Convolution

Quora

Ask or Search Quora | Ask Question

Convolutional Neural Networks (CNNs) +3

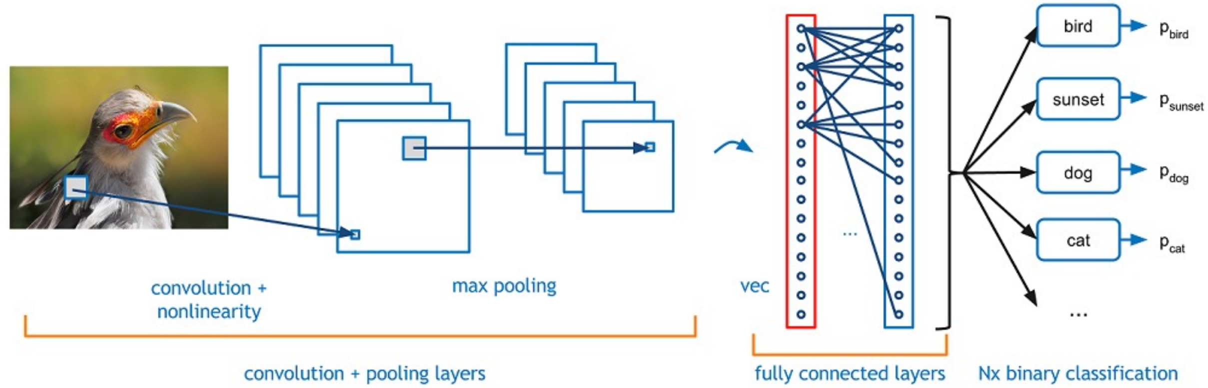
What is a convolutional neural network?

Answer | Request | Follow 76 | Comment | Share 2 | Downvote

UNDERSTANDING CONVOLUTIONAL NEURAL NETWORKS FOR NLP

When we hear about Convolutional Neural Network (CNNs), we typically think of Computer Vision. CNNs were responsible for major breakthroughs in Image Classification and are the core of most Computer Vision systems today, from Facebook's automated photo tagging to self-driving cars.

A Beginner's Guide To Understanding Convolutional Neural Networks



patterns; with more than one, you could look for patterns of patterns. Take the case of image recognition, which tends to rely on a contraption called a "convolutional neural net." (These were elaborated in a seminal 1998 paper whose lead author, a Frenchman named Yann LeCun, did his postdoctoral research in Toronto under Hinton and now directs a huge A.I. endeavor at

digital input and send output to other nodes. Layers upon layers of these nodes make up so-called convolutional neural networks, which, with sufficient training data, have become better and better at identifying images.

Inside a Convolutional Neural Network

But how does this filtering work? The secret is in the addition of two new types of layers: convolutional and pooling layers. We'll break the process down below, using the example of a network designed to do just one thing: determine whether a picture contains a grandma or not.

Next, Zweig and co optimized their own deep-learning systems based on convolutional neural networks with varying number of layers, each of which processes a different aspect of speech. They then used the

Got a tip? Let us know.

News Video Events Crunchbase

DISRUPT NY Find out how startups are transforming real estate investing at Disrupt NY Get

convolutional neural networks

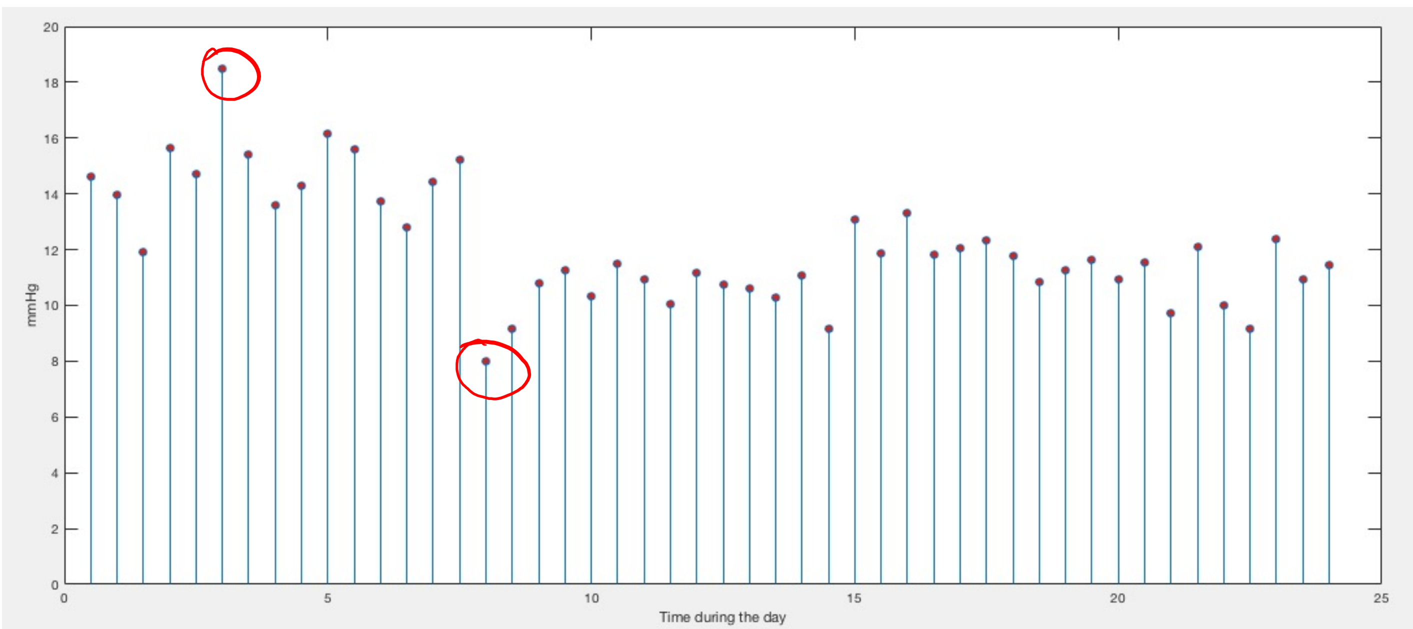
Convolution

$$(f \star g)(t) = \int_{-\infty}^{\infty} f(t')g(t - t')dt'$$

Is Convolution that convoluted?

Let us start with discrete 1D signals !

For example, my systolic blood pressure measured with a wearable over 24 hours on 1/16/2016:



$$s[n]$$

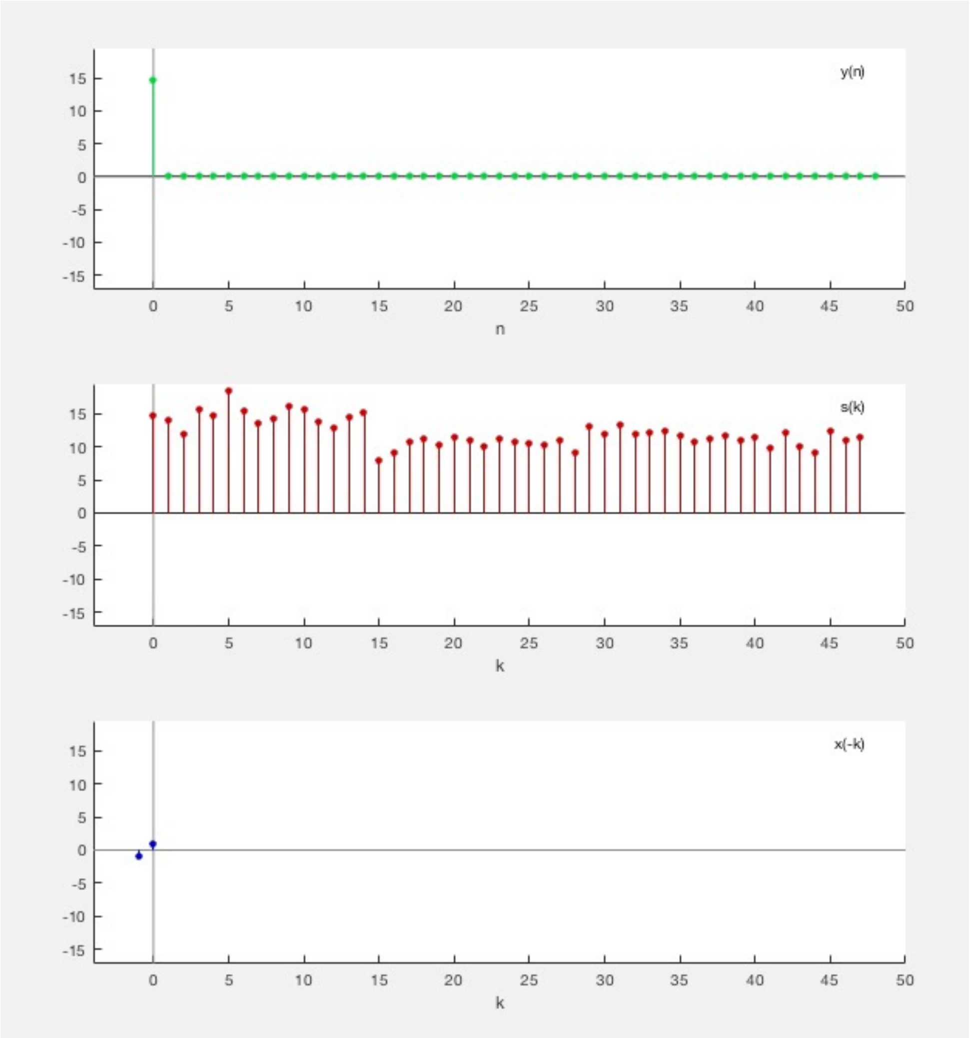
Interested in the largest drop?

