Digital Biosignal Processing Notes

December 2020

1 Sampling of continuous time signals

- Sampling can be thought of as multiplication by the impulse train modulator $s(t) = \sum_{-\infty}^{\infty} \delta(t - nT)$

2 Frequency domain representation of sampling

- Fourier transform of impulse train modulator is $S(J\Omega) = \frac{1}{T}\sum_{-\infty}^{\infty} \delta(\Omega - k\frac{2\pi}{T})$ where $\frac{2\pi}{T}$ can be written as Ω_s
- $X_s(J\Omega) = X_c(J\Omega) * S(J\Omega)$ where $*$ is the convolution operator $\rightarrow X_s(J\Omega) = \frac{1}{T}\sum_{-\infty}^{\infty} X_c(J(\Omega - k\Omega_s))$ by the sifting property of the delta function $(x(t) * \delta(t - \Delta) = x(t - \Delta))$
- Sampling corresponds to periodisation in Fourier domain

3 Nyquist-Shannon theorem

- $\Omega_s \geq 2\Omega_n$ for the signal to be reconstructable from the samples

4 Reconstruction of sampled signals

- If we choose sufficient sampling frequency we can go back using sinc functions which are equivalent to rect windows in the frequency domain
- $h_r(t) = \frac{\sin(\frac{\pi}{T}t)}{\frac{\pi}{T}t}$
- Higher sampling rates allow us to use non-ideal filters when reconstructing from the sampled signal

5 Discrete time signals

- Discrete frequencies of sinusoids belong to the interval $\omega \ \epsilon \ [-\pi, \pi]$ or $[0, 2\pi]$
- Given the sinusoidal sequence $x\,[n] = A\cos(\omega_0 n + \phi)$, the frequencies $\omega_0 = 0$ and $\omega_0 = 2\pi$ are indistinguishable
- Properties of Discrete time signals
  - Time invariance: if x is the output sequence of y then $x_1\,[n] = x[n-n_0]$ produces at output $y_1\,[n] = y[n-n_0]$
    - Down sampler $y[n] = x[Mn]$ is not time invariant because if you shift the signal, the samples that you are keeping may be completely different
  - Memoryless, Linear, Causality, Stability (bounded input / bounded output BIBO)

6 LTI signals

- Properties
  - Commutative, distributive, associative, stability BIBO
  - LTI is stable if and only if impulse response is absolutely summable $B_y = \sum_{-\infty}^{+\infty} |h\,[n]| < \infty$
  - LTI is causal if and only if $h\,[n] = 0$ for all $n < 0$
  - LTI is FIR (Finite Impulse Response) if impulse response has finite bounds
    * FIR LTI filters are stable
    * They are not necessarily causal
- A complex exponential going into an LTI system is just being multiplied by a complex value
- Pure harmonics are eigenvectors for LTI systems
  - If $x[n] = e^{J\omega_0 n}$, then $y[n]$ will be the input x convolved with the impulse response, so $\sum_k x[k]h[n-k]$ which is equivalent to the impulse response convolved with the input; $\sum_k h[k]x[n-k]$
  - Substituting in $x[n] = e^{J\omega_0 n}$ we obtain $\sum_k h[k]e^{J\omega_0(n-k)} = e^{J\omega_0 n}\sum_k h[k]e^{-J\omega_0 k} = x[n] * H(e^{J\omega_0})$
  - $H(e^{J\omega_0})$ depends on the impulse response and ω_0
  - $H(e^{J\omega_0})$ is the eigenvalue for the system
  - $x[n]$, a pure harmonic in the case, is shown to be an eigenvalue for this system
- Complex exponentials are the eigenfunctions of LTI systems
  - A frequency response shows what is the factor the multiplies an input of a certain frequency
- All frequency responses of discrete time LTI systems are periodic
- LTI system applied to $x(n) = A\cos(\omega_0 n + \phi)$
  - $H(e^{J\omega_0}) = A|H(e^{J\omega_0})| * \cos(\omega_0 n + \phi + \angle H(e^{J\omega_0}))$
  - Sinusoidal sequences are only modified in amplitude and phase when processed by an LTI system $H(e^{J\omega_0})$ can be fully determined by its behaviours from $-\pi$ to π

