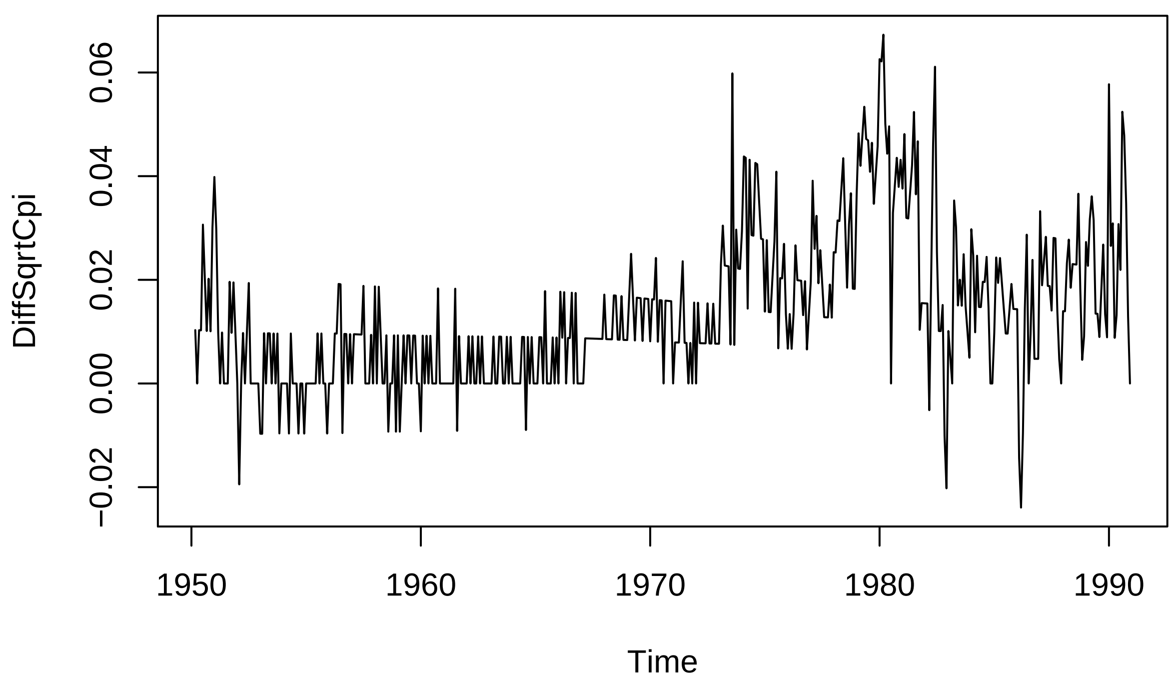
Selvstudiegang 3

Opgave 1: Long memory

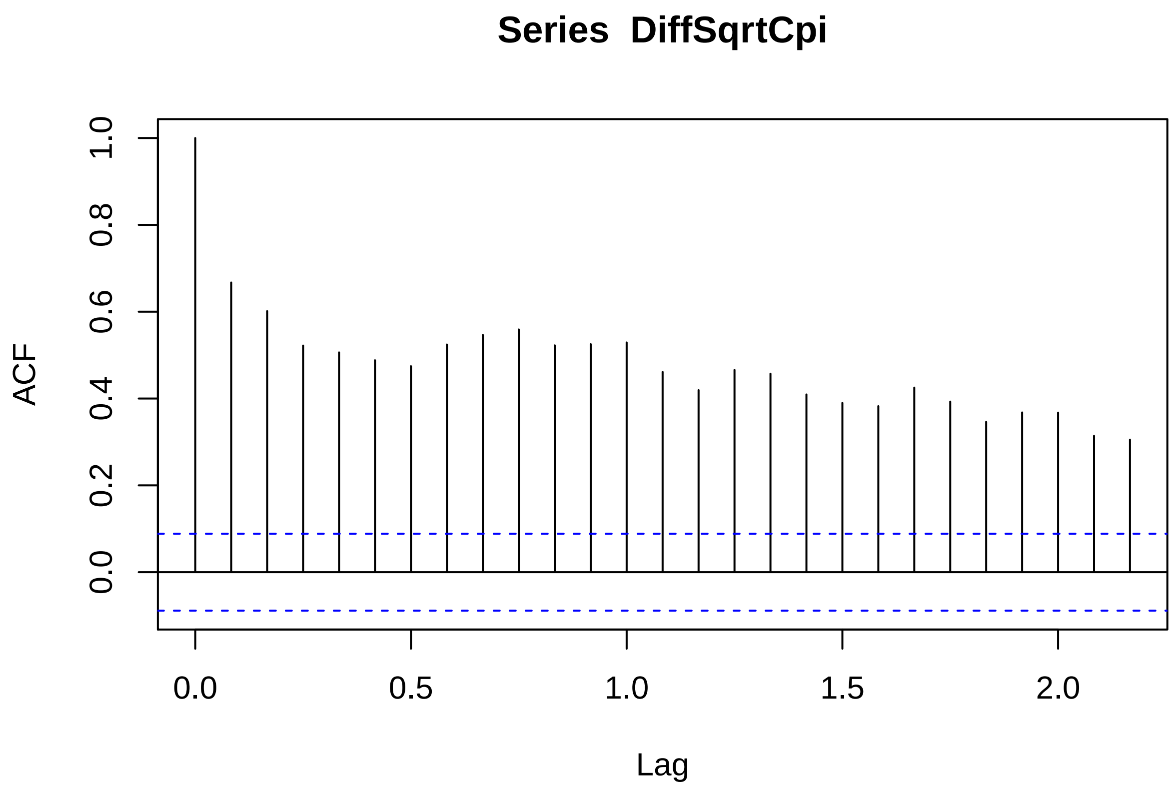
Fra datasættet `Mishkin` i **R**-pakken `Ecdat` betragtes ændringerne i kvadratroden af forbrugerprisindekset `cpi`.

1. Beregn denne variabel og kald den `DiffSqrtCpi`. Plot tidsrækken og dens ACF. Er der tegn på *long-memory*? I bekræftende fald, hvordan ses det?

```
DiffSqrtCpi <- diff(sqrt( Mishkin[,5] ))

plot(DiffSqrtCpi , type = "l")
```

