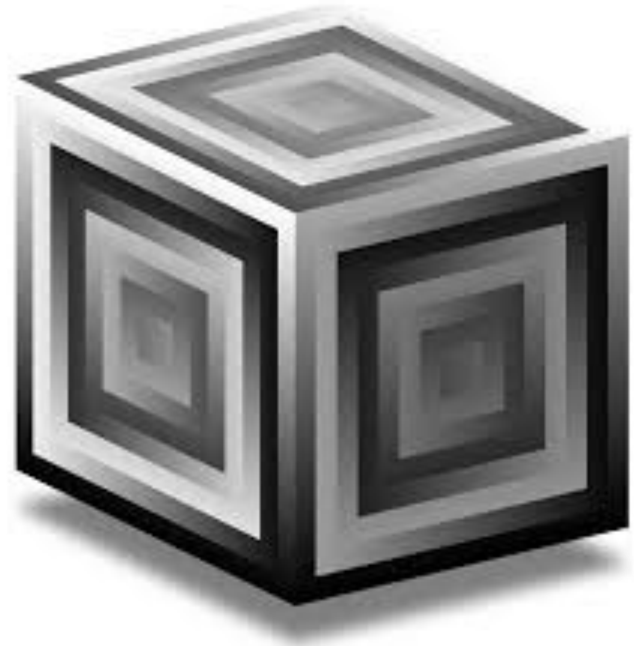
FFT

Topics Addressed

- What is the FFT?
- Frequency bins
- Non-periodic signals
- Spectral Leakage
- Windowing
- Short-Time Fourier Transform
- Additional Topics



FFT

- The Fast Fourier Transform is a quicker way to produce the results of the Discrete Fourier Transform. FFT needs N that are powers of 2.
- The Discrete Fourier Transform:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N}$$

OR

$$X_k = \sum_{n=0}^{N-1} x_n \left[\cos\left(2\pi k \frac{n}{N}\right) - i \sin\left(2\pi k \frac{n}{N}\right)\right]$$

How many X_k?

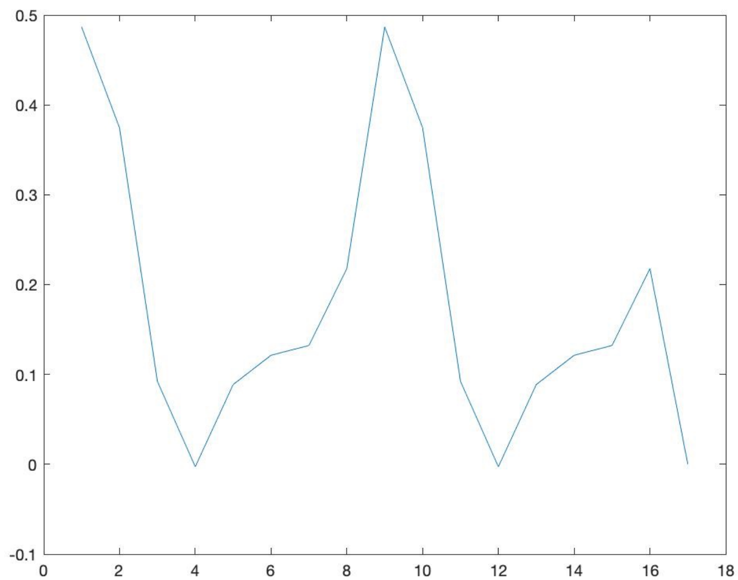
- Recall that the result of the DFT/FFT is a complex number whose real and imaginary part contain the information for the amplitude and phase of some frequency, also referred to as a "frequency bin"
- Recall that the frequency of bin $X_k = (k/N) * f_s$
- The FFT is conducted from $k = 0$ up to $k = N - 1$. These are the complete sinusoids we can capture across the samples we have.
    - This gives you the frequency bins 0, $\left(\frac{1}{N}\right) * f_s$, $\left(\frac{2}{N}\right) * f_s$, … , $\left(\frac{\frac{N}{2}}{N}\right) * f_s$, … , $\left(\frac{N-1}{N}\right) * f_s$
    - $\left(\frac{\frac{N}{2}}{N}\right) * f_s = \frac{1}{2} * f_s$ is the Nyquist frequency. For analysis, we can ignore frequency bins above $\frac{1}{2} * f_s$
    - Bins are separated by $\left(\frac{1}{N}\right) * f_s$

The FFT on some random periodic samples

Amplitude: $\frac{2}{N}\sqrt{\mathrm{Re}(X_k)^2 + \mathrm{Im}(X_k)^2}$

Phase: $\tan^{-1}(\frac{\mathrm{Im}(X_k)}{\mathrm{Re}(X_k)})$

$x = [0.487, 0.375, 0.092, -0.003, 0.089, 0.121, 0.132, 0.218]$



k	X_k	Amplitude	Phase
0	1.6	0.2*	0
1	0.796 + 0.0799i	0.2	0.1
2	0.3510 – 0.1918i	0.1	-0.5
3	0	0	-1.8707
4	0	0*	0
5	0	0	1.8707
6	0.3510 + 0.1918i	0.1	0.5
7	0.796 - 0.0799i	0.2	-0.1

*X_0 and $X_{N/2}$ actually have a different equation amplitude: $\frac{1}{N}\sqrt{\mathrm{Re}(X_k)^2 + \mathrm{Im}(X_k)^2}$

