ST565: Time Series HW7

Amirhosein "Emerson" Azarbakht *azarbaka@oregonstate.edu*

Question 1

1. Derive the spectrum for an MA(1) process.

$$\gamma(k) = \begin{cases} \beta_1^2\sigma^2 + \sigma^2 & \text{for k} = 0 \\ \beta_1\sigma^2 & \text{for k} = 1 \\ 0 & \text{for k} >= 2 \end{cases}$$

$$f(\omega) = 1/\pi\left[\gamma(0) + 2\Sigma_{k=1}^{\infty}\gamma(k)\cos(\omega k)\right]$$

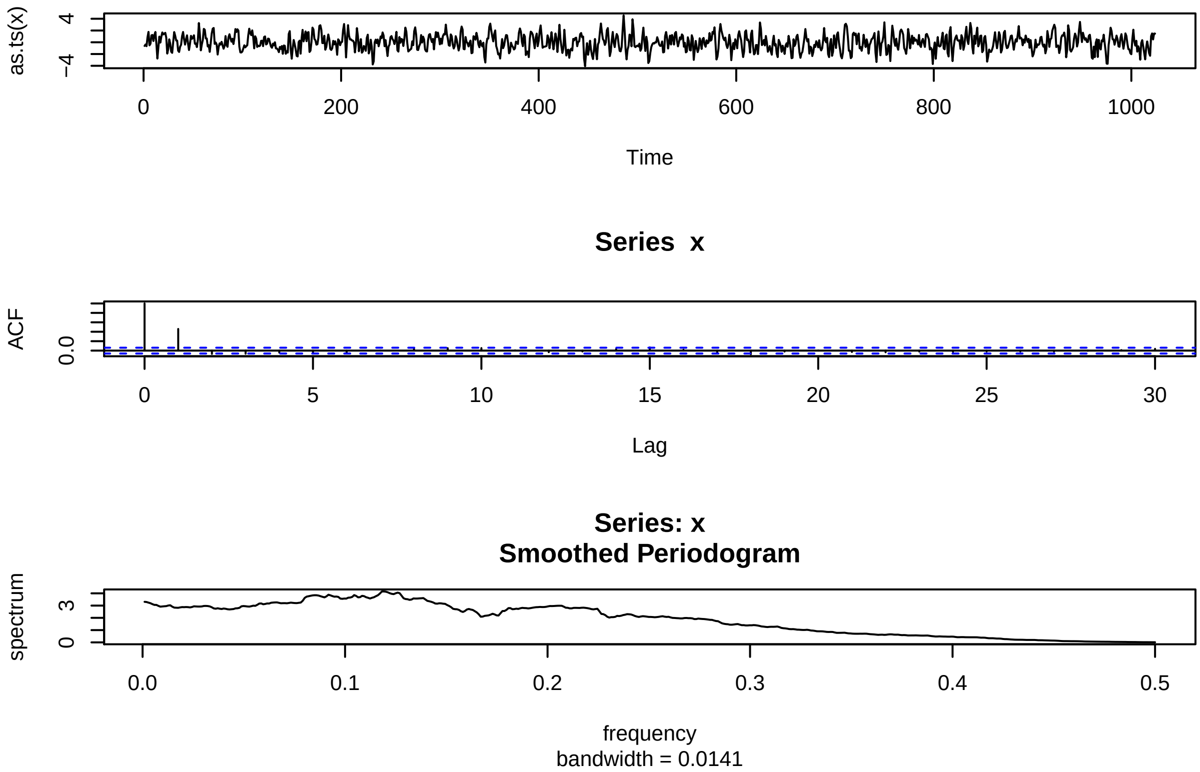
$$f(\omega) = 1/\pi\left[\beta_1^2\sigma^2 + \sigma^2 + 2[\beta_1\sigma^2]\cos(\omega)\right]$$

$$f(\omega) = 1/\pi\left[\beta_1^2\sigma^2 + \sigma^2 + 2\beta_1\sigma^2\cos(\omega)\right]$$

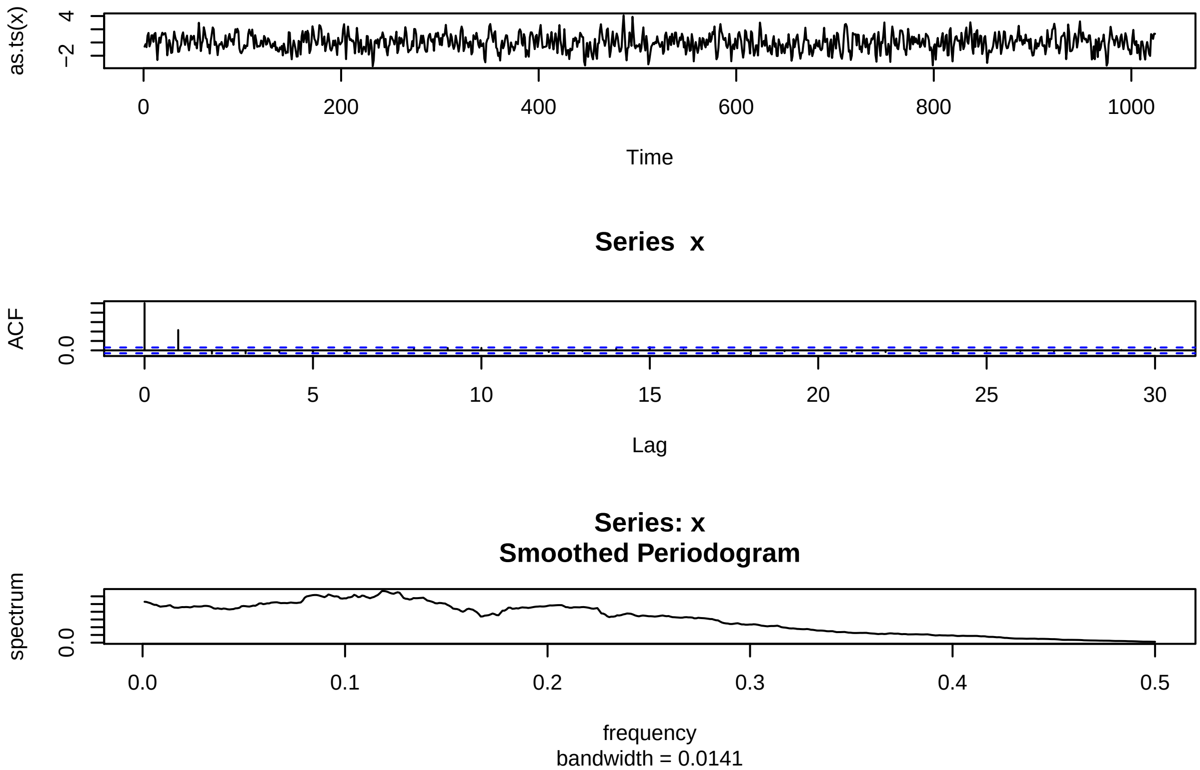
$$f(\omega) = \sigma^2/\pi\left[\beta_1^2 + 1 + 2\beta_1\cos(\omega)\right]$$

Produce a plot of the spectrum showing the shape for a few values of β.

