

Figure 113: Signal $S(t)$ and its normalized Fourier transform $\hat{S}(\omega)$. Note the large number of frequency components that make up the signal.

13.4 Time-Frequency Analysis: Windowed Fourier Transforms

The Fourier transform is one of the most important and foundational methods for the analysis of signals. However, it was realized very early on that Fourier transform based methods had severe limitations. Specifically, when transforming a given time signal, it is clear that all the frequency content of the signal can be captured with the transform, but the transform fails to capture the *moment in time* when various frequencies were actually exhibited. Figure 113 shows a proto-typical signal that may be of interest to study. The signal $S(t)$ is clearly comprised of various frequency components that are exhibited at different times. For instance, at the beginning of the signal, there are high frequency components. In contrast, the middle of the signal has relatively low frequency oscillations. If the signal represented music, then the beginning of the signal would produce high notes while the middle would produce low notes. The Fourier transform of the signal contains all this information, but there is no indication of *when* the high or low notes actually occur in time. Indeed, by definition the Fourier transform eliminates all time-domain information since you actually integrated out all time in Eq. (13.1.8).

The obvious question to arise is this: What is the Fourier transform good for in the context of signal processing? In the previous sections where the Fourier transform was applied, the signal being investigated was fixed in frequency, i.e. a sonar or radar detector with a fixed frequency ω_0 . Thus for a given signal, the frequency of interest did not shift in time. By using different measurements in time, a signal tracking algorithm could be constructed. Thus an implicit assumption was made about the invariance of the signal frequency. Ultimately, the Fourier transform is superb for one thing: characterizing *stationary* or peri-

odic signals. Informally, a stationary signal is such that repeated measurements of the signal in time yield an average value that does not change in time. Most signals, however, do not satisfy this criteria. A song, for instance, changes its average Fourier components in time as the song progresses in time. Thus the generic signal $S(t)$ that should be considered, and that is plotted as an example in Fig. 113 is a *non-stationary signal* whose average signal value does change in time. It should be noted that in our application of radar detection of a moving target, use was made of the stationary nature of the spectral content. This allowed for a clear idea of where to filter the signal $\hat{S}(\omega)$ in order to reconstruct the signal $S(t)$.

Having established the fact that the direct application of the Fourier transform provides a nontenable method for extracting signal information, it is natural to pursue modifications of the method in order to extract time and frequency information. The most simple minded approach is to consider Fig. 113 and to decompose the signal over the time domain into separate time frames. Figure 114 shows the original signal $S(t)$ considered but now decomposed into four smaller time windows. In this decomposition, for instance, the first time frame is considered with the remaining three time frames zeroed out. For each time window, the Fourier transform is applied in order to characterize the frequencies present during that time frame. The highest frequency components are captured in Fig. 114(a) which is clearly seen in its Fourier transform. In contrast, the slow modulation observed in the third time frame (c) is devoid of high-frequency components as observed in Fig. 114(c). This method thus exhibits the ability of the Fourier transform, appropriately modified, to extract out both time and frequency information from the signal.

The limitations of the direct application of the Fourier transform, and its inability to localize a signal in both the time and frequency domains, were realized very early on in the development of radar and sonar detection. The Hungarian physicist/mathematician/electrical engineer Gábor Dénes (Physics Nobel Prize in 1971 for the discovery of holography in 1947) was first to propose a formal method for localizing both time and frequency. His method involved a simple modification of the Fourier transform kernel. Thus Gábor introduced the kernel

$$g_{t,\omega}(\tau) = e^{i\omega\tau} g(\tau - t) \quad (13.4.1)$$

where the new term to the Fourier kernel $g(\tau - t)$ was introduced with the aim of localizing both time and frequency. The *Gábor transform*, also known as the *short-time Fourier transform (STFT)* is then defined as the following:

$$\mathcal{G}[f](t, \omega) = \tilde{f}_g(t, \omega) = \int_{-\infty}^{\infty} f(\tau) \bar{g}(\tau - t) e^{-i\omega\tau} d\tau = (f, \bar{g}_{t,\omega}) \quad (13.4.2)$$

where the bar denotes the complex conjugate of the function. Thus the function $g(\tau - t)$ acts as a time filter for localizing the signal over a specific window of time. The integration over the parameter τ slides the time-filtering window

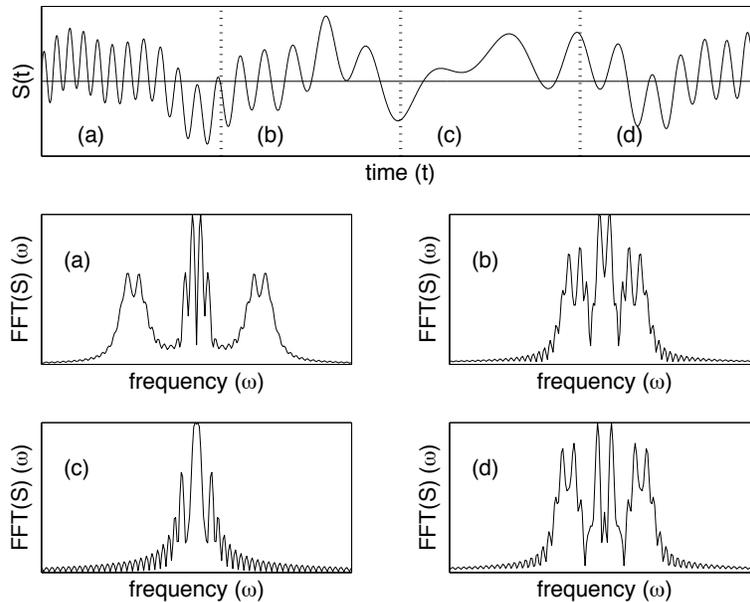


Figure 114: Signal $S(t)$ decomposed into four equal and separate time frames (a), (b), (c) and (d). The corresponding normalized Fourier transform of each time frame $\hat{S}(\omega)$ is illustrated below the signal. Note that this decomposition gives information about the frequencies present in each smaller time frame.

down the entire signal in order to pick out the frequency information at each instant of time. Figure 115 gives a nice illustration of how the time filtering scheme of Gábor works. In this figure, the time filtering window is centered at τ with a width a . Thus the frequency content of a window of time is extracted and τ is modified to extract the frequencies of another window. The definition of the Gábor transform captures the entire time-frequency content of the signal. Indeed, the Gábor transform is a function of the two variables t and ω .

A few of the key mathematical properties of the Gábor transform are highlighted here. To be more precise about these mathematical features, some assumptions about commonly used $g_{t,\omega}$ are considered. Specifically, for convenience we will consider g to be real and symmetric with $\|g(t)\| = 1$ and $\|g(\tau - t)\| = 1$ where $\|\cdot\|$ denotes the L_2 norm. Thus the definition of the Gábor transform, or STFT, is modified to

$$\mathcal{G}[f](t, \omega) = \tilde{f}_g(t, \omega) = \int_{-\infty}^{\infty} f(\tau)g(\tau - t)e^{-i\omega\tau} d\tau \quad (13.4.3)$$

with $g(\tau - t)$ inducing localization of the Fourier integral around $t = \tau$. With this definition, the following properties hold

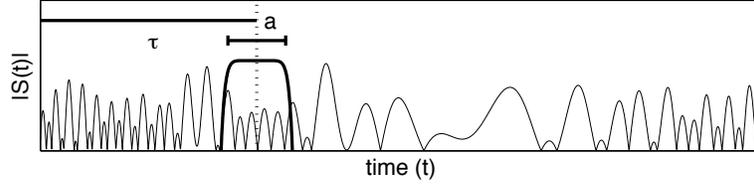


Figure 115: Graphical depiction of the Gábor transform for extracting the time-frequency content of a signal $S(t)$. The time filtering window $g(\tau - t)$ is centered at τ with width a .

