Magnitude and phase-frequency response to single tones in the auditory nerve

Jont B. Allen
Bell Laboratories, Murray Hill, New Jersey 07974

(Received 13 September 1982; accepted for publication 24 February 1983)

In this paper we describe magnitude and phase measurements obtained from primary single unit recordings in the cat auditory nerve. Levels range from threshold to 100 dB SPL, with frequencies from 0.1–30.0 kHz. The upper limit on the phase measurements was limited by the loss of neural phase locking at 4–5 kHz. For each unit, the frequency tuning curve (FTC) was measured by the method of Kiang and Moxon [M. C. Liberman, J. Acoust. Soc. Am. **63**, 442–445 (1978)] to establish the threshold frequency response of the unit. Data from several selected animals, organized by characteristic frequency (CF), are presented showing phase response, group delay, frequency tuning, and tuning slope for each CF range. The major emphasis in this paper is on the "linear" aspects of the data as characterized by the filter properties of the single unit response, however a number of nonlinear (level-dependent) effects are described. Data are presented showing the phase response normalized by the cochlear microphonic (CM) recorded at the round window membrane. This normalization simplifies the phase data since it produces a constant phase slope with respect to frequency (constant group delay) for high CF units ($f_{CF} > 1$ kHz) for frequencies more than one octave below their characteristic frequencies. A model of CM, as measured at the round window (RW), is presented and compared to experimental CM measurements. The CM model gives a reasonable fit to the experimental data above 500 Hz. Our interpretation of the CM normalization is that it removes driver and middle ear effects. In the model we assume that the CM is generated by the displacement of the basilar membrane near the round window recording site.

PACS numbers: 43.63.Pd, 43.63.Th

INTRODUCTION

It is widely accepted that the main function of the cochlea is that of an analyzing filter bank which reduces an acoustic stimulus to its various frequency components. Neural measurements indicate that the cochlear filters are very sharply tuned, having skirts with slopes greater than several hundred dB/oct. In the cat, an estimated 40 000 afferent neurons carry the resulting narrow-band hair-cell responses to the cochlear nucleus, apparently coded as a pulse-rate modulated digital code.

While the frequency selectivity of the auditory neurons has long been the subject of frequent study, only a few studies of neural signal phase have been undertaken. One purpose of this paper is to present detailed neural magnitude and phase data which might be used to refine mathematical models of the cochlea.

The first primary auditory neural phase measurements, as a function of frequency, were made by two independent groups in 1970 (Pfeiffer and Molnar, 1970; Anderson *et al.*, 1971). Anderson *et al.* also made phase versus frequency measurements as a function of sound level. At high levels they found that the phase became level-dependent (a nonlinear effect). Unfortunately the two plots showing their level-dependent results were normalized to the phase measured at 90 dB SPL (pressure *re*: 20 μPa), which made their results difficult to interpret since their normalization level corresponded to the phase showing the greatest nonlinear effect. Their main finding [see Figs. 7(c),(d)] was that the phase was proportional to frequency with a proportionality constant which depended on the characteristic frequency (CF) of the unit being measured, with low-frequency units displaying a larger slope. In this paper we show that the phase slope $\partial\phi / \partial\omega$ is not a CF-dependent constant for a given unit, but depends on frequency in a rather complex but deterministic manner.

Pfeiffer and Molnar (1970) [see Figs. 7(a),(b)] did not report constant phase slopes, with the exception of those CFs near 1 kHz. While the Pfeiffer and Molnar results describe a frequency and CF-dependent phase slope, or group delay, their description of phase does not contain any significant level dependence.

In both of the above papers only a few phase curves are given and only general trends are evident from the plots. The nonlinear (level-dependent) results of Anderson *et al.* are interesting but not directly interpretable because of the normalization procedure used.

Arthur *et al.* (1971) measured neural phase in the presence of a rate-suppressing second tone. They found that the phase did not depend on the rate suppression effect (two-tone suppression). Also in 1971, Goldstein *et al.* published a single phase curve (with no explanation as to how it was obtained). Zwislocki and Sokolich (1974) attempted to study hair cell transduction by measuring phase in the auditory nerve of the mongolian gerbil. In their comparisons they pooled data across normal and cochlear-damaged units. Since only a few data points were presented, it is impossible to assess their results in any general way. Also, Pfeiffer *et al.* (1974) measured the phase of nonpropagating harmonic distortion products. These distortion products are now believed

 0001-4966/83/062071-22$00.80

to be the product of hair cell transducer rectification. In 1976, Arthur presented further phase measurements under rate suppression conditions. Later, Kim *et al.* (1980) published detailed neural phase measurements—in a very different form than had been previously presented. Kim *et al.* measured the phase response of a large number of units to a small number of tones, and they plotted their results as a function of $\log(f_{CF})$ (where f_{CF} is the characteristic frequency), to represent approximately the phase response as a function of longitudinal position along the basilar membrane. The Kim *et al.* single-tone phase measurements were particularly interesting to this author because of the appearance of a leading π phase inflection basal to the characteristic place, because such a phase shift was predicted by the author's second-filter model (Allen, 1980). Furthermore, in keeping with that theory, the position of the leading π phase shift found by Kim *et al.* depended on the single tone stimulus frequency.

I. METHODS

A. Surgical preparation methods

Healthy mature cats weighing between 2.0 and 5.0 kg were used. Most animals were visibly free of ear disease and parasites. Animals were first given an intramuscular injection of chlorpromazine (14 mg/kg body weight), followed by an intraperitoneal injection of sodium pentobarbitol (Nembutal) (22 mg/kg body weight). The level of anesthesia was maintained throughout the course of the experiment by periodically (e.g., every 2 h) administering 10% of the original Nembutal dosage, as required to minimally suppress muscular reflexes.

Using a dorsal approach, the bulla and septum were opened on the left side. Two silver-ball electrodes were cemented to the skull; one grounded in the bone, dorsal to the meatus, and the other was placed directly on the round window membrane. All signals were presented through a push–pull electrostatic acoustic driver as described by Sokolich (1977). Experiments were performed in an ICA double-walled sound-proof chamber. The round window electrode was used to measure tone-pip $N1$ thresholds and CM (relative to the bone electrode) as a function of frequency and level. The auditory nerve was next exposed by opening the dorsal aspect of the posterior fossa and retracting the cerebellum. After the retraction, CM and threshold tone-pip $N1$ were again measured and compared with the values measured prior to the retraction. Glass microelectrodes, filled with 3M KCl, were used to record from single auditory nerve fibers. Electrode impedances ranged from 5 to 30 MΩ (dc). The electrodes were advanced into the auditory nerve using a Burleigh "inch-worm" PZT drive. This drive provided a precise control over the electrode tip placement (e.g., $\Delta x \ll 1\,\mu$m). Under optimal recording conditions, units could frequently be held for well over 1 h. However, typical measurements never required more than 10–15 min; accordingly we frequently moved to the next unit before the previous unit was lost. Typically, a distance of 10–50 μm was traveled between units when experimental recording conditions were optimal. After 1200–2500 μm of electrode motion, the electrode was withdrawn and manually advanced until a new unit was acquired. At that point, the digital position indicator of the PZT drive was reset and the electrode was again advanced under push button, electronic control.

B. Computer methods

All data generation and collection was done using a 16-bit Data General S/200 Eclipse minicomputer running with the Data General Real Time Disk Operating System (RDOS). This system was equipped with 192 kbytes of core memory, a 96 Mbyte disk drive, two Tektronics 4010 graphic I/O terminals, a 32-bit floating-point hardware multiplier, 100 kHz 15-bit D/A and A/D multichannel hardware with programmable clock, a nine-track magnetic tape drive, a line printer, real-time plotting hardware, hardware which controlled an analog signal attenuator having a 100-dB range in 1-dB steps, and a Rockland programmable analog anti-aliasing filter (model 816, 48 dB/oct, Butterworth response). A special purpose interrupt-driven digital real-time counter (we shall refer to this hardware as the "histogram counter") was designed which was incremented by the D/A clock. The histogram counter was used to record the neural event times to within one D/A clock period, which in this series of experiments was 21 μs. Each neural event caused a computer interrupt and loaded the histogram counter value into a hardware register. Each time this interrupt was answered (within 120 μs) the histogram hardware was reset to accept the next neural spike. The histogram counter could also be cleared by the unused least significant bit (LSB) of the real-time D/A word, thereby allowing the neural histogram to be exactly synchronized to the D/A output signal. All signals presented to the cat were generated digitally and were presented via the D/A through the acoustic driver transducer. The analog attenuator was only used to determine the maximum signal level for each set of presentations (e.g., once for each tuning curve or once each phase versus frequency measurement).

C. Data collection methods

The method of Kiang and Moxon (Liberman, 1978) was used to measure the neural threshold tuning curve. A shaped, 50-ms tone burst having sound level S and frequency f was presented to the cat's ear followed by a 50-ms period of silence. The number of neural spikes $N(S)$ counted during the tone burst was determined by the histogram counter hardware, as well as the number of spikes during the silent interval $M = N(0)$. At each of the measurement frequencies, the signal level S was then stepped in multiples of 2/3 dB until $N(S_0) \simeq M + 1$. This condition specifies that the number of spikes during the driven interval $N(S)$ at level $S = S_0$ is one more than the number of spikes during the silent interval M. In this manner the threshold excitation tuning $S_0(f)$ was found as a function of frequency. As a result of some experimentation, it appears that the Kiang–Moxon neuron tuning curve paradigm is reproducible to within about ± 3 dB in level, independent of the estimated threshold value [see for example Fig. 19(b)]. Frequencies were swept from high to low with a density of 90 points per decade.

Harmonic and intermodulation distortion must be carefully controlled in the transducer and D/A to avoid spurious tuning artifacts due to harmonic distortion. It is a fact that cochlear measurements strain transducer and D/A technology in terms of distortion levels. Harmonic distortion must be at least 65 dB down and probably should be 85 dB below the primary. A DATEL model HR15B D/A followed by a simple bandlimited current to voltage converting op-amp circuit gave very low distortion conversions. Spurious responses for the D/A system were never greater than -68 dB for two tones, but were typically 85 dB below a single tone 10-V peak output. D/A "glitches" could not be measured in the unfiltered time domain waveform and slew distortion appeared to be the dominant type of distortion (Blesser, 1978). Transducer distortion set the limits on the system distortion above 80 dB SPL. At 90 dB SPL the maximum transducer harmonic distortion was at least 60 dB below the primary signals over the range of signal conditions tested.

D. Experimental methods

Neural units were found using a gated wideband digitally generated acoustic noise search stimulus. The noise level was such that it would drive any unit having a threshold below 80 dB *re*: 20 μPa. Once a unit was acquired, the neural threshold curve (FTC, or frequency tuning curve) was determined, as described above, and displayed on the computer terminal. After the CF (characteristic frequency) had been determined from the FTC, a delay compensation $T_0(f_{CF})$ was entered. The delay compensation was an empirically determined function of CF (its choice will be discussed below). Next a locus (straight line connecting two points in the log–frequency, log–pressure plane) through the suprathreshold response range in the frequency versus SPL (sound pressure level) plane was established via input cursors on the computer terminal. The sweep defined the locus of frequencies and levels used in measuring the phase. For most of the experiments reported on here, single tones were generated having a geometric frequency separation of 90 points per decade quantized to the nearest multiple of 22 Hz. At each measurement frequency a continuous tone was gated on for 2 s. During this time, neural spikes were collected and a stimulus locked (PST, post-stimulus time) histogram was formed. The neural phase was then estimated from the PST histogram by computing the phase of the Fourier transform of the period histogram at the stimulus frequency (Pfeiffer and Molnar, 1970). All measured phases have been referenced to the computer defined stimulus. Ear canal pressure magnitude and phase were measured for each animal several times during the experiment via a calibrated probe microphone. For frequencies of interest, namely frequencies below 5 or 6 kHz, the driver response was typically quite flat ($\leqslant \pm$ 3-dB deviation from a flat response). Although the magnitude and phase of ear canal pressure was recorded, ear canal corrections were *not* made to the plots presented in this paper. Such corrections were not necessary because of the uniformity of the measured driver magnitude and phase response below 6 kHz. In many figures the driver voltage required to produce a pressure of 10 dB *re*: 20 μPa is shown as a solid line (e.g., Fig. 1).

During estimation of the neural phase at each frequency, a specific CF-dependent time delay compensation (for acoustic, cochlear, and neural delays) was made. This delay compensation was equivalent to removing a linear regression line from the phase versus frequency data, since a pure delay is equivalent to a phase shift proportional to frequency. This compensation was used to *approximately* remove the various *frequency dependent* delays (dispersive delays) in the cochlea and middle ear. In mathematical terms, a linear system having a pure delay of T seconds would have a phase lag $\phi_T(\omega)$ given by

$$\phi_T(\omega) = -\omega T.$$

A phase corresponding to the constant delay $T_0(f_{CF})$ was removed from the neurally measured phase $\phi_n(\omega)$, resulting in the phase displayed in the plots $\phi(\omega)$:

$$\phi(\omega) = \phi_n(\omega) + \omega T_0, \tag{1}$$

where T_0 is the CF-dependent delay compensation, $\omega = 2\pi f$, and f is the frequency of the stimulating tone. The $T_0(f_{CF})$ values were determined empirically as those values which would nominallly keep the delay compensated phase $\phi(\omega)$ within $\pm \pi$ rad for frequencies below the CF. By the use of

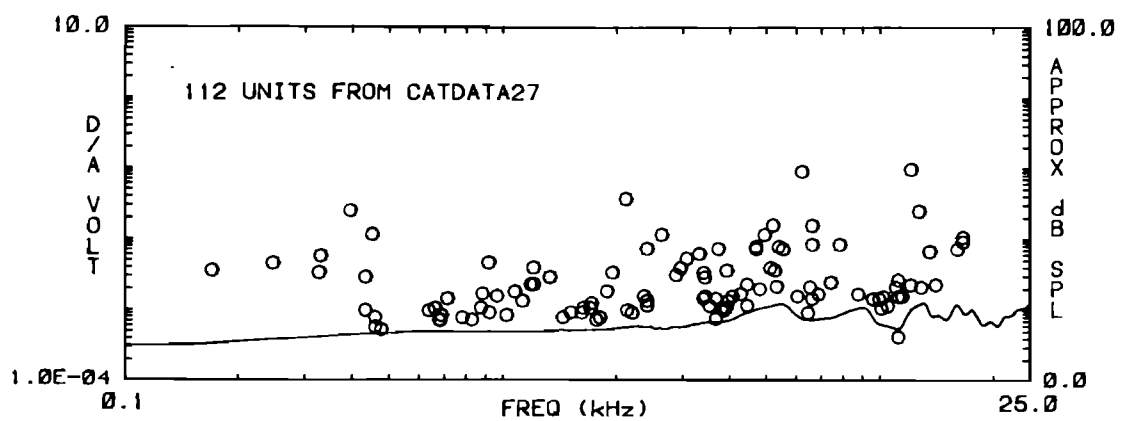

FIG. 1. This figure shows the best frequency thresholds for all units of cat 27 between 0.1 to 25.0 kHz. The left-hand ordinate is the acoustical driver voltage, plotted on a log scale. The solid line is the frequency-dependent driver voltage required to produce a pressure of 10 dB *re*: 20 μPa in the ear canal according to a calibrated probe tube measurement of the ear canal pressure.

this CF-dependent delay compensation, the phase roll was sufficiently reduced that we avoided the problem of 2π phase ambiguities. Furthermore the delay compensation allowed us to observe details that would otherwise be obscured by the large phase slopes which result from the various acoustic, cochlear, and neural delays. The original measured phase ϕ_n is easily reconstructed exactly from ϕ by use of Eq. (1) given T_0.

