Chapter 2

Windowed Fourier Transforms

Summary: Fourier series are ideal for analyzing periodic signals, since the harmonic modes used in the expansions are themselves periodic. By contrast, the Fourier integral transform is a far less natural tool because it uses periodic functions to expand nonperiodic signals. Two possible substitutes are the windowed Fourier transform (WFT) and the wavelet transform. In this chapter we motivate and define the WFT and show how it can be used to give information about signals simultaneously in the time domain and the frequency domain. We then derive the counterpart of the inverse Fourier transform, which allows us to reconstruct a signal from its WFT. Finally, we find a necessary and sufficient condition that an otherwise arbitrary function of time and frequency must satisfy in order to be the WFT of a time signal with respect to a given window and introduce a method of processing signals simultaneously in time and frequency.

Prerequisites: Chapter 1.

2.1 Motivation and Definition of the WFT

Suppose we want to analyze a piece of music for its frequency content. The piece, as perceived by an eardrum, may be accurately modeled by a function $f(t)$ representing the air pressure on the eardrum as a function of time. If the "music" consists of a single, steady note with fundamental frequency ω_1 (in cycles per unit time), then $f(t)$ is periodic with period $P = 1/\omega_1$ and the natural description of its frequency contents is the Fourier series, since the Fourier coefficients c_n give the amplitudes of the various harmonic frequencies $\omega_n = n\omega_1$ occurring in f (Section 1.4). If the music is a series of such notes or a melody, then it is not periodic in general and we cannot use Fourier series directly. One approach in this case is to compute the Fourier integral transform $\hat{f}(\omega)$ of $f(t)$. However, this method is flawed from a practical point of view: To compute $\hat{f}(\omega)$ we must integrate $f(t)$ over *all* time, hence $\hat{f}(\omega)$ contains the *total* amplitude for the frequency ω in the entire piece rather than the distribution of harmonics in each individual note! Thus, if the piece went on for some length of time, we would need to wait until it was over before computing $\hat{f}$, and then the result would be completely uninformative from a musical point of view. (The same is true, of course, if $f(t)$ represents a speech signal or, in the multidimensional case, an image or a video signal.)

Another approach is to chop f up into approximately single notes and analyze each note separately. This analysis has the obvious drawback of being

G. Kaiser, *A Friendly Guide to Wavelets*, Modern Birkhäuser Classics,
DOI 10.1007/978-0-8176-8111-1_2, © Gerald Kaiser 2011

somewhat arbitrary, since it is impossible to state exactly when a given note ends and the next one begins. Different ways of chopping up the signal may result in widely different analyses. Furthermore, this type of analysis must be tailored to the particular signal at hand (to decide how to partition the signal into notes, for example), so it is not "automatic." To devise a more natural approach, we borrow some inspiration from our experience of hearing. Our ears can hear continuous changes in tone as well as abrupt ones, and they do so without an arbitrary partition of the signal into "notes." We will construct a very simple model for hearing that, while physiologically quite inaccurate (see Roederer [1975], Backus [1977]), will serve mainly as a device for motivating the definition of the windowed Fourier transform.

Since the ear analyzes the frequency distribution of a given signal f in *real time*, it must give information about f simultaneously in the frequency domain and the time domain. Thus we model the output of the ear by a function $\tilde{f}(\omega, t)$ depending on both the frequency ω and the time t. For any fixed value of t, $\tilde{f}(\omega, t)$ represents the frequency distribution "heard" at time t, and this distribution varies with t. Since the ear cannot analyze what has not yet occurred, only the values $f(u)$ for $u \leq t$ can be used in computing $\tilde{f}(\omega, t)$. It is also reasonable to assume that the ear has a finite "memory." This means that there is a time interval $T > 0$ such that only the values $f(u)$ for $u \geq t - T$ can influence the output at time t. Thus $\tilde{f}(\omega, t)$ can only depend on $f(u)$ for $t - T \leq u \leq t$. Finally, we expect that values $f(u)$ near the endpoints $u \approx t - T$ and $u \approx t$ have less influence on $\tilde{f}(\omega, t)$ than values in the middle of the interval. These statements can be formulated mathematically as follows: Let $g(u)$ be a function that vanishes outside the interval $-T \leq u \leq 0$, i.e., such that $\operatorname{supp} g \subset [-T, 0]$. $g(u)$ will be a weight function, or *window*, which will be used to "localize" signals in time. We allow g to be complex-valued, although in many applications it may be real. For every $t \in \mathbf{R}$, define

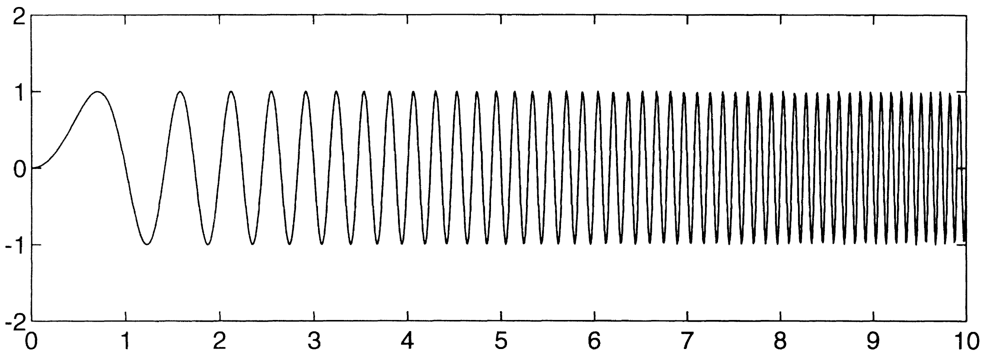
$$f_t(u) \equiv \bar{g}(u - t)\, f(u), \tag{2.1}$$

where $\bar{g}(u - t) \equiv \overline{g(u - t)}$. Then $\operatorname{supp} f_t \subset [t - T, t]$, and we think of f_t as a *localized version* of f that depends only on the values $f(u)$ for $t - T \leq u \leq t$. If g is continuous, then the values $f_t(u)$ with $u \approx t - T$ and $u \approx t$ are small. This means that the above localization is smooth rather than abrupt, a quality that will be seen to be important. We now define the *windowed Fourier transform* (WFT) of f as the Fourier transform of f_t:

$$\begin{aligned} \tilde{f}(\omega, t) \equiv \hat{f}_t(\omega) &= \int_{-\infty}^{\infty} du\ e^{-2\pi\, i\omega u}\ f_t(u) \\ &= \int_{-\infty}^{\infty} du\ e^{-2\pi\, i\omega u}\ \bar{g}(u - t)\, f(u). \end{aligned} \tag{2.2}$$

As promised, $\tilde{f}(\omega, t)$ depends on $f(u)$ only for $t - T \le u \le t$ and (if g is continuous) gives little weight to the values of f near the endpoints.

