ENGR-E 511; ENGR-E 399

"Machine Learning for Signal Processing"

Module 02: Lecture 01:

Time-Frequency Transforms

Minje Kim

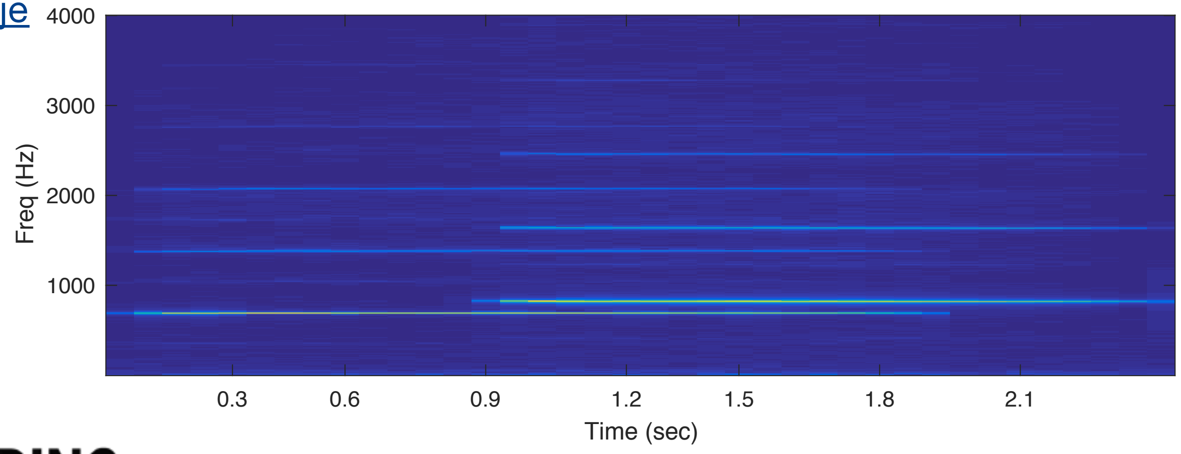
Department of Intelligent Systems Engineering

Email: minje@indiana.edu

Website: http://minjekim.com

Research Group: http://saige.sice.indiana.edu

Meeting Request: http://doodle.com/minje



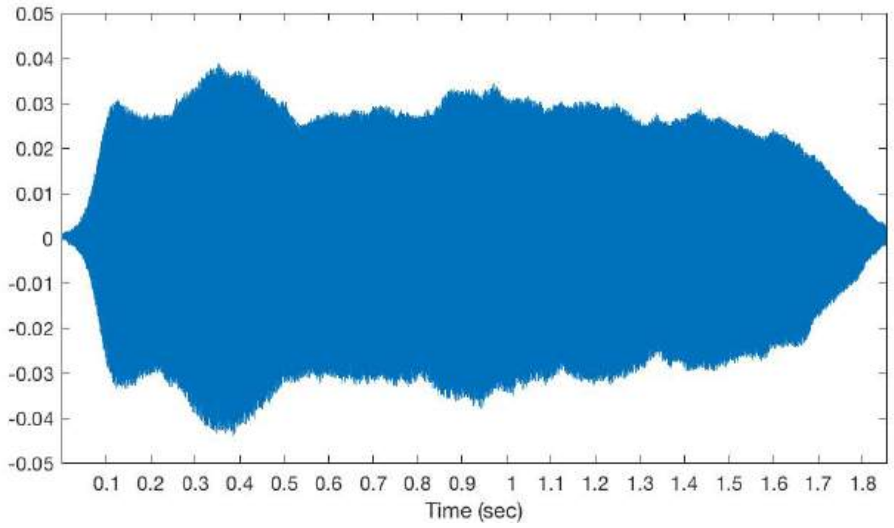
INDIANA UNIVERSITY

SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

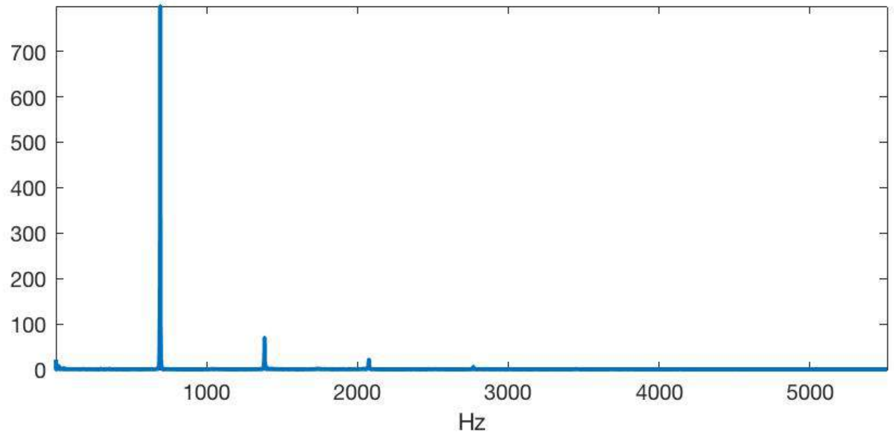
Frequency Representation of Signals

- Compactness

- How does it sound?



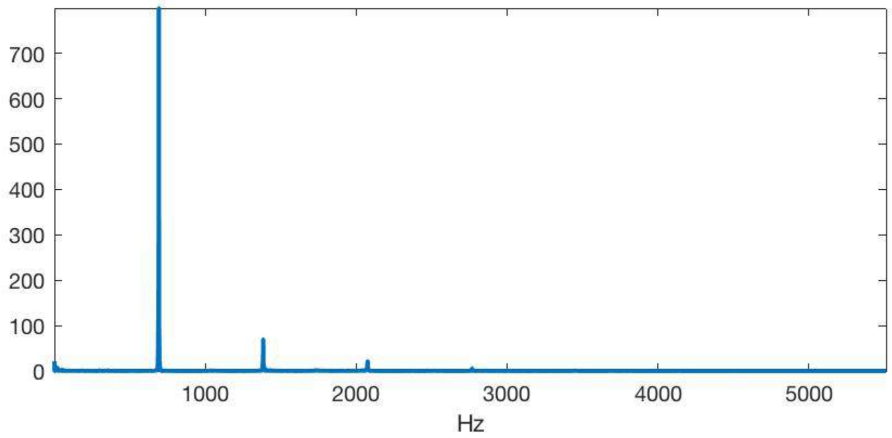
- It actually sounds like this:
- A better way to represent the signal?
  - How about this?



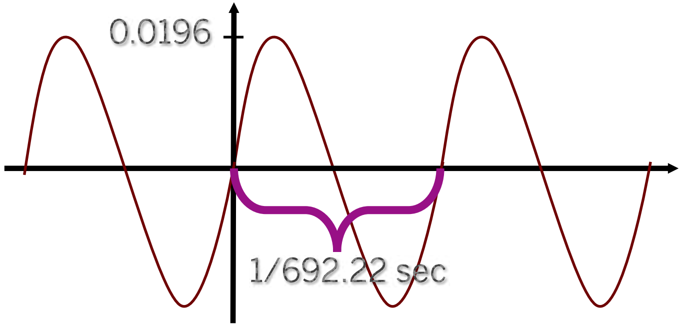
Frequency Representation of Signals

- Compactness

- I prefer the frequency representation for this signal
  - Because it says about the signal in a compact way
  - i.e. it's a mixture of four sine waves, whose frequencies and amplitudes are defined by the peaks in the spectrum



$$\begin{aligned}&{\color{blue}0.0196}\sin\left(2\pi{\color{red}692.22}t\right)\\&+{\color{blue}0.0017}\sin\left(2\pi{\color{red}1384.45}t\right)\\&+{\color{blue}0.0005}\sin\left(2\pi{\color{red}2077.21}t\right)\\&+{\color{blue}0.0001}\sin\left(2\pi{\color{red}2768.9}t\right)\\&+\cdots\end{aligned}$$



- This (manual) frequency analysis misses some other components, but is quite good

