

Chapter 2

Linear Time-Frequency Analysis

The identification and quantification of the oscillatory components present in a given signal represents a classic problem of signal processing, and one of the most successful approaches for solving it has been through time-frequency analysis. Thus, it is often useful to study the signal's structure in both time and frequency simultaneously, which can be done by considering a specifically constructed projection of the signal onto the time-frequency plane—its time-frequency representation. Such an approach proves to be especially suitable for signals containing many oscillatory components with time-varying amplitudes and/or frequencies, which is a very common scenario for real-life signals [1, 25, 37, 48, 49].

Consider, for example, the activity of the heart: although beating with frequency of 1 Hz on average, the length of the period between the consecutive beats—and therefore the frequency—is time-dependent; moreover, the corresponding variations of the beat-to-beat interval are of high clinical relevance [36]. These variations and their properties are straightforward to analyse in the time-frequency plane, where time-dependent frequency of the heart can be easily traced [9, 10, 17, 18, 25, 26, 50]. The same applies to many other real-life signals, so the TFRs are now routinely applied in almost every area of science, from image processing and finance to geophysics and the life sciences [1, 2, 25, 32–34, 37, 43, 45, 47, 49, 51, 55].

General aspects of time-frequency analysis, and the properties of the different existing TFRs, have been thoroughly discussed in a number of excellent books [1, 8, 14, 16, 21, 22, 30, 37, 46] and reviews [13, 23, 41, 43, 51]. There exist two main TFR classes: linear (such as the WFT and WT [1, 16, 37]) and quadratic (such as the Wigner-Ville [54, 56], Rihaczek [42] and Choi-Williams [12] distributions, see [8, 22, 23] for a more comprehensive list). Although each of these has its own advantages, only the WFT and WT will be considered in this work. This is because, in contrast to quadratic representations, these types offer an easy and straightforward way of extracting and reconstructing individual components present in the signal (see Chap. 3), which property will be of crucial importance in what follows.

The present chapter presents a thorough discussion of the WFT and WT, their implementation, properties and related issues, providing all the information that is needed to understand them and apply them effectively. Particular attention is paid to

the practical aspects, but the necessary mathematical theory is also reviewed. The basic definitions, such as the AM/FM component and the analytic signal, as well as the need for time-frequency analysis, are first discussed in Sect. 2.1. The WFT and WT are then defined in Sect. 2.2, which also reviews their basic properties and illustrates the difference between the two. The resolution properties of the WFT and WT are considered in Sect. 2.3, while Sect. 2.4 discusses various aspects of “real-life” TFR usage, namely: signal preprocessing that should be performed; frequency discretization and how to do it appropriately; boundary distortions due to finite signal length; and limits on the frequencies that can be studied reliably in the WFT and WT. The results of the whole chapter are summarized in Sect. 2.5.

Remark 2.0.1 It should be noted, that many advanced methods based on different TFRs have been developed over the last few decades. Examples are the wavelet bispectral analysis [28, 29], wavelet coherence [20, 33, 35] and phase coherence [3, 5, 44] etc. However, it seems in principle impossible to review appropriately all such techniques, given their large number, so they are not considered here. Rather, the emphasis in the present Chapter is placed on the basic time-frequency representations, which provide the foundation for more sophisticated methods, including nonlinear mode decomposition developed later.

2.1 AM/FM Components and the Analytic Signal

One of the basic notions of time-frequency analysis is the *AM/FM component* (or simply *component*), which is defined as a function of time t of form

$$x(t) = A(t) \cos \phi(t) \quad (\forall t : A(t) > 0, \nu(t) \equiv \phi'(t) > 0). \quad (2.1)$$

The time-dependent values $A(t)$, $\phi(t)$ and $\nu(t) \equiv \phi'(t)$ are then called the instantaneous amplitude, phase and frequency of the component (2.1) (for a more detailed discussion of their definitions and related issues see [6, 7, 40]).

Given that the signal is known to be of the form (2.1), the natural question is how to find its associated $A(t)$, $\phi(t)$ and $\nu(t)$. The most convenient way of doing this is the analytic signal approach. However, before considering it, a few additional notions should be introduced. Thus, for an arbitrary function $f(t)$, its Fourier transform (FT), positive and negative frequency parts, time-average and standard deviation will be denoted as $\hat{f}(\xi)$, $f^+(t)$, $f^-(t)$, $\langle f(t) \rangle$ and $\text{std}[f(t)]$, respectively:

$$\begin{aligned} \hat{f}(\xi) &\equiv \int_{-\infty}^{\infty} f(t) e^{-i\xi t} dt \Leftrightarrow f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\xi) e^{i\xi t} d\xi = \langle f(t) \rangle + f^+(t) + f^-(t), \\ f^+(t) &\equiv \frac{1}{2\pi} \int_{0+}^{\infty} \hat{f}(\xi) e^{i\xi t} d\xi, \quad f^-(t) \equiv \frac{1}{2\pi} \int_{-\infty}^{0-} \hat{f}(\xi) e^{i\xi t} d\xi, \\ \langle f(t) \rangle &= \frac{\int f(t) dt}{\int dt}, \quad \text{std}[f(t)] \equiv \sqrt{\langle [f(t)]^2 \rangle - [\langle f(t) \rangle]^2}, \end{aligned} \quad (2.2)$$

where, here and in what follows, the integrals are taken over $(-\infty, \infty)$ if unspecified (or, in practice, over the full time duration of $f(t)$). A simple example is $f(t) = a + b \cos \nu t = a + be^{i\nu t}/2 + be^{-i\nu t}/2$, for which one has $\hat{f}(\xi)/2\pi = a\delta(\xi) + b\delta(\xi - \nu)/2 + b\delta(\xi + \nu)/2$, $\langle f(t) \rangle = a$, $f^\pm(t) = be^{\pm i\nu t}/2$ and $\text{std}[f(t)] = b/\sqrt{2}$. Note, that if $f(t)$ is real, then $f(t) - \langle f(t) \rangle = 2 \text{Re } f^\pm(t)$ and $\hat{f}(\xi) = [\hat{f}(-\xi)]^* \Rightarrow f^+(t) = [f^-(t)]^*$, where the star denotes complex conjugation.

For a given signal $s(t)$ (which is always assumed to be real in this work), its doubled positive frequency part is called its *analytic signal* and will be denoted as $s^a(t)$:

$$s^a(t) \equiv 2s^+(t) \quad (s(t) = \langle s(t) \rangle + \text{Re}[s^a(t)]). \quad (2.3)$$

The analytic signal is complex, so its dynamics can easily be separated into amplitude and phase parts. For signals represented by a single component (2.1), the analytic amplitude and phase $A^a(t)$, $\phi^a(t)$ match closely the true amplitude and phase $A(t)$, $\phi(t)$, thus providing an easy way to estimate them:

$$A(t) \approx A^a(t) \equiv |s^a(t)|, \quad \phi(t) \approx \phi^a(t) \equiv \arg[s^a(t)]. \quad (2.4)$$

The approximate equality (2.4) will be called the analytic approximation, and it can alternatively be formulated as $[A(t) \cos \phi(t)]^+ \approx A(t)e^{i\phi(t)}/2$. As stipulated by the Bedrosian theorem [4], this approximation is exact when the spectrum of $A(t)$ lies lower than the spectrum of $e^{i\phi(t)}$, and there are no intersections between the two. For example, in the case of amplitude modulation only, $s(t) = A(t) \cos(\nu t + \varphi)$, (2.4) gives the exact amplitude and phase if all the spectral content of $A(t)$ lies lower than ν , i.e. $\hat{A}(\xi \geq \nu) = 0$. Usually, however, there is a small discrepancy between the true and analytic amplitude/phase (considered in detail in Sect. 7.1), but it is often very small. Thus, there is arguably still nothing better than the analytic signal approach for amplitude and phase estimation in the case of a single AM/FM component (2.1) [52] (except, maybe, the recently proposed direct quadrature method [24]).

However, real-life signals rarely consist of only one component, and they usually also contain noise. In this case, the analytic signal will represent a mix of the amplitude and phase dynamics of all components contained in the signal (additionally corrupted by noise), so their individual parameters cannot be recovered from it. One should therefore employ more sophisticated techniques, able to distinguish the different components within a single time-series. This can be done by using the TFR-based approaches that will be described below. The main assumption behind them (and time-frequency analysis more generally) is that the signal is represented by a sum of AM/FM components, each of which satisfies the analytic approximation (2.4), plus some noise $\eta(t)$:

$$s(t) = \sum_i x_i(t) + \eta(t) = \sum_i A_i(t) \cos \phi_i(t) + \eta(t), \quad (2.5)$$

$$\forall t, i : A_i(t) > 0, \phi_i'(t) > 0, [A_i(t) \cos \phi_i(t)]^+ \approx A_i(t)e^{i\phi_i(t)}/2.$$

Although signal representation (2.5) is not unique, in practice one aims at the sparsest among such representations, i.e. the one characterized by the smallest number of components $x_i(t)$. It is also important to note, that the most accurate estimates of the components' parameters obtainable using any TFR-based method are the corresponding analytic estimates (2.4), so the (best achievable) goal of such methods is to extract the separate analytic signals $x_i^a(t)$ for each of the chosen $x_i(t)$ in (2.5).

The above discussion was related to problems that arise while considering multi-component signals in the time domain. To make the story complete, the problems that arise when trying to treat such signals in the frequency domain should also be addressed. In practice, any signal can be represented by its Fourier expansion (discrete FT), i.e. as the sum of tones—monochromatic signals $A_i \cos(\nu_i t + \varphi_i)$ —with different constant amplitudes A_i , phase shifts φ_i and frequencies ν_i . The same is obviously true for a single AM/FM component, and its Fourier expansion is directly related to the spectrum of the corresponding amplitude and frequency modulations:

$$\begin{aligned}
 x(t) &= A \left(1 + \sum_a r_a \cos(\nu_a t + \varphi_a) \right) \cos(\nu t + \varphi + \sum_b r_b \sin(\nu_b t + \varphi_b)) \\
 &= / \text{expressing } \cos \phi = (e^{i\phi} + e^{-i\phi})/2, \ e^{ia \sin \phi} = \sum_{n=-\infty}^{\infty} J_n(a) e^{in\phi} / \\
 &= A e^{i(\nu t + \varphi)} \left[1 + \sum_a \frac{r_a (e^{i(\nu_a t + \varphi_a)} + e^{-i(\nu_a t + \varphi_a)})}{2} \right] \prod_b \sum_{n_b} J_{n_b}(r_b) e^{i(n_b \nu_b t + n_b \varphi_b)} + c.c. \\
 &= \sum_{\{n_b\}=\{n_1, n_2, \dots\} \in \mathbb{Z}} \left[\tilde{A}_{\{n_b\}} e^{i\tilde{\varphi}_{\{n_b\}}} e^{i\tilde{\nu}_{\{n_b\}} t} + \sum_a \frac{r_a}{2} \tilde{A}_{\{n_b\}} (e^{i(\tilde{\varphi}_{\{n_b\}} + \varphi_a)} e^{i(\tilde{\nu}_{\{n_b\}} + \nu_a) t} \right. \\
 &\quad \left. + e^{i(\tilde{\varphi}_{\{n_b\}} - \varphi_a)} e^{i(\tilde{\nu}_{\{n_b\}} - \nu_a) t}) \right] + c.c., \tag{2.6}
 \end{aligned}$$

where

$$\begin{aligned}
 \tilde{A}_{\{n_b\}} &\equiv A \prod_b J_{n_b}(r_b) = A J_{n_1}(r_1) J_{n_2}(r_2) \dots, \\
 \tilde{\nu}_{\{n_b\}} &\equiv \nu + \sum_b n_b \nu_b, \quad \tilde{\varphi}_{\{n_b\}} \equiv \varphi + \sum_b n_b \varphi_b, \tag{2.7}
 \end{aligned}$$

and $J_n(r_b) = (-1)^n J_{-n}(r_b)$ denote Bessel functions of the first kind, while *c.c.* stands for the complex conjugate of the preceding expression. Thus, any component (2.6) can be represented as a sum of tones with frequencies $|\nu + \sum_b n_b \nu_b|$ and $|\nu \pm \nu_a + \sum_b n_b \nu_b|$ (for all possible combinations of $a \in \mathbb{N}$ and $\{n_b\} = \{n_1, n_2, \dots\} \in \mathbb{Z}$). But, obviously, not any sum of tones can be represented as the AM/FM component. What make this possible in (2.6) are the specific relationships between the tones' amplitudes, phases and frequencies, which encode the corresponding amplitude and frequency modulations of the component.

Clearly, the representation of $x(t)$ (2.6) as a single entity is more compact and meaningful than its representation as the sum of tones, which corresponds to its FT. Thus, for a multi-component signal $s(t)$ (2.5), each component is encoded as a set of

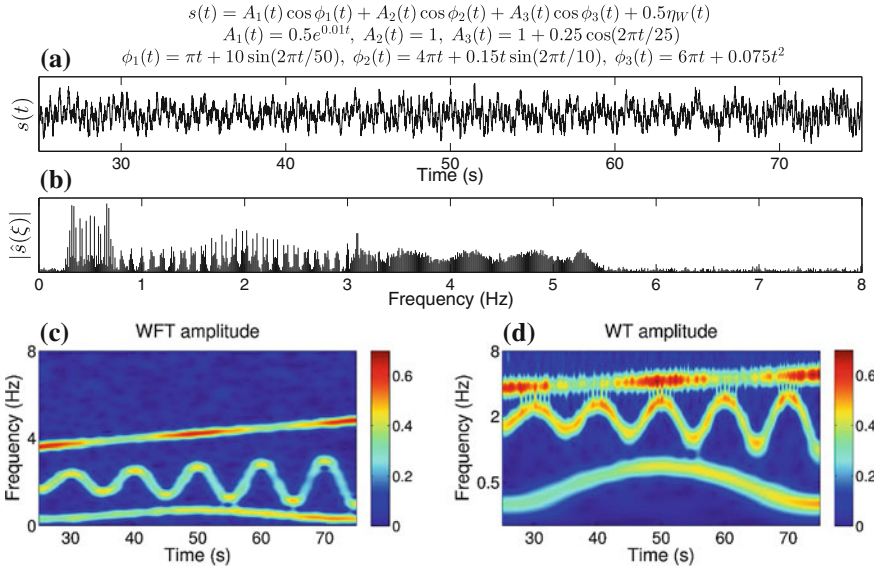


Fig. 2.1 Different representations of a signal composed of three AM/FM components and corrupted by noise $0.5\eta_W(t)$, where $\eta_W(t)$ denotes Gaussian white noise of unit deviation; the parameters of the components are described above the *top panel*. **a** Signal in the time domain; **b** signal in the frequency domain, given by its Fourier transform; **c**, **d** signal in the time-frequency domain, given by its WFT and WT (see Sects. 2.2.1 and 2.2.2 below), respectively. Note that for the WT (**d**) the frequency scale is logarithmic

entries in the associated frequency representation $\hat{s}(\xi)$, but it is usually unclear which entries correspond to what component, and which ones are attributable to noise. This is illustrated in Fig. 2.1, which shows an example of signal representation in the time, frequency, and time-frequency domains, with the latter being given by its WFT and WT, to be discussed below. As can be seen, although all representations by definition contain the same amount of information about the signal, in the case of Fig. 2.1 the most readily interpretable view of this information is provided in the time-frequency domain.

2.2 Time-Frequency Representations (TFRs)

As discussed in the previous section, instead of studying a signal in either one of the time ($s(t)$) or frequency ($\hat{s}(\xi)$) domains, it is often more useful to study its properties in time and frequency simultaneously. This can be done by considering specifically constructed projections of the signal onto a time-frequency plane, i.e. TFRs. Such an approach gives the possibility of tracking the evolution of the signal's spectral content in time, which is typically represented by variations of the amplitudes and frequencies of the components from which the signal is composed.

