Chapter 11
Designing Your Own Wavelet

To achieve effective signal signature extraction, Chap. 10 introduced several quantitative measures for selecting appropriate base wavelets from a pool of available wavelet families, such as Daubechies, Myer, and Morlet wavelets. This chapter introduces a complimentary technique focusing on wavelet customization. The goal is to design a wavelet that is specifically adapted to the signal of interest. Because such a customized wavelet would have a higher degree of matching with the signal than other wavelets, the effectiveness of signature extraction will improve.

11.1 Overview of Wavelet Design

Researchers have studied various techniques for designing base wavelets. In the late 1980s to early 1990s, Daubechies' work has led to the publication of orthonormal (Daubechies 1988) and biorthonormal (Cohen et al. 1992) base wavelets with compact support. These wavelets are independent of the signal to be analyzed. Tewfik et al. (1992) have developed cost functions for finding the optimal orthonormal wavelet basis to represent a specified signal within a finite number of scales. Their work has been extended by assuming band limited signals and finding the optimal M-band wavelet basis within a finite number of scales (Gopinath et al. 1994), for representing a desired signal. During the same period, Aldroubi and Unser (1993) proposed a method to match a wavelet basis to a desired signal by either projecting the desired signal onto an existing wavelet basis, or transforming the wavelet basis under certain conditions such that the error norm between the desired signal and the new wavelet basis is minimum. Recently, Chapa and Rao (2000) have developed two sets of equations for designing a wavelet directly from a signal of interest. The first set of equations derives expressions for continuously matched wavelet spectrum amplitudes, whereas the second set provides a direct discrete algorithm for calculating approximations to the optimal complex wavelet spectrum. By formulating wavelet design as a constrained optimization problem and then solving it by converting the optimization problem into an iterative line-search problem through a first-order parameterization of the perfect reconstruction

R.X. Gao and R. Yan, *Wavelets: Theory and Applications for Manufacturing*,
DOI 10.1007/978-1-4419-1545-0_11,

constraint, a signal-adapted, biorthogonal filter banks of finite length was constructed by Lu and Antoniou (2001). Later, Shark and Yu (2003) proposed a genetic algorithm-based design method to construct orthonormal wavelet filter banks with an optimal shift-invariant property. On the basis of a generalized Mexican-hat function, the authors also designed a new class of continuous wavelets for arbitrary transient signals (Shark and Yu 2006), where signal matching is achieved by minimizing the spectral difference between the reference signal and the generalized Mexican-hat wavelet. Gupta et al. (2005a) have proposed to construct wavelets that are matched to a given signal in the statistical sense. The main idea is to first estimate a high-pass wavelet filter from the statistics of the signal, and then obtain a FIR/IIR biorthogonal perfect reconstruction filter bank. This leads to the construction of a statistically matched wavelet. The authors have also designed both biorthogonal and semiorthogonal wavelet from a signal by maximizing projection of the signal onto successive scaling subspaces while minimizing energy of the signal in the wavelet subspace (Gupta et al. 2005b). Using the same idea, Guido et al. (2006) designed a spikelet wavelet that has shown improved performance on pattern recognition of signals corresponding to neural action potentials of H1, a motion sensitive neuron in the fly's visual system. These prior efforts motivate our study of application-specific base wavelets for improved signature extraction in signals related to manufacturing.

11.2 Construction of an Impulse Wavelet

Considering that base wavelets available in the literature (e.g., provided by MATLAB) are developed primarily from a mathematical point of view without reference to a specific physical system – although in real-world applications, signals to be analyzed are generally produced by physical systems – it would be interesting, from an intellectual pursuit point of view, to study how to construct a customized base wavelet from the physical phenomena being analyzed. Naturally, such construction process will have to satisfy the mathematical requirement for designing a base wavelet. With this in mind, we introduce an impulse wavelet designed for analyzing vibration signals measured from rolling bearings, which are widely used in manufacturing machines.

Generally, a base wavelet must satisfy the conditions as described in Chaps. 3 and 4 to ensure that a signal's wavelet transformation does not result in loss of information so that the signal can be properly reconstructed from the corresponding wavelet coefficients. Mathematically, such a reconstruction exists if a scaling function $\phi(t)$, which satisfies the following *dilation* equation (Burrus et al. 1998; Cui et al. 1994), can be found as:

$$\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) = \sum_n h_n \phi(t-n) \tag{11.1}$$

In (11.1), h_n is a set of scaling coefficients applied to $\phi(t - n)$. Equation (11.1) indicates that the dilated version of $\phi(t)$ can be written as the sum of translated versions that are scaled by the coefficients h_n. Furthermore, it indicates that a scaling function at one scale can be constructed from a number of scaling functions at a previous scale. In general, the construction of a base wavelet starts from the scaling function that satisfies (11.1). Such scaling function is then used to derive the base wavelet.

Assuming an impulsive input is applied to a rolling bearing, a corresponding output signal can be defined by the convolution integral in the continuous form as (Inman 1996; Lutes and Sarkani 1997):

$$x(t) = \int_0^t R(\tau)h(t-\tau)d\tau \tag{11.2}$$

where $R(\tau)$ denotes the impulsive input, and $x(t)$ denotes the output signal.

