FIFTH INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
DECEMBER 15-18, 1997
ADELAIDE, SOUTH AUSTRALIA

Distinguished Keynote Paper

DEVELOPMENTS IN DIGITAL ANALYSIS TECHNIQUES FOR DIAGNOSTICS OF BEARINGS AND GEARS

R. B. Randall
Associate Professor
School of Mechanical and Manufacturing Engineering
The University of New South Wales
Sydney 2052, Australia

ABSTRACT

For many years it has been recognised that so-called "envelope analysis" is a very powerful tool in the detection and particularly diagnosis of rolling element bearings. A recent quantitative study of the effects of masking signals has shown the benefits of performing the envelope analysis digitally, because of the better quality of filtering and the improved flexibility. This paper demonstrates the considerable enhancements which can be achieved by analysing the squared envelope signal, provided that the ratio of bearing to background signal can be made greater than unity, and discusses methods of achieving this by optimum bandpass filtering and self adaptive noise cancellation to remove masking by gear signals. Techniques are also discussed for dealing with very short signals, and varying signal paths such as for planet bearings in epicyclic gearboxes.

Developments are also proceeding in the field of gear diagnostics, made possible by the availability of new digital signal processing packages and for example direct downloading of vibration signals from a DAT recorder. An example is the increase in the application of time-frequency analysis techniques such as wavelet analysis and variants of the Wigner-Ville distribution. The paper discusses a number of techniques and compares the use of direct transmission error measurement with externally measured vibration signals.

1. INTRODUCTION

Vibration analysis for diagnostic purposes was previously done almost exclusively using dedicated FFT analysers, where the user was constrained by the functions, transform size etc built in by the manufacturer. In recent years, however, there has been a considerable increase in the computing power and memory size of personal computers, as well as in the availability of powerful signal processing packages such as Matlab® [1], which gives individual users a great deal more flexibility in applying diagnostic techniques not available in individual analysers. At the same time it gives researchers a great deal more flexibility in developing new algorithms,

because of the speed with which these can be written and tested. Even the previously difficult problem of obtaining digitised signals using appropriate antialiasing filters has been solved by the availability of multiple channel digital (DAT) tape recorders from which it is possible to download extremely long digital signals selected by previewing the even longer signals recorded (a typical DAT tape contains more than a gigabyte, and newly available versions 25 gigabytes).

It is thus timely to give some information about diagnostic techniques which make use of these currently available tools, and this paper endeavours to do that for the specific cases of rolling element bearings and gears. It will be found that demodulation, both amplitude and phase demodulation, is a useful part of a number of techniques in both areas, and so the paper describes how this can be done very simply digitally.

2. ROLLING ELEMENT BEARING DIAGNOSTICS

In particular in the early stages, faults in rolling element bearings give rise to a series of high frequency bursts as the fault impacts at regular intervals with other bearing elements under load. Where the load at the fault is constant (eg an outer race fault with steady radial load) the high frequency bursts will be reasonably uniform, and represent impulse responses excited by the shocks. The frequency content of such pulses will be dominated by all natural frequencies within the range of the excitation force pulse, which can be tens of kiloherz for incipient small faults, but as will be shown below, they tend to contain very little information at the repetition frequencies or so-called "ballpass frequencies".

Where the load on the fault varies periodically, eg with inner race or rolling element faults with steady radial load, the whole signal will be amplitude modulated at the frequency with which the fault passes through the load zone, eg the shaft speed for inner race faults, and cage speed for rolling element faults [2]. Modulation also occurs as a result of periodic variation in signal transmission path, even for constant load.

Figure 1 is a simulation of acceleration signals from a single outer race fault, with and without variations in pulse spacing. For simplicity, only one resonance frequency is assumed to be excited, for example the first in a series, separated out by lowpass filtration. Even with no spacing variation (Fig. 1(a-b)), the spectrum contains almost no information at the first few harmonics of the ballpass frequency, although this information is to be found in the spacing of the harmonics in the vicinity of the resonance frequency. However, in practice the spacing of the bursts is not uniform, even with constant shaft speed, because of variable slip, and there is usually some random fluctuation, which can give the situation illustrated in Fig. 1(d-e). This can be explained by the fact that the formulas for ballpass frequencies contain a $\cos\phi$ term, where ϕ is the load angle from the radial, and this is something which varies within each revolution as the ratio of radial to axial load changes. The cage keeps the rolling elements roughly evenly spaced, but some slip must occur. As illustrated in Fig. 1(d-e), the random spacing fluctuation gives a smearing of higher order harmonics to the extent that they can bridge over in the vicinity of the resonance frequency, so that the spectrum of the raw signal gives no information at all about the repetition period. The slip may only be a fraction of a percent, but for example in the case of slow speed machines such as paper machines, where shaft speeds are a few Hz, and resonances excited are several kHz, the spectrum appears completely noise-like.

