Early detection of pitting failure in gears using a spectral kurtosis analysis

CHEMSEDDINE RAHMOUNE\textsuperscript{a} AND DJAMEL BENAZZOUZ

Solid Mechanics and System Laboratory (LMSS), University M’hamed Bougara, 35000 Boumerdes, Algeria

Received 14 May 2012, Accepted 30 June 2012

Abstract – Conventional methods (i.e. time, frequency and cepstrum) can routinely be used to reveal fault-indicating information in the vibration signal. In recent years, wavelet analysis, which can lead to a clear identification of the nature of faults, is widely used to describe rotating machine condition. The capability of this method in the detection of any abnormality can be further improved when its low-order frequency moments are considered. This paper presents the use of the fast kurtogram in the early detection and condition monitoring of pitting fault. For this purpose, a dynamic model of a simple stage gearbox (with and without defects) is used. Then, the pinion’s vibration displacement is analyzed by using a fast kurtogram method. This method is suitable for such diagnosis and gives valuable information about the presence and effects of the pitting tooth defect.

Key words: Pitting failure / gears / spectral kurtosis / faults detection

1 Introduction

Gear systems are used to transfer rotation or power from one shaft to another in desired ratios and high efficiency. These factors can be satisfactorily achieved if there is no fault in the gears. Whenever a defect occurs in a gear system (e.g. pitting, abrasive wear, bending fatigue cracks) the performance of the gears deteriorates. Transmission of motion and power, therefore, cannot be transferred as demanded. As a result, occurrence of serious defects becomes inevitable. Therefore, it is important that gear faults (especially pitting) can be detected as early as possible.

When a pitting damage is detected in a gear system, the reasons behind the initiation of this damage should be found and necessary precautions should be taken. If the pitting fault cannot be detected at its early stage of development, catastrophic failures may result in gear systems.

Over the last three decades, vibration monitoring has widely been the most used method for gearbox condition monitoring and unquestionably, it contains the most information. The signal processing methods for gearbox health monitoring can be classified into time domain analysis, frequency domain analysis and time-frequency domain analysis. Most of these works have been confined to use conventional methods such as time domain and frequency domain techniques [1–3], signal demodulation techniques [4,5] and cepstrum [6,7], which may well detect abnormality and indicate the presence of faults without giving enough information about them.

Nevertheless, combined time-frequency representation gives signal energy as a function of both time and frequency and can be performed with either constant or varying time-frequency resolution. Various time–frequency methods such as Short Time Fourier Transform (STFT) [8,9], Wigner–Ville distribution (WV) [10], Choi-Williams distribution (CWD) [11], Instantaneous Power Spectrum distribution (IPS) [12,13], Smoothed Instantaneous Power Spectrum distribution (SIPS) [13] and Continuous Wavelet Transform (CWT) [14–19], have been used extensively to analyze vibration signals and extract useful diagnostic information.

Of these methods, the STFT and CWT perform a linear decomposition of the analyzed signal, and therefore they do not cause any interference. The STFT employs a constant window size during the analysis and, therefore leads to a constant time–frequency resolution. However, the CWT performs a decomposition of the analyzed signal into a set of waves (or wavelets), which are derived from a single wavelet, and wavelets at different frequencies are generated by introducing dilation into the analyzing wavelet. A large window is used for low frequency estimates with poor time resolution, whereas the window automatically narrows at high frequencies, improving time resolution of the transform, but the frequency resolution

\textsuperscript{a} Corresponding author: rahmoune.ch@gmail.com
deteriorates according to the uncertainty principle [20]. Therefore, the wavelet transform provides a good compromise between localization and frequency resolution.

From a failure detection point of view, the analysis generally requires a comparison of time–frequency maps representing good and faulty conditions for the whole $t–f$ plane, which is a tedious work due to the increased dimensionality.

Moreover, revelations of fault symptoms critically depend on the severity of the damage and indications in the two-dimensional time–frequency map may not be discernible at the early stages of fault development. However, the spectral kurtosis, which is one of the high-order spectral cumulants, is an effective tool for characterizing the dynamic behavior of the gearbox system.

