Filtering Geophysical Data: Be careful!

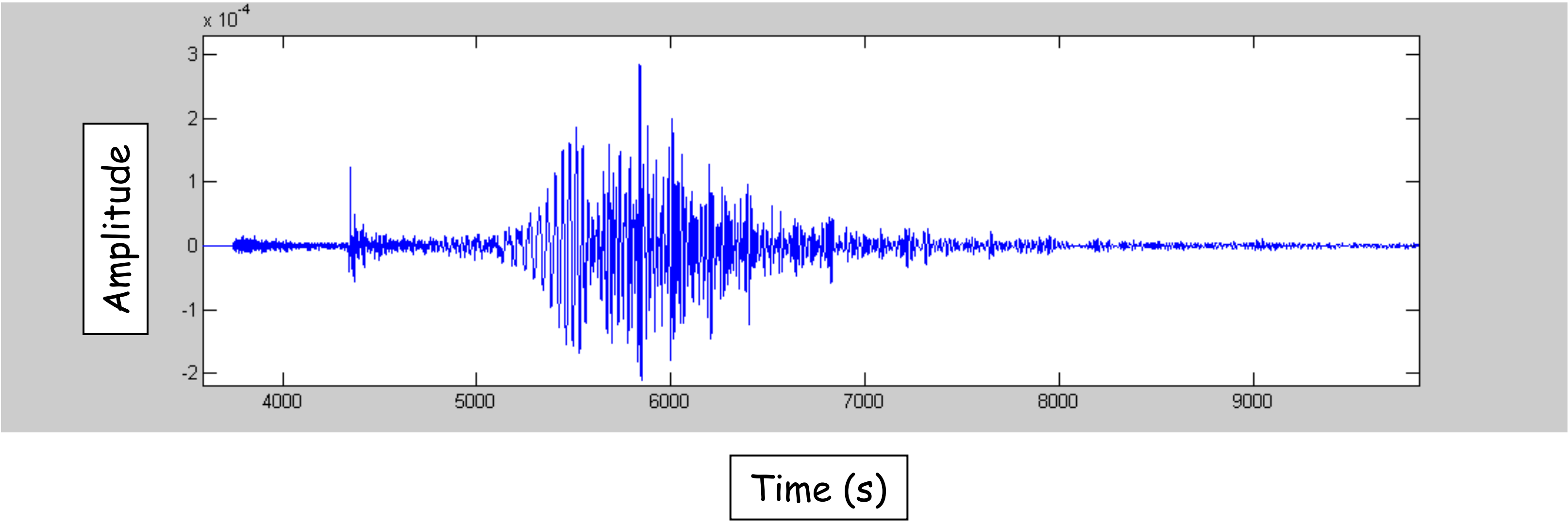
- Filtering: basic concepts
- Seismogram examples, high-low-bandpass filters
- The crux with causality
- Windowing seismic signals
  - Various window functions
  - Multitaper approach
  - Wavelets (principle)

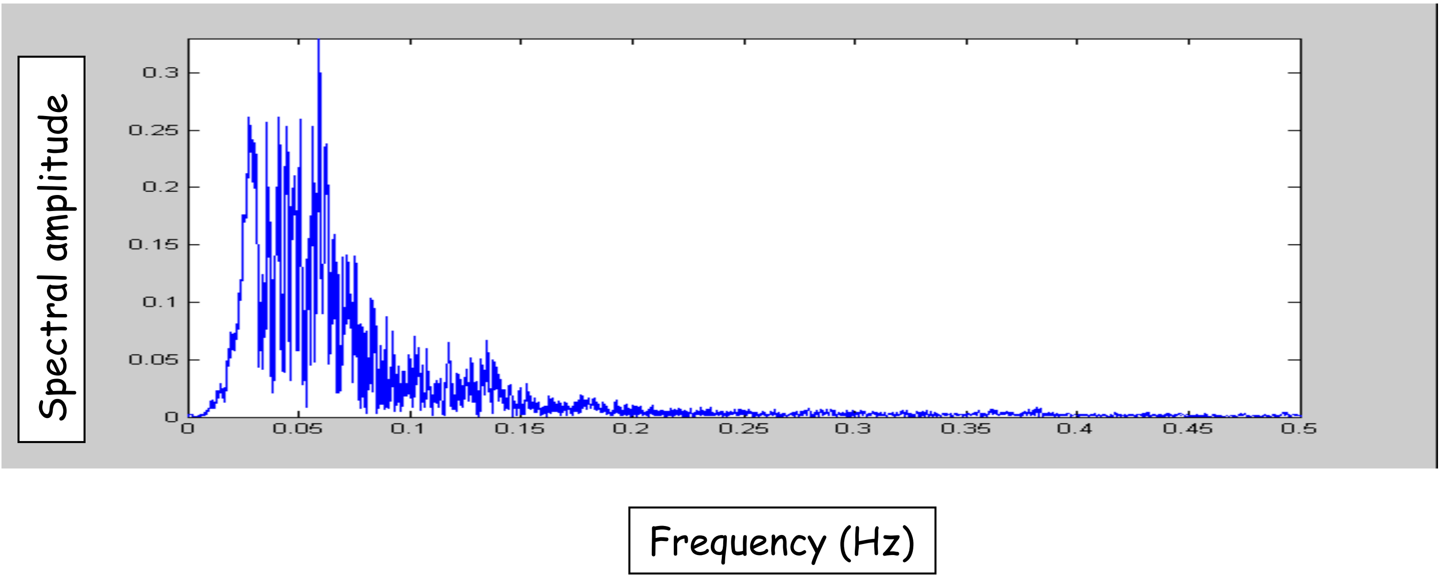
Scope: Understand the effects of filtering on time series (seismograms). Get to know frequently used windowing functions.

Why filtering

1. Get rid of unwanted frequencies
2. Highlight signals of certain frequencies
3. Identify harmonic signals in the data
4. Correcting for phase or amplitude characteristics of instruments
5. Prepare for down-sampling
6. Avoid aliasing effects

A seismogram





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

Cutoff frequency

the **cutoff or corner frequency** is the frequency either above which or below which the power output of a circuit, such as a line, amplifier, or filter, is reduced to $1/2$ of the passband power; the half-power point. This is equivalent to a voltage (or amplitude) reduction to 70.7% of the passband. This happens to be close to $-3dB$, and the cutoff frequency is frequently referred to as the $-3dB$ point. It is also referred to as the knee frequency, due to a frequency response curve's physical appearance.

Cut-off and slopes in spectra

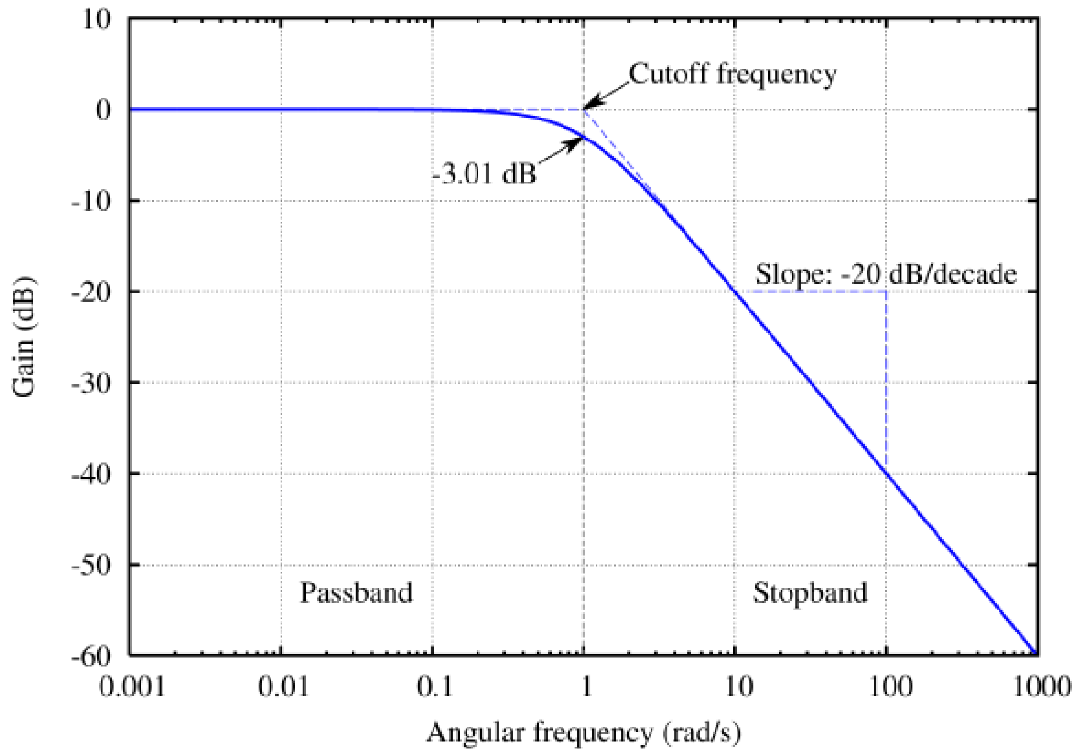
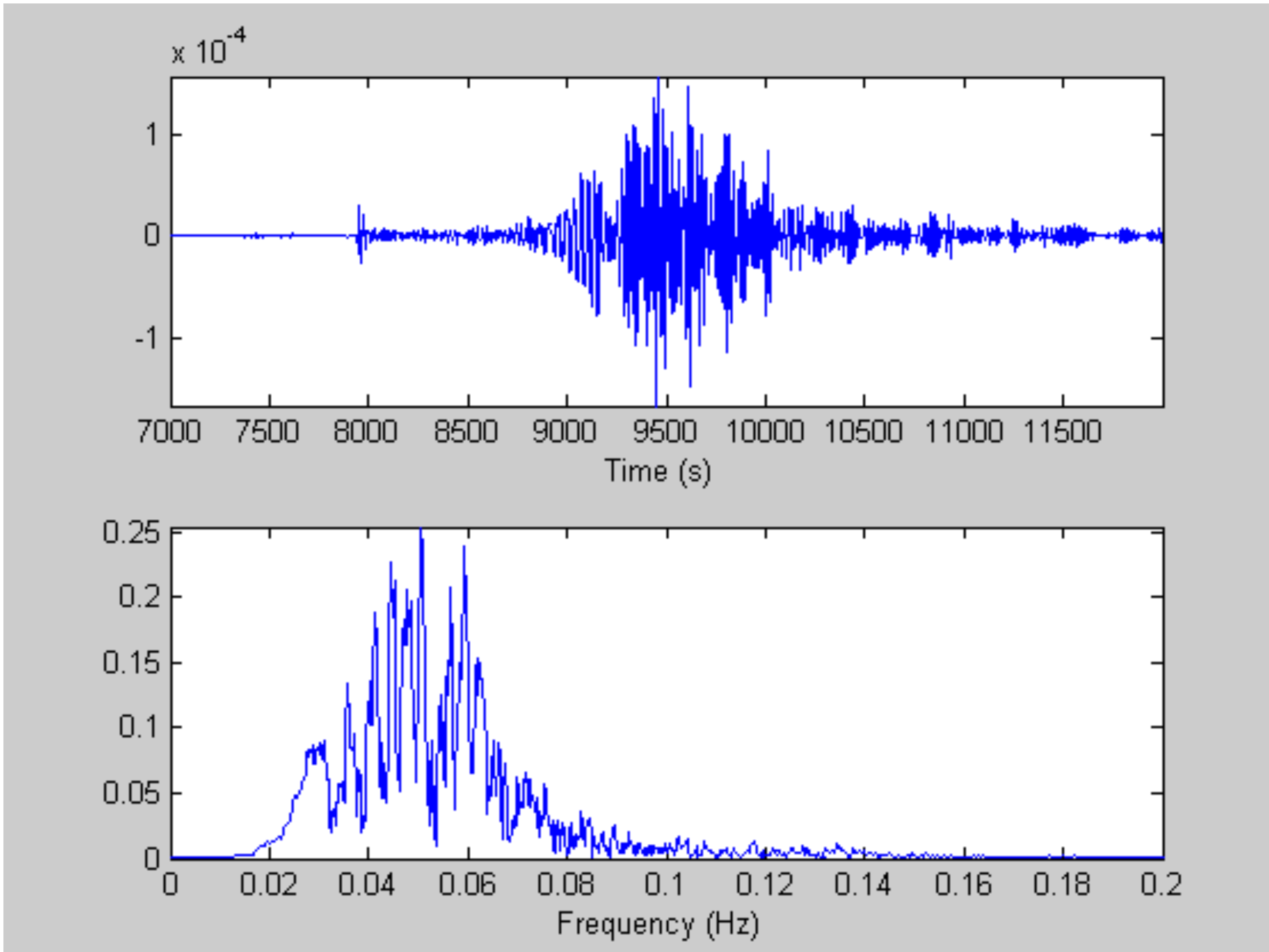
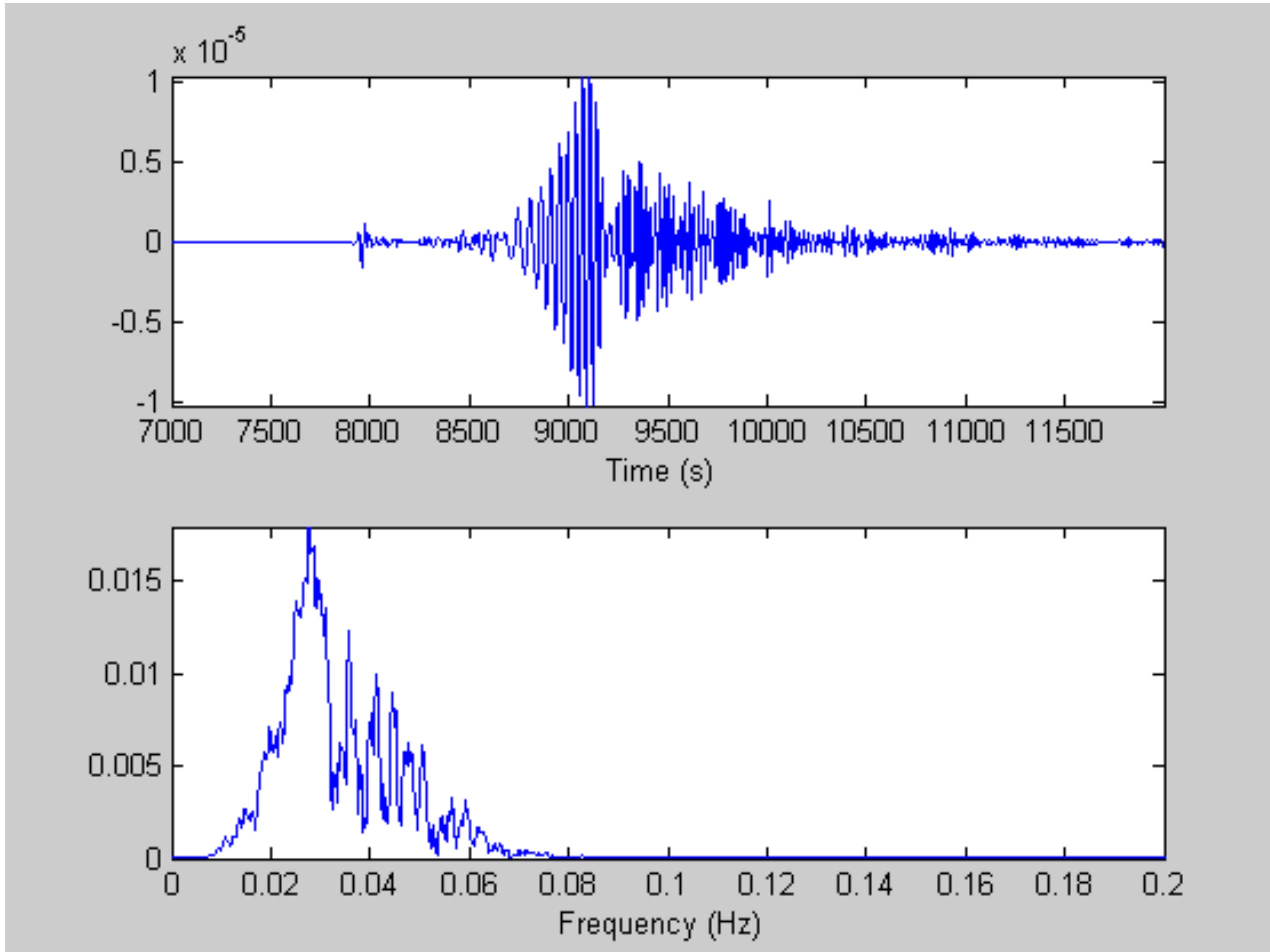


Figure 5.2: Amplitude response of a highpass Butterworth filter, showing the passband, stopband and the cutoff frequency. (The slope $-20dB$ per decade is the same as $-6dB$ per octave, equivalent to a slope of $-s^1$ [as one order of magnitude in amplitude is equal to $20db$].)