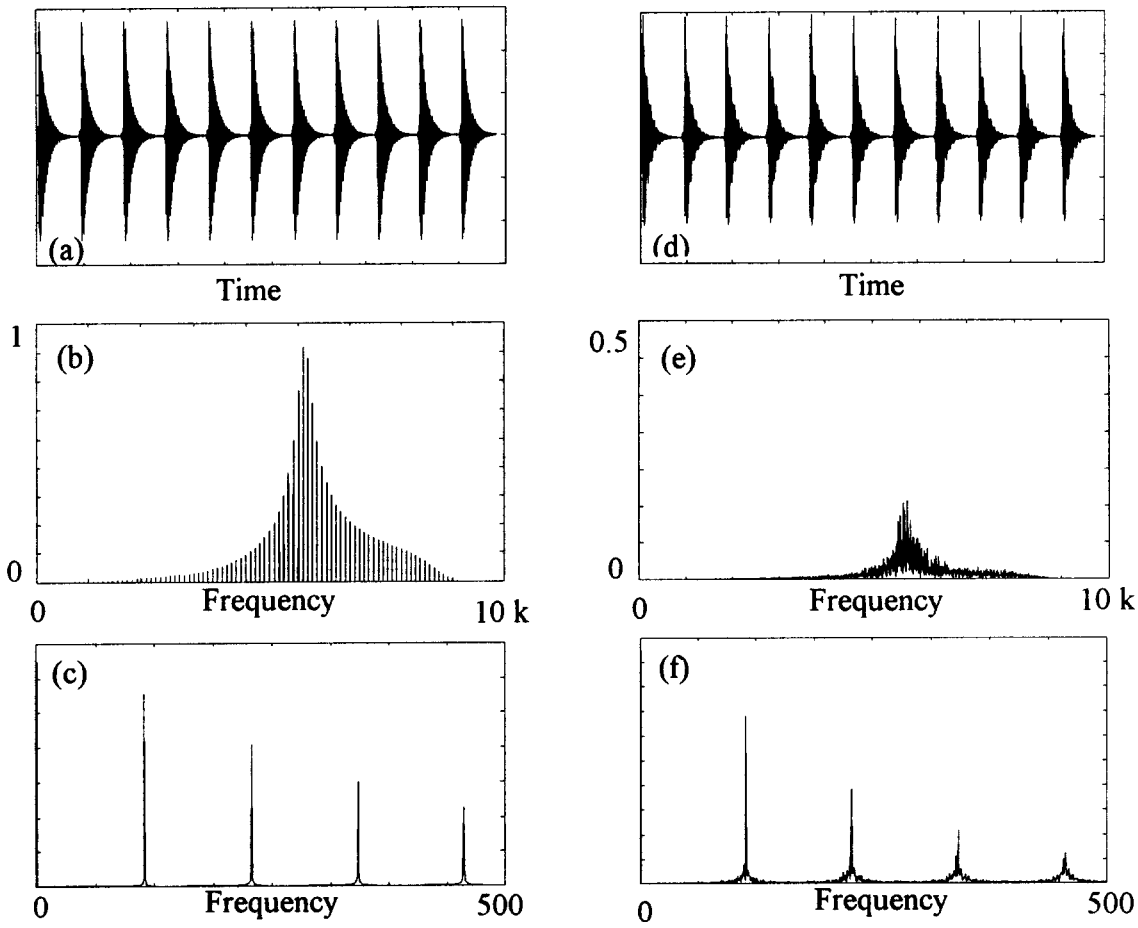

Figure 1 Simulated bearing fault signals with and without random spacing variation.
(a,b,c) Without spacing variation. (d,e,f) With spacing variation.
(a,d) Original time signals. (b,e) Spectra of raw time signals.
(c,f) Envelope spectra (note relative change in frequency scales).

Use is made below of this property to separate bearing from gear components, the latter being phase-locked to shaft speeds.

Figure 1(c, f) illustrate that by frequency analysing the amplitude demodulated bearing signal (the so-called "envelope signal") the burst repetition frequencies become apparent, even with spacing variations. For this reason, envelope analysis has become one of the most powerful techniques used in the diagnostics of rolling element bearings, but in most cases the envelope signal is obtained by adding analogue hardware, even when the envelope signal is to be analysed digitally. The envelope signal can be formed by bandpass or highpass filtering the original signal to remove masking signals, and then rectifying and smoothing it, typically with an RC smoothing circuit. This smoothing operation is not really required, as it only removes high frequency components which are automatically excluded in the final envelope analysis in a lower frequency range, and it can limit the rate at which the envelope signal can decay between pulses. This paper shows how all these operations can better be carried out by digital techniques in an FFT analyser or by using a signal analysis package such as Matlab®.

2.1 Amplitude Demodulation

The signal could of course be rectified as in the analogue case, by making the sign of all terms positive, but this is only efficient if the signal is not to be bandpass filtered. In the latter case there are two ways in which the bandpass filtration and amplitude demodulation can be carried out, both giving a reduction in sampling frequency appropriate to the frequency range of the modulation signal, and thus of the record size for a given record length in seconds. Both make use of the fact that a signal with a one-sided spectrum, a so-called "analytic signal", is complex, but its imaginary part is the Hilbert transform of the real part [3]. An analytic signal can also be expressed in the form $A(t)e^{j\phi(t)}$ where $A(t)$ is an amplitude modulating function (plus DC offset so that it is always positive), and $e^{j\phi(t)}$ is a rotating unit vector with a modulated phase function, representing fluctuations around a carrier frequency in the centre of the band. Thus $\phi(t)$ can be more specifically represented as $2\pi f_c t + \phi_m(t)$, where $\phi_m(t)$ is the phase modulation of the carrier frequency f_c. If the bandpass filtered analytic signal is frequency shifted by an amount f_k, it becomes $A(t)e^{j(\phi(t)-2\pi f_k t)}$, which still has the same amplitude function $A(t)$ and thus the same envelope. If f_k is made equal to f_c the phase will be equal to the modulating function $\phi_m(t)$ but this is mainly of interest for phase demodulation as discussed below.

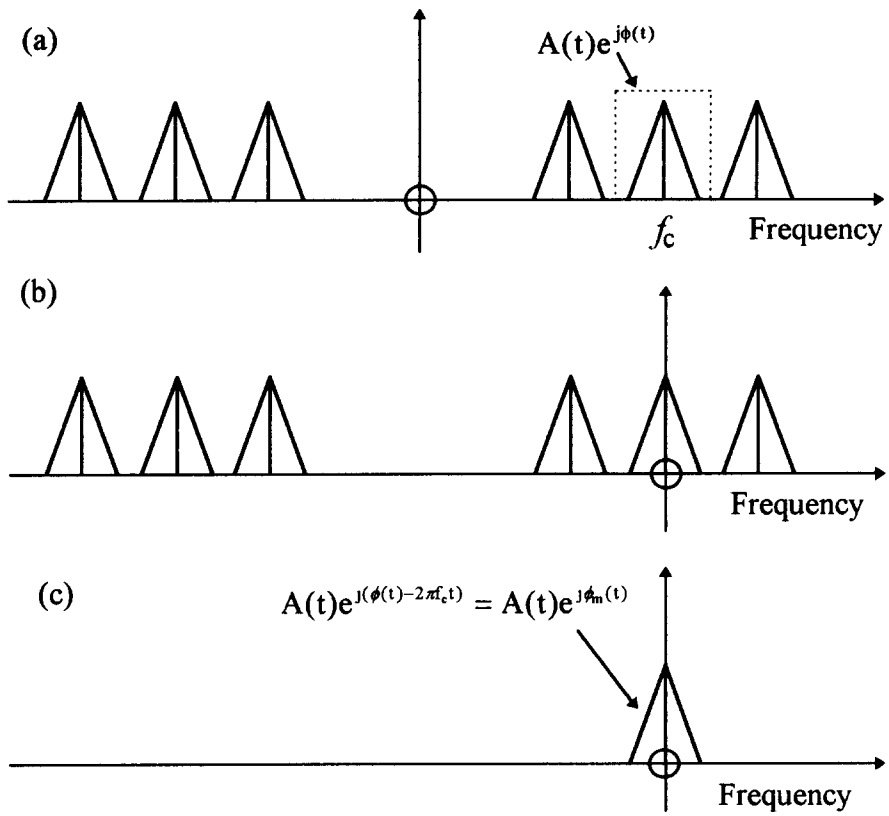

Figure 2 Effect of zoom processor. (a) Original spectrum
(b) Frequency shift to zoom centre frequency f_c
(c) Spectrum after lowpass filtration.

The first demodulation method can be used in conjunction with FFT analysers which have a real-time zoom processor. In these, the signal is first frequency shifted in the time domain (by multiplication by $e^{-j2\pi f_k t}$), and then lowpass filtered around the new zero frequency corresponding to the zoom band [3]. Figure 2 illustrates how this results in a complex time signal whose magnitude is the same as that of the analytic signal which would be formed by bandpass filtering the resonance to be demodulated, and taking only the positive frequency part. It will be seen below that this has the same amplitude or envelope signal as the real bandpass filtered signal, except for a possible scaling factor. The amplitude of the zoom processor output signal can thus be re-analysed in a lower frequency band (corresponding to the lower sampling frequency) to produce the required envelope spectrum. If the FFT analyser does not carry out this procedure internally, it is often possible to read the complex zoom processor output signal into a connected PC, and read its magnitude in again as a real signal to perform the envelope analysis. It should be noted that the lowpass filtration is done with filters of antialiasing quality, (typ. 120 dB/octave) thus allowing bearing signals to be separated from adjacent masking components which can be much larger. This is not the case normally with analogue envelopers.