Very late in the series of experiments reported here, we switched to a 1-s tone duration rather than the 2-s tones used during the early part of the experiment. This allowed a faster measurement time, with a noticeable (but small) increase in estimated phase variance. A systematic estimate of the phase variance as a function of the tone duration and frequency was not investigated. For the case of the 2-s average, the phase variance was clearly very small (less than $\pi/20$) under most conditions of measurement as estimated by the variation in the closely spaced (in frequency) measurements. Some small time-dependent effects were noticed, but they seemed to be very much smaller than systematic nonlinear phase effects which we shall discuss in Sec. VII.

II. ASSESSMENT OF ANIMAL VARIABILITY

In order to assure that the data being presented are typical, it is necessary to establish some general measures of neural response for each animal.

In Fig. 1 we show the thresholds at CF for all 112 units measured in cat 27. The solid line running along the lower part of the plot is the computed driver voltage required to produce 10 dB *re*: 20 μPa at the probe microphone. From this plot we see that the minimum thresholds are a uniform 10 dB (*re*: 20 μPa) over frequency. The left-hand ordinate in this plot is D/A volts. From an independent calibration at 200 Hz, 10 V on the D/A corresponds to 100 dB SPL. For convenience this ordinate is given on the right of the figure. Thus the animal threshold sensitivity is about 10 dB SPL. Liberman (1978) reports minimum thresholds near 0 dB SPL and even lower thresholds in chamber-raised cats. It is not clear what this 10 dB difference is due to. All of our levels are peak values whereas rms is usually quoted. This makes our levels 3 dB greater than Libermans, assuming he used an rms scale. Differences in the tuning curve paradigm (Liberman used a 0 tuning curve criterion, while we used a criterion of 1) will give rise to threshold differences of about 6 dB. Chamber noise could also raise the thresholds, however, the chamber noise is known to be frequency-dependent (it increases at low frequencies). Thus chamber noise as a possible source of error is inconsistent with the neural thresholds which are relatively constant over frequency.

Figure 2(a) shows smoothed tuning curve data for cat 47, while Fig. 2(b) shows the tuning curve slopes expressed in dB/oct. The tuning data have been smoothed by the use of

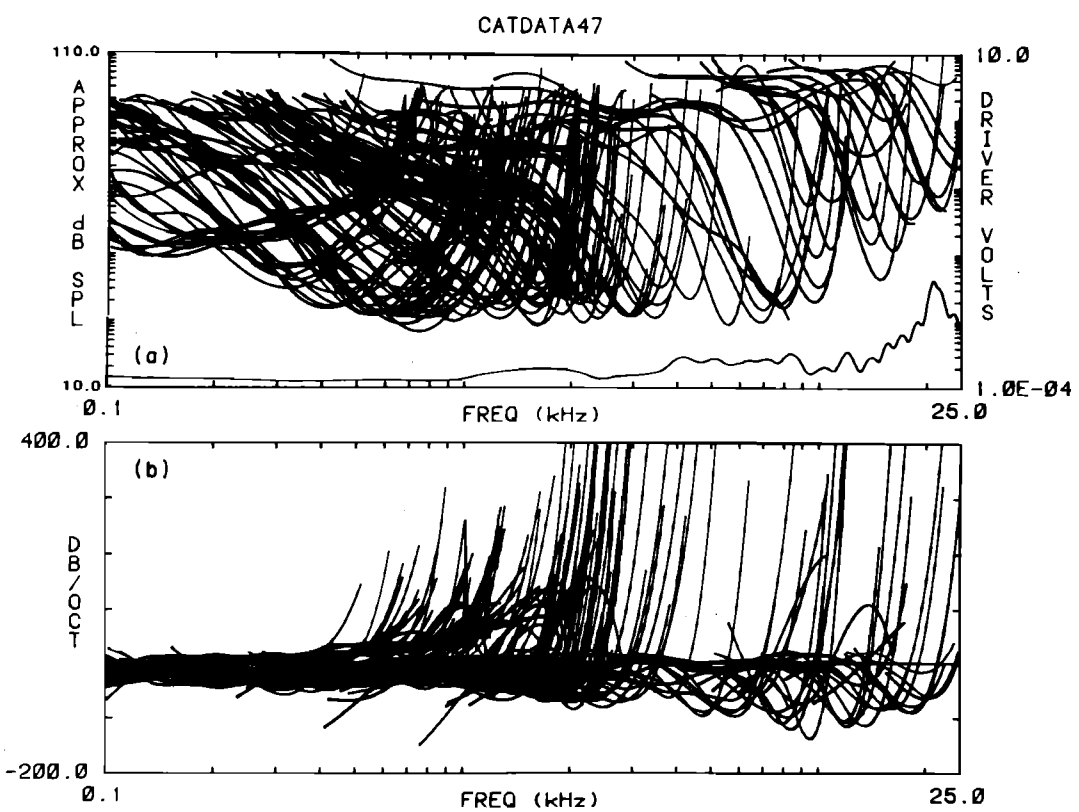

FIG. 2. (a) Each tuning curve was fitted by a least-square spline fit, as discussed in the text. The maximum error between the spline fit and the raw data were never greater than 8 dB (or the data were not plotted) using this procedure. Figures 19(a) and 20(a) display both the raw data and the spline fit for comparison. The upper panel (a) shows the spline fit tuning curves. The solid line along the bottom of the panel is the driver voltage required to produce a pressure of 10 dB *re*: 20 μPa in the ear canal. The unit threshold is near 30 dB *re*: 20 μPa. (b) The lower panel gives each tuning curve slope with respect to frequency, in units of dB/oct. Positive slopes correspond to frequencies above CF and negative slopes to frequencies less than CF. This animal is considered to be in the normal group but shows fairly broad tuning for CFs below 1 kHz as characterized by the magnitude of the slopes immediately above and below CF.

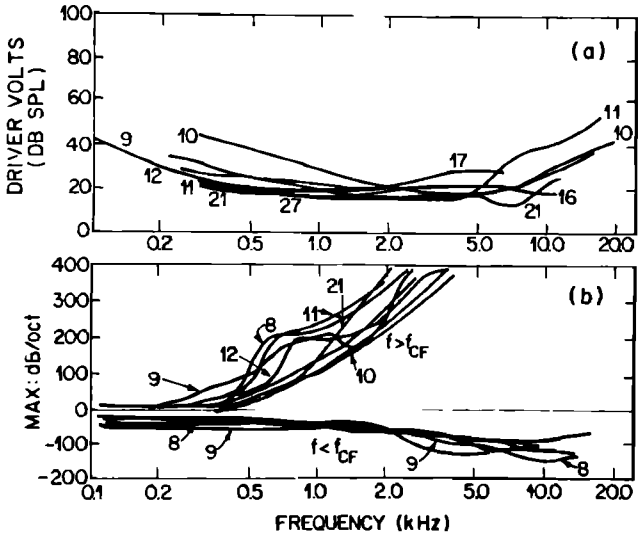

FIG. 3. In this figure we compare thresholds (a) and maximum absolute slopes (b) for a number of our lowest threshold animals measured in the study. The lower panel (b) consists of two families of curves, maximum slopes above CF (positive slopes) and maximum absolute slopes below CF (negative slopes). Using a threshold criterion, all animals, with the possible exceptions of 10 and 11, are considered to be normal. Note the large slope variability for frequencies above CF ($f > f_{CF}$) between 0.5 to 2 kHz.

least-squares fitted cubic splines (Fox *et al.*, 1976) which made practical the computation of approximate tuning curve slopes. The spline fits are smoothing polynomials whose slopes are exactly computable. In forming the spline fit to the original frequency tuning curves (FTCs), the error was computed between the actual data and the smoothed spline fit. If the error was greater than 8 dB, the data in the neighborhood of the large error were not used. Very little data (i.e., less than 2%) were rejected by this procedure. The truncated thresholds above CF in Fig. 2(a) for units above 10 kHz are examples of truncated data. The truncations frequently affected the slope estimates above 300 dB/oct. The data of Fig. 2 are a marginal representative of those animals having the lowest thresholds. Later, we show examples of raw FTC data along with the spline fits [Figs. 19(a) and 20(a)].

In Fig. 3 we summarize three characterizing measures for a number of animals. Figure 3(a) shows the locus of minimum thresholds (in terms of the driver transducer voltage) for six different animals. The numerical label indicates the animal identification number. In Fig. 3(b) we display the maximum slopes, in dB/oct, heuristically estimated from threshold tuning curve slope data [such as seen in Fig. 2(b)] obtained from each animal. The magnitudes of the slopes are consistent with the slope estimates of Evans and Wilson (1971). No special significance should be placed on our specific measure of the maximum slopes since it was largely based on where each slope estimate showed a significant error. The positive slopes correspond to frequencies above the best frequency (CF) while the negative slopes correspond to frequencies below CF. When slight damage was observed in the high-frequency region ($f > 2$ kHz), as indicated by elevated thresholds, the maximum negative slope (below CF) seemed to be a very sensitive indicator; thus threshold elevations at CF and the maximum slope below CF seem to be correlated for high-CF units (as the threshold increases the slope below CF decreases in magnitude).

When an animal showed thresholds and slopes in the range of those seen in Fig. 3, and when the number of recorded units was greater than 90, the data were defined to be "normal." Data from different animals were never pooled, although such a procedure seems justifiable based on the similarities seen across animals.

III. GROUP DELAY MEASUREMENTS

Because of the very large range of delays within the cochlea, it is quite difficult to display the measured phase data while simultaneously displaying the detailed phase variations as a function of frequency. For this reason is is useful to plot the group delay rather than the phase. The group delay is defined as

$$\tau_g(\omega) = -\frac{\partial \phi(\omega)}{\partial \omega}, \tag{2}$$

where $\phi(\omega)$ is the phase at radian frequency $\omega = 2\pi f$. For linear systems, under very general conditions, the group delay has the physical interpretation as the delay that a tone burst envelope undergoes in passing through the linear system (Papoulis, 1962, p. 135). To the extent that the basilar membrane mechanical system acts as a delay line (in the spirit of traveling waves on the basilar membrane) the group delay may be thought of as cochlear delay. From probe tube measurements in the ear canal, it was found that the acoustic drive system introduced 0.45 ms of delay independent of frequency up to 10 kHz. Neural propagation delays are frequently quoted as being about 0.7–1.0 ms (Anderson *et al.*, 1971; Eggermont, 1979). Thus delays of more than 1.0–1.15 ms are assumed to be cochlear in origin. From the above estimates we shall see that cochlear delays for CFs greater than 2 kHz appear to be small (in the noise).

Regardless of its interpretation, the group delay is useful because it greatly reduces the range of data. *This allows us to effectively display data from different CF ranges on one plot.*

The upper panel, Fig. 4(a), shows all the FTC's for animal 9 and shows the locus of each stimulus (dashed lines) used in the phase measurements. In Fig. 4(b), we show the group delay as derived from the neurally measured phase data in cat 9. The purpose of Fig. 4(b) is to show the range of the group delay over the CF range (as outlined by the heavy lines). From this figure it is clear that the phase slopes are highly correlated with frequency above 400 Hz. High CF units have small phase slopes and small phase slope variation, while low-frequency units show large delays (phase slopes) and large delay variation. The range of group delay is a well defined function of CF. Occasional outliers (less than five traces for cat 9), which were removed from the figure, could be identified with a number of problems, such as temporary loss of the unit, errors in the spline fits to the phase due to too few points, numerical sensitivity in computing the derivative from the fitted data, and occasional locking to the CM at very high sound pressure levels during the data collection process (e.g., due to poor recording conditions).

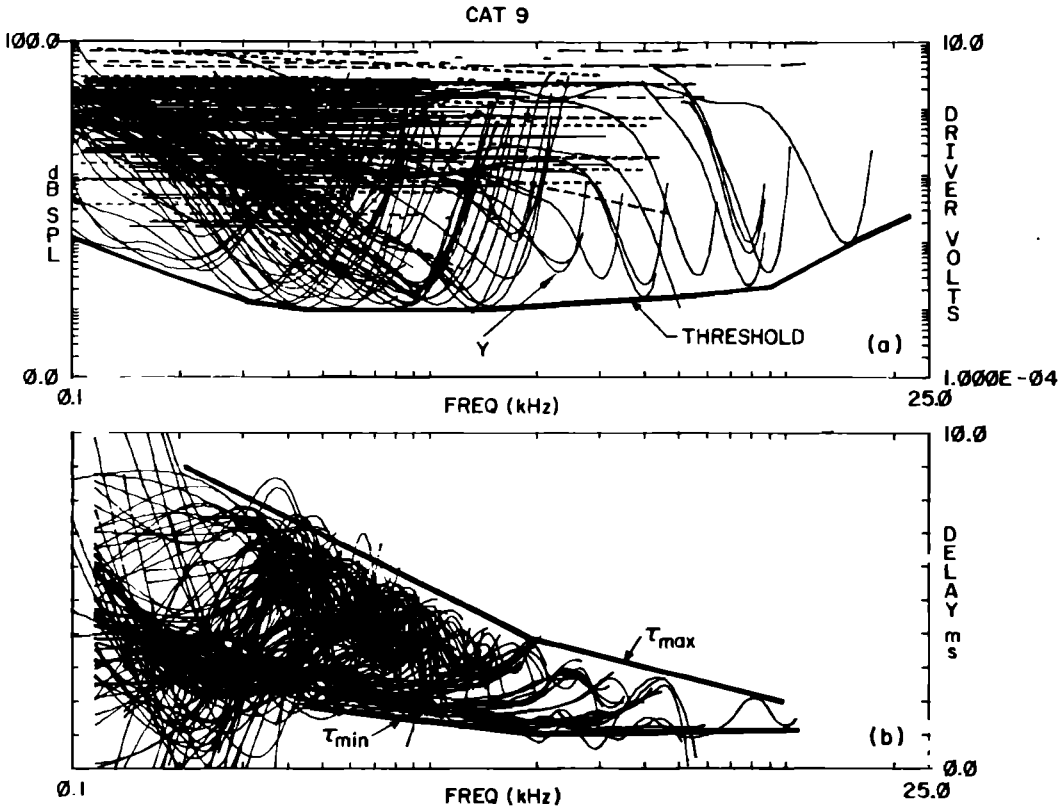

FIG. 4. (a) The upper panel of this figure shows all FTCs for animal 9. Only a few high-frequency units ($f > 2$ kHz) were measured in this animal, but the few units in this frequency range are well spaced in frequency and show normal thresholds. (b) The lower panel shows the phase slope (group delay) for all units. From this figure we see that the group delay spans a very well defined region which ranges from up to 8 ms for frequencies of 200 Hz to about 2 ms at 5 kHz. From ear canal measurements, 0.45 ms of delay is due to the acoustic driver sound delivery tube. Another 0.9 ± 0.2 ms is probably due to neural propagation delay. The remainder is cochlear delay.

In Fig. 5 we summarize the measured group delay data. The smoothed curves, as a function of frequency, are plotted with each CF range shown as a separate curve. These curves have been traced by hand from the families of group delay data, and in that sense they represent nominal values of the original family of curves.

An observation of note with regard to the data of Figs. 4 and 5 is that the group delay decreases with increasing frequency for frequencies less than $f_{CF}/2$. For example see the arrow just above 2 kHz, Fig. 5(c), for curve labeled $f_{CF} = 4.5$ kHz. This arrow is at the local minimum in the group delay for that CF range.

A. Discussion of group delay data

If one compares Rhode's measurements, cochlear model calculations, and the present neural measurements, then our neural measurements are at variance with present cochlear theory (see Sec. IIIA3 below) and Rhode's mechanical measurements, both of which show monotonically increasing group delays for frequencies below CF. We shall show that the reason for this difference is the middle ear transfer function.

1. Rhode's mechanical group delay data

Below CF Rhode's (1978) BM phase data is frequently idealized as being well approximated by two straight lines (Rhode, 1973) (on a linear frequency scale). From this description Rhode's group delay data would be constant to the first break frequency, beyond which the group delay would increase. Above CF, in the cutoff region, according to the idealization, the group delay would become zero.

