Accurate Tuning Curves in a Cochlear Model

James M. Kates

Abstract—Most practical models of cochlear mechanics are based on approximations to the wave equation in the cochlea. These approximations engender compromises in the accuracy with which the cochlear motion can be reproduced. In this paper, tuning curves are compared for two cochlear models, one based on a cascade of low-pass filter sections and the other based on a cascade of filter sections derived from a one-dimensional transmission line. The filters in the two simulations are designed to give comparable latency in the neural response to inputs at different frequencies, and the simulations include an active gain-control mechanism to adjust the characteristics of each section with changes in the input signal level. The resultant simulated tuning curves and neural outputs indicate that the modified transmission-line approximation yields a more accurate cochlear model.

I. INTRODUCTION

In a recent paper, Kates [1] described a cochlear model based on a cascade of digital low-pass filter sections. This model offered the advantages of computational efficiency and numerical stability. The model also incorporated a compressive nonlinearity to reproduce the changes in cochlear filter shape and gain with variations in signal level, and included the ability to simulate the changes in cochlear behavior due to outer hair-cell and inner hair-cell damage representative of hearing loss. However, the auditory filters and associated tuning curves in the model were much broader than those measured physiologically in animals (Kiang and Moxon [2]; Kiang [3]) or psychophysically in humans (Nelson and Freyman [4]; Nelson and Fortune [5]). These broad auditory filters were the consequence of adjusting the cascaded low-pass filter sections in the model to have group delays comparable to those found in the human cochlea. The objective of the work presented in this paper is to provide more accurate tuning curves while preserving the advantages of the existing model. Results are compared for two cochlear models, one based on the cascade of low-pass filter sections and the other based on a cascade of filters derived from a one-dimensional transmission line, with both systems adjusted to have group-delay characteristics similar to the human cochlea.

A set of neural tuning curves for the cat are shown in Fig. 1 (Kiang [3]). The corresponding curves for the human ear can be approximated by dividing the indicated frequencies by a factor of three. It is these tuning curves that the cochlear model should reproduce. The curves are characterized by a narrow but rounded peak at the characteristic frequency, a very steep high-frequency slope, a response zero about one octave below the peak, and a flat tail extending to the low-frequency cutoff of the middle-ear transfer function. The sensitivity of the tail is lowest at the highest frequencies and increases as the characteristic frequency decreases. The amplitude ratio from tip to tail is approximately 60 dB at the highest frequencies and decreases as the characteristic frequency decreases. Thus the slope of the low-frequency portion of the tuning curve near the tip is determined by the drop in sensitivity from tip to tail over the approximate octave frequency separation from the tip to the zero.

The approach to modeling waves propagating in the forward direction in the cochlea used by Kates [1] is a cascade of active low-pass filter sections (Lyon and Mead [6]). The output at a specific location on the cochlear partition is the cascade of all the filter sections that precede it. Thus, the justification for this approach is provided by Lyon and Mead [7] in their solution of the two-dimensional wave equation in the cochlea. The active low-pass filters provide unity gain at low frequencies, gain greater than one in the vicinity of the resonance frequency, and attenuation for higher frequencies. None of the active filter sections is highly tuned; rather, the system achieves its overall tuning sharpness through the buildup of the contributions of many sections in cascade in an effect termed pseudoresonance (Holmes and Cole [8]). The active system also adjusts the quality factor (Q) of each filter in response to the output signal level, thereby providing a mechanism for the hypothesized outer hair-cell behavior (Ashmore [9]) in changing the cochlear filter shape with signal level (Møller [10]; Sellick et al. [11]) and compressing the output signal dynamic range (Robles et al. [12]; Sellick et al. [11]).
The analog two-pole low-pass filter implementation of Lyon and Mead [6] provides cochlear filters and tuning curves having shapes similar to those measured in the ear. However, the filter group delay was found by Kates [1] to be excessive when simulating the human cochlea. Kates reduced the group delay by adding zeroes to the filter transfer function. This resulted in simulated cochlear filters at threshold and tuning curves that are much broader than those shown in Fig. 1, and which have high-frequency slopes that are less than half those measured in mammalian cochleas. The tuning curves from this model are reproduced in Fig. 2. Thus the low-pass filter model does not appear to adequately match the cochlea; too much group delay is generated for a given cochlear filter bandwidth.

