

Chapter 7

Wavelet Integrated with Fourier Transform: A Unified Technique

Fourier transform-based spectral analysis has been widely applied to processing signals, such as vibration and acoustic signals (Mori et al. 1996; Tandon and Choudhury 1999; Cavacece and Introini 2002), acquired from manufacturing systems. Because of noise contamination and signal interference, the constituent components of interest may be submerged in the signal and difficult to be revealed through a spectral analysis (Ho and Randall 2000). Furthermore, events occurred in the manufacturing system may be transient in nature, for example, the initiation and propagation of surface spalling in a ball bearing (Gao and Yan 2006; Orhan et al. 2006). As another example, the process of metal removal can be viewed as consisting of multiple, individual transient events in which a single chip of metal is removed (Ge et al. 2004; Obikawa and Shinozuka 2004; Byrne and O'Donnell 2007; Malekian et al. 2009). On the one hand, given the global analysis nature, it is difficult to apply the Fourier transform for localizing these transient events. On the other hand, the Fourier transform can identify a signal's frequency components, from which a specific event (e.g., localized bearing defects, which have distinct characteristic frequencies at inner raceway, outer raceway, or rolling element itself) can be detected. Leveraging the capability of wavelet transform in transient signal analysis, this chapter introduces a *unified* time–scale–frequency analysis technique through spectral postprocessing on the data set extracted by wavelet transforms to enhance the effectiveness of signal representation and identification.

7.1 Generalized Signal Transformation Frame

Fourier transform and wavelet transform were originated from different theoretical platforms, and each technique analyzes a signal from a different perspective. Specifically, the Fourier transform depicts the energy concentration of constituent frequency components within the signal, whereas the wavelet transform presents the similarity between the signal being analyzed and the base wavelet, in the time–scale domain. To enable cross-domain unification of the two techniques for signal analysis, a common signal transformation platform needs to be established first, which is the focus of this chapter.

Let us first define a function $W_{1,0}(t)$ within a certain time interval, or *support*, expressed as $[0, L)$, where the symbol L represents the width of support. The function $W_{1,0}(t)$ is called the *base template* function for signal analysis. Next, we define $W_{s,u}(t)$, which is a derived version of $W_{1,0}(t)$. Comparing to $W_{1,0}(t)$, the magnitude of $W_{s,u}(t)$ is *scaled* by *scale* s , where $s \geq 0$ is an integer number, and its location along the time axis is *shifted* by *time* u , with $u \in R$, and R represents the set of all real numbers. The function $W_{s,u}(t)$ is called the *derived template* function, at scale s and time u , and is supported within the time interval $[u, u + sL]$. Generally, $W_{s,u}(t)$ can be expressed in terms of the *base template* function $W_{1,0}(t)$ as:

$$W_{s,u}(t) = \frac{1}{\sqrt{s}} W_{1,0}\left(\frac{t-u}{s}\right) \quad (7.1)$$

where $1/\sqrt{s}$ is a factor for purpose of normalization. Specifically, it ensures that the following relationship between the *derived template* function $W_{s,u}(t)$ and the *base template* function $W_{1,0}(t)$ is always satisfied:

$$\int_{-\infty}^{\infty} W_{s,u}^2(t) dt = \int_{-\infty}^{\infty} W_{1,0}^2(t) dt \equiv \|W_{1,0}(t)\|^2 \quad (7.2)$$

The physical significance of (7.2) is that all the derived template functions preserve the same amount of energy as the base template function does.

In a linear signal space, the set of all *derived template* functions $\{W_{s,u}(t): s \geq 0, u \in R\}$ forms a *continuous frame* Γ_c , spanned by scale s and time u . In accordance with the discrete data sampling process, where data points are taken at time instances $u = mkL$ from a derived template function with a scale factor of $s = k$, a discrete expression of the derived template function is obtained as:

$$W_{k,m} = \frac{1}{\sqrt{k}} W_{1,0}\left(\frac{t - mkL}{k}\right) \quad (7.3)$$

where $W_{k,m}(t)$ is a simplified expression for $W_{k,mkL}(t)$. In the above notation, k or $k^{-1} \in N$, $m \in Z$, and k and m represent a discrete version of the continuous parameters s and u . The notation $k^{-1} \in N$ corresponds to $s < 1$. The set $\{W_{k,m}(t): k \text{ or } k^{-1} \in N, m \in Z\}$, with N being the set of all nonnegative integers and Z being the set of all integers, forms a *discrete frame* Γ_d , spanned by the parameters k and m . The *continuous frame* Γ_c (or *discrete frame* Γ_d) provides a *generalized frame* for signal transformation, and is defined as *complete* in the linear signal space, if any signal function $x(t)$ can be expressed in it as (Kaiser 1994):

$$x(t) = \int_0^\infty \int_{-\infty}^\infty C(s,u) W_{s,u}(t) ds du \quad (7.4)$$

or, in the case of *discrete frame* Γ_d :

$$x(t) = \sum_{k=1}^{\infty} \sum_{m=-\infty}^{\infty} C(k, m) W_{k,m}(t) \quad (7.5)$$

In (7.4) and (7.5), the coefficient of the functions, $C(s, u)$ or $C(k, m)$, can be considered as a *measure* function, which expresses the extent to which the signal function $x(t)$ is correlated to the *derived template* function $\{W_{s,u}(t): s \geq 0, u \in R\}$ of scale s , at the specific time u .

Under the discrete frame Γ_d , the significance of the *measure* function expressed in (7.5) can be further illustrated, considering that there exists a complete *orthogonal* set $\{W_{k,m}(t)\}$ within such a frame. The *orthogonal* identity states that:

$$\int_{-\infty}^{\infty} W_{k_1,m_1}(t) W_{k_2,m_2}(t) dt = \begin{cases} \int_{-\infty}^{\infty} W_{k_1,m_1}^2(t) dt; & \text{for } k_2 = k_1, m_2 = m_1 \\ 0; & \text{otherwise} \end{cases} \quad (7.6)$$

where k_1 and $k_2 \in \{k\}$, and m_1 and $m_2 \in \{m\}$. Using the identity of (7.6), multiplying the two sides of (7.5) by $W_{k_1,m_1}(t)$, and integrating over the time interval $(-\infty, \infty)$, we have:

$$\begin{aligned} \int_{-\infty}^{\infty} x(t) W_{k_1,m_1}(t) dt &= \int_{-\infty}^{\infty} \sum_k \sum_m C(k, m) W_{k,m}(t) W_{k_1,m_1}(t) dt \\ &= C(k_1, m_1) \int_{-\infty}^{\infty} W_{k_1,m_1}^2(t) dt \end{aligned} \quad (7.7)$$

With $k = k_1$ and $m = m_1$, rearranging (7.7) yields,

$$C(k, m) = \frac{\int_{-\infty}^{\infty} x(t) W_{k,m}(t) dt}{\int_{-\infty}^{\infty} W_{k,m}^2(t) dt} = \frac{\int_{-\infty}^{\infty} x(t) W_{1,0}\left(\frac{t-mkL}{k}\right) dt}{\int_{-\infty}^{\infty} W_{1,0}^2(t) dt} \quad (7.8)$$

The equation indicates that the *measure* function $C(k, m)$ expresses the *correlation* (or *similarity*) between the signal function $x(t)$ and the *derived template* function $W_{k,m}(t)$ of scale k and at time mkL . This concept can be expanded to view a signal transformation operation, such as Fourier or wavelet transform, as a *correlation operation* between a signal and a *template* function, and the result expresses the measures of correlation between the two functions. When a *template* function derived from a *base template* function has a large value of $C(k, m)$, or correlation, with a signal feature at certain scale k and time mkL , the *template* function is said to have a good *match* with that corresponding feature. As a result, the feature will be effectively extracted by this particular *template* function. Signal components that are of little correlation to the *template* function will show small or no correlation measures, and thus be suppressed in the analysis. A signal may show different correlation measures with different *base template* functions. In Table 7.1, several

