) \xrightarrow{\text{average}} \overline{P}(m)$

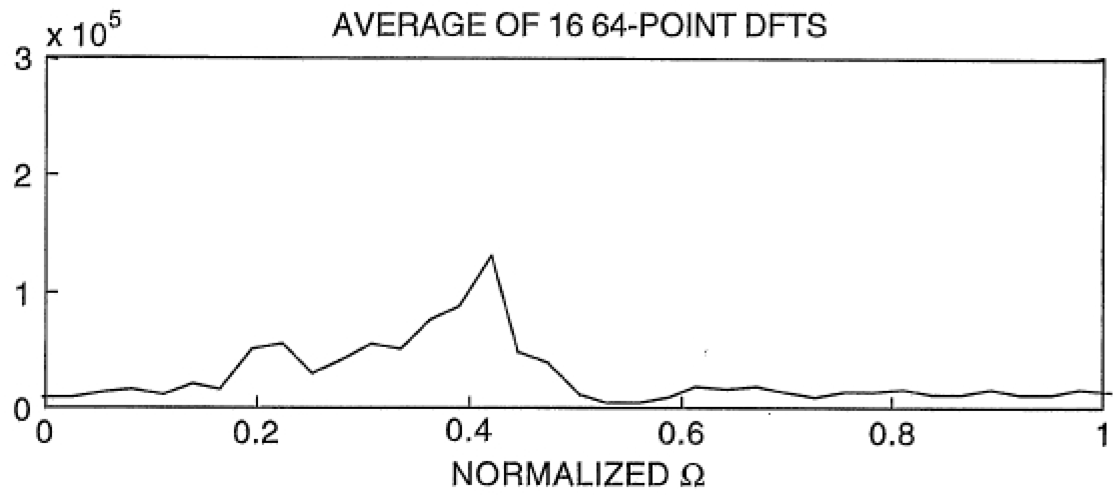
Signal averaging – example 3

DFT (magnitude) of EEG data from previous slide



Raw data is divided into 16 segments, each containing 64 points

Average in the frequency domain



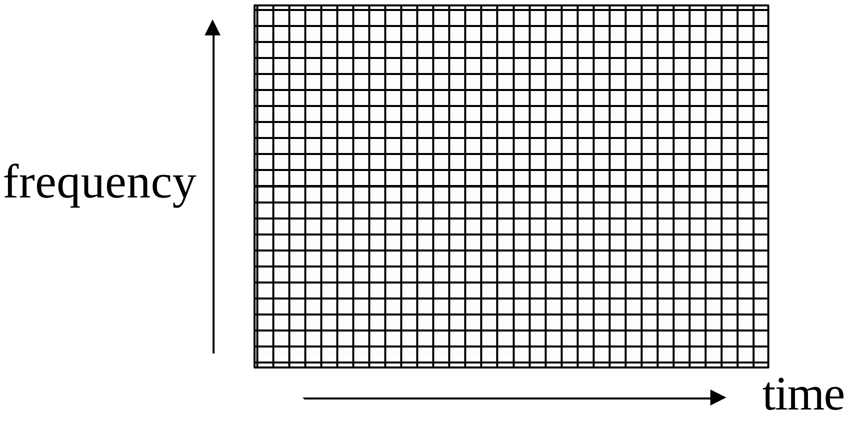
Time-frequency analysis

To capture the "local" frequency characteristics of the signal, the short-time Fourier transform (STFT) can be used and is defined as

$$X(w,a) = \int_{-\infty}^{\infty} x(t)g(t-a)e^{-jwt}dt$$

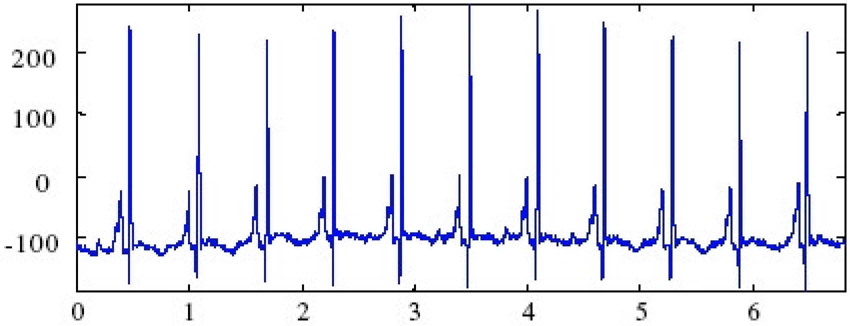
where $g(t)$ is a window function which has a limited time span. a is the amount of shift of the window function, therefore we can obtain the FT of the signal and know where in time it occurs

The result of STFT is a 2D matrix whose elements are the coefficients at corresponding frequency ω and time-shift a

