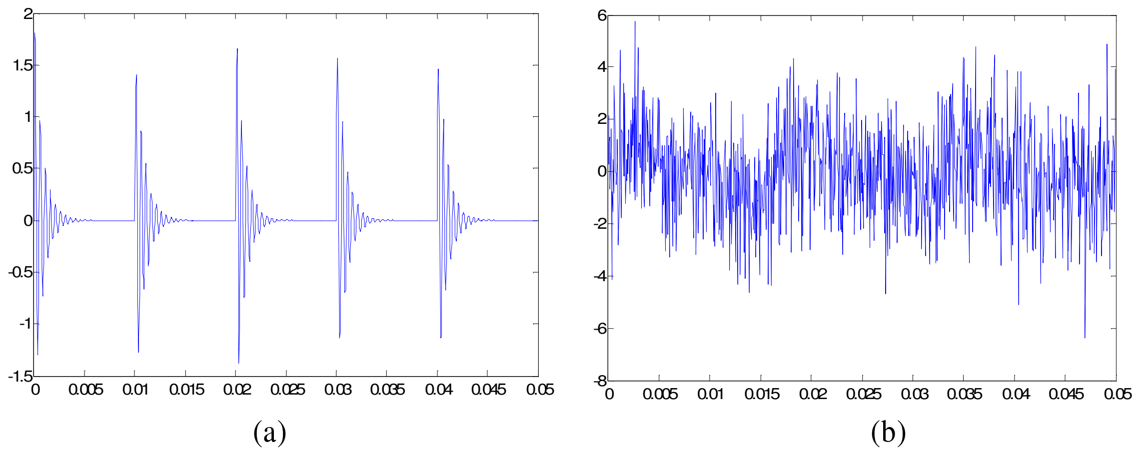
Envelope Analysis

By Jaafar Alsalaet
College of Engineering – University of Basrah 2012

1. Introduction

Envelope detection aims to identify the presence of repetitive pulses (short duration impacts) occurring due to faults in some machinery parts such as roller bearings and gears. They are often subjected to different excitations which may cause degradation over time and subsequently lead to catastrophic machinery failure if the faults are not detected early. For this reason majority of the vibratory analysis applications are concerned in the early detection of bearings anomalies. There are four main parts in a rolling bearing, the outer race, inner race, rollers and cage [1]. Consequently, there are four major frequencies associated with the bearing faults known as Bearing Characteristic Frequencies (BCF) [1]. Any defect in one (or more) of the bearing parts will excite short duration pulses when the rollers pass over the defect. These impulses will, in turn, excite the bearing natural frequencies as a consequence of impacting [1]. The repetition rate (frequency) of these impulses depends on the rotating speed, bearing geometry and the location of the defect. The duration of these impulses is very short as compared with the interval between them (period); therefore, the energy is distributed at a very low level over the period. As a result, the impulses can easily be obscured by noise or other frequency components. Fig. 1.1 demonstrates the simulated vibration signature of a ball bearing with defective outer-race producing pulse repetition rate (BCF) of 100 Hz and resonance frequency of 2800 Hz [2]. The pulses are completely obscured when the background noise and low-frequency components are added to the signal.



(a) (b)

Figure 1.1 Bearing Vibration Signature, (a) pure pulses, (b) pulses with low-frequency components and random noise

To isolate the repetitive pulses from the obscuring noise and other low-frequency components, several techniques are employed.

2. Shock Pulse Method (SPM)

This method was introduced in 1970s by SKF. It was the only bearing condition-monitoring item in SKF catalogue that time [3]. This technique offers an improved method of stethoscope (sometimes known as screwdriver method). It appeared to perform as a listening device, but reality attained a far greater complexity than what was intuitively imagined. It had borrowed from Acoustic Emissions (AE) method, and was listening in on stress waves, although the expression might not yet have been coined at that point in time. The SPM approach measures the mechanical shock speed, measuring the compression wave produced when rolling-element and race interact and with damage or failing lubrication, eventually, collide. When contact occurs, a surface reaction manifests as a compression wave and travels at the indicated speed of sound within the given material.

