Statistics 726	Assignment 4	Solutions

1. The easy way to approach these problems is to use the fact that if $Y(t)$ is a filtered version of $X(t)$ and the transfer function of the filter is $A(\lambda)$ then
$$f_{YY}(\lambda) = |A(\lambda)|^2 f_{YY}(\lambda).$$
In the examples below, the series of interest is related to white noise by a simple filtering relationship.

   (a) Here $Y(t)$ is produced from white noise by applying a differencing filter. The differencing filter has impulse response defined by
$$a(u) = \begin{cases} 1 & u = 0, \\ -1 & u = 1, \\ 0 & \text{otherside}, \end{cases}$$
and transfer function defined by
$$\begin{aligned} A(\lambda) &= \sum_{u=0}^{\infty} a(u) e^{-i\lambda u} \\ &= e^{-i\lambda 0} - e^{-i\lambda 1} \\ &= 1 - e^{-i\lambda} \\ &= e^{-i\lambda/2} \left(e^{i\lambda/2} - e^{-i\lambda/2}\right) \\ &= e^{-i\lambda/2}\, 2 \sin \lambda/2. \end{aligned}$$
This means that
$$|A(\lambda)|^2 = 4 \sin^2 \lambda/2,$$
and so
$$f_{YY}(\lambda) = \frac{2\sigma^2}{\pi} \sin^2 \lambda/2.$$

   (b) A seasonal differencing filter (with seasonal period s) has impulse response
$$a(u) = \begin{cases} 1 & u = 0, \\ -1 & u = s, \\ 0 & \text{otherside}. \end{cases}$$

By similar reasoning to part (a), the transfer function of this filter is

$$A(\lambda) = e^{-i\lambda s/2}\, 2\sin \lambda s/2.$$

(You should try graphing $|A(\lambda)|$ for $s = 12$.) This means that the power spectrum of seasonally differenced noise is

$$f_{YY}(\lambda) = \frac{2\sigma^2}{\pi}\ \sin^2 \lambda s/2.$$

(c) From the notes, the spectrum of an MA(1) series is

$$f_{XX}(\lambda) = \frac{\sigma^2}{2\pi}(1 + \theta^2 + 2\theta\cos\lambda)$$

and so the spectrum for a differenced MA(1) series must be

$$f_{YY}(\lambda) = \frac{2\sigma^2}{\pi}\ (1 + \theta^2 + 2\theta\cos\lambda)\ \sin^2 \lambda s/2.$$

(d) The general ARMA(p, q) series is defined by

$$(1 - \phi_1 L^1 - \cdots - \phi_p L^p)Y(t) = (1 + \theta_1 L^1 + \cdots + \theta_q)\varepsilon.$$

This is a statement about the relationship between two filtered series:

$$\mathcal{A}[Y](t) = \mathcal{B}[\varepsilon](t).$$

In the frequency domain this means that

$$|A(\lambda)|^2 f_{YY}(\lambda) = |B(\lambda)|^2 f_{\varepsilon\varepsilon}(\lambda)$$

or

$$f_{YY}(\lambda) = \frac{|B(\lambda)|^2}{|A(\lambda)|^2} f_{\varepsilon\varepsilon}(\lambda).$$

The individual transfer functions are:

$$A(\lambda) = 1 - \phi_1 e^{-i\lambda} - \cdots - \phi_p e^{-i\lambda p}$$

and

$$B(\lambda) = 1 + \theta_1 e^{-i\lambda} + \cdots + \theta_q e^{-i\lambda q}$$

so that

$$f_{YY}(\lambda) = \frac{\sigma^2}{2\pi}\ \frac{|1 - \phi_1 e^{-i\lambda} - \cdots - \phi_p e^{-i\lambda p}|^2}{|1 + \theta_1 e^{-i\lambda} + \cdots + \theta_q e^{-i\lambda q}|^2}.$$

Or, in terms of the AR polynomial $\theta(z)$ and MA polynomial $\phi(z)$,

$$f_{YY}(\lambda) = \frac{\sigma^2}{2\pi}\ \frac{|\phi(e^{-i\lambda})|^2}{|\theta(e^{-i\lambda})|^2}.$$

2. Figure 1 shows the European wheat price series. The variability of the series is clearly related to its mean level. This dependence can be removed by taking logs (this is very common for price series). The logged series is shown in figure 2.

