A neural-counting model incorporating refractoriness and spread of excitation. II. Application to loudness estimation

Gerard Lachs

Pennsylvania State University, Department of Electrical Engineering, University Park, Pennsylvania 16802

Malvin Carl Teich

Columbia University, Department of Electrical Engineering, New York, New York 10027
(Received 2 June 1980; accepted for publication 2 December 1980)

In a previous paper [Teich and Lachs, J. Acoust. Soc. Am. 66, 1738–1749 (1979)] we demonstrated that an energy-based neural counting model incorporating refractoriness and spread of excitation satisfactorily described the results of pure-tone intensity discrimination experiments. In this paper, we show that the identical linear filter refractoriness model (LFRM) also provides proper results for pure-tone loudness estimation experiments at all stimulus levels. In particular, as the stimulus intensity increases from very low to moderate values, the model predicts that the slope of the intensity discrimination curve will climb from 1/2 toward 1, whereas the slope of the loudness function will gradually decline below 1 in this same region. For sufficiently high values of the stimulus intensity, the slopes calculated from a simplified (crude saturation) version of the model are found to be $1-1/4N$ for the intensity discrimination curve and $1/2N$ for the loudness function. The quantity N is the number of poles associated with the tuned-filter characteristic of the individual neural channels; it is the only important free parameter in the model. Appropriate values for N appear to lie between 2 and 4, providing an asymptotic slope for the intensity discrimination curve bounded by 7/8 and 15/16 (the near miss to Weber's Law), and an asymptotic slope for the loudness function bounded by 1/4 and 1/8. The results follow from the assumption that the neural concomitant of loudness is the number of impulses observed on a collection of parallel neural channels during a fixed observation time. Our calculations are supported by Hellman and Zwislocki's [J. Acoust. Soc. Am. 33, 687–694 (1961)] observation of unit slope for the loudness function at low intensities and provide a theoretical foundation, based on spread of excitation, for Stevens' power law at high intensities.

PACS numbers: 43.66.Ba, 43.66.Cb, 43.66.Fe, 43.66.Lj [BS]

INTRODUCTION

In Part I of this series of papers (Teich and Lachs, 1979, denoted I), we demonstrated 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 particular, we showed that the experimentally observed "near miss" to Weber's Law could be theoretically supported by a model incorporating essentially a single free parameter (N) at high levels of the baseline intensity. In that region, the slope m of the intensity discrimination curve was calculated from a simplified (crude saturation) model to be

$$m = 1 - 1/4N, \tag{1}$$

where N is the number of poles associated with the tuned linear filter characteristic of the individual neural channels. To fit the broad variety of existing data for 1-kHz tones, we found that satisfactory values for N were integers such that $2 \leq N \leq 4$, corresponding to $\frac{7}{8} \leq m \leq \frac{15}{16}$.

Having examined the performance of the linear filter refractoriness model (LFRM) for intensity discrimination, we naturally ask whether it can also provide a sensible theoretical basis for pure-tone loudness estimation. Fortunately, we can carry out such a study quite easily. We simply assign the average total number of impulses observed on a set of parallel neural channels during an unspecified, but fixed, counting interval or observation time [see Eq. (10) in I], as the neural concomitant of loudness, as suggested earlier by Békésy (see Schubert, 1978), Fletcher and Munson (1933), and others. We demonstrate that this conjunction does in fact provide good agreement between our predictions and existing data, thereby suggesting that a simple neural mechanism underlies loudness; indeed Wever (1949) long ago sought such a neural basis. Equally important, perhaps, is the demonstration that a single theoretical mechanism can be viewed as mediating both pure-tone intensity discrimination and loudness.

A related approach was undertaken by McGill and Goldberg (1968a, 1968b) some years ago. In the course of studying the near miss to Weber's Law, these authors developed a single-channel Poisson neural-counting model incorporating saturation of the stimulus energy. In their analysis, the detected neural count rate n took the form of a fractional-power compression of the stimulus energy ($n = \alpha E^p$, $0 < p \leq 1$), where α is a constant and p is about 0.2. Thus loudness, the perceptual concomitant of the neural count rate, took this same form. This outcome provided the necessary agreement with the extensive body of research that indicated a power-law growth of loudness with stimulus intensity or energy at high levels (Richardson and Ross, 1930; Stevens, 1955), but failed to provide a rationale for the growth of loudness at low and moderate stimulus levels. We note that Marks (1979) has discussed the relationship of the power law to other loudness scales from a rather broad perspective.

 0001-4966/81/030774-09$00.80

Stevens' law for the loudness L of a 1-kHz pure tone is of the form (Stevens, 1955, 1970)

$$L = kI^{0.3}, \tag{2}$$

where k is a constant and I is the stimulus intensity. However, this relationship is clearly unsatisfactory at levels below about 35 dB SPL (Hellman and Zwislocki, 1961). A number of authors including S. S. Stevens, Ekman, and Scharf and J. C. Stevens (see Scharf and Stevens, 1961) therefore suggested an alternate empirical relationship of the form

$$L = k(I - I_0)^{\gamma}, \tag{3}$$

where I_0 is a "threshold" stimulus value designed to bring the zero of the physical scale into coincidence with the zero of the physiological scale, and γ is a constant ~0.3. Lochner and Burger (1961) used Hellman and Zwislocki's (1961) loudness measurements to suggest that the relationship ought, instead, to be of the form

$$L = k(I^{\gamma} - I_0^{\gamma}), \tag{4}$$

with $\gamma \simeq 0.27$. More complex forms for the loudness function have also been hypothesized (Hellman and Hellman, 1975). Using an altogether different approach, Zwicker, Zwicker and Scharf, and others have developed geometrical methods for calculating loudness based on the spread of excitation along the basilar membrane (see Scharf, 1978, pp. 224–227). Scharf (1978) has written an excellent and comprehensive review chapter summarizing the current state of affairs in loudness estimation.

In the following, we derive the loudness function based on the LFRM, describe its qualitative features, and demonstrate that it is in accord with experimental data. In the course of our excursion, we recover Stevens' power law at high intensities as well as Garner's (1948), Zwislocki's (1969), and Hellman's (1976) unit slope at low intensities.

