Compensating for speed variation by order tracking with and without a tacho signal

M.D. Coats and R.B. Randall, School of Mechanical and Manufacturing Engineering, University of New South Wales, Sydney 2052 Australia

SYNOPSIS

Most machines have a small amount of speed variation with variations in load, and this has to be compensated for by order tracking when time synchronous averaging (TSA) is to be carried out. Some machines, such as wind turbines with doubly fed induction generators, have a much more widely varying speed and then some form of order tracking is necessary even to carry out bearing diagnostics. Order tracking normally requires some sort of tacho signal to allow resampling in the rotation angle domain. This paper describes a method whereby angular resampling can be carried out with and without a tacho signal, by progressive iteration, even in the case of large speed variations. Measurements were made on a gear test rig with a faulty bearing, with speed varying over a number of different ranges up to ±25%. A once-per-rev tacho signal and a 900-per-rev encoder signal were available, but it is demonstrated that it is often possible to extract a shaft speed related signal from the signal itself to use for order tracking. Note that the first harmonic of rotation speed is well separated from the next (the second harmonic) and is often clear of noise. This can often be used for a first iteration, and as the spread of higher harmonics is progressively reduced, they can be used to improve the results by progressive iteration.

1. INTRODUCTION

Even nominally constant speed machines, such as those driven by induction motors, have a small amount of speed variation with variations in load. Order tracking has to be used to compensate for this when time synchronous averaging (TSA) is to be carried out, for example for gear analysis. Some machines, such as wind turbines with doubly fed induction generators, have a much more widely varying speed, typically up to ±30%, and then some form of order tracking is necessary even to carry out bearing diagnostics. Order tracking normally requires some sort of tacho or shaft encoder signal to allow resampling in the rotation angle domain, giving a fixed number of samples per revolution of a particular shaft (at the same time nullifying the speed variations of all geared shafts). This paper describes a method whereby angular resampling can be carried out with and without a tacho signal, by progressive iteration, even in the case of large speed variations. Measurements were made on a gear test rig with a faulty bearing, with speed varying over a number of different ranges up to ±25%. A once-per-rev tacho signal and a 900-per-rev encoder signal were available, but it is demonstrated that it is often possible to extract a shaft speed related signal from the signal itself to use for order tracking. The method is based on phase demodulation, giving a map of phase (rotation angle) vs time, and then allowing resampling at equal phase intervals. By the so-called Bedrosian’s conditions to prevent overlap of modulation sidebands, the maximum variation which can be allowed for is about ±30%, if demodulating the first harmonic of a harmonic series, or for example ±10% if demodulating the third harmonic. Note that the first harmonic is well separated from the next (the second harmonic) and is often clear of noise. This can be used for a first iteration, and as the spread of higher harmonics is progressively reduced, they can be used to improve the results by progressive iteration. Each time the spread is reduced, less noise is included in the band demodulated, and the phase/time map improved, allowing the progression to a higher harmonic and better results.
In addition to the results for the gear test rig, some results are also given for signals from a gas turbine engine for which the gear ratio between the high speed shaft and that on which a tacho signal generator was mounted was only known approximately.

2. PHASE DEMODULATION AND ANGULAR RESAMPLING

For a single carrier frequency $f_c$, a general amplitude and frequency modulated signal can be expressed as:

$$A_m(t) \cos(2\pi f_c t + \phi_m(t))$$

which can be interpreted as the real part of the rotating phasor

$$A_m(t) \exp\{j(2\pi f_c t + \phi_m(t))\}$$

Provided the spectrum of (2) is one-sided (positive frequencies only), it will be an analytic signal, and the imaginary part will be the Hilbert transform of the real part (the same as (1) with a sine replacing the cos). Note that the actual amplitude modulation signal (often with zero mean value) must be added to a constant positive value, so that $A_m(t)$ is always non-negative, and this constant would have to be subtracted after demodulation to regain the amplitude modulation signal.

If the (linear) phase corresponding to the carrier component ($2\pi f_c t$) is subtracted, the resulting complex time signal will have the amplitude modulation component $A_m(t)$ as its amplitude, and phase modulation component $\phi_m(t)$ as its phase. This will no longer be analytic since it has both positive and negative frequency components. The subtraction can be done either in the time domain by multiplication with the negatively rotating phasor $\exp\{-j(2\pi f_c t)\}$ (as in a zoom processor) or in the frequency domain simply by shifting the carrier frequency to zero, as illustrated in Figure 1. Note that because of the periodicity of FFT spectra, the negative frequency components must be shifted to just below the sampling frequency as illustrated in Fig. 1. Note that the frequency components must be complex, not the amplitudes illustrated.

