Journal of Sound and Vibration (2001) **244**(1), 155–167
doi:10.1006/jsvi.2000.3466, available online at http://www.idealibrary.com on IDEAL®

LETTERS TO THE EDITOR

MACHINE SIGNATURE IDENTIFICATION BY ANALYSIS OF IMPULSE VIBRATION SIGNALS

Z. CHEN AND C. K. MECHEFSKE

Department of Mechanical and Materials Engineering, The University of Western Ontario, London, Ontario, Canada N6G 5B9. E-mail: mechefske@uwo.ca

(*Received* 15 *May* 2000, *and in final form* 18 *September* 2000)

1. INTRODUCTION

With quality and efficiency being of continual importance in industry, a significant amount of research activity has been devoted to machine signature analysis (MSA) [1–3]. The major tool used for MSA has been spectral analysis. Classical spectral estimation techniques include Periodogram, Averaged Periodogram and Blackman–Tukey spectral estimation. Newer approaches to spectral analysis include a variety of parametric modelling techniques. Within this category are the rational transfer function modelling method, autoregressive (AR) PSD estimation, moving average (MA) PSD estimation, autoregressive moving average (ARMA) PSD estimation, Prony spectral density estimation and maximum likelihood method (MLM) [2, 3].

The majority of the research and development work carried out to date with regard to signal processing strategies for machine condition monitoring and diagnostics applications has focused on signals generated from steady state processes. Transient processes have been left relatively unstudied. Examples of methods that are applicable to transient processes are wavelet and short-time fast Fourier transform [1, 4]. The Prony method, originated by the French scientist Baron de Prony in 1795 [5], is also capable of analyzing transient processes [5, 6] and is inherently suitable for the study of exponentially decaying dynamic signals, such as those that develop as a result of many different types of machinery condition deterioration [6–9]. This paper explores the use of the Prony method to study such transient signals.

Similar to the well-known system identification techniques such as AR and ARMA models, the Prony model seeks to fit an exponential model, which is a linear combination of a series of exponentially decaying sinusoidal functions, to sampled data. The Prony method first determines the linear prediction parameters that fit the sampled data. Such linear prediction parameters are then used as coefficients to form a polynomial. The roots of this polynomial are finally employed to estimate the damping coefficients, the sinusoidal frequencies, the exponential amplitude and sinusoidal initial phase of each of the exponential terms.

2. NUMERICAL EXPERIMENT

2.1 NUMERICAL IMPLEMENTATION

The application of the Prony method in applications other than machine condition monitoring is well understood and documented [2, 3, 5]. The Prony modelling and

0022-460X/01/260155 + 13 $35.00/0

TABLE 1

Function parameters recaptured using the Prony method

Amplitude	Damping constant	Frequency (Hz)	Phase (rad)
5·0000000000	– 0·8000000000	10·0000000000	1·5707963268
3·0000000000	– 0·3000000000	40·0000000000	– 1·5707963268
10·0000000000	– 0·5000000000	25·0000000000	– 3·1415926536
10·0000000000	– 0·5000000000	26·0000000000	0·7853981634

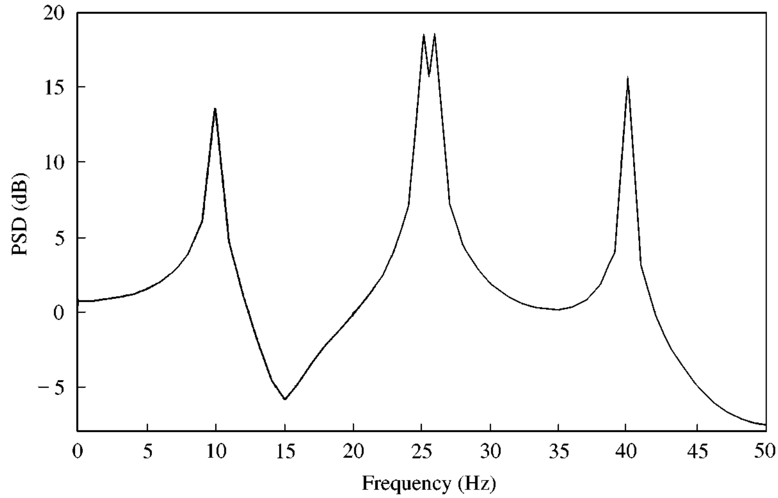

Figure 1. Spectral presentation of data generated using equation (1).

spectrum estimation procedures used in this work were coded in a FORTRAN 77 program with double precision. Some of the subroutines were borrowed from reference [5] with modifications from reference [10].

