FUNDAMENTALS OF IMAGE AND VIDEO PROCESSING

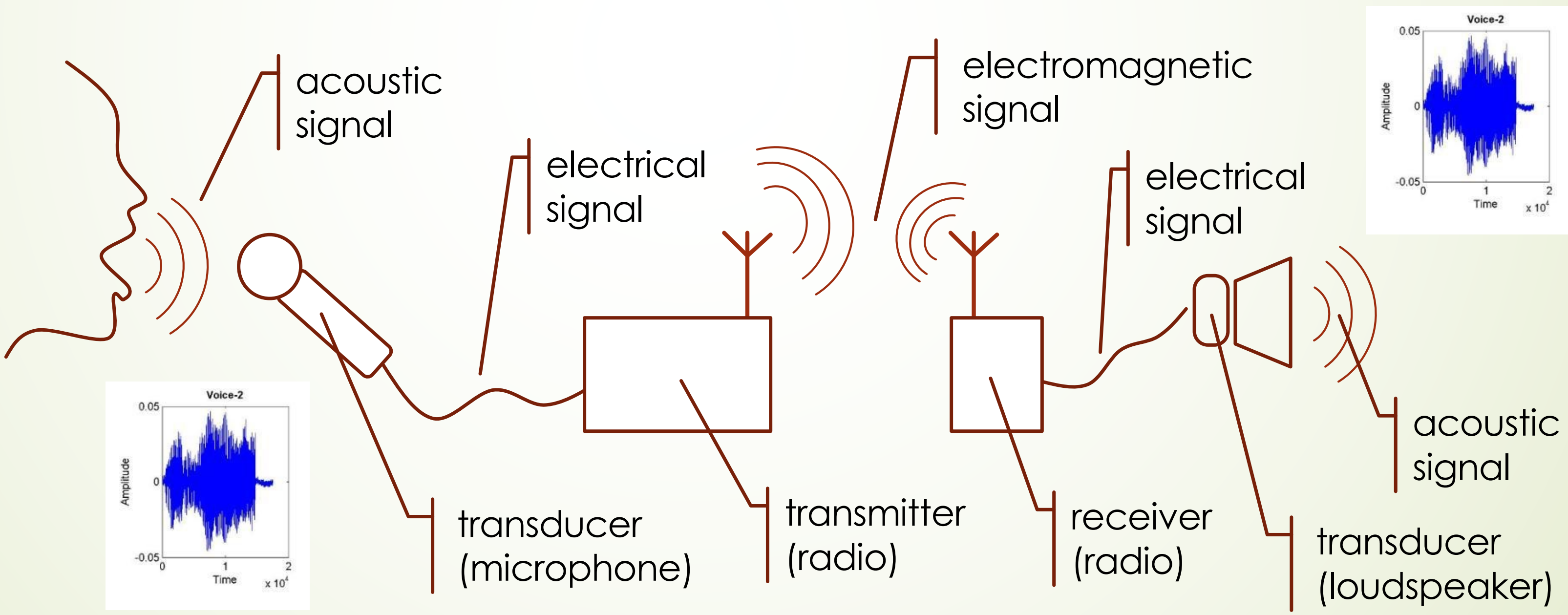
Part 1: Signals and Systems

What we'll see in this section

- What is a signal
- What is a system
  - Special case of LTI systems
- How signals and systems interact
  - Linear case and convolution
  - Non-linear case and input/output relationship
- Signals in the frequency domain and Fourier Transform
- Filters in the frequency domain

Signals

- A signal is a function that conveys information about a given phenomenon
  - It can be seen as the variation in a given domain (time, space) of some physical quantity (current, light, pressure, voltage, ...)
  - The same information can be associated to different signals that represent it

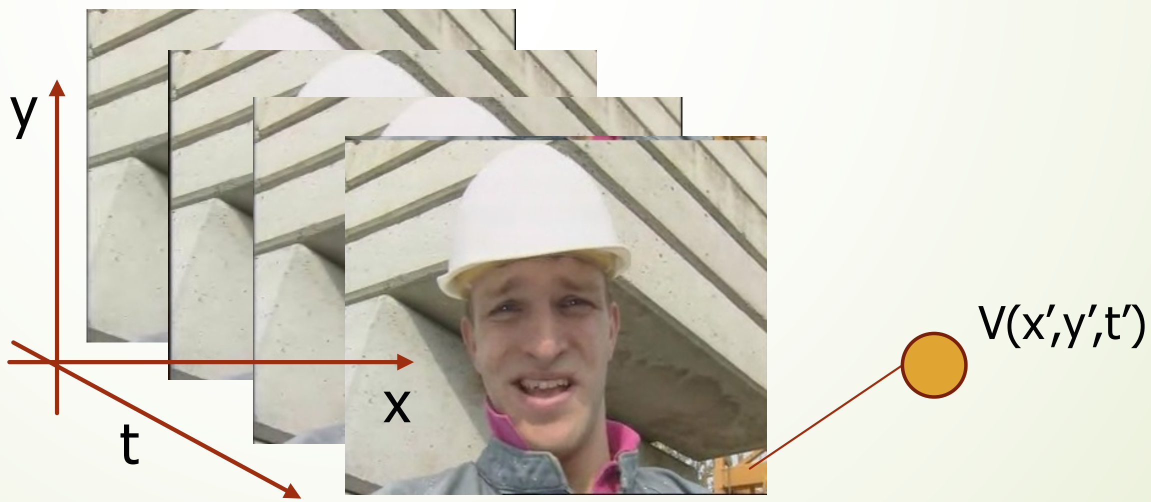


Systems

- A system is a generic equipment, hw or sw, that modifies the signal in some way
  - In the previous example, the microphone, the transmitter, the receiver and the loudspeaker are all systems, but the cable that connects the mike to the transmitter is a system as well
  - More in general, the signal can be smoothed, attenuated, amplified, cut, stored, etc.: each operation requires a system that performs it

