Cochlear nonlinearity and second filter : possible mechanism and implications

Citation for published version (APA):

Duifhuis, H. (1976). Cochlear nonlinearity and second filter : possible mechanism and implications. *Journal of the Acoustical Society of America*, 59(2), 408-423. <https://doi.org/10.1121/1.380878>

DOI:

[10.1121/1.380878](https://doi.org/10.1121/1.380878)

Document status and date:

Published: 01/01/1976

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Cochlear nonlinearity and second filter: Possible mechanism and implications*

H. Duifhuis

Institute for Perception Research IPO, Den Dolech 2, Eindhoven, The Netherlands
(Received 17 April 1975; revised 22 October 1975)

We indicate that the directional sensitivity of the hair cell together with a directional distribution of frequency over the hair cells comprise a possible physiological basis for the second filter. Tuning disparity of the first and second filter denotes the difference in tuning frequency; at a given position x the tuning frequency of the first filter is α CF, of the second CF, with $\alpha > 1$. This accounts for the asymmetry in location of two-tone suppression areas. The compressive nonlinearity is described by a ν th law device with $\nu < 1$. We analyze implications of this model for two-tone suppression, sharpening, pure-tone masking, and combination tone generation. Basic features of these phenomena are described adequately. For combination tones the propagation problem needs further study. On the basis of a comparison of literature data and theoretical predictions we estimate $\alpha = 1.2$ and $\nu = 0.6$. Regarding accurate shape of the first and second filter, the discussed data provide means for a qualitative evaluation only. Possibilities for a quantitative analysis are indicated.

Subject Classification: [43]65.20, [43]65.35, [43]65.42, [43]65.26.

INTRODUCTION

Psychoacoustical frequency selectivity appeared to be significantly sharper than mechanical traveling wave resolution along the cochlear partition (Békésy, 1960). Therefore Békésy proposed a sharpening mechanism which would enhance the selectivity as the result of lateral inhibitory interactions. This fitted in nicely with contemporary findings of lateral inhibition in other receptors. However, later auditory work demonstrated a sharp tuning of primary auditory fibers (Kiang *et al.*, 1965). Combined with the subsequent finding of a large radial innervation of the inner hair cells by primary fibers (Spoendlin, 1972), this leaves neither necessity (Small, 1959; Vogten, 1974; Zwicker, 1974a) nor an obvious physiological basis for neural sharpening.

Meanwhile, new data have been obtained on frequency selectivity of the cochlear partition (Johnstone *et al.*, 1970; Rhode, 1971; Köhlöffel, 1972; Wilson and Johnstone, 1972). Because these data showed a much higher selectivity than Békésy had observed, initial speculations suggested that no additional sharpening would be necessary (Johnstone and Taylor, 1970). However, as Evans has pointed out repeatedly (1970, 1971, 1972a, 1972b), an evaluation of a presently larger body of data teaches us that a discrepancy still exists between mechanical (basilar membrane) and neural (primary fibers) tuning. The discrepancy is particularly significant for the low-frequency slope of the tuning curve. The discrepancy can be described in terms of a second filter, which must be operative between the first mechanical filter and auditory nerve fibers.

An additional need for a second filter comes from theories on combination tones and on two-tone suppression. Goldstein (1967, p. 688) considered a three-element series arrangement of a first linear (mechanical) filter, an essential nonlinearity either in basilar membrane mechanics or in the hair cell coupling, and a second filter ("frequency selectivity weighting function") a "conceptually useful phenomenological model ... to account for the properties of the cubic combination

tones." Pfeiffer (1970) modified an analog model for cochlear microphonics by Engebretson and Eldredge (1968) to the same general scheme, i.e., a nonlinear section between two linear bandpass filters to account for two-tone inhibition as observed in single auditory nerve fibers by, e.g., Sachs and Kiang (1968). Pfeiffer also remarked that the $2f_1 - f_2$ combination tone is the major cross-product term generated by the model. The basic remaining problem was to find a physiological basis for nonlinearity and second filter.

In Sec. I we discuss the general aspects of nonlinearity and second filter. Section II presents speculations about a possible physiological basis for nonlinearity and the second filter. In Sec. III the speculations are quantified and related to sharpening and two-tone suppression. Implications for pure-tone masking and for combination tones are discussed in Secs. IV and V. It is noteworthy that, although the analysis given in Secs. III-V is based on specific assumptions made in Sec. II, the results can be generalized with little effort. In other words, the physiological basis for the second filter which is presented in Sec. II is not crucial for most of the results derived. Very similar results are predicted by more abstract models, like Pfeiffer's (1970). It is hoped, however, that the proposed physiological basis for the second filter helps to provide insight into how the ear works.

I. GENERAL CONSIDERATIONS ABOUT NONLINEARITY AND SECOND FILTER

In the course of this paper we will assume that specifications and physiological basis for the first filter are sufficiently known. In this section we discuss some general features of the nonlinearity and of the second filter.

The nonlinear element is crucial for the generation of distortion products. Since no evidence exists which suggests the temporal buildup of combination tones to differ essentially from that of pure tones (Smooenburg, 1972), we take the nonlinearity to be time invariant

(static). Goldstein (1967) termed the nonlinearity *essential* because the relative amplitude of $2f_1 - f_2$ is almost independent of the stimulus level. That implies that in a power series expansion of the nonlinearity the linear term must be negligible. Next the question arises as to whether the nonlinearity is compressive or expansive. A first general consideration is that a compressive nonlinearity is certainly more helpful than an expansive one in accounting for the considerable dynamic range of the ear. Without loss of generality we may restrict ourselves to a consideration of the so-called ν th law device, defined by $y = \text{sgn}(x) \cdot |x|^\nu$ (see Fig. 1), because over a bounded interval any continuous function can be approximated by a series of ν th power terms.¹ The distinction between compressive and expansive nonlinearity is then equivalent to the question of whether $\nu < 1$ or $\nu > 1$, respectively. Smoorenburg (1972) concludes on the basis of amplitude behavior of $2f_1 - f_2$, and, independently, on the $2f_1 - f_2$ phase or polarity, that $\nu < 1$. He found that the polarity of $2f_1 - f_2$ is opposite to the polarity of the primaries, i.e. if the primaries are described by $\cos(2\pi f_1 t)$ and $\cos(2\pi f_2 t)$, then $2f_1 - f_2$ behaves as $\cos[2\pi(2f_1 - f_2)t + \pi]$ (see also Schroeder, 1969). This is in accordance with neurophysiological data by Goldstein and Kiang (1968) as shown by Goldstein (1972b, his Fig. 4; and 1972a, in a reprocessing of unit 442-7 data). In Pfeiffer's (1970) descriptive theory of two-tone suppression the combination of compressive nonlinearity and second filter is crucial. It appears therefore attractive to assume a single time-invariant compressive nonlinearity as comprising the basis of the above-mentioned nonlinear phenomena.

Potential problems are contained in data which suggest a linearlike relation between an auditory-nerve fiber's impulse response and its tuning curve (Goldstein *et al.*, 1971; de Boer, 1967, 1969, 1973). De Boer's data show striking similarity between the tuning curve and, within some 25 dB of the threshold at CF, the Fourier transform of the impulse response of the linear elements of the system.² This suggests that the tuning curve would mirror no nonlinearities. Goldstein *et al.* (1971) show that the average response delay ex-

FIG. 1. Characteristic of the full-wave ν th law device (odd-order type), for which $R = \text{sgn}(r) \cdot |r|^\nu$. For $\nu < 1$ the characteristic represents a compressive nonlinearity, as used in this study, for $\nu > 1$ it is an expansive nonlinearity.

pected from linear minimum phase filters with tuning curve frequency selectivity is compatible with the average response delay in PST histograms of click responses. Again, a description in terms of a linear filter appears satisfactory. Two comments are in order. For both experiments it is essential to know how sensitive the tuning curve is to nonlinearities. Regarding De Boer's experiment it is then relevant whether the measuring accuracy is sufficiently high to enable reliable detection of nonlinearities. Offhand it does not seem impossible that the nonlinearity effect is of the order of magnitude of the measuring accuracy. Moreover, de Boer (1969) actually notes that the tuning curve could be slightly narrower than the linear response. The second comment concerns the Goldstein *et al.* (1971) study. It should be noted that neural responses also reflect nonlinearities of the mechanical-to-electrical transducing mechanism. Because of the constant rate-increment detection criterion used for the tuning curve, it seems plausible that the transducing mechanism effect shows up more clearly in the click response than in the tuning curve. In principle thus, neural data are contaminated by at least two seemingly essential nonlinearities (cf. Pfeiffer *et al.*, 1974, and Smoorenburg's comment on that paper). In short, we conclude that the fact that tuning as observed in auditory-nerve fibers can be described linearly for some purpose does not imply that peripheral nonlinearities do not exist. However, from the extent to which a linear description is adequate, constraints may follow for the degree of nonlinearity.

Next we consider some properties of the second filter. These are most obviously pertinent to frequency selectivity and to two-tone suppression. With regard to frequency selectivity the following constraints apply: (1) Along the basilar membrane, the tuning frequency of the second filter must follow the tuning frequency of the first filter, but the tuning frequencies are not necessarily exactly equal. This condition might suggest a physical coupling between the first and second filter. (2) Comparing neural tuning curves to mechanical tuning curves (Evans, 1972), we conclude that the second filter must have a bandwidth of at most the order of magnitude of the bandwidth of the first filter and more likely a narrower one, in order to produce the required amount of sharpening. These two constraints are in line with the requirements for a proper description of two-tone suppression (Pfeiffer, 1970). It will be shown (Sec. III) that the asymmetry in the two-tone suppression above and below CF (Sachs and Kiang, 1968) is most readily described if one assumes that at a certain point at the basilar membrane the mechanical tuning frequency is slightly higher than the tuning frequency of the second filter. We will designate the assumed difference in tuning frequency by the term *tuning disparity*.

The sequential order of nonlinearity and second filter is determined by the need to predict sufficient two-tone suppression. A second tone can reduce the specific response to a stimulus tone, but at the same time it generates additional distortion (intermodulation) products, and in general the total averaged output of the

nonlinearity is not necessarily reduced. Suppression will show up if the second filter sufficiently reduces the distortion products. This can be achieved if its bandwidth is sufficiently narrow. Clearly, then, the second filter must follow the nonlinearity.

