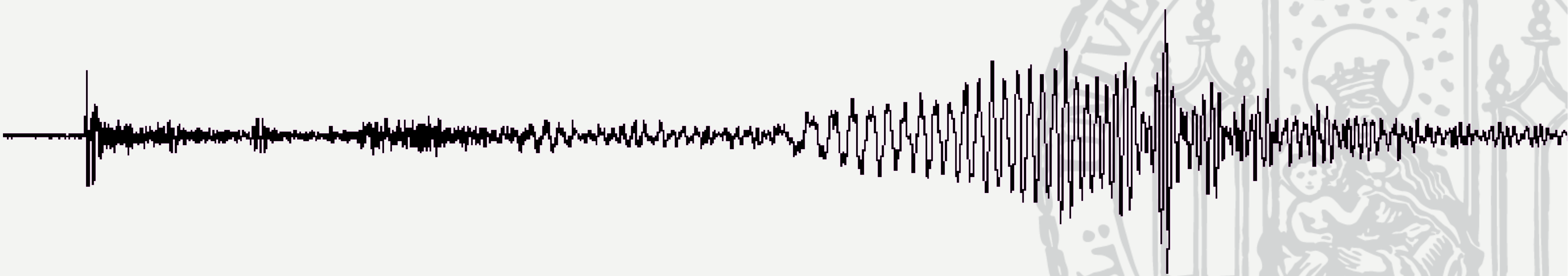
Celine Hadziioannou

Geophysical Data Analysis

L05 – Spectral analysis



Applications of the

Fourier Transform

Moving from the continuous to the discrete world.

Fourier transform

contributions to frequency ω

Add all

for each time

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

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

Add all

for each frequency

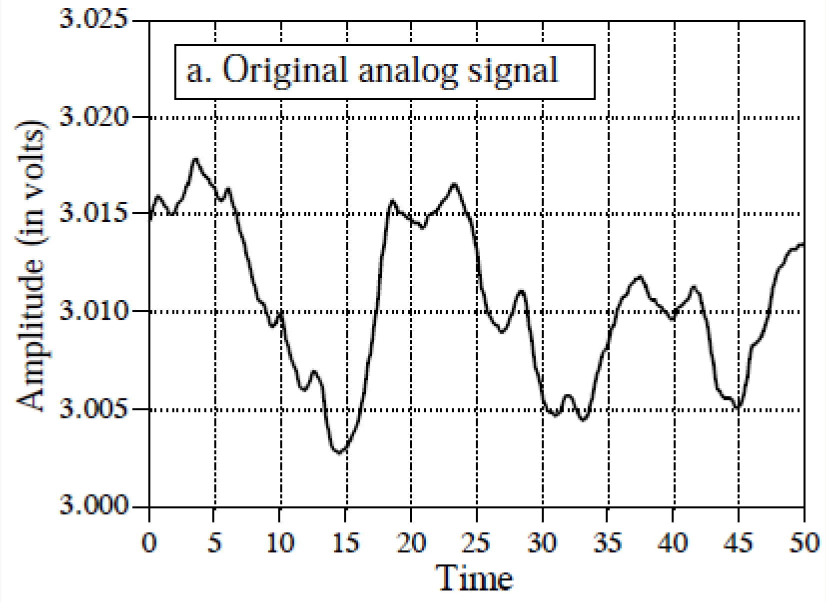
contributions to time t

Forward transform
Time domain → frequency domain
(analysis)

Inverse transform
Frequency domain → time domain
(synthesis)

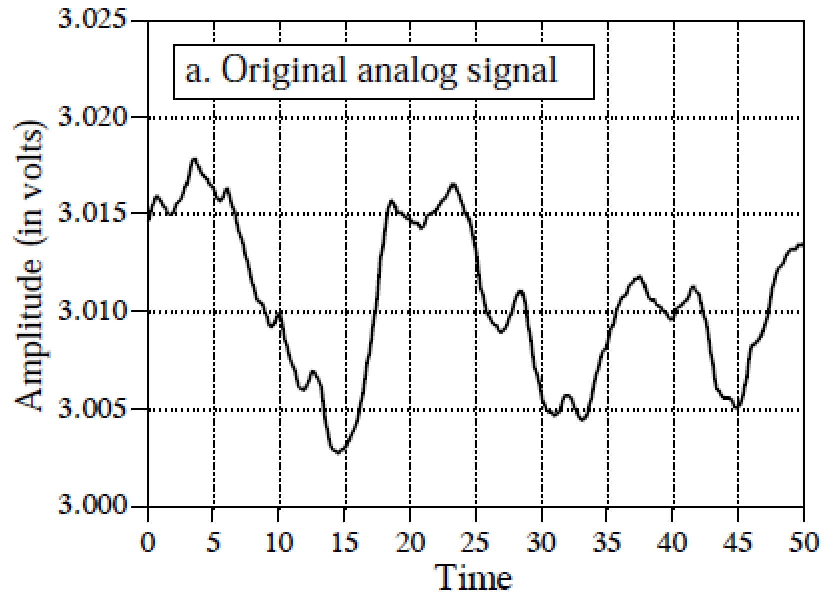
The real, digital world

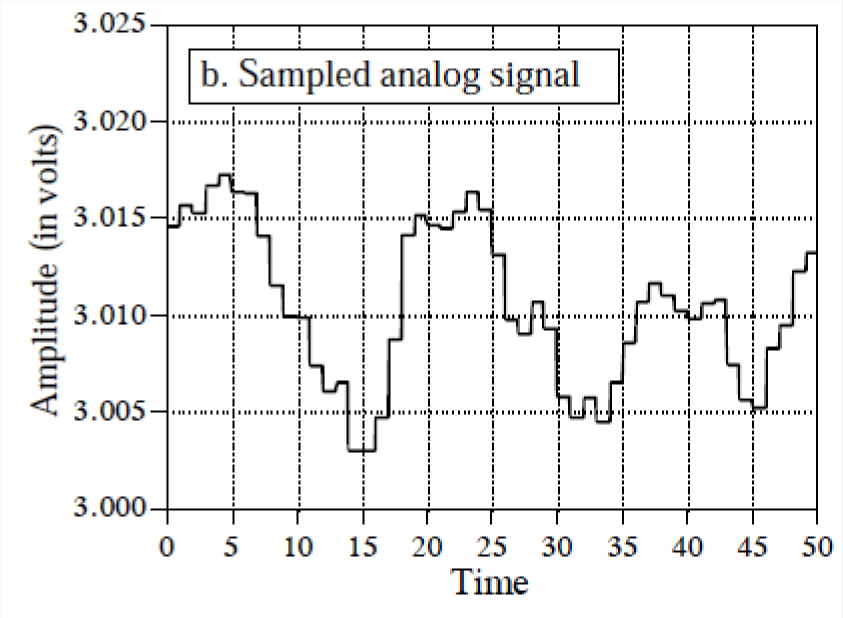
… is NEITHER periodic NOR continuous!



The real, digital world

… is NEITHER periodic NOR continuous!





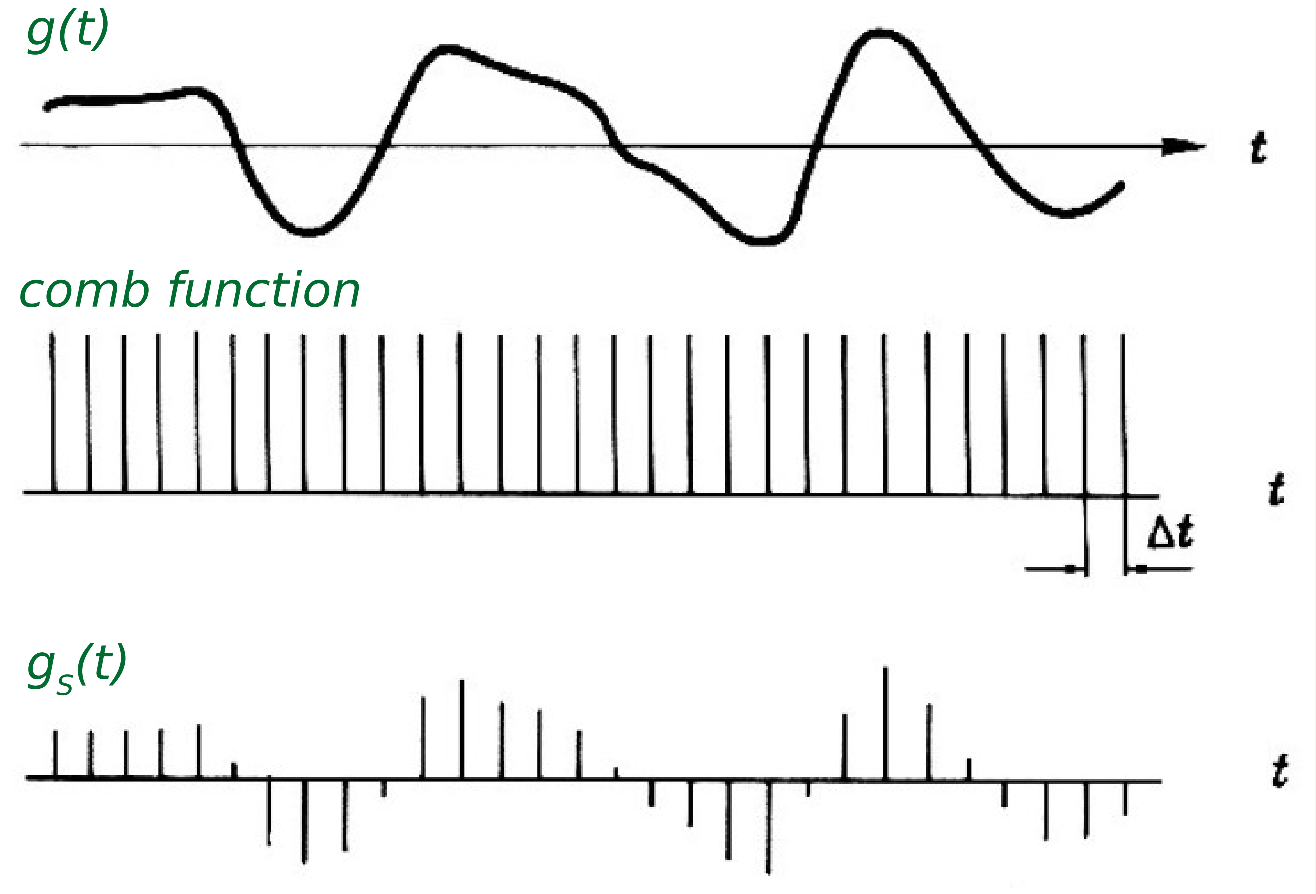
Discretize

The real, digital world

… is NEITHER periodic NOR continuous!

$$g_S(t) = g(t) \sum_{j=-\infty}^{\infty} \delta(t - j\,dt)$$

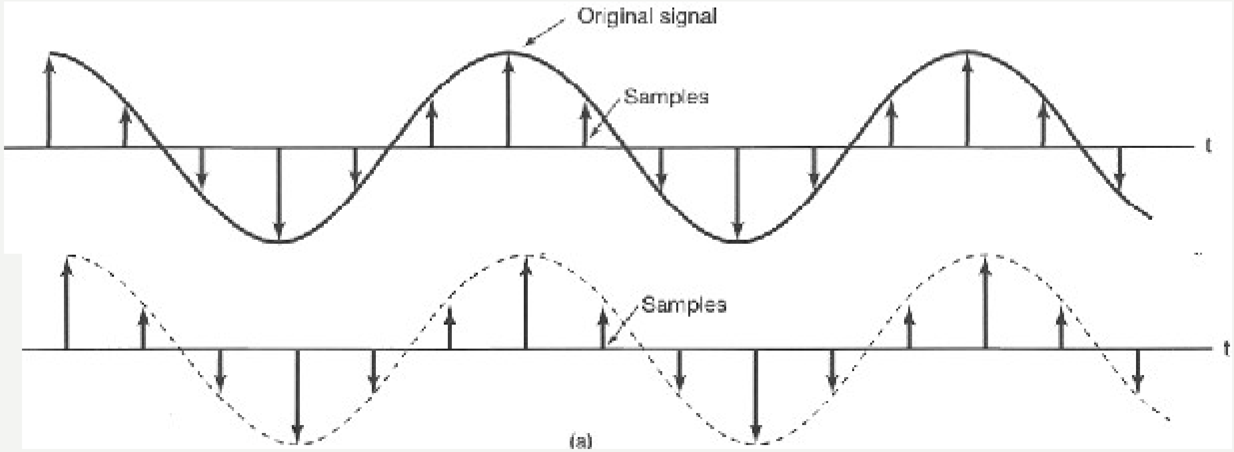
comb function



$g_s(t)$ is the digitized version of $g(t)$

time signal is multiplied with comb function.

Sampling frequency



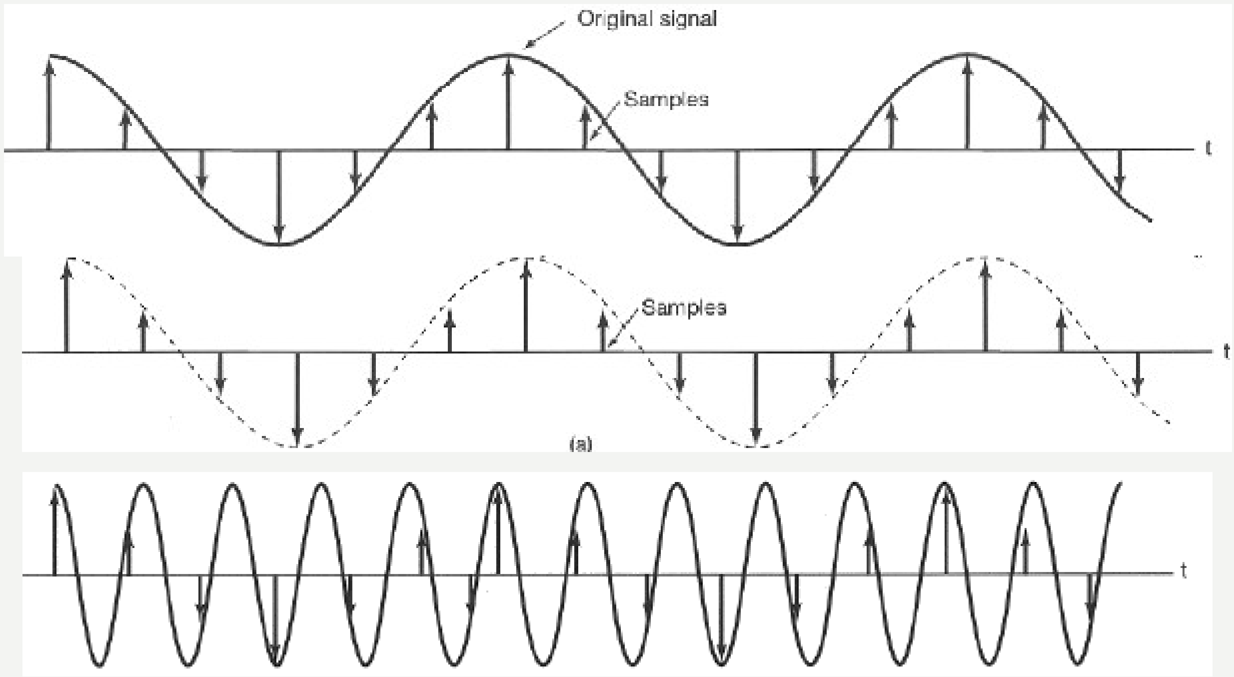
$$x[n] = x_{cont}(nT)$$

T = sampling period

ωs = 2π/T sampling frequency in [rad]

Fs = 1/T sampling frequency or 'sample rate', in [Hz]

Sampling frequency



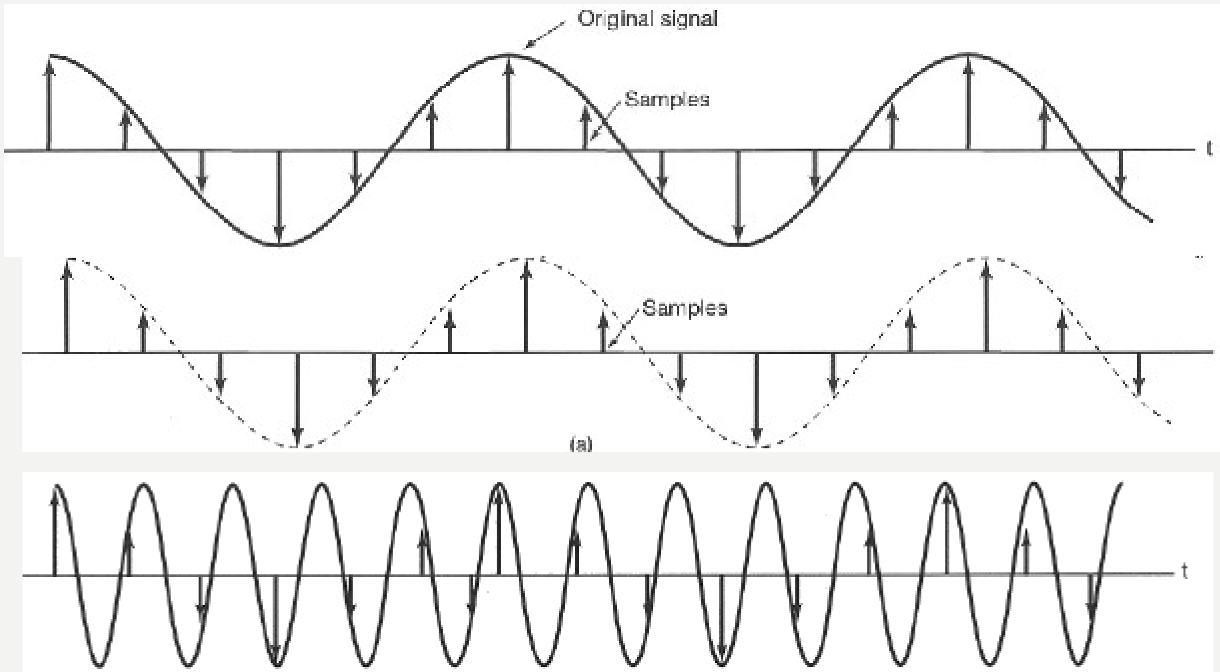
$$x[n] = x_{cont}(nT)$$

T = sampling period

ωs = 2π/T sampling frequency in [rad]

Fs = 1/T sampling frequency or 'sample rate', in [Hz]

