Cyclostationarity and Resampling for Vibroacoustic Signals

J. Leśkow

Department of Quantitative Methods in Management, WSB-NLU, Nowy Sącz, Poland

Cyclostationarity as a feature of signals was extensively studied in last decades. Publications of Serpedin or Gardner provide an important survey documenting the impact of cyclostationary models on signal analysis in telecommunication, mechanical, acoustic, and econometric signals. The aim of this paper is to introduce the concept of cyclostationarity for mechanical signals and to present the possibilities of various resampling procedures available for such signals. Recent research of Leśkow has shown applicability of bootstrap and subsampling procedures to cyclostationary models. This means, in particular, that the modern statistical mechanical signal analysis can be done without the assumption of Gaussianity and/or without specific linear filtration structure imposed on the signal. In the paper, we show that subsampling procedure proves to be a useful diagnostic tool for testing the cyclic autocorrelation structure for signals generated by vibration.

PACS: 02.50.−r, 05.20.−y

1. Introduction

Our main object of study is a signal $\{X(t) : t \in \mathbb{Z}\}$ that describes a vibrating mechanical system such as gear drive. It is already well known that cyclostationarity of such signals is a feature commonly occurring for such signals. The cyclostationarity of $\{X(t) : t \in \mathbb{Z}\}$ is a repeatable behavior of the first and second order characteristic of such signal. To be more precise, let the mean $\mu_X(t) = E(X(t))$ and the autocovariance function $B_X(t, \tau) = \mathrm{Cov}(X(t), X(t+\tau))$. The signal will be called *cyclostationary* if the mean $\mu_X(t)$ and the autocovariance function $B_X(t, \tau)$ are periodic or almost periodic in t. The mathematical framework for analysis of such signals was presented e.g. in papers of Dehay and Leśkow (1996) (see [3]). In our research, we will focus on second-order properties of the signal $\{X(t) : t \in \mathbb{Z}\}$ therefore we assume that the signal under study is mean-zero.

From the theory of the cyclostationary signals we get that the autocovariance function $B_X(t, \tau)$ can be represented as

$$B_X(t, \tau) = \sum_{\lambda \in \Lambda} a(\lambda, \tau) e^{i \lambda t},$$

where

$$a(\lambda, \tau) = \frac{1}{T} \sum_{t=1}^{T} B_X(t, \tau) e^{-i \lambda t},$$

and the set

$$\Lambda = \{\lambda : a(\lambda, \tau) \neq 0\}$$

is finite. Such representation is helpful to solve the estimation problem of $B_X(t, \tau)$ via the estimation of its complex Fourier coefficients $a(\lambda, \tau)$. When the exact period T of the autocovariance function $B_X(t, \tau)$ is known,

then the set Λ contains integer multiples of the fundamental frequency $\lambda = \frac{2\pi}{T}$. However, in analysis of various vibration signals it is not at all clear what are the elements in the set Λ. The important aspect of the present research is to show how modern statistical analysis of cyclostationary signals can help identifying the frequencies $\lambda \in \Lambda$.

Assume that a sample $\{X(1), X(2), \ldots, X(n)\}$ is available from the observed signal $X(t)$. Then a natural estimator (see [6]) of the parameter $a(\lambda, \tau)$ can be expressed as

$$\hat{a}_n(\lambda, \tau) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-\tau} X(t+\tau) X(t) e^{-i \lambda t} \\ \qquad \text{for } \tau \geq 0, \\ \frac{1}{n} \sum_{t=-\tau+1}^{n} X(t+\tau) X(t) e^{-i \lambda t} \\ \qquad \text{for } \tau < 0. \end{cases} \tag{1}$$

The estimator (1) in the sequel will be called the cyclic spectrum estimator. This estimator combines the time-lag information represented in its second argument τ and the modulating frequency information represented in its first argument λ. One has to note that the statistical analysis of the estimator (1) is fundamental to identify the cyclostationary nature of the signal X. If for all nonzero frequencies λ the parameter $a(\lambda, \tau)$ is not significantly different from zero, then it can be reasonably assumed that the original signal X is stationary and does not exhibit cyclostationarity. On the other hand, if the signal X is cyclostationary, then the parameter $a(\lambda, \tau)$ will be significantly different from zero and so should be the estimate (1). To illustrate this point, we show Figs. 1 and 2.

