

Cochlear nerve acoustic envelope response detection is improved by the addition of random-phased tonal stimuli

Kenneth R. Henry ^{a,*}, Edwin R. Lewis ^b

^a Department of Psychology, University of California, Davis, CA 95616-8686, USA

^b Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720, USA

Received 16 October 2000; accepted 30 January 2001

Abstract

We test Lowenstein's dc bias hypothesis as an alternative mechanism for the phenomenon sometimes called 'stochastic resonance'. Probe stimuli consisting of paired phase-locked tones at frequencies f_1 and f_2 (where $f_2 - f_1 = 800$ Hz, $f_1 > 4.5$ kHz) and at equal intensity were used to generate synchronous 800 Hz cochlear nerve activity (envelope responses). When a background tone of the same intensity, with a frequency halfway between f_1 and f_2 , is presented simultaneously with the probe stimulus, the envelope response amplitude typically decreases. Consistent with Lowenstein's hypothesis, however, when the intensities of the probe and background tone are near the detection threshold of the envelope response (approximately 0–20 dB sound pressure level), the simultaneous presence of the background tone often increases the amplitude of the envelope response. At these same intensity levels, when the background tone precedes the probe stimulus, it decreases the amplitude of the response to the probe stimulus. The effects of simultaneous presentation of the probe and the background tone are frequency-dependent, becoming less pronounced or reversing as the frequency of the background tone departs from those of the probe stimuli. © 2001 Elsevier Science B.V. All rights reserved.

Key words: Stochastic resonance; Phase locking; Dithering; Ac response; Dc response

1. Introduction

Cochlear detection or processing of near-threshold intensities of an acoustic stimulus often is impaired when the stimulus is presented simultaneously with a background sound. This effect, known as simultaneous masking, is especially pronounced if the spectral content of the background sound is close to, and its intensity level is equal to or higher than that of the stimulus. Simultaneous masking has been attributed primarily to

adaptation, with two-tone suppression also playing a role (Dallos and Cheatham, 1976, 1977; Harris, 1979). It is especially apparent with the cochlear nerve onset compound action potential (CAP). The onset CAP evidently corresponds to spike synchrony among a subpopulation of previously resting cochlear axons at the onset of a stimulus tone burst. In that case, masking also may be attributed to preemptive excitation of that subpopulation, leaving its members in miscellaneous states of refractoriness at the time of the onset event – a form of synchrony suppression. These various factors are believed to be responsible for the V-shaped appearance of simultaneous masker tuning curves (TCs), which are generated with tonal probe stimuli and maskers.

By contrast, for many years sensory neurophysiologists have considered the possibility that, under certain circumstances, internal or external background stimuli might enhance rather than degrade the responsiveness of a sensory unit to a very weak probe stimulus. These considerations appear to have been based on two some-

* Corresponding author. Tel.: +1 (530) 752-1854; Fax: +1 (530) 752-2087; E-mail: krhenry@ucdavis.edu

Abbreviations: ac, a rapidly changing measure (from the electrical term, 'alternating current'); C, centigrade; CAP, cochlear nerve compound action potential; CF, characteristic frequency; CV, coefficient of variation; dc, a slowly changing measure (from the electrical term, 'direct current'); DFT, discrete Fourier transform; f_1 , the lower frequency of a pair of tones; f_2 , the higher frequency of a tonal pair; rms, root mean square; RW, round window of the cochlea; SPL, sound pressure level; TC, tuning curve

what different schemes. Both schemes posit a biasing effect on the unit's otherwise silent spike trigger, giving it a substantial background firing rate. That firing rate then could be modulated by the response to a second stimulus that by itself would have been below the threshold of the spike trigger. One scheme (Lowenstein, 1956) is based on dc biasing, analogous to the way active electronic devices such as triode vacuum tubes or transistors are biased by dc voltages or currents to make them responsive to weak voltage or current signals superimposed on the bias. In response to an internal or external background stimulus, for example, one can envision a positive dc current flowing into the unit's spike trigger region, leading to background spike production. Constant current applied to traditional spike trigger models produces periodic spikes (e.g. see Agin, 1964; Stein, 1967). The other scheme (Stein, 1970; French and Stein, 1970; Stein et al., 1972; Yu and Lewis, 1989) is based on what the audio engineering community now calls dithering (Vanderkoy and Lipshitz, 1984). The background stimulus is a random ac signal (band-limited noise). The randomly fluctuating response to this stimulus frequently exceeds threshold at the unit's spike trigger, leading to a random spike train that can be modulated by a second stimulus. French and Stein (1970) demonstrated the effect with integrate-and-fire trigger models; subsequently, it was shown by Fauve and Heslot (1983) with a Schmitt trigger circuit. Yu and Lewis (1989) demonstrated it computationally with the Hodgkin–Huxley model.

In principle, a background of perfectly periodic spikes is noiseless. Observing a train of such spikes with an arbitrarily sensitive detector, one could detect an arbitrarily small perturbation of even one spike period. Owing to internal noise that must be present in any sensor operating far from 0.0 K, however, one expects even background spike trains produced by dc biasing to have some degree of randomness. This would be reflected, for example, by a non-zero coefficient of variation (CV) of the spike period. If one takes the mean spike interval to be the parameter of the spike train (i.e. the signal) that represents the dc bias, then the random deviations from that mean interval can be taken to be noise. The CV of the spike interval then becomes the ratio of the root mean square (rms) amplitude of the noise (i.e. the rms deviation of the spike interval from its mean) divided by the signal (the mean interval). Primary afferent axons of the vertebrate vestibular system exhibit a wide range of noisiness in background spike activity (spike-interval CVs ranging from the neighborhood of 0.1 to the neighborhood of 1.0 (e.g. see Goldberg and Fernandez, 1971)). Because he worked with that system, Lowenstein must have encountered a substantial population of units with low-CV spontaneous activity. In such units, the implication

of an underlying dc bias is clear. The background spike activities of vertebrate auditory units almost always are noisy, with spike-interval CVs in the neighborhood of 1.0. Thus, although the sort of dc biasing proposed by Lowenstein may well be occurring in such units, their background spike activity is strongly dithered by noise – in the manner of French and Stein (1970).

In response to acoustic stimuli comprising pure tones, mammalian inner hair cells are known to exhibit both dc and ac responses (Russell and Sellick, 1978). At the level of the eighth nerve, these evidently are reflected, respectively, in a positive shift in the mean spike rate and in a tendency for spikes to be phase-locked to a phase-shifted version of the stimulus sine wave (Kiang et al., 1965; Rose et al., 1967). Observed at this same level, ac responses to more general stimuli tend to be phase-locked to linearly-filtered versions of the stimulus waveforms (de Boer and de Jongh, 1978; Wolodkin et al., 1996; Yamada et al., 1996); dc responses to more general stimuli tend to be phase-locked to the squares of the envelopes of the linearly-filtered stimulus waveforms (Yamada and Lewis, 1999; Lewis et al., 2001). Each of the latter approximates a truly dc condition when the amplitude of the stimulus envelope is constant – as it is for a pure tone stimulus of constant amplitude. For tonal stimuli below approximately 2 kHz, ac responses typically are very much larger than dc responses. The reverse is true for tonal stimuli greater than approximately 4 kHz; in fact, ac responses in terms of phase-locked spike activity become so small that they are difficult to detect for stimulus frequencies equal to or greater than approximately 5 kHz. It follows that near-threshold broad-band stimuli, such as broad-band white noise, will produce predominantly ac responses in cochlear units with characteristic frequencies (CFs) below 2 kHz and predominantly dc responses in cochlear units with CFs above 4 kHz. In Lewis and Henry (1995), we showed that stepwise increases in the amplitude of a broad-band white noise stimulus can markedly enhance the sensitivity of both low- and high-CF gerbil cochlear units to low-frequency tonal stimuli. For the low-CF units (e.g. the 800 Hz unit of figure 10 in that paper), the result can be viewed as a consequence of dithering, a test of the French–Stein hypothesis. For the high-CF units (e.g. the 11.5 kHz unit of figure 8 in that paper), the result might be viewed as being largely a consequence of dc bias, a test of the Lowenstein hypothesis. Although it clearly has a large dc component, however, the envelope of broad-band white noise also has a random ac component. This inevitably will impose some dithering effects – as posited by the French–Stein hypothesis. Furthermore, in units exhibiting dc responses to constant-amplitude, high-frequency tonal stimuli, spike-interval CVs remain large even though the amplitude of the

stimulus envelope is constant. As it is with spontaneous spike activity, this must be a reflection of dithering by internal noise.

