UNIVERSITÀ DEGLI STUDI DI BERGAMO
Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione

Lesson 15.

Fault diagnosis III

Signal-based approaches

DATA SCIENCE AND AUTOMATION COURSE

MASTER DEGREE SMART TECHNOLOGY ENGINEERING

TEACHER
Mirko Mazzoleni

PLACE
University of Bergamo

Outline

1. Schematic of the approach
2. Spectral analysis review
3. Bearing diagnosis: problem statement
4. Application to inner race bearing diagnosis of workcenter machines
5. Bearing diagnosis: signal-based approach

Outline

1. **Schematic of the approach**
2. Spectral analysis review
3. Bearing diagnosis: problem statement
4. Application to inner race bearing diagnosis of workcenter machines
5. Bearing diagnosis: signal-based approach

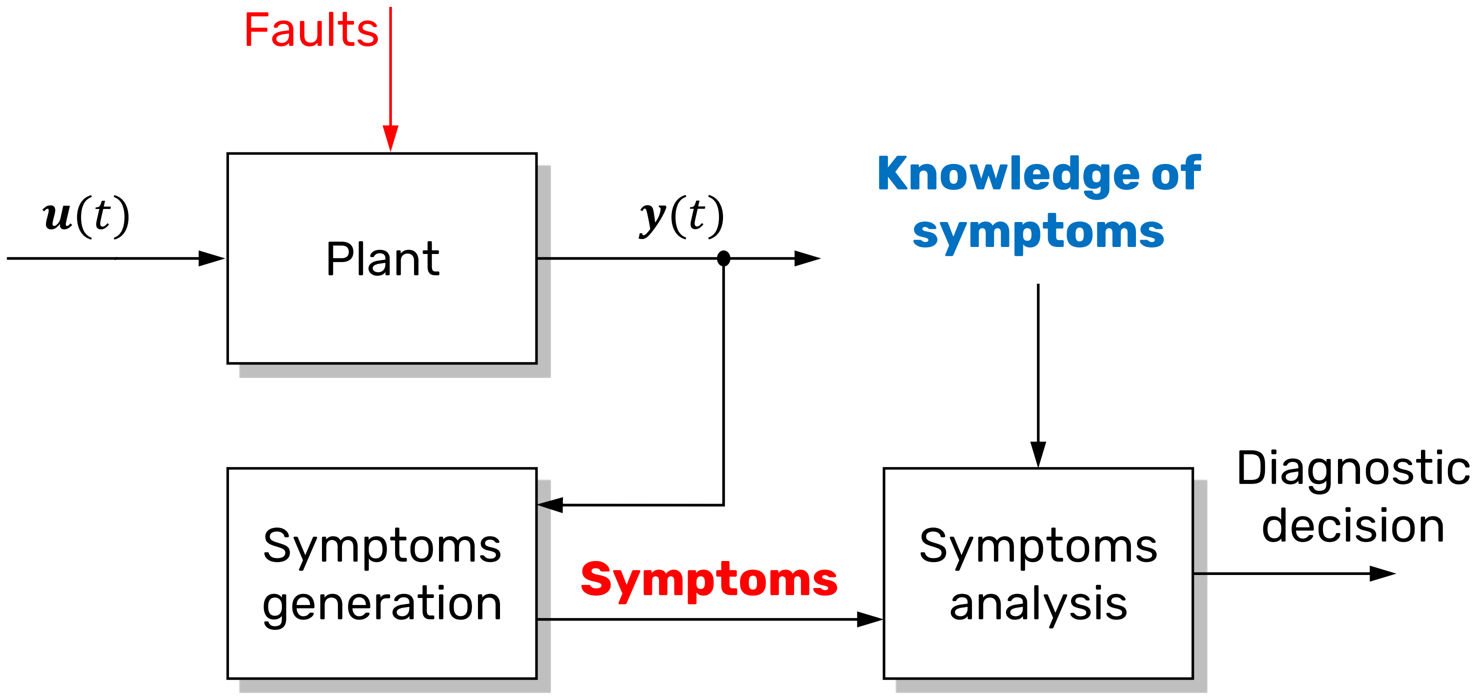
Signal-based fault diagnosis

Certain process signals carry information about the faults to be detected

- From them, **fault symptoms** are computed…
- …and compared with **prior knowledge about the faults**

Examples are:

- Diagnosis of rotating components defects
- Stator\rotor current faults

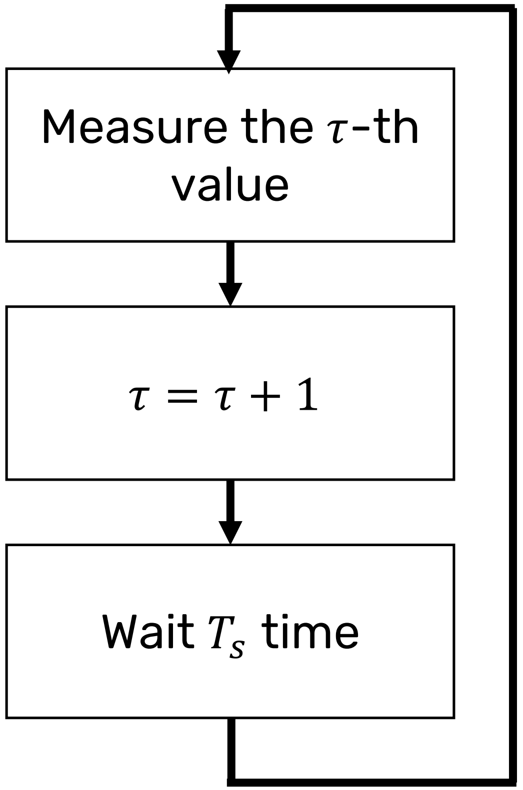


Outline

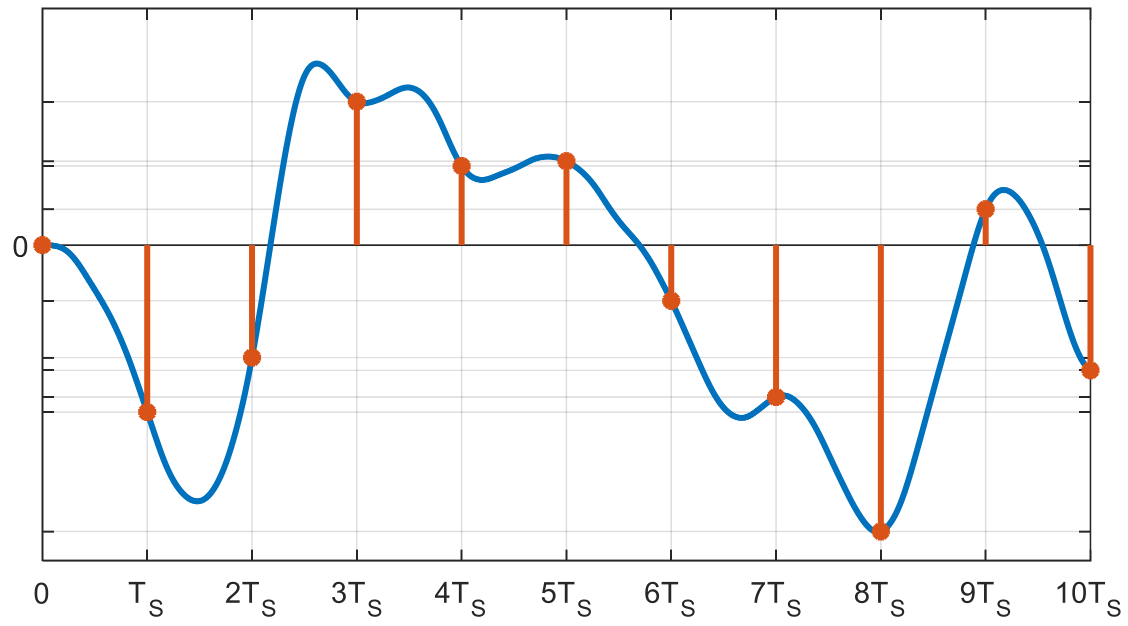
1. Schematic of the approach
2. **Spectral analysis review**
3. Bearing diagnosis: problem statement
4. Application to inner race bearing diagnosis of workcenter machines
5. Bearing diagnosis: signal-based approach

Sampling

Sampling entails acquiring samples of a continuous-time signal $s(t)$. The resulting samples are denoted as $\check{s}(\tau)$, where τ is multiple of the sampling time T_s



$$\check{s}(\tau) = s(\tau \cdot T_s)$$



Sampling

The value T_s is called **sampling time**

The **sampling frequency** f_s is the number of samples taken for unit of time

$$f_s = \frac{1}{T_s}\,[\mathrm{Hz}]$$

The sampled signal misses the information of what happened between two samples

$f_s \uparrow \Leftrightarrow$ More preserved information $\Leftrightarrow$ Higher number of data

There is a way to sample a signal **without losing** any information? It's always possible?

Sampling

A sampled signal is a sum of weighted Dirac deltas:

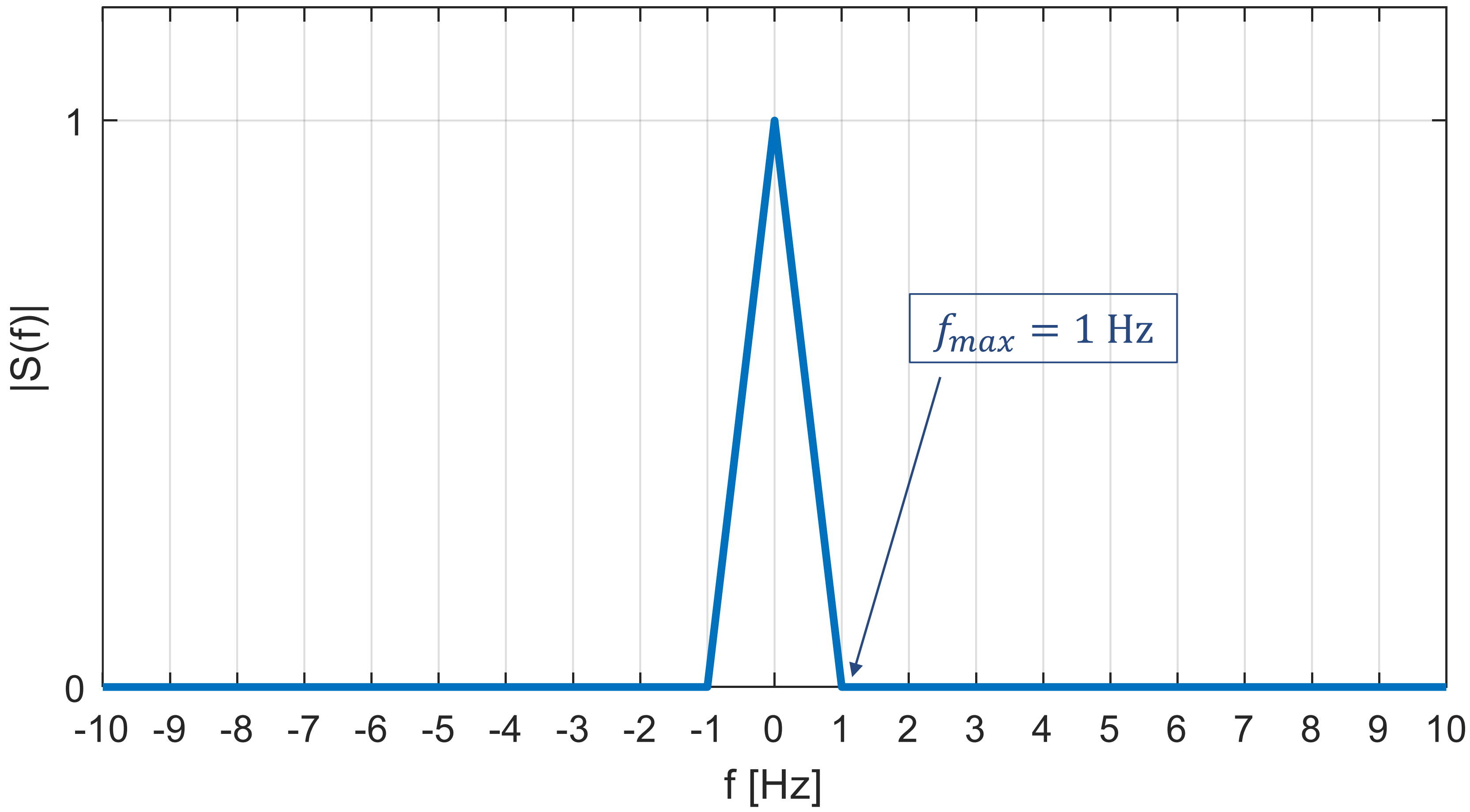
$$s_s(t) = \sum_{\tau=-\infty}^{+\infty} s(\tau \cdot T_S) \cdot \delta(t - \tau \cdot T_S)$$

Theorem

The **Fourier transform** $\mathcal{F}[\check{s}]$ of a **discrete-time** signal $s_s(t)$ is:

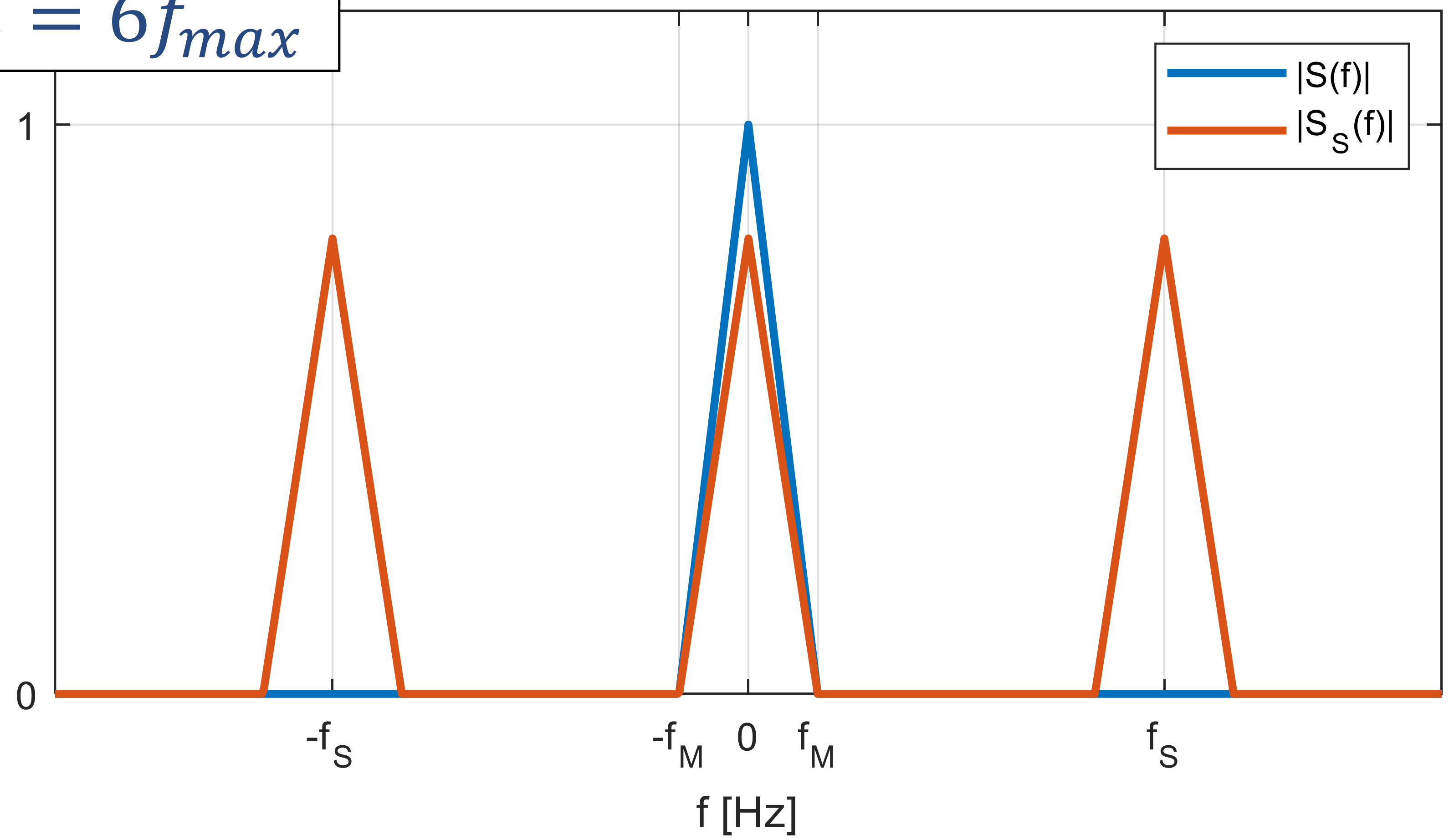
$$S(f) = \mathcal{F}[s_s](f) = f_S \cdot \sum_{\tau=-\infty}^{\tau=+\infty} \mathcal{F}[s](f - \tau \cdot f_S)$$

Example: band limited signal



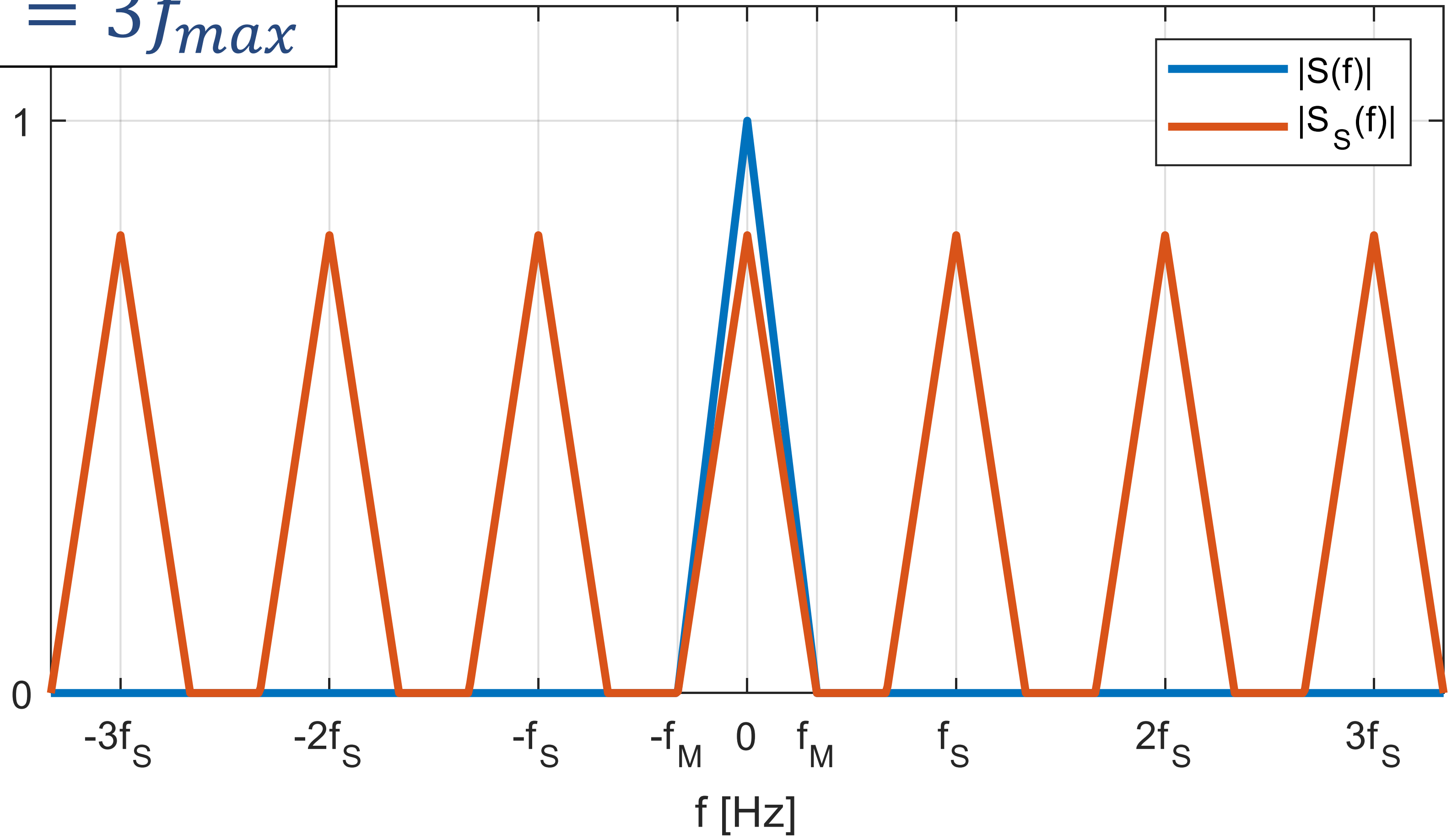
Example: band limited signal

$$f_s = 6f_{max}$$



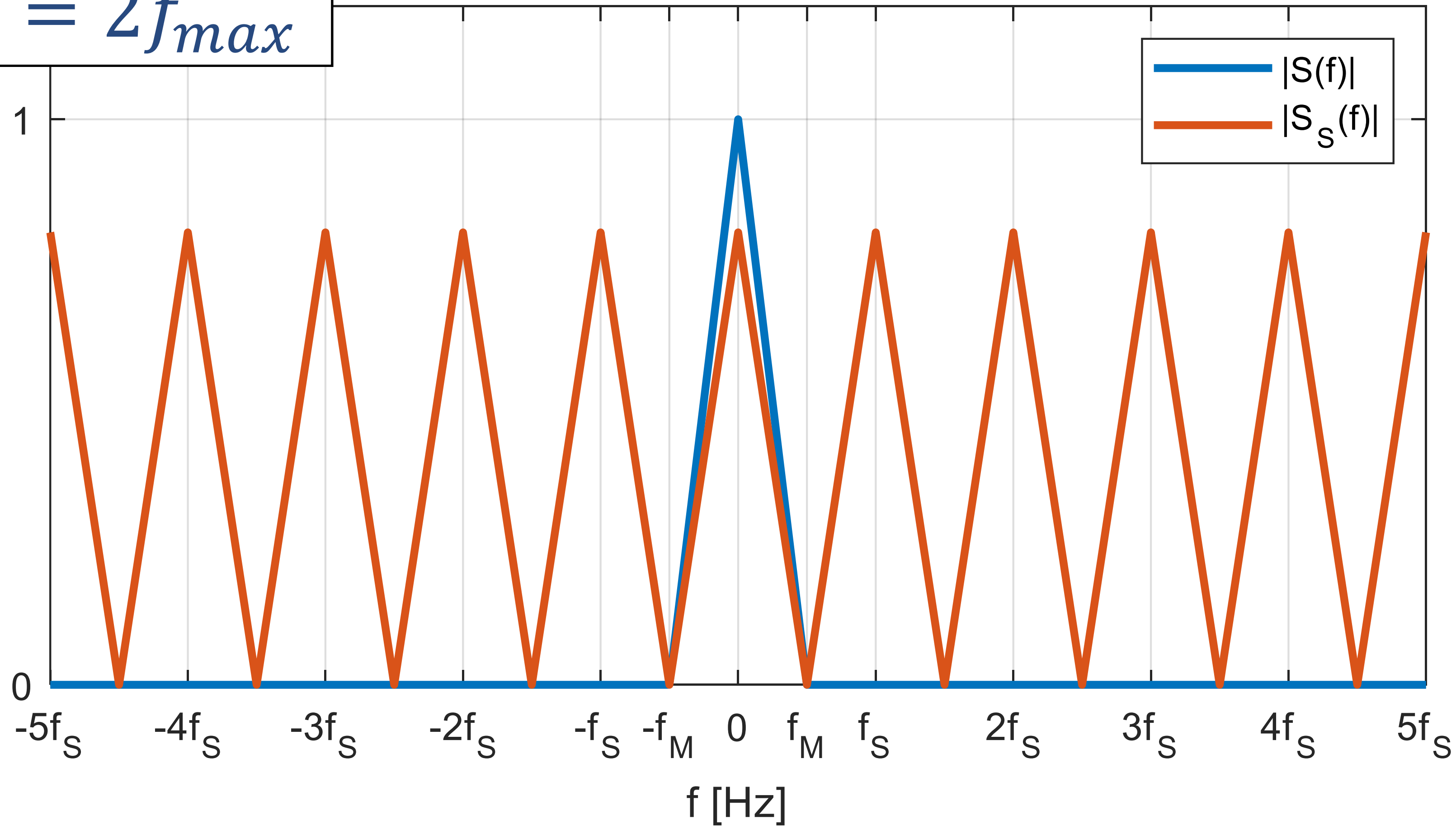
Example: band limited signal

$f_s = 3f_{max}$



Example: band limited signal

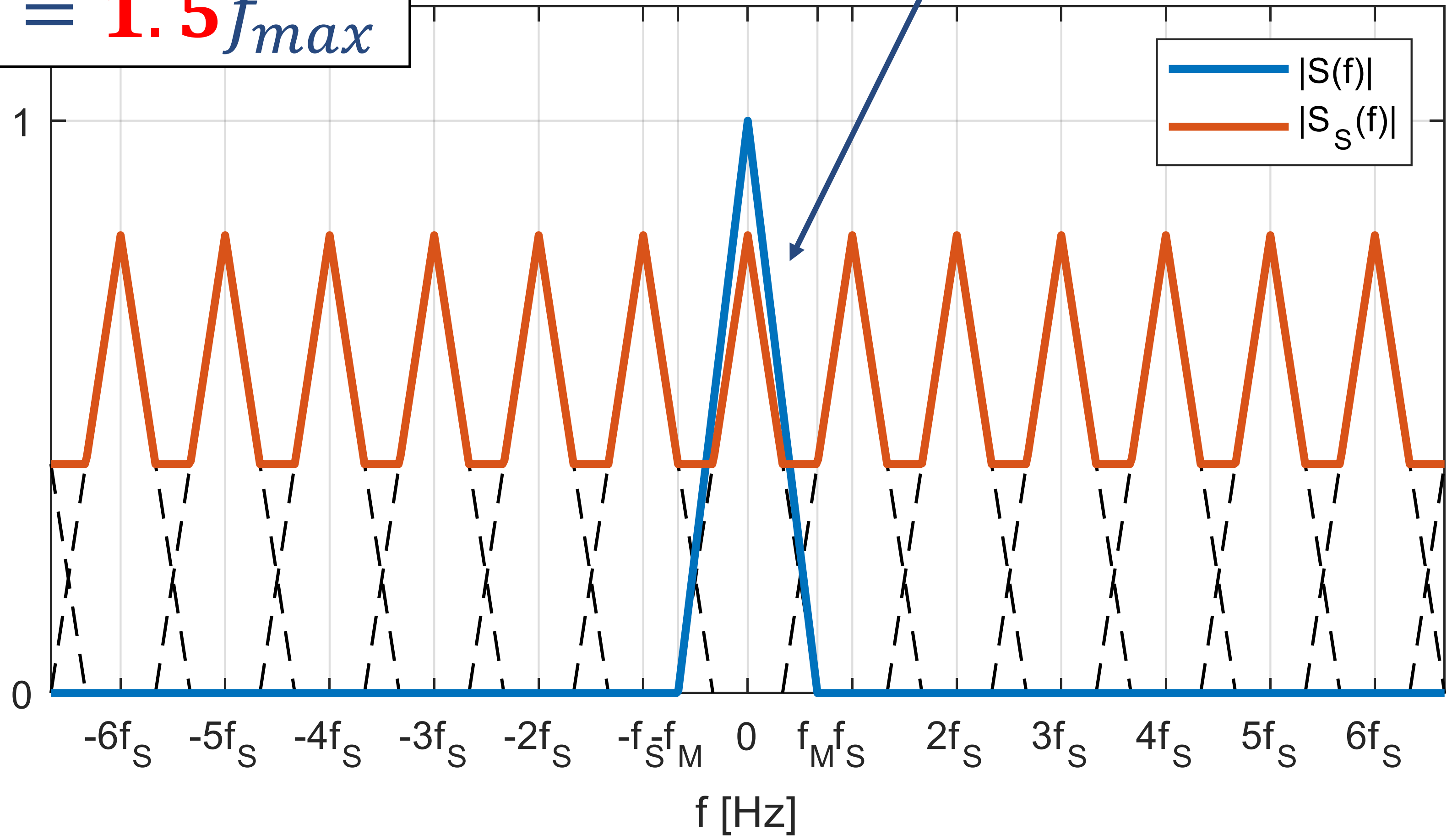
$$f_s = 2f_{max}$$



Example: band limited signal

The transform of the sampled signal is different from the continuous-time signal!

