

Basic Signal Processing (I)

Computational Social Intelligence - Lecture 16

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



Texts (see Moodle)

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

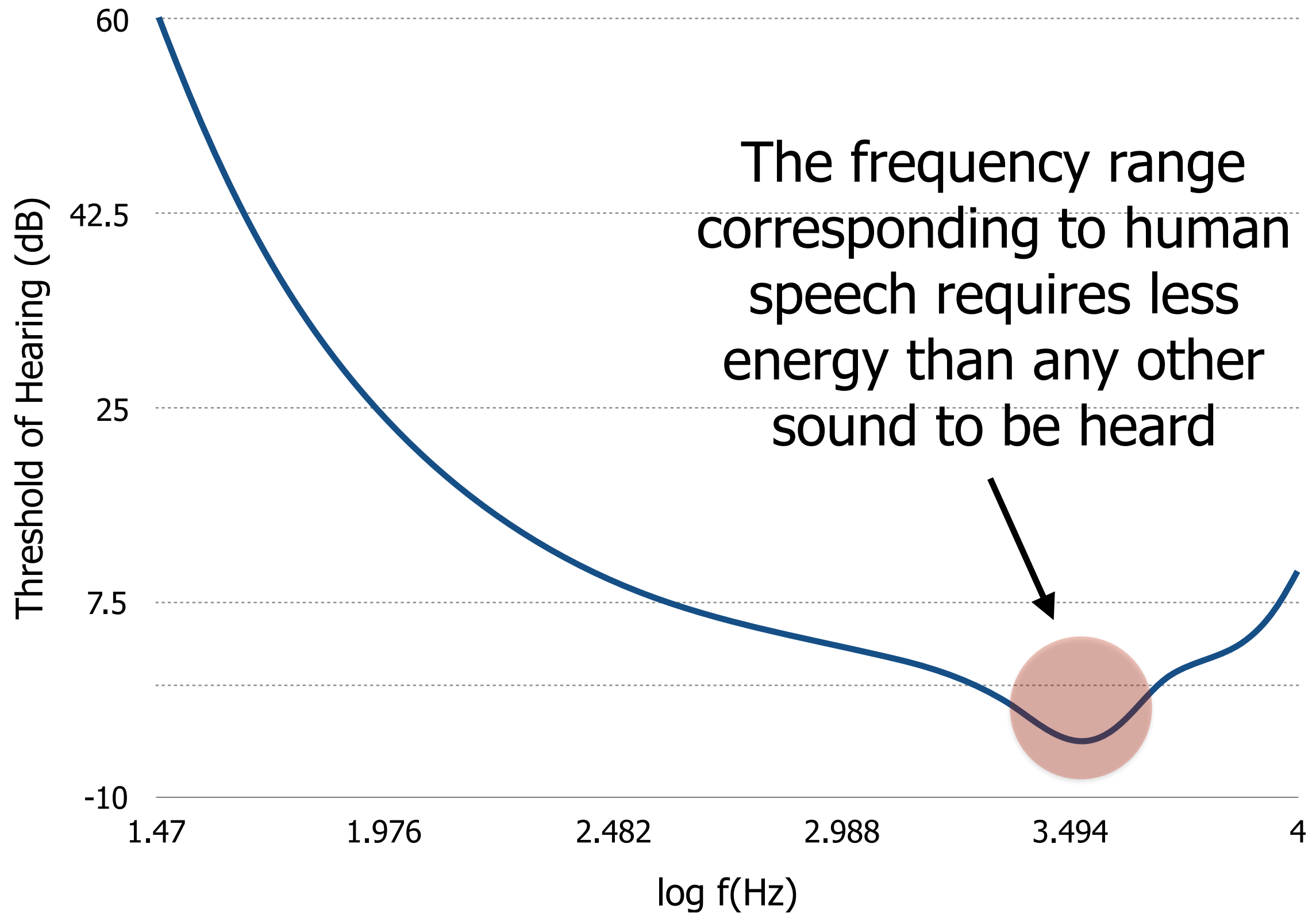
- F.Camastra and A.Vinciarelli, "Machine Learning for Audio, Image and Video Processing", Springer Verlag, Chapter 2, pp. 38-46, 2008.