![Figure 1](image)

Since Expression (2) is a product of two terms, the spectrum will be the convolution of the two spectra, whose total bandwidth will thus be the sum of the bandwidths of the two components. The bandwidth of the amplitude modulation component is directly that of $A_m(t)$, but the bandwidth of $\exp\{j(\phi_m(t))\}$ is...
theoretically infinite, even for a single modulation frequency. However, if a limit is put on dynamic range, it is meaningful to talk in terms of a finite bandwidth in terms of the bandwidth of $\phi_m(t)$.

Even for phase modulation of a single carrier frequency by a single modulation frequency, the spectrum of the modulated signal has multiple sidebands around the carrier component, as illustrated in Figure 2. The strength of the sidebands is given by Bessel functions of different orders, and argument $\beta$, this being the “modulation index”, or phase deviation around the carrier in radians. It is also the ratio of the frequency deviation $f_d$ to the modulation frequency $f_m$. In this paper we limit the number of significant sidebands as being those within 40 dB of the highest. For very small $\beta$, $\ll 1$ radian, only one pair of sidebands is required, but with very large $\beta$ it becomes more meaningful to talk in terms of frequency deviation rather than phase deviation, the frequency modulation signal (angular velocity) being the derivative of the phase modulation signal (angular displacement). Thus, if a frequency is sweeping by an amount of the order of $\pm 20\%$ of the carrier, the modulation sidebands will spread over a frequency range of the order of this amount.

![Figure 2 Sideband spectrum for single frequency modulation in terms of Bessel functions.](image)

One aim of this paper is to show how information about the instantaneous speed of a rotating machine can be extracted from the signal of a tachometer or shaft encoder attached or geared to the shaft in question. These typically give a series of pulses, in general including a zero frequency component or DC offset, as opposed to a sinusoidal signal linked to the shaft speed, and the spectrum of such a series of pulses contains a large number of harmonics of the fundamental. The amplitude of such pulses is approximately constant, and so modulation sidebands come primarily from the phase/frequency modulation. To avoid overlap of the sidebands around each harmonic, they should not spread to more than 50% of the spacing between them, and so the absolutely highest modulating sideband would be at 50% of the fundamental frequency. However, variations in speed of a shaft are represented by variations $\Delta \omega_t$ around the mean spacing of the pulses, and the corresponding phase change is thus proportional to the order of the harmonic in the spectrum of the pulse train. Since the phase $\phi_m(t)$ is in radians of the carrier frequency, this must be divided by the harmonic order of the carrier demodulated to obtain it in terms of shaft rotation angle, giving it a correspondingly higher resolution. Figure 3(a) shows the spectrum of a carrier signal with two harmonics, modulated by a single frequency at 33% of the fundamental carrier frequency. It is seen that since the spread of sidebands around the second harmonic is twice that around the first, the spread of 33% above the first harmonic is on the limit to avoid overlap with the double width coming from the second harmonic. Thus 33% is the absolute limit on frequency sweep to avoid overlap of sidebands from higher harmonics. Another way of understanding this comes from the realization that in the limit, the envelope of the sidebands for a large frequency sweep approaches the probability density function for a sinusoid as illustrated in Fig. 3(b), and thus the spread of sidebands is approximately proportional to the frequency sweep of the harmonic concerned and thus to the harmonic order.

Making this assumption, Figure 4 shows that provided there is no overlap around the lowest harmonic, the sidebands from higher harmonics will not encroach on this valid zone.
The above-mentioned limit of 33% sweep only applies when the rate of sweep is very low compared with the carrier frequency, and the allowable range of speed variation is even further restricted when the rate of change is higher. Ref. [1] reports the results of an empirical study of the effects of sweep rate on the limit of relative frequency sweep, using the above-mentioned criterion of 40 dB SNR (signal-to-noise ratio). The results are shown in Figure 5, where the x-axis is the ratio (as a percentage) of the modulation frequency to the carrier frequency, called $f_{m\%c}$, a measure of how rapidly the carrier frequency is changing. The y-axis is the allowable maximum frequency deviation, also as a percentage of the carrier frequency, and called $f_{d\%c}$. The greater the rate of change of carrier frequency, the smaller the allowable frequency deviation to avoid sideband overlap. The discontinuous nature of the chart is because the 40 dB criterion of negligibility requires the inclusion of progressively more or less sidebands in discrete steps. It is seen that once the modulation frequency is $> 33\%$ of the carrier, the allowable frequency deviation is very small.

