Discrete Fourier Transform

Institut Mines-Télécom

Discrete Fourier Transform (DFT)

The DFT of a finite signal (FS) defined on $\{0, \ldots, N-1\}$ is another FS defined on the same support $\{0, \ldots, N-1\}$:

$$\forall k \in \{0,1, \ldots, N-1\}, \qquad \hat{u}_k = \sum_{n=0}^{N-1} u_n e^{-2i\pi \frac{k}{N} n}$$

The index of $\hat{u}$ is k, but the corresponding wave frequency is k/N

Inversion theorem

The inversion formula is:

$$u_n = \frac{1}{N}\sum_{k=0}^{N-1} \hat{u}_k e^{2i\pi\frac{k}{N}n}$$

proof :

$$\frac{1}{N}\sum_{k=0}^{N-1} \hat{u}_k e^{2i\pi\frac{k}{N}n} = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{m=0}^{N-1} u_m e^{-2i\pi\frac{k}{N}m}\ e^{2i\pi\frac{k}{N}n}$$

$$= \frac{1}{N}\sum_{m=0}^{N-1} u_m \sum_{k=0}^{N-1} z^k$$

with $z = e^{2i\pi\frac{n-m}{N}}$

Inversion theorem

Given that $z = e^{2i\pi\frac{n-m}{N}}$, we find easily :

$$\sum_{k=0}^{N-1} z^k = N\delta_{n-m} \qquad \forall n, m, \in \{0,1,\dots,N-1\}$$

By replacing in the previous equation we find :

$$\frac{1}{N}\sum_{k=0}^{N-1} \hat{u}_k e^{2i\pi\frac{k}{N}n} = \frac{1}{N}\sum_{m=0}^{N-1} (u_m N\delta_{n-m}) \quad = u_n$$

In other terms, $\frac{\hat{u}(k)}{N}$ are the coefficients of a decomposition on a Fourier basis

Classical properties

$$u = \delta^m \Rightarrow \hat{u}_k = e^{-2i\pi\frac{m}{N}k}$$

$$\mathcal{F}(u \circledast_N v) = \hat{u}\,\hat{v}$$

$$\mathcal{F}(uv) = \frac{1}{N}\hat{u} \circledast_N \hat{v}$$

Circular convolution

$$\mathcal{F}(\phi u) = \hat{u}(k - k_0) \qquad \phi_n = e^{2i\pi\frac{k_0}{N}n}$$

$$\mathcal{F}(u^m) = \hat{u}_k e^{-2i\pi\frac{m}{N}n}$$

Symmetry properties

Circular permutation

Parseval "equality"

- Fourier waves are orthogonal with norm $\sqrt{N}$
- We deduce:

$$\|\hat{u}\|^2 = \sum_{k=0}^{N-1} |\hat{u}_k|^2 = \sum_{k=0}^{N-1} \sum_{n=0}^{N-1} u_n e^{-2i\pi \frac{k}{N} n} \sum_{m=0}^{N-1} \bar{u}_m e^{2i\pi \frac{k}{N} m}$$

$$= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} u_n \bar{u}_m \sum_{k=0}^{N-1} e^{-2i\pi \frac{k}{N}(n-m)} = N\|u\|^2$$

Links between DT and DTFT

- DFT is the only transform that can be computed on a computer …
- DFT can approximate DTFT under certain hypotheses

TELECOM ParisTech
Institut Mines-Télécom

Case of a finite support sequences

- Consider u defined on $\mathbb{Z}$, with finite support:

$u_n = 0 \quad \forall n \notin \{0, \dots, N-1\}$.

- Let v be the restriction of u to $\{0, \dots, M-1\}$, with $M \geq N$,

$$\hat{v}_k = \sum_{n=0}^{M-1} v_n e^{-2i\pi\frac{k}{M}n} = \hat{u}\left(\frac{k}{M}\right)$$

- Sometimes $\hat{v}$ is called M-DFT of u (zero-padding)
- We talk about M-DFT of a finite support sequence

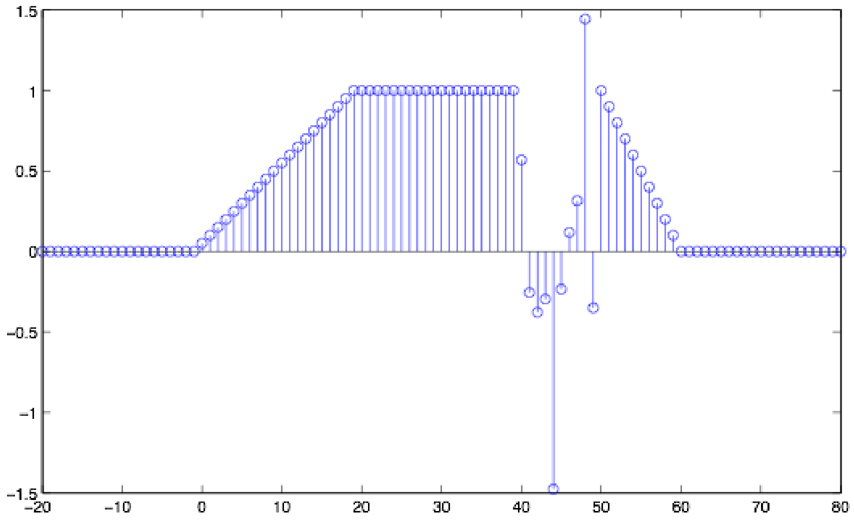
Case of a finite support series

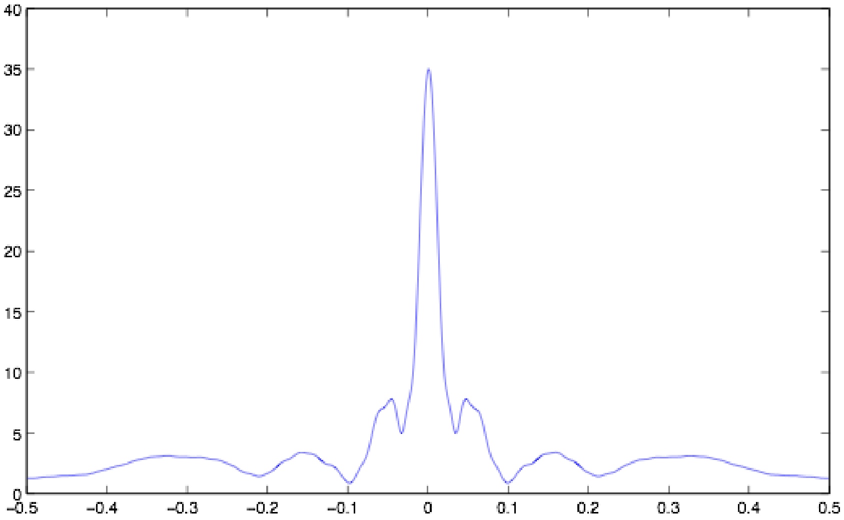
- By changing M, one can sample $\hat{u}(\nu)$ as finely as necessary
- This is equivalent to a *zero-padding*
- However, one needs only N samples of $\hat{u}(\nu)$ to perfectly know (reconstruct) $\hat{u}$

$$\left\{\hat{u}\left(\frac{k}{N}\right)\right\}_{k\in\{0,\ldots,N-1\}} \xrightarrow[\mathrm{DFT-I}]{} u \xrightarrow[\mathrm{DTFT}]{} \hat{u}(\nu)$$

- How to generalize ? When is it possible with samples at $\frac{1}{N}$ to reconstruct a function of real variable?

