# Signalbehandling for computer-ingeniører
# COMTEK-5, E22
# &
# Signalbehandling
# EIT-5, E22

# 14. Practical DFT Analysis and The Short Time Fourier Transform (STFT)

Assoc. Prof. Peter Koch, AAU

# The outline of today's lecture

- A short recap…

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

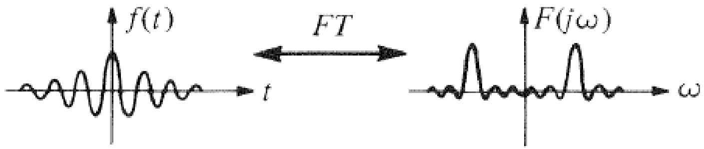
- The Short Time Fourier Transformation
  Time varying signals
  Simultaneous time-and-frequency analysis
  Heisenberg uncertainty principle

# Fourier Transform – a classification

R.A. Roberts & C.T. Mullis, ”Digital Signal Processing”, Addition-Wesley, 1987.

| | Continuous in time | Discrete in time – Periodic in frequency |
|---|---|---|
| Continuous in frequency | Fourier transform | Discrete-time Fourier transform |
| Discrete in frequency – Periodic in time | Fourier series | Discrete Fourier transform |

**Continuous in time, continuous in frequency**

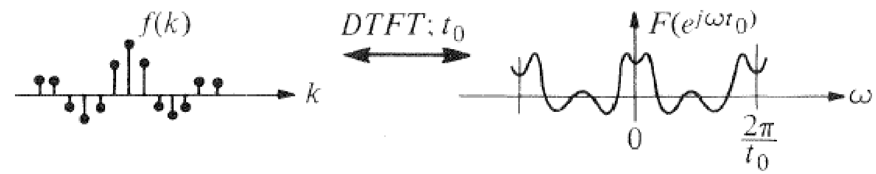



$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(j\omega)\, e^{j\omega t}\, d\omega$$

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

**Fourier transform**

**Discrete in time – Periodic in frequency, continuous in frequency**

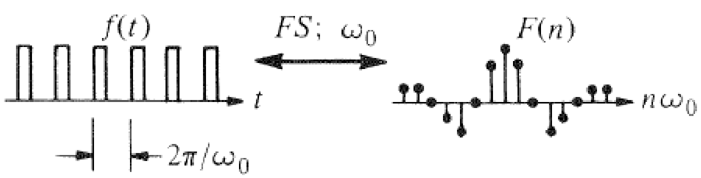



$$f(k) = \frac{t_0}{2\pi}\int_{-\pi/t_0}^{\pi/t_0} F(e^{j\omega t_0})\, e^{jk\omega t_0}\, d\omega$$

$$F(e^{j\omega t_0}) = \sum_{k=-\infty}^{\infty} f(k)\, e^{-jk\omega t_0}$$

**Discrete-time Fourier transform**

**Continuous in time, discrete in frequency – periodic in time**

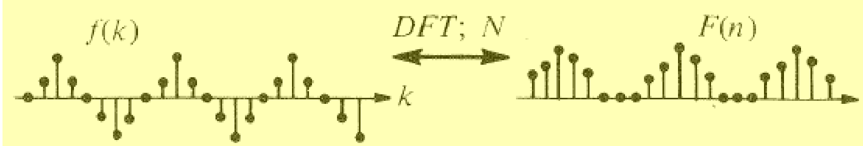



$$f(t) = \sum_{n=-\infty}^{\infty} F(n)\, e^{jn\omega_0 t}$$

$$F(n) = \frac{\omega_0}{2\pi}\int_{-\pi/\omega_0}^{\pi/\omega_0} f(t)\, e^{-jn\omega_0 t}\, dt$$

**Fourier series**

**Discrete in time – periodic in frequency, discrete in frequency – periodic in time**




$$f(k) = \frac{1}{N}\sum_{n=0}^{N-1} F(n)\, (e^{j2\pi/N})^{kn}$$

$$F(n) = \sum_{k=0}^{N-1} f(k)\, (e^{j2\pi/N})^{-kn}$$

**Discrete Fourier transform**

# Relation between the DFS and DTFT

Assume the infinite periodic sequence $\tilde{x}[n]$.

We have seen that the Discrete Fourier Series coefficients, which corresponds to the signal $\tilde{x}[n]$, can be found by sampling the DTFT. That is;

$$\tilde{X}[k] = X\left(e^{j\omega}\right)\big|_{\omega=(2\pi/N)k} = X(e^{j(2\pi/N)k}) \quad \text{for any } k$$

which corresponds to sampling $X(z)$ in $N$ equally spaced angles on the unit circle.

…and similarly for the Discrete Fourier Transform (DFT)

AALBORG UNIVERSITY
DENMARK

# Relation between DFS and DFT

Basically, what we do is that we sample the DTFT at equividistant frequency values over one period. Since we have just argued that the Fourier Series coefficients $\tilde{X}[k]$ can be represented by the DTFT, it follows that

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

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

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

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

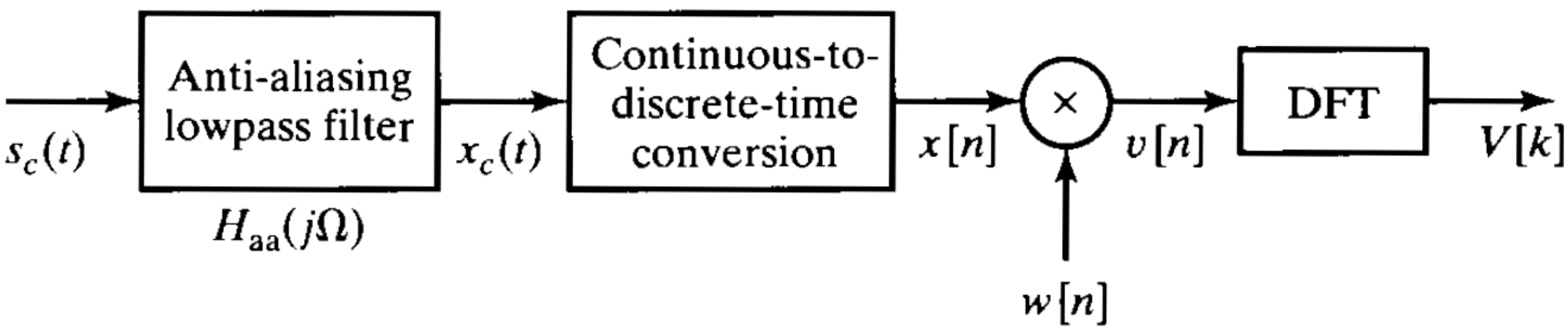
Here it is important though to notice that both $x[n]$ and $X[k]$ equals zero outside the interval $[0; N-1]$.

AALBORG UNIVERSITY
DENMARK

# Fourier Analysis using the DFT

The challenge with the DFT is that it 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*.

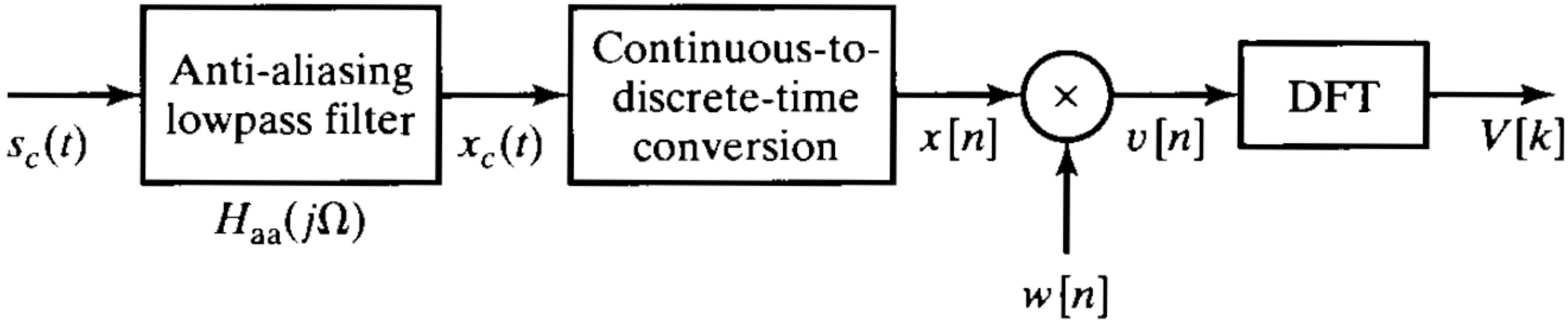



Processing steps in the discrete-time Fourier analysis of a continuous-time signal.

Now, let’s recap what happens throughout this signal chain...

AALBORG UNIVERSITY DENMARK

# 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 duration sequences by multiplying with a window function $w[n]$.

But what is a window actually, and what does it looks like in the frequency domain..??

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

AALBORG UNIVERSITY
DENMARK

# Typical window functions

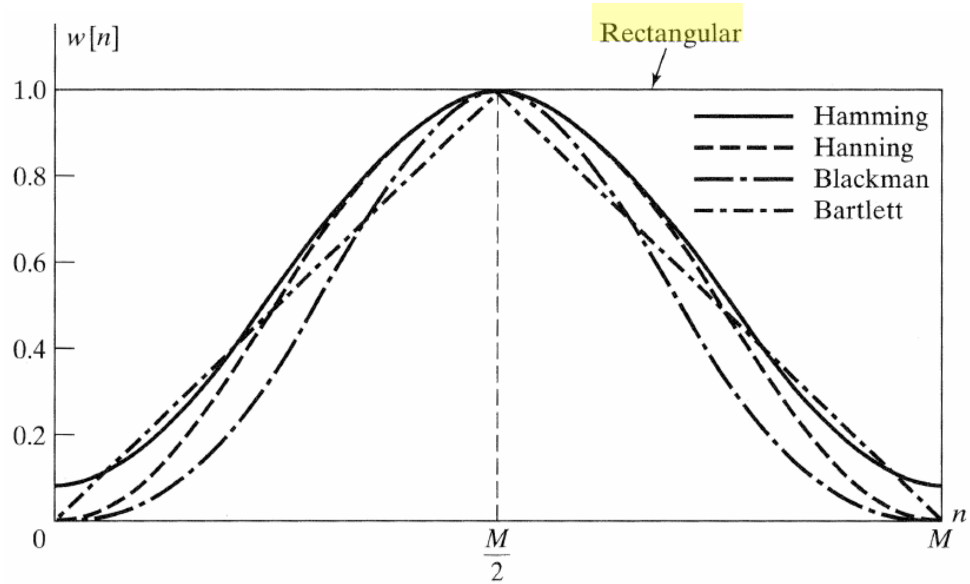
Normally, we would think of the Rectangular Window as being the most ”obvious” function to truncate a sequence.
The problem is however, that at the edges of the window, we have discontinuities which may impact negatively the overall performance of the DFT analysis.

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.

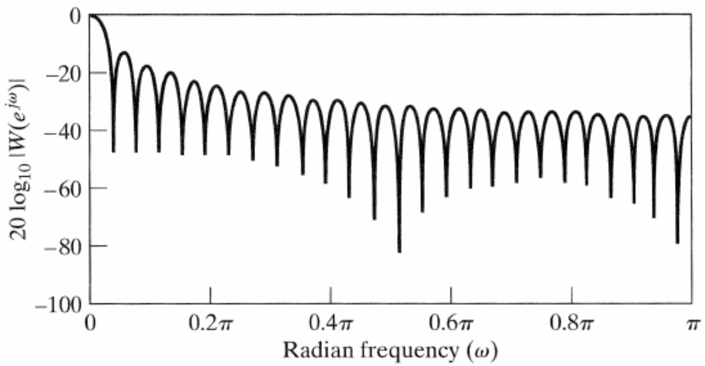
Consequently, in the frequency domain we are convolving the FT of the input sequence $x[n]$ with the FT of the window function – and to be precisely, we conduct a periodic convolution.

