A computational algorithm for computing nonlinear auditory frequency selectivity

Ray Meddis and Lowel P. O'Mard
Centre for the Neural Basis of Hearing at Essex, Department of Psychology, University of Essex, Colchester CO4 3SQ, United Kingdom

Enrique A. Lopez-Poveda
Facultad de Medicina, Universidad de Castilla-La Mancha, Campus Universitario, 02071 Albacete, Spain

(Received 26 September 2000; revised 6 March 2001; accepted 14 March 2001)

Computational algorithms that mimic the response of the basilar membrane must be capable of reproducing a range of complex features that are characteristic of the animal observations. These include complex input output functions that are nonlinear near the site's best frequency, but linear elsewhere. This nonlinearity is critical when using the output of the algorithm as the input to models of inner hair cell function and subsequent auditory-nerve models of low- and high-spontaneous rate fibers. We present an algorithm that uses two processing units operating in parallel: one linear and the other compressively nonlinear. The output from the algorithm is the sum of the outputs of the linear and nonlinear processing units. Input to the algorithm is stapes motion and output represents basilar membrane motion. The algorithm is evaluated against published chinchilla and guinea pig observations of basilar membrane and Reissner's membrane motion made using laser velocimetry. The algorithm simulates both quantitatively and qualitatively, differences in input/output functions among three different sites along the cochlear partition. It also simulates quantitatively and qualitatively a range of phenomena including isovelocity functions, phase response, two-tone suppression, impulse response, and distortion products. The algorithm is potentially suitable for development as a bank of filters, for use in more comprehensive models of the peripheral auditory system. [DOI: 10.1121/1.1370357]

PACS numbers: 43.64.Bt, 43.66.Ba [BLM]

I. INTRODUCTION

Using observations in chinchilla and guinea pig preparations, we evaluate an algorithm for simulating the nonlinear response of the cochlear partition in response to stapes motion. The algorithm is intended as a high-speed component in much larger models of the response of the whole auditory periphery to acoustic stimulation. It is not a model of cochlear mechanics, but rather a signal-processing algorithm. It is, however, an example of a recent tradition of nonlinear algorithms (e.g., Carney, 1993; Giguère and Woodland, 1994; Goldstein, 1990, 1995; Irino and Patterson, 1997; Lopez-Poveda *et al.*, 1998; Lyon, 1982; Meddis *et al.*, 1990; Robert and Eriksson, 1999). The aim of such models is to promote the evaluation of computational psychoacoustic theories based on the anatomy and physiology of the peripheral auditory system. Nonlinearity is an important characteristic of the response of the auditory partition (Rhode, 1971) and is relevant to a range of psychophysical phenomena. A high-speed nonlinear algorithm for simulating the nonlinear mechanical input to the inner hair cell transduction process is also essential to the rigorous modeling of the response of the auditory nerve (AN) and the auditory signal processing in the auditory brainstem (Winter *et al.*, 1990; Yates *et al.*, 1990).

The eventual aim is to construct a filter bank consisting of a series of such algorithms to represent as many locations along the cochlear partition as the modeller requires (cf. Lopez-Poveda and Meddis, 2000). Unfortunately, we only have detailed cochlear measurements of the input/output (I/O) functions at a limited number of sites. Below, we evaluate the algorithm at three sites for which detailed laser velocimetry observations are available. These three regions have best frequencies (BFs) of approximately 0.8, 10, and 18 kHz (Rhode and Cooper, 1996; Ruggero *et al.*, 1997; and Nuttall and Dolan, 1996). The development of a complete filter bank will be deferred so that this study can deal with the more focused question of the algorithm's ability to simulate, qualitatively and quantitatively, direct measurements of cochlear response functions. Future development of the filter bank will inevitably need to draw on observations of auditory nerve (AN) activity to fill in the gaps between the cochlear locations for which we have direct measurements. Unfortunately, nonlinear processes in the inner hair cell, the synapse, and AN dendrites intervene between the basilar membrane (BM) and AN fibers. As a result, AN data cannot be used as a *direct* test of the validity of the algorithm, even though such data will undoubtedly help tune the model once its value is established. In this study, we restrict our attention to the narrower question of whether the algorithm can generate a useful simulation of the complex, nonlinear behavior observed at those sites where physical measurements have been reported.

A particular challenge for the algorithm is the need to simulate the considerable differences in the appearance of the input/output (I/O) functions at different points along the cochlear partition [see Figs. 1(a), (c), and (e)], particularly the difference between basal and apical sites. We also need to know whether it can simulate a range of ancillary phe-

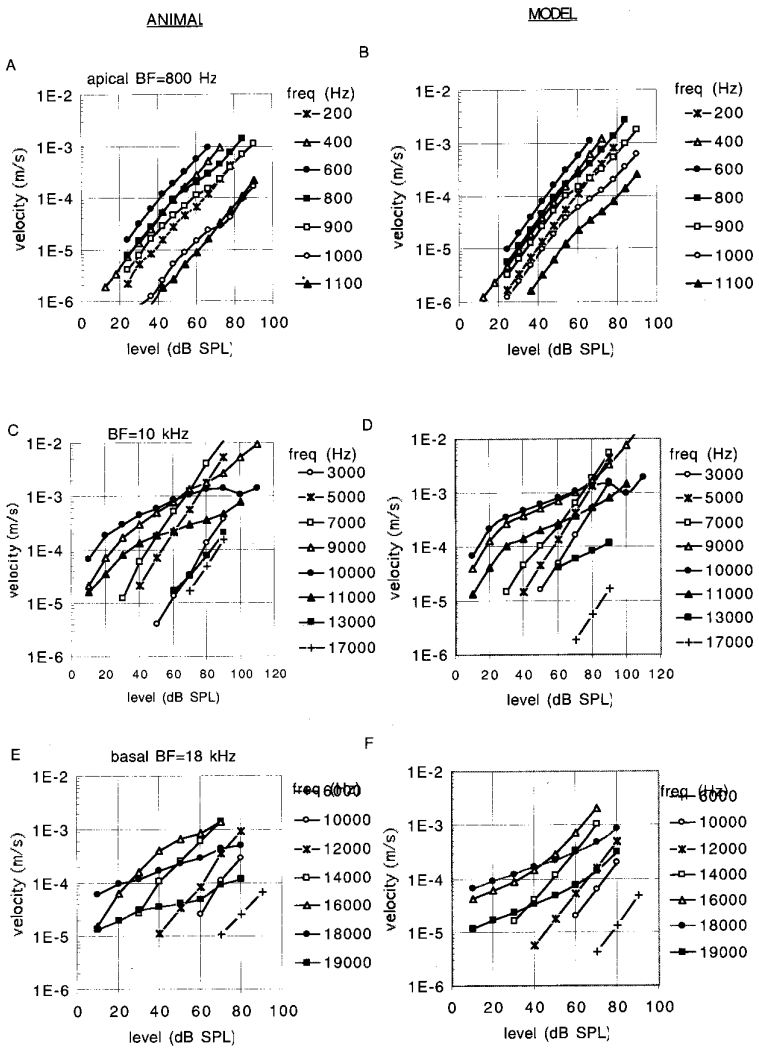

FIG. 1. Input/output functions; input is dB SPL corresponding to the stapes velocity used. Output is peak velocity of the basilar membrane (m/s). Legends show frequency of pure tone stimuli. Left column: animal observations. All data have been converted to velocity when necessary. Right column: model results corresponding to animal data immediately to left. (a) Data from Rhode and Cooper (1996, animal CH 16) for an apical site (BF ~800 Hz). (c) data from Ruggero *et al.* (1997, animal L113) (BF ~10 kHz). (e) data from Nuttal and Dolan (1996, animal GP 23-81). (b, d, and f). Corresponding algorithm results using parameters in Table I, cols 2 (CH 16), 4 and 5, respectively.

nomena including impulse responses showing frequency sweeps, filter functions that vary with signal level, two-tone suppression effects, and appropriate distortion products. All of these effects have been observed (see below) at the level of the cochlea. This model is not unique in attempting to simulate nonlinear responding, but it is probable that this is the first time that any nonlinear computational algorithm has been tested in such detail against a wide range of physical observations.

