Digital Signal Processing

ESD-5 & IV-5 (elektro), E24

11. Practical Implementation of the DFT – the Short-Time Fourier Transform

Assoc. Prof. Peter Koch, AAU

The outline of today's lecture

- A short recap of the previous lectures on DFT.

- Some considerations on practical calculation of the DFT
  - Doing DFT on an infinite length sequence
  - Spectral resolution
  - The impact of the window function

- The Short Time Fourier Transformation
  - Time varying signals
  - Simultaneous time-and-frequency analysis
  - Heisenberg's uncertainty relation, just briefly introduced…

Relation between the DFS and DTFT

Assume an infinite-length periodic sequence $\tilde{x}[n]$.

We have seen that the Discrete Fourier Series coefficients $\tilde{X}[k]$, corresponding to the sequence $\tilde{x}[n]$, can be found by sampling one period of the DTFT of the sequence x[n]. That is;

The DTFT $$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] \cdot e^{-j\omega n}$$

$$\tilde{X}[k] = X(e^{j\omega})|_{\omega=(2\pi/N)k} = X(e^{j(2\pi/N)k}) \quad 0 \le k \le N-1$$

which corresponds to sampling $X(z)$ in N equally spaced angles on the unit circle, where $X(z) = \mathcal{Z}\{x[n]\}$, $x[n]$ is an aperiodic sequence, i.e., an observable signal, and $e^{j\frac{2\pi}{N}}$ is the fundamental frequency.

…but how about the Discrete Fourier Transform (DFT)…?

AALBORG UNIVERSITY
DENMARK

Relation between DFS and DFT

If we can form the N-periodic sequence $\tilde{x}[n]$ by adding infinitely many time-shifted copies of the finite-length sequence $x[n]$ without ”time aliasing”, then the DFT can be considered identical to the DFS;

$$X[k] = \begin{cases} \tilde{X}[k], & 0 \leqslant k \leqslant N-1 \\ 0, & \text{otherwise} \end{cases}$$

…and therefore we can write the Discrete Fourier Transform as;

$$X[k] = \sum_{n=0}^{N-1} x[n] W_N^{kn} \quad \text{and}$$

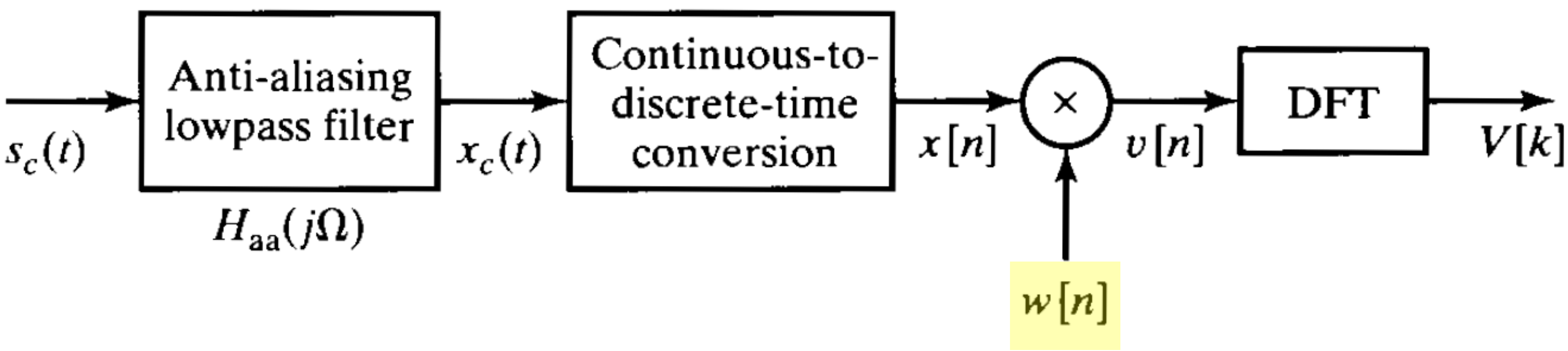
$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] W_N^{-kn}$$

Here it is important to notice that both $x[n]$ and $X[k]$ are considered equal to zero outside the interval $[0; N-1]$, despite both essentially are N-periodic.

Fourier Analysis using the DFT

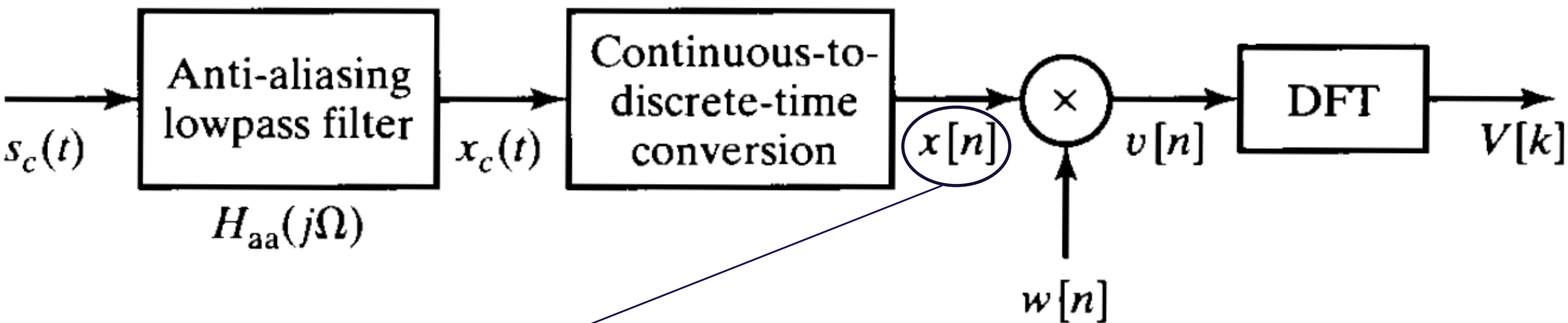
So, the DFT requires as input a finite lenght sequence which normally has to be derived from an infinite length sequence provided directly from the ADC.

In many cases this is accomplished by partitioning the signal using a finite duration "window", i.e., a sequence which is identically zero outside the interval $0..N-1$.



Now, let's recap once again what happens throughout this signal chain...

Due to sampling of $x_c(t)$, the Fourier Transform of the infinite sequence $x[n]$ is periodic in frequency



$$X(e^{j\omega}) = \frac{1}{T} \sum_{r=-\infty}^{\infty} X_c\left(j\frac{\omega}{T} + j\frac{2\pi r}{T} \right)$$

We now partition $x[n]$ into finite-length sequences by repeated multiplication with a window function $w[n]$.

But what is the consequence in the frequency domain of this multiplication..??

Well, we already discussed such topics when we addressed FIR filter design, but let's briefly recap…

Typical window functions

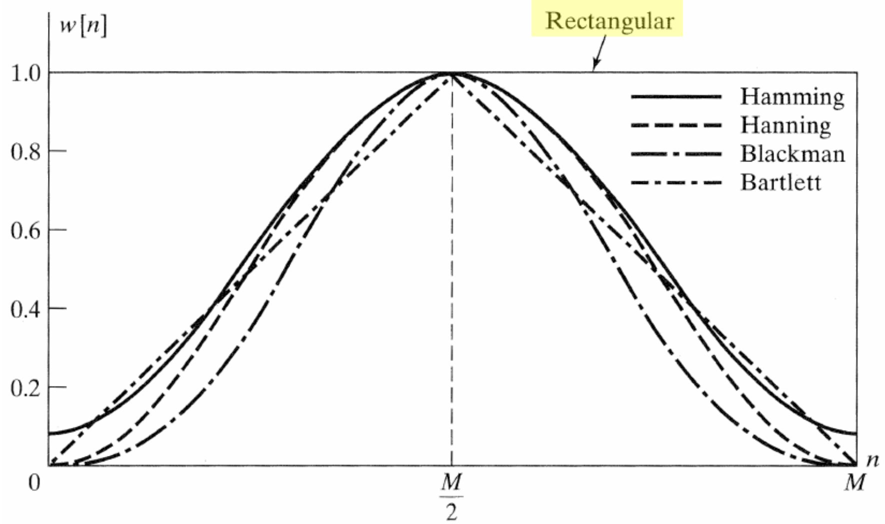
Normally, we would think of the Rectangular Window as being the most "obvious" function used when truncating a sequence.
The problem is however, that at the edges of the window, we have discontinuities which may impact negatively the overall performance of the DFT analysis – so also in terms of spectral analysis, we need to carefully select a proper window…

Hanning $(\alpha=0.5)$
& Hamming $(\alpha=0.46)$:

$$w[n]=\begin{cases}\alpha-\alpha\cos(2\pi n/M), & 0\le n\le M\\ 0, & \text{otherwise}\end{cases}$$

Blackman:

$$w[n]=\begin{cases}0.42-0.5\cos(2\pi n/M)+0.08\cos(4\pi n/M), & 0\le n\le M\\ 0, & \text{otherwise}\end{cases}$$



What happens when we multiplying with the window function..??

From the Fourier Transform theorem pair no. 7 on p. 60 in O&S 3'rd ed.;

$$7. \quad x[n]y[n] \longleftrightarrow \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\theta})Y(e^{j(\omega-\theta)})d\theta$$

So, one might think that limiting $x[n]$ to a finite duration sequence is "just a matter" of preparing the sequence for being suitable as input to the DFT, but the fact is that multiplication (significantly) impact the spectral analysis we are conducting.

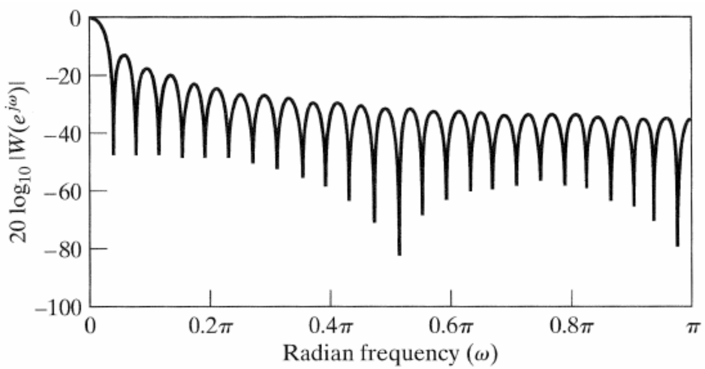
Consequently, in the frequency domain we are convolving the FT of the input sequence $x[n]$ with the FT of the window function – and to be precisely, we conduct a periodic convolution.

For this reason, it is important to study the FT of the various window functions…