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

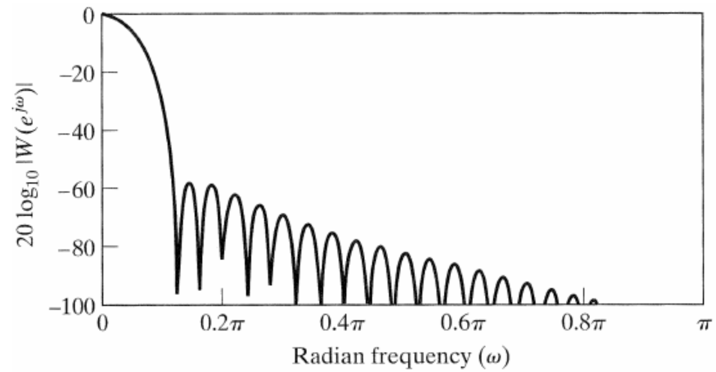
AALBORG UNIVERSITY
DENMARK

# Amplitude response of the window functions

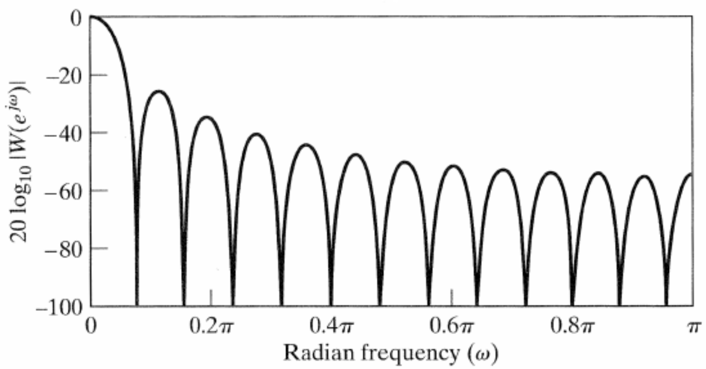
Rectangular



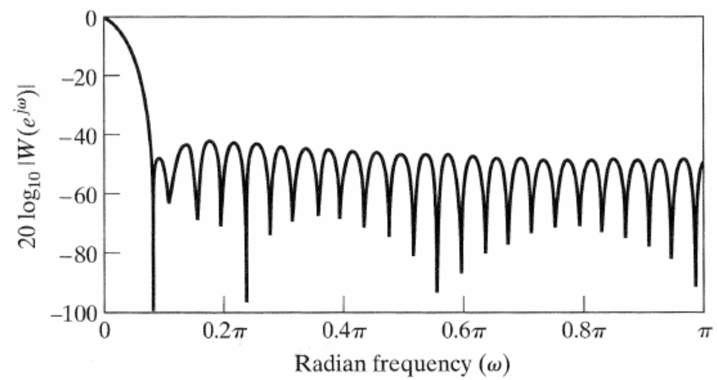

Blackman



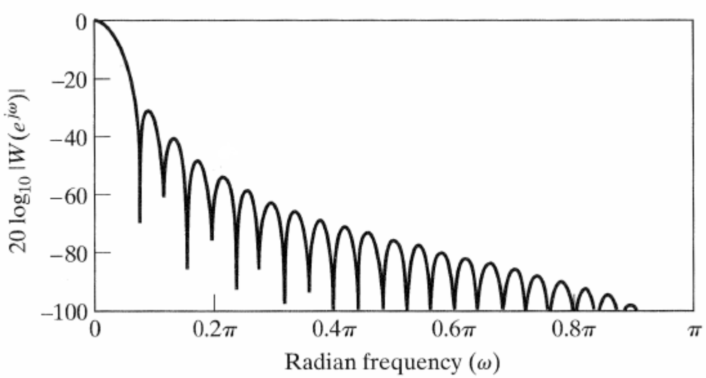

Bartlett



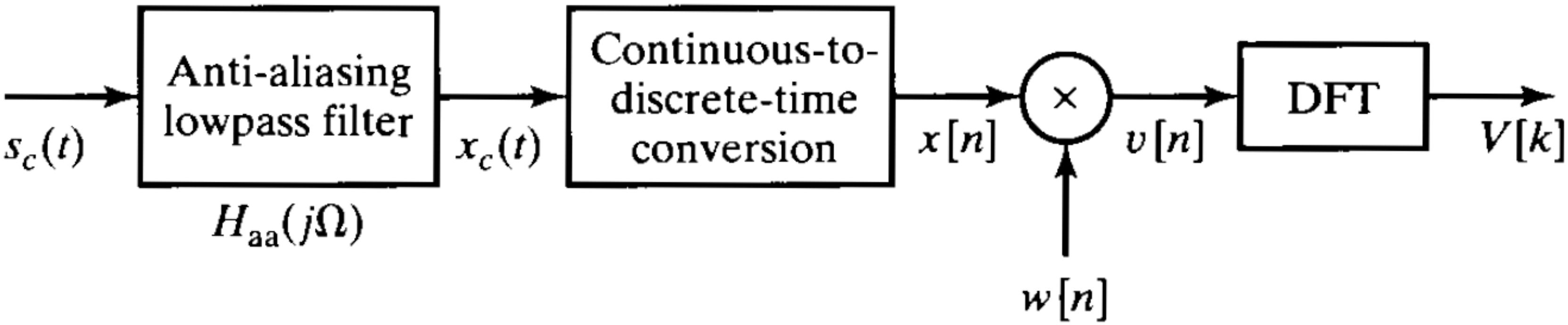

Hamming




Hanning




| 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

$$V[k]=V(e^{j\omega})\big|_{\omega=2\pi k/N}$$

Since $V[k]$ represents one period, $[-\pi;\pi]$ or $[0;2\pi]$, where $2\pi$ is the sample frequency, we can easily calculate ”the spectral accuracy”…

# Frequency resolution

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

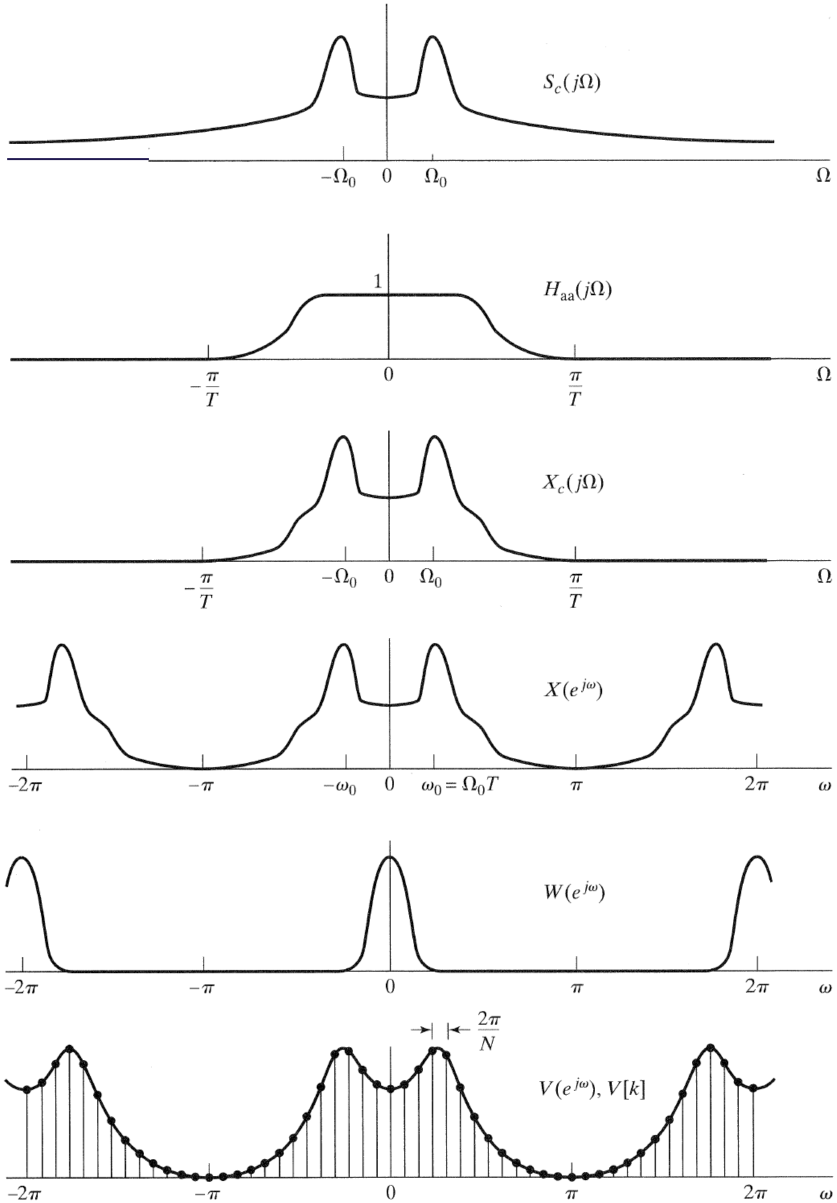
Now, using the relation $\omega = T\Omega$, then we have $T\Omega_k = {}^{2\pi k}/_{N}$ and thus;

$$\boxed{\Omega_k = {}^{2\pi k}/_{TN} \quad \Rightarrow \quad f_k = k\frac{f_s}{N} \quad 0 \le k \le N-1}$$

One important lesson to learn here is that frequency resolution is proportional with the order of the DFT, and inverse proportional with the sample frequency.

AALBORG UNIVERSITY
DENMARK

# 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 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;

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

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

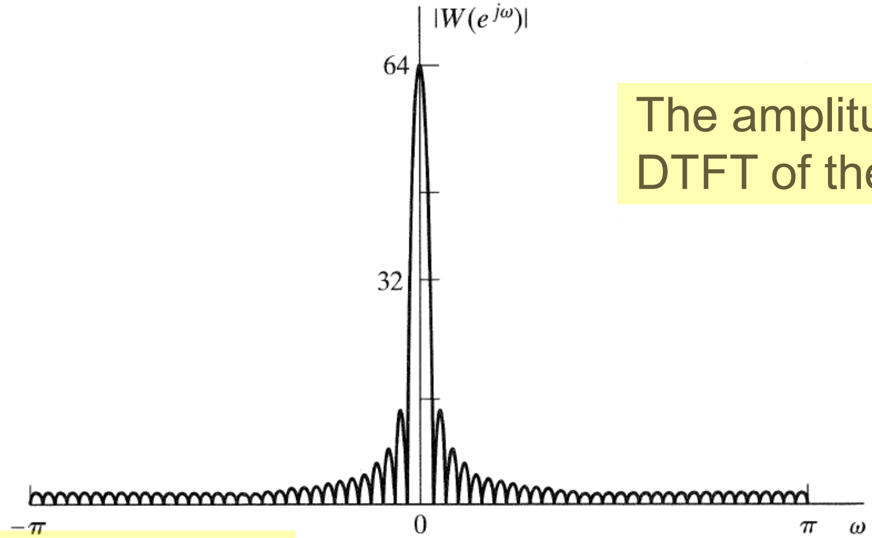
$$+ \frac{A_1}{2} e^{j\theta_1} W(e^{j(\omega-\omega_1)}) + \frac{A_1}{2} e^{-j\theta_1} W(e^{j(\omega+\omega_1)})$$

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 $A_1 = 0{,}75$

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



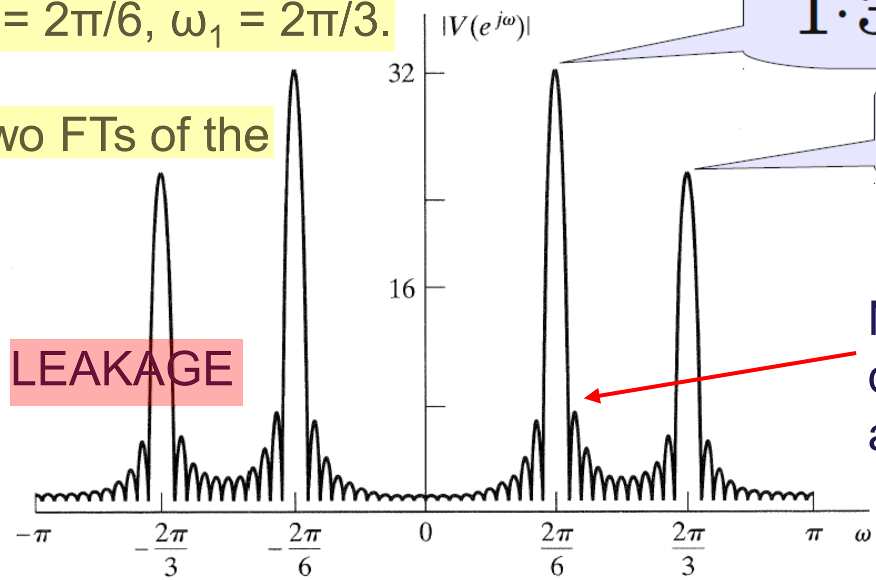

The amplitude (modulus) of the DTFT of the rectangular window

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…

$1 \cdot 32$ (one half)