An important motivation for this research comes from the observations of Winter *et al.* (1990) and Yates *et al.* (1990) at the level of the auditory nerve. They demonstrated that the differences between low- and high-spontaneous rate fibers could be understood in terms of a nonlinear contribution to their driving force. Much of this nonlinearity probably originates at the level of the BM. It follows that the appropriate AN rate-intensity functions can only be reproduced in a composite model of the auditory periphery when an appropriate nonlinearity is present prior to the inner hair cell (IHC) stage. The principal has already been explored in computational models of IHC response (Lopez-Poveda *et al.*, 1998; Schoonhoven *et al.*, 1997).

Existing nonlinear models can be characterized in various ways; some are transmission-line models (Giguère and Woodland, 1994; Lyon, 1982) others are point models (Carney, 1993; Goldstein, 1990, 1995; Irino and Patterson, 1997). Point models simulate the response of the cochlear partition at a single site while transmission models simulate the flow of energy along the length of the partition. Most use feedback to control compression while Goldstein used an explicit compression function. Most use a single processing path while Goldstein used a dual resonance approach. The model to be evaluated below is a point model using an explicit compression function in a dual resonance configuration. It is, therefore, most similar to that of Goldstein's multiple bandpass nonlinear model and both derive their ''bandpass nonlinear'' method of compression from an idea by Pfeiffer (1970). However, there are important differences to note. For example, both use a linear ''tail'' and a nonlinear ''tip'' processing path, but the present algorithm has different center frequencies (CFs)[1] for the two paths. It is this difference that produces the shift in BF of the overall system as signal intensity rises. Like Goldstein, we use an explicit compression function, but we have found it possible to avoid the use of his expansive nonlinearity function. The use of dual resonance is also not new. Schoonhoven *et al.* (1994), for ex-

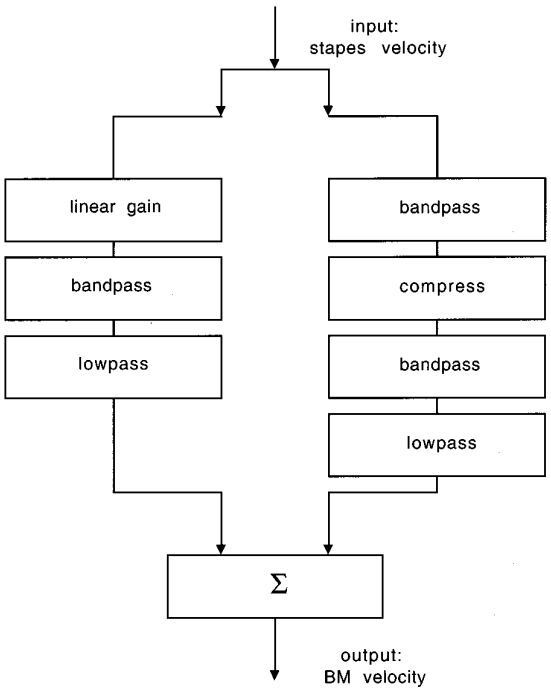

FIG. 2. Schematic diagram of the DRNL algorithm. The input is in the form of stapes velocity and is applied to two signal-processing paths: one linear and the other nonlinear. The outputs of the two paths are summed to form the output. The individual bandpass and low pass functions consist of cascades of bandpass and low pass functions (see the text and Fig. 3).

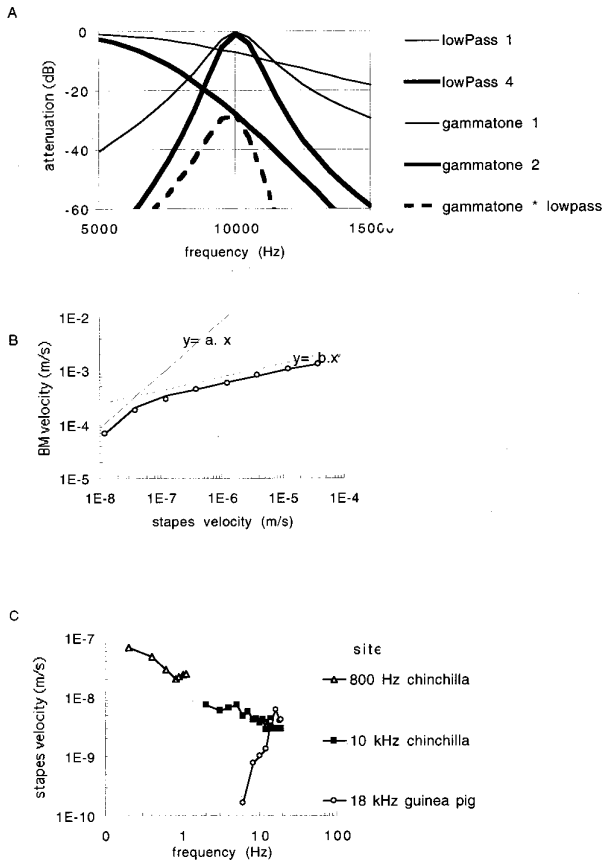

FIG. 3. (a) Examples of the gammatone and low pass filter used to construct the signal-processing units illustrated in Fig. 2. The gammatone filter shown has a CF of 10 kHz, a BW (3 dB down) of 1000 Hz and unit gain. A cascade of two gammatone filters is also shown. A low pass filter (6 dB down at 10 kHz) is also shown along with the result of four cascading identical filters. A cascade of the ''second-order'' gammatone filter and the ''fourth-order'' lowpass filter is shown to illustrate the effect of combining the components. (b) The compression effect is based on two functions (shown slightly displaced as dotted lines), one linear, y_a, and the other a power function, y_b. The individual data points are chinchilla peak velocity observations (Ruggero *et al.*, 1997, Fig. 6) and the continuous line is the model compression function, y (see the text). (c) Stapes velocity as a function of frequency used in the evaluation of the model. These data points are taken from animal measurements. White triangles: chinchilla data at the apical site (Rhode and Cooper, 1996, Fig. 4). Black squares: chinchilla data used at the ~10 kHz site (Ruggero *et al.*, 1990, Fig. 8). White circles: guinea-pig data used at the basal site (see note 3). All data points are scaled to 0-dB SPL. For modeling purposes, these are rescaled to the level of the stimulating tone.

ample, used a dual resonance model with different BFs to model threshold functions. Because they were concerned only with threshold functioning, they did not add a compression function.