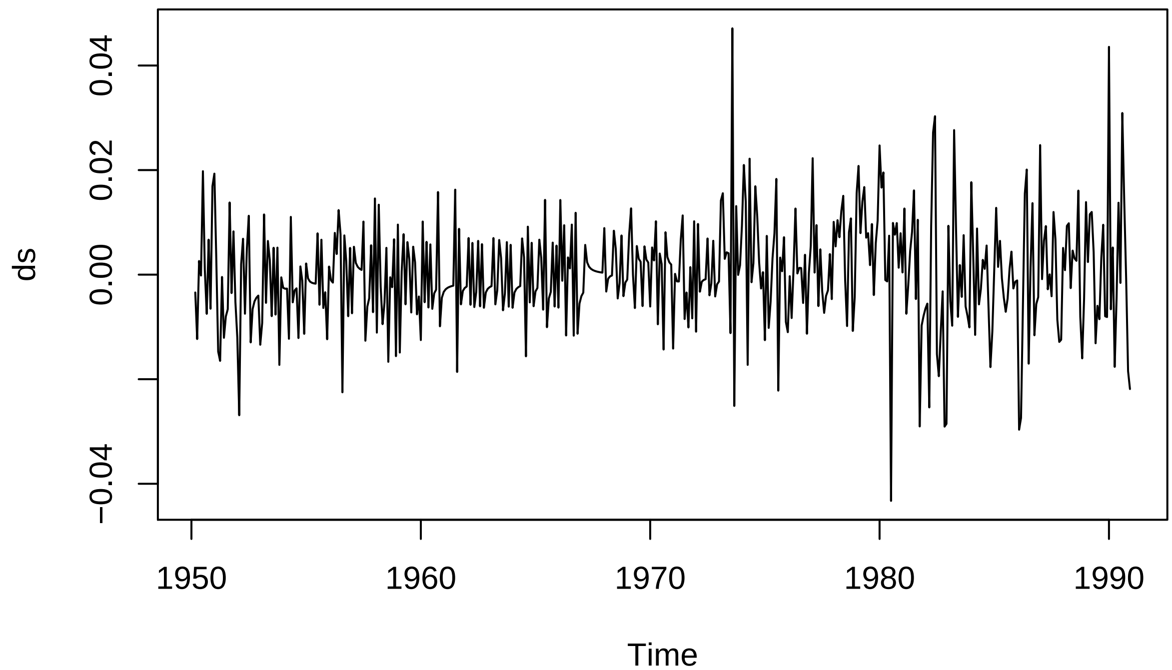


```r
acf(DiffSqrtCpi)
```



2. Estimer en passende model for tidsrækken `DiffSqrtCpi`. Dette kan gøres på forskellige måder:
   - Enten via pakken `fracdiff` samt eventuelt procedurerne i filen `Polyprocs.R` (se dokumentationen i notatet om lag polynomier `polynom.pdf`).