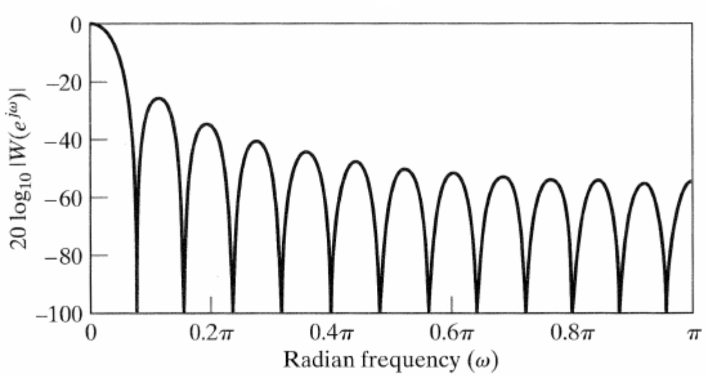
AALBORG UNIVERSITY
DENMARK

Amplitude response of the window functions

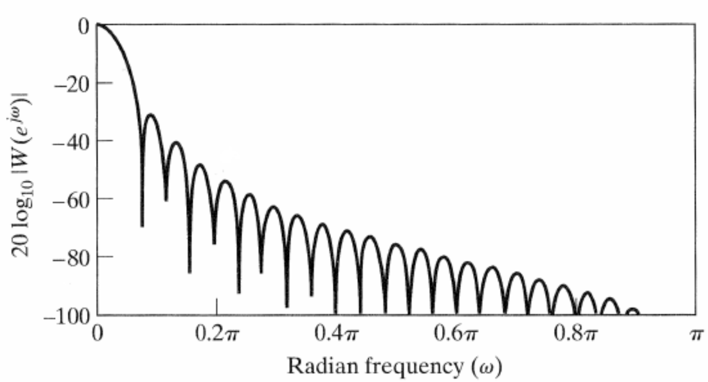
Rectangular



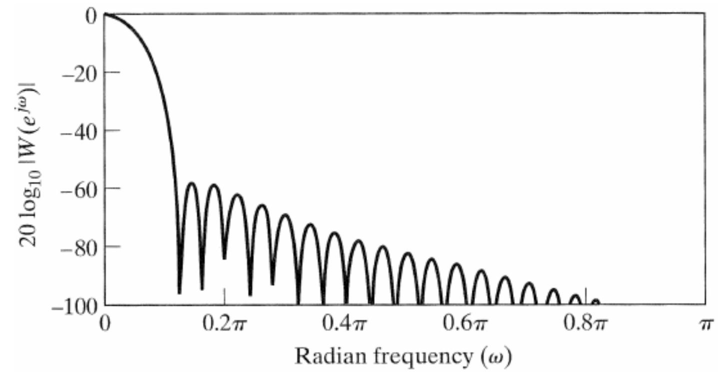
Bartlett



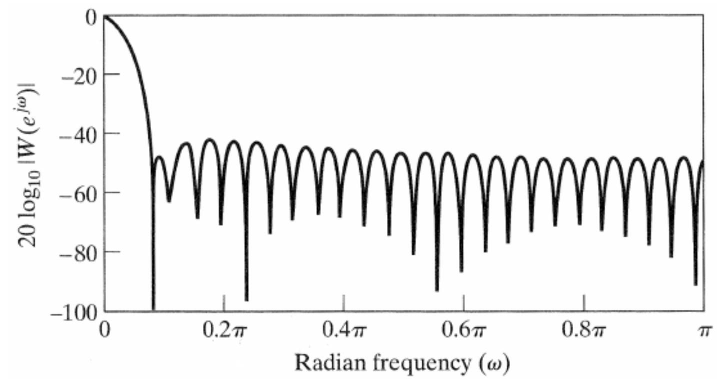
Hanning



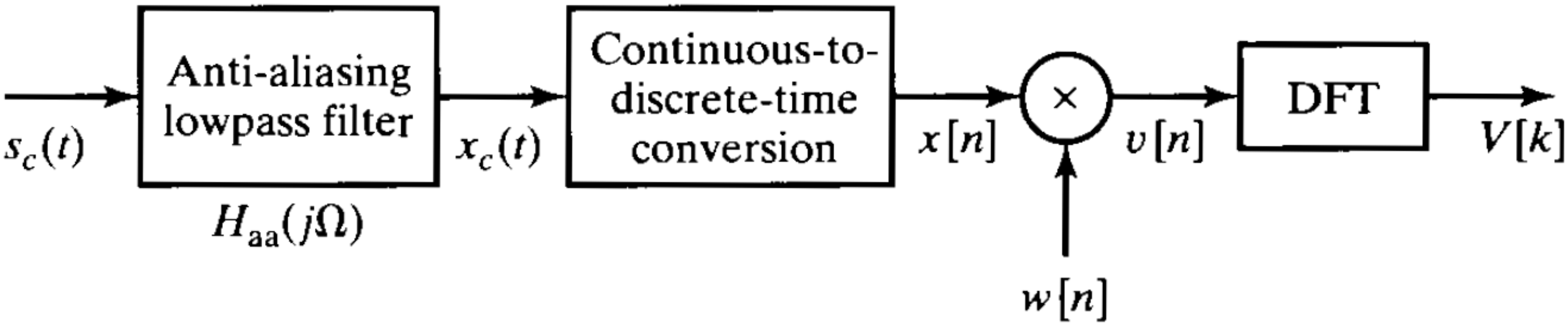
Blackman



Hamming



Type of window	Peak side lobe	Width of main lobe
Rectangular	-13 dB	$4\pi/(M+1)$
Bartlett	-25 dB	$8\pi/M$
Hanning	-31 dB	$8\pi/M$
Blackman	-57 dB	$12\pi/M$
Hamming	-41 dB	$8\pi/M$



So, according to the overall signal chain, the input of the DFT block, $v[n]$, has a spectral representation which is the periodic convolution between $X(e^{j\omega})$ and $W(e^{j\omega})$, i.e.,

$$V(e^{j\omega})=X(e^{j\omega}) * W(e^{j\omega})$$

Now, since, from a functional interpretation, the DFT block basically samples the spectrum on its input and presents it as a ”frequency discrete” representation on its output, we can state that

$$V[k] = V\left(e^{j\omega}\right)\Big|_{\omega=\frac{2\pi}{N}k}$$

Since $V[k]$, for $0 \leq k < N$, represents one period $[0; 2\pi[$, where 2π is the sample frequency, we can easily calculate ”the spectral resolution”…

Spectral resolution

Sampling the DTFT;

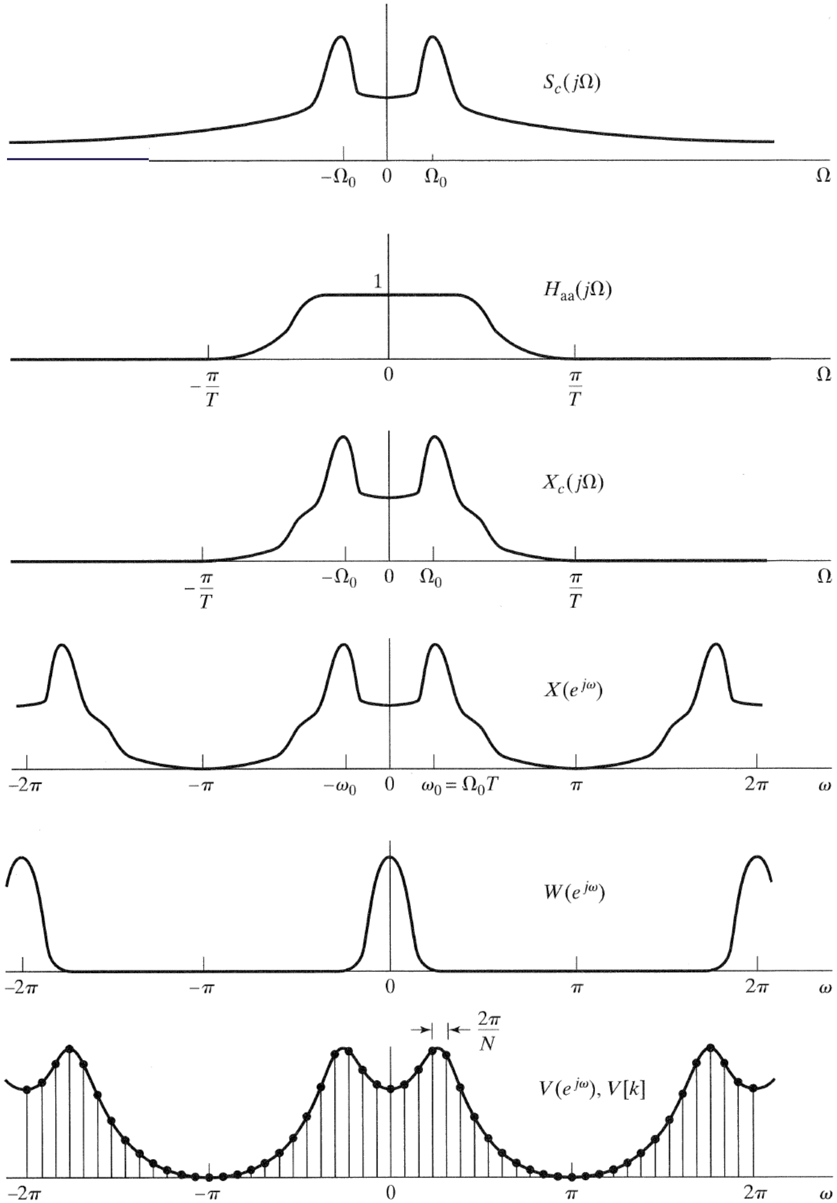
$$V[k] = V(e^{j\omega}) \quad \text{for} \quad \omega = \omega_k = \frac{2\pi}{N}k \quad 0 \leq k \leq N-1$$

The relation $\omega = T\Omega$ implies the $\omega_k = T\Omega_k = \frac{2\pi}{N}k$ and thus isolating Ω_k;

$$\boxed{\Omega_k = 2\pi f_k = \frac{2\pi}{NT}k \Rightarrow f_k = \frac{k}{NT} = k\frac{f_s}{N} = k \cdot \Delta f \quad 0 \leq k \leq N-1}$$

The important lesson learned here is that the spectral resolution (the smaller Δf, the better resolution) is proportional to the number of points N in the DFT, and inverse proportional to the sample frequency.

The overall picture - in the frequency domain



The spectrum of the ”wide band” input signal

The amplitude response of the anti-aliasing filter

The continuous-time input signal to the S/H circuit.

The discrete-time input signal

The amplitude response of the window function

The DTFT of the signal and the sampled version which is the DFT

The effect of windowing – an IMPORTANT example

In this example we consider a signal which is a sum of two sinusoids;

$$s_c(t) = A_0 \cos(\Omega_0 t + \theta_0) + A_1 \cos(\Omega_1 t + \theta_1)$$

This signal is now sampled in an ideal manner, i.e., no aliasing and no variable quantization effects;

$$x[n] = A_0 \cos(\omega_0 n + \theta_0) + A_1 \cos(\omega_1 n + \theta_1)$$

$$\text{where } \omega_0 = \Omega_0 T \text{ og } \omega_1 = \Omega_1 T$$

The sequence is now multiplied with a window function $w[n]$, $0 \leq n \leq N-1$

$$v[n] = A_0 w[n] \cos(\omega_0 n + \theta_0) + A_1 w[n] \cos(\omega_1 n + \theta_1)$$

Now we apply 1) the cosine addition formula;

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$$

and next 2) Euler's formulas for cosine and sine;

$$\cos(x) = (e^{jx} + e^{-jx})/2$$

$$\sin(x) = (e^{jx} - e^{-jx})/2j$$

AALBORG UNIVERSITY
DENMARK

The effect of windowing – an example

$$v[n]=A_0 w[n]\cos(\omega_0 n+\theta_0)+A_1 w[n]\cos(\omega_1 n+\theta_1)$$

$$v[n]=\frac{A_0}{2}w[n]e^{j\theta_0}e^{j\omega_0 n}+\frac{A_0}{2}w[n]e^{-j\theta_0}e^{-j\omega_0 n}$$

$$+\frac{A_1}{2}w[n]e^{j\theta_1}e^{j\omega_1 n}+\frac{A_1}{2}w[n]e^{-j\theta_1}e^{-j\omega_1 n}$$

Now, let's use the Fourier Transform theorem pair no. 3, O&S 3rd ed., p. 60;

$$3.\ e^{j\omega_0 n}x[n] \longleftrightarrow X(e^{j(\omega-\omega_0)})$$

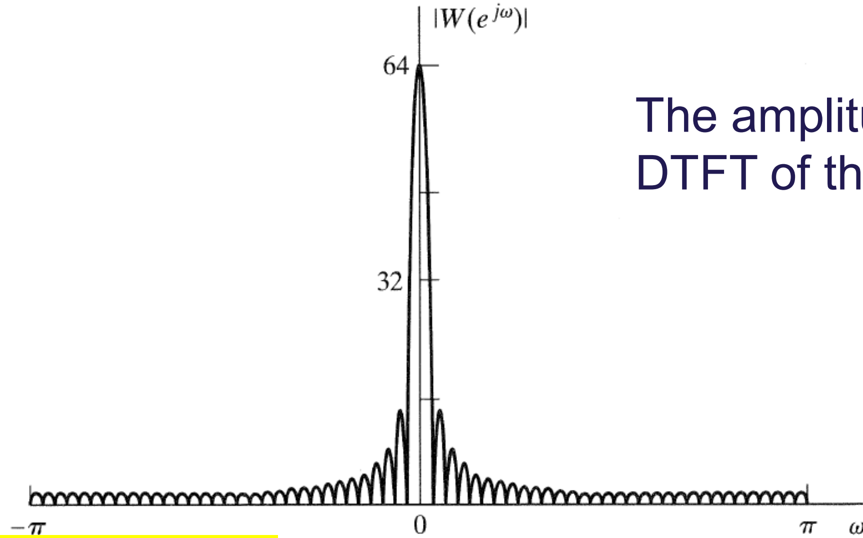
$$V(e^{j\omega})=\frac{A_0}{2}e^{j\theta_0}W(e^{j(\omega-\omega_0)})+\frac{A_0}{2}e^{-j\theta_0}W(e^{j(\omega+\omega_0)})$$

$$+\frac{A_1}{2}e^{j\theta_1}W(e^{j(\omega-\omega_1)})+\frac{A_1}{2}e^{-j\theta_1}W(e^{j(\omega+\omega_1)})$$

4 terms…!!
…and each term is
the FT of the window
- ampl and phase…!!!!

Rectangular window with length 64.
Signal amplitudes of $A_0 = 1$ og $A_1 = 0{,}75$

$\omega_0 = 2\pi/6$
$\omega_1 = 2\pi/3$

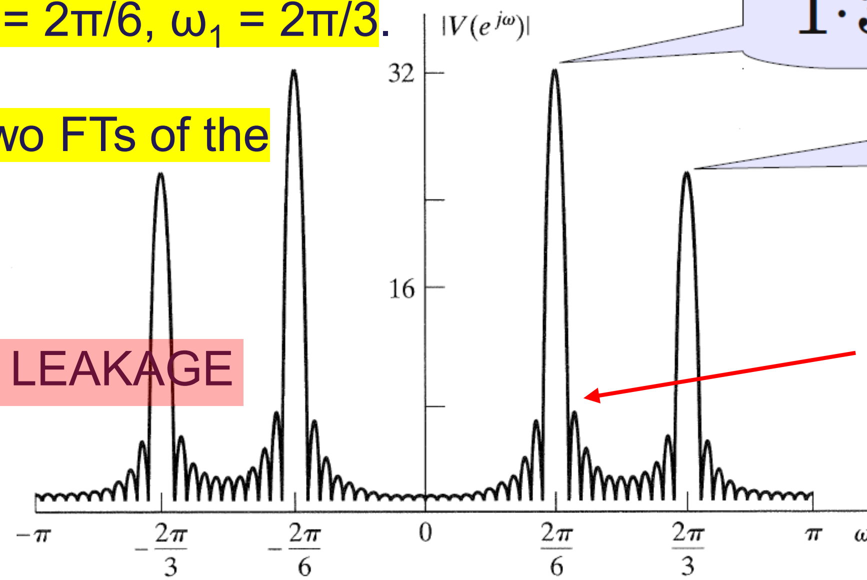


The amplitude (modulus) of the DTFT of the rectangular window

Guess you would have expected two impulses at $\omega_0 = 2\pi/6$, $\omega_1 = 2\pi/3$.