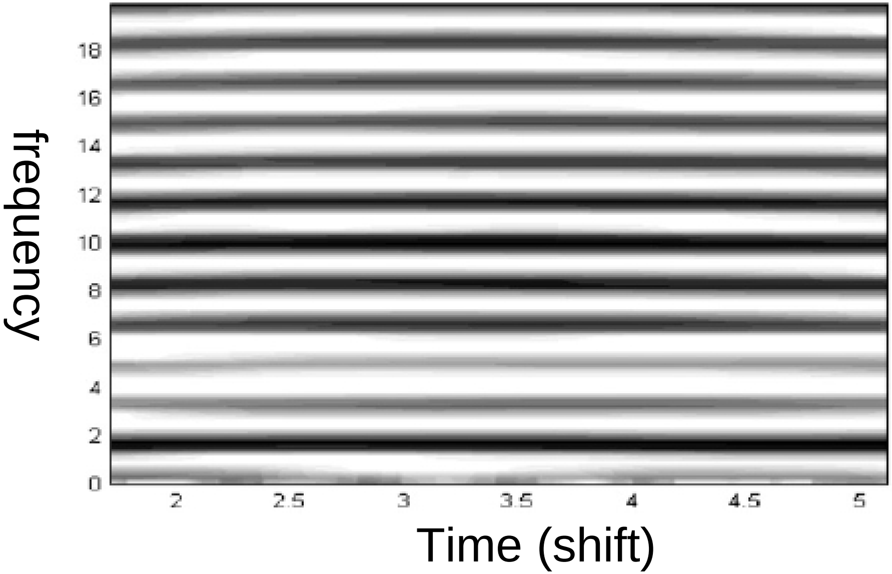


Example of STFT of ECG

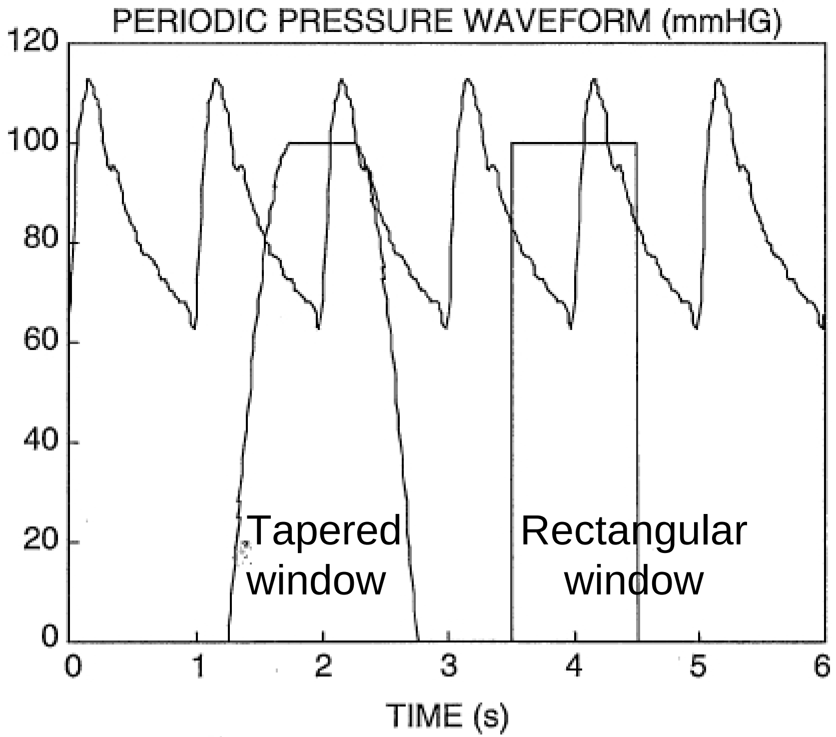
Original ECG signal



Magnitude of STFT



In practice, a sharp window is not the best choice due to the rippling effects it causes



STFT – comments

However, the short-time Fourier transform has two major shortcomings:

• The window length is fixed throughout the analysis. We are not able to capture events with different durations.

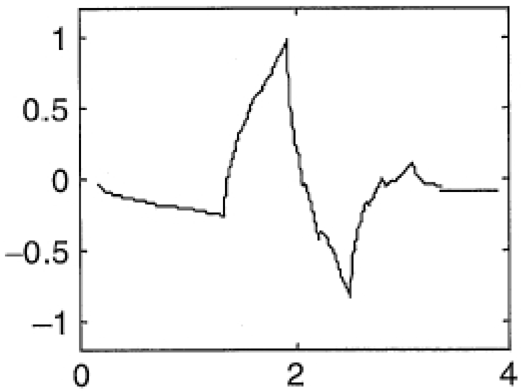
• The sinusoidal functions used in STFT to model the signal may not be the best choice. Specifically, the local features of biomedical signals may contain sharp corners that can not be modeled by the smooth shape of the sinusoidal waveforms.