In Lewis and Henry (1995), the enhancement of sensitivity to low-frequency tones often was transitory, occurring immediately after the onset of the noise stimulus, but soon being erased by adaptation to that stimulus. Its presence also required that the amplitude of the noise was not so large as to produce a saturated dc response in high-CF units or strong two-tone synchrony suppression in low-CF units (Javel, 1981; Greenwood, 1986). In fact, combining the dithering hypothesis of French and Stein with the synchrony suppression observations and theory of Javel and Greenwood, one concludes that there must be a noise level that produces optimum dithering (i.e. maximum enhancement of sensitivity to a given tonal stimulus). Although not truly a resonant effect at all, this hypothetical optimum has been labeled 'stochastic resonance' (Weisenfeld and Moss, 1995).

Interest in stochastic resonance and in the notion that processing of near-threshold stimuli can be enhanced by the presence of low levels of internally or externally generated noise has led to several experimental studies with sensory systems. Longtin et al. (1991) and Maddox (1991) suggested that phase-locked firing of auditory neurons would provide an opportunity for expressing the properties of stochastic resonance. The addition of low levels of noise has been found to improve information transfer in first-order neurons receiving input from mechanoreceptors of the crayfish (Douglass et al., 1993) and rat (Collins et al., 1996), as well as from thermoreceptors of the shark (Braun et al., 1994). The effect also has been demonstrated in psychophysical somatosensory studies (Chialvo and Apkarian, 1993; Simonotto et al., 1997). Computational and sciatic nerve models of the cochlear nerve have been used to demonstrate how the addition of noise can improve the response of a cochlear implant (Morse and Evans, 1996, 1999; Rubenstein et al., 1999; Bruce et al., 1999). Jaramillo and Wiesenfeld (1998) and Ehrenberger et al. (1999) suggested that stochastic resonance enhances the near-threshold afferent activity of the cochlear nerve, citing Brownian motion within the cochlea as a possible source of dithering noise. Ideally, one would want the noise dithering each member of a population of spike triggers to be independent of that dithering the other members. Otherwise, it would be difficult for the central nervous system to extract the effects of weak signals from the temporal pattern established by the dithering noise across the entire population. Direct evidence for the ability of acoustic noise to improve cochlear nerve population responses (whole-nerve responses) to near-threshold acoustic stimuli has been obtained from the gerbil (Henry, 1999). Whole-nerve responses to the low-

frequency modulation of high-frequency stimuli were enhanced by the presence of broad-band noise. This effect only occurred when the intensity levels of the noise ranged from 5 to 30 dB sound pressure level (SPL) and those of the stimuli were close to response threshold (−5 to 25 dB SPL).

All of the studies of the previous paragraph were focussed on the French–Stein hypothesis; dithering by the background noise was presumed to be the principal contributor to sensitivity enhancement. Nevertheless, it is possible that dc bias may also have played some role. In the experiments to be reported in this paper, the methods described in Henry (1999) are employed again, but the focus is expressly on the Lowenstein hypothesis; background noise was replaced by background tone at a frequency above the cutoff for phase-locked ac response. This removed both the random ac response to the background stimulus and the random ac component of the background stimulus envelope-leaving dc response as the source of any observed sensitivity enhancement. It also made it possible to explore rather easily the dependence of sensitivity enhancement on the frequency of the background stimulus.

2. Materials and methods

2.1. Subjects

Thirty one Mongolian gerbils (*Meriones unguiculatus*), aged 45–150 days, were screened as subjects for these experiments. Twelve of these showed a response to tonal masking similar to that seen in Fig. 3, and were not examined further. The remaining 19 gerbils showed a statistically significant increase in the amplitude of the cochlear nerve envelope response when a low intensity tone of an appropriate frequency was simultaneously presented with a low level stimulus. Each illustration in this paper represents an effect that was examined and found to be replicable in at least three animals. All the subjects were otoscopically normal and had cochlear nerve envelope response thresholds that ranged from approximately −5 to 20 dB SPL. All animals were born and maintained in a gerbil colony with restricted levels of ambient noise.

2.2. Animal preparation

Gerbils were anesthetized with a mixture of xylazine and ketamine (20 mg/kg and 60 mg/kg, intraperitoneal, respectively), and supplemental doses were administered as necessary. The cochlea was approached ventrally in a manner described previously (Henry, 1996a). Body temperature was monitored by a rectal thermocouple and actively maintained at $37 \pm 1^\circ\text{C}$.

2.3. Acoustic environment

Testing was conducted within a double-walled acoustically insulated chamber. Before data were taken, each subject remained in this chamber for approximately 30 min in order to allow its auditory periphery to become more sensitive. The inner chamber rested atop a vibration isolation table that was composed of three second-order filter stages. This structure was inside of an Industrial Acoustic Corporation 403A acoustic room. Over the range of 300 Hz–20 kHz, the total attenuation was greater than 65 dB (Henry and Lewis, 1992). The ambient noise within this double-walled acoustic chamber was below the noise level of the microphone (ER10B, see below); e.g. < -11 dB SPL at 2 kHz over a bandwidth of 3.75 Hz.

2.4. Recording technique

Approximately 1 ml of artificial perilymph was placed into the antrum of the cochlear round window (RW), into which a 100 micron diameter silver electrode was placed. The reference electrode was a stainless steel bar pressed against the roof of the mouth. RW electrical activity was amplified (10^5) and filtered (300–15 000 Hz, 48 dB/octave band edge slopes). The amplified and filtered waveform was recorded over a window of 40 ms with a resolution of 10 μ s/address. For repeated, very low intensity (< 10 dB SPL) stimuli, response-waveform averages were taken over 1000 stimulus presentations. When higher stimulus levels were used, 200 or fewer stimulus presentations were typically sufficient to obtain robust and consistent averaged responses. In order to compensate for acoustic and neural delays, a delay was introduced between the electronic triggering of each stimulus presentation and the beginning of the 40 ms recording window.

2.5. Acoustic stimuli (Fig. 1)

The stimulus period was 101 ms. The probe and background stimuli were delivered separately (see Section 2.6). Each stimulus was digitally created, shaped, mixed, and timed. After conversion to analog signal, it was subjected to power amplification, with fixed gain. The amplitude of the delivered stimulus was adjusted entirely by attenuation immediately prior to being sent to the acoustic driver. The temporal relationships of the dual tone probe stimulus and the background tone stimulus in these experiments are shown in Fig. 1. The probe stimulus comprised the sum of two sine-wave bursts, each beginning with the condensation phase and having 1 ms Gaussian rise and fall times. The frequency (f_1) of the lower-frequency sine wave was 800 Hz below that (f_2) of the higher, leading to

an 800 Hz modulation of the amplitude of envelope of the combined waveform (lower waveforms in Fig. 1A,B). The background tone (upper waveforms in Fig. 1A,B) was presented with random onset phase, so that it produced no net cochlear microphonic potential in the averaged RW response. With the exception of the experiment described in Figs. 9 and 10, the frequency of the background tone was the midpoint of the frequencies of the dual probe stimuli.