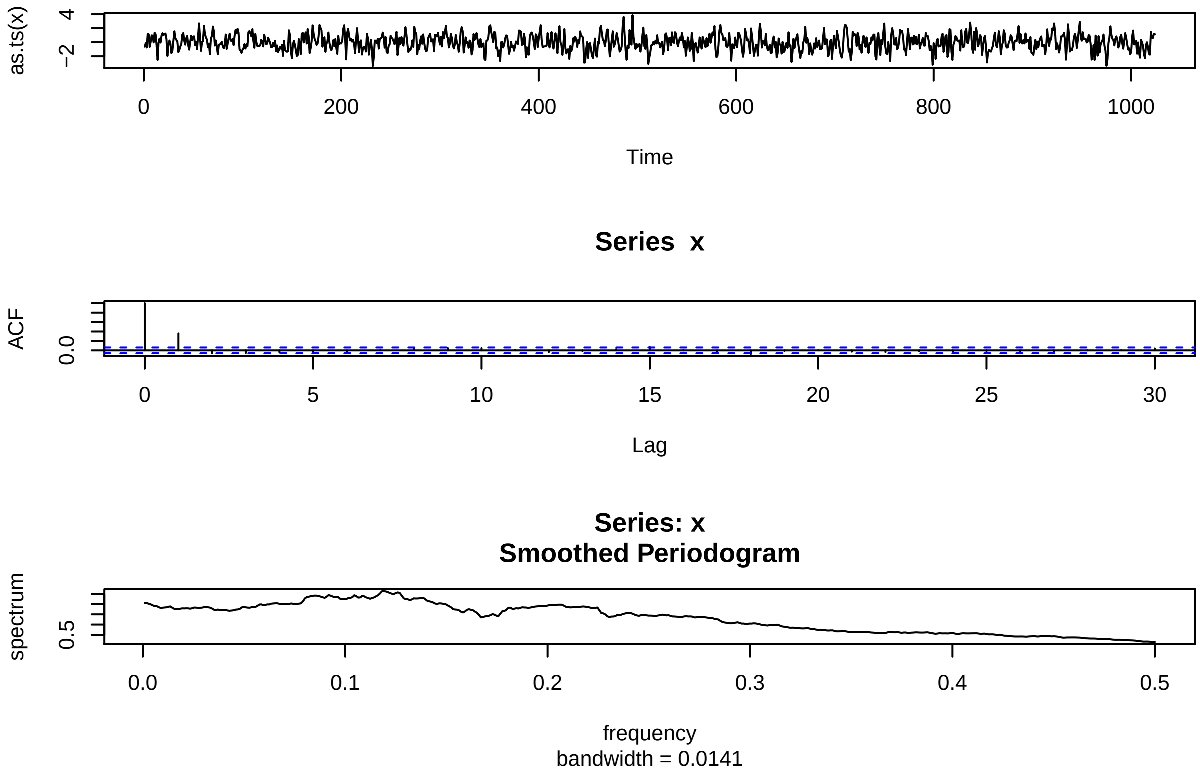
```
# MA(1)
set.seed(1)
x <- w <- rnorm(1024)
# beta = 0.9
for (t in 2:1024) x[t]<- 0.9 * w[t-1] + w[t]
layout(1:3)
plot(as.ts(x))
acf(x)
spectrum(x, span = 51, log = c("no"))
```



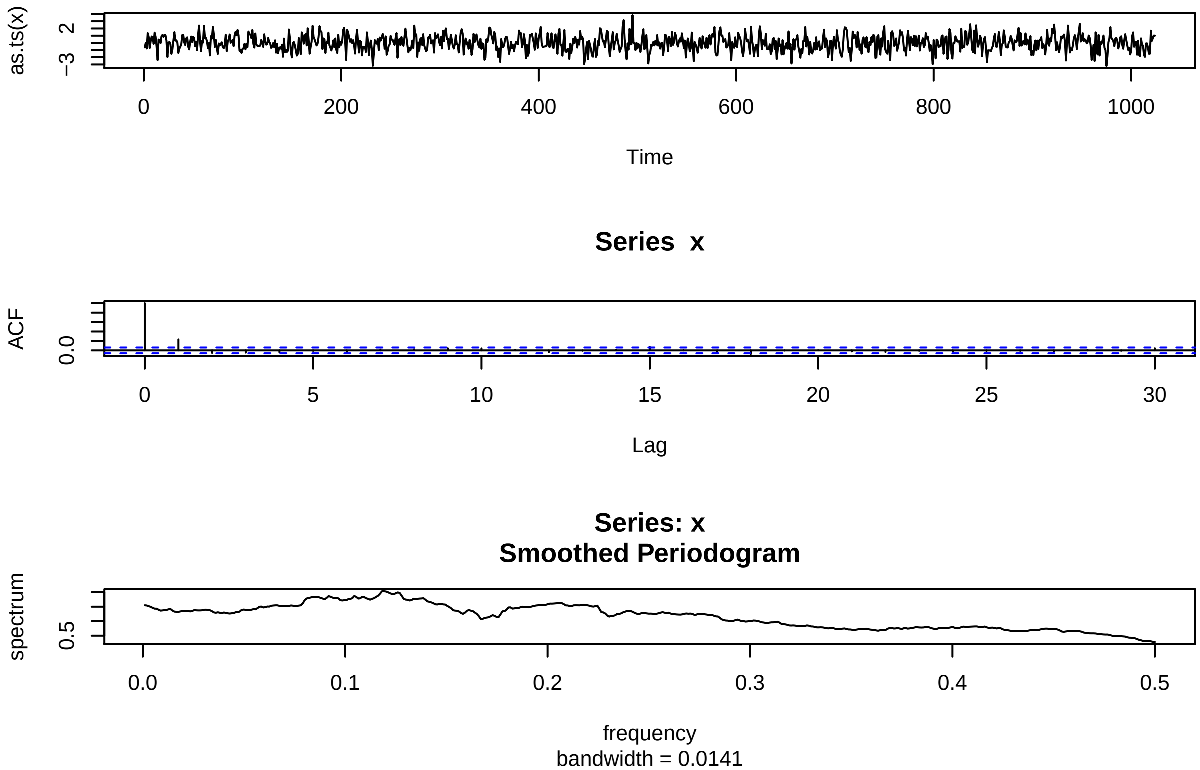

```r
# beta = 0.7
for (t in 2:1024) x[t]<- 0.7 * w[t-1] + w[t]
layout(1:3)
plot(as.ts(x))
acf(x)
spectrum(x, span = 51, log = c("no"))
```



