Basic Signal Processing (II)

Computational Social Intelligence - Lecture 17

Prof. Alessandro Vinciarelli
School of Computing Science &
Institute of Neuroscience and Psychology

http://www.dcs.gla.ac.uk/vincia
Alessandro.Vinciarelli@glasgow.ac.uk

University of Glasgow
EPSRC
Engineering and Physical Sciences Research Council
FNSNF

Texts (see Moodle)

This lecture is based on the following text (available on Moodle):

- F.Camastra and A.Vinciarelli, "Machine Learning for Audio, Image and Video Processing", Springer Verlag, Chapter 2, pp. 38-46, 2008.

Outline

- Quick Recap
- Zero Crossing Rate
- Autocorrelation
- Fourier Transform
- Conclusion

Outline

- Quick Recap
- Zero Crossing Rate
- Autocorrelation
- Fourier Transform
- Conclusion

Any short-term property that can be extracted from the signal

A function of sample "n"

$$x[k] = \sum_{n=-\infty}^{\infty} f(s[n])w[k-n]$$

A window signal

Convolution between f(s[n]) and a window

The Energy

$$f(s[n]) = \frac{(s[n])^2}{N}$$

$$f(s[n]) = \frac{|s[n]|}{N}$$

The Magnitude

Recap

- The short-term properties of a signal can be calculated through the convolution with an analysis window;
- Different functions applied to the samples lead to different properties;
- Every property is a signal that tends to remain stable in those intervals in which the signal properties are stable.

Outline

- Quick Recap
- Zero Crossing Rate
- Autocorrelation
- Fourier Transform
- Conclusion

