Hearing Research, 37 (1989) 219–240
Elsevier

HRR 01162

Transient responses to tone bursts

E.R. Lewis [1] and K.R. Henry [2]

[1] *Dept. of EECS, University of California, Berkeley, California, U.S.A. and* [2] *Dept. of Psychology, University of California, Davis, California, U.S.A.*

(Received 6 May 1988; accepted 13 September 1988)

Investigating theoretical conditions under which linearly-operating tuned structures produce click-like transient responses to onsets and offsets of trapezoidal tone bursts, we come to the following conclusions: (1) each of the four corners of the trapezoidal tone burst is capable of eliciting such a response; (2) the amplitude of the response and its dependence on the frequency of the modulated tone both depend on the phase of the modulated sinusoid at the time a corner occurs; (3) such responses will arise in structures having sufficiently steep band edges, provided that the frequency of the modulated tone is well outside the pass band of the structure – for a corner in cosine phase, the sustained slope of the low-frequency band edge must be greater than zero and that of the high-frequency band edge must be greater than 12 dB/Oct, for a corner in sine phase the sustained slope of the low-frequency band edge must be greater than 6 dB/Oct and that of the high-frequency band edge must be greater than 18 dB/Oct; (4) they will not arise in response to tone bursts whose frequencies fall within the pass band of the structure; (5) nor will they arise in response to a trapezoidal tone burst of any frequency applied to structures (such as simple microphones or drivers) following second-order dynamics and having both spectral zeros at infinity.

We present theoretically derived relationships between the amplitude of transient responses and the tone-burst frequency, not only for the corners of trapezoidal tone bursts, but also for tone bursts of more general shapes. We conclude that, owing to its extraordinarily steep high-frequency rolloff, the filter associated with each cochlear axon is well suited to extracting temporal information from onset or offset singularities in modulated tones whose frequencies are above the characteristic frequency of the filter. Applying the theory to observed onset and offset responses to high-intensity tone bursts in auditory afferents of the Mongolian gerbil, we conclude that some of the responses we observed must have been sculpted in part by cochlear nonlinearities.

Cochlear onset responses; Cochlear offset responses; Properties of tone bursts; Cochlear nonlinearities

Introduction

Owing to the conspicuous tuning properties of cochlear axons and the celebrated tonotopy of the cochlea, auditory scientists traditionally treat tonal stimuli as signals in the frequency domain. Thus, when a tonal stimulus is turned on or off, as in a tone burst, one's tendency might be to consider the spectral 'splatter' generated by the onset and offset of the burst and the locations along the cochlea that might be excited by that splatter. Although this line of reasoning could be fruitful, it has inherent difficulties. The onset or offset of a tone is an event precisely placed in time. To depict that quintessential temporal aspect of the stimulus, its spectrum must include the frequency dependence of phase as well as that of amplitude. One difficulty arises from the fact that most investigators are not facile in the art of mentally integrating phase and amplitude information to deduce temporal qualities. Another difficulty arises from the definition of 'spectrum'. Strictly speaking, the spectrum of a stimulus waveform is defined by that waveform in its entirety. Thus, for example, the amplitude spectrum of a tone burst has peaks and nodes that correspond to constructive and destructive interference of spectral components generated at each end of the burst. One could isolate part of a tone burst (e.g., its

Correspondence to: Edwin R. Lewis, Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720, U.S.A.

onset) by placing a temporal window around that part, but the nature of the resulting spectrum will depend on the width, shape and placement of that window. On the other hand, each part of the tone burst is well defined point-by-point in time. Thus, when one observes transient increases in the spike rates of cochlear axons at the onsets and offsets of tone bursts (e.g., Geisler and Sinex, 1982; Rhode and Smith, 1985), those responses might heuristically be considered in the time domain – in terms of transient excitation of the cochlear tuning structures, rather than in the frequency domain – in terms of spectral splatter.

Although cochlear tuning structures often behave nonlinearly, Geisler and Sinex (1982) have shown that the essential features of transient excitation observed in cat cochlear axons in response to onsets and offsets of tones can be explained in terms of a linear tuned structure. Consequently, we should be able to explore those responses profitably with the conventional engineering approach to linear transient analysis, employing time-domain descriptions of the stimuli. In this paper we attempt to establish a basis for this approach by presenting time-domain descriptions of tone bursts and of the responses of a generalized linear tuned structure to such stimuli. We begin with the trapezoidal tone burst and then generalize the descriptions through Taylor's theorem to other tone bursts. This approach not only provides a rational alternative to the notion of splatter, it also provides insight concerning the design of stimuli. Our results elaborate and generalize time-domain modeling studies already published by Moore (1973), Grandori (1979), Geisler and Sinex (1982), and Antonelli and Grandori (1984). Preliminary reports of our results have been presented in Lewis and Henry (1988) and Henry and Lewis (1988).

Summary of theory

In time-domain analysis, each function conventionally is defined as having only zero value prior to a particular onset (starting) time and finite values (or zero value) after that. Thus every such function possesses a singularity at its onset time (i.e., either the function itself or one of its time derivatives will exhibit a stepwise discontinuity at the onset time). In this paper we shall be concerned largely with such singularities and with the excitations they produce in linear tuned structures.

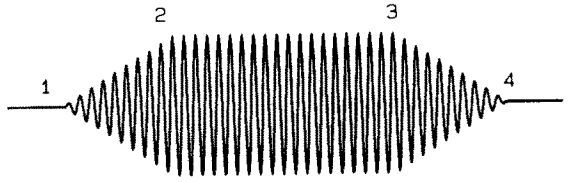

Fig. 1. A trapezoid tone burst comprises a sum of four ramp-modulated sinusoids, one beginning at each of its four corners. The ramp with its onset at point 1 provides the rising phase of the burst. The ramp with its onset at point 2 modulates a sinusoid 180 degrees out of phase with that modulated by the first ramp. The sum of the two brings the slope of the tone-burst modulation to zero; and so forth.

In this manner, a ramp function with its onset at time t_o is defined to be zero at all times prior to t_o; subsequently, its value is taken to be directly proportional to the time elapsed since onset (i.e., to $t - t_o$). The first time derivative of the ramp function exhibits a step discontinuity at the onset time. A sinusoid whose amplitude is modulated by a ramp function (i.e., a ramp-modulated sinusoid) also will possess a singularity at the onset time of the ramp. In its idealized form, the trapezoidal tone burst (Fig. 1) can be considered to comprise four ramp-modulated sinusoids, one beginning at each of its four corners (Appendix A1). Thus it possesses four singularities, each capable of exciting a tuned structure.

If the cochlear tuning structure or any other structure (such as a spectral filter or a microphone) truly responded linearly to the trapezoidal tone burst, then it would respond independently to each singularity. Thus the responses to the singularities may be analyzed one at a time. The phase of the sinusoid at the ramp onset time may vary from one singularity to another; and the nature of the responses will vary accordingly. For a linearly responding structure, the response for any onset phase will be the appropriately weighted sum of the prototypic response to a ramp-modulated sine wave (onset phase corresponds to the positive-going zero-crossing of the sinusoid) and the prototypic response to a ramp-modulated cosine wave (onset phase corresponds to the positive peak of the sinusoid) (Appendix A1). With a digital waveform synthesizer and appropriate

choices of rise, decay and plateau times, one can generate a trapezoidal tone burst with any desired phase for the ramp-modulated sinusoid added at each corner.

The response of any linearly operating structure to any stimulus waveform with a singularity will consist of a sum of components from two classes. One class derives its fundamental nature entirely from the stimulus waveform. Depending on the circumstances and the shape of the stimulus, the members of this class are known as 'steady state' response components or 'nonhomogeneous' response components. They correspond to the particular solution of the differential equation describing the dynamics of the structure. When the stimulus is a ramp-modulated sinusoid, the nonhomogeneous response components are ramp-modulated sinusoids and constant-amplitude sinusoids, both at the same frequency as the stimulus (Appendix A3). The other response class (Appendix A2) derives its fundamental nature entirely from the structure being excited. The members of this class are known as 'transient' response components or 'homogeneous' response components. Each member is a solution of the structure's differential equation in its homogeneous form (i.e., with the stimulus set equal to zero); and each corresponds to a 'natural frequency' of the structure. If the structure were struck by an impulsive stimulus (such as a click striking the ear), it would 'ring' at all of these natural frequencies at once (Guillemin, 1957).

The natural frequencies of a structure correspond to poles in its transfer function. If the structure is dynamically stable, each pole either is a negative number or is one member of a pair of complex conjugate numbers with a negative real part (Appendix A2). For the natural frequencies corresponding to both kinds of poles, the excitations produced by the singularity of a ramp-modulated sine wave and by the singularity of a ramp-modulated cosine wave are derived in Appendix A3.

Tuning of the natural-frequency excitation by ramp sinusoids

It is clear from Equations 21, 23, 28 and 30 of Appendix A3 that the degree to which a particular natural frequency of a tuned structure is excited

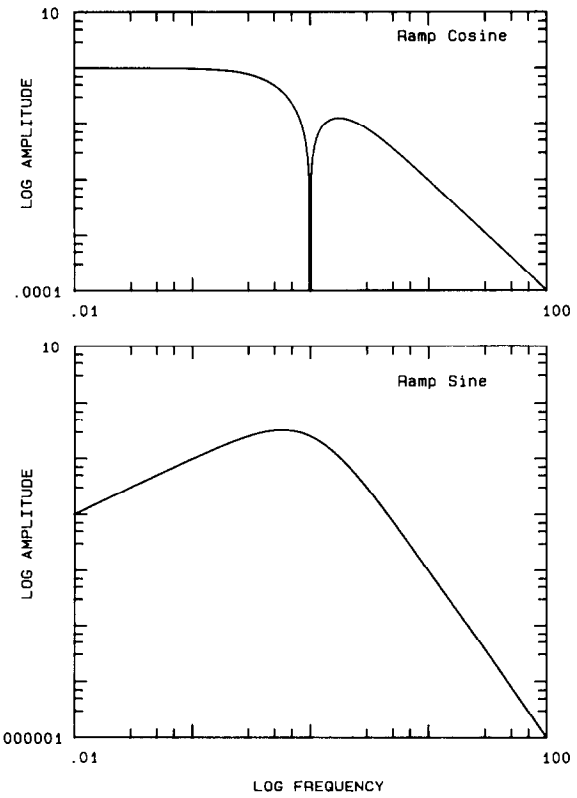

Fig. 2. Excitation of the natural frequency corresponding to a negative-real pole in a linear tuned structure, in response to the onset singularities of ramp-modulated sinusoids. Here the amplitude of the excitation is plotted against the frequency of the modulated sinusoid. For the ramp sine, the onset of the ramp modulation coincided with a zero-crossing of the modulated sinusoid; for the ramp cosine it coincided with the peak.