15	14	12	16	15	18	15	14	14	16	16	14	13	14	15	8	9	11	11	10	11
−1	1																			
	−1	1																		
		−1	1																	
			−1	1																

We are moving a sliding window and at every position we take the scalar product between the mask and the signal.

−1	−2	4	−1	4	−4	−1	0	2	0	−2	−1	1	1	−7	1	2	0	−1

Let us visualize it while in action:



We will call this operation correlation between two discrete signals and we can write it as

$$y[n] = \sum_{k=1}^{N} s[k]\, h[k-n]$$

Convolution differs from correlation by a reflection:

We take the "mask" d[k] $-1 \quad 1$

and reflect it to d[-k] $1 \quad -1$

and then we shift to
d[-(k-n)]=d[n-k]

$$y[n] = \sum_{k=1}^{N} s[k]\, h[n-k]$$

					1	−1														
15	14	12	16	15	18	15	14	14	16	16	14	13	14	15	8	9	11	11	10	11

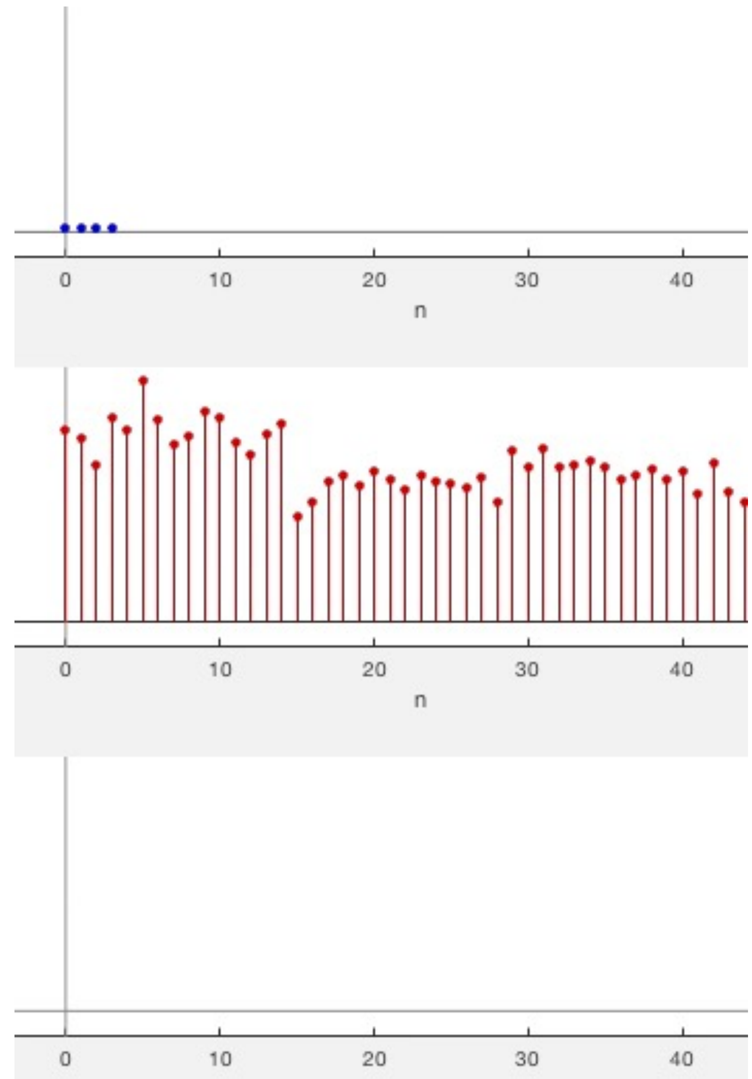
The informal term for the convolution mask is “filter mask”, its values are called “filter weights” or convolution weights.

Let us do another operation on the systolic pressure signal.

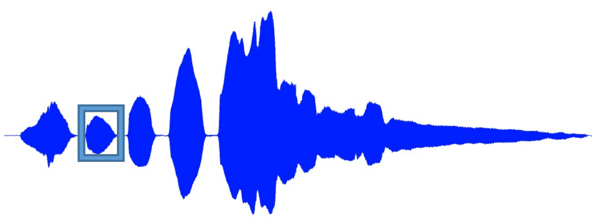
Taking the local average over a window of two hours.

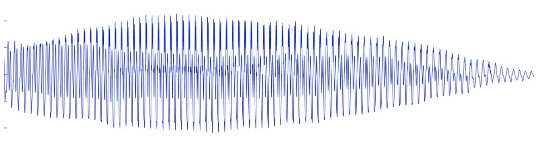
The averaging mask would look like [¼ ¼ ¼ ¼]

Does the convolution differ from correlation?

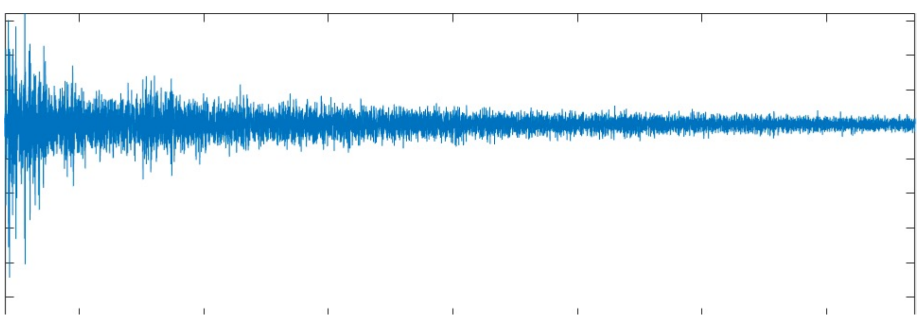


Let us move to the Academy of Music!
Convolution is implemented by the environment.