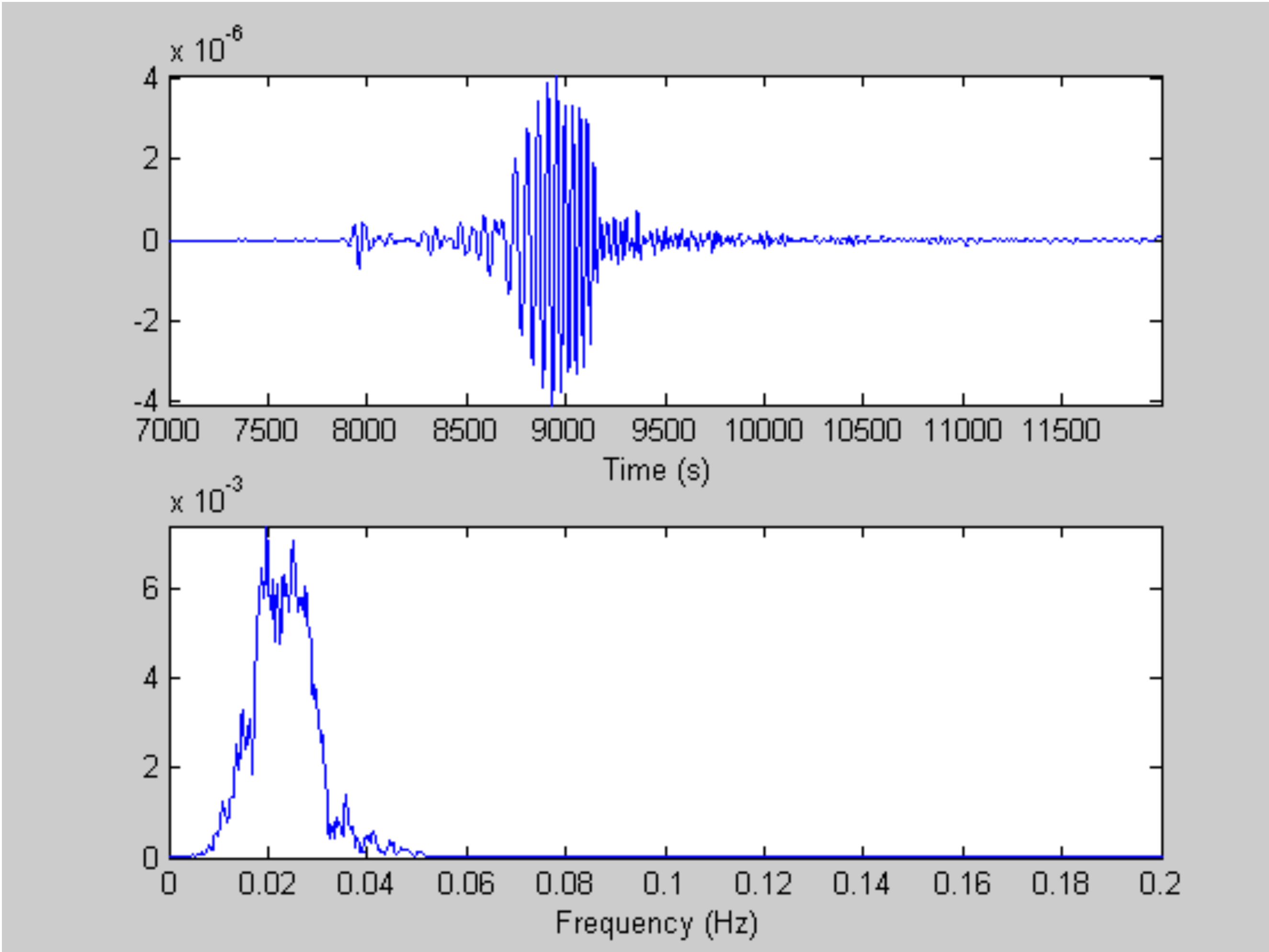
Digital Filtering



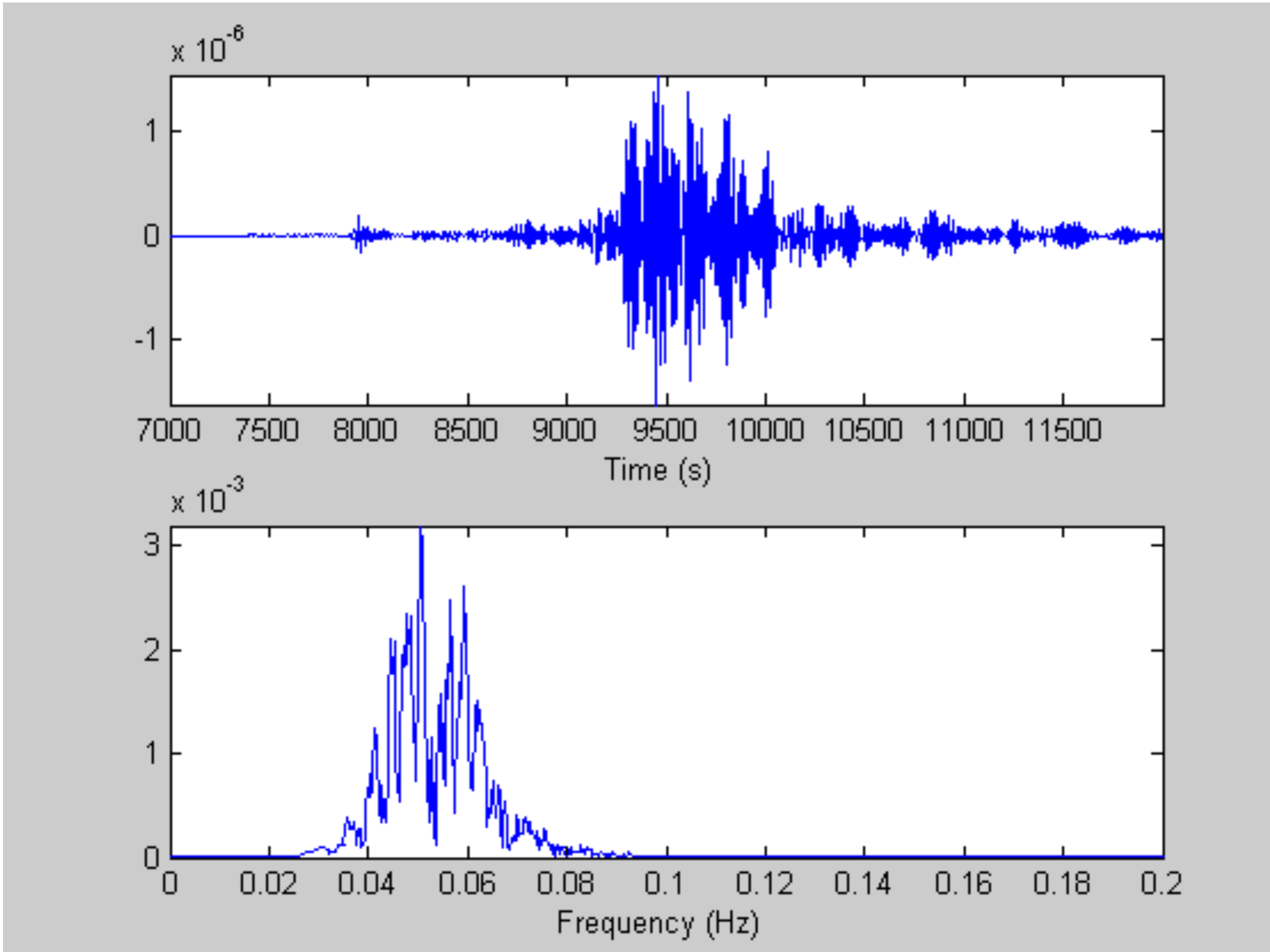
Low-pass filtering



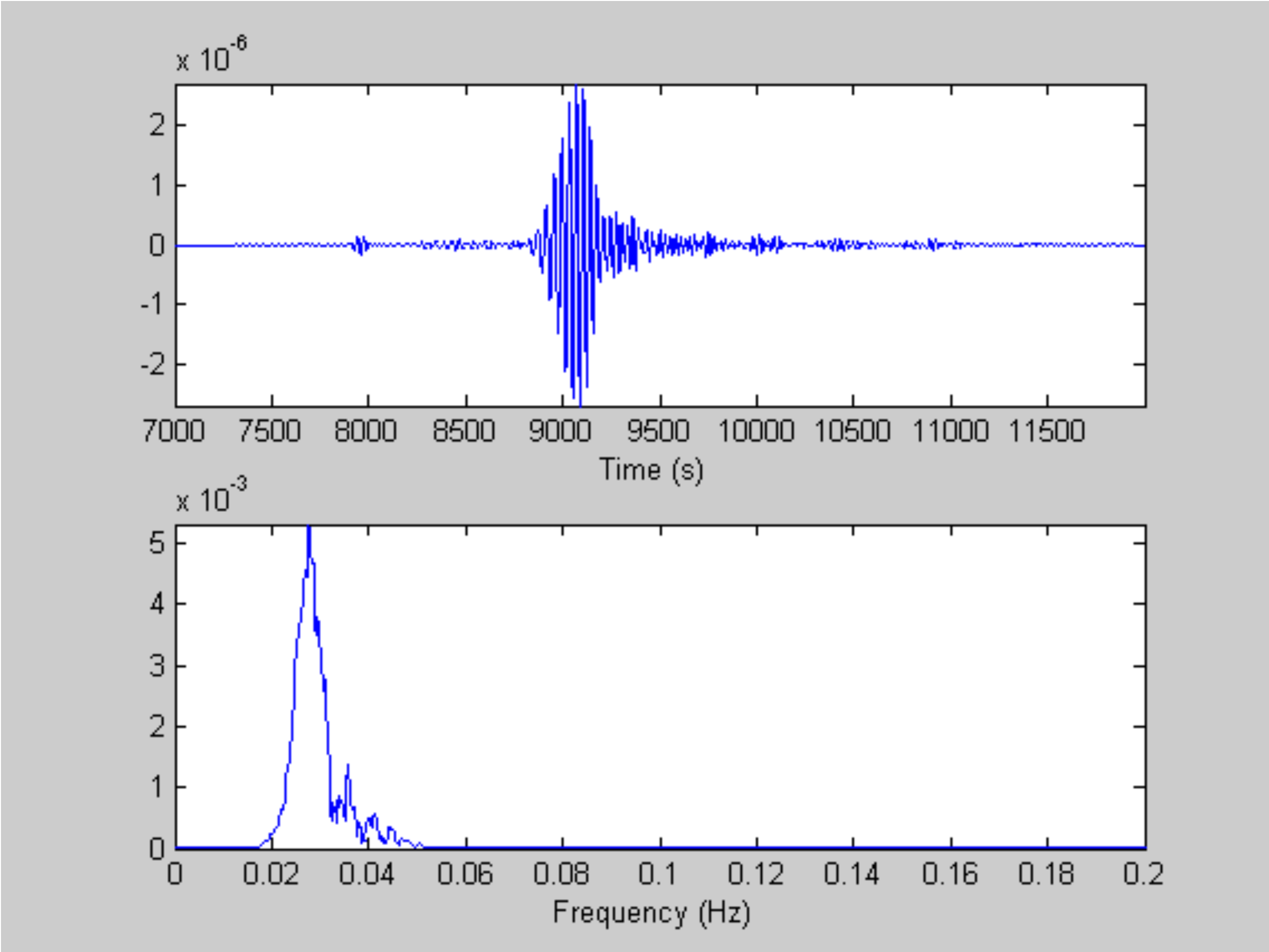
Lowpass filtering



High-pass filter

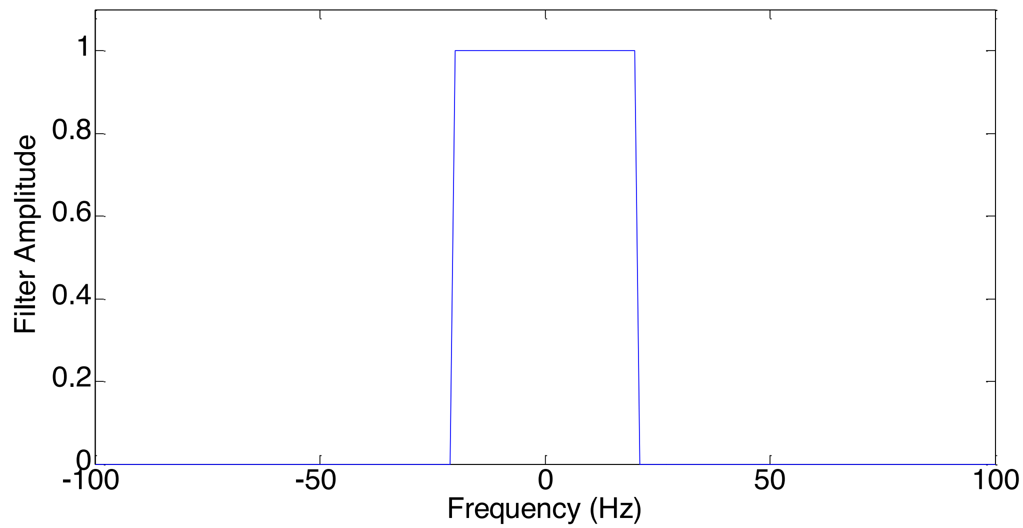


Band-pass filter



The simplemost filter

The simplemost filter gets rid of all frequencies above a certain cut-off frequency (low-pass), „box-car“

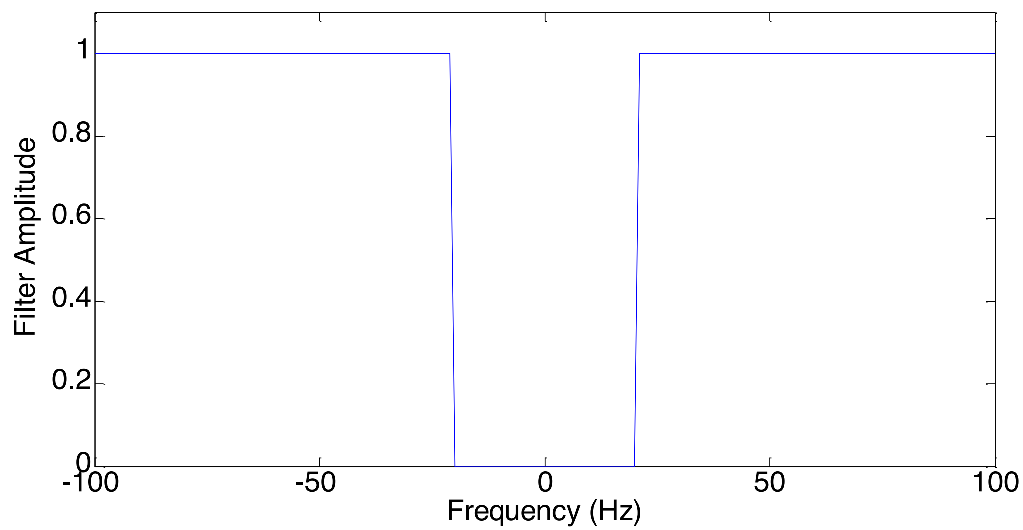


$$H_L(\omega) = \begin{cases} 1 & if \quad |\omega| \le \omega \\ 0 & if \quad |\omega| > \omega_0 \end{cases}$$

The simplemost filter

… and its brother

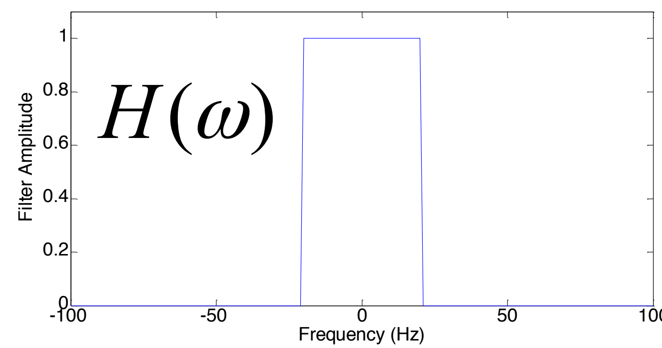
… (high-pass)



$$H_H(\omega) = 1 - H_L(\omega) = \begin{cases} 1 & if \quad |\omega| > \omega \\ 0 & if \quad |\omega| \leq \omega_0 \end{cases}$$

… let ‘s look at the consequencse

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

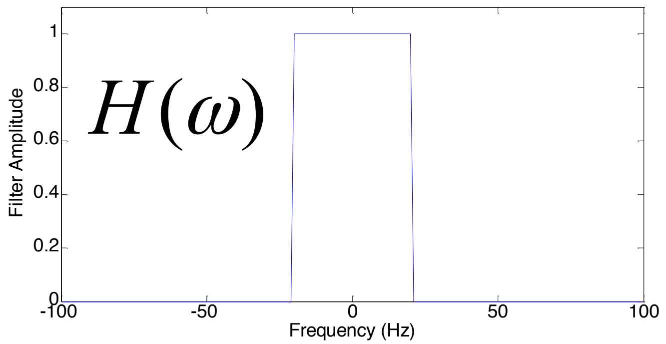


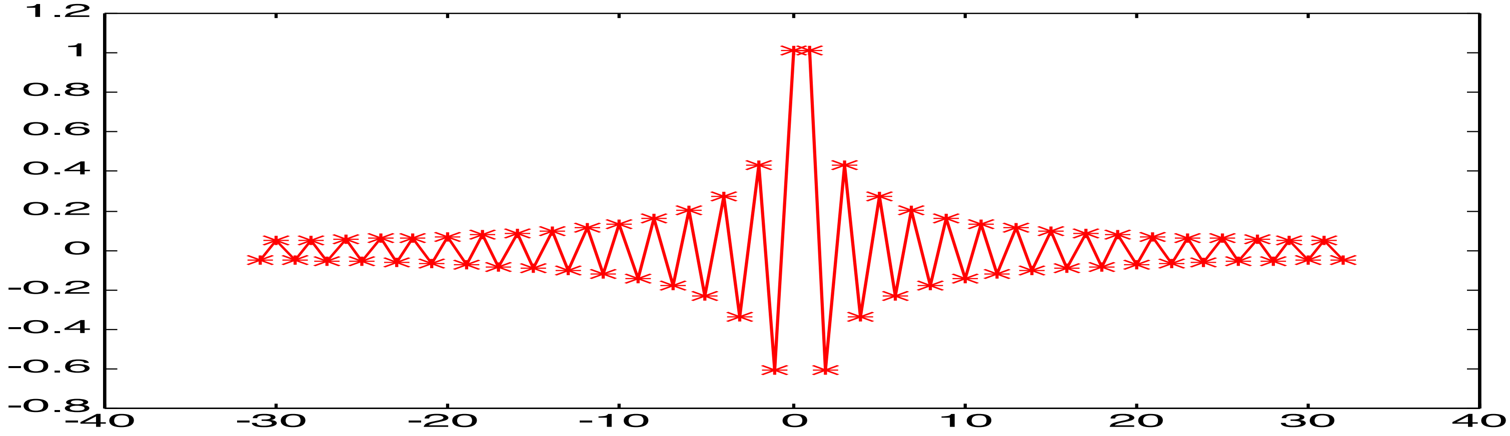
$$f^{filt}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H(\omega) F(\omega) e^{-ik\omega} dt$$

… but what does H(ω) look like in the time domain … remember the convolution theorem?

… surprise …

$$h(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H(\omega) e^{-ik\omega} dt$$





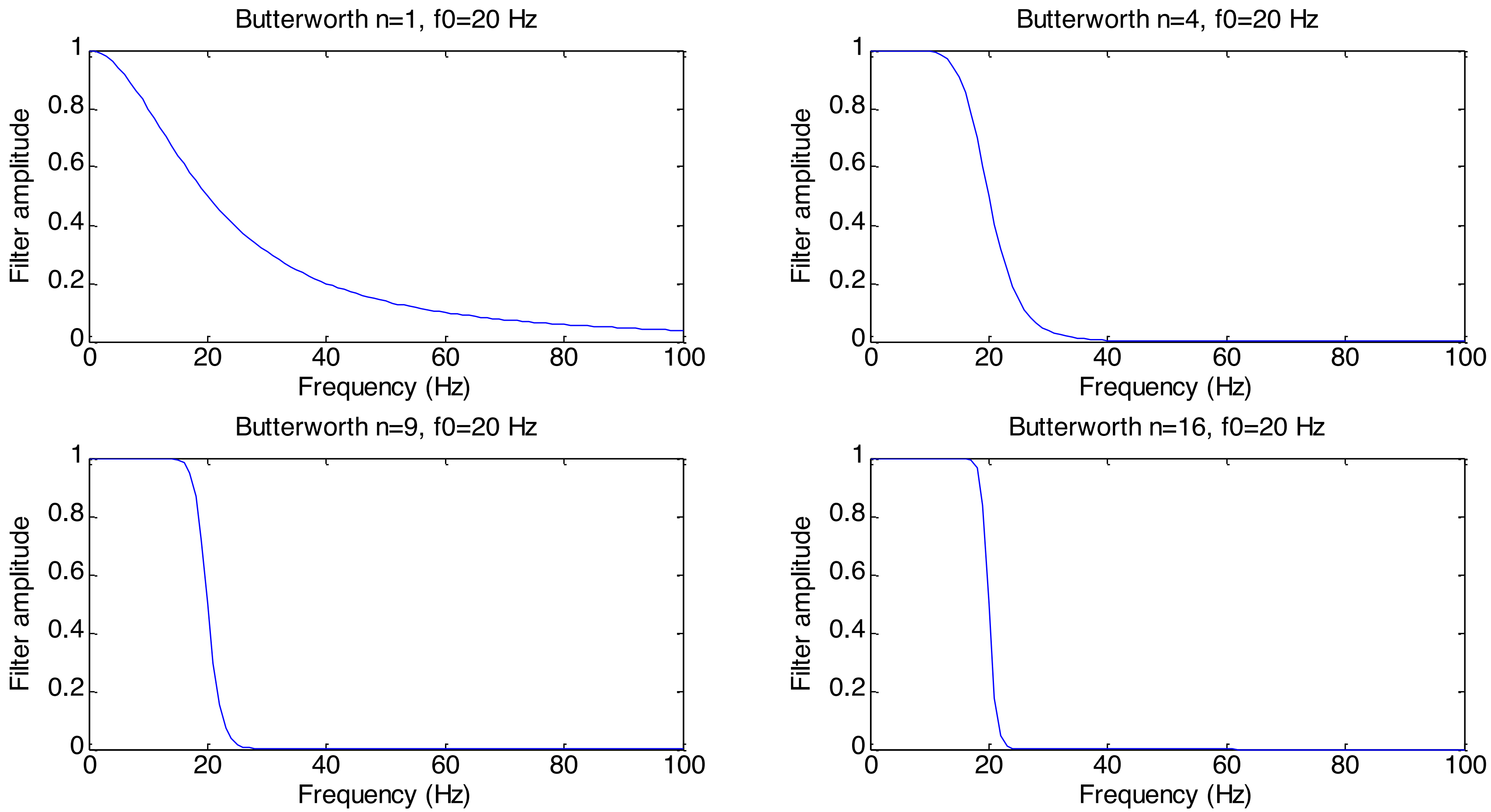
Zero phase and causal filters

Zero phase filters can be realised by

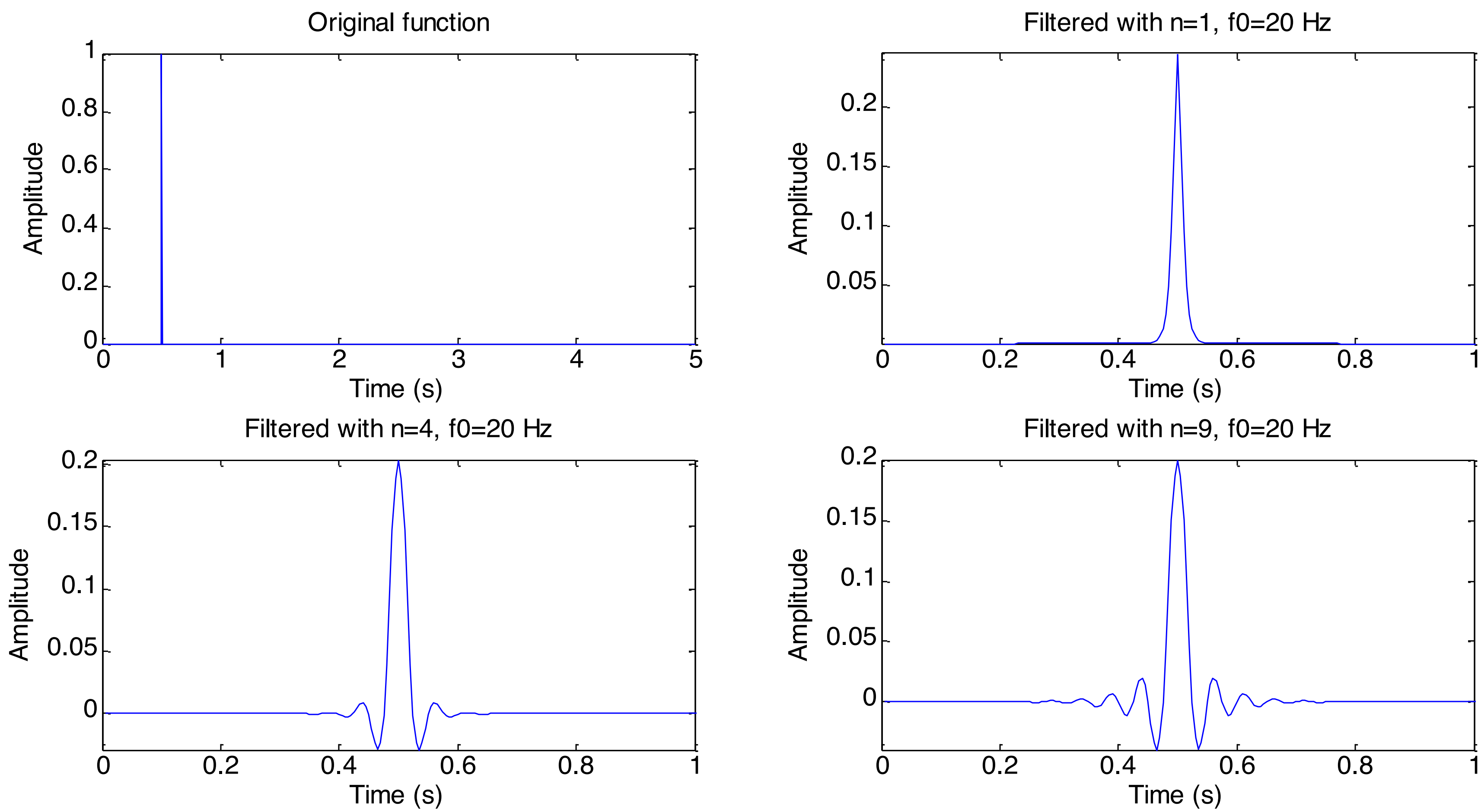
- Convolve first with a chosen filter
- Time reverse the original filter and convolve again
- First operation multiplies by $F(\omega)$, the 2nd operation is a multiplication by $F^*(\omega)$
- The net multiplication is thus $| F(w)|^2$
- These are also called two-pass filters