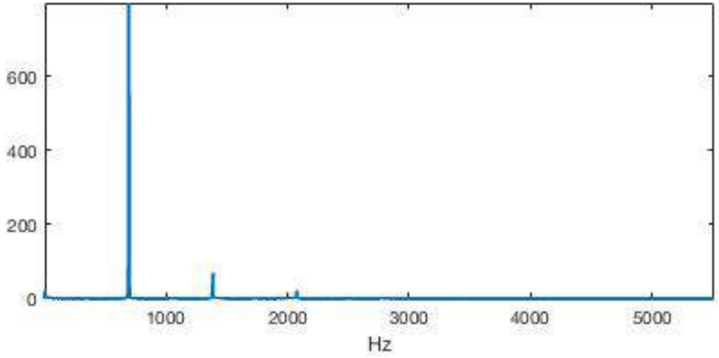
Reconstruction from 4 sinusoids

Original

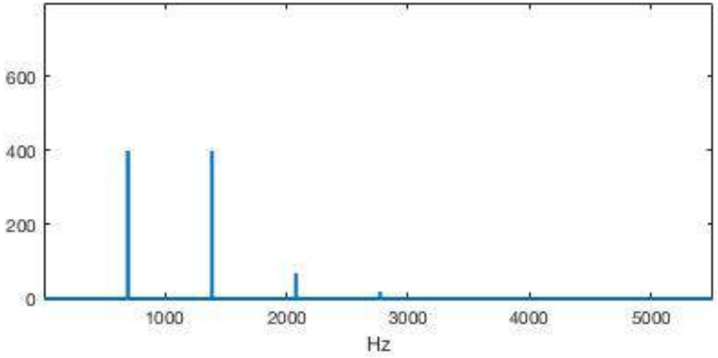
Frequency Representation of Signals

- Amplitudes of the basis functions

- Same sinusoids but with different amplitudes
- Height of the spectral peaks reflects amplitudes



$$0.0196 \sin\left(2\pi 692.22t\right) + 0.0017 \sin\left(2\pi 1384.45t\right) + 0.0005 \sin\left(2\pi 2077.21t\right) + 0.0001 \sin\left(2\pi 2768.9t\right)$$

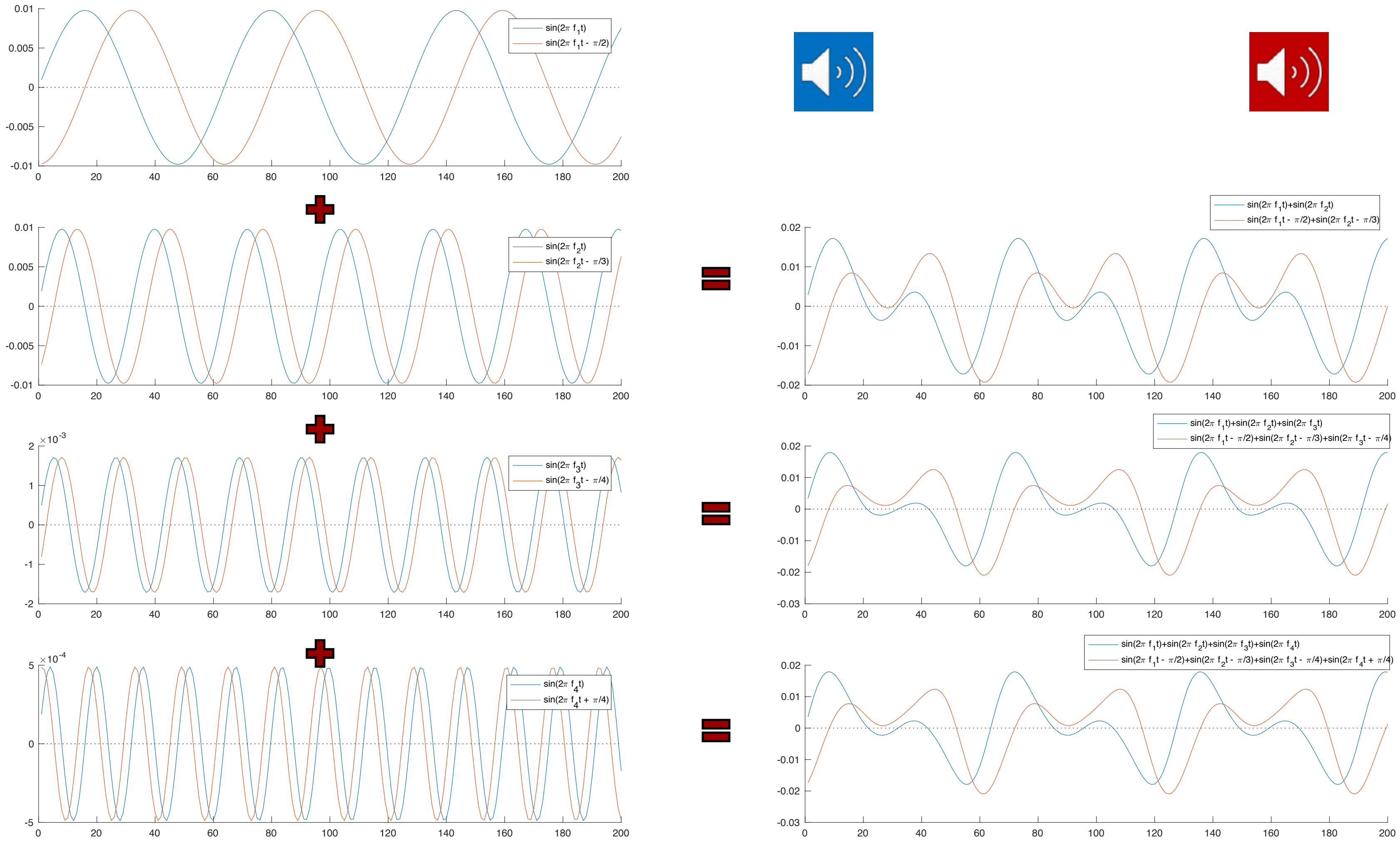


$$0.0098 \sin\left(2\pi 692.22t\right) + 0.0098 \sin\left(2\pi 1384.45t\right) + 0.0017 \sin\left(2\pi 2077.21t\right) + 0.0005 \sin\left(2\pi 2768.9t\right)$$

- Any guess as to how they would sound?
  - They sound different!

Frequency Representation of Signals

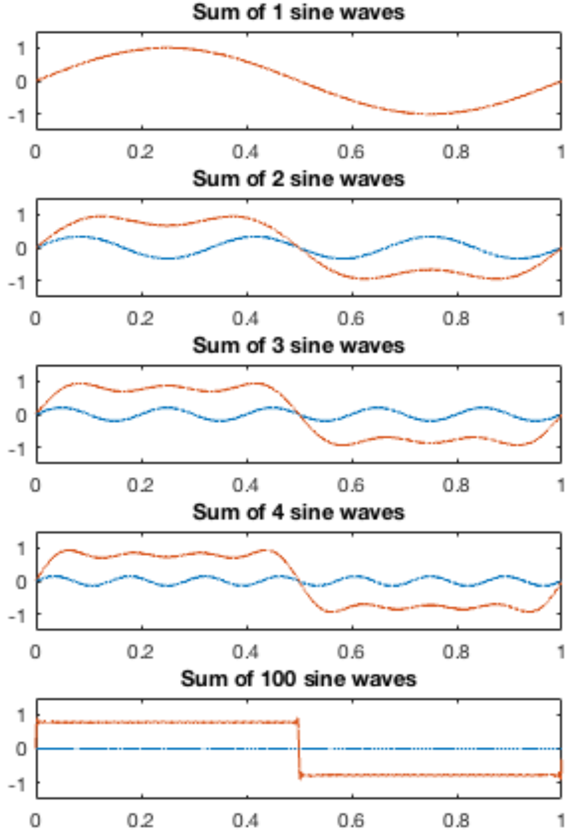
- Phase of the basis

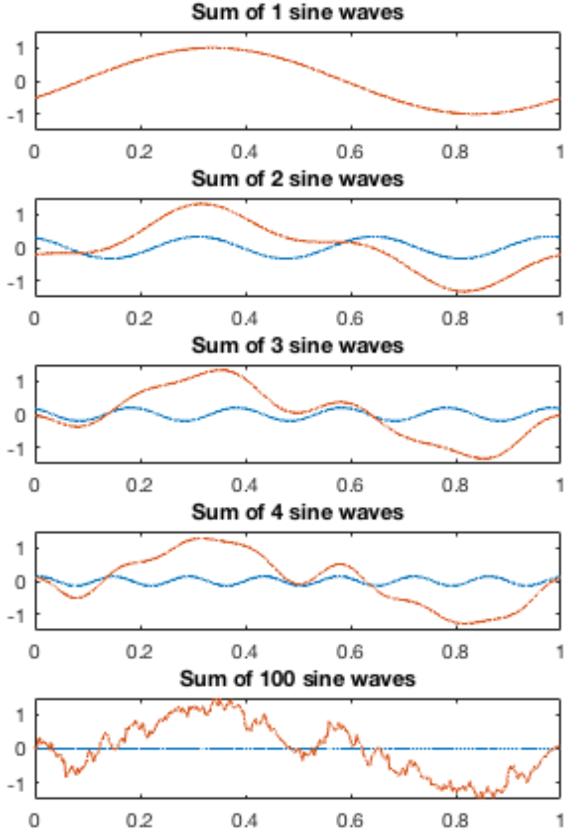


- But they sound the same!