The computational details of the model are presented in Sec. II while in Sec. III we evaluate the model against published input/output functions measured at three cochlear sites. In Sec. IV we review the response of the model to intensity changes, impulsive stimuli, and two-tone stimuli against physical measurements where these exist.

II. MODEL DESCRIPTION

The input to the algorithm is *stapes motion*, $x(t)$, and its output represents the vibration velocity,[2] $y(t)$, of a particular location along the cochlear partition. Each individual site is represented as a tuned system consisting of two parallel processes, one linear and the other nonlinear (Fig. 2). The linear path consists of a bandpass function, a low pass function, and a gain/attenuation factor, *g*, in a cascade. The nonlinear path is also a cascade consisting of a bandpass function, a compression function, a second bandpass function, and a low pass function, in that order. The output of the system is the sum of the outputs of the linear and the nonlinear paths. The complete unit is referred to below as a dual resonance nonlinear (DRNL) filter.

The three bandpass functions each consist of a cascade of two or more gammatone filters (Hartmann, 1997) with unit gain at CF. The low pass filters, likewise, consist of a cascade of second-order low pass filters. The low pass filters also have unit gain at low frequencies and a 6 dB down cutoff set to the CF of their matching bandpass functions. Figure 3(a) illustrates the shape of the gammatone and low pass filters both singly and in cascade. In the nonlinear path, the CFs and bandwidths (BW) of the two sets of gammatone filters are the same. The CF and BW of the gammatone filters in the linear path (CF_{nl} and BW_{nl}, respectively) are different from those in the nonlinear path (CF_{lin} and BW_{lin}). The parameters of the filters (including the number of elements in each cascade) for all three cochlear sites are specified in Table I.

The shape of the compressive function in the nonlinear path was chosen to agree with observations based on animal data. It is linear at low signal levels,

$$y_a[t] = a \cdot x[t],$$

TABLE I. Parameters of the DRNL algorithm used throughout the evaluation of the algorithm. The parameters were chosen to give an optimal fit to published input/output functions (see Fig. 1).

Simulated preparation	Apical (~800 Hz) CH 16	Apical (~800 Hz) CH 35	~10 kHz L113	Basal (~18 kHz) GP 23–81
Linear				
Gammatone cascade	2	2	2	2
Lowpass cascade	4	4	4	4
CF_{lin}	700 (Hz)	800	8000	15 500
BW_{lin}	130 (Hz)	800	2000	1400
Gain, g	83	25	100	500
Nonlinear				
Gammatone cascade	3	3	3	3
Lowpass cascade	3	3	3	3
CF_{nl}	750 (Hz)	850	9800	17 300
BW_{nl}	320 (Hz)	800	1400	1200
Compression				
Linear gain, a	50	100	12 000	90 000
Gain, b	0.008	0.01	0.057	0.03
Exponent, ν	0.25	0.25	0.25	0.25

where $x[t]$ is the output of the first filter in the nonlinear path and a is a parameter of the system. At higher signal levels, the response is nonlinear,

$$y_b[t] = b \cdot |x[t]|^{v} \operatorname{sign}(x[t]),$$

where b and v are parameters of the model. At all signal levels, the *smaller* result of the two functions is chosen, i.e.,

$$y[t] = \operatorname{sign}(x[t]) \times \operatorname{MIN}(a|x[t]|, b|x[t]|^{\nu}),$$

Fig. 3(b) compares this function to the response of the chinchilla BM to a 10-kHz (BF) tone. Note that a and b are used to control the gain of the nonlinear path. Their relative value determines the compression threshold.

The output of the nonlinear path is linear at very low signal levels but is otherwise compressive. The output of the linear path is, of course, linear at all signal levels. At low signal levels, when both outputs are linear, the response of the linear path to tones near CF is typically weak and the response of the nonlinear path dominates the output of the system as a whole. At higher signal levels, the output of the nonlinear path will eventually be less than the output of the linear path because the former is subject to compression. As a consequence, the output of the system is typically linear at very high signal levels. This can be seen most easily in Fig. 1(b), where the response to 800-Hz tones is linear up to 60-dB SPL, then shows a mild compression up to 75-dB SPL when it again becomes linear. This return to linearity at high signal levels cannot be seen in at the basal sites featured in Figs. 1(d) and (f) because it occurs at signal levels greater than those displayed.

The parameters shown in Table I were obtained in the following manner. It was assumed that the output of the system, at very low stimulus levels, is dominated by the activity of the nonlinear filter, at least for frequencies close to BF. However, at the highest signal levels, it was assumed that the linear filter dominates the output of the system. This is because the output from the nonlinear path is compressed at high signal levels while the output from the linear path continues to increase linearly with the signal level. The CF and BW of the nonlinear gammatone filters were, therefore, adjusted to fit the animal data at low stapes velocities close to BF. The exercise was then repeated for the parameters of the linear path by fitting the system output to the animal data at the highest stapes velocities and paying special attention to frequencies well away from the CF of the nonlinear path. The compression parameters were then adjusted to give the best match at BF.

Various trials were made using gammatone filters of different ''orders'' (i.e., the number of filters in the cascade). Third order was generally best for the nonlinear path and second order for the linear path. These orders are used at all three cochlear sites in order to limit the number of free parameters in the model. Similarly, a compression exponent of 0.25 was found to be generally satisfactory at all three sites. This parameter was, therefore, fixed at an early stage in the research. The order of the low pass filtering was also fixed to be the same at all three sites. In the linear path the cascade was fixed at four second-order low pass filters. In the nonlinear path, three low pass filters were used. Small improvements can be achieved by varying these parameters between sites; the establishment of uniformity across sites was primarily to reduce the number of degrees of freedom in fitting the model.

III. EVALUATION OF THE MODEL

A. Methods

The algorithm is implemented using digital IIR filters. It was implemented in the Development System for Auditory Modeling (DSAM), a computational environment for evaluating auditory models.[3]

The input to the system is stapes velocity (m/s) and was sampled at a rate of 100 kHz, except where stated. Sinusoidal stimulation was used throughout testing except when a 20-μs wide pulse was used to evaluate impulse responses. Rhode and Cooper (1996) and Nuttall and Dolan (1996) measured activity at the stapes and their data (for chinchilla and guinea pig, respectively) were used to convert input sound pressure levels to stapes velocity directly. When modeling the data of Ruggero *et al.* (1997), it was necessary to

use a middle-ear transfer function of a different chinchilla measured previously in the same laboratory (Ruggero *et al.*, 1990, Fig. 10). Stapes velocity was interpolated from their figure up to 15 kHz. Above this frequency, the stapes velocity for 15 kHz was used as an approximation. All three sets of stapes functions are plotted together in Fig. 3(c). Note that the first two functions (800-Hz and 10-kHz sites) are for chinchilla and the third function is from a guinea pig (18-kHz site).