$f_s = \mathbf{1.5} f_{max}$



Sampling

The spectra of a sampled band-limited signal has **infinite** (translated) **repetitions** of the original spectra **(aliases)**

- If $f_s > 2f_{max}$ such repetitions don't overlap
- If $f_s < 2f_{max}$ such repetitions overlap with each other **(aliasing problem)**

If the repetitions don't overlap around 0 there is an exact copy of the original spectra (minus an amplification). An (analog) **anti-alias filter** is placed **before** the sampling

If the aliases don't overlap, it's possible to reconstruct the original signal perfectly from its samples, using a low-pass filter

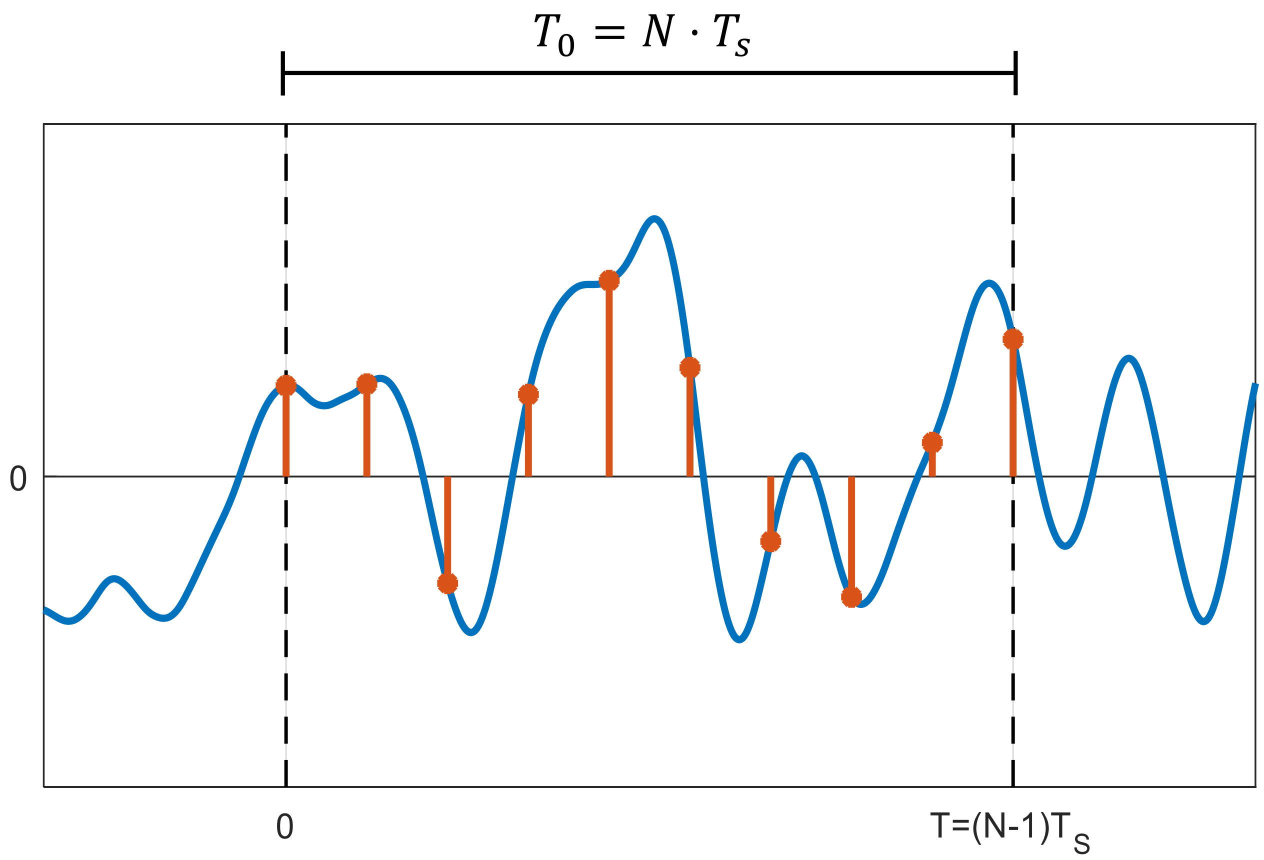
Windowing

The Fourier transform assumes that we have an **infinite amount of samples** ($\tau \to +\infty$).

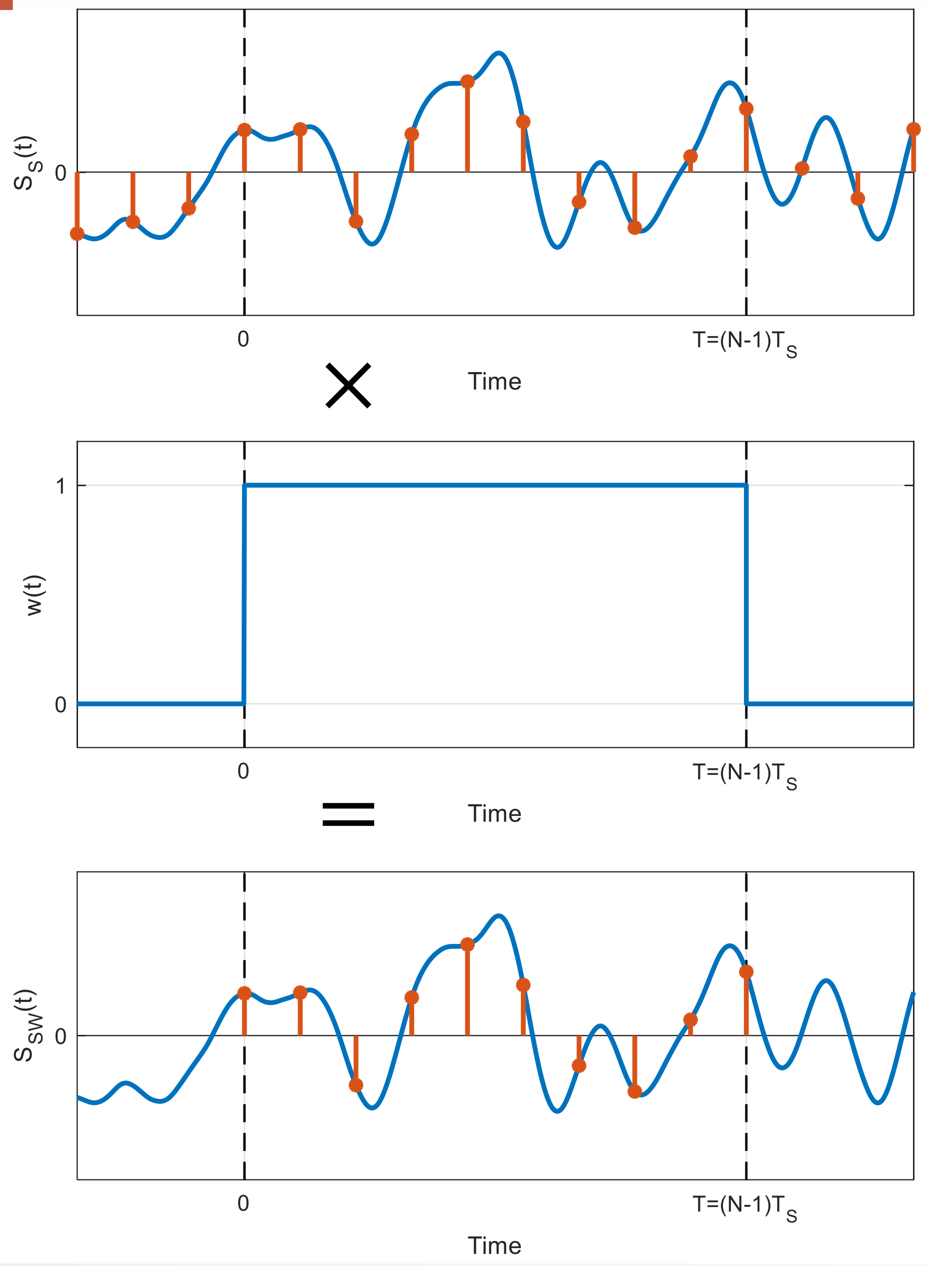
In practice, this is not possible

- Consider the case when we start to sample at time 0 and we stop at time T, by taking N samples

- We **don't have any information** on the behavior of the signal outside this time window



Windowing



$$s_s(t) = \sum_{\tau=-\infty}^{+\infty} s(\tau \cdot T_s) \cdot \delta(t - \tau \cdot T_s)$$

$$\times$$

$$w(t) = \begin{cases} 1 & 0 \leq t \leq (N-1) \cdot T_s \\ 0 & \text{otherwise} \end{cases}$$

Window function

$$=$$

$$s_{s,w}(t) = \sum_{\tau=0}^{N-1} s(\tau \cdot T_s) \cdot \delta(t - \tau \cdot T_s)$$

Fourier analysis for sampled signals

Suppose to have N samples $\breve{s}(\tau)$ of a signal $s(t)$ with a sampling frequency f_S. What can we say about the **frequency components** of $s(t)$ given the samples $\breve{s}(t)$?

We have to make some assumptions:

1. The **sampling** frequency respect the Nyquist criteria ($f_s \geq 2f_{max}$)
2. The signal $s(t)$ is **periodic** with period $T_0 = N \cdot T_S$, where $N \in \mathbb{N}_{>0}$

Since we assume that the signal is periodic, we consider the Fourier coefficients:

$$c_k = \frac{2}{T_0} \cdot \int_0^{T_0} s(t) \cdot e^{-j \cdot \frac{2\pi k t}{T_0}} dt$$

Fourier analysis for sampled signals

Since we assume that the **signal is periodic**, we consider the **Fourier coefficients:**

$$c_k = \frac{2}{T_0} \cdot \int_0^{T_0} s_{sw}(t) \cdot e^{-j \cdot \frac{2\pi k t}{T_0}} dt$$

Since we have:

$$s_{sw}(t) = \sum_{\tau=0}^{\tau=N-1} s(\tau T_S) \cdot \delta(t - \tau T_S)$$

We obtain:

$$c_k = \boxed{\frac{2}{N T_S}} \cdot \sum_{\tau=0}^{\tau=N-1} \boxed{s(\tau T_S)} e^{-j \cdot \frac{2\pi k \tau}{N}}$$

Normalization constant that is usually ignored

τ-th sample

Fourier analysis for sampled signals

Given N samples $\breve{s}(\tau)$ of a signal $s(t)$ with a sampling frequency f_s, the **discrete Fourier transform (DFT)** is given by:

$$\boxed{\mathcal{F}_D[\breve{s}](k) = \sum_{\tau=0}^{N-1} \breve{s}(\tau) \cdot e^{-j \cdot \frac{2\pi k \tau}{N}}}$$

The DFT of a signal of length N is vector of length N

(only $N/2$ useful since it is symmetric)

$\mathcal{F}_D[\breve{s}](k)$ correspond to the k-th Fourier coefficient of the **windowed** sampled signal, with the assumption that the **signal is periodic** with period $N \cdot T_S$

Fourier analysis for sampled signals

Supposing the signal as periodic, the fundamental frequency is:

$$f_0 = \frac{1}{T_0} = \boxed{\frac{1}{NT_s}}$$

«Frequency bin»

And, therefore, the k-frequency is

$$f_k = \frac{k}{T_0} = \frac{k}{NT_s} = \frac{k}{N} f_s$$

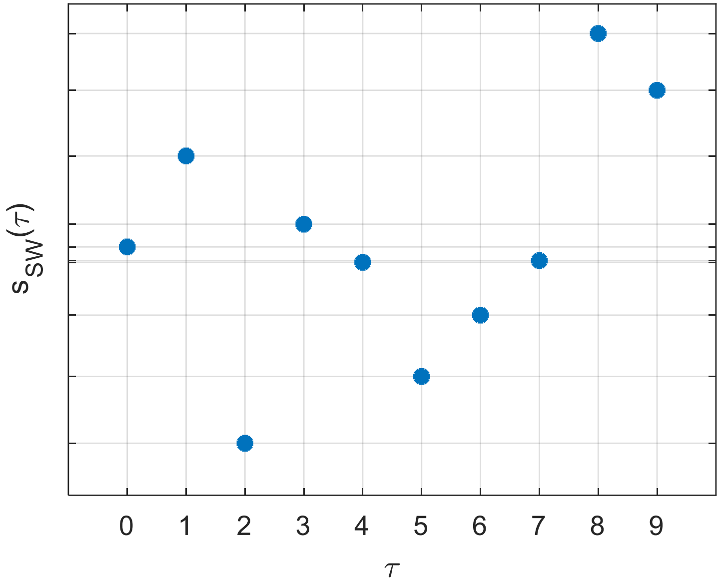
It is possible to note that

$$f_{\frac{N}{2}} = \frac{N/2}{N} f_s = \frac{f_s}{2}$$

The maximum frequency is $f_s/2$

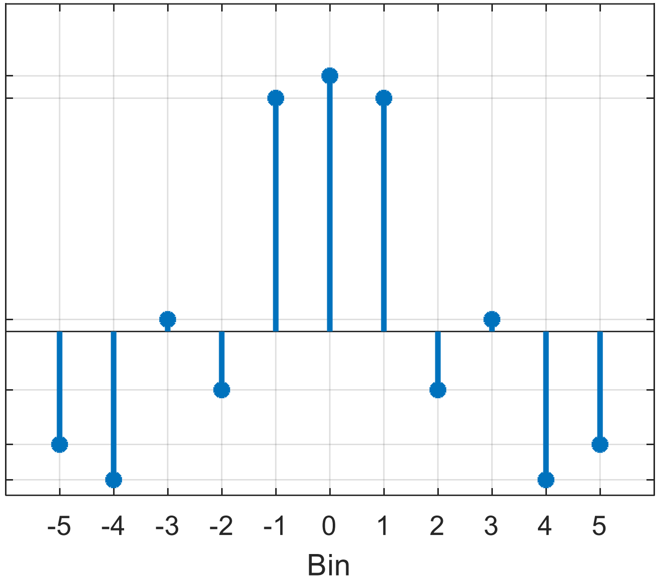
DFT summary

$N = 10$



DFT

iDFT



The **Inverse Discrete Fourier Transform (iDFT)** is used to get the sampled signal from the spectrum

$\breve{s}(0)$

$\breve{s}(1)$

$\vdots$

$\breve{s}(9)$

DFT

iDFT

$\mathcal{F}_D[\breve{s}]\left(-{}^{N}/_{2}\right)$

$\vdots$

$\mathcal{F}_D[\breve{s}](0)$

$\vdots$

$\mathcal{F}_D[\breve{s}]\left({}^{N}/_{2} - 1\right)$

FFT and leakeage

The **Fast Fourier Transform (FFT)** is an efficient algorithm that computes the DFT of a given signal