In what follows, attention will be paid exclusively to the two main linear TFRs—the WFT and the WT—because of their special suitability for the extraction of components (the latter will be discussed in more detail in the next Chapter). Note that in the present Sect. the notions of time, frequency and time-frequency resolutions of the TFR will sometimes be used; they will be further clarified in Sect. 2.3 below. There are also many classical books and reviews (e.g. [8, 22, 30, 37]) where an additional discussion of the WFT/WT and their resolution characteristics can be found.

2.2.1 Windowed Fourier Transform (WFT)

The windowed Fourier transform (WFT), also called the short-time Fourier transform or (in a particular form) the Gabor transform, is one of the oldest and thus most-investigated linear TFRs [19, 22, 37]. The WFT $G_s(\omega, t)$ of the signal $s(t)$ can be calculated either in the time domain or in the frequency domain as [27]:

$$G_s(\omega, t) = \int s^+(\tau)g(\tau - t)e^{-i\omega(\tau-t)}d\tau = \frac{1}{2\pi} \int_0^\infty e^{i\xi t} \hat{s}(\xi) \hat{g}(\omega - \xi) d\xi, \quad (2.8)$$

where $g(t)$ is the specified *window function* and $\hat{g}(\xi)$ is its FT. Without loss of generality, the $|\hat{g}(\xi)|$ is considered to have a maximum at $\xi = 0$:

$$\omega_g \equiv \operatorname{argmax} |\hat{g}(\xi)| = 0, \quad (2.9)$$

which ensures that, for a single tone $s(t) = \cos \nu t$, the amplitude of the WFT has a maximum at the frequency $\omega = \nu$ (see (2.10) below). If this is not the case, one should always set ω_g (2.9) to zero manually by considering the demodulated window function $\{g(t), \hat{g}(\xi)\} \rightarrow \{g(t)e^{-i\omega_g t}, \hat{g}(\xi + \omega_g)\}$. An example of the WFT can be seen in Fig. 2.1c.

Remark 2.2.1 Note, that the definition of the WFT (2.8) is slightly different from the conventional form. Thus, only the positive frequency part $s^+(t)$ is taken into account instead of the full $s(t)$, for the reasons that are thoroughly discussed in [27]. Basically, such modification is needed to remove the undesirable interference with negative frequencies (e.g. with $e^{-i\nu t}$ if $s(t) = \cos \nu t$), which might seriously corrupt the representation if it has low resolution in frequency. Nevertheless, in practice this interference is often very small, so that taking the full signal in (2.8) will not usually change anything significantly. Additionally, instead of the traditional $e^{-i\omega\tau}$, which corresponds to “static” phases of the WFT coefficients, there is $e^{-i\omega(\tau-t)}$ in (2.8), which corresponds to “dynamic” phases and is thus more convenient in terms of component reconstruction.

It is easy to see that, for a signal represented by a sum of tones, the corresponding WFT (2.8) takes the form

$$s(t) = \sum_n A_n \cos(\nu_n t + \varphi_n) \Rightarrow G_s(\omega, t) = \frac{1}{2} \sum_n A_n \hat{g}(\omega - \nu_n) e^{i(\nu_n t + \varphi_n)}, \quad (2.10)$$

from which one can gain a first impression of how it works. It should be noted that, although peaks in $|G_s(\omega, t)|$ are all at positive frequencies (as follows from (2.10) and the assumptions made), the WFT (2.8) is defined on the full interval $\omega \in (-\infty, \infty)$, so that one can calculate it for $\omega < 0$ as well (e.g. to trace the peak's tails). Usually, however, this is not needed.

The WFT is an invertible transform, so that the original signal in both time and frequency domains can be recovered from it as (see Sect. 7.3):

$$\begin{aligned} s^a(t) &= C_g^{-1} \int G_s(\omega, t) d\omega, \quad s(t) = \langle s(t) \rangle + \text{Re}[s^a(t)], \\ \hat{s}(\omega > 0) &= \tilde{C}_g^{-1} \int G_s(\omega, t) e^{-i\omega t} dt, \quad \hat{s}(-\omega) = \hat{s}^*(\omega), \\ C_g &\equiv \frac{1}{2} \int \hat{g}(\xi) d\xi = \pi g(0), \quad \tilde{C}_g \equiv \int g(t) dt = \hat{g}(0). \end{aligned} \quad (2.11)$$

Among the window functions $g(t)$ that can be used for calculating WFT (2.8) (see Sect. 7.2 for a list of common types and their properties), the most suitable choice seems to be the Gaussian window

$$g(t) = \frac{1}{\sqrt{2\pi} f_0} e^{-t^2/2f_0^2} \Leftrightarrow \hat{g}(\xi) = e^{-f_0^2 \xi^2/2}. \quad (2.12)$$

The time and frequency resolution of the window (2.12) are determined by its spread in time and frequency, which are controlled by its *resolution parameter* f_0 . Therefore, increasing f_0 improves the frequency resolution of the resultant WFT, but reduces its time resolution.

The Gaussian window is commonly used on account of its unique property of maximizing the “classic” time-frequency resolution of the transform (to be discussed in Sect. 2.3); the form (2.12), however, seem to have very good resolution properties in general, and not only within the traditional definitions (see Sect. 7.2). In what follows, unless otherwise specified, all simulations are performed using the Gaussian window (2.12) with $f_0 = 1$. Nonetheless, the considerations and formulas of this Chapter apply quite generally for any window function.

In numerical applications, given the signal $s(t_n)$, $n = 1, \dots, N$ sampled at frequency f_s , its WFT is calculated using the fast Fourier transform (FFT) algorithm, utilizing the frequency domain form of (2.8). The frequency axis is first divided into equidistant bins $\omega_k = (k - k_0)\Delta\omega$, where $\Delta\omega$ is the frequency bin width (its optimal choice and related issues are discussed in Sect. 2.4.2). One then computes the signal's discrete FT $\hat{s}(\xi) = \{\hat{s}(\xi_n)\}$, where $\xi_n/2\pi = (n/N - 1/2)f_s$, $n = 1, \dots, N$ are the corresponding discrete frequencies, and sets $\hat{s}(\xi_n \leq 0) = 0$. Finally, taking the

inverse FT of $\{s(\xi_n)\hat{g}(\omega_k - \xi_n)\}$ at each frequency ω_k gives the full WFT time evolution $G_s(\omega_k, t_{n=1,\dots,N})$ for that frequency bin. A complete numerical implementation, including all relevant issues, is summarized in Sect. 7.5.

2.2.2 Wavelet Transform (WT)

The (continuous) wavelet transform (WT) is the other well-known linear TFR which, in contrast to the WFT, has logarithmic frequency resolution [15, 16, 37]; in other respects the two TFRs are quite similar. The WT $W_s(\omega, t)$ of a signal $s(t)$ can be calculated as [27]:

$$W_s(\omega, t) = \int s^+(\tau) \psi^*\left(\frac{\omega(\tau - t)}{\omega_\psi}\right) \frac{\omega d\tau}{\omega_\psi} = \frac{1}{2\pi} \int_0^\infty e^{i\xi t} \hat{s}(\xi) \hat{\psi}^*(\omega_\psi \xi / \omega) d\xi, \quad (2.13)$$

where $\psi(t)$ is the chosen *wavelet function*, and

$$\omega_\psi \equiv \operatorname{argmax} |\hat{\psi}(\xi)| \quad (2.14)$$

denotes the wavelet peak frequency. Additionally, the wavelet should satisfy the *admissibility condition*

$$\hat{\psi}(0) = \int \psi(t) dt = 0, \quad (2.15)$$

the need for which will become clear below. An example of the WT can be seen in Fig. 2.1d.

Remark 2.2.2 It should be noted, that the definition (2.13) represents a modified version of the traditional one. Thus, the WT is usually defined using the scales $a(\omega)$, while in (2.13) they are already expressed through the frequencies as $a(\omega) = \omega_\psi / \omega$ (this relation establishes that, in the simplest case $s(t) = A \cos(\nu t + \varphi)$, the modulus of the WT $|W_s(\omega, t)|$ will be peaked exactly at the tone frequency $\omega = \nu$). Next, similarly to the WFT (2.8), the WT is calculated using only the positive frequency part of the signal $s^+(t)$ (this is in fact equivalent to using the full signal $s(t)$ but setting $\hat{\psi}(\xi \leq 0) = 0$, though many wavelets, called analytic, have this property from the beginning). The reasoning behind this is the same as for the case of the WFT (see Remark 2.2.1). Finally, there are a few ways of normalizing the WT, and the most commonly used one corresponds to (2.13) multiplied by $\sqrt{\omega_\psi / \omega}$. However, in the latter case the WT amplitude will be biased towards lower ω , amplifying/reducing the peaks associated with each component depending on its frequency. Such a situation does not arise for the normalization (2.13), which therefore seems to be more natural and convenient, especially in terms of the extraction and reconstruction of components discussed in Chap. 3 below. Nevertheless, the two normalizations differ

only in terms of convenience and ease of understanding: the correct usage of each will by definition give the same results.

Comparing (2.13) with (2.8), it is clear that the continuous WT can be viewed simply as the WFT with a frequency-varying window, which feature nevertheless gives rise to a different type of resolution. Thus, the analogue of (2.10) for WT (2.13) is

$$s(t) = \sum_n A_n \cos(\nu_n t + \varphi_n) \Rightarrow W_s(\omega, t) = \frac{1}{2} \sum_n A_n \hat{\psi}^*\left(\omega_\psi \frac{\nu_n}{\omega}\right) e^{i(\nu_n t + \varphi_n)}. \quad (2.16)$$

As can be seen, in contrast to the WFT, which takes into account the frequency difference $\nu - \omega$ (2.10), the WT considers the ratio ν/ω (or the difference between logarithms $\log \nu/\omega = \log \nu - \log \omega$), which amounts to the definition of the logarithmic frequency resolution. The natural frequency scale is therefore also logarithmic for the WT, in contrast to the linear scale of the WFT.

The reconstruction formulas in the case of the WT become (see Sect. 7.3)

$$\begin{aligned} s^a(t) &= C_\psi^{-1} \int_0^\infty W_s(\omega, t) \frac{d\omega}{\omega}, \quad s(t) = \langle s(t) \rangle + \text{Re}[s^a(t)], \\ \hat{s}(\omega > 0) &= \tilde{C}_\psi^{-1} \int W_s(\omega, t) e^{-i\omega t} dt, \quad \hat{s}(-\omega) = \hat{s}^*(\omega), \\ C_\psi &\equiv \frac{1}{2} \int_0^\infty \hat{\psi}^*(\xi) \frac{d\xi}{\xi}, \quad \tilde{C}_\psi \equiv \left[\int \psi(t) e^{-i\omega_\psi t} dt \right]^* = \hat{\psi}^*(\omega_\psi). \end{aligned} \quad (2.17)$$

Note that, in contrast to the WFT (2.11), the signal is reconstructed from the WT by integration of $W_s(\omega, t)$ over the logarithmic scale $d\omega/\omega = d \log \omega$, which is standard for the WT-based measures.

From (2.16) and (2.17), the need for the admissibility condition (2.15) now becomes clear. Thus, for wavelets not satisfying (2.17) the value of C_ψ in (2.17) is infinite, and signal reconstruction from the WT becomes impossible. Even more importantly, from (2.16) it follows that, if the wavelet FT does not vanish at zero frequency $\hat{\psi}(0) \neq 0$, then each component will be spread over the whole frequency range of the WT, leading to a highly corrupted representation. For example, for a single tone signal $s(t) = A \cos \nu t$, the WT amplitude approaches $|W_s(\omega, t)| = A |\hat{\psi}(\omega_\psi \nu/\omega)| \rightarrow A |\hat{\psi}(0)|$ as $\omega \rightarrow \infty$. Hence, $|\hat{\psi}(0)|$ determines the minimum level to which the WT amplitude for each component decays: if it is nonzero then even components with infinitely distant frequencies will interfere with each other. Drawing an analogy with the WFT, use of the wavelet with $\hat{\psi}(0) \neq 0$ corresponds to the use of the window with $\hat{g}(\xi \rightarrow \infty) \neq 0$ in (2.8), which is obviously inappropriate. Each wavelet should therefore be admissible, i.e. should satisfy (2.15).

Remark 2.2.3 It is often not recognized that, in contrast to the case of the WFT, rescaling of the wavelet function $\{\psi(t), \hat{\psi}(\xi)\} \rightarrow \{\psi(rt), \hat{\psi}(\xi/r)\}$ has no effect on

the resultant WT. Thus, what matters is $\hat{\psi}(\omega_{\psi}\nu/\omega)$, while direct rescaling changes both $\hat{\psi}(\xi)$ and the peak frequency $\omega_{\psi} \rightarrow r\omega_{\psi}$, having no overall effect when substituted in (2.13) and (2.16). As an illustration, the parameter k in the wavelet $\hat{\psi}(\xi) = \xi^a e^{-k\xi^b}$ is completely redundant, because this wavelet can be represented in the form $\hat{\psi}(\xi) = k^{-a/b} (k^{1/b}\xi)^a e^{-(k^{1/b}\xi)^b}$ which, up to the constant multiplier, is equivalent to $\hat{\psi}(\xi) = \xi^a e^{-\xi^b}$. To really change the time and frequency resolutions of the wavelet, one needs to change either its ω_{ψ} while trying to preserve the spread and the form of $|\hat{\psi}(\xi)|$, or *vice versa* (see Sect. 2.3.1). Furthermore, these two approaches are completely equivalent, as can be seen from (2.13), so it is redundant to define parameters controlling both the peak frequency of a wavelet and its spread.

Among the wavelet functions $\psi(t)$ used for the continuous WT (2.13) (see Sect. 7.2 for a list of common types and their properties), the most popular choice is the Morlet wavelet [38]

$$\begin{aligned}\psi(u) &= \frac{1}{\sqrt{2\pi}} (e^{i2\pi f_0 u} - e^{-(2\pi f_0)^2/2}) e^{-u^2/2}, \\ \hat{\psi}(\xi) &= e^{-(\xi - 2\pi f_0)^2/2} \left(1 - e^{-2\pi f_0 \xi}\right),\end{aligned}\tag{2.18}$$

where the *admissibility term* $\sim e^{-(2\pi f_0)^2/2}$ is needed to satisfy (2.15). The Morlet wavelet is regarded as the wavelet analogy to the Gaussian window (2.12), and is used so commonly because of the widespread belief that it has the best resolution properties. This is, however, not true. Thus, as will be discussed in Sect. 2.3 below, the wavelet (2.18) does indeed nearly maximize the “classic” time-frequency resolution (provided the admissibility term is small), but this traditional measure was originally devised for the WFT and is completely inappropriate for the WT.

Taking into account the logarithmic frequency resolution of the WT, a more “correct” wavelet analogy of the Gaussian window (2.12) would be the lognormal wavelet

$$\hat{\psi}(\xi > 0) \sim e^{-(2\pi f_0 \log \xi)^2/2}, \quad \omega_{\psi} = 1,\tag{2.19}$$

where the resolution parameter f_0 has a meaning similar to that for the Gaussian window (2.12), controlling the time and frequency resolutions of the resultant WT.

