CHAPTER

Time Meets Frequency

4

When we listen to music, we clearly "hear" the time variation of the sound "frequencies." These localized frequency events are not "pure" tones but packets of close frequencies. The properties of sounds are revealed by transforms that decompose signals over elementary functions that have a narrow localization in time and frequency. Windowed Fourier transforms and wavelet transforms are two important classes of local time-frequency decompositions. Measuring the time variations of "instantaneous" frequencies illustrates the limitations imposed by the Heisenberg uncertainty. Such frequencies are detected as local maxima in windowed Fourier and wavelet dictionaries and define a signal-approximation support. Audio-processing algorithms are implemented by modifying the geometry of this approximation support.

There is no unique definition of time-frequency energy density; all quadratic distributions are related through the averaging of a single quadratic form called the Wigner-Ville distribution. This framework gives another perspective on windowed Fourier and wavelet transforms.

4.1 TIME-FREQUENCY ATOMS

A linear time-frequency transform correlates the signal with a dictionary of waveforms that are concentrated in time and in frequency. The waveforms are called *time-frequency atoms*. Let us consider a general dictionary of atoms $\mathcal{D} = \{\phi_\gamma\}_{\gamma \in \Gamma}$, where γ might be a multiindex parameter. We suppose that $\phi_\gamma \in \mathbf{L^2}(\mathbb{R})$ and that $\|\phi_\gamma\| = 1$. The corresponding linear time-frequency transform of $f \in \mathbf{L^2}(\mathbb{R})$ is defined by

$$\Phi f(\gamma) = \int_{-\infty}^{+\infty} f(t)\, \phi_\gamma^*(t)\, dt = \langle f, \phi_\gamma \rangle.$$

The Parseval formula (2.25) proves that

$$\Phi f(\gamma) = \int_{-\infty}^{+\infty} f(t)\, \phi_\gamma^*(t)\, dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega)\, \hat{\phi}_\gamma^*(\omega)\, d\omega. \tag{4.1}$$

If $\phi_\gamma(t)$ is nearly zero when t is outside a neighborhood of an abscissa u, then $\langle f, \phi_\gamma \rangle$ depends only on the values of f in this neighborhood. Similarly, if $\hat{\phi}_\gamma(\omega)$ is negligible for ω far from ξ, then the right integral of (4.1) proves that $\langle f, \phi_\gamma \rangle$ reveals the properties of $\hat{f}$ in the neighborhood of ξ.

Heisenberg Boxes

The slice of information provided by $\langle f, \phi_\gamma \rangle$ is represented in a time-frequency plane (t, ω) by a rectangle having a position and size that depends on the time-frequency spread of ϕ_γ. Since

$$\|\phi_\gamma\|^2 = \int_{-\infty}^{+\infty} |\phi_\gamma(t)|^2 \, dt = 1,$$

we interpret $|\phi_\gamma(t)|^2$ as a probability distribution centered at

$$u_\gamma = \int_{-\infty}^{+\infty} t \, |\phi_\gamma(t)|^2 \, dt. \tag{4.2}$$

The spread around u_γ is measured by the variance

$$\sigma_t^2(\gamma) = \int_{-\infty}^{+\infty} (t - u_\gamma)^2 \, |\phi_\gamma(t)|^2 \, dt. \tag{4.3}$$

The Plancherel formula (2.26) proves that

$$\int_{-\infty}^{+\infty} |\hat{\phi}_\gamma(\omega)|^2 \, d\omega = 2\pi \|\phi_\gamma\|^2.$$

The center frequency of $\hat{\phi}_\gamma$ is therefore defined by

$$\xi_\gamma = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \omega \, |\hat{\phi}_\gamma(\omega)|^2 \, d\omega, \tag{4.4}$$

and its spread around ξ_γ is

$$\sigma_\omega^2(\gamma) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (\omega - \xi_\gamma)^2 \, |\hat{\phi}_\gamma(\omega)|^2 \, d\omega. \tag{4.5}$$

The time-frequency resolution of ϕ_γ is represented in the time-frequency plane (t, ω) by a Heisenberg box centered at (u_γ, ξ_γ), having a time width equal to $\sigma_t(\gamma)$ and a frequency $\sigma_\omega(\gamma)$. Figure 4.1 illustrates this. The Heisenberg uncertainty Theorem 2.6 proves that the area of the rectangle is at least one-half:

$$\sigma_t \, \sigma_\omega \geqslant \frac{1}{2}. \tag{4.6}$$

This limits the joint resolution of ϕ_γ in time and frequency. The time-frequency plane must be manipulated carefully because a point (t_0, ω_0) is ill-defined. There is no function that is concentrated perfectly at a point t_0 and a frequency ω_0. Only rectangles with an area of at least one-half may correspond to time-frequency atoms.

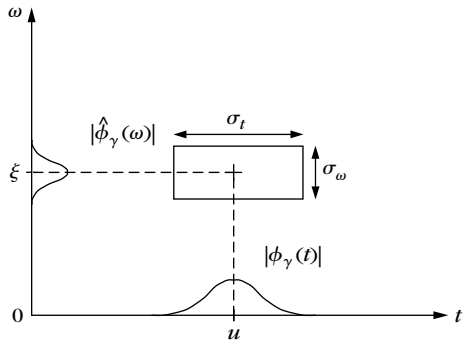

FIGURE 4.1

Heisenberg box representing an atom ϕ_γ.

Translation-Invariant Dictionaries

For pattern recognition, it can be important to construct signal representations that are translation-invariant. When a pattern is translated, its numerical descriptors are then translated but not modified. Observe that for any $\phi_\gamma \in \mathcal{D}$ and any shift u,

$$\langle f(t-u), \phi_\gamma(t)\rangle = \langle f(t), \phi_\gamma(t+u)\rangle.$$

A translation-invariant representation is thus obtained if $\phi_\gamma(t+u)$ is in $\mathcal{D}$ up to a multiplicative constant. Such a dictionary is said to be translation-invariant.

A translation-invariant dictionary is obtained by translating a family of generators $\{\phi_\gamma\}_{\gamma\in\Gamma}$ and can be written $\mathcal{D} = \{\phi_{u,\gamma}\}_{\gamma\in\Gamma, u\in\mathbb{R}}$, with $\phi_{u,\gamma}(t) = \lambda_{u,\gamma}\,\phi_\gamma(t-u)$. The resulting time-frequency transform of f can then be written as a convolution:

$$\Phi f(u,\gamma) = \langle f, \phi_{u,\gamma}\rangle = \int_{-\infty}^{+\infty} f(t)\,\lambda_{u,\gamma}\phi^*_\gamma(t-u)\,dt = \lambda_{u,\gamma} f \star \tilde{\phi}_\gamma(u),$$

with $\tilde{\phi}_\gamma(t) = \phi^*_\gamma(-t)$.

Energy Density

Let us suppose that $\phi_\gamma(t)$ is centered at $t=0$ so that $\phi_{u,\gamma}(t)$ is centered at u. Let ξ_γ be the center frequency of $\hat{\phi}_\gamma(\omega)$ defined in (4.4). The time-frequency box of $\phi_{u,\gamma}$ specifies a neighborhood of (u, ξ_γ), where the energy of f is measured by

$$P_\Phi f(u,\xi_\gamma) = |\langle f, \phi_{u,\gamma}\rangle|^2 = \left|\int_{-\infty}^{+\infty} f(t)\,\phi^*_{u,\gamma}(t)\,dt\right|^2. \tag{4.7}$$

Section 4.5.1 proves that any such energy density is an averaging of the Wigner-Ville distribution, with a kernel that depends on the atoms $\phi_{u,\gamma}$.

EXAMPLE 4.1

A windowed Fourier atom is constructed with a window g modulated by the frequency ξ and translated by u:

$$\phi_{u,\gamma}(t) = g_{u,\xi}(t) = \mathrm{e}^{i\xi t}\, g(t-u). \tag{4.8}$$

The resulting window Fourier dictionary $\mathcal{D} = \{g_{u,\xi}(t)\}_{u,\xi\in\mathbb{R}^2}$ is translation-invariant since $g_{u,\xi} = \mathrm{e}^{i\xi u}\, g_{0,\xi}(t-u)$. A windowed Fourier dictionary is also frequency-invariant because

$$\mathrm{e}^{i\omega t}\, g_{u,\xi}(t) = g_{u,\xi+\omega}(t) \in \mathcal{D}.$$

This dictionary is thus particularly useful to analyze patterns that are translated in time and frequency.

A wavelet atom is a dilation by s and a translation by u of a *mother wavelet* ψ:

$$\phi_{u,\gamma}(t) = \psi_{u,s}(t) = \frac{1}{\sqrt{s}}\,\psi\left(\frac{t-u}{s}\right). \tag{4.9}$$

A wavelet dictionary $\mathcal{D} = \{\psi_{u,s}(t)\}_{u\in\mathbb{R},s\in\mathbb{R}^+}$ is translation-invariant but also scale-invariant because scaling any wavelet produces a dilated wavelet that remains in the dictionary. A wavelet dictionary can be used to analyze patterns translated and scaled by arbitrary factors.

Wavelets and windowed Fourier atoms have well-localized energy in time while their Fourier transform is mostly concentrated in a limited-frequency band. The properties of the resulting transforms are studied in Sections 4.2 and 4.3.

4.2 WINDOWED FOURIER TRANSFORM

In 1946, Gabor [267] introduced windowed Fourier atoms to measure the "frequency variations" of sounds. A real and symmetric window $g(t) = g(-t)$ is translated by u and modulated by the frequency ξ:

$$g_{u,\xi}(t) = \mathrm{e}^{i\xi t} g(t-u). \tag{4.10}$$

It is normalized $\|g\| = 1$ so that $\|g_{u,\xi}\| = 1$ for any $(u,\xi)\in\mathbb{R}^2$. The resulting windowed Fourier transform of $f\in\mathbf{L}^2(\mathbb{R})$ is

$$Sf(u,\xi) = \langle f, g_{u,\xi}\rangle = \int_{-\infty}^{+\infty} f(t)\, g(t-u)\, \mathrm{e}^{-i\xi t}\, dt. \tag{4.11}$$

This transform is also called the *short time Fourier transform* because the multiplication by $g(t-u)$ localizes the Fourier integral in the neighborhood of $t = u$.

As in (4.7), one can define an energy density called a *spectrogram*, denoted P_S:

$$P_S f(u,\xi) = |Sf(u,\xi)|^2 = \left| \int_{-\infty}^{+\infty} f(t)\, g(t-u)\, e^{-i\xi t}\, dt \right|^2. \tag{4.12}$$

The spectrogram measures the energy of f in a time-frequency neighborhood of (u,ξ) specified by the Heisenberg box of $g_{u,\xi}$.

Heisenberg Boxes

Since g is even, $g_{u,\xi}(t) = e^{i\xi t} g(t-u)$ is centered at u. The time spread around u is independent of u and ξ:

$$\sigma_t^2 = \int_{-\infty}^{+\infty} (t-u)^2\, |g_{u,\xi}(t)|^2\, dt = \int_{-\infty}^{+\infty} t^2\, |g(t)|^2\, dt. \tag{4.13}$$

The Fourier transform $\hat{g}$ of g is real and symmetric because g is real and symmetric. The Fourier transform of $g_{u,\xi}$ is

$$\hat{g}_{u,\xi}(\omega) = \hat{g}(\omega-\xi)\, \exp[-iu(\omega-\xi)]. \tag{4.14}$$

It is a translation by ξ of the frequency window $\hat{g}$, so its center frequency is ξ. The frequency spread around ξ is

$$\sigma_\omega^2 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (\omega-\xi)^2\, |\hat{g}_{u,\xi}(\omega)|^2\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \omega^2\, |\hat{g}(\omega)|^2\, d\omega. \tag{4.15}$$

It is independent of u and ξ. Thus, $g_{u,\xi}$ corresponds to a Heisenberg box of area $\sigma_t\, \sigma_\omega$ centered at (u,ξ), as illustrated by Figure 4.2. The size of this box is independent of (u,ξ), which means that a windowed Fourier transform has the same resolution across the time-frequency plane.

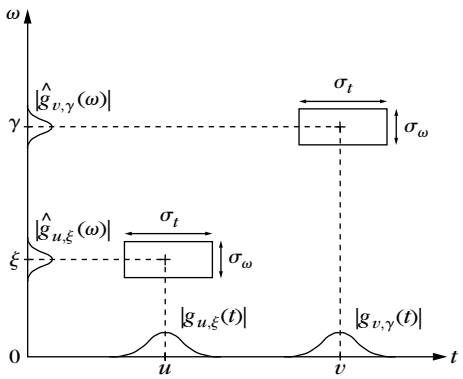

FIGURE 4.2

Heisenberg boxes of two windowed Fourier atoms, $g_{u,\xi}$ and $g_{v,\gamma}$.

EXAMPLE 4.2

A sinusoidal wave $f(t) = \exp(i\xi_0 t)$, the Fourier transform of which is a Dirac $\hat{f}(\omega) = 2\pi\delta(\omega - \xi_0)$, has a windowed Fourier transform:

$$Sf(u, \xi) = \hat{g}(\xi - \xi_0)\ \exp[-iu(\xi - \xi_0)].$$

Its energy is spread over the frequency interval $[\xi_0 - \sigma_\omega/2, \xi_0 + \sigma_\omega/2]$.

EXAMPLE 4.3

The windowed Fourier transform of a Dirac $f(t) = \delta(t - u_0)$ is

$$Sf(u, \xi) = g(u_0 - u)\ \exp(-i\xi u_0).$$

Its energy is spread in the time interval $[u_0 - \sigma_t/2, u_0 + \sigma_t/2]$.

EXAMPLE 4.4

A linear chirp $f(t) = \exp(iat^2)$ has an "instantaneous" frequency that increases linearly in time. For a Gaussian window $g(t) = (\pi\sigma^2)^{-1/4} \exp[-t^2/(2\sigma^2)]$, the windowed Fourier transform of f is calculated using the Fourier transform (2.34) of Gaussian chirps. One can verify that its spectrogram is

$$P_S f(u, \xi) = |Sf(u, \xi)|^2 = \left(\frac{4\pi\sigma^2}{1 + 4a^2\sigma^4}\right)^{1/2} \exp\left(-\frac{\sigma^2(\xi - 2au)^2}{1 + 4a^2\sigma^4}\right). \tag{4.16}$$

For a fixed time u, $P_S f(u, \xi)$ is a Gaussian that reaches its maximum at the frequency $\xi(u) = 2au$. Observe that if we write $f(t) = \exp[i\phi(t)]$, then $\xi(u)$ is equal to the instantaneous frequency, defined as the derivative of the phase: $\omega(u) = \phi'(u) = 2au$. Section 4.4.2 explains this results.

EXAMPLE 4.5

Figure 4.3 gives the spectrogram of a signal that includes a linear chirp, a quadratic chirp, and two modulated Gaussians. The spectrogram is computed with a Gaussian window dilated by $\sigma = 0.05$. As expected from (4.16), the linear chirp yields large amplitude coefficients along the trajectory of its instantaneous frequency, which is a straight line. The quadratic chirp yields large coefficients along a parabola. The two modulated Gaussians produce low- and high-frequency blobs at $u = 0.5$ and $u = 0.87$.

4.2.1 Completeness and Stability

When the time-frequency indices (u, ξ) vary across $\mathbb{R}^2$, the Heisenberg boxes of the atoms $g_{u,\xi}$ cover the whole time-frequency plane. One can expect therefore that f

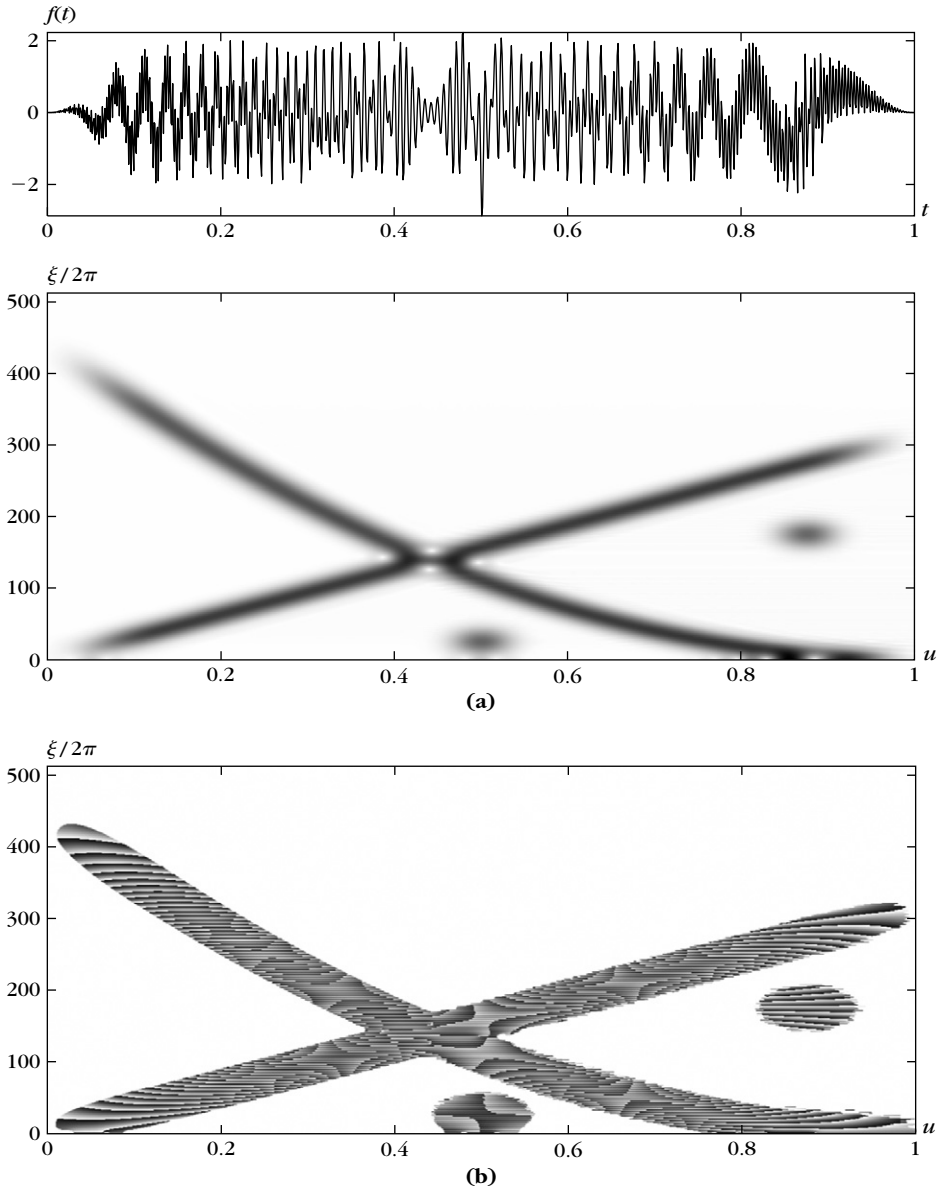

FIGURE 4.3

The signal includes a linear chirp with a frequency that increases, a quadratic chirp with a frequency that decreases, and two modulated Gaussian functions located at $t = 0.5$ and $t = 0.87$. **(a)** Spectrogram $P_S f(u, \xi)$; dark points indicate large-amplitude coefficients. **(b)** Complex phase of $Sf(u, \xi)$ in regions where the modulus $P_S f(u, \xi)$ is nonzero.

can be recovered from its windowed Fourier transform $Sf(u,\xi)$. Theorem 4.1 gives a reconstruction formula and proves that the energy is conserved.

Theorem 4.1. If $f \in \mathbf{L^2}(\mathbb{R})$, then

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Sf(u,\xi)\, g(t-u)\, e^{i\xi t}\, d\xi\, du \tag{4.17}$$

and

$$\int_{-\infty}^{+\infty} |f(t)|^2\, dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |Sf(u,\xi)|^2\, d\xi\, du. \tag{4.18}$$

Proof. The reconstruction formula (4.17) is proved first. Let us apply the Fourier Parseval formula (2.25) to the integral (4.17) with respect to the integration in u. The Fourier transform of $f_\xi(u) = Sf(u,\xi)$ with respect to u is computed by observing that

$$Sf(u,\xi) = \exp(-iu\xi) \int_{-\infty}^{+\infty} f(t)\, g(t-u)\, \exp[i\xi(u-t)]\, dt = \exp(-iu\xi)\, f \star g_\xi(u),$$

where $g_\xi(t) = g(t)\, \exp(it\xi t)$ because $g(t) = g(-t)$. Its Fourier transform therefore is

$$\hat{f}_\xi(\omega) = \hat{f}(\omega+\xi)\, \hat{g}_\xi(\omega+\xi) = \hat{f}(\omega+\xi)\, \hat{g}(\omega).$$

The Fourier transform of $g(t-u)$ with respect to u is $\hat{g}(\omega)\, \exp(-it\omega)$. Thus,

$$\frac{1}{2\pi} \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Sf(u,\xi)\, g(t-u)\, \exp(i\xi t)\, du \right) d\xi$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega+\xi)\, |\hat{g}(\omega)|^2\, \exp[it(\omega+\xi)]\, d\omega \right) d\xi.$$

If $\hat{f} \in \mathbf{L^1}(\mathbb{R})$, we can apply the Fubini theorem (A.2) to reverse the integration order. The inverse Fourier transform proves that

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega+\xi)\, \exp[it(\omega+\xi)]\, d\xi = f(t).$$

Since $\frac{1}{2\pi} \int_{-\infty}^{+\infty} |\hat{g}(\omega)|^2\, d\omega = 1$, we derive (4.17). If $\hat{f} \notin \mathbf{L^1}(\mathbb{R})$, a density argument is used to verify this formula.

Let us now prove the energy conservation (4.18). Since the Fourier transform in u of $Sf(u,\xi)$ is $\hat{f}(\omega+\xi)\, \hat{g}(\omega)$, the Plancherel formula (2.26) applied to the right side of (4.18) gives

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |Sf(u,\xi)|^2\, du\, d\xi = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\hat{f}(\omega+\xi)\, \hat{g}(\omega)|^2\, d\omega\, d\xi.$$

The Fubini theorem applies and the Plancherel formula proves that

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} |\hat{f}(\omega+\xi)|^2\, d\xi = \|f\|^2,$$

which implies (4.18). ■

The reconstruction formula (4.17) can be rewritten

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \langle f, g_{u,\xi}\rangle\, g_{u,\xi}(t)\, d\xi\, du. \tag{4.19}$$

It resembles the decomposition of a signal in an orthonormal basis; however, it is not because the functions $\{g_{u,\xi}\}_{u,\xi\in\mathbb{R}^2}$ are very redundant in $\mathbf{L^2}(\mathbb{R})$. The second equality (4.18) justifies the interpretation of the spectrogram $P_S f(u,\xi) = |Sf(u,\xi)|^2$ as an energy density because its time-frequency sum equals the signal energy.

Reproducing Kernel

A windowed Fourier transform represents a one-dimension signal $f(t)$ by a two-dimensional function $Sf(u,\xi)$. Energy conservation proves that $Sf \in \mathbf{L^2}(\mathbb{R}^2)$. Because $Sf(u,\xi)$ is redundant, it is not true that any $\Phi \in \mathbf{L^2}(\mathbb{R}^2)$ is the windowed Fourier transform of some $f \in \mathbf{L^2}(\mathbb{R})$. Theorem 4.2 gives a necessary and sufficient condition for such a function to be a windowed Fourier transform.

