

# Wavelet based demodulation of vibration signals generated by defects in rolling element bearings

C.T. Yiakopoulos and I.A. Antoniadis

National Technical University of Athens, Department of Mechanical Engineering, Machine Design and Control Systems Section, P.O. Box 64078, Athens 15710, Greece  
E-mail: antogian@central.ntua.gr

Received 11 September 2000

Revised 10 January 2002

**Abstract.** Vibration signals resulting from roller bearing defects, present a rich content of physical information, the appropriate analysis of which can lead to the clear identification of the nature of the fault. The envelope detection or demodulation methods have been established as the dominant analysis methods for this purpose, since they can separate the useful part of the signal from its redundant contents. The paper proposes a new effective demodulation method, based on the wavelet transform. The method fully exploits the underlying physical concepts of the modulation mechanism, present in the vibration response of faulty bearings, using the excellent time-frequency localization properties of the wavelet analysis. The choice of the specific wavelet family is marginal to their overall effect, while the necessary number of wavelet levels is quite limited. Experimental results and industrial measurements for three different types of bearing faults confirm the validity of the overall approach.

Keywords: Wavelets, bearing, demodulation, envelope, fault detection

## 1. Introduction

Bearings are one of the most important and frequently encountered components in the vast majority of rotating machines, their carrying capacity and reliability being prominent for the overall machine performance. Therefore, quite naturally, fault identification of roller bearings has been over the years the subject of extensive research [14]. Vibration analysis has been established as the most common and reliable analysis method. Defects or wear cause impacts at frequencies governed by the operating speed of the unit and the geometry of the bearings, which in turn excite various machine natural frequencies. Several methods exist to exploit this physical effect, based either directly on the shape of the time domain form of the signal, or on its spectral content.

From all those methods, demodulation or enveloping based methods offer a stronger and more reliable

diagnostic potential, since they are based on a more solid physical background. The corresponding physical mechanism is described in [8]. The general assumption with the enveloping approach is that a measured signal contains a low-frequency phenomenon that acts as the modulator to a high-frequency carrier signal. In bearing failure analysis, the low-frequency phenomenon is the impact caused by a small spur or crack; the high-frequency carrier is a combination of the natural frequencies of the associated rolling element or even of the machine. The goal of the enveloping is first to isolate the measured signal in a relatively narrow frequency band around a specific natural frequency using a band-pass filter and then demodulate it to produce a low-frequency signal, called the “envelope”.

Several demodulation methods have been used to identify faults in rolling element bearings. A “hardware based” approach, proposed in [3], involves the

following steps. The measured signal is passed through a band-pass filter to remove all low frequency high-amplitude signals to keep the dynamic range of the signal within the capabilities of the instruments. The band-pass-filtered signal is passed through a diode, retaining only the positive content. The rectified signal is then low-pass-filtered to remove the high frequency content. The resulting signal is the low frequency modulation with a DC component. This signal is passed through a capacitor (AC coupled) to produce the demodulated time waveform.

Alternative to "hardware based" approaches, other demodulation approaches have been also used, based on the Fourier Transform. An approach, based on the direct use of the FFT, is proposed in [6]. First, an FFT is applied to the N measured, rectangularly windowed data points, the lowest  $[(N/2) + 1]$  FFT coefficients are multiplied by two and the remaining coefficients are set to zero. Then, an inverse FFT is applied to the N modified FFT coefficients, resulting to an N point pre-envelope. The squared magnitude for each of the N pre-envelope points leads to the final envelope. The advantage of an FFT-based envelope is that, if the frequency content of the modulating signal and of the measured modulated signal do not intersect, an exact copy of the true envelope can be recovered.

A more advanced method is based on the proper combination of the FFT with the Hilbert transform [11]. The measured signal is passed through a band-pass filter, in order to isolate a specific high-frequency band, that presents in the spectrum relatively high amplitude components, corresponding presumably to a specific natural frequency of the machine. This step can be omitted in many cases with negligible effect. The band-pass-filtered signal is then converted into an analytical signal. This analytical signal is a complex signal, the real part of which is the band-pass-filtered signal and the imaginary part is the Hilbert transform of the band-pass-filtered signal. The magnitude of the analytical signal corresponds to the envelope of the measured signal.

This study presents a new demodulating approach, based on the Wavelet Transform (WT). The WT has been established as the most widespread joint time-frequency analysis tool [7,9,13] due among others, to its inherent capability to be realized in real-time in the form of a Discrete Wavelet Transform (DWT). The WT overcomes the known disability of the Fourier Transform to represent local features of the signal, such as the quite typical impulses, present in the vibration response of faulty roller bearings. It has already been

used with success in specific case studies for bearing fault detection [5,10,12], as well as for other machine components [1,2,15,16]. The purpose of using the WT in the proposed method, is to obtain the envelope of the vibration response of faulty bearings, based on the physical mechanism that generates the modulation effect and taking into full account its underlying physical concepts and major conclusions.

Part 2 of the paper summarizes the basic physical concepts describing the modulation mechanism, inherent in faulty bearing response. Part 3 performs a brief review of the basic concepts of wavelet analysis, with special emphasis on their behavior in the frequency domain. Part 4 describes the proposed method and analyses the major parameters affecting its performance. Experimental results and industrial measurements for three different types of bearing faults are provided in part 5, verifying the effectiveness of the method.

## 2. Modulation of vibration signals generated by roller bearing defects

Whenever a defect present in one surface of a bearing strikes another surface, an impact results, exciting the resonances of the bearing and of the overall mechanical system. Thus, the pulsation generated by rolling bearing defects, excites vibration at specific defect frequencies as well as a high-frequency response in the overall machine structure. A well established physical model for the vibration produced by a single point defect on the inner race of a rolling element bearing under radial load has been proposed in [8], describing the amplitude modulation of the excitation forces and the corresponding response. The model incorporates the effects of bearing geometry, shaft speed, bearing load distribution, transfer function and the exponential decay of vibration.

Initially, the impact produced by the defect striking one of the rolling elements is modeled as an ideal impulse, denoted by the Dirac function  $\delta(t)$ . The magnitude of the impulse is depended on the severity of the worn and also on the load on the defect at the time of impact. As the inner race of the bearing rotates, the impacts occur periodically at the Ball Passing Frequency Inner race frequency (BPFI), which is defined [14] by:

$$f_{BPFI} = \frac{m}{2} \left[ 1 + \frac{BD}{PD} \cos \beta \right] f_r \quad (1)$$

where  $m$  is the number of rolling elements,  $BD$  is the rolling element diameter,  $PD$  is the pitch circle

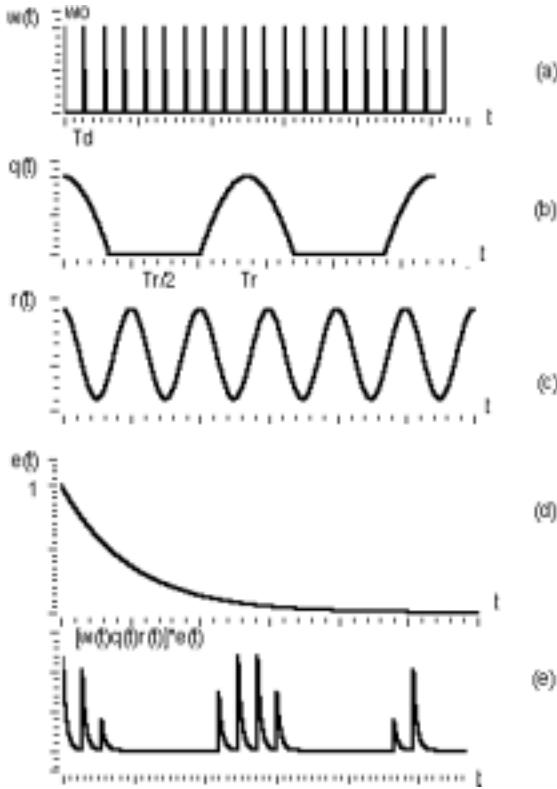


Fig. 1. Waveforms involved in the generation of the envelope of vibrations produced by an inner race defect under radial load [7]: (a) Impacts produced under a constant uniform unit load, (b) Load distribution in the bearing (Stribeck equation), (c) Transfer function between the worn and the fixed measurement point, (d) Typical response decay law, (e) Final envelope.

diameter,  $\beta$  is the contact angle and  $f_r$  is the shaft rotation frequency.