Table 7.1 Basic template functions expressed in the generalized transformation frame

	Frame $\{W_{k,m}(t)\}$	Properties
Fourier function	$W_{1,0}(t) = e^{-j2\pi t/L}, \quad t \in [0, L)$	Exponential function forms a complete orthogonal base
Haar function	$W_{1,0}(t) = \begin{cases} +1 & 0 \leq t < L/2 \\ -1 & L/2 \leq t < L \end{cases}, \quad t \in [0, L)$	Rectangular waveform forms a complete orthogonal base
Daubechies function	$W_{1,0}(t) = \psi_{1,0}^{(n)}(t), \quad t \in [0, L)$ ^a	Fractal shape forms a complete orthogonal base

^aThere are different Daubechies functions $\psi_{1,0}^{(n)}(t)$ depending on different order of n

basic template functions are expressed in the generalized signal transformation frame, and they are discussed in the following sections.

7.1.1 Fourier Transform in the Generalized Frame

A signal $x(t)$ of period T can be expressed through its Fourier transform as (Bracewell 1999):

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{-j2\pi nt/T}, \quad -\infty < t < \infty, \quad n \in N \quad (7.9)$$

where c_n is the n th-order transformation coefficient. If a single-period complex exponential function is defined as the *base template* function:

$$W_{1,0}(t) = e^{-j2\pi t/L}, \quad t \in [0, L) \quad (7.10)$$

then the corresponding *derived template* function can be expressed, per definition in (7.1), as:

$$W_{k,m}(t) = \frac{1}{\sqrt{k}} e^{-j2\pi(\frac{t}{kL} - m)}, \quad t \in [mkL, mkL + kL) \quad (7.11)$$

Using the *derived template* function, the signal $x(t)$ shown in (7.9) can be expressed, in the generalized transformation frame, as:

$$\begin{aligned} x(t) &= \sum_k \sum_m C(k, m) W_{k,m}(t) = \sum_k \sum_m \frac{1}{\sqrt{k}} C(k, m) e^{-j2\pi(\frac{t}{kL} - m)} \\ &= \sum_k \left(\sum_m \frac{1}{\sqrt{k}} C(k, m) \right) e^{-j2\pi \frac{t}{kL}} = \sum_k \bar{C}_k V_k(t) \end{aligned} \quad (7.12)$$

where the term

$$\bar{C}_k = \sum_m \frac{1}{\sqrt{k}} C(k, m) \quad (7.13)$$

represents the sum of the normalized individual measure functions corresponding to the discrete scale k in the Fourier transform, and the term

$$V_k(t) = e^{-j2\pi \frac{t}{kL}} \quad (7.14)$$

represents a periodic exponential function with a period of kL over the time interval $(-\infty, \infty)$, for a given scale k . Comparing (7.9) with (7.12), it can be seen that $k = 1/n$ when $L = T$, and $c_n = \bar{C}_k$. The expression $k = 1/n$ indicates that a higher scale (k) corresponds to a lower frequency (n), and $c_n = \bar{C}_k$ is defined only by the frequency n or scale k . From this discussion, the Fourier transform can be viewed in the generalized frame as a 1D function, defined by the scale k with the *orthogonal* base $\{V_k(t)\}$. There is no time information of the extracted signal features contained in this function. This explains why the Fourier transform does not provide time information of the extracted frequency components.

7.1.2 Wavelet Transform in the Generalized Frame

Wavelet transform decomposes a signal using finite time intervals (or *support*) at different scales, thus maintains the time location information of the signal features. Through variation of the scales of the template function used, nonstationary or *transient* features within a signal can be extracted more effectively than by using the exponential (sine and cosine) functions in the Fourier transform. The wavelet transform can be expressed as:

$$C(s, u) = \int_{-\infty}^{\infty} x(t) W_{s,u}(t) dt \quad (7.15)$$

where the term $C(s, u)$ represents the wavelet coefficients (or *measure* function, in the generalized signal transformation frame). The wavelet function $W_{s,u}(t)$ is defined by (7.1). To reduce computational load and avoid redundancy in signal representation, *discrete* instead of *continuous* wavelet transform is often employed for analyzing signals consisting of discrete data points acquired through a data acquisition process. Generally, a discrete wavelet transform discretizes a signal by using a scale of the power of 2 (Daubechies 1992; Kaiser 1994). At the scale $k = 2^n$ ($n \in \mathbb{N}$), the discrete wavelet function $W_{k,m}(t)$ ($m \in \mathbb{Z}$) has the support of $T_w = 2^n L$. Physically, the term $2^n L$ represents the time resolution, which increases linearly with the logarithm of the scale, enabling signal feature extractions at different resolutions. The frequency

resolution of wavelet transform is the inverse of the time resolution, or $1/T_w$. It increases as the scales become higher (i.e., when n increases), and is therefore well suited for analyzing slow-changing signals. At lower scales, the frequency resolution decreases, enabling analysis of fast-changing signals. Such is in contrast to the Fourier transform, which maintains a constant frequency resolution over the entire spectrum, and have a time resolution that is defined by the signal duration.

To illustrate the ability of wavelet transform in signal feature extraction, we analyze an frequency shifted keying (FSK) signal, which is commonly used for data modulation and wireless data transmission (Gibson 1999). An FSK signal is expressed as:

$$x(t) = \begin{cases} \text{square}(2\pi f_1 t) & \text{for message "1"} \\ \text{square}(2\pi f_2 t) & \text{for message "0"} \end{cases} \quad (7.16)$$

where $\text{square}(2\pi f_n t)$ represents a periodic square wave with a unit amplitude and frequency f_n . Figure 7.1 illustrates an example of such an FSK signal, where the frequency $f_1 = 30$ Hz is used to transmit the digit "1" and $f_2 = 125$ Hz is used to transmit the digit "0." The message to be transmitted is [1 0 0 1 1 0 1 0 0 0]. Such a signal $x(t)$ is nonsinusoidal and nonstationary.

To analyze this signal, we choose the Haar wavelet as the base wavelet, since its square-shaped wave form best matches the shape of the FSK signal. Given that the wavelet has a support of $L = 1$ s (refer to Table 7.1), the measure function $C(s_1, m)$ can be calculated by using (7.8). For the message "1," $C(s_1, m)$ is calculated to be $\sqrt{s_1}$, at scale $s_1 = 1/f_1$ and time $t = ms_1$. The measure $C(s_1, m)$ is zero for the message "0." Similarly, at scale $s_2 = 1/f_2$, $C(s_2, m)$ is $\sqrt{s_2}$ at time $t = ms_2$ for the message "0," and zero for a message "1." The result of such a wavelet transform operation indicates that the Haar wavelet was able to locate the time instant of the message "1" or "0" at scales s_1 and s_2 , respectively, and expresses the FSK signal $x(t)$ in the generalized signal transform frame as a single-form expression:

$$x(t) = \sum_{m=-\infty}^{\infty} C(s_1, m) \psi_{s_1, m}^{(1)}(t) + C(s_2, m) \psi_{s_2, m}^{(1)}(t) \quad (7.17)$$

