

GEAR AND BEARING DIAGNOSTICS USING NEURAL NETWORK-BASED AMPLITUDE AND PHASE DEMODULATION

E.C. Larson, D.P. Wipf, and B.E. Parker, Jr.
BARRON ASSOCIATES, INC.
3046A Berkmar Drive
Charlottesville, Virginia 22901-1444

Abstract: This paper is concerned with diagnostic methods for gears and bearings, the goals of which are to detect incipient structural and metallurgical faults, characterize their nature and severity, and isolate them to particular components. Such information is invaluable for prognostic purposes. The methodology presented herein focuses on the analysis of vibration signals induced by gears and bearings and, specifically, on a means of extracting from such vibration patterns amplitude and phase modulation signals. Such modulation terms, we conjecture, are simple indicators of the severity of a particular type of localized component defect. We show that such amplitude and phase modulation information can be extracted using a neural network computational methodology that relies on nothing more than knowledge of the bearing geometry and the frequency tones at which a given type of defect will manifest itself. We present numerical simulation examples and show that the technique requires extremely few *a priori* assumptions.

Key Words: Condition-based maintenance, Machinery diagnostics, Health monitoring, Incipient faults, Gears, Bearings, Modulation, Polynomial neural networks

Introduction: The topics addressed herein pertain to the general problem of machinery *prognostics*, i.e., estimating the remaining useful life of machinery components based on knowledge of their present condition and assumptions about future usage. The ability to diagnose particular machine components and understand the implications of diagnostic information vis-à-vis future usage of the machine is of immense practical importance in industrial, commercial, and military applications. Component failures contribute to the majority of machine operation safety incidents and are responsible for a substantial fraction of the on-going maintenance costs accruing to machinery usage. This underscores the need for *condition-based maintenance* (CBM), wherein knowledge regarding the *condition* of machinery, rather than fixed time intervals, is used to schedule maintenance, thus averting unnecessary inspections.

The quality of the diagnostic information acquired (i.e., fault detection, type and severity characterization, and component isolation) plays a paramount role in determining the efficacy of prognostication. Use of inaccurate diagnostic information in, say, a prognostic decisional algorithm driving a CBM policy for a fleet of helicopters may result in ill-fated missions. Prompt and accurate detection of structural and metallurgical faults in machinery components (e.g., gears and bearings) is an important category of such diagnostic information that is our focus herein. Almost all physical techniques for detecting the onset of such defects, e.g., fatigue cracks or surface spalling, involve the measurement of abnormal vibration signals transmitted through structural media due to the presence of the defect. Vibration sensors (e.g., accelerometers) may be used to measure these signals.

Sensitive spectral analysis and astute signal processing of such sensor data is critical, chiefly because in practical machinery operational scenarios, there are an enormous number of vibrations that interfere with one another and result in an observed signal that appears bewildering and uninformative. The main problem is therefore one of discovering the needle (i.e., spectral content that provides useful and pertinent information about an incipient fault) in the haystack (i.e., the raw vibration signal). A number of sophisticated techniques

have been proposed specifically for treating this very problem, including bispectral analysis [14], envelope spectral analysis [8], statistical moment analysis [9], time-frequency/time-scale analysis [2, 4], signal averaging [6], and signal template-based demodulation [7]. The latter approach is the one explored herein. In particular, the approach involves decomposing the vibration signal from a complex system, such as a gearbox, into amplitude modulation and phase modulation signals. The presence of such may be justified on physical grounds since amplitude and phase modulation arise, respectively, from radial and tangential forces exerted during impact between a defective region of one surface and the surface of another contacting member [11]. To present our signal demodulation approach for gears and bearings, we first work through its application to a gear problem. The method is then extended to the more complicated problem of bearing vibrations. The basic methodology and key assumptions of the proposed signal-processing approach are discussed. The computational complexity of the method is addressed, and a means of carrying out the computational tasks using neural network methods is presented.