It turns out (see Sect. 7.2) that the resolution properties of the wavelet (2.19) are generally slightly better than that of the Morlet wavelet (2.18). Apart from this, the lognormal wavelet has a variety of other advantages. Thus, it is “infinitely admissible” (in contrast to the Morlet wavelet), i.e. all its moments $\int \xi^{-n} \hat{\psi}(\xi) d\xi/\xi$ ($n \geq 0$) are finite; this allows direct reconstruction of any order time-derivatives of the component’s amplitude and phase from its WT (see Sects. 3.2 and 7.3 below). Additionally, the wavelet (2.19) is analytically tractable, allowing one to obtain C_{ψ} (2.17) and many other quantities in the explicit form. Therefore, while the considerations of the present Chapter are applicable to any wavelet type, in what follows, unless otherwise specified, all the simulations and examples will be based on the lognormal wavelet with $f_0 = 1$.

In numerical applications, given the signal $s(t_n)$, $n = 1, \dots, N$ sampled at frequency f_s , its WT is calculated using the frequency domain form of (2.13) and taking advantage of the FFT algorithm. The procedure is quite similar to that for the WFT calculation. The frequency axis is first partitioned into equilogspaced bins $\omega_k/2\pi = 2^{(k-k_0)/n_v}$, where the *number-of-voices* n_v determines the fineness of this partition and has the meaning of the number of frequency bins in each diadic interval (its selection is discussed in Sect. 2.4.2). One then computes the signal's discrete FT $\hat{s}(\xi) = \{\hat{s}(\xi_n)\}$, where $\xi_n/2\pi = (n/N - 1/2)f_s$, $n = 1, 2, \dots, N$ denote the corresponding discrete frequencies, and one sets $\hat{s}(\xi_n \leq 0) = 0$. Finally, taking the inverse Fourier transform of $\{\hat{s}(\xi_n)\hat{\psi}^*(\omega_\psi \xi_n/\omega_k)\}$ at each frequency ω_k gives the full time-evolution of the WT $W_s(\omega_k, t_{n=1,\dots,N})$ for this frequency bin. A complete numerical implementation, including all relevant issues, is summarized in Sect. 7.5.

2.2.3 Difference Between the WFT and the WT

Figure 2.2 compares the WFT and WT calculated for the same signal. As mentioned before, the main distinction between the two TFR types lies in their different kinds of frequency resolution: linear for the WFT, and logarithmic for the WT. Thus, if the signal contains two tones $\cos \nu_{1,2}t$, then the WFT takes account of their frequency difference $\nu_2 - \nu_1$, whereas what matters for the WT is their frequency ratio ν_2/ν_1 (or difference between logarithms of their frequencies). For example, if there are two tones with frequencies 0.05 and 0.125 Hz, corresponding to periods of 20

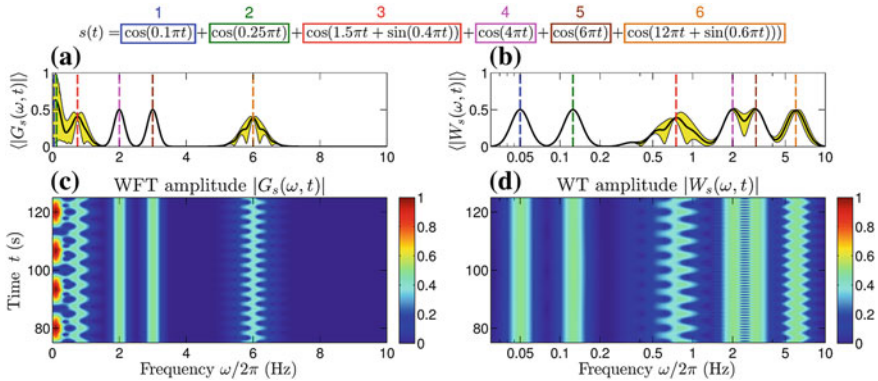


Fig. 2.2 Comparison of the WFT and WT of the same signal consisting of six components, as shown above the figure. **a, b** Mean WFT and WT amplitudes, respectively; the *yellow regions* indicate 95% ranges for the corresponding amplitudes at each frequency; the *vertical dashed lines* show mean frequencies of the components, with their *colors* being linked to the *colors of boxes* at the *top* of the figure. **c, d** Full WFT and WT amplitudes, respectively. Note, that for the WT in (**b, d**) the frequency scale is logarithmic, in contrast to the linear scale for the WFT in (**a, c**). The signal was sampled for 200 s at 50 Hz

and 8 s (components 1,2 in Fig. 2.2), then they will be much better resolved in the WT ($\nu_2/\nu_1 - 1 = 1.5$) than in the WFT ($(\nu_2 - \nu_1)/2\pi = 0.075$). At the same time, if there are two tones with frequencies 2 and 3 Hz, corresponding to periods 0.5 and ≈ 0.33 s (components 4 and 5 in Fig. 2.2), then they will be much better resolved in the WFT ($(\nu_2 - \nu_1)/2\pi = 1$ Hz) than in the WT ($\nu_2/\nu_1 - 1 = 0.5$).

On the other hand, the WFT has the same time resolution at each frequency, while that for the WT increases with frequency, being proportional to ω (see Sect. 2.3 below). This means that, for fixed window/wavelet parameters, the higher the frequency, the more time-variability is allowed for components to be reliably represented in the WT, while the WFT does not discriminate in this respect. For example, if one has a frequency-modulated component $\cos(\nu t + \sin \nu_b t)$, the WFT accounts for the value of ν_b , while the WT considers ν_b/ν . Thus, if one has a component with a mean frequency of 0.75 Hz subject to sinusoidal modulation at 0.2 Hz (component 3 in Fig. 2.2), it will be better represented in the WFT ($\nu_b/2\pi = 0.2$) than in the WT ($\nu_b/\nu \approx 0.27$). On the other hand, if one has a component with mean frequency 6 Hz and frequency modulation at 0.3 Hz (component 6 in Fig. 2.2), it will be represented more reliably in the WT ($\nu_b/\nu = 0.05$) than in the WFT ($\nu_b/2\pi = 0.3$).

Remark 2.2.4 Note, that the assessment of the quality of representation of two tones in the WFT and the WT based on the direct comparison between the corresponding $(\nu_2 - \nu_1)/2\pi$ and $\nu_2/\nu_1 - 1$ is valid only within the assumption that the WFT and the WT have the same frequency resolution properties at $\omega = 2\pi$ (see Sect. 2.3). This is true for the case of Fig. 2.2, where the WFT and WT are calculated using Gaussian window (2.12) and lognormal wavelet (2.19) with the same resolution parameters $f_0 = 1$, respectively. Likewise, in the discussion about representation of the frequency-modulated components it was assumed that the time resolution properties of the WFT and the WT are the same at $\omega = 2\pi$, which is almost true in the present case (see Sect. 2.3 below). The related considerations, however, are only approximately valid because of the difference between the fixed and frequency-dependent time resolutions of the WFT and the WT, respectively.

Obviously, the resolution parameter f_0 (or other window/wavelet parameters, if present) can always be adjusted to improve the representation of any one chosen component. However, there is often no universal f_0 suitable for all components present in the signal, so the choice between the WFT and WT depends on how many such components can in principle be represented reliably. Therefore, contrary to what is sometimes thought, the WT is not in general superior to the WFT: it just considers all on a logarithmic frequency scale, and whether or not this is more useful than the linear frequency resolution of the WFT depends on the signal structure.

Summarizing, the WT is most suitable when the ratios of frequencies of the underlying oscillations and their relative (to frequency) modulations are more consistent with each other than the corresponding frequency differences and absolute amplitude/frequency variations. In other words, the WT is preferred to the WFT when the lower-frequency components are less time-varying and closer in frequency than the components at higher frequencies. Otherwise, the WFT appears to be more appropriate (see also the related discussion and additional examples on pp. 126–134 of

[37]). Note, however, that many real signals have structure that is more suited to studies based on the WT. Additionally, one can often analyse a variety of time-series of different kinds using the same wavelet parameters, while for the WFT the window parameters should be adjusted for each particular case. Usually, the WT is also computationally cheaper due to the logarithmic frequency scale, requiring fewer bins to cover the same frequency range. On the other hand, the WFT generally has better resolution properties (see Sect. 7.2).

2.3 Time-, Frequency- and Time-Frequency Resolution

The notions of time, frequency and time-frequency resolutions have been often used in the preceding section, but were defined only briefly. In this section different resolution characteristics are considered in detail, and the definitions are made more precise.

2.3.1 General Formulation

Consider a signal $s_{\nu\nu}(t)$ ($s_{\tau\tau}(t)$) consisting of two frequency events—tones (time events—delta-peaks), so that its WFT (2.8) and WT (2.13) are

$$\begin{aligned} s_{\nu\nu}(t) &= \cos(\nu t) + \cos((\nu + \Delta\nu)t + \Delta\varphi) \\ \Rightarrow \begin{cases} G_s(\omega, t) = \frac{1}{2}[\hat{g}(\omega - \nu) + \hat{g}(\omega - \nu - \Delta\nu)e^{i\Delta\nu t}e^{i\Delta\varphi}]e^{i\nu t}, \\ W_s(\omega, t) = \frac{1}{2}[\hat{\psi}^*(\omega_\psi\nu/\omega) + \hat{\psi}^*(\omega_\psi(\nu + \Delta\nu)/\omega)e^{i\Delta\nu t}e^{i\Delta\varphi}]e^{i\nu t}, \end{cases} \end{aligned} \quad (2.20)$$

$$\begin{aligned} s_{\tau\tau}(t) &= \delta(t - \tau) + \delta(t - \tau - \Delta\tau)e^{i\Delta\varphi} \\ \Rightarrow \begin{cases} G_s(\omega, t) \cong [g(\tau - t) + g(\tau + \Delta\tau - t)e^{-i\omega\Delta\tau}e^{i\Delta\varphi}]e^{i\omega(\tau-t)}, \\ W_s(\omega, t) \cong \frac{\omega}{\omega_\psi}[\psi^*(\omega(\tau - t)/\omega_\psi) + \psi^*(\omega(\tau + \Delta\tau - t)/\omega_\psi)e^{i\Delta\varphi}]. \end{cases} \end{aligned} \quad (2.21)$$

Everywhere in this section, e.g. in (2.21), the symbol “ $\cong$ ” denotes equality up to an error associated with the difference between the WFT/WT (2.8), (2.13) and its form as calculated using the full signal $s(t)$ instead of the corresponding positive frequency part $s^+(t)$; see [27] for a detailed discussion of this issue and the quality of the approximation in (2.21). Note that, if redefining WFT/WT to use the full signal, the approximate equality will migrate from $s_{\tau\tau}(t)$ (2.21) to $s_{\nu\nu}(t)$ (2.20) [27].

Remark 2.3.1 Although in this work the original signal is assumed to be real, $s_{\tau\tau}(t)$ (2.20) represents an exception. This is because, without the phase shift introduced by $e^{i\Delta\varphi}$, the analogy between $s_{\tau\tau}(t)$ and $s_{\nu\nu}(t)$ would be incomplete while, intuitively, the time and frequency domains should have equal rights. In fact, since the WFT (2.8) and WT (2.13) are based on the positive frequency part of the signal, the

resultant TFR will remain the same if instead of $s_{\nu\nu}(t)$ (2.20) one considers $s_{\nu\nu}^+(t) = [e^{i\nu t} + e^{i(\nu+\Delta\nu)t} e^{i\Delta\varphi}]/2$. The latter has the Fourier transform $\hat{s}_{\nu\nu}^+(\xi) = [\delta(\xi - \nu) + \delta(\xi - \nu - \Delta\nu)e^{i\Delta\varphi}]/2$, which is fully analogous to $s_{\tau\tau}(t)$. Nevertheless, it turns out that in all the following considerations nothing changes if one restricts $s_{\tau\tau}(t)$ to be real, which is satisfied for $\Delta\varphi = 0, \pi$ in (2.21). For example, in Sect. 2.3.4 below, averaging over these two values ($\Delta\varphi = 0, \pi$) in the corresponding reconstruction errors will give the same result as averaging over the full range $\Delta\varphi \in [0, 2\pi]$.

It seems reasonable to define the time (frequency) resolution γ_t (γ_ω) of the transform as the reciprocal of the minimum time (frequency) difference $\Delta\tau$ ($\Delta\nu$) in $s_{\tau\tau}(t)$ ($s_{\nu\nu}(t)$) for which two delta-peaks (tones) can still be reliably resolved in the TFR:

$$\gamma_t(\nu, \tau) = 1/\Delta\tau_{\min}(\nu, \tau), \quad \gamma_\omega(\nu, \tau) = 1/\Delta\nu_{\min}(\nu, \tau). \quad (2.22)$$

However, the meaning of “reliably resolved” still remains imprecise, and will be dealt with later. Note, that the definitions (2.22) represent the most general case, where the resolutions and minimal differences are “localized”, i.e. allowed to depend on both time and frequency. This is not only useful for the WT, which is characterized by frequency-dependent resolution properties, but also allows one to consider more complicated cases, e.g. the WFT with a time-dependent resolution parameter $f_0 \rightarrow f_0(t)$.

The joint time-frequency resolution $\gamma_{\omega t}$ is most often defined as the reciprocal of the area of the minimal resolvable square $[\nu, \nu + \Delta\nu_{\min}(\nu, \tau)] \times [\tau, \tau + \Delta\tau_{\min}(\nu, \tau)]$, being equal to the product of the time and frequency resolutions. However, within such a square $\gamma_t(\nu, \tau)$ and $\gamma_\omega(\nu, \tau)$ can vary considerably, so that these variations should be taken into account to make $\gamma_{\omega t}$ more meaningful. The latter can therefore be defined as

$$\gamma_{\omega t}(\nu, \tau) = \left[\frac{\int_{\nu}^{\nu+\Delta\nu_{\min}(\nu, \tau)} \Delta\tau_{\min}(\omega, \tau) d\omega \int_{\tau}^{\tau+\Delta\tau_{\min}(\nu, \tau)} \Delta\nu_{\min}(\nu, t) dt}{\Delta\tau_{\min}(\nu, \tau) \Delta\nu_{\min}(\nu, \tau)} \right]^{-1}. \quad (2.23)$$

Note that, if neither $\Delta\nu_{\min}$ nor $\Delta\tau_{\min}$ depends on time or frequency, then one has the traditional $\gamma_{\omega t} = \gamma_\omega \gamma_t$.

The definitions (2.22) and (2.23) remain valid for any signal representation, not only the WFT and WT. For example, the time domain representation $s(t)$ has infinite time resolution, since, theoretically, two delta-peaks can be resolved for any $\Delta\tau$ in $s_{\tau\tau}(t)$, but zero frequency resolution, as both tones in $s_{\nu\nu}(t)$ have nonzero entries at almost all times, and thus cannot in principle be resolved, no matter how large $\Delta\nu$ is. On the other hand, the frequency domain representation (Fourier transform) $\hat{s}(\xi)$ has infinite frequency resolution, since (in theory) two tones can be perfectly resolved in $\hat{s}_{\nu\nu}(\xi)$ for however small frequency difference $\Delta\nu$; but, at the same time, it has zero time resolution, because the two delta-peaks are spread over all frequencies in $\hat{s}_{\tau\tau}(\xi)$.

TFRs, on the other hand, represent a “mix” of the time domain and frequency domain representations. As a result, both time and frequency resolutions are finite

for them. This can clearly be seen from the WFT and WT for the two-peak and two-tone signals (2.21), (2.20). Thus, since the window/wavelet function has non-zero supports in both time and frequency, delta-peaks in $s_{\tau\tau}(t)$ (tones in $s_{\nu\nu}(t)$) will interfere in the WFT/WT for small enough $\Delta\tau$ ($\Delta\nu$), making their accurate separation and reconstruction impossible (e.g. see the components 1, 2 in Fig. 2.2a).