Sampling frequency



Sampled version of signal could also be fitted with higher frequency signal

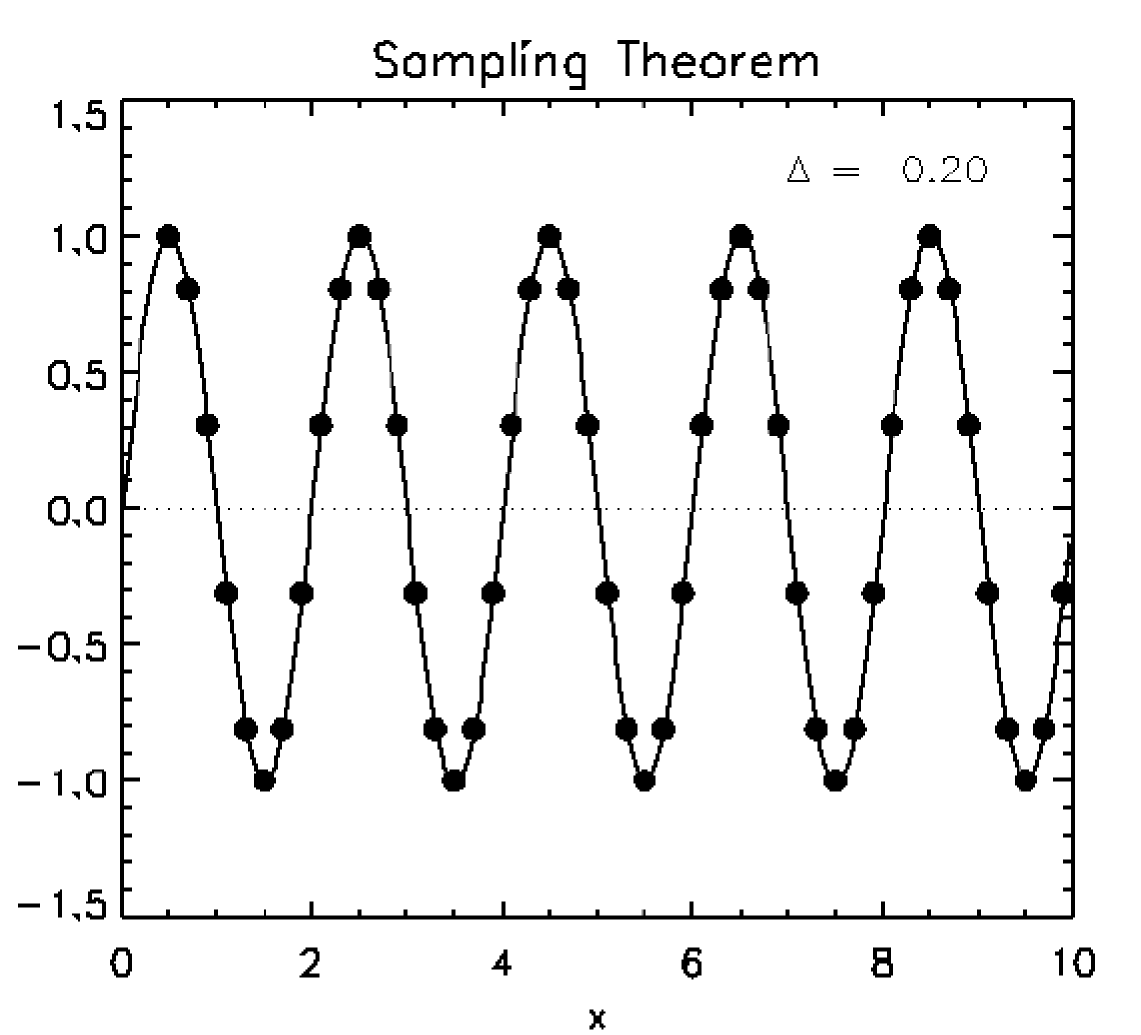
$$x[n] = x_{cont}(nT)$$

T = sampling period

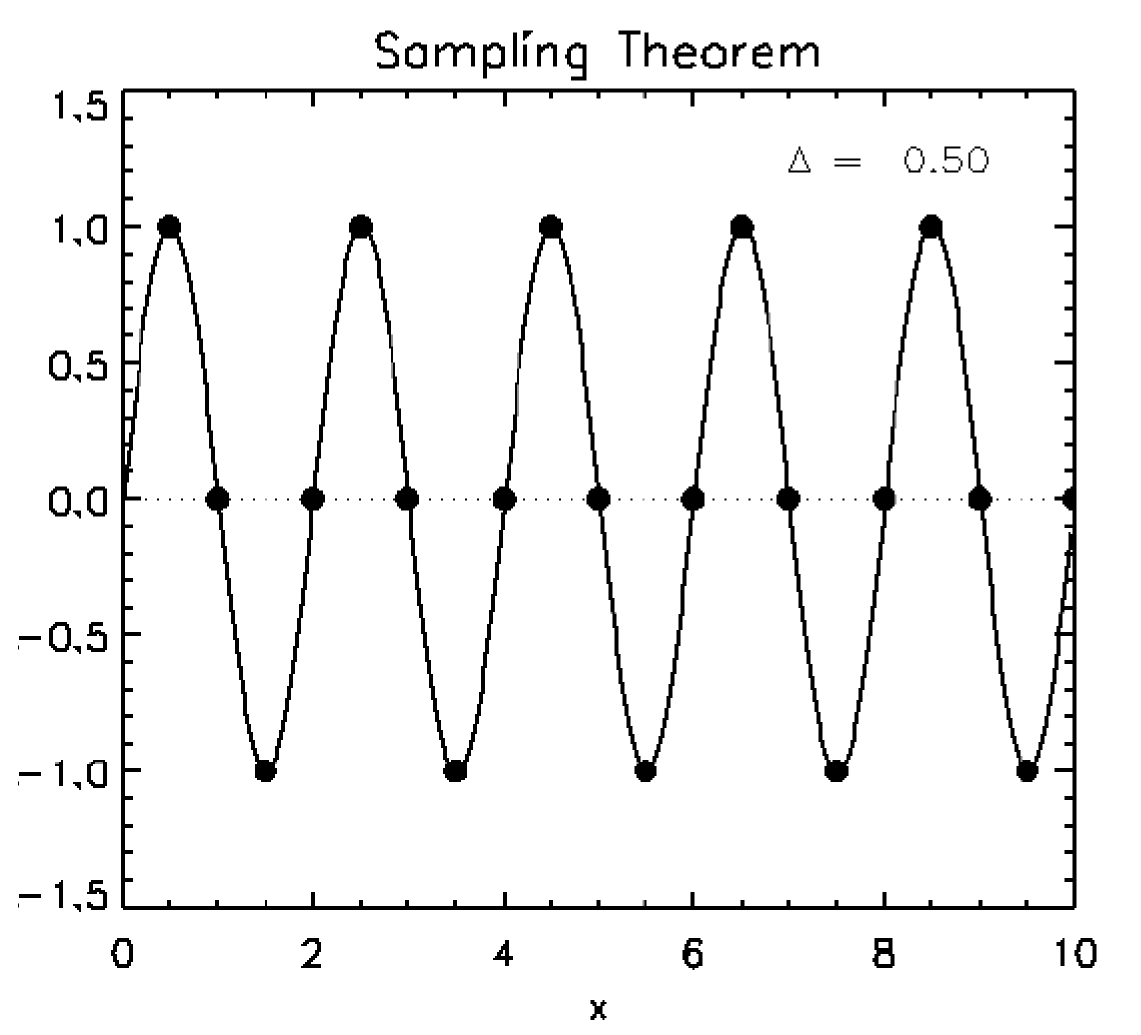
$\omega s = 2\pi/T$ sampling frequency in [rad]

$Fs = 1/T$ sampling frequency or 'sample rate', in [Hz]

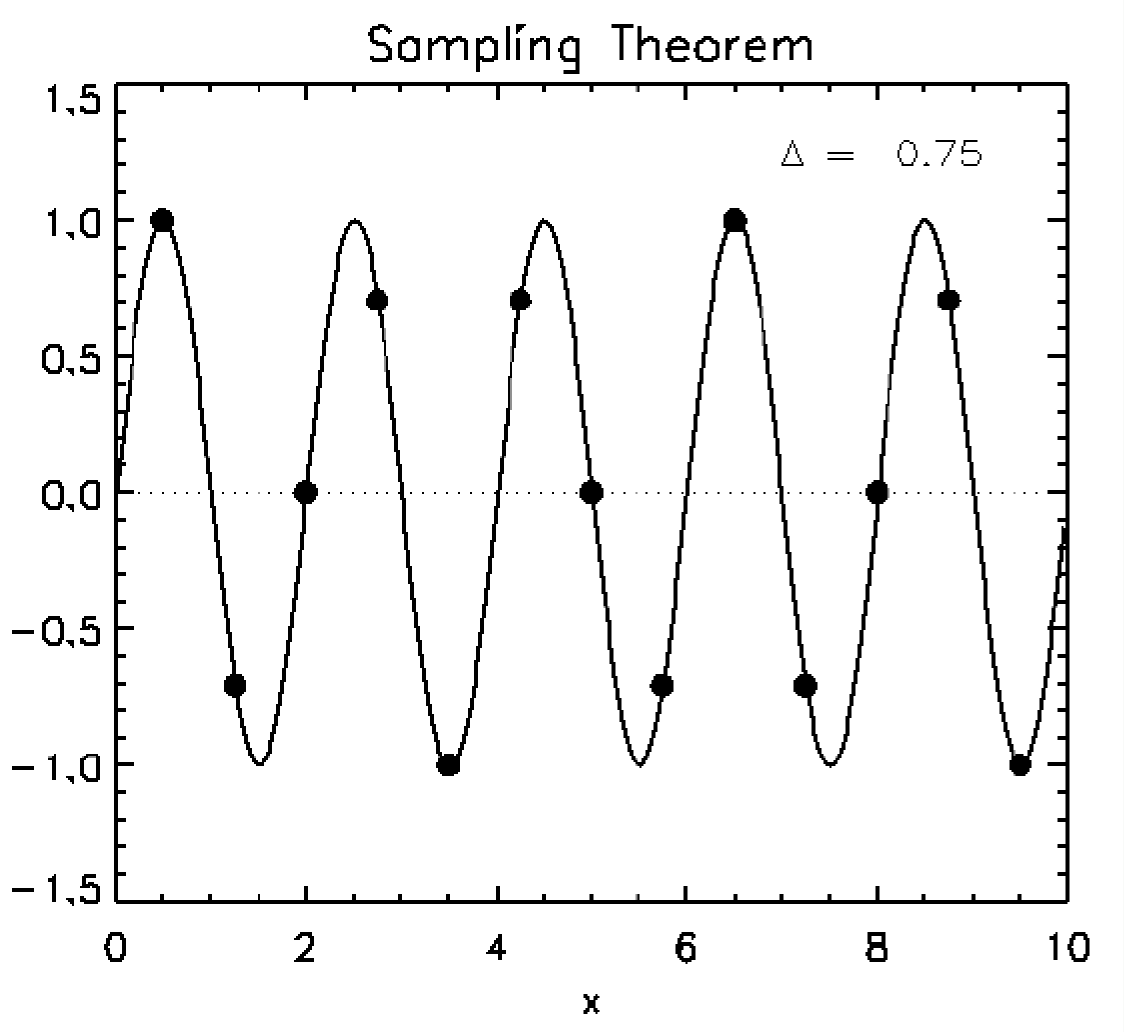
Sampling frequency



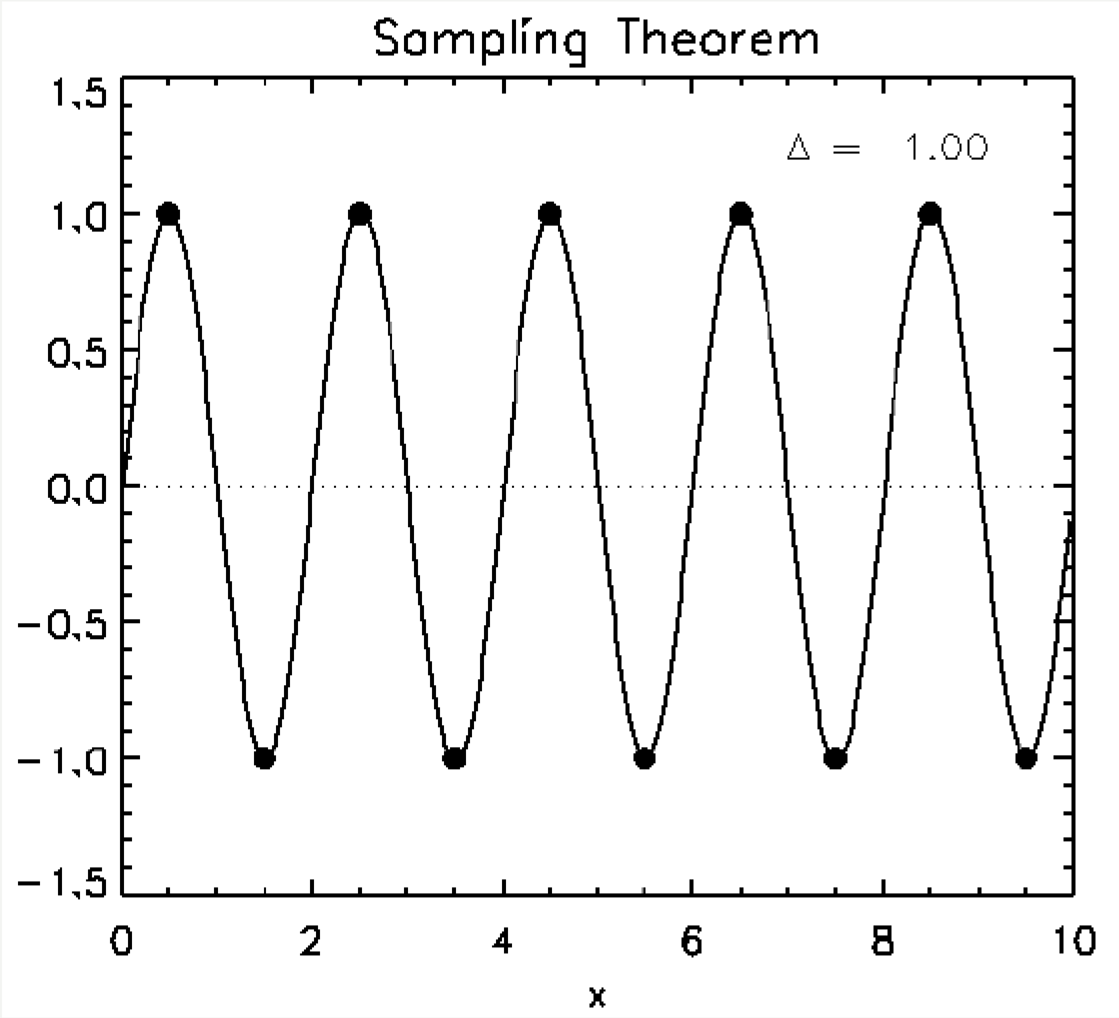
Sampling frequency



Sampling frequency

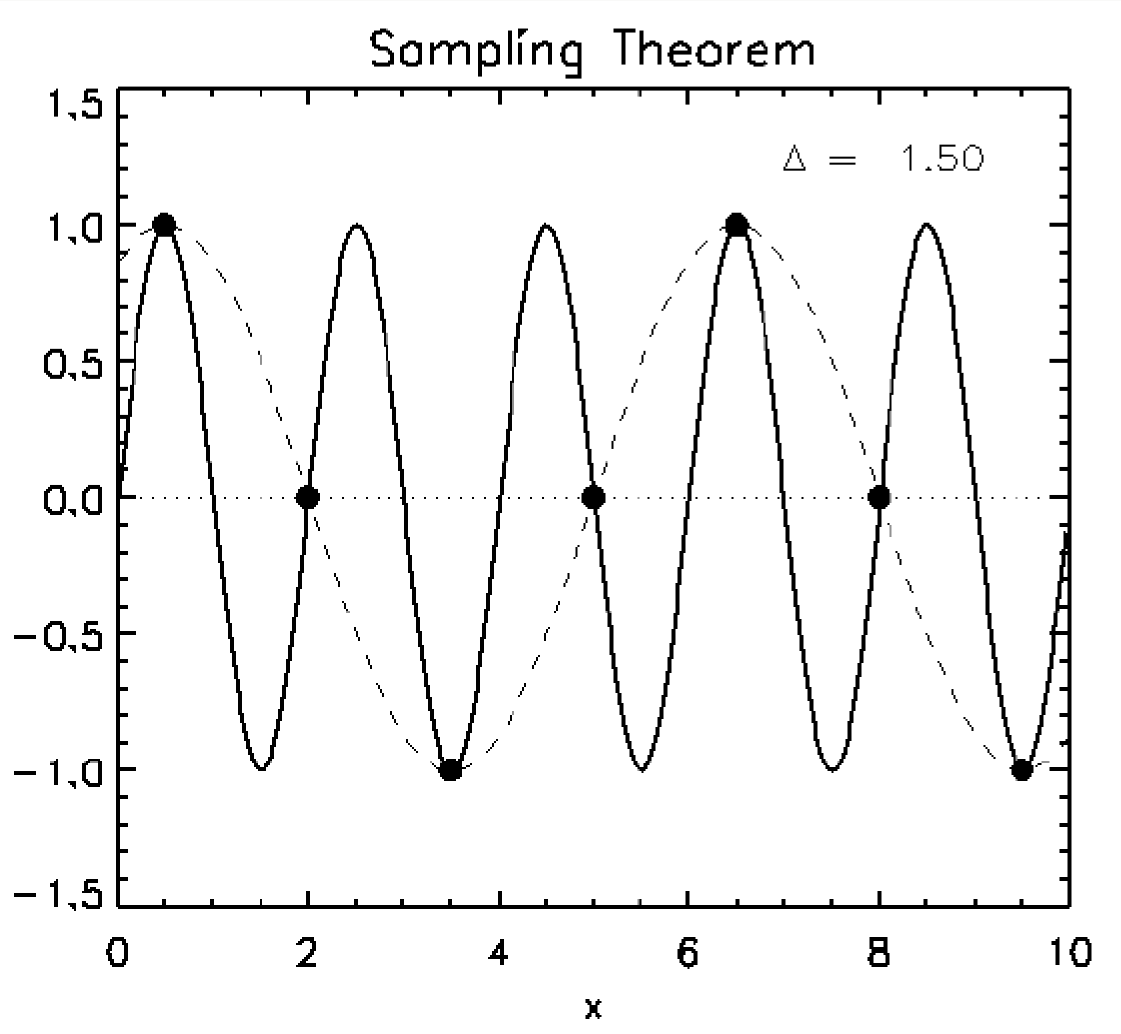


Sampling frequency



Critical sampling frequency!

Sampling frequency



Undersampled..

Sampling theorem

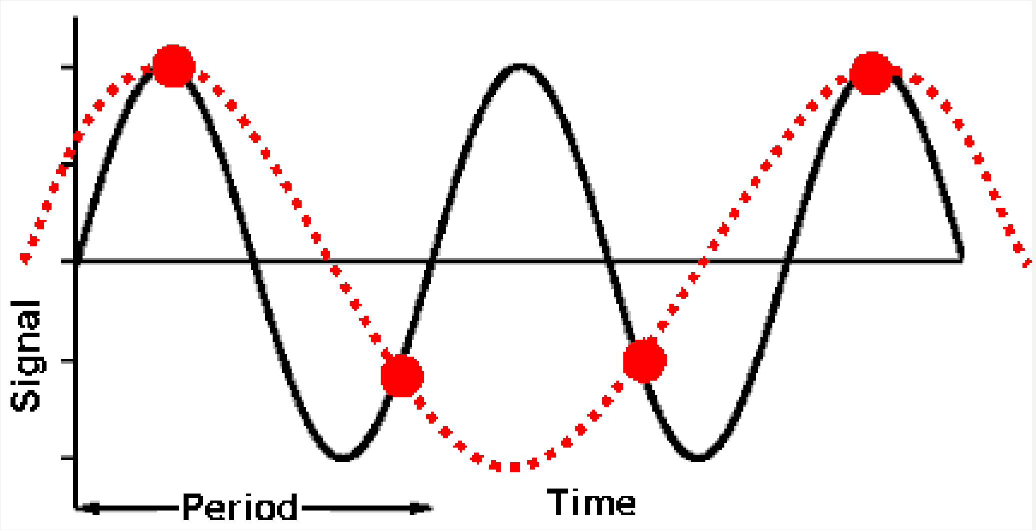
*At least **2 samples per period** are needed to correctly reproduce the **highest frequency** of a signal*

OR

a continuous signal can be properly sampled, only if it does not contain frequency components **above one-half of the sampling rate**

"Nyquist frequency"

$$f_{Ny} = \frac{1}{2\,dt}$$



*Otherwise: **aliasing** !!!*

"wrong" frequencies appear in your signal

Sampling theorem

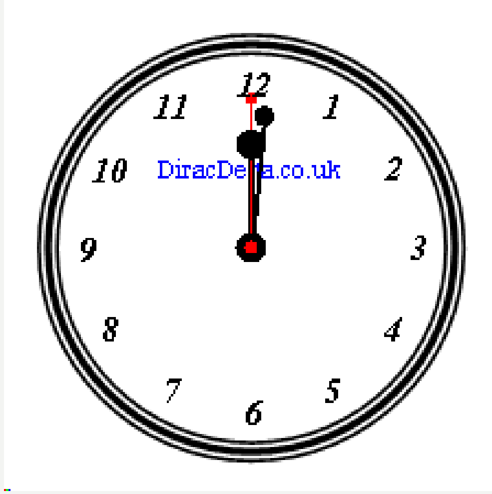
*At least **2 samples per period** are needed to correctly reproduce the **highest frequency** of a signal*

OR

a continuous signal can be properly sampled, only if it
does not contain frequency components **above one-half of the sampling rate**

“Nyquist frequency”

$$f_{Ny} = \frac{1}{2\,dt}$$



frequency of view: 50 minutes

Sampling theorem

*At least **2 samples per period** are needed to correctly reproduce the **highest frequency** of a signal*

OR

a continuous signal can be properly sampled, only if it

does not contain frequency components **above one-half of the sampling rate**

"Nyquist frequency"

$$f_{Ny} = \frac{1}{2\,dt}$$

Audio CD sampling frequency: 44.1 kHz

Why?

