M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

Chapter 9

Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform

M. C. Teich, C. Heneghan, S. M. Khanna

9.1. INTRODUCTION

In the process of hearing, sound waves travel to the eardrum (tympanic membrane) through the external ear and ear canal. The sound pressure acting on the tympanic membrane produces mechanical vibrations that are transmitted, via the ossicular chain in the middle ear, to the inner ear (cochlea). The cochlea, which is encased in a bony shell, consists of three fluid-filled canals: scala vestibuli, scala media, and scala tympani. A thin membrane (Reissner's membrane), running the length of the cochlea, separates the scala vestibuli from the scala media (middle canal). The basilar membrane forms the base of the middle canal, separating it from the scala tympani. The cochlea is coiled; it has a diameter that is widest at the base and narrowest at the apex. There are two openings in the bony shell near the base: (1) the oval window, through which the stapes drives the fluid in the scala vestibuli, and (2) the round window, which is covered by a thin membrane that accommodates the movement of fluid in the cochlea. The sensory organ of hearing (the organ of Corti) is located on the scala media side of the basilar membrane. It consists of several types of specialized cells that are organized in precise transverse and longitudinal arrangements. The transverse morphological arrangement is the same from base to apex, though the width and stiffness of the basilar membrane and the dimensions of most of the cells change over this region [1,2].

The velocity of vibration of individual cells, selected as desired, can be measured with a specially designed confocal heterodyne interferometer in response to sound applied to the ear canal. The details of the stimulus-generation and measuring

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

techniques are described elsewhere [2]. We have devoted a good deal of attention to describing the character of the dynamical vibrations of individual sensory cells in the third and fourth turns of the guinea-pig cochlea using a temporal-bone preparation excised from a freshly sacrificed animal and maintained in oxygenated tissue-culture medium [3–13]. In this chapter, we turn to a similar analysis of the recordings from the basal cochlear turn of the living cat [14].

A suitable method for studying level-dependent changes in the dynamical response of these cells is to use sinusoidal-carrier amplitude-modulated (AM) acoustic waves with low-frequency modulation. This provides an opportunity for studying the change in nonlinear dynamical response over a broad range of carrier levels, as the envelope slowly increases and decreases. The AM format is also useful because the heterodyne interferometer can measure the velocity of an object but not its absolute position.

This chapter describes the application of time–scale and time–frequency representation techniques in the analysis of cellular velocity data. The relative advantages of the two techniques are compared for several data sets. The modulation depth of the AM acoustic signal was unity and the modulation frequency was 6.1 Hz. The carrier frequency ranged from 1000 Hz to 40000 Hz, and the total duration of each data set collected was fixed at 0.16384 s (representing 16384 samples at 10-μs intervals).

In general, we also record the cellular vibration in the absence of acoustic stimuli. Using this technique we have already established the presence of spontaneous vibrations in the third turn of the guinea-pig cochlea [8,9]. We have seen similar spontaneous vibrations in the basal turn of the cat cochlea [15]. However, since the amplitude of the spontaneous components is approximately time-invariant over the period of recording, there is no particular advantage to using time-frequency or time-scale methods as compared to a purely spectral technique such as the periodogram. Accordingly, we will not discuss these results further in this chapter.

We examined both the continuous wavelet transform (CWT) and the short-time Fourier transform (STFT) of the velocity responses elicited by the AM stimuli just described. Both analysis techniques were useful in discriminating the frequency components present in the responses, though the wavelet basis for the CWT had to be carefully chosen to provide the desired frequency resolution. CWTs using a high-Q Morlet wavelet basis were found to be particularly useful for discriminating among the various response components. Octave-band–based CWTs (using low-Q Morlet, Meyer, and Daubechies 4-tap wavelets) were largely ineffective in analyzing these signals, inasmuch as their frequency resolution was too poor to distinguish among the frequency components present in the velocity responses.

9.2. METHODS

Optical access to the basal turn of the cochlea was obtained by carefully opening the cat's bulla. The measurements shown here were made after the removal of the round-window membrane. A Nikon SLWD 20X objective lens with a 19.9-mm working distance was used to view the cochlea with an optical sectioning microscope. A laser interferometer was coupled with this optical sectioning microscope, in such a way that

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

their planes of focus coincided. This allowed identification of the structure lying at the focus of the microscope/interferometer. The 2-μm-diameter laser spot on the object, seen through the microscope eyepiece, was placed on a selected cell by moving the animal's head with an x-y-z micropositioning system [2]. The sectioning depth was 10 μm so that vibrations were recorded from small regions of single cells. The interferometer measures the vibration of the object on which it is focused [2].

The measurements were carried out in a soundproof chamber. Sound generated by a high-fidelity acoustic driver was applied to the ear canal. The electrical signal applied to the acoustic driver was generated using a 486-microprocessor–based computer system coupled to a 16-bit D/A converter. The signal from the interferometer was passed through an antialiasing filter, a 16-bit A/D converter, and then stored in the hard disk of the computer system. The same system can be used to measure cellular vibration in the absence of a stimulus by simply disconnecting the sound stimulus system.

9.3. THEORY

9.3.1 The Continuous-Time Fourier Transform

Signal analysis addresses the problem of extracting information from a given signal $x(t)$ and converting it into a recognizable form. One approach to this problem is to transform $x(t)$, using an information-preserving mapping, to a different domain (viz., a dual domain), where it is easier to interpret the signal. The best-known of these approaches, perhaps, is the continuous-time Fourier transform (CFT). The relations between $x(t)$ (where t is chosen to represent time) and its dual representation $X(f)$ are

$$X(f) = \int_{-\infty}^{\infty} x(t) \exp(-j2\pi ft)\, dt \tag{9-1}$$

$$x(t) = \int_{-\infty}^{\infty} X(f) \exp(j2\pi ft)\, df \tag{9-2}$$

$X(f)$ is referred to as a spectral representation of $x(t)$, with the dual variable f defined as (global) frequency.

The CFT reveals how the energy in the signal $x(t)$ is distributed in frequency. One limitation of the CFT is that the value of $X(f)$ is affected by all values of $x(t)$ from $t = -\infty$ to $+\infty$. As a result, any particular feature in $X(f)$ cannot be linked with a specific time region of $x(t)$. The CFT provides a totally global perspective on how a signal's energy is distributed as a function of frequency; in other words, $X(f)$ is a completely nonlocal spectral representation of $x(t)$.

9.3.2 The Short-Time Fourier Transform

In many cases, the CFT is a most useful representation, particularly if $x(t)$ is stationary, or "steady," in time (for a discussion of the notion of stationarity, see

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

[16]). However, for many signals, the nature of $x(t)$ changes with time. For example, if $x(t)$ represents the vertical motion of a point on the wheel of an automobile over the course of a journey, its frequency of motion will vary depending on whether the car is moving at constant speed, accelerating, decelerating, or at rest. The CFT of this signal tells us about the range of frequencies that the motion achieved over the course of the entire journey, but it fails to provide information about the time order in which they occurred. A representation of $x(t)$ that tells us *when* a particular frequency was present provides a more useful account of the journey. To achieve this, a spectral representation that includes some explicit dependence on time is needed; a function of the form $X(f, \tau)$, where f again represents frequency and τ represents time, would be useful.

The first attempt to construct such a function was carried out by Gabor in 1946 [17]. His approach retained the frequency variable f defined by the CFT, but ensured that only values of $x(t)$ in the near vicinity of $t = \tau$ would be able to influence $X(f, \tau)$. He achieved this by multiplying the original signal $x(t)$ by window functions that are localized in time at $t = \tau$. In this way, he constructed a local spectral representation of the signal in the vicinity of time τ.

Gabor's approach was later shown to be a special case of the STFT [18,19], which is expressed as

$$\mathrm{STFT}_x^g(f, \tau) = \int_{-\infty}^{\infty} x(t) g^*(t - \tau) \exp(-j2\pi f t)\, dt \tag{9-3}$$

Here, t and τ are time variables, $x(t)$ represents the time waveform being analyzed, $g(t)$ represents a window function, f is the frequency variable, and the superscript star denotes complex conjugation. The Gaussian $[g(t) = \exp(-\beta t^2/2)]$ is a typical choice for the window function since it falls smoothly and symmetrically to zero around the time $t = 0$. Accordingly, the function $g^*(t - \tau)$, which in this case is equal to $g(t - \tau)$, is centered about the time $t = \tau$, and falls away quickly to zero for times away from τ. It is apparent that in the absence of a window function $[g(t) = 1]$, the STFT in Eq. (9-3) reduces to the CFT given in Eq. (9-1).

There are several alternative ways of expressing the STFT that are useful in different circumstances. For example, it can be written as an integral in the frequency domain, viz.,

$$\mathrm{STFT}_x^g(f, \tau) = \exp(-j2\pi f \tau) \int_{-\infty}^{\infty} X(u) G^*(u - f) \exp(j2\pi u \tau)\, du \tag{9-4}$$

where $X(u)$ and $G(u)$ represent the CFTs of $x(t)$ and $g(t)$, respectively, and u is a dummy frequency variable. Here, $\mathrm{STFT}_x^g(f, \tau)$ is seen to be a frequency-shifted version of the inverse CFT of $[X(u)G^*(u - f)]$ [compare with Eq. (9-2)]. If the function $G(u)$ is taken to represent a low-pass filter in frequency, then $X(u)G^*(u - f)$ is the CFT of $x(t)$ after filtering by a bandpass filter whose shape is simply that of $G(u)$, translated in frequency so that it is centered about f instead of 0. The factor $\exp(-j2\pi f \tau)$ in Eq. (9-4) simply frequency shifts the filtered output back down to zero frequency. The STFT can thus be viewed as the frequency-shifted output from a bank of filters $G^*(u - f)$, each with constant bandwidth but different center

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

frequency. This filter-bank interpretation of the STFT [20–22] is illustrated in Fig. 9-1(a).

The STFT can also be written as a convolution in either the time domain:

$$\mathrm{STFT}_x^g(f,\tau) = \exp(-j2\pi f\tau)[x(\tau) * g^*(-\tau)\exp(j2\pi f\tau)] \tag{9-5}$$

or in the frequency domain:

$$\mathrm{STFT}_x^g(f,\tau) = X(f) * [G^*(-f)\exp(-j2\pi f\tau)] \tag{9-6}$$

where * denotes the convolution integral operator [i.e., $u(s) = \int_{-\infty}^{+\infty} v(r)w(s-r)dr \equiv v(s) * w(s)$]. The convolution formalism represented in Eq. (9-5) identifies a limitation of the STFT. In calculating $STFT_x^g(f,\tau_0)$, the value of $x(t)$ at time $t = \tau_0$ is smeared over time by the convolution integral. Therefore, any sharp change in the value of $x(t)$ at time $t = \tau_0$ will not appear in the STFT solely at τ_0, but rather will be spread over a region of time in the vicinity of $\tau = \tau_0$. The range of time over which information is spread depends on the width (time duration) of $g^*(-t)\exp(j2\pi ft)$. (The width of a function can be defined in many ways [23]; in this chapter, we define it as the full-width at $1/e$-maximum, whether it be a time duration or a bandwidth. Alternative definitions of bandwidth would, of course, be acceptable for the transform properties we discuss provided they are dealt with consistently.)

