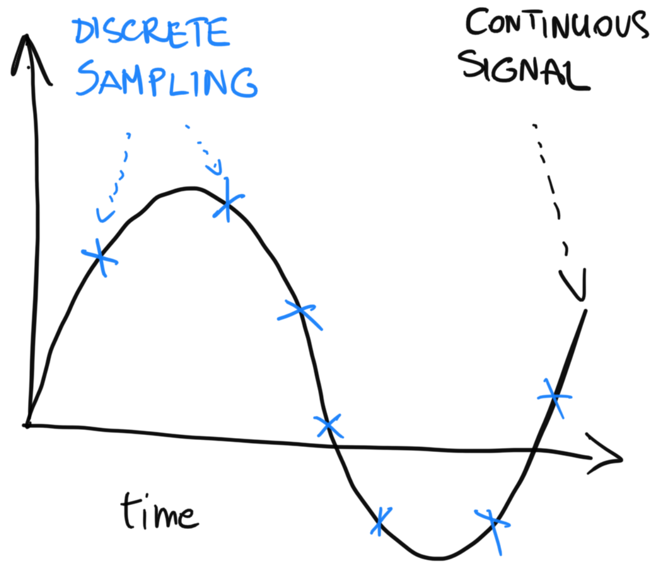
Mobile Health
Basics of Signal Processing

Cecilia Mascolo

UNIVERSITY OF CAMBRIDGE

The sensor data

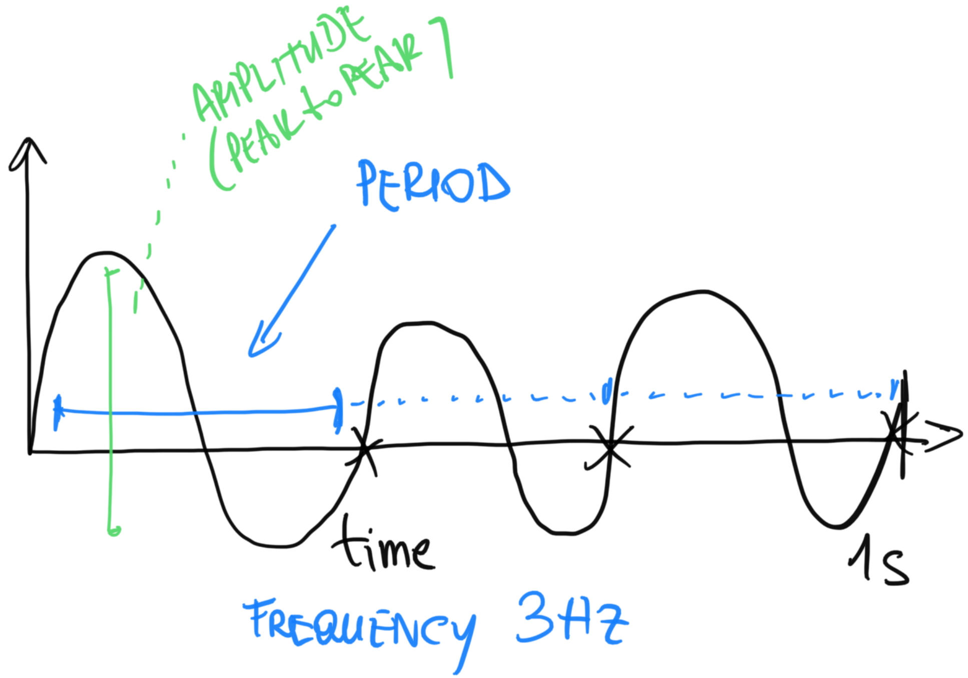
- Most of the sensor data collected by these devices' sensors is "time series data".
- A sensor is sampled at specific time intervals; data is discrete but could be seen as relatively "continuous".
- Time domain representation of signal:



UNIVERSITY OF CAMBRIDGE

The Continuous Wave

- **Frequency of continuous signal**: how many times a second the wave repeats itself.
  - Measured in **Hertz**: 3Hz means the wave repeats itself three times in one second.
- **Period**: time required to produce a complete waveform.
- (Peak to Peak) **Amplitude**: max height of the wave.

