Annals of Biomedical Engineering, Vol. 23, pp. 583–607, 1995
Printed in the USA. All rights reserved.

0090-6964/95 $10.50 + .00

Investigating Routes to Chaos in the Guinea-Pig Cochlea Using the Continuous Wavelet Transform and the Short-Time Fourier Transform

M. C. TEICH,* C. HENEGHAN,† S. M. KHANNA,‡ Å. FLOCK,§ M. ULFENDAHL,§ and L. BRUNDIN§

Departments of *Electrical Engineering, *Applied Physics, †Electrical Engineering, Columbia University, New York, NY, ‡Department of Otolaryngology, Columbia College of Physicians and Surgeons, New York, NY, §Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden

Abstract—The continuous wavelet transform (CWT) and the short-time Fourier transform (STFT) were used to analyze the time course of cellular motion in the guinea pig inner ear. The velocity responses of individual outer hair cells and Hensen's cells to amplitude modulated (AM) acoustical signals applied to the ear canal displayed characteristics typical of nonlinear systems, such as the generation of spectral components at harmonics of the carrier frequency. Nonlinear effects were particularly pronounced at the highest stimulus levels, where half-harmonic (and sometimes quarter-harmonic) components were also seen. The generation of these components was consistent with the behavior of a dynamical system entering chaos via a period-doubling route. A negative-stiffness Duffing oscillator model yielded period-doubling behavior similar to that of the experimental data. We compared the effectiveness of the CWT and the STFT for analyzing the responses to AM stimuli. The CWT (calculated using a high-Q Morlet-wavelet basis) and the STFT were both useful for identifying the various spectral components present in the AM velocity response of the cell. The high-Q Morlet wavelet CWT was particularly effective in distinguishing the lowest frequency components present in the response, since its frequency resolution is appreciably better than the STFT at low frequencies. Octave-band-based CWTs (using low-Q Morlet, Meyer, and Daubechies 4-tap wavelets) were largely ineffective in analyzing these signals, inasmuch as the frequency spacing between neighboring spectral components was far less than one octave.

Keywords—Continuous wavelet transform, Short-time Fourier transform, Cochlea, Cellular vibration, Nonlinear dynamics, Chaos, Period-doubling.

Acknowledgment—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, by the Emil Capita Foundation, by the Swedish Medical Research Council (02461), by the Magnus Bergvall Foundation, and by the Tysta Skolan Foundation. We thank B. Flock (Karolinska Institutet) for superb technical assistance in the data collection. We thank C. Herley, S. B. Lowen, and R. G. Turcott for helpful suggestions in the preparation of the manuscript.

Address correspondence to M. C. Teich, Department of Electrical Engineering, 1312 Mudd Building, 500 W. 120th Street, New York, NY 10027, U.S.A.

(Received 1July94, Revised 27Oct94, Revised, Accepted 22Nov94)

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; there are four and a half turns in the guinea pig cochlea. The coil diameter is widest at the base of the cochlea, and narrowest at the apex. There are two openings in the bony shell near the base: (i) the oval window, through which the stapes drives the fluid in the scala vestibuli, and (ii) 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 (2,23).

In the past, it has been possible only to measure vibrations at the basilar membrane. More recently, however, the use of a slit confocal microscope has made it possible to conduct vibration measurements at arbitrary positions within the organ of Corti (23), in the third and fourth turns of a special guinea pig temporal-bone preparation. This preparation is excised from the animal and kept alive by immersion in an oxygenated tissue culture medium (42). The velocity of vibration of individual cells, selected as desired, is 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 techniques have been previously described (23).

This measurement technique has served to elucidate the role of individual cells in the complex mechanical transduction process carried out in the organ of Corti. The measurements show that sensory cells (outer and inner hair cells) play an important role in this process. Measurements at the reticular lamina (which contains the tops of the sensory cells) display three types of response: (i) Cells vibrate at frequencies that are related to the frequency of the applied signal, with no net displacement. This is the *AC response*. The magnitude of the vibration velocity is frequency dependent, displaying a bandpass characteristic. The frequency of maximum velocity is defined as the characteristic frequency (CF) of the cell. The CF is highest at the base of the cochlea, and lowest at the apex (23). In the third turn of the guinea-pig cochlea, the cells respond maximally to frequencies in the range of 500–900 Hz, whereas in the fourth turn the maximum response occurs for frequencies below 500 Hz. (ii) The outer hair cells change their length when a tone is applied to the ear, and retain that length as long as the tone is present. This length change is called the motile or *DC response*. The magnitude of the motile response is dependent both on signal level and signal frequency. For a constant signal level, the length change is largest at the same frequency at which the AC response is maximum (the CF); however the frequency bandwidth over which the motile response is observed is far narrower than the AC passband (4–8,23). (iii) Some outer hair cells and Hensen's cells vibrate *spontaneously* with low amplitude at frequencies near their CF even when no acoustic signal is applied to the ear (27,28,30).

A suitable method for studying both the AC and motile responses is to use sinusoidal-carrier amplitude modulated (AM) acoustic waves with low-frequency modulation. This provides an opportunity for studying the change in cell length over a broad range of carrier levels, as the envelope 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.

An AM stimulus is a time-varying signal; accordingly specific analysis techniques are required to examine the response. This paper describes the application of time-scale and time-frequency representation techniques for the analysis of cellular velocity data, with a particular eye toward examining routes to chaos (38,39). 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 2.44 Hz. The carrier frequency ranged from 24 to 1800 Hz, and the total duration of each data set collected was fixed at 0.4096 s (representing 2048 samples at 200-μs intervals).

We examined both the continuous wavelet transform (CWT) and the short-time Fourier transform (STFT) of the velocity responses elicited by the AM stimuli described above. Both analysis techniques were useful in discriminating the frequency components present in the responses, and revealing period-doubling behavior, 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 identifying low-frequency 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 between the closely spaced frequency components present in the velocity responses.

In this paper, we show that the nonlinear dynamical characteristics of cellular motion in the cochlea (37–39) can be elucidated by using both CWT- and STFT-based analyses. The character of the cell's nonlinear response is found to depend significantly on the carrier frequency of the applied signal, relative to the innate tuning characteristics of the cell, and on the signal level. A simple nonlinear system, the Duffing oscillator with negative stiffness, exhibits some of the features seen in our experimental data, though it is clearly too idealized to serve as a model for cellular dynamics in the cochlea. We have previously reported results pertaining to the use of both the CWT and the STFT in the analysis of cellular velocity data (18–20,40,41), and provided preliminary results pertaining to chaos in the cochlea (38,39).

THEORY

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

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

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

$X(f)$ 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 $x(t)$.

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 33). 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 infor-

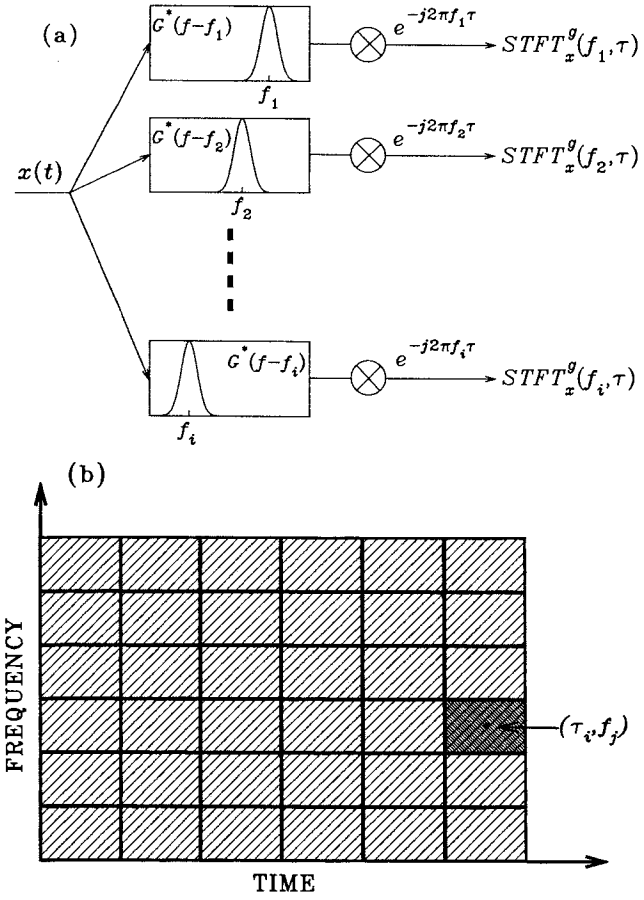


FIGURE 1. (a) Representation of the STFT in terms of filter-bank operations. The signal to be analyzed $x(t)$ is passed through a bank of filters, each with the same shape and bandwidth, but slightly different center frequencies f_i. The output from each filter is then multiplied by the factor $\exp(-j2\pi f_i\tau)$, 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.

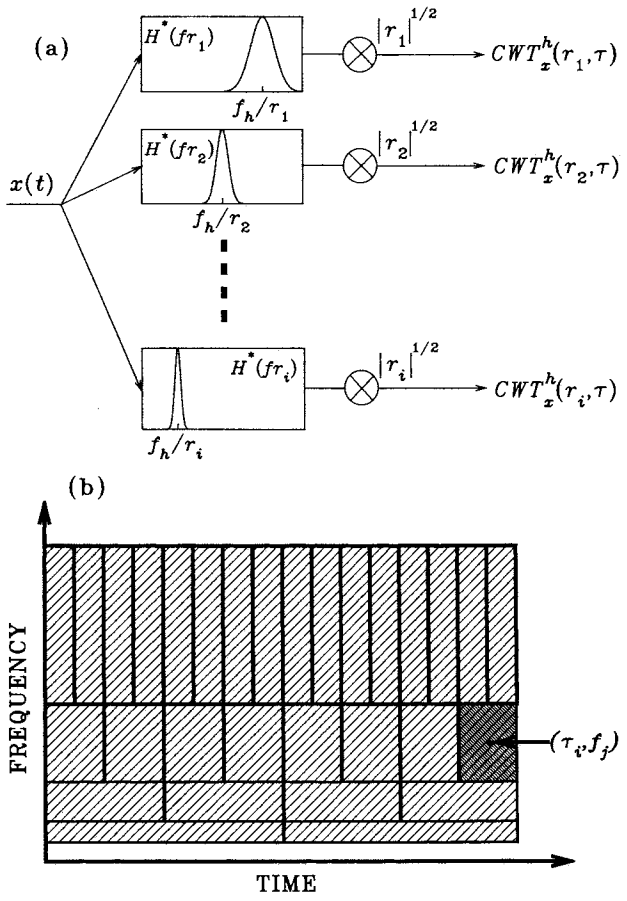


FIGURE 2. (a) Representation of the CWT in terms of filter-bank operations. The signal to be analyzed $x(t)$ is passed through a bank of filters, each of which is simply a scaled version of some prototype filter. Each filter has a fixed shape and relative bandwidth, but in absolute terms the bandwidths 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.

