A Neural-Counting Model Based on Physiological Characteristics of the Peripheral Auditory System. V. Application to Loudness Estimation and Intensity Discrimination

GERARD LACHS, SENIOR MEMBER, IEEE, RADHI AL-SHAIKH, QI BI, ROSALIE A. SAIA, MEMBER, IEEE, AND MALVIN C. TEICH, SENIOR MEMBER, IEEE

Abstract—**The psychophysical properties of a multiple-channel neural-counting model are investigated. Each channel represents a peripheral afferent fiber (or a group of such fibers) and consists of a cascade of signal-processing transformations, each of which has a physiological correlate in the auditory system. The acoustic signal (which may be a pure tone or Gaussian noise) is passed by our mathematical construct through the following series of transformations: an outer- and middle-ear transmission function, an inner-ear multiple-pole linear-filter tuning mechanism, a nonlinear receptor saturation function, and a refractoriness-modified Poisson transduction mechanism (which leads to a sub-Poisson neural spike count). Spontaneous neural activity is independently incorporated into each channel by means of an additive refractoriness-modified Poisson process. A union process at a more distal center in the nervous system is generated by a parallel collection of such channels with a density (in frequency) determined by the cochlear mapping function. The statistics of the union count (in a fixed time) are then processed at a decision center in a manner that depends on the psychophysical paradigm under consideration. This random count number is assumed to contain all of the information for the examples we consider. Our model has been used to calculate psychophysical functions for the following paradigms: pure-tone loudness estimation, pure-tone and variable-bandwidth noise intensity discrimination, and variable-bandwidth noise loudness summation. The theoretical results, which are determined in large part by spread of excitation, are in good agreement with human psychophysical data, provided that the parameters of the theoretical model are appropriately chosen. It has been found that a suitable choice of parameters is both physiologically sensible and self-consistent. As a further indication of the consistency of the model, the same general parametric dependencies as neurophysiological isointensity contours for peripheral afferent fibers in the squirrel monkey are exhibited by the single-channel theoretical count mean, which is calculated as a function of stimulus level and frequency. The single-channel count mean-to-variance ratio is in accord with laboratory data. Finally, the roles of the various components comprising our theoretical system are discussed, and our model is compared with related constructs.**

Manuscript received April 4, 1984; revised July 1984. This work was supported by the National Science Foundation under Grants BNS80-21140 and ECS82-19636.

G. Lachs is with the Department of Electrical and Electronic Systems, University of South Florida, Tampa, FL 33620.

R. Al-Shaikh was with the Department of Electrical Engineering, the Pennsylvania State University, University Park, PA 16802. He is presently with the Ministry of PT & T, Riyadh, Saudi Arabia.

Q. Bi is with the Department of Electrical Engineering, the Pennsylvania State University, University Park, PA 16802.

R. A. Saia was with the Department of Electrical Engineering, the Pennsylvania State University, University Park, PA 16802. She is presently with TRW Systems, 1 Space Park, Redondo Beach, CA 90278.

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

I. Introduction

WE DEMONSTRATED in Part I [1] of this series of papers ([1]–[4]) that pure-tone intensity discrimination could be satisfactorily described in terms of an energy-based neural-counting model incorporating refractoriness and spread of excitation. In Part II [2] we showed that the same linear filter refractoriness model (LFRM) also provided satisfactory results for fitting data obtained in pure-tone loudness estimation experiments. In Part III [3] we considered intensity discrimination and loudness summation for variable-bandwidth noise stimuli and demonstrated once again that the LFRM calculations were in accord with the trends of the data.

The common model considered in all of these papers yielded surprisingly good fits to a wide variety of psychophysical data, in spite of the fact that the LFRM is clearly incomplete in some respects. Well-known physiological effects such as receptor saturation, spontaneous neural activity, the presence of a frequency-dependent outer- and middle-ear transfer function, and nonuniform cochlear mapping are not included in the LFRM. One problem area for the LFRM is that it predicts a theoretical single-channel count mean-to-variance ratio γ_{th} that increases with stimulus level, becoming very large at high levels. This is at variance with the results of recent neurophysiological experiments showing that γ_{exp} assumes a value roughly between 1 and 2 (in the cat) for a counting time $T = 50$ ms, essentially independent of stimulus level.

In Part IV [4] of this series, a single-channel version of the LFRM was modified to include the effects of receptor saturation and spontaneous neural activity. The modified system was referred to as the "extended linear filter refractoriness model" (ELFRM). The presence of the saturation

0018-9472/84/1100-0819$01.00

function brought γ_{th} into line with the experimental value. Furthermore, the spontaneous neural activity made the ELFRM a more realistic model. The objective of Part IV was to use the single-channel extended model to generate theoretical neural firing-rate curves for an auditory fiber as a function of stimulus level and frequency (isointensity contours). By judicious choice of the parameters of the model, our results could be made to agree with experimental isointensity contours.

In this paper we incorporate the ELFRM of Part IV into a multiple-channel configuration, taking into account the frequency characteristics of the outer- and middle-ear transmission function and the cochlear mapping function. Our principal interest is in the predictive behavior of the system for several psychophysical paradigms: pure-tone loudness estimation, pure-tone intensity discrimination, and intensity discrimination and loudness summation for variable-bandwidth noise. We demonstrate that the theoretical results are in good agreement with human data for these paradigms and that the model parameters required to fit the data are physiologically sensible. We also show that a single-channel version of the model, using the same parameters used in the psychophysical calculations, leads to theoretical isointensity contours that have the same general form as neurophysiological results from peripheral fibers in squirrel monkeys. Our study permits us to demonstrate that the model parameters required to fit neurophysiological (single-channel) and psychophysical (multiple-channel) data are plausible and consistent and to determine the roles of the various components comprising the theoretical system. A preliminary account of this work was presented at the Annual Meeting of the Acoustical Society of America in Chicago, IL [5].

The point of view that we adopt is similar to that introduced by McGill [6] and Siebert [7]. We build the overall system response from the behavior of individual idealized elements in the periphery. Much has been learned about the physiology of the auditory system and neural activity in the VIIIth nerve in the past 15 years. One of our aims in this study is to determine for which paradigms neural-counting information suffices, and how well such models perform. We therefore devote our attention to processing that integrates over both place on the cochlear partition and timing information in the neural spike train. Paradigms such as frequency discrimination, the detection of structured signals buried in noise, and speech recognition could be studied in the context of our model by examining the properties of the theoretical point process rather than the neural count.

In Section II, we discuss the individual elements of the model. In Section III, theoretical formulations for loudness estimation and intensity discrimination are presented. Comparisons of the calculated results with psychoacoustic data are considered in Section IV. The sensitivity of the calculations with respect to parameter values and the role of the various components of the model are discussed in Section V. Comparison with related models is made in Section VI. Finally, the summary and conclusions are given in Section VII.

II. Elements of the Model

The multiple-channel model afferent system is shown schematically in Fig. 1. Each channel represents a peripheral afferent fiber (or a collection of such fibers at the same characteristic frequency). It consists of a cascade of signal-processing transformations, each of which has a physiological correlate in the auditory system. The acoustic signal (which may be a pure tone or Gaussian noise) passes through the following series of transformations: the outer- and middle-ear transmission function, the inner-ear multiple-pole linear-filter tuning mechanism, the nonlinear receptor saturation function, and the refractoriness-modified Poisson transduction mechanism (which leads to neural counts whose variance is less than the mean). Spontaneous counts are incorporated into each channel by means of an additive independent refractoriness-modified Poisson process. The parallel collection of these channels, with a density (in frequency) determined by the cochlear mapping function, generates a union process at a more distal center in the auditory pathway. The statistics of the union count (in a fixed counting time) are processed at a decision center, in a manner that depends on the psychophysical paradigm. This random count number is assumed to contain all of the information for the paradigms we consider.

The model differs principally from the one developed in [1]–[3] in that it includes a memoryless receptor saturation function and spontaneous events, as well as the outer- and middle-ear transmission function and the cochlear mapping function across channels. As indicated in Section I, the model considered in [4] is a single-channel version of the system shown in Fig. 1. As such, it does not incorporate the outer- and middle-ear transmission function, the cochlear mapping function, the union of neural counts, nor the decision center. These characteristics are related to psychophysical tasks.

A. Single-Channel Elements

We begin with a discussion of the single-channel elements. The mathematical description of many of the functional blocks shown in Fig. 1 has been presented in detail in [4]; the formulas are presented here with a minimum of discussion. The inner-ear multiple-pole tuning mechanism is taken to be a linear filter whose response to a pure-tone input at frequency f_T is given (see [1]) by

$$E_o = \frac{AE_i}{\left[1 + Q^2(f_T/f_0 - f_0/f_T)^2\right]^{rN}}. \tag{1}$$

Here E_o and E_i are the output and input energies of the filter, respectively, f_0 is the characteristic frequency (CF) of the neural channel, and Q is a quality factor or tuning parameter (equal to the ratio of the characteristic frequency to the 3-dB bandwidth for a single-tuned energy filter ($N = 1$)). Q itself depends on the characteristic frequency of the fiber, so we allow $Q = Q(f_0)$. As discussed subsequently, the choice $N = 3/2$ provides agreement of our calculations with experiment. The parameter r accommodates the observed asymmetric frequency–response char-

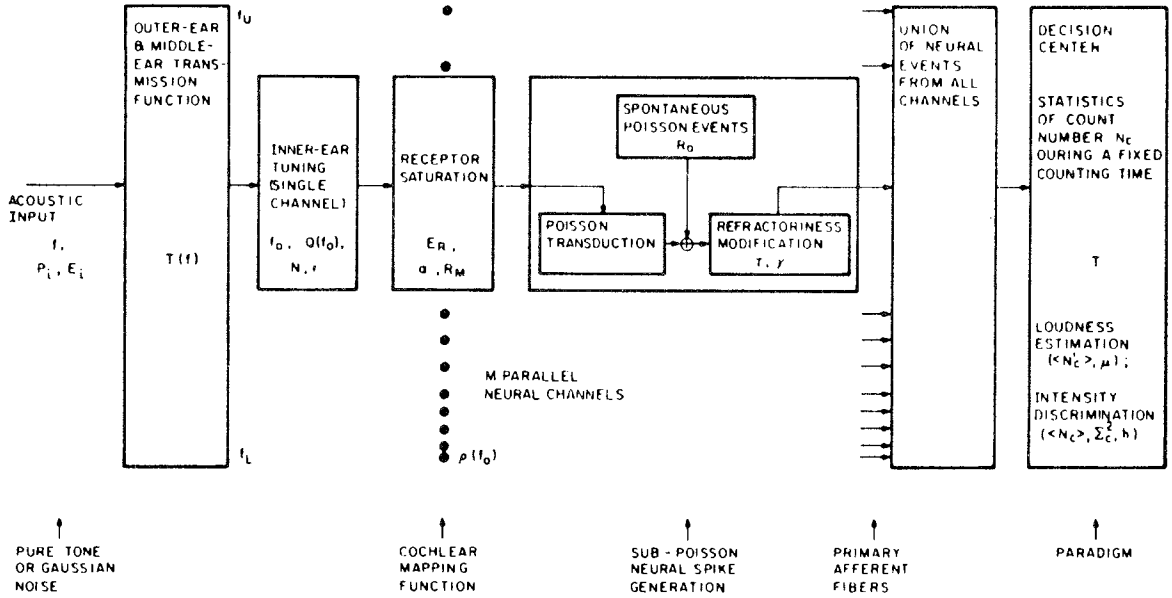


Fig. 1. Block diagram of the model afferent system used in carrying out the calculations.

acteristics of the neural fibers in a simple way. We set $r = 1$ in the region $f_T \leqslant f_0$, and $r = 2$ ($rN = 3$) in the region $f_T > f_0$. The proportionately constant A that appears in (1) is absorbed into the constant E_R [see (2) and (3)]. The linear-filter transmission function given in (1) (with $r = 1$) represents a system consisting of a cascade of masses and springs, each undergoing simple damped harmonic motion. We have used this well-known response for resonant linear systems as an empirical functional form to maintain simplicity.

