

# Investigating the Nonlinear Dynamics of Cellular Motion in the Inner Ear Using the Short-Time Fourier and Continuous Wavelet Transforms

Conor Heneghan, *Student Member, IEEE*, Shyam M. Khanna, *Senior Member, IEEE*, Åke Flock, Mats Ulfendahl, Lou Brundin, and Malvin C. Teich, *Fellow, IEEE*

**Abstract**—The short-time Fourier transform (STFT) and the continuous wavelet transform (CWT) are used to analyze the time course of cellular motion in the inner ear. The velocity responses of individual outer hair cells and Hensen's cells to sinusoidal and amplitude modulated (AM) acoustical signals applied at the ear canal display characteristics typical of nonlinear systems, including the generation of harmonic and half-harmonic components. The STFT proves to be valuable for following the time course of the frequency components generated using sinusoidal and AM input signals. The CWT is also useful for analyzing these signals; however, it is generally not as effective as the STFT when octave-band-based CWT's are used. For the transient response, the spectrogram (which is the squared magnitude of the STFT) and the octave-band-based scalogram (which is the squared magnitude of the CWT) prove equally valuable, and we have used both to study the responses of these cells to step-onset tones of different frequencies. Such analyses reveal information about the preferred vibration frequencies of cells in the inner ear and are useful for deciding among alternative mathematical models of nonlinear cellular dynamics. A modified Duffing oscillator model yields results that bear some similarity to the data.

## I. INTRODUCTION

**I**N 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  $4\frac{1}{2}$  turns in the guinea pig

Manuscript received September 6, 1992; revised August 30, 1993. 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. The associate editor coordinating the review of this paper and approving it for publication was Dr. Ahmed Tewfik.

C. Heneghan is with the Department of Electrical Engineering, Columbia University, New York, NY 10027 USA.

S. M. Khanna is with the Department of Otolaryngology, Columbia College of Physicians and Surgeons, New York, NY 10032 USA.

Å. Flock, M. Ulfendahl, and L. Brundin are with the Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden.

M. C. Teich is with the Departments of Electrical Engineering and Applied Physics, Columbia University, New York, NY 10027 USA.

IEEE Log Number 9406036.

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, although the width and stiffness of the basilar membrane and the dimensions of most of the cells change from base to apex [1], [2].

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 [2] 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 [3]. 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 measuring and stimulus-generation techniques have been previously described [2].

This measurement technique elucidates 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 two 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 band-pass 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 [2]. In the third turn of the guinea pig cochlea, the cells respond maximally to frequencies in the range 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 a motile response is observed is far narrower than the ac passband [2], [4]–[8].

A suitable method for studying the motile response is to use amplitude modulated (AM) acoustic waves with a sinusoidal carrier and low-frequency modulation. This provides an opportunity for studying the change in the cell length over a broad range of carrier levels as the envelope increases and decreases. The use of an AM format is called for 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, special analysis techniques are required to examine the response. This paper describes the application of time-frequency and time-scale representation techniques for the analysis of cellular velocity data such as that described above. The relative advantages of the two techniques are compared for several data sets. Two types of input stimuli were used: i) an AM acoustic signal with a modulation depth of unity and a modulation frequency fixed at 2.44 Hz, the carrier frequency ranging from 24 to 1800 Hz and the total duration of each collected data set fixed at 0.4096 s (representing 2048 samples at 200  $\mu$ s intervals) and ii) a sinusoidal acoustic signal, with a frequency that ranged from 24 to 1489 Hz, turned on at its positive zero crossing for a duration of 41 ms (1024 samples at 40  $\mu$ s intervals). These stimuli are well suited for investigating nonlinear dynamical behavior.

To provide a complete analysis of these signals, we have examined three representations (magnitude, squared magnitude, and phase) of both the short-time Fourier transform (STFT) and the octave-band-based continuous wavelet transform (CWT). We have compared the transforms for the different representations on selected data sets and found that the STFT magnitude was most effective in discriminating the various frequency components elicited by AM stimuli. For transient responses, the spectrogram (which is the squared magnitude of the STFT) and the scalogram (which is the squared magnitude of the CWT) both proved equally effective. Phase maps of the STFT and CWT were also examined over a small section of one of the AM responses. The resulting CWT phase map proved more readable than the magnitude or squared magnitude of the CWT. However, the STFT phase map provided information that was essentially equivalent to that offered by the magnitude or squared magnitude of the STFT. In the Discussion (Section IV), we consider some of the differences between the two transforms and identify cases where one transform might be preferred over the other.

We clearly ascertain that the nonlinear dynamical characteristics of cellular motion in the cochlea [9] can be elucidated by using both STFT- and CWT-based analyses. The character of the cell's nonlinear response is found to depend significantly

on the carrier frequency of the applied AM signal, relative to the innate tuning characteristics of the cell. We show that a simple nonlinear system (the negative-stiffness Duffing oscillator) exhibits some of the features seen in our experimental data, although it is clearly too idealized to serve as a model for cellular dynamics in the cochlea. We have previously reported preliminary results pertaining to the use of the STFT in the analysis of cellular velocity data [10]–[12], and we have also presented a preliminary version of our octave-band-based CWT analysis [13].

## II. ANALYSIS TECHNIQUES

An AM tone with a low modulation frequency is a wave of slowly varying amplitude. The AM stimulus is useful in investigating nonlinear dynamics because it permits us to observe, in a single experiment, the cell's response for a continuous and large range of acoustic intensities. The AM tone is, however, a nonstationary stimulus over the time scale of the experimental recording; a suitable representation of the response in the spectral domain must therefore be used. Although it is useful in highlighting the broad spectral features of the response, the averaged periodogram for the entire signal duration is an incomplete representation for such a signal. Similarly, the response to the step turn-on of a sinusoidal tone is clearly nonstationary.

Accordingly, we have made use of various time-frequency representations (TFR's) and time-scale representations to capture the time-varying spectral character of the cellular response. A commonly used TFR is the spectrogram (SPEC) [14], [15], which is defined as

$$\text{SPEC}_x^g(t, f) = |\text{STFT}_x^g(t, f)|^2, \quad (1)$$

with the STFT defined as

$$\text{STFT}_x^g(t, f) = \int_{-\infty}^{\infty} x(u)g^*(u-t)\exp(-j2\pi fu)du \quad (2)$$

where

- $x(t)$  time waveform being analyzed
- $g(t)$  window function in time
- $t$  time variable
- $f$  frequency variable
- $*$  complex conjugation.

A Gaussian window [ $g(t) = \exp(-t^2)$ ] provides a reasonable choice for the window function since it minimizes the time-frequency uncertainty product [14]. Specifically, a sampled version of the STFT, which is often referred to as the discrete STFT, was calculated using a summation approximation of (2):

$$\text{STFT}_x^g[n, k] = \sum_{m=0}^{L-1} x[n+m]g^*[m] \times \exp\left(\frac{-j2\pi mk}{N}\right), \quad 0 \leq k \leq N-1 \quad (3)$$

where

- $n$  discrete time index
- $k$  discrete frequency index
- $g[m]$  sampled Gaussian window with  $t$  ranging from -2 to 2:

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

The formulation presented in (3) reminds us that the discrete STFT is simply a sequence of discrete Fourier transforms 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  $L = 128$  samples was appropriate. This corresponds to 25.6 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 32 time samples for successive evaluations of the STFT provided a sufficiently detailed picture for our purposes.

We present the STFT magnitude and the spectrogram in two visual formats. The first is a 3-D representation, which is often referred to as a 3-D spectral plot. In this format, time and frequency form the bottom plane and the STFT magnitude or spectrogram is represented on a linear axis in the third dimension. The second format provides 2-D contour plots on which contours of equal STFT magnitude or spectrogram are traced on a time-frequency plane. The choice of whether to examine the STFT magnitude or the spectrogram is signal dependent. The magnitude of the STFT allows both strong and weak components to be effectively shown on the same plot; this is generally desirable for our class of signals. For other types of signals, however, plotting the spectrogram allows a weak signal component to be readily distinguished from the noise floor.

The CWT provides an alternative representation with the flexibility of trading time and frequency resolution. This flexibility can provide a more useful representation for certain signals, e.g., those with a mix of short-term impulsive events and long-term harmonic components [16]. We analyzed our data using the basic wavelet  $h(t) = \exp(jct)\exp(-t^2/2)$ , which is often referred to as the Morlet wavelet [17]. Although strictly speaking this is not an admissible wavelet basis (since  $H(0) \neq 0$ , where  $H(f)$  denotes the continuous-time Fourier transform of  $h(t)$ ), for a suitable choice of the constant  $c$  (e.g.,  $c \geq 5.0$ ), this wavelet basis is practically admissible and progressive [17], and its implementation is straightforward since an analytic expression for the wavelet basis is at hand. The CWT for continuous variables is defined as

$$\text{CWT}_x^h(r, \tau) = \frac{1}{\sqrt{|r|}} \int_{-\infty}^{\infty} x(t)h^* \left( \frac{t-\tau}{r} \right) dt \quad (5)$$

where

- $x(t)$  time waveform being analyzed
- $h(t)$  wavelet basis function
- $\tau$  time variable
- $r$  scale variable
- \* complex conjugation.

Scale is inversely related to the frequency variable used in (2).

We approximated a sampled version of the CWT using the form

$$\begin{aligned} \text{CWT}_x^h[a, b] &= \sum_{m=-\infty}^{+\infty} x[m]h_a^*[b-m] \\ &= \sum_{m=-\infty}^{+\infty} x[b-m]h_a^*[m], \end{aligned} \quad (6)$$

where  $h_a^*[m]$  represents a sampled version of the continuous function  $|r|^{-1/2}h^*(t/r)$  at a scale  $r$  corresponding to the scale index  $a$ , and  $x[m]$  represents the time waveform being analyzed. The scale index  $a$  is defined as being proportional to the logarithm of the scale  $r$ . The exact relation is  $a = V \log_2 r$ , where  $V$  is a quantity known as the voices per octave. Accordingly, scale index  $a = 0$  corresponds to scale  $r = 1$ . Explicitly,  $h_a[m]$  is calculated by sampling the Morlet wavelet in the range  $-4$  to  $4$  with  $c = 5.0$ , as used in [17]:

$$h_a[m] = \left( \frac{1}{\sqrt{2}} \right)^{\frac{a}{V}} \exp \left( jcw_m - \frac{w_m^2}{2} \right) \quad (7)$$

where

$$w_m = 4 \left[ \frac{2m - (L-1)}{L-1} \right], \quad 0 \leq m \leq L-1. \quad (8)$$

Here,  $L$  is the number of samples of  $h(t)$  taken and is defined as

$$L = N \cdot 2^{\frac{a}{V}} \text{ rounded to the nearest integer} \quad (9)$$

where

- $N$  integer sufficiently large to provide good sampling of the basis function at the smallest scale used ( $a = 1$ )
- $a$  scale index
- $V$  number of voices per octave.

The choice of  $c = 5.0$  for the Morlet wavelet gives an analysis wavelet that is octave-band in nature, i.e., the frequency width of the wavelet’s Fourier transform at a given frequency  $f_0$  is of the order of  $f_0$  itself. Such octave-band wavelet bases arise naturally in wavelets derived using two-scale equations or from orthogonal perfect-reconstruction filter banks [16].

By choosing a fixed number of voices per octave, we can consider frequency increments expressed as fractions of an octave. For example, if we decide to examine four voices per octave, that is equivalent to considering the set of frequencies  $f_0, 2^{\frac{1}{4}}f_0, 2^{\frac{2}{4}}f_0, 2^{\frac{3}{4}}f_0$ , and  $2^{\frac{4}{4}}f_0$ , which correspond to the fractional octave frequencies defined above.

From (9), it is apparent that  $L$ , which is the number of samples of the basis function, increases with  $a$ . This is equivalent to saying that the wavelet basis function is stretched by increasing the scale (i.e., by increasing  $a$ ). These longer basis functions correspond to filters with lower center frequencies and narrower bandwidths (maintaining constant relative bandwidth however), reflecting the fact that the loss in time resolution is accompanied by an increase in frequency resolution at larger scales. We have adopted the convention

that an increase in scale corresponds to a decrease in frequency. Explicitly, the scale index  $a$  is related to the logarithm of inverse frequency by

$$a = V \log_2 \left( \frac{K}{f} \right), \quad (10)$$

where

- $V$  number of voices per octave
- $K$  a constant
- $f$  frequency.

For a given CWT, we can find the exact relation between scale and frequency by analyzing a pure sinusoid of known frequency so that the proportionality constant  $K$  can be determined. This allows us to label CWT's with frequency values, as well as with the scale indices  $a$  that we use.

The scalogram (SCAL) is defined as the squared modulus of the CWT:

$$\text{SCAL}_x^h[a, b] = |\text{CWT}_x^h[a, b]|^2. \quad (11)$$

The CWT magnitude and scalogram are also presented in both 3-D and contour formats, as described earlier for the STFT and spectrogram. The phases associated with the STFT and the CWT of a data set provide additional representations, as will be illustrated subsequently in one of our examples.