Several suggestions have been advanced about a possible physiological basis for nonlinearity and second filter. Interesting results are obtained when a nonlinearity in the damping of basilar membrane motion is assumed (Kim *et al.*, 1973; Hall, 1974). A current problem, however, is the apparent disagreement in observations of nonlinearity in basilar membrane motion (Rhode, 1971, vs Johnstone *et al.*, 1970, and Wilson and Johnstone, 1972). More directly comparable data are needed. A recent study by Robertson (1974) may shed new light on this matter.

The fact that, psychophysically as well as neurophysiologically, combination tones $2f_1 - f_2$ behave primarylike (Goldstein, 1972b) suggests that the low-frequency distortion products propagate to their proper place at the cochlear partition, i.e., to the hair cell tuned to the distortion product. This may not necessarily require a propagation along the basilar membrane, although that seems most likely (e.g., Hall, 1974). An alternative medium for propagation could be the fluid in the saccular sulcus. Relevant to this are cochlear microphonic data in particular from Dallos and co-workers (1969, 1974a, 1974b). In CM there is at best a very faint indication for propagation of $2f_1 - f_2$ (Dallos and Cheatham, 1974b). However, it is still an open question how well CM reflects basilar membrane or hair cell motion. Since CM supposedly mirrors intracellular hair cell potentials (Davis, 1965), it is presumably contaminated with (part of the) transducer nonlinearities, although not to the same extent as nerve fiber data are.

II. A PHYSIOLOGICAL BASIS FOR THE SECOND FILTER

A. Introduction

Hair cells, insofar as examined, are found to be directionally sensitive. Hitherto, this has been established mostly in lateral line and labyrinth organs of certain fishes, amphibians, and reptiles (Lowenstein and Wersäll, 1959; Flock, 1965a, 1965b; Wersäll *et al.*, 1965). Recent intracellular recordings from inner ear hair cells of the alligator lizard (Weiss *et al.*, 1974) are consistent with these findings. In general hair cells are excited when the cilia are deflected towards the kinocilium or centriole. When stimulating at an angle θ to this direction (Fig. 2), the sensitivity follows approximately $\cos \theta$ as long as $-\pi/2 \leq \theta \leq \pi/2$ and approaches zero, or is negative (inhibitory), in the other half-plane (Flock, 1965a, 1965b). Since all hair cells in the mammalian organ of Corti have a similar morphological orientation, which is determined by the centriole located at the side of the stria vascularis, it is likely that hair cells in the mammalian organ of Corti have a similar directional sensitivity. Morphological polarization of both inner and outer hair cells (Wersäll *et al.*, 1965) implies³ a sensitivity in the radial direction

(Fig. 2). The uniform radial polarization and unidirectional sensitivity are, of course, consistent with the fact that CM shows no frequency doubling (in contrast with lateral line where one finds a polarization into two opposite directions, see, e.g., Flock, 1965a).

A second relevant observation is that the cochlear traveling wave sets up a radial vibration component which is maximum at a point located basalwards of the maximum of the traveling wave envelope. Also, the radial component appears to be more sharply tuned (Fig. 3) than the traveling wave pattern (Békésy, 1953a; Khanna *et al.*, 1968).

Obviously, the combination of the directional distribution of vibrations as a function of place, or, at a fixed place as a function of frequency, constitutes in combination with the directional sensitivity, a frequency weighting or filter (see Fig. 4). We assume that it constitutes the second filter. Assuming that the directional sensitivity behaves as $\cos \theta$, the selectivity of the second filter is determined by the directional distribution of vibrations over frequency at the hair cell base (cf. also Tonndorf, 1970).

B. Outline of the theory

The above considerations inspired the following more specific, tentative assumptions concerning the second filter:

(1) A characteristic frequency $CF(x)$ is assigned to the hair cell located at x mm from the base. $CF(x)$ defines the frequency that stimulates the hair cell in its most sensitive direction. $CF(x)$ may be interpreted as the tuning frequency of the second filter at x .

(2) Frequencies off CF stimulate the hair cell at an angle θ with the sensitivity axis and are less effective by a factor of $\cos \theta$ (Flock, 1965a, 1965b). We assume that θ varies monotonically from $\theta = -\pi/2$ for $f = 0$, through $\theta = 0$ at $f = CF$, to $\theta = \pi/2$ for $f \rightarrow \infty$ (Fig. 4). (An interesting variant applies if one assumes that $\theta \rightarrow -\pi/2 + \epsilon$ for $f \rightarrow 0$, with ϵ a small positive number. We will

FIG. 2. (a) Morphological polarization of inner and outer hair cells (IHC and OHC) in the mammalian organ of Corti. Dots denote the cilia (marked ci), circles the centrioles (marked ce). (After Wersäll *et al.*, 1965.) (b) Polar diagram representing the $\cos \theta$ directional sensitivity of the hair cell (HC).

FIG. 3. (a) Directional distribution of inner ear vibrations—radial, vertical, and longitudinal—as observed by Békésy in a top view of the cochlear partition. (After Békésy, 1953a). (b) Relation between excitation magnitude of the traveling wave along the cochlear partition and its radial component, for a 75 Hz stimulus, as obtained for a theoretical model study. (After Khanna *et al.*, 1968).

discuss the variant in Sec. III.) A complex stimulus yields the linear vector sum $r(t)$ of the contributions of its components (see Fig. 8).

(3) The tuning frequency of the traveling wave envelope at the hair cell at x mm from the base is $\alpha CF(x)$, with the tuning disparity factor $\alpha > 1$ (cf. Békésy, 1953a; Khanna *et al.*, 1968). In general, $\alpha = \alpha(x)$ will be a slowly varying function of x . For most of our purposes we will assume that α is a constant. The fre-

FIG. 4. Schematic representation of first filters (upper panel), directional distribution of 'hair cell vibration' (middle panels), and second filters resulting from directional sensitivity (lower panel) in both excitation (as a function of location at the cochlear partition x for fixed frequency f), and response (as a function of f at a fixed x).

FIG. 5. Block diagram of our theoretical model. Each channel (one of which is depicted) consists of a first filter $H_1(\omega; x)$; directional distribution $\theta(\omega; x)$; the time-invariant compressive nonlinearity $R = \text{sgn } r |r|^\nu$; and the directional sensitivity $\cos \theta$. Directional distribution and directional sensitivity together comprise the second filter $H_2(\omega; x)$.

quency $\alpha CF(x)$ is the tuning frequency of the first filter at x . In principle, this filter is constituted by the mechanical selectivity of the basilar membrane, plus possibly additional selectivity that might be introduced in the transformation from basilar membrane vibration to vibration at the hair cell base. In other words, we cannot (yet) claim that the first filter exclusively reflects basilar membrane selectivity. For the purpose of this paper we assume that the first filter is linear, with amplitude characteristic $H_1(f, x)$.

(4) Concerning nonlinearity: The resultant mechanical excitation at the hair cell, $r(t)$, undergoes a non-linear compression which is uniform in all directions. This means that the compression does affect the magnitude of the resultant excitation, but not the angle θ . The nonlinearity is sandwiched between the directional distribution of frequency and the hair cell's directional sensitivity. This means that it operates before the second filter becomes effective, in accordance with the requirement in Sec. I. One may think of the nonlinearity in terms of a nonlinear load to which the linear resultant stimulus at the hair cell is subjected. We will assume that the compressive nonlinearity can be described adequately with a ν th law device (Fig. 1), so that the compressed stimulus at the hair cell, $R(t)$, is given by

$$R(t) = \text{sgn}[r(t)] |r(t)|^\nu. \quad (1)$$

The above assumptions are indicated in the block diagram of Fig. 5. For comparison, the traditional nonlinearity and second filter block diagram (Pfeiffer, 1970) is presented in Fig. 6. In the course of this paper we will discuss some of the differences between the two models. At this point we remark that the major difference is that we have proposed a physiological basis for our model.

The properties outlined above meet the requirements put forward in Sec. I. Hence this theory will predict sharpening, two-tone suppression, and generation of combination tones. In Secs. III–V we will substantiate and quantify this statement.

FIG. 6. Conventional model for second filter and compressive nonlinearity (one channel depicted). The compressive nonlinearity $R = \text{sgn } r |r|^\nu$ is located between the two filters $H_1(\omega; x)$ and $H_2(\omega; x)$.

III. SHARPENING AND TWO-TONE SUPPRESSION

This section deals with sharpening caused by compressive nonlinearity and second filter, and with several aspects of two-tone suppression (rate and synchronization data).

The notation is simplified somewhat by writing r^ν for $\text{sgn}(r) |r|^\nu$ and by using the angular frequency ω instead of $2\pi f$. [Of course, $H_1(\omega; x) \equiv H_1(f; x)$, etc.]

Because we present most of our theoretical predictions in the form of average stimulating waveforms, we neglect the phase shifts produced by the filters.

A. Sharpening

We consider the effect of a stimulus tone $s(t) = A \cos \omega t$ on the hair cell CF(x). The mechanical excitation will be

$$r(t) = A H_1 \cos(\omega t), \quad (2)$$

which excites the hair cell at an angle θ to its sensitivity axis. The compressed waveform $R(t)$ will be $R(t) = r^\nu(t)$ and the stimulating waveform, i.e., the component of $R(t)$ in the direction of the sensitivity axis, $R_f(t)$, is

$$R_f(t) = R(t) \cos \theta = r^\nu(t) H_2. \quad (3)$$

The time average of the absolute value of the stimulating waveform, $E = \langle |R_f(t)| \rangle_{\text{av}}$, seems a reasonable measure for the neural effect of the stimulus. We will assume that a constant E corresponds to a constant firing rate in primary auditory-nerve fibers that innervate the hair cell CF(x). For the above tonal stimulus, E takes the form

$$E \propto (A H_1)^\nu H_2. \quad (4)$$

[The proportionality constant can be shown to equal

FIG. 7. Schematic representation of the relation between tuning curve (indicated by $H_1 H_2^{1/\nu}$) and the shape of first and second filter H_1 and H_2 , which are indicated in the figure, for $\nu = 0.6$. Note that the tuning curve is sharper than the product of H_1 and H_2 . This is due to the nonlinear compression ($\nu < 1$).