by the onset singularity of a ramp-modulated sinusoid depends on the ratio of the modulated stimulus frequency to the natural frequency being excited. In the case of the natural frequency corresponding to the negative, real pole, $z_k = -\alpha_k$, the amplitude of the excitation by the onset singularity depends on the ratio w_s/α_k, where w_s is the frequency of the modulated sinusoid in rad/s. Fig. 2 shows normalized plots of the amplitudes of excitation (K_{ak} and K_{bk} in Equations 21 and 28) for sine and cosine ramps. Notice that excitation by the onset singularity of a sine ramp is maximum when w_s/α_k is less than 1.0, and declines at a rate of 6 dB/Oct at low stimulus frequencies and 18 dB/Oct at high stimulus frequencies. The amplitude of excitation by the onset singularity of

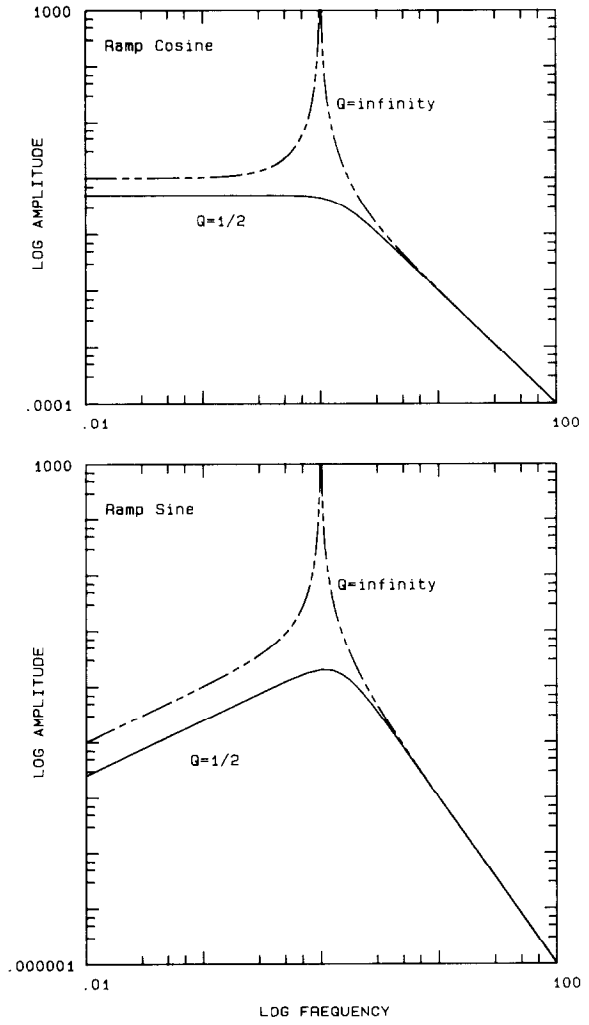

Fig. 3. Excitation of the natural frequencies corresponding to a pair of complex conjugate poles in a linear tuned structure, excited by the onset singularities of ramp-modulated sinusoids. Amplitude of excitation is plotted against frequency of the modulated sinusoid.

a cosine ramp has a sharp null where w_s/α_k equals 1.0. This corresponds to the value at which the polarity of the response reverses. At low stimulus frequencies, the response amplitude is constant; at high frequencies it declines at a rate of 12 dB/Oct.

In the case of the conjugate complex poles, $z_k = -\alpha k + i\beta_k$ and $z_k^* = -\alpha_k - i\beta_k$, the amplitude of the excitation by the onset singularity can be plotted against the ratio w_s/α_k. Fig. 3 shows normalized plots of the amplitudes of excitation (the amplitude of K_{ak} and K_{bk} in Equations 24 and 31) for sine and cosine ramps. Curves are shown for cases of natural frequencies with very high Q and very low Q, curves for intermediate values of Q will lie between these. Again, excitation by the onset singularity of a ramp-modulated sine wave declines at a rate of 6 dB/Oct at low frequencies, 18 dB/Oct at high frequencies. Excitation by the onset singularity of a ramp-modulated cosine wave is constant at low frequencies and declines at a rate of 12 dB/Oct at high frequencies.

Estimating the shape of the complete response

As we already mentioned, the complete response of a linearly-operating tuned structure to a ramp-modulated sinusoid comprises not only the excitations of natural frequencies by the onset singularity of the ramp, but also a step-modulated sinusoid and a ramp-modulated sinusoid, both at the same frequency as the modulated stimulus sinusoid. The overall response of the tuned structure to the ramp sinusoid will depend on the shape of the steady-state frequency-response characteristics of the tuned structure and on the frequency of the stimulus sinusoid relative to those characteristics. On the basis of repeated analysis of specific cases and repeated experimentation with linear analog filter devices, we offer the following conjecture:

A. If the stimulus frequency is well within a pass band of a linear tuned structure, then the sum of the natural frequency excitations and the step and ramp sinusoids will yield an overall response that approximates the stimulus waveform but with its onset apparently delayed slightly. In such cases, the natural-frequency excitations will not stand alone as markers of the ramp onset. Instead, they merely are manifested as absence of the early sinusoidal response at the stimulus frequency.

So far, we have found no exceptions to this rule. What follows is a second rule, which derives directly from the basic properties of linear analog spectral filters:

B. If the stimulus frequency is outside the pass band of a linear tuned structure, and if the asymptotic slope of the pass-band edge closest to

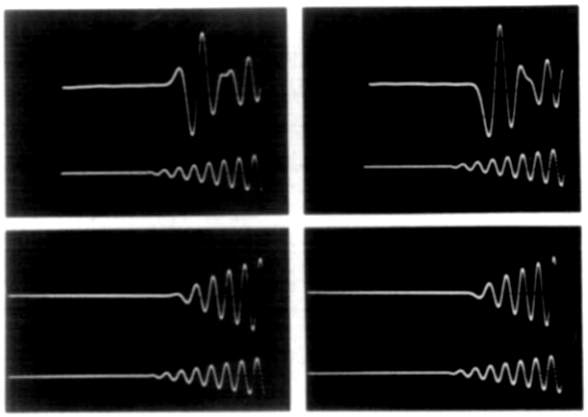

Fig. 4. Responses (upper traces in each photograph) of a specific tuned structure to ramp-modulated sine or cosine waves (lower traces), showing the dependence of response shape on the frequency of the sinusoid relative to the pass band of the structure. The ramp-sines and their responses are on the left, ramp-cosines on the right. The tuned structure was an analog filter set in the low-pass, maximally-flat (Butterworth) mode, with 96 dB/Oct rolloff at its high-frequency edge. The stimulus frequency was 10 kHz in each case. For the upper photographs, the corner frequency of the filter was 7 kHz, in the lower photographs it was 9 kHz. In the upper photographs, the response was amplified 500× relative to the stimulus; in the lower photographs it was amplified 25× relative to the stimulus.

the stimulus frequency is steeper than the corresponding rolloff in Fig. 2, then the step and ramp sinusoid response components will be selectively attenuated. In such cases, one always will be able to select frequencies for the modulated tone that will place it sufficiently far from the band edge to allow the excitations of natural-frequencies to stand alone as markers of the ramp onset. For this to occur with a ramp-modulated cosine wave, the asymptotic high-frequency rolloff must be greater than 12 dB/Oct, the asymptotic low-frequency rolloff must be greater than zero; for it to occur with a ramp-modulated sine wave, the asymptotic high-frequency rolloff must be greater than 18 dB/Oct, the asymptotic low-frequency rolloff must be greater than 6 dB/Oct.

Figs. 4 and 5 illustrate properties A and B with ramp-modulated sinusoids applied to a low-pass analog filter set for maximum flatness, with high-frequency rolloff at 96 dB/Oct. The ramp sinusoids were generated by a digital arbitrary waveform generator. Whenever the frequency of the sinusoid was within the passband of the filter, the transient excitations of the filter's own natural frequencies were absorbed into the response at the stimulus frequency. Their sole manifestation was the absence of immediate response to the stimulus ramp. When the frequency of the sinusoid was above the pass band of the filter, the stimulus sinusoid was suppressed, allowing the transient excitations of the filter's natural frequencies to emerge as a distinct sign of the onset of the ramp-modulation. We denote this phenomenon as a distinct edge response (see Rhode and Smith, 1985).

The frequency response curves of Figs. 2 and 3 tell us that the natural frequencies most excited by the onset singularity of the ramp will be those closest to the frequency of the stimulus. Fig. 6 depicts the frequency dependence of excitation of a natural frequency close to the upper band edge of a 96 dB/Oct low-pass filter when the stimulus is a ramp-cosine. Superimposed on that graph is the sinusoidal steady-state amplitude response

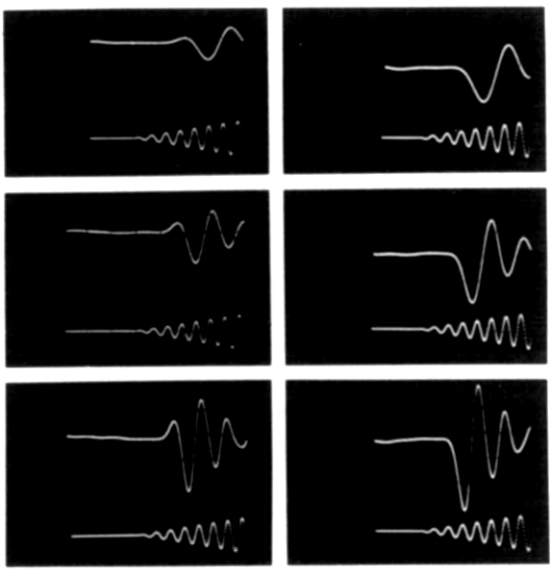

Fig. 5. Further responses (upper traces) of the tuned structure of Fig. 4 to ramp modulated sine-waves (photographs on left) and ramp-modulated cosine waves (photographs on right). In each case, the stimulus frequency was 10 kHz. The corner frequency of the filter was 4 kHz for the top pair of photographs, 5 kHz for the middle pair, and 6 kHz for the bottom pair. In each case the response was amplified 1000× relative to the stimulus.

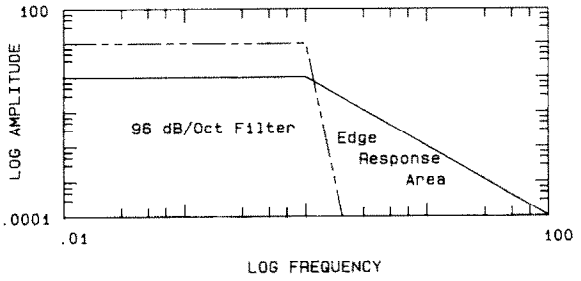

Fig. 6. Qualitative explanation of the data in Figures 4 and 5. Here the frequency dependence (solid line) of the amplitude of excitation of a low-Q natural frequency is superimposed on the frequency dependence (dashed line) of the low-pass filter. Stimulus frequency (abscissa) has been normalized with respect to the corner frequency of the filter. At frequencies well above the corner, where the filter strongly suppresses the amplitude of the stimulus sinusoid itself, the transient excitation of the natural frequency will emerge as a clear 'edge response' to the onset singularity of the ramp.

characteristic of the filter. We expect a distinct edge response to occur whenever the frequency of the stimulus sinusoid falls in the region where the amplitude graph for the natural-frequency excitation lies well above the amplitude graph for sinusoidal steady-state response of the filter. The same rule should hold for any tuned structure (including microphones, acoustic drivers, cochleas, etc.).