### III. RESULTS

#### A. Response of a Third-Turn Outer Hair Cell to AM Stimuli

To examine the behavior of the velocity response of cochlear sensory cells for stimuli of different frequencies, AM acoustic signals (with unity modulation depth and a fixed modulation frequency of 2.44 Hz) were applied to the ear canal, using different stimulus carrier frequencies. 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. The sampling frequency was fixed at 5 kilosamples/s in recording the responses to AM stimuli.

Fig. 1 shows the velocity tuning curve of a second-row outer hair 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 70 to 2000 Hz. This tuning curve displays a peak at 850 Hz.

Fig. 2 shows the response of this cell when an AM tone with a carrier frequency  $f_c = 415$  Hz (below CF) is applied to the ear. The velocity response of the cell during one cycle of the modulation envelope is displayed in Fig. 2(a). The response does not follow precisely the shape of the input envelope; a slight flattening is apparent near the center of the time waveform. In this and other such curves, the reconstructed waveform has been upsampled by a factor of 64 and interpolated by sinc functions to reduce the artifact known as "false modulation." The velocity-response STFT magnitude, which is shown in 3-D and 2-D formats in Fig. 2(b) and (c), respectively, reveals the time behavior of the spectral components. The STFT shows six spectral components at multiples of the carrier frequency. At the beginning and end of the modulation cycle, when the magnitude of the envelope is

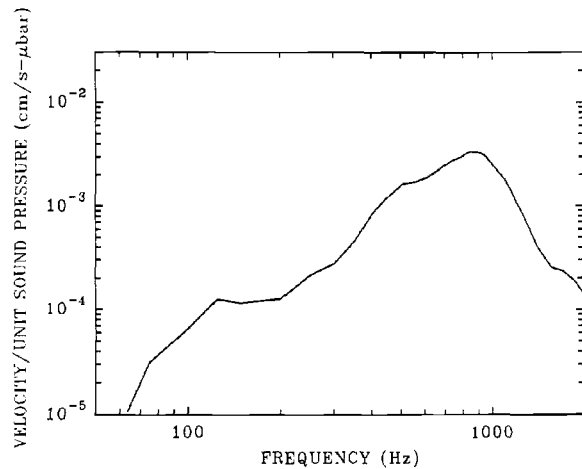


Fig. 1. Velocity tuning curve for a second-row outer hair cell (Ob061422.dt3) in the third turn of the guinea pig temporal-bone preparation. The ordinate displays the peak velocity in centimeters per second per unit sound pressure of  $1 \mu\text{bar}$  ( $1 \text{ dyne/cm}^2$ ). The CF for this cell is 850 Hz.

low, the response closely follows the input, and the response is mainly at the carrier frequency (this is most clearly seen in Fig. 2(c)). As the magnitude of the envelope increases, spectral components at  $f_c, 2f_c, 3f_c, 4f_c, 5f_c,$  and  $6f_c$  appear. The higher spectral components do not follow the time course of the input envelope precisely.

The STFT magnitude at the carrier frequency and at three of its harmonics is plotted in yet another way in Fig. 2(d). The magnitude of the second harmonic is generally greater than that of the third or fourth harmonic. Harmonic generation clearly indicates the presence of 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.

Fig. 3 shows the response of the same cell when an AM tone with a carrier frequency  $f_c = 706$  Hz (near CF) is applied. The velocity response of the cell over one modulation cycle is displayed in Fig. 3(a). Its magnitude is substantially greater than that shown in Fig. 2(a) since the carrier frequency is now near the CF of the cell. The response envelope does not follow the shape of the input envelope. The central portion of the time waveform is highly irregular in appearance. The STFT magnitude of the velocity response (Figs. 3(b) and (c)) again shows the presence of multiple spectral components, but the behavior is far more complex than that seen in Fig. 2(b) and (c). At low magnitudes of the modulation envelope, the response is only at the carrier frequency, as in Fig. 2. At higher magnitudes of the modulation envelope, harmonic components are clearly visible at  $f_c, 2f_c,$  and  $3f_c$ .

At the highest values of the stimulus envelope, however, something unusual occurs. Components are seen at three half-harmonic frequencies ( $3f_c/2, 5f_c/2,$  and  $7f_c/2$ ), as reported previously [10], [12], [18], [19]. In some cells, there is also a component present at  $f_c/2$  [12], but this feature is more-or-less absent in this particular data set. It is of interest to note that Davis *et al.* [20] and Dallos *et al.* [21], [22] observed the

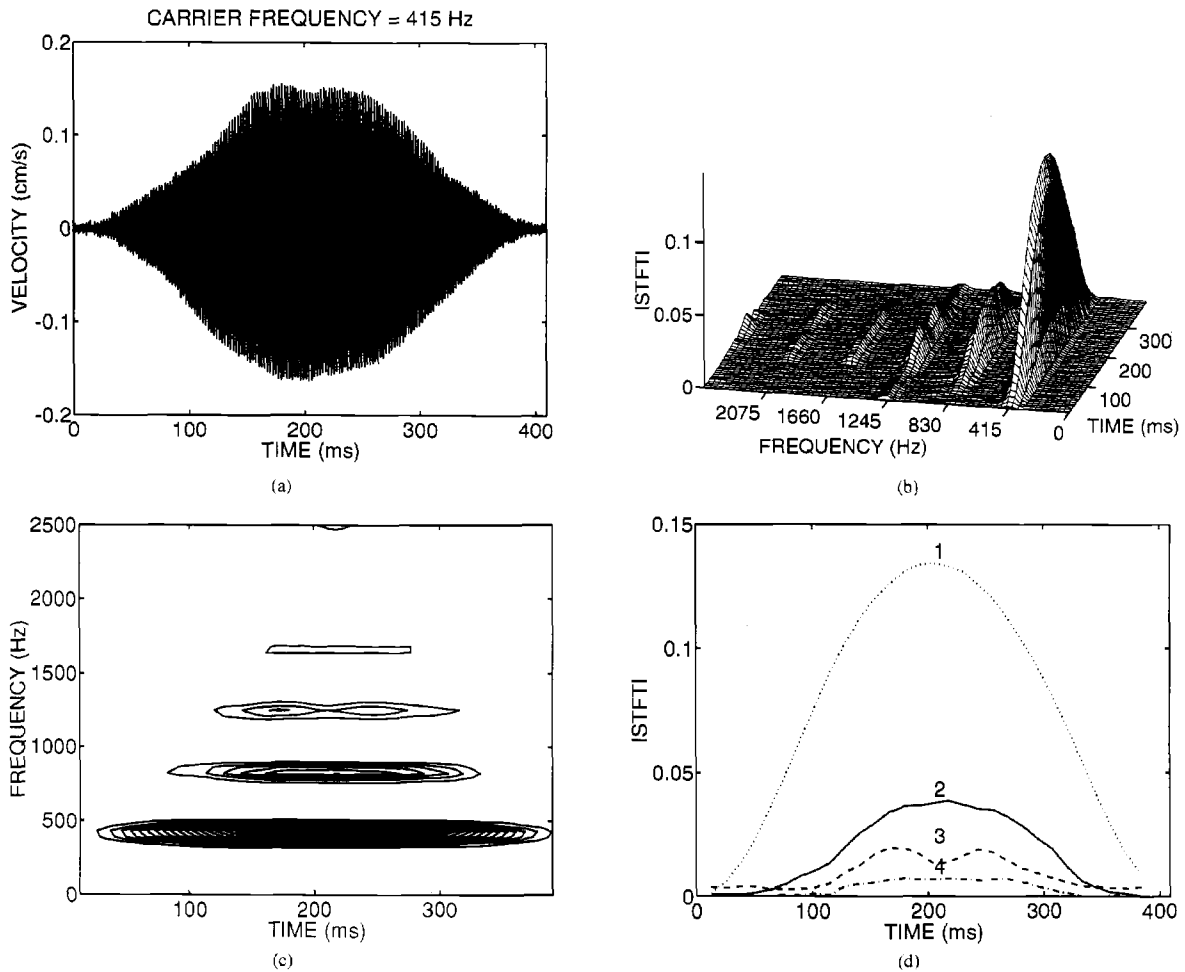


Fig. 2. Velocity response of an outer hair cell in the third turn of a guinea pig temporal-bone preparation to an AM stimulus with carrier frequency  $f_c = 415$  Hz (below CF) and modulation frequency 2.44 Hz. The velocity tuning curve for this cell is shown in Fig. 1. The highest sound pressure level occurring at the center of the input envelope was  $\approx 123$  dB: re  $0.0002$  dyne/cm<sup>2</sup>: (a) Time waveform of the velocity response (in cm/s) of the cell; (b) 3-D spectral plot of the STFT magnitude of the velocity response shown in (a). The  $x$  and  $y$  axes represent time (ms) and frequency (Hz), respectively, and STFT magnitude is plotted on a linear scale on the  $z$  axis. This plot shows spectral components at the carrier frequency  $f_c$  and at five higher harmonic frequencies; (c) same STFT magnitude as shown in (b) but now plotted in 2-D with 20 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude. Each contour curve represents 0.0064 units of STFT magnitude. The magnitude corresponding to a given contour line can be determined by counting the number of contour lines to that curve (including the curve itself) and multiplying by 0.0064; (d) STFT magnitude as a function of time at the carrier frequency  $f_c$  (indicated 1), and at its second, third, and fourth harmonics (indicated 2, 3, and 4, respectively). The shape of the carrier and each of the harmonic curves is quite different. The carrier-frequency curve is symmetrical in time around 205 ms, but the harmonic curves show a symmetry around approximately 216 ms.

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 *et al.* [23] also showed the presence of odd-fractional harmonics (which they determined to be rare in guinea pigs).

A component at dc (or possibly at nearby low frequencies) can also be clearly seen in the STFT magnitude in Fig. 3(b) and (c) near the center of the AM waveform. The dc component corresponds to the nonvibratory motile behavior of the cell in response to the envelope of the AM stimulus, as discussed earlier. Related observations have been made both in isolated outer hair cells and in the guinea pig temporal-bone preparation by Brundin *et al.* [5]–[8].

The STFT magnitude for the carrier, the second and third harmonics, and four half-harmonic frequencies is shown in Fig. 3(d) as a function of time. The amplitude of the second harmonic no longer increases monotonically with the amplitude of the stimulus. The magnitude of the third harmonic is greater than that of the second harmonic. This is opposite to the condition observed at a lower frequency (Fig. 2(d)). The harmonic curves are asymmetrical in time. Furthermore, the half-harmonics appear at times delayed with respect to the peak of the input waveform, as the amplitude of the stimulus is decreasing.

Fig. 4 shows the response of this cell when an AM tone with a carrier frequency  $f_c = 1013$  Hz (above CF) is applied to the preparation. The overall velocity response of the cell over one