Joint Amplitude/Phase Demodulation of Gear Vibrations: Consider a defect-free gear. Under normal, constant-load operation, the vibration signal, $y(t)$, emitted by the gear is periodic in the gear-mesh frequency, f_{gm} , *viz.*,

$$y(t) = \sum_{k=-\infty}^{\infty} y_k e^{ikN\omega_{\text{gm}}t} \quad (1a)$$

in which $\omega_{\text{gm}} = 2\pi f_{\text{gm}}$ and N is the number of teeth. The existence of vibrations in a defect-free gear reflects the facts that the structural materials are not infinitely rigid and that the transmission force varies slightly with the angular displacement of the teeth in contact with the load rack. It is highly advantageous, as a number of authors [6, 7] have pointed out, to work with the angular displacement, θ , of the gear rather than time, t , as the independent variable. This compensates for variations in the rotational velocity, $d\theta/dt$, of the gear and facilitates averaging the signal over many gear revolutions to improve the signal-to-noise ratio. We thus recast Eq. 1a as

$$y(\theta) = \sum_{k=-\infty}^{\infty} y_k e^{ikN\theta} \quad (1b)$$

which indicates that under normal loading conditions, a defect-free gear exhibits a vibration signal of angular period $2\pi/N$, since there are N gear-mesh events per gear revolution.

In the presence of a localized structural or metallurgical defect, a perturbed vibration signal is observed. The signal exhibits both amplitude and phase modulation and may be modeled [7] in the mathematical form

$$y(\theta) = A(\theta) \cdot \sum_{k=-\infty}^{\infty} X_k e^{ik[N\theta+B(\theta)]} \quad (2)$$

in which $A(\theta)$ and $B(\theta)$ both have period 2π . In the limiting case of an incipient fault, the perturbed signal should be only infinitesimally different from the unperturbed signal in Eq. 1b, which implies that $A(\theta) \rightarrow 1$ and $B(\theta) \rightarrow 0$ in the incipient fault limit. As a defect develops progressively over the life of a particular gear, $A(\theta)$ and $B(\theta)$ may be expected to grow slowly, but the X_k 's will remain constant (i.e., the X_k 's are independent of defect severity). It thus follows that for a defect-free gear,

$$y(\theta) = \sum_{k=-\infty}^{\infty} X_k e^{ikN\theta} \quad (3)$$

as in Eq. 1b. If the vibration signature of a gear element is measured early in life before the onset of degradation effects, the X_k 's can be computed directly via a discrete Fourier

transform (DFT). Retained knowledge of these X_k values facilitates simultaneous amplitude and phase demodulation of the perturbed signal in Eq. 2.

As a result of a localized tooth defect, however small, the measured vibration signal, $y(\theta)$ now has angular period 2π rather than $2\pi/N$, since the defective tooth makes contact with the load once per gear revolution. Moreover, even in the case of multiple tooth defects, $y(\theta)$ has angular period 2π . The prefactor, $A(\theta)$, represents an amplitude modulation and may be expanded as a discrete-tone Fourier series, *viz.*,

$$A(\theta) = \sum_{k=-\infty}^{\infty} A_k e^{ik\theta}. \quad (4)$$

Since the amplitude and phase modulation effects need not necessarily be the same at the various tooth-mesh harmonics, $A(\theta)$ is *not* necessarily equal to the amplitude of $y(\theta)$ that one would obtain from computation of the Hilbert transform. The expression

$$v(\theta) = \sum_{k=-\infty}^{\infty} X_k e^{ik[N\theta+B(\theta)]} \quad (5)$$

that accompanies $A(\theta)$ in Eq. 2 also has angular period 2π and may be expanded as a Fourier series, *viz.*,

$$v(\theta) = \sum_{r=-\infty}^{\infty} V_r e^{ir\theta}. \quad (6)$$

Note that $v(\theta)$ is of the same phase as $y(\theta)$. This important result follows from a theorem [1, 10] that states that if $f(t)$ is a low-frequency signal and $g(t)$ is a high-frequency signal lying completely above $f(t)$ in the frequency domain with no overlap, then the Hilbert transform of the product signal decomposes as

$$H[f(t)g(t)] = f(t) H[g(t)] \quad (7)$$

where the operator H denotes the Hilbert transform. It follows that if $f(t)$ (corresponding to $A(\theta)$) is in the form a real-valued signal and $g(t)$ (corresponding to $v(\theta)$) is in the form of an analytic signal (*viz.*, of the form $s(t) + iH[s(t)]$) then the instantaneous phase of $f(t)g(t)$ is identically equal to that of $g(t)$. Note that $A(\theta)$, for an incipient defect, is a narrowband signal centered at frequency zero, whereas the terms in the expression for $v(\theta)$ (Eq. 5) are narrowband signals centered at frequency kN (in which X_0 can be set to zero without loss of generality).