An alternative approach to modeling wave propagation in the cochlea is to use a lumped-parameter transmission-line model based on a one-dimensional low-frequency approximation to the cochlear wave equation (Peterson and Bogert [13]; Zwilocki [14]; Schroeder [15]; Zweig et al. [16]; Zwicker [17]; Deng and Geisler [18]). The transmission-line approach has also been extended to model an active cochlea that includes the effects of micromechanical resonances and outer-hair cell feedback (Neely and Kim [19]; Geisler [20]) in generating additional frequency-response behavior such as the "second filter" (Allen [21]). The transfer function for one section of the transmission line includes zeroes as well as poles in the inherent structure, so the group-delay behavior will differ from that of the low-pass filter approximation.

Examples of tuning curves from a nonlinear one-dimensional transmission-line cochlear model (Deng and Geisler [18]) are presented in Fig. 3. As opposed to the tuning curves of Fig. 2, which were too broad, the tuning curves of Fig. 3 have tips that are too narrow and which lack the rounding shown in Fig. 1. There is no second filter in this model, so the tuning curves lack the response zero one octave below the characteristic frequency. The low-frequency slope in the tail region is 6 dB per octave, rather than the flat tail shown in Fig. 1, since the basilar-membrane velocity is taken as the mechanical model output. Using the mechanical displacement rather than the velocity yields mechanical tuning curves that are closer in shape to neural tuning curves (Robles et al. [22]; Neely [23]), but this is not done in the Deng and Geisler [18] model nor in many other transmission-line cochlear models since it would result in inadequate low-frequency attenuation in the tuning curves. Transmission-line models also typically require a large number of sections and small time steps in the simulation, with the resultant large computational burden. Deng and Geisler [18] use 1400 sections and a 100-kHz sampling rate, as opposed to the 112 sections and 40-kHz sampling rate used by Kates [1]; reducing the number of sections in the transmission-line model can, however, lead to ripples and other instabilities in the response for low-level signals where the nonlinear system has its maximum gain (Zwicker [17]).

It appears that neither the cascade of low-pass filters nor the one-dimensional transmission line can meet the objectives of a computationally efficient model having accurate tuning curves, suitable for applications such as hearing-aid design and evaluation. A compromise approach, however, is feasible. Replace the low-pass filters in the cascade with filters that approximate the transfer function of a section of one-dimensional transmission line, driven by the pressure output from the previous section and terminated with the characteristic impedance of the following section. This approach yields a modified transmission line, with each section isolated from its neighbors rather than interacting directly with them. Certain features of cochlear behavior, such as otoacoustic emissions, can not be reproduced by this modeling approach. As in the case of the low-pass filter model, the ear is still represented phenomenologically, so there is also a limitation as to what can be deduced about the behavior of the underlying physical system. However, useful outputs can be efficiently generated to simulate normal and impaired hearing for a variety of stimuli.

The purpose of this paper is to compare the simulated cochlear behavior for the low-pass (LP) filter approximation to that of the modified transmission-line (MTL) approximation when both are adjusted to have group-delay characteristics similar to the human cochlea (Eggermont [24]). Both the LP
and MTL simulations in this paper consist of a cascade of filter sections with feedback from the cochlear motion used to adjust the filter $Q$ values, and only wave propagation in the forward direction is represented. The output of each filter section in either model is modified by a second filter (Allen [21]), also having an adaptive $Q$ value, that provides a notch in the frequency response an octave below the filter characteristic frequency and an increase in gain at the characteristic frequency (Kates [1]). The inner hair-cell neural firing behavior does not affect the feedback for the cochlear mechanics (Liberman [25]), so both models have been adjusted so that the filter control signal is derived exclusively from the motion of the cochlear partition. The inner hair-cell transduction model of Kates [1] is used for both simulations when neural firing rates are desired.