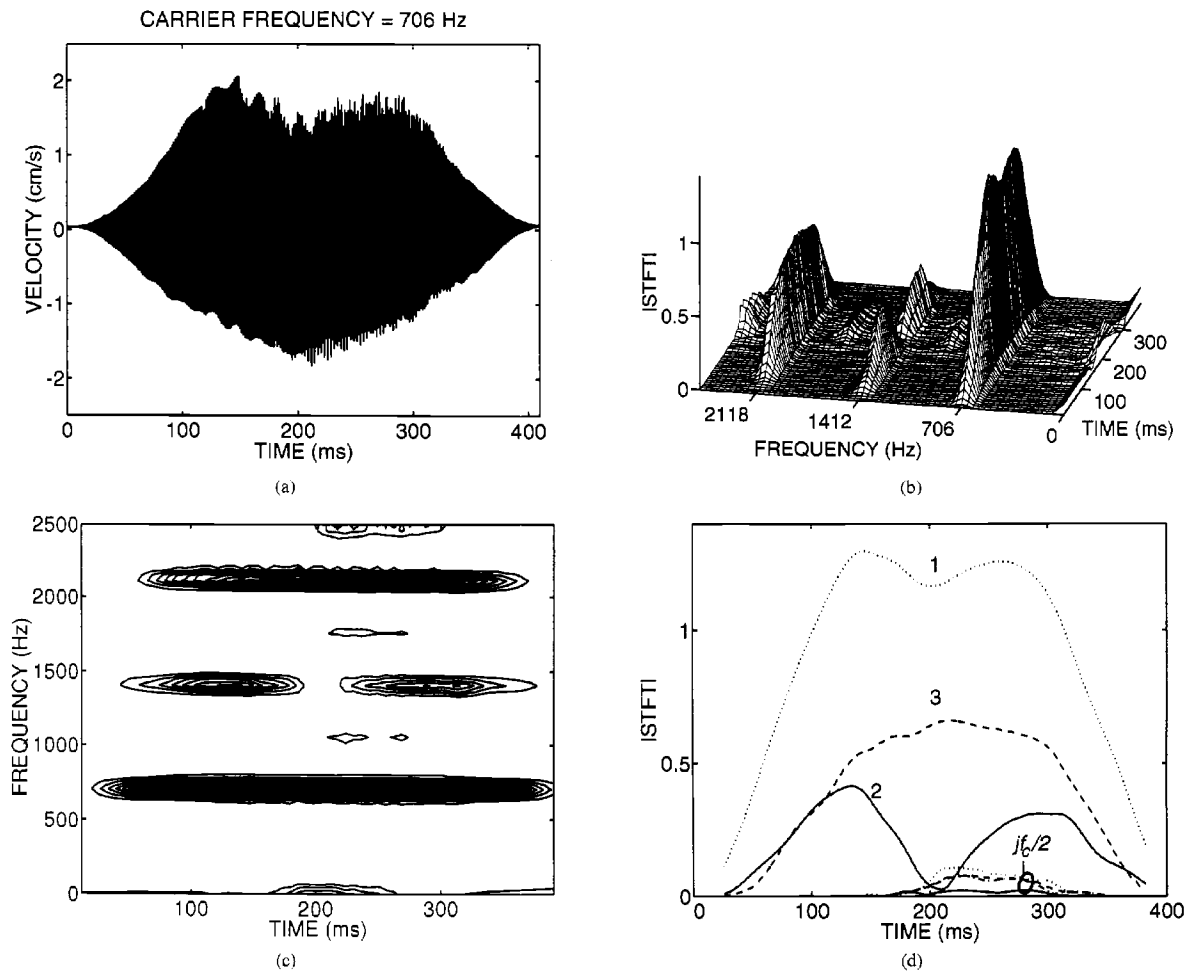


Fig. 3. Velocity response of the same outer hair cell as shown in Fig. 2 but now to an AM stimulus with carrier frequency  $f_c = 706$  Hz (near CF) and modulation frequency 2.44 Hz. The highest sound pressure level, occurring at the center of the input envelope, was  $\approx 137$  dB: re 0.0002 dyne/cm<sup>2</sup>: (a) Time waveform of the velocity response (in cm/s) of the cell; (b) STFT magnitude of the velocity response shown in (a). This plot shows spectral components at the carrier frequency  $f_c$ , at two harmonic components of the carrier ( $2f_c$  and  $3f_c$ ), at three half-harmonic components of the carrier ( $3f_c/2$ ,  $5f_c/2$ , and  $7f_c/2$ ), as well as at dc; (c) same STFT magnitude as shown in (b), but in contour format with 20 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude. Each contour curve represents 0.0634 units of STFT magnitude; (d) STFT magnitude as a function of time at the carrier frequency  $f_c$ , its second and third harmonics, and at four half-harmonics, which are denoted  $jf_c/2$  ( $f_c/2$ -solid curve,  $3f_c/2$ -dash-dot curve,  $5f_c/2$ -dashed curve,  $7f_c/2$ -dotted curve). The shapes of the carrier and each of the harmonic and half-harmonic curves are quite different. They also differ from those observed in Fig. 2(d) ( $f_c = 415$  Hz). There, the second harmonic was greater in magnitude than the third; this condition is now reversed for  $f_c = 706$  Hz. The half-harmonics are present principally between 200 and 300 ms asymmetrically with respect to the time of the peak input stimulus.

cycle of the modulation envelope is displayed in Fig. 4(a). The response follows the shape of the input envelope with little distortion. The STFT magnitude of the velocity response (Fig. 4(b) and (c)) shows the presence of only two spectral components: at  $f_c$  and at  $2f_c$ . Both of these components follow the time course of the input envelope rather faithfully. The STFT magnitudes for the carrier and its second harmonic are shown as a function of time in Fig. 4(d). The second harmonic curve is asymmetrical around the time axis. As we move to yet higher carrier frequencies (which are not shown), the component at the carrier frequency becomes increasingly dominant.

The results presented in Figs. 2–4 are typical of the 60 AM data sets associated with 24 cells (11 outer hair cells

and 13 Hensen's cells) that we have examined thus far using time-frequency representations. We have previously presented similar plots for other third-turn outer hair cells [10]–[12] and for a fourth-turn Hensen's cell [12]. In all cases, the response is reasonably, but not entirely, symmetric in time about the center of the stimulus envelope, indicating the presence of limited memory effects [11].

The CWT of a signal will in general display the same information as the STFT but with a different distribution of time-frequency resolutions [16]. Depending on the nature of the signal and the choice of analysis wavelet, however, information may be easier to discern from one or the other of these transforms. The magnitudes and phases of the STFT and CWT for data taken from the central portion of the time

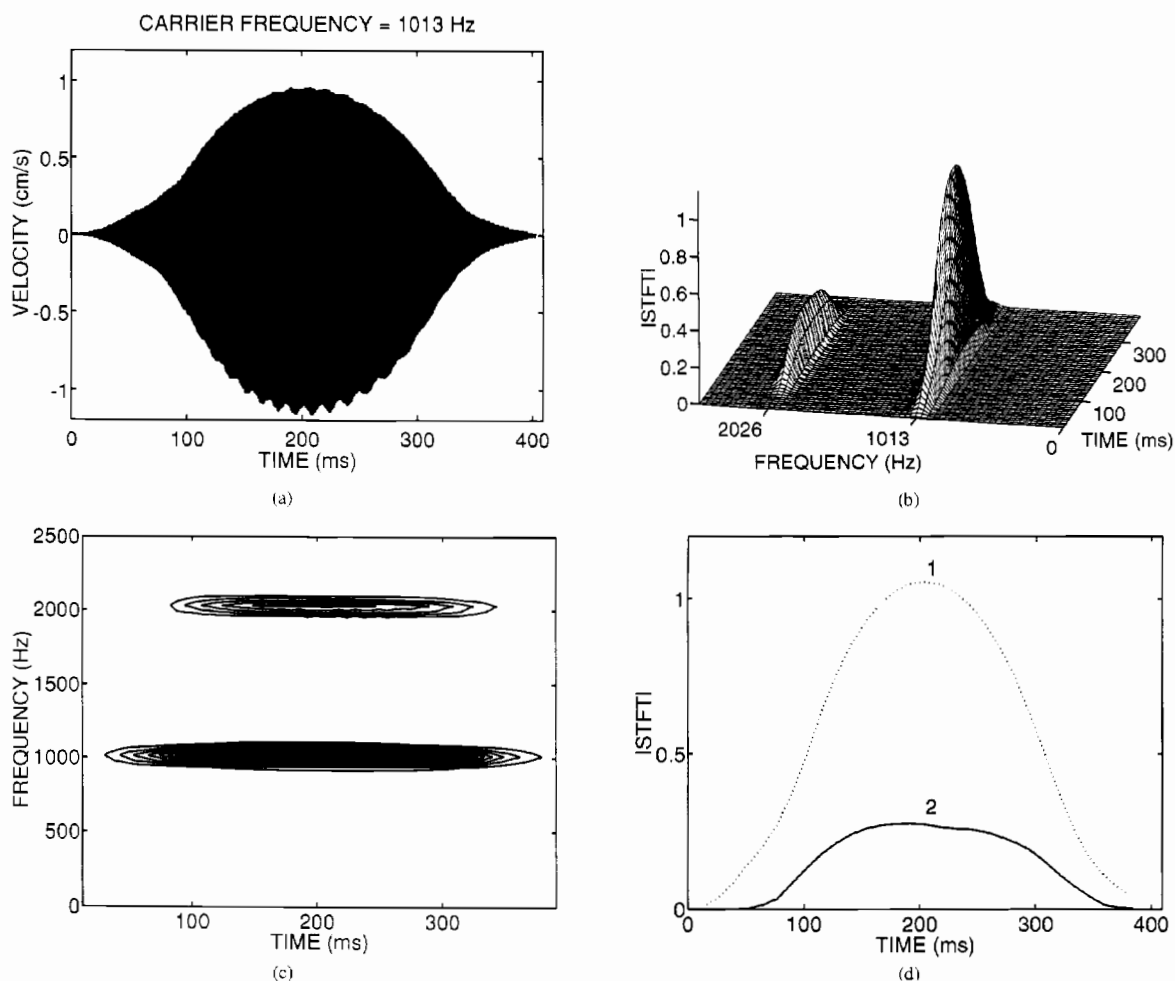


Fig. 4. Velocity response of the same outer hair cell as shown in Figs. 2 and 3 but now to an AM stimulus with carrier frequency  $f_c = 1013$  Hz (above CF) and modulation frequency 2.44 Hz. The highest sound pressure level, occurring at the center of the input envelope, was  $\approx 132$  dB; re  $0.0002$  dyne/cm<sup>2</sup>: (a) Time waveform of the velocity response (in cm/s) of the cell. The slight scalloping apparent in this time waveform is due to false modulation; (b) STFT magnitude of the velocity response shown in (a). This plot shows spectral components only at the carrier frequency  $f_c$  and at the second harmonic  $2f_c$ . Higher harmonic components may have been present but remained undetected because the sampling rate and filtering of the measurement apparatus restricted the maximum observable frequency to 2500 Hz; (c) same STFT magnitude as shown in (b) but in contour format with 20 equally spaced (in STFT magnitude) contour lines joining points of constant magnitude. Each contour curve represents 0.0502 units of STFT magnitude; (d) STFT magnitude as a function of time at the carrier frequency  $f_c$  and at its second harmonic. The shapes of the carrier and the second harmonic curves are quite different. The second harmonic curve is asymmetrical in time. This response was recorded with an upper frequency limit of 2500 Hz so that third and higher harmonics could not be observed.

waveform shown in Fig. 3(a) ( $f_c = 706$  Hz) are compared in Fig. 5. The particular segment of the time waveform analyzed is shown in Fig. 5(a).

The STFT magnitude for this time series is shown in a 3-D plot (Fig. 5(b)) and in contour format (Fig. 5(c)). The STFT shows energy primarily at the carrier frequency and at two harmonics ( $2f_c$  and  $3f_c$ ), but discernible spectral components are also present at  $3f_c/2$ ,  $5f_c/2$ ,  $7f_c/2$ , and 0 (dc). The component at  $3f_c$  is considerably larger than that at  $2f_c$ , but the component at  $2f_c$  is seen to increase significantly during the time course of the signal.

The octave-band based CWT magnitude for the same time series is shown in Fig. 5(d) and (e) in 3-D and contour format, respectively. The frequencies on the right ordinate of Fig. 5(e)

correspond to the linear scale-index axis on the left ordinate as determined by the logarithmic relation between frequency and scale index given in (10) (for this plot  $V = 4$  and  $K = 3841$  Hz). The components at  $f_c$  and  $3f_c$  are clearly visible in the 3-D plot (Fig. 5(d)) as two ridges of approximately constant magnitude. Because of the poor frequency resolution of this CWT at high frequencies, however, the component at  $2f_c$  is masked by the stronger components at  $f_c$  and  $3f_c$ . The CWT contour plot in Fig. 5(e) is difficult to interpret because of the jagged nature of the contour lines. This situation arises when the duration of the wavelet basis function is of the same order as the period of the sinusoidal component. It should be noted that the peak in the CWT of this signal occurs at a scale index of 10, corresponding to a frequency of 679 Hz, rather than