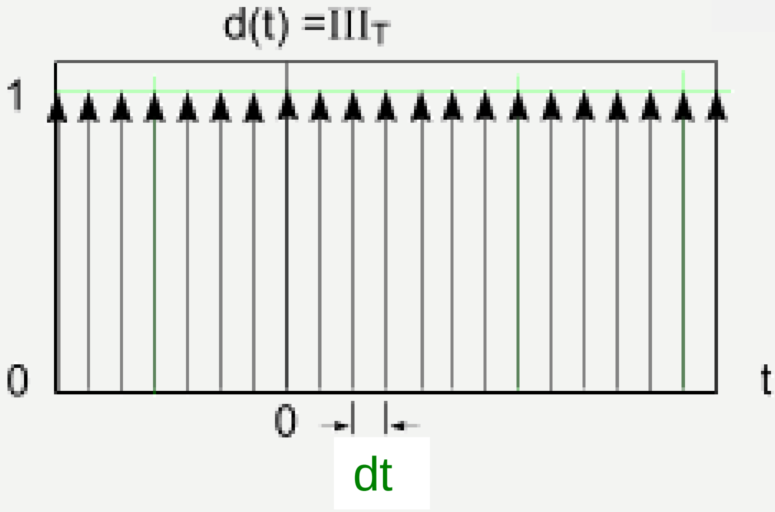
The sampling theorem

… back to the comb function …

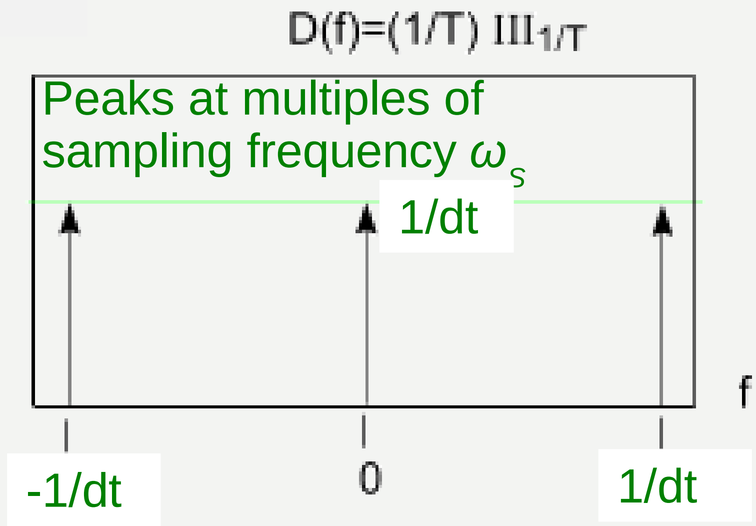
$$g_S(t) = g(t) \sum_{j=-\infty}^{\infty} \delta(t - j\,dt)$$

With the Nyquist frequency defined as

$$f_{Ny} = \frac{1}{2\,dt}$$



*xs[n] = xc(t) * time domain comb*

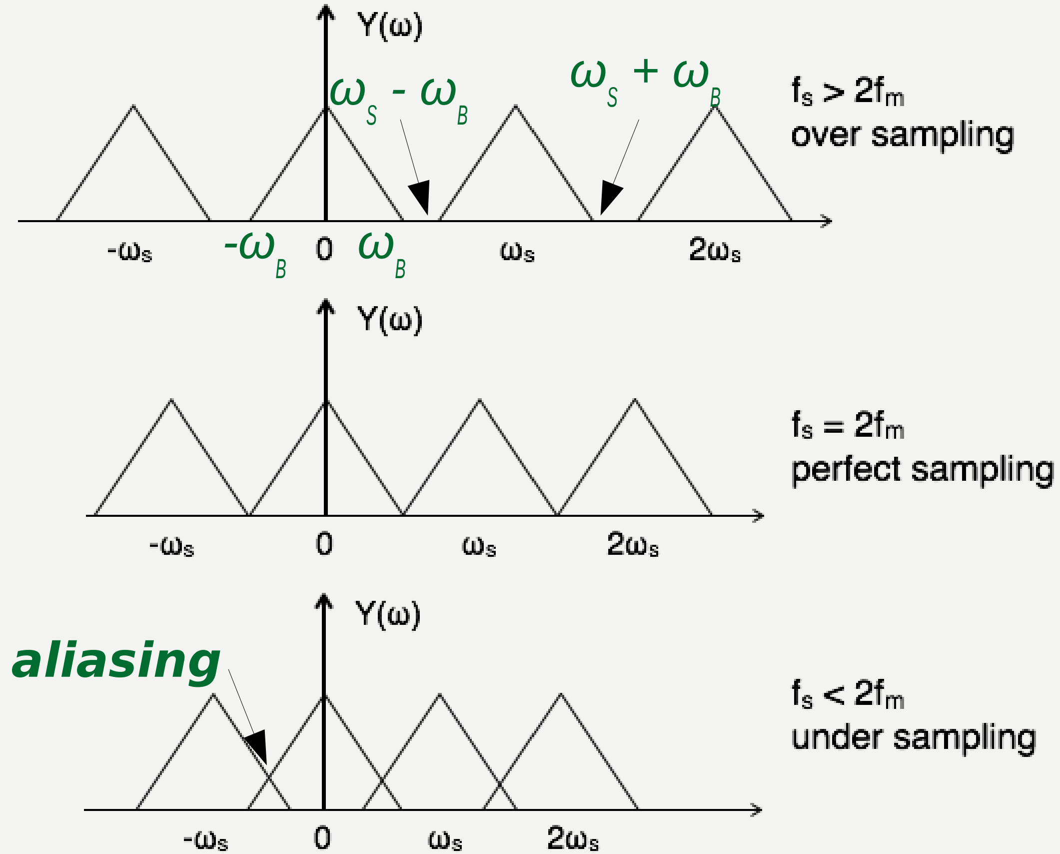


*Xs[ω] = Xc(ω) * frequency domain comb*

The sampling theorem

⇒ For the calculation of the spectrum of the frequency f, there are also contributions of frequencies f ± 2nf_{Ny}

That means dt has to be chosen such that f_{Ny} is the largest frequency contained in the signal.



The sampling theorem

⇒ For the calculation of the spectrum of the frequency f, there are also contributions of frequencies $f \pm 2nf_{Ny}$

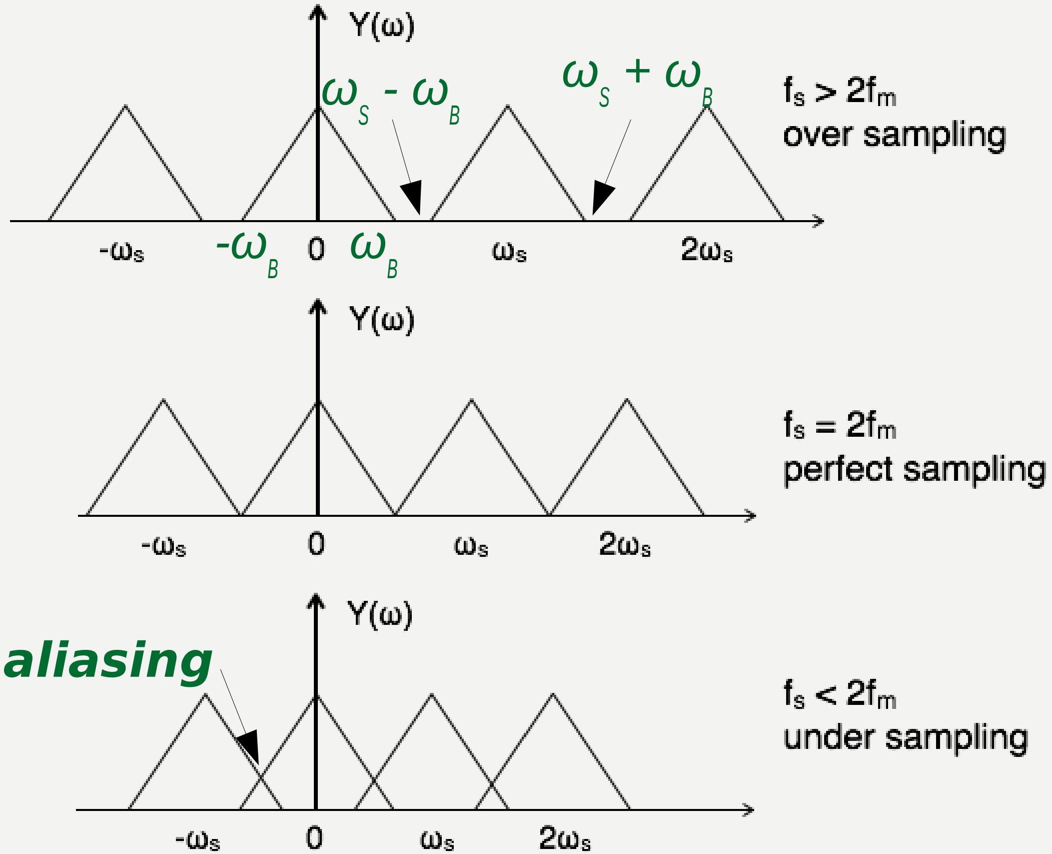
That means dt has to be chosen such that f_{Ny} is the largest frequency contained in the signal.

For a band-limited signal where the spectrum is between +- ω_B:

Don't want the 'copies' to overlap:

$$\omega_S - \omega_B > \omega_B$$

$$\omega_S > 2\omega_B$$



The sampling theorem

For the calculation of the spectrum of the frequency f, there are also contributions of frequencies $f \pm 2nf_{Ny}$

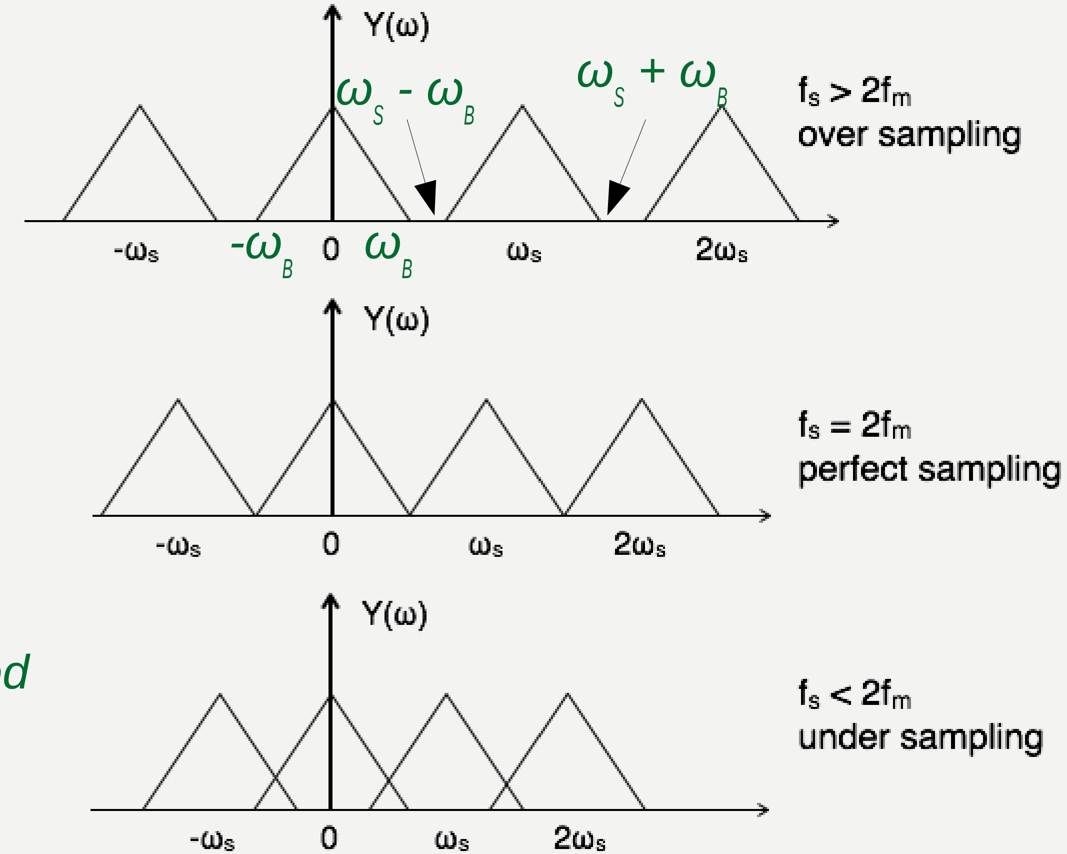
That means dt has to be chosen such that f_{Ny} is the largest frequency contained in the signal.

For a band-limited signal where the spectrum is between $\pm\ \omega_B$:

Don't want the 'copies' to overlap:

$$\omega_S - \omega_B > \omega_B$$

$$\omega_S > 2\omega_B$$

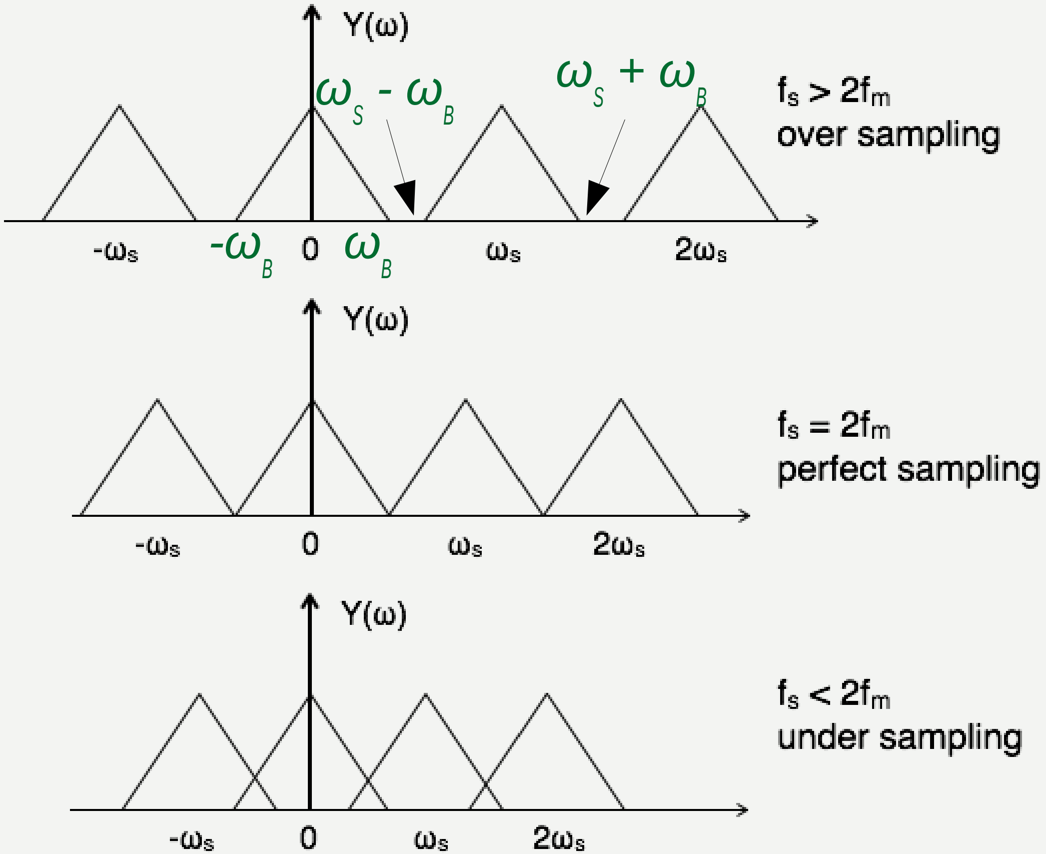


Sampling theorem: *A band-limited signal with maximum frequency ω_B can be reconstructed perfectly from evenly spaced samples if the sampling frequency ω_s satisfies $\omega_S > 2\omega_B$*

The sampling theorem

"A continuous signal can be properly sampled, only if it does not contain frequency components above one-half of the sampling rate."

What if the signal contains higher frequencies as well?

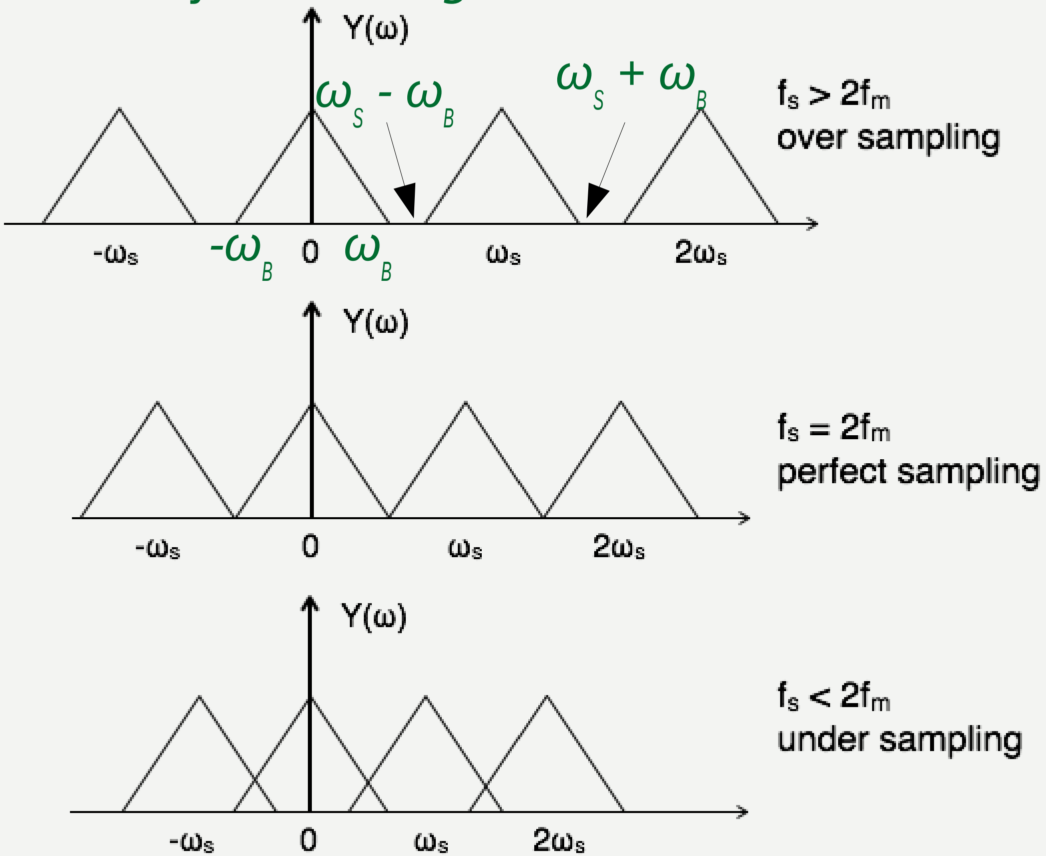


The sampling theorem

“A continuous signal can be properly sampled, only if it
does not contain frequency components above one-half of the sampling rate.”

What if the signal contains higher frequencies as well?

Solution: remove all frequencies above fNy → *filtering*



The sampling theorem

"A continuous signal can be properly sampled, only if it
does not contain frequency components above one-half of the sampling rate."
What if the signal contains higher frequencies as well?

Solution: remove all frequencies above fNy → *filtering*

Resampling (Decimating):

- Often is it useful to down-sample a time series

(e.g. from 100Hz to 1 Hz when studying surface waves)

- Then, the time series needs to be pre-processed.

  All frequencies above twice the new sampling interval have to be filtered out before !!!

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

We now have discrete signals..