Two 1 mm diameter, 0.5 m long plastic tubes were tightly sealed, along with an Etymotic ER10B probe microphone, to the right ear of the gerbil. Each tube was attached to an Etymotic ER2 acoustic driver. One tube was used to deliver the probe stimulus; the other was used to deliver the background tone. The probe microphone was used to assess the net acoustic stimulus as well as ambient sounds (primarily heart beats), all within the external auditory canal. The output of the probe microphone was continuously analyzed over an 80 dB dynamic range by a Hewlett Packard 3561A Dynamic Signal Analyzer. Although the 800 Hz f_2-f_1 acoustic envelope frequency was conspicuous in the time domain, it was not present at -40 dB (re stimulus level) in the discrete Fourier transfer (DFT) of the acoustic stimulus.

2.6. Procedure

The basic procedure consisted of obtaining RW-recorded responses in the presence of the probe stimulus alone, followed by recording responses to the probe in the presence of the background tone. Because the polarities of the phase-locked dual tone probe stimuli were not alternated, averaged cochlear microphonics were

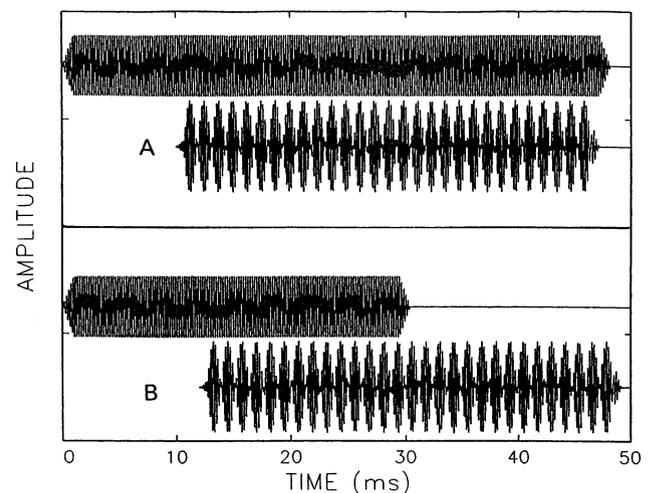


Fig. 1. Relative timing of background tones (upper traces in A and B) and probe stimuli (lower traces). In this case, the frequency of the background tone was 5.0 kHz; the probe stimulus comprised 4.6 kHz and 5.4 kHz tones of equal amplitude, both starting in condensation phase.

obtained at frequencies corresponding to f_1 and f_2 . Because the polarity of the background tone polarity was randomly varied, that stimulus did not generate an averaged cochlear microphonic.

The responses to the envelope of the probe stimulus could be readily identified by their spectral signatures of 800 Hz, corresponding to $f_2 - f_1$. The forward masking TC of an envelope response of this type consistently exhibits a sharply-tuned tip around f_1 and f_2 , demonstrating that it originates at a restricted region along the cochlea (Henry, 1996b). Such envelope responses also consistently are blocked by application of tetrodotoxin, demonstrating that they are neural in origin and correspond to activity of a population of electrically-excitable sodium channels (Henry, 1995). Therefore, we take them to reflect spike activity confined to a local population of cochlear nerve fibers and phase-locked to the envelope of the probe stimulus waveform (Henry, 1996a,b).

Although amplitude input–output functions often are obtained by presentations of random sequences of stimulus intensities, the procedures used in these experiments were designed to maintain the animal's ear in as sensitive a condition as possible when it was being tested with very low intensity stimuli. Thresholds of the envelope response were determined by stimulating the animal with a series of probe stimuli with the amplitude beginning at 30 or 35 dB SPL and decreasing in 5 dB steps to 15 dB SPL. Beyond that level, each step was accompanied by a pause of approximately 15 min in order to allow the ear to recover sensitivity loss induced by the previous stimuli. As stimulus intensity was reduced, the magnitudes of responses became increasingly difficult to evaluate visually, necessitating offline analysis. This was achieved with a DFT of the averaged response (Figs. 5 and 7). Threshold was estimated as being the lowest level of stimulus that produced DFT amplitude response, at $f_2 - f_1$, that was visibly greater than that at any other frequency (other than f_1 and f_2), over the 20 kHz range of the DFT. Thresholds estimated in this way were very close to the published behavioral thresholds for the gerbil (Ryan, 1976; Henry, 1996b). By the end of these procedures and evaluations, the gerbil had been in a very quiet environment for approximately 30–45 min. These conditions contributed to the sensitivity of the envelope response, and allowed these near-threshold effects to be more readily observed.

At this step of the procedure, utilizing a near-threshold stimulus level, the averaged RW response was recorded when the probe stimulus was presented alone and, immediately after that, when it was presented in the presence of the background tone at a given level and frequency. A DFT was performed on each averaged response and the 800 Hz ($f_2 - f_1$) envelope response

amplitude was measured. For each probe stimulus and background tone setting, this pair of recordings and measurements was repeated eight times for statistical purposes. Mean values, standard errors and *t*-tests were then obtained for the difference between the envelope response amplitude in the absence of the background tone and that in the presence of the background tone. For plots of envelope response amplitude versus stimulus level, this process was repeated with the amplitude of the probe stimulus and/or background tone increased by 5 dB. This was continued at increasing stimulus and/or background tone levels until the experiment was completed. For plots of envelope response amplitude versus the frequency of the background tone, the intensity levels of the probe stimulus and background tone were kept constant, but the frequency of the background tone was varied from one pair of recordings to the next. This was done to avoid adaptation or fatigue with any single background tone frequency.

3. Results

Fig. 2 shows typical envelope responses from the

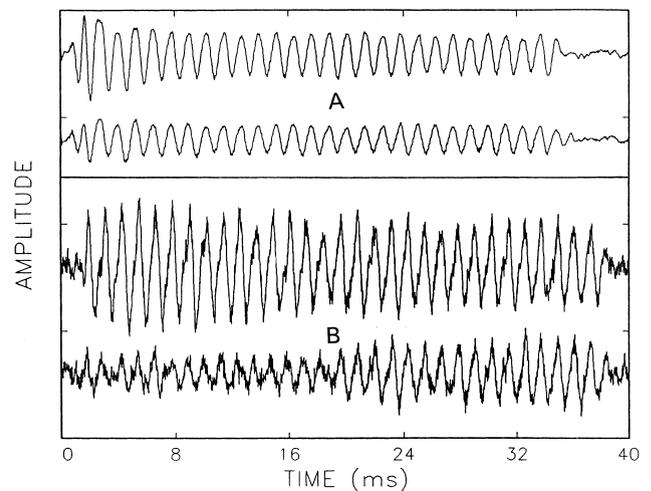


Fig. 2. A: Top waveform: 800 Hz cochlear nerve envelope response to 25 dB (SPL) probe stimulus comprising 7.6 kHz and 8.4 kHz tones. The amplitude of this response decreases over time, showing adaptation characteristic of cochlear nerve activity. Bottom waveform: response to the same probe presented in the presence of an 8 kHz, 25 dB SPL background tone (timing as in Fig. 1A) (gerbil SR052499). Reduction in amplitude is what one expects from simultaneous masking. B: Top waveform: cochlear nerve envelope response to 25 dB (SPL) probe stimulus comprising 9.6 kHz and 10.4 kHz tones. High-frequency irregularities on the waveform are cochlear microphonics. Bottom waveform: response to same probe stimulus with 25 dB, 10 kHz background tone present (timing as in Fig. 1B). The first half of the response shows the effect of simultaneous masking; the second half shows the effect of forward masking (SR112098).

