Håvard Vold
Vold Solutions
P.O. Box 8286
Cincinnati, OH 45208-0286

Jan Leuridan
LMS International
Interleuvenlaan 68
3001 Leuven, Belgium

High Resolution Order Tracking at Extreme Slew Rates Using Kalman Tracking Filters

The analysis of the periodic components in noise and vibration signals measured on rotating equipment such as car power trains, must be done more and more under rapid changes of an axle, or reference RPM. Normal tracking filters (analog or digital implementations) have limited resolution in such situations; wavelet methods, even when applied after resampling the data to be proportional to an axle RPM, must compromise between time and frequency resolution. The authors propose the application of nonstationary Kalman filters for the tracking of periodic components in such noise and vibration signals. These filters are designed to accurately track signals with a known structure among noise and signal components of different, "unknown," structure. The tracking characteristics of these filters, i.e., the predicted signal amplitude versus time values versus exact signal amplitude versus time values, can be tailored to accurate tracking of harmonics buried in other signal components and noise, even at high rates of change of the reference RPM. A key to the successful construction is the precise knowledge of the structure of the signal to be tracked. For signals that vary with an axle RPM, an accurate estimate of the instantaneous RPM is essential, and procedures to this end will also be presented.

INTRODUCTION

This article investigates the application of nonstationary Kalman filters to track harmonics and orders from signals measured on rotating equipment using a constant sample rate. This approach enables the analysis of orders even at high rates of change, or slew rate, of a reference RPM. This is important in such applications as analyzing data from a standard exterior pass-by-noise test (1SO362) that typically only lasts 2 s. Also some phenomena in rotating equipment only become apparent at operating conditions involving rapid change of speed. This can be due to nonlinearities, load dependent system dynamics as is prevalent in torsional systems, and operating requirements such as in the spindown through critical speeds of centrifuges.

Real time analog and digital tracking filters have limitations of resolution in such situations, due to transients and excessive processing requirements in digital implementations (Potter and Gribler, 1989). Resampling of recorded signals to achieve periodicity with respect to the reference RPM following by Fourier analysis may require excessive oversampling of the original signal, es-

Received September 3, 1993; Accepted May 12, 1995.
Shock and Vibration, Vol. 2, No. 6, pp. 507–515 (1995)

CCC 1070-9622/95/060507-09

pecially when tracking fractional orders. It can be nearly impossible to resample orders that are not expressible as fractions of small composite integers. Fourier analysis requires stationarity over longer time periods for good frequency resolution; and modern wavelet analysis, even when applied to resampled data, are subject to an unyielding compromise between frequency and time resolution.

The authors propose the application of nonstationary Kalman filters for the tracking of harmonic components in noise. Such filters have been employed very successfully in control and guidance systems since 1960, with particular application to avionics and navigation (Kalman, 1960; Kalman and Bucy, 1961). These filters can accurately track signals of known structure among noise and other signal components with different structure. For signals that vary with an axle RPM, an accurate estimate of the instantaneous frequency is essential; procedures to this end, based on the availability of a tachometer channel, will also be presented.

MATHEMATICAL BACKGROUND

A Kalman filter is a computational procedure that tracks a signal of known nonstationary structure, or the state of a nonstationary dynamic system from noisy measurements. The usual formulations are couched in the language of modern control theory, because the original developments and most of the engineering applications have been within the control and guidance field. We choose to develop a Kalman filter to track sine waves from signals sampled at a constant rate as an alternative to more traditional analog and digital techniques because it provides a simple formulation and a potential for tracking sine waves accurately at high slew rates. This version allows for the standard predictor form of the Kalman filter, highly useful in real time applications such as shaker control and avionics, but also for a global least squares best linear unbiased estimator of the system state, best suited for postprocessing and analysis of archival data.

The application of the Kalman filter for sine tracking is done in three stages, RPM determination, waveform tracking, and amplitude/phase determination, even though these processes may be done in lock step for real time implementation.

Waveform Tracking

Assuming that we know the RPM and the order that we want to extract from a measurement channel, the Kalman filtering consists of setting up and solving for the waveform, optionally in a recursive manner, of a sparse set of linear least squares equations, the components of which we call the structural and the data equations.