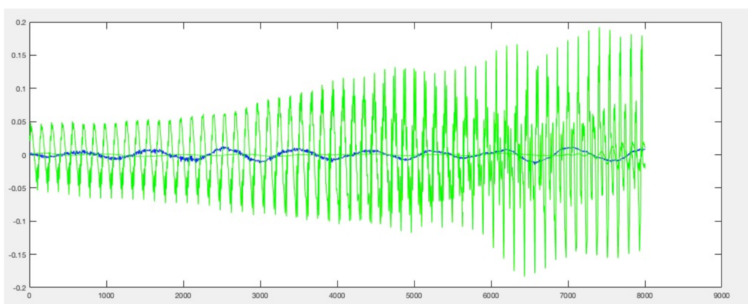


Sound is reflected creating multiple echos



Now let us take the song s[n] and convolve it with the pistol h[n]

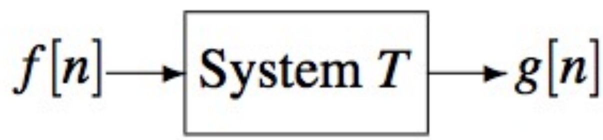
$$y[n] = \sum_{k=1}^{N} s[k]\, h[n-k]$$

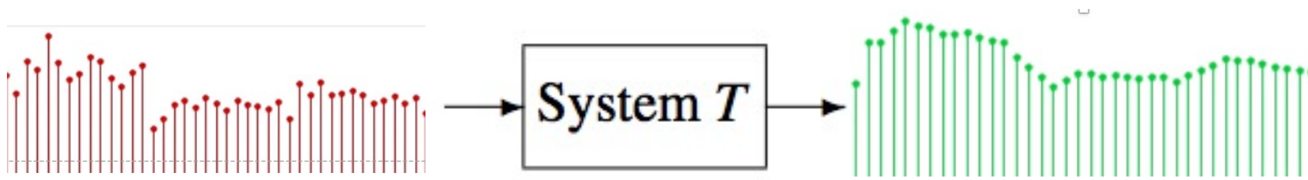


The systems' view

Consider it as a system with input and output. In the previous examples, input might have been

- Audio like a song
- Blood pressure measurements
- In general any 1D signal





Linear System

- We will say that a system is linear if

$$f_1(t) \longrightarrow \boxed{T} \longrightarrow g_1(t)$$

$$f_2(t) \longrightarrow \boxed{T} \longrightarrow g_2(t)$$

then $$af_1(t) + bf_2(t) \longrightarrow \boxed{T} \longrightarrow ag_1(t) + bg_2(t)$$

Example

$T\{f\}(t) = f(t) - f(t-1)$ is linear:

$$
\begin{aligned}
g_1(t) &= f_1(t) - f_1(t-1) \\
g_2(t) &= f_2(t) - f_2(t-1) \\
T\{af_1 + bf_2\}(t) &= [af_1(t) + bf_2(t)] - [af_1(t-1) - bf_2(t-1)] \\
&= a(f_1(t) - f_1(t-1)) + b(f_2(t) - f_2(t-1)) \\
&= aT\{f_1\}(t) + bT\{f_2\}(t)
\end{aligned}
$$

What is non-linear?

$$g(t) = \max(f(t), f(t-1))$$

Shift-invariant (or time-invariant) system

$$f(t) \longrightarrow \boxed{T} \longrightarrow g(t) \quad \text{then} \quad f(t-t_0) \longrightarrow \boxed{T} \longrightarrow g(t-t_0)$$

Examples:

- $T\{f\}(t) = f(t) - f(t-1)$ is shift-invariant.

$$\begin{aligned} T\{f(t-t_0)\} &= f(t-t_0) - f(t-t_0-1) \\ &= g(t-t_0) \end{aligned}$$

- $g(t) = tf(t)$ is not shift-invariant.

Is there a formula for describing linear shift-invariant systems?

$$f(t) \longrightarrow \boxed{\text{System } T} \longrightarrow g(t)$$

Yes, the convolution! Discrete or continuous

$$g(t) = \int_{-\infty}^{\infty} f(t')h(t-t')dt' \qquad g[n] = \sum_{k=-\infty}^{\infty} f[k]\, h[n-k]$$

What is a filter?

Linear shift-invariant systems can be written as convolutions!

$$g(t) = \int_{-\infty}^{+\infty} f(t')h(t-t')dt' = f(t) \star h(t)$$

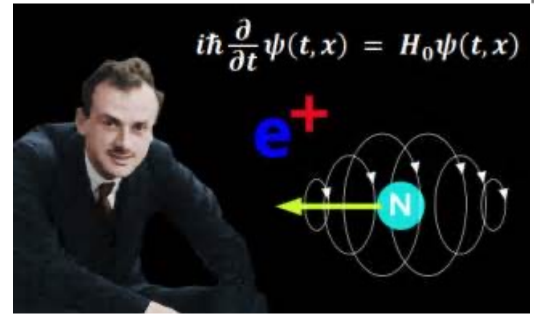
But what exactly is this h(t) or h[n] that we call filter?

Box function

$$\mathrm{rect}(t) = \begin{cases} 1\,, & |t| \le 1/2 \\ 0\,, & otherwise \end{cases}$$

$$\mathrm{rect}(t/a) = \begin{cases} 1/a\,, & |t| \le a/2 \\ 0\,, & otherwise \end{cases}$$

Dirac function



$$\delta(t) \doteq \lim_{a \to 0} \frac{1}{a} \operatorname{rect}(t/a)$$

It follows from definition of the Dirac function that

$$\delta(t) = 0 \quad \text{for all } \ t \neq 0$$

and

$$\int_{-\infty}^{+\infty} \delta(t)dt = 1$$

Absorption property

$$\int_{-\infty}^{+\infty} \delta(t) f(t) dt = f(0)$$

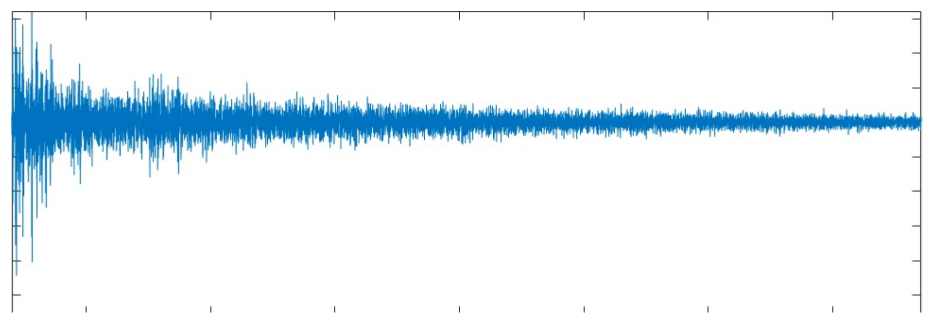
or in a more general form:

$$\int_{-\infty}^{+\infty} \delta(t) f(t_0) dt = f(t_0)$$

What happens when the input to an LSI is a Dirac?

$$\int_{-\infty}^{+\infty} \delta(t')h(t-t')dt' = h(t)$$

$$\delta(t) \longrightarrow \boxed{\text{System } T} \longrightarrow h(t)$$



A filter *h(t) or h[n]* is the response of the system to the Dirac impulse.

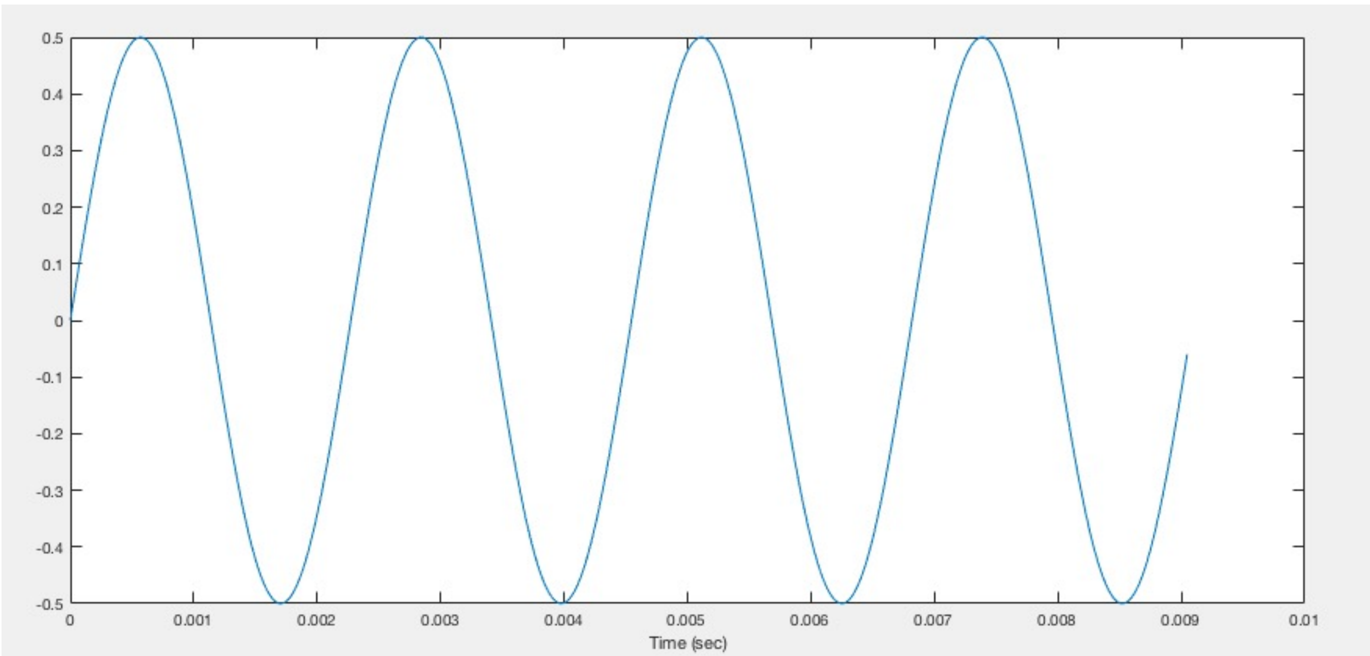
Example. Let an LSI system with impulse response defined as