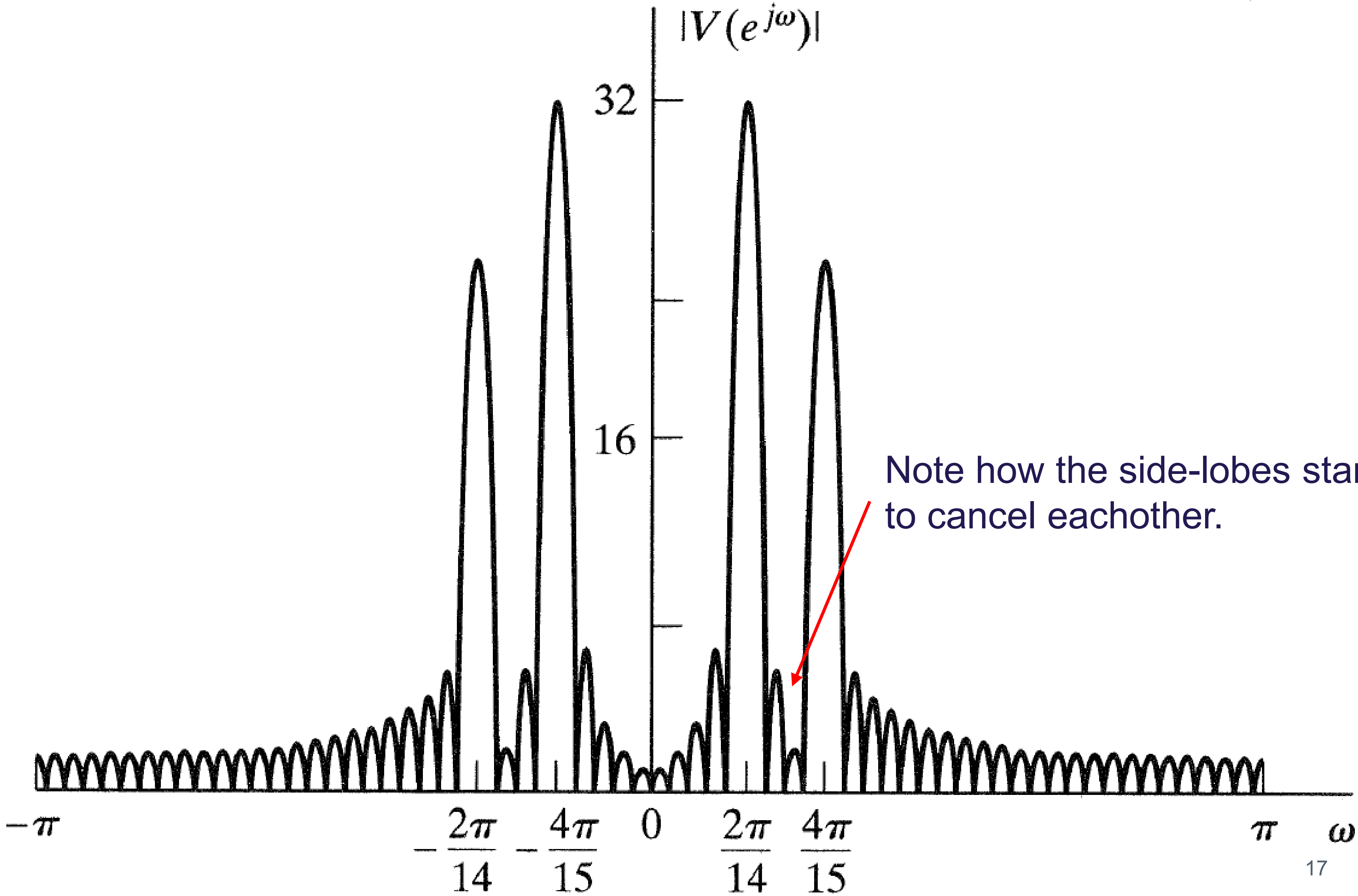
$0.75 \cdot 32$




This effect is called LEAKAGE

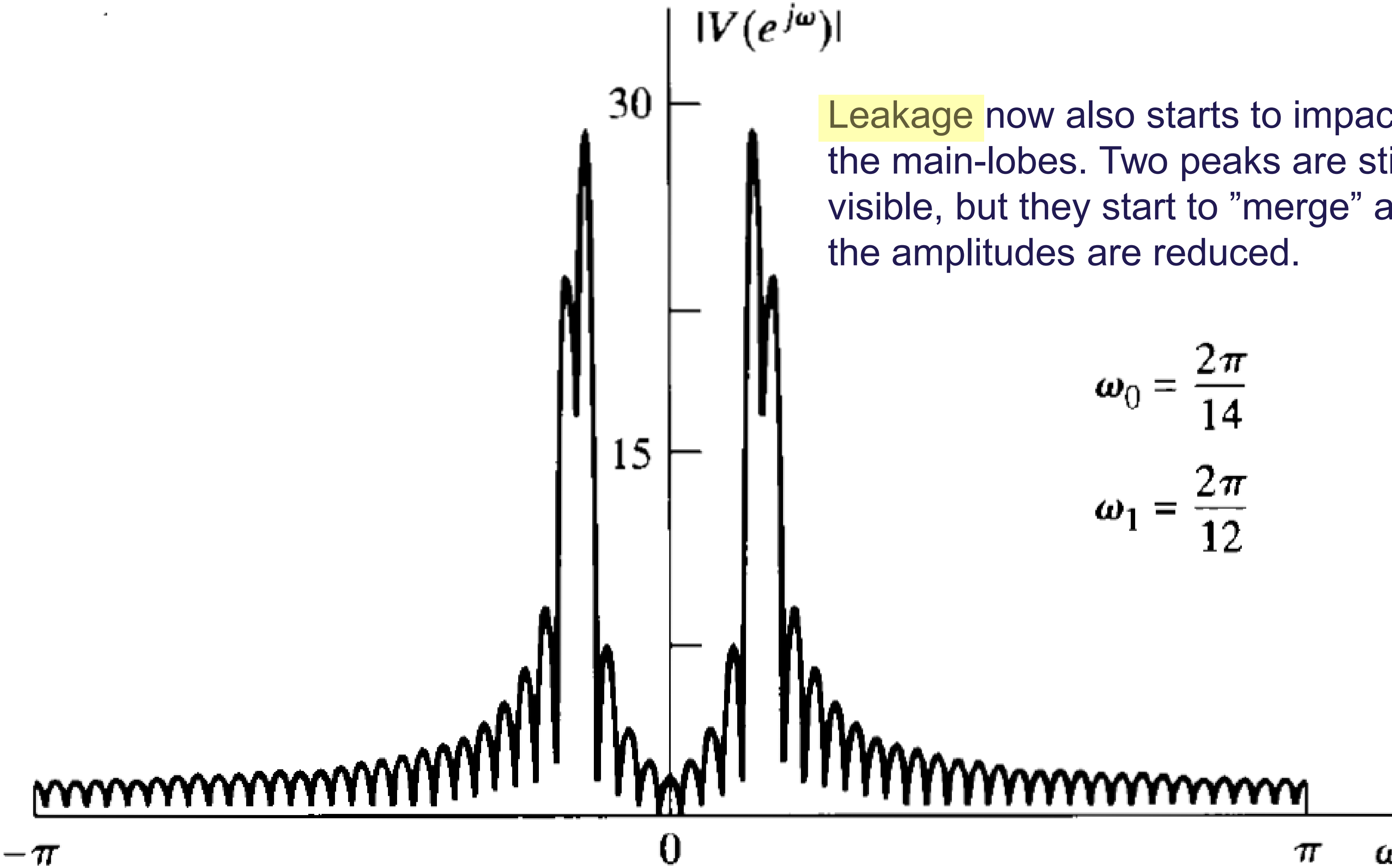
Note that the side lobes are down approx. 13 dB – just as expected…

# 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



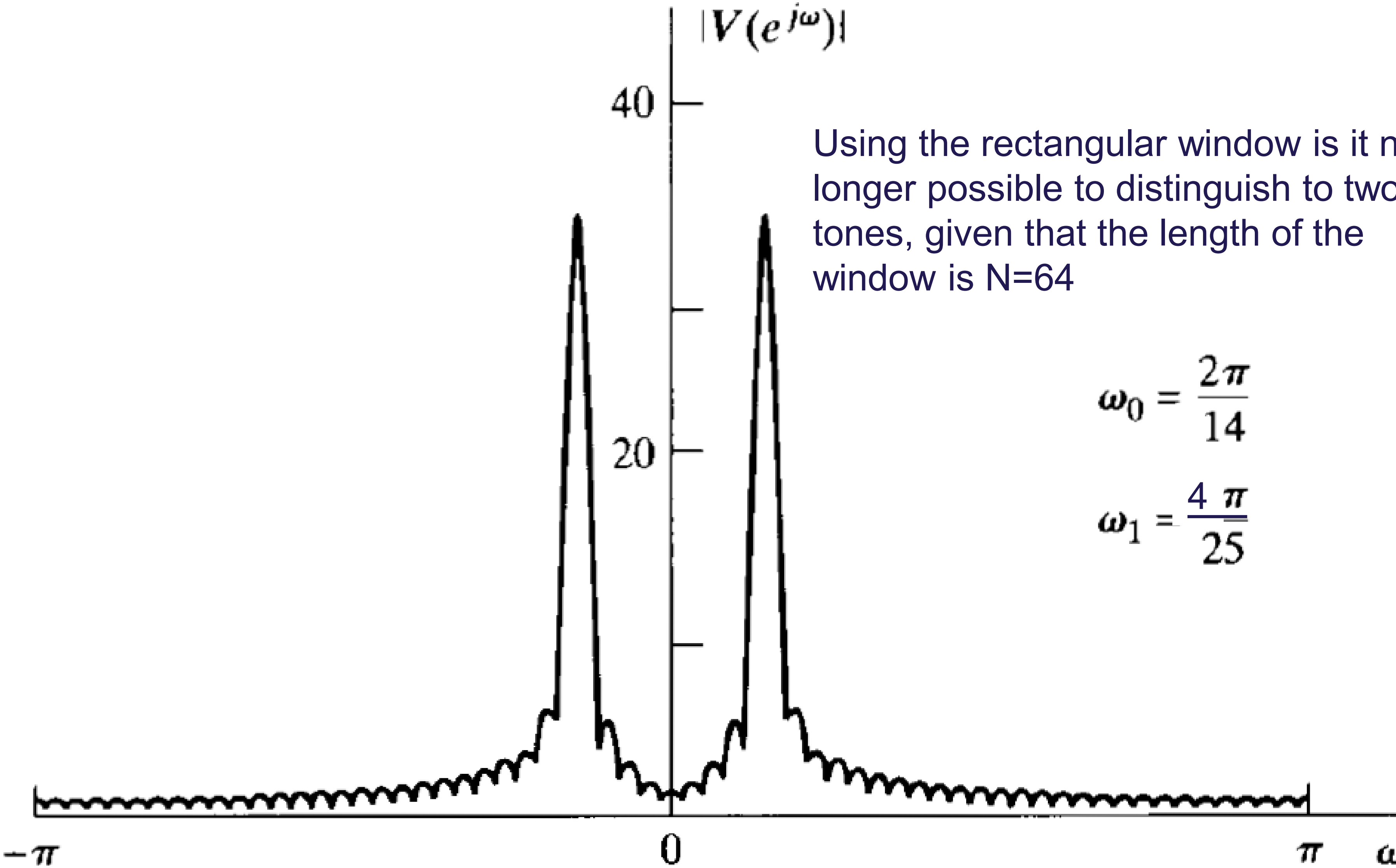

(d)

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 window's main-lobe width and thus the window length.

The leakage depends on the ratio between main-lobe amplitude and side-lobe amplitudes.

The rectangular window gives the highest possible frequency resolution but also has the largest side-lobes.

# 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:

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

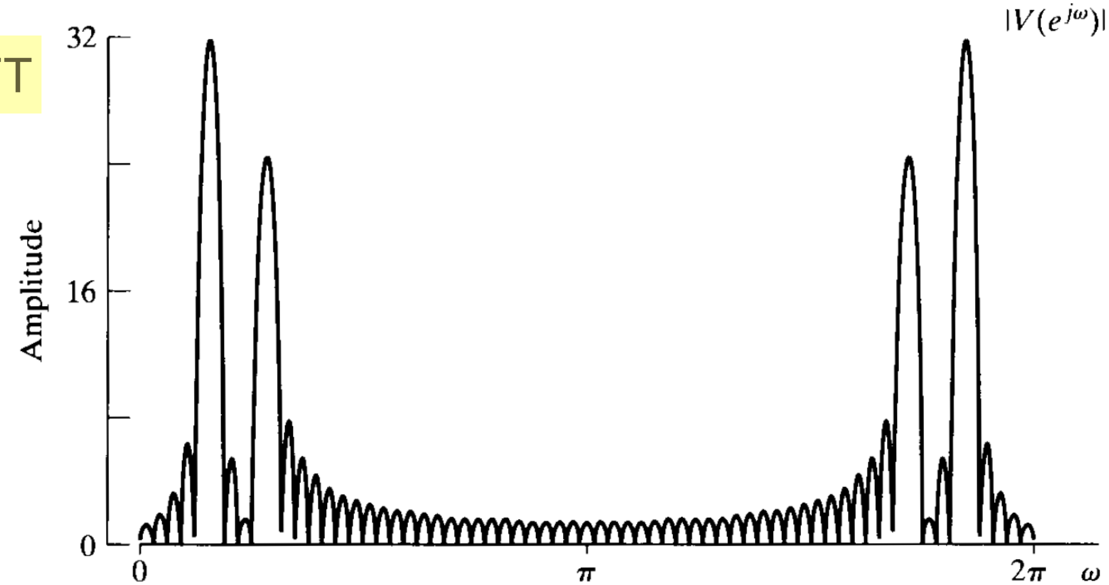
Example of a sampled (2-tone) signal, truncated using a rectangular window with length 64:

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

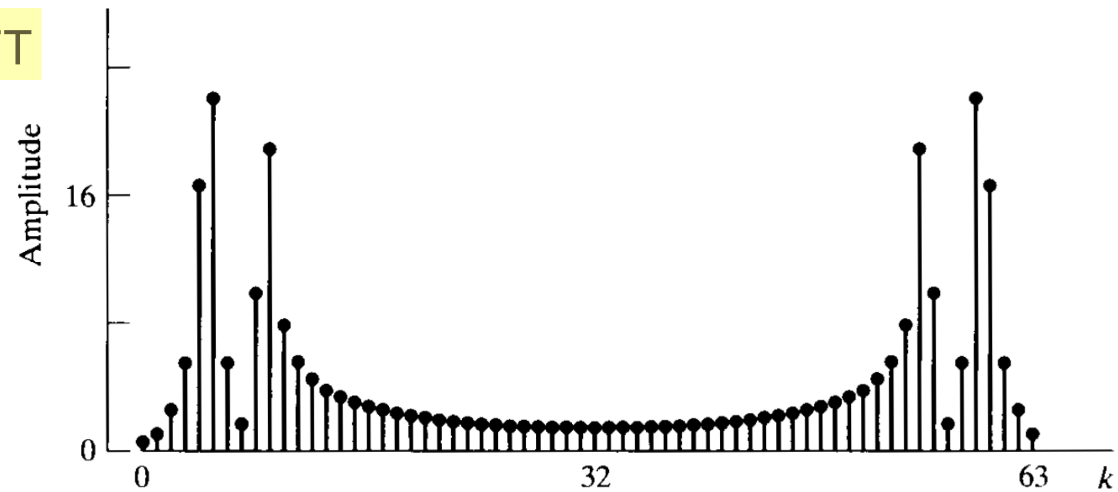
The $2^{nd}$ experíment shown on slide no. 17.