Figure 5 is based on the carrier signal being a series of impulses with all harmonics of the same strength. It also assumes that there is only one modulation frequency. It will give conservative results if $f_d$ is taken as the peak deviation of the carrier, and $f_m$ is taken as the maximum modulation frequency.
This set of criteria applies to tachometer and shaft encoder signals which are directly synchronous with the instantaneous shaft speed, and would also apply to forcing functions, such as unbalance, which are also synchronous. When the harmonics are response signals, such as vibration acceleration responses, a further limitation arises because of the response time of the system, and this is extensively studied in [2]. Basically, when a speed change is rapid, the response at a particular time has actually been excited somewhat earlier, and thus corresponds to a different shaft speed than the current one. This paper assumes that speed changes are slow enough that this limitation is negligible. This should be the case for example with wind turbines, where the inertia is so high that speed cannot change very rapidly.

It can thus be seen that provided the spread of sidebands around a given harmonic of a tachometer signal does not overlap with those from adjacent harmonics, then that carrier frequency can be phase demodulated by the procedure illustrated in Fig. 1, giving a map of phase vs time. As mentioned above, to express this in terms of rotation angle of a particular shaft, the demodulated phase would have to be divided by the harmonic order of the carrier frequency with respect to that shaft. Such a phase signal is normally sampled at uniform time intervals, but for “order” analysis (expressing “frequencies” as orders of shaft speed) the phase/time map has to be resampled at uniform intervals of phase or rotation angle, giving a constant number of samples per revolution. This is known as “angular sampling”. The corresponding “time” axis is then in terms of the number of periods of rotation. The resampling not only has to correspond to a fixed number of samples per revolution, but each record should also start at a fixed rotation angle, for example a (positive-going) zero crossing of a once-per-rev tacho signal, requiring interpolation since the actual time samples will rarely fall at these zero crossings. Linear interpolation would normally be sufficient to determine such zero crossings (or passage through a threshold level for a series of pulses), but the resampling of the actual signals should be done with a higher order interpolation. McFadden [3] shows that the best general interpolation procedure uses cubic spline interpolation, for two reasons:

1. The sidelobes which fold back into the valid measurement range because of aliasing are very small. They are even smaller if the sampling frequency is doubled before interpolation.
2. The resulting lowpass filter effect is less than with lower order interpolation. It can be made virtually negligible by doubling the sampling frequency before interpolation.

The phase demodulation procedure illustrated in Fig. 1 virtually removes the carrier frequency, and this is necessary to be able to reliably “unwrap” the phase to a continuous function of time. This is because the phase is determined as the arctangent of a number and can only be determined in the range $\pm \pi$, before unwrapping to remove “jumps” over $2\pi$. The actual phase/time map used for angular resampling, must include the linear component corresponding to the carrier frequency, but this is known exactly and can be added back in with no error.

![Figure 5](image-url)
3. COMPENSATING FOR SPEED VARIATIONS

In this section the above principles are used to show how compensation can be made for speed variations, even large speed variations, to allow bearing diagnostics in a gearbox with masking from the gear signals. The gearbox was operating with two 32-tooth spur gears at a mean shaft speed of 6 Hz, with various amounts of speed variation including ±10% and ±25%. The period of the frequency sweep was 5s, making the modulation frequency $f_m = 0.2$ Hz. The signals analysed are with 25 Nm torque load, resulting in a radial bearing load of about 250 N. One bearing had seeded local faults, and the case analysed here is for an inner race fault. A once-per-rev tacho signal was available, as was a 900-per-rev shaft encoder signal. Here, only the tacho signal was used. Figure 6 shows time signals and corresponding spectra for the case with ±10% speed variation. The acceleration signal was measured on the casing above the faulty bearing.