Structural Equation. A sine wave $x(t)$ of frequency ω with arbitrary amplitude and phase, sampled at even time increments Δt, satisfies the second-order difference equation

$$x(n\,\Delta t) - 2\cos(2\pi\omega\,\Delta t)x((n-1)\,\Delta t) + x((n-2)\,\Delta t) = 0, \quad (1)$$

which we normally write

$$x(n) - c(n)x(n-1) + x(n-2) = 0, \quad (2)$$

dropping the time increment, Δt, from the equations. We note that when the instantaneous frequency, ω, is known, Eq. (2) is a linear, frequency dependent constraint equation on the sine wave that we call the structural equation of the Kalman filter. In our application where we are tracking a sine wave of changing frequency contaminated with noise and other sinusoids, we introduce a nonhomogeneity term, $\varepsilon(n)$, that allows the sine wave to change its amplitude and phase as well as frequency slightly over the time points involved in the equation. This now becomes

$$x(n) - c(n)x(n-1) + x(n-2) = \varepsilon(n), \quad (3)$$

where $c(n) = \cos(2\pi\omega\,\Delta t)$. The right-hand side of Eq. (3) is a deterministic, but unknown term that allows deviations from a true stationary sine wave. In particular, if the sine wave is locally amplitude modulated by a slow sine of frequency ω_a, trigonometric identities show that

$$2\sin(2\pi\omega t)\sin(2\pi\omega_a t) = \cos(2\pi(\omega - \omega_a)t) - \cos(2\pi(\omega + \omega_a)t). \quad (4)$$

This equation shows that when the amplitude changes, the target wave is a superposition of two waves with slightly higher and lower frequencies. A first-order expansion of Eq. (4) gives a term that may be lumped into the right-hand side of Eq. (3). It is useful to define the measures $s_\varepsilon(n)$

as the standard deviation of the nonhomogeneity of the structural equation, Eq. (3).

Data Equation. Instead of observing $x(n)$, we measure the signal $y(n)$ that is assumed to contain both the signal that satisfies the structural equation as well as noise and other periodic components. Formally, we write this as

$$y(n) = x(n) + \eta(n), \tag{5}$$

where $\eta(n)$ is a signal component containing random noise and periodic components at other frequencies than the target signal. Here we also define $s_\eta(n)$ as the standard deviation of the nuisance component $\eta(n)$.

Least Squares Equation. We see that at any point of time n, Eqs. (3) and (5) implicitly provide linear equations for $\{x(n) \quad x(n-1) \quad x(n-2)\}$. Rearranging these equations gives us the unweighted form

$$\begin{Bmatrix} 1 & -c(n) & 1 \\ & & 1 \end{Bmatrix} \begin{bmatrix} x(n-2) \\ x(n-1) \\ x(n) \end{bmatrix} = \begin{bmatrix} \varepsilon(n) \\ y(n) - \eta(n) \end{bmatrix}, \tag{6}$$

with the structural equation as the top row and the data equation in the bottom. Writing the ratio of the standard deviation functions of the right-hand side of Eq. (6) as

$$r(n) = \frac{s_\varepsilon(n)}{s_\eta(n)}, \tag{7}$$

allows us to make the make the error in Eq. (6) isotropic by applying the weighting $r(n)$, that is,

$$\begin{Bmatrix} 1 & -c(n) & 1 \\ & & r(n) \end{Bmatrix} \begin{bmatrix} x(n-2) \\ x(n-1) \\ x(n) \end{bmatrix} = \begin{bmatrix} \varepsilon(n) \\ r(n)(y(n) - \eta(n)) \end{bmatrix}. \tag{8}$$

Roughly speaking, a least squares system weighted such that isotropic error occurs will give us the minimum variance unbiased estimates of the system state as long as the error terms have locally zero means. Deviating from this particular form of weighting will affect the estimates of system state; applied to the tracking application, it will influence the tracking charcteristics of the filter; that is, this weighting may also be used to enforce adherence to the structural equation by choosing a small value for $r(n)$, leading to a filter that is highly discriminating in frequency, but which takes a long time to converge in amplitude. Conversely, fast convergence with low frequency resolution is achieved by choosing $r(n)$ large. This corresponds also to the observation from Eq. (4) that says that rapid amplitude variations imply frequency deviations to either side of the tracked frequency.

Applying Eq. (8) to all observed time points will give us a global system of overdetermined equations for the desired waveform $x(n)$ that may be solved using standard least squares techniques such as normal equations, QR decomposition, or the singular value decomposition (SVD). Inspection of Eq. (8) shows that the normal equations are banded, allowing for very fast solution schemes.