Input/output functions were computed using 50-ms pure tones. The output from the model is given in terms of *peak velocity* (m/s) measured during the last 25 ms of the tone. For the purpose of comparison, the published animal data were converted, where necessary, from displacement measurements to velocity. The model was evaluated using *only those stimuli for which corresponding animal data exist.*

B. Input/output functions

800-Hz site. Figure 1(a) shows measurements made using chinchilla at the apex of the cochlea in a region with BF around 800 Hz (Rhode and Cooper, 1996, preparation CH16). At this site, there is only limited evidence of nonlinearity. However, it can be seen at 800, 900, 1000 Hz as a bend in the function above 50-dB SPL followed by a return to linearity at around 70-dB SPL. Figure 1(b) shows the I/O functions generated by the model using the parameters given in Table I (CH16). The nonlinear effect seen in the animal data is reproduced both quantitatively and qualitatively. At this site, the CFs of the filters in the linear and nonlinear paths are almost the same (700 and 750 Hz, respectively) and the gain of the two paths (*g* and *a*) is also roughly equal at low signal levels (see Table I). At higher signal levels, the linear path is dominant because the output from the nonlinear path is compressed. As a result, all response are linear above 80-dB SPL. Compression is only briefly visible between the compression threshold at around 50-dB SPL and the level at which the linear resonance comes to dominate all channels at around 70-dB SPL.

10-kHz site. Figures 1(c) and (d) compare the animal data as measured by Ruggero *et al.* (1997) with input/output functions of the DRNL algorithm. For clarity, only the results for multiples of 2 kHz are shown even though the complete dataset of multiples of 1 kHz were used when fitting the model. Both the animal data and the model response are compressive at 10 kHz at high signal levels while the response to low frequencies are linear at all signal levels. The animal data show a notch in the 10-kHz response at 100-dB SPL. The model reproduces this notch at the same signal intensity. Notches occur (in the model) when the outputs of the linear and nonlinear paths have similar levels but different phases. Notches are frequently seen in animal data, particularly at low BF [cf. Rhode and Cooper, 1996, Fig. 7(a)].

18-kHz site. Figure 1(e) shows data obtained in the basal turn of the guinea pig cochlea (Nuttall and Dolan, 1996, animal GP #23.81[4]) at a site with BF close to 18 kHz. Figure 1(f) shows the results from the DRNL model for the corresponding data points. The model shows a good qualitative and reasonably good quantitative match to the guinea pig

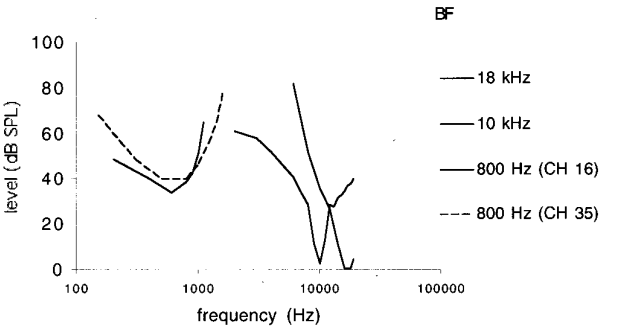

FIG. 4. Isovelocity contours for three sites computed with the algorithm and parameters given in Table I. Two contours are given for the ~800-Hz site (modeling data for animals CH 16 and 35, Rhode and Cooper, 1996). The criterion velocity is 3E-5 m/s.

data. Below BF, the functions are largely linear and above BF they are highly compressed.

The isovelocity contours for all three sites are plotted in Fig. 4 for a criterion output velocity of 3e-5 m/s (close to typical auditory nerve firing thresholds). They are not strictly comparable with each other because they are modeled on different animals from different species. However, they do show the characteristic broadening of the shape of the function at very low frequencies frequently seen in auditory nerve recordings. The high-frequency slopes may appear to be too shallow, particularly at the 10- and 18-kHz site. However, the inspection of Figs. 1(d) and (e) show that the model is underestimating the BM response at the highest frequencies. If anything, the high-frequency slope of the model result is steeper than the animal data allow.

Figure 4 shows an additional isovelocity contour derived using a second set of parameters shown in Table I. These parameters were derived by fitting I/O functions for a second animal studied by Rhode and Cooper (1996, CH 35). A simulated version of this unit will be required when evaluating the response to phase (see below).

C. Phase

In a linear system, we expect the phase to remain invariant with respect to intensity. The animal data show phase invariance at many frequencies, but not all. Phase shifts as a function of intensity are often observed for frequencies close to or above BF. The model data show a similar qualitative effect. When evaluating the phase response of the algorithm, no attempt was made to manipulate the model parameters. The parameters used are unchanged from those required to fit the I/O functions described above and are given in Table I.

Figures 5(a) and (b) compare the animal observations made by Rhode and Cooper (1996, preparation CH 35) with the algorithm results. The chinchilla phase-intensity functions were arranged by Rhode and Cooper in order of ascending frequency. The model data are presented alongside with a similar adjustment. In both animal and model data the phase is constant with respect to intensity for low frequencies. At high frequencies, an accumulating phase lag can be observed at high signal intensities. The shifts can be seen in the animal data at 800, 900, and 1000 Hz. In the model

