Signal and System
Appendix: Correlation Analysis

梁　　勝　　富
成功大學 資訊工程系
sfliang@mail.ncku.edu.tw
Office: 資訊系館 12F 65C06, Tel: Ext. 62549

Lab:神經運算與腦機介面實驗室 (3F 65301)
http://ncbci.csie.ncku.edu.tw/
2024

- Parseval's Relation

The energy of a discrete-time signal $x[n]$ is

$$E_x = \sum_{n=-\infty}^{\infty} |x[n]|^2$$

conv $x[n]*h[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]$

$$y[m] \triangleq \sum_{n=-\infty}^{\infty} x[n] \cdot x[n-m].$$

$h[n]=x[-n]$ $h[n-k]=x[k-n]$

Note that $y[n] = x[n] * x[-n]$, and in particular, $y[0] = \sum_{n=-\infty}^{\infty} x^2[n]$. Applying the convolution theorem, $y[n]$'s Fourier transform of $Y(\omega)$ can be expressed in terms of $X(\omega)$ as follows,

$$\boxed{x[-n] \xleftrightarrow{\text{DTFT}} X(-\Omega)}$$

$$Y(\omega) = X(\omega) \cdot \left(\sum_{n=-\infty}^{\infty} x[-n]e^{-j\omega n} \right) = X(\omega) \cdot X^*(\omega) = |X(\omega)|^2 .$$

Calculating the inverse DTFT at time 0, then we reach that

$$\begin{aligned}\sum_{n=-\infty}^{\infty} x^2[n] = y[0] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} Y(\omega)e^{j\omega\cdot 0}d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |X(\omega)|^2 d\omega.\end{aligned}$$

（劉奕文數位聲訊分析與合成）

Covariance

$$Cov = \sigma_{x,y} = \frac{1}{N-1}\sum_{k=1}^{N}[y(k) - \bar{y}][x(k) - \bar{x}]$$

Covariance Matrix

The covariance matrix gives the variance of the columns of the data matrix in the diagonals while the covariance between columns is given by the off-diagonals:

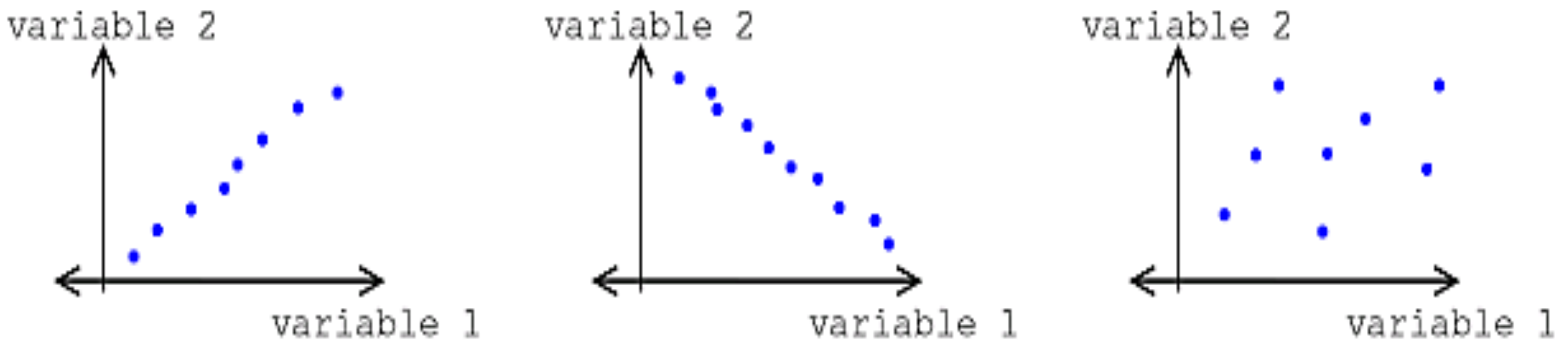
$$S = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,N} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \sigma_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N,1} & \sigma_{N,2} & \cdots & \sigma_{N,N} \end{bmatrix} \tag{19}$$

Correlation Matrix

In its usual signal processing definition, the correlation matrix is a normalized version of the covariance matrix. Specifically, the correlation matrix is related to the covariance matrix by the equation:

$$C(i,j) = \frac{C(i,j)}{\sqrt{C(i,i)\;C(j,j)}} \tag{20}$$

Correlation



Positive correlation Negative correlation Uncorrelated

http://cnx.org/content/m10673/latest/

Matlab Implementation

Correlation or covariance matrices are calculated using the `corrcoef` or `cov` functions respectively. Again, the calls are similar for both functions:

```
Rxx = corrcoef(x)
S = cov(x), or S = cov(x,1);
```

Without the additional `1` in the calling argument, `cov` normalizes by $N - 1$, which provides the best unbiased estimate of the covariance matrix if the observations are from a Gaussian distribution. When the second argument is present, `cov` normalizes by N which produces the second moment of the observations about their mean.

Matlab Implementation

COV Covariance matrix.

COV(X), if X is a vector, returns the variance. For matrices, where each row is an observation, and each column a variable, COV(X) is the covariance matrix.

COV(X,Y), where X and Y are vectors of equal length, is equivalent to COV([X(:) Y(:)]).

COV(X) or COV(X,Y) normalizes by (N-1) where N is the number of observations.

