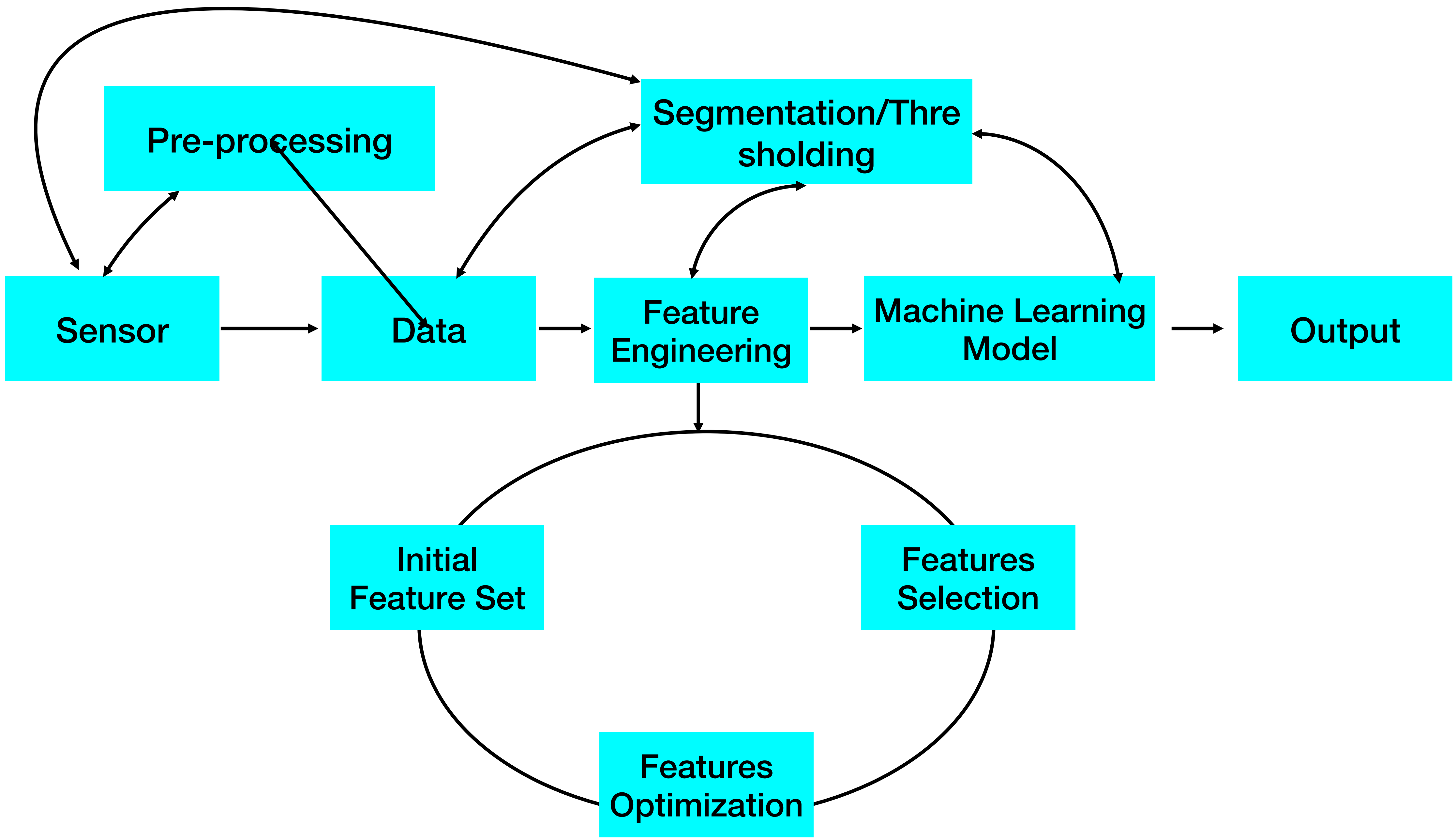
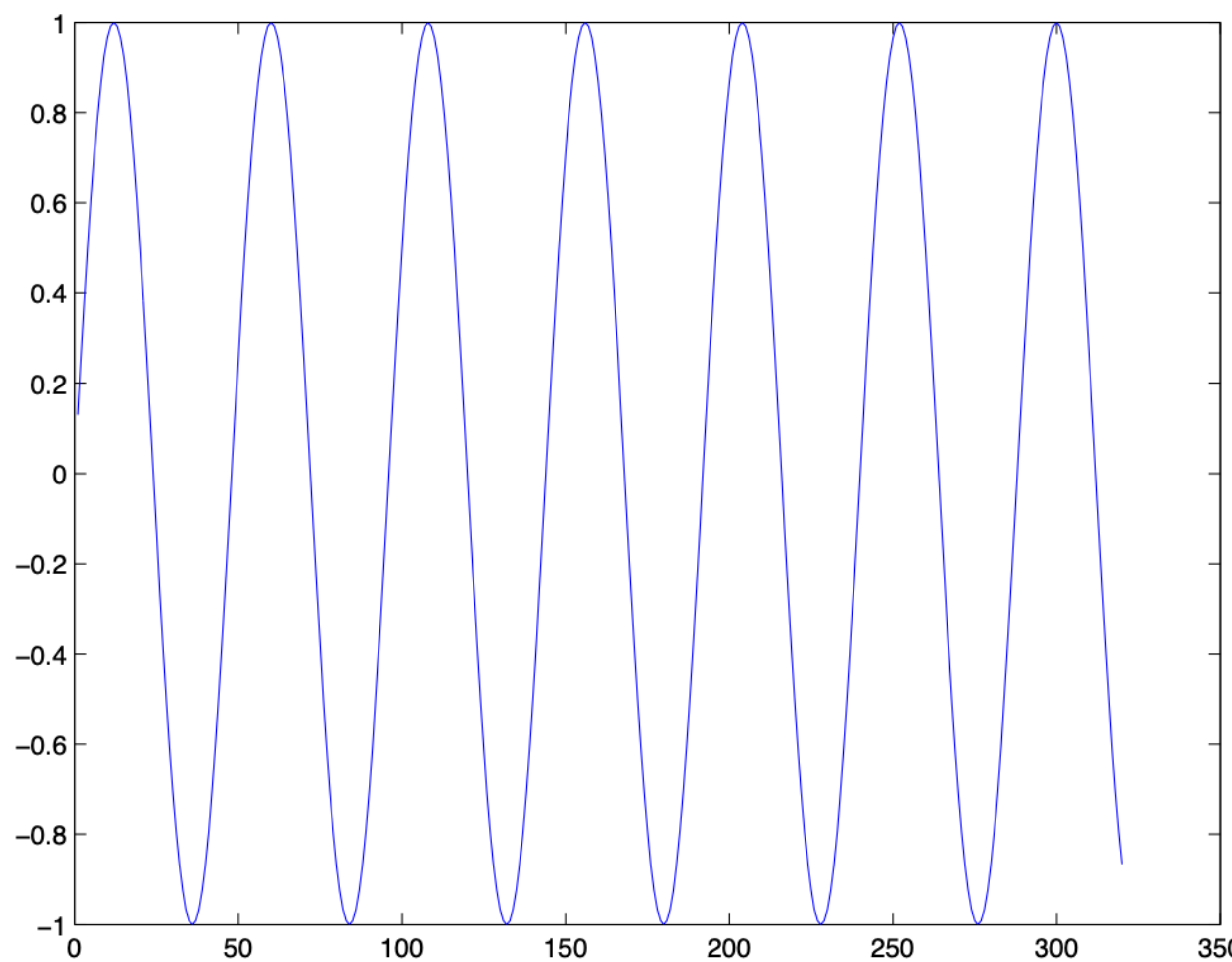
Windowing, Frequency, and Sound

