Nonlinear Modeling of Auditory-Nerve Rate Responses to Wideband Stimuli

Eric D. Young and Barbara M. Calhoun

Center for Hearing and Balance, Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, Maryland

Submitted 10 March 2005; accepted in final form 8 September 2005

Young, Eric D. and Barbara M. Calhoun. Nonlinear modeling of auditory-nerve rate responses to wideband stimuli. J Neurophysiol 94: 4441–4454, 2005. First published September 14, 2005; doi:10.1152/jn.00261.2005. The spectral selectivity of auditory nerve fibers was characterized by a method based on responses to random-spectrum-shape stimuli. The method models the average discharge rate of fibers for steady stimuli and is based on responses to ~100 noise-like stimuli with pseudorandom spectral levels in 1/8- or 1/16-octave frequency bins. The model assumes that rate is determined by a linear weighting of the spectrum plus a second-order weighting of all pairs of spectrum values within a certain frequency range of best frequency. The method allows prediction of rate responses to stimuli with arbitrary wideband spectral shapes, thus providing a direct test of the degree of linearity of spectral processing. Auditory-nerve fibers are shown to rely mainly on linear weighting of the stimulus spectrum; however, significant second-order terms are present and are important in predicting responses to random-spectrum-shape stimuli, although not for predicting responses to noise filtered with cat head-related transfer functions. The second-order terms weight the products of levels at identical frequencies positively and the products of different frequencies negatively. As such, they model both curvature in the rate versus level function and suppressive interactions between different frequency components. The first- and second-order characterizations derived in this method provide a measure of higher-order nonlinearities in neurons, albeit without providing information about temporal characteristics.

INTRODUCTION

In studies of neural coding, it is useful to have models that predict the responses of neurons to arbitrary stimuli. In the peripheral auditory system, physiological models have been developed that can duplicate accurately the responses of auditory nerve (AN) fibers (Bruce et al. 2003; Deng and Geisler 1987; Sumner et al. 2003; Zhang et al. 2001). Models of synaptic processing have extended these results into the cochlear nucleus (Hancock and Voigt 1999; Hewitt et al. 1992; Kalluri and Delgutte 2003; Rothman and Manis 2003; Rothman et al. 1993). However, such models depend on a detailed knowledge of the circuitry and the physiological properties of the neurons; it seems unlikely that similar models will be developed for higher levels of the auditory system because of the complexity of the circuits. Instead it seems more feasible to rely on “black-box” models of signal processing, using nonlinear methods derived from Wiener and Volterra analysis (Eggermont et al. 1983; Machens et al. 2004; Nelken et al. 1997; Sen et al. 2001; Versnel and Shamma 1998; Yamada and Lewis 1999; Yeshurun et al. 1989; Yu and Young 2000).

The first Wiener kernel, or reverse-correlation function (de Boer and Kuyper 1968), captures the tuning of AN fibers with low best frequencies (BFs) and accurately predicts general aspects of their responses (Carney and Yin 1988; de Boer and de Jongh 1978; Moller 1977; Temchin et al. 2005). For fibers with higher BFs and neurons in the CNS that do not phase lock to the stimulus waveform, the characterization must be based on higher-order representations, like the second Wiener kernel (Lewis et al. 2002; van Dijk and Wit 1994; Recio-Spinoso et al. 2005; Yamada and Lewis 1999), the spectrotemporal receptive field (STRF) (Aertsen et al. 1981; Kim and Young 1994; van Gisbergen et al. 1975), or a method based on second-order distortion products (van der Heijden and Joris 2003). The best studied of these is the STRF, which describes the average stimulus power spectrum as a function of time preceding spikes (Aertsen and Johannesma 1981; Eggermont 1993; Klein et al. 2000; Theunissen et al. 2001) and is directly related to the second Wiener kernel (Aertsen and Johannesma 1981; Klein et al. 2000; Lewis and van Dijk 2004). When studied with STRFs, AN fibers show tuning consistent with conventional tuning curves and inhibitory regions attributable to suppression and refractory effects (Kim and Young 1994; Lewis and van Dijk 2004).

The applicability of these characterizations is mainly limited by the degree of nonlinearity of auditory neurons; essentially it is impractical to estimate nonlinearities of a sufficiently high order to capture accurately the properties of auditory neurons, especially in the CNS (Johnson 1980). As a result, STRFs of neurons in the central auditory system change significantly according to the stimulus used to derive them (Escabi and Schreiner 2002; Nelken et al. 1997; Theunissen et al. 2000) and do not accurately predict responses to arbitrary stimuli (e.g., Eggermont et al. 1983; Machens et al. 2004; Nelken et al. 1997; Sen et al. 2001; Versnel and Shamma 1998; Yamada and Lewis 1999; Yeshurun et al. 1989; Yu and Young 2000).

One approach to the problem of nonlinearity is to separate the analyses of frequency selectivity and temporal (or envelope) characteristics. This is not generally possible for arbitrary systems, but STRFs of neurons in the AN and inferior colliculus are usually separable into independent time and frequency functions (Lewis and van Dijk 2004; Qiu et al. 2003), although not in all cases and not in auditory cortex (Depireux et al. 2001; Eggermont et al. 1981; Sen et al. 2001). In fact, models of lower auditory processing commonly separate the frequency tuning from the temporal processing of stimulus modulation (e.g., Dau et al. 1997; Delgutte et al. 1998). Of course, spectral processing (tuning) implies a corresponding temporal response function, but the representation of the auditory stimulus envelope is not determined by the frequency tuning, even in the AN (Joris and Yin 1992). Thus it seems reasonable to simplify the
task of characterizing lower-level auditory neurons by analyzing spectral and temporal processing separately.

Here we apply the method of spectral weight functions to AN fibers (Yu and Young 2000). Weight functions are equivalent to the STRF averaged over its time axis and so are a measure of the frequency-dependent component of a separable STRF. By ignoring temporal features of the stimulus/response relationship, the analysis can be extended to higher orders of nonlinearity than is feasible with the full STRF. We show here that both first- and second-order weight functions can be robustly estimated; these measure the gain of the fiber to stimuli energy in different frequency bands (1st order) or pairs of bands (2nd order). An advantage of weight functions is that they can be tested by attempting to predict the rate responses to wideband stimuli with various spectral shapes, thus providing a rigorous test of the completeness and accuracy of the model (Barbour and Wang 2003; Yu and Young 2000). The model accurately predicts the rate responses of AN fibers to both random-spectrum shape (RSS) stimuli and to noise signals filtered by head-related transfer functions (HRTFs) (Musicant et al. 1990; Rice et al. 1992). Second-order weights are necessary for accurate prediction of responses to RSS stimuli but not for responses to HRTF stimuli.

