

## Research Article

# Basilar Membrane Responses to Tones and Tone Complexes: Nonlinear Effects of Stimulus Intensity

CORSTIAEN P. C. VERSTEEGH<sup>1</sup> AND MARCEL VAN DER HEIJDEN<sup>1</sup>

<sup>1</sup>*Department of Neuroscience, Erasmus MC, P.O. Box 2040, 3000 CA, Rotterdam, The Netherlands*

Received: 20 April 2012; Accepted: 20 July 2012; Online publication: 31 August 2012

### ABSTRACT

The mammalian inner ear combines spectral analysis of sound with multiband dynamic compression. Cochlear mechanics has mainly been studied using single-tone and tone-pair stimulation. Most natural sounds, however, have wideband spectra. Because the cochlea is strongly nonlinear, wideband responses cannot be predicted by simply adding single-tone responses. We measured responses of the gerbil basilar membrane to single-tone and wideband stimuli and compared them, while focusing on nonlinear aspects of the response. In agreement with previous work, we found that frequency selectivity and its dependence on stimulus intensity were very similar between single-tone and wideband responses. The main difference was a constant shift in effective sound intensity, which was well predicted by a simple gain control scheme. We found expansive nonlinearities in low-frequency responses, which, with increasing frequency, gradually turned into the more familiar compressive nonlinearities. The overall power of distortion products was at least 13 dB below the overall power of the linear response, but in a limited band above the characteristic frequency, the power of distortion products often exceeded the linear response. Our results explain the partial success of a “quasilinear” description of wideband basilar membrane responses, but also indicate its limitations.

**Keywords:** cochlear mechanics, laser interferometry, quasilinear filtering, automatic gain control

### INTRODUCTION

The cochlea performs a spectral analysis of sounds. Different frequency components excite different longitudinal locations of the cochlear partition. The nerve fibers innervating the cochlea thus become frequency channels, each fiber coding a limited band around its characteristic frequency (CF). Frequency selectivity precedes the transduction to neural signals, as shown by the sharp tuning of mechanical responses of the basilar membrane (BM). The cochlea also performs a dynamic-range compression, mapping a 100-dB range of stimulus levels onto a much smaller range of responses (Cooper 2004). Roughly speaking, each frequency band has its own dynamic compression. Even for elementary stimuli like single tones, the combination of spectral analysis and multi-channel dynamic compression gives rise to complex stimulus–response relations, which may alternatively be described as frequency-dependent growth or intensity-dependent filtering (Robles and Ruggero 2001).

For linear systems, the response to wideband stimuli can be constructed by summing pure-tone responses. This linear approach fails for strongly nonlinear systems like the cochlea. Several studies have reported frequency tuning of BM responses to wideband sounds: white noise (De Boer and Nuttall 2000, 2002; Recio-Spinoso et al. 2009), clicks (Recio et al. 1998; Recio and Rhode 2000), and tone complexes (Rhode and Recio 2001a, b). For all of these wideband stimuli, frequency tuning was found to be comparable to the single-tone case. The effects of

*Correspondence to:* Marcel van der Heijden · Department of Neuroscience · Erasmus MC · P.O. Box 2040, 3000 CA, Rotterdam, The Netherlands. Telephone: +31 (0)10 7043567; fax: +31 (0)10 7043462; URL: <http://beta.neuro.nl/research/vanderheijden/>

sound pressure level (SPL) on tuning were similar, too: with increasing SPL, tuning became less sharp and shifted toward lower frequencies. The similarity in frequency tuning was analyzed in considerable detail by De Boer and Nuttall (2002). They compared amplitude–frequency curves between white noise and single-tone BM responses, and found the differences to be minimal—provided one is allowed to freely choose the SPLs of the stimulus pairs to be compared, thereby optimizing the match.

These findings suggest that only a modest alteration of the linear approach is needed to describe BM responses to stationary wideband stimuli. The idea is to employ a different linear filter for each SPL. This modified approach is known as intensity-dependent linear or “quasilinear” filtering (De Boer and Nuttall 2002; Recio-Spinoso et al. 2009), and it is closely related to the use of level-dependent auditory filters in psychoacoustic models (Lutfi and Patterson 1984). An explicit test of the quasilinear approach to white noise responses was provided by Recio-Spinoso et al. (2009), who estimated a linear filter from one set of white-noise BM responses and used it to successfully predict responses to independent noise stimuli having the same SPL.

The quasilinear approach becomes less straightforward when attempting to predict responses to one stimulus type (e.g., tones) from the responses of another stimulus type (e.g., noise). The similar frequency selectivity between responses to these two stimulus types suggests that quasilinearity is still a useful concept, but it is not obvious how to match the stimulus levels across stimulus types. De Boer and Nuttall (2002) tackled these questions using a specific cochlear model that incorporates saturating positive feedback by outer hair cells. Although several conclusions of that study may admit generalization beyond the model assumptions (see also De Boer 1997), it is important to analyze the relation between narrowband and wideband responses independently of specific cochlear models.

We report single-tone and wideband responses of the gerbil BM and address the following questions:

- Can wideband responses be predicted from single-tone responses (or vice versa)?
- What are the limitations of the quasilinear description?

Answering these questions required the isolation and scrutiny of the nonlinear aspects of BM responses, which yielded insights beyond the original questions.

Our stimuli were irregular tone complexes and we used Fourier analysis to study stimulus–response relations (Van der Heijden and Joris 2006; Meenderink and Van der Heijden 2011). This has several advantages over Wiener kernel analysis of noise responses. The signal-to-noise ratio is better owing to the concentration of stimulus power on a few discrete components, which allowed us to use true wideband stimuli, unlike the

“composite responses” of De Boer and Nuttall (2002); the statistical significance of phase locking to the stimulus can be tested; odd order distortion products can be separated from linear stimulus components, which is not possible in the Wiener kernel approach (Johnson 1980).

## MATERIALS AND METHODS

### Animal preparation

BM motion was measured from a single location in seven cochleae of Mongolian gerbils (*Meriones unguiculatus*; female, ~60 g), with CFs of 11.2–18.2 kHz. Methods were based on Cooper and Rhode (1992). All procedures were approved by the Erasmus MC laboratory animal committee.