Signals vs. math

- Signals can be represented as functions of one or more variables
  - An audio signal is a function of time x(t), describing the intensity of a sound at a given instant
  - An image is a function of space I(x,y), describing the luminance and/or color of a given point in space
  - A video is a function of space and time V(x,y,t), describing the luminance and/or color of a given point in space at a given instant



Systems vs math

- A system can be represented as a composite function that maps a signal into another

$$x(\cdot) \Rightarrow \boxed{\text{SYSTEM}} \Rightarrow y(\cdot) = \mathfrak{I}(x(\cdot))$$

- It would be useful to express $\mathfrak{I}(\cdot)$ in analytic form, in order to be able to predict the response of a system to any given input
  - This turns out to be difficult or even impossible in general, except for a special class of systems, called **linear and time-invariant**

Linear time-invariant (LTI) systems

- A system is linear if it fulfills the **superposition property**, i.e.:

$$\forall x_i(t):\ \mathfrak{I}\big(x_i(t)\big) = y_i(t)$$
$$\forall x_j(t):\ \mathfrak{I}\left(x_j(t)\right) = y_j(t)$$
$$\mathfrak{I}\left(\alpha x_i(t) + \beta x_j(t)\right) = \alpha y_i(t) + \beta y_j(t);\ \forall \alpha, \beta$$

- A system is time-invariant if it fulfills the **time-shift property**, i.e.:

$$\forall x(t):\ \mathfrak{I}\big(x(t)\big) = y(t)$$
$$\mathfrak{I}\big(x(t-T)\big) = y(t-T)$$

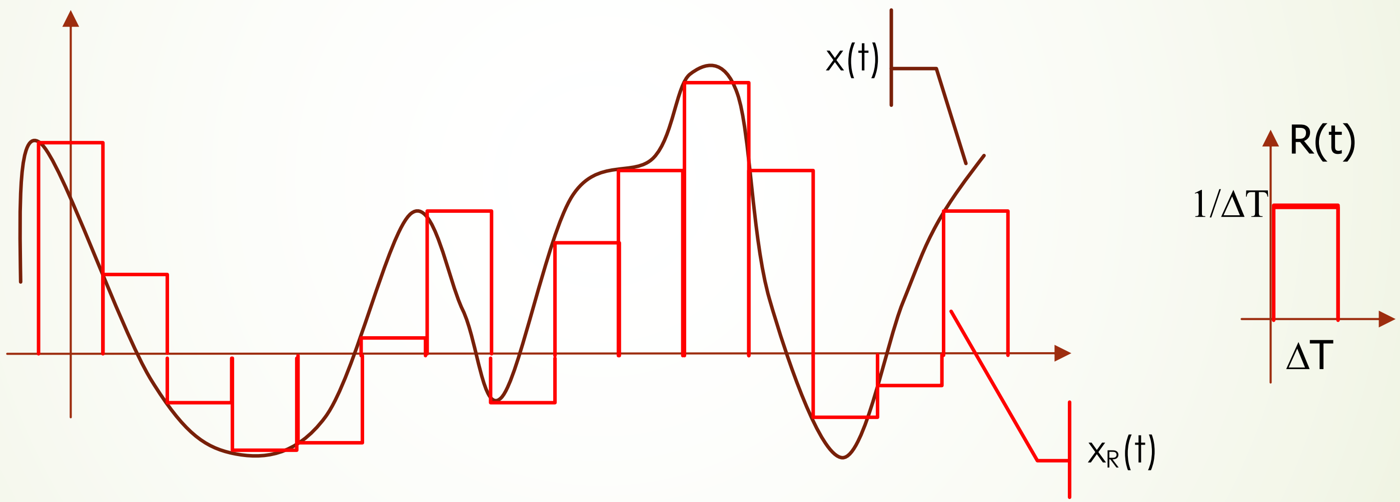
- A system that fulfills both properties is called **LTI**

Response of LTI systems

- Given an LTI system, we can calculate the response to any input signal if we know the response of the system to a single special function called **unit impulse** $\delta(t)$
  - We'll come to this important result step by step…

Response of LTI systems: step 1

- We approximate the input signal x(t) with a series of rectangular waves R(t) of fixed duration ΔT and unit area



$$x_R(t) = \sum_k x(k\Delta T) \cdot R(t - k\Delta T) \cdot \Delta T \approx x(t)$$

Response of LTI systems: step 2

- We observe the response of the system to R(t), call it $h_R(t)$
  - It is sufficient to input R(t) and measure the output $h_R(t)$
- Now, since the system is LTI, then:
  - The response to a weighted sum of inputs R(t) is the weighted sum (with same weights) of the corresponding outputs $h_R(t)$
  - A translation of R(t) by any arbitrary interval T produces the corresponding output $h_R(t)$ translated by the same time interval $h_R(t\text{-}T)$
  - Therefore:

$$y_R(t) = \sum_k x(k\Delta T) \cdot h_R\,(t - k\Delta T) \cdot \Delta T \approx y(t)$$