From (2.20) and (2.21) one can see that the interference between two tones or between two delta-peaks in the WFT does not depend on ν or τ . Therefore, for any meaningful definition of $\Delta\tau_{\min}$ and $\Delta\nu_{\min}$, they also should not depend on time or frequency. Furthermore, it is also clear that rescaling of the WFT window function $g(t) \rightarrow g(t/r) \Leftrightarrow \hat{g}(\xi) \rightarrow \hat{g}(r\xi)$ changes its resolution properties as $\{\Delta\tau_{\min}, \Delta\nu_{\min}\} \rightarrow \{r\Delta\tau_{\min}, \Delta\nu_{\min}/r\}$. This reflects the repeatedly-mentioned trade-off between time and frequency resolutions, being a manifestation of the time-frequency uncertainty principle, which excludes the possibility of simultaneous sharp localization in time and frequency [30, 37]. Thus, without changing the form of the window function, there is only the possibility of rescaling both resolutions, increasing one and decreasing the other, but not treating them separately. Their product, which in the case of the WFT is equal to the joint time-frequency resolution $\gamma_{\omega t}$ (2.23), remains fixed under such rescaling, thus representing an important characteristic of the window function: the higher it is, the better the trade-off that is possible.

Using the same arguments as for the WFT, from (2.21) and (2.20) it follows that for the WT both $\Delta\tau_{\min}$ and $\Delta\nu_{\min}$ depend on frequency, but not time, so that $\gamma_t = \gamma_t(\nu)$, $\gamma_\omega = \gamma_\omega(\nu)$. Next, one also has $\Delta\tau_{\min}(r\nu) = \Delta\tau_{\min}(\nu)/r$ and $\Delta\nu_{\min}(r\nu) = r\Delta\nu_{\min}(\nu)$, so that the time resolution of the WT increases with frequency, while the frequency resolution decreases. Nevertheless, as will be seen, the time-frequency resolution (2.23) does not depend on ν or τ , being fixed for the specified wavelet parameters.

As mentioned above (see Remark 2.2.3), the direct rescaling $\{\psi(t), \hat{\psi}(\xi)\} \rightarrow \{\psi(t/r), \hat{\psi}(r\xi)\}$ does not change anything for the WT, and to tune the resolutions at a particular frequency one needs to change the wavelet peak frequency $\omega_\psi \rightarrow r\omega_\psi$ while preserving the spreads and the forms of $\{\psi(t), \hat{\psi}(\xi)\}$ (or vice versa). If such a procedure was possible, it would lead to the same trade-off rule as for the WFT, i.e. $\gamma_t(\nu) \rightarrow \gamma_t(\nu)/r$, $\gamma_\omega(\nu) \rightarrow r\gamma_\omega(\nu)$, with $\gamma_{\omega t}$ remaining fixed. However, given restrictions such as the admissibility condition (2.15), changing the wavelet peak frequency will inevitably affect its form and/or its spread (on the linear scale). Thus, e.g. for the Morlet wavelet (2.18) a change of ω_ψ (by varying f_0) will be accompanied by a simultaneous change of its form because of the admissibility term; for the lognormal wavelet ω_ψ is fixed, while changing f_0 will change wavelet spread in frequency on the logarithmic scale, but on the linear scale $\hat{\psi}(\xi)$ will become more asymmetric, affecting the time domain form $\psi(t)$. Because of these issues, there does not seem to be a way of changing the resolution properties of the WT without altering its time-frequency resolution, in contrast to the WFT. Furthermore, $\gamma_{\omega t}$ (whichever way defined) progressively worsens with an increase of the wavelet time resolution (decrease of f_0), and it seems in principle impossible to reach a very sharp time-localization in the WT (see Sect. 7.2).

2.3.2 Classical Definitions and Their Flaws

Although $\Delta\tau_{\min}$ and $\Delta\nu_{\min}$ in (2.22) were defined respectively as the minimum time and frequency difference which can reliably be resolved in the TFR, the meaning of “reliably resolved” remains mathematically unclear. Based on how it is defined, one can characterize the resolution properties of the TFR in different ways.

The traditional approach [30, 37] implicitly assumes two delta-peaks (tones) to be well-resolved if the time (frequency) distance between them exceeds some number of standard deviations of the squared window/wavelet function in the time (frequency) domain. Within this framework, for the WFT one has

$$\begin{aligned}\Delta\nu_{\min}^{(cl)} &= k_1 \Delta\omega, \quad \Delta\tau_{\min}^{(cl)} = k_2 \Delta t, \\ \Delta\omega^2 &= E_g^{-1} \frac{1}{2\pi} \int (\omega - \omega_c)^2 |\hat{g}(\omega)|^2 d\omega, \quad \omega_c \equiv E_g^{-1} \frac{1}{2\pi} \int \omega |\hat{g}(\omega)|^2 d\omega, \\ \Delta t^2 &= E_g^{-1} \int (t - t_c)^2 |g(t)|^2 dt, \quad t_c \equiv E_g^{-1} \int t |g(t)|^2 dt, \\ E_g &\equiv \frac{1}{2\pi} \int |\hat{g}(\xi)|^2 d\xi = \int |g(t)|^2 dt \quad (\text{by Parseval's identity}),\end{aligned}\tag{2.24}$$

where $k_{1,2}$ are implicitly assumed to be the same for all window functions. Obviously, such a definition is far from universal, since for different windows different number of standard deviations are needed to resolve the two tones/delta-peaks. As an illustrative example, the window with asymptotics $|\hat{g}(\xi \rightarrow \pm\infty)| \sim |\xi|^{-1.25}$ has infinite variance of $|\hat{g}(\xi)|^2$, but can still be used and allows for an accurate resolution and reconstruction of the two tones for high enough frequency difference between them.

For the WT, the classic variance-based framework takes the form

$$\begin{aligned}\Delta\nu_{\min}^{(cl)}(\nu) &= k_1 \frac{\nu}{\omega_\psi} \Delta\omega, \quad \Delta\tau_{\min}^{(cl)}(\nu) = k_2 \frac{\omega_\psi}{\nu} \Delta t, \\ \Delta\omega^2 &= E_\psi^{-1} \frac{1}{2\pi} \int (\omega - \omega_c)^2 |\hat{\psi}(\omega)|^2 d\omega, \quad \omega_c \equiv E_\psi^{-1} \frac{1}{2\pi} \int \omega |\hat{\psi}(\omega)|^2 d\omega, \\ \Delta t^2 &= E_\psi^{-1} \int (t - t_c)^2 |\psi(t)|^2 dt, \quad t_c \equiv E_\psi^{-1} \int t |\psi(t)|^2 dt, \\ E_\psi &\equiv \frac{1}{2\pi} \int |\hat{\psi}(\xi)|^2 d\xi = \int |\psi(t)|^2 dt \quad (\text{by Parseval's identity}),\end{aligned}\tag{2.25}$$

where $k_{1,2}$ are again implicitly assumed to be the same for all wavelet functions. This approach has the same drawbacks as (2.24) for the WFT. However, in the case of the WT it is actually not appropriate at all, at least in terms of the frequency resolution. Thus, as can be seen from (2.16), the tones are represented in the WT as terms $\sim \hat{\psi}(\omega_\psi \nu / \omega)$, so that the decay of their contribution as $\omega \rightarrow \infty$, determined by the behavior of $\hat{\psi}(\xi)$ as $\xi \rightarrow 0$, will obviously have a big effect on the frequency resolution. At the same time, the usual variance $\Delta\omega^2$ (2.25) takes no account of

this fact, e.g. being invariant under $\hat{\psi}(\xi) \rightarrow \hat{\psi}(\xi + \omega_\psi)$, that makes the wavelet inadmissible (in which case tones that are infinitely distant in frequency still interfere, so that the frequency resolution becomes effectively zero). Therefore, for wavelets, it seems more appropriate to study at least the variance of $|\hat{\psi}(\omega_\psi/\xi)|^2$, but by no means that of $|\hat{\psi}(\xi)|^2$.

For both WFT and WT, the classic time-frequency resolution measure is taken as $\gamma_{\omega t}^{(cl)} = [\Delta_\omega \Delta_t]^{-1}$, with Δ_ω, Δ_t being given by (2.24) for the WFT and by (2.25) for the WT (note also the difference between $\gamma_{\omega t}^{(cl)}$ and (2.23) for the latter). It can be shown [1, 30, 37], that this measure attains its maximum for the Gaussian window (2.12) and (up to the effect of the admissibility term $\sim e^{-(2\pi f_0)^2/2}$) for the Morlet wavelet (2.18). However, as follows from the discussion above, only in the case of the WFT does the classic $\gamma_{\omega t}^{(cl)}$ make some sense, though even in this case it remains highly non-universal.

2.3.3 Notion of the Window/Wavelet ϵ -Support

Before proceeding to a reconsideration of the classic definitions, it is useful to introduce the notions of the window/wavelet ϵ -supports in frequency $[\xi_1(\epsilon), \xi_2(\epsilon)]$ and time $[\tau_1(\epsilon), \tau_2(\epsilon)]$, which will be used frequently in what follows. These ϵ -supports are defined as the widest intervals containing the $(1 - \epsilon)$ part of the total integrals of the window/wavelet function which appear in $C_{g,\psi}$ and $\tilde{C}_{g,\psi}$ (2.11), (2.17). As will be seen below, they are directly related to the accuracy with which the components can be recovered from the WFT/WT, and thus can be used effectively for quantifying it.

Considering first the WFT, for an arbitrary window function, including functions that are not always positive and can be oscillating or complex, the corresponding definitions are

$$\begin{aligned} R_g(\omega) &\equiv \frac{\int_{-\infty}^{\omega} \hat{g}(\xi) d\xi}{\int \hat{g}(\xi) d\xi} = C_g^{-1} \frac{1}{2} \int_{-\infty}^{\omega} \hat{g}(\xi) d\xi, \\ \xi_{1,2}(\epsilon) : |R_g(\xi \leq \xi_1)| &\leq \epsilon/2, \quad |1 - R_g(\xi \geq \xi_2)| \leq \epsilon/2, \\ P_g(\tau) &\equiv \frac{\int_{-\infty}^{\tau} g(t) dt}{\int g(t) dt} = \tilde{C}_g^{-1} \int_{-\infty}^{\tau} g(t) dt, \\ \tau_{1,2}(\epsilon) : |P_g(\tau \leq \tau_1)| &\leq \epsilon/2, \quad |1 - P_g(\tau \geq \tau_2)| \leq \epsilon/2. \end{aligned} \tag{2.26}$$

Evidently, $|R_g(\omega)|$ and $|1 - R_g(\omega)|$ quantify the relative parts of $\hat{g}(\xi)$ that are contained in the ranges $\xi < \omega$ and $\xi > \omega$, respectively, while the values $\xi_{1,2}(\epsilon)$ specify the limits within which the $(1 - \epsilon)$ part of the window FT resides. In the same manner, $|P_g(\tau)|$ and $|1 - P_g(\tau)|$ reflect the relative parts of $g(t)$ contained in the ranges $t < \tau$ and $t > \tau$, respectively, while $[\tau_1(\epsilon), \tau_2(\epsilon)]$ represents the region encompassing its $(1 - \epsilon)$ part. The inequalities in the definitions of $\xi_{1,2}(\epsilon)$ ($\tau_{1,2}(\epsilon)$) are needed

only if $\hat{g}(\xi)$ ($g(t)$) is not strictly positive, or complex, to ensure that the integral of the latter over any frequency (time) region containing the ϵ -support $[\xi_1(\epsilon), \xi_2(\epsilon)]$ ($[\tau_1(\epsilon), \tau_2(\epsilon)]$) will always approximate the corresponding full integral with relative error not higher than ϵ .

Considering the single tone $s(t) = \cos \nu t \Rightarrow G_s(\omega, t) = \hat{g}(\omega - \nu)e^{i\nu t}/2$, it is clear that its WFT at frequencies $\omega < \xi$ ($\omega > \xi$) will contain the $|R_g(\xi - \nu)|$ ($|1 - R_g(\xi - \nu)|$) part of the signal. Furthermore, its $(1 - \epsilon)$ part will be contained in the frequency range $[\nu + \xi_1(\epsilon), \nu + \xi_2(\epsilon)]$, so that e.g. for real $\hat{g}(\xi)$ the corresponding signal reconstructed by (2.11) from the WFT in this range will be $(1 - \epsilon) \cos \nu t$.

Likewise, for the delta-peak $s(t) = \delta(t - \tau) \Rightarrow G_s(\omega, t) \cong g(\tau - t)e^{i\omega(\tau - t)}$ the WFT at times $t < t_0$ ($t > t_0$) will contain the $\cong |1 - P_g(\tau - t_0)|$ ($\cong |P_g(\tau - t_0)|$) part of the signal, while its $(1 - \epsilon)$ part will be contained in the time interval $\cong [\tau - \tau_2(\epsilon), \tau - \tau_1(\epsilon)]$. Thus, e.g. for the real $g(t)$ the delta-peak's FT reconstructed by (2.11) from this interval will be $\hat{s}(\xi) \cong (1 - \epsilon)e^{-i\xi\tau}$.

Similarly to the case of the WFT, the ϵ -supports for the WT are defined based on (2.17) as

$$\begin{aligned} R_\psi(\omega) &\equiv \frac{\int_0^\omega \hat{\psi}^*(\xi) \frac{d\xi}{\xi}}{\int_0^\infty \hat{\psi}^*(\xi) \frac{d\xi}{\xi}} = C_\psi^{-1} \frac{1}{2} \int_0^\omega \hat{\psi}^*(\xi) \frac{d\xi}{\xi}, \\ \xi_{1,2}(\epsilon) : |R_\psi(\xi \leq \xi_1)| &\leq \epsilon/2, \quad |1 - R_\psi(\xi \geq \xi_2)| \leq \epsilon/2, \\ P_\psi(\tau) &\equiv \frac{\int_{-\infty}^\tau \psi^*(t) e^{i\omega_\psi t} dt}{\int \psi^*(t) e^{i\omega_\psi t} dt} = \tilde{C}_\psi^{-1} \int_{-\infty}^\tau \psi^*(t) e^{i\omega_\psi t} dt, \\ \tau_{1,2}(\epsilon) : |P_\psi(\tau \leq \tau_1)| &\leq \epsilon/2, \quad |1 - P_\psi(\tau \geq \tau_2)| \leq \epsilon/2. \end{aligned} \quad (2.27)$$

Like $|P_g(\tau)|$ in (2.27), $|P_\psi(\tau)|$ ($|1 - P_\psi(\tau)|$) quantifies the relative part of $\psi(t)e^{-i\omega_\psi t}$ contained at $t < \tau$ ($t > \tau$), with $[\tau_1(\epsilon), \tau_2(\epsilon)]$ specifying the interval encompassing its $(1 - \epsilon)$ part. In the same manner, $|R_\psi(\omega)|$ and $\xi_{1,2}(\epsilon)$ are related to the relative part of $\hat{\psi}(\xi)$, taken on a logarithmic scale.

However, due to the scaling nature of the WT, the relationships of (2.27) to real quantities differ slightly from the case of the WFT. Thus, for the single tone $s(t) = \cos(\nu t) \Rightarrow W_s(\omega, t) = \hat{\psi}^*(\omega_\psi \nu / \omega) e^{i\nu t}/2$ the WT at frequencies $\omega < \xi$ ($\omega > \xi$) will contain the $|1 - R_\psi(\omega_\psi \nu / \xi)|$ ($|R_\psi(\omega_\psi \nu / \xi)|$) part of the signal, while its $(1 - \epsilon)$ part will lie in the band $[\omega_\psi \nu / \xi_2(\epsilon), \omega_\psi \nu / \xi_1(\epsilon)]$.