Thus, the impacts produced by a single point defect under a constant uniform unit load can be modeled as an infinite series of impulses, shown in Fig. 1(a):

$$w(t) = w_0 \sum_i \delta(t - iT_{BPFI}) \quad (2)$$

where  $w_0$  represents the magnitude of the impulses and  $T_{BPFI} = 1/f_{BPFI}$  is the period between the impacts. The load distribution in the bearing is assumed to follow the Stribeck equation, graphically shown in Fig. 1(b):

$$q(t) = q_{\max} \left\{ 1 - \frac{1}{2\varepsilon} [1 - \cos(2\pi f_r t)] \right\}^n \quad (3)$$

for  $|Arg(2\pi f_r t)| < \vartheta_{\max}$

$q(t) = 0 \quad \text{elsewhere}$

where  $q_{\max}$  is the maximum load intensity,  $\vartheta_{\max}$  is the angular extent of the load zone,  $\varepsilon$  is the load distribution

factor and  $n$  is a constant. As the bearing rotates, the transmission path and hence the transfer function between the worn, where the impacts occur, and the fixed measurement point, vary. This rotation effect is taken into account by the introduction of a transfer function  $r(t)$ , the approximate form of which is shown graphically in Fig. 1(c). Both  $q(t)$  and  $r(t)$  are periodic, with repetition period  $T_r = 1/f_r$ .

The train of impulses  $w(t)$ , generated by a constant uniform unit load, is multiplied by the actual load  $q(t)$ , experienced by the defect, to give the actual impulses delivered at the location of the defect. These impulses are then multiplied by the amplitude of the transfer function  $r(t)$  between the defect and the fixed measurement point. Thus, the excitation on the fixed structure of the machine is described by a force in the form:

$$f(t) = w(t)q(t)r(t) \quad (4)$$

This excitation force, dependent on its location and spectral content, excites a number of machine natural frequencies. The total vibration response can be described by an equation in the form:

$$x(t) = f(t) \otimes h(t) = [w(t)q(t)r(t)] \otimes h(t) \quad (5)$$

where  $h(t)$  is the impulse response function of the entire machine and  $\otimes$  denotes the convolution operator.

In order to properly identify the fault, only the shape of the impulse sequence, as described by the excitation force pattern  $f(t)$  in Eq. (4), is necessary. Thus, the objective of the fault identification procedure, is to remove from the final response  $x(t)$  in Eq. (5) all its spectral contents resulting from the structural natural frequencies, and isolate its envelope in the following form:

$$v(t) = [w(t)q(t)r(t)] \otimes e(t) \quad (6)$$

where  $e(t)$  represents the typical decay of a resonance excited by the impacts, assumed to follow an exponential decay law with a time constant  $\tau$ , which is independent of the position at which the impact occurs (Fig. 1(d)):

$$\begin{aligned} e(t) &= e^{-t/\tau} \quad \text{for } t > 0 \\ e(t) &= 0 \quad \text{elsewhere} \end{aligned} \quad (7)$$

A typical form of the requested envelope is shown in Fig. 1(e).

A signal simulating the vibration response of a bearing with an inner race fault is shown in Fig. 2(a). This signal is generated according to Eq. (5), assuming a rotation frequency of 36 Hz, a BPFI frequency of 181 Hz and a single structural resonance at 2000 Hz. Its cor-

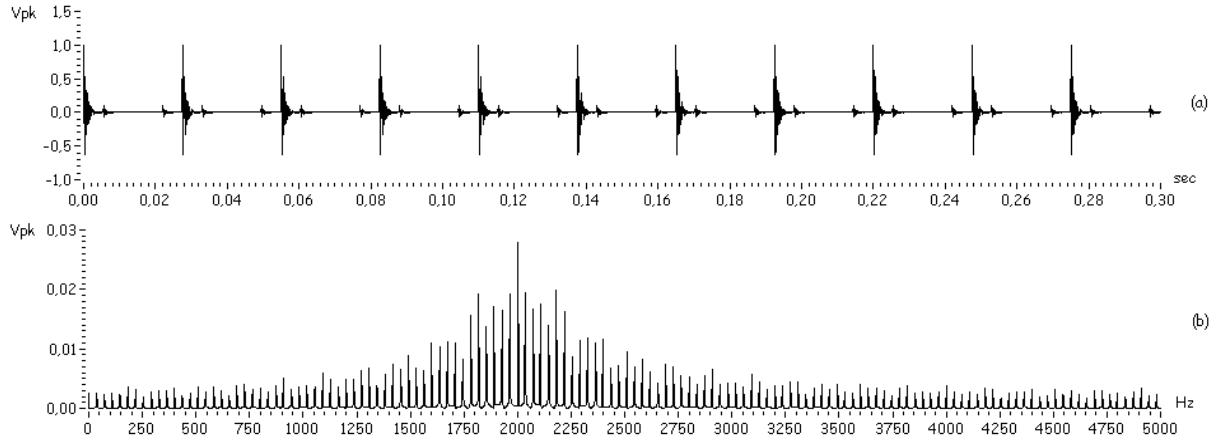


Fig. 2. Signal simulating the vibration response of a faulty bearing under an inner race fault: (a) Time waveform, (b) Spectrum (Shaft Speed: 36 Hz, BPFI: 181 Hz, Natural Frequency: 2 KHz, Sampling Frequency: 10 KHz).

responding spectrum with a sampling rate of 10 KHz is shown in Fig. 2(b). Although just a single structural resonance exists, the spectrum occupies the entire frequency band, due to the various modulation effects. This clearly indicates the difficulty of bearing fault identification, based just on FFT analysis.

For other typical bearing faults (e.g. outer race faults, ball faults, etc.), the physical modulation mechanism is essentially the same one to the inner race fault mechanism, the basic difference being in the shape of the envelope.

### 3. Wavelet analysis

The Fourier Transform (FT) is a linear expansion of the signal into sinusoidal waveforms that have infinite length in time and that are extremely localized in frequency. This results to the total loss of the time-information in the frequency domain. An improvement of the Fourier transform is given by the Windowed Fourier Transform, called the Short Time Fourier Transform (STFT). The STFT is just a series of FTs, performed on successive portions of a waveform. This approach does introduce the opportunity to identify time dependent variations in the structure of the waveform at various scales, as the window, over which the FT is computed, is moved along the longer waveform. However, a fixed window size must be used for a given STFT, and in order to obtain good resolution for the frequencies that compose the signal, a long window is required. The STFT retains the time information but has strong time-frequency resolution limitations. If

shorter windows are chosen, then one will have a higher time resolution but a coarser frequency resolution. On the other hand, if longer windows are chosen, then one will have a higher frequency resolution but a coarser time resolution.

To overcome the limitations of the fixed resolution of the STFT in frequency and time domains, a new method, based on wavelets, has been developed [7,9]. The wavelet transform (WT) is a mathematical tool that permits the decomposition of a signal in terms of elementary contributions, called wavelets. Time-domain wavelets are simple oscillating amplitude functions of time, that have large fluctuating amplitudes during a restricted time period and very low amplitude or zero amplitude outside this time period.

The wavelets are obtained from a single function  $\psi(t)$  by translation and dilation:

$$\psi(\alpha, \tau)^{(t)} = \frac{1}{\sqrt{\alpha}} \psi\left(\frac{t-\tau}{\alpha}\right) \quad (8)$$

where  $\alpha$  is the so-called scaling parameter,  $\tau$  is the time localization parameter and  $\psi(t)$  is called the ‘‘mother wavelet’’. The parameters of translation  $\tau \in R$ , and dilation  $\alpha > 0$ , may be continuous or discrete.