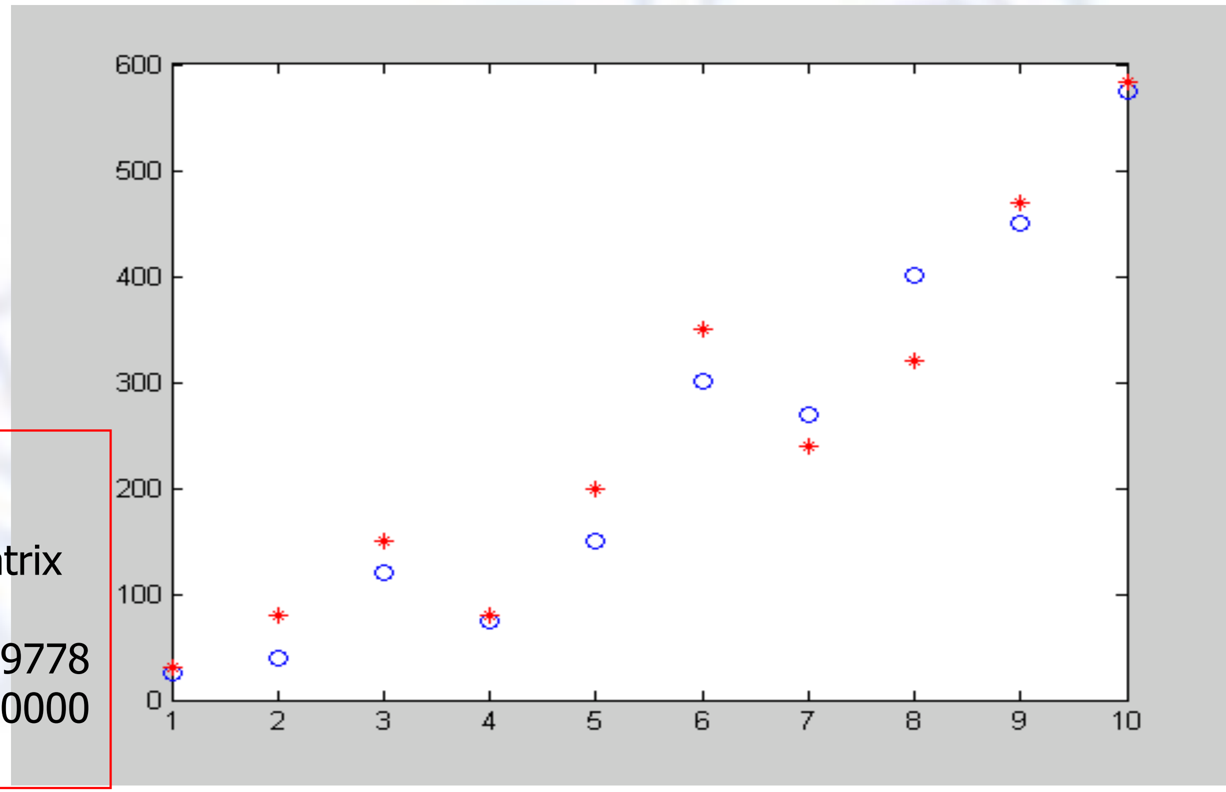
COV(X,1) or COV(X,Y,1) normalizes by N and produces the second moment matrix of the observations about their mean.

Matlab Implementation

```
CORRCOEF Correlation coefficients.
   R=CORRCOEF(X) calculates a matrix R of correlation coefficients for
   an array X, in which each row is an observation and each column is a
   variable.

   R=CORRCOEF(X,Y), where X and Y are column vectors, is the same as
   R=CORRCOEF([X Y]).

   If C is the covariance matrix, C = COV(X), then CORRCOEF(X) is
   the matrix whose (i,j)'th element is

          C(i,j)/SQRT(C(i,i)*C(j,j)).
```

Example

```
x(:,1)=[25, 40, 120, 75, 150, 300, 270, 400, 450, 575]';
x(:,2)=[30, 80, 150, 80, 200, 350, 240, 320, 470, 583]';
plot(x(:,1),'bo')
hold on
plot(x(:,2),'r*')
cov(x,1)
corrcoef(x)
```



Results:

Covariance matrix

1.0e+004 *

3.2237	3.0050
3.0050	2.9299

Correlation matrix

1.0000	0.9778
0.9778	1.0000

Autocorrelation 自相關

- Autocorrelation provides a description of how similar a waveform is to itself as various time shifts (time lags).