Outline

- Introduction
- Time Domain Processing
- Short-Term Analysis
- Conclusions

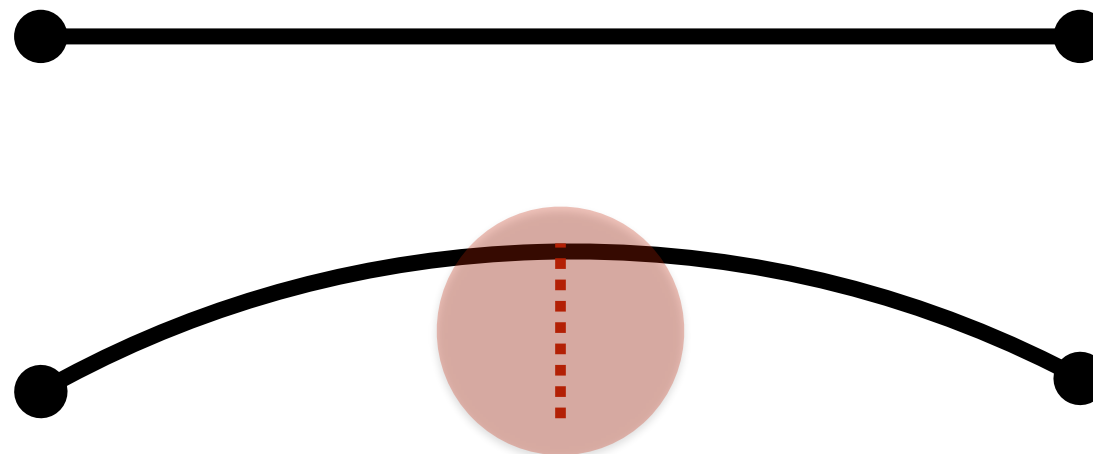
Outline

- Introduction
- Time Domain Processing
- Short-Term Analysis
- Conclusions



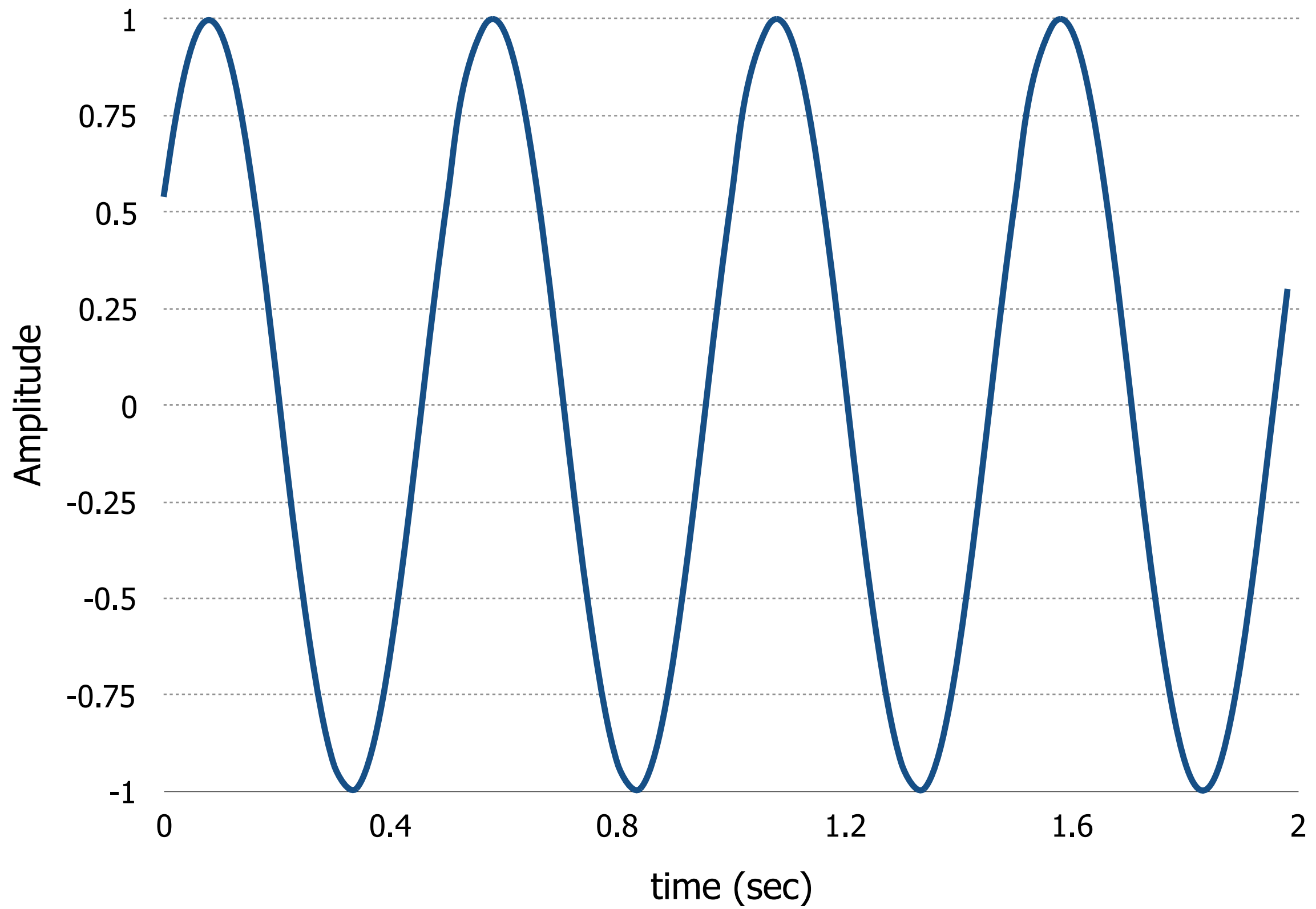
A microphone measures
at regular time steps
the oscillations of an
elastic membrane

When there is no
acoustic wave, the
membrane is in
equilibrium

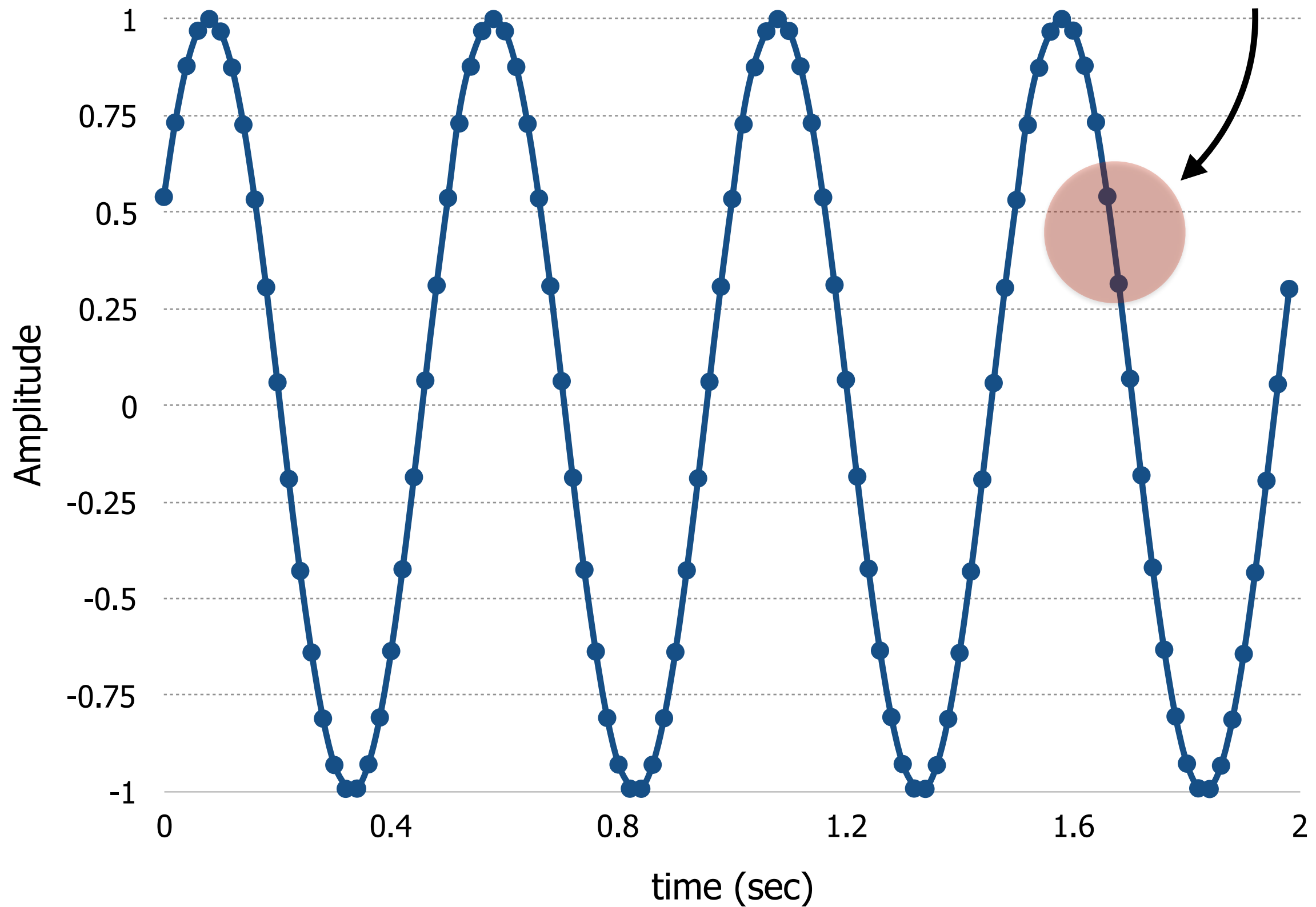


When there is an
acoustic wave, the
membrane oscillates

The amplitude of the
oscillation is
proportional to the
amplitude of the
acoustic wave



The signal is unknown
between two samples



The sample "k"

The signal at time "kT"