The WT of a finite energy signal  $x(t)$  with the analyzing wavelet  $\psi(t)$  is the convolution of  $x(t)$  with a scaled and conjugated wavelet:

$$\begin{aligned} W(\alpha, \tau) &= x(t) \otimes \frac{1}{\sqrt{\alpha}} \psi\left(-\frac{t}{\alpha}\right) \\ &= x(t) \otimes \frac{1}{\sqrt{\alpha}} \psi^*\left(\frac{t}{\alpha}\right) \\ &= \frac{1}{\alpha} \int_{-\infty}^{\infty} x(\tau) \psi^*\left(\frac{t-\tau}{\alpha}\right) d\tau \end{aligned} \quad (9)$$

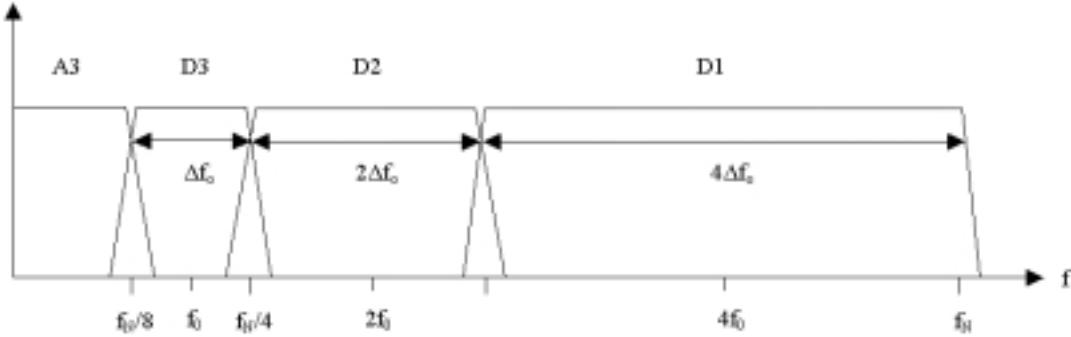


Fig. 3. Presentation of a three level wavelet analysis in the frequency domain. The signal is decomposed in a low resolution signal A3 and three detail functions D1–D3, with corresponding spectra shown.

where  $*$  denotes the complex conjugate.

Expression Eq. (9) can take the following alternative form:

$$\begin{aligned} W(\alpha, \tau) &= \sqrt{\alpha} \int_{-\infty}^{\infty} X(f) \Psi^*(2\pi\alpha f) e^{i2\pi f \tau} df \\ &= \sqrt{\alpha} F^{-1}\{X(f) \Psi^*(2\pi\alpha f)\} \end{aligned} \quad (10)$$

where  $X(f)$  and  $\Psi(f)$  are the Fourier transforms of  $x(t)$  and  $\psi(t)$  respectively, and  $F^{-1}$  denotes the Inverse Fourier Transform.

Equations (9) and (10) show that the wavelet analysis is a time-frequency analysis, or, more properly, a time-scaled analysis. In particular, Eq. (10) shows that the WT acts as also as filter.

There exist many methods [9] to compute in practice the WT of a waveform. They can be divided in two major classes: I) Methods based on the numerical computation of the Continuous Wavelet Transform (CWT), II) Methods using specially designed filters, that generate a highly efficient Discrete Wavelet Transform (DWT), also known as “Multiresolution Decomposition”.

The CWT is not computationally efficient. The information it displays at closely spaced scales or at closely time points is highly correlated and therefore unnecessarily redundant for many applications. The DWT provides a non-redundant, highly efficient wavelet representation, that can be implemented with a simple recursive filter scheme. The DWT produces only as many coefficients as the number of samples within the original signal, without the loss of any information at all.

A vibration waveform can be decomposed into its DWT coefficients through a simple recursive filter scheme, that consists of a high pass filter and a low pass filter, whose filter coefficients are uniquely determined by the particular wavelet shape used in the analysis.

Different wavelet shapes are associated with different filter coefficient sequences. Regardless of the wavelet used, the filters that produce the detail functions and the low resolution signal in a DWT have a variable bandwidth that depends on the center frequency of those filters. Figure 3 schematizes the frequency spectrum for each scale in a typical three-level DWT. The figure shows that each successive detail function in a DWT (D1, D2, D3) has a spectrum with a center frequency at  $f_{o,j}$  (level  $j = 1, 2, 3, \dots$ ) and a bandwidth  $\Delta f_{o,j}$  half than that of the previous detail function. Thus, the frequency resolution improves by a factor of 2 for each successively larger scale in a DWT and the time resolution correspondingly decreases by a factor of 2. Conversely, the time resolution improves by a factor of 2 at successively larger scales and the frequency resolution correspondingly decreases by a factor of 2.

The center frequency  $f_{o,j}$  and bandwidth  $\Delta f_{o,j}$  of the  $j$ th wavelet's spectrum become:

$$f_{o,j} = \frac{f_N}{2^j} + \frac{f_N}{2^{j+1}}, \quad j = 1, 2, 3, \dots \quad (11a)$$

$$\Delta f_{o,j} = \frac{f_N}{2^{j-1}} - \frac{f_N}{2^j}, \quad j = 1, 2, 3, \dots \quad (11a)$$

where  $f_N$  is half the sampling frequency of the signal (Nyquist frequency).

At the wavelet analysis stage shown in Fig. 3, the DWT analysis has four basic parts, namely a first detail function D1 that captures the high frequencies between 1/2 the Nyquist frequency and the Nyquist frequency, a second detail function D2 that captures the intermediate frequencies between 1/4 the Nyquist frequency and 1/2 the Nyquist frequency, a third detail function D3 that captures the intermediate frequencies between 1/8 the Nyquist frequency and 1/4 the Nyquist frequency and a low resolution signal A3 that captures the low frequencies between 0 and 1/8 the Nyquist frequency.

