A METHODOLOGY FOR THE DETECTION AND DIAGNOSTIC OF LOCALIZED FAULTS IN GEARS AND ROLLING BEARINGS SYSTEMS

METODOLOGÍA PARA LA DETECCIÓN Y DIAGNÓSTICO DE FALLAS LOCALIZADAS EN SISTEMAS DE ENGRANAJES Y RODAMIENTOS

César San Martín¹ Edgar Estupiñán² Daniel San Martín³

Recibido 10 de octubre de 2008, aceptado 18 de noviembre de 2009

Received: October 10, 2008 Accepted: November 18, 2009

RESUMEN

En este trabajo se presenta una metodología para detectar fallas incipientes en máquinas rotatorias. La metodología está basada en el análisis de cicloestacionariedad, la cual está presente en las señales de vibración generadas por máquinas rotatorias. De particular interés son las componentes cicloestacionarias de segundo orden y de órdenes superiores, puesto que contienen información relevante, que puede ser usada para detección temprana de fallas en rodamientos y transmisiones de engranajes. La primera etapa de la metodología consiste en la separación de las componentes de primer orden de la señal de vibración, para posteriormente centrar el análisis en la señal residual, la cual contiene las componentes cicloestacionarias de órdenes superiores. Luego, la señal residual es digitalmente filtrada y demodulada, considerando el rango de frecuencia de mayor importancia. Finalmente, la señal residual demodulada es autocorrelacionada, obteniendo una señal donde las componentes espectrales generadas por la presencia de una posible falla localizada pueden ser efectivamente detectadas. La metodología es validada analizando mediciones experimentales de vibraciones para dos casos particulares. El primero es la detección de una grieta en uno de los dientes de un sistema de transmisión y, el segundo, la detección de una picadura en la pista interna de un sistema de rodamientos. Los resultados muestran que el método propuesto para el monitoreo de condición de máquinas rotatorias es una herramienta útil en las tareas de diagnóstico de fallas, el cual complementa los análisis con técnicas de diagnóstico tradicionales.

Palabras clave: Análisis de cicloestacionariedad, diagnóstico de fallas, análisis de vibraciones, monitoreo de condición.

ABSTRACT

In this work, an effective methodology to detect early stage faults in rotating machinery is proposed. The methodology is based on the analysis of cyclostationarity, which is inherent to the vibration signals generated by rotating machines. Of a particularly interest are the second and higher orders cyclostationary components since they contain valuable information, which can be used for the early detection of faults in rolling bearings and gear systems. The first step of the methodology consists in the separation of the first-order periodicity components from the raw signal, in order to focus the analysis in the residual part of the signal, which contains the second and higher order periodicities. Then, the residual signal is filtered and demodulated, using the frequency range of highest importance. Finally, the demodulated residual signal is auto-correlated, obtaining an enhanced signal that may contain clear spectral components related to the presence of a prospective localized fault. The methodology is validated analyzing experimental vibration data for two different cases. The first case is related to the detection of a crack in one of the teeth of a gearbox system and the second case is related to the detection of a pitfall in the inner race of a rolling bearing. The results show that the proposed method for the condition monitoring of rotating machines is a useful tool for the tasks of fault diagnosis, which can complement the analysis made using traditional diagnostic techniques.

Keywords: Cyclostationary analysis, fault diagnosis, vibration analysis, condition monitoring.

¹ Departamento de Ingeniería Eléctrica. Universidad de La Frontera. Casilla 54-D. Temuco, Chile. E-mail: csmarti@ufro.cl

² Escuela Universitaria de Ingeniería Mecánica. Universidad de Tarapacá. Casilla 6-D. Arica, Chile. E-mail: eestupin@uta.cl

³ Laboratorio de Vibraciones Mecánicas. Universidad Técnica Federico Santa María. Sede Concepción. Casilla 457. Concepción, Chile.