7 Moving average filter

- $h[n] = \frac{1}{n_1+n_2+1}\sum_{k=-n_1}^{n_2}\delta[n-k]$
- $H(e^{J\omega}) = \sum_{n=-n_1}^{n_2} h[n]e^{J\omega n}$
- If $n_1 = 0$ the filter is causal
  - $H(e^{J\omega})$ can be written (using sum of geometric series formula and Euler's formula) as $\frac{1}{m_2+1}\frac{\sin[\omega(n_2+1)/2]}{\sin(\omega/2)}e^{-J\omega n_2/2}$
- Moving average filters are low pass - they get rid of high frequency components

8 Discrete Time Fourier Transform

- Definition of discrete time Fourier transform $X\left(e^{j\omega}\right) = \sum_{n=-\infty}^{\infty} x\left[n\right] e^{-j\omega n}$
- Continuous function of frequency ω
- Periodic in frequency with period 2π
- Corresponds to continuous FT of sampled original signal (if sampling rate > Nyquist) with the frequency axis normalised to π instead of half the sampling frequency
- $x[n]$ can be recovered using the inversion formula $x\left[n\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X\left(e^{j\omega}\right) e^{j\omega n} d\omega$
- If $x[n]$ is finite length then we don't need the DTFT to reconstruct $x[n]$ exactly
- Properties of DTFT (time reversal, Parseval's theorem, time and frequency shift, linearity, convolution property)

9 Difference equations

- Derivatives in discrete time are just subtractions between consecutive samples and can be written in the form $\sum_{k=0}^{N} a_k y[n-k] = \sum_{m=0}^{M} b_m y[n-m]$
- We can take the FT of both sides and rearrange to get $H(e^{j\omega}) = \frac{\sum_{m=0}^{M} b_m e^{-j\omega m}}{\sum_{k=0}^{N} e^{-j\omega k}}$
- Accumulator (discrete version of integration)
  - Difference equation is $y[n] - y[n-1] = x[n]$
  - $H(e^{j\omega}) = \frac{1}{1-e^{-j\omega}}$

10 Difference between discrete time transform of discrete time signal and continuous time transform of sampled continuous signal

- A sampled continuous time signal is still defined in the continuous time domain
- $\mathcal{F}\{x_s(t)\} = \sum_{n=-\infty}^{+\infty} x[n]e^{-J\Omega nT}$
- $X(e^{-J\omega}) = \sum_{n=-\infty}^{+\infty} x[n]e^{-J\omega n}$
- These two are the same if $\omega = \Omega_s T$
- $\Omega_s = \frac{2\pi}{T}$ so we can see that the periodicity of ω is 2π

11 Changing sampling rate

- Can be done directly in discrete time
- Down sampling
  - $x_d[n] = x[nM]$
- Up sampling

12 Discrete Fourier Transform

- $X[k] = \sum_{-\infty}^{\infty} x[n] e^{-J\frac{2\pi k}{N}n}$, $k = 0, \ldots, N-1$
- Discrete in both time and frequency domain
- Basically the DTFT made no longer redundant but kept invertible
- Zero padding the original time domain signal regulates the resolution of the signal returned by the DFT algorithm (otherwise it returns the minimum for invertibility)
- Increasing N is equivalent to 0 padding in the time domain

13 Power Spectrum Density

- PSD[k] $= \frac{1}{N}|X[k]|^2 = \frac{1}{N}\left|\sum_{n=0}^{N-1} x[n] e^{-J\frac{2\pi k}{N}n}\right|^2$
- Power of signal per unit frequency