.. back to the Fourier Transform

The Fourier transform

Whatever we do on the computer with data will be based on the discrete Fourier transform.

discrete

$$F_k = \frac{1}{N} \sum_{j=0}^{N-1} f_j \, e^{-2\pi i k j / N}$$

$$f_k = \sum_{j=0}^{N-1} F_j \, e^{2\pi i k j / N}$$

$$k = 0, 1, ..., N-1$$

continuous

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

$$f(t) = \int_{-\infty}^{\infty} f(\omega) e^{i\omega t} d\omega$$

Discrete Fourier Transform

- *j measures time in units of sampling interval*

$j = k\Delta t$ *up to maximum time* $T = N\Delta t$

- *k measures frequency in intervals of the sampling frequency*

$\Delta f = 1/T$

up to maximum of sampling frequency

$f_{max} = N\Delta f = 1/\Delta t$

- *angular frequency:* $\omega_k = \frac{2\pi k}{T} = \frac{2\pi k}{N\Delta t} = 2\pi k\Delta f$

$$F_k = \frac{1}{N}\sum_{j=0}^{N-1} f_j\, e^{-2\pi ikj/N}$$

$$f_k = \sum_{j=0}^{N-1} F_j\, e^{2\pi ikj/N}$$

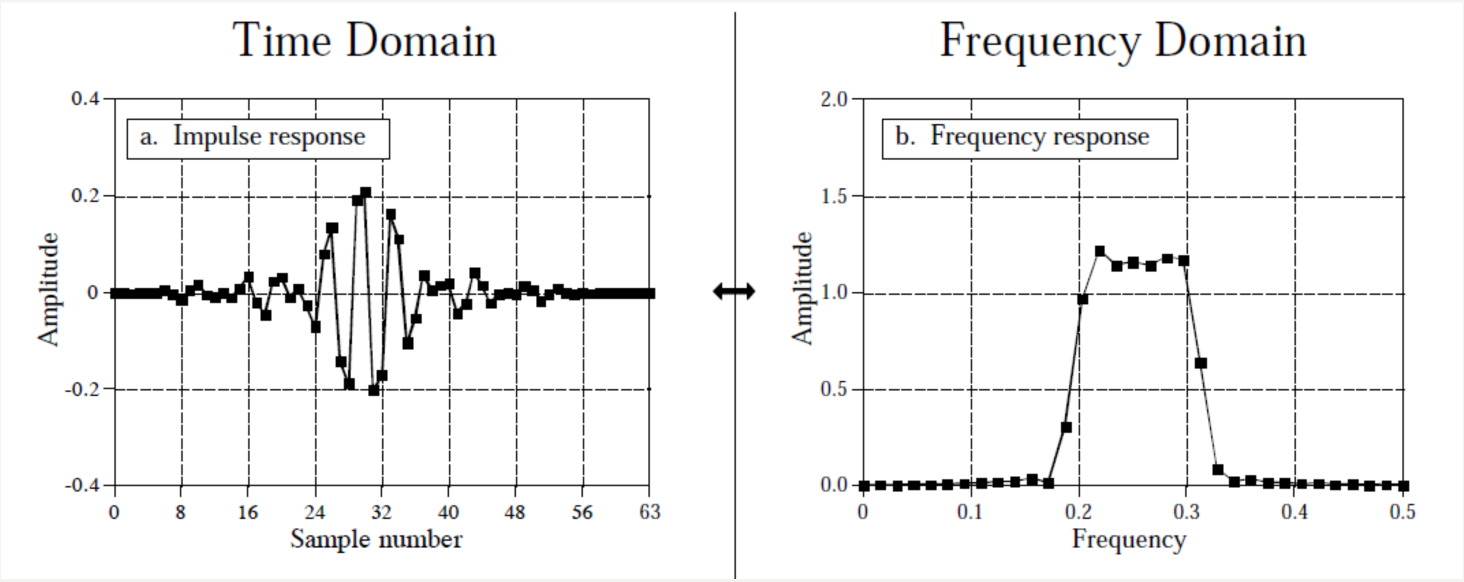
$$k = 0, 1, ..., N-1$$

$\Rightarrow$ *increase sampling density by zero padding*

$\Rightarrow$ *increase frequency resolution by using longer signal*

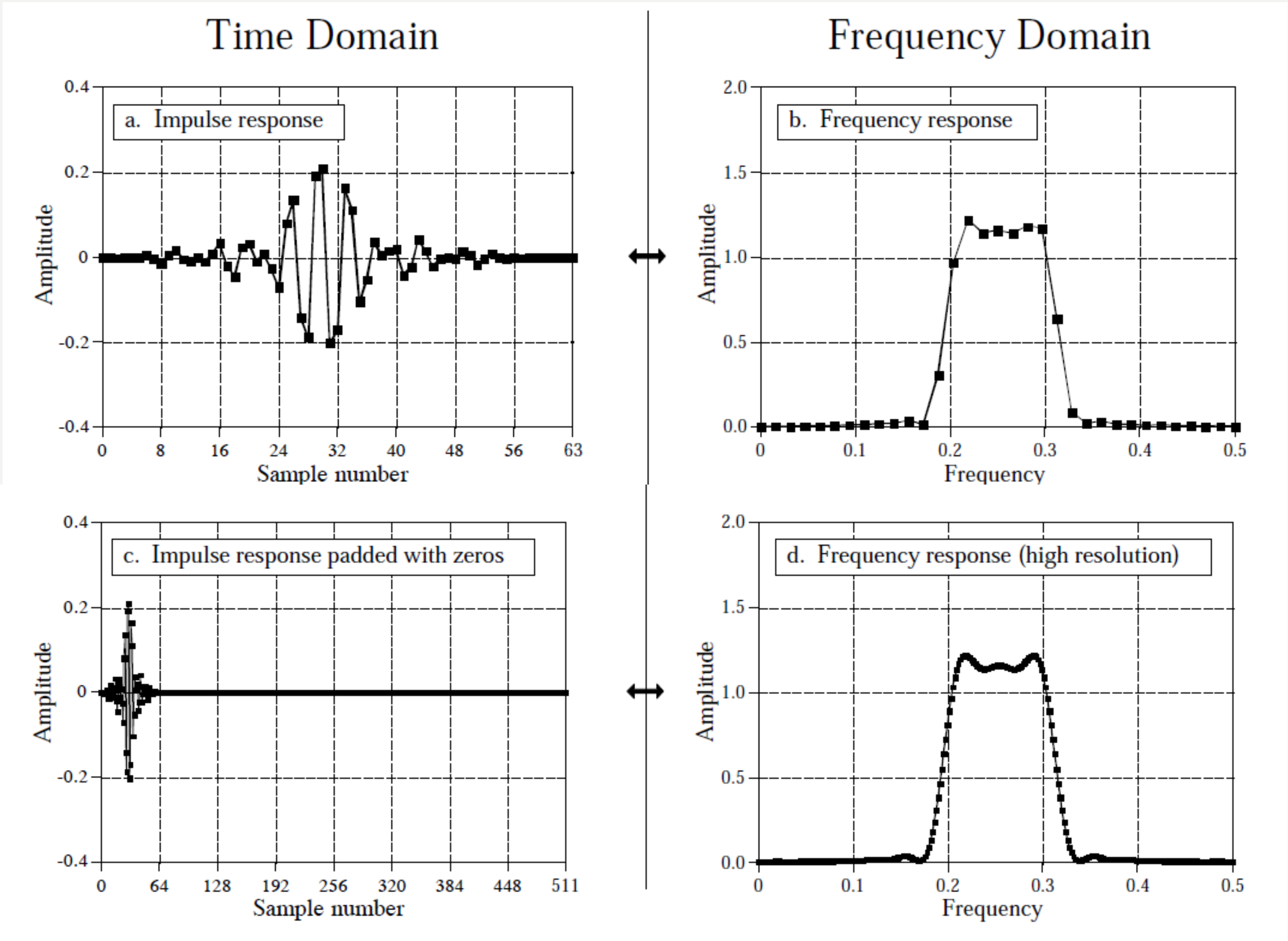
Zero padding

Increase frequency resolution by **zero-padding**



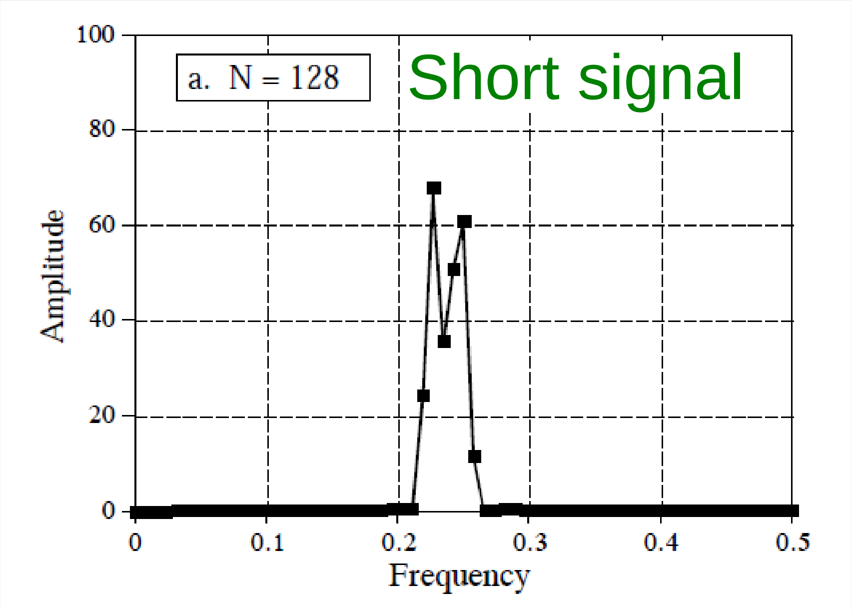
Zero padding

Increase frequency resolution by **zero-padding**



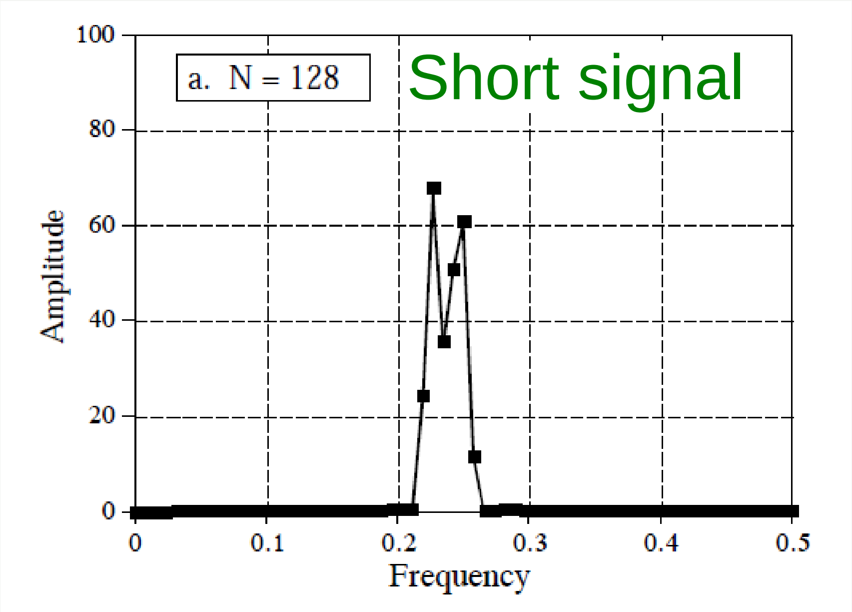
Signal length

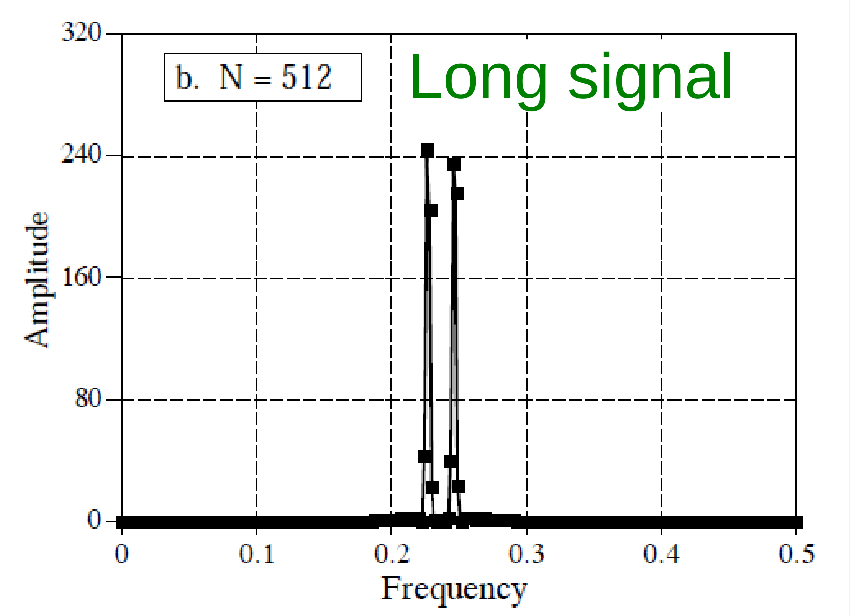
Signal with two sinusoids with very similar frequency:



Signal length

Signal with two sinusoids with very similar frequency:

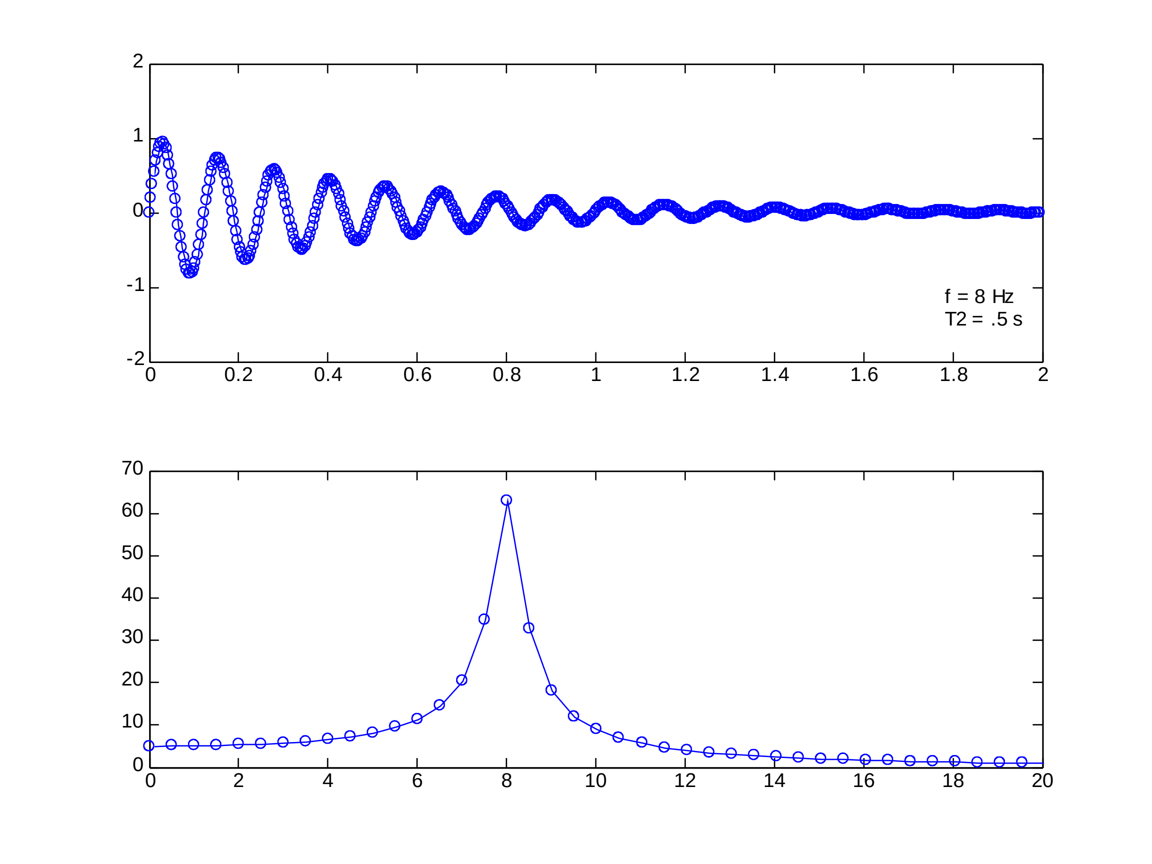




The longer the DFT, the better the ability to separate closely spaced features
→ **improve frequency resolution**