$$h(t) = (1/2)(\delta(t+1) - \delta(t-1))$$

$$\begin{aligned}
g(t) &= \int_{-\infty}^{+\infty} f(t')h(t-t')dt' \\
&= \int_{-\infty}^{+\infty} f(t-u)h(u)du \\
&= \int_{-\infty}^{+\infty} f(t-u)(1/2)(\delta(u+1) - \delta(u-1))du \\
&= (1/2)\int_{-\infty}^{+\infty} f(t-u)\delta(u+1)du - (1/2)\int_{-\infty}^{+\infty} f(t-u)\delta(u-1)du \\
&= (1/2)(f(t+1) - f(t-1))
\end{aligned}$$

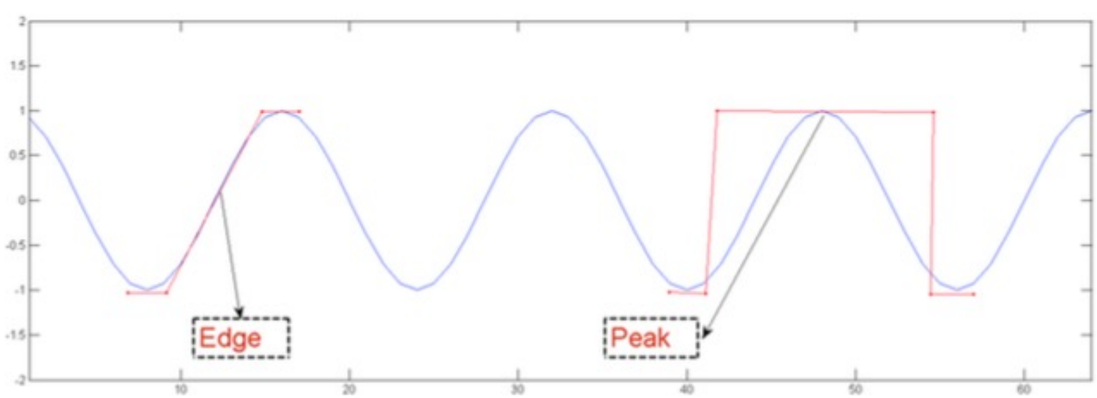
The Fourier Transform

Back to music: sine wave 440Hz

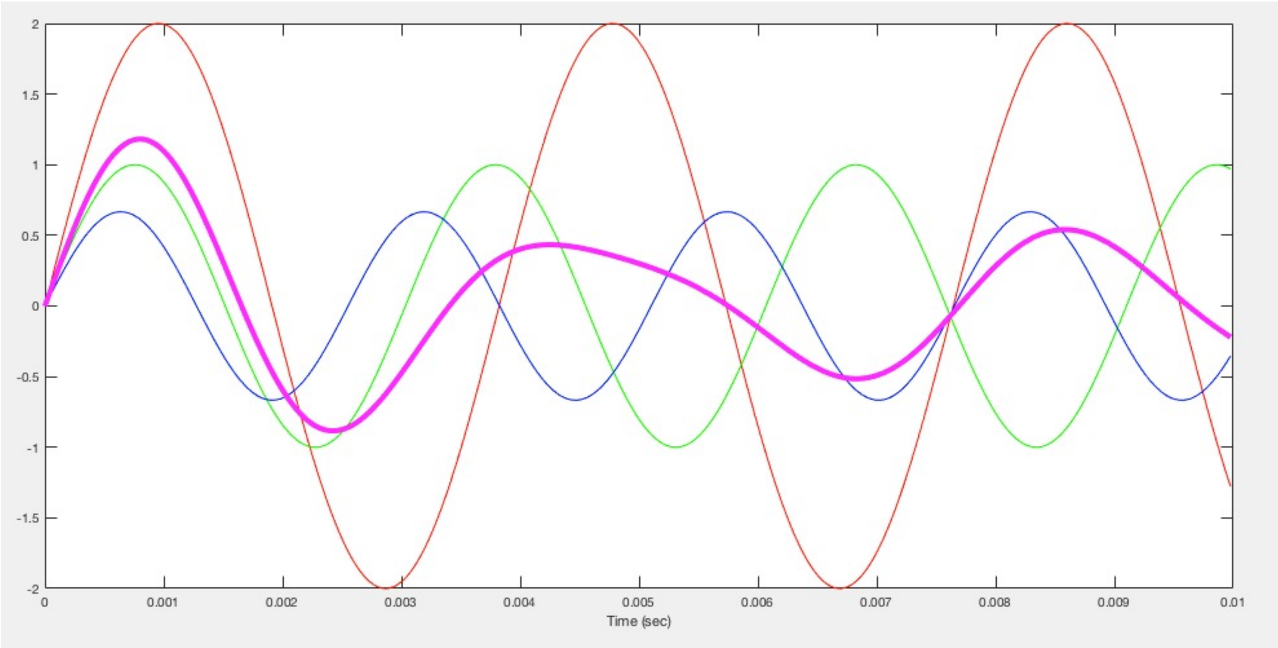


Or images….

2D Signal $f(x, y) = \cos\left(\frac{2\pi x}{16}\right)$

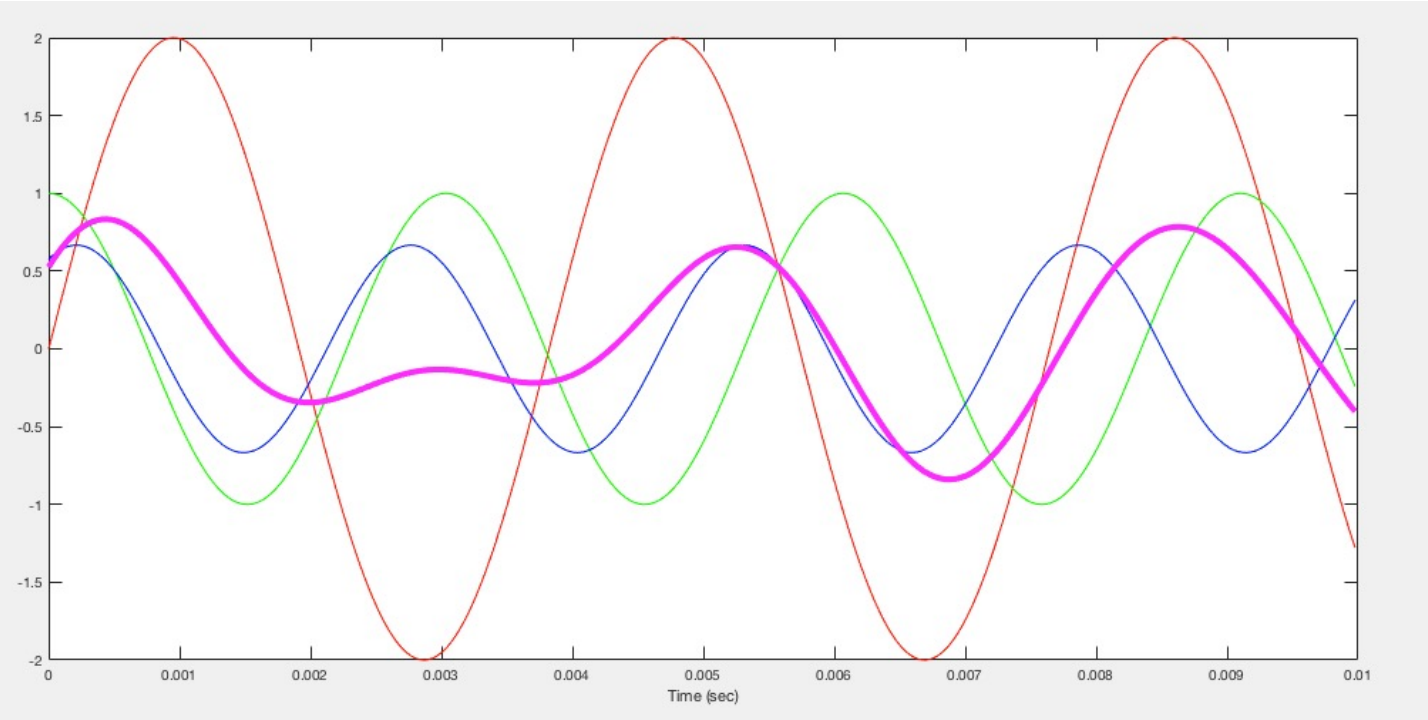


Play three notes together: C-major chord consists of a superposition of C (262Hz), E(330HZ) and G(392Hz)