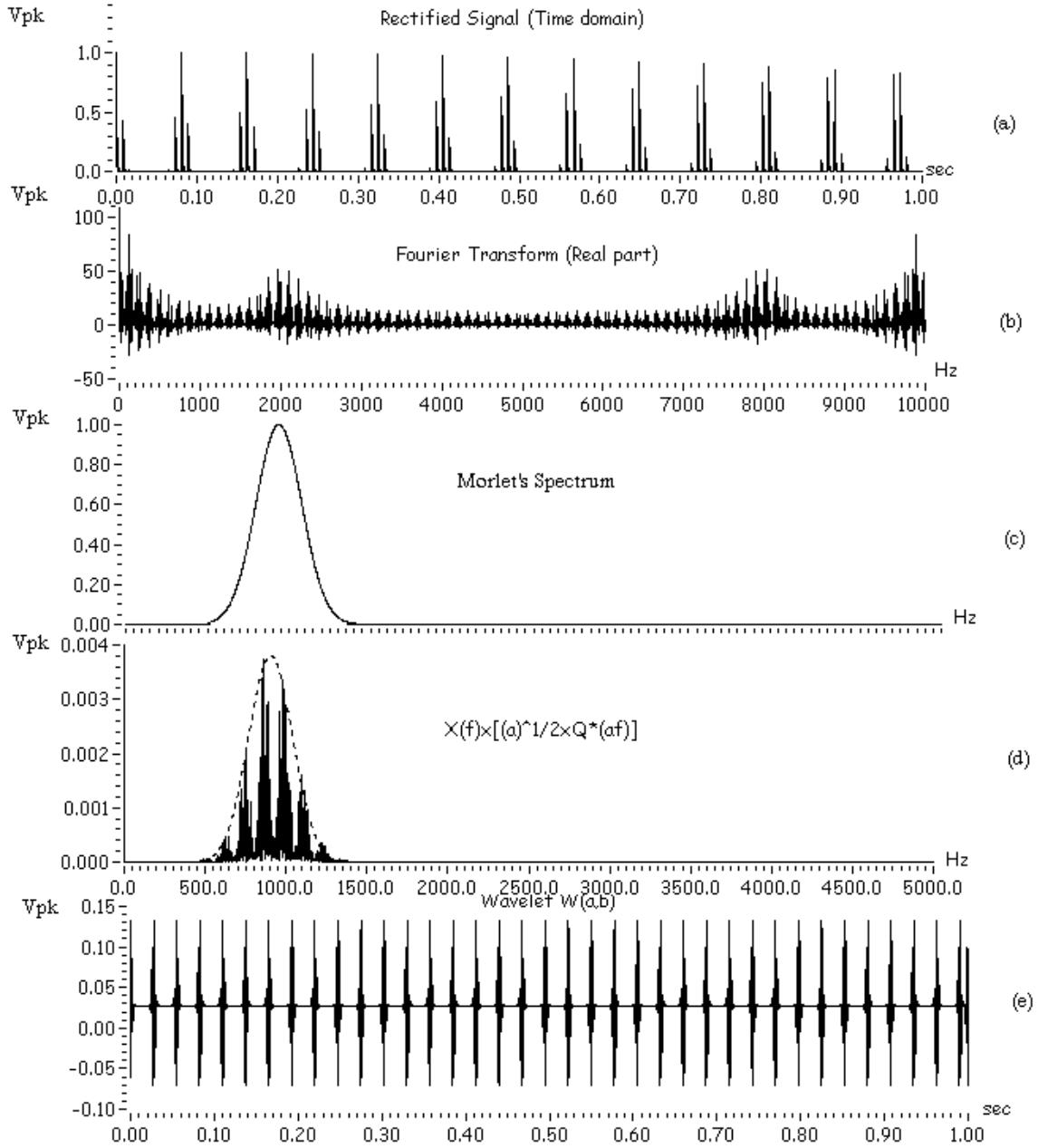


Fig. 4. Indicative application of the WT in the frequency domain: (a) Original waveform of the signal, (b) Fourier transform of the signal, (c) Fourier transform of the Morlet wavelet, (d) Application of the wavelet transform operation of Eq. (10) in the frequency domain, (e) Result of the wavelet transform in the time domain.

Several wavelet families have been developed [7] to define the exact shape of the wavelet  $\psi(t)$ . For demonstration purposes of the basic concepts of wavelet analysis, the complex-valued Morlet wavelet is used as a typical example. The Morlet wavelet is defined in the time domain as a sinusoidal wave multiplied by a Gaussian function:

$$(t) = c_j e^{-\sigma_j^2 t^2} e^{i 2\pi f_{0,j} t} \quad (12)$$

where  $c_j$  is a positive parameter,  $\sigma_j$  determines the width of the wavelet and hence the width of the frequency band and  $f_{0,j}$  is the center frequency of the band. The parameter  $c_j$  is equal to:

$$c_j = \sigma_j / \sqrt{\pi} \quad (13)$$

The Fourier transform of the Morlet's wavelet is given by:

$$\Psi(f) = e^{-\frac{\pi^2}{\sigma_j^2}(f-f_{o,j})^2} \quad (14)$$

where  $\Psi(f) = \Psi^*(f)$ , since  $\Psi(f)$  is real.

This wavelet has a Gaussian shape in the frequency domain, where the center frequency  $f_{o,j}$  is given by Eq. (11a) and the corresponding frequency range for level  $j$  is:

$$\left[ f_{o,j} - \frac{\sigma_j}{2}, f_{o,j} + \frac{\sigma_j}{2} \right] = \left[ \frac{f_N}{2^j}, \frac{f_N}{2^{j-1}} \right] \quad (15)$$

The parameter  $\sigma_j$  becomes:

$$\sigma_j = \frac{f_N}{2^j} \quad (16)$$

The procedure for the calculation of the wavelet transform in the frequency domain using Eq. (10) is demonstrated in Fig. 4. The Morlet wavelet is applied. Figure 4(a) indicates the original signal and Fig. 4(b) the real part of its Fourier Transform. Figure 4(c) indicates the Fourier Transform  $\Psi^*(2\pi\alpha f)$  of the Morlet wavelet, generated according to Eq. (15) for a given center frequency  $f_{o,j}$  (level  $j$ ) and Nyquist frequency  $f_N$ . Figure 4(d) indicates the product  $X(f)\Psi^*(2\pi\alpha f)$ , representing the band-pass filtering of the time-domain signal. Finally, Fig. 4(e) represents the inverse Fourier transform of the filtered signal. This waveform is the wavelet transform of the input signal for a given level  $j$ .

Although the calculation of the Wavelet Transform in practice is performed directly in the time domain using the DWT transform, the different presentation shown in Fig. 4 has been chosen to clearly illustrate the effect of the wavelet transform in the frequency domain.

#### 4. Wavelet based demodulation

The efficiency of the WT can be fully exploited in the demodulation of vibration signals, resulting from bearing defects. The development of the proposed enveloping approach is based on the fact that the measured signal, as described in part 2, contains a low-frequency component, which acts as the modulator to a high-frequency carrier signal. The goal of the enveloping approach is to isolate the low-frequency information of the measured signal that contains the percussive frequencies caused by the bearing defect. Thus, the high-frequency carrier signal, which contains the natural frequencies of the associated race or rolling el-

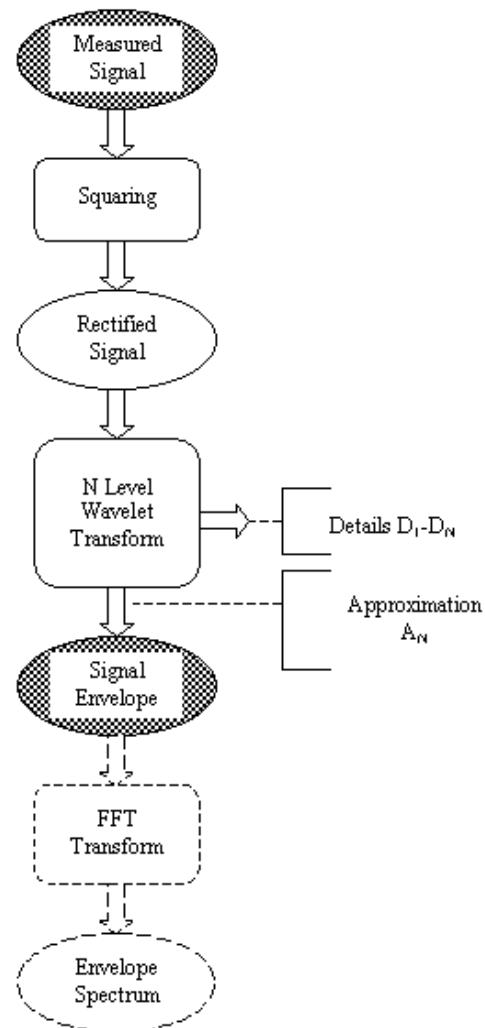


Fig. 5. Schematic presentation of the proposed demodulation method.

ement, is 'filtered' and drawn away of the measured signal. Figure 5 illustrates schematically the proposed approach.

The vibration, measured by an accelerometer mounted on the casing of the machine near the bearing, is first squared to obtain the absolute value of the modulated signal. The squaring procedure offers a number of advantages. First, a waveform is produced, possessing only positive values, as does the final form of the expected envelope. Then, squaring escalates the differences of the peak variations, in order that the peaks are more discrete in the signal. Finally, as shown [4], the squaring procedure, is able to transfer the most important frequency content of the signal to lower frequency bands.