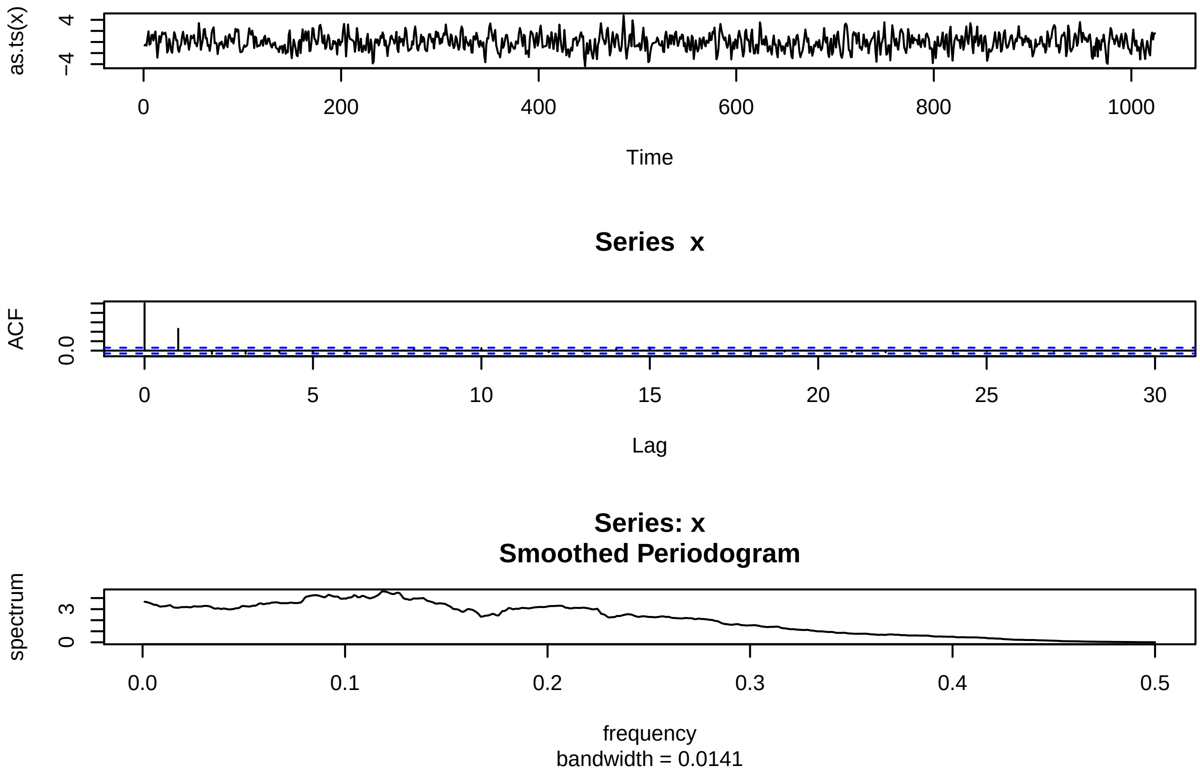

```
# beta = 0.5
for (t in 2:1024) x[t]<- 0.5 * w[t-1] + w[t]
layout(1:3)
plot(as.ts(x))
acf(x)
spectrum(x, span = 51, log = c("no"))
```



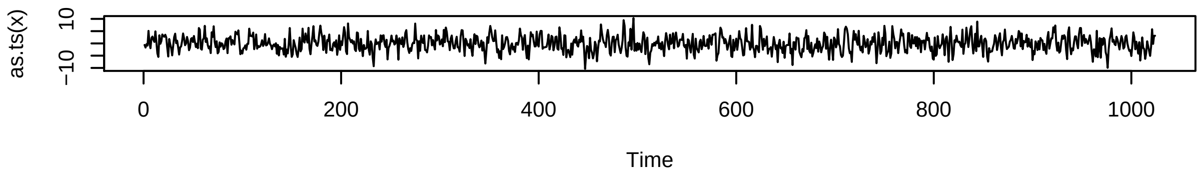

```
# beta = 0.3
for (t in 2:1024) x[t]<- 0.3 * w[t-1] + w[t]
layout(1:3)
plot(as.ts(x))
acf(x)
spectrum(x, span = 51, log = c("no"))
```

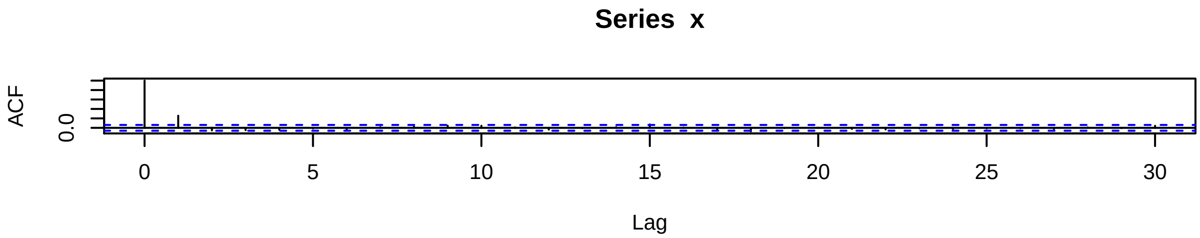


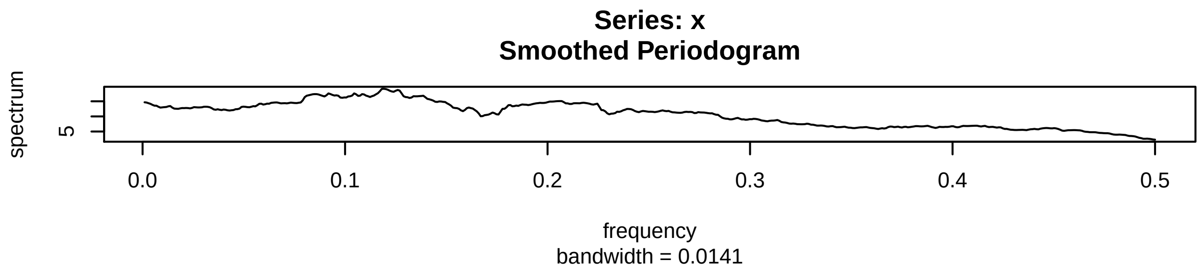

```r
# beta = 1
for (t in 2:1024) x[t]<- 1 * w[t-1] + w[t]
layout(1:3)
plot(as.ts(x))
acf(x)
spectrum(x, span = 51, log = c("no"))
```




```
# beta = 3
for (t in 2:1024) x[t]<- 3 * w[t-1] + w[t]
layout(1:3)
plot(as.ts(x))
acf(x)
spectrum(x, span = 51, log = c("no"))
```







2. Show that if X_t and Y_t are independent, stationary processes with power spectral density functions $f_x(\omega)$ and $f_y(\omega)$, then $V_t = X_t + Y_t$ is also stationary with power spectral density $f_v(\omega) = f_x(\omega) + f_y(\omega)$

$$f(\omega) = 1/\pi\big[\gamma(0) + 2\Sigma_{k=1}^{\infty}\gamma(k)\cos(\omega k)\big]$$

X_t and Y_t are independent, so,

$$\sigma_v^2 = \sigma_x^2 + \sigma_y^2$$

$$\begin{gathered}\gamma_v(h) = \Sigma_{j=1}^{h}\sigma_v^2\cos(\omega_j h)\\ = \Sigma_{j=1}^{h}(\sigma_x^2 + \sigma_y^2)\cos(\omega_j h)\\ = \Sigma_{j=1}^{h}\sigma_x^2\cos(\omega_j h) + \Sigma_{j=1}^{h}\sigma_y^2\cos(\omega_j h)\\ = \gamma_x(h) + \gamma_y(h)\end{gathered}$$

$$\begin{gathered}f_v(\omega) = 1/\pi\big[\gamma(0) + 2\Sigma_{k=1}^{\infty}\gamma_v(k)\cos(\omega k)\big]\\ = 1/\pi\big[(\gamma_x(0) + \gamma_y(0)) + 2\Sigma_{k=1}^{\infty}(\gamma_x(k) + \gamma_y(k))\cos(\omega k)\big]\\ = 1/\pi\big[(\gamma_x(0) + 2\Sigma_{k=1}^{\infty}\gamma_x(k)\cos(\omega k)\big] + 1/\pi\big[(\gamma_y(0) + 2\Sigma_{k=1}^{\infty}\gamma_y(k)\cos(\omega k)\big]\\ f_x(\omega) + f_y(\omega)\end{gathered}$$