I. LOUDNESS MODEL

A. Pure-tone loudness estimation

Figure 1 is a block diagram for the LFRM and is reproduced from paper I, which described the elements of the model in detail. Only the output of the decision center is altered, to deal with loudness estimation rather than intensity discrimination. Associated with the decision center is a scalar loudness random variable $\mathcal{L}$, given by

$$\mathcal{L} = \kappa X. \tag{5}$$

The quantity X is a discrete random variable, well represented by a Gaussian distribution, expressing the total number of neural counts from all channels in the fixed counting time; κ is a constant. The notation is identical to that used in I.

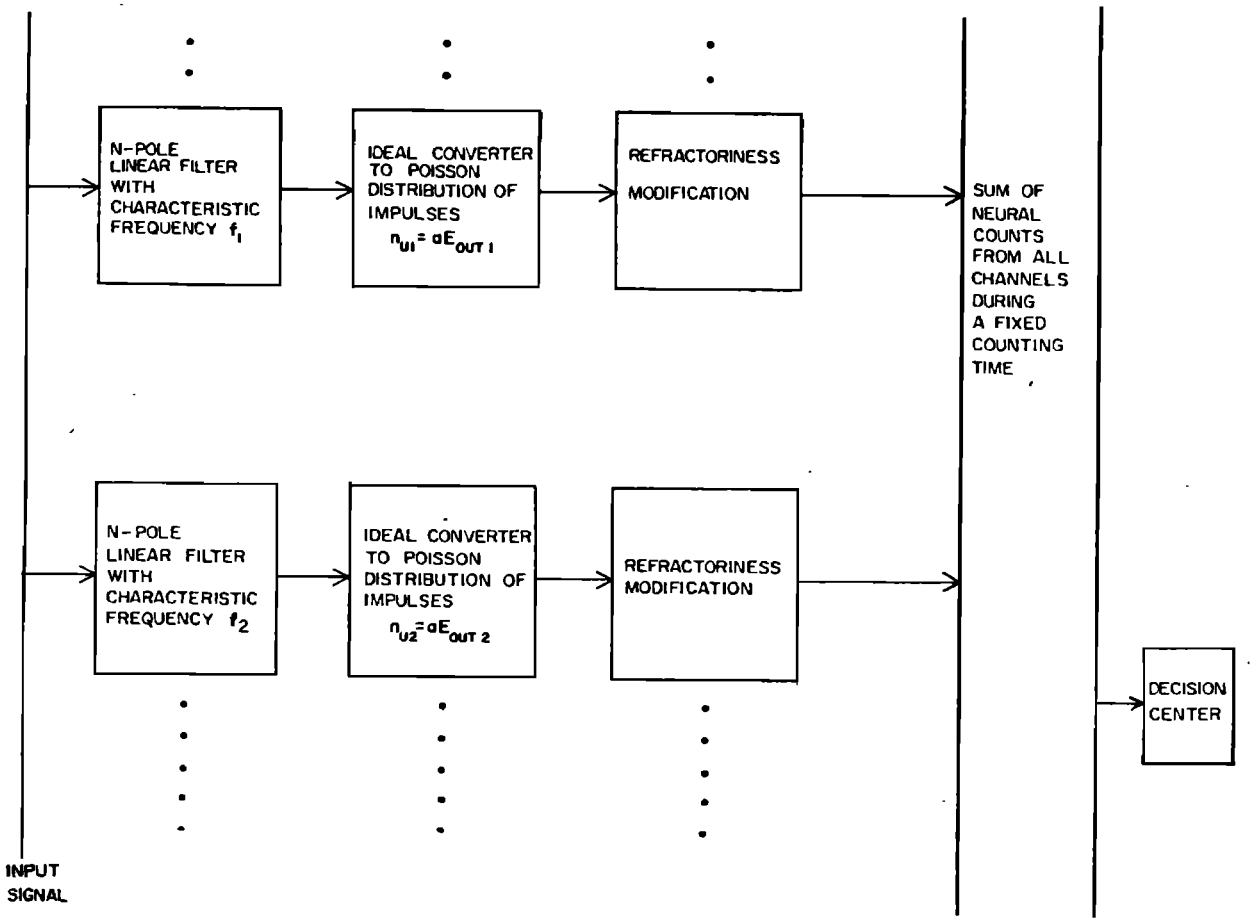

FIG. 1. Block diagram for linear filter refractoriness model (LFRM).

A variety of clever and interesting methods have been used to obtain loudness estimations from human subjects. Anyone collecting such data finds that the magnitude estimates tend to jump about unpredictably from trial to trial. In almost all studies, the results are based not on a single observation but rather on an average of two or more observations by each subject. Some researchers have constructed models in which the variability within a single subject emitting responses at different times plays a key role, whereas others have taken quite the opposite point of view and presented data that are averaged not only over observations for a given subject, but over many subjects as well. We will attempt to reconcile our model with both kinds of averaging.

In our characterization, we assume that the estimates are statistically independent from trial to trial, and that the loudness function L is obtained by forming the expectation $\mathrm{E}(\cdot)$ of Eq. (5), i.e.,

$$L = \mathrm{E}(\mathcal{L}) = \kappa \mathrm{E}(X) = \kappa \bar{N}_c . \tag{6}$$

Here $\bar{N}_c$ is the overall mean neural count given by Eq. (10) in I. Though we assume independence for simplicity in this paper, it must be emphasized that a considerable body of evidence demonstrates the existence of pronounced sequential effects in magnitude estimation (Cross, 1973; Ward, 1973; Green and Luce, 1974; Jesteadt *et al.*, 1977).

It is clear that for any *single* loudness observation, Eq. (5) will be satisfied. Since the variance of the overall neural count Σ_c^2 is described by Eq. (11) in I, the variance of $\mathcal{L}$ is

$$\mathrm{var}(\mathcal{L}) = \kappa^2 \Sigma_c^2 . \tag{7}$$

This provides an additional measure of the validity of our model, and also points up the intimate relationship between intensity discrimination and loudness estimation. Each provides a window on the underlying process. Nevertheless, it has long been assumed that discrimination data are very stable; the direction over the years has therefore been to seek loudness information from discrimination data, thereby removing the "contaminants" associated with human estimation. We find, interestingly, that both the discrimination and loudness predictions emerging from the LFRM are consistent with experimental data.