The output of this filter is fed to a memoryless receptor saturation function. A linear-filter/nonlinearity cascade is simpler to analyze than is a nonlinear filter. Two distinct functional forms for the receptor saturation have been investigated (see [4]). The first has a mean neural-firing rate without refractoriness modification, $\bar{n}_u(f_0)$, given by [8]

$$\bar{n}_u(f_0) = TR_M\left\{1 - \exp\left[\frac{-R_0}{R_M}(1 + E_o/E_R)^{\theta}\right]\right\}. \quad (2)$$

Here, R_0 and R_M are the minimum and maximum firing rates, respectively, in spikes/s (before refractoriness modification), T is the counting time, E_R is an energy reference level that governs the channel threshold, and θ is a parameter controlling the maximum slope of the response. The quantity R_0 is selected to match the spontaneous rate, since the refractoriness modifications at sufficiently low levels will be negligible. R_M may be calculated (see [4]) from the observed maximum firing rate of the neural channel R_m (which occurs after refractoriness modification).

The second saturation function that we examined is described by the expression (see [4], Appendix)

$$\bar{n}_u(f_0) = T\left\{R_0 + \frac{\alpha(R_m - R_0)\ln(1 + E_o/E_R)}{1 + \alpha\left(\dfrac{R_m - R_0}{R_M - R_0}\right)\ln(1 + E_o/E_R)}\right\}. \quad (3)$$

Here, α governs the maximum slope of the response (analogous to θ in (2)), R_m is the maximum observed firing rate in spikes/s after refractoriness modification, and the parameters R_0, R_M, E_R, and T have the same interpretation as in (2). It has been shown in the Appendix of [4] that $R_M = \sqrt{\gamma}\, R_m$. The stimulus energy or intensity E_o that appears in both (2) and (3) is given in (1).

The receptor output feeds an ideal nonparalyzable dead-time-modified Poisson transduction mechanism. (For convenience, the additive spontaneous Poisson events have been incorporated into the receptor saturation function.) The effects of dead time (refractoriness) on a Poisson process are well understood [9], [10] and have been discussed at some length in [1]. When the dead time is fixed, the expressions for the modified count mean and count variance are, respectively,

$$\bar{n}_c(f_0) = \frac{\bar{n}_u(f_0)}{[1 + (\tau/T)\bar{n}_u(f_0)]} \quad (4)$$

and

$$\sigma_c^2(f_0) = \frac{\bar{n}_u(f_0)}{[1 + (\tau/T)\bar{n}_u(f_0)]^3}. \quad (5)$$

Here τ is the dead time, and T is the counting time. When using these formulas to calculate psychophysical functions, we take T to be the integration time at the decision center (see below). The count mean-to-variance ratio γ is obtained by simply dividing (4) by (5), i.e.,

$$\gamma = [1 + (\tau/T)\bar{n}_u(f_0)]^2. \quad (6)$$

The quantity $\bar{n}_u(f_0)$ is obtained from (1), together with either (2) or (3).

B. Multiple-Channel Elements

We now turn our attention to the set of M parallel channels (see Fig. 1). For simplicity and concreteness, we take the outer- and middle-ear energy transmission function $T(f)$ to be the human threshold-of-hearing curve

(actually, the threshold-of-hearing curve reflects the behavior of the entire auditory system, but we use this function as an approximation for the function $T(f)$). We have obtained a rather simple empirical form for $T(f)$ (which is denoted $T_{\mathrm{dB}}(f)$ when expressed in dB) from the data of Scharf [11]

$$T_{\mathrm{dB}}(f) = -\{a[\ln(f)]^2 + b[\ln(f)] + c\}. \quad (7)$$

Numerical estimates for the constants a, b, and c were obtained by fitting Scharf's data with two quadratic equations in $\ln(f)$, one on each side of the minimum threshold (taken to be at 1450 Hz), and assuming zero slope at the junction. The resulting constants associated with (7) are displayed in footnote b of Table I. To convert $T_{\mathrm{dB}}(f)$ into $T(f)$, we use the simple relation

$$T(f) = 10^{[T_{\mathrm{dB}}(f)/10]}. \quad (8)$$

This element of the model plays an important role principally for broadband stimuli.

The dependence of Q on f_0 is characterized by an empirical relation of the form

$$Q(f_0) \approx d_1 + d_2 \ln(f_0). \quad (9)$$

This is drawn from the neurophysiological dependence measured in chamber-raised cats [12]. Values for the constants d_1 and d_2 are displayed in footnote d of Table I. Similar behavior is observed for cats that are not specially raised.

The cochlear mapping function is incorporated into the model by means of a nonuniform (in characteristic frequency) weighting of the number of neural channels (see Fig. 1). The quantity $\rho(f_0)$ is a suitably normalized neural-fiber density function, reflecting the tonotopic organization of the cochlea. A simple form that we used is

$$\rho(f_0) = k/f_0. \quad (10a)$$

This reflects the roughly logarithmic distribution of characteristic frequency along the cochlear partition, as revealed by neuroanatomical studies [13], [14]. We can determine the normalization constant k by assuming that

$$\int_{f_L}^{f_U} \rho(f_0)\, df_0 = M. \quad (10b)$$

Here f_L and f_U represent the lower and upper frequency limits, respectively, of the characteristic frequencies of human peripheral auditory fibers, and M is the total number of these fibers. Values are given in footnote b of Table I. As a shorthand notation we refer to calculations using this uniform-in-log-frequency distribution as "with fiber density." When $\rho(f_0)$ is chosen to be independent of f_0 (uniform-in-linear-frequency distribution), the results are referred to as "without fiber density." All of the calculations reported in [1]–[3] were carried out without fiber density. As will be seen subsequently, the cochlear mapping function plays an important role in determining the values of the model parameters.

The union operation forms a point process that is the superposition of the constituent single-channel processes. The decision center extracts those statistics of the union point process appropriate for a particular paradigm (e.g., the modified count mean for loudness estimation or the count mean and variance for intensity discrimination). The roles of these two operations are identical to those assumed in [1]–[3].

The sum of the neural counts from the individual channels during the fixed counting time T is a random variable N_c. Its mean and variance, denoted $\langle N_c \rangle$ and Σ_c^2, respectively, are given by

$$\langle N_c \rangle \approx \int_{f_L}^{f_U} \bar{n}_c(f_0)\rho(f_0)\, df_0 \quad (11)$$

and

$$\Sigma_c^2 \approx \int_{f_L}^{f_U} \sigma_c^2(f_0)\rho(f_0)\, df_0. \quad (12)$$

Equations (11) and (12) are approximate because we have used integration to provide the overall count mean and variance. This is satisfactory because the characteristic frequencies of the peripheral fibers essentially form a continuum.

The count statistics that are essential for those psychophysical paradigms of interest to use are $\langle N_c \rangle$, Σ_c^2, and a modified version of $\langle N_c \rangle$ that will be described below.

III. Theoretical Formulations for Psychoacoustic Paradigms

A. Pure-tone Loudness Estimation

In the course of carrying out our study it became apparent that the total mean number of counts $\langle N_c \rangle$ represented in (11), which includes a contribution from spontaneous events, would not suffice for fitting loudness data at low stimulus levels. In loudness estimation experiments, a subject is asked to associate a numerical value with the magnitude of a stimulus; he or she is expected to generate a loudness function with value zero in the absence of sound. It is therefore not unreasonable to suppose that spontaneous counts should not contribute to the loudness (they were omitted from the loudness estimation model used in [2]). The observable for loudness estimation is therefore considered to be a modified sum of counts, denoted as $\langle N_c' \rangle$, in which spontaneous events are subtracted at each channel before integration. There is also neurophysiological justification for ignoring the spontaneous events inasmuch as the presence of a stimulus seems to inhibit their generation [15]. We do not imply that spontaneous activity is to be ignored altogether, or that the detection system necessarily differentiates between spontaneous and driven events. In fact, the forms that we have chosen for (2) and (3) exhibit a functional dependence on R_0, and the fluctuations of the neural count therefore contain within them the fluctuations of spontaneous activity.

In analogy with (11) we define $\langle N_c' \rangle$ as the total modified mean neural count. If the spontaneous count, and the effects of dead time on it, are small $\langle N_c' \rangle$ can be approxi-

mated as

$$\langle N_c' \rangle \approx \int_{f_L}^{f_U} [\bar{n}_c(f_0) - R_0'] \rho(f_0)\, df_0. \tag{13}$$

The parameter R_0' is the spontaneous count that emerges when E_o is set equal to 0 in (2) and (3). Thus

$$R_0' = TR_M[1 - \exp(-R_0/R_M)] \tag{14a}$$

for exponential receptor saturation, and

$$R_0' = R_0 T \tag{14b}$$

for logarithmic receptor saturation.

Any single observation of the scalar loudness random variable λ will be given by the total modified neural count, which varies from trial to trial (see [2]). Thus

$$\lambda = \mu N_c' \tag{15}$$

where μ is a constant. The experimental psychophysical data that we attempt to fit almost always involve averages over repeated trials, however. If we assume that these trials are statistically independent (see [2]), the average provides a pure-tone loudness function L given by

$$L = \mu \langle N_c' \rangle. \tag{16}$$

B. Pure-Tone Intensity Discrimination

The usual intensity-discrimination paradigm makes use of a two-interval forced-choice (2IFC) procedure. The subject is sequentially presented with two short stimulus bursts at different levels. The level of one burst is fixed while the level of the other burst is adjusted until the subject makes a correct decision on a prespecified percentage of trials. This stimulus level difference is measured as a function of the base stimulus level.

An appropriate quantitative measure for pure-tone intensity discrimination is provided by the detection distance h given by [1]

$$h = \frac{\langle N_{cs} \rangle - \langle N_{cw} \rangle}{(\Sigma_{cs}^2 + \Sigma_{cw}^2)^{1/2}}. \tag{17}$$

The quantities $\langle N_{cs} \rangle$ and Σ_{cs}^2 are obtained from $\langle N_c \rangle$ and Σ_c^2 in (11) and (12), respectively; the subscripts s and w refer to the stronger and weaker stimuli. The parameter h is selected to satisfy the specified criterion for the probability of a correct decision. As an example, let's suppose that the level of the stronger input is fixed, and the level of the weaker input is adjusted to satisfy (17) with some particular value of h. This procedure provides one data point. Other data points are obtained by choosing different values for the level of the stronger input (keeping h constant) and repeating the process by varying the weaker level until (17) is satisfied.

Equation (17) is based in part on the normal approximation as described in [1]. This simplification permits error probabilities to be expressed solely in terms of the count mean and variance, which we can calculate from our model. It is interesting to note that it is immaterial whether primed or unprimed variables are used in the numerator of (17) since

$$\langle N_{cs} \rangle - \langle N_{cw} \rangle = \langle N_{cs}' \rangle - \langle N_{cw}' \rangle. \tag{18}$$

The denominator, on the other hand, retains its dependence on spontaneous activity. According to this equation, therefore, intensity discrimination (and indeed any psychophysical paradigm that depends on the probability of counts) will implicitly retain a dependence on spontaneous activity. We have not determined the extent of this dependence, though it is likely to be more important at low stimulus levels.