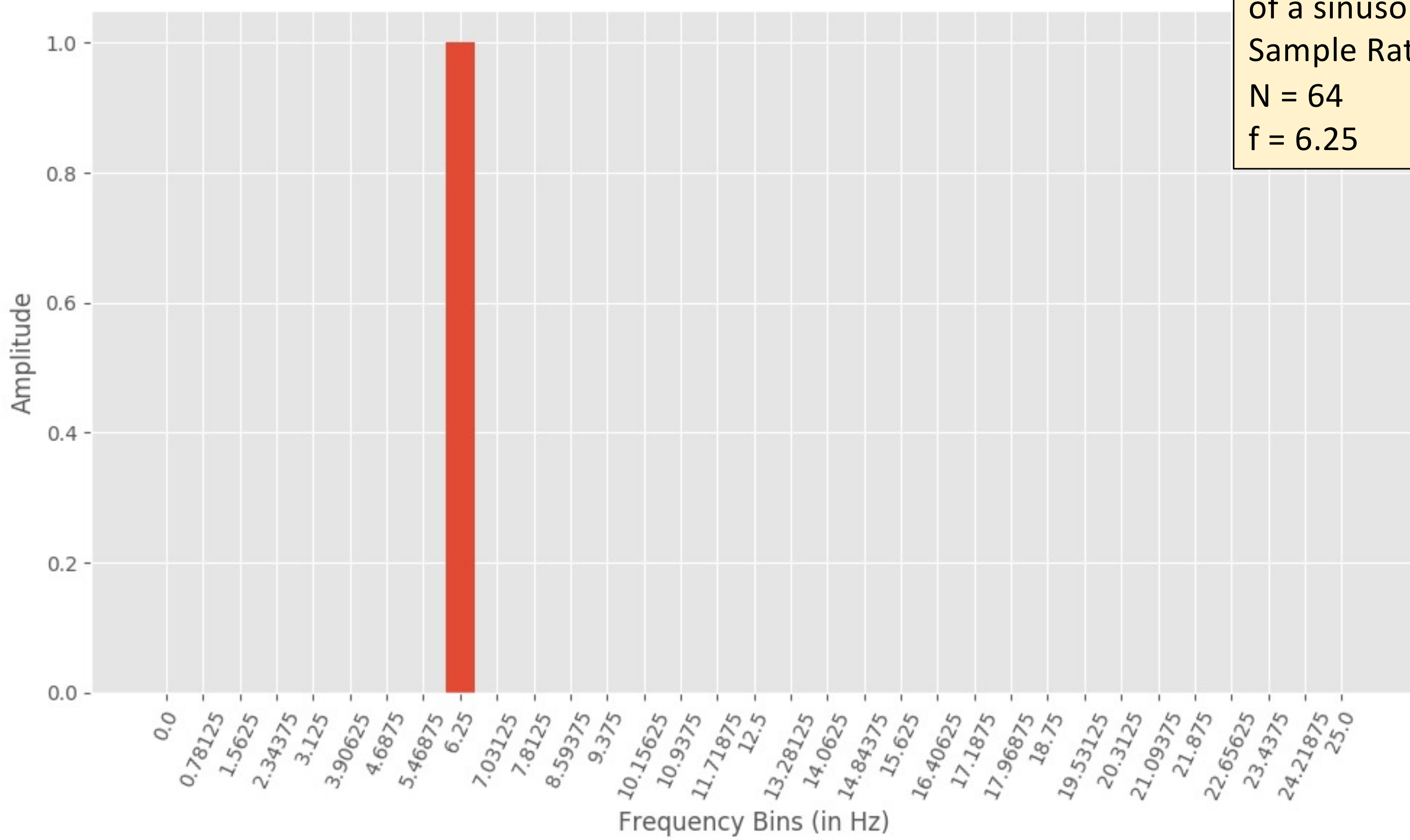
Graph of the three partials: https://www.desmos.com/calculator/locczeg0dr

Example

From the previous slide, what are the frequencies of bins $k = 0, 1, 2$ when $f_s = 16$Hz? Which bin has the Nyquist frequency?

We need to calculate 0, $\left(\frac{1}{N}\right) * f_s$, $\left(\frac{2}{N}\right) * f_s$. So the answer is 0Hz, 2Hz, 4Hz, respectively. Our Nyquist Frequency is 8Hz so $k = 4$.

FFT on $\sin(2\pi(6.25)t)$



The magnitude graph of a sinusoid at 6.25Hz
Sample Rate = 50
N = 64
f = 6.25

A reminder about key assumptions

- The DFT can accurately determine the frequency, phase, and amplitude of a signal with the following conditions:
  - Sampled
  - **Periodic**
  - The N samples must be from complete periods of the signal
  - The samples should not be from aliased signals or the aliasing should be minimal/suppressed