The Butterworth Filter (Low-pass, 0-phase)

$$\left|F_L(\omega)\right| = \frac{1}{1+(\omega/\omega_c)^{2n}}$$

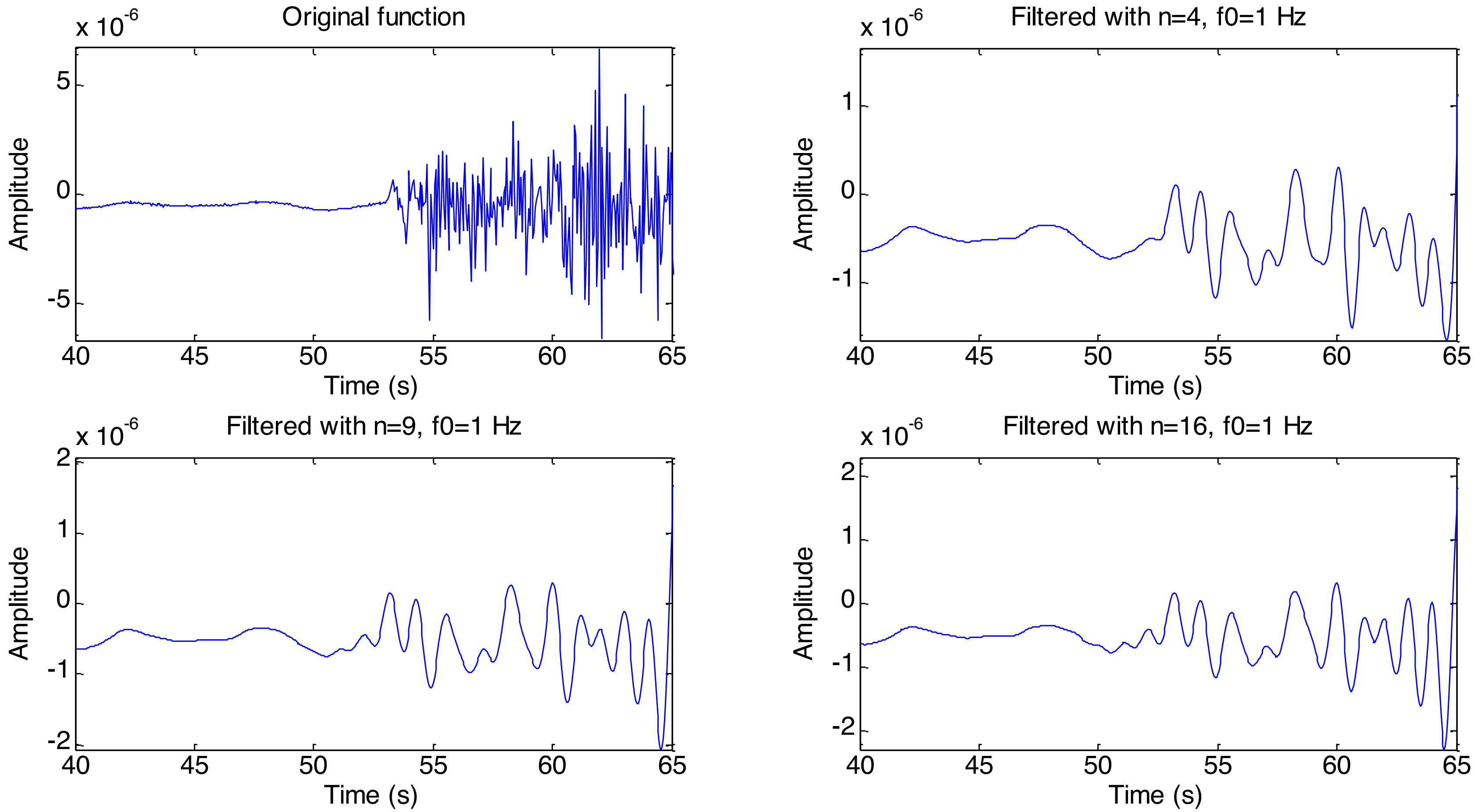


… effect on a spike …



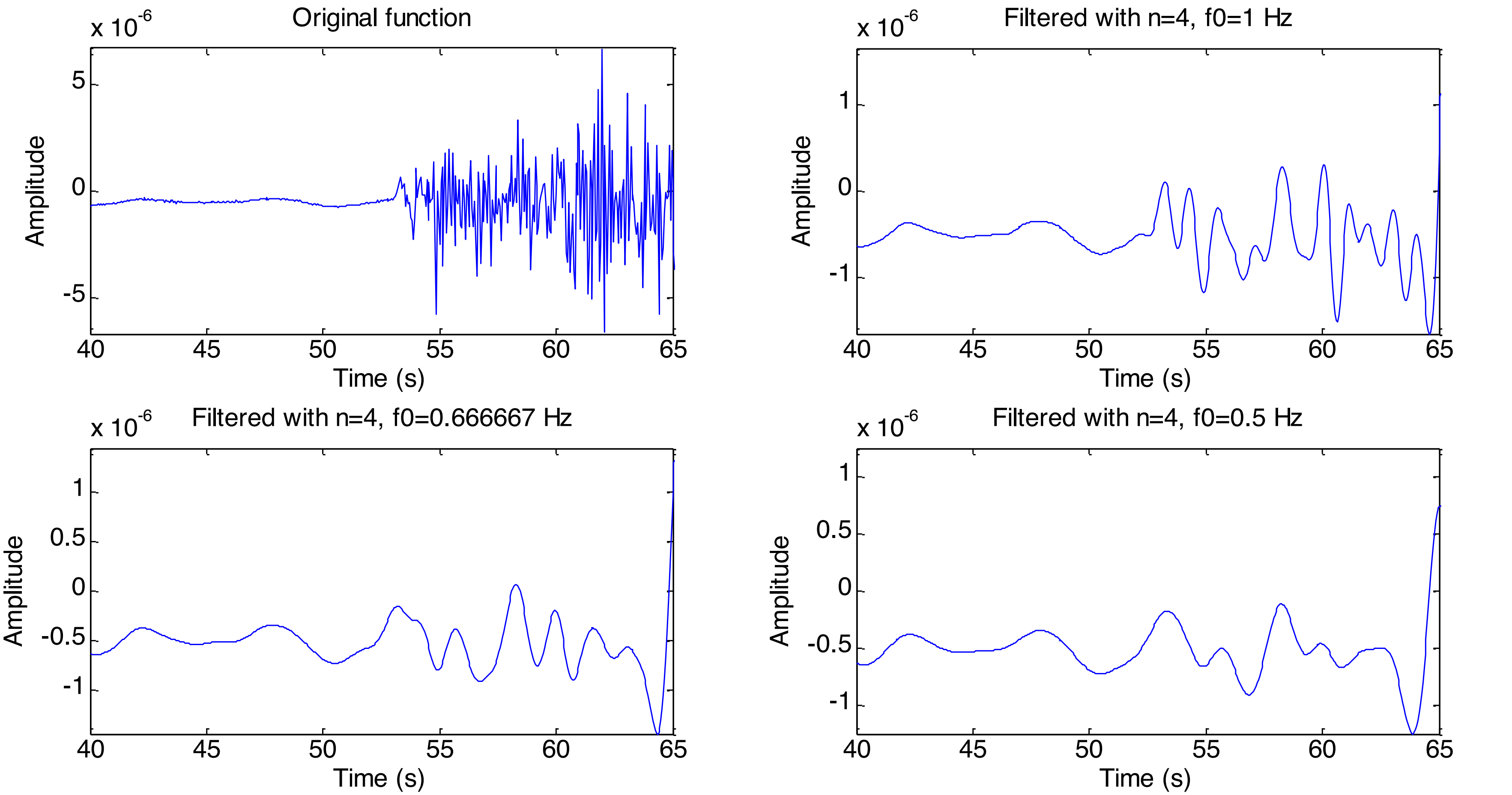
… on a seismogram …

… varying the order …



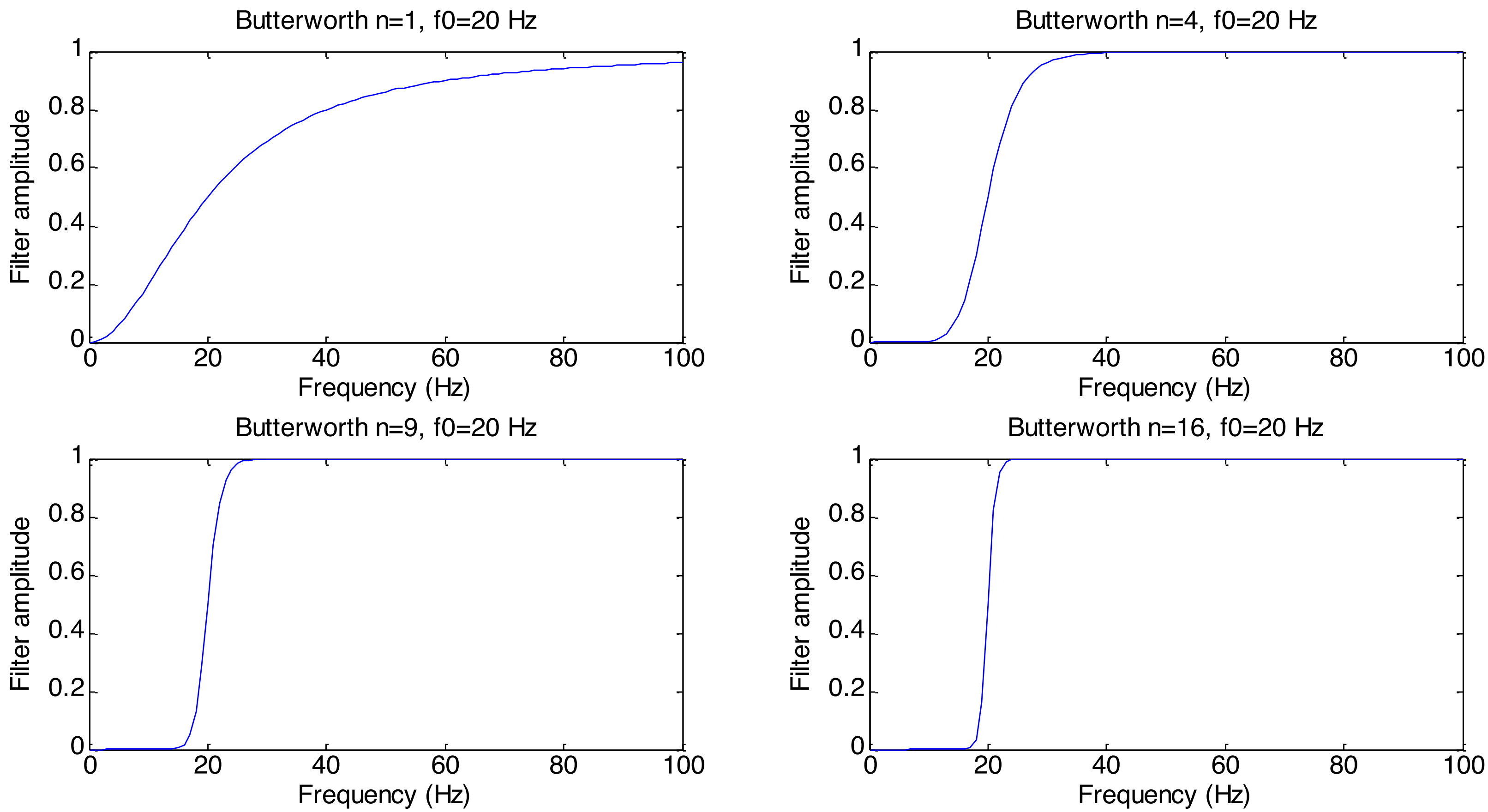
… on a seismogram …

… varying the cut-off frequency…

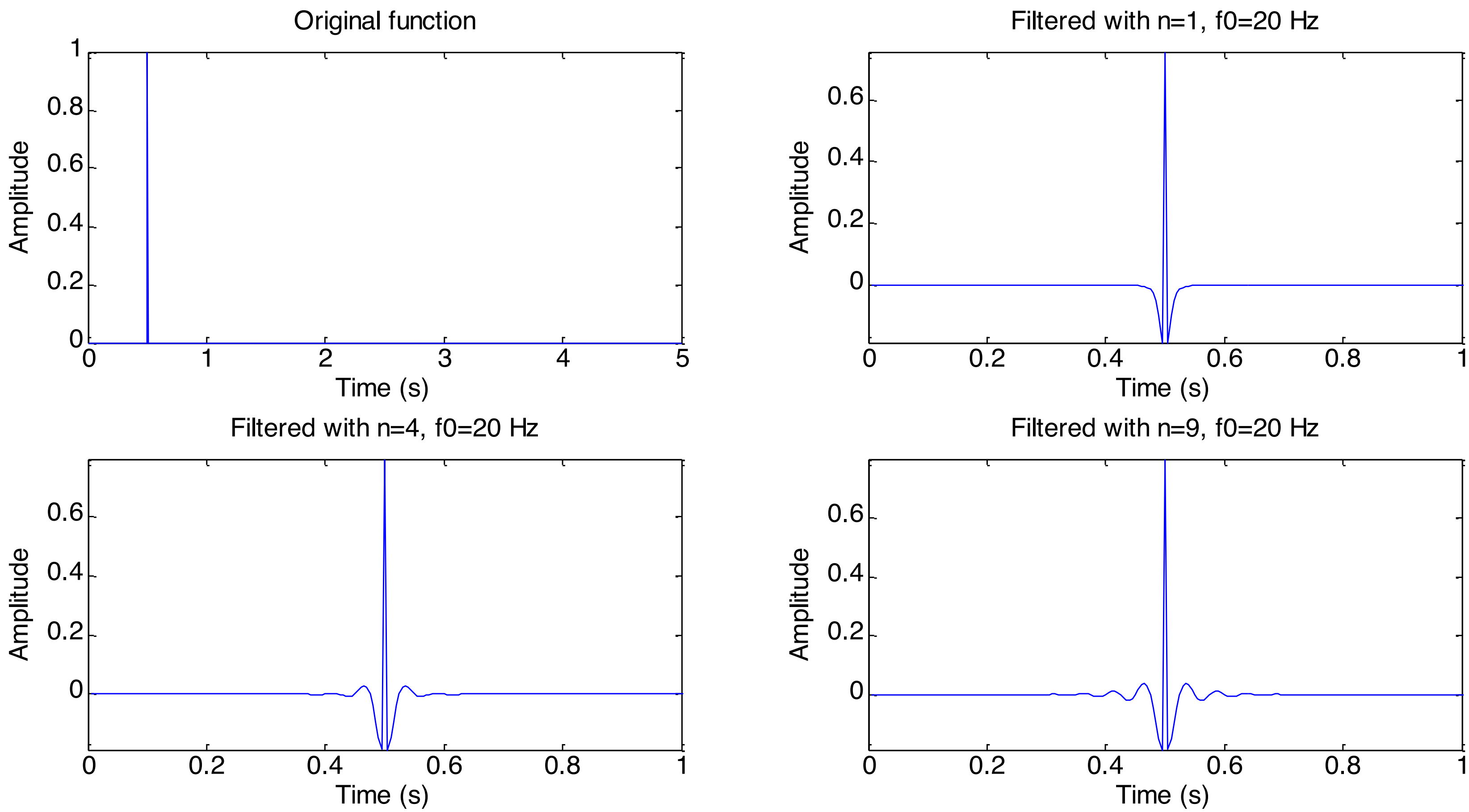


The Butterworth Filter (High-Pass)

