Chapter 4

REVIEW OF VIBRATION ANALYSIS TECHNIQUES

In this chapter, a review is made of some current vibration analysis techniques used for condition monitoring in geared transmission systems. The perceived advantages, disadvantages, and the role each of these techniques may play in the diagnosis of safety critical failure modes is discussed. A summary of the findings is then made to establish which techniques to pursue further, and to identify any deficiencies which need to be addressed.

4.1 TIME DOMAIN ANALYSIS

4.1.1 Waveform analysis

Prior to the commercial availability of spectral analysers, almost all vibration analysis was performed in the time domain. By studying the time domain waveform using equipment such as oscilloscopes, oscillographs, or 'vibrographs', it was often possible to detect changes in the vibration signature caused by faults. However, diagnosis of faults was a difficult task; relating a change to a particular component required the manual calculation of the repetition frequency based on the time difference observed between feature points.

4.1.2 Time domain signal metrics

Although detailed study of the time domain waveform is not generally used today, a number of simple signal metrics based on the time domain waveform still have widespread application in mechanical fault detection; the simplest of these being the peak and RMS value of the signal which are used for overall vibration level measurements.

4.1.2.1 Peak

The peak level of the signal is defined simply as half the difference between the maximum and minimum vibration levels:

$$peak = \tfrac{1}{2}\left(\max\left(x(t)\right) - \min\left(x(t)\right)\right) \tag{4.1}$$

4.1.2.2 RMS

The RMS (Root Mean Square) value of the signal is the normalised second statistical moment of the signal (standard deviation):

$$RMS = \sqrt{\tfrac{1}{T}\int_0^T \left(x(t) - \bar{x}\right)^2 dt} \tag{4.2}$$

where T is the length of the time record used for the RMS calculation and $\bar{x}$ is the mean value of the signal:

$$\bar{x} = \tfrac{1}{T}\int_0^T x(t)dt \tag{4.3}$$

For discrete (sampled) signals, the RMS of the signal is defined as:

$$\begin{aligned} RMS &= \sqrt{\tfrac{1}{N}\sum_{n=0}^{N-1}\left(x(n) - \bar{x}\right)^2} \\ \bar{x} &= \tfrac{1}{N}\sum_{n=0}^{N-1} x(n) \end{aligned} \tag{4.4}$$

The RMS of the signal is commonly used to describe the ‘steady-state’ or ‘continuous’ amplitude of a time varying signal.

4.1.2.3 Crest Factor

The crest factor is defined as the ratio of the peak value to the RMS of the signal:

$$Crest\ Factor = \frac{peak}{RMS} \tag{4.5}$$

The crest factor is often used as a measure of the ‘spikiness’ or impulsive nature of a signal. It will increase in the presence of discrete impulses which are larger in amplitude than the background signal but which do not occur frequently enough to significantly increase the RMS level of the signal.

4.1.2.4 Kurtosis

Kurtosis is the normalised fourth statistical moment of the signal. For continuous time signals this is defined as:

$$Kurtosis = \frac{\frac{1}{T}\int_0^T \left(x(t) - \bar{x}\right)^4 dt}{\left(RMS\right)^4} \tag{4.6}$$

For discrete signals the kurtosis is:

$$Kurtosis = \frac{\frac{1}{N}\sum_{n=0}^{N-1}\left(x(n) - \bar{x}\right)^4}{\left(RMS\right)^4} \tag{4.7}$$

The kurtosis level of a signal is used in a similar fashion to the crest factor, that is to provide a measure of the impulsive nature of the signal. Raising the signal to the fourth power effectively amplifies isolated peaks in the signal.

4.1.3 Overall vibration level

The most basic vibration monitoring technique is to measure the overall vibration level over a broad band of frequencies. The measured vibration level is trended against time as an indicator of deteriorating machine condition and/or compared against published vibration criteria for exceedences. The measurements are typically peak (4.1) or RMS (4.2) velocity recordings which can be easily made using a velocity transducer (or integrating accelerometer) and an RMS meter.

Because the peak level is not a statistical value, it is often not a reliable indicator of damage; spurious data caused by statistically insignificant noise may have a significant effect on the peak level. Because of this, the RMS level is generally preferred to the peak level in machine condition monitoring applications.

Trending of overall vibration level may indicate deteriorating condition in a simple machine, however it provides no diagnostic information and will not detect faults until they cause a significant increase in the overall vibration level. Localised faults in complex machinery may go undetected until significant secondary damage or catastrophic failure occurs.

4.1.4 Waveshape metrics

The overall vibration level provides no information on the wave form of the vibration signal. With a number of fault types, the shape of the signal is a better indicator of damage than the overall vibration level. For example, faults which produce short term impulses such as bearing faults and localised tooth faults, may not significantly alter the overall vibration level but may cause a statistically significant change in the shape of the signal.