Theorem 4.2. Let $\Phi \in \mathbf{L^2}(\mathbb{R}^2)$. There exists $f \in \mathbf{L^2}(\mathbb{R})$ such that $\Phi(u,\xi) = Sf(u,\xi)$, if and only if,

$$\Phi(u_0,\xi_0) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \Phi(u,\xi)\, K(u_0,u,\xi_0,\xi)\, du\, d\xi, \tag{4.20}$$

with

$$K(u_0,u,\xi_0,\xi) = \langle g_{u,\xi}, g_{u_0,\xi_0}\rangle. \tag{4.21}$$

Proof. Suppose that there exists f such that $\Phi(u,\xi) = Sf(u,\xi)$. Let us replace f with its reconstruction integral (4.17) in the windowed Fourier transform definition:

$$Sf(u_0,\xi_0) = \int_{-\infty}^{+\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} Sf(u,\xi)\, g_{u,\xi}(t)\, du\, d\xi\right) g^*_{u_0,\xi_0}(t)\, dt. \tag{4.22}$$

Inverting the integral on t with the integrals on u and ξ yields (4.20). To prove that the condition (4.20) is sufficient, we define f as in the reconstruction formula (4.17):

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \Phi(u,\xi)\, g(t-u)\, \exp(i\xi t)\, d\xi\, du$$

and show that (4.20) implies that $\Phi(u,\xi) = Sf(u,\xi)$. ■

Ambiguity Function

The reproducing kernel $K(u_0, u, \xi_0, \xi)$ measures the time-frequency overlap of the two atoms $g_{u,\xi}$ and g_{u_0,ξ_0}. The amplitude of $K(u_0, u, \xi_0, \xi)$ decays with $u_0 - u$ and $\xi_0 - \xi$ at a rate that depends on the energy concentration of g and $\hat{g}$. Replacing $g_{u,\xi}$ and g_{u_0,ξ_0} by their expression and making the change of variable $v = t - (u + u_0)/2$ in the inner product integral (4.21) yields

$$K(u_0, u, \xi_0, \xi) = \exp\left(-\frac{i}{2}(\xi_0 - \xi)(u + u_0)\right) Ag(u_0 - u, \xi_0 - \xi), \tag{4.23}$$

where

$$Ag(\tau, \gamma) = \int_{-\infty}^{+\infty} g\left(v + \frac{\tau}{2}\right) g\left(v - \frac{\tau}{2}\right) \mathrm{e}^{-i\gamma v}\, dv \tag{4.24}$$

is called the *ambiguity function* of g. Using the Parseval formula to replace this time integral with a Fourier integral gives

$$Ag(\tau, \gamma) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{g}\left(\omega + \frac{\gamma}{2}\right) \hat{g}\left(\omega - \frac{\gamma}{2}\right) \mathrm{e}^{i\tau\omega}\, d\omega. \tag{4.25}$$

The decay of the ambiguity function measures the spread of g in time and of $\hat{g}$ in frequency. For example, if g has a support included in an interval of size T, then $Ag(\tau, \omega) = 0$ for $|\tau| \geqslant T/2$. The integral (4.25) shows that the same result applies to the support of $\hat{g}$.

4.2.2 Choice of Window

The resolution in time and frequency of the windowed Fourier transform depends on the spread of the window in time and frequency. This can be measured from the decay of the ambiguity function (4.24) or more simply from the area $\sigma_t\, \sigma_\omega$ of the Heisenberg box. The uncertainty Theorem 2.6 proves that this area reaches the minimum value $1/2$, if, and only if, g is a Gaussian. The ambiguity function $Ag(\tau, \gamma)$ is then a two-dimensional Gaussian.

Window Scale

The time-frequency localization of g can be modified with a scaling. Suppose that g has a Heisenberg time and frequency width, respectively, equal to σ_t and σ_ω. Let $g_s(t) = s^{-1/2} g(t/s)$ be its dilation by s. A change of variables in the integrals (4.13) and (4.15) shows that the Heisenberg time and frequency width of g_s are, respectively, $s\sigma_t$ and σ_ω/s. The area of the Heisenberg box is not modified, but it is dilated by s in time and compressed by s in frequency. Similarly, a change of variable in the ambiguity integral (4.24) shows that the ambiguity function is dilated in time and frequency, respectively, by s and $1/s$:

$$Ag_s(\tau, \gamma) = Ag\left(\frac{\tau}{s}, s\gamma\right).$$

The choice of a particular scale s depends on the desired resolution trade-off between time and frequency.

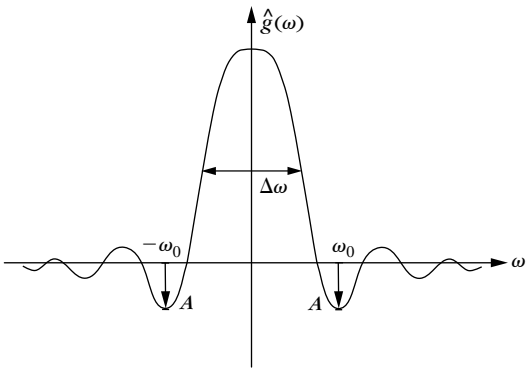

FIGURE 4.4

The energy spread of $\hat{g}$ is measured by its bandwidth $\Delta\omega$ and the maximum amplitude A of the first side lobes, located at $\omega = \pm\omega_0$.

Finite Support

In numerical applications, g must have a compact support. Theorem 2.7 proves that its Fourier transform $\hat{g}$ necessarily has an infinite support. It is a symmetric function with a main lobe centered at $\omega = 0$, which decays to zero with oscillations. Figure 4.4 illustrates its behavior. To maximize the frequency resolution of the transform, we must concentrate the energy of $\hat{g}$ near $\omega = 0$. The following three important parameters evaluate the spread of $\hat{g}$:

- The root mean-square bandwidth $\Delta\omega$, which is defined by

$$\frac{|\hat{g}(\Delta\omega/2)|^2}{|\hat{g}(0)|^2} = \frac{1}{2}.$$

- The maximum amplitude A of the first side lobes located at $\omega = \pm\omega_0$ in Figure 4.4. It is measured in decibels:

$$A = 10 \log_{10} \frac{|\hat{g}(\omega_0)|^2}{|\hat{g}(0)|^2}.$$

- The polynomial exponent p, which gives the asymptotic decay of $|\hat{g}(\omega)|$ for large frequencies:

$$|\hat{g}(\omega)| = O(\omega^{-p-1}). \tag{4.26}$$

Table 4.1 gives the values of these three parameters for several windows g having a support restricted to $[-1/2, 1/2]$ [293]. Figure 4.5 shows the graph of these windows.

To interpret the three frequency parameters, let us consider the spectrogram of a frequency tone $f(t) = \exp(i\xi_0 t)$. If $\Delta\omega$ is small, then $|Sf(u, \xi)|^2 = |\hat{g}(\xi - \xi_0)|^2$ has energy concentrated near $\xi = \xi_0$. The side lobes of $\hat{g}$ create "shadows" at $\xi = \xi_0 \pm \omega_0$, which can be neglected if A is also small.

Table 4.1 Frequency Parameters of Five Windows g

Name	$g(t)$	$\Delta\omega$	A	p
Rectangle	1	0.89	-13db	0
Hamming	$0.54 + 0.46\cos(2\pi t)$	1.36	-43db	0
Gaussian	$\exp(-18t^2)$	1.55	-55db	0
Hanning	$\cos^2(\pi t)$	1.44	-32db	2
Blackman	$0.42 + 0.5\cos(2\pi t) + 0.08\cos(4\pi t)$	1.68	-58db	2

Note: Supports are restricted to $[-1/2, 1/2]$. *The windows are normalized so that* $g(0) = 1$ *but* $\|g\| \neq 1$.

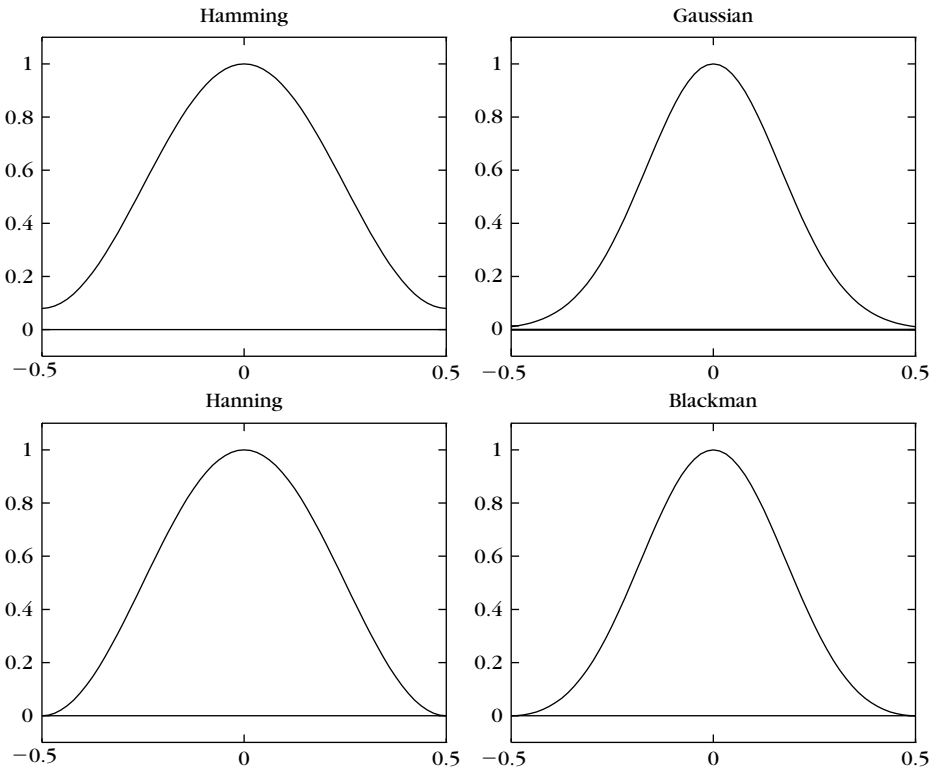

FIGURE 4.5

Graphs of four windows g with supports that are $[-1/2, 1/2]$.

If the frequency tone is embedded in a signal that has other components of much higher energy at different frequencies, the tone can still be detected if $\hat{g}(\omega - \xi)$ attenuates these components rapidly when $|\omega - \xi|$ increases. This means that $|\hat{g}(\omega)|$ has a rapid decay, and Theorem 2.5 proves that this decay depends on

the regularity of g. Property (4.26) is typically satisfied by windows that are p times differentiable.

4.2.3 Discrete Windowed Fourier Transform

The discretization and fast computation of the windowed Fourier transform follow the same ideas as the discretization of the Fourier transform described in Section 3.3. We consider discrete signals of period N. The window $g[n]$ is chosen to be a symmetric discrete signal of period N with unit norm $\|g\| = 1$. Discrete windowed Fourier atoms are defined by

$$g_{m,l}[n] = g[n-m] \exp\left(\frac{i2\pi ln}{N}\right).$$

The discrete Fourier transform (DFT) of $g_{m,l}$ is

$$\hat{g}_{m,l}[k] = \hat{g}[k-l] \exp\left(\frac{-i2\pi m(k-l)}{N}\right).$$

The discrete windowed Fourier transform of a signal f of period N is

$$Sf[m,l] = \langle f, g_{m,l}\rangle = \sum_{n=0}^{N-1} f[n]\, g[n-m] \exp\left(\frac{-i2\pi ln}{N}\right), \tag{4.27}$$

For each $0 \leqslant m < N$, $Sf[m,l]$ is calculated for $0 \leqslant l < N$ with a DFT of $f[n]g[n-m]$. This is performed with N FFT procedures of size N, and therefore requires a total of $O(N^2 \log_2 N)$ operations. Figure 4.3 is computed with this algorithm.

Inverse Transform

Theorem 4.3 discretizes the reconstruction formula and the energy conservation of Theorem 4.1.

Theorem 4.3. If f is a signal of period N, then

$$f[n] = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{l=0}^{N-1} Sf[m,l]\, g[n-m] \exp\left(\frac{i2\pi ln}{N}\right) \tag{4.28}$$

and

$$\sum_{n=0}^{N-1} |f[n]|^2 = \frac{1}{N} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} |Sf[m,l]|^2. \tag{4.29}$$

This theorem is proved by applying the Parseval and Plancherel formulas of the discrete Fourier transform, exactly as in the proof of Theorem 4.1 (Exercise 4.1). The energy conservation (4.29) proves that this windowed Fourier transform defines a tight frame, as explained in Chapter 5. The reconstruction formula (4.28) is rewritten

$$f[n] = \frac{1}{N} \sum_{m=0}^{N-1} g[n-m] \sum_{l=0}^{N-1} Sf[m,l] \exp\left(\frac{i2\pi ln}{N}\right).$$

The second sum computes, for each $0 \leqslant m < N$, the inverse DFT of $Sf[m, l]$ with respect to l. This is calculated with N FFT procedures, requiring a total of $O(N^2 \log_2 N)$ operations.

A discrete windowed Fourier transform is an N^2 image $Sf[l, m]$ that is very redundant because it is entirely specified by a signal f of size N. The redundancy is characterized by a discrete reproducing kernel equation, which is the discrete equivalent of (4.20) (Exercise 4.1).

4.3 WAVELET TRANSFORMS

To analyze signal structures of very different sizes, it is necessary to use time-frequency atoms with different time supports. The wavelet transform decomposes signals over dilated and translated wavelets. A wavelet is a function $\psi \in \mathbf{L^2}(\mathbb{R})$ with a zero average:

$$\int_{-\infty}^{+\infty} \psi(t)\, dt = 0. \tag{4.30}$$

It is normalized $\|\psi\| = 1$ and centered in the neighborhood of $t = 0$. A dictionary of time-frequency atoms is obtained by scaling ψ by s and translating it by u:

$$\mathcal{D} = \left\{ \psi_{u,s}(t) = \frac{1}{\sqrt{s}}\, \psi\left(\frac{t-u}{s}\right) \right\}_{u \in \mathbb{R}, s \in \mathbb{R}^+}$$

These atoms remain normalized: $\|\psi_{u,s}\| = 1$. The wavelet transform of $f \in \mathbf{L^2}(\mathbb{R})$ at time u and scale s is

$$Wf(u, s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t)\, \frac{1}{\sqrt{s}}\, \psi^*\left(\frac{t-u}{s}\right) dt. \tag{4.31}$$

Linear Filtering

The wavelet transform can be rewritten as a convolution product:

$$Wf(u, s) = \int_{-\infty}^{+\infty} f(t)\, \frac{1}{\sqrt{s}}\, \psi^*\left(\frac{t-u}{s}\right) dt = f \star \bar{\psi}_s(u), \tag{4.32}$$

with

$$\bar{\psi}_s(t) = \frac{1}{\sqrt{s}}\, \psi^*\left(\frac{-t}{s}\right).$$

The Fourier transform of $\bar{\psi}_s(t)$ is

$$\widehat{\bar{\psi}}_s(\omega) = \sqrt{s}\, \hat{\psi}^*(s\omega). \tag{4.33}$$

Since $\hat{\psi}(0) = \int_{-\infty}^{+\infty} \psi(t)\, dt = 0$, it appears that $\hat{\psi}$ is the transfer function of a band-pass filter. The convolution (4.32) computes the wavelet transform with dilated band-pass filters.

Analytic Versus Real Wavelets

Like a windowed Fourier transform, a wavelet transform can measure the time evolution of frequency transients. This requires using a complex analytic wavelet, which can separate amplitude and phase components. The properties of this analytic wavelet transform are described in Section 4.3.2, and its application to the measurement of instantaneous frequencies is explained in Section 4.4.3. In contrast, real wavelets are often used to detect sharp signal transitions. Section 4.3.1 introduces elementary properties of real wavelets, which are developed in Chapter 6.

4.3.1 Real Wavelets

Suppose that ψ is a real wavelet. Since it has a zero average, the wavelet integral

$$Wf(u,s) = \int_{-\infty}^{+\infty} f(t)\, \frac{1}{\sqrt{s}}\, \psi^*\left(\frac{t-u}{s}\right) dt$$

measures the variation of f in a neighborhood of u proportional to s. Section 6.1.3 proves that when scale s goes to zero, the decay of the wavelet coefficient characterizes the regularity of f in the neighborhood of u. This has important applications for detecting transients and analyzing fractals. This section concentrates on the completeness and redundancy properties of real wavelet transforms.

EXAMPLE 4.6

Wavelets equal to the second derivative of a Gaussian are called *Mexican hats*. They were first used in computer vision to detect multiscale edges [487]. The normalized Mexican hat wavelet is

$$\psi(t) = \frac{2}{\pi^{1/4}\sqrt{3\sigma}} \left(\frac{t^2}{\sigma^2} - 1\right) \exp\left(\frac{-t^2}{2\sigma^2}\right). \tag{4.34}$$

For $\sigma = 1$, Figure 4.6 plots $-\psi$ and its Fourier transform:

$$\hat{\psi}(\omega) = \frac{-\sqrt{8}\,\sigma^{5/2}\,\pi^{1/4}}{\sqrt{3}}\, \omega^2\, \exp\left(\frac{-\sigma^2\omega^2}{2}\right). \tag{4.35}$$

Figure 4.7 shows the wavelet transform of a piecewise regular signal on the left and, almost everywhere, singular on the right. The maximum scale is smaller than 1 because the support of f is normalized to $[0, 1]$. The minimum scale is limited by the sampling interval of the discretized signal used in numerical calculations. When the scale decreases, the wavelet transform has a rapid decay to zero in the regions where the signal is regular. The isolated singularities on the left create cones of large-amplitude wavelet coefficients that converge to the locations of the singularities. This is further explained in Chapter 6.

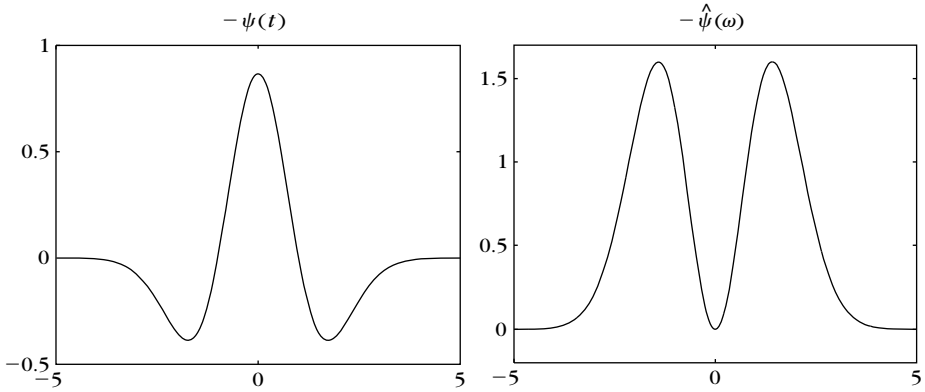

FIGURE 4.6

Mexican-hat wavelet (4.34) for $\sigma = 1$ and its Fourier transform.

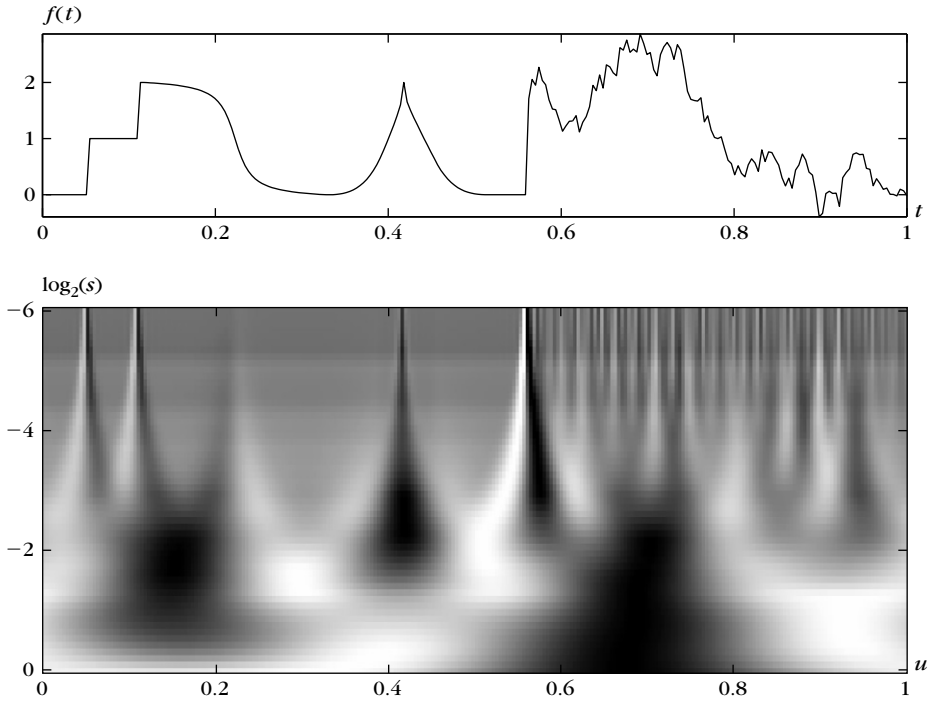

FIGURE 4.7

Real wavelet transform $Wf(u, s)$ computed with a Mexican-hat wavelet (4.34). The vertical axis represents $\log_2 s$. Black, gray, and white points correspond, respectively, to positive, zero, and negative wavelet coefficients.

A real wavelet transform is complete and maintains an energy conservation as long as the wavelet satisfies a weak admissibility condition, specified by Theorem 4.4. This theorem was first proved in 1964 by the mathematician Calderón [132] from a different point of view. Wavelets did not appear as such, but Calderón defines a wavelet transform as a convolution operator that decomposes the identity. Grossmann and Morlet [288] were not aware of Calderón's work when they proved the same formula for signal processing.

Theorem 4.4: *Calderón, Grossmann and Morlet.* Let $\psi \in \mathbf{L^2}(\mathbb{R})$ be a real function such that

$$C_\psi = \int_0^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega}\, d\omega < +\infty. \tag{4.36}$$

Any $f \in \mathbf{L^2}(\mathbb{R})$ satisfies

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} wf(u,s)\, \frac{1}{\sqrt{s}}\, \psi\left(\frac{t-u}{s}\right) du\, \frac{ds}{s^2}, \tag{4.37}$$

and

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} |wf(u,s)|^2\, du\, \frac{ds}{s^2}. \tag{4.38}$$

Proof. The proof of (4.38) is almost identical to the proof of (4.18). Let us concentrate on the proof of (4.37). The right integral $b(t)$ of (4.37) can be rewritten as a sum of convolutions. Inserting $wf(u,s) = f \star \bar{\psi}_s(u)$ with $\psi_s(t) = s^{-1/2}\, \psi(t/s)$ yields

$$\begin{aligned} b(t) &= \frac{1}{C_\psi} \int_0^{+\infty} wf(.,s) \star \psi_s(t)\, \frac{ds}{s^2} \\ &= \frac{1}{C_\psi} \int_0^{+\infty} f \star \bar{\psi}_s \star \psi_s(t)\, \frac{ds}{s^2}. \end{aligned} \tag{4.39}$$

The "." indicates the variable over which the convolution is calculated. We prove that $b = f$ by showing that their Fourier transforms are equal. The Fourier transform of b is

$$\hat{b}(\omega) = \frac{1}{C_\psi} \int_0^{+\infty} \hat{f}(\omega)\, \sqrt{s}\, \hat{\psi}^*(s\omega)\, \sqrt{s}\, \hat{\psi}(s\omega)\, \frac{ds}{s^2} = \frac{\hat{f}(\omega)}{C_\psi} \int_0^{+\infty} |\hat{\psi}(s\,\omega)|^2\, \frac{ds}{s}.$$

Since ψ is real we know that $|\hat{\psi}(-\omega)|^2 = |\hat{\psi}(\omega)|^2$. The change of variable $\xi = s\omega$ thus proves that

$$\hat{b}(\omega) = \frac{1}{C_\psi} \hat{f}(\omega) \int_0^{+\infty} \frac{|\hat{\psi}(\xi)|^2}{\xi}\, d\xi = \hat{f}(\omega).$$

■

The theorem hypothesis

$$C_\psi = \int_0^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega}\, d\omega < +\infty$$

is called the wavelet *admissibility condition*. To guarantee that this integral is finite, we must ensure that $\hat{\psi}(0) = 0$, which explains why wavelets must have a zero average. This condition is nearly sufficient. If $\hat{\psi}(0) = 0$ and $\hat{\psi}(\omega)$ is continuously differentiable, then the admissibility condition is satisfied. One can verify that $\hat{\psi}(\omega)$ is continuously differentiable if ψ has a sufficient time decay:

$$\int_{-\infty}^{+\infty} (1 + |t|)\, |\psi(t)|\, dt < +\infty.$$

Reproducing Kernel

Like a windowed Fourier transform, a wavelet transform is a redundant representation with a redundancy characterized by a reproducing kernel equation. Inserting the reconstruction formula (4.37) into the definition of the wavelet transform yields

$$Wf(u_0, s_0) = \int_{-\infty}^{+\infty} \left(\frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} Wf(u, s)\, \psi_{u,s}(t)\, du \frac{ds}{s^2} \right) \psi^*_{u_0,s_0}(t)\, dt.$$

Interchanging these integrals gives

$$Wf(u_0, s_0) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} K(u, u_0, s, s_0)\, Wf(u, s)\, du\, \frac{ds}{s^2}, \tag{4.40}$$

with

$$K(u_0, u, s_0, s) = \langle \psi_{u,s}, \psi_{u_0,s_0} \rangle. \tag{4.41}$$

The reproducing kernel $K(u_0, u, s_0, s)$ measures the correlation of two wavelets, $\psi_{u,s}$ and ψ_{u_0,s_0}. The reader can verify that any function $\Phi(u, s)$ is the wavelet transform of some $f \in \mathbf{L^2}(\mathbb{R})$ if and only if it satisfies the reproducing kernel equation (4.40).