To address the above shortcomings $\Rightarrow$ Wavelet Transform

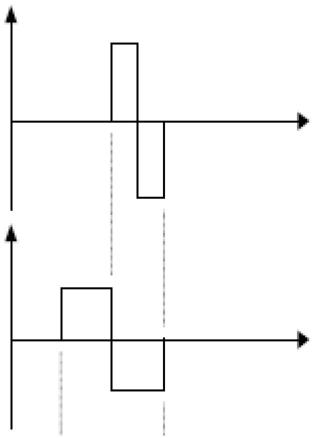
Wavelet transform (WT)

Some commonly used wavelets for processing of biomedical signals

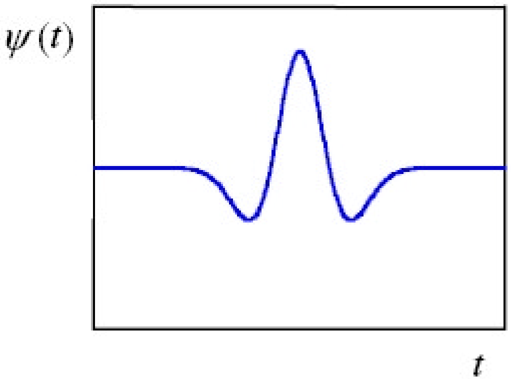
The Daubechies order 4 (db4) wavelet



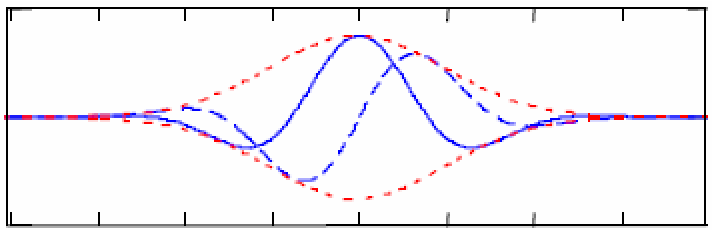
The Haar wavelets



The Maxican hat wavelet



The Morlet wavelet with $\omega_0 = 2$

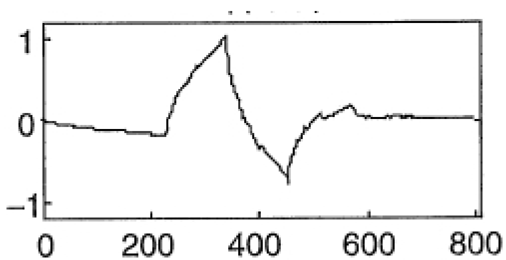


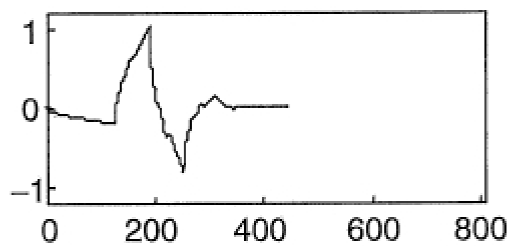
Wavelet transform

To address the problem of fixed window widths, the concept of “scaling” is used

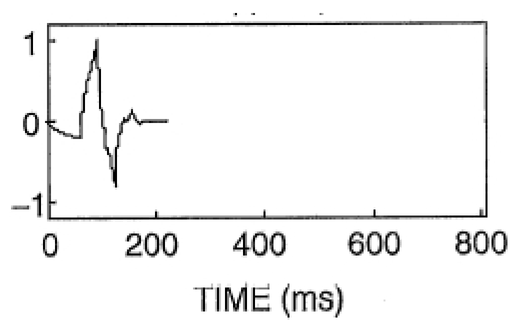
The mother wavelet is scaled in time to create a series of “high-frequency” components as an analogy to the harmonics in sinusoidal decomposition (Fourier series)

Original wavelet





Scaled by a factor of 1/2



Scaled by a factor of 1/4

Wavelet transform

The continuous wavelet transform of $x(t)$ can be expressed as

$$C(a,s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t) j\ * (\frac{t-a}{s}) dt$$

where a is the shifting factor and s is the scaling factor

$j\ (t)$ is the mother wavelet

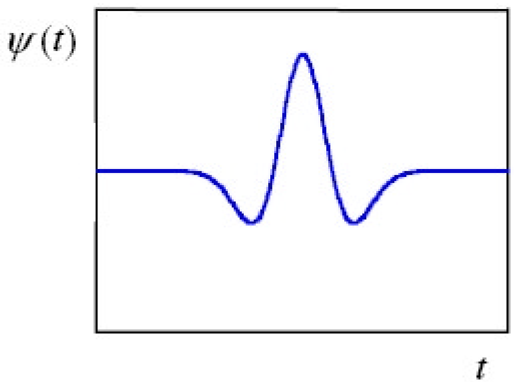
The WT coefficients $C(a,s)$

- measure the similarity between the wavelet basis functions and the input waveform $x(t)$