Response of LTI systems: step 3

- We improve the approximations of step 1 and 2 by sending $\Delta t \rightarrow 0$.
  - In the limit, R(t) becomes a function with null duration and infinite amplitude, but still unit area. This function is just theoretical and is called **unit impulse** (or **Dirac delta**). It is represented with a vertical arrow.

$$\lim_{\Delta t \to 0} R(t) = \delta(t)$$

  - Accordingly, in the limit $h_R(t)$ becomes the response to a unit impulse, called **impulse response** h(t)
  - x(t) becomes an infinite series of consecutive impulses, and should be re-written in integral form. The same holds for $y_R(t)$:

$$\lim_{\Delta t \to 0} x_R(t) = \int_{-\infty}^{+\infty} x(\tau)\delta(t-\tau)d\tau = x(t)$$

$$\boxed{\lim_{\Delta t \to 0} y_R(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)d\tau = y(t)}$$

Convolution theorem

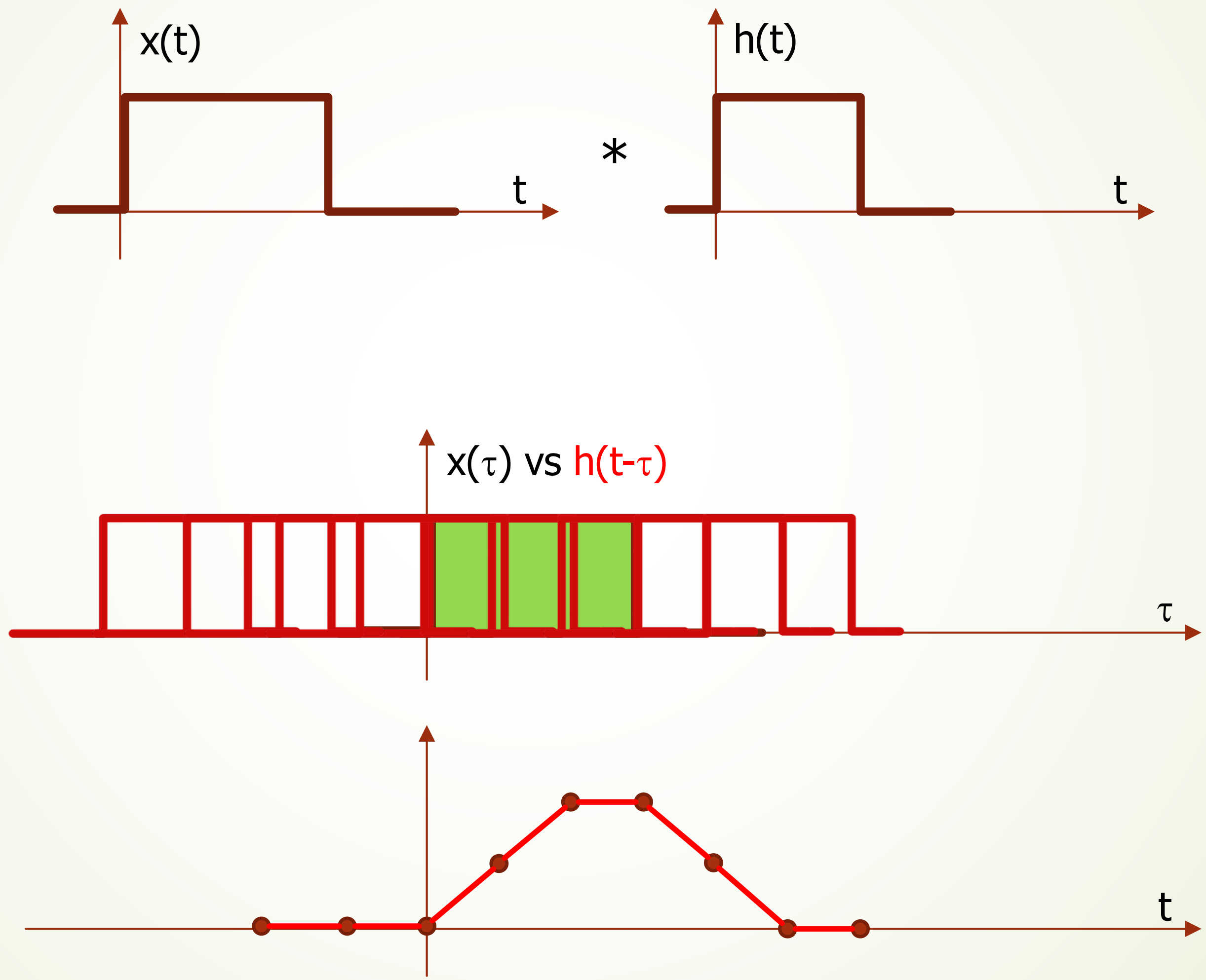
- The above result is very important
  - Given an LTI system with impulse response h(t), the response to a generic input signal x(t) will be:

$$\boxed{y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)d\tau = x(t) * h(t)}$$

