T-61.181 Biomedical Signal Processing

# EEG Signal Processing

*Jan-Hendrik & Jan*

7th October, 2004

# Goals

- Extraction of clinically valuable information.
- Facilitating visual inspection.
- Extracting relevant features for classification tasks.
- Automation of standard analysis.
- Artifact removal.
- Understand the underlying mechanism.

# Content

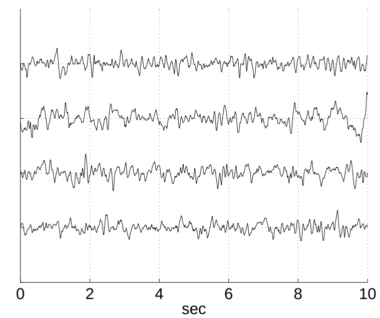
- Modeling EEG Signals
  - Stochastic vs. Deterministic
  - Gaussianity
  - Stationarity
  - Linear Models
  - Nonlinear Model
- Artifact Removal
  - Different Types

## Stochastic vs. Deterministic Signals

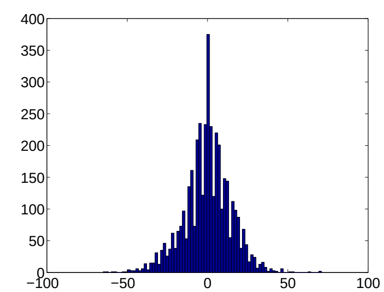
- Depends on the level of modeling.
- Even if the pure biological EEG source is deterministic, Amplifier, Digitalization add noise.
- Finding a quantitative answer to the question (DVV), but issue is not settled.
- Similar concepts.
- Seizure studies with nonlinear dynamic systems assumption and descriptors measuring ”chaoticness”.

# Stochastic Models

- Of which form is the joint distribution $p(\mathbf{x};\boldsymbol{\theta}) = p(x(0), \cdots, x(N-1);\boldsymbol{\theta})$ ?
- Nonparametric Approach:
  - Compute Amplitude histrogram
    Problem: one realization $\rightarrow$ stationarity, ergodicity: (ensemble mean = sample mean)
  - Guess the structure
- Parametric Approach:
  - Assume a parametric Form a priori possibly based on physiological insights.
  - Use data to estimate parameter $\boldsymbol{\theta}$.
- Ever-changing properties of the EEG require a highly complex PDF to account for different brain states




$\Downarrow$



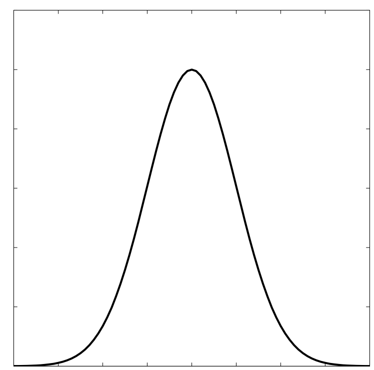
# Hopefully everything is Gaussian

- Using Gaussian PDF as model is from an engeneering point attractive.

$$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \mathbf{C}) = (2\pi)^{-N/2}|\mathbf{C}|^{-1} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{m}_x)^T \mathbf{C}^{-1}(\mathbf{x}-\mathbf{m}_x)\right)$$

$$\mathbf{m}_x = E[\mathbf{x}] \qquad \mathbf{C}_x = E\left[(\mathbf{x}-\mathbf{m}_x)(\mathbf{x}-\mathbf{m}_x)^T\right]$$

- Plausibility: EEG results form summation of a large number of individual oscillators
  $\rightarrow$ allows Central limit theorem
  But they are not independent as required of CLT



- 

| Gaussian | non Gaussian |
|---|---|
| synchronized activities | asynchronous firing |
| alpha rhythm, deep sleep | mental tasks, REM |

- Statistical tests for Gaussianity rely on strong assumptions themself.