Global versus Predictor Solutions. The banded form of the least squares equation, Eq. (8), entails that an incremental solution is readily available. To this end, one simply solves the equations up to time n and uses the estimated value of $x(n)$ as a state estimate for that time point. To obtain an estimate for time in the next time step, one simply adds the next block of equations for time $n + 1$ as in Eq. (8), and using a recursive argument, obtains the estimate for $x(n + 1)$ from the new equations and the old estimate of $x(n)$. This is the standard procedure from the modern control procedures where a current state estimate is needed for proper closing of the control loop. This formulation of Kalman filtering may be found in a number of textbooks, see e.g. Giordano and Hsu (1985) and Goodwin and Sang Sin (1984).

It may be argued from a probability theory that the interior points of the time series $x(n)$ are determined with less statistical variability than the end points. From a practical point of view this translates into developing a smoothed estimate or prediction of the system state at time n by including data points for up to time $n + p$, where p is a nonnegative integer. The specific application will determine the trade-off between the lag p of the prediction, and the variability of the prediction. For many applications, including postprocessing of field data, the optimal procedure would be to include the entire data set in a global estimation. Figure 1 shows the prediction error from a

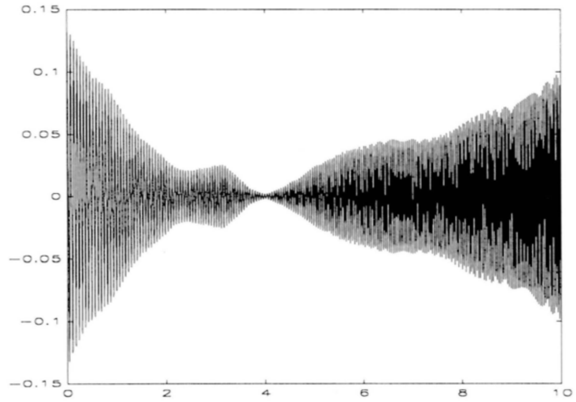

FIGURE 1 Waveform prediction error with Gaussian noise.

track of a sine wave whose frequency doubles in the analysis window. The signal is contaminated with Gaussian noise whose standard deviation is the same as the amplitude of the sine. Figure 2 shows the prediction error when the contaminants are harmonics of the tracked wave. Both of these figures illustrate the higher variability at the end points.

Amplitude/Phase Determination

For the purpose of order tracking, the filtered waveform is most conveniently described in terms of amplitude and phase with respect to a reference such as a tachometer channel. Because the tracked signal in our formulation is a sine wave $x(n)$ with known frequency as a function of time locally at time n, the cosine and sine components are given through the equation

$$x(n) = \left\{ \cos(2\pi \sum_{k=0}^{n} \omega(k)\,\Delta t) \quad \sin(2\pi \sum_{k=0}^{n} \omega(k)\,\Delta t \right\} \begin{bmatrix} a(n) \\ b(n) \end{bmatrix}. \tag{9}$$

Assuming that the amplitude and phase are locally constant, a least squares overdetermined system of equations may be constructed by considering Eq. (9) for a number of time points neighboring n. Strictly speaking, only two equations are necessary; but, considering the impreciseness of the tracking prediction, at least enough time points to cover a complete period should be used. These equations are known from experimental dynamics as the Vandermonde equations, even though our formulation considers frequency variation over the data points used in the estimation (Vold et al., 1982).

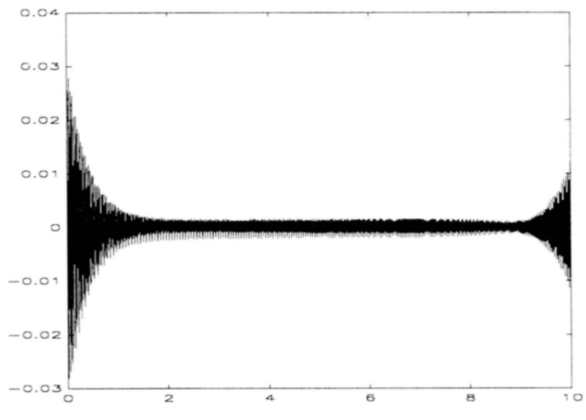

FIGURE 2 Waveform prediction error with harmonic noise.

RPM Determination

Given that the Kalman filter has a high capacity for tracking a target sine wave contaminated with high levels of noise, it becomes crucial that the instantaneous frequency, or RPM of the system be estimated with a high degree of precision, lest we track the wrong frequency. The normal way of estimating frequency is to use a tachometer or any other transducer that gives a clean signal with a periodic waveform that is directly related to the target sine wave. A typical signal would be the ignition pulses in a combustion engine.