Figure 3 shows an example from a faulty paper mill bearing analysed in this way using a Bruel & Kjaer Analyser type 2035, which performs both the demodulation and the envelope analysis. Figure 3(a) is a comparison of baseband spectra with and without the fault, and shows that the major change is at 5.4 kHz, even though the shaft speed is less than 2 Hz, and ballpass

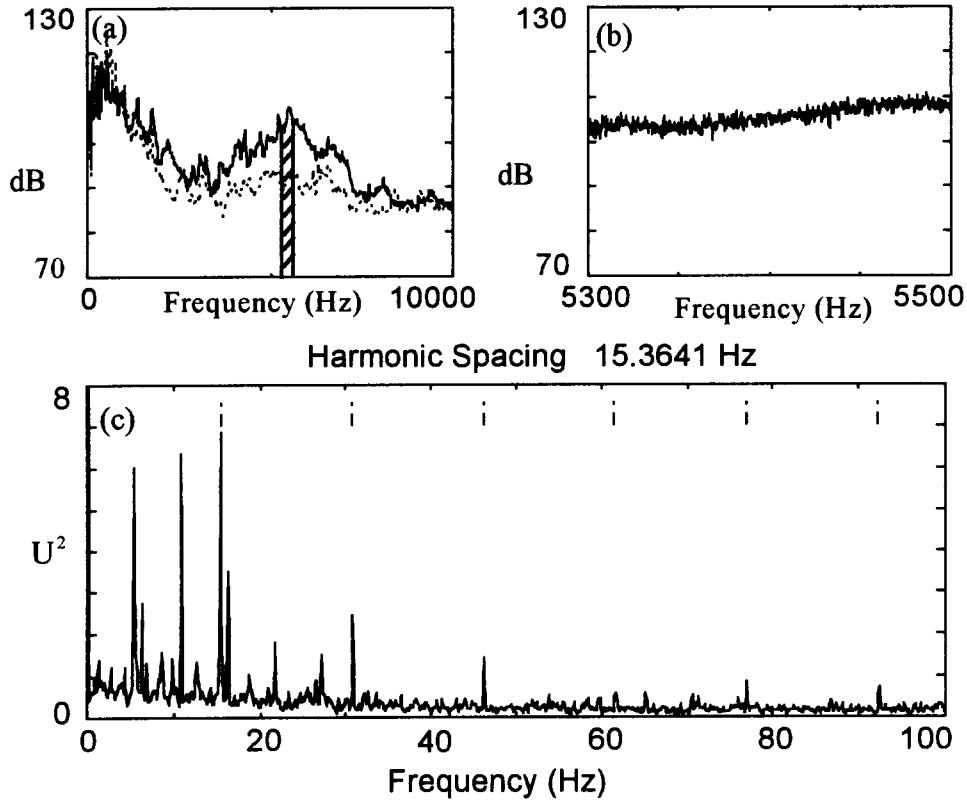

Figure 3 Envelope analysis by the zoom FFT method for a paper mill bearing.
(a) Comparison of baseband spectra with and without the fault.
(b) Zoom analysis in the demodulation band indicated in (a).
(c) Envelope (power) spectrum showing harmonics of BPFO.

frequency, outer race (BPFO) 15.5 Hz. Figure 3(b) shows the autospectrum of the zoom band centred on 5.4 kHz which was demodulated to obtain the envelope spectrum in Fig. 3(c). The zoom spectrum is a typical case where the high order harmonics have merged so that the bearing fault spectrum appears like white noise. On the other hand, the envelope spectrum clearly shows the harmonics of BPFO, even in the presence of pulses from a pneumatic lubricator (at about 5.5 Hz) which were present even with the bearing in good condition.

When a digital signal processing package is used, basically the same result can be achieved by a different procedure. An initial large transform is normally required to encompass the high resonance frequencies to be demodulated, at the same time as providing sufficient resolution in the spectrum to resolve the modulation sidebands (ie a length equal to several periods of the modulation sidebands, the lowest frequency typically being cage speed, or 40% of shaft speed). As illustrated in Fig. 4, the band to be demodulated is then extracted into a smaller buffer and inverse transformed to a complex time signal. Even though the frequency shift and bandpass filtering have been done in the frequency domain, the finally resulting complex time signal and its envelope are very similar to what is produced by the zoom technique. The separation in the frequency domain is perhaps even better, depending on what window is used for the initial transform, but even the common Hanning window gives complete separation over a wide dynamic range in a small number of lines. In Fig. 4(b) the situation is illustrated where the

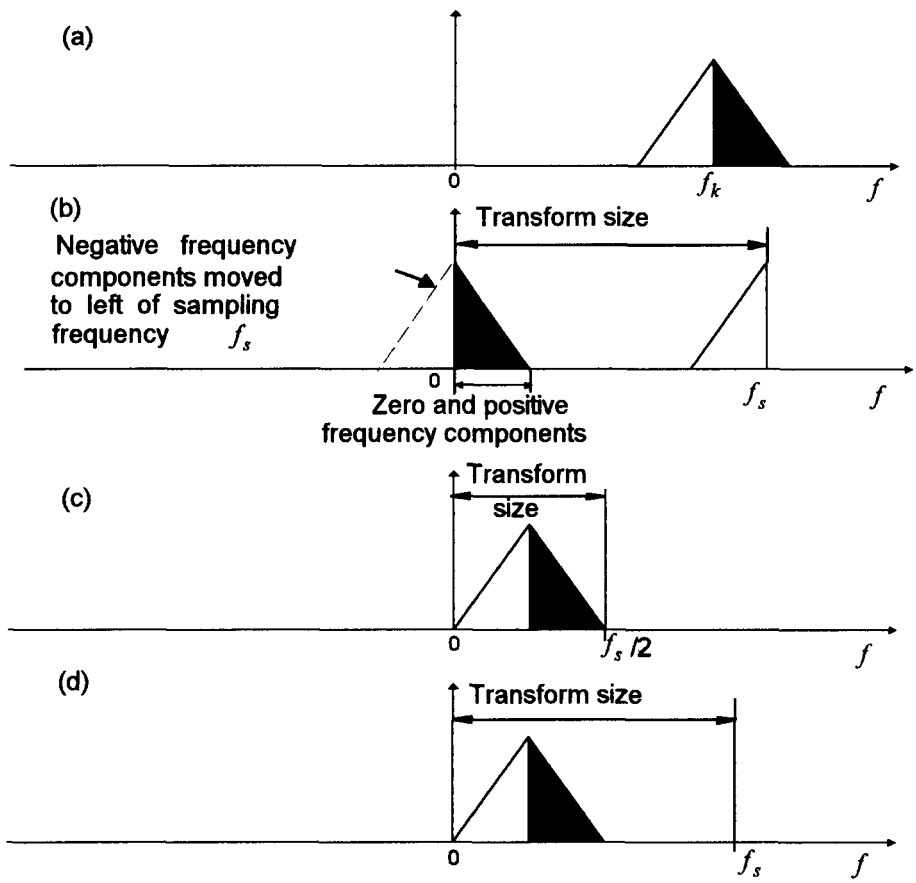

Figure 4 Block shift procedure for selection of frequency band for inverse transformation
(a) Original spectrum of one-sided bandpass section
(b) Frequency shift by f_k (centre of passband)
(c) Frequency shift by amount corresponding to lower passband limit (half size transform)
(d) Frequency shift by amount corresponding to lower passband limit (full size transform)