$$s[k] = s(kT)$$

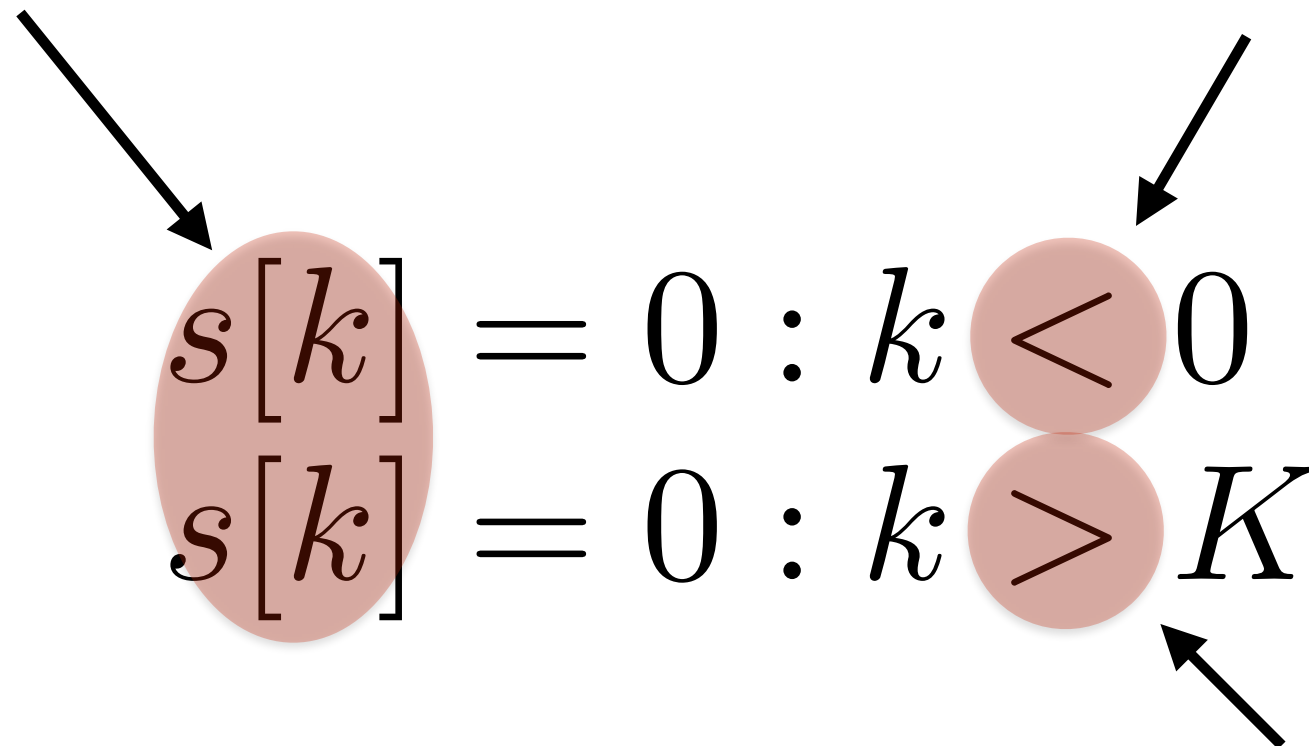
$$k = -\infty, \dots, \infty$$

The index of the sample

The sampling period
(time interval between
consecutive samples)

There are no infinite signals in reality

All samples with $k < 0$ are conventionally null

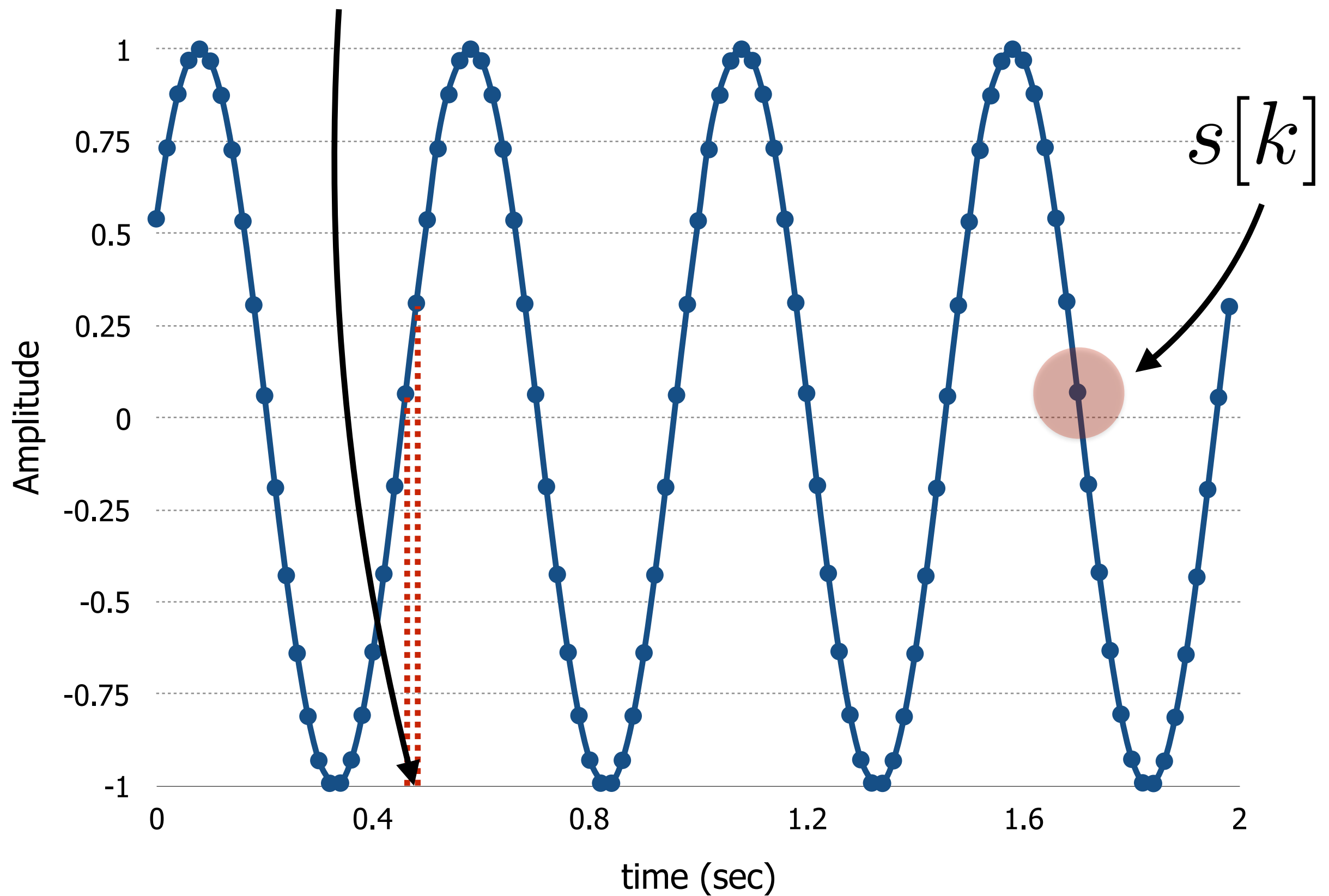


The diagram shows two equations for a signal $s[k]$. The first equation is $s[k] = 0 : k < 0$, where the expression $k < 0$ is enclosed in a red oval. An arrow points from the text 'There are no infinite signals in reality' to this oval. The second equation is $s[k] = 0 : k > K$, where the expression $k > K$ is enclosed in a red oval. Two arrows point from the text 'All samples with $k < 0$ are conventionally null' to the first oval, and another arrow points from the text 'All samples with $k < K$ (where K is the time the last measurement is done) are conventionally null' to this second oval.

$$s[k] = 0 : k < 0$$
$$s[k] = 0 : k > K$$

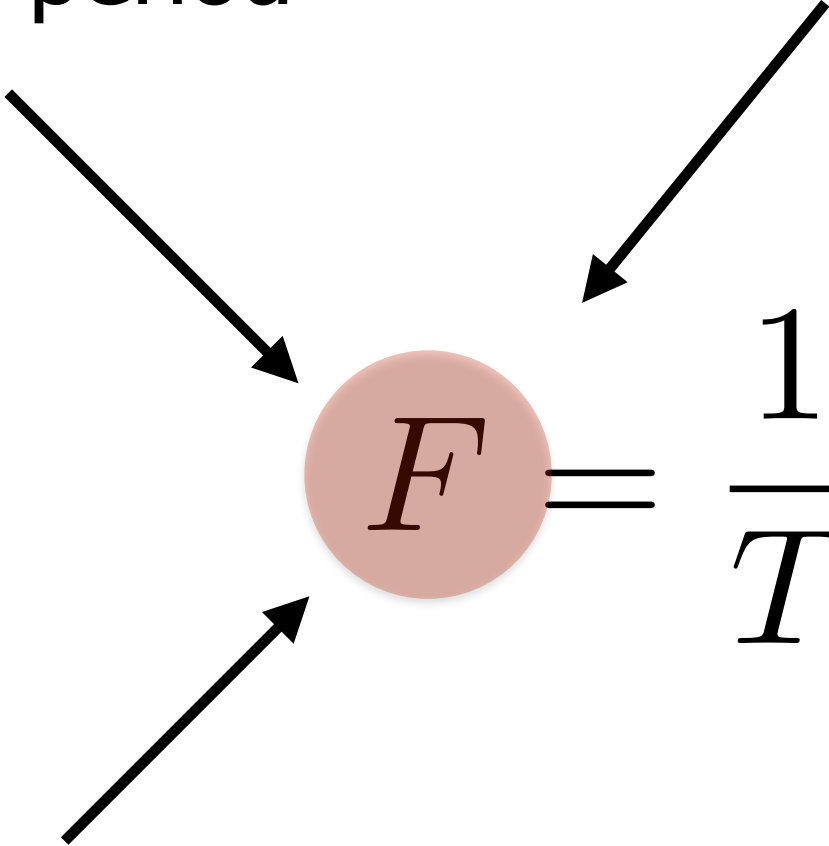
All samples with $k < K$
(where K is the time
the last measurement is
done) are
conventionally null

The sampling period



The sampling frequency
is the inverse of the
sampling period

It must be at least twice
as much as the highest
frequency expected to
be observed in the
signal


$$F = \frac{1}{T}$$

It is the number of
samples per unit of time

Recap

- A signal is a measurable physical quantity that changes over time and it is continuous;
- A digital signal is a sequence of physical measurements collected at regular time steps (the sampling period) and it is discrete;
- Nothing can be said about what happens in the signal between two consecutive samples.