Crest factor (4.5) or kurtosis (4.6) are often used as non-dimensional measures of the shape of the signal waveform. Both signal metrics increase in value as the ‘spikiness’ of the signal increases (i.e., as the signal changes from a regular continuous pattern to one containing isolated peaks). Kurtosis, being a purely statistical parameter, is usually preferable to crest factor in machine condition monitoring applications for the same reasons that RMS is preferable to peak. However, crest factor is in more widespread use because meters which record crest factor are more common (and more affordable) than kurtosis meters.

Because of the non-dimensional nature of the crest factor and kurtosis values, some assessment of the nature of a signal can be made without trend information. Both waveshape metrics will give a value of 0.0 for a DC signal and 1.0 for a square wave. For a pure sine wave, the crest factor will be $\sqrt{2}$ =1.414 and the kurtosis will be 1.5. For normally distributed random noise, the kurtosis will be 3.0 and the crest factor will be approximately 3 (note that because the crest factor is not a statistical measure, its value in the presence of random noise will vary).

Trending of the waveshape metrics can also be used to help identify deteriorating condition. However, the trend of these values may be misleading in some cases; faults which produce a small number of isolated peaks (such as the initial stages of bearing damage) may cause an increase in the crest factor and kurtosis but, as the damage becomes more widely spread, a large number of impulses may occur causing both the crest factor and kurtosis to decrease again. Both the kurtosis and crest factor will decrease if the number of pulses increase (increasing the RMS value of the signal) without an increase in the individual pulse height.

As with the overall vibration level, the waveshape metrics will not detect faults unless the amplitude of the vibration from the faulty component is large enough to cause a significant change in the total vibration signal. This limits their use to components whose vibration signature forms a significant portion of the measured overall vibration.

4.1.5 Frequency band analysis

Often, the fault detection capability using overall vibration level and/or waveshape metrics can be significantly improve by dividing the vibration signal into a number of frequency bands prior to analysis. This can be done with a simple analogue band-pass filter between the vibration sensor and the measurement device. The rationale behind the use of band-pass filtering is that, even though a fault may not cause a significant change in overall vibration signal (due to masking by higher energy, non-fault related vibrations), it may produce a significant change in a band of frequencies in which the non-fault related vibrations are sufficiently small. For a simple gearbox, with judicious selection of frequency bands, one frequency band may be dominated by shaft vibrations, another by gear tooth-meshing vibrations, and another by excited structural resonances; providing relatively good coverage of all gearbox components.

4.1.6 Advantages

Meters for recording overall vibration levels, crest factor and/or kurtosis are readily available, relatively cheap and simple to use. Because of this, they can be a very cost effective method of monitoring simple machine components which are relatively cheap and easily replaceable but perform a critical role (for example small pumps and generators). The time domain signal metrics may detect the imminent failure of these components allowing replacement prior to total failure; although the damaged component may be beyond repair by this time, the component replacement cost is generally insignificant compared to the potential cost of catastrophic failure (secondary damage, loss of utility, etc.).

4.1.7 Disadvantages

For more complex or costly machines, it is generally preferable to detect damage at an early stage to allow the machine to be repaired rather than replaced. This requires techniques which are more sensitive to changes in the vibrations of individual components and which can provide at least some diagnostic capabilities.

4.1.8 Applicability to safety critical failure modes

Simple time domain signal metrics, even with the use of band pass filtering, do not provide any diagnostic information and, therefore, cannot be used to distinguish any of the safety critical failure modes from other failure modes.

For very simple safety critical systems, overall vibration level and/or kurtosis level (in combination with oil debris and/or temperature monitoring) may be useful as part of a cost effective failure detection system.

4.2 SPECTRAL ANALYSIS

Spectral (or frequency) analysis is a term used to describe the analysis of the frequency domain representation of a signal. Spectral analysis is the most commonly used vibration analysis technique for condition monitoring in geared transmission systems and has proved a valuable tool for detection and basic diagnosis of faults in simple rotating machinery [65,67]. Whereas the overall vibration level is a measure of the vibration produced over a broad band of frequencies, the spectrum is a measure of the vibrations over a large number of discrete contiguous narrow frequency bands.

The fundamental process common to all spectral analysis techniques is the conversion of a time domain representation of the vibration signal into a frequency domain representation. This can be achieved by the use of narrow band filters or, more commonly in recent times, using the discrete Fourier Transform (DFT) of digitised data. The vibration level at each 'frequency' represents the vibration over a narrow frequency band centred at the designated 'frequency', with a bandwidth determined by the conversion process employed.

For machines operating at a known constant speed, the frequencies of the vibrations produced by the various machine components can be estimated (as per the model described in Chapter 2) therefore, a change in vibration level within a particular frequency band can usually be associated with a particular machine component. Analysis of the relative vibration levels at different frequency bands can often give an indication of the nature of a fault, providing some diagnostic capabilities.