In the discrete form, the impulsive input $R(\tau)$ is sampled at $R(n)$, and the output signal can be obtained as:

$$x(t) = \sum_n R(n)h(t-n) \tag{11.3}$$

In (11.2) and (11.3), the symbol $h(\bullet)$ represents the impulse response of the rolling bearing. Considering the discrete form of convolution expression, the similarity between (11.3) and (11.1) becomes apparent: the output $x(t)$ in (11.3) can be viewed as the sum of translated versions of the impulse response $h(\bullet)$ that are scaled by the input $R(n)$. If the impulse response satisfies (11.1), then it can be used to form a scaling function that contains relevant information on the underlying dynamics of the bearing being monitored. Subsequently, the scaling function can be used to construct a base wavelet for analyzing vibration signals measured from the bearing. Because of the nature of such a derivation, it is expected that the base wavelet presents a more direct and meaningful decomposition of the bearing signal than the standard wavelets commonly found in the literature.

To construct the base wavelet, several impulse responses of a ball bearing have been taken through hammer strikes, as shown in Fig. 11.1. The corresponding spectra are shown in Fig. 11.2. It is seen that the frequency components below 1,500 Hz are consistent in terms of their magnitudes, but those above 1,500 Hz have varied. The magnitudes of these high frequency components are smaller than those of the lower frequency components.

Because of their relatively small magnitude and random behavior, the high frequency components were treated as noise and removed from the signal with a low pass filter. The cutoff frequency for the filter was chosen to be 1,500 Hz, as the spectral components of each impulse were stable below this frequency. As seen in Fig. 11.3 where the original and filtered signals are shown, the filter is effective in removing noise from the impulse response and retaining the frequency constant of the original signal.

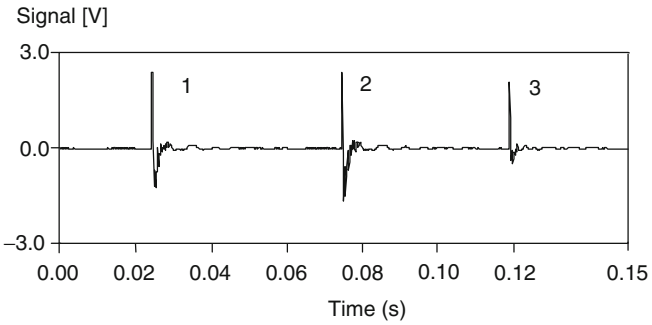

Fig. 11.1 Waveform of three consecutive impulse responses from a ball bearing

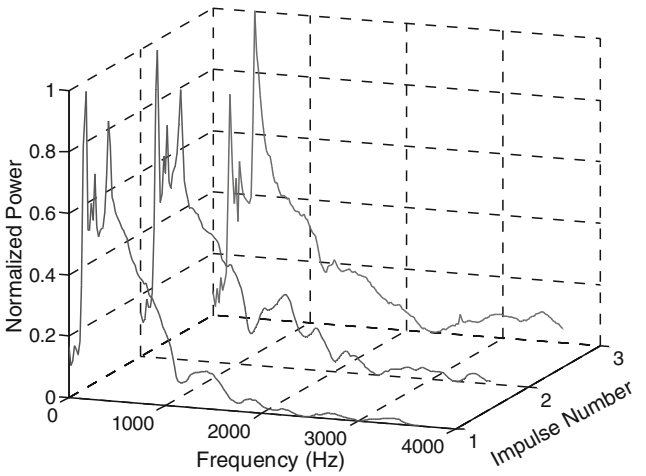

Fig. 11.2 Spectra of three consecutive impulse responses from a ball bearing

In order for the filtered impulse response shown in Fig. 11.3 to be used as a scaling function, it must satisfy the dilation equation, for which the length of the support interval of the signal must be at least one. A function with a support interval of less than one would have a gap between $h_n\phi(t–n)$ and $h_{n+1}\phi(t–n–1)$, within which nothing contributes to the sum on the right hand side of (11.1). Consequently, such a function would not satisfy the dilation equation. To address this issue, the impulse response is first dilated such that its support is greater than one. After dilation, the coefficients h_n are determined from a recursive relationship that is derived from the dilation equation.

As an example of the procedure of satisfying the dilation equation, a standard Daubechies scaling function $\phi(t)$ (Fig. 11.4) is illustrated below. It should be noted first that an explicit expression for the Daubechies scaling function does not exist (Daubechies 1992).