Example

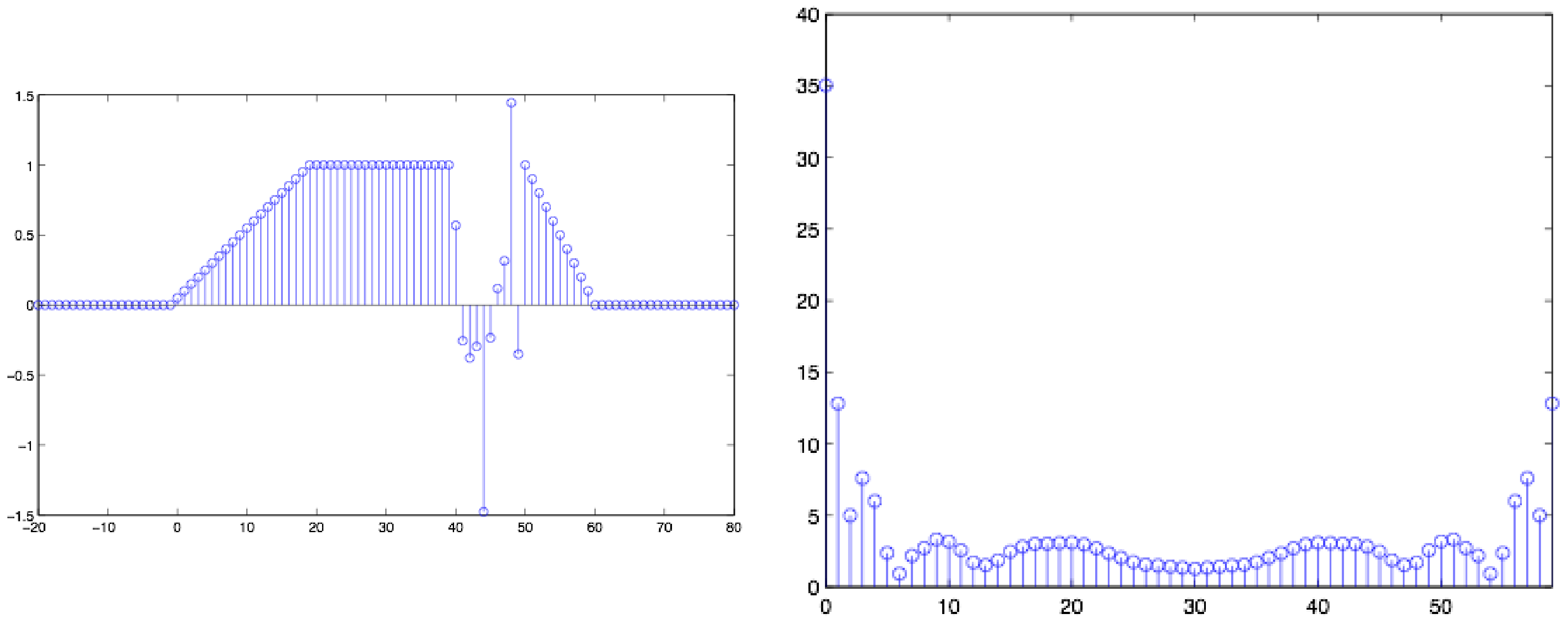




Signal on Z (left) and its DTFT (right)

Support {0, ..., 59}

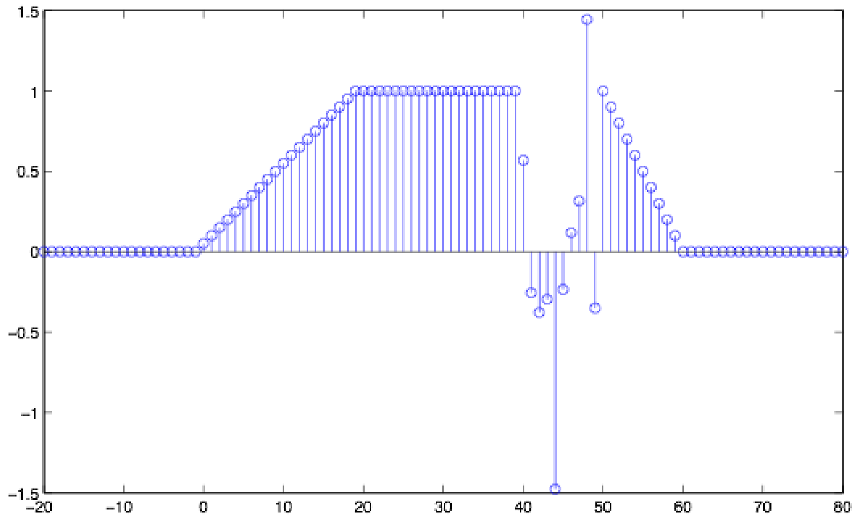
Example

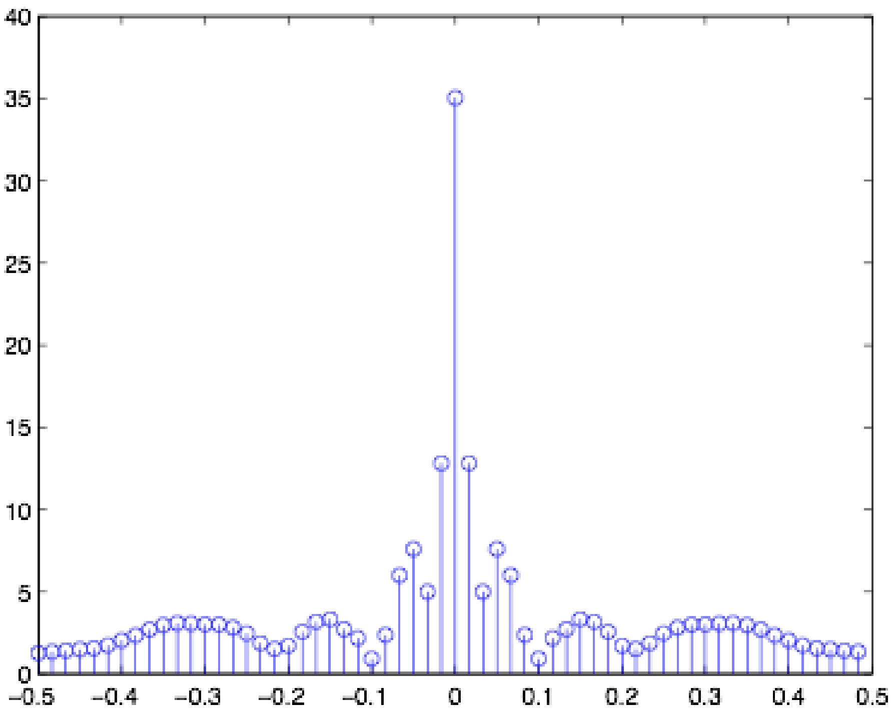


Right: 60-DFT (indexed by k) of the non null part of the signal on left.

