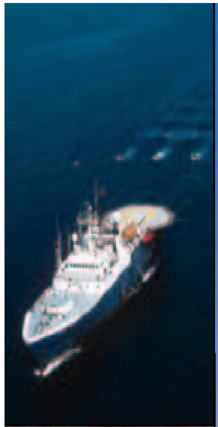
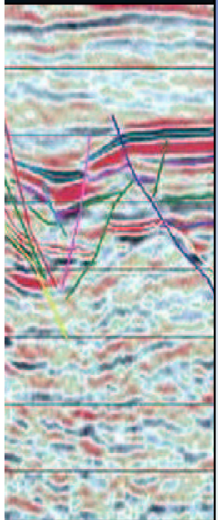
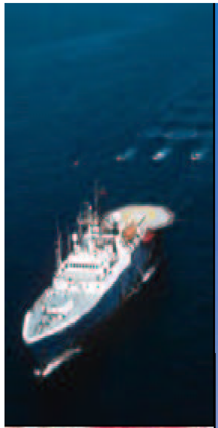
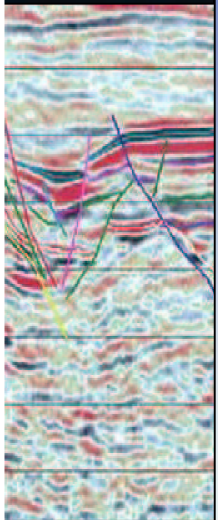



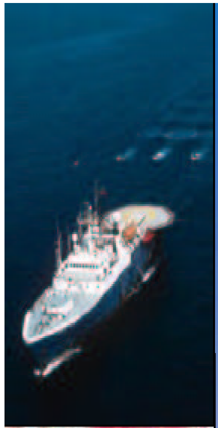

Geophysical Data Processing

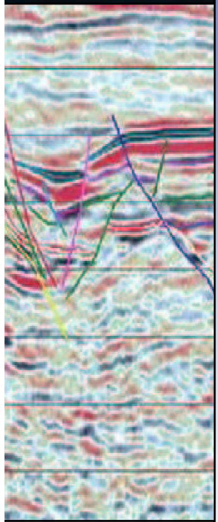
- **Digitization**
  - Dynamic range
- **Spectral analysis**
  - Fourier analysis
  - Sampling
  - Spectra in space and time domain
- **Waveform processing**
  - Convolution
  - Deconvolution
  - Correlation
  - Digital Filtering

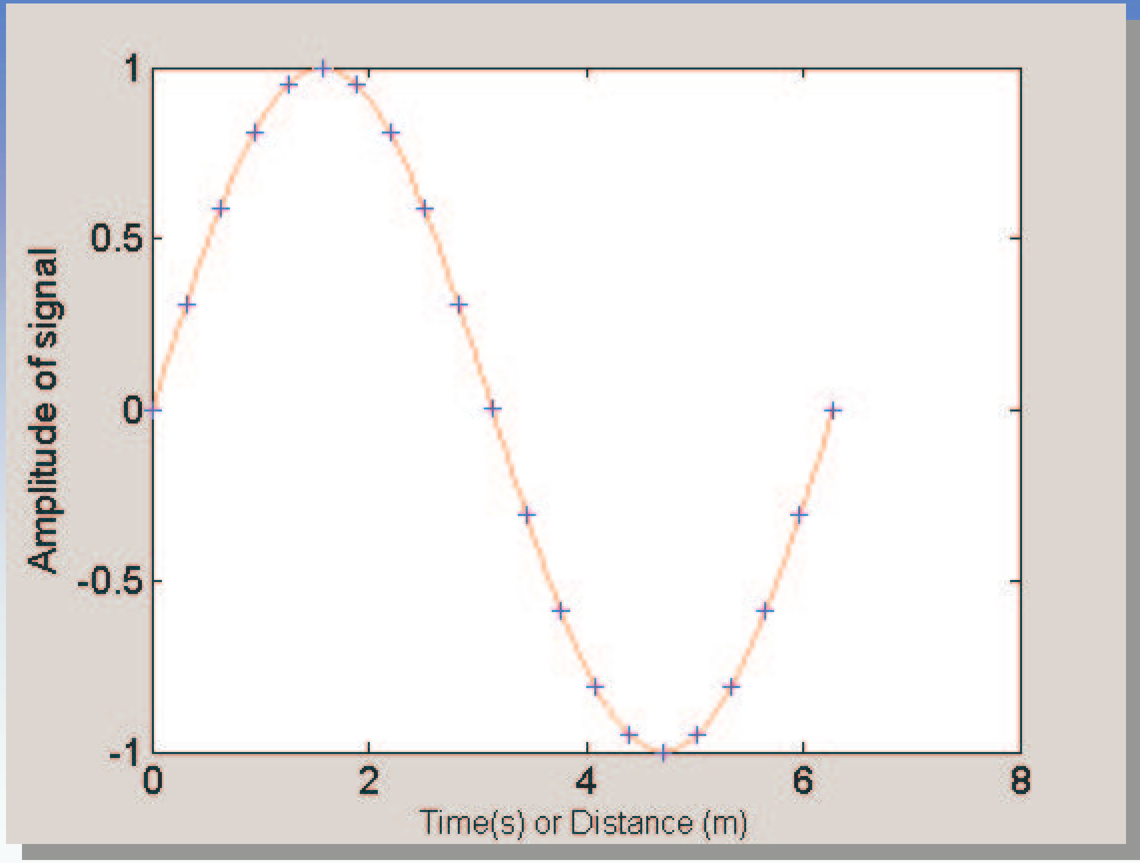

Digitization

- What happens when I digitize a signal (ground motion, temperature, etc.) in space and/or time?
- What are the consequences of a particular sampling rate on the information content?
- How can we treat (process, transform) the observed signals to extract relevant information?

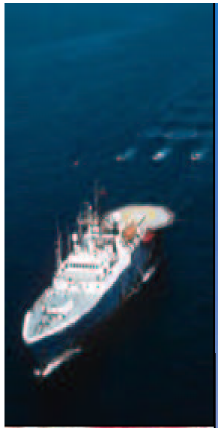
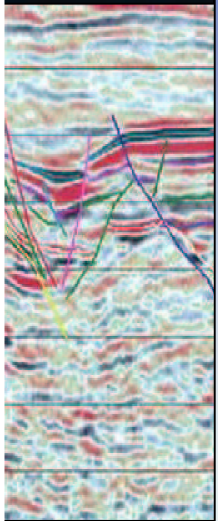




Digitization

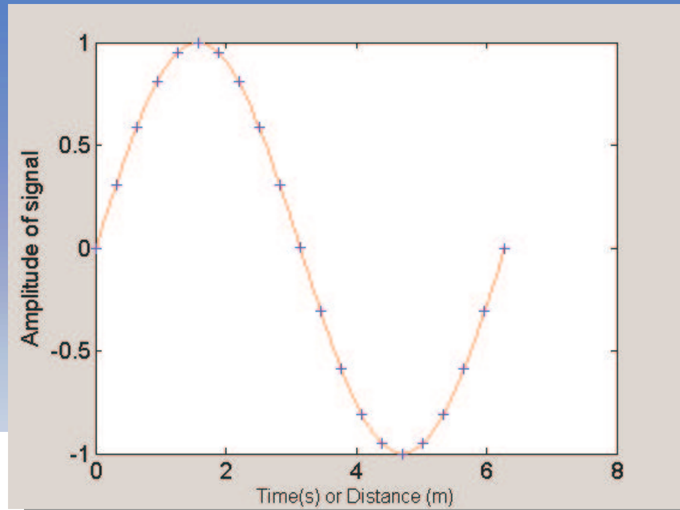


Analog and digital (+) representation of a sinusoidal function


Wavelength, period, etc.

The most important concepts you will ever need in processing observations are spatial and temporal frequencies

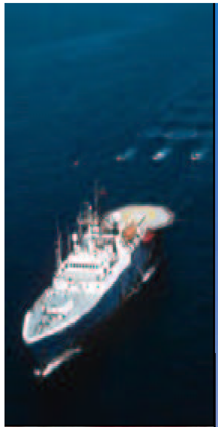
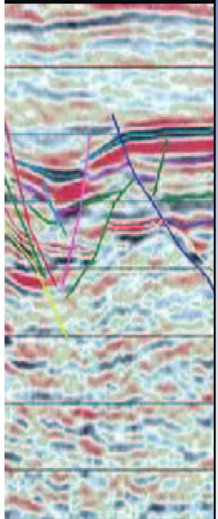


T period
f frequency
ω angular frequency

$T=1/f$
$\omega=2\pi f$

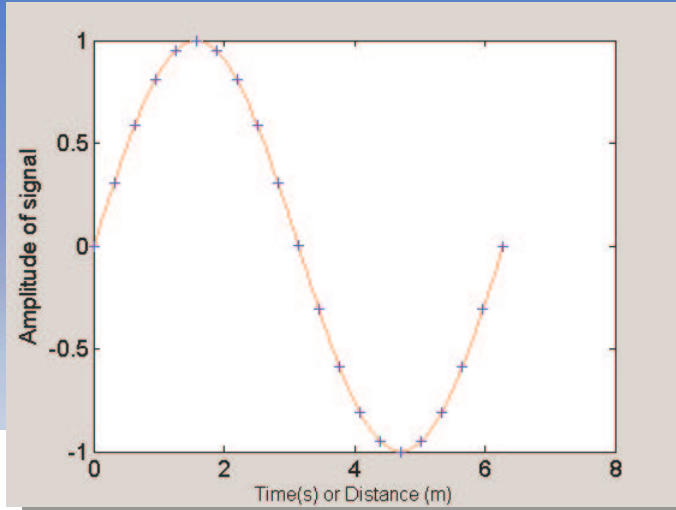
temporal frequencies

Harmonic oscillation in time:
$f(t) = A \sin(\omega t) = A \sin(2\pi f t) = A \sin((2\pi/T)\, t)$
A motion amplitude


Wavelength, period, etc.

... and the space analogue ...



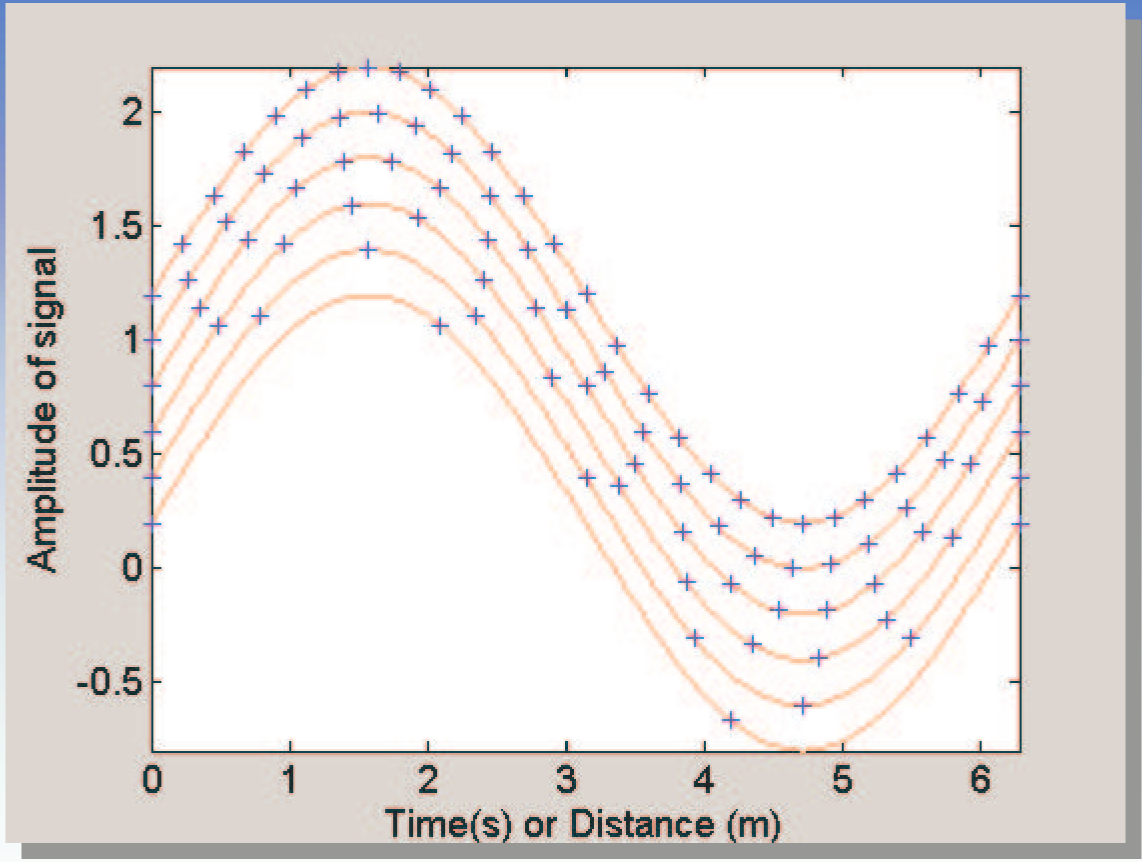
λ wavelength
k spatial wavenumber

$k=2\pi/\lambda$

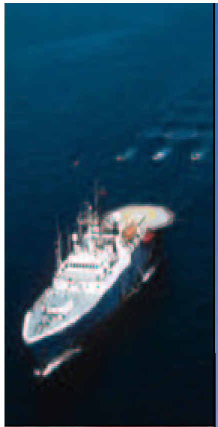
spatial frequencies

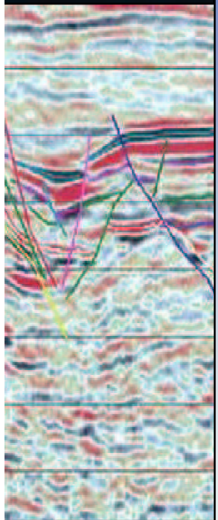
Harmonic oscillation in time:
$f(x) = A \sin(kx) = A \sin((2\pi/\lambda)\, x)$
A motion amplitude

Sampling rate

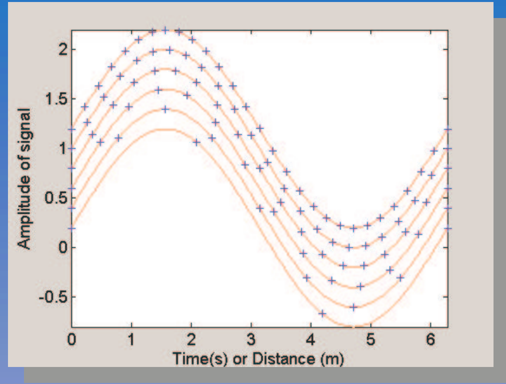


Sampling frequency, sampling rate is the number of sampling points per unit distance or unit time. Examples?

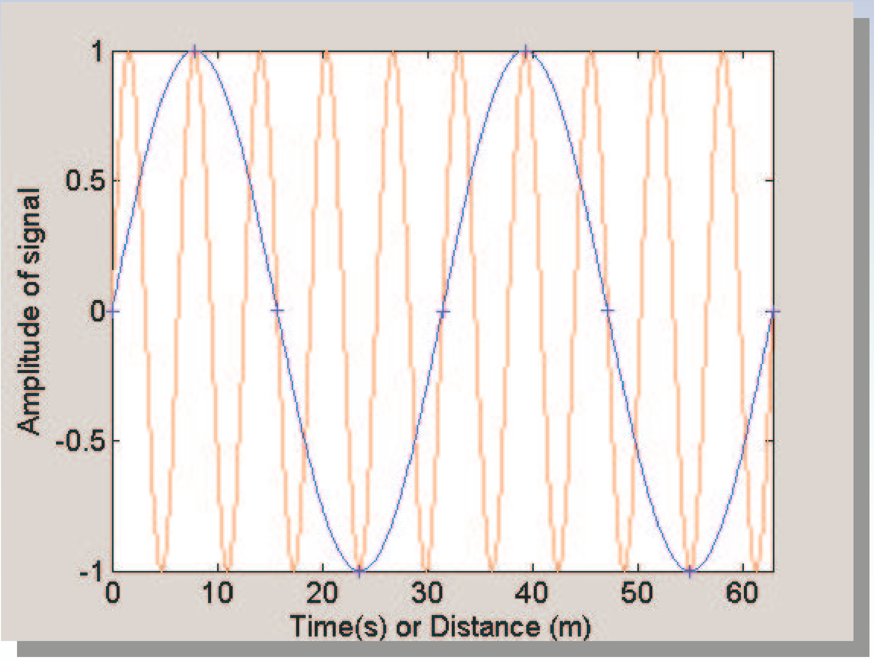




Nyquist Frequency (Wavenumber, Interval)



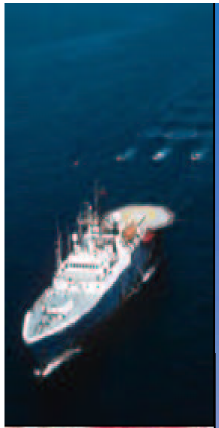
The frequency half of the sampling rate dt is called the Nyquist frequency $f_N=1/(2dt)$. The distortion of a physical signal higher than the Nyquist frequency is called aliasing.

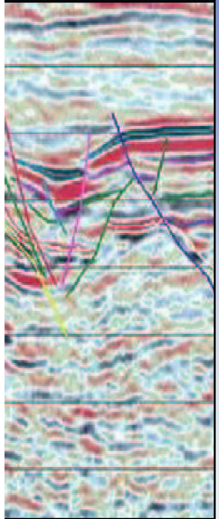


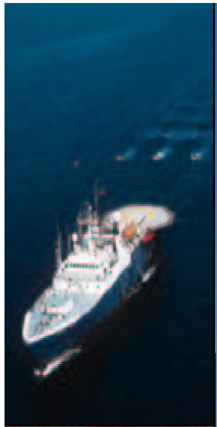
The frequency of the physical signal is > f_N is sampled with (+) leading to the erroneous blue oscillation.

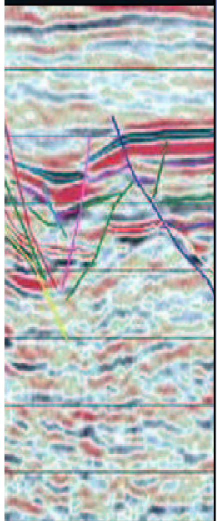
What happens in space? How can we avoid aliasing?

A cattle grid









Dynamic range

What is the precision of the sampling of our physical signal in amplitude?

Dynamic range: the ratio between largest measurable amplitude A_{max} to the smallest measurable amplitude A_{min}.