Clearly, if the rolloff of the steady-state tuning curve of a particular structure were not sufficiently steep, the amplitude graph for natural-frequency excitation would never diverge from the amplitude graph for the filter's steady-state response. In that case, no distinct edge response would occur at any frequency. This result is illustrated in Figs. 7 and 8, which show the responses of a second-order resonant filter with both zeros at infinity (high-frequency rolloff at 12 dB/Oct, low-frequency rolloff at 0 dB/Oct) to ramp-cosines at various frequencies. Although these results were derived analytically with the aid of an algebraic manipulation program (VAXIMA), they are representative of the situations in many microphones and drivers with second-order dynamics. Ramp-modulated sinusoids will not produce edge responses in such devices.

Finally, where edge responses do occur in linearly-operating tuned structures, they comprise the combined excitations of all the natural frequencies in the pass bands, each weighted according to its distance from the frequency of the modulated stimulus sinusoid. Again, it is not a simple task to

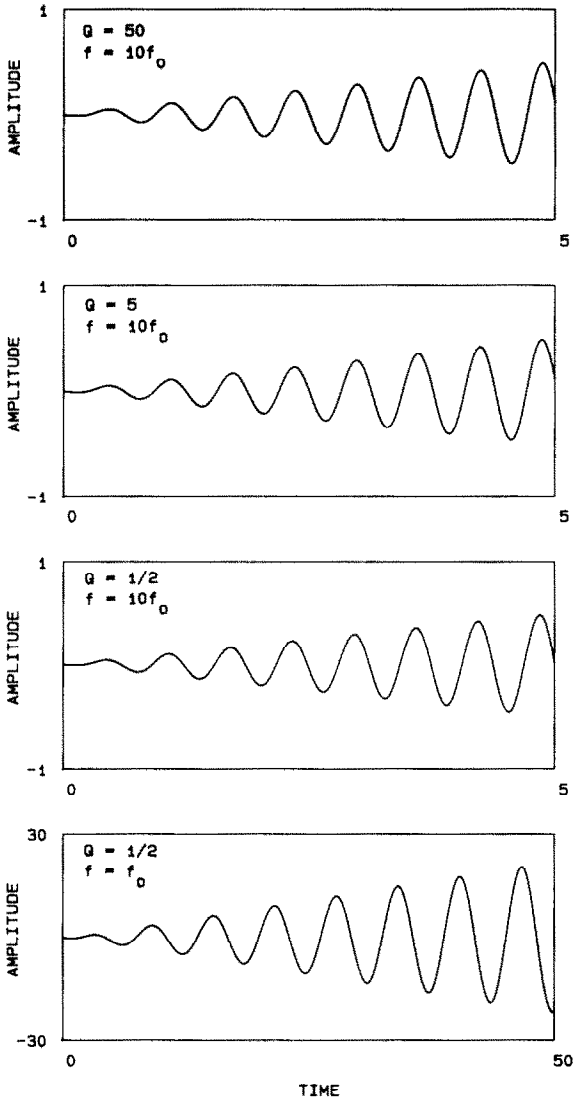

Fig. 7. Computed responses of a second-order resonant (two-pole) filter with both zeros at infinity to ramp cosines at the resonance frequency (bottom panel) and at 10× the resonance frequency (upper three panels). The Q of the filter is displayed in the upper left of each panel. Regardless of the Q of the filter (i.e., regardless of its bandwidth), and regardless of the frequency of the stimulus, ramp modulations of sinusoids at frequencies equal to or greater than the resonance frequency produce no edge responses. As predicted by the linear theory, at 12 dB/Oct the rolloff of the filter is not sufficiently steep to selectively suppress the sinusoidal response relative to the natural-frequency excitation.

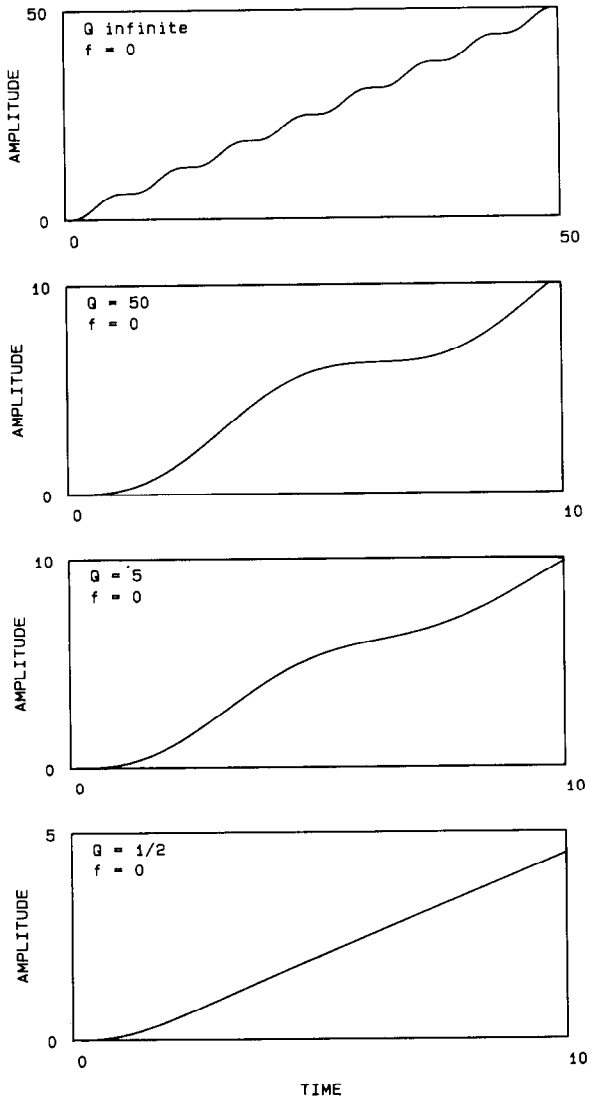

Fig. 8. The sequence of Fig. 7 continued for frequencies below the resonance frequency. Even as the modulated frequency approaches zero, the second-order filter fails to produce a distinct edge response. If one or both zeros of the filter had been placed at the origin, then a distinct edge response would have appeared as the modulated frequency approached zero.

deduce precisely the shape of such complex edge responses; but we can state a simple rule of thumb that is applicable to linearly-operating tuned structures with some zeros at infinity (i.e., with a high-frequency band edge with indefinitely sustained rolloff). Thus it would apply to linear operation of the filter of Fig. 6 and evidently to linear operation of the mammalian cochlea as well as that of the hearing organs of non-mammalian vertebrates:

C. When an edge response is produced in a linearly-operating tuned structure with n zeros at infinity (i.e., with indefinitely-sustained high-frequency rolloff at 6n dB/Oct), the nth temporal derivative (d^n/dt^n) is the lowest order derivative not initially equal to zero in the edge-response waveform. For a given pass-band, increasing n will result in the initial portion of the edge response becoming more deeply inflected; the onset of the edge response therefore will appear to be increasingly delayed. This apparent delay would become conspicuous if the edge response were observed through a threshold device, such as a spike generator (i.e., there would be no response at all until the edge response reached the threshold). *

To be convinced of the validity of this rule of thumb, imagine all of the partial-fraction terms corresponding to an edge response being recombined over a common denominator. Since its denominator includes all of the poles of the individual terms of the partial-fraction expansion, this combined response term can be factored into two components – one with either a negative-real pole or a pair of conjugate complex poles, arbitrarily chosen from among those in the partial-fraction expansion, the other containing the rest of the poles. The factor containing the single pole or pole-pair can be considered to be the Laplace transform of an input (whose shape is that of an impulse-excitation to the corresponding natural frequency). The remaining term then would represent a tuned structure through which that natural frequency excitation must pass, and which therefore would reshape the response waveform. Each zero at infinity in that remaining term will represent a separate temporal integration of the initial part of the waveform; and each such integration

* It is important to recall that a true linear time delay is a device with the capacity to store an infinite amount of information (all of the values of a time function over the entire duration of the delay). A linear system of finite order can store only a finite number of values of a time function. Combining such a system with a threshold device (which is inherently nonlinear) does not increase its memory capacity.

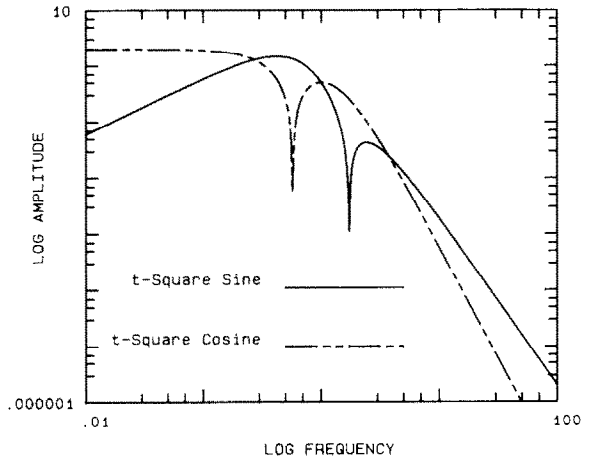

Fig. 9. Excitation of the natural frequency corresponding to a negative-real pole in a linear tuned structure, in response to the onset singularities of sinusoids modulated by *t*-squared. The amplitude of the excitation is plotted against the frequency of the modulated sinusoid.

removes another nonzero temporal derivative at the beginning of the response.

Modulation by functions other than ramps

A wide variety of tone-burst envelopes should be representable by a Taylor's series that combines terms in various powers of the time elapsed since a single onset time, with each term assigned an appropriate sign and an appropriate amplitude parameter. This approach is presented in Appendix A4. Each term in the series would have its own singularity at the onset time; and a linearly operating structure would respond independently to each of these singularities. For each generic term – e.g., step modulation, ramp modulation, time-squared modulation, time-cubed modulation, and so forth, one can derive plots similar to those of Figs. 2 and 3, showing the dependence of the amplitude of natural-frequency excitation by the onset singularity on the frequency of the sinusoid being modulated.

Fig. 9 shows the excitation of the natural frequency corresponding to a negative real pole, α_k, as a function of the ratio w_s/α_k (recall that w_s is the frequency of the modulated sinusoid in rad/s). Notice that the high-frequency rolloff of the excitation in response to the onset singularity of the sine wave modulated by *t*-squared has the same asymptotic slope as the excitation in response to the onset singularity of a ramp-modulated sine wave; the asymptotic slope of the high-frequency rolloff for the onset singularity of a cosine wave modulated by *t*-squared is 12 dB/Oct greater than that for the ramp-modulated cosine wave. This pattern continues indefinitely: when the degree, k, of the t^k modulation increases from an odd integer to the subsequent even integer, the asymptotic slope of the high-frequency rolloff for natural-frequency excitation by the onset singularity of a modulated sine wave does not change; that for a modulated cosine wave increases by 12 dB/Oct. When k increases from an even integer to the subsequent odd integer, it is the asymptotic slope of the high-frequency rolloff for excitation by the onset singularity of the modulated sine wave that increases by 12 dB/Oct, while that for the modulated cosine wave does not change. Curves of the sort shown in Fig. 9 can be constructed for various values of k, for natural frequencies corresponding to negative real poles and for natural frequencies corresponding to complex conjugate poles. The algebraic manipulations required for generation of the plotted functions are easily carried out with available software.

