Applied Time Series Analysis
SS 2014 – Week 12

Marcel Dettling

Institute for Data Analysis and Process Design

Zurich University of Applied Sciences

marcel.dettling@zhaw.ch

http://stat.ethz.ch/~dettling

ETH Zürich, May 12, 2014

Applied Time Series Analysis
SS 2014 – Week 12

Spectral Analysis

Idea: Time series are interpreted as a combination of cyclic components, and thus, a linear combination of harmonic oscillations.

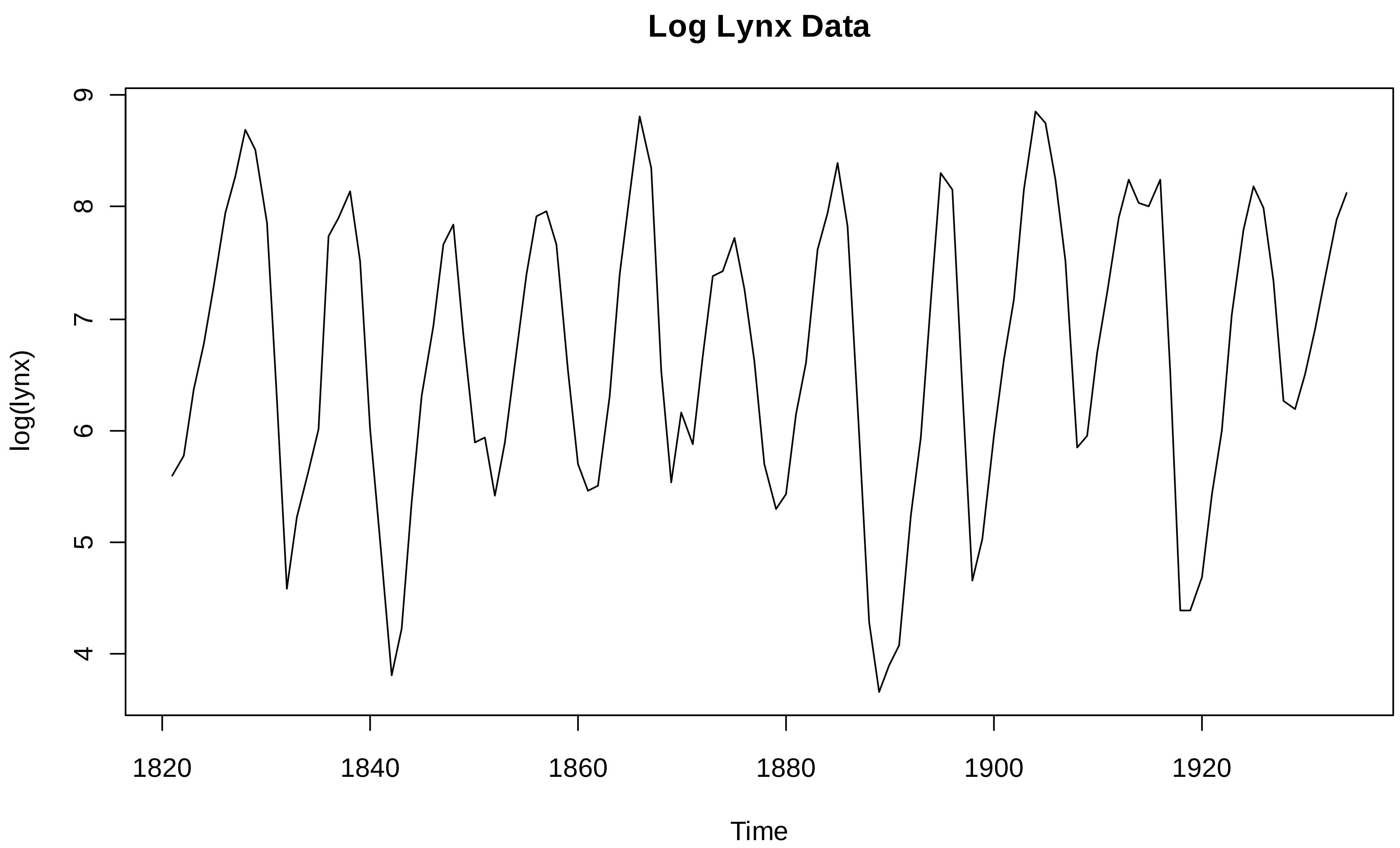
Why: As a descriptive means, showing the character and the dependency structure within the series.

What: It is in spirit, but also mathematically, closely related to the correlogram