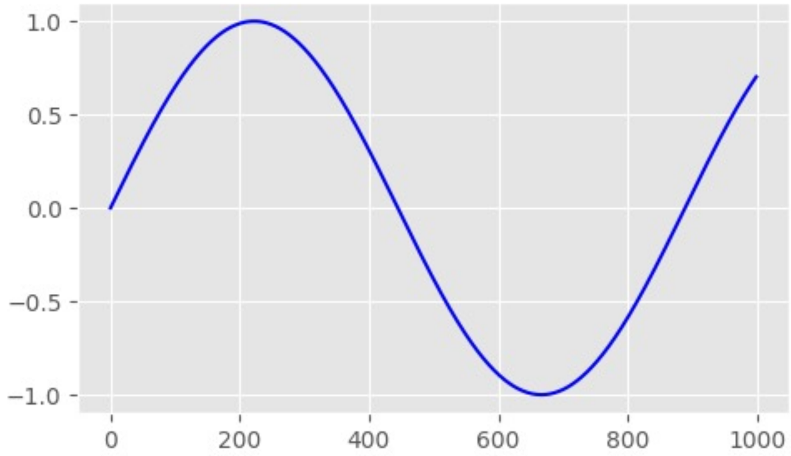
- The fact that the DFT assumes periodicity is most problematic. Nearly all audio is not perfectly periodic unless you are dealing with triangle waves, sawtooth waves, etc.

What if your signal is not periodic…

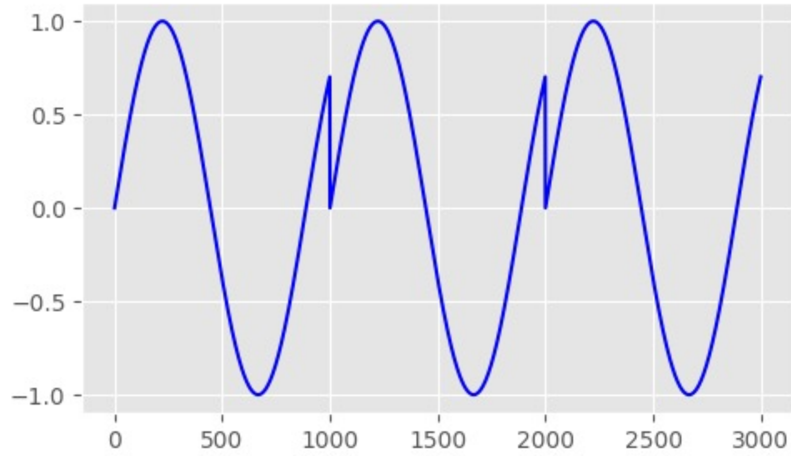
FFT on aperiodic signals

- The FFT/DFT is designed to work on a **periodic** number of samples and accurately capture their phase/amplitude.
- How FFT views an **aperiodic** number of samples:

Discontinuities produce rich harmonics!