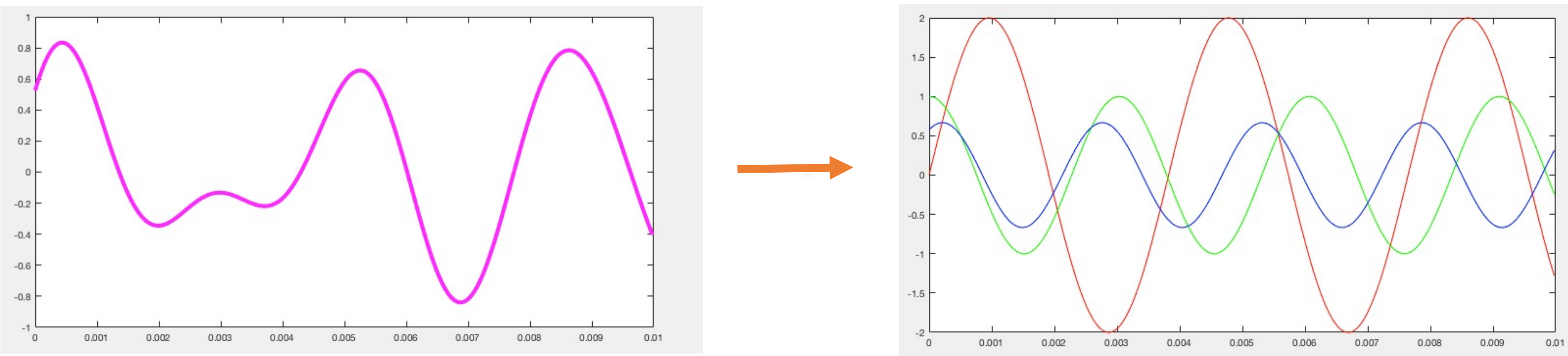
$$A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t) + A_3 \sin(2\pi f_3 t)$$

We could also change the phases:

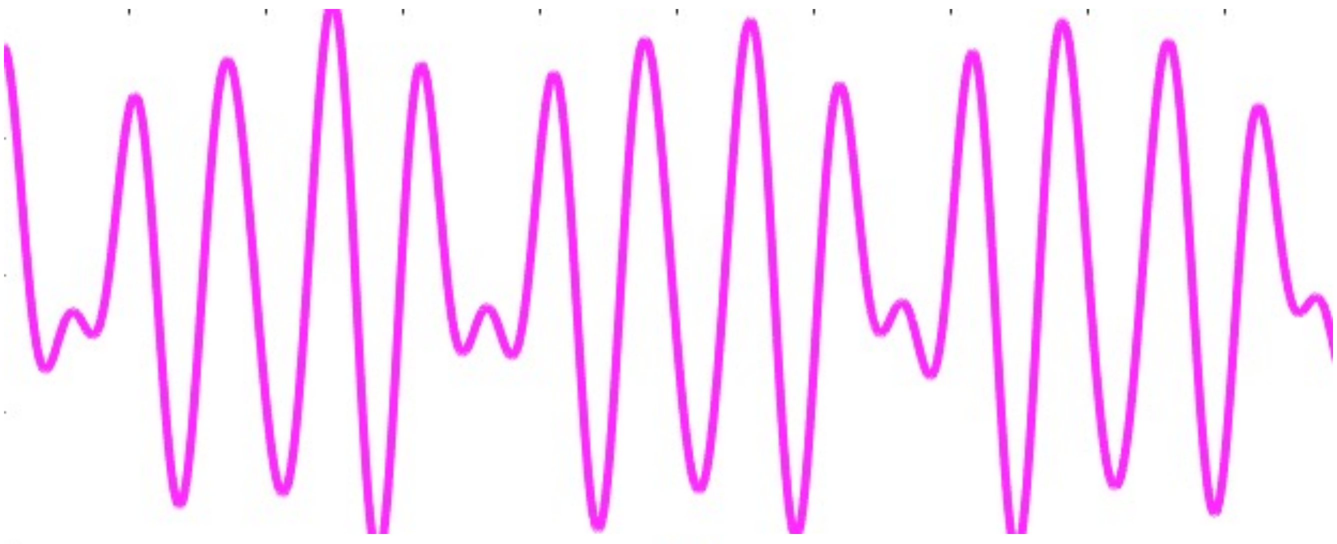


$$A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t + \frac{\pi}{2}) + A_3 \sin(2\pi f_3 t + \frac{\pi}{3})$$

Can we solve the inverse problem?



Yes, we can decompose a function into "sinusoids" if it is periodic: Fourier series



$$f(t) = \sum_{n=-\infty}^{\infty} F[n] e^{j2\pi \frac{n}{T} t}$$

The complex Fourier coefficients F[n] can be found with the Fourier Transform

$$F[n] = \int_{-T/2}^{T/2} f(t) e^{-j2\pi \frac{n}{T} t} dt$$

$j^2=-1$

Quick reminder on complex numbers:

- $a + jb \in \mathbb{C}$, $j^2 = -1$
- $e^{j\omega t} = \cos(\omega t) + j\sin(\omega t)$.
- $a + jb = re^{j\theta}$ with $r = \sqrt{a^2 + b^2}$, $\theta = atan2(b, a)$

Why we need complex numbers?

To account for the phase

$$f(t) = \sum_{n=-\infty}^{\infty} F[n] e^{j2\pi \frac{n}{T} t}$$

or

$$f(t) = \sum_{n=-\infty}^{\infty} A_k(\cos(2\pi \frac{n}{T} t + \phi_k)$$

Let us agree on a notation:

$$F[n] = \int_{-T/2}^{T/2} f(t) e^{-j2\pi \frac{n}{T} t} dt$$

Definition of the Fourier Transform:

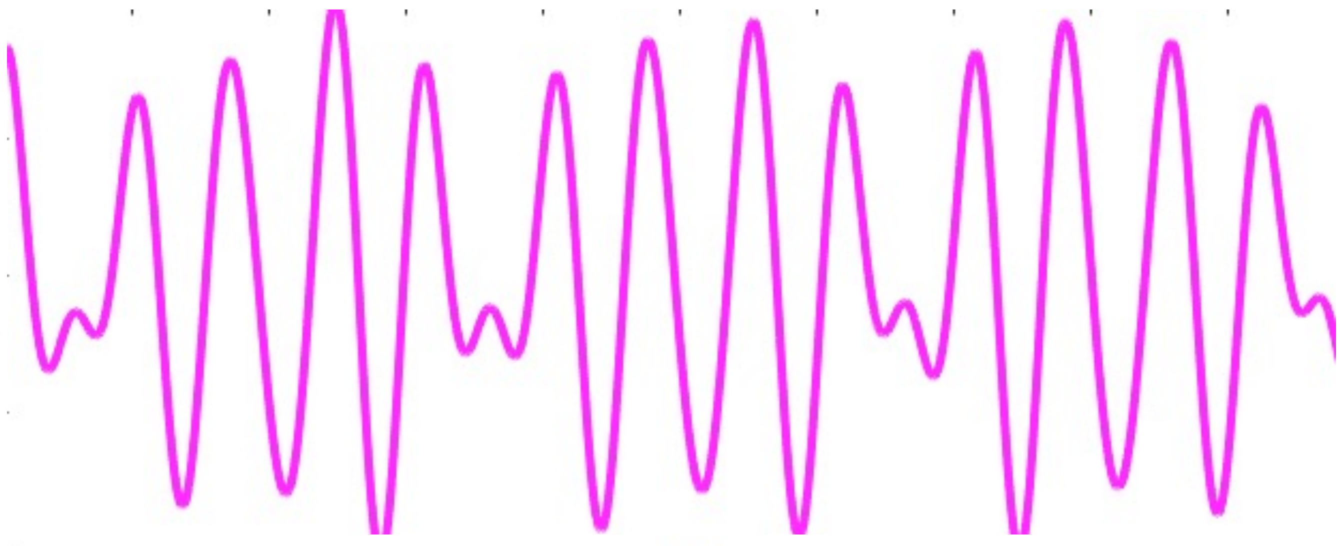
$$f(t) \circ\!\!-\!\!\bullet F(\omega) = \mathcal{F}\{f(t)\}$$