This paper presents the use of fast kurtogram in the early detection and advancement monitoring of gear tooth pitting fault. For this purpose, a dynamic model of one stage gear box system (with and without defects) is proposed. The fast kurtogram method is used to characterize the signature of a pitting tooth defect. This method is suitable for such diagnosis and gives valuable information about the presence and effects of tooth pitting defect.

2 Theoretical background: spectral kurtosis and kurtogram

Kurtogram was first introduced by Antoni and Randall [21], which comes from spectral kurtosis (SK) [21]. The SK method selects kurtosis as a measure of distance between an arbitrary random process and a Gaussian one with the aim of detecting the existence of transients in a signal. The SK value of the signal is obtained by calculating the kurtosis of each frequency component contained in the signal. SK finally represents the transient characteristics of the signal as a function of frequency. SK can indicate not only transient components in the signal but also their locations in the frequency domain, and therefore overcome the disadvantages of the power spectral density in detecting and characterizing transients signal.

Considering the Wold-Cramer decomposition of a non stationary signal, we can define a signal $y(t)$ as the response of a system with time varying impulse response $h(t, s)$ excited by a signal $x(t)$. Then can be expressed as [23]

$$y(t) = \int_{-\infty}^{+\infty} e^{2\pi j f t} H(t, f) dX(f)$$

where $H(t, f)$ is the time varying transfer function of the considered system and can be interpreted as the complex envelope of signal $y(t)$ at frequency $f$ and $dX(f)$ is the spectral process associated with $x(t)$.

When $H$ is time stationary and independent of process $x$, we obtain a conditionally non stationary process (CNS). SK is based on the fourth order spectral cumulant of a CNS process.

$$C_{\Delta y}(f) = S_{\Delta y}(f) - 2S_{2y}(f)$$

and

$$S_{ny}(f) = (|F_y(\tau, f)|^4)$$

$F_y(\tau, f)$ is estimated using the short time fourier transform:

$$F_y(\tau, f) = \int_{-\infty}^{+\infty} y(t) u^*(t - \tau) e^{-j2\pi ft} dt$$

$$= \langle y(t), w(t - \tau) e^{-j2\pi ft} \rangle$$

where $u^*(t - \tau)$ represents the window function and superscript $*$ denotes the conjugation.

The STFT is actually an inner product operation $(y(t), u(t - \tau)e^{-j2\pi ft})$ of signal $y(t)$ and filter $w(t - \tau)e^{-j2\pi ft}$, which is the triangle basis function $e^{-j2\pi ft}$ enveloped by window function $w(t - \tau)$.

Then, the SK is generated by normalizing the fourth order cumulant given in Equation (2). It is a measure of the peakness of the probability density function

$$k_y(f) = \frac{C_{\Delta y}(f)}{S_{2y}(f)} = \frac{S_{\Delta y}(f)}{S_{2y}(f)} - 2$$

The SK estimation values for particular frequency $f$ is implemented by calculating the STFT of signal $y(t)$ given by Equation (4). It is noted that, the selected window length and the percent of overlap used for the STFT calculation significantly influence the value of SK. Therefore, for the purpose of selecting the optimal window length the spectral kurtosis is calculated for several different window lengths. The implementation result of these calculations is the so-called kurtogram. The window parameters for which the spectral kurtosis has its maximum value determine the band pass filter parameters. Therefore, kurtogram is a 2D map and presents values of SK calculated for various parameters of frequency $f$ and bandwidth $Bw$.

A more detailed explanation of kurtogram can be found in references [21,23,24].