In Fig. 6(a) we show two mechanically measured response curves as measured by Rhode (1978) in squirrel monkeys. In the lower right panel, Fig. 6(d), we show Rhode's group delay measurements. Note that while the group delay is somewhat noisy, it is approximately constant up to 5 kHz, and then increases up to f_{CF}.

2. Model group delay data

The neural delay data of Fig. 5 are also at odds with present one-and two-dimensional theories of basilar membrane motion. Cochlear theory stipulates that for frequencies sufficiently below CF (e.g., 0.5–1.0 oct) the group delay is independent of frequency, and only depends on CF (or equivalently place). Applying the well-known WKB approximation method to a one-dimensional transmission-line model for frequencies below CF (e.g., Zweig *et al.*, 1976; Allen, 1979).

$$\tau_g(x) = \int_0^x \frac{dx}{c(x)}$$

$$c(x) = [HK(x)/2\rho]^{1/2}, \quad \omega \ll \omega_{CF},$$

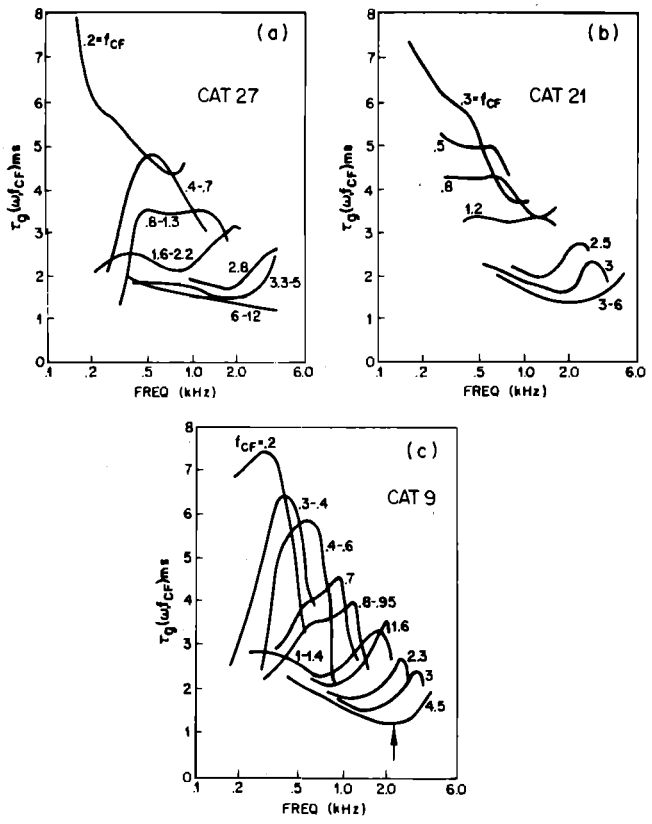

FIG. 5. In this figure we show hand traces of the phase slope (group delay) as function of frequency with CF (in kHz) as the parameter. These curves were generated by sorting the tuning data in CF groups of one octave. All the group delays for each octave group were then plotted, and a line was drawn by hand through the group of curves which seemed to best represent that octave family of delay. As is well known, low-frequency units show much greater delays than high-frequency units, with the maximum delay, for any CF range, being at or slightly above CF. Of special interest are the local minima just below CF [see arrow in 5(c), curve labeled 4.5] which is not present in cochlear measurements or in cochlear model calculations.

where $c(x)$ is the phase velocity on the basilar membrane, H is the scala height, $K(x)$ is the basilar membrane stiffness, ρ is scala fluid density, and x is the positional coordinate along the BM. If we assume, as is common, that

$$K(x) = K_0 e^{-2ax},$$

then

$$\tau_g(x) = (e^{ax} - 1)/ac(0).$$

If we define $f_{CF}(x) = [K(x)/m]^{1/2}$ [m is the BM mass and $f_{CF}(x)$ is the resonant frequency of the unloaded BM at place x] then

$$\tau_g(x) = [f_{CF}(0) - f_{CF}(x)]/ac_0 f_{CF}(x),$$

where $f_{CF}(0)$ is the maximum CF on the BM at the stapes ($x = 0$). For all cases of interest here $f_{CF}(x) \ll f_{CF}(0)$. Thus

$$\tau(x) \simeq f_{CF}(0)/[ac_0 f_{CF}(x)].$$

According to the above theory this equation is valid for group delays measured below CF. The equation indicates that the measured delay below CF is independent of frequency for a given place [constant $f_{CF}(x)$], and that the delay below f_{CF}, in the FTC tail, is inversely proportional to f_{CF}. The above equation does not account for acoustic, middle ear, and neural propagation delays, which are present in our experimental results.

The neural group delay data of Fig. 5 are in disagreement with the theory described above, and with Rhode's BM phase measurements, as shown in Figs. 6(b), and (d), because below f_{CF} the neural group delay decreases with frequency rather than being constant (or increasing). Rhode's data, and cochlear theory, are in agreement since in both cases τ_g is constant below f_{CF}. For CFs between 2–4 kHz, the magnitude of the neural delay *change* is about 1.0 ms between 200

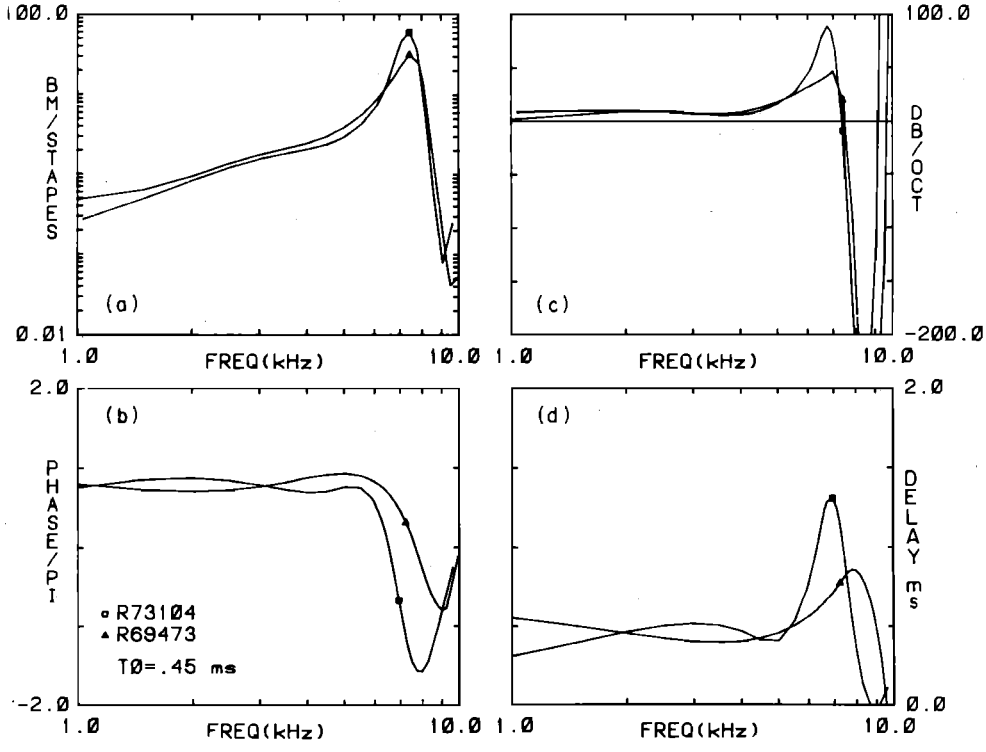

FIG. 6. (a) In the upper left panel we compare two separate mechanical response curves measure from squirrel monkey [Rhode's animals 69–473 (70 dB), 73-104 (65 dB)]. Rhode's phase for animal 69-473 is given in Zweig *et al.* Fig. 4(b). (b) In the lower left panel of the figure we show the measured phase data for each case. The phase data have been modified by the removal of 0.45 ms of delay. By modifying the basilar membrane phase in this way, the phase becomes constant for frequencies below 5 kHz. The constant phase region will be contrasted to neural data which have been similarly modified. (c) Slopes in dB/oct. (d) Group delay computed from measured phase data. Only cubic spline fits to the raw data are shown in the figures.

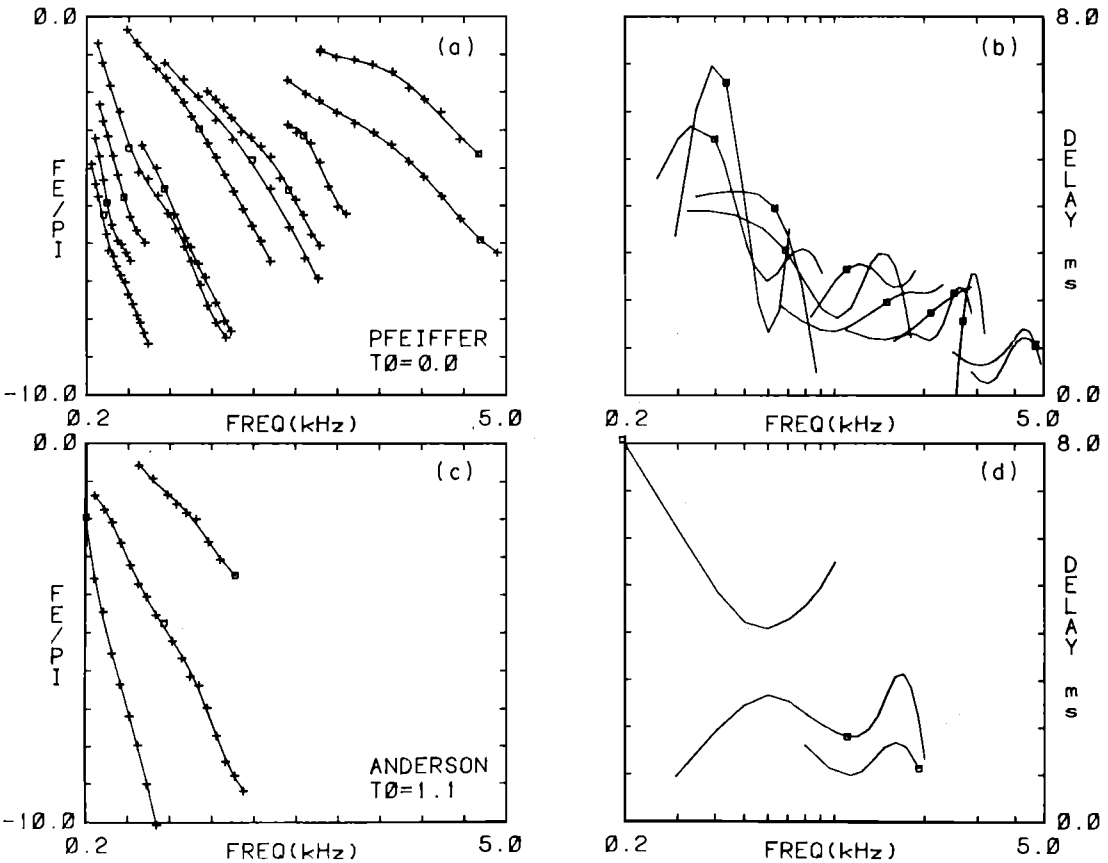

FIG. 7. (a), (b) Phase and group delay replotted from data of Pfeiffer and Molnar (1970) and (c), (d) Anderson *et al.* (1971). The square indicates the CF for each measurement. The solid lines are spline fits to the data.

Hz and $f_{CF}/2$. This change represents a significant (100%) change in delay with frequency.

3. Middle ear effects

Rhode (1978) normalized his results to the malleus displacement while we reference our results to the driver volts, which is equivalent, within a 0.45-rms delay, to ear canal pressure. Thus the middle ear transfer admittance (velocity of stapes/pressure in ear canal), loaded by the cochlear input impedance, could account for the observed difference between his observations and ours. In a later section we shall compare the neural phase to the cochlear microphonic phase and we will argue that the middle ear transfer function does account for the observed difference.

4. Group delay and phase from previous measurements

In Fig. 7(a) we replot the phase data of Pfeiffer and Molnar (1970) and in Fig. 7(c) the phase data of Anderson *et al.* (1971). Also shown is the group delay as computed from their data as found by fitting cubic splines to their data points. The group delay data of Pfeiffer and Molnar [Fig. 7(b)] and of Anderson *et al.* [Fig. 7(d)] are similar to the data of Fig. 5 since they found 7-ms delays at 300 Hz and 1–2-ms delays above 1 kHz. The square symbol in each case indicates the CF of the unit measured.

IV. PHASE DATA MEASUREMENTS

In this section we present data (Figs. 9–15) where we separate the neural phase data into approximate octave CF groups. Each plot displays *all* the phase data for the specified animal in the CF range indicated. In each CF range four plots will be given.

(a) The frequency tuning curves for all units in the CF range for which phase was measured. Superimposed are the sweep loci used in the phase measurements. The driver voltage corresponding to 10 dB *re*: 20 μPa is shown as the solid line in the lower portion of the plot once for each animal.

(b) The phase, compensated by a specific flat delay T_0 as specified in panel (d). The phase axes are in multiples of π rad.

(c) Slopes for each FTC in dB/oct.

(d) The group delay [Eq. (2)]. The dashed line indicates the delay compensation T_0 for panel (b) above. The numerical value of T_0 is also indicated in the upper right corner of the figure. The group delay plot is independent of T_0 since T_0 is not subtracted but is just indicated for reference on the delay plot.

The scale on each group delay curve has been chosen independently, however the plot ordinate range is always 5 ms. The frequencies at which the delay compensation T_0 intersects the group delay curve correspond, by definition, to a minimum (or a maximum) in the corresponding delay compensated phase curve plot since these zero points are inflec-

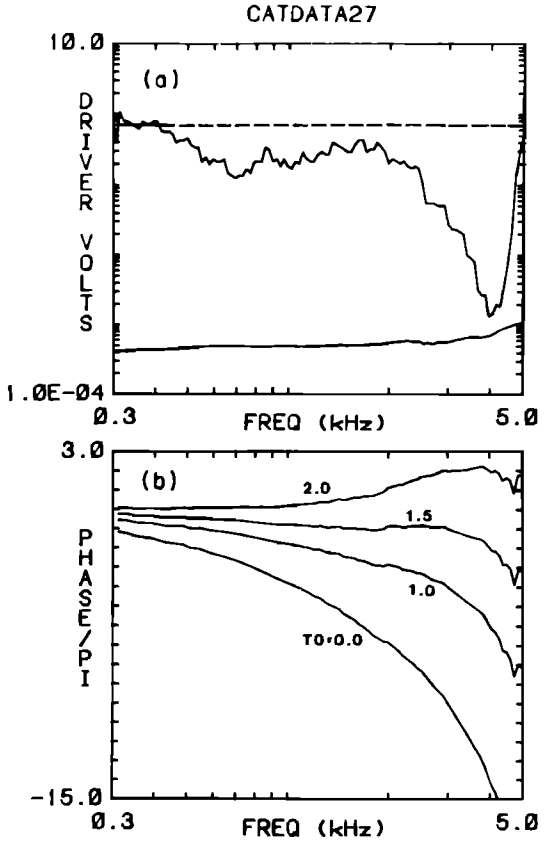

FIG. 8. Raw phase data for several different T_0 values showing the effect of various T_0 values on the plotted phase. The dashed line in (a) shows the locus of the sweep amplitudes and frequencies over which the phase was measured.

tion points of the compensated group delay $\phi(\omega) = \int^{\omega}[\tau_g(\omega) - T_0]d\omega$. Thus the shape of the compensated phase plot will depend on the choice of T_0. In Fig. 8 we show the effect of several T_0 values on a single phase response. By the proper choice of T_0, it is possible to keep the phase contained within a small range of values (e.g., $T_0 = 2$ ms).