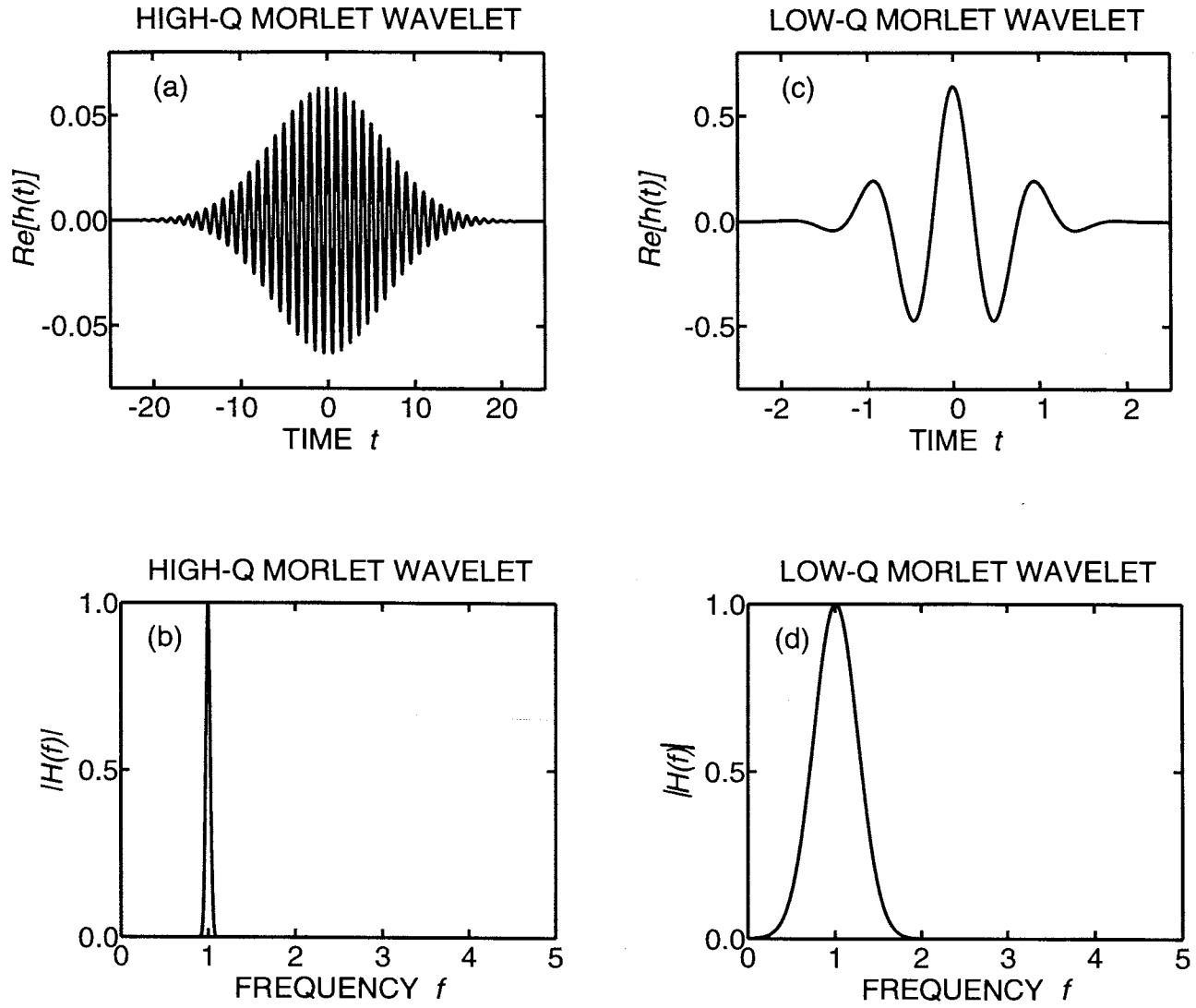


FIGURE 3. Four wavelet bases in the time domain and their respective CFTs. (a) Real part of a high-Q Morlet wavelet, $h(t) = \exp(jct)\exp(-\alpha t^2/2)$, with $c = 2\pi$ and $\alpha = 0.0254$. There are many (>40) oscillations within the envelope of the wavelet, and its overall duration is approximately 40. (b) Magnitude of the CFT of the wavelet shown in Panel a. The relative bandwidth of this function is narrow (=0.072). 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 $\alpha = 2.54$. There are relatively few oscillations within the envelope of the wavelet, and its overall duration is approximately 4. (d) Magnitude of the CFT of the wavelet shown in Panel c. The relative bandwidth of this function is 0.72. It is symmetric about the frequency $f = 1$. (e) Real part of a Meyer wavelet calculated as outlined in Ref. 44 (there is no closed-form expression for this wavelet in the time domain). As for the Morlet wavelet pictured in Panel 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 Panel 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) (44). 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 Panel g. The relative bandwidth of the first bandpass region is 1.5. This function has multiple lobes in the frequency domain.

mation about the time order in which they occurred. A representation of $x(t)$ that tells us *when* a particular frequency was present would provide 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 (16). His approach retained the frequency variable f defined by the CFT, but insured 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.

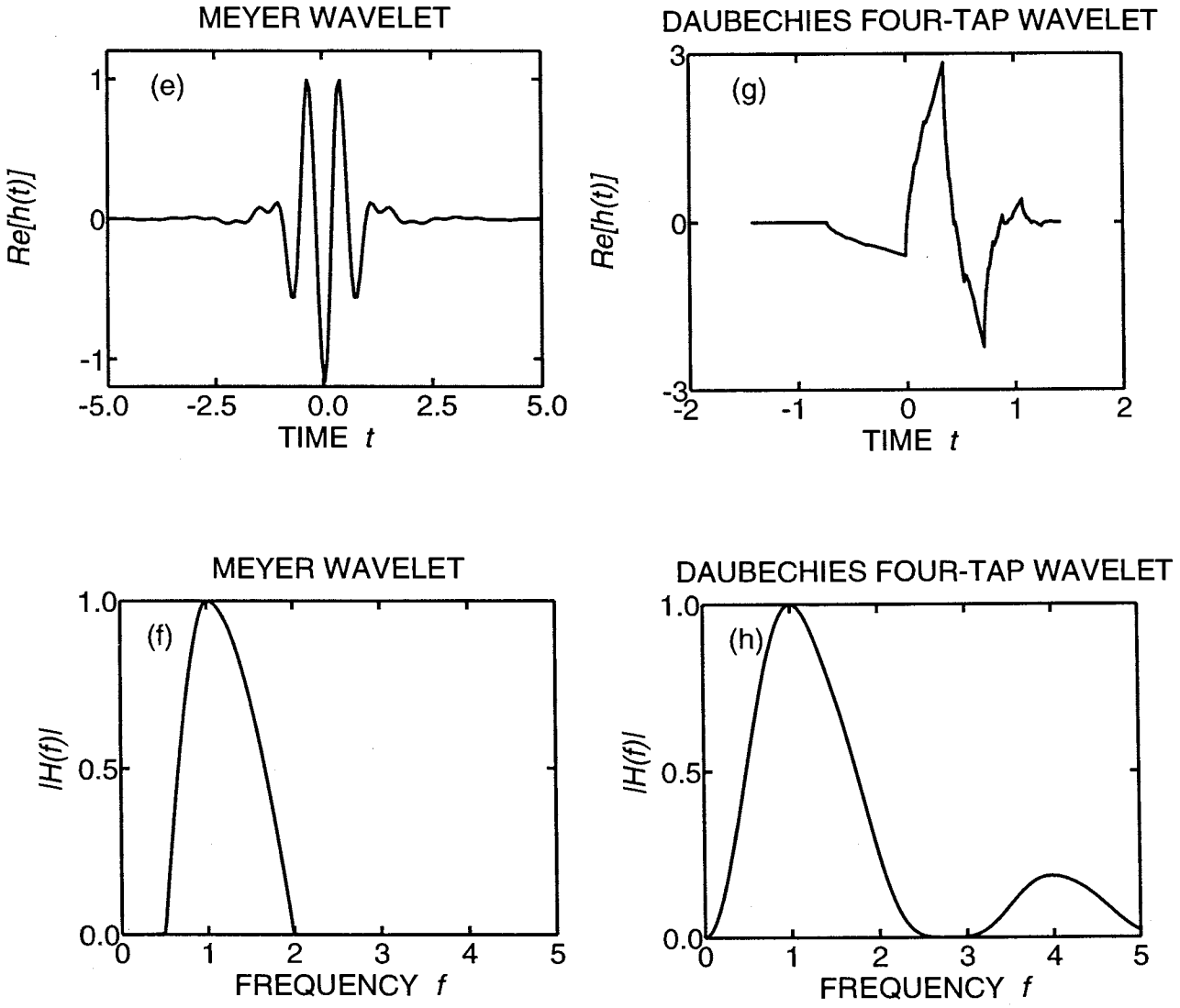


FIGURE 3. Continued.

Gabor's approach was later shown to be a special case of the short-time Fourier transform (STFT) (22,32), which is expressed as:

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

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 asterisk 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. 3 reduces to the CFT given in Eq. 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.*:

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

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. 2). If the function $G(u)$ is taken to represent a lowpass 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. 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 frequency. This filter-bank interpretation of the STFT (43,44) is illustrated in Fig. 1a.

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

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

or in the frequency domain:

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

where the on-line asterisk 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. 5 identifies a limitation of the STFT. In calculating $\mathrm{STFT}_x^g(f,\tau_0)$, 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 (35); in this paper, 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. 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 we would be delighted 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. 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 (35) (*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 once. Thus, $g(t)$ 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. 1b by drawing regions in the τ-f plane where a set of functions $g^*(\tau\text{-}\tau_i)\exp(-j2\pi f_j\tau)$ 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 (43,44).

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, $\Delta f = C/\Delta t$ 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

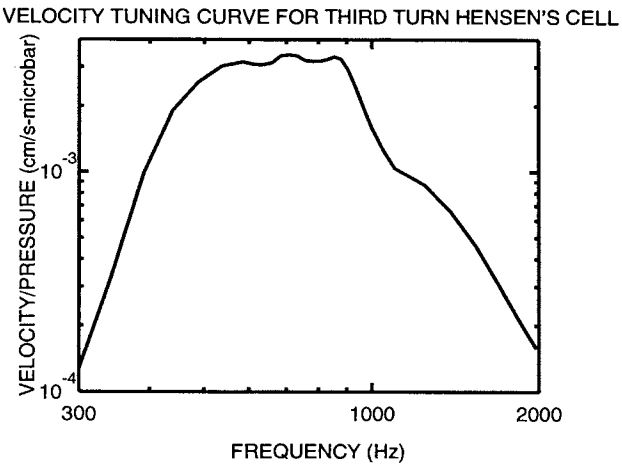


FIGURE 4. Velocity tuning curve for a Hensen's cell (file number 0b181153.dt1) in the third turn of the guinea pig temporal-bone preparation. The ordinate displays the peak velocity in cm/s per unit sound pressure of 1 μbar (1 dyne/cm^2). The CF for this cell is 693 Hz, but the tuning is rather broad.

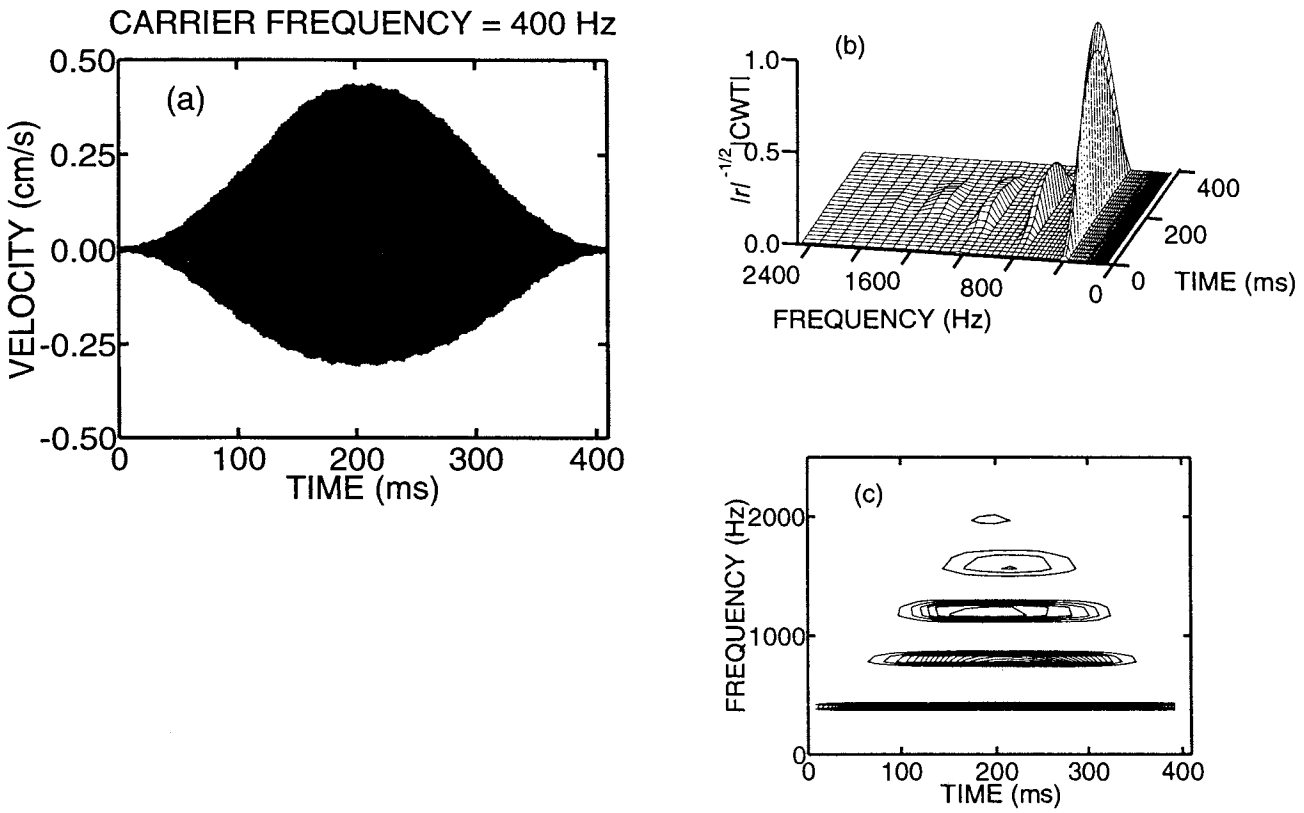


FIGURE 5. Velocity response of a Hensen's cell (file number 0b181444.dt1) in the third turn of the guinea pig temporal-bone preparation to an AM stimulus with carrier frequency f_c = 400 Hz (below CF) and modulation frequency 2.44 Hz. The velocity tuning curve for this cell is shown in Fig. 4. The highest sound pressure level, occurring at the center of the input envelope was ≈116 dB:re 0.0002 dyne/cm^2. (a) Time waveform of the velocity response (in cm/s) of the cell. The waveform is asymmetric in amplitude. (b) 3D plot of the modified CWT magnitude (*viz.*, the CWT multiplied by the factor of $|r|^{-1/2}$) of the velocity response shown in Panel a). The high-*Q* Morlet wavelet basis shown in Fig. 3a and 3b was used. The *x* and *y* axes represent time (ms) and frequency (Hz) 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 four higher harmonic frequencies, $2f_c$, $3f_c$, $4f_c$ and $5f_c$. (c) Same modified CWT magnitude as shown in Panel b, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (d) 3D plot of the STFT magnitude of the velocity response shown in Panel a. Spectral components are present at the carrier frequency f_c and at five higher harmonics, $2f_c$, $3f_c$, $4f_c$, $5f_c$ and $6f_c$. (e) Same STFT magnitude as shown in Panel d, but now plotted in 2D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude. (f) 3D plot of the modified CWT magnitude of the velocity response shown in Panel a. The low-*Q* Morlet wavelet basis 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 Panel f, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (h) 3D plot of the modified CWT magnitude of the velocity response shown in Panel a. The Meyer wavelet basis was used. As in Panel f, this plot shows a large-bandwidth spectral component centered at the carrier frequency f_c and higher frequency spectral components that cannot be clearly resolved. (i) Same modified CWT magnitude as shown in Panel h, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (j) 3D plot of the modified CWT magnitude of the velocity response shown in Panel a. The Daubechies 4-tap wavelet basis was used. As in Panels f and h, this plot shows a large-bandwidth spectral component at the carrier frequency f_c and higher frequency spectral components that cannot be clearly resolved. (k) Same modified CWT magnitude as shown in Panel j, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude.