The function is 1 when the signal crosses the horizontal axis

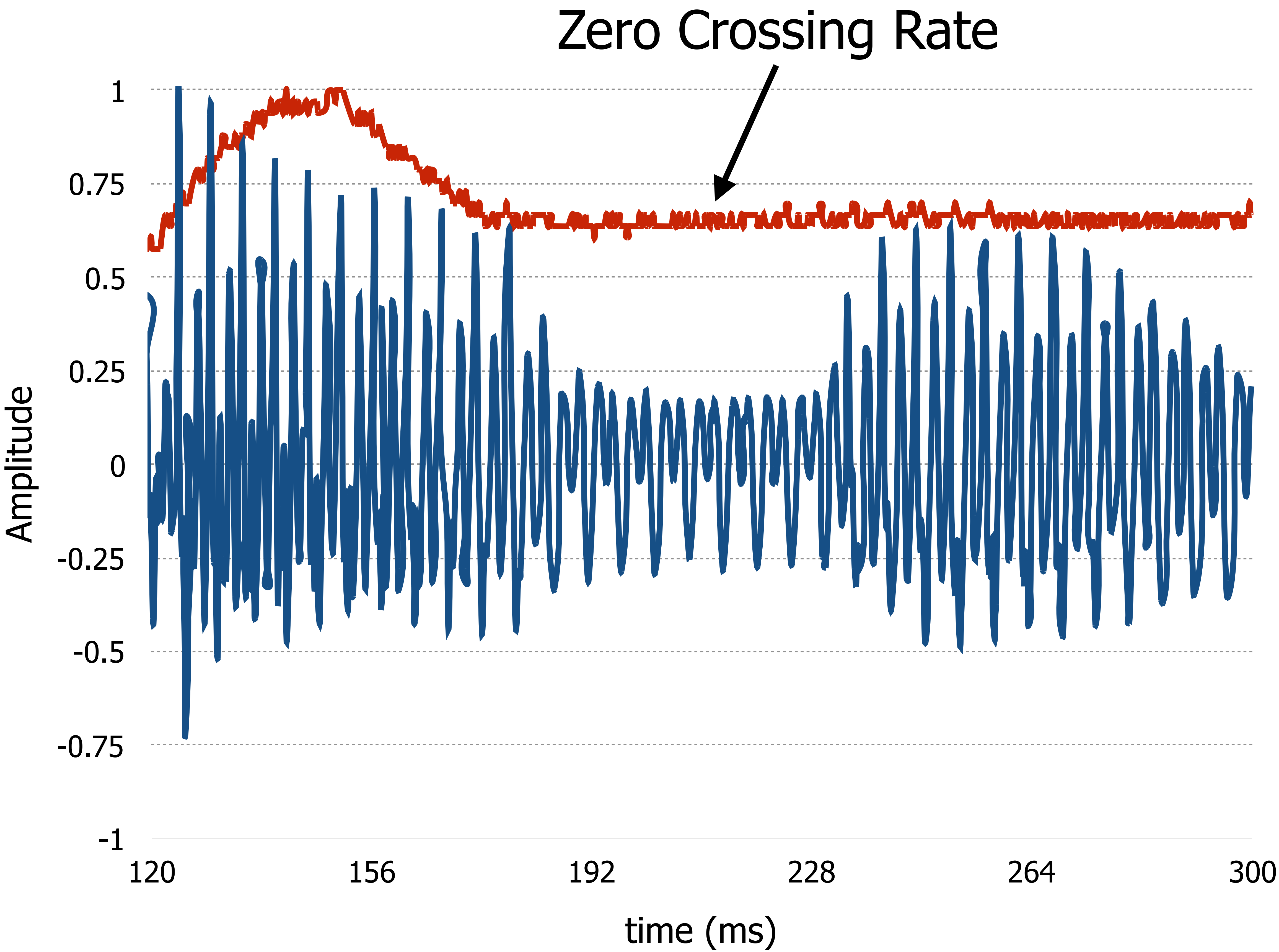
$$g(s[k], s[k-1]) =$$

$$= \begin{cases} 0 & : \quad s[k], s[k-1] \geq 0 \\ 0 & : \quad s[k], s[k-1] \leq 0 \\ 1 & : \quad s[k] > 0, s[k-1] < 0 \\ 1 & : \quad s[k] < 0, s[k-1] > 0 \end{cases}$$

The Zero Crossing Rate

$$Z[k] = \sum_{n=-\infty}^{\infty} \frac{g(s[n], s[n-1])}{2N} w[k-n]$$

The ZCR is half the number of times the signal crosses the horizontal axis in the window



Zero Crossing Rate
Amplitude
1
0.75
0.5
0.25
0
-0.25
-0.5
-0.75
-1
120
156
192
228
264
300
time (ms)

The ZCR when the signal is a <u>sinusoid of frequency "f"</u>

The sinusoid frequency

$$Z[k] = \frac{2f}{F}$$

The sampling frequency

Outline

- Quick Recap
- Zero Crossing Rate
- Autocorrelation
- Fourier Transform
- Conclusion

The parameter "m" is called the lag

The Autocorrelation

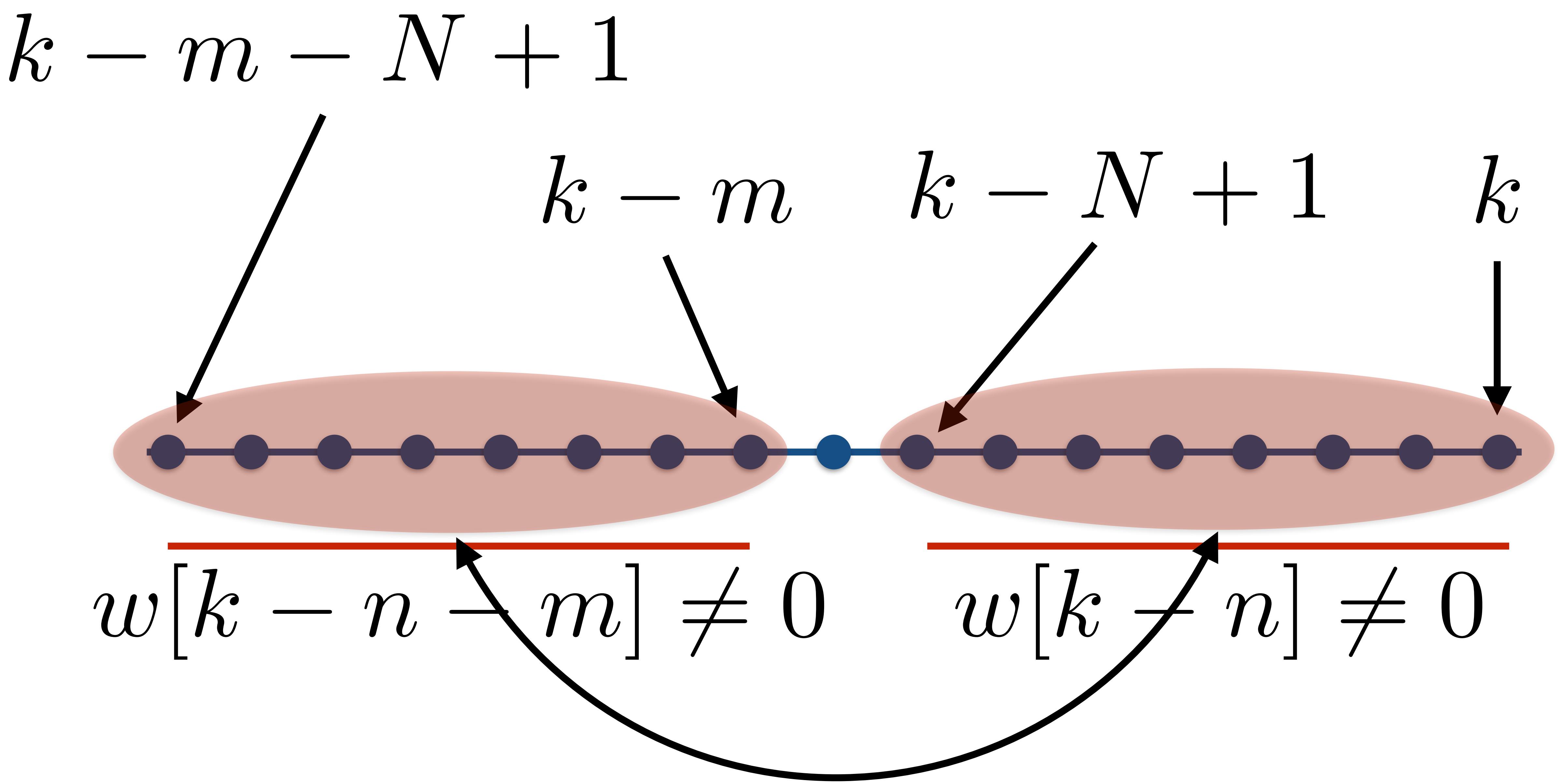
$$R_m[k] = \sum_{n=-\infty}^{\infty} s[n]w[k-n]s[n+m]w[k-n-m]$$