Animals were anesthetized with intraperitoneal injections of ketamine/xylazine (initial dose, 80 and 12 µg/g body weight, respectively; subsequent doses, 1/3 of that each hour). A heating pad kept the body temperature at 37 °C. The head was glued to a head holder. A tracheotomy allowed the animal to breathe freely. The pinna of the left ear, its cartilaginous material, and the tissue slightly posterior from the ear canal were removed, exposing the bony rim of the ear canal and the superior mastoid chambers of the left bulla. Careful opening of the bulla with a scalpel, avoiding high intensity sounds, granted a clear view on the stapes, and BM through the round window. Middle ear muscles were left intact. After tearing the round window membrane using an insect pin, reflective beads (silver coated microspheres, 25 µm, 1.0–1.2 g/cm<sup>3</sup> (hollow glass), or ~1.3 g/cm<sup>3</sup> (PMMA); Nanoparticulate Surface Adhesion Ltd., Loanhead, UK) were introduced and allowed to settle on the BM. These beads have been shown to have negligible effect on BM motion (Cooper 1999). A glass cover slip overlying the round window area stabilized the air–fluid interface.

Experiments were carried out in a 45-min window within 2 h after tearing the round window, except experiment RG11349 where single-tone data were recorded 1 h after multitone data. All single-tone responses showed near-CF nonlinearities starting below 40 dB SPL, indicating a good cochlear condition (Ren and Nuttall 2001; Ren et al. 2011).

### Recording system

A personal computer running custom software in MATLAB (The MathWorks, Natick, MA, USA) calculated stimuli and sent them to a 24-bit D/A-channel (RX6; Tucker-Davis Technologies (TDT), Alachua, FL, USA) at 111.6 kHz. Signals were fed through a programmable attenuator (PA5; TDT) and an amplifier (SA1; TDT), and played over a speaker (CF1; TDT) connected to a custom-built sound delivery probe, which was sealed to

the ear canal with Vaseline. Stimuli were compensated for the acoustic transfer of the sound delivery system using a cavity measurement and varied less than 4 dB in the 5–25 kHz range. Output spectra were monitored during the experiments using a probe microphone (40AG, G.R.A.S., Holte, Denmark).

A single-point laser vibrometer (OFV-534; Polytec, Waldbronn, Germany) with an illumination unit (VIB-A-510; Polytec), connected to a velocity decoder (VD-06; Polytec), measured a single bead's vibrations through a 5× microscope objective (M Plan Apo 5×, NA=0.14,  $f=40$ ; Mitutoyo, Veenendaal, The Netherlands). Decoder output was sampled by a 24-bit A/D-converter (RX6; TDT) at 111.6 kHz and stored on hard disk. A double-walled sound-proof booth (Acoustair, Moerkapelle, The Netherlands) and an optical table (Newport, Irvine, CA, USA) isolated the experiments from ambient noise.

## Stimuli

Each experiment started with measuring stapes velocity in response to a series of 80-dB SPL single tones in the 5–25-kHz range (700-ms duration including 2-ms rise/fall, presented each 1,000 ms; 0.01 octave steps; two repetitions). Stapes response was assumed linear; the reflex of the middle ear muscles has little effect above 5 kHz (Møller 1965; Rosowski et al. 2006). Stapes velocity, measured by pointing the laser beam to a reflective bead placed near the footplate, was used as a reference for the input to the cochlea. Thus, the phase data (Figs. 2 and 3) relate BM motion toward scala vestibuli to inward stapes motion.

Using the same stimulus protocol, single-tone BM responses were measured from 5 to 25 kHz in 0.05 octave steps at 0–80 dB SPL in 10-dB steps. The order of the tones was randomized. Amplitudes and phases of the responses were evaluated using Fourier analysis. CF was determined from the peak of the lowest SPL velocity–frequency curve, normalized to stapes motion.

The multitone stimuli consisted of 40 frequency components, irregularly spaced over 5–25 kHz, such that combination tones up to the third order never coincided with any of the 40 primary components ( $f_i \neq f_j$ ,  $f_i \neq f_j \pm f_k$ ,  $f_i \neq f_j \pm f_k \pm f_l$ ; see also Victor 1979; Van der Heijden and Joris 2006; Meenderink and Van der Heijden 2011). In order to avoid any spectral splatter in the Fourier analysis of the responses, the stimulus waveform must contain an integer number of cycles of each of the primary frequencies. Thus, all  $f_k$  must be integer multiples of a fundamental frequency,  $f_k = f_0 n_k$ . This makes the problem of finding the frequencies  $f_k$  equivalent to finding positive integers  $n_k$  obeying the same inequalities as the  $f_k$ . This nontrivial numerical problem is akin to that of the “Golomb ruler” (Bloom and Golomb 1977). We solved it by “brute force”, i.e., by testing random sequences  $\{n_k\}$ . The number of inequalities  $N_{\text{ineq}}$  grows rapidly with the

number of primaries (as a third-order polynomial). For 40 components,  $N_{\text{ineq}}=44,280$ , and it is unavoidable that sets of numbers  $\{n_k\}$  simultaneously obeying these 44,280 inequalities contain large values; typically,  $\max(n_k) \approx 125,000$ . Given the ~25-kHz range of our stimuli, this implies that the fundamental frequency  $f_0$  is in the order of 1/5 Hz, corresponding to a 5-s cycle.

Within each experiment, the frequencies and starting phases of the stimulus components were fixed, making the stimuli identical within one experiment. The multitone stimuli were varied across experiments. Duration of all multitone stimuli was three times its fundamental cycle of ~5 s, amounting to 15.05 s. The amplitude was equal for all stimulus components. Intensity was increased in 10-dB steps from 0 to 80 dB SPL per component. When playing the multitone stimulus at 80 dB SPL per tone, acoustic distortion, measured in a cavity using a ¼” microphone (Brüel and Kjaer, Naerum, Denmark) and evaluated as described in the text addressing Figure 8, was ~46 dB below the stimulus components, independent of frequency in the 5–25 kHz range.

## Analysis

Response components were evaluated using Fourier analysis. For both single-tone and multitone responses, there is an intrinsic difficulty in estimating the noise floor at a given frequency. In BM recordings, this is particularly bothersome, because floor effects at low SPLs are easily confounded with compressive growth of the response. We therefore tested individual response components for their phase locking to the stimulus. Each 15-s response waveform was divided in 10 equally long, non-overlapping sections. To analyze a given section, the samples of the other sections were set to zero, preserving the total number of samples in the waveform. The section under investigation was windowed using 25-ms cosine-squared ramps and the phases were computed from an FFT of the entire waveform. For each spectral component, a Rayleigh test ( $p < 0.001$ ) of its 10 phase values determined whether the data point was accepted.