Segmentation

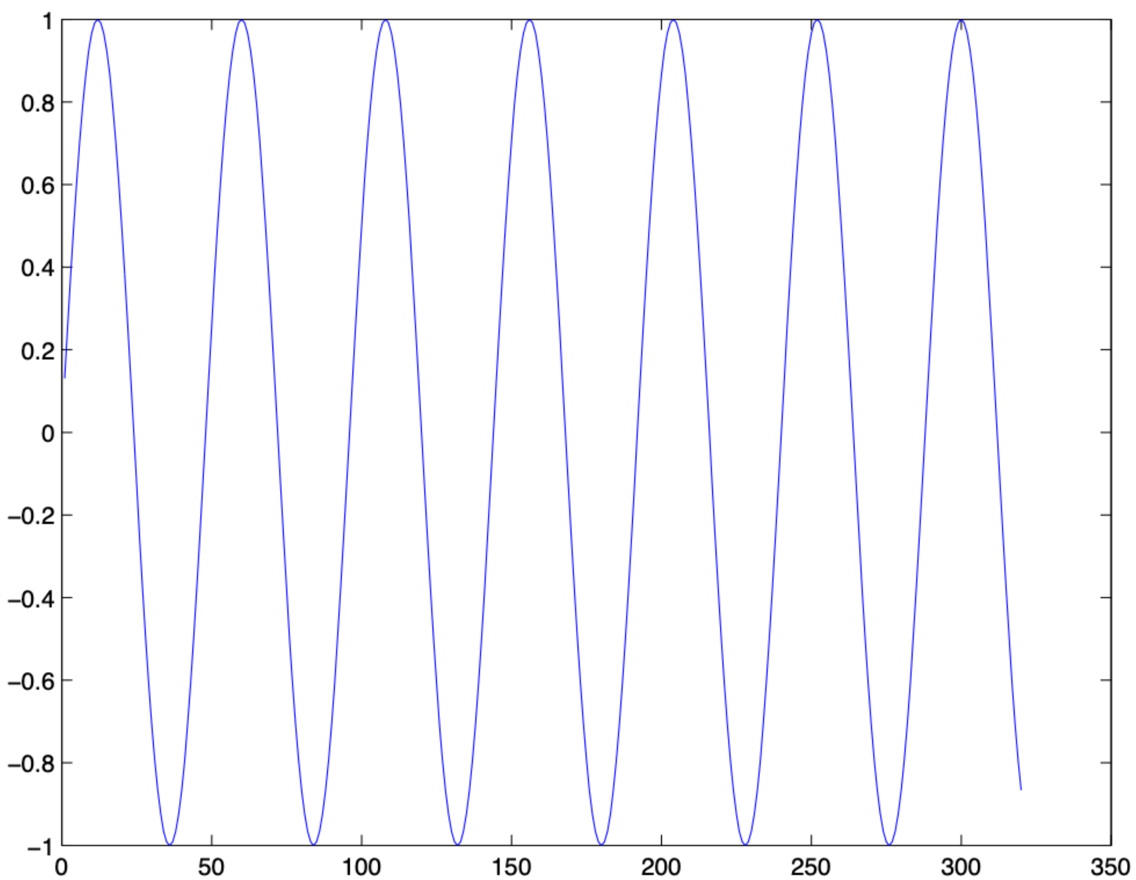


Frequency Domain

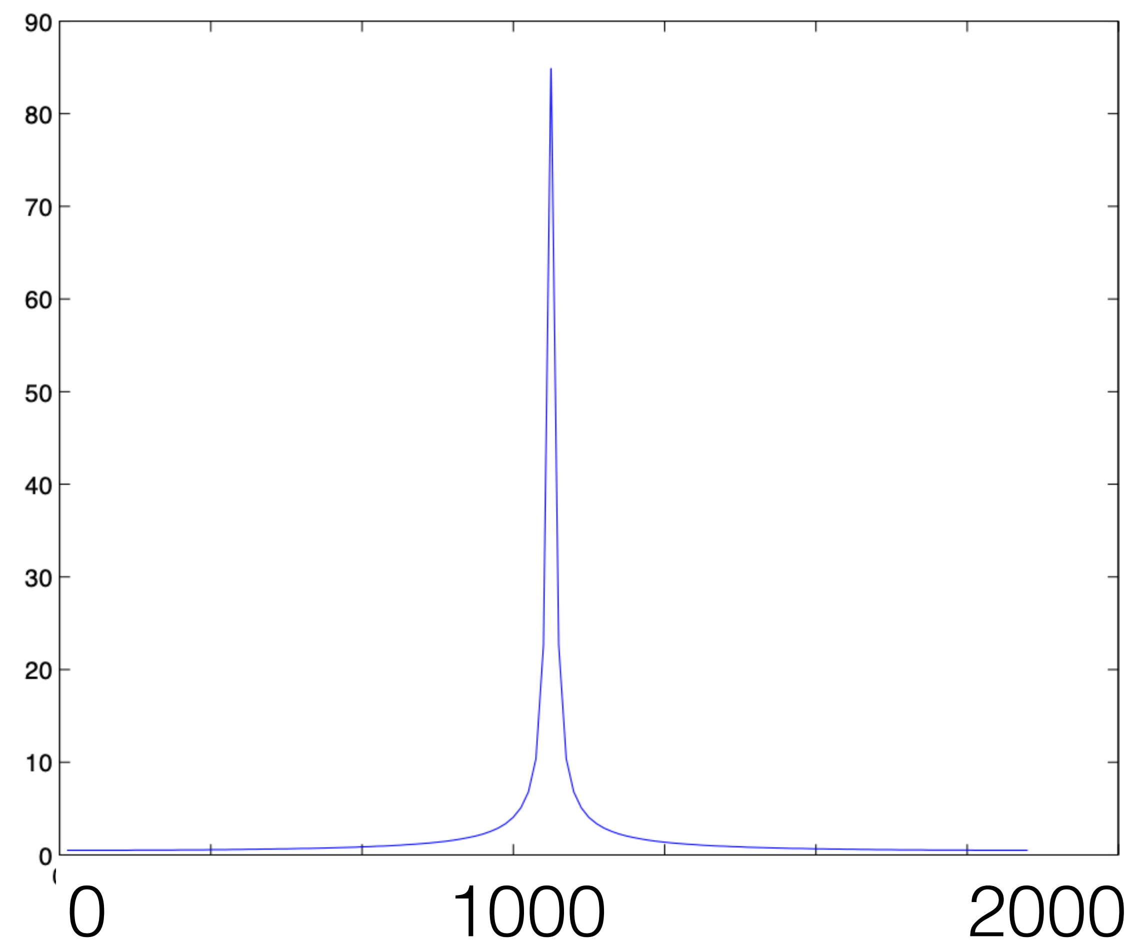


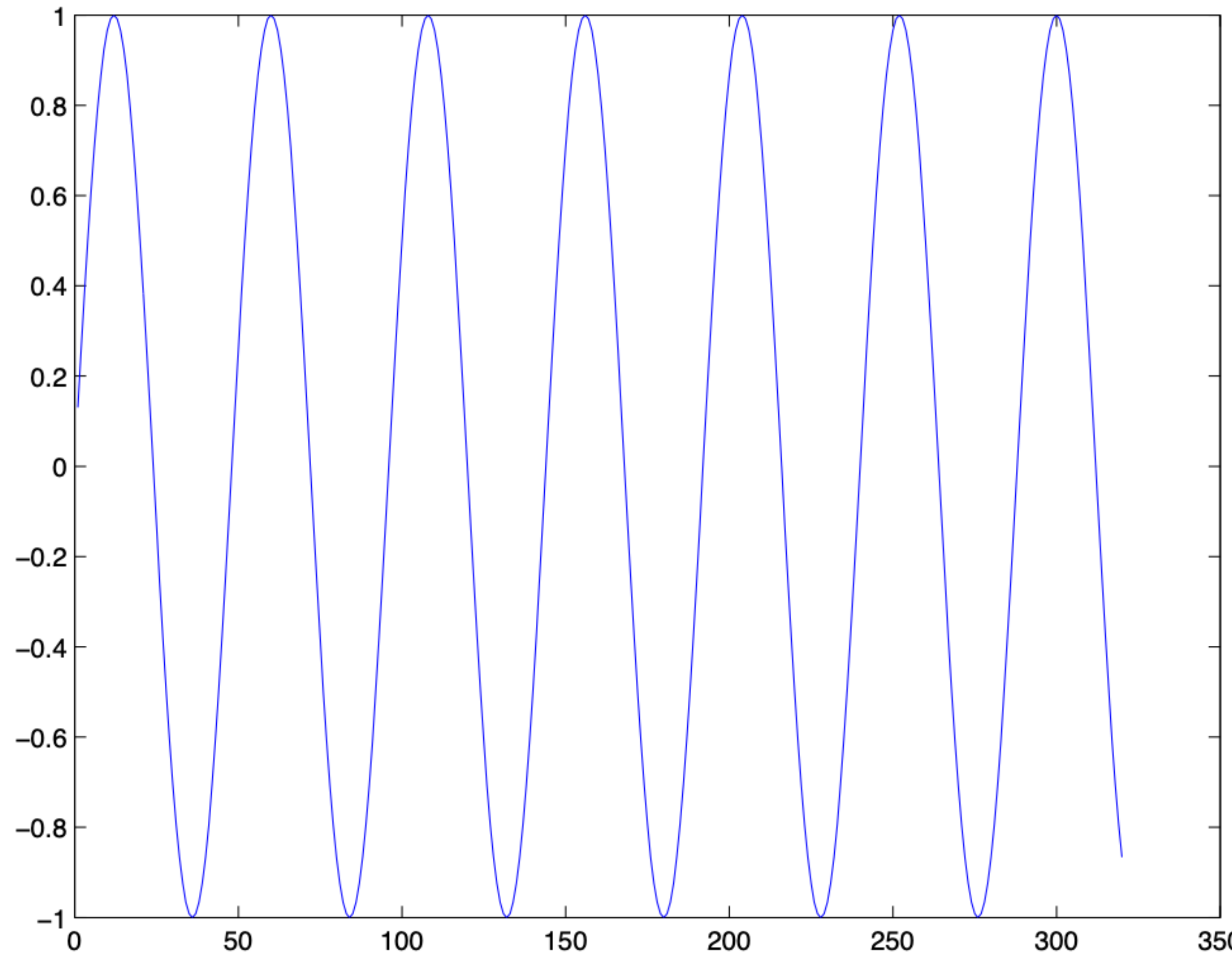
1kHz