2.2. VALIDATION OF ROUTINES

In this section, examples are presented for validation of the computer program. Four deterministic functions were combined to generate data sets, which in turn were used as input to the program for computation of a Prony model that then describes the data set. All the data sets generated using the deterministic functions consist of 64 points. Computation results collected include function parameters recaptured and spectral figures. The first example is

$$x(t) = 10\exp(-0{\cdot}5t)\cos(2\pi 25t + \pi) + 10\exp(-0{\cdot}5t)\cos(2\pi 26t + \pi/4)$$
$$+ 5\exp(-0{\cdot}8t)\cos(2\pi 10t + \pi/2) + 3\exp(-0{\cdot}3t)\cos(2\pi 40t + 3\pi/2). \quad (1)$$

The parameters and the spectrum computed by the prony method are shown in Table 1 and Figure 1. The order of the prony model used was 8. The peaks in the figure appear at frequencies of 10·0, 25·0, 26·0 and 40·0 Hz respectively. All parameters are accurately recaptured.

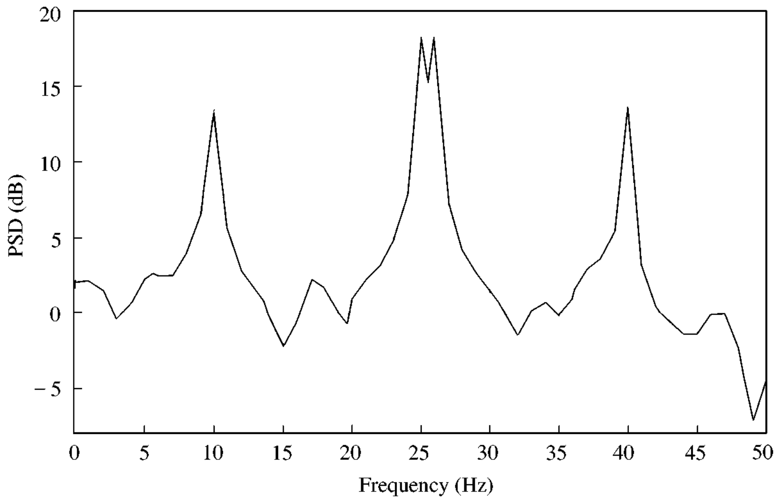

Figure 2. Spectral presentation of simulated data generated using equation (2).

2.3. SIMULATION

The simulated data were generated according to the following function, which consists of four exponential components and a stationary Gaussian white noise process. The data consists of 256 points:

$$x(t) = 10\exp(-0{\cdot}5t)\cos(2\pi 25t + \pi) + 10\exp(-0{\cdot}5t)\cos(2\pi 26t + \pi/4) + 5\exp(-0{\cdot}8t)\cos(2\pi 10t + \pi/2) + 3\exp(-0{\cdot}3t)\cos(2\pi 40t + 3\pi/2) + W(t). \quad (2)$$

Here, $W(t)$ is a Gaussian white noise process with zero mean and r.m.s. of 1·0. The spectral plot computed by using the program is presented in Figure 2 where the order of the Prony model is 32. The peaks in the figure are clearly visible at frequencies of 10·0, 25·0, 26·0 and 40·0 Hz as expected.

3. METHODOLOGY ENHANCEMENT

In engineering practice, many factors can affect the accuracy of measurement and computation results, which do not exist in the laboratory environment. To ensure the efficiency and the effectiveness of the proposed technique, it is necessary to enhance the methodology in order to increase its robustness. To this end, a few measures have been taken and are described in this section.

3.1. MINIMIZING THE NUMERICAL DEVIATION

It is commonly understood that wherever numerical computation is involved, numerical error exists. It is, however, possible to reduce such errors through steps such as careful selection of numerical algorithm and routines. In this work, the following steps were taken: (1) increase of the program precision—double precision was used; (2) use of longer digit platform machine—32 bit; (3) test and comparision of different algorithms; (4) test and comparison of different subroutines.

The critical root finding routines were tested and compared extensively and the "three-stage variable shifting iteration" by Jenkins and Traub [11, 12] was proven to be the best in handling different situations. For the least-squares linear prediction estimation method, the covariance method was adopted. The "fast algorithm to solve the covariance normal equation" by Morf *et al.* and Marple [5] was employed. To examine the robustness of the program, a number of experimental data sets of various lengths were analyzed and the computation results were compared. It was concluded after the comparisons that the accuracy was satisfactory. The numerical results are the same up to eight decimal points. Presented below is one of those examples. The input data in this example was cut in half iteratively until a minimum data length was reached. Minimum data length means the reciprocal of the lowest dominant frequency in the sample data. From the figures, one sees that the results computed by using the minimum data length are the same as those when using a full set of data. They are, in fact, virtually identical.