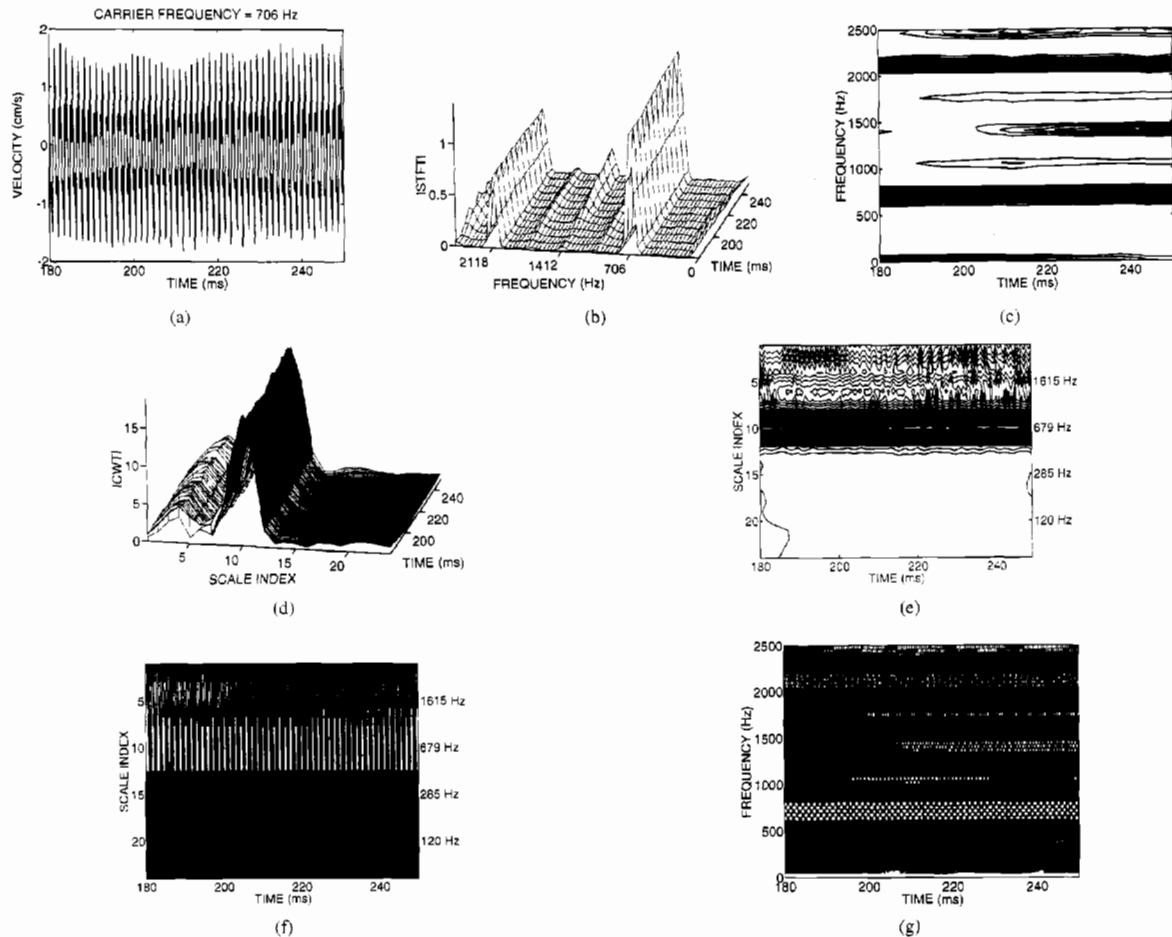


Fig. 5. Comparison of the STFT and the octave-band-based CWT for analyzing cellular velocity responses to an AM stimulus. The signal is the central portion of the velocity response shown in Fig. 3(a): (a) Time waveform of a portion of the velocity response (in centimeters per second) for times between 180 and 250 ms; (b) STFT magnitude of the velocity response shown in (a). This plot, which is an expanded view of the central portion of Fig. 3(b), shows spectral components at seven frequencies:  $0$ ,  $f_c$ ,  $3f_c/2$ ,  $2f_c$ ,  $5f_c/2$ ,  $3f_c$ , and  $7f_c/2$ ; (c) same STFT magnitude as shown in (b), but in contour format with 20 equally spaced contour lines, each representing 0.0615 units of STFT magnitude; (d) CWT magnitude of the velocity response shown in (a). The  $x$  and  $y$  axes represent time (ms) and scale index, respectively, and CWT magnitude is plotted on a linear scale on the  $z$  axis. This plot shows two significant components centered at scale indices of 10 and 3; (e) same CWT magnitude as shown in (d) but now in contour format with 20 equally spaced contour lines joining points of constant magnitude. Each contour line represents 0.85 units of CWT magnitude; (f) phase of the CWT of the velocity response shown in a gray-scale format. The phase is mapped to the range  $[-\pi, \pi]$ , corresponding to the transition from absolute black ( $-\pi$ ) to absolute white ( $\pi$ ). In areas where the CWT magnitude falls below a certain threshold, the phase is not plotted. This appears as gray areas on the figure. Comparison with the CWT-magnitude contour plot in (e) reveals that the thresholded phase map provides a crisper representation for this data set; (g) phase of the STFT of the velocity response shown in a gray-scale format. As in (f), the phase is mapped to the range  $[-\pi, \pi]$ , which again corresponds to the transition from absolute black ( $-\pi$ ) to absolute white ( $\pi$ ). In areas where the STFT magnitude falls below a certain threshold, the phase is not plotted. These appear as gray areas on the figure. The thresholded STFT phase map shown here contains essentially the same information as the STFT magnitude contour plot shown in (c), clearly illustrating the time course of the seven frequency components present.

706 Hz, which is the applied carrier frequency. This difference arises because the scale index is quantized to integer values, and a scale index of 10 corresponds to the frequency nearest to the actual frequency of the signal (a scale index of nine maps to a frequency of 807 Hz, which lies farther from the carrier frequency than 679 Hz). The use of more voices per octave in calculating the CWT would allow a more precise estimate of the frequency of vibration to be obtained but at the expense of added computational complexity.

The discriminability provided by the octave-band-based CWT magnitude plots in Fig. 5(d) and (e) is clearly seen

to be inferior to that provided by the STFT magnitude plots presented in Fig. 5(b) and (c), respectively. This issue will be discussed further in Section IV.

Since we have used a complex wavelet for calculating the CWT, it is also of interest to examine what information may be discerned from the phase of the CWT for this time series. The phase is shown in Fig. 5(f), with  $-\pi$  plotted as absolute black and  $\pi$  plotted as absolute white. Only regions of the transform with CWT magnitude greater than a selected threshold are plotted, thereby eliminating clutter arising from the phases of nonsignificant signal components. The component at  $f_c$



shows up clearly at a scale index  $a = 10$ . It registers as a set of bands spaced at time intervals equal to the period of the carrier component. For smaller scales (higher frequencies), the component at  $3f_c$  is also readily visible (scale index  $a = 3$ ), with a periodicity that is one third of the fundamental period (i.e., there are three phase transitions from  $-\pi$  to  $\pi$  at this scale for every transition at the carrier frequency). Components at frequencies between  $f_c$  and  $3f_c$  are also discernible, starting at a time of about 230 ms. This feature, corresponding to the appearance of components between  $f_c$  and  $3f_c$ , was not readily visible in the octave-band-based CWT magnitude plots. Indeed, the CWT phase map is more useful than the CWT magnitude alone, inasmuch as it combines some magnitude information (determined by the chosen threshold) with phase information. Nevertheless, we cannot clearly distinguish the three separate frequency components at  $3f_c/2$ ,  $2f_c$ , and  $5f_c/2$ , which are evident in the STFT's shown in Fig. 5(b) and (c).

In Fig. 5(f), the bands in the phase map corresponding to the fundamental component at  $f_c$  are seen to be spread over five scale indices, ranging from 8 to 12. At any fixed time, the phase of this signal transform remains essentially constant across these five scale indices. To understand why this occurs, (5) is recast in the form of an inverse Fourier transform:

$$\text{CWT}_x^h(r, \tau) = \sqrt{|r|} \int_{-\infty}^{\infty} X(f) H^*(fr) \exp(j2\pi f\tau) df \quad (12)$$

where  $X(f)$  and  $H(f)$  are the continuous-time Fourier transforms of  $x(t)$  and  $h(t)$ , respectively. The CWT of a pure sinusoid of frequency  $f_0$  can be readily calculated from (12) and shown to be equal to  $\sqrt{r} [H^*(f_0r) \exp(j2\pi f_0\tau) - H^*(-f_0r) \exp(-j2\pi f_0\tau)]$ . If  $|H^*(f_0r)| \gg |H^*(-f_0r)|$ , then the phase of the CWT is essentially the phase of the dominant component  $H^*(f_0r) \exp(j2\pi f_0\tau)$ . For the Morlet wavelet we have used,  $H(f)$  is purely real; therefore, the phase of the CWT for a single sinusoid of frequency  $f_0$  is well approximated by  $2\pi f_0\tau$  in the regions where the approximation  $|H^*(f_0r)| \gg |H^*(-f_0r)|$  holds (this is the case for  $f_0 = 706$  Hz and  $r$  corresponding to scale indices near 10). However, the relatively high threshold used for the display in Fig. 5(f) limits our view of the constant phase behavior to a range of scale indices between 8 and 12.

It should be noted that the data used to calculate the CWT was upsampled and interpolated in order to stretch out the scale axis. This is permissible since the original signal was well bandlimited at 2500 Hz.

For completeness, we display the phase associated with the STFT of the same time series in Fig. 5(g). Again,  $-\pi$  is plotted as absolute black and  $\pi$  as absolute white, and an STFT-magnitude threshold is selected, below which the phase is not plotted. The STFT phase plot retains the behavior of the STFT-magnitude contour plot shown in Fig. 5(c), clearly distinguishing and exhibiting the time course of the seven significant frequency components. The presence of the white bar at zero frequency indicates that the dc component of this signal has a phase of  $\pi$ , reflecting the negative mean of the time waveform in Fig. 5(a).

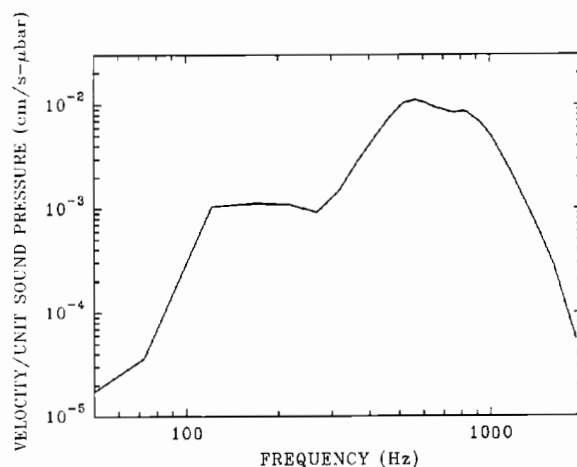


Fig. 6. Velocity tuning curve for a Hensen's cell (06051530.dt1) in the third turn of the guinea pig temporal-bone preparation. The ordinate displays the peak velocity in centimeters per second per unit sound pressure of 1  $\mu$ bar (1 dyne/cm<sup>2</sup>). The CF for this cell is 550 Hz.

The phase shifts between adjacent time bins and frequency bins in the STFT give rise to a checkered pattern in place of the banded pattern associated with the CWT phase plot in Fig. 5(f), for which the phase changes only along the time axis. The origin of the phase changes in frequency for the STFT can be understood by rewriting (2) as an inverse Fourier transform:

$$\text{STFT}_x^g(t, f) = e^{-j2\pi ft} \int_{-\infty}^{\infty} X(u) G^*(u - f) \exp(j2\pi ut) du \quad (13)$$

where  $X(u)$  and  $G(u)$  are the continuous-time Fourier transforms of  $x(t)$  and  $g(t)$ , respectively. The prefactor  $e^{-j2\pi ft}$ , which is present in (13) but not in (12), shows that the phase of the transform varies linearly with frequency at any fixed time  $t$ .

From all of the measures presented in Fig. 5, it is apparent that the STFT is superior to the octave-band-based CWT that we have used, as is explained further in Section IV.

### B. Response of a Third-Turn Hensen's Cell to the Step Turn-On of a Tone

To examine the behavior of the transient velocity response of cochlear sensory cells for stimuli of different frequencies, step-onset tones were applied to the ear canal. The response of a nonlinear system to a transient stimulus provides complementary information to that provided by the response to an AM tone. The sampling frequency was fixed at 25 ksamples/s in recording the responses to step-onset stimuli.

We proceed to detail the response of a Hensen's cell in the third turn of the guinea pig temporal-bone preparation using both CWT and STFT analysis techniques. Fig. 6 shows the velocity tuning curve for this cell. It is broadly peaked with its maximum value at 562 Hz (CF). The three frequencies for which the time waveforms are shown are as follows: below CF ( $f_{\text{stim}} = 122$  Hz; Fig. 7(a)), at CF ( $f_{\text{stim}} = 562$  Hz; Fig. 8(a)), and above CF ( $f_{\text{stim}} = 1489$  Hz; Fig. 9(a)). The transient

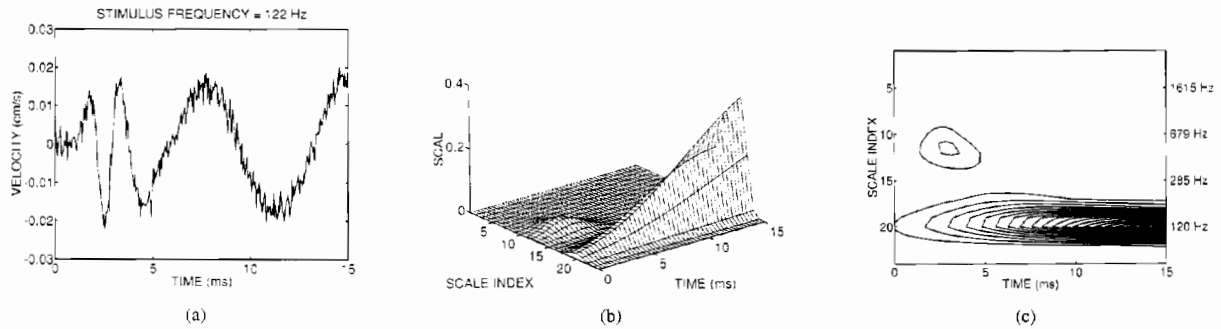


Fig. 7. Velocity response of a Hensen's cell in the third turn of a guinea pig temporal-bone preparation to the step turn-on of a sinusoidal stimulus of frequency  $f_{stim} = 122$  Hz (below CF). The tuning curve for this cell is shown in Fig. 6. The sound pressure level of the sinusoidal stimulus was  $\approx 105$  dB: re  $0.0002$  dyne/cm<sup>2</sup>: (a) Time waveform of the velocity response (in centimeters per second) of the cell; (b) scalogram (3-D plot) of the velocity response shown in (a). The  $x$  and  $y$  axes represent time (ms) and scale index, respectively, and scalogram magnitude is plotted on a linear scale on the  $z$  axis; (c) same scalogram as shown in (b) but now plotted in contour format with 20 equally spaced contour lines, each representing 0.0195 units of scalogram magnitude. This plot reveals the presence of an initial component at  $\approx 300$  Hz, which dies away to leave a steady-state component at the stimulus frequency.