According to Eq. (9-6), a sharp spectral feature at frequency $f = f_0$ will similarly be blurred by convolution with $G^*(-f)\exp(-j2\pi ft)$.

The ability to resolve fine features in either the time or frequency domains is referred to as the time or frequency resolution of the transform operation, respectively. From the discussion in the preceding paragraph, it is apparent that the time and frequency resolutions of the STFT are dependent on the widths of the functions $g(t)$ and $G(f)$. These widths are denoted by Δt and Δf, respectively, and it would be ideal if both of these quantities could go to zero.

This is not possible. To illustrate this, consider choosing $g(t)$ as short in time as possible. In the limit, we obtain $g(t) = \delta(t)$, where $\delta(t)$ is the Dirac delta function, which transforms to $G(f) = 1$. Substitution into Eq. (9-3) then shows that $\mathrm{STFT}_x^g(f,\tau) = x(\tau)\exp(-j2\pi f\tau)$, which is simply the original signal $x(t)$ translated down in frequency. This STFT has therefore exactly preserved the time information in the signal $x(t)$, but it provides no frequency information whatsoever. This is because the width of the function $G(f)$ is infinite.

The inability to simultaneously access information at arbitrarily small values of Δt and Δf is an inherent property of the transform. In fact, by using the Schwarz inequality for any function $g(t)$, it can be shown that an uncertainty principle ensues [23], i.e., that $\Delta t\Delta f = C$, where C is a nonzero constant whose precise value depends on the definition of width that is selected. Thus once $g(t)$ is chosen, the time and frequency resolutions of the STFT are fixed for all values of t and f. This is shown schematically in Fig. 9-1(b) by drawing regions in the τ-f plane where a set of functions $g^*(\tau - \tau_i)\exp(-j2\pi f_j\tau)$ are concentrated, since it is functions of this form that set the time and frequency resolutions of the STFT. These regions are illustrated as rectangles of fixed area and dimensions for all values of τ_i and f_j, and are said to tile the time-frequency plane [20–22].

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

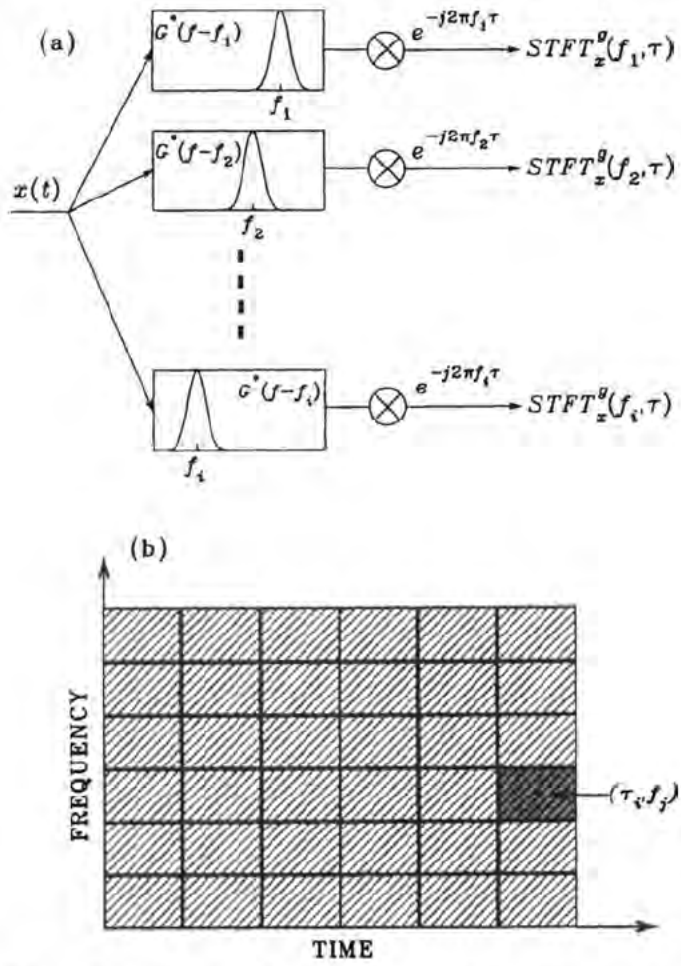


Figure 9-1 (a) Representation of the STFT in terms of filter-bank operations. The signal $x(t)$ to be analyzed is passed through a bank of filters, each with the same shape and bandwidth, but slightly different center frequencies f_i. The output from each filter is then multiplied by the factor $\exp(-j2\pi f_i\tau)$, which has the effect of shifting the output down to zero frequency and thereby providing the envelope. (b) Tiling of the time–frequency plane by the STFT. The rectangles centered at (τ_i, f_j) represent regions of the time–frequency plane where the functions $g^*(\tau-\tau_i)\exp(-j2\pi f_j\tau)$ are concentrated. These rectangles therefore also indicate the time and frequency resolution of the STFT.

9.3.3 The Continuous Wavelet Transform

A characteristic of the STFT is that both the time and frequency resolutions of the transform are fixed over the entire time–frequency plane. The time resolution Δt is fixed for the function $g(t)\exp(-j2\pi ft)$, whatever the value of f. As a result,

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

$\Delta f = C/\Delta t$ is also fixed over the entire time–frequency plane. In certain circumstances it is desirable to relax this restriction. Consider, for example, a signal with a mixture of short-lived high-frequency events that are closely spaced in time together with long-duration low-frequency components that are closely spaced in frequency. A suitable transform for this signal would have sufficient time resolution to distinguish the brief high-frequency events, and at the same time, enough frequency resolution to separate the closely spaced low-frequency components. These two aims are incompatible with the STFT since the time and frequency resolutions are both fixed.

One possible approach is to calculate two STFTs with different choices of $g(t)$: a short-lived $g(t)$ with a small value of Δt for good time resolution, and a long-lived $g(t)$ for good frequency resolution. An alternative solution is to use a representation that has variable time–frequency resolution over the (τ, f) plane, chosen in such a way that it provides good time resolution at high frequencies and good frequency resolution at low frequencies. One such representation is the continuous wavelet transform [19–22,24,25]. The CWT is expressed as

$$\mathrm{CWT}_x^h(r,\tau) = \frac{1}{\sqrt{|r|}} \int_{-\infty}^{\infty} x(t) h^*\left(\frac{t-\tau}{r}\right) dt \tag{9-7}$$

where t and τ are time variables, $x(t)$ is the time waveform being analyzed, $h(t)$ is the wavelet basis function, and r is a variable known as scale. As with the STFT, the CWT can also be expressed as an integral in the frequency domain [compare with Eq. (9-4)]:

$$\mathrm{CWT}_x^h(r,\tau) = \sqrt{|r|} \int_{-\infty}^{\infty} X(u) H^*(ur) \exp(j2\pi u\tau) du \tag{9-8}$$

or as a convolution in the time domain [compare with Eq. (9-5)]:

$$\mathrm{CWT}_x^h(r,\tau) = \frac{1}{\sqrt{|r|}} \left[x(\tau) * h^*\left(\frac{-\tau}{r}\right) \right] \tag{9-9}$$

where, as before, $X(u)$ and $H(u)$ denote the CFTs of $x(t)$ and $h(t)$, respectively.

The prefactor of $\sqrt{|r|}$ that appears in Eq. (9-8) illustrates that the standard CWT does not map equal-amplitude sinusoids of different frequencies to CWTs of the same magnitude; rather, it suppresses low-scale (high-frequency) components relative to those at high-scale (low-frequency). To facilitate comparison of the CWT results with those obtained with the STFT, it is useful to eliminate this difference. We therefore generally plot $|r|^{-1/2}|\mathrm{CWT}|$, which we refer to as the modified CWT. This has no effect on the time–frequency resolution characteristics of our analysis.

To understand how the CWT differs from the STFT, consider Eq. (9-9), which is the formulation of the CWT as a convolution in the time domain. As with the STFT, the value of $x(t)$ at $t = \tau_0$ is smeared over a time equal to the width of the function $h(\tau/r)$. In this case, however, the width of $h(\tau/r)$ is not fixed, but rather depends on the value of r. As an example, $h(2t)$ has half the width of $h(t)$, while $h(t/2)$ has twice the width of $h(t)$. The larger the value of r, the wider the function $h(\tau/r)$. Since time resolution depends on the width of this function, the following situation obtains: as r decreases, $h(\tau/r)$ becomes narrower in time so that the time resolution

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

improves. Conversely, as r increases, the time resolution is degraded, but the frequency resolution is simultaneously enhanced because the quantity $\Delta t \Delta f$ must be maintained constant. This is also apparent from Eq. (9-8), where $H(ur)$ becomes narrower as r increases, thus improving the frequency resolution. The reason the variable r is called scale is that it stretches and contracts the function $h(\tau/r)$. The net result is that such a transform is in fact useful for analyzing the kind of mixed signal discussed earlier.

The CWT is strictly defined as a time–scale representation; however, it often proves easier to interpret CWTs in terms of time and frequency rather than time and scale. A short-lived function (r small) inherently contains high frequencies, so that r is inversely related to frequency. For a given wavelet transform, the mapping $f = K/r$ can be used, allowing the CWT of a signal to be interpreted in terms of frequency rather than scale. This mapping is discussed further under the section that deals with the details of implementation (section 9.3.5).

As with the STFT, a filter-bank interpretation [20–22] can be invoked for the CWT, as illustrated in Fig. 9-2(a). In this case, the CWT is obtained by filtering the original signal by a bank of filters with fixed *relative* bandwidth rather than fixed absolute bandwidth, as for the STFT. The relative bandwidth ($\mathrm{BW_{rel}}$) of a filter (or function) is defined as the absolute bandwidth (Δf) of the bandpass region surrounding the filter's center frequency divided by the center frequency (f_h) itself. It is the inverse of the Q-factor:

$$\mathrm{BW_{rel}} = \frac{\Delta f}{f_h} = \frac{1}{Q} \tag{9-10}$$

For the CWT filter bank illustrated in Fig. 9-2(a), the relative bandwidth of the filters remains fixed, since both the absolute bandwidth and the center frequency of the functions $H(ur)$ vary in inverse proportion to r. The tiling of the time–frequency plane associated with the CWT is shown in Fig. 9-2(b); it consists of rectangles of fixed area but variable shape. At low frequencies, the rectangles are broad in time but narrow in frequency, since for large r, the time resolution is poor and the frequency resolution is good. The converse is true at high frequencies.