But in fact we get two FTs of the window function…

This effect is called LEAKAGE

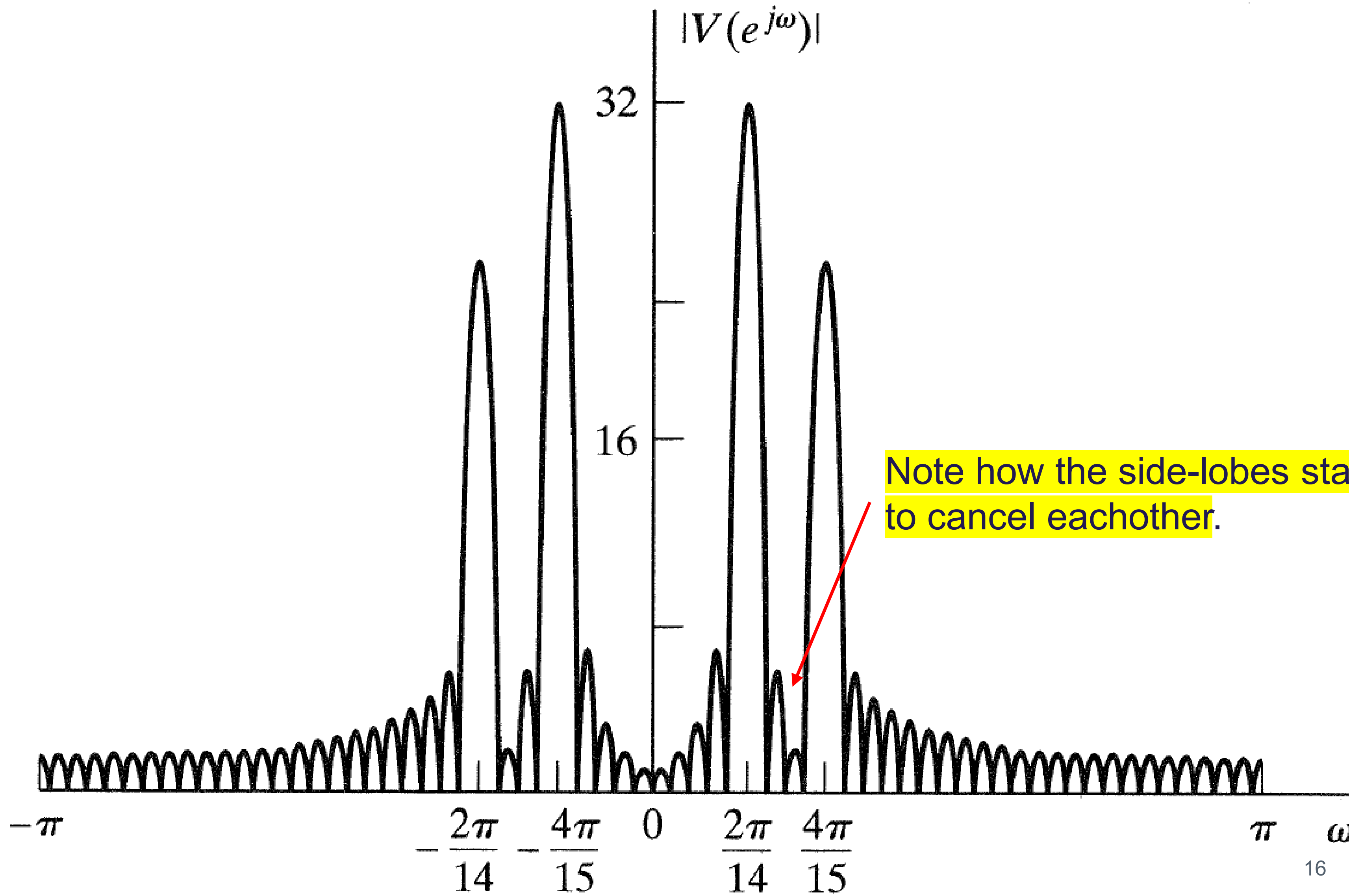


$1 \cdot 32$ (one half)

$0.75 \cdot 32$

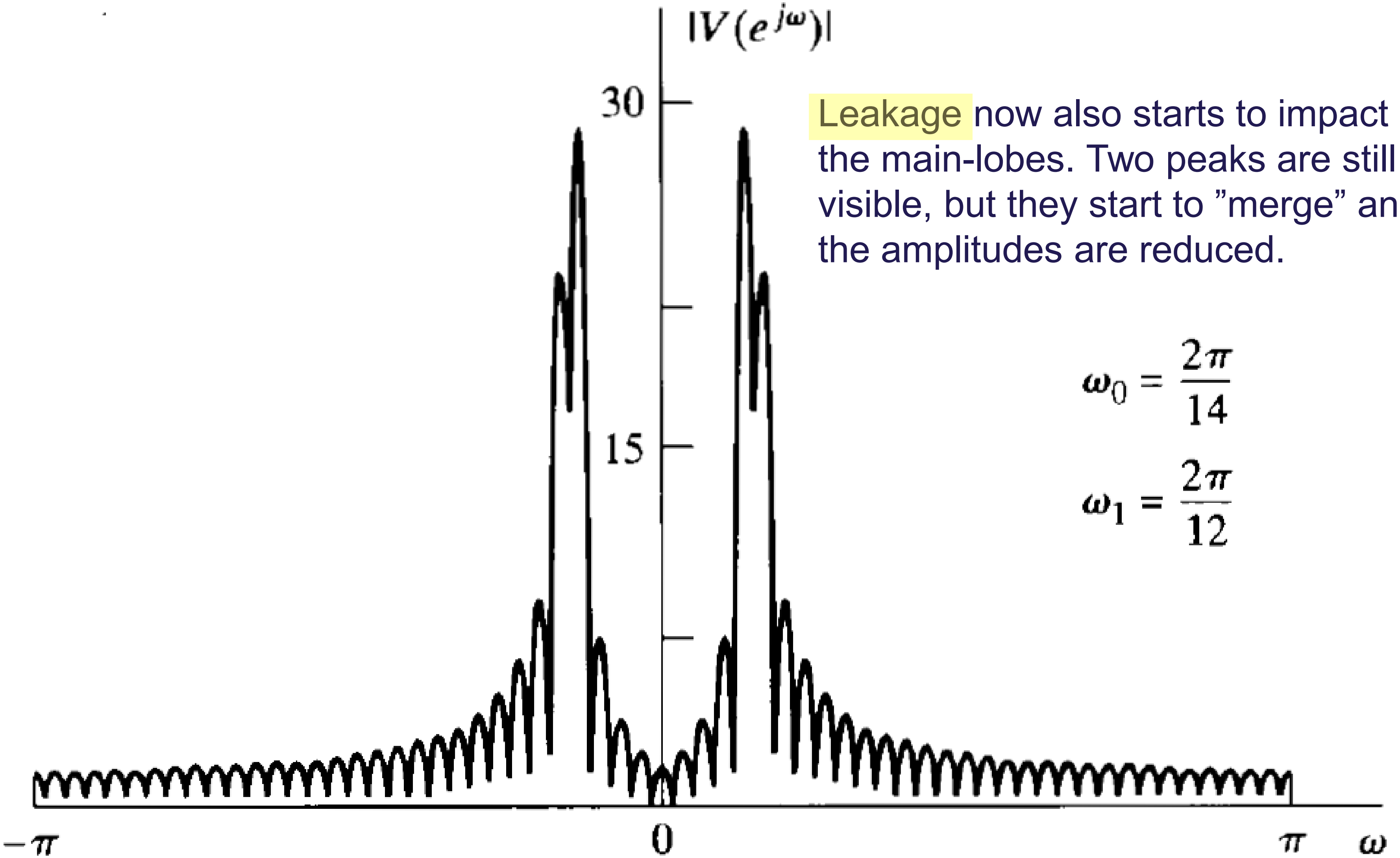
Note that the side lobes are down approx. 13 dB – just as expected…

Now, let's try to move the two cosine signals a bit closer to eachother



Note how the side-lobes start to cancel eachother.

…and even closer



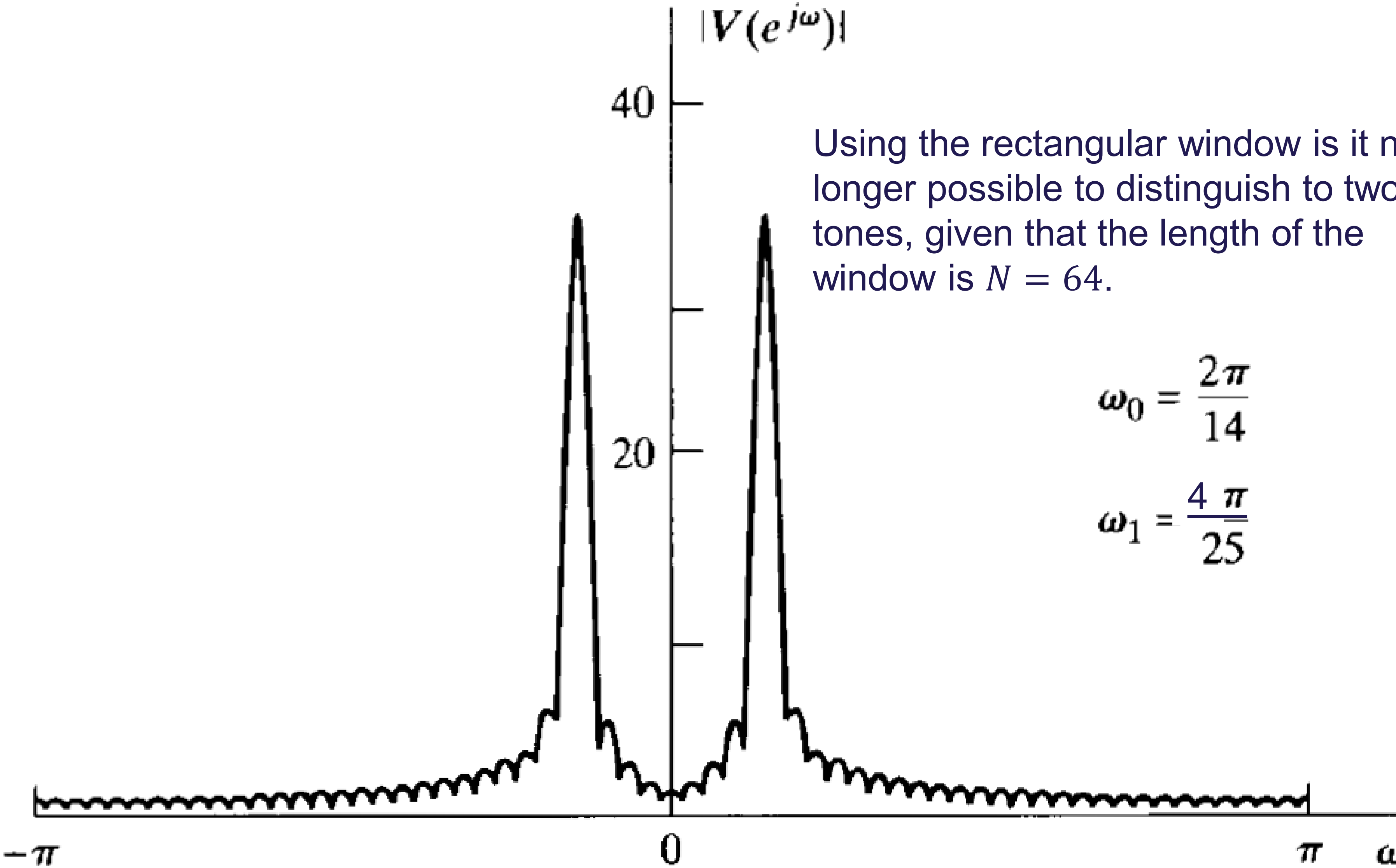
(d)

Leakage now also starts to impact the main-lobes. Two peaks are still visible, but they start to ”merge” and the amplitudes are reduced.

$$\omega_0 = \frac{2\pi}{14}$$

$$\omega_1 = \frac{2\pi}{12}$$

…and finally very close together.



(e)

Using the rectangular window is it no longer possible to distinguish to two tones, given that the length of the window is $N = 64$.

$$\omega_0 = \frac{2\pi}{14}$$

$$\omega_1 = \frac{4\pi}{25}$$

Some important considerations

- The **effective frequency resolution** depends on the width of the window-function main-lobe
- The **amount of leakage** depends on the ratio between the main-lobe amplitude and the side-lobe amplitudes
- The **rectangular window** gives the highest possible frequency resolution, but at the same time it also generates the larges amount of leakage

IMPORTANT…!!

AALBORG UNIVERSITY
DENMARK

The DFT is a ”spectral sampling” of the DTFT
What happens if we don’t sample in the right frequencies..??

Discrete time frequencies

$$\omega_k = 2\pi k / N$$

corresponds to the continuous time frequencies:

$$\Omega_k = (2\pi k)/(NT)$$

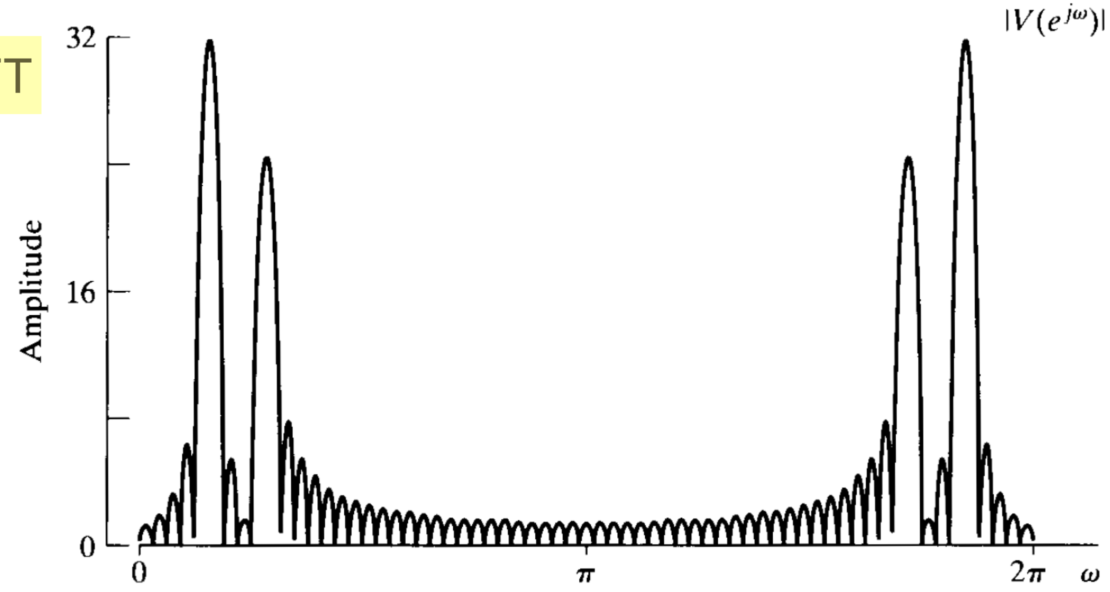
Example of a sampled (2-tone) signal, truncated using a rectangular window with length 64:

$$v[n] = \begin{cases} 1\cos\left(\frac{2\pi}{14}n\right) + 0{,}75\cos\left(\frac{4\pi}{15}n\right), & 0 \le n \le 63 \\ 0, \text{ otherwise} \end{cases}$$

The 2^{nd} experíment shown on slide no. 16.