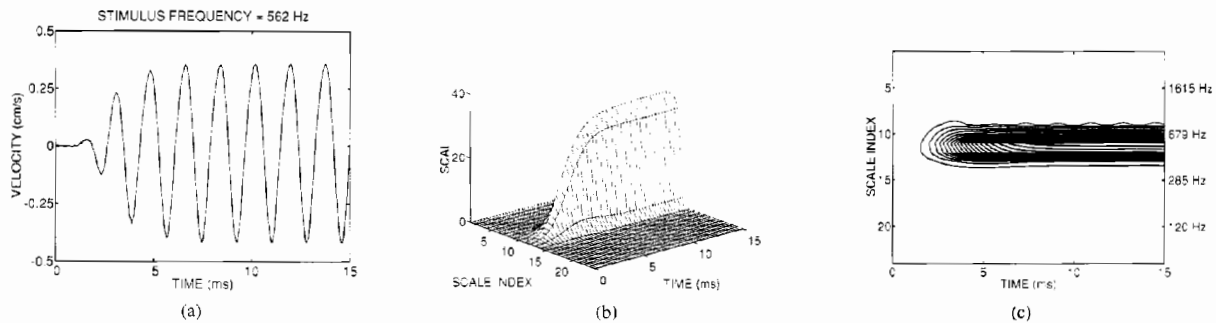


Fig. 8. Velocity response of the same Hensen's cell as shown in Fig. 7 but now to the step turn-on of a sinusoidal stimulus of frequency  $f_{stim} = 562$  Hz (at CF). The sound pressure level of the sinusoidal stimulus was  $\approx 114$  dB: re  $0.0002$  dyne/cm<sup>2</sup>: (a) Time waveform of the velocity response of the cell; (b) scalogram of the velocity response shown in (a); (c) same scalogram as shown in (b) but in contour format with 20 equally spaced contour lines, each representing 1.87 units of scalogram magnitude. This plot principally shows a steady-state component at the stimulus frequency.

response time of the sound-generation system is negligibly short at these low frequencies.

As we have reported previously [24], the qualitative characteristics of the transient response change with frequency. When the stimulus frequency is below the CF of the cell, the transient vibration amplitude monotonically approaches its final value, while its instantaneous frequency decreases toward the stimulus frequency. When the stimulus frequency is well above the CF, there is a substantial amplitude overshoot, while the instantaneous vibration frequency starts out low and then increases to the stimulus frequency. Finally, for stimulus frequencies in the vicinity of CF, there may be either a slight overshoot or a steady approach to the final amplitude (both are observed), and the instantaneous frequency remains relatively constant.

These observations can be made clearer by examining the CWT magnitude, the CWT phase, or the scalogram of the transient responses. Since there is little difference in these representations for the transient data, we chose the latter for diversity. The octave-band-based scalograms are displayed in 3-D format in Figs. 7–9(b) and in contour format in Figs. 7–9(c) [13]. In these figures, the scalogram is constructed

over six octaves with  $V = 4$  voices per octave so that an increase of 4 in the scale index is equivalent to a halving in frequency. The value of  $N$  (the length, in samples, of the lowest-scale wavelet basis function) was set equal to 40. We use the same scale-index axis as in Fig. 5(d)–(f), as the same parameters are used to calculate the CWT in both cases.

The scalogram clearly shows how energy is diverted into frequencies that are pulled toward the CF at the onset of vibration before reaching steady-state behavior when the energy remains concentrated at the stimulus frequency. In Fig. 7(b) and (c), for example, the stimulus frequency is at 122 Hz, but there is an initial concentration of response at  $\approx 300$  Hz. This corresponds to the decrease in instantaneous frequency observed in the time waveform. A similar pattern is observed in Fig. 9(b) and (c), in which the early transient displays a concentration of the response close to 750 Hz (again in the vicinity of the CF) for a driving frequency of 1489 Hz. The substantial magnitude of the scalogram for this peak, in comparison with its eventual steady-state value at the stimulus frequency, reflects the overshoot evident in the time waveform. It is also apparent from the tuning curve in Fig. 6 that the steady-state response at 750 Hz far outweighs that at 1489

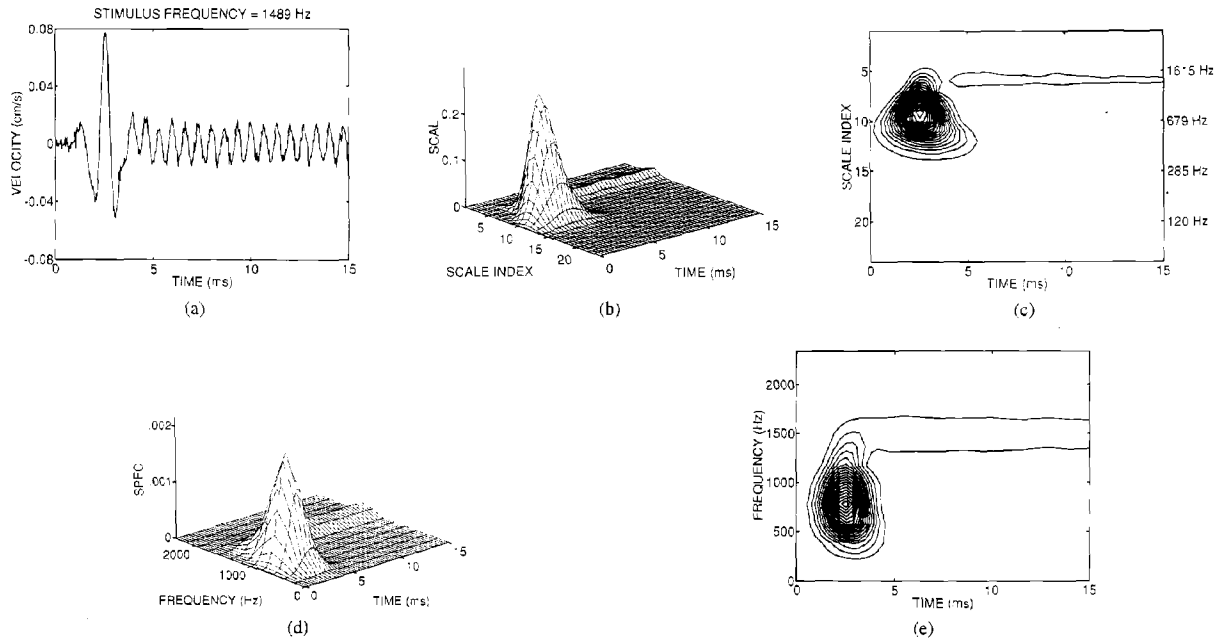


Fig. 9. Velocity response of the same Hensen's cell as shown in Figs. 7 and 8 but now to the step turn-on of a sinusoidal stimulus of frequency  $f_{stim} = 1489$  Hz (well above CF). The sound pressure level of the sinusoidal stimulus was  $\approx 109$  dB; re  $0.0002$  dyne/cm<sup>2</sup>: (a) Time waveform of the velocity response (in centimeters per second) of the cell; (b) scalogram of the velocity response shown in (a); (c) same scalogram as shown in (b) but in contour format with 20 equally spaced contour lines, each representing 0.012 units of scalogram magnitude. This plot reveals the presence of an initial component at  $\approx 750$  Hz, which dies away to leave a steady-state component at the stimulus frequency; (d) spectrogram of the velocity response shown in (a). The  $x$  and  $y$  axes represent time (ms) and frequency (Hz), respectively, with spectrogram magnitude plotted on a linear scale on the  $z$  axis. The spectrogram closely resembles the scalogram shown in (b); (e) same spectrogram as shown in (d) but in contour format with 20 equally spaced contour lines, each representing 0.000094 units of spectrogram magnitude. Like the scalogram shown in (c), there is a response at  $\approx 750$  Hz, which dies away to leave a steady-state component at the stimulus frequency.

Hz. Finally, the response to a stimulus of 562 Hz, which is shown in Fig. 8(b) and (c), reveals only a response at the stimulus frequency with no apparent initial transient at any other frequency since 562 Hz is at CF.

It is of interest to compare the scalogram and the spectrogram of the transient response. For a stimulus frequency of 1489 Hz, this comparison is provided by the spectrogram shown in Fig. 9(d) and (e). These figures are quite similar in character to the scalogram shown in Fig. 9(b) and (c); both display an initial burst of energy at a frequency pulled towards the CF. Although the scalogram and spectrogram provide approximately equivalent performance for this example, the scalogram provides a more compressed view of the signal's behavior because of the manner in which frequency is mapped to a logarithmic scale. Nevertheless, from Fig. 9(c) and (e), it is clear that in this case, both yield a good estimate of the center frequency of the cell's initial vibration (750 Hz). For the particular octave-band-based wavelet that we have used, the scalogram has slightly better frequency resolution (and slightly worse time resolution) than the spectrogram.

There is an important distinction between the CWT (or scalogram) and the STFT (or spectrogram) in mapping sinusoidal signals. Unlike the STFT, the CWT does not map equal-magnitude sinusoidal components of different frequencies to transforms of equal magnitude. For instance, comparing the ratio of the roughly steady-state sinusoidal waveforms in Fig. 8(a) and 9(a) (which is  $\approx 0.35/0.01 = 35$ ) to the ratio of

the scalograms of these steady-state components in Fig. 8(b) and 9(b) (which is  $\approx 40/0.01 = 4000$ ) illustrates that the scalogram magnitudes do not reflect the square of the time-waveform magnitude ratio (since  $4000 \neq (35)^2$ ). This can be understood in terms of the scale-dependent prefactors in (5) and (12).

The attentive reader will notice the absence of the harmonic and half-harmonic frequencies in the transient response. These were observed when AM stimuli were used (see Section III-A). The reason is straightforward: the sound pressure levels used in recording the transient responses were considerably lower than the values used in the AM experiments. Nonlinear effects are more pronounced at higher acoustic intensities.

### C. 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 [9]–[12], [18], [19], [25], we can consider analogies to other well-studied nonlinear systems. Keilson *et al.* [26]–[28] have already analyzed the responses of a class of bilinear oscillators and compared the results with experimental findings [9].

Here, we consider a negative-stiffness Duffing oscillator model [29]–[31] defined by the equation

$$\ddot{x}(t) + \alpha\dot{x}(t) - \beta x(t) + \gamma x^3(t) = F(t) \quad (14)$$

where

$x(t)$	position
$t$	time
$\alpha, \beta, \gamma$	constants
$F(t)$	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 CF, the value of which is determined by suitable choice of the parameters  $\alpha$ ,  $\beta$ , and  $\gamma$ . Motion between the wells gives rise to another, lower, characteristic frequency. For comparison with the cellular velocity data considered in Sections III-A and B, we choose  $F(t)$  to be either a slowly modulated AM signal or a tone with step turn-on (with suitably chosen initial conditions).

Fig. 10(a) shows the velocity response  $\dot{x}$  for such a system driven by a unity modulation-depth AM signal with carrier frequency 706 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  $f_m$  are 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 850 (as suggested by the experimental tuning curve of Fig. 1) by choosing  $\alpha = 2000$ ,  $\beta = 1.43 \times 10^7$ , and  $\gamma = 7.13 \times 10^6$ .  $F_0$  was set equal to  $4.6 \times 10^6$ .

The response does not follow the shape of the input very closely, and a sudden irregularity appears in the time waveform at time  $t = 210$ . Indeed, the model response shown here bears some qualitative similarity to the corresponding data shown in Fig. 3(a). The STFT magnitude of this velocity response (Fig. 10(b), (c), and (d)) shows the presence of multiple spectral components, not unlike those seen in the data (compare with Fig. 3(b), (c), and (d)). The time development of the fundamental and harmonic components are particularly similar. However, in the vicinity of time  $t = 210$ , the model response jumps somewhat suddenly and briefly to encompass a large range of frequencies particularly centered about the half-harmonic frequencies at  $f_c/2$ ,  $3f_c/2$ , and  $5f_c/2$ . This is a consequence of the subharmonic cascade (period-doubling [32]) 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 towards chaotic behavior. Similar features are also present in the cellular velocity data at very large acoustic intensities (see Fig. 3(b) and (c) and [9], [10], [12], [18], [19]). However, the data do not generally exhibit quite as broad a range of frequencies as the model, at least for the model parameters illustrated. In addition, the data exhibit a dc displacement in response to the stimulus envelope that is absent in the Duffing model.

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 that are similar to those seen in Fig. 2. For carrier frequencies well above the CF, the response is mostly at the carrier frequency as in Fig. 4.