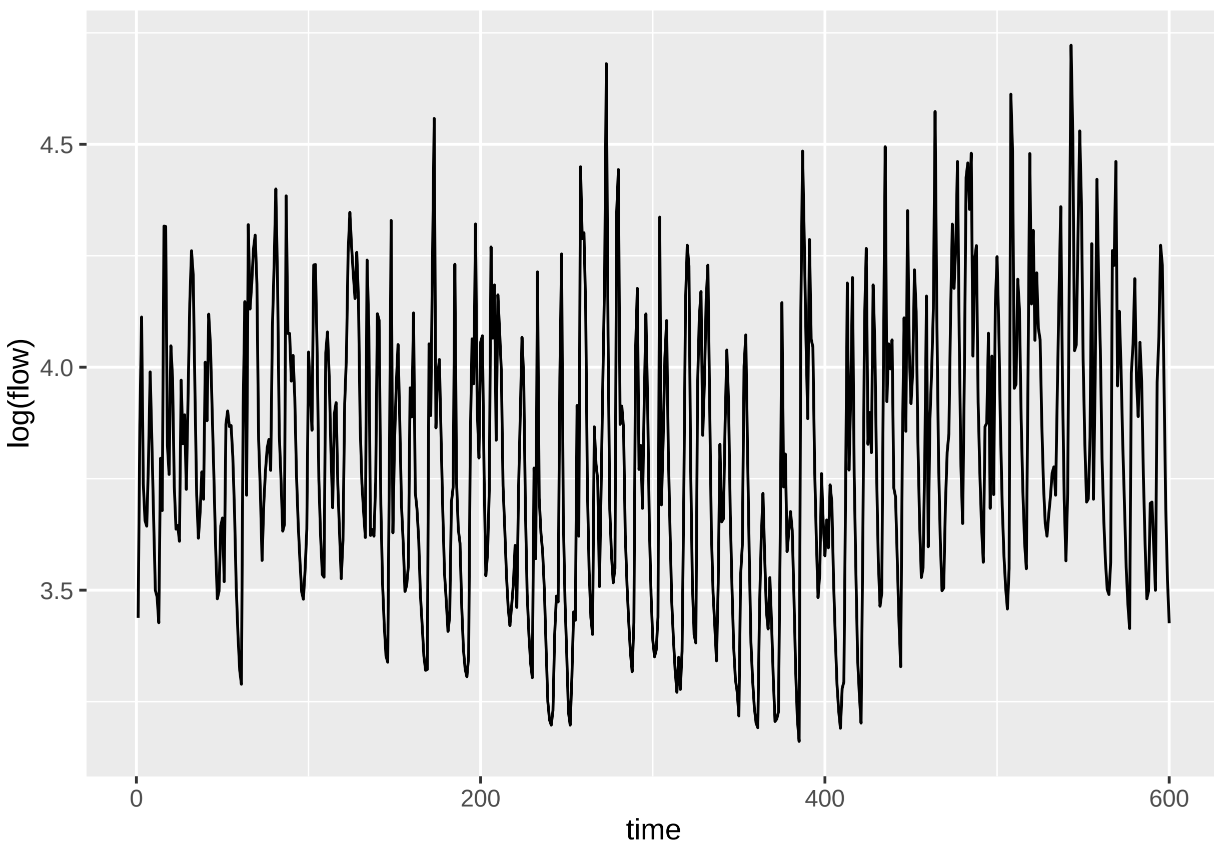
Hence, $f_v(\omega) = f_x(\omega) + f_y(\omega)$.

Question 2

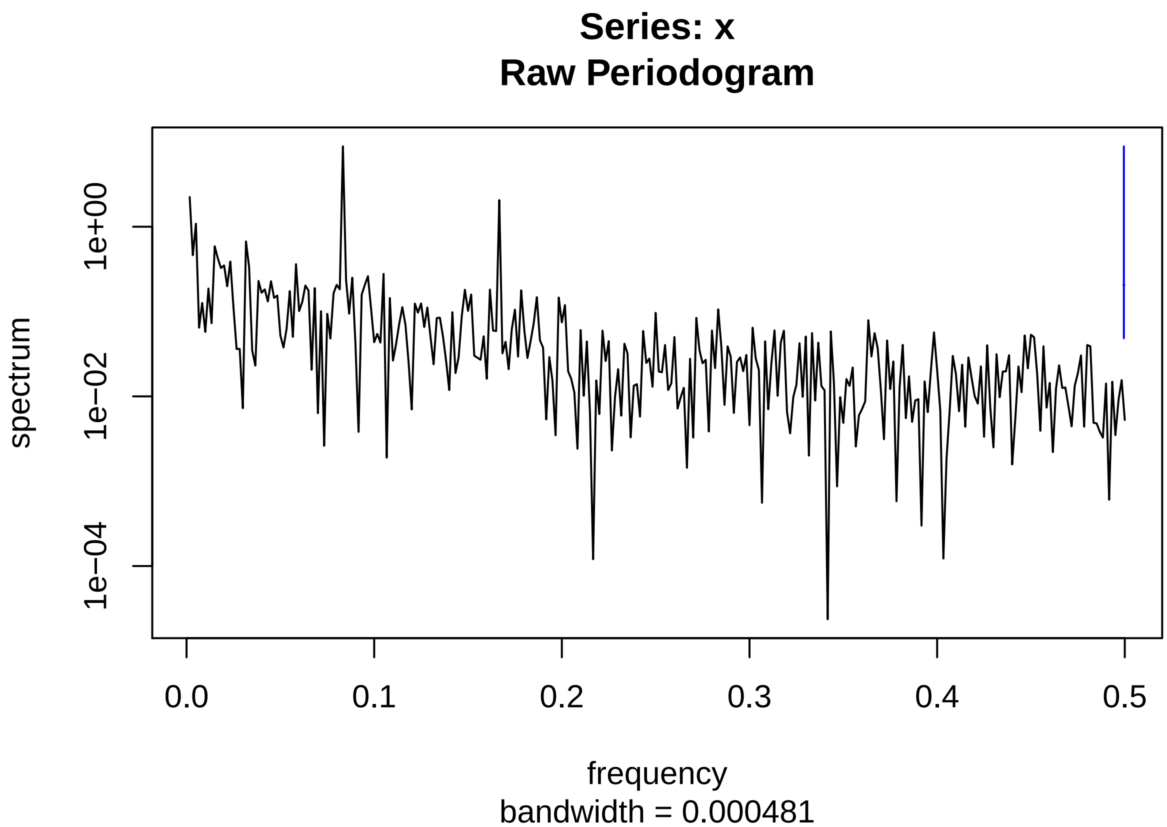
The data.frame flow_df contains the average monthly river flow m^3/sec in the Mckenzie river at Mckenzie Bridge, Oregon. (I got this from http://robjhyndman.com/tsdldata/askew/askew7.dat who quotes the source: Hipel and Mcleod (1994)) The column time contains a simple time index, the number of months since the start of the record. The column date contains a decimal representation of the date, i.e. 1911.750 is October 1911.