Product of two samples at distance "m" from one another

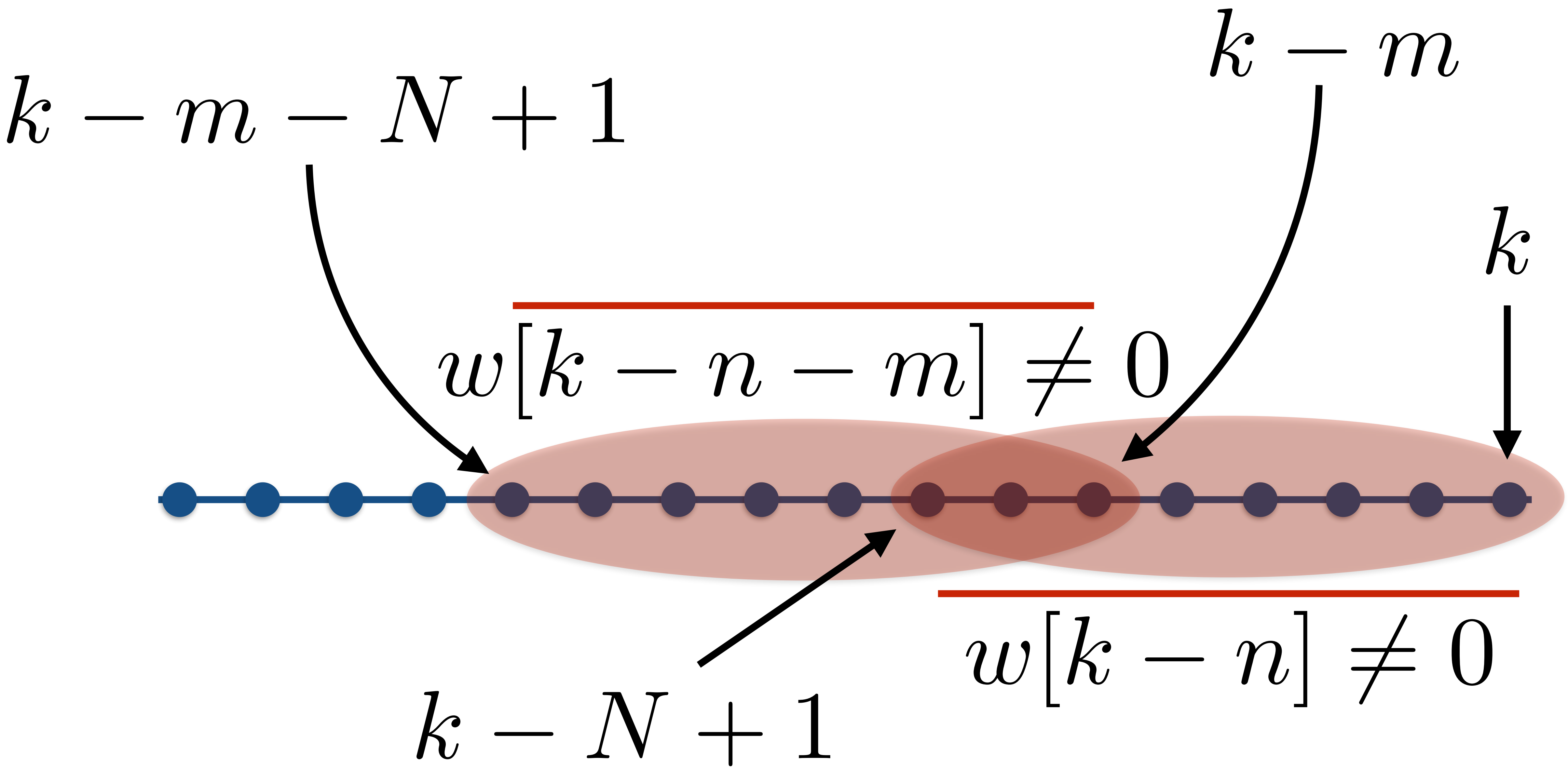
Index range where
"w[k-n]" is different
from zero