1. The energy is bounded by the Schwarz inequality so that

$$|\tilde{f}_g(t, \omega)| \leq \|f\| \|g\| \quad (13.4.4)$$

2. The energy in the signal plane around the neighborhood of (t, ω) is calculated from

$$|\tilde{f}_g(t, \omega)|^2 = \left| \int_{-\infty}^{\infty} f(\tau) g(\tau - t) e^{-i\omega\tau} d\tau \right|^2 \quad (13.4.5)$$

3. The time-frequency spread around a Gábor window is computed from the variance, or second moment, so that

$$\sigma_t^2 = \int_{-\infty}^{\infty} (\tau - t)^2 |g_{t,\omega}(\tau)|^2 d\tau = \int_{-\infty}^{\infty} \tau^2 |g(\tau)|^2 d\tau \quad (13.4.6a)$$

$$\sigma_\omega^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} (\nu - \omega)^2 |\tilde{g}_{t,\omega}(\nu)|^2 d\nu = \frac{1}{2\pi} \int_{-\infty}^{\infty} \nu^2 |\tilde{g}(\nu)|^2 d\nu \quad (13.4.6b)$$

where $\sigma_t \sigma_\omega$ is independent of t and ω and is governed by the Heisenberg uncertainty principle.

4. The Gábor transform is linear so that

$$\mathcal{G}[af_1 + bf_2] = a\mathcal{G}[f_1] + b\mathcal{G}[f_2] \quad (13.4.7)$$

5. The Gábor transform can be inverted with the formula

$$f(\tau) = \frac{1}{2\pi} \frac{1}{\|g\|^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{f}_g(t, \omega) g(\tau - t) e^{i\omega\tau} d\omega dt \quad (13.4.8)$$

where the integration must occur over all frequency and time-shifting components. This double integral is in contrast to the Fourier transform which requires only a single integration since it is a function, $\hat{f}(\omega)$, of the frequency alone.

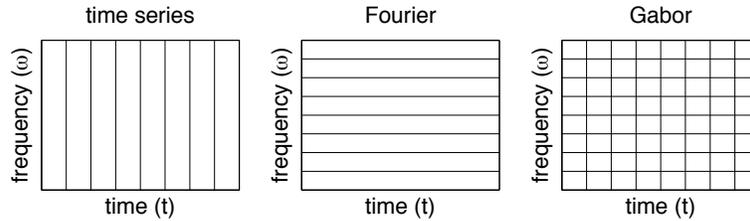


Figure 116: Graphical depiction of the difference between a time series analysis, Fourier analysis and Gabor analysis of a signal. In the time series method, good resolution is achieved of the signal in the time domain, but no frequency resolution is achieved. In Fourier analysis, the frequency domain is well resolved at the expense of losing all time resolution. The Gabor method, or short time Fourier transform, is constructed to give both time and frequency resolution. The area of each box can be constructed from $\sigma_t^2 \sigma_\omega^2$.

Figure 116 is a cartoon representation of the fundamental ideas behind a time series analysis, Fourier transform analysis and Gabor transform analysis of a given signal. In the time series method, good resolution is achieved of the signal in the time domain, but no frequency resolution is achieved. In Fourier analysis, the frequency domain is well resolved at the expense of losing all time resolution. The Gabor method, or short-time Fourier transform, trades away some measure of accuracy in both the time and frequency domains in order to give both time and frequency resolution simultaneously. Understanding this figure is critical to understanding the basic, high-level notions of time-frequency analysis.

In practice, the Gabor transform is computed by discretizing the time and frequency domain. Thus a discrete version of the transform (13.4.2) needs to be considered. Essentially, by discretizing, the transform is done on a lattice of time and frequency. Thus consider the lattice, or sample points,

$$\nu = m\omega_0 \quad (13.4.9a)$$

$$\tau = nt_0 \quad (13.4.9b)$$

where m and n are integers and $\omega_0, t_0 > 0$ are constants. Then the discrete version of $g_{t,\omega}$ becomes

$$g_{m,n}(t) = e^{i2\pi m\omega_0 t} g(t - nt_0) \quad (13.4.10)$$

and the Gabor transform becomes

$$\tilde{f}(m, n) = \int_{-\infty}^{\infty} f(t) \bar{g}_{m,n}(t) dt = (f, g_{m,n}) . \quad (13.4.11)$$

Note that if $0 < t_0, \omega_0 < 1$, then the signal is over-sampled and time frames exist which yield excellent localization of the signal in both time and frequency.

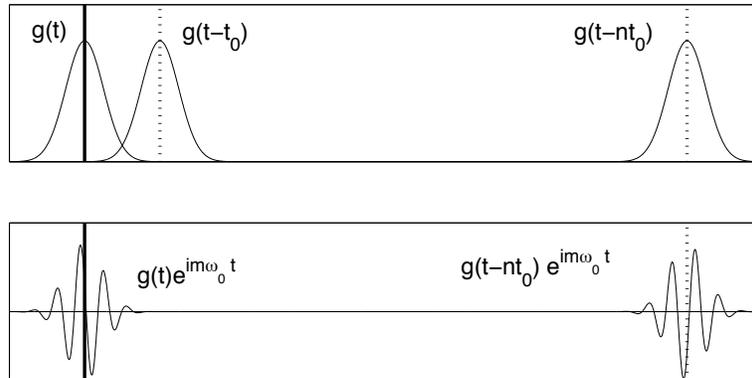


Figure 117: Illustration of the discrete Gabor transform which occurs on the lattice sample points Eq. (13.4.9). In the top figure, the translation with $\omega_0 = 0$ is depicted. The bottom figure depicts both translation in time and frequency. Note that the Gabor frames (windows) overlap so that good resolution of the signal can be achieved in both time and frequency since $0 < t_0, \omega_0 < 1$.

If $\omega_0, t_0 > 1$, the signal is under-sampled and the Gabor lattice is incapable of reproducing the signal. Figure 117 shows the Gabor transform on lattice given by Eq. (13.4.9). The overlap of the Gabor window frames ensures that good resolution in time and frequency of a given signal can be achieved.

Drawbacks of the Gabor (STFT) transform

Although the Gabor transform gives a method whereby time and frequency can be simultaneously characterized, there are obvious limitations to the method. Specifically, the method is limited by the time filtering itself. Consider the illustration of the method in Fig. 115. The time window filters out the time behavior of the signal in a window centered at τ with width a . Thus when considering the spectral content of this window, any portion of the signal with a wavelength longer than the window is completely lost. Indeed, since the Heisenberg relationship must hold, the shorter the time filtering window, the less information there is concerning the frequency content. In contrast, longer windows retain more frequency components, but this comes at the expense of losing the time resolution of the signal. Figure 118 provides a graphical description of the failings of the Gabor transform. Specifically the trade offs that occur between time and frequency resolution, and the fact that high-accuracy in one of these comes at the expense of resolution in the other parameter. This is a consequence of a fixed time filtering window.

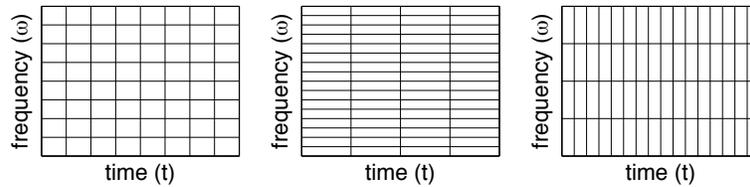


Figure 118: Illustration of the resolution trade-offs in the discrete Gabor transform. The left figure shows a time filtering window that produces nearly equal localization of the time and frequency signal. By increasing the length of the filtering window, increased frequency resolution is gained at the expense of worse time resolution (middle figure). Decreasing the time window does the opposite: time resolution is increased at the expense of poor frequency resolution (right figure).

Other short-time Fourier transform methods

The Gabor transform is not the only windowed Fourier transform that has been developed. There are several other well-used and highly developed STFT techniques. Here, a couple of these more highly used methods will be mentioned for completeness [27].