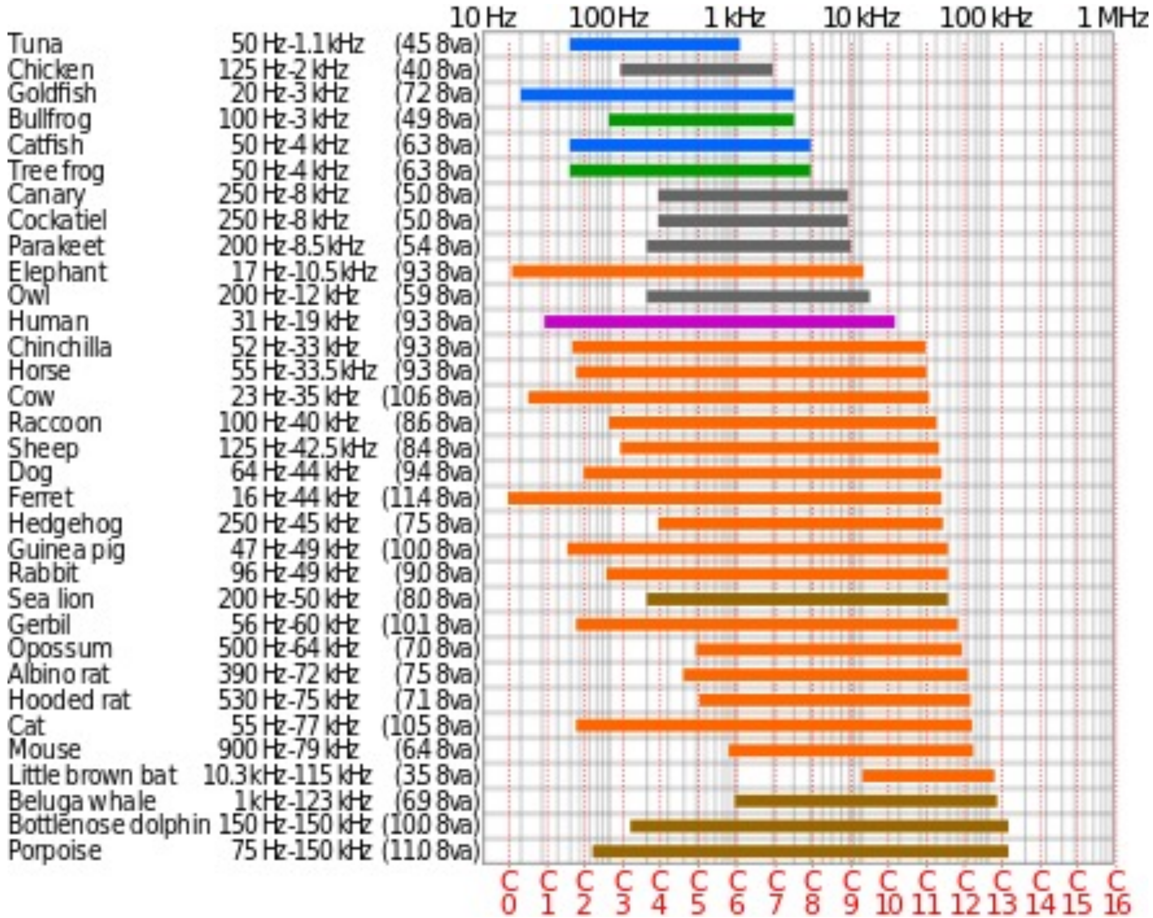


UNIVERSITY OF CAMBRIDGE

Frequencies variations

- Humans hear 20Hz to 20kHz

- Cats hear 55Hz to 79kHz



UNIVERSITY OF CAMBRIDGE

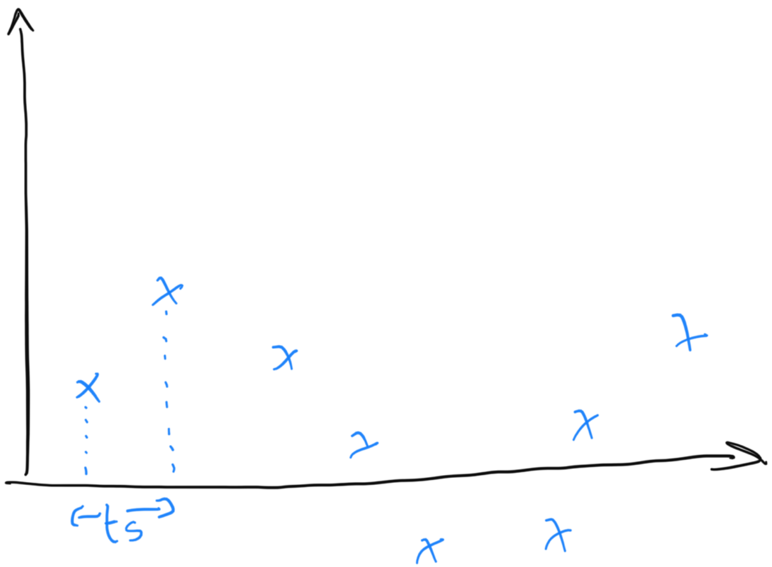
Digital Sample of Continuous Signal

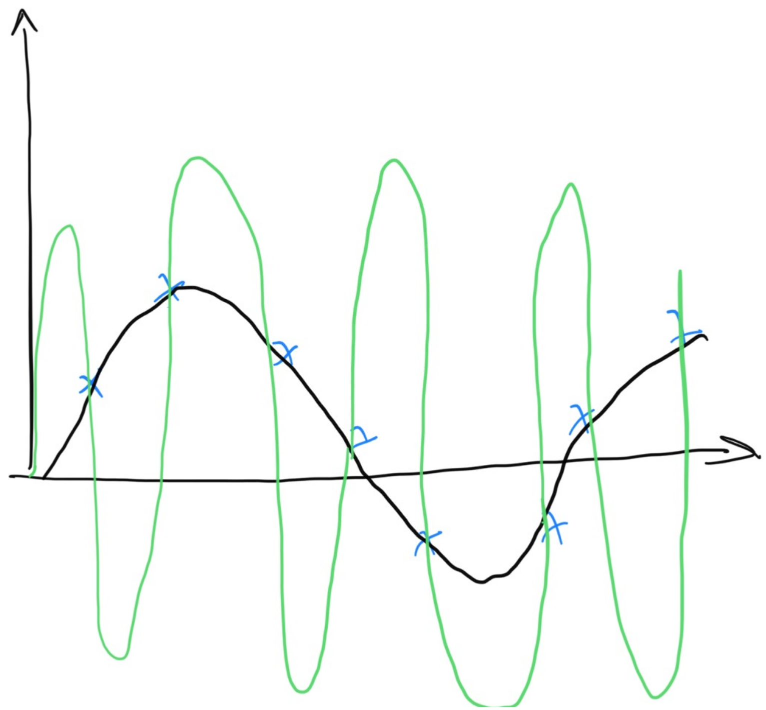
- How do we make sure our digital sampling of a continuous signal records “the important” characteristics of the signal?

- A little bit of signal processing recap ☺

Discrete Sampling of a Signal & **Aliasing**

t_s is the sampling rate can fit curves at two different frequencies (at least)