Where:
- engineering
- economics
- biology/medicine

Applied Time Series Analysis
SS 2014 – Week 12

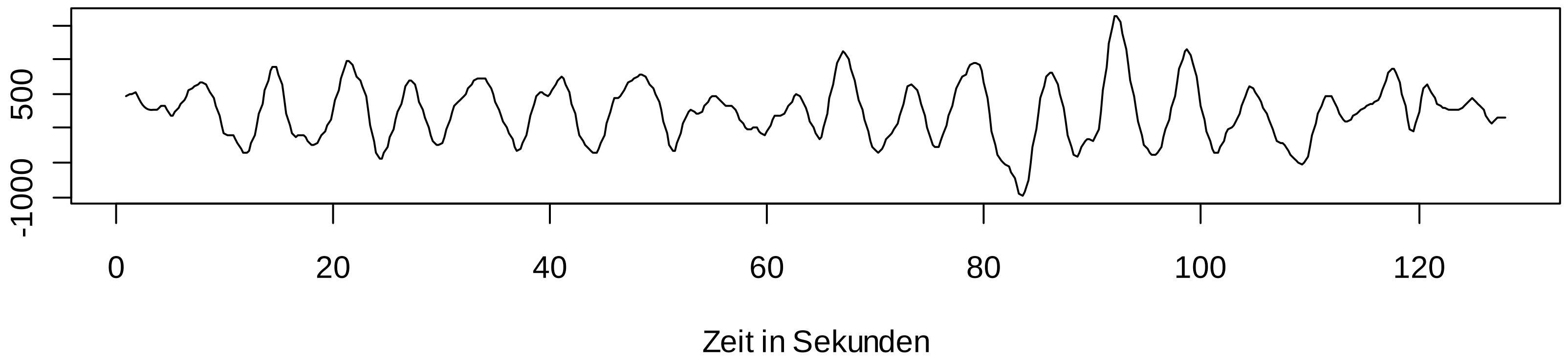
Lynx Data



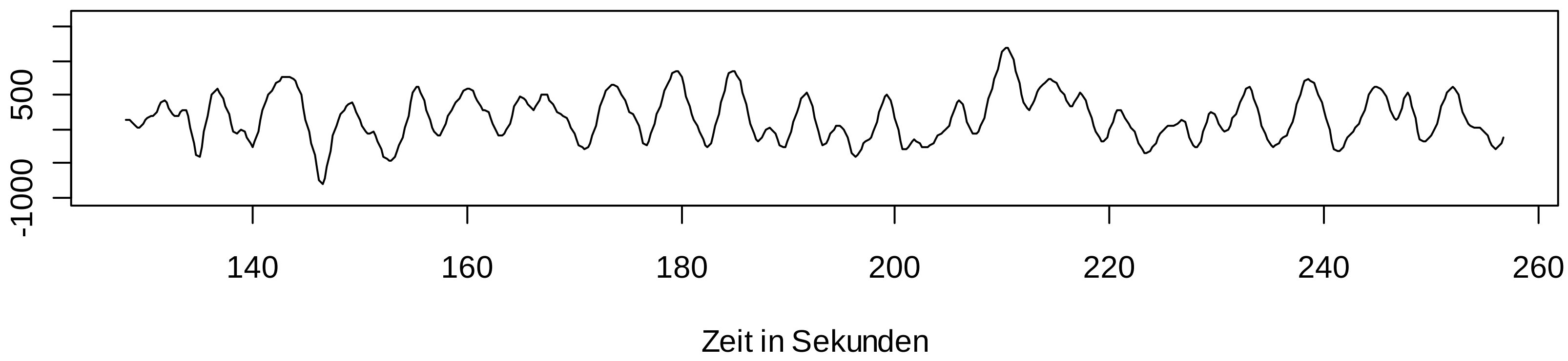
Applied Time Series Analysis
SS 2014 – Week 12

Ocean Wave Data

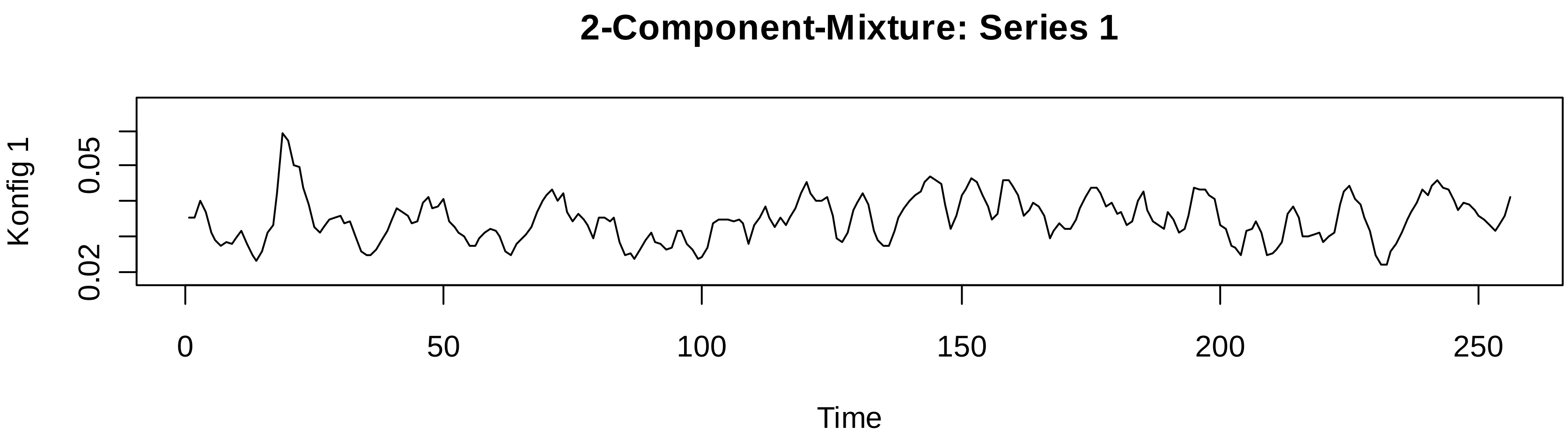
Ocean Wave Height Data, Part 1

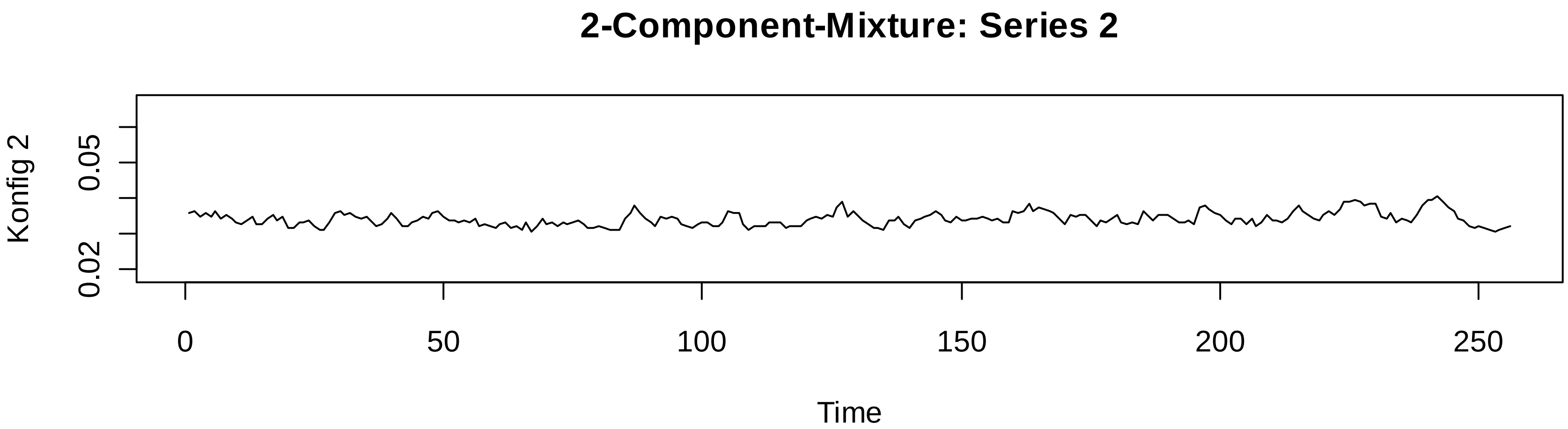


Ocean Wave Height Data, Part 2



2-Component-Mixture Data





Harmonic Oscillations

The most simple periodic functions are sine and cosine, which we will use as the basis of our analysis.

A harmonic oscillation has the following form:

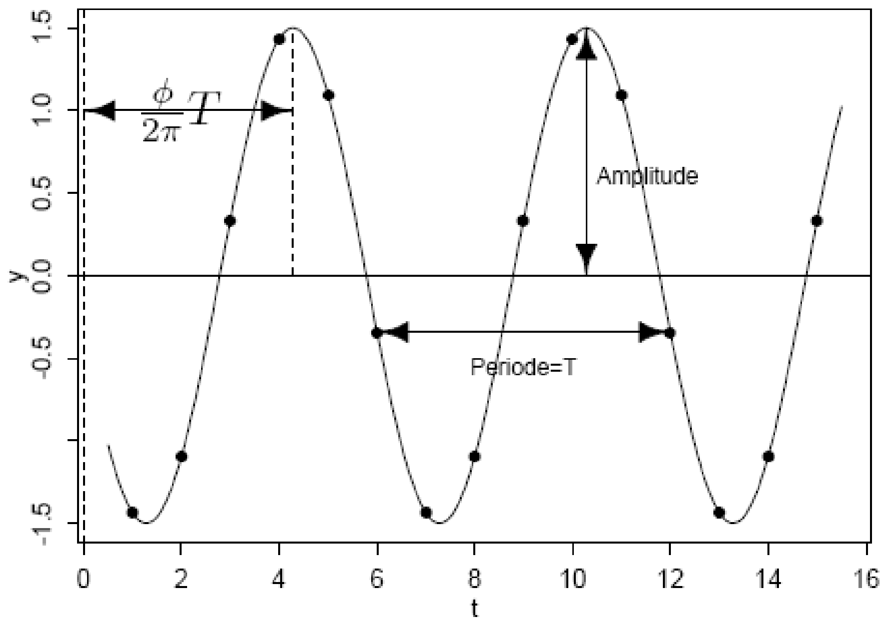
$$y(t) = \alpha \cos(2\pi\nu t) + \beta \sin(2\pi\nu t)$$