CONVOLUTION INTEGRAL

  - This means that, for an LTI system, it is sufficient to know the impulse response to analytically calculate the response to any input signal
  - Nevertheless, the above integral should be hard to solve…
    - *In general, it will be feasible with simple functions (mathematical or graphical methods) or with numeric computation (we'll see later)*

Example of convolution

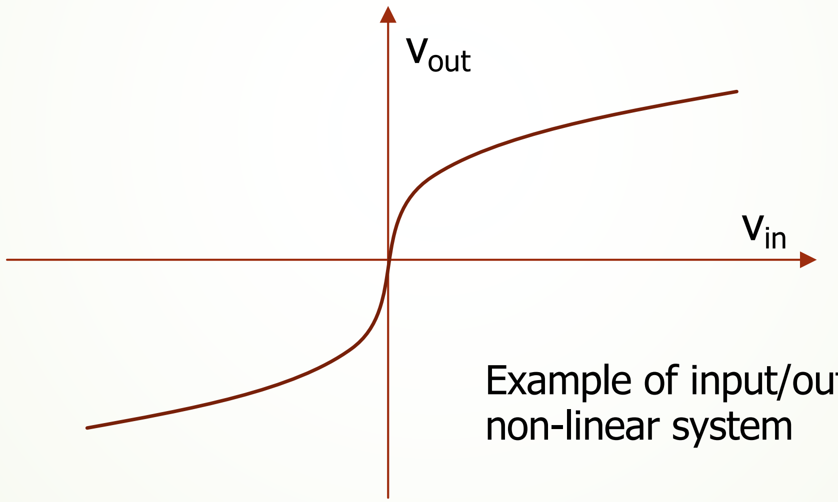


And if the system is not LTI?

- If the system is not LTI, impulse response exists, but it does not describe the behavior of the system for a general input
  - Since we miss the two fundamental properties (superposition and time-shift invariance), convolution theorem does not hold!
- We can just represent the behavior of the system at a specific point in the domain: the so-called **Input/Output Relationship**

Input/output relationship

- The input/output relationship determines the output value for a given input value
  - It is not a function of time or space, but it simply expresses the behavior of the system at a specific point in the input domain (time, space, …)

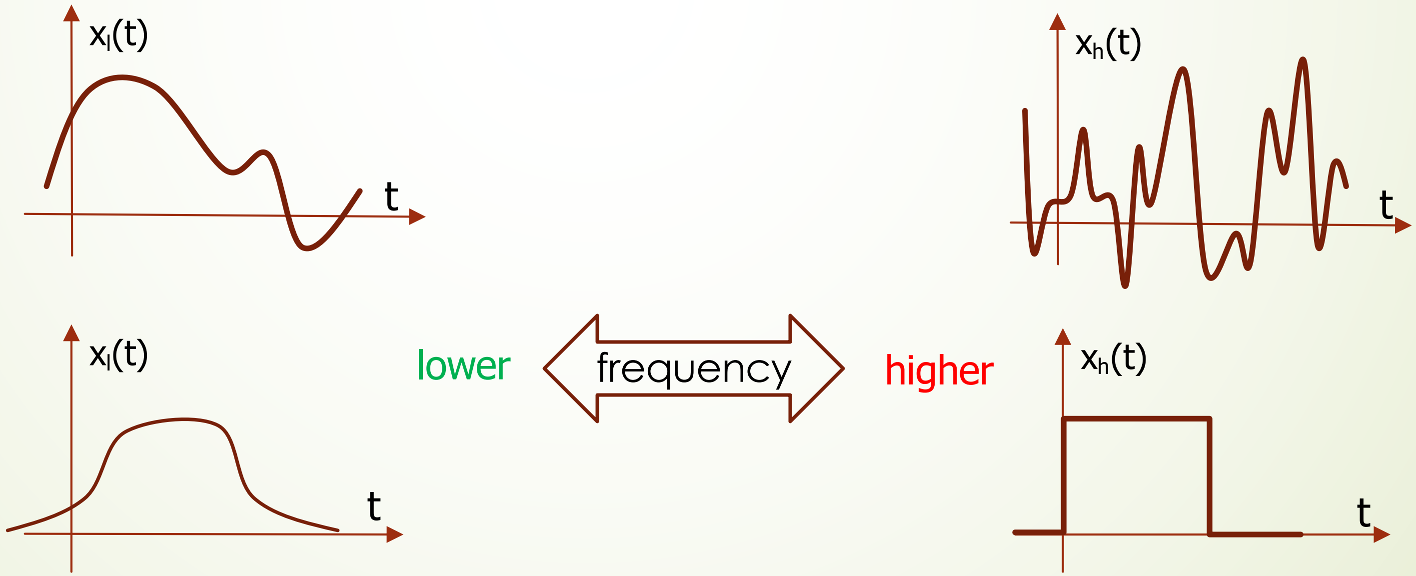


Example of input/output relationship of a non-linear system

- *NB. For an LTI system, the input/output relationship exists and it is simply a straight line*