9.3.4 Wavelet Bases

Many different functions $h(t)$ can be used as prototypes in forming a CWT. Much work in recent years has focused on the issue of defining and using different wavelet bases for a variety of purposes. In this chapter we examine three particular wavelet bases—the Morlet, Meyer, and Daubechies 4-tap. It is useful to calculate CWTs using these different bases, with the aim of determining which are most useful for analysis purposes.

All wavelet basis sets should satisfy an "admissibility condition" [20, 22, 24], which states that if $h(t)$ is a wavelet basis for $\mathcal{L}^2$ (the set of square-integrable functions) then

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

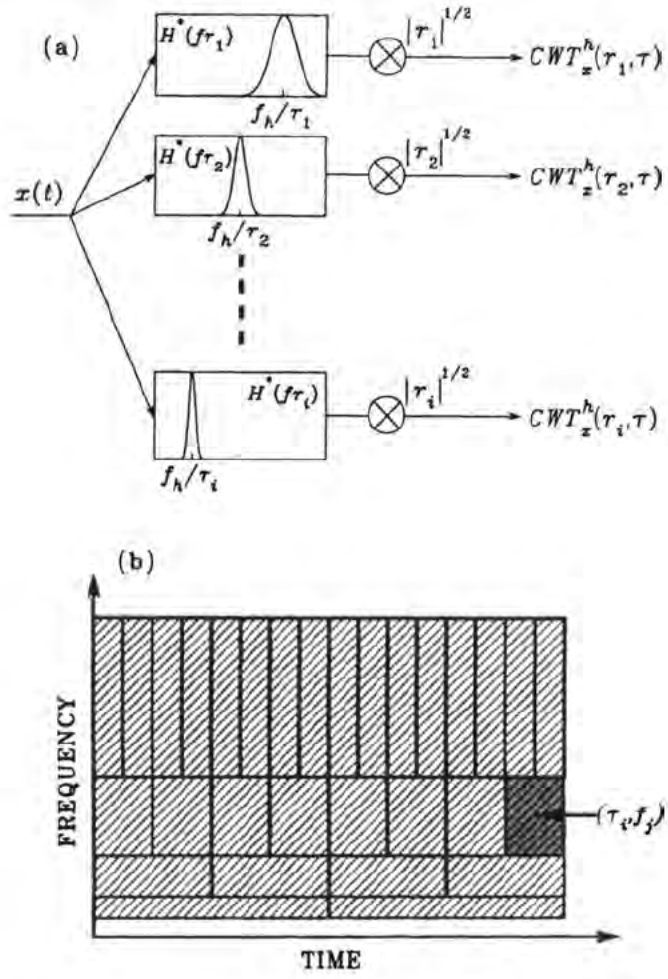


Figure 9-2 (a) Representation of the CWT in terms of filter-bank operations. The signal $x(t)$ to be analyzed is passed through a bank of filters, each of which is simply a scaled version of some prototype filter. Each filter has a fixed shape and relative bandwidth, but in absolute terms the bandwidths of the filters increase as the center frequency increases. The center frequency of the analysis filters is inversely proportional to scale r_i. The output from each filter is then multiplied by the gain factor $\sqrt{|r_i|}$. Since this gain factor increases with scale r, low-frequency components are accentuated with respect to high-frequency components. (b) Tiling of the time–frequency plane by the CWT. The rectangles centered at (τ_i, f_j) represent regions of the time–frequency plane where the functions $h^*([\tau - \tau_i]/r_j)$ are concentrated, with $f_j = K/r_j$. These rectangles also indicate the time and frequency resolution of the CWT. The CWT cannot resolve components at DC, since the basis functions do not extend to zero frequency. The figure shown here represents a dyadic grid, i.e., the rectangles double in length as we move along the frequency axis, but of course their areas are all the same. Our choice of a factor of 2 is illustrative; other ratios could equally well be used.

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

$$\int_{-\infty}^{+\infty} \frac{|H(f)|^2}{f} df < \infty \tag{9-11}$$

A consequence of the admissibility condition is that $H(0) = 0$ (otherwise $|H(f)|^2/f$ diverges at $f = 0$). In general, it is also true that $H(f) \to 0$ as $|f| \to \infty$. Combined, these two conditions show that the CFTs of wavelet basis functions represent bandpass filters since they remove components at both low and high frequencies.

In Fig. 9-3 we show four different wavelet bases in the time domain, along with the magnitudes of their CFTs in the frequency domain. For ease of comparison, we have normalized the CFTs so that their maxima always have unity magnitude and lie at $f = 1$. The Morlet and Meyer wavelet bases [Figs. 9-3(a), (c), and (e)] are complex and only the real part of the wavelet is plotted; the Daubechies 4-tap wavelet [Fig. 9-3(g)] is purely real.

Figure 9-3(a) shows the real part of the single-sided Morlet wavelet [24] given by $h(t) = \exp(jct)\exp(-\alpha t^2/2)$ [with $c = 2\pi$ and $\alpha = 0.0151$], and the magnitude of its Fourier transform [Fig. 9-3(b)], which is given by $\sqrt{2\pi/\alpha}\,\exp[-(2\pi f - c)^2/2\alpha]$. For the values of c and α that we have chosen here, $h(t)$ is a windowed sinusoid and $H(f)$ is a narrow bandpass filter centered at $f = 1$. The relative bandwidth of this wavelet is readily calculated to be

$$\mathrm{BW}_{\mathrm{rel}} = \frac{2\sqrt{2\alpha}}{c} \tag{9-12}$$

which equals 0.055 for this choice of parameters. Consequently, the Q-factor for this wavelet is $Q = 1/0.055 = 18.2 \gg 1$, which is why it is referred to as a high-Q Morlet wavelet in the remainder of this chapter.

Figures 9-3(c) and (d) also show a Morlet wavelet, but this time with $c = 2\pi$ and $\alpha = 0.151$. As in Fig. 9-3(b), this Morlet wavelet is also a bandpass function; however, its relative bandwidth ($= 0.55$) is much larger than that for the previous choice of parameters, and the number of oscillations in the time domain is lower than in Fig. 9-3(a). This wavelet is referred to as a low-Q Morlet wavelet in the text and figures that follow since its Q-factor is only 1.82.

Figures 9-3(e) and (f) show the Meyer wavelet similarly normalized (for details concerning the construction of the Meyer wavelet see [22]). Its structure is similar in both the frequency and time domains to the Morlet wavelet shown in Figs 9-3(c) and (d). However, unlike the Morlet wavelet, there are no free parameters that can be used to alter the Meyer wavelet's relative bandwidth. This is a consequence of the constraints under which the Meyer wavelet is constructed. The relative bandwidth of the Meyer wavelet is 1.18.

Finally, in Figs 9-3(g) and (h) we show a wavelet generated by an infinite iteration of a 4-tap finite impulse response (FIR) filter proposed by Daubechies (for details pertaining to the construction of this wavelet see [21,22]). Unlike the previous three wavelets, this basis is not symmetric, it has multiple peaks in the frequency domain, and there are no closed-form expressions for $h(t)$ or $H(f)$. The effect of these properties is discussed in section 9-4, which includes a CWT calculated using the Daubechies 4-tap wavelet. Like the Meyer wavelet, the relative bandwidth of this wavelet is also fixed, as a result of the manner in which it is calculated. The relative bandwidth of the Daubechies 4-tap wavelet (where the center frequency of

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

the filter is defined as the frequency at which the maximum of the first bandpass region occurs) is 1.5.

The principle feature of the Morlet wavelet of interest to us is that its relative bandwidth is easily adjusted by choice of the parameters c and α. We later show that this allows us complete flexibility in setting the CWT to have a desired frequency resolution at any particular frequency. However, the Morlet wavelet, unlike the Meyer and Daubechies wavelets, has two theoretical limitations. Though these are worthy of mention, they have no bearing on the usefulness of this basis set for our purposes. First, the Morlet wavelet does not strictly satisfy the admissibility condition since $H(0) \neq 0$. However for $\mathrm{BW}_{\mathrm{rel}} \leq 0.8$, the value of $H(0)$ is close to zero, and the Morlet wavelet is deemed to be practically admissible. For the Morlet wavelets shown in Figs 9-3(a) and (c), $\mathrm{BW}_{\mathrm{rel}} \approx 0.055$ and 0.55, respectively. Second, the Morlet wavelet cannot be used as the prototype wavelet $h(t)$ to create an orthonormal basis for $\mathcal{L}^2$ of the form $\{h_{ij}(t) = 2^{i/2}h(2^i t - j)\}$, $i, j \in \mathcal{Z}$ (the set of natural numbers) [20–22, 25]. The ability to form such a basis is central to the design of wavelet bases for use in perfect reconstruction filter banks.

For the functions shown in Figs 9-3(c)–(h) (low-Q Morlet, Meyer, and Daubechies 4-tap wavelets), the relative bandwidth of the wavelet's CFT is of the order of unity. Such wavelets are loosely termed octave-band, since the functions $H(f)$ and $H(2f)$, whose center-frequencies are separated by an octave, are just about far enough apart in frequency to be resolved. This octave-band property arises naturally in wavelets that are designed to satisfy two-scale equations [21,22]. Both the Meyer and Daubechies wavelets satisfy such equations, and can be rigorously used as prototype wavelets to form an orthonormal basis for $\mathcal{L}^2$ of the form indicated above. However, while orthonormal-basis-generating wavelets do provide a useful set of functions for constructing efficient wavelet series expansions of $x(t)$ [22], they are not always suited for use as a CWT basis, as our examples will show.