Period Determination. To determine the instantaneous frequency from a tachometer channel, the customary procedure is to identify a recurrent event that defines the end of a period. Such events could be consecutive downcrossings of the mean, possibly with a certain hysteresis level to prevent premature indications from noise or higher harmonics. Experience indicates that each type of tachometer signal warrants some user setup to select the best period detection procedure.

The sampling interval Δt of the tachometer channel dictates directly the standard deviation of the estimate of the period because the error in the determination of any time domain event is uniformly distributed on an interval with length Δt. It is clearly desirable to sample the tach channel at very high frequencies to obtain a small statistical variability in the period estimation. If a high sampling rate is expensive or awkward to obtain, one possibility is to use sampling rate interpolation based on finite impulse response filters (see e.g., Oppenheim and Schafer, 1975). The usage of interpolation filters is predicated by the sampling theorem that says that any band limited signal that has been sampled at a frequency higher than twice the maximum frequency present in the signal may be reconstructed in the time domain to any desired degree of time resolution, albeit at computational expense. The noise floor of the tachometer channel should be considered when choosing an interpolation rate because the time resolution is only a component in the accuracy of the period estimation.

To illustrate the effect of sampling rates on the period estimates, a periodic signal that doubles in frequency in 10 s was analyzed for downcrossings at a coarse sampling rate of 200 Hz, corresponding to $\Delta t = 0.005$ s. The estimated period length is plotted in Fig. 3. The data were then resampled at 1,600 Hz ($\Delta t = 0.000625$ s) using a digital interpolation filter. Figure 4 shows the improvement in the period estimate with the increased time resolution.

Spline Smoothing of Period Estimates. Because all mechanical systems have some inertia and flexibility associated with them, it is reasonable to assume that the period will have certain smoothness properties as a function of time. We expect the period to be a continuous function, and in most cases also to have a continuous first derivative. Cubic splines satisfy these requirements, so we propose to subdivide the observation window into appropriate intervals and performing a least squares fit of a cubic spline to the estimated period function. Cubic splines are quite flexible interpolants and also do not have the numerical instabilities associated with high-order polynomials. For a thorough discussion on splines see de Boor (1978).

Other smoothing techniques are also possible, including moving average filters and even Kalman filtering, but our experience indicates that the spline smoothing is often sufficient. Figure 5 shows a spline function with three segments interpolating the period estimate of Figure 3 with the coarse time resolution.

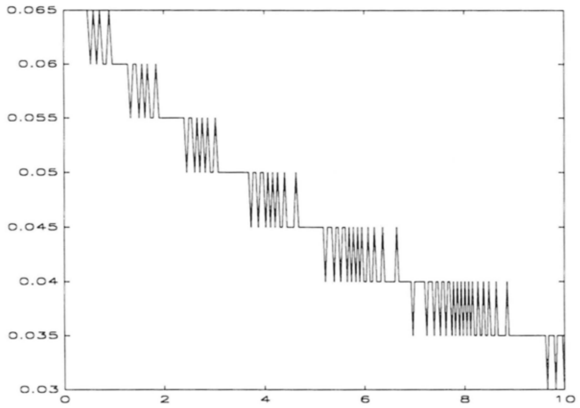

FIGURE 3 Period estimate with coarse sampling (Δt).

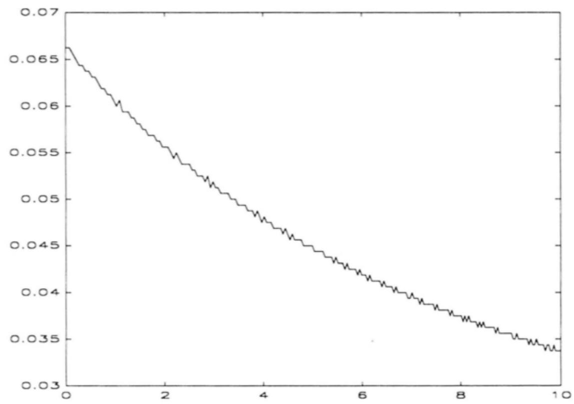

FIGURE 4 Period estimate with fine sampling ($\Delta t/8$).