Since the phase of $v(\theta)$ is equal to that of $y(\theta)$, which is an observable signal, it follows that, given the X_k 's, we have sufficient information to compute $B(\theta)$ uniquely. The solution procedure begins with the phase modulation, $B(\theta)$, which, like $A(\theta)$, is of angular period 2π , *viz.*,

$$B(\theta) = \sum_{q=1}^{\infty} B_q \sin(q\theta - \beta_q) \quad (8)$$

where the B_q 's are real-valued. $B(\theta)$ appears in the expression for $v(\theta)$ in Eq. 5 in a factor of the form

$$W_k(\theta) \equiv e^{ikB(\theta)} \quad (9)$$

which, as a periodic function of angular period 2π , may be expressed as a Fourier series

$$W_k(\theta) = \sum_{h=-\infty}^{\infty} M_h(k, \underline{B}, \underline{\beta}) e^{ih\theta} \quad (10)$$

in which $M_h(k, \underline{B}, \underline{\beta})$ may be computed simply by taking the DFT of W_k . The Fourier coefficients depend not only on the Fourier index, h , but also the Fourier amplitudes and

phases (collectively denoted as \underline{B} and $\underline{\beta}$ respectively) of $B(\theta)$ and the gear-mesh harmonic index, k . Note that in the incipient fault limit, $B(\theta) \approx 0$, from which it follows that $M_h(k, \underline{B}, \underline{\beta}) \approx \delta_{h,0}$.

Having hypothesized the functional form of $B(\theta)$, i.e., \underline{B} and $\underline{\beta}$, the $M_h(k, \underline{B}, \underline{\beta})$'s can be computed, and we obtain an expression relating it and the X_k 's to the V_r 's, *viz.*,

$$\sum_{r=-\infty}^{\infty} V_r e^{ir\theta} = \sum_{k=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} X_k M_h(k, \underline{B}, \underline{\beta}) e^{i(kN+h)\theta}. \quad (11)$$

The V_r 's are thus computable as

$$V_r = \sum_{k=-\infty}^{\infty} X_k M_{r-kN}(s, \underline{B}, \underline{\beta}). \quad (12)$$

Having computed all of the V_r 's, we can thus reconstruct the partial signal $\hat{v}(\theta)$, which is analytic and therefore has a well-defined instantaneous phase, which we denote as $\hat{\phi}_v(\theta)$. The caret denotes an estimate in that we have surmised $B(\theta)$, and the results reflect the accuracy of this approximation or educated guess. To judge whether the surmised $B(\theta)$ is accurate, the estimated phase, $\hat{\phi}_v(\theta)$ is required to match the actual phase, $\phi_v(\theta)$, which can be inferred directly from the observed vibration signal, $y(\theta)$. The degree to which the instantaneous phases over the angular interval $(0, 2\pi)$ coincide reflects the accuracy of the hypothesized phase modulation signal, $B(\theta)$. The discrepancy, as a *functional* of the function $B(\theta)$, may be quantified by way of the squared error over $(0, 2\pi)$, *viz.*,

$$J[B] = \int_0^{2\pi} [\hat{\phi}_v(\theta) - \phi_v(\theta)]^2 d\theta. \quad (13)$$

To determine $B(\theta)$ requires a search algorithm to find optimal values for \underline{B} and $\underline{\beta}$. Since \underline{B} and $\underline{\beta}$ both are infinite sets, however, it is feasible to proceed by approximating \underline{B} and $\underline{\beta}$ as finite Fourier series and gradually including additional harmonics in more refined approximations. We later show, following the succeeding discussion on the extension of the methodology to bearings, how this numerical procedure can be implemented using neural network algorithms to enhance computational speed while retaining excellent accuracy.

As in [7], the amplitude signal, $A(\theta)$, may then be computed at the very end of the entire procedure as $y(\theta)/v(\theta)$, once the optimal $B(\theta)$ functional form has been found and $v(\theta)$ has been computed.

Joint Amplitude/Phase Demodulation of Bearing Vibrations: In a defect-free roller bearing under constant loading conditions, the vibration signal, $y(t)$, assumes the same form as in the gear problem, *viz.*,

$$y(\theta) = \sum_{k=-\infty}^{\infty} y_k e^{ikN\theta} \quad (14)$$

where N is the number of balls and θ is the angular displacement of the cage. The perturbed signal that arises in the presence of a localized structural defect also assumes the form

$$y(\theta) = A(\theta) \cdot \sum_{k=-\infty}^{\infty} X_k e^{ik[N\theta+B(\theta)]} \quad (15)$$

as in Eq. 2. Unlike in the gear problem, however, $A(\theta)$ and $B(\theta)$ have more complicated frequency-domain properties. They are both biperiodic with angular periods $2\pi\gamma_q$ and $2\pi\gamma_d$,