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

DTFT




DFT

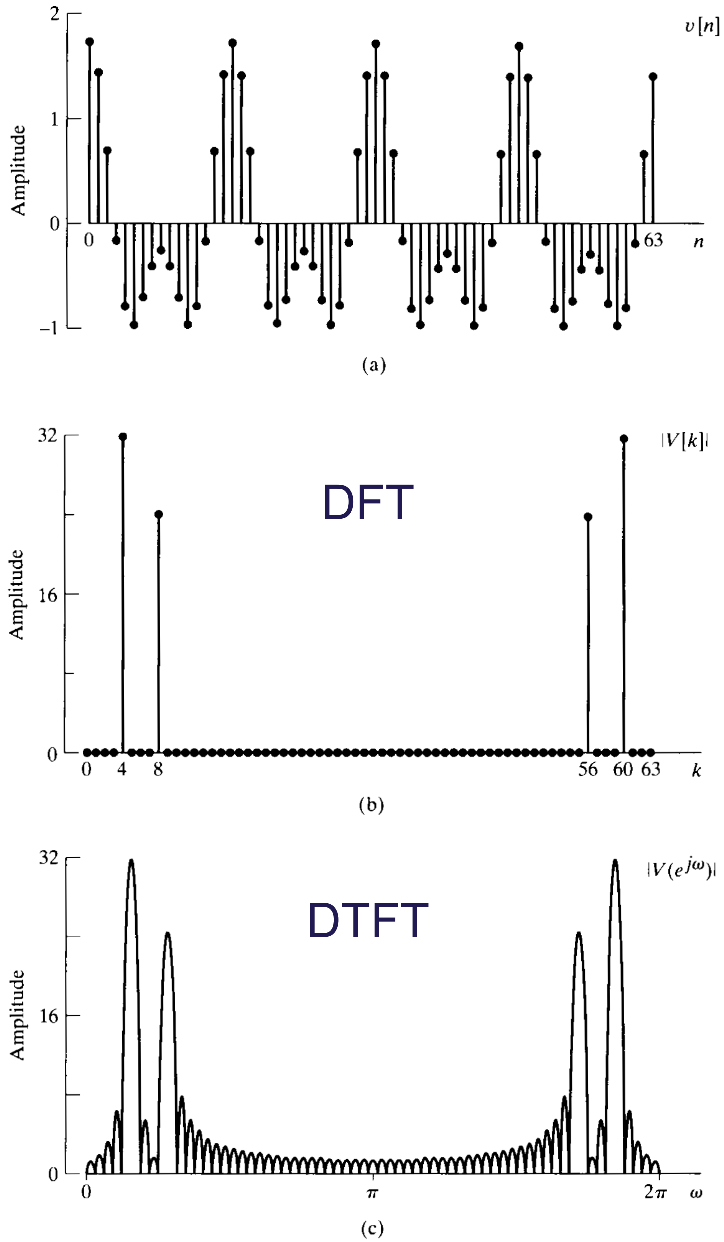



Note the difference in amplitude at the two peaks between the DTFT and the DFT.