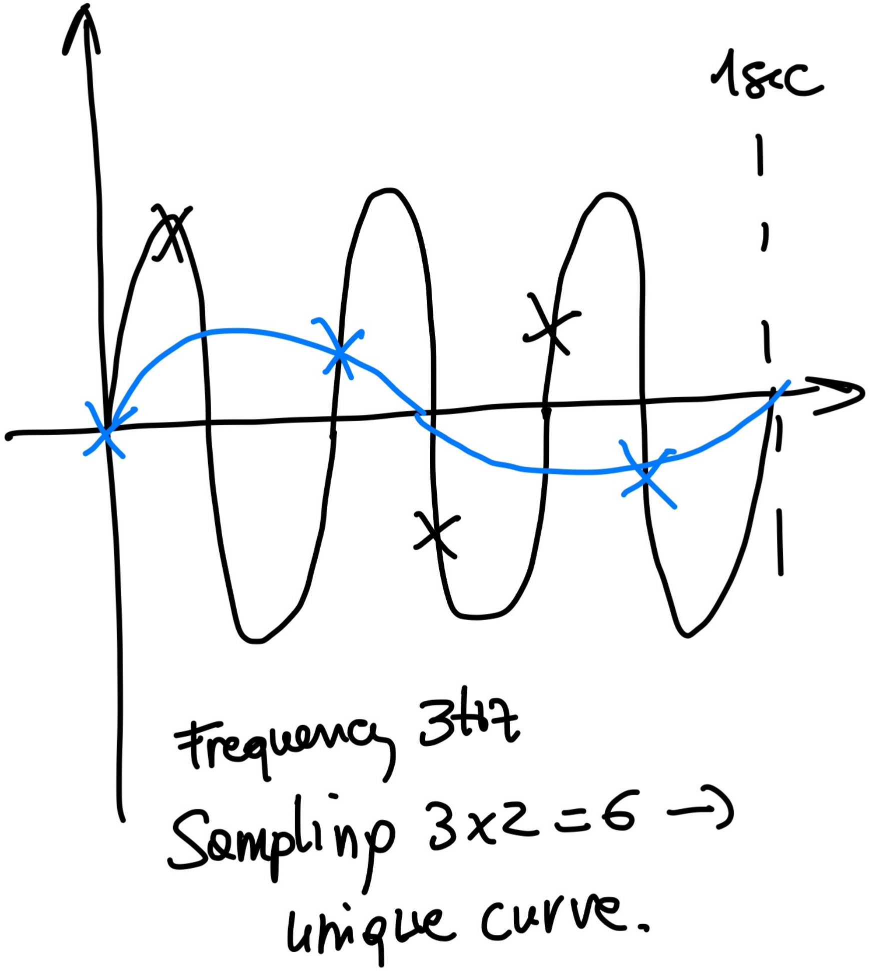


UNIVERSITY OF CAMBRIDGE

Nyquist Theorem in Practice

- To avoid aliasing the sampling frequency should be **double** the maximum frequency to be captured in the signal.
  - **Nyquist sampling rate**: rate you need to sample (at least) to have no aliasing
  - Intuition: sample twice per period!
  - Exception (result of compressed sensing): if the signal is sparse you can sample randomly and get away without sampling at twice per period!

Intuition of why the Nyquist rate works…



UNIVERSITY OF CAMBRIDGE

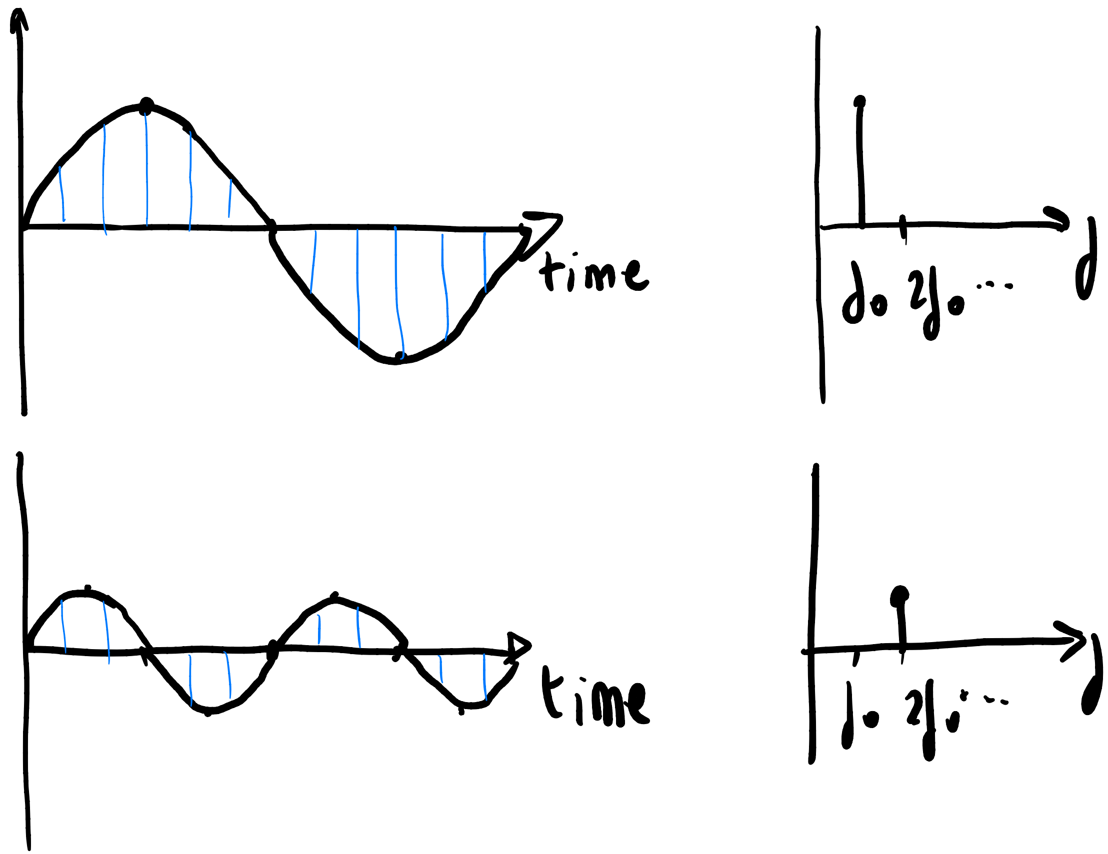
Question for you!

- Humans hear at max 22 KHz

- What is the Nyquist rate we should sample at (at least)?

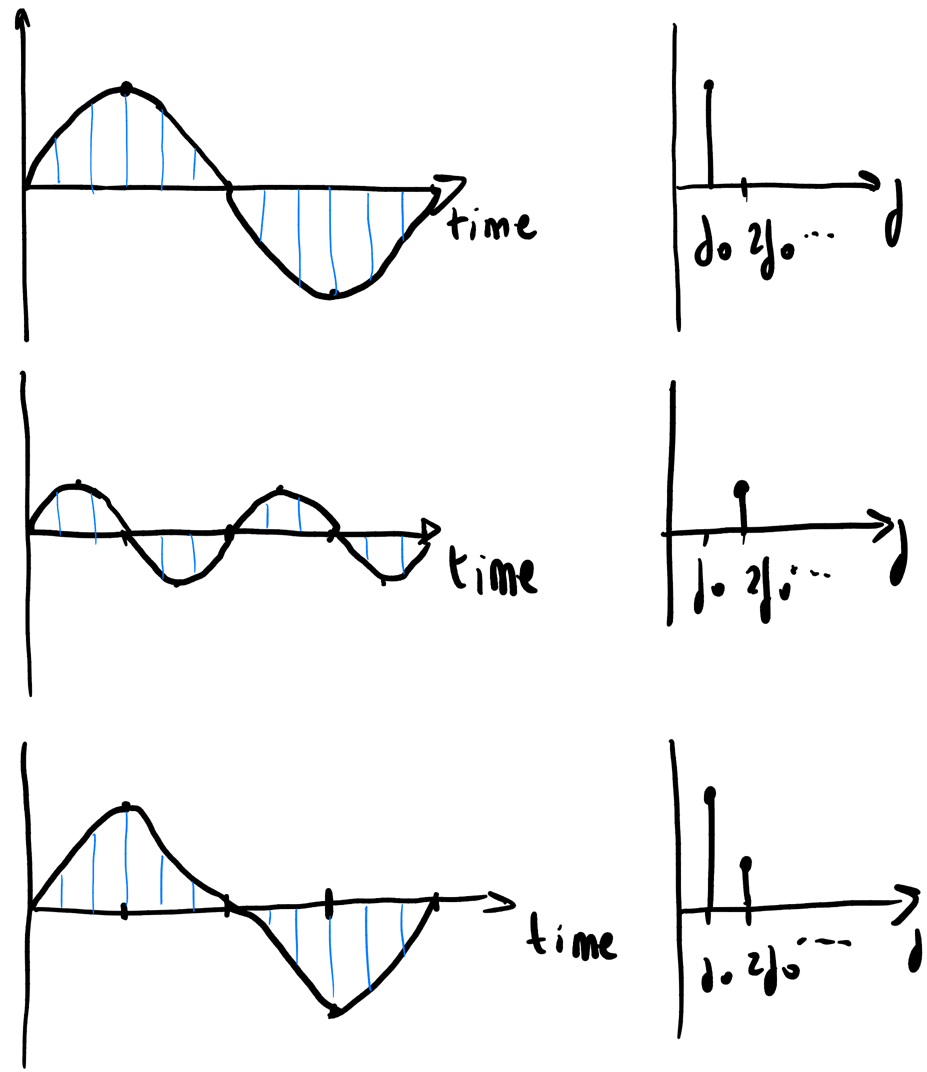
Representation of signal in frequency domain