The *Zak transform* is closely related to the Gabor transform. It is also called the *Weil-Brezin* transform in harmonic analysis. First introduced by Gelfand in 1950 as a method for characterizing eigenfunction expansions in quantum mechanical systems with periodic potentials, it has been generalized to be a key mathematical tool for the analysis of Gabor transform methods. The Zak transform is defined as

$$\mathcal{L}_a f(t, \omega) = \sqrt{a} \sum_{n=-\infty}^{\infty} f(at + an) e^{-i2\pi n\omega} \quad (13.4.12)$$

where $a > 0$ is a constant and n is an integer. Two useful and key properties of this transform are as follows: $\mathcal{L}f(t, \omega + 1) = \mathcal{L}f(t, \omega)$ (periodicity) and $\mathcal{L}f(t + 1, \omega) = \exp(i2\pi\omega)\mathcal{L}f(t, \omega)$ (quasi-periodicity). These properties are particularly important for considering physical problems placed on a lattice.

The *Wigner-Ville Distribution* is a particularly important transform in the development of radar and sonar technologies. Its various mathematical properties make it ideal for these applications and provides a method for achieving great time and frequency localization. The Wigner-Ville transform is defined as

$$\mathcal{W}_{f,g}(t, \omega) = \int_{-\infty}^{\infty} f(t + \tau/2) \bar{g}(t - \tau/2) e^{-i\omega\tau} d\tau \quad (13.4.13)$$

where this is a standard Fourier kernel which transforms the function $f(t +$

$\tau/2)\bar{g}(t - \tau/2)$. This transform is nonlinear since $\mathcal{W}_{f_1+f_2, g_1+g_2} = \mathcal{W}_{f_1, g_1} + \mathcal{W}_{f_1, g_2} + \mathcal{W}_{f_2, g_1} + \mathcal{W}_{f_2, g_2}$ and $\mathcal{W}_{f+g} = \mathcal{W}_f + \mathcal{W}_g + 2\Re\{\mathcal{W}_{f, g}\}$.

Ultimately, alternative forms of the STFT are developed for one specific reason: to take advantage of some underlying properties of a given system. It is rare that a method developed for radar would be broadly applicable to other physical systems unless it were operating under the same physical principles. Regardless, one can see that specialty techniques exist for time-frequency analysis of different systems.

13.5 Time-Frequency Analysis and Wavelets

The Gábor transform established two key principles for joint time-frequency analysis: *translation* of a short-time window and *scaling* of the short-time window to capture finer time resolution. Figure 115 shows the basic concept introduced in the theory of windowed Fourier transforms. Two parameters are introduced to handle the *translation* and *scaling*, namely τ and a . The shortcoming of this method is that it trades off accuracy in time (frequency) for accuracy in frequency (time). Thus the fixed window size imposes a fundamental limitation on the level of time-frequency resolution that can be obtained.

A simple modification to the Gábor method is to allow the scaling window (a) to vary in order to successively extract improvements in the time resolution. In other words, first the low-frequency (poor time resolution) components are extracted using a broad scaling window. The scaling window is subsequently shortened in order to extract out higher-frequencies and better time resolution. By keeping a catalogue of the extracting process, both excellent time and frequency resolution of a given signal can be obtained. This is the fundamental principle of *wavelet theory*. The term wavelet means little wave and originates from the fact that the scaling window extracts out smaller and smaller pieces of waves from the larger signal.

Wavelet analysis begins with the consideration of a function known as the *mother wavelet*:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (13.5.14)$$

where $a \neq 0$ and b are real constants. The parameter a is the scaling parameter illustrated in Fig. 115 whereas the parameter b now denotes the translation parameter (previously denoted by τ in Fig. 115). Unlike Fourier analysis, and very much like Gábor transforms, there are a vast variety of mother wavelets that can be constructed. In principle, the mother wavelet is designed to have certain properties that are somehow beneficial for a given problem. Thus depending upon the application, different mother wavelets may be selected.

Ultimately, the wavelet is simply another expansion basis for representing a given signal or function. Thus it is not unlike the Fourier transform which represents the signal as a series of sines and cosines. Historically, the first wavelet was constructed by Haar in 1910 [28]. Thus the concepts and ideas of

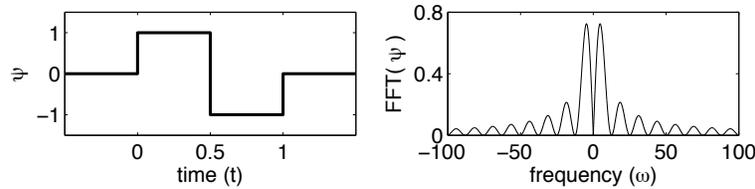


Figure 119: Representation of the compactly supported Haar wavelet function $\psi(t)$ and its Fourier transform $\hat{\psi}(\omega)$. Although highly localized in time due to the compact support, it is poorly localized in frequency with a decay of $1/\omega$.

wavelets are a century old. However, their widespread use and application did not become prevalent until the mid-1980s. The Haar wavelet is given by the piecewise constant function

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \\ 0 & \text{otherwise} \end{cases} . \quad (13.5.15)$$

Figure 119 shows the Haar wavelet step function and its Fourier transform which is a sinc like function. Note further that $\int_{-\infty}^{\infty} \psi(t) dt = 0$ and $\|\psi(t)\|^2 = \int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1$. The Haar wavelet is an ideal wavelet for describing localized signals in time (or space) since it has compact support. Indeed, for highly localized signals, it is much more efficient to use the Haar wavelet basis than the standard Fourier expansion. However, the Haar wavelet has poor localization properties in the frequency domain since it decays like a sinc function in powers of $1/\omega$. This is a consequence of the Heisenberg uncertainty principle.

To represent a signal with the Haar wavelet basis, the translation and scaling operations associated with the mother wavelet need to be considered. Depicted in Fig. 119 and given by Eq. (13.5.15) is the wavelet $\psi_{1,0}(t)$. Thus its translation is zero and its scaling is unity. The concept in reconstructing a signal using the Haar wavelet basis is to consider decomposing the signal into more generic $\psi_{m,n}(t)$. By appropriate selection of the m and n , finer scales and appropriate locations of the signal can be extracted. For $a < 1$, the wavelet is a compressed version of $\psi_{1,0}$ whereas for $a > 1$, the wavelet is a dilated version of $\psi_{1,0}$. The scaling parameter a is typically taken to be a power of two so that $a = 2^j$ for some integer j . Figure 120 shows the compressed and dilated Haar wavelet for $a = 0.5$ and $a = 2$, i.e. $\psi_{1/2,0}$ and $\psi_{2,0}$. The compressed wavelet allows for finer scale resolution of a given signal while the dilated wavelet captures low-frequency components of a signal by having a broad range in time.

The simple Haar wavelet already illustrates all the fundamental principles of the wavelet concept. Specifically by using scaling and translation, a given signal or function can be represented by a basis of functions which allows for

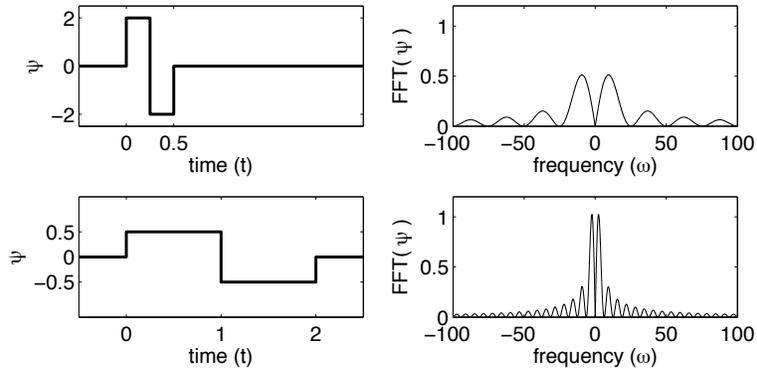


Figure 120: Illustration of the compression and dilation process of the Haar wavelet and its Fourier transform. In the top row, the compressed wavelet $\psi_{1/2,0}$ is shown. Improved time resolution is obtained at the expense of a broader frequency signature. The bottom row, shows the dilated wavelet $\psi_{2,0}$ which allows it to capture lower-frequency components of a signal.

higher and higher refinement in the time resolution of a signal. Thus it is much like the Gábor concept, except that now the time window is variable in order to capture different levels of resolution. Thus the large scale structures in time are captured with broad time-domain Haar wavelets. At this scale, the time resolution of the signal is very poor. However by successive rescaling in time, a finer and finer time resolution of the signal can be obtained along with its high-frequency components. The information at the low and high scales is all preserved so that a complete picture of the time-frequency domain can be constructed. Ultimately, the only limit in this process is the number of scaling levels to be considered.