The SPM meter has an accelerometer of 32kHz natural frequency used to capture the **stress waves**. A fault-free bearing, that is well designed and properly selected for this application, is used to generate the “Carpet Value” or baseline relative to RPM and shaft diameter. Empiric testing established a table of such normal noise level expectations, or a quantifiable “carpet value” measured in dB. A brand new bearing entering service already has an expected Shock Pulse value measured in dB. Also, it would generate a regular and unvarying noise to our ear.

Even at that point in time, minor noises from impulses will rise above the average carpet noise. The maximum observed excursion reflects the number of shocks from all sources. A phenomenon such as the slight “catching” of an out-of-round imperfection would generate a rise of the maximum (peak value) rather than in the overall noise. Both the carpet level and maximum value develop over time. This yields not two, but rather, three values: the carpet, the peak, and the difference between them. Bearing deterioration progressing from an incipient fault to a serious one will often see [3]:

- relatively stable RMS amplitudes, eventually followed by a rise
- a contrasting sharp rise of peak amplitudes, followed by a decline
- a quick rise and then a fall of the Crest Factor.

Recent SPM meters involve not only level detection but also shock pulse spectrum for bearing defect frequency confirmation.

2. Crest Factor Development

Crest factor is the ratio of the peak value (positive or negative) in a signal to its RMS value. It ranges from 1 for square waves to 1.414 for pure sinusoidal signals and higher values for signals with pulses or short duration activities. Trending of this factor for roller bearings shows an increase of Crest factor (rise of short duration pulses with respect to RMS value) then eventual drop as the fault develops and gets broader. However, there is no vibration analysis software package currently trends Crest factor.

3. Kurtosis Factor

The kurtosis factor or forth statistical moment is given by the following equation:

$$K = \frac{\frac{1}{N}\sum_{i=0}^{N-1}(x_i - \bar{x})^4}{\left(\frac{1}{N}\sum_{i=0}^{N-1}(x_i - \bar{x})^2\right)^2}$$

High kurtosis factor indicates the presence of repeated impulses. It lends itself to the detection of whether the spectrum contains small peaks distributed along a broad frequency range or several peaks positioned at certain locations. This factor is combined with more advanced techniques later (discrete wavelet transform) to identify bearing problems.

4. Spike Energy

Spike energy (SE) measurement was proposed by IRD Mechanalysis (currently Entek IRD). This method employs the overall vibration in the frequency range 5 – 50kHz to assess the condition of rolling elements bearings. The IRD accelerometer type 970, which has a natural frequency of 27kHz, is utilized to collect vibration signal.

Currently, the SE measurement is based on a bandpass filtered signal with lower cut-off (highpass filter) set at 100Hz, 200Hz, 500Hz, 1kHz, 2kHz or 5kHz. While the upper frequency (lowpass filter) is permanently set at 65kHz [3]. The highpass filter selection depends on the machine speed and it tends to remove the low-frequency components resulting from common problem such as unbalance and misalignment which may obscure the pulses resulting from a bearing defect. Usually acceleration peak-peak value in g's (1 g = 10 m/s^2) is employed as monitoring parameter and, hence, denoted as gSE. The following table shows the recommended filter setting for gSE measurement as advised by Entek IRD:

Filter setting	Speed range (RPM)
500 Hz	0 to 100
1000 Hz	100 to 1000
2000 Hz	1000 to 1500
5000 Hz	1500 and above

Many users have observed the fluctuation of Spike Energy amplitude, when observed over time (whether using analog or modern digital meters). Unlike the quickly reacting classical parameters, gSE levels show a laborious rise time and slow decay. This feature is inherent to gSE processing. What it means: cumulative or multiple event impulse energy is required to push the gSE level up. A period of quiescence is needed to settle back down. The delay period is required to prevent the amplitude modulation in the signal from imposing fluctuations in the measured gSE value. Recently, the decay time becomes proportional to the frequency manifest within the signal. despite the fact that gSE is a good indicator for bearing fault, it can easily be compromised. Also, measurement must be trended to identify the acceptable limits.