Time Domain



Frequency Domain

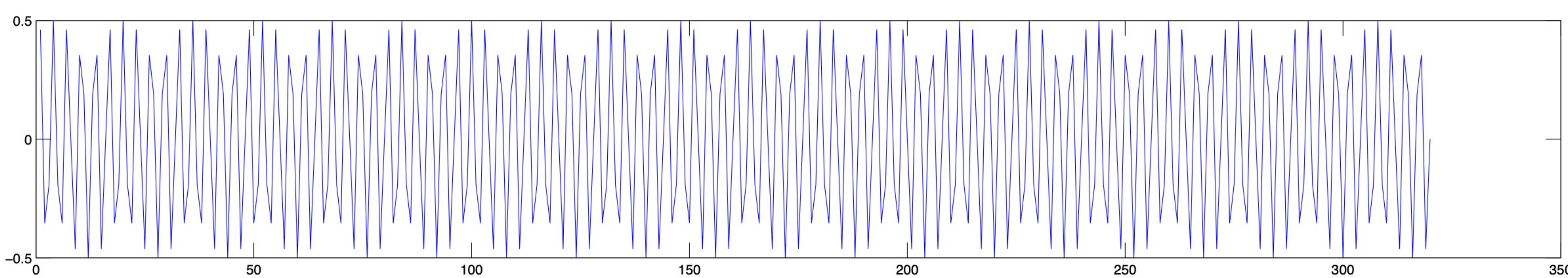




1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
0
50
100
150
200
250
300
350