where

$$\gamma_q = f_{\text{cage}}/f_{\text{load-pass}} \quad (16a)$$

$$\gamma_d = f_{\text{cage}}/f_{\text{defect}} \quad (16b)$$

The biperiodic form is due to the physical circumstance that the perturbative excitation force, $e(t)$, resulting from contact between a surface defect and other members of the bearing assembly, is the product of an impact amplification function, $q(t)$, and an impact event function, $d(t)$ [8]. The latter function is a train of impulses that indicate the instants at which the defective surface region contacts other members. The spacing between the impulses is determined by the frequency, f_{defect} , characteristic of the particular type of defect. The three most common examples are

$$f_{\text{BPIR}} = \frac{1}{2}Nf_{\text{shaft}} \left(1 + \frac{d_r}{d_p} \cos \alpha \right) \quad \text{for an inner-race defect;} \quad (17a)$$

$$f_{\text{BPOR}} = \frac{1}{2}Nf_{\text{shaft}} \left(1 - \frac{d_r}{d_p} \cos \alpha \right) \quad \text{for an outer-race defect;} \quad (17b)$$

$$f_{\text{roller spin}} = \frac{1}{2}f_{\text{shaft}} \left(\frac{d_p}{d_r} \right) \left(1 - \frac{d_r^2}{d_p^2} \cos^2 \alpha \right) \quad \text{for a rolling-element defect,} \quad (17c)$$

where d_r is the diameter of the rolling elements, d_p is the pitch diameter, and α is the contact angle. The cage frequency, f_{cage} , is equal to f_{BPOR}/N (assuming that the outer race is stationary and the inner race is rotating at the shaft frequency) and is analogous to the gear-mesh frequency in the gear problem.

The impact amplification function, $q(t)$, reflects the fact that an impact event involving a rolling element in the load zone is more consequential than an impact involving an unloaded rolling element. $q(t)$ is therefore periodic at the frequency $f_{\text{load-pass}}$, which is the frequency at which the defective surface passes through the load zone, *viz.*,

$$f_{\text{load-pass}} = f_{\text{shaft}} \quad \text{for an inner-race defect;} \quad (18a)$$

$$f_{\text{load-pass}} = 0 \quad \text{for an outer-race defect;} \quad (18b)$$

$$f_{\text{load-pass}} = f_{\text{cage}} \quad \text{for a rolling-element defect.} \quad (18c)$$

All of the various frequencies in the above equations may be expressed as products of the cage frequency times a factor that depends solely on the bearing geometry. Hence, γ_q and γ_d are independent of variations in the shaft rotational speed.

The periodicity of $A(\theta)$ and $B(\theta)$ depends strongly on the type of defect. For an outer-race defect, $\gamma_q \rightarrow \infty$ and $\gamma_d = f_{\text{cage}}/f_{\text{BPOR}} = 1/N$. An outer-race defect is therefore equivalent dynamically to a gear defect, as treated in the preceding section. For a rolling-element defect, $\gamma_q = 1$, but $\gamma_d = f_{\text{cage}}/f_{\text{roller spin}}$. The frequency ratio $f_{\text{cage}}/f_{\text{roller spin}}$ is equal to

$$\frac{f_{\text{cage}}}{f_{\text{roller spin}}} = \left(\frac{d_r}{d_p} \right) \left(1 + \frac{d_r}{d_p} \cos \alpha \right)^{-1} \quad (19)$$

which is generally an irrational number. For an inner-race defect, γ_d and γ_q are both irrational. The presence of multiple frequencies and the irrationality of one or both γ 's makes the mathematics of gear and bearing problems significantly different. If there are multiple faults of different types, three or more γ 's may arise.

To demonstrate the demodulation procedure for a bearing, we focus on the special case of an inner-race defect, wherein the vibration signal assumes a triperiodic Fourier series form, *viz.*

$$y(\theta) = \sum_{k=-\infty}^{\infty} \sum_{d=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} Y_{k,d,q} e^{i[kN+\gamma_d d+\gamma_q q]\theta} \quad (20)$$

with $Y_{k,d,q} \rightarrow Y_{k,0,0} \delta_{d,0} \delta_{q,0}$ in the incipient fault limit. Similarly, $v(\theta)$ becomes

$$v(\theta) = \sum_{r=-\infty}^{\infty} \sum_{d=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} V_{r,d,q} e^{i[r+\gamma_d d + \gamma_q q]\theta}. \quad (21)$$