9.3.5 STFT and CWT Implementation

A sampled version of the STFT, often referred to as the discrete STFT, was calculated using a summation approximation of Eq. (9-3):

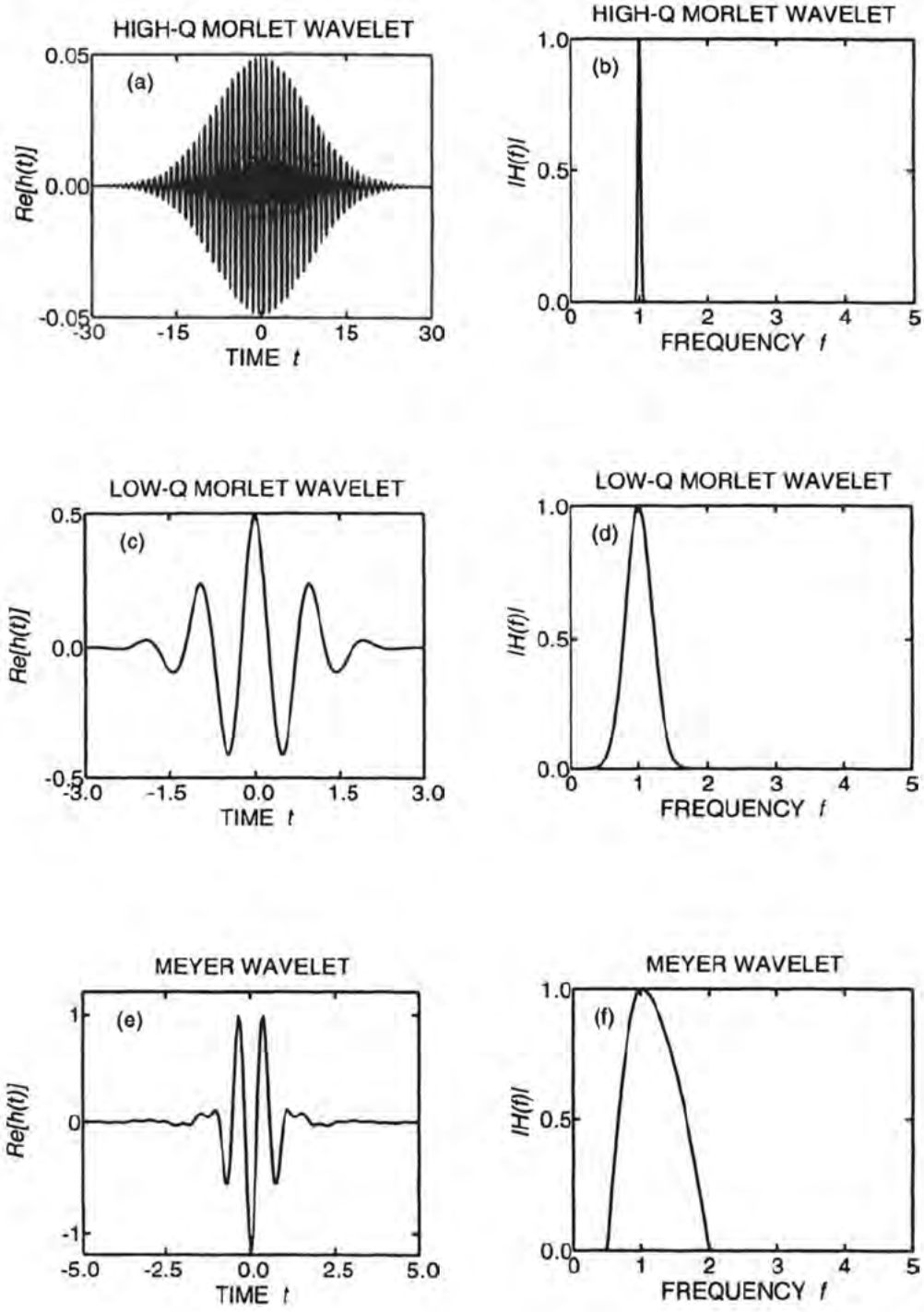
$$\mathrm{STFT}_x^g[k, n] = \sum_{m=0}^{L-1} x[n+m]g^*[m]\exp\left(\frac{-j2\pi mk}{N}\right), \qquad 0 \leq k \leq N-1 \tag{9-13}$$

where k is the discrete frequency index, n is the discrete time index, L is the window length in samples, and $g[m]$ is chosen to be samples of a Gaussian window $g(t) = \exp(-\beta t^2/2)$, with $g(t)$ falling to e^{-4} at the sampled endpoints:

$$g[m] = \exp\left[-\left(-2 + \frac{4m}{L-1}\right)^2\right], \qquad 0 \leq m \leq L-1 \tag{9-14}$$

The formulation presented in Eq. (9-13) reminds us that the discrete STFT is simply a sequence of discrete Fourier transforms (DFTs) of the windowed signal segments.

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.



M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

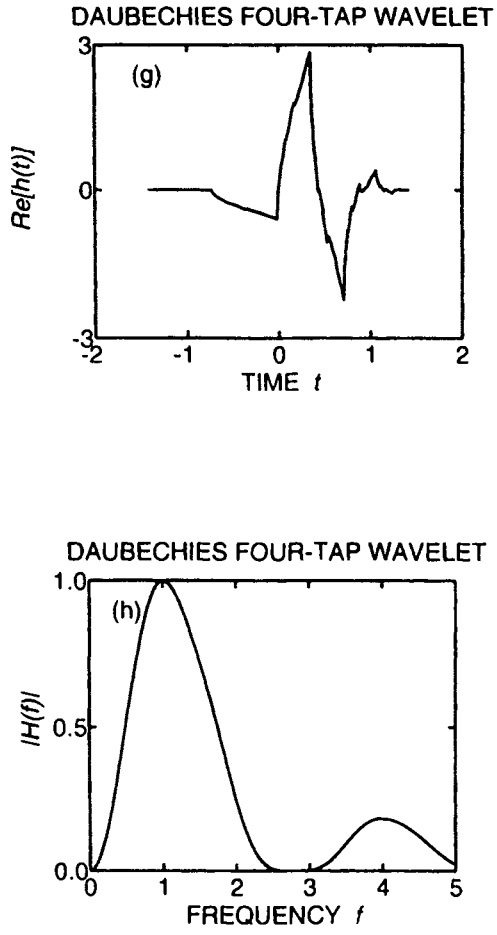


Figure 9-3 Four wavelet bases in the time domain and their respective CFTs. (a) Real part of a high-Q Morlet wavelet, $h(t) = \exp(jct)\exp(-\alpha t^2/2)$, with $c = 2\pi$ and $\alpha = 0.0151$. There are many (> 50) oscillations within the envelope of the wavelet, and its overall duration is approximately 60. (b) Magnitude of the CFT of the wavelet shown in (a). The relative bandwidth of this function is narrow ($= 0.055$). It is symmetric about the frequency $f = 1$. (c) Real part of a low-Q Morlet wavelet, $h(t) = \exp(jct)\exp(-\alpha t^2/2)$, with $c = 2\pi$ and $\alpha = 0.151$. There are relatively few oscillations within the envelope of the wavelet, and its overall duration is approximately 6. (d) Magnitude of the CFT of the wavelet shown in (c). The relative bandwidth of this function is 0.55. It is symmetric about the frequency $f = 1$. (e) Real part of a Meyer wavelet calculated as outlined in [22] (there is no closed-form expression for this wavelet in the time domain). As for the Morlet wavelet pictured in (c), there are relatively few oscillations within the envelope. The overall duration of the wavelet is approximately 5. (f) Magnitude of the CFT of the wavelet shown in (e). The relative bandwidth of this function is approximately 1.18. The function is asymmetric about the frequency $f = 1$. (g) Time-waveform of a wavelet calculated using an iterative procedure based on a 4-tap filter with coefficients as proposed by Daubechies (the Daubechies 4-tap filter) [22]. This wavelet has no closed-form expression, is purely real, and is not symmetric in time. (h) Magnitude of the CFT of the wavelet shown in (g). The relative bandwidth of the first bandpass region is 1.5. This function has multiple lobes in the frequency domain.

Once a window length L is chosen, the time–frequency uncertainty product is fixed—a good rule of thumb is to choose the window length so that the signal appears "relatively stationary" within it. For example, in the analysis of AM responses, as discussed subsequently, a window length $L = 128$ samples was appropriate. This corresponds to 1.28 ms at the 100 kHz sampling rate used in recording our data. The value of N (which sets the number of discrete frequencies at which the STFT is sampled in the frequency domain) was chosen equal to L. The STFT was not evaluated for all values of n; moving the time window through 256 time samples for successive evaluations of the STFT provided a sufficiently detailed picture for our purposes. For our particular choice of window length and sampling rate, the time resolution of the STFT was $\Delta t = 0.64$ ms, and the frequency resolution was $\Delta f = 1989.4$ Hz. The uncertainty product $\Delta t \Delta f = C = 4/\pi \approx 1.27$.

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

We present the STFT magnitude in two visual formats. The first is a three-dimensional representation, often referred to as a 3-D spectral plot. In this format, time and frequency form the bottom plane, and the STFT magnitude is represented on a linear axis in the third dimension. The second format provides 2-D contour plots, on which contours of equal STFT magnitude are traced on a time–frequency plane.

To calculate the CWT, we implemented the fast-CWT algorithm proposed by Jones and Baraniuk [26]. This technique avoids carrying out a time-consuming direct time-convolution of the data with the scaled wavelet time waveform; instead we express the time-convolution as multiplication in the frequency domain. Efficient algorithms are then exploited to carry out the calculation.

To see explicitly how this is done, reconsider Eq. (9-8), in which the CWT is written as an integral in the frequency domain:

$$\mathrm{CWT}_x^h(r,\tau) = \sqrt{|r|}\left[\int_{-\infty}^{\infty} X(u)H^*(ur)\exp(j2\pi u\tau)du\right] \tag{9-15}$$

where r is scale, τ is time, and u represents frequency. The term inside the square brackets is the inverse CFT of $X(u)H^*(ur)$ [compare with Eq. (9-2)], which suggests that an inverse fast Fourier transform (FFT) can be used to evaluate the CWT. Specifically, consider $x[n]$ as a well-sampled version of the continuous-time function $x(t)$ with sampling time equal to T_x, and $h[n]$ as a well-sampled version of $h(t)$, normalized as shown in Fig. 9-3, with a sampling time equal to T_h. There are N_x samples of $x[n]$ and N_h samples of $h[n]$. The discrete-time Fourier transforms (DTFTs) of $x[n]$ and $h[n]$ are denoted by $X(e^{j\omega})$ and $H(e^{j\omega})$, respectively, where ω represents digital angular frequency [18]. The sampling theorem tells us that for well-sampled signals, $X(e^{j\omega}) = X(f)/T_x$ at $f = \omega/2\pi T_x$ over the range $\omega = [-\pi, \pi]$. Since we assume that $x(t)$ is well-sampled, $X(f)$ is essentially zero for $f > 1/2T_x$, which allows us to write Eq. (9-15) in terms of the DTFTs of $x[n]$ and $h[n]$:

$$\mathrm{CWT}_x^h(r,\tau) = \sqrt{|r|}\left[\frac{T_x}{2\pi}\int_{-\pi}^{+\pi} X(e^{j\omega})H^*(e^{j\omega r})\exp(j\omega\tau/T_x)\,d\omega\right] \tag{9-16}$$

Restricting ourselves to evaluating the CWT at a discrete set of $r = [r_i]$ and $\tau = nT_x$ brings us to a sampled version of the CWT:

$$\mathrm{CWT}_x^h[r_i,n] = \sqrt{|r_i|}\,\frac{T_x}{2\pi}\int_{-\pi}^{+\pi} X(e^{j\omega})H^*(e^{j\omega r_i})\exp(j\omega n)\,d\omega \tag{9-17}$$

which is recognizable as a multiple of the inverse DTFT of the function $Y(e^{j\omega}, r_i) = X(e^{j\omega})H^*(e^{j\omega r_i})$. This inverse DTFT is efficiently implemented by the inverse FFT

$$\mathrm{CWT}_x^h[r_i,n] = \sqrt{|r_i|}\,\frac{T_x}{M}\sum_{k=-M/2}^{M/2-1} Y[k,r_i]\exp(j2\pi kn/M) \tag{9-18}$$

where $Y[k, r_i] = Y(e^{j\omega}, r_i)$ evaluated at $\omega = 2\pi k/M$, with k ranging from $-M/2$ to $M/2 - 1$. Since efficient computation algorithms exist for the inverse FFT, only two questions remain: how to calculate $Y[k, r_i]$, and what is an appropriate value of M?