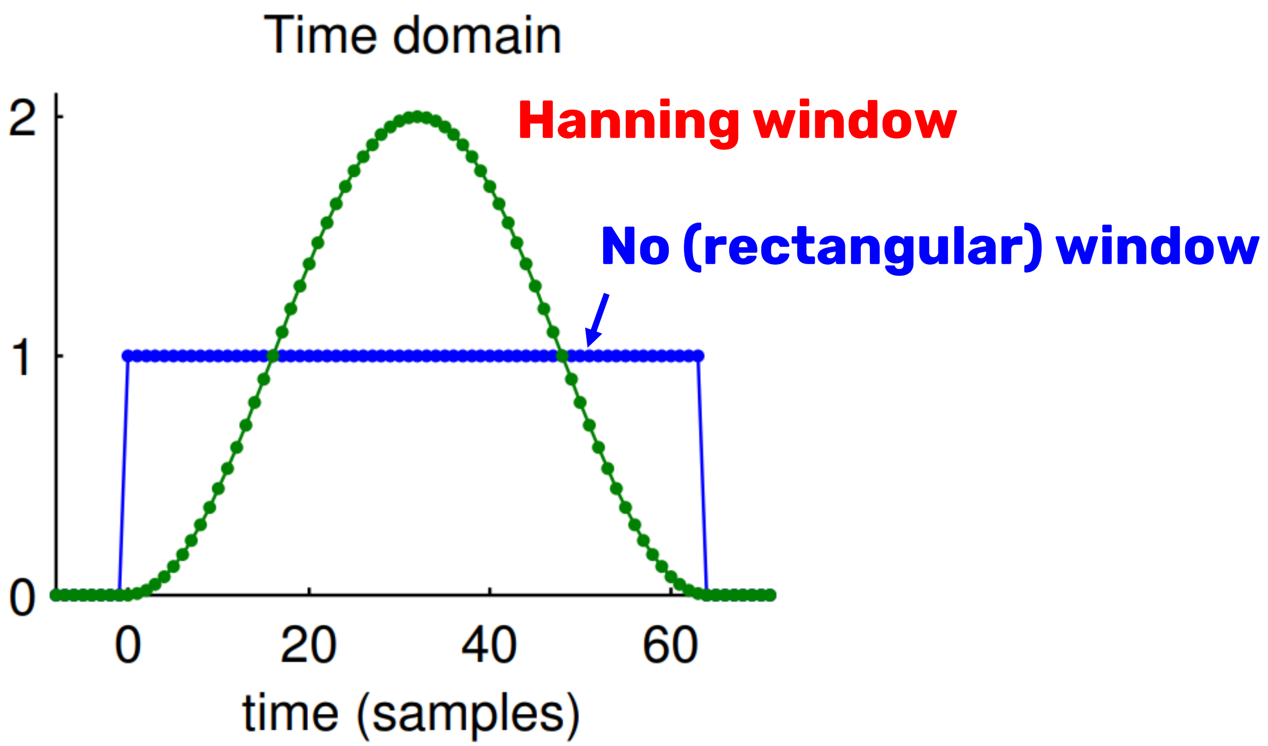
It always returns the complex N length vector:

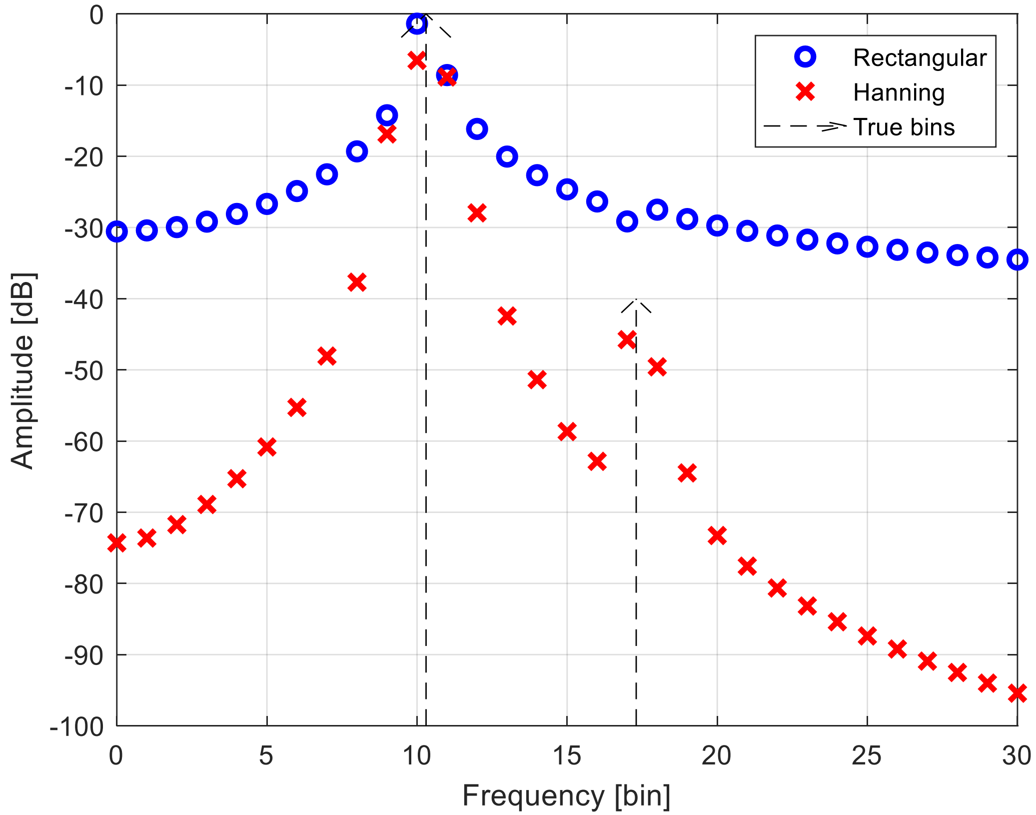
$$[\mathcal{F}_D[\check{s}](0) \quad \cdots \quad \mathcal{F}_D[\check{s}](N-1)]$$

If you apply the FFT to a **non periodic** signal, some **spurious frequencies** appear in the spectrum **(leakage)**

FFT and leakeage

To attenuate this problem, you can **window the signal** with special windowing functions (e.g. **Hanning window**)





Outline

1. Schematic of the approach

2. Spectral analysis review

3. **Bearing diagnosis: problem statement**

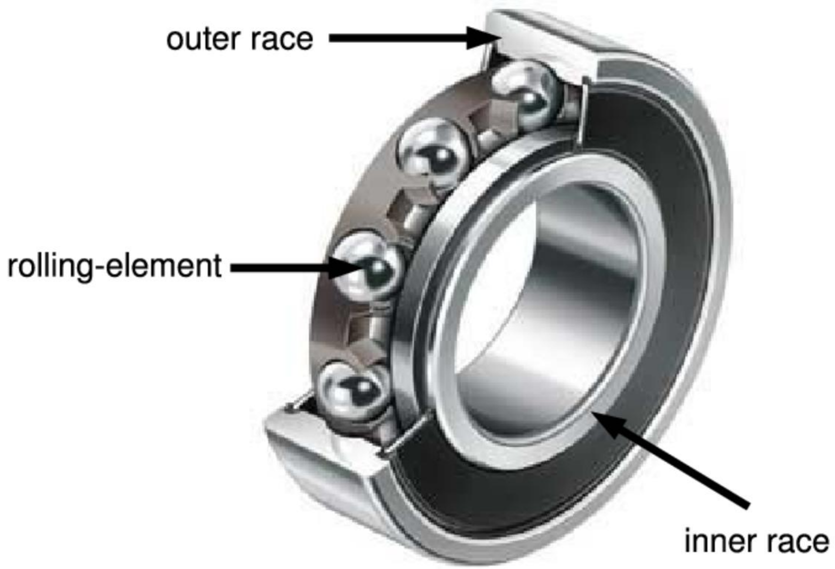
4. Application to inner race bearing diagnosis of workcenter machines

5. Bearing diagnosis: signal-based approach

Rolling bearings

Rolling bearings are mechanical components whose function is to **interpose** between machine parts in mutual **rotation** and to **limit their friction**

They are one of the **most widely used** elements in machines and their **failure** one of the **most frequent** reasons for machine breakdown



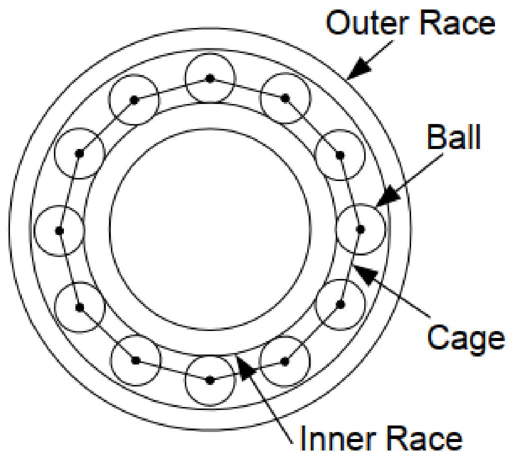
Rolling elements can have **different geometries**: spheres, needles, cylindrical, tapered or barrel rollers

The choice of **the type of rolling element** depends on the **application**, the load to which it is subjected and its direction

Rolling bearings

They are composed by several components:

- **Races:** the surfaces on which the bearing rolls. The load placed on the bearing is supported by this contact surface. In general, the inner ring rests on the shaft, while the outer ring rests on the bearing housing



- **Rolling elements:** the rolling elements are constructed in such a way as to allow also their rotation (simultaneous rotation around their axis and around the axis of the bearing)

- **Cage:** separates the rolling elements at a regular interval, holds them in position between the internal and external races and allows their free rotation

Rolling bearings defects

- **Galling**: type of wear that occurs due to the friction that occurs when moving two materials, that are compressed on each other.
- **Spalling:** process by which metal is broken into small fragments (spalls).
- **Peeling:** formation of coarse irregularities on the surface of the coating applied to a metal surface.
- **Pitting:** corrosion that presents itself as localized attacks, by creating small holes on the surface of the metal.
- **Scoring:** surface damage caused by debris accumulated in the bearing under improper lubrication conditions or excessive loads.

Rolling bearings defects

- **Smearing:** superficial damage that occurs due to the presence of small debris between the bearing components due to the breakage of the lubricant film or due to slippage of the elements.
- **Fracture and cracks:** fracture of the elements can be caused by excessive or impulsive load acting locally on the component considered.
- **Denting:** it occurs when the debris, made up of small metal particles, are in the contact area between the rotating element and the track.
- **Fretting:** wear occurs due to repeated slipping between two surfaces. Fretting occurs both on the mounting surfaces and on the contact surfaces between the tracks and the rotating elements.
- **Creep:** is a phenomenon in which slipping between two mounting surfaces creates a play.

Rolling bearings defects

- **Seizure:** when the bearing overheats quickly during rotation, the bearing changes color. After overheating the tracks, the rotating elements and the cage slowly begin to melt and deform, accumulating more and more damage.
- **True brinelling:** occurs when the load on the bearing is greater than the elastic limit of the bearing material.
- **False brinelling:** it looks similar to true brinelling but is due to vibrations. For example, during transport, vibrations can cause the rolling elements to move and therefore leave indentations on the tracks.
- **Flaking:** occurs when small particles of material detach from the surface of the rolling track or from the rolling body due to fatigue, forming rough and irregular areas

Rolling bearings defects

The causes of rolling bearing defects can be of a different nature: **wear** (e.g. due to lack of lubricant\maintenance), **fatigue**, excessive **loads**, presence of **debris**,…

We can classify bearing defects into two major classes

- **Localized defects:** a crack or an incision
- **Distributed defects:** misalignments, eccentricity of the races or rolling elements

Localized defects are often **indicators of failures in progress**. Their monitoring is crucial

When a rolling element «steps over» a (locally) damaged element, it is like it **gets «hit» by an impulsive input**. If we could measure a **vibration**, we will see «hits» in correspondence to the fault

Rolling bearings defects

Analytical formulas exist to describe the **base frequencies excited** by localized defects

BallPass Frequency Inner race

Fault on the inner race

$$\mathrm{BPFI} = \frac{n f_r}{2}\left\{1 + \frac{d}{D}\cos\phi\right\}$$

BallPass Frequency Outer race

Fault on the outer race

$$\mathrm{BPFO} = \frac{n f_r}{2}\left\{1 - \frac{d}{D}\cos\phi\right\}$$

Ball (roller) Speed Frequency

Fault on rolling elements

$$\mathrm{BSF(RSF)} = \frac{D}{2d} f_r\left\{1 - \left(\frac{d}{D}\cos\phi\right)^2\right\}$$

Foundemental Train Frequency

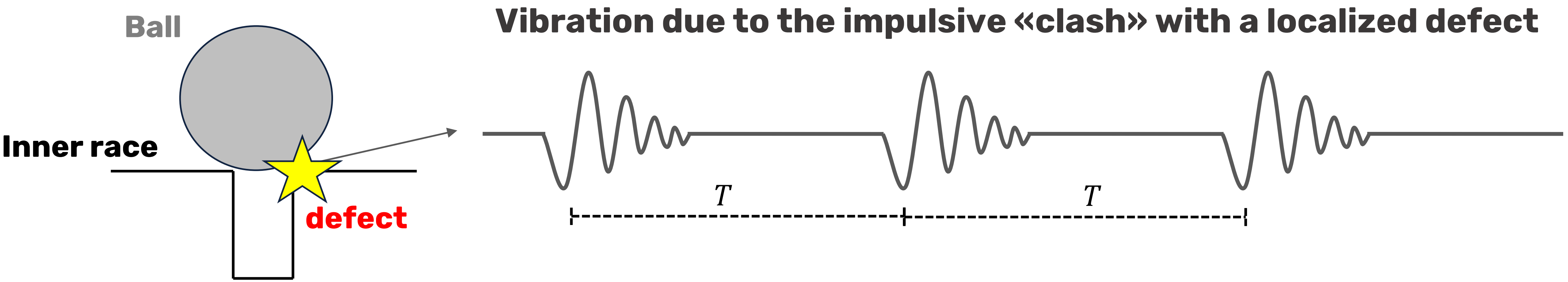
Fault on cage

$$\mathrm{FTF} = \frac{f_r}{2}\left\{1 - \frac{d}{D}\cos\phi\right\}$$

- f_r: shaft speed [
- ϕ: angle of the load from radial plane
- d: single rolling element diameter
- D: average bearing diameter (pitch)
- n: number of rolling elements of the bearing