The following pictures are showing the time–frequency analysis of signals based on cyclostationary models. For a simple example such as above it is clear that the sta-
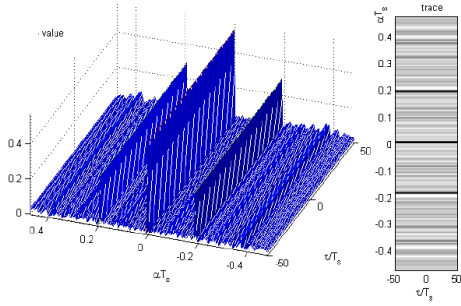
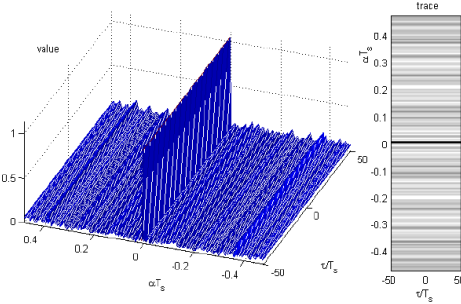Fig. 1. Amplitude modulated signal $X(t) = f(t)Z(t)$ cyclic spectrum and its trace.

Fig. 2. Stationary background noise $Z(t)$ and its cyclic spectrum.

tionary background noise is not showing any significant frequency beyond the zero frequency. For the amplitude modulated signal we clearly see an additional line beyond the zero frequency corresponding to the frequency.

For cyclostationary signals, however, we have also available the spectral analysis based on the concept of bifrequency plane analysis. The cyclostationary signals $\{X(t) : t \in \mathbb{Z}\}$ admit also the spectral decomposition

$$X(t) = \int_0^{2\pi} e^{i\xi t} Z(d\xi). \tag{2}$$

The formula (2) is defining the spectral bimeasure as

$$R((a,b] \times (c,d]) = E[(Z(b) - Z(a))(Z(d) - Z(c))], \tag{3}$$

with a spectral support

$$S = \bigcup_{\lambda \in \Lambda} \{(\xi_1, \xi_2) \in (0, 2\pi]^2 : \xi_2 = \xi_1 \pm \lambda\}. \tag{4}$$

The above spectral approach to cyclostationary signals allows to consider the spectral density estimator

$$\hat{G}_n(\nu, \omega) = \frac{1}{2\pi n} \sum_{t=1}^n \sum_{s=1}^n K_n(s - t) X_t X_s e^{-i\nu t} e^{i\omega s}. \tag{5}$$

Areas for which the estimator is defined form the set of parallell lines depicted in the figures above.

The main task of this article is to provide a modern statistical tool for the cyclostationary signal analysis without invoking the assumption of Gaussianity or a specific

parametric form of the signal. Such approach is possible in view of recent research of the author dedicated to resampling procedures for cyclostationary signals (see e.g. [6] or [7]). In what follows, we will briefly review state-of-the art results on resampling for cyclostationary signals and will show a specific application in cyclic spectrum estimation and frequency identification. The introduced method will allow to identify the frequencies with a statistical significance.

2. Resampling methods for cyclostationary signals

The idea of resampling for signals and time series has been vigorously investigated in the last three decades. Useful surveys of these results can be found e.g. in the monograph of Lahiri [5] or in the survey paper of Zoubir and Iskander [9]. In the context of cyclostationary signals the pioneering papers [6] and [7] have provided useful discussions of specific resampling techniques like periodic block bootstrap and subsampling. Resampling, in general, is based on recalculating the value of the estimate on the subset of data selected from the original sample. To be more specific, assume that $\{X(1), X(2), \ldots, X(n)\}$ is a sample from a signal $\{X(t) : t \in \mathbb{Z}\}$. Let θ be a parameter corresponding to the signal, for example mean, autocovariance, higher order moments, quantiles etc. Moreover, let $\hat{\theta}_n$ be an estimate of θ based on the sample $\{X(1), X(2), \ldots, X(n)\}$. One of the fundamental tasks of statistical inference for signals is to build confidence intervals and test for the parameter θ using the estimate $\hat{\theta}_n$. To complete such a task, classical statistical inference is using normal distribution to approximate variability of the estimate. More modern statistical techniques propose various forms of resampling to complete such task without resorting to normality or other asymptotic distribution of $\hat{\theta}_n$. In principle, a subset $\{X(i_1), X(i_2), \ldots, X(i_n)\}$ is drawn from the initial sample $\{X(1), X(2), \ldots, X(n)\}$. Such drawing can be done in a random manner (see e.g. periodic block bootstrap described in [7]) or in a deterministic way (see e.g. subsampling technique described in [6]). On such a subsample $\{X(i_1), X(i_2), \ldots, X(i_n)\}$ a new value of the estimate $\hat{\theta}_n^{*1}$ is calculated. Repeating such process B times we are getting B values $(\hat{\theta}_n^{*1}, \ldots, \hat{\theta}_n^{*B})$, respectively. Using these values we can provide confidence intervals and tests for the initial parameter θ provided that the considered resampling method is statistically adequate. For more discussion on that see e.g. [5]. Here, we will focus on subsampling and its application to cyclostationary signals.