1. Estimate the spectrum of the logarithm of flow. Make sure you show evidence you experimented with the amount of smoothing, but you need only show your final plot.

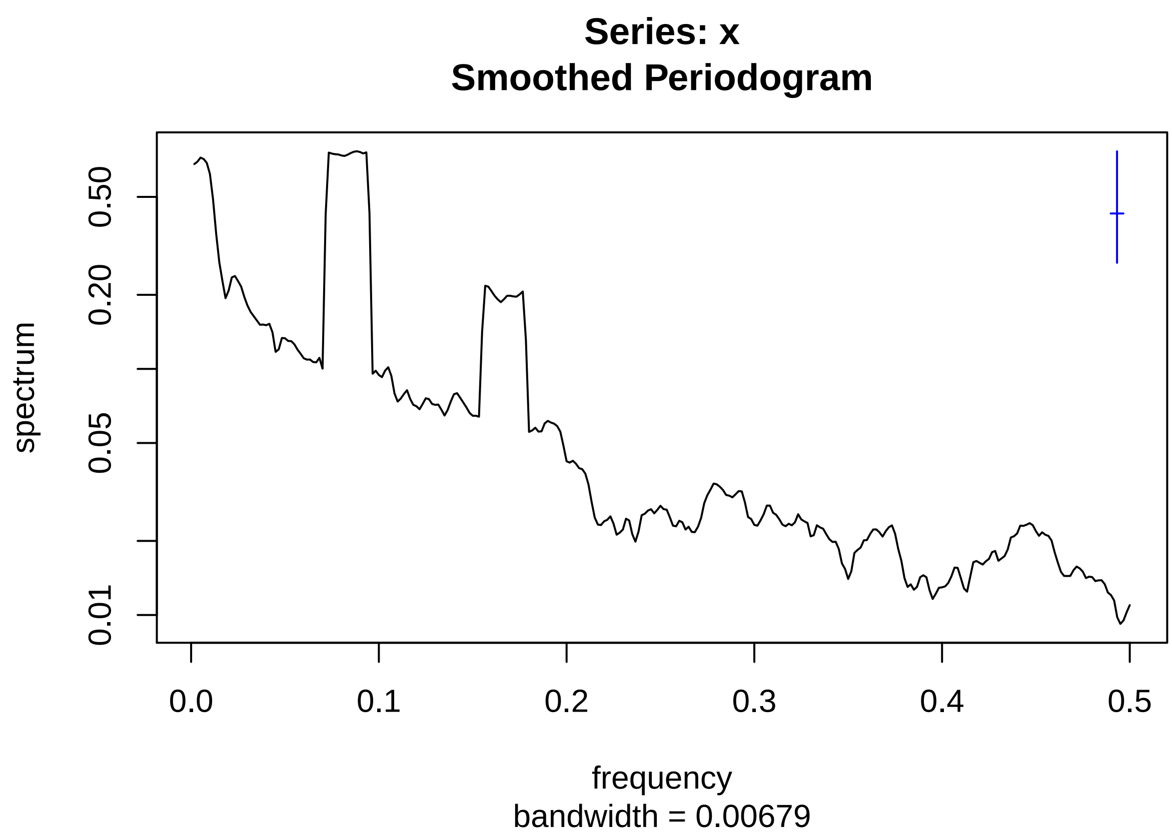
```
library(ggplot2)
load(url("http://stat565.cwick.co.nz/data/flow_df.rda"))
qplot(time, log(flow), data = flow_df, geom = "line")
```




```
attach(flow_df)
flow_log <- log(flow)
# flow_log
spectrum(flow_log)
```




```
# freq = 1, corresponds to 1 cycle per year

## try some spans
# spectrum(flow_log, taper = 0)
# spectrum(flow_log, spans = 5, taper = 0)
# spectrum(flow_log, spans = 10, taper = 0)
spectrum(flow_log, spans = 15, taper = 0)
```




```r
# looks good
# spectrum(flow_log, spans = 20, taper = 0)
# spectrum(flow_log, spans = 23, taper = 0)
# spectrum(flow_log, spans = 50, taper = 0)
```

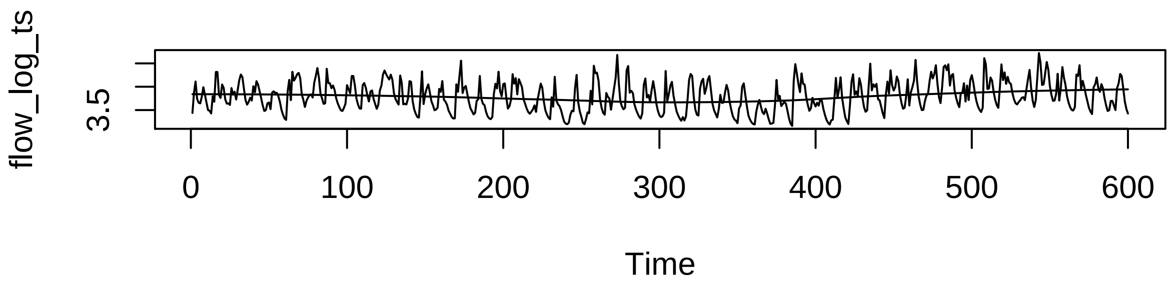
---

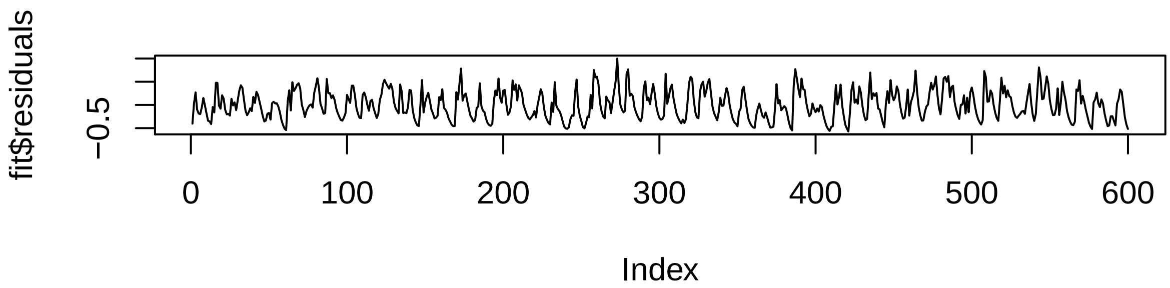
2. Fit a smooth trend to the logarithm of flow, and estimate the spectrum of the residuals. How does this spectrum differ from the one in part 1.?

```r
attach(flow_df)
```