In the first figure of the series, Fig. 9, we show all units for animal 9 having CFs between 176 and 293 Hz. This animal was somewhat unusual in that the low-frequency units were quite sharply tuned in comparison with many other animals. As described, the phase shown here has been compensated for $T_0 = 5.5$ ms of delay. Note that the compensated phase lies between 0 and -2π rad. The variations in the phase are largely due to level dependent effects. Figure 9(c) shows the FTC slopes in dB/oct.

In Fig. 10 we show all units for cat 9 having CFs between 321 and 399 Hz. Again, for this CF range, these units are more sharply tuned than those for most animals measured. Most animals have a shallow FTC slope above CF in this CF range as may be seen in Fig. 2 (animal 47) for CFs between 0.4–1.0 kHz and in Fig. 12(c) (0.7–1.2 kHz). For Fig. 10(c) a shallow FTC slope region exists over a limited region between 0.4–0.7 kHz. This effect is easily seen in the FTC slope plot but is not readily identified from the FTC data directly. The group delay data drops very sharply to small (and even negative) values below 200 Hz. Perhaps at this frequency the basilar membrane has its lowest resonant frequency and below 200 Hz traveling waves cease. Because of the decrease in the group delay, the delay compensated phase shows a positive slope (when $\tau_g < T_0$).

In Fig. 11 we show all data, again from cat 9, for CFs in the range of 404–666 Hz. In this CF range the group delay minimum [Fig. 10(d)] at 200 Hz is very pronounced. Again, due to the minimum, the delay compensated phase (b) has a positive slope for frequencies near 200 Hz.

Figure 12 shows all units for animal 27 having CFs between 433 and 713 Hz. Note that this is the same octave CF range as the previous curve Fig. 11 which was for cat 9. The largest slopes in the FTCs are all in the range of 1.1–2 kHz. The FTC slopes immediately above CF (0.7–1.1 kHz) are between 20 and 50 dB/oct, modest slopes for CFs in this range. These slopes should be compared to the previous figure where the slopes above CF immediately become as large as 200 dB/oct. Based on these two figures, and similar data in other animals, in this CF range there seems to be an animal correlated dependence between the CF and the frequency of the high slope region above CF. The phase [Fig. 12(b)] is fairly smooth about a flat delay T_0 of 4 ms.

A most interesting feature of Fig. 12(d) is the large variation in the group delay for frequencies below 400 Hz. Above 400 Hz the delay is tightly clustered. Below 400, the delay data scatters. A good many of the delay curves become negative in this region, indicative of a well-defined spectral zero in the transfer function. For animal 9, the delay is highly variable in a similar manner [Fig. 11(d)] below 200 or 300 Hz. These delays below 400 Hz were highly level dependent.

Such highly variable phase behavior seems consistent with the concept of the cochlear traveling wave having reached the end of the basilar membrane, giving rise to reflections, and thus standing waves (which would appear as poles and zeros in the response). From unpublished model studies, it seems possible that reflections might occur for frequencies below the lowest resonance of the BM. The observed variable group delay data below 400 Hz might therefore be due to apex reflections.

Figure 13 shows all phase units for animal 27 between 828 and 1511 Hz. For these units the CF has approached the frequency region where sharp high-frequency slopes occur in this animal. The slopes of 30 dB/oct for frequencies below CF are still quite modest. The phase [Fig. 13(b)] is more nearly a constant when 3 ms of delay is removed than the phase of previous figures. Note the large variability in phase slope [Fig. 13(d)] below 400–200 Hz as noted in the previous figure.

In Fig. 14 we present data (cat 27) for all units in the 1.6–3.2-kHz range. These units are in the transition region between broadly and sharply tuned responses. Units F and G [Fig. 14(a)] are beginning to show "tails," and thus steep slopes in their tuning curves below CF. Note the maximum negative slopes [Fig. 14(c)] for these two units of about -70 dB/oct. The group delay [Fig. 14(d)] shows a local minimum at a frequency just below CF. For a delay correction of $T_0 = 2.1$ ms, $\tau_g - T_0$ becomes negative between 1–2 kHz.

In Fig. 15 we show units for animal 27 between 3.3 kHz and 4.9 kHz (phase locked neural response is lost abruptly above 4–5 kHz). For units in this CF range, FTC slopes [Fig.

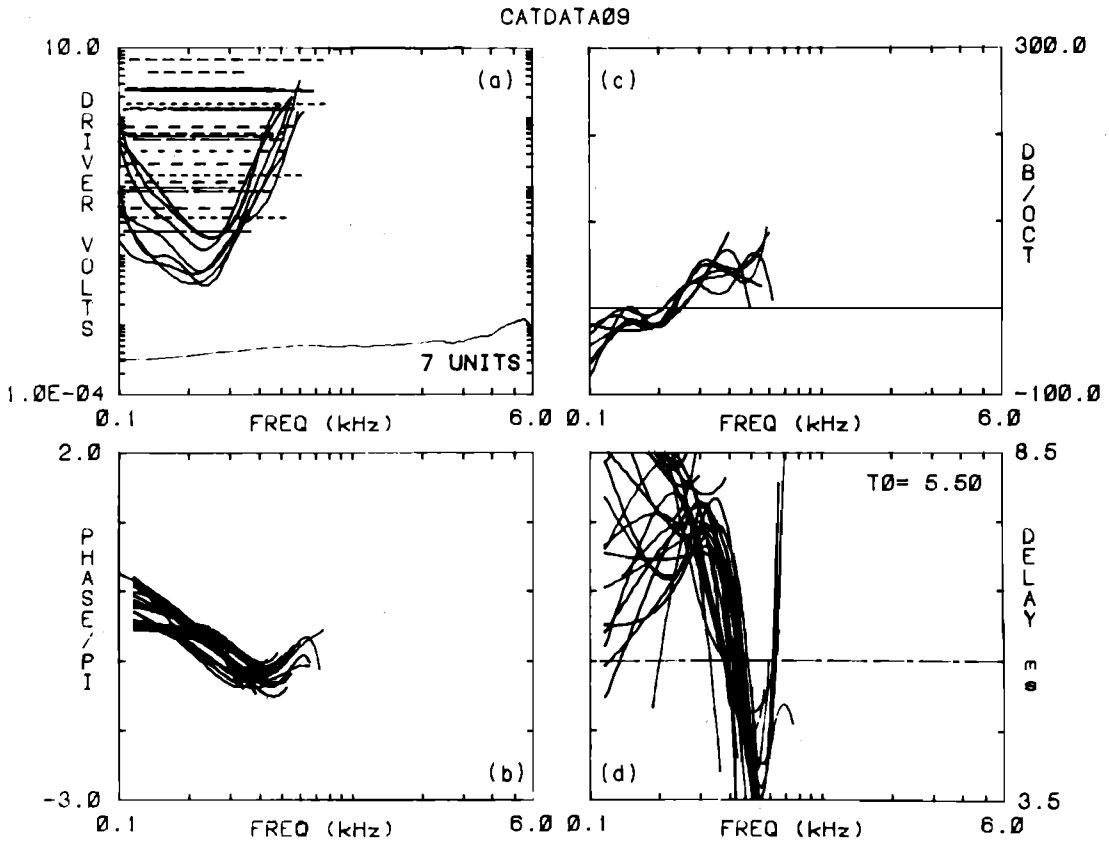

FIG. 9. In the sequence of Figs. 9–15 we divide the neural data into one octave frequency bands according to the CFs of the units. Data are shown for two different animals, (Figs. 9, 10, 11 are cat 9, and Figs. 12, 13, 14, 15 are cat 27). In each figure four panels are shown which display (a) tuning curves, (b) delay compensated phase (in units of π rad), (c) tuning curve slopes, and (d) group delay. The tuning curves (a) are from all units measured for the animal in the specific CF interval. Only spline smoothed data are shown. The dashed lines through the tuning curves indicate sweeps for which phase of (b) was measured. The measured phase has been delay compensated by delay T_0 as specified in the upper right-hand corner of the group delay plot (d). The value of T_0 is also indicated by a horizontal dashed line across the group delay plot (d). All phase data (b) have been modified by the same delay T_0, but T_0 varies with the CF octave range. Only phase deviations from the constant delay T_0 are shown in (b). When the group delay $\tau_g(\omega)$ (d) falls below T_0, the compensated phase slope, by definition, must be positive.

15(c)] above and below CF are large. Below CF, slopes approach -100 dB/oct, while above CF slopes surpass 300 dB/oct. The group delay data [Fig. 15(d)] for this CF range have a distinct (but shallow) local minimum near 2 kHz. Thus the delay compensated phase [Fig. 15(b)] shows a positive inflection for frequencies where $\tau_g(\omega) - T_0$ is negative.

A. Discussion

Kim *et al.* (1980) have measured neural phase for a few fixed tones over a large number of units, and by use of the cochlear map, and knowing the CF of each unit, they have (approximately) mapped out the phase along the BM for each of the tones. Thus Kim *et al.* measured the phase response as a function of place [$\log(f_{CF})$] for a given tone amplitude and frequency, while we measured over level and frequency for a given place (fixed f_{CF}). The Kim *et al.* data differ from ours in that they found very distinct π phase shifts about 5.0 mm basal to the CF place of the exciting tone, while we find either no such phase shifts or more gradual π phase shifts, one octave below CF (see Fig. 15). Thus while the Kim *et al.* π phase shift seems to be qualitatively in agreement with the present data, the shapes of the phase curves seem to be different. One important measurement difference between the Kim *et al.* data and the data presented here is that they used 45 dB SPL tones and averaged for 40 s, while most of our data were taken between 60–90 dB SPL and we averaged (in forming the PST histogram) for 2 s. At present we have no explanation for the apparent differences between the two sets of results.

Based on the similarities between the group delay calculated from both the Pfeiffer and Molnar data and from the Anderson *et al.* data (Fig. 7), their phase data seem to be consistent with the data presented here.

V. NEURAL PHASE REFERENCED TO COCHLEAR MICROPHONIC PHASE

Pfeiffer and Molnar (1970) compared the cochlear microphonic phase from three different turns of the cochlea to the single-unit phase. From this comparison they found that the CM and single-unit phase were somewhat similar.

In Fig. 16(a) we show the CM as recorded on the round window membrane, normalized by the ear canal pressure. The ordinate is in millivolts/Pascal. Pure tones were used in making these measurements. The signal value plotted is the magnitude of the Fourier component of the CM at the stimulus frequency. If the cochlear microphonic were indepen-

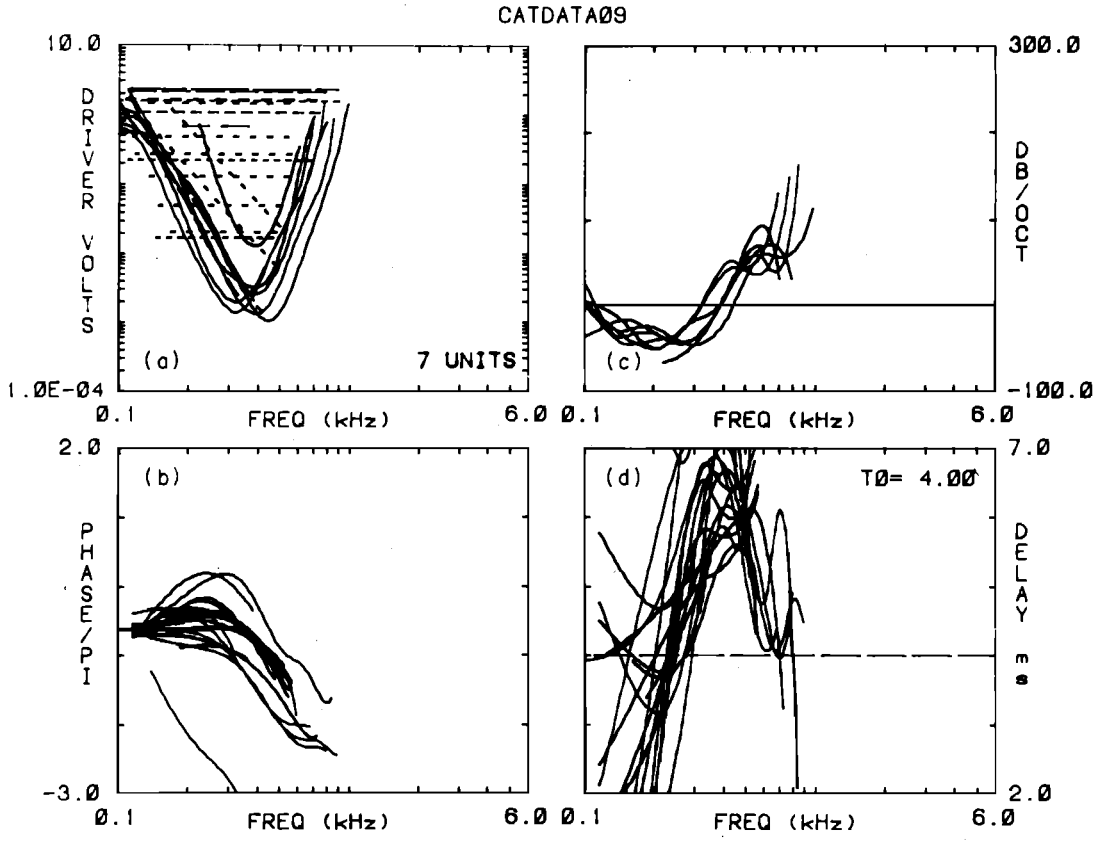

FIG. 10. (a) Magnitude FTC plot, (b) phase, (c) FTC slopes, and (d) group delay with $T_0 = 4$ ms for cat 9. This data represent *all* units for this animal having (as estimated from the raw data, not the spline fit) between 300 and 400 Hz. The FTC response (a),(c) is more sharply tuned than that for most other animals in this frequency range.

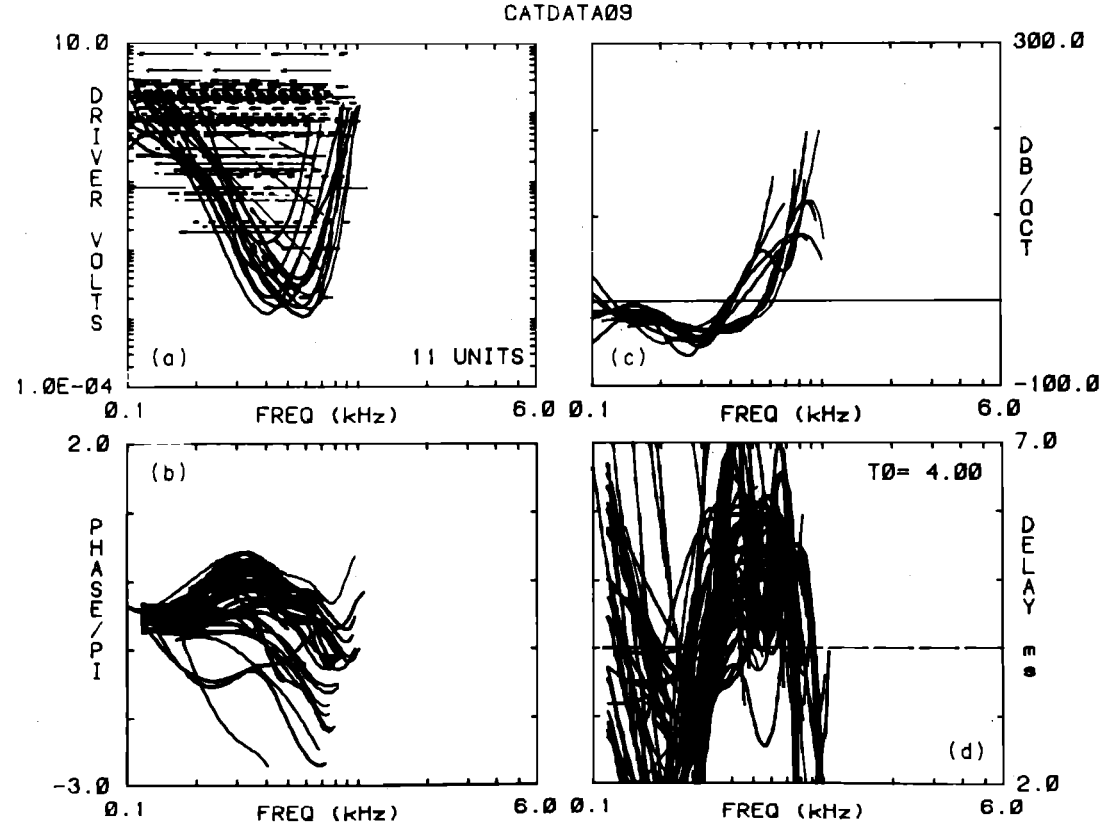

FIG. 11. The animal presented here is cat 9 and the CF range is between 400–800 Hz. As in the previous figure the tuning response (a) for this animal is sharper than that for other animals. (b) Note the sharply defined leading π phase shift in the delay compensated phase response ($T_0 = 4.0$ ms). (d) Note the regions between 150 and 300 Hz where the group delay is much smaller than at other frequencies. This region of small delay has moved up in frequency with the CF, in comparison to Fig. 10(d) for example. (c) The absolute FTC slopes for this animal in this CF range are large compared with many other animals.