C. Intensity Discrimination and Loudness Summation for Variable-Bandwidth Noise

The detection of signals other than pure tones is governed by many individual channels that contribute to the union process even at low stimulus levels. As is evident from Fig. 1, the functions $T(f)$, $Q(f_0)$, and $\rho(f_0)$ will be most important under these stimulus conditions. Furthermore there is the question of external stimulus fluctuations and the role that they play [6], [7], [16], [17]. We consider this latter issue first.

In [3] we pointed out that stimulus intensity fluctuations diminish in importance for detection systems exhibiting large values of the time-bandwidth product (in comparison with unity) because of the time averaging of the fluctuations. The averaging (for each afferent fiber in the counting time T) may be sufficient to eradicate almost all variability in the energy [6], leading back to the dead-time-modified Poisson process used to derive the pure-tone results. Studies of the photon-counting detection of natural light (which is an exponential-noise-driven Poisson process) show that the output events may be accurately represented as a simple Poisson counting process (rather than a doubly stochastic counting process) as long as the product of counting time and bandwidth (rad/s) is greater than about ten [18], [19].

If we consider that the bandwidth of a system is expressible as $\Delta\omega \equiv 2\pi\Delta f = 2\pi f_0/Q$, then the time-bandwidth product $T\Delta\omega \approx 2\pi T f_0/Q$. If $T \approx 0.1$ s and $Q \approx 40$, then at low characteristic frequencies, where stimulus fluctuations exhibit their maximum effect, $T\Delta\omega \approx \pi f_0/200$. Even for a fiber with a CF as low as 200 Hz, there is a good deal of averaging occurring. Saturation and refractoriness further diminish the importance of residual fluctuations at moderate and high stimulus levels [3]. Our treatment of intensity discrimination and loudness summation with noise stimuli therefore ignores the effect of such fluctuations. This is the point of view adopted by Siebert [7], [16].

The calculations for intensity discrimination and loudness summation follow from the results presented in [3]. The stimulus that we consider is bandlimited flat-spectrum noise characterized by the spectral density

$$S_i(f) = \begin{cases} \eta, & f_1 \leq f \leq f_2 \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

The quantity η is the stimulus noise energy spectral density, and f_1 and f_2 are the lower and upper cutoff frequencies, respectively. When the noise consists of short bursts, as in all of the experiments cited here, its magnitude may be expressed equivalently in terms of energy (E), power (P), or intensity (I). After passing through the outer- and middle-ear transmission function, the energy spectral density presented to the inner ear takes the form

$$S(f) = S_i(f)T(f). \tag{20}$$

The output of the asymmetric linear filter associated with an individual neural channel is then [3]

$$E(f_0) = A_1 \int_{f_L}^{f_0} \frac{S(f)\,df}{\left[1 + Q^2(f/f_0 - f_0/f)^2\right]^N} + A_1 \int_{f_0}^{f_U} \frac{S(f)\,df}{\left[1 + Q^2(f/f_0 - f_0/f)^2\right]^{rN}}. \tag{21}$$

Here A_1 is a constant that is absorbed into E_R, and $r = 2$ for the second integral, where $f \geq f_0$ as discussed earlier. The remainder of the calculations follow those for pure tones. $E(f_0)$, obtained in (21) above, is used in the receptor saturation formulas (2) or (3). Ultimately, results are obtained for $\langle N_c \rangle$, Σ_c^2, and $\langle N_c' \rangle$, and intensity discrimination and loudness are calculated by using (17) and (16), respectively.

IV. Comparison of Theory with Psychoacoustic Data

A. Pure-Tone Loudness Estimation

In Fig. 2, we present loudness-estimation data (versus pure-tone stimulus intensity in dB) for tone bursts at 100 Hz, 1000 Hz, and 3000 Hz. The open circles are experimental data points. The data at 100 Hz are adapted from Hellman and Zwislocki [20, Fig. 2]; they were obtained monaurally and represent an average over nine listeners. The data at 1000 Hz are adapted from Hellman and Zwislocki [21, Fig. 1]; they were obtained binaurally and represent an average over several studies. The data at 3000 Hz are adapted from Hellman [22, Fig. 1]; they were obtained binaurally and represent an average over ten listeners. The solid curves are theoretical loudness functions calculated on the basis of (16). The logarithmic saturation function (3) was used in all cases (exponential

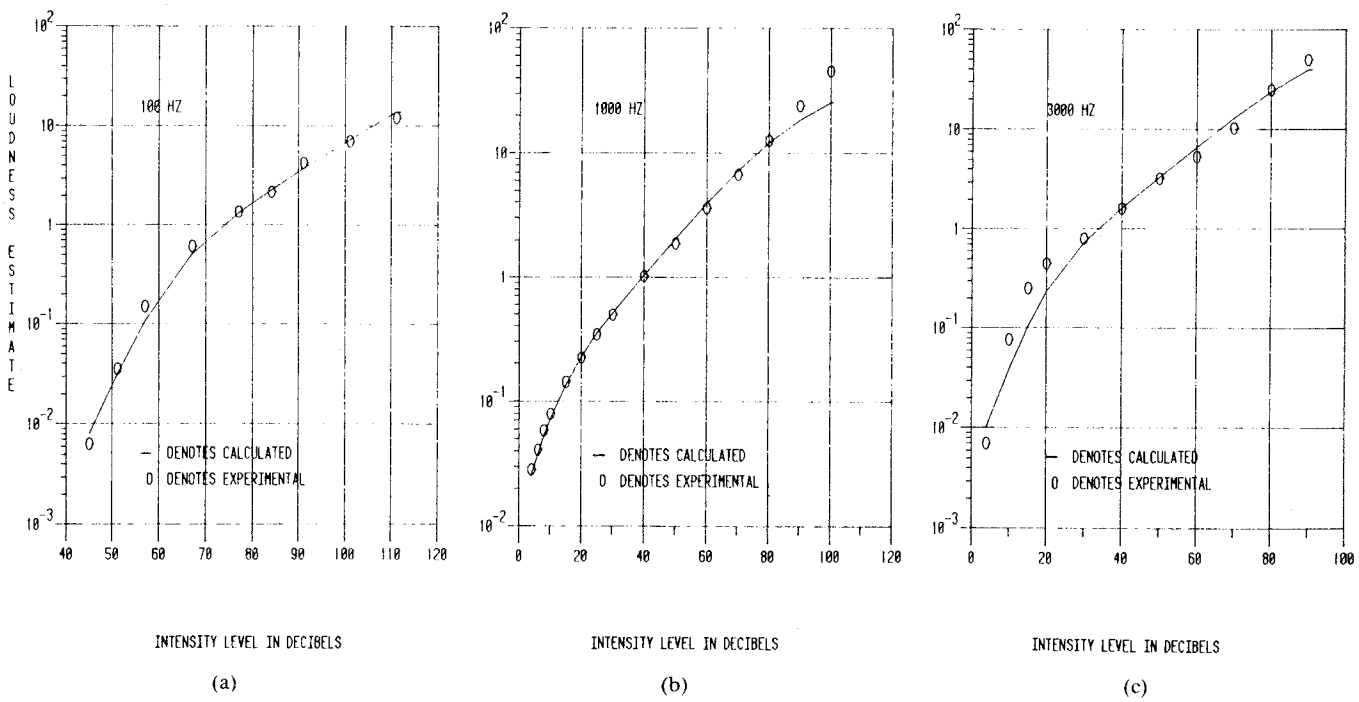


Fig. 2. Plot of loudness estimate versus pure-tone stimulus intensity (dB) for three frequencies. Open circles denote experimental measurements. Solid curves are theoretical loudness functions calculated on the basis of the model afferent system shown in Fig. 1 using the logarithmic saturation function. Parameters for the theoretical curves are presented in Table I. The fits of theory to data are seen to be quite good, with behavior similar to Stevens' power law emerging at high stimulus levels. (a) Data at 100 Hz (adapted from Hellman and Zwislocki, [20, Fig. 2]); they were obtained monaurally and represent an average over nine listeners. (b) Data at 1000 Hz (adapted from Hellman and Zwislocki [21, Fig. 1]); they were obtained binaurally and represent an average over several studies. (c) Data at 3000 Hz (adapted from Hellman [22, Fig. 1]); they were obtained binaurally and represent an average over ten listeners.

TABLE I
MODEL PARAMETER VALUES USED TO FIT EXPERIMENTAL DATA IN FIGS. 2–6 USING LOGARITHMIC FORM FOR RECEPTOR SATURATION FUNCTION[a,b,c]

Fig. No.	Paradigm	f_T or center freq.	E_R	$Q(f_0)$[d]	μ[e]	h[f]	Results matched at[g]
2a	Pure-tone	100 Hz	1.44×10^{6}	40–90	6.6×10^{-4}	–	–
2b	loudness	1000 Hz	6.07	40–90	2.1×10^{-4}	–	–
2c	estimation	3000 Hz	56.7	40–90	7.6×10^{-4}	–	–
3	Pure-tone intensity discrimination	1000 Hz	8.0	40–90	–	3.87	43 dB SL
4a	Noise intensity discrimination (ΔF = 200 Hz)	1000 Hz	750	40–90	–	3.87	40 dB SPL
4b	Noise intensity discrimination (ΔF = 800 Hz)	1000 Hz	3000	40–90	–	3.87	50 dB SPL
5	Loudness summation	1420 Hz	52.4	2.80	–	–	–
6b	Spike-rate isointensity contours[h]	2100 Hz	100	50	–	–	–

[a] The following single-channel parameters are identical for all figures (except for the single-fiber isointensity contours as noted in footnote h below): $N = 3/2$, $r = 2$, $\alpha = 0.5$, $R_0 = 2$, $R_M = 150$ (corresponding to $R_m = 122$), $\tau = 1.3$–1.5 ms, and $\gamma = 1.5$.

[b] The following multiple-channel parameters are identical for all paradigms. For the outer- and middle-ear transmission function $T_{dB}(f)$ (7), when $f_T < 1450$ Hz: $a = 8.4410$, $b = -122.89$, $c = 447.28$; when $f_T > 1450$ Hz: $a = 13.245$, $b = -192.79$, $c = 701.70$. This function is used only for noise stimuli. For the lower and upper limits of the characteristic frequencies f_0 of the individual channels ($f_L \leq f_0 \leq f_U$): $f_L = 50$ Hz; $f_U = 15$ kHz. For the cochlear mapping function $\rho(f_0)$ (10): $k = 5260$, when the number of neural channels $M = 30000$ (which is the approximate number of peripheral afferent fibers in the human auditory system). The central integration time T is taken to be 0.1 s.

[c] The parameters A in (1) and A_1 in (21) are absorbed into E_R and hence carry no significance.

[d] For the variation of Q with characteristic frequency (9): d_1 and d_2 were chosen such that $Q(50\text{ Hz}) = 40$ and $Q(15\text{ kHz}) = 90$, viz., $d_1 = 5.7068$, $d_2 = 8.7661$. This was satisfactory for all paradigms except loudness summation. See footnote h for the value used for single-fiber isointensity contours.

[e] μ is a free parameter for loudness estimation. It brings the calculated loudness function into accord with the subject's scale, essentially moving it vertically. Since μ is determined at the central decision center, its value is governed by effects all along the auditory pathway (e.g., thinning of the union process).

[f] h is a free parameter for intensity discrimination. It has been adjusted to bring the calculated intensity discrimination functions into approximate consonance with the experimental curves (E_R also plays some role in this). The value reported below cannot be taken too seriously since it is appropriate for the union process rather than for the process at the decision center where h operates. If the central counting process is a thinned version of the peripheral process N_c, as may be expected, the required value of h will be smaller than that specified in this column.