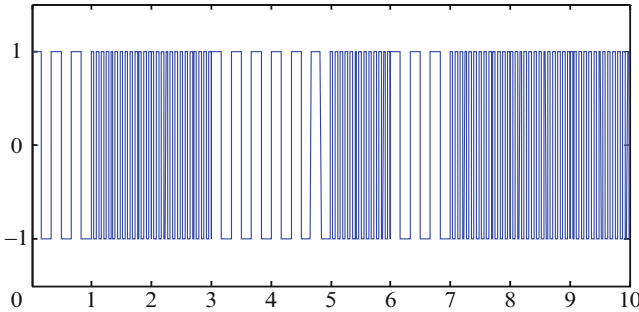


Fig. 7.1 A FSK signal $x(t)$

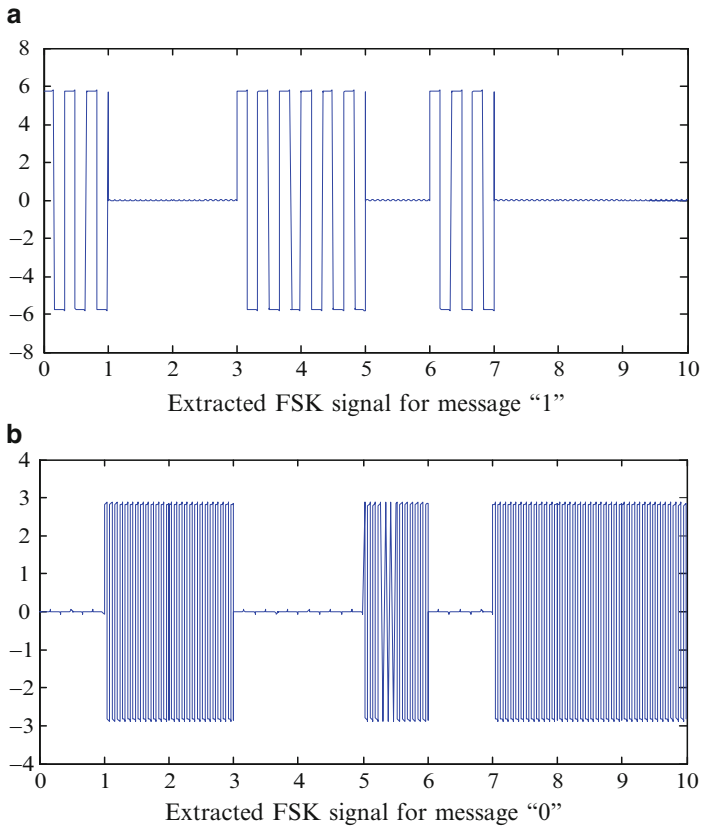


Fig. 7.2 Extracted FSK signals. (a) Extracted FSK signal for message “1” and (b) extracted FSK signal for message “0”

The results are shown in Fig. 7.2, where the messages “1” and “0” can be clearly separated into two different scales. In comparison, it is not feasible to use the Fourier transform to specify at which time which message (1 or 0) is transmitted. The sine and cosine template functions in the Fourier transform do not match the square waveform of the FSK signal, and consequently, the frequency component of the FSK signal will be spread out in a broad spectrum, especially when the message “0” and “1” are randomly transmitted.

7.2 Wavelet Transform with Spectral Postprocessing

As a time–scale domain technique, the wavelet transform utilizes template functions of different time resolutions at different scales to extract “transient” features embedded in a signal. Such transient features can be generated by the interactions between the rolling elements in bearing and a localized defect (e.g., surface spalling) on the surface of the raceway. As the rolling elements periodically roll over

the localized defect, the “transient” feature will reoccur at a fundamental frequency f_0 , which is a function of the bearing rotational speed. Such a relationship will be reflected in a wavelet transform of the bearing vibration signal, in that the measure function $C(s_1, m)$ will retain the same fundamental frequency along the time axis, at one of its scales (s_1). As a result, the spectral feature of the transient signal is retained in the wavelet transform, although it is not explicitly expressed. Because of masking by noise and other signals with similar spectral characteristics that appear at the same scale, it can be difficult to rely on wavelet transform alone to identify such hidden patterns.

Such constraint of the wavelet transform can be compensated for by subsequently applying the Fourier transform to the measure function $C(s_1, m)$ resulting from the wavelet transform. Such a postspectral technique reveals the specific frequency location of the transient features, and presents a *unified* approach to transient signal processing. The following section explains how such a postspectral method is realized.

7.2.1 Fourier Transform of the Measure Function

In a complete linear signal space, the wavelet-extracted data set at *scale* s can be expressed as:

$$x_s(t) = \int_{u=-\infty}^{u=\infty} C_s(u)W_s(t-u)du = C_s(t) \otimes W_s(t) \quad (7.18)$$

where the symbol \otimes represents the *convolution* operation between the measured function $C_s(u)$ and the wavelet function $W_s(u)$. To perform Fourier transform on the data set, the Fourier transform of the measure function $C_s(u)$ is first derived. For this purpose, the wavelet transform defined in (7.15) at a fixed scale s is rewritten as:

$$C_s(u) = \int_{-\infty}^{\infty} x(t)W_s(t-u)dt \quad (7.19)$$

In (7.19), the terms $C_s(u)$ and $W_s(t-u)$ represent their respective counterparts in (7.15), $C(s, u)$ and $W_{s,u}(t)$, with a fixed scale s . Through a normalization operation, $\|W_{1,0}(t)\|^2$ in (7.2) is set as 1 for simplicity. With respect to time u , the Fourier transform of $C_s(u)$, denoted as $\tilde{C}_s(f)$, is derived as:

$$\tilde{C}_s(f) = \tilde{x}(f)\tilde{W}_{s,u}(f) \quad (7.20)$$

where the symbol $\tilde{x}(f)$ expresses the Fourier transform of the signal $x(t)$. The symbol $\tilde{W}_{s,u}(f)$ expresses the Fourier transform of the wavelet function $W_{s,u}(t)$, which is derived as:

$$\begin{aligned}
\tilde{W}_{s,u}(f) &= \int_{-\infty}^{\infty} W_{s,u}(t) e^{-j2\pi ft} dt = \int_{-\infty}^{\infty} \frac{1}{\sqrt{s}} W_{1,0}\left(\frac{t-u}{s}\right) e^{-j2\pi ft} dt \\
&= \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} W_{1,0}\left(\frac{t-u}{s}\right) e^{-j2\pi ft} \left[s \cdot d\left(\frac{t-u}{s}\right)\right] \\
&= \sqrt{s} \int_{-\infty}^{\infty} W_{1,0}\left(\frac{t-u}{s}\right) e^{-j2\pi fs\left(\frac{t-u}{s} + \frac{u}{s}\right)} d\left(\frac{t-u}{s}\right) \\
&= \sqrt{s} e^{-j2\pi fu} \int_{-\infty}^{\infty} W_{1,0}\left(\frac{t-u}{s}\right) e^{-j2\pi fs\left(\frac{t-u}{s}\right)} d\left(\frac{t-u}{s}\right) \\
&= \sqrt{s} e^{-j2\pi fu} \tilde{W}_{1,0}(sf)
\end{aligned} \tag{7.21}$$

Combining (7.21) with (7.20) yields:

$$\tilde{C}_s(f) = \tilde{x}(f) \sqrt{s} e^{-j2\pi fu} \tilde{W}_{1,0}(sf) \tag{7.22}$$

Let $\tilde{W}_{1,0}^*(sf) = e^{-j2\pi fu} \tilde{W}_{1,0}(sf)$, (7.22) can be further expressed as:

$$\tilde{C}_s(f) = \sqrt{s} \tilde{x}(f) \tilde{W}_{1,0}^*(sf) \tag{7.23}$$

where the superscript * denotes the conjugate operator.