For the calculation of explained variance of the phase predictions of the gain control model, we used the circular variance (Fisher 1993) defined as

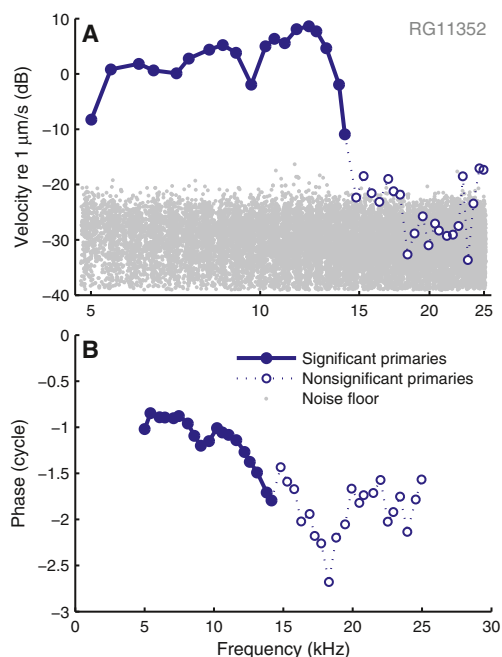
$$\sigma^2 = 1 - \left| \frac{1}{N} \sum_{k=1}^N \exp(i\phi_k) \right|,$$

where  $|\cdot|$  denotes the absolute value.

## RESULTS

### Rayleigh test

Figure 1 shows the amplitude and phase of a multitone response. The Rayleigh test clearly distinguishes be-



**FIG. 1.** Using the Rayleigh test to determine significance of phase locking to stimulus components of analog recordings. *Solid lines and solid symbols* mark the subset of data meeting the  $p < 0.001$  criterion; *dotted lines and open symbols* indicate failure to meet it. **A** and **B** BM amplitude and phase responses to a 20-dB SPL multitone stimulus.

tween significant responses (*solid circles*, above the noise floor) and nonsignificant responses (*open circles*, in the noise floor). The nonsignificant responses above 15 kHz show erratic phase behavior, consistent with random data. Insignificant components have been discarded from all data presented below.

### Isointensity curves

Normalized responses are shown in Figures 2 and 3 for two experiments (CF=11.2 and 17.0 kHz). The shapes of the isointensity curves were very similar between single-tone data (upper panels) and multitone data (lower panels). Low-SPL responses were linear, as reflected by overlying curves. As SPL increased, both single-tone and multitone responses showed similar compressive nonlinear growth ( $< 1$  dB/dB) of components near CF, visible as a reduction of the normalized magnitude (panels A and C of Figs. 2 and 3). This reduction of sensitivity with increasing SPL was accompanied by small, but systematic phase changes (panels B and D of Figs. 2 and 3).

For frequencies well above CF, the responses showed a high-frequency plateau, reflected by a flat, SPL-independent gain and a constant phase approximately equal to the middle-ear phase. The plateau extended to lower frequencies with increasing SPL.

Well below CF ( $< 0.7 \times \text{CF}$ ), the multitone responses showed a slightly expansive nonlinear growth, visible as an increased sensitivity at high SPLs ( $> 70$  dB SPL). The irregularities in amplitude and phase around 10 kHz in Figure 2 originate from the middle-ear responses and probably reflect the transition from piston-like to rocking motion of the stapes (de La Rochefoucauld et al. 2008; Ravicz et al. 2008).

Compression started at lower SPLs for multitone stimuli than for single tones, and, for example, the 80-dB SPL multitone curves showed a larger reduction in sensitivity near CF than any of the single-tone curves. Since the intensity of the multitone stimuli is specified as SPL per tone, the response to individual components was clearly affected by the presence of other components. For the cochlea with its entangled compression and frequency selectivity, it is not a priori clear how to account for such across component influences. We start by quantifying the systematic differences between single-tone and multitone responses.

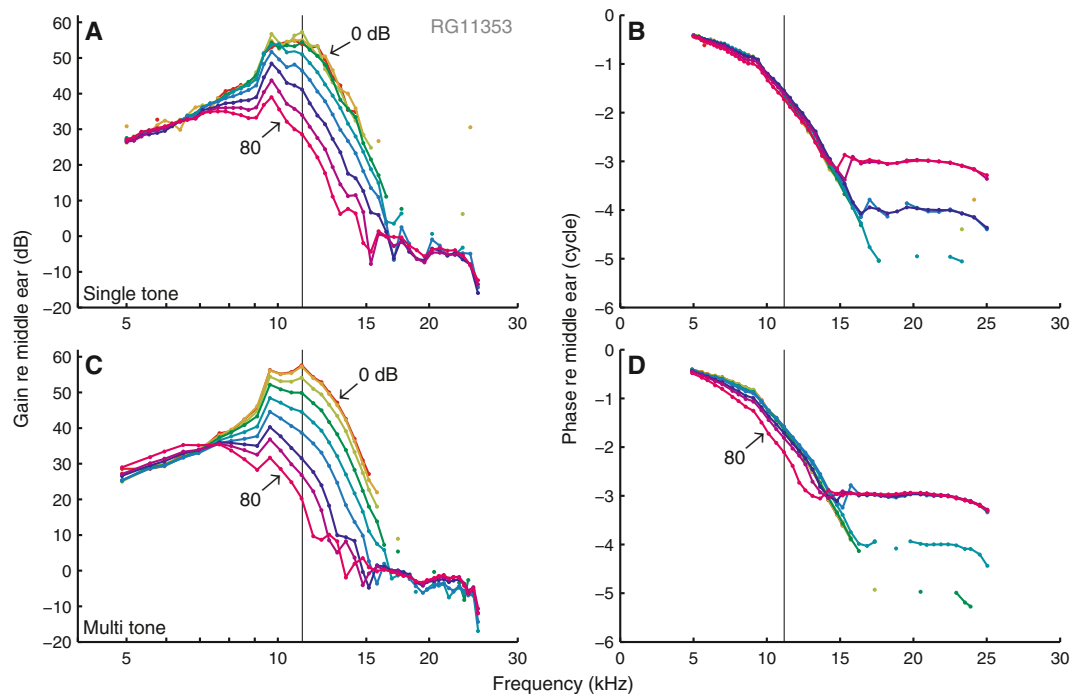
### Equivalent SPL

The similar shapes of single-tone and multitone curves (Figs. 2 and 3) suggest a description of the SPL dependence in terms of a single parameter, which we call “effective level.” By assumption, its value determines the responses to all frequency components, independent of whether they are presented consecutively or simultaneously. We aim at matching the SPLs of the two stimulus conditions (single tone and multitone) that correspond to the same effective level. Specifically, given a single-tone isointensity curve, we determined the SPL of best matching multitone response (“equivalent multitone SPL”).