Figures 3 and 4 are examples that illustrate the robustness of the program. Figure 5 is a spectral plot calculated by using the input vibration data shown in Figure 3. Figure 4 presents the input data which is one-sixteenth of that in Figure 3. Figure 6 is the corresponding spectral plot computed by using the short input data shown in Figure 4. It is easy to observe that both spectral plots are virtually identical. This indicates that the program is able to offer accurate output with minimum input data. Several other sets of experimental data were used to test the procedure with similar results.

3.2. SINGULAR VALUE DECOMPOSITION

Background noise is, and will always be an issue accompanying any type of data acquisition and processing. The reduction of noise is sometime critically important in order to extract useful information from the sampled data. It would thus be beneficial to include a data reduction technique with the program. There exist several approaches for noise reduction that can be used: FILTERING which includes analog filtering and digital filtering, AVERAGING and numerical methods. The filtering techniques are integrated systems whereas the averaging technique usually requires many trials and is therefore not suitable here as the focus of this investigation is non-repeatable short data samples. It appears that the appropriate technique to be incorporated into the program is the numerical approach. The Singular Value Decomposition (SVD) is a proven numerical technique that can be adopted for noise reduction by truncating decomposition terms of the data matrix.

The procedure for applying the SVD technique is: first calculating the singular values of the data matrix; rearranging the singular value in descending order and finally truncating the eigenvectors associated with the small eigenvalues. If a signal consists of m components with background noise, then the m eigenvectors associated with the m largest singular values primarily span the m major components while the remaining smaller singular values mainly span the noise components.

Presented in this subsection are two examples to demonstrate the effectiveness of the application of the SVD in noise reduction. Figure 7 is a spectral plot calculated with simulated signal data in additive noise and Figure 8 is the corresponding spectral content computed by using the same set of data with the application of SVD. One clearly sees that the latter is much clearer in terms of revealing the dominant frequencies. Figures 9 and 10 show another similar example, computed by using experimental data with additive noise, to prove the effectiveness of the SVD in reducing the effect of additive noise.

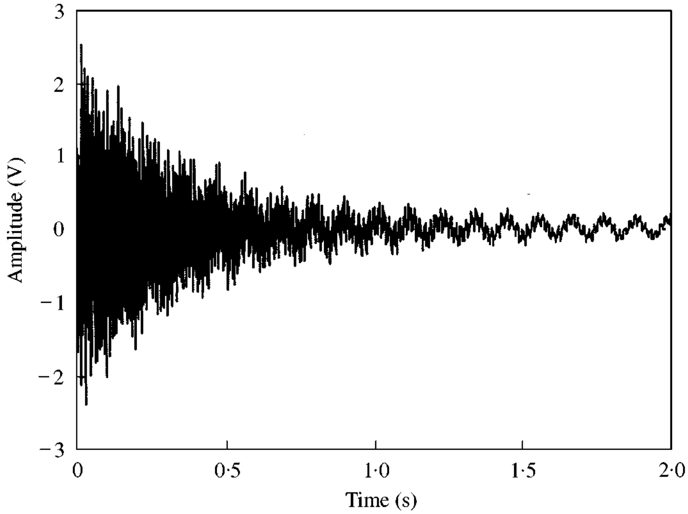

Figure 3. Impact response vibration data from a cantilever beam (full data set).

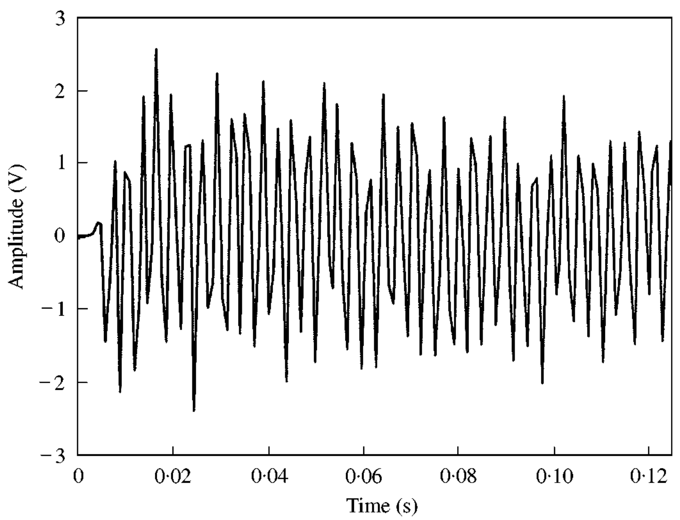

Figure 4. Vibration input data (one-sixteenth of that in Figure 3).