The wavelet basis can be accessed via an integral transform of the form

$$(Tf)(\omega) = \int_t K(t, \omega) f(t) dt \quad (13.5.16)$$

where $K(t, \omega)$ is the kernel of the transform. This is equivalent in principle to the Fourier transform whose kernel are the oscillations given by $K(t, \omega) = \exp(-i\omega t)$. The key idea now is to define a transform which incorporates the mother wavelet as the kernel. Thus we define the *continuous wavelet transform (CWT)*:

$$\mathcal{W}_\psi[f](a, b) = (f, \psi_{a,b}) = \int_{-\infty}^{\infty} f(t) \bar{\psi}_{a,b}(t) dt \quad (13.5.17)$$

Much like the windowed Fourier transform, the CWT is a function of the dilation parameter a and translation parameter b . Parenthetically, a wavelet is

admissible if the following property holds:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (13.5.18)$$

where the Fourier transform of the wavelet is defined by

$$\hat{\psi}_{a,b} = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} e^{-i\omega t} \psi\left(\frac{t-b}{a}\right) dt = \frac{1}{\sqrt{|a|}} e^{-ib\omega} \hat{\psi}(a\omega). \quad (13.5.19)$$

Thus provided the admissibility condition (13.5.18) is satisfied, the wavelet transform can be well defined.

As an example of the admissibility condition, consider the Haar wavelet (13.5.15). Its Fourier transform can be easily computed in terms of the sinc-like function:

$$\hat{\psi}(\omega) = ie^{-i\omega/2} \frac{\sin^2(\omega/4)}{\omega/4}. \quad (13.5.20)$$

Thus the admissibility constant can be computed to be

$$\int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega = 16 \int_{-\infty}^{\infty} \frac{1}{|\omega|^3} \left| \sin \frac{\omega}{4} \right|^4 d\omega < \infty. \quad (13.5.21)$$

This then shows that the Haar wavelet is in the admissible class.

Another interesting property of the wavelet transform is the ability to construct new wavelet bases. The following theorem is of particular importance.

Theorem: *If ψ is a wavelet and ϕ is a bounded integrable function, then the convolution $\psi \star \phi$ is a wavelet.*

In fact, from the Haar wavelet (13.5.15) we can construct new wavelet functions by convolving with for instance

$$\phi(t) = \begin{cases} 0 & t < 0 \\ 1 & 0 \leq t \leq 1 \\ 0 & t \geq 1 \end{cases} \quad (13.5.22)$$

or the function

$$\phi(t) = e^{-t^2}. \quad (13.5.23)$$

The convolutions of these functions ϕ with the Haar wavelet ψ (13.5.15) are produced in Fig. 121. These convolutions could also be used as mother wavelets in constructing a decomposition of a given signal or function.

The wavelet transform principle is quite simple. First, the signal is split up into a bunch of smaller signals by translating the wavelet with the parameter b over the entire time domain of the signal. Second, the same signal is processed

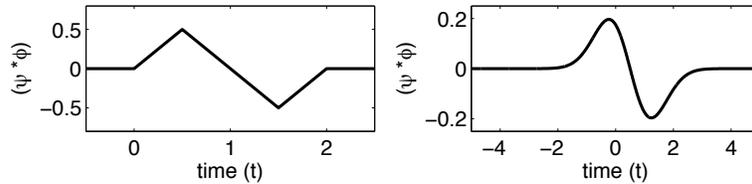


Figure 121: Convolution of the Haar wavelet with the functions (13.5.22) (left panel) and (13.5.23) (right panel). The convolved functions can be used as the mother wavelet for a wavelet basis expansion.

at different frequency bands, or resolutions, by scaling the wavelet window with the parameter a . The combination of translation and scaling allows for processing of the signals at different times and frequencies. Figure 121 is an upgrade of Fig. 116 that incorporates the wavelet transform concept in the time-frequency domain. In this figure, the standard time-series, Fourier transform and windowed Fourier transform are represented along with the multi-resolution concept of the wavelet transform. In particular, the box illustrating the wavelet transform shows the multi-resolution concept in action. Starting with a large Fourier domain window, the entire frequency content is extracted. The time window is then scaled in half, leading to finer time resolution at the expense of worse frequency resolution. This process is continued until a desired time-frequency resolution is obtained. This simple cartoon is critical for understanding wavelet application to time-frequency analysis.

Example: The Mexican Hat Wavelet. One of the more common wavelets is the Mexican hat wavelet. This wavelet is essentially a second moment of a Gaussian in the frequency domain. The definition of this wavelet and its transform are as follows:

$$\psi(t) = (1 - t^2)e^{-t^2/2} = -d^2/dt^2 \left(e^{-t^2/2} \right) = \psi_{1,0} \quad (13.5.24a)$$

$$\hat{\psi}(\omega) = \hat{\psi}_{1,0}(\omega) = \sqrt{2\pi}\omega^2 e^{-\omega^2/2} \quad (13.5.24b)$$

The Mexican hat wavelet has excellent localization properties in both time and frequency due to the minimal time-bandwidth product of the Gaussian function. Figure 123 (top panels) shows the basic Mexican wavelet function $\psi_{1,0}$ and its Fourier transform, both of which decay in t (ω) like $\exp(-t^2)$ ($\exp(-\omega^2)$). The Mexican hat wavelet can be dilated and translated easily as is depicted in Fig. 123 (bottom panel). Here three wavelets are depicted: $\psi_{1,0}$, $\psi_{3/2,-3}$ and $\psi_{1/4,6}$. This shows both the scaling and translation properties associated with any wavelet function.

To finish the initial discussion of wavelets, some of the various properties of

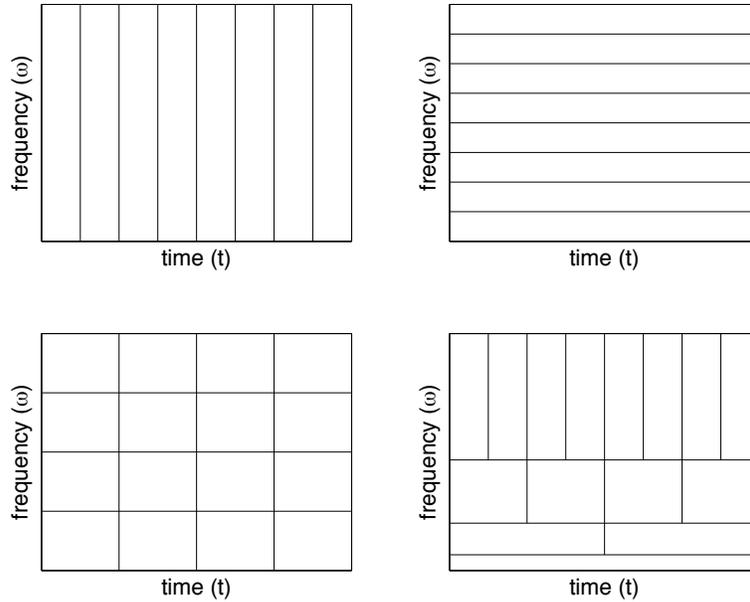


Figure 122: Graphical depiction of the difference between a time series analysis, Fourier analysis, Gabor analysis and wavelet analysis of a signal. This figure is identical to Fig. 116 but with the inclusion of the time-frequency resolution achieved with the wavelet transform. The wavelet transform starts with a large Fourier domain window so that the entire frequency content is extracted. The time window is then scaled in half, leading to finer time resolution at the expense of worse frequency resolution. This process is continued until a desired time-frequency resolution is obtained.