METHODS

Animal care and stimulus protocol

The data were obtained from nine experiments in healthy adult cats (3–4 kg). Cats were anesthetized with ketamine (40 mg/kg im) and xylazine (2 mg im). An intravenous line and tracheal tube were inserted. The animal was maintained at a surgical plane of anesthesia throughout the experiment with pentobarbital sodium (=5–15 mg/h iv as needed to prevent withdrawal reflexes). Atropine (0.1 mg im) was given to control secretions. The bulla was vented with a length of PE-90 tubing. The cat’s temperature was maintained at 38.5°C, and intravenous fluids were given. For recording, the posterior fossa was opened and the cerebellum retracted to expose the AN. Glass micropipettes filled with 3 M NaCl (10–30 MΩ) were used to record from single AN fibers. The procedures were approved by the Johns Hopkins Animal Care and Use Committee.

Single fiber recording was done in a sound-proofed (IAC) chamber. The acoustic system consisted of an electrostatic or a dynamic driver coupled to the ear through a 12.5-cm earbar. The system was calibrated in situ prior to the experiment using a probe tube placed within 2 mm of the ear drum. Typical calibrations have been shown previously (Rice et al. 1995). The electrostatic driver has a bandwidth of 30–40 kHz, whereas the dynamic driver provides significant energy only ≦10 kHz. The dynamic driver was used to provide stimuli at high sound levels in five experiments; only fibers with BF ≦10 kHz were studied in those experiments.

When a fiber was isolated, its BF and threshold were determined using an automated tuning curve algorithm. A series of RSS stimuli, described in the following text, were then presented. The stimuli were 400 ms in duration and were presented once per second across a range of sound levels in 10-dB steps. In one experiment, the same set of stimuli was repeated multiple times, but usually each stimulus was presented only once. The spontaneous rate (SR) was determined as the average rate during the last 400 ms of the interstimulus interval for the lowest-level RSS presentation. The rate responses to the RSS stimuli were used to construct the weight-function model described in the following text.

In one experiment, a test stimulus set consisting of a sample of noise (bandwidth: 10 Hz to 50 kHz) filtered with cat HRTFs (from Rice et al. 1992) was presented. Stimuli filtered with 99 different HRTFs were presented along with the unfiltered noise, and the resulting rate responses were used to test the ability of the weight-function model to predict responses to stimuli other than those used to generate the model. HRTF stimuli were also 400 ms in duration and were presented across a range of sound levels in 10-dB steps; the sound levels were adjusted to cover a similar range as the RSS stimuli, although it is not possible to exactly match levels. The sampling rate of the HRTF stimuli was adjusted to shift the spectra nominally at 10 kHz to the BF of the fiber. For example, a fiber with BF 10 kHz would receive HRTF stimuli with sampling rate 100 kHz (the nominal value), but a fiber with BF 5 kHz would receive the same stimuli with sampling rate 50 kHz. This manipulation guarantees that the spectral shapes and sound levels near the fiber’s BF vary maximally across stimuli.

Weight function model

The analyses in this paper are based on the following model for the average discharge rate $r$ of a fiber over the duration of the stimulus (Yu and Young 2000)

$$
 r = R_0 + \sum_{i=m}^{m+40} w_i S(f_i) + \sum_{i=m}^{m+40} \sum_{j=m}^{j=m+40} w_{ij} S(f_i) S(f_j) 
$$

$S(f)$ is the stimulus power in a bin centered on frequency $f$, measured in decibels relative to a reference level, which changes according to the overall sound level of the stimulus. In the model, rate is the sum of a constant $R_0$, which is the rate response to a stimulus with all frequency bins set to 0 dB, plus a first-order weighting of the spectrum, the first summation in Eq. 1, plus a second-order weighting of the spectrum, the second summation. The first-order weights $w_i$ are a measure of the gain of the neuron for energy at frequencies near $f_i$. Appropriate for a gain, the units of $w_i$ are s/0 dB). The second-order weights $w_{ij}$ measure the response of the fiber to quadratic terms like the energy-squared at a particular frequency [e.g., $w_{55}$ for $S(f_5)^2$] or the product of the energy at two different frequencies. Higher-order terms could be added, but it is not feasible to estimate them with the data obtained here. The weights are zero for frequency bins sufficiently far from BF, so it is only necessary to consider the model over some range of frequency bins $m_1$ to $m_2$. Note that the second-order weights are symmetric, $w_{ij}$ is the same as $w_{ji}$, so only one of each such second-order weight pair is included in the summation.

For all calculations using Eq. 1, the dB levels of the stimulus $S(f)$ were corrected for the acoustic calibration by setting the 0-dB reference level for the stimuli to the SPL of the frequency bin at BF with $S(\text{SPL}) = 0$ and then correcting the values in other bins according to the calibration $C(f)$ [C(f) is the sound level of a tone of frequency $f$ at a fixed rms voltage signal into the speaker]. In other words, the decibel values of the stimulus $S(f)$ were corrected by adding $C(f) - C(0)$, the difference between the calibration at $f$ and the calibration at BF. In practice, this adjustment has only a small effect.

The parameters of the weight-function model were fit to data consisting of the rate responses to a set of 95–99 RSS stimuli, described in the next section. Each of the stimuli has a different second-order weight pair is included in the summation.

$$
 E = \frac{1}{df} \sum_{k=1}^{n} \left[ \frac{S_k - r(S_k)}{\sigma_k^2} \right]^2 
$$

$r_k$ is the actual rate in response to the $k$th stimulus $S_k$, and $r(S_k)$ is the rate predicted by Eq. 1 for the $k$th stimulus. The variances $\sigma_k^2$ should be the variances of the rates $r_k$. Because multiple repetitions of the stimuli were usually not obtained, estimates of the rate variances...
are not available; therefore Poisson variances were assumed, $\hat{\sigma}_k^2 = \hat{r}_k/T$, where $T$ is the duration of the stimulus presentation. To avoid dividing by zero in Eq. 2, $\hat{\sigma}_k^2$ was not allowed to be $<1$, corresponding to 0.4 spikes/s. df is the number of degrees of freedom of the estimate. When $M$ parameters are being estimated, df = $K - M$, where $K$ is the number of stimuli. The degrees of freedom is important in comparing fits of first- and second-order models. Because the second-order model has more parameters, it is bound to fit better; however, dividing by df partially compensates for that effect. If the model fit the data exactly except for additive Gaussian noise in the rates (i.e., $\hat{r}_k = r(S_j) + n_k$ where $n_k$ has mean 0 and variance $\bar{\sigma}_k^2$), then $E$ in Eq. 2 would have mean value 1 and SD $\sqrt{2/df}$. Estimation of the weights by minimization of error $E$ is done by the method of normal equations (Press et al. 1986).