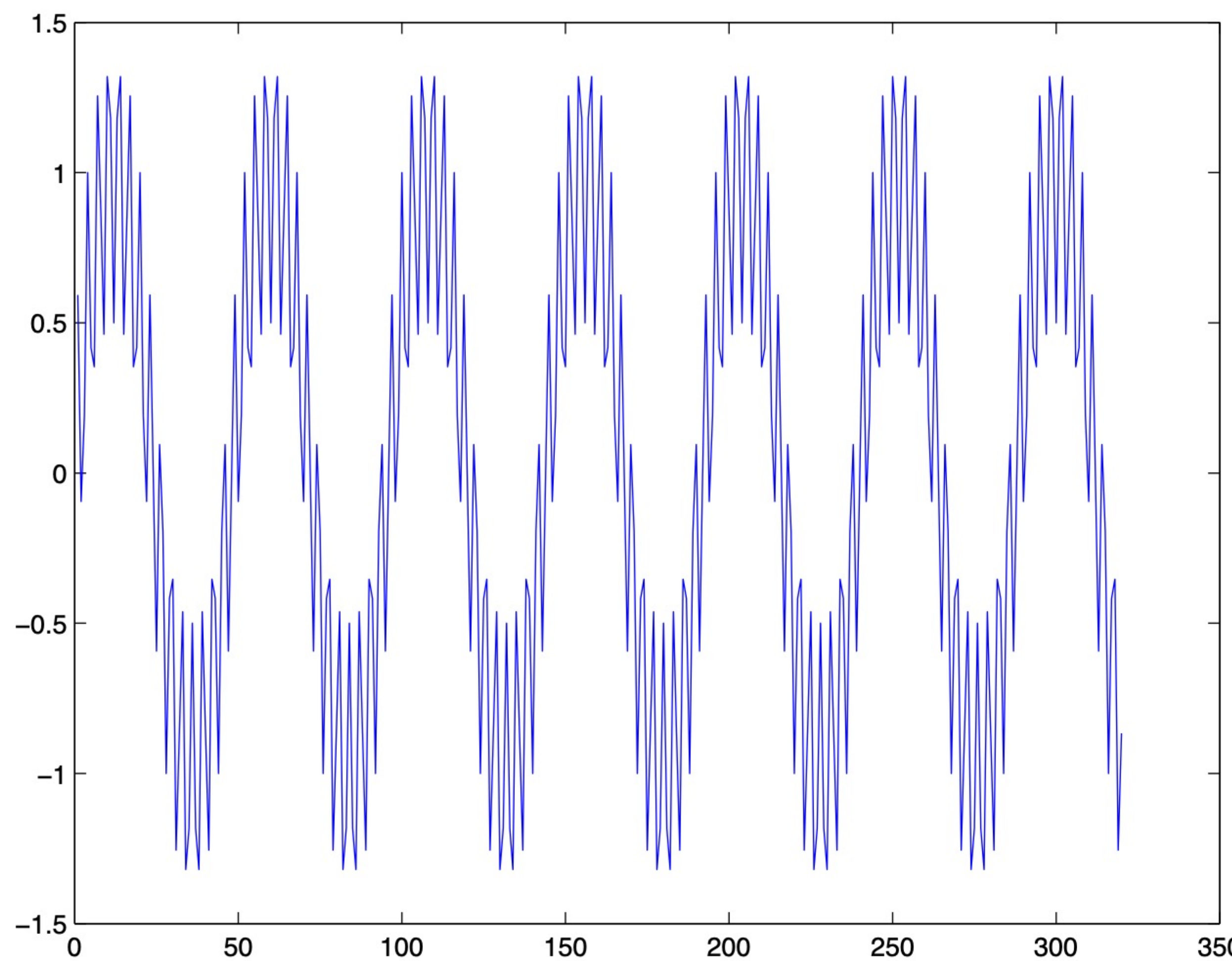
1kHz

+



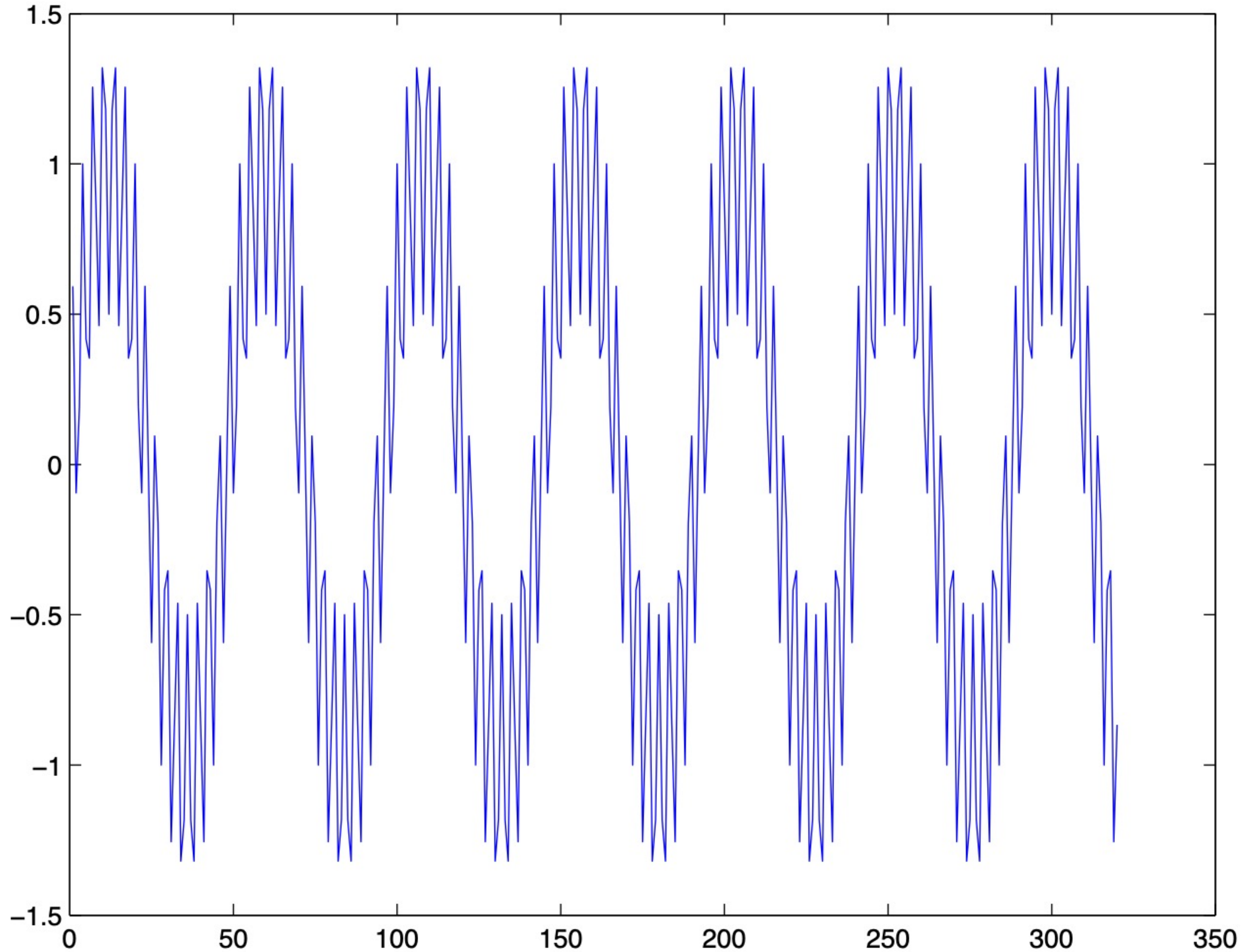
0.5
0
−0.5
0
50
100
150
200
250
300
350

15kHz

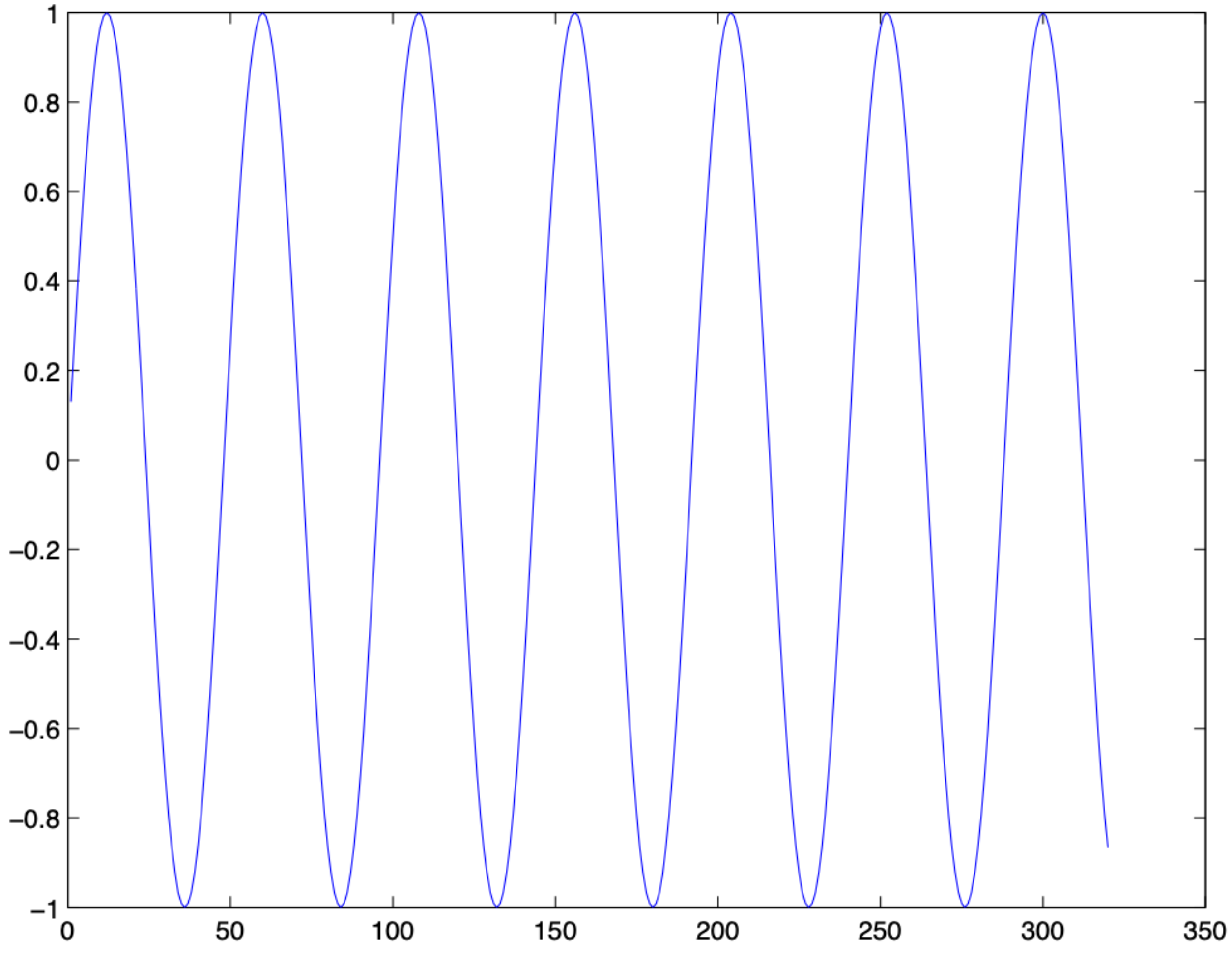


1kHz + 15kHz