Outline

- Introduction
- **Time Domain Processing**
- Short-Term Analysis
- Conclusions

The convolution
between the signal "s"
and the window "w"

The sum extends over
all samples of the signal

The diagram shows the convolution equation $y[k] = \sum_{n=-\infty}^{\infty} s[n]w[k-n]$. The summation symbol \sum is enclosed in a red circle, with ∞ above and $n=-\infty$ below it. The terms $s[n]$ and $w[k-n]$ are also enclosed in red circles. Arrows point from descriptive text to these elements: 'The signal' points to $s[n]$, 'The window' points to $w[k-n]$, and 'The sum extends over all samples of the signal' points to the summation symbol. 'The convolution between the signal "s" and the window "w"' points to the entire equation.

$$y[k] = \sum_{n=-\infty}^{\infty} s[n]w[k-n]$$

The signal

The window

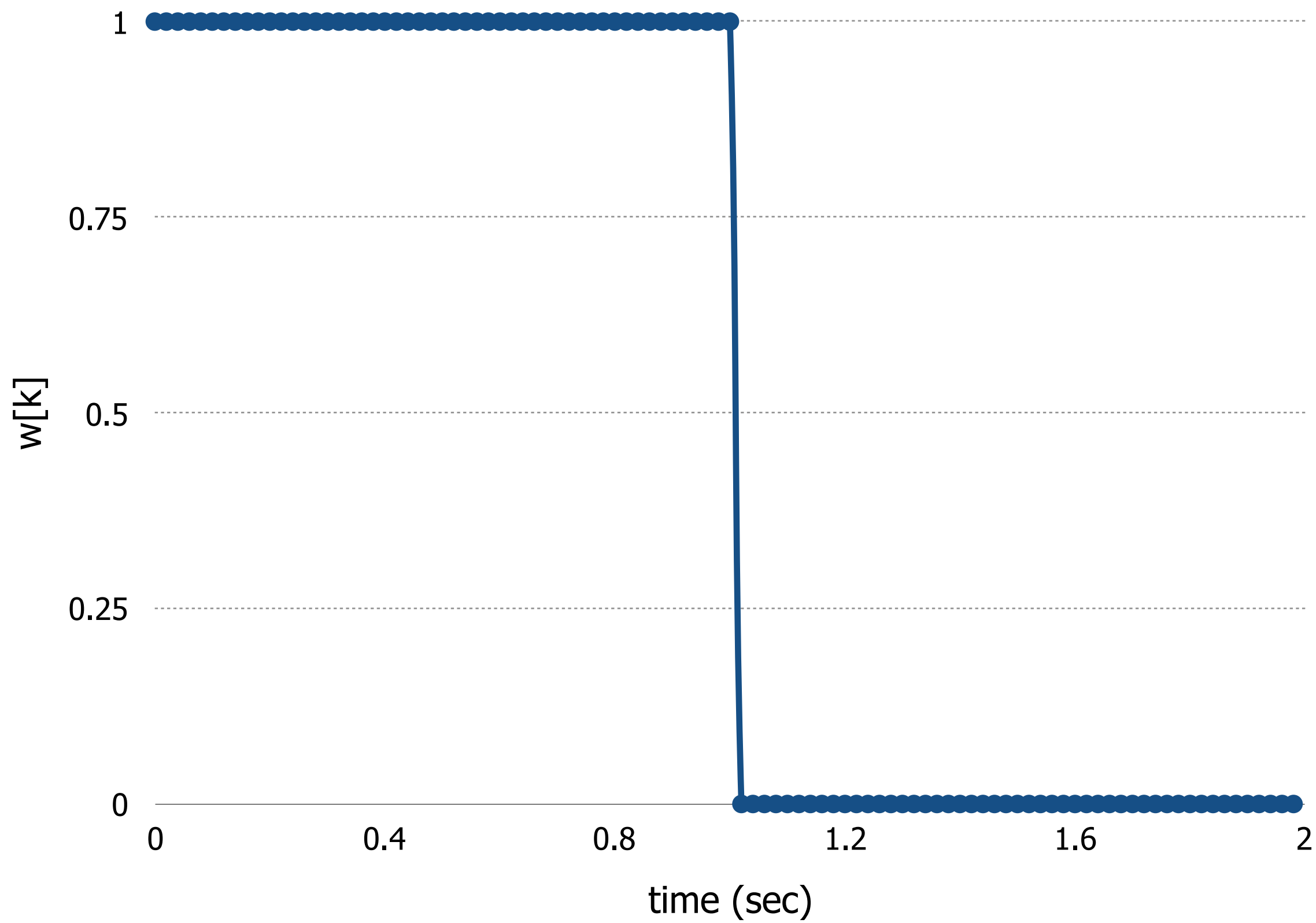
The rectangular window

$$w[n] = \begin{cases} 1 & : 0 \leq n \leq N-1 \\ 0 & : n < 0 \\ 0 & : n > N-1 \end{cases}$$

When the index is negative, the window signal is null

The window length

When the index is larger than the length, the window signal is null



The convolution
between the signal "s"
and the window "w"

The sum extends over
all samples of the signal

$$y[k] = \sum_{n=-\infty}^{\infty} s[n] w[k-n]$$

The window

The signal

Any product in which
"n" is less than or equal
to "k" is null



$$k - n \geq 0 \Rightarrow n \leq k$$

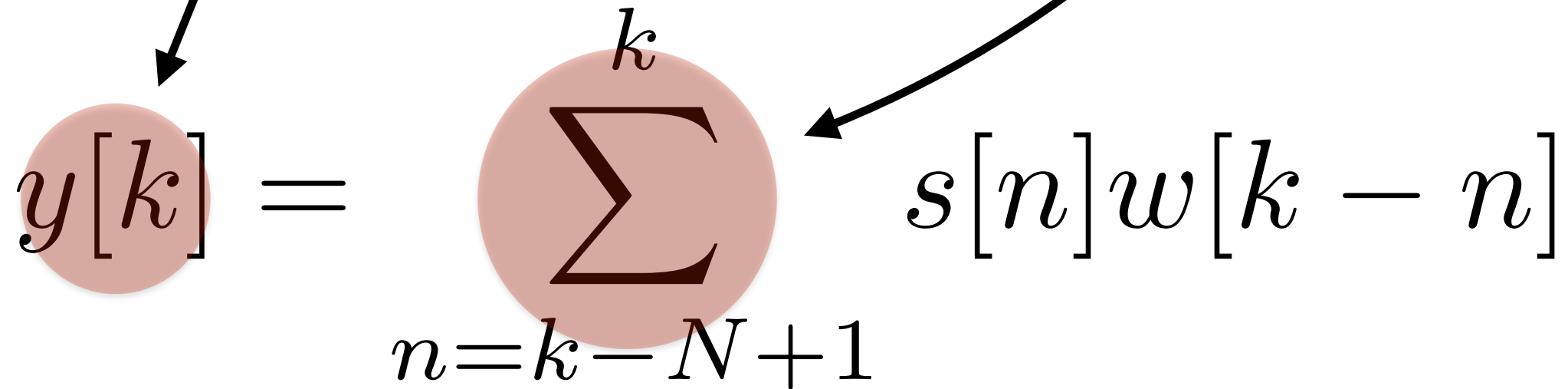
$$k - n \leq N - 1 \Rightarrow n \geq k - N + 1$$