   The logged series is clearly nonstationary, so before estimating the power spectrum it is important to remove the trend apparent in the series. If this is not done, the trend will appear in the spectrum as a "peak" in the spectrum close to frequency 0.

   There are a number of possible ways to remove the trend. I've chosen to do it by fitting a smooth curve throught the data will lowess and substructing the curve out. I used a smoothing fraction of 0.2 so that the smooth follows the data reasonably closely. The code used to do this was:

   ```
   lw = log(wheat)
   dtlw = lw - lowess(lw, f = .3)$y
   ```

   and plots of the smooth and the residuals from the smooth are shown in figures 3 and 4.

   Figures 5 and 6 show spectrum estimates with spans of 3 and 5 respectively for the residual series. There are possible low frequency peaks at about 0.016, 0.062, 0.139, 0.435 and 0.477. The lowest frequency peak is clear cut, the others are not.

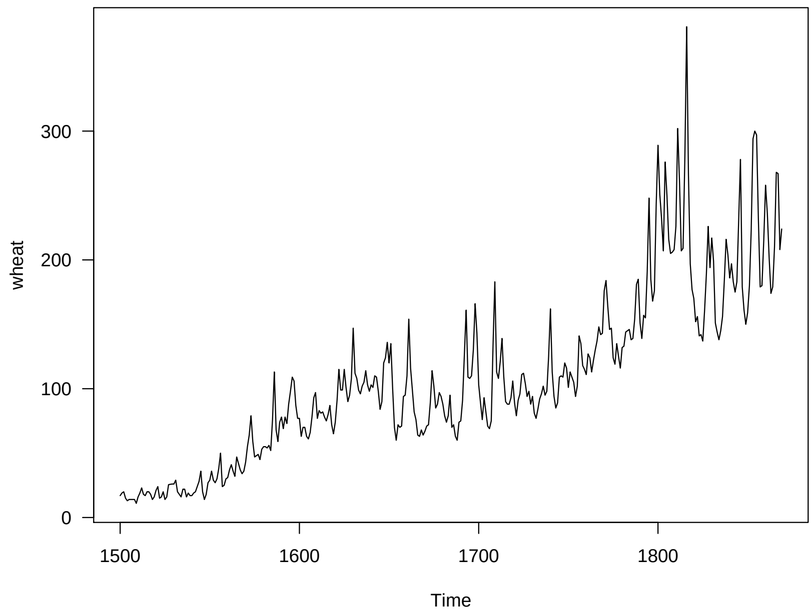

Figure 1: The European wheat index series.

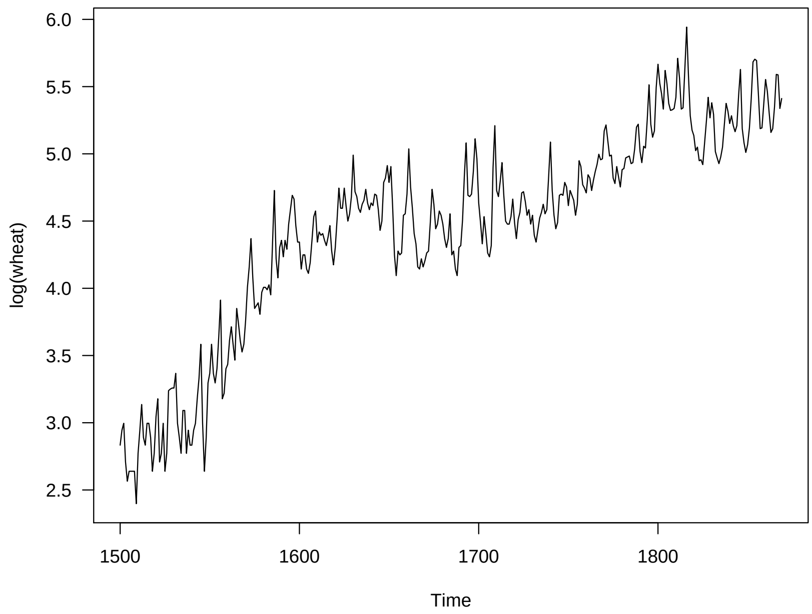

Figure 2: The logged European wheat index series.