EXAMPLES

We will illustrate the application of the Kalman order tracking to both analytical and acquired data. In the section treating the analytical data

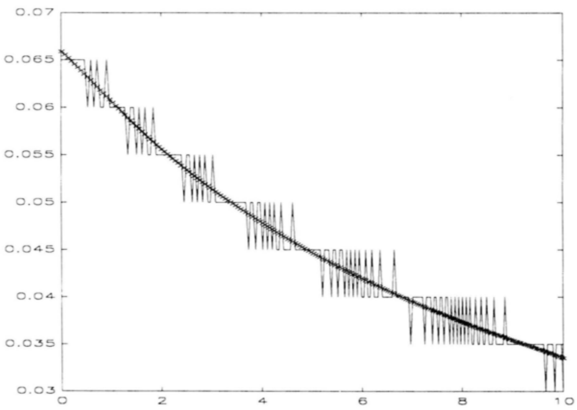

FIGURE 5 Period estimate with spline interpolant.

we will demonstrate various phenomena both for fast Fourier transform (FFT) techniques as well as for the Kalman filtering. Because we know the "right" answers for the analytical data, this is where we can develop intuition about the application of the Kalman filtering.

In the section with experimental automotive data we will restrict the exposition to a graphical comparison of the FFT and the Kalman results and some tentative conclusions.

Analytical Experiments

This section will look at the properties of the Kalman filter from several points of view. By inspection of Eq. (8) one can see that the inhomogeneity term $\varepsilon(n)$ has locally a zero mean as long as the amplitude of the target sine wave is constant. This translates into excellent tracking properties of signals with slowly varying amplitude almost independently of slew rate. As long as the RPM function is well determined then, the limiting factor would be amplitude tracking.

Analytical Square Wave. This example uses an analytical square wave that has harmonics of odd orders only and where each order has constant amplitude. The functional form in continuous time is given by

$$s(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \cos\left(2\pi \int_0^t \omega(t)\, dt\right). \quad (10)$$

A sampled version of this square wave with the first seven harmonics is shown in Fig. 6.

We now track the first harmonic of one case of this square wave where the frequency doubles in 10 s. Figure 7 shows the relative error in the tracking with weighting 1/10, giving a filter with broader frequency range acceptance than the filter shown in Fig. 8 that was generated with the weighting $r(n) = 1/50$. The latter choice of weighting is clearly superior for this example where the amplitude of the target harmonic is constant.

The next experiment was conducted with a constant frequency Kalman filter applied to the square wave of Fig. 6. This constant frequency was selected such that it would be crossed by the first two odd harmonics of this signal. Two runs were made, one with weighting 1/60 for rapid convergence, and one with weighting 1/150 for better frequency resolution. These two runs are plotted in Fig. 9 where it is seen that the data biased (1/60) weighting gives a higher dynamic range and better definition of the order crossings

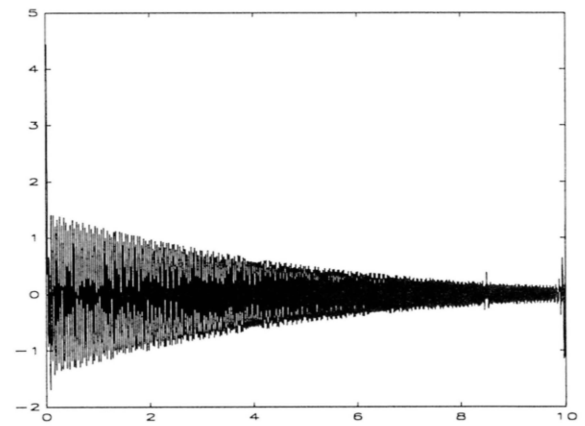

FIGURE 7 Percent tracking error with data bias ($r(n) = 1/10$).

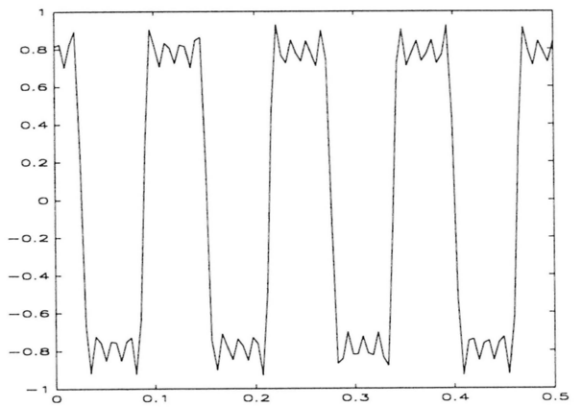

FIGURE 6 Sampled square wave with seven harmonics.