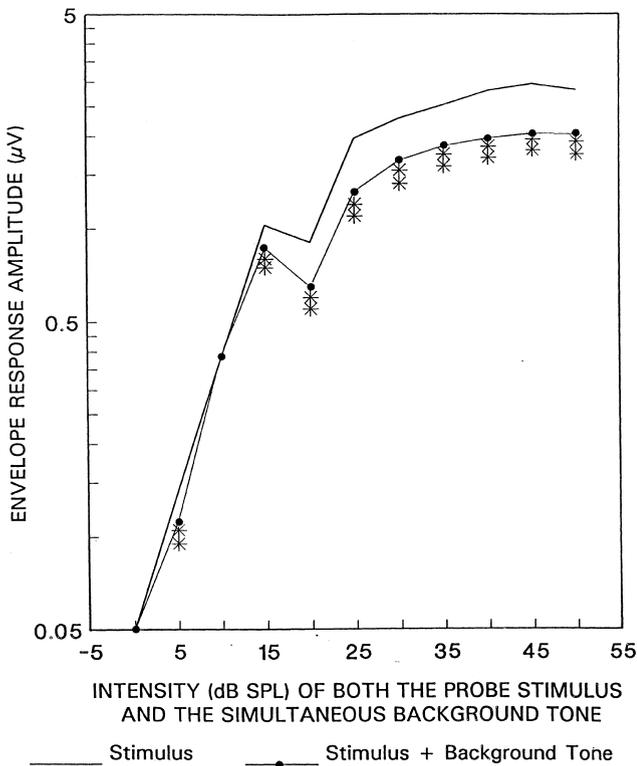


Fig. 3. Envelope response amplitude vs. stimulus amplitude, with and without background tone (timing relationship of Fig. 1A) in one gerbil subject. At intensities equal to or greater than 15 dB SPL, the background tone decreases the amplitude of the envelope response. The simultaneous masking effect is inconsistent at lower intensities. This near-threshold response pattern was seen in 12 of the 31 gerbils screened for this study. In this and subsequent illustrations, the presence of a single * indicates that the background tone altered the amplitude of the envelope response with a two-tailed $P < 0.05$; ** indicates a two-tailed $P < 0.005$ (SR052499).

gerbil cochlear nerve (RW recordings). The probe stimulus amplitude for each of the four traces was 25 dB SPL. For Fig. 2A, the probe stimulus comprised 7.6 kHz and 8.4 kHz components; for Fig. 2B it comprised 9.6 kHz and 10.4 kHz components. The recording in the upper trace in each panel shows the response to the probe stimulus alone. That in the lower trace shows the response to the same probe stimulus presented with a background tone stimulus at 25 dB SPL and at a frequency midway between those of the probe components (i.e. 8 kHz for Fig. 2A, 10 kHz for Fig. 2B). The timing of the background tone for Fig. 2A was that shown in Fig. 1A (i.e. the background tone stimulus was present throughout the duration of the probe stimulus). At the beginning of the upper trace in Fig. 2A, one sees adaptation in the response to the probe alone; in the lower trace, one sees conspicuous reduction in amplitude (simultaneous masking) of the probe stimulus response. The timing of the background tone for Fig. 2B was that shown in Fig. 1B (the background tone ended midway through the presentation of the probe stimulus).

Again, the upper trace shows some adaptation to the probe alone. The first half of the lower trace in Fig. 2B shows reduction in amplitude (simultaneous masking); and this reduction in amplitude continues (as forward masking) in the second half of the trace. Toward the end of the trace, one sees a gradual recovery from forward masking. The envelope response is generated by activity of cochlear nerve axons having CFs close to the two tones that compose the probe stimulus (Henry, 1996b); therefore, it is not surprising that a third (background) tone, with a frequency halfway between those of f_2 and f_1 and a random phase relationship with the stimulus, is capable of both simultaneous and forward masking the envelope response.

This masking effect is not uniform across the dynamic range of the envelope response. The background tone reduces envelope response amplitude by a fairly consistent proportion when the probe stimulus and the tone are both more than approximately 15 or 20 dB above the envelope detection threshold. This relationship breaks down, however, when the intensity of the stimulus and tone are closer to threshold. The input-output

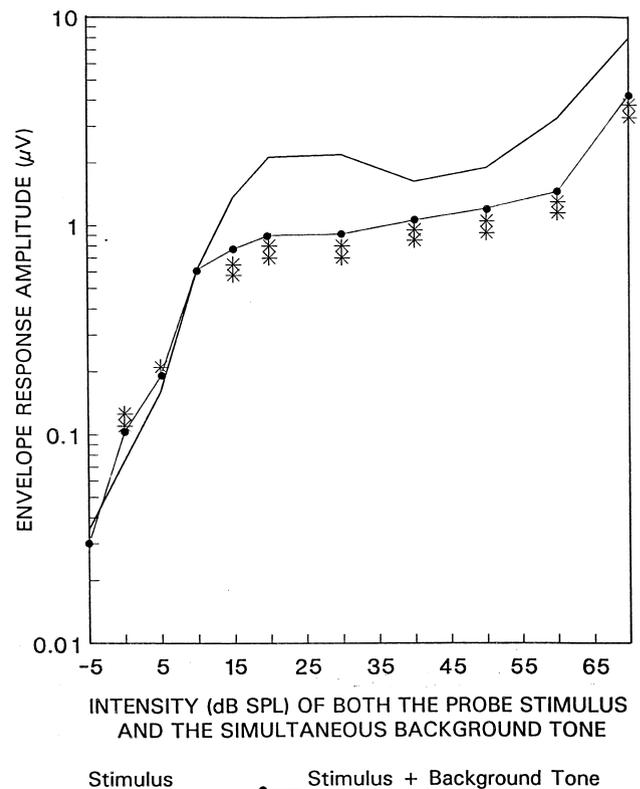


Fig. 4. Envelope response amplitude vs. stimulus amplitude, with and without background tone (timing relationship of Fig. 1A) (same conditions as those of Fig. 3, in a different gerbil subject). As in Fig. 3, a simultaneous masking effect was seen when the probe stimulus and extra tone intensities were greater than 15 dB SPL. In this gerbil, however, the effect was reversed at levels of 0 and 5 dB SPL. Variations of this near-threshold response pattern were seen in 19 of the 31 gerbils screened for this study (SR052899).

curve of Fig. 3 shows a pattern in response to near-threshold stimuli that was found in 40% (12/31) of the gerbils screened for this paper. In these animals, the introduction of the background tone at near-threshold levels resulted either in a very small decrement or in no change at all in the amplitude of the envelope response.

In the majority (19/31) of the gerbils, simultaneous presentation of the background tone and the probe stimulus resulted in an increase in the amplitude of the envelope response. Fig. 4 shows an input–output curve with a near-threshold feature that is typical of these animals. Fig. 5 illustrates the means used to quantify these changes. These effects were not observed in cochlear microphonics recorded from the same electrode. Fig. 6 illustrates this from another gerbil. Increases in the amplitude of the envelope response were observed only when the probe and background tone stimuli were simultaneously present. When the timing of the two stimuli was that of Fig. 1B, only reduction of the envelope response amplitude (forward

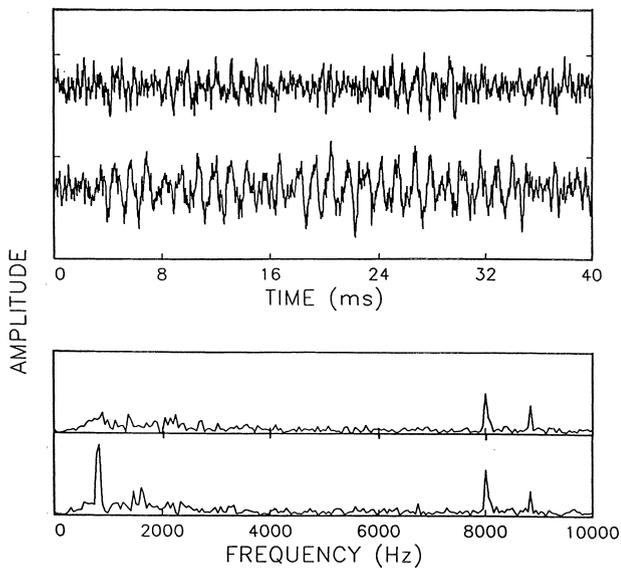


Fig. 5. Time- and frequency-domain presentations of a conspicuous example of enhancement (timing relationships of Fig. 1A). Probe stimulus comprised 8.0 and 8.8 kHz tones; frequency of background tone was 8.4 kHz; probe stimulus and background stimulus amplitudes were 5 dB SPL. Top trace, upper box: averaged RW response in absence of the background tone. Bottom trace, upper box: averaged RW response in presence of background tone. Top and bottom traces, lower box: amplitude components of the DFTs of the top and bottom traces in the upper box, respectively. Cochlear microphonic responses to the 8.0 and 8.8 kHz probe stimulus components are obvious in both DFTs and seem not to have been altered by the background tone. Because it was presented with random phase, the background tone produced no (8.4 kHz) spectral peak in the DFT of the averaged response. In the presence of the background stimulus, the 800 Hz envelope of the probe stimulus produced a conspicuous spectral peak (13 μ V in this case). In the absence of the background tone, that peak appears to be buried in the noise (SR103098).