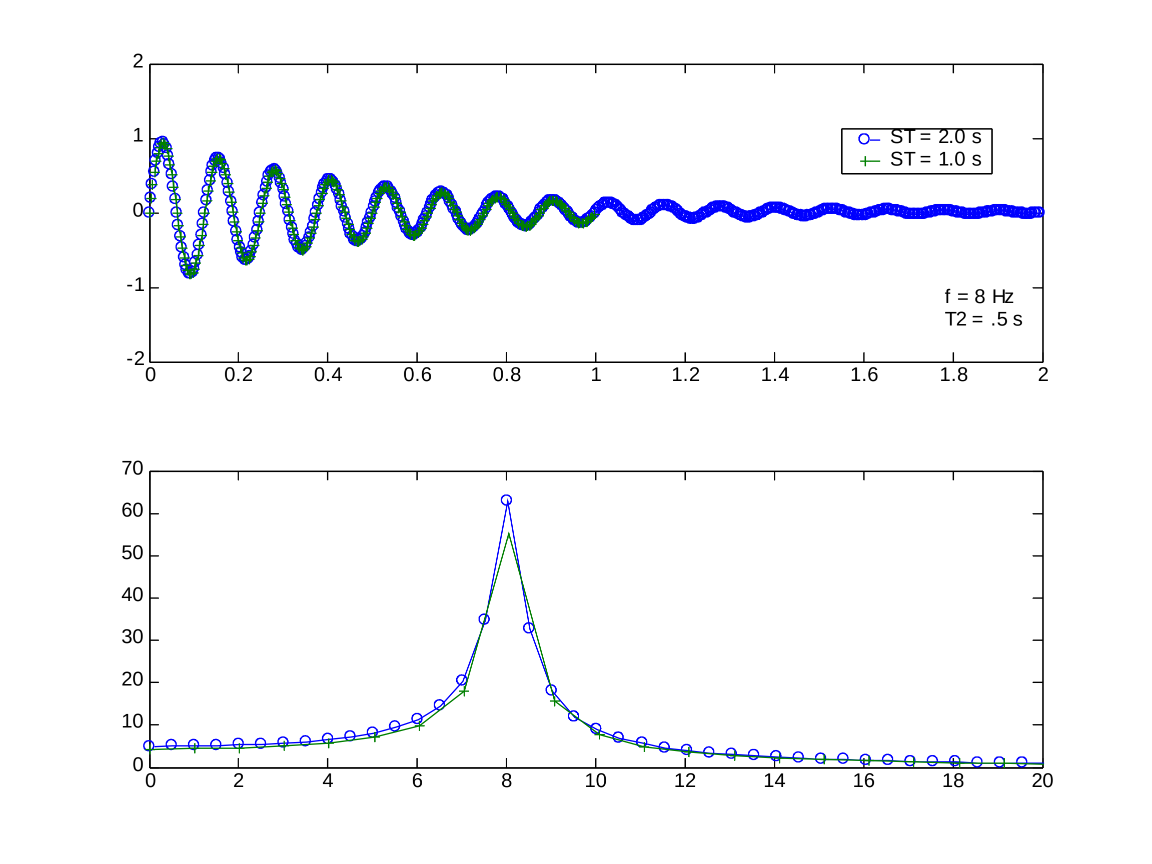
Signal length

Effect of changing sampling duration



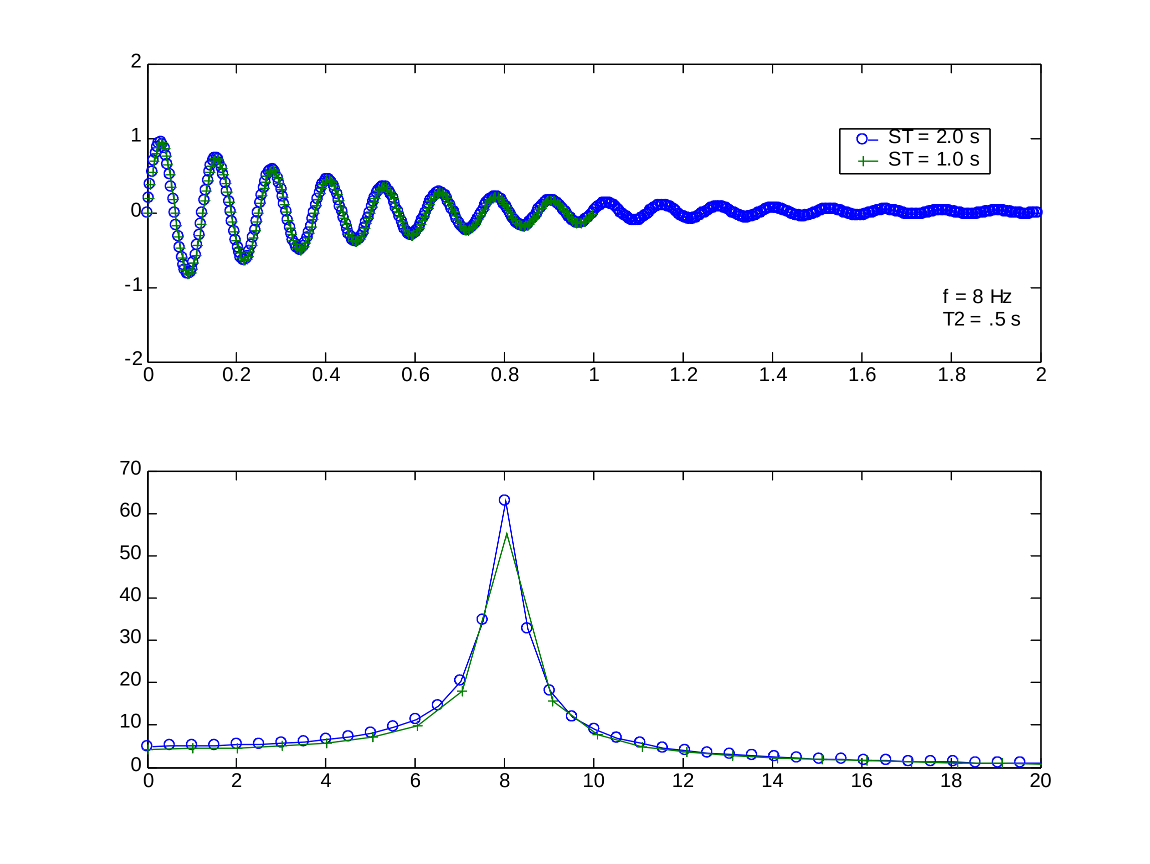
Signal length

Effect of changing sampling duration



Signal length

Effect of changing sampling duration



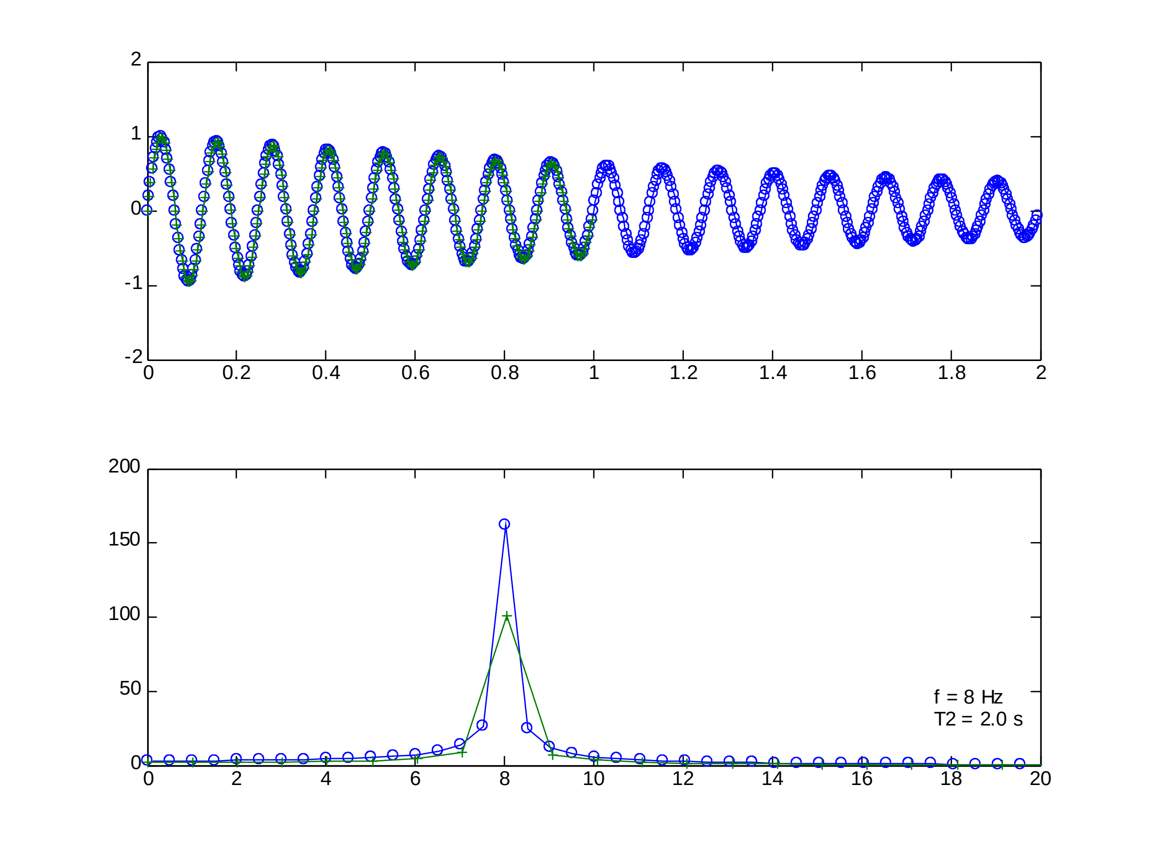
Reducing the sampling duration:

+ Lowers the frequency resolution

+ Does not affect the range of frequencies you can measure

Signal length

Effect of changing sampling duration



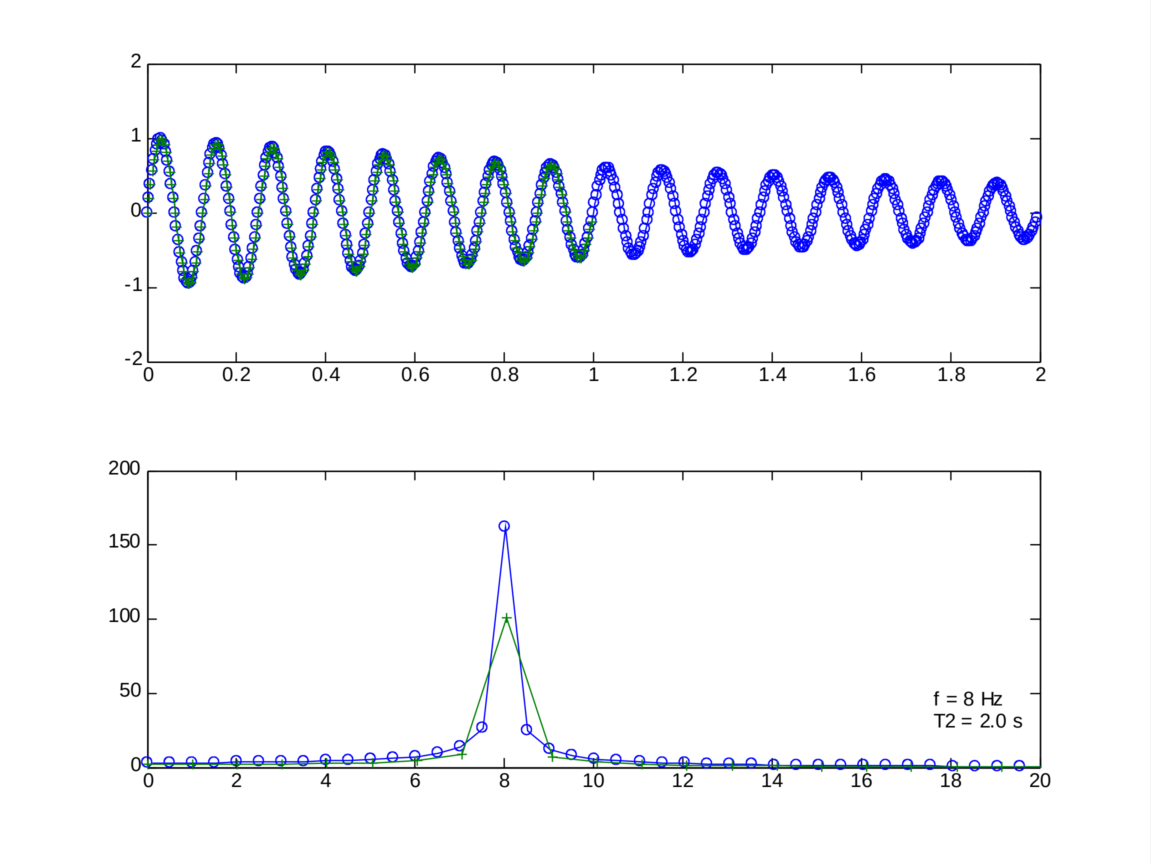
Reducing the sampling duration:

+ Lowers the frequency resolution

+ Does not affect the range of frequencies you can measure

Signal length

Effect of changing sampling duration



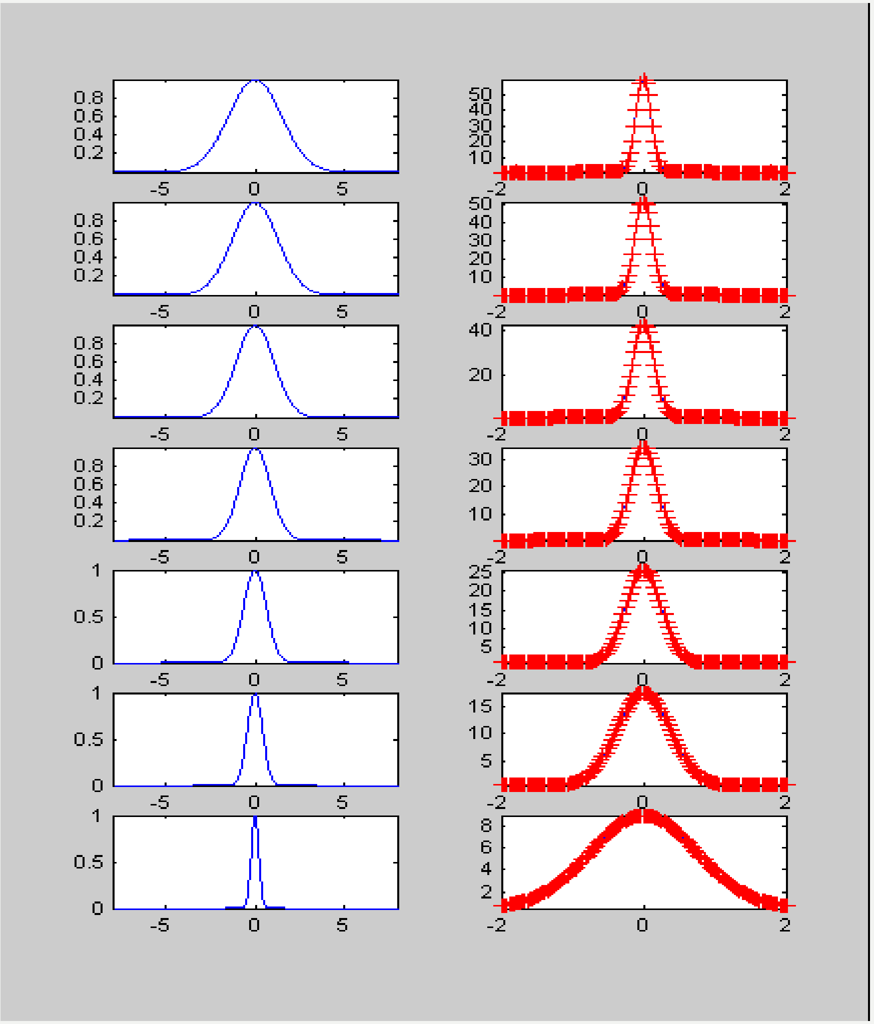
Reducing the sampling duration:

+ Lowers the frequency resolution

+ Does not affect the range of frequencies you can measure

Spectral leakage

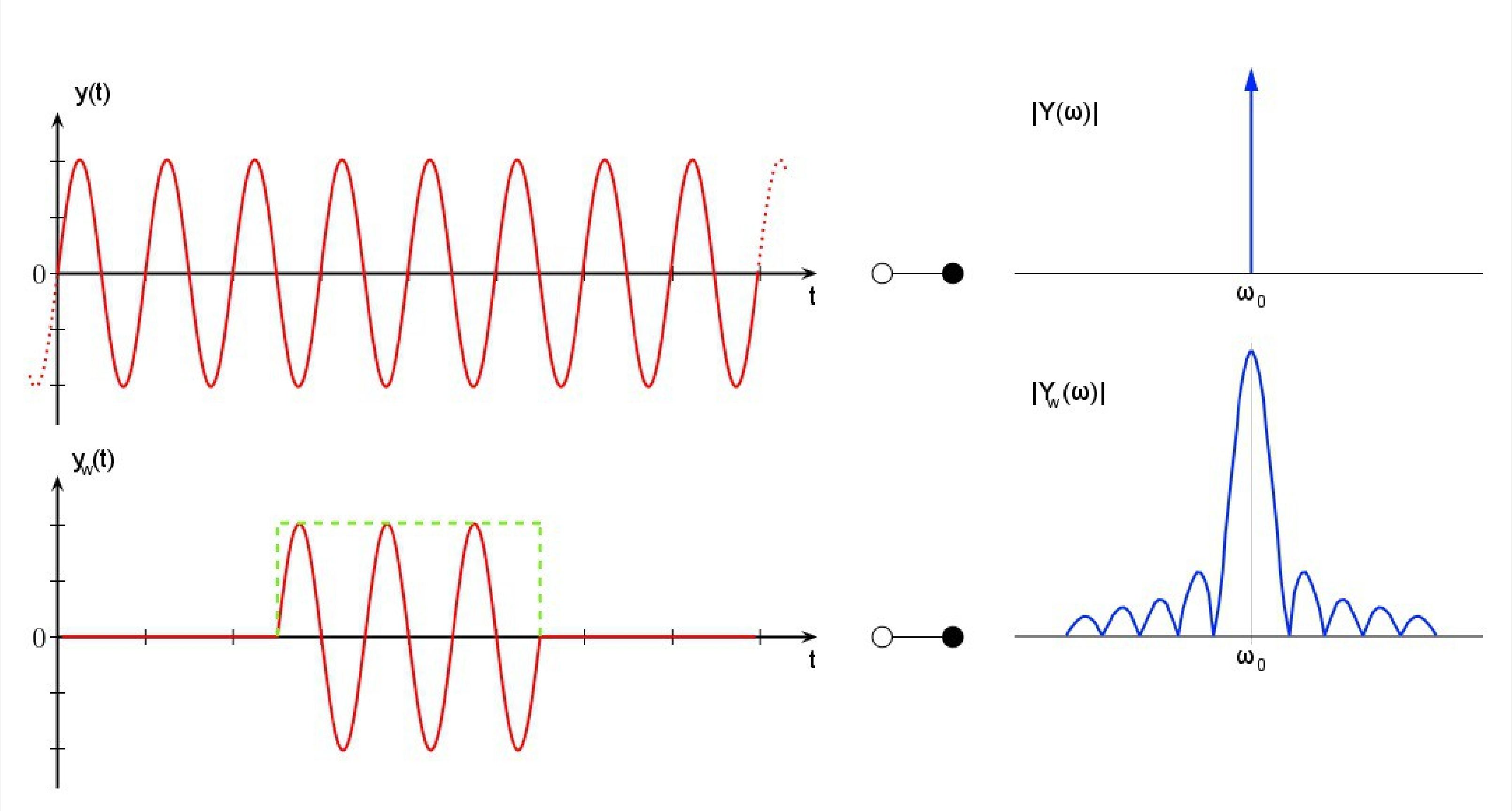
Narrowing physical signal



Widening frequency band

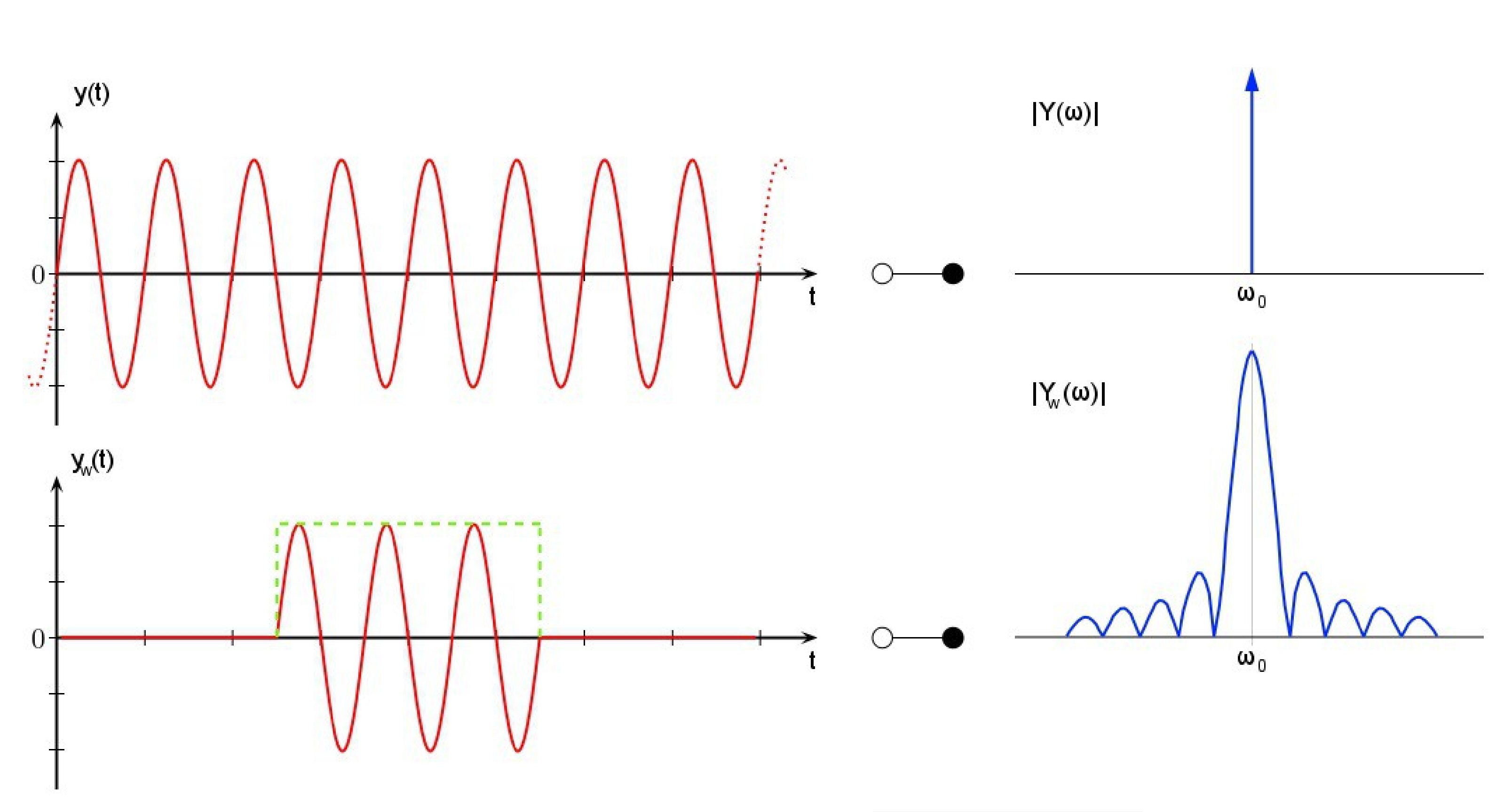
Spectral leakage

The finite length of our **data** pose the problem of spectral leakage → Gibbs phenomenon!



Spectral leakage

The finite length of our **data** pose the problem of spectral leakage → Gibbs phenomenon!