Scaling Function

When $Wf(u, s)$ is known only for $s < s_0$, to recover f we need a complement of information that corresponds to $Wf(u, s)$ for $s > s_0$. This is obtained by introducing a *scaling function* ϕ that is an aggregation of wavelets at scales larger than 1. The modulus of its Fourier transform is defined by

$$|\hat{\phi}(\omega)|^2 = \int_1^{+\infty} |\hat{\psi}(s\omega)|^2 \frac{ds}{s} = \int_\omega^{+\infty} \frac{|\hat{\psi}(\xi)|^2}{\xi}\, d\xi, \tag{4.42}$$

and the complex phase of $\hat{\phi}(\omega)$ can be arbitrarily chosen. One can verify that $\|\phi\| = 1$, and we can derive from the admissibility condition (4.36) that

$$\lim_{\omega \to 0} |\hat{\phi}(\omega)|^2 = C_\psi. \tag{4.43}$$

The scaling function therefore can be interpreted as the impulse response of a low-pass filter. Let us denote

$$\phi_s(t) = \frac{1}{\sqrt{s}}\, \phi\left(\frac{t}{s}\right) \text{ and } \bar{\phi}_s(t) = \phi^*_s(-t).$$

The low-frequency approximation of f at scale s is

$$Lf(u,s)=\left\langle f(t),\frac{1}{\sqrt{s}}\phi\left(\frac{t-u}{s}\right)\right\rangle=f\star\bar{\phi}_s(u). \tag{4.44}$$

With a minor modification of Theorem 4.4's, proof it can be shown that (Exercise 4.3)

$$f(t)=\frac{1}{C_\psi}\int_0^{s_0} wf(.,s)\star\psi_s(t)\,\frac{ds}{s^2}+\frac{1}{C_\psi s_0}Lf(.,s_0)\star\phi_{s_0}(t). \tag{4.45}$$

EXAMPLE 4.7

If ψ is the second-order derivative of a Gaussian with a Fourier transform given by (4.35), then the integration (4.42) yields

$$\hat{\phi}(\omega)=\frac{2\sigma^{3/2}\pi^{1/4}}{\sqrt{3}}\sqrt{\omega^2+\frac{1}{\sigma^2}}\,\exp\left(-\frac{\sigma^2\omega^2}{2}\right). \tag{4.46}$$

Figure 4.8 displays ϕ and $\hat{\phi}$ for $\sigma=1$.

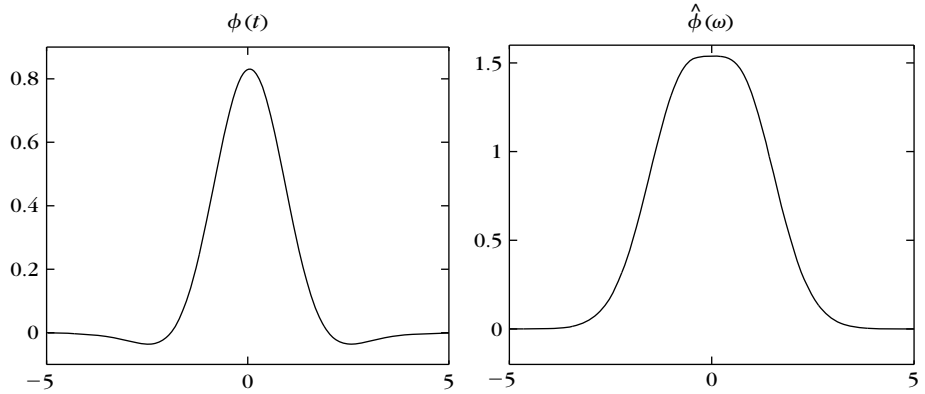

FIGURE 4.8

Scaling function associated to a Mexican-hat wavelet and its Fourier transform calculated with (4.46).

4.3.2 Analytic Wavelets

To analyze the time evolution of frequency tones, it is necessary to use an analytic wavelet to separate the phase and amplitude information of signals. The properties of the resulting analytic wavelet transform are studied next.

Analytic Signal

A function $f_a \in \mathbf{L}^2(\mathbb{R})$ is said to be *analytic* if its Fourier transform is zero for negative frequencies:

$$\hat{f}_a(\omega) = 0 \quad \text{if } \omega < 0.$$

An analytic function is necessarily complex but is entirely characterized by its real part. Indeed, the Fourier transform of its real part $f = \mathrm{Re}[f_a]$ is

$$\hat{f}(\omega) = \frac{\hat{f}_a(\omega) + \hat{f}_a^*(-\omega)}{2},$$

and this relation can be inverted:

$$\hat{f}_a(\omega) = \begin{cases} 2\hat{f}(\omega) & \text{if } \omega \geqslant 0 \\ 0 & \text{if } \omega < 0 \end{cases} \tag{4.47}$$

The analytic part $f_a(t)$ of a signal $f(t)$ is the inverse Fourier transform of $\hat{f}_a(\omega)$ defined by (4.47).

Discrete Analytic Part

The analytic part $f_a[n]$ of a discrete signal $f[n]$ of size N is also computed by setting the negative frequency components of its discrete Fourier transform to zero. The Fourier transform values at $k = 0$ and $k = N/2$ must be carefully adjusted so that $\mathrm{Re}[f_a] = f$ (Exercise 3.4):

$$\hat{f}_a[k] = \begin{cases} \hat{f}[k] & \text{if } k = 0,\ N/2 \\ 2\hat{f}[k] & \text{if } 0 < k < N/2 \\ 0 & \text{if } N/2 < k < N \end{cases} \tag{4.48}$$

We obtain $f_a[n]$ by computing the inverse DFT.

EXAMPLE 4.8

The Fourier transform of

$$f(t) = a\cos(\omega_0 t + \phi) = \frac{a}{2}\Big(\exp[i(\omega_0 t + \phi)] + \exp[-i(\omega_0 t + \phi)]\Big)$$

is

$$\hat{f}(\omega) = \pi a\Big(\exp(i\phi)\,\delta(\omega - \omega_0) + \exp(-i\phi)\,\delta(\omega + \omega_0)\Big).$$

The Fourier transform of the analytic part computed with (4.47) is $\hat{f}_a(\omega) = 2\pi a\,\exp(i\phi)\,\delta(\omega - \omega_0)$ and therefore

$$f_a(t) = a\,\exp[i(\omega_0 t + \phi)]. \tag{4.49}$$

Time-Frequency Resolution

An analytic wavelet transform is calculated with an analytic wavelet ψ:

$$wf(u,s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t)\, \frac{1}{\sqrt{s}}\, \psi^*\left(\frac{t-u}{s}\right) dt. \tag{4.50}$$

Its time-frequency resolution depends on the time-frequency spread of the wavelet atoms $\psi_{u,s}$. We suppose that ψ is centered at 0, which implies that $\psi_{u,s}$ is centered at $t = u$. With the change of variable $v = \frac{t-u}{s}$, we verify that

$$\int_{-\infty}^{+\infty} (t-u)^2\, |\psi_{u,s}(t)|^2\, dt = s^2\, \sigma_t^2, \tag{4.51}$$

with $\sigma_t^2 = \int_{-\infty}^{+\infty} t^2\, |\psi(t)|^2\, dt$. Since $\hat{\psi}(\omega)$ is zero at negative frequencies, the center frequency η of $\hat{\psi}$ is

$$\eta = \frac{1}{2\pi} \int_0^{+\infty} \omega\, |\hat{\psi}(\omega)|^2\, d\omega. \tag{4.52}$$

The Fourier transform of $\psi_{u,s}$ is a dilation of $\hat{\psi}$ by $1/s$:

$$\hat{\psi}_{u,s}(\omega) = \sqrt{s}\, \hat{\psi}(s\omega)\, \exp(-i\omega u). \tag{4.53}$$

Its center frequency therefore is η/s. The energy spread of $\hat{\psi}_{u,s}$ around η/s is

$$\frac{1}{2\pi} \int_0^{+\infty} \left(\omega - \frac{\eta}{s}\right)^2 \left|\hat{\psi}_{u,s}(\omega)\right|^2 d\omega = \frac{\sigma_\omega^2}{s^2}, \tag{4.54}$$

with

$$\sigma_\omega^2 = \frac{1}{2\pi} \int_0^{+\infty} (\omega - \eta)^2\, |\hat{\psi}(\omega)|^2\, d\omega.$$

Thus, the energy spread of a wavelet time-frequency atom $\psi_{u,s}$ corresponds to a Heisenberg box centered at $(u,\, \eta/s)$, of size $s\sigma_t$ along time and σ_ω/s along frequency. The area of the rectangle remains equal to $\sigma_t\, \sigma_\omega$ at all scales but the resolution in time and frequency depends on s, as illustrated in Figure 4.9.

An analytic wavelet transform defines a local time-frequency energy density $P_W f$, which measures the energy of f in the Heisenberg box of each wavelet $\psi_{u,s}$ centered at $(u, \xi = \eta/s)$:

$$P_W f(u,\xi) = |Wf(u,s)|^2 = \left|Wf\left(u, \frac{\eta}{\xi}\right)\right|^2. \tag{4.55}$$

This energy density is called a *scalogram*.

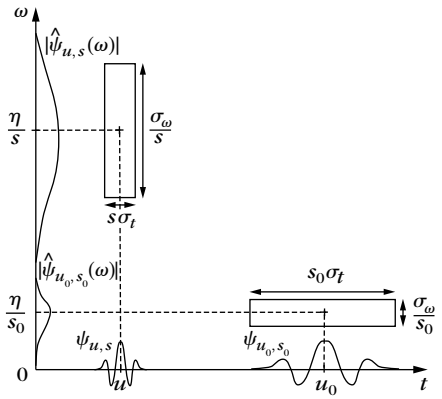

FIGURE 4.9

Heisenberg boxes of two wavelets. Smaller scales decrease the time spread but increase the frequency support, which is shifted toward higher frequencies.

Completeness

An analytic wavelet transform of f depends only on its analytic part f_a. Theorem 4.5 derives a reconstruction formula and proves that energy is conserved for real signals.

Theorem 4.5. For any $f \in \mathbf{L^2}(\mathbb{R})$,

$$Wf(u,s) = \frac{1}{2} Wf_a(u,s). \tag{4.56}$$

If $C_\psi = \int_0^{+\infty} \omega^{-1} |\hat{\psi}(\omega)|^2 \, d\omega < +\infty$ and f is real, then

$$f(t) = \frac{2}{C_\psi} \mathrm{Re}\left[\int_0^{+\infty} \int_{-\infty}^{+\infty} Wf(u,s)\, \psi_s(t-u)\, du\, \frac{ds}{s^2} \right], \tag{4.57}$$

and

$$\|f\|^2 = \frac{2}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} |Wf(u,s)|^2 \, du\, \frac{ds}{s^2}. \tag{4.58}$$

Proof. Let us first prove (4.56). The Fourier transform with respect to u of

$$f_s(u) = wf(u,s) = f \star \bar{\psi}_s(u)$$

is

$$\hat{f}_s(\omega) = \hat{f}(\omega)\, \sqrt{s}\, \hat{\psi}^*(s\omega).$$

Since $\hat{\psi}(\omega) = 0$ at negative frequencies, and $\hat{f}_a(\omega) = 2\hat{f}(\omega)$ for $\omega \geqslant 0$, we derive that

$$\hat{f}_s(\omega) = \frac{1}{2} \hat{f}_a(\omega)\, \sqrt{s}\, \hat{\psi}^*(s\omega),$$

which is the Fourier transform of (4.56).

With the same derivations as in the proof of (4.37), one can verify that the inverse wavelet formula reconstructs the analytic part of f:

$$f_a(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} W f_a(u,s)\, \psi_s(t-u)\, \frac{ds}{s^2}\, du. \tag{4.59}$$

Since $f = \mathrm{Re}[f_a]$, inserting (4.56) proves (4.57).

An energy conservation for the analytic part f_a is proved as in (4.38) by applying the Plancherel formula:

$$\int_{-\infty}^{+\infty} |f_a(t)|^2\, dt = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} |W_a f(u,s)|^2\, du\, \frac{ds}{s^2}.$$

Since $W f_a(u,s) = 2W f(u,s)$ and $\|f_a\|^2 = 2\|f\|^2$, equation (4.58) follows. ■

If f is real, the change of variable $\xi = 1/s$ in the energy conservation (4.58) proves that

$$\|f\|^2 = \frac{2}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} P_W f(u,\xi)\, du\, d\xi.$$

It justifies the interpretation of a scalogram as a time-frequency energy density.

Wavelet Modulated Windows

An analytic wavelet can be constructed with a frequency modulation of a real and symmetric window g. The Fourier transform of

$$\psi(t) = g(t)\, \exp(i\eta t) \tag{4.60}$$

is $\hat{\psi}(\omega) = \hat{g}(\omega - \eta)$. If $\hat{g}(\omega) = 0$ for $|\omega| > \eta$, then $\hat{\psi}(\omega) = 0$ for $\omega < 0$. Therefore, ψ is analytic, as shown in Figure 4.10. Since g is real and even, $\hat{g}$ is also real and symmetric. The center frequency of $\hat{\psi}$ is therefore η and

$$|\hat{\psi}(\eta)| = \sup_{\omega \in \mathbb{R}} |\hat{\psi}(\omega)| = \hat{g}(0). \tag{4.61}$$

A *Gabor wavelet* $\psi(t) = g(t)\, e^{i\eta t}$ is obtained with a Gaussian window:

$$g(t) = \frac{1}{(\sigma^2 \pi)^{1/4}}\, \exp\left(\frac{-t^2}{2\sigma^2}\right). \tag{4.62}$$

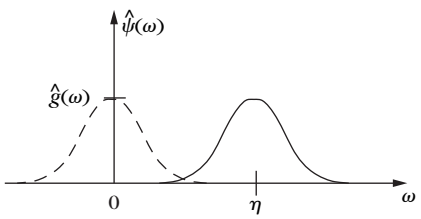

FIGURE 4.10

Fourier transform $\hat{\psi}(\omega)$ of a wavelet $\psi(t) = g(t)\, \exp(i\eta t)$.

The Fourier transform of this window is $\hat{g}(\omega) = (4\pi\sigma^2)^{1/4} \exp(-\sigma^2\omega^2/2)$. If $\sigma^2\eta^2 \gg 1$ then $\hat{g}(\omega) \approx 0$ for $|\omega| > \eta$. Thus, such Gabor wavelets are considered to be approximately analytic.

EXAMPLE 4.9

The wavelet transform of $f(t) = a\ \exp(i\omega_0 t)$ is

$$Wf(u,s) = a\sqrt{s}\,\hat{\psi}^*(s\omega_0)\ \exp(i\omega_0 u) = a\,\sqrt{s}\,\hat{g}(s\omega_0 - \eta)\ \exp(i\omega_0 u).$$

Observe that the normalized scalogram is maximum at $\xi = \omega_0$:

$$\frac{\xi}{\eta} P_W f(u,\xi) = \frac{1}{s}|Wf(u,s)|^2 = a^2 \left|\hat{g}\left(\eta\left(\frac{\omega_0}{\xi} - 1\right)\right)\right|^2.$$

EXAMPLE 4.10

The wavelet transform of a linear chirp $f(t) = \exp(iat^2) = \exp[i\phi(t)]$ is computed for a Gabor wavelet with a Gaussian window given in (4.62). By using the Fourier transform of Gaussian chirps (2.34), one can verify that

$$\frac{|Wf(u,s)|^2}{s} = \left(\frac{4\pi\sigma^2}{1 + 4s^2a^2\sigma^4}\right)^{1/2} \exp\left(\frac{-\sigma^2}{1 + 4a^2s^4\sigma^4}(\eta - 2asu)^2\right).$$

As long as $4a^2s^4\sigma^4 \ll 1$ at a fixed time u, the renormalized scalogram $\eta^{-1}\xi P_W f(u,\xi)$ is a Gaussian function of s that reaches its maximum at

$$\xi(u) = \frac{\eta}{s(u)} = \phi'(u) = 2\,a\,u. \tag{4.63}$$

Section 4.4.3 explains why the amplitude is maximum at the instantaneous frequency $\phi'(u)$.

EXAMPLE 4.11

Figure 4.11 displays the normalized scalogram $\eta^{-1}\xi P_W f(u,\xi)$, and the complex phase $\Theta_W(u,\xi)$ of $Wf(u,s)$, for the signal f of Figure 4.3. The frequency bandwidth of wavelet atoms is proportional to $1/s = \xi/\eta$. The frequency resolution of the scalogram is therefore finer than the spectrogram at low frequencies but coarser than the spectrogram at higher frequencies. This explains why the wavelet transform produces interference patterns between the high-frequency Gabor function at the abscissa $t = 0.87$ and the quadratic chirp at the same location, whereas the spectrogram in Figure 4.3 separates them well.

4.3.3 Discrete Wavelets

Let $\bar{f}(t)$ be a continuous time signal defined over $[0, 1]$. Let $f[n]$ be the discrete signal obtained by a low-pass filtering of $\bar{f}$ and uniform sampling at intervals N^{-1}.

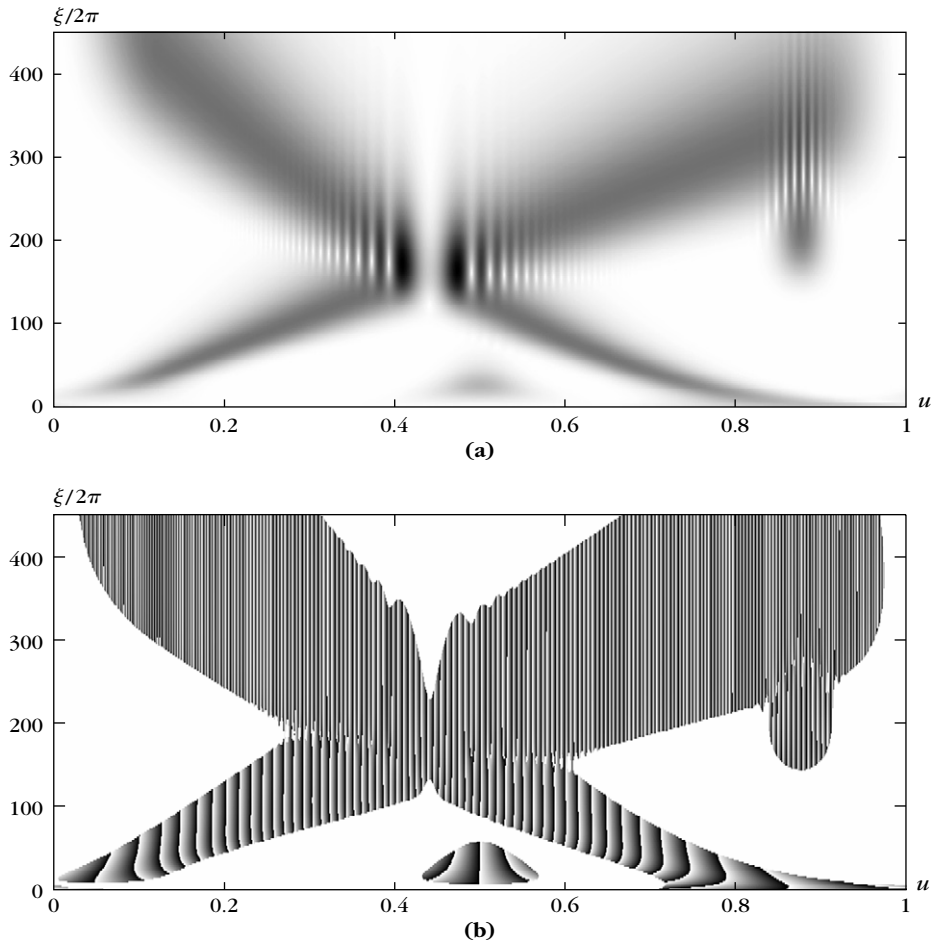

FIGURE 4.11

(a) Normalized scalogram $\eta^{-1}\xi P_W f(u,\xi)$ computed from the signal in Figure 4.3; dark points indicate large-amplitude coefficients. **(b)** Complex phase $\Theta_W(u,\xi)$ of $Wf(u,\eta/\xi)$, where the modulus is nonzero.

Its discrete wavelet transform can only be calculated at scales $N^{-1} < s < 1$, as shown in Figure 4.7. It is calculated for $s = a^j$, with $a = 2^{1/v}$, which provides v intermediate scales in each octave $[2^j, 2^{j+1})$.

Let $\psi(t)$ be a wavelet with a support included in $[-K/2, K/2]$. For $1 \leqslant a^j \leqslant N K^{-1}$, a discrete wavelet scaled by a^j is defined by

$$\psi_j[n] = \frac{1}{\sqrt{a^j}}\,\psi\left(\frac{n}{a^j}\right).$$

This discrete wavelet has Ka^j nonzero values on $[-N/2, N/2]$. The scale a^j is larger than 1; otherwise, the sampling interval may be larger than the wavelet support.

Fast Transform

To avoid border problems, we treat $f[n]$ and the wavelets $\psi_j[n]$ as periodic signals of period N. The discrete wavelet transform can then be written as a circular convolution with $\bar{\psi}_j[n] = \psi_j^*[-n]$:

$$Wf[n, a^j] = \sum_{m=0}^{N-1} f[m]\, \psi_j^*[m-n] = f \circledast \bar{\psi}_j[n]. \tag{4.64}$$

This circular convolution is calculated with the fast Fourier transform algorithm, which requires $O(N \log_2 N)$ operations. If $a = 2^{1/v}$, there are $v \log_2(N/(2K))$ scales $a^j \in [2N^{-1}, K^{-1}]$. The total number of operations to compute the wavelet transform over all scales therefore is $O(vN(\log_2 N)^2)$ [408].

To compute the scalogram $P_W[n, \xi] = |Wf[n, \frac{\eta}{\xi}]|^2$, we calculate $Wf[n, s]$ at any scale s with a parabola interpolation. Let j be the closest integer to $\log_2 s / \log_2 a$, and $p(x)$ be the parabola such that

$$p(j-1) = Wf[n, a^{j-1}], \quad p(j) = Wf[n, a^j], \quad p(j+1) = Wf[n, a^{j+1}].$$

A second-order interpolation computes

$$Wf[n, s] = p\left(\frac{\log_2 s}{\log_2 a}\right).$$

Parabolic interpolations are used instead of linear interpolations in order to more precisely locate the ridges defined in Section 4.4.3.