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

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…

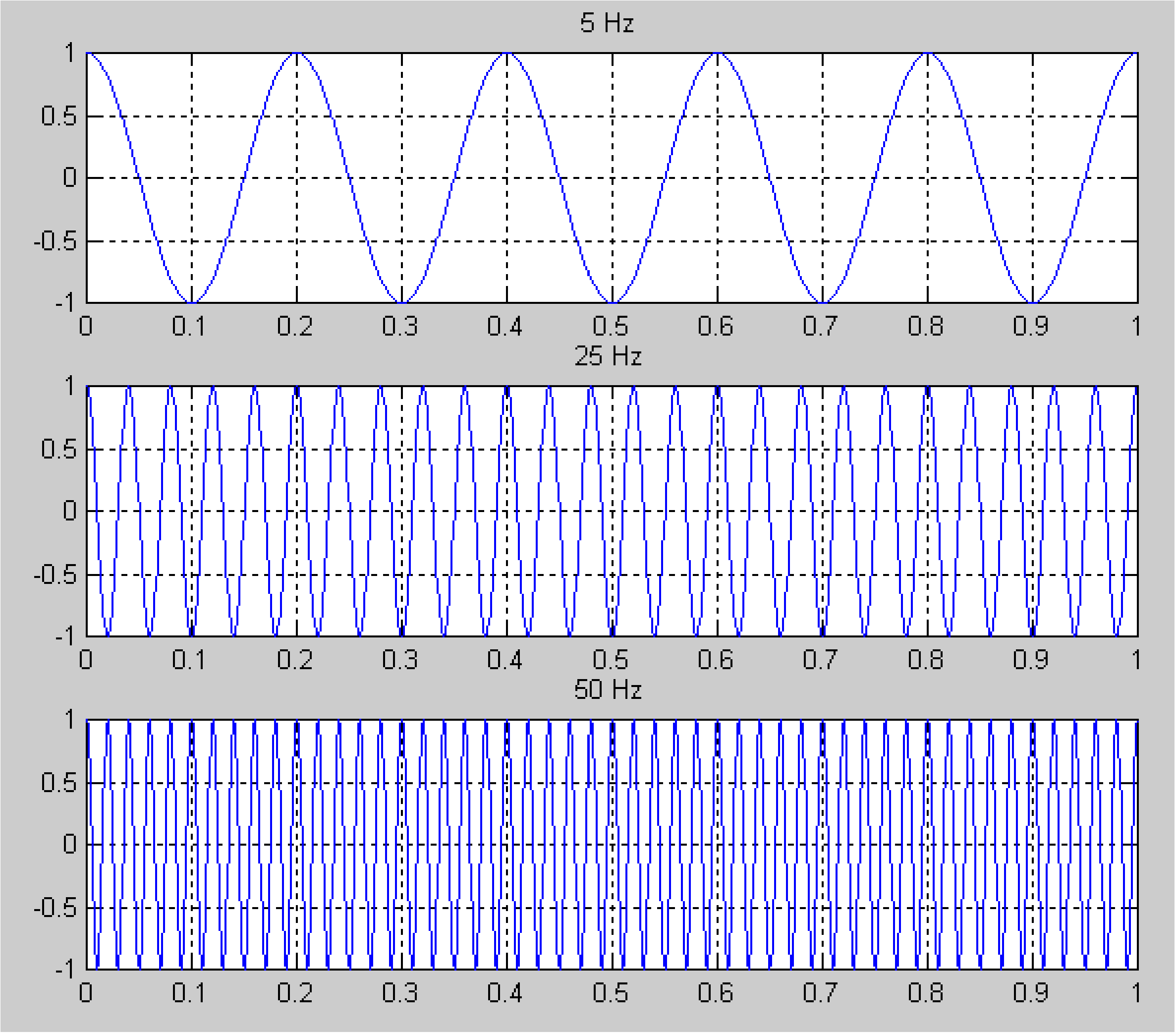

**…before we discuss the Short-Time Fourier Transform (STFT)**

# A motivating example – simple sinusoids

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

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

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

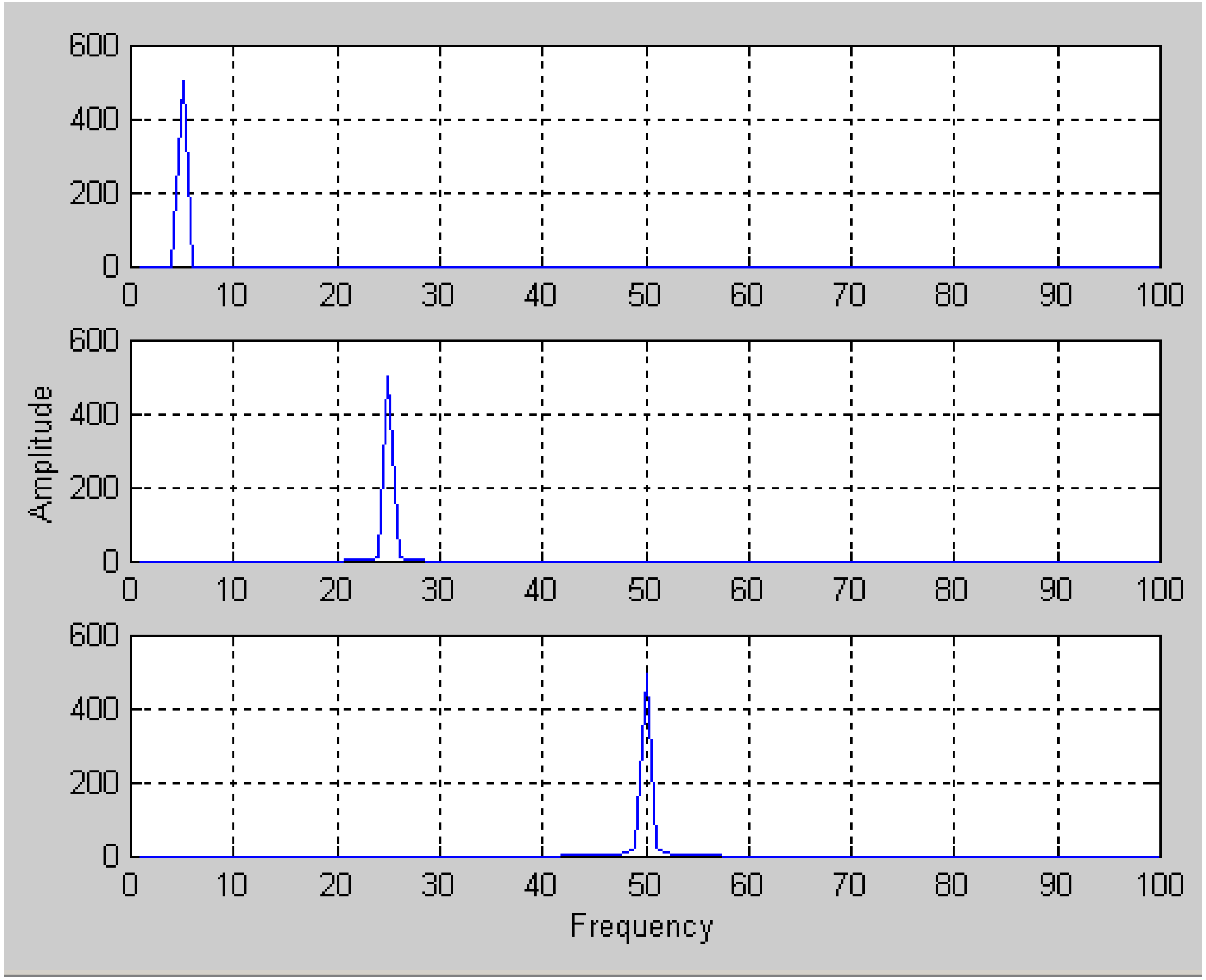


# The Amplitude Responses

$F_1(\Omega)$

$F_2(\Omega)$

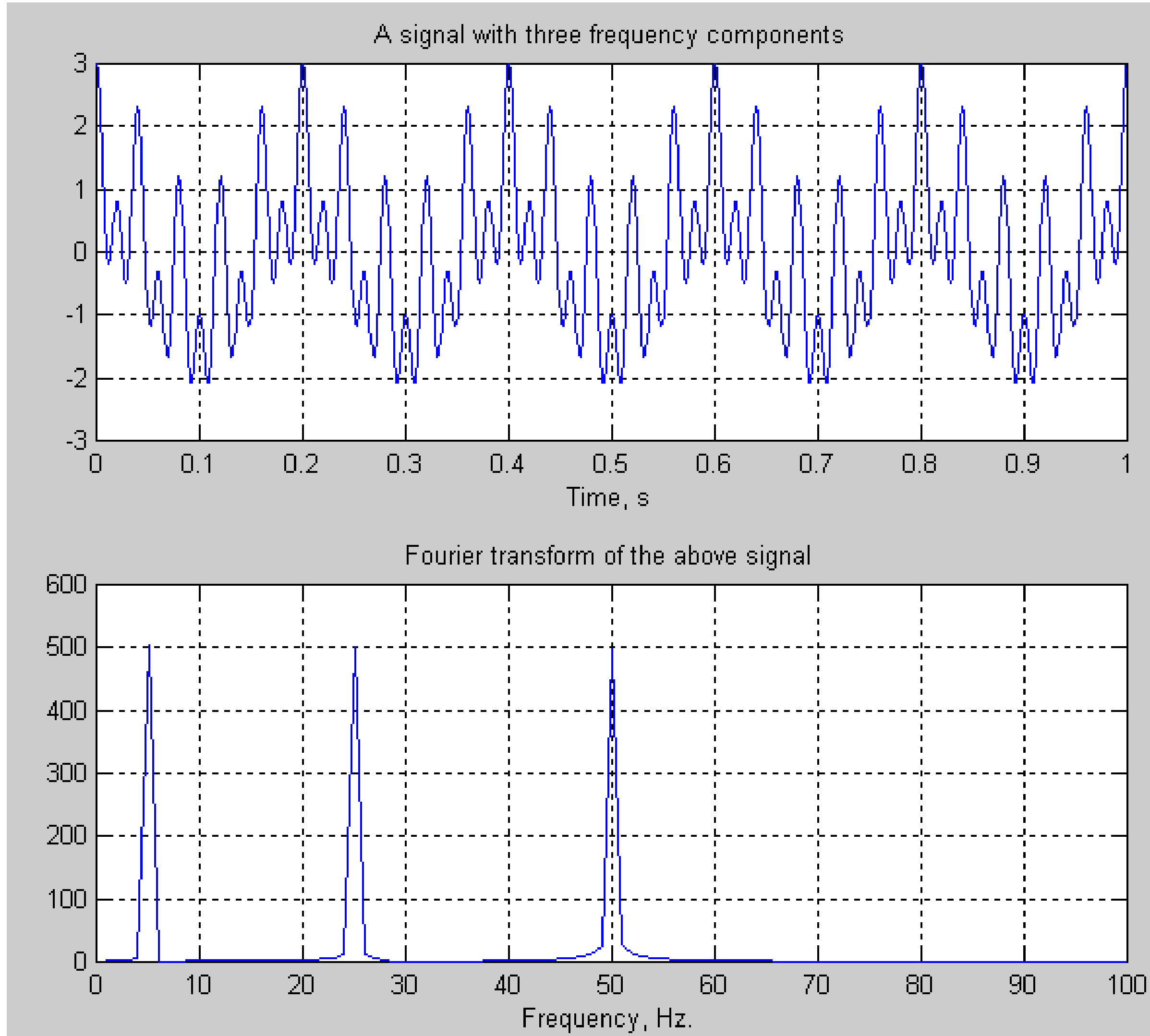
$F_3(\Omega)$



# Now, let's add the three sinusoids

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

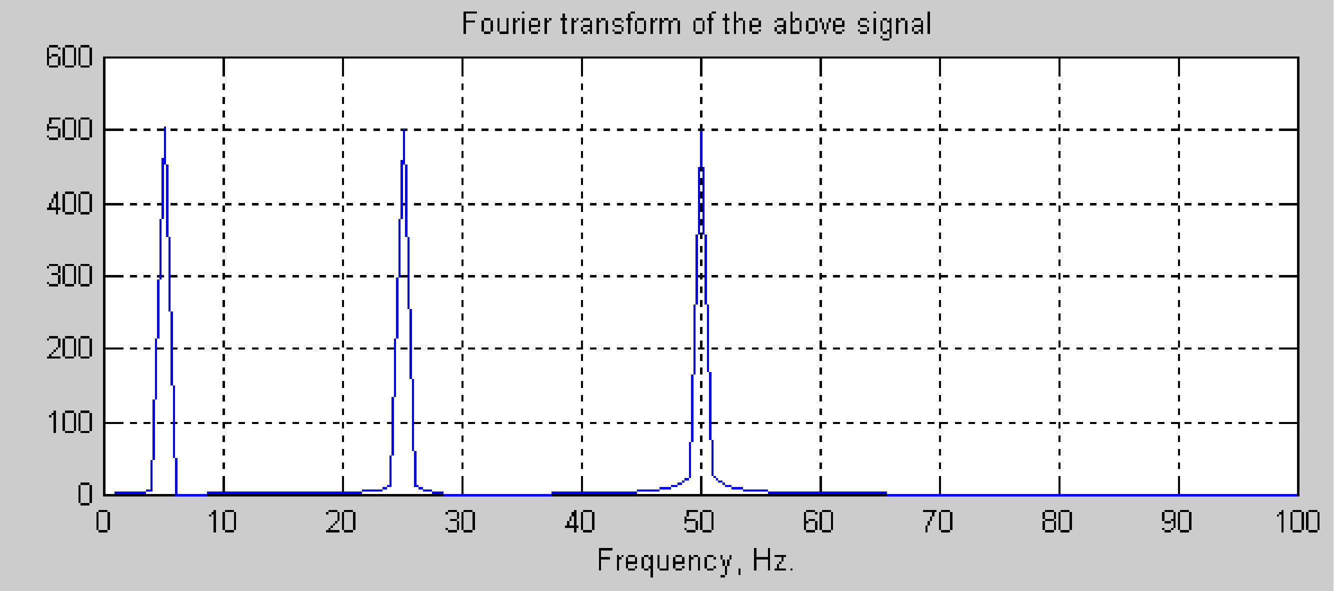
$F_4(\Omega)$




AALBORG UNIVERSITY
DENMARK

# The sum of sinusoids is a Time-Invariant signal

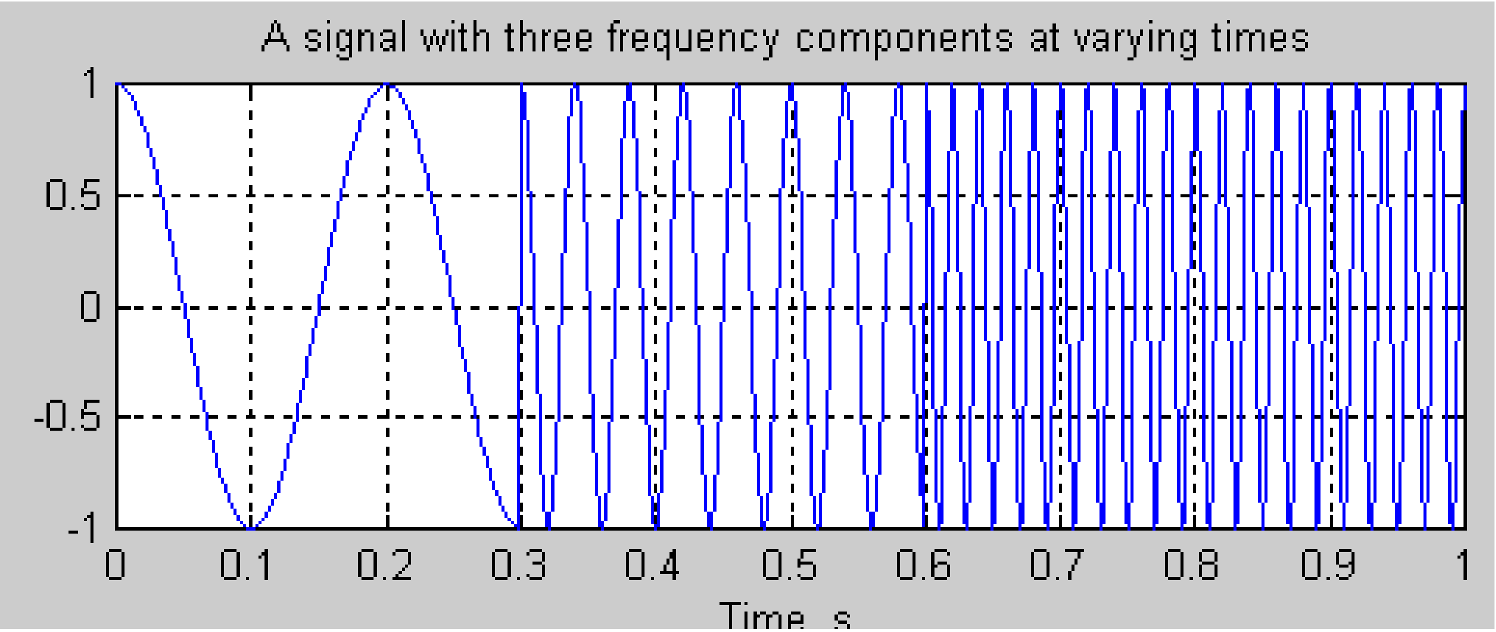
The three frequency components, are present at all times!




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

AALBORG UNIVERSITY
DENMARK

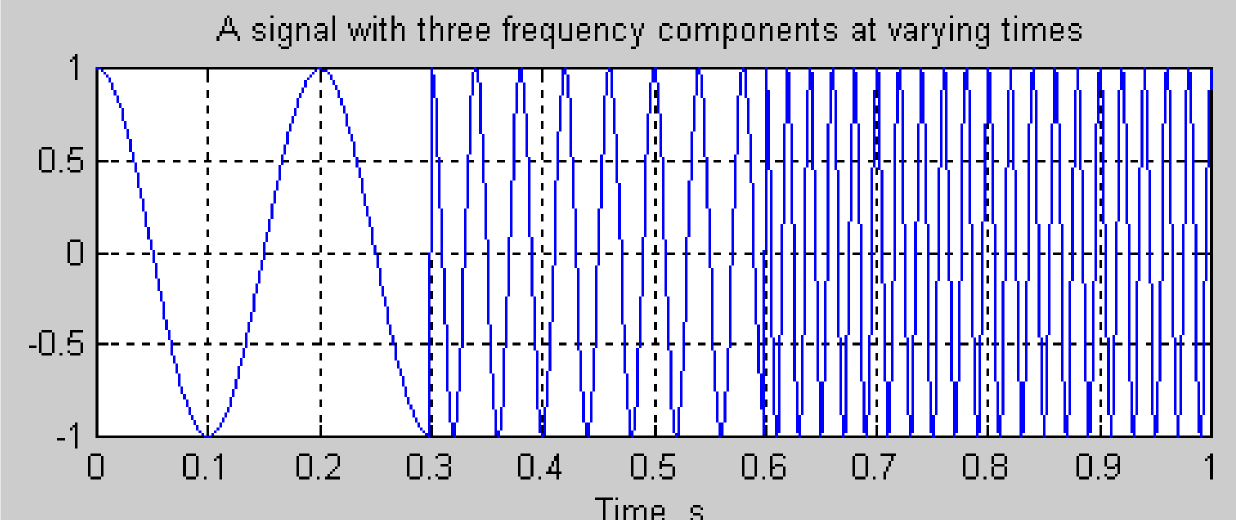
## Now, let's append the three sinusoids such that they occur distinctive in time



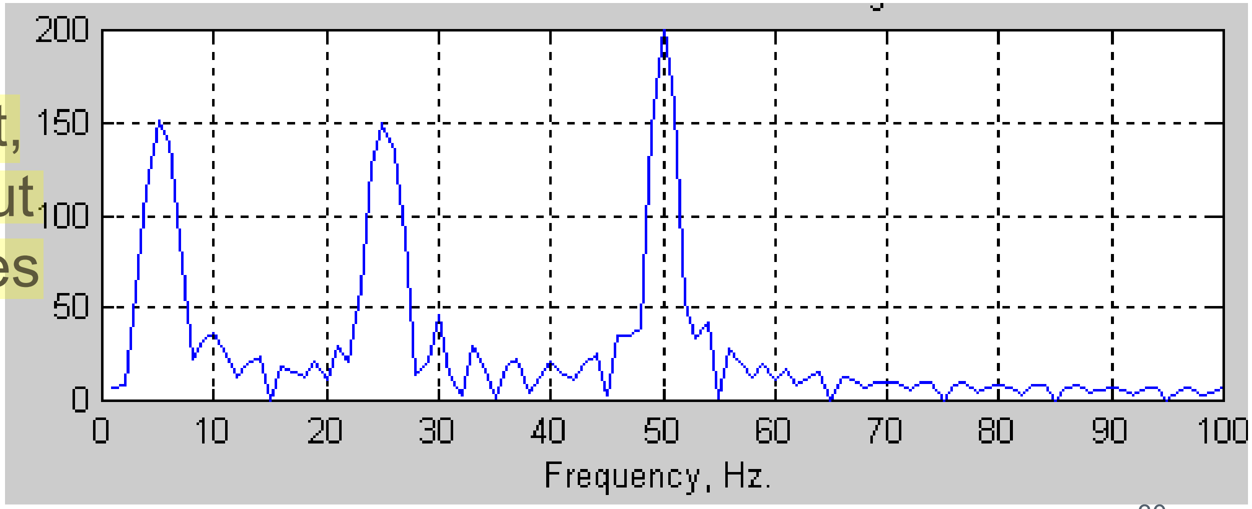

Depending on WHEN you perform the Fourier Analysis, you will see different results…