The SDs of the weight estimates were estimated using bootstrap, in which the calculation of weights was repeated for sets of 100 stimulus/response pairs chosen at random with replacement from the full set (Efron and Tibshirani 1993). Two-hundred bootstrapped weight sets were computed, and the mean ± SD of the weight estimates were calculated from those results. The bootstrap calculation gives an elevated estimate of SD for this problem, ~2.5 times larger than direct estimates of SD calculated in seven cases where the RSS stimulus sets were presented multiple times. For this reason, the bootstrap SDs are used here only for illustration and not to qualify weights for inclusion in the model.

Simulations show that this estimation method usually gives unbiased estimates of the weights. However, two sources of error were found. One source of bias in $R_0$ and error in the weights is rate clipping. Stimuli that drive the discharge rate to zero would produce negative rates if that were possible. Because the model does not include rectification at zero rate, the estimate of $R_0$ increases to better fit the clipped rates and the second-order weights are also distorted. In AN fibers, few rates are actually driven to zero, and so clipping is not a major source of error in these data. Compensation for clipping can be done by deleting zero-rate cases from the data before fitting; however, this was not necessary here.

A second source of bias in $R_0$ is estimating fewer weights than are actually needed, i.e., making $m_1$ too large or $m_2$ too small in Eq. 1. If an important second-order weight is deleted from the diagonal (where $f_j = f_k$), then the fitting process changes $R_0$ to compensate and better fit the overall average rate. The first- and second-order weights are still estimated without bias in this case. Low-BF fibers have the widest (along the log-frequency axis) weight functions, and it was sometimes difficult to include enough second-order weights in their models because the RSS stimulus sets were not large enough; this effect accounts for the fact that $R_0$ estimates for low-BF neurons were consistently biased by a few spikes/s, which is comparable to the SD of the $R_0$ estimates. Because of this bias, low-BF neurons (<3 kHz) are not shown in Fig. 7.

The range of frequencies used in estimating weights ($m_1$ to $m_2$ in Eq. 1) was set to 0.5–1 octave below BF and 0.5 octave above BF for the first-order weights and to 3/16–3/8 octave above and below BF for the second-order weights. As the number of weights included in the calculation increases past about half the number of stimuli presented (i.e., $M > K/2$), the bootstrap SD of the estimates increases rapidly, indicating that the calculation is unstable and the fits unreliable. The range of frequencies included in the model ($m_1$ to $m_2$) was set as wide as possible within this constraint ($M < K/2$). The number of second-order weights is proportional to the square of the bandwidth of the weight set, so the second-order weight limits were more restrictive than the first-order limits; however, the second-order weight range usually included the range over which first-order weights were significantly different from zero, from the bootstrap calculation, the major exception being low-BF neurons. In the section of RESULTS on prediction of responses (Fig. 8) a method for further refining the range $m_1$ to $m_2$ is described.

### RSS stimuli

The RSS stimuli are designed to estimate the parameters of Eq. 1. Each stimulus consists of a sum of tones equally spaced logarithmically at 1/64 octave. Tones of successive frequencies are gathered into groups of either four or eight tones to form frequency bins of width 1/16 or 1/8 octave, respectively. All the tones within a bin have the same amplitude, and the energy of a bin $S(j)$ is the decibel value corresponding to the sum of the energies of the tones contained in the bin. Because the frequencies are logarithmically spaced, the overall stimulus is not periodic and sounds noise-like. For the same reason, the overall spectrum of the stimulus has a $1/f$ shape, e.g., for a stimulus with 0 dB in all bins. To avoid a click at time 0, the starting phases of the tones were randomized.

The total bandwidths of the stimuli were 1, 2.13, or 8.5 octaves; the 1-octave stimuli were used in five experiments where high sound levels were studied (narrow bandwidth was needed to maximize the sound level). When 1- and 2.13-octave stimuli were used, the sampling rate of the stimulus was adjusted to place the BF of the fiber ~2/3 of the way between the lowest bin frequency and the highest.

The stimulus levels $S(j)$ were chosen as follows. The RSS stimulus set consists of 100 stimuli. Consider a matrix $S$ whose columns are the dB values of individual stimuli

$$S = \begin{pmatrix} \frac{1}{S_1} & \cdots & \frac{1}{S_{100}} \end{pmatrix}$$

The center frequencies of the bins $f_j$ vary down the columns of this matrix. To compute the decibel values, rows were generated as random vectors with 0 mean. The matrix was orthogonalized, using the orth() function in Matlab, so that $S \cdot S^T = (\text{const} \cdot I)$. The zero-mean property of the rows is preserved by this manipulation. The SD of the rows was set to 12 dB by multiplying by a constant. Although the orthogonalization is not necessary, it improves the robustness of the first-order weight estimation by making the estimation of weights at different frequencies independent. The values in the resulting matrix are the stimulus levels in each frequency bin, in dB relative to the reference level for the stimulus (as defined for Eq. 1). Note that the resulting values of $S(j)$ are roughly normally distributed around 0 dB.

One stimulus was generated for each column in Eq. 3. In practice, between one and five columns were set to 0 dB to provide an independent estimate of $R_0$, as will be discussed in Fig. 7. Thus the actual RSS stimulus set consisted of 95–99 stimuli. Each stimulus was generated by adding sinusoids, in random phase, with the amplitudes set by converting the decibel values in $S$ to linear values (as $10^{0.028}$). Stimuli were generated to have duration 400 ms at a nominal sampling rate of 100 kHz. When the sampling rate was changed to position the spectrum relative to a fiber’s BF, stimuli were extended in time, if necessary, by making them periodic. Although the periodic stimuli have a discontinuity (click) at each period, the clicks were not audible and the fibers did not show click responses. For the 8.5-octave-wide stimuli, frequency bin centers ranged from ~100 Hz to 38 kHz and the sampling rate was fixed at 100 kHz. For the variable-sampling rate stimuli, the sampling rate was always at least four times the highest frequency component in the stimulus.

### RESULTS

#### Properties of weight functions

Weight functions of AN fibers show an excitatory area (positive weights) near BF that may or may not be accompanied by an inhibitory area (negative weights) at lower or higher frequencies. The first-order weight functions for the example unit in Fig. 1, $B$ and $C$, were taken at sound levels separated by 20 dB. In both cases, there is a peak weight near 3 spikes/(s dB)
the second-order terms improve the fit at the highest rates; at
the higher level (Fig. 1E), the second-order terms improve the
overall fit, shown by a smaller vertical scatter around the line
of equality of the data points relative to the ○ points. The
legends in these figures give the values of error $E$ (Eq. 2) for
the fits. As is typical, $E$ is smaller with the second-order terms.