Recall that the estimator $\hat{a}_n(\lambda, \tau)$, defined in (1), is the central point of our investigation here. This estimator is based on the sample $\{X(1), X(2), \ldots, X(n)\}$ from the time series $\{X(t) : t \in \mathbb{Z}\}$. Let $b(n)$ be any sequence tending to infinity with n, such that $b(n) < n$. In the sequel, we will simply write b in place of $b(n)$. Take also any $t \in \{1, \ldots, n - b + 1\}$ and define:

$$\hat{a}_{n,b,t}(\lambda,\tau) = \frac{1}{b}\sum_{j=1}^{b-\tau} X(t-1+j+\tau)$$

$$\times X(t-1+j)\mathrm{e}^{-\mathrm{i}\lambda j} \qquad (6)$$

as the estimator of the cyclic spectrum $a(\lambda,\tau)$ based on the subsample $\{X(t), X(t+1),\ldots, X(t+b-1)\}$. We can interpret t as a starting point for recalculating our estimate on the block of the size b. The classical statistical result for the estimator $\hat{a}_n(\lambda,\tau)$ (see (1)), developed in [3] provides an approximation of variability of that estimator via some limiting normal distribution. However, the parameters of such limiting distribution are very difficult to calculate in practice and in recent years subsampling was proven to be more reliable technique (see [6] for a detailed discussion).

To simplify our discussion, let us focus on the absolute value $|\hat{a}_n(\lambda,\tau)|$ of the estimate of the cyclic spectrum. From the subsampling procedure we get the $n-b+1$ values of the estimate (6) calculated for $n-b+1$ samples. We further consider absolute values from such sequence and we are ordering such values to get $\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$ quantiles from them, where α is a desired significance level. Let us call such quantiles as $c_{n,b}\left(\frac{\alpha}{2}\right)$ and $c_{n,b}\left(1-\frac{\alpha}{2}\right)$, respectively. Then the subsampling-based confidence interval for the cyclic spectrum will have the form

$$\left(|\hat{a}_n(\lambda,\tau)| - \frac{c_{n,b}(1-\frac{\alpha}{2})}{\sqrt{n}}, |\hat{a}_n(\lambda,\tau)| - \frac{c_{n,b}(\frac{\alpha}{2})}{\sqrt{n}}\right), \quad (7)$$

where $\alpha \in (0,1)$ is a significance level.

From the confidence interval derived in (7) is it also quite straightforward to derive tests for significant values of the cyclic spectrum. If, for a fixed λ and all τ, zero belongs to the interval (7), then the cyclic spectrum $a(\lambda,\tau)$ is insignificant, that is for such λ there is no significant nonzero value of the cyclic spectrum. On the other hand, if for a fixed λ and at least one τ we will show that zero is NOT inside the interval (7) then for such λ the cyclic spectrum $a(\lambda,\tau)$ is significant. In such a case we are also identifying the frequency λ as significant. In the next section, we will show how this technique applies to vibroacoustic signals.

3. Applications to mechanical signals

Vibrations of machines form an important part of contemporary methods of diagnostics. Very frequently the observed signals are mixtures of various impulses resulting from different physical phenomena. Therefore, studying the *frequency signature* of signals is becoming a central part of signal analysis (see e.g. [2]). In our study, we focus on mixtures of signals that lead naturally to cyclostationary models studied in previous sections. There is already large body of research dedicated to such signals (see e.g. [8, 9]) but, according to the author's knowledge, the resampling methods like moving block bootstrap or subsampling were not applied yet to the problem of identifying relevant frequencies of such signals. In what follows, we present new diagnostic tools of identify-
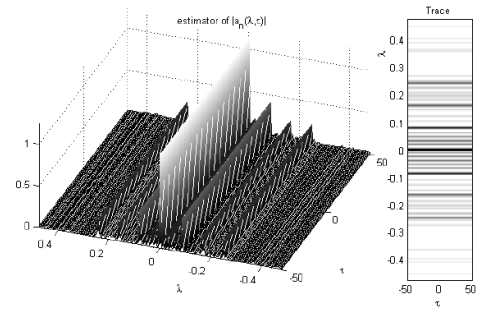