The amplitude function becomes

$$A(\theta) = \sum_{d,q} A_{d,q} e^{i(\gamma_d d + \gamma_q q)\theta} \quad (22)$$

as in Eq. 4, and the phase modulation function becomes

$$B(\theta) = \sum_{d=1}^{\infty} \sum_{q=1}^{\infty} B_{q,d} \sin [(\gamma_d d + \gamma_q q)\theta - \beta_{q,d}]. \quad (23)$$

$W_k(\theta)$ becomes

$$W_k(\theta) = \sum_{h_d=-\infty}^{\infty} \sum_{h_q=-\infty}^{\infty} M_{h_d,h_q}(k, \underline{B}, \underline{\beta}) e^{i(\gamma_d h_d + \gamma_q h_q)\theta} \quad (24)$$

with

$$M_{h_d,h_q}(k, \underline{B}, \underline{\beta}) = \delta_{h_d,0} \delta_{h_q,0} \quad (25)$$

in the incipient fault limit. The $M_{h_d,h_q}(k, \underline{B}, \underline{\beta})$'s can be computed via a DFT, as in the gear case. However, since $W_k(\theta)$ is no longer periodic over the interval $(0, 2\pi)$, the DFT is somewhat less clean numerically. This means that either a very large Fourier window must be used or that $M_{h_d,h_q}(k, \underline{B}, \underline{\beta})$ may have to be computed analytically by appealing to a Bessel function expansion of $W_k(\theta)$. This is done by substituting Eq. 23 into Eq. 9, which yields

$$W_k(\theta) = \prod_{d=1}^{\infty} \prod_{q=1}^{\infty} e^{ikB_{d,q} \sin[(\gamma_d d + \gamma_q q)\theta - \beta_{q,d}]} \quad (26a)$$

$$= \prod_{d=1}^{\infty} \prod_{q=1}^{\infty} \sum_{s=-\infty}^{\infty} J_s(kB_{d,q}) e^{-ikB_{d,q} \sin \beta_{d,q}} \cdot e^{is(\gamma_d d + \gamma_q q)\theta} \quad (26b)$$

in which the mathematical identity

$$e^{ix \sin \theta} = \sum_{s=-\infty}^{\infty} J_s(x) e^{is\theta} \quad (27)$$

was invoked. It follows from inspection of Eq. 26 that the Fourier coefficients are

$$M_{h_d,h_q}(k, \underline{B}, \underline{\beta}) = \sum_s J_s(kB_{d,q}) e^{-ikB_{d,q} \sin \beta_{d,q}} \delta_{sd,h_d} \delta_{sq,h_q} \quad (28)$$

in which the summation is over all common factors, s , that divide both h_d and h_q .

Having computed $M_{h_d,h_q}(k, \underline{B}, \underline{\beta})$, the X_k 's can be computed by comparing the left- and right-hand sides of

$$\sum_{r,d,q} V_{r,d,q} e^{i[r+\gamma_d d + \gamma_q q]\theta} = \sum_{k,d,q} X_k M_{h_d,h_q}(k, \underline{B}, \underline{\beta}) e^{i(kN + \gamma_d h_d + \gamma_q h_q)\theta}. \quad (29)$$

Since the γ 's are irrational, it follows that the only combinations of indices that enable the exponential factors to agree are in pairs of corresponding terms on the two sides of Eq. 29, such that $r = kN$, $d = h_d$, and $q = h_q$. It follows that the X_k 's can be computed readily from the relation

$$V_{k,d,q} = X_k M_{d,q}(k, \underline{B}, \underline{\beta}) \quad (30)$$

which, unlike in the gear case, becomes exact owing to the irrationality of the γ 's. This result actually makes the irrationality of γ helpful, not a hindrance, since it simplifies the computational burden. For the gear problem, this simplification is also valid if N is large, in which case N is "almost irrational" in the sense that if it is written as p/q , where p and q are integers, p or q must be large.

The instantaneous phases of $y(\theta)$ and $v(\theta)$, as in the gear application, are required to be equal, and a search algorithm must be used, as before. Once an optimal functional form for $B(\theta)$ is found, the amplitude, $A(\theta)$, is computed at the very end.

Amplitude/Phase Demodulation Neural Network: Numerical solution of the vibration amplitude and phase demodulation problem can be realized by means of neural networks, as described herein. *GNOSIS* (*Generalized Networks for Optimal Synthesis of Information Systems*) [5, 12, 13] is a software package developed by Barron Associates, Inc., for synthesis and implementation of a wide variety of artificial neural network paradigms that employ various alternative types of basis functions. Specifically, a *GNOSIS* network can be designed to automate the algorithmic computation procedure of the flowchart in Fig. 1, which involves a search algorithm.