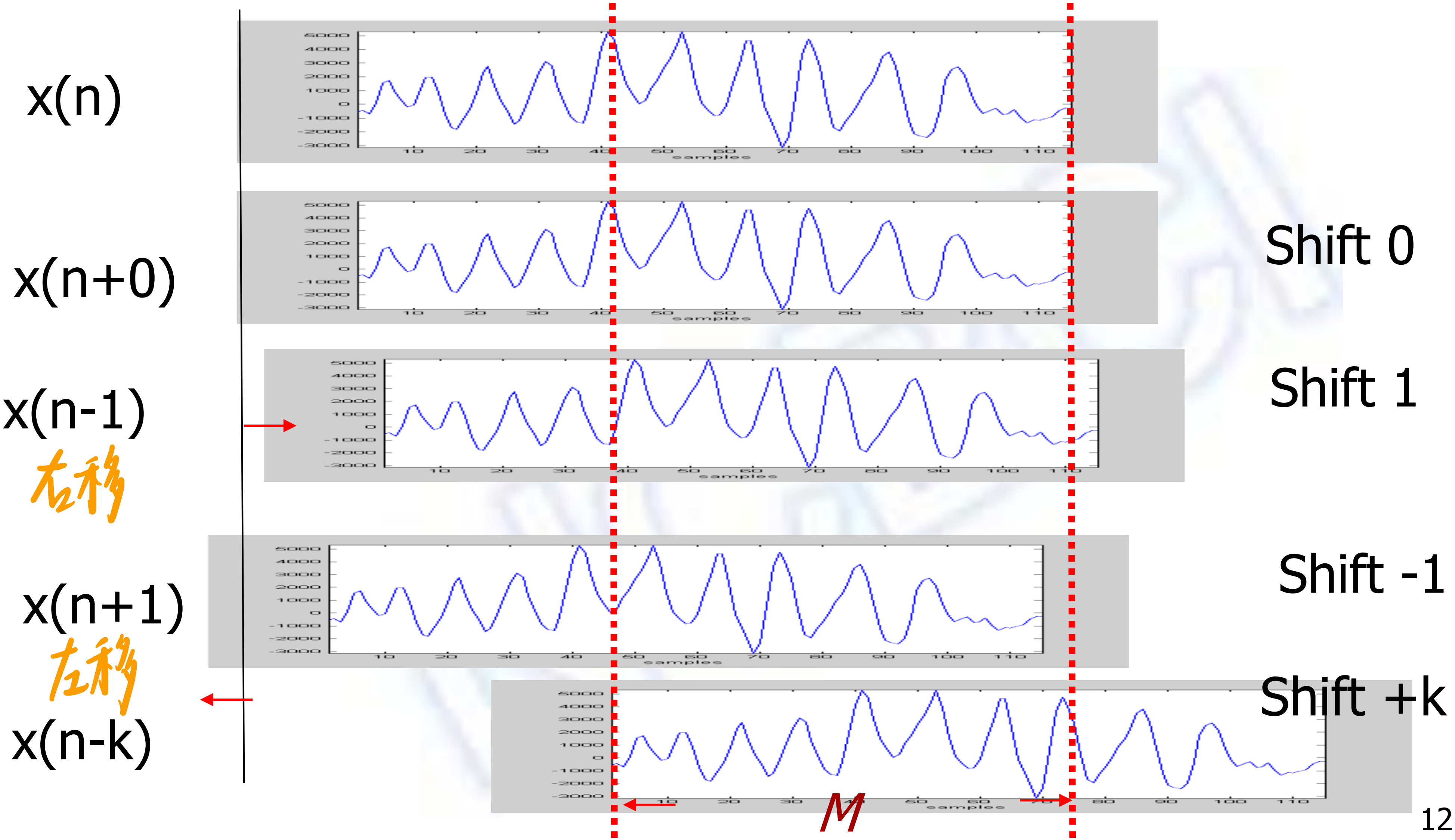
$$r_{xx}(k)=\sum_{n=1}^{M}x(n)\cdot x(n+k)\ \text{ or}$$

$$r_{xx}(k)=\frac{1}{M}\sum_{n=1}^{M}x(n)\cdot x(n+k)$$

M is the range of the available overlapped data

(P.78, Biomedical Signal Analysis, Rangaraj M. Rangayyan, Wiley, ISBN: 0-471-20811-62.)
(P.46-50, Biosignal and Biomedical Image Processing, John L. Semmlow, Marcel Dekker, ISBN: 0-8247-4803-4 3.)

Autocorrelation



Autocorrelation

- Usually the correlation at zero lag is 1(maximum value).
- The autocorrelation must be symmetric about $k=0$.

Crosscorrelation

- Crosscorrelation provides a description of how similar a waveform is to the other waveform as various time shifts (time lags).

$$r_{xy}(k) = \sum_{n=1}^{M} x(n) \cdot y(n+k), \text{ or}$$

$$r_{xy}(k) = \frac{1}{M} \sum_{n=1}^{M} x(n) \cdot y(n+k)$$

Autocorrelation

- Matlab Implementation xcorr

[C,LAGS] = XCORR(...) returns a vector of lag indices (LAGS).

C = XCORR(A,B), where A and B are length M vectors (M>1), returns the length 2*M-1 cross-correlation sequence C. If A and B are of different length, the shortest one is zero-padded.

XCORR(A), when A is a vector, is the auto-correlation sequence.