4.2.1 Conversion to the frequency domain

The frequency domain representation of a signal can be described by the Fourier Transform [67] of its time domain representation

$$X(f)=\int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}\,dt\,. \tag{4.8}$$

The inverse process (Inverse Fourier Transform [67]) can be used to convert from a frequency domain representation to the time domain

$$x(t)=\int_{-\infty}^{\infty} X(f)e^{j2\pi ft}\,df\,. \tag{4.9}$$

There are a number of limitations inherent in the process of converting vibration data from the time domain to the frequency domain.

4.2.1.1 Bandwidth-time limitation

All frequency analysis is subject to a bandwidth-time limitation (often called the Uncertainty Principle [27,67] due to the analogous concepts in quantum mechanics, enunciated by Werner Heisenberg in 1927).

Frequency analysis made with bandwidth of B hertz for each measurement and a duration in time of T seconds has a bandwidth-time limitation of:

$$BT \geq 1 \tag{4.10}$$

If an event lasts for T seconds, the best measurement bandwidth (the minimum resolvable frequency) which can be achieved is $1/\mathrm{T}$ hertz. If an analysing filter with a bandwidth of B hertz is used, at least $1/\mathrm{B}$ seconds will be required for a measurement.

The measurement uncertainty due to the bandwidth-time limitation imposes a resolution restriction on the frequency conversion. To resolve frequencies separated by B hertz at least 1/B seconds of data must be taken.

4.2.1.2 FFT Analysers

Most modern spectrum analysers use the Fast Fourier Transform (FFT) [25], which is an efficient algorithm for performing a Discrete Fourier Transform (DFT) [61,67] of discrete sampled data.

The Discrete Fourier Transform is defined as [61]

$$X(m) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j2\pi \frac{mn}{N}} \ , \tag{4.11}$$

and the Inverse Discrete Fourier Transform [61] is

$$x(n) = \sum_{m=0}^{N-1} X(m) e^{j2\pi \frac{mn}{N}} \quad . \tag{4.12}$$

The sampling process used to convert the continuous time signal into a discrete signal can cause some undesirable effects.

4.2.1.2.1 Aliasing

Frequencies which are greater than half the sampling rate will be aliased to lower frequencies due to the stroboscope effect. To avoid aliasing, an analogue low-pass 'anti-aliasing' filter is used prior to sampling to ensure that there are no frequencies above half the sampling rate.

4.2.1.2.2 Leakage

When applying the FFT, it is assumed that the sampled data is periodic with the time record. If this is not the case, spurious results can arise from discontinuities between the start and end points of the time record. This 'leakage' is normally compensated for by applying a smooth window function which has zero values at the start and end of the time record. This entails a resolution trade-off since it effectively reduces the time duration of the signal. For a simple machine, the time record can be synchronised with the rotation of the machine, ensuring that the major vibration components are periodic within the time record; this is difficult to achieve with complex machines due to the large number of non-harmonically related frequencies.

4.2.1.2.3 Picket Fence Effect

The picket fence effect is a result of the discrete frequency nature of the FFT. Where a frequency does not lie on one of the discrete frequency lines, the amplitude will be reduced. If the frequency is well separated from other frequency components, a correction can be made by curve fitting the samples around the peak. Windowing reduces the effect due to the increase in bandwidth caused by the windowing process.

4.2.1.3 Speed variations

The ability to resolve frequency components is not only related to the bandwidth-time limitation but also to the stability of the vibration signal over the analysis period. For FFT analysers, the resolution imposed by the bandwidth-time limitation is constant for all frequencies, however, the frequency of vibration signals due to the mechanically linked rotating components in a geared transmission system will vary proportionally with variations in the rotational speed of the machine, imposing a resolution limitation which is a constant percentage of the frequency.

Even with ‘constant’ speed machines, some drift in operating speed over time is likely to occur and, in some cases, may cause frequency variations (and uncertainty) which are greater than those due to the bandwidth-time limitations. For example, performing an FFT on one second of data would give a spectrum with a resolution of one hertz, however, a one percent speed variation over the analysis period would cause a 5 hertz uncertainty at a frequency of 500 hertz.

4.2.1.3.1 Synchronous sampling

The effects of speed variations can be overcome to a certain extent by the use of ‘synchronous sampling’, in which the sampling rate of the analyser is linked to the speed of the machine. However, this adds further complication to the monitoring process as it requires a speed sensor attached to the machine being monitored, a frequency multiplier to convert the speed sensor signal into a ‘clock pulse’ signal suitable for driving the signal analyser, and often needs an external anti-aliasing filter to avoid aliasing problems (although almost all modern FFT analysers have in-built anti-aliasing filters, when they are driven from an ‘external clock’ these are often bypassed or have inappropriate frequencies due to the unknown external clock frequency).

4.2.2 Fault detection

4.2.2.1 Spectral comparison

The most common spectral analysis technique employed for machine condition monitoring is spectral comparison, where a baseline power (magnitude squared) spectrum is taken under well defined normal operating conditions with the machine in known good condition (preferably soon after commissioning). This ‘baseline’ spectrum is used as a reference for subsequent power spectra taken at regular intervals throughout the machine life under similar operating conditions (Mathew [47]). The comparison is usually done on a logarithmic amplitude scale, with increases of 6-8 dB considered to be significant and changes greater than 20dB from the baseline considered serious (Randall [66]).