resolution to separate the closely spaced low-frequency components. These two aims are incompatible with the STFT—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 (CWT) (9,17,22,34,43,44). The CWT is expressed as:

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

where t and τ are time variables, $x(t)$ is the time waveform

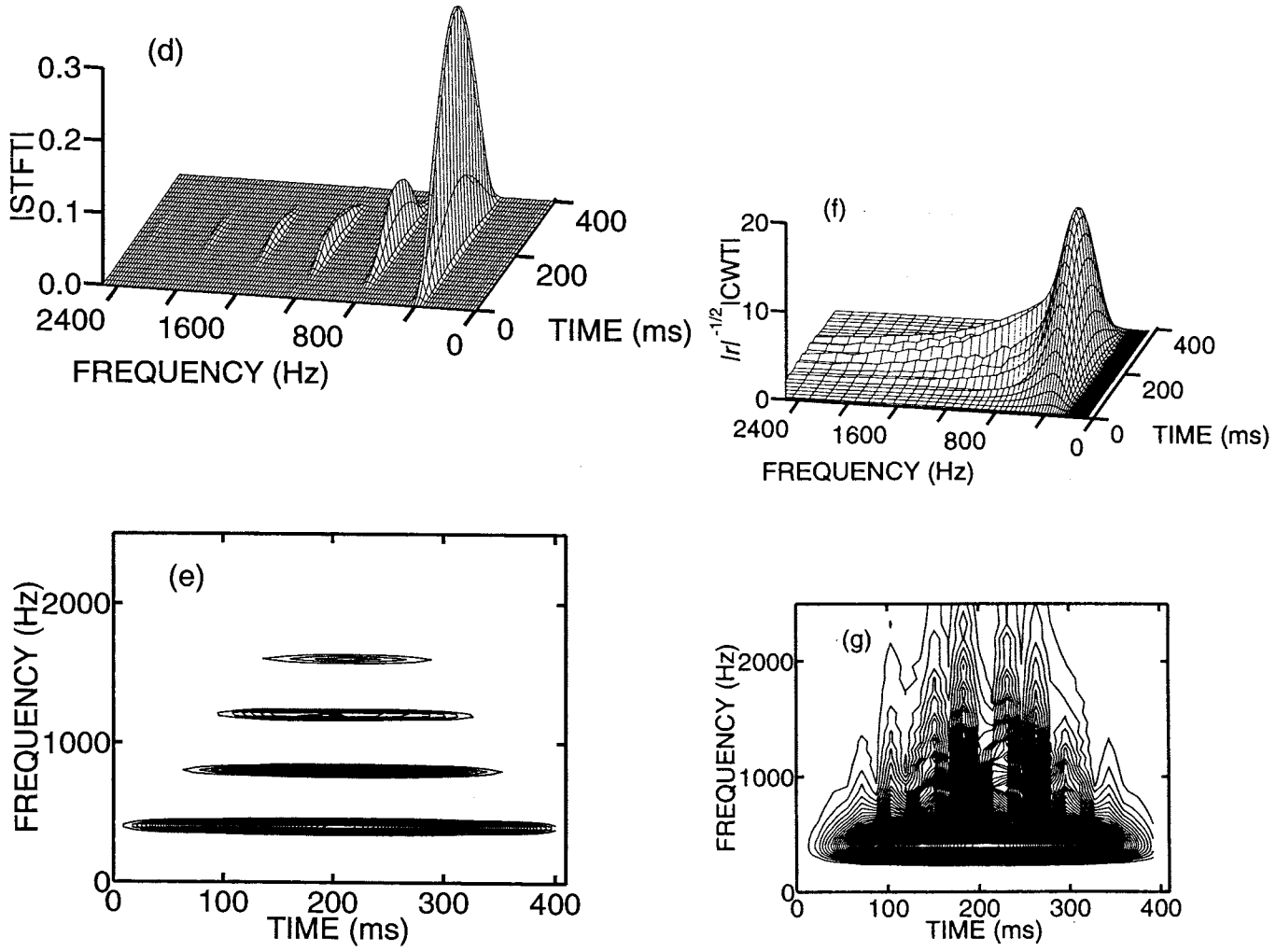


FIGURE 5. Continued.

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. 4):

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

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

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

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. 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, 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 follow-

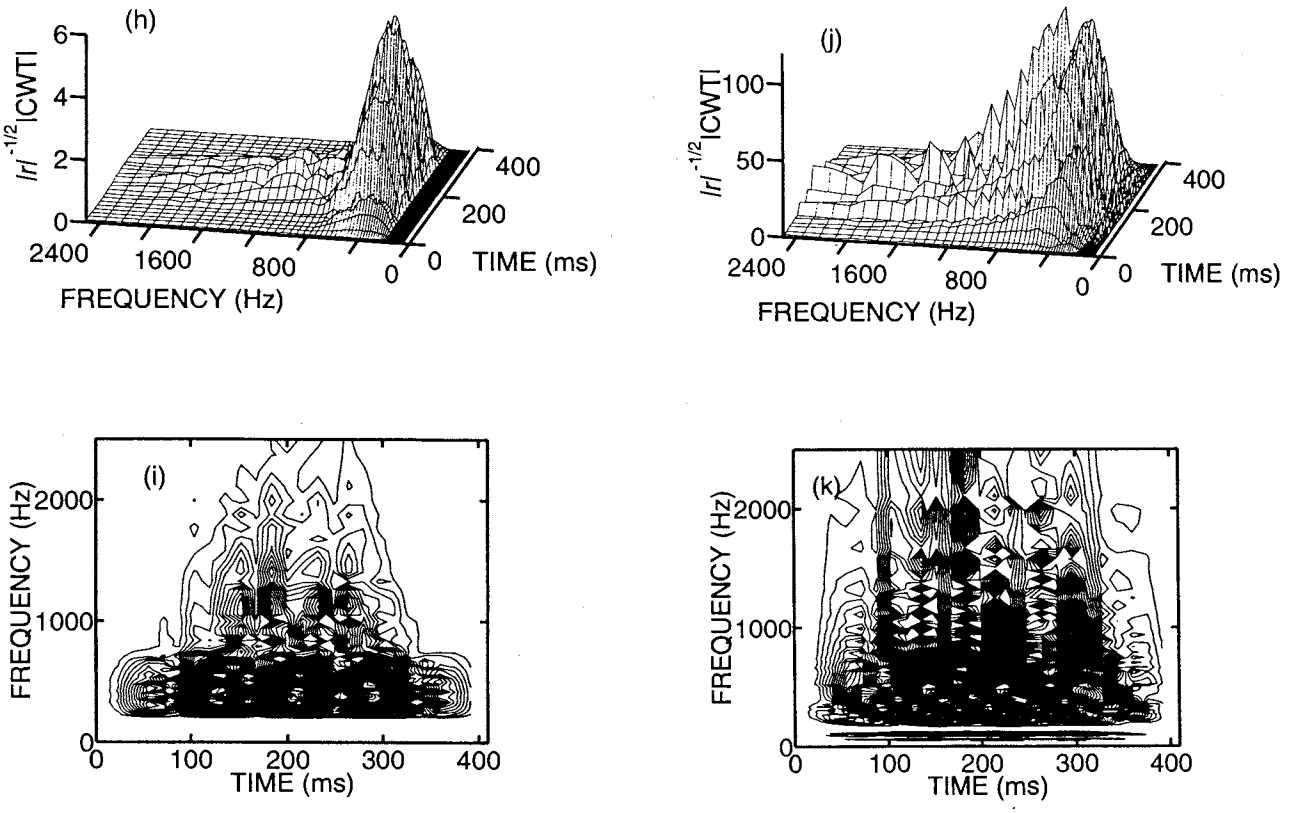


FIGURE 5. Continued.

ing situation obtains: as r decreases, $h(\tau/r)$ becomes narrower in time so that the time resolution 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. 8, where $H(ur)$ becomes narrower as r increases, hence 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 above.

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.

As with the STFT, a filter-bank interpretation (43,44) can be invoked for the CWT, as illustrated in Fig. 2a. 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 (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:

$$BW_{rel} = \frac{\Delta f}{f_h} = \frac{1}{Q}. \tag{10}$$

For the CWT filter-bank illustrated in Fig. 2a, 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. 2b; 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.

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 paper 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" (17,34,44), which states that if $h(t)$ is a wavelet basis for $\mathcal{L}^2$ (the set of square-integrable functions), then:

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

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. 3, we show four different wavelet bases in the time domain, along with the magnitudes of the 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 (Fig. 3a, c, and e) are complex and only the real part of the wavelet is plotted; the Daubechies 4-tap wavelet (Fig. 3g) is purely real.

Figure 3a shows the real part of the single-sided Morlet wavelet (17) given by $h(t) = \exp(jct)\exp(-\alpha t^2/2)$ [with $c = 2\pi$ and $\alpha = 0.0254$], and the magnitude of its Fourier transform (Fig. 3b), which is given by

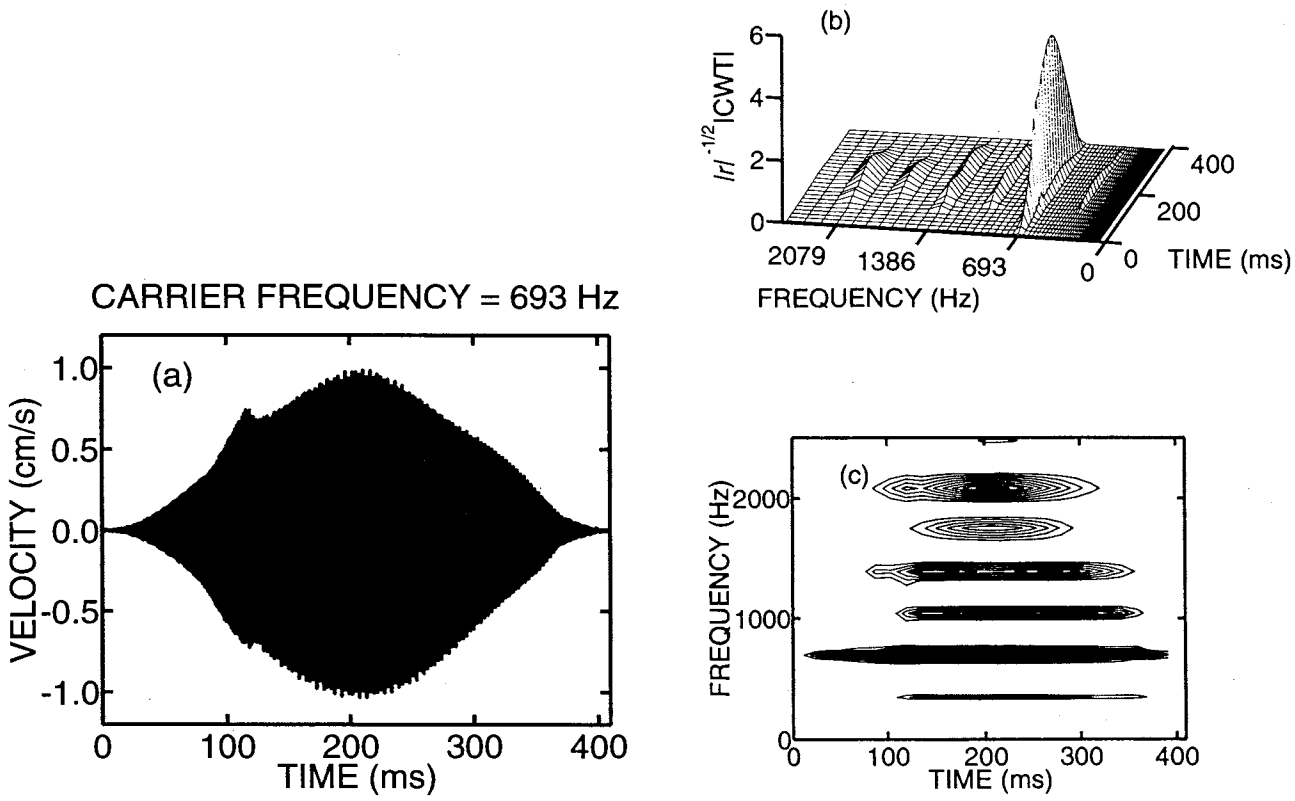


FIGURE 6. Velocity response of the same Hensen's cell as shown in Fig. 5, but now to an AM stimulus with carrier frequency f_c = 693 Hz (at CF) and modulation frequency 2.44 Hz. The highest sound pressure level, occurring at the center of the input envelope was ≈130 dB:re 0.0002 dyne/cm^2. (a) Time waveform of the velocity response (in cm/s) of the cell. There is a distinct chink in the response at t = 115 ms. (b) 3D plot of the modified CWT magnitude of the velocity response shown in Panel a. The high-Q Morlet wavelet basis was used. This plot shows spectral components at the carrier frequency f_c and at two higher harmonic frequencies, $2f_c$, and $3f_c$. Components at three half-harmonic frequencies ($f_c/2$, $3f_c/2$, and $5f_c/2$) are also present, and their sudden appearance corresponds in time to the chink apparent in the time waveform at t = 115 ms. (c) Same modified CWT magnitude as shown in Panel b, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (d) 3D plot of the STFT magnitude of the velocity response shown in Panel a. Spectral components are present at the carrier frequency f_c, at two higher harmonics ($2f_c$ and $3f_c$), and at four half-harmonic frequencies ($f_c/2$, $3f_c/2$, $5f_c/2$, and $7f_c/2$). (e) Same STFT magnitude as shown in Panel d, but now plotted in 2D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude.