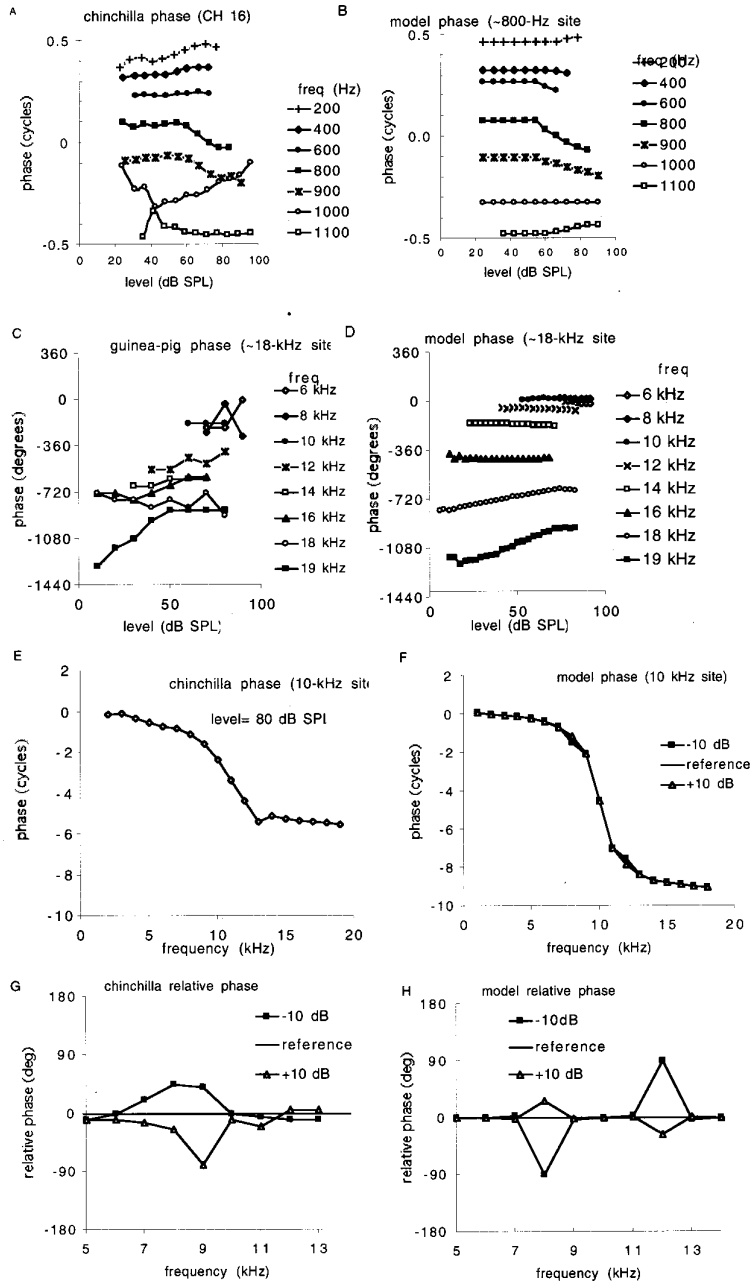

FIG. 5. A comparison of phase characteristics for animal and model data. (a) and (b) Phase as a function of the signal level at the apical, ~800-Hz site for chinchilla (Rhode and Cooper, 1996, Fig. 6, preparation CH 16) and model data, respectively. (c) and (d) Phase as a function of signal level at the basal, ~18-kHz site for a guinea pig. (Nuttall and Dolan, 1996, Fig. 8, preparation, GP#23-81) and model data, respectively. (e) and (f) Phase as a function of *frequency* at the 10-kHz site for chinchilla (Ruggero *et al.*, 1997, preparation L113, Fig. 13) and model data, respectively. Model parameters are given in Table I. The model data were computed using a 0.2-Pa, 20-μs-wide reference click (i.e., 80 dB *re* 20 μPa). Data are also shown for clicks 10 dB more and less intense than the reference. (g) and (h) Relative phase as a function of signal intensity for chinchilla (Ruggero *et al.*, 1997, preparation L113, Fig. 14). Two signal intensities are shown 10 dB above and below an 80-dB SPL reference. The corresponding model data are based on (f).

results the phase shifts can be seen at 800, 900, and 1100 Hz. At low intensities and close to BF, the model phase reflects the phase of the nonlinear path. At high intensity, the phase for all channels reflects the phase of the dominant linear path. The model phase response disagrees with the animal data at 1 kHz. This discrepancy matches the failure to reproduce the I/O function exactly in Fig. 1(b) at the same frequency.

A similar effect can be seen in the phase/intensity function for a more basal site in Fig. 5(c) (Nuttall and Dolan, 1996). The phase does not change significantly as a function of the signal level for frequencies at or below 14 kHz. Above this frequency, phase shifts as a function of intensity. Figure 5(d) shows the model response, which has a similar shift in phase at 19 kHz.

Figures 5(e) and (f) show how phase changes as a function of *frequency* at the 8-kHz site for the chinchilla and the model (Ruggero *et al.*, 1997). The chinchilla data were collected using pure tones. For computational convenience, the model data were evaluated using clicks. A 0.2-Pa, 20 μs-wide reference click (i.e., 80 dB *re* 20 μPa) was used for this purpose along with two other clicks; one 10 dB greater and one 10 dB smaller than the reference. No attempt was made to optimize the model's parameters to improve the fit. The slope of the model function is about three times that of the chinchilla data. This probably indicates that the filter bandwidths in the nonlinear path are too narrow. The total phase lag at high frequencies is also greater in the model response. The discontinuity in the animal function at 14 kHz suggests

that some of this difference may be the result of different methods used to unwrap phase.

The chinchilla data also show increasing phase lag with the stimulus level in a region below BF with a tendency in the opposite direction above BF. Figure 5(c) shows the phase response to tones at 70 and 90-dB SPL expressed relative to the response to a reference click at 80-dB SPL. No variation in phase lag is observed at BF for any intensity. The model shows a similar pattern with a change over around BF, except that the sign is reversed (i.e., increasing intensity produces less phase lag below BF). The reason for this discrepancy is unclear. Unfortunately, Ruggero *et al.* do not give full details of their phase unwrapping algorithm. Moreover, the idealized model clicks may not constitute the ideal comparison method.

D. Two-tone suppression

Two-tone suppression (2TS) was first observed at the level of the AN, where ''it has been shown that average discharge rate in response to a tone at a fiber's BF can be reduced by the addition of a second tone of appropriate frequency and sound pressure level'' (Abbas and Sachs, 1976). Ruggero *et al.* (1992) did find an analogous suppression in the velocity response of the BM. Cooper and Rhode (1996) did not find any overall suppression at an apical site but did observe a reduction in the response of the *spectral component* at BF when a second tone was added. Geisler and Nuttall (1997) also report no overall suppression at a basal site. These results leave open the question of how AN observations relate to BM observations and whether a second mechanism is required to complete the picture. However, the question for the present evaluation is simply whether the algorithm can simulate the observations made at the level of the BM.

The data of Cooper and Rhode (1996) are particularly useful in this respect because they measured both I/O and 2TS functions at the same site. Their measurement paradigm was replicated almost exactly in the evaluation of the algorithm. Figure 6(a) shows the chinchilla data (preparation CH35) at a site with an 800-Hz BF. In this experiment, a probe tone (f1) at 800 Hz (=BF) was presented at the same time as a suppressor tone (f2) at 1000 Hz. Cooper and Rhode did not observe a reduction in the amplitude of the peak displacement of Reisner's membrane, but they did observe a reduction in the *spectral component* of the displacement at f1 when f2 was added. Suppression is expressed as the ratio of the Fourier component at f1 in the presence and the absence of the f2 suppressor. Figure 6(a) shows the amount of suppression observed during an analysis using all combinations of levels (L1 and L2) of f1 and f2. Figure 6(b) shows the corresponding model functions using the parameters given in Table I (CH 35) for the 800-Hz site. The algorithm results agree with those of Cooper and Rhode in that they did not show any net suppression (not shown).

Ruggero *et al.* (1992) did observe total suppression at a site with a BF of 8.6 kHz. The peak BM velocity was observed to decrease when the suppressor was added. Figure 6(d) shows how the algorithm can simulate their main finding. In this example a probe tone is presented at BF (10 kHz)