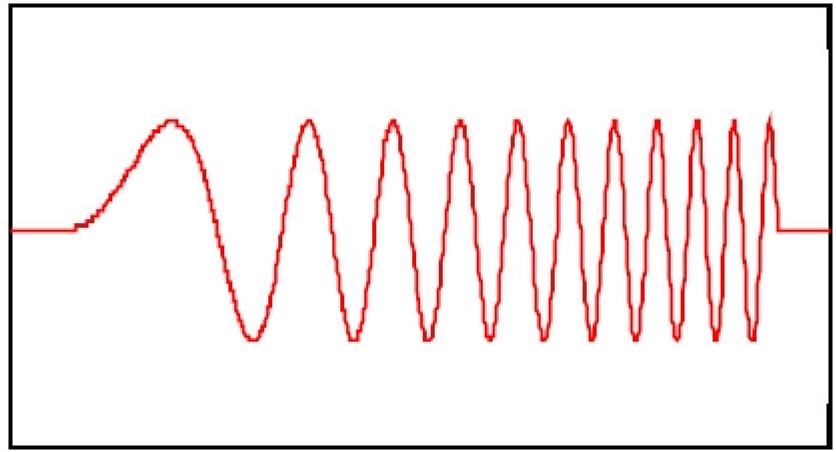
- is a function of the shifting factor and the scaling factor (2D)

Wavelet transform – example

The Maxican hat wavelet



Chirp signal



Wavelet transform coefficients

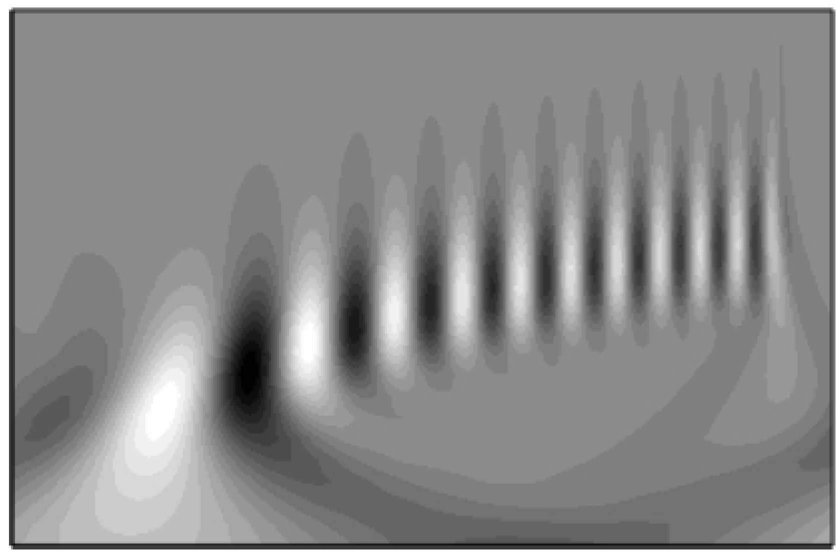
small

High frequency

Scale (s)

Low frequency

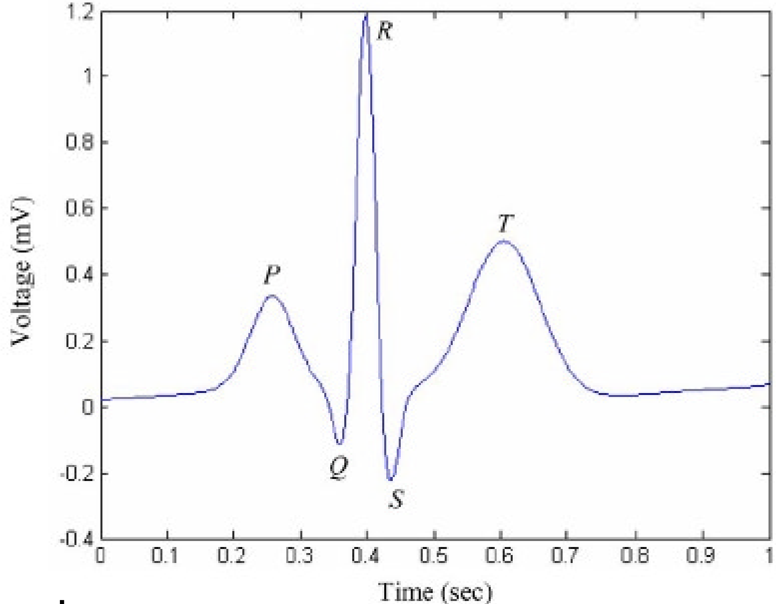
large



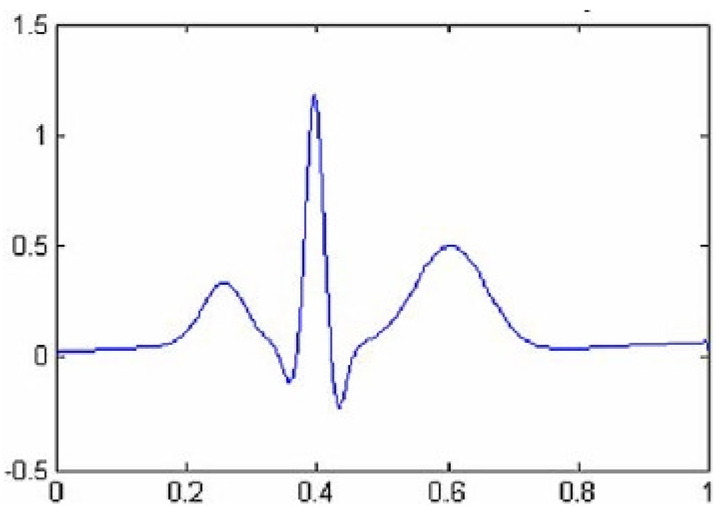
Shift (a)