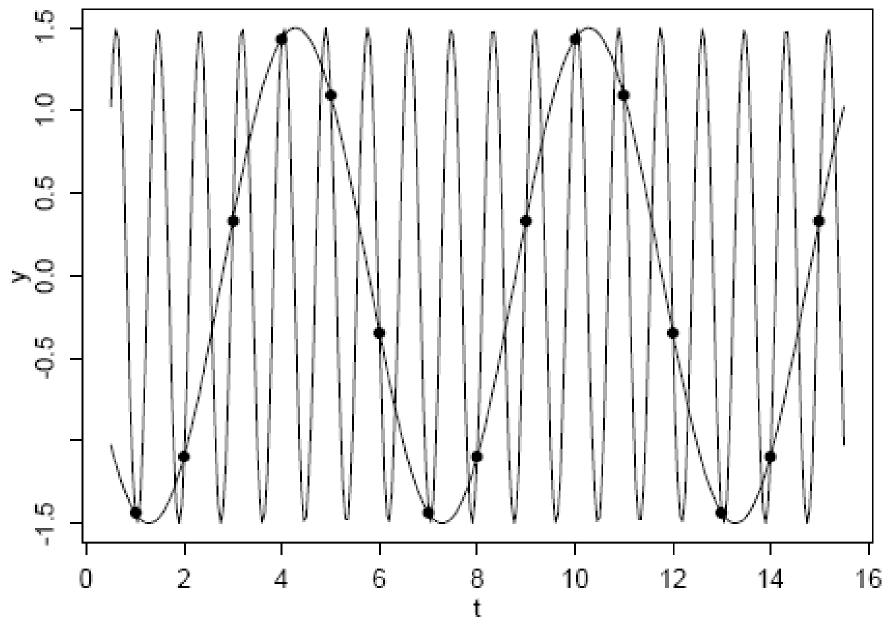
For the derivation, **see the blackboard…**

- In discrete time, we have aliasing, i.e. some frequencies cannot be distinguished (→ see next slide).
- The periodic analysis is limited to frequencies between 0 and 0.5, i.e. things we observe at least twice.

Applied Time Series Analysis
SS 2014 – Week 12

Aliasing





Regression Model & Periodogram

We try to write a time series with a regression equation containing sine and cosine terms at the fourier frequencies.

→ **see the blackboard**

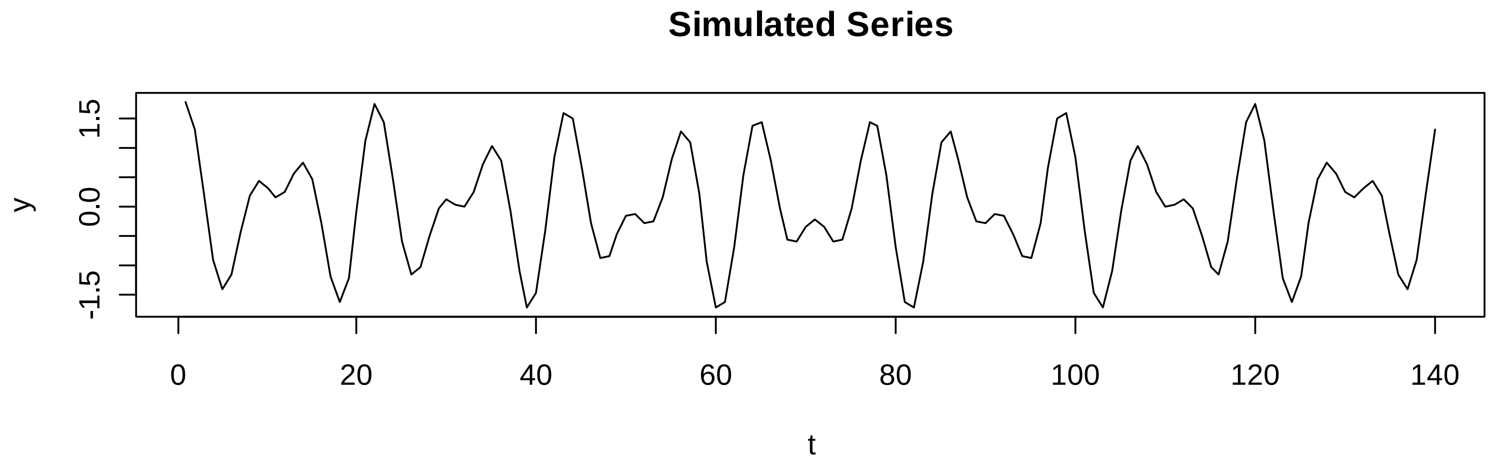
The most important frequencies within the series, which when omitted, lead to pronounced increase in goodness-of-fit.

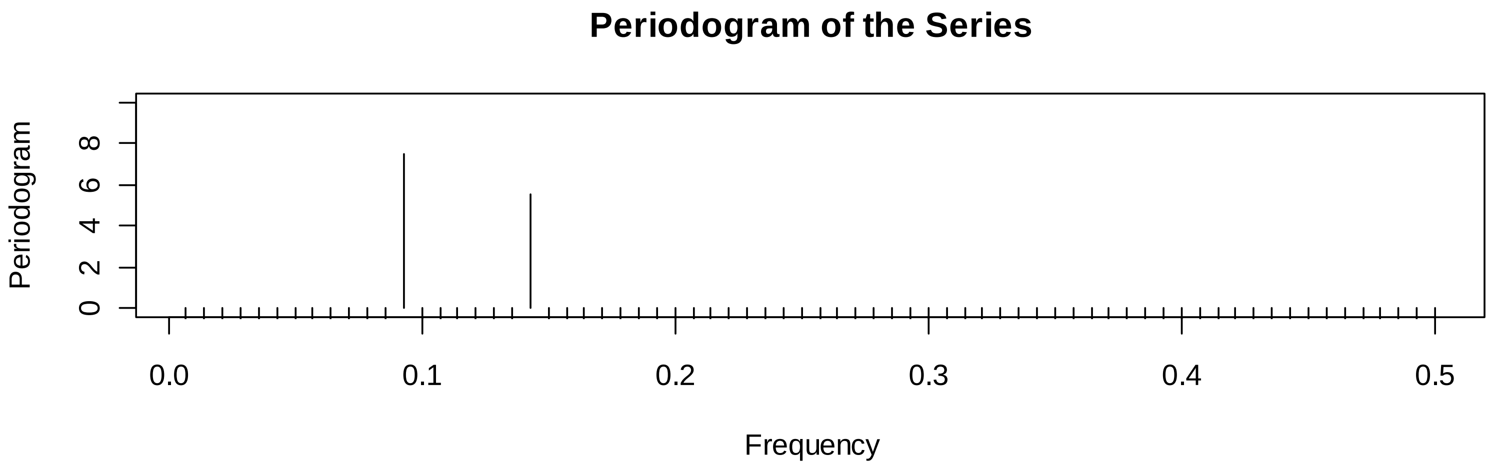
- This idea is used as a proxy for the periodogram,
  → **see the blackboard…**

- However, if the „true“ frequency is not a fourier frequency, we have leakage (→ see next 2 slides).