The pattern of weights in Fig. 1, B and C, is typical of AN
fibers. In this case, the stimuli had frequency bins 1/8 octave
wide and the BF of the neuron was high (16.3 kHz), so
significant weights only occupy one or two bins near BF. At
lower BFs, the weight functions are broader, on logarithmic
axes, consistent with the behavior of AN tuning curves (Evans
1977; Liberman 1978). Figure 2 shows an example for a fiber
with BF 1.4 kHz. In addition to being broader, the first-order
weight function has a smaller peak amplitude, \( \approx 1.5 \) spikes/s dB. The second-order weights show the same general patterns
as in Fig. 1.

The reference level for the RSS stimuli is important in that
the weight model is a local approximation of the neuron’s
properties; $R_0$ and the weights change with the reference level.
The relationship of the weight model and the discharge rates is
illustrated in Fig. 3, which plots discharge rate in response to
the RSS stimuli (ordinate) versus the sound level of a tone in
the frequency bin at BF (abscissa). Data are shown for a
low-SR fiber (Fig. 3A) and a high SR fiber (Fig. 3B); in each
case, data are shown at three reference sound levels, identified
by the symbols. The straight lines show a part of the fit to the
data, consisting only of $R_0 + w_{BF,BF}S(f_{BF})$. For some cases, the
effect of the second-order weight at BF is also shown by the
curved line, a plot of $R_0 + w_{BF,BF}S(f_{BF}) + w_{BF,BF}S^2(f_{BF})$. Of
course, the lines do not fit the data points accurately because
they are only a portion of the overall fit. Nevertheless the shift


![Fig. 1](https://example.com/fig1.png)

**Fig. 1.** Examples of weight functions for a fiber with best frequency (BF) 16.3 kHz and spontaneous rate (SR) 8/s. A: discharge rates in response to 100 random-spectrum shape (RSS) stimuli, 400 ms in duration; -- -, SR. **B.** 1st-order (left) and 2nd-order (right) weights for the responses in A. Weights are plotted vs. the bin center frequencies, in octaves relative to the BF of the fiber, determined from the tuning curve. For the 1st-order weights, the error bars are \( \pm 1 \) SD as determined by bootstrap (as discussed in METHODS, these estimates are probably inflated). For the 2nd-order weights, the × mark the weights that are >1 bootstrap SD away from 0. The 2nd-order weights are symmetric, so weights are only plotted on and below the diagonal. Sound levels are the dB SPL of a single tone in the stimulus at the 0-dB reference level, called dB/component. **C:** weights for the same fiber at 20 dB higher stimulus level. **D** and **E:** quality of the fits for the weight functions in **B** and **C** respectively. Rates predicted by the best-fitting model (ordinate) are plotted vs. actual rates (abscissa), ○, best-fitting 1st-order model; ×, best-fitting 2nd-order model. Insets: errors of the 2 fits (Eq. 2).

at BF with smaller weights at adjacent frequencies. In this case,
the frequency of the center of the RSS bin nearest BF (i.e., the
tuning curve BF) was slightly below BF, so the peak weight,
which is plotted at the bin center, appears to be below BF. The
second-order weights show positive (green) values on the diagonal
of the weight function, i.e., terms like $w_{BF,BF}S(f_{BF})^2$;
such terms fit the curvature of the rate response to stimulus
level. Negative weights (red) appear off the diagonal. The
largest negative and positive weight are usually at or near BF.

The quality of the fit of the model to the data are shown in
Fig. 1, D and E, which plots the rates predicted by the models
against the actual rates. The fits were done twice, once with
only $R_0$ and the first-order summation in the model (a “1st-
order model”), and then again with the full model (a “2nd-
order model”). Both $R_0$ and $\{w_i\}$ are slightly different between
first- and second-order fits. At the lower sound level (Fig. 1D),

![Fig. 2](https://example.com/fig2.png)

**Fig. 2.** Same as B–E of Fig. 1 for a lower BF unit (BF = 1.4 kHz, SR = 10/s). Notice that the frequency axis for the 2nd-order weights extends further than in Fig. 1.
of model parameters with reference level is clear. $R_0$ is the rate at the center point of the lines and increases monotonically with level in each case. $w_{BF}$ is the slope of the lines and changes most with level for the high SR fiber in Fig. 3B. $w_{BF,BF}$ models the curvature of the lines and is largest near threshold in these examples.

Three aspects of the RSS analysis can be seen in Fig. 3 (Calhoun et al. 1998). First, the first-order weights provide a best linear fit to the rate-level function, most clearly seen for the highest two levels in Fig. 3B. Second, the second-order weights on the diagonal of the weight matrix (like $w_{BF,BF}$) model departures of the rate response from linearity. This is clearly illustrated at the lowest reference level in Fig. 3A (Δ), where about half the data points are below threshold and the rate response has a clear upward curvature. Third, the RSS stimuli do not recapitulate the static rate-versus-level function of the fiber. Instead, there is an adjustment of the dynamic range of the fiber at each reference level. The adjustment can be seen by comparing the points for the 19 and 39 dB reference levels in Fig. 3A (× and ○). For a particular sound level on the abscissa, the points for the 39 dB reference level are systematically below those for the 19 dB level (except for 2 points). Presumably, this behavior reflects cochlear suppression. The dynamic range of the rate response is preserved, to some extent, at higher reference levels by suppressing responses to tones at all frequencies. Consistent with this hypothesis, the suppression-like effect is apparent in the low-SR fiber in Fig. 3A and not in the high-SR fiber in Fig. 3B.

The large second-order terms near threshold are produced by the rectification of rate at zero in Fig. 3A and at SR in Fig. 3B. These second-order terms are not the focus of later sections on second-order weights. Instead, it is the smaller second-order weights illustrated by the curves at 11 and 31 dB in Fig. 3B, where the fiber is well away from threshold. Notice that rates do not show a hard rectification at high sound levels as they do at threshold. The data at the highest levels in Fig. 3 are typical in that saturation occurs as a decrease in the slope of the rate functions at high levels and not as a true saturation with negative curvature.

**First-order weights at various sound levels**

First-order weight functions are qualitatively the same in all AN fibers. The major quantitative trends are illustrated in Figs. 1 and 2: weight functions broaden, on a logarithmic frequency axis, and have lower amplitudes at lower BF. There are three additional trends: first, weight functions saturate, meaning a decrease in weight amplitude, at higher reference sound levels; second, the frequency of the peak weight shifts toward lower frequencies at very high sound levels; third, at higher sound levels there is a consistent small inhibitory weight at frequencies above BF.