centre of the demodulation band is shifted to zero frequency (as for the zoom technique). Though this is necessary for phase demodulation, it has been shown above that the envelope signal is unaffected by the amount of frequency shift, and therefore the procedures of Fig. 4(c) and (d) may be used, these being much simpler. It is shown below that forming the (squared) envelope of the complex time signal corresponds in the frequency domain to a convolution of the spectrum with its complex conjugate. Using FFT procedures, the convolutions are circular, and to avoid any chance of wraparound effects it is best to zero pad the spectral segment up to double its size, as in Fig. 4(d). This means that the transform size and sampling frequency are double what they otherwise would be, though it does not change the resolution of the final envelope analysis. In most cases, however, a resonance peak is demodulated, and the concentration of the spectral information in the centre of the band means that wraparound effects are negligible and the procedure of Fig. 4(c) can be used. Analysis using this procedure on the same paper mill signal as Fig. 3 is included below in Fig. 6.

2.2 Masking by Extraneous Signals

In a study published elsewhere [4], the masking effects of extraneous random and discrete frequency signals were studied in detail. The results of the mathematical analysis can be explained graphically as in Fig. 5. Figure 5(a) shows the (amplitude) spectrum of an analytic signal $f(t)$ which contains modulated bearing signals at two carrier frequencies, and thus having sidebands with a spacing equal to the bearing fault frequencies. It also contains a section of narrow band masking noise, and a single discrete frequency masking tone (though in practice there would normally be several of each). The squared envelope of $f(t)$ is given by $f(t)\cdot f^*(t)$ whose spectrum is given by the convolution theorem as the convolution of $F(f)$ with $F^*(-f)$, the Fourier transforms of $f(t)$ and $f^*(t)$, respectively. In this convolution operation, the second term is first wrapped around so that the complex conjugate frequency components line up with their positive frequency counterparts, and are then multiplied and

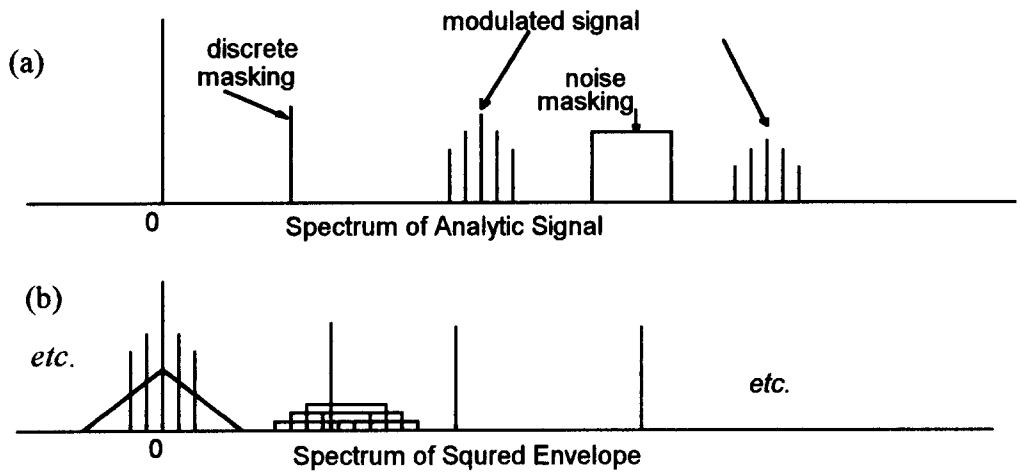

Figure 5 Generation of the spectrum of the squared envelope of analytic signal $f(t)$ with Fourier spectrum $F(f)$. (a) Amplitude spectrum of $f(t)$ containing two modulated components, a discrete tone and a narrow band noise. (b) The result of the convolution of $F(f)$ with $F^*(-f)$ with components at all difference frequencies and a triangular noise spectrum around zero frequency.

integrated for each frequency displacement of one spectrum with respect to the other. Thus at zero frequency, the result is a sum of squared amplitudes, and is proportional to the mean square value of the signal. Each time the displacement is such that two discrete frequency components line up, a component will be found in the spectrum of the squared envelope signal, and thus at each of the sideband spacings corresponding to the modulating frequency. Discrete frequency masking signals will only mask the result if they have a small spacing in the original spectrum from some other tone; most difference frequency terms will be at higher frequency than the bearing frequencies. Narrow band noise terms, however, wherever they are located in the spectrum, will add to the triangular noise spectrum illustrated around zero frequency, and this provides the dominant masking of the bearing frequencies. Note that the bearing frequency components are always present additively in the spectrum, even when they may appear masked in the original time signal.

What is the difference between the envelope signal and its squared value of which we have just discussed the spectrum? Both will fluctuate with the same basic periodicity, although as shown in [4] the relative level of harmonics will vary somewhat. This would not affect the diagnostic capability significantly. Since the various components are present in the spectrum of the envelope squared signal according to the square of their values in the original signal, the signal to noise ratio of bearing to extraneous components is approximately the square of its value in the original signal. In other words, provided it can be made greater than unity, there will be an advantage in analysing the squared envelope signal rather than the envelope directly.

Fig. 6 illustrates a number of the above points. Figure 6(a) shows the envelope spectrum of the rectified signal, which despite considerable masking still contains the harmonics of BPFO at about the same level as in Fig. 6(b), where the masking has been considerably reduced by bandpass filtering in the frequency range 5.4-5.6 kHz, where the bearing fault gave the biggest change. Note the similarity of this result to that of Fig. 3(c). This bandpass filtration gives such an improvement in signal to noise ratio that a considerable advantage is given by analysing the squared envelope signal as shown in Fig. 6(c). The BPFO components now dominate completely over those from the lubricator.

2.3 Self Adaptive Noise Cancellation

Where it is not possible to find a part of the spectrum where the bearing signal dominates completely, so that the signal to noise ratio cannot be made greater than unity simply by bandpass filtration, this will often be because of masking by gear noise, comprising discrete frequency sidebands spaced around gearmesh harmonics. In such a case, use can often be made of so-called "self adaptive noise cancellation" to remove or reduce the gear signals and thus improve the signal to noise ratio of the bearing signals.