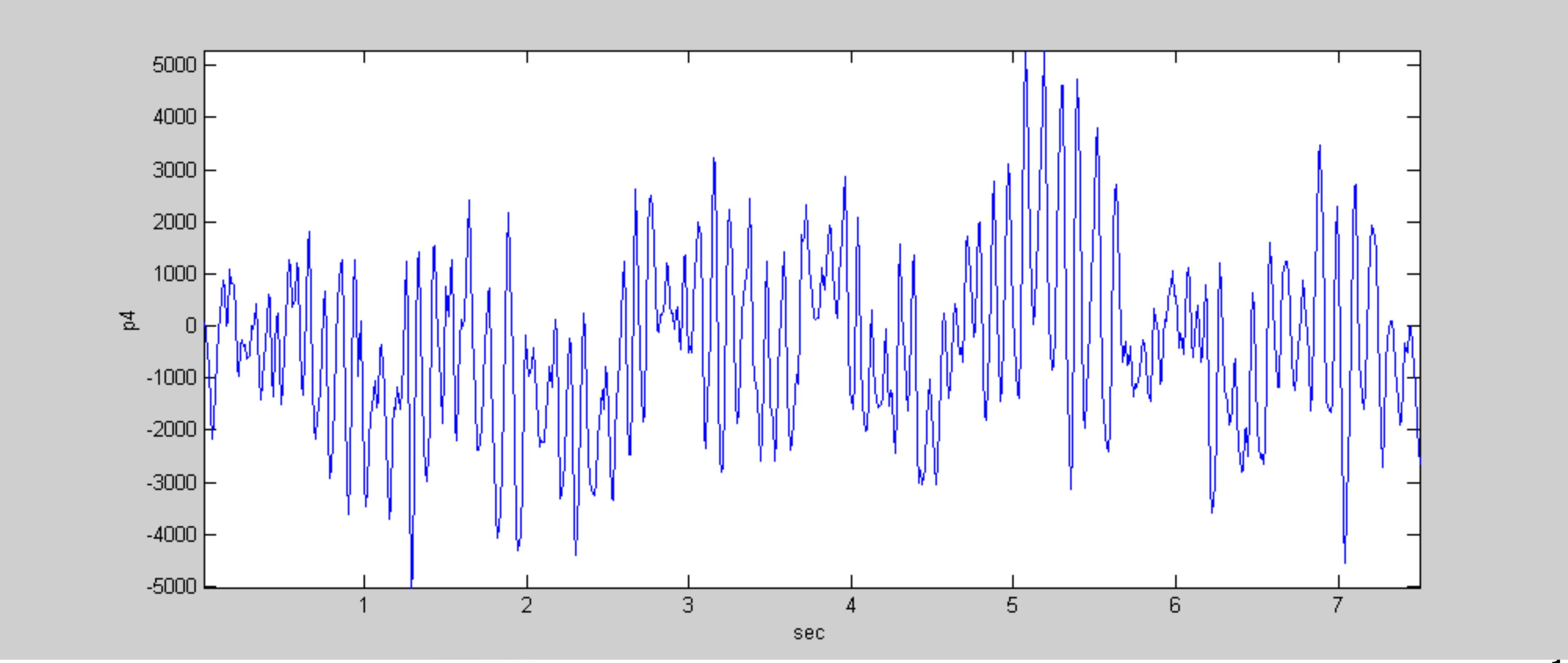
XCORR(...,SCALEOPT), normalizes the correlation according to SCALEOPT:

'coeff'   - normalizes the sequence so that the auto-correlations at zero lag are identically 1.0.

(P.46-50, Biosignal and Biomedical Image Processing, John L. Semmlow, Marcel Dekker, ISBN: 0-8247-4803-4 3.))

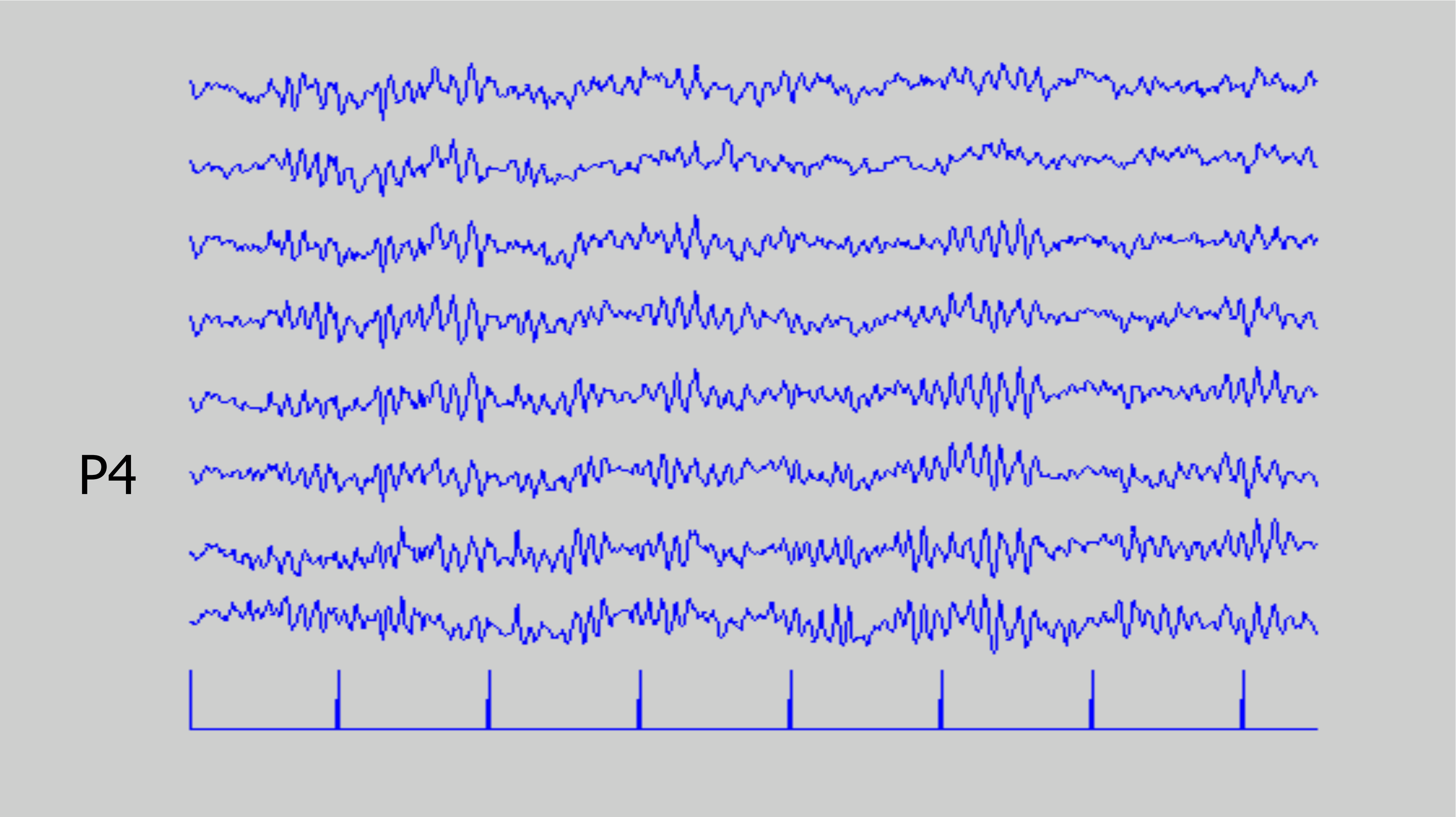
EEG Rhythm Detection

- How to develop a method to detect the EEG rhythms automatically?