$$
\begin{aligned}
k - n &\geq 0 && \Rightarrow && n \leq k \\
k - n &\leq N - 1 && \Rightarrow && n \geq k - N + 1 \\
k - n - m &\geq 0 && \Rightarrow && n \leq k - m \\
k - n - m &\leq N - 1 && \Rightarrow && n \geq k - m - N + 1
\end{aligned}
$$

Index range where
"w[k-n-m]" is different
from zero



when "m" is greater than "N-1" ("N" is the length of the window), there is no overlapping between analysis windows



when “m” is less than “N” (“N” is the length of the window), there is overlapping between analysis windows

The lag is zero

The Autocorrelation

Only when the window is rectangular

$$R_0[k] = \sum_{n=-\infty}^{\infty} s[n]s[n]w[k-n]w[k-n] = \sum_{n=-\infty}^{\infty} (s[n])^2 w[k-n] = NE[k]$$

The autocorrelation with lag zero is the energy

Recap

- The <u>Zero Crossing Rate</u> provides information about the <u>frequency</u> that carries <u>most energy</u>;
- The <u>autocorrelation changes</u> depending on how <u>periodic</u> is the signal;
- The <u>inverse of the distance between the maxima</u> of the autocorrelation accounts for the <u>frequency</u> that carries <u>most energy</u>;

Outline

- Quick Recap
- Zero Crossing Rate
- Autocorrelation
- Fourier Transform
- Conclusion

Disclaimer

The Section "Fourier Transform" is not part of the exam (it is presented only for information). However, in case of interest, further information is available in the following text:

- F.Camastra and A.Vinciarelli, "Machine Learning for Audio, Image and Video Processing", Springer Verlag, Appendix B, pp. 38-51, 2008.

Discrete Fourier Transform

The sum is over the samples of an analysis window

$$X[k] = \sum_{n=0}^{N-1} s[n] \cdot e^{-i2\pi \frac{k}{N} n}$$

The weights of the sum are complex exponentials

Discrete Fourier Transform

The sum is over the samples of an analysis window

$$X[k] = \sum_{n=0}^{N-1} s[n] \cdot e^{-i 2\pi \frac{k}{N} n} =$$

$$= \sum_{n=0}^{N-1} s[n] \left\{ \cos\left(2\pi \frac{k}{N} n\right) - i \sin\left(2\pi \frac{k}{N} n\right) \right\}$$

The weights are periodic functions

This exponential is equal to 1

$$X[k+N] = \sum_{n=0}^{N-1} s[n] \cdot e^{-i2\pi \frac{k+N}{N} n} =$$

$$= \sum_{n=0}^{N-1} s[n] \cdot e^{-i2\pi \frac{k}{N} n} e^{-i2\pi n} =$$

$$= \sum_{n=0}^{N-1} s[n] \cdot e^{-i2\pi \frac{k}{N} n} = X[k]$$

"X[k]" is periodic of period "N"

Inverse Fourier Transform

The sum is over the samples of a period of the DFT

$$s[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k] \cdot e^{i2\pi \frac{k}{N} n} =$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} X[k] \left\{ \cos\left(2\pi \frac{k}{N} n\right) + i \sin\left(2\pi \frac{k}{N} n\right) \right\}$$