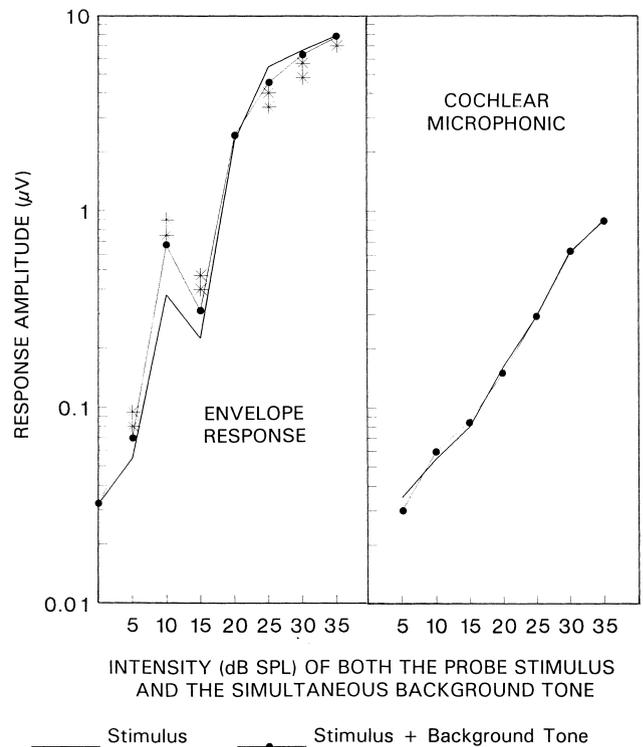


Fig. 6. Left-hand panel: envelope response amplitude vs. probe stimulus intensity, in presence and in absence of background tone, timing relationship of Fig. 1A (same gerbil subject as Fig. 5). Right-hand panel: amplitude of the 8.0 kHz cochlear microphonic component under the same two conditions. When the probe stimulus level was 0 dB SPL, the amplitude of the envelope response was not measurable, and was assigned a value equal to that of the background noise. This was not changed by the presence of the 10 kHz background tone. At slightly higher probe stimulus levels, the presence of the background tone conspicuously increased the amplitude of the envelope response. At still higher levels, it decreased it. By contrast, the cochlear microphonic amplitudes were unaffected by the 8.4 kHz background tone (SR103098).

masking) was observed during the time that the background tone was off (Figs. 7 and 8).

The ability of the background tone to enhance the responsiveness of the cochlea to the probe stimulus envelope at low stimulus levels, as reflected in Figs. 4–8, was only observed in gerbils who had been adapted to the quiet environment in which they were tested. This is why initial testing was begun with the lowest intensities and then carried successively to the highest.

The left-hand panels of Figs. 9 and 10 illustrate the frequency-specific nature of enhancement. Increases in the envelope response were observed only when the frequency of the background tone was close to those of the probe stimulus components. This frequency-specificity is more pronounced when the intensity levels of the stimulus and the background tones are very low. When the timing relationship of Fig. 1B was used, the forward masking seen during the second half of the probe stimulus also depended strongly on the frequency

of the background (masking) tone (right-hand panels in Figs. 9 and 10).

4. Discussion

The carrier frequencies of the probe and background stimuli employed in this study all were equal to or greater than 4.6 kHz. At those frequencies, the components of hair cell receptor potential at the carrier frequencies (i.e. those commonly called ‘ac receptor poten-

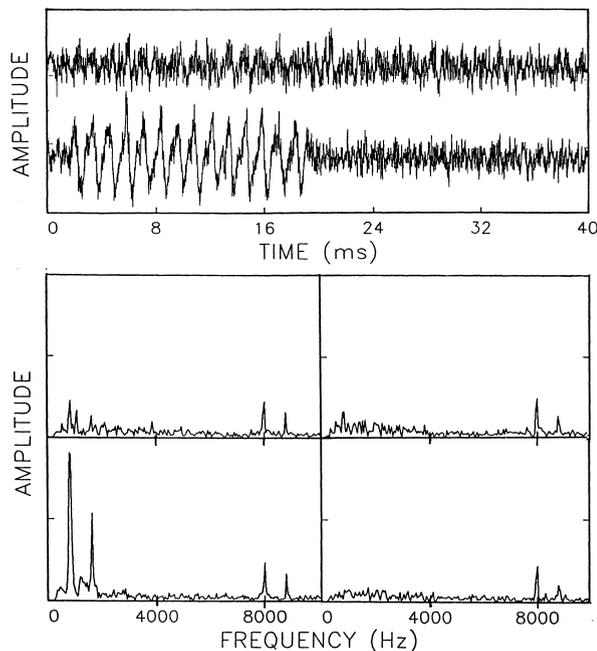


Fig. 7. Time- and frequency-domain presentations of a conspicuous example of enhancement (timing relationship of Fig. 1B). For this subject, the envelope response threshold was 18 dB SPL. Probe stimulus comprised 8.0 and 8.8 kHz tones; frequency of background tone was 8.4 kHz; probe stimulus and background stimulus amplitudes were 25 dB SPL. Top trace, upper box: averaged RW response in absence of the background tone. Bottom trace, upper box: averaged RW response in presence of background tone. Bottom box: DFTs of waveforms in upper box. Upper left frame: DFT of first half of response to probe stimulus alone. Upper right frame: DFT of second half of same response. The 800 Hz envelope response component shows adaptation over time (from 0.225 to 0.16 μ V). Lower left-hand frame: DFT of averaged RW response to probe stimulus in presence of background tone (first half of lower waveform in top box). Lower right-hand frame: DFT of the averaged RW response immediately after the background tone was terminated (second half of lower waveform in top box). During presentation of the background tone there was a large (0.91 μ V) 800 Hz envelope response, accompanied by its harmonic – a substantial (0.62 μ V) 1.6 kHz component. Immediately following the tone, the 800 Hz component was reduced below its levels in the late response to the probe stimulus alone (upper right-hand frame of the bottom panel), indicating a forward masking effect of the background tone (0.16 μ V in the non-masked condition of the upper right-hand panel vs. 0.11 μ V in the forward masked condition of the lower right-hand panel (SR111198)).