TELECOM ParisTech
Institut Mines-Télécom

Example

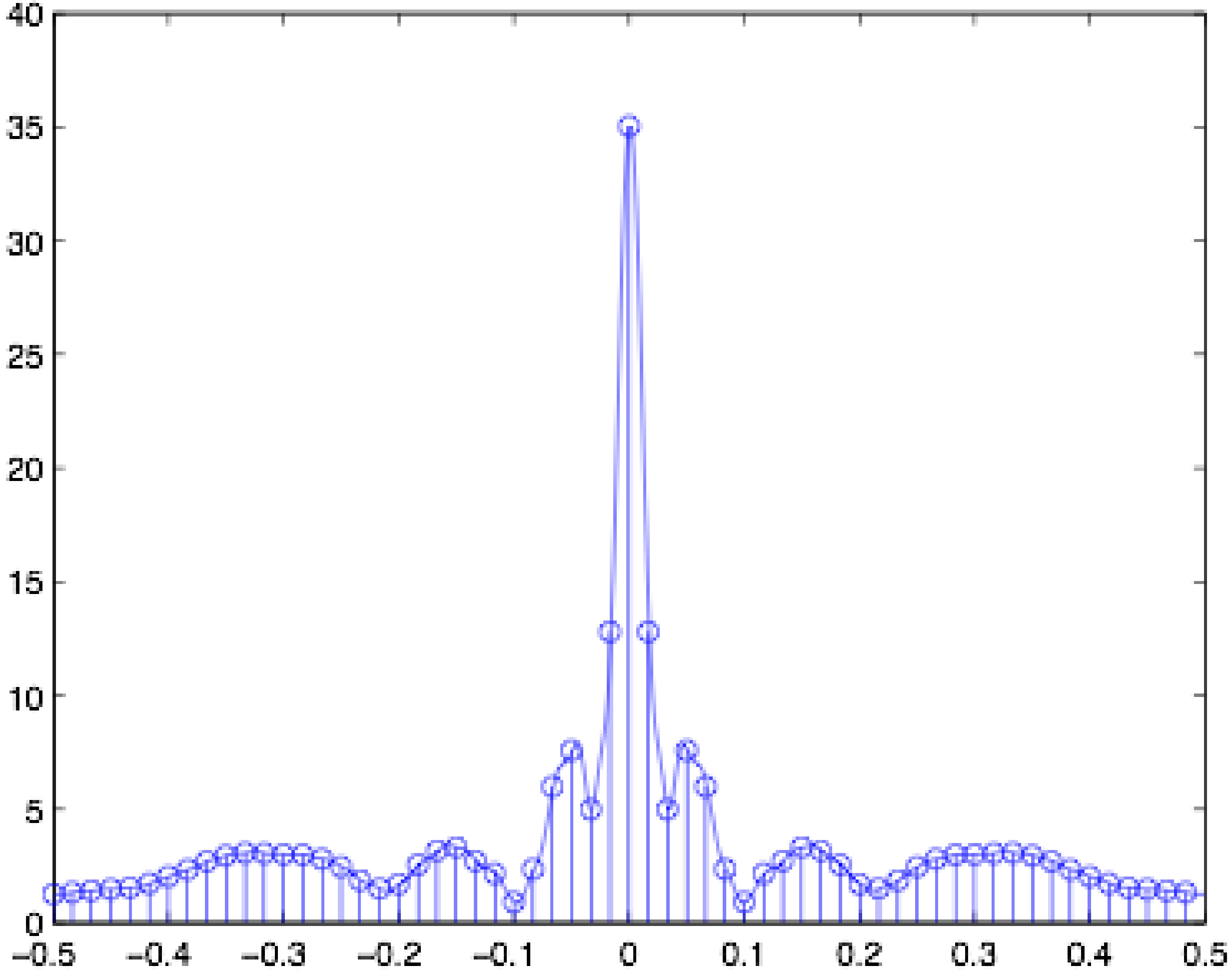




Indexed by k/M and periodized by period 1

to remain in the interval $\left[-\frac{1}{2}, \frac{1}{2}\right[$

Institut Mines-Télécom



Superposition of DFT and DTFT. As $M \geq N$ the DFT is a *perfect* sampling of the DFTF

Wave frequency determination

- We observe N samples of a wave signal.
- From these samples, we want to find the wave frequency

Wave frequency calculation

$\forall n \in \mathbb{Z}, u_n = e^{2i\pi\nu_0 n}$ i.e., $u = \phi$, FW at frequency ν_0

We can only observe a finite number of samples; we have

$$u^T = \phi w$$

where w is a finite support sequence:

$$w_n = \begin{cases} 1 & \text{if } n \in \{0, \dots N-1\} \\ 0 & \text{otherwise} \end{cases}$$

Rectangular window

Wave frequency calculation

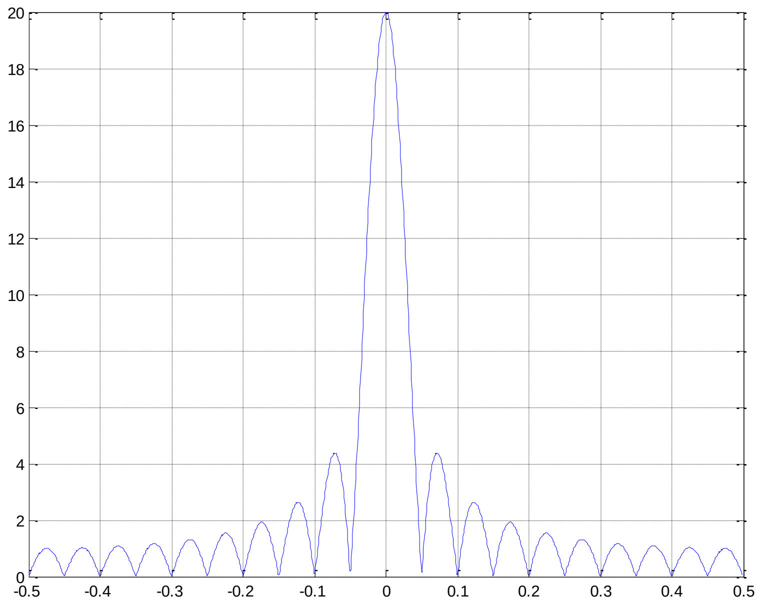
- The DTFT of u^T:

$$\mathcal{F}(u^T) = \mathcal{F}(\phi w) = \hat{w}(\nu - \nu_0)$$

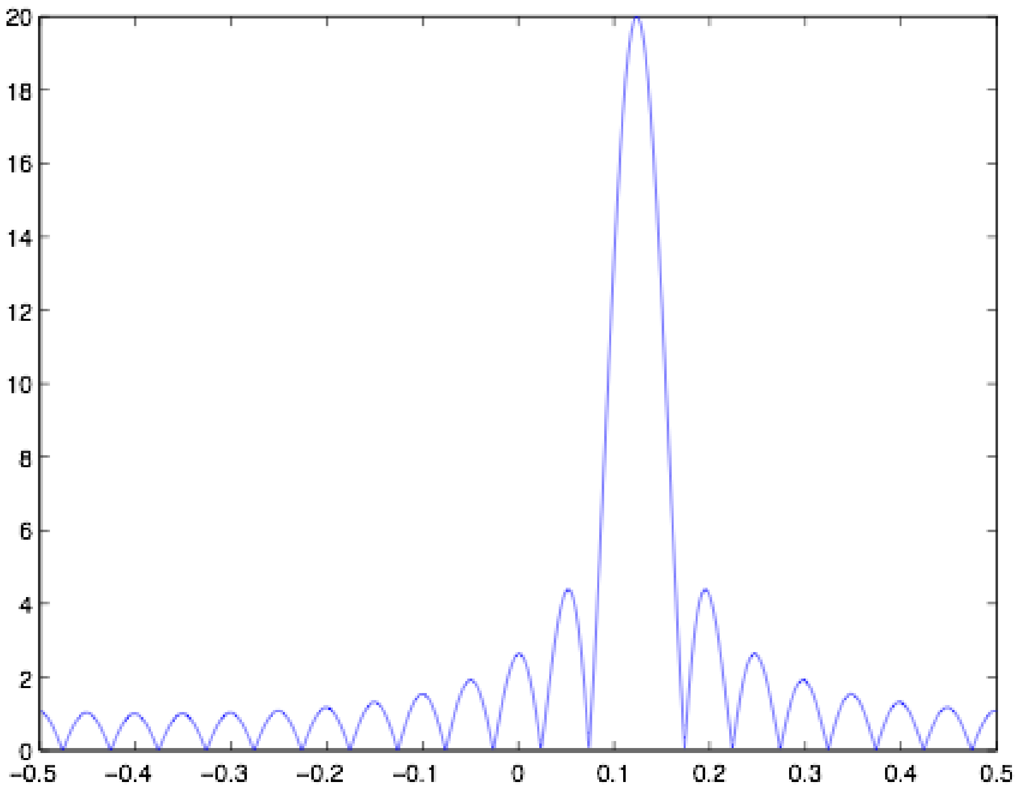
We find:

$$|\hat{w}(\nu)| = \frac{|\sin \pi N \nu|}{|\sin \pi \nu|}$$

Then, ν_0 is the position of the maximum of $\mathcal{F}(u^T)$



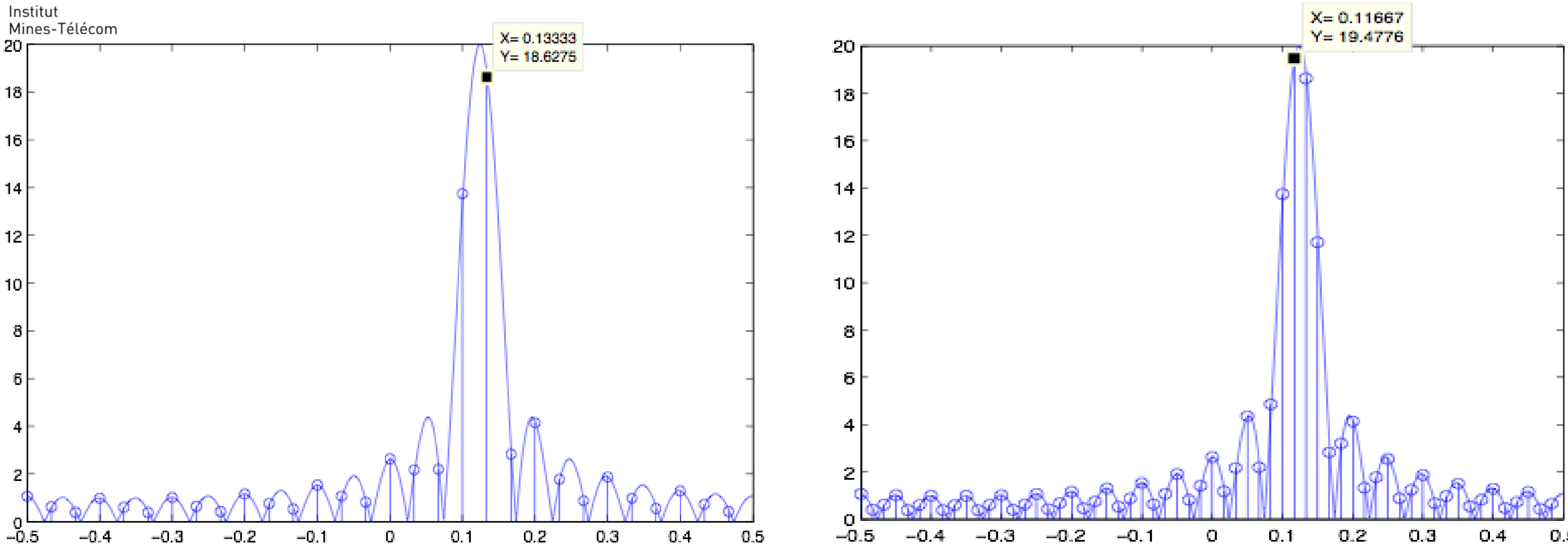
$|\hat{w}(\nu)|$ for N=20



DTFT of a wave of frequency 0,123
trunkated at 20 samples

Wave frequency calculation

- Problem : we cannot compute $\widehat{u^T}(\nu)$, but only its samples at $1/M$
- The position of the maximum will therefore be known with a precision related to the order M of the DFT and rather than to the duration of the observation, N



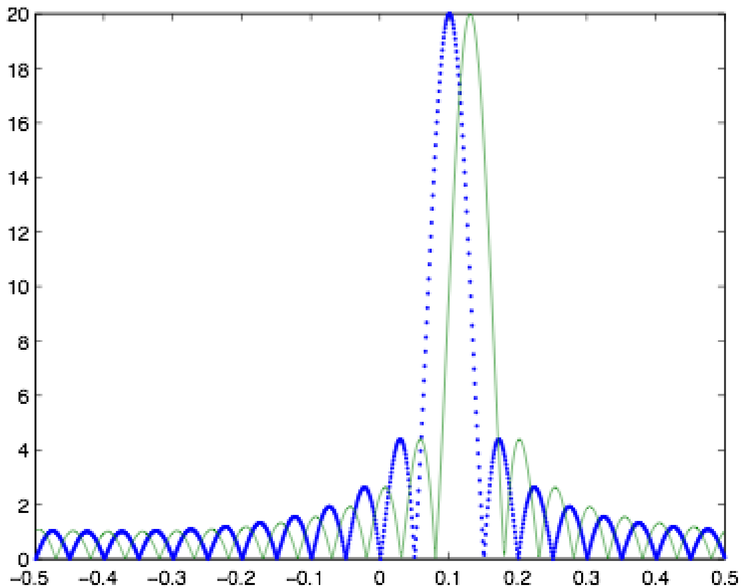
- 30-DFT (left) and 60-DFT (right)
- The precision for the frequency computation is $1/M$
- We select the index k for which the DFT is maximum

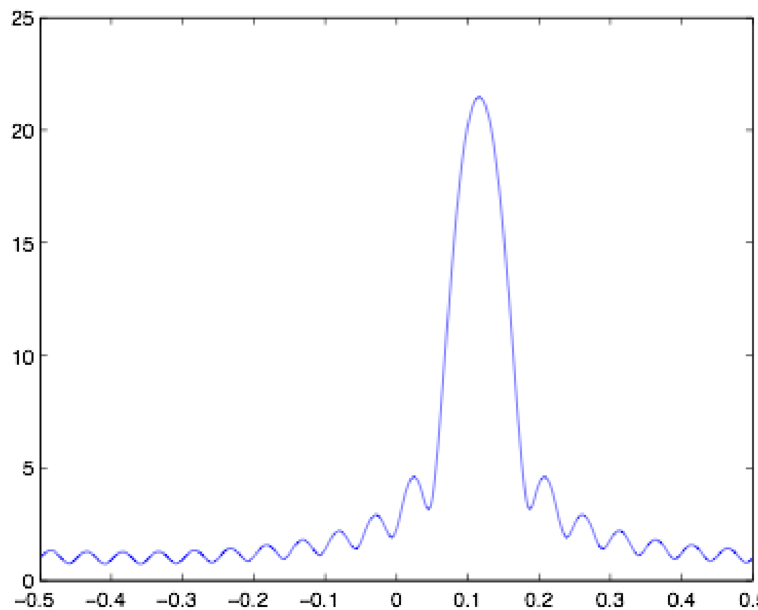
Separation of two frequencies (waves)