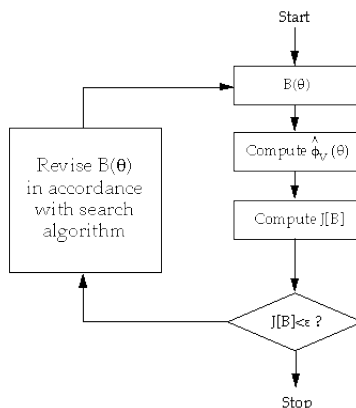


Figure 1: Search Algorithm Procedure Realized Via *GNOSIS*

The *GNOSIS* network estimates the function $\hat{\phi}_v$ over the angular range $0 \leq \theta < 2\pi$ by optimizing the phase modulation, $B(\theta)$, in such a way that the objective function

$$J[B] = \int_0^{2\pi} \left[\hat{\phi}_v(\theta) - \phi_v(\theta) \right]^2 d\theta \quad (31)$$

is minimized globally. Several numerical simulation tests were performed to construct *GNOSIS* algorithms to implement the flowchart algorithm in Fig. 1.

Several synthetic vibration signals of the form

$$y(\theta) = A(\theta) \cdot \text{Re} \left\{ \sum_{k=1}^K X_k e^{ik[N\theta + B(\theta)]} \right\}$$

$$= A(\theta) \cdot \sum_{k=1}^K \{(\text{Re } X_k) \cos(k[N\theta + B(\theta)]) - (\text{Im } X_k) \sin(k[N\theta + B(\theta)])\} \quad (32)$$

were generated for a gear with $N = 10$ teeth. K denotes the maximum number of gear-mesh harmonic terms included in the simulated signal. Synthetic $A(\theta)$ and $B(\theta)$ functions were computed as

$$A(\theta) = \sum_{p=0}^P [(\text{Re } A_p) \cos p\theta - (\text{Im } A_p) \sin p\theta] \quad (33a)$$

$$B(\theta) = \sum_{q=1}^Q [(\text{Re } B_q) \cos q\theta - (\text{Im } B_q) \sin q\theta] \quad (33b)$$

where P and Q denote the maximum number of terms in $A(\theta)$ and $B(\theta)$ respectively.

The analytic continuation, $y_a(\theta)$, of the observed signal, $y(\theta)$ was computed via the Hilbert transform, *viz.*

$$y_a(\theta) = y(\theta) + i H[y(\theta)] = |y(\theta)| e^{i\phi_y(\theta)}. \quad (34)$$

However, instead of computing the $\phi_y(\theta)$ directly, the *GNOSIS* networks were designed to model the real and imaginary parts of the normalized signal, *viz.*,

$$\cos \phi_y(\theta) = y(\theta)/|y(\theta)| \quad (35a)$$

$$\sin \phi_y(\theta) = H[y(\theta)]/|y(\theta)|. \quad (35b)$$

The normalized signals, unlike $\phi_y(\theta)$ itself, are continuous functions of θ , which facilitates the *GNOSIS* computations.

A feedforward network structure suitable for modeling the normalized signals can be realized using polynomial basis functions with trigonometric and linear post-transformations. A block diagram of the basic network structure is depicted in Fig. 2.

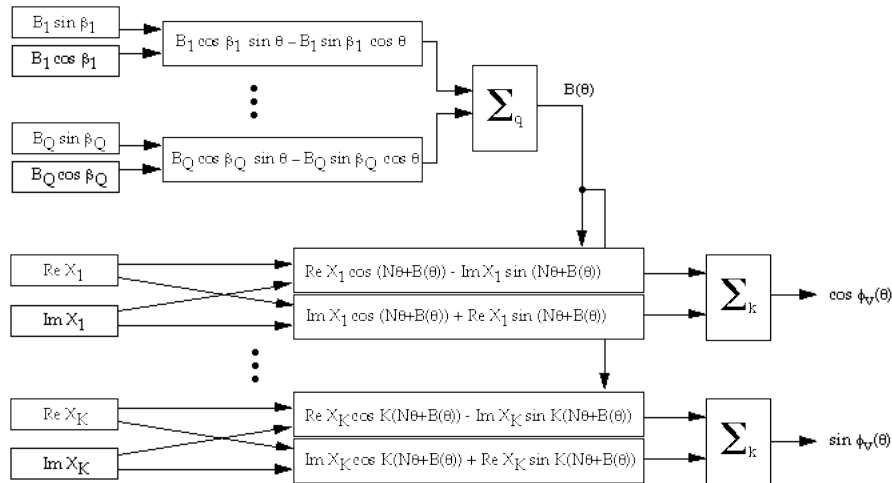


Figure 2: *GNOSIS* Network for Vibration Demodulation

The neural network algorithm for modeling the vibration signal assumed the *structural form* of a phase-modulated $v(\theta)$ signal; the Levenberg-Marquardt algorithm [12, 13] within *GNOSIS* was used to fit coefficients to the training signal. This approach is very efficient computationally. The versatility of *GNOSIS*, with respect to implementing alternative