The trends in weight functions were studied by averaging weights across fibers, after gathering fibers into three groups by BF: 1–3, 3–10, and 10–30 kHz. The similarity of weights across fibers is illustrated in Fig. 4, which shows weight functions for individual medium-SR fibers (gray), along with their averages (black lines with unfilled circles). The weight functions were aligned on their tuning-curve BF for averaging. The center frequency of the bin containing the largest weight is always near BF (except at very high sound levels) but is not always exactly at BF. However, the trends discussed in this section do not change if weight functions are aligned on the peak weight instead of the tuning curve BF. The weights plotted here were estimated from responses to RSS stimuli with 1/16-octave bins, providing better frequency resolution than in Figs. 1 and 2. The tendency of weight functions to be lower-amplitude but broader in extent at lower BFs is clear in Fig. 4.

Figure 5 shows averaged first-order weights for fibers in three SR groups. The amplitudes of weights are small for reference levels near threshold (at the bottom of the figure); weights increase in amplitude with reference sound level (moving upward in the figure) and ultimately decrease again at high sound levels. Large weights occur at relatively low reference sound levels in high SR fibers (right) and at successively higher levels in medium and low SR fibers. The decrease in weights at high sound levels is consistent with the saturating nature of AN rate responses (Sachs and Abbas 1974). Data are available at levels up to 80 dB SPL/component; at higher levels, the mean weights continue to decrease in amplitude toward zero (not shown).

**FIG. 3.** Discharge rate vs. sound level at BF for the RSS stimuli at 3 reference sound levels given in the legends as dB/component. A: data from a low-SR fiber, BF = 5.7 kHz, SR = 1.8/s. B: data from a high-SR fiber, BF = 9.0 kHz, SR = 55/s. Each point is the average rate in response to 4 repetitions of a 0.4-s RSS stimulus plotted against the sound level of a tone in the bin containing BF (dB/component). The sound levels include the effect of the reference level [i.e., they are $S(BF) +$ reference level]. The lines are portions of the weight-function model as explained in the text. The line styles identify the reference level to which the line belongs, defined in the legends. Only first-order partial fits (lines) are shown for the top 2 sound levels in A. These data are from an early experiment in which there were only 25 stimuli in the RSS set and each was repeated 4 times; there were 8 bins/octave.
A second property of first-order weights illustrated in Fig. 5 is the tendency of the peak of the weight function to shift toward lower frequencies at high sound levels. The vertical dashed lines show BF; for all three SR groups, the peak weight shifted below BF at levels higher than 40- to 50-dB/component. This is about 70–80 dB above the threshold of the most sensitive fibers.

Finally, significant negative weights were often seen at frequencies above BF in individual fibers (e.g., Fig 1C); such negative weights were consistently seen in the averages, in 15/23 cases in Fig. 5, mostly at high sound levels (>10 dB/component). These are often statistically significantly different from zero. Similar inhibitory regions have been reported in STRFs (Kim and Young 1994; Lewis and van Dijk 2004).

The peak sensitivity to stimulus intensity is found in high SR fibers at low sound levels but shifts to medium and low SR fibers at higher levels (Fig. 6). The shift is clearest for the high BF fibers (Fig. 6A) but is evident for lower BFs as well (Fig. 6B). The weight at BF increases rapidly at low sound levels, peaks at a level that depends on the SR group, and then declines toward zero. Of course, at the highest levels the largest weight is below BF, so it is not plotted in Fig. 6; however, the decline of peak weight at high levels occurs regardless of which weight is plotted. Weights are significantly smaller in the low BF group (1–3 kHz, gray lines in Fig. 6B) than for the two higher BF groups.

The bandwidth of a weight function is a measure of the frequency selectivity of the fiber for the components of a broadband stimulus. Because weights have a relatively narrow dynamic range, it is not feasible to compute Q10s, the usual measure of bandwidth; instead, the widths of weight functions at half the peak weight were computed as the octave difference of upper and lower half-gain points (Fig. 6C). Q10 and halfwidths increase with sound level for fibers in the two higher BF ranges. Halfwidths also seem to increase for low BF fibers (gray dashed lines with open symbols), but the data are noisier for the low BF group. Ignoring the low SR fibers, the halfwidths are larger in low BF units and increase more rapidly with sound level.

**Nonlinearity of weight functions**

Second-order weights are necessary only if they improve fits to data and predictions of responses to other stimuli. However, there is a simple argument that the second-order terms are needed. In the model of Eq. 1, the rate response to the all-0-dB stimuli [i.e., those with \( S(f) = 0 \) for all \( f \)] is \( R_0 \). If the second-order terms are not important and can be dropped from
Eq. 1, then the average of the rates across the entire RSS stimulus set is also $R_0$, because of the way the RSS stimuli were generated. To see this, consider the average rate $\langle r_k \rangle$ predicted by a model containing first-order terms only

$$\langle r_k \rangle = \frac{1}{K} \sum_{i=1}^{K} \left[ R_0 + \sum_{i} w_i S_i(f_i) \right]$$

(4)

$S_i$ and $r_k$ are the $k$th stimulus and the rate in response to it and $K$ is the total number of stimuli. As described in METHODS, the last summation in Eq. 4 is zero by the way the stimulus set was computed, i.e., the average level across stimuli is zero within each frequency bin. Thus the overall average rate $\langle r_k \rangle$ should equal $R_0$ if the second-order terms in the model can be ignored.

Figure 7 shows a comparison of the rates in response to the all-0-dB stimuli (abscissa) to the overall average rates ($\times$) and the estimates of $R_0$ from the second-order model ($\bigcirc$). The latter are close to the all-0-dB rates, as expected. However the average rates are systematically larger than both $R_0$ and the rates to the all-0-dB stimuli. The differences are statistically significant (see the caption) and are caused by the contribution of the second-order terms. For this analysis, only fibers with BF$s$ from 3 to 30 kHz were used; lower BF fibers yield a biased estimate of $R_0$ as discussed in METHODS. Fibers with rates near threshold or saturation were also excluded (see caption) to eliminate nonlinearities from those causes. The result does not change qualitatively if low BF fibers and fibers at threshold and saturation are included.