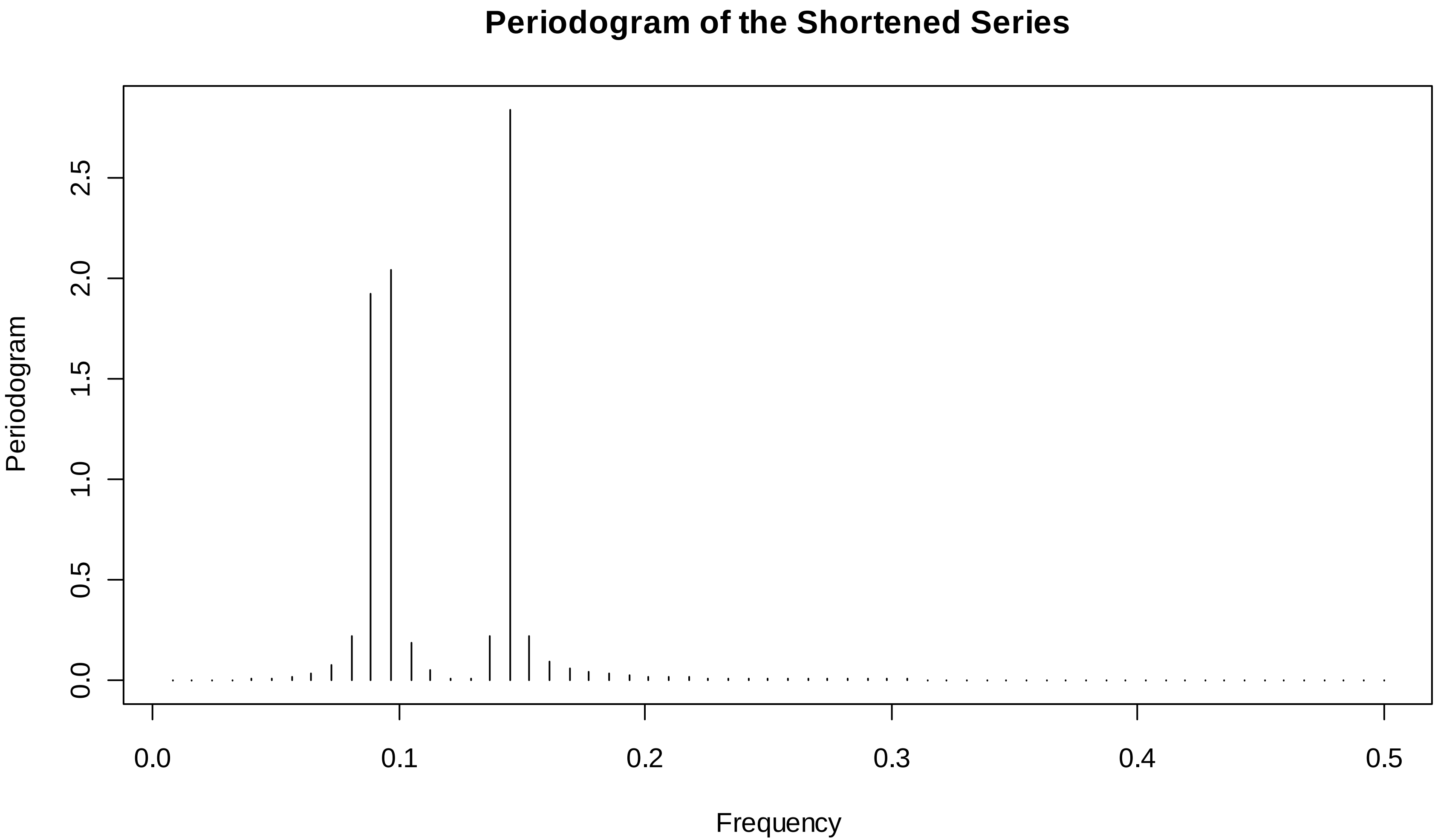
Applied Time Series Analysis
SS 2014 – Week 12

Periodogram of a Simulated Series





Periodogram of the Shortened Series



Properties of the Periodogram

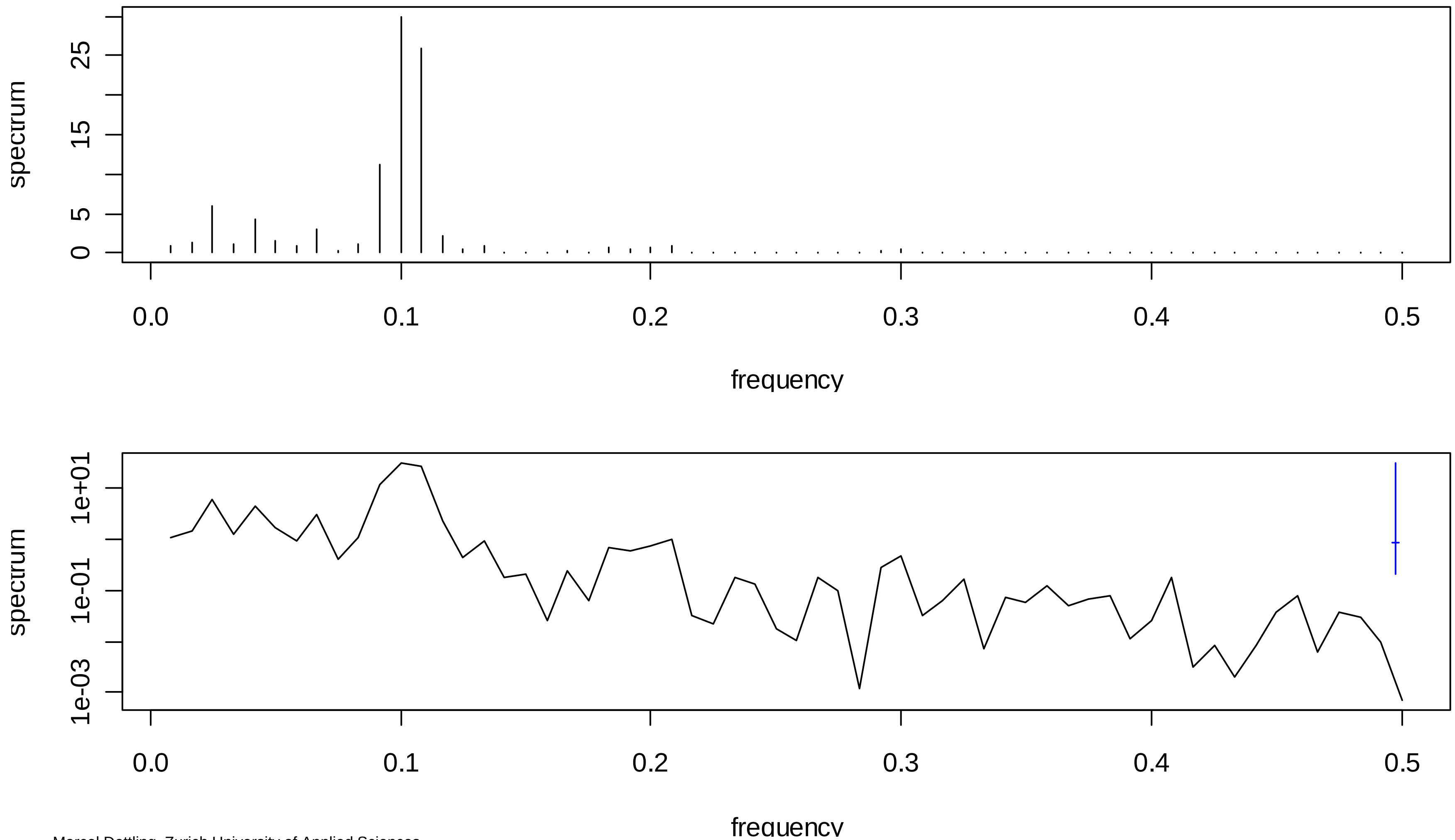
Periodogram and correlogram are mathematically equivalent, the former is the **fourier transform** of the latter.