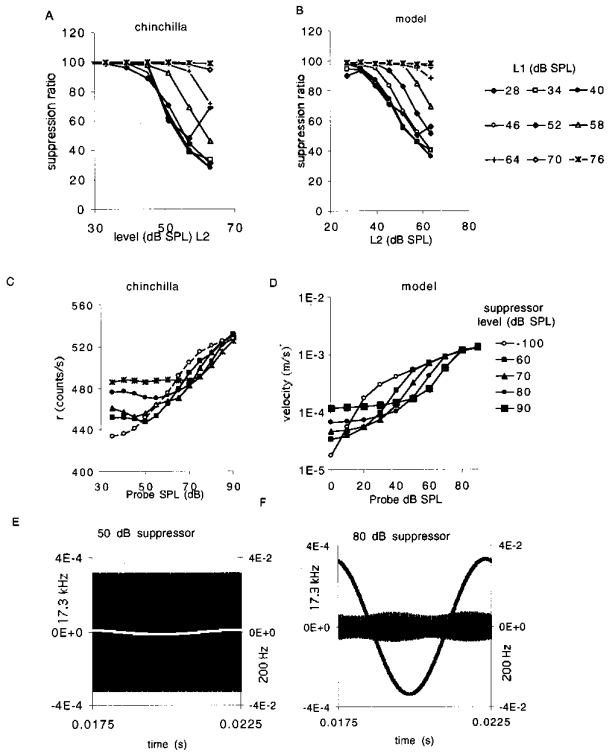

FIG. 6. Two-tone suppression. (a) and (b) Suppression ratio (suppressed/unsuppressed f1 Fourier component) for two tones (f1=800 Hz=BF and f2=1000 Hz) as a function of the level of f2 (L2). Each function represents a fixed level (L1) of f1. Animal data are taken from Cooper and Rhode (1996, Fig. 3, preparation CH 35). The model data use the parameters for CH 35 given in Table I. (c) and (d) Intensity functions for a BF probe tone alone and in the presence of a 1.2*BF suppressor tone at four suppressor levels (60–90-dB SPL). The animal observations are the average rates of gamma radiation. The model data are based on the parameters given in Table I and are for a 10-kHz BF site. (e) and (f) The effect of a 200-Hz suppressor tone (50-dB and 80-dB SPL, respectively) on the velocity response to a CF (17.3 kHz, 30-dB SPL) probe tone using the model parameters given in Table I for the 18-kHz site. The velocity has been processed using a very narrow filter at the suppressor and probe-tone frequencies and reflects the spectral component at that frequency. As the suppressor tone increases in amplitude, the velocity response to the probe tone decreases. The left *y* axis shows the scale for the probe tone; the right *y* axis shows the scale for the suppressor tone. The model data are similar to measurements reported by Geisler and Nuttall (1997, Fig. 5).

at a range of signal levels either alone or in the presence of a suppressor tone (12 kHz=1.2 BF) at four signal levels between 60- and 90-dB SPL, respectively). The suppressor tone clearly produces a reduction in the peak velocity of the output. Note that this is a net reduction in response level, and not just a reduction at the level of the spectral component. A similar pattern can be observed in the chinchilla data, where the *y* axis represents the raw physical measurement (''*r* counts'' or gamma-radiation energy) when using the Mössbauer method. Note that the algorithm parameters used are the same as those given in Table I (~10 kHz site) and are based on a different animal. This may explain differences in the signal level required to produce suppression.

Geisler and Nuttall (1997) reported observations at a basal site with a BF of around 17 kHz. They measured the effect of very low-frequency suppressors (f2) when added to

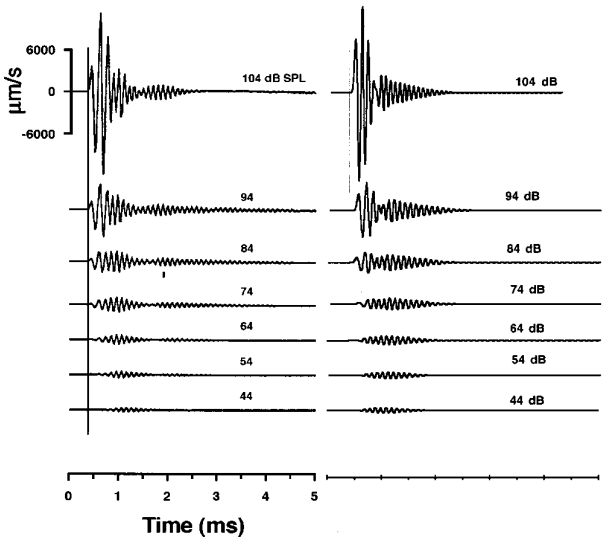

FIG. 7. (a) Basilar membrane velocity response to rarefaction clicks presented at several intensities (44–104 dB SPL), reproduced from Recio *et al.* (1998, Fig. 2). (b) Algorithm output in response to 20-μs clicks at corresponding signal levels.

high-frequency (f1) stimuli presented at BF. Like Cooper and Rhode (1996), they also found no net suppression of the combined waveform but did observe suppression of the magnitude of the spectral component at f1. The response of the algorithm was computed using the parameters given in Table I for the 18-kHz site using their paradigm.

Geisler and Nuttall do not report comprehensive I/O measurements for their animal data and the parameters used are based on the model for the Nuttall and Dolan (1996) data. The algorithm produced no net suppression when the suppressor was added and this is similar to the Geisler and Nuttall data in this respect. Figures 6(e) and (f) shows the effect of a 200-Hz suppressor at two signal levels on the Fourier component at 17.3 kHz. The resulting figures are very similar to those of Geisler and Nuttall [1997, Figs. 5(a) and (d)]. Note that there is strong suppression at the BF frequency at two different points in the suppressor cycle, soon after both the highest and lowest suppressor-tone velocities.[5] This is a characteristic feature of the Geisler and Nuttall data, although there is an asymmetry in their compression that is not present in the model output.

E. Impulse response

Figure 7 shows the impulse response data from the chinchilla investigated using 20-μs wide clicks at different signal levels (Recio *et al.*, 1998). Comparable model results are shown alongside. The parameters of the model are the same as those used above (see Table I). In both the animal and the model data, it is possible to see at high signal levels a low-frequency rapid response that is not visible at low signal levels. This results in an apparent increase in the frequency of the impulse response during the first millisecond after the click. This low-frequency component has a very low amplitude at low click levels but increases linearly with click level. The higher-frequency component is prominent at low click intensities but shows little increase in amplitude over a 60-dB increase in click level. In the model, these two components are the outputs of two separate processing paths, linear and nonlinear. The linear path has a greater bandwidth and a lower CF resulting in an impulse response that peaks earlier and has a lower instantaneous frequency. The nonlinear path has a smaller bandwidth and a higher CF. This produces an impulse response that peaks later, lasts longer, and has a higher instantaneous frequency.

There are some differences between the animal and model impulse responses. In particular, the animal impulse response lasts considerably longer and shows amplitude modulations missing in the model response. This might be explained by the fact that the model parameters in Table I were tuned to the data from another animal in another study. Alternatively, it might be the case that the animal preparation has a third very narrowly tuned resonance at around BF that is not represented in the model. It should be noted that the model uses a perfect click stimulus that is not achievable using loudspeakers. The character of the model response changes when nonidealized clicks are used containing oscillations following the leading pulse (not shown). This alone may explain the discrepancies between the model and the animal observations.

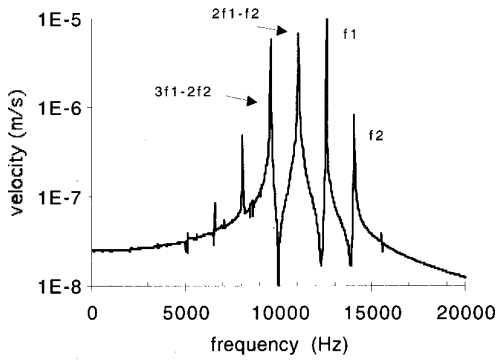

FIG. 8. Frequency spectrum of the response of the algorithm to two-tone stimuli (f1 = 12.5 kHz and f2 = 14 kHz, both at 50-dB SPL). Distortion products can be seen at 2f1-f2 and 3f1-f2. These results are comparable to measurements made of the response of the basilar membrane (Robles *et al.*, 1990, Fig. 2).