In addition to the argument of Fig. 7, the importance of second-order terms is suggested by the quality of fits to the responses to RSS stimuli. Figure 8 shows the results of a cross-validation test in which the RSS stimulus set was divided into two groups: the first 60 stimuli were used to derive a weight-function model and the remaining 40 stimuli were used to test the adequacy of the model. Testing with stimuli that were not included in calculating the model provides a test of the ability of the model to generalize and allows identification of model overfitting, i.e., adjusting model parameters to fit noise in the data. The frequency range covered by the weights included in the model (i.e., the range $m_1$ to $m_2$ in Eq. 1) was chosen using the algorithm illustrated in Fig. 8, A and B. Beginning with only one weight at BF, weights were added alternately on the low- and high-frequency side of BF. The number of second-order weights was also increased, by using a square of second-order weights the sides of which cover the same frequency range as the first-order weights. Both first- and second-order models were computed for each number of
same test data. The medians are significantly different (rank-sum test, Eq. 5 of the fraction of variance (\( f_v \)) as the error measure, where \( f_v \) is defined as

\[
 f_v = 1 - \frac{\sum_{i=1}^{K} (\hat{r}_k - r_i(S_i))^2}{\sum_{i=1}^{n} (\hat{r}_k - \bar{r})^2}
\]  

(5)

where \( \hat{r}_k \) is the rate in response to the \( k \)th stimulus, \( r_i(S_i) \) is the value predicted by the model, and \( \bar{r} \) is the mean value of \( \hat{r}_k \). \( f_v \) has a maximum of 1 when the model fit is perfect and decreases as the error increases; it can go negative for large error, where the mean rate \( \bar{r} \) is a better fit than the model.

The fraction of variance captured by the second-order model is significantly better than that captured by the first-order model (Fig. 8D, see legend for statistics). However, the quality of the fit varies. There are a number of cases of poor fits in Fig. 8D (\( f_v \) values <0.2). These cases are mostly responses at high sound levels where the rates are close to saturation. In this case, the noise in the rate estimates is a significant portion of the rate variation and the model cannot fit this random component.

In practice, the maximum likely value of \( f_v \) when fitting test data is \([1 - \sigma_r^2/(\sigma_r^2 + \sigma_n^2)]\), where \( \sigma_r^2 \) is the variance in estimating the rates (i.e., the variance of the noise in \( \hat{r}_k \)) and \( \sigma_n^2 \) is the variance of the rates (the denominator sum in Eq. 5, divided by \( K - 1 \), assuming no noise in the rates). As a test of the noise limit in these data, \( f_v \) was corrected by subtracting an estimate of \( \sigma_n^2 \) from both numerator and denominator in Eq. 5 (using the Poisson estimate \( \hat{r}_k/K \) for \( \sigma_n^2 \)). With the correction, the peak in the distribution of \( f_v \) for the second-order model (i.e., the peak at 0.6–0.8 in Fig. 8D) is centered on 1. Thus

weights and the errors of the fits (Eq. 2) were plotted as in Fig. 8B. The left plot shows errors in the fit to the model data for first- and second-order models; these are typical in that the error is smaller for the second-order model and both fits improve over the first few weights and then saturate, usually with a shallow local minimum around four first-order weights (arrow). The right plot shows errors for the test data (i.e., for the 40 RSS stimuli not included in computing the model). Here the error is also smaller for the second-order model, but the error of the second-order model increases rapidly beyond three first-order weights (typically 3–6 weights). The increase suggests that the second-order model is overfitting the data with the larger number of weights.

For each set of responses, the number of weights included was set by the first local minimum in the error function for the second-order model fit (the arrow in Fig. 8B, left) or, if there was no local minimum, by the smallest number of weights giving an error within 20% of the overall minimum. This method usually yields an error for the test data (Fig. 8B, right) near the minimum error, so models chosen in this way are called “best-fitting” models in the following text. Figure 8A compares the best-fitting model to a model computed with a wider bandwidth. As is typical, the differences between the weights in the two models are small. Best-fitting models had between three and five first-order weights, for RSS stimuli with 1/8-octave bins.

The second-order best-fitting models provided a clearly better fit to the test data, compared with the first-order models. Figure 8C shows a comparison of error \( E \) for first- and second-order models, meaning best-fitting models evaluated with test data. The second-order models have less error in the majority of cases and the difference is significant (\( P \approx 0 \), see legend).

The quality of the fit was also better using fraction of variance (\( f_v \)) as the error measure, where \( f_v \) is defined as}

\[
f_v = 1 - \frac{\sum_{i=1}^{K} (\hat{r}_k - r_i(S_i))^2}{\sum_{i=1}^{n} (\hat{r}_k - \bar{r})^2}
\]  

(5)
evaluating the fits of the second-order model appears to be limited by the noise in the data for fv values around 0.6–0.8.

Properties of second-order weight functions

Second-order weights have a particular shape that is illustrated by the examples in Figs. 1, 2, and 8 and the averaged second-order weights in Fig. 9. Weights are positive on the diagonal of the weight matrix (green squares), i.e., for terms like $w_{BF,BF}S(f_{BF})^2$. The positive diagonal is surrounded by negative weights (red squares) for terms involving the product of two different frequencies. This shape is observed in all BF ranges and SR groups. As sound level changes, the behavior shown in Fig. 9 is typical. At low levels, positive weights on the diagonal dominate. At higher levels, the negative weights off-diagonal increase in magnitude. At very high levels, the weights become smaller, like the first-order weights, reflecting a slight advantage for the second-order models, depending on the number of weights determined by the analysis of Fig. 8. The weight-function model was computed using the reference level of the lower-level RSS stimulus set. In some cases, the points lie near the line of equality between actual and model predictions, with RSS binwidths of 1/8 and 1/16 octave. The shapes are qualitatively the same, which suggests that the positive diagonal is very narrow in its frequency extent, in particular that it is $\leq$1/16 octave in width. The actual shape of the second-order weights near the diagonal is probably not adequately revealed in the present data.

Prediction of rate responses to HRTF spectral shapes

The responses to HRTF stimuli were used to test the ability of the weight-function model to predict response rates. These stimuli were computed by filtering a sample of broadband noise with transfer functions of the cat head and ear for sound sources at varying azimuth and elevation (Musicant et al. 1990; Rice et al. 1992). HRTF spectra are different from the RSS spectra in that they are smoother in shape except at spectral notches. HRTF and RSS spectra do have the common feature of a general low-pass shape, so that the average overall spectral levels behave roughly as $1/f$ in both stimulus sets.

For the weight-function modeling, the energy in the HRTF spectra was gathered into frequency bins corresponding to the RSS bins. The resulting spectra were corrected for the acoustical calibration and expressed as dB relative to the reference level for the weight-function model. In general the RSS and HRTF stimuli were presented at a number of different reference sound levels. For each HRTF data set, a weight function model was computed from the two RSS data sets with reference levels nearest (within 10 dB) the average HRTF sound level at the fiber’s BF. The reference level for the whole calculation was the reference level of the lower-level RSS stimulus set. In some cases, only one RSS set within the 10-dB window was available. The weight-function model was computed using the number of weights determined by the analysis of Fig. 8. The model was then applied to compute the discharge rates for the HRTF stimuli. In general, the two-level models did a better job of predicting HRTF responses than the one-level models (not shown).