[g] The parameter h was adjusted to approximately match the intensity discrimination calculations and data along the abscissa. The theoretical curve was then finely adjusted in the horizontal direction to provide precise matching at the stimulus levels specified below.

[h] For these squirrel-monkey isointensity contours, we chose $R_m = 130$, $T = 1$ (rather than 0.1) because the firing rate is measured in spikes/s, and $Q(2100\text{ Hz}) = 50$, which lies in the region between 40 and 90.

saturation led to almost identical results). The fits of the curves to the data points are very good. Numerical values for the parameters used in these calculations are presented in Table I. $N = 3/2$ was used throughout. The transmission function $T(f)$ was not used because the stimulus is a single frequency in each case. This accounts for the rather large value of the parameter E_R when $f_T = 100$ Hz.

B. Pure-Tone Intensity Discrimination

In Fig. 3, we show a typical set of intensity-discrimination data collected for a pure-tone stimulus at 1 kHz (open circles). These particular data are adapted from Luce and Green [23, Fig. 3], though similar results have been reported by many researchers (see [1]). Intensity discrimination data may be displayed in a number of ways [24]; we plot the Weber fraction, $\log(\Delta I/I)$, versus I (dB). (To make contact with the usual notation, we replace E_{is} by I and $E_{is} - E_{iw}$ by ΔI.) The solid curve represents calculations based on (17), using $N = 3/2$ and the logarithmic saturation function. (Exponential saturation again led to similar results.) The theoretical and experimental values were precisely matched at 43 dB SL. Numerical values for the parameters are given in Table I. The fit of the theory to the data is quite good, representing something close to the near miss to Weber's law at high stimulus levels (see [1]).

C. Intensity Discrimination for Variable-Bandwidth Noise

Data similar to that shown in Fig. 3, but for variable-bandwidth noise, are presented in Fig. 4. The Weber fraction ($\Delta I/I$ in dB) is plotted versus the power ratio per 1/3 octave of the stimulus burst (I dB re 0.0002 μbar). The experimental data (open circles) are adapted from Bos and de Boer [25, Fig. 5]. The Gaussian stimulus bursts had a flat energy spectral density, with a (geometric mean) frequency of 1 kHz and a duration of 125 ms. A very weak broadband background noise (-30 dB re I) was also present during the experiments. In Fig. 4(a), the noise bandwidth ΔF is 200 Hz, whereas in Fig. 4(b) it is 800 Hz. The solid curves represent theoretical intensity-discrimination functions calculated on the basis of (17), using (19)–(21) (the weak broadband background noise was ignored in the theory). The calculations were matched to the data at 40 dB SPL in Fig. 4(a) and at 50 dB SPL in Fig. 4(b). Numerical value for the parameters are given in Table I. Again, the model calculations follow the data quite well. The behavior for these relatively narrowband stimuli resembles that for the pure tone (Fig. 3); both exhibit a near miss to Weber's law at high stimulus levels.

D. Loudness Summation for Variable-Bandwidth Noise

These data are usually presented in the form of sound pressure level for an equivalently loud pure tone or prespecified band of noise, versus the half-power bandwidth of the (flat-spectrum bandlimited Gaussian) noise stimulus ΔF (see [3]). In Fig. 5, the open circles represent measurements for which the comparison was a band of noise of center frequency 1420 Hz and half-power bandwidth 210 Hz. The data are adapted from Zwicker, Flottorp, and Stevens [26, Fig. 8], and represent values of ΔF ranging from 35 Hz to 7.5 kHz. The solid curves represent the theoretical calculations based on (16), using (19)–(21). The logarithmic saturation function was used for the calculation (exponential saturation gave similar results). Parameters for the theoretical curves are shown in Table I. Here too, the theory follows the general shapes of the loudness summation data, but we have had to use a substantially lower value of Q than for the other paradigms. The use of larger values of Q in our model yields curves that rise too steeply with increasing ΔF. We conjecture in Section VII that critical-band phenomena may have some bearing on this issue.

Because the stimulus in this case is broadband, the middle-ear transfer function $T(f)$ was essential to the calculation. The theoretical curves in Fig. 5 fit the data far better than do the calculations presented in [3, Fig. 3]. The principal reason for this is the use of $T(f)$. Without this transmission function, the loudess summation effect grows rapidly and monotonically with ΔF. The results for $\Delta F =$ 3.5 kHz and 7.5 kHz are then far too large, and the gentle maximum in the 30 dB calculation never appears (see [3]). Similar measurements have recently been made with complexes of pure tones [27].

E. Neurophysiological Isointensity Contours

Strictly speaking, the discussion in this subsection should be dealt with elsewhere since it pertains to animal neurophysiological data rather than human psychophysical data. Nevertheless we wish to explicitly show that the mean spike firing rate for an individual mammalian peripheral fiber, stimulated by a pure tone of variable frequency and intensity, is not too dissimilar from the results predicted by our model using $N = 3/2$. In Fig. 6(a), we present experimental isointensity contours collected by Rose *et al.* [28] for a (squirrel-monkey) afferent fiber with a characteristic frequency of 2100 Hz. In Fig. 6(b), we show the theoretical curves based on (a single channel of) the model afferent system, using $N = 3/2$. The lowest theoretical isointensity contour is at a level of 30 dB and the highest is at 80 dB (the curves are in steps of 10 dB). Parameter values are given in Table I (see footnote h). The value of Q ($= 50$) was chosen to lie in the region between 40 and 90. Though the calculated and experimental curves do not agree well in detail, they have the same general parametric dependences. Isointensity contours calculated at other frequencies have about the same level of agreement with data.

V. Role of Parameter Values and Various Elements of the Model

We now examine the values assigned to the parameters in the model and show that they are both physiologically

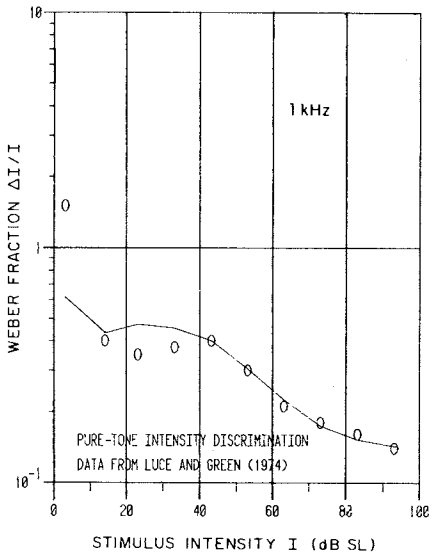


Fig. 3. Weber fraction ($\log \Delta I/I$) plotted against stimulus intensity (I dB SL) for pure-tone intensity discrimination. Experimental data are represented by open circles (adapted from Luce and Green [23, Fig. 3]). Data were obtained at 1 kHz. Solid curve is the intensity discrimination function calculated on the basis of the model afferent system shown in Fig. 1. It follows general trend of the data and represents something close to the near miss to Weber's law at high stimulus levels. Parameters for the theoretical curve are presented in Table I.

plausible and self-consistent. The interplay among the cochlear mapping function and the inner-ear tuning parameters N and Q is then discussed. This is followed by an examination of the roles assumed by the various components of the theoretical system and a discussion of their relative importance in provding agreement with the data.

A. Parameter Values

There are many parameters specified in the model afferent system (see Fig. 1) and it is important to establish their association with empirical values. We begin with the single-channel parameters listed in footnote a of Table I. Values for α, R_0, and R_m have been selected to accord with typical single-fiber measurements (e.g., rate functions and isointensity contours) in the cat [15], [29] and squirrel monkey [28]. The count mean-to-variance ratio γ has been set at 1.5 in accordance with the measurements of Teich and Khanna in the cat [30]. R_M and τ are determined from R_m and γ (see [4, Appendix]). Values for human auditory fibers are likely to be similar.

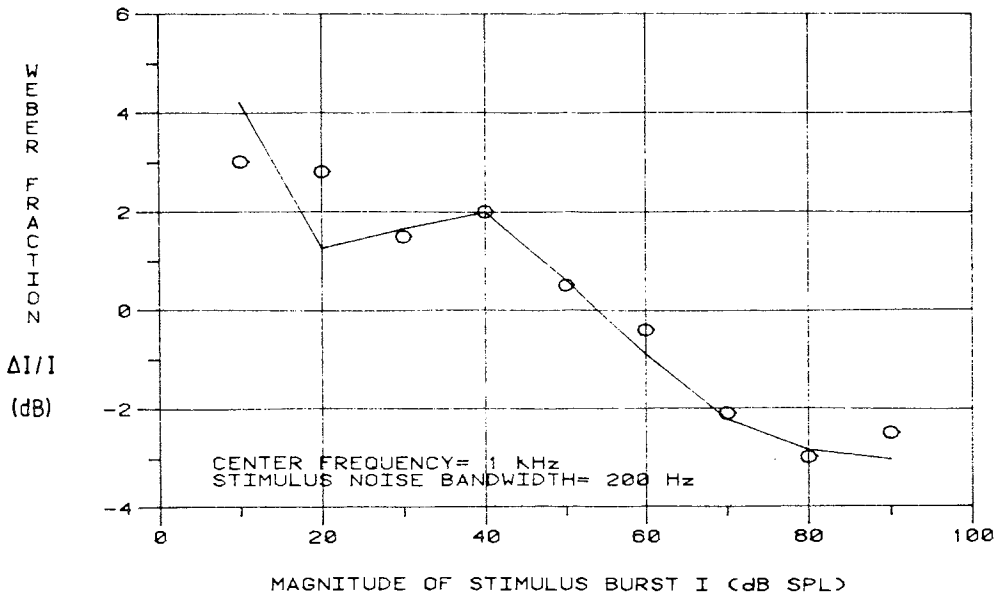
The multiple-channel parameters a, b, and c associated with T_{dB} (see footnote b of Table I), as well as the range of f_0 (between f_L and f_U), and the central integration time T [6], are determined from behavioral measurements in humans. The number of primary fibers M, and the logarithmic form for the cochlear mapping distribution, are determined from neuroanatomical studies in humans [13], [14], [31]. (It is interesting to note that recent neuromagnetic measurements confirm that the logarithmic tonotopic organization present in the cochlea is carried through to the auditory cortex [32], [33].) Taking the number of peripheral afferent fibers in the auditory system to be $M \approx 30000$, with $f_L = 50$ Hz and $f_U = 15$ kHz, we find from (10a) and (10b) that $k \approx M/[\ln(f_U/f_L)] \approx 30000/[\ln(300)] \approx 5260$.

The role of the inner-ear linear-filter parameters N, r, and $Q(f_0)$ is more subtle and quite critical. Essentially all of the psychophysical data, and the squirrel–monkey isointensity contours as well, can be fit by using $N = 3/2$ when the cochlea mapping function is included in the model. This half-integer value of N suggests that the sound pressure level probably plays an important role in the mechanism of auditory transduction. Indeed, the fact that the neural point process in primary auditory fibers displays firing behavior that is linked to the phase of the pressure waveform (phase locking) supports this notion. Nevertheless, for convenience we frame our model in terms of energy and do not concern ourselves with phase information at this juncture. The asymmetry parameter r was inserted in an *ad hoc* manner to account for the steeper rolloff of neural tuning curves on the high-frequency side relative to the low-frequency side. Its value does not affect the computed results significantly as long as $r \geq 2$. The value $r = 2$ provides sufficient attenuation for frequencies above the CF so it was used throughout.