the wavelets are listed. To begin, consider the time-frequency resolution and its localization around a given time and frequency. These quantities can be calculated from the relations:

$$\sigma_t^2 = \int_{-\infty}^{\infty} (t - \langle t \rangle)^2 |\psi(t)|^2 dt \quad (13.5.25a)$$

$$\sigma_\omega^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} (\omega - \langle \omega \rangle)^2 |\hat{\psi}(\omega)|^2 d\omega \quad (13.5.25b)$$

where the variances measure the spread of the time and frequency signal around $\langle t \rangle$ and $\langle \omega \rangle$ respectively. The Heisenberg uncertainty constrains the localization of time and frequency by the relation $\sigma_t^2 \sigma_\omega^2 \geq 1/2$. In addition, the CWT has the following mathematical properties

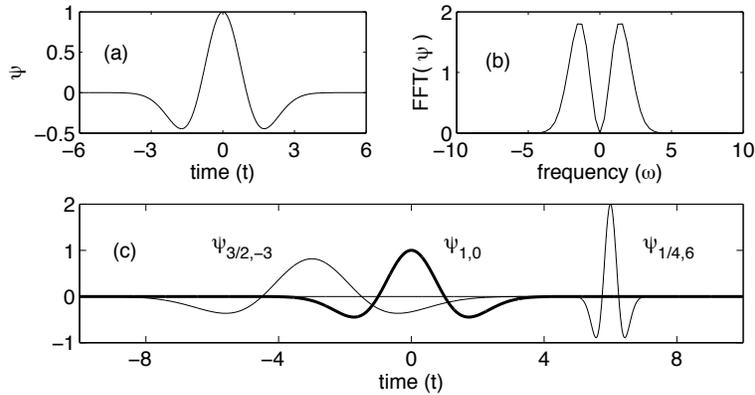


Figure 123: Illustration of the Mexican hat wavelet $\psi_{1,0}$ (top left panel), its Fourier transform $\hat{\psi}_{1,0}$ (top right panel), and two additional dilations and translations of the basic $\psi_{1,0}$ wavelet. Namely the $\psi_{3/2,-3}$ and $\psi_{1/4,6}$ are shown (bottom panel).

1. **Linearity:** the transform is linear so that

$$\mathcal{W}_\psi(\alpha f + \beta g)(a, b) = \alpha \mathcal{W}_\psi(f)(a, b) + \beta \mathcal{W}_\psi(g)(a, b)$$

2. **Translation:** the transform has the translation property

$$\mathcal{W}_\psi(T_c f)(a, b) = \mathcal{W}_\psi(f)(a, b - c)$$

where $T_c f(t) = f(t - c)$.

3. **Dilation:** the dilation property follows

$$\mathcal{W}_\psi(D_c f)(a, b) = \frac{1}{\sqrt{c}} \mathcal{W}_\psi(f)(a/c, b/c)$$

where $c > 0$ and $D_c f(t) = (1/c)f(t/c)$.

4. **Inversion:** the transform can be inverted with the definition

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{W}_\psi(f)(a, b) \psi_{a,b}(t) \frac{db da}{a^2}$$

where it becomes clear why the admissibility condition $C_\psi < \infty$ is needed.

To conclude this section, consider the idea of discretizing the wavelet transform on a computational grid. Thus the transform is defined on a lattice so that

$$\psi_{m,n}(x) = a_0^{-m/2} \psi(a_0^{-m} x - nb_0) \quad (13.5.26)$$

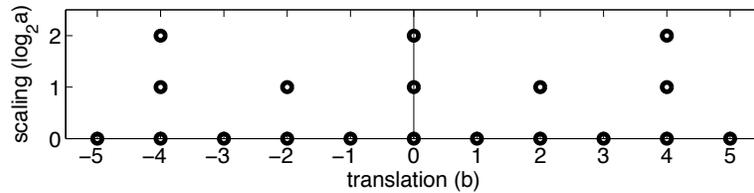


Figure 124: Discretization of the discrete wavelet transform with $a_0 = 2$ and $b_0 = 1$. This figure is a more formal depiction of the multi-resolution discretization as shown in Fig. 122.

where $a_0, b_0 > 0$ and m, n are integers. The *discrete wavelet transform* is then defined by

$$\begin{aligned} \mathcal{W}_\psi(f)(m, n) &= (f, \psi_{m,n}) \\ &= \int_{-\infty}^{\infty} f(t) \bar{\psi}_{m,n}(t) dt \\ &= a_0^{-m/2} \int_{-\infty}^{\infty} f(t) \bar{\psi}(a_0^{-m}t - nb_0) dt. \end{aligned} \quad (13.5.27)$$

Futhermore, if $\psi_{m,n}$ are complete, then a given signal or function can be expanded in the wavelet basis:

$$f(t) = \sum_{m,n=-\infty}^{\infty} (f, \psi_{m,n}) \psi_{m,n}(t). \quad (13.5.28)$$

This expansion is in a set of wavelet frames. It still needs to be determined if the expansion is in terms of a set of basis functions. It should be noted that the scaling and dilation parameters are typically taken to be $a_0 = 2$ and $b_0 = 1$, corresponding to dilations of 2^{-m} and translations of $n2^m$. Figure 124 gives a graphical depiction of the time-frequency discretization of the wavelet transform. This figure is especially relevant for the computational evaluation of the wavelet transform. Further, it is the basis of fast algorithms for multi-resolution analysis.

13.6 Multi-Resolution Analysis and the Wavelet Basis

Before proceeding forward with wavelets, it is important to establish some key mathematical properties. Indeed, the most important issue to resolve is the ability of the wavelet to actually represent a given signal or function. In Fourier analysis, it has been established that any generic function can be represented by a series of cosines and sines. Something similar is needed for wavelets in order to make them a useful tool for the analysis of time-frequency signals.

The concept of a wavelet is simple and intuitive: construct a signal using a single function $\psi \in L^2$ which can be written $\psi_{m,n}$ and that represents binary dilations by 2^m and translations of $n2^{-m}$ so that

$$\psi_{m,n} = 2^{m/2}\psi(2^m(x - n/2^m)) = 2^{m/2}\psi(2^m x - n) \quad (13.6.29)$$

where m and n are integers. The use of this wavelet for representing a given signal or function is simple enough. However, there is a critical issue to be resolved concerning the *orthogonality* of the functions $\psi_{m,n}$. Ultimately, this is the primary issue which must be addressed in order to consider the wavelets as basis functions in an expansion. Thus we define the orthogonality condition as

$$(\psi_{m,n}, \psi_{k,l}) = \int_{-\infty}^{\infty} \psi_{m,n}(x)\psi_{k,l}(x)dx = \delta_{m,k}\delta_{n,l} \quad (13.6.30)$$

where δ_{ij} is the Dirac delta defined by

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \quad (13.6.31)$$

where i, j are integers. This statement of orthogonality is generic, and it holds in most function spaces with a defined inner product.

The importance of orthogonality should not be underestimated. It is very important in applications where a functional expansion is used to approximate a given function or solution. In what follows, two examples are given concerning the key role of orthogonality.