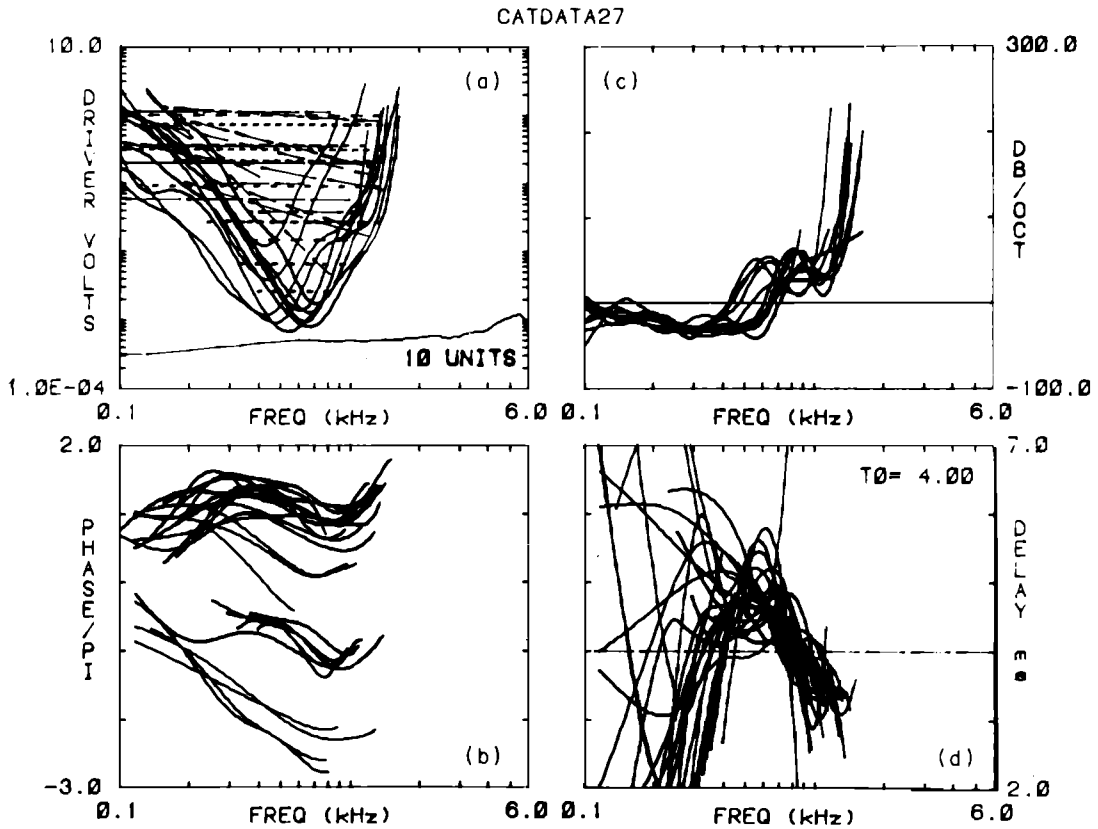

FIG. 12. These data are for animal 27 and are from the same CF range as the previous figure, 400–800 Hz. (a) The tuning curves are not as sharply tuned as those of Fig. 11(a), and the phase shifts below CF [Fig. 12(b), 0.2–0.3 kHz] are not well defined. From this comparison [Figs. 11(a), (b) and 12(a),(b)] there seems to be a correlation between the sharpness of tuning and the leading π phase jump below CF. Such a correlation would be expected from the spectral zero model of transduction (Allen, 1980). (c) Note the region ($f = 0.7$–1.2 kHz) where the FTC slopes are shallow (e.g., 25 dB/oct). Above 1.2 kHz the slopes become quite steep (200 dB/oct).

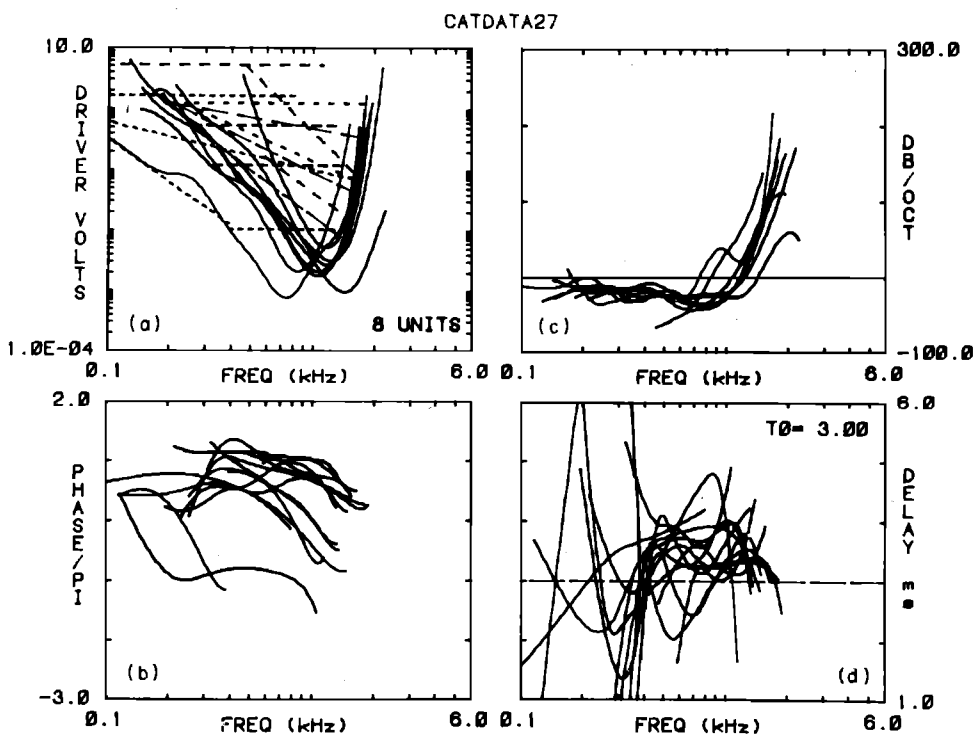

FIG. 13. (a) FTC tuning, (b) phase, (c) FTC slopes, and (d) group delay for animal 27 in the 0.8–1.6 kHz CF range.

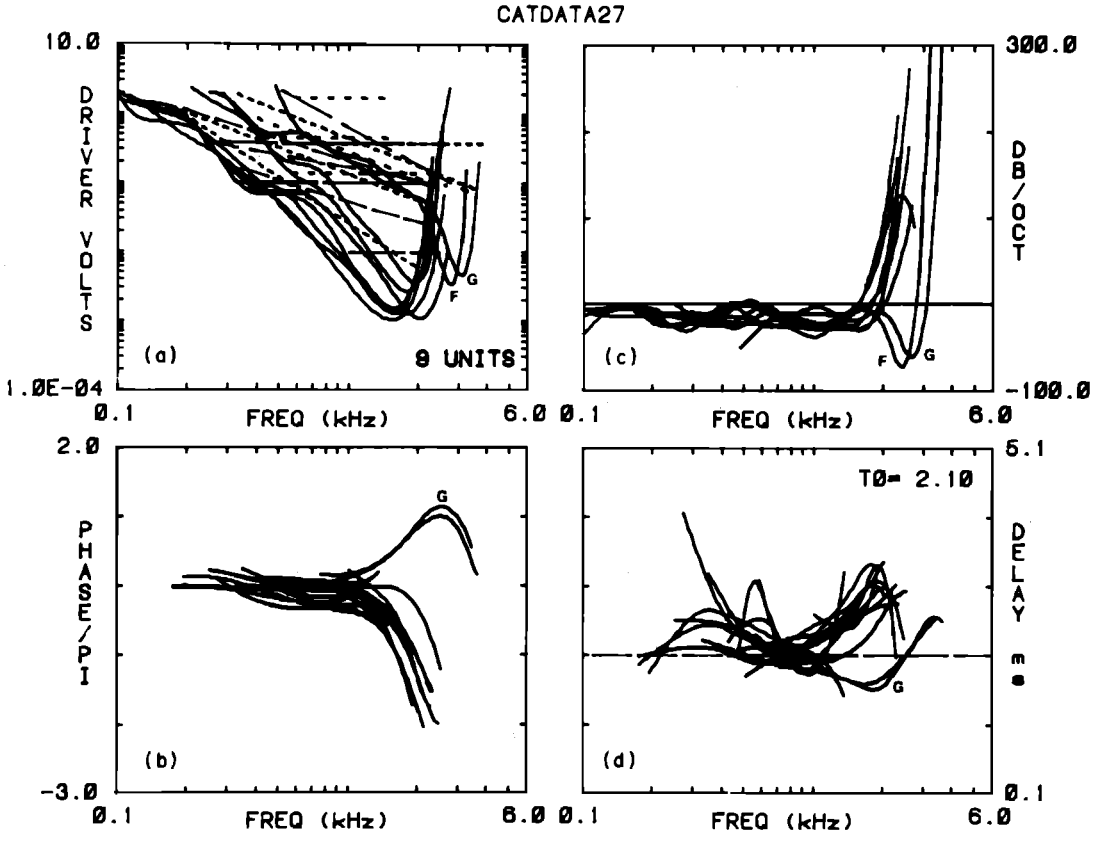

FIG. 14. (a) FTC tuning, (b) phase, (c) FTC slopes and (d) group delay for animal 27 in the 1.6–3.2 kHz CF range. The two units labeled (F,G) which have the highest CFs for this group [Fig. 14(a)] are much more sharply tuned than the other units of the group. For one of these units we obtain phase data [labeled G, Fig. 14(a)]. Note that the most sharply tuned unit shows the most clearly defined π leading phase shifts below CF of all the units.

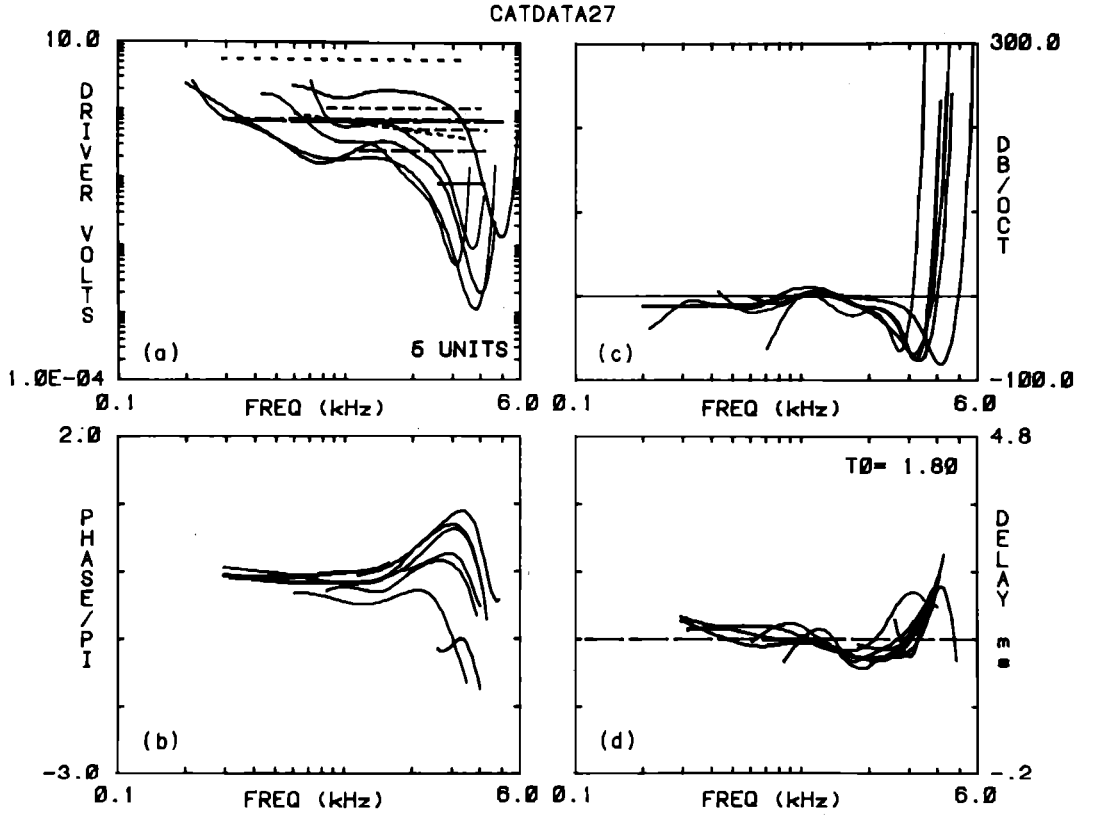

FIG. 15. (a) Tuning, (b) phase, (c) tuning slope, and (d) group delay for all units of animal 27 between 3.2 and 6 kHz, and (d) group delay. All units show (a) sharp tuning and (b) leading π phase shifts one octave below CF. The group delay (d) shows a shallow but clear minimum in the 2-kHz frequency region. FTC slopes (c) are off scale at 300 dB/oct and approach − 100 dB/oct below CF.