# Spectral analysis of Time-varying signals

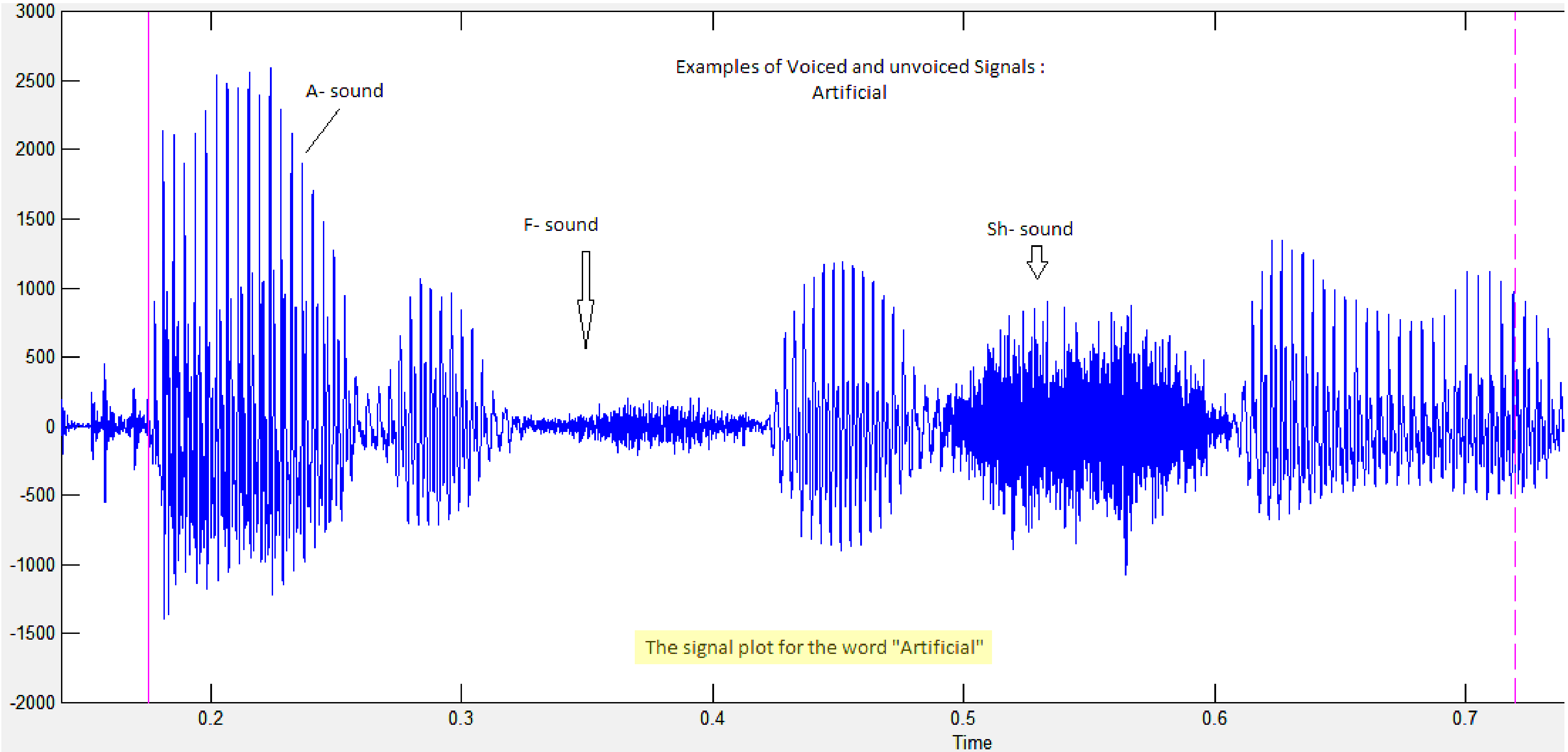
The three frequency components are NOT present at all times!




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



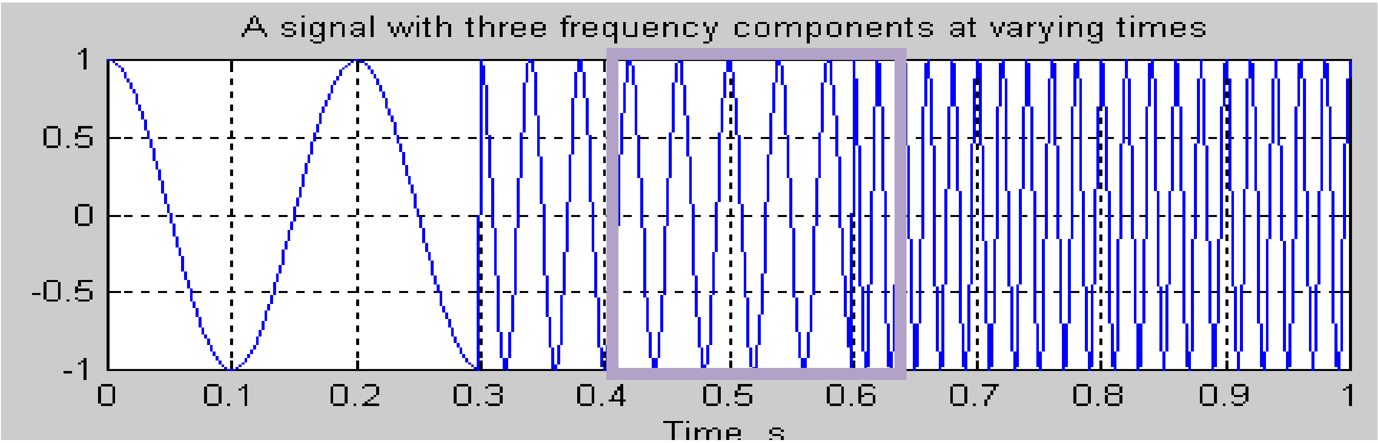
# This is the normal situation for real-life signals…




AALBORG UNIVERSITY
DENMARK

# Therefore, we need (yet) another Fourier Transform

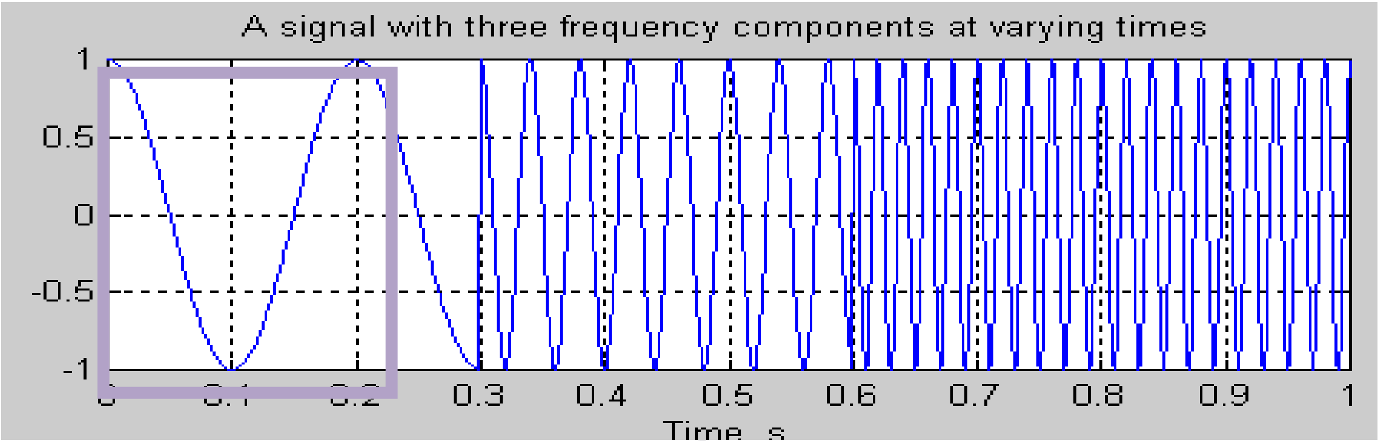
- Segment the signal into short time intervals (i.e., short enough for the signal to be considered time-invariant) and then calculate the FT of each segment.
- Each FT provides the spectral information of a separate time-slice of the signal, thus providing simultaneous time and frequency information.




AALBORG UNIVERSITY
DENMARK

# 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 FFT of the truncated signal, save results.
(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