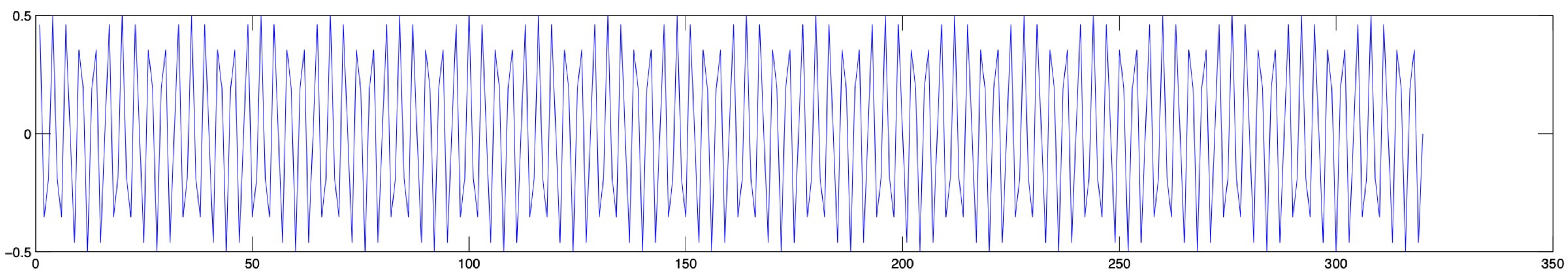
Unfiltered Data



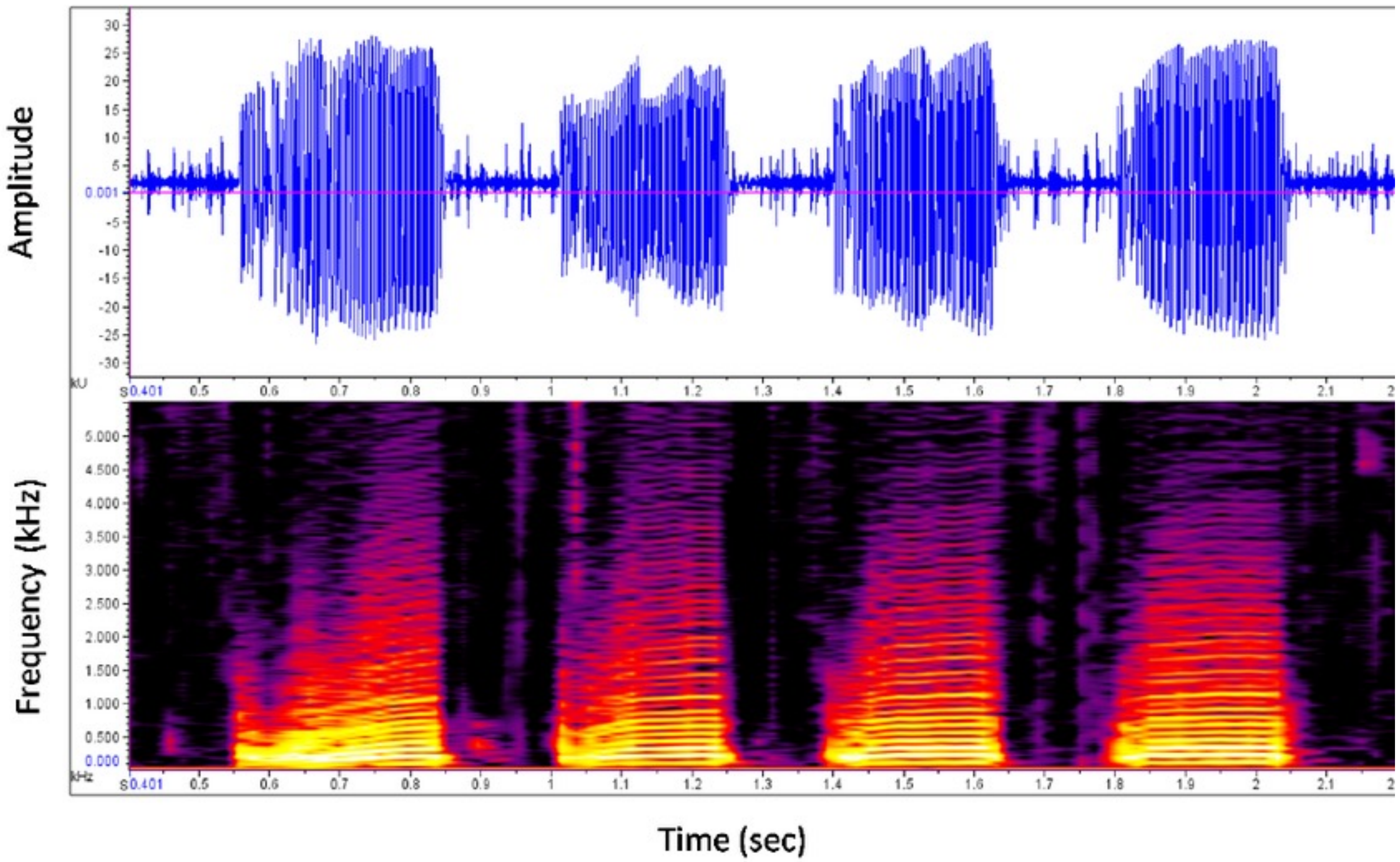
Low Pass Filter



High Pass Filter

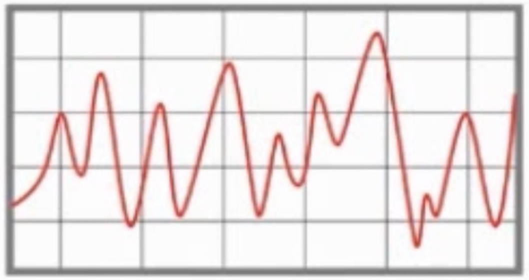


Audio