B. Explicit expression for the loudness function

The loudness function in Eq. (6) is expressed in terms of the overall mean neural count $\bar{N}_c$. From Eq. (10) of I, we have

$$\bar{N}_c = \frac{A'E}{2\pi} \int_{\omega_1}^{\omega_2} \frac{d\omega_0}{[1 + Q^2(\omega_T/\omega_0 - \omega_0/\omega_T)^2]^N + EA'(\tau/T)}, \tag{8}$$

where ω_T is the circular frequency of the stimulus (test tone) in rad/sec, ω_0 is the characteristic or center frequency of the neural channel in rad/sec, N is the number of poles associated with the linear filter for each channel, Q is the ratio of the characteristic frequency to the 3-dB bandwidth for a single-pole filter, E is the input energy of the test tone, τ is the dead-time interval, T is the counting-time interval, and A' is a constant. When the stimulus consists of short tone bursts, as is the case in all of the experiments considered here, it is immaterial whether its magnitude is expressed in terms of energy E or in terms of intensity I. The limits of integration are chosen to agree with the frequency limits over which peripheral auditory fibers are assumed to respond. In this paper $\omega_1 = 2\pi(50)$ and $\omega_2 = 2\pi(15\,000)$ are used as limits, but this choice is not critical. The factor $1/2\pi$ that appears in Eq. (8) is equivalent to the assumption of a uniform fiber density of 1 fiber per Hz. It should be noted that the variable of integration is the characteristic frequency of the filter and not the frequency of the incident tone which is fixed at ω_T.

C. Method of computation

To describe how the various parameters in Eq. (8) were selected to match a set of experimental data, it is useful to define the integral J,

$$J(E, A'\tau/T, \omega_T, Q, N) = \int_{\omega_1}^{\omega_2} \frac{d\omega_0}{[1 + Q^2(\omega_T/\omega_0 - \omega_0/\omega_T)^2]^N + EA'(\tau/T)}, \tag{9}$$

and the constants

$$B_1 = \kappa A'/2\pi , \tag{10a}$$

$$B_2 = A'(\tau/T) . \tag{10b}$$

Then, from Eqs. (6), (8), (9), and (10), the loudness function is

$$L = B_1 E J(E, B_2, \omega_T, Q, N) . \tag{11}$$

We begin with two coordinates of the data. Let $\{L_1, E_1\}$ and $\{L_2, E_2\}$ represent the lowest and highest experimental loudness and energy values, respectively. We then have

$$L_1 = B_1 E_1 J(E_1, B_2, \omega_T, Q, N) \tag{12a}$$

and

$$L_2 = B_1 E_2 J(E_2, B_2, \omega_T, Q, N) . \tag{12b}$$

Selecting specific values for ω_T and Q, we obtain B_2 from the solution of

$$\frac{L_2}{L_1} = \frac{E_2}{E_1} \frac{J(E_2, B_2, \omega_T, Q, N)}{J(E_1, B_2, \omega_T, Q, N)}, \tag{13}$$

with the aid of an IBM 370/3033 computer using numerical integration and the IMSL routine called ZBRENT. The constant B_1 is then obtained by substituting the computed value of B_2 into Eq. (12a). These values are then employed as initial values in a minimum-mean-square difference-parameter optimization procedure to determine the final values of B_1 and B_2. This was accomplished by employing the IMSL routine called ZXSSQ to minimize the normalized quantity

$$\sum_{\substack{\text{all data} \\ \text{points}}} \left(\frac{L_{\text{exp}} - L_{\text{th}}}{L_{\text{exp}}} \right)^2 ,$$

where L_{exp} represents the experimental loudness estimate and L_{th} represents the theoretical loudness value

computed with the aid of Eqs. (9)–(11). Finally, the theoretical pure-tone loudness function is obtained from Eqs. (9)–(11), using the optimized values of B_1 and B_2, in terms of the parameters N, ω_T, and Q. We discuss the dependence of L on these parameters in Sec. IIC.

We next deal with the qualitative behavior of the loudness function at low and high values of the stimulus intensity.

D. Qualitative features of the loudness function

The loudness function is plotted on double-logarithmic coordinates ($\log L$ versus $\log E$) to match experimental data, which are usually graphed as log loudness estimate versus sound pressure level or sensation level (SPL in dB *re* 0.0002 dyne/cm^2 or SL). We examine the behavior of this function at both low and high values of the stimulus energy, and compare it with the intensity discrimination curve presented in paper I.

For E sufficiently small, the quantity $EA'(\tau/T)$ in Eq. (9) is negligible (there is no refractoriness), so that the integral J is independent of E. Therefore, using Eq. (11), and the fact that B_1, B_2, ω_T, Q, and N do not depend on E, we have

$$L = B_1 E J(B_2, \omega_T, Q, N). \tag{14}$$

In this limit, therefore,

$$\log L = \log B_1 J(B_2, \omega_T, Q, N) + \log E. \tag{15}$$

This equation describes a straight line loudness curve of unity slope, in accord with Garner's (1948) postulate that loudness is proportional to sound intensity near threshold. Zwislocki's (1969) analytic theory of temporal summation also requires such a proportionality. Indeed, by eliminating a number of sources of bias, Hellman and Zwislocki (1961) obtained reproducible loudness curves near threshold, and showed that the loudness function approaches unity slope at the lowest level, 4 dB SL, at which they could obtain estimates. This result may be understood as follows: In the region of low E where refractoriness is absent, L is proportional to E since only one, or perhaps a small number of channels, contribute to the neural count and the LFRM predicts that the overall neural count random variable will be Poisson. As discussed in paper I, Poisson statistics lead to the deVries–Rose square-root law for intensity discrimination in this region.