The LP cochlear model briefly described in the next section of the paper is identical to that of Kates [1], and the reader is referred to that paper for a more detailed description. The MTL cochlear model is described in the following section of the paper. The model exposition is followed by a comparison of auditory filter shapes, tuning curves, and impulse responses, as well as simulated neural responses to a synthetic speech stimulus. In all examples, the MTL model gives more accurate results.

II. LOW-PASS FILTER MODEL

The low-pass filter cochlear simulation is the model developed by Kates [1]. The model consists of a cascade of low-pass filter sections transformed from the analog to the digital domain via the bilinear transformation, with each prototype analog filter section consisting of two zeroes and three poles as given below:

$$H_i(s) = \frac{1 + (\mu + 1/Q_i)(s/\omega_i) + b(\mu/Q_i)(s/\omega_i)^2}{1 + s/\omega_i} \frac{1 + s/\omega_i}{1 + s/\omega_i}$$

The maximum filter $Q$ values increase with increasing frequency, ranging from 0.28 at 100 Hz to 0.45 at 10 kHz on a linear cochlear distance scale. All of the zeroes and poles in the filters are real, giving rise to broad filter gain-versus-frequency characteristics. This filter design represented a compromise between the desired frequency response and group delay, with the sharpness of the filters reduced in order to get latency comparable to that measured in the human cochlea. The net traveling-wave motion at a given location on the cochlear partition is given by

$$G_k(z) = \prod_{i=1}^{k} H_i(z)$$

where $H_i(z)$ is the transfer function of each filter in the cascade. This output is then modified by a one-pole high-pass filter, having a cutoff frequency two octaves below the section characteristic frequency, that provides the velocity transformation. A total of 112 LP filter sections are used to model the cochlea from 16 kHz to 100 Hz.

The second filter, consisting of a complex zero pair over a complex pole pair, operates on the velocity to give the cochlea mechanical output. The analog second filter is given by

$$F(s) = \frac{1 + s/\omega_0 Q_0 + (s/\omega_0)^2}{1 + s/\omega_p Q_p + (s/\omega_p)^2}.$$  (3)

For the second filter, $Q_0 = 2Q_p$ and $\omega_0 = \omega_p/2$, thus giving zero group delay at low frequencies while providing a peak at the filter characteristic frequency and a notch an octave below. This filter was also transformed into the digital domain via the bilinear transformation. The maximum $Q$ of the second filter is also frequency-dependent, varying according to the relationship $Q_p = 1.5(1 + f)$ for $f$ in kHz.

The feedback system adjusts the $Q$ values of both the low-pass filters and the second filters in response to the signal level. The simulated cochlear amplifier gives a maximum of 60 dB of gain for an input at or below threshold, and this gain is reduced to 0 dB for inputs at or above 100 dB SPL. Thus for inputs between 0 and 100 dB SPL, the outputs range between 60 and 100 dB SPL, providing a compression ratio of about 2.5:1. The control signal is the peak velocity output from the second filter over a narrow region to either side of the given frequency tap on the simulated cochlea, and the traveling-wave and second-filter filter $Q$ values are reduced linearly from 100 percent to 10 percent of their maximum values as the output increases in dB. The $Q$ scale factor is thus $r = (100 - y)/(100 - 60)$, where $y$ is the output in dB and $r$ is constrained to lie within the range $0 \leq r \leq 1$. The $Q$ value is given by $Q = (0.1 + 0.9r)Q_{max}$. The $Q$ adjustment has an instantaneous attack, and the release time is set to 100 ms.

III. MODIFIED TRANSMISSION-LINE MODEL

The modified transmission-line model also uses the filter cascade, but with the filter transfer functions now chosen to represent isolated sections of a one-dimensional transmission line. A single section of the equivalent analog circuit is shown in Fig. 4; the transmission-line section (Peterson and Bogert [13]; Zwischenolzki [14]; Schroeder [15]) has been isolated from its neighbors and loaded by the input impedance of the adjacent section (Eysholdt and Melfatt [26]; Ambikairajah et al. [27]). Eysholdt and Melfatt [26] demonstrated that replacing the input impedance of the adjacent section with an approximate characteristic impedance has a minimal effect on the transfer function from the base to the apex of the cochlea, and they derived a transfer function having one pair each of relatively high-$Q$ poles and zeroes combined with
a pair each of low-Q poles and zeroes. The high-Q poles and zeroes will dominate the transfer function behavior in the vicinity of the section characteristic frequency, so a reasonable simplification is just a zero pair over a pole pair, as in (3) for the second filter, transformed into the digital domain via the bilinear transformation. The relationship of the zero and pole parameters, however, must be adjusted to give behavior characteristic of the transmission-line transfer function, and this results in setting \( \omega_0 = 1.15\omega_0 \) and \( Q_0 = 1.50Q_p \) to give a peak followed at a higher frequency by a notch.