Note: (a) The condition $\operatorname{supp} g \subset [-T, 0]$ was imposed mainly to give a physical motivation for the WFT. In order for the WFT to make sense, as well as for the reconstruction formula (Section 2.3) to be valid, it will only be necessary to assume that $g(u)$ is square-integrable, i.e. $g \in L^2(\mathbf{R})$. (b) In the extreme case when $g(u) \equiv 1$ (so $g \notin L^2(\mathbf{R})$), the WFT reduces to the ordinary Fourier transform. In the following we merely assume that $g \in L^2(\mathbf{R})$.



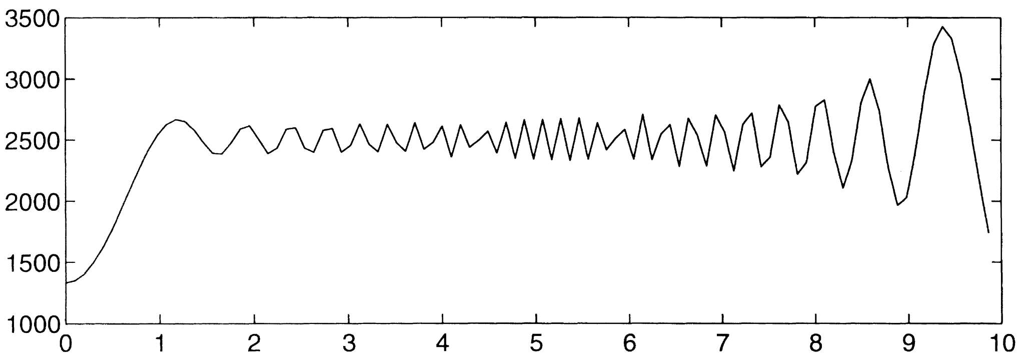


Figure 2.1. *Top*: The chirp signal $f(u) = \sin(\pi u^2)$. *Bottom*: The spectral energy density $|\hat{f}(\omega)|^2$ of f.

If we define

$$g_{\omega,t}(u) \equiv e^{2\pi\, i\omega u}\, g(u - t), \tag{2.3}$$

then $\|g_{\omega,t}\| = \|g\|$; hence $g_{\omega,t}$ also belongs to $L^2(\mathbf{R})$, and the WFT can be expressed as the inner product of f with $g_{\omega,t}$,

$$\tilde{f}(\omega, t) = \langle\, g_{\omega,t}\,, f\,\rangle \equiv g^*_{\omega,t} f, \tag{2.4}$$

which makes sense if both functions are in $L^2(\mathbf{R})$. (See Sections 1.2 and 1.3 for the definition and explanation of the "star notation" $g^* f$.) It is useful to think of $g_{\omega,t}$ as a "musical note" that oscillates at the frequency ω inside the *envelope* defined by $|g(u-t)|$ as a function of u.

Example 2.1: WFT of a Chirp Signal. A *chirp* (in radar terminology) is a signal with a reasonably well defined but steadily rising frequency, such as

$$f(u) = \sin(\pi u^2). \tag{2.5}$$

In fact, the *instantaneous* frequency $\omega_{\text{inst}}(u)$ of f may be defined as the derivative of its phase:

$$2\pi\omega_{\text{inst}}(u) \equiv \partial_u(\pi u^2) = 2\pi u\,. \tag{2.6}$$

Ordinary Fourier analysis hides the fact that a chirp has a well-defined instantaneous frequency by integrating over all of time (or, practically, over a long time period), thus arriving at a very broad frequency spectrum. Figure 2.1 shows $f(u)$ for $0 \le u \le 10$ and its Fourier (spectral) energy density $|\hat{f}(\omega)|^2$ in that range. The spectrum is indeed seen to be very spread out.

We now analyze f using the window function

$$g(u) \equiv \begin{cases} 1+\cos(\pi u) & -1 \le u \le 1 \\ 0 & \text{otherwise,} \end{cases} \tag{2.7}$$

which is pictured in Figure 2.2. (We have centered $g(u)$ around $u=0$, so it is not causal; but $g(u+1)$ is causal with $\tau = 2$.)

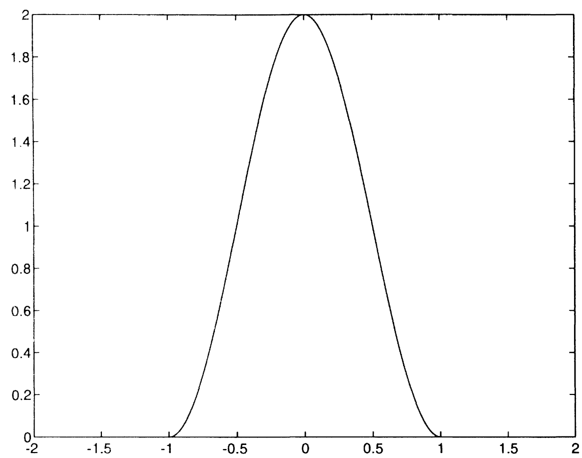


Figure 2.2. The window function $g(u)$.

Figure 2.3 shows the localized version $f_3(u)$ of $f(u)$ and its energy density $|\hat{f}_3(\omega)|^2 = |\tilde{f}(\omega,3)|^2$. As expected, the localized signal has a well-defined (though not *exact!*) frequency $\omega_{\text{inst}}(3) = 3$. It is, therefore, reasonably well localized both