Frequency Representation of Signals

- Phase of the basis functions



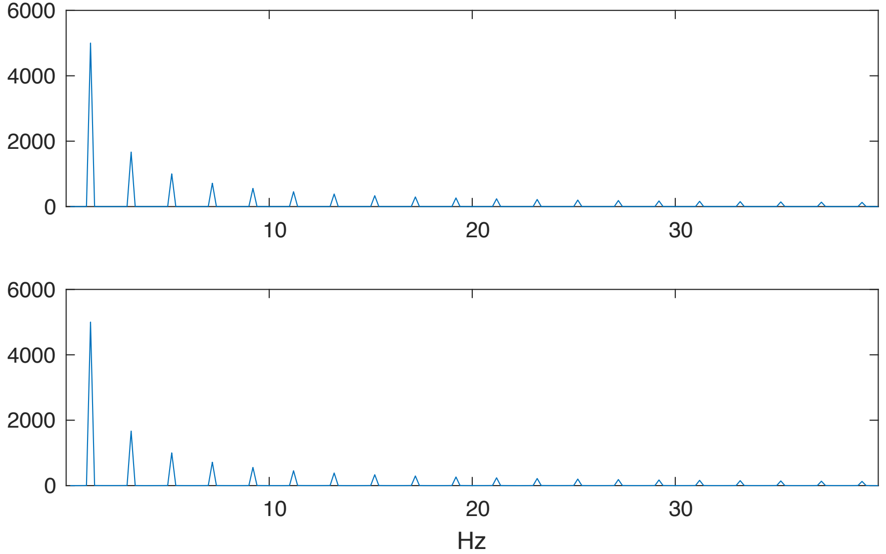


- Now they sound different!

Frequency Representation of Signals

- Magnitude spectra

- So, the “magnitude” spectrum representation is somehow misleading



Box Wave

Box Wave with random phase

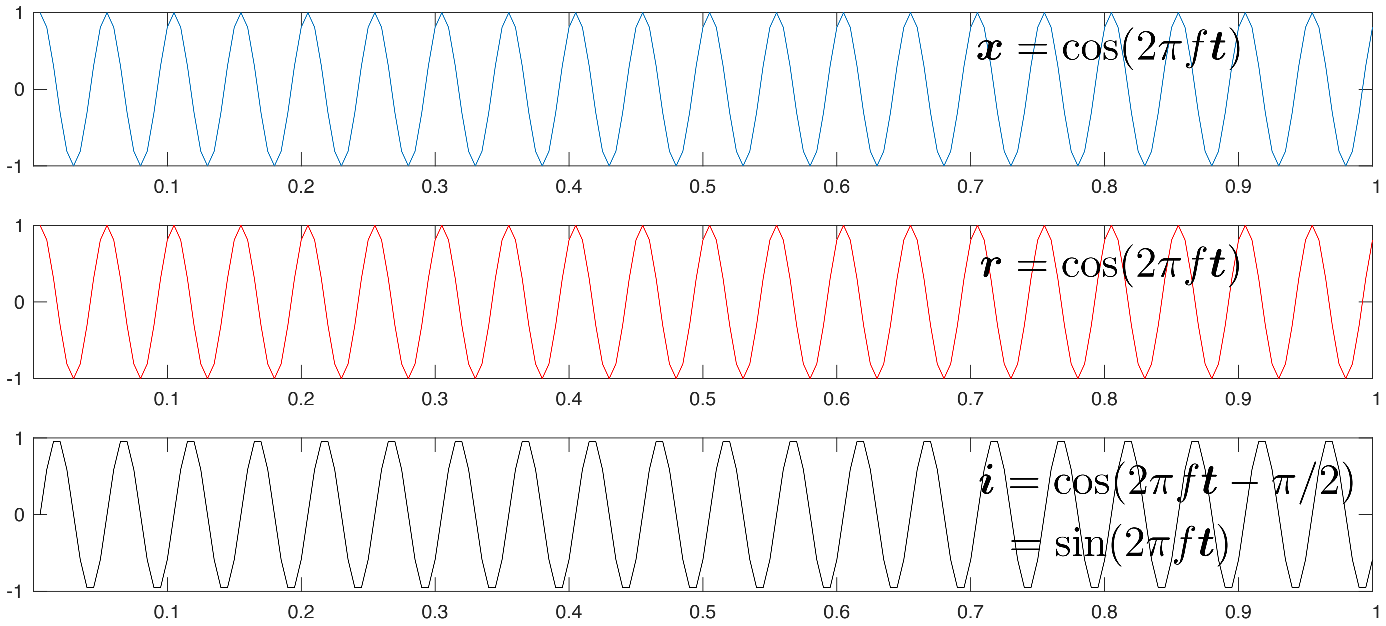
- The magnitudes of the box wave and its random-phase version are the same
- Representing a signal with a sum of sinusoids
  - Needs period, amplitude, and phase of each of the sinusoids
- How do we find this information?

Discrete Fourier Transform

- Basis vectors with different phase

- We are projecting a signal $\boldsymbol{x}$ onto basis vectors $\boldsymbol{r}$ and $\boldsymbol{i}$

$$\boldsymbol{x} = \cos(2\pi f \boldsymbol{t}) \quad f = 20Hz, \ \boldsymbol{t} = [0, 0.0001, 0.0002, \cdots, 1]$$

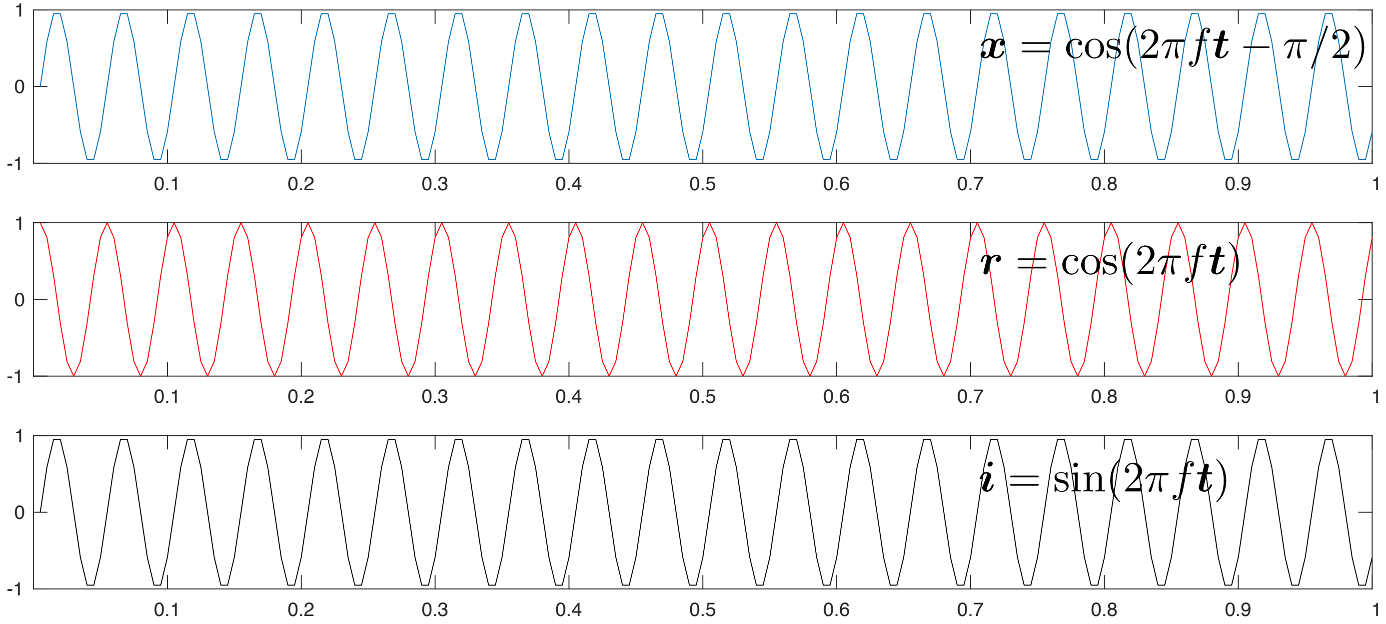


$$\boldsymbol{x} \cdot \boldsymbol{r} = 100$$
$$\boldsymbol{x} \cdot \boldsymbol{i} = 0$$