Rolling bearings diagnostic

Local faults in rolling element bearings produce a **series of impacts** which repeat (almost) **periodically** at a rate dependent on bearing **geometry** and **rotation speed**



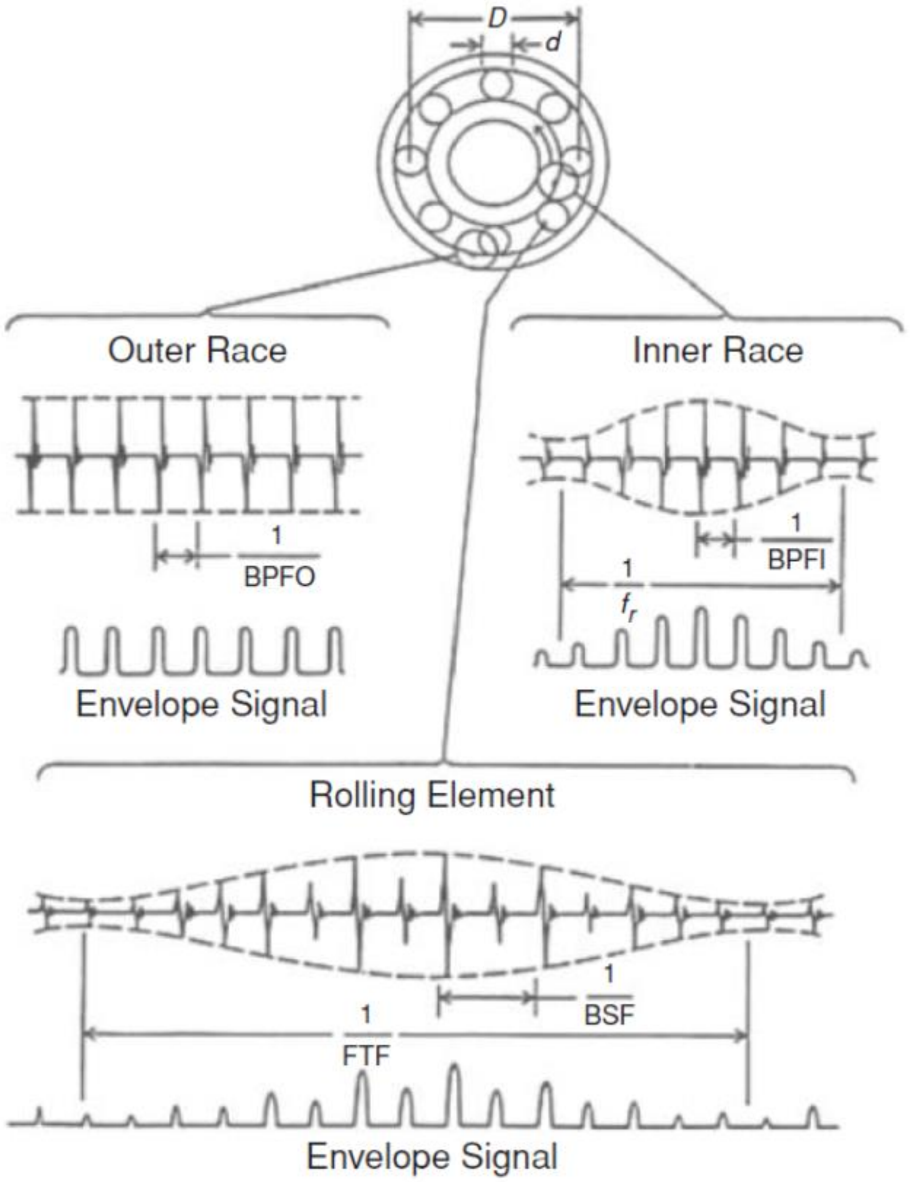
The **diagnostic information** of interest is contained in the **repetition frequency of the impact series**, and not in the frequency spectrum resulting from the impacts, as this would usually be a sum of the resonance frequencies excited

Rolling bearings diagnostic

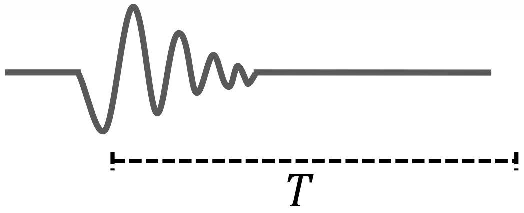
In the **outer race** case, the fault will always be subject to the same load, and the impacts will have equal amplitudes

In the **inner race** case, the fault will rotate in and out of the loaded region, causing modulation of the impact amplitudes by the rotation shaft speed

Envelope analysis allows to extract a signal useful for diagnostic



Rolling bearings diagnostic



Due to the **nature of the fault**, the best way to detect fault-generated **vibrations** is to use an **accelerometer**

However, vibration signals are often **severely corrupted** by strong levels of background noise, encompassing all **other vibration sources** in the machine under inspection.

This problem can be formulated as that of **detecting transient signals** in strong additive noise

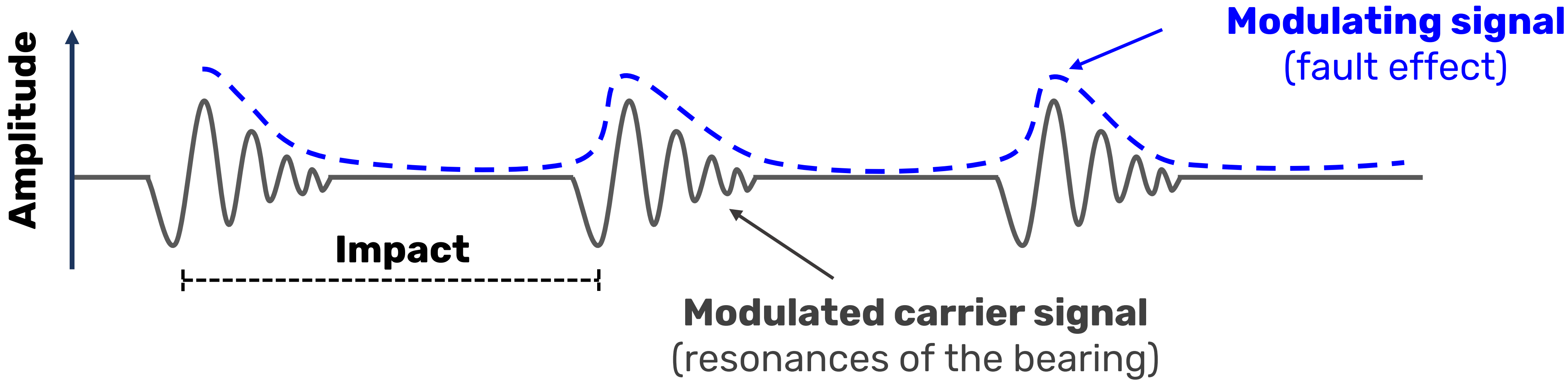
Frequency analysis based on the FFT **on the raw vibration** signal **is not useful** due to:

1. The before mentioned noise
2. Random **fluctuations in the shaft speed**, which compromise the repeatability of the fault impulses responses *(if the speed changes, the fault frequencies change too)*

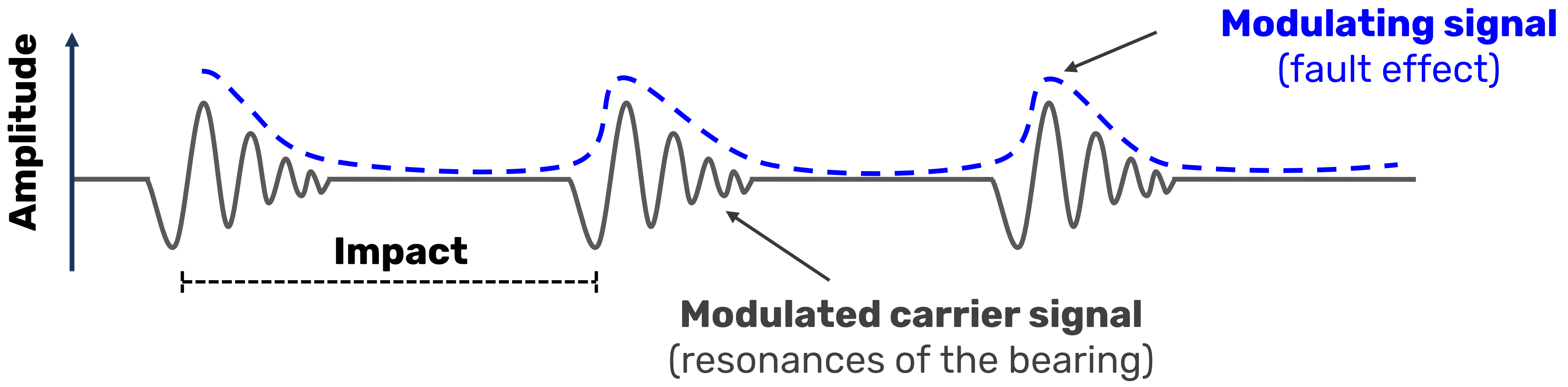
Envelope analysis

One way to enhance the fault frequencies is by using the **envelope analysis**

As we saw, **impactive faults** excite the **structural resonances** of bearing, simply **amplifying** standard operational vibrations. This variation effect on the amplitude of the natural frequency is known as **amplitude modulation**



Envelope analysis



The aim of envelope analysis (also known as amplitude demodulation) is to **reconstruct** the **modulating signal** from the measured **modulated signal.**

Then, a **frequency analysis** of the resulting modulating signal can be performed to **evaluate the presence** of the fault frequency (and its harmonics)

Outline

1. Schematic of the approach
2. Spectral analysis review
3. Bearing diagnosis: problem statement
4. **Application to inner race bearing diagnosis of workcenter machines**
5. Bearing diagnosis: signal-based approach

Bearing diagnosis of workcenters machines

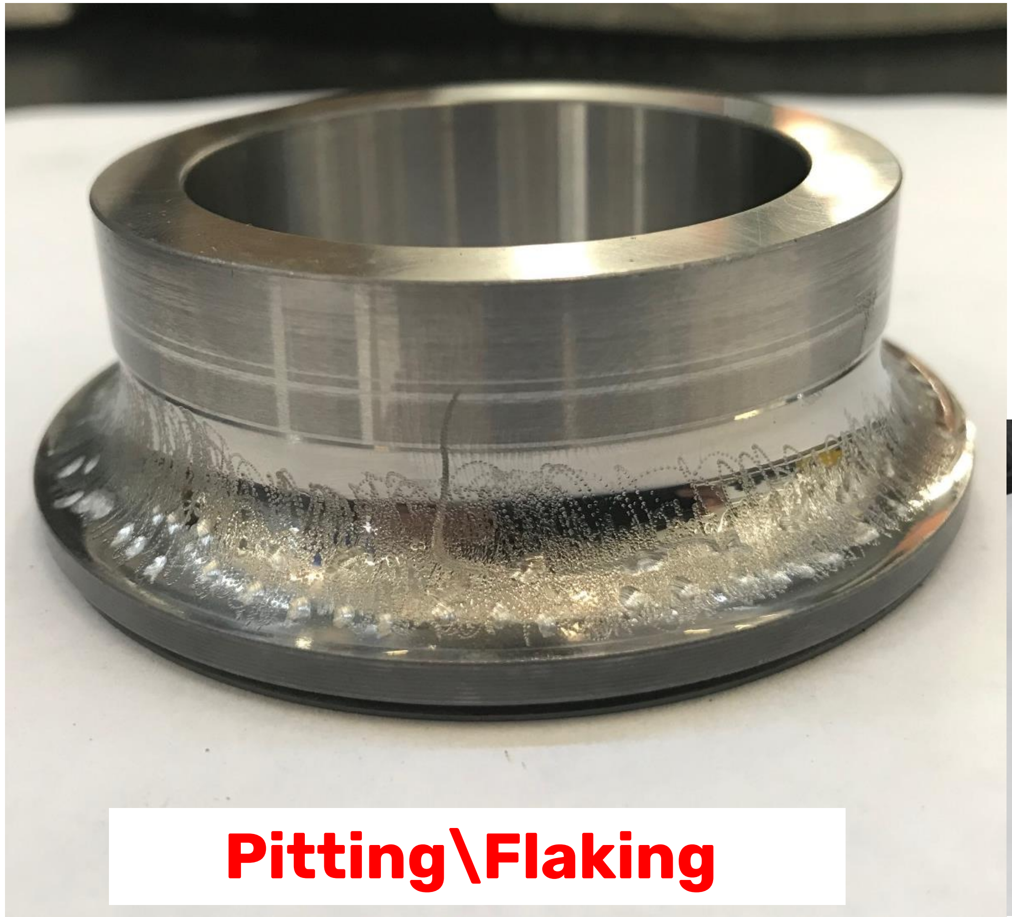
AIM: monitoring the **bearing** of a control center (CNC machine) vertical shaft

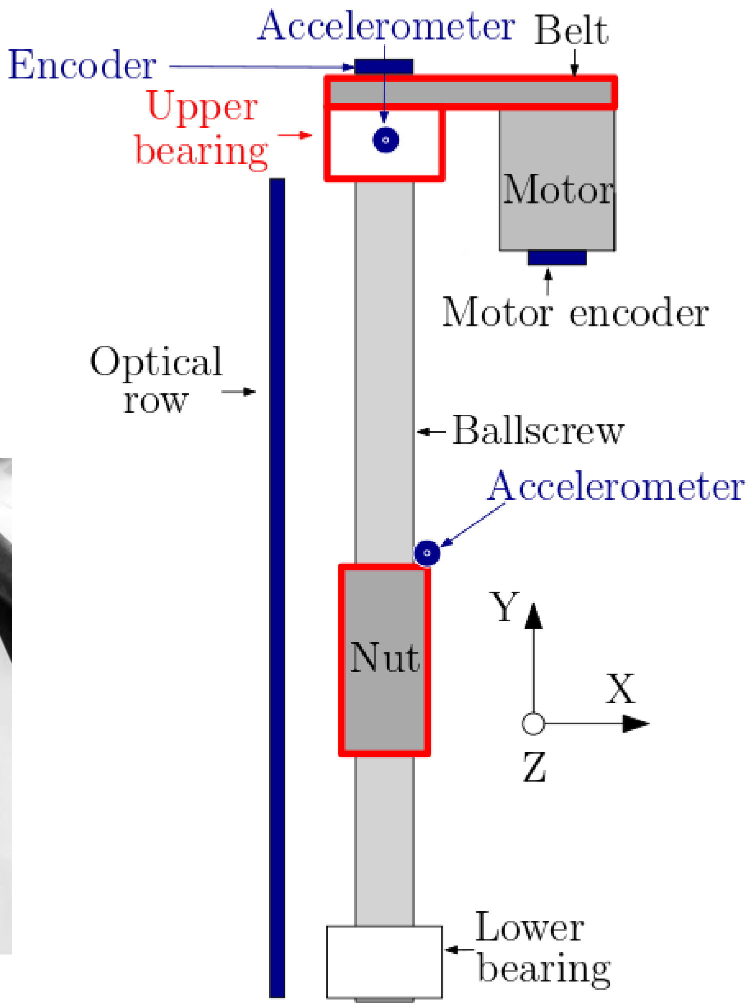
This kind of machine is use for **highly precision** manufacturing operations

These centers can work on the three Euclidean axes, and also on several rotational ones (where the **spindle** is oriented)

mandelli

Bearing diagnosis of workcenters machines





Bearing diagnosis of workcenters machines

Accelerometer on bearing housing

Chassis NI CompactDAQ 9184

NI 9230 module

For the development of the algorithm for the diagnosis of the bearing, the Y axis of the accelerometer is considered.