3.3. MODEL ORDER SELECTION

The model order selection is an extremely important part of parametric approaches and in industrial applications. Improper selection of model order will cause either loss of information or introduction of extra peaks in the spectra. Either case may lead to incorrect or inaccurate judgements. There has been extensive research on model order selection carried out over the last few decades with many different criteria described. These include the Akaike Information Criterion (AIC) [13–15], the Minimum Descriptive Length (MDL) criterion [16, 17] and the Hannan criterion [18]. These criteria are based on or partially based on the maximum likelihood principle. A thorough comparison of the various model order selection criteria has been carried out by the authors of this paper [19, 20].

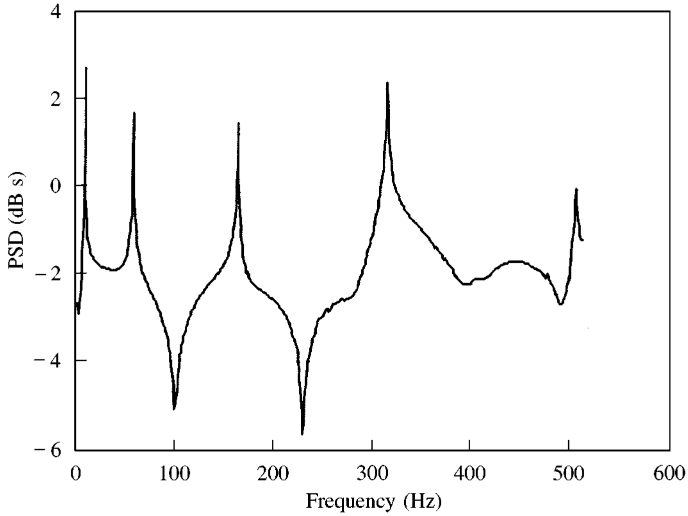

Figure 5. Spectral plot using full set of input data shown in Figure 3.

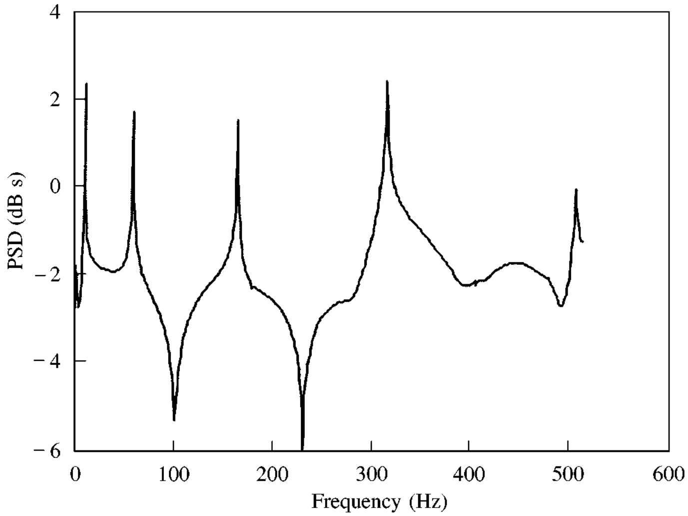

Figure 6. Spectral plot using the short data set shown in Figure 4.

4. LABORATORY TESTS

4.1. CANTILEVER BEAMS

Simple experiments were conducted to illustrate the application of the Prony model to transient vibration signals. The structures tested were two cantilever beams, which were made of mild steel with properties as follows: density of 7860 kg/m^3, elasticity modulus of 200 GPa. The dimensions of the beams are 25·12 mm × 4·61 mm × 643 mm and 37·63 mm × 3·13 mm × 323 mm. Vibrations were generated by impact testing the two beams. The vibration data for the long beam impact response is presented in Figures 3 and 4.

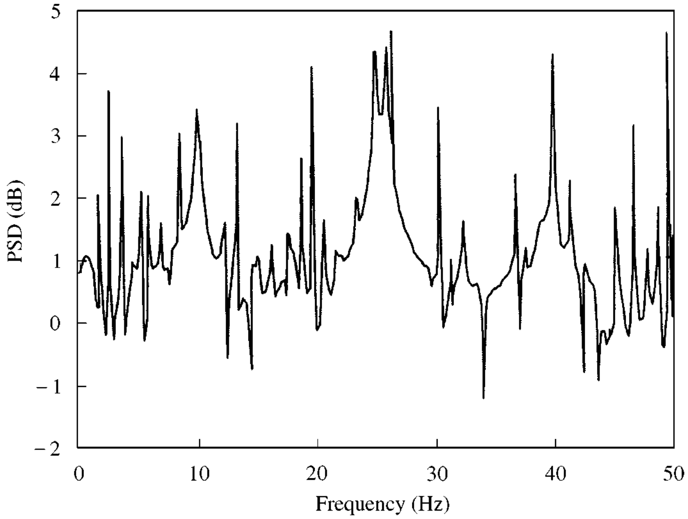

Figure 7. Spectrum of a sample signal [data from equation (1)] in additive noise (without SVD filtering).