Rules A, B and C apply to t^k modulation as well as to ramp modulation, but the rolloff slopes specified in Rule B would have to be adjusted for the particular value of k. For example, with $k = 0$ (corresponding to a step-modulated sinusoid), the slopes for the cosine wave remain the same, those for the sine wave become 6 dB/Oct on both low- and high-frequency sides. Thus, as expected, the modeling observations of Antonelli and Grandori (1984; see their Fig. 7) with pulsed tone bursts are consistent with these rules.

Under certain circumstances, such as studies of whole nerve (compound) action potentials, it is desirable to randomize the phase of the modulated sinusoid at the time of onset singularity, so that the phase varies from one tone burst to the next. When the phase of the modulated sinusoid at the time of modulation onset is not a precise multiple of 1/4 cycle (so that the modulated sinusoid is neither pure sine nor pure cosine, but a combination of the two), then: (1) the sharp nulls in Figs. 2 and 9 will not be present; (2) cosine excitation will prevail at low stimulus frequencies; and (3) the excitation with less-steep high-frequency rolloff

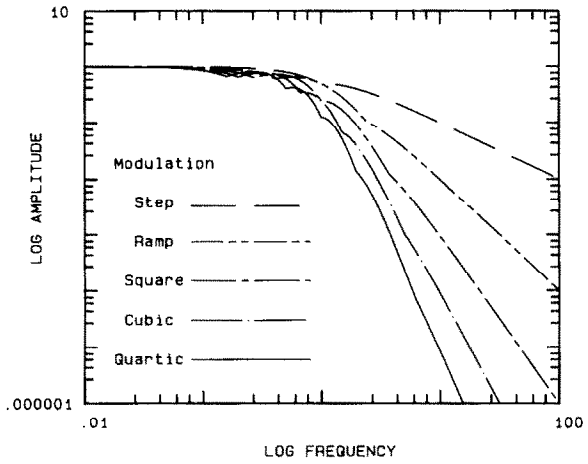

Fig. 10. Excitation of the natural frequency corresponding to a negative-real pole in a linear tuned structure, in response to the onset singularities of sinusoids modulated by various powers of t. The frequency dependence of amplitude is approximated here for onsets at random phases of the modulated sinusoid (see text for explanation).

will prevail at high frequencies. Thus, for ramp-sinusoids of this type, cosine excitation will prevail at high frequencies (see Figs. 2 and 3); for t-square sinusoids of this type, sine excitation will prevail at high frequencies (see Fig. 9). In Figs. 10 and 11, we have incorporated conditions (1) through (3) into the graph for each degree (k) by plotting the excitation (sine or cosine) that is larger at each value of frequency. The results are

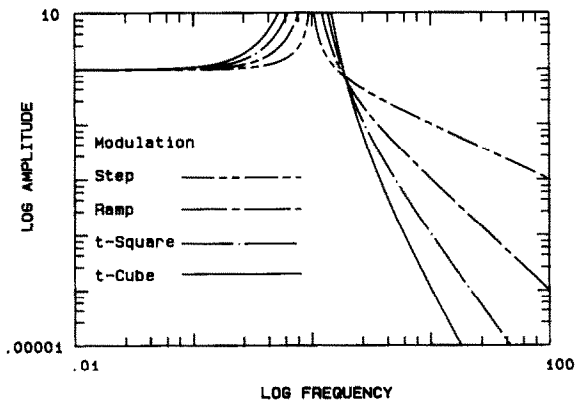

Fig. 11. Excitation of the natural frequencies corresponding to a complex conjugate pair of poles in a linear tuned structure, in response to the onset singularities of sinusoids modulated by various powers of t. As in Fig. 10, the frequency dependence of amplitude is approximated here for onsets at random phases of the modulated sinusoid.

shown for excitation of the natural frequency corresponding to a negative real pole, and for excitation of the natural frequencies corresponding to a very high-Q pair of complex conjugate poles. Notice that the amplitude of the slope of the high-frequency rolloff begins at 6 dB/Oct for the step-modulated sinusoid and increases by 6 dB/Oct for each increase of 1.0 in the degree (k) of the t^k modulation.

The situation in Figs. 4 and 5 might be considered a crude model of the cochlea, depicting responses by axons with different CFs to the onset singularity of one tone. We now can see that the excitation profile over those axons would depend markedly on the shape of the modulation. Thus we have the interesting proposition that the auditory nerve might encode the shapes of modulation waveforms in the profile of edge responses of axons, with an important part of that profile being mapped over axons whose CFs are different from the frequency of the stimulus sinusoid.

Application to some experimental results

Methods

60 to 200-day old Mongolian gerbils (*Meriones unguiculatus*), reared in an acoustically controlled room, were examined to insure that they were free from aural cholesteatoma and otitis. Healthy animals were pretranquilized with chlorprothixene (5 mg/kg, i.m.) and anesthetized 30 to 160 min later with ketamine (40 mg/kg, i.p.). Supplemental doses of ketamine (10 mg/kg, i.p.) were administered as needed. The surgical approach was that first published by Chamberlain (1977). The left pinna and associated muscles were removed, leaving the left ear canal exposed an unimpeded. The left bulla was opened, exposing the round-window antrum; and part of the floor of the antrum was removed to expose the cochlear nerve of the left ear. Single cochlear afferent axons within 0.25 mm of the surface of the nerve were penetrated with glass micropipettes filled with 3 M NaCl and having impedances greater than 50 mΩ. Spike activity was recorded simultaneously with the stimulus and stimulus trigger. The stimulus was a trapezoid tone burst, approximately 20 ms in duration, with the onset occurring at random phases of the modulated sine wave. It was moni-

tored with an Etymotic ER7C probe microphone inserted in the ear canal. The output of the probe microphone was calibrated with a Bruel and Kjaer sound pressure level meter and analyzed on-line with a Hewlett-Packard 3561A Dynamic Signal Analyzer. Both instruments were used to calibrate the stimulus amplitude over the entire frequency range employed in the experiment. Throughout the experiments, the stimuli all were well within the linear operating range of the Yamaha piezoelectric acoustic driver; thus the Dynamic Signal Analyzer revealed no amplitude-dependence in the normalized shape of the spectral profile of the acoustic driver output. Also, throughout the experiment, the core temperature of the animal (sensed by rectal probe) was maintained between 36° and 39°C; and the level of anesthesia was maintained at a level sufficient to eliminate reflexive movement to tail pinch.

The CF of each unit was determined from on-line auditory presentations of spike responses to 20-ms trapezoidal tone bursts presented at a rate of 5/s. The CF was bracketed increasingly narrowly as the amplitude of the stimulus was reduced in steps, beginning at approximately 60 dB SPL. The threshold in each case was taken to be the minimum sound pressure level at CF at which the experimenter could hear clear spike responses to each tone burst. Once CF had been determined, the stimulus amplitude was increased to a level approximately 50 dB above threshold at CF, and the frequency, amplitude and the slopes of the trapezoidal modulation ('rise' and 'decay' times of the tone burst) were varied in steps.

Experimental results and discussion

Figs. 12 and 15–17 display peristimulus time histograms of tone burst responses from three axons that are representative of our sample to date (well over 100 axons). The axon of Fig. 12 exhibited a threshold of approximately 20 dB SPL at its CF (4.7 kHz). In response to trapezoidal tone bursts with 0.3 ms rise and decay (r/d) times at frequencies well above CF, it exhibited conspicuous tendencies to produce spikes synchronized with the onsets and offsets of the tone bursts. However, at frequencies close to or equal to CF, it produced no such responses, even when the r/d times were reduced drastically. This pattern (onset and offset responses at frequencies well above CF, none at CF) was repeated often in our sample of cochlear axons. It might be explained by the shape of the response of a linear tuning structure to a ramp-modulated sinusoid. If the band edge of the structure's tuning curve is

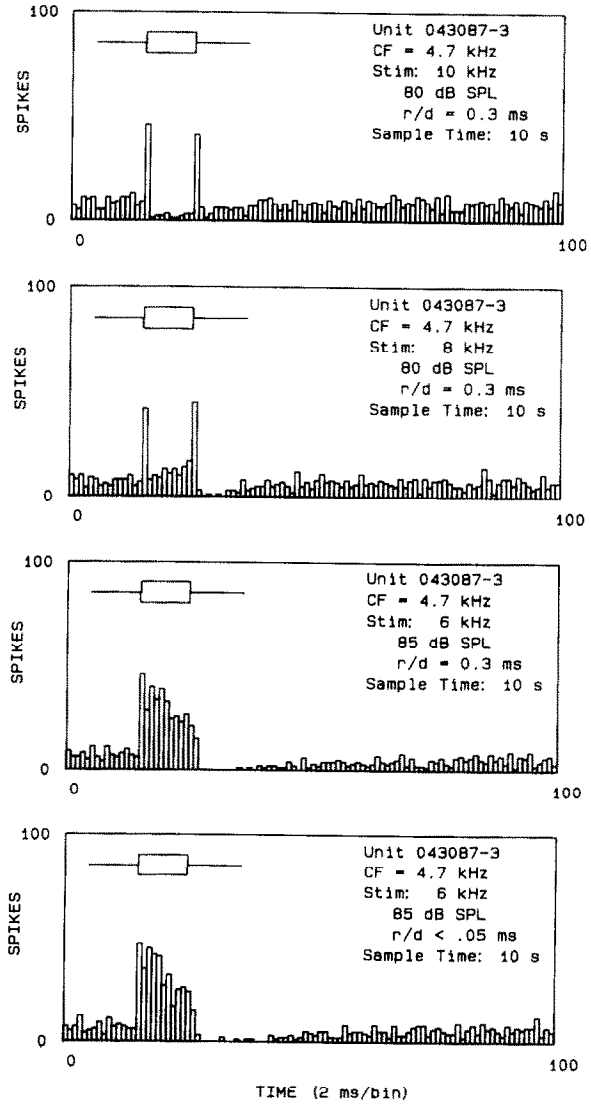

Fig. 12. Peristimulus time histograms of the responses of a gerbil cochlear afferent axon to trapezoidal tone bursts of random phase. The CF of the axon and the parameters of the stimulus are presented in the upper right of each panel (r/d = rise and decay times, amplitude is given for the tone-burst plateau). The timing of the tone burst is depicted by the diagram in the upper left. The tone burst was presented at a rate of 5 per s; thus the histogram in each case shows the entire stimulus cycle.

sufficiently steep (e.g., greater than 12 dB/Oct for ramp-modulated cosine waves), and the modulated frequency of the stimulus falls well above the structure's pass band, then the transient excitation of the structure's own natural frequencies will emerge as distinct (edge) responses at the onset of each ramp-modulated sinusoid. On the other hand, when the modulated stimulus frequency falls within the structure's tuning band, the transient excitation of the structure's natural frequencies will not emerge as distinct (edge) responses.