- 90% of all one second intervals could be concidered Gaussian

## Covariance matrix

- Correlation function: $r_x(n_2, n_1) = r_x(n_1, n_2) = E\left[x(n_1)x(n_2)\right]$
- Correlation matrix:

$$\begin{aligned}
\mathbf{R}_x &= E\left[\mathbf{x}\mathbf{x}^T\right] \\
&= \begin{bmatrix} r_x(0,0) & r_x(0,1) & \cdots & r_x(0,N-1) \\ r_x(1,0) & r_x(1,1) & \cdots & r_x(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(N-1,0) & r_x(N-1,1) & \cdots & r_x(N-1,N-1) \end{bmatrix}
\end{aligned}$$

- Covariance matrix: $\mathbf{C}_x = E\left[(\mathbf{x}-\mathbf{m}_x)(\mathbf{x}-\mathbf{m}_x)^T\right] = \mathbf{R}_x - \mathbf{m}_x\mathbf{m}_x^T$
- In the Gaussian model the covariance matrix contains the essential information on the signal properties.
- $\mathbf{C}_x$ is without assumptions on its form difficult to estimate form one signal realization $\rightarrow$ stationary process, slowly changing correlation

# Stationarity

- Statistical properties are time invariant.
- *strictly stationary:*
  $\forall h \in \mathbb{R} : p(x(0), \ldots, x(N-1)) = p(x(0+h), \ldots, x(N-1+h))$
- *wide-sense stationary*:
  $m_x(n) = m_x$ and $r_x(k) = E\left[x(n)x(n-k)\right]$
- *strict stationary* $\Longrightarrow$ *wide-sense stationarity*
- For Gaussian: *wide-sense stationarity* $\Longrightarrow$ *strict stationary*
- $\mathbf{R}_x = \begin{bmatrix} r_x(0) & r_x(-1) & \cdots & r_x(-(N-1)) \\ r_x(1) & r_x(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_x(-1) \\ r_x(N-1) & \cdots & r_x(1) & r_x(0) \end{bmatrix}$
  is symmetric ($r_x(1) = r_x(-1)$) and Toeplitz

## Power Spectral Density (PSD)

- $S_x(e^{\jmath\omega}) = \sum_{k=-\infty}^{\infty} r_x(k) e^{-\jmath\omega k}$
- Related to Gaussian distribution: $S_x(e^{\jmath\omega})$ relies only on $r_x(k)$
- Assumes *stationarity*.
- Used for normal *spontaneous* activity, but only for *short* intervals.
- Can be efficiently estimated via FFT: $\hat{S}_x(e^{\jmath\omega}) = \frac{1}{N}|X(e^{\jmath\omega})|^2$

## Non-stationarity

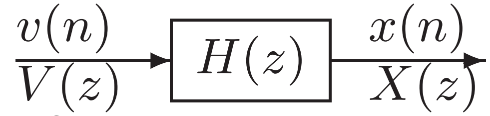
- Mean, correlation function and higher-order moments are time varying.
- Major EEG non-stationarities:
  1. Slow time-varying properties:
     e.g.: gradually changing wakefulness → $\alpha$-rhythm varies slowly
     - Apply analysis to consecutive overlapping, "sliding" windows (stFFT)
     - Use parametric approaches and adaptive filter
  2. Abruptly changing activity:
     e.g.: closing eyes
     - Decompose signal into variable length, quasistationary segments - Spectral analysis of these segments
  3. Transient waveforms:
     e.g.: K-complex, vertex waves, spikes
     - Event detection
     - No spectral analysis but characterized by wavefrom parameter (amplitude & duration)
     - Wavelet analysis (convolute parameterized carrier wave with signel)

## Non-Gaussian signals

- Study higher-order moments of the univariate amplitude distribution $E\left[(x(n)-m_x)^k\right], k=3,4,\cdots$
- *skewness* (k=3): degree of deviation from symmetry of a Gaussian PDF
- *kurtosis* (k=4): peakedness of PDF near $m_x$
- difficult to estimate because prone to outlier
- *Bispectrum:*
  - two-dimensional Fourier transform of $c_x(k_1,k_2)=E\left[x(n)x(n-k_1)x(n-k_2)\right]$
  - Displays PSD as function of two frequencies (interrelations between frequencies)
  - Degree of Gaussianity

# Linear stochastic models

- *Phenomenological* model since no prior anatomical or physiological information is incorporated.
- Not explaining the underlying mechanisms.
- Clinically useful model parameters.
- Signal is composed of different narrow-band components.
- Computationally efficient.
- Deviation between AR model and signal $\rightarrow$ epilepsy.
- EEG simulator.




$\sigma_v^2$ $\quad$ $S_x(z) = H(z)H(z^{-1})\sigma_v^2$

- EEG as the output of a linear system driven by (Gaussian) *white noise.*
- The filter $H(z)$ spectrally shapes the noise.
- System parameter estimated by fitting model to signal using (MSE).
- Parameter are useful features for classification.