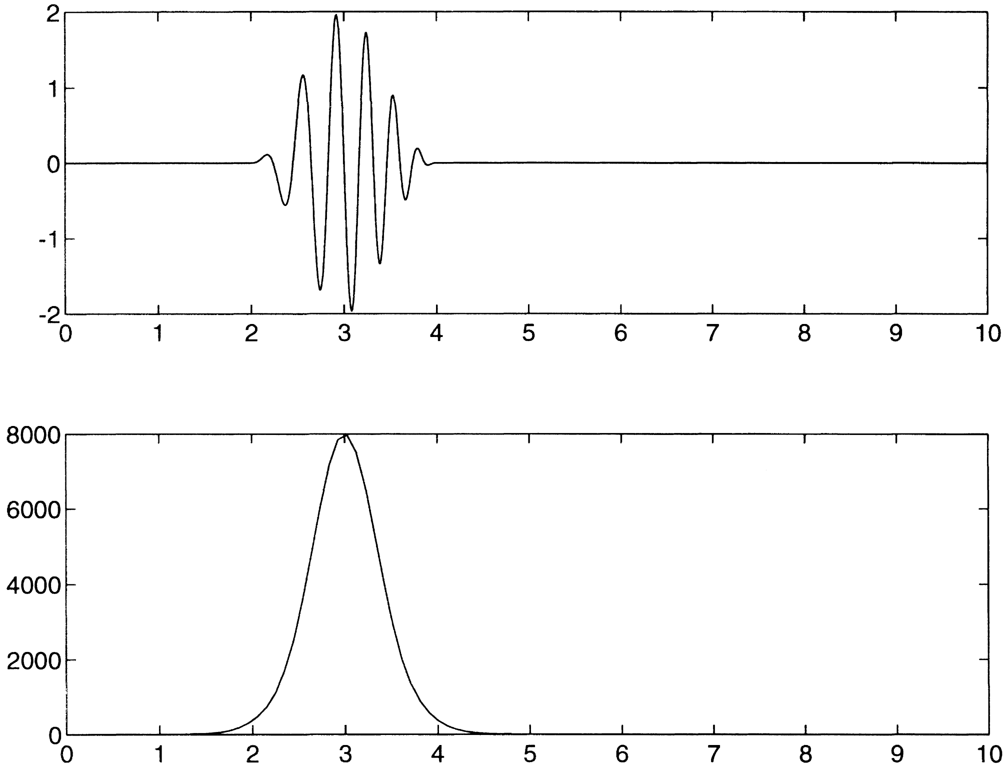


Figure 2.3. *Top*: The localized version $f_3(u)$ of the chirp signal in Figure 2.1 using the window $g(u)$ in Figure 2.2. *Bottom*: The spectral energy density of f_3 showing good localization around the instantaneous frequency $\omega_{\text{inst}}(3) = 3$.

in time and in frequency. Figure 2.4 repeats this analysis at $t = 7$. Now the energy is localized near $\omega_{\text{inst}}(7) = 7$.

Figures 2.5 and 2.6 illustrate frequency resolution. In the top of Figure 2.5 we plot the function

$$\begin{aligned} h(u) &= \operatorname{Re}\left[g_{2,4}(u) + g_{4,6}(u)\right] \\ &= g(u-4)\cos(4\pi u) + g(u-6)\cos(8\pi u), \end{aligned} \tag{2.8}$$

which represents the real part of the sum of two "notes": One centered at $t = 4$ with frequency $\omega = 2$ and the other centered at $t = 6$ with frequency $\omega = 4$. The bottom part of the figure shows the spectral energy density of h. The two peaks are essentially copies of $|\hat{g}(\omega)|^2$ (which is centered at $\omega = 0$) translated to $\omega = 2$ and $\omega = 4$, respectively. This is repeated in Figure 2.6 for $h = \operatorname{Re}\left[g_{2,4} + g_{3,6}\right]$. These two figures show that the window can resolve frequencies down to $\Delta\omega = 2$ but not down to $\Delta\omega = 1$.

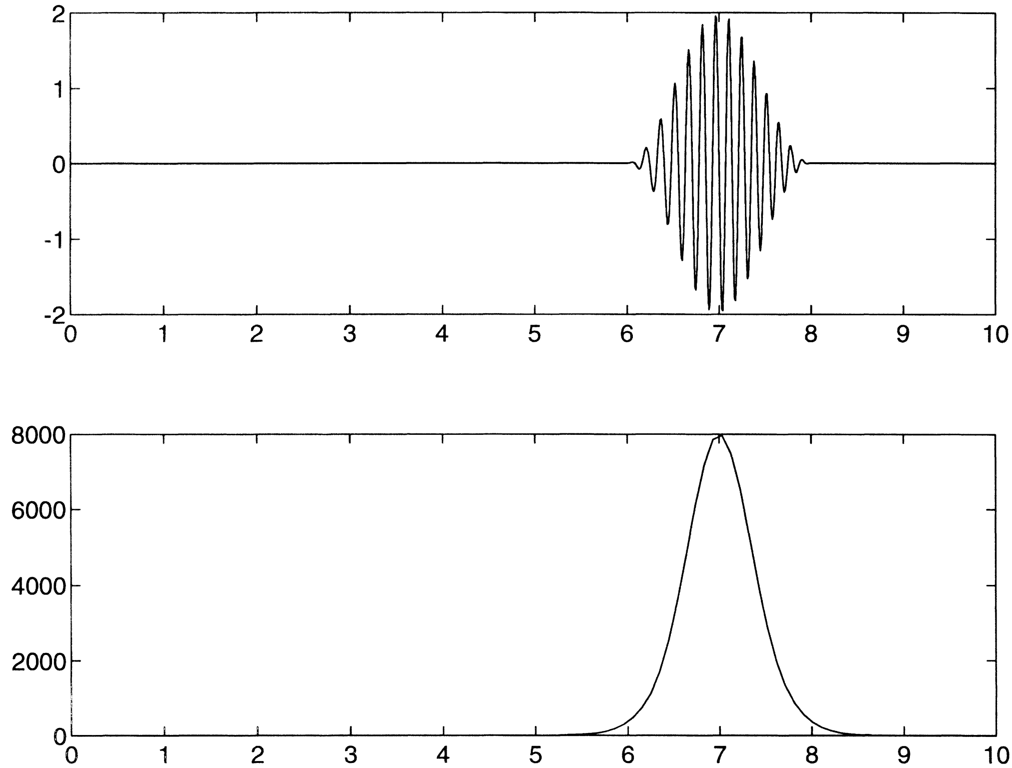


Figure 2.4. *Top*: The localized version $f_7(u)$ of the chirp signal using the window $g(u)$. *Bottom*: The energy density of f_7 showing good localization around the instantaneous frequency $\omega_{\text{inst}}(7) = 7$.