The models did a good job of predicting rate responses to HRTF stimuli (Fig. 10). The examples in Fig. 10, A and B, show comparisons of predicted and actual rates for two fibers with good fits; results are shown for both the first- and second-order models, identified in the legend. In all cases, the points lie near the line of equality between actual and model rates. Unlike the result with RSS stimuli, the first- and second-order models performed about equally well, although there was a slight advantage for the second-order models, depending on how the error was measured. Figure 10C shows the chi-square error $E$ (Eq. 2) for the first- versus second-order models. The second-order models are significantly better (see legend). However, the second-order models were not significantly better when evaluated using $fv$ (Fig. 10D). The $fv$ values for both fits in Fig. 10D are comparable to the results for the second-order model with RSS test stimuli in Fig. 8D. Again, the $fv$ values that cluster around 0.6–0.8 are at the maximum allowed by the variance of the data. For these fibers, the second-order
sound levels of the stimuli. Many of the cases with large errors are caused by this constant offset. The advantage of the second-order model is mainly in these large-error cases.

DISCUSSION

Nearly linear spectral weighting in AN fibers

These results show that AN fibers can be approximated to have a nearly linear dependence of discharge rate on the spectral content of a broadband stimulus. “Nearly linear” means that a first-order weighting of the dB spectrum captures much of the rate variation, but a small second-order correction can improve the fit (Figs. 8 and 10). The second-order correction consists of excitatory and inhibitory terms. The excitatory (diagonal) terms provide a quadratic model of the curvature of rate functions away from a strictly linear dependence on stimulus level. The inhibitory (off diagonal) terms model the inhibitory interactions of differing frequencies. The negative coefficients of the second-order weights vary with sound level (Fig. 9) and presumably reflect cochlear suppressive interactions (Delgutte 1990; Sachs and Abbas 1976; Sachs and Young 1980; Wong et al. 1998).

The near-linearity of AN weight functions provides an explanation for previous results in which the rate responses to vowel stimuli were studied (Le Prell et al. 1996; May et al. 1996, 1998). In those data, discharge rate was approximately linearly proportional to the dB level of the frequency component nearest BF. This behavior is expected if the first-order weights account for most of the rate response and if the largest weight is at BF. Moreover, the levels of adjacent stimulus components vary slowly with frequency in vowel spectra, so the effects of summing over a range of frequency components near BF would produce a small effect and would not be noticeable in the analysis done by May and colleagues.

First-order weight functions show small negative weights above BF. Presumably these negative weights also reflect cochlear suppression. As is the case for second-order weights (Fig. 9), the negative first-order weights are seen at higher sound levels. However, the effects of cochlear suppression are present in the first-order weights in more ways than as negative weights. As was pointed out in discussing Fig. 3A, the rate response can be smaller, for a given S(BF), when the overall stimulus energy is higher. This is especially so in low SR fibers, which show the effects of suppression more strongly. This effect was previously reported for vowel responses in AN fibers (May et al. 1996; Wong et al. 1998). Presumably the higher overall level of the components of the broadband stimulus suppresses responses to individual frequency components, with the result that a dynamic response is observed over a wider range of levels.

The dynamic range of AN fibers can be defined as the range over which weights are substantially nonzero. Figure 6, A and B, shows that this range is quite wide. For the high SR fibers, the range extends over 80 dB, much wider than the range usually quoted for rate versus level functions for tones (Evans and Palmer 1980; Sachs and Abbas 1974). The fact that AN fibers have wider dynamic ranges for broadband stimuli than for tones has been shown previously in experiments with broadband noise (Heinz and Young 2004; Schalk and Sachs 1980) and envelope modulation (Joris and Yin 1992; Smith and Brachman 1982). Again this effect seems to be a result of cochlear suppression.
Weight-function model

The weight-function model has two important features: first, it uses a logarithmic measure of frequency, and second, it is based on a logarithmic (dB) measure of stimulus energy. The logarithmic spacing of bins on the frequency axis is justified by the approximately logarithmic spacing of frequencies on the basilar membrane (Greenwood 1961; Keithley and Schreiber 1987; Liberman 1982). The use of logarithmic frequency bins also means that the RSS stimuli have amplitude spectra with an average 1/f shape. This is a reasonable stimulus design as many natural sounds, like speech and HRTF-filtered noise, have such 1/f spectra on average.

The justification of the decibel measure of stimulus energy (as opposed to, say, a linear measure of sound pressure in pascals or sound intensity in watts) was originally that it works well as an empirical description of AN rate responses: a plot of discharge rate versus the decibel stimulus spectrum level at BF produces an approximately linear relationship over 20–30 dB for vowels (Conley and Keilson 1995; May et al. 1996) and for HRTF-filtered noise (May and Huang 1997). Further support is provided by the recent result that neurons in the inferior colliculus respond more strongly to pseudorandom stimuli based on a logarithmic, as opposed to linear, intensity scale (Escabi and Schreiner 2002). The results shown here demonstrate that the dependence is not simply linear, that higher-order correction terms are needed to obtain the best fit. However, the results of prediction experiments with the HRTF stimuli show that the second-order corrections can be a relatively small effect in AN fibers.

The weight-function model takes the form of a Taylor-series approximation of the nonlinear rate function of the neuron. As such, it is valid only over a limited, but unknown, range near the center-point of the approximation, which in the present case is an all-zero-dB RSS stimulus of a particular sound level. Although the model works well in AN fibers with a 12-dB SD of the bin levels, such is not the case in neurons in the DCN (Yu and Young 2000), where good approximation can be achieved only with a smaller stimulus range (down to 3 dB) (L. J. Reiss, S. Bandyopadhyay, and E. D. Young, unpublished data). Ultimately, the usefulness of the weight-function model is limited to those neurons for which a low-order Taylor approximation is sufficient. In fact, one use of the weight-function method, which is enabled by the ease with which it can be used to predict rate responses to arbitrary wideband spectra, is to classify neurons according to their degree of linearity of spectral processing (Yu and Young 2000).

There is a risk in using log amplitudes for $S(f)$ in that doing so departs from the theoretical foundations provided by the Wiener-Volterra theory. The use of logarithmic bin amplitudes is empirically motivated in this work by the studies quoted in the preceding text and by the fact that the weight model works well in fitting and predicting AN rate data. One speculation about the logarithmic model is that it represents an approximation to the compressive input/output function of the basilar membrane. In this view, the input signal to the AN fiber is a compressed function of the stimulus, so that a logarithmic or other compressive measure of stimulus energy is appropriate in a model of rate responses. This question remains as a matter for future work.