4.2.2.2 Spectral trending

In addition to spectral comparison, various forms of spectral trending [47,59,66] can be used to give some indication of the rate of fault progression. In its simplest form, spectral trending involves the trending of the changes in amplitude of all (or a number of selected) spectral lines over time. For complex machines, this can often involve a large amount of data, resulting in information overload due to the large number of significant spectral lines [47]. In an attempt to simplify the detection process, a number of parameters based on the spectrum have been proposed which provide statistical measurements of spectral differences. Mechefske and Mathew [59] give an overview of spectral parameters and a comparison of their detection and diagnostic performance for a number of bearing faults. They found that a number of these parameters performed well in the detection of the faults but that none of the parameters provided significant diagnostic information.

4.2.2.3 Spectral masks

Spectral masks are a method of spectral comparison sometimes employed to identify and evaluate changes in the signature spectrum [47,66], with allowances made for variation in operating condition. A spectral mask is derived from the baseline spectrum by adding an allowable tolerance limit to the logarithmic amplitude. To allow for variations in speed, the constant bandwidth spectrum is sometimes converted to a constant percentage bandwidth spectrum [66], with the percentage bandwidth being determined by the estimated speed differences which can occur between recordings (note that this is different to, and would be expected to be much larger than, the speed variations during the recording time described in Section 4.2.1.3).

Once a spectral mask is defined, comparison of individual recordings is made with reference to the mask to identify exceedences.

4.2.3 Fault diagnosis

Even for relatively simple machines, the vibration spectrum can be quite complex due to the multiple harmonic structures of the vibration from various components in combination with the transmission path effects (as detailed in Chapter 2). This makes detailed diagnostic analysis of an individual spectrum very difficult. The diagnostic process is simplified when performed in conjunction with spectral comparison and/or trending; typically, only the frequencies identified as having significant changes are analysed in detail for diagnostic purposes.

Randall [65] provides details of the expected spectral differences associated with various gear faults, and Su and Lin [74] provide similar information for bearing faults.

Distributed faults which cause significant change in the mean amplitude of the vibration at discrete frequencies, such as heavy wear and unbalance, should be relatively simple to diagnose using spectral analysis, as they would simply translate to changes in a few associated frequency lines in the spectrum.

Faults which cause low frequency sinusoidal modulations, such as an eccentric or misaligned gear, may also be diagnosed as they will translate to increases in the sidebands surrounding the tooth meshing frequency and harmonics.

Very localised faults, such as tooth cracking or spalling, are not easily diagnosed (and may not even be detected) as the short term impulsive vibrations produced translate to a large number of low amplitude lines in the spectrum (McFadden [56] and Randall [65]). An example where the diagnostic information available using spectral analysis was not sufficient to detect a fatigue crack in a gear (resulting in a fatal helicopter accident) is provided by McFadden [54].

4.2.4 Advantages

A number of companies manufacture and/or supply high quality FFT analysers at a reasonable price. In addition to marketing analysers, several of these companies also provide comprehensive after sales support in the form of literature and training in diagnostic methods using their equipment.

Because of the fairly widespread use of spectral analysis over a number of years, there is a fairly comprehensive collection of literature on its use for machine fault diagnosis (e.g., Randall [67] and Braun [15]).

4.2.5 Disadvantages

The major disadvantage with spectral analysis lies in its complexity. Even with the amount of literature available, specialist skills are still required to exploit the diagnostic capabilities of spectral analysis. When dealing with complex machines or with localised faults such as gear tooth faults, even expert analysts find diagnosis difficult.

4.2.6 Applicability to safety critical failure modes

For relatively simple machines, and those where the first few harmonics of the shaft vibration frequencies can be clearly identified (i.e., can be well separated from other vibration frequencies within the limits of bandwidth and/or speed variations), diagnosis of shaft related faults (fracture, unbalance, misalignment and bent shaft) should be quite simple with spectral analysis, by trending of the amplitudes of the shaft related vibrations or use of spectral masks.

The other safety critical faults identified in the previous chapter all produce impulsive signals and, as was demonstrated by McFadden [54], these faults cannot be reliably diagnosed (or, in some cases, even detected) using spectral analysis.

4.3 SYNCHRONOUS SIGNAL AVERAGING

Stewart [73] showed that with ‘time synchronous averaging’ the complex time-domain vibration signal from a machine such as a helicopter transmission could be reduced to estimates of the vibration for individual shafts and their associated gears. The synchronous average for a shaft is then treated as if it were a time domain vibration signal for one revolution of an individual, isolated shaft with attached gears.