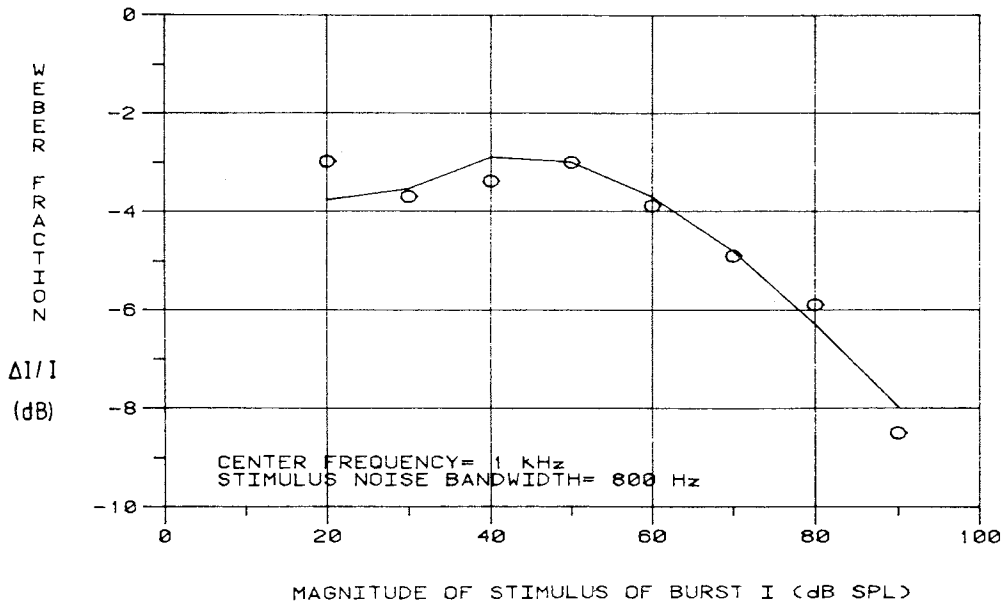
The values and range of the tuning parameter $Q(f_0)$ turn out to be crucial, particularly for intensity discrimination. Increasing Q causes the intensity discrimination curve [$\log(\Delta I/I)$ versus $\log I$] to descend more rapidly. The values reported in Table I have been found to provide the best fit to the intensity discrimination data. Pure-tone loudness estimation is not very sensitive to the choice of Q, provided that it is sufficiently large. Thus intensity discrimination and pure-tone loudness estimation data can be fit with the same $Q(f_0)$. As we indicated earlier, we have used an empirical form for $Q(f_0)$ [see (9)] drawn from Liberman's neurophysiologically determined dependence in the cat [12]. His measurements for $Q_{40\mathrm{dB}}$ can be converted to equivalent values for the parameter Q in (1), by using the relation

$$Q = \frac{[10^{4/N} - 1]^{1/2}(2Q_{40\mathrm{dB}})(2Q_{40\mathrm{dB}} + 1)}{(4Q_{40\mathrm{dB}} + 1)}. \quad (22)$$

Equation (22) is very similar to (21) in [2], where we converted Kiang's $Q_{10\mathrm{dB}}$ to Q. Using Liberman's $Q_{40\mathrm{dB}}$ values for the cat, the conversion yields $Q(50\ \mathrm{Hz, cat}) \approx 14$ and $Q(15\ \mathrm{kHz, cat}) \approx 37$. The exact values depend on precisely which bandwidth he used in determining $Q_{40\mathrm{dB}}$. To obtain good fits to human psychophysical data, we had to choose slightly larger values of Q. These are $Q(50\ \mathrm{Hz, human}) = 40$ and $Q(15\ \mathrm{kHz, human}) = 90$, leading to $d_1 = 5.7068$ and $d_2 = 8.7661$. For the squirrel-monkey isointensity contours (Fig. 6(b)), we used $Q(2100\ \mathrm{Hz, monkey}) = 50$, which lies between 40 and 90. Much lower values of Q (≈ 2.80) were required to fit loudness summa-



(a)



(b)

Fig. 4. Weber fraction ($\Delta I/I$ in dB) plotted against power ratio per 1/3 octave of the stimulus burst (I dB re 0.0002 μbar) for variable-bandwidth noise intensity discrimination. Experimental data are represented by open circles (adapted from Bos and de Boer [25, Fig. 5]). Data obtained using Gaussian noise stimulus with a (geometric mean) frequency of 1 kHz and duration of 125 ms. A very weak broadband background noise (-30 dB re I) was also present. Solid curves are the intensity discrimination functions calculated on the basis of the model afferent system shown in Fig. 1. Again, they follow the trends of the data and represent a behavior similar to the near miss to Weber's law at high stimulus levels. (a) Noise bandwidth $\Delta F = 200$ Hz. Theory is matched to experiment at 40 dB SPL. (b) Noise bandwidth $\Delta F = 800$ Hz. Theory is matched to experiment at 50 dB SPL. Parameters for the theoretical curves are presented in Table I.

tion data [34] (see Table I and Fig. 5), as indicated in the previous section.

Aside from Q, the parameter that varies substantially from one paradigm to another is the scaling factor E_R (it ranges from ≈ 6 to $\approx 10^6$; see Table I). This is a free parameter of the model, although some of its variation results from substantial changes in auditory sensitivity with stimulus frequency (the outer- and middle-ear transmission function was ignored for pure-tone calculations). For loudness estimation, changing E_R causes the computed curves

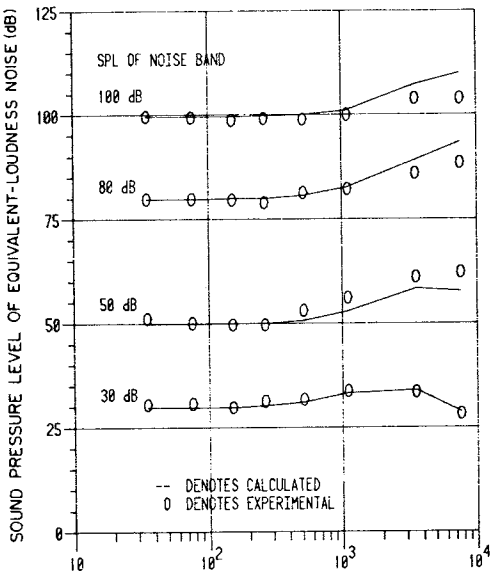
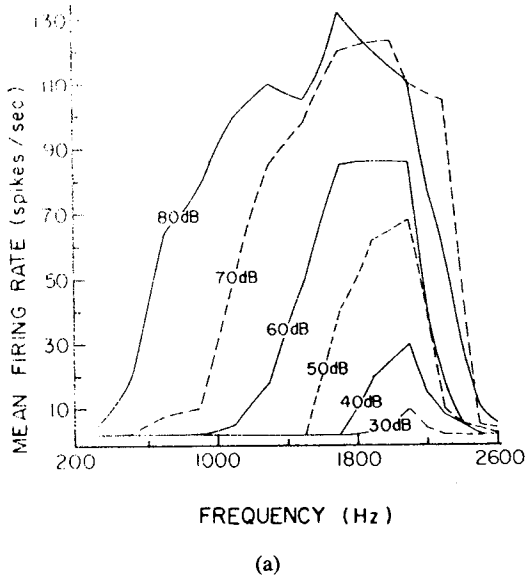


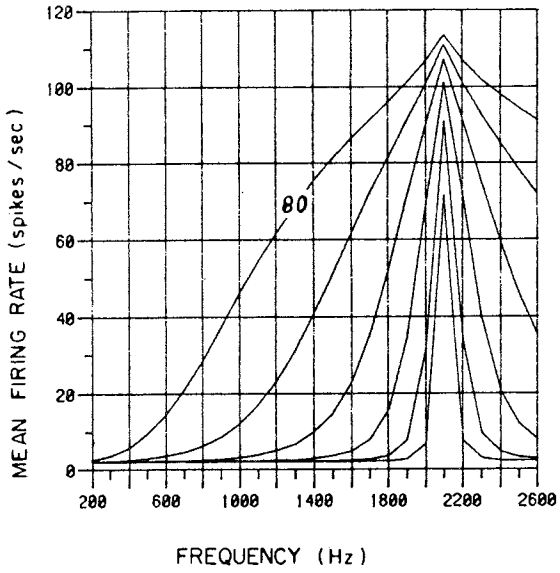
Fig. 5. Loudness summation effect (LSE) versus stimulus noise bandwidth. Experimental data are represented by open circles (adapted from Zwicker, Flottorp, and Stevens [26, Fig. 8]). Data show dependence of loudness on the bandwidth ΔF of flat-spectrum Gaussian noise of center frequency 1420 Hz. Total noise energy was maintained constant for each curve. Subjects adjusted the SPL of a band of noise, of center frequency 1420 Hz and bandwidth 210 Hz, to match the loudness at each noise bandwidth ΔF. Solid curves represent theoretical calculations based on the model afferent system shown in Fig. 1. Parameters for the theoretical curves are presented in Table I. Theoretical calculations follow the general trends exhibited by the experimental LSE.

to move horizontally, essentially without change of shape. It also has a modest role in adjusting the position of the intensity discrimination curves, and in altering the loudness summation curves. Since these latter data are referenced to specified preset stimulus intensities, the locations of the maximum loudness summation effect (LSE) will vary with E_R. The values for E_R listed in Table I represent the least mean-square-difference fit for the loudness summation data. Larger values produced an exaggerated LSE at 80 and 100 dB SPL, together with a diminished LSE at 30 dB SPL. Lower values produced the opposite. The parameters A in (1) and A_1 in (21) are absorbed into E_R, and therefore have no significance.

The scaling factor μ is a loudness estimation parameter [see (16)]. It moves the calculated loudness functions vertically, so that they can be brought into conjunction with the scale chosen by the subject. Since μ is presumably determined at the decision center, rather than peripherally, its value is governed in part by more distal effects (e.g., thinning and/or scaling of the point process before it reaches the decision center). For this reason, its absolute value means little, though it is interesting to observe that it assumes roughly the same value for the three sets of pure-tone loudness estimation data (see Table I). Neither intensity discrimination nor loudness summation depend on μ, since these paradigms are essentially self-normalizing.



(a)



(b)

Fig. 6. Mean firing rate (count mean in spikes/s) for an individual peripheral fiber stimulated by a pure tone of variable frequency and intensity (isointensity contours). (a) Neurophysiological data collected by Rose *et al.* [28] for an afferent fiber with a characteristic frequency of 2100 Hz in the squirrel monkey. (b) Results calculated on the basis of a single channel of the model afferent system shown in Fig. 1. The lowest isointensity contour shown is 30 dB and the highest is 80 dB (the curves are in steps of 10 dB). Parameters for the theoretical calculations are presented in Table I; the value of Q (= 50) was chosen to be in the range of most of our psychophysical data. Note that although they do not agree in detail, calculated and experimental curves have same general parametric dependences.

The quantity h is a parameter for intensity discrimination [see (17)]. It has been adjusted to bring the calculated intensity discrimination curves into approximate consonance with the experimental curves. E_R also plays a role in this. A slight additional adjustment in the horizontal position of the calculated curves is made in order to provide precise matching at the stimulus levels specified in Table I.

Though the same value of h suffices for all three intensity discrimination curves, this value cannot be taken too seriously because its choice has again been based on calculations with the peripheral union process rather than with the decision center process, where h is presumably operative. If the central process is a thinned version of the peripheral process N_c, as may be expected, the required value of h will be smaller than that specified in the table. All loudness paradigms are, of course, independent of h.