Similar to (7.19), in the case of *discrete* wavelet transform, the discrete measure function $C_k(m)$ at a fixed scale k can be expressed as:

$$C_k(m) = \int_{-\infty}^{\infty} x(t) W_k(t - mkL) dt \tag{7.24}$$

The corresponding Fourier transform of $C_k(m)$ is expressed as:

$$\tilde{C}_k(f) = \tilde{x}(f) \tilde{W}_{k,m}(f) \tag{7.25}$$

In (7.25), $\tilde{W}_{k,m}(f)$ is derived as follows:

$$\begin{aligned}
\tilde{W}_{k,m}(f) &= \int_{-\infty}^{\infty} W_{k,m}(t) e^{-j2\pi ft} dt = \int_{-\infty}^{\infty} \frac{1}{\sqrt{k}} W_{1,0}\left(\frac{t - mkL}{k}\right) e^{-j2\pi ft} dt \\
&= \frac{1}{\sqrt{k}} \int_{-\infty}^{\infty} W_{1,0}\left(\frac{t - mkL}{k}\right) e^{-j2\pi ft} \left[k \cdot d\left(\frac{t - mkL}{k}\right)\right] \\
&= \sqrt{k} \int_{-\infty}^{\infty} W_{1,0}\left(\frac{t}{k} - mL\right) e^{-j2\pi f\left(\frac{t}{k} - mL + mL\right)} d\left(\frac{t}{k} - mL\right) \\
&= \sqrt{k} e^{-j2\pi fmkL} \int_{-\infty}^{\infty} W_{1,0}\left(\frac{t}{k} - mL\right) e^{-j2\pi fk\left(\frac{t}{k} - mL\right)} d\left(\frac{t}{k} - mL\right) \\
&= \sqrt{k} e^{-j2\pi fmkL} \int_{-\infty}^{\infty} W_{1,0}\left(\frac{t}{k} - mL\right) e^{-j2\pi fk\left(\frac{t}{k} - mL\right)} d\left(\frac{t}{k} - mL\right) \\
&= \sqrt{k} e^{-j2\pi fmkL} \tilde{W}_{1,0}(kf)
\end{aligned} \tag{7.26}$$

As a result, (7.25) is given by:

$$\tilde{C}_k(f) = \tilde{x}(f)\tilde{W}_{k,m}(f) = \tilde{x}(f)\sqrt{k}e^{-j2\pi fmkL}\tilde{W}_{1,0}(kf) \quad (7.27)$$

Let $\tilde{W}_{1,0}^*(kf) = \tilde{W}_{1,0}(kf)e^{-j2\pi fmkL}$, (7.27) can be further expressed as:

$$\tilde{C}_k(f) = \sqrt{k}\tilde{x}(f)\tilde{W}_{1,0}^*(kf) \quad (7.28)$$

Equations (7.23) and (7.28) illustrate that the Fourier transform of the measure function at scale s (for continuous transform) or k (for discrete transform) can be viewed as the original signal $x(t)$ passing through a data filter, which is a contracted (by a frequency factor of s or k) and amplified (by a factor of \sqrt{s} or \sqrt{k}) version of the filter represented by the base wavelet function. Such an operation establishes the link between measured function and data filtering, and is of significance in wavelet transform-based signal analysis.

7.2.2 Fourier Transform of Wavelet-Extracted Data Set

With the Fourier transform of the measure function obtained, the Fourier transform of the extracted (or reconstructed) data set from the continuous wavelet transform, denoted $x_s(t)$ as shown in (7.18), can be derived as:

$$\begin{aligned} \tilde{x}_s(f) &= \tilde{C}_s(f)\tilde{W}_s(f) \\ &= \sqrt{s}\tilde{x}(f)\tilde{W}_{1,0}^*(sf)\sqrt{s}\tilde{W}_{1,0}(sf) \\ &= s\tilde{x}(f)|\tilde{W}_{1,0}(sf)|^2 \end{aligned} \quad (7.29)$$

In case of a *discrete* wavelet transform, the Fourier transform of the data set $x_k(t)$ is obtained by setting $s = k$ and $u = mkL$ in (7.18), and its Fourier transform is expressed as:

$$\begin{aligned} \tilde{x}_k(f) &= \tilde{C}_k(f)\tilde{W}_k(f) \\ &= \sqrt{k}\tilde{x}(f)\tilde{W}_{1,0}^*(kf)\sqrt{k}\tilde{W}_{1,0}(kf) \\ &= k\tilde{x}(f)|\tilde{W}_{1,0}(kf)|^2 \end{aligned} \quad (7.30)$$

This indicates that the Fourier transform of the extracted data set $x_k(t)$ at scale k can be viewed as the Fourier transform of the original signal $x(t)$ passing through a low-pass filter and the filter being represented by the transfer function $|\tilde{W}_{1,0}(kf)|^2$. If the template function at scale k correlates well with the “transient” feature of the signal $x(t)$ in the time domain, then its Fourier transform will contain a strong “disturbance” component in its spectrum. As a result, the filter $|\tilde{W}_{1,0}(kf)|^2$ will extract the

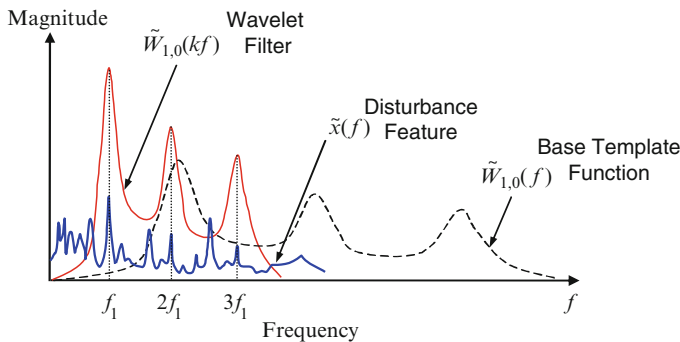


Fig. 7.3 Illustration of the filtering effect of wavelet transform

“disturbance” signal features from the original signal $x(t)$ at the scale k , as shown in (7.29) and (7.30). Because of a lower degree of correlation between this filter and other constituent components in the signal, other components will be attenuated at the scale k .

The filtering effect of postspectral processing of a wavelet transformed data series is illustrated in Fig. 7.3. The “transient” feature, represented by the solid thick line, is shown to have a fundamental frequency f_1 , characterized by a magnitude peak at f_1 and its harmonics ($2f_1$ and $3f_1$) in the frequency spectrum. When an appropriate base wavelet is selected (or designed) to decompose this signal (Holm-Hansen et al. 2004), a base template function will exist at a certain scale k where a high degree of correlation between the template function and the “transient” features can be identified. If the data set resulting from such a wavelet transform is subsequently processed by the Fourier transform, the result will be a data spectrum similar to that of the “transient” signal, with its major frequency components at f_1 , $2f_1$, and $3f_1$, respectively. Such a postspectral analysis can be viewed as filtering the data set in the frequency domain denoted by $\tilde{W}_{1,0}(kf)$.