Self adaptive noise cancellation (SANC) is a further development of adaptive noise cancellation (ANC) [5], for which it is necessary to have two signals, one (the primary signal) containing two components to be separated, and the other (the reference signal) coherent with one component only. An adaptive filter is caused to act on the reference signal, adapting its linear transfer function in such a way as to make the filter output most similar to the coherent component in the primary signal. The other component can then be obtained by subtraction. Adaptive noise cancellation has been applied to bearing diagnostics in [6, 7]. Self adaptive noise cancellation can be used when there is only a primary signal available, but where the two

components have different characteristics, in particular where one is made up of discrete frequency components and the other random. As shown above, the degree of randomness in normal bearing signals satisfies this requirement with respect to gear signals, which are locked to shaft speeds. Even if there is some speed fluctuation, this can be removed by "tracking", which involves resampling the signals at equal shaft angle increments, as determined by

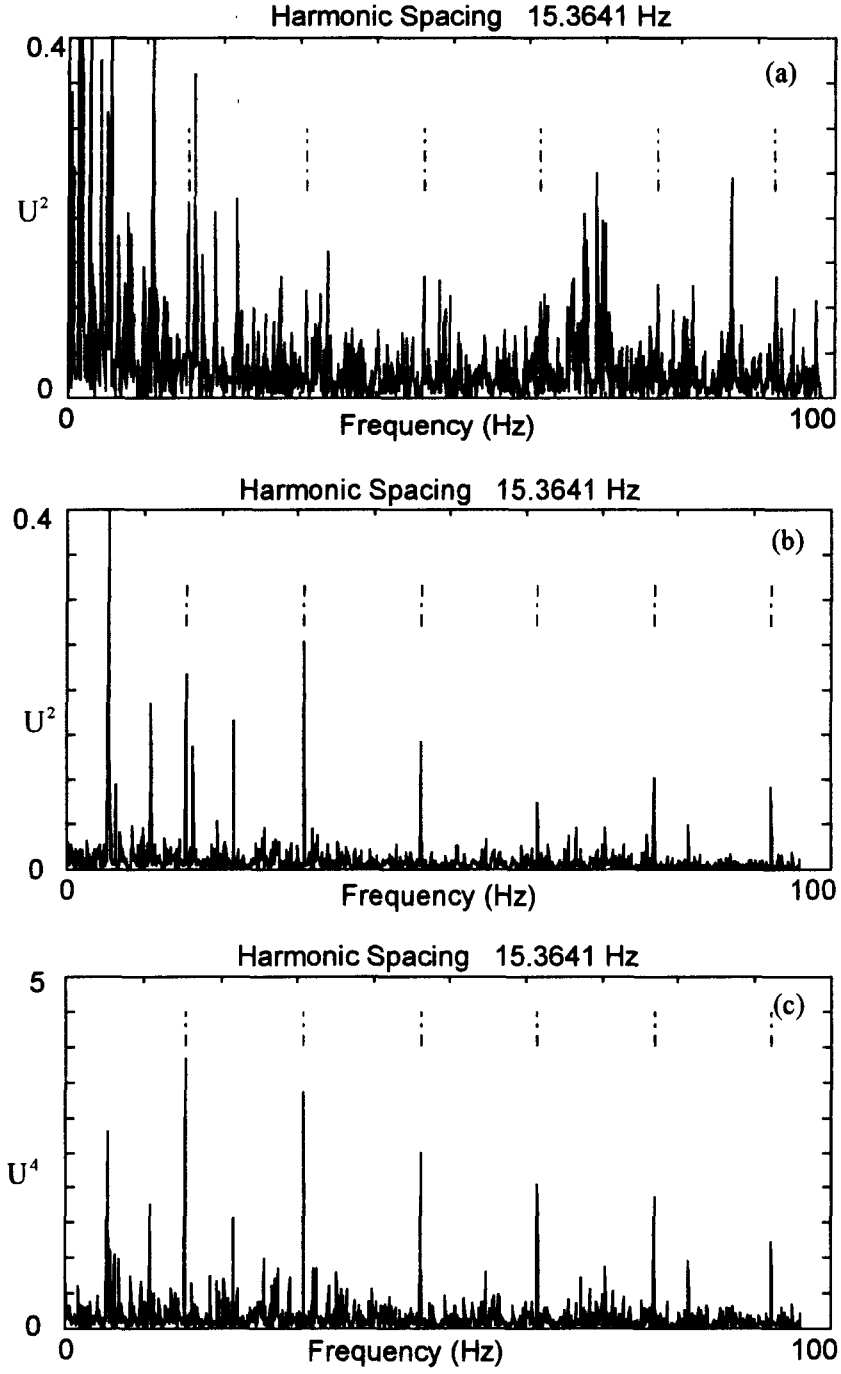

Figure 6 Envelope spectra from the paper mill bearing of Fig. 3, obtained using digital signal processing, with harmonics of BPFO indicated by a special cursor.
(a) (Power) spectrum obtained by rectification of full bandwidth signal.
(b) (Power) spectrum obtained by block shift of band from 5.2 - 5.6 kHz.
(c) (Power) spectrum of squared envelope of same signal as in (b).

analysing a shaft tachometer signal at the same time. Tracking is built into some analysers, but can also be done by interpolation routines in for example Matlab®. As explained in [8], the reference signal can then be made a delayed version of the primary signal, so that the gear component is still correlated, but the random bearing component uncorrelated. Ref.[8] gives details of the correct choice of parameters to use in optimal SANC.

Figure 7 shows an example of the benefits given by SANC in diagnosing inner and outer race faults in a planet gear bearing in a helicopter gearbox. This compares envelope spectra both before and after the application of SANC to a signal bandpass filtered in a band where the maximum change in level occurred, but where this was only 6 dB because of masking by gear noise. The result after SANC enhances the bearing components greatly, both outer race (BPFO) and inner race (BPFI). No tacho signal was available for tracking, and neither could use be made of another procedure recommended in [9] to improve signal/noise ratio by special windowing of the time signals to extract the part of the signal corresponding to the passage of just the faulty planet bearing.

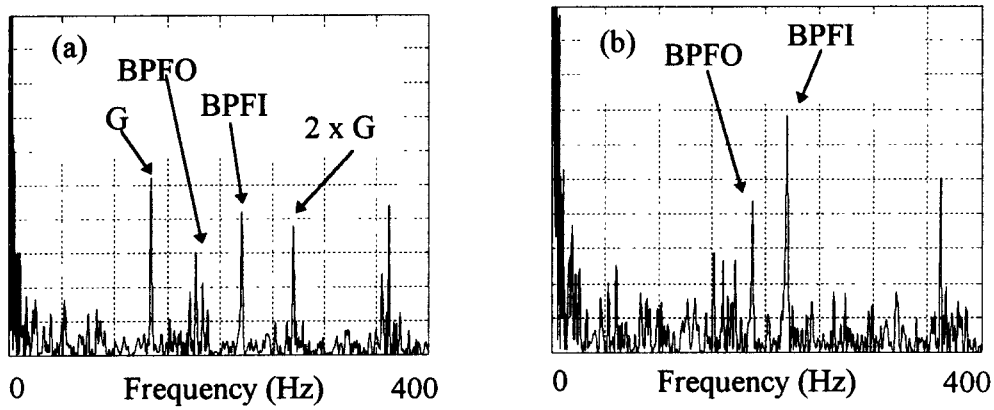

Figure 7 Use of SANC to remove masking by gear noise. G is a gear related masking component and 2 x G its second harmonic. BPFO and BPFI defined in text.
(a) Envelope spectrum obtained by optimum bandpass filtration in range 13-16 kHz.
(b) Spectrum of squared envelope signal after SANC.