## ARMA

- $x(n) = \underbrace{-\sum_{k=1}^{p} a_k x(n-k)}_{\text{AR}} + b_0 v(n) + \underbrace{\sum_{l=1}^{q} b_l v(n-l)}_{\text{MA}}$
- $z$-Transfrom:

$$H(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + \cdots + b_q z^{-q}}{1 + a_1 z^{-1} + \cdots + a_p z^{-p}}$$

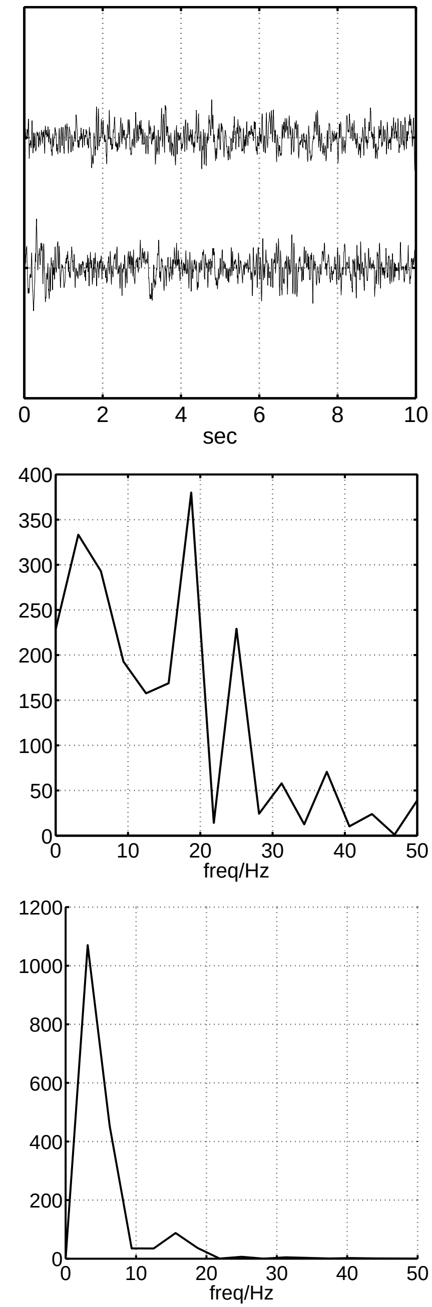
- Knowing the parameter $a_1, \cdots, a_p$, $b_0, \cdots, b_q$ the power spectrum is

$$\begin{aligned} S_x(e^{\jmath\omega}) &= |H(e^{\jmath\omega})|^2 \sigma_v^2 \\ &= \left| \frac{b_0 + b_1 e^{-\jmath\omega} + \cdots + b_q e^{-\jmath\omega q}}{1 + a_1 e^{-\jmath\omega} + \cdots + a_q e^{-\jmath\omega p}} \right|^2 \sigma_v^2 \end{aligned}$$

- Main characteristics: roots (spectral valleys) and poles (spektral peaks)

## Auto Regression (AR)

- $q = 0, b_0 = 1 \rightarrow$
  $x(n) = -\sum_{k=1}^{p} a_k x(n-k) + v(n)$
- All pole filter. $S_x(e^{\jmath\omega}) = \frac{1}{|A(e^{\jmath\omega})|}\sigma_v^2$
- $a_k$ are compact description of eeg stages; contain spectral information.
- $p$ determines number of peaks presented in AR-PSD.
- $a_k$ are obtained by solving linear matrix equation $\mathbf{X}_i \mathbf{a} = \mathbf{x}_o$.
  MSE solution is $\mathbf{a} = (\mathbf{X}_i^T \mathbf{X}_i)^{-1} \mathbf{X}_i^T \mathbf{x}_o$.
- Less computation then general ARMA.



## AR

- *Time-varying AR modeling* for nonstationary signals:

  $x(n) = -\sum_{k=1}^{p} a_k(n)x(n-k) + v(n)$

  $a_k(n)$ have to be estimated by an adaptive algorithm.

- Poisson distributed $\delta$-impulse as input to model *transient* events.

- *Multivariate AR models:*

  Study spatial interaction between different regions of the brain.