Only products in which
"n" is greater than or
equal to "k-N+1" is null

The convolution
between the signal "s"
and the window "w"

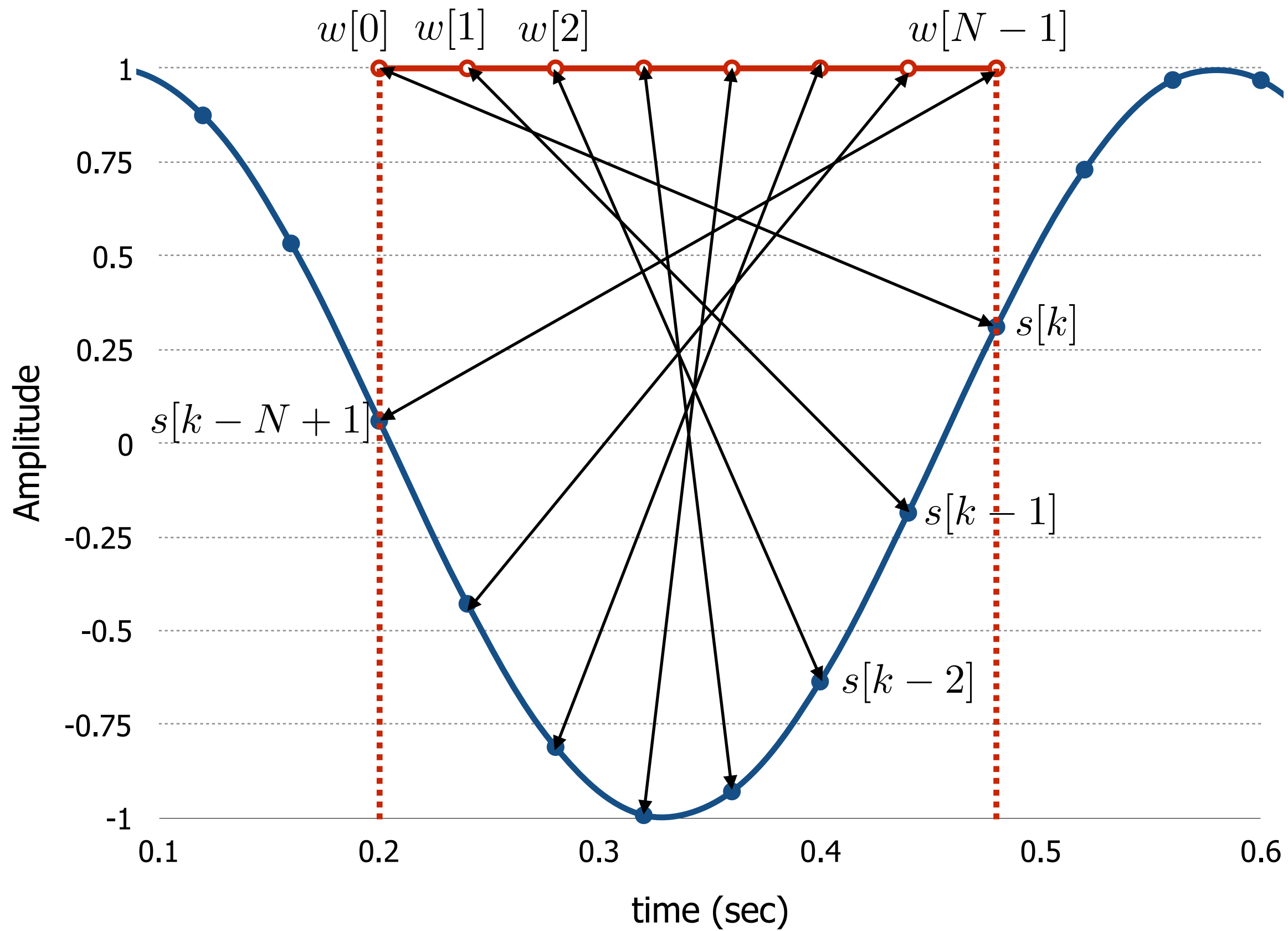
The sum extends over a
finite number of
samples



The diagram shows the convolution equation $y[k] = \sum_{n=k-N+1}^k s[n]w[k-n]$. The term $y[k]$ is enclosed in a light red circle, and the summation symbol Σ is also enclosed in a light red circle. An arrow points from the text "The convolution between the signal 's' and the window 'w'" to the $y[k]$ circle. Another arrow points from the text "The sum extends over a finite number of samples" to the Σ circle. The summation limits are $n=k-N+1$ at the bottom and k at the top.

$$y[k] = \sum_{n=k-N+1}^k s[n]w[k-n]$$

$$\begin{aligned}
y[k] &= \sum_{n=k-N+1}^k s[n]w[k-n] = \\
&= s[k-N+1]w[N-1] + \\
&+ s[k-N+2]w[N-2] + \dots \\
&+ \dots + s[k-1]w[1] + s[k]w[0]
\end{aligned}$$



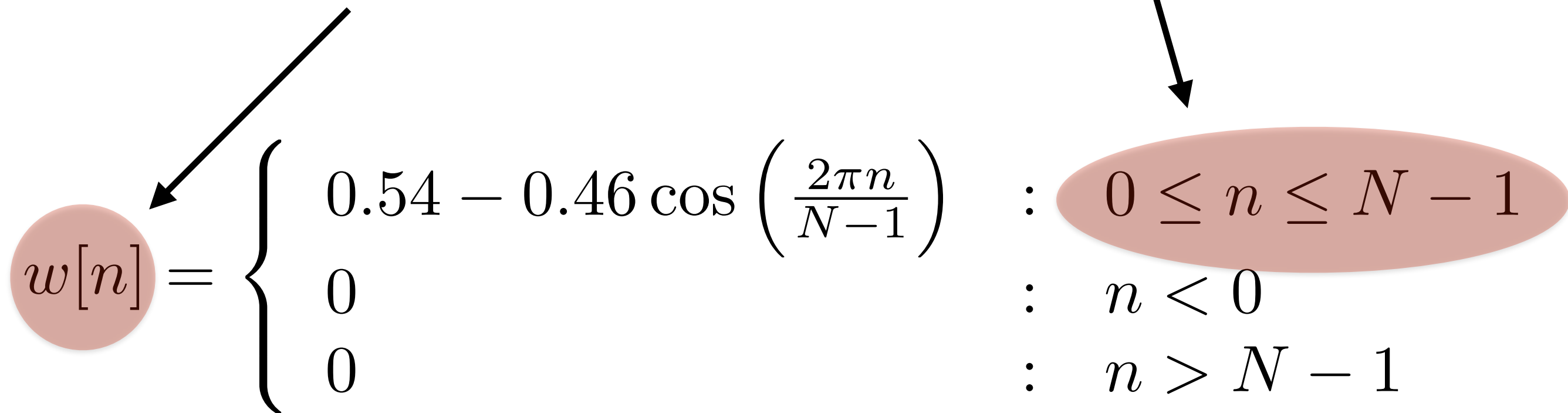
The convolution between the signal "s" and the rectangular window "w"