2.4 Envelope Analysis of Short Records

In some cases, the record length of signal available for analysis is limited by extraneous factors. A typical case is that mentioned above of planet gear bearings in epicyclic gearboxes, where the time is limited by the time of passage of one bearing past an externally mounted transducer. Another is in the analysis of rail vehicle bearings using track mounted accelerometers. Figure 8 shows a result from [10] where it was shown that considerably better resolution could be obtained with a short record by using the maximum entropy technique [11], although almost as much benefit could be obtained by the simple expedient of spectrum interpolation by massive zero padding of the time signal.

3. GEAR DIAGNOSTICS

Some years ago, the author proposed a model for interpreting gear vibrations as a modulation (both amplitude and phase modulation) of an otherwise uniform gear mesh signal [12]. The

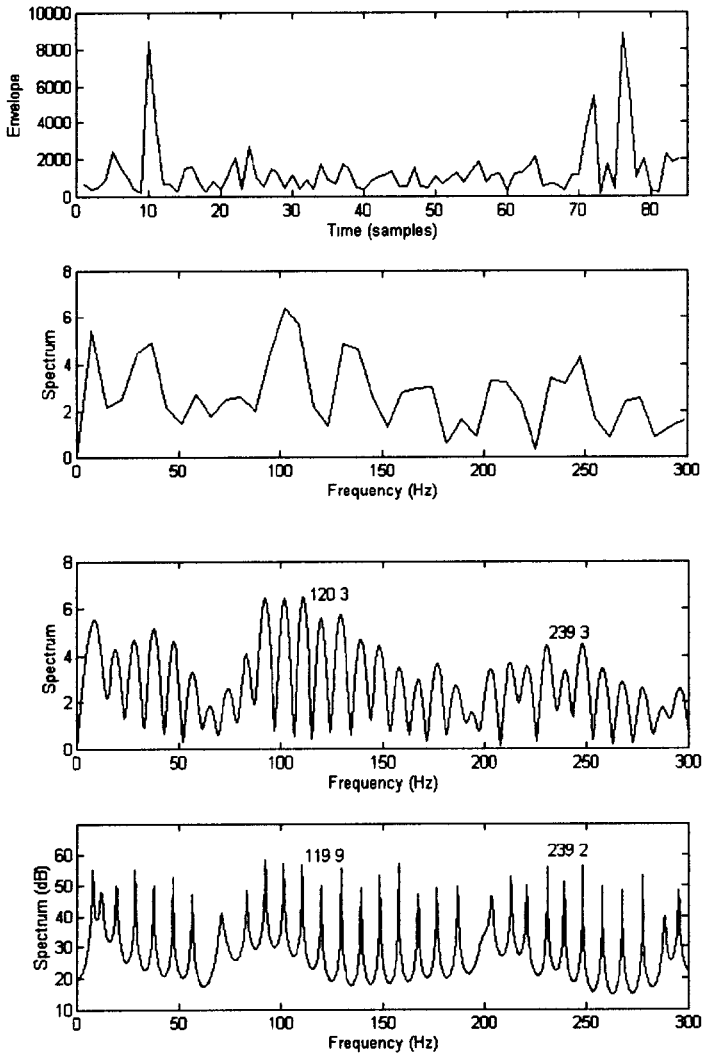

Figure 8 Envelope spectra of short time recording (1.29 revolutions) just spanning the first two sections where the inner race fault passes the loading zone.
(a) Envelope signal after bandpass filtration in frequency range 2.7 ~ 3.3 kHz,
(b) Envelope spectrum without interpolation, (c) Envelope spectrum with interpolation, (d) Maximum entropy envelope spectrum

basis of the amplitude modulation concept is that the gear transmission error, which is generally acknowledged as the fundamental source of gear vibration and noise, has some components due to geometric profile errors, but others due to tooth deflection under load, which results in a modulation effect with load. The phase modulation was initially thought of somewhat naively as due to speed fluctuations, but has later been shown to be orders of magnitude greater than the torsional vibration of the gears themselves. Mark [13, 14] was one of the first to make a detailed mathematical model of the generation of gear vibrations from transmission error, and his model can be interpreted as an amplitude modulation of the mean tooth mesh signal by functions having the periodicity of rotation of the two gears in mesh. Mark himself preferred to explain the resulting sidebands as a "DFT effect", as errors associated with each tooth were decomposed into elemental errors of different order such as spacing, slope, curvature etc, for which there was only one value per tooth, and so the resulting spectrum of each term had a period equal to the toothmesh frequency, which was the

"sampling frequency". Consequently, low harmonics tend to be reflected around the toothmesh frequency and its harmonics, giving the appearance of sidebands. The fact that the actual spectrum is not periodic is explained by its subsequent modification by "mesh transfer functions", which have the properties of lowpass filters associated with the contact ratio, and with double the rolloff rate at high frequencies for helical gears. This is presumably just another way of interpreting the same phenomena.

As a result of either of these models, effects such as uniform (or in fact the mean) wear show up at the tooth mesh frequency and its harmonics, while non-uniformly distributed faults show up at other harmonics of the two shaft speeds, in particular as sidebands around the gearmesh frequency and its harmonics. Distributed effects such as distortion due to heat treatment, and non uniformly distributed wear would tend to show up in a small number of strong sidebands, while local effects such as spalls and cracked teeth would be more impulsive and thus show up as more uniformly distributed sidebands throughout the spectrum.

3.1 Toothmesh Demodulation

Even though such local effects are quite often detectable in the spectrum and cepstrum [15], other cases have been reported where local defects were not apparent in the amplitude modulation of the toothmesh signal [16]. In Ref. [16], however, McFadden showed that by demodulating the toothmesh signals for both amplitude and phase, a crack could show up more definitely in one than the other, and in the case reported was initially clearer in the phase than the amplitude modulation signal, as shown in Fig. 9 (from Ref. [16]). It should be noted that to obtain a result such as that in Fig. 9, it is necessary to carry out the following steps, all of which can be achieved with a signal processing package (although perhaps easier and faster in an analyser which incorporates them).

1) Signal resampling using the tracking techniques mentioned above. This is to ensure that the digitised signals are phaselocked with the shaft rotation, as even one sample spacing represents 360° of phase shift at the sampling frequency and thus 144° at 40% of it, which is normally the maximum frequency being analysed.
2) Synchronous averaging synchronised with the rotation speed (and starting phase) of the particular gear being analysed. This can be repeated for each gear in the train.
3) Amplitude and phase demodulation of the toothmesh signal (or one of its harmonics) by either the zoom demodulation technique, as illustrated in Fig. 2, or the block shift technique illustrated in Fig. 4(b).