$$|F_H(\omega)| = 1 - \frac{1}{1+(\omega/\omega_c)^{2n}}$$

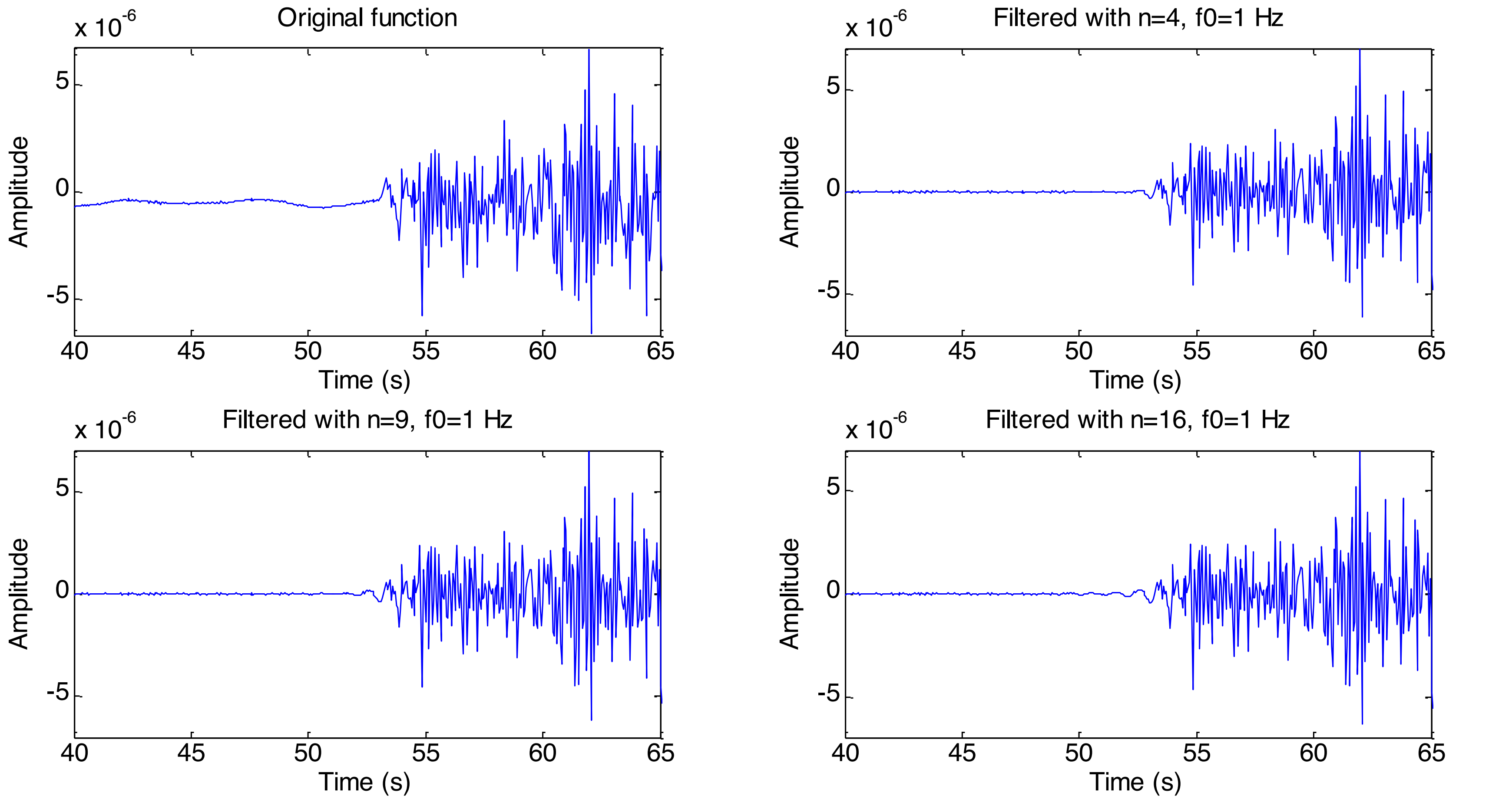


… effect on a spike …



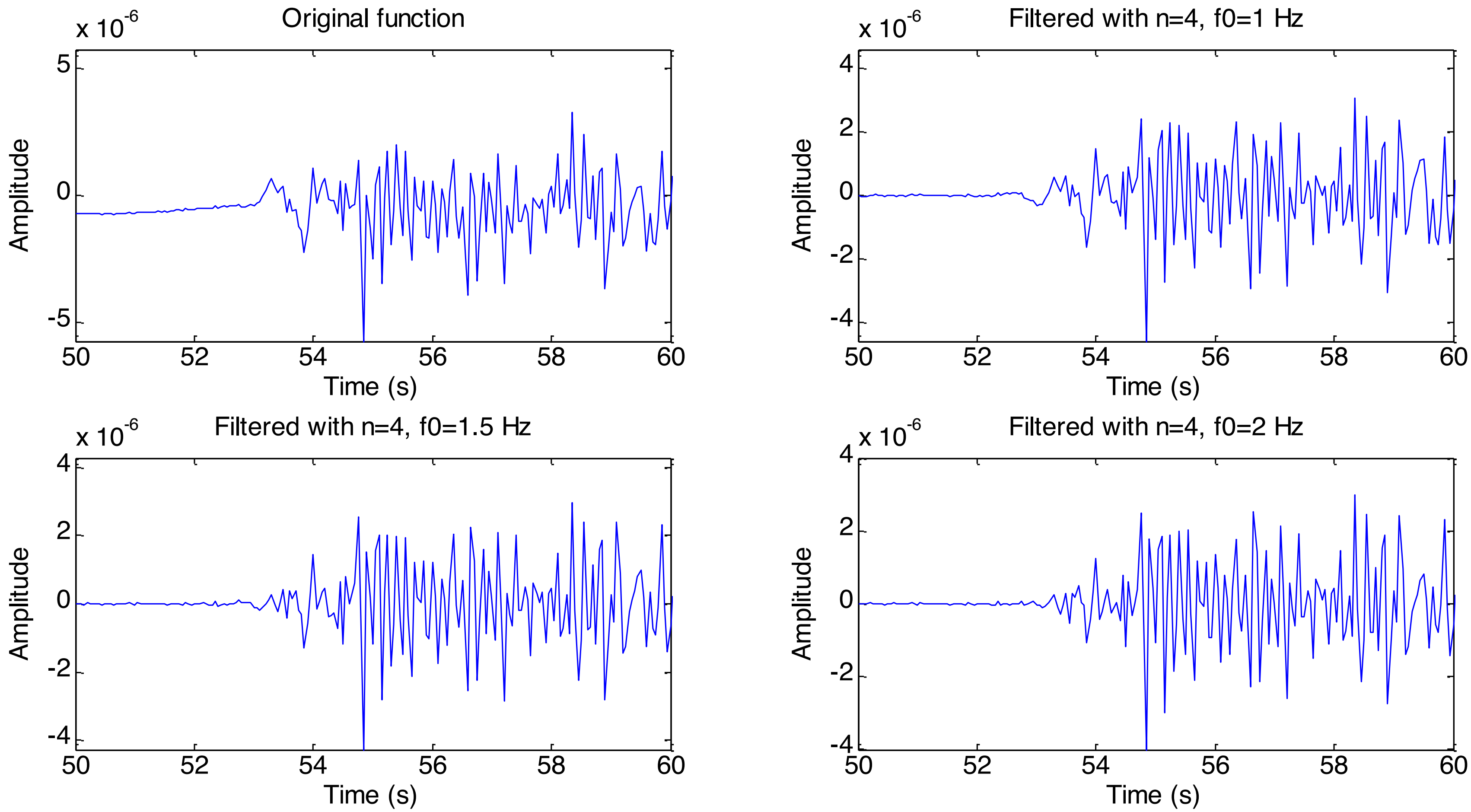
… on a seismogram …

… varying the order …



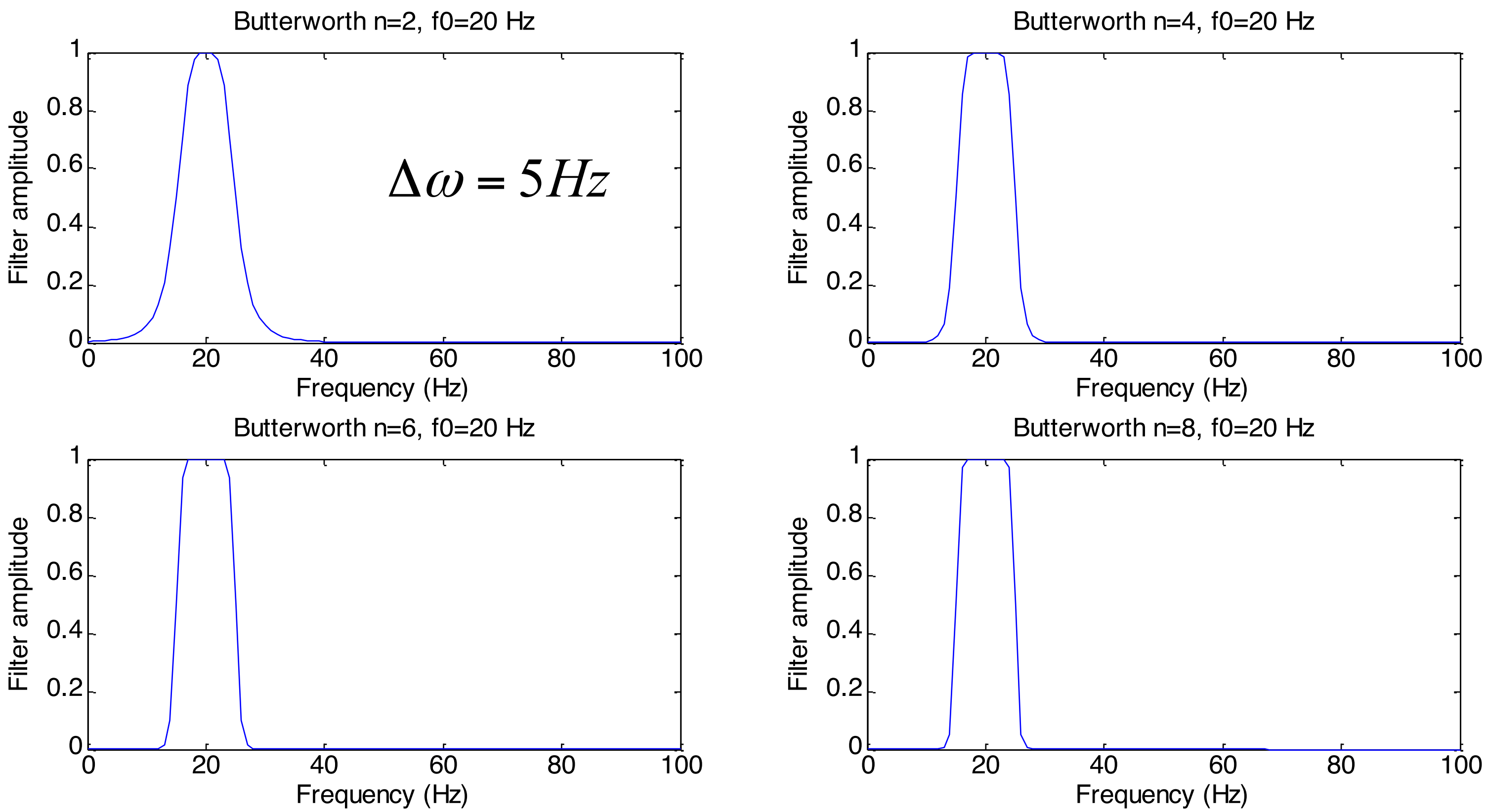
… on a seismogram …

… varying the cut-off frequency…

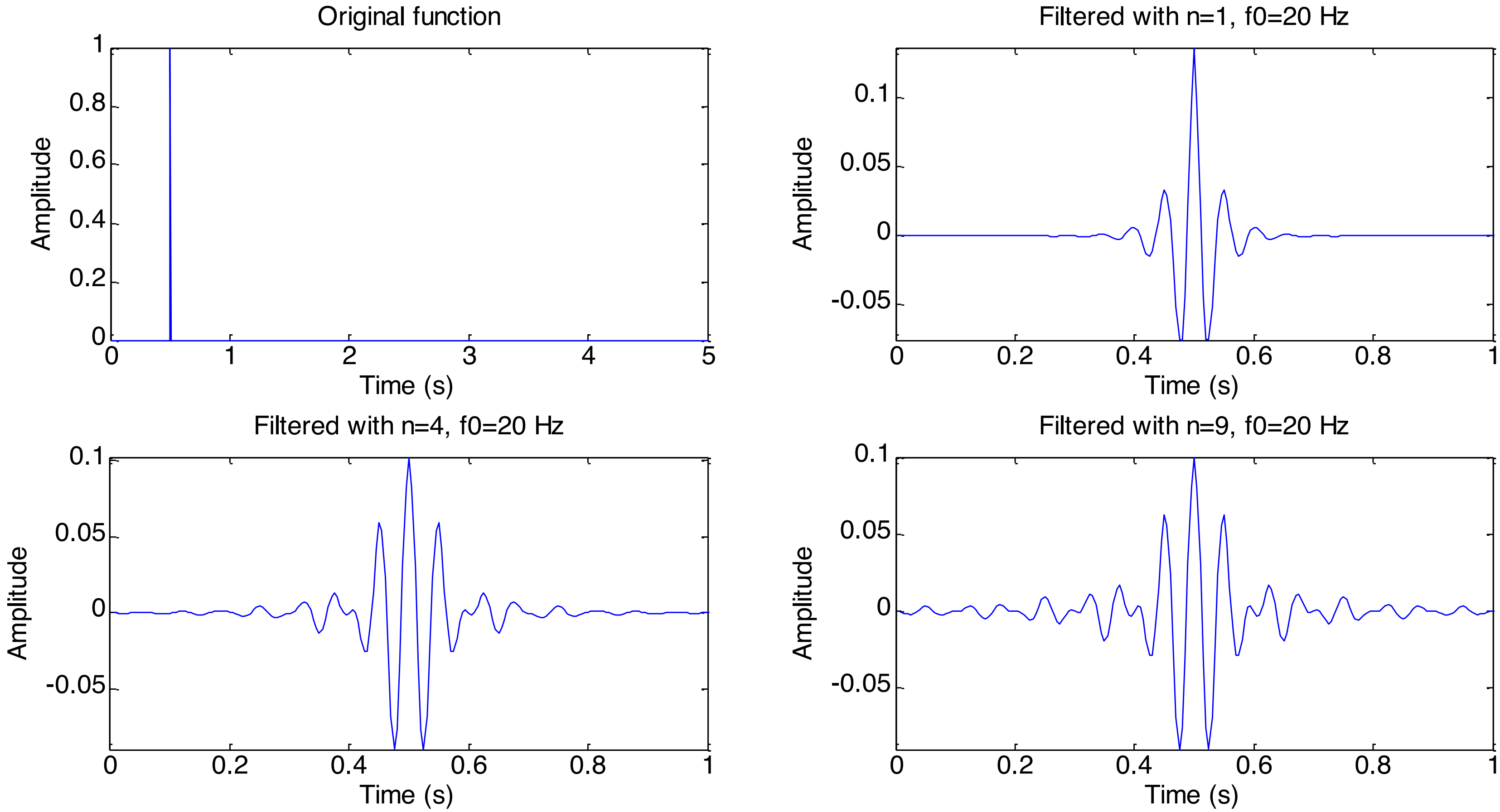


The Butterworth Filter (Band-Pass)

$$|F_{BP}(\omega)| = 1 - \frac{1}{1 + \left[(\omega - \omega_b)/\Delta\omega\right]^{2n}}$$