- Signals can be represented in
  - Time domain
  - Frequency domain



Complex Signals

- Signals are composed of more than one frequency
- Example sum of signals:



UNIVERSITY OF CAMBRIDGE

Discrete Fourier Transform (DFTs)

- DFTs map the time domain graph into a frequency domain graph
  - They tell us which frequencies are important in the signal
- How? By correlating with various sin/cos waves at different frequencies!

Correlation

$$\sum_{i=0}^{N} x(i)\,y(i)$$

UNIVERSITY OF CAMBRIDGE

Intuition on DFTs

- DFTs **correlate** a wave with sin() and cos() waves at different frequencies.

function

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-i2\pi kn/N}$$

analyzing function

$$e^{-i\theta} = \cos\theta - i\sin\theta$$

$$\theta = 2\pi kn/N$$

DFTs as Correlations of Sinusoids

$$X(k) = \underbrace{\sum_{n=0}^{N-1} x(n)\cos\left(\frac{2\pi k n}{N}\right)}_{\text{CORRELATION with COS}} - i \underbrace{\sum_{n=0}^{N-1} x(n)\sin\left(\frac{2\pi k n}{N}\right)}_{\text{CORRELATION WITH SIN}}$$

Workings of DFTs (assume N=4)

- The result for each X(k) is a complex number which represents a vector

- The magnitude of the vector (ie of the two components) is the amplitude in the frequency domain

$$
\begin{aligned}
X(0) = {} & x(0)\cos(2\pi\cdot 0\cdot 0/4) - j\,x(0)\sin(2\pi\cdot 0\cdot 0/4) \\
& + x(1)\cos(2\pi\cdot 1\cdot 0/4) - j\,x(1)\sin(2\pi\cdot 1\cdot 0/4) \\
& + x(2)\cos(2\pi\cdot 2\cdot 0/4) - j\,x(2)\sin(2\pi\cdot 2\cdot 0/4) \\
& + x(3)\cos(2\pi\cdot 3\cdot 0/4) - j\,x(3)\sin(2\pi\cdot 3\cdot 0/4)
\end{aligned}
$$

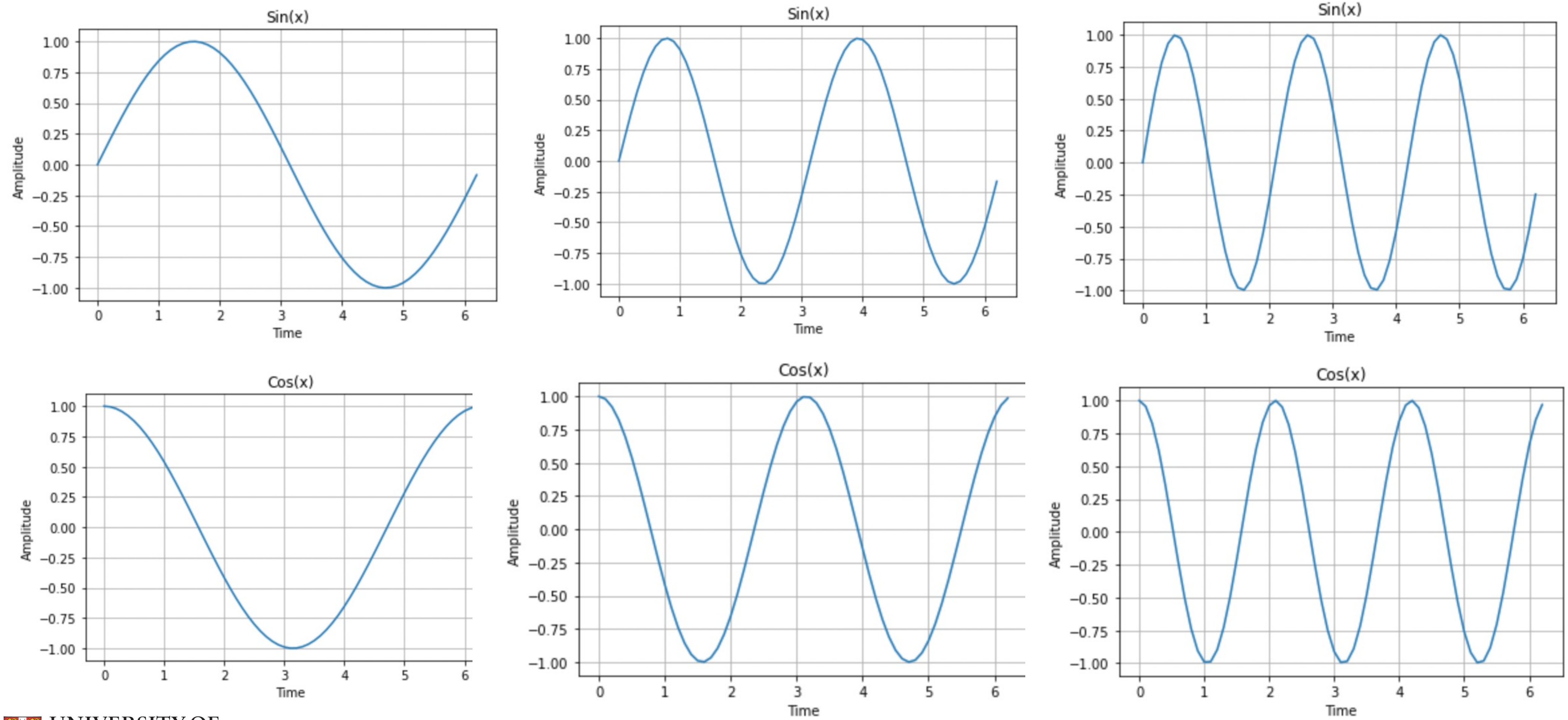
$$
\begin{aligned}
X(1) = {} & x(0)\cos(2\pi\cdot 0\cdot 1/4) - j\,x(0)\sin(2\pi\cdot 0\cdot 1/4) \\
& + x(1)\cos(2\pi\cdot 1\cdot 1/4) - j\,x(1)\sin(2\pi\cdot 1\cdot 1/4) \\
& + x(2)\cos(2\pi\cdot 2\cdot 1/4) - j\,x(2)\sin(2\pi\cdot 2\cdot 1/4) \\
& + x(3)\cos(2\pi\cdot 3\cdot 1/4) - j\,x(3)\sin(2\pi\cdot 3\cdot 1/4)
\end{aligned}
$$