   - Eller via pakken `arfima` som tilpasser fraktionelle modeller og giver mere rimelige estimater af usikkerheden. Se *Example 5.1 Redux* på denne side som inspiration: https://www.stat.pitt.edu/stoffer/tsa4/Rexamples.htm

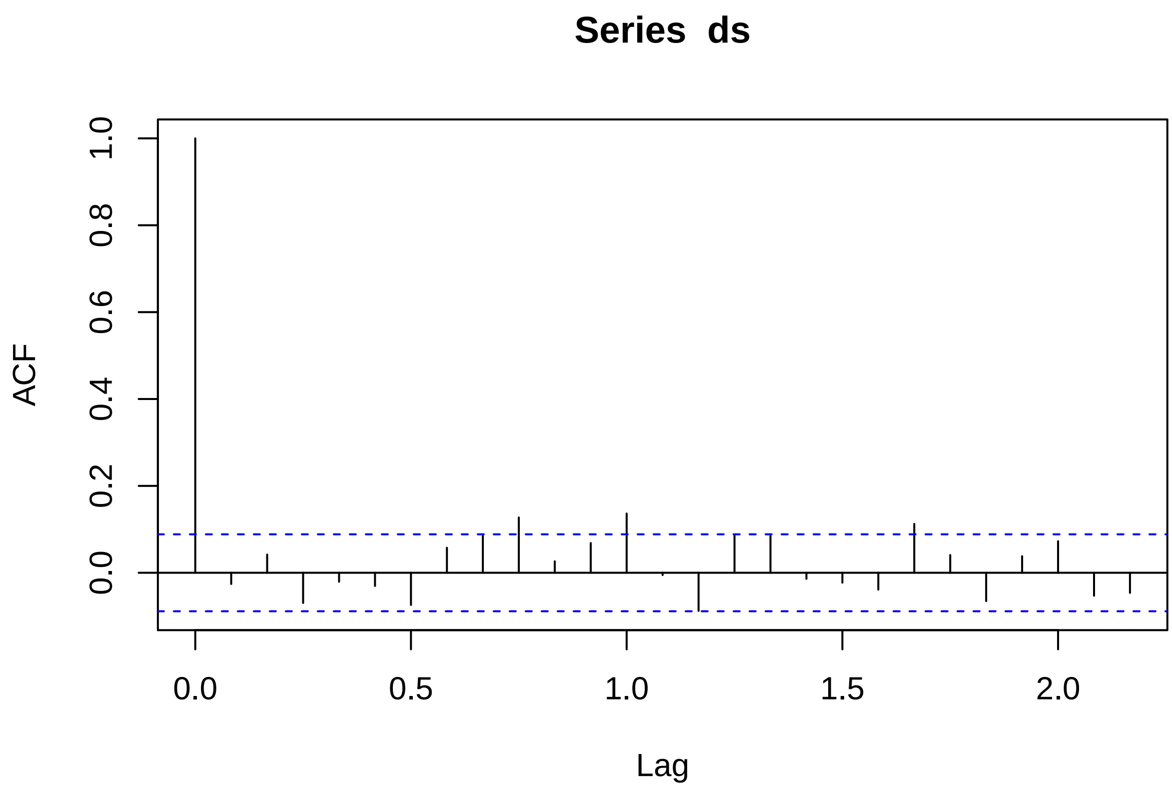
```r
fd <- fracdiff(DiffSqrtCpi)

ds <- diffseries(DiffSqrtCpi , fd$d)

plot(ds)
```