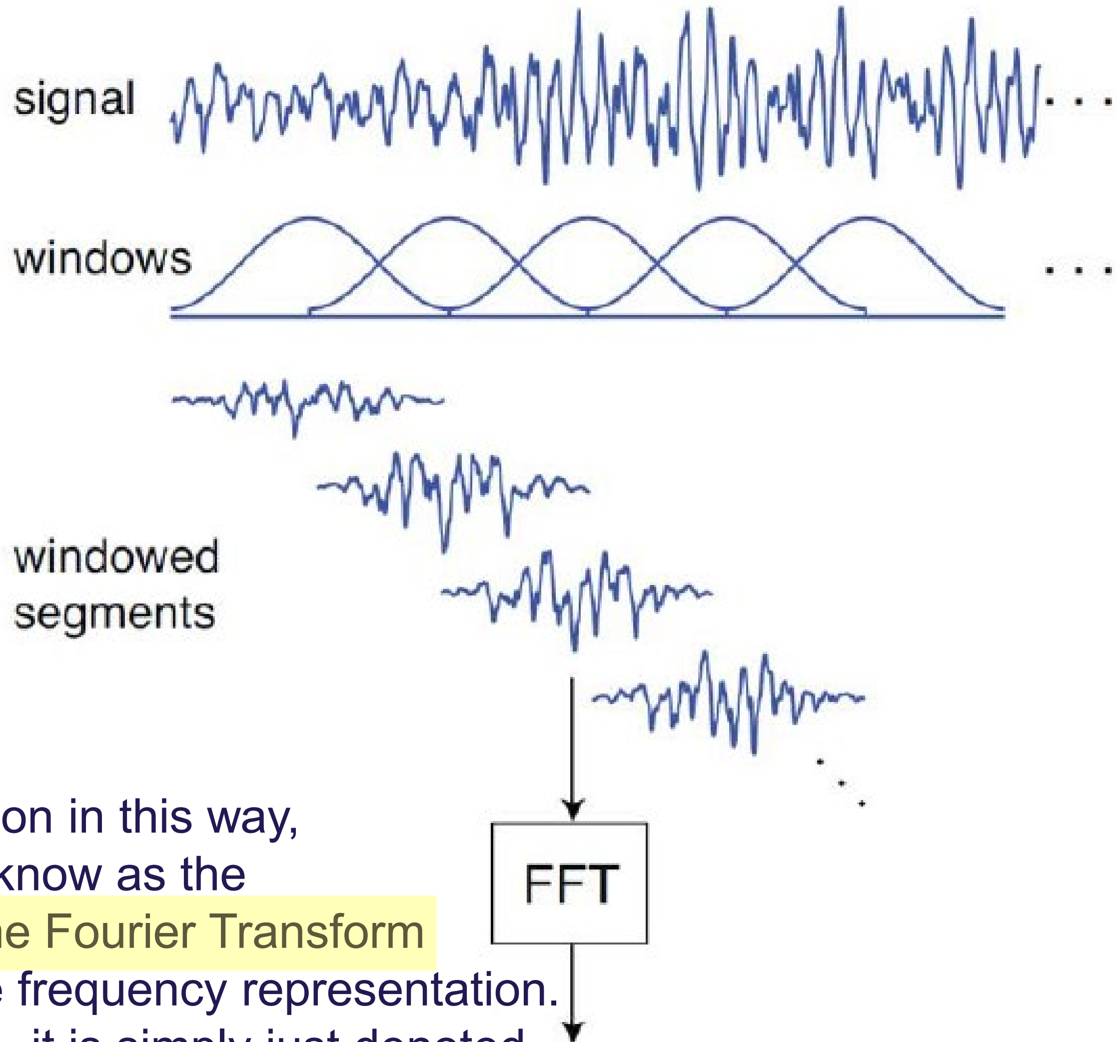
Window function

Located at the beginning of the segment

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

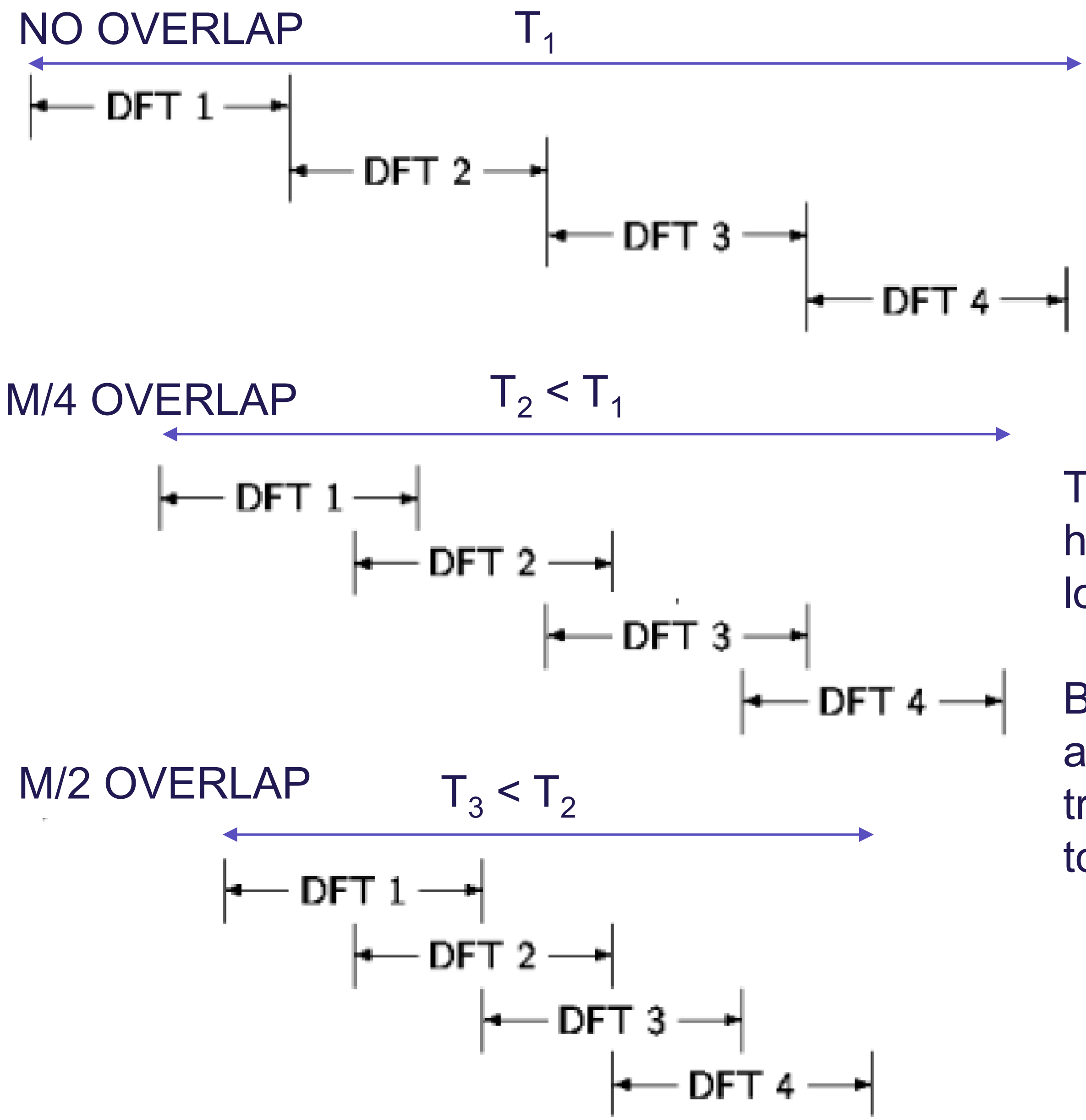
$$\sum_{m=0}^{M-1} x[n+m]w[m]e^{-j\lambda m}$$

# 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…

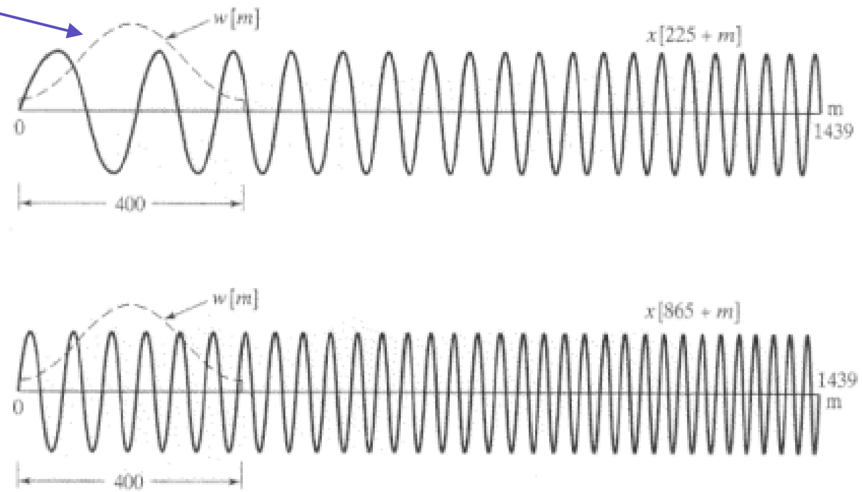
But at the same time, we also obtain a smoother transition from one segment to the next.

# An example - a linear frequency sweep

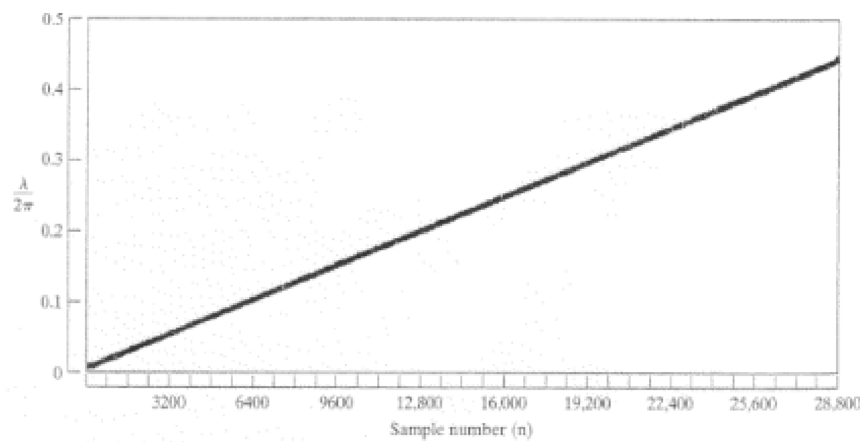
$$x[n] = \cos(\omega_0 n^2),\ \omega_0 = 2\pi \cdot 7.5 \cdot 10^{-6}$$

Hamming window
with length N=400

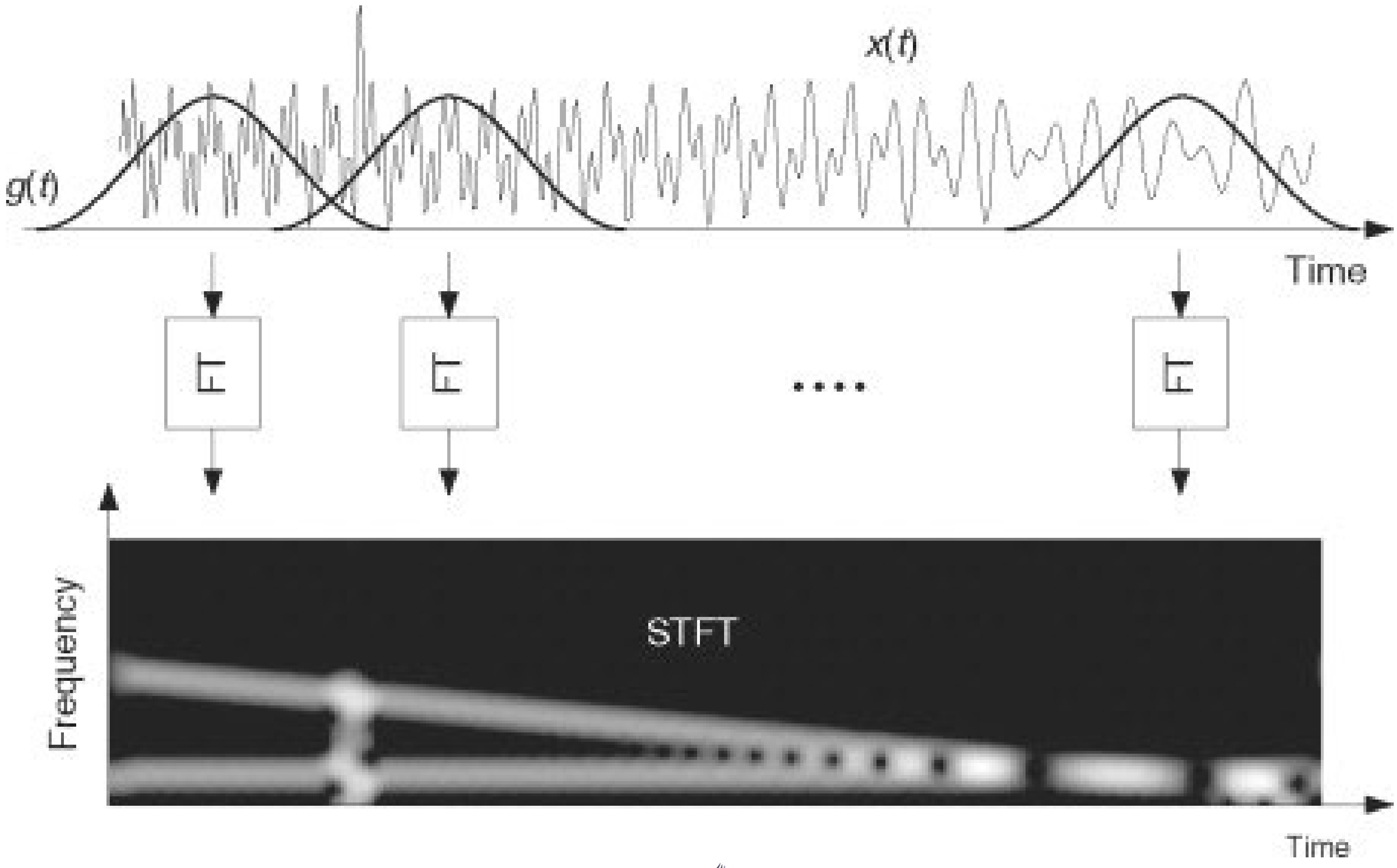
Time:




Time/
frequency:

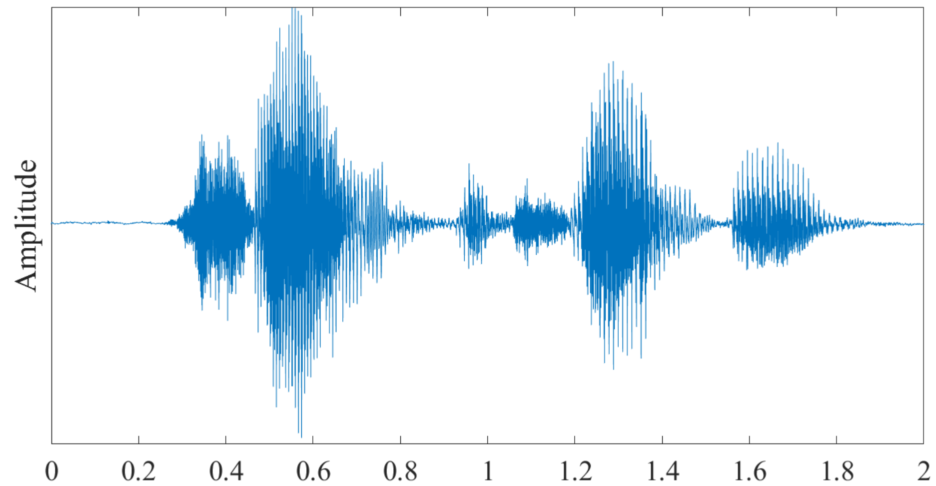


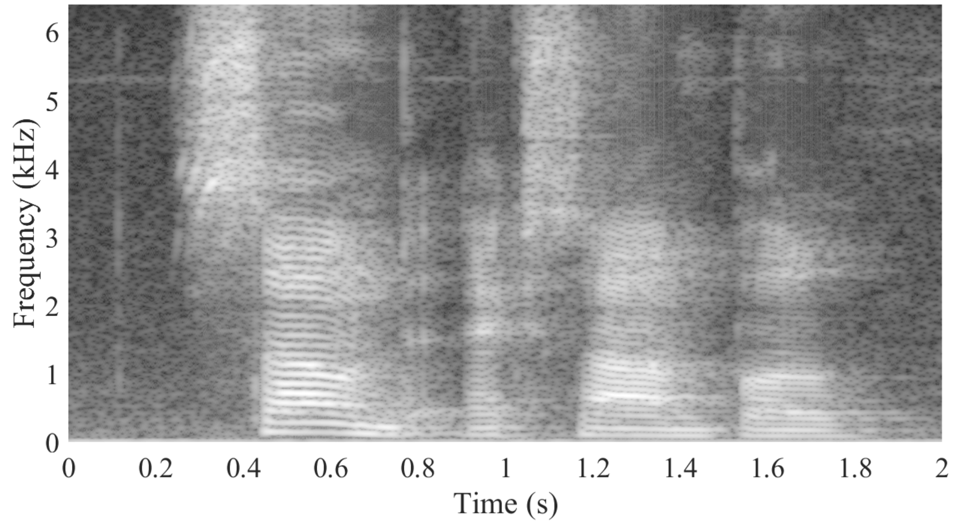
# Simultaneous Time- and Frequency representation - The Spectrogram




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

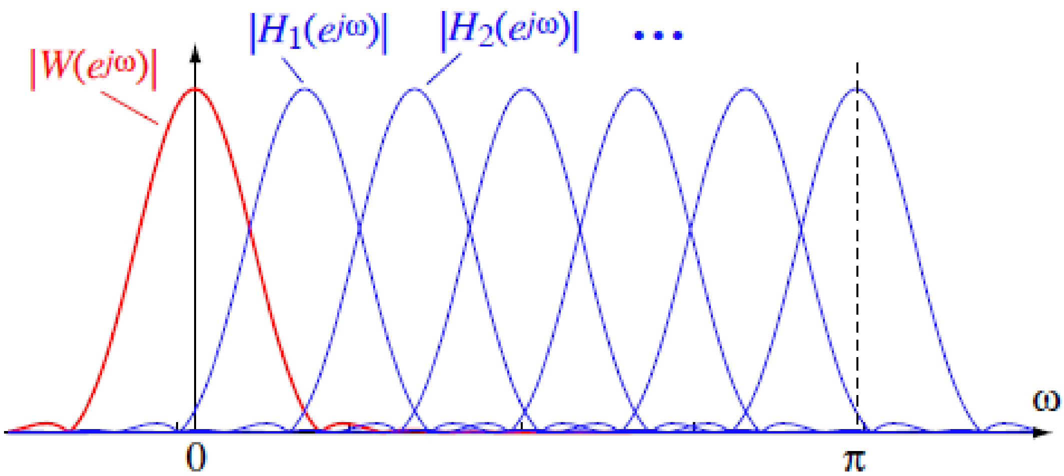
$$h_\lambda[n] = w[-n]e^{j\lambda n}.$$

we see that the time-dependent Fourier transform as a function of $n$ with $\lambda$ 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 lowpass filter, i.e., bandpss filters with different center frequencies $\lambda_k$ and amplitude responses defined by the window…