Recall that the trapezoid tone burst comprises four ramp-modulated sinusoids. In principle, one expects a distinct response to the onset singularity of each of them (Fig. 13). However, when the periods of the excited natural frequencies are comparable to the r/d times of the trapezoid, then the edge responses of the tuning structure will merge (Fig. 14). In order to separate these responses, we applied tone bursts with longer r/d times. Fig. 15 shows the responses of a cochlear axon with a threshold less than 10 dB SPL at CF (6.6 kHz). The stimulus in this case was a 10 kHz tone burst with r/d times ranging from 2 to 9 ms. In the upper two panels, we see distinct responses to the first two corners of the tone burst (corners 1 and 2 in Fig. 1), with the response to the first corner being by far the larger. In the upper panel, it is not clear whether corner 3 or corner 4 is responsible for the strong response at the end of the trapezoid; but the response seems to anticipate corner 4, implying that it is caused by corner 3. This causal relationship is clarified in the subsequent panels, in which the response clearly anticipates corner 4 and exhibits a nearly constant latency (approximately 2 ms) with respect to corner 3. In terms of the slopes of its constituent ramp-sinusoids, a 90 dB SPL trapezoid with 9 ms r/d times is the same as a 60 dB SPL trapezoid with 0.3 ms r/d times. Recall that the phase of the trapezoidally-modulated sinusoid was random. Therefore, taken over the entire sample of approximately 50 tone bursts for each panel, the four corners of the trapezoid represented essentially identical sets of ramp-modulated sinusoids. From the linear theory alone there is no explanation for the selective responsiveness to corners 1 and 3 in the top two panels and to corner 3 alone

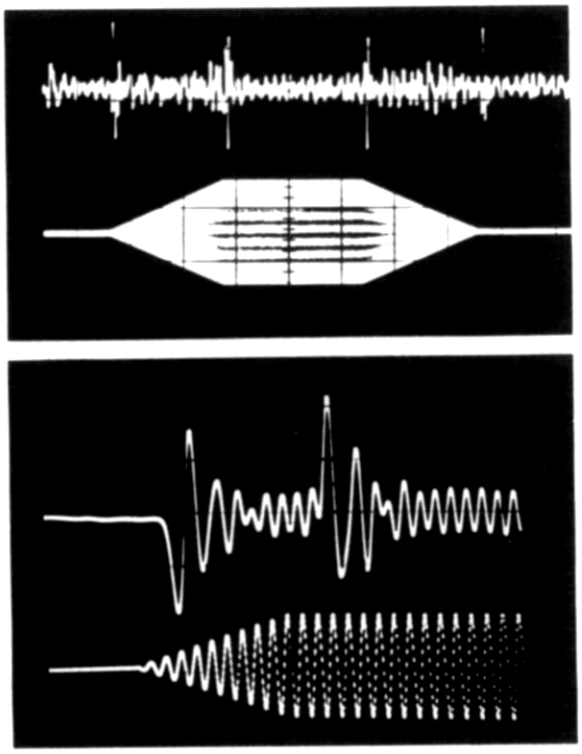

Fig. 13. Responses of a specific tuned structure (analog filter, set for maximally-flat, low-pass operation, 96 dB/Oct rolloff) to a trapezoid tone burst (see also Grandori, 1979; Geisler and Sinex, 1982). For the top photograph, the filter corner frequency was 5 kHz; the stimulus frequency was 10 kHz; and the response (upper trace) was amplified 5000× relative to the stimulus. For the top photograph, a 24 dB/Oct high-pass filter (corner frequency = 1 kHz) was added in cascade with the low-pass filter. This prevented low-frequency noise components from obscuring the edge responses. For the bottom photograph, the stimulus frequency was 10 kHz; the filter corner frequency was 6.6 kHz; and the response (upper trace) was amplified 500× relative to the stimulus (lower trace).

in the bottom two panels. The linear theory alone also cannot account for the conspicuous difference in latency between the response to corner 1 and that to corner 3.

Additionally, the linear theory does not explain another response pattern seen in several cochlear axons in our sample. That pattern comprised distinct edge responses for high-intensity tone bursts at frequencies close to or equal to CF. The responses of such an axon are shown in Figs. 16 and 17. This cochlear afferent exhibited a threshold of approximately 20 dB SPL at its CF (2.6 kHz). At low to moderate stimulus intensities at CF, the

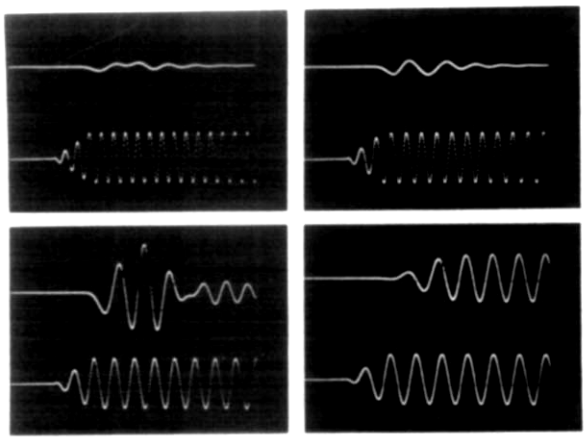

Fig. 14. Responses of a low-pass filter with corner frequency at 4.7 kHz to a series of tone bursts similar to those used in the experiments of Fig. 12. The analog filter again was set for maximally-flat low-pass operation with 96 dB/Oct rolloff. Stimulus frequencies: upper left = 10 kHz; upper right = 8 kHz; lower left = 6 kHz; lower right = 4.7 kHz. For the photographs at 6, 8 and 10 kHz, the response (upper trace) was amplified 25× relative to the stimulus. For the photograph at 4.7 kHz, the response was amplified 2.5× relative to the stimulus.

axon exhibited the classic 'primary-type' response (upper panel in Fig. 16). As the stimulus intensity was increased, clear edge responses emerged. However, while the edge response at the onset of the tone burst was well synchronized, that at the offset was not. In this respect, the offset response at CF was conspicuously different from that to tone bursts well above CF (Fig. 17).

The linear theory presented earlier in this paper forces us to invoke nonlinear mechanisms in our attempts to explain the response patterns of Figs. 12 and 15–17. Among other things, the linear theory demonstrates conclusively that one must not dismiss these patterns as mere consequences of 'spectral splatter'. Regardless of how one might define 'splatter', the time domain analysis demonstrates that it must be the same for all four corners of the trapezoid tone burst with random phase. Why then would an axon be selectively sensitive to corner 3; and why would the latency in the response to corner 3 be so much longer than that to corner 1? We found the selective sensitivity to corner 3 (relative to corners 2 and 4) not only among individual axons, but also in the whole-nerve compound action potential (CAP). However, the CAP response generally reflects greatest

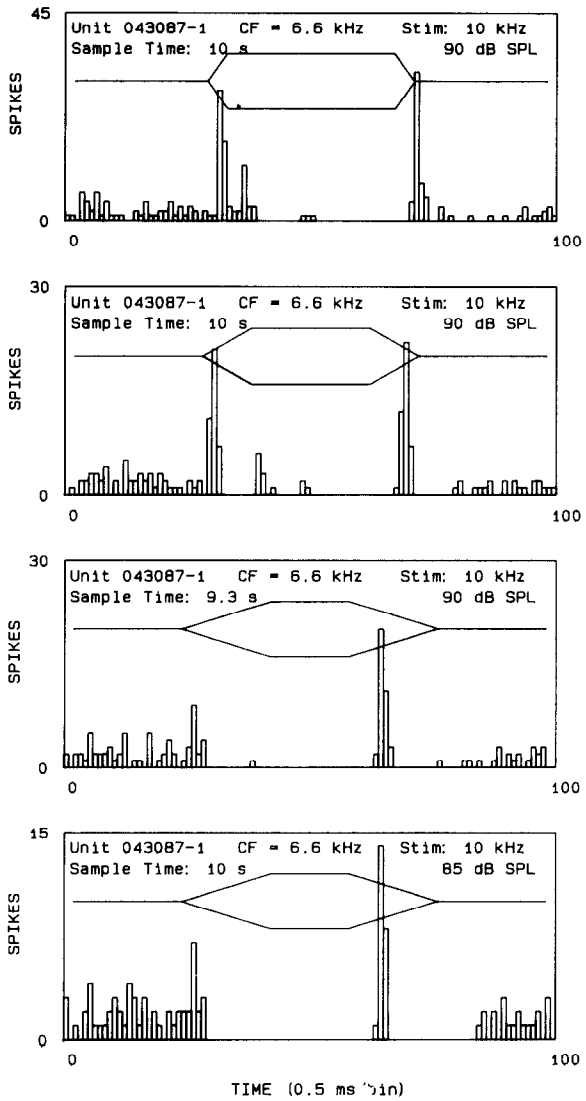

Fig. 15. Peristimulus time histograms showing responses of a gerbil cochlear afferent axon to tone bursts with rise and decay times designed to sort the edge responses to the four corners of a trapezoidal tone burst (the onsets of the four component ramps). The timing of the tone burst is depicted by the figure inserted in each panel. The tone burst was repeated at a rate of 5 per s; therefore the histogram in each case displays only one-fourth (50 ms) of the complete stimulus cycle. Between 500 Hz and 20 kHz, the sound level within the acoustical chamber was less than 1.4×10^{-6} Pa/Hz$^{1/2}$ in the absence of the animal. No effort was made to control for cardiovascular, respiratory, or gastro-intestinal sounds emitted by the animal itself. The latencies of responses to corner 1 typically were less than 1 ms. The latencies for second-order auditory neurons in the gerbil are reported to be approximately 7 to 9 ms (Woolf and Ryan, 1985).

sensitivity to corner 1. Since the CAP apparently corresponds to near-synchronous responses of many axons, the selectivity of the CAP response to corner 3 (relative to corners 2 and 4) implies that this aspect of Fig. 15 is representative of a large population of cochlear axons. According to the results in the lower two panels of Fig. 15, the presence of the portion of the tone burst that

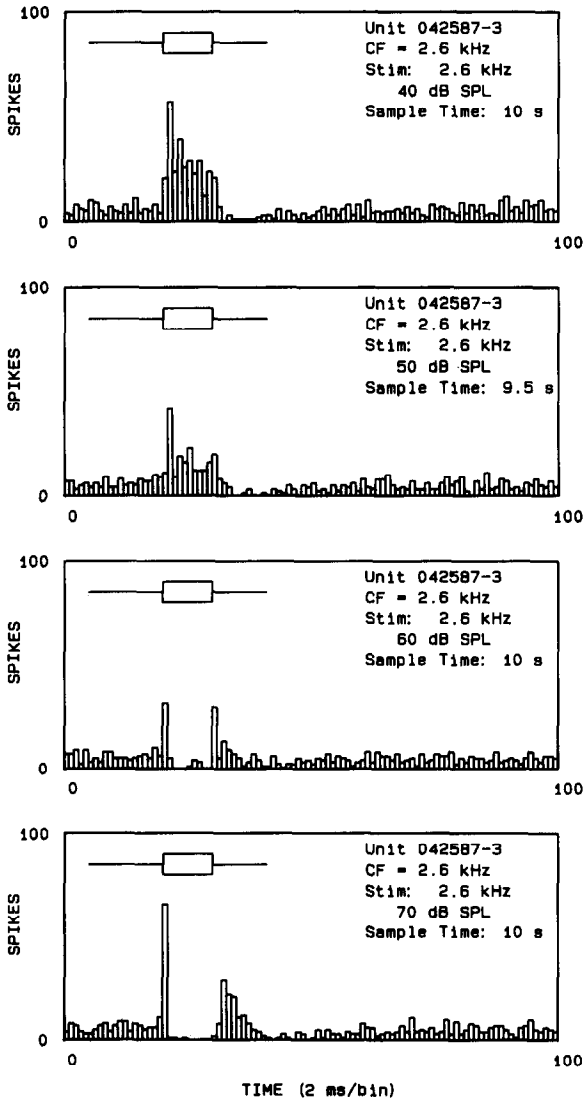

Fig. 16. Peristimulus time histograms showing responses of a gerbil cochlear afferent axon to tone bursts at CF. A few of the axons in our sample showed onset and offset responses to tone bursts at CF; others, such as that of Fig. 12, did not. See caption of Fig. 15 for discussion relevant to apparent suppression to spike rates below spontaneous level.