Fig. 11(a) shows the velocity response  $\dot{x}$  of the model system to the step turn-on of a tone with frequency  $f_{\text{stim}} = 1489$ . The driving function was of the form

$$F(t) = F_0 \sin[2\pi f_{\text{stim}}(t - t_0)]u(t - t_0)$$

where

$f_{\text{stim}}$	stimulus frequency
$u(t)$	unit step function
$t_0$	measured time delay for the production of experimental data (1 ms)
$F_0$	a constant.

The characteristic frequency of the (small-amplitude, undamped) velocity response was set at 562 (as suggested by the experimental tuning curve of Fig. 6) by choosing  $\alpha = 2000$ ,  $\beta = 6 \times 10^6$ , and  $\gamma = 3 \times 10^6$ .  $F_0$  was set equal to  $4.6 \times 10^5$ . Except for the driving frequency, these parameters are not too different from those used to generate Fig. 10.

The scalogram of this time waveform is shown in Fig. 11(b) and (c). Fig. 11(c) shows an initial peak of energy at a frequency of about 700, which is pulled substantially away from the stimulus frequency of 1489 toward the CF. These plots should be compared with those in Fig. 9(b) and (c). Some features are shared in common; however, the principal distinction is that in Fig. 11(b) and (c), the scalogram magnitude of the initial low-frequency transient peak is far lower than at the stimulus frequency, whereas in Fig. 9(b) and (c), the opposite is true. This reflects the significant overshoot in the time waveform shown in Fig. 9(a), which is absent in the model response shown in Fig. 11(a). This feature is a result of the broader bandwidth of the model's tuning in comparison with the cellular system.

The modified Duffing oscillator does not seem to be a terribly good model for cellular dynamics in the cochlea; it is clearly too simple and too idealized. 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 may contain some of the features of the Duffing system, but then again, many other nonlinear systems with quite different characteristics also share these features. More work is clearly needed to describe in detail the nature of the nonlinearities in the cochlea and to develop more appropriate models.

#### IV. DISCUSSION

We have already seen in Fig. 5 that the STFT and the octave-band-based CWT have quite different frequency- and time-resolution characteristics. It is these characteristics that determine the relative advantages of their use for the analysis of a given signal.

In the STFT analysis of the AM responses presented here, the frequency resolution (which is defined as the full width at  $1/e$ -maximum of the window's Fourier transform) was 99.5 Hz at all frequencies from 0 to 2500 Hz (the Nyquist

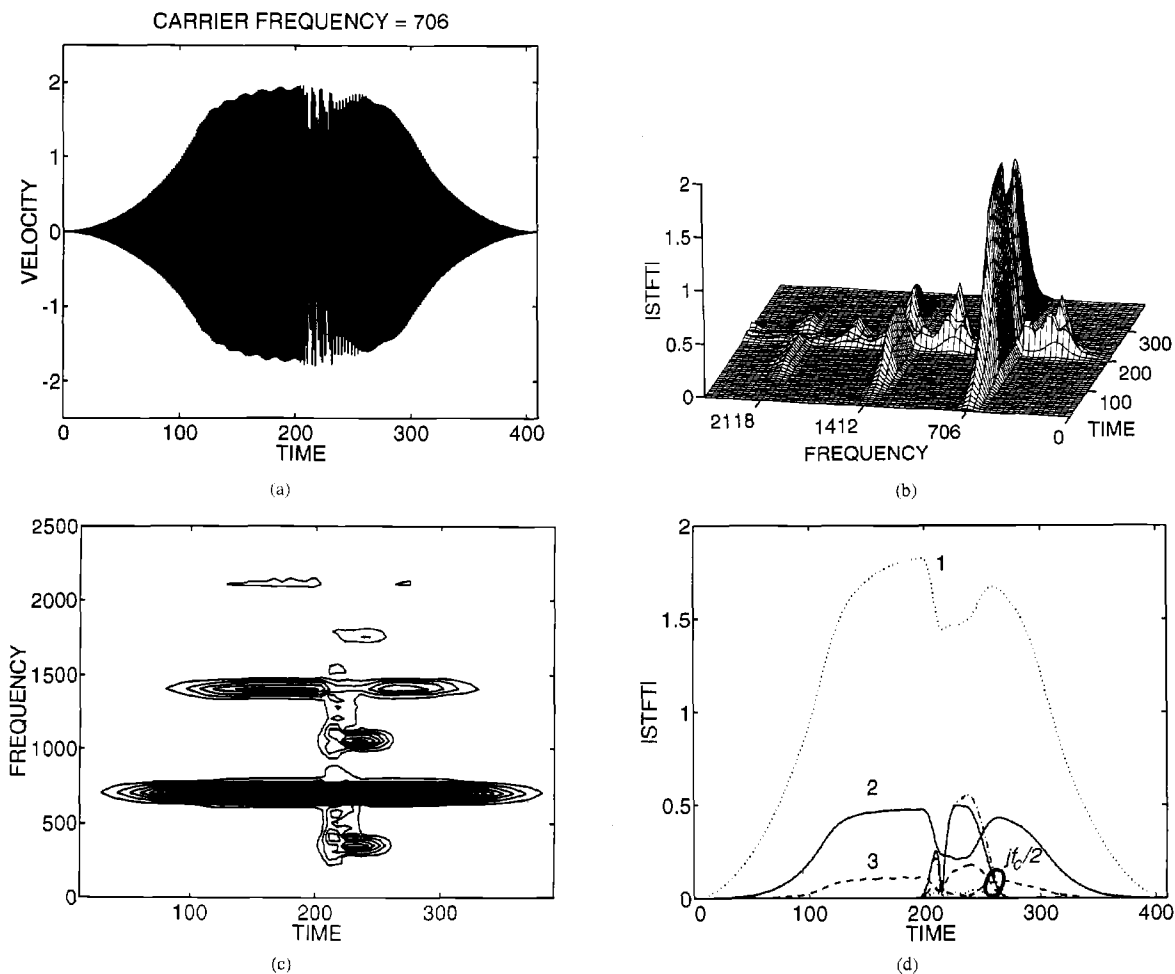


Fig. 10. Velocity response of a model system based on the negative-stiffness Duffing oscillator (as described in Section III-C) to an AM stimulus with carrier frequency  $f_c = 706$  (near CF) and modulation frequency 2.44. The characteristic frequency of the model system has been set at 850: (a) Time waveform of the velocity response of the model system; (b) STFT magnitude of the velocity response shown in (a). The  $x$  and  $y$  axes represent time and frequency, respectively, both in arbitrary units, and STFT magnitude is plotted on a linear scale on the  $z$  axis. This plot shows spectral components at the carrier frequency  $f_c$  and at two harmonic frequencies ( $2f_c$  and  $3f_c$ ). In addition, at the time of the sudden irregularity in the waveform shown in (a), spectral broadening appears briefly, centered about the half-harmonic components at  $f_c/2$ ,  $3f_c/2$ ,  $5f_c/2$ , and  $7f_c/2$ . There is no displacement component at dc; (c) same STFT magnitude as shown in (b) but in contour format with 20 equally spaced contour lines, each representing 0.087 units of STFT magnitude; (d) STFT magnitude as a function of time at the carrier frequency  $f_c$ , its second and third harmonics, and at four half-harmonics, denoted  $f_c/2$  ( $f_c/2$ -solid curve,  $3f_c/2$ -dash-dot curve,  $5f_c/2$ -dashed curve,  $7f_c/2$ -dotted curve). The shapes of the carrier and each of the harmonic and half-harmonic curves are different. The half-harmonics occur in a narrow time window between 200 and 270, cotemporally with the notches in the STFT magnitudes of the harmonic components.

frequency), whereas the time resolution (which is defined as the full width at  $1/e$ -maximum of the window function in time) was 12.8 ms at all frequencies. Consequently, events taking place within a time scale shorter than this resolution time could not be clearly distinguished. The spectrogram also has a characteristic time-frequency resolution.

The time and frequency resolutions of the CWT do, in contrast, vary with scale, though, as for the STFT, the product of the time and frequency resolutions remains fixed. The frequency resolution is worst at small scales and improves (i.e., decreases in width) with increasing scale (whatever the choice of wavelet basis). Since scale is proportional to inverse frequency (see (10)), this means that CWT frequency resolution is best for low frequencies.

This is, of course, the fundamental difference between the STFT and the CWT. For the STFT, the time and frequency resolutions are constant for all frequencies, whereas for the CWT, the time and frequency resolutions are continuously changing as functions of frequency. The choice of window length is critical when implementing the STFT, inasmuch as it fixes the time-frequency resolution thereafter. For a particular data set, there is usually an optimum choice, which can generally be found by trial and error. For the Morlet wavelet-based CWT, on the other hand, the critical choice is the frequency at which a desired time or frequency resolution is achieved, as governed by the choice of the parameter  $c$  in (7).

For an octave-band-based CWT, we can estimate the magnitudes of these frequency resolutions by considering the

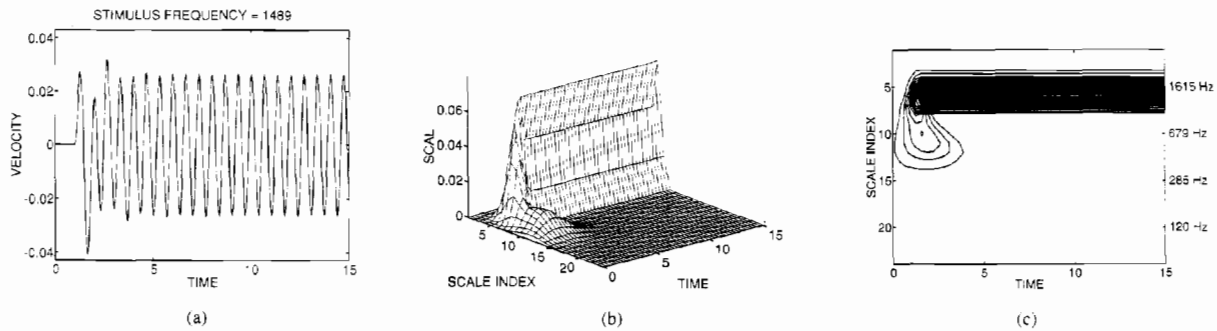


Fig. 11. Velocity response of a model system based on the modified Duffing oscillator (as described in Section III-C) to the step turn-on of a sinusoidal stimulus of frequency  $f_{stim} = 1489$  (well above CF). The characteristic frequency of the model system has been set at 550: (a) Time waveform of the velocity response of the model system; (b) scalogram of the velocity response shown in (a). The  $x$  and  $y$  axes represent time and scale index, respectively, both in arbitrary units, and scalogram magnitude is plotted on a linear scale on the  $z$  axis; (c) same scalogram as shown in (b) but in contour format with 20 equally spaced contour lines, each representing 0.0034 units of scalogram magnitude. This plot reveals the presence of an initial component at  $\approx 700$ , which dies away quickly to leave a steady-state component at the stimulus frequency.

relation between the CWT and its discrete counterpart, the discrete wavelet transform (DWT). In the DWT, a signal is divided into octave bands using filter banks composed of high- and low-pass half-band filters [16]. The lowest scale of the transform corresponds to extracting those signal components whose digital frequencies lie in the range  $[\pi/2, \pi]$ . Successively higher scales correspond to the frequency bands  $[\pi/4, \pi/2]$ ,  $[\pi/8, \pi/4]$ ,  $\dots$ ,  $[\pi/2^{m+1}, \pi/2^m]$ ,  $\dots$ . The frequency resolution at the lowest scale is essentially  $\pi/2$  (which is the width of the corresponding frequency band at the lowest scale), the frequency resolution at the next highest scale is  $\pi/4$  (the width of the band  $[\pi/4, \pi/2]$ ), and so on. Since an octave-band-based CWT interpolates between the octave scales of the DWT, the frequency resolution of the CWT improves in a similar manner, but gradually rather than in discrete jumps as for the DWT. Not surprisingly then, the frequency resolution of an octave-band-based CWT is essentially one spectral component per octave of frequency, as implied by the name. The scalogram also has a characteristic time-frequency resolution.

The effect of the diminished frequency resolution at high frequencies is evident in our analysis of the cellular velocity responses to AM stimuli, where there are multiple spectral components present within the same octave band. Consider the results presented in Fig. 5, where we have explicitly compared the STFT and CWT for the same data. The STFT magnitude (Fig. 5(b) and (c)) clearly shows the presence of three significant spectral components (at 706, 1412, and 2118 Hz) since the separation of these components is much larger than the 99.5-Hz frequency resolution of our STFT analysis. The octave-band-based CWT of the same data, which is shown in Fig. 5(d) and (e), does not clearly distinguish between the components at 1412 and 2118 Hz. This is because the frequency separation of these spectral components is comparable with the frequency resolution of the CWT at these scales (the full width at  $1/e$ -maximum of the wavelet basis function's Fourier transform is 1590 Hz at a scale index equal to 4, which corresponds to a frequency of 1920 Hz according to (10)). Of course, the benefit of the CWT is that the time

resolution at low scales is greatly enhanced. For instance, the time resolution (which is defined as the full width at  $1/e$ -maximum of the wavelet basis function) at scale index 4 is 0.8 ms, which is a substantial improvement over the 12.8 ms for the STFT at the same frequency. For AM responses, however, this increased time resolution provides no significant advantage since it turns out that there are no brief events in the observed time waveform within those time scales. Note that the product of the frequency and time resolutions is  $4/\pi = 1.27$  for both the Morlet wavelet-based CWT and the STFT.