3 One stage spur gear model

Figure 1 shows the proposed dynamic model of one stage gear transmission. This is a two parameters model, involving stiffness and damping, with torsional and lateral vibration, and has 6 degrees of freedom: the motor rotational angle: $\theta_m$, pinion and wheel rotational angles respectively: $\theta_1$ and $\theta_2$, the load rotational angle: $\theta_c$, radial displacement of flexible shafts: $y_1$ and $y_2$ respectively. It is modeled as a pair of rigid disks connected by a spring damper set along the line of contact (for simplicity inter-tooth friction is ignored) and mounted on flexible shafts; the shafts are supported by rolling element bearing. Because inter-tooth friction is ignored, the vibration in the $x$ direction is a free response and will disappear due to inherent damping. Differential equations can be obtained.
The duration of contact depends on the contact ratio varying from one and two pairs of teeth in contact [25].

As a result, the gear mesh stiffness will fluctuate around a mean value $k_m$ (Fig. 2).

Thus, gear mesh stiffness is periodic with the period $T_{eng} = 60/N_1Z_1$ and can be approximated as $[7, 25]$:

$$k_t(t) = \begin{cases} 
  k_{\text{max}} & 0 \leq t \leq (\varepsilon - 1)T_{eng} \\
  k_{\text{min}} & (\varepsilon - 1)T_{eng} \leq t \leq T_{eng} 
\end{cases}$$

where $N_1$ is the pinion rotational speed in rpm, and $Z_1$ is the tooth numbers.

Fourier development of $k_t(t)$ yields:

$$k_t(t) = k_m + \frac{\Delta k}{\pi} \sum_{i=1}^{\infty} \frac{1}{i} \left[ \sin(2i\pi(\varepsilon - 1)) \cos \frac{2i\pi t}{T_{eng}} + (1 - \cos(2i\pi(\varepsilon - 1))) \sin \frac{2i\pi t}{T_{eng}} \right]$$

with:

$$k_m = k_{\text{max}}(\varepsilon - 1) + (2 - \varepsilon)k_{\text{min}}$$

$$\Delta k = k_{\text{max}} - k_{\text{min}}$$

5 Pitting defects modeling

It has been established that gear tooth failure will induce amplitude and phase changes in vibration, which in turn can be represented by magnitude and phase changes in gear mesh stiffness [26].

Undamaged configuration of the mesh stiffness corresponds to 0% phase change and 0% amplitude reduction. Amplitude and phase changes are applied on gear mesh stiffness from these reference values to simulate surface pitting.

An amplitude modulation of the gear mesh signal is expected from these defect-induced changes (Fig. 3). In fact, the new gear mesh signal, $k_t(t)$, resulting from the defect modeling can be expressed by:

$$k_d(t) = k_t(t)(1 - d(t))$$

where $d(t)$ is the modulating function.

The tooth failure induced variations in gear mesh stiffness used for the simulations are given in Figure 4. From this figure, it can be observed that as the size of the pitting surface grows, the total mesh stiffness when the pitting tooth is in meshing becomes much lower.

6 Numerical simulations and results

For simulation, it is necessary to write the model of the mechanical system in state space representation. In this order, each second-order differential equation is written in the form of two first order differential equations. Thus, 12 nonlinear first order differential equations with

\begin{align*}
I_m\ddot{\theta}_m &= T_m - k_{\text{c}1}(\theta_m - \theta_1) - c_{\text{c}1}(\dot{\theta}_m - \dot{\theta}_1) \\
I_1\ddot{\theta}_1 &= k_{\text{c}1}(\theta_m - \theta_1) + c_{\text{c}1}(\dot{\theta}_m - \dot{\theta}_1) - R_31k_1 \\
&\quad \times (R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
&\quad - R_31c_1(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
I_2\ddot{\theta}_2 &= -k_{\text{c}2}(\theta_2 - \theta_r) - c_{\text{c}2}(\dot{\theta}_2 - \dot{\theta}_r) + R_32k_2 \\
&\quad \times (R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
&\quad + R_32c_1(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
I_r\ddot{\theta}_r &= -T_r + k_{\text{c}2}(\theta_2 - \theta_r) + c_{\text{c}2}(\dot{\theta}_2 - \dot{\theta}_r) \\
m_1\ddot{y}_1 &= k_1(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
&\quad + c_1(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) - k_1\dot{y}_1 - c_1\ddot{y}_1 \\
m_2\ddot{y}_2 &= k_2(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) - k_2\ddot{y}_2 - c_2\ddot{y}_2 \\
\end{align*}