- The duration of observation affects the frequency resolution
- We consider a mixture of two Fourier waves, observed over N samples:

$$u_n = A_0 e^{2i\pi\nu_0 n} + A_1 e^{2i\pi\nu_1 n}$$

Frequency resolution

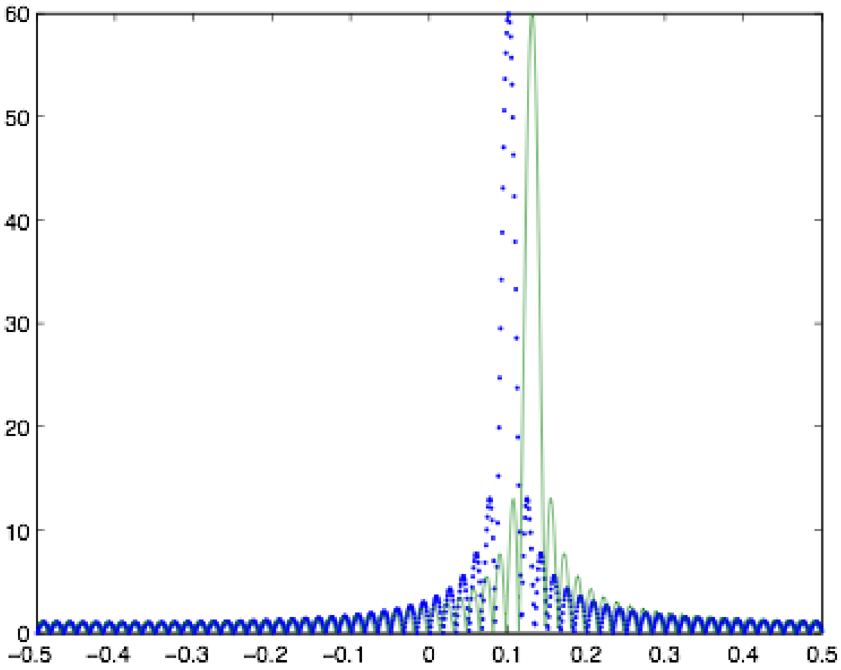


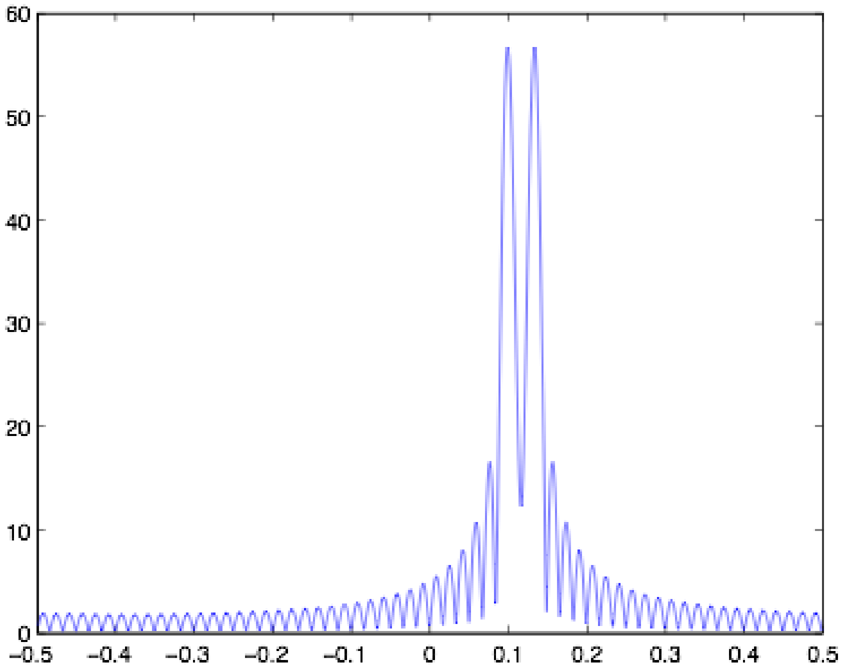


$N = 20$. Left : the 2 DTFT; right: their sum.
We cannot distinguish two close frequency waves (at less than $1/N$)

TELECOM ParisTech
Institut Mines-Télécom

Frequency resolution





$$N = 60$$

We can now distinguish the two frequencies

Frequency resolution

- The DFT of the 2 waves mixture is :

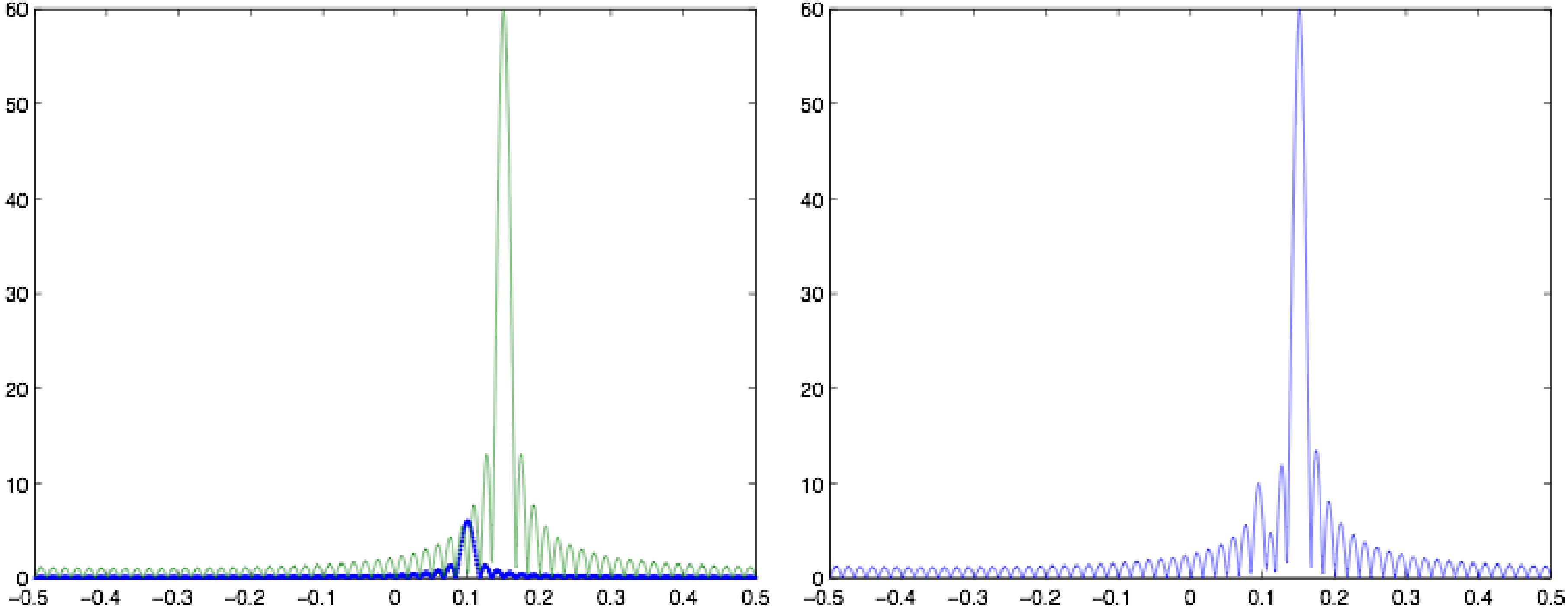
$$A_0 w\left(\frac{k}{M} - \nu_0\right) + A_1 w\left(\frac{k}{M} - \nu_1\right)$$