$$F : \mathbb{R} \to \mathbb{C}$$

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt,$$

where ω denotes the frequency. This definition is sometimes called non-unitary Fourier transform, with angular frequency (ω is referred to as angular frequency, and $s = \frac{n}{T}$ is the ordinary frequency in Hz such that $\omega = 2\pi s$).

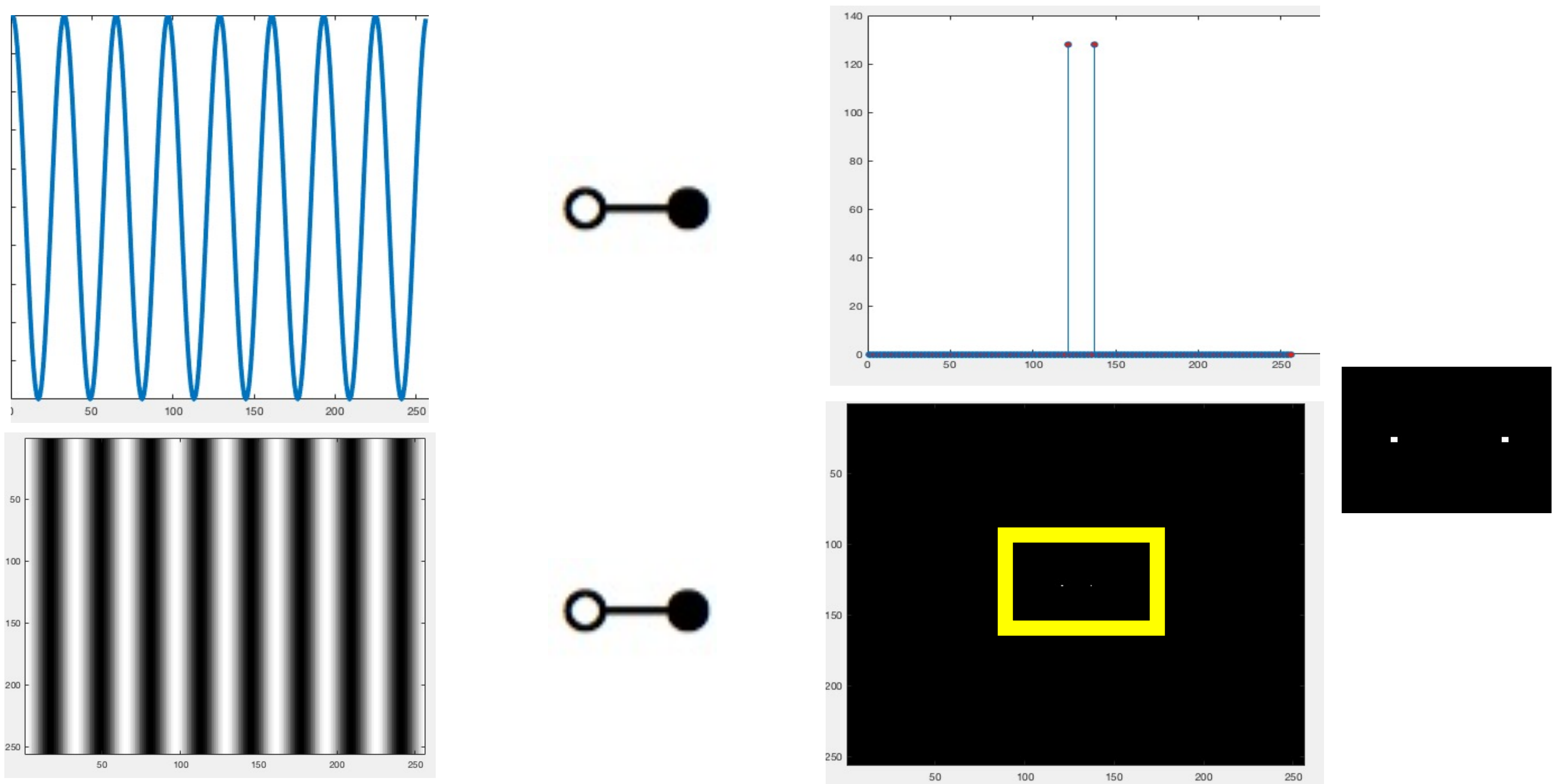
We learnt about the Fourier series of periodic signals



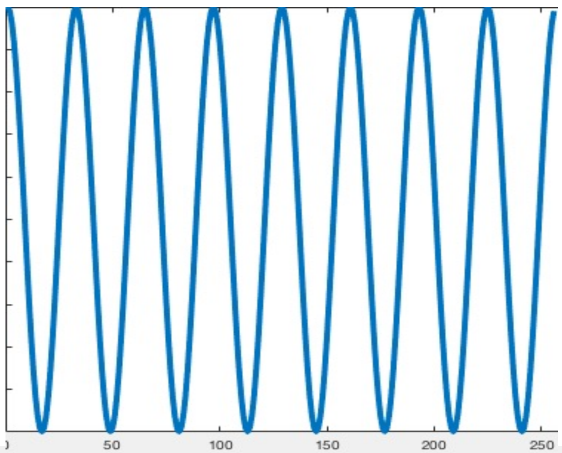
$$f(t) = \sum_{n=-\infty}^{\infty} F[n] e^{j2\pi \frac{n}{T} t}$$

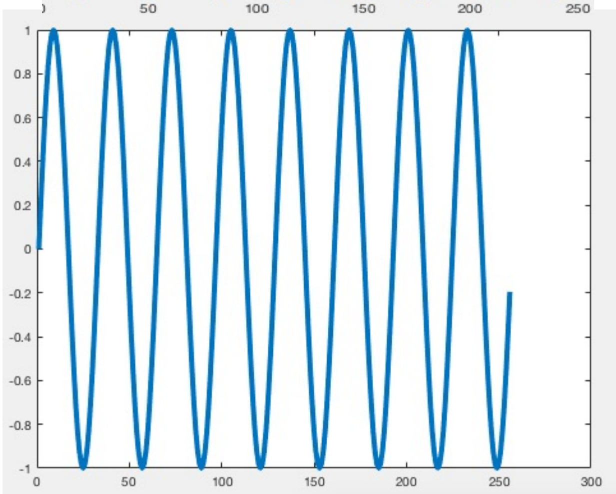
Some facts about the Fourier Transform

Fourier Transform of a Cosine

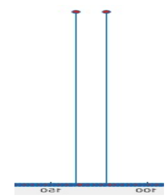


Do sine and cosine have the same Fourier transform





Only in the magnitude |F[n]|!



They differ in the phase !

$$\cos(\omega_0 t) \quad \circ\!\!-\!\!\bullet \quad 2\pi.\tfrac{1}{2}(\delta(\omega-\omega_0)+\delta(\omega+\omega_0))$$

$$\sin(\omega_0 t) \quad \circ\!\!-\!\!\bullet \quad 2\pi.\tfrac{1}{2j}(\delta(\omega-\omega_0)-\delta(\omega+\omega_0))$$

Fourier of the harmonic exponential

$$e^{j\omega_0 t} \circ\!\!-\!\!\bullet\ 2\pi\delta(\omega - \omega_0)$$

Two phase-related facts

Shift theorem

$$f(t-t_0) \circ\!\!-\!\!\bullet\ F(\omega)e^{-j\omega t_0}$$

Example: for $f(t)$ even, $f(t-\frac{T}{2}) \circ\!\!-\!\!\bullet\ F(\omega)e^{-j\omega\frac{T}{2}}$, where $F(\omega)$ is real.