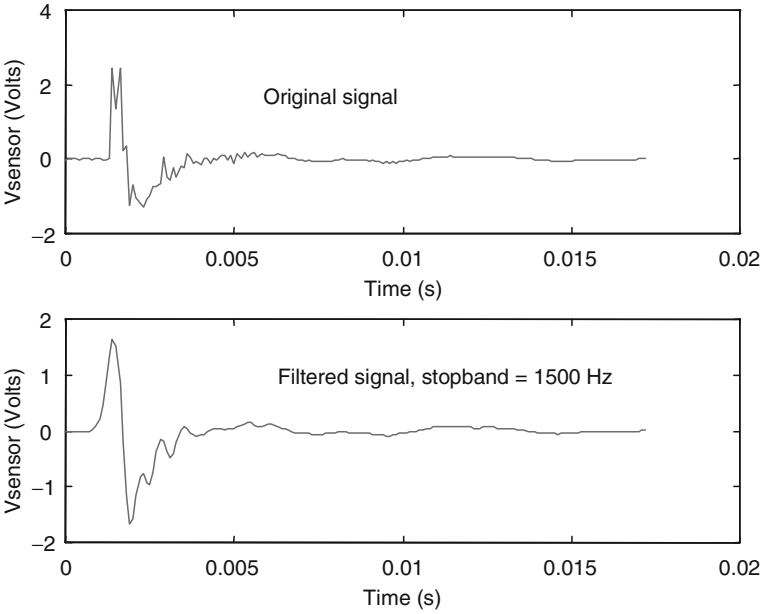

Fig. 11.3 Impulse response of a ball bearing

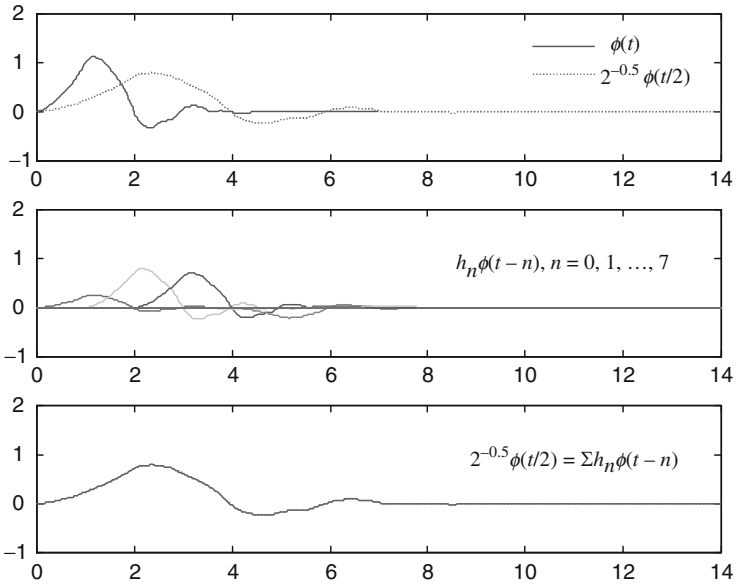

Fig. 11.4 The Daubechies scaling function

After the sequence of coefficients h_n is obtained, they are used to form an FIR filter denoted by H^*. The FIR filters H^* and G^* are quadrature mirror filters if, for a signal $x(t)$:

$$||H^*x(t)|| + ||G^*x(t)|| = ||x(t)|| \tag{11.4}$$

Together, H^* and G^* form a pair of reconstruction filters for the wavelet decomposition of a signal. This process, called *deconstruction*, is implemented via the adjoints of H^* and G^*, which are denoted by H and G, respectively (Kaiser 1994). For the Daubechies scaling function $\phi(t)$ shown in Fig. 11.4a, the filter coefficients are as follows: $h_n = \{0.2304, 0.7148, 0.6309, -0.0280, -0.1870, 0.0308, 0.0329, -0.0106\}$, $n = 0, 1, \ldots, 7$ (Misiti et al. 1997). In Fig. 11.4b, the scaled and translated version of $\phi(t)$ (i.e., $h_n\phi(t-n)$ for $n = 0, 1, 2, \ldots, 7$) is shown. Since $\phi(t)$ is a valid scaling function and h_n are valid filter coefficients, the dilation equation is satisfied, as shown in Fig. 11.4c.

The above procedure is repeated for the impulse response as shown below in Fig. 11.5. It should be noted that the impulse response $\phi(t)$ here is a function of the bearing dynamics, not an exact solution to the dilation equation. However, a set of filter coefficients h_n can be determined such that the impulse response approximately satisfies the dilation equation.

The filter coefficients can be calculated such that the dilation equation is satisfied exactly at integer values of t. The solution is recursive: each h_n, for $n > 0$, can be

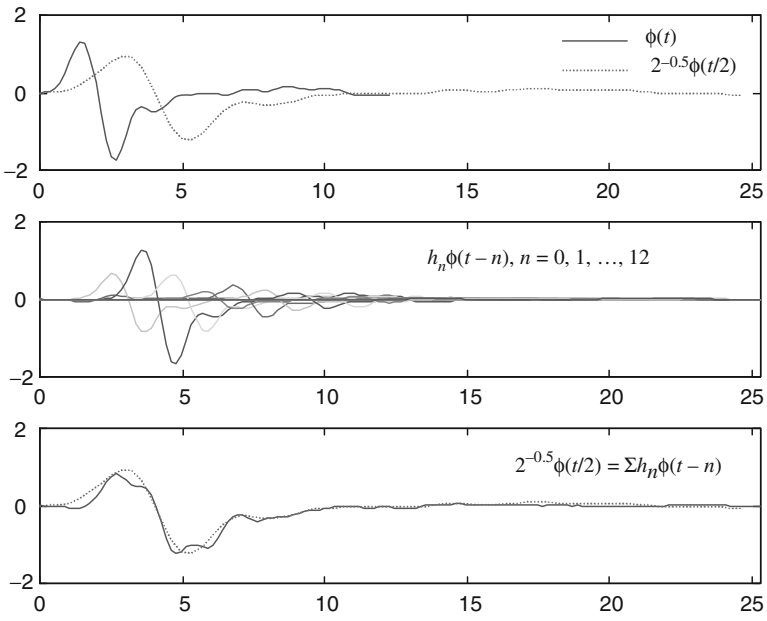

Fig. 11.5 The impulse scaling function obtained from the ball bearing

explicitly determined as a function of $\phi(t)$ and h_0, h_1, $h_2, \ldots, h_{n-1}$. The first coefficient, h_0, is simply a function of $\phi(t)$. These solutions are obtained by evaluating the dilation equation at integer values of t. For $t = 1$, (11.1) gives $\phi(1/2)/\sqrt{2} = h_0\phi(1)$. The terms $h_1\phi(0)$, $h_2\phi(-1)$, etc., do not appear because $\phi(t) = 0$ for $t \leq 0$. (Recall that ϕ is compactly supported.) Thus, h_0 is determined by:

$$h_0 = \frac{2^{-1/2}\phi(1/2)}{\phi(1)} \tag{11.5}$$

Similarly, for $t = 2$, (11.1) gives:

$$\frac{1}{\sqrt{2}}\phi(1) = h_0\phi(2) + h_1\phi(1) \tag{11.6}$$

For $t = 3$:

$$\frac{1}{\sqrt{2}}\phi\left(\frac{3}{2}\right) = h_0\phi(3) + h_1\phi(2) + h_2\phi(1) \tag{11.7}$$

For $t = N + 1$:

$$\frac{1}{\sqrt{2}}\phi\left(\frac{N+1}{2}\right) = h_0\phi(N+1) + h_1\phi(N) + h_2\phi(N-1) + \cdots + h_N\phi(1) \tag{11.8}$$

Equations (11.5)–(11.8) determine a recursive definition for each filter coefficient h_n. With the first coefficient h_0 given by (11.5), the remaining coefficients are given by:

$$h_n = \frac{2^{-1/2}\phi((n+1)/2) - \sum_{k=0}^{n-1} h_k\phi(n+1-k)}{\phi(1)}, \quad \text{for } n \geq 1 \tag{11.9}$$

Since the scaling function $\phi(t)$ is given by the impulse response of the bearing, each of the filter coefficients h_n can be readily determined from (11.5) and (11.9). Furthermore, note that the dilation equation can be written as:

$$\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) = h_0\phi(t) + \sum_{j=1}^{N} h_j\phi(t-j) \tag{11.10}$$

Since h_n is given by (11.9), the dilation equation can be rewritten as:

$$\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) = \frac{\sum_{j=1}^{N}\left[2^{-1/2}\phi((j+1)/2) - \sum_{k=0}^{j-1} h_k\phi(j+1-k)\right]\phi(t-j)}{\phi(1)} + h_0\phi(t) \tag{11.11}$$

Collecting terms yields the following form for the dilation equation:

$$\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) = \frac{\sum_{j=1}^{N}\left[2^{-1/2}\phi((j+1)/2)\phi(t-j)\right]}{\phi(1)} - \frac{\sum_{j=1}^{N}\sum_{k=0}^{j-1} h_k\phi(j+1-k)\phi(t-j)}{\phi(1)} + h_0\phi(t) \tag{11.12}$$

Note that only the second term on the right hand side of (11.12) contains filter coefficients h_n, which are determined by (11.5) and (11.9). Equation (11.12) serves to illustrate the interesting relationship that the dilation equation imposes between h_n and $\phi(t)$. Particularly, the expression given by (11.12) shows that the dilated version of the scaling function is related not only to $\phi(t-n)$ scaled by the filter coefficient h_n, but also to $\phi(t-n)$, scaled by $\phi(t)$, which is evaluated at integer values of t. The recursive relationship given by (11.9) gives h_n such that the dilation equation is satisfied at $x = \{0, 1, 2, \ldots\}$. At other points, the sum on the right hand side of (11.1) might differ from the left hand side. In practical treatment of an impulse scaling function such as shown in Fig. 11.5a, (11.5) and (11.9) are first used to obtain an initial set of filter coefficients. These coefficients are then optimized by minimizing the following error function:

$$E_{\text{rms}} = \sqrt{\frac{1}{T}\int_0^T \left(\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) - \sum_n h_n\phi(t-n)\right)^2 dt} \tag{11.13}$$

The error E_{rms} is a scalar valued function of the vector of filter coefficients h_n, and the optimization is accomplished by finding the vector which minimizes E_{rms}. Since E_{rms} is a measure of how well the dilation equation is satisfied, the vector h_n minimizing E_{rms} is the best set of filter coefficients that can be obtained from $\phi(t)$. Using this technique, the filter coefficients are determined to be: $h_n = \{-0.0529, 0.4897, 0.9601, 0.4848, 0.1467, 0.2653, 0.1723, 0.1295, 0.1208, 0.0495, -0.0182, -0.0255, 0.0131\}$, for $n = 0, 1, \ldots, 12$. The translated and scaled versions of $\phi(t)$ corresponding to these h_n (i.e., $h_n\,\phi(t-n)$) are plotted in Fig. 12.5b. As indicated by Fig. 12.5c, the impulse response is an approximate solution to the dilation equation ($E_{\text{rms}} = 0.0984$). The low pass filter coefficients derived from this scaling function $\phi(t)$ can then be used to determine the corresponding wavelet $\psi(t)$ (Young 1993; Mallat 1998). The coefficients for the high pass reconstruction filter G^* are determined from (11.4). The wavelet is evaluated by upsampling G^*, convolving it with H^*, and then iteratively repeating this procedure:

$$H^*_{n+1} = \Uparrow G^* * H^*_n \tag{11.14}$$

where $\Uparrow$ is a dyadic up-sampling operator. Thus, after N iterations, $\psi(t) \cong H^*_{N+1}$. Figure 11.6 shows the result of four iterations of (11.14), which produced a

customized wavelet, based on the impulse response of the rolling bearing structure. The set of FIR filters based on $\psi(t)$ and synthesis based on $\phi(t)$ are given in Table 11.1, where the filters have been normalized to have a norm of $1/\sqrt{2}$.