  - If $A_0 = A_1$, the two peaks can be separated if their lobes are separated by a half-amplitude
  - The amplitude of the lobe depends on the window: for the stair window it is $\frac{1}{N}$
  - The condition is, in this case: $N \geq \frac{1}{|\nu_0 - \nu_1|}$

Very different amplitudes, masking and windowing problems

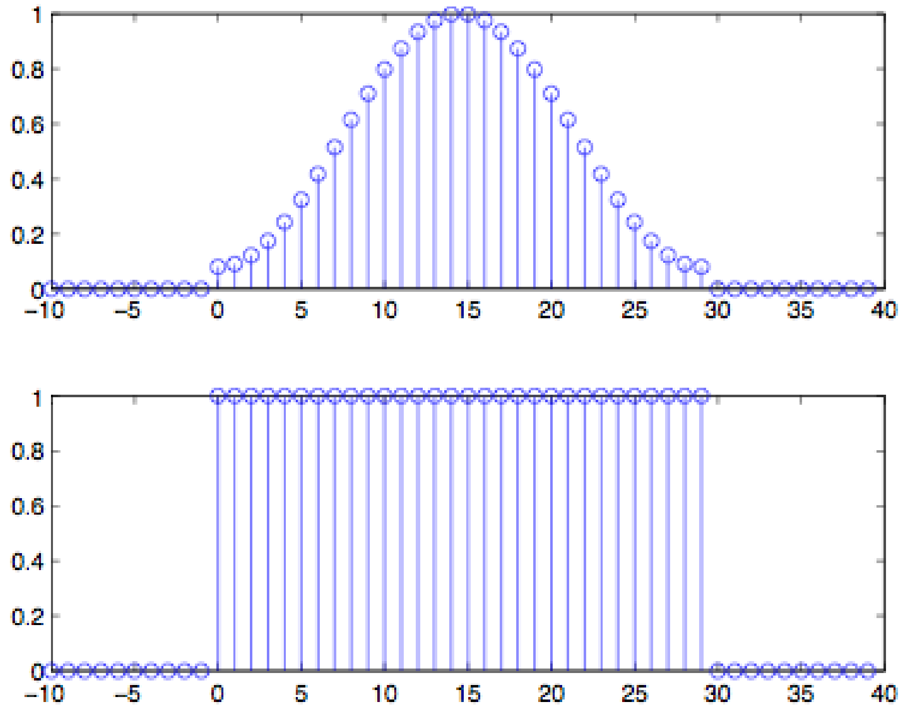
- In order to improve frequency resolution, one has to increase the number of observed samples N, if possible.
- This does not depend on the order M of the DFT
  - We still need to insure $M \geq N$
- What happens if the two waves have very different amplitudes?

TELECOM ParisTech
Institut Mines-Télécom



Left: DTFT of two waves.
Right: DTFT of their sum.
Secondary peaks mask the second wave.