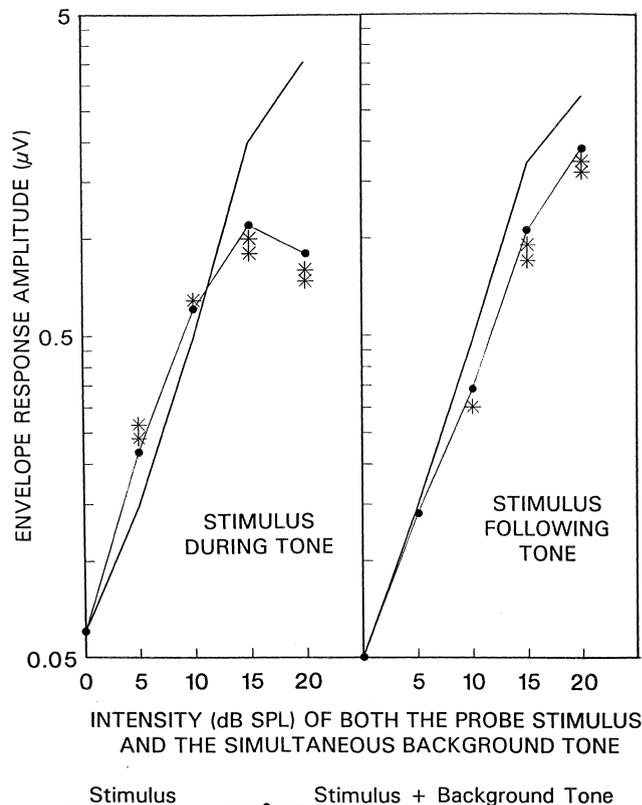


Fig. 8. Left-hand panel: envelope response amplitude vs. probe stimulus intensity, in presence and in absence of background tone, timing relationship of Fig. 1B. The amplitudes of the probe stimulus and the background tone were the same. Left panel: amplitude of the envelope response vs. stimulus intensity during first half of probe stimulus presentation, with and without simultaneous presentation of the background tone. The presence of the background tone increased the amplitudes of the 800 Hz envelope responses generated by 5 and 10 dB (SPL) probe stimuli, and reduced the amplitudes generated by 15 and 20 dB stimuli. Right panel: amplitude of the envelope response vs. stimulus intensity during second half of probe stimulus presentation, with and without presentation of the background tone during the first half of the stimulus. The response following the background tone shows a forward masking effect (SR112098).

tials’) are expected to be highly attenuated; the envelope components (i.e. those commonly called ‘dc receptor potentials’) are expected to dominate. Therefore, the receptor potential in response to the probe stimulus is expected to be a distorted 800 Hz waveform (a distorted version of the envelope of the lower traces in Fig. 1A,B). The receptor potential in response to the background tone (averaged over many presentations of that tone with random phase) is expected to be a dc pulse (duration approximately 48 ms for the upper trace in Fig. 1A, approximately 30 ms for that in Fig. 1B). By definition of linearity, dc receptor potentials are consequences of a non-linear process akin to rectification. This non-linearity may be attributable to the inner hair cell mechanoelectric transduction process (Huds-

peth and Corey, 1977). It also may reflect processes in the outer hair cell (Zenner, 1993). Whatever its source, a Taylor's series description of it would contain only even-order terms.

One expects to see these same response patterns at the level of spike production in the cochlear nerve. At very low stimulus levels, the dc pulse in response to the background tone somehow has produced enhancement of the 800 Hz envelope response to the probe stimulus. By definition of linearity, enhancement of an 800 Hz signal by the presence of a dc signal reflects non-linear interaction. We cannot be certain about the locus of this interaction. As mentioned in Section 1, tests with tetrodotoxin (Henry, 1995) and with forward and simultaneous masking strongly imply that when it is viewed in the averaged RW potential, as it was done here, the 800 Hz envelope response reflects only a synchronized 800 Hz modulation of spike activity over a subpopulation of cochlear axons with CFs close to carrier frequencies (f_1 and f_2) of the probe stimulus. There is no evidence of a cochlear microphonic component in the envelope response. This strongly implies that the non-linear interaction of the ac and dc components occurs after the hair cell mechano-electric transduction process.

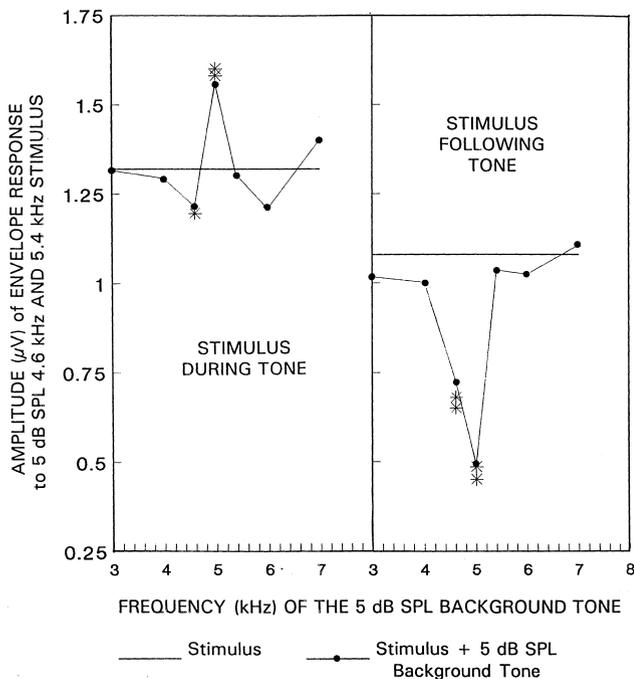


Fig. 9. Envelope response amplitude vs. background tone frequency, timing relationship of Fig. 1B. Probe stimulus: 4.6 kHz and 5.4 kHz, 5 dB SPL. Background tone: 5 dB SPL. When the background tone and probe were simultaneously present, the response amplitude was increased only when the background tone was centered on the two stimulus frequencies (i.e. when the background tone frequency was 5 kHz). When the background tone preceded the probe stimulus, 5 kHz background tone was most effective at forward masking the response (SR032299).

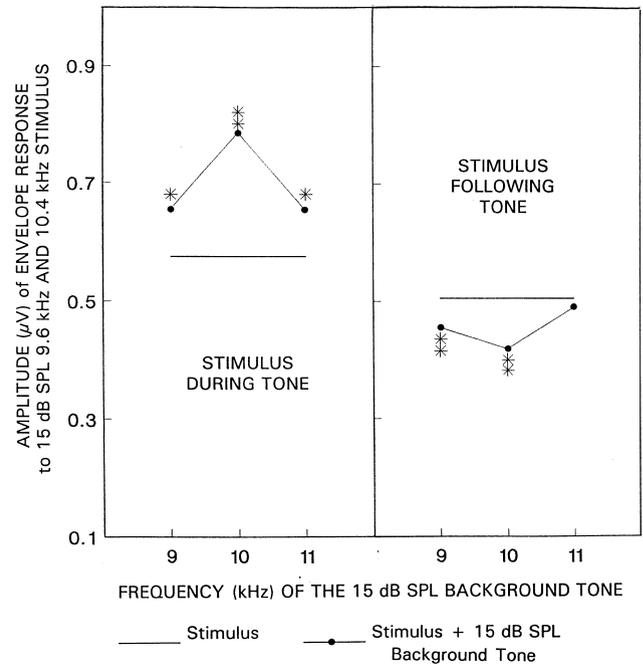


Fig. 10. Envelope response amplitude vs. background tone frequency. Probe stimulus: 9.6 kHz and 10.4 kHz, 15 dB SPL. Background tone: 15 dB SPL. The background tone was most effective as an enhancer (left panel) and as a forward masker (right panel) when its frequency was midway between those of the probe stimulus (SR111998).

There seem to be two obvious possibilities: (1) the locus of non-linear interaction is the presynaptic part of the hair cell afferent synapse, and (2) the locus of non-linear interaction is the spike trigger of the primary afferent axon. At either locus, one can imagine a dc bias on the membrane voltage that is well below threshold (well below threshold for transmitter release or well below threshold for spike generation). A weak 800 Hz component added to that bias would have little or no effect on either transmitter release or spike production. On the other hand, a sufficiently large positive dc component added to the bias could translate the membrane potential to a level close to threshold or above it. In that case, the weak 800 Hz component would begin to have an effect – producing either periodic excursions of membrane potential to super-threshold values, or producing 800 Hz modulation of a potential already above threshold. Enhancement of the 800 Hz envelope response thus could be attributable to either or both of the following effects: (1) the dc response to the background tone allows previously silent units to join the subpopulation of units responding to the probe stimulus. (2) The dc response to the background tone increases the amplitude of the 800 Hz envelope response of units already responding to the probe stimulus.