4.3.1 Fundamental principle

The fundamental principle behind synchronous signal averaging is that all vibration related to a shaft, and the gears on that shaft, will repeat periodically with the shaft rotation (see Chapter 2, Sections 2.2.1 and 2.2.2). By dividing the vibration signal into contiguous segments, each being exactly one shaft period in length, and ensemble averaging a sufficiently large number of segments, the vibration which is periodic with shaft rotation will be reinforced and vibrations which are not periodic with the shaft rotation will tend to cancel out; leaving a signal which represents the average vibration for one revolution of the shaft. Figure 4.1 illustrates how this process might be performed on a continuous time signal from a gearbox, using a tacho multiplier to calculate each rotational period of the shaft.

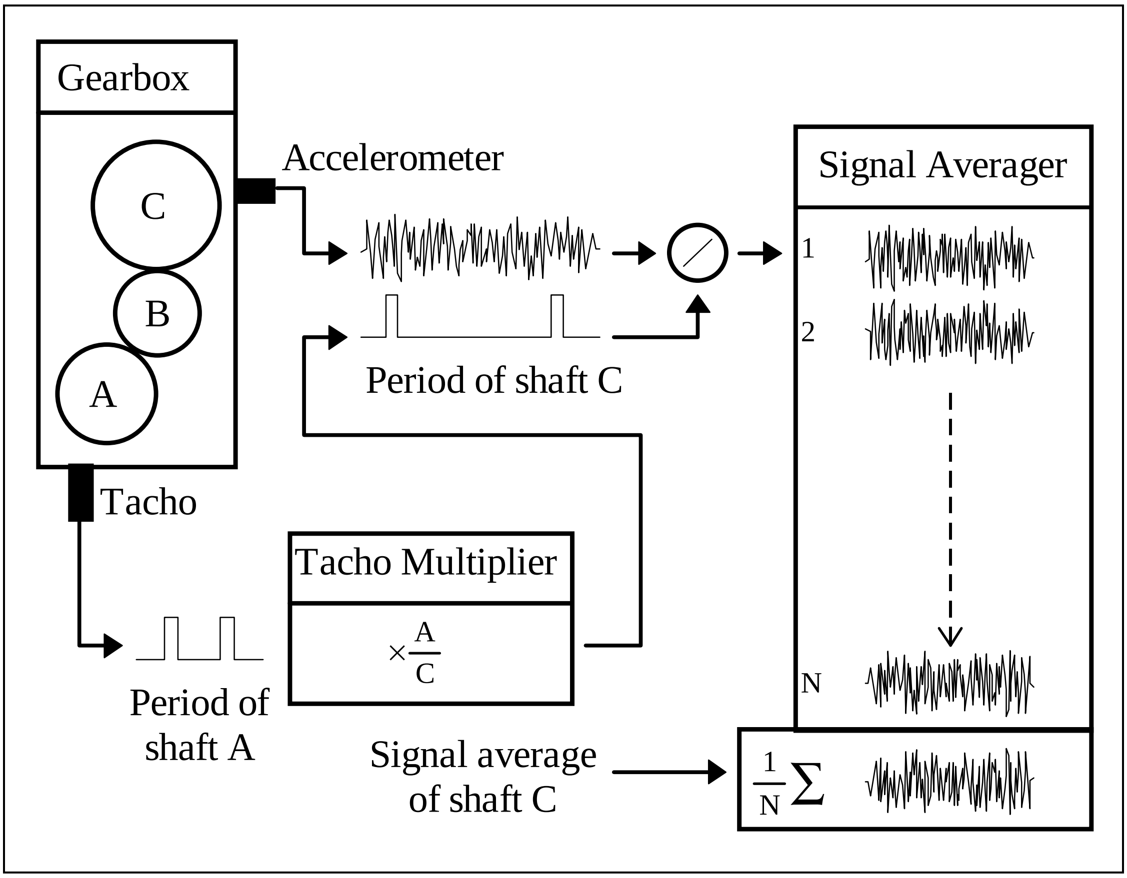

Figure 4.1 Synchronous signal averaging

4.3.2 Synchronous signal averaging of discrete signals

The process illustrated in Figure 4.1 assumes the vibration signal being averaged is a continuous time signal. In practice, the signal averaging process usually takes place on a discrete sampled signal (e.g., via an analogue-to-digital converter in a PC) and, in addition to defining the start and end points of the shaft rotation, some mechanism is needed to ensure that the sample points are at equally spaced angular increments of the shaft and that these are at the same angular position for each revolution of the shaft. That is, the sampling must be coherent with the rotation of the shaft. Originally, Stewart [73] used a phase-locked frequency multiplier however, McFadden [58] showed that far greater accuracy and flexibility could be achieved using digital resampling of time sampled vibration based on a reference derived from a simultaneously time sampled tacho signal. This method will be expanded upon in later chapters.

4.3.3 Terminology

It should be noted that the terminology related to this process is somewhat confused. The same process has been referred to as 'time domain averaging' (McFadden [53] and Braun [14]), 'time synchronous averaging' (Stewart [73]), 'coherent rotational signal averaging' (Swansson, et al [78]) and 'synchronous averaging' (Succi [75]).