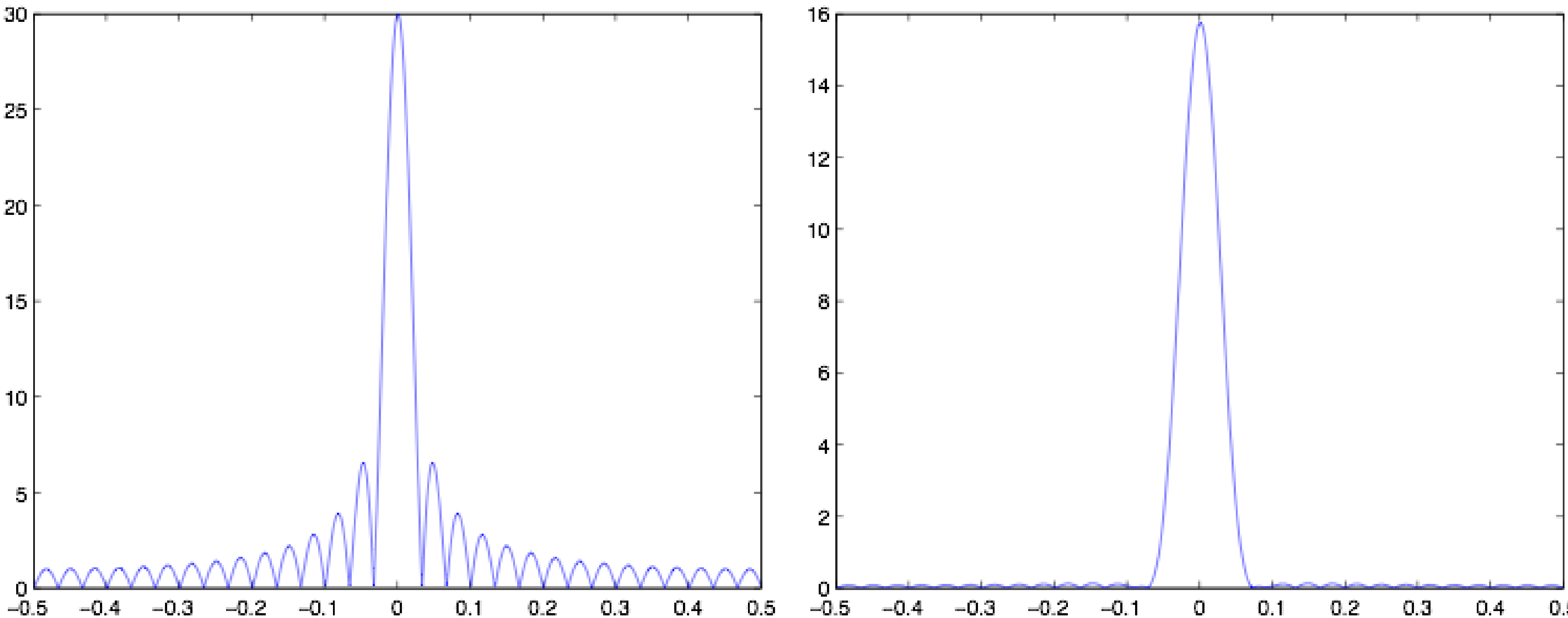
Choice of the window shape



Top: Hamming window.
Bottom: stair (or rectangular) window
Size 30 in both cases

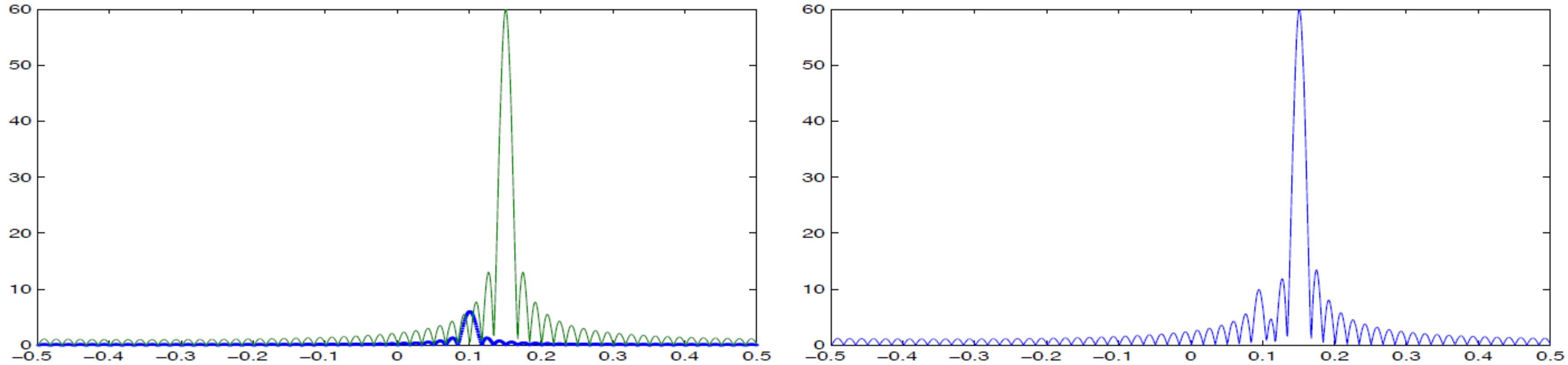
TELECOM ParisTech
Institut Mines-Télécom

Choice of the window shape

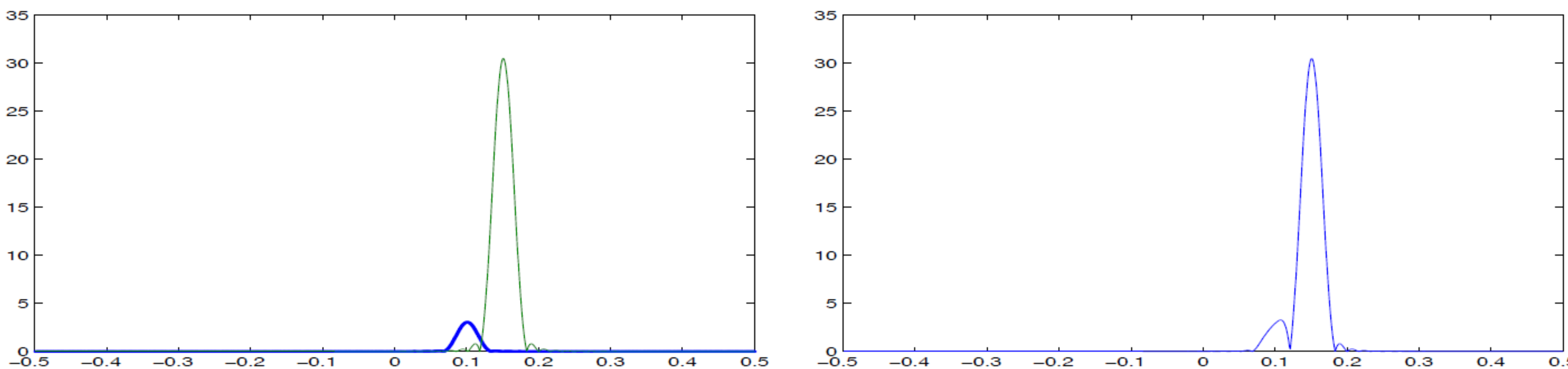


Left : DTFT of a stair window of size 30.
Right : DTFT of a Hamming window.

TELECOM



Multiplication by the Rectangular window



Multiplication by the Hamming window

Frequency analysis: conclusion

- Calculation of the frequency for a Fourier wave: precision $= 1/M$
- Separation of 2 waves with the same amplitude : $|\Delta\nu| \geq 1/N$
- Separation of 2 waves with very different amplitudes : depends on the ratio between the amplitude of the principal and secondary lobes
  - This does not depend on N , but on the *shape* of the window

TELECOM ParisTech
Institut Mines-Télécom

The spectrogram

- The idea of the spectrogram is to locally analyze the frequency content of a signal.
- Around each signal sample, we keep a window on which we compute a DTFT (through a DFT)
- For a signal *u* and a window *w* centered in zero:

$$\forall n \in \mathbb{Z}, \forall \nu \in \left[-\frac{1}{2}, \frac{1}{2}\right[, \qquad U(n,\nu) = \sum_{m\in\mathbb{Z}} u_m w_{m-n} e^{-2i\pi\nu m}$$

We can compute directly the samples of $U(n,\nu)$ through

$$\forall n \in \mathbb{Z}, \forall k \in \{0, \dots M-1\},$$

$$U\left(n, \frac{k}{M}\right) = \sum_{m\in\mathbb{Z}} u_m w_{m-n} e^{-2i\pi\frac{k}{M}m}$$

The spectrogram

- For n (time index) fixed, we have the formula
$$\forall \nu \in \left[-\frac{1}{2},\frac{1}{2}\right[, \qquad U(n,\nu) = \sum_{m\in\mathbb{Z}} u_m w_{m-n} e^{-2i\pi\nu m}$$

- This means that $U(n,\nu)$ is the DTFT of u multiplied by the window$w$ translated by n : frequency analysis around the time instant n

The spectrogram

- For a fixed frequency, we have the formula:

$$U(n,\nu_0) = \sum_m u_m w_{m-n} e^{-2i\pi\nu_0(n-m)} e^{-2i\pi\nu_0 n}$$

$$|U(n,\nu_0)| = \left| \sum_m u_m w_{m-n} e^{-2i\pi\nu_0(n-m)} \right| = |< u, \psi^n >|$$

$$\psi = w\phi^{\nu_0}$$

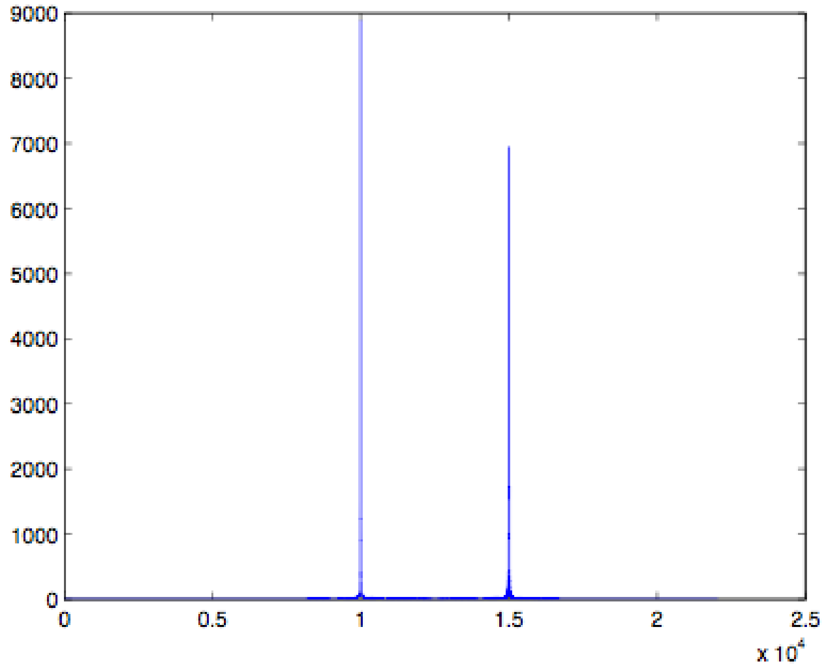
Scalar product between u and ψ^n: similitude between u and a « wave » truncated around n and with frequency ν_0