INTRODUCTION

Vibration signals generated by rotating machines may be considered as non-stationary processes that present periodic (i.e. cyclic) variations in the time domain in some of its statistics [1], which is a main characteristic of those type of signals named cyclostationary signals. A vibratory signal $x(t)$ is said to be n th order cyclostationary with period T if its n th order moment exist and is periodic in the time domain with the period T . Typical examples of first order periodicity (FOP) vibration signals are generated by rotating machines with misaligned couplings and/or unbalanced rotors, whereas, modulated vibratory signals generated by wear mechanisms, friction and impact forces are some examples of second order periodicity (SOP) processes. In order to analyze FOP signals and to extract the required information for the fault detection tasks, the classical spectral analysis is an adequate a practical tool that may be used for most of these cases. However, when SOP signals have to be analyzed (e.g. signals with amplitude and/or frequency modulations), the analysis should be carried out using more sophisticated tools, in order to be able to identify variations in the statistics of the signals, containing meaningful information of the system under analysis [2]. In some cases, demodulation techniques may be satisfactorily used to analyze SOP signals, as long as, either the resonant zones or the main frequency ranges of the expected faults can be known in advance. However, the efficacy of the demodulation techniques diminishes when the signal contains higher orders of cyclostationarity, as well as, random noise components. The basic idea behind the theory of cyclostationary analysis is to apply an appropriate quadratic transformation to a SOP signal in order to obtain a modified signal of FOP [3]. Then, the modified signal can be analyzed with traditional diagnostic techniques applied to the mechanical components under study.

In this framework, the methodology proposed in this study, incorporates time-frequency analysis of SOP signals in combination with traditional techniques typically used in fault detection (e.g. spectral analysis, enveloping analysis, etc). First, the components of FOP are reduced using an adaptive filtering method. In this way, a residual signal containing the SOP components is also obtained. Then, the residual signal is filtered in order to highlight the SOP components of the signal. In this stage, the cutting frequencies of the filter are estimated by using a time-frequency transformation based on the cyclic autocorrelation function. Finally, the filtered residual signal is demodulated and auto-correlated, obtaining a resultant signal with useful information for fault diagnostic purposes. In summary, from the time-frequency analysis, appropriate filters are configured, and then, using an enveloping detector

the residual signal is demodulated, and finally, in order to improve the signal to noise ratio (SNR), a matched filter based in the autocorrelation function is used.

Figure 1. a) Picture of an induced fault (of 10mm length) on the tooth surface of the pinion of a single-stage spur gear transmission. b) Picture of an induced localized fault in the inner race of a radial ball bearing.

This procedure has been tested with two cases using experimental vibration data from two test rigs used to simulate faults in gears and rolling bearings respectively (Figure 1). The results show that the proposed methodology is an effective tool for the early detection and diagnostic of faults in rotating machinery.

This work is organized as follows: firstly, the principles of cyclostationarity and the basics of the proposed method is presented and validated using experimental vibration data of a faulty gearbox and a faulty rolling bearing; finally, the main conclusions are drawn.

BASICS OF CYCLOSTATIONARITY AND PROPOSED METHOD

A well detailed tutorial on the principles of cyclostationarity, focused on mechanical applications is given in [4]. However for the completeness of this work, and to address the use of cyclostationarity towards the proposed method, the basics of cyclostationarity are included here. A non-stationary signal can be considered as cyclostationary with FOP and SOP components, only if its moments of first and second order are periodic, in other words, if the moments satisfy the equations (1) and (2) [5]:

$$E[x(t_1)] = E[x(t_1 + T)], \quad (1)$$

$$E[x(t_1)x(t_2)] = E[x(t_1 + T)x(t_2 + T)], \quad (2)$$

where, E is the expected operator and T is the period or cycle of the signal $x(t)$. An auto-correlation function with variation in time can be associated to the signal $x(t)$, which is given by:

$$r_x(t, \tau) = E \left[x(t + \beta\tau) x(t - \beta\tau) \right], \quad (3)$$

where, τ is the time lag and β satisfies: $\beta + \bar{\beta} = 1$. The function of equation (3) is also known as the instantaneous auto-correlation function (ACF).