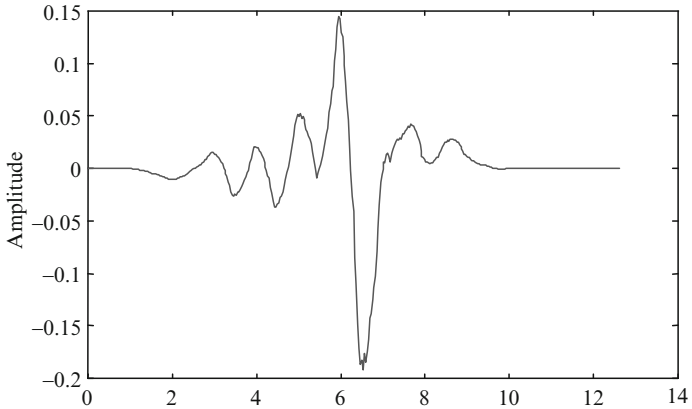

Fig. 11.6 The wavelet derived from the impulse response

Table 11.1 Normalized filter coefficients

Deconstruction		Reconstruction	
Low pass	High pass	Low pass	High pass
0	0.0274	−0.0274	0
0.0068	0.2532	0.2532	−0.0068
−0.0132	−0.4964	0.4964	−0.0132
−0.0094	0.2507	0.2507	0.0094
0.0256	−0.0759	0.0759	0.0256
0.0625	0.1372	0.1372	−0.0625
0.0670	−0.0891	0.0891	0.0670
0.0891	0.0670	0.0670	−0.0891
0.1372	−0.0625	0.0625	0.1372
0.0759	0.0256	0.0256	−0.0759
0.2507	0.0094	−0.0094	0.2507
0.4964	−0.0132	−0.0132	−0.4964
0.2532	−0.0068	0.0068	0.2532
−0.0274	0	0	0.0274

11.3 Impulse Wavelet Application

As an application example, the impulse wavelet is used to diagnose bearing defect. Figure 11.7a shows a vibration signal measured on a SKF 6220 ball bearing with a 0.25-mm hole on its inner raceway. This signal is sampled at 10 kHz, and the rotating speed of the bearing is 600 rpm (i.e., corresponding to 10 Hz rotating frequency). Based on geometric dimensions of the bearing and the rotating speed (Harris 1991), the defect characteristic frequency on the inner raceway of such bearing is $f_{\mathrm{BPFI1}} = 58.6$ Hz. As illustrated in Fig. 11.7b, such a defect-related frequency component is not seen in its power spectral density (PSD) resulted from the Fourier transform.

Utilizing the wavelet integrated with Fourier transform technique, which is described in Chap. 7, the same vibration signal is first analyzed by the wavelet transform. The impulse wavelet, developed from the impulse response of the rolling bearing as described above, is used as the base wavelet. Fourier transform is then performed on the wavelet coefficients obtained from the wavelet transform to expose explicitly the related frequency components. Figure 11.8 illustrates the resulting wavelet coefficients and their corresponding PSD. It is seen that the defect-related frequency component f_{BPFI1} at 58.6 Hz is clearly shown in the spectrum, thus verifying the existence of a localized inner raceway defect.

To demonstrate the signature extraction capability of the designed impulse wavelet for bearing defect diagnosis, a comparison study is carried out, where