Fourier expansions. Consider representing an even function $f(x)$ over the domain $x \in [0, L]$ with a cosine expansion basis. By Fourier theory, the function can be represented by

$$f(x) = \sum_{n=0}^{\infty} a_n \cos \frac{n\pi x}{L} \quad (13.6.32)$$

where the coefficients a_n can be constructed by using inner product rules and orthogonality. Specifically, by multiplying both sides of the equation by $\cos(m\pi x/L)$ and integrating over $x \in [0, L]$, i.e. taking the inner product with respect to $\cos(m\pi x/L)$, the following result is found:

$$(f, \cos m\pi x/L) = \sum_{n=0}^{\infty} a_n (\cos n\pi x/L, \cos m\pi x/L). \quad (13.6.33)$$

This is where orthogonality plays a key role, the infinite sum on the right hand side can be reduced to a single index where $n = m$ since the cosines are orthogonal to each other

$$(\cos n\pi x/L, \cos m\pi x/L) = \begin{cases} 0 & n \neq m \\ L & n = m \end{cases}. \quad (13.6.34)$$

Thus the coefficients can be computed to be

$$a_n = \frac{1}{L} \int_0^L f(x) \cos \frac{n\pi x}{L} dx \quad (13.6.35)$$

and the expansion is accomplished. Moreover, the cosine basis is complete for even functions and any signal or function $f(x)$ can be represented, i.e. as $n \rightarrow \infty$ in the sum, the expansion converges to the given signal $f(x)$.

Eigenfunction expansions: The cosine expansion is a subset of the more general eigenfunction expansion technique that is often used to solve differential and partial differential equations problems. Consider the nonhomogeneous boundary value problem

$$Lu = f(x) \quad (13.6.36)$$

where L is a given self-adjoint, linear operator. This problem can be solved with an eigenfunction expansion technique by considering the associated eigenvalue problem of the operator L :

$$Lu_n = \lambda_n u_n. \quad (13.6.37)$$

The solution of (13.6.36) can then be expressed as

$$u(x) = \sum_{n=0}^{\infty} a_n u_n \quad (13.6.38)$$

provided the coefficients a_n can be determined. Plugging in this solution to (13.6.36) yields the following calculations

$$\begin{aligned} Lu &= f \\ L(\sum a_n u_n) &= f \\ \sum a_n Lu_n &= f \\ \sum a_n \lambda_n u_n &= f. \end{aligned} \quad (13.6.39)$$

Taking the inner product of both sides with respect to u_m yields

$$\begin{aligned} (\sum a_n \lambda_n u_n, u_m) &= (f, u_m) \\ \sum a_n \lambda_n (u_n, u_m) &= (f, u_m) \\ a_m \lambda_m &= (f, u_m) \end{aligned} \quad (13.6.40)$$

where the last line is achieved by orthogonality of the eigenfunctions $(u_n, u_m) = \delta_{n,m}$. This then gives $a_m = (f, u_m)/\lambda_m$ and the eigenfunction expansion solution is

$$u(x) = \sum_{n=0}^{\infty} \frac{(f, u_n)}{\lambda_n} u_n. \quad (13.6.41)$$

Provided the u_n are a complete set, this expansion is guaranteed to converge to $u(x)$ as $n \rightarrow \infty$.

Orthonormal wavelets

The preceding examples highlight the importance of orthogonality for representing a given function. A wavelet ψ is called orthogonal if the family of functions $\psi_{m,n}$ are orthogonal as given by Eq. (13.6.30). In this case, a given signal or function can be uniquely expressed with the doubly infinite series

$$f(t) = \sum_{n,m=-\infty}^{\infty} c_{m,n} \psi_{m,n}(t) \quad (13.6.42)$$

where the coefficients are given from orthogonality by

$$c_{m,n} = (f, \psi_{m,n}). \quad (13.6.43)$$

The series is guaranteed to converge to $f(t)$ in the L^2 norm.

The above result based upon orthogonal wavelets establishes the key mathematical framework needed for using wavelets in a very broad and general way. It is this result that allows us to think of wavelets philosophically as the same as the Fourier transform.

Multi-Resolution Analysis (MRA)

The power of the wavelet basis is its ability to take a function or signal $f(t)$ and express it as a limit of successive approximations, each of which is a finer and finer version of the function in time. These successive approximations correspond to different resolution levels.

A *multi-resolution analysis*, commonly referred to as an MRA, is a method that gives a formal approach to constructing the signal with different resolution levels. Mathematically, this involves a sequence

$$\{V_m : m \in \text{integers}\} \quad (13.6.44)$$

of embedded subspaces of L^2 that satisfies the following relations:

1. The subspaces can be embedded in each other so that

$$\text{Course} \cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \cdots V_m \subset V_{m+1} \cdots \text{Fine}$$

2. The union of all the embedded subspaces spans the entire L^2 space so that

$$\cup_{m=-\infty}^{\infty} V_m$$

is dense in L^2

3. The intersection of subspaces is the null set so that

$$\cap_{m=-\infty}^{\infty} V_m = \{0\}$$

4. Each subspace picks up a given resolution so that $f(x) \in V_m$ if and only if $f(2x) \in V_{m+1}$ for all integers m .
5. There exists a function $\phi \in V_0$ such that

$$\{\phi_{0,n} = \phi(x - n)\}$$

is an orthogonal basis for V_0 so that

$$\|f\|^2 = \int_{-\infty}^{\infty} |f(x)|^2 dx = \sum_{-\infty}^{\infty} |(f, \phi_{0,n})|^2$$

The function ϕ is called the *scaling function* or *father wavelet*.

If $\{V_m\}$ is a multiresolution of L^2 and if V_0 is the closed subspace generated by the integer translates of a single function ϕ , then we say ϕ generates the MRA.

One remark of importance: since $V_0 \subset V_1$ and ϕ is a scaling function for V_0 and also for V_1 , then

$$\phi(x) = \sum_{-\infty}^{\infty} c_n \phi_{1,n}(x) = \sqrt{2} \sum_{-\infty}^{\infty} c_n \phi(2x - n) \quad (13.6.45)$$

where $c_n = (\phi, \phi_{1,n})$ and $\sum_{-\infty}^{\infty} |c_n|^2 = 1$. This equation, which relates the scaling function as a function of x and $2x$ is known as the *dilation equation*, or *two-scale equation*, or *refinement equation* because it reflects $\phi(x)$ in the refined space V_1 which as the finer scale of 2^{-1} .

Since $V_m \subset V_{m+1}$, we can define the orthogonal complement of V_m in V_{m+1} as

$$V_{m+1} = V_m \oplus W_m \quad (13.6.46)$$

where $V_m \perp W_m$. This can be generalized so that

$$\begin{aligned} V_{m+1} &= V_m \oplus W_m \\ &= (V_{m-1} \oplus W_{m-1}) \oplus W_m \\ &\vdots \\ &= V_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_m \\ &= V_0 \oplus (\oplus_{n=0}^m W_n) . \end{aligned} \quad (13.6.47)$$

As $m \rightarrow \infty$, it can be found that

$$V_0 \oplus (\oplus_{n=0}^{\infty} W_n) = L^2 . \quad (13.6.48)$$

In a similar fashion, the resolution can rescale upwards so that

$$\oplus_{n=-\infty}^{\infty} W_n = L^2 . \quad (13.6.49)$$

Moreover, there exists a scaling function $\psi \in W_0$ (the mother wavelet) such that

$$\psi_{0,n}(x) = \psi(x - n) \quad (13.6.50)$$

constitutes an orthogonal basis for W_0 and

$$\psi_{m,n}(x) = 2^{m/2}\psi(2^m x - n) \quad (13.6.51)$$

is an orthogonal basis for W_m . Thus the mother wavelet ψ spans the orthogonal complement subset W_m while the scaling function ϕ spans the subsets V_m . The connection between the father and mother wavelet are shown in the following theorem.