3. Er der ud fra autokorrelationsfunktionen af den fraktionelt differensede tidsrække (som selvfølgelig er fremkommet ved fraktionel differensning ved brug af den estimerede d fra tidligere spm.) kort og eller lang hukommelse i denne?

   (Vink: Brug `diffseries` fra pakken `fracdiff`.)

```r
acf(ds)
```



Der er kort

Opgave 2: Kointegration og ECM

1. Betragt følgende såkaldte *autoregressive distributed lag* (ADL) model, hvor y_t og x_t er to tidsrækker, og w_t er en hvid støj, og hvor $|\phi_1| < 1$:

$$y_t = \alpha + \phi_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + w_t.$$

Spm1: Vis at

$$\sum_{s=0}^{\infty} \frac{\partial y_{t+s}}{\partial x_t} = \frac{\beta_0 + \beta_1}{1 - \phi_1}.$$

Denne størrelse kaldes langsigtsmultiplikatoren (*the long run multiplier*). Dvs. langsigts påvirkningen af x_t på y-processen.

Spm2: Vis at ADL modellen kan skrives som følgende fejlkorrektionsmodel (ECM),

$$\nabla y_t = \beta_0 \nabla x_t - (1 - \phi_1)(y_{t-1} - \mu - \beta_2 x_{t-1}) + w_t,$$

hvor $\mu = \frac{\alpha}{1-\phi_1}$ og hvor $\beta_2 = \frac{\beta_0+\beta_1}{1-\phi_1}$ er langsigtsmultiplikatoren. *VINK: Start med at omskrive ECM og nå frem til ADL.*

Spm3: Hvilken simpel ligevægtsrelation ville opnås mellem y og x på langt sigt, hvis vi antog at $w_t = 0$, og at værdierne af $y_t = y_{t-1} = y$ og $x_t = x_{t-1} = x$ var faste?

2. Download filen `ibr.Rdata` fra Moodle. Filen indeholder *Interest and Bond Rates (IBR)* data fra USA. `AAA` er månedlige *corporate bond yields* (Triple A obligationer) i procenter og `US3MT` er månedlige *three-month US Treasury Bill rates* (skatkammerbeviser) i procenter. `DAAA` og `DUS3MT` er førstedifferenserne af disse to variable.

```r
load("ibr.Rdata")
head(ibr)
```

```
##       OBS DUS3MT  DAAA US3MTBIL  AAA
## 1 1948:01   0.02  0.00     0.97 2.86
## 2 1948:02   0.02 -0.01     0.99 2.85
## 3 1948:03   0.01 -0.02     1.00 2.83
## 4 1948:04   0.00 -0.05     1.00 2.78
## 5 1948:05   0.00 -0.02     1.00 2.76
## 6 1948:06   0.00  0.00     1.00 2.76
```

```
**Spm1:** Vis at `AAA` og `US3MT` kointegrerer.
```

```r
lm.reg <- lm(ibr$AAA ~ ibr$US3MT)
lm.err <- lm.reg$residuals

adf.test(lm.err , k = 0)
```

```
## Warning in adf.test(lm.err, k = 0): p-value smaller than printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  lm.err
## Dickey-Fuller = -4.0476, Lag order = 0, p-value = 0.01
## alternative hypothesis: stationary
```