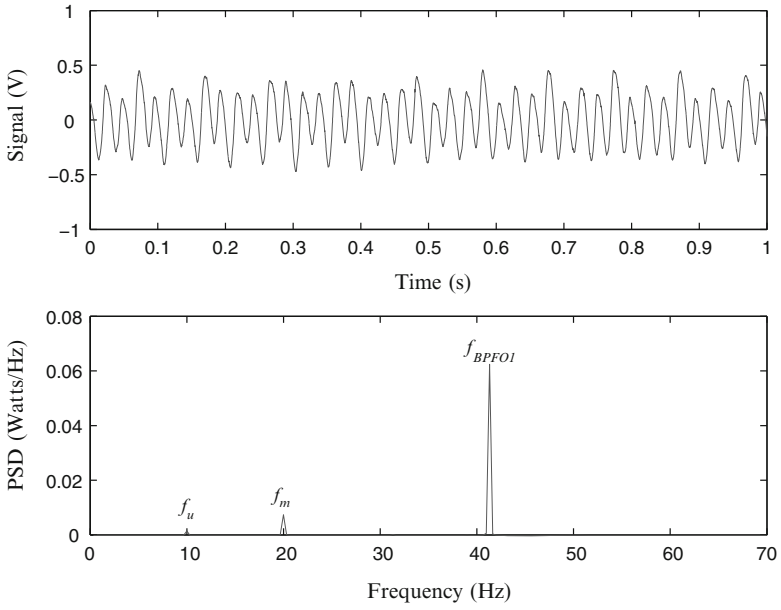

Fig. 11.7 Vibration signal and its PSD from a defective bearing

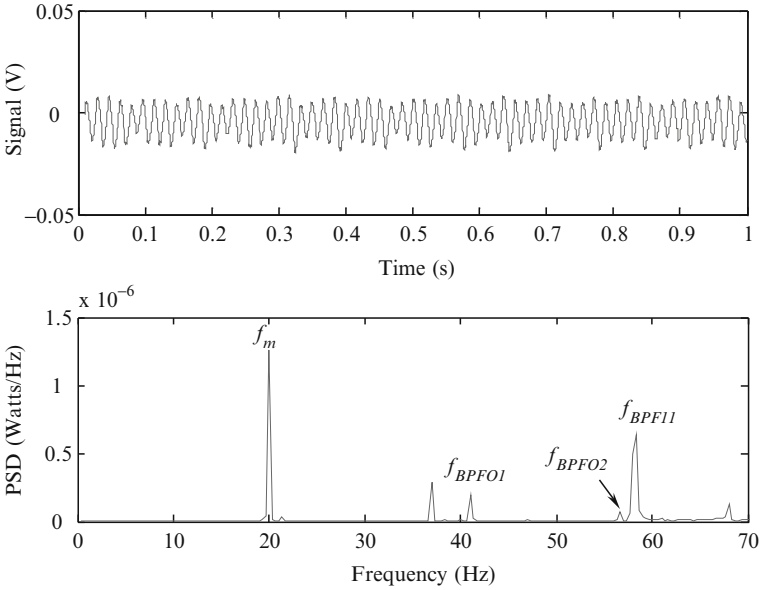

Fig. 11.8 Wavelet-integrated Fourier spectrum using the customized impulse wavelet

five standard base wavelets from the literature: Daubechies 2 and 4 wavelets, Coiflets 1, Symlets 3, and Biorthogonal 2.2 (Daubechies 1992; Lou and Loparo 2004; Zhang et al. 2005) are used to analyze the vibration signal. The upper parts of Figs. 11.9–11.13 are intermediate results (i.e., wavelet coefficients) of the integrated wavelet-Fourier transform analysis, and the lower parts of these figures are their corresponding PSDs. It is shown that all the five standard base wavelets can identify the defect-related frequency component, and the results are shown in the lower parts of Figs. 11.9–11.13.

In the spectra of Figs. 11.9–11.13, there exists a frequency component f_{BPFO2} at 56.5 Hz, which has a distinct magnitude. Such a frequency component is identified as from the ball rotation of another bearing in the support structure (Yan et al. 2009). To enable a quantitative performance comparison of the developed impulse wavelet and other five standard base wavelets, a signal-to-noise ratio measure is introduced, which is the amplitude ratio between the defect frequency f_{BPFI1} and the adjacent frequency f_{BPFO2}. As listed in Table 11.2, the impulse wavelet has shown the highest signal-to-noise ratio in detecting the defect-characteristic frequency of $f_{\mathrm{BPFI1}} = 58.6$ Hz. This result can be attributed to the nature of this customized wavelet, which is derived from the actual impulse response of the bearing structure. The direct link to the dynamics of the bearing and thus inherent better match to the bearing signature than a standard wavelet has made it more effective in exposing the constituent features for defect identification.

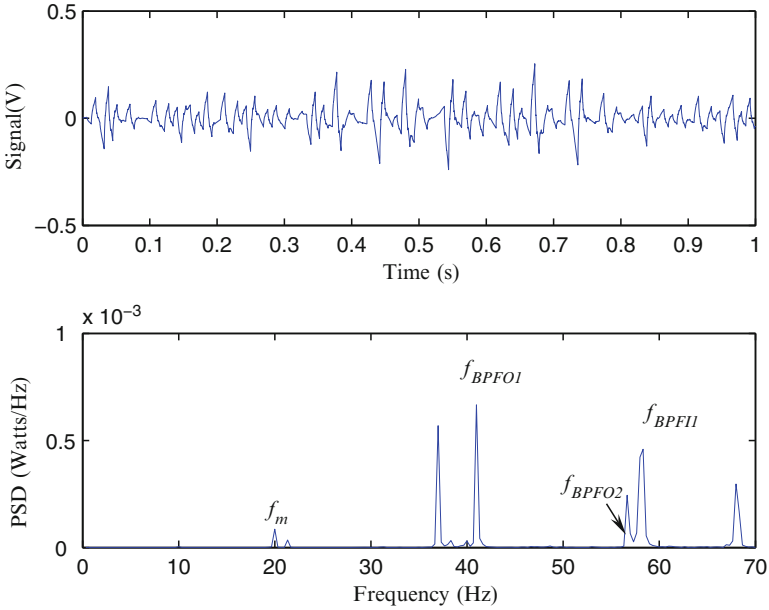

Fig. 11.9 Wavelet-integrated Fourier spectrum results using Daubechies 2 (Db2) wavelet