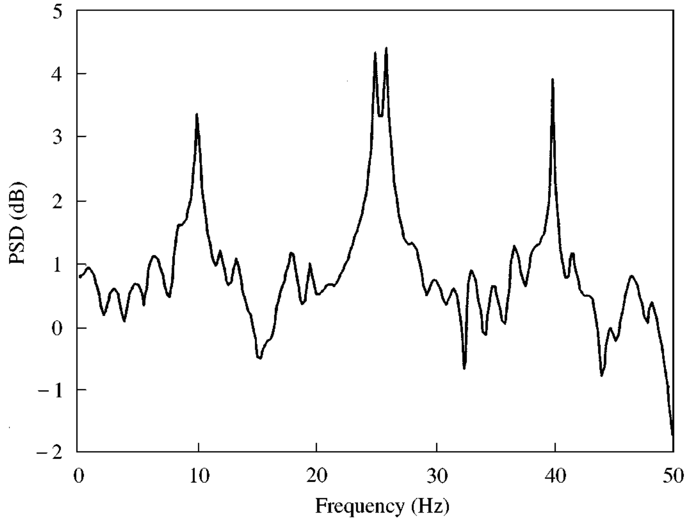

Figure 8. Spectrum of a sample signal [data from equation (1)] in additive noise (with SVD filtering).

Figures 5 and 6 show the frequency spectra computed using the procedure described in this paper. The peaks in Figures 5 and 6 are located at 9·2, 57·3, 163·3 and 314.1 Hz. These locations coincide with calculated theoretical values and finite element analysis results. Similar agreement with theoretical and FEA results was found after analysis of the data from the impact test of the short cantilever beam as well. For the Prony method, one trail (trace) was always enough for the program to produce a clear spectral plot. Such an advantage would be extremely valuable in situations where repeated signals are difficult or even impossible to collect.

To further test the program, the clamped base of the cantilever beams was slightly loosened purposely and impact vibration testing was again conducted. The frequency

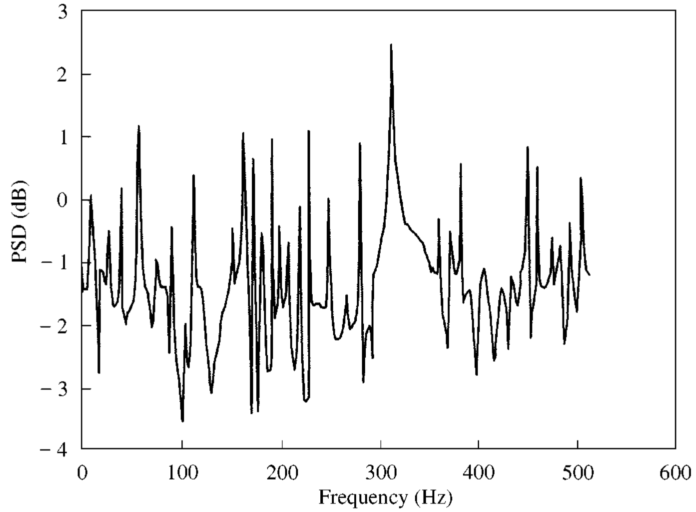

Figure 9. Spectrum of an experimental signal in additive noise (without SVD filtering).

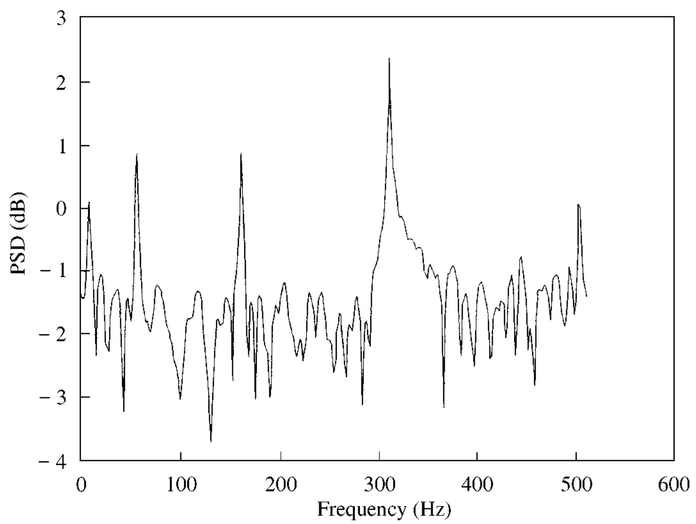

Figure 10. Spectrum of an experimental signal in additive noise (with SVD filtering).

spectral content from one of the tests is presented in Figure 11. The dominant frequencies and the overall pattern of the frequency spectrum have changed. This suggests the potential application of the Prony method in Machine Condition Monitoring where changes in equipment operating condition need to be detected through the analysis of short transient vibration data.