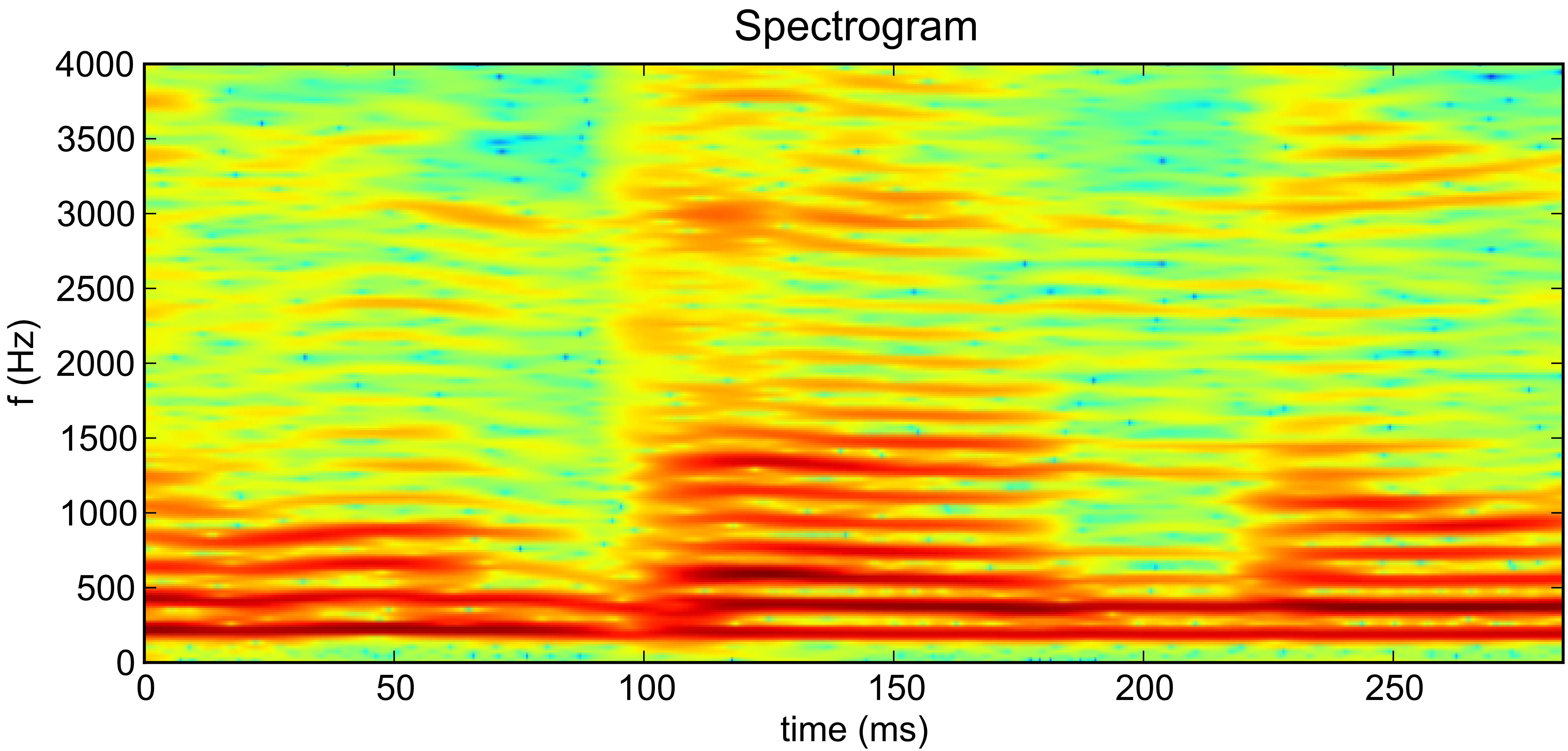
The weights are periodic functions

Sample "n" is the signal at time "nT" (T is the sampling period)

$$s(nT) = \frac{1}{N}\sum_{k=0}^{N-1} X[k]\cdot e^{i2\pi\frac{k}{N}nT} =$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} X[k]\cdot e^{i2\pi\frac{k}{N}\frac{1}{F}n}$$

The sampling rate



Spectrogram
4000
3500
3000
2500
2000
1500
1000
500
0
f (Hz)
0
50
100
150
200
250
time (ms)

Recap

- Discrete Fourier Transform and its inverse represent the <u>same information in two different ways</u> (they are equivalent);
- The spectrogram shows the <u>energy distribution</u> across the frequencies;
- The Discrete Fourier Transform <u>changes</u> depending on the <u>property</u> of the signal.

Outline

- Quick Recap
- Zero Crossing Rate
- Autocorrelation
- Fourier Transform
- Conclusion

Conclusions

- Speech signals can be analysed in the time domain through convolution operations;
- In most cases, the processing takes place in the frequency domain (after performing Fourier transform);
- The main reason for analysing speech is that it is the main form of communication between people.