The modified transmission-line implementation, like the low-pass filter version, is an active system, with adjustments to the filter \( Q \) values changing the filter shapes and gains. The same second filter functions with adaptive \( Q \) values and frequency dependence are used in the MTL version as for the LP version. The MTL output is taken as the pressure output shown in Fig. 4 passed through the second filter, so the pressure-to-velocity transformation provided in the LP model has been eliminated in the MTL model. The resultant adaptive system is shown in the block diagram of Fig. 5, where the peak displacement over a narrow frequency region is used to control the MTL section filters and the second filters. The maximum filter \( Q_p \) value for the MTL section ranges from 0.5 at the low-frequency end of the simulated cochlea to 8.0 at the high-frequency end on a linear cochlear distance scale.

The variation of \( Q \) with output signal level has also been changed from that used in the LP model in response to the change in the cochlear filter shapes. The MTL rule for adjusting the \( Q \) values with signal level, determined empirically, uses a scale factor \( r = 0.25d + 0.75d^3 \), where \( d = (120 - y)/60 \) and \( y \) is the cochlear mechanical output in dB. The factor \( r \) is constrained to lie between 0 and 1. The MTL section filter \( Q_p \) is given by \( Q_p = (0.3 + 0.7r)Q_{\text{max}} \), and the second filter \( Q_p \) is given by \( Q_p = (0.1 + 0.9r)Q_{\text{max}} \). This functional variation of \( Q_p \) with level gives a nearly uniform 2.5:1 compression ratio in the cochlear output for inputs ranging from 0 to 100 dB SPL, as shown in Fig. 6 for the 1-kHz tap. This behavior, with compression occurring over a broad range of signal levels, is consistent with physiological measurements of cochlear motion (Sellick et al. [11]). The mechanical output in the figure is relative to that produced by an input at 0 dB SPL; add 60 dB to get the corresponding dB SPL at output. The same instantaneous attack and 100-ms release times are used in the MTL model as were used in the LP model.

The MTL rule for adaptively adjusting the filter \( Q \) is the result of two effects: the excitation pattern in the cochlea and the dependence of the overall gain on the transfer function of each section. The net gain at the place of maximum response for a sinusoidal input is the product of the gains in the filter cascade for approximately twenty filter sections just higher in frequency. At auditory threshold, every filter section is set to maximum \( Q \) and hence provides maximum gain. If the signal level is increased to be slightly above threshold, only the filters very close to the characteristic place have outputs intense enough to reduce the filter \( Q \) values, so the change in \( Q \) with an increase in the signal level (proportional to \( \frac{dx}{dy} \)) is at a high rate in order to provide the 2.5:1 compression ratio. As the signal level increases, however, a greater number of filters will begin to have outputs above threshold, and the \( Q \) in each of these filters will begin to be adaptively reduced as well. More sections being affected means that a smaller change in \( Q \) is needed in each one to provide the same total amount of compression at the place of maximum response. The rule for adjusting the filter \( Q \) takes this into account by decreasing the rate of reduction in the filter \( Q \) values as the signal level, and with it the region of upward spread of excitation, is increased. By the time high signal levels are reached, all the sections that can affect the gain at the place or maximum response are being excited. The overall gain is no longer being affected by the increase in the spread of excitation, so the rate of reduction of filter \( Q \) with the increase in output level is set to be much lower at high signal levels than at low signal levels to give the same compression ratio.