In general, it is possible to assume that a vibration signal $x(t)$ is composed of FOP, SOP and random noise, as shown in equation (4).

$$x(t) = x_{FOP}(t) + x_{SOP}(t) + n(t). \quad (4)$$

Considering that the focus of the analysis is on the SOP components of the signal $x(t)$, the first stage of the procedure consists of using an LMS (least mean square) adaptive filter [6], in order to reduce the FOP components from the signal to be analyzed. In this way, the FOP components are separated from the raw signal and a residual signal (i.e. error signal) containing the SOP components and random noise is obtained. If a typical vibration signal containing amplitude modulations (i.e. SOP components), is assumed, the residual signal can be expressed as in equation (5).

$$x_{SOP}(t) = \left(b + \sum_{i=1}^N \cos 2\pi f_{c_i} t + n(t) \right) \cos 2\pi f_0 t, \quad (5)$$

where, $i = 1, 2, \dots, N$ is the number of modulation signals, f_{c_i} and f_0 are the modulating and modulated signal respectively, b is a constant and $n(t)$ is white noise with unknown variance. Since the main interest here, is to extract the information of the modulating signals from the signal that includes the SOP components (x_{SOP}), a simple demodulator (i.e. a low pass filter with cut frequency f_0) can be used. In order to obtain a good estimation of the cut frequency f_0 , it is used a time-frequency distribution of the Cohen type [7], which is given by equation (6):

$$C(t, f) = \iint \Phi(u - t, \theta - f) r_x(u, \tau) e^{-j2\pi\tau\theta} d\tau d\theta, \quad (6)$$

where, Φ is an arbitrary function (kernel) and r_x corresponds to the instantaneous ACF given by equation (3). The type of time-frequency distribution is determined by the selected function Φ . For instance, if Φ is equal to 1, the Wigner-Ville distribution is obtained, which is given by (7), whereas, if Φ is a cubic function type, that helps to reduce the frequency cross terms, the Zhao-Atlas-Marks (ZAM) distribution is obtained [8], which is given by (8):

$$WVD(t, f) = \int r_x(t, \tau) e^{-j2\pi\tau f} d\tau, \quad (7)$$

$$Z(t, f) = \int \phi(\tau) r_x(t, \tau) e^{-j2\pi\tau f} d\tau. \quad (8)$$

Finally, in order to enhance the SOP signal, the auto correlation function of the filtered signal is computed.

Table 1. Main characteristic frequencies of the gearbox.

Component	Frequency [Hz]
Motor (rotational speed)	17
Main mesh frequency	289
i th harmonic of mesh frequency	$i \times 289$
Pinion (rotational speed)	17
Gear (rotational speed)	10.32

The auto correlation function of a signal $x_c(t)$ is given by:

$$E \left[\{x_c(t) + n(t)\} \{x_c(t - \tau) + n(t - \tau)\} \right] = E \left[x_c(t) x_c(t - \tau) \right] \quad (9)$$

In summary, the main steps of the proposed method are listed below:

- *LMS adaptive filtering*: to separate the FOP components from the measurement vibration signal, obtaining a residual signal with the SOP components.
- *Time frequency transformation*: the estimation of f_0 (required for the further digital filtering), is done by using the ZAM distribution.
- *Digital filtering*: the residual signal is filtered using the cutting frequencies identified in the previous stage.
- *Noise reduction*: the filtered residual signal and specially the SOP components are enhanced by using the autocorrelation function (matched filter).
- *Detection and diagnosis of faults*: the spectrum of the enhanced residual signal is analyzed, looking for spectral components that might be related to the presence of a fault.

EXPERIMENTAL VALIDATION

In this section, the proposed method is validated by analyzing experimental vibration data from two different cases (see Figures 2 and 3). The first case corresponds to the detection of a fault in a one-stage gearbox, and the second case corresponds to the detection of a localized pitfall in a rolling bearing.

Case 1: A faulty one-stage gearbox

In this case, the experimental data is taken from a test rig, which consists of an asynchronous electrical motor controlled by a frequency converter and coupled to a single-stage spur gear transmission through a flexible coupling. The pinion has 17 teeth and the wheel has 28 teeth. The system is under a constant load, which is supplied by a

DC generator, as it is illustrated in the sketch of Figure 4. The rotational frequencies and mesh frequencies of the gearbox are listed in the Table 1.