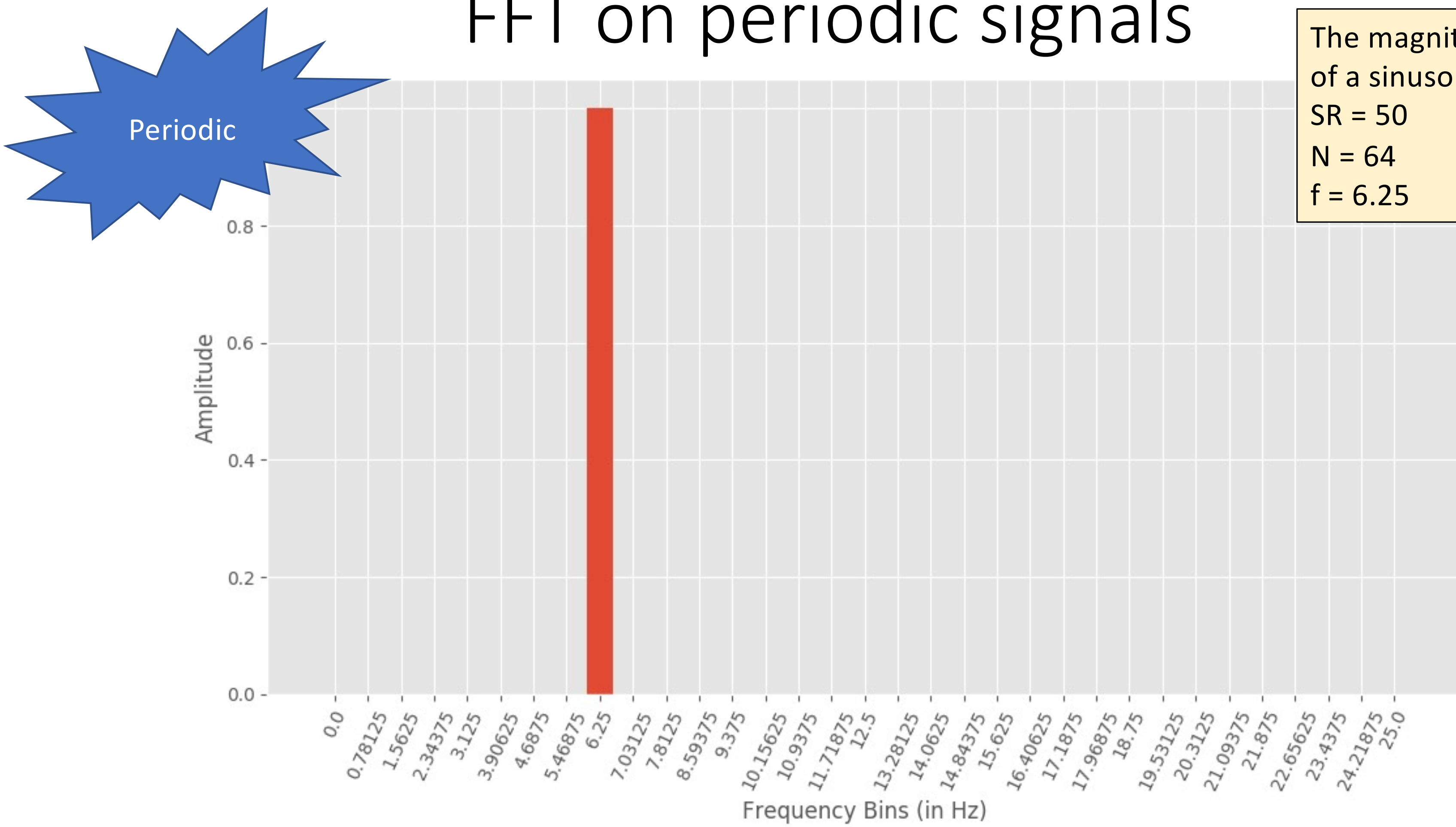


Aperiodic signal



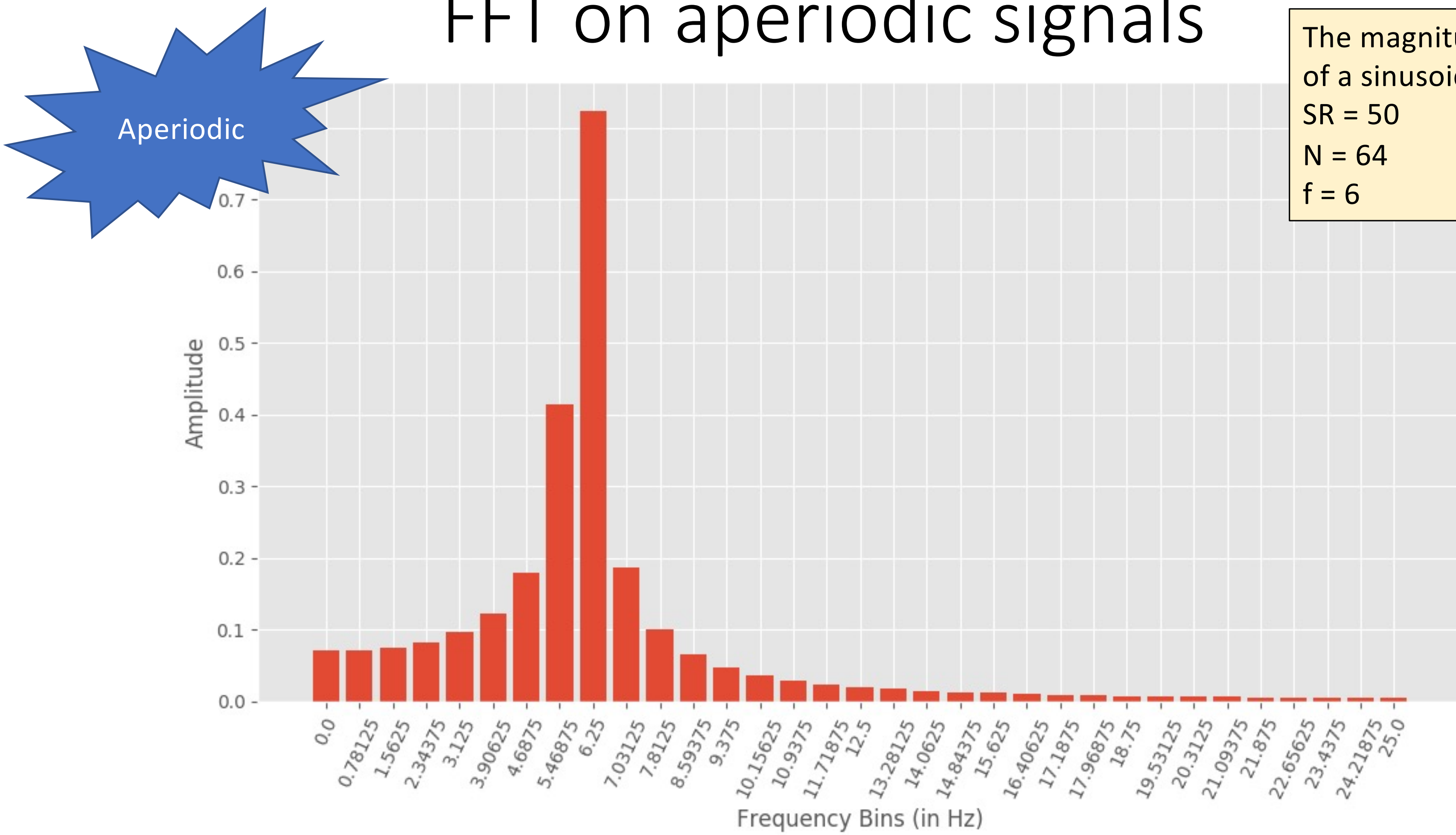
Periodic aperiodic signal

FFT on periodic signals



The magnitude graph of a sinusoid at 6.25Hz
SR = 50
N = 64
f = 6.25

FFT on aperiodic signals



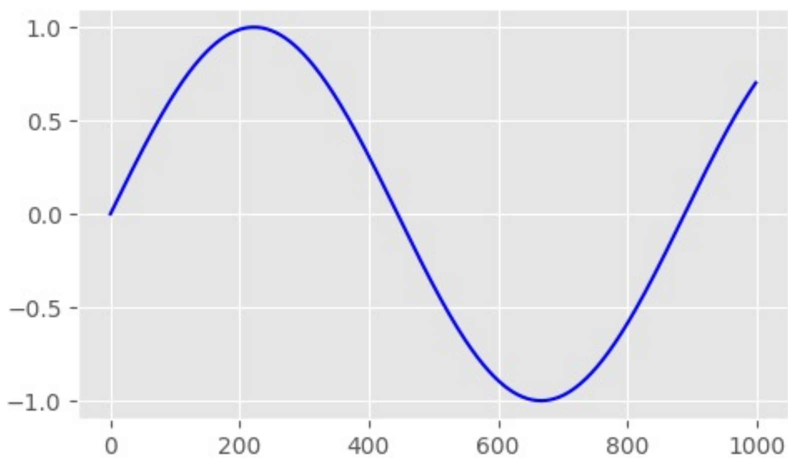
The magnitude graph of a sinusoid at 6.25Hz
SR = 50
N = 64
f = 6

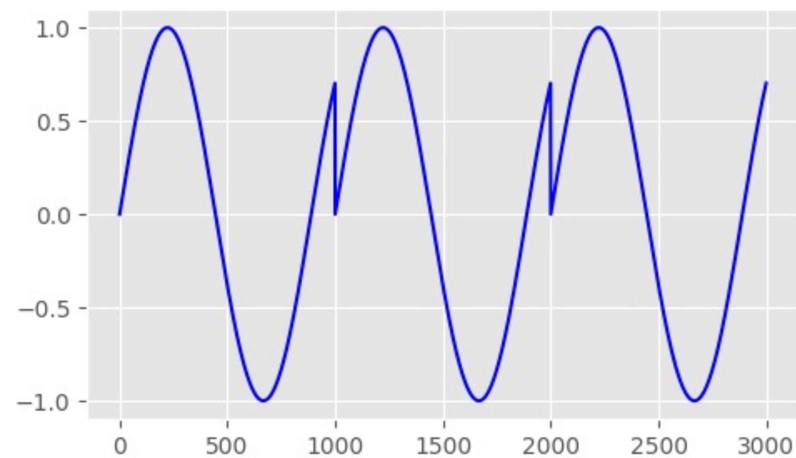
FFT on aperiodic signals

- Aperiodic samples are treated as periodic samples in an FFT/DFT. Many times this leads to sharp discontinuities at the edges.
- The discontinuities lead to rich harmonics that spread out across the entire frequency spectrum.