→ **see the blackboard** for the derivation

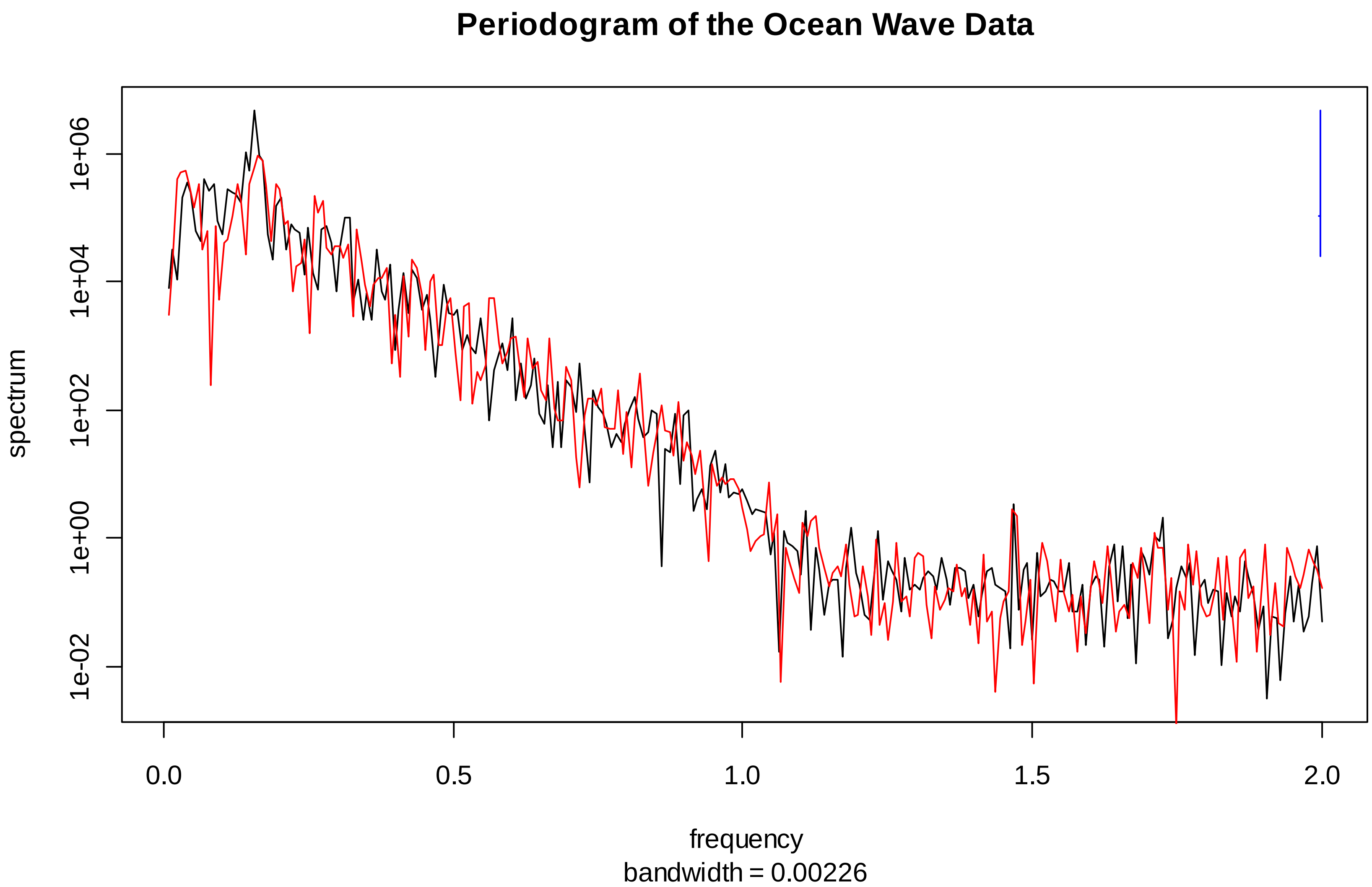
Note: this is a reason why we divided by 1/n in the ACV.

- $I(\nu_k)$ or $\log(I(\nu_k))$ are plotted against k/n
- Estimates seem rather instable and noisy
- On the log-scale, most frequencies are present
- It seems as if smoothing is required for interpretation.