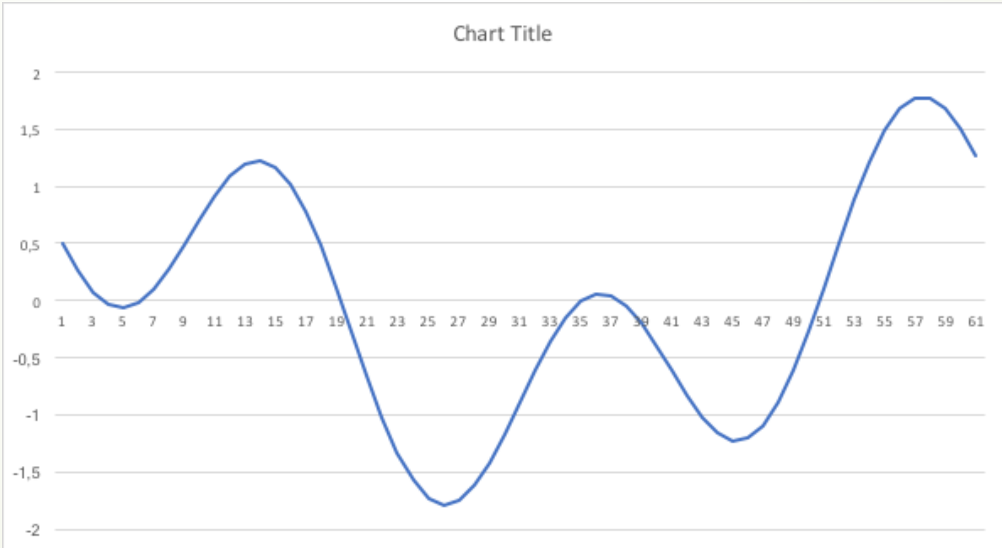
The concept of frequency in signals

- In physics, the concept of frequency is typically connected to periodic events (e.g., the oscillations of a pendulum)
- In signals the concept is somewhat larger. Higher frequencies are associated to signals that:
  - Present **more variations per domain unit** → similar to periodicity
  - Present **steeper transitions** → less intuitive



How to measure frequency content

- The previous definition is qualitative
- In many cases we need to precisely calculate the frequency content of a signal



What's the frequency content of the above signal?

How to measure frequency content

- Since we know exactly the frequency of a sinusoid, it would be nice to express a generic signal in terms of sinusoids
  - If I could represent a generic signal as a combination of sinusoids, its frequency content would become straightforward
- This turns out to be feasible, thanks to **Fourier** and his transform

Jean-Baptiste Joseph Fourier (1768-1830)

The Fourier Transform (FT)

- Fourier demonstrated that a signal can be decomposed in an infinite series of sinusoidal waves, with varying amplitude, frequency and phase
- For any given frequency, the FT provides:
  - **amplitude** → how much that frequency is present in the signal
  - **phase** → which is the offset of that frequency in the signal
  - To represent both, we need a **complex function**

The Fourier Transform (FT)

- To calculate how much a given frequency f_0 is present in a signal x(t), we need to "correlate" the signal with the sinusoid at f_0:
  - 1. Multiply the signal by a complex sinusoid at frequency f_0
  - 2. Integrate over the whole domain

$$\boxed{X(f_0) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi f_0 t} dt}$$

where $e^{-j2\pi f_0 t} = cos2\pi f_0 - jsin2\pi f_0$ (Eulero formula)

  - $X(f_0)$ is a complex number with:
    - $|X(f_0)|$ → modulus, intensity of frequency f_0 in x(t)
    - arg $X(f_0)$ → phase, offset of frequency f_0 in x(t)
  - *NB. If x(t) doesn't contain the frequency f_0, $X(f_0)$ will be zero*

From time to frequency and back

- Extending the above computation to the whole frequency domain we obtain the **forward transform**:

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi f t} dt$$

FOURIER TRANSFORM

- It is a continuous function of *f*, returning all the (possibly infinite) frequency components of x(t)
- It is also possible to reconstruct the signal x(t) from its frequency representation (**inverse transform**)
  - We need to sum up all the frequency components, each one with the amplitude and phase given by the corresponding FT value

$$x(t) = \int_{-\infty}^{+\infty} X(f) e^{j2\pi f t} df$$

INVERSE TRANSFORM

A few typical examples: sinusoids

- If we apply the above definition to a sinusoidal signal, it is clear that only the frequency corresponding to the input will "correlate", all the others will output a null value
  - Accordingly, the transform of a sinusoidal signal at frequency f_0 will be an impulse in f_0 and zero elsewhere
  - The difference between a cosine and a sine will just be a phase shift

$$\mathcal{F}(cos2\pi f_0 t) = \frac{1}{2}[\delta(f - f_0) + \delta(f + f_0)]$$
$$\mathcal{F}(sin2\pi f_0 t) = \frac{1}{2j}[\delta(f - f_0) - \delta(f + f_0)]$$

NB. Due to the mathematical model of the transform, the frequency representation of real signals presents a symmetry, then there is always a matching "negative frequency" counterpart (with no physical meaning)

A few typical examples: impulses vs. constants

- Strange as it may appear, it comes out that an impulse signal (Dirac delta) contains all the frequencies with the same amplitude and shift.
    - Accordingly, its transform is a constant in the frequency domain
- Conversely, the transform of a constant is an impulse in frequency
    - This is more intuitive, if we think to a constant as a sinusoid at frequency zero

$$\mathcal{F}\big(\delta(t)\big) = 1$$

$$\mathcal{F}(1) = \delta(f)$$

NB. Intuition 1: an impulse is the steepest possible variation: as such, it contains all possible frequencies