The multitone data were first interpolated to a fine grid of SPLs and frequencies. Next, minimization of the sum of squared differences led to the (interpolated) multitone curve that best matched the single-tone curve. Figures 4A, B show examples of single-tone curves together with the best matching multitone curves. These examples illustrate the good match in “filter shapes” between the single-tone and multitone data. The correlation between entire families of single-tone data (such as shown in Figs. 2A and 3A) and their matched multitone counterparts was  $0.991 \pm 0.005$  ( $N=7$ ). This near equivalence in frequency tuning is a nontrivial observation in itself (see also De Boer and Nuttall 2002), which also justifies the parametric description of the level dependency in terms of equivalent SPL.

Figure 4C shows the relation between single-tone SPLs and their equivalent multitone SPLs for seven ears. The linearity of low-level ( $< 30$ – $40$  dB SPL) single-tone responses causes them to match well with *any* of





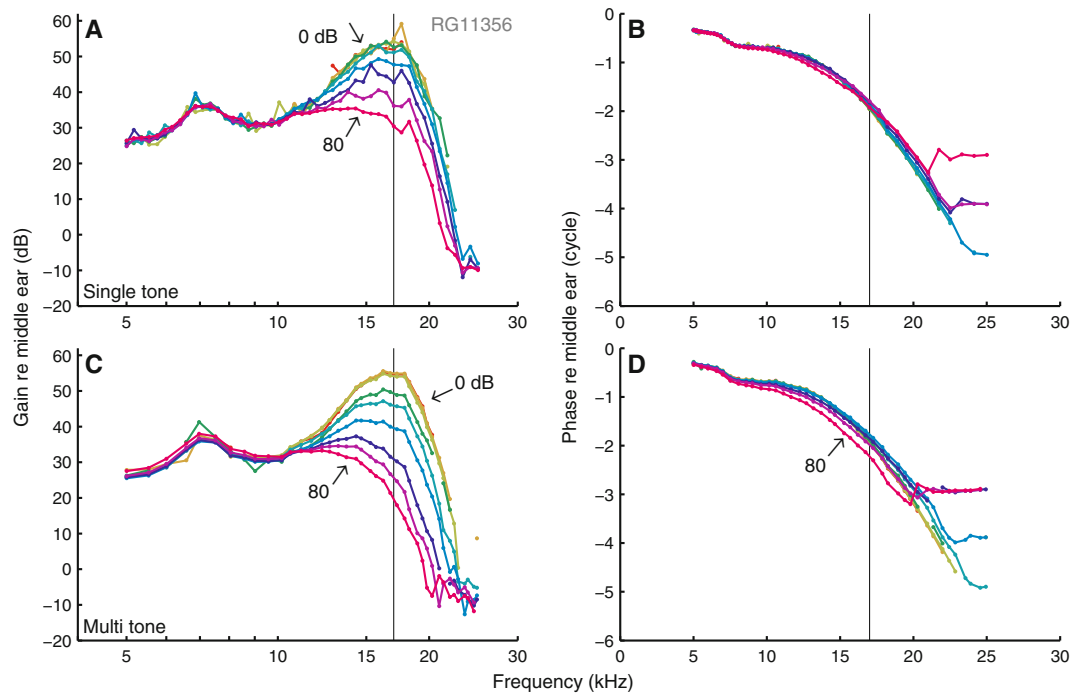
**FIG. 2.** Isointensity curves. BM amplitude (A, C) and phase (B, D), expressed *re* stapes motion, in response to single-tone and multitone stimuli, respectively. Vertical black lines indicate CF. SPL per component was varied in 10-dB steps as indicated in the graph. CF=11.2 kHz.

the low-level (<10–20 dB SPL) multitone responses. In this linear range (shaded area in Fig. 4C), the concept of an equivalent SPL became ambiguous. Outside the linear range, however, the correspondence was systematic. As expected from Figures 2 and 3, the equivalent multitone SPL grew monotonically with single-tone SPL, and the latter was higher than the former. Linear fits yielded slopes of  $0.96 \pm 0.11$  ( $N=7$ ).

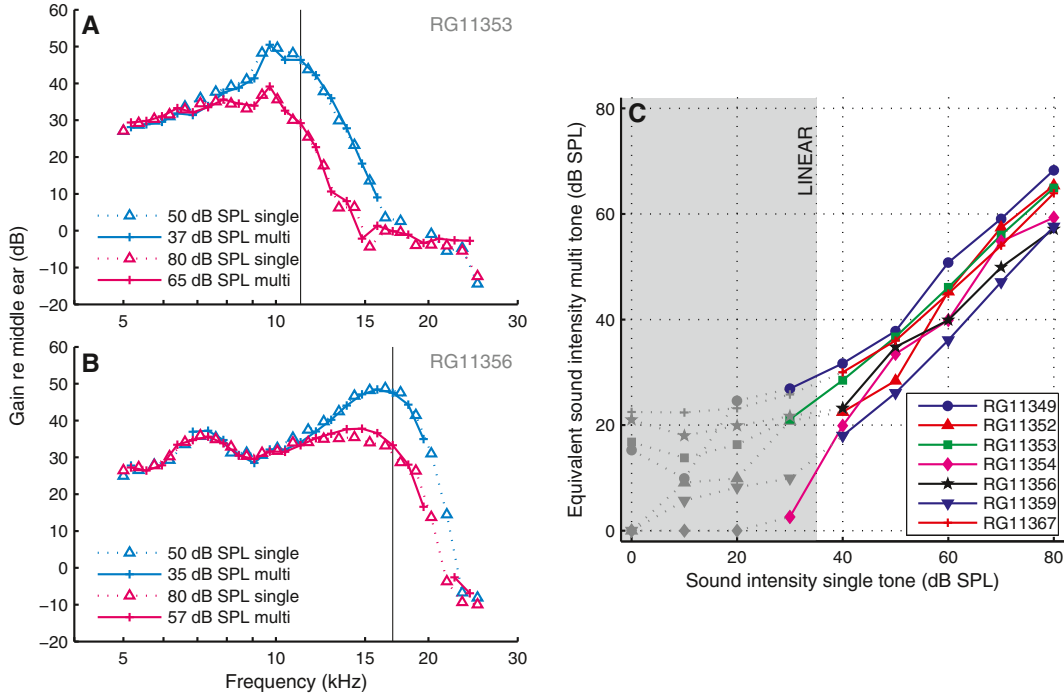
Equivalent SPL is therefore well described by a uniform offset  $\Delta$ :

$$SPL_{\text{multi}} = SPL_{\text{single}} - \Delta \quad (1)$$

This offset, however, varied considerably across experiments ( $\Delta=5.7\text{--}25.1$  dB;  $16.4 \pm 4.1$  dB). Since the total power in the 40-tone complex is 16 dB above the