An additional feature of the MTL model concerns the isolation amplifier between the modified transmission-line sections shown in Fig. 4. The data of Kiang and Moxon [2], and of Kiang [3] reproduced in Fig. 1, show extended tails of the tuning curves, which are consistent with displacement output in the cochlear model (Robles et al. [22]). The tuning curves also show tip-to-tail ratios that are greater at high frequencies than at low frequencies even though the tip thresholds do not vary significantly. In order to duplicate this latter effect, the signal input to the model at the base (high-frequency end) of
the cochlea is attenuated by 30 dB, and the isolation amplifier is used to provide a gain of 0.6 dB per section for the 120 sections used to model the cochlea from 16 kHz to 50 Hz. The increase in gain as frequency decreases compensates for the lower $Q$ values, holding the peak output levels nearly constant while reducing the tip-to-tail ratios in accordance with the physiological data.

IV. MODEL COMPARISONS

A. Filter Shapes and Tuning Curves

The LP and MTL modeling approaches produce very different auditory filters. Fig. 7 presents the auditory filters at the 1-kHz characteristic place for the maximum and minimum $Q$ values in the models. The LP model (dashed line) gives broad filters, having a $Q_{10}$ of about 1.5 for the maximum filter $Q$ values as compared to a $Q_{10}$ of about 3.5 for the MTL model (solid line). While the MTL model result is still lower than the typical $Q_{10}$ values of 5 to 10 obtained for human subjects for low-level simultaneously-masked psychophysical tuning curves (PTC's) at 1 kHz (Nelson and Fortune [5]), it is clearly much closer to the desired value than the LP model result. The high-frequency slopes over the first octave also differ substantially, with the LP model giving about 35 dB/oct and the MTL model giving about 85 dB/oct in comparison with human slopes in excess of 150 dB/oct for simultaneous masking (Nelson and Fortune [5]). The high-frequency slopes of the MTL model decrease with increasing frequency more than in measured tuning curves, and this is probably the result of using only a pair of poles in the simplified transfer function; an additional real pole would restore the slopes at high frequencies, but the filter sensitivity is already so low that this would have little practical effect. Thus the MTL model gives a much better overall approximation to the auditory filter.
shape at threshold than the LP model, although the actual cochlear filters are sharper still.

At the minimum filter Q values, Fig. 7 also shows important differences between the LP and MTL model filters. In particular, the high-frequency slope of the LP model filter is reduced, which runs counter to the masking data cited above (Nelson and Fortune [5]). The MTL model shows a parallel shift of the high-frequency edge of the filter as the signal level increases, which agrees with the human data. In addition, the MTL model shows a more dramatic downward shift of the peak frequency of the filter as the Q values are decreased, and this agrees more closely with measurements of cochlear motion with increases in signal level (Johnstone et al. [28]).

A comparison of tuning curves for the LP and MTL models is shown in Fig. 8. A tuning curve is produced by adjusting the input amplitude as the frequency is varied in order to maintain a constant output level. The tuning curves are shown inverted, relative to the physiological data of Fig. 1, in order to facilitate comparison with the auditory filters of Fig. 7. The tuning curves of Fig. 8 have steeper slopes at high frequencies than the auditory filters shown in Fig. 7 since increasing the signal level as the test frequency moves away from the characteristic frequency lowers the filter Q values, resulting in lower gain in the filters and requiring even higher signal levels to maintain the constant output. The MTL model gives a much sharper tuning curve, having a slope of about 135 dB/oct over the first octave, in comparison with a slope of about 55 dB/oct for the LP model. The \( Q_{10} \) for the MTL model tuning curve is about 4.5, which is still below the desired human value but which is a good deal closer than the \( Q_{10} \) value of 1.8 for the LP model. Thus the MTL model gives more accurate tuning curves than the LP model.

A set of tuning curves for the MTL model is presented in Fig. 9, and should be compared with the physiological tuning curves presented in Fig. 1. This comparison shows the effect of the Q variation with frequency and the gain increment for each MTL section in reproducing a family of tuning curves. Important features that match at least qualitatively in the measurements and simulation are the approximate sharpness of the tuning curves, the decrease of the \( Q_{10} \) of the tuning curves with decreasing frequency, the decrease in tip-to-tail ratio with decreasing frequency, and the low-frequency extent of the tails of the high-frequency tuning curves. The decrease in sensitivity in the tails of the tuning curves in both figures...
at low frequencies is due to the low-frequency attenuation of the middle ear transfer function, which has also been included in the cochlear model.