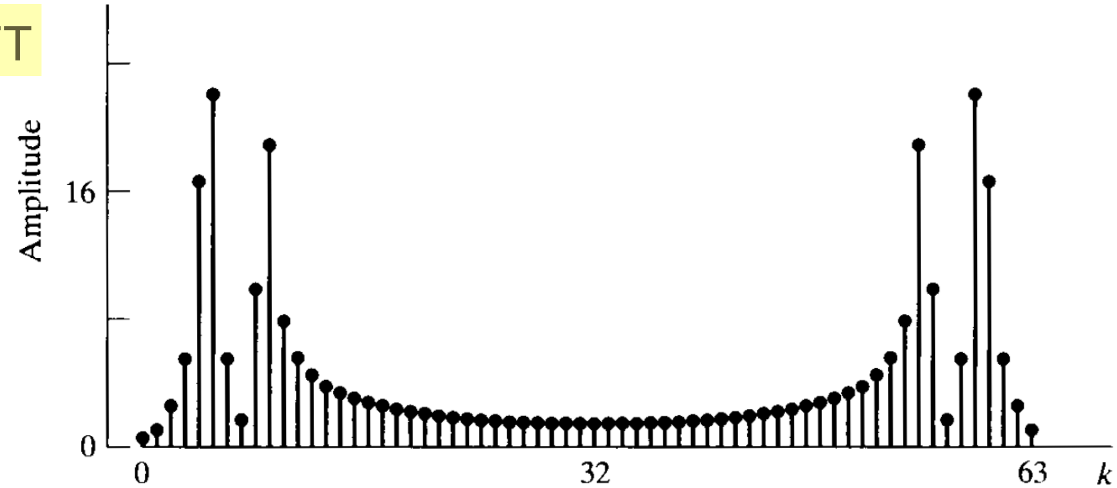
What we would like to have, and what we actually get…
DTFT vs. DFT if the frequencies don’t match

DTFT



Note the difference in amplitude at the two peaks between the DTFT and the DFT.

DFT

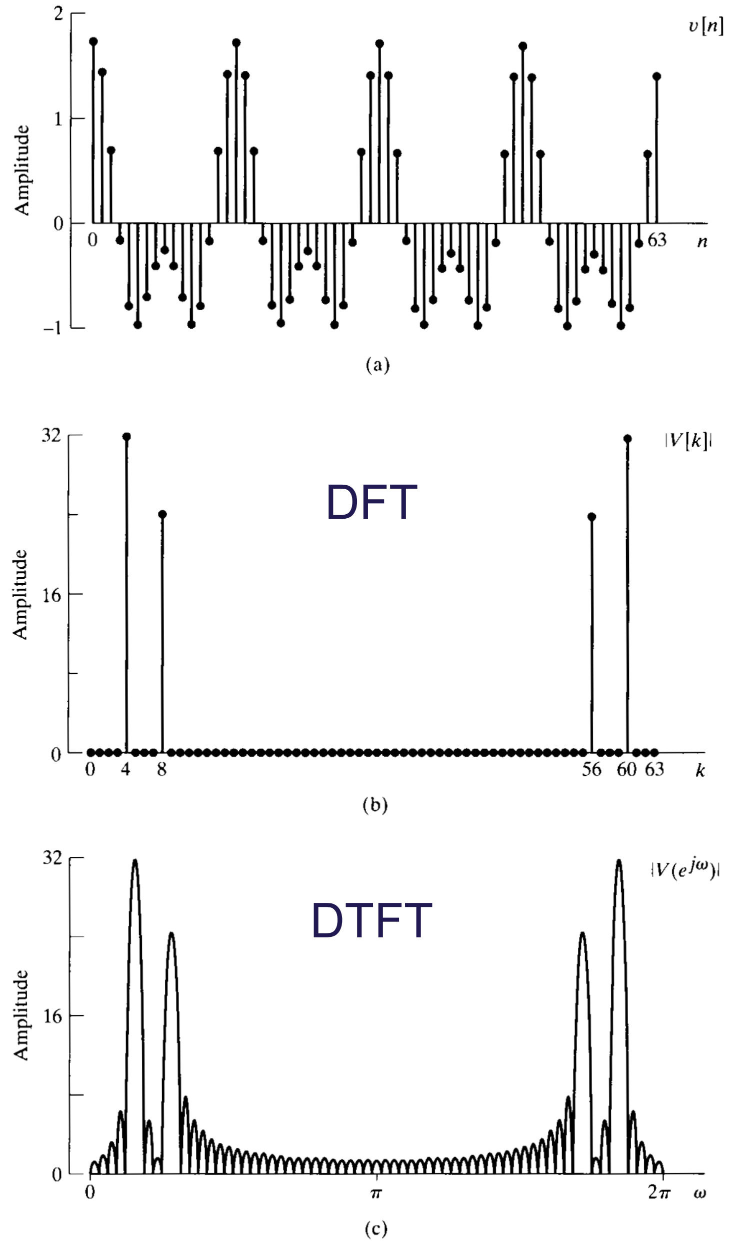


The frequency resolution is too low to capture ”all” the details…

Or, our sampled frequencies are not ”correctly” located.

Now, we move the frequencies a bit…

$$v[n] = \begin{cases} \cos\left(\frac{2\pi}{16}n\right) + 0.75\cos\left(\frac{2\pi}{8}n\right), & 0 \le n \le 63, \\ 0, & \text{otherwise,} \end{cases}$$



Discrete Fourier analysis of the sum of two sinusoids for a case in which the Fourier transform is zero at all DFT frequencies except those corresponding to the frequencies of the two sinusoidal components. (a) Windowed signal. (b) Magnitude of DFT. (c) Magnitude of discrete-time Fourier transform ($|V(e^{j\omega})|$).

For this particular two-tone sequence and the given window length, $N = 64$, we have a perfect match, and thus we get a correct sampling of the two tones…

In any real-life situation, that is extremely unlikely to happen…!

Furthermore, the ”perfect sampling” also now hides the impact from the window function, i.e., we sample all the zero-crossings…!

AALBORG UNIVERSITY
DENMARK

It's now time for a…

…before we discuss the Short-Time Fourier Transform (STFT)