NB. Intuition 2: this also explains why impulses are so representative of a system (convolution theorem): they excite all frequencies responses

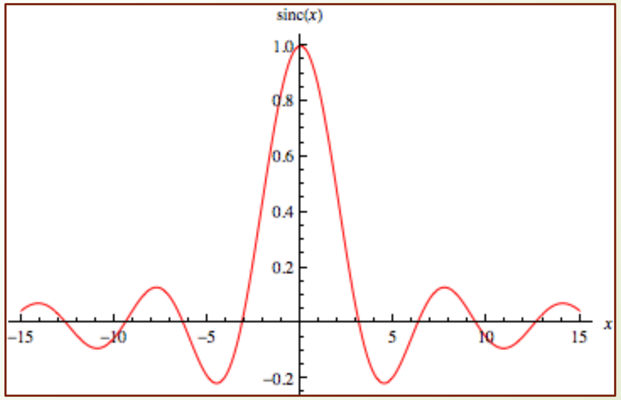
A few typical examples: rectangles

- Another function we came across is the rectangle R(t), let see its Fourier counterpart:

$$\boxed{\mathcal{F}\left(R_T(t)\right) = T\,\frac{sin(\pi f T)}{\pi f T} \triangleq sinc(fT)}$$

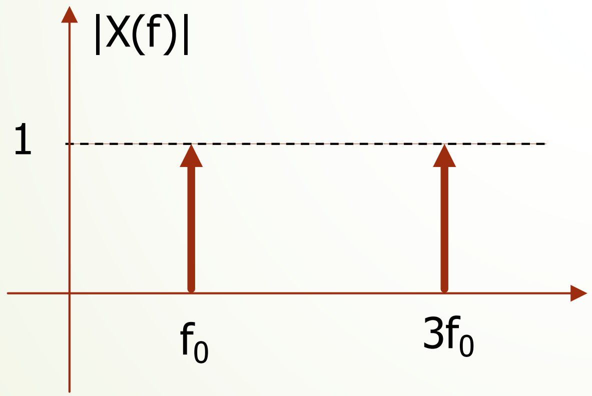
where $R_T(t)$ is a rectangular function of duration T and unit amplitude, centered on t=0

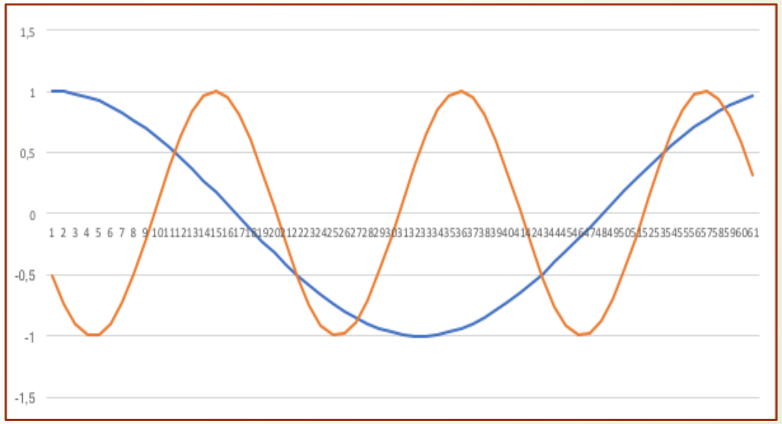
NB. Also in this case the steep transitions produce infinite frequency components, but with a damping factor proportional to 1/T



Back to our example

- If we calculate the FT of the signal in slide 16, we simply obtain 2 impulses of unit amplitude
  - The first represents a cosine a frequency f_0 and 0 phase
  - The second is a cosine at frequency $3f_0$ and phase $2\pi/3$

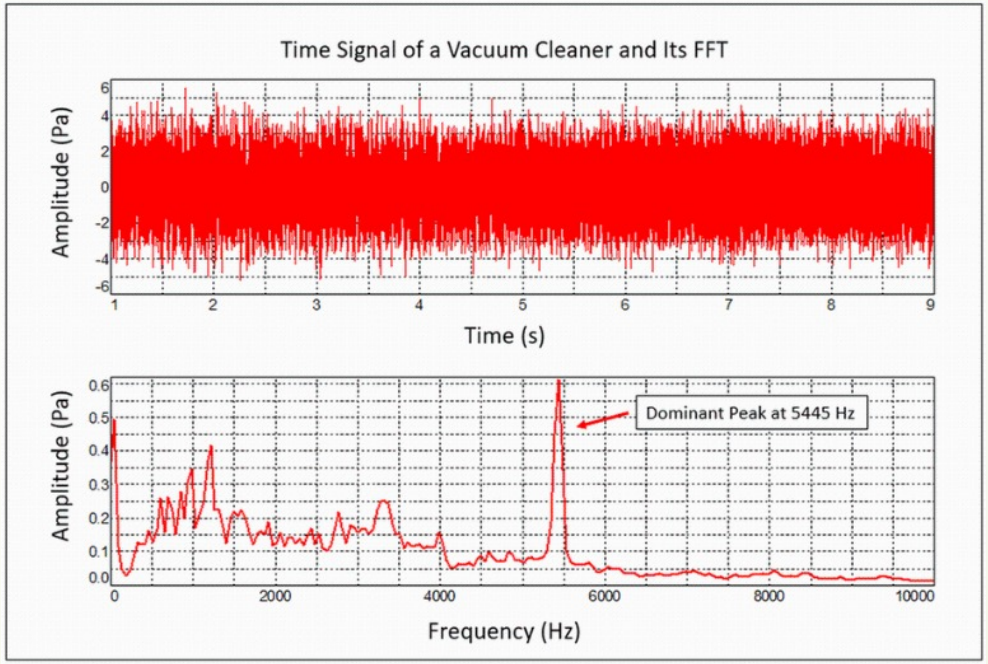




The signal is the sum of two sinusoidal waves

More in general

- FT gives an immediate feeling of the frequency content of an unknown generic signal