Forkaster nulhypotesen om at residualerne ikke er stationære, altså kointegrerer `AAA` og `US3MT`

```
**Spm2:** Hvis $y_t=$`AAA`, og $x_t=$`US3MT` estimér da modellen
og tolk koefficienterne.
```

```r
yl <- embed(ibr$AAA , 2)
xl <- embed(ibr$US3MT , 2)

lm.reg.ADL <- lm(yl[,1] ~ yl[,2] + xl[,1] + xl[,2])
summary(lm.reg.ADL)
```

```
##
## Call:
## lm(formula = yl[, 1] ~ yl[, 2] + xl[, 1] + xl[, 2])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.65903 -0.06086 -0.00771  0.06684  0.96201
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.044779   0.017127   2.615  0.00915 **
## yl[, 2]      0.971022   0.005046 192.439  < 2e-16 ***
## xl[, 1]      0.279967   0.014044  19.935  < 2e-16 ***
```

```
## xl[, 2]      -0.247889   0.014649 -16.922  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1633 on 619 degrees of freedom
## Multiple R-squared:  0.9971, Adjusted R-squared:  0.9971
## F-statistic: 7.214e+04 on 3 and 619 DF,  p-value: < 2.2e-16
```

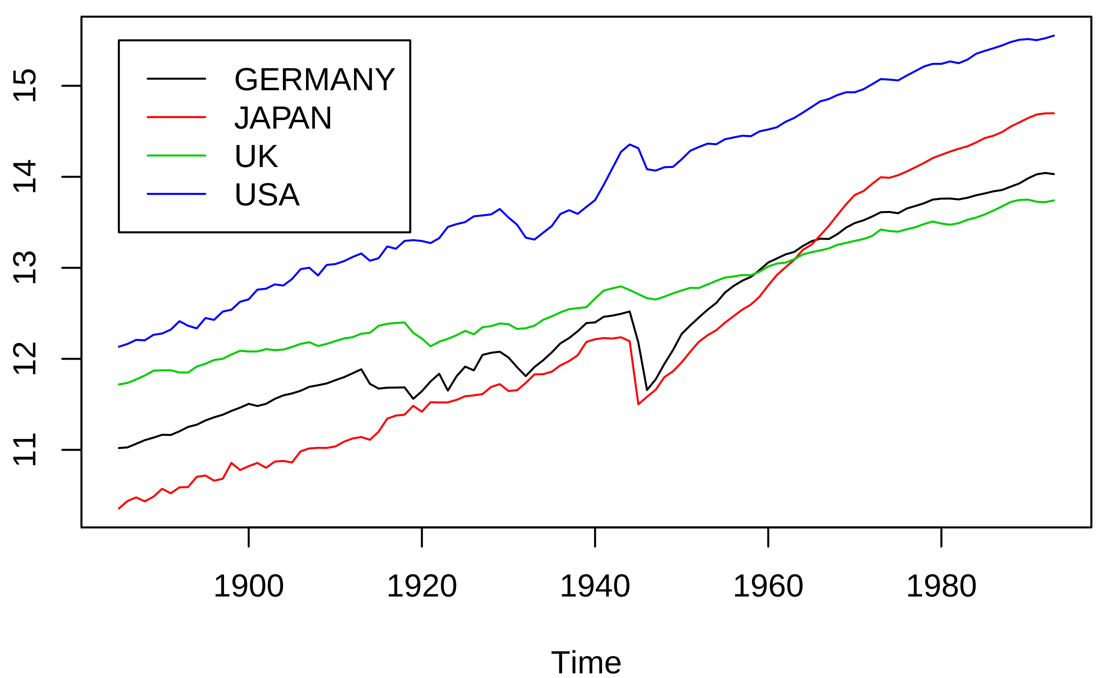
Opgave 3: Kointegration med flere tidsrækker