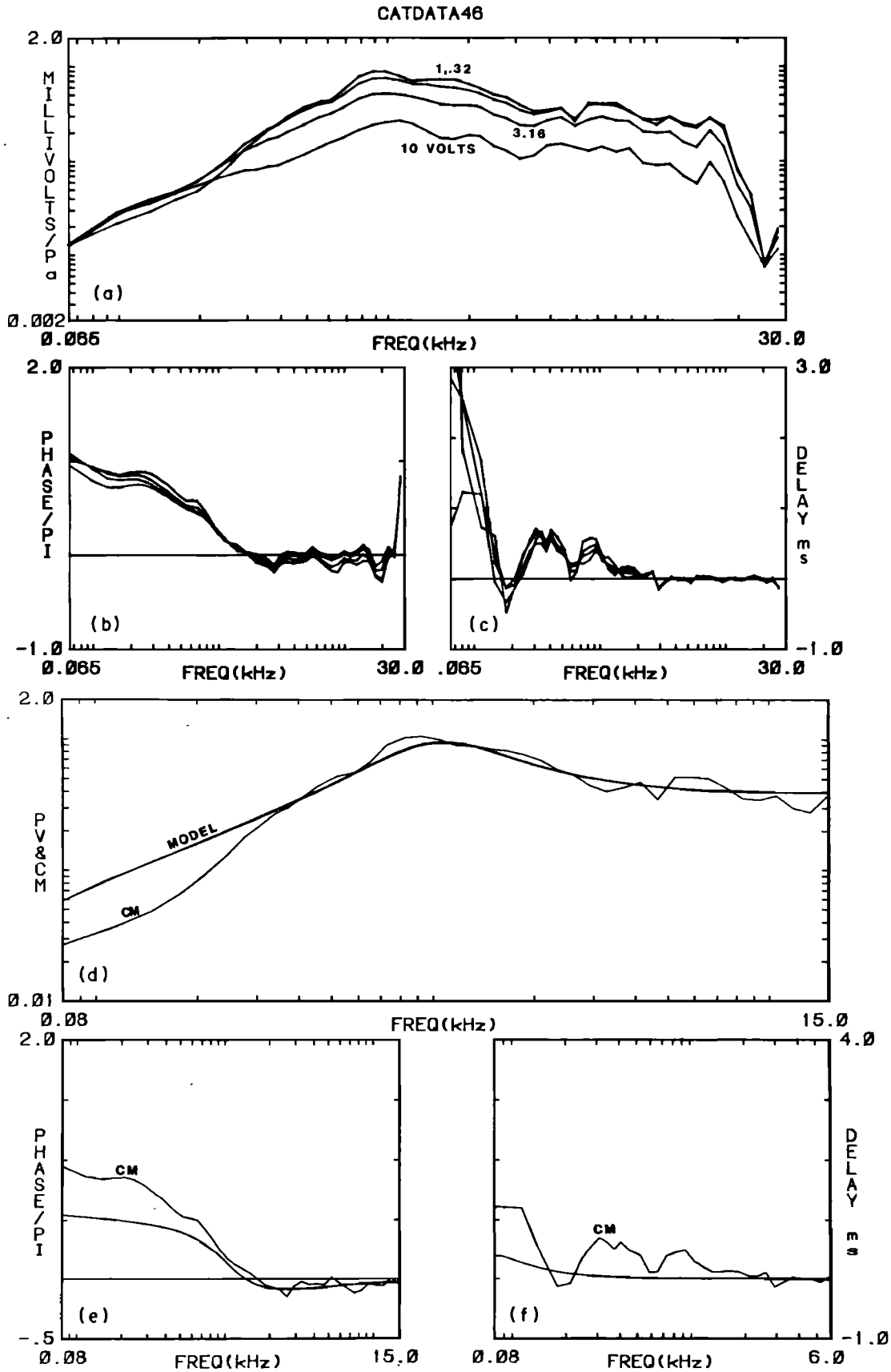

FIG. 16. Round window CM magnitude (a), phase (b), and group delay (c) after normalization by ear canal pressure. (d) Using the middle ear model of Lynch *et al.* (1982) we modeled the CM by assuming that it is proportional to the BM displacement at the base. In Fig. 16(d) we show the model CM response compared to the measured data. In 16(e), (f) we compare the phase and group delay. The parameters used in the Lynch *et al.* model (see Lynch *et al.*, 1982, p. 123, Fig. 23) were: $C_{AL} = 0.43 \times 10^{-9}$ cm^5/dyn, $M_t = 35.0$ g/cm^4, $R_{AL} = 0.05 \times 10^6$ dyn-s/cm^5, $M_V = 22.0$ g/cm^4, $R_c = 0.3 \times 10^6$ dyn-s/cm^5, $R_0 = 0.28 \times 10^6$ dyn-s/cm^5, $M_0 = 2250.0$ g/cm^4, $C_{RW} = 10^{-8}$ cm^5/dyn. Lynch's choice of M_s was 3.3. We assume that the larger value of M_s used here is necessary because the middle ear bones are intact for our CM measurements. In Lynch's experiments only the stapes remained after the surgical preparation.

dent of level then each of the curves, taken at 10-dB SPL intervals, would superimpose. However, because of the CM saturation effect, after normalization, the CM data taken at higher levels is smaller as a result of its being divided by the larger ear canal pressure. Below about 80 dB *re*: 20 μPa, the CM (*re*: ear canal pressure) becomes level independent. The lowest curve in Fig. 16(a) corresponds to 10 V into the driver, or about 100 dB *re*: 20 μPa sound pressure.

In Fig. 16(b) we show the phase of the CM at the driven frequency, and in Fig. 16(c) the group delay is shown. One feature of these data which is beyond our present understanding is the oscillation seen in the group delay at, and below, 1 kHz. This feature seems to indicate a weak reflection in the third turn of the cochlea.

From Fig. 16(b) we see that the CM phase is close to zero for frequencies above 2.0 kHz. Correspondingly the group delay is near zero in this frequency range. We interpret this phase and delay data as a demonstration of the fact that the electrode pickup must be local. If the electrode were summing voltages from points other than just the first few millimeters along the cochlea, then the cochlear delays would affect the round window CM phase response for frequencies above 1 kHz.

A. Cochlear microphonic model

Our interpretation of this data may be summarized by the following model. We assume that the microphonic signal generators are proportional to the basilar membrane displacement, presumably through shearing of the outer hair cells. (The exact nature of this generation mechanism is unimportant for the present discussion.) We also assume that only a few millimeters of the basilar membrane contribute to the CM signal at the round window.

In order to check the above assumptions we have calculated the basilar membrane displacement at the stapes by use of the middle ear model of Lynch *et al.* (1982). Several of the middle ear parameters were slightly changed from Lynch's values to improve the fit to the CM data. Because the ossicles were in place, the stapes mass was increased from 3.3 to 35 g/cm^4. From the model one may calculate the pressure across the basilar membrane at the stapes (base), relative to the ear canal pressure. The resonant frequency of the basilar membrane in the base is above 10 kHz. For frequencies less than 10 kHz we assume that the basilar membrane impedance near the stapes is stiffness dominated. Thus the BM displacement near the stapes is proportional to the pressure across it (by Hook's law). Using the Lynch *et al.* model we calculate the pressure at the stapes across the basilar membrane, which is proportional to the BM displacement, and therefore to the cochlear microphonic. In Figs. 16(d), (e), and (f) we compare round window CM magnitude, phase, and group delay data to the model. In general the fit is good for frequencies above 500 Hz. The main features which do not fit are the large delay peaks seen in the CM group delay below 1 kHz. Given the simple nature of the model cochlear input impedance of Lynch *et al.*, we would not expect this feature to be present in the calculated group delay.

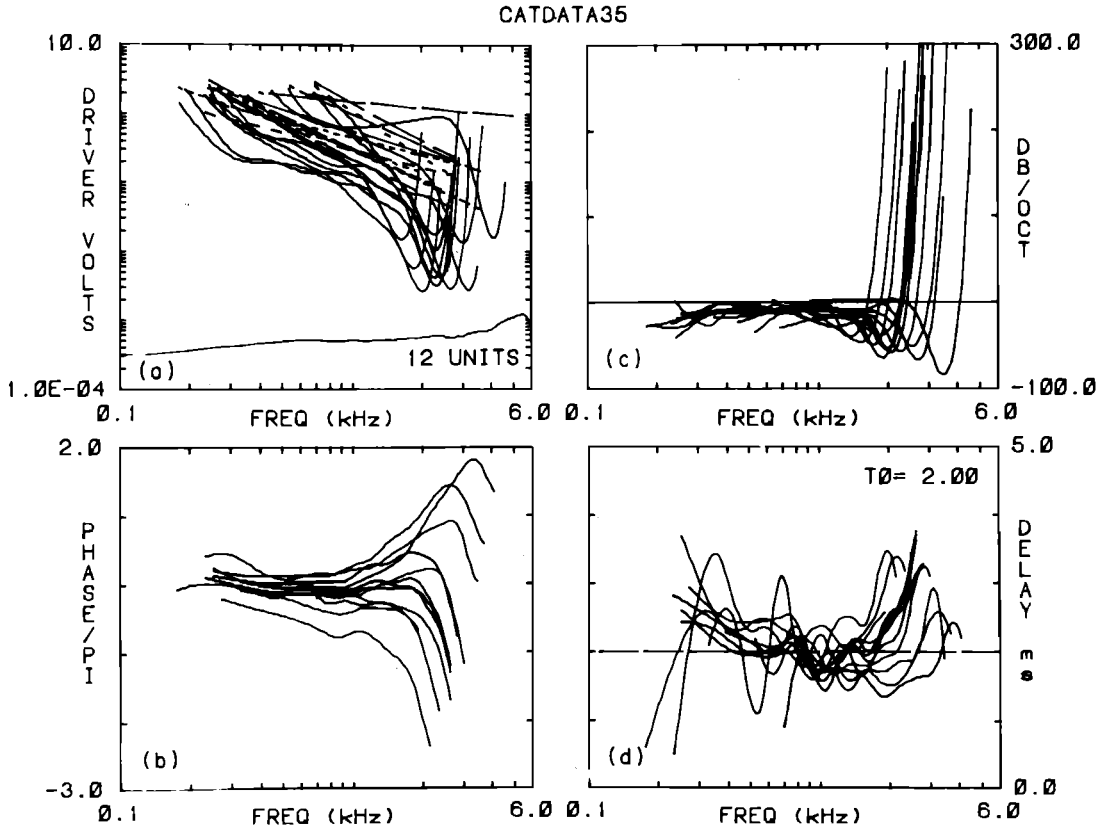

FIG. 17. Magnitude (a), phase (b), slope (c), and group delay (d) data from cat 35. For the phase data, the delay compensation was 2 ms. This figure should be compared to Fig. 18 where the same phase and magnitude data have been normalized by the CM magnitude and phase.

In our model calculation we did not include the incudo-malleolar joint or the middle ear and bulla air cavities. Because the incudo-malleolar joint has been ignored, we would expect the model to deviate from measurements near and above 10 kHz.

B. CM normalization

Armed with this interpretation of the round window CM data, we may now interpret the neural phase, referenced to the CM, as the neural phase referenced to the basilar membrane displacement at the stapes. This normalization is convenient for many reasons, the most important here being that it removes the middle ear and driver transfer functions from the results. It would also remove middle ear nonlinear effects.

In Fig. 17 we present some neural data from cat 35 for all units which had CFs between 2.5 and 4.0 kHz. The phase data has been compensated for 2.0 ms of delay. Note the leading phase shifts in the 1–3-kHz region of Fig. 17(b). Note also the gradual decrease in the group delay between 300 Hz and 1.5 kHz.

In Fig. 18 we show the effect of normalizing the neural data by CM magnitude and phase. In Figs. 18(a) and (c) we see that the slopes are near zero below 800 Hz due to the normalization. In Figs. 18(b) and (d) the leading phase shift is gone and the group delay is a constant 1.2 ms below about 1.4 kHz. These phase plots should be compared to those of Rhode, Fig. 6(b). It is presently unclear what the effect of CM normalization would be for low CF units ($f_{CF} < 1$ kHz) since for low frequencies both the CM and the neural phase are quite level dependent. We did not attempt to account for level effects here since we used only one nominal CM curve for all neural normalization. It would seem that the proper way to do the normalization is to measure the CM simultaneously with the neural phase so that common level-dependent effects would cancel.

C. Discussion

The correction of the neural phase by the CM phase removes middle ear transfer function effects from the neural phase and group delay. After being corrected, the neural phase is qualitatively similar to Rhode's measured phase of basilar membrane motion.

These data also make certain "second filter" models as being unlikely explanations of cochlear sharpening, such as the model proposed by this author (Allen, 1980), or other two-component cancellation models (Zwislocki and Sokolich, 1974), since such models would have a phase behavior quite unlike that found here. This point is one of the main conclusions of this paper since it clearly rules out the "second filter" model proposed by this author (Allen, 1980).

VI. LEVEL-DEPENDENT PHASE EFFECTS

In this section we summarize numerous nonlinear (NL) effects seen in the data. A phase measurement is defined here as being nonlinear when it depends on the sound level of the

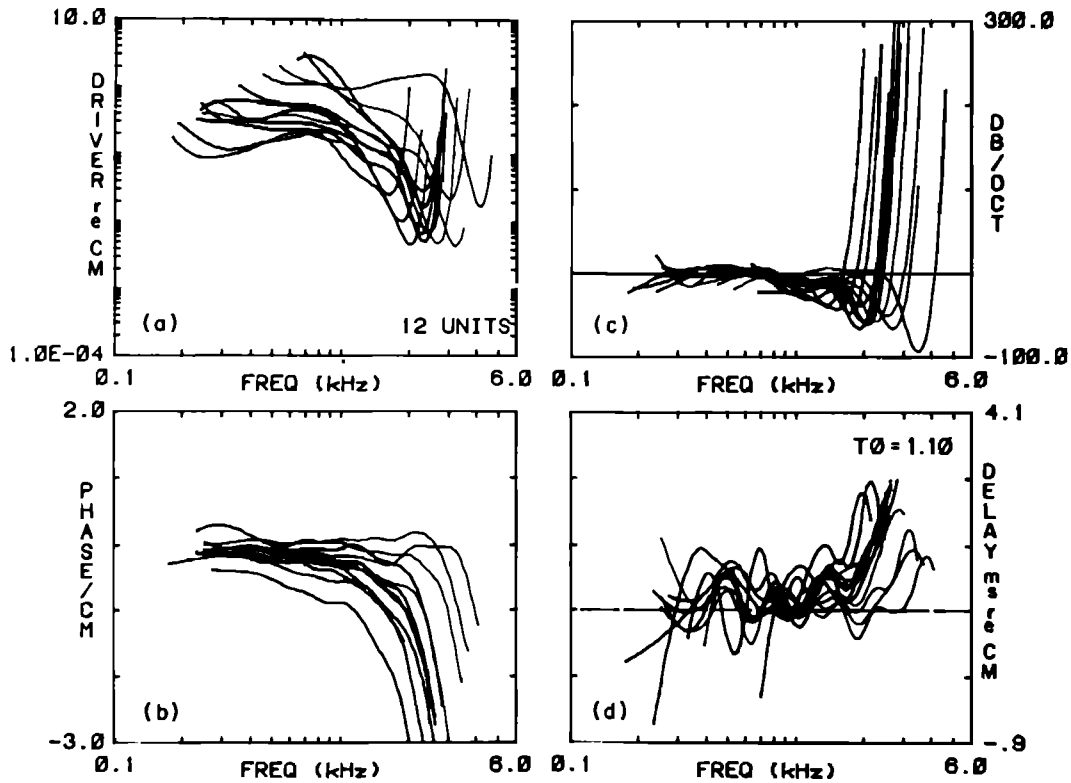

FIG. 18. These data are the same as the previous figure. However in this case we have normalized the data by the CM as measured at the round window membrane (1 V to acoustic driver). This normalization removes the local minimum in the group delay below the CF, and therefore it also removes the region of leading phase just below CF. In this plot the delay compensation was $T_0 = 1.1$ ms. This figure is the sole figure in this paper which presents data normalized by the CM response.