We will see that the vectors $g_{\omega,t}$, parameterized by all frequencies ω and times t, form something analogous to a basis for $L^2(\mathbf{R})$. Note that the inner product $\langle\, g_{\omega,t}\,, f\,\rangle$ is well defined for every choice of ω and t; hence the *values* $\tilde{f}(\omega, t)$ of $\tilde{f}$ are well defined. By contrast, recall from Section 1.3 that the value $f(u)$ of a "function" $f \in L^2(\mathbf{R})$ at any single point u is in general *not* well defined, since f can be modified on a set of measure zero (such as the one-point set $\{u\}$) without changing as an element of $L^2(\mathbf{R})$. The same can be said about the Fourier transform $\hat{f}$ as a function of ω. This remark already indicates that windowed Fourier transforms such as $\tilde{f}(\omega, t)$ are better-behaved than either the corresponding signals $f(t)$ in the time domain or their Fourier transforms $\hat{f}(\omega)$ in the frequency domain. Another example of this can be obtained by applying the Schwarz inequality to (2.4), which gives

$$|\tilde{f}(\omega, t)| = |\langle\, g_{\omega,t}, f\,\rangle| \le \|g_{\omega,t}\|\,\|f\| = \|g\|\,\|f\|, \tag{2.9}$$

showing that $\tilde{f}(\omega, t)$ *is a bounded function,* since the right-hand side is finite and independent of ω and t. Thus a *necessary* condition for a given function

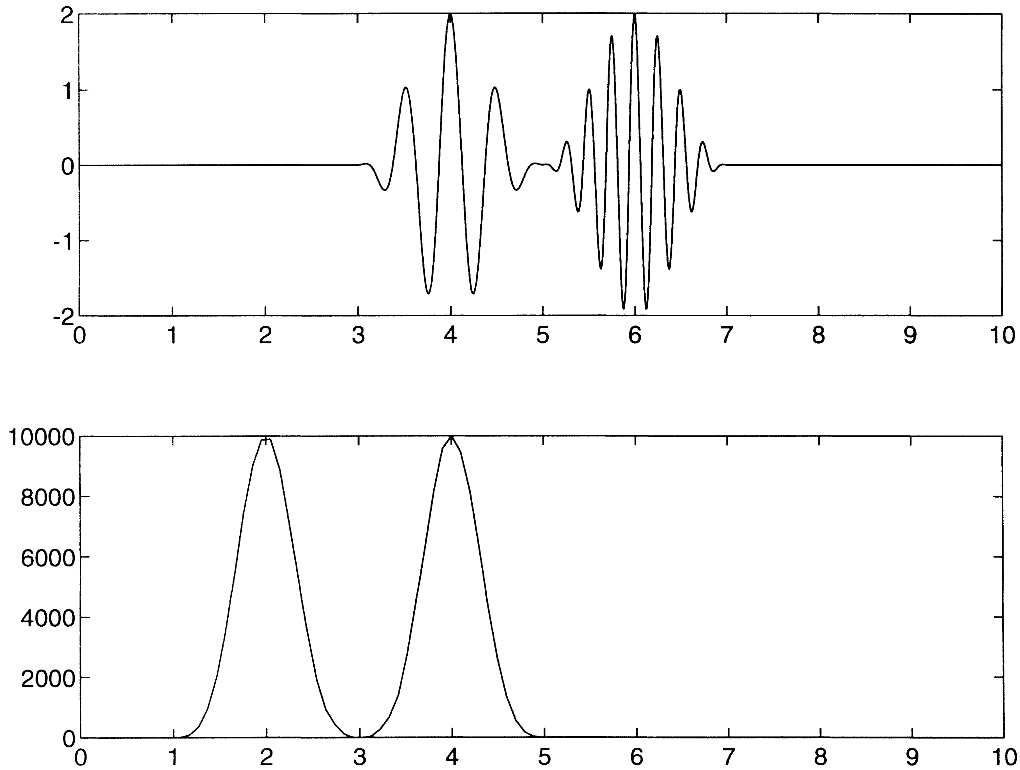


Figure 2.5. *Top*: Plot of $h = \mathrm{Re}\,[g_{2,4} + g_{4,6}]$ for the window in (2.7). *Bottom*: The spectral energy density of h showing good frequency resolution at $\Delta\omega = 2$.

$h(\omega, t)$ to be the WFT of some signal $f \in L^2(\mathbf{R})$ is that h be bounded. However, this condition turns out to be far from *sufficient.* That is, not every bounded function $h(\omega, t)$ is the WFT $\tilde{f}(\omega, s)$ for some time signal $f \in L^2(\mathbf{R})$. In Section 2.3, we derive a condition that is sufficient as well as necessary.

2.2 Time-Frequency Localization

A remarkable aspect of the ordinary Fourier transform is the symmetry it displays between the time domain and the frequency domain, i.e., the fact that the formulas for $f \mapsto \hat{f}$ and $h \mapsto \check{h}$ are identical except for the sign of the exponent. It may appear that this symmetry is lost when dealing with the WFT, since we have treated time and frequency very differently in defining $\tilde{f}$. Actually, the WFT is also completely symmetric with respect to the two domains, as we now show. By Parseval's identity,

$$\tilde{f}(\omega, t) = \langle\, g_{\omega,t}\,, f\,\rangle = \langle\, \hat{g}_{\omega,t}\,, \hat{f}\,\rangle, \tag{2.10}$$