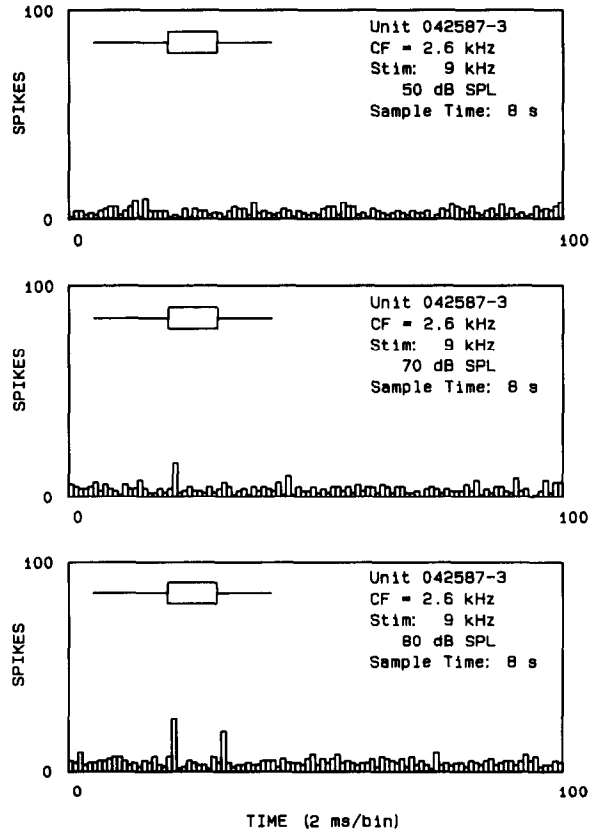

Fig. 17. Peristimulus time histograms for the axon of Fig. 16, stimulated at a frequency 1.8 octaves above CF. The offset response is considerably sharper than it was at CF.

occurs prior to corner 3 somehow has modified the sensor and/or its afferent axon in such a way that it now is responsive to a stimulus to which it previously was insensitive.

So far in the CAP we have not seen conspicuous latency differences in the responses to corners 1 and 3. Nevertheless, its presence in Fig. 15 implies a marked change in the dynamics of the tuning structure in that case, between the time of the response to corner 1 and the time of the response to corner 3. Our third rule of thumb (C) regarding the shapes of edge responses tells us that one possibility is an increase in the number of zeros at infinity. Another possibility would be a downward shift in the natural frequencies excited by the onset singularity of the ramp-sinusoid. Regardless of its underlying causes, some sort of change in the dynamics of the tuning structure definitely is implied by the data of Fig. 15. If

sensitivity to corner 3 in all of the panels in Fig. 15 is a consequence of the same modification that apparently produces it in the lower two panels, then the change in dynamics that leads to the latency difference might be concomitant with the shift in sensitivity that allows the system to respond to corner 3. If that were the case, we would have a phenomenon in the ear (concomitant shifts in sensitivity and dynamics in response to applied stimuli) apparently analogous to that first discovered by Fuortes and Hodgkin (1964) in invertebrate photoreceptors and subsequently observed by Baylor et al. (1974) and others in photoreceptors of vertebrates. It is possible that in the ear the modification involves a shift in the trade-off between spectral resolution and temporal resolution (e.g., see Lewis, 1987).

All of these issues await further investigation.

Final speculation

It is clear that precise temporal resolution is a key ingredient in spatial localization of sound, and that the temporal resolution (a few tens of microseconds) implied by psychophysical measurements in humans (Hafter et al., 1980) is remarkable considering the apparent timing jitter in VIIIth-nerve axons. Of course one can imagine considerable enhancement of the temporal signal/noise ratio by ensemble averaging over large populations of auditory axons. In this regard, it is interesting to note a paper published by Katz and Schmitt in 1940, in which they show that spikes being conducted in neighboring unmyelinated axons tend to become synchronized if they are not too far apart and the axonal conduction velocities are not too different. The spike in the faster axon is slowed by the presence of the spike in the slower axon, and that in the slower axon is accelerated by the presence of the spike in the faster axon. In this manner, the spikes tend to come together in space and time. Extending this result to a bundle of unmyelinated fibers, one can imagine that bundle serving as a synchrony amplifier. Unless nodes of Ranvier are aligned, it seems unlikely that this phenomenon could occur among myelinated fibers. On the other hand, it might well occur among unmyelinated fibers at various locations in the auditory system.

Appendices

A1. Decomposition of a trapezoidal tone burst

A trapezoidal tone burst (Fig. 1) can be decomposed into a series of four ramp-modulated sine functions and four ramp-modulated cosine functions, as follows (see Grandori, 1979, for a similar decomposition of a pulsed tone burst):

$$y(t) = y_1(t) + y_2(t) \tag{1a}$$

$$\begin{aligned} y_1(t) = m_a[&(t-t_1)\sin(w_s t)u(t-t_1) \\ &-(t-t_2)\sin(w_s t)u(t-t_2) \\ &-(t-t_3)\sin(w_s t)u(t-t_3) \\ &+(t-t_4)\sin(w_s t)u(t-t_4)] \end{aligned} \tag{1b}$$

$$\begin{aligned} y_2(t) = m_b[&(t-t_1)\cos(w_s t)u(t-t_1) \\ &-(t-t_2)\cos(w_s t)u(t-t_2) \\ &-(t-t_3)\cos(w_s t)u(t-t_3) \\ &+(t-t_4)\cos(w_s t)u(t-t_4)] \end{aligned} \tag{1c}$$

$$w_s = 2\pi f_s \tag{1d}$$

$$\begin{aligned} u(t-t_i) &= 1 \quad t \geq t_i \\ &= 0 \quad t < t_i \end{aligned} \tag{1e}$$

where $y(t)$ is the overall tone burst; f_s is the frequency of the tone (in Hz); and

$$t_1 < t_2 \leq t_3 < t_4 \tag{1f}$$

$$t_2 - t_1 = t_4 - t_3 \tag{1g}$$

When $t_2 = t_3$, the trapezoidal tone burst becomes a tent-shaped tone burst (see Geisler and Sinex, 1982). The parameters m_a and m_b determine both the phase, ϕ, of the sinusoid during the tone burst (relative to $\cos[w_s t]$) and the magnitude, M, of the slopes of the rising and falling skirts of the tone burst:

$$M = \left(m_a^2 + m_b^2\right)^{1/2} \tag{2a}$$

$$\phi = \tan^{-1}[m_a/m_b] \tag{2b}$$

If the cochlear tuning structure or any other structure (such as a spectral filter or a microphone) responded linearly to the tone burst, then it would respond independently to each of the four sine ramps in $y_1(t)$ and each of the four cosine ramps in $y_2(t)$. Define $x_a(t-t_i)u(t-t_i)$ to be the (linear) response of the tuned structure to the ramp $(t-t_i)\sin[w_s t]u(t-t_i)$ and $x_b(t-t_i)u(t-t_i)$ to be its (linear) response to the ramp $(t-t_i)\cos[w_s t]u(t-t_i)$, where $u(t-t_i)$ is zero for all values of t less than or equal to t_i and 1.0 for all values of t greater than or equal to t_i. If a tuned structure responded linearly, then the response, $x_1(t)$, to the stimulus component $y_1(t)$ would be

$$\begin{aligned} x_1(t) = m_a\{ & x_a(t-t_1)u(t-t_1) \\ & - x_a(t-t_2)u(t-t_2) \\ & - x_a(t-t_3)u(t-t_3) \\ & + x_a(t-t_4)u(t-t_4)\} \end{aligned} \tag{3a}$$

the response, $x_2(t)$ to $y_2(t)$ would be

$$\begin{aligned} x_2(t) = m_b\{ & x_b(t-t_1)u(t-t_1) \\ & - x_b(t-t_2)u(t-t_2) \\ & - x_b(t-t_3)u(t-t_3) \\ & + x_b(t-t_4)u(t-t_4)\} \end{aligned} \tag{3b}$$

and the response, $x(t)$, of the tuned structure to the complete stimulus, $y(t)$, would be

$$x(t) = x_1(t) + x_2(t) \tag{4}$$

Therefore, we can characterize the linear response $x(t)$ completely if we know the two basic component responses $x_a(t)$ and $x_b(t)$.

A2. Predicted linear responses to impulses

The theories of Laplace transforms and partial-fraction expansion allow us to describe the overall linear operation of any structure as a finite or infinite series of terms, each representing an elementary time-domain operation. Each term corresponds to a 'natural frequency' of the structure. Although, theoretically, a structure may exhibit multiple occurrences of the same natural frequency, the probability of this happening in a real structure is zero. Furthermore, even if it did happen, the dynamic behavior of the structure would be no different if the set of identical natural frequencies were replaced by a set of distinct natural frequencies whose differences were infinitesimal. Therefore, we can describe the linear operation (including time delay) of any structure (including a traveling wave structure) as a series of terms each corresponding to a different natural frequency.