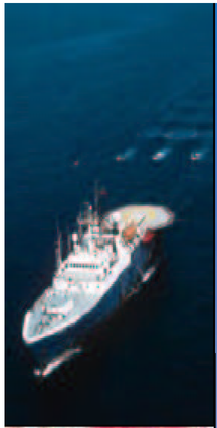
The unit is Decibel (dB) and is defined as the ratio of two power values (and power is proportional to amplitude square)

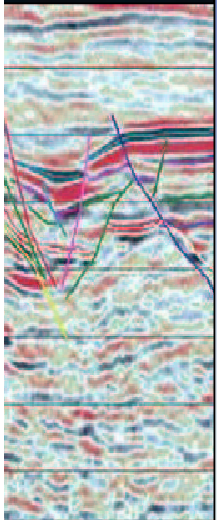
In terms of amplitudes

Dynamic range = $20 \log_{10}(A_{max}/A_{min})$ dB

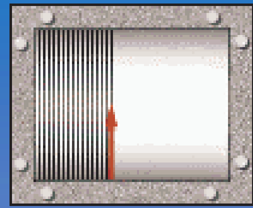
Example: with 1024 units of amplitude (A_{min}=1, A_{max}=1024)

$$20 \log_{10}(1024/1) \text{ dB} \approx 60 \text{ dB}$$

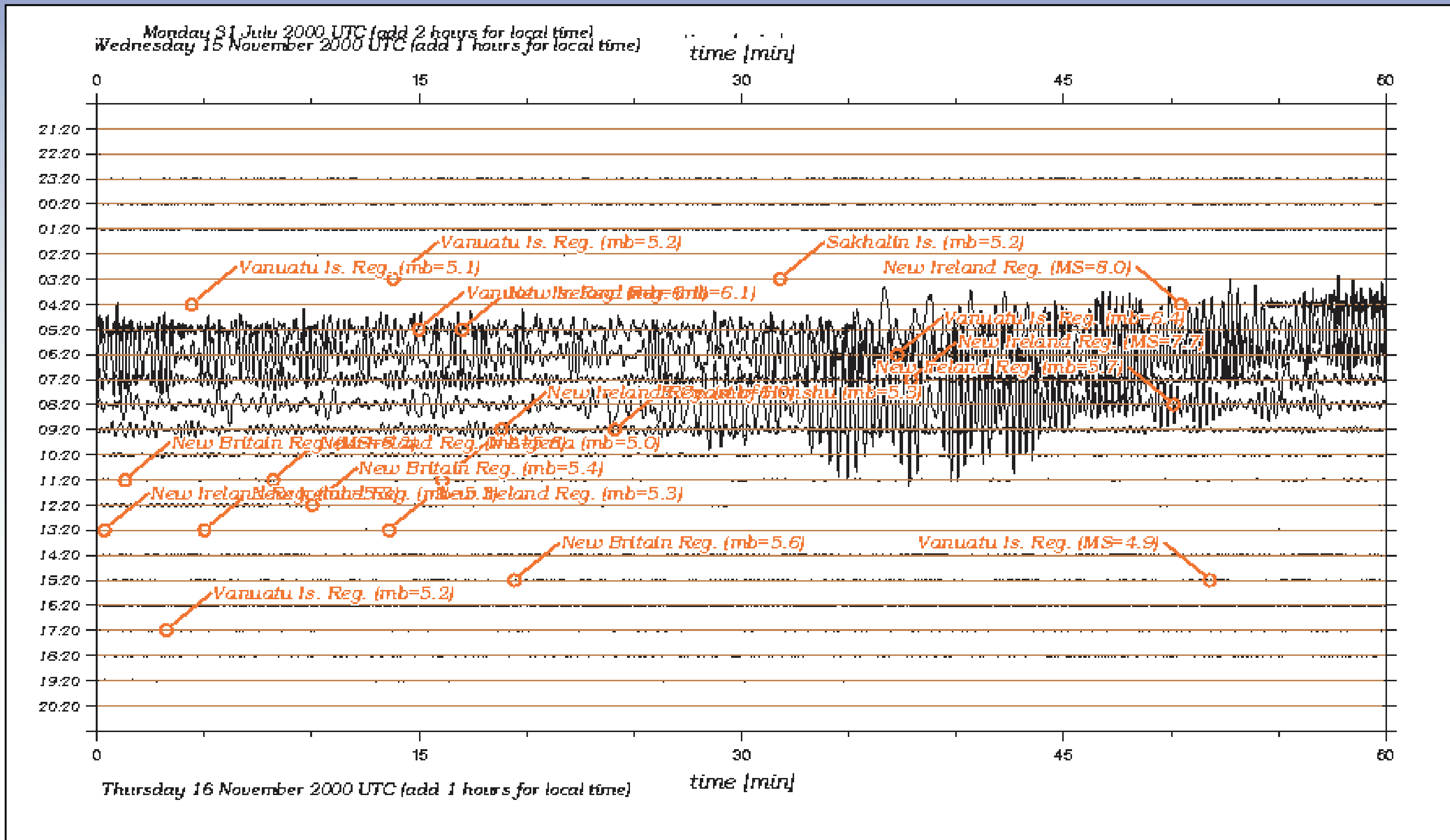


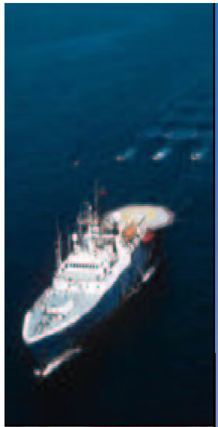
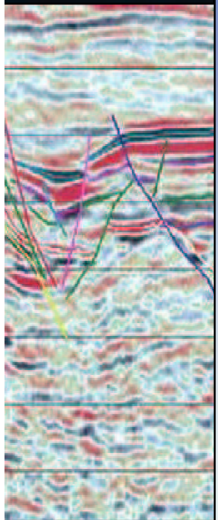


Signal and Noise



Almost all signals contain noise. The signal-to-noise ratio is an important concept to consider in all geophysical experiments. Can you give examples of noise in the various methods?



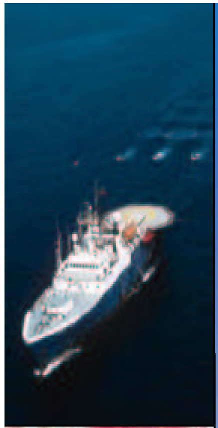



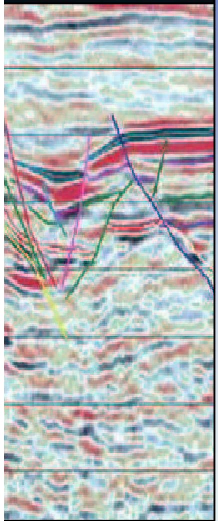

Spectral Analysis

Spectral analysis is so central to all natural (and other) sciences that it's role can not be **overemphasized**!

With spectral analysis we can get answers to questions like:

- What (spatial or temporal) frequencies are contained in my signal?
- Is there a periodic signal in my observations?
- Do I have to account for the response of my measurement instrument (e.g. seismometer) to obtain the physical signal?
- Do I have to filter the observations first before seeing the physical signal?
- and, and, and …





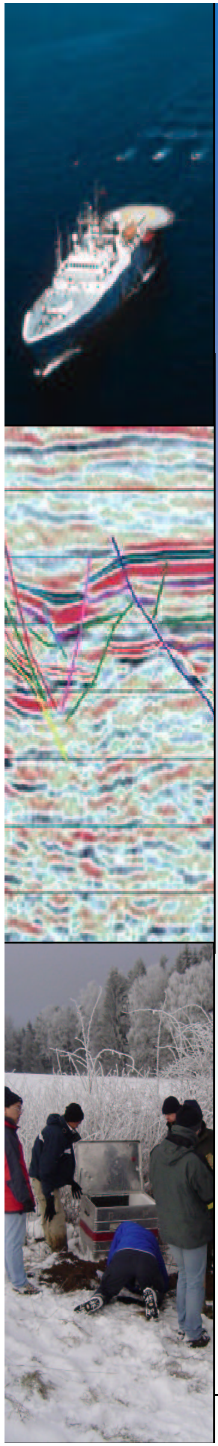
Harmonic Analyis – Spectral Synthesis

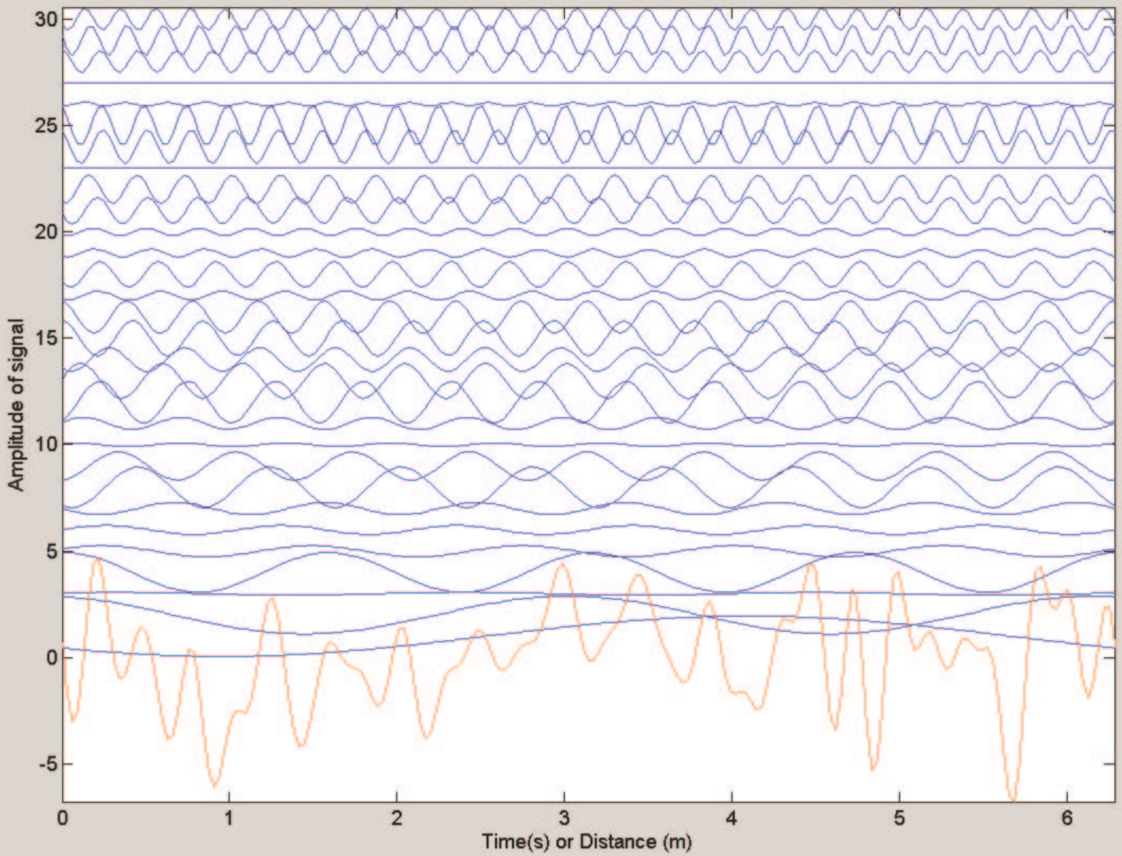
At the heart of spectral analyis is an extremely powerful concept, that is one of the most important theorems in mathematical physics:

Any arbitrary signal can be obtained by superposition of harmonic (sinusoidal) signals.

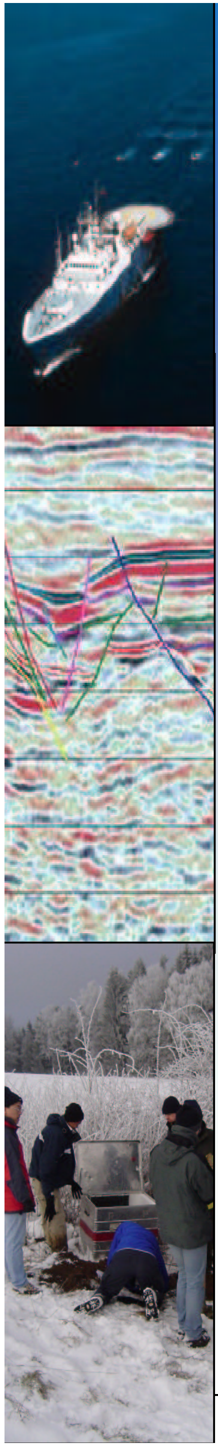
Furthermore: the representation of physical systems in time and space or in frequency and wavenumber domain is equivalent! There is no loss of information when going from one space to the other and back.

Spectral synthesis





The red trace is the sum of all blue traces!

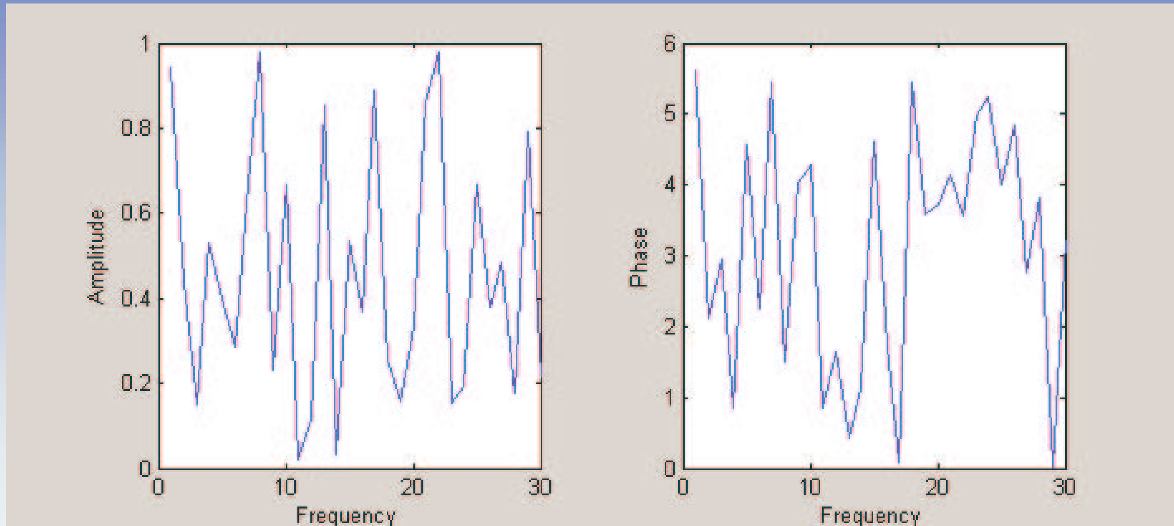


The spectrum

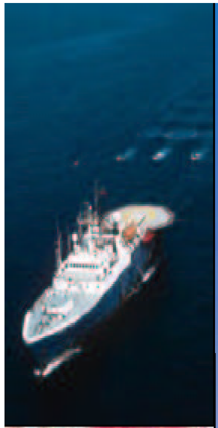
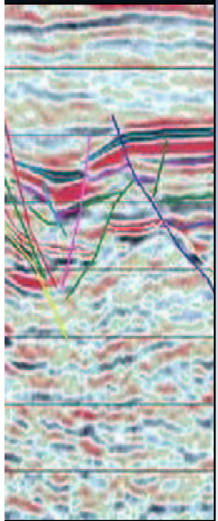
Amplitude spectrum

Phase spectrum

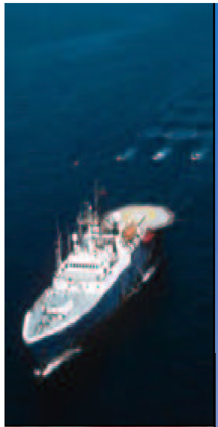
Fourier space

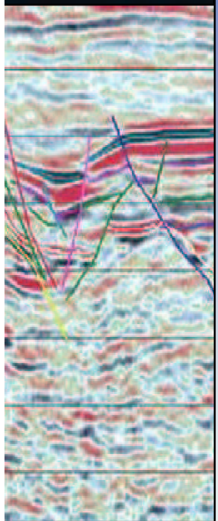


Physical space


Fourier decomposer





Let's get formal!

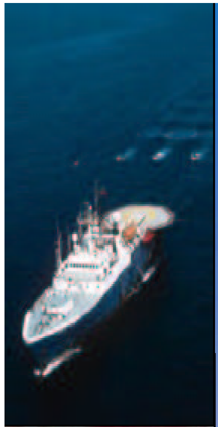
A sinusoidal function is represented as (a amplitude, λ wavelength):

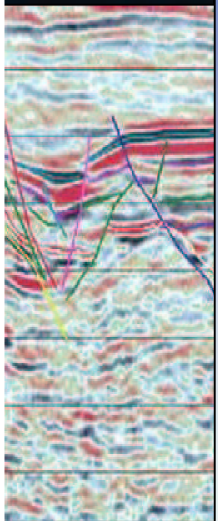
$$y = a \sin\left(\frac{2\pi x}{\lambda}\right)$$

Ignoring phase shifts an arbitrary signal (a_0 at both ends) can be obtained by superposition of

$$f(x) = a_0 + \sum_n a_n \sin\left(2\pi x \frac{n}{2L}\right) \qquad n = 1, \infty$$

Here L is the length of the space (or time) domain. Note the sequence of wavelengths (or periods): 2L, L, 2/3L, L/2 …





The Fourier components

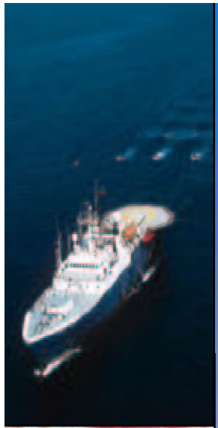
The amplitudes (a_n) of the signal (Fourier components) can be obtained by integrating over the signal:

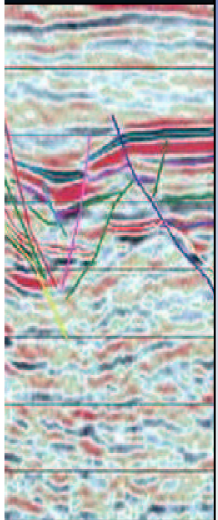
$$a_0 = \frac{1}{L}\int_0^L f(x)dx$$

Average value of signal

$$a_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}dx$$

Spectral component





Fourier: Space and Time

Space	
x	space variable
L	spatial wavelength
$k=2\pi/\lambda$	spatial wavenumber
F(k)	wavenumber spectrum

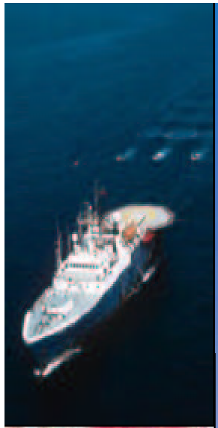
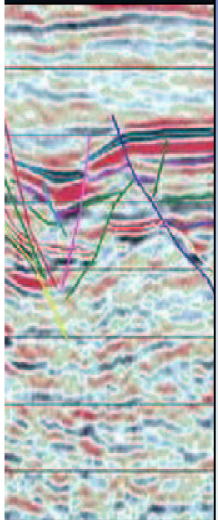
Time	
t	Time variable
T	period
f	frequency
$\omega=2\pi f$	angular frequency

Fourier integrals

With the complex representation of sinusoidal functions e^{ikx} (or ($e^{i\omega t}$) the Fourier transformation can be written as:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) e^{-ikx} dx$$

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{ikx} dx$$


The Fourier Transform
discrete vs. continuous

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

continuous

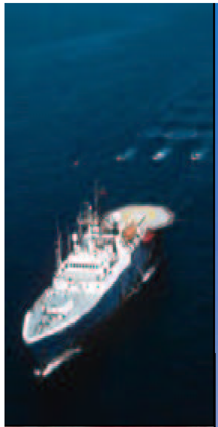
$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) e^{-ikx} dx$$

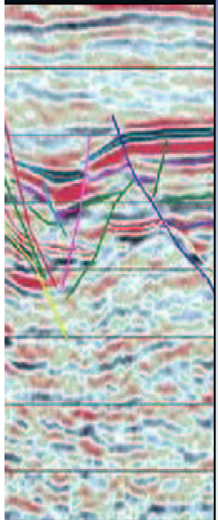
$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{ikx} dx$$

discrete

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

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





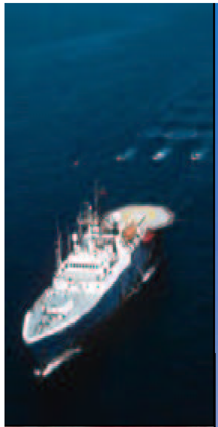
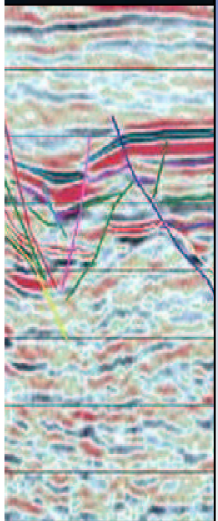
The Fast Fourier Transform (FFT)