Discrete Fourier Transform

- Basis vectors with different phase

$$f = 20Hz,\ \boldsymbol{t} = [0, 0.0001, 0.0002, \cdots, 1]$$



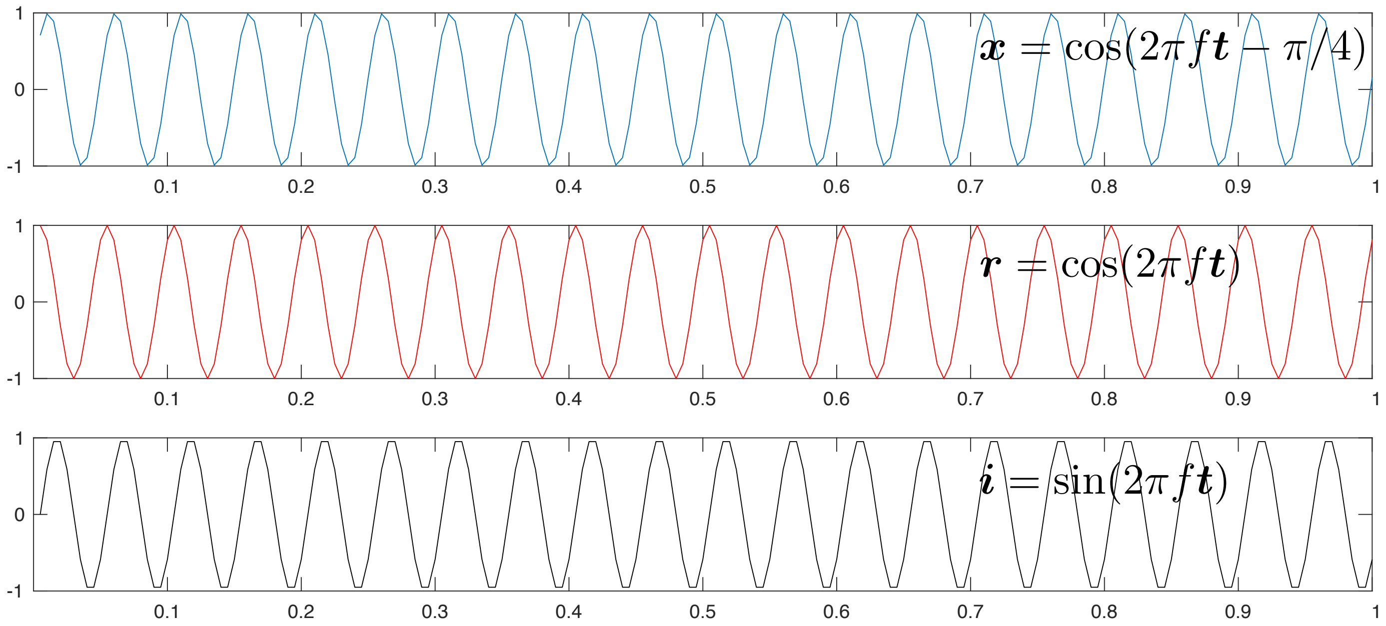
$$\boldsymbol{x} \cdot \boldsymbol{r} = 0$$

$$\boldsymbol{x} \cdot \boldsymbol{i} = 100$$

Discrete Fourier Transform

- Basis vectors with different phase

$$f = 20Hz, \ \boldsymbol{t} = [0, 0.0001, 0.0002, \cdots, 1]$$



$$\boldsymbol{x} \cdot \boldsymbol{r} = 70.71$$

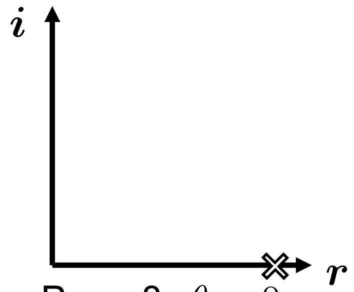
$$\boldsymbol{x} \cdot \boldsymbol{i} = 70.71$$

$$\sqrt{70.71^2 + 70.71^2} = 100$$

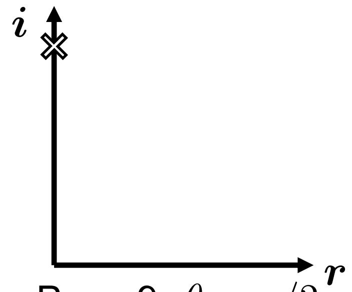
Discrete Fourier Transform

- Basis vectors with different phase

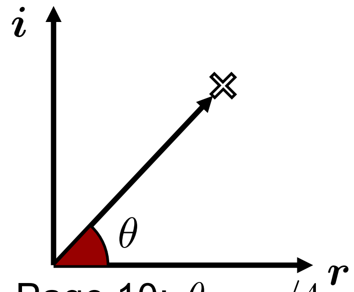
- We defined two orthogonal basis vectors $\boldsymbol{r}$ and $\boldsymbol{i}$
  - Which are the same cosine functions with a phase shift by $\pi/2$



Page 8: $\theta = 0$



Page 9: $\theta = \pi/2$



Page 10: $\theta = \pi/4$

- Two coefficients: $X_r(f) = \boldsymbol{x} \cdot \boldsymbol{r} = \sum_n x[n] \cos(2\pi f \frac{n}{N})$

$$X_i(f) = \boldsymbol{x} \cdot \boldsymbol{i} = \sum_n x[n] \sin(2\pi f \frac{n}{N})$$

- Or, using Euler's rule:

$$X(f) = \boldsymbol{x} \cdot \boldsymbol{r} - j\boldsymbol{x} \cdot \boldsymbol{i} = \sum_n x[n] \left\{ \cos\left(2\pi f \frac{n}{N}\right) - j \sin\left(2\pi f \frac{n}{N}\right) \right\}$$

$$= \sum_n x[n] \exp\left\{ -j\left(2\pi f \frac{n}{N}\right) \right\}$$

- So, why are we doing this?
- The projection preserves information
  - It's onto two basis vectors that are with the same frequency with my input signal
  - Basis vectors are orthogonal (shifted by $\pi/2$)
  - Coefficients preserve the amplitudes (100)
  - Coefficients preserve the phase of the signal $\theta = \pi/4$
- This procedure introduces a new representation

  $\boldsymbol{x}$ versus $(\boldsymbol{X}_r(f), \boldsymbol{X}_i(f))$

Discrete Fourier Transform

- Definition

- We just analyzed a sinusoid with a fixed frequency
  - By using a pair of cos and sin functions with a matching frequency
- What if x is not a pure sinusoid, e.g. a weighted sum of multiple sinusoids?
  - We repeat this procedure for various pre-defined pairs of cos/sin basis by varying f

$$X[f] = \sum_{n=0}^{N-1} x[n] \exp\left\{ -j\left(2\pi f \frac{n}{N}\right)\right\}$$

  - It's called **discrete Fourier transform**
  - The magnitude of the coefficient tells the amount of the contribution of the given frequency
  - Its phase encodes the phase shift of the sinusoid

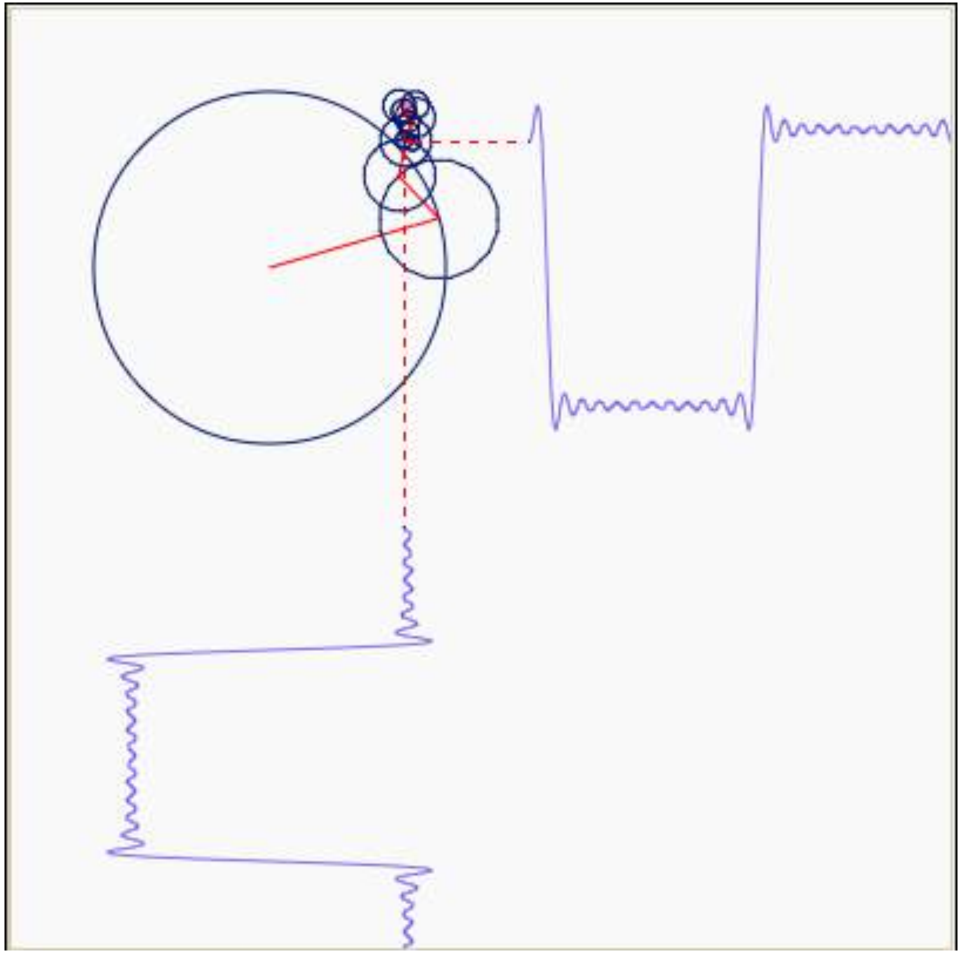
$$\angle(X[f]) = \arctan\left(\frac{\sum_n x[n]\sin(2\pi f \frac{n}{N})}{\sum_n x[n]\cos(2\pi f \frac{n}{N})}\right)$$