where, the vertical radial stiffness of the input and the output bearings are constant and equal, i.e., $k_1 = k_2 = k_r$; also the damping coefficients of the input and the output bearings are constant and equal, i.e., $c_1 = c_2 = c_r$; the torsional stiffness values of the input and the output flexible coupling are constant and equal, i.e., $k_{\text{c}1} = k_{\text{c}2} = k_{\text{c}}$; the damping coefficients of the input and the output flexible coupling are constant and equal, i.e., $c_{\text{c}1} = c_{\text{c}2} = c_{\text{c}}$; the mesh damping coefficient, $c_t$, is assumed to be proportional to the total mesh stiffness, $k_t$: $c_t = \mu k_t$, where $\mu$ is set to be constant.

4 Gear mesh stiffness evolution

The teeth in a healthy gear in good running condition will deflect under load. The meshing process is always varying from one and two pairs of teeth in contact [25]. The duration of contact depends on the contact ratio $\varepsilon$. From Figure 1 as follows:

\begin{align*}
I_m\ddot{\theta}_m &= T_m - k_{\text{c}1}(\theta_m - \theta_1) - c_{\text{c}1}(\dot{\theta}_m - \dot{\theta}_1) \\
I_1\ddot{\theta}_1 &= k_{\text{c}1}(\theta_m - \theta_1) + c_{\text{c}1}(\dot{\theta}_m - \dot{\theta}_1) - R_31k_1 \\
&\quad \times (R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
&\quad - R_31c_1(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
I_2\ddot{\theta}_2 &= -k_{\text{c}2}(\theta_2 - \theta_r) - c_{\text{c}2}(\dot{\theta}_2 - \dot{\theta}_r) + R_32k_2 \\
&\quad \times (R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
&\quad + R_32c_1(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
I_r\ddot{\theta}_r &= -T_r + k_{\text{c}2}(\theta_2 - \theta_r) + c_{\text{c}2}(\dot{\theta}_2 - \dot{\theta}_r) \\
m_1\ddot{y}_1 &= k_1(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) \\
&\quad + c_1(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) - k_1\dot{y}_1 - c_1\ddot{y}_1 \\
m_2\ddot{y}_2 &= k_2(R_31\dot{\theta}_1 - R_32\dot{\theta}_2 - y_1\sin\alpha + y_2\sin\alpha) - k_2\ddot{y}_2 - c_2\ddot{y}_2 \\
\end{align*}
Fig. 2. Time variation of stiffness gear mesh $k(t)$ of the pair of gears when the gear teeth are perfect (that is, have no pitting).

Fig. 3. Modulating function.

Fig. 4. Tooth failure induced variations in gear mesh stiffness with 30% phase change.
time varying coefficients are obtained. These equations are written such that each equation contains the time derivative of only one variable.

Solving the new system of equations, the vibration response, that is, the pinion’s vibration displacement as a function of time can be obtained. The vibration displacement response for the healthy gearbox, with 0% pitting, is shown in Figure 5.

First, a tooth pitting defect on a tooth of wheel type 1 is modeled by 30% phase change and 2% amplitude reduction of gear mesh stiffness operated in modulating function $d(t)$ (Fig. 3). The pinion’s vibration displacement in presence of tooth pitting defect is shown in Figure 6. It can be observed that it is not possible to detect a gear tooth pitting from the direct analysis of waveform of the pinion’s vibration displacement.
The presence of a tooth pitting defect causes repetitive-like transient in pinion’s vibration displacement because of phase change and amplitude reduction of the mesh stiffness when the failure tooth comes into contact. These transients present two important features which are:

1. Wide frequency.
2. Small energy: the energy created by the defect is very small, and typically occupies less than one thousandth of the total signal energy.

Because the signals generated by a defective gear have the characteristics mentioned above, it is difficult to recognize their faults through simple conventional methods.