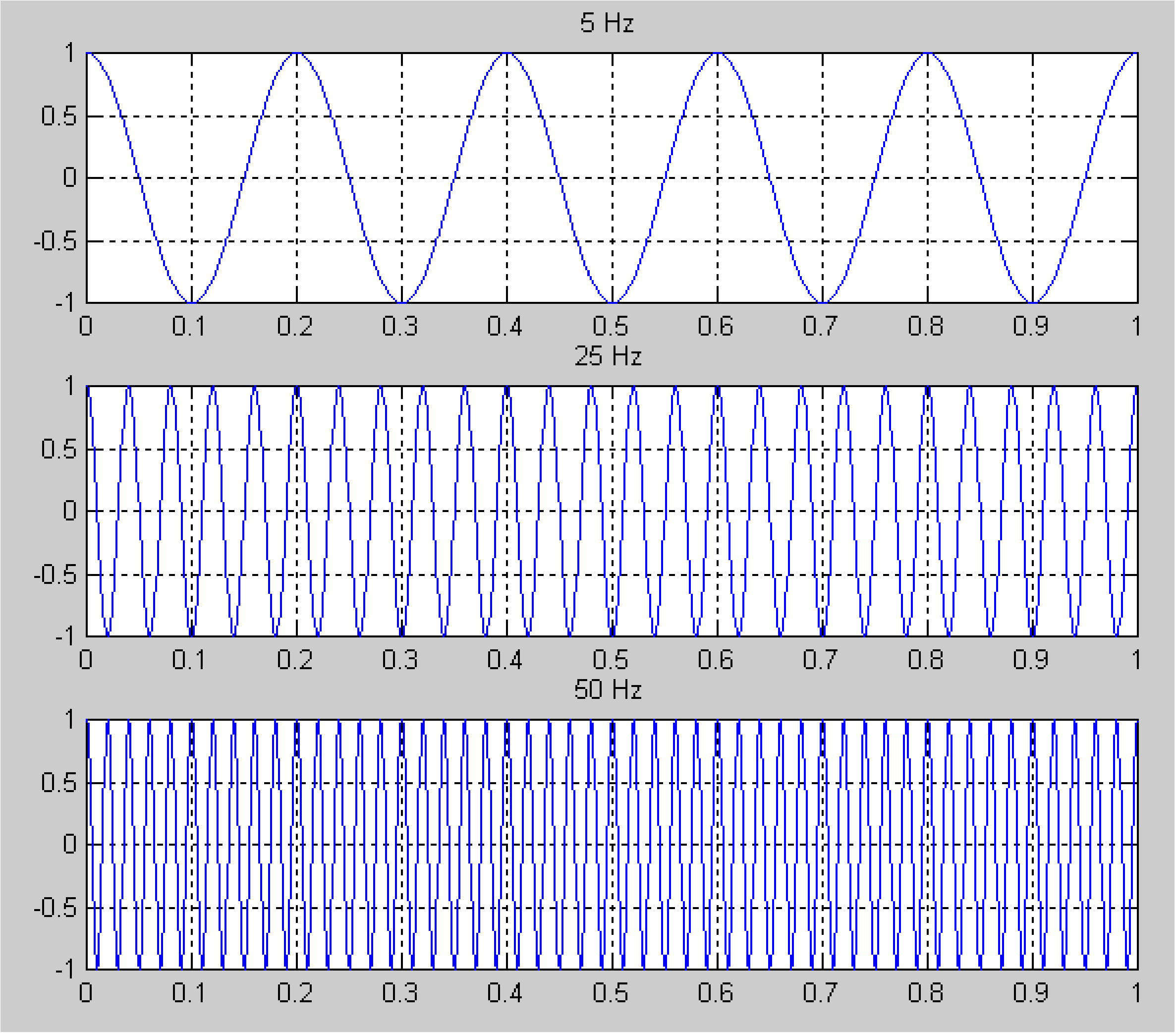
AALBORG UNIVERSITY
DENMARK

A motivating example – simple sinusoids

$$f_1(t) = \cos(2\pi \cdot 5 \cdot t)$$

$$f_2(t) = \cos(2\pi \cdot 25 \cdot t)$$

$$f_3(t) = \cos(2\pi \cdot 50 \cdot t)$$

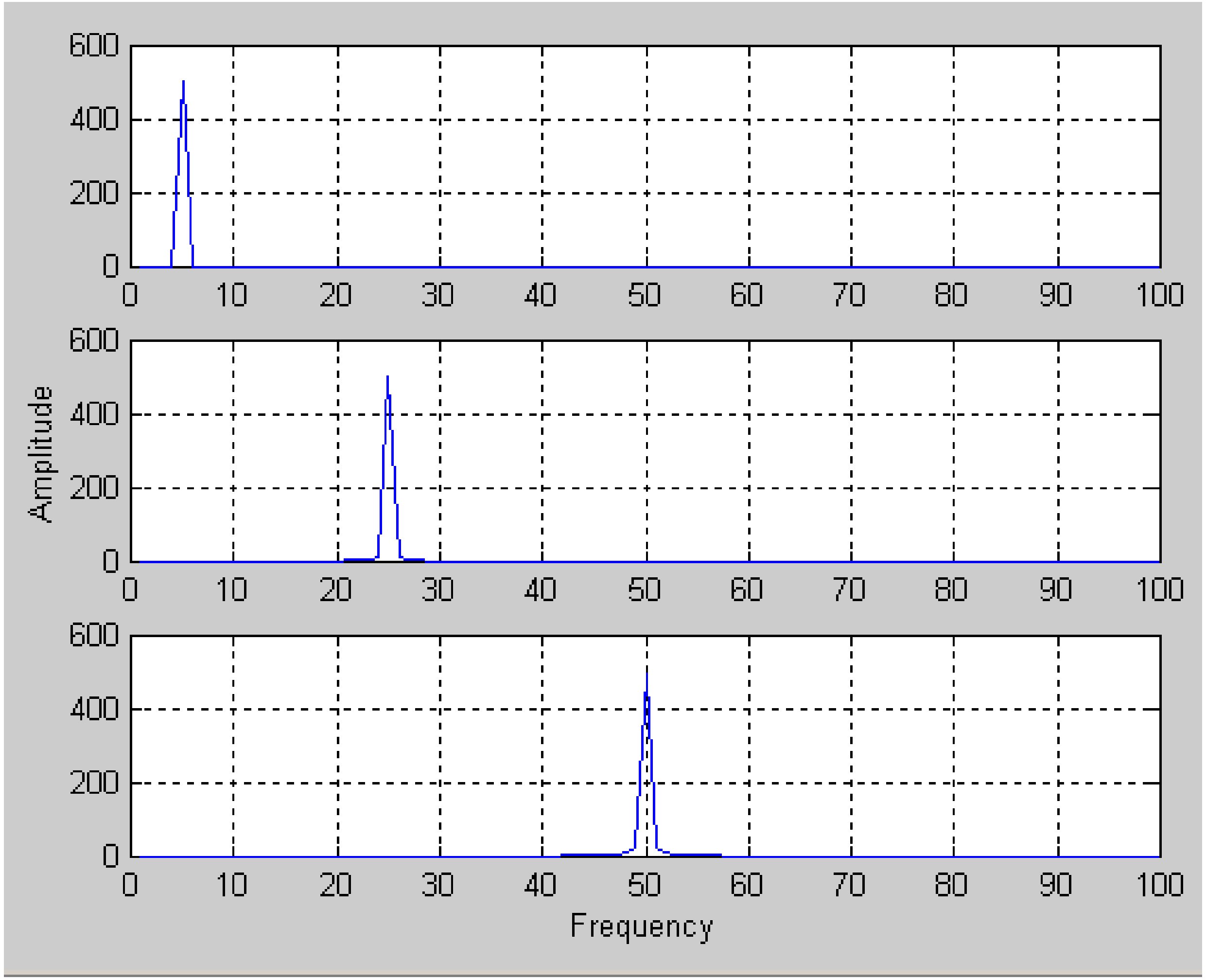


The Amplitude Responses

$F_1(\Omega/2\pi)$

$F_2(\Omega/2\pi)$

$F_3(\Omega/2\pi)$

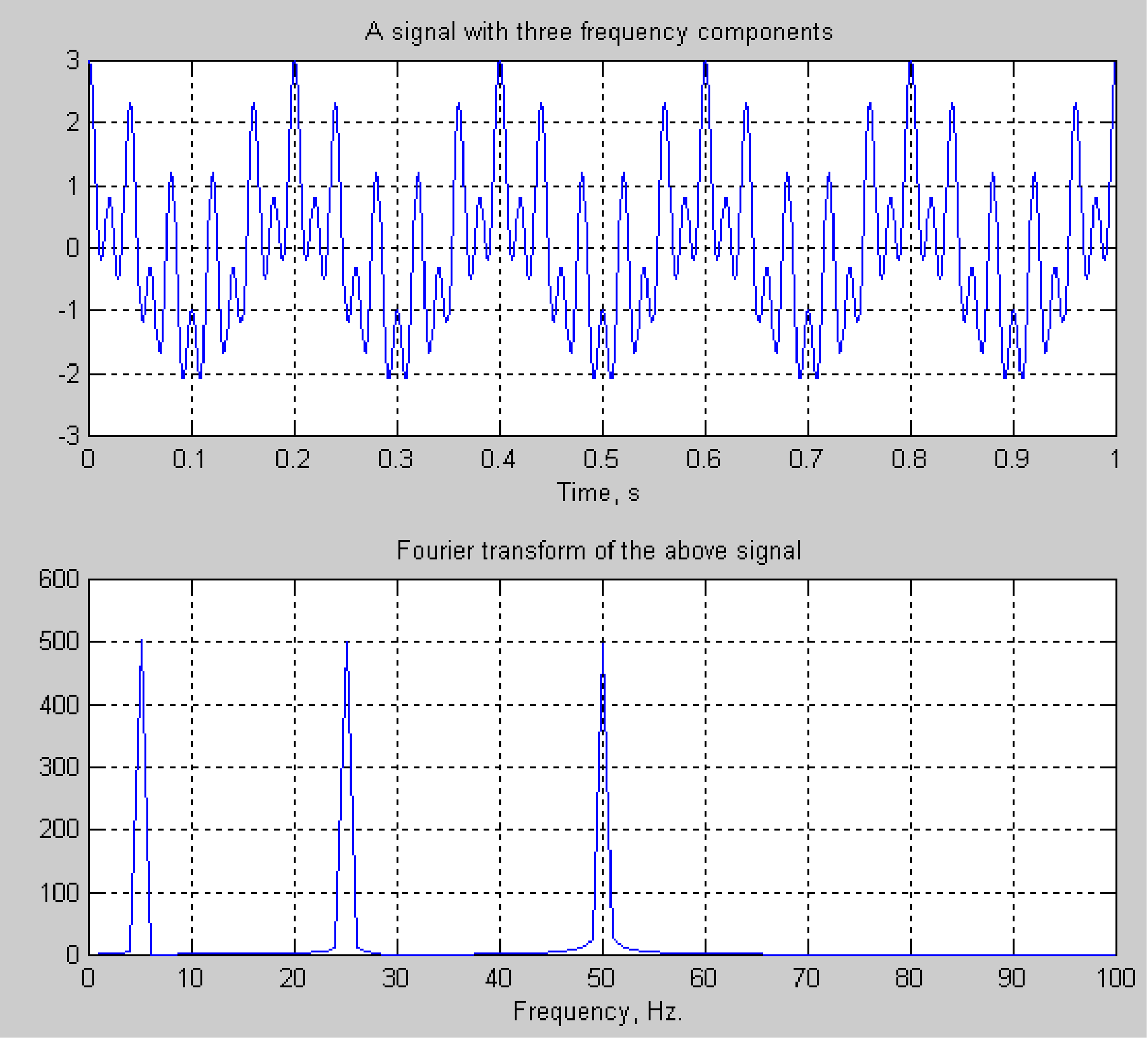


AALBORG UNIVERSITY
DENMARK

Now, let's add the three sinusoids

$$f_4(t) = \cos(2\pi \cdot 5 \cdot t) + \cos(2\pi \cdot 25 \cdot t) + \cos(2\pi \cdot 50 \cdot t)$$

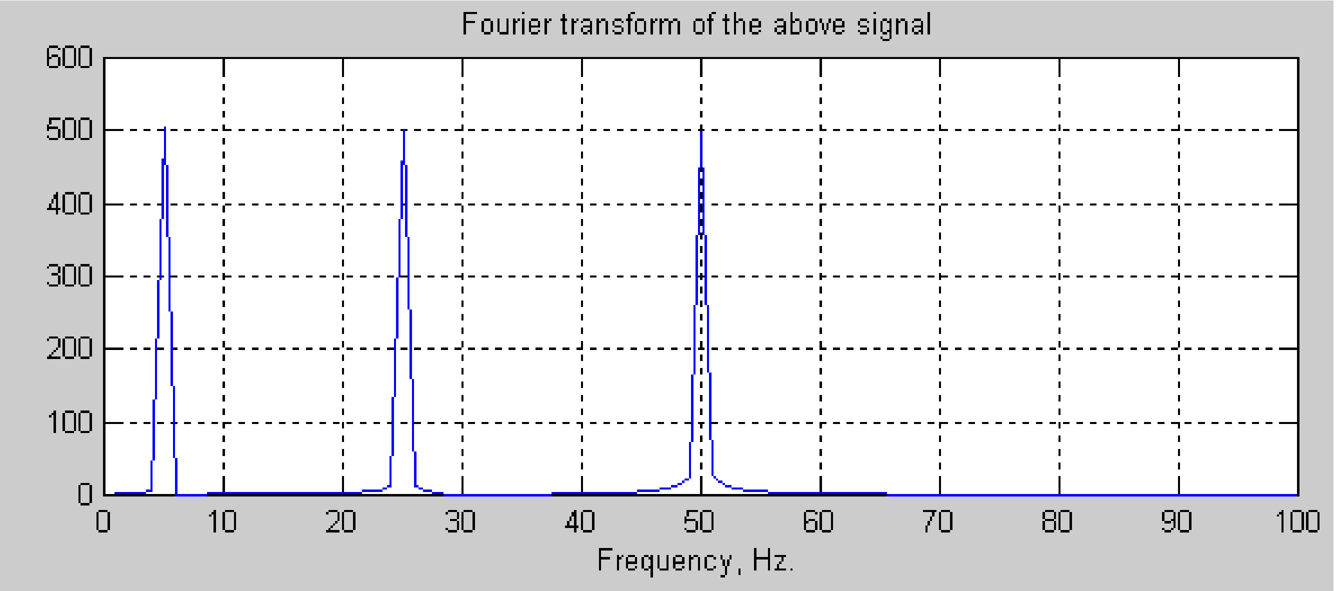
$F_4(\Omega/2\pi)$



AALBORG UNIVERSITY
DENMARK

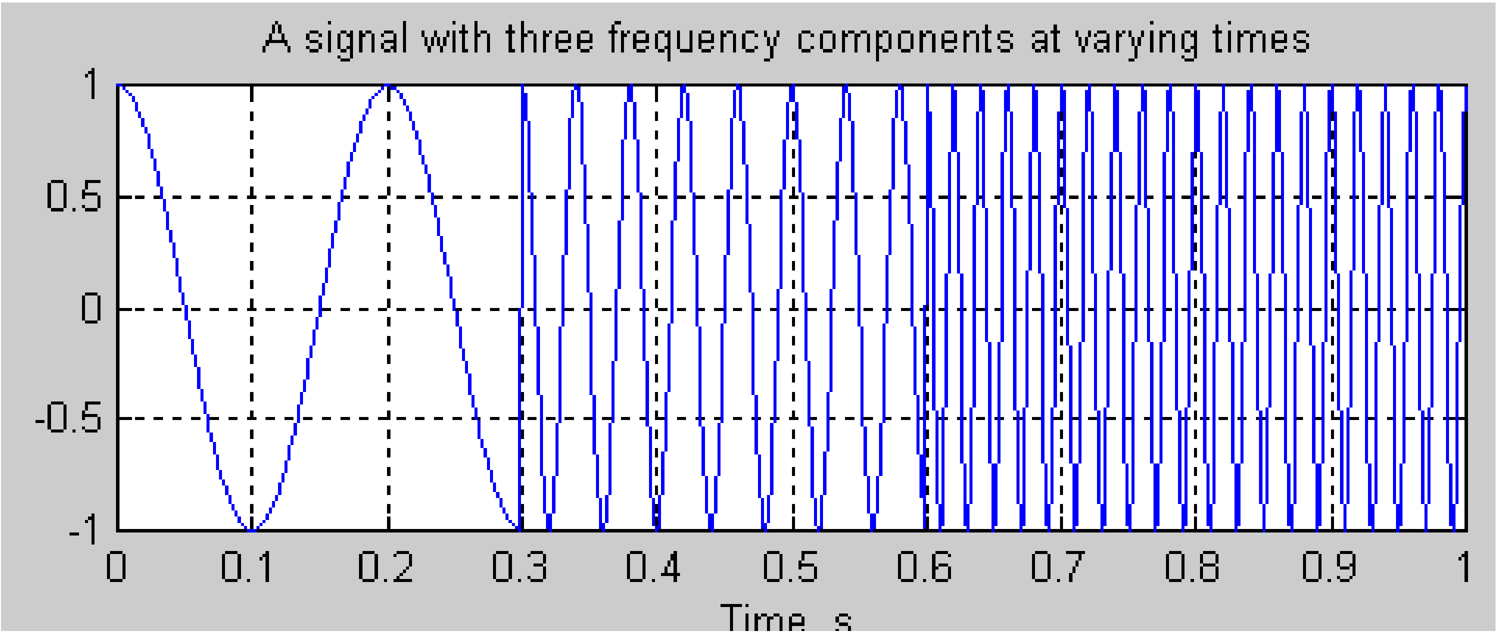
The sum of sinusoids is a Time-invariant signal

The three frequency components, are present at all times!



No matter when we perform the Fourier Analysis, we get the same result – the spectra is therefore also time-invariant.