For the delta-function $s(t) = \delta(t - \tau) \Rightarrow W_s(\omega, t) \cong \frac{\omega}{\omega_\psi} \psi^*\left(\frac{\omega(\tau - t)}{\omega_\psi}\right)$ the WT spread in time will vary for different ω . At each frequency the part of the delta-function's total FT contained in the WT at $t < t_0$ ($t > t_0$) will be $\cong |1 - P_\psi(\omega(\tau - t_0)/\omega_\psi)|$ ($\cong |P_\psi(\omega(\tau - t_0)/\omega_\psi)|$), while its $(1 - \epsilon)$ part will reside in the interval $\cong [\tau - \omega_\psi \tau_2(\epsilon)/\omega, \tau - \omega_\psi \tau_1(\epsilon)/\omega]$.

The quantities (2.26), (2.27) are very convenient and will be used extensively below, not only in the present section. For simplicity, $\tau_{1,2}(\epsilon)$ and $\xi_{1,2}(\epsilon)$ denote the respective ϵ -supports both for the window function in the WFT and for the wavelet

function in the WT. The meaning will always be clear from the context. Note that the full supports of the window/wavelet in time ($g(t)$, $\psi(t)$) and frequency ($\hat{g}(\xi)$, $\hat{\psi}(\xi)$), whether finite or not, are $[\tau_1(0), \tau_2(0)]$ and $[\xi_1(0), \xi_2(0)]$, respectively.

2.3.4 Reconsidered Definitions

A more universal and appropriate approach (than the traditional variance-based one) is to regard two components as being reliably resolved if they can each be accurately identified and reconstructed from the signal's TFR (i.e. can be recovered with a relative error not exceeding some threshold). Consider the WFT of the two-tone signal (2.20), from which one wants to find the individual analytic signals $x_{\nu\nu;1}^a(t) = e^{i\nu t}$, $x_{\nu\nu;2}^a(t) = e^{i(\nu+\Delta\nu)t} e^{i\Delta\varphi}$ for each of the two tones. At any time t , this can be done by first dividing the frequency range at some $\omega = \omega_x(t)$ into two parts, each responsible for a separate tone, and then integrating the WFT over the corresponding frequency ranges in the same way as in (2.11). This will give the reconstructed analytic signals $\tilde{x}_{\nu\nu;1,2}^a(t)$ which, using (2.20) and (2.26), can be represented as

$$\begin{aligned}\tilde{x}_{\nu\nu;1}^a(t) &= \int_{-\infty}^{\omega_x(t)} G_s(\omega, t) d\omega \\ &= \frac{C_g^{-1}}{2} e^{i\nu t} \left[\int_{-\infty}^{\omega_x(t)} \hat{g}(\omega - \nu) d\omega + e^{i(\Delta\nu t + \Delta\varphi)} \int_{-\infty}^{\omega_x(t)} \hat{g}(\omega - \nu - \Delta\nu) d\omega \right] \\ &= e^{i\nu t} \left[(1 - R_g(\nu - \omega_x(t))) + R_g(\omega_x(t) - \nu - \Delta\nu) e^{i(\Delta\nu t + \Delta\varphi)} \right], \\ \tilde{x}_{\nu\nu;2}^a(t) &= \int_{\omega_x(t)}^{\infty} G_s(\omega, t) d\omega \\ &= \frac{C_g^{-1}}{2} e^{i\nu t} \left[\int_{\omega_x(t)}^{\infty} \hat{g}(\omega - \nu) d\omega + e^{i(\Delta\nu t + \Delta\varphi)} \int_{\omega_x(t)}^{\infty} \hat{g}(\omega - \nu - \Delta\nu) d\omega \right] \\ &= e^{i\nu t} \left[R_g(\nu - \omega_x(t)) + (1 - R_g(\omega_x(t) - \nu - \Delta\nu)) e^{i(\Delta\nu t + \Delta\varphi)} \right],\end{aligned}\tag{2.28}$$

where $R_g(x) = 1 - R_g(-x)$ is defined in (2.26).

Obviously, the reconstruction errors $x_{\nu\nu;1,2}^a(t) - \tilde{x}_{\nu\nu;1,2}^a(t)$ generally depend on the phase-shift $\Delta\varphi$. Therefore, in the corresponding expressions one should take the average over $\Delta\varphi$, which will be denoted as $\langle \dots \rangle_{\Delta\varphi}$. The relative errors of each tone's reconstruction $\varepsilon_{\nu\nu;1,2}(\nu, t, \Delta\nu)$ then become

$$\begin{aligned}\varepsilon_{\nu\nu;1,2}^2(\nu, t, \Delta\nu) &\equiv \frac{\langle |x_{\nu\nu;1,2}^a(t) - \tilde{x}_{\nu\nu;1,2}^a(t)|^2 \rangle_{\Delta\varphi}}{\langle |x_{\nu\nu;1,2}^a(t)|^2 \rangle_{\Delta\varphi}} \\ &= |R_g(\nu - \omega_x(t))|^2 + |R_g(\omega_x(t) - \nu - \Delta\nu)|^2.\end{aligned}\tag{2.29}$$

Note that, in the present case, averaging over $\Delta\varphi$ and time-averaging will give the same results; however, in general the TFR resolution properties can depend on time, and taking the mean over phase-shifts allows one to localize these errors at each t .

The minimum resolvable frequency difference $\Delta\nu_{\min}(\nu, t)$ can be defined as the minimum $\Delta\nu$ in (2.29) for which the total error $\varepsilon_{\nu\nu}(\nu, t, \Delta\nu)$ is still smaller than some accuracy threshold ϵ_r :

$$\Delta\nu \geq \Delta\nu_{\min}(\nu, t) : \varepsilon_{\nu\nu}(\nu, t, \Delta\nu) \equiv [\varepsilon_{\nu\nu;1}^2(\nu, t, \Delta\nu) + \varepsilon_{\nu\nu;2}^2(\nu, t, \Delta\nu)]^{1/2} \leq \epsilon_r. \quad (2.30)$$

It can be expressed through the ϵ_r -support of the window in frequency (2.26). Thus, consider the WFT with real, positive and symmetric $\hat{g}(\omega)$, e.g. a Gaussian (2.12). Then it follows from (2.20) that the minimum WFT amplitude between the peaks corresponding to two tones will always appear at $\omega = \nu + \Delta\nu/2$ (unless these two peaks are merged into a single one at some times, which might happen if $\Delta\nu$ is too small). Therefore, in practice the respective frequency regions of the tones will be separated exactly at their average frequency (see Sect. 3.2.2 below), so that one should use $\omega_x(t) = \nu + \Delta\nu/2$ when estimating the errors (2.29); it can also be shown that, in the present case, these errors are minimized by such a choice of $\omega_x(t)$. The overall reconstruction error is then

$$\varepsilon_{\nu\nu}(\Delta\nu) = [2|R_g(-\Delta\nu/2)|^2 + 2|R_g(-\Delta\nu/2)|^2]^{1/2} = 2|R_g(-\Delta\nu/2)|, \quad (2.31)$$

and, taking into account that $\xi_1(\epsilon) = -\xi_2(\epsilon)$ due to the assumed window symmetry, it follows from (2.30) and (2.26) that the frequency difference for which two tones are recovered with inaccuracy ϵ is exactly equal to the ϵ -support of $\hat{g}(\xi)$.

For other window forms (e.g. asymmetric $\hat{g}(\xi)$) the situation becomes more complicated, but one can still expect to get an overall error of around ϵ when $\Delta\nu = \xi_2(\epsilon) - \xi_1(\epsilon)$. Note that the above considerations hold for reasonably small ϵ , so that $\Delta\nu$ is high enough and there are always two distinct peaks in the WFT amplitude; otherwise, if the peaks are merged at certain times, the actual reconstruction errors will be larger than (2.29).

The case of two delta-peaks (2.21) is closely similar to that of two tones, so the same considerations apply, with just $\xi_{1,2}(\epsilon) \rightarrow \tau_{1,2}(\epsilon)$. Hence, setting ϵ_r as the maximum allowable reconstruction error for which two tones/delta-peaks can still be regarded as resolved, the minimum resolvable time-delay $\Delta\tau_{\min}$ and frequency difference $\Delta\nu_{\min}$, and the other resolution parameters based on them, for the WFT take the forms

$$\begin{aligned} \Delta\nu_{\min} &= \xi_2(\epsilon_r) - \xi_1(\epsilon_r), \quad \Delta\tau_{\min} \cong \tau_2(\epsilon_r) - \tau_1(\epsilon_r), \\ \gamma_\omega &\equiv \Delta\nu_{\min}^{-1}, \quad \gamma_t \equiv \Delta\tau_{\min}^{-1}, \quad \gamma_{\omega t} \equiv \gamma_\omega \gamma_t = [\Delta\nu_{\min} \Delta\tau_{\min}]^{-1}, \end{aligned} \quad (2.32)$$

where $\tau_{1,2}(\epsilon)$ and $\xi_{1,2}(\epsilon)$ are defined in (2.26). Generally, accurate reconstruction might reasonably be assumed as being at 95 % precision, so one can set $\epsilon_r = 0.05$ in (2.32). The resolution characteristics of different windows are listed in Sect. 7.2.

The same approach straightforwardly extends to the WT case, where one applies similar considerations in terms of (2.27). Thus, it can be shown that two tones with frequency ratio $\frac{\nu+\Delta\nu}{\nu} = 1 + \Delta\nu/\nu = \frac{\xi_2(\epsilon)}{\xi_1(\epsilon)}$ are reconstructed from the WT with an overall relative error of around ϵ . This estimate is exact if $\hat{\psi}(\xi)$ is real, positive and symmetric on a logarithmic scale (such as the lognormal wavelet (2.19)), in which case the tones will always be separated at $\omega_x(t) = \exp[\log \nu + \log(\nu + \Delta\nu)] = \sqrt{\nu(\nu + \Delta\nu)}$. For the resolution of two delta-peaks, it follows from (2.21) that one should consider the ϵ -supports corresponding to $\psi^*(\omega t/\omega_\psi)$, so that the related error will be different at each frequency ω , characterized by ϵ calculated from $\omega \Delta\tau/\omega_\psi = \tau_2(\epsilon) - \tau_1(\epsilon)$. Hence, the resolution parameters (2.22), (2.23) for the WT are

$$\begin{aligned} \Delta\nu_{\min}(\nu) &= \nu \left(\frac{\xi_2(\epsilon_r)}{\xi_1(\epsilon_r)} - 1 \right), \quad \Delta\tau_{\min}(\nu) \cong \frac{\omega_\psi}{\nu} (\tau_2(\epsilon_r) - \tau_1(\epsilon_r)), \\ \gamma_\omega(\nu) &\equiv [\Delta\nu_{\min}(\nu)]^{-1}, \quad \gamma_t(\nu) = [\Delta\tau_{\min}(\nu)]^{-1}, \\ \gamma_{\omega t} &\cong \left[\omega_\psi (\tau_2(\epsilon_r) - \tau_1(\epsilon_r)) \log \frac{\xi_2(\epsilon_r)}{\xi_1(\epsilon_r)} \right]^{-1}, \end{aligned} \quad (2.33)$$

where $\tau_{1,2}(\epsilon)$ and $\xi_{1,2}(\epsilon)$ are defined in (2.27), and ϵ_r denotes the maximum allowable reconstruction error, which can be set to $\epsilon_r = 0.05$, similarly to that in (2.32). The resolution characteristics of different wavelets are listed in Sect. 7.2.

Summarizing, in contrast to the classic resolution measures (2.24), (2.25), the quantities in (2.32), (2.33) are very universal and have clear physical meaning, being related directly to the accuracy with which two time or frequency events can be recovered from the resultant TFR. In the context of the present work, where the main topic is the TFR-based decomposition of the signal (see Chaps. 3 and 4), such an approach seems to be the most relevant.

2.4 Practical Issues

In theory, one has infinite time and frequency scales, and both of these variables are continuous. In practice, however, everything is finite and discrete, which has specific consequences in terms of the resultant TFRs. In this section, the issues that arise while dealing with real signals are reviewed and studied.

2.4.1 Signal Preprocessing

To obtain a reliable TFR, an initial preprocessing of the signal should be performed. It consists of eliminating trends, followed by bandpass filtering in the frequency band of interest (for which the TFR is to be calculated). These two steps are considered below, with their effects being illustrated in Fig. 2.3.

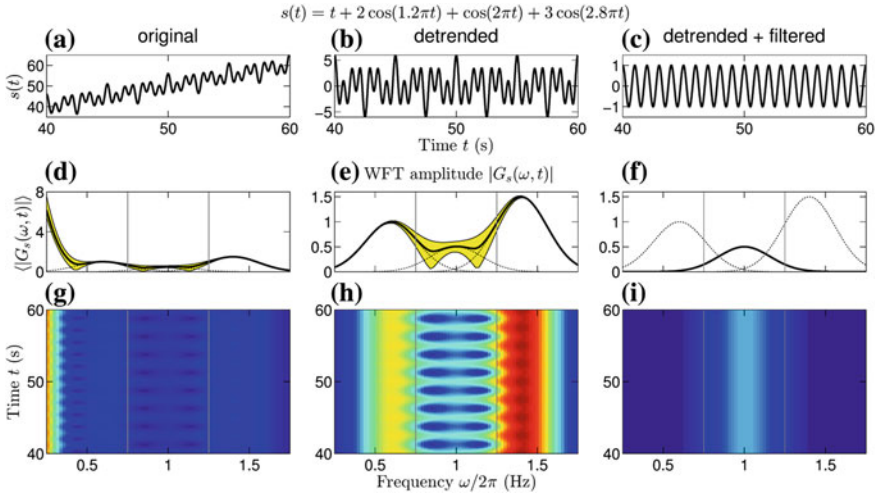


Fig. 2.3 Preprocessing and its effects on the resultant WFT (the case of the WT is qualitatively similar). **a** The original signal, consisting of the trend and three tones (shown at the top). **b** The same signal after detrending. **c** The same signal after detrending plus filtering in the band $[0.75, 1.25]$ Hz. **d, e, f** The corresponding mean WFT amplitudes, with their 95 % ranges being indicated by yellow regions; dotted lines show the mean amplitudes for the WFTs of each tone (out of 3) separately. **g, h, i** The corresponding WFT amplitudes in the time-frequency plane, with the respective color ranges being the same as y-axis limits in **(d, e, f)**. It is assumed that one is interested in calculating the TFR for the frequency range $[0.75, 1.25]$ Hz, which is indicated by gray vertical lines in panels **(d–i)**, but for completeness the resultant transforms are shown within wider ranges. The signal was sampled at 100Hz for 100 s

2.4.1.1 Removing the Trends

If the signal contains a trend-like term (i.e. a term of the form Kt or more generally Kt^α), it can seriously corrupt the resultant TFR and complicate its interpretation. This is because trends make non-negligible contributions to the signal's spectral power in a wide frequency band. Thus, the FT of a trend existing for time T will be $\int_0^T Kte^{-i\xi t} dt = \frac{KT}{\xi}e^{-i\xi T} + \frac{K}{\xi^2}(e^{-i\xi T} - 1)$, which is proportional to the overall time of the trend and decays slowly with ξ . In general, any order contributions Kt^n , $n \in \mathbb{N}$ will have $\text{FT} \sim iT^n e^{-i\xi T} / \xi$ in the first order over ξ^{-1} . Therefore, it is clear from the frequency domain form of (2.8) and (2.13) that trends might seriously affect the representation of the other components in the TFR, as illustrated in Fig. 2.3d, g.