- We call this spread “spectral leakage”
  - Spectral leakage can make it difficult to determine the original frequency.
  - Spectral leakage makes it **impossible** to determine the original amplitude and phase.

How To Interpret This Information?

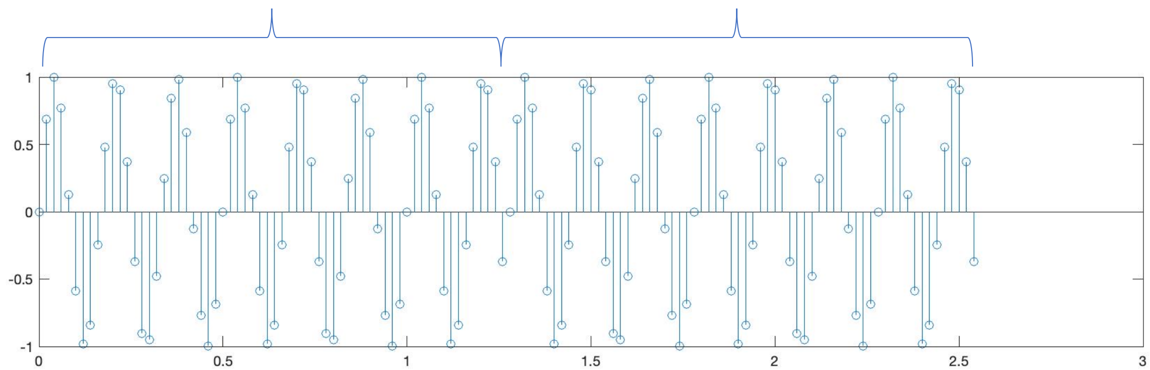
- The DFT gives a set of harmonic cosines that can approximate any periodic function. We **assume** we are approximating the function on the right, unfortunately.
- Therefore, we need to use **many** cosine waves to represent a signal with discontinuities. Hence, the spectral leakage.





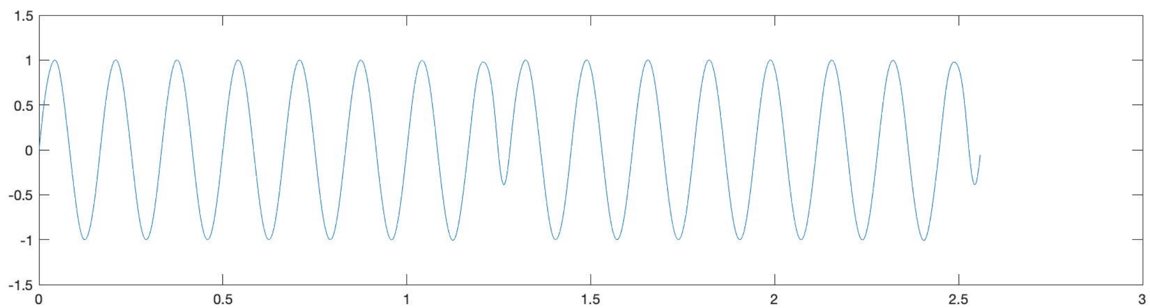
How To Interpret This Information?