Datafilen `GNP.RData` indeholder årlige data for bruttonationalproduktet (GNP) for fire lande i perioden 1885–1993. Det kan være en fordel at lave denne til en mulitvariat tidsrække (`mts` object). Dette gøres via `ts`-kommandoen med en `matrix`/`data.frame` som input, f.eks:

```
load("GNP.RData")
GNP2 <- ts(GNP[, 6:9], start = 1885, frequency = 1)
```

1. Plot de fire tidsrækker (i logaritmer) i samme figur. Og kommentér graferne.

```
ts.plot( GNP2 , col = 1:4) # Er logget i forvejen
#legend(x = 1885 , y = 15.5, colnames(GNP[2:5])  , fill = palette()[1:4] )
legend(x = 1885 , y = 15.5, colnames(GNP[2:5]) , col = 1:4 , lty = 1 )
```



2. Test om de fire tidsrækker hver for sig har en unit root.

```r
df <- data.frame( GER = adf.test(GNP2[,1] , k = 0)$p.value ,
                  JAP = adf.test(GNP2[,2] , k = 0)$p.value ,
                  UK  = adf.test(GNP2[,3] , k = 0)$p.value ,
                  USA = adf.test(GNP2[,4] , k = 0)$p.value ) ; df
```

```
##         GER       JAP        UK     USA
## 1 0.5917497 0.8056927 0.7361467 0.22711
```

Unit root i alle

3. Undersøg om de fire tidsrækker (i logaritmer) kointegrerer. Brug Engle-Granger metoden, og lad de amerikanske data være responsvariablen i regressionen.

```r
lm.err.mts <- lm(GNP2[,4] ~ GNP2[,1] + GNP2[,2] + GNP2[,3])$residuals

adf.test(lm.err.mts , k = 0)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  lm.err.mts
## Dickey-Fuller = -2.1238, Lag order = 0, p-value = 0.5253
## alternative hypothesis: stationary
```

```r
# Det herunder er ikke hvad der er tænkt med opgaven
df <- data.frame(GER = numeric(4) , JAP = numeric(4) , UK = numeric(4) , USA = numeric(4))
rownames(df) <- c("GER" , "JAP" , "UK" , "USA")

for(i in 1:4){
  for(j in 1:4){
    lm.err <- lm(GNP2[,i] ~ GNP2[,j])$residuals
    df[i,j] <- adf.test(lm.err , k = 0)$p.value
  }
}
```

```
## Warning in adf.test(lm.err, k = 0): p-value smaller than printed p-value

## Warning in adf.test(lm.err, k = 0): p-value smaller than printed p-value

## Warning in adf.test(lm.err, k = 0): p-value smaller than printed p-value

## Warning in adf.test(lm.err, k = 0): p-value smaller than printed p-value
```

```r
df
```

```
##            GER        JAP        UK       USA
## GER 0.01000000 0.07061032 0.2132333 0.6269782
## JAP 0.07286237 0.01000000 0.3698125 0.7188888
## UK  0.20495869 0.35783154 0.0100000 0.6274462
## USA 0.62363936 0.70628220 0.6188106 0.0100000
```

```r
df < 0.05 # TRUE hvis nulhypotesen om ikke stationaritet forkastes, og der "er" kointegration
```

```
##       GER   JAP    UK   USA
## GER  TRUE FALSE FALSE FALSE
## JAP FALSE  TRUE FALSE FALSE
## UK  FALSE FALSE  TRUE FALSE
## USA FALSE FALSE FALSE  TRUE
```

Ingen kointegration da der kun forkastes når en tidsrække bliver sammenlignet med sig selv