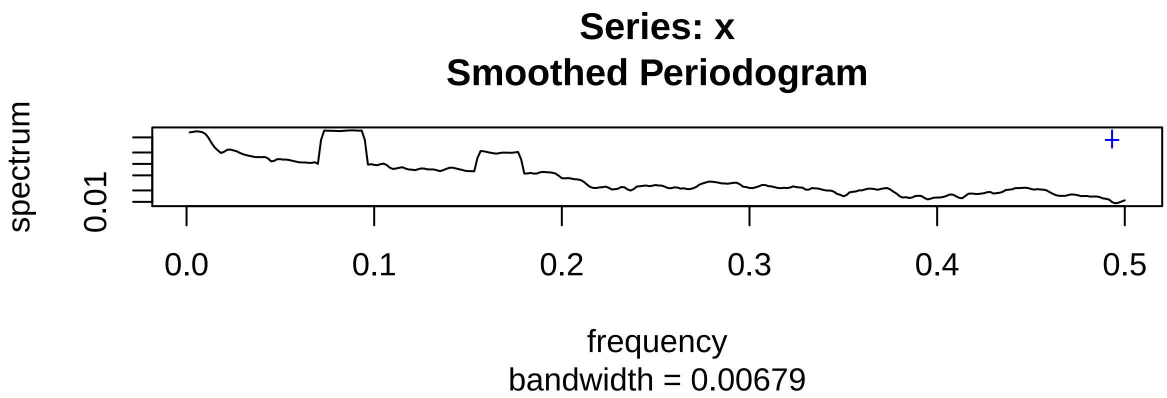
```
## The following objects are masked from flow_df (pos = 3):
##
##     date, flow, time
```

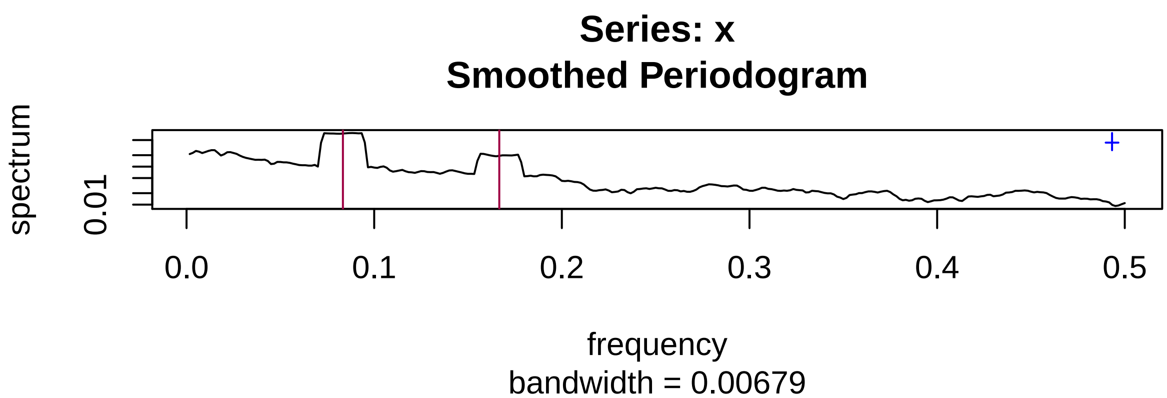
```r
fit <- loess(flow_log ~ time, method = "loess")

par(mfrow = c(2,1))
flow_log_ts <- ts(flow_log, start = 1, frequency = 1)
plot(flow_log_ts)
lines(fit$fitted)
plot(fit$residuals, type = "l")
```






```r
par(mfrow = c(2,1))
spectrum(flow_log, spans = 15, taper = 0)
spectrum(fit$residuals, spans = 15, taper = 0)

abline(v = c(1/12, 2/12), col = "#9E0142")
```






```
#
# abline(v = c(1/12, 2/12, 3/12, 4/12), col = "#9E0142")

str(flow_log_ts)
```

```
##  Time-Series [1:600] from 1 to 600: 3.44 3.87 4.11 3.74 3.66 ...
```

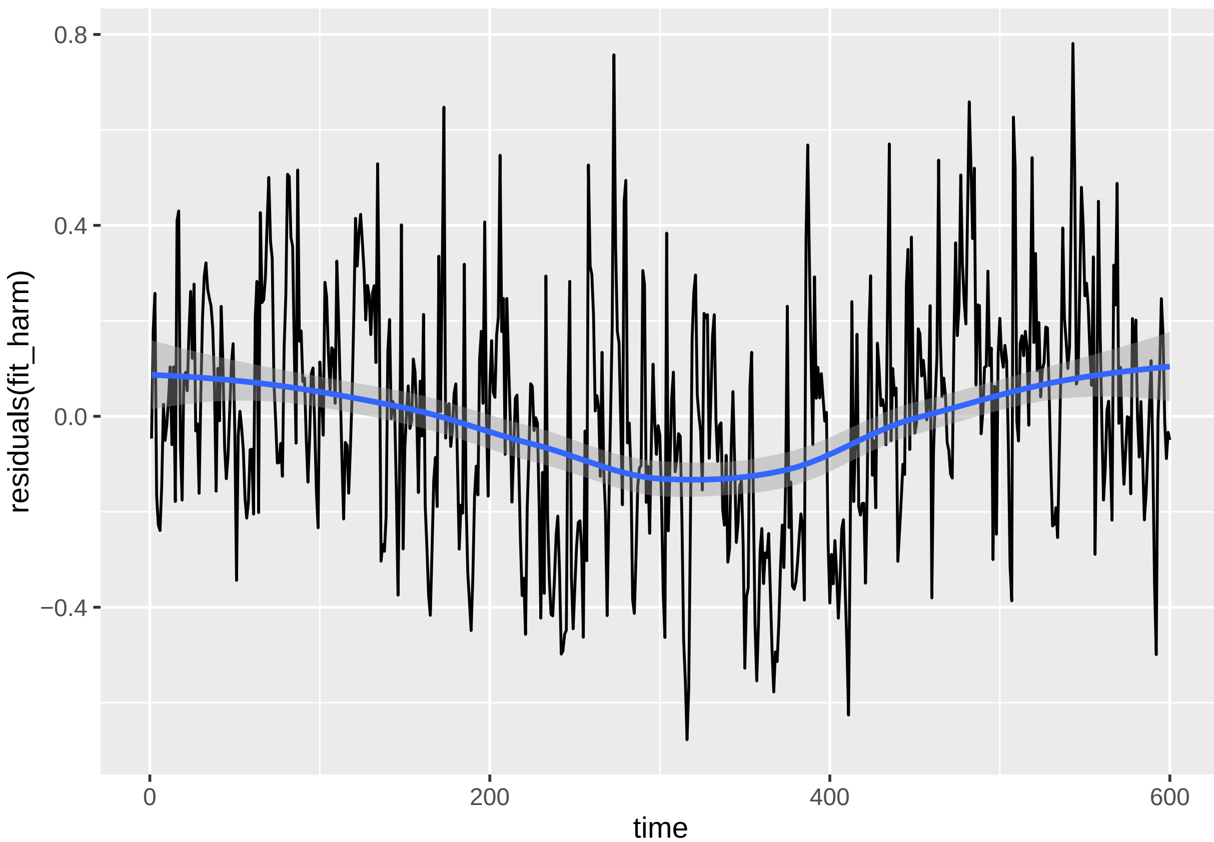
The spectrum of the residual differs from the spectrum of the original data (logarithm of flow) in the first left-most frequencies. This is not surprising as the smoothed trend model we fit earlier, is pretty much a horizontal line/curve, and removing the trend would result in a similar residual series with similar frequencies, except for the level shift. This level shift difference between the two spectrums is visible in the two spectrum plots' Y-axis tick marks.

---

3. Fit a harmonic regression to the residuals from 2 using the estimated spectrum to choose the number and frequencies of the periodic components.