7.3 Application to Bearing Defect Diagnosis

This section illustrates how the unified time–scale–frequency analysis technique described earlier can be applied to rolling bearing defect diagnosis. A custom-designed bearing test bed, as shown in Fig. 7.4, is set up to provide an experimental platform for evaluating the developed method. Axial and radial loads on the test bearing are applied through a hydraulic system, and the bearing rotation speed is varied by controlling the DC drive motor. Commercially available accelerometers (model 8624) are placed on the housing for vibration measurement with a data sampling frequency of 10 kHz. A deep-groove ball bearing of type 6220 with a seeded structural defect serves as the test bearing. The defect is implemented as a 0.25-mm diameter hole drilled on the inner raceway of the bearing, simulating the

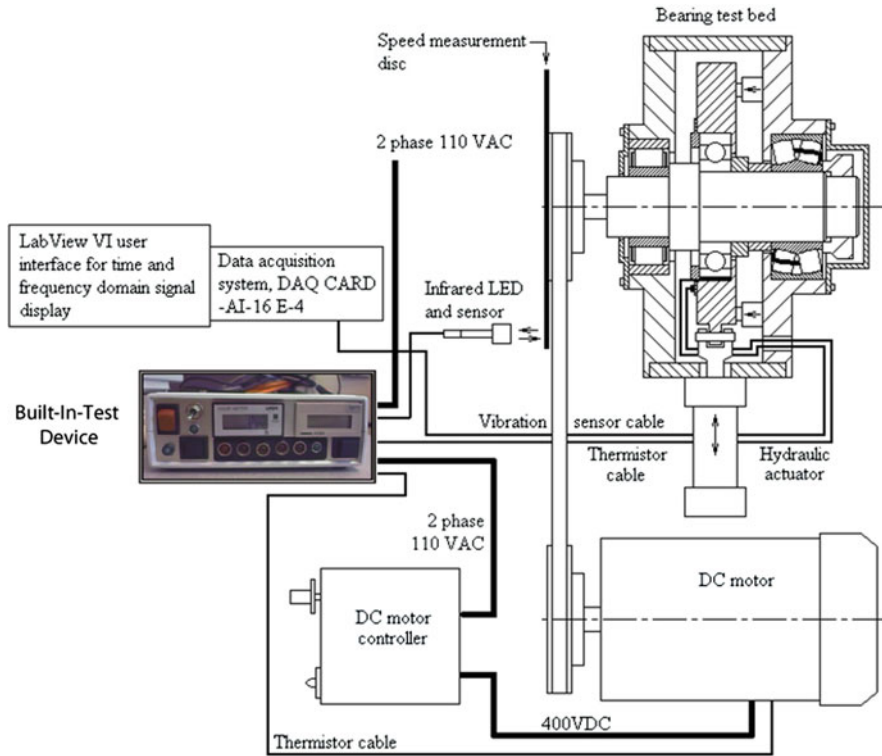


Fig. 7.4 Bearing test bed with hydraulic load application capability

condition of a surface spall. The relationship between the bearing rotational frequency f_r and the characteristic frequencies associated with defect-induced vibrations can be determined analytically as a function of the defect location, for example, on the inner raceway (f_{BPFI}), outer raceway (f_{BPFO}) or a rolling element (f_{BSF}) (Harris 1991). Specifically for the test bearing, these characteristic frequencies are calculated as $f_{BPFI} \approx 5.86f_r$, $f_{BPFO} \approx 4.1f_r$, $f_{BSF} \approx 5.3f_r$, respectively (SKF 1996). By identifying the existence of these characteristic frequencies and/or their combinations, the existence of bearing structural defects can be determined.

In the experimental evaluation, following aspects are studied:

1. The effectiveness of the unified time–scale–frequency analysis technique in extracting defect features (i.e., characteristic frequencies) from bearing vibration signals is compared to that of the Fourier transform and the discrete wavelet transform when it is applied alone.
2. The effectiveness of the new technique at different wavelet decomposition levels.
3. The effectiveness of the new technique under varying bearing operating conditions, such as the radial load, axial load, and shaft rotational speed.

7.3.1 Effectiveness in Defect Feature Extraction

To establish a basis for objective comparison, vibration signals are measured on both a defect-free (i.e., *healthy*) and a defective bearing of the same model (SKF 6220), under the same operation conditions: shaft speed $f_r = 600$ rpm (corresponding to 10 Hz rotational frequency), axial load of 7,038 N, and radial load of 17,468 N. Figures 7.5 and 7.6 shows the two signals in the time and frequency domains, respectively, while the related frequency resolution is approximately 0.3 Hz.

As shown in Fig. 7.6, both spectra indicate the existence of two dominant frequency components: (1) ball rotation (f_{BPFO1} , ball passing frequency on outer raceway, with the subscript “1” referring to the 6220 bearing) and (2) bearing misalignment (f_m). Ball rotation-related vibration has a peak value at the fundamental frequency $f_{\text{BPFO1}} = 41$ Hz, which is equal to $4.1 f_r$. Misalignment-related vibration has a characteristic frequency of $f_m = 20$ Hz ($= 2f_r$). In addition to these two major components, components related to bearing imbalance are also identified at the frequency $f_u = 10$ Hz (identical to f_r). However, visual comparison of the two spectra reveals no difference between them, as the characteristic frequency of $f_{\text{BPFI1}} = 58.6$ Hz, related to the inner raceway defect, is not recognized.

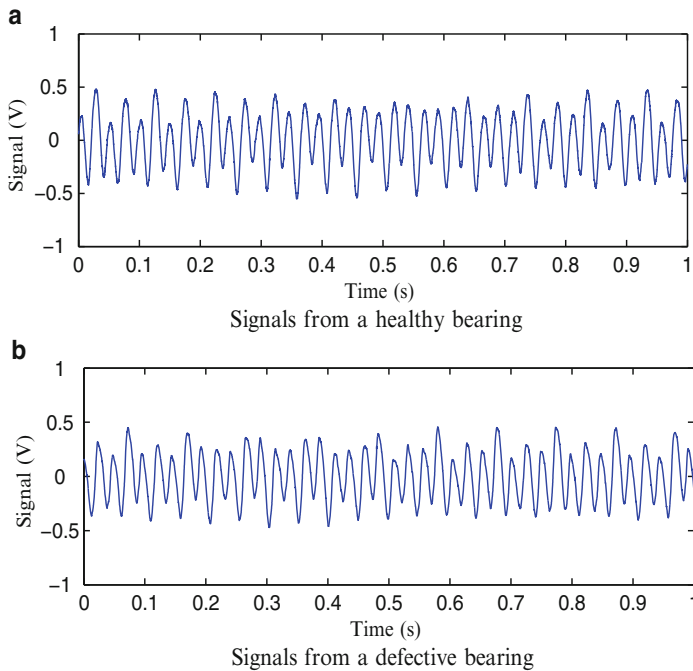


Fig. 7.5 Time domain signals from a healthy and a defective bearing. (a) Signals from a healthy bearing and (b) signals from a defective bearing

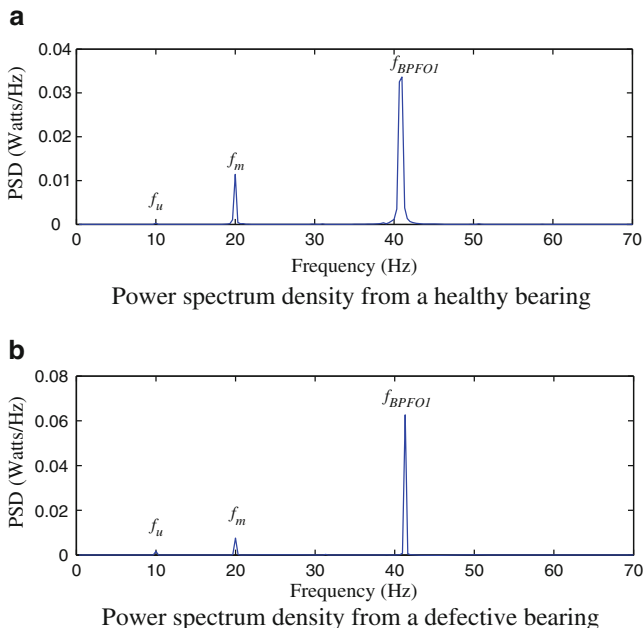


Fig. 7.6 Results of frequency domain analysis of the healthy and defective bearing. (a) Power spectrum density from a healthy bearing and (b) power spectrum density from a defective bearing

This illustrates that Fourier transform, when applied alone, may not be effective in detecting the existence of bearing structural defect.