FIG. 8. Vectorial summation of the first-filter responses, $r_1(t)$ and $r_2(t)$, to the components of a two-tone stimulus. Tone 1 is at CF, tone 2 is off CF. The resultant $r(t)$ is compressed to $R(t)$, which yields the effective $R_f(t)$ at the direction of sensitivity. Note that θ_r depends on time because $f_1 \neq f_2$.

$(1/\pi)B(\frac{1}{2}, \frac{1}{2} + \frac{1}{2}\nu)$, where B is the beta function]. A constant value for E is obtained when the right-hand part of Eq. (4) is constant, or when

$$A = A(\omega) \propto 1/(H_1 H_2^{1/\nu}). \quad (5)$$

Since the tuning curve is also determined by a constant rate increment criterion, Eq. (5) gives our prediction for the shape of the tuning curve. Because $\nu < 1$, the tuning curve is sharper than just the product of H_1 and H_2 . Figure 7 shows examples of Eq. (5). We remark that the relatively flat low-frequency tail of the tuning curve can be predicted by assuming that for $\omega \rightarrow 0$ the angle $\theta(\omega)$ approaches $-\pi/2 + \epsilon$, so that $H_2 = \cos \theta(\omega)$ approaches a constant value greater than 0. For very low frequencies the tuning curve then parallels the first filter.

B. Constant tone at CF

First we consider the effect of a second tone on the average response to a constant tone at CF, as studied by Sachs and Kiang (1968) and Sachs (1969). The constant tone was termed the CTCF tone by these authors.

Let the CTCF stimulus be $s_1(t) = A_1 \cos \omega_1 t$ and the second variable tone $s_2(t) = A_2 \cos \omega_2 t$. The resulting mechanical excitation consists of the components (Fig. 8)

$$r_1(t) = A_1 H_{11} \cos \omega_1 t, \quad \text{at } \theta_1 = 0,$$

and

$$r_2(t) = A_2 H_{12} \cos \omega_2 t, \quad \text{at } \theta_2 = \theta_2(\omega_2) \quad (6)$$

(the notation H_{ij} is used to designate the effect of filter i on stimulus j). From Fig. 8 we see that the stimulating waveform can be written as

$$R_f(t) = r_f(t) |r(t)|^{\nu-1}, \quad (7)$$

where $r(t)$ is the vector sum of $r_1(t)$ and $r_2(t)$, and $r_f(t)$ the component of $r(t)$ in the sensitivity direction. If the variable tone approaches the CTCF in frequency, then the resultant mechanical excitation occurs primarily in the sensitivity direction, so that $r_f(t) \approx r(t)$, and $R_f(t) \approx |r_1(t) + r_2(t)|^\nu$. In this case the average stimulating waveform, $E(s_1 + s_2)$, increases monotonically when increasing the amplitude of the variable tone (see Fig. 9, the $f_1 = f_2$ curve). For small values of A_2 the increase is marginal; for large values of A_2 the average stimulating waveform follows A_2^ν , i.e., the second tone response. The transition occurs at A_{20} , where $A_2 = A_1$.

FIG. 9. Average response of the model of Fig. 5 to a two-tone stimulus. Tone 1, the CTCF, is fixed. The amplitude of tone 2 is the independent variable. Suppression occurs in the shaded area. The right-hand part of the figure shows (schematized) vector diagrams for situations 1-3, indicated in the main figure. At point 4 the average response asymptotically approaches the line at $\arctan \nu$, which represents the response to tone 2 alone. Parameters: $\theta_1 = 0$; $\theta_2 \rightarrow \pi/2$; $A_{20} = A_1$; $A_{21} = A_1/H_{11}/H_{12}$; $A_{22} = A_1 H_{11} H_{21} / (H_{12} H_{22})$; $A_{23} = A_1 H_{11} H_{21}^{1/\nu} / (H_{12} H_{22}^{1/\nu}) = A_1 H_{11} / (H_{12} H_{22}^{1/\nu})$.

The monotonic increase implies that we find no suppression if the variable tone approaches CF.

If the variable tone is significantly off CF, so that $\cos \theta_2 \approx 0$, then Eq. (7) modifies to

$$R_f(t) = r_f(t) [\gamma_1^2(t) + \gamma_2^2(t)]^{(1-\nu)/2}. \quad (8)$$

Because the exponent of the bracket term is negative, an increase of the variable tone $r_2(t)$ will reduce $R_f(t)$ (see Fig. 9, inset 3). The suppression becomes significant when the response of the first filter to the variable tone exceeds the CTCF response ($A_2 H_{12} > A_1 H_{11}$). This is indicated with the transition value A_{21} , where $A_2 = A_1 H_{11} / H_{12}$. For very large values of the amplitude of the variable tone the contribution of this tone in the sensitivity direction, $r_2(t) \cos \theta_2$, will no longer be negligible. Ultimately the second tone will dominate the average stimulating waveform (Fig. 9, point 4). This starts at A_{22} , where $A_2 = A_1 H_{11} / (H_{12} H_{22})$. At A_{23} the activity in response to the variable tone equals the CTCF response, and we leave the suppression area. A_{23} is defined by $A_2 = A_1 H_{11} / (H_{12} H_{22}^{1/\nu})$.

FIG. 10. Theoretical two-tone suppression areas (shaded). The suppression areas are bounded by first filter and tuning curve. Parameter values for first and second filter are indicated in the inset. Additional values: $\nu = 0.8$; $\alpha = 1.4$. Suppression conditions: CTCF is 0 dB at 1 kHz; suppression in shaded areas $> 20\%$.

FIG. 11. Theoretical two-tone suppression as in Fig. 10, but with the fixed tone at α CF at 60 dB. The isorate criterion is the same as in Fig. 10. Note that the suppression areas have virtually disappeared, whereas the isorate contour ("tuning curve") is shifted upward.

Thus, the lower boundaries of the suppression areas, defined by A_{21} values, follow the inverted first filter, and the upper boundaries, A_{23} , parallel the tuning curve.⁴

Figure 10 shows an example of the two-tone suppression areas for the schematized first and second filters specified in the inset. In the shaded areas the suppression is $\geq 20\%$. Other parameter values are $\nu = 0.8$, and $\alpha = 1.4$. The latter values are discussed in Sec. VI. Because of (1) the asymmetry of the first filter (the high-frequency slope is much steeper than the low-frequency slope); (2) the proper choice of α ; and (3) the proper relative location of upper and lower boundaries of the suppression areas, one obtains the two suppression areas asymmetrically around CF, in accordance with the neurophysiological data (Sachs and Kiang, 1968; Arthur *et al.*, 1971). The relative location of the upper and lower boundaries is not a free parameter but follows directly from the theory. For this example one can show that at CF the upper boundary is approximately 10 dB below the lower boundary. This relative location excludes, of course, suppression around CF.⁵

C. Constant tone at α CF

Next we consider two-tone suppression where the fixed tone is off CF. As an example we will place it at α CF (CT α CF). This value is of particular interest, since it bears some relation to broad-band responses of the first filter.

Figure 11 shows results for this situation. The parameter values are the same as in Fig. 10. Apparently only one significant suppression area is left, viz., above CT α CF. The major effect of two-tone suppression is seen in the upward shift, due to the CT α CF, of the threshold isorate contour ("tuning curve"), which determines the upper boundary of the suppression area. Figure 10 shows that a tone at α CF is a very effective suppressor. (We have examined the CT α CF case only with the use of a computer simulation of the model.)

The results shown in Fig. 11 may explain why Evans (1974) observed little or no suppression in the form of suppression areas in an experiment where he used

FIG. 12. Synchronization coefficients as a function of stimulus level. Data from Rose *et al.* (1974) (AVCN unit 71-213-2) are given in (a), predictions from the theory which is extended with a saturation mechanism, are given in panel (b). The response to tone 1 alone is denoted by σ_ω ; σ_{ω_1} and σ_{ω_2} are elements from the two-tone response. Parameters: (a) $f_1 = 0.8$ kHz; $f_2 = 1.2$ kHz; CF = 1.2 kHz; amplitude of tone 2 is 30 dB SPL; ordinate values: $20^{10} \log A_{1j}$, respectively; $20^{10} \log A_{1j}$ as defined by Rose *et al.*; (b) $\cos \theta_1 = 0.5$; $\cos \theta_2 = 0$; $C = \langle p(t)^2 \rangle_{av}$ in response to tone 1 at amplitude $A_1 H_{11} = 1$; $\nu = 0.6$; scaling factors: ordinate $2fr_{\max} = 40$ dB, abscissa $A_2 = 1 = 37.5$ dB.

white noise for the fixed stimulus. The response of the first filter (at CF) to the noise may be considered an amplitude and phase modulated CT α CF stimulus. Evans's observation that the isoresponse contours shift upwards with increase in noise level, a phenomenon also described by Kiang *et al.* (1965, Fig. 9.5), reflects—in our opinion—a clear effect of two-tone suppression.

D. Damping power

In a recent study Rose *et al.* (1974)⁸ give a description of the two-tone suppression effect, introducing the concept of attenuating or *damping power*. The damping power is defined by the factor by which the primary synchronization coefficient in the neural response to a low-frequency tone lags the primary amplitude, and it is expressed in decibels. At threshold the damping power is zero by definition, at saturation it increases linearly with tone level. The synchronization coefficients σ can be defined as the amplitudes in the Fourier spectrum of the period histogram in response to the stimulus. The primary σ 's occur at the stimulus frequencies. For a two-tone stimulus the data are described by the algorithm that the largest of the two damping powers is effective on both components. Figure 12(a) shows data for their unit 71-213-2 with the CTCF at 1.2 kHz and a variable tone at 0.8 kHz. At small amplitudes of the variable tone the response to the CTCF is constant. When the variable tone saturates, its damping power increases. Eventually it dominates, and the CTCF response is suppressed.