Wavelet transform – reconstruction

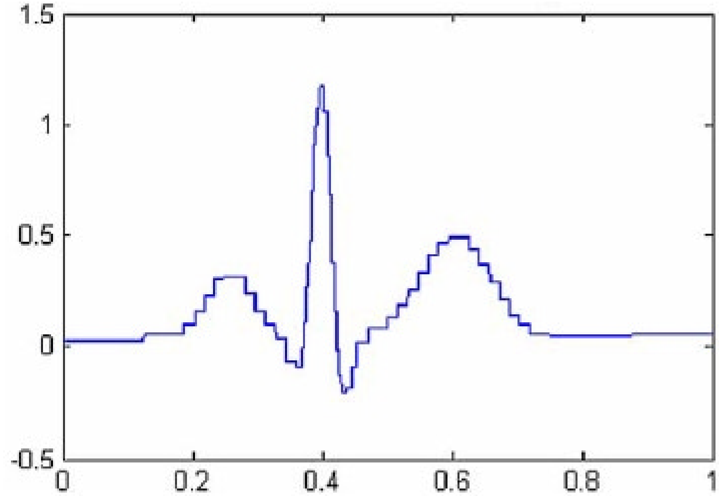
Original ECG signal



Reconstructed waveform using Daubechies order 8 (db8) wavelet



Reconstructed waveform using Haar wavelet



Wavelet transform – summary

The basis functions in WT are the shifted and scaled versions of the mother wavelet

Every choice of mother wavelet gives a particular WT $\Rightarrow$ very important to choose most suitable mother wavelet for a particular task.

Rules of thumb:

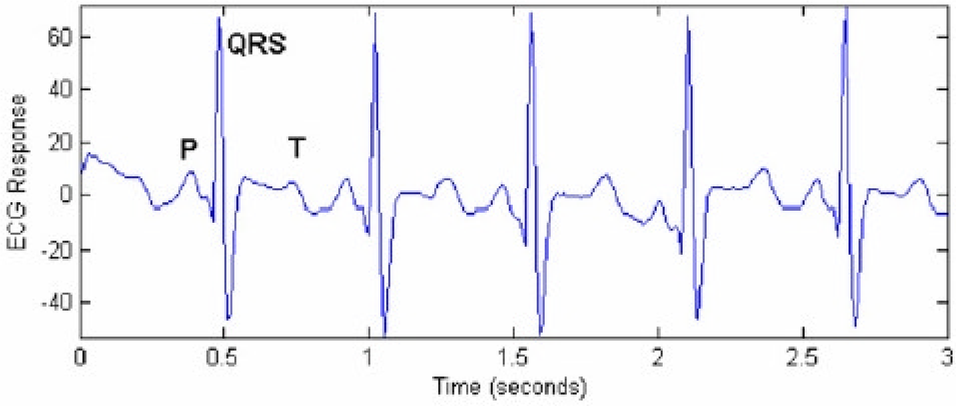
(a) use complex mother wavelets for complex signals.

(b) Mother wavelet resembles the general shape of the signal to be analyzed

In practice, discrete wavelet transforms dealing with discrete signals are implemented by using digital low-pass and high-pass filters to decompose the signal into a series of "approximation" and "detail" components