We are proposing that the enhancement effect shown in this paper requires a cascade of two non-linearities –

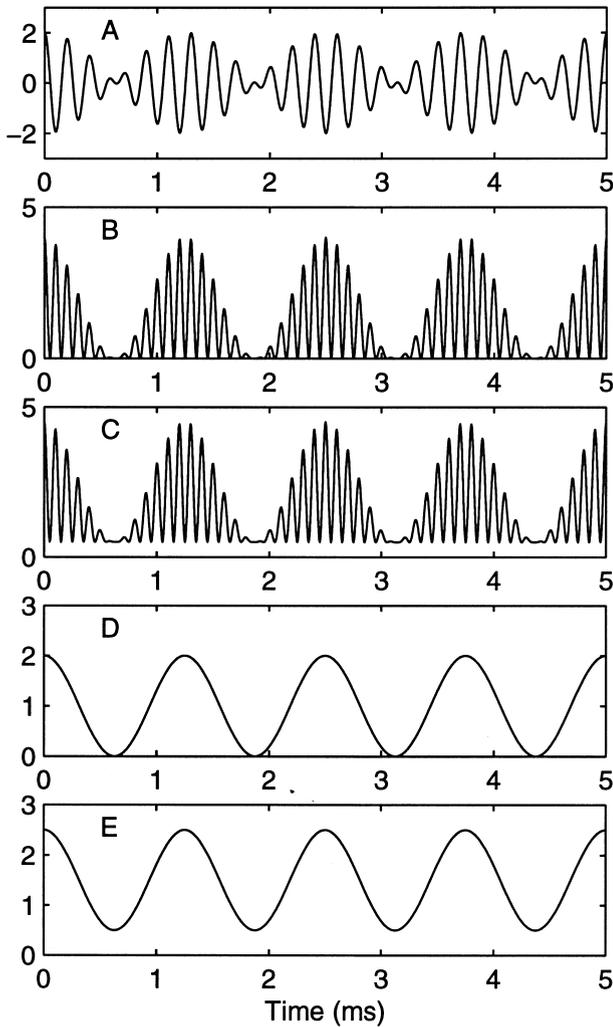


Fig. 11. Simulated experiment: a probe stimulus (A) is applied as input to a square-law non-linearity in the absence (B) and presence (C) of a background tone. Being subjected to low-pass filtering, the waveforms of B and C are transformed to those of D and E, respectively (see text for details).

the first being a rectifying (even-order) non-linearity that extracts a positive-only, envelope-following response from a zero-mean acoustic waveform, and the second being a threshold non-linearity. The mechanisms posited to underlie enhancement according to the French–Stein hypothesis and the common theories of stochastic resonance include just one non-linearity – a threshold non-linearity.

A reviewer of the first draft of this paper pointed out that, at low levels, the slopes of all of the log–log stimulus response curves in this paper are 2, corresponding to a square-law non-linearity (see also Goodman et al., 1982). We have found this as well with our Wiener series analyses of gerbil high-CF cochlear units. In such a unit, the post-stimulus time histogram of cochlear nerve spike activity in response to a repeated complex temporal waveform (a segment of band-limited

noise) follows the square of the envelope of a linearly-filtered version of that waveform (Lewis et al., 2001). The filter function is given by the Wiener analysis. Taking the even-order non-linearity of the previous paragraph to be square-law, we can construct a simple simulation of its response to our experimental stimuli. Fig. 11A shows a probe stimulus comprising ongoing 4.6 kHz and 5.4 kHz tones, each of unity amplitude, summed in cosine phase. Fig. 11B shows the probe stimulus squared. Fig. 11C shows the result of summing the probe stimulus with a 5.0 kHz background tone (also of unity amplitude), squaring the result, and averaging over the full range of relative phases of the background tone. Thus, one can imagine all three sinusoids converging on the hair bundle of an inner hair cell, and being subjected together to the (putatively) square-law non-linearity of the transduction process. Assuming that the waveforms of Fig. 11B,C subsequently are subjected (in the hair cell and, perhaps, the initial segment

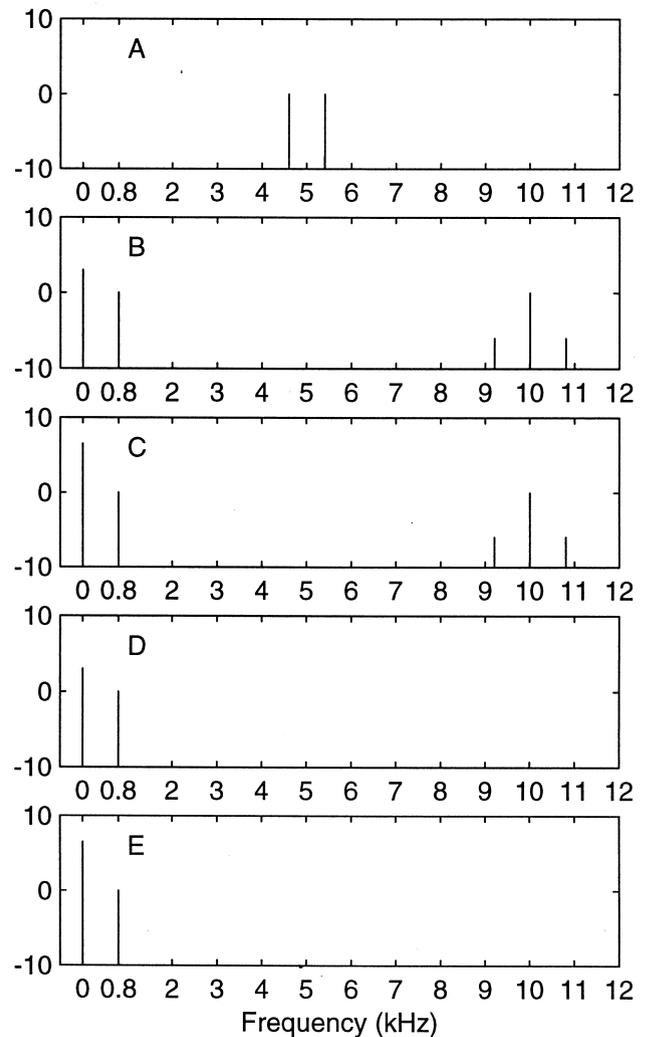


Fig. 12. Amplitude spectra of the waveforms in Fig. 11 (see text).

of the afferent axon) to low-pass filtering, passing the DC and 800 Hz components and rejecting the high-frequency components, one can easily construct the resulting waveforms (Fig. 11D,E).

The amplitude spectra of waveforms in Fig. 11A–E are presented in Fig. 12A–E. The ordinate is given in dB re, a unit-amplitude sine wave. Notice that the net effect of adding the randomly-phased background tone is a 50% (3.52 dB) increase in the dc component of the output of the squaring process. The energy of the 800 Hz component is not changed, nor is that of any of the three components in the vicinity of 10 kHz. Viewing Fig. 11D (simulated probe only) and Fig. 11E (simulated probe plus background), imagine a second non-linearity, in this case a strong one, imposing a threshold-like effect at 2.0 units. Clearly, in that case, the presence of the background tone would enhance the 800 Hz envelope response conspicuously.