14 Z transform

- Fourier vs Laplace transforms
  - Fourier transform $X(J\Omega) = \int_{-\infty}^{+\infty} x(t)e^{-J\Omega t}dt$
  - Laplace transform $X(s) = \int_{-\infty}^{+\infty} x(t)e^{-st}dt$ where $s = \sigma + J\Omega$
- Discrete time equivalent of Laplace transform = z transform
  - $X(z) = \sum_{-\infty}^{\infty} x[n]z^{-n}$
  - The DTFT is defined as the Laplace transform where $z = e^{J\omega}$ (|z| = 1)
    - This is the unit circle in the Im{z} / Re{z} plane
- Region of convergence of z transform
  - Converges when $|X(z)| = \left|\sum_{-\infty}^{\infty} x[n] z^{-n}\right| < \infty$
  - $\left|\sum_{n=-\infty}^{\infty} x[n] z^{-n}\right| \leq \sum_{n=-\infty}^{\infty} |x[n]| |z|^{-n}$
  - If converges at a certain z, it will converge on all other points with the same absolute value for z, around the unit circle in the z plane
- Properties of z transform
  - Time delay property: delay is multiplying by z^{-n_0} where n_0 is the delay
  - Convolution property: Convolution is z domain is multiplication in the discrete time domain and vice versa

15 LTI systems and the z transform

- $Y(z) = H(z)\,X(z)$ where $H(z)$ is the system function
- $H(z)$ calculation for a system defined with a difference equation
  - We can apply the z transform to a system defined by the difference equation $\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{M} b_k y[n-k]$
  - This gives $H(z) = \frac{\sum_{k=0}^{M} b_k z^{-k}}{\sum_{k=0}^{N} a_k z^{-k}}$
- Recursive way to write difference equations
  - $y[n] - \sum_{k=1}^{N} a_k y[n-k] = \sum_{k=0}^{M} b_k x[n-k]$
  - This kind of system is used alongside the z transform to develop digital filters

16 Z transform and difference equations

- General difference equation $y[n] = \sum_{k=1}^{N} a_k y\,[n-k] + \sum_{k=0}^{M} b_x x\,[n-k]$
- The first summation is the "recursive part" that looks at the previous outputs
- Take z transform of whole equation and rearrange to find system function $H(z) = \frac{\sum_{k=0}^{M} b_k z^{-k}}{1-\sum_{k=1}^{N} a_k z^{-k}}$
- You need to choose the coefficients a and b and order M and N to obtain the correct system function
- Alternate useful form for H(z) is $H(z) = \frac{b_0 \prod_{k=1}^{M}(1-C_k z^{-1})}{a_0 \prod_{k=1}^{N}(1-d_k z^{-1})}$ where c_k are the zeroes and d_k the poles

17 System function and stability

- An LTI system is stable if and only if its impulse response is absolutely summable meaning $B_y = \sum_{-\infty}^{+\infty} |h\,[n]| < \infty$
- The transfer function of the system is $H\left(e^{j\omega}\right) = \sum_{-\infty}^{\infty} h\,[n]\,e^{-j\omega n}$
- Therefore the stability condition on the impulse response implies that the an LTI system is stable if $\left|H\left(e^{j\omega}\right)\right| < \infty$ which means that $|H(z)| < \infty$ for $z = e^{j\omega}$
- This shows that the system is stable if the system function does not have poles on the unit circle in the z plane

18 Finite Impulse Response (FIR) filters

- If N = 0, there is no recursive part and
  - $H(z) = \sum_{k=0}^{M} b_k z^{-k}$
  - $y\,[n] = \sum_{k=0}^{M} b_x x\,[n-k]$
  - Impulse response $h\,[n] = b_n\ ,\ n = \ 0, \ldots, M$
- FIR filters have no poles (z=0 is an exception) – they are inherently stable
- $y\,[n] = \sum_{k=0}^{M} h[k] x\,[n-k]$
- $Y\,[k] = H\,[n]\,X[k]$
- FIR can be designed to have an exact linear phase

19 Digital derivative filters

- 2 point difference (approximates 1^{st} derivative)
  - Linear phase filter, introducing delay of $\frac{1}{2}$ a sample and approximating the derivative operator
- 2^{nd} derivative
  - Linear phase filter introducing delay of 1 sample and approximating the 2^{nd} derivative operator