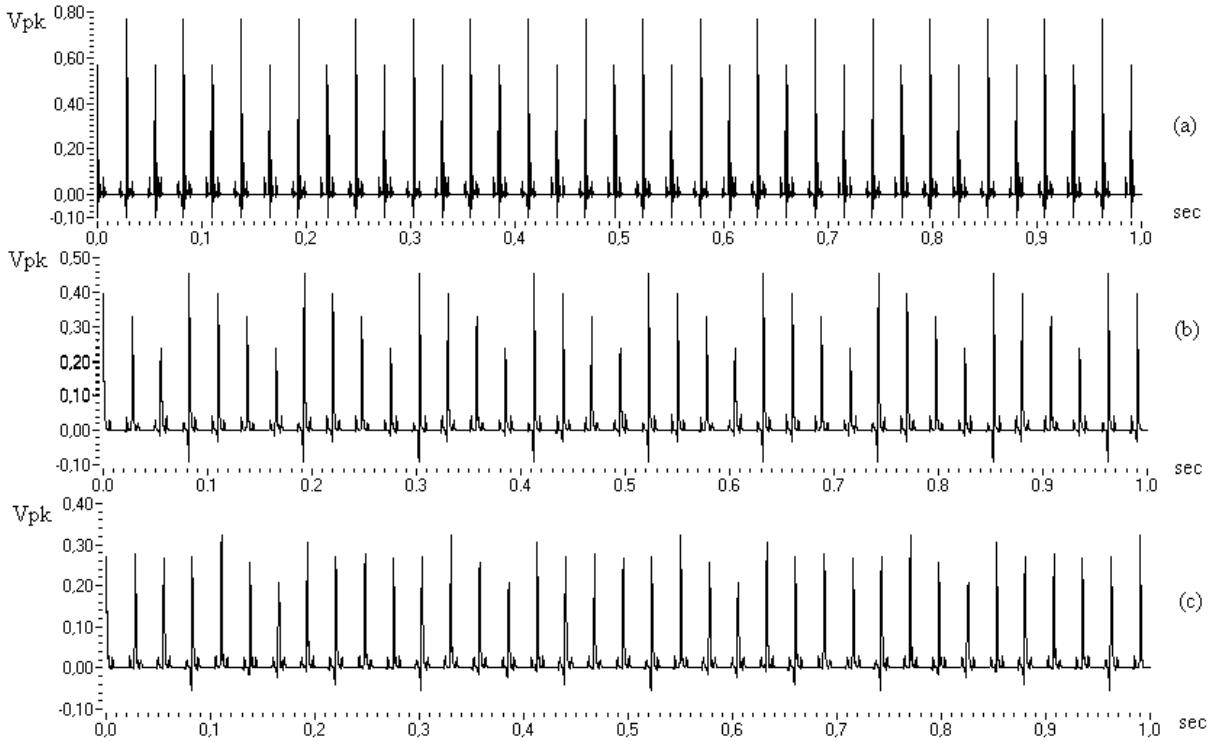


Fig. 6. Low resolution approximations A1, A2, A3 of the signal of Fig. 2, using the Daubechies wavelet family of order 2.

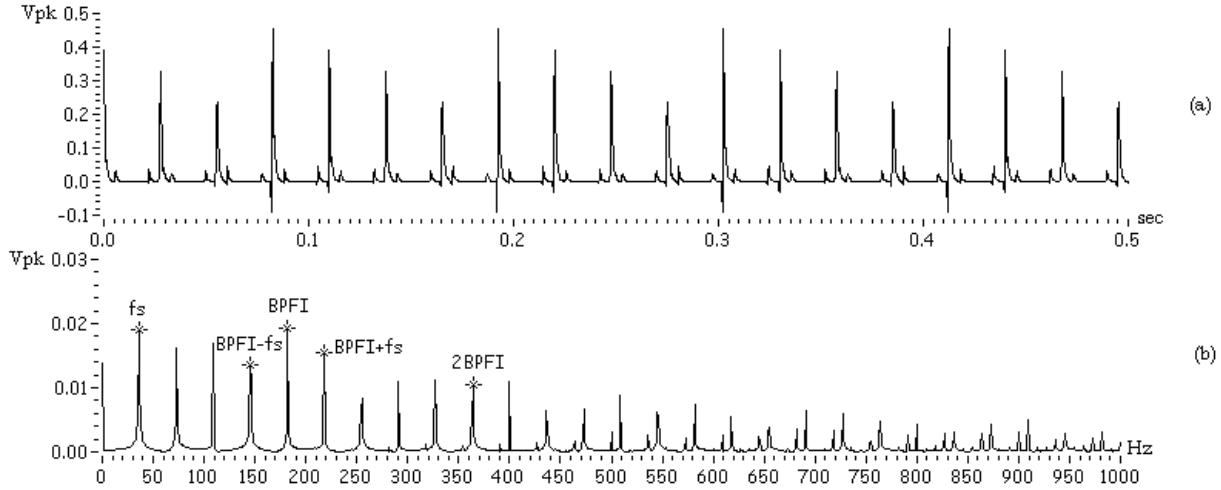


Fig. 7. Details and spectrum of the low resolution approximation A2 of Fig. 2.

Then, using an  $N$  level wavelet transform, the rectified signal is decomposed into its approximation and detail waveforms. The approximation waveform  $A_N$ , which contains the low-frequency components of the signal, is the requested envelope. Optionally, it can be further processed using alternative signal processing

techniques like the FFT transform, in order to derive other specific signal features.

The selection of the proper decomposition-level  $N$  is critical for the method. It depends on the extent of the low-frequency region, where the characteristic bearing frequencies are expected to appear and on the sampling

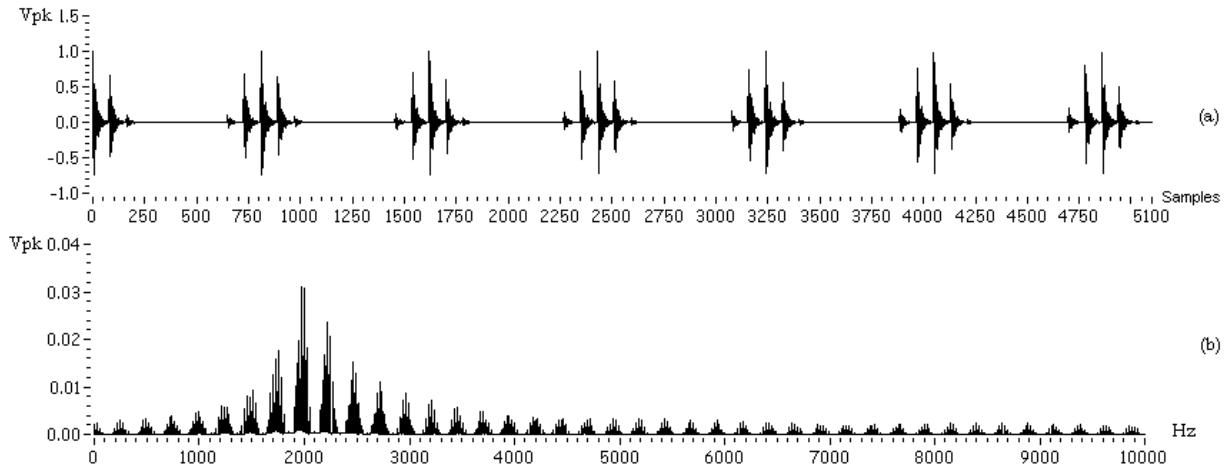


Fig. 8. Alternative signal simulating the vibration response of a faulty bearing under an inner race fault: (a) Time waveform, (b) Spectrum (Shaft Speed: 24.5 Hz, BPFI: 245 Hz, Natural Frequency: 2 KHz, Sampling Frequency: 20 KHz).