(From Section 4.4 of Biomedical Signal Analysis, Rangaraj M. Rangayyan, Wiley, ISBN: 0-471-20811-62.)

8-channel EEG signals



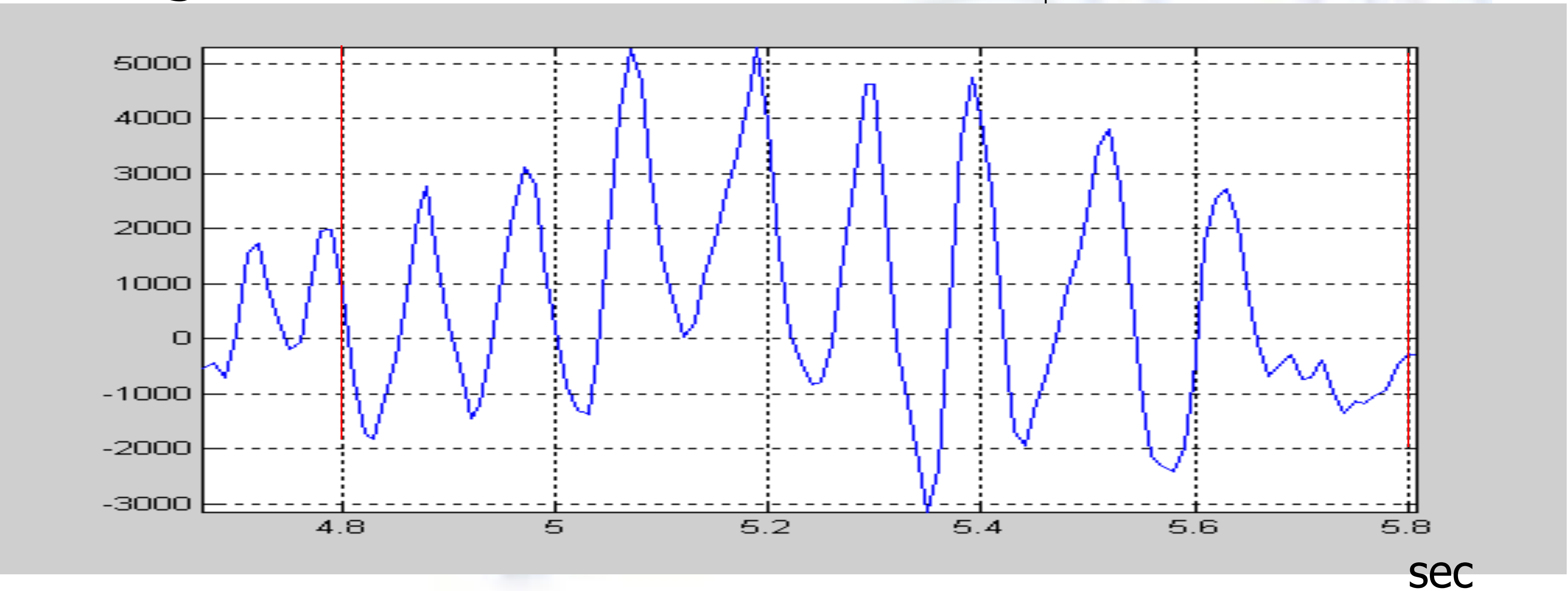
EEG Rhythm Detection

- To detect the rhythm of P4 channel
  - Manual calculation
  - Autocorrelation

Manual calculation

```
figure
plot(t(467:581),eegp4(467:581))
grid
axis tight
```

The EEG segment of the p4 during 4.67-5.81 sec.