At somewhat higher stimulus levels, from about 10 to 35 dB, saturation from refractoriness sets in, driving the slope of the intensity discrimination curve toward unity in accordance with the calculations of van der Velden (1949) and Bouman *et al.* (1963). This decrease of discriminability is manifested in the loudness function as a gradual decrease in the slope below unity. It is in this region that the most detailed information about the relative importance of various saturation mechanisms may be available.

At levels above 35 dB, channels with characteristic frequencies near the stimulus frequency are largely saturated. Nevertheless, the loudness function continues to grow because of spread of excitation. As demonstrated in I, the intensity discrimination curve exhibits the near miss to Weber's Law in this region; improved discrimination arises from the increasing response of strongly driven channels whose characteristic frequencies are far removed from the test-tone frequency. Because spread of excitation is the dominant mechanism at high levels, the details of the saturation are not important there. Thus we obtained a simple analytic expression for the slope of the intensity-discrimination curve in this region by means of the crude saturation model. We now proceed by employing this same tactic for loudness. Using Eq. (18) of I, assuming a sufficiently high level for E, we obtain

$$E(X) \approx C_3 E^{1/2N}, \tag{16}$$

where C_3 is a constant (see I). Defining a new constant $B_3 = \kappa C_3$, and using Eq. (6), we obtain

$$L \approx B_3 E^{1/2N}, \tag{17}$$

which we rewrite as

$$\log L = (1/2N) \log E + \log B_3. \tag{18}$$

The result in Eq. (18) is the promised power law for the loudness function at high levels of stimulus intensity, with slope

$$p = 1/2N. \tag{19}$$

According to our model, energy presented to the auditory system is nonlinearly encoded into frequency spread in a way that matches the wide range of energy values in the outer stimulus world to the limited processing capacity of the inner neural world. From a neurophysiological point of view, the power law may be understood as arising from the filter characteristic of the skirts of the tuned neural channels.

It is useful to compare the predictions of the LFRM with those of the McGill–Goldberg (1968a, 1968b) single-channel neural counting model. At high stimulus levels, both models produce the same functional form for the slope m of the intensity discrimination curve expressed in terms of the slope p of the loudness function

$$m = 1 - p/2. \tag{20}$$

A moment's reflection provides the underlying reason. Both the McGill–Goldberg model and the crude-saturation version of the LFRM (for the unsaturated channels) assume that the underlying statistics are precisely Poisson. In the former case, the power-law relationship is introduced by the hypothesized fractional power-law saturation of the neural count rate ($n \sim E^p$, $0 < p \leq 1$), whereas in the latter case it is introduced by the hypothesized spread of excitation across the bank of N-pole linear filters ($\overline{N_c} \sim E^{1/2N}$). For the LFRM, the fine structure of the saturation plays a role only at intermediate stimulus intensities (10–35 dB), whereas for the McGill–Goldberg model it is the crucial determinant that provides both straight-line intensity discrimination and loudness functions for *all* values of stimulus intensity.

II. COMPARISON WITH EXPERIMENTAL DATA

A. General description

Figures 2–5 present a broad range of loudness data. The solid and open dots represent data points obtained by a number of researchers using various loudness-estimation techniques. All ordinates bear the label "loudness estimate," but this subsumes a broad range of experimental conditions. Some figures present data collected with a reference standard, and some present data collected without a standard. Data are sometimes binaural, sometimes monaural, sometimes averaged over many subjects, and sometimes collected for a single subject. The test-tone frequency ranges from 100 Hz to 3 kHz. We deal only with magnitude estimation data; neither magnitude production nor magnitude balance data are considered in this paper.

The abscissa is labeled "intensity in dB" in all figures. It is not important in our method of calculation that the experimental data presented in Fig. 2 are in terms of sensation level (SL), while those presented in Figs. 3–5 are in terms of sound-pressure level (SPL).

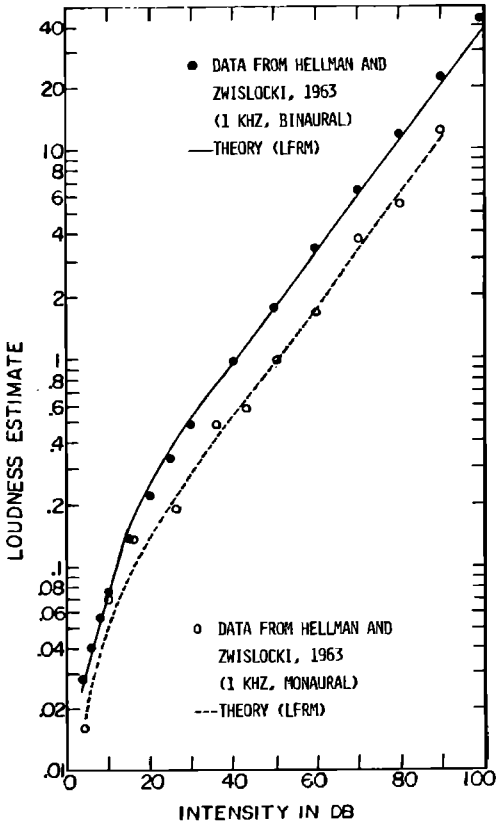

FIG. 2. Plot of loudness estimate versus stimulus intensity in dB. Experimental data (solid dots) adapted from Fig. 1 of the paper by Hellman and Zwislocki (1963). These data were obtained binaurally at 1 kHz, with a reference standard of unity at 40 dB SL, and represent an average over several studies. Solid curve is the LFRM loudness function with $N=2$. Experimental data (open dots) adapted from Fig. 8 of the paper by Hellman and Zwislocki (1963). These data were obtained monaurally at 1 kHz, with no reference standard, and represent an average over nine listeners. Dashed curve is the LFRM loudness function with $N=2$.