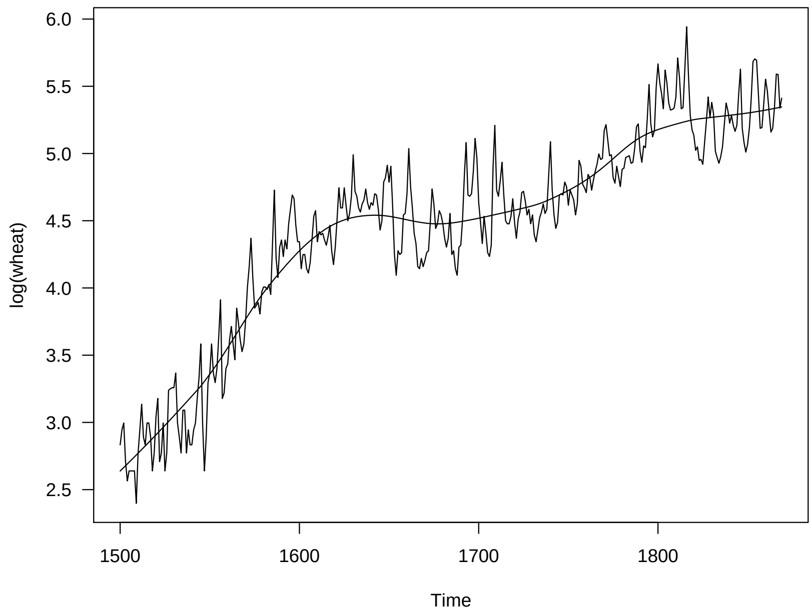

Figure 3: A lowess curve fitted to the logged European wheat index series.

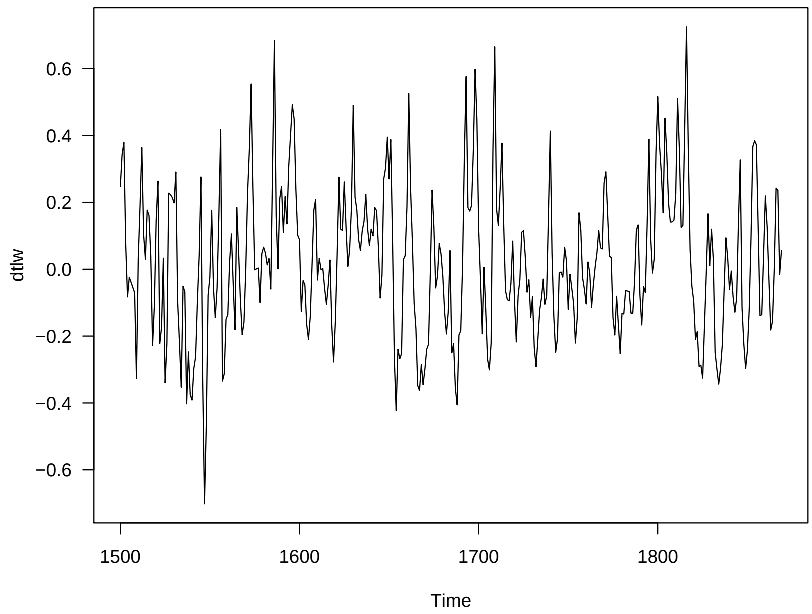

Figure 4: Residuals from the lowess curve fit.