T. Mila, Siemens Community Article, community.sw.siemens.com, 2019

FT vs LTI systems response

- Coming back to systems, we have seen that we can calculate the response to an LTI systems provided that we know its impulse response
  - This operation (convolution) could be difficult to calculate
- It can be demonstrated that in the Fourier domain the **convolution becomes a simple product**!

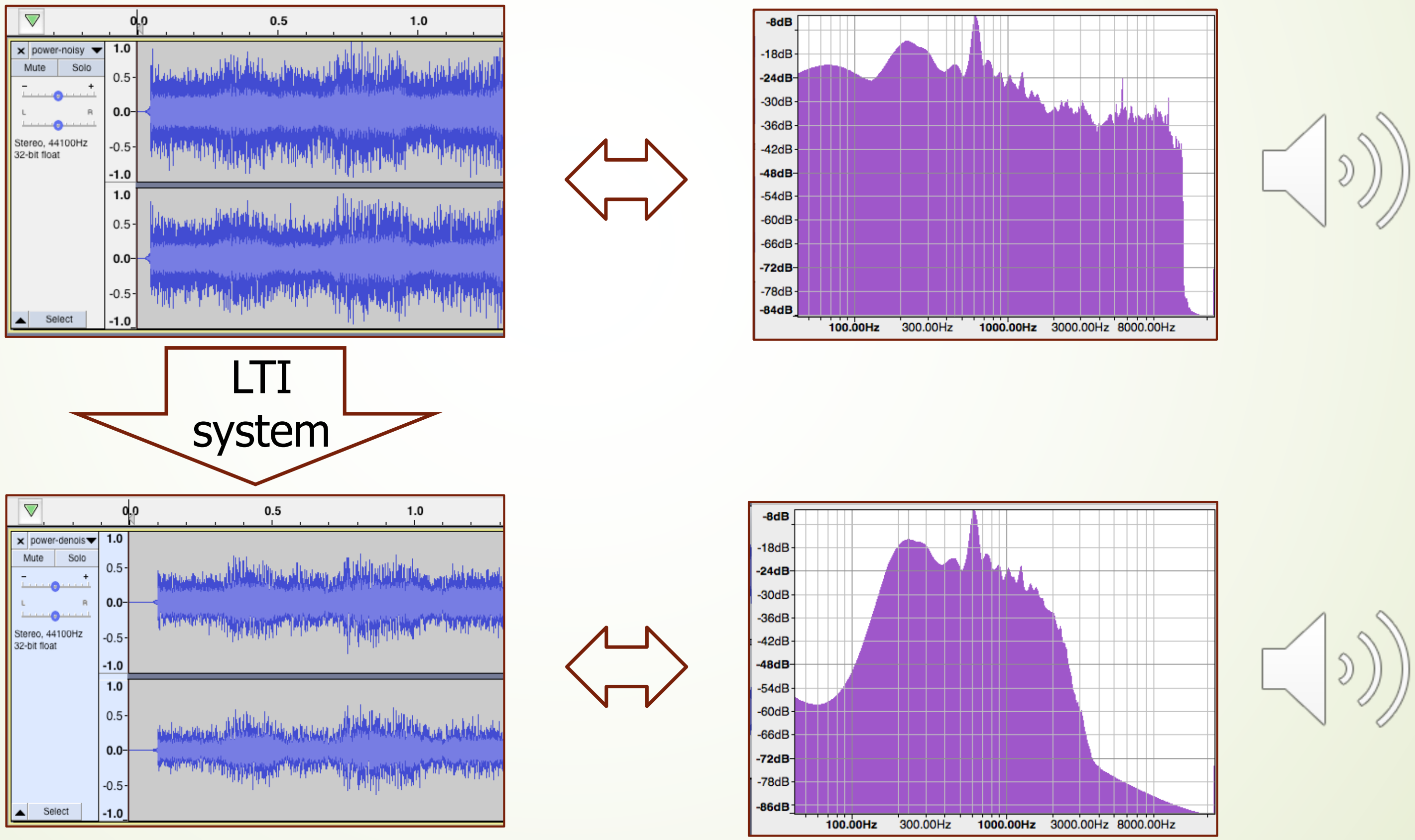
$$\boxed{\mathcal{F}(x(t) * h(t)) = \mathcal{F}\big(x(t)\big) \cdot \mathcal{F}\big(h(t)\big) = X(f) \cdot H(f) = Y(f)}$$

- The transform of the impulse response h(t) is called **frequency response** H(f) of the LTI system, and it completely determines the behavior of the system