Measurements in single auditory fibers of the cat, chinchilla, rat, and squirrel monkey (see [4]) tell us that the spontaneous rate R_0 can vary substantially from fiber to fiber. The same is true of R_M, γ, α, E_R, $Q(f_0)$, N, and r. Appropriate distributions for these parameters could be constructed and incorporated into a more complex version of this model. We have chosen the simpler route of employing parallel channels that are identical in all respects, except for characteristic frequency. Evidently this simplification is satisfactory for the paradigms we have considered here. To investigate the role of such variations, however, we have examined the dependence of our results on the specific values used for R_0 and R_m. We carried out calculations for a theoretical collection of "high spontaneous" units with $R_0 = 20$ and $R_m = 200$. γ was fixed at 1.5 so that τ became 0.7 ms. α, N, and r remained the same as in Table I. The results for loudness did not change, but the results for intensity discrimination did. To match the data in this case, we had to set $Q(50\text{ Hz}) = 90$ and $Q(15\text{ kHz}) = 150$.

In short, most of the parameters used in our model have values comparable with their physiological counterparts and are consistent across paradigms. Some variation is expected, of course, because of model simplifications and species differences. Therefore, in spite of the large number of parameters in the model, it essentially has only two *free* parameters: E_R and $Q(f_0)$. μ and h are essentially fixed, and their values have little physical significance. The most critical parameters in providing good fits to the data are N and $Q(f_0)$, both of which are associated with the inner-ear tuning mechanism. The calculations that we have carried out should not, therefore, be viewed as merely an exercise in curve fitting.

B. Interplay Among $\rho(f_0)$, N, and $Q(f_0)$

The loudness functions (Fig. 2) predicted by the model using $N = 3/2$ are quite good. We carried out the same calculations without fiber density to examine the interplay among the cochlear mapping function and the values of N and Q required to fit the data. Good fits were obtained using $N = 2$ and a much lower value of Q. This is similar to the result obtained in [2], where the saturation function and cochlear mapping function were absent, and where $N = 2$ and Q was fixed at 18.7 (and was not a function of f_0).

This can be understood as follows. At low and moderate stimulus intensities, only one or a few channels near the stimulus frequency contribute to the overall neural count (see [2]), and the form of $\rho(f_0)$ is not important. At high levels, channels with characteristic frequencies near the stimulus frequency are largely saturated, and the loudness function continues to grow principally because of spread of excitation. In this region, the crude saturation model is satisfactory (see [1], [2]); it leads to a loudness function described by Stevens' power law [35], with slope $p = 1/2N$. It is here that $\rho(f_0)$ plays a crucial role. The origin of the $1/2N$ value for the slope in the crude saturation model is spread of excitation, primarily in the direction of increasing characteristic frequency. This is exemplified by the f_0^{2N} term in the denominator of (1). When fiber density is included, this term goes as f_0^{2N+1}, leading to Stevens' power law with a slope $p = 1/(2N + 1)$. In this simple scheme the uniform-in-log-frequency fiber density behaves like an additional pole in the system, lowering the slope of the loudness function. Thus, good fits can be obtained to high-intensity loudness data (which has a slope of $\approx 1/4$) by using $N = 3/2$ with fiber density or $N = 2$ without fiber density. If the experimental loudness function had a slope of $1/3$ instead of $1/4$, we would require $N = 1$ with fiber density or $N = 3/2$ without it.

For pure-tone intensity discrimination, the fit of the theory to the data is again quite good using $N = 3/2$, as is evident from Fig. 3; something close to the near miss to Weber's law [36], [37] emerges at high stimulus levels. In [1], we carried out calculations without fiber density, with $N = 2$, and with substantially lower Q, obtaining reasonable agreement with experiment. The situation is similar to that for loudness estimation: incorporating the cochlear mapping function requires a decrease in N by $1/2$ and a substantial increase in Q.

For noise intensity discrimination and loudness summation, calculations were carried out without fiber density in [3]. The best-fitting values of N and Q were 2 and 6.2, respectively. The computed curves followed the trend of the experimental data, but the fits of theory to data were not nearly as good as those shown in Figs. 4 and 5, using $N = 3/2$ and the outer- and middle-ear transmission function.

Finally we consider the squirrel-monkey isointensity contours displayed in Fig. 6. These were fit by a single channel version of the model displayed in Fig. 1, using $N = 3/2$ and $Q = 50$. The reader may recall that [4] was devoted to the response of individual neural fibers. All of the theoretical calculations explicitly shown there used $N = 2$ and $Q \approx 8$. The fits to the data were somewhat better than those displayed in Fig. 6(b). It seems that isointensity contours can be best fit by using $N = 2$ with low Q, or $N = 3/2$ with substantially higher Q. Only the $N = 3/2$ calculations are consistent with loudness estimation experiments when fiber density is included, however.

There is obviously a strong interplay among $\rho(f_0)$, N, and Q. Incorporating the cochlear mapping function generally requires a reduction of the value of N by $1/2$ and a substantially higher value of Q to fit the data.

C. Saturation, Refractoriness, and Spread of Excitation

A brief discussion regarding the roles that saturation, refractoriness, and spread of excitation play may be useful. The details of the saturation function turn out not to be an important determinant of the outcome of our calculations, provided that a reasonable saturation function is used. Both the logarithmic and the exponential forms provided good results. At high stimulus levels, in particular, the theoretical system is essentially transparent to the behavior of the receptor saturation function.

The two key high-intensity pure-tone laws, Stevens' law and the near-miss to Weber's law, arise from spread of excitation associated with the inner-ear linear filters. It is this element in the theory that appears to permit the auditory system as a whole to respond over a range of stimulus intensities exceeding 100 dB, while the individual channels experience their entire growth, R_0 to R_m, over a much smaller range.

To investigate the importance of refractoriness in the theory, we carried out a set of calculations identical to those described earlier, except for removing the effects of refractoriness altogether. This was achieved by setting $\tau = 0$ ($\gamma = 1.0$, $R_M = R_m$), leaving the (logarithmic) saturation function and all other parameters of the model intact. The loudness estimation and loudness summation results were both nearly unchanged, provided that E_R and μ were readjusted to fit the data. Pure-tone intensity discrimination data could also be fit in the absence of refractoriness, but the value of Q had to be increased somewhat (we used $Q(50\ \text{Hz}) = 60$ and $Q(15\ \text{kHz}) = 120$). Noise intensity discrimination results were not recomputed for $\tau = 0$. It appears that refractoriness, like saturation, is not an critical determinant of the character of the theoretical functions. The important thing is that at least one of them is present.

It is now not difficult to understand why the simpler model we considered in [1]–[3] worked so well (the LFRM included refractoriness but excluded saturation). The spread of excitation across the bank of linear filters gives rise to power-function loudness and near-miss intensity discrimination. In the simple LFRM scheme, saturation is provided by the refractoriness function (rectangular hyperbolic rather than logarithmic or exponential). The role of dead time, *per se*, was not crucial in arriving at these laws. This is apparent from Figs. 2, 4, and 5 of [1], where the calculations for $\tau/T = 0.05$, 0.005, and 0.0005 display essentially the same slopes on a log–log plot.

Nevertheless, it is useful and appropriate to incorporate refractoriness (and/or an effect like it such as phase locking) into the model, in addition to saturation. It provides a mechanism for reconciling the model calculations for count mean-to-variance ratio with single-fiber neurophysiological data. Typical values for γ_{exp} are 1.5 [30], demonstrating that the underlying peripheral point process is usually sub-Poisson. One may achieve this most easily by using the LFRM, but the system presented here provides a more complete theory in which the separate receptor saturation function limits γ to a value that is in accord with experiment.

VI. Comparison with Related Models

We now compare and contrast the model afferent system presented in Fig. 1 with a number of related models. We discuss several theoretical systems that have predictive capabilities for both loudness and intensity discrimination.

A. McGill–Goldberg–Penner Model

Our model evolved from the neural-counting formalism presented by McGill in 1967 [6] and modified by Penner in 1972 [38]. McGill considered the energy fluctuations of a superposed acoustic tone and interfering Gaussian noise. He derived the properties of a doubly stochastic Poisson counting process driven by these energy fluctuations, and associated the result with the mass flow of neural information in the auditory pathway. His calculations predicted Weber's Law for tone-in-noise and noise-in-noise intensity discrimination, in accord with experiment, and the deVries-Rose square-root law for tone-in-tone intensity discrimination. This latter result is at variance with experiment (see [1]). To eliminate this problem, and to provide a rudimentary model for loudness, McGill and Goldberg [36], [37] inserted into the model an E^p power-law saturation function ($0 < p \leqslant 1$), acting on the energy before Poisson transduction. This provided a power-law loudness function, and a near-miss-to-Weber's-law intensity discrimination function (a term coined by McGill and Goldberg) for pure tones, in accord with experiment. Penner [38] was concerned with the response of the McGill–Goldberg model to noise stimuli, and further modified it by considering a collection of parallel channels. Each channel consisted of an independent Poisson generator driven by a power–law function of the energy. The number of neural counts, rather than the energy, was summed. Our model replaces the power–law receptor saturation function by a form that is more consistent with physiological data (it includes spontaneous counts), introduces spread of excitation by incorporating a linear filter in each of the multiple channels, includes the effects of refractoriness, includes the cochlear mapping function, and incorporates the frequency dependence of the outer- and middle-ear transmission function. It ignores stimulus fluctuations, however.

B. Siebert Model

The theoretical framework that our model resembles most is that first discussed by Siebert [7] in 1968. He envisioned a system consisting of independent multiple channels, each with a linear filter driving a Poisson converter through a saturating nonlinearity. Spontaneous counts were included. A cochlear mapping function and a linear-filter middle-ear transmission function were incorporated into the model. In subsequent studies [16], [39], he developed a related construct based on nonlinear transduction to a nonhomogeneous Poisson process. This permitted

stimulus phase information to be retained in the neural point process.

There are a number of ways in which our model differs from Siebert's 1968 framework: 1) In addition to a saturating nonlinearity prior to Poisson transduction, we assume that a *refractoriness* modification follows the transduction. 2) We associate a simple *cascaded tuned-circuit* linear-filter transfer function with each channel. 3) We construct our model with *physiological* elements, avoiding optimal detection schemes. Thus we use a suboptimal *unweighted* sum of counts in a fixed counting time to arrive at a decision. A further and less important distinction concerns the computational methods. We carry out integrations in the frequency domain so that the role of the various elements on the overall system response can be determined. Siebert's variable of integration is distance along the cochlear partition so that the effects of fiber density, for example, are intrinsic to the integrand.

We have discussed item 1) in the previous section; incorporating refractoriness has little bearing on calculated psychophysical functions but permits the theoretical single-fiber counting process to exhibit $\gamma > 1$, in accordance with neurophysiological data.

The most crucial distinction that causes the two models to offer different predictions for psychophysical laws is item 2). In Fig. 1 we used a multiple-pole linear-filter inner-ear transmission characteristic for each channel, much as we did for the LFRM and for the crude saturation model (see [1]). It is this function (in particular the value of N) that leads to theoretical predictions of Stevens' loudness law and the near-miss discrimination law. Siebert [7] also used a tuned transmission function, but it consisted of a pure 10th-power frequency rolloff on the low-frequency side, and 20th-power frequency rolloff on the high-frequency side of the center frequency. His model predicted a roughly logarithmic loudness law, and Weber's discrimination law at high stimulus intensities. We would obtain these results as well, were we to use such steep rolloffs in our model. (It is not clear, however, whether Siebert's model would produce the near miss to Weber's law if he used rolloffs comparable to those we use.) Indeed, in 1974 Goldstein [40] demonstrated that Siebert's logarithmic loudness function becomes a power–law loudness function if a segment of the characteristic tuning in each channel is made less steep. Neither Siebert's nor Goldstein's models deal with the parameter Q. We have found that the details of the linear-filter transmission function must be dealt with carefully since they govern the character of the theoretical loudness and intensity discrimination functions.