Sample "k" of the convolution is the sum of the samples between "k-N+1" and "k"

$$y[k] = \sum_{n=k-N+1}^k s[n]$$

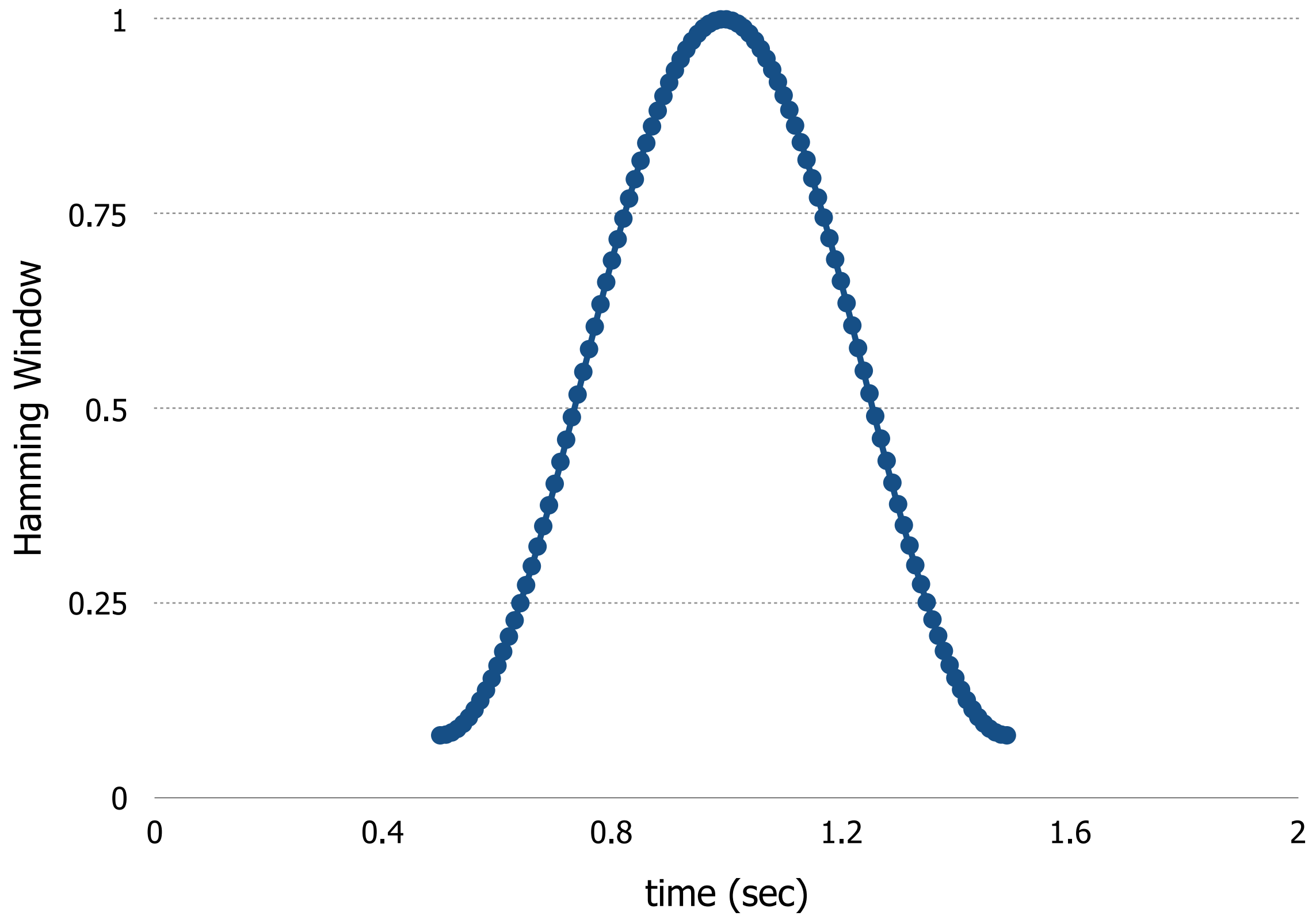
Only samples for which
"n" is between 0 and
"N-1" are not null

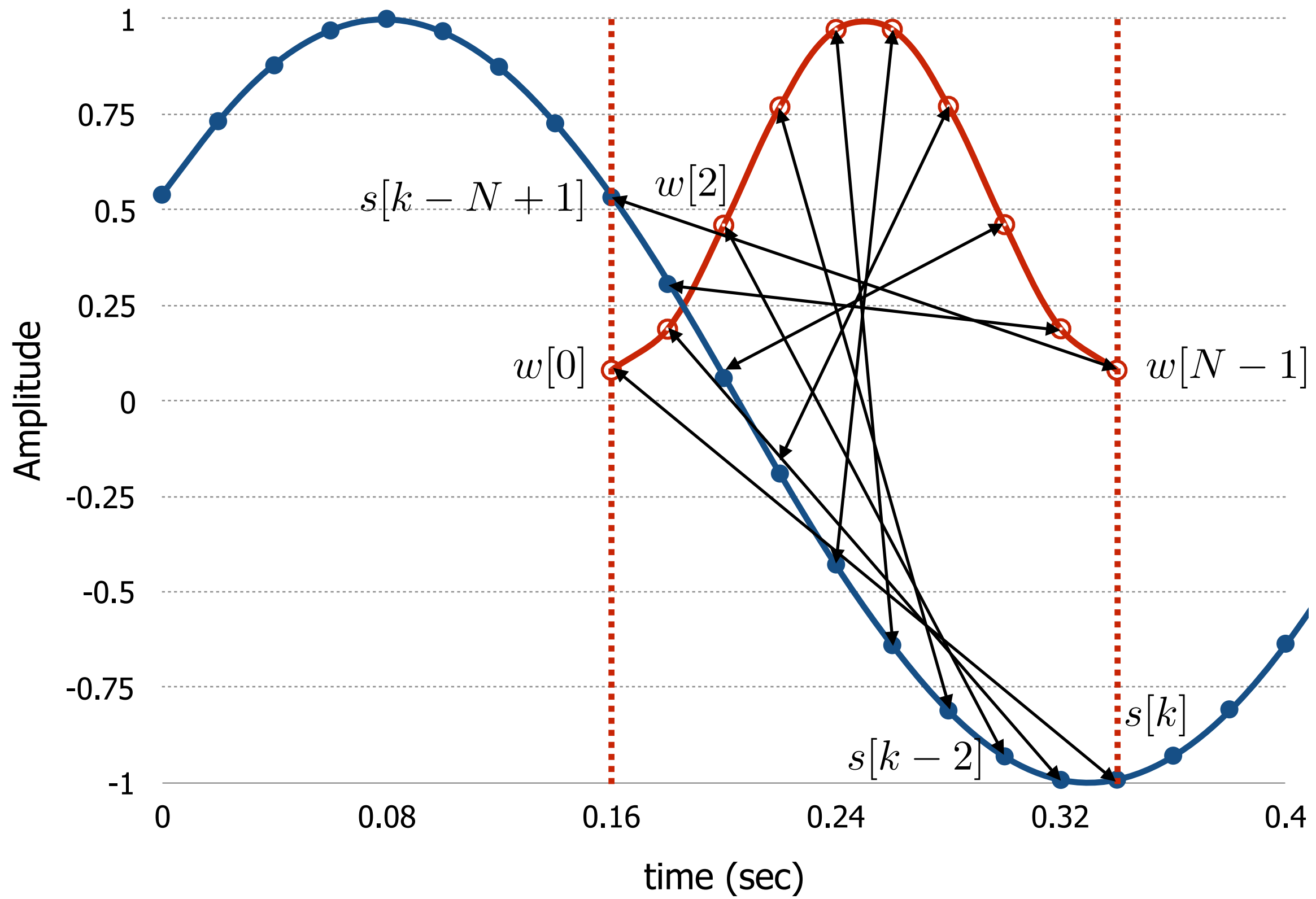
The Hamming window



The diagram illustrates the definition of the Hamming window $w[n]$. A red circle containing $w[n]$ is on the left. An arrow points from the text "The Hamming window" to this circle. Another arrow points from the text "Only samples for which 'n' is between 0 and 'N-1' are not null" to a red oval containing the inequality $0 \leq n \leq N - 1$. The equation for $w[n]$ is presented as a piecewise function with three cases, each preceded by a colon.