Importance of second-order terms

Although the need for the second-order terms in modeling AN fibers seems clear, they were not important in predicting rate responses to the HRTF spectra. There are two possible reasons for the difference between the RSS and HRTF spectra in this regard. One reason is that the range of variation of the RSS stimuli around the reference level is larger (SD $\approx$ 12 dB) than for the HRTF spectra (SD $\approx$ 9 dB). Smaller-amplitude stimulus variations produce rate responses that are better fit by low-order spectral weight models (L. J. Reiss, S. Bandyopadhyay, and E. D. Young, unpublished data). A second possible reason is that the HRTF stimuli have much smoother spectra than the RSS stimuli. Whereas the RSS stimuli have uncorrelated values in adjacent frequency bins, the HRTF spectra generally vary slowly with frequency except at notches. In addition, the notches are substantially reduced in amplitude by the filtering inherent in grouping HRTF stimuli into bins for the weight-function model. Smoother spectra are likely to lead to a cancellation of positive and negative terms in the second-order summation, reducing its importance. Thus RSS spectra are a more challenging test of the model, even though it is computed from them in the first place.

It is instructive to relate the first- and second-order weights to the orders of nonlinearity defined in the Wiener-Volterra theory. First-order representations are those that respond to the stimulus waveform like a linear filter. They are generally estimated by reverse correlation, which gives the first Wiener kernel directly (de Boer and de Jongh 1978; de Boer and Kuyper 1968). Second-order representations are computations like the STRF that model responses to the stimulus envelope (Theunissen et al. 2001) or power spectrum (Aertsen and Johannesma 1981). The STRF can be derived from the second Wiener kernel, so that it is a second-order characterization in the Wiener-Volterra sense (Aertsen and Johannesma 1981; Klein et al. 2000; Lewis and van Dijk 2004). It can be shown that the first-order weights used here are equivalent to the STRF integrated along its time axis, i.e., to the STRF collapsed on its frequency axis. Of course, the STRF uses linear stimulus energy, not logarithmic as in the weight model. However, STRFs computed from stimuli with logarithmically defined stimulus levels have similar properties to those defined from linear stimulus levels (Escabi et al. 2003). The difference in the definition of the stimulus, while important mathematically, does not change the qualitative nature of empirical STRFs by much. Thus first-order weights are analogous to a second-order representation in the Wiener-Volterra theory, with the reservations stated above. In the same way, the second-order weights are analogous to a fourth-order representation in the Wiener-Volterra theory.

In the central auditory system, neurons become noticeably nonlinear. As discussed in the INTRODUCTION, methods such as the STRF that weight the stimulus spectrum linearly have difficulty in predicting responses to arbitrary stimuli (e.g., Eggermont et al. 1983; Machens et al. 2004; Yeshurun et al. 1989) and give different receptive fields for different stimulus sets (e.g., Theunissen et al. 2000). These failures are important because if a receptive field model does not predict the responses to stimuli accurately, it is not clear to what extent it is a meaningful description of the responses of the neuron.
Moreover even in the inferior colliculus, there are units that do not have a significant STRF (Escabi and Schreiner 2002), suggesting that such units must be characterized with higher-order models. Results in auditory cortex, where responses are not well-fit with first-order weight functions, also suggest the potential of further work with higher-order characterizations (Barbour and Wang 2003). By allowing higher-order model construction, the weight function method may be useful for studying the spectral selectivity of such units.

**Measures of frequency selectivity**

The first-order weights in Fig. 5 provide a measure of the frequency selectivity of AN fibers responding to a broadband stimulus. This measure differs from those derived using tones (i.e., tuning curves or response maps) because it includes interactions among the frequency components of the broadband stimulus used to obtain it. The first-order weights share with revcor filters and STRF measures of AN tuning the property that the bandwidth of the tuning increases with sound level (Fig. 6C) (Carney and Yin 1988; Kim and Young 1994; Lewis and Henry 1994; Möller 1977; Recio-Spinoso et al. 2005).

It is useful to compare the increase with that measured for basilar membrane tuning using tones (Rhode 1984; Ruggero et al. 1997). Figure 11 shows a comparison of the halfwidths of weight functions (from Fig. 6C) with the halfwidths of two published gain functions for the chinchilla basilar membrane (Ruggero et al. 1997). The basilar membrane gain functions are comparable to weight functions in that both are gains in units of response per stimulus level and the halfwidths are computed in the same way (described in the caption of Fig. 6C). The difference between response/decibel in one case and response/pascal in the other can be shown not to change the shape of the gain functions (i.e., gain vs. frequency) as long as the basilar membrane data for each gain function are taken at a fixed sound level. The basal difference in bandwidth between cat and chinchilla cochleas is removed by plotting the widths in relative terms, i.e., in terms of octaves. On these log-log axes, the rate of widening of basilar membrane tuning with sound level (the slope of the functions in Fig. 11) is more than twice that of the weight functions (see the legend).

Because the basilar membrane response in chinchilla has little or no even-order nonlinearity (Recio et al. 1997), the basilar membrane tuning is presumably mostly first order in the Wiener sense. Thus the comparison in Fig. 11 is between a first- and second-order characterization of peripheral tuning, and the difference in the growth of bandwidth with level may reflect some aspect of first- versus second-order components of the system. Consistent with this result, the bandwidths of STRFs for AN fibers are narrower than those of revcor functions and increase less with sound level (Kim and Young 1994). Taken together, these data suggest that tuning is different for responses that are predominantly first order, such as low-BF AN fibers or basilar membrane responses at all BFs, and responses that are predominantly second-order, such as high BF AN fibers. For this to be true, there must be a narrowing of effective tuning between the basilar membrane and AN responses in high-BF fibers.

The weight function method provides a view of AN spectral sensitivity that complements other approaches. Its principal advantages are the possibility of predicting rate responses to arbitrary wideband spectra and the possibility of studying higher order nonlinearities than is feasible with reverse-correlation methods. The corresponding disadvantage is that no information is provided about temporal characteristics of neurons. As long as neurons’ responses are separable, however, the temporal and spectral sensitivities can be studied separately.

**ACKNOWLEDGMENTS**

B. Cranston, J. Wong, and R. Miller assisted in some experiments. Preparation of the manuscript was aided by comments from members of the Neural Encoding Lab.

Present address of B.M. Calhoun: Scientific Learning, Oakland, CA.

**GRANTS**

This work was supported by National Institute of Deafness and Other Communications Disorders Grant DC-00115 and fellowship DC-00202.

**REFERENCES**