To avoid this, one should remove trends before doing any time-frequency analysis. To do so, one can subtract a simple linear fit of the data, which will eliminate the $\sim \xi^{-1}$ spectral contribution of any term $\sim t^n$, $n \geq 1$, changing it to $\sim \xi^{-2}$ for $n > 1$. It might be better, however, to subtract a third order polynomial fit, which will in addition fully remove the trends $\sim t^2$, t^3 and reduce the spectral power of the higher order terms to $\sim \xi^{-4}$; it will also eliminate to a large extent the step-increases in mean value that are sometimes present in real data. This approach is very simple

and introduces minimal undesirable distortions to the signal's spectrum, as can be seen by comparing (d, g) and (e, h) in Fig. 2.3. Third order is suggested because such a polynomial has at most 3 zero crossings and thus cannot model more than one oscillation during the whole time-series; and since oscillations having less than one cycle over the signal time length cannot in principle be reliably studied in the TFR (see Sect. 2.4.4 below), one does not lose anything by filtering them out.

Remark 2.4.1 Sometimes detrending is performed by subtracting a moving average from the signal, which additionally filters out the low-frequency spectral content. However, due both to the properties of such a filter and its associated boundary effects, this procedure usually introduces more unwanted spectral distortions than subtraction of a simple polynomial fit followed by bandpass filtering (see below). In general, however, there exist many different approaches for trend removal [11], with the preferred choice being application dependent.

2.4.1.2 Bandpass Filtering

Using time-frequency analysis for a given signal, one is usually interested only in a particular frequency range (e.g. containing a chosen AM/FM component). At the same time, due to the peak broadening resulting from the TFR's finite frequency resolution, the frequency content of the signal *outside* the given range can significantly influence the TFR *inside* that range. This effect can be especially prominent when the spread of the window/wavelet function in the frequency domain is large. Therefore, one should always filter the signal in the frequency range of interest $[\omega_{\min}, \omega_{\max}]$ before or during the application of the TFR, i.e. set $\hat{s}(\xi < \omega_{\min}) = \hat{s}(\xi > \omega_{\max}) = 0$ in (2.8) and (2.13) (or use some other filter to remove spectral content lying outside the considered frequency range).

This issue is illustrated in Fig. 2.3e, h, where the WFT of the tone in a frequency range of interest (indicated by gray vertical lines) is seriously corrupted by the other two tones that lie near in frequency. As seen from Fig. 2.3f, i, filtration of the signal only within the band considered solves this problem, allowing for an accurate representation of the corresponding component. Note that, performed alone, band-pass filtering does not fully remove the influence of a trend in the frequency range under consideration. That is why trends should be removed first, as described in the previous subsection.

Remark 2.4.2 It should be noted, that even if the instantaneous frequency of the AM/FM component lies within the considered frequency range $[\omega_{\min}, \omega_{\max}]$, some of the related tones responsible for its amplitude/frequency modulation (see Eq. 2.6 and its discussion) might lie outside this range. In this case they will be filtered out, which will spoil the representation of the corresponding component to some extent. However, the probability that the related spectral content lies outside the considered band is not higher than the probability that there exists some unrelated components which, if unfiltered, might affect the components of interest to the same or a greater

extent. Additionally, real signals are usually corrupted by noise, which can be viewed as many undesirable tones. Therefore, despite the possible drawbacks, it is generally preferable to filter than not to do so. Note also that there are many other filters [39, 53] that can be used instead of a simple bandpass filtering.

2.4.2 Frequency Discretization

In theory one has a continuous frequency variable ω , for which TFRs are calculated. In practice, however, the frequency axis is discretized, being partitioned into bins centered at the chosen discrete values $\omega = \omega_k$, and the WFT/WT is calculated only for these frequencies. As discussed previously (see Sect. 2.2), the WFT and WT have linear and logarithmic frequency resolutions, so that the discretization should also be performed on linear and logarithmic scales, respectively. It is convenient to take $\omega_k = (k - k_0)\Delta\omega$ (frequency bins $[(k - k_0 - 1/2)\Delta\omega, (k - k_0 + 1/2)\Delta\omega]$) for the WFT and $\omega_k/2\pi = 2^{(k-k_0)/n_v}$ (frequency bins $2\pi[2^{(k-k_0-1/2)/n_v}, 2^{(k-k_0+1/2)/n_v}]$) for the WT. The question is then what effects the choices of $\Delta\omega$ and n_v have on the resultant TFRs, and how to select these parameters appropriately.

Theoretically, the frequency resolution of the TFR is determined only by the window/wavelet properties (see Sect. 2.3). However, due to frequency discretization there also appears numerical frequency resolution, determined by the widths of the frequency bins, i.e. the choice of $\Delta\omega$ for the WFT or the number-of-voices n_v for the WT. It imposes an upper bound on the effective frequency resolution of the transform, which is equal to the minimum among the theoretical and numerical resolutions. Thus, if $\Delta\omega$ or $1/n_v$ is chosen to be too large, a few peaks in the TFR amplitude might be merged into one frequency bin, leading to an inability to distinguish between them and lowering the effective frequency resolution, as illustrated in Fig. 2.4. On the other hand, if $\Delta\omega$ or $1/n_v$ is selected to be too small, it cannot improve the frequency resolution beyond the theoretical maximum, but it will increase the computational cost due to requiring the calculation of the TFR at more frequencies. Therefore, $\Delta\omega$ and n_v should be selected so as to retain the original, theoretical frequency resolution, predicted from the chosen window/wavelet parameters. This maximizes the effective resolution and occurs when the numerical frequency resolution is the same as or better than the theoretical one. Obviously, the optimal values of $\Delta\omega$ and n_v depend on the chosen form of window/wavelet and its properties (see Fig. 2.4). Thus, while for $f_0 = 1$ a frequency step of $\Delta\omega/2\pi = 0.08$ is sufficient to distinguish between two nearby but theoretically resolved components (Fig. 2.4b), for $f_0 = 4$ it is already insufficient (Fig. 2.4c).

To avoid reducing the original frequency resolution by discretization, for the WFT the width of a frequency bin $\Delta\omega$ should be smaller than the theoretical $\Delta\nu_{\min}$ (see Sect. 2.3); while for the WT one has the same criterion but on a logarithmic frequency scale. In other words, one should break the minimum resolvable frequency difference (WFT) or ratio (WT) into a large enough number of segments. However, one is now interested in preserving the TFR's appearance more generally, i.e. not

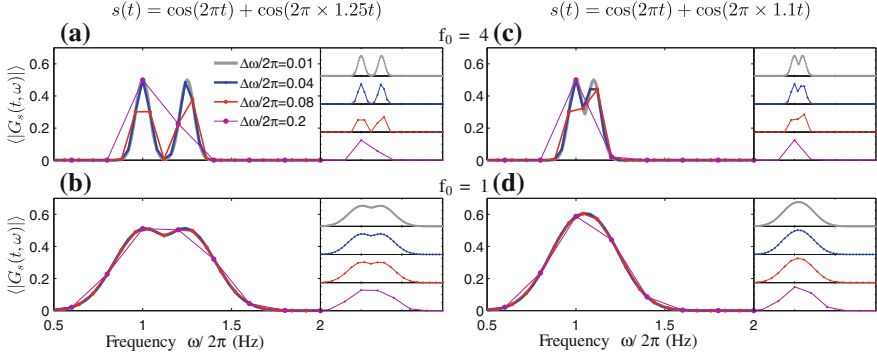


Fig. 2.4 Time-averaged WFT amplitudes for the signals $s(t) = \cos(2\pi t) + \cos(2.5\pi t)$ (a, b) and $s(t) = \cos(2\pi t) + \cos(2.2\pi t)$ (c, d), calculated using different resolution parameters f_0 (corresponding to rows) and frequency bin widths $\Delta\omega$; in each case, the *small insets* on the right show pictures for the considered $\Delta\omega$ individually. The situation for the WT is qualitatively the same

only for the reliably resolved components, but in relation to *any* components that can be distinguished in the TFR (even where they are substantially corrupted by interference). Thus, the optimal frequency bin widths should be smaller than $\Delta\nu_{\min}$ (2.32) and (2.33), determined for a relatively high error threshold ϵ_r ; the latter is therefore set to an extreme value of $\epsilon_r = 0.5$. Based on these considerations, one can select an appropriate $\Delta\omega$ or n_v as

$$\begin{aligned} \text{WFT: } \Delta\omega &= \frac{\xi_2(0.5) - \xi_1(0.5)}{N_b} \quad (\approx 1.35/f_0 N_b \text{ for Gaussian window (2.12)}), \\ \text{WT: } n_v &= \frac{N_b \log 2}{\log \xi_2(0.5) - \log \xi_1(0.5)} \quad (\approx 3.23 f_0 N_b \text{ for lognormal wavelet (2.19)}), \end{aligned} \quad (2.34)$$

where $N_b > 1$ denotes the chosen number of bins into which the 50%-support of the window/wavelet FT is divided (in the following $N_b = 10$ is used by default), and $\xi_{1,2}(\epsilon)$ are as defined in (2.26) and (2.27) for the WFT and WT, respectively.

The criterion (2.34) is based on (2.32) and (2.33), which are related to the resolution of two equal-amplitude tones. However, it is clear that, in order to be resolved, tones with different amplitudes (e.g. $s(t) = \cos \nu t + 0.1 \cos(\nu + \Delta\nu)t$) should have a bigger frequency difference than the tones of equal amplitudes. Therefore, $\Delta\omega$ and n_v are sufficient to resolve *any* two components if they are separated theoretically, as desired.

Remark 2.4.3 It should be noted, that even with the choice (2.34) it is still possible that at certain times close peaks will appear in the WFT/WT amplitude that are separated theoretically, but not practically, being merged due to insufficient numerical resolution. However, as mentioned previously, if the bin widths (2.34) are not small enough to resolve two tones, most of the time they will behave as a single component

in the TFR, and it will be impossible to investigate them separately. Hence, attributing the corresponding peaks to one or few components will not make much difference.

2.4.3 Boundary Effects and Padding

Theoretically, one integrates over an infinite time or frequency axis while computing the WFT (2.8) and WT (2.13), but in practice the signal has a finite time duration and sampling frequency. Consequently, the resultant TFR becomes ill-defined near the signal's time borders (when t is close either to 0 or to the overall time-length T). Irrespectively of how this problem is tackled, it often leads to distortions of the TFR near both signal ends—*boundary effects*.

2.4.3.1 Padding and Its Schemes

Although this is not really the case in practice (see below), suppose for now that the WFT and WT are calculated using the convolution in the time domain in (2.8) and (2.13), respectively. The signal's positive part $s^+(\tau)$ is thus multiplied by $g(\tau-t)e^{-i\omega(\tau-t)}$ (WFT) or $\psi^*(\omega(\tau-t)/\omega_\psi)$ (WT) and integrated over τ . Therefore, one needs to devise a rule by which the integration can be performed when τ lies outside the signal's time limits, i.e. when $\tau < 0$ or $\tau > T$. Generally, this can be done by continuing the signal beyond its original time interval $[0, T]$ in some way, e.g. setting $s^+(\tau \notin [0, T]) = 0$ in (2.8), (2.13). In practice, the signal is padded at both ends using a particular convention, then the convolution is calculated with this padded signal, and after this only the part lying within the original time limits is taken. An appropriate padding can be constructed by many different schemes, and the most common/useful ones are listed below, with their effect on the resultant TFR being illustrated in Fig. 2.5. Note, that padding should be performed *after* the initial signal preprocessing, discussed in Sect. 2.4.1, has been completed.

Zero padding, when one pads the signal with zeros at both ends, represents the simplest and most predictable form of padding. This scheme effectively sets $s(\tau < 0) = s(\tau > T) = 0$ in (2.8), (2.13), “cutting” the convolution when τ goes beyond the signal's time limits. Its effect is illustrated in Fig. 2.5a, e: as can be seen, this type of padding does not introduce any new behavior (in contrast to a few other types, see below), always leading to a gradual decay of the TFR amplitude towards the time ends. This allows one to derive the expressions for boundary errors that arise in the case of zero padding, which will be done in the following.

Periodic padding is constructed by periodic continuation of the signal. It leads to quite unpredictable boundary effects, with the TFR near the time borders being affected by the signal's behavior at both ends. For example, as illustrated in Fig. 2.5b, the tone occurring at the beginning (end) has its “phantom” tail near the time end (beginning) in the WFT. Consequently, this type of padding does not represent the

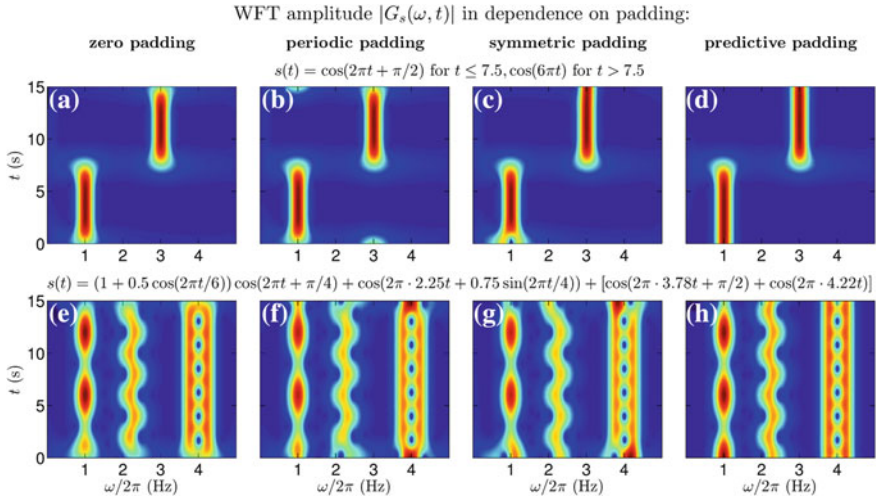


Fig. 2.5 Examples of the WFT amplitudes calculated using each of the four padding schemes discussed in the text (*columns*) for two different signals (*rows*): **a, e** zero padding; **b, f** periodic padding; **c, g** symmetric padding; **d, h** predictive padding. The effects of different padding strategies on the WT are qualitatively the same. The signals are indicated in the figure, and they were sampled at $f_s = 100$ Hz for $T = 15$ s

preferred choice, unless one knows that the signal is fully periodic over the whole interval, which almost never occurs in practice.

Symmetric padding whose effects are shown in Fig. 2.5c, g, is performed by reflecting the signal along $t = 0$ and $t = T$. It localizes boundary inaccuracies at each edge, thus solving the problem of the end-to-end influence that occurs for periodic padding. Indeed, comparing (c, g) and (b, f) of Fig. 2.5, it can be seen that the “phantom tails” near the boundaries have been removed. However, the effect of symmetric padding depends to a large extent on the phases of oscillations at $t = 0, T$, which are quite unpredictable. Thus, if the initial/end phase is not equal to zero or π , symmetric padding introduces a phase jump and leads to a splitting of the single peak in the TFR amplitude into several peaks near the time borders. As an example, in Fig. 2.5b the first oscillation has a phase shift of $\pi/2$ at $t = 0$ and thus is doubled, while the second one has zero shift at $t = T$ and thus is well represented.

Predictive padding, as implied by its name, continues the signal beyond its time limits by inferring/forecasting its past/future behavior. Thus, one tries to predict the values of the signal by assuming some model of the process generating the time-series and fitting data to this model to find its parameters. In the present case, one is mainly interested in continuing the signal in such a way as to best represent in the TFR its existing characteristics near the boundaries, and not in finding the unknown signal behavior beyond the available data, which is generally impossible. Because time-frequency analysis is devoted to studying the oscillatory properties of the data, it seems that the most relevant approach would be to forecast the signal based on

its local spectral characteristics at both ends. A scheme for doing this is discussed in Sect. 7.4, while the effect of such padding is shown in Fig. 2.5d, h. As can be seen, for both signals in Fig. 2.5 predictive padding almost completely eliminates the boundary distortions. However, it does not represent an “ultimate cure”, and for a complicated signal some errors might remain; they depend on the signal’s structure and thus are hard to estimate in general. Nevertheless, in terms of its ability to reduce boundary effects, predictive padding usually outperforms all other schemes, and it is therefore used by default in this work.