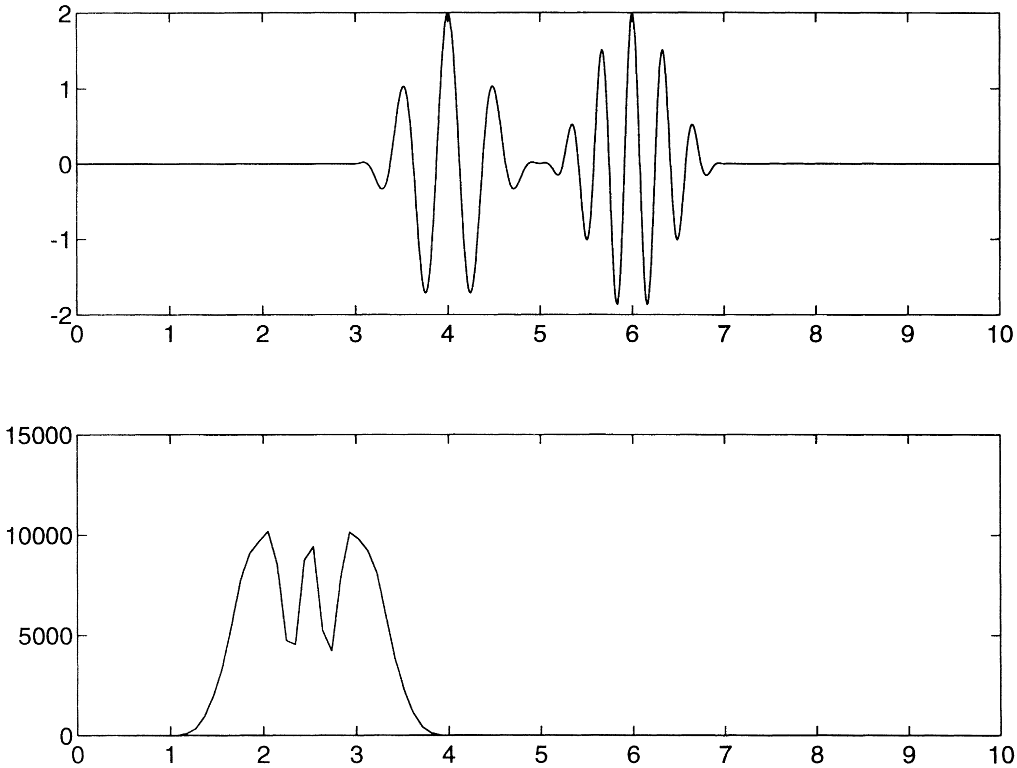


Figure 2.6. *Top*: Plot of $h = \mathrm{Re}\,[g_{2,4} + g_{3,6}]$ for the window in (2.7). *Bottom*: The spectral energy density of h showing poor frequency resolution at $\Delta\omega = 1$.

and the right-hand side is computed to be

$$\begin{aligned} \tilde{f}(\omega, t) &= e^{-2\pi\, i\omega t} \int_{-\infty}^{\infty} d\nu\; e^{2\pi\, i\nu t}\; \bar{\hat{g}}(\nu - \omega)\, \hat{f}(\nu) \\ &= e^{-2\pi\, i\omega t} \left(\bar{\hat{g}}(\nu - \omega)\, \hat{f}(\nu) \right)^{\vee} (t), \end{aligned} \tag{2.11}$$

where $\bar{\hat{g}}(\nu - \omega) \equiv \overline{\hat{g}(\nu - \omega)}$. Equation (2.11) has almost exactly the same form as (2.2) but with the time variable u replaced by the frequency variable ν and the time window $g(u - t)$ replaced by the frequency window $\hat{g}(\nu - \omega)$. (The extra factor $e^{-2\pi\, i\omega t}$ in (2.11) is related to the "Weyl commutation relations" of the *Weyl–Heisenberg group*, which governs translations in time and frequency.) Thus, from the viewpoint of the frequency domain, we begin with the signal $\hat{f}(\nu)$ and "localize" it near the frequency ω by using the window function $\hat{g}$: $\hat{f}_\omega(\nu) \equiv \bar{\hat{g}}(\nu - \omega)\hat{f}(\nu)$; we then take the inverse Fourier transform of $\hat{f}_\omega$ and multiply it by the extra factor $e^{-2\pi\, i\omega t}$, which is a modulation due to the translation of $\hat{g}$ by ω in the frequency domain. If our window g is reasonably well localized in frequency as well as in time, i.e., if $\hat{g}(\nu)$ is small outside a small frequency band in addition to $g(t)$ being small outside a small time interval, then (2.11) shows that the WFT gives a local *time-frequency analysis* of the signal f in the sense that it provides accurate information about f simultaneously in the time domain and in the frequency domain. However, all functions, including windows, obey the *uncertainty principle*, which states that *sharp localizations in time and in frequency are mutually exclusive.* Roughly speaking, if a nonzero function $g(t)$ is small outside a time-interval of length T and its Fourier transform is small outside a frequency band of width Ω, then an inequality of the type $\Omega T \geq c$ must hold for some positive constant $c \sim 1$. The precise value of c depends on how the widths T and Ω of the signal in time and frequency are measured. For example, suppose we normalize g so that $\|g\| = 1$. Let us interpret $|g(t)|^2$ as a "weight distribution" of the window in time (so the total weight in time is $\|g\|^2 = 1$) and $|\hat{g}(\omega)|^2$ as a "weight distribution" of the window in frequency (so the total weight in frequency is $\|\hat{g}\|^2 = \|g\|^2 = 1$, by Plancherel's theorem). The "centers of gravity" of the window in time and frequency are then

$$t_0 \equiv \int_{-\infty}^{\infty} dt \ \cdot t\,|g(t)|^2, \qquad \omega_0 \equiv \int_{-\infty}^{\infty} d\omega \ \cdot \omega\,|\hat{g}(\omega)|^2, \tag{2.12}$$

respectively. A common way of defining T and Ω is as the *standard deviations* from t_0 and ω_0:

$$T^2 \equiv \int_{-\infty}^{\infty} dt \ (t - t_0)^2 |g(t)|^2, \qquad \Omega^2 \equiv \int_{-\infty}^{\infty} d\omega\, (\omega - \omega_0)^2 |\hat{g}(\omega)|^2. \tag{2.13}$$

With these definitions, it can be shown that $4\pi\Omega \cdot T \geq 1$, so in this case $c = 1/4\pi$.[†]

Let us illustrate the foregoing with an example. Choose an arbitrary positive constant a and let g be the *Gaussian* window

$$g(t) = (2a)^{1/4}\, e^{-\pi a t^2}. \tag{2.14}$$

Then $\|g\| = 1$,

$$\hat{g}(\omega) = (2/a)^{1/4}\, e^{-\pi\omega^2/a}\,, \tag{2.15}$$

and

$$t_0 = \omega_0 = 0, \qquad T = \sqrt{\frac{1}{4\pi a}}\,, \qquad \Omega = \sqrt{\frac{a}{4\pi}}. \tag{2.16}$$

† This is the *Heisenberg* form of the uncertainty relation; see Messiah (1961). Contrary to some popular opinion, it is a general property of functions, not at all restricted to quantum mechanics. The connection with the latter is due simply to the fact that in quantum mechanics, if t denotes the position coordinate of a particle, then $2\pi\hbar\omega$ is interpreted as its momentum (where $\hbar$ is Planck's constant), $|f(t)|^2$ and $|\hat{f}(\omega)|^2$ as its probability distributions in space and momentum, and T and $2\pi\hbar\Omega$ as the uncertainties in its position and momentum, respectively. Then the inequality $2\pi\hbar\Omega \cdot T \geq \hbar/2$ is the usual Heisenberg uncertainty relation.