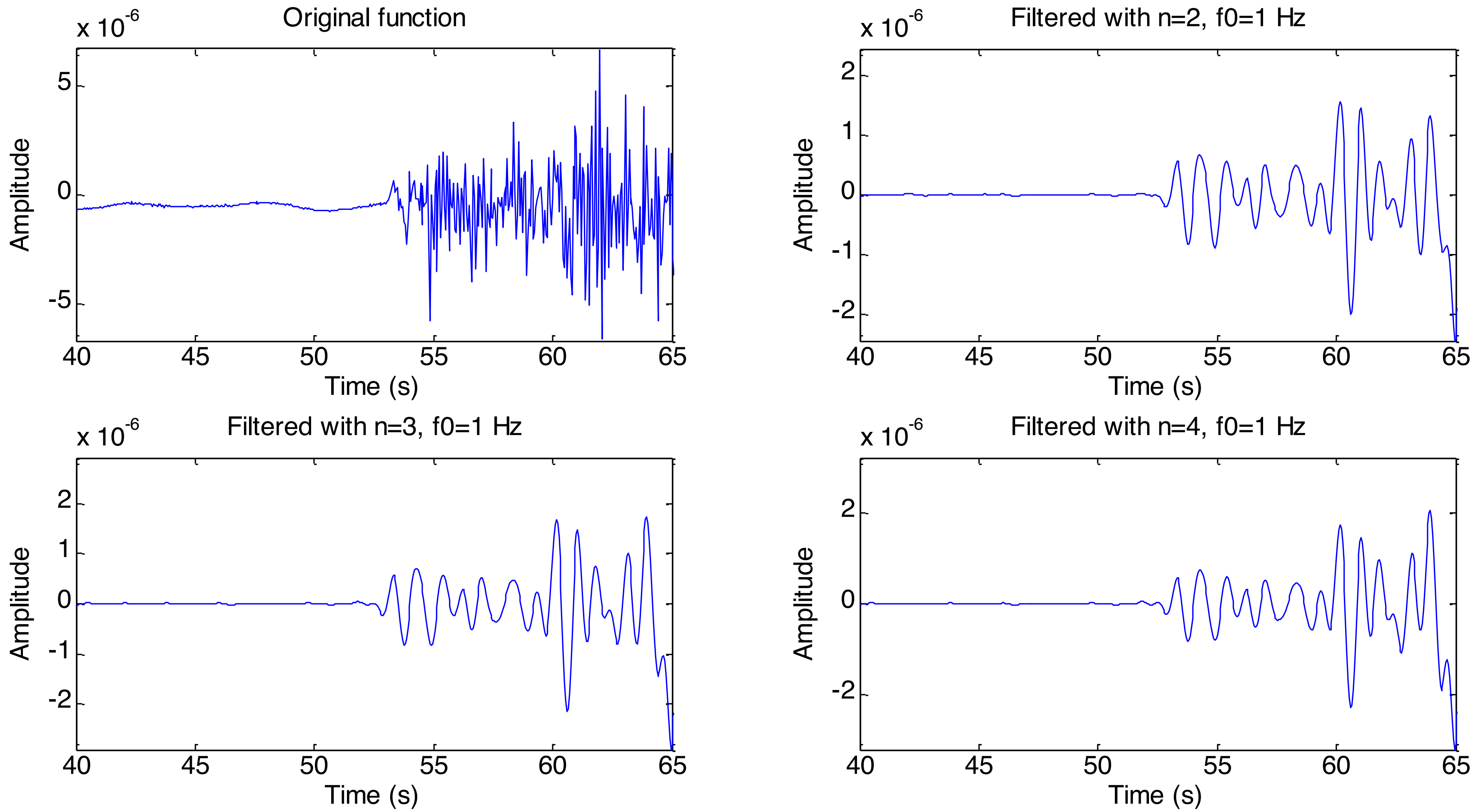


… effect on a spike …



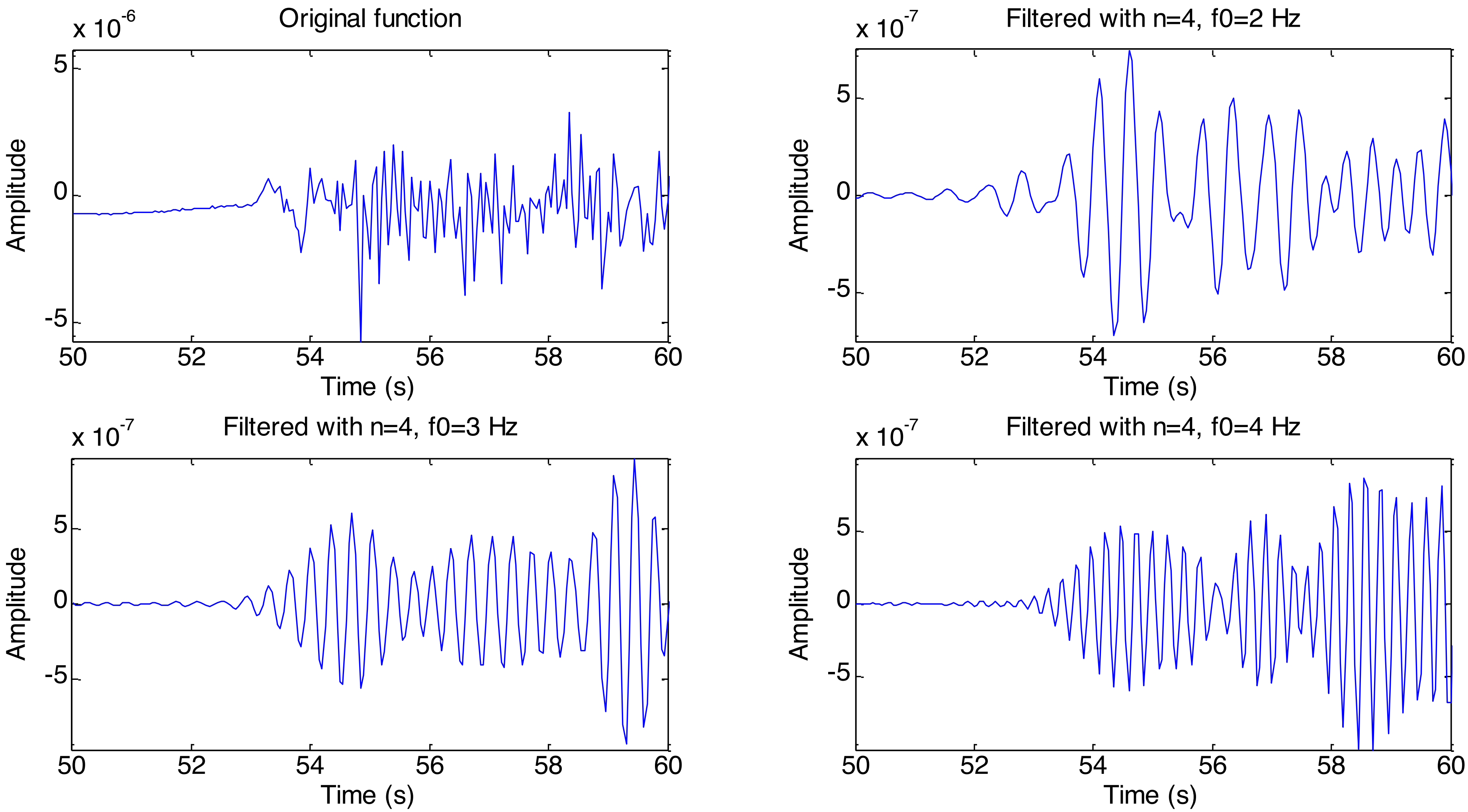
… on a seismogram …

… varying the order …

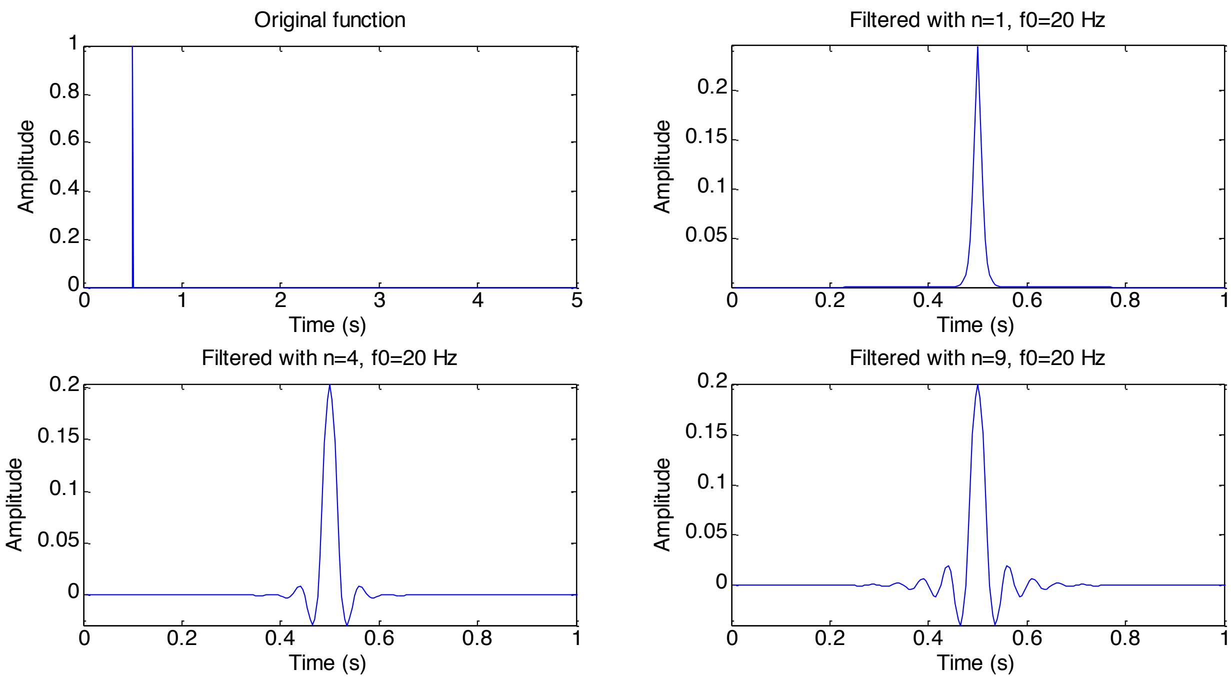


… on a seismogram …

… varying the cut-off frequency…



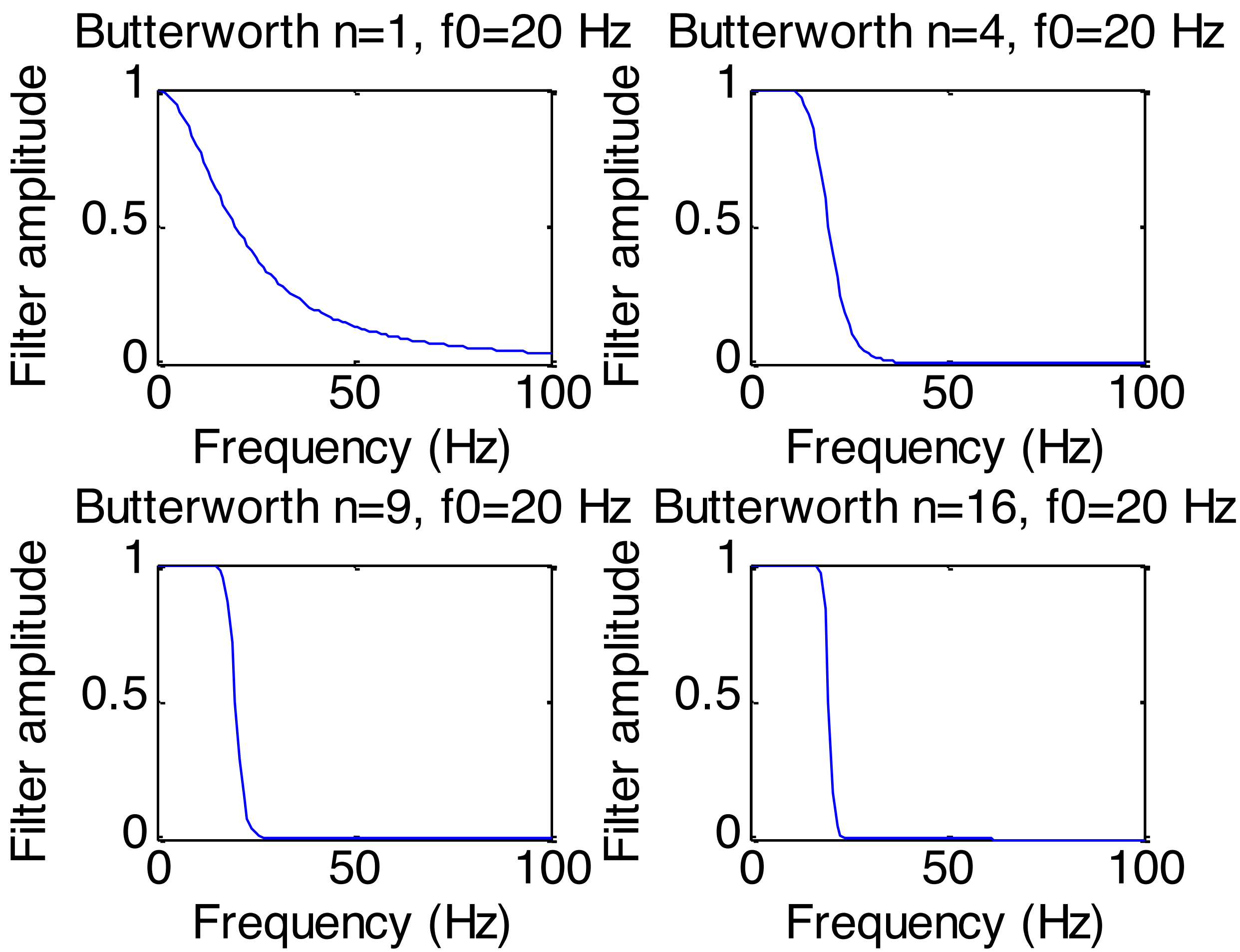
Zero phase and causal filters



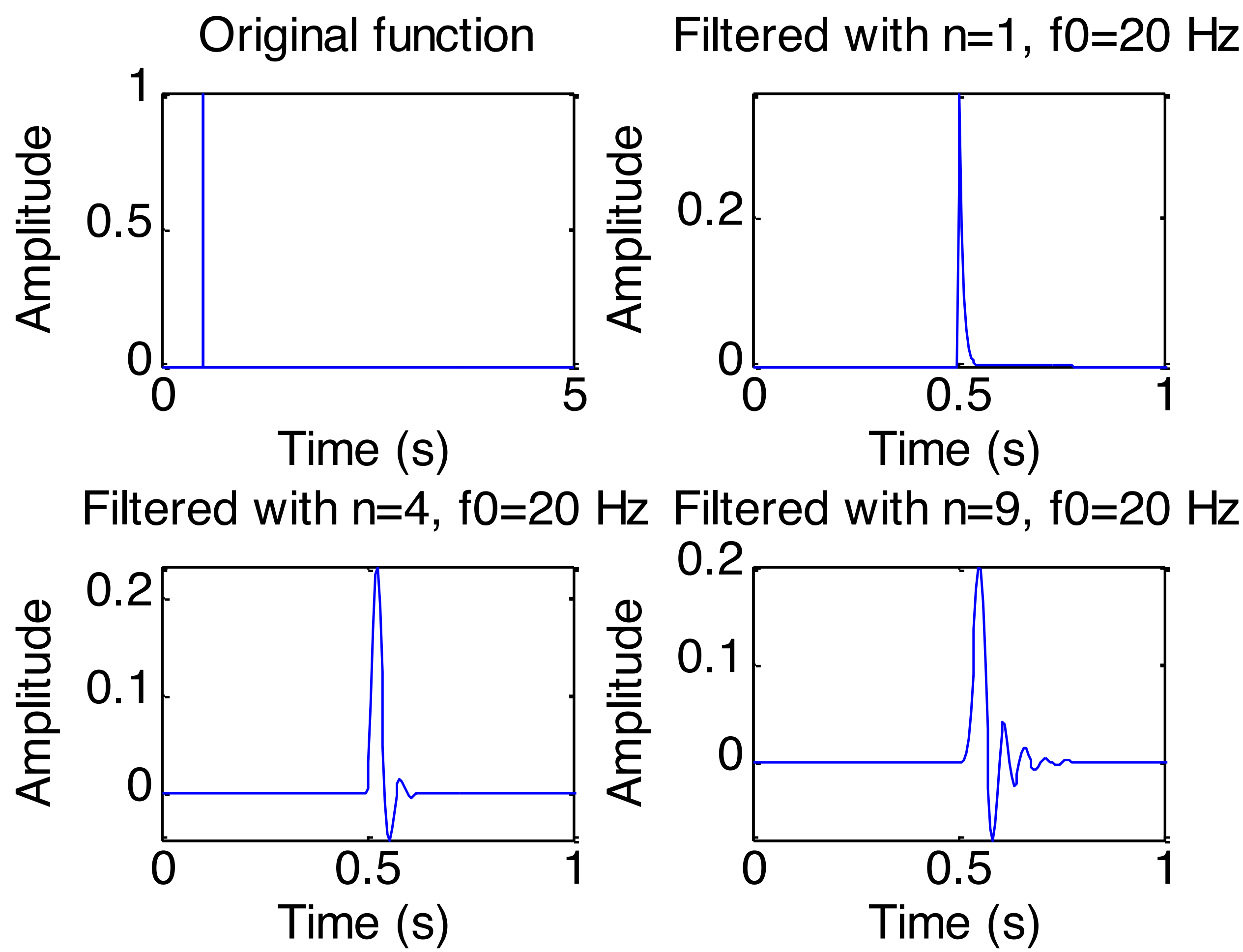
When the phase of a filter is set to zero (and simply the amplitude spectrum is inverted) we obtain a **zero-phase filter**. It means a peak will not be shifted.

Such a filter is **acausal**. Why?

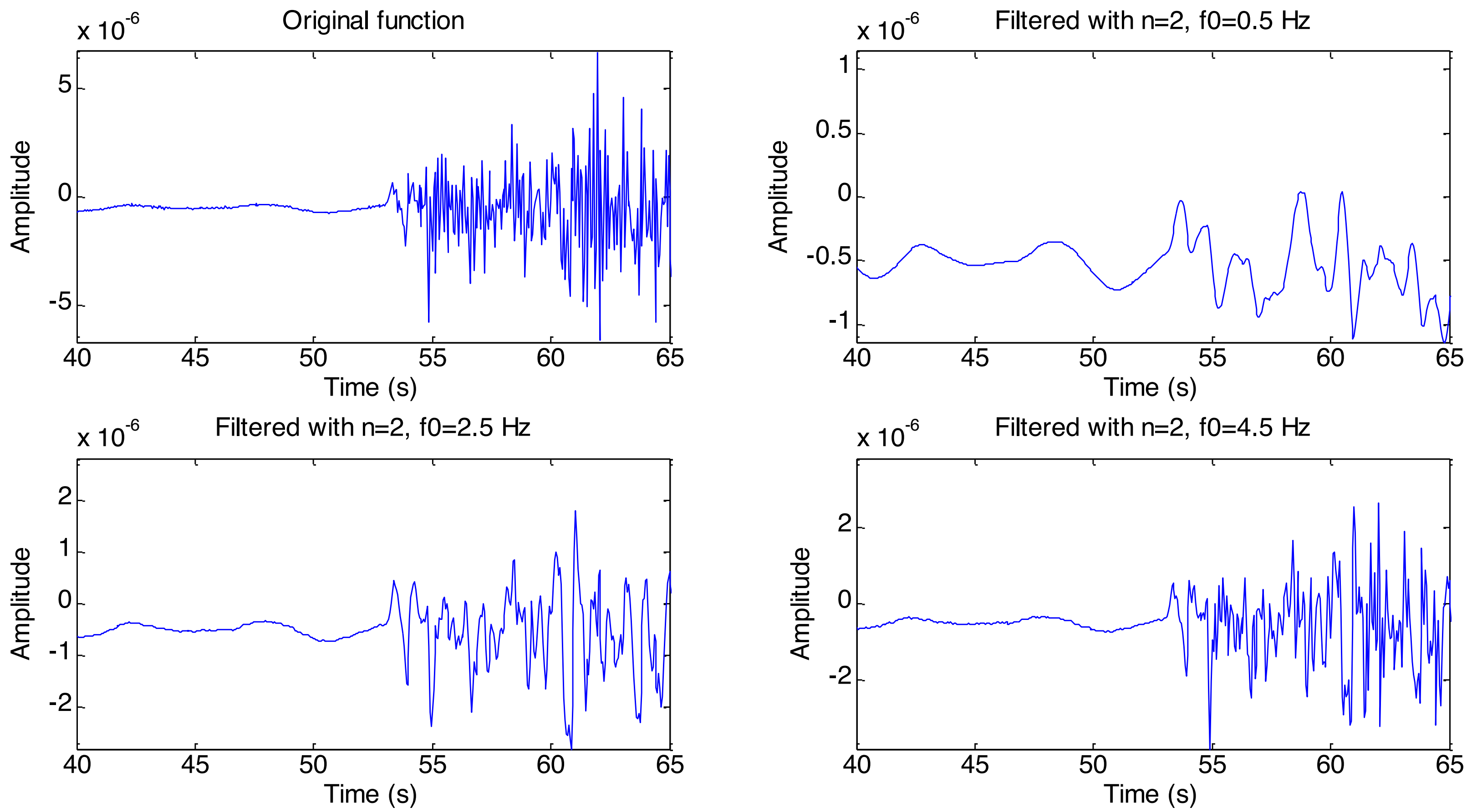
Butterworth Low-pass (20 Hz) on spike



(causal) Butterworth Low-pass (20 Hz) on spike



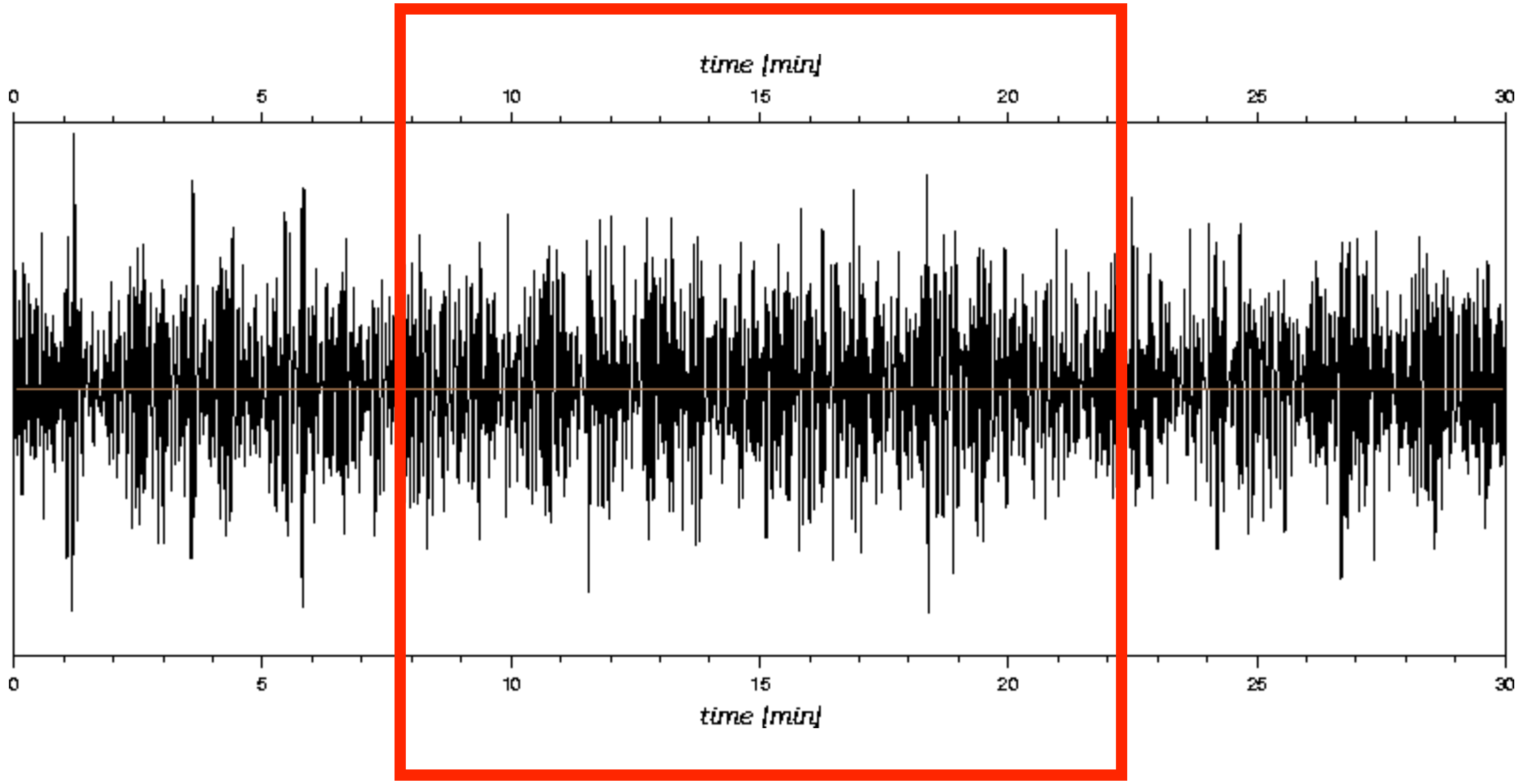
Butterworth Low-pass (20 Hz) on data



Other windowing functions

- So far we only used the Butterworth filtering window
- In general if we want to extract time windows from (permanent) recordings we have other options in the time domain.
- The key issues are
  - Do you want to preserve the main maxima at the expense of side maxima?
  - Do you want to have as little side lobes as posible?

Example



Possible windows

Plain box car (arrow stands for Fourier transform):

$$w_R(t) = \begin{Bmatrix} 1, & |t| \le M \\ 0, & |t| > M \end{Bmatrix} \Rightarrow \qquad W_R(f) = 2M\,\frac{\sin 2\pi f M}{2\pi f M}$$

Bartlett

$$w_R(t) = \begin{Bmatrix} 1 - \dfrac{|t|}{M}, & |t| \le M \\ 0, & |t| > M \end{Bmatrix} \Rightarrow \qquad W_B(f) = M\left(\frac{\sin 2\pi f M}{\pi f M}\right)^2$$

Possible windows

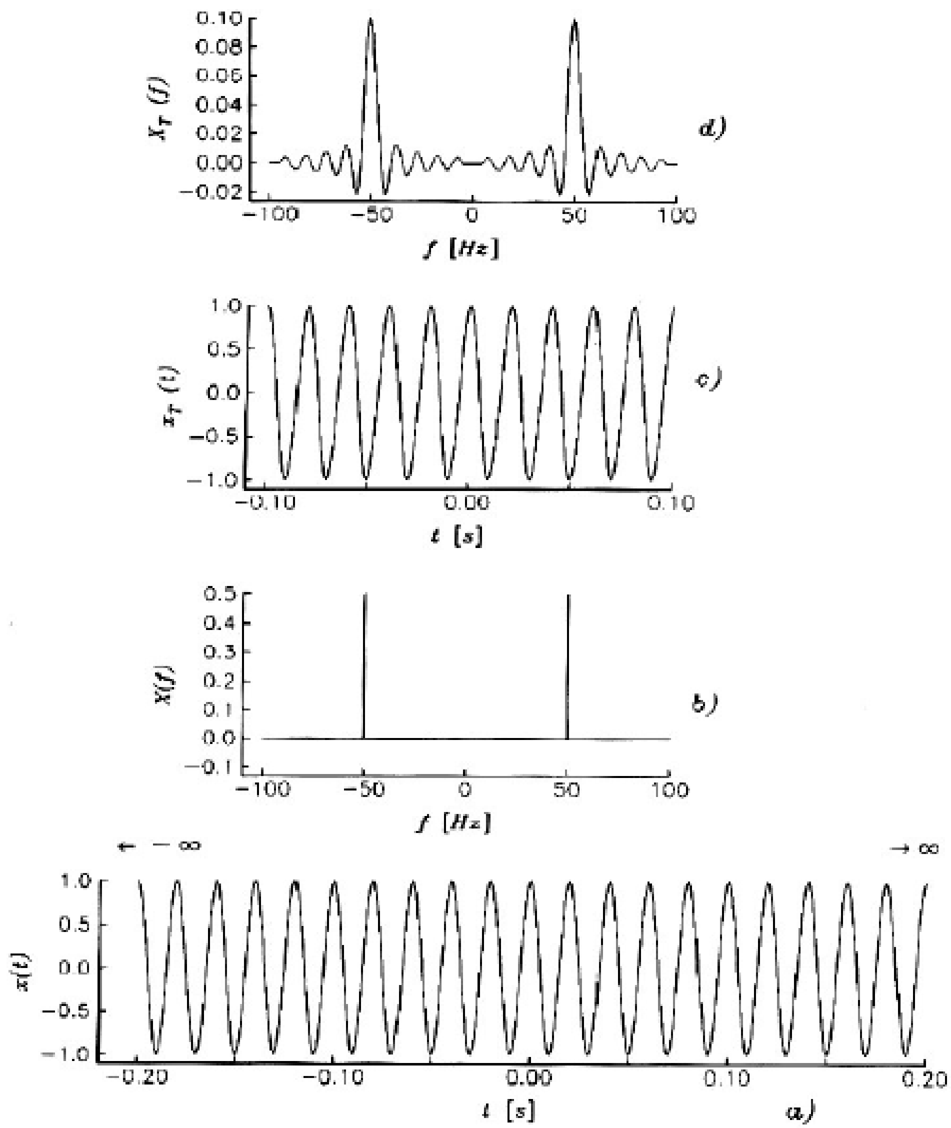
Hanning

$$w_H(t) = \begin{cases} \frac{1}{2}(1+\cos\frac{\pi t}{M}), & |t| \le M \\ 0, & |t| > M \end{cases} \Rightarrow \qquad W_H(f) = M\,\frac{\sin 2\pi fM}{2\pi fM}\,\frac{1}{1-(2\pi fM)^2}$$