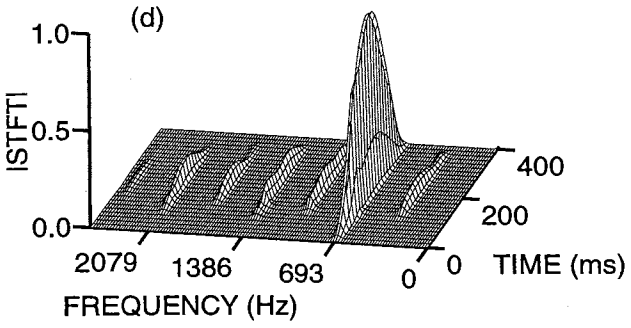
$\sqrt{2\pi/\alpha}\exp[-(2\pi f - c)^2/2\alpha]$. For the values of c and α that we have chosen here, $h(t)$ is a windowed sinusoid and $H(f)$ is a narrow bandpass filter centered at $f = 1$. The relative bandwidth of this wavelet is readily calculated to be:

$$BW_{rel} = \frac{2\sqrt{2\alpha}}{c}, \tag{12}$$

which equals 0.072 for this choice of parameters. Consequently, the Q-factor for this wavelet is $Q = 1/0.072 = 13.9 \gg 1$, which is why it is referred to as a high-Q Morlet wavelet in the remainder of this paper.

Figure 3c and d also shows a Morlet wavelet, but this time with $c = 2\pi$ and $\alpha = 2.54$. As in Fig. 3b, this Morlet wavelet is also a bandpass function; however its relative bandwidth ($=0.72$) 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. 3a. This wavelet is referred to as a low-Q Morlet wavelet in the text and figures that follow since its Q-factor is only 1.39.

Figure 3e and f shows the Meyer wavelet similarly normalized (for details concerning the construction of the Meyer wavelet see 44). Its structure is similar in both the frequency and time domains to the Morlet wavelet shown in Fig. 3c 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.



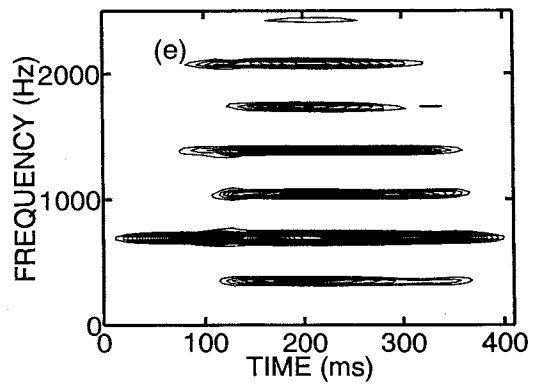


FIGURE 6. Continued.

Finally, in Fig. 3g 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 43 and 44). 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 the Results section, 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 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. Firstly, the Morlet wavelet does not strictly satisfy the admissibility condition since $H(0) \neq 0$. However, for $BW_{rel} \leq 0.8$, 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 Fig. 3a and c, $BW_{rel} \approx 0.072$ and 0.72, respectively. Secondly, 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) (9,34,43,44). 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 Fig. 3c–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 (43,44). 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-

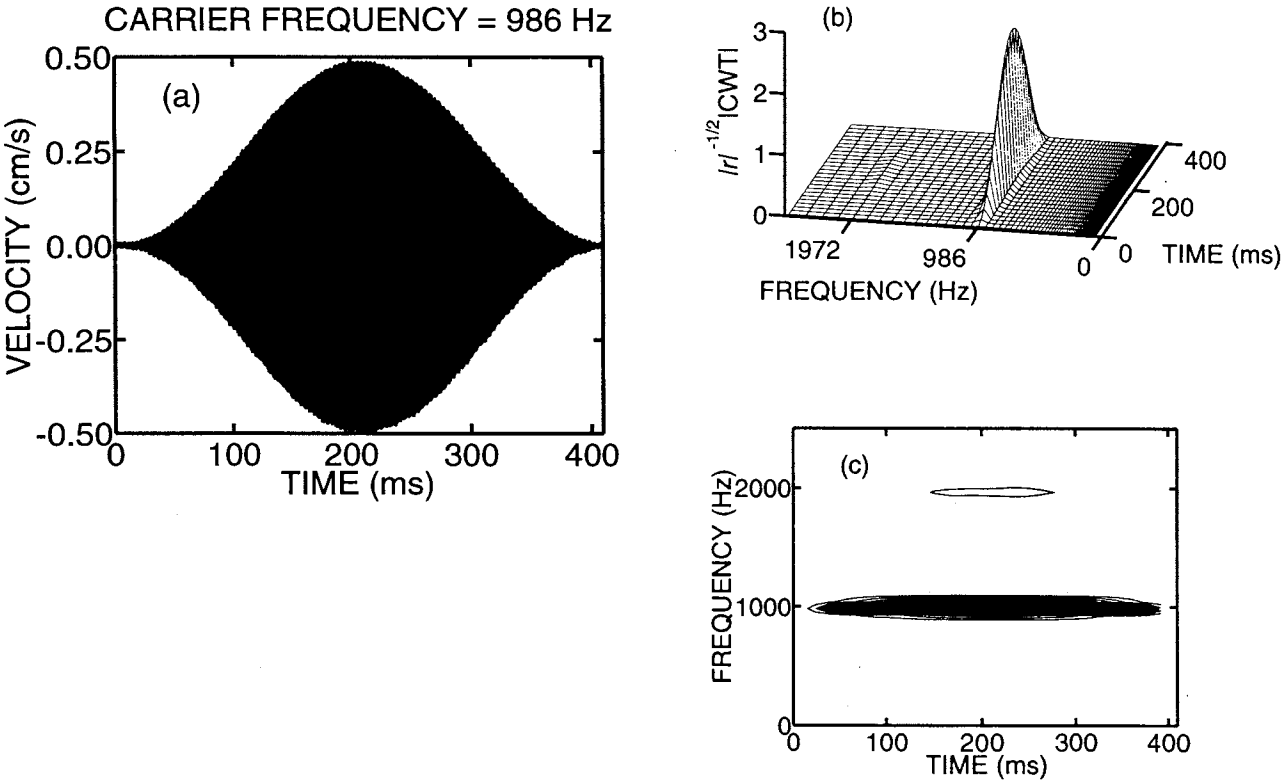


FIGURE 7. Velocity response of the same Hensen's cell as shown in Figs. 5 and 6, but now to an AM stimulus with carrier frequency f_c = 986 Hz (above CF) and modulation frequency 2.44 Hz. The highest sound pressure level, occurring at the center of the input envelope was ≈124 dB:re 0.0002 dyne/cm^2. (a) Time waveform of the velocity response (in cm/s) of the cell. The response follows the envelope of the input closely. (b) 3D plot of the modified CWT magnitude of the velocity response shown in Panel a. The high-Q Morlet wavelet basis was used. This plot shows spectral components at the carrier frequency f_c and at $2f_c$. (c) Same modified CWT magnitude as shown in Panel b, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (d) 3D plot of the STFT magnitude of the velocity response shown in Panel a. This plot shows a spectral component at the carrier frequency f_c, along with a just-visible component at $2f_c$. (e) Same STFT magnitude as shown in Panel d, but now plotted in 2D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude.

basis-generating wavelets do provide a useful set of functions for constructing efficient wavelet series expansions of $x(t)$ (44), they are not always suited for use as a CWT basis, as our examples will show.

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. 3:

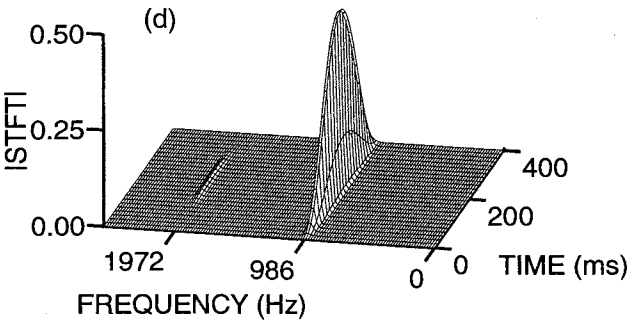
$$\mathrm{STFT}_x^g[k,n] = \sum_{m=0}^{L-1} x[n+m]g^*[m]\exp\left(\frac{-j2\pi mk}{N}\right),$$

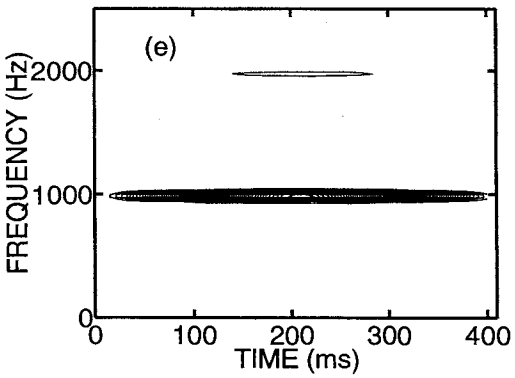
$$0 \leq k \leq N-1, \qquad (13)$$

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:

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

The formulation presented in Eq. 13 reminds us that the discrete STFT is simply a sequence of discrete Fourier transforms (DFTs) of the windowed signal segments. 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 $L = 256$ samples was appropriate. This corresponds to 51.2 ms at the 5-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 80 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 $\Delta t = 25.6$ ms, and the frequency resolution was $\Delta f = 49.7$ Hz. The uncertainty product $\Delta t \Delta f = C = 4/\pi \approx 1.27$.

We present the STFT magnitude in two visual formats. The first is a three-dimensional representation, often referred to as a 3D 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 2D 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 (24). 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. 8, in which the CWT is written as an integral in the frequency domain:

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

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. 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. 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 (32). 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. 15 in terms of the DTFTs of $x[n]$ and $h[n]$:

$$\mathrm{CWT}_x^h(r,\tau) =$$

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

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,

$$\mathrm{CWT}_x^h[r_i,n] =$$

$$\sqrt{|r_i|}\,\frac{T_x}{2\pi}\int_{-\pi}^{+\pi} X(e^{j\omega})H^*(e^{j\omega r_i})\exp(j\omega n)d\omega, \tag{17}$$

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:

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

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? 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 32).

To choose an appropriate value of M, we must insure 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. 15. This is insured 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 two greater than $N_x + \max(r_i)N_h - 1$ is chosen, so that power-of-two 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. 15, we must insure that aliased versions of $H(e^{j\omega r_i})$ are not brought into the range $[-\pi,\pi]$.

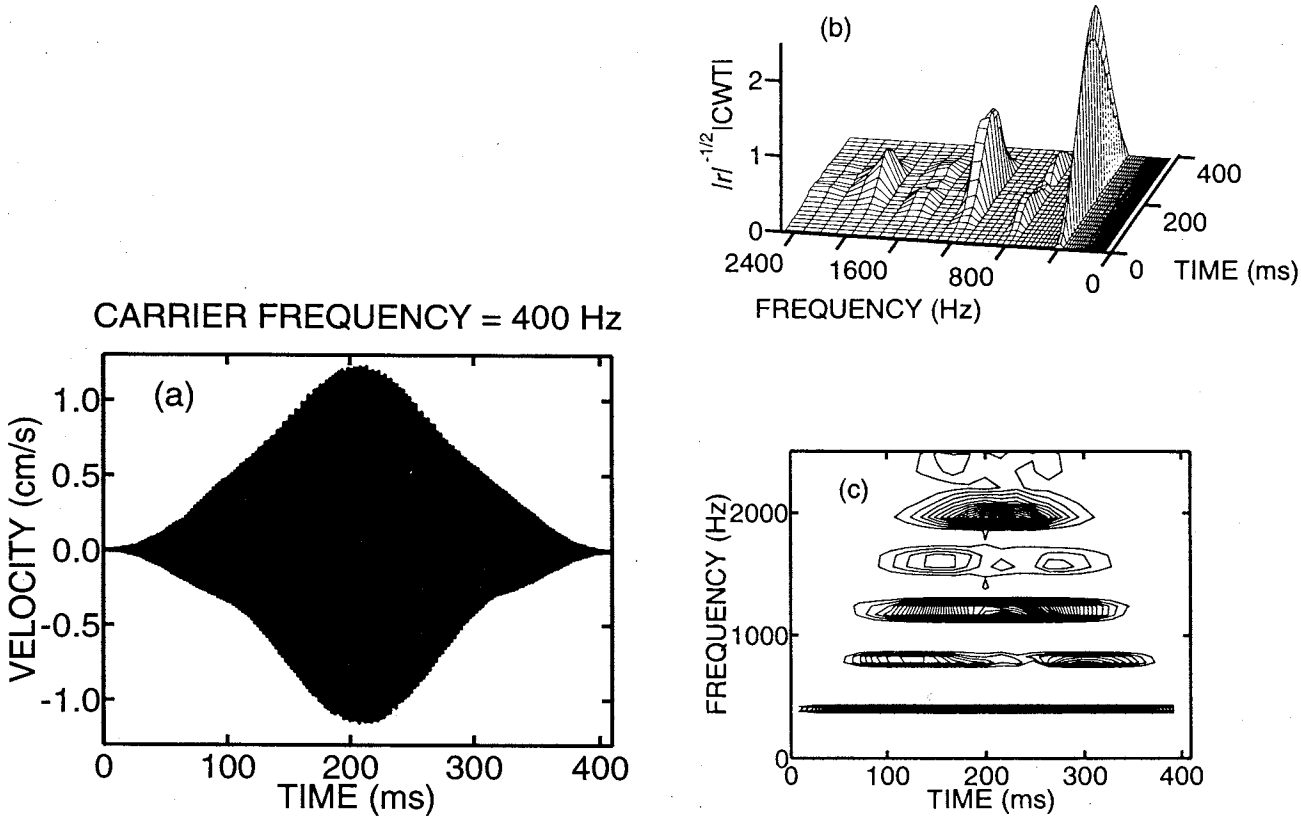
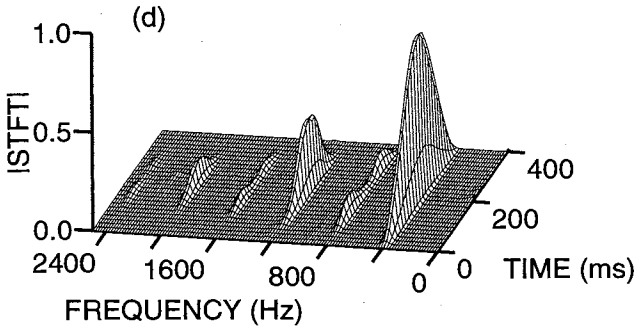
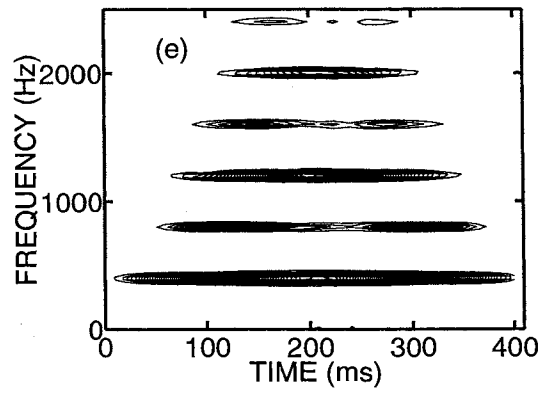