$$
\begin{aligned}
X(2) = {} & x(0)\cos(2\pi\cdot 0\cdot 2\ 4) - j\,x(0)\sin(2\pi\cdot 0\cdot 2/4) \\
& + x(1)\cos(2\pi\cdot 1\cdot 2\ 4) - j\,x(1)\sin(2\pi\cdot 1\cdot 2/4) \\
& + x(2)\cos(2\pi\cdot 2\cdot 2/4) - j\,x(2)\sin(2\pi\cdot 2\cdot 2/4) \\
& + x(3)\cos(2\pi\cdot 3\cdot 2/4) - j\,x(3)\sin(2\pi\cdot 3\cdot 2/4)
\end{aligned}
$$

$$
\begin{aligned}
X(3) = {} & x(0)\cos(2\pi\cdot 0\cdot 3/4) - j\,x(0)\sin(2\pi\cdot 0\cdot 3/4) \\
& + x(1)\cos(2\pi\cdot 1\cdot 3/4) - j\,x(1)\sin(2\pi\cdot 1\cdot 3/4) \\
& + x(2)\cos(2\pi\cdot 2\cdot 3/4) - j\,x(2)\sin(2\pi\cdot 2\cdot 3/4) \\
& + x(3)\cos(2\pi\cdot 3\cdot 3/4) - j\,x(3)\sin(2\pi\cdot 3\cdot 3/4)
\end{aligned}
$$

UNIVERSITY OF CAMBRIDGE

Sin and Cos Waves at different frequencies

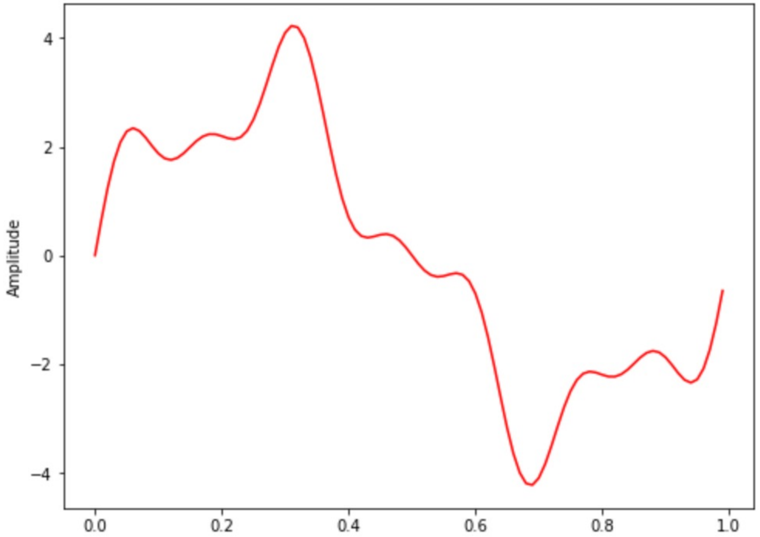


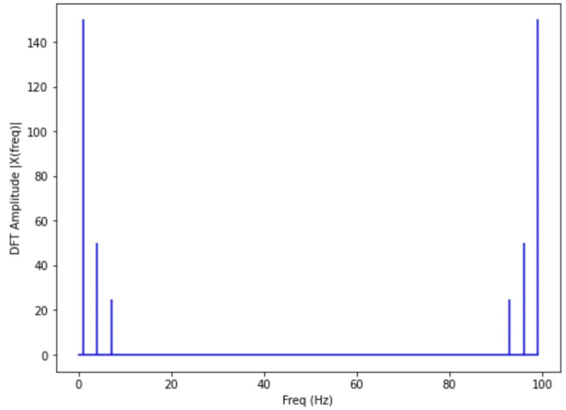
UNIVERSITY OF CAMBRIDGE

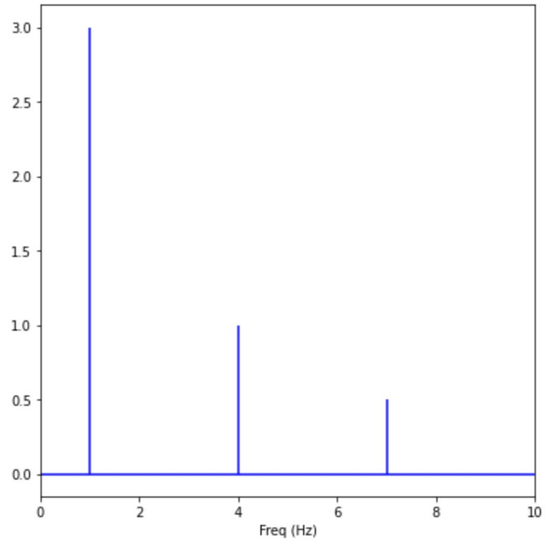
DFT plot

DFT for this wave showing the three frequencies and amplitude.

Wave composed of 3 sine waves with frequencies 1 Hz, 4 Hz, and 7 Hz, amplitudes 3, 1 and 0.5
Sampling rate 100Hz







UNIVERSITY OF CAMBRIDGE

Example from https://pythonnumericalmethods.berkeley.edu/notebooks/chapter24.02-Discrete-Fourier-Transform.html

Fast Fourier Transforms

- Fast way of calculating DFT
  - Intuition: Splitting the computation of even and odd n in the formula recursively and parallelize