Concerning item 3), we mention that although there is strong resemblance between the structure of Siebert's model and ours, the point of view is somewhat different. Siebert [7], [39] has devoted a great deal of effort to determining theoretical limits for the minimum variance achievable by various observables. He has, for example, shown that human frequency discrimination is considerably less acute than that predicted by the Cramer–Rao bound. He concludes that the processing of the information carried on the VIIIth nerve to higher auditory stations must be suboptimal. Our approach has been to deal with a structured peripheral model, with elements closely allied to known physiological function. Both approaches, of course, can provide useful and complementary information.

The single specific observable of interest to us is the unweighted sum of counts from all channels, in the fixed counting time T. Integration over a fixed time is a relatively simple process. Using N_c as the statistic corresponds to suboptimal processing since the occurrence-time information carried by the individual channels is averaged over the time T. In his 1968 work [7], Siebert considered multiple observables and obtained approximate results for the simultaneous optimal processing of the number of neural events in each individual channel. Later, in 1970 [39], he compared some of his results with human psychophysical performance and showed that the auditory system behaved suboptimally. In a subsequent study carried out in 1973 [16], that dealt principally with threshold auditory detection, he concluded that "the central processor does not make optimum use of the incoming pattern information, but bases its decision on simply a running estimate of the average firing rate in which the effect of past firing dies out exponentially with age." This is the suboptimal counting approach used by McGill [6] and in our model as well.

C. Whitfield Model

In 1967 Whitfield [41] developed an interesting qualitative model for several aspects of audition based principally on spread of excitation. He considered each neural channel as being in one of two states in binary fashion: either stimulated or not stimulated by the external signal. The number of contributing channels in his model increases with the stimulus intensity as does the perceived loudness. The increase is governed by the frequency response characteristics of the individual channels. Of course, a simple model of this kind ignores the details of the receptor saturation function and the stochastic character of the neural response. The cochlear mapping function and middle-ear transmission function are also omitted. The crude saturation model introduced in [1] is somewhat similar in character to Whitfield's model for pure-tone stimuli; presum-ably, therefore, his model would also lead to Stevens' law and the near-miss law, if the tuning characteristics of the channels were appropriately chosen. These laws are not in accord with experiment for weak stimulus levels, however, so that the usefulness of the Whitfield model is likely to be restricted to high intensities. Whitfield further postulated the existence of complex but unspecified processing in the central nervous system, to resolve the dilemma posed by Viemeister [42]. This relates to the observation that Weber's law obtains in the presence of band-reject background noise which should in theory act to prevent the spread of excitation. It is possible to explain both the loudness and intensity discrimination laws, and Viemeister's result, by considering the model afferent system in Fig. 1, in conjunction with the logarithmic satura-

tion function. This provides Weber's law at large stimulus levels, even for a single channel (see [4]). The spread of excitation produces the near miss.

D. Zwicker Model

Zwicker introduced a single-band excitation pattern model (EPM) that has proved useful for studying intensity discrimination [43]–[45] and loudness [34], [46]. In particular, Zwicker and Scharf [34] constructed an empirical model that accurately reproduces the loudness summation effect; agreement with experiment is very good. Their calculations make use of an excitation pattern derived from masking experiments [46], critical bands, and graphical procedures. The loudness within a critical band is assumed to be governed by power–law behavior for large signals; the 0.23 slope of the loudness function used by them is in close agreement with the value of about $\frac{1}{4}$ predicted by our model for $N = 3/2$ (when the cochlear mapping function is included). Although we do not explicitly consider critical bands in our formulation, it is possible that their effects are intrinsically manifested in the low value of Q that we must use to fit loudness summation data. In any case, we do not divide the frequency response characteristics of the auditory system into discrete bands as they do.

E. Florentine–Buus Model

Zwicker's excitation pattern model has recently been extended to a multiple-band version by Florentine and Buus [47]; performance is determined by information in all critical bands. This provides a substantially improved basis for computing intensity discrimination, and their procedure turns out to be very successful. Indeed, their formulation should be suitable for dealing with all of the psychophysical paradigms that we consider. The distinction between the EPM and the model afferent system shown in Fig. 1 lies not so much in the predictions of the models, as both are successful in this regard, but rather in their structures. We have attempted to reconcile psychophysical observations with the physiology of the peripheral auditory system, examining the role of each of the elements of the model along the way. The EPM, on the other hand, draws its structure from empirical excitation patterns.

F. Luce–Green Model

A number of other researchers have constructed detailed neural-counting and neural-timing models based on the point process observed from primary auditory fibers. Luce and Green [48] considered the successive spike interarrival times (IAT's) from first-order neurons. As with the McGill–Goldberg–Penner formulation, their model is predicated on a power–law relationship between count rate and energy; a number of other specialized conditions are also required (see [1] for further discussion). As indicated earlier, it is not likely that all of the timing information is used in the simple psychophysical paradigms that we have considered. Nevertheless, the Luce–Green model brings to mind the fact that the time required for the sum of neural counts to achieve a preset value is an observable closely related to the (inverse of the) number of neural events recorded in a fixed counting time T. The system shown in Fig. 1 could be adapted to accommodate such processing, but we have not carried out calculations of this nature.

VII. Summary and Conclusions

We have constructed a mathematical model for psychoacoustic performance using the physiological elements of the peripheral auditory system. The acoustic energy of the stimulus (pure tone or noise) is assumed to pass through a filter that characterizes the middle-ear transmission function, and then to excite a parallel bank of identical neural channels. Each channel represents a peripheral afferent fiber (or a collection of such fibers) and consists of a cascade of elements. Spontaneous neural activity is independently incorporated into each channel. The parallel collection of these channels has a density (in frequency) determined by the cochlear mapping function, and generates a union process at a more distal center in the nervous system. The statistics of the union count (in a fixed time) are then processed at a decision center in a manner that depends on the psychophysical paradigm. This random count number is assumed to contain all of the information for the examples we consider.

The system has been investigated in the context of four psychoacoustic paradigms: pure-tone loudness estimation, pure-tone and variable-bandwidth noise intensity discrimination, and variable-bandwidth noise loudness summation. In all cases, the theoretical results are in good agreement with human psychoacoustic data, provided that the parameters of the theoretical model are appropriately chosen.

The model has many parameters but almost all of these are fixed at values that are physiologically plausible and consistent (even though they have been selected from a mixture of data applicable to humans, monkeys, and cats). The loudness parameter μ and the intensity discrimination parameter h have little physical significance. There are two parameters for which the values vary substantially. These are the reference energy E_R and the quality factor $Q(f_0)$. An essential finding of our work is that we are able to obtain theoretical psychophysical functions that fit loudness and intensity discrimination data, for both tones and noise, using a neural counting model with two free parameters. The noise paradigms appear to require higher values of E_R and, at least for loudness summation, a substantially lower value of Q. It is as if the system widens its bandwidth in response to a broadband stimulus, and is consequently less sensitive because of the increased noise. This behavior is also associated with ideal energy detection [6], where the detection process increases its bandwidth in order to pay attention to the broadband energy characteristics of a pure-noise stimulus. An alternative hypothesis to explain this observation is that noise excites the system incoherently, resulting in a lower conversion efficiency.

The tuned linear-filter transmission function associated with each channel turns out to be a critical determinant of psychophysical behavior, primarily for strong pure-tone and narrowband stimuli. It permits the spread of excitation across the bank of channels that leads to the pure-tone laws: Stevens' power law for loudness and the near-miss to Weber's law for intensity discrimination. Theoretically these laws are a direct consequence of the assumed transfer-function characteristic of the individual neural channels, coupled with the cochlear mapping function. The two principal parameters associated with the linear filters, N and $Q(f_0)$, are most important. The *details* of saturation, Poisson transduction, and refractoriness are not important in this strong stimulus range.

The ideal Poisson converter is the stochastic element of the model. It, together with saturation and refractoriness, are psychoacoustically important primarily at low and intermediate stimulus levels and for broadband stimuli, viz., when spread of excitation is not dominant. Psychophysical functions calculated with either saturation or refractoriness alone turn out to be acceptable. The principal distinction relates to neurophysiological data. A Poisson process with presaturation, in the absence of refractoriness, leads to a neural count mean-to-variance ratio that is identically unity at all stimulus levels. A refractoriness-modified Poisson process without presaturation leads to $\gamma \gg 1$ at high stimulus levels. The use of both effects together, however, can provide $\gamma \approx 1.5$ at all stimulus levels, in accordance with experiment [30]. The form used for the saturation function is not particularly important. The logarithmic form has the advantage of unequivocally producing Weber's law in band-reject noise, thereby providing accord with Viemeister's [42] measurements. The sum of the neural counts in a fixed counting time appears to be a suitable decision variable. It is suboptimal and yields a detection law with the proper parametric dependences [6], [16].

What are the special conditions required for the model calculations to accord with psychoacoustic data? In the first place we have seen that spontaneous counts must be subtracted to obtain proper loudness functions. This is not difficult to justify (especially if we consider the neural spike train at higher centers in the auditory system). A more interesting condition is the substantially lower value of Q required to fit loudness summation, a matter that we have already discussed. To fit the psychoacoustic data we have had to choose $N = 3/2$. This requires that, within the confines of our model, the inner-ear tuning mechanism have a three-pole structure. Although the mechanisms of tuning in the inner ear are still not fully understood, one plausible explanation is via a cascade of filtering mechanisms, one or more of which is responsive to pressure [49], [50].

We have compared our model with a number of related theories in rather qualitative terms (viz., those of McGill–Goldberg–Penner, Siebert, Whitfield, Zwicker, Florentine–Buus, and Luce–Green). The theoretical construct that our work resembles most closely is that developed by Siebert [7]. The essential distinction is that he uses very sharp rolloffs for the tuning of the individual neural channels. This leads to a loudness function that is approximately logarithmic in form, and to Weber's intensity discrimination law at sufficiently high stimulus levels, neither of which is in accord with experiment. Though single-fiber neural rate tuning curves do indeed have very high slopes when weakly stimulated [12], [15], response functions that are considerably less steep are observed at moderate and high excitation levels [28], [51] so that the value $N = 3/2$ may not be unreasonable in that regime. This is, in fact, evident from the broadening of the isointensity contours with increasing level, as illustrated in Fig. 6.

There remain a number of reasonably well-understood physiological characteristics of the peripheral auditory system that we have not accounted for in our model afferent system. These include the distribution of excitation along the basilar membrane, the slightly nonuniform fiber innervation density, and the large variability of parameters associated with individual VIIIth-nerve fibers (e.g., the spontaneous and maximum firing rates). Other relevant factors are symmetric versus nonsymmetric linear-filter characteristics, individual tuning of the inner-ear cascade elements, monaural versus binaural processing, and the role of relative refractoriness [10]. Furthermore, most of the paradigms with which we have dealt involve short stimulus bursts that generate neural firing transients at stimulus discontinuities [15]. We have recently made some theoretical progress in dealing with this effect when the process is Poisson and nonparalyzable dead time is present [9]. These results can be incorporated into our calculations. In terms of processing at the decision center, we have postulated simple counting. A more realistic assumption would be to incorporate filtering at this stage as well [16], though such a limitation is not likely to be serious.