What exactly is an FFT?

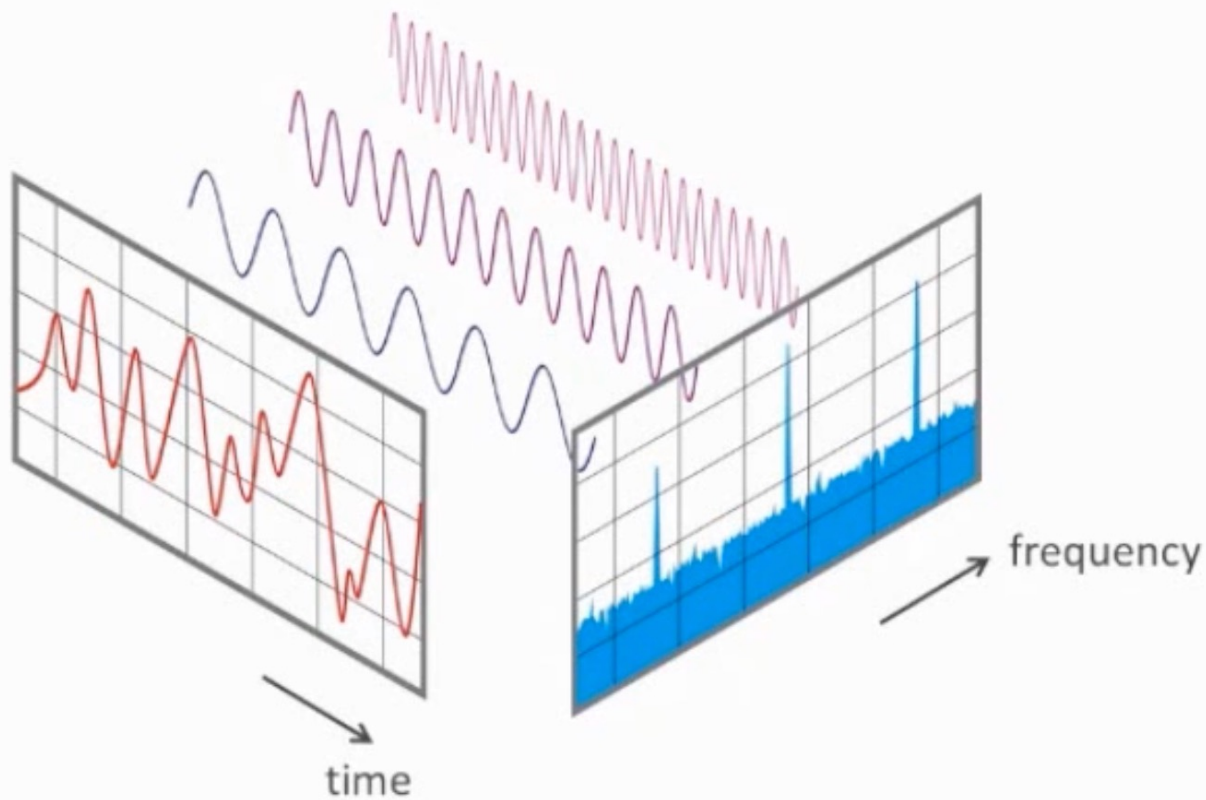
FFT calculates the spectrum of a *periodic* time signal



Credit: NTI Audio

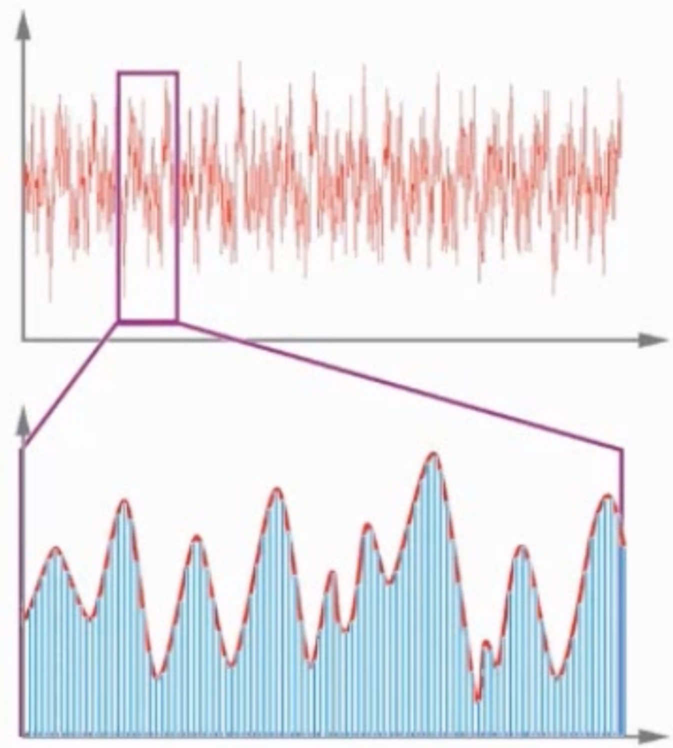
What exactly is an FFT?