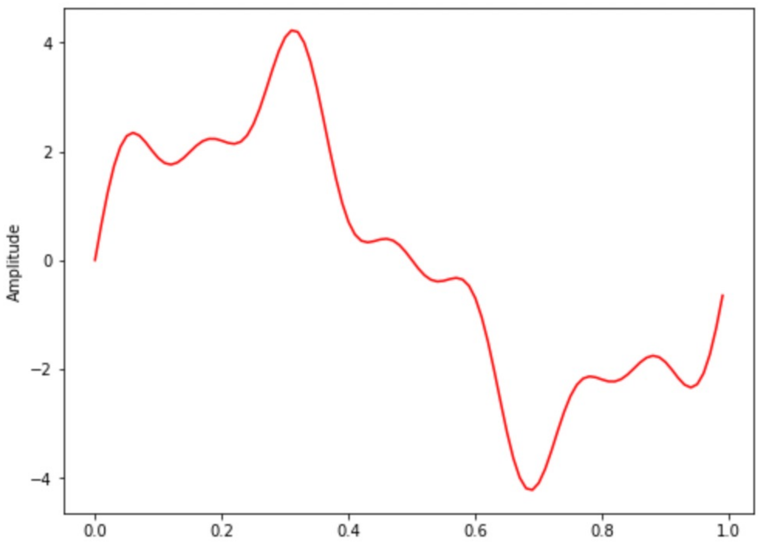
  - DFT ~ O(N^2)

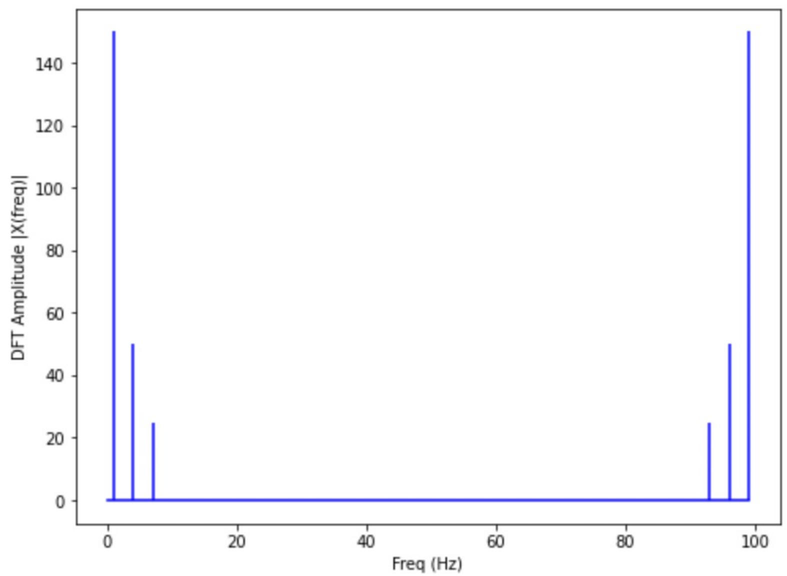
  - FFT ~ O(NlogN)

UNIVERSITY OF CAMBRIDGE

From time and frequency representations to Spectrograms

- Representation in time and frequency (separately)

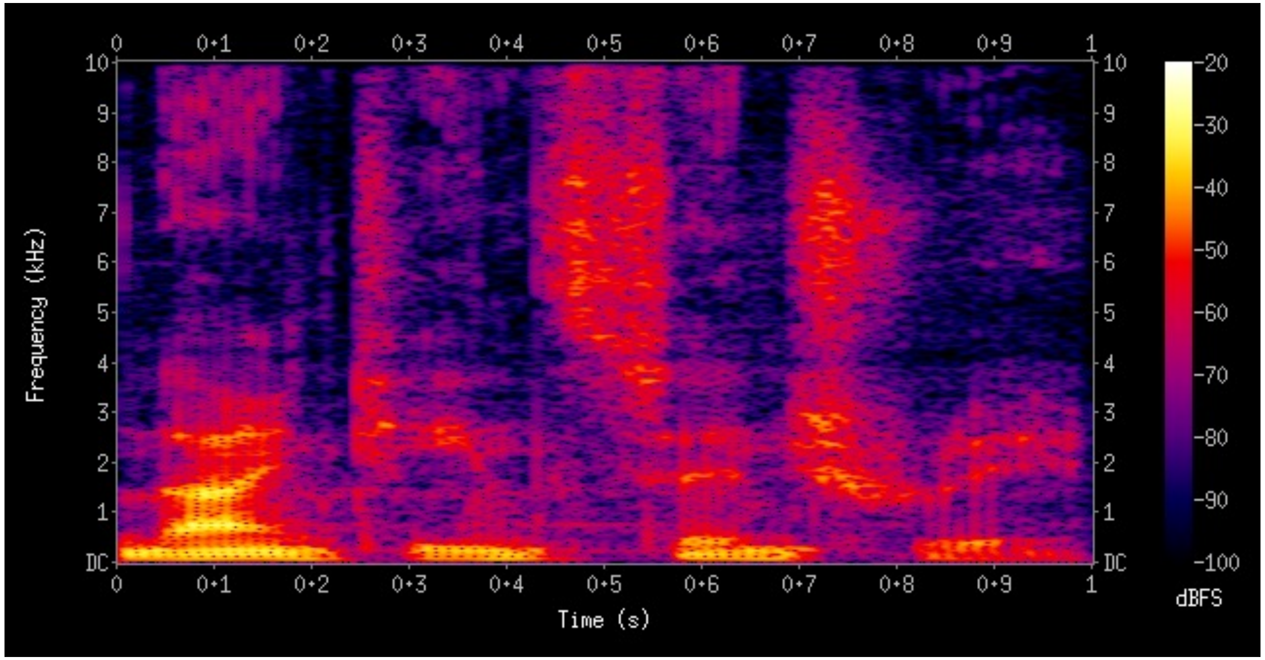




UNIVERSITY OF CAMBRIDGE

Spectrograms

- Representation in frequency and time



UNIVERSITY OF CAMBRIDGE

From Wikipedia. Word: nineteenth century

How to generate a spectrogram

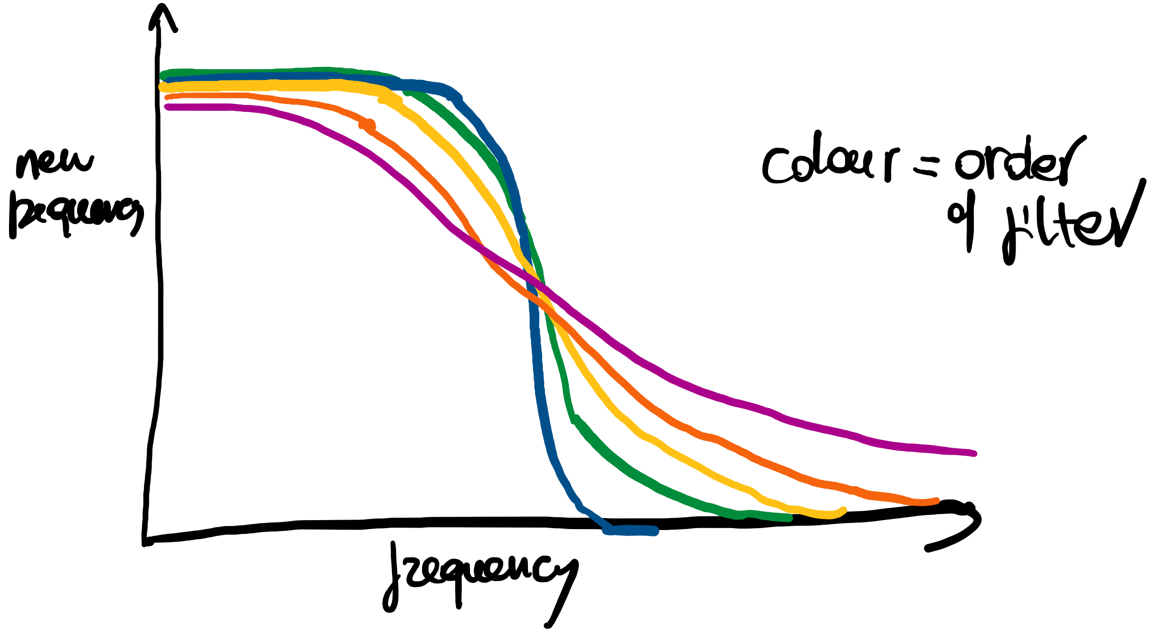
- For each time window a DFT is calculated and frequency intensity is represented with ”colours” which indicate the level.