FIGURE 8. Velocity response of the same Hensen's cell as shown in Figs. 5–7, but now to an AM stimulus with carrier frequency f_c = 400 Hz (below CF), modulation frequency 2.44 Hz, and an 8-dB increase in sound pressure level in comparison with Fig. 5. The highest sound pressure level is now ≈124 dB:re 0.0002 dyne/cm^2. (a) Time waveform of the velocity response (in cm/s) of the cell. (b) 3D plot of the modified CWT magnitude of the velocity response shown in Panel a. The high-Q Morlet wavelet basis was used. This plot shows spectral components at the carrier frequency f_c, and at four higher harmonic components ($2f_c$, $3f_c$, $4f_c$ and $5f_c$). (c) Same modified CWT magnitude as shown in Panel b, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (d) 3D plot of the STFT magnitude of the velocity response shown in Panel a. Spectral components are present at the carrier frequency f_c, and at five higher harmonic components ($2f_c$, $3f_c$, $4f_c$, $5f_c$ and $6f_c$). Near the center of the modulation cycle, where the stimulus is strongest, the components at odd harmonics ($3f_c$ and $5f_c$) are greater than those at even harmonics. (e) Same STFT magnitude as shown in Panel d, but now plotted in 2D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude.

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. 3. To provide a complete representation, but without undue oversampling, $h[n]$ is constructed by sampling $h(t)$ at a different rate $(1/T_h)$ compared with the sampling rate $1/T_x$ for $x[n]$. However, in evaluating the CWT as described above, 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) (43,44). The DWT provides a multiresolution approximation of the sequence $x[n]$. This approximation involves repeated highpass and lowpass filtering of the original signal $x[n]$, with downsampling by a factor of two after each filtering operation. The various highpass sequences are retained as a useful approximation of the signal. The number of samples contained in the complete set of highpass sequences plus the residual lowpass sequence is equal to the number in the original signal, and the highpass sequences contain nearly all the information of the original signal (however the DC component is lost as it is contained in the lowpass residual). This process decomposes the sequence $x[n]$ into various frequency bands, as represented by the set of highpass sequences. The center frequency of each band differs by a factor of two as a result of the downsampling 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 two; 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 (3D) 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 2D contour plots, on which contours of equal modified CWT magnitude are traced on a time-frequency plane.

RESULTS

We used the modified CWT and STFT to examine the behavior of the velocity response of cochlear hair cells and Hensen's cells. AM acoustic signals (with unity modulation depth and a fixed modulation frequency of 2.44 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 well above the CF, and as the stimulus intensity was varied. To investigate routes to chaos in the cellular motion, we examine velocity responses both at high and at very high peak intensities (but below the threshold for damage). The sampling frequency was fixed at 5 kilosamples/s in recording the responses.

Figure 4 shows the velocity tuning curve of a Hensen's cell in the third turn of the guinea pig temporal-bone preparation. The measured peak velocity of the cell, per unit applied sound pressure, is shown over frequencies ranging from 300 to 2000 Hz. This tuning curve displays its largest value at 693 Hz, but it is rather broad.

Responses to AM Stimuli of High Intensity

Figure 5 shows the response of the cell when an AM tone with a carrier frequency f_c = 400 Hz (below CF), and a peak sound pressure level of ≈116 dB:re 0.0002 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. 5a. The response does not follow precisely the shape of the input envelope; rather the envelope is skewed

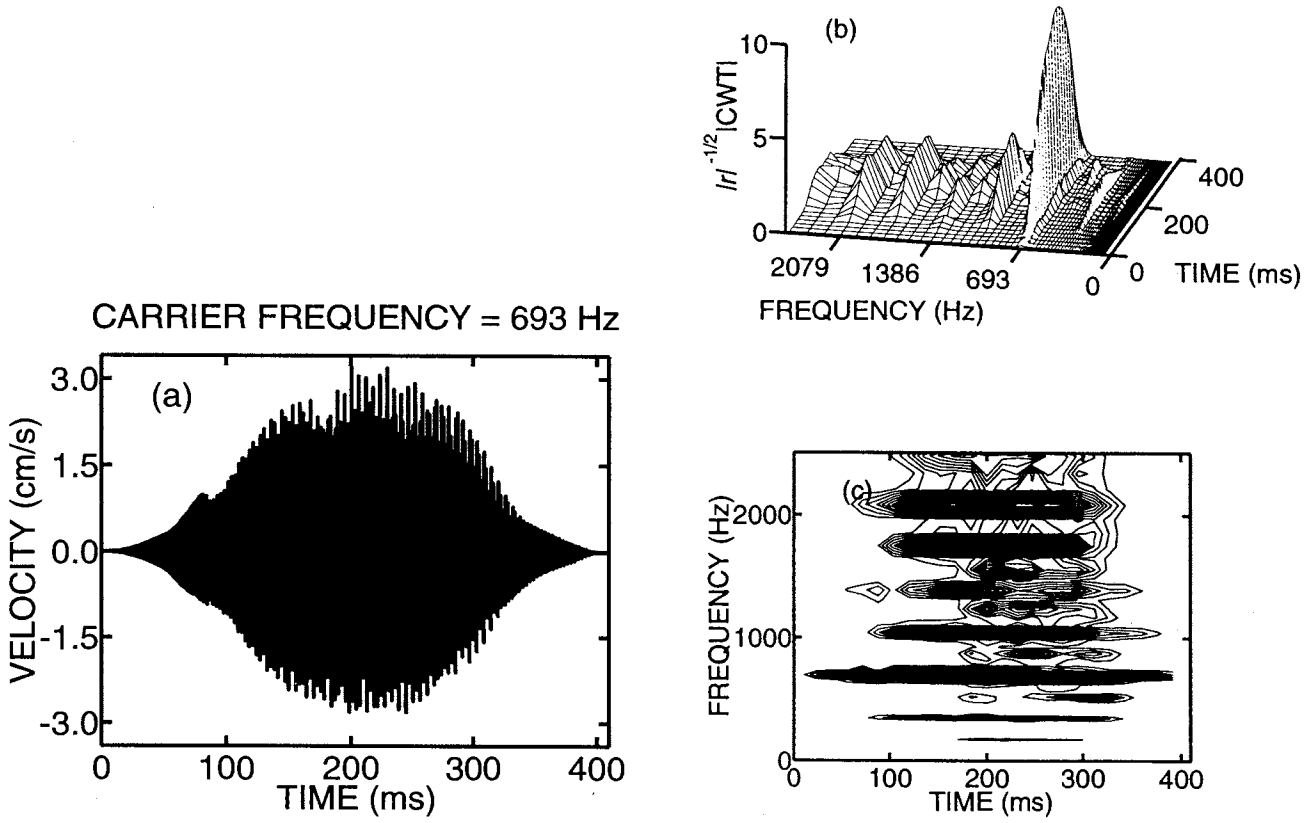


FIGURE 9. Velocity response of the same Hensen's cell as shown in Figs. 5–8, but now to an AM stimulus with carrier frequency f_c = 693 Hz (at CF), modulation frequency 2.44 Hz, and an 8-dB increase in sound pressure level in comparison with Fig. 6. The highest sound pressure level is now ≈138 dB:re 0.0002 dyne/cm^2. (a) Time waveform of the velocity response (in cm/s) of the cell. The waveform is jagged, and only approximately follows the envelope of the stimulus. A chink appears in the waveform at approximately the same velocity value as in Fig. 6a. (b) 3D plot of the modified CWT magnitude of the velocity response shown in Panel a. The high-*Q* Morlet wavelet basis was used. This plot shows the presence of many spectral components. Firstly, as for the lower sound pressure level shown in Fig. 6b, there are spectral components at the carrier frequency f_c, and at two higher harmonics ($2f_c$ and $3f_c$). Secondly, again as before, half-harmonic components are present ($f_c/2$, $3f_c/2$, $5f_c/2$, and $7f_c/2$). Finally, when the envelope is near its maximum we see the emergence of multiple quarter-harmonic components ($f_c/4$, $3f_c/4$, $5f_c/4$, $7f_c/4$, and $9f_c/4$ are clearly visible). (c) Same modified CWT magnitude as shown in Panel b, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. Note the good frequency resolution at the lowest frequency components. (d) 3D plots of the STFT magnitude of the velocity response shown in Panel a. This plot shows the same type of behavior as does the CWT, though the improved frequency resolution at high frequencies now allows us to identify additional quarter-harmonic components at $11f_c/4$ and $13f_c/4$. (e) Same STFT magnitude as shown in Panel d, but now plotted in 2D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude.

towards positive velocity values. In this, and other such curves, the reconstructed waveform has been upsampled by a factor of 4 and interpolated by sinc functions to reduce the artifact known as "false modulation."

The modified CWT magnitude of the velocity response, shown in 3D and 2D formats in Fig. 5b 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 $K = 491$. This CWT was based on the high-Q Morlet wavelet shown in Fig. 3a and b. The modified CWT shows five spectral components at multiples of 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 principally at the carrier frequency. Because the CWT is only shown at a restricted number of time points, the plot presented in Fig. 5c, and indeed all 2D plots shown subsequently, do not extend all the way to the left and right borders. As the magnitude of the envelope increases, a spectral component at $2f_c$ appears, followed later by components at $3f_c$ and $4f_c$ as the envelope increases. Then, as the envelope decreases, the spectral components disappear in inverse order to their appearance. The higher the frequency of the spectral component, the worse its frequency resolution, as expected for the CWT. Of course, this loss of frequency resolution is accompanied by an improvement 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 typical of the 24 cells we have examined.

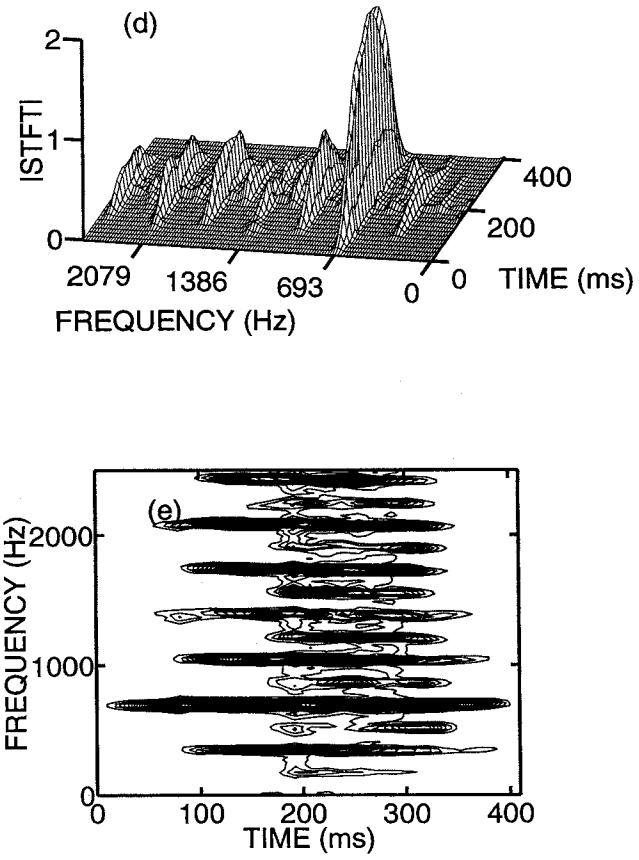


FIGURE 9. Continued.

For comparison, Fig. 5d and e shows 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 Fig. 5b and c at 693 Hz, which is the CF of the cell. Like the CWT, the STFT clearly shows spectral components at f_c, $2f_c$, $3f_c$, $4f_c$, and $5f_c$; a component at $6f_c$ is also just barely visible in the 3D plot of Fig. 5d.

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. 5a was calculated using the three other wavelet bases shown in Fig. 3c–h.