Wavelet transform – example 1

An example of WT of an ECG waveform

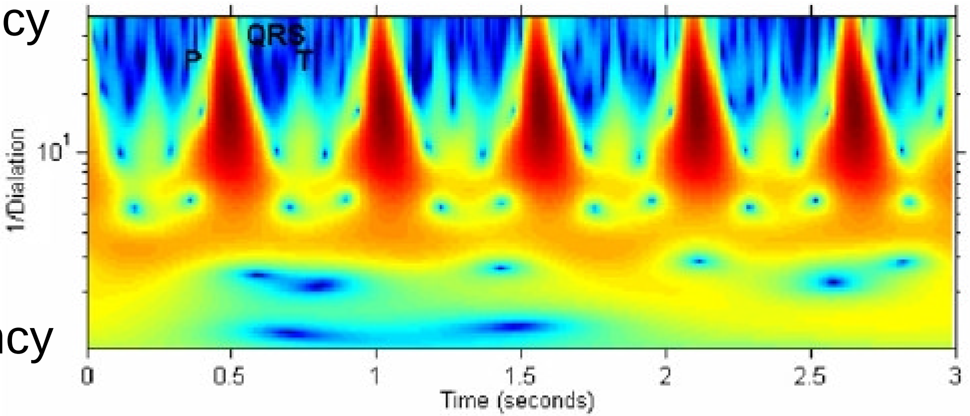


Original data

High frequency

scale

Low frequency



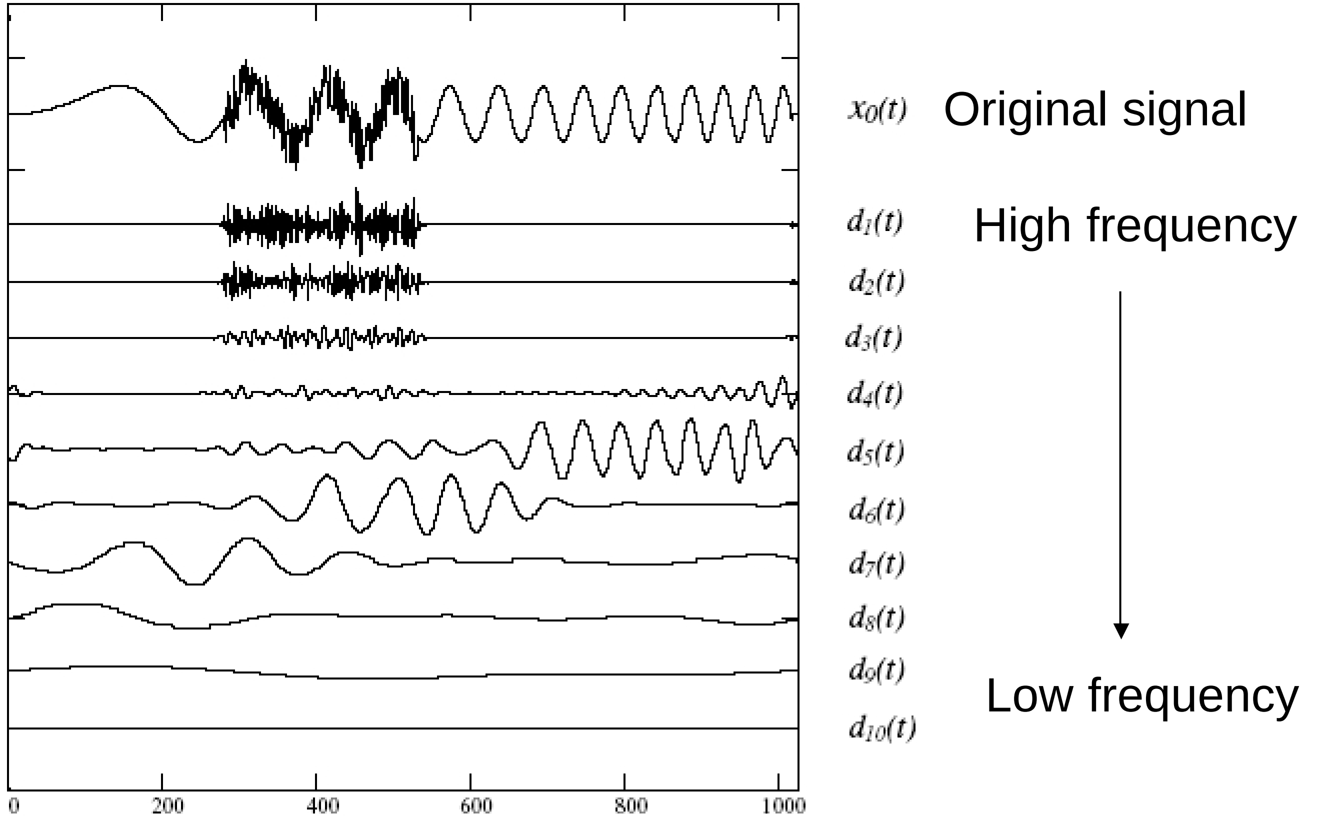
WT coefficients

Time

(Morlet wavelet)