An important point made by Rose *et al.* (1974) is that not only firing rate data but also synchronization data are pertinent to two-tone suppression (see also Hind *et al.*, 1970). However, we are in doubt as to what extent the damping power concept—admittedly intro-

duced as a descriptive concept only—helps to understand the suppression phenomenon. The point is that the damping power primarily describes saturation, which is thought to originate somewhere at the hair cell—afferent synaps level (Schroeder and Hall, 1974). As outlined in this study, we believe two-tone suppression occurs before that level. If our theory is extended by an element describing neural saturation, then the data from Rose *et al.* (1974) are readily accounted for. The following saturation function, proposed by Siebert (1972), gives a useful description:

$$fr(t) = \frac{2fr_{\max} p(t)^2}{C + \langle p(t)^2 \rangle_{av}}, \quad (9)$$

where $p(t)$ is the stimulating waveform; $p(t)^2$ the quadratic half-wave rectified stimulating waveform; $fr(t)$ the firing rate function (for a nonhomogeneous Poisson firing process); $fr_{\max}$ the saturation rate; and C a constant. [A proper choice of the averaging time window allows for the coverage of time-dependent (adaptation) effects. At present this is of secondary interest since we restrict ourselves to stationary stimuli.] Applying this saturation function to our model amounts to substituting $p(t) = R_{\#}(t)$.

Figure 12(b) gives the results from this substitution, thus predicting the data of Fig. 12(a) from Rose *et al.* (1974). We consider the agreement satisfactory.

IV. PURE-TONE MASKING

A. Introduction

In pure-tone masking we have in general four variables: probe threshold level L_P , probe frequency f_P , masker level L_M , and masker frequency f_M . [We neglect temporal effects, an obvious oversimplification, which is justified for long-duration (~ 1 sec) simultaneous masking.] The independent variable is usually a frequency, the dependent variable a level, the other two variables being fixed. In the classical experiment by Wegel and Lane (1924), L_M and f_M were kept constant. We refer to this as an iso- $L_M f_M$ experiment. In later experiments the other conditions have also been investigated (iso- $L_P f_P$: Small, 1959; Vogten, 1972, 1974a; Zwicker, 1974a; iso- $L_M f_P$: Vogten, 1972, 1974a; Verschuure *et al.*, 1974; iso- $L_P f_M$: Zwicker, 1974a; Verschuure *et al.*, 1974). Because of nonlinearities in the auditory system, the masking curves obtainable from the different types of experiments are not each other's linear transforms. This can result in differences in slopes of the masking curves. If we assume that the probe is detected at x_P , then the iso- $L_P f_P$ cases are most closely related to our treatment of the tuning curve (Sec. III A), which also considers the response at a particular place x . We will therefore, deal mainly with iso- $L_P f_P$ data.

From our theoretical point of view, the non-simultaneous-masking data are simpler to interpret than simultaneous-masking data. This is due to the nonlinearity of the system. The nonlinear response to the sum of two signals is in general more complex than the sum of the responses to the two signals. Therefore, we will start with the more recent pulsation threshold

method, followed by forward masking, and finally deal with conventional simultaneous pure-tone masking. Throughout this section we use the indices P and M for probe and masker, instead of the numerical indices used in the other sections. Furthermore, we use A for (linear) amplitude, and L for the logarithmic representation thereof (level). (The choice between A and L is largely arbitrary.) We will also write f instead of $\omega (= 2\pi f)$.

B. Pulsation threshold

A plausible explanation for the continuity effect, on which the pulsation threshold is based, is that the activity in the probe channel remains constant (Houtgast, 1974a). Let us assume that this implies a constant firing rate in the primary neurons tuned to the probe, and thus a constant E at x_P . The iso- $L_P f_P$ experiment implies that E in response to the masker equals E in response to the probe, which is constant. This is similar to the tuning-curve criterion [Eqs. (4) and (5)]. The continuity effect theory thus implies that the iso- $L_P f_P$ pulsation threshold contour must parallel the tuning curve, that is

$$A_M H_{1M} H_{2M}^{1/\nu} = \text{const} \quad (\text{at } x_P). \quad (10)$$

Houtgast's (1974a, Fig. 4.2) data are positively suggestive in this respect.

In case of iso- $L_M f_M$ masking, x_P is not fixed but becomes the variable. This requires knowledge of excitation patterns over x for fixed frequencies, which are more difficult to measure. Henceforth, we will therefore restrict ourselves to iso- $L_P f_P$ masking.

C. Forward masking

In this subsection we make a brief excursion to some aspects of temporal masking, because it seems tempting to relate forward masking data to pulsation threshold data. We will stress some difficulties which obscure this relationship.

If the forward masking pattern, i.e., probe threshold as a function of probe frequency, might be assumed to give an adequate linear map of the perstimulatory excitation pattern, then forward masking data expectedly would match pulsation threshold data in shape. Two assumptions underlie the condition specified above. The first is that the with-time-decaying forward masking is due to the recovering perstimulatory adaptation. The second assumption is that adaptation reflects the excitation pattern linearly. This interpretation meets some objections. We have remarked elsewhere (Duifhuis, 1973) that forward masking, expressed in decibel threshold shift, i.e., after a logarithmic amplitude transformation, decreases exponentially with a time constant of about 75 msec. This implies that the half-power bandwidth of the masking pattern during forward masking increases with increasing time interval. If at all, then a linear relation between excitation and forward masking can be expected for very brief intervals only. An additional complication follows from the inference that for such intervals (< 20 msec) forward masking contains a significant transient masking com-

ponent (Duifhuis, 1973) which directly reflects the decaying excitation. Furthermore, there is evidence that the net adaptation follows the square root of the excitation rather than the excitation itself.

In short, although forward masking is not contaminated by nonlinear interaction of probe and masker, it is more complicated than the pulsation threshold because of effects in the time domain. An adequate descriptive theory of these temporal effects is necessary for a quantitative comparison of forward masking and pulsation threshold data. We exclude forward masking from further discussion in this paper.

D. Simultaneous masking

We approach simultaneous masking from two points of view, the distinction between which is related to the classical difference between "place" theory and "periodicity" theory. We will show that the two approaches lead to the same predictions, and hence that simultaneous masking data provide no tool for a decision.

1. Generalized place theory

We consider again the iso- $L_P f_P$ case. Obviously the probe can be detected only in the channels responding to the probe. It is a convenient simplification to consider only the channel tuned to the probe, i.e., at x_P , as the proper representative of the responding set of channels. We refer to this simplification as the single channel hypothesis (see Sec. IV E). The probe will be detected if the average stimulating waveform at x_P in response to probe plus masker significantly exceeds the response to masker alone.

We can analyze this problem using Fig. 9, which we have replotted as Fig. 13 using the notation of this section. The heavy dashed lines represent the response to masker alone, the heavy full lines give the response to probe plus masker, as indicated. We observe that the response to probe plus masker exceeds the masker alone response for masker amplitudes below A_{M0} for $f_M = f_P$ or below A_{M2} for $f_M \neq f_P$. We define these transi-

FIG. 13. Average response of the model of Fig. 5 to probe plus masker, $E\{P+M\}$, and masker alone, $E\{M\}$, for two masker frequencies. The probe threshold occurs at the point where the difference between $E\{M\}$ and $E\{P+M\}$ becomes negligible: A_{M2} for f_M off f_P , and A_{M0} for $f_M = f_P$ (iso- $L_P f_P$ stimulus).

FIG. 14. Theoretical iso- $L_P f_P$ masking curves. Curve 1: simultaneous masking; curves 2 and 3: nonsimultaneous masking (pulsation threshold) with $\nu=0.8$ and $\nu=0.6$, respectively. First and second filter as in Fig. 7.

tion points as masking criterions. This can be summarized as $A_M H_{1M} H_{2M} = \text{constant}$ (cf. Sec. III B), so that the iso- $L_P f_P$ curve is determined by the linear combination of the first and second filter at x_P . Hence, the simultaneous masking curve is broader than the nonsimultaneous masking curve, where the second filter occurs with the power $1/\nu$ (Fig. 14). This theoretical result agrees at least qualitatively with psychophysical data from Houtgast (1974a), who also found that pulsation threshold and forward masking curves are more sharply tuned than simultaneous masking curves. If the compressive nonlinearity can be adequately described by a single ν th law device ($\nu = \text{constant}$) and if ν can be determined independently, then it might be possible to separate the first and second filter when confronting simultaneous with nonsimultaneous masking data. However, we cannot expect a high degree of accuracy of such results since the differences may be "second-order" effects (cf. Fig. 14).

It is interesting to note that for values of A_M between A_{M1} and A_{M2} , the value of $E(P+M)$ decreases with increasing A_M . This implies that the internal probe-to-masker ratio decreases not only because of the increase of the denominator, but also because of a decrease of the numerator. Hence, the perceived probe level will decrease faster than suggested by the external probe-to-masker ratio (i.e., in the stimulus). In view of the results of Sec. III this can be readily interpreted as a suppression phenomenon, which occurs only if $f_M \neq f_P$. This prediction is again in agreement with psychophysical observations as reported, e.g., by Scharf (1964) and Houtgast (1974a, 1974b).

We remark that with the use of the pulsation threshold it seems possible to determine a rather direct correlate of Figs. 9 and 13. The ratio of the A_M values at A_{M2} and A_{M1} , $A_{M2}/A_{M1} = 1/H_{2M}$, then provide a direct measure of H_2 , which can be determined as a function of f_M . The determination of ν , however, presents us with a problem. Since the pulsation tone is subjected to the same nonlinearity, the pulsation amplitude will be proportional to $E^{1/\nu}$. Therefore the slopes of $\nu - 1$ and ν modify to $1 - 1/\nu$ and 1 , and ν can be estimated only from the slope in the interval bounded by $A_M = A_{M1}$ and $A_M = A_{M2}$. Houtgast (1974a, Chap. 7; 1974b) presents data for a 1-kHz probe tone in a broadband noise

masker. These data are not in disagreement with $\nu \approx 0.6$. The noise masker is not suitable for scanning the second filter H_2 as a function of f .

In this section we have used a simple threshold criterion. It would, of course, be possible to define more advanced criteria, but only at the cost of further specifications and it would presumably lead to results differing at most only in degree.