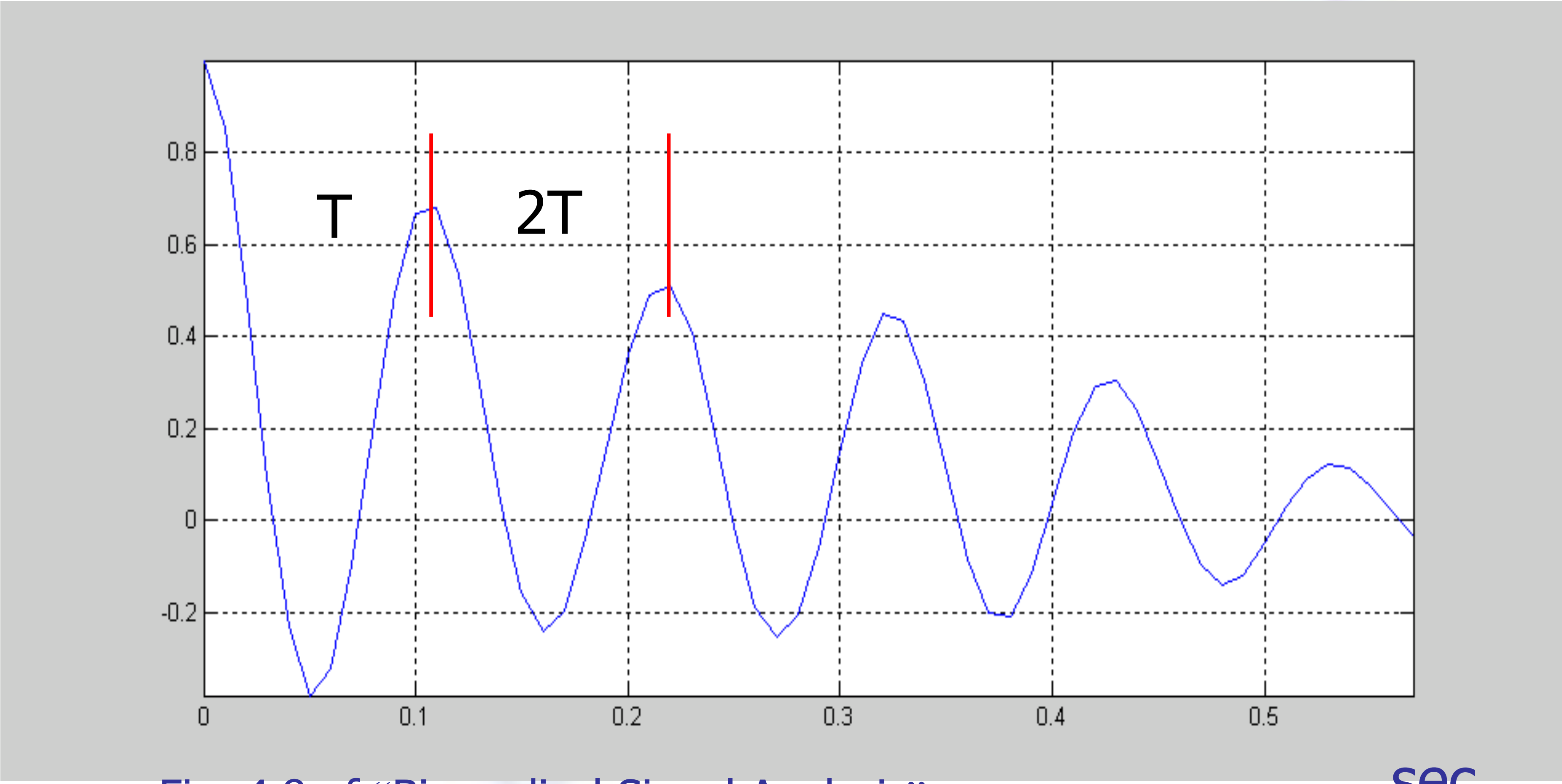
EEG Rhythm Detection by Autocorrelation

Autocorrelation of EEG segment of the p4 during 4.67-5.81 sec.

```
x=eegp4(467:581);
N=length(x);
[c lag]=xcorr(x,'coeff');
lag1= lag(N:(N+fix(N/2)));
c1=c(N:(N+fix(N/2)));
plot(lag1/fs,c1);
axis tight
grid
```

Take the autocorrelation during the lags from 0~(N/2)