In Ref. [16] the signal demodulated was an externally measured vibration acceleration signal, but as pointed out in [17], this will not necessarily represent the situation at the tooth mesh source, viz, the transmission error signal. The reason is that the difference between amplitude and phase modulation, in particular for the very low modulation indices represented here, lies primarily in the phase relationships of sidebands on either side of the carrier frequency component. These can be modified considerably by the transfer function which exists between the source function and external measurement points, and thus it should be kept in mind that the result obtained by demodulating such externally measured signals does not necessarily reflect directly what is occurring at the source. There are two ways of counteracting this difficulty, one of them not yet fully developed. The first is simply to measure the transmission error directly, rather than (or as well as) the vibration signals. As explained in [18], this can be

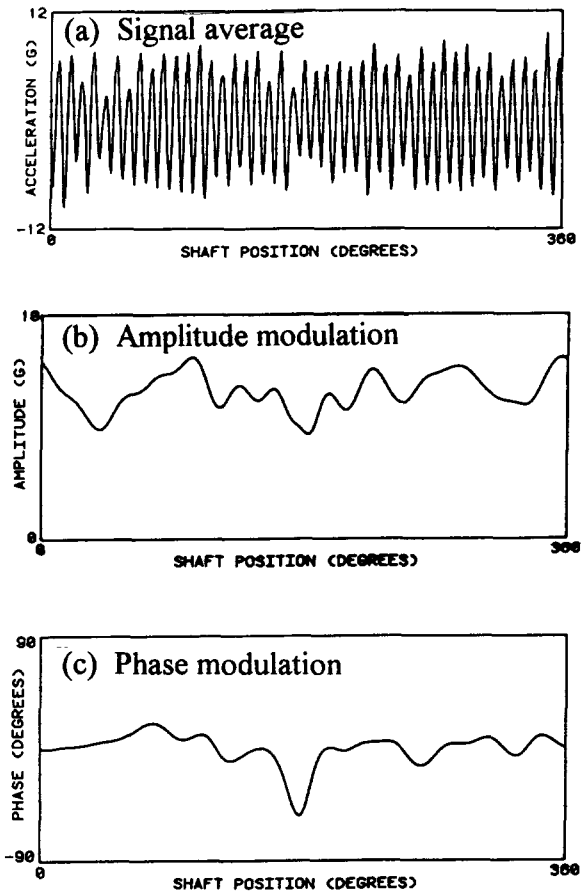

Figure 9 Detection of a fatigue crack in a gear using amplitude and phase demodulation of the tooth mesh signal. From [16].
(a) Signal average (b) Amplitude modulation signal (c) Phase modulation signal

done by phase demodulating the signals from shaft encoders mounted on each shaft involved in the mesh, as the transmission error can be interpreted as the differential torsional vibration of the two gears in mesh, with proper compensation for the gear ratio. However, this relies on being able to mount shaft encoders on both shafts, and this is not always practicable. A potential future approach, though this has not yet been demonstrated for gears, is to use inverse filtration to counter the effects of the transfer function from the source to the external acceleration signals, such as has been done in the recovery of diesel engine cylinder pressure signals from external measurements [19, 20].

3.2 Time/Frequency Methods

In recent years, a number of researchers have attempted to combine the advantages of analysis in the time and frequency domains by using a combined time/frequency approach. The basic problem is associated with the "uncertainty principle", which states that the product $\Delta f \times \Delta t = a$ constant, so that to obtain better resolution in frequency involves poorer resolution in time and vice versa.

The so-called Wigner-Ville distribution, developed originally by Wigner [21], but modified by Ville [22] to use analytic time signals, so as to avoid interference between positive and

negative frequency components, appears to give better resolution than predicted by the uncertainty principle, but suffers from interference terms located between the "true" components, which are thus difficult to distinguish. As summarised in [23], however, a number of workers have suggested modifications of the basic Wigner-Ville distribution which reduce interference terms by smoothing in both time and frequency directions, but often achieve better simultaneous resolution in time and frequency than the so-called "short time Fourier transform" where the frequency resolution is basically given by the inverse of the length of a time window which is moved in overlapping steps along the record, and the resulting spectra plotted in 3-dimensional "waterfall" form, or as contour plots. The reader is directed to Refs. [24-26] for details of a number of studies of the potential use of Wigner-Ville methods for gear diagnostics.

An alternative time/frequency method which has been applied to gear diagnostics is wavelet analysis [27], where the advantage is not so much that it gives better resolution than dictated by the uncertainty principle, but that the time resolution is better at higher frequencies because of the constant percentage bandwidth character of the resulting filters. The real advantages of wavelets are connected with their analysis/synthesis capabilities, so it is difficult to see what advantage they give in a purely analysis situation over classical 1/n-th octave analysis. In [28], McFadden and co-workers proposed a novel display technique for wavelet analysis of gear signals, a polar diagram with frequency in the radial direction vs rotational angle of the gear. This is a natural way of utilising the greater resolution available at higher frequencies. The reader is directed to Refs. [28-30] for details of a number of studies of the potential use of wavelet analysis for gear diagnostics.

4 SIMULATION TECHNIQUES

A number of researchers are attempting to apply neural networks to the automation of machine diagnostics [30, 31], but the problem is that the networks must be trained to recognise the faults which are to be detected and diagnosed, and it would not be economically viable to experience all the failures required to train them in the normal way. Thus it appears that the only effective way to train them is by using mathematical models to simulate the action of the machines being monitored, and in particular the effects of a wide range of typical faults.

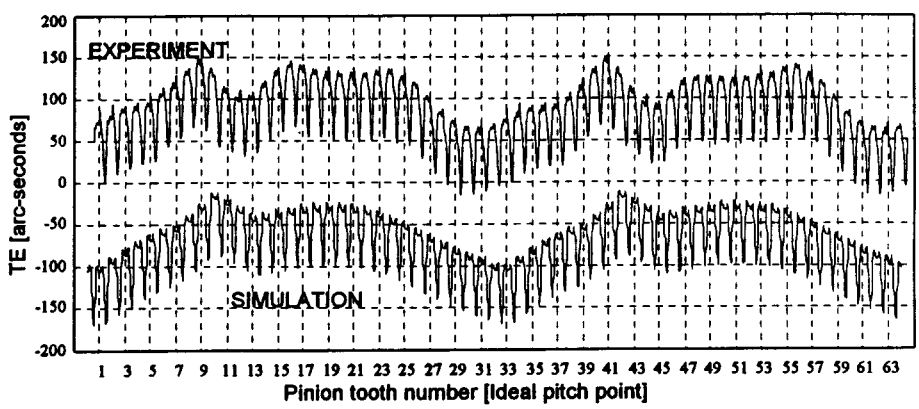

Figure 10 Comparison of measured vs simulated transmission error for a spur gear set, based on measured geometric deviations and estimated tooth pair stiffness, for a given load and alignment. From [32].

A method of simulating bearing faults was outlined in Section 2, and this was in fact used in [8] to make recommendations for the choice of suitable analysis parameters for SANC. It could presumably also be used for training neural networks.

The simulation of gear vibrations is also well developed, and is improving continuously. Figure 10 from [32] shows how closely it was possible to predict the transmission error (TE) of a spur gear from measurements of its profile errors, and an estimate of tooth pair stiffness. These were automotive quality gears where the TE is dominated by the geometric errors, but a more recent work [33] using finite element techniques to take into account the variation of tooth compliance with contact point gives a much better agreement for nylon gears, and would thus presumably also give an improved simulation for precision gears where the tooth deflection effects also dominate over geometric errors.

5 CONCLUSION

The availability of simple means of recording and digitising vibration signals, as well as analysing them using powerful signal processing packages on increasingly more powerful PCs means that users are no longer limited by what analyser manufacturers provide. The paper illustrates some current possibilities in bearing and gear analysis.