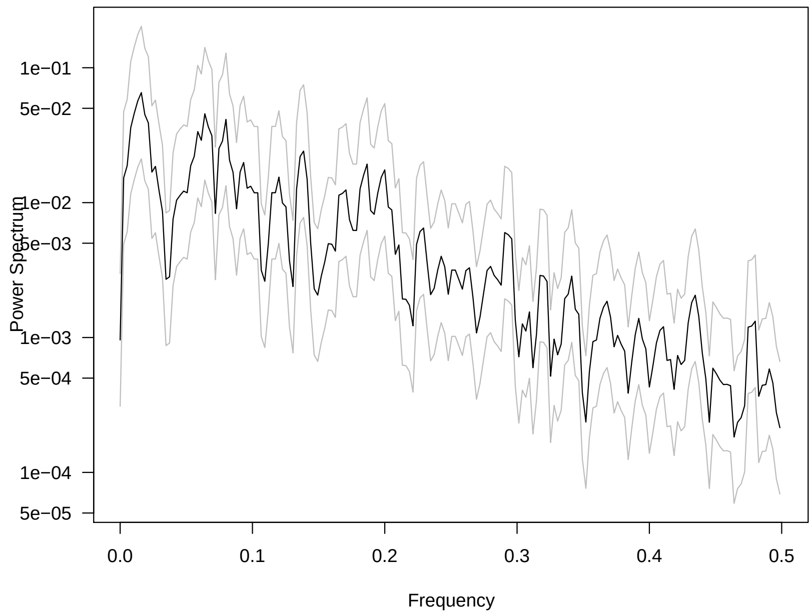

Figure 5: The power spectrum for the detrended wheat series with `span`=3.

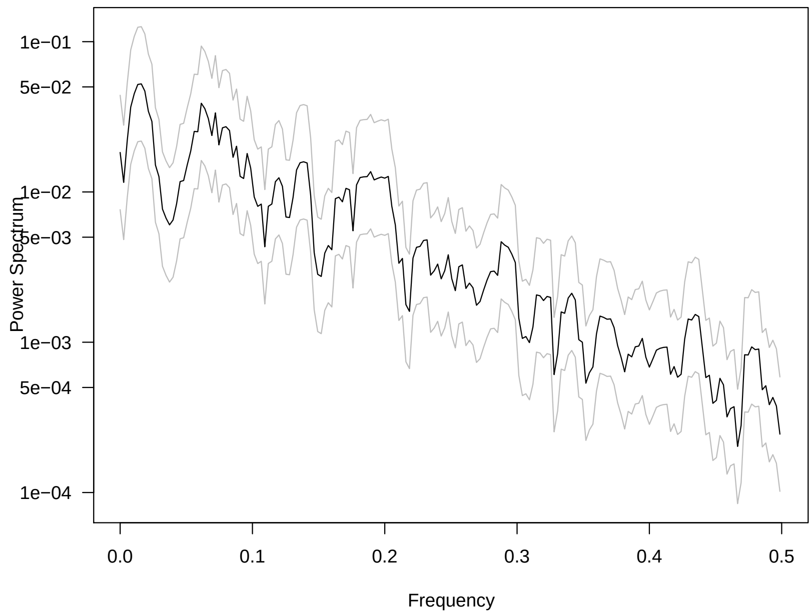

Figure 6: The power spectrum for the detrended wheat series, with `span`=5.

I did't expect you to go this far, but it is interesting to try some tapering on this series to see what effect it has. Tapering helps resolve peaks and so it may be helpful here. Figure 7 shows the estimate spectrum with 50% tapering and a span of 5.

Figure 8 shows spectrum estimate with 50% tapering and a span off 3. The gray lines show the confidence interval for a heavily smoothed estimate (span=21). The intent is to judge the significance of the peaks by whether or not they protrude above the confidence band. The peaks which do have been marked in the plot. They are at the following frequencies:

$$0.016, 0.061, 0.080, 0.165, 0.192, 0.228, 0.292, 0.320, 0.427, 0.480.$$

Obviously not all of these are really significant, but some of them bear further inspection.