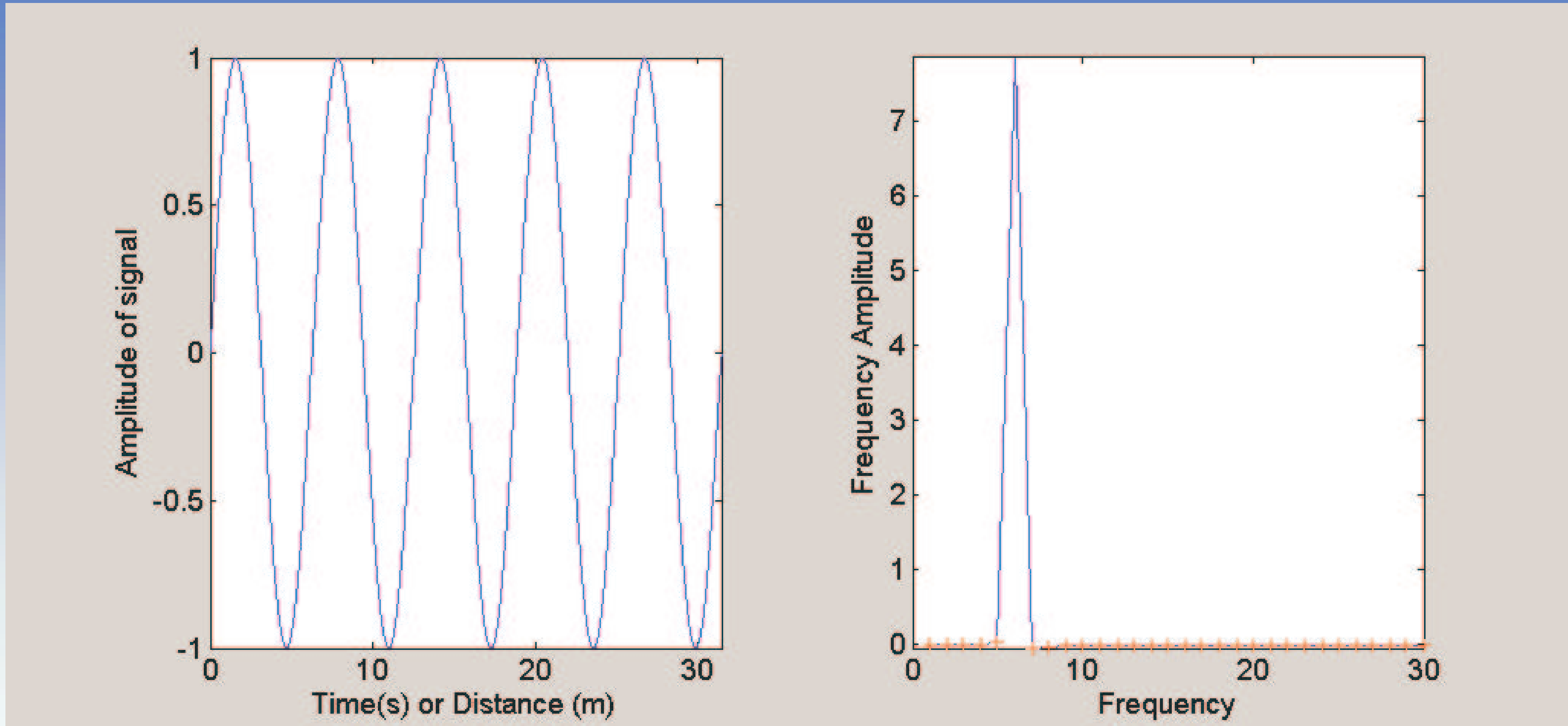
Most processing tools (e.g. octave, Matlab, Mathematica, Fortran, etc) have intrinsic functions for FFTs

Matlab FFT

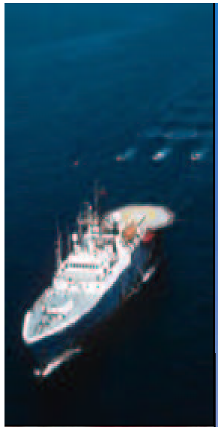
```
>> help fft

 FFT Discrete Fourier transform.
FFT(X) is the discrete Fourier transform (DFT) of vector X.  For
matrices, the FFT operation is applied to each column. For N-D
arrays, the FFT operation operates on the first non-singleton
dimension.

FFT(X,N) is the N-point FFT, padded with zeros if X has less
than N points and truncated if it has more.

FFT(X,[],DIM) or FFT(X,N,DIM) applies the FFT operation across the
dimension DIM.

For length N input vector x, the DFT is a length N vector X,
with elements
                 N
  X(k) =       sum  x(n)*exp(-j*2*pi*(k-1)*(n-1)/N), 1 <= k <= N.
                n=1
The inverse DFT (computed by IFFT) is given by
                 N
  x(n) = (1/N) sum  X(k)*exp( j*2*pi*(k-1)*(n-1)/N), 1 <= n <= N.
                k=1

See also IFFT, FFT2, IFFT2, FFTSHIFT.
```

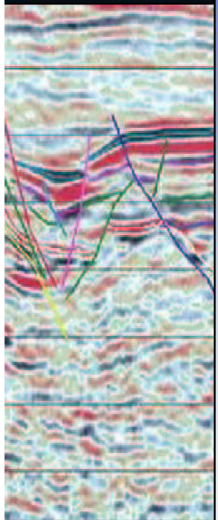
Fourier Spectra: Main Cases
harmonic signals

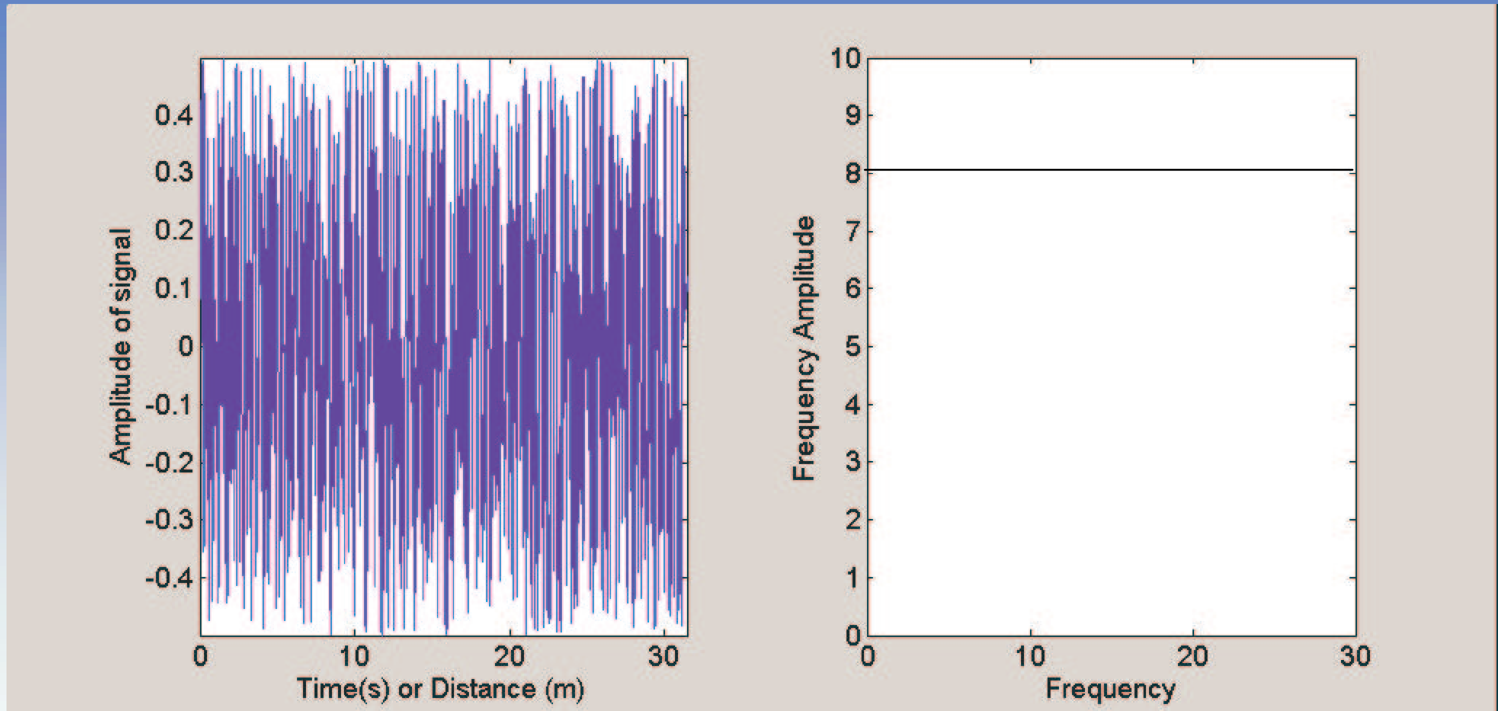



The spectrum of a (monochromatic) harmonic signal in space (or time) is a spike („Delta function") in the frequency domain.

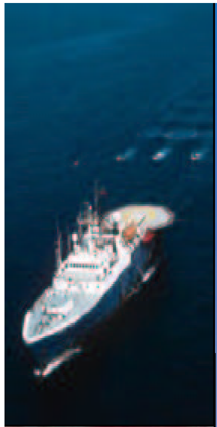


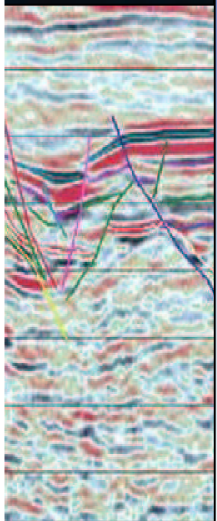


Fourier Spectra: Main Cases
random signals

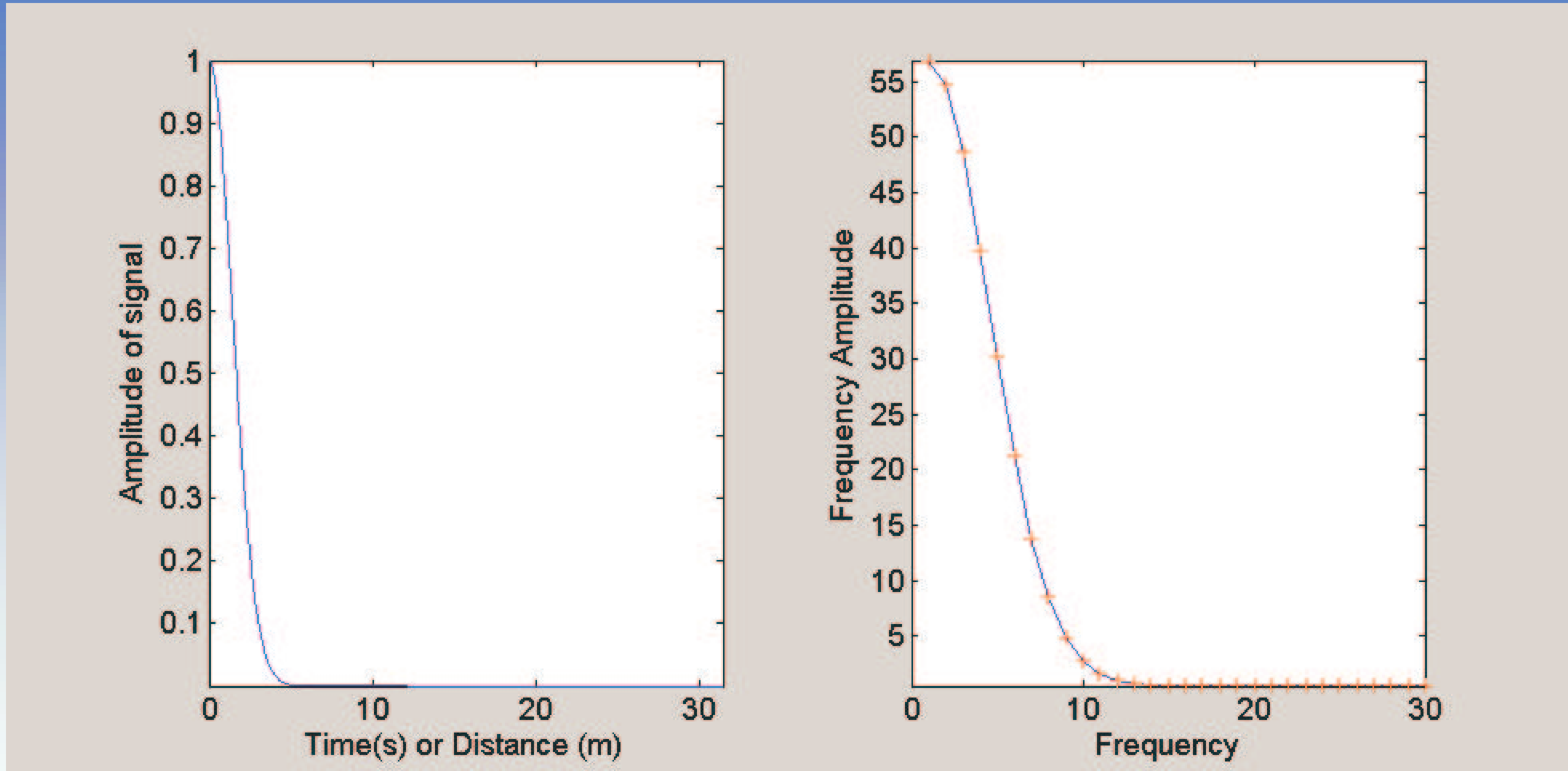


Random signals contain all frequencies. A spectrum with constant contribution of all frequencies is called a white spectrum

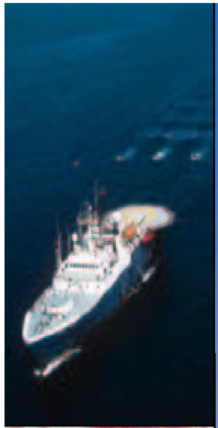


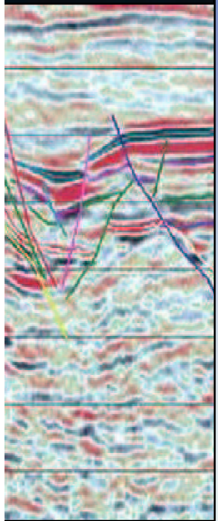


Fourier Spectra: Main Cases
Gaussian signals

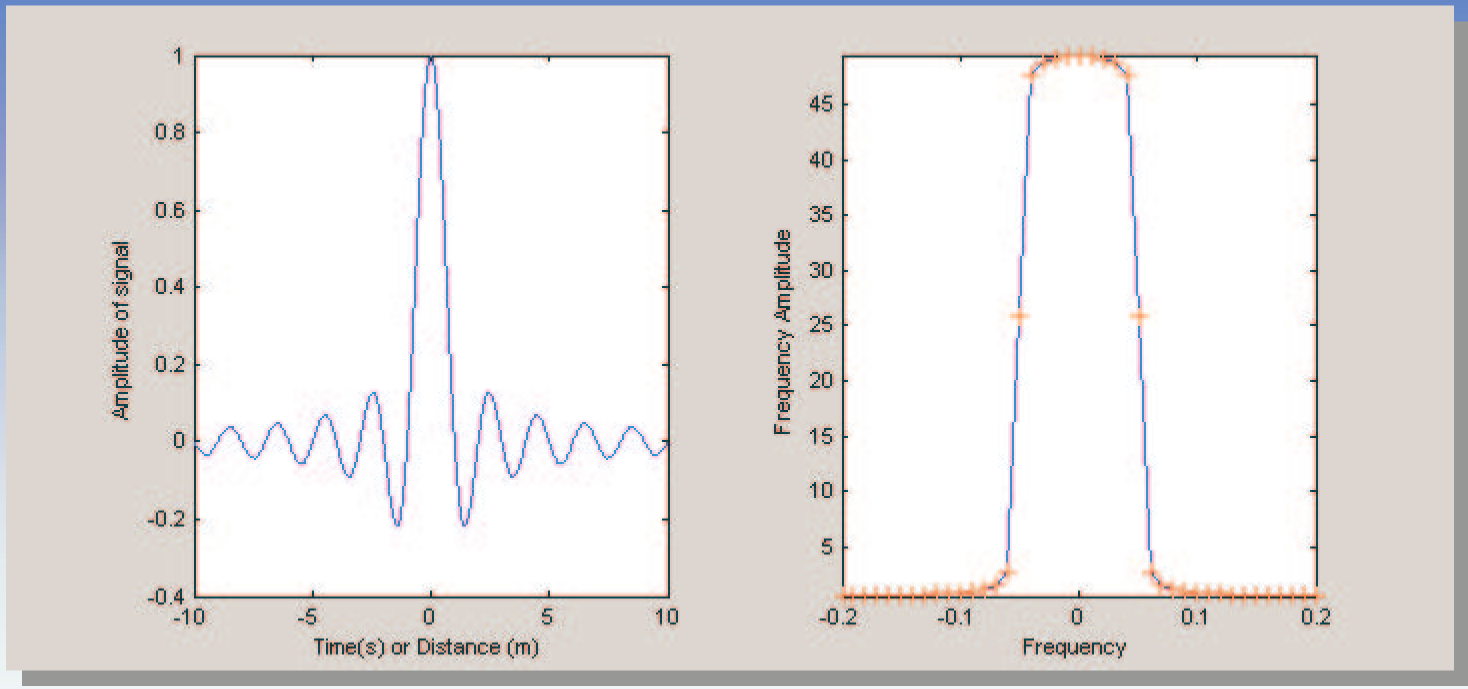


The spectrum of a Gaussian function will itself be a Gaussian function. How does the spectrum change, if I make the Gaussian narrower and narrower?