$$w[n] = \begin{cases} 0.54 - 0.46 \cos \left(\frac{2\pi n}{N-1} \right) & : 0 \leq n \leq N - 1 \\ 0 & : n < 0 \\ 0 & : n > N - 1 \end{cases}$$



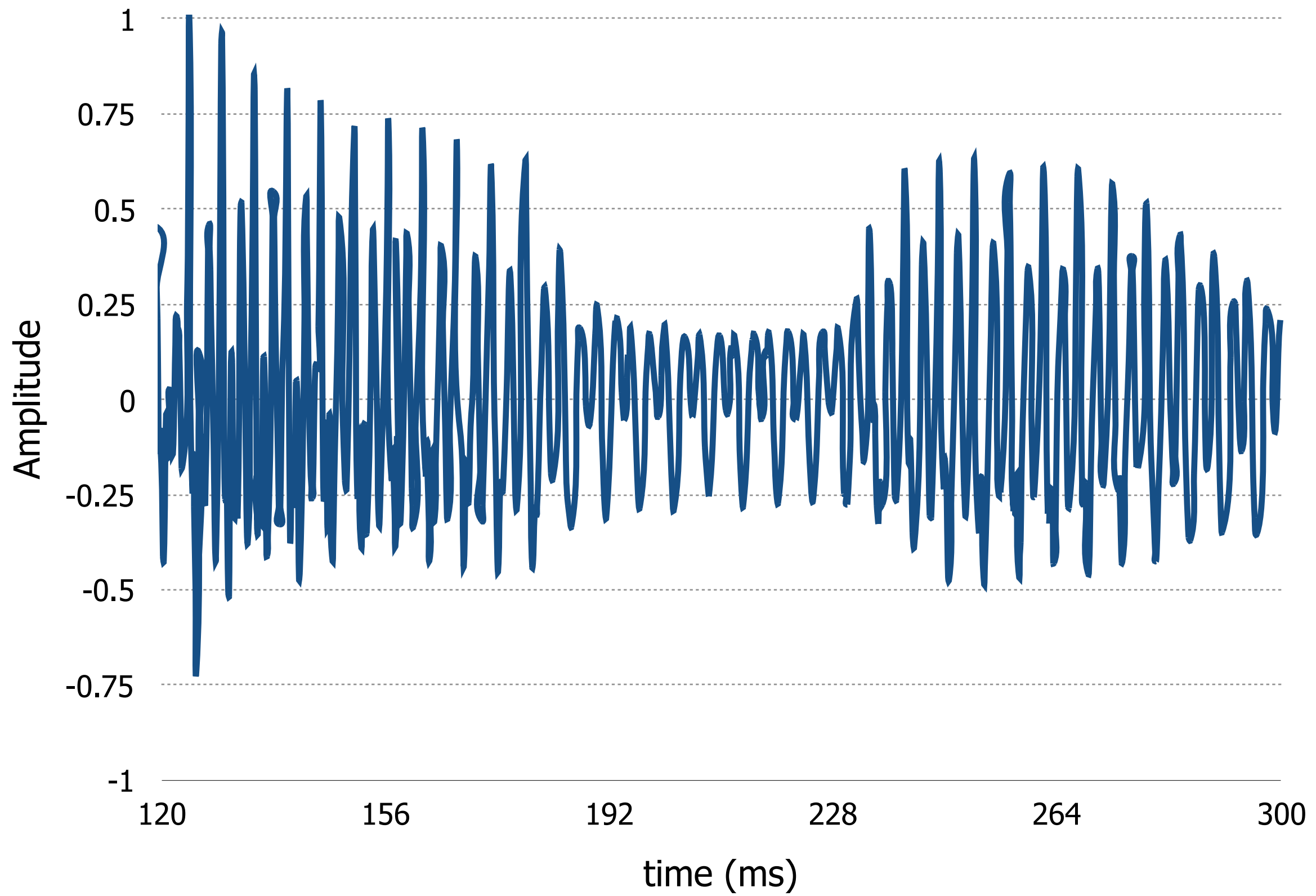


Short-Term Properties

- Any property should be measured over an interval that is short enough to ensure that the signal characteristics are stable;
- However, the interval should still be long enough to capture meaningful atomic units;
- In the case of speech, the typical interval length is 20-30 ms, the time one speaker keeps the articulators in a stable configuration.

Outline

- Introduction
- Time Domain Processing
- **Short-Term Analysis**
- Conclusions



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

A function of sample "n"

The diagram shows the equation $x[k] = \sum_{n=-\infty}^{\infty} f(s[n])w[k-n]$. The term $x[k]$ is enclosed in a red circle, with an arrow pointing to it from the text "Any short-term property that can be extracted from the signal". The term $f(s[n])$ is enclosed in a red circle, with an arrow pointing to it from the text "A function of sample 'n'". The term $w[k-n]$ is enclosed in a red circle, with an arrow pointing to it from the text "A window signal". The entire equation is annotated with the text "Convolution between f(s[n]) and a window" at the bottom left.