If the structure is dynamically stable, the ringing corresponding to each natural frequency will be damped and eventually will subside. In that case, the individual components of ringing will take one of two forms: either

$$f_i(t) = A_i \exp[-\alpha_i t] \quad \text{when } t \geq 0 \tag{5a}$$

$$f_i(t) = 0 \quad \text{when } t < 0 \tag{5b}$$

or

$$\begin{aligned} f_j(t) = {} & 2a_j \exp[-\alpha_j t]\cos[\beta_j t] - 2b_j \\ & \times \exp[-\alpha_j t]\sin[\beta_j t] \quad \text{when } t \geq 0 \end{aligned} \tag{6a}$$

$$f_j(t) = 0 \quad \text{when } t < 0 \tag{6b}$$

Each of these components is the response to an impulse stimulus occurring at $t = 0$. In the case of f_i, the natural frequency corresponds to a negative real exponential coefficient $(-\alpha_i)$, which sometimes is described as an imaginary frequency [units equal to nepers (e-foldings) per s]. In the case of f_j, the ringing corresponds to a concomitant pair of exponential coefficients that are conjugate complex numbers. Each member of the pair comprises a real frequency component, β_j (radians per s) and an imaginary frequency component, α_j (nepers per s). In the realm of Laplace transforms, these time-domain functions become

$$F_i(s) = L\{f_i(t)\} = A_i/(s+\alpha_i) \tag{7}$$

and

$$F_j(s) = L\{f_j(t)\}$$
$$= 2\left[a_j(s+\alpha_j) - b_j\beta_j\right] / \left[(s+\alpha_j)^2 + \beta_j^2\right] \quad (8)$$

where $L\{f(t)\}$ is the Laplace transform of $f(t)$. Each imaginary natural frequency in $f(t)$ is represented in the denominator polynomial of $L\{f(t)\}$ by a negative real zero, which is a negative real pole of the function $F(s) = L\{f(t)\}$. Similarly, each complex natural frequency in $f(t)$ is represented by a complex pole of $F(s)$, with the imaginary component of frequency being represented by the real component of the pole, and vice versa.

An impulse stimulus of unit amplitude is applied to a linear structure at time $t = 0$. The response, $h(t)$, will be

$$h(t) = f_1(t) + f_2(t) + f_3(t) + \ldots \quad (9)$$

where each function on the right-hand side has the same form as either $f_i(t)$ or $f_j(t)$. The Laplace transform of $h(t)$ is

$$H(s) = L\{f(t)\} = F_1(s) + F_2(s) + F_3(s) + \ldots \quad (10)$$

where each term on the right-hand side has the same form as either $F_i(s)$ or $F_j(s)$ and represents either a single negative real pole, z_i,

$$z_i = -\alpha_i \quad (11)$$

or a conjugate pair of complex poles, z_j and z_j^*,

$$z_j = -\alpha_j + i\beta_j \quad (12a)$$

$$z_j^* = -\alpha_j - i\beta_j \quad (12b)$$

of the function $H(s)$. In that case, the amplitude coefficients for the terms on the right-hand sides of Equations 9 and 10 are given by

$$A_i = (s - z_i)H(s)]_{s=z_i} \quad (13)$$

for a negative real pole, and

$$a_j = Re\left\{(s - z_j)H(s)]_{s=z_j}\right\} \quad (14a)$$

$$b_j = Im\left\{(s - z_j)H(s)]_{s=z_j}\right\} \quad (14b)$$

for a complex pole; where

$$Re\{a + ib\} = a \quad (14c)$$

and

$$Im\{a + ib\} = b \quad (14d)$$

Each of these coefficients or pairs of coefficients reflects the ability of an impulse of unit amplitude to elicit a transient response component at the corresponding natural frequency. Next we shall compare these coefficients with those obtained when the stimulus is a sine or cosine ramp rather than an impulse.

A3. Predicted linear responses to sine and cosine ramps

The Laplace transforms of sine and cosine ramps are

$$Y_a(s) = L\{Mt\ \sin[w_s t]\} = 2Mw_s s / \left(s^2 + w_s^2\right)^2 \quad (15a)$$

and

$$Y_b(s) = L\{Mt\ \cos[w_s t]\}$$
$$= M\left(s^2 - w_s^2\right) / \left(s^2 + w_s^2\right)^2 \quad (15b)$$

where M is the slope of the ramp. The Laplace transforms of the responses of the structure of the previous paragraphs to sine and cosine ramps, respectively, are

$$X_a(s) = H(s)Y_a(s) \quad (16a)$$

and

$$X_b(s) = H(s)Y_b(s) \quad (16b)$$

The poles of $X_a(s)$ and $X_b(s)$ include all of those of the function $H(s)$. Additionally, $X_a(s)$ and $X_b(s)$ have duplicate poles at iw_s and $-iw_s$. Therefore, partial-fraction expansions of the two functions can be developed in the following forms:

$$X_a(s) = q_{a1}(s)/(s^2 + w_s^2) + q_{a2}(s)/(s^2 + w_s^2)^2 + K_{a1}F_1(s) + K_{a2}F_2(s) + K_{a3}F_3(s) + \ldots \quad (17a)$$

$$X_b(s) = q_{b1}(s)/(s^2 + w_s^2) + q_{b2}(s)/(s^2 + w_s^2)^2 + K_{b1}F_1(s) + K_{b2}F_2(s) + K_{b3}F_3(s) + \ldots \quad (17b)$$

where $q_{mn}(s)$ is a polynomial in s, with degree equal to or less than 2; and the generic factors K_{ak} and K_{bk} are real or complex numbers. Our task is to find these factors K_{ak} and K_{bk}.

Translated back into the time domain functions, $x_a(t)$ and $x_b(t)$, the terms on the right sides of Equations 17a and 17b yield three classes of response: (1) a step-modulated sinusoid whose frequency is the same as that of the stimulus sinusoid (but whose phase may be different), with its step modulation occurring at the time of onset of the stimulus ramp (i.e., this component has zero amplitude prior to the onset of the stimulus ramp and constant peak-to-peak amplitude thereafter), (2) a ramp-modulated sinusoid whose frequency is the same as that of the stimulus sinusoid (but whose phase may be different), with the onset of its ramp modulation synchronized to the onset of the stimulus ramp, (3) excitation of every natural frequency of the tuned structure, with each such excitation being that which would take place in response to an impulse of appropriate amplitude occurring at the moment of onset of the stimulus ramp. Thus, taken individually, the response component corresponding to each term on the right-hand side of Equation 17a or 17b is easy to envision. On the other hand, it is not at all easy to envision the overall response that emerges from the summation of all of these components. However, by dealing with the components individually, and by considering the overall nature of the tuned structure, we shall be able to develop some insights and some rules of thumb.

We shall begin by finding the factors K_{ak} and K_{bk}, which transform the impulse response amplitudes of the tuned structure into sine and cosine ramp response amplitudes. In this way, we can determine the effectiveness of the onset singularity of the sine or cosine ramp in exciting the natural frequency corresponding to the generic pole z_k, which can be either real or complex.

Since the natural frequencies of our hypothetical structure are not duplicated, we can compute K_{ak} and K_{bk} as follows: if z_k is a negative real pole,

$$K_{ak}A_k = (s - z_k)Y_a(s)H(s)]_{s=z_k} \quad (18a)$$

and if z_k is a complex pole,

$$K_{ak}(a_k + ib_k) = (s - z_k)Y_a(s)H(s)]_{s=z_k} \quad (18b)$$

where A_k, a_k and b_k are the amplitude coefficients for the impulse responses (see Equations 13 and 14a, b). In either case,

$$K_{ak} = Y_a(s)]_{s=z_k} = 2Mz_k w_s/(z_k^2 + w_s^2)^2 \quad (19)$$

for a sine ramp; and, by similar argument,

$$K_{bk} = Y_b(s)]_{s=z_k} = 2M(z_k^2 - w_s^2)/(z_k^2 + w_s^2)^2 \quad (20)$$

for a cosine ramp.

Natural-frequency excitation by sine ramps

In the case of a sine ramp, we apply Equation 19. For excitation of the natural frequency corresponding to the negative real pole, $z_k = -\alpha_k$, we have

$$K_{ak} = -2M\alpha_k w_s/(\alpha_k^2 + w_s^2)^2 \quad (21)$$

Here, K_{ak} is a real number; and the time-domain response (to a sine ramp) corresponding to the negative real pole $-\alpha_k$ is

$$f_{ak}(t) = K_{ak}A_k \exp[-\alpha_k t] \quad \text{when } t \geq 0 \quad (22a)$$

$$f_{ak}(t) = 0 \quad \text{when } t < 0 \quad (22b)$$

This simple exponential response begins at the time of the onset of the ramp that modulates the sinusoid ($t = 0$ in this case). Thus it is the response that would be elicited by an impulse of appropriate amplitude, synchronized with the onset of the ramp.

For excitation of the natural frequencies corresponding to the complex pole, $z_k = -\alpha_k + i\beta_k$, we have

$$K_{ak} = 2M(-\alpha_k + i\beta_k)w_s/\left([-\alpha_k + i\beta_k]^2 + w_s^2\right)^2 \tag{23a}$$

which can be restated as follows:

$$K_{ak} = \left[2M/\beta_k^3\right](-\delta + i)w_s/\left([-\delta + i]^2 + \omega^2\right)^2 \tag{23b}$$

where

$$\delta = \alpha_k/\beta_k \tag{23c}$$

and

$$\omega = w_s/\beta_k \tag{23d}$$

Here, K_{ak} is a complex number:

$$K_{ak} = c_{ak} + id_{ak} \tag{24a}$$

so that the left side of Equation 18b becomes

$$K_{ak}(a_k + ib_k) = g_{ak} + h_{ak} \tag{24b}$$

where

$$g_{ak} = c_{ak}a_k - d_{ak}b_k \tag{24c}$$

and

$$h_{ak} = d_{ak}a_k + c_{ak}b_k \tag{24d}$$

Therefore, the time-domain response (to a sine ramp) corresponding to the pair of complex poles $-\alpha_k + i\beta_k$ and $-\alpha_k - i\beta_k$ is

$$f_k(t) = 2g_{ak}\exp[-\alpha_k t]\cos[\beta_k t] - 2h_{ak} \times \exp[-\alpha_k t]\sin[\beta_k t] \quad \text{when } t \geq 0 \tag{25a}$$

$$f_k(t) = 0 \quad \text{when } t < 0 \tag{25b}$$

This exponentially-decaying sinusoid is the response that would be elicited by an impulse of appropriate amplitude, synchronized with the onset of the ramp.

Although the expression for K_{ak} corresponding to the complex pole $-\alpha_k + i\beta_k$ is complicated, it simplifies when that pole corresponds to a sharply-tuned (high-Q) resonance. In that case

$$\delta \ll 1 \tag{26a}$$

and

$$K_{ak} = i\left(2/\beta_k^2\right)\omega/(\omega^2 - 1)^2 \tag{26b}$$

which is valid except when ω is close to 1.0 (i.e., the stimulus frequency, w_s, is very close to the resonance frequency, β_k). In that case

$$K_{ak} = i\left(2/\beta_k^2\right)/(\omega^2 - 1 - 2i\delta)^2 \tag{27a}$$

which becomes

$$K_{ak} = -i/2\alpha_k^2 \tag{27b}$$

when ω is 1.0. This last expression, then, applies to a high-Q (second-order) resonance excited by a ramp-modulated sine wave whose frequency is just equal to that of the resonance.