Discrete Scaling Filter

A wavelet transform computed up to a scale a^j is not a complete signal representation. It is necessary to add low frequencies $Lf[n, a^j]$ corresponding to scales larger than a^j. A discrete and periodic scaling filter is computed by sampling the scaling function $\phi(t)$ defined in (4.42):

$$\phi_J[n] = \frac{1}{\sqrt{a^j}}\, \phi\left(\frac{n}{a^j}\right) \quad \text{for} \quad n \in [-N/2, N/2].$$

Let $\bar{\phi}_J[n] = \phi_J^*[-n]$; the low frequencies are carried by

$$Lf[n, a^j] = \sum_{m=0}^{N-1} f[m]\, \phi_J^*[m-n] = f \circledast \bar{\phi}_J[n]. \tag{4.65}$$

Reconstruction

An approximate inverse wavelet transform is implemented by discretizing the integral (4.45). Suppose that $a^I = 1$ is the finest scale. Since $ds/s^2 = d\log_e s/s$ and the discrete wavelet transform is computed along an exponential scale sequence $\{a^j\}_j$ with a logarithmic increment $d\log_e s = \log_e a$, we obtain

$$f[n] \approx \frac{\log_e a}{C_\psi} \sum_{j=I}^{J} \frac{1}{a^j} Wf[., a^j] \circledast \psi_j[n] + \frac{1}{C_\psi a^j} Lf[., a^j] \circledast \phi_J[n]. \tag{4.66}$$

The "." indicates the variable over which the convolution is calculated. These circular convolutions are calculated using the FFT, with $O(vN(\log_2 N)^2)$ operations.

Analytic wavelet transforms are often computed over real signals $f[n]$ that have no energy at low frequencies. The scaling filter component is then negligible. Theorem 4.5 shows that

$$f[n] \approx \frac{2\log_e a}{C_\psi} \operatorname{Re}\left(\sum_{j=I}^{J} \frac{1}{a^j} Wf[., a^j] \circledast \psi_j[n] \right). \tag{4.67}$$

The error introduced by the discretization of scales decreases when the number v of voices per octave increases. However, the approximation of continuous time convolutions with discrete convolutions also creates high-frequency errors. Perfect reconstructions are obtained with a more careful design of the reconstruction filters (Exercise 4.3). Section 5.2.2 describes an exact inverse wavelet transform computed at dyadic scales $a^j = 2^j$.

4.4 TIME-FREQUENCY GEOMETRY OF INSTANTANEOUS FREQUENCIES

When listening to music, we perceive several frequencies that change with time. In music, it is associated to the geometric perception of "movements." This notion of instantaneous frequency remains to be defined. The time variation of several instantaneous frequencies is measured with local maxima of windowed Fourier transforms and wavelet transforms. They define a geometric time-frequency support from which signal approximations are recovered. Audio processing is implemented by modifying this time-frequency support.

4.4.1 Analytic Instantaneous Frequency

The notion of instantaneous frequency is not well defined. It can, however, be uniquely specified with the signal analytic part. A cosine modulation

$$f(t) = a\cos(\omega_0 t + \theta_0) = a\cos\theta(t)$$

has a frequency ω_0 that is the derivative of the phase $\theta(t) = \omega_0 t + \theta_0$. To generalize this notion, real signals f are written as an amplitude $a(t)$ modulated with a time-varying phase $\theta(t)$:

$$f(t) = a(t)\ \cos\theta(t), \quad \text{with} \quad a(t) \geqslant 0. \tag{4.68}$$

The *instantaneous frequency* can be defined as a positive derivative of the phase:

$$\omega(t) = \theta'(t) \geqslant 0.$$

The derivative is chosen to be positive by adapting the sign of $\theta(t)$. However, for a given $f(t)$, there are many possible choices of $a(t)$ and $\theta(t)$ to satisfy (4.68), so $\omega(t)$ is not uniquely defined relative to f.

A particular decomposition (4.68) is obtained from the analytic part f_a of f, which has a Fourier transform defined in (4.47) by

$$\hat{f}_a(\omega) = \begin{cases} 2\hat{f}(\omega) & \text{if } \omega \geqslant 0 \\ 0 & \text{if } \omega < 0 \end{cases} \tag{4.69}$$

This complex signal is represented by separating the modulus and the complex phase:

$$f_a(t) = a(t)\ \exp[i\theta(t)]. \tag{4.70}$$

Since $f = \mathrm{Re}[f_a]$, it follows that

$$f(t) = a(t)\ \cos\theta(t).$$

We call $a(t)$ the *analytic* amplitude of $f(t)$ and $\theta'(t)$ its *instantaneous frequency*; they are uniquely defined.

EXAMPLE 4.12

If $f(t) = a(t)\ \cos(\omega_0 t + \theta_0)$, then

$$\hat{f}(\omega) = \frac{1}{2}\Big(\exp(i\theta_0)\,\hat{a}(\omega - \omega_0) + \exp(-i\theta_0)\,\hat{a}(\omega + \omega_0)\Big).$$

If the variations of $a(t)$ are slow compared to the period $2\pi/\omega_0$, which is achieved by requiring that the support of $\hat{a}$ be included in $[-\omega_0, \omega_0]$, then

$$\hat{f}_a(\omega) = \hat{a}(\omega - \omega_0)\ \exp(i\theta_0),$$

so $f_a(t) = a(t)\ \exp[i(\omega_0 t + \theta_0)]$.

If a signal f is the sum of two sinusoidal waves:

$$f(t) = a\cos(\omega_1 t) + a\cos(\omega_2 t),$$

then

$$f_a(t) = a\ \exp(i\omega_1 t) + a\ \exp(i\omega_2 t) = 2a\ \cos\left(\frac{1}{2}\,(\omega_1 - \omega_2)\,t\right)\ \exp\left(\frac{i}{2}\,(\omega_1 + \omega_2)\,t\right).$$

The instantaneous frequency is $\theta'(t) = (\omega_1 + \omega_2)/2$ and the amplitude is

$$a(t) = 2a \left| \cos\left(\frac{1}{2}(\omega_1 - \omega_2)\, t\right) \right|.$$

This result is not satisfying because it does not reveal that the signal includes two sinusoidal waves of the same amplitude; it measures an average frequency value. The next sections explain how to measure the instantaneous frequencies of several spectral components by separating them with a windowed Fourier transform or a wavelet transform. We first describe two important applications of instantaneous frequencies.

Frequency Modulation

In signal communications, information can be transmitted through the amplitude $a(t)$ (amplitude modulation) or the instantaneous frequency $\theta'(t)$ (frequency modulation) [60]. Frequency modulation is more robust in the presence of additive Gaussian white noise. In addition, it better resists multipath interferences, which destroy the amplitude information. A frequency modulation sends a message $m(t)$ through a signal

$$f(t) = a \cos\theta(t) \quad \text{with} \quad \theta'(t) = \omega_0 + k\, m(t).$$

The frequency bandwidth of f is proportional to k. This constant is adjusted depending on transmission noise and available bandwidth. At reception, the message $m(t)$ is restored with a frequency demodulation that computes the instantaneous frequency $\theta'(t)$ [120].

Additive Sound Models

Musical sounds and voiced speech segments can be modeled with sums of sinusoidal *partials*:

$$f(t) = \sum_{k=1}^{K} f_k(t) = \sum_{k=1}^{K} a_k(t) \cos\theta_k(t), \tag{4.71}$$

where a_k and θ_k' vary slowly [413, 414]. Such decompositions are useful for pattern recognition and for modifying sound properties [339]. Sections 4.4.2 and 4.4.3 explain how to compute a_k and the instantaneous frequency θ_k' and reconstruct signals from this information.

Reducing or increasing the duration of a sound f by a factor α in time is used for radio broadcasting to adjust recorded sequences to a precise time schedule. A scaling $f(\alpha t)$ transforms each $\theta_k(t)$ in $\theta_k(\alpha t)$ and therefore $\theta'_k(t)$ in $\alpha\theta'_k(t)$. For sound reduction, with $\alpha > 1$ all frequencies are thus increased. To avoid modifying the values of θ_k' and a_k, a new sound is synthesized:

$$f_\alpha(t) = \sum_{k=1}^{K} a_k(\alpha\, t)\, \cos\Big(\frac{1}{\alpha}\, \theta_k(\alpha\, t)\Big). \tag{4.72}$$

The partials of f_α at $t = t_0/\alpha$ and the partials of f at $t = t_0$ have the same amplitudes and the same instantaneous frequencies, therefore the properties of these sounds are perceived as identical.

A frequency transposition with the same duration is calculated by dividing each phase by a constant α in order to shift the sound harmonics:

$$f_\alpha(t) = \sum_{k=1}^{K} b_k(t)\ \cos\Big(\theta_k(t)/\alpha\Big). \tag{4.73}$$

The instantaneous frequency of each partial is now $\theta_k{}'(t)/\alpha$. To maintain the sound properties, the amplitudes $b_k(t)$ must be adjusted so as not to modify the global frequency envelope $F(t, \omega)$ of the harmonics:

$$a_k(t) = F\Big(t, \theta_k{}'(t)\Big) \quad \text{and} \quad b_k(t) = F\Big(t, \theta_k{}'(t)/\alpha\Big). \tag{4.74}$$

Many types of sounds—musical instruments or speech—are produced by an excitation that propagates across a wave guide. Locally, $F(t, \omega)$ is the transfer function of the wave guide. In speech processing, it is called a *formant*. This transfer function is often approximated with an autoregressive filter of order M, in which case:

$$F(t, \omega) = \frac{C}{\sum_{m=0}^{M-1} c_m\, \mathrm{e}^{-im\omega}}. \tag{4.75}$$

The parameters c_m are identified with (4.74) from the a_k and the b_k are then derived with (4.74) and (4.75).

4.4.2 Windowed Fourier Ridges

The spectrogram $P_S f(u, \xi) = |Sf(u, \xi)|^2$ measures the energy of f in a time-frequency neighborhood of (u, ξ). The ridge algorithm computes the signal instantaneous frequencies and amplitudes from the local maxima of $P_S f(u, \xi)$. These local maxima define a geometric support in the time-frequency plane. Modifications of sound durations or frequency transpositions are computed with time or frequency dilations of the ridge support.

Time-frequency ridges were introduced by Delprat, Escudié, Guillemain, Kronland-Martinet, Tchamitchian, and Torrésani [66, 204] to analyze musical sounds and are used to represent time-varying frequency tones for a wide range of signals [66, 289].

The windowed Fourier transform is computed with a symmetric window $g(t) = g(-t)$ that has a support equal to $[-1/2, 1/2]$. The Fourier transform $\hat{g}$ is a real symmetric function. We suppose that $|\hat{g}(\omega)| \leqslant \hat{g}(0)$ for all $\omega \in \mathbb{R}$, and that $\hat{g}(0) = \int_{-1/2}^{1/2} g(t)\, dt$ is on the order of 1. Table 4.1 listed several examples of such windows. The window g is normalized so that $\|g\| = 1$. For a fixed scale $s, g_s(t) = s^{-1/2} g(t/s)$

has a support of size s and a unit norm. The corresponding windowed Fourier atoms are

$$g_{s,u,\xi}(t) = g_s(t-u)\, e^{i\xi t},$$

and the resulting windowed Fourier transform is

$$Sf(u,\xi) = \langle f, g_{s,u,\xi}\rangle = \int_{-\infty}^{+\infty} f(t)\, g_s(t-u)\, e^{-i\xi t}\, dt. \tag{4.76}$$

Theorem 4.6 relates $Sf(u,\xi)$ to the instantaneous frequency of f.

Theorem 4.6. Let $f(t) = a(t)\, \cos\theta(t)$. If $\xi \geqslant 0$, then

$$\langle f, g_{s,u,\xi}\rangle = \frac{\sqrt{s}}{2}\, a(u)\, \exp(i[\theta(u) - \xi u])\, \Big(\hat{g}(s[\xi - \theta'(u)]) + \varepsilon(u,\xi)\Big). \tag{4.77}$$

The corrective term satisfies

$$|\varepsilon(u,\xi)| \leqslant \varepsilon_{a,1} + \varepsilon_{a,2} + \varepsilon_{\theta,2} + \sup_{|\omega| \geqslant s\theta'(u)} |\hat{g}(\omega)| \tag{4.78}$$

with

$$\varepsilon_{a,1} \leqslant \frac{s\,|a'(u)|}{|a(u)|}, \quad \varepsilon_{a,2} \leqslant \sup_{|t-u| \leqslant s/2} \frac{s^2\,|a''(t)|}{|a(u)|}, \tag{4.79}$$

and if $s\,|a'(u)|\,|a(u)|^{-1} \leqslant 1$, then

$$\varepsilon_{\theta,2} \leqslant \sup_{|t-u| \leqslant s/2} s^2 |\theta''(t)|. \tag{4.80}$$

If $\xi = \theta'(u)$, then

$$\varepsilon_{a,1} = \frac{s\,|a'(u)|}{|a(u)|}\, \left|\hat{g}'\Big(2\, s\, \theta'(u)\Big)\right|. \tag{4.81}$$

Proof. Observe that

$$\begin{aligned}
\langle f, g_{s,u,\xi}\rangle &= \int_{-\infty}^{+\infty} a(t)\, \cos\theta(t)\, g_s(t-u)\, \exp(-i\xi t)\, dt \\
&= \frac{1}{2}\int_{-\infty}^{+\infty} a(t)\, \Big(\exp[i\theta(t)] + \exp[-i\theta(t)]\Big)\, g_s(t-u)\, \exp[-i\xi t]\, dt \\
&= I(\theta) + I(-\theta).
\end{aligned}$$

We first concentrate on

$$\begin{aligned}
I(\theta) &= \frac{1}{2}\int_{-\infty}^{+\infty} a(t)\, \exp[i\theta(t)]\, g_s(t-u)\, \exp(-i\xi t)\, dt \\
&= \frac{1}{2}\int_{-\infty}^{+\infty} a(t+u)\, e^{i\theta(t+u)}\, g_s(t)\, \exp[-i\xi(t+u)]\, dt.
\end{aligned}$$

This integral is computed by using second-order Taylor expansions:

$$a(t+u) = a(u) + t\,a'(u) + \frac{t^2}{2}\,\alpha(t) \;\text{ with }\; |\alpha(t)| \leqslant \sup_{h\in[u,t+u]} |a''(h)|$$
$$\theta(t+u) = \theta(u) + t\,\theta'(u) + \frac{t^2}{2}\,\beta(t) \;\text{ with }\; |\beta(t)| \leqslant \sup_{h\in[u,t+u]} |\theta''(h)|.$$

We get

$$\begin{aligned}
&2\;\exp\Big(-i(\theta(u)-\xi u)\Big)\,I(\theta)\\
&= \int_{-\infty}^{+\infty} a(u)\,g_s(t)\;\exp\Big(-it(\xi-\theta'(u))\Big)\;\exp\Big(i\frac{t^2}{2}\beta(t)\Big)\,dt\\
&\quad + \int_{-\infty}^{+\infty} a'(u)\,t\,g_s(t)\;\exp\Big(-it(\xi-\theta'(u))\Big)\;\exp\Big(i\frac{t^2}{2}\beta(t)\Big)\,dt\\
&\quad + \frac{1}{2}\int_{-\infty}^{+\infty} \alpha(t)\,t^2\,g_s(t)\;\exp\Big(-i(t\xi+\theta(u)-\theta(t+u))\Big)\,dt.
\end{aligned}$$

A first-order Taylor expansion of $\exp(ix)$ gives

$$\exp\Big(i\frac{t^2}{2}\beta(t)\Big) = 1 + \frac{t^2}{2}\beta(t)\,\gamma(t),\;\text{ with } |\gamma(t)| \leqslant 1. \tag{4.82}$$

Since

$$\int_{-\infty}^{+\infty} g_s(t)\;\exp[-it(\xi-\theta'(u))]\,dt = \sqrt{s}\,\hat{g}(s[\xi-\theta'(u)])\,,$$

inserting (4.82) in the expression of $I(\theta)$ yields

$$\begin{aligned}
&\left|I(\theta) - \frac{\sqrt{s}}{2}a(u)\;\exp[i(\theta(u)-\xi u)]\,\hat{g}(\xi-\theta'(u))\right|\\
&\qquad \leqslant \frac{\sqrt{s}\,|a(u)|}{4}(\varepsilon^{+}_{a,1} + \varepsilon_{a,2} + \varepsilon_{\theta,2}),
\end{aligned} \tag{4.83}$$

with

$$\varepsilon^{+}_{a,1} = \frac{2|a'(u)|}{|a(u)|}\left|\int_{-\infty}^{+\infty} t\,\frac{1}{\sqrt{s}}\,g_s(t)\;\exp[-it(\xi-\theta'(u))]\,dt\right|, \tag{4.84}$$

$$\varepsilon_{a,2} = \int_{-\infty}^{+\infty} t^2\,|\alpha(t)|\,\frac{1}{\sqrt{s}}\,|g_s(t)|\,dt, \tag{4.85}$$

$$\begin{aligned}
\varepsilon_{\theta,2} &= \int_{-\infty}^{+\infty} t^2\,|\beta(t)|\,\frac{1}{\sqrt{s}}\,|g_s(t)|\,dt\\
&\quad + \frac{|a'(u)|}{|a(u)|}\int_{-\infty}^{+\infty} |t^3|\,|\beta(t)|\,\frac{1}{\sqrt{s}}\,|g_s(t)|dt.
\end{aligned} \tag{4.86}$$

Applying (4.83) to $I(-\theta)$ gives

$$|I(-\theta)| \leqslant \frac{\sqrt{s}\,|a(u)|}{2}\,|\hat{g}(\xi+\theta'(u))| + \frac{\sqrt{s}\,|a(u)|}{4}(\varepsilon^{-}_{a,1} + \varepsilon_{a,2} + \varepsilon_{\theta,2}),$$

with

$$\varepsilon_{a,1}^{-} = \frac{2|a'(u)|}{|a(u)|} \left| \int_{-\infty}^{+\infty} t \, \frac{1}{\sqrt{s}} \, g_s(t) \, \exp[-it(\xi + \theta'(u))] \, dt \right|. \tag{4.87}$$

Since $\xi \geqslant 0$ and $\theta'(u) \geqslant 0$, we derive that

$$|\hat{g}(s[\xi + \theta'(u)])| \leqslant \sup_{|\omega| \geqslant s\theta'(u)} |\hat{g}(\omega)|;$$

therefore

$$I(\theta) + I(-\theta) = \frac{\sqrt{s}}{2} \, a(u) \, \exp[i(\theta(u) - \xi u)] \Big(\hat{g}\big(s[\xi - \theta'(u)]\big) + \varepsilon(u, \xi) \Big),$$

with

$$\varepsilon(u, \xi) = \frac{\varepsilon_{a,1}^{+} + \varepsilon_{a,1}^{-}}{2} + \varepsilon_{a,2} + \varepsilon_{\theta,2} + \sup_{|\omega| \geqslant s|\theta'(u)|} |\hat{g}(\omega)|.$$

Let us now verify the upper bound (4.79) for $\varepsilon_{a,1} = (\varepsilon_{a,1}^{+} + \varepsilon_{a,1}^{-})/2$. Since $g_s(t) = s^{-1/2} g(t/s)$, a simple calculation shows that for $n \geqslant 0$

$$\int_{-\infty}^{+\infty} |t|^n \, \frac{1}{\sqrt{s}} \, |g_s(t)| \, dt = s^n \int_{-1/2}^{1/2} |t|^n \, |g(t)| \, dt \leqslant \frac{s^n}{2^n} \, \|g\|^2 = \frac{s^n}{2^n}. \tag{4.88}$$

Inserting this for $n = 1$ in (4.84) and (4.87) gives

$$\varepsilon_{a,1} = \frac{\varepsilon_{a,1}^{+} + \varepsilon_{a,1}^{-}}{2} \leqslant \frac{s \, |a'(u)|}{|a(u)|}.$$

The upper bounds (4.79) and (4.80) of the second-order terms $\varepsilon_{a,2}$ and $\varepsilon_{\theta,2}$ are obtained by observing that the remainder $\alpha(t)$ and $\beta(t)$ of the Taylor expansion of $a(t+u)$ and $\theta(t+u)$ satisfy

$$\sup_{|t| \leqslant s/2} |\alpha(t)| \leqslant \sup_{|t-u| \leqslant s/2} |a''(t)|, \quad \sup_{|t| \leqslant s/2} |\beta(t)| \leqslant \sup_{|t-u| \leqslant s/2} |\theta''(t)|. \tag{4.89}$$

Inserting this in (4.85) yields

$$\varepsilon_{a,2} \leqslant \sup_{|t-u| \leqslant s/2} \frac{s^2 \, |a''(t)|}{|a(u)|}.$$

When $s \, |a'(u)| |a(u)|^{-1} \leqslant 1$, replacing $|\beta(t)|$ by its upper bound in (4.86) gives

$$\varepsilon_{\theta,2} \leqslant \frac{1}{2} \Big(1 + \frac{s \, |a'(u)|}{|a(u)|} \Big) \sup_{|t-u| \leqslant s/2} s^2 |\theta''(t)| \leqslant \sup_{|t-u| \leqslant s/2} s^2 |\theta''(t)|.$$

Let us finally compute ε_a when $\xi = \theta'(u)$. Since $g(t) = g(-t)$, we derive from (4.84) that

$$\varepsilon_{a,1}^{+} = \frac{2|a'(u)|}{|a(u)|} \left| \int_{-\infty}^{+\infty} t \, \frac{1}{\sqrt{s}} \, g_s(t) \, dt \right| = 0.$$

We also derive from (2.22) that the Fourier transform of $t\frac{1}{\sqrt{s}}g_s(t)$ is $i\,s\,\hat{g}'(s\omega)$, so (4.87) gives

$$\varepsilon_a = \frac{1}{2}\varepsilon_{a,1}^{-} = \frac{s|a'(u)|}{|a(u)|}\,|\hat{g}'(2s\theta'(u))|. \qquad \blacksquare$$

Delprat et al. [204] give a different proof of a similar result when $g(t)$ is a Gaussian, using a stationary phase approximation. If we can neglect the corrective term $\varepsilon(u,\xi)$, we will see that (4.77) enables us to measure $a(u)$ and $\theta'(u)$ from $Sf(u,\xi)$. This implies that the decomposition $f(t) = a(t)\cos\theta(t)$ is uniquely defined. By reviewing the proof of Theorem 4.6, one can verify that a and θ' are the analytic amplitude and instantaneous frequencies of f.

The expressions (4.79, 4.80) show that the three corrective terms $\varepsilon_{a,1}, \varepsilon_{a,2}$, and $\varepsilon_{\theta,2}$ are small if $a(t)$ and $\theta'(t)$ have small relative variations over the support of window g_s. Let $\Delta\omega$ be the bandwidth of $\hat{g}$ defined by

$$|\hat{g}(\omega)| \ll 1 \quad \text{for } |\omega| \geqslant \Delta\omega. \tag{4.90}$$

The term

$$\sup_{|\omega| \geqslant s|\theta'(u)|} |\hat{g}(\omega)| \text{ of } \varepsilon(u,\xi)$$

is negligible if

$$\theta'(u) \geqslant \frac{\Delta\omega}{s}.$$

Ridge Points

Let us suppose that $a(t)$ and $\theta'(t)$ have small variations over intervals of size s and that $\theta'(t) \geqslant \Delta\omega/s$ so that the corrective term $\varepsilon(u,\xi)$ in (4.77) can be neglected. Since $|\hat{g}(\omega)|$ is maximum at $\omega = 0$, (4.77) shows that for each u the spectrogram $|Sf(u,\xi)|^2 = |\langle f, g_{s,u,\xi}\rangle|^2$ is maximum at $\xi(u) = \theta'(u)$. The corresponding time-frequency points $(u, \xi(u))$ are called *ridges*. At ridge points, (4.77) becomes

$$Sf(u,\xi) = \frac{\sqrt{s}}{2}\,a(u)\,\exp(i[\theta(u) - \xi u])\left(\hat{g}(0) + \varepsilon(u,\xi)\right). \tag{4.91}$$

Theorem 4.6 proves that the $\varepsilon(u,\xi)$ is smaller at a ridge point because the first-order term $\varepsilon_{a,1}$ becomes negligible in (4.81). This is shown by verifying that $|\hat{g}'(2s\theta'(u))|$ is negligible when $s\theta'(u) \geqslant \Delta\omega$. At ridge points, the second-order terms $\varepsilon_{a,2}$ and $\varepsilon_{\theta,2}$ are predominant in $\varepsilon(u,\xi)$.