Theorem: If $\{V_m\}$ is a MRA with scaling function ϕ , then there is a mother wavelet ψ

$$\psi(x) = \sqrt{2} \sum_{-\infty}^{\infty} (-1)^{n-1} \bar{c}_{-n-1} \phi(2x - n) \quad (13.6.52)$$

where

$$c_n = (\phi, \phi_{1,n}) = \sqrt{2} \int_{-\infty}^{\infty} \phi(x) \bar{\phi}(2x - 1) dx. \quad (13.6.53)$$

That is, the system $\psi_{m,n}(x)$ is an orthogonal basis of L^2 .

This theorem is critical for what we would like to do. Namely, use the wavelet basis functions as a complete expansion basis for a given function $f(x)$ in L^2 . Further, it explicitly states the connection between the *scaling function* $\phi(x)$ (father wavelet) and *wavelet function* $\psi(x)$ (mother wavelet). It is only left to construct a desirable wavelet basis to use. As for wavelet construction, the idea is to build them to take advantage of certain properties of the system so that it gives an efficient and meaningful representation of your time-frequency data.

13.7 Spectrograms and the Gábor transforms in MATLAB

The aim of this lecture will be to use MATLAB's fast Fourier transform routines modified to handle the Gábor transform. The Gábor transform allows for a fast and easy way to analyze both the time and frequency properties of a given signal. Indeed, this windowed Fourier transform method is used extensively for analyzing speech and vocalization patterns. For such applications, it is typical to produce a *spectrogram* that represents the signal in both the time and frequency domain. Figures 125 and 126 are produced from the vocalization patterns in time-frequency of a human saying "do re mi" and a humpback whale vocalizing to other whales. The time-frequency analysis can be used to produce speech

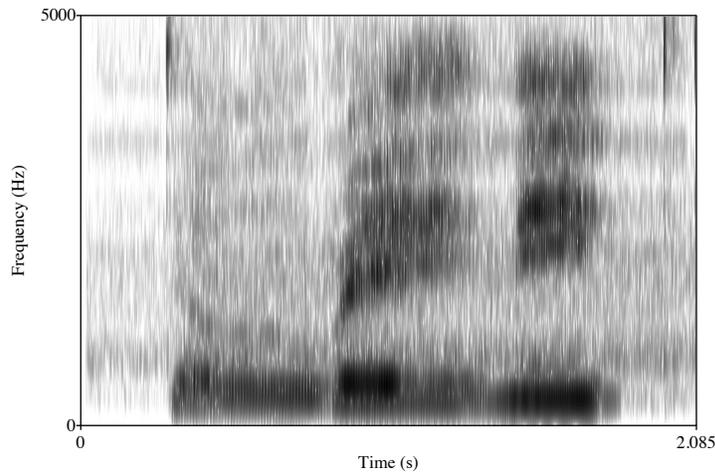


Figure 125: Spectrogram (time-frequency) analysis of a male human saying “do re mi.” The spectrogram is created with the software program Praat, which is an open source code for analyzing phonetics.

recognition algorithms given the characteristic signatures in the time-frequency domains of sounds. Thus spectrograms are a sort of fingerprint of sound.

To understand the algorithms which produce the spectrogram, it is informative to return to the characteristic picture shown in Fig. 115. This demonstrates the action of an applied time filter in extracting time localization information. To build a specific example, consider the following MATLAB code that builds a time domain (t), its corresponding Fourier domain (ω), a relatively complicated signal ($S(t)$), and its Fourier transform ($\hat{S}(\omega)$).

```
clear all; close all; clc

L=10; n=2048;
t2=linspace(0,L,n+1); t=t2(1:n);
k=(2*pi/L)*[0:n/2-1 -n/2:-1]; ks=fftshift(k);

S=(3*sin(2*t)+0.5*tanh(0.5*(t-3))+ 0.2*exp(-(t-4).^2)...
+1.5*sin(5*t)+4*cos(3*(t-6).^2))/10+(t/20).^3;
St=fft(S);
```

The signal and its Fourier transform can be plotted with the commands

```
figure(1)
subplot(3,1,1) % Time domain
```

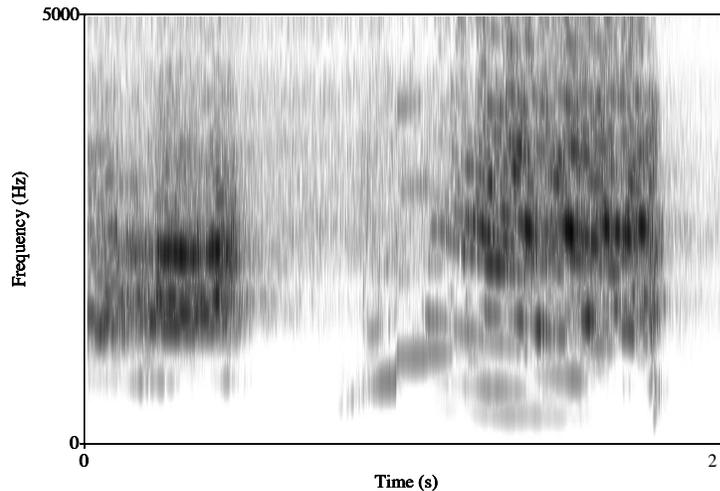


Figure 126: Spectrogram (time-frequency) analysis of a humpback whale vocalization over a short period of time. The spectrogram is created with the software program Praat, which is an open source code for analyzing phonetics.

```

plot(t,S,'k')
set(gca,'FontSize',[14]),
xlabel('Time (t)'), ylabel('S(t)')

subplot(3,1,2) % Fourier domain
plot(ks,abs(fftshift(St))/max(abs(St)),'k');
axis([-50 50 0 1])
set(gca,'FontSize',[14])
xlabel('frequency (\omega)'), ylabel('FFT(S)')

```

Figure 127 shows the signal and its Fourier transform for the above example. This signal $S(t)$ will be analyzed using the Gábor transform method.

The simplest Gábor window to implement is a Gaussian time-filter centered at some time τ with width a . As has been demonstrated, the parameter a is critical for determining the level of time-resolution versus frequency resolution in a time-frequency plot. Figure 128 shows the signal under consideration with three filter widths. The narrower the time-filtering, the better resolution in time. However, this also produces the worst resolution in frequency. Conversely, a

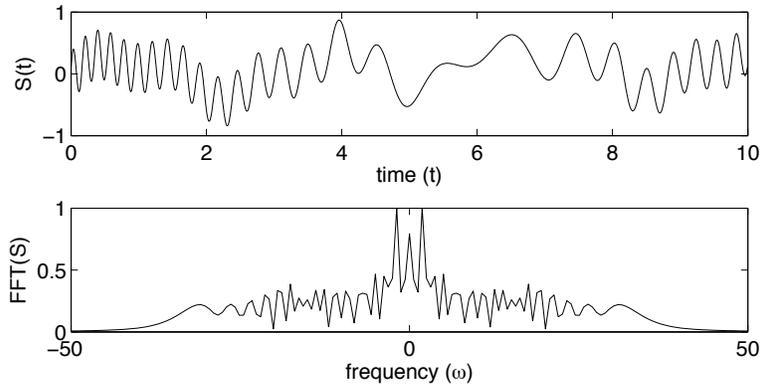


Figure 127: Time signal and its Fourier transform considered for a time-frequency analysis in what follows.

wide window in time produces much better frequency resolution at the expense of reducing the time resolution. A simple extension to the existing code produces a signal plot along with three different filter widths of Gaussian shape.

```

figure(2)
width=[10 1 0.2];
for j=1:3
    g=exp(-width(j)*(t-4).^2);
    subplot(3,1,j)
    plot(t,S,'k'), hold on
    plot(t,g,'k','Linewidth',[2])
    set(gca,'FontSize',[14])
    ylabel('S(t), g(t)')
end
xlabel('time (t)')

```

The key now for the Gábor transform is to multiply the time filter Gábor function $g(t)$ with the original signal $S(t)$ in order to produce a windowed section of the signal. The Fourier transform of the windowed section then gives the local frequency content in time. The following code constructs the windowed Fourier transform with the Gábor filtering function

$$g(t) = e^{-a(t-b)^2}. \quad (13.7.54)$$

The Gaussian filtering has a width parameter a and translation parameter b . The following code constructs the windowed Fourier transform using the Gaussian with $a = 2$ and $b = 4$.