2. Generalized periodicity theory

The alternative to detecting and identifying a supra-threshold probe tone on the basis of local rate information (place theory), is to use the temporal information which is characteristic for the probe tone. Information about the stimulus waveform is to some extent preserved in neural action potentials (see Rose *et al.*, 1967, for data; Siebert, 1972, and Duifhuis, 1972, for theoretical descriptions). The neural temporal information is conveniently described by synchronization coefficients σ , as defined in Sec. III D. Detection criterions can be based either on absolute values, or on ratios of σ . We will use a ratio criterion. Consideration of the ratio of synchronization coefficients makes us less sensitive to the decrease of synchronization with increasing frequency. The question arises as to how the ratio of the synchronization coefficients of probe and masker $\rho_{PM} = \sigma_{f_P}/\sigma_{f_M}$, depends on place x . It is possible to show that

$$\rho_{PM}(x) \approx \frac{A_P H_{1P}(x) H_{2P}(x)}{A_M H_{1M}(x) H_{2M}(x)}, \quad (11)$$

or that $\rho_{PM}(x)$ depends on the ratio of "excitations" produced by probe and masker. Figure 15 shows that if the logarithmic excitation patterns are parallel, then the ratio ρ_{PM} is maximum at x_P . If also, as depicted in Fig. 15, $x_P > x_M$, or $f_P < f_M$, then ρ_{PM} is constant and equal to the maximum value for all $x \geq x_P$, and vice versa. This implies that, as long as the masker excitation has slopes not steeper than the probe excitation (cf. Sec. IV E), $\rho_{PM}(x)$ is maximum at $x = x_P$. Integra-

FIG. 15. Probe and masker "excitation" AH_1H_2 as a function of x . For the depicted situation we have $f_M > f_P$, and the ratio k (Eq. 11) of probe and masker excitation is maximum at and beyond $x = x_P$. In the opposite case, i.e., $f_M < f_P$ and $x_M = x_P$, k is maximum at and below $x = x_P$.

tion over x , or summation over a number of fibers, then does not change the expected value of ρ_{PM} , it can only reduce the variance of ρ_{PM} . The above considerations justify the simplification of considering the synchronization ratio ρ_{PM} in the channel tuned to the probe (at x_p). If, as was suggested, the probe threshold is specified by $\rho_{PM} = \text{constant}$, then from Eq. (11) it follows directly that for iso- $L_P f_P$ masking $A_M H_{1M} H_{2M} = \text{constant}$ at x_p . This is the same result as obtained with the generalized place theory. In other words, the generalized place theory and the generalized periodicity theory worked out above predict the same shapes for the iso- $L_P f_P$ masking curve.

E. Discussion of the single-channel hypothesis

It should be noted that the results in this section apply only in the situation where "internal noise" can be neglected with respect to the probe plus masker activity (but note the following paragraph). Modifications have to be made when probe and masker level approach threshold, in particular for simultaneous masking (see Vogten 1974b).

The simplification of considering only the one probe channel as being relevant for probe detection in iso- $L_P f_P$ masking should be regarded with some caution. If probe and masker excitation patterns are represented adequately by Fig. 15, and if masker slopes are not steeper than probe slopes, then the simplification is justified as follows. Optimum probe detection would require a matched weighting function across the channels. This gives maximum weight to the channel in which signal-to-noise ratio (internal + external noise) is maximum (cf. De Boer and Bos, 1962, and Siebert, 1968). Taking more channels into account has little or no effect on the expected value of probe response and it has a second-order effect on probe detection (cf. Duifhuis, 1973). However, the condition that masker slopes are not steeper than probe slopes appears to be in disaccordance with our assumption of tuning disparity. Tuning disparity predicts an increase in the high-frequency slope at αf_P (cf. Fig. 14), although actually the increase will be more gradual than schematically depicted. This implies that for $f_P < f_M$ the maximum signal-to-(external)-noise ratio occurs at αf_P . If the probe is well above its threshold-in-quiet level, then the internal signal-to-noise ratio may also be maximum at αf_P . If probe detection were indeed determined by activity in the channel at αf_P , then the high-frequency part of the iso- $L_P f_P$ masking curve would become steeper. In the example given in Fig. 14 the less steep slopes between f_P and αf_P would disappear and the steeper slopes indicated for $f > \alpha f_P$ would extend down to f_P . In other words, taking the tuning disparity into account, an optimum detection criterion would predict a leftward shift of the high-frequency slope of the iso- $L_P f_P$ masking curve with increase in probe level. However, because the high-frequency slopes are very steep anyway, it remains questionable whether these effects are significant. Thus, the single probe channel approach appears to be justified for low probe levels and it will provide a reasonable first approximation at moderate and higher probe levels.

V. COMBINATION TONES

A. Introduction

It has been understood for some time that aural combination tones of the type $nf_1 - (n-1)f_2$, i.e., the odd-order combination tones, reflect an essential cochlear nonlinearity (Goldstein, 1967; Smoorenburg, 1972; Hall, 1972; see also Sec. I). Smoorenburg (1972, 1974) has examined predictions of a ν th law device—the nonlinearity introduced in Sec. I—for combination tones. We will present and extend some of his results.

The ν th law device is chosen for the compressive nonlinearity for the convenience of analysis. Other compressive nonlinearities can be approximated by polynomials or series of ν th law terms.¹ We first present some general remarks about even- and odd-order nonlinearities. In Secs. VB and VC we analyze certain amplitude and phase properties of odd-order combination tones.

The ν th law half-wave rectifiers (Fig. 16a) have received considerable attention in the literature (see Feuerstein, 1957). The response of the rectifier to a two-tone stimulus of frequencies f_1 and f_2 is formulated in terms of a double Fourier series, which contains a dc component, components at the stimulus frequencies and their harmonics, and intermodulation products, or in general components at frequencies $mf_1 + nf_2$. Odd-order products obtain for $m+n$ is odd, even-order for $m+n$ is even. It is straightforward to show that the responses of full-wave odd and even rectifiers [Figs. 16(b) and 16(c)] contain only odd or only even products, respectively (at twice the half-wave response amplitudes). Thus, the characteristic proposed in Fig. 1 produces only odd-order combination tones. Any modification in the symmetry of the rectifier characteristic can be accounted for by a decomposition into the sum of an odd-order and an even-order rectifier. This means that the generation of difference tones ($f_2 - f_1$) can be predicted by introducing asymmetry in the nonlinear characteristic of Fig. 1. Since the analysis is not fundamentally different, we restrict ourselves to the treatment of odd-order distortion products.

B. Amplitudes of odd-order combination tones

If the two component frequencies f_1 and f_2 are contiguous, so that $\theta_2 \approx \theta_1 = 0$ at x_{f_2} , then the resultant excitation becomes $R_\nu(t) = [r_1(t) + r_2(t)]^\nu$ (cf. Sec. III B and Fig. 8). This results in a situation which is equal to the one treated by Pfeiffer (1970) and Smoorenburg (1972,

FIG. 16. Characteristics of half-wave rectifier (a), full-wave odd-order rectifier (b), and full-wave even-order rectifier (c) with compressive nonlinearity.

1974). It can be shown that for this case as well as for the case $\theta_2 \neq 0$ (primary frequencies further apart) the following asymptotic behavior follows for the amplitude of the $2f_1 - f_2$ combination tone

$$\begin{aligned} R_{2f_1-f_2} &\propto r_1^2 r_2^{\nu-2} & \text{if } r_2 \gg r_1 \\ &\propto r_1^{\nu-1} r_2 & \text{if } r_2 < r_1, \end{aligned} \quad (13)$$

where r_1 and r_2 are the amplitudes of $r_1(t)$ and $r_2(t)$, respectively [cf. Eq. (6)]. In Fig. 17 we have plotted the $2f_1 - f_2$ amplitude as a function of the primary amplitudes (full lines). Significant deviations from the asymptotes are restricted to an approximate 20 dB interval around $r_1 = r_2$. [The theory of ν th law rectifiers gives an exact analytical expression for $R_{2f_1-f_2}$ over the entire amplitude range. A noteworthy aspect is that in general the asymptotes do not intersect exactly at $r_1 = r_2$.]

If the primary frequencies f_1 and f_2 are related harmonically, then higher-order intermodulation products of the same frequency will be generated. If, e.g., $f_2/f_1 = \frac{4}{3}$, then we have $2f_1 - f_2 = (2+4k)f_1 - (1+3k)f_2$. Since the higher-order terms become decreasingly significant, this is only a second-order effect.

It should be noted that the results given above apply to a two-tone stimulus, which means that measurement of these amplitudes should be passive, i.e., noninterfering with primaries. This condition can be met in physiological measurements of $2f_1 - f_2$. In psychophysical experiments the probe tone and the cancellation tone are subject to the same essential nonlinearity, which results in a different value for the amplitudes of distortion products.

With the pulsation threshold one has for the probe tone

$$R_p \propto r_p^\nu. \quad (14)$$

The pulsation threshold condition, $R_p = R_{2f_1-f_2}$, yields [as follows from Eq. (13)]

$$\begin{aligned} r_p &\propto r_1^{2/\nu} r_2^{1-2/\nu} & \text{if } r_2 \gg r_1 \\ &\propto r_1^{-1/\nu} r_2^{1/\nu} & \text{if } r_2 < r_1. \end{aligned} \quad (15)$$

Note that r_p follows the primaries linearly as long as $r_1/r_2 = \text{constant}$. The results of Eq. (15) are also given in Fig. 17 (dotted lines). Compared with the results of passive measurement [Eq. (13)] all slopes have been increased by a factor of $1/\nu$.