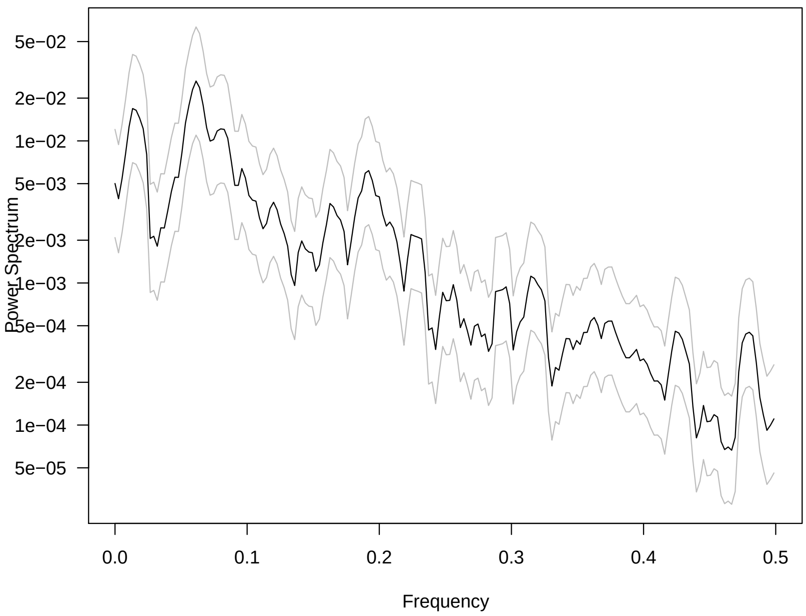

Figure 7: The power spectrum for the detrended wheat series, with 50% tapering and `span`=5.

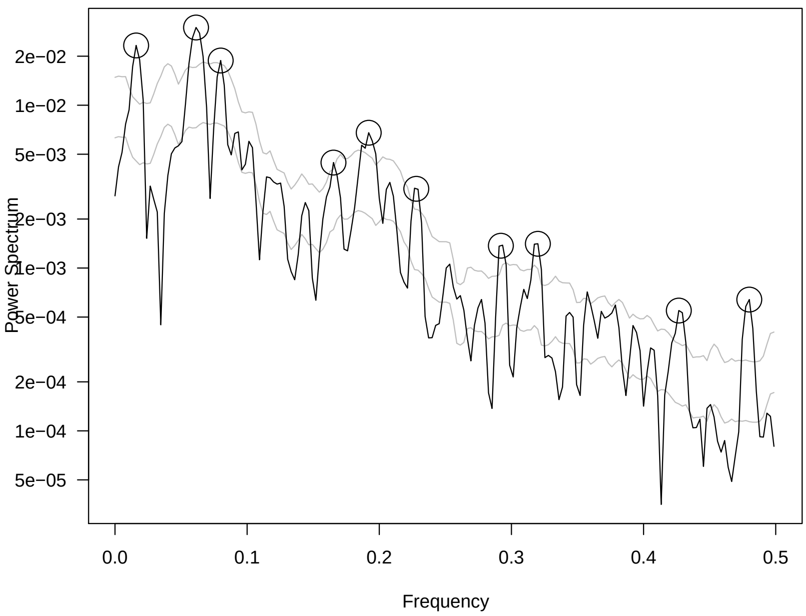

Figure 8: The power spectrum for the detrended wheat series.

3. (I'm not really answering the question here — that was pretty routine. There are some comments at the end of the answer. This is just having a bit of fun . . .)

   When doing spectral analysis on time series it is important to carry out smoothing with several different bandwidths or spans. Not to search for the "best" amount of smoothing, but because different amounts of smoothing will reveal different things about the series. In the case of the Berlin and Vienna temperatures, light smoothing will do the best job of showing off the peaks and heavier smoothing will show off the structure elsewhere in the spectrum.

   Figures 9 and 10 show power spectra for the Berlin and Vienna series computed with a smoothing span of 3. The series have also been tapered 5% to help make the peaks stand out.

   The spectra are very similar. Both show a very large peak corresponding to a yearly cycle (frequency = 1/12), with a hint of another peak at an even lower frequency of about 0.01, corresponding to a cycle of roughly 8 years. There is also evidence of a long-term trend in both series.

   The major difference between the two spectra is the presence of a six-monthly harmonic (frequency = 2/12) in the Vienna series. This corresponds to a real difference in termperature pattern in the two series; Vienna warms up a little more quickly than Berlin in springtime. The monthy average temperatures for the two cities, normalized to the interval [0,1] are shown in figure 11.

   The coherence fo the two series is shown in figure 12. The one-year and eight-year cycles in the two series appear to be coherent, but the long-term (frequency = 0) components are less coherent. This is also clear by looking at the series. For example, the long-term trend in Vienna has been toward lower summer temperatures, while summer temperature in Berlin have been fairly constant. (You can see this by applying the `monthplot` function to the two series.)