![Figure 6 Signals with ±10% speed variation](image)

From Fig. 6(b) it can be seen that the first three harmonics do not have overlapping sidebands and can be used for demodulation. In this case the response acceleration has very weak first and second harmonics (the machine had recently been refurbished including balancing and realignment), but the third harmonic appears equally separated as that of the tacho signal. The acceleration response signal is obviously amplitude modulated, this being explained by the fact that forcing components such as gearsmesh harmonics were passing through fixed resonance frequencies. It is worth remarking that this does not affect the phase modulation, as these are separated in Expression (2). Another way of understanding this is to recognise that for a modulated sinusoid such as (1), the zero crossings represent phase increments of $180^\circ$, and this is unaffected by any amplitude modulating function which is always positive.

Figure 7 shows the results of using the third harmonic of the two signals to perform order tracking. In both cases, the spectrum of the corrected tacho signal is shown (cf Fig. 6(b)). The frequency axis is scaled in terms of the mean speed of 6 Hz, but it is actually an order axis, which can be obtained by division by 6. It is seen that even though the tacho signal does a better job, the correction of 7(b) is still sufficient to allow envelope analysis of bearing faults, where only the first few harmonics are required. A second iteration using say the 9th harmonic of Fig. 7(b) could have been used to greatly improve the correction, and this (and further iterations as discussed below) would be required for gear analysis.
Fig. 7  Spectra of order tracked tacho signal (a) using 3\textsuperscript{rd} harmonic of tacho signal (b) using 3\textsuperscript{rd} harmonic of acceleration signal

Fig. 8 shows the averaged spectrum of the squared envelope of the acceleration signal, demodulated in a high frequency band which maximised the kurtosis of the transmitted signal. The frequency axis is once again based on mean rotational speed. The first two harmonics of BPFI (ballpass frequency, inner race) are seen, along with low harmonics of shaft speed (6 Hz), and sidebands spaced at this frequency around BPFI. It should be noted that in the direct FFT spectrum of the envelope, there were modulation sidebands spaced at a mean value of 0.2 Hz because of the amplitude modulation seen in Fig. 6(c), but these are hidden because of the coarser resolution of the averaged spectrum.

In this case, the signal itself did not contain a strong first harmonic, but this is unusual. Figure 9 shows the time signals and spectra (cf Fig. 6) for ±25% speed variation.
It is now clear that only the first harmonic (of the tacho signal) is separated from other harmonics, but when this was used for order tracking the order spectra of Figure 10 resulted (scaled in terms of mean frequency).

The low harmonics of the acceleration signal are now well separated, and the peak near 192 Hz corresponds to the gearmesh frequency, surrounded by modulation sidebands at shaft speed. These are still a little smeared, but by comparison with the equivalent harmonics of the tacho signal, it appears that most of the speed variation has already been removed, and the residual sidebands are due to amplitude modulation (a fixed frequency of 0.2 Hz in the time domain, but smeared in the angle domain). Even so, the correction given by this single stage of order tracking was sufficient to allow bearing diagnostics and a similar result was achieved to that in Fig. 8.
However, it seems that this amplitude modulation may provide a limit to what can be achieved by synchronous averaging for gear analysis, and this is the subject of ongoing research.

3. IMPROVEMENT BY FURTHER ITERATION

Where speed variation is small, so that the amplitude modulation effect is negligible, further iteration can be used to improve the phase/time map, using higher and higher harmonics, possibly of the signal itself.

This was studied in Ref.[4], from which some results are given here. The experimental data utilised in that case was captured from a Larzac gas turbine engine at DSTO Laboratories, Melbourne, and the measurement data was supplied by DSTO. A vibration signal was captured from an accelerometer mounted on the casing of the high pressure (HP) stage compressor. A once-per-rev tachometer was mounted on the low pressure (LP) stage main shaft, providing a reference signal for the low pressure shaft, and this was first used to remove the harmonics of the LP shaft, which did not extend to very high frequencies. A sixty pulse-per-rev tachometer was mounted on the accessory gearbox shaft, providing a reference signal in proportion to the HP shaft of interest. The gear ratio supplied by the manufacturer between the accessory gearbox shaft and the HP shaft was 0.1861. This was only sufficiently accurate to allow partial removal of the first few harmonics of the HP shaft by synchronous averaging, leaving the higher order harmonics in place.