Hence for Gaussian windows, the Heisenberg inequality becomes an *equality*, i.e., $4\pi\,\Omega\,T = 1$. In fact, equality is attained *only* for such windows and their translates in time or frequency (see Messiah [1961]). Note that Gaussian windows are not causal, i.e., they do not vanish for $t > 0$.

Because of its quantum mechanical origin and connotation, the uncertainty principle has acquired an aura of mystique. It can be stated in a great variety of forms that differ from one another in the way the "widths" of the signal in the time and frequency domains are defined. (There is even a version in which T and Ω are replaced by the *entropies* of the signal in the time and frequency domains; see Zakai [1960] and Białynicki–Birula and Mycielski [1975]). However, all these forms are based on one simple fundamental fact: *The precise measurements of time and frequency are fundamentally incompatible, since frequency cannot be measured instantaneously.* That is, if we want to claim that a signal "has frequency ω_0," then the signal must be observed for at least one period, i.e., for a time interval $\Delta t \geq 1/\omega_0$. (The larger the number of periods for which the signal is observed, the more meaningful it becomes to say that it has frequency ω_0.) Hence we cannot say with certainty exactly *when* the signal has this frequency! It is this basic incompatibility that makes the WFT so subtle and, at the same time, so interesting. The solution offered by the WFT is to observe the signal $f(t)$ over the length of the window $\bar{g}(u-t)$ such that the time parameter t occurring in $\tilde{f}(\omega, t)$ is no longer *sharp* (as was u in $f(u)$) but actually represents a *time interval* (e.g., $[t-T, t]$, if $\operatorname{supp} g \subset [-T, 0]$). As seen from (2.11), the frequency ω occurring in $\tilde{f}(\omega, t)$ is not sharp either but it represents a *frequency band* determined by the spread of the frequency window $\overline{\hat{g}}(\nu - \omega)$. The WFT therefore represents a mutual compromise where both time and frequency acquire an approximate, *macroscopic* significance, rather than an exact, *microscopic* significance. Roughly speaking, *the choice of a window determines the dividing line between time and frequency*: Variations in $f(t)$ over time intervals much longer than T show up in the time behavior of $\tilde{f}(\omega, t)$, while those over time intervals much shorter than T become Fourier-transformed and show up in the frequency behavior of $\tilde{f}(\omega, t)$. For example, an elephant's ear can be expected to have a much longer T-value than that of a mouse. Consequently, what sounds like a rhythm or a flutter (time variation) to a mouse may be perceived as a low-frequency tone by an elephant. (However, we remind the reader that the WFT model of hearing is mainly academic, as it is incorrect from a physiological point of view!)

2.3 The Reconstruction Formula

The WFT is a real-time replacement for the Fourier transform, giving the *dynamical* (time-varying) frequency distribution of $f(t)$. The next step is to find a replacement for the inverse Fourier transform, i.e., to *reconstruct* f from $\tilde{f}$. For this purpose, note that since $\tilde{f}(\omega, t) = \hat{f}_t(\omega)$, we can apply the inverse Fourier transform with respect to the variable ω to obtain

$$\bar{g}(u-t)\, f(u) \equiv f_t(u) = \int_{-\infty}^{\infty} d\omega\; e^{2\pi\, i\omega u}\, \tilde{f}(\omega, t). \tag{2.17}$$

We cannot recover $f(u)$ by dividing by $\bar{g}(u-t)$, since this function may vanish. Instead, we multiply (2.17) by $g(u-t)$ and integrate over t:

$$\int_{-\infty}^{\infty} dt\; |g(u-t)|^2\, f(u) = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\omega\; e^{2\pi\, i\omega u}\, g(u-t)\tilde{f}(\omega, t). \tag{2.18}$$

But the left-hand side is just $\|g\|^2\, f(u)$, hence

$$f(u) = C^{-1} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} d\omega\; e^{2\pi\, i\omega u}\, g(u-t)\tilde{f}(\omega, t), \tag{2.19}$$

where we have set $C \equiv \|g\|^2$. This makes sense if g is any nonzero vector in $L^2(\mathbf{R})$, since then $0 < \|g\| < \infty$. By the definition of $g_{\omega,t}$, (2.19) can be written

$$f(u) = C^{-1} \iint d\omega\, dt\; g_{\omega,t}(u)\tilde{f}(\omega, t), \tag{2.20}$$

which is the desired reconstruction formula.

We are now in a position to combine the WFT and its inverse to obtain the analogue of a resolution of unity. To do so, we first summarize our results using the language of vectors and operators (Sections 1.1-1.3). Given a window function $g \in L^2(\mathbf{R})$, we have the following *time-frequency analysis* of a signal $f \in L^2(\mathbf{R})$:

$$\tilde{f}(\omega, t) = \langle\, g_{\omega,t}, f\,\rangle \equiv g^*_{\omega,t} f\,. \tag{2.21}$$

The corresponding *synthesis* or reconstruction formula is

$$f = C^{-1} \iint d\omega\, dt\; g_{\omega,t}\, \tilde{f}(\omega, t)\,, \tag{2.22}$$

where we have left the u-dependence on both sides implicit, in the spirit of vector analysis (f and $g_{\omega,t}$ are both vectors in $L^2(\mathbf{R})$). Note that the complex exponentials $e_\omega(u) \equiv e^{2\pi\, i\omega u}$ occurring in Fourier analysis, which oscillate forever, have now been replaced by the "notes" $g_{\omega,t}(u)$, which are *local* in time if g has compact support or decays rapidly. By substituting (2.21) into (2.22), we obtain

$$f = C^{-1} \iint d\omega\, dt\; g_{\omega,t}\, g^*_{\omega,t}\, f \tag{2.23}$$

for all $f \in L^2(\mathbf{R})$, hence

$$C^{-1} \iint d\omega\, dt\; g_{\omega,t}\, g^*_{\omega,t} = I, \tag{2.24}$$

where I is the identity operator in $L^2(\mathbf{R})$. This is analogous to the resolution of unity in terms of an orthonormal basis $\{\mathbf{b}_n\}$ as given by (1.42) but with the sum $\sum_n$ replaced by the integral $C^{-1}\iint d\omega\, dt$. Equation (2.24) is called a *continuous resolution of unity* in $L^2(\mathbf{R})$, with the "notes" $g_{\omega,t}$ playing a role similar to that of a basis.† This idea will be further investigated and generalized in the following chapters. Note that due to the Hermiticity property (H) of the inner product,

$$f^* g_{\omega,t} = \overline{g^*_{\omega,t} f} = \overline{\tilde{f}(\omega,t)}, \tag{2.25}$$

so that (2.24) implies

$$\begin{aligned} \|f\|^2 \equiv f^* f &= C^{-1} \iint d\omega\, dt\; f^* g_{\omega,t}\, g^*_{\omega,t} f \\ &= C^{-1} \iint d\omega\, dt\; |\tilde{f}(\omega,t)|^2. \end{aligned} \tag{2.26}$$