The principle of synchronising the averaging with some other process (in this case the rotational frequency of a shaft) is fundamental to the technique; whether it is performed on continuous or discrete signals. As was seen above, when the process is performed on discrete signals the sampling must be coherent with the rotation of the shaft (hence the term 'coherent rotational signal averaging' used by Swansson, et al). Note that the process can be performed in the time or frequency domain (as long as the frequency domain averaging is performed on the complex frequency domain representation). The term 'time domain averaging' ([53] and [14]) was used to distinguish the technique from that of averaging of amplitude or power spectra to reduce variance in spectral analysis (Randall [67]).

To properly describe the process when applied to discrete signals, it should probably be referred to as 'rotationally coherent synchronous signal averaging'. However, this is quite clumsy and therefore the technique will normally be referred to as 'synchronous signal averaging' in the remainder of this thesis (the rotational coherency being implied when the technique is applied to discrete data).

4.3.4 Angle domain and shaft orders

Because the synchronous signal average is based on the rotation of the shaft rather than time, it is no longer correct to refer to it as a 'time' domain signal (although this is often done). Throughout the remainder of this thesis, a signal resulting from a synchronous signal averaging process (or any other angular based process) will be said to be in the **angle domain**. The angle domain is expressed in radians (or degrees) of revolution of the shaft (2π radians = 360 degrees = 1 revolution).

In the spectrum (Fourier transform) of an angle domain signal the frequency is expressed in **shaft orders** rather than Hertz (or RPM), where 1 shaft order = 1 cycle per revolution of the shaft.

4.3.5 Signal enhancements

In addition to treatment of a synchronous signal average as a simplified vibration signal (i.e., time domain analysis and spectral analysis can be applied), its periodic nature allows far more scope for signal manipulation than does a conventional vibration signal. The Fourier transform of a periodic signal is a pure line spectrum not subject to leakage (Section 4.2.1.2.2) or the picket fence effect (Section 4.2.1.2.3), and for which ideal filtering can be used; that is, one or more frequency lines (here representing shaft orders) can be completely removed from the spectrum without causing discontinuities when the signal is translated back to the angle domain.

This allows various signal enhancement techniques, specifically designed for the treatment of synchronous signal averages, and a number of related signal metrics to be used as an aid to fault detection and diagnosis.

4.3.5.1 Stewart’s Figures of Merit

Stewart developed a number of non-dimensional parameters based on the synchronous signal average, which he termed ‘Figures of Merit’ [73]. These were originally defined as a hierarchical group, with which Stewart described a procedure for the detection and partial diagnosis of faults.

4.3.5.1.1 FM0

The zero order figure of merit, FM0, was proposed as a general purposes fault detector to be applied to all signal averages. It is calculated simply by dividing the peak-to-peak value of the angle domain signal by the sum of the amplitudes of the tooth-meshing frequencies and harmonics. In simple terms, FM0 is a relative measure of the peak deviation of the signal from that defined purely by the mean gear tooth meshing vibration. Stewart claimed this to have good detection capabilities for

a) localised faults (such as tooth breakage, pitting, and spalling), bearing instability and misalignment as these increased the peak-to-peak level with little change in tooth-meshing level, and

b) heavy wear, as this produces little change in the peak-to-peak level with a reduction in tooth-meshing level.

4.3.5.1.2 Mesh-specific figures of merit (FM1, FM2 and FM3)

In the cases where FM0 indicated a significant change in the average, higher order 'mesh-specific' figures of merit (FM1, FM2 and FM3) were used to detect various patterns in the signal average indicative of certain types of faults.

FM1 is the relative measure of the low frequency (first and second order) modulation to the tooth-meshing amplitudes. This can be calculated for individual tooth-meshes (by dividing the sum of the two upper and lower sideband amplitudes by the tooth-meshing amplitude) or for all tooth-mesh related vibration (by dividing the sum of the two upper and lower sidebands of all tooth-mesh related frequencies by the sum of the amplitudes of the tooth-mesh related frequencies). FM1 will respond to misalignment, eccentricity, swash or shaft failures.

FM2 was designed specifically to detect single tooth damage, such as fracture or chipping, in multi-mesh gears. This is done by matching a pattern, consisting of spikes spaced at the mesh angular separation, with the envelope of the signal average. The matching is done using a circular matched filter (cross-correlation) of the pattern and the envelope. The ratio of the kurtosis (fourth statistical moment) of the matched filter output to the kurtosis of the envelope is used as the detection parameter. If there are impulses in the signal which are correlated to the angular separation between the mesh points, the matched filter operation emphasises these, forcing the kurtosis of the matched filter output above that of the envelope (giving FM2 > 1). Where no significant correlation of impulses to angular separation between the mesh points exists, the matched filter output tends to be flat with a kurtosis less than that of the envelope (giving FM2 < 1).