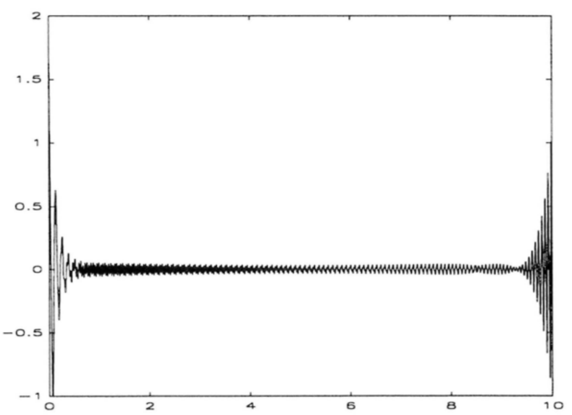

FIGURE 8 Percent tracking error with structural bias, ($r(n) = 1/50$).

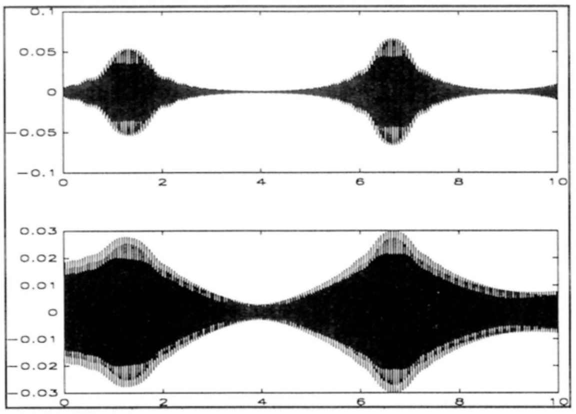

FIGURE 9 Harmonic run-through of constant filters with weighting 1/60 and 1/150.

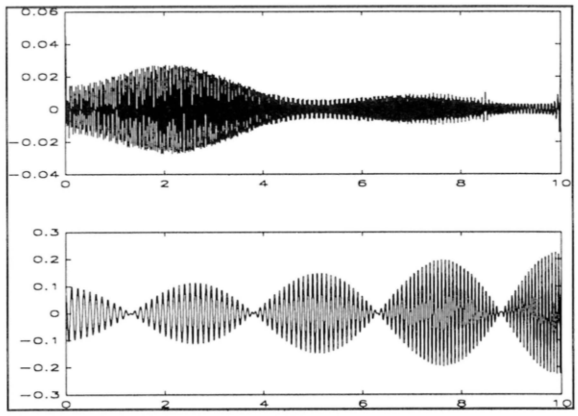

FIGURE 11 Tracking error of first harmonic with weighting 1/10 (top) and 1/100.

than the structure biased equation that predominantly shows the transients of slow convergence.

The square wave of Fig. 6 was then amplitude modulated with an offset squared low frequency sine as shown in Fig. 10. This data then has the form of Eq. (4), implying that the each harmonic is the superposition of two sine waves of frequencies slightly above and below the nominal frequency. The lowest harmonic waveform was tracked with weighting 1/10 and 1/100, respectively, and the tracking error plotted in Fig. 11. The data weighted (1/10) filter clearly gives the best results, due to its rapid convergence. For this particular data set where the only contaminants are the nuisance harmonics of the target sine wave, sharpness of the filter in frequency is not really required nor desirable because the harmonics are well separated in frequency and we do want to capture the energy in the amplitude modulation side bands. Adding white noise into the data would require a sharper filter, and hence a trend toward more structural weighting. The equivalent tracking errors with 50% Gaussian noise is shown in Fig. 12. In this plot we see the leakage of energy from the side bands with the data weighting (1/10) dominating the waveform prediction error; whereas the sharper filter with weighting 1/100 has an error roughly the same as for the signal without the broadband Gaussian noise.

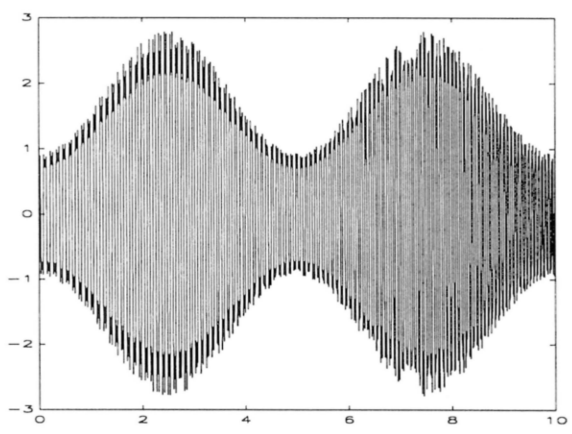

FIGURE 10 Modulated square wave.