Result of Autocorrelation

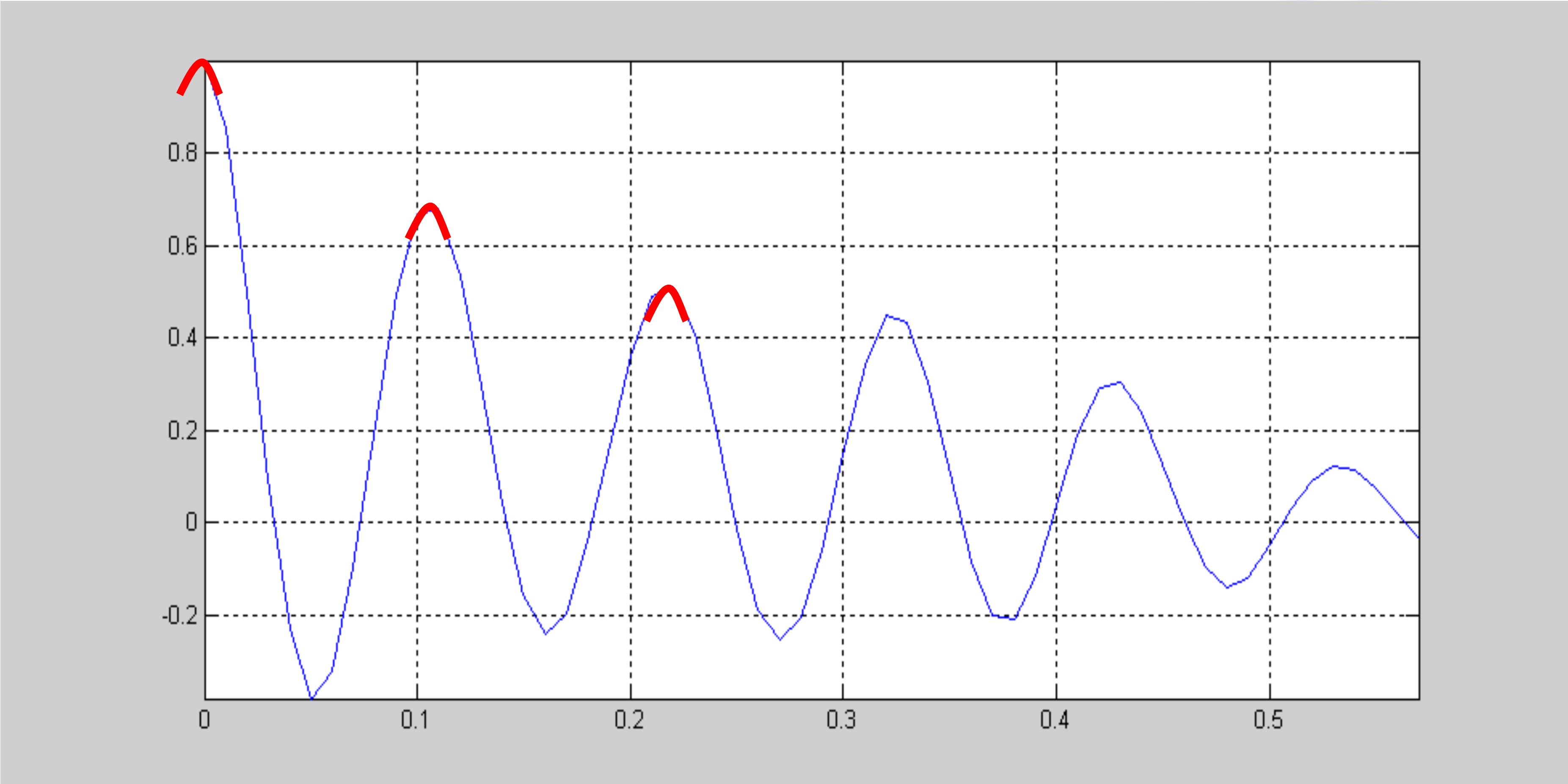


sec

Fig. 4.8 of "Biomedical Signal Analysis"

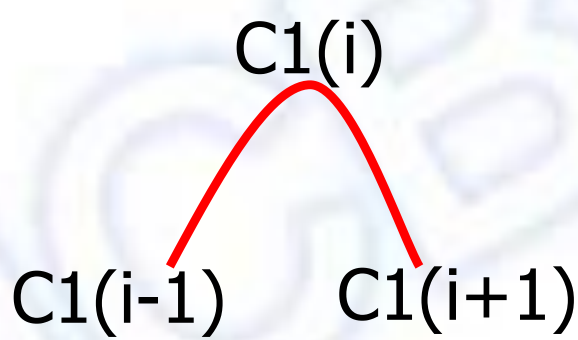
Peak Search

A simple peak-search algorithm may be applied to the ACF to detect the shift period



Peak Search

```
peak=0;
for i=2:length(c1)-1
 if c1(i)>0 & c1(i-1)<c1(i) & c1(i+1)<c1(i) & peak<c1(i)
  index=i;
  peak=c1(i);
 end
end

T= (index-1)/fs
f=1/T
```