We may interpret

$$\rho(\omega,t) \equiv C^{-1} |\tilde{f}(\omega,t)|^2$$

as the energy density per unit area of the signal in the time-frequency plane. But area in that plane is measured in cycles, hence $\rho(\omega,t)$ *is the energy density per cycle* of f. Equation (2.26) shows that if a given function $h(\omega,t)$ is to be the WFT of *some* time signal $f \in L^2(\mathbf{R})$, i.e., $h = \tilde{f}$, then h must necessarily be square-integrable in the *joint time-frequency domain.* That is, h must belong to the Hilbert space $L^2(\mathbf{R}^2)$ of all functions with finite norms and inner products defined by

$$\begin{aligned} \|h\|^2_{L^2(\mathbf{R}^2)} &= \iint d\omega\, dt\; |h(\omega,t)|^2, \\ \langle\, h_1, h_2 \,\rangle_{L^2(\mathbf{R}^2)} &= \iint d\omega\, dt\; \bar{h}_1(\omega,t)\, h_2(\omega,t). \end{aligned} \tag{2.27}$$

Equation (2.26) plays a role similar to the Plancherel formula, stating that $\|f\|^2_{L^2(\mathbf{R})} = C^{-1}\|\tilde{f}\|^2_{L^2(\mathbf{R}^2)}$. Since the norm determines the inner product by the polarization identity (1.92), (2.26) implies a counterpart of Parseval's identity:

$$\langle\, f_1, f_2 \,\rangle_{L^2(\mathbf{R})} = C^{-1} \langle\, \tilde{f}_1, \tilde{f}_2 \,\rangle_{L^2(\mathbf{R}^2)} \tag{2.28}$$

for all $f_1, f_2 \in L^2(\mathbf{R})$. (This can also be obtained directly by substituting (2.24) into $f_1^*\, f_2 = f_1^*\, I\, f_2$.)

To simplify the notation, we now normalize g, so that $C = 1$.

---

† Recall that no such resolution of unity existed in terms of the functions e_ω associated with the ordinary Fourier transform, since $e_\omega \notin L^2(\mathbf{R})$; see (1.96) and the discussion below it.

2.4 Signal Processing in the Time-Frequency Domain

The WFT converts a function $f(u)$ of one variable into a function $\tilde{f}(\omega, t)$ of two variables without changing its total energy. This may seem a bit puzzling at first, and it is natural to wonder where the "catch" is. Indeed, not *every* function $h(\omega, t)$ in $L^2(\mathbf{R}^2)$ is the WFT of a time signal. That is, the space of all windowed Fourier transforms of square-integrable time signals,

$$\mathcal{F} \equiv \{\tilde{f} : f \in L^2(\mathbf{R})\}, \tag{2.29}$$

is a *proper subspace* of $L^2(\mathbf{R}^2)$. (This means simply that it is a subspace of $L^2(\mathbf{R}^2)$ but not equal to the latter.) To see this, recall that if $h = \tilde{f}$ for some $f \in L^2(\mathbf{R})$, then h is necessarily *bounded* (2.9). Hence any square-integrable function $h(\omega, t)$ that is unbounded cannot belong to $\mathcal{F}$, and such functions are easily constructed. Thus, being square-integrable is a *necessary* but not *sufficient* condition for $h \in \mathcal{F}$. The next theorem gives an extra condition that is both necessary and sufficient.

Theorem 2.2. *A function $h(\omega, t)$ belongs to $\mathcal{F}$ if and only if it is square-integrable and, in addition, satisfies*

$$h(\omega', t') = \iint d\omega\, dt\, K(\omega', t' \,|\, \omega, t) h(\omega, t)\,, \tag{2.30}$$

where

$$\begin{aligned} K(\omega', t' \,|\, \omega, t) &\equiv g^*_{\omega', t'}\, g_{\omega, t} \equiv \langle\, g_{\omega', t'}\,, g_{\omega, t}\,\rangle = \int_{-\infty}^{\infty} du\ \bar{g}_{\omega', t'}(u)\, g_{\omega, t}(u) \\ &= \int_{-\infty}^{\infty} du\ e^{-2\pi\, i(\omega' - \omega)u}\, \bar{g}(u - t')\, g(u - t). \end{aligned} \tag{2.31}$$

Proof: Our proof will be simple and informal. First, suppose $h = \tilde{f} \in \mathcal{F}$. Then h is square-integrable since $\mathcal{F} \subset L^2(\mathbf{R}^2)$, and we must show that (2.30) holds. Now $h(\omega, t) = \tilde{f}(\omega, t) = g^*_{\omega, t} f$, hence (2.24) implies (recalling that we have set $C = 1$)

$$\begin{aligned} h(\omega', t') = g^*_{\omega', t'}\, If &= \iint d\omega\, dt\ g^*_{\omega', t'}\, g_{\omega, t}\, g^*_{\omega, t}\, f \\ &= \iint d\omega\, dt\ K(\omega', t' \,|\, \omega, t) h(\omega, t)\,, \end{aligned} \tag{2.32}$$

so (2.30) holds. This proves that the two conditions in the theorem are necessary for $h \in \mathcal{F}$. To prove that they are also sufficient, let $h(\omega, t)$ be any square-integrable function that satisfies (2.30). We will *construct* a signal $f \in L^2(\mathbf{R})$ such that $h = \tilde{f}$. Namely, let

$$f \equiv \iint d\omega\, dt\ g_{\omega, t}\, h(\omega, t)\ . \tag{2.33}$$

(Again, this is a *vector* equation!) Then (2.30) implies

$$\begin{aligned}\|f\|^2 &= \iint d\omega'\, dt' \iint d\omega\, dt\ \bar{h}(\omega', t')\,\langle\, g_{\omega',t'}\,, g_{\omega,t}\,\rangle\, h(\omega, t)\\ &= \iint d\omega'\, dt' \iint d\omega\, dt\ \bar{h}(\omega', t')\, K(\omega', t' \,|\, \omega, t)\, h(\omega, t)\\ &= \iint d\omega\, dt\ |h(\omega, t)|^2 = \|h\|^2_{L^2(\mathbf{R}^2)} < \infty\,,\end{aligned} \tag{2.34}$$