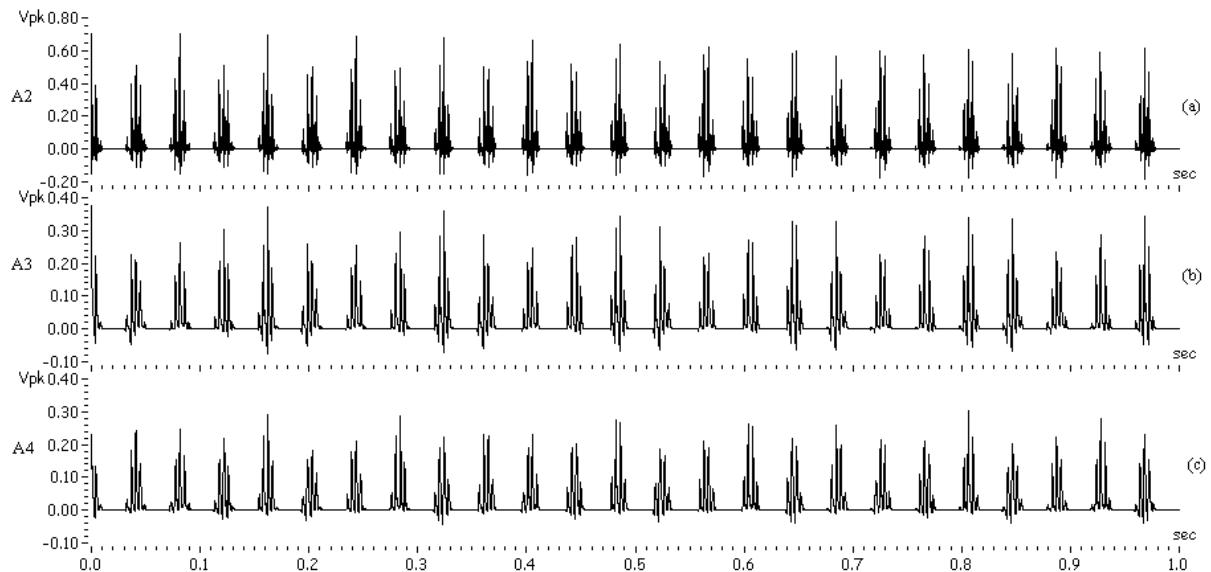


Fig. 9. Low resolution approximations  $A_2$ ,  $A_3$ ,  $A_4$  of the signal of Fig. 12, using the Daubechies wavelet family of order 2.

rate. The first factor is known by the geometry of the bearings that are monitored and on the shaft speed. Characteristic calculation formula, similar in form to Eq. (1), exist in the literature [14]. The selection of the sampling rate determines the total frequency bandwidth of the monitored signal. This bandwidth should be selected as high as necessary, in order to include a number of structural natural frequencies, excited by the characteristic impulses of the bearing defect. Thus, the measured signal includes all the relevant information necessary for allowing the fault features to be properly

exposed. The proper decomposition level  $N$  is then subsequently selected in such a way, that the frequency content of the approximation waveform  $A_N$  completely covers the low frequency region, without intersecting with the frequency band dominated by the structural resonances.

The procedure for the selection of the proper decomposition level is shown in two characteristic examples. In the first example, the signal of Fig. 2 is used. Wavelet decomposition is performed at three different levels  $N = 1, 2$  and  $3$ . The resulting approximations  $A_1$ ,  $A_2$  and

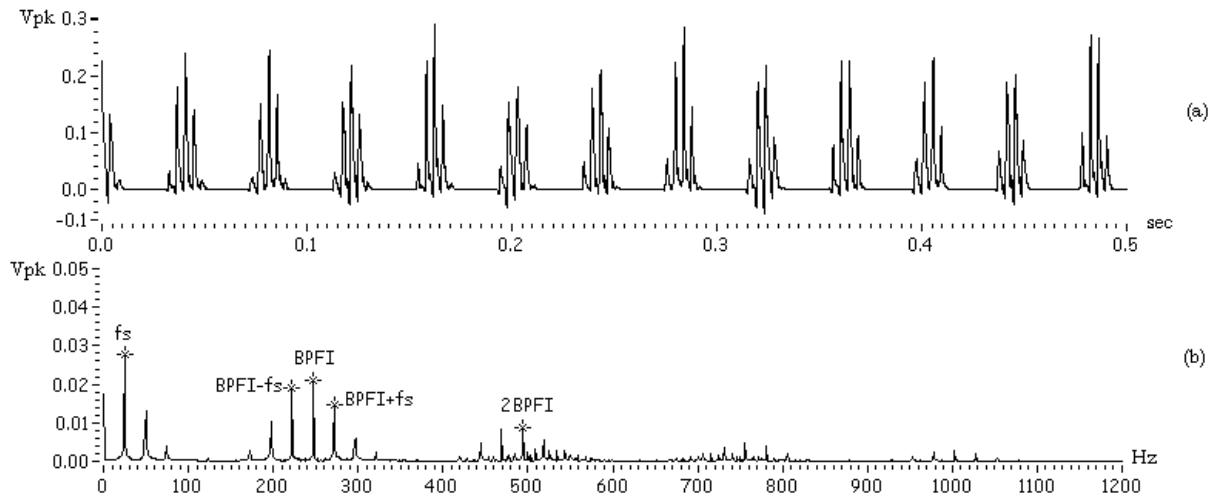


Fig. 10. Details and spectrum of the low resolution approximation A3 of Fig. 8.

A3 respectively, are shown in Fig. 6. Approximation A1 covers the range between 0 Hz and 2500 Hz, half the Nyquist frequency of 5000 Hz. Since the structural resonance of 2000 Hz is also present in this range, the level A1 contains additional high frequency components, which render the demodulation procedure inefficient. Approximations A2 and A3 cover respectively the ranges 0–1250 Hz and 0–625 Hz. Both isolate well the low frequency region of the characteristic bearing fault frequencies from the structural resonance region and as shown in Fig. 6, both are practically equivalent as envelope estimators. Thus, for computational purposes, only analysis up to level  $N = 2$  is necessary in this case. A corresponding detail view of the waveform A2 and its Fourier analysis is shown in Fig. 7. The BPFI frequency of 181 Hz and its modulation by the shaft rotation frequency of 36 Hz, both characterizing the fault, clearly appear now in the spectrum.

The signal used in the second example and its corresponding spectrum are shown in Fig. 8. Wavelet decomposition is performed at three different levels  $N = 2, 3$  and 4 and the resulting approximations A2, A3 and A4 respectively, are shown in Fig. 9. Since approximation A2 covers also in this case the range 0–2500 Hz, where the structural resonance of 2000 Hz is also present, it is inappropriate as envelope estimator. Both approximations A3 and A4, covering respectively the ranges 0–1250 Hz and 0–625 Hz, isolate well the low frequency region from the structural resonance region. Thus, analysis only up to level  $N = 3$  is necessary in this case. A corresponding detail view of the waveform A3 and its spectrum is shown in Fig. 10. The

BPFI frequency and the shaft rotation frequency characterizing the fault, clearly appear again in the spectrum. Compared to the analysis of the first example, although the structural resonance and the characteristic bearing fault frequencies are in the same region, an additional analysis level was necessary, since the sampling frequency was doubled.

The selection of the specific wavelet family to be used has a marginal effect on the method, since the primary effect of the wavelet analysis with respect to the proposed method is to isolate the low frequency component of the signal, preserving its specific local features in time. However, the Daubechies wavelet of order 2 presents a slightly better behavior, indicating a better high frequency component isolation.

## 5. Experimental results

Three characteristic cases are presented, each one being typical of a different type of bearing fault. Case A presents an inner race fault, case B presents an outer race fault and case C a rolling element fault. The measurements in cases A and C were conducted on a machinery fault simulator manufactured by SpectaQuest, in order to study signatures of common machinery faults. The measurements in Case B were conducted on a fan motor at the industrial installations of Aluminium of Greece S.A.