Modulation theorem

$$f(t)e^{j\omega_0 t} \circ\!\!-\!\!\bullet\ F(\omega-\omega_0)$$

Multiplying by a complex exponential causes a shift in the frequency domain.

Let us look at some more Fourier transforms: Dirac

Recall the absorption property

$$\int_{-\infty}^{\infty} \delta(t-t_0) f(t)dt = f(t_0)$$

Then we have:

$$\int_{-\infty}^{\infty} \delta(t) e^{-j\omega t} dt = 1$$

We will remember the two following results:

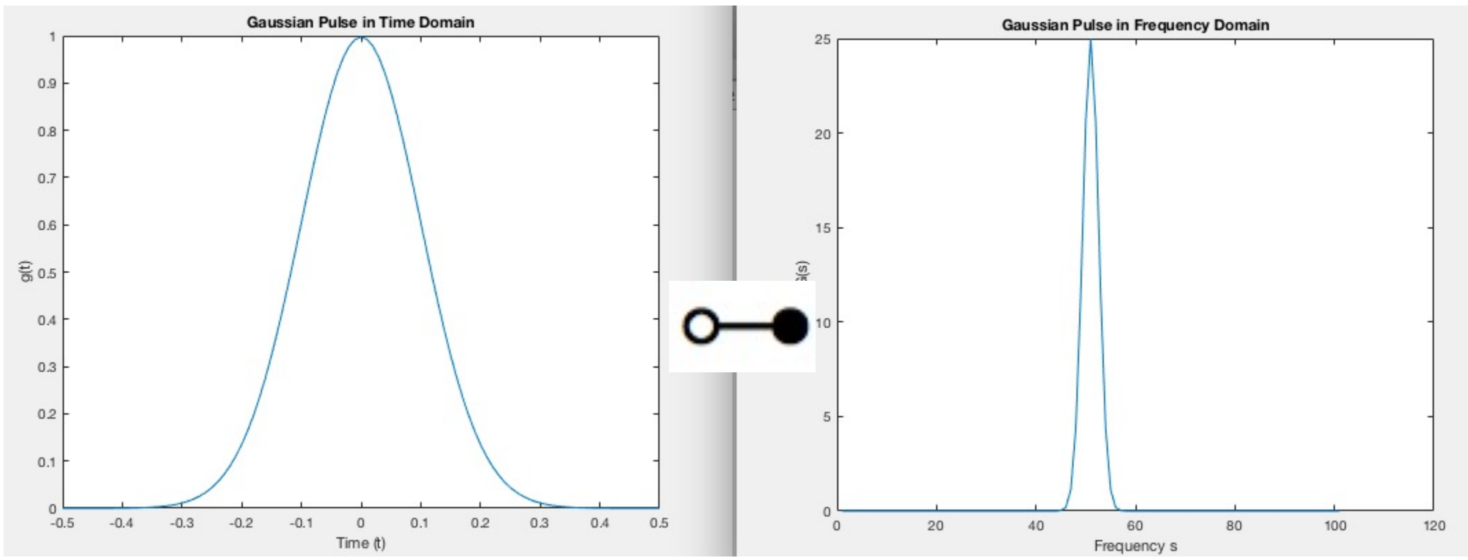
$$\begin{array}{rcl} \delta(t) & \circ\!\!-\!\!\bullet & 1 \\ 1 & \circ\!\!-\!\!\bullet & 2\pi\delta(t) \quad \text{(DC-component)} \end{array}$$

Fourier of the comb function $\text{Ш}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)$:

$$\sum_{n=-\infty}^{\infty} \delta(t - nT) \circ\!\!-\!\!\bullet \frac{1}{|T|} \sum_{n=-\infty}^{\infty} \delta(s - \frac{n}{T})$$

(or $\frac{2\pi}{|T|} \sum_{n=-\infty}^{\infty} \delta(\omega - \frac{2\pi n}{T})$ when using the non-unitary Fourier Transform with angular frequency).

Fourier Transform of the Gaussian Function



$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad \circ\!\!-\!\!\bullet \quad F(s) = \frac{1}{\sigma\sqrt{2\pi}} e^{-2\pi^2\sigma^2 s^2}$$

Convolution Theorem

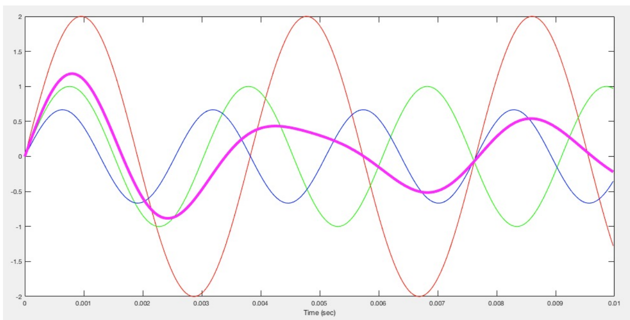
Convolution in time or space means multiplication in frequency

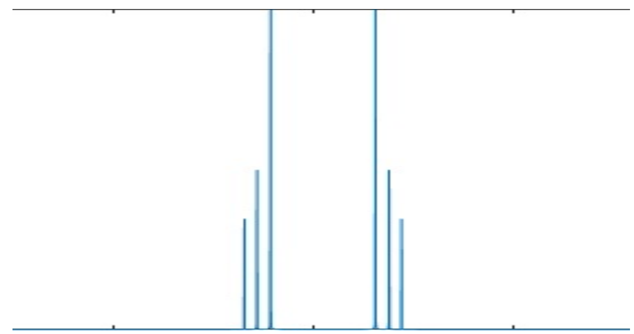
$$\begin{array}{ccccc}
f(t) & \rightarrow & \boxed{h(t)} & \rightarrow & g(t) = \int_{-\infty}^{\infty} g(t')h(t-t')dt' = f(t) * h(t) \\
\circ\!\!-\!\!\bullet & & \circ\!\!-\!\!\bullet & & \circ\!\!-\!\!\bullet \\
F(\omega) & & H(\omega) & & G(\omega)?
\end{array}$$

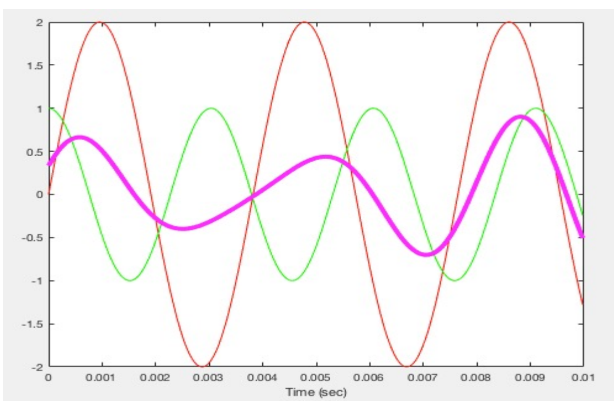
$$f(t) * h(t) \circ\!\!-\!\!\bullet F(\omega)H(\omega)$$

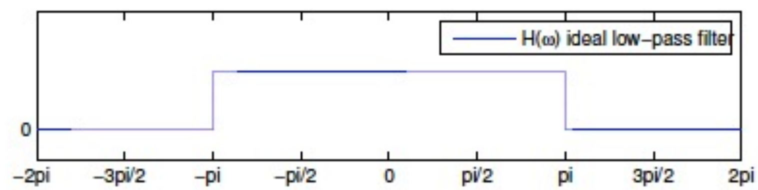
Convolution Theorem: Low Pass Filtering

$$f(t) * h(t) \circ\!\!-\!\!\bullet\ F(\omega)H(\omega)$$









Convolution Theorem II

Convolution in time or space means multiplication in frequency

$$f(t)h(t) \circ\!\!-\!\!\bullet\; F(\omega) * H(\omega)$$

Discrete signals and Discrete Fourier Transform

Discrete time signals have still continuous Fourier transforms (CFT)!
We call it also DTFT (Discrete Time Fourier Transform).

$$f[n] \circ\!\!-\!\!\bullet \sum_{n=0}^{L-1} f[n]e^{-j\omega n}$$

Discrete time signals have continuous and periodic Fourier transforms with period 2π!
That's why we consider only the interval $(-\pi,\pi]$ when we talk about the CFT of a discrete signal!