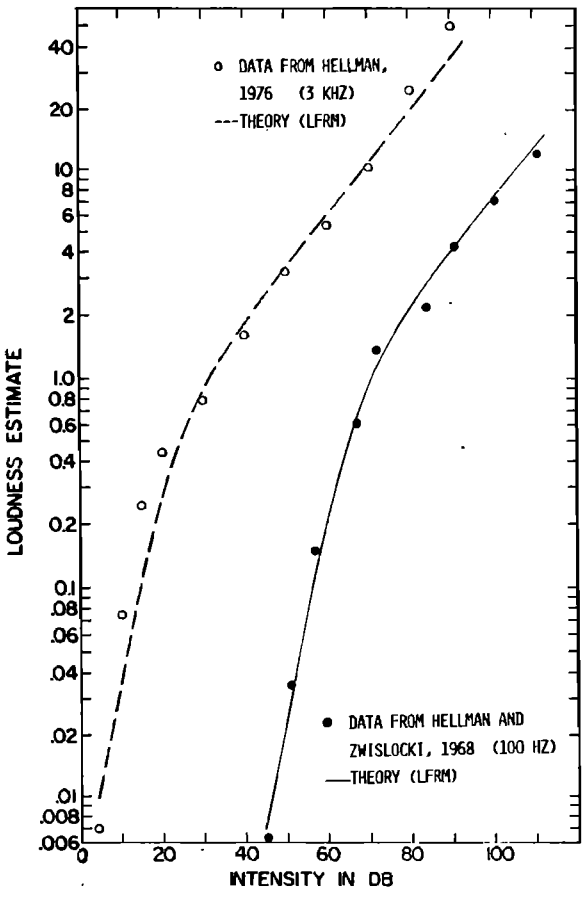

FIG. 3. Plot of loudness estimate versus stimulus intensity in dB. Experimental data (solid dots) adapted from Fig. 2 of the paper by Hellman and Zwislocki (1968). These data were obtained monaurally at 100 Hz, with no reference standards, and represent an average over nine listeners. Solid curve is the LFRM loudness function with $N=2$. Experimental data (open dots) adapted from Fig. 1 of the paper by Hellman (1976). These data were obtained binaurally at 3 kHz, with no reference standards, and represent an average over ten listeners. Dashed curve is the LFRM loudness function with $N=2$.

This is because the stimulus intensity E in Eq. (8) is always multiplied by the constant A', and the process of selecting the constants B_1 and B_2 (see Sec. IC) internally adjusts A' to yield an optimal fit to the data, regardless of the scale used for the stimulus intensity. An explicit value for A' can be obtained only if either the dead-time ratio τ/T or the loudness constant κ is known.

The solid and dashed curves in Figs. 2–5 represent calculations based on the LFRM, where the constants B_1 and B_2 are selected by the procedure described in Sec. IC. In most cases the fit was obtained by minimizing the mean-square of the normalized difference between the experimental and theoretical loudness values. In Fig. 4, however, the fit was effected by matching the experimental data at the endpoints.

The theoretical curves were calculated using a nonsymmetric frequency response function for the individual neural channels. The response below the characteristic frequency of each channel was chosen to behave as an N-pole linear filter, whereas the response above

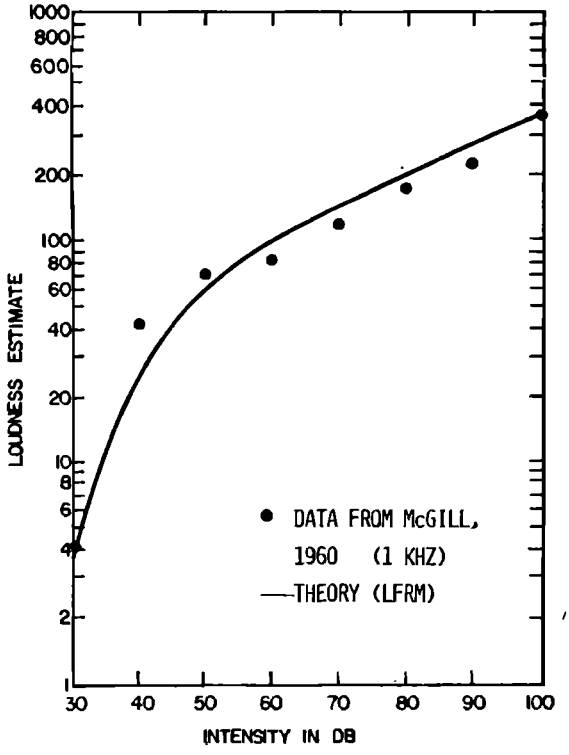

FIG. 4. Plot of loudness estimate versus stimulus intensity in dB. Experimental data (solid dots) adapted from Fig. 1 of the paper by McGill (1960). These data were obtained binaurally at 1 kHz, with reference standards presented at the beginning of a data run. Data are for a single subject. Solid curve is the LFRM loudness function with $N=4$, fitted at the endpoints.

the characteristic frequency was chosen to behave as a $2N$-pole linear filter. We were motivated in this choice by the masking data collected by Egan and Hake (1950) and by Zwicker (1974). Nevertheless, as we discovered in our study of intensity discrimination (see I), the results differ only insignificantly from those obtained using a symmetric N-pole linear filter response function.

B. Specific comparisons

The data in Fig. 2 are adapted from Figs. 1 and 8 of the paper by Hellman and Zwislocki (1963). The solid dots in Fig. 2 represent an average over several studies, for the binaural presentation of a pure tone at 1 kHz, with a reference standard of 1 at 40 dB SL. Above 30 dB SL, the data obey a power function with slope 0.27. In the vicinity of the threshold, the experimentally observed slope approaches 1. The open dots shown in Fig. 2, by contrast, were obtained without standards for a group of 9 listeners. The stimulus was again a pure tone at 1 kHz, but was monaurally presented. Above 30 dB SL, the data obey a straight line with a slope that is slightly greater than 0.27. The solid dots in Fig. 3 are adapted from Fig. 2 of the paper by Hellman and Zwislocki (1968). They represent magnitude estimates averaged over 9 subjects, for the monaural presentation of a pure tone at 100 Hz, without reference standards. The authors did not report a value for the slope of the measured loudness function. Similarly, the open dots in Fig. 3 are adapted from Fig. 1 of the paper by Hellman (1976). They represent magnitude estimates averaged over 10 subjects, for the binaural presentation of a pure tone at 3 kHz, without reference standards. The author reported a slope of 0.285 for this data.