References

- Agin, D., 1964. Hodgkin–Huxley equations: Logarithmic relation between membrane current and frequency of repetitive activity. *Nature* 201, 625.
- Braun, H.A., Wissing, H., Schafer, K., Hirsch, M.C., 1994. Oscillation and noise determine signal transduction in shark multimodal sensory cells. *Nature* 367, 270–273.
- Bruce, I.C., White, M.W., Irlicht, L.S., O’Leary, S.J., Clark, G.M., 1999. The effects of stochastic neural activity in a model predicting intensity perception with cochlear implants: Low rate stimulation. *IEEE Trans. Biomed. Eng.* 46, 1393–1404.
- Chialvo, D.R., Apkarian, A.V., 1993. Modulated noisy biological dynamics: Three examples. *J. Stat. Phys.* 70, 375–391.
- Collins, J.C., Imhoff, T.T., Grigg, P., 1996. Noise-enhanced information transmission in rat SA1 cutaneous mechanoreceptors via aperiodic stochastic resonance. *J. Neurosci.* 16, 642–645.
- Dallos, P., Cheatham, M.A., 1976. Compound action potential (AP) tuning curves. *J. Acoust. Soc. Am.* 59, 591–597.
- Dallos, P., Cheatham, M.A., 1977. Analog of two tone suppression in whole nerve response. *J. Acoust. Soc. Am.* 62, 1048–1052.
- de Boer, E., de Jongh, H.R., 1978. On cochlear encoding: potentialities of the reverse-correlation technique. *J. Acoust. Soc. Am.* 63, 115–135.
- Douglass, J.K., Wilkens, L., Pantazelou, E., Moss, F., 1993. Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance. *Nature* 365, 337–340.
- Ehrenberger, K., Felix, D., Svozil, K., 1999. Stochastic resonance in cochlear signal transduction. *Acta Otolaryngol.* 119, 166–170.
- Fauve, S., Heslot, F., 1983. Stochastic resonance in a bistable system. *Phys. Lett.* 97A, 5–8.
- French, A.S., Stein, R.B., 1970. A flexible neural model using integrated circuits. *IEEE Trans. Biomed. Eng.* 17, 248–253.
- Goldberg, J.M., Fernandez, C., 1971. Physiology of peripheral neurons innervating semicircular canals of the squirrel monkey. I. Resting discharge and response to constant angular accelerations. *J. Neurophysiol.* 34, 635–660.
- Goodman, D.A., Smith, R.L., Chamberlain, S.C., 1982. Intracellular and extracellular responses in the organ of Corti of the gerbil. *Hear. Res.* 7, 161–179.
- Greenwood, D.D., 1986. What is ‘synchrony suppression’? *J. Acoust. Soc. Am.* 79, 1857–1872.
- Harris, D.M., 1979. Action potential suppression tuning curves and thresholds: comparison with single fiber data. *Hear. Res.* 1, 133–154.
- Henry, K.R., 1995. Auditory nerve neurophonic recorded from the round window of the Mongolian gerbil. *Hear. Res.* 90, 176–184.
- Henry, K.R., 1996a. Auditory nerve neurophonic produced by the frequency difference of two simultaneously presented tones. *Hear. Res.* 99, 151–159.
- Henry, K.R., 1996b. Tuning curves of the difference tone auditory nerve neurophonic. *Hear. Res.* 99, 160–167.
- Henry, K.R., 1999. Noise improves the transfer of near-threshold, phase-locked activity of the cochlear nerve: Evidence for stochastic resonance? *J. Comp. Physiol. A* 184, 577–584.
- Henry, K.R., Lewis, E.R., 1992. One-tone suppression in the cochlear nerve of the gerbil. *Hear. Res.* 63, 1–6.
- Hudspeth, A.J., Corey, D.P., 1977. Sensitivity, polarity, and conductance change in the response of vertebrate hair cells to controlled mechanical stimuli. *Proc. Natl. Acad. Sci. USA* 74, 2407–2411.
- Jaramillo, F., Wiesenfeld, K., 1998. Mechano-electrical transduction assisted by Brownian motion: A role for noise in the auditory system. *Nat. Neurosci.* 1, 384–388.
- Javel, E., 1981. Suppression of auditory nerve responses. I. Temporal analysis, intensity effects and suppression contours. *J. Acoust. Soc. Am.* 69, 1735–1745.
- Kiang, N., Watanabe, T., Thomas, E.C., Clark, L.F., 1965. *Discharge Patterns of Single Fibers in the Cat’s Auditory Nerve*. MIT Press, Cambridge, MA.
- Lewis, E.R., Henry, K.R., 1995. Nonlinear effects of noise on phase-locked cochlear-nerve responses to sinusoidal stimuli. *Hear. Res.* 92, 1–16.
- Lewis, E.R., Henry, K.R., Yamada, W.M., 2001. What Predictions from the Wiener Series Reveal about Cochlear Signal Processing. Annual Midwinter Meeting of the Association for Research for Otolaryngology, Abst. 24, 102.
- Longtin, A., Bulsara, A., Moss, F., 1991. Time-interval sequences in bistable systems and the noise-induced transmission of information by sensory neurons. *Phys. Rev. Lett.* 67, 656–659.
- Lowenstein, O., 1956. Peripheral mechanisms of equilibrium. *Br. Med. Bull.* 12, 114–118.
- Maddox, J., 1991. Towards the brain-computer’s code? *Nature* 352, 469.
- Morse, R.P., Evans, E.F., 1996. Enhancement of vowel coding for cochlear implants by addition of noise. *Nat. Med.* 2, 928–932.
- Morse, R.P., Evans, E.F., 1999. Additive noise can enhance temporal coding in a computational model of analogue cochlear implant stimulation. *Hear. Res.* 133, 107–119.
- Rose, J.E., Brugge, J.F., Anderson, D.J., Hind, J.E., 1967. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neurophysiol.* 30, 769–793.
- Rubenstein, J.T., Wilson, B.S., Finley, C.C., Abbas, P.J., 1999. Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation. *Hear. Res.* 127, 108–118.
- Russell, I.J., Sellick, P.M., 1978. Intracellular studies of hair cells in the mammalian cochlea. *J. Physiol.* 338, 176–206.
- Ryan, A.F., 1976. Hearing sensitivity of the Mongolian gerbil, *Meriones unguiculatus*. *J. Acoust. Soc. Am.* 59, 1222–1226.
- Simonotto, E., Riani, E., Roberts, M., Twitty, J., Moss, F., 1997. Visual perception of stochastic resonance. *Phys. Rev. Lett.* 78, 1186–1189.
- Stein, R.B., 1967. The frequency of nerve action potentials generated by applied currents. *Proc. R. Soc. B* 167, 64–86.

- Stein, R.B., 1970. The role of spike trains in transmitting and distorting sensory signals. In: Schmitt, F.O. (Ed.), *The Neurosciences*. Rockefeller University Press, New York, pp. 597–604.
- Stein, R.B., French, A.S., Holden, A.V., 1972. The frequency response, coherence, and information capacity of two neuronal models. *Biophys. J.* 12, 295–322.
- Vanderkoy, J., Lipshitz, P., 1984. Resolution below the least significant bit in digital systems with dither. *J. Audiol. Eng. Soc.* 32, 106–113.
- Weisenfeld, K., Moss, F., 1995. Stochastic resonance and the benefits of noise: From ice age to crayfish and SQUIDS. *Nature* 373, 33–36.
- Wolodkin, G., Yamada, W.M., Lewis, E.R., Henry, K.R., 1996. Spike rate models for auditory fibers. In: Lewis, E.R., Long, G.R., Lyon, R.F., Narins, P.M., Steele, C.R., Hecht-Poinar, E. (Eds.), *Diversity in Auditory Mechanics*. World Scientific Press, Singapore, pp. 104–110.
- Yamada, W.M., Lewis, E.R., 1999. Predicting the temporal responses of non-phase-locking bullfrog auditory units to complex acoustic waveforms. *Hear. Res.* 130, 155–170.
- Yamada, W.M., Wolodkin, G., Lewis, E.R., Henry, K.R., 1996. Wiener-kernel analysis and the singular-value decomposition. In: Lewis, E.R., Long, G.R., Lyon, R.F., Narins, P.M., Steele, C.R., Hecht-Poinar, E. (Eds.), *Diversity in Auditory Mechanics*. World Scientific Press, Singapore, pp. 111–118.
- Yu, X., Lewis, E.R., 1989. Studies with spike initiators: Linearization by noise allows continuous signal modulation in neural networks. *IEEE Trans. Biomed. Eng.* 36, 36–43.
- Zenner, H.P., 1993. Possible roles of outer hair cell d.c. movements in the cochlea. *Br. J. Audiol.* 27, 73–77.