Periodogram of the Log Lynx Data

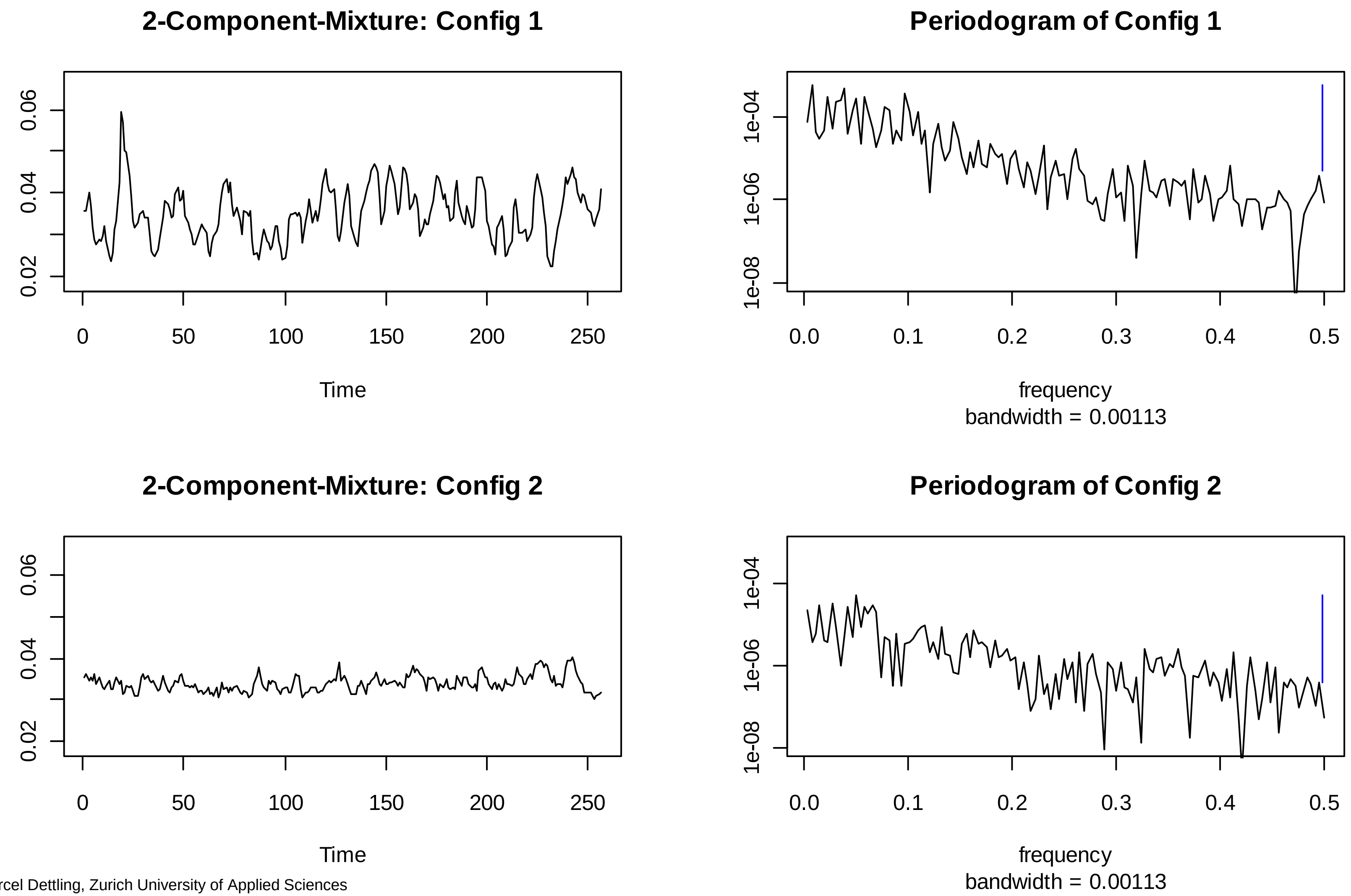


Periodogram of the Ocean Wave Data



Applied Time Series Analysis
SS 2014 – Week 12

Periodogram of the 2-Component-Mixture



The Spectrum

Observed time series	⇔	Stochastic process
Empirical ACF	⇔	Theoretical ACF
Periodogram	⇔	Spectrum

There is a link between ACF and periodogram/spectrum

$$f(\nu) = \sum_{k=-\infty}^{+\infty} \gamma(k) \cos(2\pi\nu k)$$

and

$$\gamma(k) = \int_{-0.5}^{+0.5} f(\nu) \cos(2\pi\nu k) d\nu$$

respectively. The spectrum is thus the Fourier transformation of the ACV.

What's the Spectrum Good For?

Theorem: **Cramer Representation**

Every stationary process can be written as the limit of a linear combination consisting of harmonic oscillations with random, uncorrelated amplitudes.

- The spectrum characterizes the variance of all these random amplitudes.
- Or vice versa: $\int_{\nu_1}^{\nu_2} f(\nu)d\nu$ is the variance between the frequencies that make the integration limits.
- The spectrum takes only positive values. Thus, not every ACF sequence defines a stationary series.

Applied Time Series Analysis
SS 2014 – Week 12

A Few Particular Spectra

- **White noise**
  → the spectrum is constant over all frequencies.

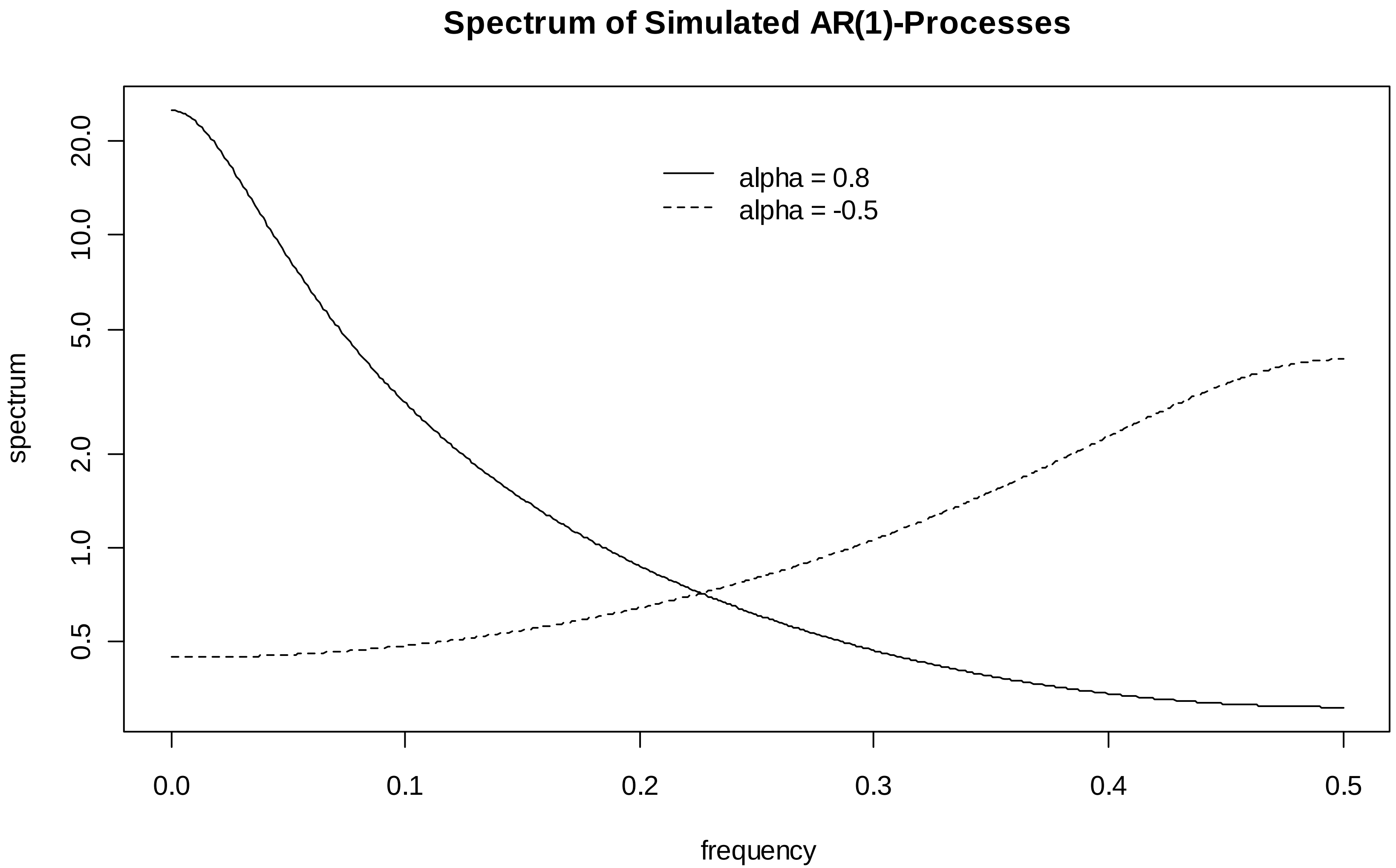
- **AR(1),** see next slide
  → already quite a complicated function α_1