NB. this is true only for LTI systems

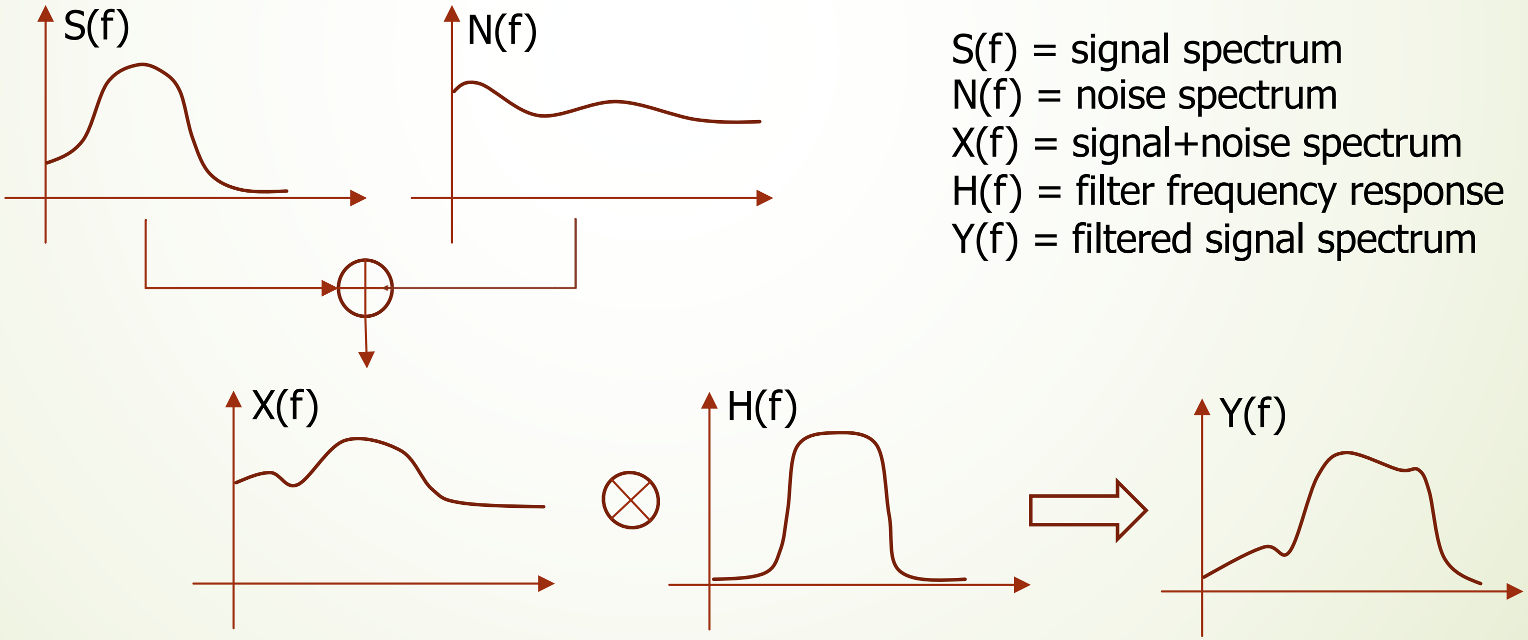
Let’s apply what we’ve seen so far

- Imagine you have a sound immersed in noise and you want to design a system to reduce noise while preserving the sound



Linear filters

- How did we achieve this?
  - Typically signals have different **frequency spectra**
  - Linear filters modify frequency spectrum components
  - In the example, voice and noise cover different frequencies, then, we can try to attenuate the noise where the voice is less relevant



S(f) = signal spectrum
N(f) = noise spectrum
X(f) = signal+noise spectrum
H(f) = filter frequency response
Y(f) = filtered signal spectrum

Linear filters

- The filter in the previous slide is called **bandpass** (BPF)
  - It cuts (attenuates) lower and higher frequencies, and leaves almost unchanged the intermediate ones
  - We can place the lower/upper cuts as to preserve the wanted signal (voice) and remove as much as possible the unwanted signal (noise)
- In the frequency domain (Fourier) its behavior is easy understood:

$$Y(f) = X(f) \cdot H(f)$$

- In the time domain it is much harder:

$$y(t) = x(t) * h(t)$$

where h(t) is the inverse transform of H(f), typically, a sort of *sinc* function

Linear filters

- Exactly the same way it is possible to define various types of filters useful for different applications:
  - LOW-PASS FILTER (LPF): cuts higher frequencies leaving lower ones unchanged
  - HIGH-PASS FILTER (HPF): cuts lower frequencies leaving higher ones unchanged
  - ALL-PASS FILTER: leaves signal unchanged
  - NOTCH FILTER: cuts a given part of the spectrum from the signal leaving the rest unchanged

- Typical usage:
  - Noise removal (LPF, BPF)
  - Interference removal (BPF)
  - Signal separation (BPF)
  - Highlighting variations (HPF)

What we've learned in this section

- Signals convey information, systems are used to manipulate them
- LTI systems are particularly easy to handle, they response can be calculated based on impulse response and convolution theorem
- Non-LTI systems can only be studied in terms of input/output relationship (on a fixed time instant)
- Signals can be decomposed in sinusoids, thanks to the Fourier Transform: this allows understanding their frequency content
- Fourier transform is reversible: I can easily switch from time to frequency domain and vice versa without loosing information
- In the Fourier domain it is much easier to understand the behavior of systems (filters), but this applies only to LTI systems