**FIG. 3.** Layout as Figure 2. Different ear. CF=17.0 kHz.



**FIG. 4.** Similar frequency tuning of single-tone and multitone responses. **A, B** For two experiments, two isointensity single-tone responses and their equivalent-level multitone responses are shown. *Vertical black lines* mark CF. CF=11.2 (**A**) and 17.0 (**B**). **C** Relation between single-tone SPL and equivalent multitone SPL for seven ears. The low SPL, linear range is marked by *gray markers, dotted lines, and a gray area*.

SPL per tone, the fact that  $\Delta$  often exceeded 16 dB indicates that across-component interaction (suppression) can exceed compression (“auto-suppression”) of single-tone responses. There was a positive correlation between CF and  $\Delta$  ( $R=0.863$ ,  $p=0.013$ ). The fixed, 5-kHz low cutoff of our stimuli introduced a correlation between CF and the number of below-CF components, so part of the spread in  $\Delta$  may come from the dominance of low-side suppression at high SPLs (low-side suppression grows steeply with SPL (Rhode and Cooper 1993; Cooper 1996; Rhode 2007b)). Alternatively, the correlation of CF and  $\Delta$  may indicate that the amount of compression varied with CF. Our data do not allow a further analysis of the spread in  $\Delta$ .

### Predicting Single-Tone from Multitone Responses

The similarity between single-tone and multitone sensitivity curves (Figs. 2 and 3) suggests that one can be predicted from the other. The matching of the responses (Fig. 4) depended on a single parameter, the effective level, which, however, did not have an obvious relation to the *stimulus* intensity. For instance, if it were equal to the total SPL of the stimulus, the effective levels of single-tone and multitone stimuli should differ by a fixed 16 dB. The alternative assumption, namely, that effective level depends on the local *response* magnitude, leads to an automatic gain control scheme as shown in Figure 5A. The collection of (interpolated) multitone sensitivity

curves at various SPLs serves as the realizable filter shapes of a variable-gain filter. The gain setting of the filter is controlled by the magnitude of its own *output*, with higher output magnitudes invoking lower gains. In this scheme, the multitone sensitivity data play a double role: they provide both the collection of filter shapes and the association of output magnitude with individual shapes. For the output magnitude, we took the mean square BM displacement denoted  $\langle D^2 \rangle$ . Note that the scheme does not involve any fitting parameters. The single-tone responses follow unambiguously from the multitone responses.

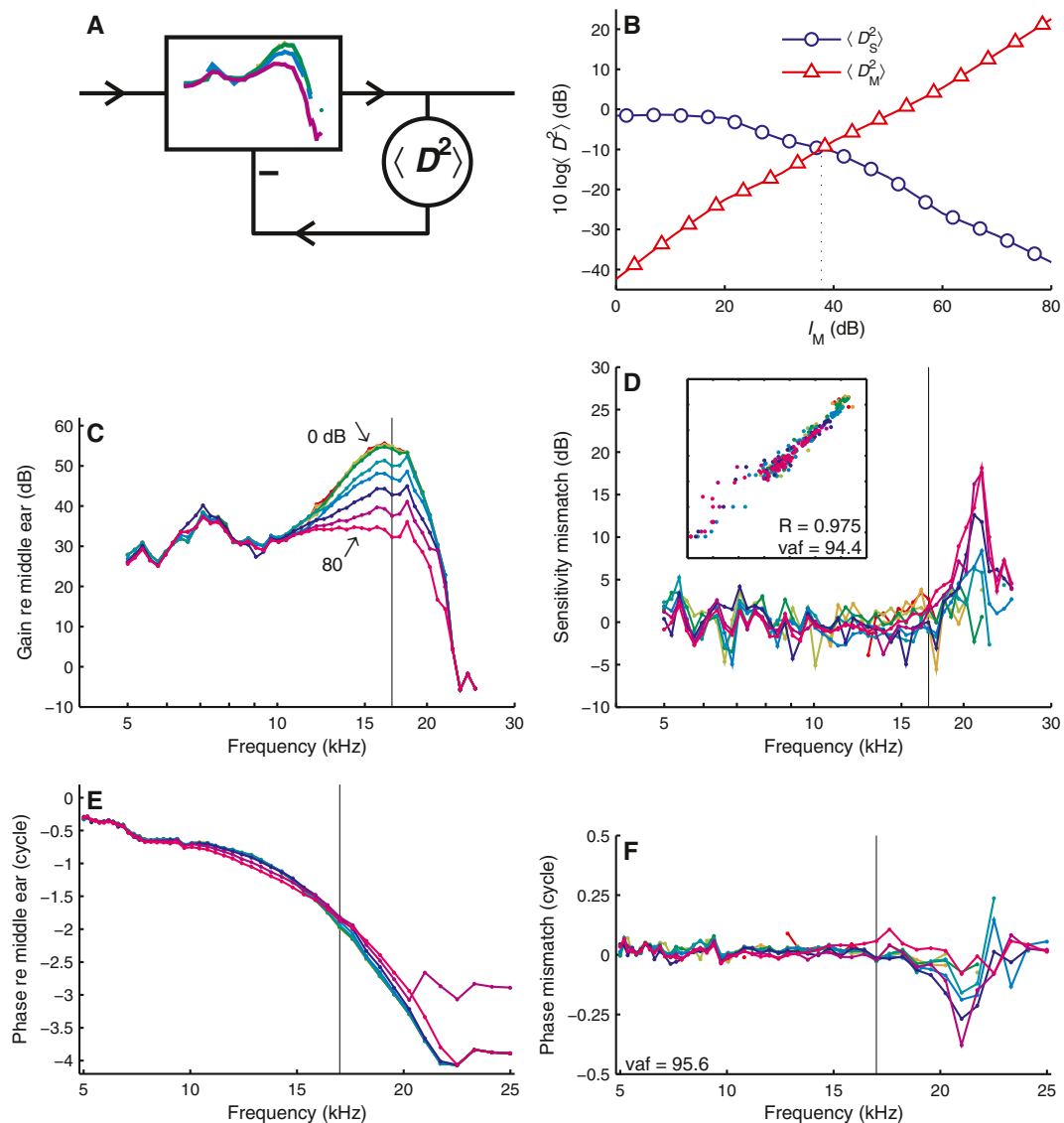
The predictions of the single-tone responses were obtained as follows. The amplitude and phase of the multitone responses (Figs. 3C, D) were cast into an SPL-dependent complex (stapes to BM) transfer function