5. High Frequency Detection

This technique was introduced by SKF as a competitive to Spike Energy and SPM. It is based on measuring the peak of the signal in the frequency range 5 to 60kHz. Unlike gSE, there is no further processing to the signal other than bandpass filtering. This technique provides reasonable performance as a condition detection system. It had been employed in some SKF products such as Microlog portable vibration analyzer.

6. Envelope Spectrum Processing (ESP)

Envelope may be defined as the outer shape of the signal. When a high-frequency signal is amplitude modulated with a low-frequency signal, the envelope would be waveform of the latter signal. Envelope detection is the process of demodulation in order to extract the low-frequency content. For bearing vibration, envelope detection is useful in identifying the intensity of the pulses and also in finding the repetition rate of these pulses. Repetition rate is related to the bearing characteristics frequencies (BPFI, BPFO, BSP and FTF) and can be found by spectrum analysis of the demodulated signal.

Enveloping has special importance in identifying bearing faults in the earlier stages. The process of extracting the envelope is shown in Fig. 6.1. The signal is bandpass filtered to remove the low-frequency content (related to common machinery problems) and very high frequency noise. Then, the signal is rectified and low-pass filtered or integrated to obtain the envelope [6].

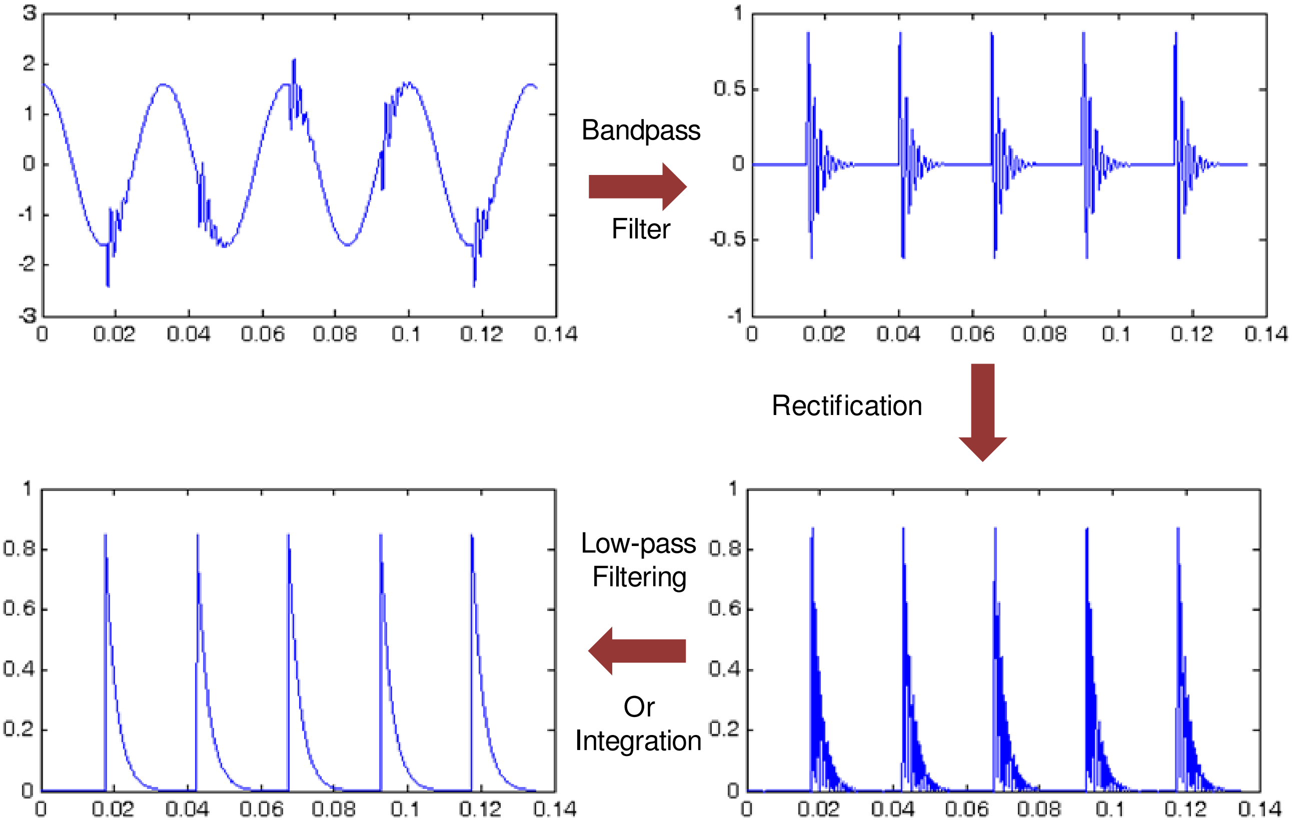


Figure 6.1 Envelope Extraction