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

We can write $Y[k, r_i]$ as $X[k]H^*[k, r_i]$, where $X[k] = X(e^{j\omega})$ evaluated at $\omega = 2\pi k/M$, and $H^*[k, r_i] = H^*(e^{j\omega r_i})$, also evaluated at $\omega = 2\pi k/M$. $X[k]$ is now simply the FFT of the sequence $x[n]$, and $H[k, r_i]$ is the chirp z-transform of the sequence $h[n]$. Efficient algorithms exist for both the FFT and the chirp z-transform, allowing us to evaluate both $X[k]$ and $H[k, r_i]$, which in turn leads us to $Y[k, r_i]$ and ultimately to $\mathrm{CWT}_x^h[r_i, n]$ via the inverse FFT. (For a complete discussion of the chirp z-transform, see pp. 623–628 of [18].)

To choose an appropriate value of M, we must ensure that the frequency multiplication of $H(e^{j\omega r})$ and $X(e^{j\omega r})$ really gives us the desired linear convolution in time from Eq. (9-15). This is ensured by selecting M greater than the combined lengths of the sequences $x[n]$ and the longest wavelet basis we use (which is N_h multiplied by $\max(r_i)$, since the largest value of r will produce the most stretched wavelet function). In practice, the next highest power of 2 greater than $N_x + \max(r_i)N_h - 1$ is chosen, so that power-of-2 FFTs can be used.

There is a remaining subtlety in the fast-CWT algorithm. At the outset, we assumed that $H(f) = 0$ for all values of $|f| > 1/2T_x$. Therefore, to correctly carry out the multiplication of Eq. (9-15), we must ensure that aliased versions of $H(e^{j\omega r_i})$ are not brought into the range $[-\pi, \pi]$. Therefore, we set $H(e^{j\omega r_i}) = 0$ for $|\omega r_i| > \pi$. Since we are using $\omega = 2\pi k/M$, this implies $Y[k, r_i] = 0$ for $k > M/2r_i$. As k ranges from $-M/2$ to $M/2 - 1$, this only occurs for $r_i < 1$.

As stated earlier, scale r is inversely related to frequency. For convenience in interpreting CWTs, we have mapped scale to frequency using the mapping $f = K/r$ as indicated earlier. We choose this mapping to assign a given scale r_i to a frequency f_i equal to the center frequency of the filter $H(fr_i)$. The proportionality constant K is evaluated by obtaining the center frequency of the function $H(f)$, since this corresponds to $H(fr_i)$ at $r_i = 1$. Therefore, we must find what $H(f)$ the samples $h[n]$ represent. The set $h[n]$ is constructed by taking N_h samples of the continuous function $h(t)$, normalized as shown in Fig. 9-3. To provide a complete representation, but without undue oversampling, $h[n]$ is constructed by sampling $h(t)$ at a different rate $(1/T_h)$ as compared to the sampling rate $1/T_x$ for $x[n]$. However, in evaluating the CWT as described previously, the $h[n]$ are taken as a set of samples at rate $1/T_x$. This means that $h[n]$ represents samples of a continuous function $h(tT_h/T_x)$. The center frequency of this function's CFT, which is $|T_x/T_h|\, H(fT_x/T_h)$, occurs at $f = T_h/T_x$ rather than at $f = 1$. This center frequency T_h/T_x corresponds to scale $r = 1$. Substituting these values into $f = K/r$, we find that $K = T_h/T_x$. This allows us to map the CWTs on a time–frequency plane, but we remind the reader that this is only an interpretational convenience; strictly speaking, the CWT is a time–scale representation.

A word is in order describing the link between the CWT and its discrete counterpart, the discrete wavelet transform (DWT) [21,22]. The DWT provides a multiresolution approximation of the sequence $x[n]$. This approximation involves repeated high-pass and low-pass filtering of the original signal $x[n]$, with downsampling by a factor of 2 after each filtering operation. The various high-pass sequences are retained as a useful approximation of the signal. The number of samples contained in the complete set of high-pass sequences plus the residual low-pass sequence is equal to the number in the original signal, and the high-pass sequences contain

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

nearly all the information of the original signal (however the DC component is lost as it is contained in the low-pass residual). This process decomposes the sequence $x[n]$ into various frequency bands, as represented by the set of high-pass sequences. The center frequency of each band differs by a factor of 2 as a result of the down-sampling factor used in calculating the DWT.

In a similar manner, the CWT decomposes the original signal $x(t)$ into an infinite set of time sequences $\mathrm{CWT}_x^h(r, \tau)$, also distributed in various frequency bands across the time–frequency plane. However, unlike the DWT, the various frequency bands are not constrained to differ by a factor of 2; rather, they can be evaluated at arbitrary values of r. Essentially, the CWT is an interpolated version of the DWT in which a decomposition at arbitrary scales can be examined. This is analogous to considering the CFT as an interpolated version of the FFT. The DWT and FFT may alternatively be viewed as sampled versions of the CWT and CFT, respectively.

As in the case of the STFT, we use two visual formats for the CWT. The first is a three-dimensional (3-D) representation, in which time and frequency form the bottom plane, and CWT magnitude (multiplied by the factor of $|r|^{-1/2}$ for ease of comparison with the STFT) is represented on a linear axis in the third dimension. The second format provides 2-D contour plots, on which contours of equal modified CWT magnitude are traced on a time–frequency plane.

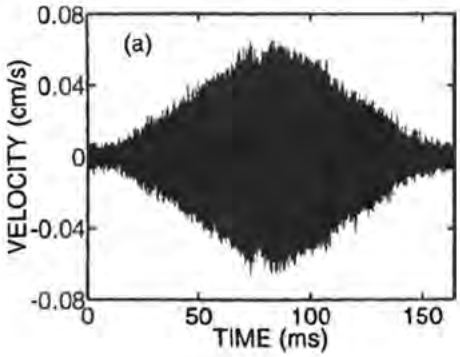
9.4. RESULTS

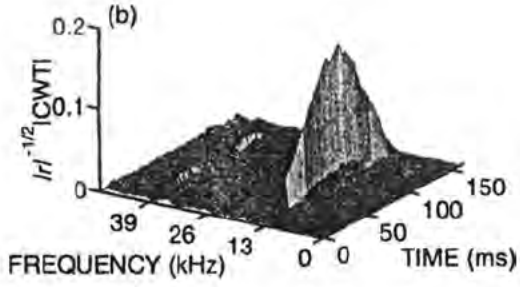
We used the modified CWT and STFT to examine the behavior of the velocity response of auditory sensory cells in response to the AM acoustic signals described in section 9.1. These AM signals (with unity modulation depth and a fixed modulation frequency of 6.1 Hz) were applied to the ear canal, using different carrier frequencies and levels. This allowed us to examine the nonlinear velocity response of a given cell as the carrier frequency was altered from well below to above the CF, and as the stimulus intensity was varied. The sampling frequency was fixed at 100 kilosamples/s in recording the responses. The data reported here were collected at the level of the basilar membrane, just below the first row of outer hair cells, at the edge of the tunnel region of the cochlea. The CF was 36 kHz, a typical value for cells in the basal turn of the cat cochlea.

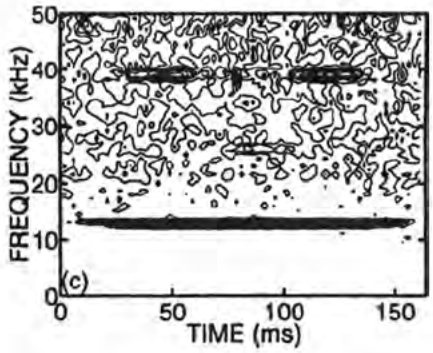
Figure 9-4 shows the response of the cell when an AM tone with a carrier frequency $f_c = 13$ kHz (well below CF), and a peak sound pressure level of ≈ 108 dB:re .0002 dyne/cm^2, is applied to the ear. The velocity response of the cell during one cycle of the modulation envelope is displayed in Fig. 9-4(a). It roughly reflects the underlying shape of the modulation envelope.

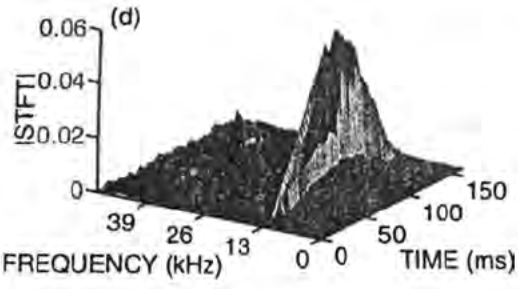
The modified CWT magnitude of the velocity response, shown in 3-D and 2-D formats in Figs 9-4(b) and (c), respectively, reveals the time behavior of the signal at different scales, which are represented as different frequencies as discussed earlier. The mapping from scale to frequency was accomplished using $f = K/r$ with $K = 1.27 \times 10^4$. This CWT is based on the high-Q Morlet wavelet shown in Figs 9-3(a) and (b). The modified CWT shows three spectral components at multiples of

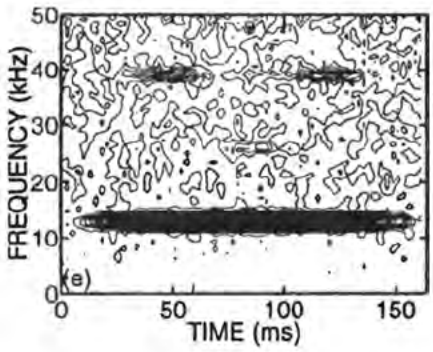
M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.