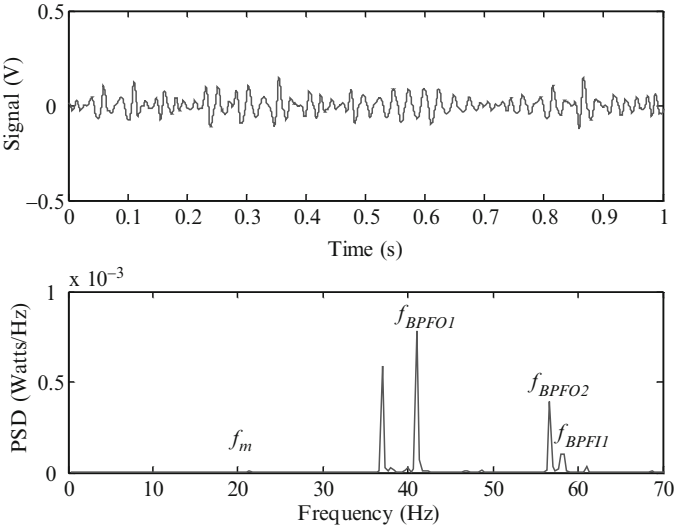

Fig. 11.10 Wavelet-integrated Fourier spectrum results using Daubechies 4 (Db4) wavelet

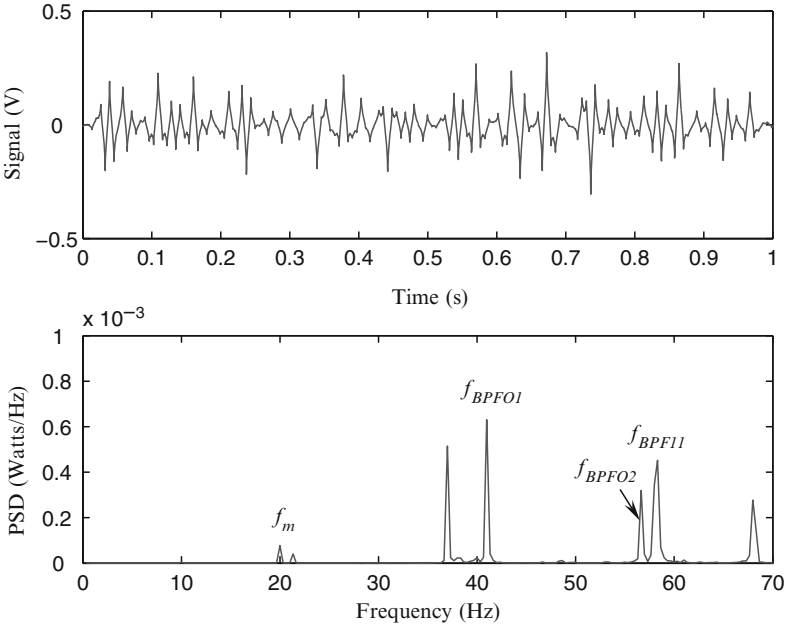

Fig. 11.11 Wavelet-integrated Fourier spectrum results using Coiflets 1 (Coif1) wavelet

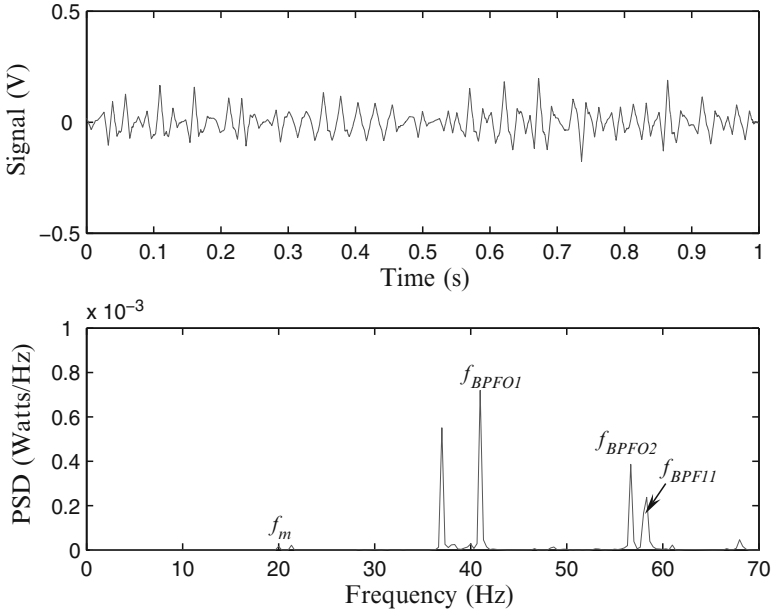

Fig. 11.12 Wavelet-integrated Fourier spectrum results using Symlets 3 (Sym3) wavelet