FFT calculates the spectrum of a *periodic* time signal



Credit: NTI Audio

FFT Calculation – Step 1



1. Time signal

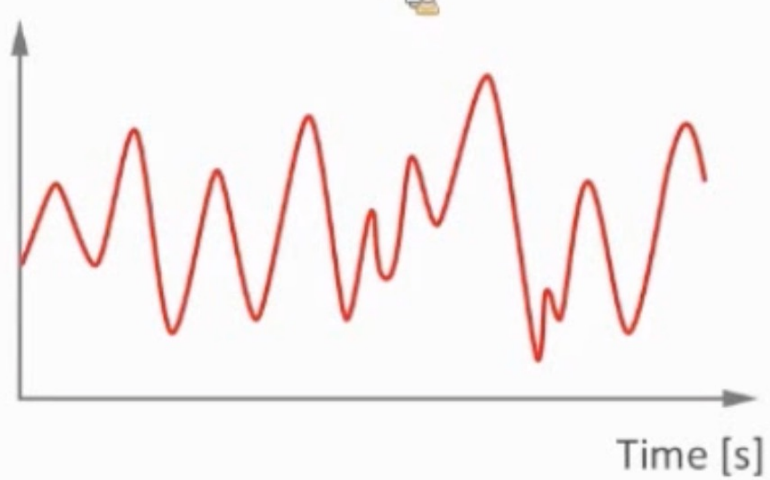
2. Sampling
   ‘Size’ or ‘blocklength’ = number of samples e.g. 512, 1024, 2048, 4096, 8192, …

Credit: NTI Audio

FFT Calculation – Step 2

$$f_m = \sum_{k=0}^{2n-1} x_k e^{-\frac{2\pi i}{2n} mk}$$

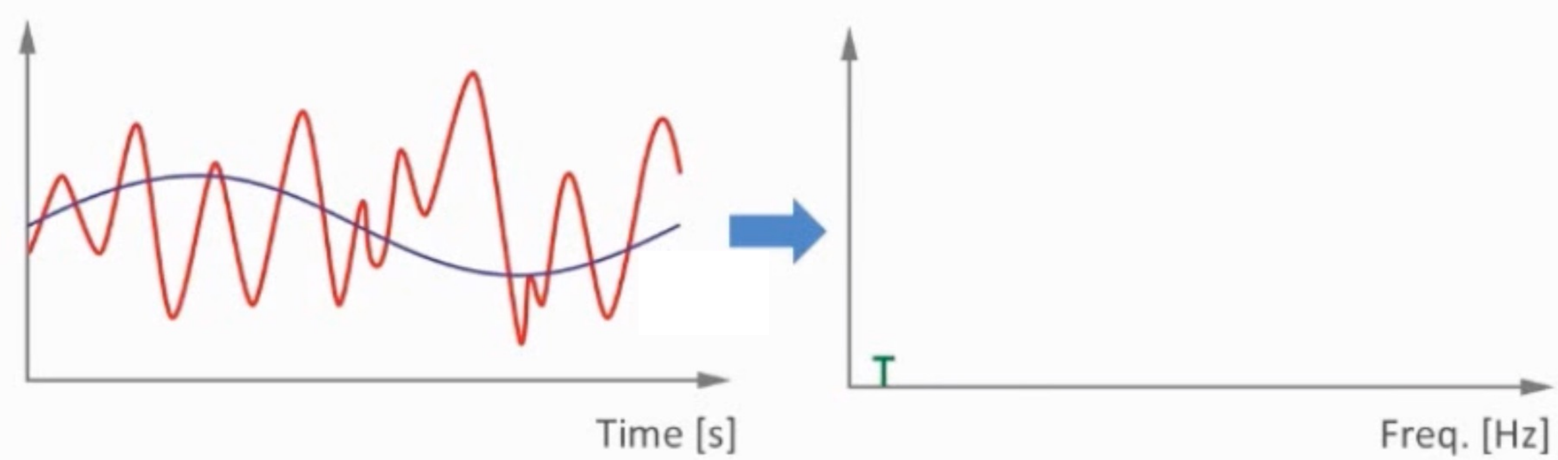
→ *Sum of Sine / Cosine waves*



Credit: NTI Audio

FFT Calculation – Step 2

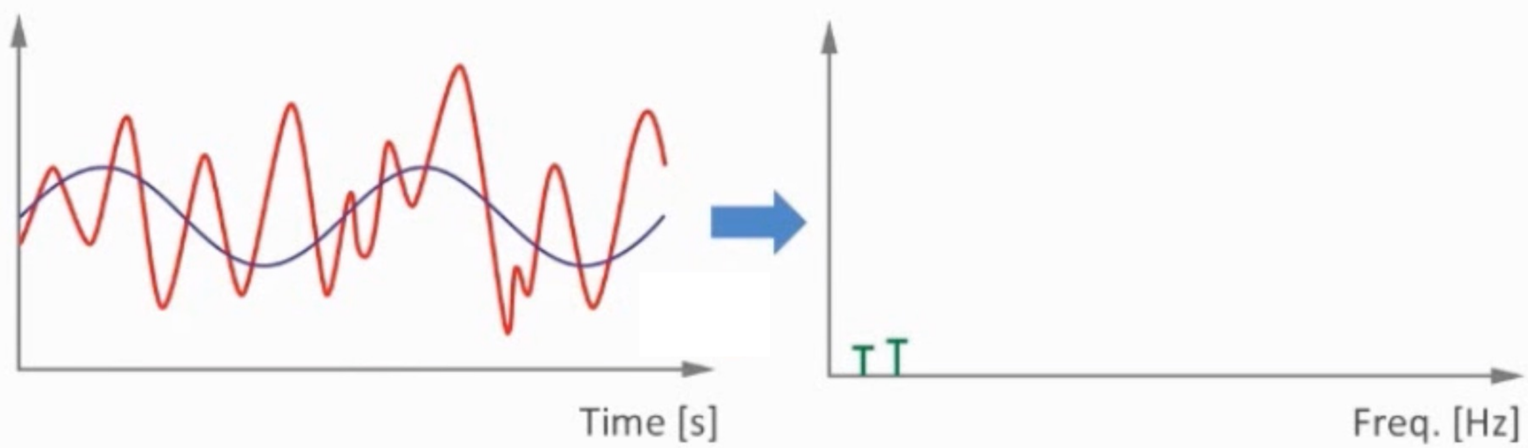
$$f_m = \sum_{k=0}^{2n-1} x_k e^{-\frac{2\pi i}{2n} mk}$$ $\rightarrow$ *Sum of Sine / Cosine waves*



Credit: NTI Audio

FFT Calculation – Step 2

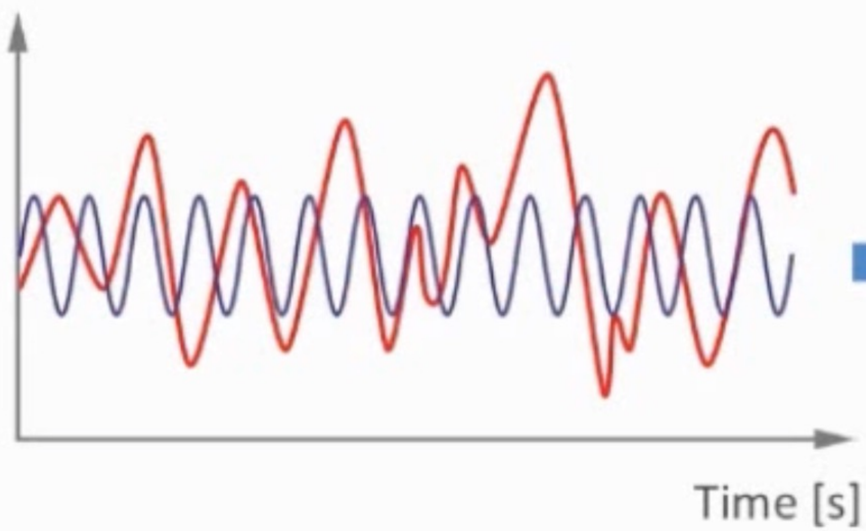
$$f_m = \sum_{k=0}^{2n-1} x_k e^{-\frac{2\pi i}{2n}mk}$$ $\rightarrow$ *Sum of Sine / Cosine waves*