   Figures 13 and 14 show the gain and phase of the filter which best matches the Vienna series to the Berlin. Notice that the relative phases corresponding to the six-monthly components of the series are different. This again can be related to the differences shown in figure 11.

   Finally, the inpulse response for the best fitting filter is shown in figure 15. The degree of smoothing does not affect the impulse response as much as it does the frequency domain parameter.

   These basic features of the cross spectral analysis are fairly insensitive to the amount of smoothing used. This is particularly true of the impule response, which is very stable indeed.

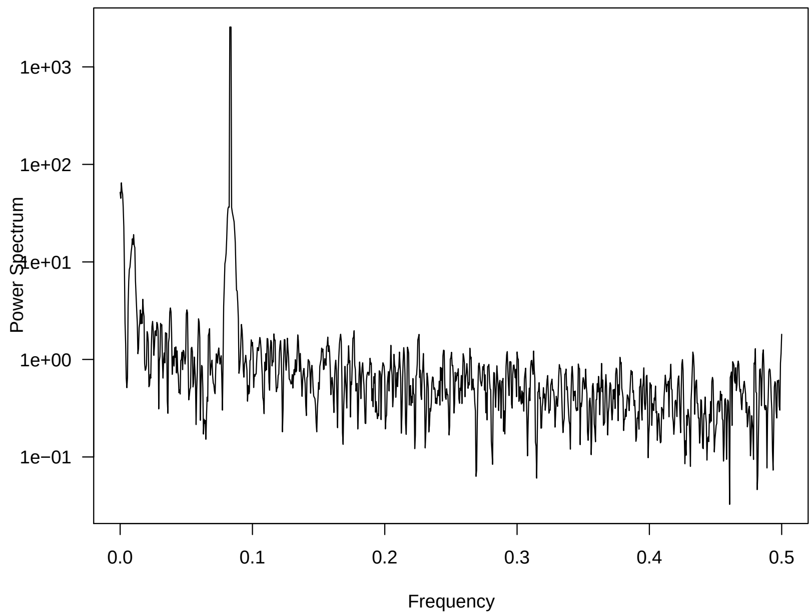

Figure 9: The spectrum for the Berlin temperatures computed using a smoothing span of 3.

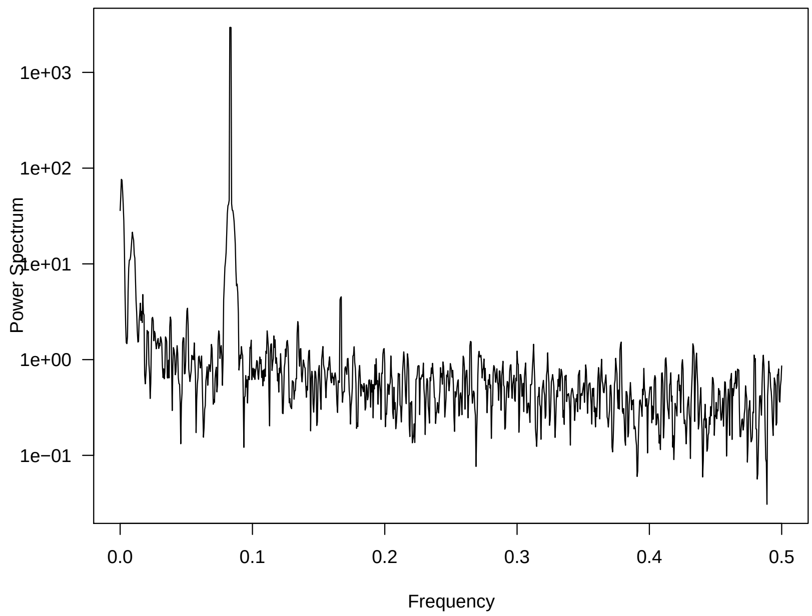

Figure 10: The spectrum for the Vienna temperatures computed using a smoothing span of 3.