When a local fault of the cracked tooth type occurs in one of the gears of the system, it is expected to have a vibration signature containing amplitude modulations of the fundamental gear mesh frequency and its harmonics with a modulating frequency equal to the rotational frequency of the faulty gear [9]. Therefore, in this particular case, if spectral components at a frequency of 17 Hz and its multiples are identified in the spectrum of the vibration signal, it can be associated to a fault in the pinion. In contrast, if the spectral components are at the frequency of 10.32 Hz and its multiples, the fault can be associated to the gear. The following analysis is done for vibration data taken from the test rig with a faulty pinion.

The vibration data were acquired from two piezoelectric accelerometers mounted on the supports A and B, shown in Figure 4, and using a data acquisition system configured with a sampling frequency of 30 kHz. The waveform time and the frequency spectrum of the acquired vibration signal are shown in Figures 2a and 2b respectively. The main mesh frequency and some of its harmonics can be identified from the spectrum of Figure 2b.

Figure 2. Waveform time and spectrum of the raw vibration signal-case I.

Figure 3. Waveform time and spectrum of the raw vibration signal-case II.

Figure 4. Sketch of the test rig of case I-a-1-stage gearbox with fault in the pinion.

In order to separate the FOP components, a LMS adaptive filter with 500 coefficients and a learning rate of 0.01ms was used. The waveform time and the spectrum of filtered signal, which contains the FOP cyclostationary components, are shown in Figure 5. In the same manner, the waveform time and the spectrum of the residual signal, which contains the SOP cyclostationary components and signal noise, are shown in Figure 6. From the spectra of Figures 5 and 6, it can be observed that the FOP components are predominated, when they are compared to the other components, which is generally expected [1].

Analyzing the spectrum of Figure 6, two possible resonant zones of the system can be identified, with the range frequencies between 1.200 to 2.600 Hz and between 2.800 to 5.800 Hz approximately. To complement the analysis, and before of filtering the signal containing the SOP components, the ZAM transform was applied to the residual signal, obtaining the time-frequency distribution

shown in Figure 7. Despite the frequency resolution is a bit coarse ($\Delta f \approx 208$ Hz), it is enough to visualize the variation in time of the two main resonant zones. It can be observed, that the resonant zones are excited approximately every 0.003 s, which corresponds to a close value of the fundamental mesh frequency (289 Hz). In Figure 7, can be seen that the impulsive variations are clearer defined in the second frequency range (2800 - 5800 Hz), therefore, these frequencies are selected as the cutting frequencies of the further filtering stage of the residual signal.

In order to filter the residual signal, a finite impulse response (FIR) filter was used. The implemented bandwidth filter has 400 coefficients with the low cut frequency of 2500 Hz and the high cut frequency of 6500 Hz. The waveform time and the spectrum of the filtered signal are shown in Figure 8. This signal has the typical pattern of an amplitude-modulated signal found in a mechanical system. Therefore, if a demodulation process is applied to the filtered signal, the results are shown in Figure 9.

Figure 5. Waveform time and spectrum of the filtered signal (FOP components) - case I.

The demodulation technique applied is as follows: first, a high pass filter with cut frequency of 2800 Hz is applied; second, the signal is rectified and the mean value of the signal is subtracted and third, a low pass filter with cut frequency of 200 Hz is applied. Both of the filters used for the demodulation are of the infinitive impulse response (IIR), and with 5 coefficients. From Figure 9, a fault in the pinion can be confirmed since clear spectral

components at 17 Hz and its first harmonics are presented in the spectrum. Additionally, in order to enhance the main components of interest in the residual signal the auto-correlation function can be used and the result is shown in Figure 10. This last step in the methodology could be avoided, as in this case, where the spectral components were already identified with the filtered and demodulated residual signal (Figure 11), however, in other cases where the vibration signals contain higher noisy components, the use of the auto-correlation function is very useful to clean the signal and therefore, should be include it in the analysis.

Case II: A faulty rolling bearing

In this case, the method is applied to experimental vibration data taken from a test rig with an incipient localized fault in the inner race of one of the bearings that support the shaft. The test rig consists of an asynchronous electrical motor controlled by a frequency converter, which drives a rotor shaft supported by two radial ball bearings. A schematic drawing of the test rig is shown in Figure 11. A static load can be indirectly applied to the bearings by using a tensor pulley system mounted in the centre of the shaft. The vibration data were acquired using piezoelectric accelerometers mounted on the supports A and B, and using a data acquisition system with a sampling frequency of 30 kHz. The vibration data analyzed for this case is for a faulty bearing located at the motor side (bearing A). The main fault bearing frequencies for the bearing under study are listed in the Table 2.