Automotive Example

We will look here at experimentally acquired data and analyze these with both traditional digital FFT based methods and with the Kalman filtering approach.

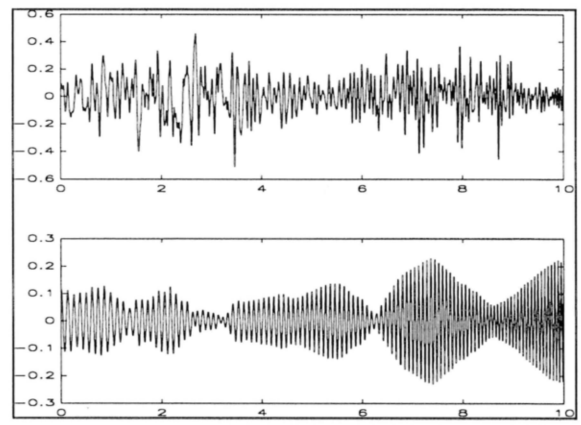

FIGURE 12 Tracking error of first harmonic of square wave with 50% Gaussian noise with weighting 1/10 (top) and 1/100 (bottom).

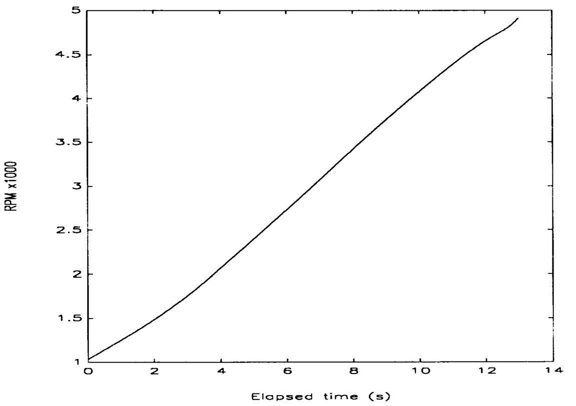

FIGURE 13 RPM of vehicle as function of elapsed time.

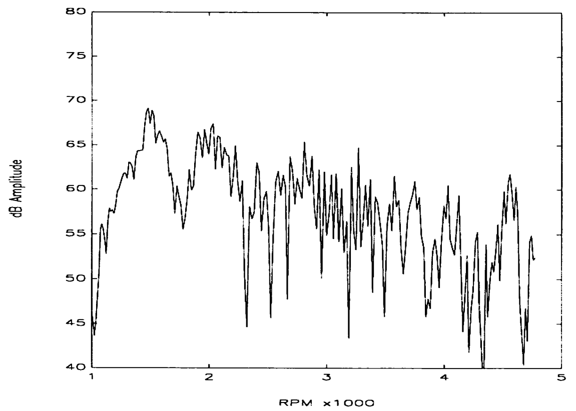

FIGURE 15 Decibel amplitude of fourth order (FFT based).

Vehicle Runup. The vehicle in question was powered by a four cylinder combustion engine and was taken through the speed range of 1000–4900 RPM in roughly 13 s as shown in Fig. 13.

A microphone was used as the response transducer, and using a FFT-based postprocessing system we extracted the level of the second order and fourth order as shown in Figs. 14 and 15. This plot was made using 80% overlap processing that tends to smear the frequency information, but also smoothes out the jitter due to short FFT block lengths.

We then tracked using the Kalman filter technique with the second and the fourth order as waveforms, and then, using the Vandermonde equation, Eq. (9), calculated amplitude as a function of RPM. The weighting function was set to 1/200 as a compromise between amplitude convergence and frequency discrimination. The orders are plotted in Figs. 16 and 17.

We notice a good overall agreement between the orders as estimated with FFTs and Kalman filters, but with a clear edge in resolution and dynamic range with the Kalman approach, especially for the fourth order.

CONCLUSIONS

Under rapid changes of operating conditions, the accurate analysis of harmonic and order components from noise and vibration measurements on rotating equipment remains a challenge for the automotive engineer.