Sampling frequency: 12800 Hz

Acceleralation range: $\pm 50 \text{ g}$

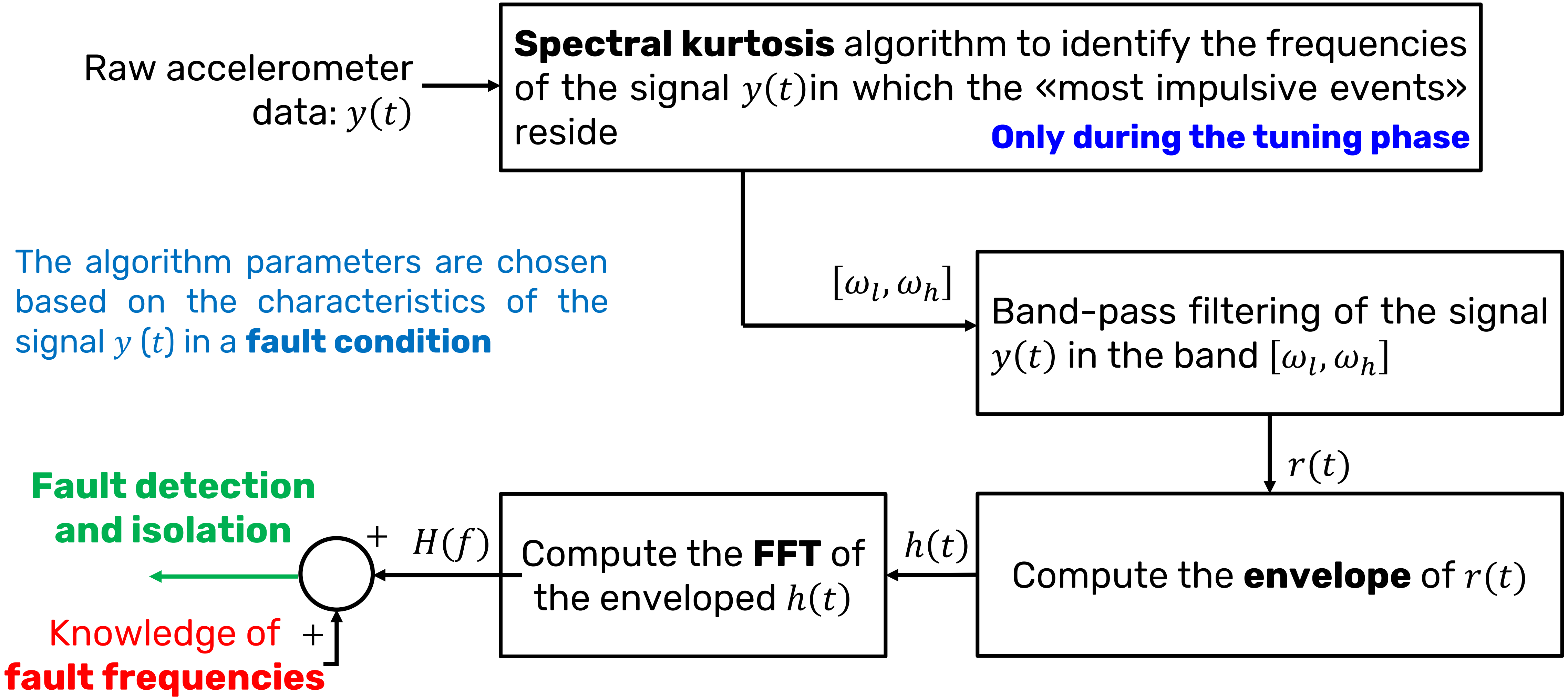
Frequency range: $\pm[0.5 - 3000]\text{Hz}$

Piezoelectric accelerometer

Outline

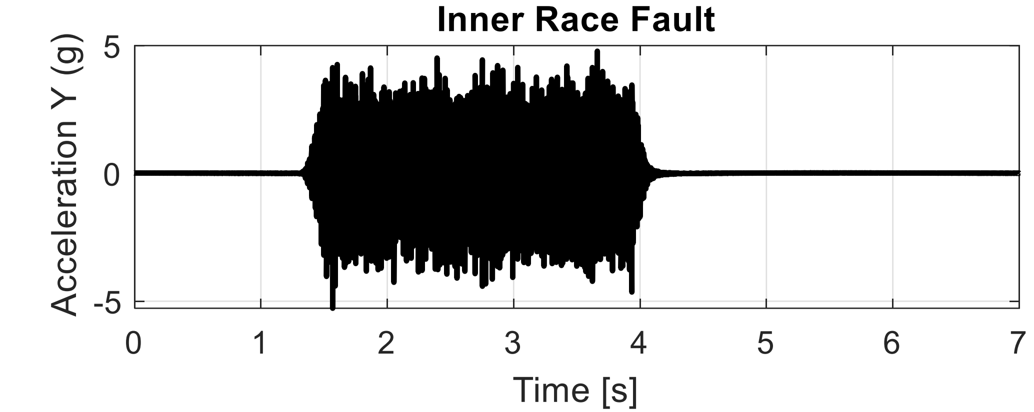
1. Schematic of the approach
2. Spectral analysis review
3. Bearing diagnosis: problem statement
4. Application to inner race bearing diagnosis of workcenter machines
5. **Bearing diagnosis: signal-based approach**

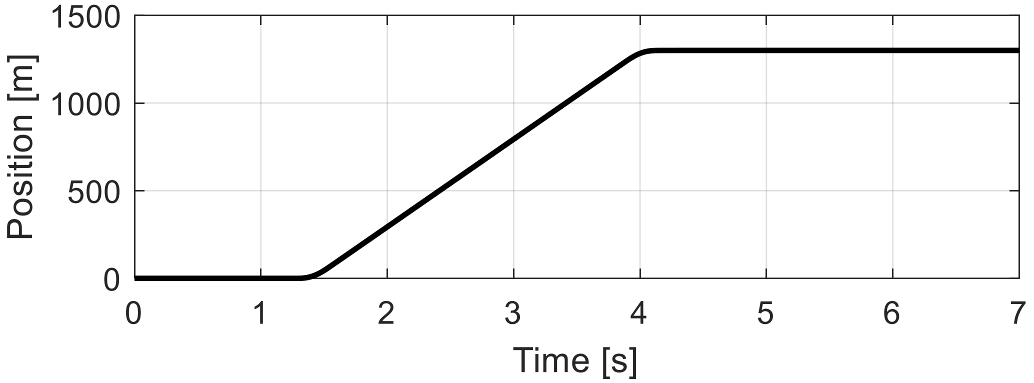
Bearing diagnosis of workcenters machines

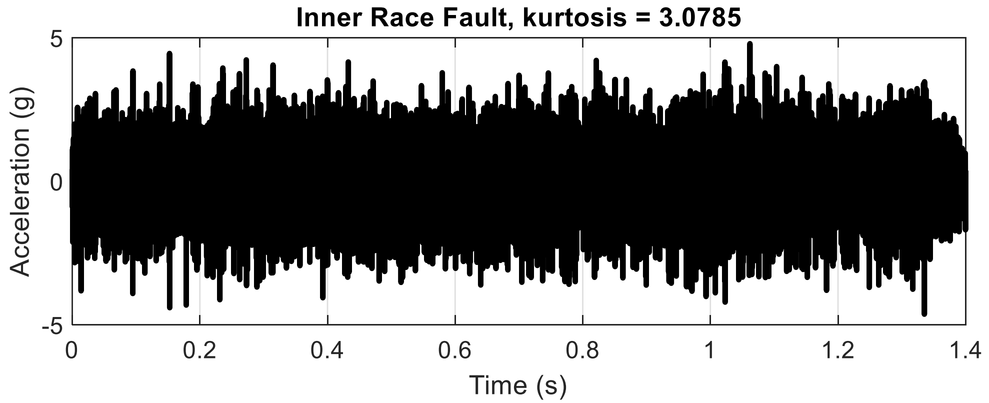


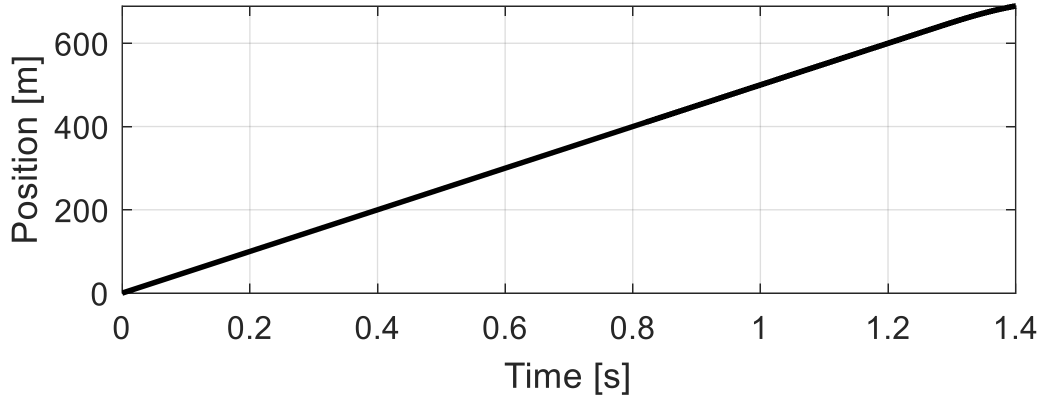
Step 1: constant speed data

It is necessary to maintain only the vibration signal section that corresponds to a **constant rotational speed** of the shaft





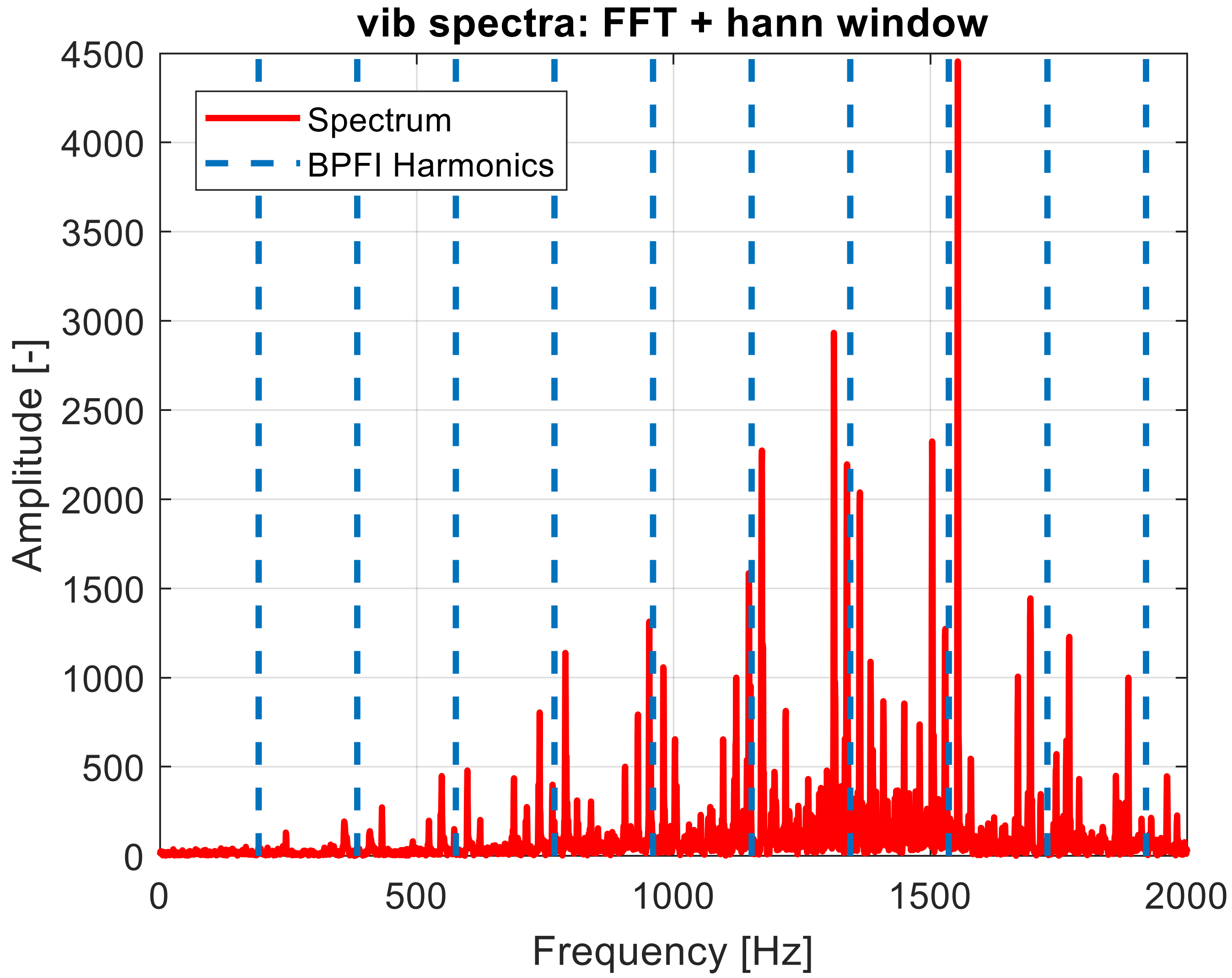




Example: application of the FFT to raw data

The direct application of the FFT to the signal $y(t)$, at constant speed, **does not lead** to immediate recognition of the fault