In the context of the LFRM with N restricted to integer values, the best fit to all of the data described above is obtained for $N=2$, as represented by the solid and dashed curves in Figs. 2 and 3. It is evident that the curves produced by the LFRM are in good accord with the data; the relevant parameters are presented in Table I. From Eq. (17) it is clear that at high levels of the stimulus intensity, the theoretical loudness function L will be related to the intensity by a power-law function with an exponent of $\frac{1}{4}$. Thus the slope on a log–log plot will be 0.25 [see Eq. (18)], in good agreement with the experimental values.

We have also obtained very good fits for other experimental data which we do not present because of space limitations. In particular, we find (using endpoint fits) that optimal theoretical loudness functions are obtained when $N=2$, for experimental data from Fig. 1 of Scharf and Stevens (1961), Fig. 5 of Lochner and Burger (1962), and Figs. 2–4 of Hellman and Zwislocki (1963).

All of the data discussed to this point were obtained by averaging over multiple subjects, and all were best fit by $N=2$. By contrast, the data in Figs. 4 and 5 are

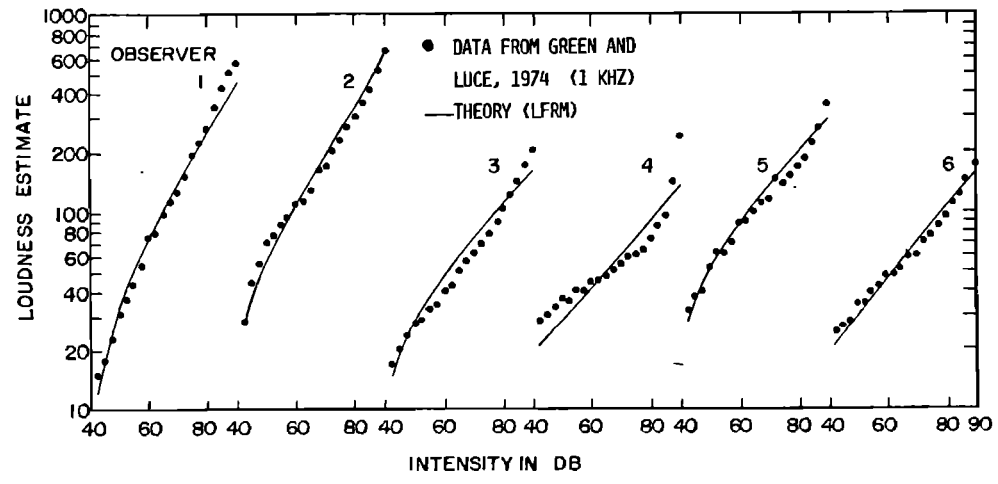

FIG. 5. Plot of loudness estimates versus stimulus intensity in dB. Experimental data (solid dots) adapted from Fig. 4 of the paper by Green and Luce (1974). These data were obtained binaurally at 1 kHz, with no reference standard. Data are shown for six observers. Solid curves are best-fitting LFRM loudness functions, which have N values that are either 2 or 3 (see Table I). The range of stimulus intensity for each curve is 42.5 to 90 dB SPL.

TABLE I. Parameters for LFRM theoretical curves in Figs. 2–5.

Fig. no.	Identification	Stimulus frequency (Hz)	Best N	Single-pole Q	B_1	B_2
2	solid curve	1000	2	18.7	3.323×10^{-4}	7.586×10^{-2}
2	dashed curve	1000	2	18.7	2.466×10^{-4}	1.125×10^{-1}
3	solid curve	100	2	18.7	7.110×10^{-8}	3.240×10^{-7}
3	dashed curve	3000	2	18.7	3.837×10^{-5}	7.916×10^{-3}
4	solid curve	1000	4	11.2	1.056×10^{-4}	1.122×10^{-4}
5	observer 1	1000	2	18.7	2.719×10^{-5}	3.574×10^{-5}
5	observer 2	1000	2	18.7	1.249×10^{-4}	1.794×10^{-4}
5	observer 3	1000	3	13.7	5.223×10^{-5}	1.288×10^{-4}
5	observer 4	1000	3	13.7	3.270×10^{-3}	2.334×10^{-2}
5	observer 5	1000	3	13.7	1.151×10^{-4}	1.638×10^{-4}
5	observer 6	1000	3	13.7	1.997×10^{2}	1.433×10^{4}

for individual listeners, which appear to exhibit a broader range of slopes (Stevens and Guirao, 1964).

The data points in Fig. 4 are adapted from Fig. 1 (subject R, data collected 1 August 1957) of the paper by McGill (1960). The dots were obtained by means of a measurement paradigm in which the subject marked a point along a 6-in.-long line, to estimate the loudness of the tone. The 1-kHz pure-tone stimulus was presented binaurally with reference standards provided at the beginning of a data run. The best fitting loudness function derived from the LFRM, based on fitting the endpoints, is shown as the solid curve in Fig. 4, and corresponds to the choice $N=4$ (see Table I). The value $N=4$ was also found to provide the best fit to the data presented in Fig. 2 of the paper by McGill (1960). In most cases there was only a modest difference between theoretical curves generated using the optimization technique to obtain B_1 and B_2 and those generated by selecting B_1 and B_2 to fit the endpoints. The endpoint matching procedure was employed in Fig. 4 to show that this method can also provide a good fit to experimental data. All other graphed results are based on the optimization procedure.

Figure 5 shows a comparison between the LFRM-based computations (solid curves) and the experimental data (dots) reported by Green and Luce (1974). The dots in Fig. 5 are obtained from the measured numerical data values originally used to generate Fig. 4 in Green and Luce's paper, and describe experimental loudness functions for six individual observers. The dots represent experimental estimates for 1-sec duration binaurally presented 1-kHz pure tones, without a reference standard. The best-fitting loudness functions derived from the LFRM are shown as the solid curves in Fig. 5; the optimal value of N is either 2 or 3 (see Table I). In this case, the minimum-mean-square fits were substantially better than the endpoint fits. We observe that the agreement of theory and data is not as good as that shown previously. We offer no explanation as to why this is so. Improvement could be obtained, of course, were N not restricted to integer values.