The ridge frequency gives the instantaneous frequency $\xi(u) = \theta'(u)$ and the amplitude is calculated by

$$a(u) = \frac{2\,|Sf(u,\xi(u))|}{\sqrt{s}\,|\hat{g}(0)|}. \tag{4.92}$$

Let $\Theta_S(u,\xi)$ be the complex phase of $Sf(u,\xi)$. If we neglect the corrective term, then (4.91) proves that ridges are also stationary phase points:

$$\frac{\partial\Theta_S(u,\xi)}{\partial u} = \theta'(u) - \xi = 0.$$

Testing the stationarity of the phase locates the ridges more precisely.

Multiple Frequencies

When the signal contains several spectral lines having frequencies sufficiently apart, the windowed Fourier transform separates each of these components and the ridges detect the evolution in time of each spectral component. Let us consider

$$f(t) = a_1(t)\cos\theta_1(t) + a_2(t)\cos\theta_2(t),$$

where $a_k(t)$ and $\theta_k'(t)$ have small variations over intervals of size s and $s\theta_k'(t) \geqslant \Delta\omega$. Since the windowed Fourier transform is linear, we apply (4.77) to each spectral component and neglect the corrective terms:

$$\begin{aligned} Sf(u,\xi) = {} & \frac{\sqrt{s}}{2}\, a_1(u)\, \hat{g}(s[\xi - \theta_1'(u)])\, \exp(i[\theta_1(u) - \xi u]) \\ & + \frac{\sqrt{s}}{2}\, a_2(u)\, \hat{g}(s[\xi - \theta_2'(u)])\, \exp(i[\theta_2(u) - \xi u]). \end{aligned} \tag{4.93}$$

The two spectral components are discriminated if for all u

$$\hat{g}(s|\theta_1'(u) - \theta_2'(u)|) \ll 1, \tag{4.94}$$

which means that the frequency difference is larger than the bandwidth of $\hat{g}(s\omega)$:

$$|\theta_1'(u) - \theta_2'(u)| \geqslant \frac{\Delta\omega}{s}. \tag{4.95}$$

In this case, when $\xi = \theta_1'(u)$, the second term of (4.93) can be neglected and the first term generates a ridge point from which we may recover $\theta'_1(u)$ and $a_1(u)$ by using (4.92). Similarly, if $\xi = \theta_2'(u)$, the first term can be neglected and we have a second ridge point that characterizes $\theta_2'(u)$ and $a_2(u)$. The ridge points are distributed along two time-frequency lines, $\xi(u) = \theta_1'(u)$ and $\xi(u) = \theta_2'(u)$. This result is valid for any number of time-varying spectral components, as long as the distance between any two instantaneous frequencies satisfies (4.95). If two spectral lines are too close, they interfere, thus destroying the ridge pattern.

Time-Frequency Ridge Support

The number of instantaneous frequencies is typically unknown. The ridge support Λ therefore is defined as the set of all (u, ξ)—the local maxima of $|Sf(u,\xi)|^2$ for u fixed and ξ varying and points of stationary phase $\partial\Theta_S(u,\xi)/\partial u \approx 0$. This support is often reduced by removing small-ridge amplitudes $|Sf(u,\xi)|$ that are mostly dominated by the noise, or because smaller ridges may be "shadows" of other instantaneous frequencies created by the side lobes of $\hat{g}(\omega)$.

Let $\{g_{s,u,\xi}\}_{(u,\xi)\in\Lambda}$ be the set of ridge atoms. For discrete signals, there is a finite number of ridge points, that define a frame of the space $\mathbf{V}_\Lambda$ they generate. A ridge signal approximation is computed as an orthogonal projection of f on $\mathbf{V}_\Lambda$. Section 5.1.3 shows that it is obtained with the dual frame $\{\tilde{g}_{\Lambda,u,\xi}\}_{(u,\xi)\in\Lambda}$ of $\{g_{s,u,\xi}\}_{(u,\xi)\in\Lambda}$ in $\mathbf{V}_\Lambda$:

$$f_\Lambda = \sum_{(u,\xi)\in\Lambda} Sf(u,\xi)\, \tilde{g}_{\Lambda,u,\xi}. \tag{4.96}$$

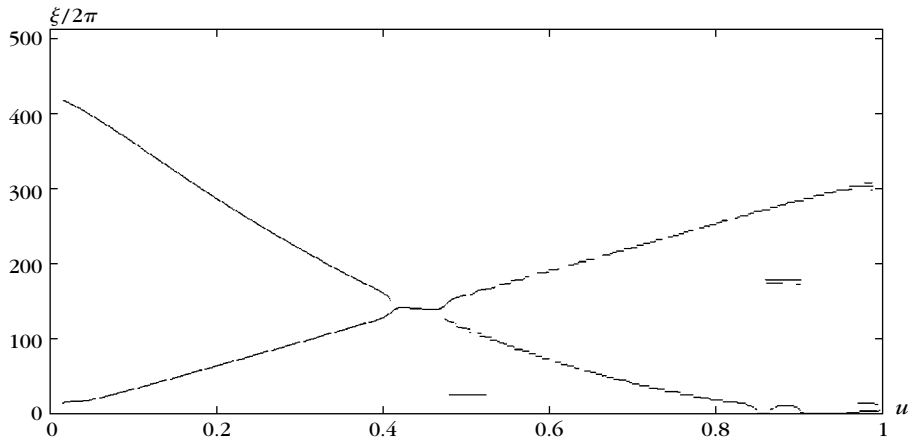

FIGURE 4.12

Support of larger-amplitude ridges calculated from the spectrogram in Figure 4.3. These ridges give the instantaneous frequencies of the linear and quadratic chirps and of low- and high-frequency transients at $t = 0.5$ and $t = 0.87$.

The dual-synthesis algorithm of Section 5.1.3 computes this orthogonal projection by inverting the symmetric operator

$$Lh = \sum_{(u,\xi)\in\Lambda} \langle h, g_{s,u,\xi}\rangle\, g_{s,u,\xi}. \tag{4.97}$$

The inversion requires iteration of this operator many times. If there are only a few ridge points, then (4.97) is efficiently computed by evaluating the inner product and the sum for just $(u, \xi) \in \Lambda$. If there are many ridge points, it is more efficient to compute the full windowed Fourier transform $Sh(u, \xi) = \langle h, g_{s,u,\xi}\rangle$ with the FFT algorithm (described in Section 4.2.3); set all coefficients to zero for $(u, \xi) \notin \Lambda$ and apply the fast inverse windowed Fourier transform over all coefficients. The normalization factor N^{-1} in (4.28) must be removed (set to 1) to implement (4.97).

Figure 4.12 displays the ridge support computed from the modulus and phase of the windowed Fourier transform shown in Figure 4.3. For $t \in [0.4, 0.5]$, the instantaneous frequencies of the linear chirp and the quadratic chirps are close, the frequency resolution of the window is not sufficient to discriminate them. As a result, the ridges detect a single average instantaneous frequency.

Time-Scaling and Frequency Transpositions

A reduction of sound duration by a factor α is implemented, according to the deformation model (4.72), by dilating the ridge support Λ in time:

$$\Lambda_\alpha = \{(u, \xi) \,:\, (\alpha u, \xi) \in \Lambda\}. \tag{4.98}$$

The windowed Fourier coefficients $c(u,\xi)$ in Λ_α are derived from the modulus and phase of ridge coefficients

$$\forall (v,\xi) \in \Lambda_\alpha, \quad c(v,\xi) = |Sf(\alpha v,\xi)|\, e^{i\Theta_S(\alpha v,\xi)/\alpha}. \tag{4.99}$$

The scaled signal is reconstructed from these coefficients, with the dual-synthesis algorithm of Section 5.1.3, as in (4.96):

$$f_\alpha = \sum_{(v,\xi)\in\Lambda_\alpha} c(v,\xi)\, \tilde{g}_{\Lambda_\alpha,v,\xi}.$$

Similarly, a sound transposition is implemented according to the transposition model (4.73) by dilating the ridge support Λ in frequency:

$$\Lambda_\alpha = \{(u,\xi) \ : \ (u,\alpha\xi) \in \Lambda\}. \tag{4.100}$$

The transposed coefficient amplitudes $|c(u,\xi)|$ in Λ_α are calculated with (4.74). At any fixed time u_0, the ridge amplitudes at all frequencies $\{a(u_0,\xi) = |Sf(u_0,\xi)\}_{(u_0,\xi)\in\Lambda}$ are mapped to transposed amplitudes $\{b(u_0,\eta)\}_{(u_0,\eta)\in\Lambda_\alpha}$ at frequencies $\eta = \xi/\alpha$ by computing a frequency envelope. The resulting ridge coefficients are

$$\forall (u,\eta) \in \Lambda_\alpha, \quad c(u,\eta) = b(u,\eta)\, e^{i\Theta_S(u,\alpha\eta)/\alpha}. \tag{4.101}$$

The transposed signal is reconstructed with the dual-synthesis algorithm of Section 5.1.3:

$$f_\alpha = \sum_{(u,\eta)\in\Lambda_\alpha} c(u,\eta)\, \tilde{g}_{\Lambda_\alpha,u,\eta}.$$

Choice of Window

The measurement of instantaneous frequencies at ridge points is valid only if the size s of the window g_s is sufficiently small so that the second-order terms $\varepsilon_{a,2}$ and $\varepsilon_{\theta,2}$ in (4.79) and (4.80) are small:

$$\sup_{|t-u|\leqslant s/2} \frac{s^2\,|a_k''(t)|}{|a_k(u)|} \ll 1 \ \text{ and } \ \sup_{|t-u|\leqslant s/2} s^2 |\theta_k''(t)| \ll 1. \tag{4.102}$$

On the other hand, the frequency bandwidth $\Delta\omega/s$ must also be sufficiently small to discriminate consecutive spectral components in (4.95). The window scale s therefore must be adjusted as a trade-off between both constraints.

Table 4.1 listed the spectral parameters of several windows of compact support. For instantaneous frequency detection, it is particularly important to ensure that $\hat{g}$ has negligible side lobes at $\pm\omega_0$, as illustrated by Figure 4.4. The reader can verify with (4.77) that these side lobes "react" to an instantaneous frequency $\theta'(u)$ by creating shadow maxima of $|Sf(u,\xi)|^2$ at frequencies $\xi = \theta'(u) \pm \omega_0$. The ratio of the amplitude of these shadow maxima to the amplitude of the main local maxima at $\xi = \theta'(u)$ is $|\hat{g}(\omega_0)|^2\, |\hat{g}(0)|^{-2}$. They can be removed by thresholding or by testing the stationarity of the phase.

EXAMPLE 4.13

The sum of two parallel linear chirps

$$f(t) = a_1 \cos(bt^2 + ct) + a_2 \cos(bt^2) \tag{4.103}$$

has two instantaneous frequencies $\theta_1'(t) = 2bt + c$ and $\theta_2'(t) = 2bt$. Figure 4.13 gives a numerical example. The window g_s has enough frequency resolution to discriminate both chirps if

$$|\theta_1'(t) - \theta_2'(t)| = |c| \geqslant \frac{\Delta\omega}{s}. \tag{4.104}$$

Its time support is small enough if

$$s^2 \, |\theta_1''(u)| = s^2 \, |\theta_2''(u)| = 2\, b\, s^2 \ll 1. \tag{4.105}$$

Conditions (4.104) and (4.105) prove that there exists an appropriate window g, if and only if,

$$\frac{c}{\sqrt{b}} \gg \Delta\omega. \tag{4.106}$$

Since g is a smooth window with a support $[-1/2, 1/2]$, its frequency bandwidth $\Delta\omega$ is on the order of 1. The linear chirps in Figure 4.13 satisfy (4.106). Ridges are computed with the truncated Gaussian window of Table 4.1, with $s = 0.5$.

EXAMPLE 4.14

The hyperbolic chirp

$$f(t) = \cos\left(\frac{\alpha}{\beta - t}\right)$$

for $0 \leqslant t < \beta$ has an instantaneous frequency

$$\theta'(t) = \frac{\alpha}{(\beta - t)^2},$$

which varies quickly when t is close to β. The instantaneous frequency of hyperbolic chirps goes from 0 to $+\infty$ in a finite time interval. This is particularly useful for radars. Such chirps are also emitted by the cruise sonars of bats [204].

The instantaneous frequency of hyperbolic chirps cannot be estimated with a windowed Fourier transform because for any fixed window size instantaneous frequency varies too quickly at high frequencies. When u is close enough to β, then (4.102) is not satisfied because

$$s^2 |\theta''(u)| = \frac{s^2 \alpha}{(\beta - u)^3} > 1.$$

Figure 4.14 shows a signal that is a sum of two hyperbolic chirps:

$$f(t) = a_1 \cos\left(\frac{\alpha_1}{\beta_1 - t}\right) + a_2 \cos\left(\frac{\alpha_2}{\beta_2 - t}\right), \tag{4.107}$$

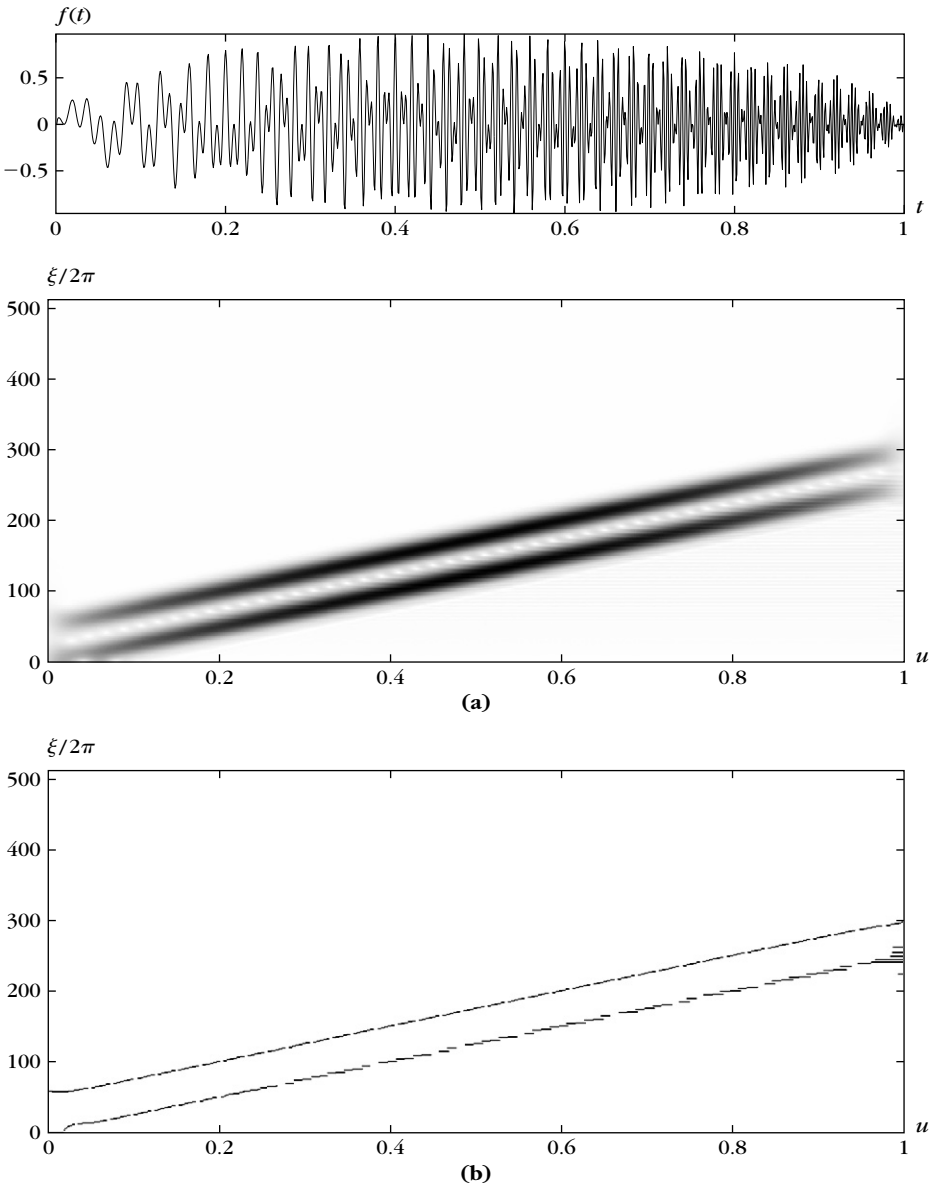

FIGURE 4.13

Sum of two parallel linear chirps: **(a)** Spectrogram $P_S f(u, \xi) = |Sf(u, \xi)|^2$. **(b)** Ridge support calculated from the spectrogram.

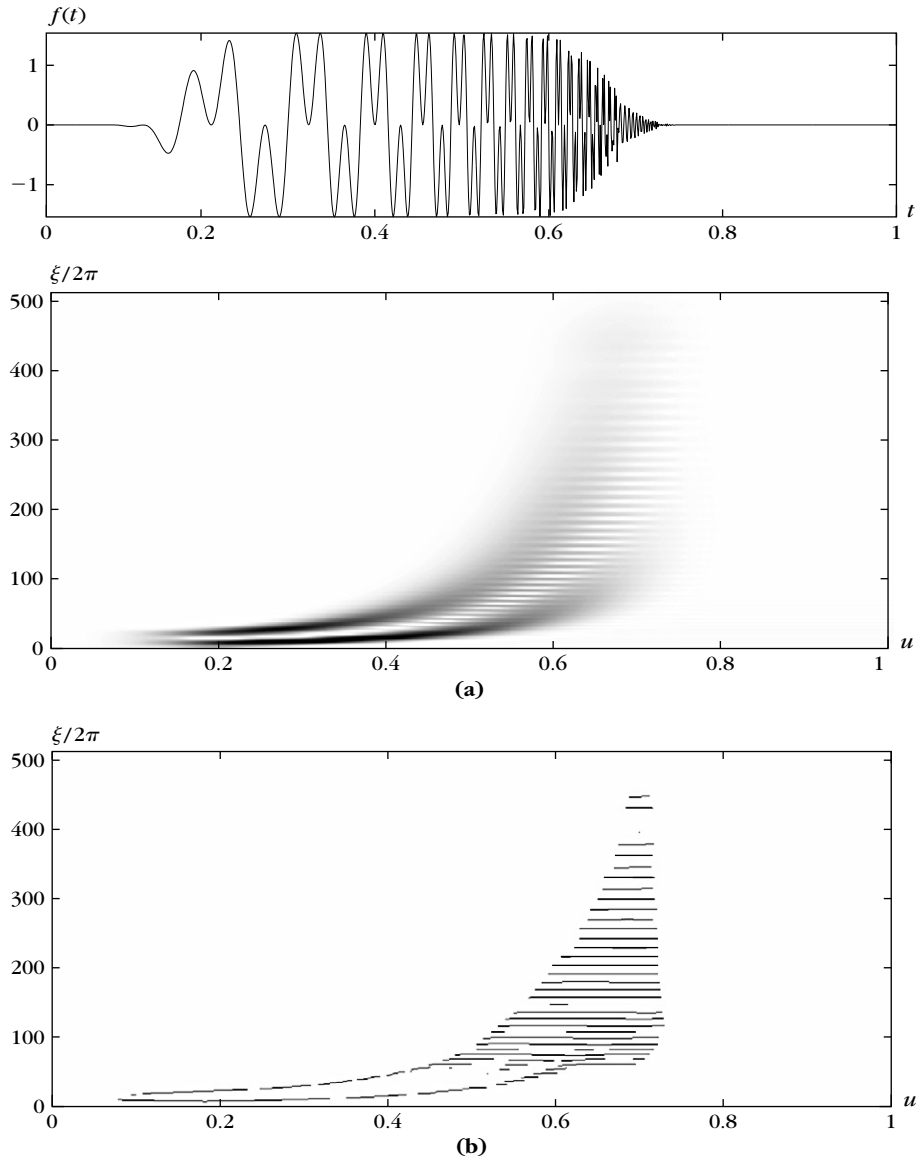

FIGURE 4.14

Sum of two hyperbolic chirps: **(a)** Spectrogram $P_S f(u, \xi)$. **(b)** Ridge support calculated from the spectrogram.

with $\beta_1 = 0.68$ and $\beta_2 = 0.72$. At the beginning of the signal, the two chirps have close instantaneous frequencies that are discriminated by the windowed Fourier ridge computed with a large window. When getting close to β_1 and β_2, the instantaneous frequency varies too quickly relative to window size. The resulting ridges cannot follow the instantaneous frequencies.

4.4.3 Wavelet Ridges

Windowed Fourier atoms have a fixed scale and thus cannot follow the instantaneous frequency of rapidly varying events such as hyperbolic chirps. In contrast, an analytic wavelet transform modifies the scale of its time-frequency atoms. The ridge algorithm of Delprat et al. [204] is extended to analytic wavelet transforms to accurately measure frequency tones that are rapidly changing at high frequencies.

An approximately analytic wavelet is constructed in (4.60) by multiplying a window g with a sinusoidal wave:

$$\psi(t) = g(t)\ \exp(i\eta t).$$

As in the previous section, g is a symmetric window, with a support equal to $[-1/2, 1/2]$, and a unit norm $\|g\| = 1$. Let $\Delta\omega$ be the bandwidth of $\hat{g}$ defined in (4.90). If $\eta > \Delta\omega$, then

$$\forall \omega < 0, \quad \hat{\psi}(\omega) = \hat{g}(\omega - \eta) \ll 1.$$

The wavelet ψ is not strictly analytic because its Fourier transform is not exactly equal to zero at negative frequencies.

Dilated and translated wavelets can be rewritten as

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\,\psi\left(\frac{t-u}{s}\right) = g_{s,u,\xi}(t)\exp(-i\xi u)\,,$$

with $\xi = \eta/s$ and

$$g_{s,u,\xi}(t) = \sqrt{s}\,g\left(\frac{t-u}{s}\right)\ \exp(i\xi t).$$

The resulting wavelet transform uses time-frequency atoms similar to those of a windowed Fourier transform (4.76); however, in this case scale s varies over $\mathbb{R}^+$ while $\xi = \eta/s$:

$$Wf(u,s) = \langle f, \psi_{u,s}\rangle = \langle f, g_{s,u,\xi}\rangle \exp(i\xi u).$$

Theorem 4.6 computes $\langle f, g_{s,u,\xi}\rangle$ when $f(t) = a(t)\cos\theta(t)$, which gives

$$Wf(u,s) = \frac{\sqrt{s}}{2}\,a(u)\ \exp[i\theta(u)]\Big(\hat{g}(s[\xi - \theta'(u)]) + \varepsilon(u,\xi)\Big). \tag{4.108}$$

The corrective term $\varepsilon(u,\xi)$ is negligible if $a(t)$ and $\theta'(t)$ have small variations over the support of $\psi_{u,s}$ and if $\theta'(u) \geqslant \Delta\omega/s$.