- It has its inverse transform as well:

$$x[n] = \frac{1}{N}\sum_{f=0}^{N-1} X[f] \exp\left\{ j\left(2\pi f \frac{n}{N}\right)\right\}$$

Discrete Fourier Transform

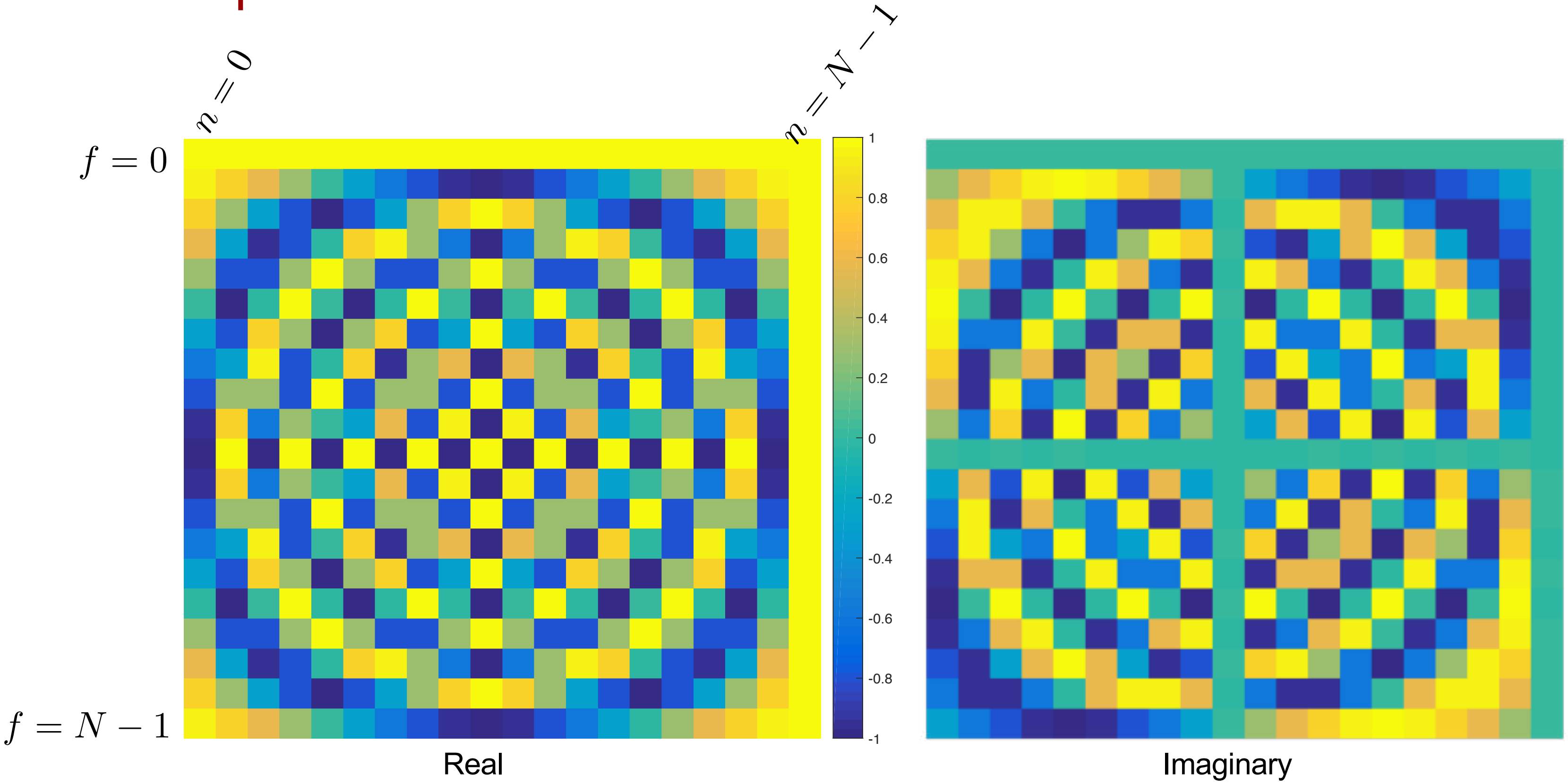
- Box wave example



http://i.imgur.com/BuO2INb.gif

Discrete Fourier Transform

- As a matrix multiplication

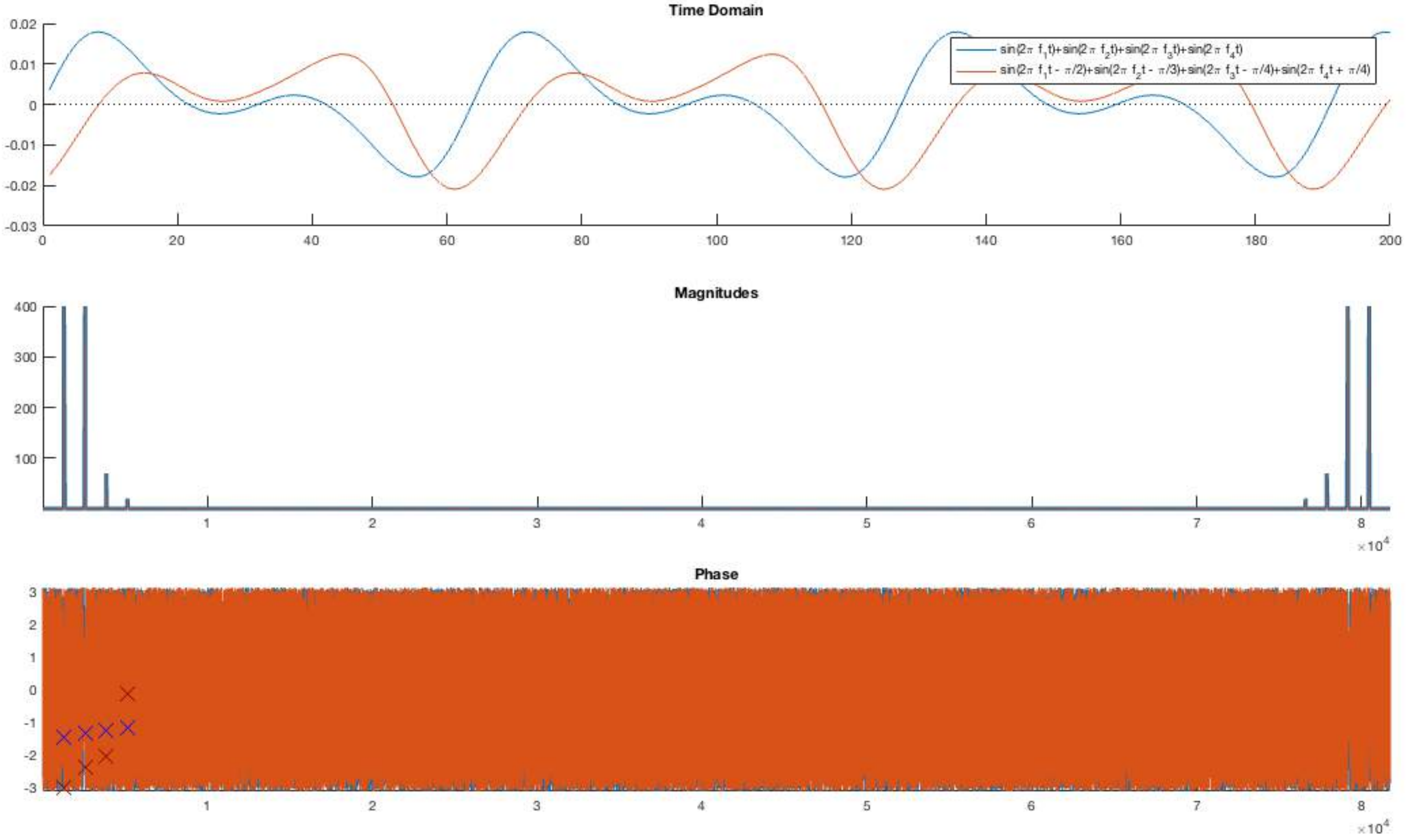


$$X_r(f) = \sum_n x[n]\cos(2\pi f\frac{n}{N})$$

$$X_i(f) = \sum_n x[n]\sin(2\pi f\frac{n}{N})$$