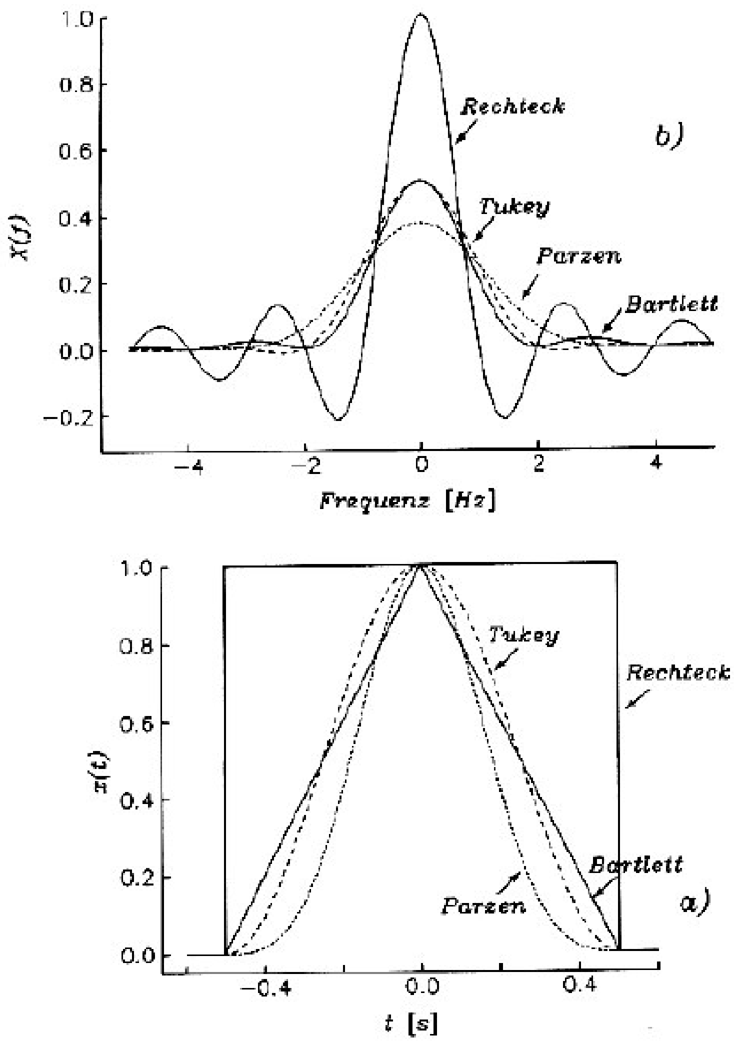
The spectral representations of the boxcar, Bartlett (and Parzen) functions are:

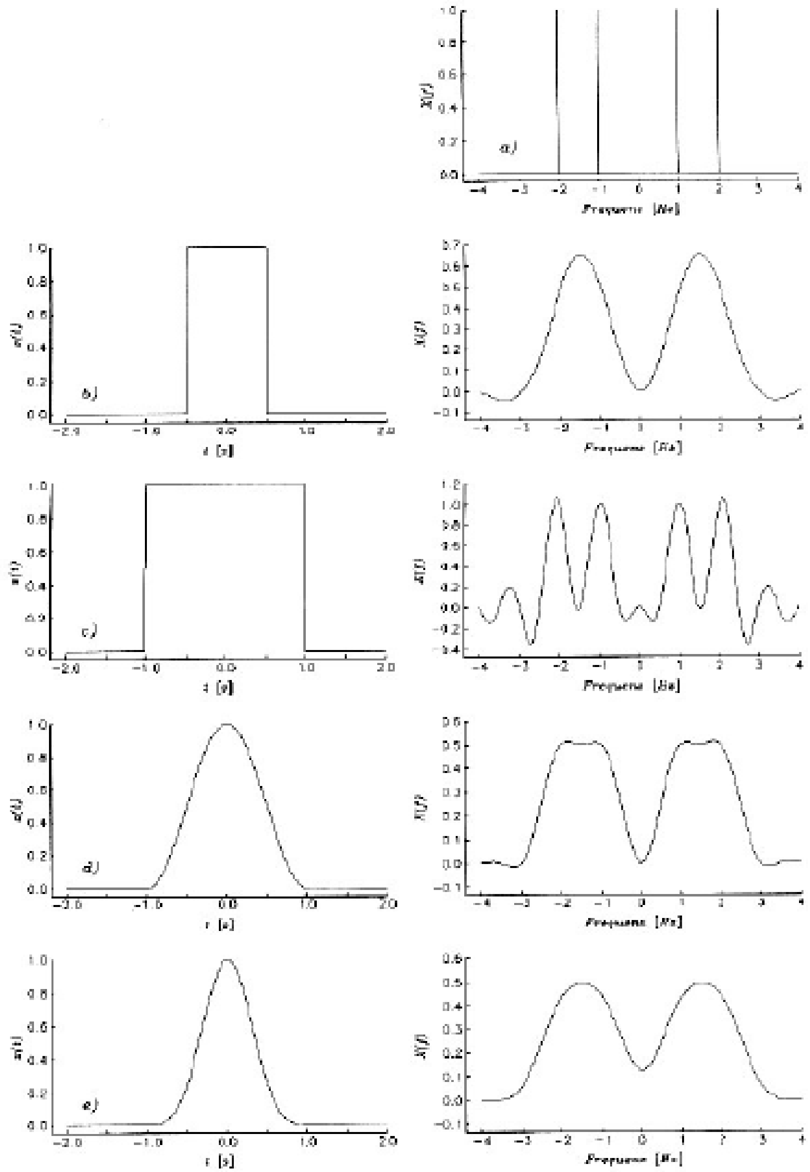
$$W(f) = \left(\frac{\sin 2\pi fM}{2\pi fM}\right)^n \qquad n = 1,2,4; \qquad M = T/2$$

Examples


0.10
0.08
0.06
0.04
0.02
0.00
−0.02
$X_T(f)$
−100
−50
0
50
100
f [Hz]
d)
1.0
0.5
0.0
−0.5
−1.0
$x_T(t)$
−0.10
0.00
0.10
t [s]
c)
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
$X(f)$
−100
−50
0
50
100
f [Hz]
b)
← −∞
→ ∞
1.0
0.5
0.0
−0.5
−1.0
$x(t)$
−0.20
−0.10
0.00
0.10
0.20
t [s]
a)

Examples

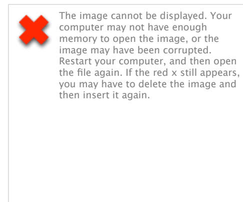

Rechteck
Tukey
Parzen
Bartlett
b)
X(f)
Frequenz [Hz]
Tukey
Rechteck
Bartlett
Parzen
a)
x(t)
t [s]



The Gabor transform: t-f misfits

phase information:

- can be measured reliably
- ± linearly related to Earth structure
- physically interpretable



amplitude information:

- hard to measure (earthquake magnitude often unknown)
- non-linearly related to structure

$$\hat{u}(\omega, t) := G(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(\tau)\, g(\tau - t)\, e^{-i\omega\tau} d\tau$$

[t-ω representation of synthetics, u(t)]

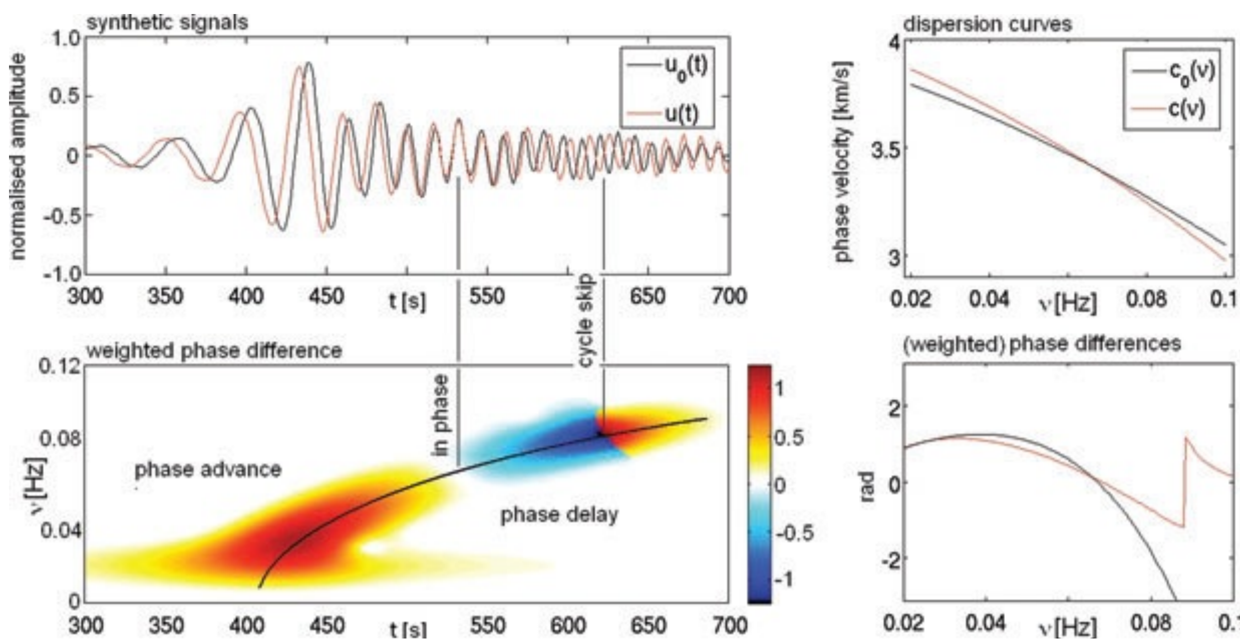
$$\hat{u}_0(\omega, t) := G(u_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u_0(\tau)\, g(\tau - t)\, e^{-i\omega\tau} d\tau$$

[t-ω representation of data, $u_0(t)$]

The Gabor time window

The Gaussian time windows is given by

$$g_\sigma(t) = \frac{1}{2\sqrt{\pi\sigma}} e^{-\frac{t^2}{4\sigma}}$$



Example

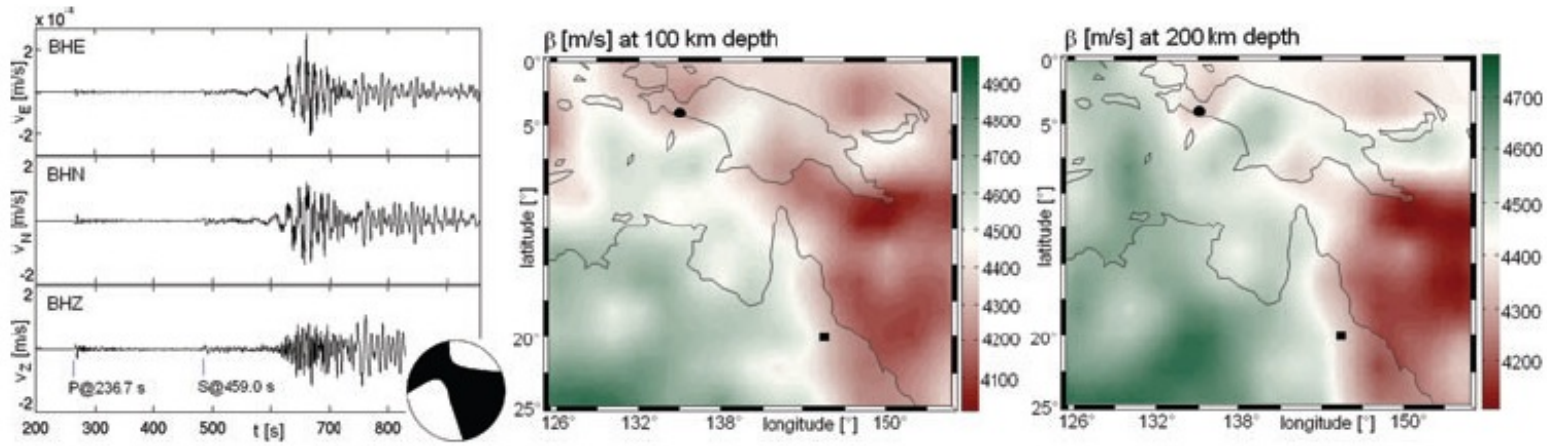


Figure 3. Left-hand panel: unprocessed velocity seismograms of the West Irian event (1993 June 12) recorded at the permanent station CTAO, located in NW Australia. The CMT solution is visualized in the lower right-hand panel of the BHZ channel recording. Central panel: model of the S-wave velocity β at the depth of 100 km. The maximum lateral variations $\Delta\beta$ reach 10 per cent of the background value. The source and receiver locations are plotted as a circle (•) and a square (****), respectively. Right-hand panel: the same as in the central one but at the depth of 200 km. The lateral variations are smaller than at 100 km depth and finally vanish below 350 km.

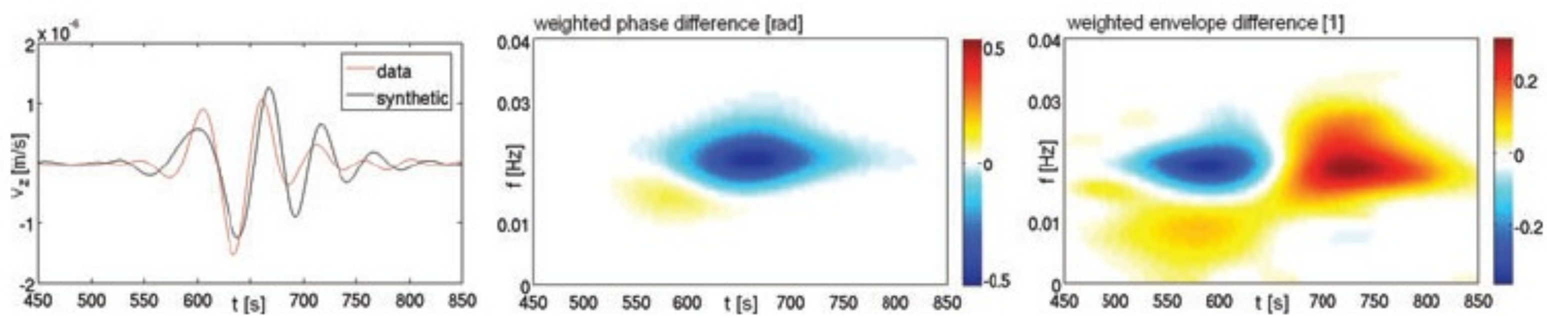


Figure 4. Left-hand panel: comparison of vertical-component surface wave trains, low-pass filtered with a cut-off frequency of 0.02 Hz (50 s). The synthetic is plotted in black and the data in red. Centre panel: weighted phase difference in time–frequency space. Both positive and negative phase differences are observable. Right-hand panel: weighted envelope difference in time–frequency space.

Multitaper

Goal: „*obtaining a spectrum with little or no bias and small uncertainties*“. problem comes down to finding the right tapering to reduce the bias (i.e, spectral leakage).

In principle we seek:

$$\hat{S}(f) = \left| \sum_{t=0}^{N-1} x(t)a(t)\, e^{-2\pi i f t} \right|^2,$$

where $a(t)$ is a series of weights called a taper.

This section follows Prieto eet al., GJI, 2007. Ideas go back to a paper by Thomson (1982).

Multi-taper Principle

$$\hat{S}(f) = \left| \sum_{t=0}^{N-1} x(t)a(t)\, e^{-2\pi i f t} \right|^2$$

- Data sequence x is multiplied by a set of orthgonal sequences (tapers)
- We get several single periodograms (spectra) that are then averaged
- The averaging is not even, various weights apply
- Tapers are constructed to optimize resistance to spectral leakage
- Weighting designed to generate smooth estimate with less variance than with single tapers

Spectrum estimates

We start with

$$\hat{S}(f) = |Y(f)|^2 = \left| \sum_{t=0}^{N-1} x(t)a(t)\, e^{-2\pi i f t} \right|^2.$$

with

$$\sum_{t=0}^{N-1} |a(t)|^2 = 1 \qquad\qquad A(f) = \sum_{t=0}^{N-1} a(t)\, e^{-2\pi i f t}$$

To maintain total power.

Condition for optimal tapers

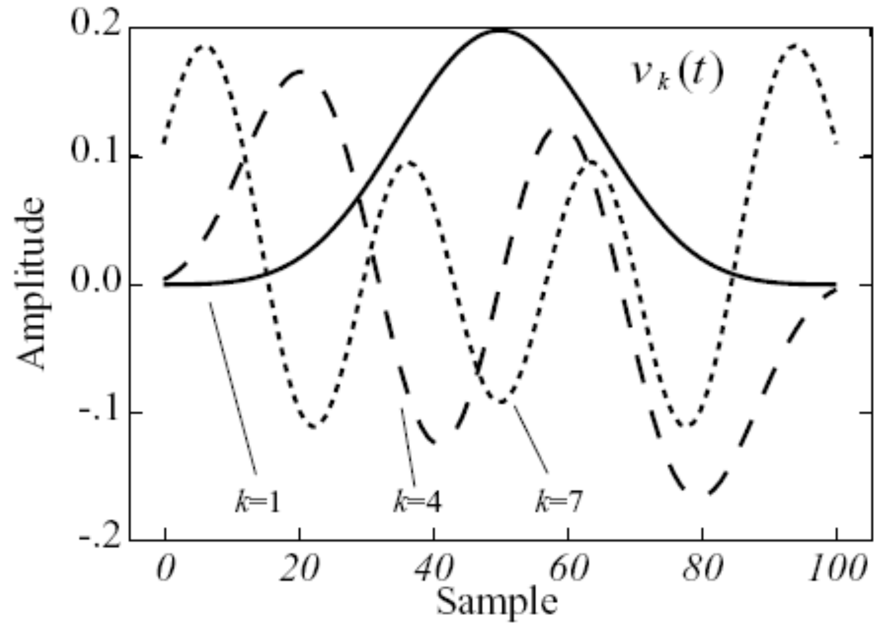
N is the number of points, W is the resolution bandwith (frequency increment)

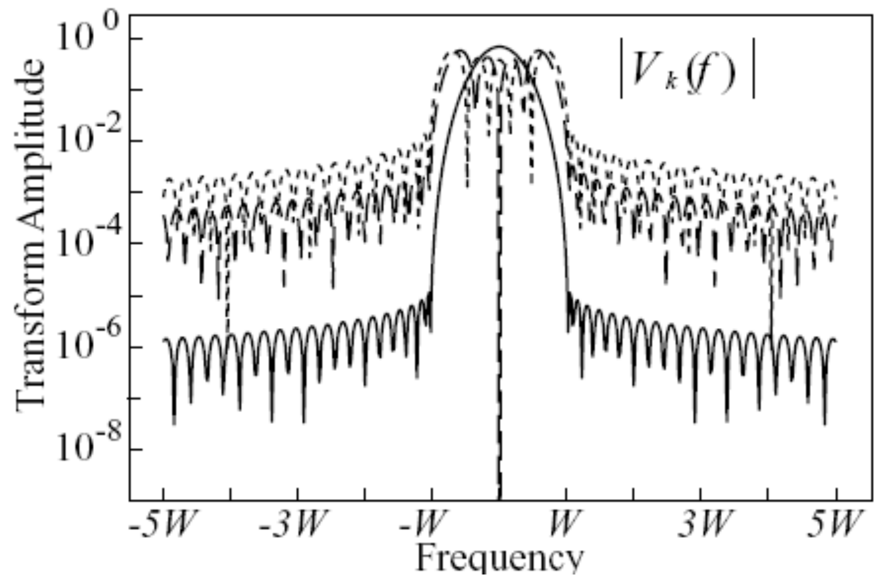
$$\lambda(N, W) = \frac{\int_{-W}^{W} |A(f)|^2 \, \mathrm{d}f}{\int_{-\frac{1}{2}}^{\frac{1}{2}} |A(f)|^2 \, \mathrm{d}f}$$

One seeks to maximize λ the fraction of energy in the interval (–W,W). From this equation one finds a ‘s by an eigenvalue problem -> Slepian function

Slepian functions

The tapers (Slepian functions) in time and frequency domains





Final assembly

The objective of this method is to estimate the spectrum $S(f)$ by using K of the Slepian sequences to obtain the k *eigencomponents*:

$$Y_k(f) = \sum_{t=0}^{N-1} x(t) v_k(t) e^{-2\pi i f t}$$

Slepian sequences (tapers)

and a set of K eigenspectra as in (7):