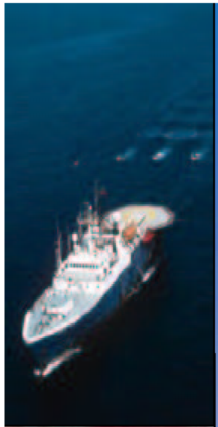


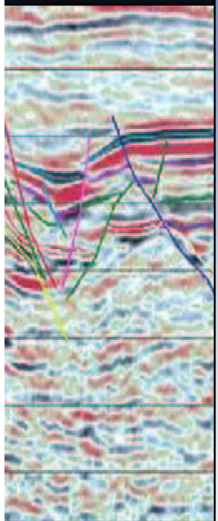
Fourier Spectra: Main Cases
Transient waveform



A transient wave form is a wave form limited in time (or space) in comparison with a harmonic wave form that is infinite

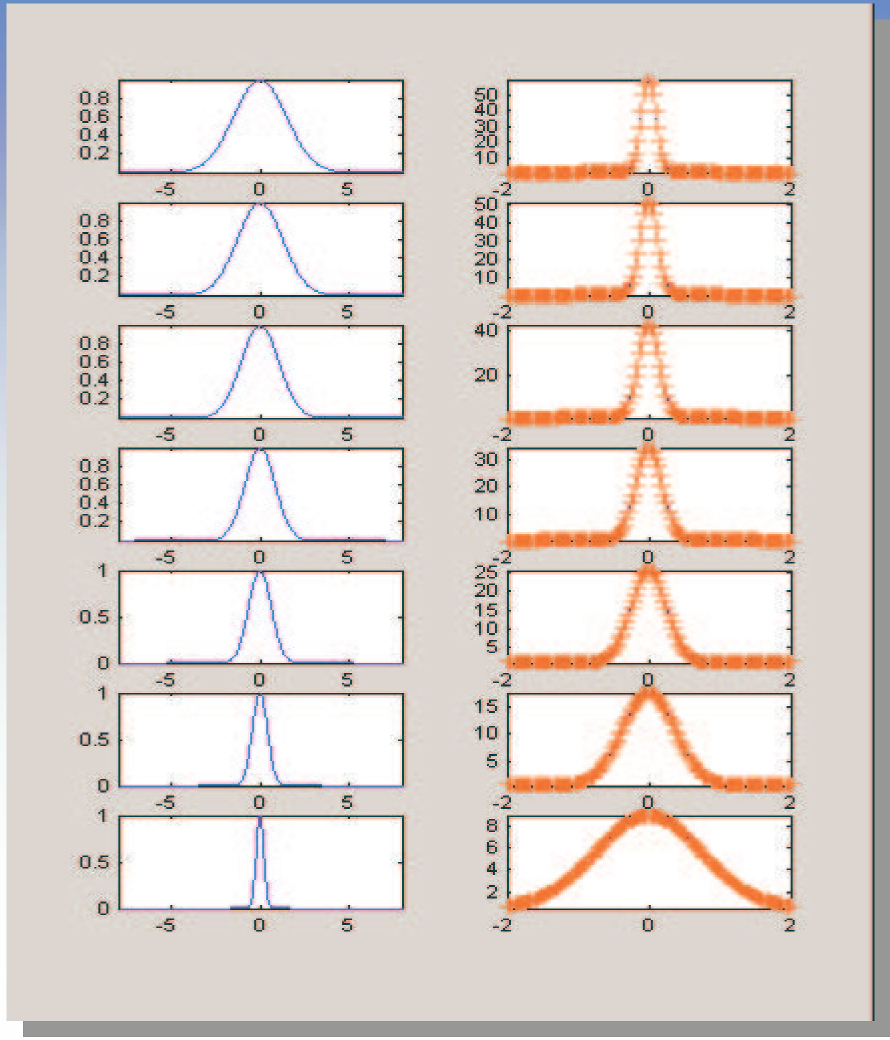
Puls-width and Frequency Bandwidth



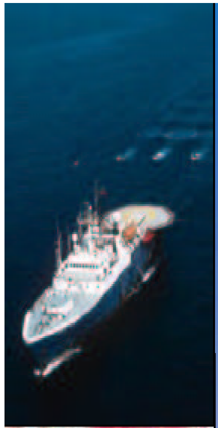


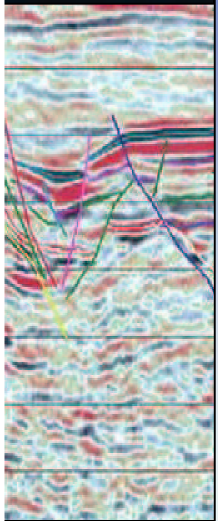
time (space) | spectrum

Narrowing physical signal

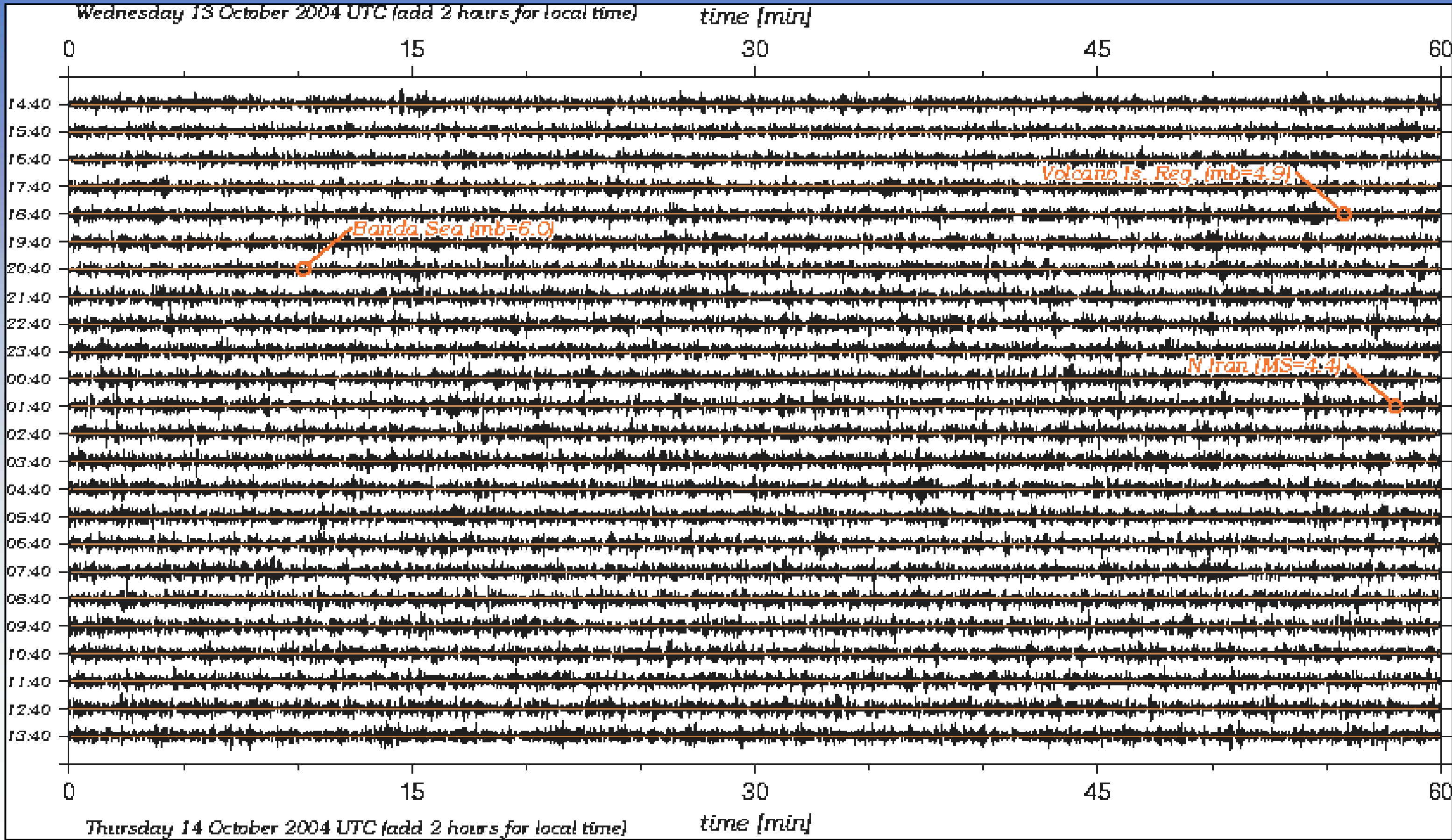


Widening frequency band

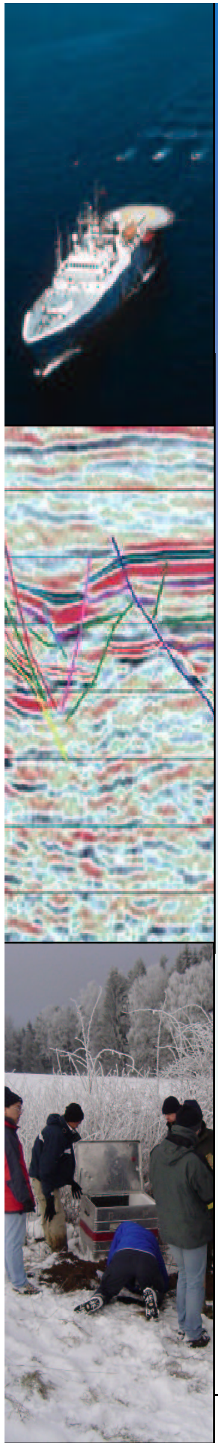




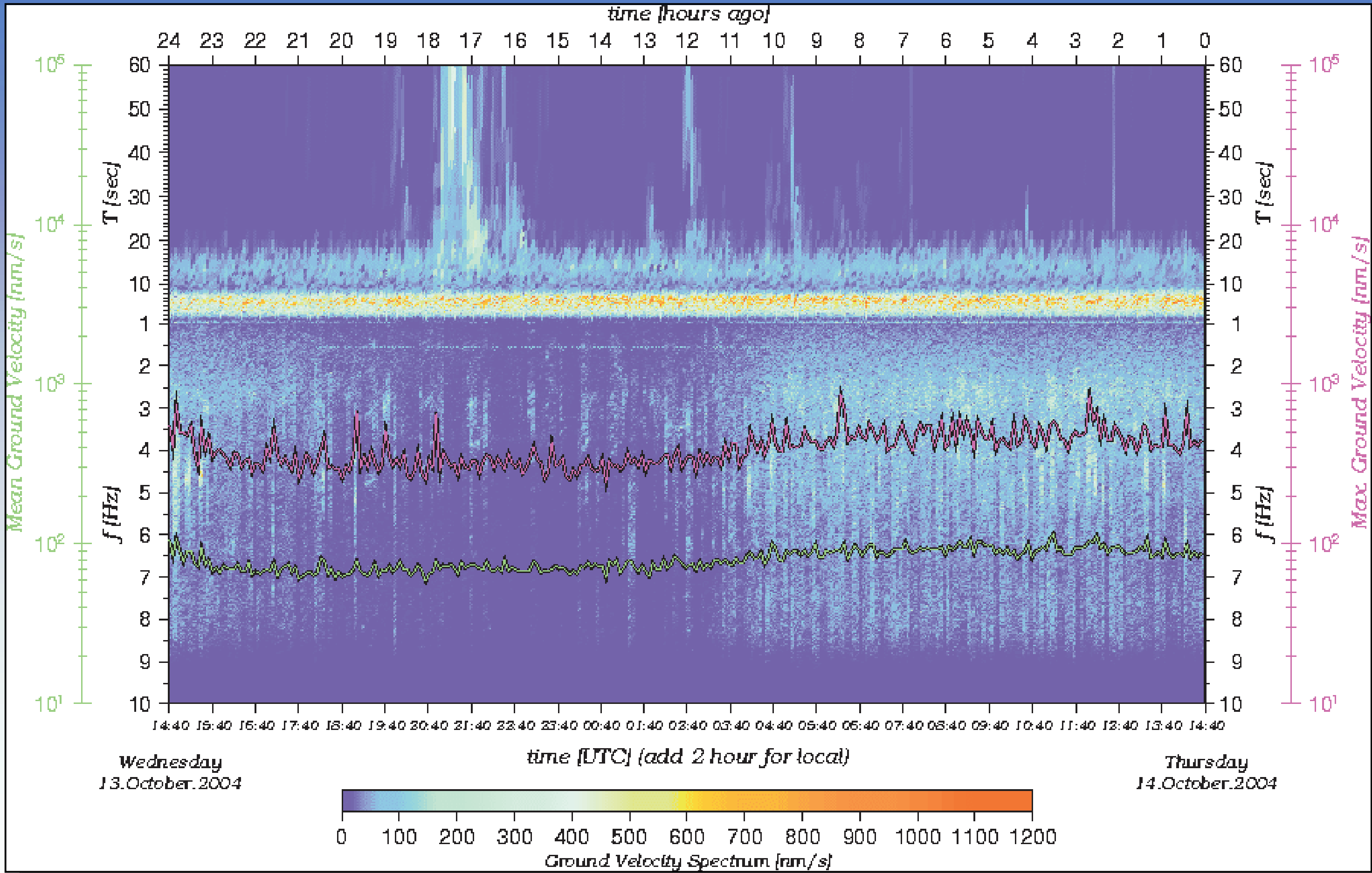
Spectral analysis: an Example



24 hour ground motion, do you see any signal?

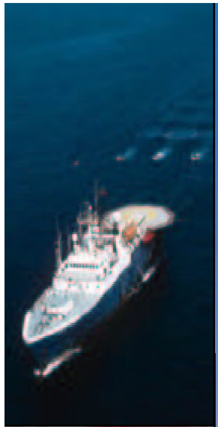


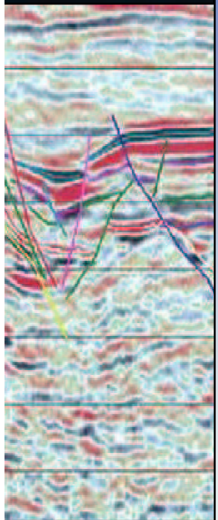
Seismo-Weather

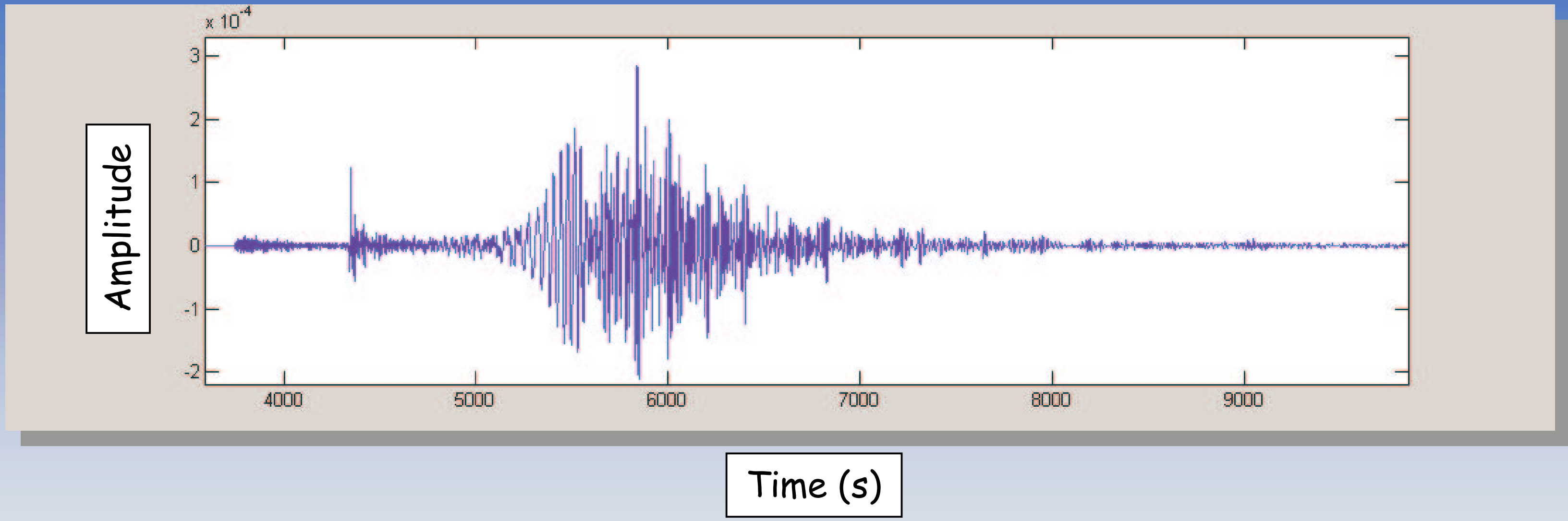


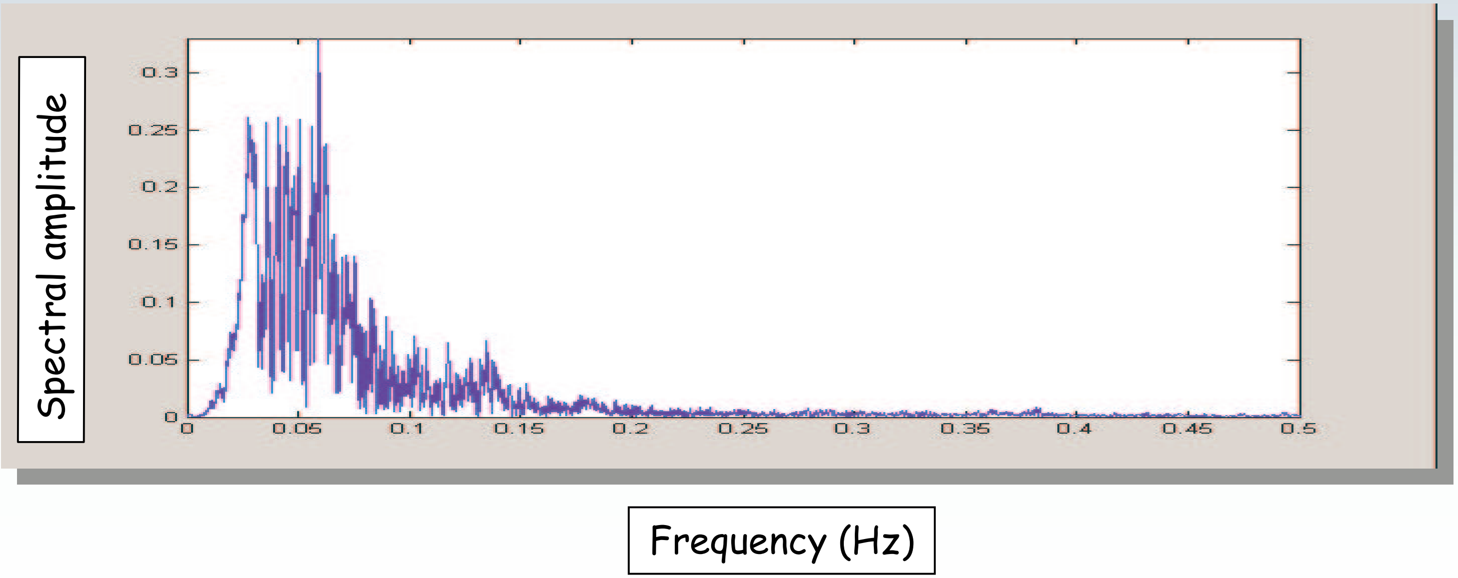
Running spectrum of the same data

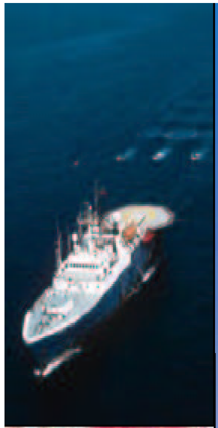
A seismogram

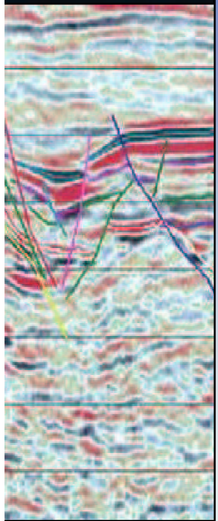








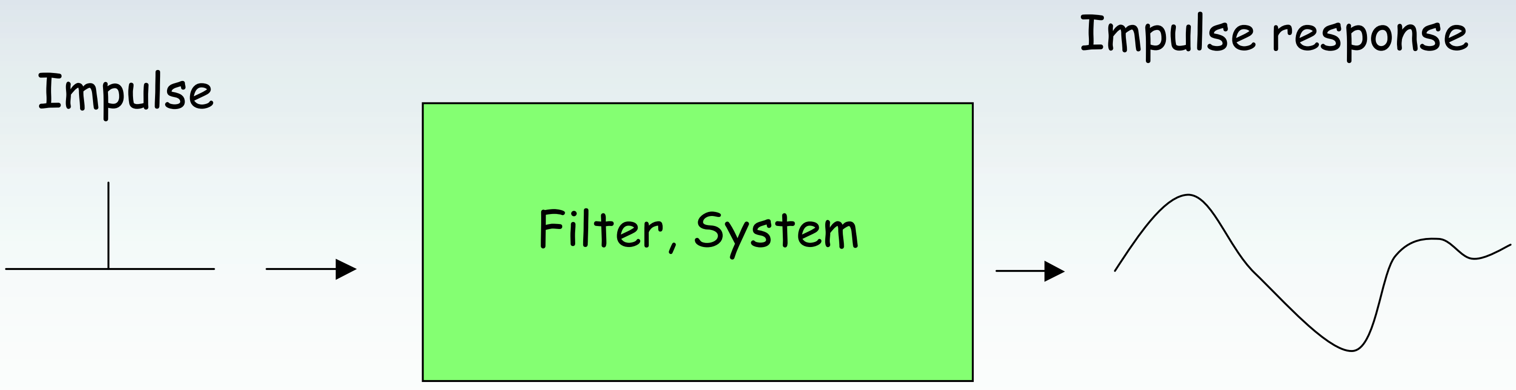




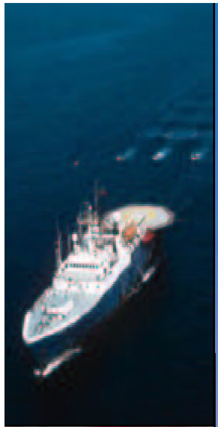
Waveform processing

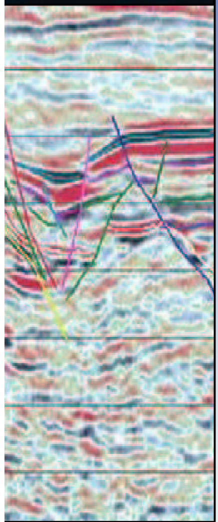
How can (must) we treat our digitized observations to extract information on our physical signals? This question leads us to the concepts of (de-) convolution, (auto-,cross-) correlation and filtering.