Now, let's append the three sinusoids such that they occur distinctive in time

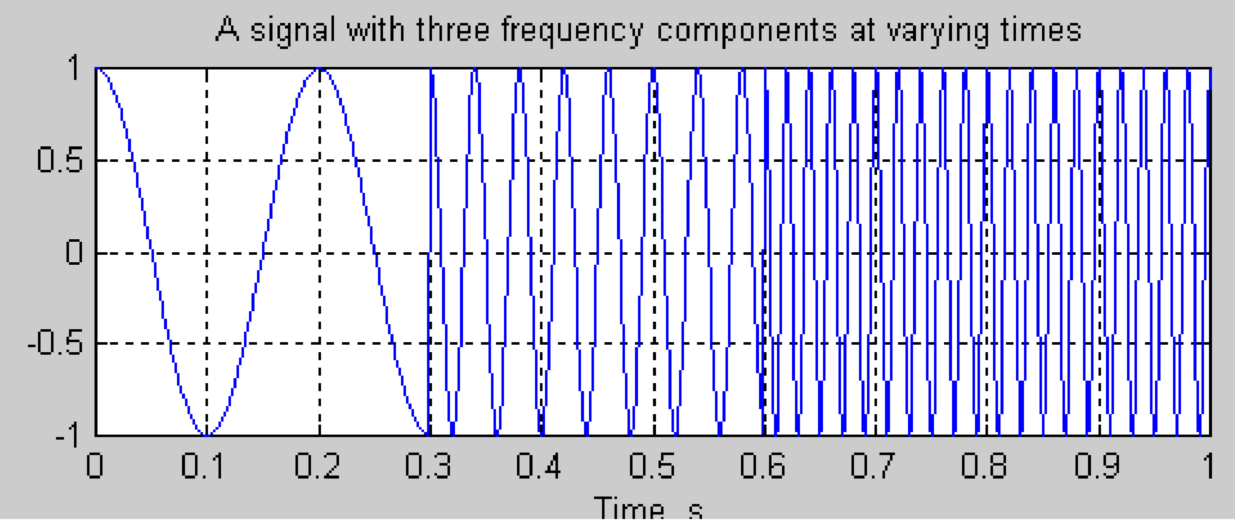


Depending on WHEN we perform the Fourier Analysis, we will see different results…

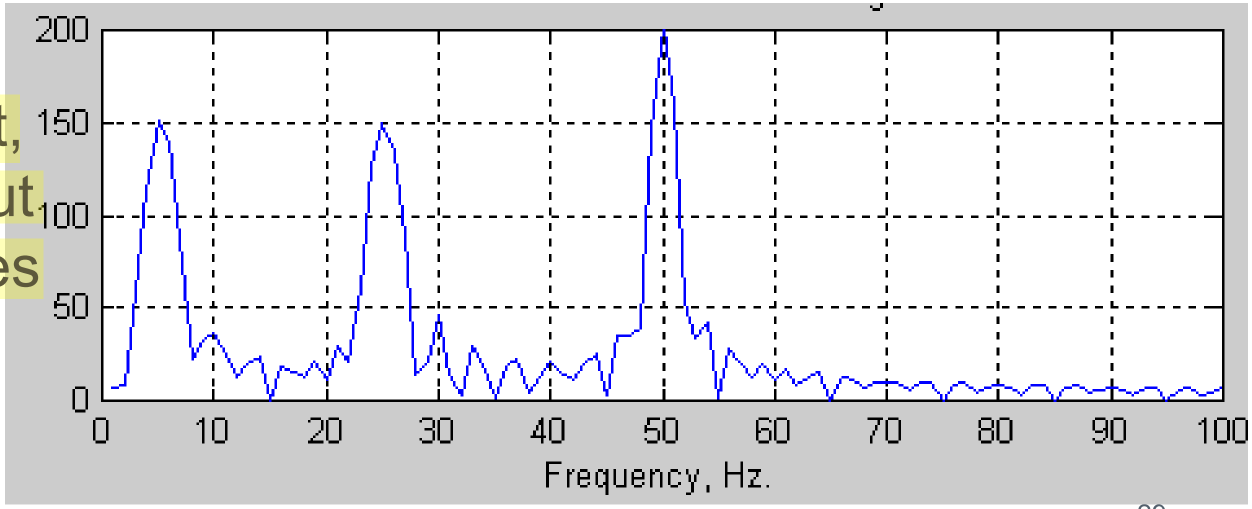
AALBORG UNIVERSITY
DENMARK

Spectral analysis of Time-varying signals

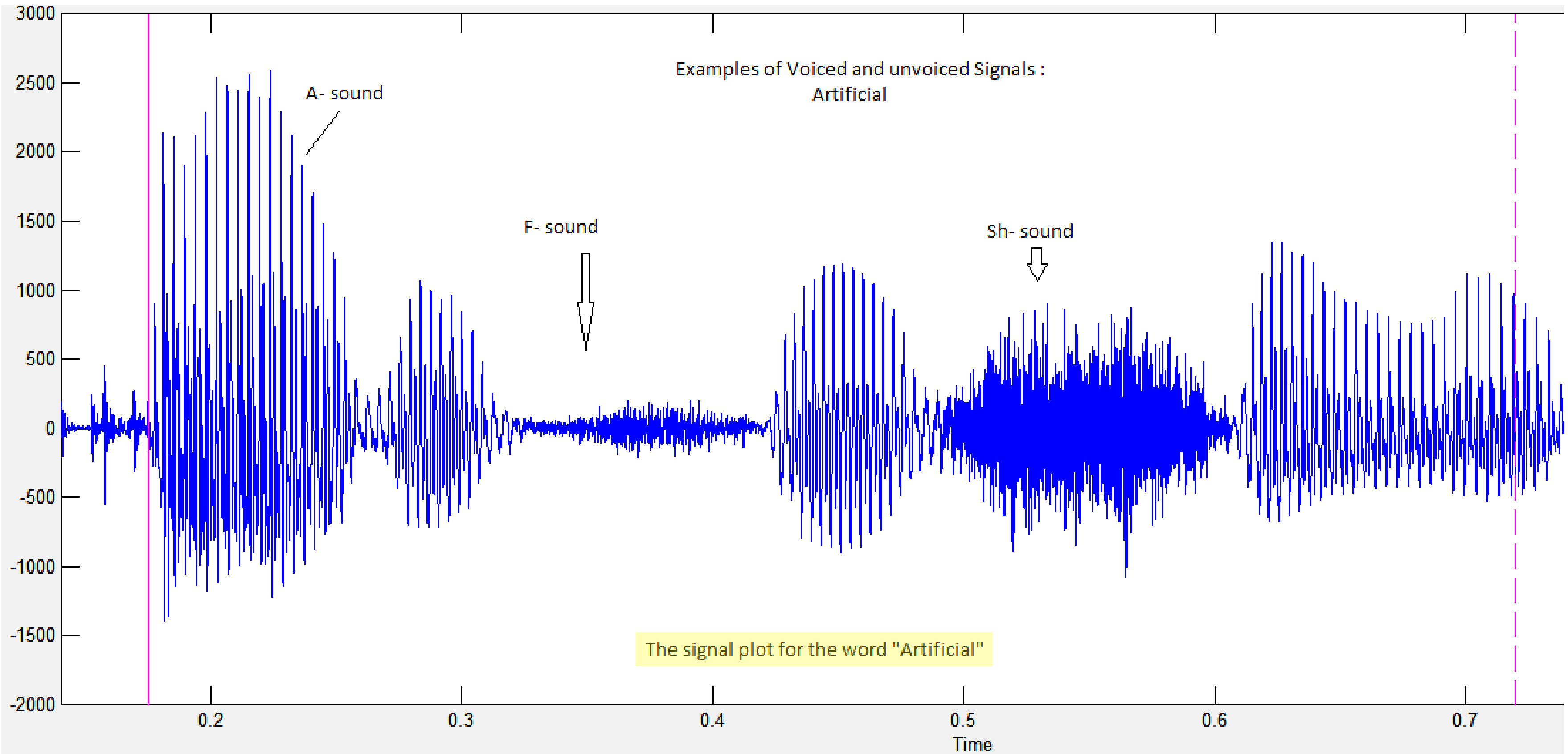
The three frequency components are NOT present at all times!



Perfect knowledge of which frequencies exist, but no information about where these frequencies are located in time!



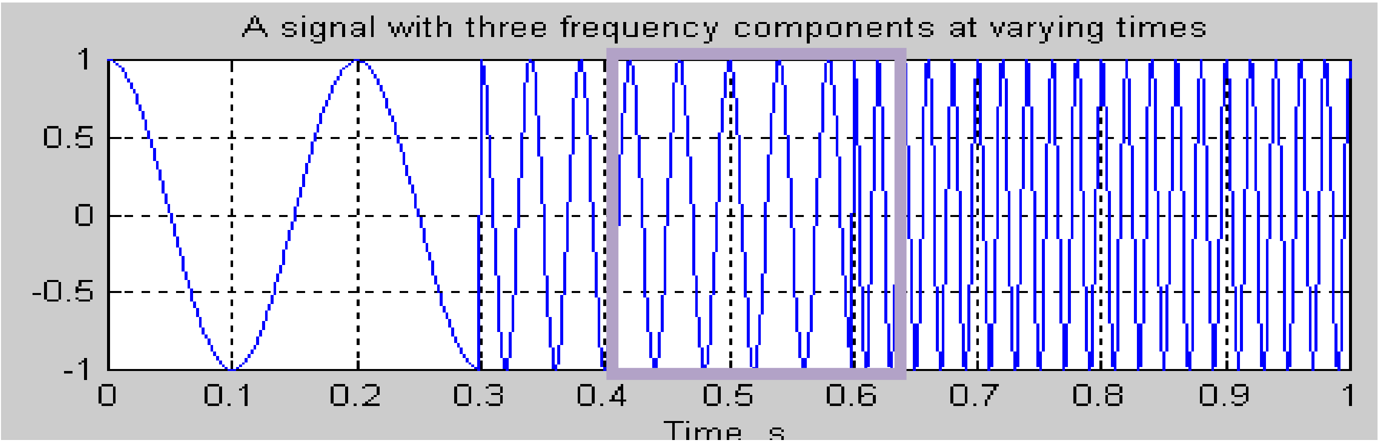
This is the normal situation for real-life signals…



AALBORG UNIVERSITY
DENMARK

Therefore, we need (yet) another Fourier Transform

- Segment the signal into short time intervals (i.e., short enough for the signal to be considered time-invariant) and then calculate the FT of each segment.
- Each FT provides the spectral information of a separate time-slice of the signal, thus providing simultaneous time and frequency information.



Short-Time Fourier Transform, STFT
- also known as the Time Dependent FT

(1) Choose a window function of finite length

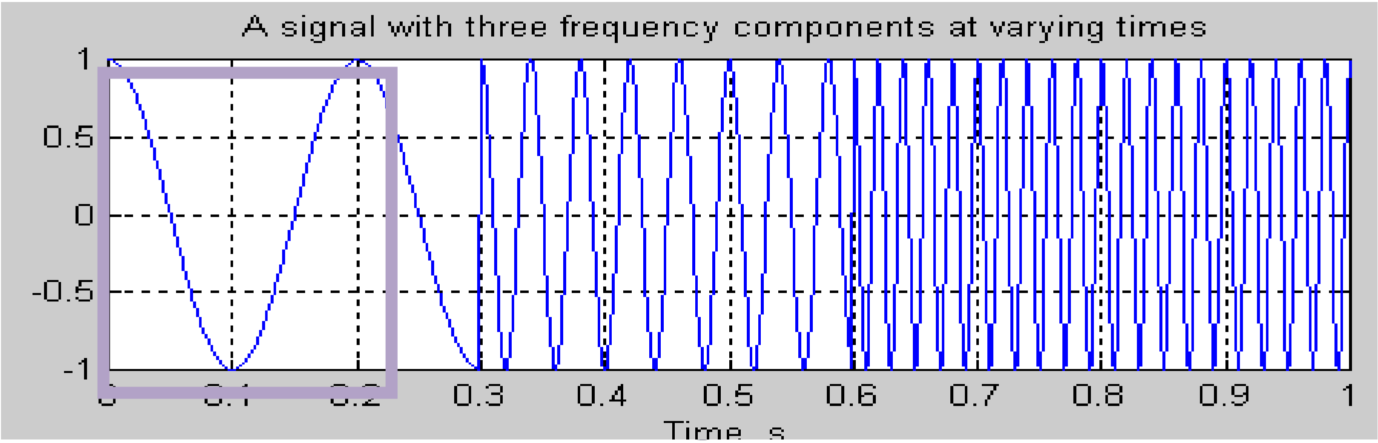
(2) Place the window on top of the signal at $t = 0$

(3) Truncate the signal using this window

(4) Compute the DTFT of the truncated signal

(5) Incrementally slide the window to the right

(6) Go to step 3, until window reaches the end of the signal



Definition of STFT

The STFT is a 2D function…!!!!

STFT of $x[n]$ here defined as the DTFT for each segment starting at time n

Signal to be analyzed - represented as the segment starting at time n.

The usual complex exponential function

$$X[n, \lambda) = \sum_{m=-\infty}^{\infty} x[n+m]w[m]e^{-j\lambda m}$$

Discrete Time parameter

Continuous Frequency parameter

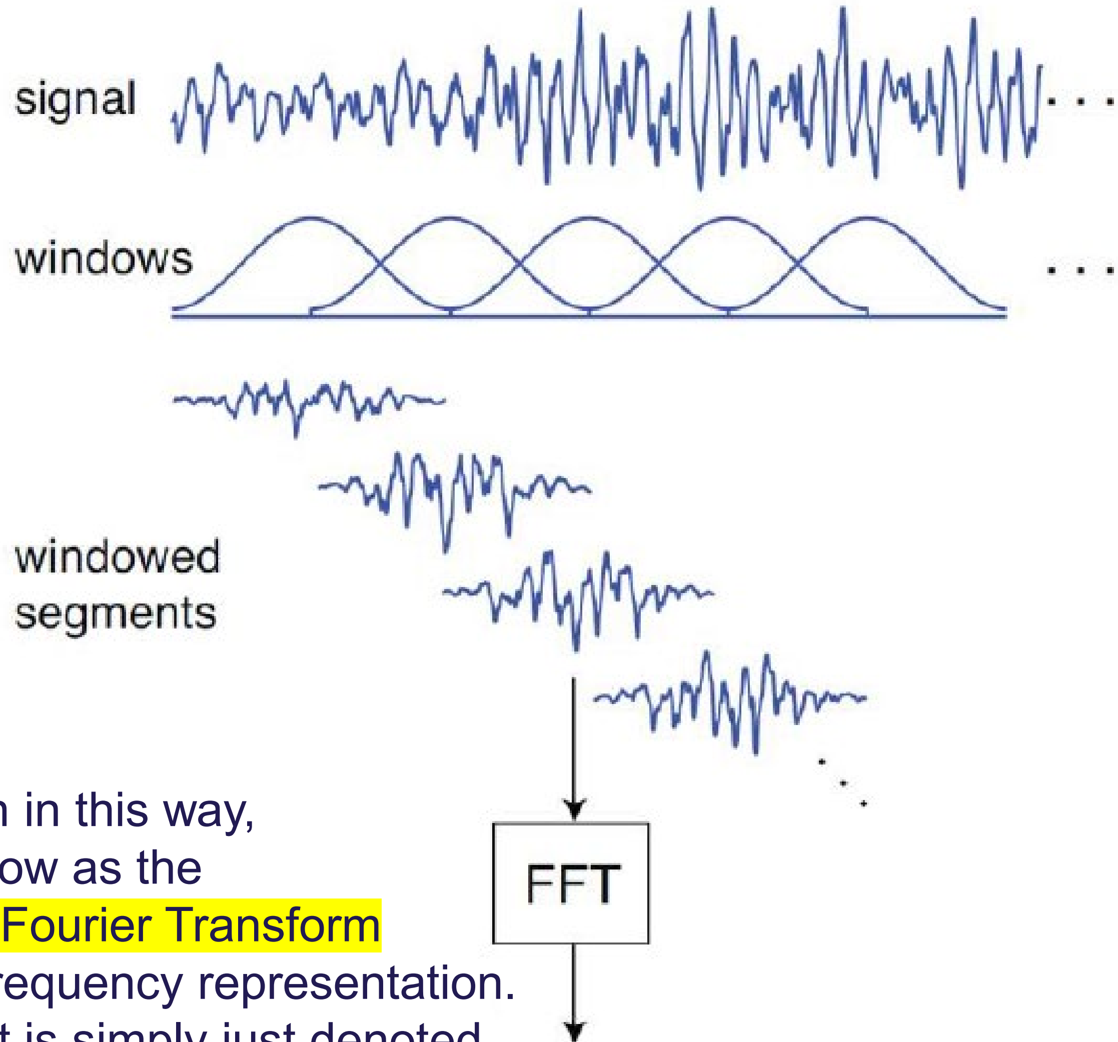
Window function

Located at the beginning of the sequence, $n = 0$

For every n, calculate X for $0 \leq \lambda < 2\pi$ (or any other interval of length 2π).
Since $w[m] = 0$ outside $0 \leq m \leq M-1$, the sum reduces to