Results

- T=0.11 sec
- f=1/T=9.09 Hz

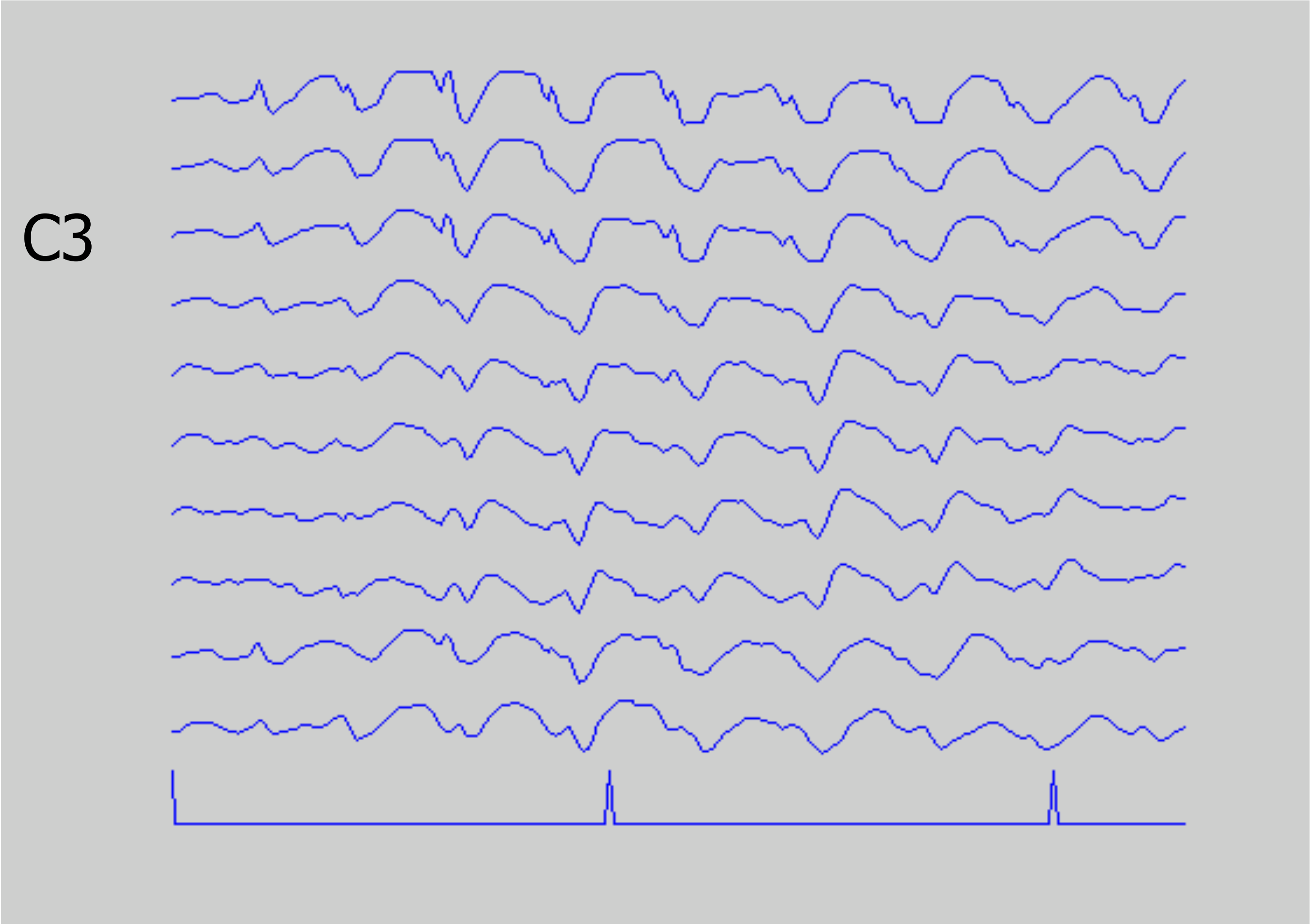
The rhythm of EEG at P4 is α

Template Matching

- Generate event template (a wave shape) and use the template to extract the signals portions corresponding to the event by cross-correlation.
- Example: detect spike-and-wave complexes in an EEG signal. (Assume a sample segment of a spike-and wave complex is available)

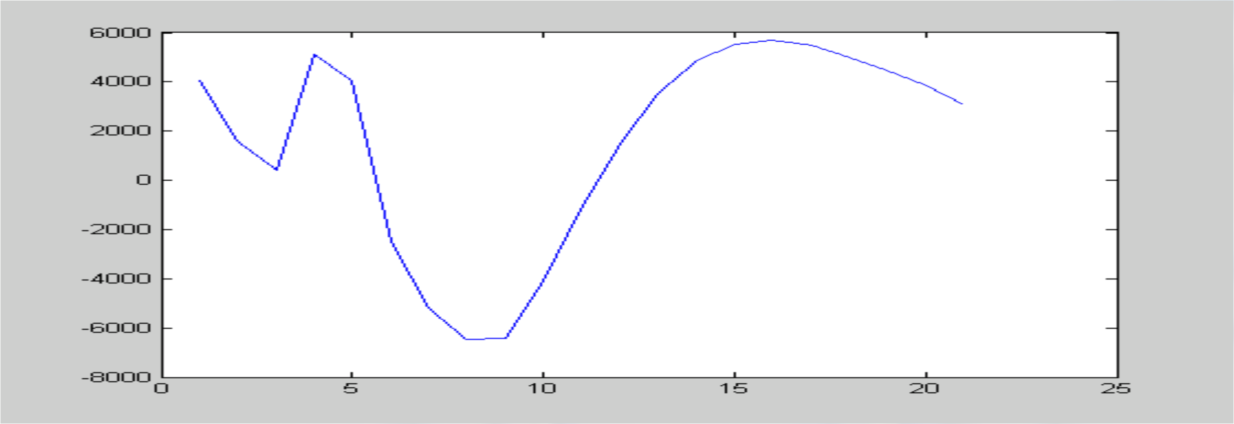
(3.3.1 &4.4.2, Biomedical Signal Analysis, Rangaraj M. Rangayyan, Wiley, ISBN: 0-471-20811-62.)

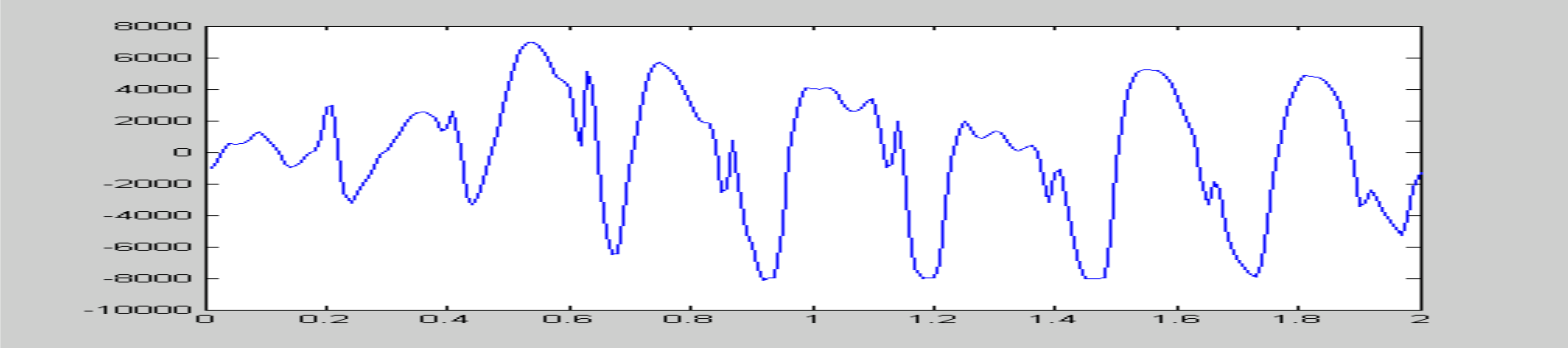
10-channel EEG signals



Spike-and-Wave Detection

Spike-and-wave: a sharp spike followed by a wave with a frequency of about 3Hz





Spike-and-Wave Detection

Using the EEG of c3 between 0.6 s to 0.82 s to detect the spike and wave patterns

```
x=eegc3(1:200);
y=eegc3(60:80);
N=length(x);
[c lag]=xcorr(x,y);
c=c/max(c)
lag1= lag(N:(2*N-1));
c1=c(N:(2*N-1));
plot(lag1/fs,c1);
axis tight
grid
```

Spike-and-Wave Detection