The central concept is the response of a system to an implusive input; the impulse response

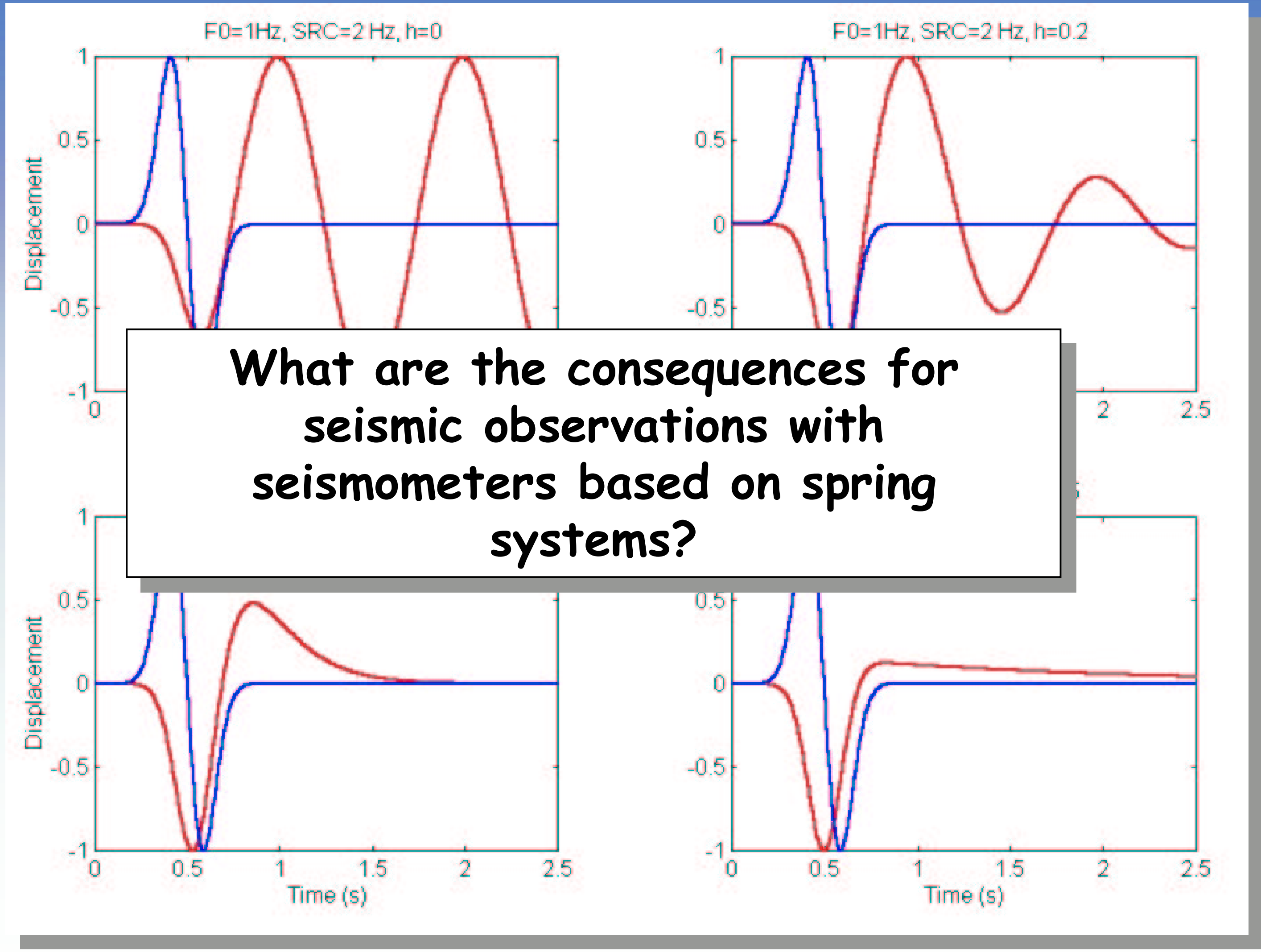


Can you give examples?

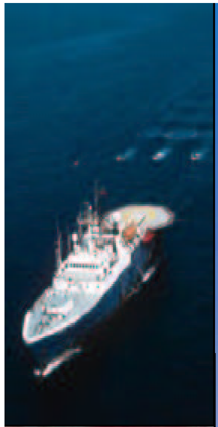
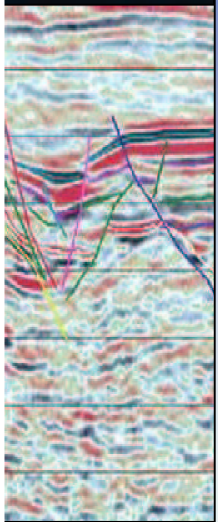




Impulse response of a seismometer



What are the consequences for seismic observations with seismometers based on spring systems?


Discrete Convolution

Convolution is the mathematical description of the change of waveform shape after passage through a filter (system).

There is a special mathematical symbol for convolution (*):

$$y(t) = g(t) * f(t)$$

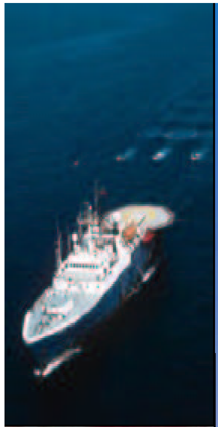
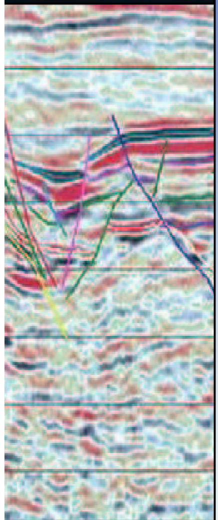
Here the impulse response function g is convolved with the input signal f. g is alse named the „Green's function"

$$y_k = \sum_{i=0}^{m} g_i f_{k-i}$$

$$k = 0,1,2,\ldots,m+n+1$$

$$g_i \qquad i = 0,1,2,\ldots.,m$$

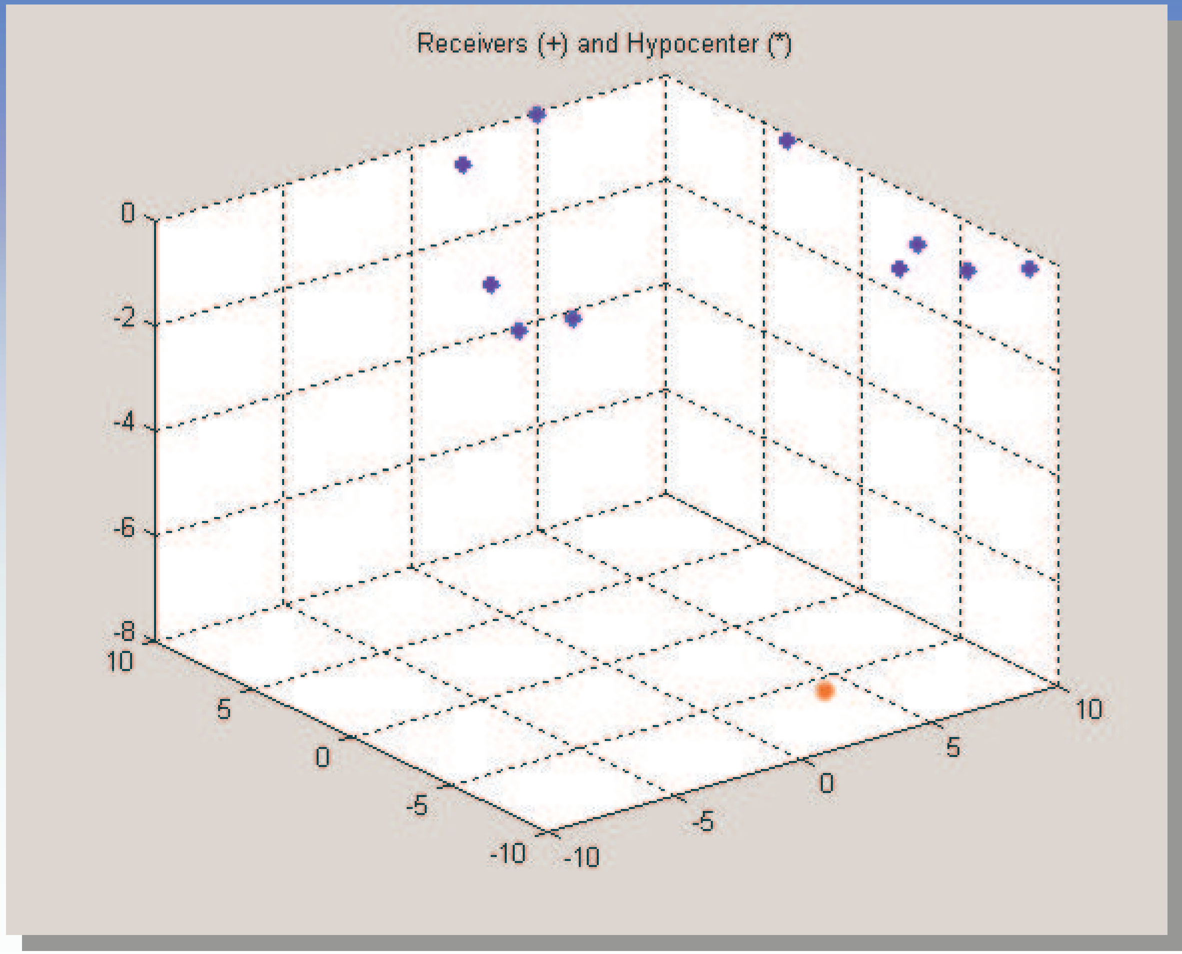
$$f_j \qquad j = 0,1,2,\ldots.,n$$

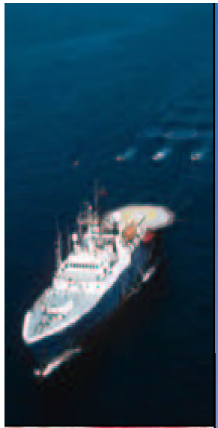



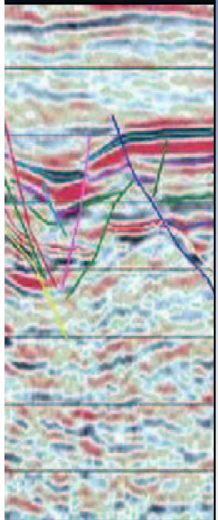

Convolution Example (Matlab)

```
>> x
x =
     0     0     1     0
>> y
y =
     1     2     1
>> conv(x,y)
ans =
     0     0     1     2     1     0
```

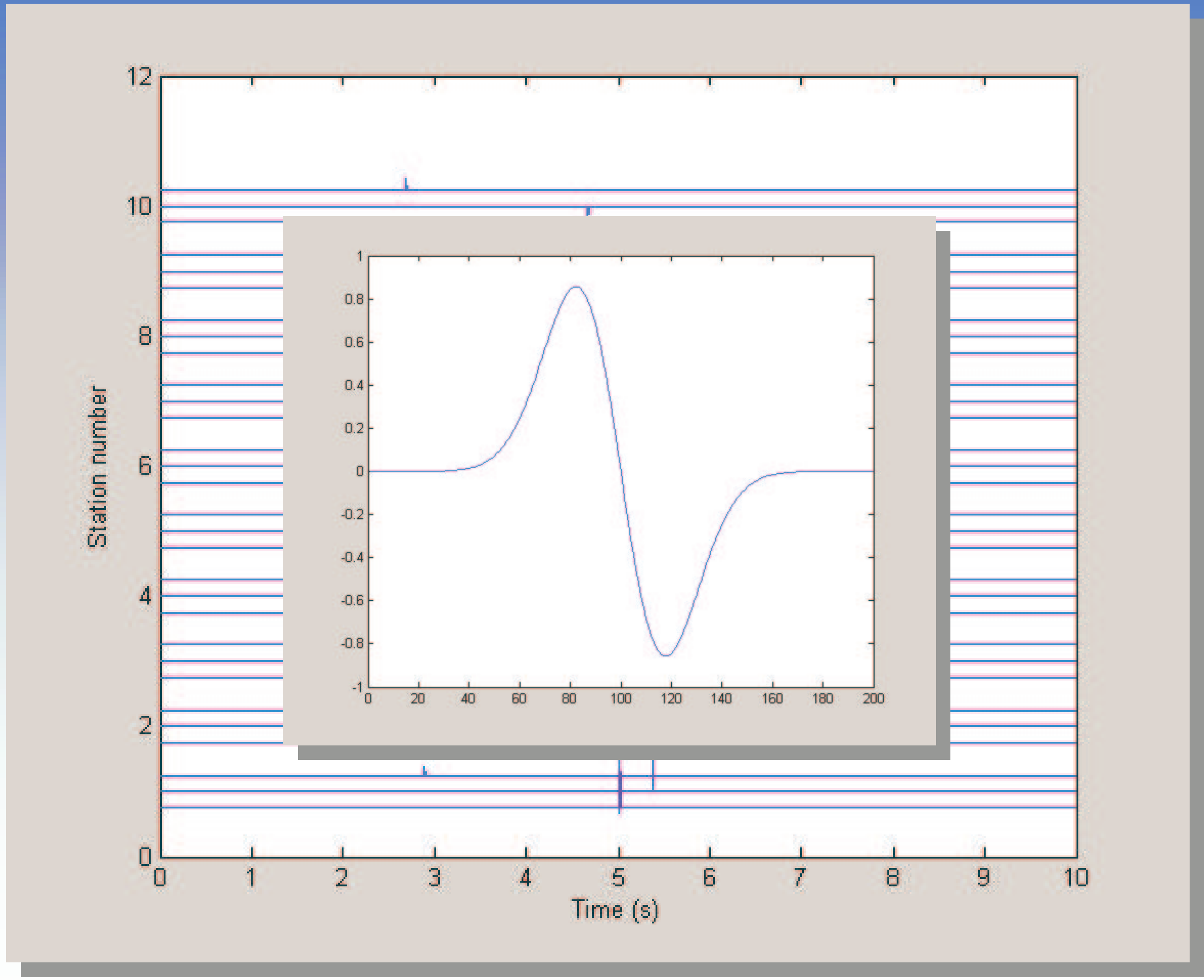
Convolutional model:
seismograms



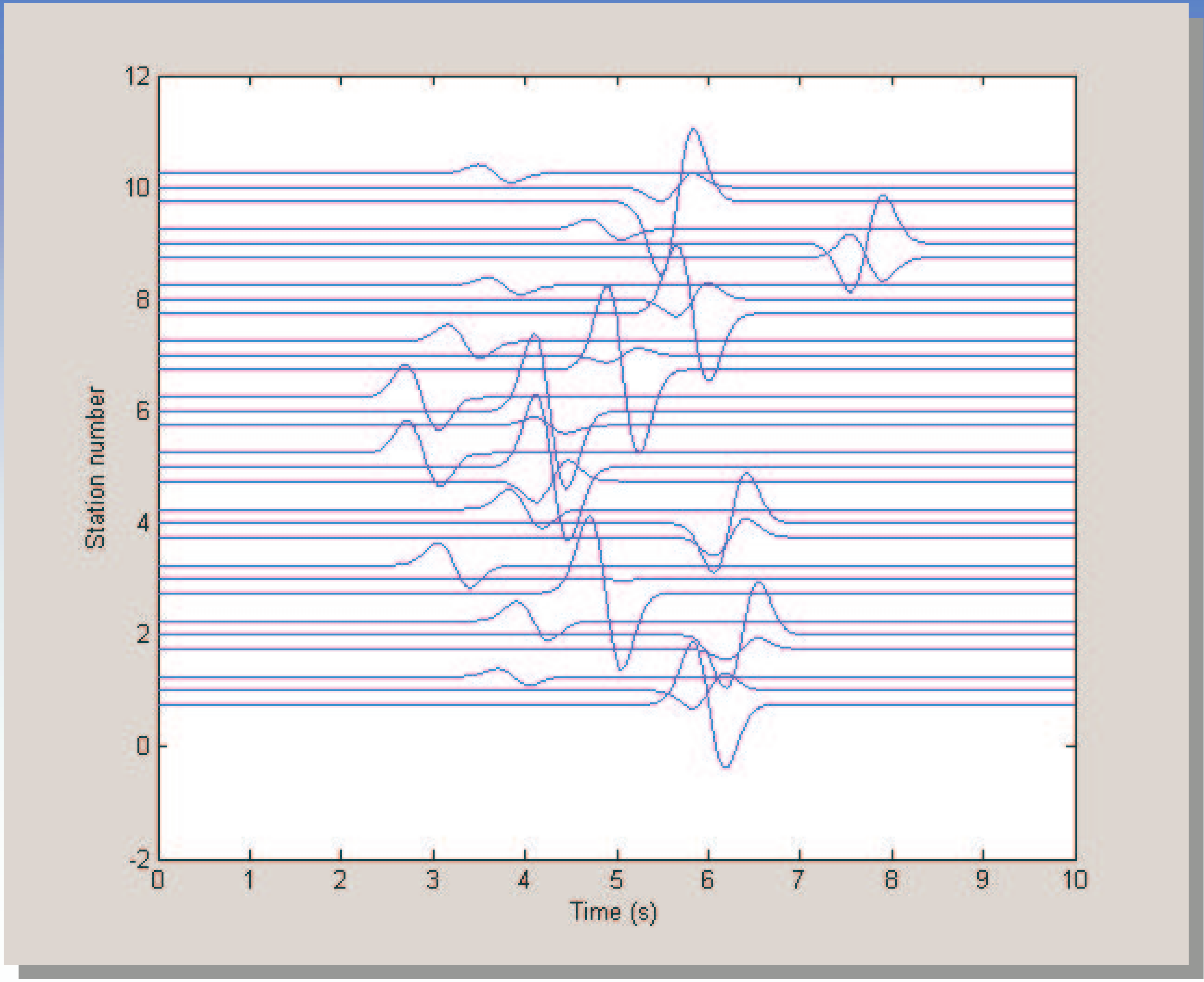


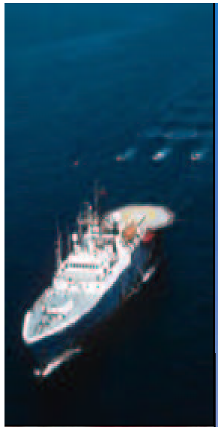


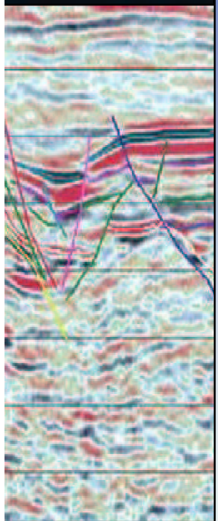
The seismic *impulse response*



The filtered response

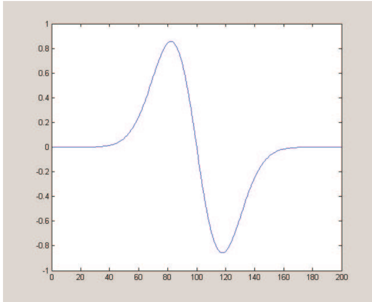






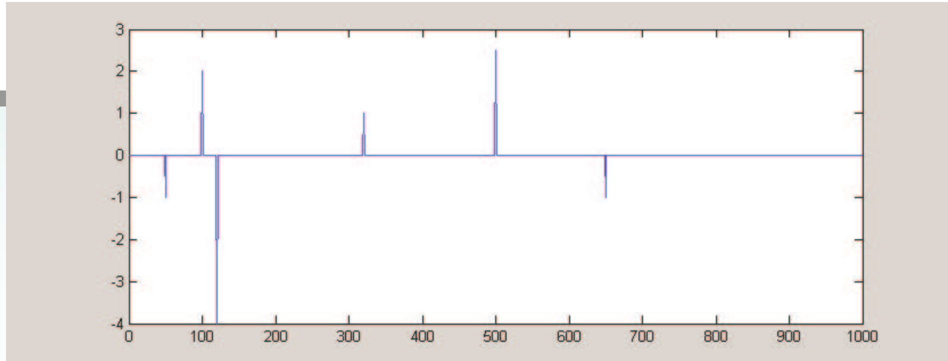
1D convolutional model of a seismic trace