The same signals are then analyzed using discrete wavelet transform, with the Daubechies 2 wavelet as the base wavelet. Figure 7.7 illustrates the wavelet coefficients of the vibration signals at the decomposition level 7, which has a corresponding frequency range of 39–78 Hz, thus covering the defect characteristic frequency of $f_{BPFI} = 58.6$ Hz. As seen in Fig. 7.7, the wavelet coefficients for the defective bearing have shown more dynamical variations than that of the healthy bearing. To quantify their difference, the root-mean-square (RMS) values of the two wavelet coefficients are calculated. It is found that the RMS value of the defective bearing (56 mV) is about 145% larger than that of the healthy bearing (22.8 mV). Although such an increase can be used as an indicator of structure defect in the bearing, it has the limitation that proper threshold needs to be set up a priori, to determine the quantitative extent that distinguishes a healthy bearing from a defective one. Another limitation of the wavelet transform is that the wavelet coefficients do not provide any indication on the specific location of the defect in the bearing, since it does not explicitly reveal the characteristic frequencies from the bearing.

Next, the bearing signal (as shown in Fig. 7.5) is analyzed using the unified time–scale–frequency method. For this purpose, the wavelet coefficients of the signal (shown in Fig. 7.7) is postprocessed using the Fourier transform.

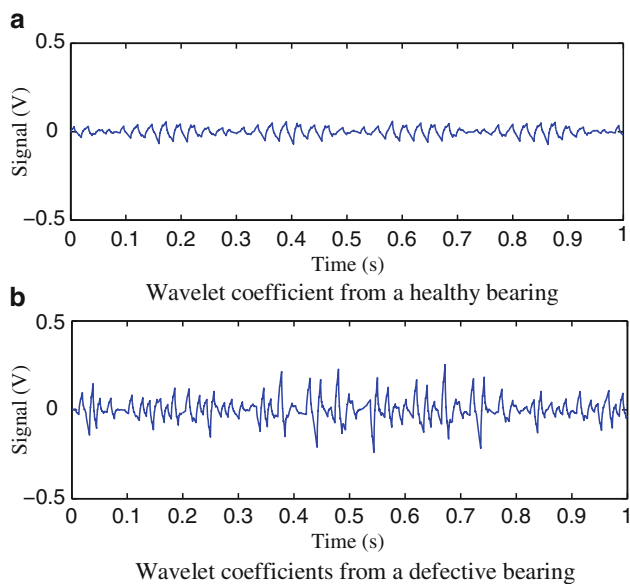


Fig. 7.7 Wavelet decomposition of bearing signals at decomposition level 7. (a) Wavelet coefficient from a healthy bearing and (b) wavelet coefficients from a defective bearing

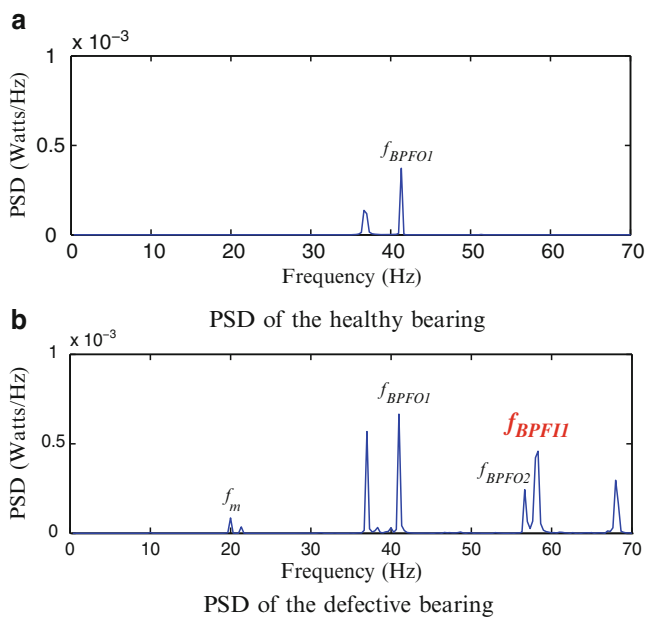


Fig. 7.8 Unified analysis for defect feature extraction at decomposition level 7. (a) PSD of the healthy bearing and (b) PSD of the defective bearing

Comparing the spectra of the healthy (Fig. 7.8a) and defective bearings (Fig. 7.8b), it is seen that the inner raceway defect can be clearly identified by its characteristic frequency at $f_{\text{BPF11}} = 58.6$ Hz. No distinctive peak is seen in the spectrum of the healthy bearing at this frequency. The spectrum further indicates several other major peaks at $f_m = 20$ Hz, $f_{\text{BPFO1}} = 41$ Hz, and $f_{\text{BPFO2}} = 56.5$ Hz. These are reflective of misalignment (at 20 Hz) of the defective bearing and rotational characteristic of other bearing. For example, $f_{\text{BPFO1}} = 41$ Hz is the ball passing frequency of the type 6220 bearing, and $f_{\text{BPFO2}} = 56.5$ Hz is found to be related to ball rotation of a different bearing (cylindrical bearing type 2322 with its vibration component indicated by the subscript “2”). This is based on the parameters of $Z = 14$, $D = 33.5$ mm, and $d_m = 175$ mm, and the characteristic frequency of the type 2322 bearing is calculated as $f_{\text{BPF12}} = 83.4$ Hz, $f_{\text{BPFO2}} = 56.5$ Hz, $f_{\text{BSF2}} = 50.1$ Hz. This bearing structurally supports the rotating shaft in the bearing test bed. Comparing with the Fourier analysis and the wavelet transform, the new, *unified* time–scale–frequency technique has shown to be more effective in extracting bearing defect features. In that it not only reveals the existence of a localized bearing defect, but also the defect characteristic frequency that is indicative of its specific location (e.g., inner raceway).

7.3.2 Selection of Decomposition Level

When evaluating the *unified* technique, a particular decomposition level (e.g., level 7) is chosen for the wavelet transform. The selection of an appropriate level is based on the signal sampling rate (or frequency) f_{sample} and the defect characteristic frequency f_{char} . The relationship is expressed as:

$$\frac{f_{\text{sample}}}{2^{L+1}} \leq f_{\text{Char}} \leq \frac{f_{\text{sample}}}{2^L} \quad (7.31)$$

where L denotes the wavelet decomposition level. As an example, when the sampling frequency is $f_{\text{sample}} = 10$ kHz, the frequency range associated with decomposition level $L = 7$ is calculated as 39–78 Hz. In Table 7.2, the frequency ranges covered by each of the decomposition levels under a 10 kHz sampling rate are shown. The essence of finding the best-suited decomposition level when wavelet transforming a dynamic signal is to ensure that its frequency range $[f_{\text{sample}}/2^{L+1}, f_{\text{sample}}/2^L]$ covers the characteristic frequency of structural defect in the bearing with the highest likelihood, if such a defect exists.

Table 7.3 lists the best-suited decomposition levels for analyzing bearing vibration signals specifically related to a localized inner raceway defect, under various bearing rotational (or shaft) speeds. Since at 600 rpm, the defect characteristic frequency $f_{\text{BPF11}} = 58.6$ Hz falls within the frequency range of 39–78 Hz, decomposition level 7 is chosen initially for data analysis.