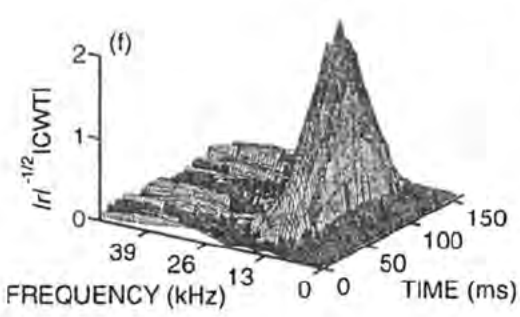


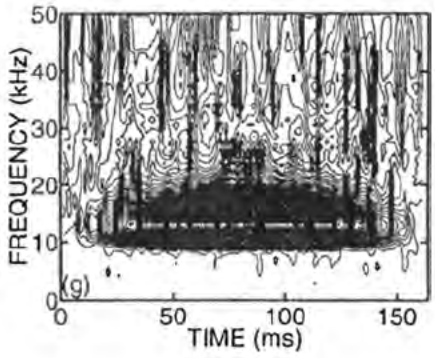


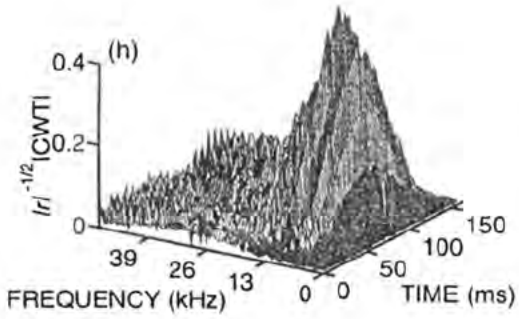


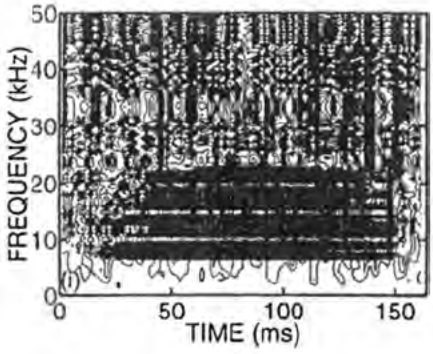


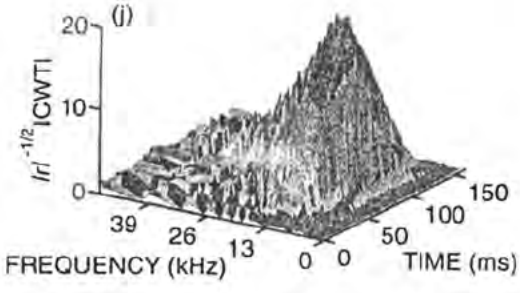
M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

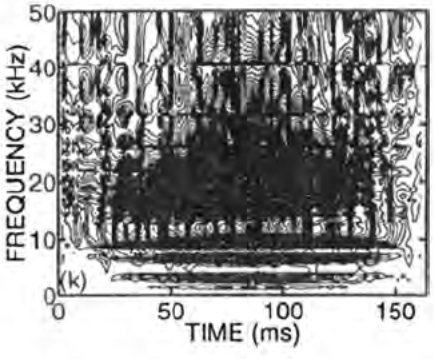












M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

Figure 9-4 Velocity response at the level of the basilar membrane, under the first row of outer hair cells in the basal turn of the living cat cochlea (file number 38260710.mat). The CF of the cell was 36 kHz, and this response is to an AM stimulus with carrier frequency $f_c = 13$ kHz (well below CF) and modulation frequency 6.1 Hz. The highest sound pressure level, occurring at the center of the input envelope was $\approx$ 108 dB:re .0002 dyne/cm^2. (a) Time waveform of the velocity response (in cm/s) of the cell. (b) 3-D plot of the modified CWT magnitude (viz., the CWT multiplied by the factor $|r|^{-1/2}$) of the velocity response shown in (a). The high-Q Morlet wavelet basis shown in Figs 9-3(a) and (b) was used. The x and y axes represent time (ms) and frequency (kHz), respectively, and modified CWT magnitude is plotted on a linear scale on the z axis. This plot shows spectral components at the carrier frequency f_c and at two higher harmonic frequencies, $2f_c$ and $3f_c$. (c) Same modified CWT magnitude as shown in (b), but now plotted in 2-D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (d) 3-D plot of the STFT magnitude of the velocity response shown in (a). Spectral components are present at the carrier frequency f_c and at two higher harmonics, $2f_c$ and $3f_c$. (e) Same STFT magnitude as shown in (d), but now plotted in 2-D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude. (f) 3-D plot of the modified CWT magnitude of the velocity response shown in (a). The low-Q Morlet wavelet basis shown in Figs 9-3(c) and (d) was used. This plot shows a large-bandwidth spectral component centered at the carrier frequency f_c. Energy is present at higher frequencies but the resolution is very poor. (g) Same modified CWT magnitude as shown in (f), but now plotted in 2-D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (h) 3-D plot of the modified CWT magnitude of the velocity response shown in (a). The Meyer wavelet basis shown in Figs 9-3(e) and (f) was used. As in (f), this plot shows a large-bandwidth spectral component centered at the carrier frequency f_c, and energy at higher frequencies that cannot be clearly resolved. (i) Same modified CWT magnitude as shown in (h), but now plotted in 2-D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (j) 3-D plot of the modified CWT magnitude of the velocity response shown in (a). The Daubechies 4-tap wavelet basis shown in Figs 9-3(g) and (h) was used. As in (f) and (h), this plot shows a large-bandwidth spectral component at the carrier frequency f_c and energy at higher frequencies that cannot be clearly resolved. (k) Same modified CWT magnitude as shown in (j), but now plotted in 2-D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude.

the carrier frequency. At the beginning and end of the modulation cycle, when the magnitude of the envelope is low, the response closely follows the input and is exclusively at the carrier frequency. As the magnitude of the envelope increases, a spectral component at $3f_c$ appears, and right at the center of the modulation envelope, where the sound pressure is greatest, a component at $2f_c$ is just discernible above the noise floor. These components subsequently disappear as the envelope decreases. The spectral component at f_c is more clearly resolved in frequency as compared to the component at $3f_c$, as expected for the CWT, where frequency resolution

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

improves at lower frequencies. Of course, this improvement of frequency resolution is accompanied by a degradation in time resolution. The presence of harmonic generation clearly indicates nonlinearity in the cellular response. The generation of multiple harmonic components for carrier frequencies below CF is also typical of sensory cells in the guinea-pig cochlea [12,13]. In particular, we have often seen a tendency for odd-harmonic components to be more significant than their neighboring even-harmonic components (e.g., see Fig. 8 of [13]).

For comparison, Figs 9-4(d) and (e) show the STFT magnitude of the velocity response, calculated as described earlier. The time and frequency resolution of the STFT were chosen to match those of the CWT shown in Figs 9-4(b) and (c) at a frequency of 36 kHz. Like the CWT, the STFT also shows spectral components at f_c, $2f_c$, and $3f_c$, with the component at $2f_c$ barely visible.

As expected, the spectral widths represented in the STFT are constant whatever the frequency of the component; unlike the CWT, the frequency resolution is fixed for all frequencies. Similarly, the time resolution remains constant at all frequencies.

To evaluate the relative merits of different wavelet bases, the CWT of the same velocity-response waveform shown in Fig. 9-4(a) was calculated using the three other wavelet bases shown in Figs 9-3(c)–(h).

Figures 9-4(f) and (g) display the modified CWT magnitude calculated using the low-Q Morlet wavelet basis displayed in Figs 9-3(c) and (d). The proportionality constant K, mapping scale to frequency, was 1.27×10^3. These plots show the presence of a component at approximately 13 kHz, but little else besides. The component at $3f_c$ is not apparent, since the frequency resolution of this CWT is so poor at that frequency. This is because the relative bandwidth of this wavelet basis is a factor of ten larger than that of the high-Q Morlet wavelet, so the frequency resolution at any given frequency is diminished by a factor of ten. For instance, at the frequency 39 kHz ($=3f_c$ for this example), the frequency resolution of the high-Q Morlet wavelet is 2155 Hz; that of the low-Q Morlet wavelet is 21550 Hz, so that it is not surprising that no clearly resolved component can be distinguished at that high frequency. The improved time resolution accruing at higher frequencies is of little advantage for our purposes, as the poor frequency resolution makes it hard to ascribe events in time to a particular frequency component. Such information would be useful in determining the underlying dynamics of the response.

Figures 9-4(h) and (i) show the modified CWT magnitude calculated using the Meyer wavelet basis shown in Fig. 9-3(e) and (f). The proportionality constant K, mapping scale to frequency, was 6.26×10^3. The results are similar to those for the low-Q Morlet wavelet, as expected. However, there is more roughness in the CWT; this reflects the roughness of the absolute value of the Meyer wavelet, whose real part is displayed in Fig. 9-3(e).

Finally, Figs 9-4(j) and (k) show the modified CWT magnitude calculated using the Daubechies 4-tap wavelet basis shown in Figs 9-3(g) and (h). The proportionality constant K, mapping scale to frequency, was 281. Unlike the previous three wavelets, this wavelet is real and asymmetric. The CWT is therefore also real, and is capable of switching between positive and negative values. As a consequence, the absolute value of the CWT often goes to zero, leading to a highly scalloped structure for the CWT magnitude. A further difficulty in using the Daubechies 4-tap wavelet basis as an

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

analysis tool is the presence of multiple bandpass regions in the CFT of the wavelet basis, as is evident in Fig. 9-3(h). This allows the energy from a single frequency component to enter the CWT at a variety of different frequencies. In the CWT shown here, for example, the energy from the component at 13 kHz appears in the CWT magnitude not only at 13 kHz, but also at $13/4 = 3.25$ kHz and $13/8 = 1.625$ kHz (the component at 13 kHz arises when the main lobe of the analysis wavelet is at 13 kHz, the component at 3.25 kHz appears when the second lobe of the analysis wavelet is at 13 kHz, and the component at 1.625 kHz appears when the third lobe of the analysis wavelet is at 13 kHz).

The most useful wavelet basis for analyzing the responses recorded in our experiments is the high-Q Morlet wavelet. This is because the frequency separation of the spectral components we need to resolve is relatively small compared to the absolute values of the frequencies themselves. The Morlet wavelet basis works well since we can control its relative bandwidth by adjusting the ratio $\sqrt{\alpha}/c$. The other wavelet bases we have investigated are unsuitable as analysis tools for the class of data examined here since they are octave-band in nature. In the further data analysis that follows, therefore, we restrict ourselves to use of the high-Q Morlet wavelet CWT and the STFT.