20 Group delay

- The group delay is the negative derivative of $\angle H\left(e^{j\omega}\right)$ with respect to ω
- If $H\left(e^{j\omega}\right) = e^{-j\omega\alpha}$ then the group delay grd $\left[H\left(e^{j\omega}\right)\right] = \alpha$
- Linear phase systems: $H\left(e^{j\omega}\right) = \left|H\left(e^{j\omega}\right)\right| e^{-j\omega\alpha}$, $|\omega| < \pi$
- Generalised form: $H\left(e^{j\omega}\right) = A\left(e^{j\omega}\right) e^{-j\omega\alpha + j\beta}$

21 Linear phase FIR filters

- Symmetry / Asymmetry often means it is a linear phase FIR filter
- For example, h[n] = h[M − n] is linear phase filter with group delay M/2

22 Infinite Impulse Response (IIR) Filters

- Using a recursive difference equation is often easier the convolution operation and enables you to apply the same system function
- Also, using convolution forces you to truncate the output
- If there is a recursive part in the difference equation then the filter has an infinite impulse response
- General expression of IIR: $y\left[n\right] = \sum_{k=1}^{N} a_k y\left[n-k\right] + \sum_{k=0}^{M} b_k x\left[n-k\right]$

23 Z transform of moving average filter

- MA filter can be written as $y[n] = \frac{1}{N}\sum_{i=0}^{N-1} x[n-i]$ where N is the length of the average
- $Y(z) = Z\left\{y[n]\right\} = \sum_{n=0}^{\infty} y[n]z^{-n}$
- Using the linearity of the z transform and the time shifting property of the z transform, $H(z)$ can be written as $\frac{1}{N}\sum_{i=0}^{N-1} z^{-i}$
- This can be further simplified using the formula for the sum of a geometric series, giving $\frac{1}{N}\frac{z^N - 1}{z^{N-1}(z-1)}$
- This shows that the filter has N-1 poles and zeros
  - Poles 0
  - Zeros $e^{2\pi jn/N}|n \in [1, N-1]$

24 Digital integrator

- Pole on unit circle – unstable
- Infinite sum of h[n] sums to infinity – unstable
- Unstable when frequency is 0Hz

25 Structures of IIR filters

- Block diagrams
- System function can also be represented as a cascade of 1^{st} order filters by factorising the numerator and denominator

26 Design of IIR filters

- Bilinear transformation
  - Allows to pass from the Laplace transform of a filter impulse response to the z-transform
  - s domain is Laplace domain and is defined in Ω
  - z domain is defined in ω
  - $s = \frac{2}{T_d}\left(\frac{1-z^{-1}}{1+z^{-1}}\right)$
  - $\Omega = \frac{2}{T_d}\tan\left(\frac{\omega}{2}\right)$

27 Discrete time random signals

- Model that describes signals statistically as opposed to on the basis of individual recordings
- A random process can be defined as a collection of random variables
- At each discrete time instant n, x_n is a sample / single realisation from a random process

28 Cumulative probability distribution function

- $F_{x_n}(\alpha, n) = P\{x_n \leq \alpha\}$ where n is a time instant and alpha is the independent variable

29 Probability density function

- $f_{x_n}(\alpha, n) = \frac{\partial F_{x_n}(\alpha,n)}{\partial \alpha}$

30 Mean of a random variable

- Mean of a random variable is the expected value, defined as $m_{x_n} = \mathcal{E}\{x_n\} = \int_{-\infty}^{+\infty} \alpha f_{x_n}(\alpha, n)\ d\alpha$
- $\mathcal{E}\{g(x_n)\} = \int_{-\infty}^{+\infty} g(\alpha) f_{x_n}(\alpha, n)\ d\alpha$

31 Joint cumulative probability distribution function

- $F_{x_n}(\alpha_1, \alpha_2, n, m) = P\{x_n \leq \alpha_1, x_m \leq \alpha_2\}$

32 Joint probability density function

- $f_{x_n}(\alpha_1, \alpha_2, n, m) = \frac{\partial^2 F_{x_n}(\alpha_1, \alpha_2, n, m)}{\partial \alpha_1 \partial \alpha_2}$

33 Statistically independent

- Means that $F_{x_n, x_m} = F_{x_n} F_{x_m}$
- Joint probability distribution function is equal to 2 single pdf's multiplied together
- Statistically independent signals are uncorrelated but not all uncorrelated signals are statistically independent