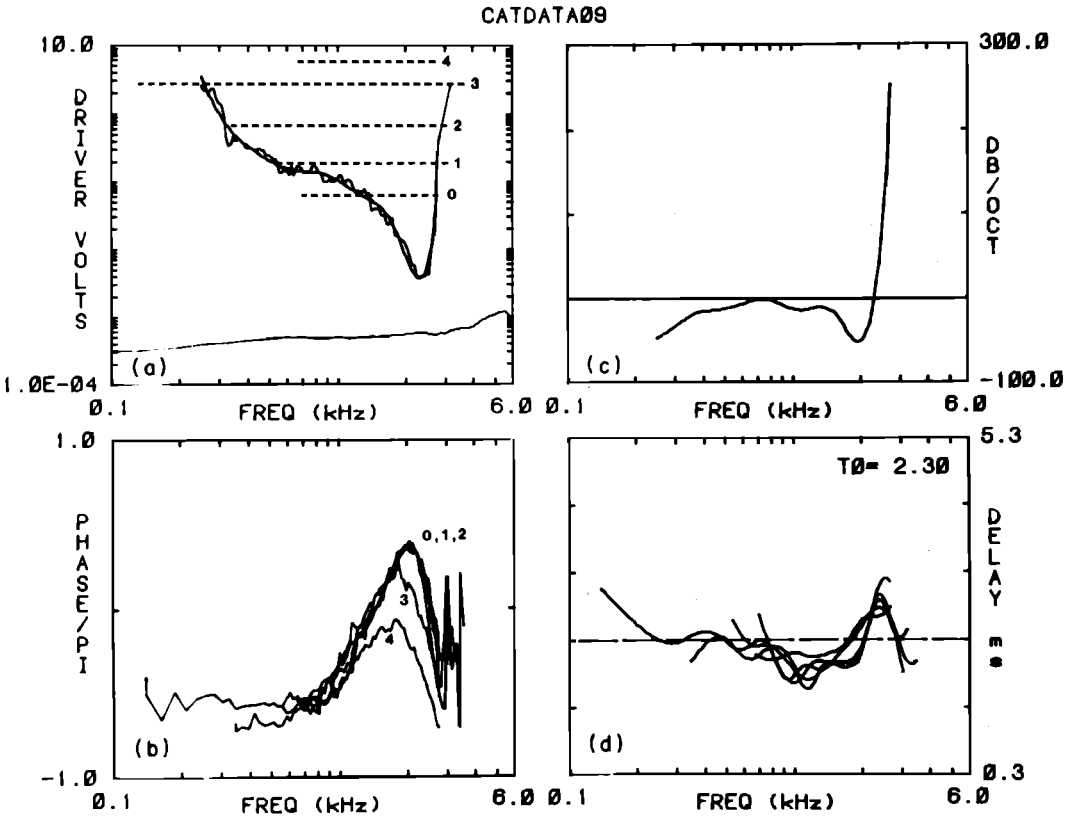

FIG. 19. We show here [animal 9, unit Y, Fig. 4(a)] the effect of sound pressure level on the phase of a neuron having a CF of 2.2 kHz. For levels below 80 dB SPL (1 V) the phase is invariant to tone level. For levels above 80 dB SPL, the sharp cutoff region shifts down in frequency, as estimated from the phase (b) in the frequency region near CF having the largest group delay (indicated by the phase curves labeled 3 and 4). For Fig. 19(a) the raw FTC data are superimposed on the spline fit. In 19(b) individual phase data points have been connected by straight lines, giving an indication of the density of frequency points and the variability of the phase measurement. As before, the phase ordinate is in terms of π rad. The two right-hand panels (c),(d) are slopes, as calculated by spline fits (not shown) to the corresponding raw data. The nonlinear effect of the CF shifting down in frequency with increasing SPL has frequently been noted previously under a number of very different experimental conditions as discussed in the text.

stimulus tone. In general other types of cochlear nonlinear effects are known to be present, such as hair cell transducer rectification, two-tone effects, or noise masking nonlinearities. However in this set of experiments we only look at single-tone steady-state phase as a function of the independent variables frequency and sound pressure level.

All of the effects discussed have been seen many times in all "normal" animals (with the exception of the data of Fig. 22), and the examples shown are believed to be typical. The most important independent variable, besides level, is the CF of the unit. The different types of nonlinear effects are specific to the CF range of observation. The types of NL effects can roughly be divided into three CF ranges, below 1 kHz, between 1–5 kHz, and effects in the low-frequency tail below 5 kHz for units having their CF above 5 kHz.

The first, and perhaps the most interesting nonlinear effect, is the level-dependent shift of the cutoff frequency as reflected in the phase at high sound levels. The effect of level on phase is shown in Fig. 19(b) (the spline fit is not shown here, only the connected raw data points are given) where neural phase is plotted as a function of frequency ($T_0 = 2.3$ ms) for various sound pressure levels. The unit presented here is shown in Fig. 4(a) where it is labeled Y (CF = 2.3 kHz). The unit had a threshold that was elevated by 10 dB above most of the units of cat 9, and had a CF in the transition range between the broadly tuned low-frequency units and sharply tuned high CF units. According to our calibration, the unit CF had a threshold of 25 dB SPL, with the animal's overall sound level threshold, below 2.0 kHz, being 20 dB SPL. On this scale, the nonlinear effects appeared between 75 and 90 dB SPL, with 75 dB showing no effect and 90 dB showing a small effect (curve 3 is at 90 dB SPL). At 95 dB the effect of the nonlinearity is pronounced in the frequency neighborhood of the CF. The nonlinear phase effect for this unit is small below 1200 Hz. From the phase data, it appears that the sharp cutoff region (those frequencies above CF for which the FTC slope is large) has shifted down in frequency as the sound level increases above 75 dB.

For units which have CFs below 1 kHz, level-dependent effects can be very complicated. In this CF range nonlinear effects are varied and animal dependent. In Fig. 20 we see an extension of what happened in Fig. 19, namely the $\pi/2$ phase shift below CF [Fig. 20(b)] is absorbed by the large phase shift associated with the high-frequency band edge, as the level increases. In Fig. 20(a), we show both the raw FTC data as well as the spline fit.

In Fig. 21(b) we show phase (raw data only) as a function of frequency and level for three units. All three units show level-independent phase near CF. The group delay [Fig. 21(d)] decreases with increasing sound level. We call points

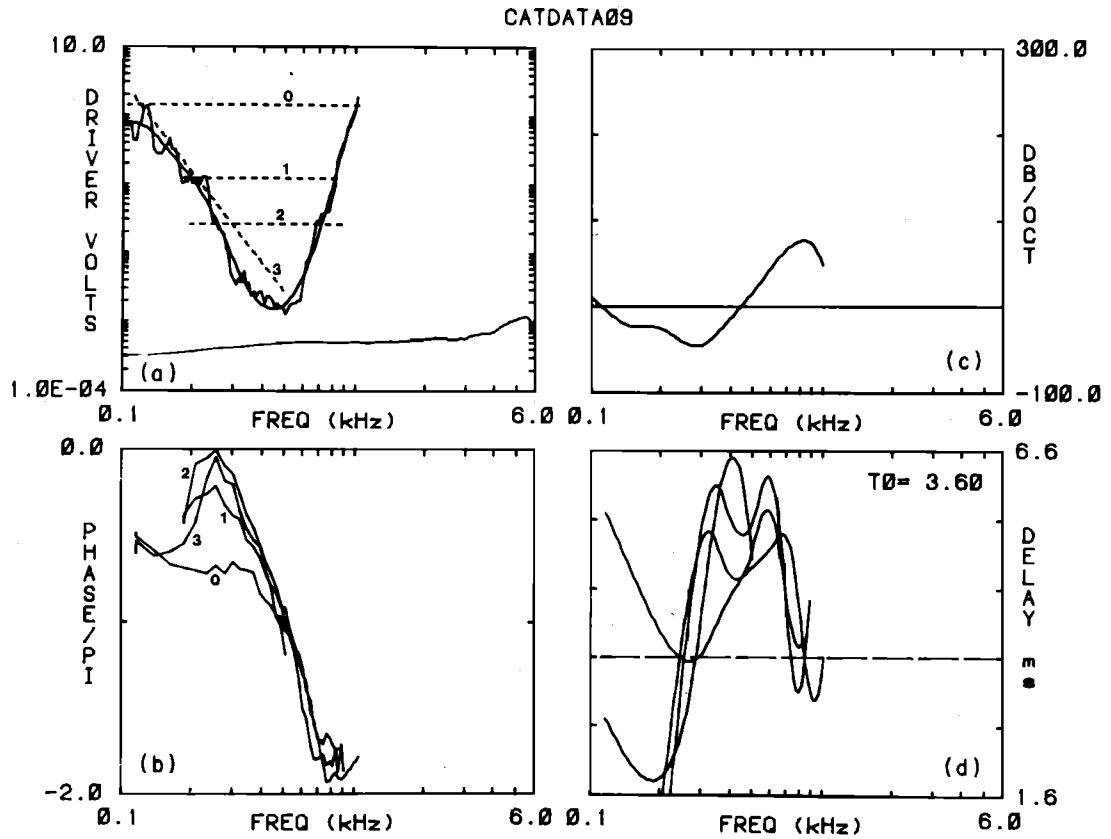

FIG. 20. The nonlinear phase effects below 1.0 kHz are much more complicated than the phase data for CFs greater than 1 kHz. In this figure we show a simple low-frequency extension of the effect presented in the previous figure. In cat 9, with its relatively sharp low-frequency tuning, this effect was quite common. It, as well as the effect of the previous figure, were the most common nonlinear effects observed across animals. While the NL effect of Fig. 19 only appeared at high sound levels, in general the NL effects below 1.0 kHz could be observed at much lower levels, as is seen in this figure.

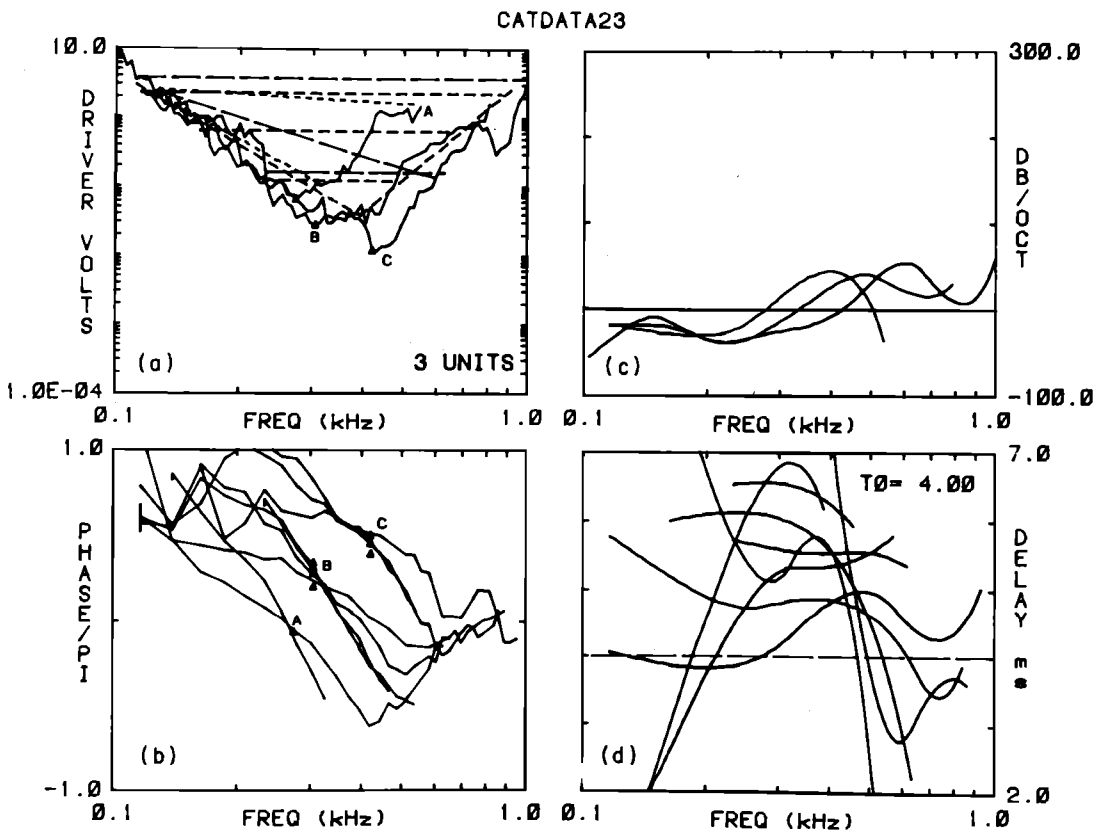

FIG. 21. It has previously been observed (Anderson *et al.*, 1971) that all phase responses meet at the CF, as a function of level. This property was found in our experiments for units having CFs below 1 or 2 kHz. For units above this range, the level independent "pivot" points no longer tracked the CF and remained below about 1 Hz. As the CF increased, the concept of a "pivot" point degenerated with the phase becoming linear (level independent) over a large range of frequencies. Note the rather random looking behavior of the group delay (d) for these three units as compared to the stylistic behavior of the phase.

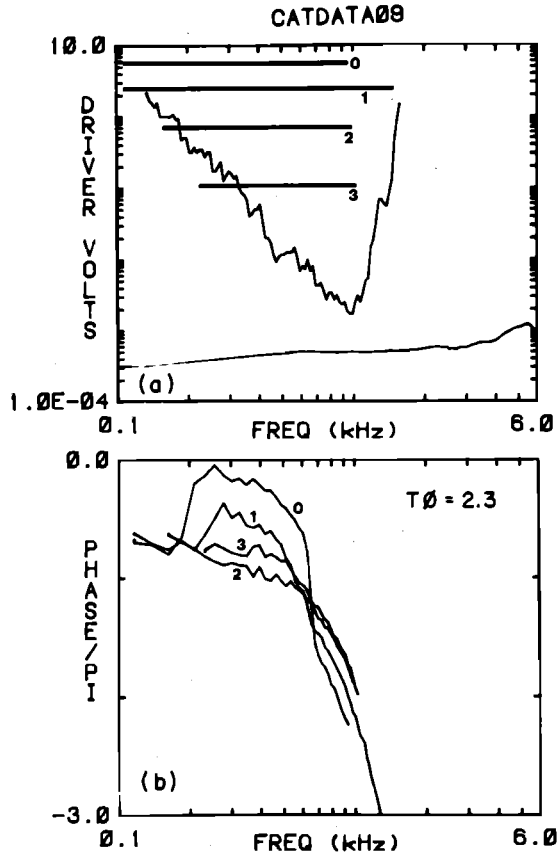

FIG. 22. The nonlinear effect seen in this figure was only seen in two animals (cats 8 and 9) very early in the series of experiments. The π phase shifts between 200 and 600 Hz for levels greater than 90 dB SPL are similar to the "peak splitting" rate-level nonlinearity observed by Kiang *et al.* (1969), which we associate with the sharp null in the rate of firing as SPL is increased. Note how the NL effect spreads beyond the CF, but does not appear there as a π phase shift. Only raw data is shown in this figure.

where a family of curves all intercept at a common point (where they are level independent) a "pivot point." In each case in this figure the CF (arrowhead) is very close to the pivot point. As the CF increases the concept of the pivot point degenerates and, when such points exist, they were well below the CF of the unit, and always below 1.5 kHz. Low-frequency pivot points were previously observed by Anderson *et al.* (1970).

In Fig. 22 we show a nonlinear effect which was observed in two animals (cats 8,9) and then only under very limited conditions (very high pressure levels, frequencies below 1.5 kHz, for units having CFs below 2 kHz). It is mentioned here only in that it seems to relate to the rate-level nonlinearity described by Kiang *et al.* (1969) (sometimes referred to as "peak splitting"). Neural phase is an excellent measure of this nonlinear effect because of the rapid change in phase as the signal level passes through the "notch." The phase effect shown in Fig. 22(b) was only seen in two animals and has not been observed since.

One nonlinear effect (Fig. 23) is *always* observed for any unit having its CF > 5 kHz, for stimulus frequencies below 5 kHz. This effect is a small ($\pi/2$) phase shift which seems to be only weakly dependent on frequency. Since this phase nonlinearity can only be measured in the "tail" region of high CF units, it may only be measured at high levels (80–90 dB SPL). The NL region spreads as the level is increased above 90 dB SPL. One possible explanation for this NL effect might include a change in the cochlear input impedance and thus a changing load on the middle ear transfer impedance.