Original signal samples



2 periods of $N = 64$ samples at $f_s = 50$ lasting $\frac{N}{f_s} = 1.28$ seconds

Cosine reconstruction

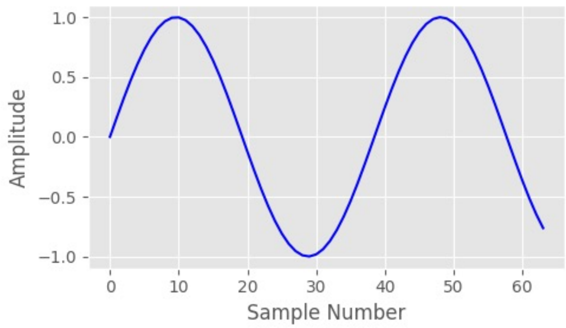


Windows

- Spectral leakage is a problem that can be combatted by the use of windows.
- What is a window? The idea here is to point-wise multiply an N sampled audio signal by a window of the same length to remove the discontinuities at the edges.
- Many different types of windows:
  - Rectangular
  - Hamming
  - Hann (Raised sinusoid)
  - Sine

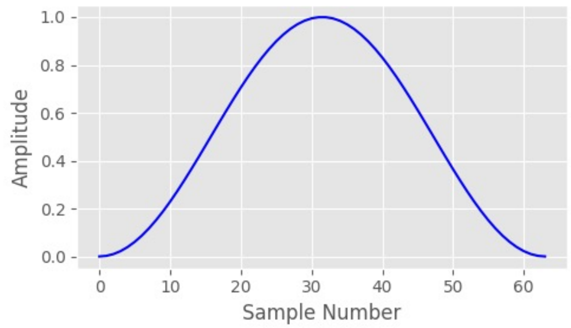
Windowing in Action

- Windowing is just pointwise amplitude multiplication *before* taking the FFT of the signal.



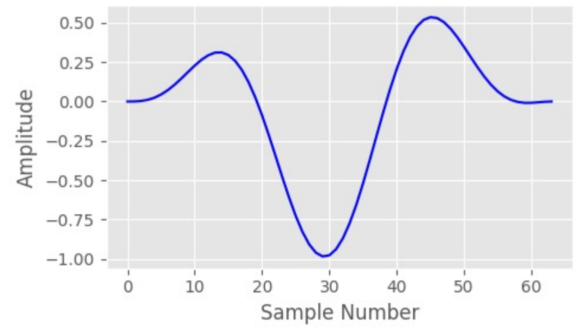
Aperiodic signal

A periodic sinusoids create discontinuities at the edges leading to spectral leakage.



Hann window

A window approaches zero near the edges to smooth out discontinuities at the edges.



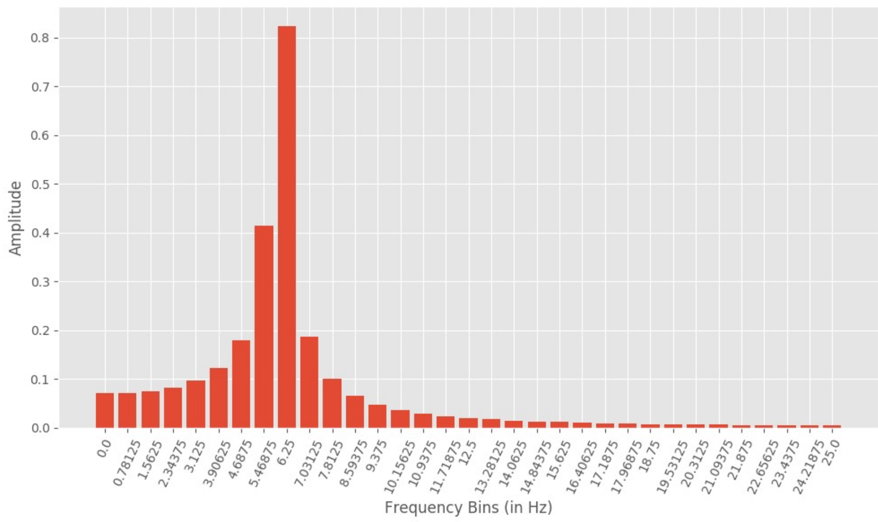
Output

Our originally aperiodic signal is now periodic!