- The X axis is the time
- The Y axis is the frequency

UNIVERSITY OF CAMBRIDGE

Filtering of the signal

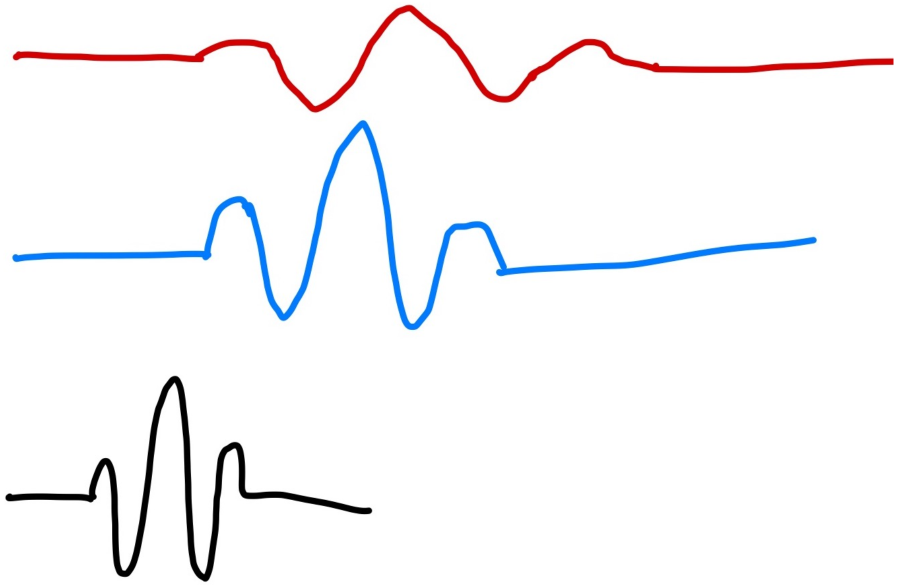
- Band-pass filter: an operation which ensures that only certain frequencies are kept in the signal
  - Often used to eliminate specific signals: eg heart signals from respiration signals.
- Butterworth Filter:



DWT: Discrete Wavelet Transforms

- DFT or STFT (short-time Fourier Transforms: FTs on a portion of the time segment) fail to capture time and frequency dependencies well.
  - Only known what frequency in an interval.
  - DFTs decompose the signal into sinusoidal basis functions of different frequencies.
- Discrete Wavelet Transforms (DWT) decompose a signal in orthogonal wavelet basis functions. These functions are nonzero over only part of the total signal length.
- DWTs are dilated, translated and scaled versions of a ”base” function (wavelet).