As we show elsewhere, the usefulness of the CWT analysis technique for our class of AM responses can be enhanced enormously by relaxing the restriction that it be based on octave-band wavelets. This approach turns out to be particularly useful for investigating low-frequency subharmonic components that emerge as the cell is driven harder, and its vibration undergoes period-doubling behavior en route to chaos [33]–[35].

For the analysis of transient responses, the scalogram will often provide a superior measure to the spectrogram because of its improved time resolution. In our examples, nevertheless, the performance of both is rather similar since only one impulsive event (the initial transient response to the step turn-on) occurs within the time resolution of the spectrogram. If the response featured two or more impulsive events separated by less than the time resolution of the spectrogram, the ability of the scalogram to provide finer time resolution would make it particularly useful. It is worthy of mention that both the scalogram and the spectrogram provide a good estimate of the characteristic frequency of the initial response of the cell, which is something that could not be readily discerned from the time waveform alone.

## V. CONCLUSION

The different time-frequency resolution characteristics for the STFT and the octave-band-based CWT make the STFT more suitable for analyzing the responses to AM signals, which change relatively slowly in time. In this paper, we have restricted our comparison to octave-band-based CWTs:

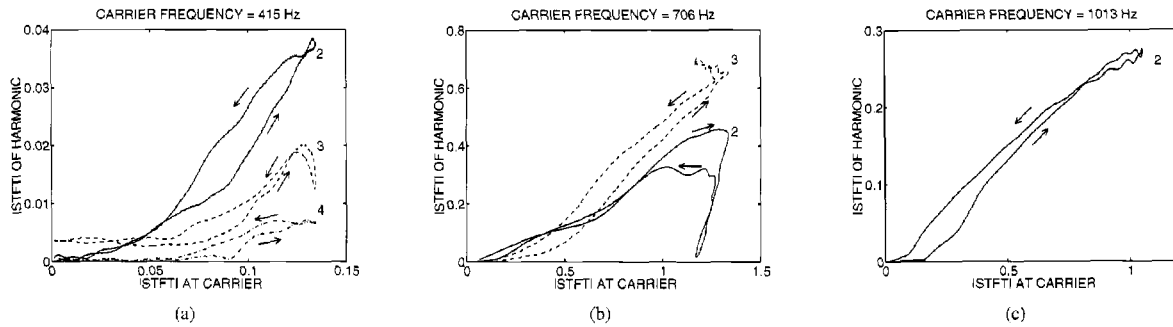


Fig. 12. Relationship between the STFT magnitudes of harmonics of the response and that at the carrier frequency  $f_c$  for the velocity responses below, near, and above CF. Arrows indicate the direction of increasing time: (a) STFT magnitude of the second, third, and fourth harmonics (denoted 2, 3, and 4, respectively) as a function of the STFT magnitude at  $f_c$ , for  $f_c = 415$  Hz (data obtained from Fig. 2(d)); (b) STFT magnitude of the second and third harmonics as a function of the STFT magnitude at  $f_c$ , for  $f_c = 706$  Hz (data obtained from Fig. 3(d)); (c) STFT magnitude of the second harmonic as a function of the STFT magnitude at  $f_c$ , for  $f_c = 1013$  Hz (data obtained from Fig. 4(d)).

we consider elsewhere the use of more general wavelet bases in analyzing this class of data [33]–[35].

For transient responses, the scalogram is generally a superior tool to the spectrogram, although in the examples we have illustrated, they are about equally useful. The response to the step turn-on of a tone allows us to generate a scalogram that acts as a ‘signature’ for the system under investigation. We have also used the scalogram as a visual tool for comparing various model responses. We conducted a preliminary investigation comparing the scalograms of transient responses from different driven oscillator models with the scalograms of the observed transients. We have found that the damped linear oscillator produces scalograms that are markedly different from those we have observed experimentally. The behavior of a damped nonlinear oscillator, such as the modified Duffing model, also differs from our measured results but matches them far more closely than the damped linear oscillator, particularly at frequencies above the CF.

Analysis of the AM velocity responses with the STFT has allowed us to both confirm and extend earlier findings based on the FFT analysis of pure-tone responses [2], as detailed below. These observations are based on 60 AM data set taken from outer hair cells and Hensen’s cells in both the third and fourth turns of the guinea pig cochlea.

- For carrier frequencies below the CF, multiple harmonic components are present in the response waveform [2], [9]–[12], [33]–[35]. In both third- and fourth-turn measurements, the component at the stimulus frequency usually has the largest peak magnitude; however, in many data sets, harmonic components above the carrier frequency achieve higher peak magnitudes.
- For carrier frequencies near the CF, the spectrum is once again dominated by the component at the stimulus frequency, and multiple harmonic components are present. We previously showed that in the fourth turn, for pure-tone stimuli at sound pressure levels of the order of 100 dB, even harmonic components are usually greater than their neighboring odd harmonics [36]. The STFT analysis of our AM responses augments this earlier result. At lower sound pressure levels (where the AM envelope

is small), even harmonics are indeed more pronounced, but at higher sound pressure levels, the odd harmonics become dominant (see Figs. 3(b)–(d), Fig. 2 in [11], and Figs. 2 and 5 in [12]). Additionally, in the third turn, we have seen numerous examples of the appearance of half-harmonic components (see, e.g., Fig. 3(b)–(d) and Fig. 5 in [12]) and quarter-harmonic components at the highest peak acoustic intensities for both outer hair cells and Hensen’s cells [33]–[35].

- For carrier frequencies above the CF, the spectral component at the carrier frequency is dominant for both third- and fourth-turn cells. Harmonic and half-harmonic components generally diminish as the carrier frequency increases relative to the CF.
- In some sensory cells, in both the third and fourth turns, low-level steady components can sometimes be discerned. These components arise from spontaneous cellular vibrations that occur both in the absence of a stimulus [12], [37], [38], and in its presence [25].

In addition to confirming earlier pure-tone results, the STFT offers a dramatic picture of both the time- and level-dependent dynamics of these cells, which remain hidden when using more conventional FFT-based analyses. Fig. 3(b)–(d) provide clear examples of the power of the STFT in illustrating the cellular dynamics. The magnitudes of the harmonic components are seen to vary in complex ways with time. This cannot be discerned with FFT analysis, which averages over the time course of the response.

The data in Figs. 2(d), 3(d), and 4(d) can be converted into plots of STFT magnitude for various harmonics as a function of STFT magnitude for the carrier, as shown in Fig. 12(a)–(c), respectively. The magnitude of the harmonics for the same carrier level differs, depending on whether the AM envelope is increasing or decreasing, i.e., there is considerable hysteresis in the response. Furthermore, the STFT magnitude of a given harmonic, for a fixed value of the STFT magnitude at  $f_c$ , is dependent on the value of  $f_c$ . For example, for an STFT magnitude of 1 at  $f_c$ , the second-harmonic magnitude is slightly higher for  $f_c = 706$  Hz (Fig. 12(b)) than for  $f_c = 1013$  Hz (Fig. 12(c)).

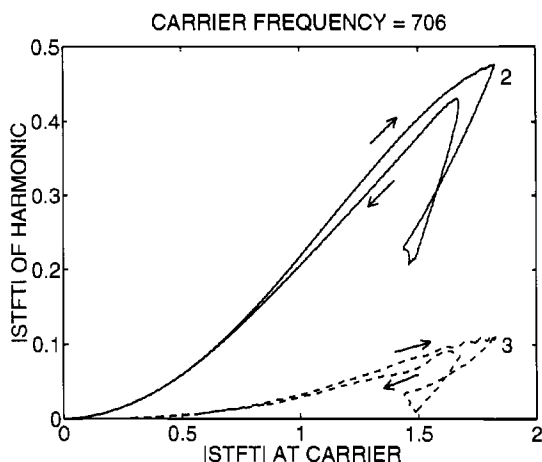


Fig. 13. STFT magnitude of the second and third harmonics as a function of the STFT magnitude at  $f_c$ , for the model response shown in Fig. 10(d) ( $f_c = 706$ ). The third harmonic curve lies below the second harmonic curve, unlike the cellular response shown in Fig. 12(b) ( $f_c = 706$  Hz).

The same sort of plot for the Duffing oscillator model shown in Fig. 13 differs from the cellular results in that the third harmonic curve lies below the second harmonic curve, and it exhibits hysteresis principally at high signal levels.

As a final indication of the value of the STFT, we reiterate that the four half-harmonic components illustrated in Fig. 3(d) do not appear symmetrically in time with respect to the peak of the envelope but rather appear just after the peak acoustic intensity has been reached. This shows the time-dependent nature of the response and rules out the possibility of a static, memoryless nonlinearity as a suitable model for understanding cellular nonlinear dynamics.

In short, time-frequency and time-scale techniques provide a powerful set of tools for observing the spectrum of the cellular velocity response as a function of time for different acoustic intensities and carrier frequencies. This in turn helps us achieve an increased understanding of cochlear dynamics and places constraints on the types of systems that can be used to model cochlear function.

#### ACKNOWLEDGMENT

The authors wish to thank B. Flock (Karolinska Institutet) for superb technical assistance in the data collection. They would also like to thank M. Vetterli and C. Herley for helpful suggestions in the preparation of the manuscript.

#### REFERENCES

- [1] G. von Békésy, *Experiments in Hearing*. New York: McGraw-Hill, 1960; Huntington, NY: Krieger, 1980.
- [2] International Team for Ear Research (ITER), *Cellular Vibration and Motility in the Organ of Corti*, *Acta Otolaryngologica* (Stockholm), Supplement 467, pp. 1-279, 1989.
- [3] M. Ulfendahl, Å. Flock, and S. M. Khanna, "A temporal bone preparation for the study of cochlear micromechanics at the cellular level," *Hearing Res.*, vol. 40, pp. 55-64, 1989.
- [4] W. E. Brownell, C. R. Bader, D. Bertrand, and Y. de Ribaupierre, "Evoked mechanical responses of isolated cochlear outer hair cells," *Sci.*, vol. 227, pp. 194-196, 1985.
- [5] L. Brundin, Å. Flock, and B. Canlon, "Sound-induced motility of isolated cochlear outer hair cells is frequency-specific," *Nature*, vol. 342, pp. 814-816, 1989.
- [6] L. Brundin, Å. Flock, S. M. Khanna, and M. Ulfendahl, "The tuned displacement response of the hearing organ is generated by the outer hair cells," *Neurosci.*, vol. 49, pp. 607-616, 1992.
- [7] ———, "Frequency-specific position shift in the guinea pig organ of Corti," *Neurosci. Lett.*, vol. 128, pp. 77-80, 1991.
- [8] B. Canlon, L. Brundin, and Å. Flock, "Acoustic stimulation causes tonotopic alterations in the length of outer hair cells from the guinea pig hearing organ," *Proc. Nat. Acad. Sci. (USA)*, vol. 85, pp. 7033-7035, 1988.
- [9] M. C. Teich, S. M. Khanna, and S. E. Keilson, "Nonlinear dynamics of cellular vibrations in the organ of Corti," *Acta Otolaryngologica* (Stockholm), Supplement 467, pp. 265-279, 1989.
- [10] C. Heneghan *et al.*, "Time-frequency representations of the dynamical motion of cellular structures in the organ of Corti," in *Abstracts Sixteenth Midwinter Res. Mtg. Assoc. Res. Otolaryngology*, Feb. 7-11, 1993, p. 85, Abstract No. 338.
- [11] M. C. Teich *et al.*, "Analysis of dynamical motion of sensory cells in the organ of Corti using the spectrogram," in *Biophysics of Hair Cell Sensory Systems* (H. Duifhuis, J. W. Horst, P. van Dijk and S. van Netten, Eds.). Singapore: World Scientific, 1993, pp. 272-279.
- [12] C. Heneghan, M. C. Teich, S. M. Khanna, and M. Ulfendahl, "Nonlinear dynamical motion of cellular structures in the cochlea," *Proc. SPIE*, vol. 2036, pp. 183-197, 1993.
- [13] M. C. Teich, C. Heneghan, S. M. Khanna, and M. Ulfendahl, "Characterizing the transient response of cellular structures in the organ of Corti using the wavelet transform," in *Abstracts Sixteenth Midwinter Res. Mtg. Assoc. Res. Otolaryngology*, Feb. 7-11, 1993, p. 84, Abstract No. 335.
- [14] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989, pp. 713-726.
- [15] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Mag.*, vol. 9, no. 2, pp. 21-67, Apr. 1992.
- [16] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, no. 4, pp. 14-38, Oct. 1991.
- [17] A. Grossmann, R. Kronland-Martinet, and J. Morlet, "Reading and understanding continuous wavelet transforms," in *Wavelets: Time-Frequency Methods and Phase Space* (J. M. Combes, A. Grossmann and Ph. Tchamitchian, Eds.). New York: Springer-Verlag, 1989/1990, pp. 2-20, 2nd ed.
- [18] M. C. Teich *et al.*, "Chaos in the cochlea," in *Abstracts Fourteenth Midwinter Res. Mtg. Assoc. Res. Otolaryngology*, Feb. 3-7, 1991, p. 50, Abstract No. 155.
- [19] M. C. Teich *et al.*, "Chaotic vibrations of outer hair cells and Hensen's cells in the cochlea," in *Abstracts Fifteenth Midwinter Res. Mtg. Assoc. Res. Otolaryngology*, Feb. 2-6, 1992, p. 17, Abstract No. 41.
- [20] H. Davis, B. E. Gernandt, and W. P. Covell, "Aural microphonics in the cochlea of the guinea pig," *J. Acoust. Soc. Amer.*, vol. 21, pp. 502-510, 1949.
- [21] P. J. Dallos and C. O. Linnell, "Subharmonic components in cochlear-microphonic potentials," *J. Acoust. Soc. Amer.*, vol. 40, pp. 4-11, 1966.
- [22] ———, "Even-order subharmonics in the peripheral auditory system," *J. Acoust. Soc. Amer.*, vol. 40, pp. 561-564, 1966.
- [23] P. J. Dallos, "On the generation of odd-fractional harmonics," *J. Acoust. Soc. Amer.*, vol. 40, pp. 1381-1391, 1966.
- [24] S. M. Khanna, S. E. Keilson, M. C. Teich, and M. Ulfendahl, "Transient vibrations of cellular structures in the organ of Corti," in *Abstracts Fifteenth Midwinter Res. Mtg. Assoc. Res. Otolaryngology*, Feb. 2-6, 1992, p. 18, Abstract No. 45.
- [25] S. E. Keilson, 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 Sixteenth Midwinter Res. Mtg. Assoc. Res. Otolaryngology*, Feb. 7-11, 1993, p. 84, Abstract No. 336.
- [26] S. E. Keilson, M. C. Teich, and S. M. Khanna, "Models of nonlinear vibration. I. Oscillator with bilinear resistance," *Acta Otolaryngologica* (Stockholm), Supplement 467, pp. 241-248, 1989.
- [27] M. C. Teich, S. E. Keilson, and S. M. Khanna, "Models of nonlinear vibration. II. Oscillator with bilinear stiffness," *Acta Otolaryngologica* (Stockholm), Supplement 467, pp. 249-256, 1989.
- [28] S. E. Keilson, M. C. Teich, and S. M. Khanna, "Models of nonlinear vibration. III. Oscillator with bilinear mass," *Acta Otolaryngologica* (Stockholm), Supplement 467, pp. 257-264, 1989.
- [29] A. H. Nayfeh and D. T. Mook, *Nonlinear Oscillations*. New York: Wiley-Interscience, 1979, pp. 161-224.