As an alternative to the above scheme of envelope extraction, the Hilbert transform can be used to obtain more accurate envelope. Low-pass filtering and integration may smooth out the peaks of the signal and impacts the amplitude detection process. Hilbert transform has a unique feature of detecting the envelope without affecting the peaks. Envelope by Hilbert transform can be found as follows:

1. The signal is bandpass filtered as usual to remove unwanted frequencies

2. The Hilbert transform of the filtered signal is obtained as follows;

   - Take the FFT of the signal
   - Cancel out the negative spectrum (the frequency components beyond 0.5 x Number of points) and double the positive spectrum. (Digital Signal Processing, Oppenheim and Schafer)

- Take the inverse FFT to obtain the Hilbert transform. Hilbert transform is an analytical time signal with real part the same as the input signal and imaginary part as the complementary function of it.

3. Calculate the envelope $A(t)=\sqrt{x_R^2(t)+x_I^2(t)}$, where $x_R(t)$ is the real part of the Hilbert transform and $x_I(t)$ is the imaginary part of it. (Not that $x_R(t)$ is exactly the same as the input real time signal).

This scheme is used as the envelope detection scheme in HiDAC-8 portable vibration analyzer. The peak-peak value of the resulting signal represents the gSE reading when filters are properly selected. The signal may be then FFT transformed to obtain the envelope spectrum. Fig. 6.2 shows a typical envelope spectrum for a faulty ball bearing. The bearing characteristic frequency will be shown as a peak in the spectrum (50 Hz in this example). There will be multiples of this frequency due to impulsive nature of the envelope waveform as shown in the figure below.

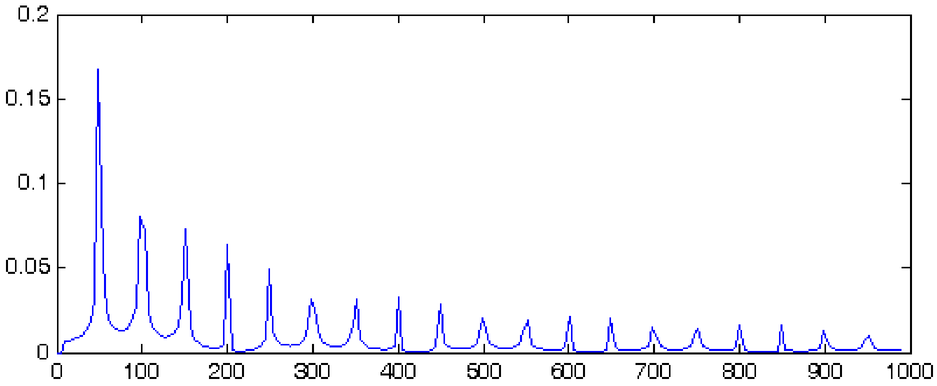
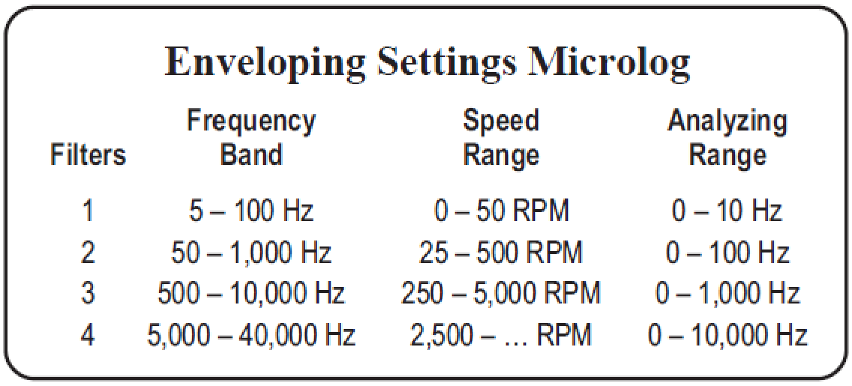