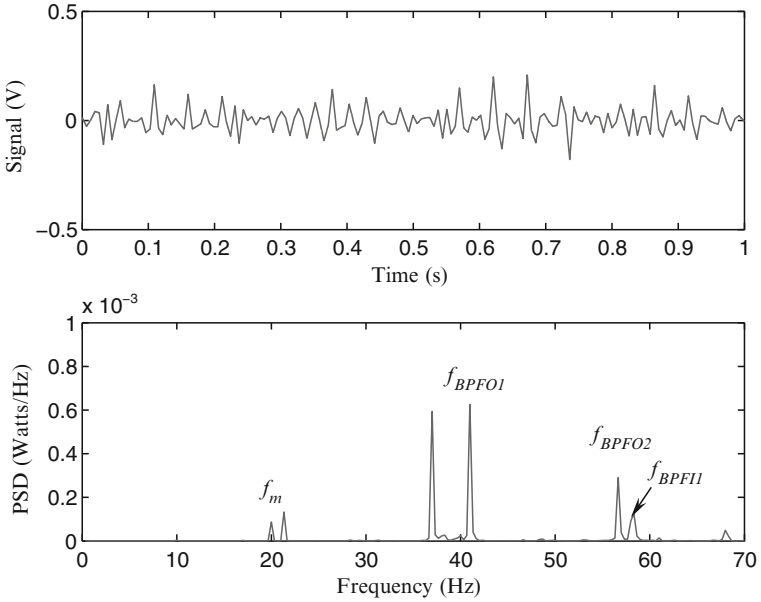

Fig. 11.13 Wavelet-integrated Fourier spectrum results using Biorthogonal 2.2 (Bior2.2) wavelet

Table 11.2 Comparison of signal-to-noise ratios for different base wavelets

Base wavelet	f_{BPFI1}/f_{BPFO2}
Impulse	9.55
Db2	1.88
Db4	0.27
Coif1	1.41
Sym3	0.62
Bior2.2	0.43

11.4 Summary

This chapter introduces the procedure to design a wavelet based on the dynamics of the physical system being analyzed. Using the impulse response of a rolling bearing system, an impulse wavelet has been constructed for defect-induced signature extraction. Experimental study has verified the effectiveness of the impulse wavelet in identifying bearing localized defect of the bearing in its inner raceway, as illustrated in the comparative study involving five standard wavelets from the literature. Although the impulse wavelet development is based on a specific type of bearing, the analytical procedure described in this chapter should be applicable to the analysis of other types of mechanical systems.

11.5 References

Aldroubi A, Unser M (1993) Families of multiresolution and wavelet spaces with optimal properties. Numer Funct Anal Optim 14:417–446

Burrus CS, Gopinath R, Guo H (1998) Introduction to wavelets and wavelet transforms: a primer. Prentice Hall, Englewood Cliffs, NJ

Chapa JO, Rao RM (2000) Algorithm for designing wavelets to match a specified signal. IEEE Trans Signal Process 48(12):3395–3406

Cohen A, Daubechies I, Feauveau JC (1992) Biorthogonal bases of compactly supported wavelets. Commun Pure Appl Math 45:485–560

Cui CK, Montefusco L, Puccio L (1994) Wavelet: theory, algorithms, and applications. Academic, New York

Daubechies I (1988) Orthonormal bases of compactly supported wavelets. Commun Pure Appl Math 41:909–996

Daubechies I (1992) Ten lectures on wavelets. Society of Industrial and Applied Mathematics, Pennsylvania, PA

Gopinath RA, Odegard JE, Burrus CS (1994) Optimal wavelet representation of signals and wavelet sampling theorem. IEEE Trans Circuits Syst II Analog Digital Signal Process 41:262–277

Guido RC, Slaets JFW, Koberle R, Almeida LOB, Pereira JC (2006) A new technique to construct a wavelet transform matching a specified signal with applications to digital, real time, spike, and overlap pattern recognition. Digit Signal Process 16:22–44

Gupta A, Joshi SD, Prasad S (2005a) A new approach for estimation of statistically matched wavelet. IEEE Trans Signal Process 53(5):1778–1793

Gupta A, Joshi SD, Prasad S (2005b) A new method of estimating wavelet with desired features from a given signal. Signal Processing 85:147–161

Harris T (1991) Rolling bearing analysis. Wiley, New York

Inman D (1996) Engineering vibration. Prentice Hall, Englewood Cliffs, NJ.

Kaiser G (1994) A Friendly Guide to Wavelets. Birkhauser, Boston, MA

Lou X, Loparo KA (2004) Bearing fault diagnosis based on wavelet transform and fuzzy inference. Mech Syst Signal Process 18:1077–1095.

Lu WS, Antoniou A (2001) Design of signal-adapted biorthogonal filter banks. IEEE Trans Circuits Syst I Fundam Theory Appl 48:90–102.

Lutes L, Sarkani S (1997) Stochastic analysis of structural and mechanical vibrations. Prentice Hall, Englewood Cliffs, NJ

Mallat S (1998) A wavelet tour of signal processing. Academic, Boston, MA

Misiti M, Misiti Y, Oppenheim G, Poggi J (1997) Wavelet toolbox for use with Matlab. The Math Works, Inc., Natick, MA

Shark L, Yu C (2003) Design of optimal shift-invariant orthonormal wavelet filter banks via genetic algorithm. Signal Processing 83:2579–2591

Shark L, Yu C (2006) Design of matched wavelets based on generalized Mexican-hat function. Signal Processing 86:1451–1469

Tewfik AH, Sinha D, Jorgensen P (1992) On the optimal choice of a wavelet for signal representation. IEEE Trans Inf Theory 38:747–765

Yan R, Gao R, Wang C (2009) Experimental evaluation of a unified time-scale-frequency technique for bearing defect feature extraction. ASME J Vib Acoust 131:041012-1-12

Young R (1993) Wavelet theory and its applications. Kluwer Academic Publishers, Boston, MA

Zhang S, Mathew J, Ma L, Sun Y (2005) Best basis-based intelligent machine fault diagnosis. Mech Syst Signal Process 19:357–370