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

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

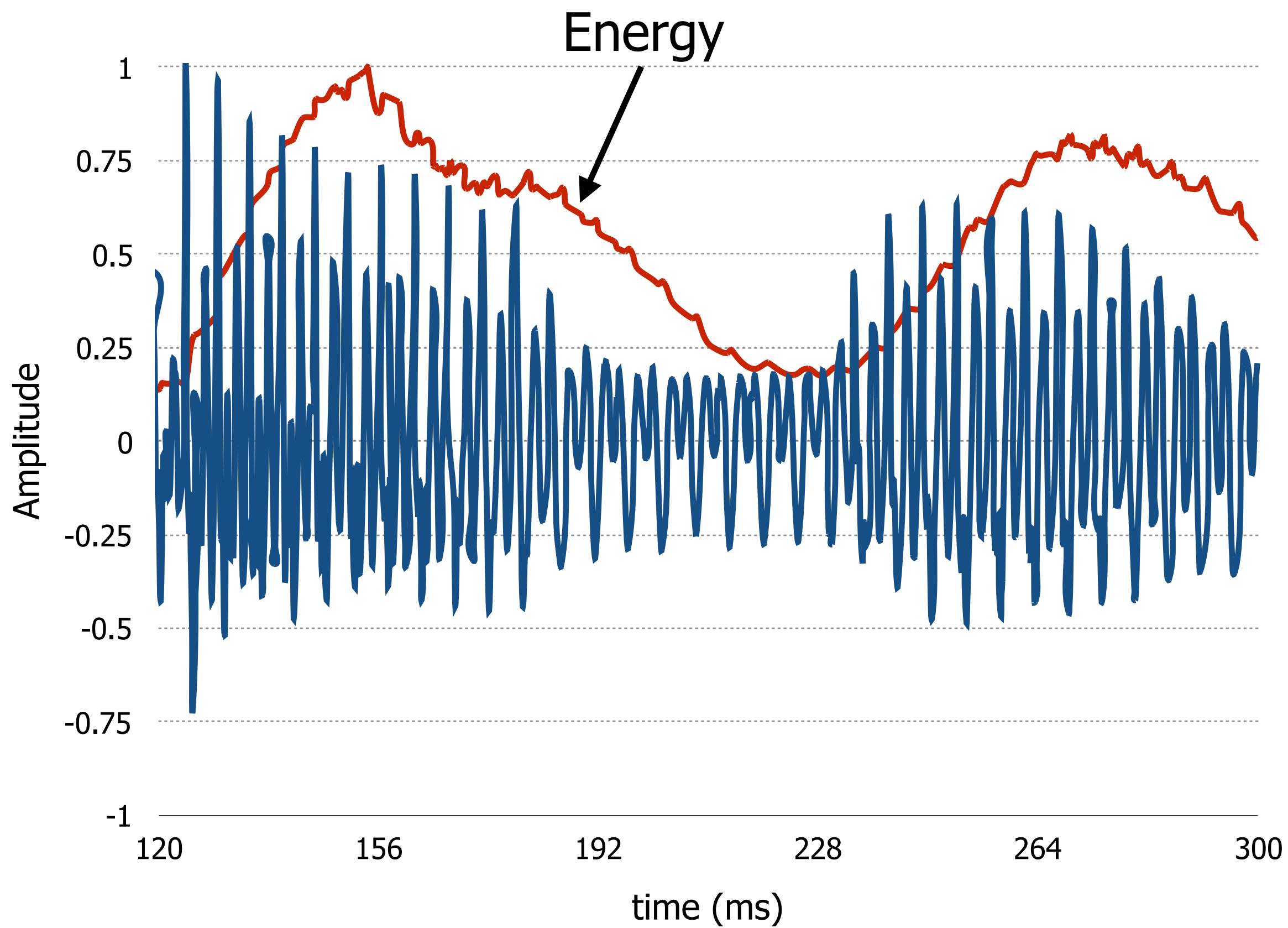
A window signal

The function is the square of the sample divided by the length of the window

The Energy

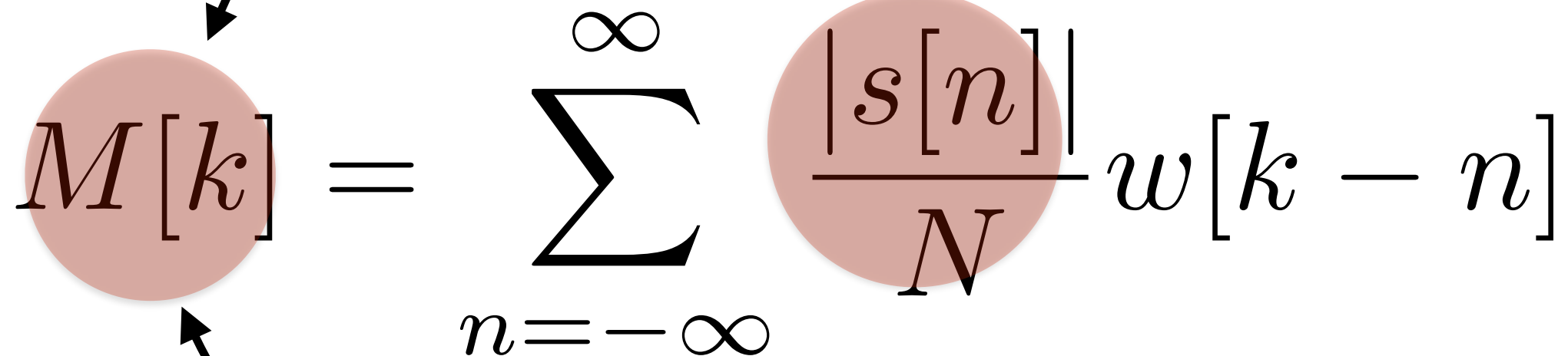
$$E[k] = \sum_{n=-\infty}^{\infty} \frac{(s[n])^2}{N} w[k - n]$$

The energy is the average of the samples' squares in the window ending at sample "k"



The function is the
absolute value of the
sample divided by the
length of the window

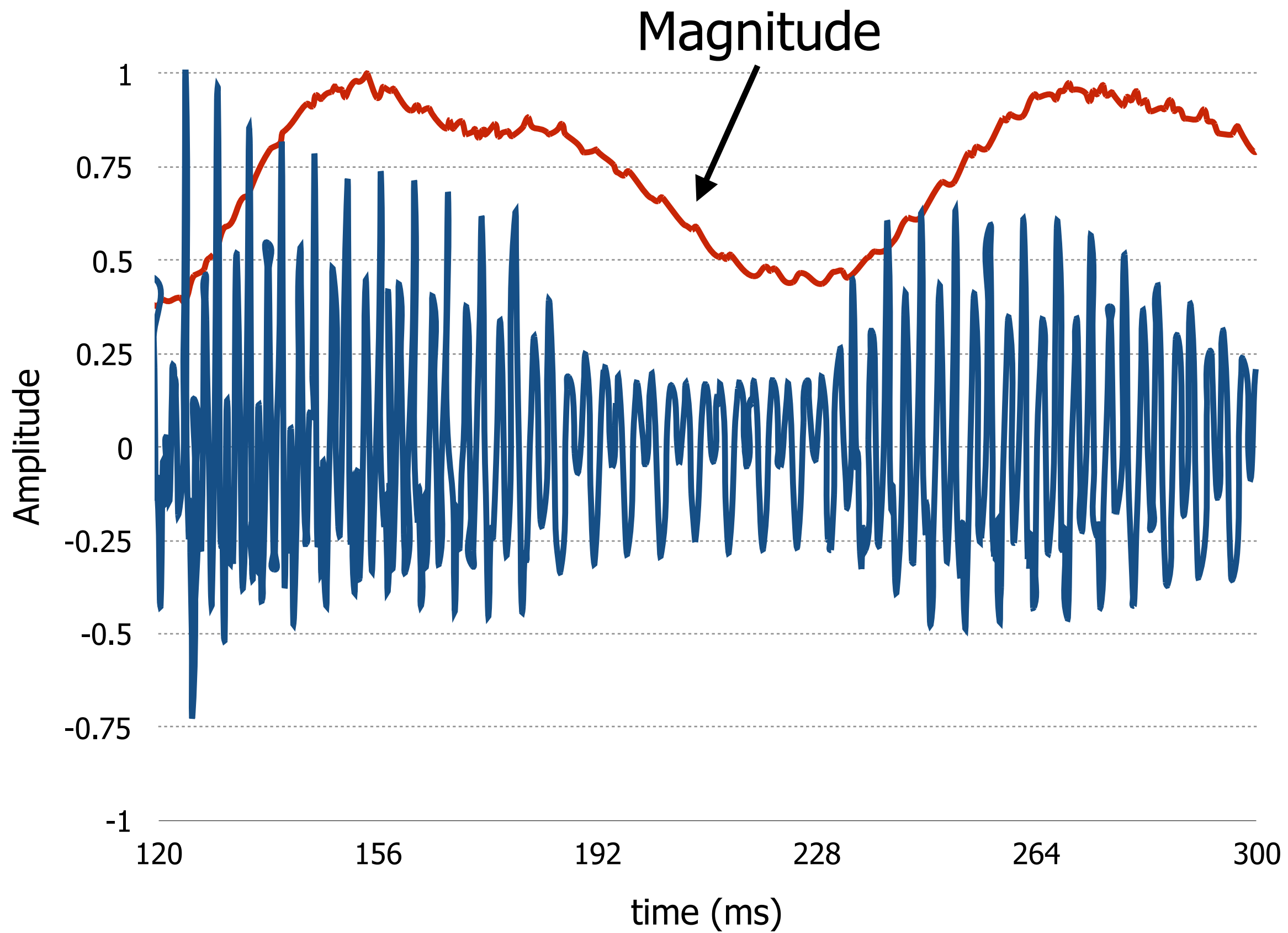
The Magnitude



The diagram shows the formula for calculating the magnitude $M[k]$ at sample k . The formula is
$$M[k] = \sum_{n=-\infty}^{\infty} \frac{|s[n]|}{N} w[k-n]$$
 Annotations include: an arrow from "The Magnitude" pointing to the $M[k]$ term; an arrow from "The function is the absolute value of the sample divided by the length of the window" pointing to the $\frac{|s[n]|}{N}$ term; and a curved arrow from "The magnitude is the average of the samples' absolute value in the window ending in sample 'k'" pointing to the entire summation.

$$M[k] = \sum_{n=-\infty}^{\infty} \frac{|s[n]|}{N} w[k-n]$$

The magnitude is the
average of the samples'
absolute value in the
window ending in
sample "k"



Outline

- Introduction
- Time Domain Processing
- Short-Term Analysis
- **Conclusions**

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.