A number of methods were tested to improve the estimation of the gear ratio, including curve fitting the instantaneous frequency and progressive phase of the HP shaft speed and the tacho signal over various lengths of record, and using various harmonics of each. The criterion used to judge the accuracy of the estimate was termed the “Separation Index” (SI), defined as the ratio of the total mean square value of the separated deterministic components to the mean square value of the residual signal after subtracting them. Assuming that these two components are independent, use of the correct ratio should maximise the numerator of this quotient, and simultaneously minimise the denominator. The actual value of the Index would of course be data dependent, and would vary for each measurement point and operating condition, but for a particular measurement should be maximised when the optimal separation of deterministic and random components is achieved.

An alternative approach, corresponding to that described in this paper, was also applied, which did not make reference to the gear ratio between the HP shaft and tacho signals, but simply attempted to remove frequency modulation of higher and higher harmonics of the shaft speed, so as to improve the phase/time map. In one version of this, the tacho signal was used in a first iteration, simply to remove as much as possible of the speed variation. After a first iteration using the tacho signal, the 141th harmonic of the HP shaft speed was reasonably separated from adjacent harmonics and could be used for a further iteration to obtain a better result. It was then found that virtually the same result could be achieved without using the tacho signal, but simply demodulating harmonics of the HP shaft itself. Successive iterations were done using the 1st, 43rd, and again the 141st harmonic to obtain the best possible result. Table 1 shows a comparison of the best results achieved by the different methods, and Figure 11 compares spectra of the original and residual signals for two of the methods listed in the table. The figures show the high orders around 58, since lower orders were virtually completely removed. It is likely that what appear to be harmonics in the residual signal spectra are actually narrow band noise peaks caused by minor random amplitude and phase modulation of the actual deterministic harmonics of shaft speed. For example in Ref. [5] it is pointed out that bladepass frequencies in turbines arise from interactions between the blades and casing via a turbulent fluid, and so are not completely deterministic.
Table 1 Effectiveness of separation of discrete harmonic components from background noise

<table>
<thead>
<tr>
<th>Method</th>
<th>Separation Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gear Ratios</td>
<td></td>
</tr>
<tr>
<td>Manufacturer’s Ratio</td>
<td>0.0649</td>
</tr>
<tr>
<td>Frequency domain Approach</td>
<td>0.2211</td>
</tr>
<tr>
<td>Phase Domain Approach (best)</td>
<td>0.5025</td>
</tr>
<tr>
<td><strong>Extraction Approach</strong></td>
<td></td>
</tr>
<tr>
<td>Using tachometer signal</td>
<td>0.6917</td>
</tr>
<tr>
<td>Without using tachometer</td>
<td>0.7205</td>
</tr>
</tbody>
</table>

Figure 11  Spectra of original and residual signals  (a) Using the linearised phase (SI 0.5025)  
(b) Using signal harmonics – no tacho (SI 0.7205)

Later a trial and error method was used to find the optimum gear ratio, and this gave even better separation than the methods listed in the table. However, it is interesting that the best ratio was found to be different for two separate measurements at different speeds/loads, and both were different (to seven significant figures) from the true gear ratio, later obtained from the manufacturer. This leads to a discussion of the use of such a gear ratio in performing order analysis using a tacho signal from a shaft other than that on which the tacho is mounted. This effectively assumes that the transmission is rigid, whereas the actual ratio of the number of rotations of the two (or more) shafts is influenced by the elastic deflection of the components and in principle also by the transmission error and even the possible loss of tooth contact, in particular under light load.

**CONCLUSION**

This paper has shown that it is possible to compensate for speed variations of a machine with or without the need for a speed sensor such as a tachometer or shaft encoder. The instantaneous speed information can often be extracted from the signal itself. Even speed variations of the order of 25-30% can be compensated by first demodulating the first harmonic of shaft speed, and residual errors can usually be reduced by iteration, through demodulation of progressively higher harmonics. For large speed variations,
some limits may be imposed by the fact that signals become amplitude modulated by the fact that different shaft orders pass through fixed resonance frequencies, and it is more difficult to remove deterministic components whose amplitude is not constant. These methods give promise that it will become easier to perform diagnostics of machines such as wind turbines, with widely varying speed and load.

ACKNOWLEDGMENTS

This work was supported in part by the Australian Research Council, through Discovery Grant P0771204, and by Australia’s Defence Science and Technology Organisation (DSTO).

REFERENCES