Several **signal preprocessing** stages are required to bring out the fault symptoms

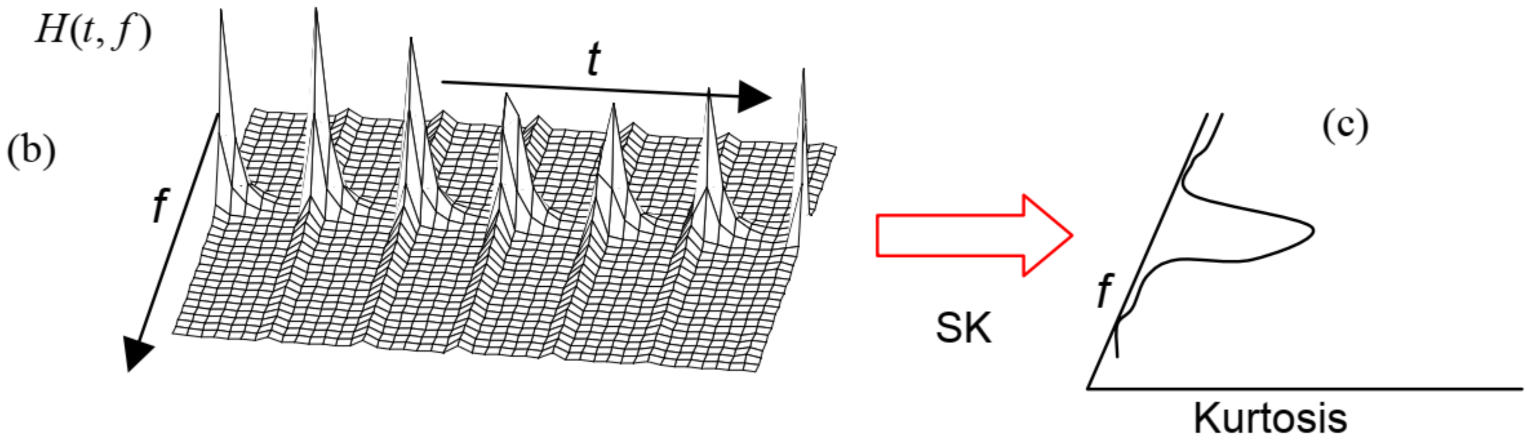


Step 2: spectral kurtosis

The **spectral kurtosis** algorithm is used to select the frequency bands that exhibit «more impulsivity» or «energy»

The goal is to **isolate the impulsive fault components**, by filtering the signal in the frequency bands found by the SK algorithm

Spectral kurtosis (SK) is obtained by **computing the kurtosis for each frequency** (or frequency window) in a **time-frequency** diagram (spectrogram)



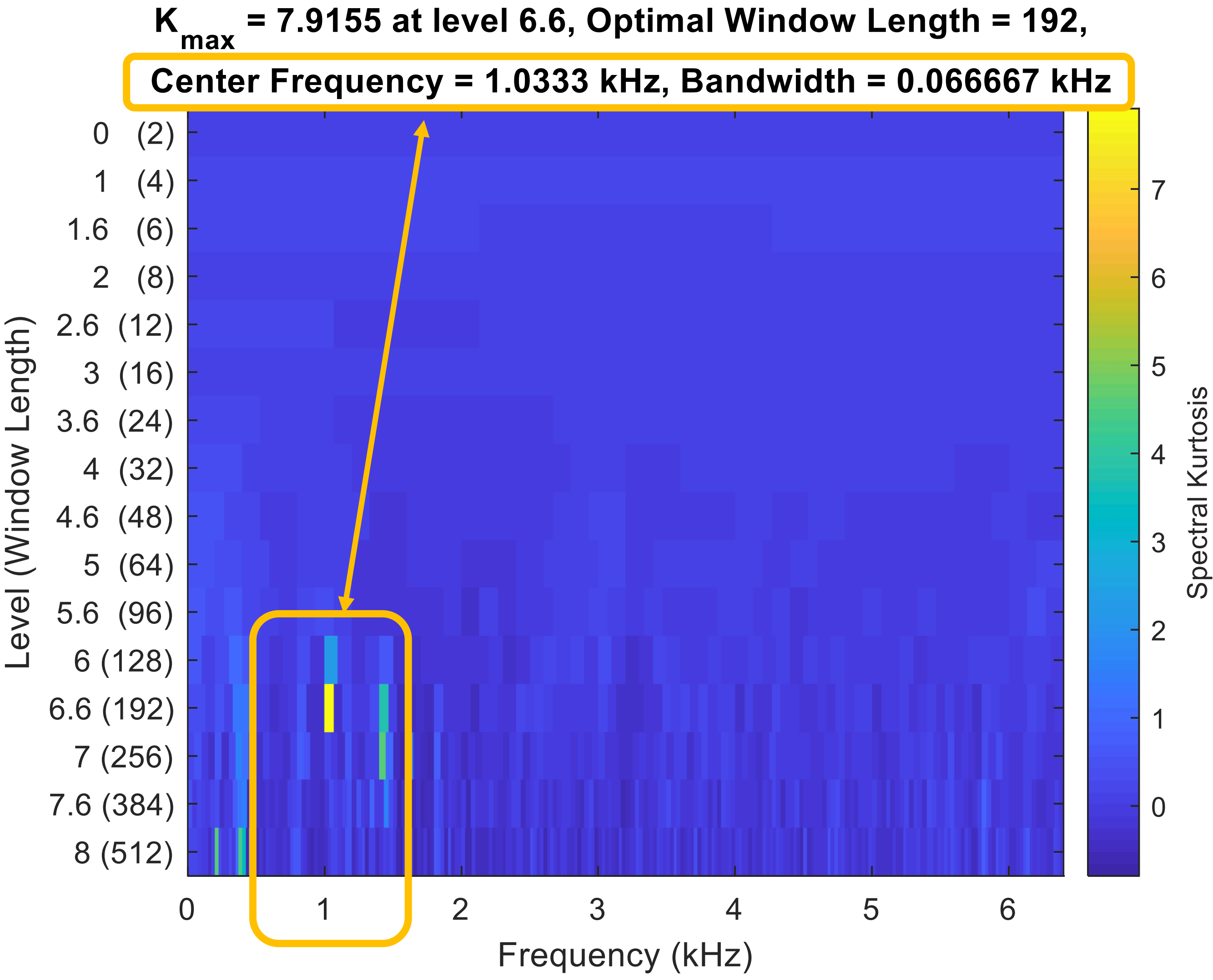
The MatLab **`kurtogram`** algorithm computes the SK at different «frequency resolutions», then selecting the one with the highest kurtosis value

Example: spectral kurtosis

- By applying the SK algorithm to the signal $y(t)$, we obtain a graph showing the kurtosys value as the **frequency** and **frequency resolution** (level) vary

- The algorithm also returns an optimal level of **frequency band** in which to filter

- These values should be taken only as an **indicative suggestion** for a correct selection of the filter band. The chosen value $[\omega_l, \omega_h]$ is then **fixed for all tests**

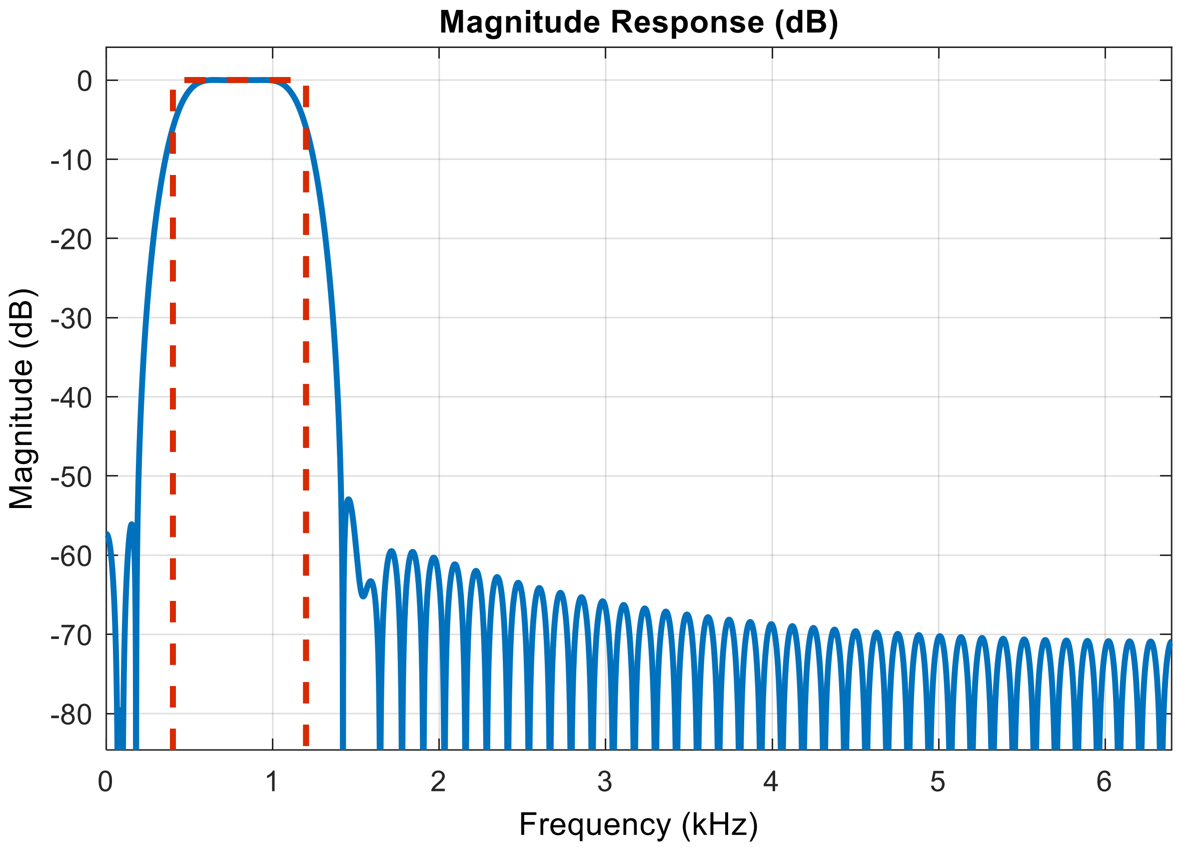


Step 3: band-pass filtering

The $y(t)$ signal is **band-pass filtered** using the values $[\omega_l, \omega_h]$ selected with the help of the SK, obtaining the $r(t)$ signal

It is possible, for simplicity, to use a **FIR filter** of an appropriate order based on the required level of attenuation

In the example, we chose a FIR filter of order 100 centered at $f_c = 800\text{ Hz}$ and bandwith $b = 800\text{ Hz}$



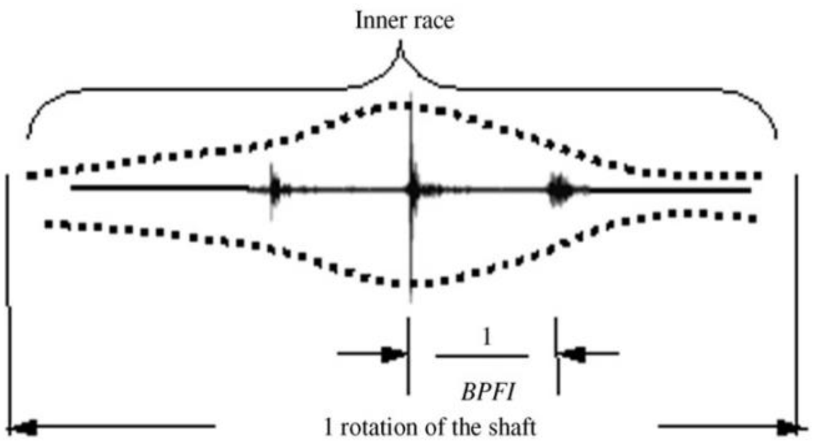
Step 4: envelope analysis

Compute the **envelope** $h(t)$ of the filtered band-pass signal $r(r)$

In the case of an **inner race fault**, the accelerometric signal is modulated by the **rotation speed** of the shaft (which runs at the same speed as the inner race)

The **fault information** is contained in the «time» between the pulses $\left(\frac{1}{\text{BPFI}}\,\text{s}\right)$ rather than in their frequency content

The computation of the signal envelope allows to be **more robust** compared to variations in the «distance» between the pulses (due to changes in contact angle, loads, ...)

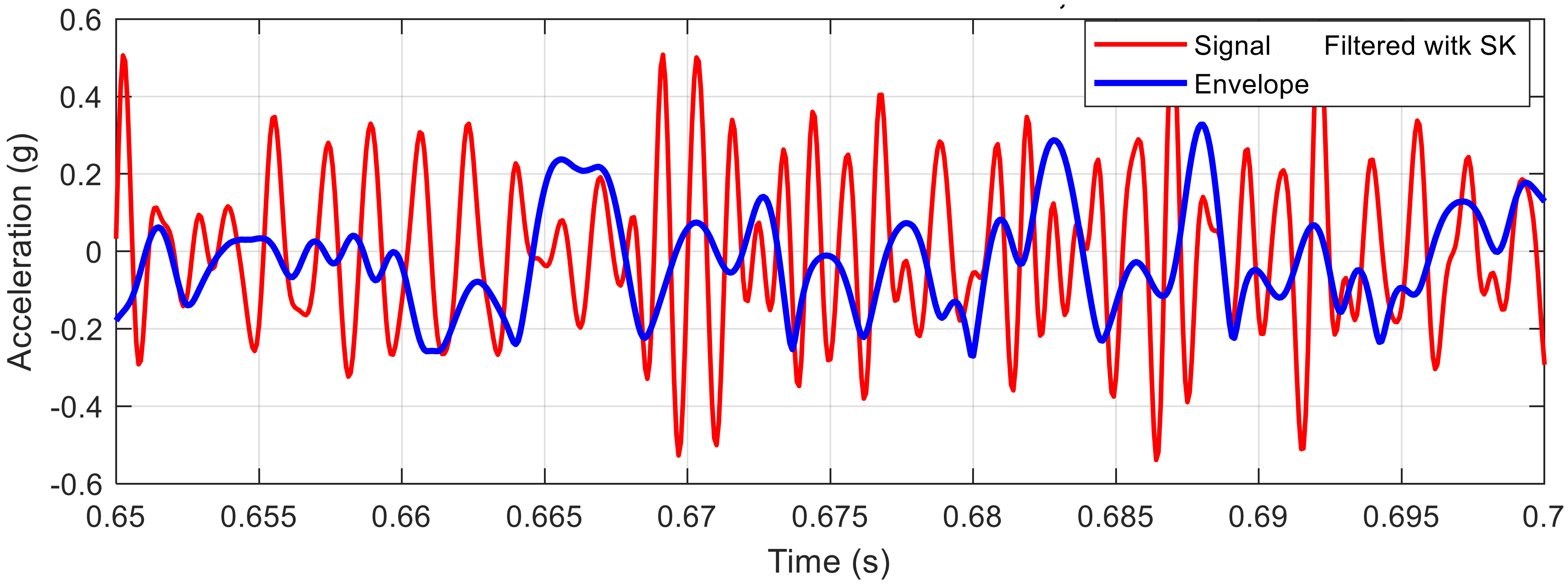


Accelerometric signal in the presence of an inner race fault

Step 4: envelope analysis

The plot shows the envelope $h(t)$ computed on the signal $r(t)$

This signal $h(t)$ can be used for **computing the FFT** and diagnosing the fault