Figure 5f and g displays the modified CWT magnitude calculated using the low-Q Morlet wavelet basis displayed in Fig. 3c and d. The proportionality constant K, mapping scale to frequency, was 49. The spectral resolution in these figures is significantly degraded in comparison with both the high-Q Morlet wavelet modified CWT and the STFT shown earlier. The component at f_c is still apparent, but none of the harmonic components can be clearly resolved. This is because the relative bandwidth of this wavelet basis is a factor of 10 larger than that of the high-Q Morlet wavelet, so the frequency resolution at any given frequency is diminished by a factor of 10. At a frequency $f_c = 400$ Hz, for example, the frequency resolution of the high-Q Morlet wavelet is 28.6 Hz; that of the low-Q Morlet wavelet is 286 Hz. In the latter case, the spectral component at 400 Hz therefore still has a relatively large skirt at 800 Hz. Thus, the modified CWT magnitude at 800 Hz contains residual contribution from the 400 Hz component as well as the 800 Hz component itself. At higher frequencies, of course, the frequency resolution is worse: at 800 Hz it is 572 Hz, whereas at 1200 Hz it is 860 Hz. The improved time resolution accruing at higher frequencies is of little advantage, however, since the poor frequency resolution makes it hard to ascribe time events to a particular frequency component.

Figure 5h and i shows the modified CWT magnitude calculated using the Meyer wavelet basis shown in Fig. 3e and f. The proportionality constant K, mapping scale to frequency, was 313. The results are similar to those for the

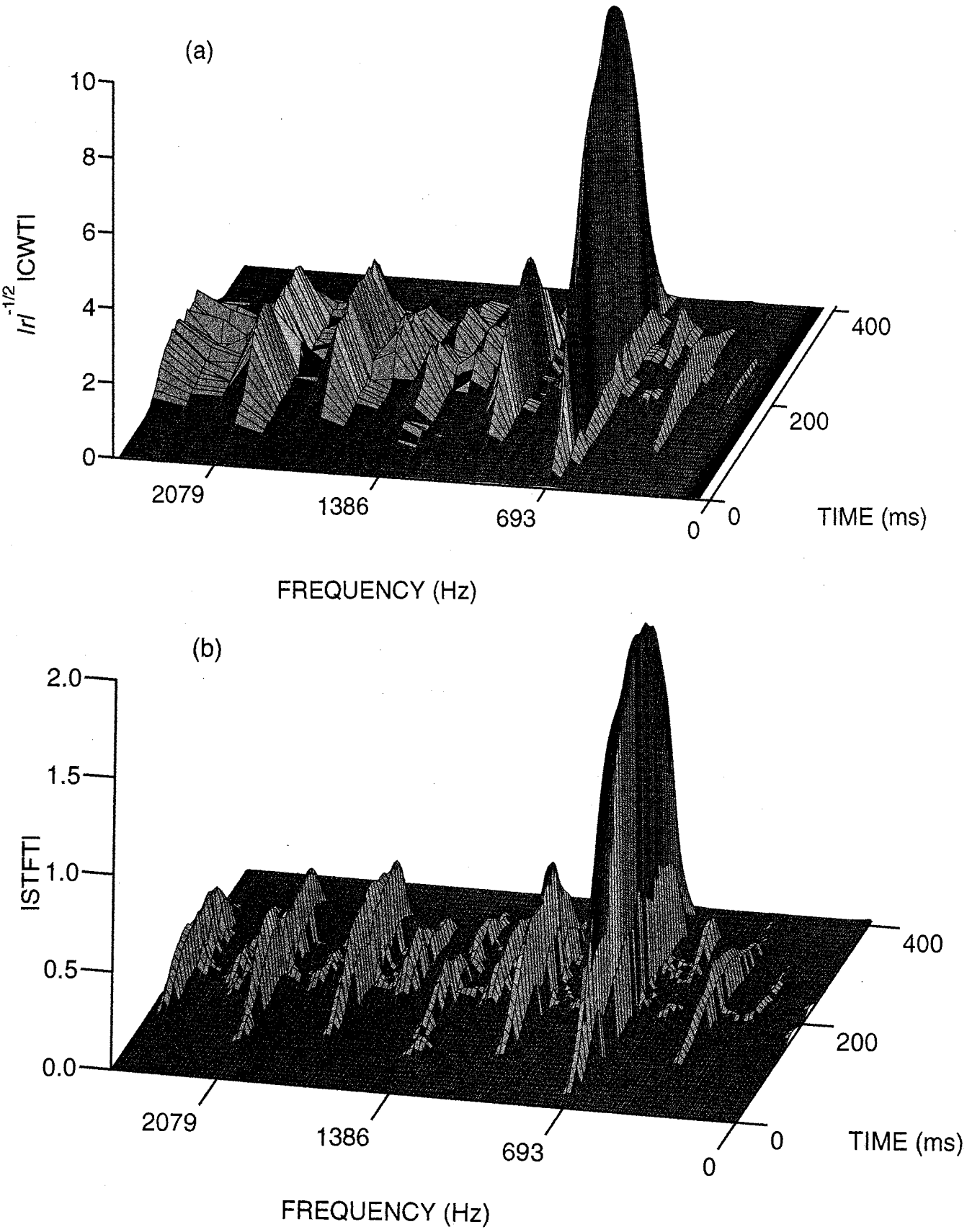


FIGURE 10. (a) Magnified view of the modified CWT magnitude shown in Fig. 9b with more time resolution points plotted. (b) Magnified view of the STFT magnitude shown in Fig. 9d with more time resolution points plotted. The relative magnitudes of the various spectral components, and their onset and offset times, should be compared.

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. 3e.

Finally, Fig. 5j and k shows the modified CWT magnitude calculated using the Daubechies 4-tap wavelet basis shown in Fig. 3g and h. The proportionality constant K, mapping scale to frequency, was 14. 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 analysis tool is the presence of multiple bandpass regions in the CFT of the wavelet basis, as is evident in Fig. 3h. 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 400 Hz appears in the CWT magnitude not only at 400 Hz, but also at 400/4=100 Hz and 400/8=50 Hz (the component at 400 Hz arises when the main lobe of the analysis wavelet is at 400 Hz, the component at 100 Hz appears when the second lobe of the analysis wavelet is at 400 Hz, and the component at 50 Hz appear when the third lobe of the analysis wavelet is at 400 Hz).

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 unsuit-

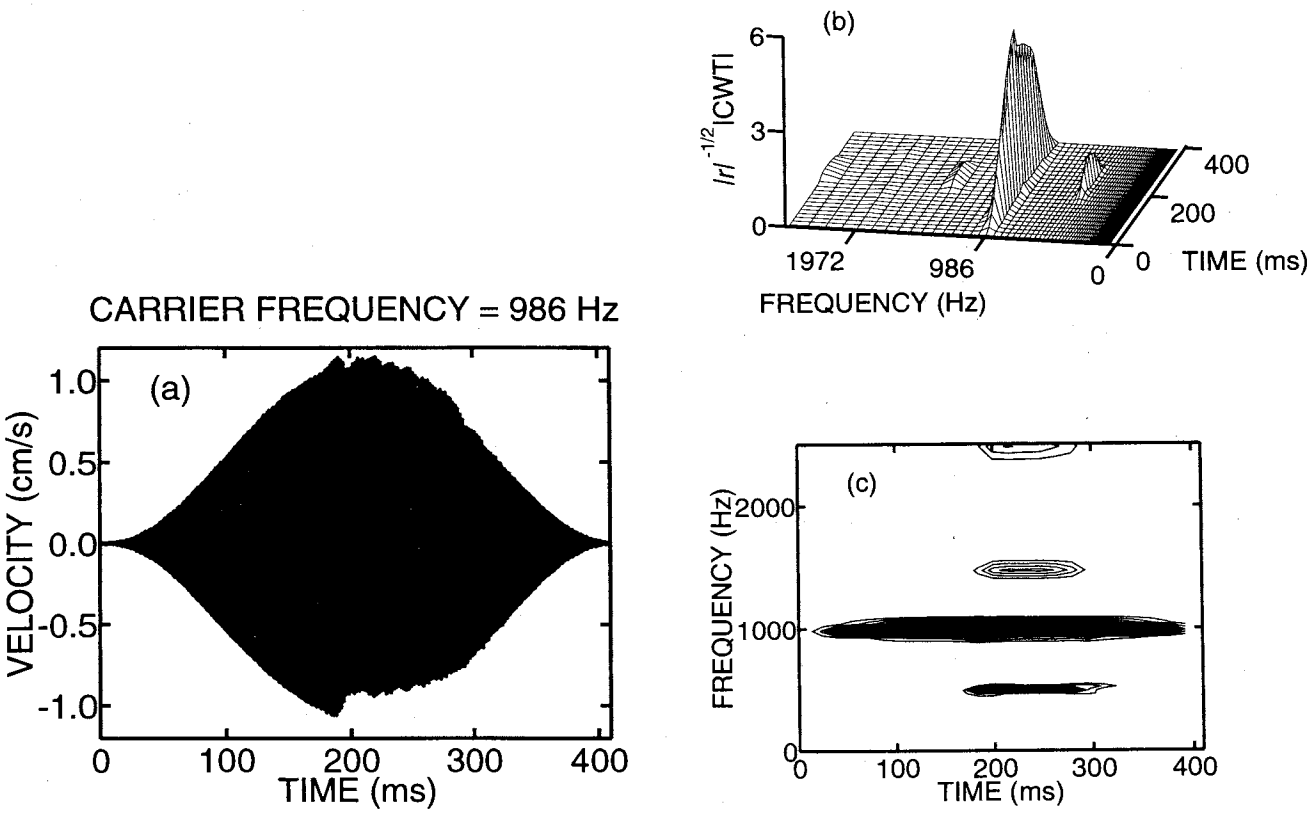


FIGURE 11. Velocity response of the same Hensen's cell as shown in Fig. 5–10, but now to an AM stimulus with carrier frequency f_c = 986 Hz (above CF), modulation frequency 2.44 Hz, and an 8-dB increase in sound pressure level in comparison with Fig. 7. The highest sound pressure level is now ≈132 dB:re 0.0002 dyne/cm^2. (a) Time waveform of the velocity response (in cm/s) of the cell. There is a dip in the response at t = 190 ms. (b) 3D plot of the modified CWT magnitude of the velocity response shown in Panel a. The high-Q Morlet wavelet basis was used. This plot shows spectral components at the carrier frequency f_c and at three half-harmonic frequencies ($f_c/2$, $3f_c/2$, and $5f_c/2$). The appearance of the half-harmonic components corresponds in time to the dip in the waveform at t = 190 ms. (c) Same modified CWT magnitude as shown in Panel b, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (d) 3D plot of the STFT magnitude of the velocity response shown in Panel a. Spectral components are present at the carrier frequency f_c, at the second harmonic $2f_c$, and at three half-harmonic frequencies ($f_c/2$, $3f_c/2$, and $5f_c/2$). (e) Same STFT magnitude as shown in Panel d, but now plotted in 2D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude.

able as analysis tools for the class of data examined here since they are octave-band in nature. In the data analysis that follows, therefore, we restrict ourselves to use of the high-Q Morlet wavelet CWT and the STFT.

Figure 6 shows the response of the same cell when an AM tone with a carrier frequency f_c = 693 Hz (at CF), and a peak sound pressure level of ≈130 dB:re 0.0002 dynes/cm^2, is applied to the ear. The velocity response of the cell over one modulation cycle is displayed in Fig. 6a. The average magnitude of the velocity is about a factor of 3 greater than that shown in Fig. 5a since the carrier frequency is now at the CF of the cell and the sound pressure level is ≈14 dB higher. (The frequency response of the acoustic transducer used in these experiments was not flat, so that the sound pressure level for a constant input voltage varies with frequency.)

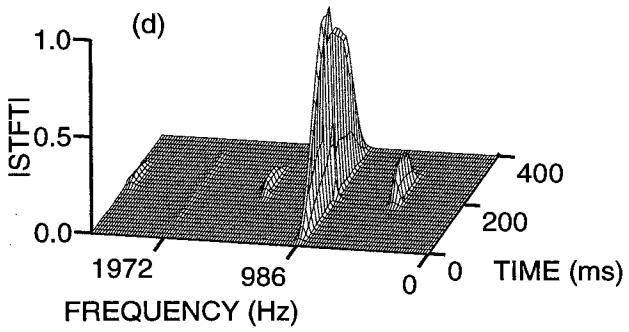
The response envelope does not follow the shape of the input envelope exactly; in particular it exhibits a chink at time t = 115 ms. The modified CWT magnitude of the velocity response (Fig. 6b and c), calculated using the high-Q Morlet wavelet, again shows the presence of multiple spectral components, but the behavior is more complex than that manifested at lower frequencies in Fig. 5b and c. At low values of the modulation envelope, the response is only at the carrier frequency. At slightly higher values, however, harmonic components emerge at $2f_c$ and $3f_c$. Shortly after the appearance of the harmonics, components are seen at three half-harmonic frequencies ($f_c/2$, $3f_c/2$, and $5f_c/2$); their appearance corresponds in time to the chink apparent in the time waveform presented in Fig. 6a. We have previously observed half-harmonic components in the vibration of outer hair cells in the third turn of the guinea pig cochlea (18–20,41). It is of interest to note that Davis *et al.* (13) and Dallos and Linnell (10,11) long ago observed the presence of half-harmonic components in guinea pig cochlear-microphonic potentials and in the sound field in front of the tympanic membrane when high-intensity pure-tone stimuli were applied to the ear canal. Dallos (12) also observed the presence of odd-fractional harmonics (which they determined to be rare in guinea pigs).

Figure 6d and e shows the STFT magnitude of the velocity response. Again, the STFT and modified CWT magnitudes display similar information. In this particular case, however, the STFT was able to discern a component at $7f_c/2$, which the modified CWT did not show.

Figure 7 shows the response of this cell when an AM tone with a carrier frequency f_c = 986 Hz (above CF), and a peak sound pressure level of ≈124 dB:re 0.0002 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. 7a. The response follows the shape of the input envelope with practically no distortion, and has about the same magnitude as that shown in Fig. 5a.