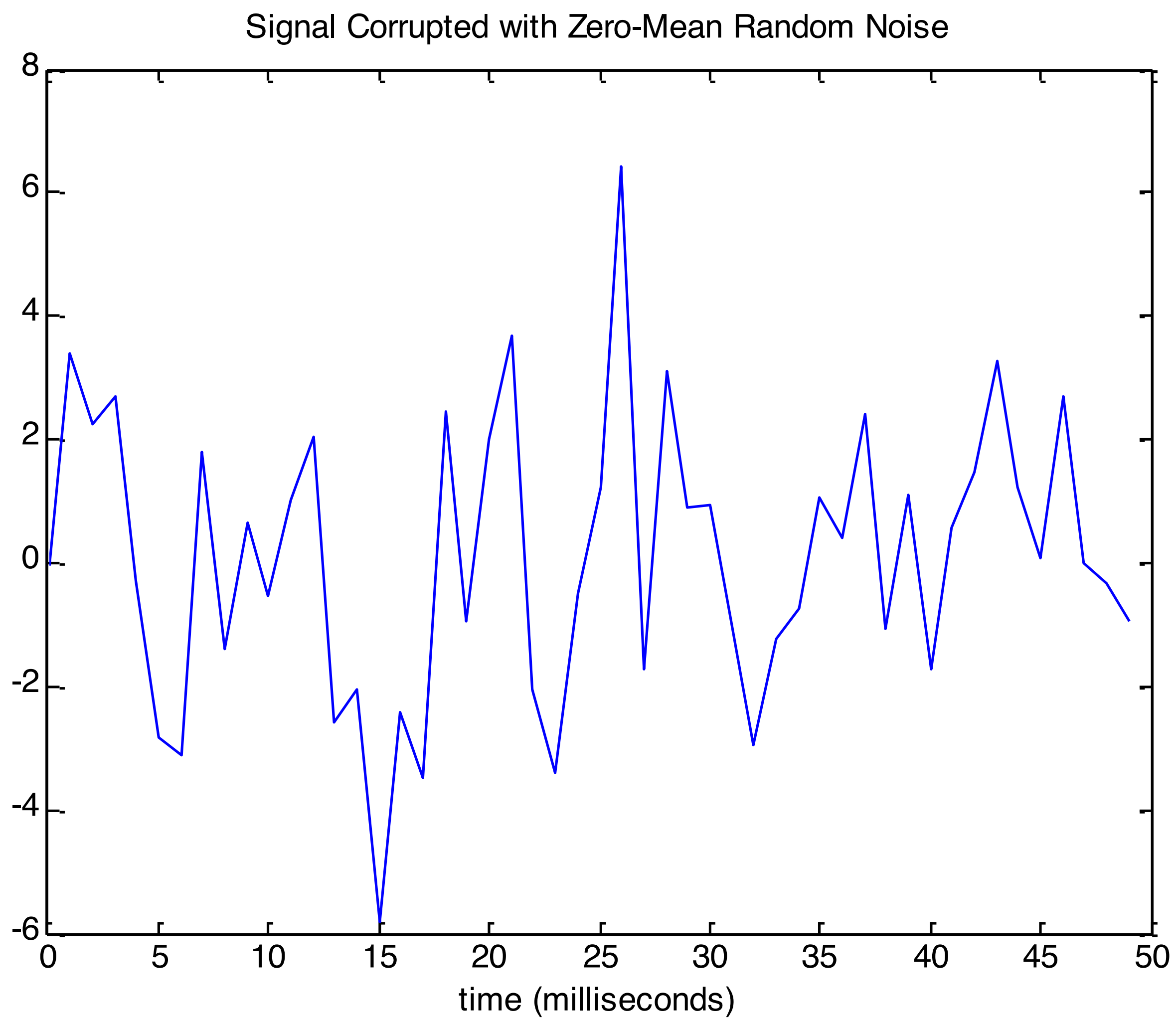
$$\hat{S}_k(f) = |Y_k(f)|^2$$

from which we can form the mean spectrum

$$\bar{S}(f) = \frac{1}{K} \sum_{k=1}^{K} \hat{S}_k(f).$$

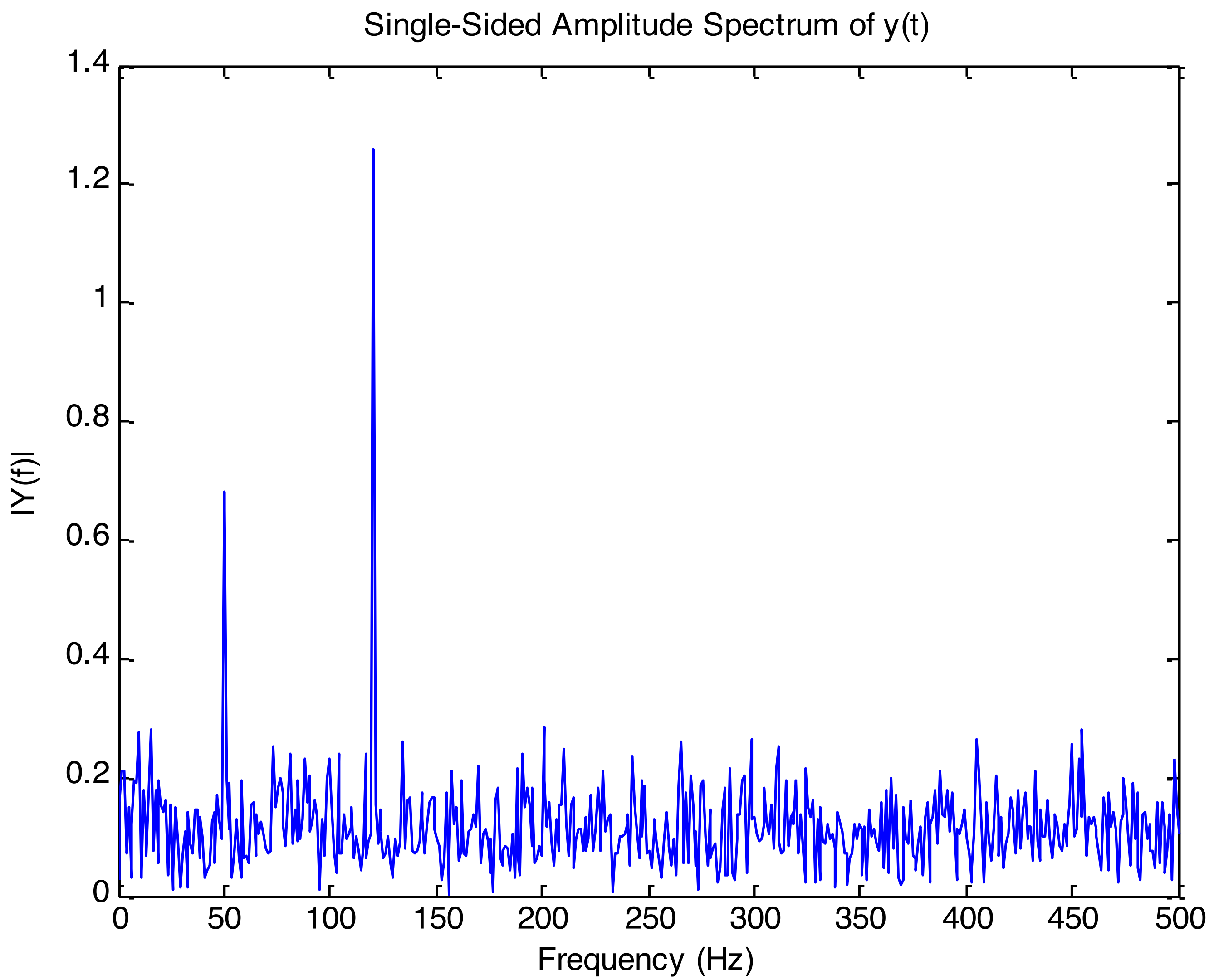
Final averaging of spectra

Example

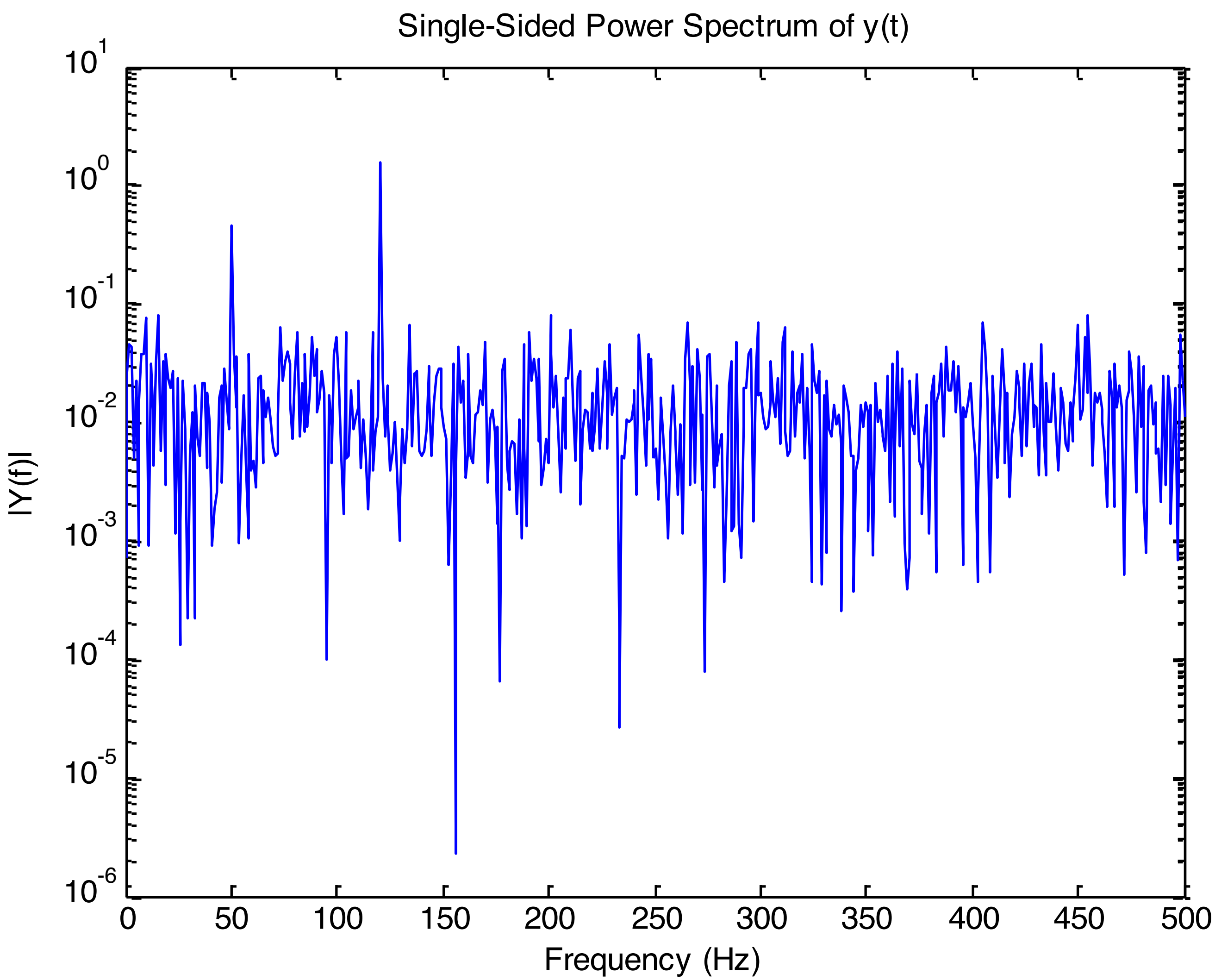


Signal Corrupted with Zero-Mean Random Noise
8
6
4
2
0
-2
-4
-6
0
5
10
15
20
25
30
35
40
45
50
time (milliseconds)

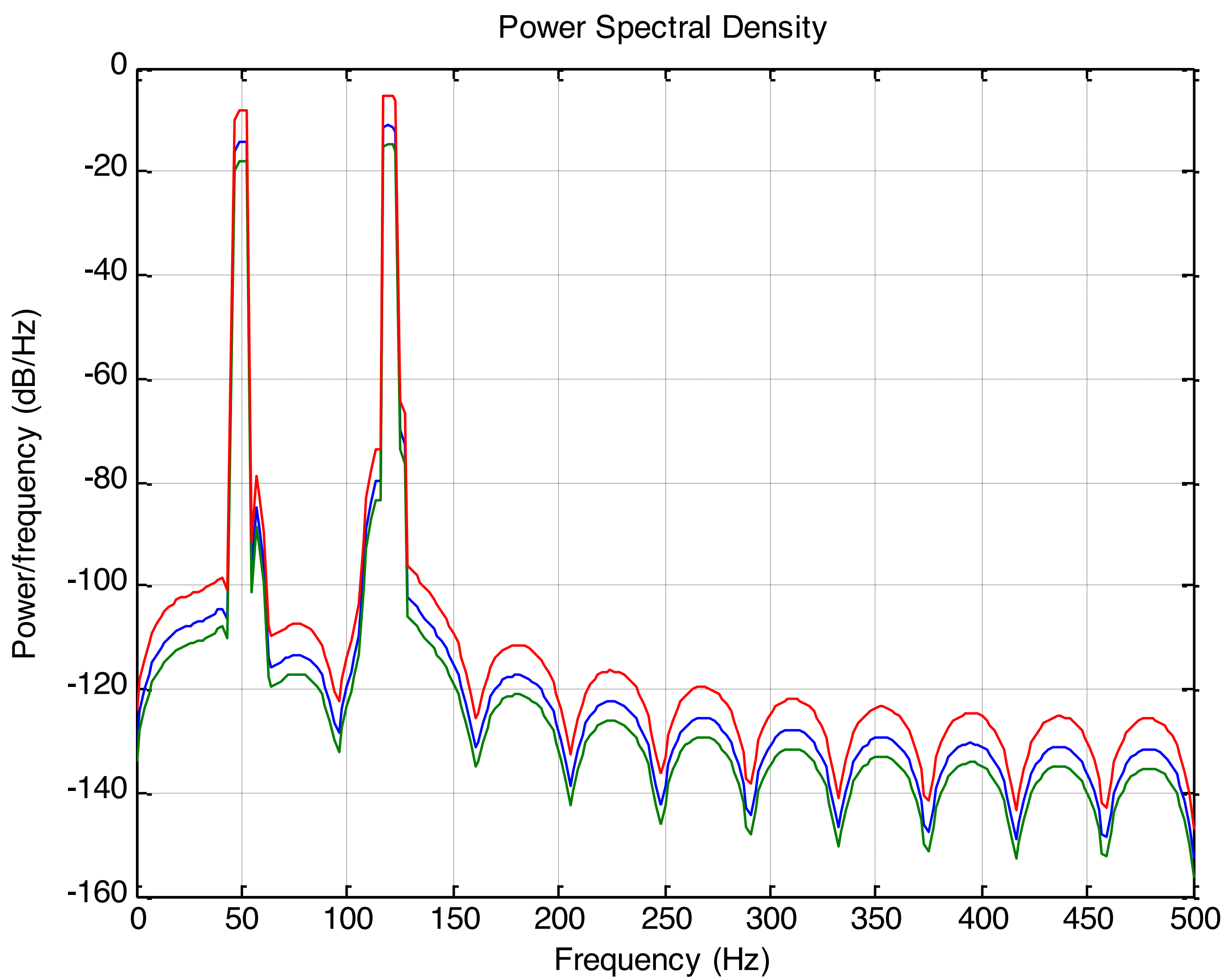
Classical Periodogram



… and its power …



… multitaper spectrum …



Wavelets – the principle

Motivation:

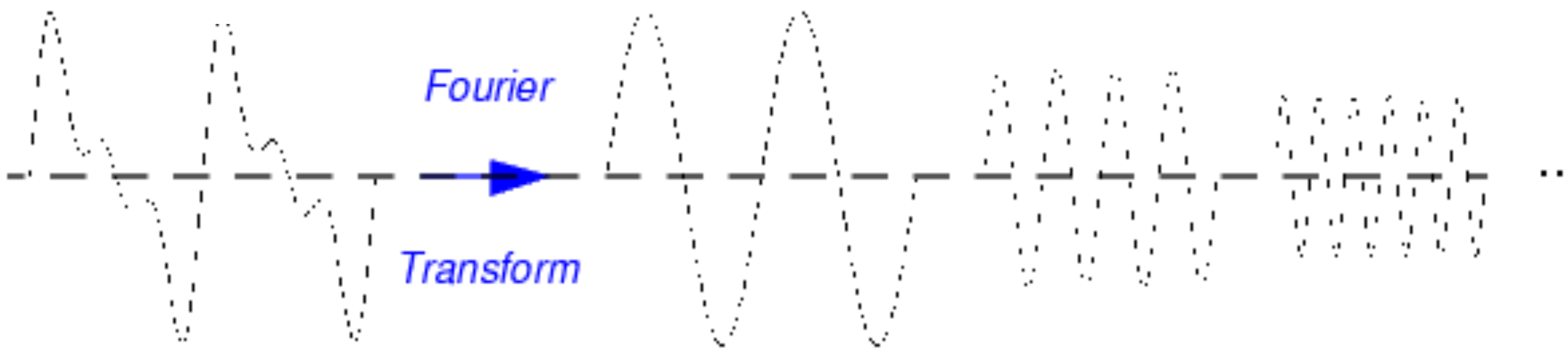
- Time-frequency analysis
- Multi-scale approach
- „when do we hear what frequency?“

Continuous vs. local basis functions

Sine Wave

Wavelet (db10)

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



Signal

Constituent sinusoids of different frequencies

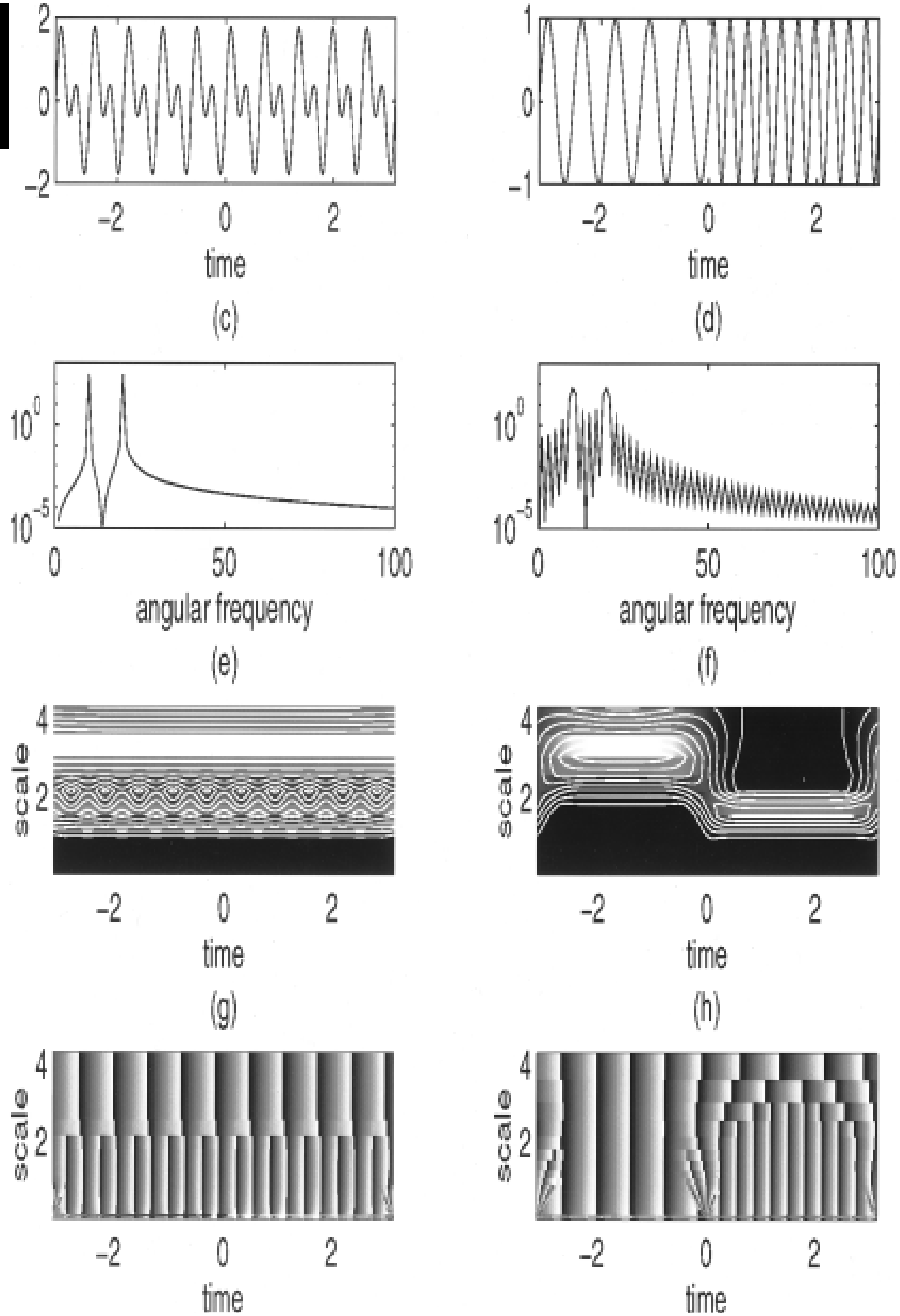


Figure 1. Spectral and wavelet analysis of two signals. The first signal (Figure 1a) consists of superposition of two frequencies (sin $10t$ and sin $20t$), and the second (Figure 2b) consists of the same two frequencies, each applied separately over half of the signal duration. Figures 1c and 1d show the Fourier spectra of the signals (i.e., $|f(\omega)|^2$ versus ω) for Figures 1a and 1b, respectively. Figures 1e and 1f show the magnitude of their wavelet transforms, and Figures 1g and 1h show the phase of their wavelet transforms (using Morlet wavelet). Notice the instability in calculation of phase at small scales where the modulus of wavelet transforms is very small.