types of basis functions and structuring networks in innovative ways pursuant to the needs of specialized applications, provides a strong incentive for adopting the neural network paradigm for the present application.

A network was trained on the signal $y(\theta)$ with the angular interval $(0, 2\pi)$ discretized into $2^{11} = 2,048$ sample points. Between ten and twenty iterations were required for convergence of the computed $B(\theta)$. Fig. 3 provides a comparison of the actual and *GNOSIS*-computed $B(\theta)$ signals for a particular sample problem with $K = 6$, $P = 4$ and $Q = 7$. In the figure, the computed $B(\theta)$ is sufficiently accurate that it is visually indistinguishable from the actual signal at the scale shown.

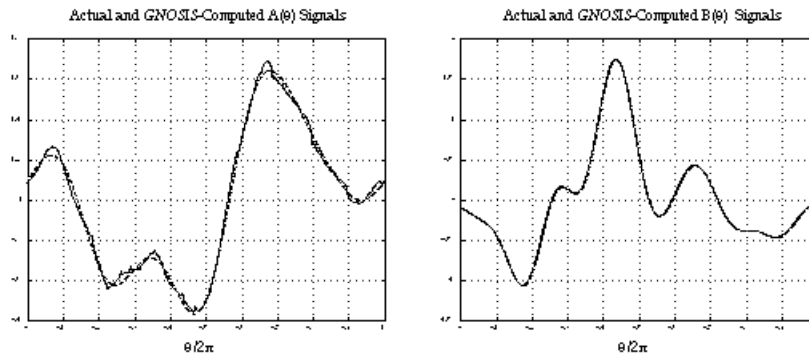


Figure 3: $A(\theta)$ (left) and $B(\theta)$ (right) Signals

Note that the X_k 's are included in the network structure simply as additional degrees of freedom independent of the parameters defining $B(\theta)$. Very significantly, it was found that no *a priori* knowledge of their values was necessary. This implies that diagnostic knowledge from the early life of the gear element, as previously discussed, is actually *not* necessary after all!

A second major assumption that was made in the introductory discussion was that the $A(\theta)$ and $B(\theta)$ signals must both be sufficiently narrowband to the extent that the conditions of Eq. 7 are satisfied. Specifically, this requires that the bandwidths satisfy the inequality

$$\mathcal{B}_A + (\delta B + 1)\mathcal{B}_B < f_{\text{gm}} \quad (36)$$

where \mathcal{B} denotes bandwidth and δB is the peak absolute value of $B(\theta)$, i.e., the maximum phase deviation introduced by $B(\theta)$. Eq. 36 follows from Carson's rule [3] and the fact that $v_1(\theta)$ contains the lowest frequency components of $v(\theta)$. It is important to note that in the case of $B(\theta)$, Eq. 36 not only imposes a bandwidth restriction, but also a magnitude restriction.

Fig. 4 displays the spectrum of $v(\theta)$ for the particular test case in Fig. 3. The spectrum contains significant modulation spreading into the baseband region occupied by $A(\theta)$, in violation of the conditions of Eq. 36. However, as exemplified by the good accuracy in Fig. 3, the assumption of narrowband $A(\theta)$ and $B(\theta)$ can also be lifted! Once $v(\theta)$ is computed, $A(\theta)$ can be obtained from $A(\theta) = y(\theta)/v(\theta)$. The left plot in Fig. 3 displays the actual (solid) and computed (dashed) $A(\theta)$ signals; the visible discrepancy is attributable to the failure of Eq. 7 to hold in the present case. The phase of the $v(\theta)$ signal computed in this first pass, however, can be used to apply the algorithm in a second pass to obtain incrementally better $B(\theta)$, X_k 's, and $A(\theta)$. The process can be repeated until a self-consistent $A(\theta)$ is obtained. Alternatively, the computed $A(\theta)$ can be low-pass filtered to obtain a better approximation.