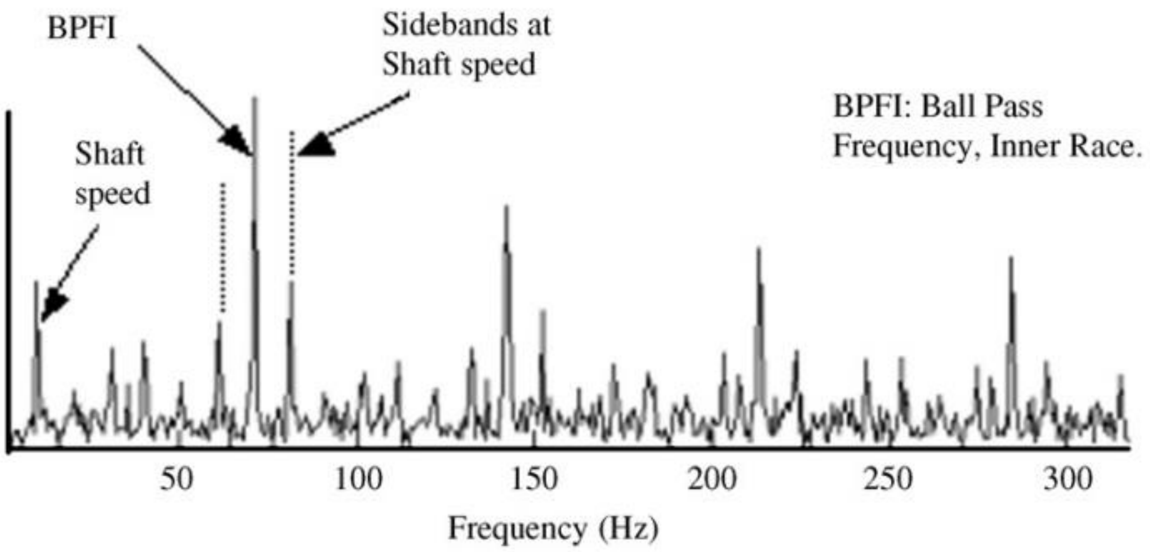


Step 5: envelope analysis

The FFT $H(f)$ of the envelope $h(t)$ is compute to **diagnose the fault**

The aim is to observe the presence of a **high amplitude** value in the typical **fault frequencies** (BPFI and multiples), and to evaluate its energy **compared to** the case of a **healthy bearing**

It is very probable, in the case of an inner race fault, to observe a **modulation** due to the frequency of rotation of the bearing shaft (and the inner race) at the frequencies of the fault



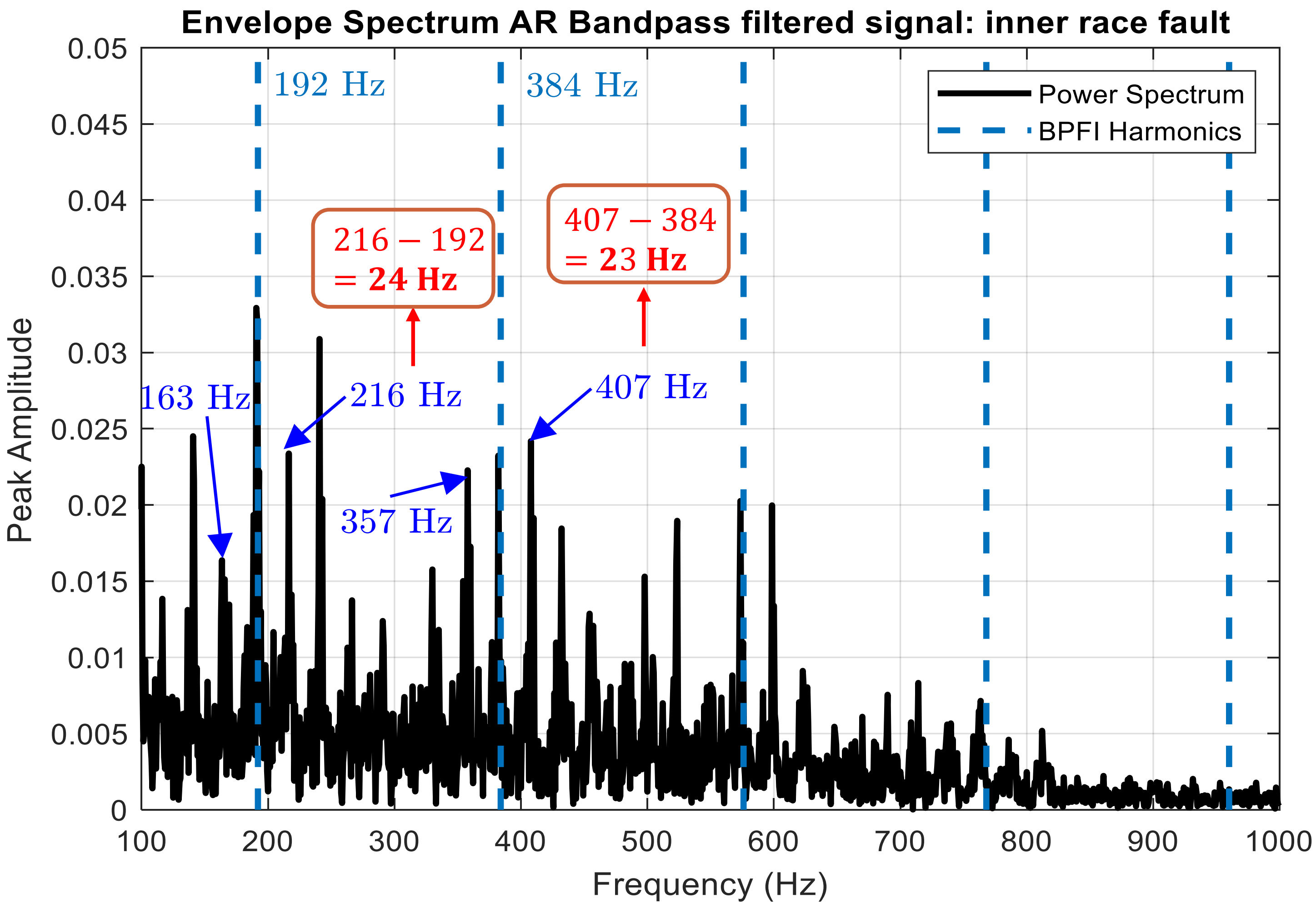
Example: envelope analysis

BPFI fault frequencies are clearly visible

Note the **modulation of the BPFI** frequencies with the **rotation frequency** of the bearing

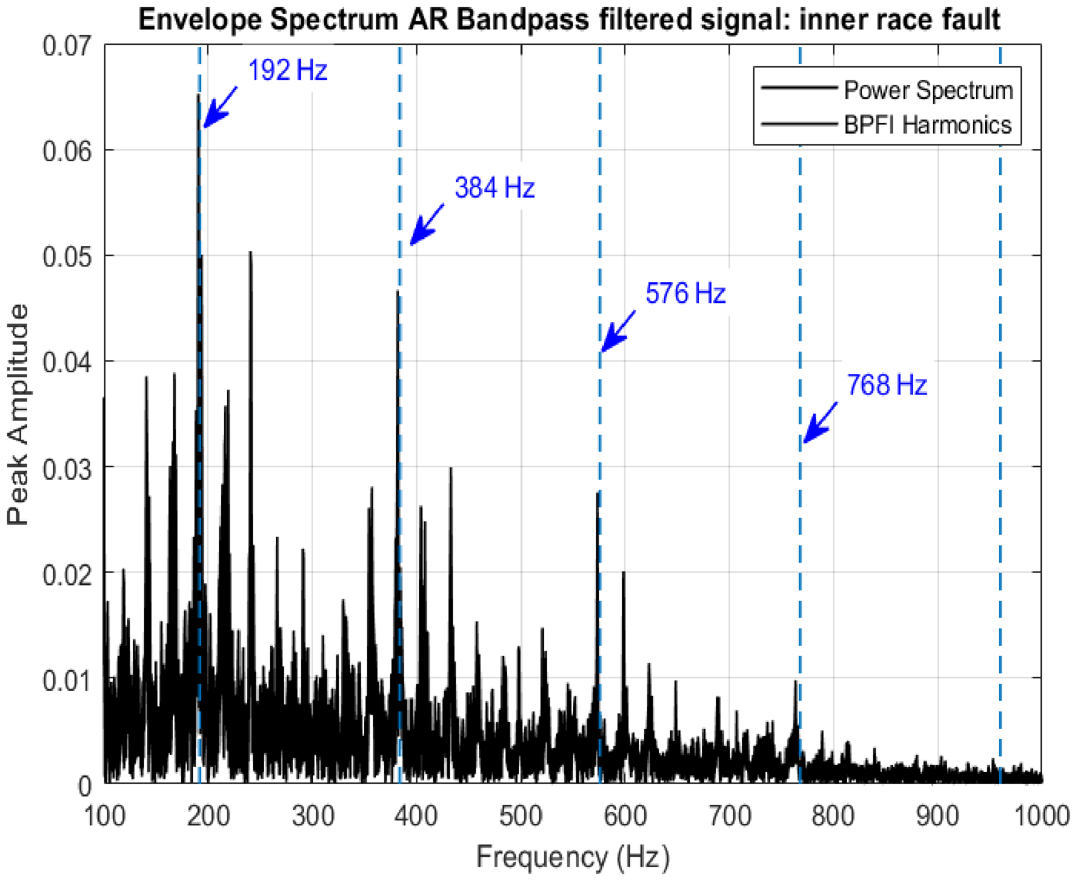
$$\omega = \frac{30000\ \text{mm/min}}{20\ \text{mm/round}} = 1500\ \text{round/min}$$

$$f_m = \frac{\omega}{60\ \text{s/min}} = \frac{1500\ \text{round/min}}{60\ \text{s/min}} \boxed{= 25\ \textbf{Hz}}$$

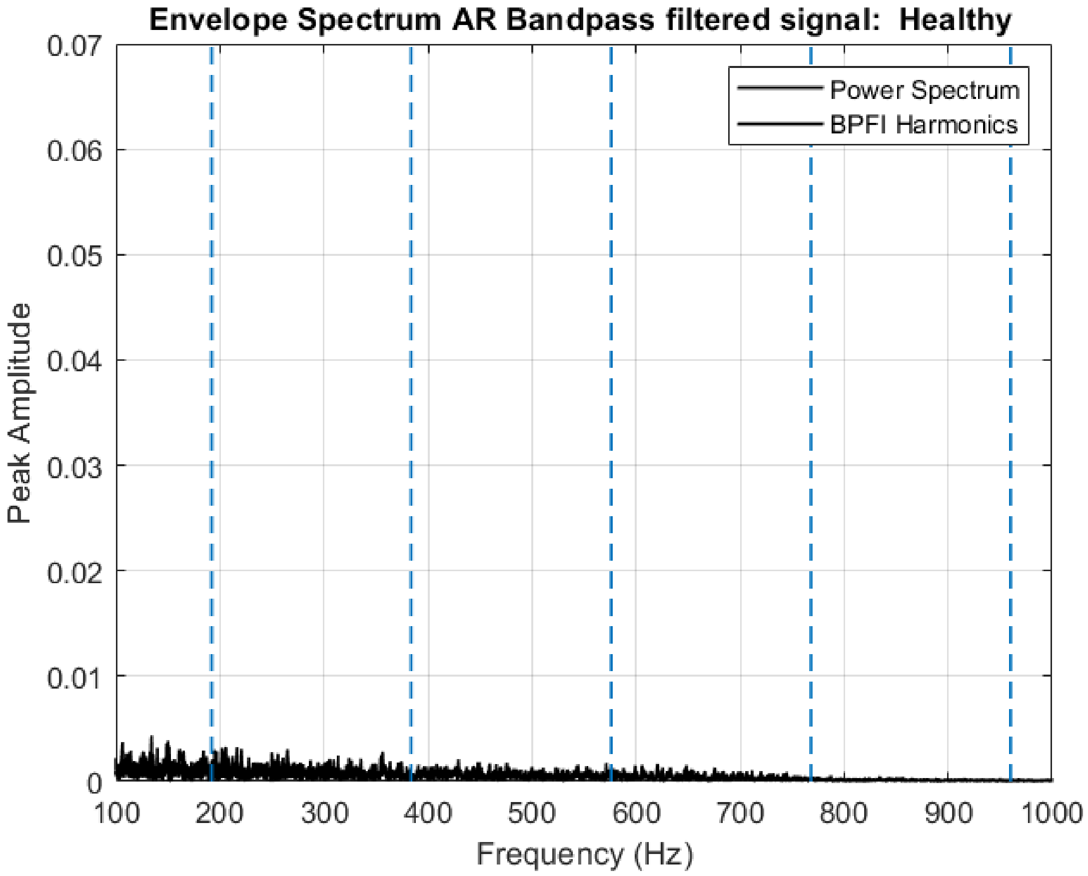


Example: envelope analysis

Inner race fault bearing



Healthy bearing



References

1. Randall, Robert B., and Jerome Antoni. "Rolling element bearing diagnostics—A tutorial." Mechanical systems and signal processing 25.2 (2011): 485-520.
2. Randall, Robert Bond. Vibration-based condition monitoring: industrial, aerospace and automotive applications. John Wiley & Sons, 2011.
3. Scheffer, Cornelius, and Paresh Girdhar. Practical machinery vibration analysis and predictive maintenance. Elsevier, 2004.
4. Antoni, Jérôme, and R. B. Randall. "The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines." Mechanical systems and signal processing 20.2 (2006): 308-331.