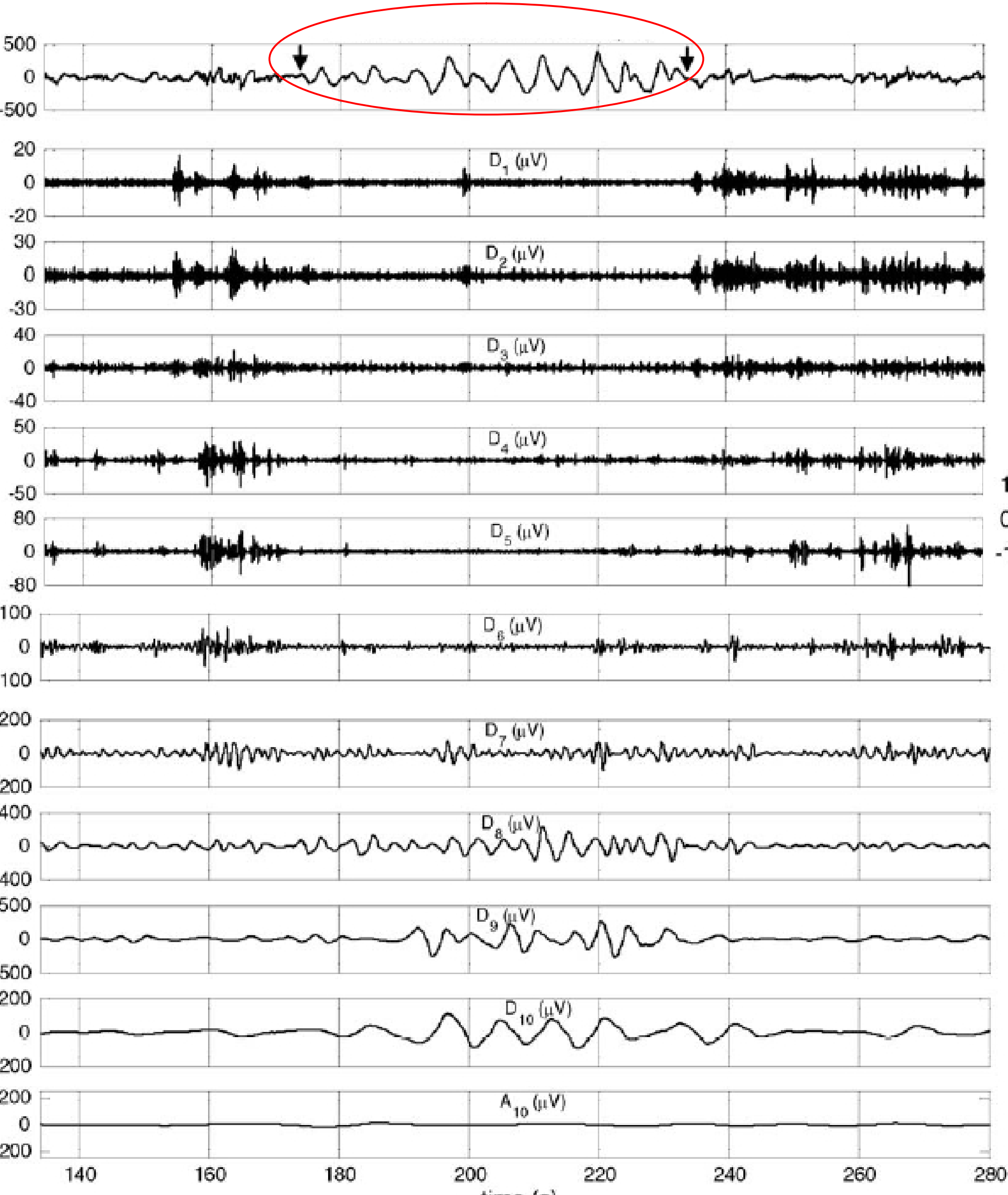
Wavelet transform – example 2

Decomposition of a chirp signal containing a short burst of noise into different levels of details d_1-d_{10}



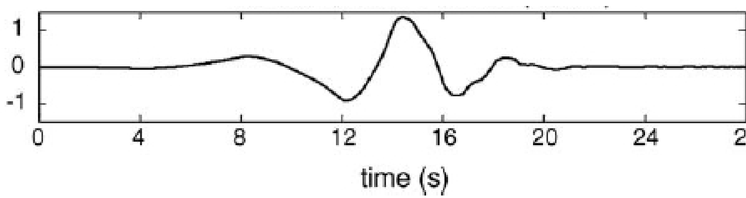
(Daubechies order 20 wavelet)

Example 3: analysis of electrooculogram (EOG) using wavelet



EOG – slow eye movements

Mother wavelet – Daubechies order 4 (db4)



Most slow eye movements signal is in details 8-10

Wavelet transform – comments

- Time-frequency decomposition of input signal
- Both short duration, high frequency and longer duration, lower frequency information can be captured simultaneously
- Particularly useful in the analysis of transient, aperodic and non-stationary signals
- Variety of wavelet functions is available, which allows signal processing with the most appropriate wavelet
- Applications of WT in biomedical signal and image processing: filtering, denoising, compression, and feature extraction

Artificial neural networks (ANN)

The human brain is a complex, non-linear, highly parallel information processing system

The human brain consists of 100 billions of brain cells (neurons) that are highly interconnected

ANNs are computational methods inspired by the formation and function of biological neural structures