4.2. LOW SPEED FAULTY BEARING TESTS

The next laboratory experiment was conducted to apply the Prony method to vibration signals collected from machinery operating at low rotating speed. Bearing fault detection

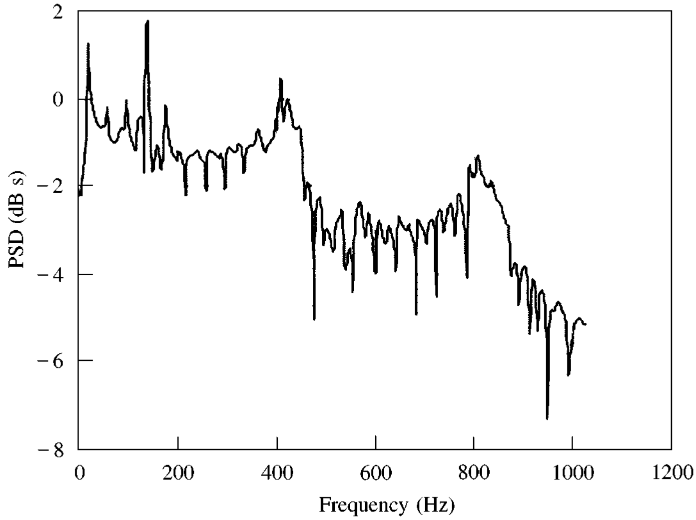

Figure 11. Frequency content of the impact vibration signal from the loosened cantilever beam.

has been a research subject for years and many satisfactory results have been achieved. Most of this research has focused on the machine's operation frequencies (defect frequencies) and therefore a sufficient length of data is needed which may be difficult to collect in the cases where the rotation speed of the bearings is low. Mechefske and Mathew [21–24] applied a modelling approach to deal with this problem and achieved promising results. This section looks at the low-speed bearing fault issue from a different perspective, investigating the high-frequency content. The incentive for such an investigation is the following. For high-frequency investigation, only a very small length of data is needed. This could provide the ability to detect the occurrence of a fault or a change in operating condition. This is particularly important in circumstances where long data samples are difficult or even impossible to collect. This also provides the opportunity for real time application.

Applying the Prony method for bearing fault detection serves the purpose of demonstrating the capability of this method in detecting condition changes using only short length vibration signals. Artificial faults in the form of a groove cut across the full width of the outer race were created with widths ranging from 0·2 to 2·0 mm in 10 increments of 0·2 mm and depths ranging from 0·1 to 1·0 mm in increments of 0·1 mm respectively. The faults were located at the center of the load zone. Vibration signals are collected through an accelerometer placed on the bearing housing in the center of the load zone and close to the fault location. Each time a roller runs over the groove, a small impact signal is generated. Low-frequency investigation analyzes the frequency at which these impact signals occur. This study, by applying the Prony method, looks at analyzing each of these impact signals individually and tries to extract useful information from these short impulses.

Before the bearing was installed, its natural frequency was measured by impact test performed on the bearing hanging without other constraints. The first natural frequency of the bearing is 1200 Hz. This will provide useful information as to approximately what frequency range one should be looking at when analyzing the impulse generated by the rollers. Unlike the cantilever beam cases, one should not expect the natural frequencies to shift significantly when artificial faults are created or the groove is widened because such faults are rather minor in terms of the integrity of the structure as a whole. Instead, one should focus on the energy input to generate such signals.

Figure 12 shows a typical signal collected from the bearing structure operating with artificial faults, in which frequencies below 20 Hz have been filtered out. The Prony method was applied to analyze each impulse of these signals. Figure 13 is the spectral result after applying the Prony method to one impulse corresponding to a fault groove of 0·4 mm. Figure 14 is one that corresponds to a groove of 1·0 mm while Figure 15 is the result associated with a fault groove of 2·0 mm. Notice that the spectral peaks are located at around 1300 Hz instead of 1200 Hz. This is because the bearing is now clamped by the bearing housing and the structure becomes more rigid. As was expected, the frequencies