FM3 was designed to detect sub-harmonic meshing (i.e., significant sinusoidal vibration at a lower frequency than that of the tooth-meshing) and was claimed to be an indicator of parametric excitation and heavy wear. The signal average is low-pass filtered to include only frequencies up to and including the highest tooth-meshing fundamental frequency (e.g., filter set to 1.1 times the highest tooth-meshing fundamental). FM3 is the ratio of the zero-crossing count of a signal consisting only of the mesh related frequencies to the zero-crossing count of the low-pass filtered signal. Where there is strong sub-harmonic activity, the zero-crossing count of the filtered signal will be lower than that of the mesh related frequencies, causing an increase in the FM3 value.

4.3.5.1.3 FM4

If the change in the signal average could not be attributed to any of the predefined mesh-specific faults, FM4 (or 'bootstrap reconstruction') was used to detect other mesh related faults.

Stewart [73] reasoned that if one could define the expected frequency content for the vibration from a particular shaft (the regular signal), then all other vibration represents the deviation (the residual signal) from the expected signal. He proposed that the regular signal would normally include all tooth-meshing frequencies and their harmonics, plus their immediately adjacent sidebands. The residual signal is simply calculated by converting to the frequency domain, eliminating all components defined by the regular signal, and converting back to the angle domain.

Two parameters were proposed based on the residual signal:

a) FM4A: the kurtosis (4.6) of the residual, which will respond to impulsive signals such as those produced by pitting, spalling and tooth cracks.

b) FM4B: the ratio of the RMS of the residual to that of the original signal, which will respond to distributed faults which cause an increase in non-mesh related vibration.

The inclusion of the upper and lower sidebands in the regular signal meant that faults causing once per revolution modulations would not be detected using FM4. However, these are covered by FM1 and Stewart found that the removal of the sidebands from the residual increased sensitivity to other faults.

Experimental investigation by Stewart [73] showed that FM0 was not a very sensitive detector of localised tooth faults, such as tooth cracking, and he suggested that FM4 be used as a more sensitive general detector.

Because of the sensitivity of FM4 as a general fault detector, and its ease of implementation in comparison to the more specific fault detectors (FM1, FM2 and FM3), a number of experimental investigations based on Stewart's work have concentrated on enhancement techniques related to FM4 (e.g., McFadden [54] and Zakrajsek [84]).

4.3.5.2 Trend analysis

Stewart [73] also proposed a method of trending signal averages by calculating and trending the kurtosis and RMS of the difference between a baseline signal average and subsequent signal averages (from data taken from the same transducer location under the same operating conditions). The signal averages at different times need to be aligned in the angle domain before the difference signal is calculated. In cases were an absolute position reference is not available, the angular alignment is done by using a circular matched filter (cross-correlation) of the two signals; the location of the maximum correlation defining the angular offset between the two signals.

The kurtosis and RMS of the difference signal are interpreted in a similar fashion to FM4A and FM4B respectively.

4.3.5.3 Narrow-band Envelope Analysis

In the process of examining an in-service gear fatigue failure, McFadden [54] recognised the importance of phase modulation in the diagnosis of cracks and, based on this, developed a method of including phase information in a signal enhancement technique.

Assuming that the sideband structure surrounding a strong tooth-meshing harmonic was predominantly due to the modulation of the harmonic, McFadden [54] proposed that by narrow bandpass filtering about the selected tooth-meshing harmonic, removing the tooth-meshing harmonic, and calculating the angle domain envelope, a signal could be obtained which contains contributions from both the amplitude and phase modulations. The kurtosis of this signal was used as a measure of localised damage.

McFadden [54] showed that, in the early stages of cracking, this parameter had a better response than Stewart's FM4A.

4.3.5.4 Demodulation

McFadden [56] further developed the technique based on narrow bandpass filtering by using demodulation to extract an estimate of both the amplitude and phase modulation signals. He showed that a cracked tooth displays an amplitude drop with simultaneous phase change as the tooth comes in to contact. By displaying the amplitude and phase modulations simultaneously as a polar plot, these characteristic modulations could be seen as loops.

This particular development is of great significance, as it allows a distinction between tooth cracking and other localised tooth faults, such as pitting or spalling, to be made.

4.3.6 Advantages

The main advantages of synchronous signal averaging is that it allows a complex vibration signal to be reduced to a number of much simpler signals, each of which is an estimate of the vibration from a single shaft and its associated gears. The resultant signals are purely periodic and can be enhanced using ideal filtering.

The signal parameters developed by Stewart [73] and McFadden [54,56] further simplify the analysis task.

4.3.7 Disadvantages

Very few pieces of equipment are currently available which accurately implement synchronous signal averaging, and the ones that are available are very expensive. A number of analysers have 'time synchronous averaging' capabilities however, these generally have the capability of synchronising the start of each ensemble, but no method of ensuring coherent rotational sampling.

Most research equipment implementing synchronous signal averaging, such as the one used in this research project, are constructed and programmed by the researchers themselves based on PC's, analogue-to-digital converters and anti-aliasing filters. Accurate tacho signals and phase-locked frequency multipliers or digital interpolation techniques are also needed.