The best-fitting integer N values, along with the resulting values of B_1 and B_2, are presented in Table I for all figures. In all cases, we have arbitrarily set $Q_{10dB}=12.5$ which requires that the single-pole Q used in Eq. (9) be

$$Q=(10^{1/N}-1)^{1/2}\cdot 2Q_{10dB}\cdot\frac{(2Q_{10dB}+1)}{(4Q_{10dB}+1)}. \tag{21}$$

This explains how we arrived at what appear to be highly specific values of Q in Table I. The choice of parameter values is discussed in the next section.

C. Choice of parameter values

We have seen in Sec. ID that the theoretical loudness function depends principally on the number-of-poles parameter N. In particular, the slope of the function at high intensities depends only on N whereas the slope at low intensities is unity, independent of all parameters in the model. From the discussion in Sec. IIB, the value $N=2$, which provides $p=\frac{1}{4}$ and $m=\frac{7}{8}$, seems to play a special role (see Table I). Indeed, the associated 12 dB/oct slope of the linear-filter transfer function is in reasonable accord with the masking data collected by Zwicker (1974). Values of the slope of this magnitude may perhaps be reconciled with the (much greater) slopes of the neural tuning curves measured at the primary afferent fibers by recalling that the latter depend strongly on the excitation level. This point has been made particularly forcefully by Zwicker (1974) in his excellent contribution. Furthermore, if the slope of the loudness function is task dependent, as has been suggested by Marks (1979) and others, this is easily accomodated in our model in terms of an adaptive filter. It is also possible that loudness estimation involves processes more central than the peripheral auditory system (see footnote 4 in paper I).

The dependence of L on other system parameters (ω_T, $Q, A', \tau/T, \kappa$) is more subtle. These parameters will principally affect the shape of the loudness function at intermediate stimulus levels, as well as the absolute position (as opposed to the slope).

The quantity ω_T is a property of the stimulus, and our theoretical loudness functions are in accord with all of the experimental data that we attempted to fit (the range was 100 Hz to 3 kHz). The shape of L will depend on

ω_T, and we are able to predict this dependence. With regard to Q, we find the following: For $N=2$ and arbitrary ω_T, the computer-generated LFRM results are independent of Q_{10dB} for $Q_{10dB} \geq 7.5$. This was explicitly demonstrated by driving Q_{10dB} as high as 25. We observe that the theoretical loudness function begins to diverge from the experimental data when Q_{10dB} is chosen below about 7.5; in paper I, the pure-tone intensity discrimination curve was found to be independent of Q_{10dB} for the two values employed, $Q_{10dB}=3$ and $Q_{10dB}=10$. From a mathematical point of view, we have assumed throughout that all poles of the linear filter have identical center frequencies. A small difference could change the model parameters substantially.

In our treatment, the three parameters A', τ/T, and κ are not specified since they are subsumed in the (two) constants B_1 and B_2 that are fit to the experimental data [see Eq. (10)]. Interestingly, we may ultimately be able to extract explicit values for A', τ/T, and κ by making use of the additional information contained in the experimental variance of the loudness estimate [see Eq. (7) and Green and Luce, 1974]. This would provide us with three equations and three unknowns. We must keep in mind, however, that the values of the model constants B_1, B_2, and Q which produce a good fit to the empirical data with the simple LFRM presented here (see Table I) will change when the details of the peripheral auditory system are incorporated. The saturation of the peripheral receptors in the intermediate intensity region provides an example, as indicated in the Conclusion. To provide an additional measure of the validity of our model, it will be particularly valuable to simultaneously extract the relevant parameters from intensity discrimination data (see I).

III. CONCLUSION

It is clear that the LFRM provides satisfactory results both for pure-tone intensity discrimination and loudness estimation at all levels of the stimulus intensity. It therefore provides a single theoretical framework, within which intensity discrimination and loudness estimation are perceived as representing two manifestations of a common underlying neural basis. It would be desirable to collect intensity discrimination and loudness data from the same set of subjects to ascertain whether N is indeed a constant of the subject.

Much work remains to be done. To begin with, we must obtain the theoretical loudness function in the presence of background noise, and for stimuli other than a pure tone. We are just now developing the appropriate mathematical tools to carry out this task for Gaussian noise stimuli (Vannucci and Teich, 1981). Preliminary results indicate that the behavior is in accord with experiment.

And, of course, we should incorporate into our model the myriad known physiological characteristics of the peripheral auditory system. These include the nonuniformly distributed acoustic energy along the basilar membrane associated with the envelope of the traveling acoustic wave, the approximately logarithmic relationship between distance along the basilar membrane and best frequency, the nonlinear receptor response, the nonuniform fiber innervation density, and the nonlinear (with stimulus intensity) active fiber density arising from the spread in range over which different fibers initiate firing. Other relevant factors are symmetric versus nonsymmetric linear filter characteristics, the use of critical bands, and monaural versus binaural processing.

We shall report the results of a more complete study incorporating these components at a future date, but we mention here a number of preliminary findings. One rather curious outcome is that the assumption of a uniform-in-linear-frequency fiber density provides a better match of theory with data than does a uniform-in-log-frequency fiber density. As if to echo this result, employing this latter density in the crude saturation model provides a logarithmic, rather than (the proper) power-law, growth of the theoretical loudness function with stimulus energy (see also Siebert, 1968). Both densities lead to the near miss to Weber's Law for intensity discrimination, however, as pointed out in I. We have also incorporated a nonlinear element in our model, to account for the saturation of the hair-cell response, along with refractoriness and spread of excitation. The result is a depression of the theoretical loudness function at low and intermediate stimulus energies, leaving the power-law behavior intact at high stimulus levels where spread of excitation is the dominant mechanism. This result is relatively independent of the details of the saturation function, and is in accord with our expectations. We did not use critical bands in calculating the loudness function, though it appears they should be incorporated in our model if we wish to address the masking of tones by tones.

Lastly, we point out that we have implicitly ignored any distinction between binaural and monaural hearing. Perhaps this is not unreasonable inasmuch as the only class of binaural experiments that we have considered deals with a stimulus that is symmetrically presented to both ears. For information in the general case, the reader is directed to the work of Falmagne *et al.* (1979). These authors present a modern probabilistic treatment of binaural loudness summation within the framework of the forced-choice paradigm.