B. Impulse Response

Increasing the signal level causes a corresponding decrease in the filter Q values, resulting in decreased latency and decreased filter ringing. This effect was illustrated in the rat cochlea by Möller [10], who cross-correlated the neural response with the noise excitation as the noise level was increased from 37 to 77 dB SPL; the results for a fiber most sensitive to 3 kHz is shown in Fig. 10. The comparable output for the LP model was shown (Kates [1], Fig. 14) to agree qualitatively with the neural data.

Impulse response results for the MTL model are presented in Fig. 11. The output frequency of 1 kHz for the human simulation was chosen to correspond to the same distance along the cochlea as for the 3-kHz fiber monitored in the rat. The time scales should be adjusted in accordance with the frequencies in comparing the figures. The impulse responses were generated by first exciting the model with 45 ms of noise, followed by a 5-ms gap, and then an impulse at the same amplitude as the noise level. The noise burst causes the adaptive mechanical system to adjust the cochlear filter Q values under conditions similar to those used by Möller. The short gap allows the noise response to decay to nearly zero at mid and high frequencies, after which the impulse is used to extract the system response. Since the Q feedback adjustments have a 100-ms release time in the models (Kates [1]), there will be very little change in the system parameters during the gap or the impulse, so the impulse response is equivalent to that obtained using cross-correlation.

The agreement of the MTL model output with the rat data is quite good on a qualitative level. Increasing the signal level reduces the amount of ringing in the cochlear filter impulse response and reduces the time delay to the peak of the output waveform. The number of cycles of ringing are similar at each signal level in the rat and simulated human responses. The sharp cochlear filters in the MTL model have not resulted in excessive latency or ringing in comparison with the measured data, so the MTL model has an accurate temporal response along with the accurate frequency response.

C. Response to Synthesized Speech

The synthetic speech stimulus is the syllable /da/ digitally generated by a Klatt synthesizer (Klatt [29]). The first three formants of the vowel portion of the syllable are at 0.7, 1.2, and 2.4 kHz, respectively. The consonant portion of the syllable is represented by formant transitions that occur during the first 50 ms, with the starting frequencies of the first three formants being 0.5, 1.6, and 2.8 kHz, respectively. The formant frequencies are held constant during the second 50 ms of the stimulus, and the total stimulus duration is 100 ms. The fundamental frequency is 120 Hz during the first 50 ms of the stimulus and drops to 116 Hz during the second 50 ms. The average power in the signal is about 65 dB SPL at the onset of the stimulus and increases to about 69 dB SPL at the end. This synthetic speech stimulus was used by Miller and Sachs [30] to generate cat auditory-nerve firing patterns over a broad range of nerve-fiber center frequencies, and their data has also been analyzed by Shamma [31] and Secker-Walker and Searle [32].

The time-frequency response on the cat auditory nerve to the speech stimulus is shown in Fig. 12 (Shamma [31]). The responses have been averaged over all the fibers measured at
each frequency, so a large range of thresholds and spontaneous rates are included in the figure; each curve, however, is based on a small number of fibers and therefore represents a more limited dynamic range. The fibers can be divided into groups based on the phase-locked firing behavior (Shamma [31]; Secker-Walker and Searle [32]). Fibers having characteristic frequencies below about 300 Hz respond primarily to the fundamental frequency of the speech, giving a temporal indication of the glottal pulses. Fibers between 400 and 1200 Hz have peak separations that correspond to the period of the first formant, and fibers between 1200 and 2000 Hz have peak separations that correspond to the period of the second formant. Between 2000 and 3000 Hz there is some indication of firing synchronized to the third formant that appears to decrease for increasing time, and above 3000 Hz the response is again primarily to the glottal pulses.