Ridge Detection

Instantaneous frequency is measured from ridges defined over the wavelet transform. The normalized scalogram defined by

$$\frac{\xi}{\eta} P_W f(u,\xi) = \frac{|Wf(u,s)|^2}{s} \quad \text{for } \xi = \eta/s$$

is calculated with (4.108):

$$\frac{\xi}{\eta} P_W f(u,\xi) = \frac{1}{4} a^2(u) \left| \hat{g}\left(\eta\left[1 - \frac{\theta'(u)}{\xi}\right]\right) + \varepsilon(u,\xi) \right|^2 .$$

Since $|\hat{g}(\omega)|$ is maximum at $\omega = 0$, if we neglect $\varepsilon(u,\xi)$, this expression shows that the scalogram is maximum at

$$\frac{\eta}{s(u)} = \xi(u) = \theta'(u). \tag{4.109}$$

The corresponding points $(u, \xi(u))$ are called *wavelet ridges*. The analytic amplitude is given by

$$a(u) = \frac{2\sqrt{\eta^{-1}\xi P_w f(u,\xi)}}{|\hat{g}(0)|}. \tag{4.110}$$

The complex phase of $Wf(u,s)$ in (4.108) is $\Theta_W(u,\xi) = \theta(u)$; at ridge points,

$$\frac{\partial \Theta_W(u,\xi)}{\partial u} = \theta'(u) = \xi. \tag{4.111}$$

When $\xi = \theta'(u)$, the first-order term $\varepsilon_{a,1}$ calculated in (4.81) becomes negligible. The corrective term is then dominated by $\varepsilon_{a,2}$ and $\varepsilon_{\theta,2}$. To simplify the expression, we approximate the supremum of a'' and θ'' in the neighborhood of u by their value at u. Since $s = \eta/\xi = \eta/\theta'(u)$, (4.79) and (4.80) imply that these second-order terms become negligible if

$$\frac{\eta^2}{|\theta'(u)|^2} \frac{|a''(u)|}{|a(u)|} \ll 1 \text{ and } \eta^2 \frac{|\theta''(u)|}{|\theta'(u)|^2} \ll 1. \tag{4.112}$$

The presence of θ' in the denominator proves that a' and θ' must have slow variations if θ' is small but may vary much more quickly for large instantaneous frequencies.

Multispectral Estimation

Suppose that f is a sum of two spectral components:

$$f(t) = a_1(t)\cos\theta_1(t) + a_2(t)\cos\theta_2(t).$$

As in (4.94), we verify that the second instantaneous frequency θ'_2 does not interfere with the ridge of θ'_1 if the dilated window has a sufficient spectral resolution at ridge scale $s = \eta/\xi = \eta/\theta'_1(u)$:

$$\hat{g}(s|\theta_1'(u) - \theta_2'(u)|) \ll 1. \tag{4.113}$$

Since the bandwidth of $\hat{g}(\omega)$ is $\Delta\omega$, this means that

$$\frac{|\theta_1'(u) - \theta_2'(u)|}{\theta'_1(u)} \geqslant \frac{\Delta\omega}{\eta}. \tag{4.114}$$

Similarly, the first spectral component does not interfere with the second ridge located at $s = \eta/\xi = \eta/\theta'_2(u)$ if

$$\frac{|\theta_1'(u) - \theta_2'(u)|}{\theta'_2(u)} \geqslant \frac{\Delta\omega}{\eta}. \tag{4.115}$$

To separate spectral lines that have close instantaneous frequencies, these conditions prove that the wavelet must have a small octave bandwidth $\Delta\omega/\eta$. The bandwidth $\Delta\omega$ is a fixed constant on the order of 1. The frequency η is a free parameter that is chosen as a trade-off between the time-resolution condition (4.112) and the frequency bandwidth conditions (4.114) and (4.115).

Figure 4.15 displays the ridges computed from the normalized scalogram and the wavelet phase shown in Figure 4.11. The ridges of the high-frequency transient located at $t = 0.87$ have oscillations because of interference with the linear chirp above. The frequency-separation condition (4.114) is not satisfied. This is also the case in the time interval $[0.35, 0.55]$, where the instantaneous frequencies of the linear and quadratic chirps are too close.

Ridge Support and Processing

The wavelet ridge support Λ of f is the set of all ridge points (u, s) in the time-scale plane or $(u, \xi = \eta/s)$ in the time-frequency plane, corresponding to local maxima of

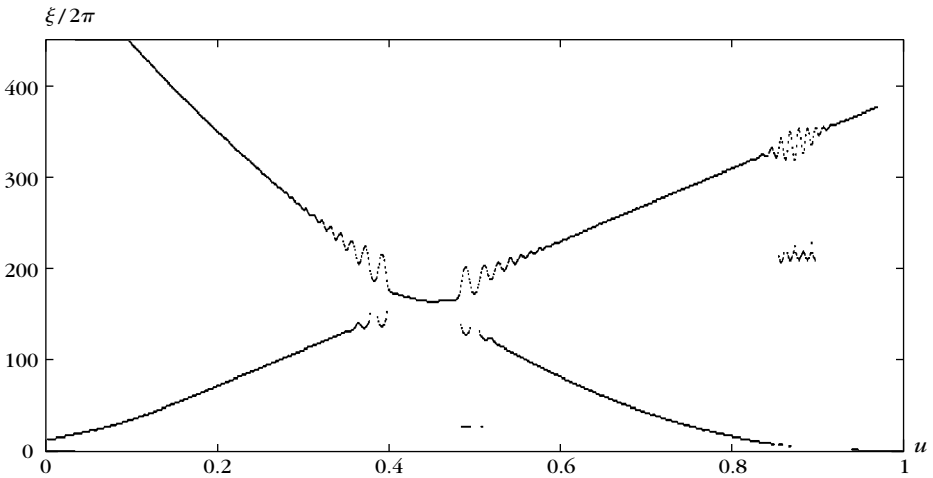

FIGURE 4.15

Ridge support calculated from the scalogram shown in Figure 4.11; compare with the windowed Fourier ridges in Figure 4.12.

$|Wf(u,s)|/s$ for a fixed u and s varying, where the complex phase $\Theta_W(u,s)$ nearly satisfies (4.111).

As in the windowed Fourier case, an orthogonal projection is computed over the space $\mathbf{V}_\Lambda$ generated by the ridge wavelets $\{\psi_{u,s}\}_{(u,s)\in\Lambda}$ by using the dual wavelet frame $\{\tilde{\psi}_{\Lambda,u,s}\}_{(u,s)\in\Lambda}$:

$$f_\Lambda = \sum_{(u,s)\in\Lambda} Wf(u,s)\,\tilde{\psi}_{\Lambda,u,s}. \tag{4.116}$$

It is implemented with the dual-synthesis algorithm of Section 5.1.3 by inverting the symmetric operator

$$Lh = \sum_{(u,s)\in\Lambda} \langle h, \psi_{u,s}\rangle\, \psi_{u,s}, \tag{4.117}$$

which is performed by computing this operator many times. When there are many ridge points, instead of computing this sum only for $(u,s)\in\Lambda$, it may require less operations to compute $wf(u,s)$ with the fast wavelet transform algorithm of Section 4.3.3. All coefficients $(u,s)\notin\Lambda$ are set to zero, and the fast inverse wavelet transform algorithm is applied. The inverse wavelet transform formula (4.66) must be modified by removing the renormalization factor a^{-j} and a^{-J} in the sum (set them to 1) to implement the operator (4.117).

The same as in the windowed Fourier case, modifications of sound durations or frequency transpositions are computed by modifying ridge support. A reduction of sound duration by a factor α transforms ridge support Λ into:

$$\Lambda_\alpha = \{(u,s) \,:\, (\alpha u, s)\in\Lambda\}. \tag{4.118}$$

A sound transposition is implemented by modifying the scales of the time-scale ridge support Λ, which defines:

$$\Lambda_\alpha = \{(u,s) \,:\, (u, s/\alpha)\in\Lambda\}. \tag{4.119}$$

The wavelet coefficients over these supports are derived from the deformation model (4.72) or (4.74), similar to (4.99) and (4.101) for the windowed Fourier transform. Processed signals are recovered from the modified wavelet coefficients and modified supports with the dual-synthesis algorithm of Section 5.1.3.

EXAMPLE 4.15

The instantaneous frequencies of two linear chirps

$$f(t) = a_1\cos(b\,t^2 + c\,t) + a_2\cos(b\,t^2)$$

are not precisely measured by wavelet ridges. Indeed,

$$\frac{|\theta'_2(u) - \theta'_1(u)|}{\theta_1{}'(u)} = \frac{c}{b\,t}$$

converges to zero when t increases. When it is smaller than $\Delta\omega/\eta$, the two chirps interact and create interference patterns like those shown in Figure 4.16. The ridges follow these interferences and do not properly estimate the two instantaneous frequencies, as opposed to the windowed Fourier ridges shown in Figure 4.13.

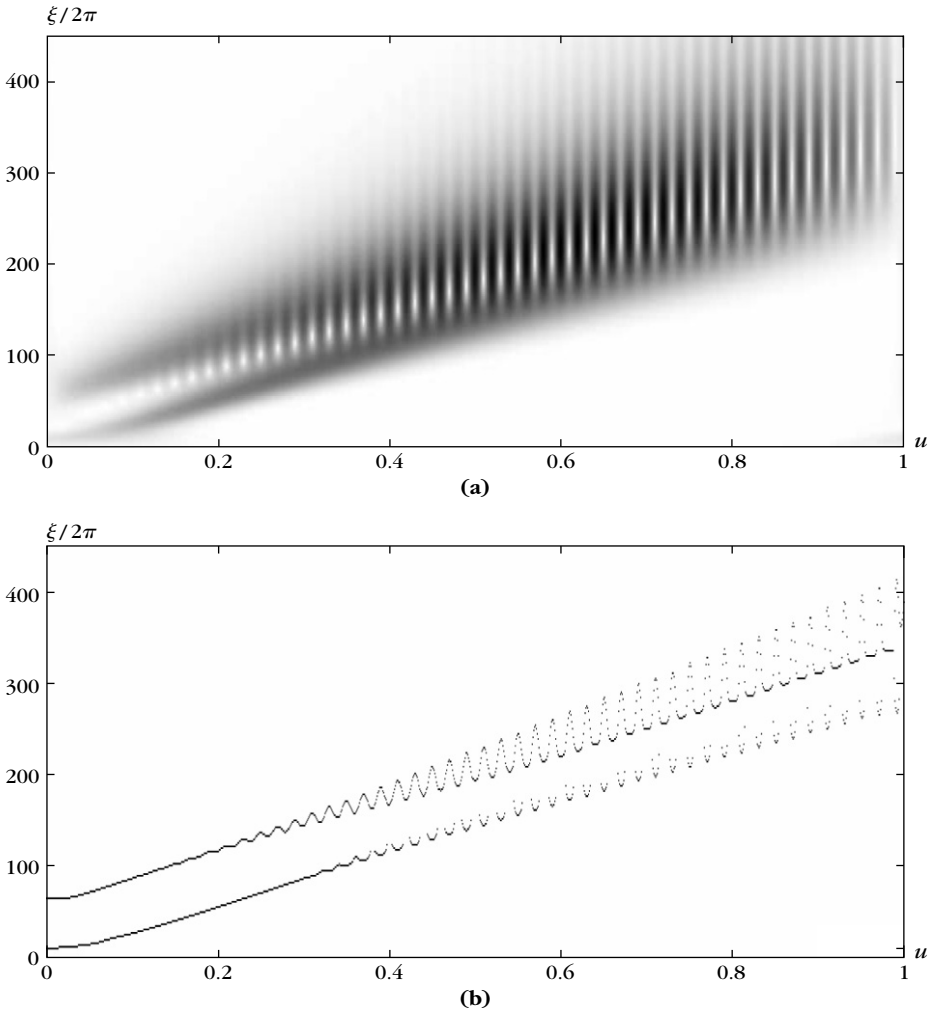

FIGURE 4.16

(a) Normalized scalogram $\eta^{-1}\xi P_W f(u, \xi)$ of two parallel linear chirps shown in Figure 4.13. **(b)** Wavelet ridges.

EXAMPLE 4.16

The instantaneous frequency of a hyperbolic chirp

$$f(t) = \cos\left(\frac{\alpha}{\beta - t}\right)$$

is $\theta'(t) = \alpha\,(1-t)^{-2}$. Wavelet ridges can measure this instantaneous frequency if the time-resolution condition (4.112) is satisfied:

$$\eta^2 \ll \frac{\theta'(t)^2}{|\theta''(t)|} = \frac{\alpha}{|t-\beta|}.$$

This is the case if $|t-\beta|$ is not too large.

Figure 4.17 displays the scalogram and the ridges of two hyperbolic chirps

$$f(t) = a_1 \cos\left(\frac{\alpha_1}{\beta_1 - t}\right) + a_2 \cos\left(\frac{\alpha_2}{\beta_2 - t}\right),$$

with $\beta_1 = 0.68$ and $\beta_2 = 0.72$. As opposed to the windowed Fourier ridges shown in Figure 4.14, the wavelet ridges follow the rapid time modification of both instantaneous frequencies. This is particularly useful in analyzing the returns of hyperbolic chirps emitted by radar or sonar. Several techniques have been developed to detect chirps with wavelet ridges in the presence of noise [151, 455].

Better Is More Sparse

The linear and hyperbolic chirp examples show that the best transform depends on the signal time-frequency property. All examples also show that when the time-frequency transform has a resolution adapted to the signal time-frequency properties, the number of ridge points is reduced. Indeed, if signal structures do not match dictionary time-frequency atoms, then their energy is diffused over many more atoms, which produces more local maxima. Sparsity therefore appears as a natural criterion to adjust the resolution of time-frequency transforms.

Section 12.3.3 studies sparse time-frequency decompositions in very redundant Gabor time-frequency dictionaries, including windowed Fourier and wavelet atoms with a computationally more intense matching pursuit algorithm.

4.5 QUADRATIC TIME-FREQUENCY ENERGY

Wavelet and windowed-Fourier transforms are computed by correlating the signal with families of time-frequency atoms. The time and frequency resolution of these transforms is limited by the time-frequency resolution of the corresponding atoms. Ideally, one would like to define a density of energy in a time-frequency plane with no loss of resolution.

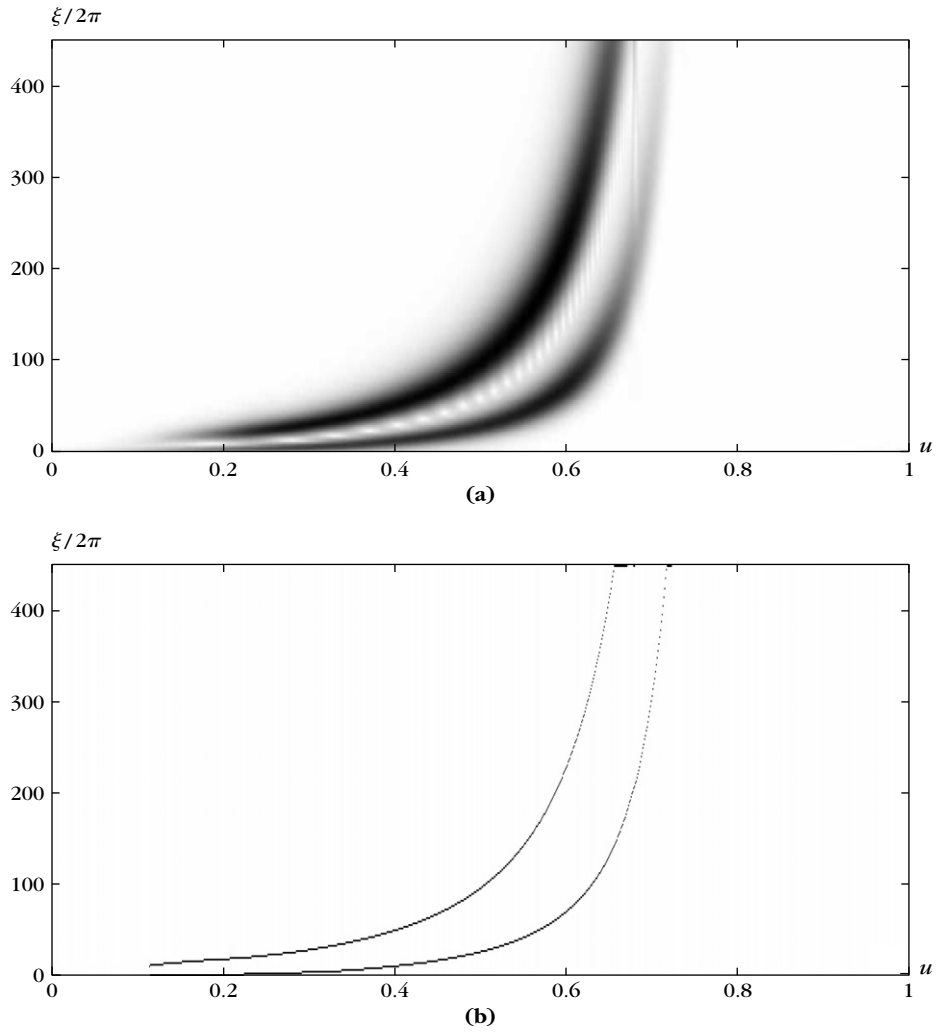

FIGURE 4.17

(a) Normalized scalogram $\eta^{-1}\xi P_W f(u,\xi)$ of two hyperbolic chirps shown in Figure 4.14. **(b)** Wavelet ridges.

The Wigner-Ville distribution is a time-frequency energy density computed by correlating f with a time and frequency translation of itself. Despite its remarkable properties, the application of the Wigner-Ville distribution is limited by the existence of interference terms. Such interferences can be attenuated by time-frequency averaging but this results in a loss of resolution. It will be proved that the spectrogram, the scalogram, and all squared time-frequency decompositions can be written as a

time-frequency averaging of the Wigner-Ville distribution, which gives a common framework to relate the transforms.

4.5.1 Wigner-Ville Distribution

To analyze time-frequency structures, in 1948 Ville [475] introduced in signal processing a quadratic form that had been studied by Wigner [484] in a 1932 article on quantum thermodynamics:

$$P_V f(u,\xi) = \int_{-\infty}^{+\infty} f\left(u+\frac{\tau}{2}\right) f^*\left(u-\frac{\tau}{2}\right) e^{-i\tau\xi}\, d\tau. \tag{4.120}$$

The Wigner-Ville distribution remains real because it is the Fourier transform of $f(u+\tau/2)f^*(u-\tau/2)$, which has a Hermitian symmetry in τ; time and frequency have a symmetric role. This distribution can also be rewritten as a frequency integration by applying the Parseval formula:

$$P_V f(u,\xi) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{f}\left(\xi+\frac{\gamma}{2}\right) \hat{f}^*\left(\xi-\frac{\gamma}{2}\right) e^{i\gamma u}\, d\gamma. \tag{4.121}$$

Time-Frequency Support

The Wigner-Ville transform localizes the time-frequency structures of f. If the energy of f is concentrated in time around u_0 and in frequency around ξ_0, then $P_V f$ has its energy centered at (u_0, ξ_0), with a spread equal to the time and frequency spread of f. This property is illustrated by Theorem 4.7, which relates the time and frequency support of $P_V f$ to the support of f and $\hat{f}$.

Theorem 4.7.

- If the support of f is $[u_0 - T/2, u_0 + T/2]$, then for all ξ the support in u of $P_V f(u,\xi)$ is included in this interval.
- If the support of $\hat{f}$ is $[\xi_0 - \Delta/2, \xi_0 + \Delta/2]$, then for all u the support in ξ of $P_V f(u,\xi)$ is included in this interval.

Proof. Let $\bar{f}(t) = f(-t)$; the Wigner-Ville distribution is rewritten as

$$P_V f(u,\xi) = \int_{-\infty}^{+\infty} f\left(\frac{\tau+2u}{2}\right) \bar{f}^*\left(\frac{\tau-2u}{2}\right) e^{-i\xi\tau}\, d\tau. \tag{4.122}$$

Suppose that f has a support equal to $[u_0 - T/2, u_0 + T/2]$. The supports of $f(\tau/2+u)$ and $\bar{f}(\tau/2-u)$ are then, respectively,

$$[2(u_0-u)-T,\, 2(u_0-u)+T] \text{ and } [-2(u_0+u)-T,\, -2(u_0+u)+T].$$

The Wigner-Ville integral (4.122) shows that $P_V f(u,\xi)$ is nonzero if these two intervals overlap, which is the case only if $|u_0 - u| < T$. Support of $P_V f(u,\xi)$ along u is therefore included in the support of f. If the support of $\hat{f}$ is an interval, then the same derivation based on (4.121) shows that support of $P_V f(u,\xi)$ along ξ is included in support of $\hat{f}$. ■

EXAMPLE 4.17

Theorem 4.7 proves that the Wigner-Ville distribution does not spread the time or frequency support of Diracs or sinusoids, unlike windowed Fourier and wavelet transforms. Direct calculations yield

$$f(t) = \delta(u - u_0) \implies P_V f(u, \xi) = \delta(u - u_0) \tag{4.123}$$

$$f(t) = \exp(i\xi_0 t) \implies P_V f(u, \xi) = \frac{1}{2\pi}\, \delta(\xi - \xi_0) \tag{4.124}$$

EXAMPLE 4.18

If f is a smooth and symmetric window, then its Wigner-Ville distribution $P_V f(u, \xi)$ is concentrated in a neighborhood of $u = \xi = 0$. A Gaussian $f(t) = (\sigma^2\pi)^{-1/4} \exp(-t^2/(2\sigma^2))$ is transformed into a two-dimensional Gaussian because its Fourier transform is also Gaussian (2.32), and one can verify that

$$P_V f(u, \xi) = \frac{1}{\pi} \exp\left(\frac{-u^2}{\sigma^2} - \sigma^2\xi^2\right). \tag{4.125}$$

In this particular case, $P_V f(u, \xi) = |f(u)|^2 |\hat{f}(\xi)|^2$.

The Wigner-Ville distribution has important invariance properties. A phase shift does not modify its value:

$$f(t) = e^{i\theta} g(t) \implies P_V f(u, \xi) = P_V g(u, \xi). \tag{4.126}$$

When f is translated in time or frequency, its Wigner-Ville transform is also translated:

$$f(t) = g(t - u_0) \implies P_V f(u, \xi) = P_V g(u - u_0, \xi) \tag{4.127}$$

$$f(t) = \exp(i\xi_0 t) g(t) \implies P_V f(u, \xi) = P_V g(u, \xi - \xi_0) \tag{4.128}$$

If f is scaled by s and thus $\hat{f}$ is scaled by $1/s$, then the time and frequency parameters of $P_V f$ are also scaled, respectively, by s and $1/s$

$$f(t) = \frac{1}{\sqrt{s}}\, g\left(\frac{t}{s}\right) \implies P_V f(u, \xi) = P_V f\left(\frac{u}{s}, s\xi\right). \tag{4.129}$$

EXAMPLE 4.19

If g is a smooth and symmetric window, then $P_V g(u, \xi)$ has its energy concentrated in the neighborhood of $(0, 0)$. The time-frequency atom

$$f_0(t) = \frac{a}{\sqrt{s}} \exp(i\theta_0)\, g\left(\frac{t - u_0}{s}\right) \exp(i\xi_0 t)$$

has a Wigner-Ville distribution that is calculated with (4.126), (4.127), and (4.128):

$$P_V f_0(u, \xi) = |a|^2\, P_V g\left(\frac{u - u_0}{s}, s(\xi - \xi_0)\right). \tag{4.130}$$

Its energy therefore is concentrated in the neighborhood of (u_0, ξ_0), on an ellipse with axes that are proportional to s in time and $1/s$ in frequency.