ACKNOWLEDGEMENTS

The author would like to thank Yujin Gao and Dominique Ho for providing figures used in the paper. Much of the research work presented here was supported by Hamersley Iron, Dampier, Aeronautical and Maritime Research Laboratories, Melbourne, and BTR Engineering, Sydney.

REFERENCES

[1] MATLAB® (Windows version 5.1, 1997), The MathWorks Inc. Natick MA, USA.

[2] McFadden, P. and Smith, D. (1984) "Model for the vibration produced by a single point defect in a rolling element bearing", *Journal of Sound and Vibration*, **96**(1), pp. 69-82.

[3] Randall, R.B. (1987), *Frequency Analysis*, 3rd Edition, Bruel and Kjaer, Copenhagen.

[4] Randall, R.B and Gao, Y. (1996), "Masking Effects in Digital Envelope Analysis of Faulty Bearing Signals", *6th Int. Conf. on Vib. in Rot. Machinery*, IMechE, Oxford,pp351-359.

[5] Widrow, B., et.al, (1985) *Adaptive Signal Processing*, Prentice-Hall Inc., Englewood Cliffs, N.J. 07632.

[6] Chaturvedi, G.K. and Thomas, D.W., (1982) "Bearing Fault Detection Using Adaptive Noise Cancelling", *ASME, J. of Mech. Design*, Vol.**104**, Apr. '82, pp.280-283.

[7] Tan, C.C., (1987) "An Adaptive Noise Cancellation Approach for Condition Monitoring of Gear Box Bearings", *Inter. Tribology Conf. 1987*, Melbourne, 2-4 Dec. '87, pp.360-365.

[8] Ho, D. and Randall, R.B. (1997) "Effects of Time Delay, Order of FIR Filter and Convergence Factor on Self Adaptive Noise Cancellation". ICSV'5, Adelaide, Dec., '97.

[9] Randall, R.B. and Li, Y.J., (1995) "Modified Envelope Analysis for Diagnostics of Planetary Gear Bearings". *Machine Vibration*,.Vol.**3**, pp185-191, Springer, London.

[10] Randall, R. B., Gao, Y. and Ford, R. "Envelope Analysis of Short Records for Bearing Diagnostics," IEAust, Vibration and Noise Symposium, Newcastle, October, 1995.

[11] Burg, J. P. "Maximum Entropy Spectral Analysis," Ph.D. Dissertation, Stanford University, Stanford, CA, 1975.

[12] Randall, R.B. (1982) "A New Method of Modeling Gear Faults". *ASME J.Mech.Design*, Vol.**104**, pp 259-267.

[13] Mark, W.D. (1978), "Analysis of the Vibratory Excitation of Gear Systems: Basic Theory", *J. Acoust. Soc. Am.*, Vol. **63**, p1409-1430.

[14] Mark, W.D. (1979), "Analysis of the Vibratory Excitation of Gear Systems: II: Tooth Error Representations, Approximations, and Application", *J. Acoust. Soc. Am.*, Vol. **66**, p1758-1787.

[15] Randall, R.B., (1997) "Advanced Machine Diagnostics", *Shock & Vibration Digest*, Vol. **29,** No.1, Jan/Feb '97, pp 6-30.

[16] McFadden, P.D. (1986), "Detecting Fatigue Cracks in Gears by Amplitude and Phase Demodulation of the Meshing Vibration", *Journal of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. **108**, pp165-170.

[17] Randall , R.B, and Sidahmed, M., (1995) "Gear Transmission Error Measurement as a Diagnostic Tool". *2nd Int. Conf. on Acoustical and Vibratory Surveillance Methods and Diagnostic Techniques*. CETIM, Senlis, France, Oct. '95.

[18] Sweeney, P.J. and Randall, R.B. (1996), "Gear Transmission Error Measurement Using Phase Demodulation", *Proc Instn Mech Engrs, Part C: Journal of Mechanical Engineering Science*, Vol. **210**, No. 3, p201-213.

[19] Lyon, R. H. and Kim, J.T., (1988) "Reduced parameter set descriptions for system and event identification." *Robotics and Computer-Integrated Manufacturing* **4**(3/4), 447-455..

[20] Cassini, G. D., D'Ambrogio, W. andSestieri, A. (1996) "Frequency domain vs. cepstrum technique for machine diagnostics and input waveform reconstruction." *Proceedings of 21st International Seminar on Modal Analysis*, K. U. Leuven, Belgium, 835-846..

[21] Wigner, E.P., (1932) "On the Quantum Correction for Thermodynamic Equilibrium." *Physical Review*, **40**, pp749-759.

[22] Ville, J., (1948) "Theorie et Applications de la Notion de Signal Analytique," *Cables et Transmission*, **2 A**, pp61-74.

[23] Boashash, B., (ed.) *Time-Frequency Signal Analysis: Methods and Applications*, Longman Cheshire Halsted Press, 1992.

[24] Forrester, D., "Time-Frequency Analysis in Machine Fault Detection," in *Time-Frequency Signal Analysis: Methods and Applications*, ed Boashash, B., Longman Cheshire Halsted Press, 1992, pp406-444.

[25] Wang, W.J. and McFadden, P.D. (1993) "Early Detection of Gear Failure by Vibration Analysis. 1. Calculation of the Time-Frequency Distribution." *Mech. Systems & Signal Processing* 7(3):193-203, May.

[26] Oehlmann, H. et al, (1997), "A Method for Analysing Gearbox Faults using Time-Frequency Representations". *Mech. Systems & Signal Processing* **11**(4), pp529-545.

[27] Newland, D.E.,(1993) *An Introduction to Random Vibrations, Spectral and Wavelet Analysis.* 3rd Ed., Longmans Scientific, Harlow, UK

[28] Lin, S.T. and McFadden, P.D., (1995) "Vibration Analysis of Gearboxes by the Linear Wavelet Transform." *Second Int. Conf. on Gearbox Noise, Vibration and Diagnostics* IMechE, London

[29] Staszewski, W.J. and Tomlinson, G.R., (1994) "Application of the Wavelet Transform to Fault Detection in a Spur Gear." *Mech. Systems & Signal Processing* **8**(3):289-307, May.

[30] Paya, B.A., Esat, I.I., Badi, M.N.M., (1997) "Artificial Neural Network Based Fault Diagnostics of Rotating Machinery using Wavelet Transforms as a Preprocessor." *Mech. Systems & Signal Processing* **11**(5), September, pp751-765

[31] Zhang, S., Ganeson, R., Xistris, G.D., (1996) "Self-organising Neural Networks for Automated Machinery Monitoring Systems". *Mech. Systems and Signal Processing*, **10**(5), September, , pp517-532

[32] Sweeney, P.J. (1994), *Gear Transmission Error Measurement and Analysis*, PhD dissertation, University of New South Wales.

[33] Du, S., Randall, R.B. and Kelly, D.W. (1997), "Modelling of Spur Gear Mesh Stiffness and Static Transmission Error", submitted to *Proc Instn Mech Engrs, Part C: Journal of Mechanical Engineering Science*, Feb., 1997.