Most psychophysical measurements of combination tones have used the cancellation method. In that case a third tone is added to the stimulus in order to null the subjective response to the combination tone. We consider the cancellation of $2f_1 - f_2$ for the case where $\cos\theta_1 \approx \cos\theta_2 \approx \cos\theta_c$ (the index c refers to the cancellation tone). For this condition it is possible to show⁷ that cancellation can occur if the cancellation tone amplitude, represented by r_c , follows:

$$r_c \propto r_2 \quad \text{if } r_2 + r_c \ll r_1$$

FIG. 17. Theoretical amplitudes of the combination tone $2f_1 - f_2$ as a function of the level of the primary amplitudes, for $r_1 = r_2$ in (a), $r_2 = \text{constant}$ in (b), and $r_1 = \text{constant}$ in (c). The full lines represent the internally generated component. The dashed-dotted lines give the amplitudes as to be expected from the cancellation method, and the dotted lines give the predicted pulsation threshold measures for the $2f_1 - f_2$ amplitude. Vertical shifts are introduced arbitrarily in order to separate the curves. One horizontal log unit equals one vertical log unit. Asymptotic slopes are indicated; we used $\nu = 0.6$ in this figure.

$$\propto r_1^2 r_2^{-1} \quad \text{if } r_1 + r_c \ll r_2. \quad (16)$$

Note that the nonlinearity does not show up in the cancellation amplitude (at least not in its asymptotic slopes). The results of Eq. (16) are also given in Fig. 17 (dashed-dotted lines).

The differences between the slopes of r_p and r_c as functions of r_1 and r_2 provide means of estimating ν . Smoorenburg (1974) shows a comparison of pulsation threshold and cancellation amplitude data for the same subject. The qualitative differences observed agree with the differences between Eqs. (15) and (16). There are, however, some marked discrepancies. Both Eqs. (15) and (16) predict that the combination tone grows linearly with the combined primary level ($r_1/r_2 = \text{constant}$). The data show a slope which for the pulsation threshold is consistently lower than for the cancellation amplitude (Smoorenburg, 1974, Fig. 5), especially for small frequency ratios f_2/f_1 . The cancellation slope also appears to be below unity. It appears that this problem is related to the question of propagation, already mentioned at the end of Sec. I. If we assume that distortion products propagate distortion-free from the place where they are generated to the point tuned to the frequency of the distortion product, and if we assume further that at that place the distortion product is again subjected to the compressive nonlinearity, then we would arrive at the following results. The $2f_1 - f_2$ response at $x_{2f_1-f_2}$ would be [cf. Eq. (13)].

$$\begin{aligned} R_{2f_1-f_2}(x_{2f_1-f_2}) &\propto r_1^{2\nu} r_2^{\nu-2\nu} & \text{if } r_2 \gg r_1 \\ &\propto r_1^{\nu-2\nu} r_2^\nu & \text{if } r_2 < r_1. \end{aligned} \quad (17)$$

We have assumed that the direct effect of the primaries at $x_{2f_1-f_2}$ can be neglected. For the pulsation threshold one finds, assuming that the pulsation threshold r_p^* is determined at $x_{2f_1-f_2}$ [cf. Eq. (15)]

$$\begin{aligned} r_p^* &\propto r_1^2 r_2^{\nu-2}, \quad r_2 \gg r_1, \\ &\propto r_1^{\nu-1} r_2, \quad r_2 < r_1. \end{aligned} \quad (18)$$

We observe that r_p^* parallels $R_{2f_1-f_2}$ as given in Eq. (13) and Fig. 17. In case of cancellation one might suspect the interaction at $x_{2f_1-f_2}$ to occur predominantly between the propagated distortion product and the cancellation tone, the latter being affected mainly by f_1 . Then the cancellation tone contribution is

$$R_c(x_{2f_1-f_2}) \propto r_c^\nu \quad \text{if } r_1 \ll r_c \\ \propto r_c r_1^{\nu-1} \quad \text{if } r_1 \gg r_c \quad (19)$$

at $x_{2f_1-f_2}$. Equation (19) can be interpreted as follows: For small frequency differences $f_2 - f_1$, in general the second condition will be fulfilled. For larger differences between the primary frequencies the effect of f_1 on the cancellation tone will become insignificant, and the upper relation applies. The latter case is similar to Eq. (18). However, if f_2 approaches f_1 , then

$$r_c^* \propto r_1^{\nu+1} r_2^{\nu-2\nu} \quad \text{if } r_2 \gg r_1 \\ \propto r_1^{\nu-2\nu+1} r_2^\nu \quad \text{if } r_2 \ll r_1. \quad (20)$$

We observe that r_c^* follows the combined primary amplitude with the exponent $\nu^2 - \nu + 1$. Since $(\nu - 1)^2 > 0$ for all $\nu \neq 1$, we have $\nu^2 - \nu + 1 > \nu$ for $\nu \neq 1$. This means that the cancellation amplitude slope is always steeper than the pulsation threshold slope when presented as a function of the combined primary level. In a number of aspects the results derived here, taking account of propagation, agree better with psychophysical data than the results of Eqs. (15) and (16). However, because of the simplistic treatment of propagation we advocate a cautious interpretation of the results of this analysis.

We have assumed a nonlinearity in the mechanics at the hair-cell location. This raises the question of how the presumably relatively low-energetic distortion products are coupled back into the mechanism in order to allow for significant propagation. A first comment on this question is that it is not clear how far the distortion products have to be coupled back into the system. Although there is much circumstantial evidence for propagation to the hair cell, evidence for propagation along the basilar membrane (that is, of the distortion products) appears to be lacking (Wilson and Johnstone, 1972). The second comment regards the alternative location of the nonlinearity. Since the location should be between first and second filter (cf. Sec. I) it might also be located in the cochlear impedance (membrane and/or fluid). Evidence against such a more peripheral location of the nonlinearity, which is responsible for combination tone generation, might be found in psychophysical data by Smoorenburg (1972). A subject with a narrow-band hearing loss (presumably neural or hair cell) reported no aural combination tones when either of the primaries was presented within the dip.

In conclusion, propagation remains an important problem. The results of the present analysis tend to add to the evidence in support of the propagation hypothesis, but cannot be considered conclusive.

C. Some phase effects

As long as the propagation problem remains unsolved, it will be difficult to interpret psychophysically measured phases of combination tones. Goldstein (1972b) considers as one of the basic problems the difference in phase behavior of combination tones as a function of level when measured psychophysically or physiologically. In the first case there is a systematic change, in the neural data the phase is constant.

We will leave these problems as they are, and mention some consequences of our theory for phase effects. For a two-component stimulus it can be shown that the phase of the combination tone $mf_1 \pm nf_2$ simply is given by $m\phi_1 \pm n\phi_2$, except for a possible additional factor π which occurs if the minimum of m and n is odd. The latter addition rule applies only in case of compressive nonlinearity ($\nu < 1$) and odd-order combination tones ($m+n=\text{odd}$). Related rules can be derived for three or more component combination tones. They are, e.g., relevant to the phase effect reported by Buunen *et al.* (1974).

VI. DISCUSSION

The questions of the second filter and of cochlear nonlinearities have been receiving increasing attention over the past few years. Section I presented a number of references to those papers that were directly relevant to our study. [It is worthwhile noting that much of the data and questions referred to has been summarized in a good review paper by Møller (1972)]. In this section we first discuss the physiological basis of our theory, in the course of which additional references are given to some alternative theories. Second, we discuss some of the features of the "second filter + compressive nonlinearity" theory in general and of our theory in particular.

The two basic phenomena on which our second filter theory is based, i.e., the directional sensitivity of the hair cell and the directional distribution of vibration at the hair cell, appear to be fairly well established qualitatively. Thus, it appears plausible that in fact such a frequency weighting mechanism is operative. However, it remains to be established whether this mechanism can quantitatively account for the necessary sharpening. No such direct basis, unfortunately, underlies our assumption that the relevant nonlinear compression occurs at the hair cell base. Using this assumption we arrived at functional predictions which in some situations are undistinguishable from theories assuming nonlinear basilar membrane impedance. Depending on the question of how and to what extent the generated distortion products are coupled back into basilar membrane motion, direct measurement of distortion products in basilar membrane motion might provide a tool for deciding between the possibilities mentioned. The two possible locations for the sources of low-level combination tones mentioned above have already been stated explicitly by Goldstein (1967). We remark that a third possibility, viz., that nonlinearities arise at both (or even more) places, also deserves further attention (cf.

Weber and Mellert, 1975).

The directional distribution of cochlear vibration as effective at the hair cell bases, appears a worthwhile point for further investigation. A hardware model study by Zwicker (1974b) is of interest in this respect. In his scale model he observed a radial component in fluid flow between tectorial membrane and cuticular plates, but unfortunately with a tuning disparity $\alpha < 1$. This conflicts with our assumption on this point (Sec. II B) and therefore does not predict the data that can be accounted for by $\alpha > 1$ as, e.g., the relative location of low-frequency and high-frequency two-tone suppression areas (see Fig. 10).

Our theory provides no clue as to a possible functional difference and/or interaction of inner and outer hair cells as e.g., proposed by Zwislocki and Sokolich (1973, 1974) and, with a different interaction hypothesis, by Evans (1974). Since little evidence exists in support of a direct interaction between the inner and outer hair cell's afferent fibers (Spoendlin, 1974), we do not consider this to be a point to the disadvantage of our theory.

Regarding the hair cell's directional sensitivity, we emphasize that, although the *phenomenon* appears to be well established, little is known about the *mechanism* which is responsible for it. In particular, this leaves open the possibility that the metabolic situation of the hair cell is involved. Therefore, the reversible effects of antimetabolic and ototoxic drugs and hypoxia on the tuning curve, which effects are believed to provide direct evidence for the existence of a physiologically vulnerable second filter, (Evans, 1974; Evans and Klinke, 1974), might be considered as supporting our second filter theory.