VII. SUMMARY

In this paper we have considered the question of phase at the level of the primary auditory nerve. Measurements of

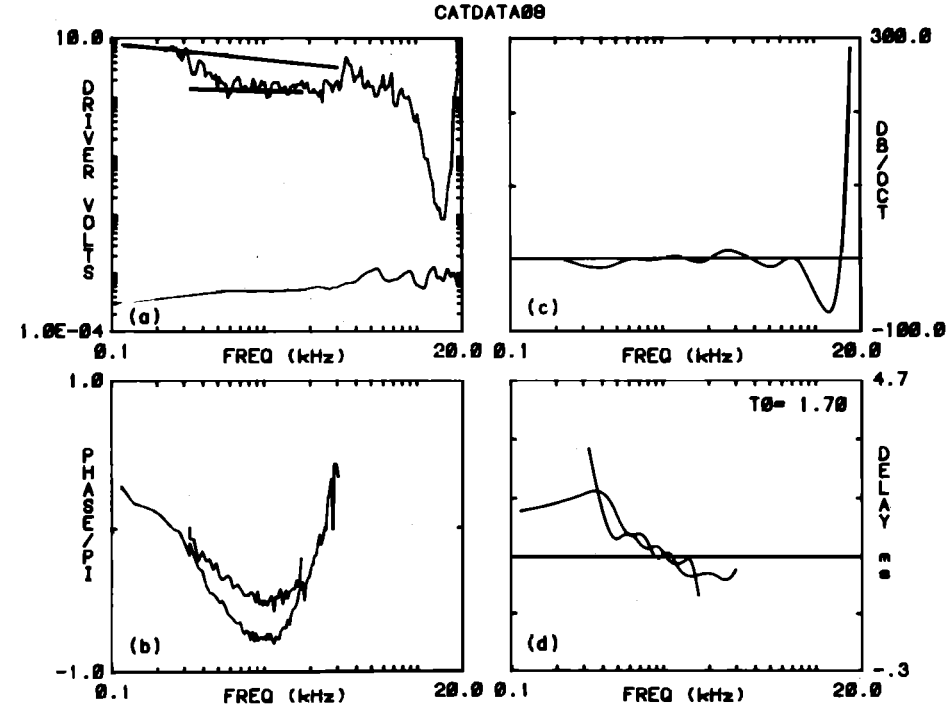

FIG. 23. One final nonlinear effect is a phase shift in the tail of high-frequency units. The effect seems to be CF independent for CF > 5 kHz. Because it can only be observed in the tail of high CF units, it may be measured only at very high sound levels.

phase locking (neural phase response, estimated by Fourier transforms of neural period histograms) as a function of frequency and level in cat primary units have been presented in some detail. As was shown, a great deal of information may be obtained from the phase locked response as estimated from the Fourier transform of the stimulus-synchronized, neural period-histogram response. Phase data were measured, as a function of frequency and level, in more than 1000 primary fibers from more than 47 mature cats.

To interpret the phase measurements three different transformations were performed on the measured data. *First*, a flat (frequency-independent) delay was subtracted from the neural phase data in order to remove the large phase roll which obscures interesting detail. *Second*, the slope of the phase was estimated providing a measure of the neural (cochlear) group delay. *Third*, neural phase was normalized by the cochlear microphonic (CM) phase measured at the round window (RW) membrane.

Four observed effects are most notable.

(a) For all high-frequency units ($f_{CF} > 1$ kHz) and for frequencies less than f_{CF}, we found a phase shift (relative to ear canal pressure) having a group delay inversely proportional to frequency. The source of this phase shift was found to be due to the cochlear input impedance loading the middle ear transfer impedance since it also appeared in the CM measurements and could be removed by referencing the neural phase to the CM phase (see Figs. 17, 18). For these units the group delay tends to sharply increase near the CF where the unit response goes into cutoff. The combination of these two effects produced a CF-dependent local minimum in the group delay about one octave below CF ($f \approx f_{CF}/2$).

(b) At stimulus levels above 75 dB SPL, for high CF's ($f_{CF} > 1$ kHz), the large phase shifts (large group delay) associated with the sharp high-frequency band edge of the cochlear filters were found to shift down in frequency with increasing level (see Fig. 19).

(c) For low-frequency units ($f_{CF} < 1$ kHz), very complicated level-dependent phase was observed (Figs. 9–13, 20–22). The CM phase measurements also displayed significant level-dependent effects below 1 kHz.

(d) For frequencies above 1 kHz and levels below 75 dB SPL, the phase was almost always independent of level (Fig. 19).

The phase measured in the auditory nerve was found to have two components. The *first* was cochlear in nature, being similar to the mechanical phase measurements of Rhode. For high CFs ($f_{CF} > 1$ kHz) this component has a constant phase slope (group delay) up to about $f_{CF}/2$, at which point the phase slope increases. For lower CFs ($f_{CF} < 1$ kHz) this picture is not as clear since the group delay changes dramatically with CF while the delay-compensated phase remains approximately constant over most of the passband of the unit.

The *second* phase component is due to the middle ear transfer function. For each animal we were able to estimate this phase component from the phase of the round window (RW) membrane potential (cochlear microphonic). We presented and discussed a simple CM model which assumes that CM generation is proportional to basilar membrane displacement, and that the potential on the round window is coupled only to those generators in the very high-frequency region of the cochlea, namely those CM generators proximal to the round window.

This model did a good job of matching the CM magnitude and phase assuming some changes in the parameters of the Lynch *et al.* middle ear model. The model did not accurately match the large oscillations found in the CM group delay below 1 kHz [Fig. 16(c)]. The reason for this seems to be due to the over-simplified form of the cochlear input impedance assumed by the Lynch *et al.* model.

When the two phase components, cochlear and middle ear, are added together, the resulting group delay decreases with increasing frequency. Since the group delay increases at the CF, there is a local minimum in the group delay just below CF. This local minimum can give rise to leading phase shifts in the delay compensated phase for certain choices of delay compensation (Fig. 17(b)].

A. Neural phase and the second filter

If the neural phase and the BM phase were shown to be similar, then the requirements of a second filter would be weakened. In these measurements we have found that the neural group delay relative to the CM is constant below CF. This result rules out two-component cancellation models (Zwislocki and Sokolich, 1974) and the spectral zero model (Allen, 1980) of cochlear sharpening.

The Kim *et al.* neural measurements showed spread of excitation along the BM to single tones. They found π phase shifts basal to the tone place. Thus their measurements of phase and our present measurements seem to be in disagreement. At present there is no obvious explanation for this difference.

B. Nonlinear phase effects

As previously mentioned, the CM group delay shows large oscillations for frequencies below 1 kHz. These standing wavelike oscillations are always observed in our CM measurements. As a function of level, the peak group delay decreases as the level is increased. Below 1 kHz, the CM phase is usually nonlinear. For high sound levels the total phase shift between 60 and 1000 Hz is less than that at low sound levels. This nonlinear effect is not well represented by the data of Fig. 16(b) which is relatively level independent. No attempt has been made to remove $N1$ from the CM data. At low levels $N1$ could be affecting these measurements.

The neural phase measurements seem to be linear (level-independent) over a very large range of levels and frequencies, namely above 1 kHz and below 70 dB SPL. Since our phase measurements necessarily stop at 4 kHz, we do not know if this range extends into the above 4 kHz range. The general trend however is that the phase shows less level dependence as the frequency is increased.

Above 70 dB SPL, for frequencies above 1 kHz, two effects were noted. First the CF seems to shift down with increased level. This nonlinear (level-dependent) effect is consistent with the concept of a sharp linear cutoff low-pass filter band edge shifting down in frequency at high sound pressure levels. Such "CF shifts" have been observed by a number of investigators at many experimental levels, for ex-

ample by Rhode (1978) in his BM velocity measurements, by Evans (1977) and Møller (1977) in their reverse correlation neural measurements, by Russell and Sellick (1978) in their receptor potential experiments, by Cody and Johnston (1980) in their ganglion cell measurements, and by Dallos (1973) in his cochlear microphonic measurements. While only speculations are possible at present, it seems that the observed shifts in phase at large sound levels are the same as the nonlinear cochlear mechanical effects discussed in the review papers of Hall (1980, 1981) and Kim *et al.* (1980). If the nonlinearity is the one modeled by Hall and Kim, then the phase data presented in this paper give important quantification of, and insight into, the mechanical nonlinearities present in the cochlea. The origin of the NL effect seems to be at the mechanical level because of the strong frequency dependence of the nonlinearity, assuming the mechanics determines the frequency dependence, as is commonly believed. This nonlinear effect was observed many times on most animals for CFs above 1 kHz at levels above 70 dB SPL. It appears to be present, to some degree, in all units in this range (namely CF > 1 kHz, SPL > 70 dB SPL). This frequency shift was never more than 1/2 octave. A second high-frequency, high level effect was seen where the phase in the tails of very high-frequency units ($f_{CF} > 5$ kHz) shifted in the leading direction, independent of CF, with increasing sound level.

For frequencies below 1 kHz the neural phase is significantly more level dependent. Below some critical frequency, which is animal dependent, the phase slope is erratic, with the group delay frequently becoming negative. One reasonable explanation for this is apical reflection of the traveling wave on the basilar membrane. The critical frequencies for these apparent reflections ranged between 100–400 Hz. For frequencies below 1 kHz, the only general trend of the phase nonlinearities seemed to be one of decreasing group delay with increasing level. Frequently, for CFs in this range, the phase near CF was level independent. This effect was reported previously by Anderson *et al.* (1971).

Finally we make an observation about the tuning curve threshold data for CFs below 1 kHz. For about half of the animals measured we found an intermediate slope region of about 50 dB/oct between the CF and the sharp high-frequency cutoff, where the slopes were greater than 100 dB/oct.

VIII. CONCLUDING REMARKS

This paper has only begun to focus on what seems to be a rich area, namely cochlear neural phase response to single frequency tones. One important question which remains unexplored is single unit phase under conditions of a second, sub-threshold suppressor tone. Since we know that single unit threshold rate response can be altered by the presence of a suppressor tone, phase, in the presence of a suppressing tone, should be studied. If, as found by Arthur (1976), phase is invariant to a rate suppressing tone, then perhaps two-tone rate suppression is not mechanical in origin. On the other hand, (assuming the mechanical elements determine the neural frequency selectivity) if two-tone suppression nonlinear phase effects change in a frequency-dependent manner in the frequency region near CF, then two-tone suppression is likely to be mechanical, having its counterpart at the cochlear mechanical level.

ACKNOWLEDGMENTS

I would like to thank the following people for their generous help: D. A. Berkley, D. Bock, M. J. Galley, J. L. Hall, S. M. Khanna, D. G. B. Leonard, J. Palin, F. Pirz, A. M. Ranade, G. Sokolich, and J. Tondorf. I would like to especially thank J. L. Flanagan for his special support of this work. The measurements described in this paper were carried out at the Fowler Memorial Laboratory of Columbia University at the College of Physicians and Surgeons, New York City.

Allen, J. B. (1979). "Cochlear models—1978," in *Models of the Auditory System and Related Signal Processing Techniques*, edited by M. H. Hoke and E. de Boer, Scand. Audiol. Suppl. 9, pp. 1–16.

Allen, J. B. (1980). "Cochlear micromechanics—A physical model of transduction," J. Acoust. Soc. Am. 68, 1660–1670.

Allen, J. B. (1981). "Cochlear modeling—1980," ICASSP-81, Proc. of IEEE ASSP Meeting, Atlanta, GA.

Allen, J. B. and Sondhi, M. M. (1979). "Cochlear macromechanics—Time domain solutions," J. Acoust. Soc. Am. 66, 123–132.

Anderson, D. J., Rose, J. E., Hind, J. E., and Brugge, J. F. (1971). "Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: Frequency and intensity effects," J. Acoust. Soc. Am. 49, 1131–1138.

Arthur, R. M. (1976). "Harmonic analysis of two-tone discharge patterns in cochlear nerve fibers," Biol. Cybernetics 22, 21–31.

Arthur, R. M., Pfeiffer, R. R., and Suga, N. S. (1971). "Properties of 'two-tone inhibition' in primary auditory neurons," J. Physiol. 212, 593–609.

Blesser, B. A. (1978). "Digitization of audio: A comprehensive examination of theory, implementation, and current practice," J. Audio Eng. Soc. 26, 739–771.

Cody, A. R., and Johnston, B. M. (1980). "Single auditory neuron response during acute acoustic trauma," Hear. Res. 3, 3–16.

Dallos, P. (1973). "Cochlear potentials and cochlear mechanics," in *Basic Mechanisms of Hearing*, edited by A. M. Møller (Academic, New York), pp. 335–376.

Eggermont, J. J. (1979). "Compound action potentials: Tuning curves and delay times," in *Models of the Auditory Signal Processing Techniques*, edited by M. Hoke and E. de Boer, Scand. Audiol. Suppl. 9, pp. 129–140.

Evans, E. F., and Wilson, J. P. (1971). "Frequency sharpening of the cochlea: The effective bandwidth of cochlear nerve fibers," Seventh International Congress on Acoustics, Budapest, 453–456.

Evans, E. F. (1977). "Frequency selectivity at high signal levels of single units in cochlear nerve and nucleus," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 185–192.

Fox, P. A., Hall, A. D., and Schryer, N. L. (1976). "Computing science technical report #47—The PORT mathematical subroutine library," (Fortran code and documentation Bell Laboratories Computing Information Service, 600 Mountain Ave., Murray Hill, NJ 07974).

Geisler, C. D., Rhode, W. S., and Kennedy, D. T. (1974). "Response to tonal stimuli of single auditory nerve fibers and their relationship to basilar membrane motion in the squirrel monkey," J. Neurophysiol. 37, 1156–1172.

Goldstein, J. L., Baer, T., 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 Educational Consultants Inc., Baltimore, MD), pp. 133–141.

Hall, J. L. (1980). "Cochlear models: Evidence in support of mechanical nonlinearity and a second filter (A review)." Hear. Res. 2, 455–464.

Hall, J. L. (1981). "Observations on a nonlinear model for motion of the basilar membrane," in *Hearing Research and Theory*, Vol. 1 (Academic, New York), pp. 1–61.

Kiang, N. Y. S., Baer, T., Marr, E. M., and Demont, D. (**1969**). "Discharge rates of single auditory-nerve fibers as functions of tone level," J. Acoust. Soc. Am. **46**, 106.

Kim, D. O., Molnar, C. E., and Matthews, J. W. (**1980**). "Cochlear mechanics: Nonlinear behavior in two-tone responses as reflected in cochlear-nerve-fiber responses and in ear-canal sound pressure," J. Acoust. Soc. Am. **67**, 1704–1721.

Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," J. Acoust. Soc. Am. **63**, 442–455.

Lynch, T. J., Nedzelnitsky, V., and Peake, W. T. (**1982**). "Input impedance of the cochlea in cat," J. Acoust. Soc. Am. **72**, 108–130.

Møller, A. R. (**1977**). "Frequency selectivity of single auditory-nerve fibers in response to broadband noise stimuli," J. Acoust. Soc. Am. **62**, 135–142.

Papoulis, A. (**1962**). *The Fourier Integral and its Applications* (McGraw-Hill, New York).

Pfeiffer, R. R., and Molnar, C. E. (**1970**). "Cochlear nerve fiber discharge patterns: Relationship to the cochlear microphonic," Science **167**, 1614–1616.

Pfeiffer, R. R., Molnar, C. E., and Cox, J. R. Jr. (**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), pp. 323–331.

Rhode, W. S. (**1978**). "Observations on Cochlear Mechanics," J. Acoust. Soc. Am. **64**, 158–175.

Rhode, W. S. (**1973**). "An Investigation of post-mortem cochlear mechanics using the Mössbauer Effect," in *Basic Mechanisms of Hearing*, edited by A. R. Møller (Academic, New York).

Russell, I. J., and Sellick, P. M. (**1978**). "Intracellular studies of hair cells in the mammalian cochlea," J. Physiol. **284**, 261–290.

Sokolich, W. G. (**1977**). "Improved acoustic system for auditory research," J. Acoust. Soc. Am. Suppl. 1 **62**, S12.

Zweig, G., Lipes, R., and Pierce, J. R. (**1976**). "The cochlear compromise," J. Acoust. Soc. Am. **59**, 975–982.

Zwislocki, J. J., and Sokolich, W. G. (**1974**). "Neuro-mechanical frequency analysis in the cochlea," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt (Springer, Berlin).