Figure 6. Waveform time and spectrum of the residual signal (SOP components and noise)-case I.

Figure 7. ZAM time-frequency distribution of the residual signal-case I.

It has been shown by several studies that similarly to the case of faulty gears, localized faults in bearings generate spectral sidebands around the resonant frequencies, which are related to the source frequency of the fault [10]. However, when a fault is located in the inner race, it is a complete challenge to detect it in an early stage, due to the low amplitude of the spectral vibration components related to this fault (BPFI), which may be hidden by the background noise of the signal and by the cyclostationary components of FOP. The waveform time and the frequency spectrum of the vibration signal taken from the accelerometer mounted on the bearing A, are shown in Figures 3a and 3b, respectively. From the waveform time, some impulsive events can be identified, which seem to be modulated and periodic, however, it is not possible either from the waveform or the spectrum, to identify precisely their periodicity and/or frequency of repetition, which should be equal to the BPFI frequency listed in Table 2.

Table 2. Main fault frequencies of the radial ball bearing.

Component	Frequency [Hz]
Shaft rotational speed	10
Ball pass frequency, inner race (BPFI)	44.26
Ball pass frequency, outer race (BPFO)	25.74

Figure 8 (a)

Figure 8. Residual signal: band-pass filtered (2800-6500 Hz)-case I.

(a)

(b)

Figure 9. Residual signal: filtered and demodulated-case I.

(a)

(b)

Figure 10. Residual signal: filtered, demodulated and auto-correlated-case I.

Figure 11. Sketch of the test rig of case II-bearing with a fault in the inner race.

Figure 12. Waveform time and spectrum of the filtered signal (FOP components)-case II.

Figure 13. Waveform time and spectrum of the residual signal (SOP components and noise)-case II.

Following with the application of the proposed method, the FOP components were separated from the raw signal using an adaptive filtering scheme and using the same parameters for the filter used in case I. The waveform time and the spectrum of the filtered signal, which contains the FOP cyclostationary components, are shown in Figure 12 and the waveform time and the spectrum of the residual signal, which contains the SOP cyclostationary components and signal noise, are shown in Figure 13.

In contrast to the results obtained in case I for the filtered and residual signals (see Figures 5 and 6), in this case, the cyclostationary components of SOP are predominated when they are compared to the FOP components in the signal. This behavior can be due to several factors: the modulation of the zone load over the localized fault in the inner race (i.e. the inner race is rotating at the rotational shaft speed), slip motion between the rolling elements and the races, and random vibration components generated by friction mechanisms (e.g. the friction in the -tensor-pulley system).

In order to identify an appropriated frequency range for the further filtering stage, the ZAM transform was applied to the residual signal, obtaining the time-frequency distribution shown in Figure 14 and computed with a $\Delta f \approx 110$ Hz. In this figure, it can be identified the presence of short duration events in the range of frequency between 800 to 2000 Hz, which may involve resonant frequencies of the bearing races. Therefore, this frequency range is selected for the configuration of band-pass filter used to filter the residual signal.

Figure 14. ZAM time-frequency distribution of the residual signal-case II.

Figure 15. Residual signal: band-pass filtered (800-2.000 Hz)-case II.

The waveform time and the spectrum of the filtered signal, using a FIR filter with 400 coefficients, are shown in Figure 5. Even though, the impulsive events are notorious in the waveform time of the filtered signal, it is still not possible to determine clearly their periodicity. However, when the filtered signal is demodulated the fault frequencies at BPFI (44 Hz) can be precisely identified from the spectrum, as shown in Figure 16.

Figure 16. Residual signal: filtered and demodulated-case II.

Finally, and in order to reconfirm the results obtained, the auto-correlation function is applied to the demodulated signal in order to enhance even more the main components of interest, as it can be seen in the results shown in Figure 17. In this way, the diagnostic of the fault is confirmed and it is very precise since the frequency of the main spectral component found is very close to the theoretical value of the BPFI frequency. With the results obtained from the analysis of these two cases, it is shown the effectiveness of the proposed methodology when it is applied to the detection and diagnostic of localized faults, particularly in gear systems and rolling bearings.