Figure 6.2 Envelope Spectrum for a Faulty Roller Bearing

The maximum frequency of the envelope spectrum (Fmax) is normally set at 1:5 of the highpass filter setting or sometimes at different ratio (1:2 for example). This is because that the bearing characteristic frequencies are far less than resonance frequencies which are smoothed out during envelope detection. It is important to select the appropriate filter setting for envelope spectrum processing such that no details are discarded. The following settings, which are provided by *SKF*, are used in some vibration analyzers and were adopted in HiDAC-8 portable machinery fault diagnostic platform as predefined settings.



Enveloping Settings Microlog

Filters	Frequency Band	Speed Range	Analyzing Range
1	5 – 100 Hz	0 – 50 RPM	0 – 10 Hz
2	50 – 1,000 Hz	25 – 500 RPM	0 – 100 Hz
3	500 – 10,000 Hz	250 – 5,000 RPM	0 – 1,000 Hz
4	5,000 – 40,000 Hz	2,500 – ... RPM	0 – 10,000 Hz

Other settings which are used in *Rockwell-Automation* EnPACK2500 data collector are given in the following table:

Bandpass filter	Machine speed (RPM)
0.6 – 1.25 kHz	0 to 500
1.25 – 2.5 kHz	250 to 1000
2.5 – 5 kHz	500 to 2500
5 – 10 kHz	1250 to 5000
10 – 20 kHz	2500 and above

Some vibration analyzers (including HiDAC-8) offer flexibility in the selection of the lower and upper cut-off frequencies to meet specific analysis requirements.

7. Cepstrum Analysis

Cepstrum analysis is referred to a range of techniques all involving functions which can be considered as *spectrum of the log spectrum*. Presumably, because it was a spectrum of a spectrum, the word *cepstrum* comes from *spectrum*, and likewise the following terms are derived in the same manner;

Quefrency from Frequency

Rahmonics from Harmonics

However, the distinctive feature of the cepstrum is not that it is a spectrum of a spectrum, but rather than the logarithmic conversion of the original spectrum. In fact, the most commonly used definition of the cepstrum nowadays is as the inverse Fourier transform of the logarithmic power spectrum [7]. The last definition of the power cepstrum may be expressed as follows

$$C_{AA}(\tau) = F^{-1}\{\log\ S_{AA}(f)\}$$

in which the two sided power spectrum, $S_{AA}(f)$, of a time signal g(t) is given by:

$$S_{AA}(f) = \overline{\left|F\{g(t)\}\right|^2}$$

Where the bar means averaging over a number of records to improve reliability.

The parameter τ in the definition is actually time, although it is referred to as quefrency. It can better be thought as a “delay time” or “periodic time” rather than absolute time. Since cepstrum can detect harmonic vibration components as single quefrency component, the application of the cepstrum analysis in vibration diagnosis is required for machines with gear mesh, anti-friction bearing and other sources of high frequency vibration containing harmonically related components.

8. Wavelet Transform

Wavelet transform is a method of converting a function (signal) from one form into another for the purpose of making the features of the signal more amenable to study. It has the ability of representing the signals in both time and frequency. Given a function *x(t)* which is square integrable, the Continuous Wavelet Transform (CWT) coefficient *W*(*a,b*) is the inner product of *x(t)* and a normalized, dilated and translated wavelet function $\psi_{a,b}(t)$ [8];

$$W(a,b) = \left\langle x(t), \psi_{a,b}(t) \right\rangle = \int_{-\infty}^{\infty} x(t)\psi^*_{a,b}(t)\,dt$$

Where the shifted and dilated wavelet function is given by: $\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$

Wavelet transform has been extensively applied in diagnostic analysis during the last 15 years. Due to its localization in both time and frequency, it has outperformed FFT analysis in many applications such as in the analysis of non-stationary and transient vibration signals. CWT is useful when high temporal resolution is required at all frequencies, such as in the analysis of gears and bearings vibration, to capture short-event pulses. Fig. 7.1 below shows a 3D wavelet scalogram for a faulty gearbox.