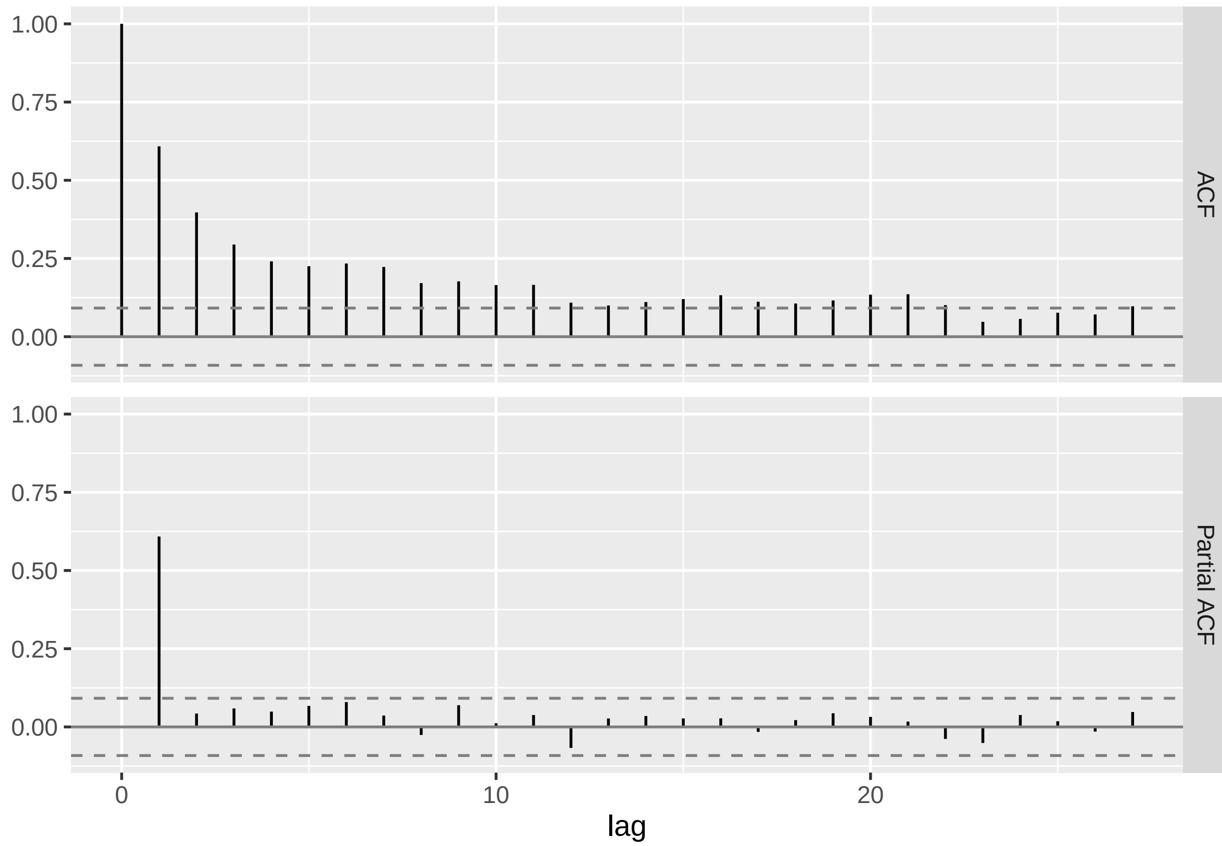
```
par(mfrow = c(1,1))



periodic <- function(x, frequency = 1, order = 1){
```

```
  do.call(cbind, lapply(1:order, function(ord){
    cbind(cos(2*pi*ord*frequency*x), sin(2*pi*ord*frequency*x))
  }))
}

fit_harm <- lm(flow_log ~ time + periodic(time, freq = 1/12, order = 2), data = flow_df)

qplot(time, residuals(fit_harm), data = flow_df, geom = "line") + geom_smooth()
```

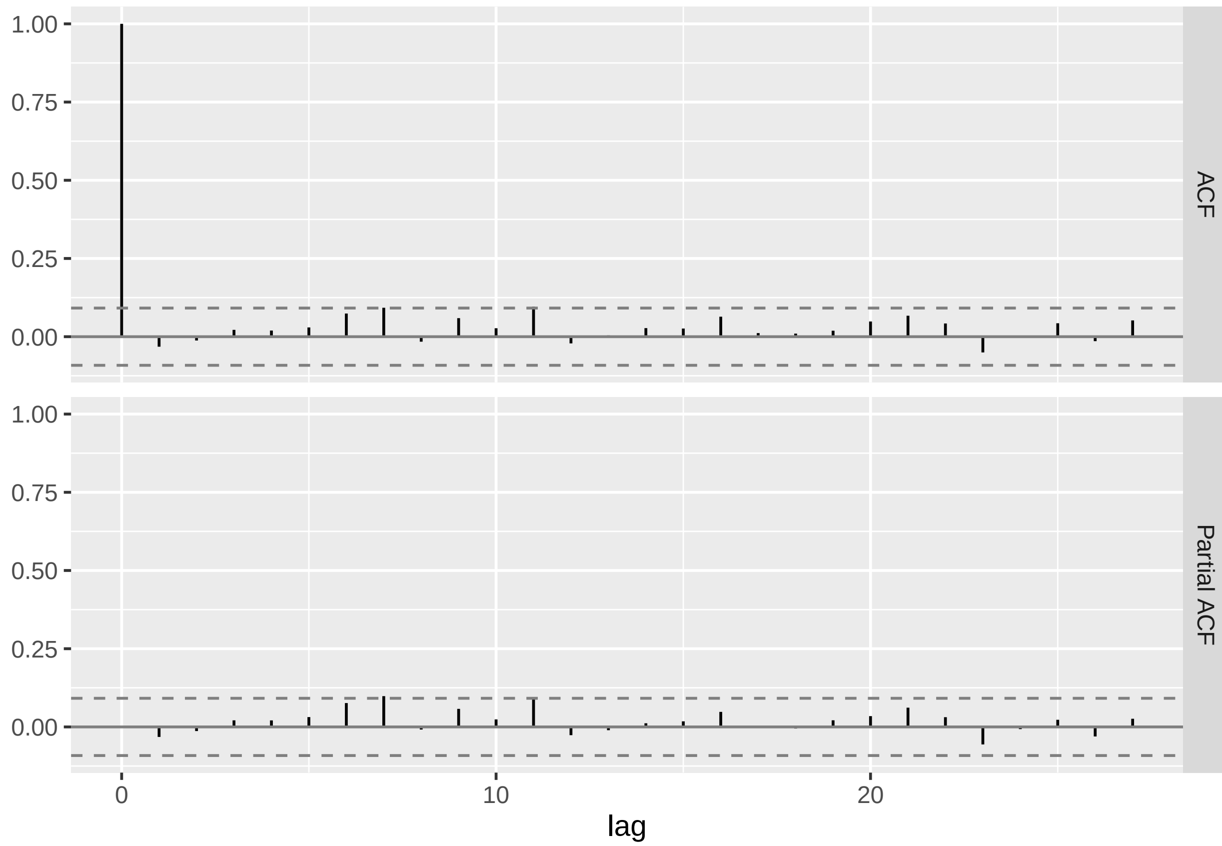



```
source(url("http://stat565.cwick.co.nz/code/get_acf.R"))
examine_corr(residuals(fit_harm))
```

```
## Warning: closing unused connection 6 (http://stat565.cwick.co.nz/code/
## get_acf.R)
```