$$H_{\text{BM}}(f, I_M), \quad (2)$$

where  $f$  is the frequency and  $I_M$  is the SPL per tone of the multitone stimuli. This equation expresses the assumption that at each fixed value of  $I_M$  the filter is linear. Linear interpolation was used to estimate  $H_{\text{BM}}$  in between the measured grid of frequencies and intensities. The next step is to apply these transfer functions to single-tone stimuli. Denoting the middle ear transfer function (relating pressure in the ear canal to stapes motion) as

$$H_{\text{ME}}(f), \quad (3)$$

the mean square displacement of BM vibration in response to a single tone of amplitude  $A$  (expressed *re*



**FIG. 5.** Predicting single-tone responses from multitone responses. **A** Automatic gain control model. The input passes a variable-gain filter that reproduces the measured multitone responses. The overall magnitude of the output,  $\langle D^2 \rangle$ , is fed back to the filter as a gain control, setting the shape of the filter. **B** Predicting a single-tone response. Passing a 50-dB near-CF single tone to all possible filters produces  $\langle D_S^2 \rangle$  as a function of the multitone level  $I_M$  associated with the respective filters. For the filter setting to be self-consistent,  $\langle D_S^2 \rangle$  must match the  $\langle D_M^2 \rangle$  associated

with the candidate filter. **C, E** Model predictions of the data in Figure 3 for normalized amplitude and phase, respectively. **D, F** Difference between predictions and data plotted versus frequency for amplitude and phase, respectively. Each curve represents a different SPL. *Inset* in **D** scatter plot of model predictions versus data for amplitude, with correlation  $R$  and variance accounted for (vaf) indicated. Variance accounted for (vaf) in phase indicated in **F**.

20  $\mu$ Pa) and frequency  $f$  is predicted to be

$$\langle D_S^2 \rangle = \frac{1}{2} |A H_{ME}(f) H_{BM}(f, I_M) / 2\pi f|^2. \quad (4)$$

At this stage, it is not clear which value of the equivalent multitone level  $I_M$  should be applied to a given single-tone stimulus. The blue circles in Figure 5B show an example of  $\langle D_S^2 \rangle$  as a function of  $I_M$ . It is a descending function because the single tone has fixed amplitude and is subjected to a gain that is negatively controlled by  $I_M$ .

For a multitone stimulus presented at  $I_M$  dB SPL per tone, the amplitude of the tones is

$$A = 10^{I_M/20} \quad (5)$$

and the mean square displacement in response to this multitone stimulus adds up to

$$\langle D_M^2 \rangle = \frac{1}{2} \sum_{k=1}^{40} |10^{I_M/20} H_{ME}(f_k) H_{BM}(f_k, I_M) / 2\pi f_k|^2. \quad (6)$$

The red triangles in Figure 5B show  $\langle D_M^2 \rangle$  as a function of  $I_M$ . It is an ascending function because BM

displacement grows with the SPL of a multitone stimulus. The true control parameter in the feedback scheme is the mean square displacement of the BM. ( $I_M$  merely serves to label the control parameter in a way that provides a direct link to the multitone data.) Therefore, the correct choice of  $I_M$  in the single-tone case (or any stationary stimulus) is that value that, when applied to the stimulus, yields the same mean square BM displacement as a multitone stimulus presented at  $I_M$  decibels. In formulas, one has to find the  $I_M$  for which  $\langle D_S^2(I_M) \rangle = \langle D_M^2(I_M) \rangle$ . This amounts to determining the intersection of the two curves in Figure 5B. For each single tone having frequency  $f$  and sound level  $L$ , this procedure yielded an equivalent multitone level  $I_M(f, L)$ . The multitone transfer functions  $H_{BM}(f, I_M(f, L))$  then provided the prediction of the response amplitude and phase for that single tone. The resulting predictions are shown in Figures 5C, E.

Figure 5D compares measured and predicted single-tone normalized amplitudes. For this ear, the predictions explained 94.4 % of the variance. For the seven ears tested, the average correlation between data and predictions was  $R=0.955\pm 0.037$ ; the explained variance,  $91.2\pm 6.9$  %. The largest prediction errors occurred above CF, where automatic gain control overestimated high-SPL response magnitudes. Figure 5F compares measured and predicted phases. Using the circular variance metric (see Materials and Methods section), the explained variance of the phase data of this ear was 95.6 %. For the seven ears tested, the explained variance was  $91.8\pm 3.4$  %. Again, the largest deviations occurred above CF at the highest SPLs, where the gain control scheme predicted too much phase lag. An alternative approach, in which the filter was controlled by BM velocity (rather than displacement) yielded only minimal ( $0.12\pm 0.23$  dB) changes in the predictions; the limited frequency span of our data prevents a critical comparison between the methods. Our choice for BM displacement was motivated by its consistency as a predictor of suppression (Fig. 2 of Cooper 1996).