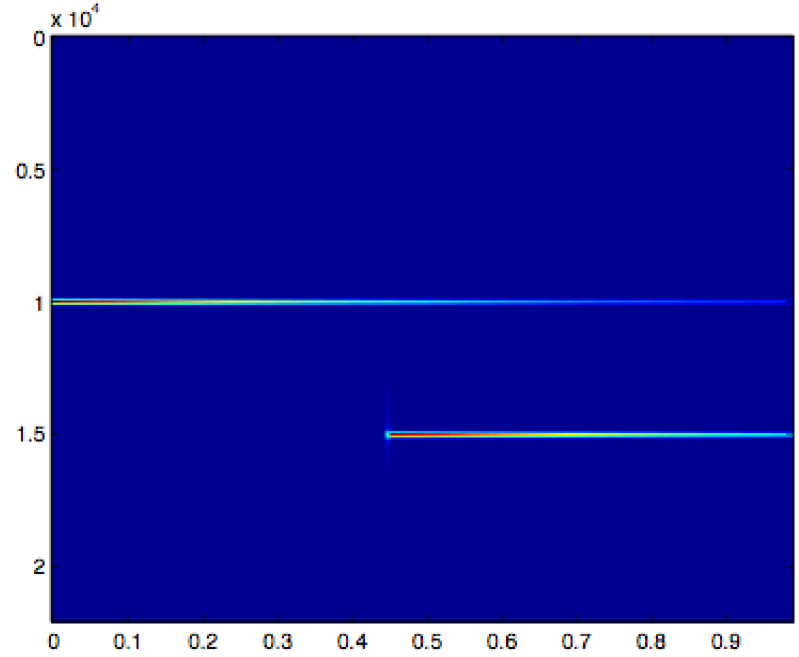
TELECOM ParisTech
Institut Mines-Télécom

Spectrogram display

- Since we have real signals, the module of the DTFT is symmetrical.
- Time axis is x and frequency axis is y (in Hz)
- We use a logarithmic scale for the module, otherwise certain frequencies will « smash » the others (ear sensitivity is btw logarithmic)

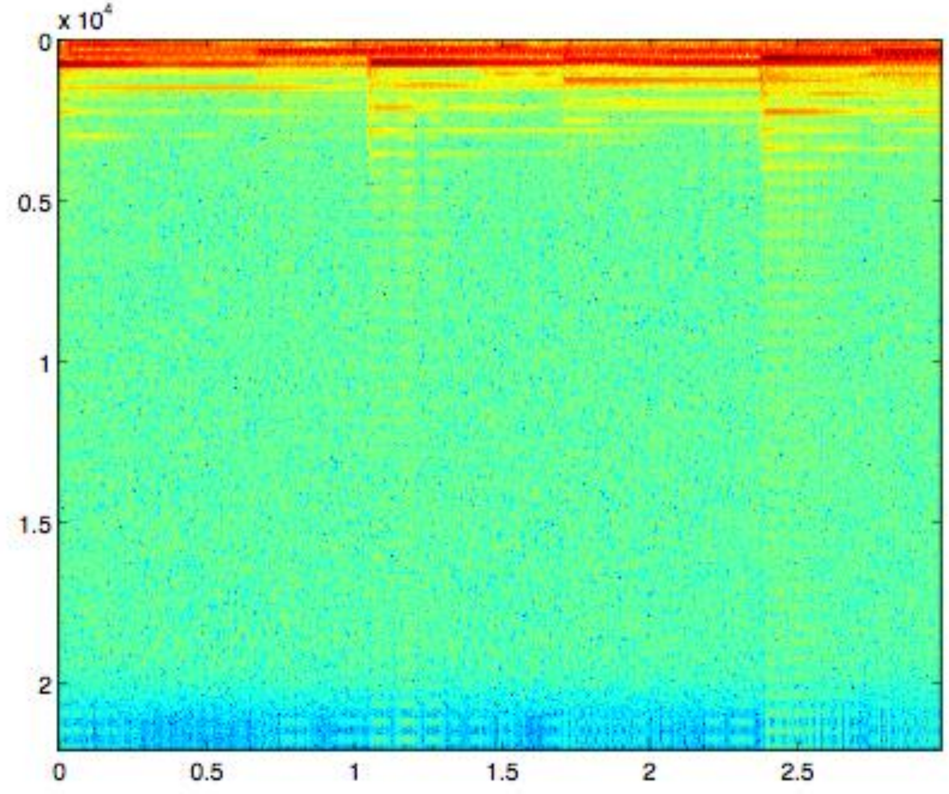
Example





Left: DTFT; right: spectrogram of the same signal.
The DTFT does not allow to know at which time instant arrives the wave at 15000Hz.

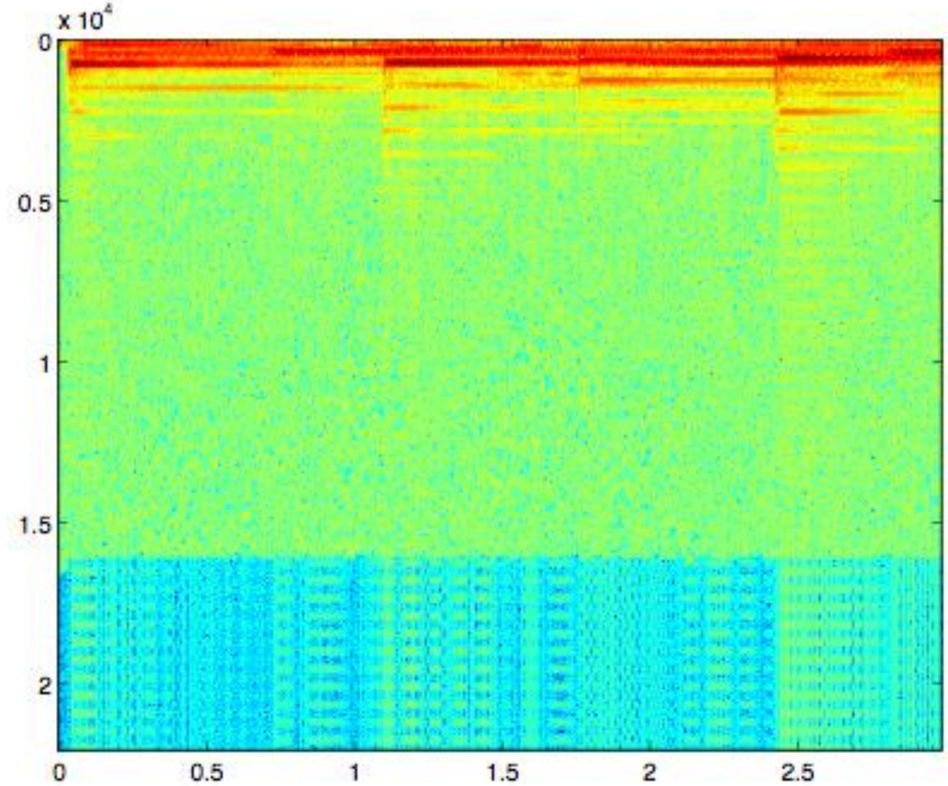
Observation of a sound on the spectrogram



Piano: Original

TELECOM ParisTech
Institut Mines-Télécom

Spectrogram of the mp3-encoded version



MP3: 128kbits/s