- **ARMA (p,q)**
  → the characteristic polynoms determine the spectrum

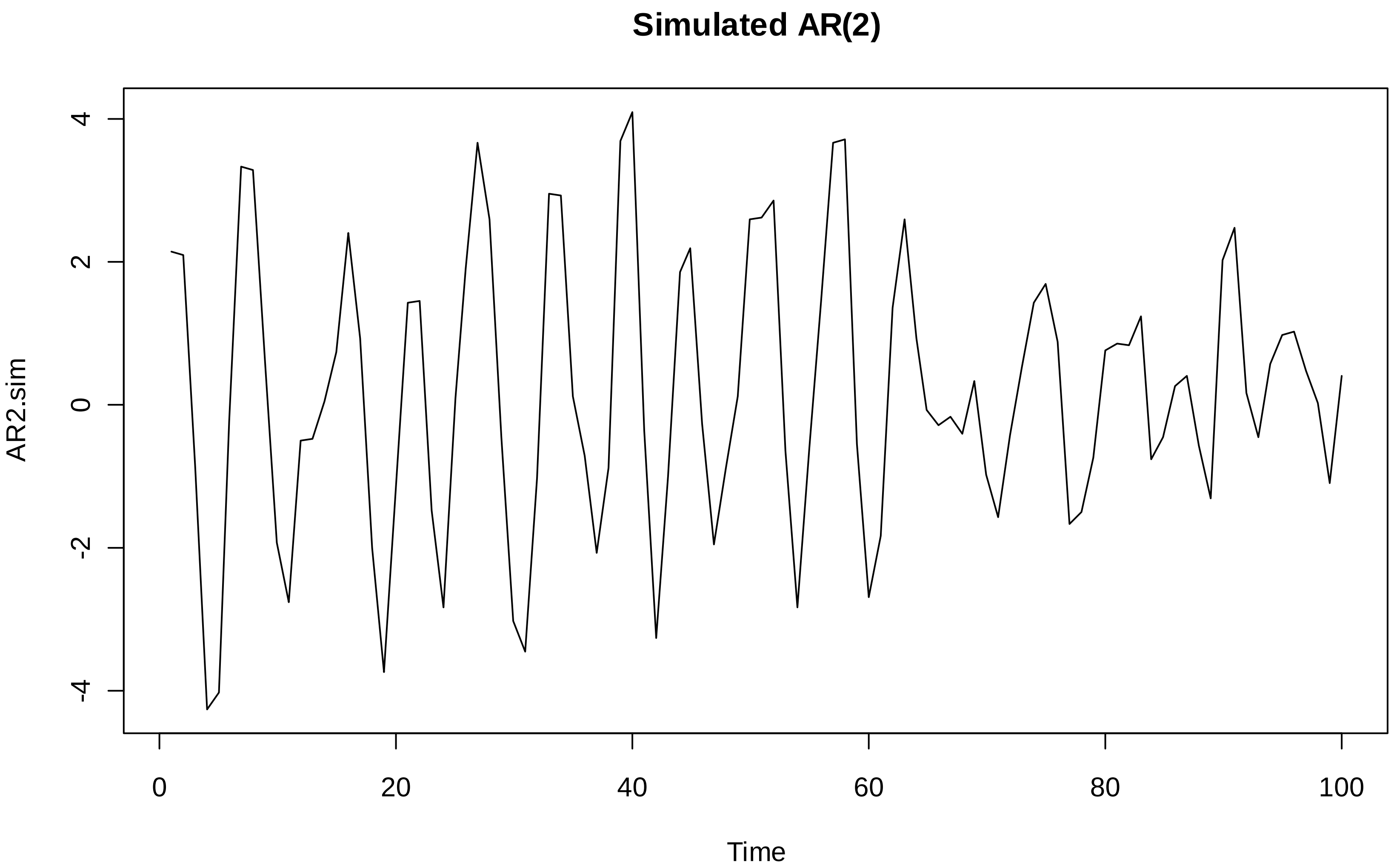
$$f(\nu) = \sigma_E^2 \frac{|\Theta(\exp(-i2\pi\nu))|}{|\Phi(\exp(-i2\pi\nu))|}$$

- Note: to generate m maxima in the spectrum, we require an AR-model, where the order is at least $2m$.