In all cases, the measuring device is based on a Pentium II/266 MHz portable computer, equipped with a PCMCIA DAQCard-1200 data acquisition card from

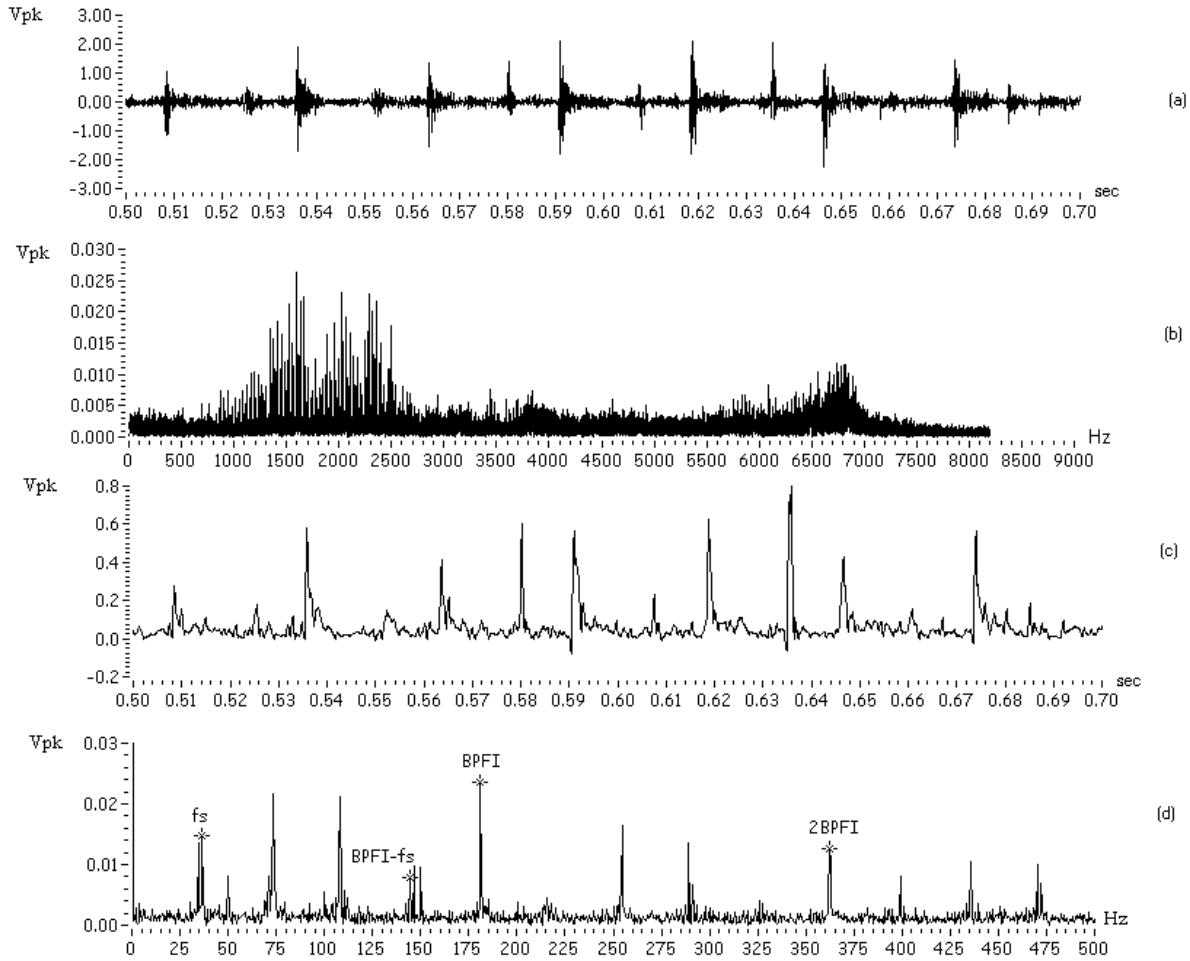


Fig. 11. Measurements and analysis results of the bearing of Case A, representing an inner race fault: (a) Measured signal, (b) Spectrum of the measured signal, (c) Envelope predicted by the proposed approach, (d) Spectrum of the envelope.

National Instruments. This is an 8 channel software-configurable 12-bit data acquisition card, with a total sampling rate capacity of 100 KHz. A B&K type 8325 accelerometer is used, with a sensitivity of 97.3 mV/g and a dynamic range of 1 Hz to 10 KHz. The code of the algorithm that is used in the data acquisition procedure and signal analysis has been developed under the LabVIEW programming environment of National Instruments. The wavelet transform of the measured signal is accomplished at the MATLAB wavelet toolbox.

The bearing examined in Case A consists of 8 balls, has a ball diameter equal to 0.2813 inches, a pitch diameter equal to 1.1228 inches and a contact angle equal to 0 deg. The rotor speed is 36.52 Hz and the sampling frequency is 16384 KHz. Figure 11(a) illustrates the measured signal and Fig. 11(b) the corresponding spectrum. Although a “spiky” behavior is observable

in the signal and a number of modulation indicating side-bands are observable in the spectrum, the source and the nature of the fault cannot be identified without further processing. Figure 11(c) indicates the envelope produced by the demodulation procedure proposed in this paper. A three level wavelet analysis is performed and the approximation function A3 is retained, corresponding to the range 0–1024 Hz. The envelope clearly follows with great accuracy the shape and the local features of the spikes of the original signal, without containing its high frequency components, clearly demonstrating the value of the WT. For a more clear identification of the nature of the fault, Fig. 11(d) presents the spectrum of the envelope. The spectrum of the demodulated signal in Fig. 11(d), compared to the spectrum of the original signal in Fig. 11(b), presents a far more clear structure, revealing peaks to the rotor

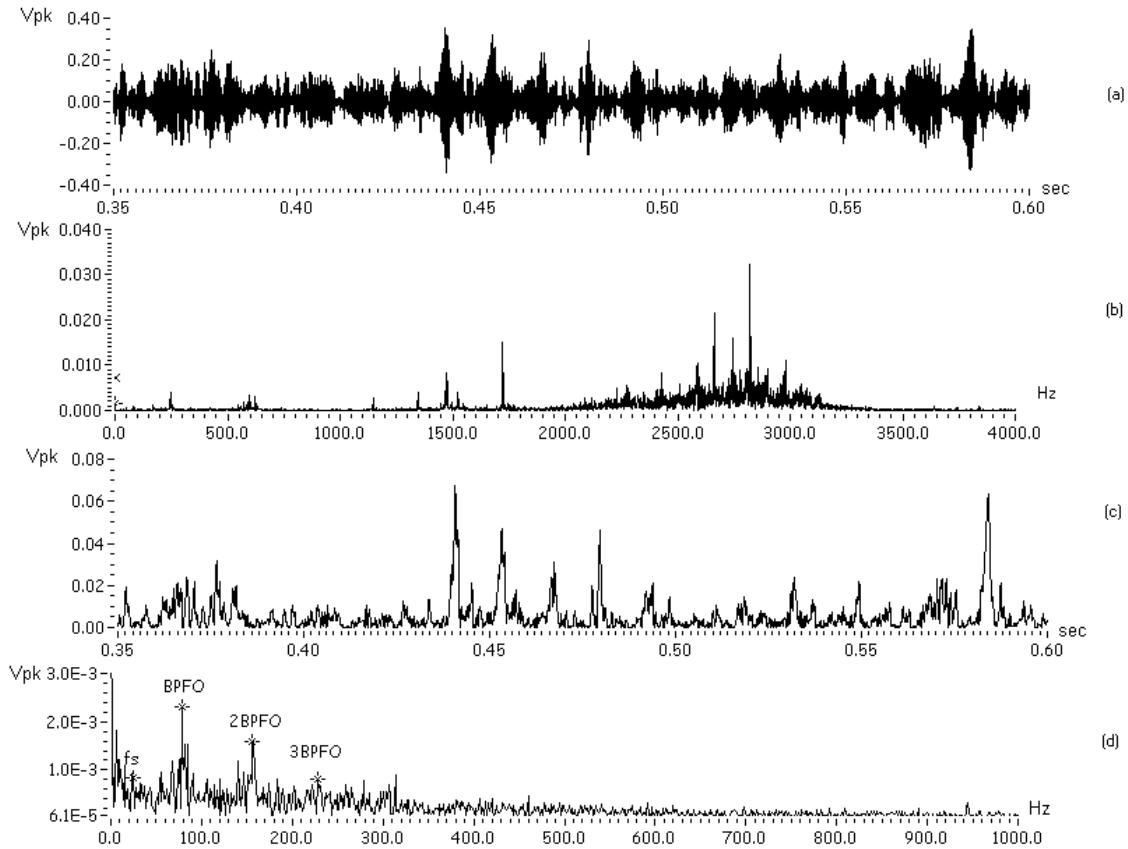


Fig. 12. Measurements and analysis results of the bearing of Case B, representing an outer race fault: (a) Measured signal, (b) Spectrum of the measured signal, (c) Envelope predicted by the proposed approach, (d) Spectrum of the envelope.

speed  $f_s$ , the characteristic bearing frequency BPFI ( $= 181.1$  Hz) and its second harmonic  $2^*\text{BPFI}$ . Therefore, the nature of the fault can be clearly identified.