CONCLUSIONS AND FUTURE ASPECTS

In this work, it has been proposed a practical procedure based on the cyclostationary analysis of vibration signals, which can be used for the early detection of localized faults in mechanical components, such as gears and bearings. Vibration signals can be assumed as a combination of FOP and SOP components.

Figure 17. Residual signal: filtered, demodulated and auto-correlated-case II.

FOP signals are generated by rotating machines with misaligned couplings and/or unbalanced rotors, whereas, SOP components correspond to the modulated vibratory signals generated by wear mechanisms, friction and impact forces.

To analyze FOP signals and to extract the required information for the fault detection tasks, the classical spectral analysis is an appropriate tool. For SOP signals the analysis should be carried out using more sophisticated

tools. In this work, a procedure to analyze SOP signal is presented and tested in two practical cases. From the time-frequency analysis, appropriate filters are configured in order to obtain the SOP signal and then, using an enveloping detector the useful signal is obtained. To improve the signal to noise ratio (SNR), a matched filter based in the autocorrelation function is used.

The results of applied the proposed method has been validated with the analysis of vibration data taken from two different laboratory test rigs. In order to extent this results to industrial applications, some of the further aspects of the present work include the analysis of vibration data taken from industrial rotating machinery, such as, multi-stage gearboxes and rotors supported on rolling bearings. Additionally, the procedure can be adapted and modified to be applied in cases where the early detection of localized faults is even more challenged, as for instance, detection of faults in the rolling elements of bearings and in general in mechanical systems with variable load and speed conditions. Finally, when more than one fault occur is possible to obtain the detection of several faults in time-frequency analysis, but in this case, several SOP components are obtained. Nevertheless, this aspect is part of the future research work.

ACKNOWLEDGMENTS

This paper was partially supported by Universidad de La Frontera, DIUFRO DI08-0048.

REFERENCES

- [1] J. Antoni, F. Bonnardot, A. Raad and M. El Badaoui. "Cyclostationary modelling of rotating machine vibration signals". *Mechanical Systems and Signal Processing*. Vol. 18 N° 6, pp. 1285-1314. 2004.
- [2] A.C. McCormick and A.K. Nandy. "Cyclostationarity in Rotating Machine Vibrations". *Mechanical Systems and Signal Processing*. Vol. 12, Issue 2, pp. 225-242. March 1998.
- [3] F. Bonnardot, R.B. Randall and F. Guillet. "Extraction of second-order cyclostationarity sources-Application to vibration analysis". *Mechanical System and Signal Processing*. Vol. 19, Issue 6, pp. 1230-1244. 2005.
- [4] Jérôme Antoni. "Cyclostationarity by examples". *Mechanical Systems and Signal Processing*. Vol. 23, Issue 4, pp. 987-1036. May 2009.
- [5] C. Capdessus, M. Sidahmed and J.L. Lacoume. "Cyclostationary Processes: Applications in Gear faults early Diagnosis". *Mechanical Systems and Signal Processing*. Vol. 14, Issue 3, pp. 371-385. 2000.
- [6] S.K. Lee and P.R. White. "The enhancement of impulsive noise and vibration signals for fault detection in rotating and reciprocating machinery". *Journal of Sound and Vibration*. Vol. 217 N° 3, pp. 485-505. October 1998.
- [7] A.D. Poularikas. "The Transforms and Applications". *Handbook*. Second Edition. Boca Raton, Florida, USA. 2000.
- [8] Yunxin Zhao, Les E. Atlas and Robert J. Marks. "The Use of Cone-Shaped Kernels for Generalized Time-Frequency Representations of Nonstationary Signals". *IEEE Transactions on Acoustics, Speech, and Signal Processing*. Vol. 38, pp. 1084-1091. 1990.
- [9] R.B. Randall. "A new method of modeling gear faults". *Journal of Mechanical Design*. Vol. 104, pp. 259-267. 1982.
- [10] Y.T. Su and S.J. Lin. "On initial fault detection of a tapered roller bearing: frequency domain analysis". *Journal of Sound and Vibration*. Vol. 155, pp. 75-84. 1992.