Given all of the complexities touched on above, and the certainty that there are many phenomena in the peripheral auditory system that we either do not yet understand or do not know how to properly describe, it is surprising indeed that the simple model outlined here performs so well. But it is also gratifying that it does.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation. We are grateful to D. M. Green, R. Hellman, R. D. Luce, and B. Scharf for valuable suggestions and for providing us with detailed experimental loudness data.

Bouman, M. A., Vos, J., and Walraven, P. (1963). "Fluctuation theory of luminance and chromaticity discrimination," J. Opt. Soc. Am. 53, 121–128.
Cross, D. V. (1973). "Sequential dependencies and regression in psychophysical judgements," Percept. Psychophys. 14, 547–552.
Egan, J. P., and Hake, H. W. (1950). "On the masking pattern of a simple auditory stimulus," J. Acoust. Soc. Am. 22, 622–630.
Falmagne, J. C., Iverson, G., and Marcovici, S. (1979). "Binaural 'loudness' summation: Probabilistic theory and data," Psychol. Rev. 86, 25–43.
Fletcher, H. F., and Munson, W. A. (1933). "Loudness, its definition, measurement and calculation," J. Acoust. Soc. Am. 5, 82–108.
Garner, W. R. (1948). "The loudness of repeated short tones," J. Acoust. Soc. Am. 20, 513–527.
Green, D. M., and Luce, R. D. (1974). "Variability of magnitude estimates: A timing analysis," Percept. Psychophys. 15, 291–300.
Hellman, R. P. (1976). "Growth of loudness at 1000 and 3000 Hz," J. Acoust. Soc. Am. 60, 672–679.
Hellman, W. S., and Hellman, R. P. (1975). "Relation of the loudness function to the intensity characteristic of the ear," J. Acoust. Soc. Am. 57, 188–192.
Hellman, R. P., and Zwislocki, J. (1961). "Some factors affecting the estimation of loudness," J. Acoust. Soc. Am. 33, 687–694.
Hellman, R. P., and Zwislocki, J. J. (1963). "Monaural loudness function at 1000 cps and interaural summation," J. Acoust. Soc. Am. 35, 856–865.
Hellman, R. P., and Zwislocki, J. J. (1968). "Loudness determination at low sound frequencies," J. Acoust. Soc. Am. 43, 60–64.
Jesteadt, W., Luce, R. D., and Green, D. M. (1977). "Sequential effects in judgements of loudness," J. Exp. Psychol.: Hum. Percept. Perform. 3, 92–104.
Lochner, J. P. A., and Burger, J. A. (1961). "Form of the loudness function in the presence of masking noise," J. Acoust. Soc. Am. 33, 1705–1707.
Lochner, J. P. A., and Burger, J. F. (1962). "Pure-tone loudness relations," J. Acoust. Soc. Am. 34, 576–581.
Marks, L. E. (1979). "A theory of loudness and loudness judgements," Psychol. Rev. 86, 256–285.
McGill, W. J. (1960). "The slope of the loudness function: A puzzle," in *Psychological Scaling: Theory and Application*, edited by H. Gulliksen and S. Messick (Wiley, New York), pp. 67–81; reprinted in *Sensation and Measurement—Papers in Honor of S. S. Stevens*, edited by H. R. Moskowitz, B. Scharf, and J. C. Stevens (Reidel, Dordrecht, 1974), pp. 295–307.
McGill, W. J., and Goldberg, J. P. (1968a). "Pure-tone intensity discrimination and energy detection," J. Acoust. Soc. Am. 44, 576–581.
McGill, W. J., and Goldberg, J. P. (1968b). "A study of the near-miss involving Weber's law and pure-tone intensity discrimination," Percept. Psychophys. 4, 105–109.
Richardson, L. F., and Ross, J. S. (1930). "Loudness and telephone current," J. Gen. Psychol. 3, 288–306.
Scharf, B. (1978). "Loudness," in *Handbook of Perception*, Vol. IV, *Hearing*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 187–242.
Scharf, B., and Stevens, J. C. (1961). "The form of the loudness function near threshold," in *Proceedings of the Third International Congress on Acoustics*, Vol. I, edited by L. Cremer (Elsevier, Amsterdam), pp. 80–82.
Schubert, E. D. (1978). "History of research on hearing," in *Handbook of Perception*, Vol. IV, *Hearing*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 59–63.
Siebert, W. M. (1968). "Stimulus transformations in the peripheral auditory system," in *Recognizing Patterns*, edited by P. A. Kolers and M. Eden (MIT, Cambridge, MA), pp. 104–133.
Stevens, J. C., and Guirao, M. (1964). "Individual loudness functions," J. Acoust. Soc. Am. 36, 2210–2213.
Stevens, S. S. (1955). "The measurement of loudness," J. Acoust. Soc. Am. 27, 815–829.
Stevens, S. S. (1970). "Neural events and the psychophysical law," Science 170, 1043–1050.
Teich, M. C., and Lachs, G. (1979). "A neural counting model incorporating refractoriness and spread of excitation. I. Application to intensity discrimination," J. Acoust. Soc. Am. 66, 1738–1749. This paper is denoted as I in the text.
van der Velden, H. A. (1949). "Quanteuse verschijnselen bij het zien," Nederlands Tijdschrift voor Natuur Kunde 15, 147–151.
Vannucci, G., and Teich, M. C. (1981). "Dead-time-modified photocount mean and variance for chaotic radiation," J. Opt. Soc. Am. 71, 164–170.
Ward, L. M. (1973). "Repeated magnitude estimations with a variable standard: Sequential effects and other properties," Percept. Psychophys. 14, 193–200.
Wever, E. G. (1949). *Theory of Hearing* (Wiley, New York).
Zwicker, E. (1974). "On a psychoacoustical equivalent of tuning curves," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer–Verlag, Heidelberg), pp. 132–141.
Zwislocki, J. (1969). "Temporal summation of loudness: An analysis," J. Acoust. Soc. Am. 46, 431–441.