Table 7.2 Frequency range associated with each decomposition level at a 10-kHz sampling rate

Decomposition level	Frequency range (Hz)	Decomposition level (<i>L</i>)	Frequency range (Hz)
1	2,500–5,000	5	156–312
2	1,250–2,500	6	78–156
3	625–1,250	7	39–78
4	312–625	8	19–39

Table 7.3 Best suited decomposition level for inner raceway defect frequency detection

Shaft speed (rpm)	f_{BPFI1} (Hz)	Decomposition level	Frequency range (Hz)
300	29.3	8	19–39
600	58.6	7	39–78
900	87.9	6	78–156
1,200	117.2	6	78–156
1,500	146.5	6	78–156

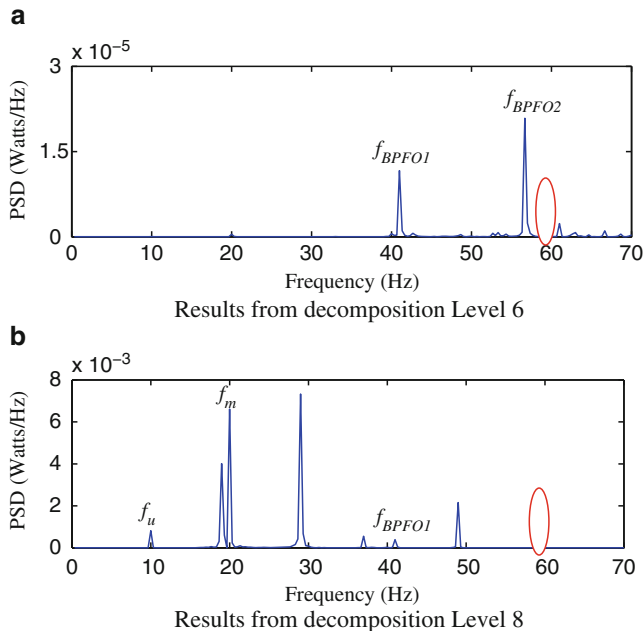


Fig. 7.9 Unified analysis using Daubechies 2 wavelet at levels 6 and 8. (a) Results from decomposition level 6 and (b) results from decomposition level 8

The importance of choosing proper decomposition level is illustrated in Fig. 7.9, where the results of decomposing defective bearing signal at levels 6 and 8 are shown. It is seen that none of the two levels (combined with the postspectral analysis) are able to identify the defect characteristic frequency at $f_{BPFI1} = 58.6$ Hz, due to the

mismatch between their respective frequency range (level 6 at 78–156 Hz and level 8 at 19–39 Hz) and the characteristic frequency of the inner raceway defect (at 58.6 Hz).

7.3.3 Effect of Bearing Operation Conditions

To investigate the effectiveness of the *unified* time–scale–frequency analysis method in defect feature extraction under varying bearing operating conditions, three groups of experiments are designed and conducted using a type 6220 ball bearing with a seeded defect.

7.3.3.1 Variation of Radial Load

The effect of radial loads on defect feature extraction is illustrated through the experimental results shown in Fig. 7.10, where four levels of radial loads are presented. It is noted that, as the radial load has progressively increased from 4,367 to 26,202 N, the peak value of the bearing defect frequency of f_{BPFI1} (= 58.6 Hz) has grown by 609.5%, as given in Table 7.4. Such an increase can be explained by the increased preload applied by the rolling elements of the bearing to the defect on the raceway. An increased preload enhances the impacts when the rolling elements roll over the defect, leading to increased amplitude of the defect feature.

7.3.3.2 Variation of Axial Load

Increase in the axial load has also shown to lead to an increase of the defect feature amplitude, as is evident when comparing the three different axial load conditions in Fig. 7.11. For example, when the axial load applied to the bearing increases from 0 to 4,192 N, the defect signal amplitude at defect characteristic frequency f_{BPFI1} has increased by 4.7%, from 3.18×10^{-3} to 3.33×10^{-3} W/Hz, as listed in Table 7.5. This is because the application of axial load on the bearing increases the extent of the bearing load zone distribution, resulting in an increased number of defect impacts within the load zone (Harris 1991). Such an increase enhances the overall energy content of the defect signal at its feature frequency, and is reflected by the increased defect feature amplitude.

7.3.3.3 Variation of Rotational Speed

The power spectral density graphs in Fig. 7.12 illustrate the effect of bearing rotational speed on the defect signal strength. As the speed increased from 300 to 1,200 rpm, the defect frequency amplitude increased by 71.2%, as listed in

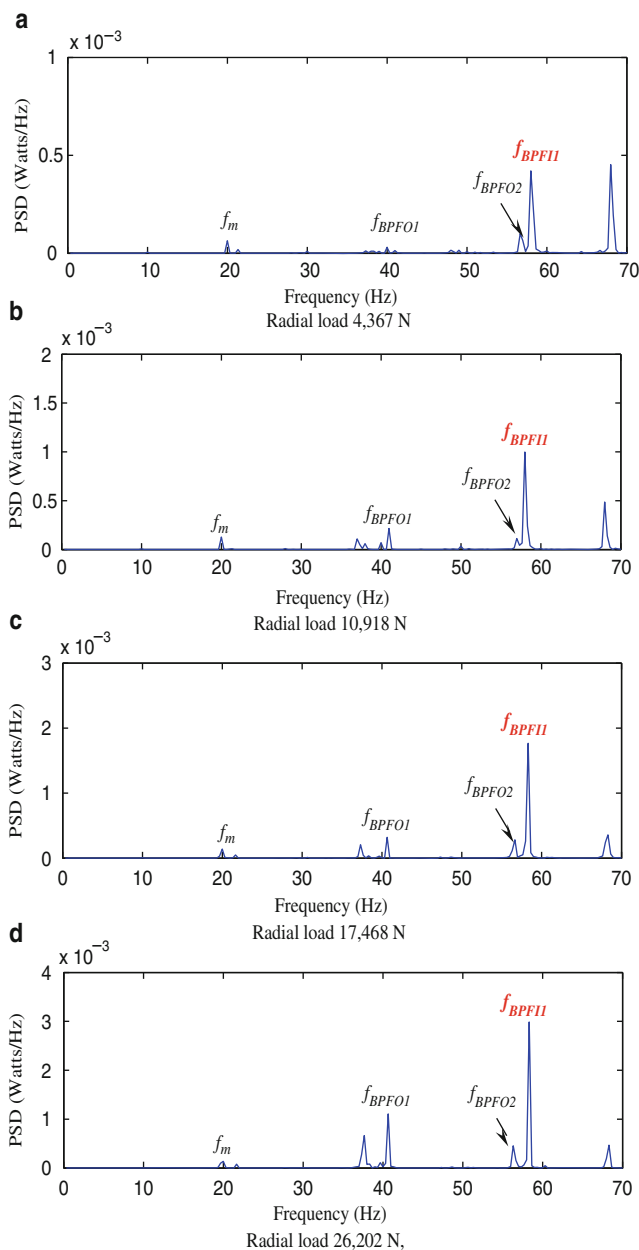


Fig. 7.10 Effect of the radial load on defect feature amplitude. (a) Radial load 4,367 N, (b) radial load 10,918 N, (c) radial load 17,468 N, and (d) radial load 26,202 N

Table 7.4 Effect of radial load on the defect signal amplitude (f_{BPF11})

Radial load (N)	PSD 10^{-3} (W/Hz)	Percentage of increase
4,367	0.42	—
10,918	0.99	135.7
17,468	1.77	321.4
26,202	2.98	609.5