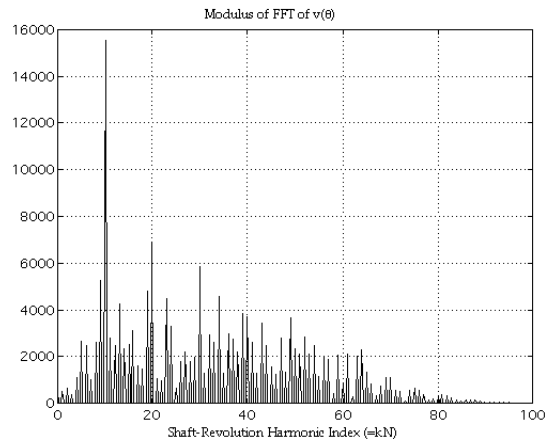


Figure 4: Fourier Spectrum of $y(\theta)$ Signal

Conclusion: We have presented a diagnostic neural network algorithm for decomposing gear and bearing vibration signals into amplitude and phase modulation components. Surprisingly few *a priori* assumptions are necessary with the approach.

References:

- [1] Bedrosian, E., "A product theorem for Hilbert transforms," *Proc. of the IEEE*, Vol. 51, 1963, pp. 868-869.
- [2] Brotherton, T., T. Pollard, and D. Jones, "Applications of time-frequency and time-scale representations to fault detection and classification," *Proc. IEEE-SP Int'l. Symp. on Time-Frequency and Time-Scale Analysis*, Oct. 4-6, 1992, Victoria, BC, Canada.
- [3] Couch, II, Leon W., *Digital and Analog Communication Systems*, 4th ed., 1993, p. 332. (MacMillan: New York)
- [4] Forrester, B.D., "Use of the Wigner-Ville distribution in helicopter fault detection," *ASSPA '89, Signal Processing, Theories, Implementations, and Applications* (R.F. Barrett, ed.), Adelaide, Australia, Apr. 17-19, 1989, pp. 78-82.
- [5] Barron Associates, Inc., *Generalized Networks for Optimal Synthesis of Information Systems*, Users' Manual, Version 2.1, Oct. 1996.
- [6] Jong, J., J. McBride, J. Jones, T. Fiorucci, and T. Zoladz, "Synchronous phase averaging method for machinery diagnostics," *Integrated Monitoring, Diagnostics, and Failure Prevention - Proc. Joint Conf. 50th Mtg. of the MFPT*, Mobile, AL, Apr. 22-26, 1996, pp. 441-451. (Vibration Institute: Willowbrook, IL)
- [7] Ma, J. and C.J. Li, "A new approach to gear demodulation and its application to defect detection," *Advanced Materials and Process Technology for Mechanical Failure Prevention - Proc. 48th Mtg. of the MFPT*, Wakefield, MA, Apr. 19-21, 1994, pp. 43-55. (Vibration Institute: Willowbrook, IL)
- [8] McFadden, P.D. and J.D. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing," *J. of Sound and Vibration*, Vol. 96, No. 1, 1984, pp. 69-82.
- [9] McFadden, P.D. and J.D. Smith, "A signal processing technique for detecting local defects in a gear from the signal average of the vibration," *Proc. Inst. Mech. Engrs.*, Vol. 199, No. C4, 1985, pp. 287-292.

- [10] Rihaczek, A.W., "Hilbert transforms and the complex representation of real signals," *Proc. of the IEEE*, Vol. 54, 1966, pp. 434-435.
- [11] Swanson, D.C., private communication.
- [12] Ward, D.G., B.E. Parker, Jr., and R.L. Barron, *Principles of Function Estimation Using Artificial Neural Networks*, Barron Associates, Inc. Technical Memorandum, Sept. 1993.
- [13] Ward, D.G., "Generalized networks for complex function modeling," *Proc. IEEE Systems, Man & Cybernetics (SMC-94) Conf.*, Oct. 2-5, 1994.
- [14] Zoladz, T., E. Earhart, and T. Fiorucci, *Bearing Defect Signature Analysis Using Advanced Nonlinear Signal Analysis in a Controlled Environment*, CDDF Final Report, Project No. 93-10, NASA Technical Memorandum 108491, May 1995.

Acknowledgments: This investigation was funded by the Office of Naval Research under Contract N00014-95-C-0413. The authors thank Dr. Thomas M. McKenna for his kind support of this work. Gratitude is also expressed to Dr. C.J. Li of the Rensselaer Polytechnic Institute and Dr. J. Ma of the New York Institute of Technology for their inspiration for, and critical review of, our work. The authors, however, assume full responsibility for any errors in the present work.