Some maths

A wavelet can be defined as

$$\Psi^{a,b}(t) = |a|^{-1/2}\Psi\left(\frac{t-b}{a}\right)$$

With the transform pair:

$$W_\Psi(f)(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t)\Psi\left(\frac{x-b}{a}\right)dt$$

$$f(t) = C_\Psi \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\langle f, \Psi^{a,b}\right\rangle \Psi^{a,b}(t)a^{-2}dadb$$

[redacted]

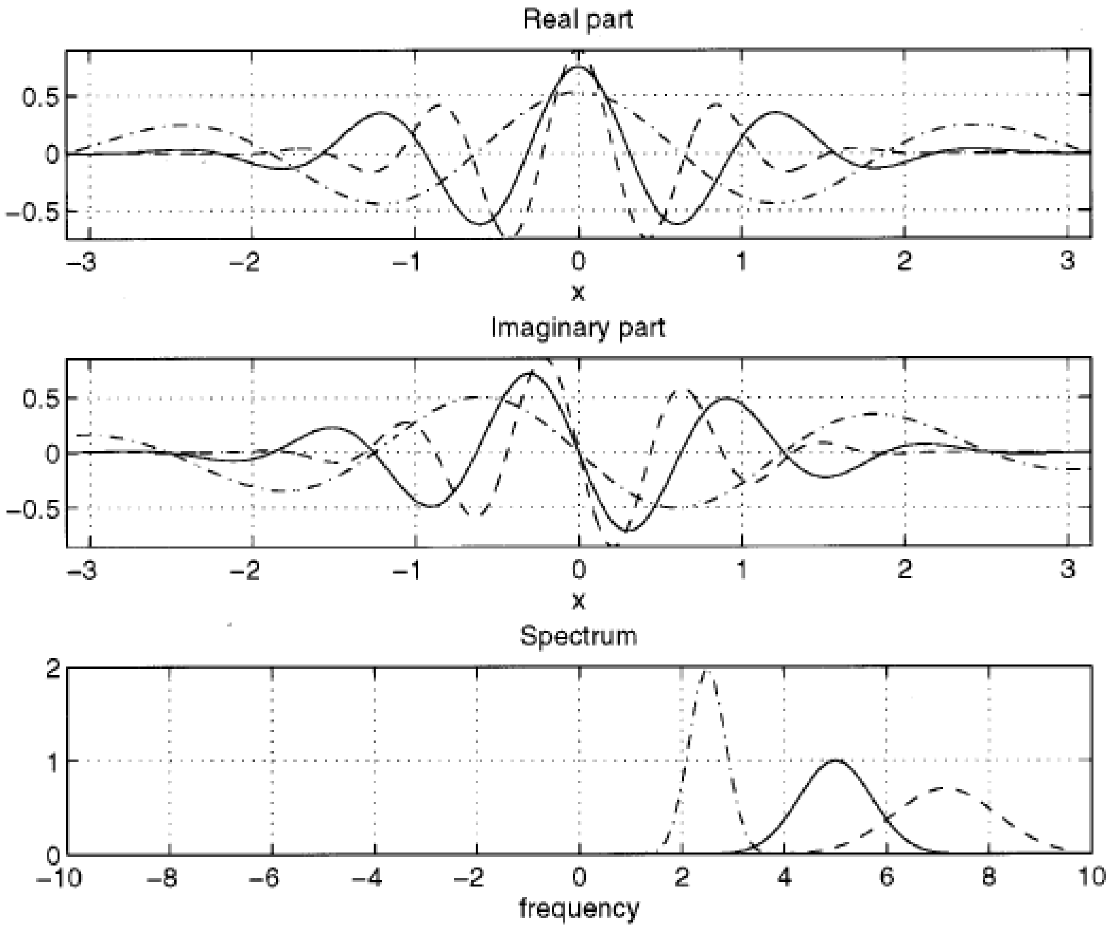
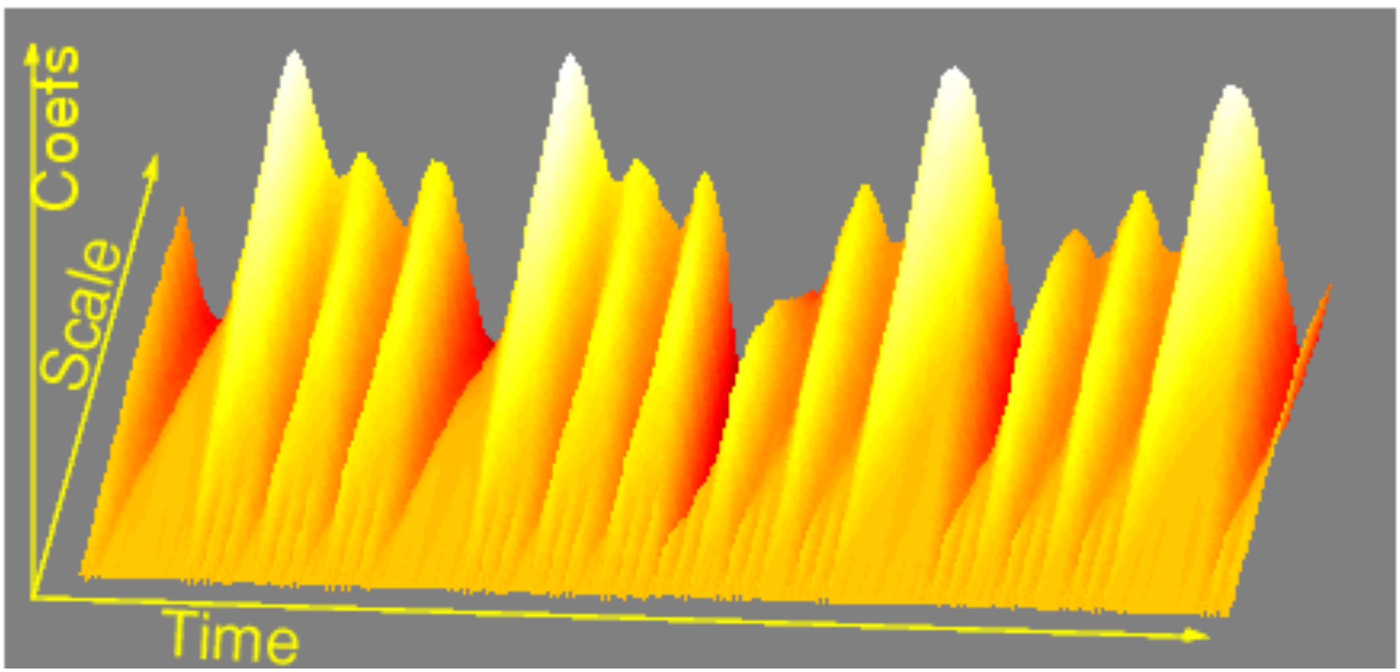
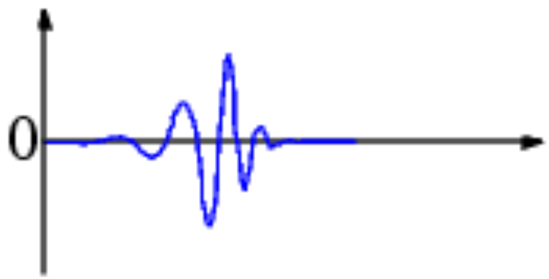


Figure 4. Real and imaginary parts of the Morlet wavelet ($\omega_0 = 5$) and its Fourier spectrum for different scales: $\lambda < 1$ (dashed lines), $\lambda = 1$ (solid lines), and $\lambda > 1$ (dash-dotted lines). Notice the effect of dilation on the wavelet and the corresponding change in its Fourier spectrum. When the wavelet dilates, its Fourier transform contracts, and vice versa. Also notice that the Fourier transform of the Morlet wavelet is supported entirely on the positive-frequency axis.

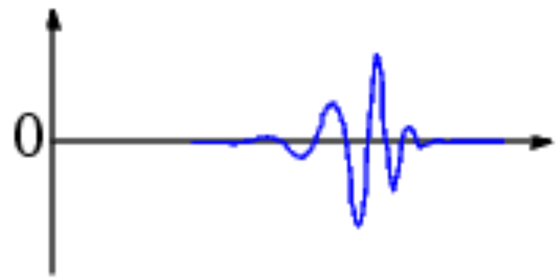
Resulting wavelet representation



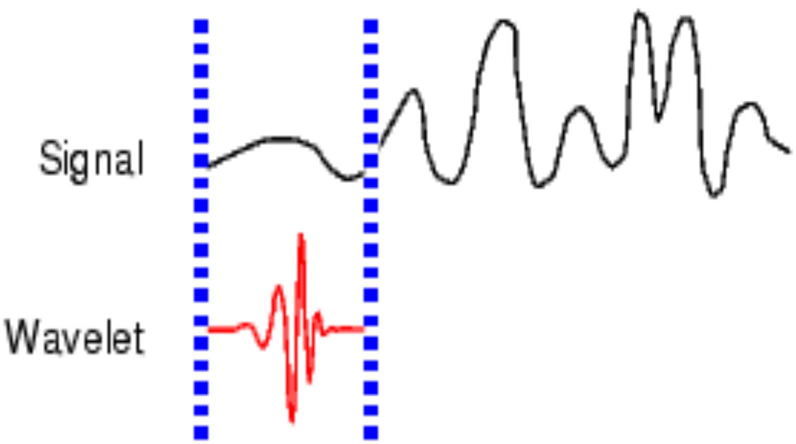
Shifting and scaling



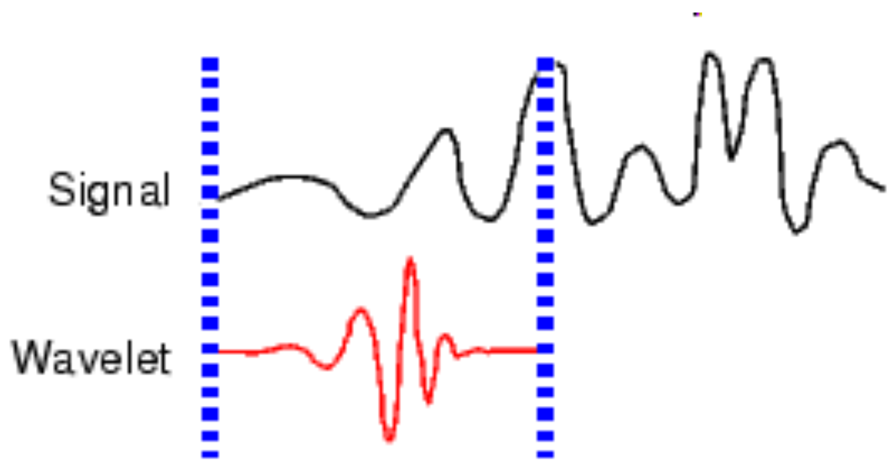
Wavelet function
$\psi(t)$



Shifted wavelet function
$\psi(t-k)$



C = 0.0102



C = 0.2247

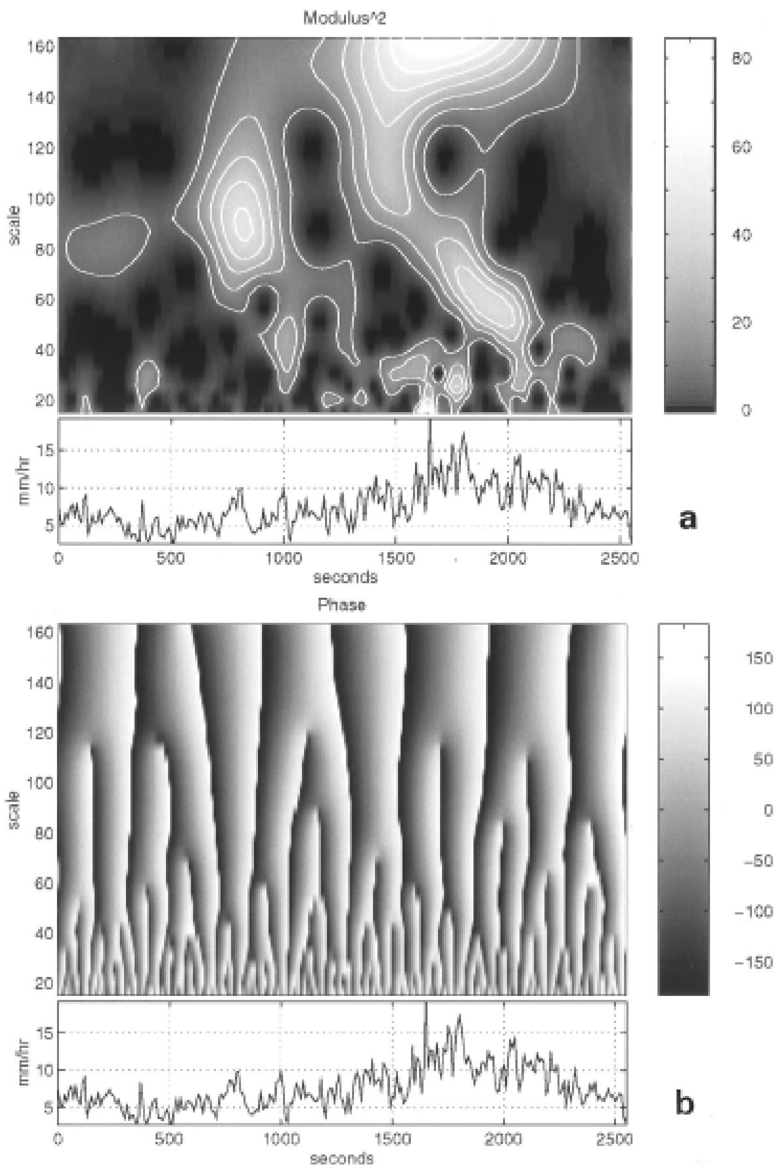
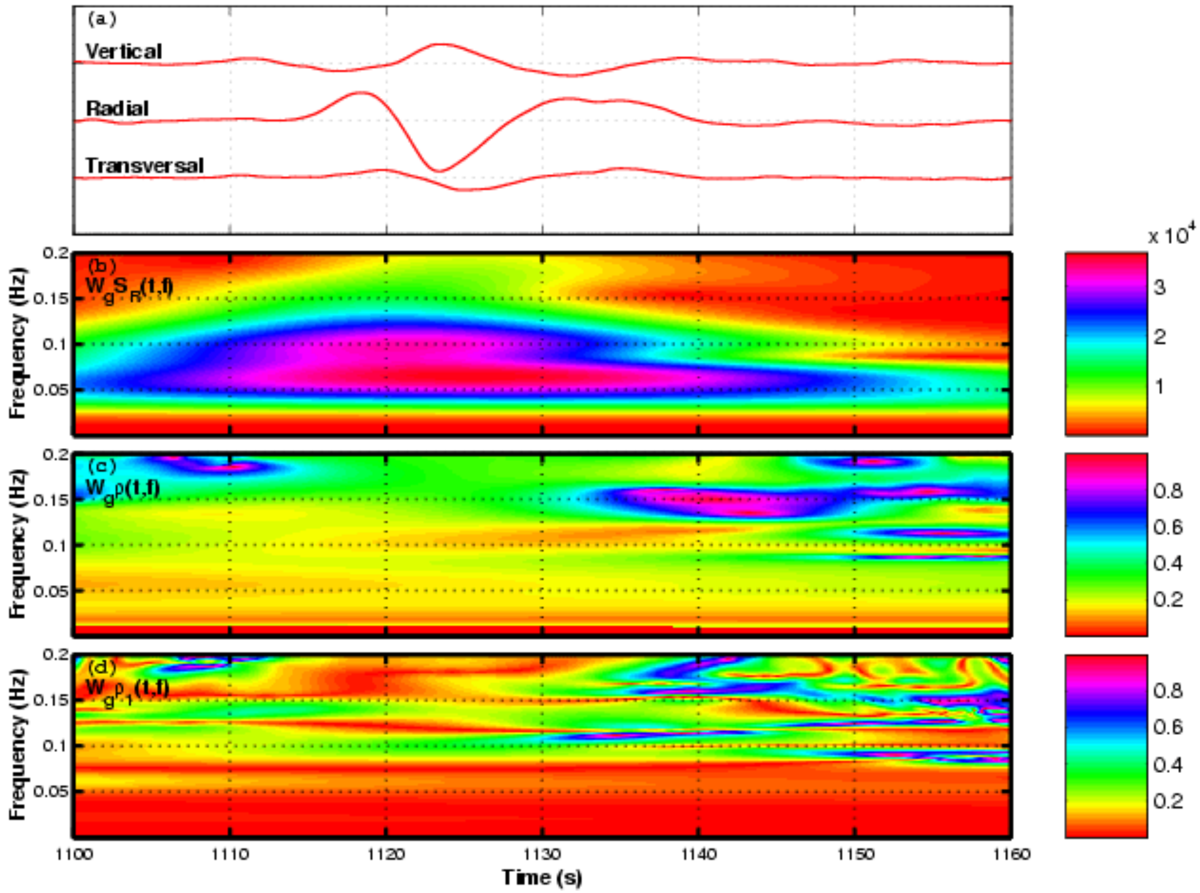


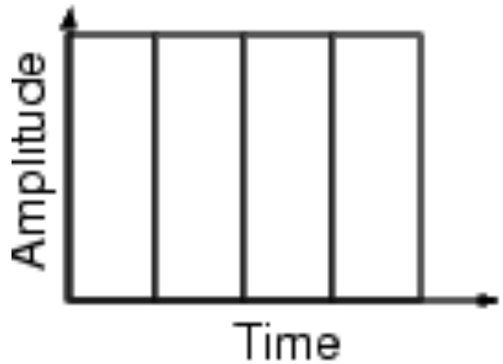
Figure 5. Analysis of temporal rainfall using the Morlet wavelet. The data were collected every 10 s on May 3, 1990, over Iowa City, Iowa, using an optical rain gage. (a) Square of the modulus or scalogram, that is, $|Wf(\lambda, t)|^2$, and (b) phase of $Wf(\lambda, t)$. The rainfall intensity is shown at the bottom of each figure. The scalogram clearly shows the presence of multiscale features and also some embedding of small-scale features within large-scale features. The phase plot shows the convergence of lines of constant phase to singularities (see discussion in text). (Reprinted by permission of Academic Press.)

Application to seismograms

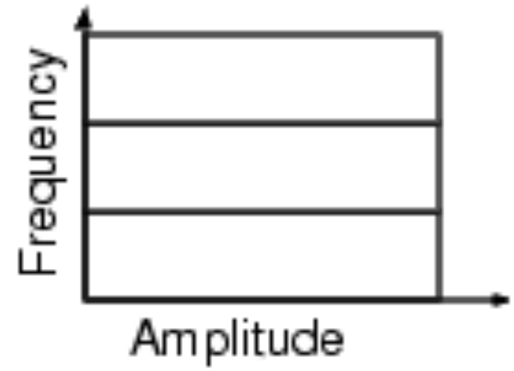


http://users.math.uni-potsdam.de/~hols/DFG1114/projectseis.html

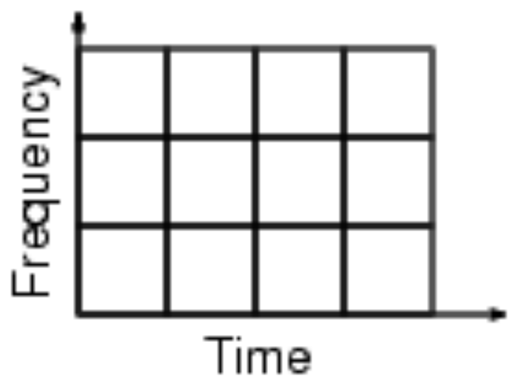
Graphical comparison



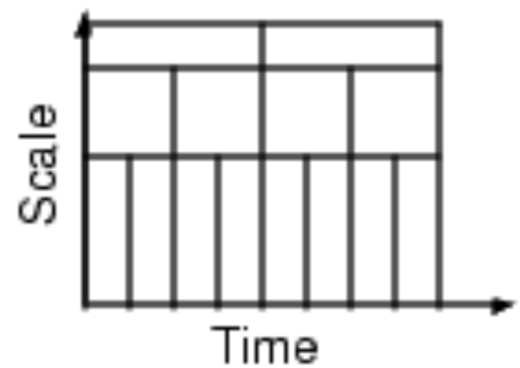
Time Domain (Shannon)



Frequency Domain (Fourier)



STFT (Gabor)



Wavelet Analysis

Summary

- Filtering is not necessarily straight forward, even the fundamental operations (LP, HP, BP, etc) require some thinking before application to data.
- The form of the filter decides upon the changes to the waveforms of the time series you are filtering
- For seismological applications filtering might drastically influence observables such as travel times or amplitudes
- „Windowing“ the signals in the right way is fundamental to obtain the desired filtered sequence

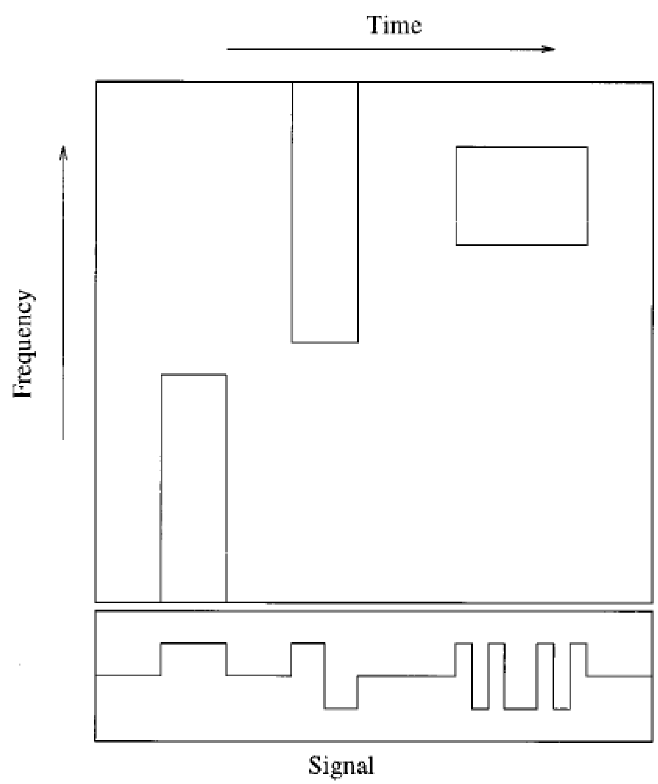


Figure 3. Schematic of time-frequency plane decomposition using wavelet packets.