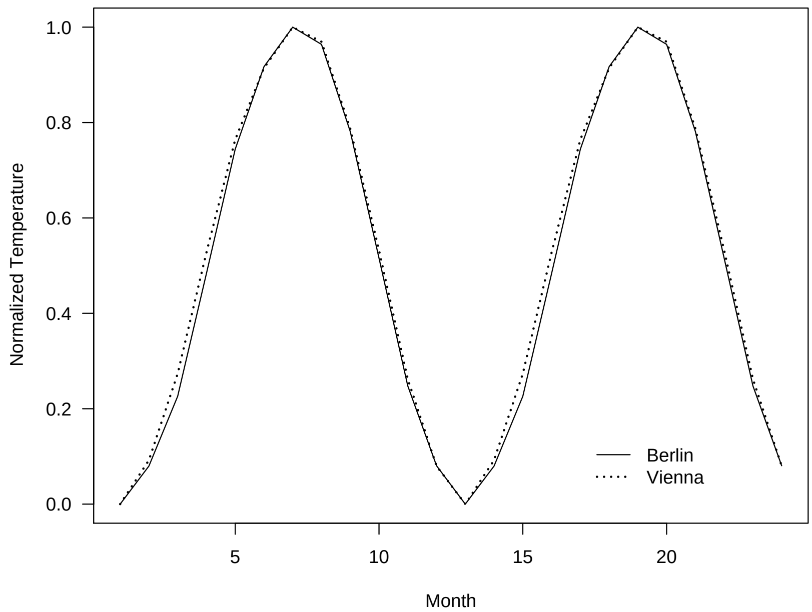

Figure 11: The seasonal pattern of temperatures in Berlin and Vienna.

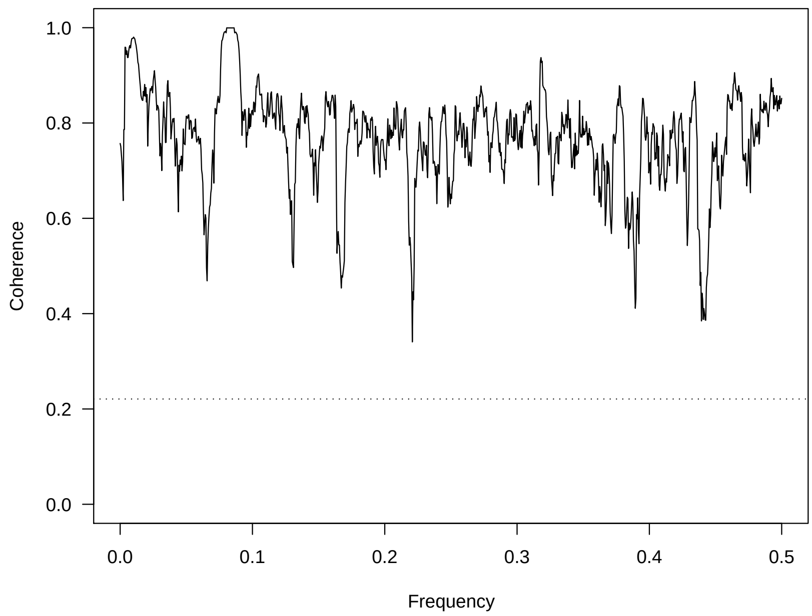

Figure 12: The coherence for the Berlin and Vienna temperatures computed using a smoothing span of 13.