- [30] S. De Souza-Machado, R. W. Rollins, D. T. Jacobs, and J. L. Hartman, "Studying chaotic systems using microcomputer simulations and Lyapunov exponents," *Amer. J. Phys.*, vol. 58, pp. 321-329, 1990.
- [31] J.-C. Bacri, U. d'Ortona, and D. Salin, "Magnetic-fluid oscillator: observation of nonlinear period doubling," *Phys. Rev. Lett.*, vol. 67, pp. 50-53, 1991.
- [32] P. Bergé, Y. Pomeau, and C. Vidal, *Order Within Chaos*. Paris: Hermann, 1984; New York: Wiley-Interscience, 1986, pp. 210-220.
- [33] C. Heneghan, M. C. Teich, S. M. Khanna, and M. Ulfendahl, "Analysis of nonlinear cellular dynamics in the cochlea using the continuous wavelet transform and the short-time Fourier transform," in *Proc. IEEE-Signal Processing Int. Symp. Time-Frequency Time-Scale Analysis*, Oct. 25-28, 1994.
- [34] M. C. Teich, C. Heneghan, S. M. Khanna, and M. Ulfendahl, "Investigating cellular vibrations in the cochlea using the continuous wavelet transform and the short-time Fourier transform," in *Proc. Sixteenth Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Nov. 1-6, 1994.
- [35] M. C. Teich *et al.*, "Investigating routes to chaos in the inner ear using the continuous wavelet transform and the short-time Fourier transform," to be published in *Ann. Biomed. Eng.*
- [36] S. M. Khanna, M. Ulfendahl, and Å. Flock, "Waveforms and spectra of cellular vibrations in the organ of Corti," *Acta Otolaryngologica* (Stockholm), Supplement 467, pp. 189-193, 1989.
- [37] S. E. Keilson, S. M. Khanna, M. Ulfendahl, and M. C. Teich, "Spontaneous cellular vibrations in the guinea-pig cochlea," *Acta Otolaryngologica* (Stockholm), vol. 113, pp. 591-597, 1993.
- [38] S. M. Khanna, S. E. Keilson, M. Ulfendahl, and M. C. Teich, "Spontaneous cellular vibrations in the guinea-pig temporal-bone preparation," *Brit. J. Audiology*, vol. 27, pp. 79-83, 1993.



**Conor Heneghan** (S'93) received the B.E. degree from University College Dublin, Ireland, in electronic engineering in 1990. He received the M.S. degree in electrical engineering from Columbia University, New York, NY, in 1992.

His research interests include signal processing, nonlinear dynamics, and information collection and processing in the auditory system. He is presently a candidate for the Ph.D. degree at Columbia University.

Mr. Heneghan was the recipient of the 1990 Traveling Bursary Award from the National University of Ireland and the 1991 Millman Award from Columbia University.



**Shyam M. Khanna** (M'57-SM'74) was born in Agra, India, on May 10, 1932. He received the B.S. degree from Lucknow University in 1951 and the Ph.D. degree from the City University of New York in 1970.

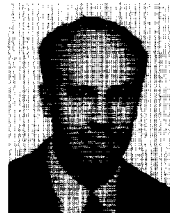
From 1953 to 1955, while at Pye Ltd. in Cambridge, England, he developed test instruments. From 1955 to 1958, he designed "line-of-sight" and "scatter" communication systems at Canadian Westinghouse Ltd., Hamilton, Canada. From 1958 to 1961, as a systems development engineer with the ITT Corporation, Nutley, NJ, he coined and developed an airborne Doppler-radar navigation system. From 1961 to 1964, he was with the IBM Corporation, investigating communication and coding in the auditory system. In 1964, he joined Columbia University's College of Physicians and Surgeons in New York City, where he is Professor of Auditory Biophysics. He has been Director of the Fowler Memorial Laboratory of Hearing Research since 1983. Starting in 1968, he began to use homodyne interferometry and time-averaged holography to study the mechanical vibrations in the ear. Since 1983, he has led an international team of scientists to develop and use sensitive optical instruments for the noninvasive visualization of cellular structures and the measurement of cellular vibrations at the subatomic level in the inner ear. His research interests include the function of the external, middle, and inner ear, as well as the auditory nerve. He is the author of more than 125 research papers and coeditor of a recent book on the use of optical techniques in the life sciences.

Dr. Khanna is a Fellow of the Acoustical Society of America (ASA) and President of the International Society on Optics Within Life Sciences (OWLS).



**Åke Flock** was born in Gothenburg, Sweden. He received the M.D. degree in 1963 and defended his Ph.D. thesis in 1965 at the Karolinska Institute Medical School, Stockholm. He became Associate Professor of Experimental Otolaryngology in 1967 at the same institute. He is now Professor and Chairman of the Department of Physiology and Pharmacology, Karolinska Institute. He works with mechano-electrical transduction in the organ of hearing, in particular, with the notion that sensory cells are motile and contribute actively to the motion of the organ. He worked as a postdoctoral student at Bell Telephone Laboratories from 1965 to 1967 and has been a Visiting Professor at Harvard Medical School; Baylor College of Medicine; Marine Biological Laboratories, Woods Hole; University of Hawaii; University of Sussex, Great Britain; Max Planck-Institute, Göttingen, Germany; and Columbia University, New York. He has authored or co-authored some 175 scientific publications.

Among several missions for the faculty, Dr. Flock is a member of the Board of the Nobel Assembly deciding on the Nobel Prize in Medicine. He has received nine research prizes, including the Prize in Medicine awarded by the Royal Swedish Academy of Sciences in 1981, the Fernström Research Award of the Karolinska Institute in 1982, the Reese Award for Hearing in the USA in 1986 when it was first given (for excellence in hearing research during the last ten years), the Caliceti Prize in Bologna, Italy, in 1989 (for last four years of best research), and the 1989 Award of Merit from the American Association for Research in Otolaryngology (which was the first time it was given to a non-American). He is a fellow of the Bärány Society, the Collegium Otolaryngologicum, The Association for Research in Otolaryngology, the Inner Ear Biology Group, the International Union of Physiology, the Scandinavian Society for Electron Microscopy, the Swedish Physiological Society, and the Swedish Society for Physicians. He has done *ad hoc* reviewing for the National Institutes of Health, USA, the National Science Foundation, USA, and the Royal Society, Great Britain. He has served as a member of the Editorial Board for *Hearing Research* since it started in 1979.



**Mats Ulfendahl** was born in Stockholm, Sweden. He received the B.M. degree from the Karolinska Institute Medical School, Stockholm, in 1983, and the Ph.D. degree in physiology from the Karolinska Institute in 1989.

Since receiving the doctorate, he has been collaborating with Shyam Khanna on vibration measurements in the hearing organ, both at Columbia University in New York and at the Karolinska Institute in Stockholm. In 1991, he became Associate Professor in the Department of Physiology and Pharmacology at the Karolinska Institute. As a member of the faculty, he is now teaching and is involved in a number of research projects in the area of auditory physiology.



**Lou Brundin** was born in Stockholm, Sweden. She received the M.D. degree in 1987 and defended her Ph.D. thesis in physiology in 1991 at the Karolinska Institute, Stockholm.

Her work concerns the participation of sensory cells of the hearing organ in frequency-specific functions: as mechanical sensors, generators of mechanical displacement, and producers of electrical receptor potentials. She has been a member of the teaching faculty at the Department of Physiology and Pharmacology, Karolinska Institute. Working with Professor Shyam Khanna with laser heterodyne interferometry, she has been appointed Associate Research Scientist at Columbia University, New York, NY. During 1992 and 1993, she worked as a postdoctoral researcher with Professor Ian Russell at the University of Sussex, England, supported by fellowships from, among others, the European Science Foundation, the Karolinska Institute, and the Swedish Society for Otolaryngology.

Dr. Brundin has been the coauthor of numerous scientific publications, one of which was awarded the Alvarenga Prize by the Swedish Association for Physicians for the best scientific paper in the field of Medicine for the year 1990. She has presented oral and poster contributions at national and international meetings. She is a member of the Swedish Association for Physicians, the Scandinavian Physiological Society, the Swedish Otolaryngological Society, and the Inner Ear Biology Group.



**Malvin C. Teich** (S'62-M'66-SM'72-F'89) was born in New York City. He received the S.B. degree in physics from the Massachusetts Institute of Technology, Cambridge, in 1961, the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 1962, and the Ph.D. degree in quantum electronics from Cornell University, Ithaca, NY, in 1966.

After receiving the doctorate, he joined MIT Lincoln Laboratory, Lexington, MA, where he was engaged in work on coherent infrared detection. In 1967, he became a member of the faculty in the Department of Electrical Engineering, Columbia University, New York, NY, where he is now teaching and pursuing his research interests in the areas of quantum optics, photonics, fractal point processes in physical and biological systems, and information transmission in biological sensory systems. He served as Chairman of the Department from 1978 to 1980. He is also a member of the faculty in the Department of Applied Physics and a member of the Columbia Radiation Laboratory. He has authored or coauthored some 200 technical publications and holds one patent. He is the coauthor of *Fundamentals of Photonics* (Wiley, 1991).

Dr. Teich was the recipient of the IEEE Browder J. Thompson Memorial Prize Award for his paper "Infrared Heterodyne Detection" in 1969, and in 1981, he received the Citation Classic Award of the Institute for Scientific Information for this work. He was awarded a Guggenheim Fellowship in 1973. In 1992, he was honored with the Memorial Gold Medal of Palacký University in the Czech Republic. He is a Fellow of the American Physical Society, the Optical Society of America, the American Association for the Advancement of Science, and the Acoustical Society of America. He is a member of Sigma Xi, Tau Beta Pi, the Association for Research in Otolaryngology, and the New York Academy of Sciences. He served as a member of the Editorial Advisory Panel for the journal *Optics Letters* from 1977 to 1979 and as a Member of the Editorial Board of the *Journal of Visual Communication and Image Representation* from 1989 to 1992. He is currently Deputy Editor of *Quantum Optics: Journal of the European Optical Society B*, a Member of the Editorial Board of the journal *Jemná Mechanika a Optika*, and a Member of the Scientific Board of the Czech Academy of Sciences' Institute of Physics.