Instantaneous Frequency

Ville's original motivation for studying time-frequency decompositions was to compute the instantaneous frequency of a signal [475]. Let f_a be the analytical part of f obtained in (4.69) by setting $\hat{f}(\omega)$ to zero for $\omega < 0$. We write $f_a(t) = a(t)\,\exp[i\theta(t)]$ to define the instantaneous frequency $\omega(t) = \theta'(t)$. Theorem 4.8 proves that $\theta'(t)$ is the "average" frequency computed relative to the Wigner-Ville distribution $P_V f_a$.

Theorem 4.8. If $f_a(t) = a(t)\,\exp[i\theta(t)]$, then

$$\theta'(u) = \frac{\int_{-\infty}^{+\infty} \xi\, P_V f_a(u,\xi)\, d\xi}{\int_{-\infty}^{+\infty} P_V f_a(u,\xi)\, d\xi}. \tag{4.131}$$

Proof. To prove this result, we verify that any function g satisfies

$$\begin{aligned}\iint \xi\, g\Big(u+\frac{\tau}{2}\Big)\, g^*\Big(u-\frac{\tau}{2}\Big)\, \exp(-i\tau\xi)\, d\xi\, d\tau \\ = -\pi i \Big[g'(u)\, g^*(u) - g(u)\, g^{*'}(u)\Big].\end{aligned} \tag{4.132}$$

This identity is obtained by observing that the Fourier transform of $i\xi$ is the derivative of a Dirac, which gives an equality in the sense of distributions:

$$\int_{-\infty}^{+\infty} \xi\, \exp(-i\tau\xi)\, d\xi = -i\, 2\pi\, \delta'(\tau).$$

Since $\int_{-\infty}^{+\infty} \delta'(\tau)\, h(\tau) d\tau = -h'(0)$, inserting $h(\tau) = g(u+\tau/2)\, g^*(u-\tau/2)$ proves (4.132). If $g(u) = f_a(u) = a(u)\,\exp[i\theta(u)]$, then (4.132) gives

$$\int_{-\infty}^{+\infty} \xi\, P_V f_a(u,\xi)\, d\xi = 2\pi\, a^2(u)\, \theta'(u).$$

We will see in (4.136) that $|f_a(u)|^2 = (2\pi)^{-1} \int_{-\infty}^{+\infty} P_V f_a(u,\xi)\, d\xi$, and since $|f_a(u)|^2 = a(u)^2$, we derive (4.131). ■

This theorem shows that for a fixed u the mass of $P_V f_a(u,\xi)$ is typically concentrated in the neighborhood of the instantaneous frequency $\xi = \theta'(u)$. For example, a linear chirp $f(t) = \exp(iat^2)$ is transformed into a Dirac located along the instantaneous frequency $\xi = \theta'(u) = 2au$:

$$P_V f(u,\xi) = \delta(\xi - 2au).$$

Similarly, the multiplication of f by a linear chirp $\exp(iat^2)$ makes a frequency translation of $P_V f$ by the instantaneous frequency $2au$:

$$f(t) = \exp(iat^2)\, g(t) \implies P_V f(u,\xi) = P_V g(u, \xi - 2au). \tag{4.133}$$

Energy Density

The Moyal [379] formula proves that the Wigner-Ville transform is unitary, which implies energy-conservation properties.

Theorem 4.9: *Moyal.* For any f and g in $\mathbf{L}^2(\mathbb{R})$,

$$\left|\int_{-\infty}^{+\infty} f(t)\, g^*(t)\, dt\right|^2 = \frac{1}{2\pi}\int\!\!\int P_V f(u,\xi)\, P_V g(u,\xi)\, du\, d\xi. \tag{4.134}$$

Proof. Let us compute the integral

$$\begin{aligned} I &= \int\!\!\int P_V f(u,\xi)\, P_V g(u,\xi)\, du\, d\xi \\ &= \int\!\!\int\!\!\int\!\!\int f\left(u+\frac{\tau}{2}\right) f^*\left(u-\frac{\tau}{2}\right) g\left(u+\frac{\tau'}{2}\right) g^*\left(u-\frac{\tau'}{2}\right) \\ &\quad \exp[-i\xi(\tau+\tau')]\, d\tau\, d\tau'\, du\, d\xi. \end{aligned}$$

The Fourier transform of $h(t)=1$ is $\hat{h}(\omega)=2\pi\delta(\omega)$, which means that we have a distribution equality $\int \exp[-i\xi(\tau+\tau')]d\xi = 2\pi\delta(\tau+\tau')$. As a result,

$$\begin{aligned} I &= 2\pi\int\!\!\int\!\!\int f\left(u+\frac{\tau}{2}\right) f^*\left(u-\frac{\tau}{2}\right) g\left(u+\frac{\tau'}{2}\right) g^*\left(u-\frac{\tau'}{2}\right)\delta(\tau+\tau')\, d\tau\, d\tau'\, du \\ &= 2\pi\int\!\!\int f\left(u+\frac{\tau}{2}\right) f^*\left(u-\frac{\tau}{2}\right) g\left(u-\frac{\tau}{2}\right) g^*\left(u+\frac{\tau}{2}\right) d\tau\, du. \end{aligned}$$

The change of variable $t = u+\tau/2$ and $t' = u-\tau/2$ yields (4.134). ■

One can consider $|f(t)|^2$ and $|\hat{f}(\omega)|^2/(2\pi)$ as energy densities in time and frequency that satisfy a conservation equation:

$$\|f\|^2 = \int_{-\infty}^{+\infty} |f(t)|^2\, dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |\hat{f}(\omega)|^2\, d\omega.$$

Theorem 4.10 shows that these time and frequency densities are recovered with marginal integrals over the Wigner-Ville distribution.

Theorem 4.10. For any $f \in \mathbf{L}^2(\mathbb{R})$,

$$\int_{-\infty}^{+\infty} P_V f(u,\xi)\, du = |\hat{f}(\xi)|^2 \tag{4.135}$$

and

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} P_V f(u,\xi)\, d\xi = |f(u)|^2. \tag{4.136}$$

Proof. The frequency integral (4.121) proves that the one-dimensional Fourier transform of $g_\xi(u) = P_V f(u,\xi)$, with respect to u, is

$$\hat{g}_\xi(\gamma) = \hat{f}\left(\xi+\frac{\gamma}{2}\right)\hat{f}^*\left(\xi-\frac{\gamma}{2}\right).$$

We derive (4.135) from the fact that it is

$$\hat{g}_\xi(0) = \int_{-\infty}^{+\infty} g_\xi(u)\, du.$$

Similarly, (4.120) shows that $P_V f(u,\xi)$ is the one-dimensional Fourier transform of $f(u+\tau/2)f^*(u-\tau/2)$ with respect to τ, where ξ is the Fourier variable. Its integral in ξ therefore gives the value for $\tau = 0$, which is the identity (4.136). ■

This theorem suggests interpreting the Wigner-Ville distribution as a joint time-frequency energy density. However, the Wigner-Ville distribution misses one fundamental property of an energy density: positivity. Let us compute, for example, the Wigner-Ville distribution of $f = \mathbf{1}_{[-T,T]}$ with the integral (4.120):

$$P_V f(u,\xi) = \frac{2\sin\Big(2(T-|u|)\xi\Big)}{\xi}\, \mathbf{1}_{[-T,T]}(u).$$

It is an oscillating function that takes negative values. In fact, one can prove that translated and frequency-modulated Gaussians are the only functions with positive Wigner-Ville distributions. As we will see in the next section, to obtain positive energy distributions for all signals, it is necessary to average the Wigner-Ville transform, thus losing some time-frequency resolution.

4.5.2 Interferences and Positivity

At this point, the Wigner-Ville distribution may seem to be an ideal tool for analyzing the time-frequency structures of a signal. This, however, is not the case because of interferences created by the transform's quadratic properties. Interference can be removed by averaging the Wigner-Ville distribution with appropriate kernels that yield positive time-frequency densities. However, this reduces time-frequency resolution. Spectrograms and scalograms are examples of positive quadratic distributions obtained by smoothing the Wigner-Ville distribution.

Cross Terms

Let $f = f_1 + f_2$ be a composite signal. Since the Wigner-Ville distribution is a quadratic form,

$$P_V f = P_V f_1 + P_V f_2 + P_V[f_1, f_2] + P_V[f_2, f_1], \tag{4.137}$$

where $P_V[h,g]$ is the cross Wigner-Ville distribution of two signals:

$$P_V[h,g](u,\xi) = \int_{-\infty}^{+\infty} h\left(u+\frac{\tau}{2}\right) g^*\left(u-\frac{\tau}{2}\right) e^{-i\tau\xi}\, d\tau. \tag{4.138}$$

The interference term

$$I[f_1, f_2] = P_V[f_1, f_2] + P_V[f_2, f_1]$$

is a real function that creates nonzero values at unexpected locations of the (u, ξ) plane.

Let us consider two time-frequency atoms defined by

$$f_1(t) = a_1 \, e^{i\theta_1} \, g(t-u_1) \, e^{i\xi_1 t} \text{ and } f_2(t) = a_2 \, e^{i\theta_2} \, g(t-u_2) \, e^{i\xi_2 t},$$

where g is a time window centered at $t = 0$. Their Wigner-Ville distributions computed in (4.130) are

$$P_V f_1(u,\xi) = a_1^2 P_V g(u-u_1, \xi-\xi_1) \text{ and } P_V f_2(u,\xi) = a_2^2 P_V g(u-u_2, \xi-\xi_2).$$

Since the energy of $P_V g$ is centered at $(0, 0)$, the energy of $P_V f_1$ and $P_V f_2$ is concentrated in the neighborhoods of (u_1, ξ_1) and (u_2, ξ_2), respectively. A direct calculation verifies that the interference term is

$$I[f_1, f_2](u,\xi) = 2a_1 a_2 \, P_V g(u-u_0, \xi-\xi_0) \cos\Big((u-u_0)\Delta\xi - (\xi-\xi_0)\Delta u + \Delta\theta\Big)$$

with

$$u_0 = \frac{u_1+u_2}{2}, \quad \xi_0 = \frac{\xi_1+\xi_2}{2}$$
$$\Delta u = u_1 - u_2, \quad \Delta\xi = \xi_1 - \xi_2$$
$$\Delta\theta = \theta_1 - \theta_2 + u_0 \, \Delta\xi.$$

The interference term is an oscillatory waveform centered at the middle point (u_0, ξ_0). This is quite counterintuitive because f and $\hat{f}$ have very little energy in the neighborhood of u_0 and ξ_0. The frequency of the oscillations is proportional to the Euclidean distance $\sqrt{\Delta\xi^2 + \Delta u^2}$ of (u_1, ξ_1) and (u_2, ξ_2). The direction of the oscillations is perpendicular to the line that joins (u_1, ξ_1) and (u_2, ξ_2). Figure 4.18 on the next page displays the Wigner-Ville distribution of two atoms obtained with a Gaussian window g. Oscillating interference appears at the middle time-frequency point.

This figure's example shows that interference $I[f_1, f_2](u, \xi)$ has some energy in regions where $|f(u)|^2 \approx 0$ and $|\hat{f}(\xi)|^2 \approx 0$. Such interferences can have a complicated structure [26, 302], but they are necessarily oscillatory because the marginal integrals (4.135) and (4.136) vanish:

$$\int_{-\infty}^{+\infty} P_V f(u,\xi) d\xi = 2\pi |f(u)|^2, \quad \int_{-\infty}^{+\infty} P_V f(u,\xi) du = |\hat{f}(\xi)|^2.$$

Analytic Part

Interference terms also exist in a real signal f with a single instantaneous frequency component. Let $f_a(t) = a(t) \, \exp[i\theta(t)]$ be its analytic part:

$$f = \mathrm{Re}[f_a] = \frac{1}{2}(f_a + f_a^*).$$

Theorem 4.8 proves that for fixed u, $P_V f_a(u,\xi)$ and $P_V f_a^*(u,\xi)$ have an energy concentrated, respectively, in the neighborhood of $\xi_1 = \theta'(u)$ and $\xi_2 = -\theta'(u)$. Both

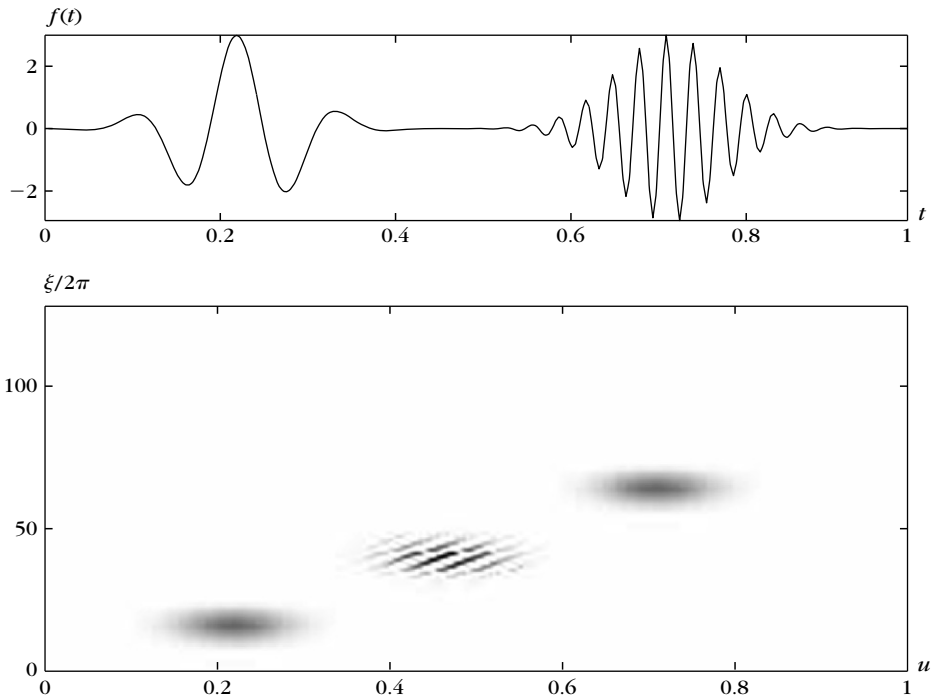

FIGURE 4.18

Wigner-Ville distribution $P_V f(u, \xi)$ of two Gabor atoms (*top*); the oscillating interferences are centered at the middle time-frequency location.

components create an interference term at the intermediate zero frequency $\xi_0 = (\xi_1 + \xi_2)/2 = 0$. To avoid this low-frequency interference, we often compute $P_V f_a$ as opposed to $P_V f$.

Figure 4.19 displays $P_V f_a$ for a real signal f that includes a linear chirp, a quadratic chirp, and two isolated time-frequency atoms. The linear and quadratic chirps are localized along narrow time-frequency lines; they are spread on wider bands by the spectrogram shown in Figure 4.3 and the scalogram shown in Figure 4.4, respectively. However, interference terms create complex oscillatory patterns that make it difficult to detect the existence of the two time-frequency transients at $t = 0.5$ and $t = 0.87$, which clearly appear in the spectrogram and the scalogram.

Positivity

Since interference terms include positive and negative oscillations, they can be partially removed by smoothing $P_V f$ with a kernel K:

$$P_K f(u, \xi) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_V f(u', \xi')\, K(u, u', \xi, \xi')\, du'\, d\xi'. \qquad \textbf{(4.139)}$$

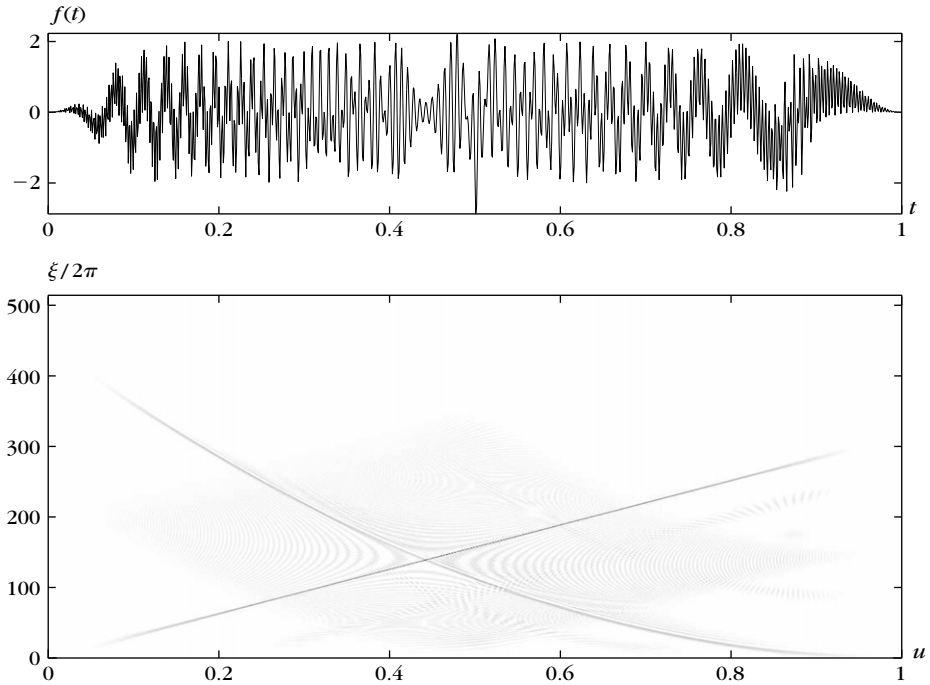

FIGURE 4.19

The Wigner-Ville distribution $P_V f_a(u, \xi)$ (*bottom*) of the analytic part of the signal (*top*).

The time-frequency resolution of this distribution depends on the spread of the kernel K in the neighborhood of (u, ξ). Since interferences take negative values, one can guarantee that all interferences are removed by imposing that this time-frequency distribution remain positive $P_K f(u, \xi) \geqslant 0$ for all $(u, \xi) \in \mathbb{R}^2$.

The spectrogram (4.12) and scalogram (4.55) are examples of positive time-frequency energy distributions. In general, let us consider a family of time-frequency atoms $\{\phi_\gamma\}_{\gamma \in \Gamma}$. Suppose that for any (u, ξ) there exists a unique atom $\phi_{\gamma(u,\xi)}$ centered in time and frequency at (u, ξ). The resulting time-frequency energy density is

$$Pf(u, \xi) = |\langle f, \phi_{\gamma(u,\xi)} \rangle|^2.$$

The Moyal formula (4.134) proves that this energy density can be written as a time-frequency averaging of the Wigner-Ville distribution:

$$Pf(u, \xi) = \frac{1}{2\pi} \int\int P_V f(u', \xi')\ P_V\, \phi_{\gamma(u,\xi)}(u', \xi')\ du'\, d\xi'. \tag{4.140}$$

The smoothing kernel is the Wigner-Ville distribution of the atoms:

$$K(u, u', \xi, \xi') = \frac{1}{2\pi} P_V\, \phi_{\gamma(u,\xi)}(u', \xi').$$

The loss of time-frequency resolution depends on the spread of the distribution $P_V\phi_{\gamma(u,\xi)}(u',\xi')$ in the neighborhood of (u,v).

EXAMPLE 4.20

A spectrogram is computed with windowed Fourier atoms:

$$\phi_{\gamma(u,\xi)}(t) = g(t-u)\,\mathrm{e}^{i\xi t}.$$

The Wigner-Ville distribution calculated in (4.130) yields

$$K(u,u',\xi,\xi') = \frac{1}{2\pi}\,P_V\phi_{\gamma(u,\xi)}(u',\xi') = \frac{1}{2\pi}\,P_V\,g(u'-u,\xi'-\xi). \tag{4.141}$$

For a spectrogram, the Wigner-Ville averaging (4.140) is therefore a two-dimensional convolution with $P_V\,g$. If g is a Gaussian window, then $P_V\,g$ is a two-dimensional Gaussian. This proves that averaging $P_V f$ with a sufficiently wide Gaussian defines a positive energy density. The general class of time-frequency distributions obtained by convolving $P_V f$ with an arbitrary kernel K is studied in Section 4.5.3.

EXAMPLE 4.21

Let ψ be an analytic wavelet with a center frequency that is η. The wavelet atom $\psi_{u,s}(t) = s^{-1/2}\psi((t-u)/s)$ is centered at $(u,\xi=\eta/s)$, and the scalogram is defined by

$$P_W f(u,\xi) = |\langle f,\psi_{u,s}\rangle|^2 \quad \text{for } \xi = \eta/s.$$

Properties (4.127, 4.129) prove that the averaging kernel is

$$K(u,u',\xi,\xi') = \frac{1}{2\pi}\,P_V\psi\left(\frac{u'-u}{s}, s\xi'\right) = \frac{1}{2\pi}\,P_V\psi\left(\frac{\xi}{\eta}(u'-u),\,\frac{\eta}{\xi}\xi'\right).$$

Positive time-frequency distributions totally remove the interference terms but produce a loss of resolution. This is emphasized by Theorem 4.11 described by Wigner [485].

Theorem 4.11: *Wigner.* There is no positive quadratic energy distribution Pf that satisfies

$$\int_{-\infty}^{+\infty} Pf(u,\xi)\,d\xi = 2\pi\,|f(u)|^2 \text{ and } \int_{-\infty}^{+\infty} Pf(u,\xi)\,du = |\hat f(\xi)|^2. \tag{4.142}$$

Proof. Suppose that Pf is a positive quadratic distribution that satisfies these marginals. Since $Pf(u,\xi)\geqslant 0$, the integrals (4.142) imply that, if the support of f is included in an interval I, then $Pf(u,\xi)=0$ for $u\notin I$. We can associate to the quadratic form Pf a bilinear distribution defined for any f and g by

$$P[f,g] = \frac{1}{4}\Big(P(f+g) - P(f-g)\Big).$$

Let f_1 and f_2 be two nonzero signals having their support in two disjoint intervals I_1 and I_2 so that $f_1 f_2 = 0$. Let $f = a f_1 + b f_2$:

$$Pf = |a|^2 Pf_1 + ab^* P[f_1, f_2] + a^* b P[f_2, f_1] + |b|^2 Pf_2.$$

Since I_1 does not intersect I_2, $Pf_1(u, \xi) = 0$ for $u \in I_2$. Remember that $Pf(u, \xi) \geqslant 0$ for all a and b, so necessarily $P[f_1, f_2](u, \xi) = P[f_2, f_1](u, \xi) = 0$ for $u \in I_2$. Similarly, we can prove that these cross terms are zero for $u \in I_1$; therefore

$$Pf(u, \xi) = |a|^2 Pf_1(u, \xi) + |b|^2 Pf_2(u, \xi).$$

Integrating this equation and inserting (4.142) yields

$$|\hat{f}(\xi)|^2 = |a|^2 |\hat{f}_1(\xi)|^2 + |b|^2 |\hat{f}_2(\xi)|^2.$$

Since $\hat{f}(\xi) = a\hat{f}_1(\xi) + b\hat{f}_2(\xi)$, it follows that $\hat{f}_1(\xi)\hat{f}_2^*(\xi) = 0$. But this is not possible because f_1 and f_2 have a compact support in time and Theorem 2.7 proves that $\hat{f}_1$ and $\hat{f}_2$ are $\mathbf{C}^\infty$ functions that cannot vanish on a whole interval. Thus, we conclude that one cannot construct a positive quadratic distribution Pf that satisfies the marginals (4.142). ■

4.5.3 Cohen's Class

While attenuating the interference terms with a smoothing kernel K, we may want to retain certain important properties. Cohen [177] introduced a general class of quadratic time-frequency distributions that satisfy the time translation and frequency modulation invariance properties (4.127) and (4.128). If a signal is translated in time or frequency, its energy distribution is translated just by the corresponding amount. This was the beginning of a systematic study of quadratic time-frequency distributions obtained as a weighted average of a Wigner-Ville distribution [8, 26, 178, 301].