DWT: Discrete Wavelet Transforms



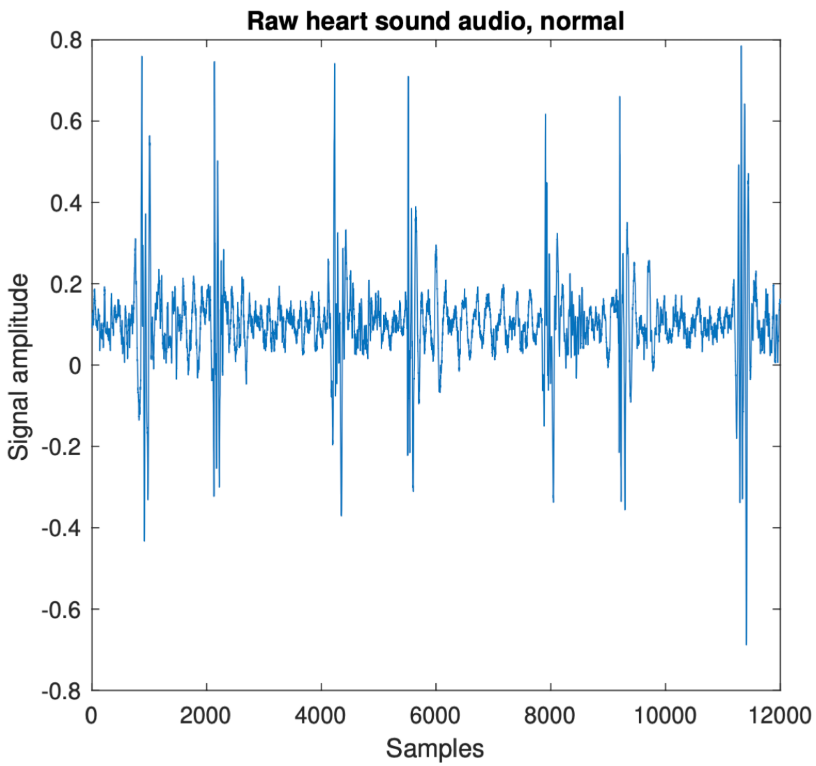
UNIVERSITY OF CAMBRIDGE

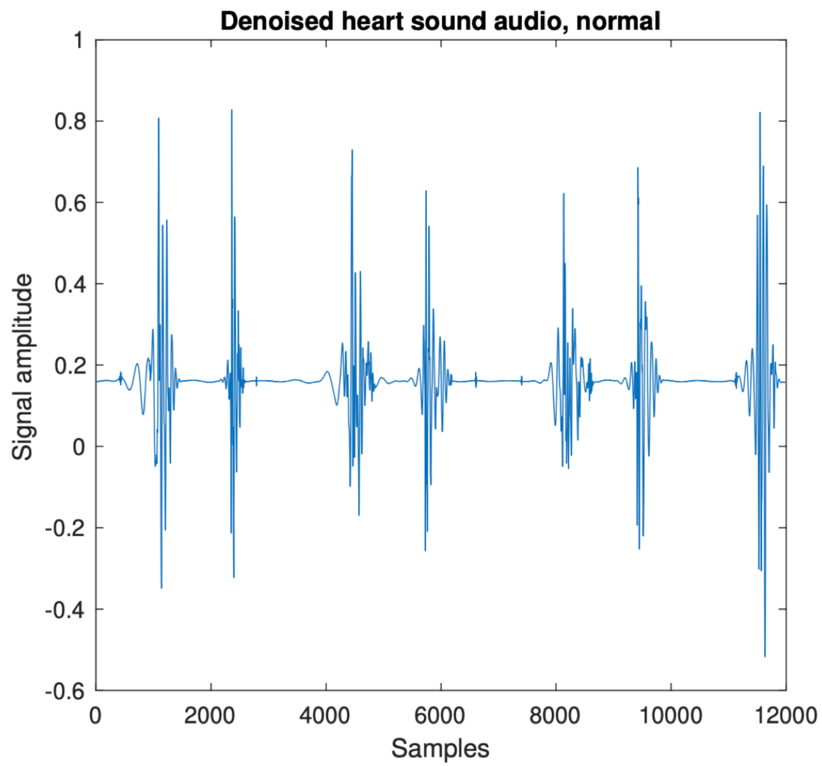
Wavelets: intuition

- The various wavelets (with different frequency) are passed through a signal and highlight different regions of the signal, respectively.

- This process is very good for denoising a signal (highlighting the important frequencies).

Denoising with DWT



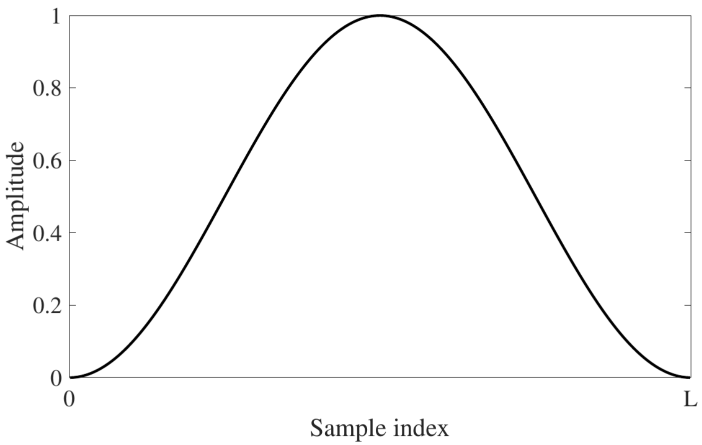


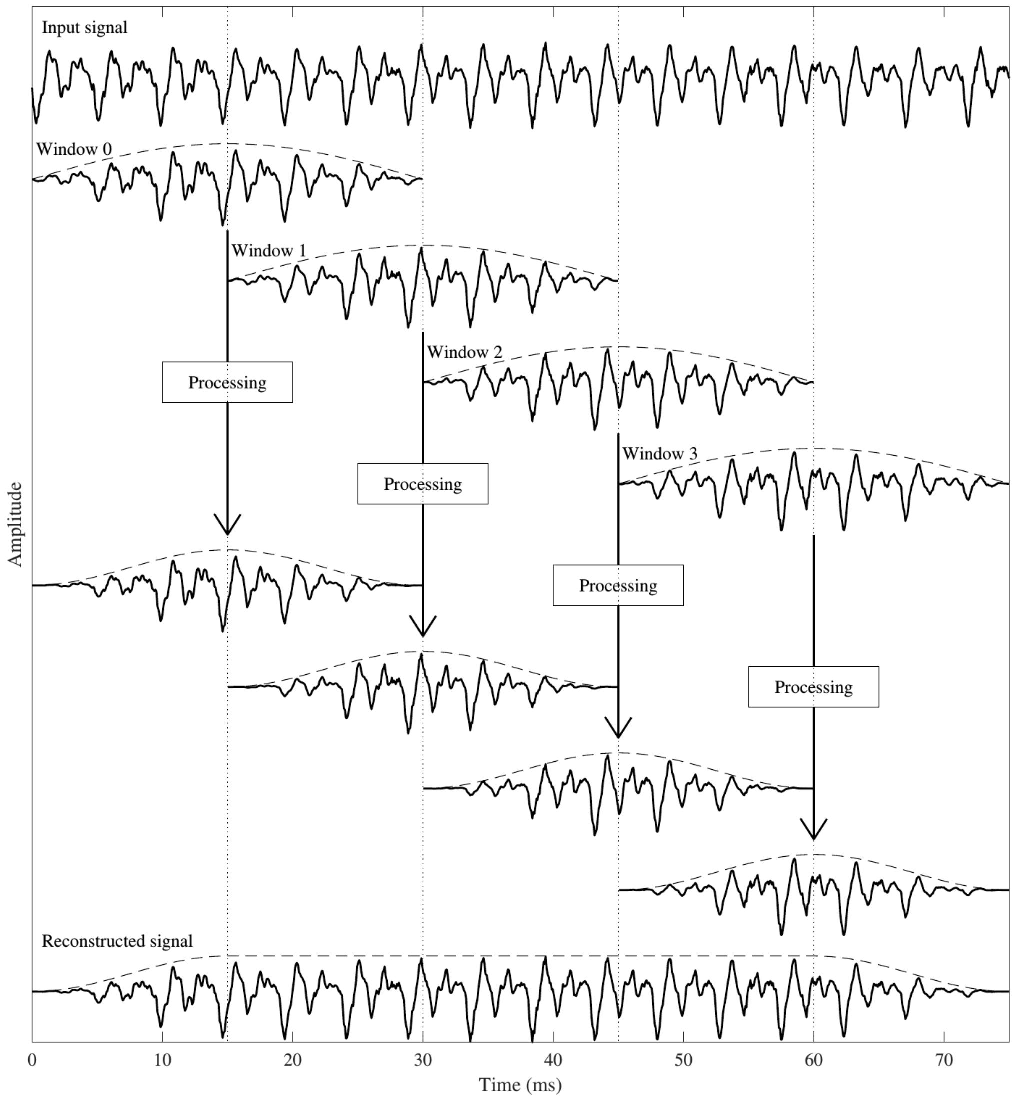
UNIVERSITY OF CAMBRIDGE

Bondareva, E., Han, J., Bradlow, W., & Mascolo, C. (2021). Segmentation-free Heart Pathology Detection Using Deep Learning. Graphs from presentation at 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC).

Windowing

- Used to isolate a small portion of the signal.
- Often a filer is applied to smoothen start and end discontinuities.
- Example of filter function (Hann).





Figures from https://wiki.aalto.fi/display/ITSP/Windowing CC

Questions