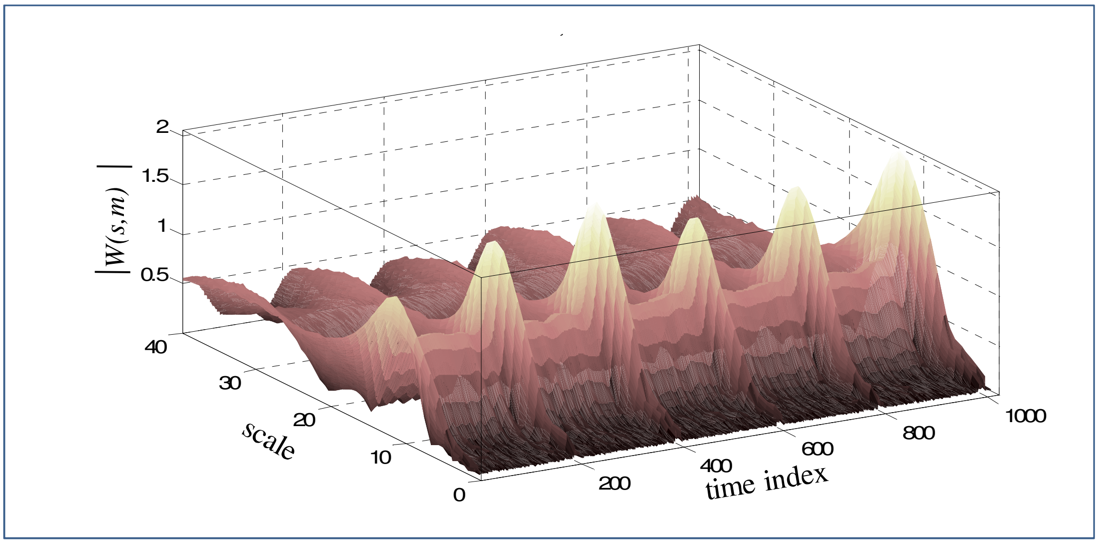


Figure 7.1 Three Dimensional Wavelet Scalogram

Application of wavelet transform in bearing signature analysis offers many advantages over the traditional analysis techniques. The non-stationary and impulsive signals can be better analyzed by using the time-frequency distribution of the wavelet transform. Moreover, the CWT can be used to reduce noise in vibration signal by smoothing and thresholding [8].

References

1. Gridhar, P. and C. Scheffer, *Practical Machinery Vibration Analysis and Predictive Maintenance*, IDC Technologies , Newnes Publications 2004, ISBN 07506 6275 1.

2. Djebala, A., Nouredine, O. and N. Hamzaoui, *Detection of Rolling Bearing Defects Using Discrete Wavelet Analysis*, Springer Science+Business Media, Meccanica, Vol. 43, 2008, pp. 339-348.

3. Frank Cagnon, *Spike Energy Diagnosis and Similar Techniques*, VibraK Consultant, 2006.

4. *Detecting Rolling Element Bearing Faults with Vibration Analysis*, Mobius Institute, online article, http://www.mobiusinstitute.com.

5. Nathan Weller, *Acceleration Enveloping-Higher Sensitivity Earlier Detection*, GE Energy, http://www.ge-mcs.com/download/orbit-archives/2001-2005/2nd_quarter_2004/2q04accelenvel.pdf.

6. R. B. Randall, *Frequency Analysis*, Brüel and Kjaer 1987.

7. Rao, R. M. and A. S. Bopardikar, *Wavelet Transform; Introduction to Theory and Applications*, Addison Wesley Longman Publication, July 1998, ISBN 0-201-63463-5.