To overcome this problem, we propose to use the Fast Kurtogram. It is a technique well suited to the analysis of transient in signal. The paradigm relies on the assertion that each type of transient is associated with an optimal (frequency/frequency resolution) dyad \( \{f; Bw\} \) which maximizes its kurtosis, and hence its detection. For our application, the method is based on the following steps:

1. Calculate the fast kurtogram of the vibration signal: The abnormally high value of the spectral kurtosis in the kurtogram indicates abnormal transients occurring in a signal and indicates in which frequency band(s) these abnormal transients occur,
2. Design the band pass filter based on the results of the kurtogram: i.e. determine the dyad (frequency \( f \)/frequency resolution \( Bw \)) with abnormally high value of the spectral kurtosis,
3. Analyze the filtered signal by using the fast Fourier transform.

Since the pitting tooth should generate impulses at one of the characteristic frequencies of the gear, analysis will be focused on looking for such periodic excitations in the signal.

Figure 7 shows the kurtogram of pinion’s vibration displacement for a healthy gear; we can observe a maximum value of the spectral kurtosis \( K_{\text{max}} = 0.6 \) at dyad \( (170.6641 \text{ Hz}/341.3281 \text{ Hz}) \).

In the presence of a tooth pitting defect on the wheel type 1 (2% reduction in the wheel mesh stiffness and 30% phase change), the kurtogram (Fig. 8) has an abnormally high value of the spectral kurtosis \( K_{\text{max}} = 1.4 \) at dyad \( (2943.9551 \text{ Hz}/255.966 \text{ Hz}) \). It clearly reveals the presence of sharp abnormal shocks in the signal. The corresponding envelope magnitude is shown in Figure 9a and its envelope spectrum in Figure 9b. Note that the envelope spectrum immediately indicates a fault.
Fig. 8. Fast kurtogram of the signal shown in Figure 6. The maximum is reached at dyad (2943.9551 Hz/255.966 Hz).

Fig. 9. (a) Magnitude of the envelope which maximizes the kurtogram, together with its 0.1% significance threshold. (b) Envelope spectrum as provided by the Fourier transforms.
frequency component related to the rotational tooth frequency \( f_1 = 43.83 \text{ Hz} \) and its harmonics.

To study the relation between % amplitude reductions in wheel mesh stiffness and the maximum value of spectral kurtosis and its location in the kurtogram, we simulate the amplitude reduction from 3% to 6%, and gain the graphics of kurtogram. Kurtograms for each case are shown in Figure 10.

From Figure 10, we can observe that the maximum value of spectral kurtosis increases with aggravation of the default, but stays at the same location in the kurtogram i.e. at dyad \((2943.9551 \text{ Hz} / 255.966 \text{ Hz})\).

Figure 11 shows the amplitude evolution of the characteristics frequencies components as function of the default evolution. We notice that the amplitude of the characteristic frequency component increases with the defect increase.

### 7 Conclusion

In this paper, a dynamic model of simple stage gear box has been developed to examine the pinion’s vibration displacement in the presence of defect such as tooth pitting. This defect is modeled by a phase shift and an amplitude fall of the gear mesh stiffness. Simulation results have shown that it is not possible to detect a gear tooth pitting from the direct analysis of waveform of the pinion’s vibration displacement. However, we showed that, the use of the fast kurtogram allows us to detect tooth pitting fault. In fact the fast-kurtogram method determines the dyad \((f, B_w)\) that maximizes the spectral kurtosis. Filtering in the corresponding band frequency is used to obtain the envelope. Comparing between the obtained envelope spectrum for a healthy gear and that obtained for a faulty gear allows us to conclude the presence of a tooth crack defect, which will result in the manifestation of
new frequency components related to the rotational tooth frequency and its harmonics. On the other hand the amplitude of these harmonics increases with respect to the increase defect level. Such observation can be used to determine the threshold that allows to decide when to stop the production line before that the gravity of failure engenders serious or dramatic problems.

References


Fig. 11. Amplitude evolution of the characteristic frequency component based on the evolution of the defect.