Figure 9-5 shows the response of the same cell when an AM tone with a carrier frequency $f_c = 22$ kHz (below CF), and a peak sound pressure level of ≈ 107 dB:re .0002 dynes/cm^2, is applied to the ear. The velocity response of the cell over one modulation cycle is displayed in Fig. 9-5(a). Figures 9-5(b) and (c) show the modified CWT magnitude of the velocity response. In this case a component at $2f_c$ can be distinguished; the presence of components at higher harmonics cannot be established, as the sampling frequency does not allow us to view components above 50 kHz.

Figures 9-5(d) and (e) show the STFT magnitude of the velocity response. Again, the STFT and modified CWT magnitudes display similar information.

Figure 9-6 shows the response of this cell when an AM tone with a carrier frequency $f_c = 36$ kHz (at CF), and a peak sound pressure level of ≈ 99 dB:re .0002 dynes/cm^2, is applied to the preparation. The velocity response of the cell over one cycle of the modulation envelope is displayed in Fig. 9-6(a). The response does not follow the shape of the input envelope, but rather shows substantial amplitude variations. We have often observed similarly uneven time waveform responses in the guinea-pig cochlea for AM signals with carrier frequencies near CF (e.g., see Fig. 9(a) of [13] and Fig. 3(a) of [12]; these responses often contain harmonic, half-harmonic, and quarter-harmonic components). Figures 9-6(b) and (c) show the modified CWT magnitude of the velocity response calculated using the high-Q Morlet wavelet. The modified CWT shows that the amplitude of the component at f_c varies significantly with time near the center of the modulation envelope. The STFT magnitude of the velocity response [Figs 9-6(d) and (e)] shows a similar effect. This variation in time of the response at f_c is likely accompanied by the presence of energy at higher harmonic frequencies which we did not measure.

Similar results are obtained for carrier frequencies above the CF, at least up to the highest frequency we examined, which was 40 kHz.

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

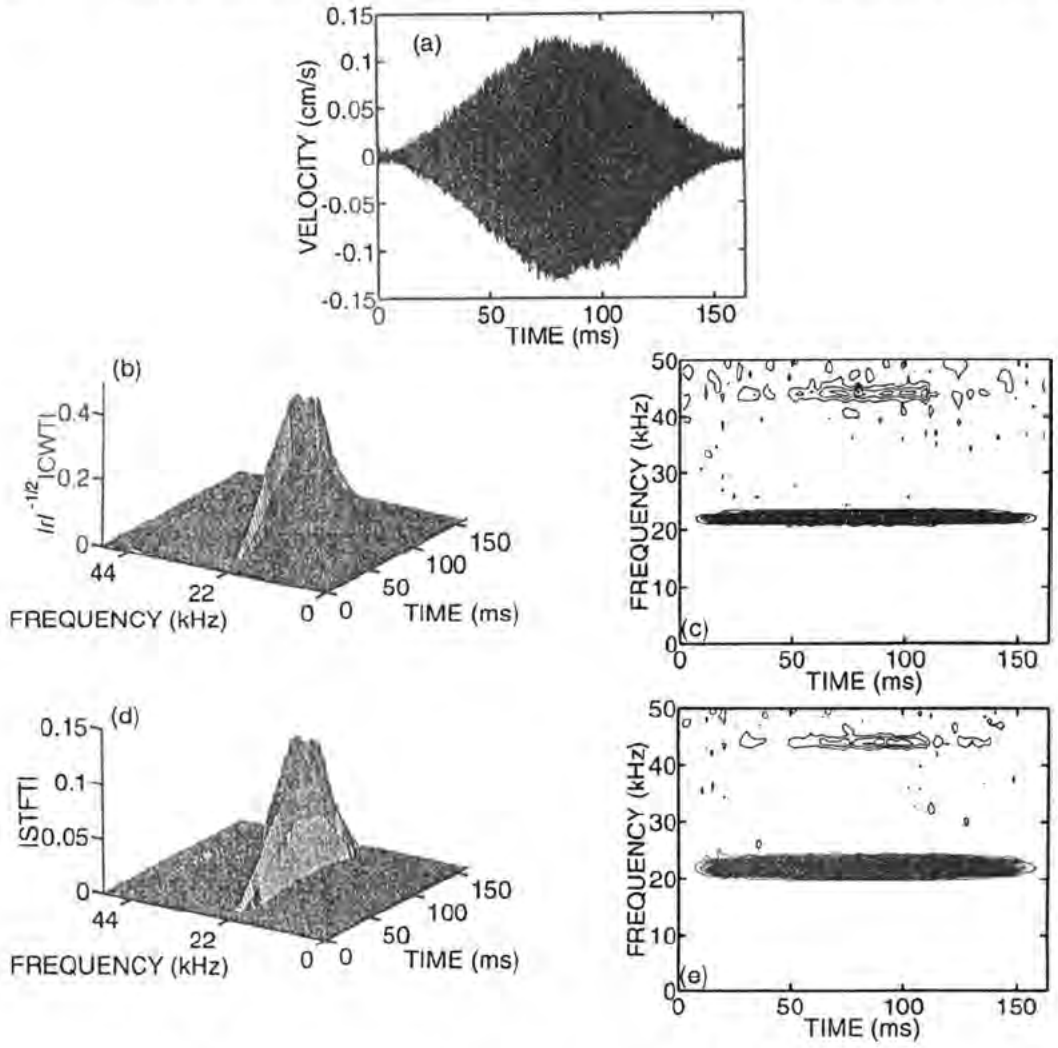


Figure 9-5 Velocity response of the same cell as shown in Fig. 9-4, but now to an AM stimulus with carrier frequency $f_c = 22$ kHz (below CF) and modulation frequency 6.1 Hz. The highest sound pressure level, occurring at the center of the input envelope was $\approx$ 107 dB:re .0002 dyne/cm^2. (a) Time waveform of the velocity response (in cm/s) of the cell. (b) 3-D plot of the modified CWT magnitude of the velocity response shown in (a). The high-Q Morlet wavelet basis shown in Figs 9-3(a) and (b) was used. This plot shows spectral components at the carrier frequency f_c and at the second harmonic, $2f_c$. (c) Same modified CWT magnitude as shown in (b), but now plotted in 2-D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (d) 3-D plot of the STFT magnitude of the velocity response shown in (a). Spectral components are present at the carrier frequency f_c, and at the second harmonic, $2f_c$. (e) Same STFT magnitude as shown in (d), but now plotted in 2-D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude.

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

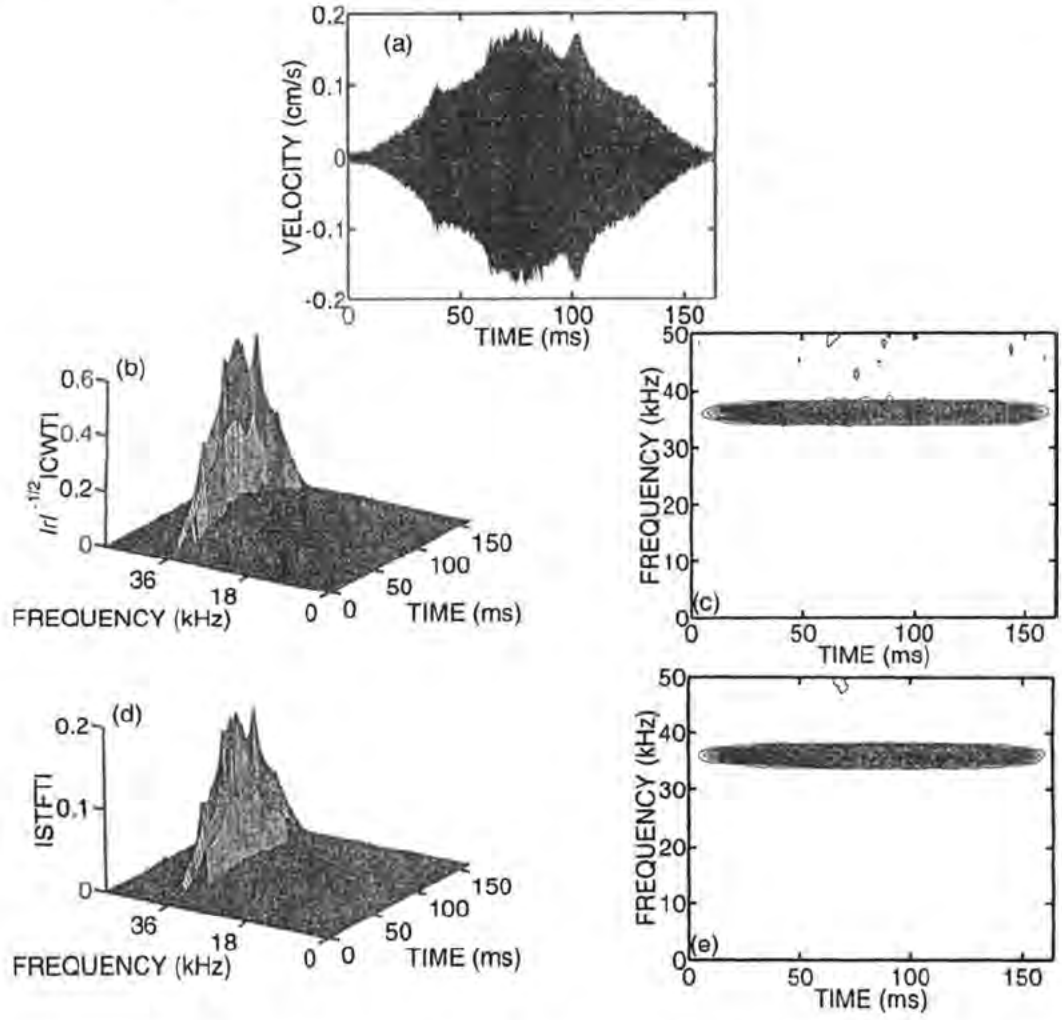


Figure 9-6 Velocity response of the same cell as shown in Figs 9-4 and 9-5, but now to an AM stimulus with carrier frequency $f_c = 36$ kHz (at CF) and modulation frequency 6.1 Hz. The highest sound pressure level, occurring at the center of the input envelope, was $\approx$ 99 dB re .0002 dyne/cm^2. (a) Time waveform of the velocity response (in cm/s) of the cell. (b) 3-D plot of the modified CWT magnitude of the velocity response shown in (a). The high-Q Morlet wavelet basis shown in Figs 9-3(a) and (b) was used. This plot shows the irregular variation with time of the spectral component at f_c, particularly near the center of the modulation envelope. (c) Same modified CWT magnitude as shown in (b), but now plotted in 2-D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (d) 3-D plot of the STFT magnitude of the velocity response shown in (a). This plot also shows the time-variation of the spectral component at the carrier frequency f_c. (e) Same STFT magnitude as shown in (d), but now plotted in 2-D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude.