The remaining question is how many values to pad. At the beginning of this section it was assumed for simplicity that the WFT/WT are calculated using the time domain forms of (2.8), (2.13). Such an approach, however, is computationally very expensive, requiring $O(N)$ convolutions for each $t_{n=1,\dots,N}$ and thus being of $O(N^2)$ in cost. In practice, therefore, the WFT and WT are calculated using the frequency domain convolutions in (2.8) and (2.13) which, utilizing the FFT algorithm, can be performed in $O(N \log N)$ computations (see Sect. 7.5). However, the discrete FT of the signal $\hat{s}(\xi_n)$, which is used in numerical convolution, represents the periodic spectrum estimate, being an exact FT of the periodically continued signal: $\hat{s}(\xi_n) = \int s(\text{mod}_T t) e^{-i\xi_n t} dt$. Thus, the absence of padding is in practice completely equivalent to periodic padding.

One therefore needs to add enough points at both signal ends to guarantee that, within the original signal time-length, the effects of implicit periodic continuation due to FFT-based convolution are small. This minimum number of padded values will obviously depend on the spread of the window/wavelet in time: the larger it is, the more values one should pad so that most of the window/wavelet is convoluted with padded values and not with the periodically-continued other end of the signal. Since the effective window/wavelet length in time (up to a predefined tolerance ϵ) can be expressed through its ϵ -supports $\tau_{1,2}(\epsilon)$ (2.26), (2.27), the minimum number of points $n_1^{(\min)}$ ($n_2^{(\min)}$) with which one should pad the signal for $t < 0$ ($t > T$) can be determined as

$$\begin{aligned} \text{WFT: } n_{1,2}^{(\min)}(\epsilon) &= f_s |\tau_{1,2}(\epsilon)|, \\ \text{WT: } n_{1,2}^{(\min)}(\epsilon) &= \frac{\omega_\psi}{\omega_{\min}} f_s |\tau_{1,2}(\epsilon)|, \end{aligned} \tag{2.35}$$

where f_s is the sampling frequency of the signal, and $\omega_{\min}$ denotes the minimum frequency for which the TFR is calculated. Thus, for the WT the wavelet is rescaled at each frequency, so the number of points needed to assure the specified precision ϵ also scales; to guarantee the accuracy for all frequencies, the maximum number of points is taken, which corresponds to the lowest frequency.

Additionally, the FFT algorithm requires the total length of the signal in samples to be a power of two. Therefore, in practice one should simultaneously assure both the power-of-two points and criterion (2.35), which can be done by padding the signal from originally N to N_p points, with n_1 values to the left ($t < 0$) and n_2 to the right ($t > T$), given by

$$N_p = \text{NextPowerOfTwo}[N + n_1^{(\min)}(\epsilon) + n_2^{(\min)}(\epsilon)] = N + n_1 + n_2,$$

$$n_{1,2} = \frac{n_{1,2}^{(\min)}(\epsilon)}{n_1^{(\min)}(\epsilon) + n_2^{(\min)}(\epsilon)} (N_p - N), \quad (2.36)$$

where $n_{1,2}^{(\min)}(\epsilon)$ are as defined in (2.35); by default the precision $\epsilon = 0.01$ is used.

2.4.3.2 Error Estimates and the Cone-of-influence

As discussed above, for periodic, symmetric and predictive padding the boundary effects very much depend on the signal's structure. For zero padding, however, they are quite universal and can thus be estimated. A simple and straightforward way to do so is to quantify the relative boundary inaccuracies $\epsilon_b(\omega, t)$ through the difference between the theoretical (2.10), (2.16) and “practical” WFT/WT of the single-tone signal $s(t) = \cos(\nu t + \varphi)$, calculated at the tone frequency ν . Then in the case of zero padding one obtains

WFT:

$$\begin{aligned} \epsilon_b(\nu, t) &\equiv \frac{|G_s(\nu, t) - \tilde{G}_s(\nu, t)|}{|G_s(\nu, t)|} \\ &= \frac{\left| \frac{1}{2} \int e^{i(\nu\tau + \varphi)} g(\tau - t) e^{-i\nu(\tau - t)} d\tau - \frac{1}{2} \int_0^T e^{i(\nu\tau + \varphi)} g(\tau - t) e^{-i\nu(\tau - t)} d\tau \right|}{\left| \frac{1}{2} \int e^{i(\nu\tau + \varphi)} g(\tau - t) e^{-i\nu(\tau - t)} d\tau \right|} \\ &\leq \frac{\left| \int_{-\infty}^0 g(\tau - t) d\tau \right| + \left| \int_T^\infty g(\tau - t) d\tau \right|}{\left| \int g(\tau) d\tau \right|} \\ &= |P_g(-t)| + |P_g(t - T)|, \end{aligned} \quad (2.37)$$

WT:

$$\begin{aligned} \epsilon_b(\nu, t) &\equiv \frac{|W_s(\nu, t) - \tilde{W}_s(\nu, t)|}{|W_s(\nu, t)|} \\ &= \frac{\left| \frac{1}{2} \int e^{i(\nu\tau + \varphi)} \psi^*(\nu(\tau - t)/\omega_\psi) \frac{\nu d\tau}{\omega_\psi} - \frac{1}{2} \int_0^T e^{i(\nu\tau + \varphi)} \psi^*(\nu(\tau - t)/\omega_\psi) \frac{\nu d\tau}{\omega_\psi} \right|}{\left| \frac{1}{2} \int e^{i(\nu\tau + \varphi)} \psi^*(\nu(\tau - t)/\omega_\psi) \frac{\nu d\tau}{\omega_\psi} \right|} \\ &\leq \frac{\left| \int_{-\infty}^0 \psi^*(u - \nu t/\omega_\psi) e^{i\omega_\psi u} du \right| + \left| \int_T^\infty \psi^*(u - \nu t/\omega_\psi) e^{i\omega_\psi u} du \right|}{\left| \int \psi^*(u - \nu t/\omega_\psi) e^{i\omega_\psi u} du \right|} \\ &= |P_\psi(-\nu t/\omega_\psi)| + |P_\psi(\nu(t - T)/\omega_\psi)|, \end{aligned} \quad (2.38)$$

where $\tilde{G}_s(\omega, t)$ and $\tilde{W}_s(\omega, t)$ denote respectively the WFT and WT obtained in practice, while $P_{g,\psi}(\tau)$ are as defined in (2.26), (2.27). Note that $\epsilon_b(\omega, t)$ (2.37),

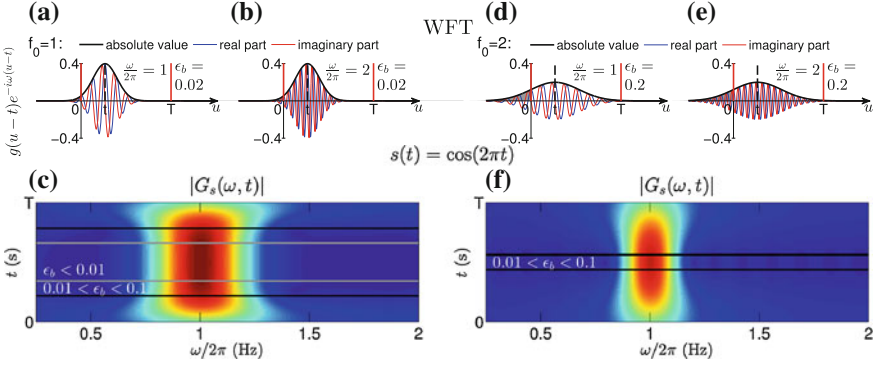


Fig. 2.6 **a, b, d, e** Examples of the time domain Gaussian window functions (2.12) (multiplied by the corresponding oscillatory part), with which the signal is convolved while constructing the WFT (2.8), for different resolution parameters f_0 and frequencies ω ; *thick vertical lines* indicate the time limits of the signal (*solid red*) and the time t at which the window is centered (*dashed black*); the area under the window modulus outside the time limits is *filled with gray* (note that, for a Gaussian window (2.12), $|g(t)| = g(t)$). **c, f** WFTs for the single tone signal $s(t) = \cos(2\pi t)$ in the same time limits $[0, T]$ as in **(a, b, d, e)**, with *black and gray lines* enclosing the time-frequency regions where boundary errors (2.37) are small: $\epsilon_b < 0.1$ and $\epsilon_b < 0.01$, respectively. The signal was sampled at $f_s = 100$ Hz for $T = 7.5$ s, and in **(c, f)** padding with zeros was used

(2.38) does not depend on the phase-shift φ of the tone, as would be the case for the symmetric and periodic padding schemes.

It is clear that the boundary distortions in the TFR manifest themselves when a non-negligible portion of the window $g(\tau - t)e^{-i\omega(\tau - t)}$ or wavelet $\psi^*(\omega(\tau - t)/\omega_\psi)$, by which the signal $s^+(\tau)$ is multiplied and integrated in (2.8), (2.13), lies outside the time limits. This is true for any padding but, in the case of padding with zeros, the related errors take the simple forms (2.37), (2.38). From the latter it can be seen, that the dependence of the boundary inaccuracies on time and frequency is different for the WFT and WT, as illustrated in Figs. 2.6 and 2.7.

In the case of the WFT (Fig. 2.6), the boundary inaccuracies are independent of the frequency ω and depend only on time t : $\epsilon_b(\omega, t) = \epsilon_b(t)$. This is because ω controls the effective number of oscillations within the window, but not its spread in time, which is determined by the resolution parameter f_0 . Hence, for each frequency the relative part of $g(\tau - t)$ lying outside the time limits (gray-shaded areas in Fig. 2.6a, b, d, e) is the same, so that for some (central) time range the WFT is well-behaved at all frequencies, as shown in Fig. 2.6c, f.

For the WT (Fig. 2.7), on the other hand, the wavelet is rescaled at each frequency ω , and so is the part of the wavelet outside the time limits (gray-shaded areas in Fig. 2.7a, b, d, e). As a result, the boundary errors depend on both frequency and time, and the region where the WT coefficients are determined with some predefined accuracy in terms of boundary effects takes the form of a cone, as shown in Fig. 2.7c, f.

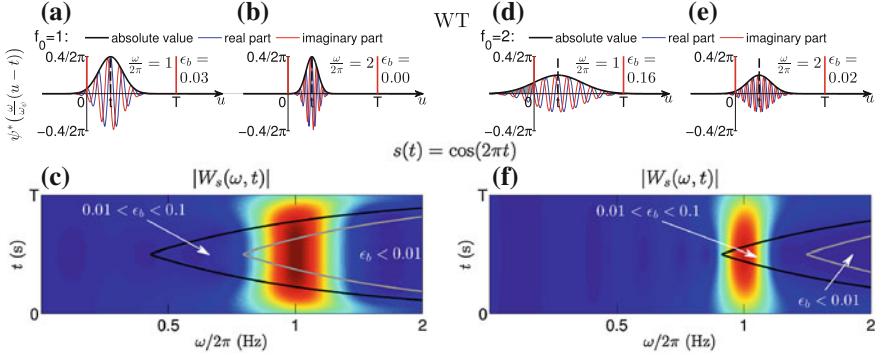


Fig. 2.7 Same as Fig. 2.6, but for the WT (2.13) with lognormal wavelet (2.19). **a, b, d, e** the wavelet function with which the signal is convoluted while constructing the WT (2.13); the area under the wavelet modulus outside the time limits is *filled with gray* (note that $|\psi(t)| \approx \psi(t)e^{-i\omega_\psi t}$ in the present case). **c, f** WTs for the single tone signal $s(t) = \cos(2\pi t)$ within the same time limits $[0, T]$ as in **(a, b, d, e)**. The signal was sampled at $f_s = 100$ Hz for $T = 7.5$ s, and in **(c, f)** padding with zeros was used

The time-frequency region where the boundary errors are small in the current TFR is called the *cone-of-influence* (although this region has a truly conic form only for the WT):

$$\textbf{Cone-of-influence } \{\omega, t\}_{coi}^{(\epsilon)} : (\omega, t) \in \{\omega, t\}_{coi}^{(\epsilon)} \Leftrightarrow \epsilon_b(\omega, t) \leq \epsilon, \quad (2.39)$$

where ϵ is the chosen accuracy threshold. Examples of the cones-of-influence for $\epsilon = 0.1$ and $\epsilon = 0.01$ have already been presented in Figs. 2.6c, f and 2.7c, f, where they are enclosed by black and gray lines, respectively. Based on the error estimates (2.37), (2.38), the cones-of-influence for the WFT and WT can be expressed through the corresponding window/wavelet ϵ -supports in time (2.26), (2.27) as

$$\begin{aligned} \textbf{WFT: } \{\omega, t\}_{coi}^{(\epsilon)} &= \left\{ [\omega_{\min}, \omega_{\max}], [\delta t_1(\epsilon), T - \delta t_2(\epsilon)] \right\}, \\ \delta t_1(\epsilon) &\leq -\tau_1(\epsilon), \quad \delta t_2(\epsilon) \leq \tau_2(\epsilon), \\ \textbf{WT: } \{\omega, t\}_{coi}^{(\epsilon)} &= \left\{ [\omega_{\min}, \omega_{\max}], [\delta t_1(\epsilon, \omega), T - \delta t_2(\epsilon, \omega)] \right\}, \\ \delta t_1(\epsilon, \omega) &\leq -\omega_\psi \tau_1(\epsilon) / \omega, \quad \delta t_2(\epsilon, \omega) \leq \omega_\psi \tau_2(\epsilon) / \omega, \end{aligned} \quad (2.40)$$

where $[\omega_{\min}, \omega_{\max}]$ is the frequency range in which the WFT/WT is calculated (the restrictions on this range and related issues are discussed in Sect. 2.4.4 below).

Because boundary effects can greatly influence the TFR behavior, especially for the WT, it is recommended that all characteristics (e.g. the WFT/WT mean amplitudes) be calculated using only TFR coefficients inside the cone-of-influence. The most appropriate padding scheme in this case is zero padding, as the corresponding boundary effects have a universal and well-defined form. In fact, the boundary errors

(2.37), (2.38) and the cone-of-influence (2.40) were rigorously estimated exclusively for this type of padding.

However, if one wants to extract some AM/FM component from the signal's TFR (the ways of doing this will be discussed in Chap. 3), then it should obviously be extracted for all time, and the consideration cannot be restricted to the cone-of-influence only. Thus, if all the TFR coefficients are to be used, then zero padding does not represent a good choice, because the WFT/WT near the boundaries will *surely* contain considerable errors. In this case predictive padding would be the most suitable, as it usually has the best performance in terms of reducing boundary effects.

2.4.4 TFR Frequency Range

For completeness, the restrictions on the frequency range $[\omega_{\min}, \omega_{\max}]$ over which to calculate the TFR, i.e. how high/low in frequency one can in principle go, should also be discussed. Consider a signal $s(t)$ sampled at f_s Hz for T seconds. Then one has

$$\omega_{\min}/2\pi \geq 1/T, \quad \omega_{\max}/2\pi \leq f_s/2. \quad (2.41)$$

The restriction on $\omega_{\max}$ follows from the Nyquist theorem, which states that oscillations with frequencies higher than half of the sampling frequency cannot be represented in a discrete time signal. The constraints on $\omega_{\min}$, on the other hand, are based on simple logic: it is clear that for a particular oscillation to be reliably studied (generally by any method), there should be at least one of its cycles within the signal.