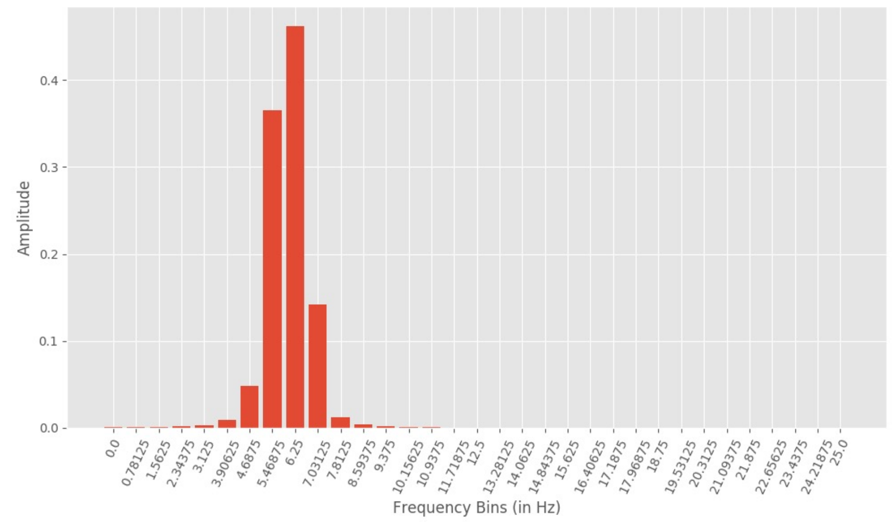
Window Comparison

- Here is the Hann window applied to our earlier example of an aperiodic sine wave of frequency 6Hz. Notice that the windowed version reduces the spread (i.e., width) of the spectral leakage.

Without windowing

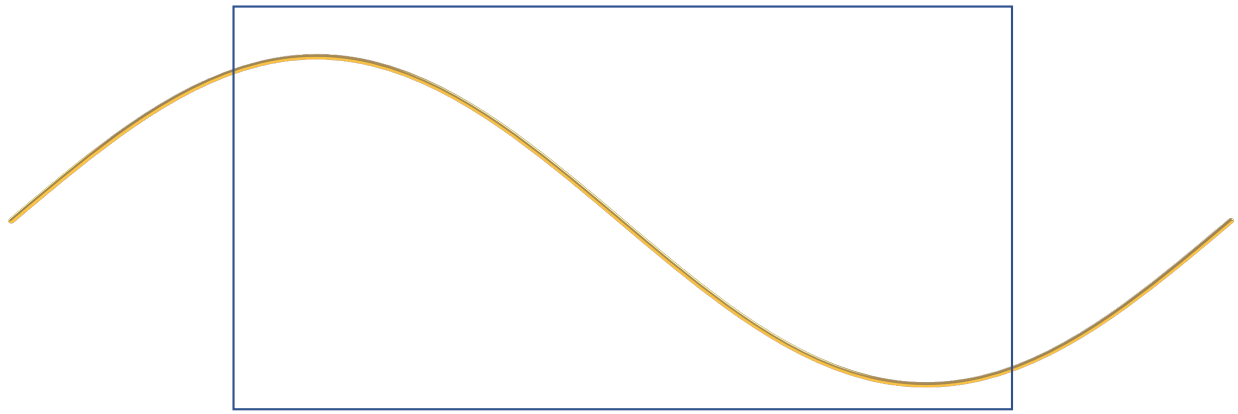


With windowing



Rectangular Windows

- Sampling a signal is equivalent to using a rectangular window on the signal.
- Rectangular windows like any window alters the shape of the signal, here by creating discontinuities at the edges.
- Generally want to avoid rectangular windows for audio processing or analysis. Therefore, we rarely use just the FFT of a sample of signals (i.e., a rectangular window). Instead, we take the FFT of a sample of signals after windowing with a Hann/Hamming/Sine window.