F. Distortion products

Figure 8 shows the spectrum of the response of the model to a two-tone stimulus. The parameters of the pure tones are based on an experiment by Robles *et al.* (1990). They used two tones (f1 and f2) whose frequencies were chosen such that distortion products 2f1-f2 and 3f1-2f2 would be close to the BF of the site on the basilar membrane being measured. Their data showed clear frequency components in the response at f1, f2, 2f1-f2 and 3f1-2f2. The model was evaluated using two, 50-dB SPL test tones with frequencies of 12.5 and 14 kHz, respectively. The parameters of the DRNL model are given in Table I (~10-kHz site). The model output gives a distortion product at the 2f1-f2 frequency of 11 kHz and at the 3f1-2f2 frequency of 9.5 kHz. Other distortion products are also visible.

In the model, many distortion products are generated by the compression function in the nonlinear path. However, most of these are attenuated by the second bandpass filter (after the compression function) and are lost to the output. Pfeiffer (1970) first suggested this arrangement. It has frequently been used since then, most notably by Goldstein (1990 and 1995), who has also explored the strength of combination tones as a function of the level of the primaries.

IV. DISCUSSION

The signal-processing algorithm reproduces a wide range of phenomena that are desirable in a nonlinear filter bank to be used in more general models of the auditory periphery. The I/O functions show compression near BF but are linear at low and sometimes very high frequencies. As a consequence, the I/O functions cross at high signal levels and the BF of the system shifts as signal intensity increases. At least, this is true at the basal but not at the most apical sites, where both the animal and the algorithm show largely linear responses with only small shifts in BF. The model shows phase changes with intensity at frequencies above BF. Appropriate two-tone suppression effects mirror the results of animal observations. The impulse response is characterized by a shift in frequency with time. Distortion products are limited to a range of frequencies close to the site's BF. Most importantly, the model can produce reasonably good *quantitative* fits to the animal data.

While the algorithm has been tested against observations of motion in the cochlear partition, it is important to stress that this is *not* a model of the mechanics of the system. It is a high-speed simulation of cochlear response and is not intended as an explanation of cochlear phenomena. Its primary value will rest with its potential to act as a component in much larger-scale simulations of the response of the auditory periphery to arbitrary sounds. However, the similarities between the animal and model data do permit us to conclude that the animal responds *as if* a linear and nonlinear resonance were both involved. It is also the case that recent direct measurements of the apical cochlear partition in guinea pig have found two modes of vibration, not dissimilar to those implemented above when modeling chinchilla data below 1 kHz (Hemmert *et al.*, 2000). The model offers a close quantitative fit most of the time. However, there are some difficulties in reproducing, exactly, the full range of effects in the impulse response. Some I/O functions and phase characteristics could not be replicated exactly. These failures may indicated the need for more than two resonances. The Recio *et al.* (1998) impulse response measurements at 84- and 94-dB SPL are indicative of a very narrowly tuned resonance whose impulse response substantially outlasts that of the model. The Nuttall and Dolan data give another example of this. Their detailed isovelocity functions show a number of irregularities that could never be simulated with a double resonance (Nuttall and Dolan, 1996, Fig. 2). Their results show two minima close to CF (17 and 18 kHz, respectively) that could not be attributed to measurement error. The DRNL algorithm is a good match to most of the data published at these three sites but it could still be improved given a larger database.

There are limits as to what might be expected of a point model of cochlea filtering. Distortion products are known to migrate from their point of origin to be detected at remote locations. As a consequence, measurements on the BM must reflect the sum of such indirect contributions as well as purely local effects. A point model can only estimate the local effects. A desirable future development of the model would be to find a computational procedure for simulating the migration of distortion products among simulated sites.

Our aim in this study was to show that the basic algorithm is robust enough to be considered a useful candidate for inclusion in a filter bank arrangement. However, it is not possible to use the parameters in Table I as the basis of such a filter bank. The animal data are limited to a small number of sites along the cochlear partition within the human hearing range and these data were collected from two species, chinchilla and guinea pig. Given the technical difficulties of making these measurements, it is unlikely that many more locations at the apical end of the partition will be subjected to measurement in the near future and other routes will be required to develop a full filterbank representing all sites along the partition.

Auditory nerve data are not suitable for evaluating the model directly because additional nonlinearities intervene between the basilar membrane and the AN fibers. The I/O functions observed at the level of the AN fiber are characterized by firing thresholds and saturation of firing rates that are consequences of process that occur subsequent to mechanical frequency selectivity. Nevertheless, we might be able to use AN data to help define the parameters of the DRNL if used in conjunction with a good model of the inner hair cell and the IHC/AN synapse. Studies using REVCOR methods with low BF AN fibers for reconstructing the frequency response characteristics have shown frequency sweeps in the derived impulse responses (Carney *et al.*, 1999; Lin and Guinan, 2000). These parallel those found in BM measurements (see above) and these could be used to help define some of the details of the filter bank. Similarly, it is likely that AN measurements using off-BF frequencies can be used to infer some aspects of the I/O functions of the cochlear partition using methods indicated by Winter *et al.* (1990) and Yates *et al.* (1990).

Human psychophysical data can also be used to help define the way that parameters of the algorithm need to vary across cochlear sites. Moore *et al.* (1999), and Plack and Oxenham (1998, 2000) have used masking techniques to measure the degree of nonlinearity at different stimulating frequencies. As with the animal data, compression is severe at high frequencies but much weaker at very low frequencies. The transition appears to occur between 1 and 3 kHz. At its most severe, the slope of the compression is rarely less than 0.3. These two items of data suggest useful constraints on a filter bank to simulate mechanical frequency selectivity in the human cochlea. Alcantara and Moore (2000) have recently shown that a two-filter approach similar to that proposed above can be used to successfully model psychophysical frequency–selectivity measures of intensity-dependent changes in bandwidth. This suggests that our algorithm may benefit from the availability of extensive data on human fre-

quency selectivity in the design of a full filter bank suitable for use in models of human psychophysics but consistent with direct measurements of mechanical frequency selectivity in animals.

[1]CF is used throughout to represent the center frequency of digital filters that are components of the model. This is to distinguish it from BF which is used to specify observed best frequencies either in animal data or the response of the DRNL system as a whole when being compared to animal data.

[2]The model was explored using both displacement and velocity. Velocity was finally chosen because the threshold for compression at a particular site changed less with respect to signal frequency when velocity was used. However, opinions vary on this topic and the use of velocity in this project does not preclude the use of displacement as the metric of choice when the issue is decided using empirical methods. Published displacement observations were converted to velocity using the formula, $v = 2\pi fD$, where v is velocity, f is frequency, and D is the displacement measure. Velocity is always expressed as meters per second (ms^{-1}).

[3]O'Mard and Meddis (1997). The software platform and an application that can be used to run all of the tests in this paper is available through the internet at ''http://www.essex.ac.uk/psychology/hearinglab/dsam/index.htm.''

[4]The data were taken from files available in American Institute of Physics Auxiliary Publication Service (JASA PAPS) archive using AIP Document No. PAPS JASMA-99-1556. For further details, see Nuttall and Dolan (1996) or on the internet at ''http://ftp.aip.org/epaps/journ_acoust_soc/E-JASMA-99-1556-disk/.''

[5]Stapes velocity at 200 Hz was set to a value of 2e-6 m/s at 0-dB SPL. This is an arbitrary value as no measured stapes velocity was available at a frequency 200 Hz for the guinea pig. The value chosen is the one that would be required to yield a degree of suppression comparable to that observed by Geisler and Nuttall (1997). Note that the parameters in Table I are based on a different animal and a close quantitative correspondence is not expected nor claimed.