The accuracy and resolution obtained with most digital and analog tracking filters will de-

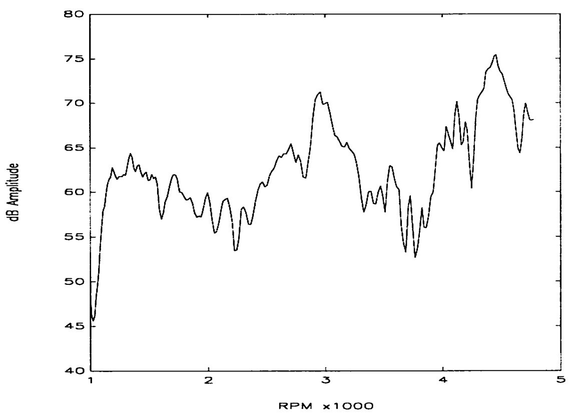

FIGURE 14 Decibel amplitude of second order (FFT based).

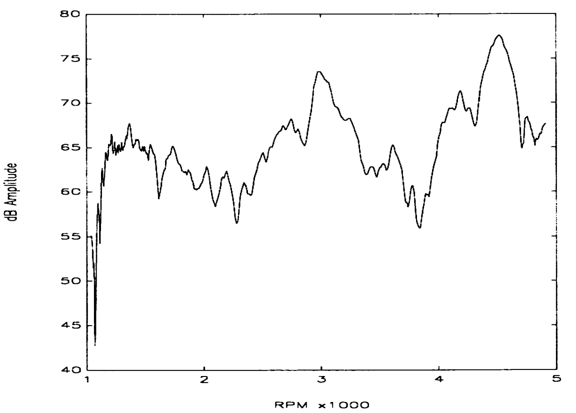

FIGURE 16 Decibel amplitude of second order (Kalman, filtering).

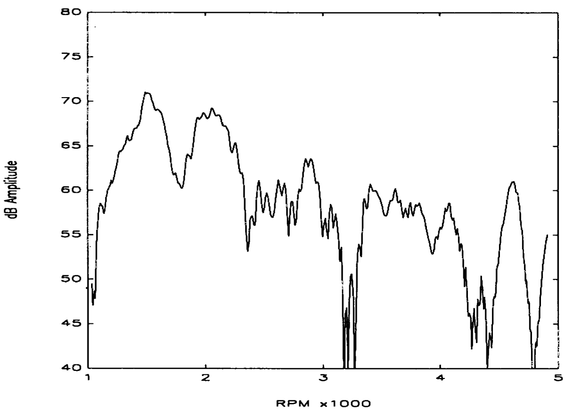

FIGURE 17 Decibel amplitude of fourth order (Kalman filtering).

grade quickly as the rate of change of a reference axle or RPM increases. The authors have developed a procedure for accurate analysis of harmonic and order components under such conditions. This procedure entails the transient recordings at constant sample rate of the signals, followed by analysis using an adaptive, nonstationary filter based on a Kalman filter formulation.

The tracking characteristics of these filters, i.e. the predicted signal amplitude versus time values versus exact signal amplitude versus time, can be tailored to track harmonics accurately with rapidly changing amplitude and phase over time, and under conditions of high slew rate. Also, as the analysis procedure is a true filtering method, it enables the extraction of order and harmonic information at a speed that equals the original sample rate of the signals.

REFERENCES

de Boor, C., 1978, A Practical Guide to Splines, Springer–Verlag, New York.

Giordano, A. A., and Hsu, F. M., 1985, Least Square Estimation with Applications to Digital Signal Processing, Wiley Interscience, New York.

Goodwin, G., and Sang Sin, K., 1984, Adaptive Filtering, Prediction and Control, Prentice–Hall, Englewood Cliffs, NJ.

Kalman, R. E., and Bucy, R. S., 1961, "New Results in Linear Filtering and Prediction Theory," *Transactions of the American Society of Mechanical Engineering, Journal of Basic Engineering,* Vol. 83, pp. 95–108.

Kalman, R. E., 1960, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the American Society of Mechanical Engineering, Journal of Basic Engineering,* Vol. 82, pp. 35–45.

Oppenheim, A. V., and Schafer, R. W., 1975, Digital Signal Processing, Prentice–Hall, Englewood Cliffs, NJ.

Potter, R., and Gribler, M. 1989, "Computer Order Tracking Obsoletes Older Methods," Proceedings of the 1989 SAE Noise and Vibration Conference, Traverse City, MI, pp. 63–67.

Vold, H., Kundrat, J., Rocklin, G. T., and Russell, R., 1982, "A Multi-Input Modal Estimation Method For Mini-Computers," SAE Paper Number 820194.