Further work needs to be done on methodologies for determining the parameters defining the synchronous signal averaging process, such as the number of averages and the sampling accuracy required. These issues will be discussed in subsequent chapters.

4.3.8 Applicability to safety critical failure modes

Synchronous signal averaging appears to be applicable in the detection of all the safety critical failure modes except bearing failures, which were considered to be a minor contributory factor in the area of safety.

From the brief descriptions given above, it would appear that McFadden's demodulation technique is the only one which provides discrimination between safety critical tooth fractures and other localised tooth faults such as pitting and spalling.

4.4 CEPSTRAL ANALYSIS

The power cepstrum is the power spectrum of the logarithm of the power spectrum and the complex cepstrum is the spectrum of the logarithm of the complex spectrum [16]. Both the power cepstrum and the complex cepstrum result in a time domain signal, which in the terminology of cepstral analysis is the quefrency domain, and give a measure of periodic structures in the spectrum. The usefulness of the cepstrum is in the fact that a series of harmonically related structures reduce to predominantly one 'quefrency' at the reciprocal of the harmonic spacing. This allows faults which produce a number of low-level harmonically related frequencies, such as bearing and localised tooth faults, to be detected. In this respect, it has advantages over spectral analysis in the detection of safety critical faults such as tooth cracking however, it still does not provide any distinction between these and less serious faults such as pitting and spalling.

Cepstral analysis has proved to be a useful tool in the detection of bearing faults; as bearing faults produce a series of impulses which excite structural resonances, the periodicity of the excitation is commonly evident in the 'quefrency' domain but, in the frequency domain, it appears as a number of low-level sidebands (separated by the frequency of the impulses and centred about each of the resonant frequencies) which are often difficult to detect.

Cepstral analysis is not very useful in the analysis of synchronously averaged signals because, even though the signal is periodic in the angle domain, and a pure line spectrum in the frequency domain, it looses its periodicity when translated to the quefrency domain and manipulation will introduce discontinuities.

4.5 ADAPTIVE NOISE CANCELLATION

Adaptive noise cancellation (ANC) has been used to increase the effective signal-to-noise ratio for bearing fault detection [79]. Deterministic methods for attenuating non-synchronous signal components (such as synchronous signal averaging) cannot be used for bearings because of the uncertainty in the rotational periodicities due to slippage and skidding. Bearing faults often go undetected due to the low level of the fault signature in relation to vibration from other components in the gearbox. However, it has been shown in laboratory experiments (e.g., Swansson and Favaloro [76]) that even simple time domain techniques, such as RMS, Crest Factor and Kurtosis, provide good bearing fault detection capabilities if the signal-to-noise ratio is sufficiently high. In an operational gearbox, bearing fault detection capabilities can be increased by increasing the signal-to-noise ratio of the fault signature rather than increasing the complexity of the diagnostic techniques.

ANC is implemented by using two transducers; one in close proximity to the bearing being monitored and the other remote from the bearing (usually closer to the source of the major interfering vibrations or ‘noise’). It is assumed that where there is a bearing fault, the vibration signal from the monitoring transducer contains the fault signature plus ‘noise’ and the signal from the reference transducer contains only the ‘noise’. The two ‘noise’ signals are assumed to be from the same source and correlated, with the difference being due to transmission path effects. An adaptive filter is used, minimising the difference between the two signals, which simulates the transmission path effects between the two transducers. The adaptive filter is applied to the ‘noise’ signal from the reference transducer, giving an estimate of the ‘noise’ signal at the monitoring transducer location, which is then subtracted from the signal from the monitoring transducer to provide an estimate of the bearing fault vibration signature.

The advantage of using ANC is that it requires no a-priori knowledge of the vibration frequencies in the interfering ‘noise’ signal. However, to give good results it requires a monitoring transducer for each bearing (or group of bearings in close proximity) and a remote reference transducer (although it may be possible to use a monitoring transducer for a remote bearing as the reference transducer).

ANC has no advantage in the monitoring of shafts and gears, for which synchronous signal averaging provide greater attenuation of non-synchronous vibration.

4.6 SUMMARY

Of the techniques discussed above:

a) Overall vibration level, crest factor and kurtosis monitoring of the time domain vibration signal do not provide any diagnostic information but may have limited application in fault detection in simple safety critical accessory components.

b) Spectral analysis may be useful in the detection and diagnosis of shaft faults.

c) Synchronous signal averaging has the potential of greatly simplifying the diagnosis of shaft and gear faults (i.e., the safety critical failures) by providing significant attenuation of non-synchronous vibrations and signals on which ideal filtering can be used. Further development needs to done on the implementation of synchronous averaging techniques and the analysis of results.

d) Cepstral analysis and adaptive noise cancellation mainly have application in bearing fault detection and diagnosis.

Based on the above, and the priorities placed on the safety critical failure modes in Chapter 3, the remainder of this thesis will concentrate on the synchronous signal averaging technique, firstly on further investigation, development, and testing of the process itself and subsequently on analysis methods to improve diagnostic capabilities for the safety critical failure modes.