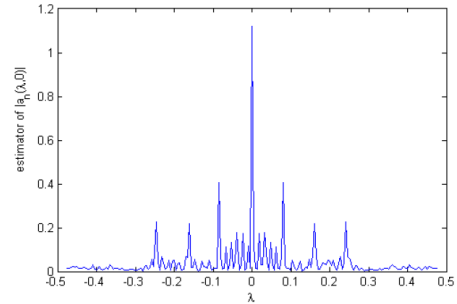Fig. 3. Cyclic spectrum and trace for the properly working gear drive.

Fig. 4. Cross-section of the cyclic spectrum for properly working gear drive — $\hat{a}_T(\lambda,0)$.
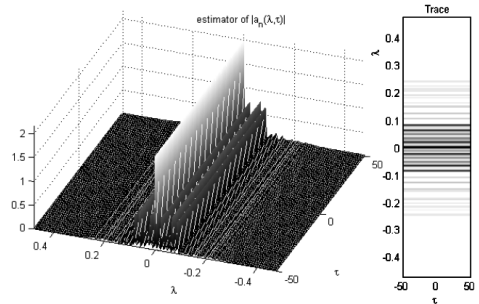

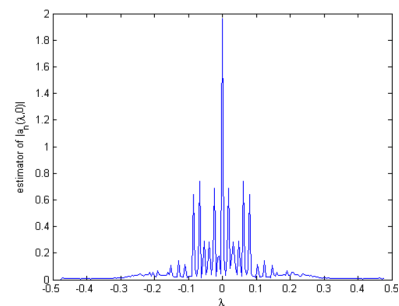Fig. 5. Cyclic spectrum and trace for the case of surface fatigue.

Fig. 6. Cross-section of the cyclic spectrum in the case of surface fatigue — $\hat{a}_T(\lambda,0)$.

ing *statistically significant* frequencies of cyclostationary signals using resampling methods. Our signals come from vibromechanics. For such signals we study three important features: the cyclic spectrum, the trace in frequency domain and statistically significant frequencies using subsampling.

We will apply our statistical tools to diagnostics of gear drives where pitting process induced by the surface fatigue should be monitored. The traditional method of diagnosing the pitting process in mechanical gear drive is based on visual inspection and is described by the Polish norm PN=91/M-88506. Essentially, the pitting process corresponds to the surface fatigue occurring in gear drives. In order to check our method, the sensors were placed on two gear drives. The first was running properly and the second has exhibited the pitting process related to surface fatigue beyond the norm. The visual inspection is, therefore, replaced by statistical estimation of the signal based on the cyclostationary model. We can clearly see that the frequency signature of the signal for the properly working gear drive is significantly different from the frequency signature in the gear drive exhibiting the pitting process due to surface fatigue. Thus, we can reinforce our traditional visual-inspection based tools with modern statistical diagnostics that is based on cyclostationary signals.

As an example, in Figs. 3 and 4 we show the vibromechanical signal induced by the properly working gear drive.

The second example is a vibromechanical signal for gear drive exhibiting surface fatigue and pitting, shown in Figs. 5 and 6.

Analyzing the pictures above we see that the frequency signature for the properly working gear drive is completely different from the case when there is a surface fatigue. Our approach can be used to identify the surface fatigue.

References

[1] J. Antoni, *Mechanical Systems and Signal Processing*, doi:10.1016/j.ymssp.2008.10.010 (2009).

[2] J. Antoni, F. Bonnardot, A. Raad, M. el Badaoui, *Mechanical Systems and Signal Processing*, **18**, 1020 (2004).

[3] D. Dehay, J. Leśkow, *J. Appl. Probab.* **33**, 1077 (1996).

[4] W. Gardner, A. Napolitano, L. Paura, *Signal Process.* **86**, 639 (2006).

[5] S. Lahiri, *Resampling Methods for Dependent Data*, Springer Verlag, 2003.

[6] J. Leśkow, *J. Time Series Anal.* **29**, 253 (2008).

[7] J. Leśkow, R. Synowiecki, *Metrika* **71**, (3) (2009).

[8] E. Serpedin, F. Panduru, I. Sari, G.B. Giannakis, *Signal Process.* **85**, (2005).

[9] A.M. Zoubir, D.R. Iskander, *IEEE Sign. Process. Mag.* **10**, (2007).