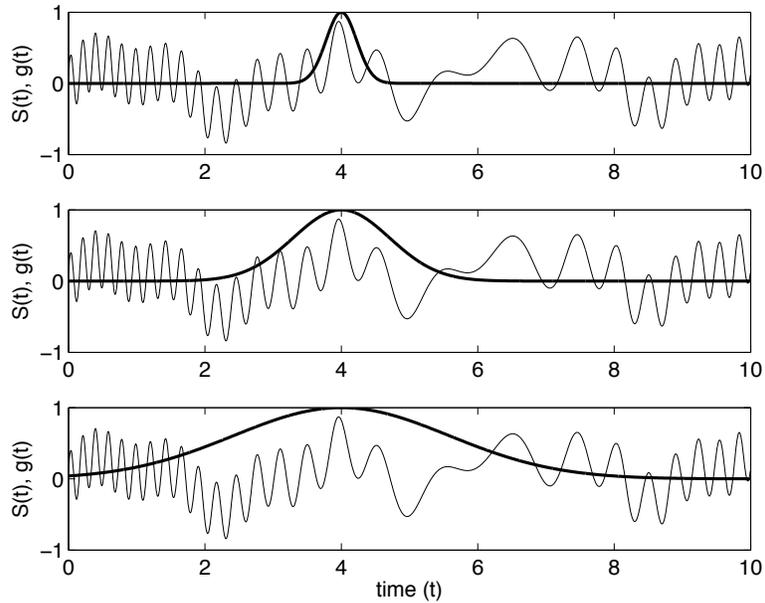


Figure 128: Time signal $S(t)$ and the Gábor time filter $g(t)$ (bold lines) for three different Gaussian filters: $g(t) = \exp(-10(x - 4)^2)$ (top), $g(t) = \exp(-(x - 4)^2)$ (middle), and $g(t) = \exp(-0.2(x - 4)^2)$ (bottom). The different filter widths determine the time-frequency resolution. Better time resolution gives worse frequency resolution and vice-versa due to the Heisenberg uncertainty principle.

```

figure(3)
g=exp(-2*(t-4).^2);
Sg=g.*S;
Sgt=fft(Sg);

subplot(3,1,1), plot(t,S,'k'), hold on
    plot(t,g,'k','Linewidth',[2])
    set(gca,'FontSize',[14])
    ylabel('S(t), g(t)'), xlabel('time (t)')
subplot(3,1,2), plot(t,Sg,'k')
    set(gca,'FontSize',[14])
    ylabel('S(t)g(t)'), xlabel('time (t)')
subplot(3,1,3), plot(kS,abs(fftshift(Sgt))/max(abs(Sgt)),'k')
    axis([-50 50 0 1])
    set(gca,'FontSize',[14])
    ylabel('FFT(Sg)'), xlabel('frequency (\omega)')

```

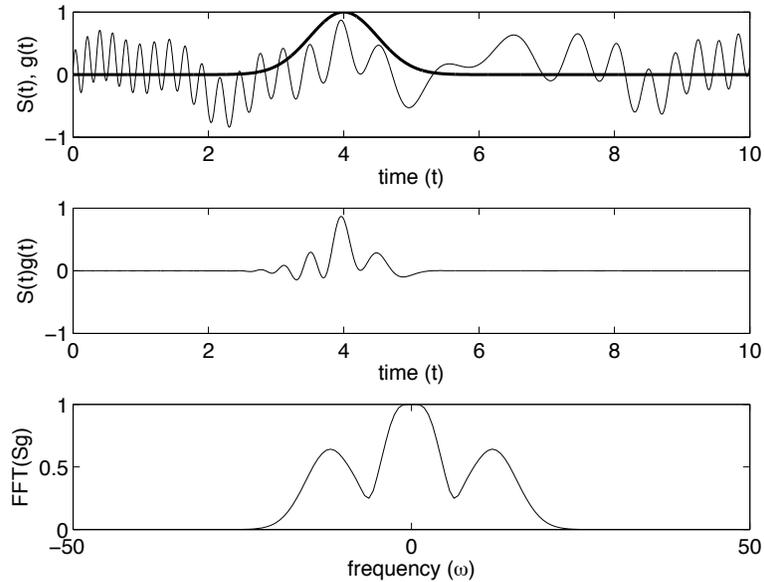


Figure 129: Time signal $S(t)$ and the Gábor time filter $g(t) = \exp(-2(x - 4)^2)$ (bold line) for a Gaussian filter. The product $S(t)g(t)$ is depicted in the middle panel and its Fourier transform $\hat{S}g(\omega)$ is depicted in the bottom panel. Note that the windowing of the Fourier transform can severely limit the detection of low-frequency components.

Figure 129 demonstrates the application of this code and the windowed Fourier transform in extracting local frequencies of a local time window.

The key to generating a spectrogram is to now vary the position b of the time filter and produce spectra at each location in time. In theory, the parameter b is continuously translated to produce the time-frequency picture. In practice, like everything else, the parameter b is discretized. The level of discretization is important in establishing a good time-frequency analysis. Specifically, finer resolution will produce better results. The following code makes a dynamical *movie* of this process as the parameter b is translated.

```
figure(4)
Sgt_spec=[];
tslide=0:0.1:10
for j=1:length(tslide)
    g=exp(-2*(t-tslide(j)).^2); % Gabor
    Sg=g.*S; Sgt=fft(Sg);
    Sgt_spec=[Sgt_spec; abs(fftshift(Sgt))];
    subplot(3,1,1), plot(t,S,'k',t,g,'r')
```

```

subplot(3,1,2), plot(t,Sg,'k')
subplot(3,1,3), plot(ks,abs(fftshift(Sgt))/max(abs(Sgt)))
axis([-50 50 0 1])
drawnow
pause(0.1)
end

```

This movie is particularly illustrative and provides an excellent graphical representation of how the Gábor time-filtering extracts both local time information and local frequency content. It also illustrates, as the parameter a is adjusted, the ability (or inability) of the windowed Fourier transform to provide accurate time and frequency information.

The code just developed also produces a matrix **Sgt_spec** which contains the Fourier transform at each slice in time of the parameter b . It is this matrix that produces the spectrogram of the time-frequency signal. The spectrogram can be viewed with the commands

```

pcolor(tslide,ks,Sgt_spec. '), shading interp
set(gca, 'Ylim', [-50 50], 'FontSize', [14])
colormap(hot)

```

Modifying the code slightly, a spectrogram of the signal $S(t)$ can be made for three different filter widths $a = 5, 1, 0.2$ in Eq. (13.7.54). The spectrograms are shown in Fig. 130 where from left to right the filtering window is broadened from $a = 5$ to $a = 0.2$. Note that for the left plot, strong localization of the signal in time is achieved at the expense of suppressing almost all the low-frequency components of the signal. In contrast, the right most figure with a wide temporal filter preserves excellent resolution of the Fourier domain but fails to localize signals in time. Such are the tradeoffs associated with a fixed Gábor window transform.

13.8 MATLAB Filter Design and Wavelet Toolboxes

The applications of filtering and time-frequency analysis are so ubiquitous across the sciences, that MATLAB has developed a suite of toolboxes that specialize in these applications. Two of these toolboxes will be demonstrated in what follows. Primarily, screenshots will give hints of the functionality and versatility of the toolboxes.

The most remarkable part of the toolbox is that it allows for entry into high level signal processing and wavelet processing almost immediately. Indeed, one hardly needs to know anything to begin the process of analyzing, synthesizing, and manipulating data to one's own ends. For the most part, each of the toolboxes allows you to begin usage once you upload your signal, image or data. The only drawback is cost. For the most part, many academic departments have access to the toolboxes. And if you are part of a university environment,