So much for the discussion of the physiological basis of the model. Next we discuss model predictions. Although a good understanding of the model deserves a thorough analysis,⁷ we have tried to present a more heuristical approach in this, already comprehensive, paper. In line herewith, we have restricted ourselves by and large to qualitative verifications, but at several points we have indicated possibilities for quantitative verifications. Where we did not pursue a quantitative verification this was also because of lack of relevant and directly comparable data. This is particularly so for first and second filter shape, and hence for the distribution of θ over x or over f . We can be a little more specific about the tuning disparity factor α and the exponent of the compressive nonlinearity ν . We started most of our calculations with the somewhat arbitrary choice of $\alpha = 1.4$ and $\nu = 0.8$. Assuming that our interpretation of two-tone suppression is correct, α is simply the ratio of the "best frequency" of the high-frequency suppression area αCF and CF . The neurophysiological data by Sachs and Kiang (1968) do not always give an accurate value for αCF , but a conservative interpretation of their data allows the statement that $1.1 < \alpha < 1.6$ for CF 's ranging from about 1 to 20 kHz. The data seems to cluster about $\alpha = 1.2$. We find no indications for a systematic change of α with CF . From the paper by Arthur *et al.* (1971) we estimate $\alpha = 1.22$ at $CF = 8$ kHz. The psychophysical cor-

relate of two-tone suppression, as measured by Houtgast (1974a) yields $\alpha = 1.22$ at $CF = 1$ kHz. Thus, the estimate of $\alpha = 1.2$ appears to be more reasonable than the original choice of $\alpha = 1.4$. For the power of the nonlinearity Smoorenburg (1974) proposes $\nu = 0.6$, given as a qualitative fit for combination tone level data, and fitted quantitatively to psychophysical two-tone suppression data. We have adopted this value for ν in most of our later calculations. In combination with Siebert's (1972) firing rate function [Eq. (9)] it gives a very good fit to the two-tone suppression data by Rose *et al.* (1974), as we showed in Fig. 12. The value for ν initially chosen by Pfeiffer (1970), i.e., $\nu = \frac{1}{3}$, appears to be somewhat too low.

As we have pointed out in Eqs. (5) and (10), the value of ν also determines a certain amount of extra sharpening. The compressive nonlinearity increases the effective slopes of the second filter by a factor of $1/\nu$, which factor becomes increasingly important as ν goes from 1 to 0. Thus, the total amount of sharpening $H_2^{1/\nu}(f, x)$ is sensitive to changes in ν (cf. Fig. 14). We have mentioned already that ν could be obtained directly from Figs. 9 and 13. However, psychophysical measures of E are necessarily affected by the same nonlinearity, so that the observed slopes will be increased by a factor of $1/\nu$, which leaves ν only in the slope $1 - 1/\nu$ in the range between A_{M1} and A_{M2} (Fig. 13). Houtgast (1974a, 1974b) presents data which are not in disaccordance with $\nu = 0.6$. Smoorenburg (1974, Fig. 7) gives directly comparable data from an experiment where the masker level instead of the probe level was kept constant, from which data he concludes $\nu = 0.6$ as we have mentioned above. Available data, so far,⁵ appear to be consistent with a compressive nonlinearity that can be described with a single ν th law characteristic with $\nu = 0.6$. However, one should keep in mind the possibility that an approximation with more ν th law terms will turn out to be more precise.¹

We have seen in Sec. IV that our theory predicts a difference in sharpness of simultaneous masking and pulsation threshold curves. The pulsation threshold curve benefits from sharpening by the second filter and additional sharpening due to the compressive nonlinearity, whereas the simultaneous masking curve reflects only sharpening by the second filter. Also, the remarks in Sec. IV E apply primarily to simultaneous masking and may introduce additional differences (see, e.g., Vogten, 1974b). These differences may be described in terms of two-tone interaction. The theory explicitly shows the relation between the two phenomena. The manner in which the theory accounts for two-tone suppression is basically similar to Pfeiffer's (1970) theory. What is new is the introduction of the tuning disparity which accounts for the asymmetric location of low-frequency and high-frequency suppression areas. Also the description of combination tones is basically identical to Smoorenburg's (1972, 1974) theory. A major problem about combination tones, which remains unsolved, is how propagation occurs and whether it may account for the discrepancy noted by Goldstein (1972b) between the psychophysically and neurophysiologically measured phase of the combination tone. Also the pos-

sible functional relation between two-tone suppression and combination tones had been pointed out by Pfeiffer (1970) and Smoorenburg (1974). Our study underlines these ideas, giving an explicit formulation of the relation between the phenomena, including pure tone masking, and, moreover, presents a physiological basis for the theory. We have indicated possible experiments for further verification.

Finally, we speculate that the tuning disparity factor α might be related to the critical band (cf. Greenwood *et al.*, 1974).

ACKNOWLEDGMENTS

The ideas leading to the theory which has been presented in this paper have resulted and benefited from discussions with several colleagues from different laboratories. The notion that directional sensitivity of the hair cell in combination with directional distribution of vibrations in the cochlea provides sharpening arose a couple of years ago in a discussion with B. L. Cardozo at our institute. This idea received new impulses from a course by L. F. Frishkopf at M. I. T., bringing literature on directional sensitivity of the hair cell to my attention. The mathematical analysis benefited from earlier work by G. F. Smoorenburg (Institute for Perception, Soesterberg) and from suggestions by A. J. Breimer of our institute, who brought the early literature on ν th law rectifiers to my attention. H. G. ter Morsche from the Mathematics Department of the Eindhoven University of Technology helpfully provided me with the theorem referred to in Ref. 1.

C. A. A. J. Greebe and B. L. Cardozo of our institute and E. de Boer from the University of Amsterdam are acknowledged for comments on earlier drafts of this paper.

*Some preliminary results from this study have been presented at the symposium on "Facts and Models in Hearing," Tutzing 22–26 April 1974 (see Duifhuis, 1974).

¹This can be considered a generalization of Weierstrass's theorem for integer exponents. The proof for noninteger exponents goes back to Bernstein (1922).

²It should be noted that this similarity exists for the tuning curve near threshold. Isorate contours at higher levels may deviate significantly in shape.

³This conclusion is at variance with Békésy's (1953b) observations of directional sensitivity under mechanical stimulation of the tectorial membrane.

⁴A precise analytical derivation of this result requires more caution, since $r_1(t)$ and $r_2(t)$ depend on time differently. It can be shown, however, that as long as ω_1 and ω_2 are uncorrelated, the simplification to consider amplitudes leads to the proper results.

⁵As noted by M. B. Sachs, an effect not accounted for by the theory in its present form (one ν th law nonlinearity) is that in neural data the slope indicated by $\nu - 1$ in Fig. 9 (i.e., between A_{21} and A_{22}) appears to depend on the frequency of tone 2, tone 1 being the CTCF (Sachs, 1969).

⁶The data in Rose *et al.* (1974) are from AVCN units in the cat, but the authors judge the results to be also representative of data from primary units.

⁷Those readers interested in mathematical details of the theory are requested to contact the author.

- Arthur, R. M., Pfeiffer, R. R., and Suga, N. (1971). "Properties of 'Two-Tone Inhibition' in Primary Auditory Neurons," *J. Physiol.* **22**, 593–609.
- Békésy, G. von (1953a). "Description of Some Mechanical Properties of the Organ of Corti," *J. Acoust. Soc. Am.* **25**, 770–785.
- Békésy, G. von (1953b). "Shearing Microphonics Produced by Vibrations Near Inner and Outer Hair Cells," *J. Acoust. Soc. Am.* **25**, 786–790.
- Békésy, G. von (1960). *Experiments in Hearing* (McGraw-Hill, New York).
- Bernstein, S. N. (1922). "Sur l'ordre de la meilleure approximation des fonctions continues par des polynômes," *Académie Royale de Belgique, Classe des Sciences, Mémoires, Deuxième Série, Tome IV*, 1–103.
- Boer, E. de, and Bos, C. E. (1962). "On the concept of the critical band," *Proceedings of the 4th ICA, Copenhagen*, paper H 13.
- Boer, E. de (1967). "Correlation Studies Applied to the Frequency Resolution of the Cochlea," *J. Auditory Res.* **7**, 209–217.
- Boer, E. de (1969). "Reverse Correlation. II," *Proc. K. Ned. Akad. Wet.*, series C **72**, 129–151.
- Boer, E. de (1973). "On the Principle of Specific Coding," *Trans. ASME, J. Dyn. Sys. Meas. Control* **262**–273.
- Buunen, T. J. F., Festen, J. M., Bilsen, F. A., and Brink, G. van den (1974). "Phase Effects in a Three-Component Signal," *J. Acoust. Soc. Am.* **55**, 297–303.
- Dallos, P. J. (1969). "Combination Tone $2f_1 - f_2$ in Microphonic Potentials," *J. Acoust. Soc. Am.* **46**, 1437–1444.
- Dallos, P., Cheatham, M. A., and Ferraro, J. (1974a). "Cochlear Mechanics, Nonlinearities, and Cochlear Potentials," *J. Acoust. Soc. Am.* **55**, 597–605.
- Dallos, P., and Cheatham, M. A. (1974b). "Cochlear Microphonic Correlates of Cubic Difference Tones," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).
- Davis, H. (1965). "A Model For Transducer Action in the Cochlea," *Cold Spring Harbor Symp. Quant. Biol.* **30**, 181–190.
- Duifhuis, H. (1972). *Perceptual Analysis of Sound*, Doctoral dissertation (Eindhoven Univ. of Technology).
- Duifhuis, H. (1973). "Consequences of Peripheral Frequency Selectivity for Nonsimultaneous Masking," *J. Acoust. Soc. Am.* **54**, 1471–1488.
- Duifhuis, H. (1974). "An Alternative Approach to the Second Filter," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).
- Engelbreton, A. M., and Eldredge, D. H. (1968). "Model for Nonlinear Characteristics of Cochlear Potentials," *J. Acoust. Soc. Am.* **44**, 548–554.
- Evans, E. F. (1970). "Narrow 'Tuning' of the Responses of Cochlear Nerve Fibres Emanating from the Exposed Basilar Membrane," *J. Physiol.* **208**, 75P–76P (A).
- Evans, E. F., and Wilson, J. P. (1971). "Frequency Sharpening of the Cochlea: The Effective Bandwidth of Cochlear Nerve Fibers," *Proceedings of the 7th ICA, Budapest*, paper 23H4.
- Evans, E. F. (1972a). "The Frequency Response and Other Properties of Single Fibres in the Guinea-Pig Cochlear Nerve," *J. Physiol.* **226**, 263–287.
- Evans, E. F. (1972b). "Does Frequency Sharpening Occur in the Cochlea?" *Symposium on Hearing Theory 1972, IPO Eindhoven, Holland*.
- Evans, E. F. (1974). "Auditory Frequency Selectivity and the Cochlear Nerve," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).
- Evans, E. F., and Klinke, R. (1974). "Reversible Effects of Cyanide and Furosemide on the Tuning of Single Cochlear Nerve Fibres," *J. Physiol.* **242**, 129P–131P (A).
- Feuerstein, E. (1957). "Intermodulation Products for ν -Law