The finite length limits the frequency spacing $\Delta f = 1/T$

The sampling interval Δt *limits the maximum meaningful frequency f*

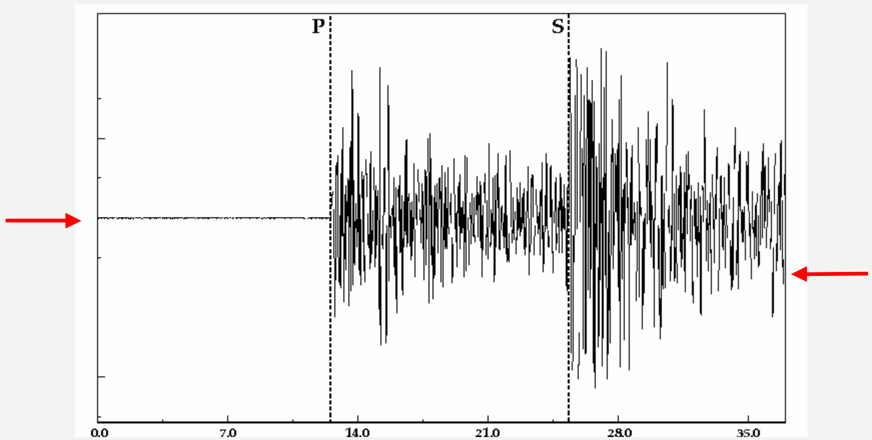
Spectral leakage, windowing & tapering

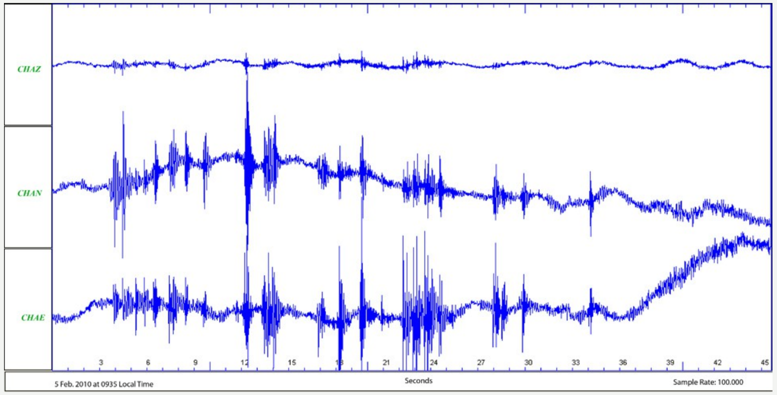
Care must be taken when extracting time windows while estimating spectra!

The **narrower** the window, the **wider** the Fourier transform, the more leakage.

The Fourier transform does not like ***discontinuities****!*

And the FFT assumes periodicity, both ends of the time series must have the same value.





Spectral leakage, windowing & tapering

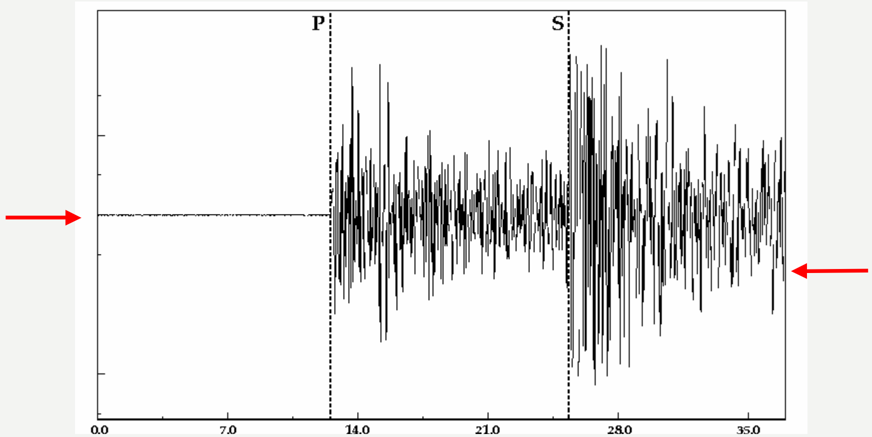
Care must be taken when extracting time windows while estimating spectra!

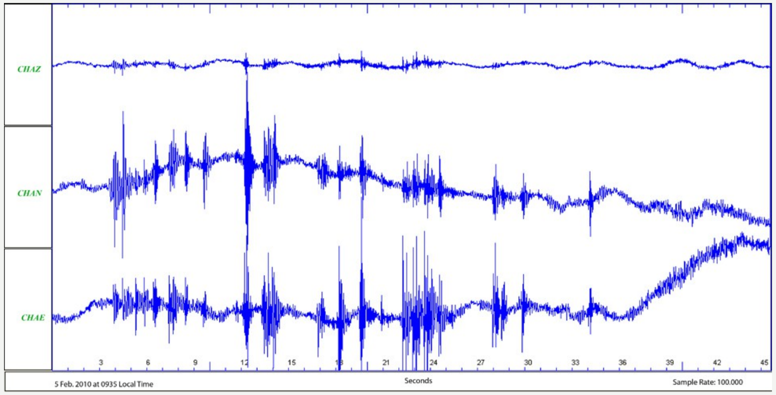
The **narrower** the window, the **wider** the Fourier transform, the more leakage.

*The Fourier transform does not like **discontinuities**!*

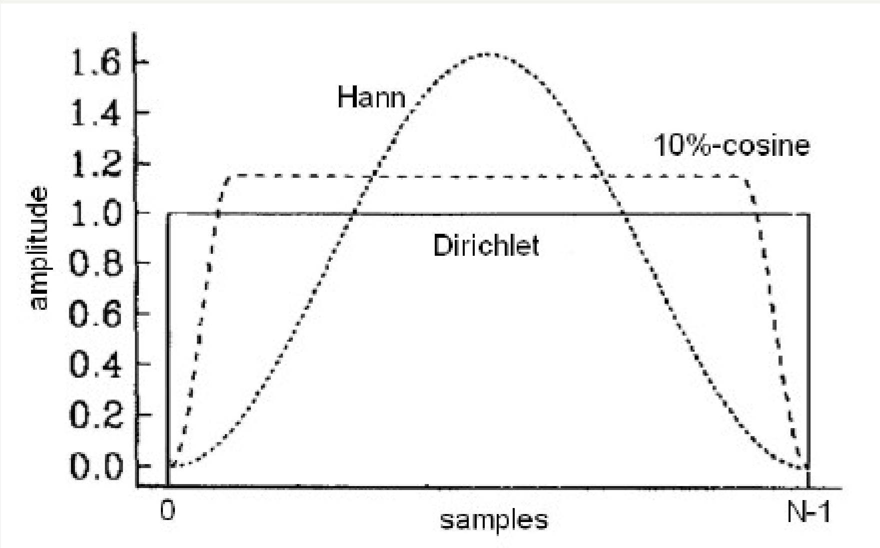
And the FFT assumes periodicity, both ends of the time series must have the same value.

⟹ achieved by **tapering**



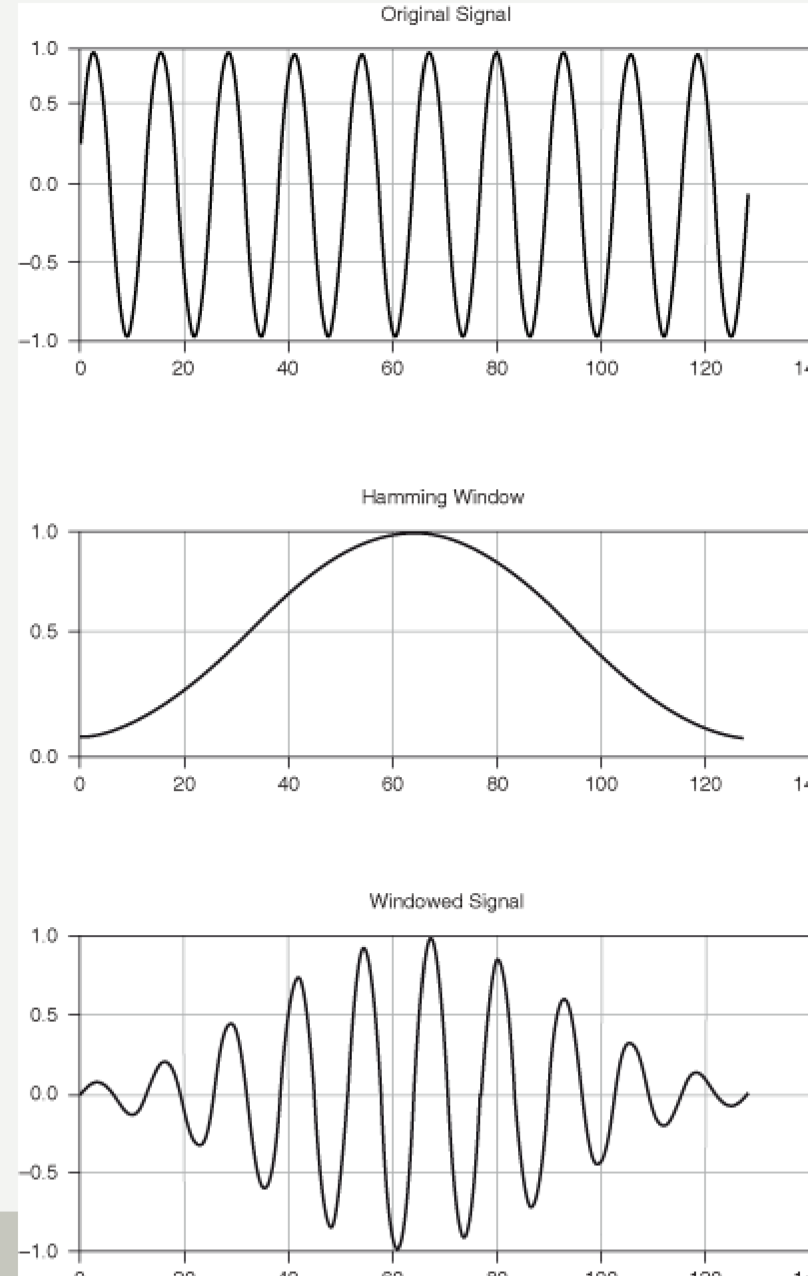


Taper



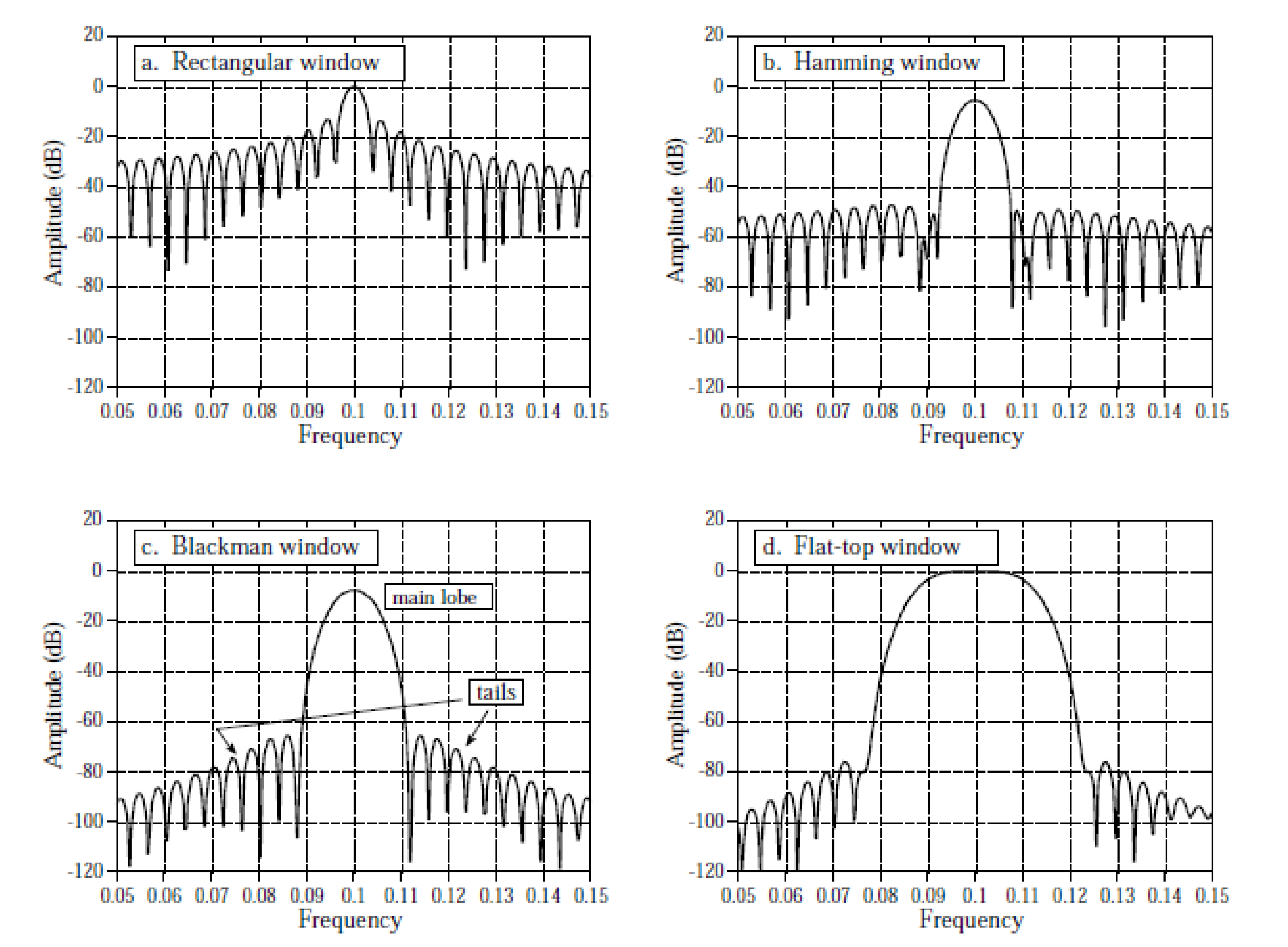
cosine taper with ratio a

$$c(t) = \begin{cases} \frac{1}{2}\left(1 - \cos\frac{\pi}{a}t\right) & \text{for } 0 \le t \le a \\ 1 & \text{for } a \le t \le (1-a) \\ \frac{1}{2}\left(1 - \cos\frac{\pi}{a}(1-t)\right) & \text{for } (1-a) \le t \le 1 \end{cases}$$



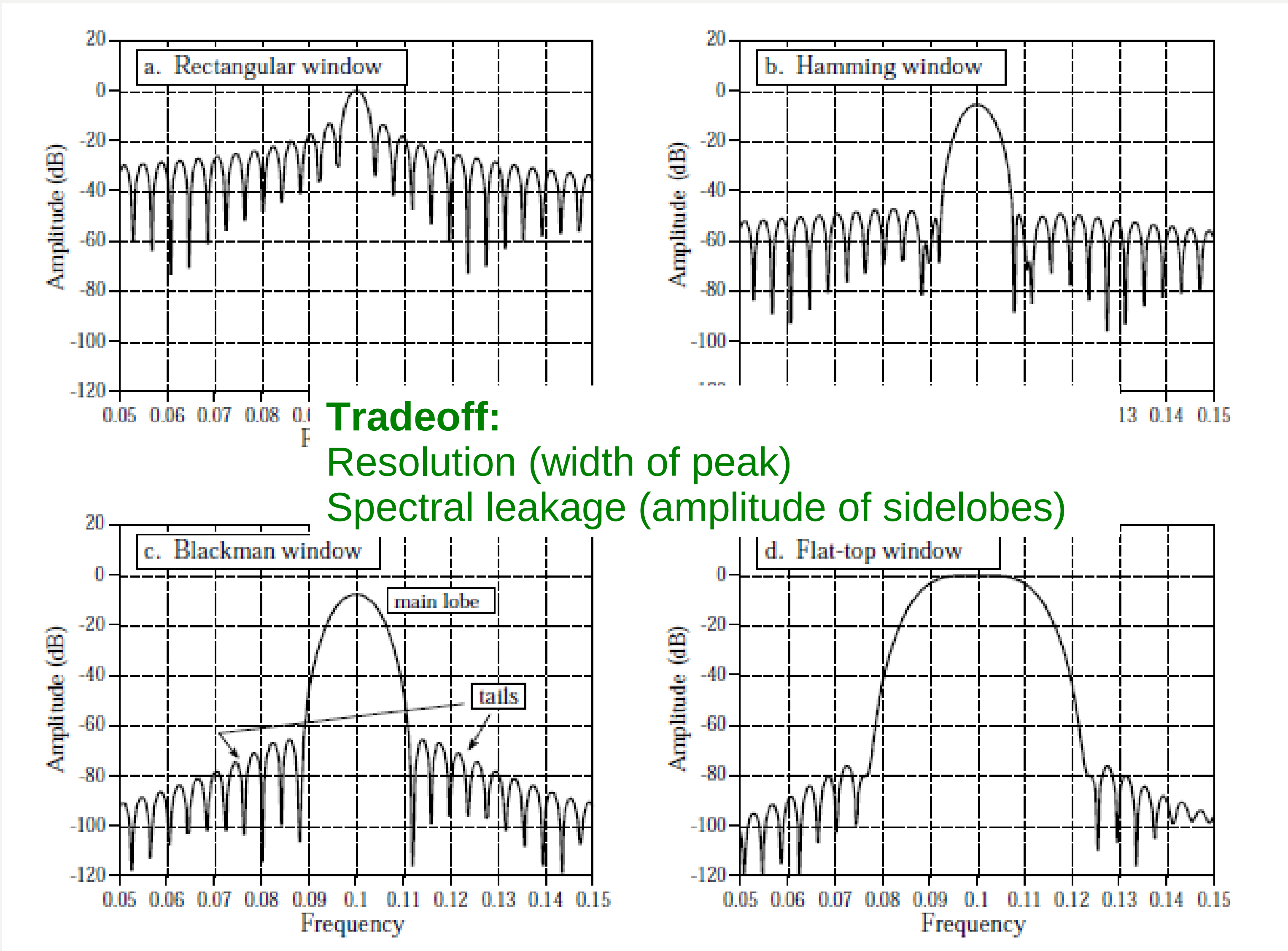
Taper

Effect of different windowing functions:



Taper

Effect of different windowing functions:



Tradeoff:
Resolution (width of peak)
Spectral leakage (amplitude of sidelobes)

Good practice

Preprocessing

1. Filter the analog record to avoid aliasing.

2. Digitise such that the Nyquist frequency lies above the highest frequency in the original data.

3. Window to appropriate length.

4. Detrend (i.e. remove the best-fitting line)

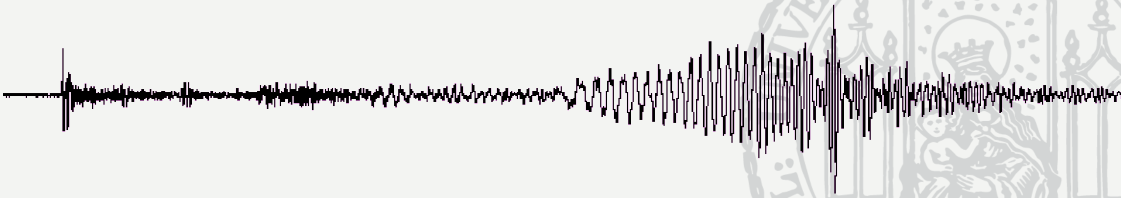
5. Taper to smooth the ends of the record to avoid Gibbs phenomenon.

6. Pad with zeros to smooth the spectrum and/or to lengthen the record to avoid spectral leakage.

Celine Hadziioannou

Geophysical Data Analysis

L05 – Spectral analysis



Applications of the

Fourier Transform

Moving from the continuous to the discrete world.

Fourier transform

contributions to frequency ω

Add all

for each time

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

$$f(t)=\int_{-\infty}^{\infty} F(\omega)e^{-i\omega t}d\omega$$

Add all

for each frequency

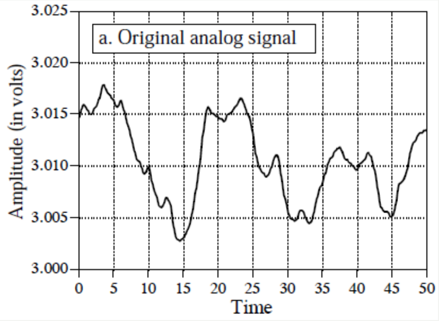
contributions to time t

Forward transform
Time domain → frequency domain
(analysis)

Inverse transform
Frequency domain → time domain
(synthesis)

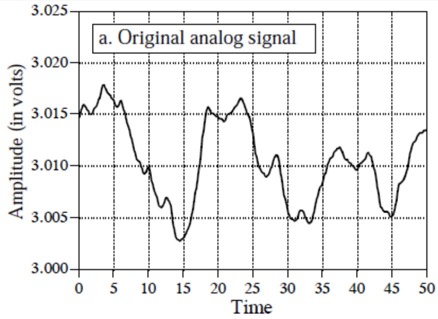
The real, digital world

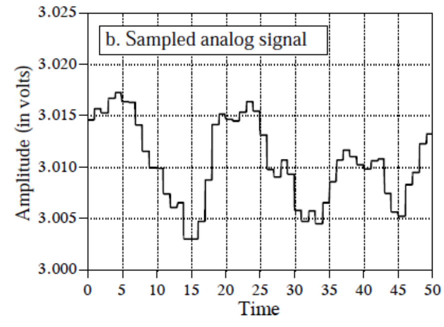
… is NEITHER periodic NOR continuous!