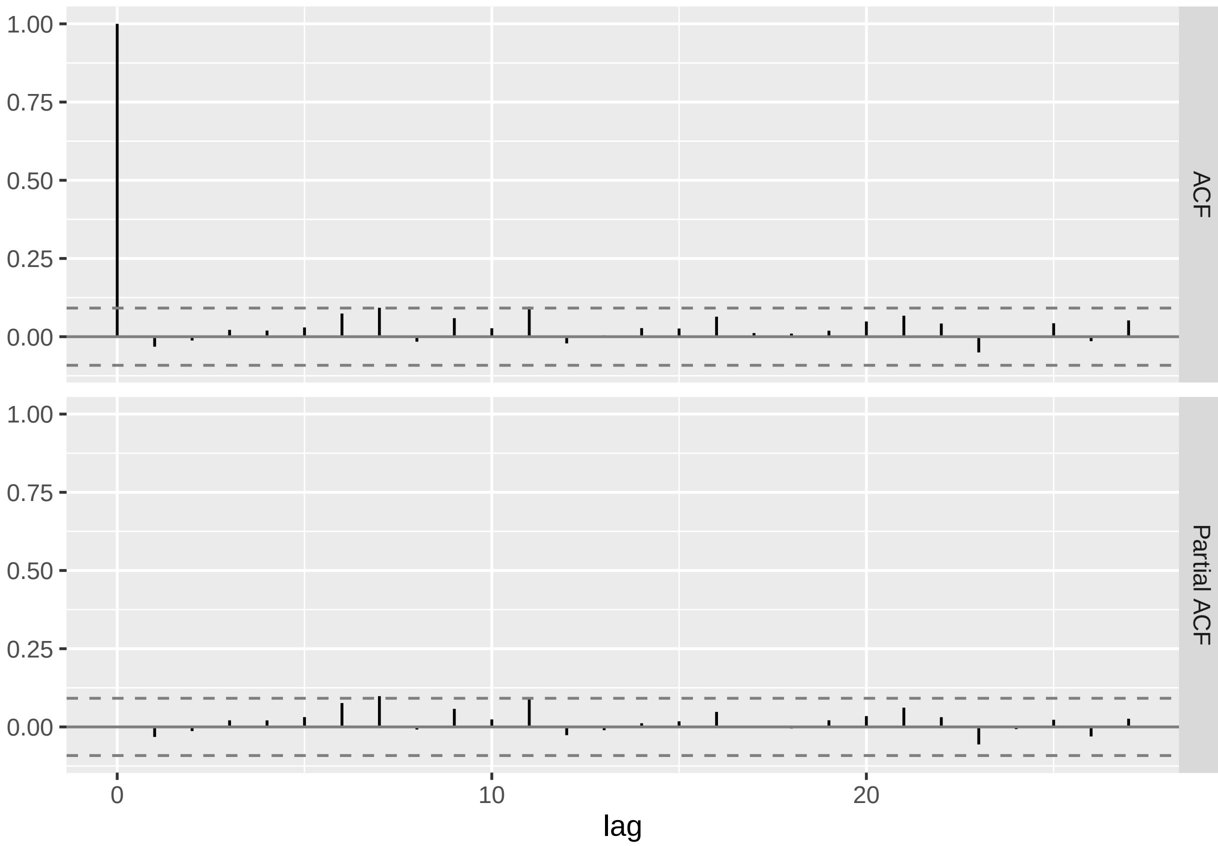

```
library(nlme)
fit_gls <- gls(flow_log ~ time + periodic(time, freq = 1/12, order = 2), data = flow_df,
               correlation = corARMA(p = 1, q = 0))
examine_corr(residuals(fit_gls, type = "normalized"))
```



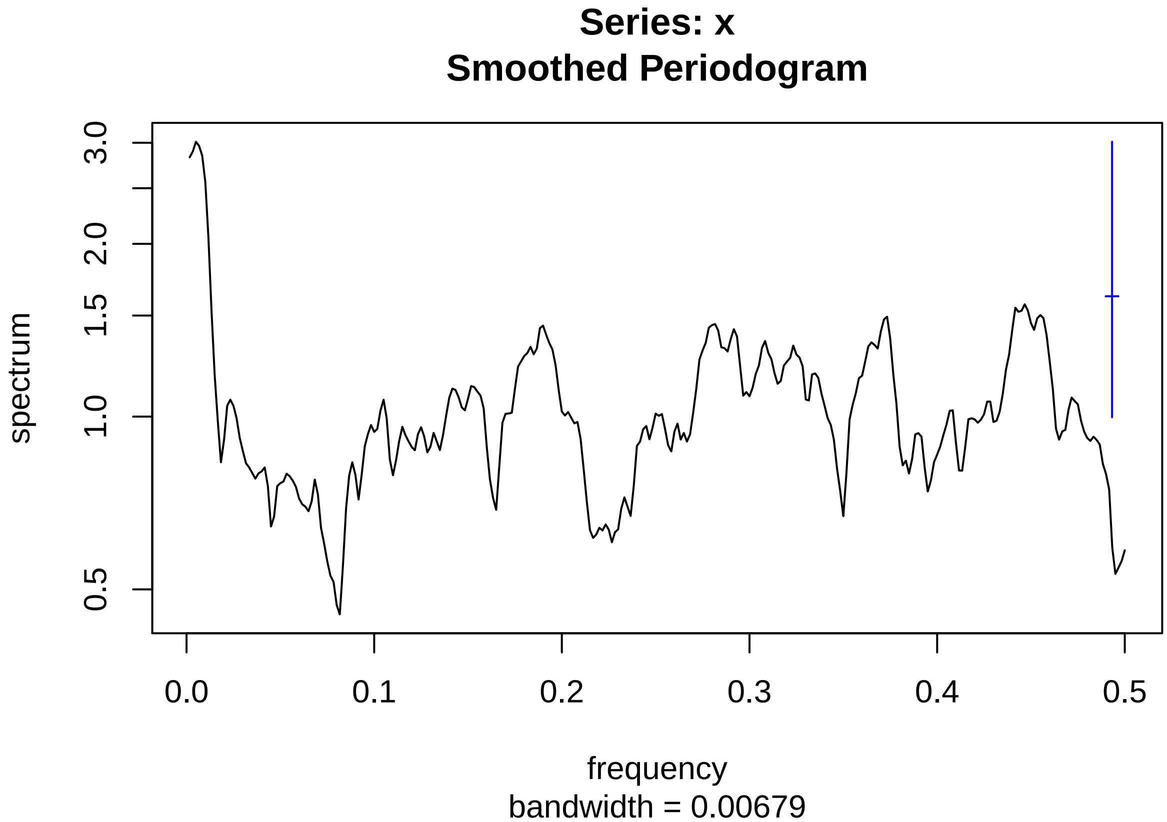
—

4. Examine the residuals from the harmonic regression using both the ACF/PACF and periodogram. Is there any evidence of remaining autocorrelation?

```
# examine the residuals
examine_corr(residuals(fit_gls, type = "normalized"))
```




```
spectrum(residuals(fit_gls, type = "normalized"), span = 15) # looks good
```



Residuals from the linear regression model shows AR(1) temporal correlation remaining. So, I fit a gls model with AR(1) component. The residuals from this gls harmonic regression model with an AR(1) component looks good, because it looks like white noise, which shows there's no evidence of remaining autocorrelation.