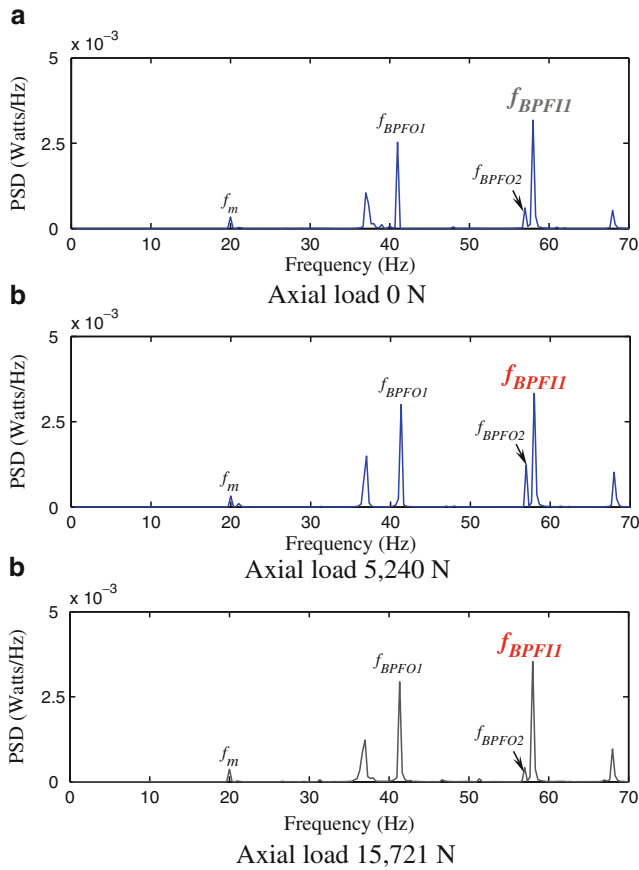


Fig. 7.11 Effect of axial load on defect feature amplitude. (a) Axial load 0 N, (b) axial load 5,240 N, and (c) axial load 15,721 N

Table 7.5 Effect of axial load on the defect feature amplitude (f_{BPF11})

Axial load	PSD 10^{-3} (W/Hz)	Percentage of increase
0	3.18	—
5,240	3.33	4.7
15,721	3.54	11.3

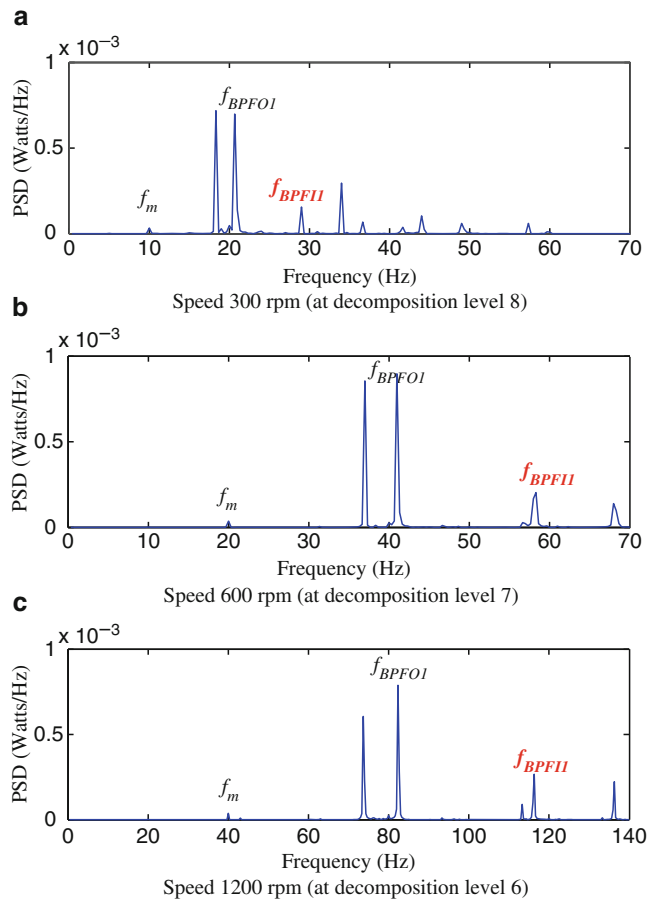


Fig. 7.12 Effect of bearing speed on defect amplitude. (a) Speed 300 rpm (at decomposition level 8), (b) speed 600 rpm (at decomposition level 7), and (c) Speed 1,200 rpm (at decomposition level 6)

Table 7.6 Effect of speed on the defect signal amplitude (f_{BPFI1})		
Shaft speed (rpm)	PSD 10^{-4} (W/Hz)	Percentage of increase
300	1.56	—
600	2.03	30.1
1,200	2.67	71.2

Table 7.6. The speed-related defect feature amplitude increase can be explained by the fact that with the increase of the speed, the number of impacts per bearing revolution increases proportionally, hence the total amount of defect impact energy also increases, leading to increased peak amplitude at the defect characteristic frequency f_{BPFI1} .

7.4 Summary

This chapter introduces a *unified* time–scale–frequency analysis technique based on the combination of discrete wavelet transform with frequency domain postprocessing. The effectiveness of this technique in improving bearing defect diagnosis is then investigated. A localized defect of 0.25 mm in diameter at the inner raceway of a type 6220 bearing has been successfully detected, under various bearing operation conditions. It is shown that the Fourier-transform-based spectral analysis technique alone is not reliable to detect the transient components that are characteristic of localized bearing defect, whereas wavelet transform alone does not explicitly identify the specific location of the defect. Thus, the *unified* technique combines the advantages of both the time and frequency domain analyses and provides more information on the defect feature than each of the techniques employed individually. In addition to bearing defect diagnosis, the new technique provides a powerful tool for the detection, extraction, and identification of weak “defect” features submerged in vibration signals in a wide range of manufacturing equipment

7.5 References

- Bracewell, R (1999) The Fourier transform and its applications, 3rd edn. McGraw-Hill, New York
- Byrne G, O'Donnell GE (2007) An integrated force sensor solution for process monitoring of drilling operations. *CIRP Ann Manuf Technol* 56(1):89–92
- Cavacece M, Introini A (2002) Analysis of damage of ball bearings of aeronautical transmissions by auto-power spectrum and cross-power spectrum. *ASME J Vib Acoust* 124(2):180–185
- Daubechies I (1992) Ten lectures on wavelets. SIAM, Philadelphia, PA
- Gao R, Yan R (2006) Non-stationary signal processing for bearing health monitoring. *Int J Manuf Res* 1(1):18–40
- Ge M, Du, R, Zhang GC, Xu YS (2004) Fault diagnosis using support vector machine with an application in sheet metal stamping operations. *Mech Syst Signal Process* 18(1):143–159
- Gibson J (1999) Principle of digital and analog communication, 2nd edn. Prentice Hall, Inc, Upper Saddle River, NJ
- Harris TA (1991) Rolling bearing analysis, 3rd edn. Wiley, New York
- Ho D, Randall RB (2000) Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mech Syst Signal Process*, 14(5):763–788
- Holm-Hansen BT, Gao R, Zhang L (2004) Customized wavelet for bearing defect detection. *ASME J Dyn Syst Meas Control* 126(6):740–745
- Kaiser G (1994) A friendly guide to wavelets. Birkhäuser, Boston, MA
- Malekian M, Park SS, Jun M (2009) Tool wear monitoring of micro-milling operations. *J Mater Process Technol* 209:4903–4914
- Mori K, Kasashima N, Yoshioka T, Ueno Y (1996) Prediction of spalling on a ball bearing by applying the discrete wavelet transform to vibration signals. *Wear* 195(1–2):162–168
- Obikawa T, Shinozuka J (2004) Monitoring of flank wear of coated tools in high speed machining with a neural network ART2. *Int J Mach Tools Manuf* 44:1311–1318
- Orhan S, Aktürk N, Çelik V (2006) Vibration monitoring for defect diagnosis of rolling element bearings as a predictive maintenance tool: comprehensive case studies. *NDTE Int* 39:293–298
- SKF Company (1996) SKF bearing maintenance handbook. SKF Company, Denmark
- Tandon T, Choudhury A (1999) A review of vibration and acoustic measurement methods for the defection of defects in rolling element bearings. *Tribol Int* 32:469–480