The bearing examined in Case B is of type 6324MC3 manufactured by SKF. The rotor speed is about 1,500 rpm. The sensor is mounted near the bearing at the horizontal direction. The sampling rate used is 20 KHz and the number of samples is 32,768. The same type of analysis as in Case A is performed, using a 3 level wavelet analysis. The results are presented in Fig. 12. The low-frequency information transferred to the spectrum of the demodulated signal reveals peaks to the rotor speed  $f_s$ , the characteristic Ball Passing Frequency Outer race BPFO ( $= 78.12$  Hz) and its second and third harmonic, confirming also in this Case the validity of the approach.

The last Case C presented, was accomplished with the same type of bearing as in Case A, exhibiting now a fault on the rolling elements. The rotor speed is about 1,450 rpm and the values of the rest of the measurement parameters are the same as in Case A. The results

of the analysis are shown in Fig. 13, using again a 3 level WT. The spectrum of the original signal presents a broad band spectrum, indicating a rather strong modulation. Only the spectrum of the envelope allows the identification of the fault, revealing peaks at the rotor speed  $f_s$ , the characteristic Ball Spin Frequency BSF ( $\text{BSF} = 1.871 \times \text{Rotor speed}$ ) and its second harmonic.

## 6. Conclusion

The excellent time-frequency localization capabilities of the wavelet transform, enhanced by the squaring preprocessing phase of the signal, are able to exhibit the underlying physical modulation mechanism, present in the vibration response of faulty bearings, leading to a new effective demodulation procedure. Key element for the effectiveness of this demodulation procedure is first that the choice of the specific wavelet family to perform the analysis has a marginal effect. Also, for the typical defect frequencies, resonance regions, and

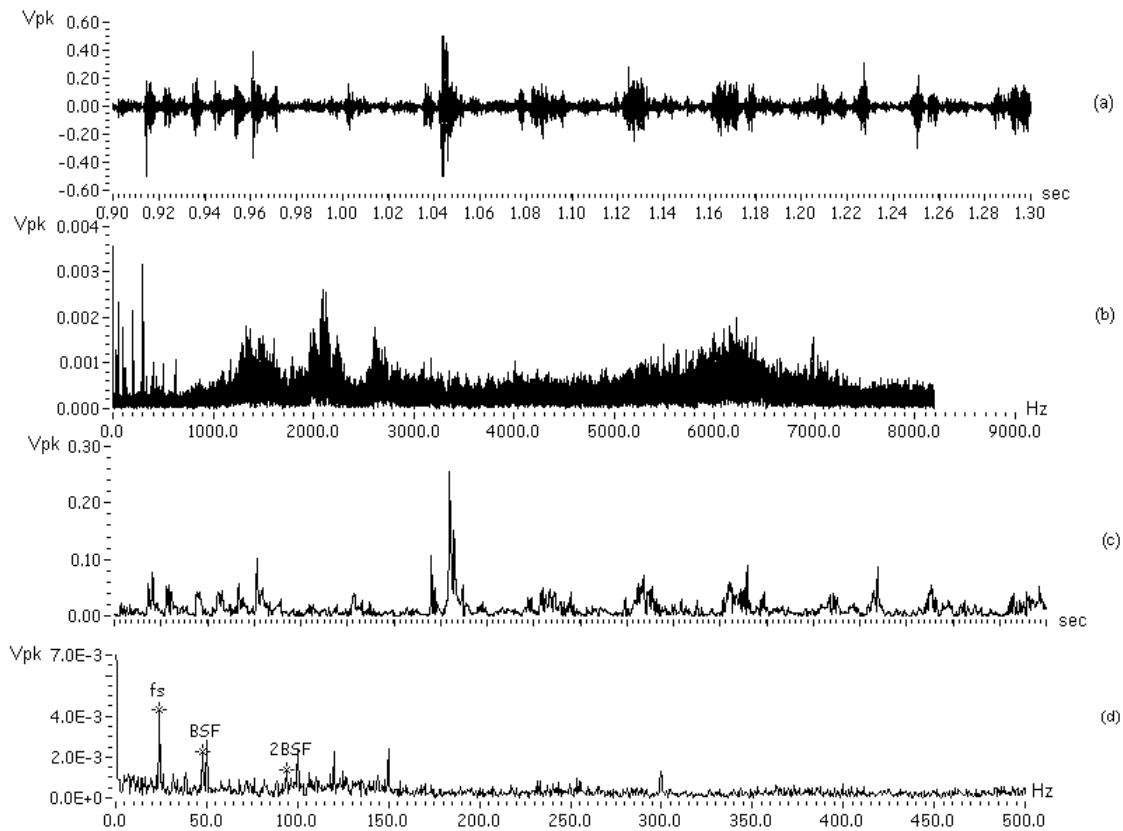


Fig. 13. Measurements and analysis results of the bearing of Case C, representing a ball fault: (a) Measured signal, (b) Spectrum of the measured signal, (c) Envelope predicted by the proposed approach, (d) Spectrum of the envelope.

sampling frequencies encountered in faulty bearing response, the number of the necessary wavelet levels can be quite limited in practice, a typical value being three levels of approximation. The above facts render the overall procedure quite simple conceptually and fast computationally, taking into account the efficiency of the DWT. The experimental results clearly confirm the effectiveness of the proposed method.

## References

- [1] D. Boulahbal, F.M. Golnaraght and F. Ismail, Amplitude and phase wavelet for the detection of cracks in geared systems, *Mechanical Systems and Signal Processing* **13**(3) (1999), 423–436.
- [2] G. Dalpiaz, A. Rivola and R. Rubini, Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears, *Mechanical Systems and Signal Processing* **14**(3) 2000, 387–412.
- [3] R.L. Eshleman, Comments on rolling element bearing analysis, *Vibrations* **13**(2) (1997), 11–17.
- [4] D. Ho and R.B. Randall, Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals, *Mechanical Systems and Signal Processing* **14**(5) (2000), 763–788.
- [5] J.C. Li and J. Ma, Wavelet decomposition of vibrations for detection of bearing-localized defects, *NDT&E International* **30**(3) (1997), 143–149.
- [6] D.E. Lyon, Thoughts on 1996 CMVA presenters' questions: 'Peakness' methods for bearing fault diagnosis, *Vibrations* **14**(3) (1998), 5–11.
- [7] S. Mallat, *A wavelet tour of signal processing*, Academic Press, CA, 1998.
- [8] P.D. McFadden and J.D. Smith, Model for the Vibration produced by a single point defect in a rolling element bearing, *Journal of Sound and Vibration* **96**(1) (1984), 69–82.
- [9] M. Misiti, Y. Misiti, G. Oppenheim and J.M. Poggi, *Wavelet toolbox*, The MathWorks Inc., 1996–1997.
- [10] K. Mori, N. Kasahima, T. Yoshioka and Y. Ueno, Prediction of spalling on a ball bearing by applying the discrete wavelet transform to vibration signals, *Wear* **195** (1996), 162–168.
- [11] R.B. Randall, *Frequency Analysis*, (3rd ed.), Brüel & Kjaer, 1987.
- [12] R. Rubini and U. Meneghetti, Use of the wavelet transform for the diagnosis of incipient faults in ball bearings, in: *Proc. of the 3rd International Conference on Acoustical and Vibratory surveillance methods and diagnostic techniques*, Senlis, France, 1998, pp. 371–378.

- [13] W.J. Staszewski, Wavelet based compression and feature selection for vibration analysis, *Journal of Sound and Vibration* **211**(5) (1998), 735–760.
- [14] N. Tandon and A. Choudhury, A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings, *Tribology International* **32** (1999), 469–480.
- [15] W.J. Wang and P.D. McFadden, Application of wavelets to gearbox vibration signals for fault detection, *Journal of Sound and Vibration* **192**(5) (1996), 927–939.
- [16] T. Yilmaz and A. Kodal, An analysis on coaxial jet flows using different decomposition techniques, *Journal of Fluids and Structures* **14** (2000), 359–373.