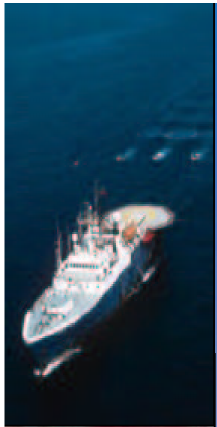
The seismogram of a layered medium can also be calculated using a convolutional model ...

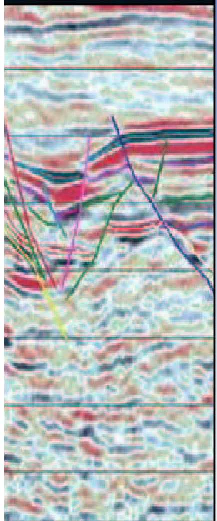


$$u(t) = s(t) * r(t) + n(t)$$

u(t)	seismogram
s(t)	source wavelet
r(t)	reflectivity



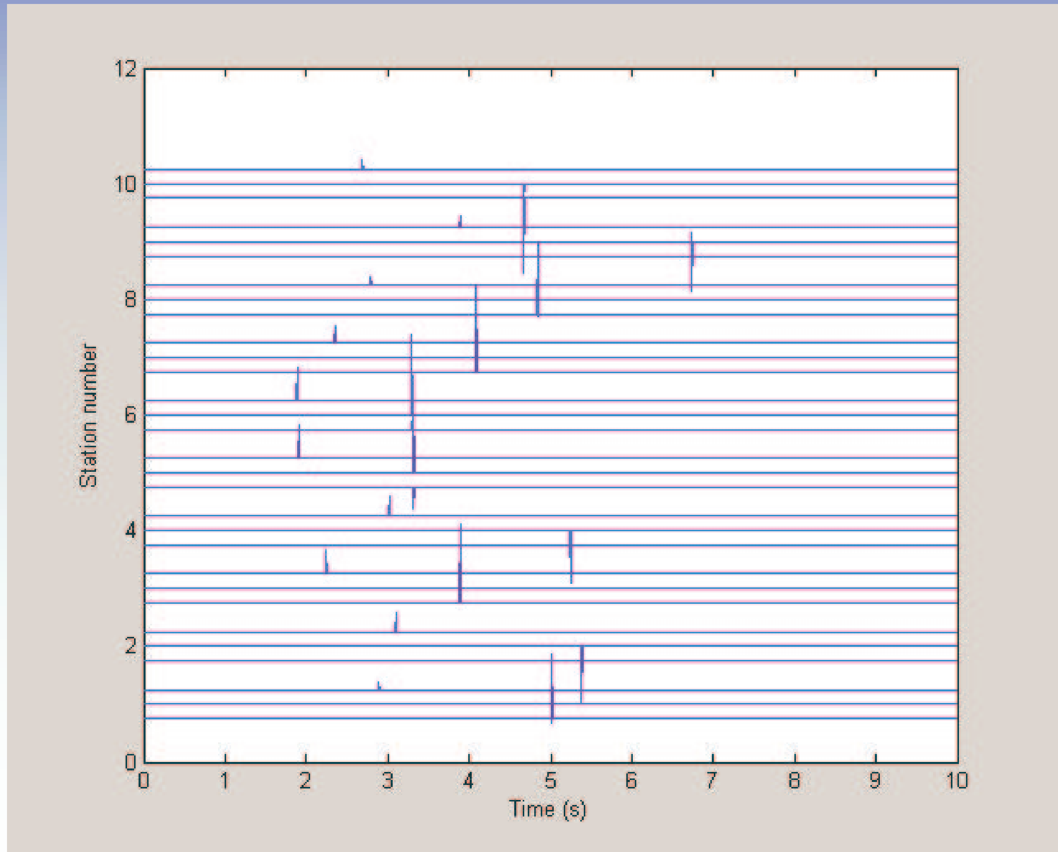


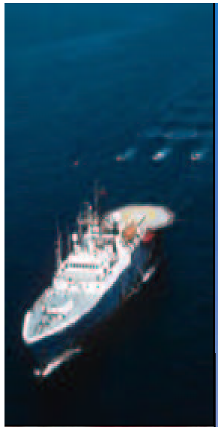
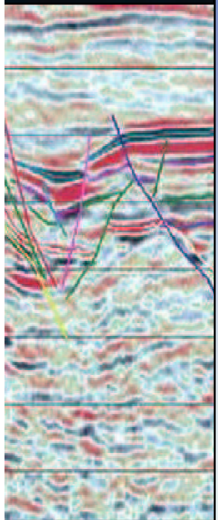


Deconvolution

Deconvolution is the inverse operation to convolution.

When is deconvolution useful?




Correlation

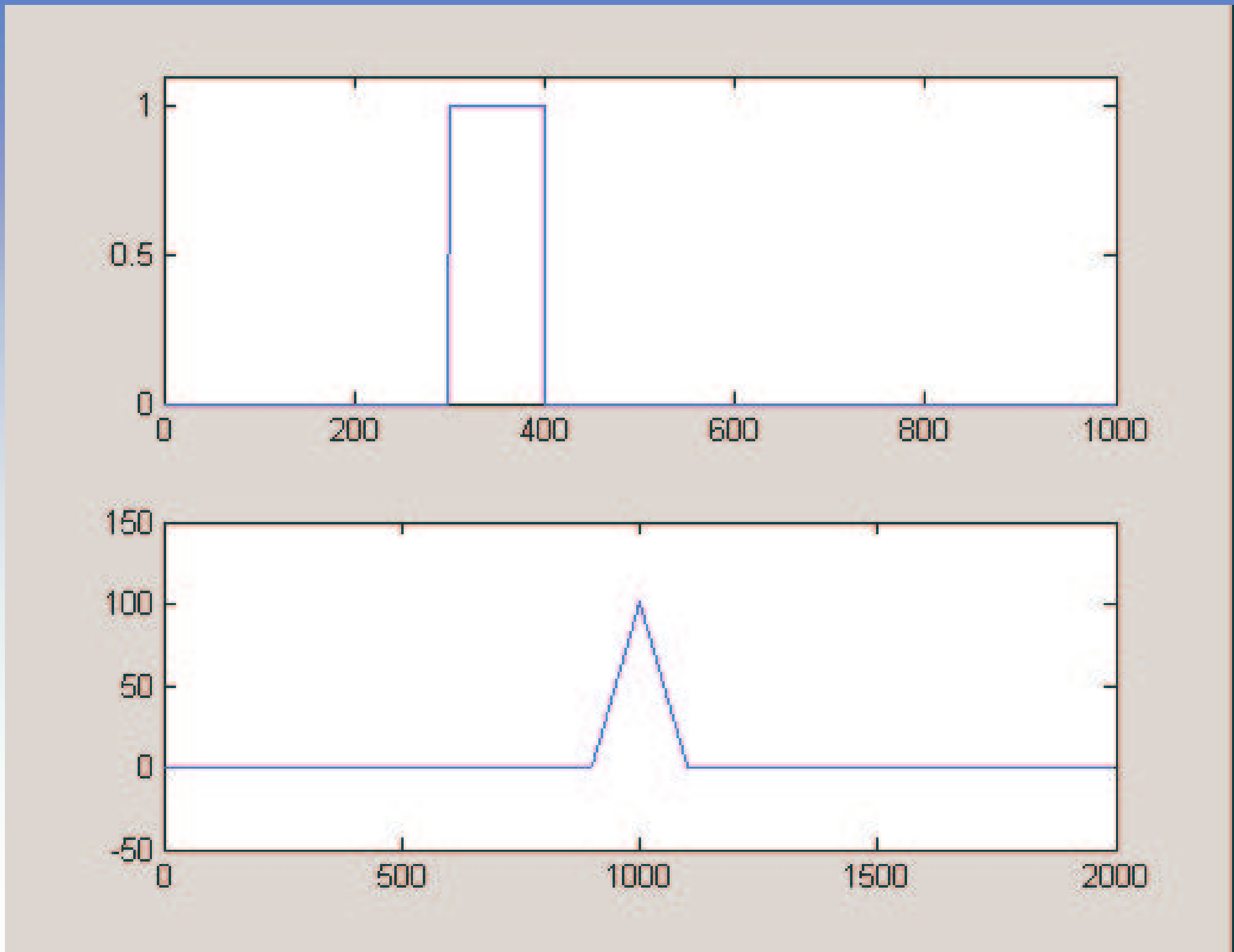
Correlation plays a central role in the study of time series. In general, correlation gives a quantitative estimate of the degree of similarity between two functions.

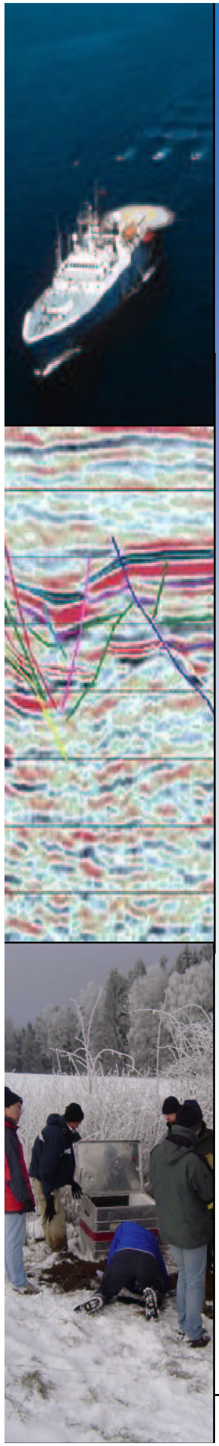
The correlation of functions g and f both with N samples is defined as:

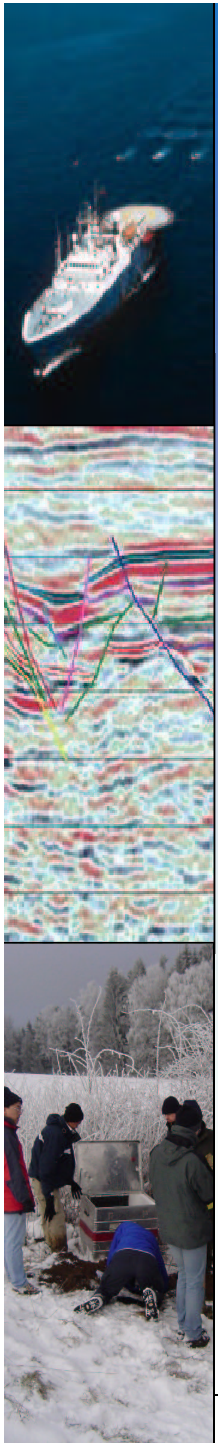
$$r_k = \frac{1}{N} \sum_{i=0}^{N-k-1} g_i \, f_{k+i}$$