### SPL-dependent changes in sensitivity and phase

Figure 6 shows the effects of SPL on sensitivity and phase, expressed as changes *re* the low-SPL (linear) responses. The use of a low-SPL reference highlights and quantifies any deviations from linearity and facilitates the comparison with suppression data (Cooper 1996). The sensitivity graphs may be viewed as input/output curves with a 1 dB/dB reference subtracted. Linear growth yields horizontal lines, and expansive and compressive growth yield positive and negative slopes, respectively. Single-tone and multitone data are shown together, but horizontally shifted

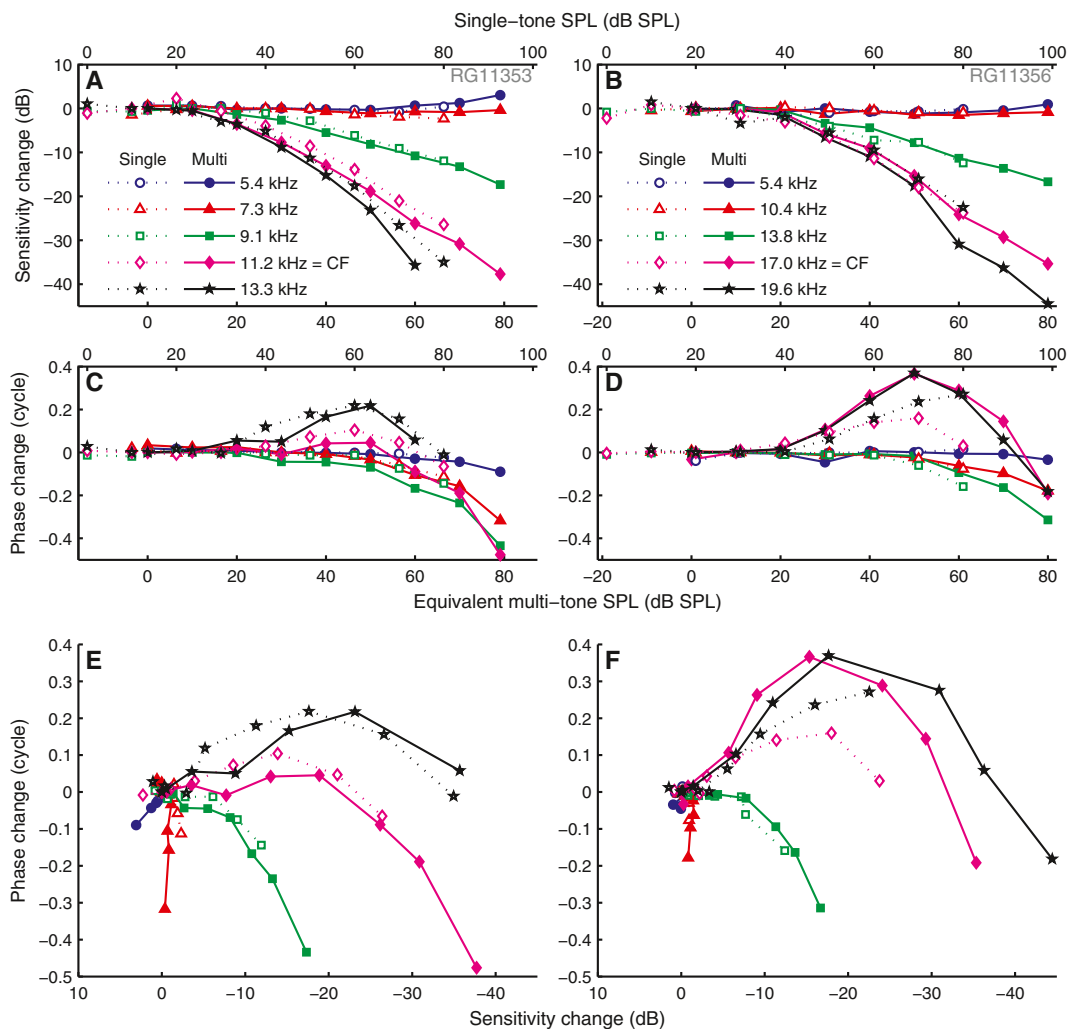
to match their effective SPLs. Well below CF, both single-tone and multitone responses showed little change in sensitivity or phase for SPLs up to  $\sim 70$  dB SPL. The high-SPL multitone stimuli, however, showed an expansive growth (Figs. 6A–D, *blue curves*). For somewhat higher frequencies, a slight phase lag occurred at the highest SPLs, without any change in sensitivity (Figs. 6A–D, *red triangles*). For frequencies closer to (but still well below) CF, sensitivity started to drop at moderate SPLs, while phase started lagging at high ( $>70$  dB) SPLs only (Figs. 6A–D, *green squares*). Near CF, sensitivity started to decrease at even lower SPLs, and dropped steeply with SPL. Phase varied nonmonotonically; the mid-SPL lead diminished with SPL and turned into a lag at high SPLs (Figs. 6A–D, *magenta diamonds*). Finally, above CF (yet below the plateau, which is not considered), the largest reductions in sensitivity occurred, accompanied by phase leads exceeding those at CF. Again, phase leads decreased at the highest SPLs (Figs. 6A–D, *black stars*).

After matching effective levels, the SPL dependence of sensitivity became similar between single-tone and multitone responses (Figs. 6A, B, *dashed lines with open symbols* versus *solid lines with solid symbols*). The phase changes of the single-tone and multitone responses also shared the same trends, but the match was less perfect (Figs. 6C, D).

An alternative way of visualizing the nonlinear effects of SPL is to plot phase change directly against sensitivity change (cf. Cooper 1996, Fig. 5). This removes effective SPL from the graph and allows a direct comparison of single-tone and multitone responses. Each frequency is now represented by a curve (Figs. 6E, F) that starts in the origin (zero change in sensitivity and phase) and moves away from it with increasing SPL. The effects described for Figures 6A–D are now seen as a lengthening of the curves and simultaneous counter-clockwise rotation around the origin with increasing frequency. Although the single-tone and multitone responses showed some differences, the overall pattern was the same.

The rotation effect was systematic and robust. Figure 7 shows phase change versus sensitivity change for all 40 frequency components of the multitone responses. All four ears show the peculiar nonlinear effect, mentioned above, of an initial expansive nonlinearity with a slight SPL-induced phase lag (Fig. 7C, *blue 5.0 kHz*) that gradually changes to a larger phase lag without a sensitivity change (Fig. 7C, *black 7.1 kHz*). At still higher frequencies, the nonlinearity adopts the more familiar compressive character. Importantly, neither the expansive nonlinearity nor the “phase-only” nonlinearity is an isolated phenomenon. They are merely stages in a gradually transforming nonlinear pattern.