The real, digital world

… is NEITHER periodic NOR continuous!





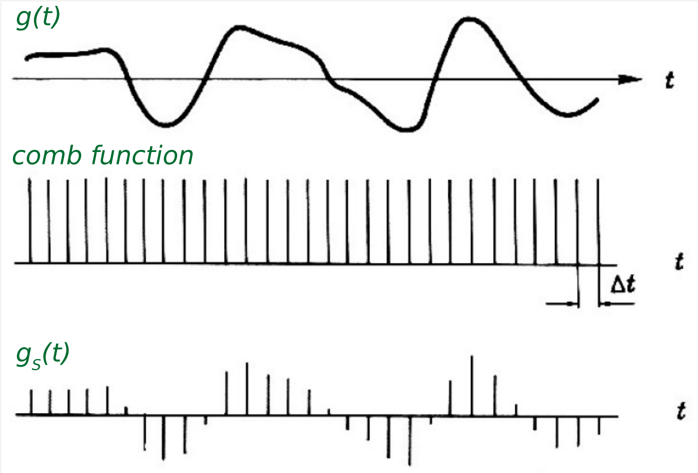
Discretize

The real, digital world

… is NEITHER periodic NOR continuous!

$$g_S(t) = g(t) \sum_{j=-\infty}^{\infty} \delta(t - j\,dt)$$

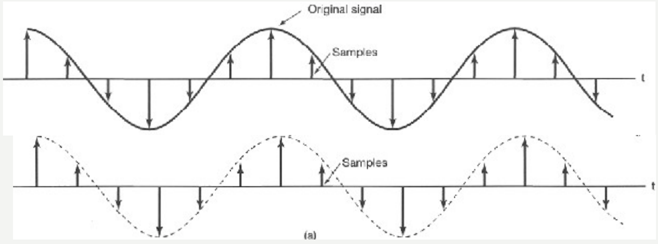
comb function



$g_s(t)$ is the digitized version of $g(t)$

time signal is multiplied with comb function.

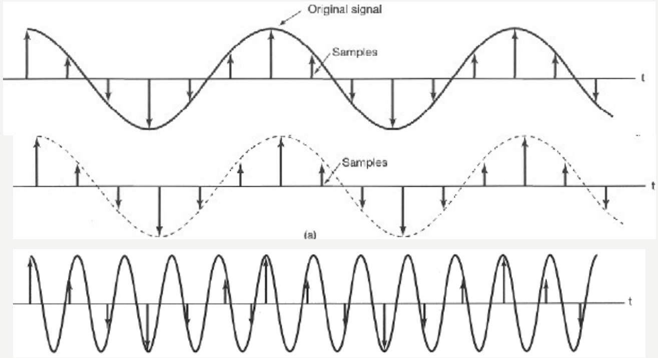
Sampling frequency



$$x[n] = x_{cont}(nT)$$

T = sampling period
ωs = 2π/T sampling frequency in [rad]
Fs = 1/T sampling frequency or 'sample rate', in [Hz]

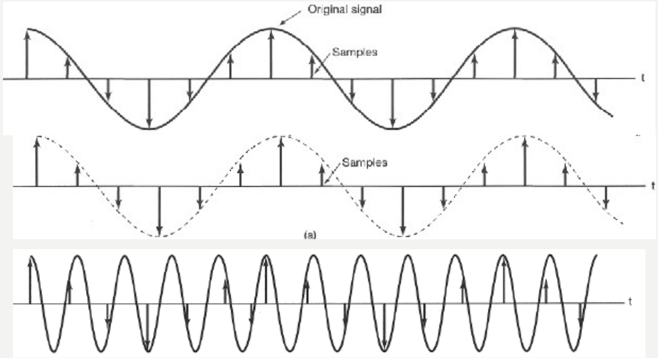
Sampling frequency



$$x[n] = x_{cont}(nT)$$

T = sampling period
ωs = 2π/T sampling frequency in [rad]
Fs = 1/T sampling frequency or 'sample rate', in [Hz]

Sampling frequency



Sampled version of signal could also be fitted with higher frequency signal

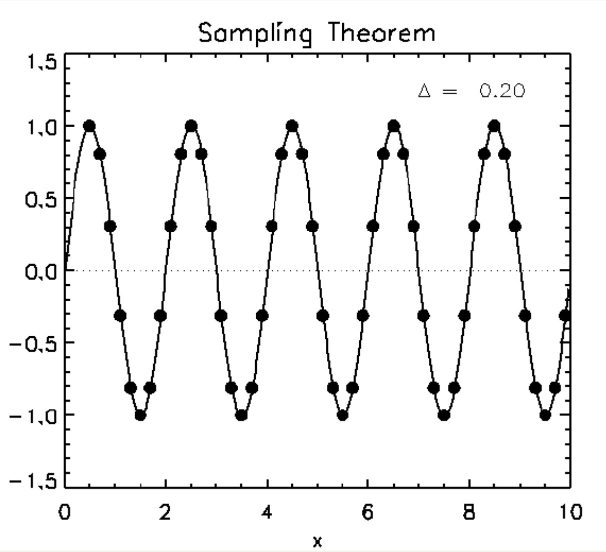
$$x[n] = x_{cont}(nT)$$

T = sampling period

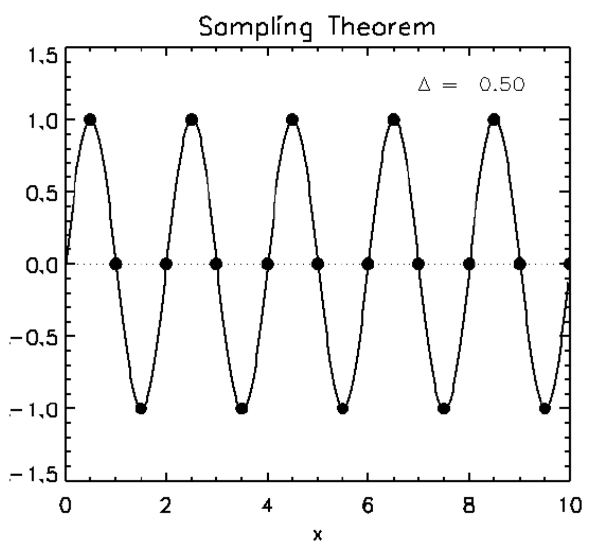
ωs = 2π/T sampling frequency in [rad]

Fs = 1/T sampling frequency or 'sample rate', in [Hz]

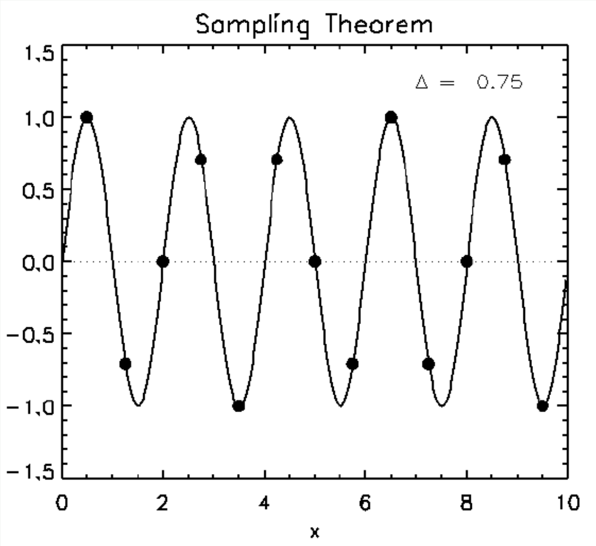
Sampling frequency



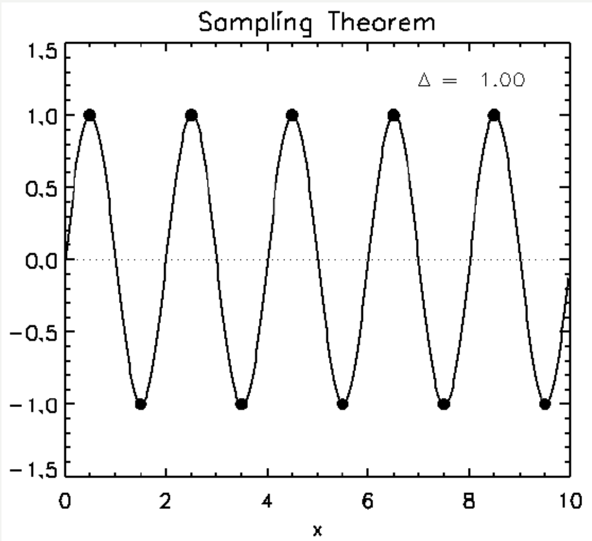
Sampling frequency



Sampling frequency

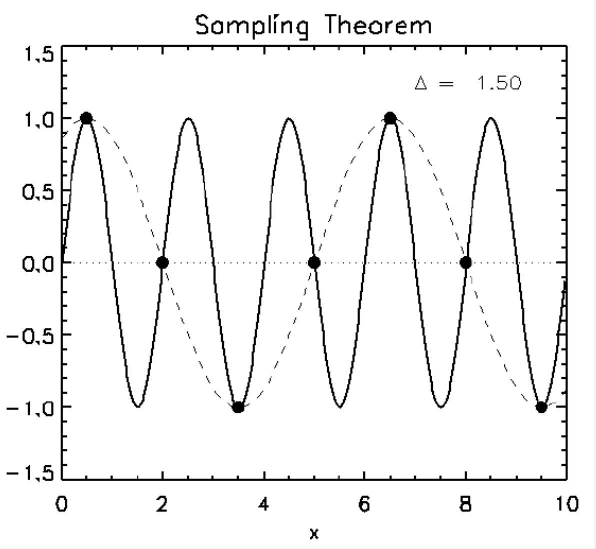


Sampling frequency



Critical sampling frequency!

Sampling frequency



Undersampled..

Sampling theorem

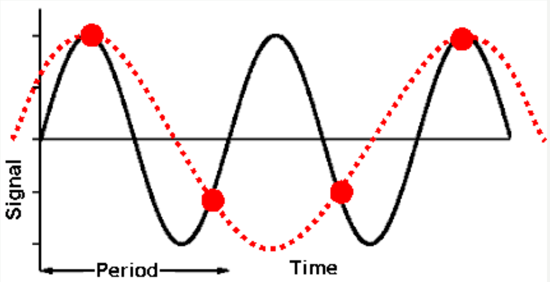
*At least **2 samples per period** are needed to correctly reproduce the **highest frequency** of a signal*

OR

a continuous signal can be properly sampled, only if it does not contain frequency components **above one-half of the sampling rate**

"Nyquist frequency"

$$f_{Ny} = \frac{1}{2\,dt}$$



*Otherwise: **aliasing** !!!*

"wrong" frequencies appear in your signal

Sampling theorem

At least ***2 samples per period*** *are needed to correctly reproduce the* ***highest frequency*** *of a signal*

OR

a continuous signal can be properly sampled, only if it

does not contain frequency components **above one-half of the sampling rate**

"Nyquist frequency"

$$f_{Ny} = \frac{1}{2\,dt}$$

frequency of view: 50 minutes

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Sampling theorem

At least ***2 samples per period*** *are needed to correctly reproduce the* ***highest frequency*** *of a signal*

OR

a continuous signal can be properly sampled, only if it

does not contain frequency components **above one-half of the sampling rate**

"Nyquist frequency"

$$f_{Ny} = \frac{1}{2\,dt}$$

Audio CD sampling frequency: 44.1 kHz

Why?

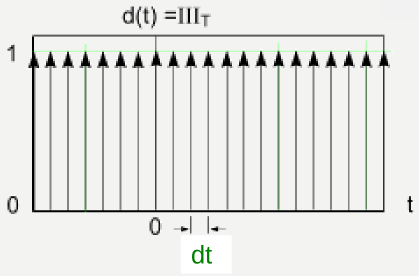
The sampling theorem

… back to the comb function …

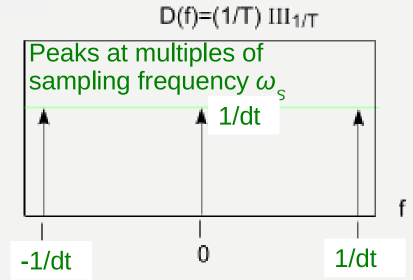
$$g_S(t) = g(t) \sum_{j=-\infty}^{\infty} \delta(t - j\,dt)$$

With the Nyquist frequency defined as

$$f_{Ny} = \frac{1}{2\,dt}$$



*xs[n] = xc(t) * time domain comb*

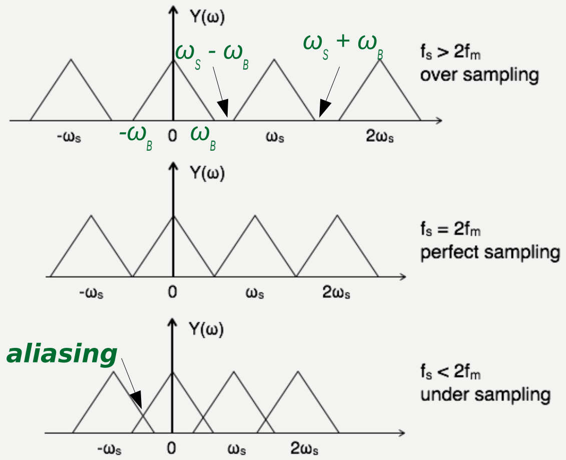


*Xs[ω] = Xc(ω) * frequency domain comb*

The sampling theorem

⇒ For the calculation of the spectrum of the frequency f, there are also contributions of frequencies f ± 2nf_{Ny}

That means dt has to be chosen such that f_{Ny} is the largest frequency contained in the signal.



The sampling theorem

⇒ For the calculation of the spectrum of the frequency f, there are also contributions of frequencies $f \pm 2nf_{Ny}$

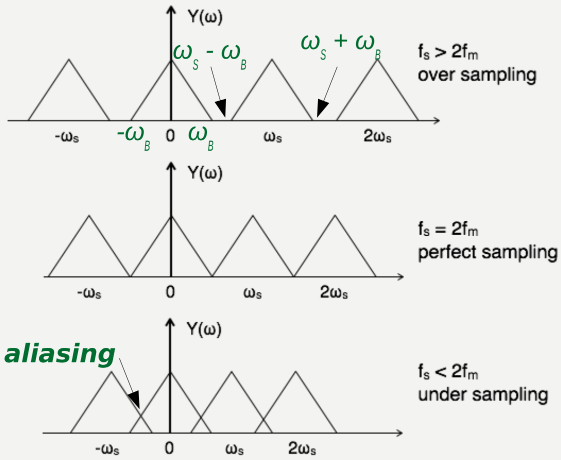
That means dt has to be chosen such that f_{Ny} is the largest frequency contained in the signal.

For a band-limited signal where the spectrum is between $\pm\ \omega_B$:

Don't want the 'copies' to overlap:

$$\omega_S - \omega_B > \omega_B$$

$$\omega_S > 2\omega_B$$



The sampling theorem

For the calculation of the spectrum of the frequency f, there are also contributions of frequencies $f \pm 2nf_{Ny}$

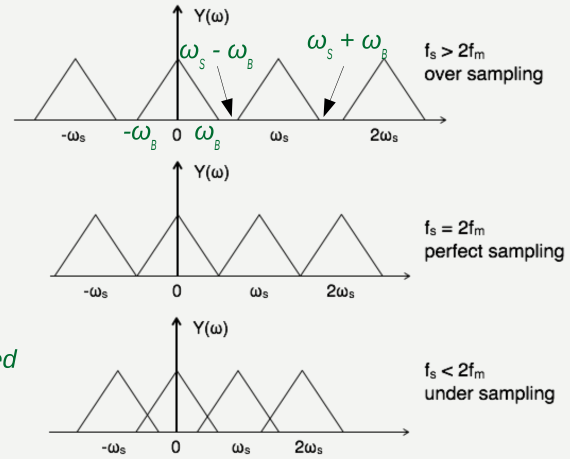
That means dt has to be chosen such that f_{Ny} is the largest frequency contained in the signal.

For a band-limited signal where the spectrum is between +- ω_B:

Don't want the 'copies' to overlap:

$\omega_S - \omega_B > \omega_B$

$\omega_S > 2\omega_B$

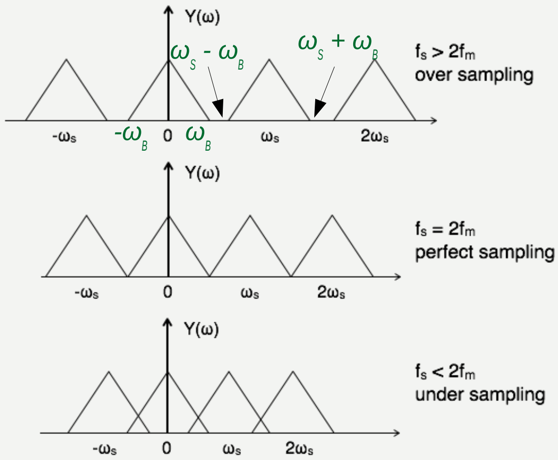


Sampling theorem: *A band-limited signal with maximum frequency ω_B can be reconstructed perfectly from evenly spaced samples if the sampling frequency ω_S satisfies $\omega_S > 2\omega_B$*

The sampling theorem

"A continuous signal can be properly sampled, only if it
does not contain frequency components above one-half of the sampling rate."

What if the signal contains higher frequencies as well?

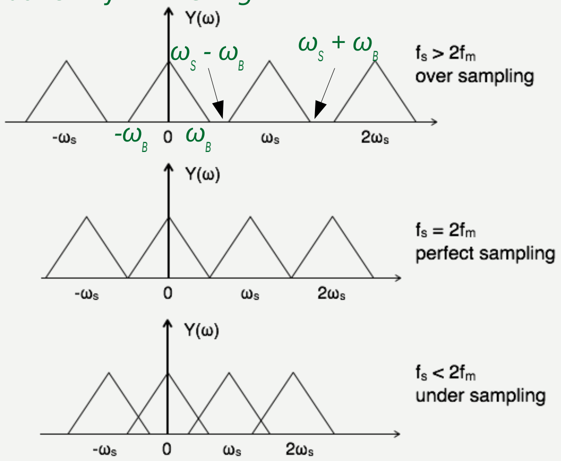


The sampling theorem

"A continuous signal can be properly sampled, only if it
does not contain frequency components above one-half of the sampling rate."

What if the signal contains higher frequencies as well?

Solution: remove all frequencies above fNy → *filtering*



LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

The sampling theorem

"A continuous signal can be properly sampled, only if it
does not contain frequency components above one-half of the sampling rate."
What if the signal contains higher frequencies as well?

Solution: remove all frequencies above fNy → *filtering*

Resampling (Decimating):

- Often is it useful to down-sample a time series

(e.g. from 100Hz to 1 Hz when studying surface waves)

- Then, the time series needs to be pre-processed.

  All frequencies above twice the new sampling interval have to be filtered out before !!!

We now have discrete signals..