34 Autocorrelation

- $\phi_{xx}[n, m] = \mathcal{E}\{x_n, x_m\}$

35 Stationary random processes

- A complete statistical characterisation of a random process would imply specification of all possible pdf's (any number / combination of alphas), which is impossible
- A random process is stationary of order N if its joint pdf's up to order N do not depend on time shifts
- A random process is strict-sense stationary if it is stationary for any order N

36 Stationary of order 2

- $m_{\vec{x}_n} = \mathcal{E}\left\{\vec{x}_n\right\} = m_x$
- Mean does not depend on time
- Autocorrelation function $\phi_{xx}[n, m]$ depends only on the time difference

37 Wide sense stationarity (WSS)

- A weaker condition the stationarity of order 2
- A process for which the following properties hold:
  - $m_{\vec{x}_n} = \mathcal{E}\left\{\vec{x}_n\right\} = m_x$
  - $\phi_{xx}[n, n+l] = \phi_{xx}[l]$
- ϕ_{xx} of a WSS process is symmetric
- If $l = 0$, ϕ_{xx} is the variance of the process

38 Time average of random processes

- In practice, we have a limited observation view of the random process (on realisation and limited samples)
- Move from vertical comparisons (between realisations) to horizontal comparisons (within the same realisation)
- $\langle x_n \rangle = \lim_{L \to \infty} \frac{1}{2L+1} \sum_{n=-L}^{L} x_n$

39 Autocorrelation sequence

- $\langle x_n, x_{n+m} \rangle = \lim_{L\to\infty} \frac{1}{2L+1} \sum_{n=-L}^{L} x_n x_{n+m}$

40 Ergodicity

- Means vertical and horizontal analysis equivalent
- Entire information about process contained in 1 realisation
- Ergodicity is described specifically for a property
- Mean horizontally the same as vertical average $\to$ mean is ergodic
- To apply Fourier analysis to random signals, we need them to be WSS and ergodic for the mean and autocorrelation

41 Autocorrelation of ergodic process

- The autocorrelation function of 0 for an ergodic process corresponds to the power of each of the process realisations
- If the process also has 0 mean, we have:
  - $\lim_{L\to\infty} \frac{1}{2L+1} \sum_{n=-L}^{L} x^2[n] = \phi_{xx}[0] = {\sigma_{x_n}}^2$

42 In practice, infinite time samples are impossible

- $\hat{m}_x = \frac{1}{L} \sum_{n=0}^{L-1} x[n]$ approximates the mean
- As L is not infinite, $\frac{1}{L} \sum_{n=0}^{L-1} x_1[n]$ and $\frac{1}{L} \sum_{n=0}^{L-1} x_2[n]$ will be different (x_1 and x_2 are two different realisations of the random process)
- $\hat{m}_x$ itself is a random variable, but on average it is the real value (the estimate is unbiased)
- ${\sigma_{\hat{m}_x}}^2 = \frac{{\sigma_x}^2}{L}$

43 The power spectrum density of a WSS random process is the DTFT of the autocorrelation function

- $\Phi_{xx}\left(e^{J\omega}\right) = \mathcal{E}\left\{\lim_{L\to\infty} \frac{1}{2L+1} \left|X_L\left(e^{J\omega}\right)\right|^2\right\}$
- $\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{xx}\left(e^{J\omega}\right) d\omega = \mathcal{E}\left\{\lim_{L\to\infty} \frac{1}{2L+1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|X_L\left(e^{J\omega}\right)\right|^2 d\omega \right\}$
- By Parseval's relation $= \mathcal{E}\left\{\lim_{L\to\infty} \frac{1}{2L+1} \sum_{n=-L}^{L} x^2[n]\right\} =$ power
- If a process is ergodic for autocorrelation, we can use the time average to compute DTFT from the autocorrelation, but for this infinite samples are needed