Two of the obvious limitations of our model involve the assumptions of a refractoriness-modified Poisson process and simple neural counting. These assumptions seem to be satisfactory for the simple paradigms with which we have dealt. However, we know that the single-fiber auditory neural point process exhibits substantial phase locking [52], which illustrates that the single-fiber process is not a simple Poisson modified by refractoriness [30]. Furthermore the individual channels may not be independent, and the contributions of their events to the central process may not be a simple union. In the context of the *counting* statistics, however, the refractoriness-modified Poisson and the true process may be quite similar in character [30]. However, the extraction of phase and frequency information almost certainly requires a scheme more complex than simple counting. It is reasonable to expect that experiments involving paradigms such as frequency discrimination and speech recognition will require a model in which the spike timing, and perhaps place information as well, is retained at the decision center. As an example, Srulovicz and

Goldstein [53] recently constructed a detailed probabilistic model using the "central spectrum" to represent the psychophysics of frequency discrimination.

Perhaps we should conclude with a brief indication of possible directions for a next-generation psychoacoustic model. It might be useful to consider one or more of the following: 1) using pressure rather than energy as the stimulus variable. This will provide a point process that captures phase information; 2) using a measure such as the spectral density of the point process, rather than simply the count number. This will permit timing information to be transmitted; 3) using critical bands in the system, to allow for the apparent adaptation of the detection process to the bandwidth of the stimulus; and 4) incorporating the effects of stimulus fluctuations into the model, which may be important at low intensities.

REFERENCES

[1] M. C. Teich and G. Lachs, "A neural-counting model incorporating refractoriness and spread of excitation. I. Application to intensity discrimination," *J. Acoust. Soc. Amer.*, vol. 66, pp. 1738–1749, 1979.

[2] G. Lachs and M. C. Teich, "A neural counting model incorporating refractoriness and spread of excitation. II. Application to loudness estimation," *J. Acoust. Soc. Amer.*, vol. 69, pp. 774–782, 1981.

[3] M. C. Teich and G. Lachs, "A neural counting model incorporating refractoriness and spread of excitation. III. Application to intensity discrimination and loudness estimation for variable bandwidth noise stimuli," *Acustica*, vol. 53, pp. 225–236, 1983.

[4] G. Lachs, R. A. Saia, and M. C. Teich, "A neural-counting model based on physiological characteristics of the peripheral auditory system. IV. Application to response of individual neural fibers," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, pp. 964–972, 1983.

[5] M. C. Teich, G. Lachs, and R. A. Saia, "A neural-counting model incorporating refractoriness and spread of excitation: Role of the peripheral auditory system in intensity discrimination," (abstract), *J. Acoust. Soc. Amer.*, vol. 71, p. S18, 1982.

[6] W. J. McGill, "Neural counting mechanisms and energy detection in audition," *J. Math. Psychol.*, vol. 4, pp. 351–376, 1967.

[7] W. M. Siebert, "Stimulus transformations in the peripheral auditory system" in *Recognizing Patterns*, P. A. Kolers and M. Eden, Eds. Cambridge, MA: MIT Press, 1968, pp. 104–133.

[8] J. J. Zwislocki, "On intensity characteristics of sensory receptors: A generalized function," *Kybernetik*, vol. 12, pp. 169–183, 1973.

[9] P. R. Prucnal and M. C. Teich, "Refractory effects in neural counting processes with exponentially decaying rates," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, pp. 1028–1033, 1983.

[10] M. C. Teich and P. Diament, "Relative refractoriness in visual information processing," *Biol. Cybern.*, vol. 38, pp. 187–191, 1980.

[11] B. Scharf, "Critical bands and the loudness of complex sounds near threshold," *J. Acoust. Soc. Amer.*, vol. 31, pp. 365–370, 1959.

[12] M. C. Liberman, "Auditory-nerve response from cats raised in a low-noise chamber," *J. Acoust. Soc. Amer.*, vol. 63, pp. 442–455, 1978.

[13] H. F. Schuknecht, "Neuroanatomical correlates of auditory sensitivity and pitch discrimination in the cat," in *Neural Mechanisms of the Auditory and Vestibular System*, G. L. Rasmussen and W. Windle, Eds. Springfield, IL: Charles C. Thomas, 1960.

[14] H. Spoendlin, "Innervation densities of the cochlea," *Acta Otolaryngol.*, vol. 73, pp. 235–248, 1972.

[15] N. Y.-S. Kiang, T. Watanabe, E. C. Thomas, and L. F. Clark, *D. charge Patterns of Single Fibers in the Cat's Auditory Nerve*. Cambridge, MA: Mass. Inst. Technology, 1965, research monograph no. 35.

[16] W. M. Siebert, "What limits auditory performance?," in *IVth Int. Biophysics Congr.*, USSR Academy of Sciences, Pushchino, 1973, pp. 399–413.

[17] P. R. Prucnal and M. C. Teich, "An increment threshold law for stimuli of arbitrary statistics," *J. Math. Psychol.*, vol. 21, pp. 168–177, 1980.

[18] G. Lachs, "Approximate photocount statistics for coherent and chaotic radiation of arbitrary spectral shape," *J. Appl. Phys.*, vol. 42, pp. 602–609, 1971.

[19] G. Lachs and R. Koval, "High frequency signal processor for the simulation of superposed coherent and chaotic radiation," *J. Appl. Phys.*, vol. 43, pp. 2918–2921, 1972.

[20] R. P. Hellman and J. J. Zwislocki, "Loudness determination at low sound frequencies," *J. Acoust. Soc. Amer.*, vol. 43, pp. 60–64, 1968.

[21] ———, "Monaural loudness function at 1000 cps and interaural summation," *J. Acoust. Soc. Amer.*, vol. 35, pp. 856–865, 1963.

[22] R. P. Hellman, "Growth of loudness at 1000 and 3000 Hz," *J. Acoust. Soc. Amer.*, vol. 60, pp. 672–679, 1976.

[23] R. D. Luce and D. M. Green, "Neural coding and psychophysical discrimination data," *J. Acoust. Soc. Amer.*, vol. 56, pp. 1554–1564, 1974.

[24] D. W. Grantham and W. A. Yost, "Measures of intensity discrimination," *J. Acoust. Soc. Amer.*, vol. 72, pp. 406–410, 1982.

[25] C. E. Bos and E. de Boer, "Masking and discrimination," *J. Acoust. Soc. Amer.*, vol. 39, pp. 708–715, 1966.

[26] E. Zwicker, G. Flottorp, and S. S. Stevens, "Critical bandwidth in loudness summation," *J. Acoust. Soc. Amer.*, vol. 29, pp. 548–557, 1957.

[27] E. Zwicker and Y. Yamada, "Lautstaerke von Tonkomplexen in Abhaengigkeit von der Bandbreite, vom Schallpegel und von zugefuegtem Breitbandrauschen," *Acustica*, vol. 53, pp. 26–30, 1983.

[28] J. E. Rose, J. E. Hind, D. J. Anderson, and J. F. Brugge, "Some effects of stimulus intensity on response of auditory nerve fibers in the squirrel monkey," *J. Neurophysiol.*, vol. 34, pp. 685–699, 1971.

[29] N. Y.-S. Kiang, "Peripheral neural processing of auditory information," in *Handbook of Physiology: The Nervous System, Vol. 3, Sensory Processes*, I. Darian Smith, Ed. Bethesda, MD: Amer. Physiol. Soc., 1984, pp. 639–674.

[30] M. C. Teich and S. M. Khanna, "Pulse-number distribution for the neural spike train in the cat's auditory nerve," *J. Acoust. Soc. Amer.*, vol. 77, Feb. 1985.

[31] A. T. Rasmussen, "Studies of the eighth cranial nerve of man," *Laryngoscope*, vol. 50, pp. 67–83, 1940.

[32] G. L. Romani, S. J. Williamson, and L. Kaufman, "Tonotopic organization of the human auditory cortex," *Science*, vol. 216, pp. 1339–1340, 1982.

[33] G. L. Romani, S. J. Williamson, L. Kaufman, and D. Brenner, "Characterization of the human auditory cortex by the neuromagnetic method," *Exp. Brain Res.*, vol. 47, pp. 381–393, 1982.

[34] E. Zwicker and B. Scharf, "A model of loudness summation," *Psychol. Rev.*, vol. 72, pp. 3–26, 1965.

[35] S. S. Stevens, "The measurement of loudness," *J. Acoust. Soc. Amer.*, vol. 27, pp. 815–829, 1955.

[36] W. J. McGill and J. P. Goldberg, "Pure-tone intensity discrimination and energy detection," *J. Acoust. Soc. Amer.*, vol. 44, pp. 576–581, 1968.

[37] ———, "A study of the near-miss involving Weber's Law and pure-tone intensity discrimination," *Percept. Psychophys.*, vol. 4, pp. 105–109, 1968.

[38] M. J. Penner, "Neural or energy summation in a Poisson counting model," *J. Math. Psychol.*, vol. 9, pp. 286–293, 1972.

[39] W. M. Siebert, "Frequency discrimination in the auditory system: Place or periodicity mechanisms?," *Proc. IEEE*, vol. 58, pp. 723–730, 1970.

[40] J. L. Goldstein, "Is the power law simply related to the driven spike response rate from the whole auditory nerve?," in *Sensation and Measurement*, H. R. Moskowitz *et al.*, Eds. Dordrecht-Holland: Reidel, 1974, pp. 223–229.

[41] I. C. Whitfield, *The Auditory Pathway*. London: Arnold, 1967.

[42] N. F. Viemeister, "Intensity discrimination of noise in the presence of band-reject noise," *J. Acoust. Soc. Amer.*, vol. 56, pp. 1594–1600, 1974.

[43] E. Zwicker, "Die elementaren Grundlagen zur Bestimmung der Informationskapazitaet des Gehoers," *Acustica*, vol. 6, pp. 365–381, 1956.

[44] ———, "Masking and psychological excitation as consequences of the ear's frequency analysis," in *Frequency Analysis and Periodicity Detection in Hearing*, R. Plomp and G. F. Smoorenburg, Eds. Leiden-Holland: Sijthoff, 1970, pp. 376–396.

[45] D. Maiwald, "Berechnung von Modulationsschwellen mit Hilfe eines Funktionsschemas," *Acustica*, vol. 18, pp. 193–207, 1967.

[46] E. Zwicker, "Ueber psychologische and methodische Grundlagen

der Lautheit," *Akust Beih. (Acustica)*, vol. 1, pp. 237–258, 1958.

[47] M. Florentine and S. Buus, "An excitation-pattern model for intensity discrimination," *J. Acoust. Soc. Amer.*, vol. 70, pp. 1646–1654, 1981.

[48] R. D. Luce and D. M. Green, "A neural timing theory for response times and the psychophysics of intensity," *Psychol. Rev.*, vol. 79, pp. 14–57, 1972.

[49] S. M. Khanna, "Interpretation of the sharply tuned basilar membrane response observed in the cochlea," in *Hearing and Other Senses: Presentations in Honor of E. G. Wever*, R. R. Fay and G. Gourevitch, Eds. Groton, CT: Amphora, 1983, pp. 65–86.

[50] A. C. Crawford and R. Fettiplace, "An electrical tuning mechanism in turtle cochlear hair cells," *J. Physiol. (London)*, vol. 312, pp. 377–412, 1981.

[51] A. R. Møller, "Use of pseudorandom noise in studies of frequency selectivity: The periphery of the auditory system," *Biol. Cybern.*, vol. 47, pp. 95–102, 1983.

[52] D. H. Johnson, "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," *J. Acoust. Soc. Amer.*, vol. 68, pp. 1115–1122, 1980.

[53] P. Srulovicz and J. L. Goldstein, "A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum," *J. Acoust. Soc. Amer.*, vol. 73, pp. 1266–1276, 1983.