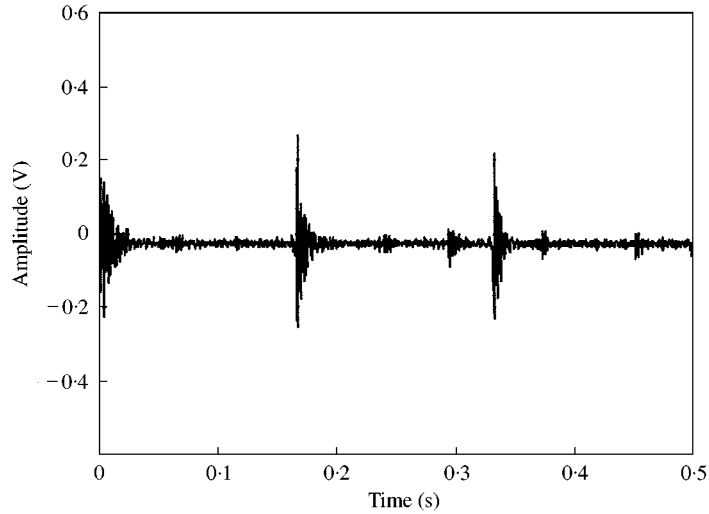

Figure 12. Typical vibration signal collected from faulty bearings.

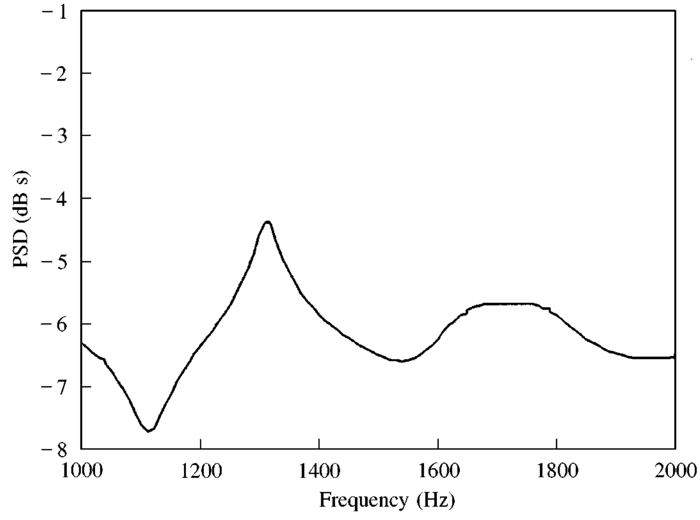

Figure 13. Spectral content of signal from faulty bearing (groove: 0·4 mm).

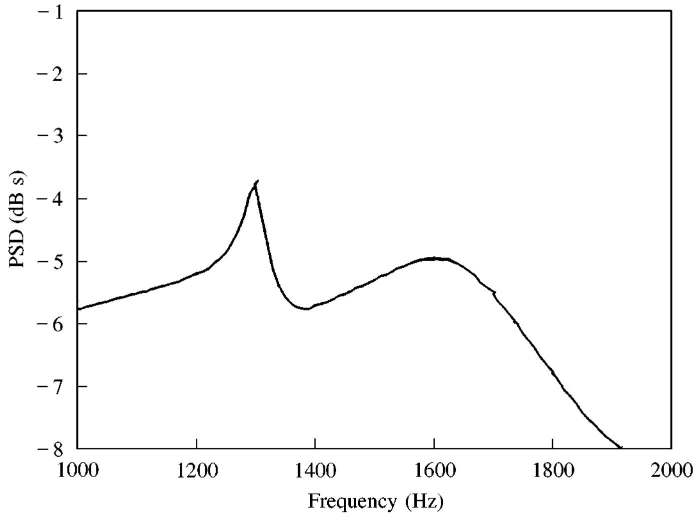

Figure 14. Spectral content of signal from faulty bearing (groove: 1·0 mm).

Figure 15. Spectral content of signal from faulty bearing (groove: 2·0 mm).

change very little with width change of the fault grooves. However, the spectral intensity of the signals changes visibly with the widening of the grooves as one compares Figures 13–15. Such changes can certainly be used as an indicator of fault worsening in condition monitoring. More studies need to be conducted in order to acquire a thorough understanding of the Prony parameters associated with different fault conditions.

5. CONCLUDING REMARKS

Investigations have been conducted which explore the application of the Prony method to the analysis of transient vibration signals. Numerical validation, computer simulation and laboratory testing, including cantilever beams and bearing test rigs, have been conducted to examine the Fortran 77 program that implements the Prony modelling procedures. From the preliminary results presented in this paper, a few concluding remarks can be drawn.