  $\mathbf{x}(n) = -\sum_{k=1}^{p} \mathbf{A}_k \mathbf{x}(n-p) + \mathbf{v}(n)$

  $\mathbf{v}$ : uncorrelated channel noise $\sigma_{v_1}^2, \cdots, \sigma_{v_M}^2$

  $\mathbf{A}_k$ : $(M \times M)$ for $M$ channels

  $\rightarrow$ *spatial correlation*

$$\begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \end{array}\begin{array}{c} \begin{array}{ccc} c_1 & c_2 & c_3 \end{array} \\ \left[\begin{array}{ccc} a_{k_1} & \square & \square \\ \square & a_{k_2} & \square \\ \square & \square & a_{k_3} \end{array}\right] \end{array} = \mathbf{A}_k$$

# Nonlinear EEG Models

- Goal: Understanding the underlying generation process.
- Either strong regularization or based on neurophysiological facts, reflecting how different neuron populations interact
- Nonlinear Model of one cortical neuron population in early 1970's. Later extended to multiple coupled populations for the purpose of seizure detection.
- *"Usefulness for the design of signal processing methods yet to be demonstrated."*

# One Neuron Population

Two interacting subpopulations:

phyramidal cells & positive and negetive feedback interneurons.

Average pulse density $\rightarrow$ LTI systems:

$h_e(t) = Aate^{-at}u(t), \qquad h_i(t) = Bbte^{-bt}u(t)$

$A, B$: max amplitude; $a, b$ *lumped-parameter* (dendrite average time delay)

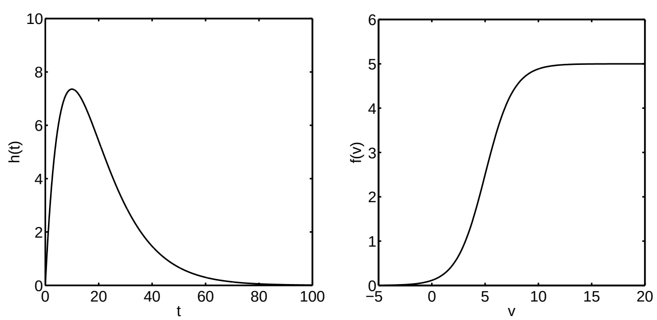
$\bigoplus$: cell soma/axon hilloc; $C_k$: av. number of synaptic contacts

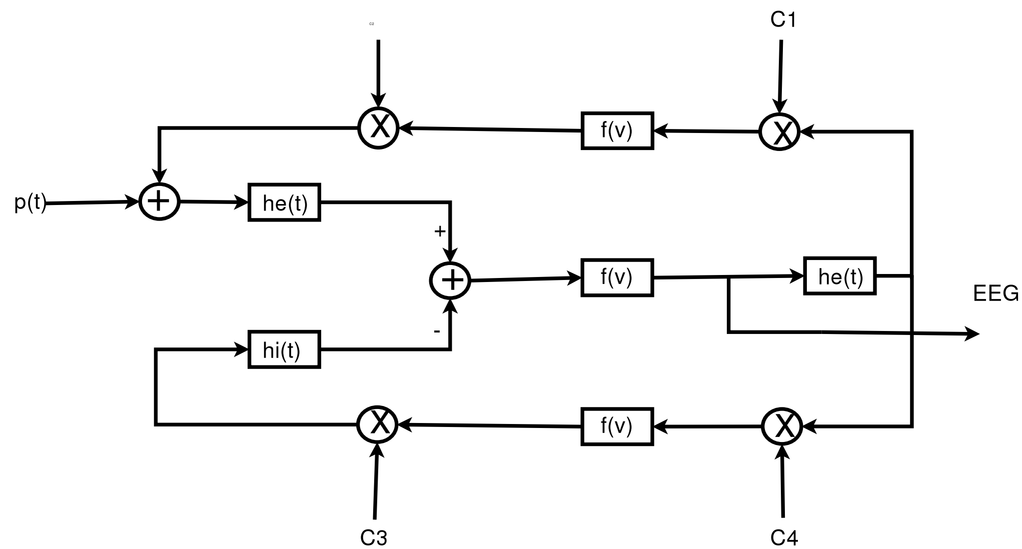
$p(t)$ neighboring populatoins (stochastic process)

av. post.-potential $\rightarrow$ pulse density

Static nonlinearity:

$f(v) = \frac{2e_0}{1+e^{r(v_0-v)}}$



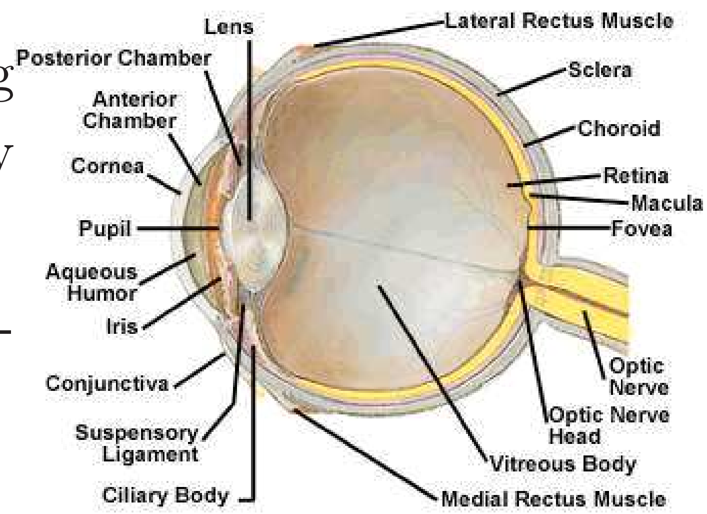



# Artifacts in EEG

- Can be of *physiological* or of *technical* origin.
  Easier to deal with because of there different nature.
- 50Hz alternating current; digitalization.
- Eye movement & blinks; cardiac activity; muscle activity; respiration; skin potential.

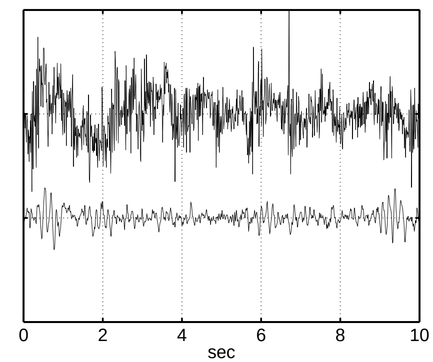
## Eye movements

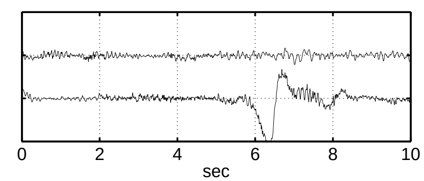
- EOG: potential difference between *cornea* and *retina*.
- Voltage amplitude proportional to angle of gaze.
- Proximity of sensor to the eyes. Direction of the movement.
- Can be mixed with slow EEG.
- Prominent in REM sleep.
- Eyelid blinks → abruptly changeling waves (high frequency), substantially larger than background.
- "Pure" EOG reference for artifact cancellation.
  $\hat{s}(t) = x(t) - \hat{n}(t) = s(t) + (n(t) - \hat{n}(t))$

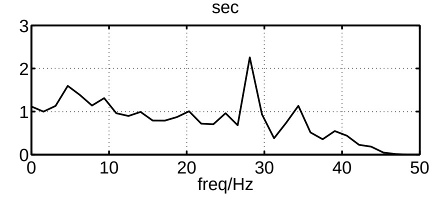


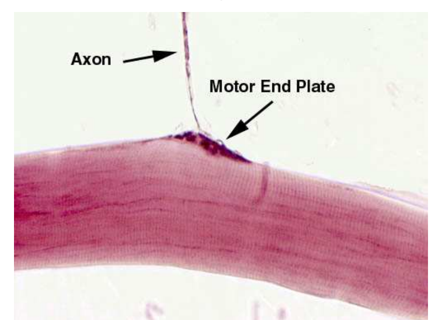
# Muscle activity

- Measured with EMG.
- Recordings during wakefulness.
- Tongue movement; swallowing, grimacing, chewing,
- Shape depends on the degree of muscle contraction:
  - weak contraction → low-amplitude spike train.
  - increasing contraction → decrease in inter-spike distance (colored noise)
- Occurs less in sleep
- Contrast to eye movements, the spectral properties overlap with beta band (15-30Hz).
- Difficult to get pure reference signal.



## Cardiac activity

- The Amplitude is usually low on the scalp (1-2 $\mu V$) compared to EEG ($20 - 100\mu V$).
- Still hampers certain electrodes.
- Effect depends on the probands anatomy.
- Regular pattern of heart beat helps revealing it. (Arrhythmias)
- Can be mistaken for epileptic waveforms.
- Reference ECG for cancellation.

## Electrodes and equipment

- Moving electrodes change DC contact potential ("electrode-pop"). Abrupt change of base line level, followed by gradual return to original baseline.
- Misinterpreted as sharp waves.
- Amplifier noise
- Amplitude clipping by A/D converter
- Insufficient isolation $\rightarrow$ 50/60Hz power line.

## Artifact Processing

- Rejection or Cancellation.
- $x(t) = s(t) + n(t)$ vs. $x(t) = s(t)n(t)$
- Additive noise is preferred due to simplicity and optimal estimation techniques.
- Linear filtering e.g.:
  - low-pass filter for EMG activity
  - but bursts of EMG spikes could be smoothed into alpha waves
  - sharp waves get distorted