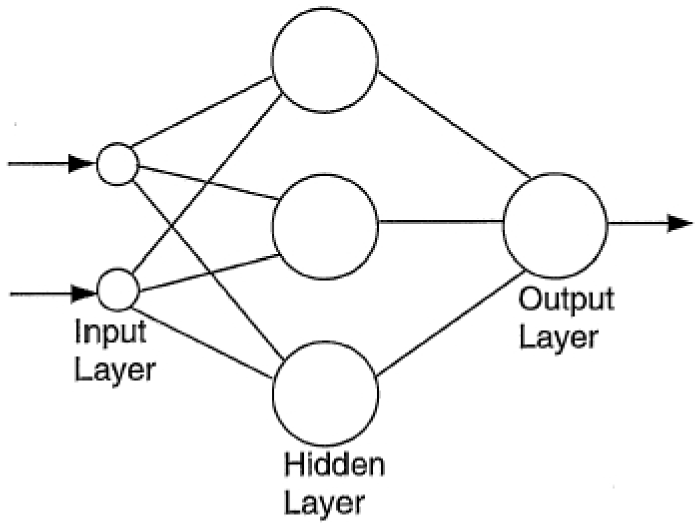
ANNs consist of much less number of neurons

ANNs are designed to learn from examples to recognize certain inputs and to produce a particular output for a given input

ANNs are commonly used for pattern detection and classification

Artificial neural networks

Simple example of a multilayer neural network



Input layer: number of neurons equals to number of inputs

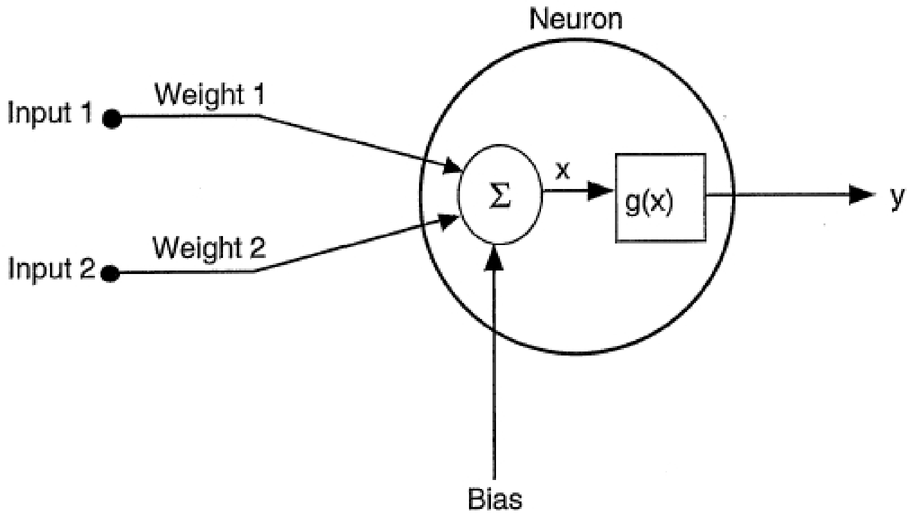
Output layer: number of neurons equals to number of outputs

There can be an arbitrary number of hidden layers, each of which can have an arbitrary number of neurons

The connections between neurons are represented in “weights”

Artificial neural networks

Relationship between the inputs and the output of a neuron



$$x = bias + \sum_i input_i \cdot weight_i$$

$$y = g(x)$$

The function g(x) is called activation function to mimic the excitation of a biological neuron. Some examples of activation function include

Threshold function (T = threshold)

$$y = g(x) = \begin{cases} 1 & \text{if } x > T \\ 0 & \text{if } -T < x < T \\ -1 & \text{if } x < -T \end{cases}$$

Sigmoid function

$$y = g(x) = \frac{1}{1+e^{-x}}$$

Artificial neural networks

Training of the ANN (supervised learning)

• The goal of training is to teach it how to determine the particular class each input pattern belongs to

• The weights are initialized by random numbers

• Each example with a known class is fed into the ANN and the weights are updated iteratively to produce a better classification

• In backpropagation neural networks, the weights of the neural network are updated through the propagation of error (difference between the output and the target) from the output backward to the neurons in the hidden layers and input layers

• The training process terminates when the total error at the output no longer improves or a preset number of iterations has been passed

• Once the weights are found, the ANN is uniquely defined and ready to be used to analyze unknown data inputs

Note that ANNs require many training examples to learn a pattern. If the number of examples is not sufficient for a problem $\Rightarrow$ overfit the training examples without learning the true patterns

Artificial neural networks

- ANNs do not need any assumption about the distribution of data
- ANNs are nonlinear, therefore are suitable for solving complex nonlinear classification and pattern recognition problems
- ANNs with <u>unsupervised learning</u> (the outputs for given input examples are not known) have also been used in biosignal processing (ex. self-organizing feature maps networks)

References

- Biomedical Signal and Image Processing, by Kayvan Najarian and Robert Splinter
  - Ch5: Wavelet Transform
  - Ch7.7: Neural Networks
- Biomedical Signal and Image Processing, by Kayvan Najarian and Robert Splinter
  - Ch5: Wavelet Transform

It suffices to stand in awe at the structure of the world, insofar as it allows our inadequate senses to appreciate it.

- *Albert Einstein*