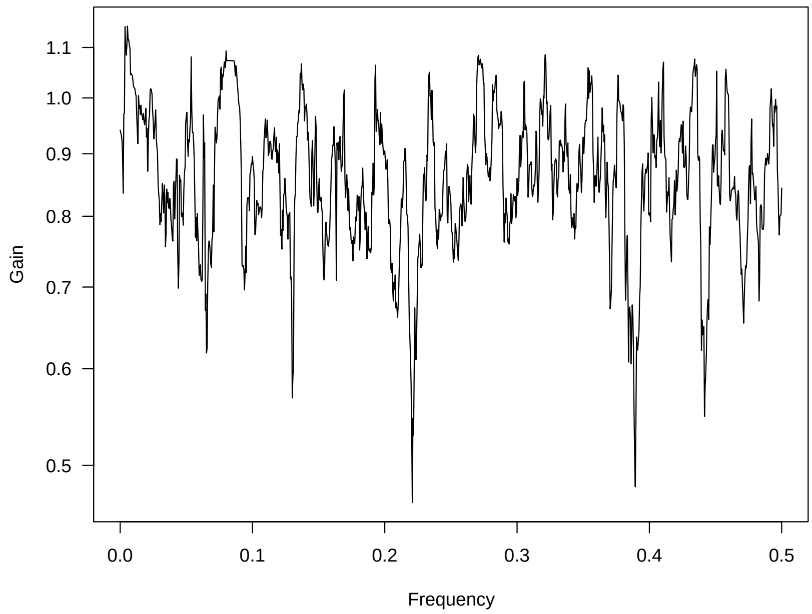

Figure 13: The gain for the Vienna series over the Berlin one (computed using a smoothing span of 13).

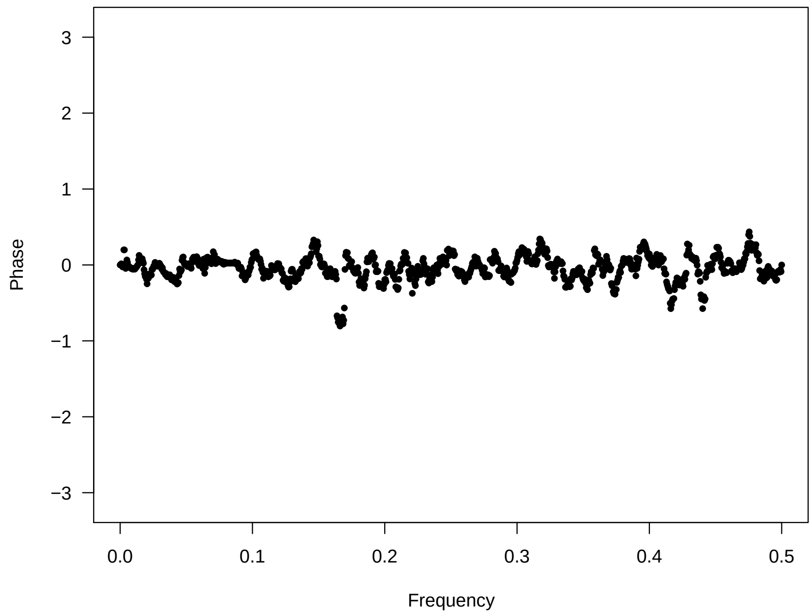

Figure 14: The gain for the Vienna with respect to the Berlin one (computed using a smoothing span of 13).

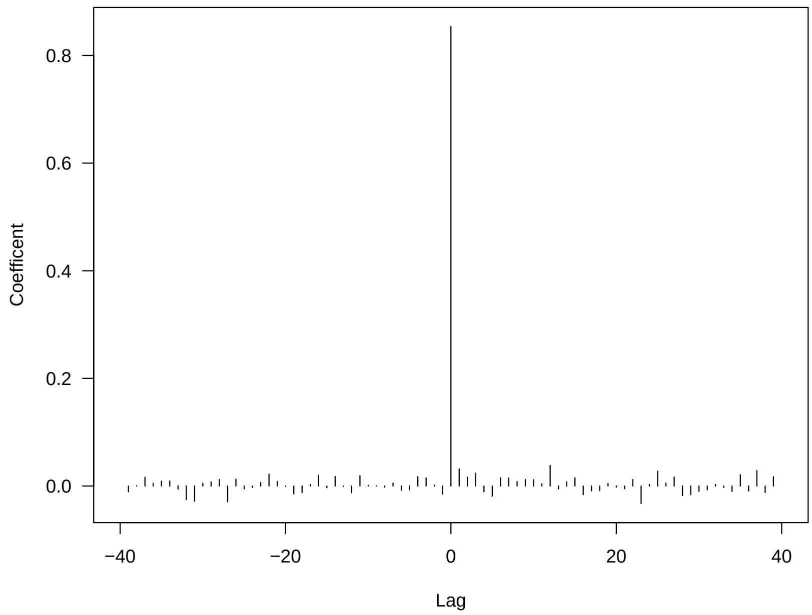

Figure 15: The impulse response for the best-fitting filter which matches the Vienna series to the Berlin one (computed using a smoothing span of 13).