Abbas, P. J., and Sachs, M. B. (**1976**). ''Two-tone suppression in auditory-nerve fibers: extension of a stimulus-response relationship,'' J. Acoust. Soc. Am. **59**, 112–122.

Alcantara, J. I., and Moore, B. C. J. (**2000**). ''A psychoacoustic measure of level-dependent shifts in maximum cochlear excitation,'' Br. J. Audiol. **34**, 105.

Carney, L. H. (**1993**). ''A model for the responses of low-frequency auditory-nerve fibers in cat,'' J. Acoust. Soc. Am. **93**, 401–417.

Carney, L. H., McDuffy, M. J., and Shekhter, I. (**1999**). ''Frequency glides in the impulse responses of auditory-nerve fibers,'' J. Acoust. Soc. Am. **105**, 2384–2391.

Cooper, N. P., and Rhode, W. S. (**1996**). ''Two-tone suppression in apical cochlear mechanics,'' Aud. Neurosci. **3**, 123–134.

Geisler, C. D., and Nuttall, A. L. (**1997**). ''Two-tone suppression of basilar membrane vibrations in the base of the guinea pig cochlea using 'low-side' suppressors,'' J. Acoust. Soc. Am. **102**, 430–440.

Giguère, C., and Woodland, P. C. (**1994**). ''A computational model of the auditory periphery for speech and hearing research. I. Ascending path,'' J. Acoust. Soc. Am. **95**, 331–342.

Glasberg, B. R., Moore, B. C. J., and Stone, M. A. (**1999**). ''Modelling changes in frequency selectivity with level,'' in *Psychophysics, Physiology and Models of Hearing*, edited by T. Dau, V. Hohmann, and B. Kollmeier (World Scientific, River Edge, NJ).

Goldstein, J. L. (**1990**). ''Modeling rapid wave form compression on the basilar membrane as multiple-band-pass-nonlinearity filtering,'' Hear. Res. **49**, 39–60.

Goldstein, J. L. (**1995**). ''Relations among compression, suppression, and combination tones in mechanical responses of the basilar membrane: data and MBPNL model,'' Hear. Res. **89**, 52–68.

Hartmann, W. M. (**1997**). *Signals, Sound and Sensation* (Springer-Verlag, Berlin).

Hemmert, W., Zenner, H., and Gummer, A. W. (**2000**). ''Three-dimensional motion of the organ of Corti,'' Biophys. J. **78**, 2285–2297.

Irino, T., and Patterson, R. D. (**1997**). ''A time–domain level-dependent auditory filter: The gammachirp,'' J. Acoust. Soc. Am. **101**, 412–419.

Lin, T., and Guinan, J. J. (**2000**). ''Auditory-nerve-fiber responses to high-level clicks: Interference patterns indicate that excitation is due to the combination of multiple drives,'' J. Acoust. Soc. Am. **107**, 2615–2630.

Lopez-Poveda, E. A., O'Mard, L. P., and Meddis, R. (**1998**). ''A revised computational inner-hair cell model,'' in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).

Lopez-Poveda, E. A., and Meddis, R. (**2000**) ''A computational model for simulating basilar-membrane nonlinearity in subjects with normal and impaired hearing,'' International Hearing Aid Research Conference, PA9, Lake Tahoe, California.

Lyon, R. F. (**1982**). ''A computational model of filtering, detection, and compression in the cochlea,'' IEEE Proceedings, Paris, May 1982, pp. 1282–1285.

Meddis, R., Hewitt, M. J., and Shackleton, T. (**1990**). ''Nonlinearity in a computational model of the response of the basilar membrane,'' in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos and C. D. Geisler (Springer-Verlag, Berlin).

Moore, B. C. J., Vickers, D. A., Plack, C. J., and Oxenham, A. J. (**1999**). ''Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism,'' J. Acoust. Soc. Am. **106**, 2761–2776.

Nuttall, A. L., and Dolan, D. F. (**1996**). ''Steady state sinusoidal velocity responses of the basilar membrane in guinea pig,'' J. Acoust. Soc. Am. **99**, 1556–1565.

O'Mard, L. P., and Meddis, R. (**1997**). ''Computer exploration of the auditory system with LUTear,'' Assoc. Res. Otolaryngol. Abs. **457**, p. 115.

Pfeiffer, R. R. (**1970**). ''A model for two-tone inhibition of single cochlear-nerve fibers,'' J. Acoust. Soc. Am. **48**, 1373–1378.

Plack, C. J., and Oxenham, A. J. (**1998**). ''Basilar membrane nonlinearity and the growth of forward masking,'' J. Acoust. Soc. Am. **103**, 1598–1608.

Plack, C. J., and Oxenham, A. J. (**2000**). ''Basilar membrane nonlinearity estimated by pulsation threshold,'' J. Acoust. Soc. Am. **107**, 501–507.

Recio, A., Rich, N. C., Narayan, S. S., and Ruggero, M. A. (**1998**). ''Basilar-membrane responses to clicks at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **103**, 1972–1989.

Rhode, W. S. (**1971**). ''Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique,'' J. Acoust. Soc. Am. **49**, 1218–1231.

Rhode, W. S., and Cooper, N. P. (**1996**). ''Nonlinear mechanics in the apical turn of the chinchilla cochlea *in vivo*,'' Aud. Neurosci. **3**, 101–121.

Robert, A., and Eriksson, J. L. (**1999**). ''A composite model of the auditory periphery for simulating responses to complex sounds,'' J. Acoust. Soc. Am. **106**, 1852–1864.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1990**). ''Two-tone distortion products in the basilar membrane of the chinchilla cochlea,'' in *Lecture Notes in Biomathematics*, edited by S. Levin, P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, Berlin).

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). ''Basilar membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.

Ruggero, M. A., Rich, R. L., and Shivapuja, B. G. (**1990**). ''Middle-ear response in the chinchilla and its relationship to mechanics at the base of the cochlea,'' J. Acoust. Soc. Am. **87**, 1612–1629.

Ruggero, M. A., Robles, L., and Rich, N. C. (**1992**). ''Two-tone suppression in the basilar membrane of the cochlea: mechanical basis of auditory nerve suppression,'' J. Neurophysiol. **68**, 1087–1099.

Schoonhoven, R., Keizer, J., Versnel, H., and Prijs, V. F. (**1994**). ''A dual filter model describing single-fiber responses to clicks in the normal and noise-damaged cochlea,'' J. Acoust. Soc. Am. **95**, 2104–2121.

Schoonhoven, R., Prijs, V. F., and Frijns, J. H. M. (**1997**). ''Transmitter release in inner hair cell synapses: a model analysis of spontaneous and driven rate properties of cochlear nerve fibres,'' Hear. Res. **113**, 247–260.

Winter, I. M., Robertson, D., and Yates, G. K. (**1990**). ''Diversity of characteristic frequency rate-intensity functions in guinea pig auditory nerve fibers,'' Hear. Res. **45**, 191–202.

Yates, G. K., Winter, I. A., and Robertson, D. (**1990**). ''Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range,'' Hear. Res. **45**, 203–220.