Natural-frequency excitation by cosine ramps

In the case of a cosine ramp, we apply Equation 20. For excitation of the natural frequency corresponding to the negative real pole, $z_k = -\alpha_k$, we have

$$K_{bk} = M\left(\alpha_k^2 - w_s^2\right)/\left(\alpha_k^2 + w_s^2\right)^2 \tag{28}$$

Here, K_{bk} is a real number; and the time-domain response (to a cosine ramp) corresponding to the negative real pole $-\alpha_k$ is

$$f_{bk}(t) = K_{bk}A_k\exp[-\alpha_k t] \quad \text{when } t \geq 0 \tag{29a}$$

$$f_{bk}(t) = 0 \quad \text{when } t < 0 \tag{29b}$$

Again, this is the response that would be elicited by an impulse of appropriate amplitude, synchronized with the onset of the ramp.

For excitation of the natural frequencies corresponding to the complex pole, $z = -\alpha_k + i\beta_k$, we have

$$K_{bk} = M\left\{(-\alpha_k + i\beta_k)^2 - w_s^2\right\} / \left([-\alpha_k + i\beta_k]^2 + w_s^2\right)^2 \quad (30a)$$

which can be restated as follows:

$$K_{bk} = \left[M/\beta_k^3\right]\left\{(-\delta + i)^2 - w_s^2\right\} / \left([-\delta + i]^2 + \omega^2\right)^2 \quad (30b)$$

where δ and ω are defined in Equations 23c and d. Here, K_{bk} is a complex number:

$$K_{bk} = c_{bk} + id_{bk} \quad (31a)$$

so that

$$K_{bk}(a_k + ib_k) = g_{bk} + h_{bk} \quad (31b)$$

where

$$g_{bk} = c_{bk}a_k - d_{bk}b_k \quad (31c)$$

and

$$h_{bk} = d_{bk}a_k + c_{bk}b_k \quad (31d)$$

Therefore, the time-domain response (to a cosine ramp) corresponding to the complex pole $-\alpha_k + i\beta_k$ is

$$f_k(t) = 2g_{bk}\exp[-\alpha_k t]\cos[\beta_k t] - 2h_{bk} \times \exp[-\alpha_k t]\sin[\beta_k t] \quad \text{when } t \geq 0 \quad (32a)$$

$$f_k(t) = 0 \quad \text{when } t < 0 \quad (32b)$$

Again, this is the response that would be elicited by an impulse of appropriate amplitude, synchronized with the onset of the ramp.

When the pole z_k corresponds to a high-Q resonance,

$$\delta \ll 1 \quad (33a)$$

and

$$K_{bk} = -\left(1/\beta_k^2\right)(\omega^2 + 1)/(\omega^2 - 1)^2 \quad (33b)$$

which is valid except when the stimulus frequency (w_s) is very close to the resonance frequency (β_k). In that case,

$$K_{bk} = -\left(1/\beta_k^2\right)(\omega^2 + 1)/(\omega^2 - 1 - 2i\delta)^2 \quad (34a)$$

When the ramp-modulated stimulus frequency is equal to the resonance frequency (i.e., $\omega = 1.0$),

$$K_{bk} = 1/2\alpha_k^2 \quad (34b)$$

A4. Onsets and offsets with other shapes

Consider a tone modulated by the general function $R(t)u(t - t_i)$, which is zero for all values of t less than t_i. If $R(t)$ is n times differentiable for all values of t greater than t_i, then one can expand it into a Taylor's series with terms up to degree n; and that series will approximate $R(t)u(t - t_i)$ in the interval $t_i < t < \infty$.

$$R(t)u(t - t_i) = \left[M_o + M_1(t - t_i) + M_2(t - t_i)^2 + \ldots + M_n(t - t_i)^n + \text{remainder}\right]u(t - t_i) \quad (35)$$

Invoking the principle of superposition, we can deduce the response of a linear filter to each term, then add all such responses to yield an estimate of the response to the complete waveform. For that purpose we need the Laplace transforms of

$$f_{ak}(t) = \begin{cases} M_{ak}t^k \sin[wt] & \text{when } t \geq 0 \\ 0 & \text{when } t < 0 \end{cases} \quad (36a)$$

and

$$f_{bk}(t) = \begin{cases} M_{bk}t^k \cos[wt] & \text{when } t \geq 0 \\ 0 & \text{when } t < 0 \end{cases} \quad (36b)$$

These can be found in published tables of Laplace transforms (e.g., see Roberts and Kaufman, 1966).

However, in their generic forms they are complicated functions, tedious to manipulate. With the availability of algebraic analysis programs, one can generate and display these transforms on a computer – by beginning with the Laplace transforms for step-modulated sine and cosine waves

$$g_{a0}(t) = \begin{matrix} \sin[wt] & \text{when } t \geq 0 \\ 0 & \text{when } t < 0 \end{matrix} \quad (36c)$$

$$L\{g_{a0}\} = w/(s^2 + w^2) \quad (36d)$$

$$g_{b0}(t) = \begin{matrix} \cos[wt] & \text{when } t \geq 0 \\ 0 & \text{when } t < 0 \end{matrix} \quad (36e)$$

$$L\{g_{b0}\} = s/(s^2 + w^2) \quad (36f)$$

and applying the following general property of the Laplace transform:

$$L\{tf(t)\} = -dF(s)/ds \quad (36g)$$

where

$$F(s) = L\{f(t)\} \quad (36h)$$

For $k = 0$ to 4 we have the following: Step-modulated sinusoid beginning at $t = 0$,

sine $$L\{f_{a0}(t)\} = M_{a0}w/(s^2 + w^2) \quad (37a)$$

cosine $$L\{f_{b0}(t)\} = M_{b0}s/(s^2 + w^2) \quad (37b)$$

Ramp-modulated sinusoids beginning at $t = 0$,

sine $$L\{f_{a1}(t))\} = 2M_{a1}ws/(s^2 + w^2)^2 \quad (37c)$$

cosine $$L\{f_{b1}(t))\} = M_{b1}(s^2 - w^2)/(s^2 + w^2)^2 \quad (37d)$$

Sinusoids modulated by t^2 beginning at $t = 0$,

sine $$L\{f_{a2}(t)\} = 2M_{a2}w(3s^2 - w^2)/(s^2 + w^2)^3 \quad (37e)$$

cosine $$L\{f_{b2}(t)\} = 2M_{b2}s(s^2 - 3w^2)/(s^2 + w^2)^3 \quad (37f)$$

Sinusoids modulated by t^3 beginning at $t = 0$,

sine $$L\{f_{a3}(t)\} = 24M_{a3}sw(s^2 - w^2)/(s^2 + w^2)^4 \quad (37g)$$

cosine $$gaL\{f_{b3}(t)\} = 6M_{b3}(s^4 - 6w^2s^2 + w^4)/(s^2 + w^2)^4 \quad (37h)$$

Sinusoids modulated by t^4 beginning at $t = 0$,

sine $$L\{f_{a4}(t)\} = 24M_{a4}w(5s^4 - 10w^2s^2 + w^4)/(s^2 + w^2)^5 \quad (37i)$$

cosine $$L\{f_{b4}(t)\} = 24M_{b4}s(s^4 - 10w^2s^2 + 5w^4)/(s^2 + w^2)^5 \quad (37j)$$

The transient response corresponding to each modulation term on the right-hand side of Equation 35 and each term in the partial-fraction expansion of a linear filter's transfer function will take the form of an impulse response beginning at t_i, with an amplitude computable by application of Equation 19 or 20. Thus, to find the frequency-dependent amplitude factors for transient responses to abruptly initiated (step-modulated) sine and cosine waves, we substitute the right-hand sides of Equation 37a and 37b for $Y(s)$ in Equations 19 and 20; for sine and cosine waves modulated by the factor $(t - t_i)^2u(t - t_i)$, we substitute the right-hand sides of Equations 37e and 37f for $Y(s)$ in Equations 19 and 20; and so forth. Having done this, one derives expressions for the dependence of the amplitude of natural-frequency excitation on the frequency of the modulated sinusoid.

Acknowledgements

We are grateful to David A. Feld, Eva Poinar, John M. Price and Robert J. Sweet for assistance with various aspects of this work. The research

was supported by grants NS 12359 and NS 21573 from the National Institute of Neurological and Communicative Disorders and Stroke.

References

Antonelli, A. and Grandori, F. (1984) Some aspects of the auditory nerve responses evoked by tone bursts. Br. J. Audiol. 18, 117–126.

Baylor, D.A., Hodgkin, A.L. and Lamb, T.D. (1974) The electrical responses of turtle cones to flashes and steps of light. J. Physiol. (London) 242, 685–727.

Chamberlain, S.C. (1977) Neuroanatomical aspects of the gerbil inner ear: light microscope observations. J. Comp. Neurol. 171, 193–204.

Fuortes, M.G.F., and Hodgkin, A.L. (1964) Changes in time scale and sensitivity in the ommatidia of limulus. J. Physiol. (London) 172, 239–263.

Geisler, C.D. and Sinex, D.G. (1982) Responses of primary auditory fibers to brief tone bursts. J. Acoust. Soc. Am. 72, 781–794.

Grandori, F. (1979) Interpretation of the whole-nerve action potential off-effect in response to tone bursts. Audiology 18, 109–118.

Guillemin, E.A. (1957) Synthesis of Passive Networks, Wiley, New York, pp. 1–4.

Hafter, E.R., Dye, R.H., Jr. and Nuetzel, J.M. (1980) Lateralization of high-frequency stimuli on the basis of time and intensity. In: G. van den Brink and F.A. Bilsen (Eds.), Physiological and Behavioral Studies in Hearing, Delft University Press, Delft, The Netherlands, pp. 393–400.

Henry, K.R. and Lewis, E.R. (1988) Cochlear nonlinearities implied by the differences between transient responses to onsets and offsets of a tone burst. In: J.P. Wilson and D.T. Kemp (Eds.), Mechanics of Hearing, Plenum Press, New York, in press.

Katz, B., and Schmitt, O.H. (1940) Electric interaction between two adjacent nerve fibers. J. Physiol. (London) 97, 471–488.

Lewis, E.R. (1987) Speculations about noise and the evolution of vertebrate hearing. Hear. Res. 25, 83–90.

Lewis, E.R. and Henry, K.R. (1988) Cochlear axon responses to tonal offsets: near linear effects. In: H. Duifhuis, J.W. Horst and H.P. Wit (Eds.), Basic Issues in Hearing, Academic Press, London, pp. 177–184.

Moore, B.C.J. (1973) Frequency difference limens for short-duration tones. J. Acoust. Soc. Am. 54, 610–619.

Rhode, W.S. and Smith, P.H. (1985) Characteristics of tone-pip response patterns in relationship to spontaneous rate in cat auditory nerve fibers. Hear. Res. 18, 159–168.

Roberts, G.E., and Kaufman, H. (1966) Table of Laplace Transforms. Saunders, New York.

Woolf, N.K. and Ryan, A.F. (1985) Ventral cochlear nucleus discharge characteristics in the absence of outer hair cells. Brain Res. 342, 205–218.