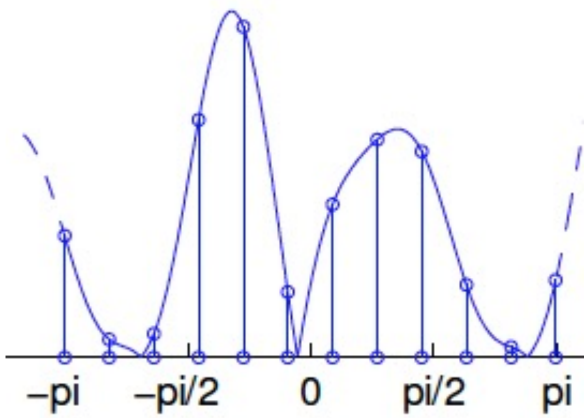
-

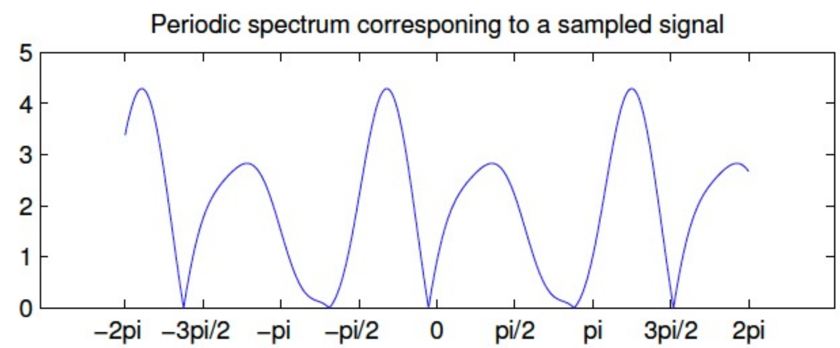
Example. Let $f[n]$ a discrete time signal and $h[n]$ a discrete time signal defined as:

$$h[n] = \begin{cases} 1, & n = 0 \\ -1, & n = 1 \\ 0, & \text{elsewhere} \end{cases}$$

$$H(\omega) = \sum_n h[n]e^{-j\omega n} = h[0]e^{-j\omega \cdot 0} + h[1]e^{-j\omega \cdot 1} = 1 - e^{-j\omega}$$

What does discrete frequency mean?

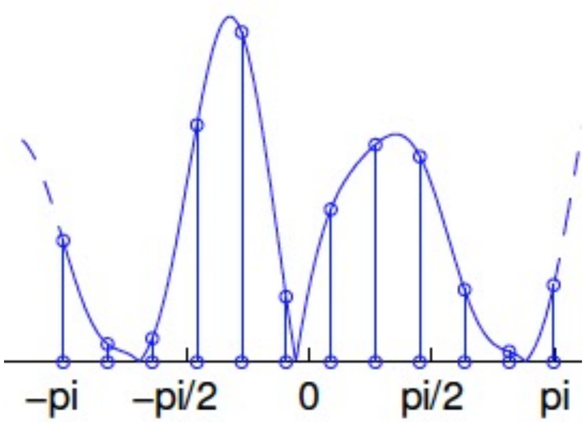


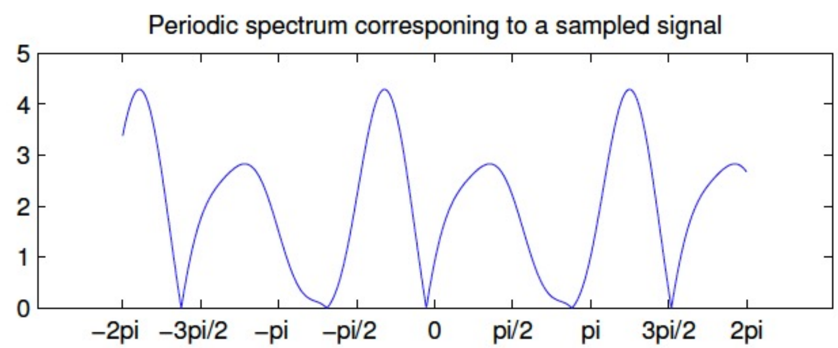


What does sampling in the frequency domain correspond to in the time/space domain? It corresponds to a **convolution**:

- FREQUENCY DOMAIN: multiplication with $\sum \delta(\omega - \frac{2\pi k}{L})$

- TIME DOMAIN: convolution with with $\sum \delta[n - kL]$, equivalent to **replication of the signal at multiples of its length**.





How is the DFT of a discrete time finite signal different than the DTFT (CFT)?

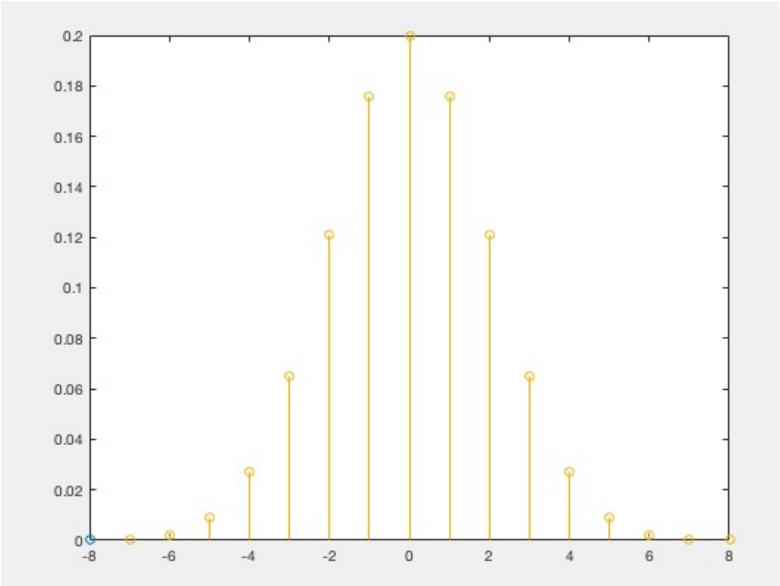
Sampling of the frequency domain only means replicating the original signal.

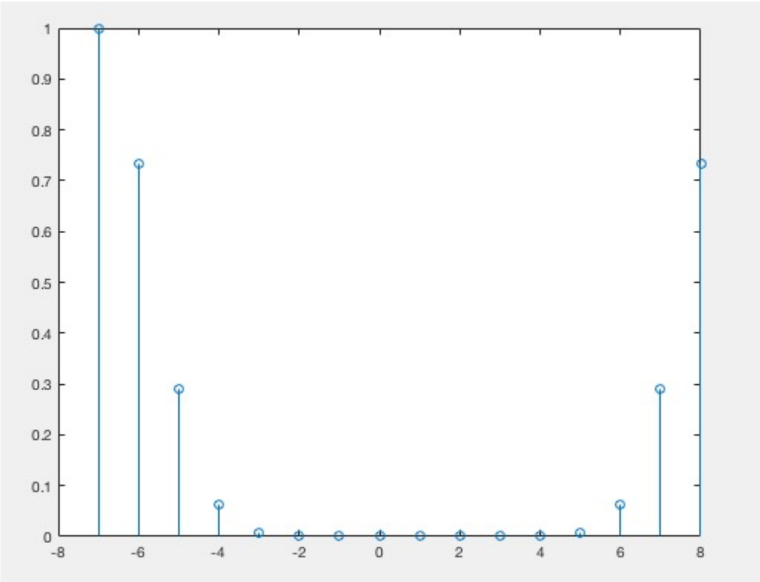
Some MATLAB….

```
support = 16;
sigma = 2;
d = -floor(support/2)+1:floor(support/2);
g = (1/(sigma*sqrt(2*pi)))*exp(-d.^2/(2*sigma^2));

figure(1)
stem(d,g);

figure;
stem(d,abs(fft(g)));
```





fftshift… performs a shift by L/2 in the frequency domain which is $e^{-j\pi Ln/(2L)}$ in time domain corresponding to multiplying signal with $(-1)^n$ which means –h[1], h[2], -h[3], …