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

9.5. DISCUSSION

It is apparent from the foregoing results, and the data analyses presented in Figs 9-4–9-6, that the CWT and STFT have different frequency- and time-resolution properties. It is these properties, and their relationship to the characteristics of the signal itself, that determine the relative advantages of the two techniques for analyzing a given signal.

The time and frequency resolutions of the CWT vary with scale, though their product remains fixed. Whatever the choice of wavelet basis, the frequency resolution is worst at small scales, and improves (i.e., decreases in width) with increasing scale. Since frequency is proportional to inverse scale [20–22], this means that CWT frequency resolution is best for low frequencies, and worst for high frequencies.

Given this fundamental property of the CWT, one can inquire whether a desired frequency resolution Δf can be selected at any arbitrary frequency f. This is equivalent to asking whether the relative bandwidth ($\Delta f/f$) of a wavelet basis function can be freely chosen. Many wavelet bases (including the Meyer and Daubechies 4-tap wavelets) were designed to provide a useful multiresolution framework onto which a signal could be efficiently decomposed [22], and therefore satisfy two-scale equations. As a consequence, their relative bandwidths are fixed (near unity) and cannot be freely chosen. These wavelets (and others like them) were not primarily meant for use in signal analysis.

The Morlet wavelet, however, was designed as an analysis tool, and its relative bandwidth can be freely set by choosing the value of $\sqrt{\alpha}/c$. Explicitly, given a desired frequency resolution Δf at the frequency f, set $\sqrt{\alpha}/c$ by using Eq. (9-12), i.e.,

$$\frac{\sqrt{\alpha}}{c} = \frac{\Delta f}{2\sqrt{2}f} \tag{9-19}$$

The choice of c and α individually is unimportant; rather it is the ratio $\sqrt{\alpha}/c$ that determines the relative bandwidth.

STFT analysis, on the other hand, requires that the frequency resolution and time resolution remain constant at all frequencies. The frequency resolution is determined by the choice of the window function $g(t)$. In our case, the window function is a sampled version of $g(t) = \exp(-\beta t^2/2)$, so that the frequency resolution (at all frequencies) is given by $\sqrt{2\beta}/\pi$. In our discrete implementation of the STFT, β was a function of the sampled window length L and the sampling time T_x, in accordance with $\beta = 32/L^2T_x^2$. For $L = 128$ and $T_x = 1/100\,000$, which were used in calculating the STFTs shown here, $\beta = 1.95 \times 10^7$ and the frequency resolution of the STFT was $\Delta f = 1989.4$ Hz at all frequencies.

With this kind of control over the time and frequency resolutions of both the CWT and the STFT, the final choice of which technique to use is signal-dependent. From a signal-analysis point of view, the CWT fares best when the required frequency resolution (or spacing between spectral components) varies as inverse frequency. On the other hand, the STFT is best suited to situations where the spectral components are linearly spaced in frequency. Thus, the CWT is the preferred tool when the analysis requires good frequency resolution at low frequencies together

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

with good time resolution for impulsive (high-frequency) events. The STFT is appropriate when the required frequency resolution (and time resolution) remains fixed across the time–frequency plane.

Other time–frequency analysis techniques can be used aside from the CWT and the STFT. Indeed it can be shown that the absolute squares of the CWT and the STFT are members of more general classes of quadratic time–frequency representations (the affine class for the absolute square of the CWT, and the Cohen class for the absolute square of the STFT [19]). Indeed, other wavelet-based analysis techniques exist. Of particular interest are wavelet packets [25,27], which provide a more general tiling of the time-frequency plane than does the CWT. Rather than being restricted to good frequency resolution at low frequencies and poor frequency resolution at high frequencies, wavelet packets permit good frequency resolution to be achieved at arbitrary analysis frequencies. The product of the time and frequency resolutions, of course, remains fixed. Yet other signal-dependent wavelet techniques have also been developed [27].

9.6. CONCLUSION

Time–frequency analysis has been found to be useful for the analysis of cochlear cellular responses to AM acoustic stimuli. The time course of cellular vibration in the inner ear of the living cat has been successfully studied using both the STFT and an appropriately chosen CWT. Similar results, but with lower carrier-to-noise ratios, were obtained with the round-window membrane intact. Both analysis techniques reveal the time course of the nonlinear dynamics.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research under Grant N00014-92-J-1251, by the National Institutes of Health through NIDCD Program Project Grant DC00316, and by the Emil Capita Foundation. We thank C. Herley, S. B. Lowen, and R. G. Turcott for helpful suggestions in the preparation of the manuscript.

REFERENCES

[1] G. von Békésy, *Experiments in Hearing*. New York, NY: McGraw-Hill, 1960; Huntington, NY: Krieger, 1980.

[2] International Team for Ear Research (ITER), "Cellular vibration and motility in the organ of Corti, *Acta Otolaryngologica* (Stockholm) Supplement 467, pp. 1–279, 1989.

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

[3] M. C. Teich, S. M. Khanna, and S. E. Keilson, "Nonlinear dynamics of cellular vibrations in the organ of Corti," *Acta Otolaryngologica* (Stockholm) Supplement 467, pp. 265–279, 1989.

[4] M. C. Teich, S. E. Keilson, S. M. Khanna, L. Brundin, M. Ulfendahl, and Å. Flock, "Chaos in the cochlea." In *Abstracts 14th Midwinter Res. Mtg. Assoc. Res. Otolaryngology,* February 3–7, p. 50, Abstract No. 155, 1991.

[5] M. C. Teich, S. E. Keilson, S. M. Khanna, L. Brundin, M. Ulfendahl, and Å. Flock, "Chaotic vibrations of outer hair cells and Hensen's cells in the cochlea." In *Abstracts 15th Midwinter Res. Mtg. Assoc. Res. Otolaryngology,* February 2–6, p. 17, Abstract No. 41, 1992.

[6] M. C. Teich, C. Heneghan, S. M. Khanna, Å. Flock, L. Brundin, and M. Ulfendahl, "Analysis of dynamical motion of sensory cells in the organ of Corti using the spectrogram." In *Biophysics of Hair Cell Sensory Systems.* H. Duifhuis, J. W. Horst P.van Dijk, and S. van Netten (eds.), Singapore: World Scientific, pp. 272–279, 1993.

[7] C. Heneghan, M. C. Teich, S. M. Khanna, and M. Ulfendahl, "Nonlinear dynamical motion of cellular structures in the cochlea." *Proc. SPIE,* vol. 2036, pp. 183–197, 1993.

[8] S. E. Keilson, S. M. Khanna, M. Ulfendahl, and M. C. Teich, "Spontaneous cellular vibrations in the guinea-pig cochlea," *Acta Otolaryngologica* (Stockholm), vol. 113, pp. 591–597, 1993.

[9] S. M. Khanna, S. E. Keilson, M. Ulfendahl, and M. C. Teich, "Spontaneous cellular vibrations in the guinea-pig temporal-bone preparation," *Br. J. of Audiol.*, vol. 27, pp. 79–83, 1993.

[10] C. Heneghan, M. C. Teich, S. M. Khanna, and M. Ulfendahl, "Analysis of nonlinear cellular dynamics in the cochlea using the continuous wavelet transform and the short-time Fourier transform." In *Proc. IEEE-Signal Processing Int. Symp. Time–Frequency and Time–Scale Analysis*, October 25–28, 1994.

[11] M. C. Teich, C. Heneghan, S. M. Khanna, and M. Ulfendahl, "Investigating cellular vibrations in the cochlea using the continuous wavelet transform and the short-time Fourier transform." In *Proc. 16th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, November 1–6, 1994.

[12] C. Heneghan, S. M. Khanna, Å. Flock, M. Ulfendahl, L. Brundin, and M. C. Teich, "Investigating the nonlinear dynamics of cellular motion in the inner ear using the short-time Fourier transform and the continuous wavelet transform." *IEEE Trans. Signal Proc.*, vol. 42, pp. 3335–3352, 1994.

[13] M. C. Teich, C. Heneghan, S. M. Khanna, Å. Flock, M. Ulfendahl, and L. Brundin, "Investigating routes to chaos in the guinea-pig cochlea using the continuous wavelet transform and the short-time Fourier transform," *Ann. Biomed. Eng.*, vol. 23, pp. 583–607, 1995.

[14] S. M. Khanna, C. Heneghan, and M. C. Teich, "Dynamical motion of cellular structures in the basal turn of the living cat cochlea." In *Abstracts 17th Midwinter Res. Mtg. Assoc. Res. Otolaryngology,* February 6–10, p. 88, Abstract No. 351, 1994.

M. C. Teich, C. Heneghan, and S. M. Khanna, "Analysis of Cellular Vibrations in the Living Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform," in Time-Frequency and Wavelets in Biomedical Engineering, edited by M. Akay (IEEE Press, New York, 1997), ch. 9, pp. 243-269.

[15] M. C. Teich, C. Heneghan, and S. M. Khanna, "Spontaneous cellular vibrations in the basal turn of the living cat cochlea." In *Abstracts 17th Midwinter Res. Mtg. Assoc. Res. Otolaryngology,* February 6–10, p. 88, Abstract No. 352, 1994.

[16] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 1984.

[17] D. Gabor, "Theory of communication," *J. IEE,* vol. 93, pp. 429–457, 1946.

[18] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, pp. 713–726, 1989.

[19] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time–frequency signal representations," *IEEE Signal Proc. Mag.,* vol. 9, no. 2, pp. 21–67, April 1992.

[20] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Proc. Mag.,* vol. 8, no. 4, pp. 14–38, October 1991.

[21] M. Vetterli and C. Herley, "Wavelets and filter banks: theory and design," *IEEE Trans. Signal Proc.,* vol. 40, pp. 2207–2232, 1992.

[22] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[23] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics.* New York: Wiley, pp. 921–924, 1991.

[24] A. Grossmann, R. Kronland-Martinet, and J. Morlet, "Reading and understanding continuous wavelet transforms." In *Wavelets: Time–Frequency Methods and Phase Space.* J. M. Combes, A. Grossmann, and Ph. Tchamitchian (eds.), New York, NY: Springer-Verlag, pp. 2–20, 2nd edn., 1989/1990.

[25] C. K. Chui, *Wavelet Analysis and its Applications.* Boston, MA: Academic Press, 1992.

[26] D. L. Jones and R. G. Baraniuk, "Efficient approximation of the continuous wavelet transform," *Electronics Letts.,* vol. 27, pp. 748–750, 1991.

[27] C. Herley, J. Kovačević, K. Ramchandran, and M. Vetterli, "Tilings of the time–frequency plane: construction of arbitrary orthogonal bases and fast tiling algorithms," *IEEE Trans. Signal Proc.,* vol. 41, pp. 3341–3359, 1993.