Discrete Fourier Transform

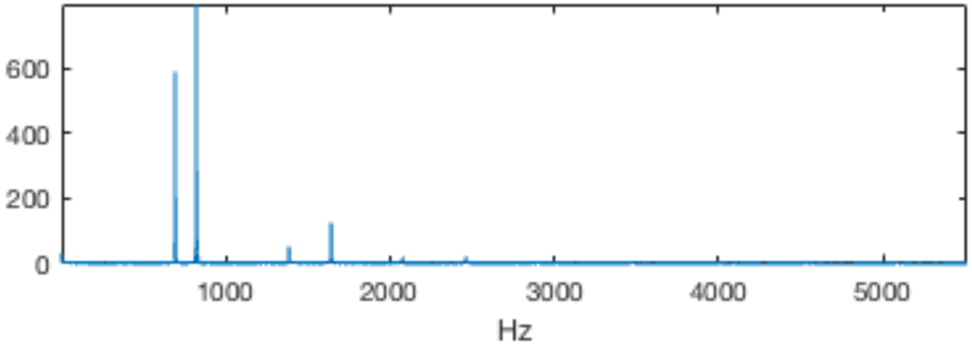
- Magnitudes VS phase



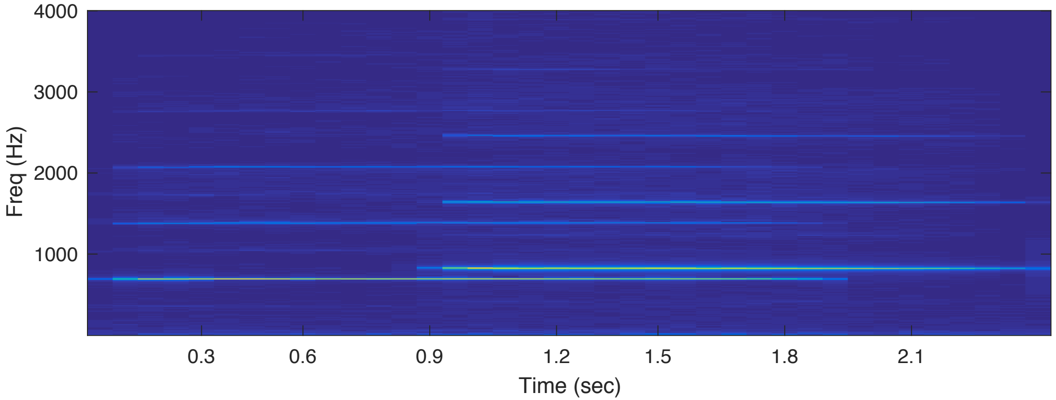
Short-Time Fourier Transform

- Time-frequency representation

- Two notes: 
- Do you like this representation?

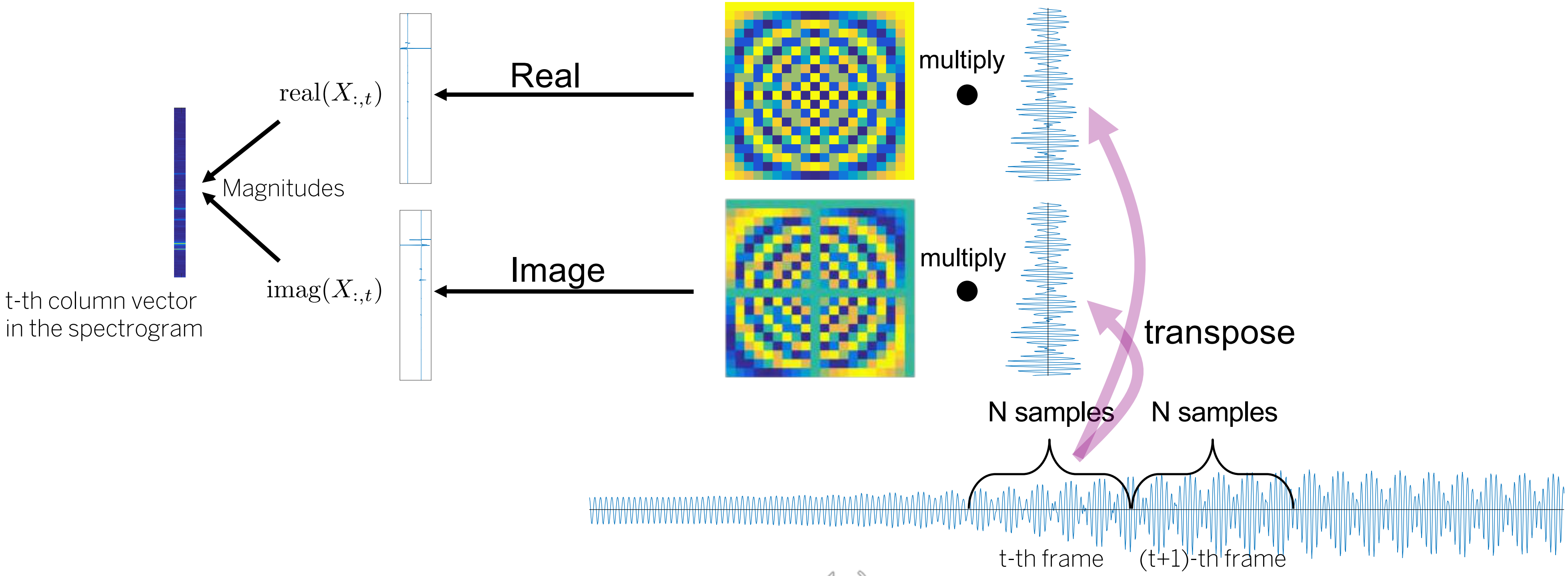


- We need some temporal structure in the figure



Short-Time Fourier Transform
- The mechanics

- What's the magic?
  - Slice the signal into pieces and then apply DFT one by one