- Biased Wave Rectifier for Multiple Frequency Input," Q. Appl. Math. 15, 183-192.
- Flock, Å. (1965a). "Electron Microscopic and Electrophysiological Studies on the Lateral Line Canal Organ," Acta Otolaryngol. Suppl. 199.
- Flock, Å. (1965b). "Transducing Mechanisms in the Lateral Line Organ Receptors," Cold Spring Harbor Symp. Quant. Biol. 30, 133-145.
- Goldstein, J. L. (1967). "Auditory Nonlinearity," J. Acoust. Soc. Am. 41, 676-689.
- Goldstein, J. L., and Kiang, N. Y. S. (1968). "Neural Correlates of the Aural Combination Tone $2f_1 - f_2$," Proc. IEEE 56, 981-992.
- Goldstein, J. L., Baer, Th., and Kiang, N. Y. S. (1971). "A Theoretical Treatment of Latency, Group Delay, and Tuning Characteristics for Auditory Nerve Responses to Clicks and Tones," in *Physiology of the Auditory System*, edited by M. B. Sachs (National Education Consultants, Baltimore).
- Goldstein, J. L. (1972a). "Neural Phase Locking to Combination Tones Predicted by Simple Transducer Models," J. Acoust. Soc. Am. 52, 142 (A).
- Goldstein, J. L. (1972b). "Evidence from Aural Combination Tones and Musical Tones Against Classical Temporal Periodicity Theory," Symposium on Hearing Theory 1972, IPO Eindhoven, Holland.
- Greenwood, D. D., Merzenich, M. M., and Roth, G. L. (1974). "Two-tone suppression and combination-tone driving," J. Acoust. Soc. Am. 56, S22 (A).
- Hall, J. L. (1972). "Auditory Distortion Products $f_2 - f_1$ and $2f_1 - f_2$," J. Acoust. Soc. Am. 51, 1863-1871.
- Hall, J. L. (1974). "Combination Tones $2f_1 - f_2$ and $f_2 - f_1$ in a Nonlinear Model of the Basilar Membrane," 8th International Congress of Acoustics, London, Vol. I, p. 170.
- Hind, J. E., Rose, J. E., Brugge, J. F., and Anderson, D. J. (1970). "Two-Tone Masking Effects in Squirrel Monkey Nerve Fibers," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden).
- Houtgast, T. (1974a). *Lateral Suppression in Hearing*, Doctoral dissertation (Free University, Amsterdam).
- Houtgast, T. (1974b). "Lateral Suppression and Loudness Reduction of a Tone in Noise," Acustica 30, 214-221.
- Johnstone, B. M., and Taylor, K. J. (1970). "Mechanical Aspects of Cochlear Function," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden).
- Johnstone, B. M., Taylor, K. J., and Boyle, A. J. (1970). "Mechanics in the Guinea Pig Cochlea," J. Acoust. Soc. Am. 47, 504-509.
- Khanna, S. M., Sears, R. E., and Tonndorf, J. (1968). "Some properties of Longitudinal Shear Waves: A Study by Computer Simulation," J. Acoust. Soc. Am. 43, 1077-1084.
- Kiang, N. Y. S., Watanabe, T., Thomas, E. C., and Clark, L. F. (1965). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT Press, Cambridge, MA).
- Kiang, N. Y. S., and Moxon, E. C. (1974). "Tails of Tuning Curves of Auditory-Nerve Fibers," J. Acoust. Soc. Am. 55, 620-630.
- Kim, D. O., Molnar, C. E., and Pfeiffer, R. R. (1973). "A System of Nonlinear Differential Equations Modeling Basilar-Membrane Motion," J. Acoust. Soc. Am. 54, 1517-1529.
- Kohliöffel, L. U. E. (1972). "A Study of Basilar Membrane Vibrations, III. The Basilar Membrane Frequency Response Curve in the Living Guinea Pig," Acustica 27, 82-89.
- Lowenstein, O., and Wersäll, J. (1959). "A Functional Interpretation of the Electron-Microscopic Structure of the Sensory Hairs in the Cristea of the Elasmobranch *Raja Clavata* in Terms of Directional Sensitivity," Nature 184, 1807-1808.
- Möller, A. R. (1972). "Coding Sound in Lower Levels of the Auditory System," Q. Rev. Biophys. 5, 59-155.
- Pfeiffer, R. R. (1970). "A Model for Two-Tone Inhibition of Single Cochlear Nerve Fibers," J. Acoust. Soc. Am. 48, 1373-1378.
- Pfeiffer, R. R., Molnar, C. E., and Cox, J. R. (1974). "The Representation of Tones and Combination Tones in Spike Discharge Patterns of Single Cochlear Nerve Fibers," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).
- Rhode, W. S. (1971). "Observations of the Vibration of the Basilar Membrane in Squirrel Monkeys Using the Mössbauer Technique," J. Acoust. Soc. Am. 49, 1218-1231.
- Robertson, D. (1974). "Cochlear Neurones: Frequency Selectivity Altered by Perilymph Removal," Science 186, 153-155.
- Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (1967). "Phase-locked Responses to Low-Frequency Tones in Single Auditory Nerve Fibers of the Squirrel Monkey," J. Neurophysiol. 30, 769-793.
- Rose, J. E., Gibson, M. M., Kitzes, L. M., and Hind, J. E. (1974). "Observations on Phase-Sensitive Neurones of the Anteroventral Cochlear Nucleus of the Cat: Nonlinearity of Cochlear Output," J. Neurophysiol. 37, 218-253.
- Sachs, M. B., and Kiang, N. Y. S. (1968). "Two-Tone Inhibition in Auditory-Nerve Fibers," J. Acoust. Soc. Am. 43, 1120-1128.
- Sachs, M. B. (1969). "Stimulus-Response Relation for Auditory-Nerve Fibers: Two-Tone Stimuli," J. Acoust. Soc. Am. 45, 1025-1036.
- Scharf, B. (1964). "Partial Masking," Acustica 14, 16-23.
- Schroeder, M. R. (1969). "Relation between Critical Bands in Hearing and the Phase Characteristics of Cubic Difference Tones," J. Acoust. Soc. Am. 46, 1488-1492.
- Schroeder, M. R., and Hall, J. L. (1974). "Model for the Mechanical to Neural Transduction in the Auditory Receptor," J. Acoust. Soc. Am. 55, 1055-1060.
- Siebert, W. M. (1968). "Stimulus Transformations in the Peripheral Auditory System," in *Recognizing Patterns*, edited by P. A. Kolars and M. Eden (MIT Press, Cambridge, MA).
- Siebert, W. M. (1972). "What Limits Auditory Performance?" Proc. 4th Conf. Int. Union of Pure and Appl. Biophys., Moscow.
- Small, A. M. (1959). "Pure-Tone Masking," J. Acoust. Soc. Am. 31, 1619-1625.
- Smoorenburg, G. F. (1972). "Combination Tones and their Origin," J. Acoust. Soc. Am. 52, 615-632.
- Smoorenburg, G. F. (1974). "On the Mechanisms of Combination Tone Generation and Lateral Inhibition in Hearing," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).
- Spoendlin, H. (1972). "Innervation Densities of the Cochlea," Acta Oto-Laryngol. 73, 255-248.
- Spoendlin, H. (1974). "Neuroanatomy of the cochlea," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).
- Tonndorf, J. (1970). "Cochlear Mechanics and Hydrodynamics," in *Foundations in Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York).
- Verschuure, J., Rodenburg, M., and Maas, A. J. J. (1974). "Frequency Selectivity and Temporal Effects of the Pulsation Threshold Method," 8th International Congress on Acoustics, London, Vol. I, p. 131.
- Vogten, L. L. M. (1972). "Pure-Tone Masking of a Phase-Locked Tone Burst," IPO-Ann. Prog. Rep. 7, 5-16.
- Vogten, L. L. M. (1974a). "Pure-Tone Masking; New Results from a New Method," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).
- Vogten, L. L. M. (1974b). "Low-Level Pure-Tone Masking and Two-Tone Suppression," IPO-Ann. Prog. Rep. 9, 22-31.
- Weber, R., and Mellert, V. (1975). "On the Nonmonotonic

- Behavior of Cubic Distortion Products in the Human Ear," J. Acoust. Soc. Am. **57**, 207-214.
- Wegel, R. L., and Lane, C. E. (1924). The Auditory Masking of one Pure Tone by Another and its Probable Relation to the Dynamics of the Inner Ear," Phys. Rev. **23**, 266-285.
- Weiss, T. F., Mulroy, M. J., and Altmann, D. W. (1974). "Intracellular Responses to Acoustic Clicks in the Inner Ear of the Alligator Lizard," J. Acoust. Soc. Am. **55**, 606-619.
- Wersäll, J., Flock, Å., and Lundquist, P.-G. (1965). "Structural Basis for Directional Sensitivity in Cochlear and Vestibular Sensory Receptors," Cold Spring Harbor Symp. Quant. Biol. **30**, 115-132.
- Wilson, J. P., and Johnstone, J. R. (1972). "Capacitive Probe Measures of Basilar Membrane Vibration," Symposium on Hearing Theory 1972, IPO Eindhoven, Holland.
- Zwicker, E. (1974a). "On a Psychoacoustical Equivalent of Tuning Curves," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).
- Zwicker, E. (1974b). "A 'Second Filter' Established Within the Scala Media," in *Facts and Models in Hearing* edited by E. Zwicker and E. Terhardt (Springer, Berlin).
- Zwislocki, J. J., and Sokolich, W. G. (1973). "Velocity and Displacement Responses in Auditory-Nerve Fibers," Science **182**, 64-66.
- Zwislocki, J. J., and Sokolich, W. G. (1974). "Neuromechanical Frequency Analysis in the Cochlea," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).