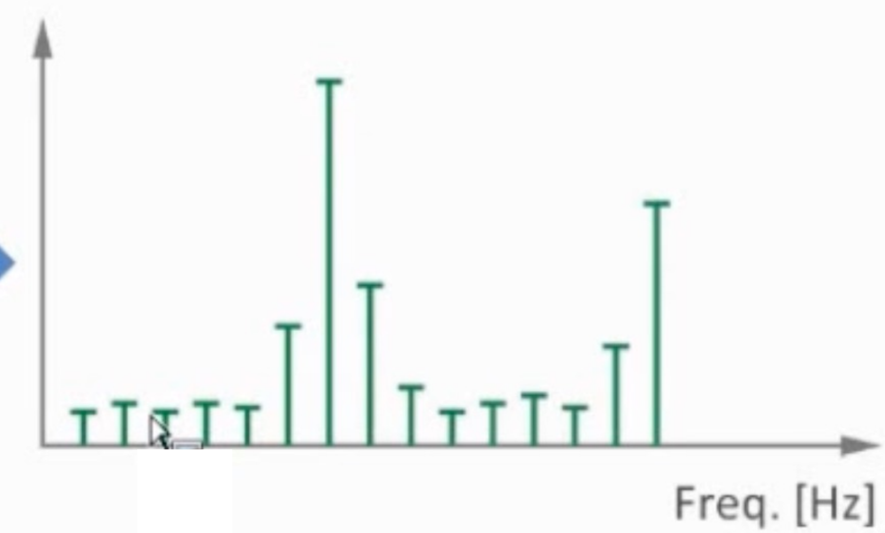


Credit: NTI Audio

FFT Calculation – Step 2

$$f_m = \sum_{k=0}^{2n-1} x_k e^{-\frac{2\pi i}{2n} mk} \quad \rightarrow \textit{Sum of Sine / Cosine waves}$$



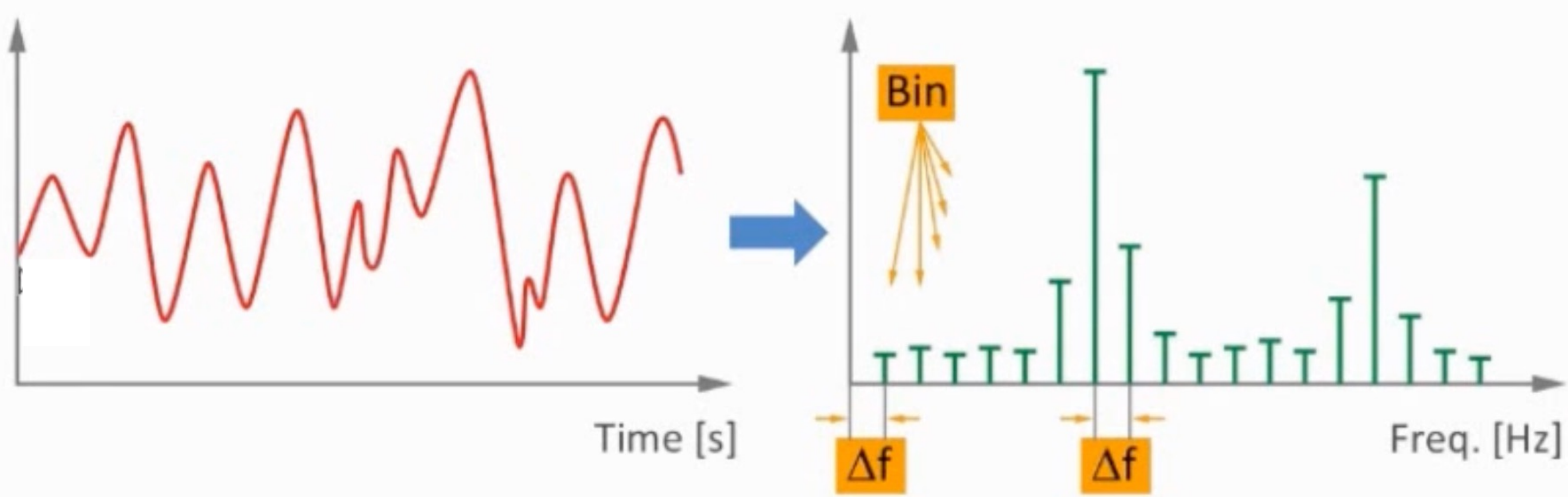


Credit: NTI Audio

FFT Calculation – Step 2

$$f_m = \sum_{k=0}^{2n-1} x_k e^{-\frac{2\pi i}{2n}mk}$$

→ *Sum of Sine / Cosine waves*



Credit: NTI Audio

Implications – D, Δf

The sampling rate and the blocklength give the

$\Rightarrow$ **Duration** per block

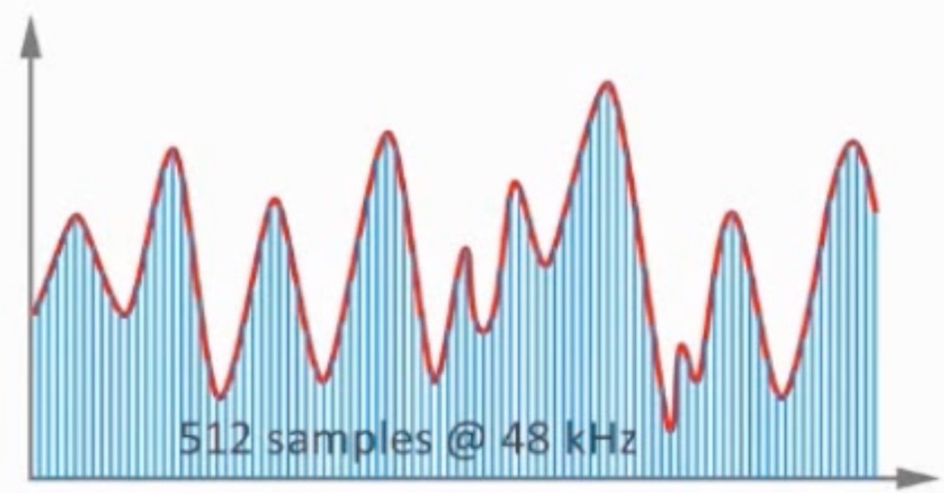
$\Rightarrow$ **Frequency resolution** of the FFT

Example

- Sampling rate: 48 kHz
- Blocklength: 512 samples

$$\Rightarrow \quad D = \frac{512}{48000} = \mathbf{10.67}\ [{}^{ms}/_{block}]$$

$$\Rightarrow \quad \Delta f = \frac{1}{10.67} = \mathbf{93.75}\ [Hz]$$



Credit: NTI Audio

Ono et al. 2013

Features Used

- 8192-point FFT
    - 4096 valid points
- 0 to 48 kHz
- **400 peaks from this data as features**