Figure 7b and c shows the modified CWT magnitude of the velocity response calculated using the high-Q Morlet wavelet. The modified CWT shows components at f_c and $2f_c$. The envelope of the component at the carrier frequency closely follows the envelope of the input stimulus.

The STFT magnitude of the velocity response (Fig. 7d and e), on the other hand, shows the presence of a dominant component at f_c and a small component at $2f_c$.



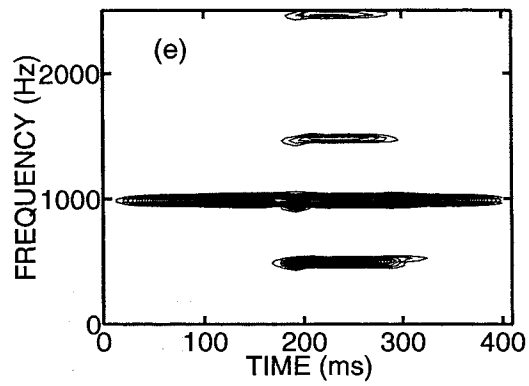


FIGURE 11. Continued.

Responses to AM Stimuli of Very High Intensity

Our principal interest here is the investigation of routes to chaos in the inner ear. Since these are more apparent at higher stimulus levels, we increased the peak sound pressure level attained by the AM stimulus by 8 dB at all carrier frequencies and repeated the set of measurements.

Figure 8 shows the response of the same cell as described in the previous section, when an AM tone with a carrier frequency f_c = 400 Hz (below CF) is applied to the preparation. The peak sound pressure level of the stimulus is now ≈124 dB:re 0.0002 dyne/cm^2. The velocity response of the cell over one modulation cycle is displayed

in Fig. 8a. Its magnitude is approximately 2.5 times greater than the corresponding velocity-time curve shown in Fig. 5a, as would be expected for a linear system under the influence of an 8-dB increase in sound pressure level. However, the system is anything but linear. The response envelope does not follow the input envelope. Though there is less asymmetry in the time waveform than evident in Fig. 5a, the CWT and STFT reveal a substantially increased level of nonlinearity.

The modified CWT magnitude of the velocity response (Fig. 8b and c), calculated using the high-Q Morlet wavelet, shows the presence of five spectral components: at f_c, $2f_c$, $3f_c$, $4f_c$, and $5f_c$. Substantial dips in the components at $2f_c$ and $4f_c$ are visible in the central region of the modulation envelope, where the sound pressure level is greatest. Unlike the corresponding figures at a lower intensity (Fig. 5b and c), the third harmonic and fifth harmonic components are stronger than their neighboring even harmonic components. We have previously reported this tendency of odd-harmonic components to dominate at the highest sound pressure levels (18,20,40). The STFT magnitude of the velocity response (Fig. 8d and e) shows the same five components as seen in the modified CWT, plus a component at $6f_c$ that is not clearly visible in the latter.

Figure 9 shows the response of this same cell to an AM tone with a carrier frequency f_c = 693 Hz (at CF) and a peak sound pressure level ≈138 dB:re 0.0002 dyne/cm^2. The velocity response of this cell over one modulation cycle is displayed in Fig. 9a. The waveform is highly irregular. Its magnitude is again approximately 2.5 times greater than that observed in Fig. 6a, and exhibits a chink at about the same value of velocity (which now occurs earlier). This illustrates the repeatability of the nonlinear response and its dependence on level. The modified

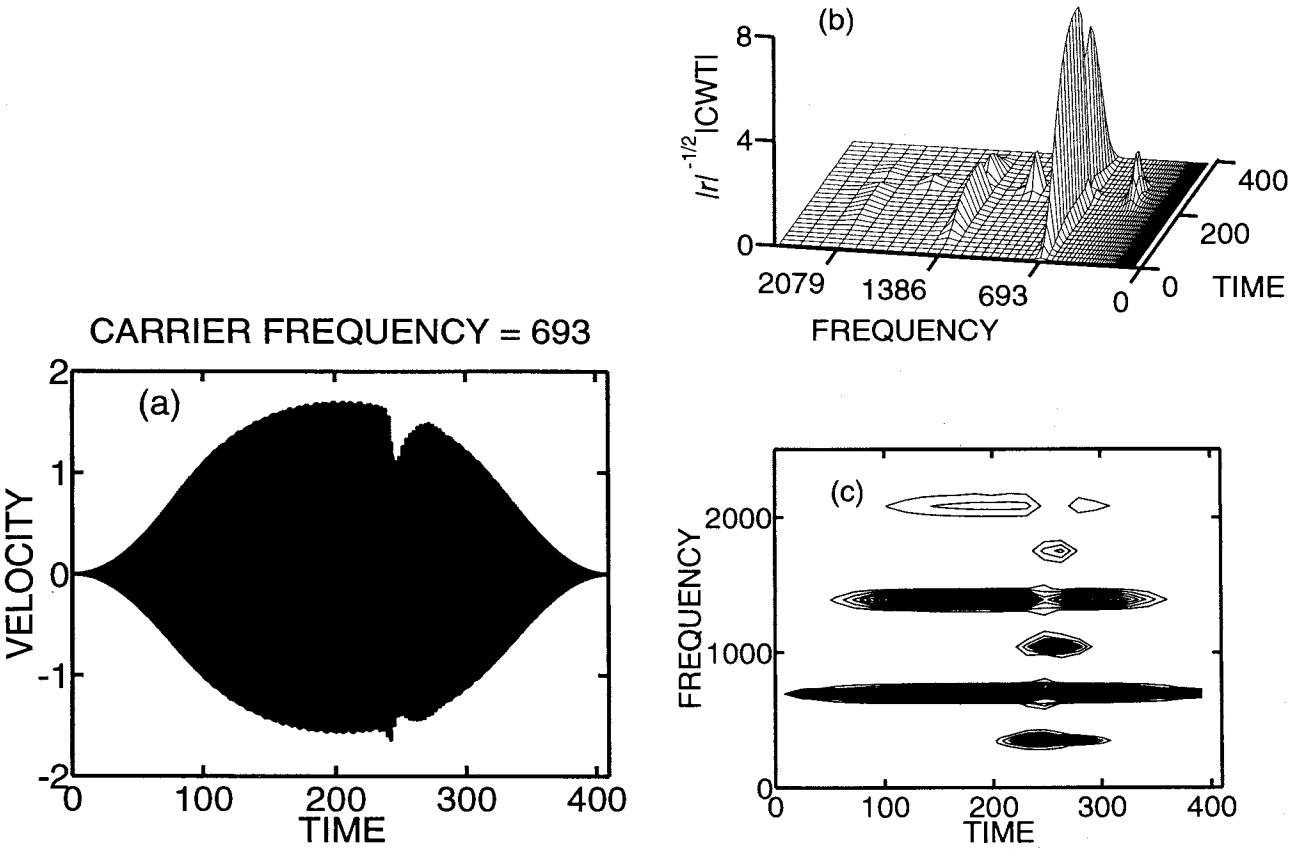


FIGURE 12. Velocity response of a model system based on the Duffing oscillator with negative stiffness (as described in text) to an AM stimulus with carrier frequency f_c = 693 and modulation frequency 2.44. The characteristic frequency of the model system has been set at 693. (a) Time waveform of the velocity response of the model system. There is a dip in the response at t = 240. (b) 3D plot of the modified CWT magnitude of the velocity response shown in Panel a. The high-Q Morlet wavelet basis was used. This plot shows spectral components at the carrier frequency f_c, at two harmonic frequencies ($2f_c$ and $3f_c$), and at three half-harmonic frequencies ($f_c/2$, $3f_c/2$, and $5f_c/2$). The appearance of the half-harmonic components corresponds in time to the dip apparent in the time waveform at t = 240. (c) Same modified CWT magnitude as shown in Panel b, but now plotted in 2D format with 80 equally spaced (in modified CWT magnitude) contour lines joining points of constant magnitude. (d) 3D plot of the STFT magnitude of the velocity response shown in Panel a. Spectral components are present at the carrier frequency f_c, at two harmonic frequencies ($2f_c$ and $3f_c$), and at three half-harmonic frequencies ($f_c/2$, $3f_c/2$, and $5f_c/2$). (e) Same STFT magnitude as shown in Panel d, but now plotted in 2D format with 80 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude.

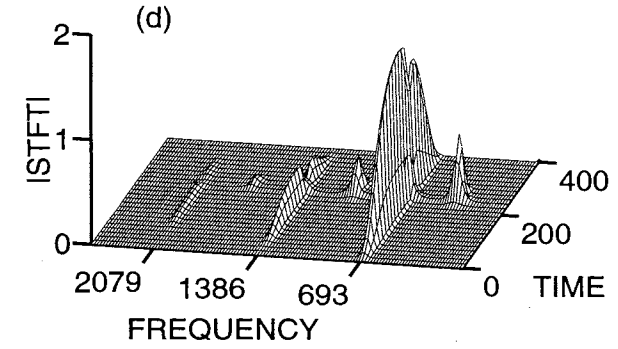
CWT magnitude of the velocity response (Fig. 9b and c) shows a large number of components at different scales. As illustrated in Fig. 6a at a sound pressure level that is 8 dB lower, clearly defined harmonic components (at f_c, $2f_c$, and $3f_c$), and half-harmonic components (at $f_c/2$, $3f_c/2$, $5f_c/2$, and $7f_c/2$), are visible. In addition, however, a large number of components at quarter-harmonic frequencies ($f_c/4$, $3f_c/4$, $5f_c/4$, $7f_c/4$, and $9f_c/4$) have emerged; these were not present at the lower stimulus level. The improved frequency resolution of the CWT at low frequencies is especially useful in clearly resolving the components at $f_c/4$, $f_c/2$, $3f_c/4$, and f_c. We clearly observe period-doubling behavior, which often serves as a prelude to deterministic chaos as discussed further in the section dealing with the negative stiffness Duffing oscillator.

The STFT magnitude of the velocity response illustrated in Fig. 9d and e shows similar results. However, the superior frequency resolution of the STFT at high frequencies is beneficial for identifying further quarter-harmonic components at $11f_c/4$ and $13f_c/4$.

The relative magnitudes of the various spectral components for this stimulus level and frequency are most clearly discerned in Fig. 10a and b. These are magnified versions of Fig. 9b and d, respectively, illustrated in color with more time resolution points plotted. It is clear from this figure that the half-harmonic components are similar in magnitude to the third-harmonic, and larger than the second-harmonic. Furthermore, the quarter-harmonic components are seen to emerge at about $t = 180$ ms, just prior to the envelope reaching its peak. They are asymmetrically skewed toward the falling portion of the envelope, and persist to about $t = 320$ ms.

Finally, in Fig. 11, we show the response of the same cell, but now to an AM tone with a carrier frequency $f_c = 986$ Hz (above CF) and a peak sound pressure level of $\approx$132 dB:re 0.0002 dyne/cm^2. The velocity response of this cell over one modulation cycle is displayed in Fig. 11a. There is a notch in the velocity response at time $t = 190$ ms. The modified CWT magnitude of the velocity response (Fig. 11b and c) shows a similar notch in the carrier frequency component at time $t = 190$ ms. This dip is accompanied by the sudden appearance of three half-harmonic components (at $f_c/2$, $3f_c/2$, and $5f_c/2$). The component at the second harmonic is barely visible in the modified CWT plots.

Figure 11d and e show the STFT of the same velocity response. As with the modified CWT, half-harmonic components at $f_c/2$, $3f_c/2$, and $5f_c/2$ are visible. In Fig. 11d, the second harmonic component is also visible, albeit at a very low level. This above-CF behavior is in marked contrast to that obtained at a stimulus level 8 dB lower, where the dynamics are nearly linear as is evident in Fig. 7.



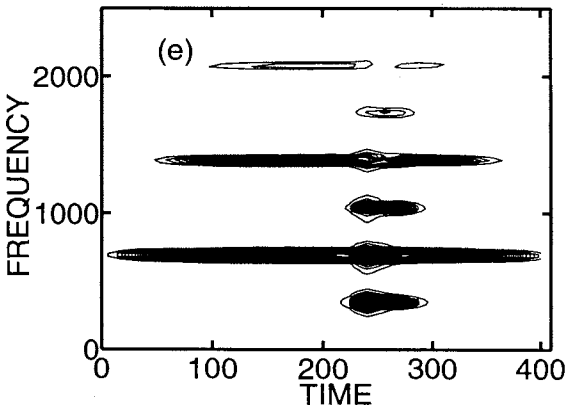


FIGURE 12. Continued.

The Negative-Stiffness Duffing Oscillator as a Model of Cochlear Dynamics

In an attempt to characterize the nonlinear mechanisms responsible for the unusual behavior detailed above and elsewhere (18–20,27,29,37–41), we can consider analogies to other well-studied nonlinear systems. Keilson *et al.* (25,26,36) have already analyzed the responses of a class of bilinear oscillators and compared the results with experimental findings (37).

Here, we consider a Duffing oscillator model with negative stiffness (1,14,15,18,20,31), defined by the equation:

$$\ddot{x}(t) + \alpha\dot{x}(t) - \beta x(t) + \gamma x^3(t) = F(t), \qquad (19)$$

where $x(t)$ is position; t is time; α, β, and γ are constants; $F(t)$ is an arbitrary forcing function of time; and the overdot indicates differentiation with respect to time. This equation represents an oscillatory system with two symmetric potential wells. Since the potential energy is a quartic function of position, this equation models, for example, the motion of a ball, subject to an arbitrary driving

force, oscillating in two valleys separated by an energy barrier.