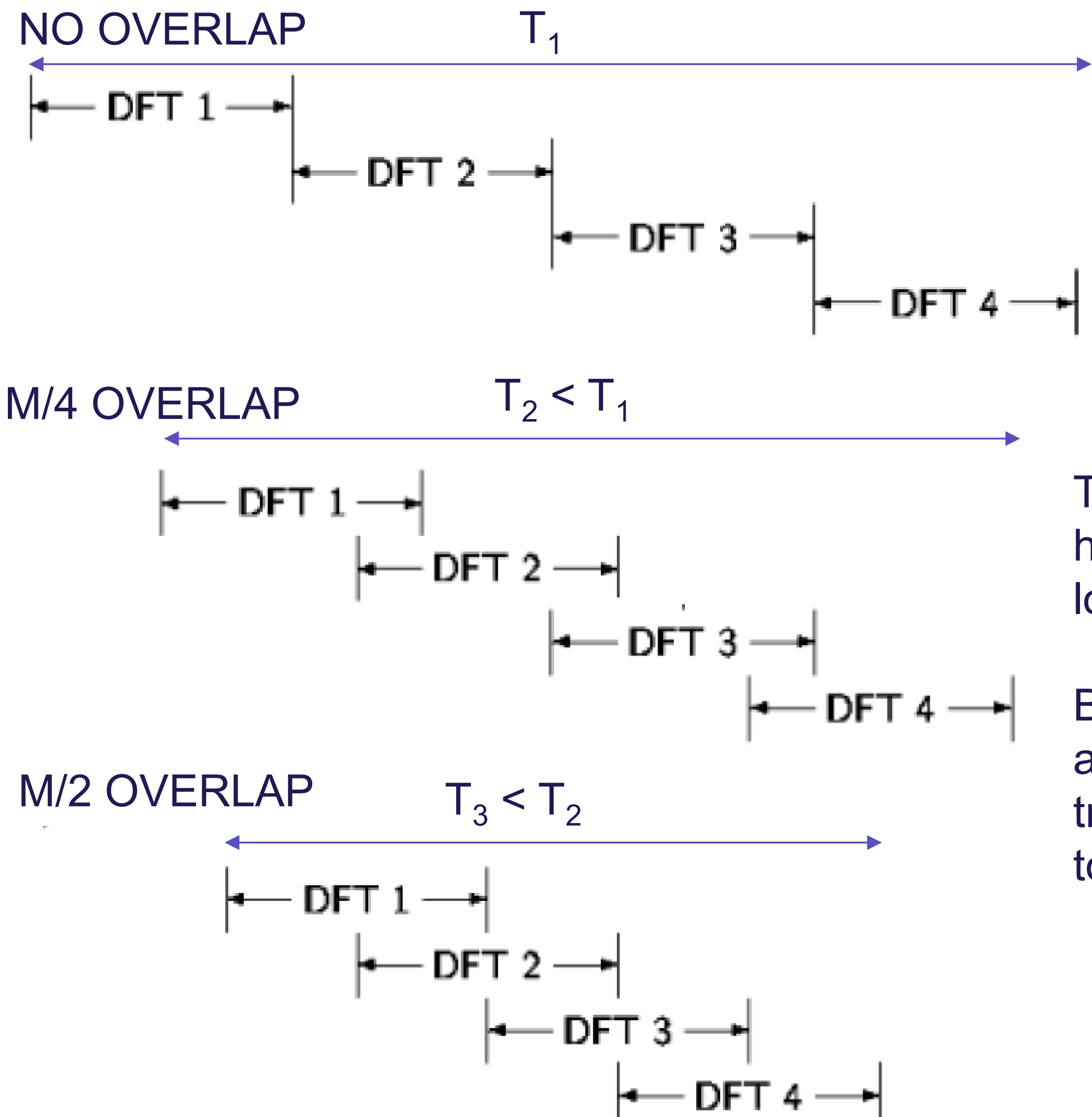
$$X[n, \lambda) = \sum_{m=0}^{M-1} x[n+m]w[m]e^{-j\lambda m}$$

The Short Time Fourier Transform - slide the window in steps of $M \geq 1$ samples



Doing the calculation in this way,
we obtain what is know as the
Discrete Short Time Fourier Transform
due to the discrete frequency representation.
In many literatures, it is simply just denoted
the Short Time Fourier Transform

The overlap in a STFT calculation



The larger the overlap, the higher the computational load per time unit…

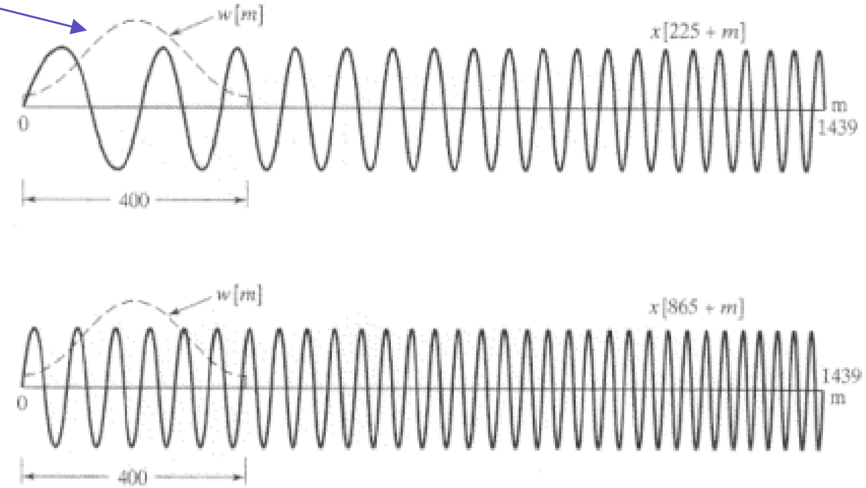
But at the same time, we also obtain a smoother transition from one segment to the next.

An example - a linear frequency sweep

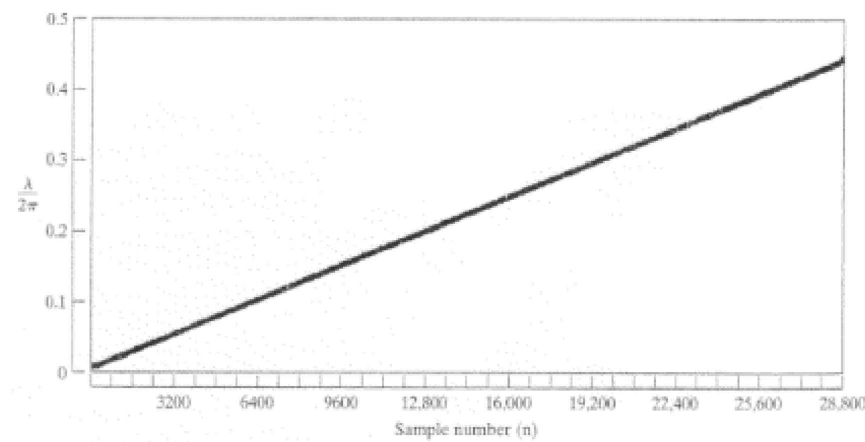
$$x[n] = \cos(\omega_0 n^2),\ \omega_0 = 2\pi \cdot 7.5 \cdot 10^{-6}$$

Hamming window
with length N=400

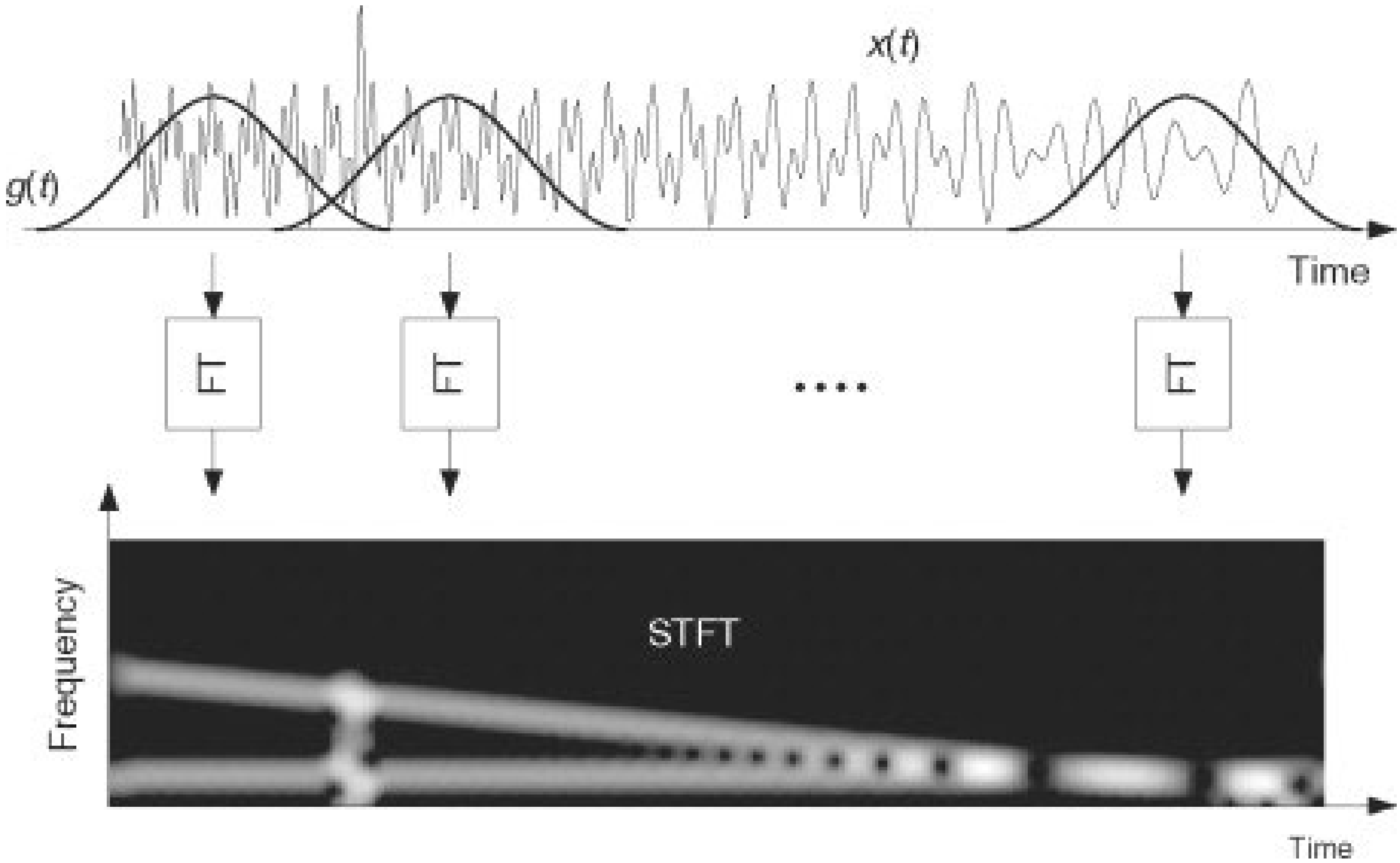
Time:



Time/
frequency:

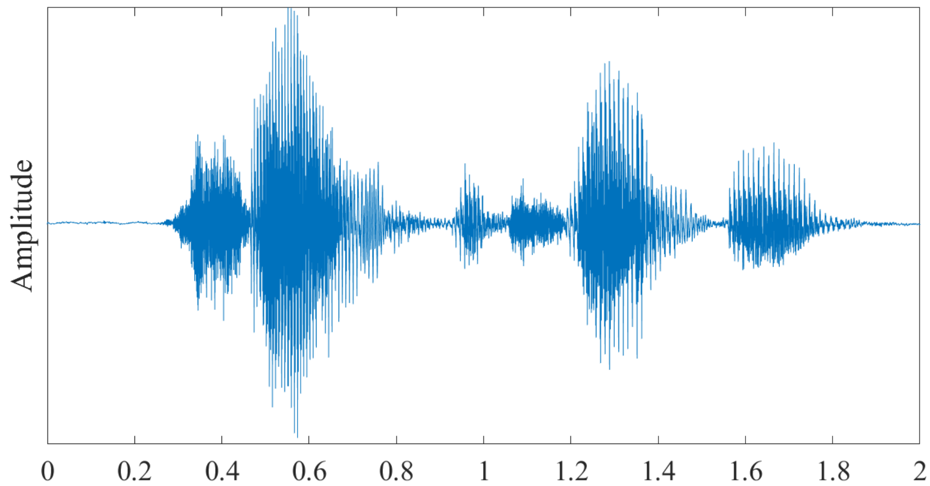


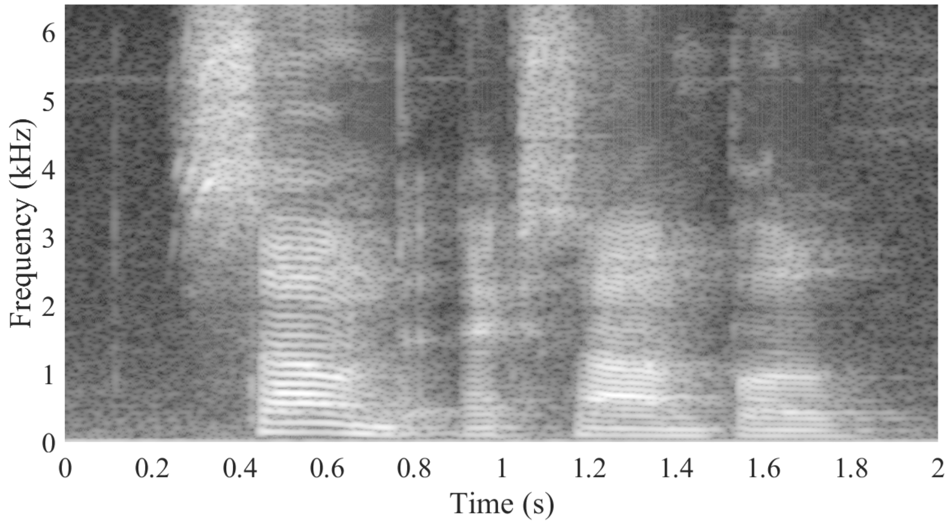
Simultaneous Time- and Frequency representation - The Spectrogram



AALBORG UNIVERSITY
DENMARK

Simultaneous Time- and Frequency representation - The spectrogram of a speech signal





STFT – an alternative interpretation

$$X[n,\lambda) = \sum_{m=-\infty}^{\infty} x[n+m]w[m]e^{-j\lambda m}$$

we make the substitution $m' = n + m$

$$X[n,\lambda) = \sum_{m'=-\infty}^{\infty} x[m']w[-(n-m')]e^{j\lambda(n-m')}.$$

can be interpreted as the convolution

$$X[n,\lambda) = x[n] * h_\lambda[n],$$

where

Modulation in Time (heterodyne)
~ Shift in Frequency

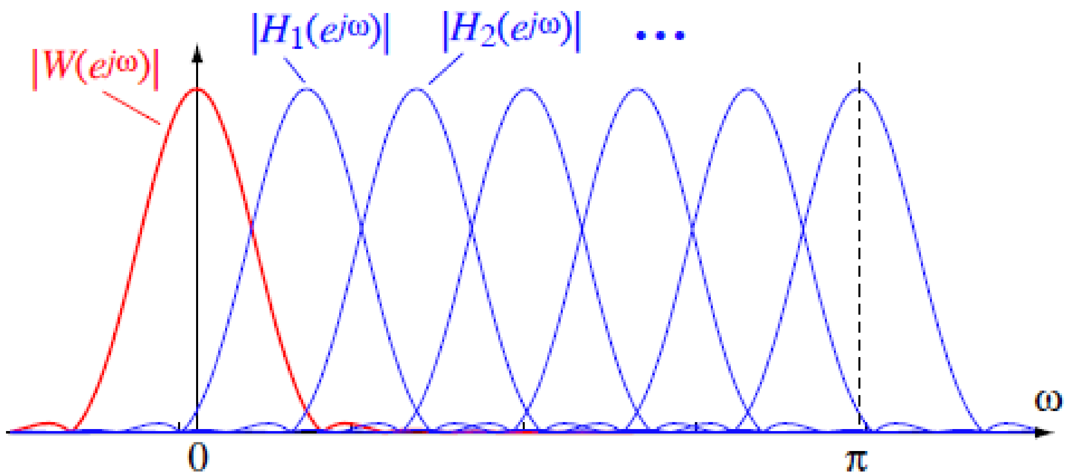
$$h_\lambda[n] = w[-n]e^{j\lambda n}.$$

we see that the time-dependent Fourier transform as a function of n with λ fixed can be interpreted as the output of a linear time-invariant filter with impulse response $h_\lambda[n]$ or, equivalently, with frequency response

$$H_\lambda(e^{j\omega}) = W(e^{j(\lambda-\omega)}).$$

Thus, the STFT can be considered as a set of N parallel and frequency shifted bandpas filters with different center frequencies λ_k and amplitude responses defined by the window function…

STFT – The Filter Bank interpretation



A filter bank consisting of N identical frequency-shifted bandpass filters

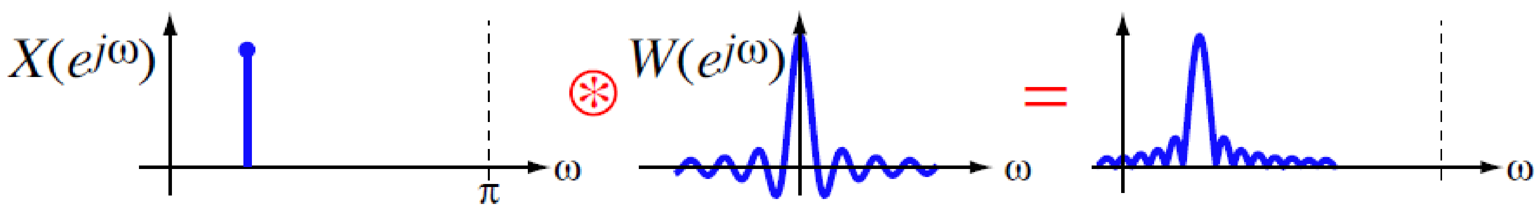
The more narrow the main-lope of the window function is, and the more values of λ we choose, i.e., the larger N, the better a spectral estimation we get…

How to choose an appropriate window - function and length

Choosing an appropriate window is not necessarily an easy and straight forward task…

This is exactly the same problem as we were facing when designing FIR filters using the window method.

- e.g. if $x[n]$ is a pure sinusoid,

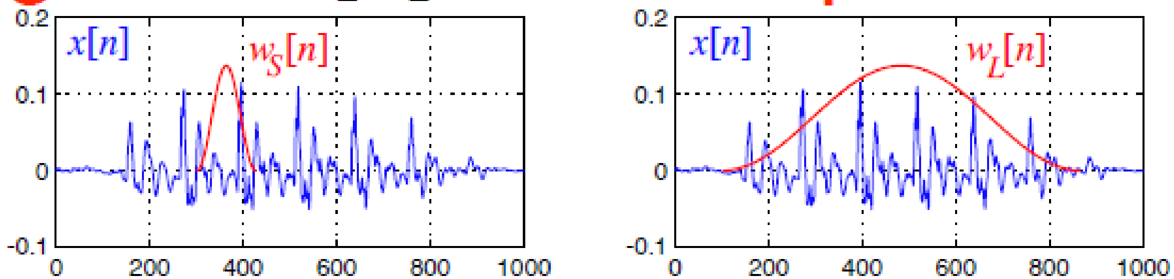


Window effect due to width of main lobe -> blurring
Window effect due to non-zero side lobes -> leakage

AALBORG UNIVERSITY
DENMARK

Choosing a window

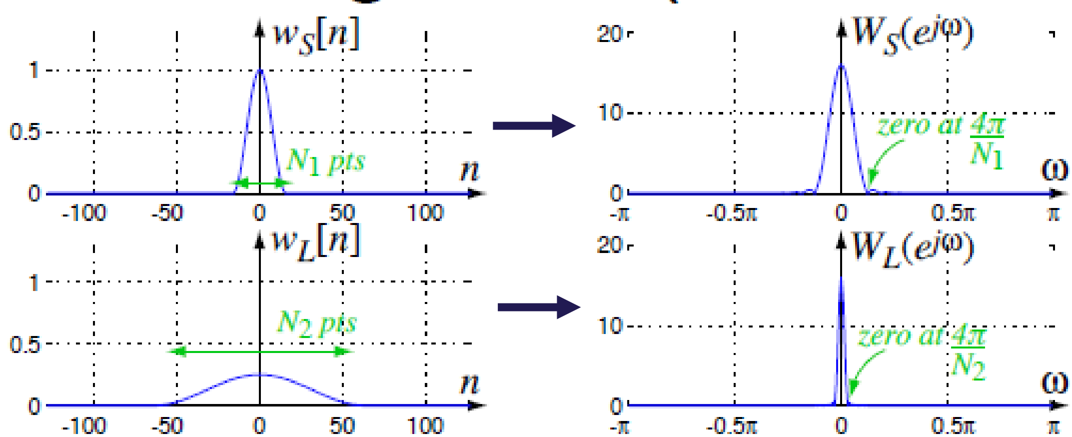
- Length of $w[n]$ sets temporal resolution



short window measures only local properties

longer window averages spectral character

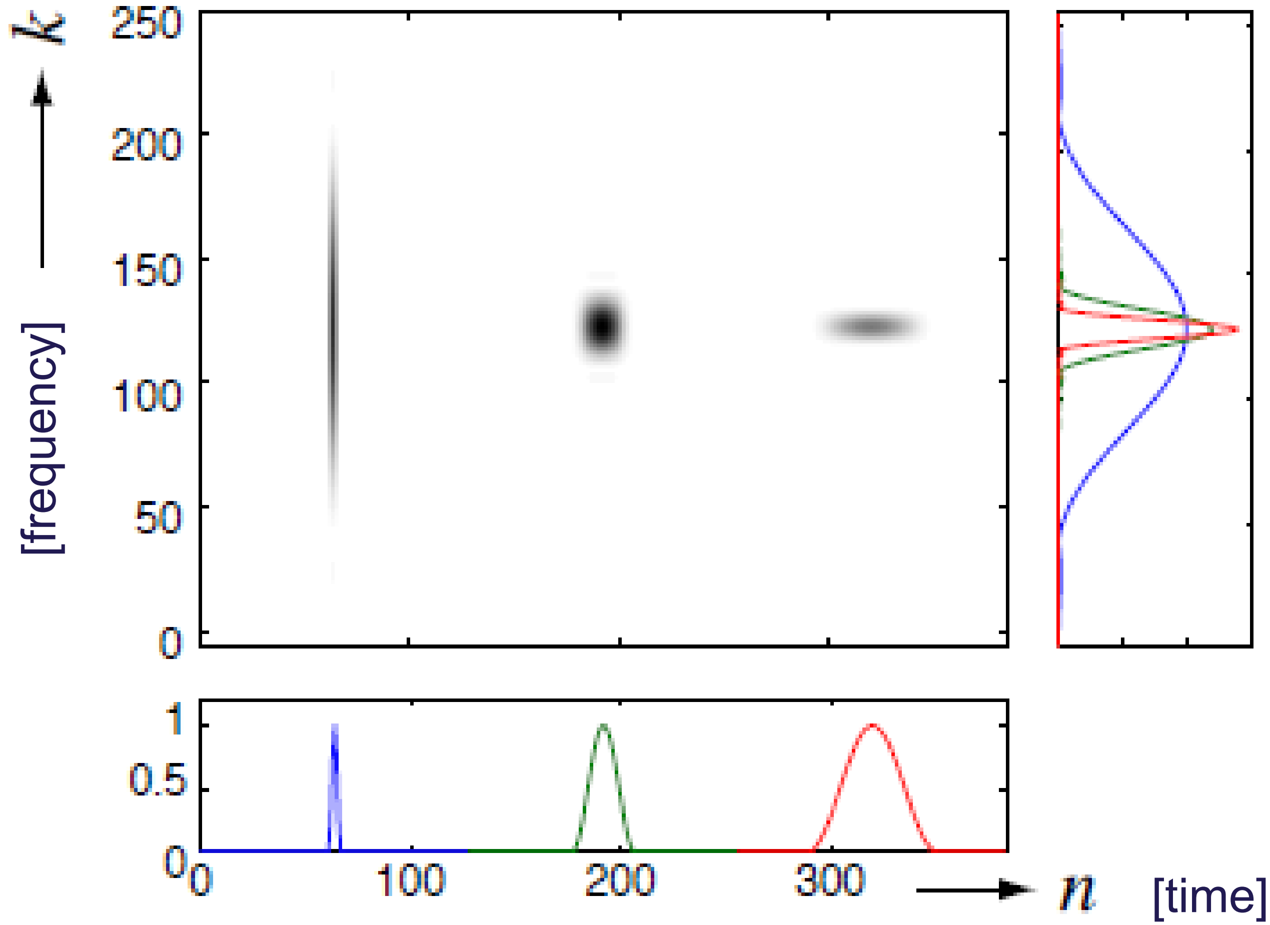
- Window length $\propto$ 1/(Mainlobe width)



shorter window → *more blurred spectrum*

- *more* time detail ↔ *less* frequency detail

Choosing a window
- here we have three windows with different length



Choosing a window

- The window should be narrow enough to ensure that the portion of the signal falling within the window is statistically time-invariant – quasi stationary.

- But … very narrow windows do not offer good localization in the frequency domain.

Wide window → good frequency resolution, poor time resolution.

Narrow window → good time resolution, poor frequency resolution.

IMPORTANT…!!

Heisenberg's Uncertainty Relation

$$\Delta t \cdot \Delta f \geq \frac{1}{4\pi}$$

Time resolution: **Frequency resolution:**

Δt and Δf CANNOT AT THE SAME TIME be made arbitrarily small

AALBORG UNIVERSITY
DENMARK

Narrowband vs. Wideband Spectral Analysis