Spectrum of AR(1)-Processes

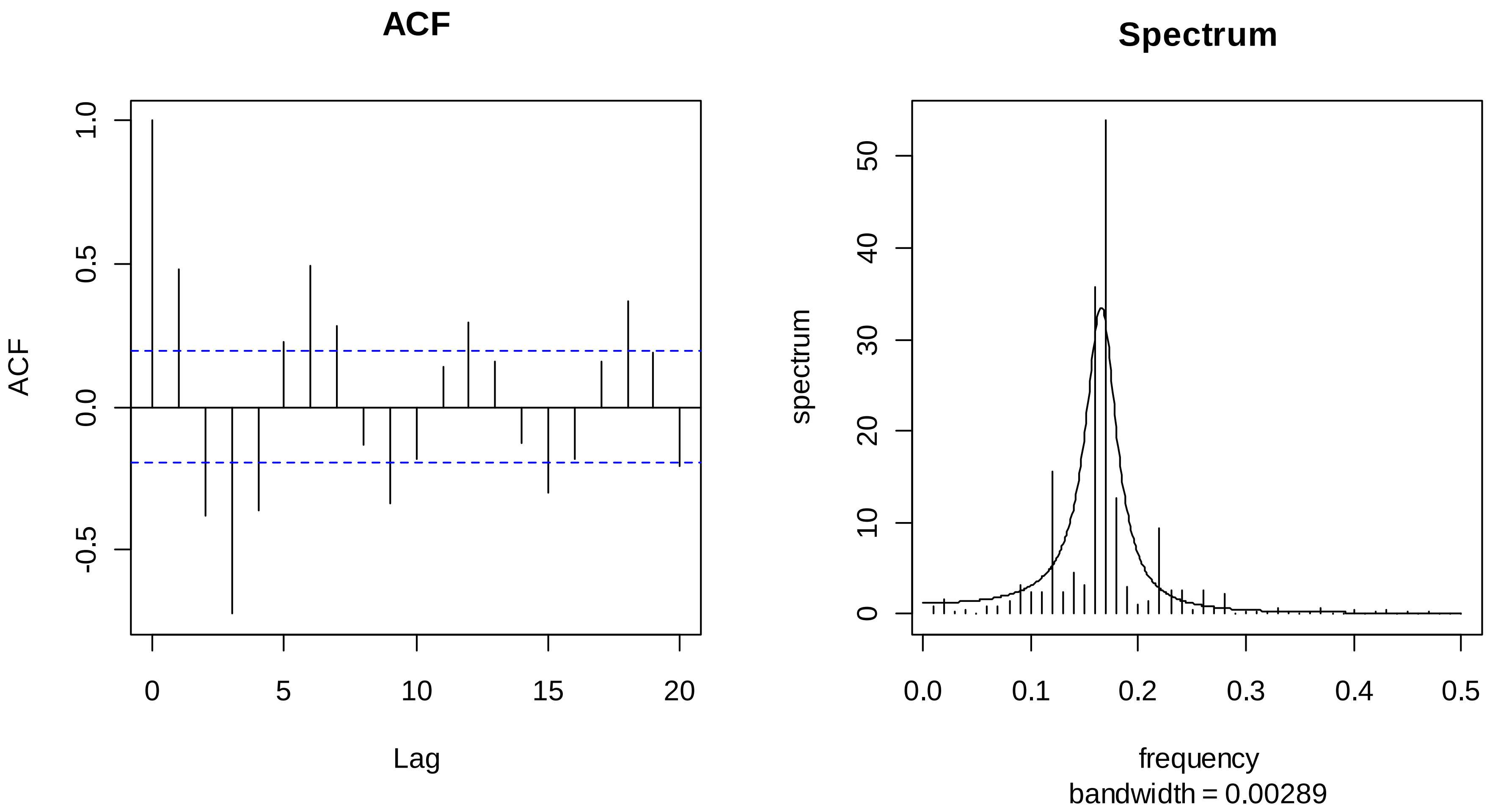


Simulated AR(2)-Process



Applied Time Series Analysis
SS 2014 – Week 12

ACF/Spectrum of Simulated AR(2)-Process



Applied Time Series Analysis
SS 2014 – Week 12

Spectral Analysis

- **Spectral analysis** is a descriptive technique, where the time series is interpreted as a linear combination of harmonic oscillations.

- The **periodogram** shows empirically, which frequencies are „important“, i.e. lead to a substantial increase in RSS when ommitted from the linear combination.

- The **spectrum** is the theoretical counterpart to the periodogram. It can also be seen as the Fourier transformation of the theoretical autocovariances.

- The periodogram is a poor **estimator** for the spectrum: it‘s not smooth and inconsistent.

Improving the Raw Periodogram

1) **Smoothing with a running mean**
   + simple approach
   - choice of the bandwith

2) **Smoothing with a weighted running mean**
   + choice of the bandwith is less critical
   - difficulties shift to the choice of weights

3) **Weighted plug-in estimation**
   + weighted Fourier trsf. of estimated autocovariances
   - choice of weights

4) **Piecewise periodogram estimation with averaging**
   + can serve as a check for stationarity, too

Improving the Raw Periodogram

5) Spectrum of an estimated model

- \+ fundamentally different from 1)-4)
- \- only works for „small“ orders p

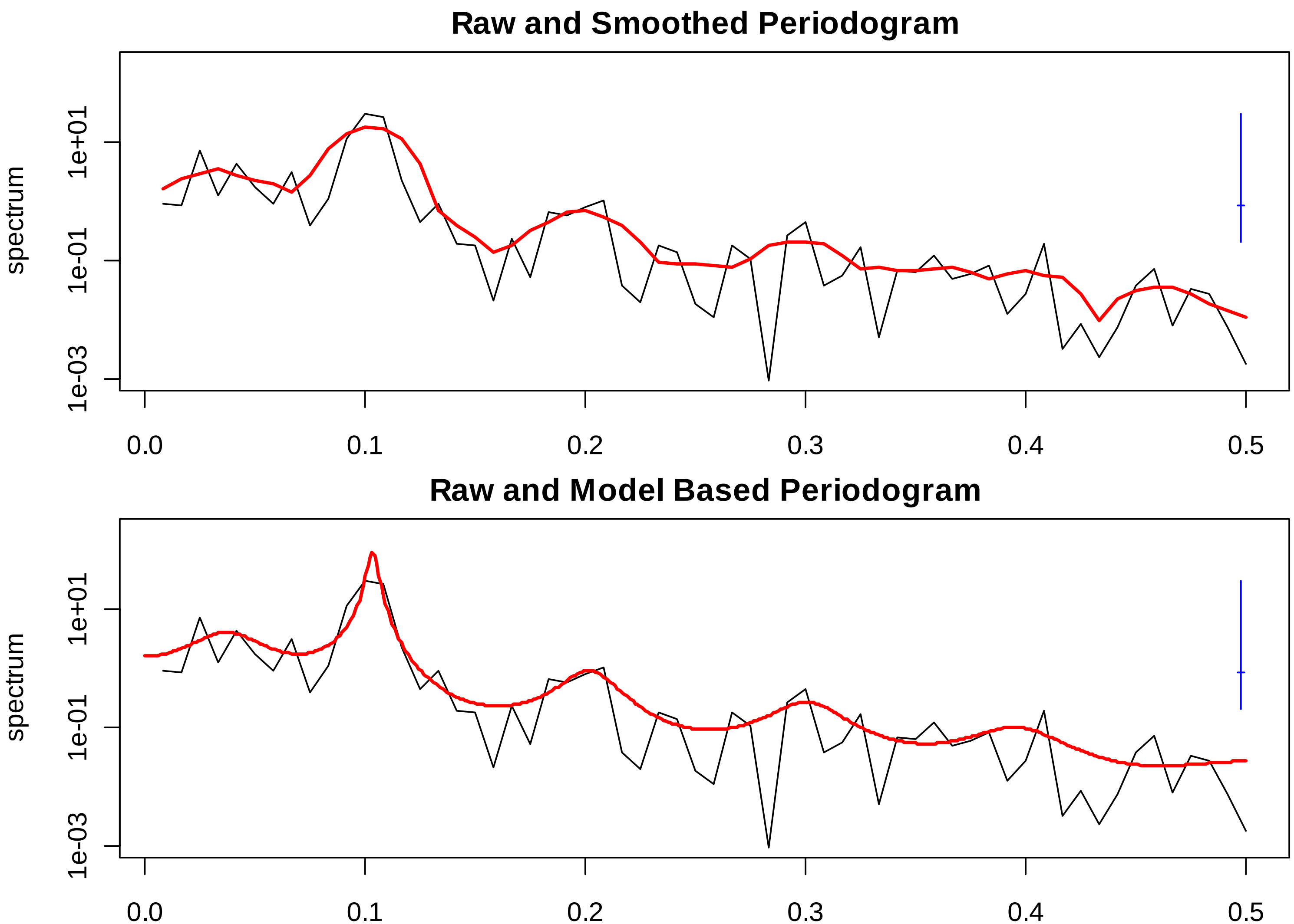
6) Tapering

- \+ further modification of periodogram estimation
- \+ reduces the bias in the periodogram
- \+ should always be applied

7) Prewhitening and Rescoloring

- \+ model fit and periodogram estimation on residuals
- \+ the effect of the model will be added again

Modified Periodogram of log(Lynx) Data



Applied Time Series Analysis
SS 2014 – Week 12

Modified Periodogram of log(Lynx) Data