Small motions within either of these wells have a clearly defined characteristic frequency, the value of which is determined by suitable choice of the parameters α, β, and γ. Motion between the wells gives rise to another, lower, characteristic frequency also determined by the parameters and by the signal level. For comparison with the cellular velocity data considered earlier, we choose $F(t)$ to be a slowly modulated AM signal.

Figure 12a shows the velocity response $\dot{x}$ for such a system driven by a unity modulation-depth AM signal with carrier frequency 693 and modulation frequency 2.44. The driving function was of the form $F(t) = F_0[1 - \cos(2\pi f_m t)]\sin(2\pi f_c t)$, where f_c and f_m represent the carrier and modulation frequencies of the AM signal, respectively, and F_0 is a constant. The characteristic frequency of the (small-amplitude, undamped) velocity response for the system was set at 693 (as suggested by the experimental tuning curve of Fig. 4), by choosing $\alpha = 2000$, $\beta = 9.48 \times 10^6$, and $\gamma = 4.74 \times 10^6$. F_0 was set equal to 5.0×10^6.

A sudden irregularity appears in the time waveform at time $t = 240$. Indeed the dip in the model response bears some qualitative similarity to the corresponding dips shown in Figs. 6a, 9a, and 11a. The modified CWT magnitude of this velocity response (Fig. 12b and c) reveals the presence of multiple spectral components. Near time $t = 240$, the model response jumps rather suddenly to encompass half-harmonic ($f_c/2$, $3f_c/2$, and $5f_c/2$), as well as harmonic frequencies ($2f_c$ and $3f_c$). This is a consequence of the subharmonic cascade (or period-doubling [3]) route to chaos that the Duffing oscillator system exhibits. As the magnitude of the driving force increases further, the spectrum broadens and the system is driven toward chaotic behavior. Similar features are also present in the cellular velocity data at very large acoustic intensities (18,20,37–39). However, the cellular data do not generally behave so abruptly, nor do they exhibit quite as broad a range of frequencies as the model, at least for the model parameters illustrated.

The behavior of the Duffing model was studied over a range of carrier frequencies and driving intensities. For carrier frequencies well below the CF, multiple harmonic components are generated, similar to those seen in Figs. 5 and 8. However, unlike the results in Fig. 8, the magnitudes of the model's harmonic components decrease monotonically with harmonic order. For carrier frequencies well above the CF, the response is mostly at the carrier frequency as in Figs. 7 and 11. The model response shown in Fig. 12 perhaps most closely resembles the data shown in Fig. 11.

We believe that the negative-stiffness Duffing oscillator is too simple and idealized to be used for modeling cellular dynamics in the cochlea. Nevertheless some features of its behavior, such as the presence of harmonic and half-harmonic spectral components, and spectral broadening at high acoustic intensities, suggest that a reasonable model of cochlear dynamics will contain some of the features of the Duffing system. However, many other nonlinear systems with quite different characteristics also share these features. We are continuing our investigation of nonlinearities in the cochlea with the goal of developing more appropriate nonlinear dynamical models.

DISCUSSION

It is apparent from the foregoing discussion and the data analyses presented in Figs. 5–11 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 (34,43,44), 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 multi-resolution framework onto which a signal could be efficiently decomposed (44), 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. 12, *i.e.*:

$$\frac{\sqrt{\alpha}}{c} = \frac{\Delta f}{2\sqrt{2}\, f}. \tag{20}$$

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 $L = 256$ and $T_x = 1/5000$, which were used in calculating the STFTs shown here, $\beta = 12207$ and the frequency resolution of the STFT was $\Delta f = 49.7$ 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. For example, Fig. 9b (or 10a) and c demonstrates a case where better frequency resolution is needed at low frequencies to clearly distinguish components at $f_c/4$, $f_c/2$, $3f_c/4$, and f_c, and to search for further period doublings at $f_c/8, f_c/16$, *etc*. On the other hand, the STFT is best suited to situations where the spectral components are linearly spaced in frequency. This is illustrated by comparing the modified CWT and STFT magnitudes in Fig. 10a and b at high frequencies. Clearly, the STFT is superior in terms of frequency resolution. Time resolution is important when abrupt time changes are contained in the signal, which does not seem to be the case for our measurements.

Thus, the CWT is the preferred tool when the analysis requires good frequency resolution at low frequencies together 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 (22)]. Indeed, other wavelet-based analysis techniques exist. Of particular interest are wavelet packets (9,21), which provide a more general tiling of the time-frequency plane than 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 (21).

CONCLUSION

The time course of cellular vibration in the guinea pig inner ear has been successfully studied using both the STFT and an appropriately chosen CWT. Both analysis techniques reveal rich and varied nonlinear dynamics inherent in the velocity responses of individual outer hair cells and Hensen's cells, in response to AM acoustical signals. Nonlinear effects appear at all stimulus frequencies (20), but have been found to be most pronounced at the highest stimulus levels near CF, where harmonic, half-harmonic, and quarter-harmonic components are all present. The generation of these components is consistent with the behavior of a dynamical system entering chaos via a period-doubling route. We have compared the behavior of a negative-stiffness Duffing oscillator with our data. An increase in the bifurcation parameter (input intensity) for the pure-tone-driven Duffing oscillator does indeed cause first the appearance of harmonic components, then half-harmonic components, and then quarter-harmonic (and possibly further inverse power-of-two harmonic) components, prior to the onset of full-blown chaos. The appearance of the half-harmonic components in the Duffing model is often quite sudden in time. These features are also characteristic of the experimental data we have presented here, strengthening the likelihood that cellular dynamics in the cochlea can be driven into a chaotic regime under sufficiently strong (and probably nonphysiological) stimulus conditions.

REFERENCES

1. Bacri, J.-C., U. d'Ortona, and D. Salin. Magnetic-fluid oscillator: observation of nonlinear period doubling. *Phys. Rev. Lett.* 67:50–53, 1991.
2. von Békésy, G. *Experiments in Hearing*. New York: McGraw-Hill, 1960; Huntington, NY: Krieger, 1980, 745 pp.
3. Bergé, P., Y. Pomeau, and C. Vidal. *Order Within Chaos*. New York: Wiley-Interscience, 1986, 329 pp.
4. Brownell, W. E., C. R. Bader, D. Bertrand, and Y. de Ribaupierre. Evoked mechanical responses of isolated cochlear outer hair cells. *Science* 227:194–196, 1985.
5. Brundin, L., Å. Flock, and B. Canlon. Sound-induced motility of isolated cochlear outer hair cells is frequency-specific. *Nature* 342:814–816, 1989.
6. Brundin, L., Å. Flock, S. M. Khanna, and M. Ulfendahl. Frequency-specific position shift in the guinea pig organ of Corti. *Neurosci. Lett.* 128: 77–80, 1991.
7. Brundin, L., Å. Flock, S. M. Khanna, and M. Ulfendahl. The tuned displacement response of the hearing organ is generated by the outer hair cells. *Neuroscience* 49:607–616, 1992.
8. Canlon, B., L. Brundin, and Å. Flock. Acoustic stimulation causes tonotopic alterations in the length of outer hair cells from the guinea pig hearing organ. *Proc. Natl. Acad. Sci. (U.S.A.)* 85:7033–7035, 1988.
9. Chui, C. K. *Wavelet Analysis and its Applications*. Boston: Academic Press, 1992, 264 pp.
10. Dallos, P. J., and C. O. Linnell. Subharmonic components

in cochlear-microphonic potentials. *J. Acoust. Soc. Am.* 40: 4–11, 1966.
11. Dallos, P. J., and C. O. Linnell. Even-order subharmonics in the peripheral auditory system. *J. Acoust. Soc. Am.* 40: 561–564, 1966.
12. Dallos, P. J. On the generation of odd-fractional harmonics. *J. Acoust. Soc. Am.* 40:1381–1391, 1966.
13. Davis, H., B. E. Gernandt, and W. P. Covell. Aural microphonics in the cochlea of the guinea pig. *J. Acoust. Soc. Am.* 21:502–510, 1949.
14. De Souza-Machado, S., R. W. Rollins, D. T. Jacobs, and J. L. Hartman. Studying chaotic systems using microcomputer simulations and Lyapunov exponents. *Am. J. Phys.* 58:321–329, 1990.
15. Drazin, P. G. *Nonlinear Systems*. Cambridge, United Kingdom: Cambridge University Press, 1992, 317 pp.
16. Gabor, D. Theory of communication. *J. IEE.* 93:429–457, 1946.
17. Grossmann, A., R. Kronland-Martinet, and J. Morlet. Reading and understanding continuous wavelet transforms. In: *Wavelets: Time-Frequency Methods and Phase Space*, edited by J. M. Combes, A. Grossmann, and Ph. Tchamitchian. New York: Springer-Verlag, 1989/1990, pp. 2–20.
18. Heneghan, C., M. C. Teich, S. M. Khanna, and M. Ulfendahl. Nonlinear dynamical motion of cellular structures in the cochlea. *Proc. SPIE.* 2036:183–197, 1993.
19. Heneghan, C., 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 Proc. Int. Symp. Time-Frequency Time-Scale Anal.* October:25–28, 1994.
20. Heneghan, C., 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. Sign. Proc.* 42:3335–3352, 1994.
21. Herley, C., J. Kovačević, K. Ramchandran, and M. Vetterli. Tilings of the time-frequency plane: construction of arbitrary orthogonal bases and fast tiling algorithms. *IEEE Trans. Sign. Proc.* 41:3341–3359, 1993.
22. Hlawatsch, F., and G. F. Boudreaux-Bartels. Linear and quadratic time-frequency signal representations. *IEEE Sign. Proc. Mag.* 9(2):21–67, 1992.
23. International Team for Ear Research (ITER). Cellular vibration and motility in the organ of Corti. *Acta Otolaryngol. (Stockholm)* 467(Suppl.):1–279, 1989.
24. Jones, D. L., and R. G. Baraniuk. Efficient approximation of the continuous wavelet transform. *Electron. Lett.* 27: 748–750, 1991.
25. Keilson, S. E., M. C. Teich, and S. M. Khanna. Models of nonlinear vibration. I. Oscillator with bilinear resistance. *Acta Otolaryngol. (Stockholm)* 467(Suppl.):241–248, 1989.
26. Keilson, S. E., M. C. Teich, and S. M. Khanna. Models of nonlinear vibration. III. Oscillator with bilinear mass. *Acta Otolaryngol. (Stockholm)* 467(Suppl.):257–264, 1989.
27. Keilson, S. E., M. C. Teich, S. M. Khanna, and M. Ulfendahl. The effect of sinusoidal stimuli on the spontaneous cellular vibrations in the guinea-pig cochlea. In: *Abstracts of the Sixteenth Midwinter Research Meeting of the Association for Research in Otolaryngology*, February 7–11, 1993, ISSN 0742-3152, Abstract No. 336, p. 84.
28. Keilson, S. E., S. M. Khanna, M. Ulfendahl, and M. C. Teich. Spontaneous cellular vibrations in the guinea-pig cochlea. *Acta Otolaryngol. (Stockholm)* 113:591–597, 1993.
29. Khanna, S. M., M. Ulfendahl, and Å. Flock. Waveforms and spectra of cellular vibrations in the organ of Corti. *Acta Otolaryngol. (Stockholm)* Supplement 467(Suppl.):189–193, 1989.
30. Khanna, S. M., S. E. Keilson, M. Ulfendahl, and M. C. Teich. Spontaneous cellular vibrations in the guinea-pig temporal-bone preparation. *Br. J. Audiol.* 27:79–83, 1993.
31. Nayfeh, A. H., and D. T. Mook. *Nonlinear Oscillations*. New York: Wiley-Interscience, 1979, 704 pp.
32. Oppenheim, A. V., and R. W. Schafer. *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989, 879 pp.
33. Papoulis, A. *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984, 576 pp.
34. Rioul, O., and M. Vetterli. Wavelets and signal processing. *IEEE Sign. Proc. Mag.* 8(4):14–38, 1991.
35. Saleh, B. E. A., and M. C. Teich. *Fundamentals of Photonics*. New York: Wiley, 1991, 966 pp.
36. Teich, M. C., S. E. Keilson, and S. M. Khanna. Models of nonlinear vibration. II. Oscillator with bilinear stiffness. *Acta Otolaryngol. (Stockholm)* 467(Suppl.):249–256, 1989.
37. Teich, M. C., S. M. Khanna, and S. E. Keilson. Nonlinear dynamics of cellular vibrations in the organ of Corti. *Acta Otolaryngol. (Stockholm)* Supplement 467(Suppl.):265–279, 1989.
38. Teich, M. C., S. E. Keilson, S. M. Khanna, L. Brundin, M. Ulfendahl, and Å. Flock. Chaos in the cochlea. In: *Abstracts of the Fourteenth Midwinter Research Meeting of the Association for Research in Otolaryngology*, February 3–7, 1991, ISSN 0742-3152, Abstract No. 155, p. 50.
39. Teich, M. C., 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 of the Fifteenth Midwinter Research Meeting of the Association for Research in Otolaryngology*, February 2–6, 1992, ISSN 0742-3152, Abstract No. 41, p. 17.
40. Teich, M. C., 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*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. van Netten. Singapore: World Scientific, 1993, pp. 272–279.
41. Teich, M. C., 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. Sixteenth Ann. Int. Conf. IEEE Eng. in Med. and Biol. Soc.*, November 1–6, 1994.
42. Ulfendahl, M., Å. Flock, and S. M. Khanna. A temporal bone preparation for the study of cochlear micromechanics at the cellular level. *Hearing Res.* 40:55–64, 1989.
43. Vetterli, M., and C. Herley. Wavelets and filter banks: theory and design. *IEEE Trans. Sign. Proc.* 40:2207–2232, 1992.
44. Vetterli, M., and J. Kovačević. *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995, 279 pp.