Section 2.1 proves that linear translation-invariant operators are convolution products. The translation-invariance properties (4.127, 4.128) therefore are equivalent to having a smoothing kernel in (4.139) be a convolution kernel:

$$K(u, u', \xi, \xi') = K(u - u', \xi - \xi'); \tag{4.143}$$

therefore

$$P_K f(u, \xi) = P_V f \star K(u, \xi) = \int\int K(u - u', \xi - \xi')\, P_V f(u', \xi')\, du'\, d\xi'. \tag{4.144}$$

The spectrogram is an example of Cohen's class distribution with a kernel in (4.141) that is the Wigner-Ville distribution of the window:

$$K(u, \xi) = \frac{1}{2\pi} P_V g(u, \xi) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} g\left(u + \frac{\tau}{2}\right) g\left(u - \frac{\tau}{2}\right) e^{-i\tau\xi}\, d\tau. \tag{4.145}$$

Ambiguity Function

The properties of the convolution (4.144) are more easily studied by calculating the two-dimensional Fourier transform of $P_V f(u, \xi)$ with respect to u and ξ. We denote by $Af(\tau, \gamma)$ this Fourier transform:

$$Af(\tau, \gamma) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_V f(u, \xi)\ \exp[-i(u\gamma + \xi\tau)]\, du\, d\xi.$$

Note that the Fourier variables τ and γ are inverted with respect to the usual Fourier notation. Since the one-dimensional Fourier transform of $P_V f(u, \xi)$, with respect to u, is $\hat{f}(\xi + \gamma/2)\hat{f}^*(\xi - \gamma/2)$, applying the one-dimensional Fourier transform with respect to ξ gives

$$Af(\tau, \gamma) = \int_{-\infty}^{+\infty} \hat{f}\left(\xi + \frac{\gamma}{2}\right)\hat{f}^*\left(\xi - \frac{\gamma}{2}\right) \mathrm{e}^{-i\tau\xi}\, d\xi. \tag{4.146}$$

The Parseval formula yields

$$Af(\tau, \gamma) = \int_{-\infty}^{+\infty} f\left(u + \frac{\tau}{2}\right) f^*\left(u - \frac{\tau}{2}\right) \mathrm{e}^{-i\gamma u}\, du. \tag{4.147}$$

We recognize the *ambiguity function* encountered in (4.24) when studying the time-frequency resolution of a windowed Fourier transform. It measures the energy concentration of f in time and in frequency.

Kernel Properties

The Fourier transform of $K(u, \xi)$ is

$$\hat{K}(\tau, \gamma) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K(u, \xi)\ \exp[-i(u\gamma + \xi\tau)]\, du\, d\xi.$$

As in the definition of the ambiguity function (4.146), the Fourier parameters τ and γ of $\hat{K}$ are inverted. Theorem 4.12 gives necessary and sufficient conditions to ensure that P_K satisfies marginal energy properties such as those of the Wigner-Ville distribution. Wigner's Theorem 4.11 proves that in this case $P_K f(u, \xi)$ takes negative values.

Theorem 4.12. For all $f \in \mathbf{L^2}(\mathbb{R})$,

$$\int_{-\infty}^{+\infty} P_K f(u, \xi)\, d\xi = 2\pi\, |f(u)|^2, \quad \int_{-\infty}^{+\infty} P_K f(u, \xi)\, du = |\hat{f}(\xi)|^2, \tag{4.148}$$

if and only if

$$\forall (\tau, \gamma) \in \mathbb{R}^2, \quad \hat{K}(\tau, 0) = \hat{K}(0, \gamma) = 1. \tag{4.149}$$

Proof. Let $A_K f(\tau, \gamma)$ be the two-dimensional Fourier transform of $P_K f(u, \xi)$. The Fourier integral at $(0, \gamma)$ gives

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_K f(u, \xi)\, \mathrm{e}^{-iu\gamma}\, d\xi\, du = A_K f(0, \gamma). \tag{4.150}$$

Since the ambiguity function $Af(\tau, \gamma)$ is the Fourier transform of $P_V f(u, \xi)$, the two-dimensional convolution (4.144) gives

$$A_K(\tau, \gamma) = Af(\tau, \gamma)\, \hat{K}(\tau, \gamma). \tag{4.151}$$

The Fourier transform of $2\pi|f(u)|^2$ is $\hat{f} \star \overline{\hat{f}}(\gamma)$, with $\overline{\hat{f}}(\gamma) = \hat{f}^*(-\gamma)$. The relation (4.150) shows that (4.148) is satisfied, if and only if,

$$A_K f(0, \gamma) = Af(0, \gamma)\, \hat{K}(0, \gamma) = \hat{f} \star \overline{\hat{f}}(\gamma). \tag{4.152}$$

Since $P_V f$ satisfies the marginal property (4.135), we similarly prove that

$$Af(0, \gamma) = \hat{f} \star \overline{\hat{f}}(\gamma).$$

Requiring (4.152) to be valid for any $\hat{f}(\gamma)$ is equivalent to requiring that $\hat{K}(0, \gamma) = 1$ for all $\gamma \in \mathbb{R}$. The same derivation applied to other marginal integration yields $\hat{K}(\xi, 0) = 1$. ■

In addition to requiring time-frequency translation invariance, it may be useful to guarantee that P_K satisfies the same scaling property as a Wigner-Ville distribution:

$$g(t) = \frac{1}{\sqrt{s}} f\left(\frac{t}{s}\right) \;\Rightarrow\; P_K g(u, \xi) = P_K f\left(\frac{u}{s}, s\xi\right).$$

Such a distribution P_K is *affine* invariant. One can verify that affine invariance is equivalent to imposing that

$$\forall s \in \mathbb{R}^+, \quad K\left(s\,u, \frac{\xi}{s}\right) = K(u, \xi), \tag{4.153}$$

therefore

$$K(u, \xi) = K(u\,\xi, 1) = \beta(u\,\xi).$$

EXAMPLE 4.22

The *Rihaczek* distribution is an affine invariant distribution whose convolution kernel is

$$\hat{K}(\tau, \gamma) = \exp\left(\frac{i\,\tau\,\gamma}{2}\right). \tag{4.154}$$

A direct calculation shows that

$$P_K f(u, \xi) = f(u)\hat{f}^*(\xi)\, \exp(-iu\xi). \tag{4.155}$$

EXAMPLE 4.23

The kernel of the *Choi-William* distribution is [161]

$$\hat{K}(\tau, \gamma) = \exp(-\sigma^2 \tau^2 \gamma^2). \tag{4.156}$$

It is symmetric and thus corresponds to a real function $K(u, \xi)$. This distribution satisfies the marginal conditions (4.149). Since $\lim_{\sigma \to 0} \hat{K}(\tau, \gamma) = 1$, when σ is small the Choi-William distribution is close to a Wigner-Ville distribution. Increasing σ attenuates the interference terms but spreads $K(u, \xi)$, which reduces the time-frequency resolution of the distribution.

Figure 4.20 shows that the interference terms of two modulated Gaussians nearly disappear when the Wigner-Ville distribution of Figure 4.18 is averaged by a Choi-William kernel having a sufficiently large σ. Figure 4.21 gives the Choi-William distribution corresponding to the Wigner-Ville distribution shown in Figure 4.19. The energy of the linear and quadratic chirps are spread over wider time-frequency bands but the interference terms are attenuated, although not totally removed. It remains difficult to isolate the two modulated Gaussians at $t = 0.5$ and $t = 0.87$, which clearly appear in the spectrogram of Figure 4.3.

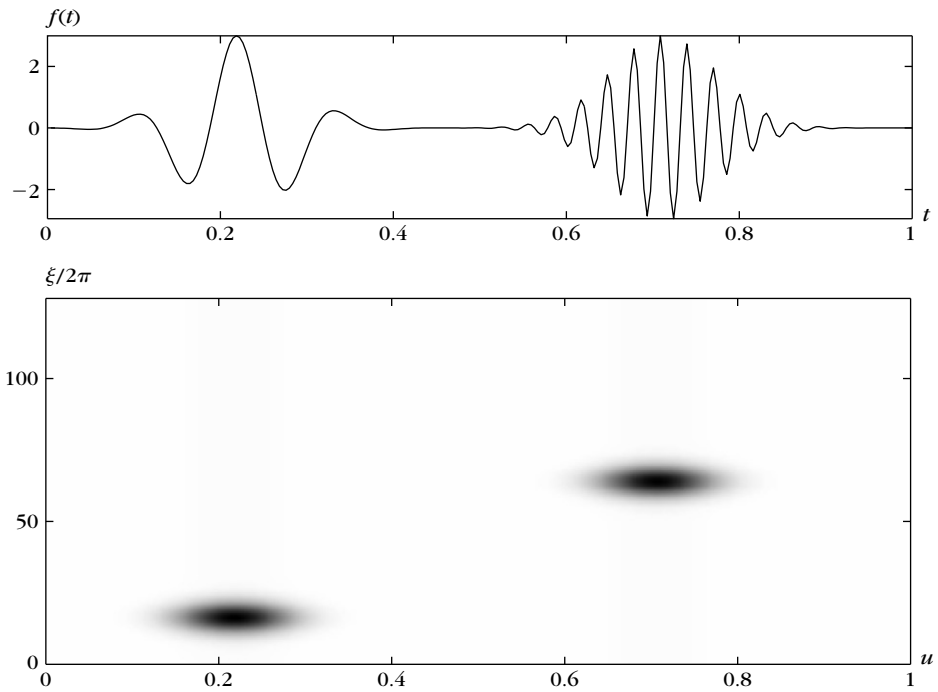

FIGURE 4.20

Choi-William distribution $P_K f(u, \xi)$ of two Gabor atoms (*top*); the interference term that appears in the Wigner-Ville distribution of Figure 4.18 has nearly disappeared.

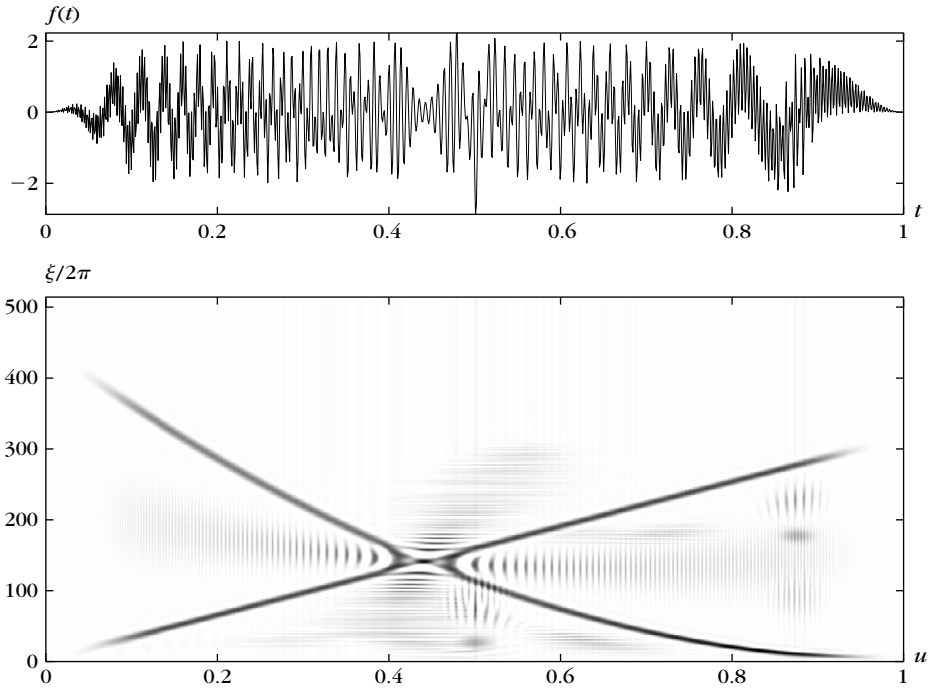

FIGURE 4.21

Choi-William distribution $P_K f_a(u, \xi)$ of signal's analytic part (*top*); the interferences remain visible.

4.5.4 Discrete Wigner-Ville Computations

The Wigner integral (4.120) is the Fourier transform of $f(u+\tau/2)f^*(u-\tau/2)$:

$$P_V f(u,\xi) = \int_{-\infty}^{+\infty} f\left(u+\frac{\tau}{2}\right) f^*\left(u-\frac{\tau}{2}\right) \mathrm{e}^{-i\tau\xi}\, d\tau. \tag{4.157}$$

For a discrete signal $f[n]$, defined over $0 \leqslant n < N$, the integral is replaced by a discrete sum:

$$P_V f[n,k] = \sum_{p=-N}^{N-1} f\left[n+\frac{p}{2}\right] f^*\left[n-\frac{p}{2}\right] \exp\left(\frac{-i2\pi kp}{N}\right). \tag{4.158}$$

When p is odd, this calculation requires knowing the value of f at half integers. Such values are computed by interpolating f by adding zeroes to its Fourier transform. This is necessary to avoid the aliasing produced by the discretization of the Wigner-Ville integral [165].

The interpolation $\tilde{f}$ of f is a signal of size $2N$ that has a DFT $\widehat{\tilde{f}}$ defined from the discrete Fourier transform $\hat{f}$ of f by

$$\widehat{\tilde{f}}[k] = \begin{cases} 2\hat{f}[k] & \text{if } 0 \leqslant k < N/2 \\ 0 & \text{if } N/2 < k < 3N/2 \\ 2\hat{f}[k-N] & \text{if } 3N/2 < k < 2N \\ \hat{f}[N/2] & \text{if } k = N/2,\, 3N/2 \end{cases}$$

Computing the inverse DFT shows that $\tilde{f}[2n] = f[n]$ for $n \in [0, N-1]$. When $n \notin [0,\, 2N-1]$, we set $\tilde{f}[n] = 0$. The Wigner summation (4.158) is calculated from $\tilde{f}$:

$$P_V f[n,k] = \sum_{p=-N}^{N-1} \tilde{f}[2n+p]\tilde{f}^*[2n-p]\, \exp\left(\frac{-i2\pi kp}{N}\right)$$

$$= \sum_{p=0}^{2N-1} \tilde{f}[2n+p-N]\tilde{f}^*[2n-p+N]\, \exp\left(\frac{-i2\pi(2k)p}{2N}\right).$$

For $0 \leqslant n < N$ fixed, $P_V f[n,k]$ is the discrete Fourier transform of size $2N$ of $g[p] = \tilde{f}[2n+p-N]\tilde{f}^*[2n-p+N]$ at the frequency $2k$. Thus, the discrete Wigner-Ville distribution is calculated with N FFT procedures of size $2N$, which requires $O(N^2 \log N)$ operations. To compute the Wigner-Ville distribution of the analytic part f_a of f, we use (4.48).

Cohen's Class

A Cohen's class distribution is calculated with a circular convolution of the discrete Wigner-Ville distribution with a kernel $K[p,q]$:

$$P_K[n,k] = P_V \circledast K[n,k]. \tag{4.159}$$

Its two-dimensional discrete Fourier transform is therefore

$$A_K[p,q] = Af[p,q]\, \hat{K}[p,q]. \tag{4.160}$$

The signal $Af[p,q]$ is the discrete ambiguity function, calculated with a two-dimensional FFT of the discrete Wigner-Ville distribution $P_V f[n,k]$. As in the case of continuous time, we have inverted the index p and q of the usual two-dimensional Fourier transform. The Cohen's class distribution (4.159) is obtained by calculating the inverse Fourier transform of (4.160). This also requires a total of $O(N^2 \log N)$ operations.

4.6 EXERCISES

4.1 [2] *Instantaneous frequency*. Let $f(t) = \exp[i\phi(t)]$.

(a) Prove that $\int_{-\infty}^{+\infty} |Sf(u,\xi)|^2 \, d\xi = 2\pi$. *Hint*: $Sf(u,\xi)$ is a Fourier transform; use the Parseval formula.

(b) Similarly, show that

$$\int_{-\infty}^{+\infty} \xi \, |Sf(u,\xi)|^2 \, d\xi = 2\pi \int_{-\infty}^{+\infty} \phi'(t) \, |g(t-u)|^2 \, dt,$$

and interpret this result.

4.2 [1] When $g(t) = (\pi\sigma^2)^{-1/4} \exp(-t^2/(2\sigma^2))$, compute the ambiguity function $Ag(\tau, \gamma)$.

4.3 [1] Prove that the approximate reconstruction formula (4.66) is exact if and only if

$$\frac{\log_e a}{C_\psi} \sum_{j=I}^{J} \frac{1}{a^j} \hat{\psi}_j[k] + \frac{1}{C_\psi \, a^J} \hat{\phi}_J[k] = 1.$$

Compute numerically the left equation value for different a when $\psi_j[n]$ is constructed from the Gabor wavelet (4.60) and (4.62) with $\sigma = 1$ and $2^J < N/4$.

4.4 [1] Let $g[n]$ be a window with L nonzero coefficients. For signals of size N, describe a fast algorithm that computes the discrete windowed Fourier transform (4.27) with $O(N \log_2 L)$ operations.

4.5 [3] Let K be the reproducing kernel (4.21) of a windowed Fourier transform: $K(u_0, u, \xi_0, \xi) = \langle g_{u,\xi}, g_{u_0,\xi_0} \rangle$.

(a) For any $\Phi \in \mathbf{L}^2(\mathbb{R}^2)$ we define

$$P\Phi(u_0, \xi_0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi(u,\xi) \, K(u_0, u, \xi_0, \xi) \, du \, d\xi.$$

Prove that P is an orthogonal projector on the space of functions $\Phi(u,\xi)$ that are windowed Fourier transforms of functions in $\mathbf{L}^2(\mathbb{R})$.

(b) Suppose that for all $(u,\xi) \in \mathbb{R}^2$ we are given $\tilde{S}f(u,\xi) = Q\Big(Sf(u,\xi)\Big)$, which is a quantization of the windowed Fourier coefficients. How can we reduce the norm $\mathbf{L}^2(\mathbb{R}^2)$ of the quantification error $\varepsilon(u,\xi) = Sf(u,\xi) - Q\Big(Sf(u,\xi)\Big)$?

4.6 [3] Prove the wavelet reconstruction formula (4.45).

4.7 [3] Prove that a scaling function ϕ defined by (4.42) satisfies $\|\phi\| = 1$.

4.8 [2] Let ψ be a real and even wavelet such that $C = \int_0^{+\infty} \omega^{-1} \hat{\psi}(\omega)\, d\omega < +\infty$. Prove that

$$\forall f \in \mathbf{L^2}(\mathbb{R}), \quad f(t) = \frac{1}{C} \int_0^{+\infty} Wf(t,s)\, \frac{ds}{s^{3/2}}. \tag{4.161}$$

4.9 [3] *Analytic continuation.* Let $f \in \mathbf{L^2}(\mathbb{R})$ be a function such that $\hat{f}(\omega) = 0$ for $\omega < 0$. For any complex $z \in \mathbb{C}$ such that $\mathrm{Im}(z) \geqslant 0$, we define

$$f^{(p)}(z) = \frac{1}{\pi} \int_0^{+\infty} (i\omega)^p \hat{f}(\omega)\, \mathrm{e}^{iz\omega}\, d\omega.$$

(a) Verify that if f is $\mathbf{C}^p$ then $f^{(p)}(t)$ is the derivative of order p of $f(t)$.
(b) Prove that if $\mathrm{Im}(z) > 0$, then $f^{(p)}(z)$ is differentiable relative to the complex variable z. Such a function is said to be *analytic* on the upper-half complex plane.
(c) Prove that this analytic extension can be written as a wavelet transform

$$f^{(p)}(x + iy) = y^{-p-1/2}\, Wf(x,y),$$

calculated with an analytic wavelet ψ that you will specify.

4.10 [2] Let $f(t) = \cos(a \cos bt)$. We want to compute precisely the instantaneous frequency of f from the ridges of its windowed Fourier transform. Find a necessary condition on the window support as a function of a and b. If $f(t) = \cos(a \cos bt) + \cos(a \cos bt + ct)$, find a condition on a, b, and c in order to measure both instantaneous frequencies with the ridges of a windowed Fourier transform. Verify your calculations with a numerical implementation.

4.11 [4] *Noise removal.* We want to suppress noise from audio signals by thresholding ridge coefficients. Implement a dual-synthesis algorithm that reconstructs audio signal approximations from windowed Fourier ridge points (4.96) or wavelet ridge points (4.116), with the conjugate-gradient inverse frame algorithm of Theorem 5.8. Study the SNR of audio denoising by thresholding the ridge coefficients. Try to improve this SNR by averaging the spectrogram values along ridges. Compare the SNR with a linear filtering estimator.

4.12 [4] *Sound duration.* Make a program that modifies the sound duration with the formula (4.72) by modifying the ridges of a window Fourier transform with (4.99) or of a wavelet transform with (4.118), and by reconstructing a signal with a dual synthesis.

4.13 [4] *Sound transposition.* Implement a sound transposition, with windowed Fourier or wavelet ridges, with the transposition model (4.73). The resulting modifications of the ridge supports are specified by (4.100) and (4.119). The amplitude of the transposed harmonics can be computed with the autoregressive model (4.75). A signal is restored with a dual-synthesis algorithm.

4.14 [4] The sinusoidal model (4.71) is improved for speech signals by adding a nonharmonic component $B(t)$ to the partials [339]:

$$F(t) = \sum_{k=1}^{K} a_k(t) \cos \phi_k(t) + B(t). \tag{4.162}$$

Given a signal $f(t)$ that is considered to be a realization of $F(t)$, compute the ridges of a windowed Fourier transform, and find the "main" partials and compute their amplitude a_k and phase ϕ_k. These partials are subtracted from the signal. Over intervals of fixed size, the residue is modeled as the realization of an autoregressive process $B(t)$, of order 10 to 15. Use a standard algorithm to compute the parameters of this autoregressive process [57]. Evaluate the audio quality of the sound restored from the calculated model (4.162). Study an application to audio compression by quantizing and coding the parameters of the model.

4.15 [2] Prove that $Pf(u,\xi) = \|f\|^{-2} |f(u)|^2 |\hat{f}(\xi)|^2$ satisfies the marginal properties (4.135) and (4.136). Why can't we apply the Wigner Theorem 4.11?

4.16 [1] Let g_σ be a Gaussian of variance σ^2. Prove that $P_\theta f(u,\xi) = P_V f \star \theta(u,\xi)$ is a positive distribution if $\theta(u,\xi) = g_\sigma(u)\, g_\beta(\xi)$ with $\sigma\beta \geqslant 1/2$. *Hint:* Consider a spectrogram calculated with a Gaussian window.

4.17 [3] Let $\{g_n(t)\}_{n\in\mathbb{N}}$ be an orthonormal basis of $\mathbf{L}^2(\mathbb{R})$. Prove that

$$\forall (t,\omega) \in \mathbb{R}^2, \quad \sum_{n=0}^{+\infty} P_V\, g_n(t,\omega) = 1.$$

4.18 [2] Let $f_a(t) = a(t)\, \exp[i\phi(t)]$ be the analytic part of $f(t)$. Prove that

$$\int_{-\infty}^{+\infty} \Big(\xi - \phi'(t)\Big)^2 P_V f_a(t,\xi)\, d\xi = -\pi\, a^2(t)\, \frac{d^2 \log a(t)}{dt^2}.$$

4.19 [4] To avoid the time-frequency resolution limitations of a windowed Fourier transform, we want to adapt the window size to the signal content. Let $g(t)$ be a window supported in $[-\frac{1}{2}, \frac{1}{2}]$. We denote by $S_j f(u,\xi)$ the windowed Fourier transform calculated with the dilated window $g_j(t) = 2^{-j/2} g(2^{-j}t)$. Find a procedure that computes a single map of ridges by choosing a "best" window size at each (u,ξ). One approach is to choose the scale 2^l for each (u,ξ) such that $|S_l f(u,\xi)|^2 = \sup_j |S_j f(u,\xi)|^2$. Test your algorithm on the linear and hyperbolic chirp signals (4.103, 4.107). Test it on the Tweet and Greasy signals in WaveLab.