The response of the LP model combined with the simulated inner hair-cell transduction (Kates [1]) is shown in Fig. 13 for the synthetic speech stimulus, and the response of the MTL model combined with the simulated inner hair-cell transduction is shown in Fig. 14 for the same stimulus. The two model responses are similar to the cat data in that phase-locked neural firing patterns occur over frequency regions corresponding to each of the formants. The boundary between the first and second formants occurs at about 1000 Hz, and the boundary between the second and third formants occurs at about 2000 Hz. Below 300 and above 3000 Hz the firing patterns represent the glottal pulses. Both models show very similar neural latency, as expected from the design constraints, with the time delay as a function of frequency in the peak neural response to the onset of the signal agreeing closely between Figs. 13 and 14.

The differences between the LP and MTL models stem from the differences in the shapes of the respective cochlear filters. The boundaries between the formant regions are much more clearly defined in Fig. 14, for the MTL model, than in Fig. 13, where the broader LP model filters allow a more gradual transition from one formant region to the next. The sharpness of the MTL model formant-region boundaries agrees more closely with the neural data of Fig. 12, which also shows very sharp demarcations of the neural fiber groups synchronized to each of the three formants in the signal. Thus the MTL model gives more accurate simulated neural firing patterns than the LP model, although both models appear to yield useful results.

V. CONCLUSIONS

A cascade of isolated filter sections gives a useful model of traveling-wave behavior in the cochlea. This modeling approach is not intended to replicate the physical details of the mechanical system, but rather seeks to duplicate the observed signal-processing characteristics for a signal propagating from base to apex. To this end, a simple filter section, comprising a zero pair over a pole pair, reproduces most of the salient transfer-function features observed in the motion of the cochlear partition. This modeling approach yields pseudoresonant response peaks in the motion of the cochlear partition, is readily combined with the second filter representing the cochlear micromechanics, is amenable to \( Q \) adjustment as a function of signal level to give adaptive filter shapes and gains, and is easily modified to model impaired hearing. A limitation in only modeling waves traveling in the forward direction is that some effects that rely on propagation in both directions within the cochlea, such as otoacoustic emissions, cannot be reproduced.

The details of the modeled cochlear behavior depend on the design of the filter sections. When the filters are adjusted to have similar group delays, the MTL model results in much
Fig. 12. Time-frequency analysis of the cat auditory nerve response evoked by the stimulus *ida*: (a) 0–25 ms, (b) 25–50 ms (from Shamma [31]).

A more accurate cochlear filter and tuning curves than the LP model. The simulated neural firing patterns are also more accurate for the MTL model. However, this difference in accuracy does not necessarily mean that one filter design approach is more valid than the other at all frequencies; rather, it indicates that given the simplified circuit implementations, the MTL model yields better results than the LP model when used to represent the entire auditory frequency range.

The choice of a modeling approach to use in an auditory simulation depends on the importance of computational efficiency as well as physical accuracy. If the computational requirements are of paramount importance and excessive latency at low frequencies is unimportant, then the two-pole low-pass filters used by Lyon and Mead [6] give the most efficient system. If overall accuracy, including latency, is more important and a slight increase in computational time can be tolerated, then the MTL model described in this paper should be used since it offers greater accuracy than the LP model for a similar computational burden. Even greater accuracy, as would be needed for simulating otoacoustic emissions, requires an appropriate physical model of the motion of the cochlear partition and the second filter rather than the phenomenological model presented in this paper, and would be accompanied by a corresponding increase in the computational burden. But for many applications, the MTL model presented in this paper represents a useful compromise between computational efficiency and physical accuracy.

REFERENCES


James M. Kates received the degrees of B.S.E.E. and M.S.E.E. from the Massachusetts Institute of Technology in 1971 and the professional degree of Electrical Engineer from M.I.T. in 1972. Since 1988 he has been with the Center for Research in Speech and Hearing Sciences of the City University of New York, where his research interests include digital signal processing for the design and evaluation of hearing aids and the study of normal and impaired auditory function. He also holds an appointment as an Adjunct Professor in the Doctoral Program in Speech and Hearing Sciences at CUNY. Previously he has worked on applied research in hearing aids (Siemens Hearing Instruments), signal processing for radar, speech, and hearing applications (SIGNATRON, Inc.), and loudspeaker design and signal processing for audio applications (Acoustic Research and CBS Laboratories).