44 Quantisation

- Error due to quantisation $e[n] = \hat{x}[n] - x[n]$
- If X_m is the dynamic of the signal, $\Delta = \frac{X_m}{2^B}$ where B is the number of bits
- The variance ${\sigma_e}^2 = \frac{\Delta^2}{12} = \frac{2^{-2B}{X_m}^2}{12}$
- If process is ergodic and stationary, ${\sigma_e}^2$ is the power of the noise
- $\Phi_{ee}\left(e^{J\omega}\right) = {\sigma_e}^2 = \frac{2^{-2B}{X_m}^2}{12}$ can be used to get the PSD if the quantisation noise

45 Assessing the properties of a random process that is limited in length

- With a limited number of samples, we can estimate the PSD
- $\widehat{\Phi}_{xx}\left(e^{J\omega}\right) = \frac{1}{2L+1}\left|X_L\left(e^{J\omega}\right)\right|^2$
- This estimate is referred to as a periodogram
- The estimated power spectrum is the DTFT of the autocorrelation function, even for estimates like the periodogram

46 Bias of the periodogram estimate

- $\mathcal{E}\left\{\widehat{\Phi}_{xx}\left(e^{J\omega}\right)\right\} = \mathcal{E}\left\{\frac{1}{2L+1}\left|X_L\left(e^{J\omega}\right)\right|^2\right\}$
- The estimate tends to the PSD of the windowed random signal
- $\mathcal{E}\left\{\widehat{\Phi}_{xx}\left(e^{J\omega}\right)\right\} = \Phi_{xx}\left(e^{J\omega}\right) * \left|W_{x+1}(e^{J\omega})\right|^2$
- $\left|W_{x+1}(e^{J\omega})\right|^2$ is the square magnitude of the window
- Windows can be rectangular, Hamming, Hanning, Nuttal etc.

47 Bartlett's method

- Method of averaged periodograms
- Reduces the variance of the periodogram in exchange for a reduction in resolution
- A final estimate of the spectrum at a given frequency is obtained by averaging the estimates from the periodograms (at the same frequency) derived from non-overlapping portions of the original series

48 Welch's method

- Extension of Bartlett's method
- Window type not necessarily rectangular
- Allows overlapping segments

49 Filtering random signals

- Assume WSS + ergodicity
- $\Phi_{yy}\left(e^{J\omega}\right) = \Phi_{xx}\left(e^{J\omega}\right) * \left|H\left(e^{J\omega}\right)\right|^2$

50 Filtering random signal with LTI filter

- $\Phi_{yy}\left(e^{J\omega}\right) = {\sigma_x}^2 * \left|H\left(e^{J\omega}\right)\right|^2$
- LTI filter can change the shape of the PSD and also change the correlation between samples

51 Linear predictor

- Prediction problem of order 1
  - Predict next value in a sequence $x[n]$
  - $\hat{x}[n] = \sum_{k=1}^{N} a_k x[n-k]$
  - $e[n] = x[n] - \hat{x}[n]$
  - Minimise squared error $\mathcal{E}\left\{e^2[n]\right\}$
  - Minimum found by solving the equation $\sum_{k=1}^{N} a_k \phi_{xx}[k-l] = \phi_x[k]$
  - This is N equations with N unknowns
  - You need to know the autocorrelation function $\phi_{xx}[l]$ for $l = 0, \ \ldots, \ N$
- The linear predictor assumes that the error is uncorrelated with the data
- The best prediction minimises the error
- Fully exploits the data and the error is the residue

52 Modelling

- Given the random process $x[n]$, if we can identify the transfer function $H\left(e^{J\omega}\right)$ that best models the process as the output of the LTI system with white noise as the input, we have obtained an estimate of the power spectrum of the process
- Autoregressive (AR) model of order N of the random process $x[n]$:
  - $x[n] = \sum_{k=1}^{N} a_k x[n-k] + u[n]$
- Power spectrum of a random process modelled by an AR model of order N: $\Phi_{xx}\left(e^{J\omega}\right) = {\sigma_u}^2 * \frac{1}{\left|1-\sum_{k=1}^{N} a_k e^{-j\omega k}\right|^2}$
  - The power spectral estimation corresponds to the optimal selection of the coefficients a_k

53 ARMA – autoregressive moving average models – can be approximated with sufficiently long AR