# STFT – The Filter Bank interpretation




A filter bank consisting of identical frequency-shifted bandpass filters

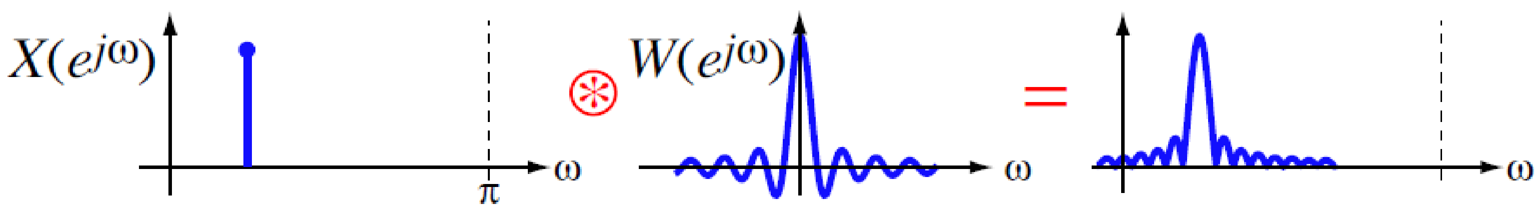
The more narrow the main-lope of the window function is, and the more values of $\lambda$ we choose, the better a spectral estimation we get…

# 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…

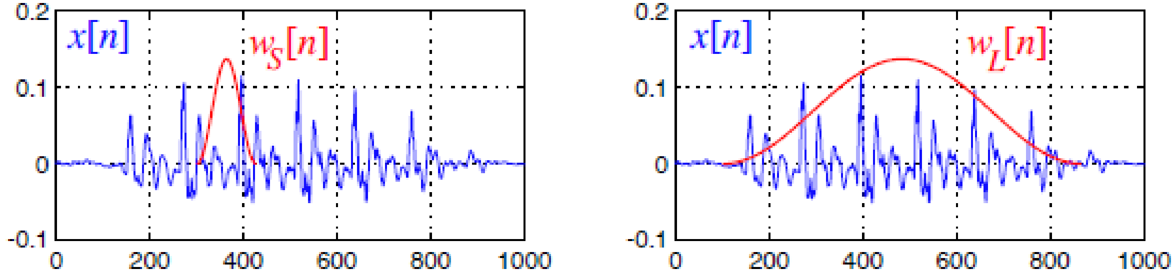
- e.g. if $x[n]$ is a pure sinusoid,




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

# Choosing a window

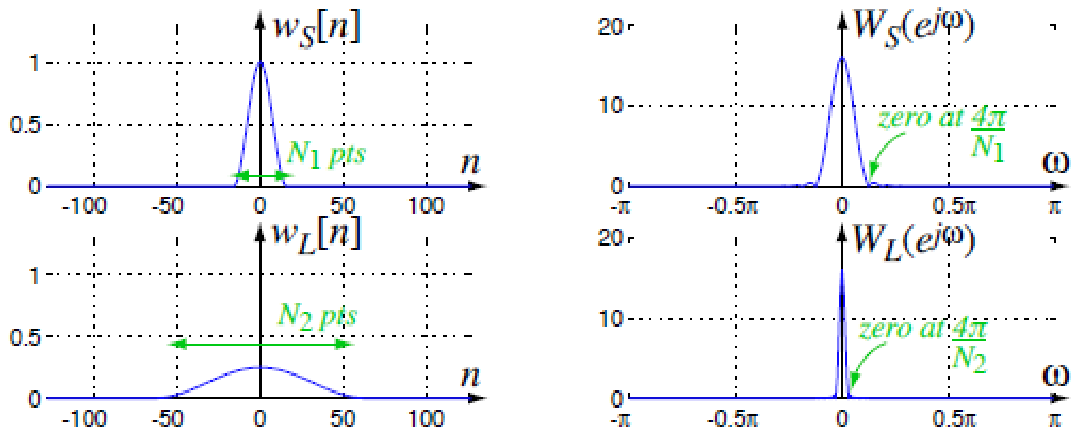
- Length of $w[n]$ sets temporal resolution




short window measures only local properties

longer window averages spectral character

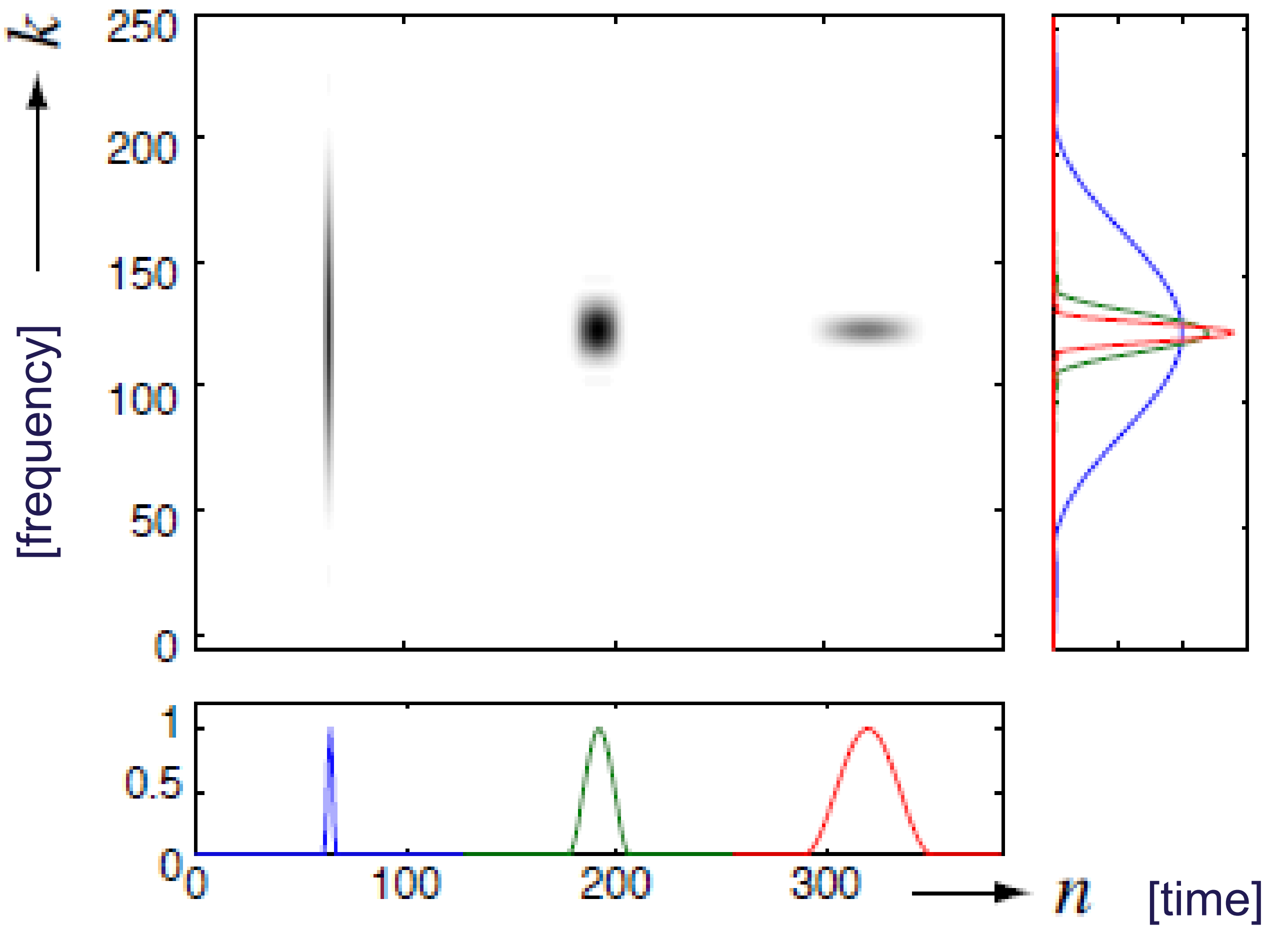
- Window length $\propto$ 1/(Mainlobe width)




shorter window → more blurred spectrum

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

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




AALBORG UNIVERSITY
DENMARK

# Choosing a window

- 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.

# Heisenberg's Uncertainty Principle

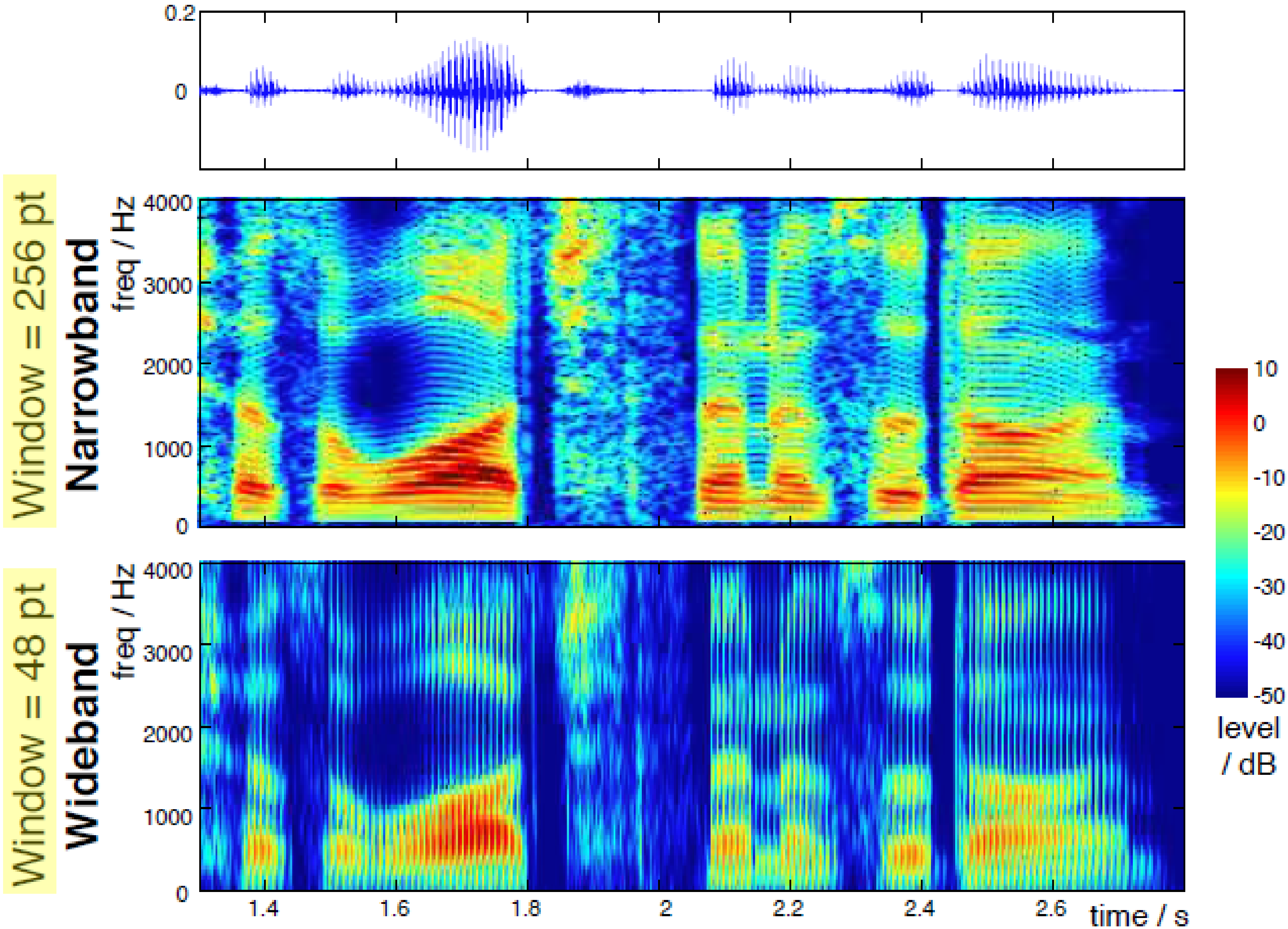
$$\Delta t \cdot \Delta f \geq \frac{1}{4\pi}$$

**Time resolution:**
How well two spikes in time can be separated from each other in the frequency domain.

**Frequency resolution:**
How well two spectral components can be separated from each other in the time domain

$\Delta t$ and $\Delta f$ cannot AT THE SAME TIME be made arbitrarily small

# Narrowband vs. Wideband Spectral Analysis



# Narrowband vs. Wideband Spectral Analysis

- For a long window $w[n]$, the result is the narrowband spectrogram, which exhibits the harmonic structure in the form of horizontal striations
- For a short window $w[n]$, the result is the wideband spectrogram, which exhibits periodic temporal structure in the form of vertical striations