However, from the statistical viewpoint, to reach any meaningful conclusions about the characteristics of the oscillatory process, such as its typical amplitude (as calculated e.g. from the time-averaged TFR amplitude), there should be at least 5-6 corresponding cycles within the signal [31], so that

$$\omega_{\min}^{(\text{stat})}/2\pi = 5/T. \quad (2.42)$$

Thus, although one can in principle estimate the properties of the components at lower frequencies, the resultant estimates might be highly untypical. Indeed, the number of cycles can be associated with a number of trials. For example, when testing the effects of some drug, one cannot base conclusions on only one subject (since the probability that the subject tested is an outlier is quite high); but if the same effects appear in 5-6 subjects, this suggests that they are quite common. Note also that, for statistical comparison of oscillatory properties between different data (e.g. as is done in [25, 45]), all these properties should be statistically meaningful in themselves, so that the related oscillations have frequencies $\geq \omega_{\min}^{(\text{stat})}$.

Finally, since the cone-of-influence for the WT contracts towards the lower frequencies (see Sect. 2.4.3), there exists also the minimal frequency $\omega_{\min}^{(\epsilon)}$ from the viewpoint of precision (while there is no such restriction in the case of the WFT).

The latter can be defined as the minimum frequency for which at least one WT coefficient is determined with accuracy ϵ in terms of the boundary effects. Using the estimates (2.40), one obtains

$$\text{WFT: } \omega_{\min}^{(\epsilon)} = -\infty; \quad \text{WT: } \omega_{\min}^{(\epsilon)} \leq \omega_{\psi}(\tau_2(\epsilon) - \tau_1(\epsilon))/T. \quad (2.43)$$

Summarizing, it is quite safe to calculate and analyse the TFR in the range $[\omega_{\min}, \omega_{\max}]$ with $\omega_{\min} > \max[2\pi/T, \omega_{\min}^{(\epsilon)}]$ and $\omega_{\max} < 2\pi f_s/2$. However, if one needs to draw conclusions about the typical oscillatory parameters of the data, they should be based on the TFR within the cone-of-influence (see Sect. 2.4.3) and on at least 5 cycles of the related oscillations, so that $\omega_{\min} \geq \max[2\pi \times 5/T, \omega_{\min}^{(\epsilon)}]$.

Remark 2.4.4 Even if one is interested in the frequency range $[\omega_{\min}, \omega_{\max}]$ (so that the signal is filtered in this range on the stage of preprocessing, see Sect. 2.4.1), it might be useful to calculate the TFR for a larger interval (up to $(-\infty, \infty)$ for the WFT and $(0, \infty)$ for the WT). This allows to trace the “tails” of the TFR amplitude peaks located in the original range, which might be needed e.g. for a reliable parameter reconstruction by the direct method (see Sect. 3.2.2). For example, in Fig. 2.3f, the considered frequency band—between the gray vertical lines—contains only the main part of the corresponding peak, with its tails occupying a wider range. If one wants to encompass the supports of all possible peaks contained in $[\omega_{\min}, \omega_{\max}]$ then, while bandpass filtering the signal in the original frequency band, its WFT/WT should be calculated within a slightly larger region $[\tilde{\omega}_{\min}, \tilde{\omega}_{\max}]$. To achieve relative accuracy ϵ , so that the new range contains the $(1 - \epsilon)$ part of the total area under any peak in the range of interest, one should use

$$\begin{aligned} \text{WFT: } [\tilde{\omega}_{\min}, \tilde{\omega}_{\max}] &= [\omega_{\min} + \xi_1(\epsilon), \omega_{\max} + \xi_2(\epsilon)], \\ \text{WT: } [\tilde{\omega}_{\min}, \tilde{\omega}_{\max}] &= \left[\omega_{\min} \frac{\xi_1(\epsilon)}{\xi_2(\epsilon)}, \omega_{\max} \frac{\xi_2(\epsilon)}{\xi_1(\epsilon)} \right], \end{aligned} \quad (2.44)$$

where the ϵ -supports $\xi_{1,2}(\epsilon)$ for the window and wavelet are defined in (2.26) and (2.27), respectively. Usually, however, this complication is not needed, unless the instantaneous frequency of some component of interest lies near $\omega_{\min}$ or $\omega_{\max}$.

2.5 Summary

This Chapter reviewed the main properties of the WFT and WT and discussed various issues related to their practical use, setting the foundations for the more advanced techniques discussed in the following Chapters. To summarize, the WFT and WT both represent particular projections of the signal onto the time-frequency plane, carried out as discussed in Sect. 2.2. They are mainly devoted to analysing signals of the form (2.5), i.e. those composed of a number of oscillatory components each of

which satisfies the analytic approximation (2.4) (so that the corresponding amplitudes and phases are well-defined).

The WFT and WT, and generally any TFR, have finite time and frequency resolutions. That is, they can represent reliably only signals whose components do not vary too fast in time and are not too close to each other in frequency. By adjusting window/wavelet parameters (such as the resolution parameter f_0 in the case of Gaussian window and lognormal wavelet), one can vary resolution properties of the resultant transform. However, high time and high frequency resolutions are mutually exclusive, and increasing one of them will inevitably decrease the other.

The main difference between the WFT and WT lies in the type of resolution that they have: while the former resolves components based on absolute differences between their frequencies, the latter takes into account the corresponding ratios (or the difference between logarithms); in effect, time resolution of the WFT remains constant at all frequencies, while that of the WT increases with frequency. The resolution properties of both transforms, and the approaches that can be used for their quantification, were thoroughly considered in Sect. 2.3. In particular, it was shown that classical resolution measures appear to be highly non-universal and can give misleading results in certain cases (especially for the WT), and a more appropriate measures were introduced.

There are also many practical issues related to the application of the WFT and WT, which were discussed in detail in Sect. 2.4. Thus, before calculating the WFT or WT of the signal, it should first be detrended and filtered in the frequency range of interest (see Sect. 2.4.1). Next, one should also choose appropriate steps of frequency discretization, and the formula (2.34) was derived for this purpose. Finally, both TFRs suffer from the distortions near time boundaries; the expressions to estimate the related errors and the padding scheme which can be used to significantly reduce them were devised in Sect. 2.4.3. Section 7.5 provides complete algorithms for calculating the WFT and WT, while the commonest windows/wavelets and their properties are listed in Sect. 7.2.

References

1. P.S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance* (IOP Publishing, Bristol, 2002)
2. L. Aguiar-Conraria, P.C. Magalhães, M.J. Soares, Cycles in politics: wavelet analysis of political time series. *Am. J. Pol. Sci.* **56**(2), 500–518 (2012)
3. A. Bandrivskyy, A. Bernjak, P.V.E. McClintock, A. Stefanovska, Wavelet phase coherence analysis: application to skin temperature and blood flow. *Cardiovasc. Eng.* **4**(1), 89–93 (2004)
4. E. Bedrosian, A product theorem for Hilbert transforms. *Proc. IEEE* **51**(5), 868–869 (1963)
5. D.S. Bloomfield, R.T.J. McAtter, B.W. Lites, P.G. Judge, M. Mathioudakis, F.P. Keenan, Wavelet phase coherence analysis: application to a quiet-sun magnetic element. *Astrophys. J.* **617**(1), 623 (2004)
6. B. Boashash, Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals. *Proc. IEEE* **80**(4), 520–538 (1992)

7. B. Boashash, Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications. *Proc. IEEE* **80**(4), 540–568 (1992)
8. B. Boashash, *Time Frequency Signal Analysis and Processing* (Elsevier, New York, 2003)
9. R.A. Carmona, W.L. Hwang, B. Torresani, Characterization of signals by the ridges of their wavelet transforms. *IEEE Trans. Signal Proc.* **45**(10), 2586–2590 (1997)
10. R.A. Carmona, W.L. Hwang, B. Torresani, Multiridge detection and time-frequency reconstruction. *IEEE Trans. Signal Proc.* **47**(2), 480–492 (1999)
11. R. Chandler, M. Scott, *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences* (Wiley, New York, 2011)
12. H.-I. Choi, W.J. Williams, Improved time-frequency representation of multicomponent signals using exponential kernels. *IEEE Trans. Acoust. Speech Signal Process.* **37**(6), 862–871 (1989)
13. L. Cohen, Time-frequency distributions: a review. *Proc. IEEE* **77**(7), 941–981 (1989)
14. L. Cohen, *Time-Frequency Analysis*, vol. 778 (Prentice Hall PTR, New Jersey, 1995)
15. I. Daubechies, The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. Theory* **36**(5), 961–1005 (1990)
16. I. Daubechies, *Ten Lectures on Wavelets*, volume 61 of CBMS-NSF Regional Conference Series in Application Mathematics, SIAM, 1992
17. I. Daubechies, J. Lu, H. Wu, Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.* **30**(2), 243–261 (2011)
18. N. Delprat, B. Escudie, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, B. Torr  sani, Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies. *IEEE Trans. Inf. Theory* **38**(2), 644–664 (1992)
19. D. Gabor, Theory of communication. *J. IEEE* **93**, 429–457 (1946)
20. A. Grinsted, J.C. Moore, S. Jevrejeva, Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Proc. Geophys.* **11**(5/6), 561–566 (2004)
21. K. Gr  chenig, *Foundations of Time-Frequency Analysis* (Springer, Berlin, 2001)
22. F. Hlawatsch, F. Auger, *Time-Frequency Analysis*, vol. 36 (Wiley, New York, 2010)
23. F. Hlawatsch, G.F. Boudreaux-Bartels, Linear and quadratic time-frequency signal representations. *IEEE Signal Proc. Mag.* **9**(2), 21–67 (1992)
24. N.E. Huang, Z. Wu, S.R. Long, K.C. Arnold, X. Chen, K. Blank, On instantaneous frequency. *Adv. Adapt. Data Anal.* **1**(02), 177–229 (2009)
25. D. Iatsenko, A. Bernjak, T. Stankovski, Y. Shiogai, P.J. Owen-Lynch, P.B.M. Clarkson, P.V.E. McClintock, A. Stefanovska, Evolution of cardio-respiratory interactions with age. *Phil. Trans. R. Soc. Lond. A* **371**(1997), 20110622 (2013)
26. D. Iatsenko, P.V.E. McClintock, A. Stefanovska, On the extraction of instantaneous frequencies from ridges in time-frequency representations of signals [preprint—[arXiv:1310.7276](https://arxiv.org/abs/1310.7276)], 2014
27. D. Iatsenko, P.V.E. McClintock, A. Stefanovska, Linear and synchrosqueezed time-frequency representations revisited: Overview, standards of use, reconstruction, resolution, concentration, and algorithms. *Dig. Signal Proc.* (in press), 2015, doi:[10.1016/j.dsp.2015.03.004](https://doi.org/10.1016/j.dsp.2015.03.004)
28. J. Jam  sek, M. Palu  s, A. Stefanovska, Detecting couplings between interacting oscillators with time-varying basic frequencies: Instantaneous wavelet bispectrum and information theoretic approach. *Phys. Rev. E* **81**(3), 036207 (2010)
29. J. Jam  sek, A. Stefanovska, P.V.E. McClintock, Wavelet bispectral analysis for the study of interactions among oscillators whose basic frequencies are significantly time variable. *Phys. Rev. E* **76**(4), 046221 (2007)
30. G. Kaiser, *A Friendly Guide to Wavelets* (Birkh  user, Boston, 1994)
31. L. Keselbrener, S. Akselrod, Selective discrete Fourier transform algorithm for time-frequency analysis: method and application on simulated and cardiovascular signals. *IEEE Trans. Biomed. Eng.* **43**(8), 789–802 (1996)
32. P. Kumar, E. Foufoula-Georgiou, Wavelet analysis for geophysical applications. *Rev. Geophys.* **35**(4), 385–412 (1997)
33. D. Labat, Recent advances in wavelet analyses: part 1. A review of concepts. *J. Hydrol.* **314**(1), 275–288 (2005)

34. D. Labat, J. Ronchail, J.L. Guyot, Recent advances in wavelet analyses: part 2. Amazon, Parana, Orinoco and Congo discharges time scale variability. *J. Hydrol.* **314**(1), 289–311 (2005)
35. J.P. Lachaux, A. Lutz, D. Rudrauf, D. Cosmelli, M. Le van Quyen, J. Martinerie, F. Varela, Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence. *Clin. Neurophysiol.* **32**(3), 157–174 (2002)
36. M. Malik, Heart rate variability. *Ann. Noninvas. Electr.* **1**(2), 151–181 (1996)
37. S. Mallat, *A Wavelet Tour of Signal Processing*, 3rd edn. (Academic Press, Burlington, 2008)
38. J. Morlet, Sampling theory and wave propagation, in *Issues on Acoustic Signal/Image Processing and Recognition*, vol. I, NATO ASI series, ed. by C.H. Chen (Springer, Berlin, 1983), pp. 233–261
39. A.V. Oppenheim, R.W. Schaffer, J.R. Buck, *Discrete-time signal processing*, vol. 2. (Prentice Hall, New Jersey, 1989)
40. B. Picinbono, On instantaneous amplitude and phase of signals. *IEEE Trans. Signal Proc.* **45**(3), 552–560 (1997)
41. S. Qian, D. Chen, Joint time-frequency analysis. *IEEE Signal Proc. Mag.* **16**(2), 52–67 (1999)
42. A.W. Rihaczek, Signal energy distribution in time and frequency. *IEEE Trans. Inf. Theory* **14**(3), 369–374 (1968)
43. E. Sejdić, I. Djurović, J. Jiang, Time-frequency feature representation using energy concentration: an overview of recent advances. *Dig. Signal Proc.* **19**(1), 153–183 (2009)
44. L.W. Sheppard, A. Stefanovska, P.V.E. McClintock, Testing for time-localised coherence in bivariate data. *Phys. Rev. E* **85**, 046205 (2012)
45. Y. Shiogai, A. Stefanovska, P.V.E. McClintock, Nonlinear dynamics of cardiovascular ageing. *Phys. Rep.* **488**, 51–110 (2010)
46. L. Stankovic, M. Dakovic, T. Thayaparan, *Time-Frequency Signal Analysis with Applications* (Artech House, Norwood, 2013)
47. A. Stefanovska, Coupled oscillators: complex but not complicated cardiovascular and brain interactions. *IEEE Eng. Med. Biol. Mag.* **26**(6), 25–29 (2007)
48. A. Stefanovska, M. Bračič, Physics of the human cardiovascular system. *Contemp. Phys.* **40**(1), 31–55 (1999)
49. A. Stefanovska, M. Bračič, H.D. Kvernmo, Wavelet analysis of oscillations in the peripheral blood circulation measured by laser Doppler technique. *IEEE Trans. Bio. Med. Eng.* **46**(10), 1230–1239 (1999)
50. G. Thakur, E. Brevdo, N.S. Fuckar, H.-T. Wu, The Synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new paleoclimate applications. *Sig. Process.* **93**(5), 1079–1094 (2013)
51. C. Torrence, G.P. Compo, A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* **79**(1), 61–78 (1998)
52. D. Vakman, On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency. *IEEE Trans. Sig. Proc.* **44**(4), 791–797 (1996)
53. M. Vetterli, J. Kovačević, *Wavelets and Subband Coding*, vol. 87 (Prentice Hall, Englewood Cliffs, 1995)
54. J.D. Ville et al., Théorie et Applications de la Notion de Signal Analytique. *Câbles et transmission* **2**(1), 61–74 (1948)
55. M. Wacker, H. Witte, Time-frequency techniques in biomedical signal analysis. A tutorial review of similarities and differences. *Method. Inform. Med.* **52**(4), 279–296 (2013)
56. E. Wigner, On the quantum correction for thermodynamic equilibrium. *Phys. Rev.* **40**(5), 749 (1932)

Nonlinear Mode Decomposition

Theory and Applications

Iatsenko, D.

2015, XXIII, 135 p. 33 illus., 13 illus. in color., Hardcover

ISBN: 978-3-319-20015-6