$$k = 0,1,2,\ldots,N-1$$

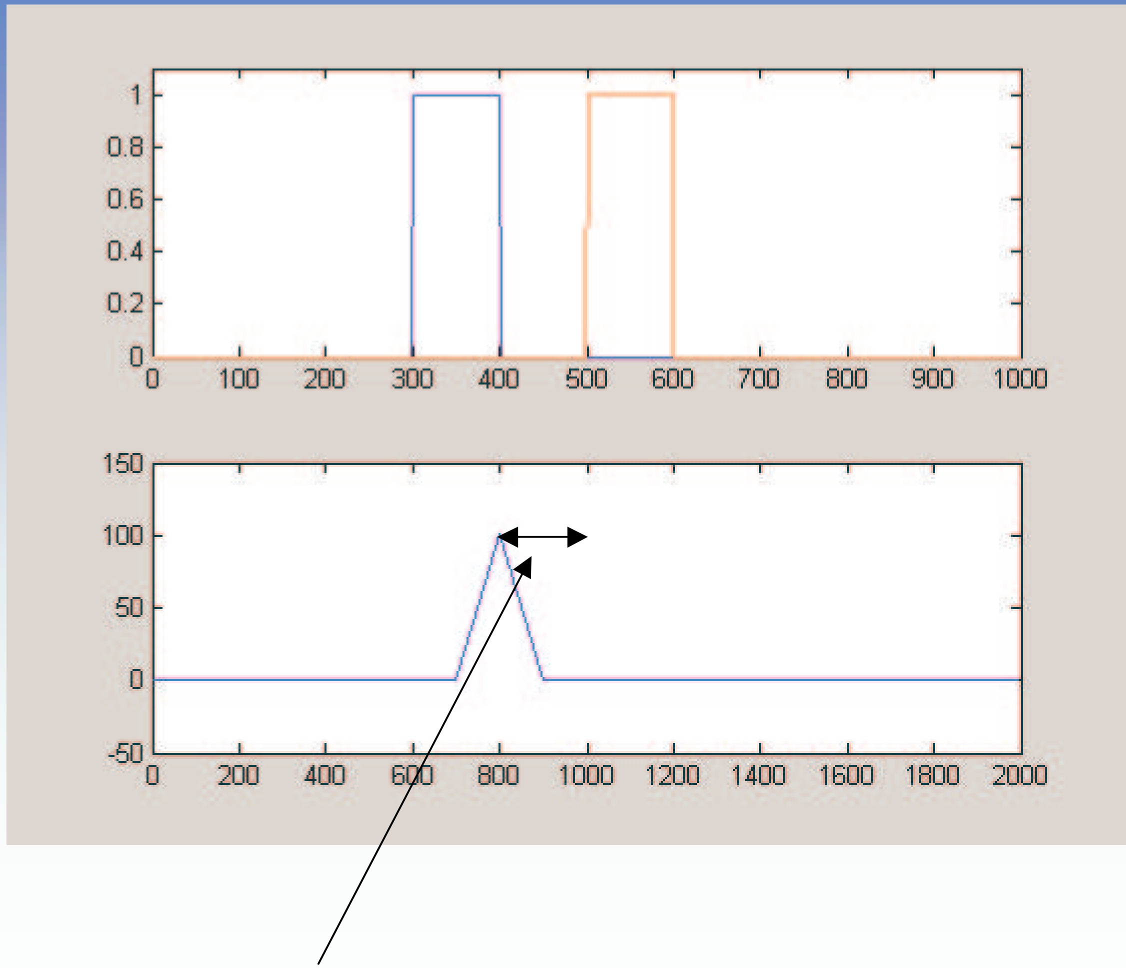
Auto-correlation



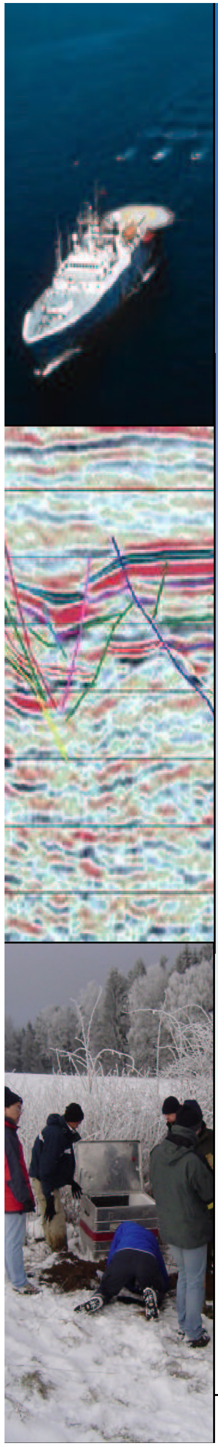




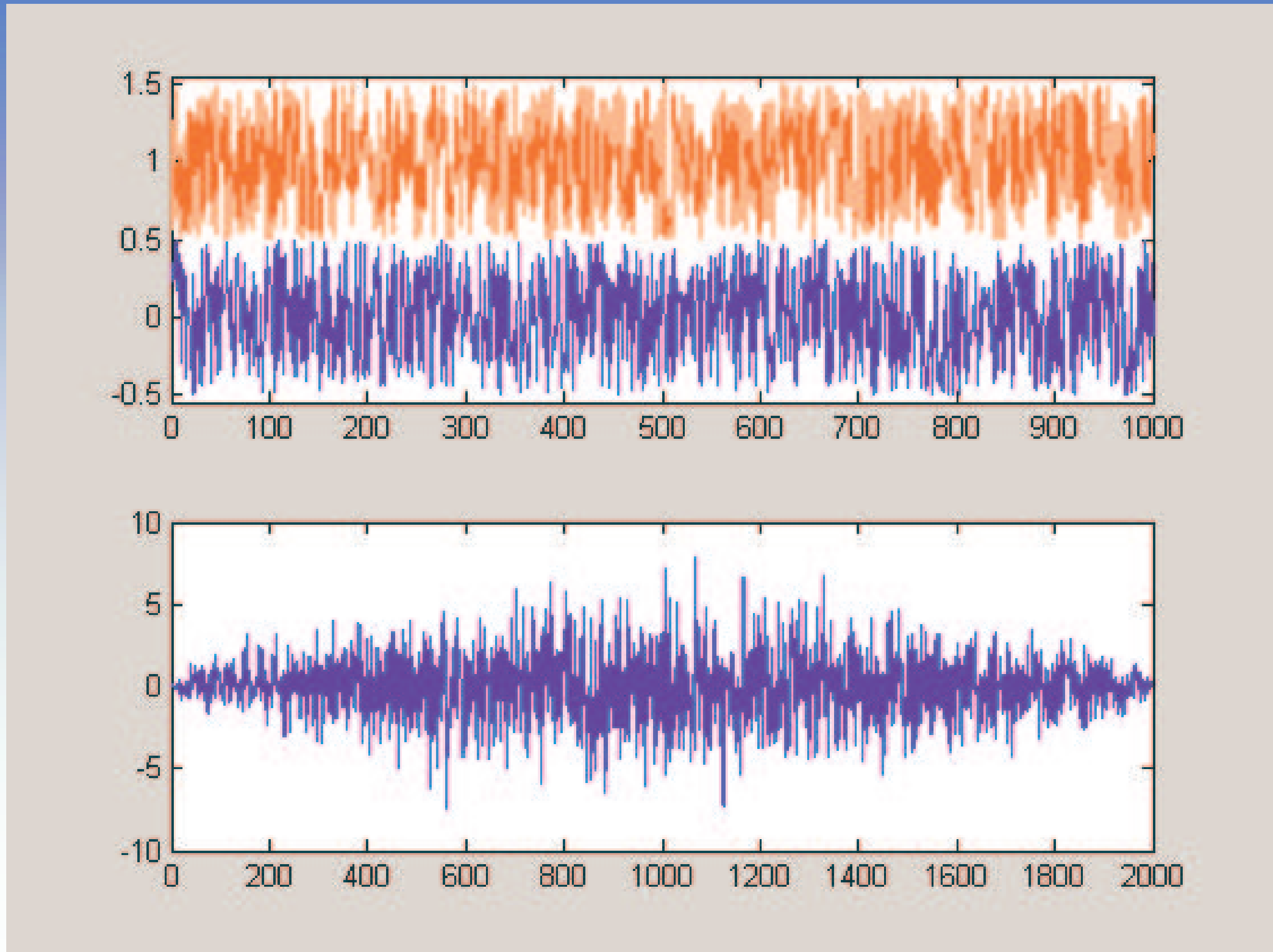
Cross-correlation



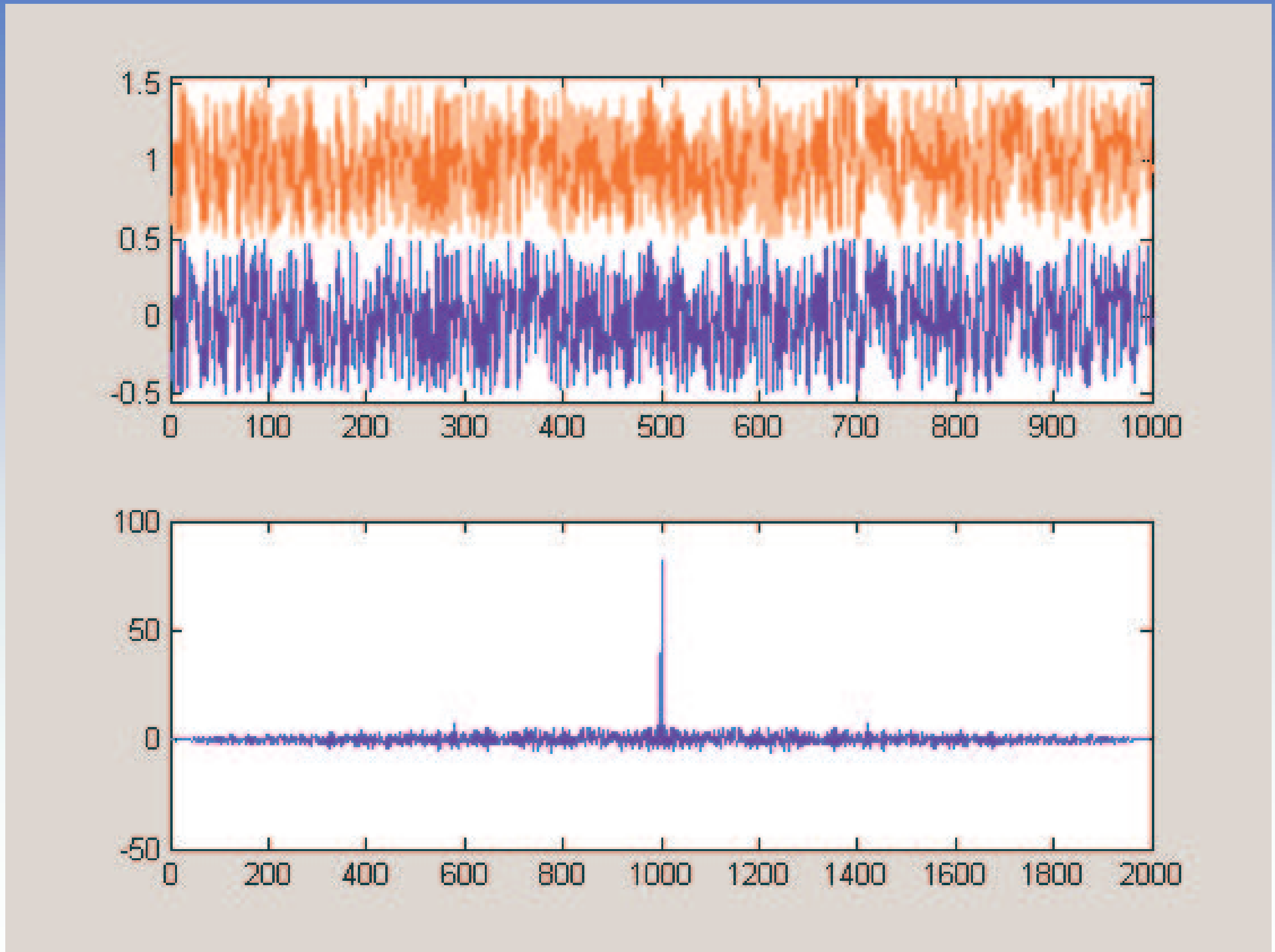
Lag between two functions

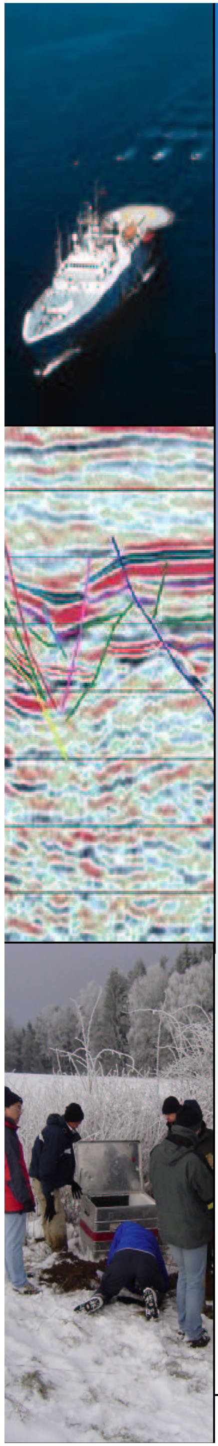


Cross-correlation
Random functions

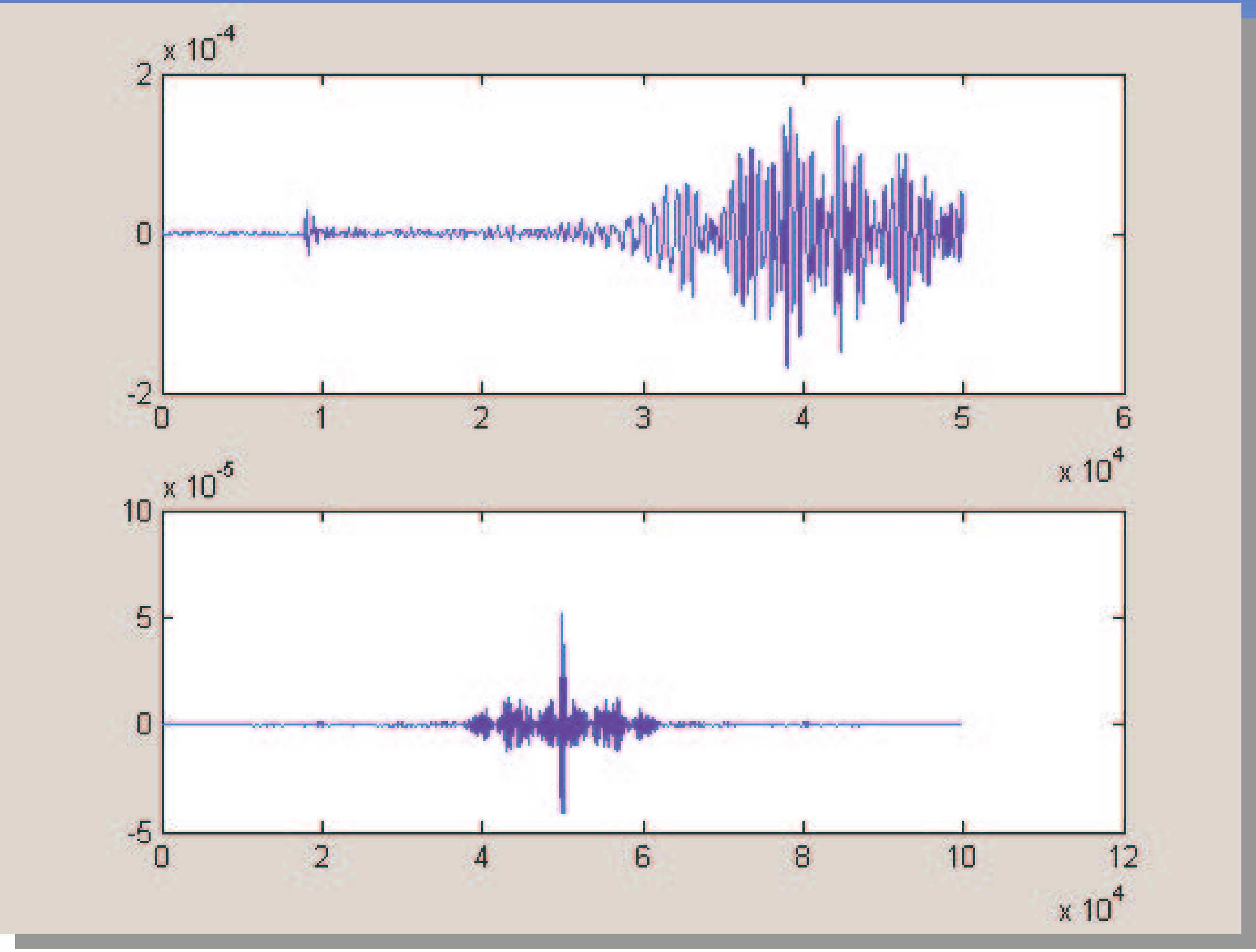


Auto-correlation
Random functions



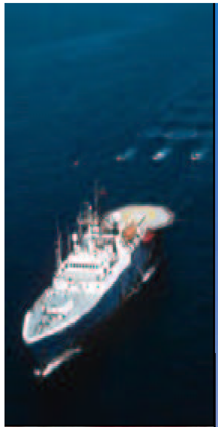


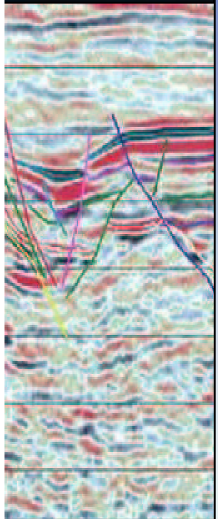
Auto-correlation
Seismic signal

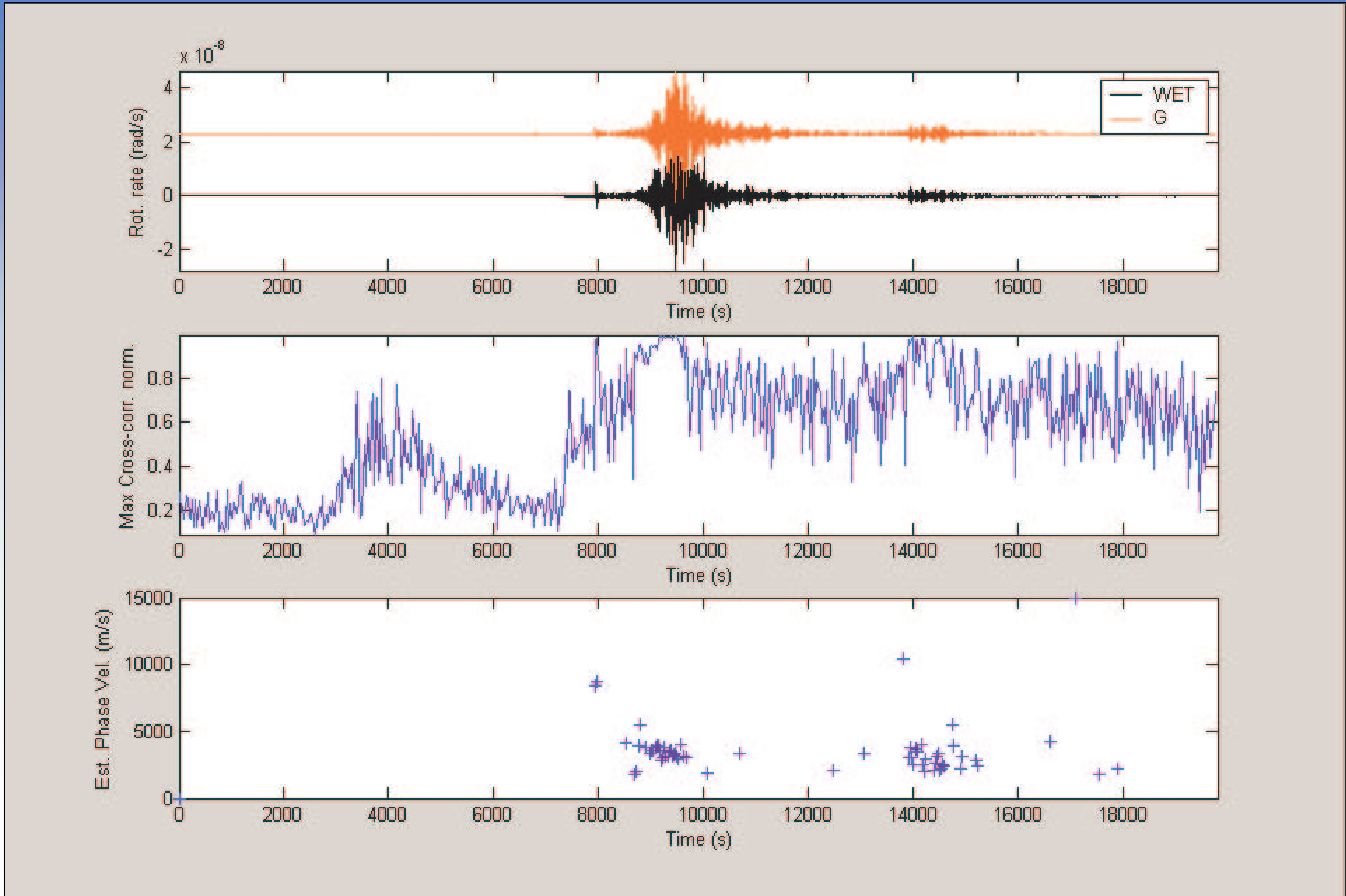


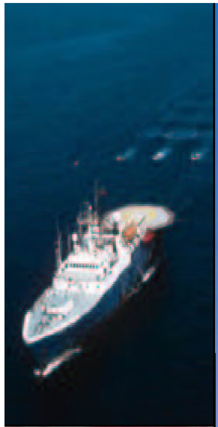
Cross-correlation

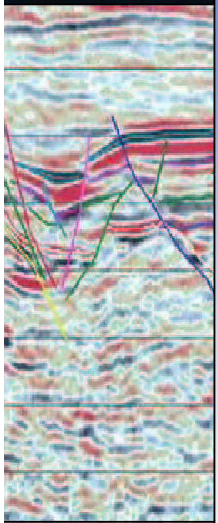
An example











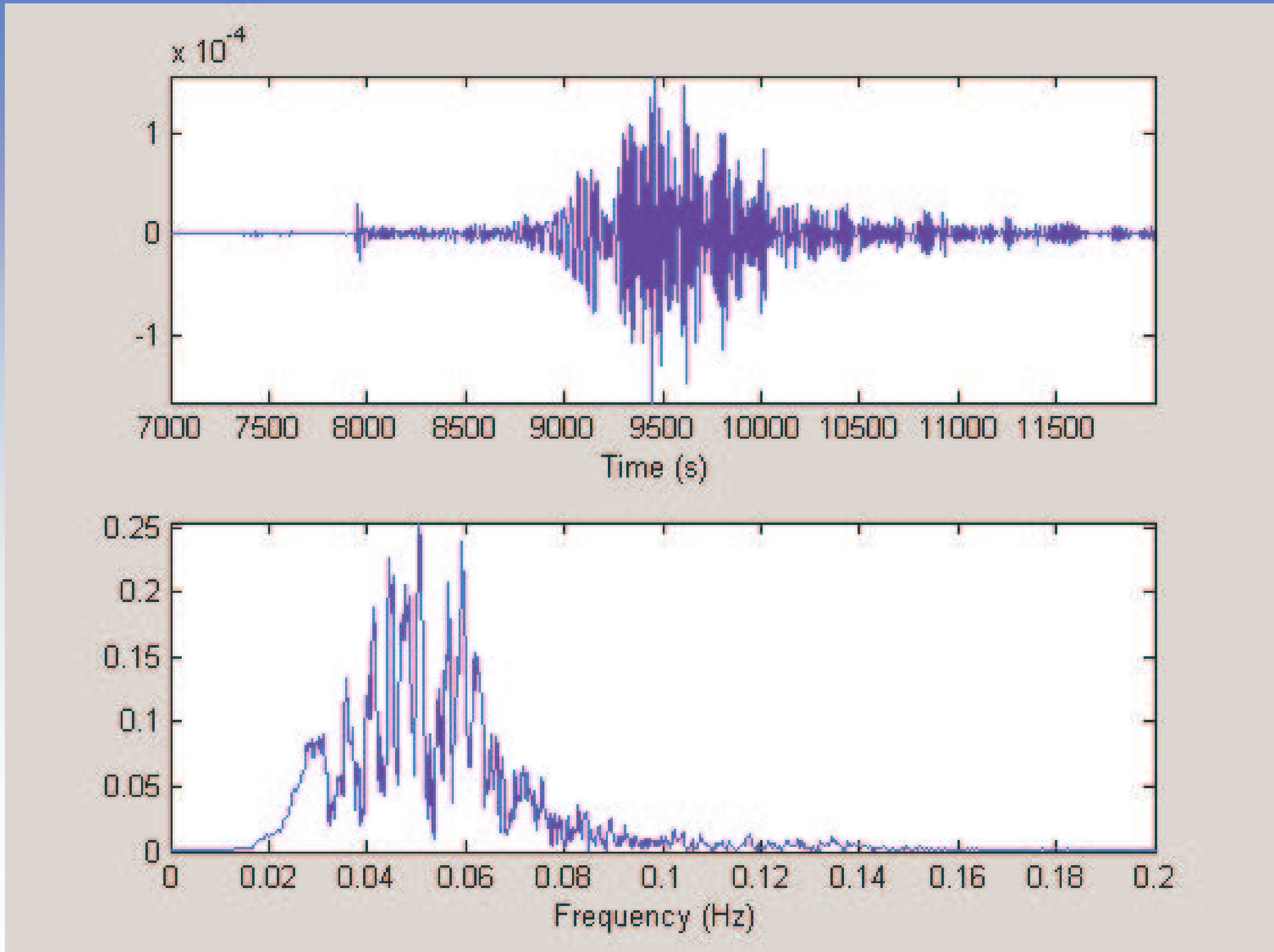
Digital Filtering

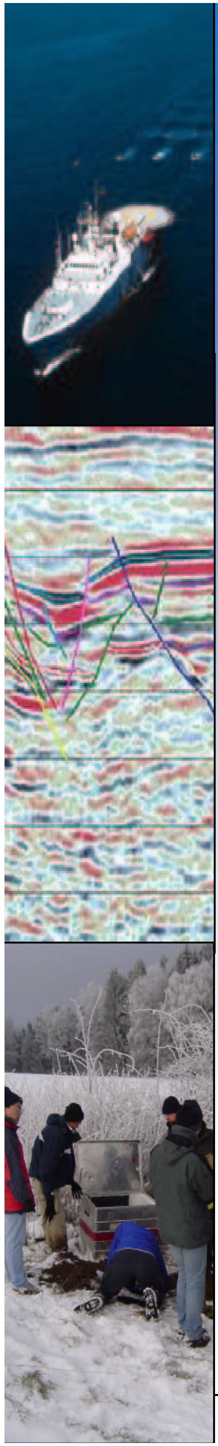
Often a recorded signal contains a lot of information that we are not interested in (noise). To get rid of this noise we can apply a filter in the frequency domain.