which shows that $f \in L^2(\mathbf{R})$. Furthermore, by (2.30),

$$\begin{aligned}\tilde{f}(\omega', t') \equiv g^*_{\omega',t'}\, f &= \iint d\omega\, dt\ g^*_{\omega',t'}\, g_{\omega,t}\, h(\omega, t)\\ &= \iint d\omega\, dt\ \ K(\omega', t' \,|\, \omega, t)\, h(\omega, t)\\ &= h(\omega', t')\,,\end{aligned} \tag{2.35}$$

hence $h = \tilde{f}$, proving that $h \in \mathcal{F}$. ■

The function $K(\omega', t' \,|\, \omega, t)$ is called the *reproducing kernel* determined by the window g, and we call (2.30) the associated *consistency condition.*

It is easy to see why not just *any* square-integrable function $h(\omega, t)$ can be the WFT of a time signal: If that were the case, then we could design time signals with arbitrary time-frequency properties and thus violate the uncertainty principle! As will be seen in Chapter 4, reproducing kernels and consistency conditions are naturally associated with a structure we call *generalized frames*, which includes continuous and discrete time-frequency analyses and continuous and discrete wavelet analyses as special cases.

Suppose now that we are given an arbitrary square-integrable function $h(\omega, t)$ that does not necessarily satisfy the consistency condition. Define $f_h(u)$ by blindly substituting h into the reconstruction formula (2.33), i.e.,

$$f_h \equiv \iint d\omega\, dt\ g_{\omega,t}\, h(\omega, t)\,. \tag{2.36}$$

What can be said about f_h?

Theorem 2.3. *f_h belongs to $L^2(\mathbf{R})$, and it is the unique signal with the following property: For any other signal $f \in L^2(\mathbf{R})$,*

$$\|h - \tilde{f}\|_{L^2(\mathbf{R}^2)} > \|h - \tilde{f}_h\|_{L^2(\mathbf{R}^2)}. \tag{2.37}$$

The meaning of this theorem is as follows: Suppose we want a signal with certain specified properties in both time and frequency. In other words, we look for $f \in L^2(\mathbf{R})$ such that $\tilde{f}(\omega, t) = h(\omega, t)$, where $h \in L^2(\mathbf{R}^2)$ is given. Theorem 2.2 tells us that no such signal can exist unless h satisfies the consistency condition.

The signal f_h defined above comes *closest*, in the sense that the "distance" of its WFT $\tilde{f}_h$ to h is a minimum. We call f_h the *least-squares approximation* to the desired signal. When $h \in \mathcal{F}$, (2.36) reduces to the reconstruction formula. The proof of Theorem 2.3 will be given in a more general context in Chapter 4.

The least-squares approximation can be used to process signals simultaneously in time and in frequency. Given a signal f, we may first compute $\tilde{f}(\omega, t)$ and then modify it in any way desirable, such as by suppressing some frequencies and amplifying others while simultaneously localizing in time. Of course, the modified function $h(\omega, t)$ is generally no longer the WFT of *any* time signal, but its least-squares approximation f_h comes closest to being such a signal, in the above sense. Another aspect of the least-squares approximation is that even when we do not purposefully tamper with $\tilde{f}(\omega, t)$, "noise" is introduced into it through round-off error, transmission error, human error, etc. Hence, by the time we are ready to reconstruct f, the resulting function h will no longer belong to $\mathcal{F}$. ($\mathcal{F}$, being a subspace, is a very *thin* set in $L^2(\mathbf{R}^2)$, like a plane in three dimensions. Hence any random change is almost certain to take h out of $\mathcal{F}$.) The "reconstruction formula" in the form (2.36) then automatically yields the least-squares approximation to the original signal, given the incomplete or erroneous information at hand. This is a kind of built-in stability of the WFT reconstruction, related to *oversampling*. It is typical of reconstructions associated with generalized frames, as explained in detail in Chapter 4.

Exercises

2.1. Prove (2.11) by computing $\hat{g}_{\omega,t}(\omega')$.

2.2. Prove (2.15) and (2.16). *Hint: If* $b > 0$ *and* $u \in \mathbf{R}$, *then*

$$\int_{-\infty}^{\infty} dt\ e^{-\pi b(t+iu)^2} = b^{-1/2}. \tag{2.38}$$

2.3. (a) Show that translation in the time domain corresponds to *modulation* in the frequency domain, and that translation in the frequency domain corresponds to modulation in the time domain, according to

$$\begin{aligned} \left(f(t-t_0)\right)^{\wedge}(\omega) &= e^{-2\pi\, i\omega t_0}\, \hat{f}(\omega), \\ \left(h(\omega-\omega_0)\right)^{\vee}(t) &= e^{2\pi\, i\omega_0 t}\, \check{h}(t). \end{aligned} \tag{2.39}$$

(b) Given a function f_1 such that $f_1(t) \approx 0$ for $|t - t_0| \geq T/2$ and $|\hat{f}_1(\omega)| \approx 0$ for $|\omega - \omega_0| \geq \Omega/2$, use (2.39) to construct a function $f(t)$ such that $f(t) \approx 0$ for $|t| \geq T/2$ and $\hat{f}(\omega) \approx 0$ for $|\omega| \geq \Omega/2$. How can this be used to extend the qualitative explanation of the uncertainty principle given at the end of Section 2.2 to functions centered about arbitrary times t_0 and frequencies ω_0?

2.4. Let $f(t) = e^{2\pi i\alpha t}$, where $\alpha \in \mathbf{R}$. Note that $f \notin L^2(\mathbf{R})$. Nevertheless, the WFT of f with respect to the Gaussian window $g(t) = e^{-\pi t^2}$ is well defined. Use (2.38) to compute $\tilde{f}(\omega, t)$, and show that $|\tilde{f}(\omega, t)|^2$ is maximized when $\omega = \alpha$. Interpret this in view of the fact that $\tilde{f}(\omega, t)$ represents the frequency distribution of f "at" the (macroscopic) time t.

http://www.springer.com/978-0-8176-8110-4

A Friendly Guide to Wavelets
Kaiser, G.
2011, XX, 300 p. 33 illus., Softcover
ISBN: 978-0-8176-8110-4
A product of Birkhäuser Basel