- Selection of the Prony method for transient vibration signal analysis is proven to be appropriate.
- The auxiliary techniques enhance the method and make it more robust.
- The test results show that the program runs well and the computations are accurate and efficient.
- It is revealed that the program has the potential to be adopted in machine condition monitoring and real time application.
- The next step to follow is to further test the program on bearing test rigs and more complicated structures as well as in industrial machine condition monitoring situations.

ACKNOWLEDGMENTS

The authors would like to express their sincere thanks to Prof. S. M. Dickinson for his help and support during the course of this study and to NSERC (National Science and Engineering Research Council of Canada) for financial support.

REFERENCES

1. S. Braun 1986 *Mechanical Signature Analysis—Theory and Applications*. New York: Academic Press Inc.
2. S. M. Kay and S. L. Marple 1981 *Proceedings of the IEEE* **69**, 1380–1419. Spectrum analysis—a modern perspective.
3. S. M. Kay 1988 *Modern Spectral Estimation*: *Theory and Application*. Englewood cliffs, NJ: Prentice-Hall Inc.
4. S. Gade and K. Gram-Hansen 1997 *Sound and Vibration*, 40–46. The analysis of nonstationary signals.
5. S. L. Marple Jr 1987 *Digital Spectral Analysis with Applications*. Englewood Cliffs, NJ: Prentice-Hall.
6. A. G. Innes 1993 Department of Electrical and Electronic Engineering, University of Tasmania, Hobart, Australia, The use of Prony method for modeling induction motor current waveforms under fault conditions.
7. C. K. Mechefske 1996 *Insight* **37**, 599–604. Correlating power transformer tank vibration characteristics to winding looseness.
8. T. Sun, Z. T. Zhang, K. T. V. Grattan and A. W. Palmer 1997 *Review of Scientific Instruments* **68**, 58–63. Analysis of double exponential fluorescence decay behavior for optical temperature sensing.
9. E. M. Dowling, R. D. DeGroat and D. A. Linebarger 1994 *IEEE Transactions on Antennas and propagation* **42**, 590–599. Exponential parameter estimation in the presence of known components and noise.
10 W. T. Vetterling 1992 *Numerical Recipes in Fortran* 77. Cambridge: Cambridge Press.
11. M. A. Jerkins and J. F. Traub 1972 *Communication of ACM* **15**, 97–99. Zeros of a complex polynomial.
12. M. A. Jerkins and J. F. Traub 1970 *Numeriske Mathemtik* **14**, 252–263. A three-stage variable-shift iteration for polynomial zeros and its relation to generalized rayleigh iteration.
13. H. Akaike 1971 *2nd International Symposium on Information Theory*, *Tsahkadsor*, *Armenia*, *USSR*, 267–281. Information theory and an extension of the maximum likelihood principle.
14. H. Akaike 1974 *IEEE Transaction on Automatic Control AC* **19**, 716–723. A new look at statistical model identification.

15. H. AKAIKE 1969 *Annals of The Institute of The Statistical Mathematics* **21**, 407–419. Power spectrum estimation through autoregressive model fitting.
16. J. RISSANEN 1983 *Automatica* **14**, 465–471. Modeling by shortest data description.
17. J. RISSANEN 1983 *The Annals of Statistics* **11**, 416–431. A universal prior for integers and estimation by minimum description length.
18. E. J. HANNEN and B. G. QUINN 1979 *Journal of Royal Statistical Society* **41**, 190–195. The determination of the order of an autoregression.
19. Z. CHEN and C. K. MECHEFSKE 2000 *CSME* (*Canadian Society of Mechanical Engineers*) *Forum* 2000, Montreal, Quebec, Canada, May 16–19. A numerical comparison of model order selection criteria performance.
20. Z. CHEN and C. K. MECHEFSKE, accepted for publication in the journal *Mechanical Systems and Signal Processing*. Model order selection: a practical approach.
21. C. K. MECHEFSKE and J. MATHEW 1992 *Mechanical Systems and Signal Processing* **6**, 297–307. Fault detection and diagnosis in low speed rolling element bearings Part I. The use of parametric spectra.
22. C. K. MECHEFSKE and J. MATHEW 1992 *Mechanical Systems and Signal Processing* **6**, 309–316. Fault detection and diagnosis in low speed rolling element bearings Part II. The use of nearest neighbor classification.
23. C. K. MECHEFSKE and J. MATHEW 1993 *Mechanical Systems and Signal Processing* **7**, 1–12. Parametric spectral estimation to detect and diagnose faults in low speed rolling element bearings: preliminary investigations.
24. C. K. MECHEFSKE and J. MATHEW 1995 *Mechanical Systems and Signal Processing* **9**, 275–286. Fault detection and diagnosis in low speed rolling element bearings using inductive inference classification.