The most important filters are:

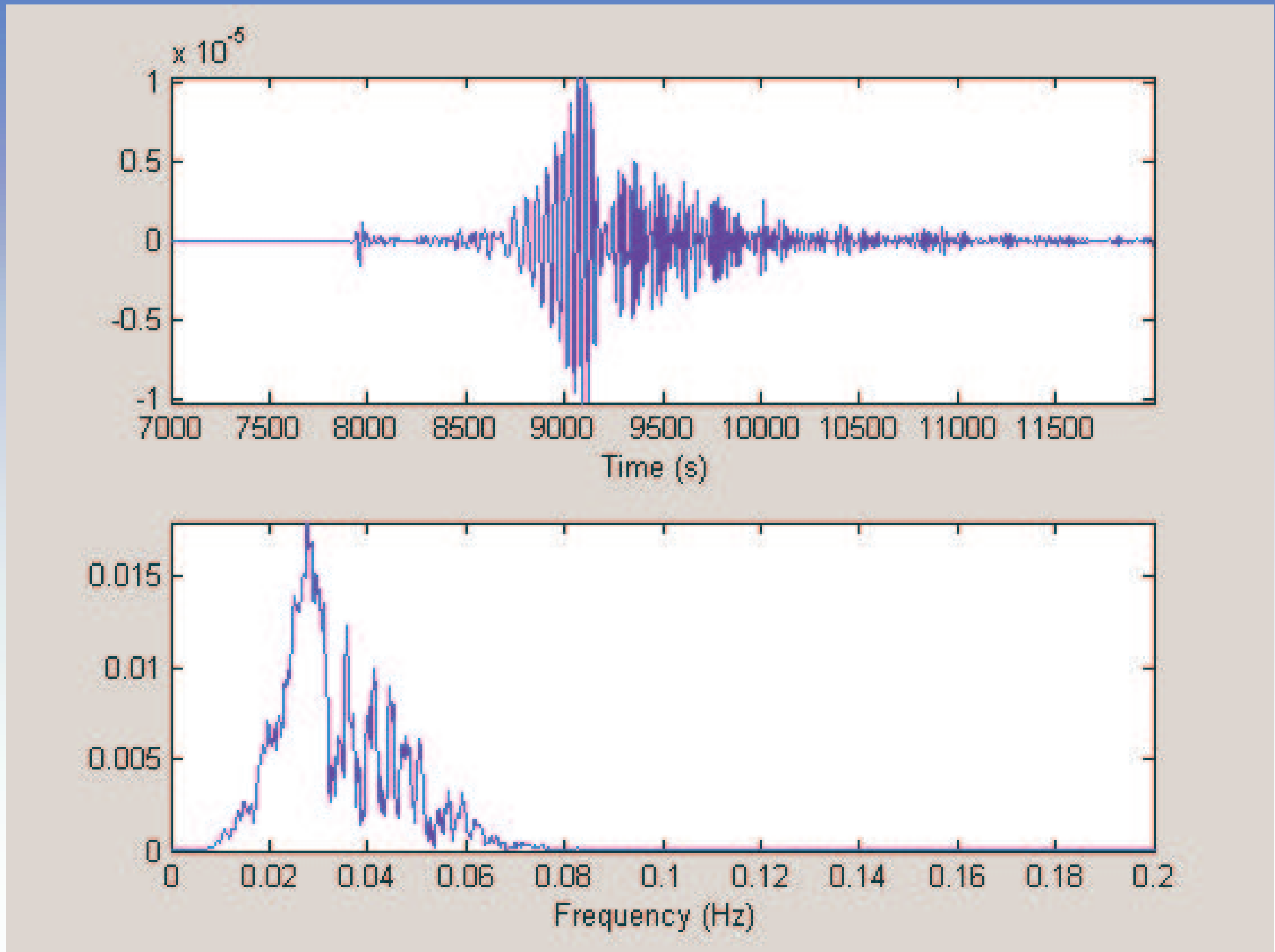
- High pass: cuts out low frequencies
- Low pass: cuts out high frequencies
- Band pass: cuts out both high and low frequencies and leaves a band of frequencies
- Band reject: cuts out certain frequency band and leaves all other frequencies

Digital Filtering

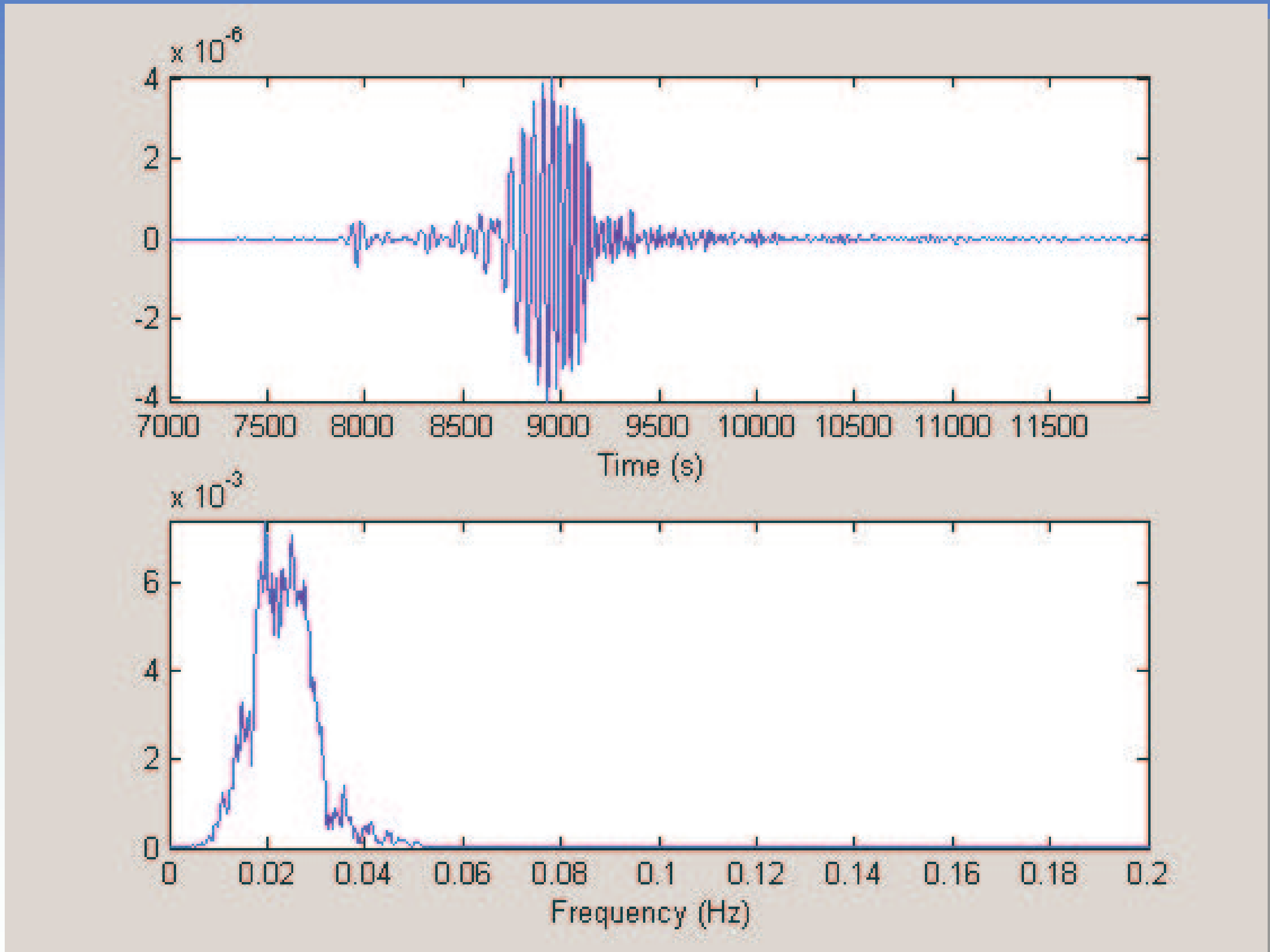


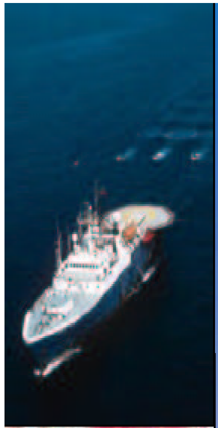


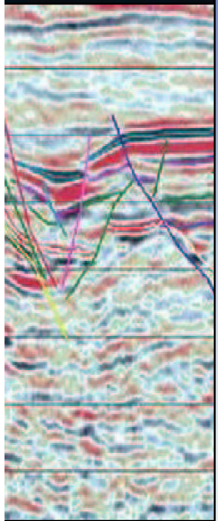
Low-pass filtering



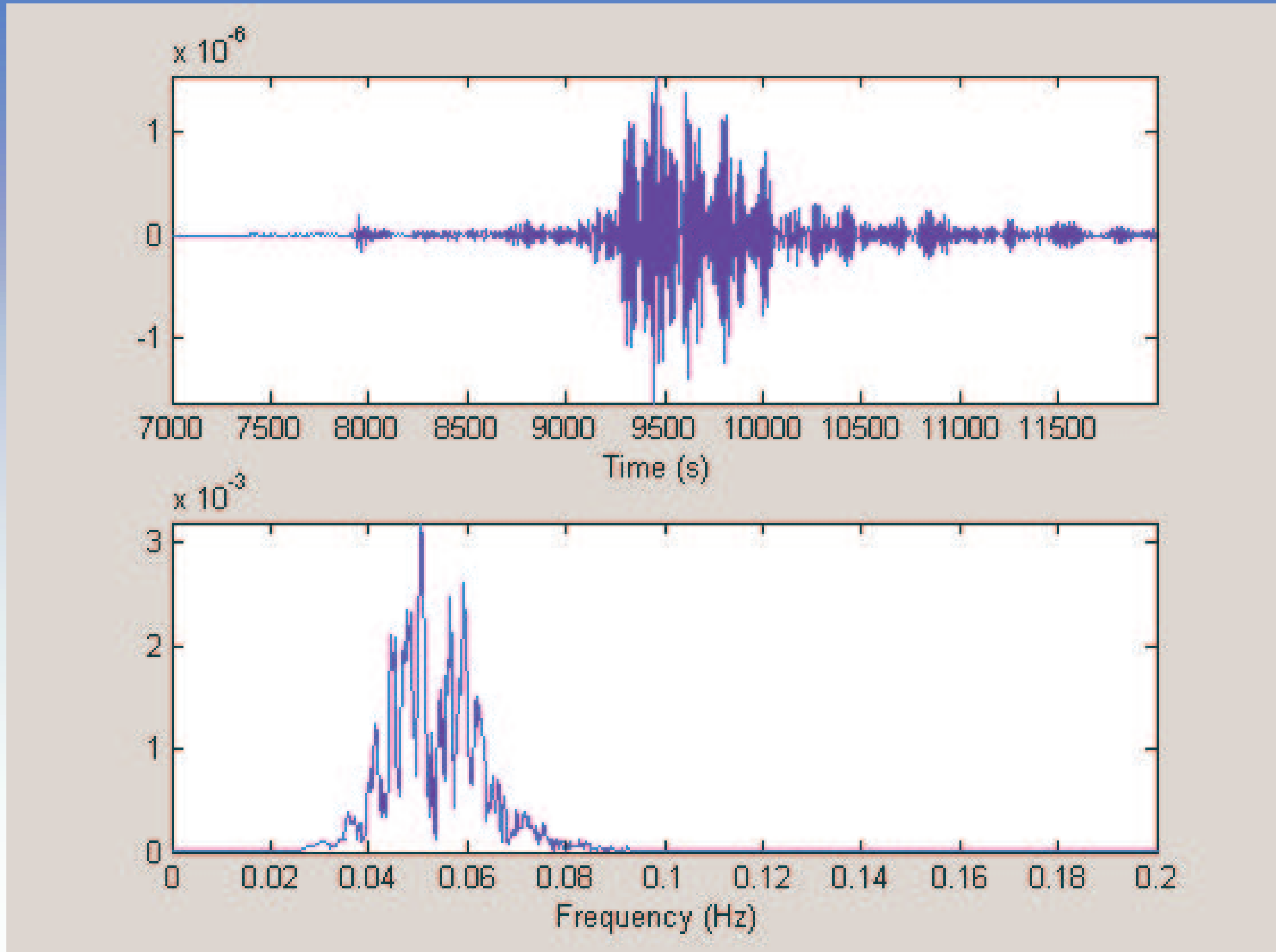
Lowpass filtering



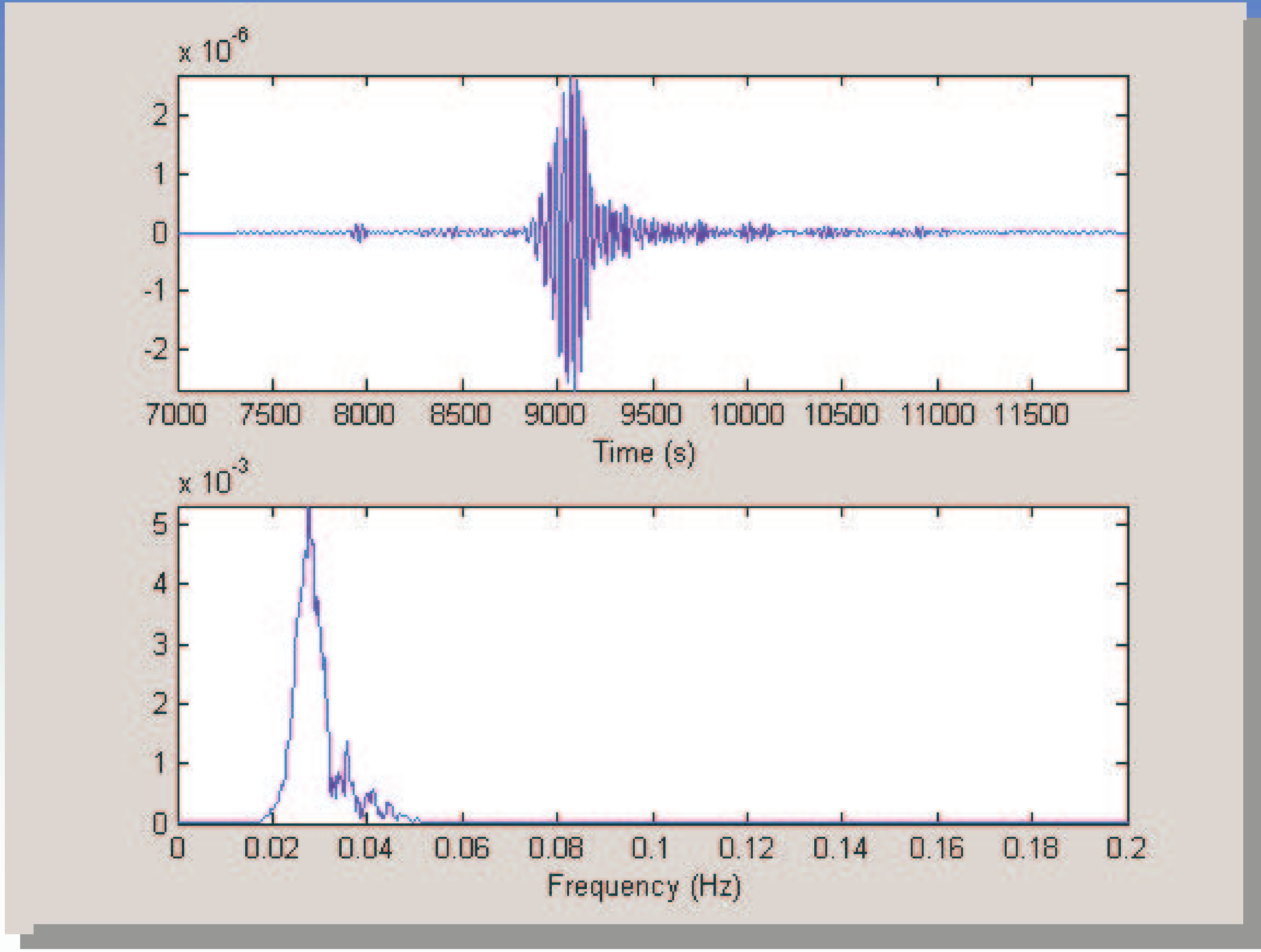


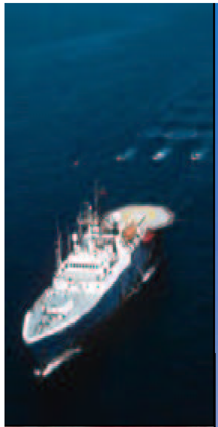


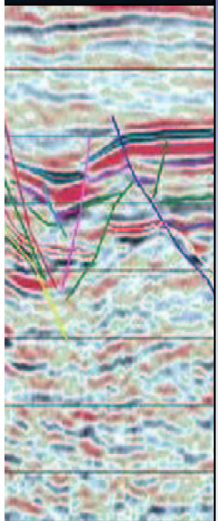
High-pass filter

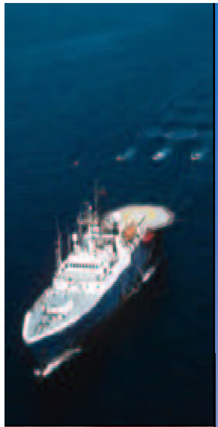


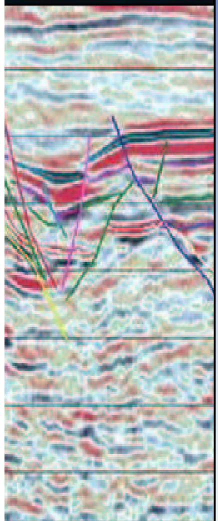
Band-pass filter











Summary

Today almost all data analysis involves the use of spectral methods and filtering.

The concepts are:

Convolution -> to obtain the reponse of a system to a particular input

Correlation -> to compare signal for their similarity

Fourier transformation – spectra - filtering -> to cut out certaqin frequencies and potentially highlight the signals of interest