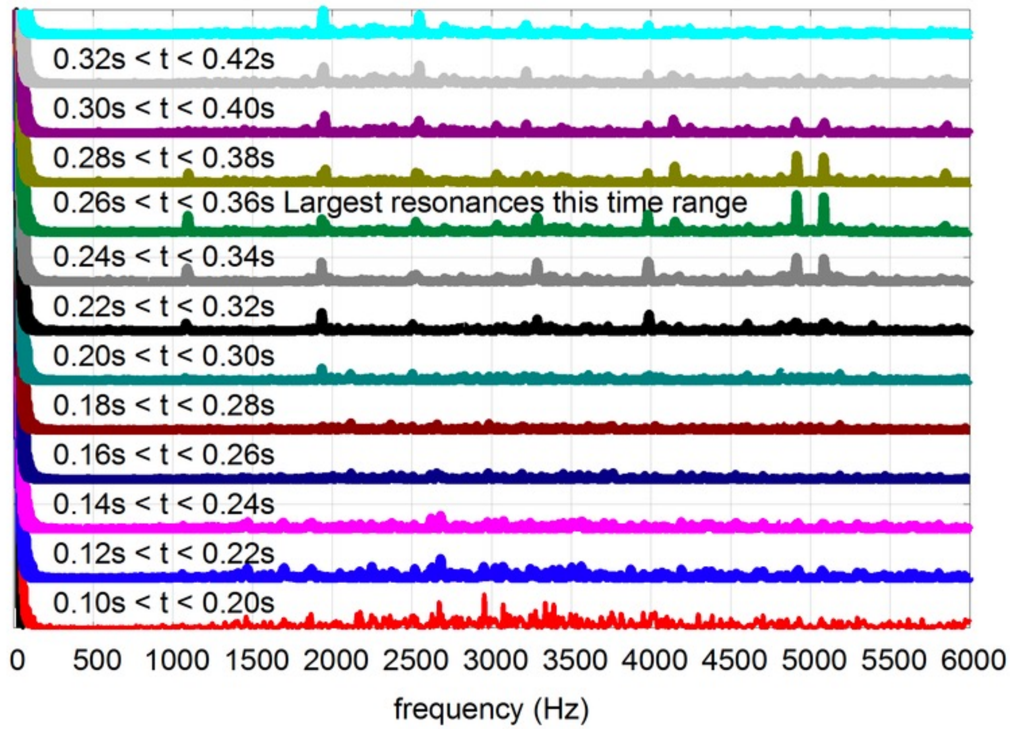
Windowing Drawbacks

- Applying a window to a signal changes fundamental components of the original signal.
  - Even a rectangular window does so if the window does not capture an integer number of periods (which we almost always do in practical situations)
- Windowing cannot completely isolate each frequency
  - Still restricted by the frequency of our frequency bins
  - Windowing reduces spectral leakage but does not remove it
- Window size causes a tension between frequency resolution and time resolution
  - Larger window – more frequency resolution, but captures more time
    - Less likely that the signal is periodic
    - More changes in the frequency content over time. FFT gives a snapshot so won't capture change.
  - Small window – good time resolution, but poorer frequency resolution
    - Less likely that the frequency bins will accurately represent the frequency spectrum because there are fewer bins spaced farther apart linearly.

Short-Time Fourier Transform

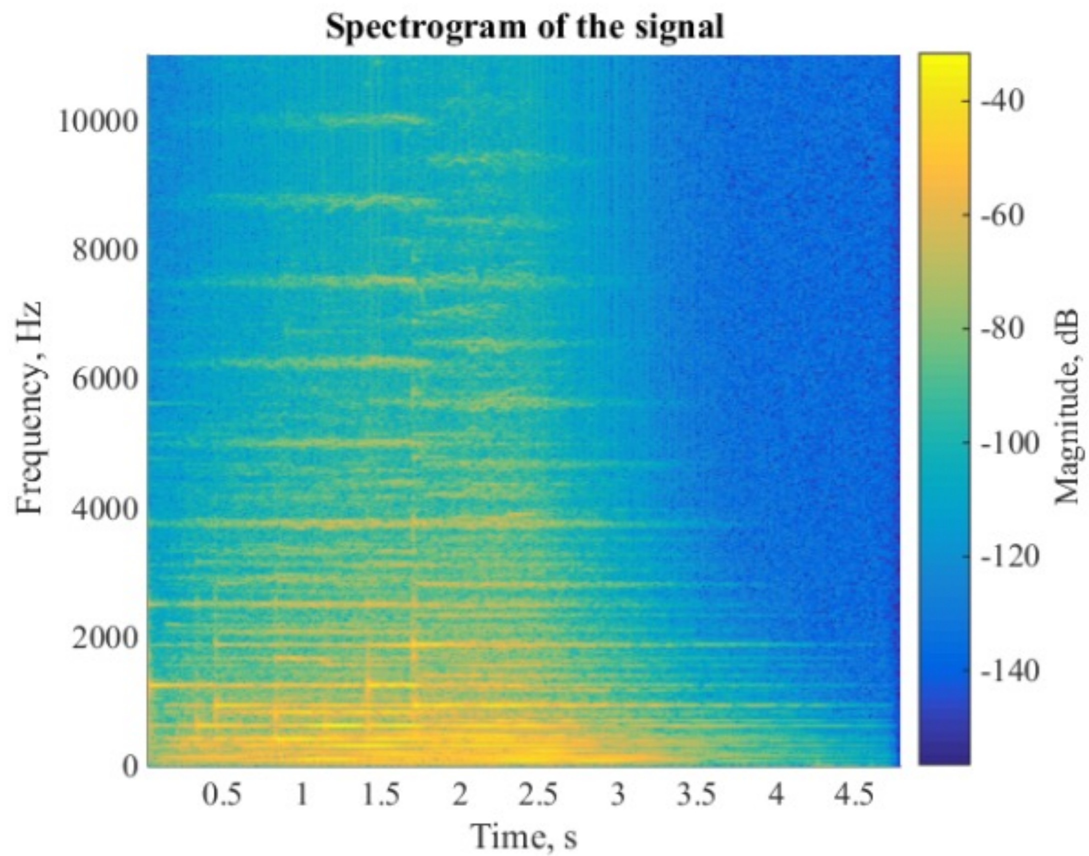
- Problem: the frequency spectrum of an audio file changes consistently over time. How can we capture the changing frequency spectrum?
  - Answer: take snapshots of the frequency spectrum over time
- Divide a long time period into shorter periods in which we calculate the FFT on each shorter period
  - Generally these periods overlap
  - Discrete and continuous versions (we care about discrete)
  - Can be inverted to original signal with inverse STFT through a process called overlap-add method (OLA). Allows for modification and processing. This is what SuperCollider does!

Short-Time Fourier Transform



Short-Time Fourier as Spectrogram

What can we determine about our signal by looking at this picture?



Additional Topics

- Windowing techniques – each window has pros and cons depending upon application
- Convolution
- Inverse Fast Fourier Transform – the process of taking the complex numbers of the FFT and reconstructing the original audio signal.
- Wavelet Transform – short-time fourier transform in which the size of the window can vary with time. Attempts to gain good time and frequency resolution across different parts of audio signal.
- So much more…