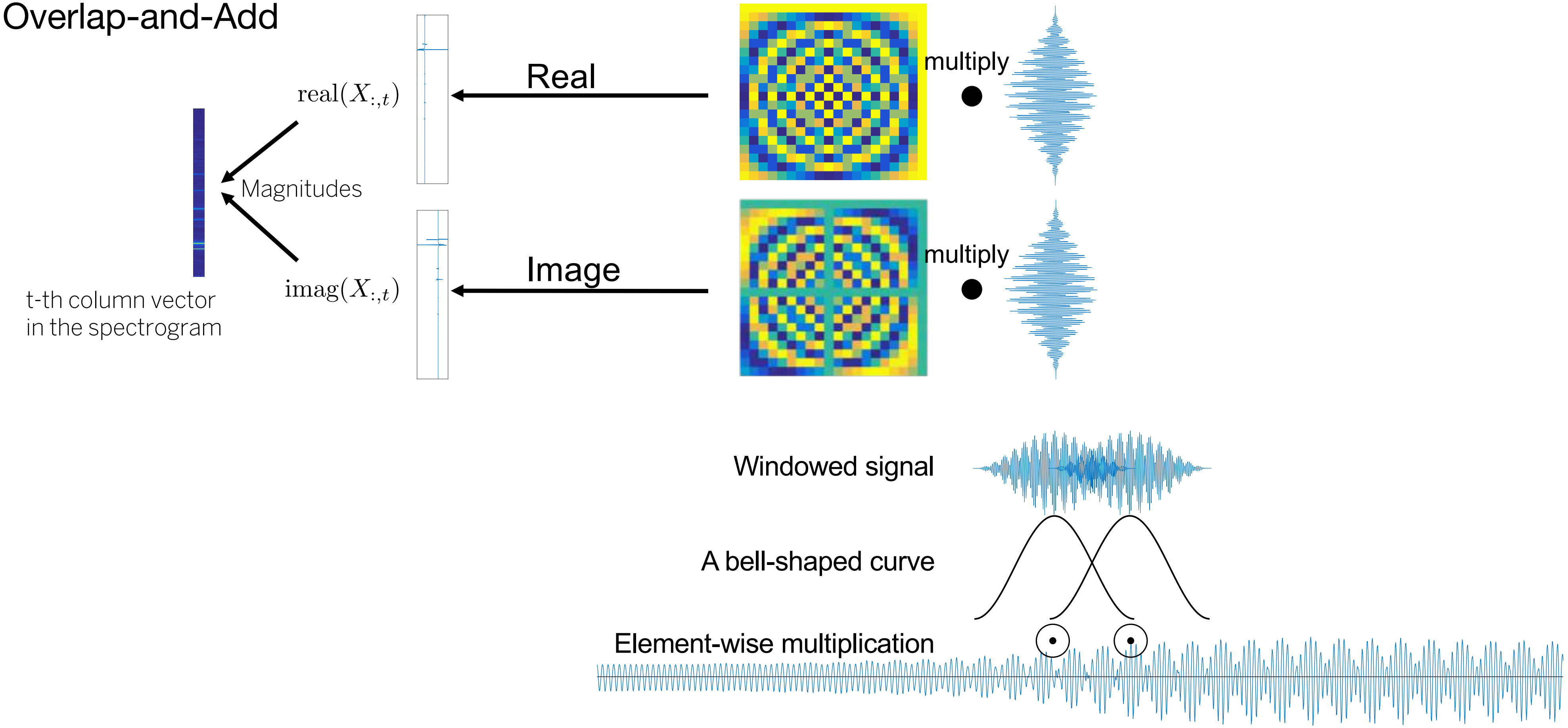


- There's discontinuity we need to take care of

Short-Time Fourier Transform

- Windowing and overlap-and-add

- Overlap-and-Add



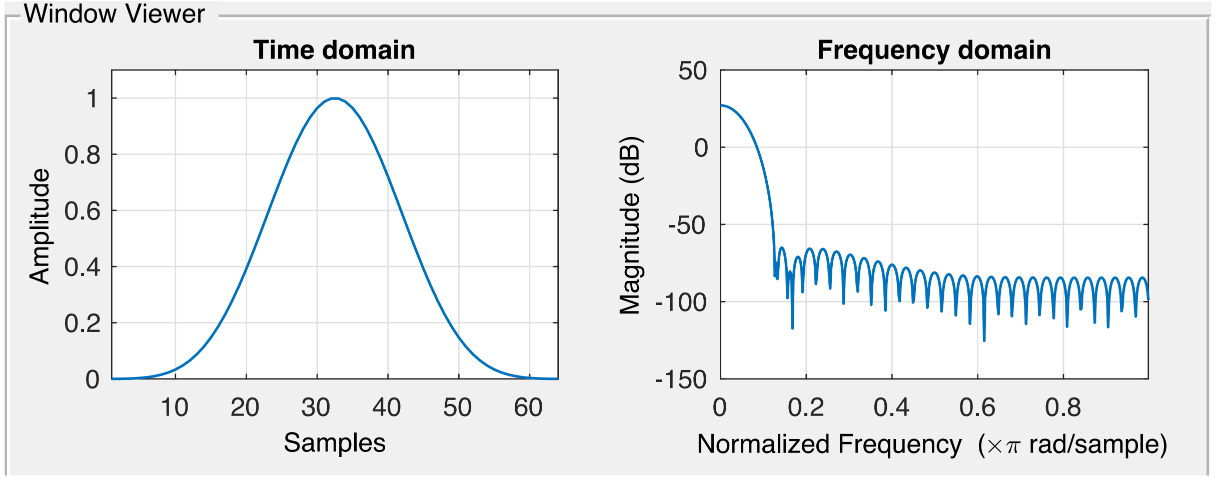
Short-Time Fourier Transform

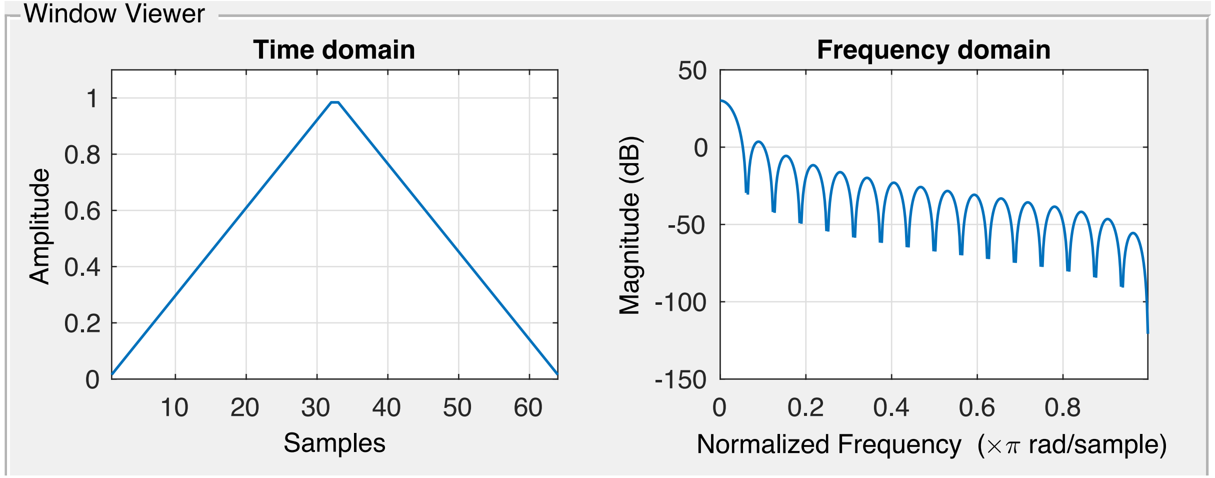
- Windowing and overlap-and-add

- The windows need to have some desired properties
  - Their sum should be 1 after the overlap-and-add

$$w_t[512+n] + w_{t+1}[n] = 1$$
$$\forall n = \{0, 1, \cdots, 511\}$$

  - Windowing should not affect too much on the DFT

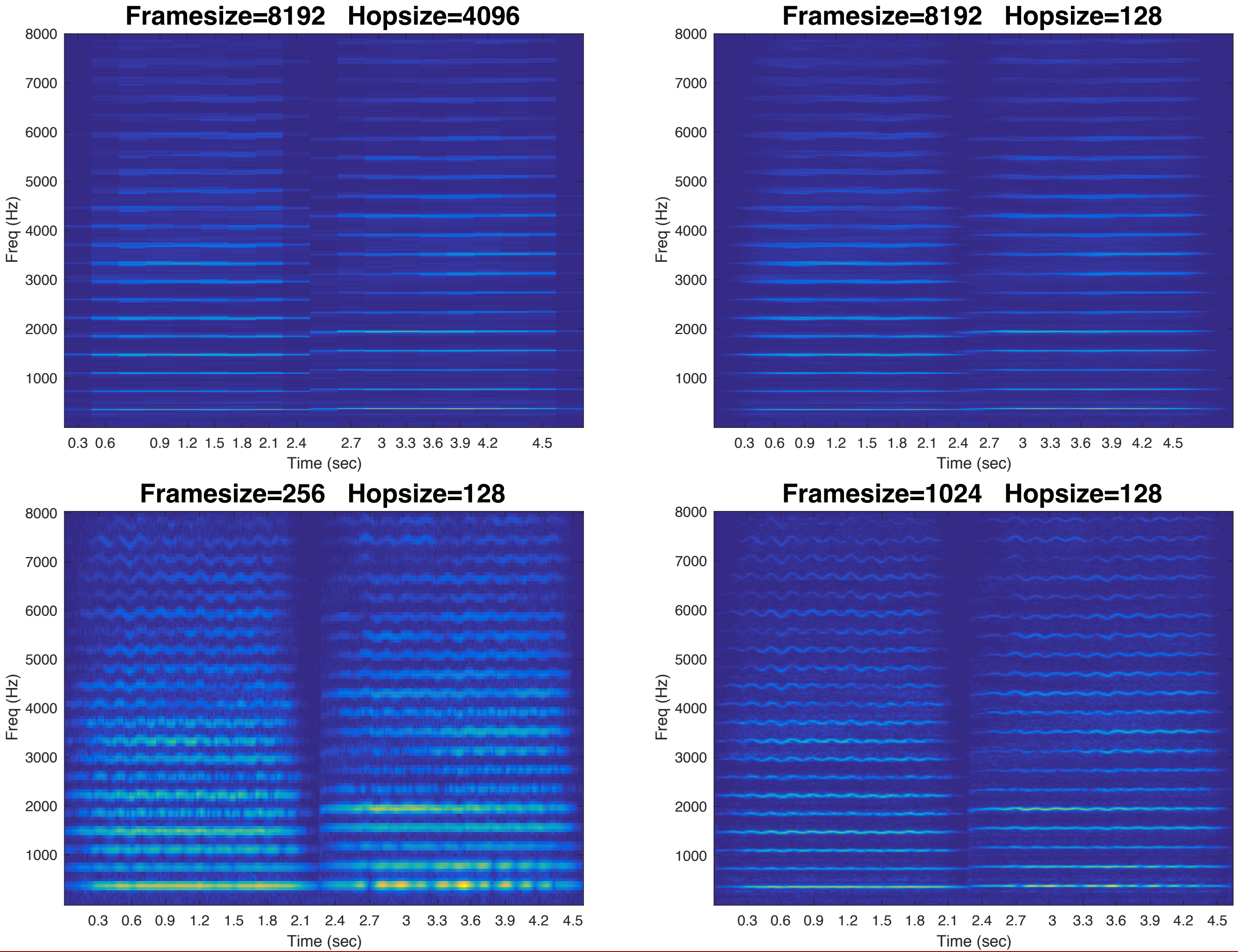




Short-Time Fourier Transform

- Resolution control

- Trade-off between
  time and frequency resolutions



- Which one do you like
  the best?