.. back to the Fourier Transform

The Fourier transform

Whatever we do on the computer with data will be based on the discrete Fourier transform.

continuous

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

$$f(t) = \int_{-\infty}^{\infty} f(\omega) e^{i\omega t} d\omega$$

discrete

$$F_k = \frac{1}{N} \sum_{j=0}^{N-1} f_j \, e^{-2\pi i k j / N}$$

$$f_k = \sum_{j=0}^{N-1} F_j \, e^{2\pi i k j / N}$$

$$k = 0, 1, ..., N-1$$

Discrete Fourier Transform

- *j measures time in units of sampling interval*

  $j = k\Delta t$ *up to maximum time* $T = N\Delta t$

- *k measures frequency in intervals of the sampling frequency*

  $\Delta f = 1/T$

  up to maximum of sampling frequency

  $f_{max} = N\Delta f = 1/\Delta t$

- *angular frequency:* $\omega_k = \frac{2\pi k}{T} = \frac{2\pi k}{N\Delta t} = 2\pi k \Delta f$

$$F_k = \frac{1}{N}\sum_{j=0}^{N-1} f_j \, e^{-2\pi ikj/N}$$

$$f_k = \sum_{j=0}^{N-1} F_j \, e^{2\pi ikj/N}$$

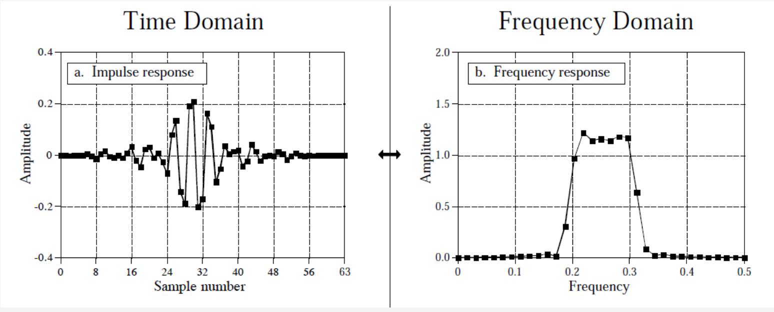
$$k = 0, 1, ..., N-1$$

⇨ *increase sampling density by zero padding*

⇨ *increase frequency resolution by using longer signal*

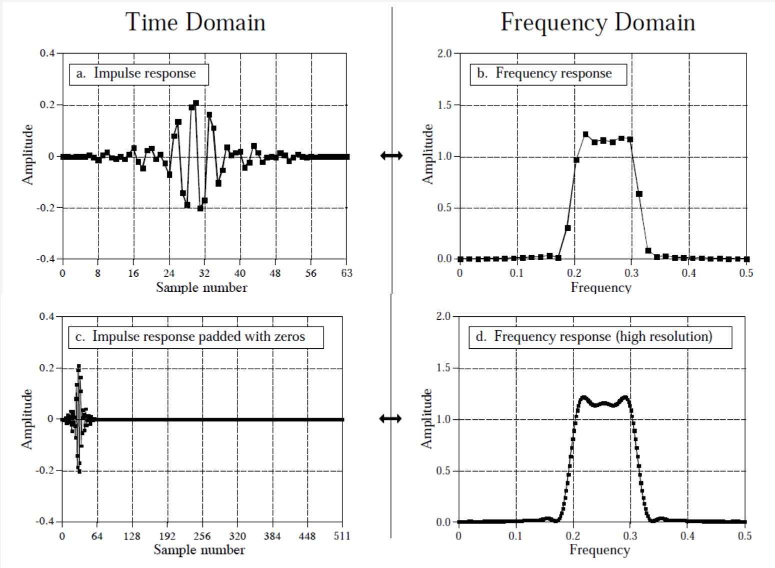
Zero padding

Increase frequency resolution by **zero-padding**



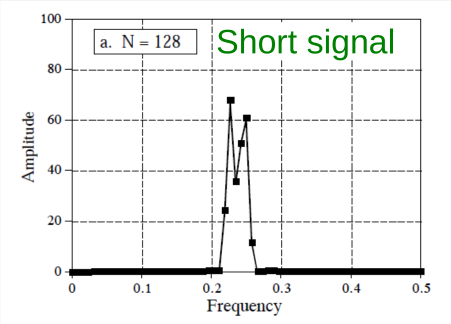
Zero padding

Increase frequency resolution by **zero-padding**



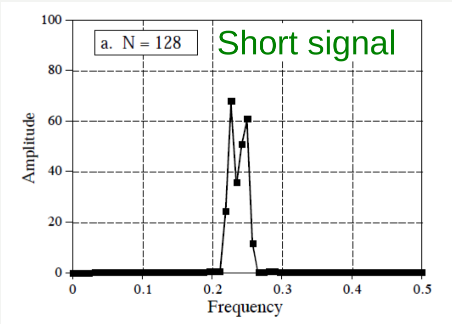
Signal length

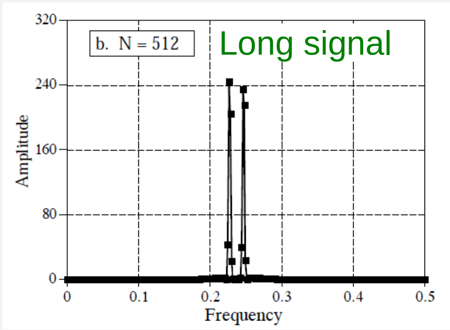
Signal with two sinusoids with very similar frequency:



Signal length

Signal with two sinusoids with very similar frequency:

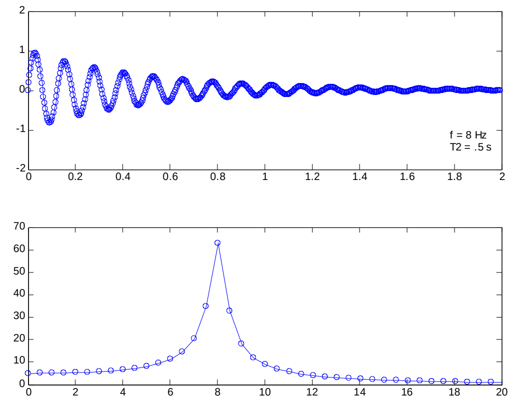




The longer the DFT, the better the ability to separate closely spaced features
→ **improve frequency resolution**

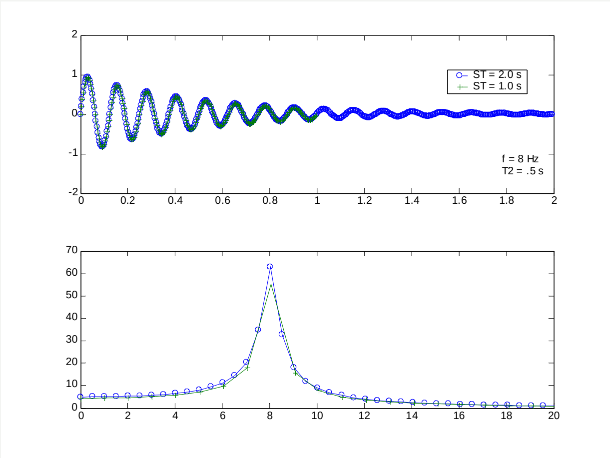
Signal length

Effect of changing sampling duration



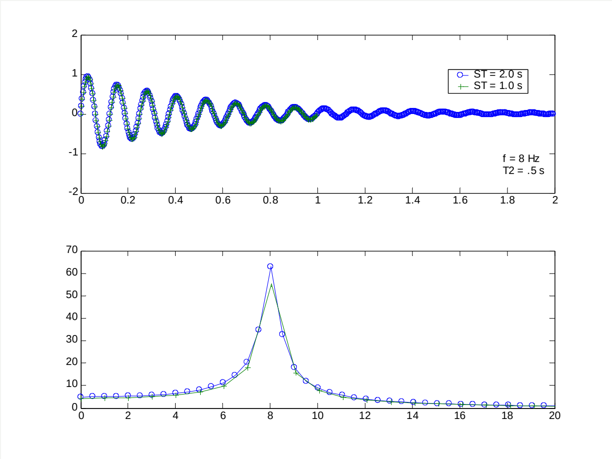
Signal length

Effect of changing sampling duration



Signal length

Effect of changing sampling duration



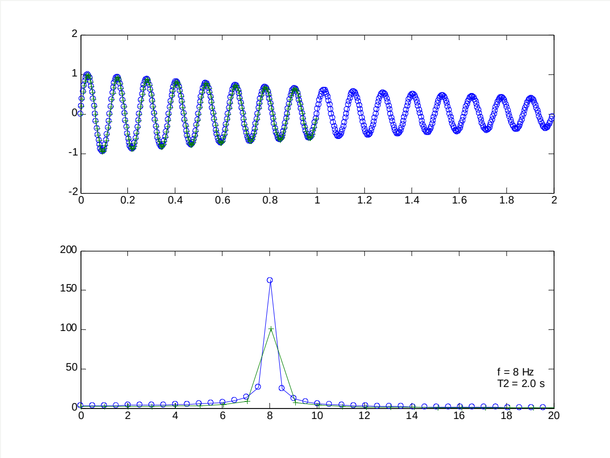
Reducing the sampling duration:

+ Lowers the frequency resolution

+ Does not affect the range of frequencies you can measure

Signal length

Effect of changing sampling duration



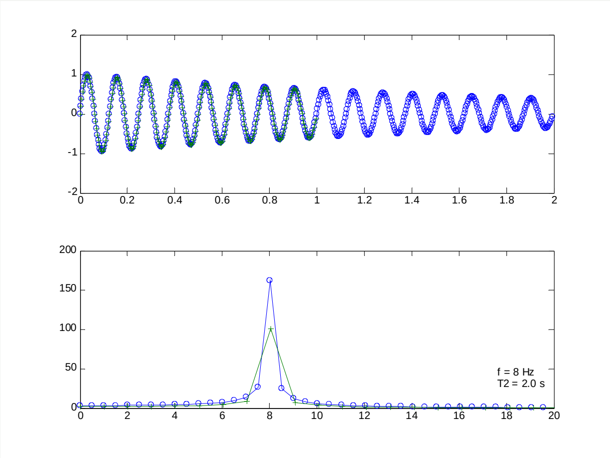
Reducing the sampling duration:

+ Lowers the frequency resolution

+ Does not affect the range of frequencies you can measure

Signal length

Effect of changing sampling duration



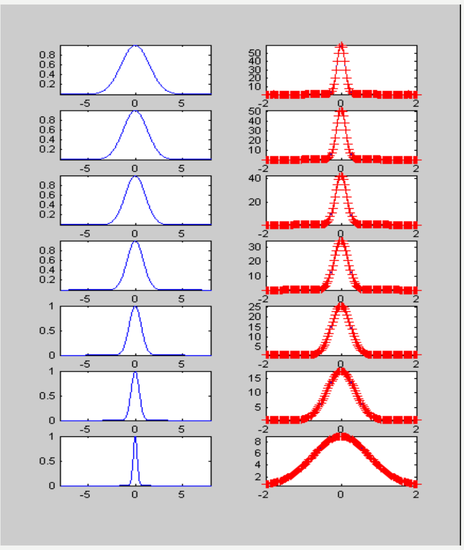
Reducing the sampling duration:

+ Lowers the frequency resolution

+ Does not affect the range of frequencies you can measure

Spectral leakage

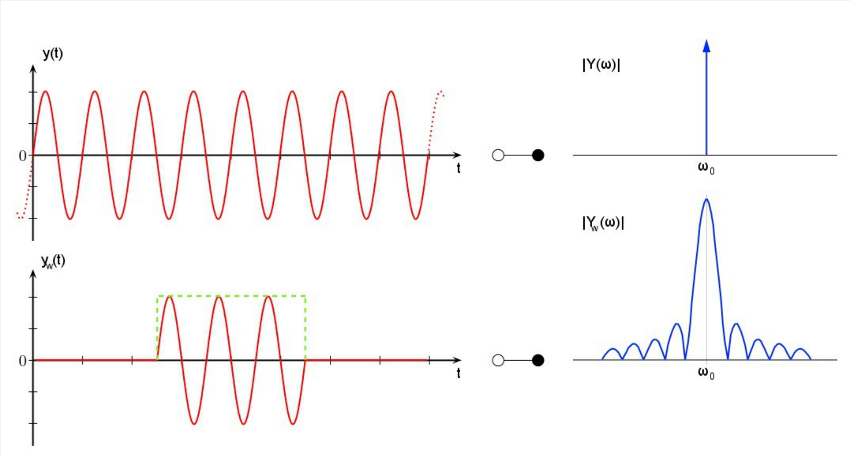
Narrowing physical signal



Widening frequency band

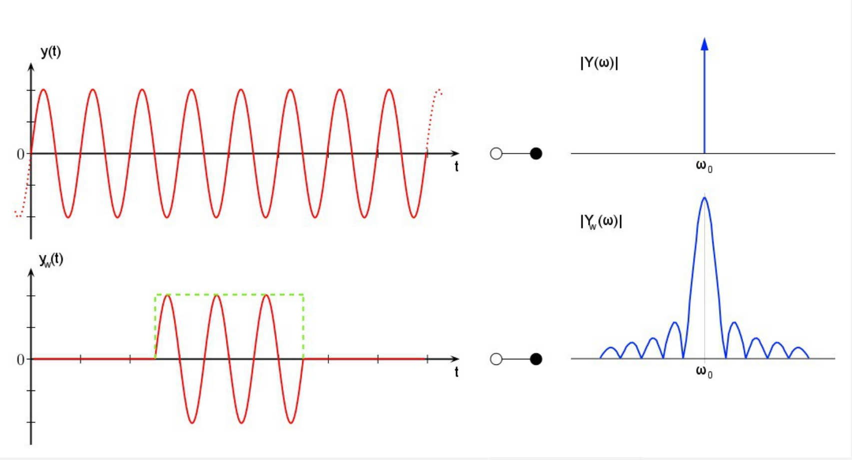
Spectral leakage

The finite length of our **data** pose the problem of spectral leakage → Gibbs phenomenon!



Spectral leakage

The finite length of our **data** pose the problem of spectral leakage → Gibbs phenomenon!



The finite length limits the frequency spacing $\Delta f = 1/T$

The sampling interval Δt *limits the maximum meaningful frequency f*

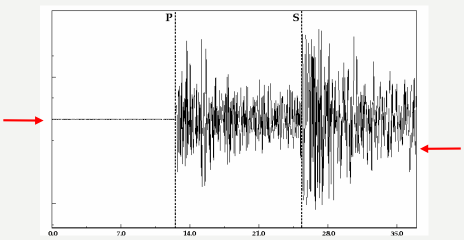
Spectral leakage, windowing & tapering

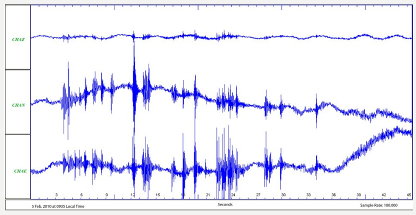
Care must be taken when extracting time windows while estimating spectra!

The **narrower** the window, the **wider** the Fourier transform, the more leakage.

*The Fourier transform does not like **discontinuities**!*

And the FFT assumes periodicity, both ends of the time series must have the same value.





Spectral leakage, windowing & tapering

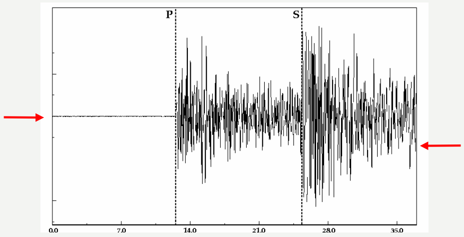
Care must be taken when extracting time windows while estimating spectra!

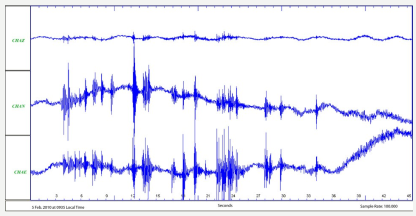
The **narrower** the window, the **wider** the Fourier transform, the more leakage.

*The Fourier transform does not like **discontinuities**!*

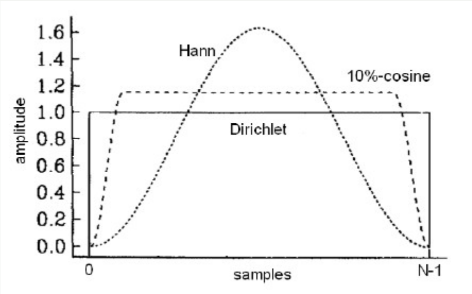
And the FFT assumes periodicity, both ends of the time series must have the same value.

⇨ achieved by **tapering**



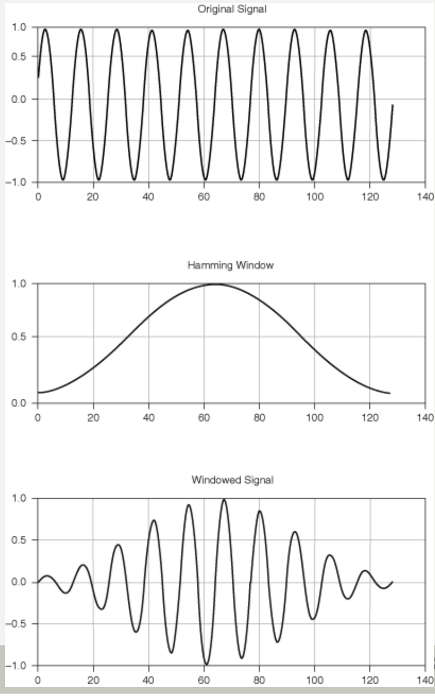


Taper



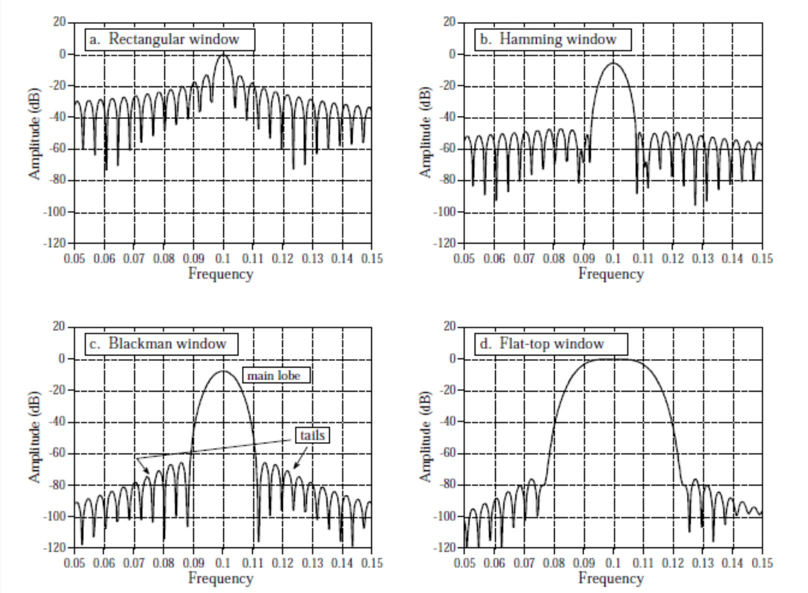
cosine taper with ratio a

$$c(t) = \begin{cases} \frac{1}{2}\left(1 - \cos\frac{\pi}{a}t\right) & \text{for } 0 \le t \le a \\ 1 & \text{for } a \le t \le (1-a) \\ \frac{1}{2}\left(1 - \cos\frac{\pi}{a}(1-t)\right) & \text{for } (1-a) \le t \le 1 \end{cases}$$



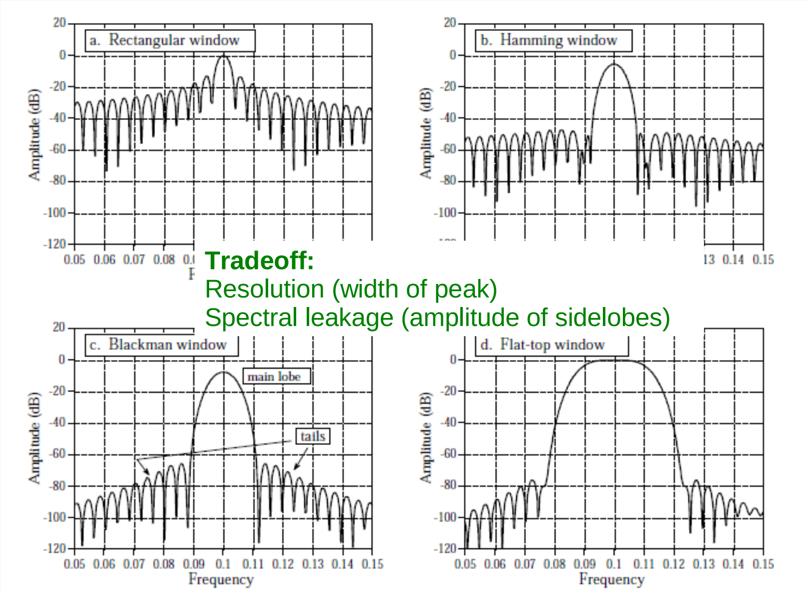
Taper

Effect of different windowing functions:



Taper

Effect of different windowing functions:



Tradeoff:
Resolution (width of peak)
Spectral leakage (amplitude of sidelobes)

Good practice

Preprocessing

1. Filter the analog record to avoid aliasing.

2. Digitise such that the Nyquist frequency lies above the highest frequency in the original data.

3. Window to appropriate length.

4. Detrend (i.e. remove the best-fitting line)

5. Taper to smooth the ends of the record to avoid Gibbs phenomenon.

6. Pad with zeros to smooth the spectrum and/or to lengthen the record to avoid spectral leakage.