Discrete Cosine Transform

- An alternative

- What if you don't really care about the perfect reconstruction?
  - As in lossy compression techniques: JPEG and AAC

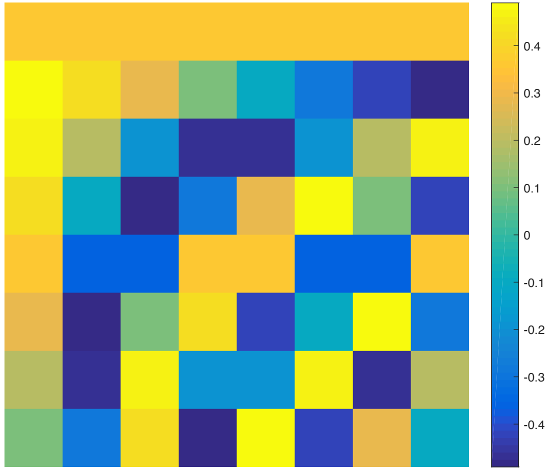
Samples

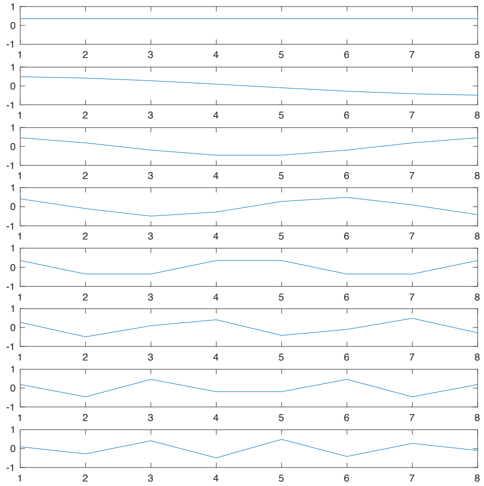
$$C(k, n) = \sqrt{\frac{1}{N}} \cos\left(\frac{\pi(2n+1)k}{2N}\right)$$

Basis vectors

n

k

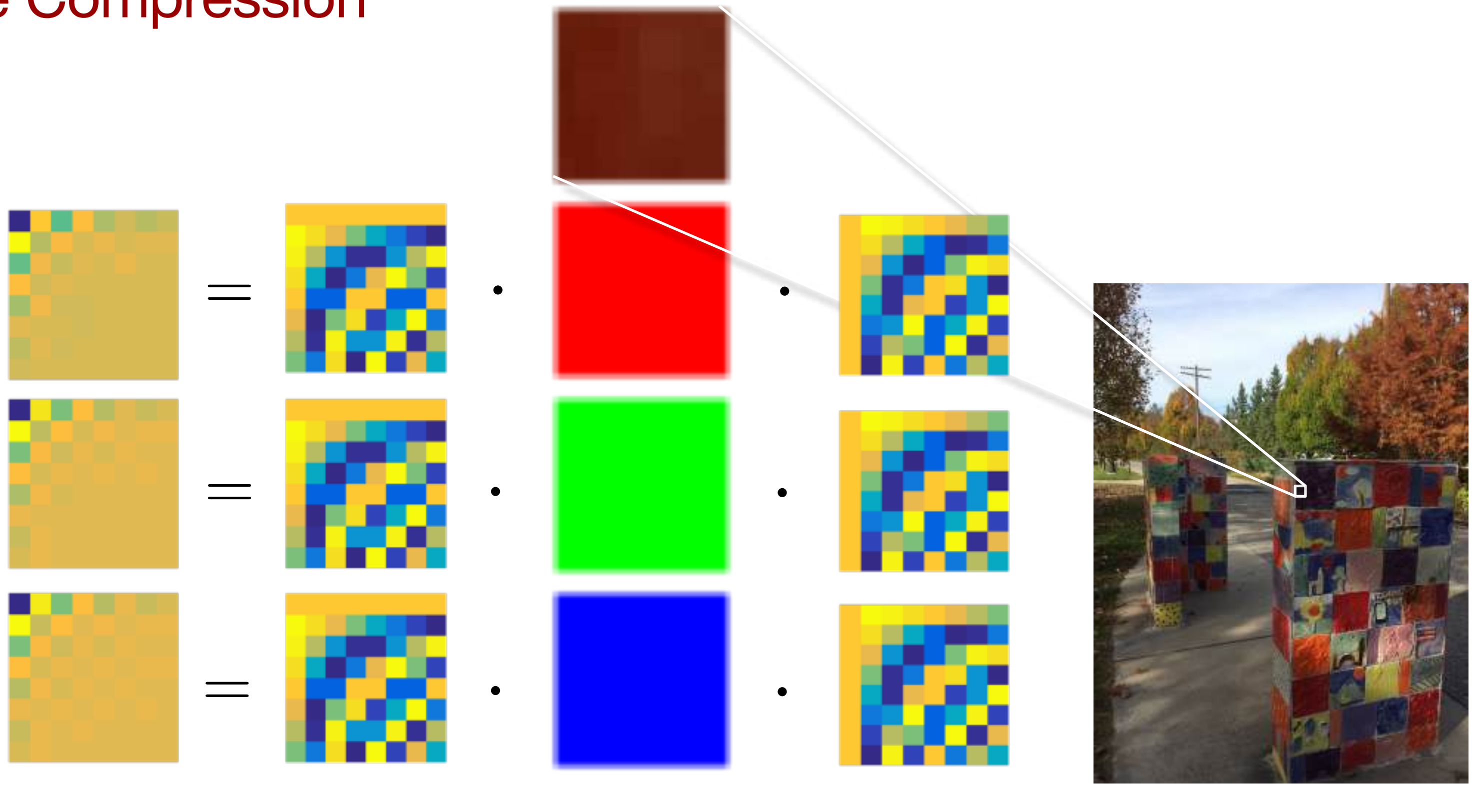




Discrete Cosine Transform

- JPEG Image Compression

- The 2D DCT

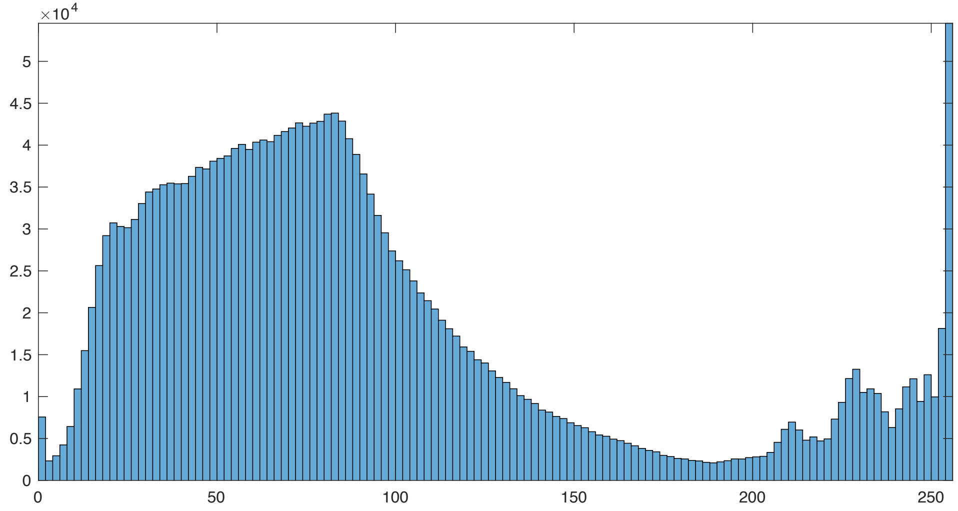


- Small patches (8X8) are with very similar values
  - A high sum
  - Low freq. components are stronger than high freq.

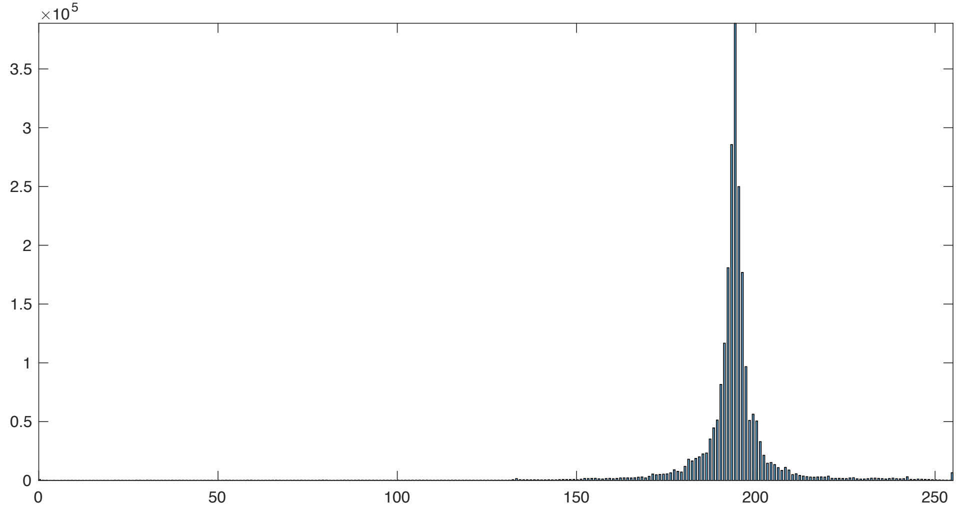
Discrete Cosine Transform

- JPEG Image Compression

- DCT coefficients are more centralized than the original pixels
  - Huffman coding can benefit



Sample distribution of the patch
Entropy: 6.53



Sample distribution of the DCT coefficients
Entropy: 4.73

  - (Remember that entropy can be seen as the average number of bits to encode the signal)

Open Questions

- Does this have anything to do with machine learning?

Reading

- Textbook Chapter 6.8 and 6.9

Thank You!