**FIG. 6.** SPL-induced sensitivity and phase changes. Columns show two different experiments. *Dashed* and *solid* lines indicate single-tone and multi-tone responses, respectively. **A–D** For a subset of frequencies (indicated in the graphs), magnitude and phase data are shown after subtraction of their low-SPL, linear values. *Upper* and *lower abscissae* indicate the SPL per component of the single-tone

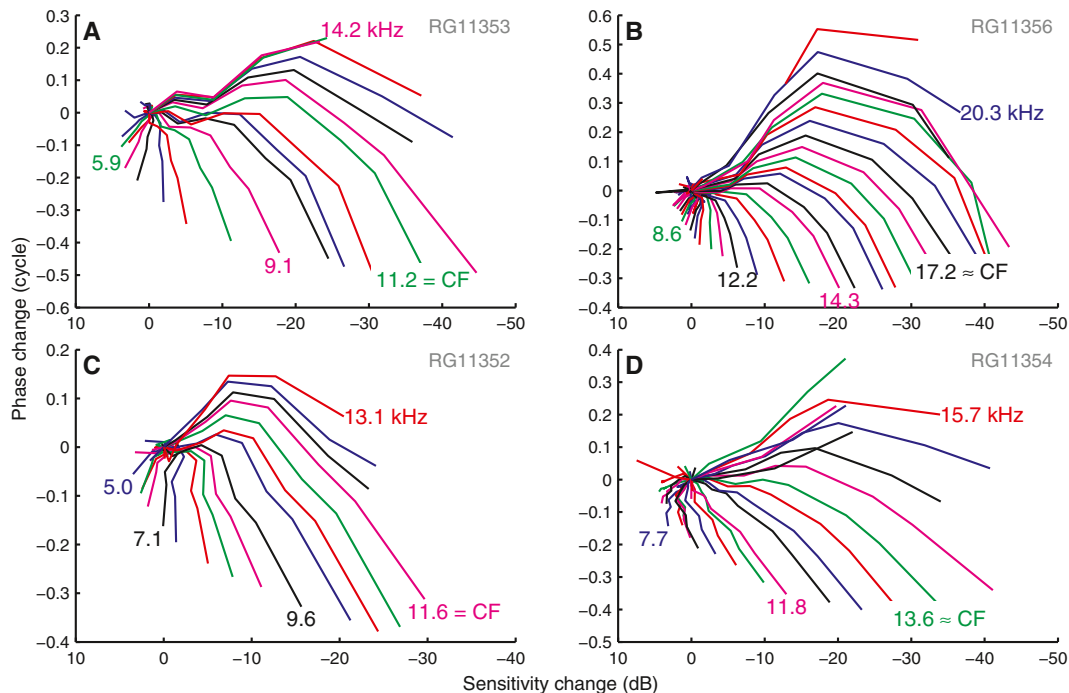
and multi-tone stimuli, respectively. Their offset was based on the matching of the effective levels (Fig. 4C). **E** The data of **A** and **C** with phase change plotted directly against sensitivity change. The *inverted abscissa* emphasizes the reduction of sensitivity with growing SPL. **F** Layout as panel **E**; data from panels **B** and **D**.

### The Limits of Quasilinearity: Distortion Products

In Recio-Spinoso et al. (2009), first-order Wiener kernels obtained from white-noise BM responses, were shown to successfully predict BM responses to independent, same-SPL noise waveforms. Because different SPLs required different linear descriptions, they described their findings as “quasi-linear filtering.” This raises the question how linear the cochlear responses to wideband stimuli really are. Two-tone stimuli can produce audible and measurable odd order distortion product (DPs). Does wideband stimulation somehow “linearize” the cochlear response or is the amount of distortion comparable to the two-tone case? Such questions cannot be answered by an analysis of white noise responses (Johnson 1980). The multi-tone stimuli of the present study, on the other

hand, allow an analysis of the third-order DPs (DP3s), since they never coincide with the primaries (see [Materials and Methods](#) section).

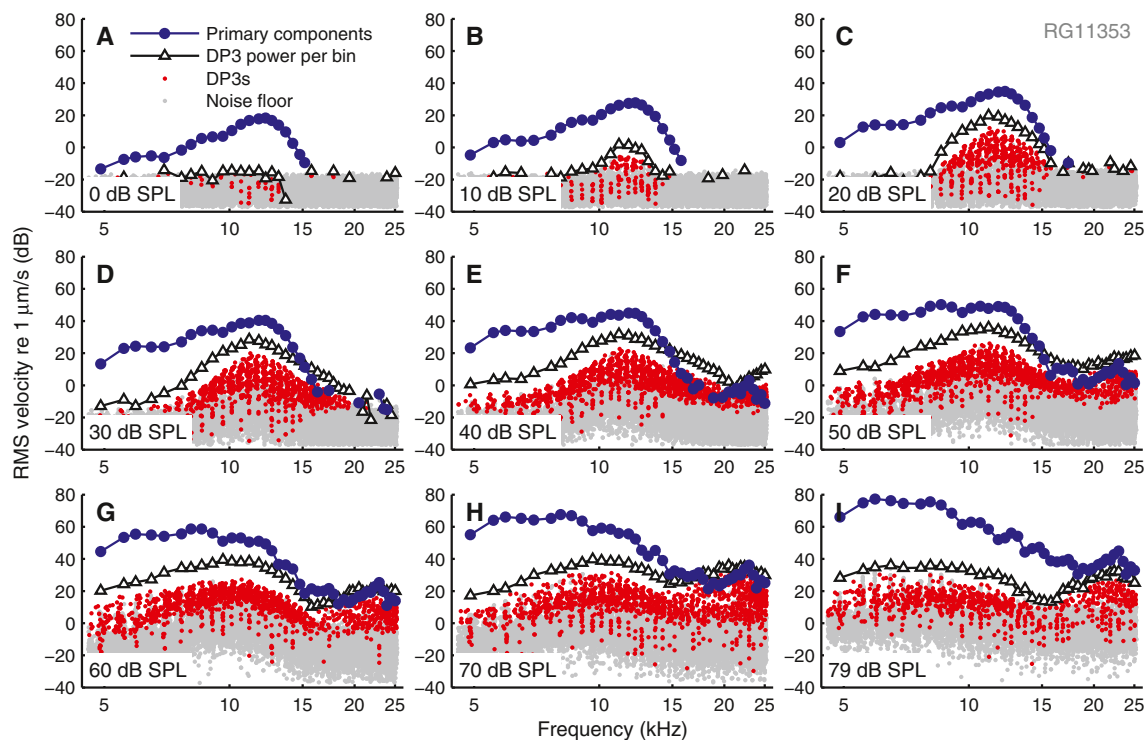
Figure 8 shows power spectra of multi-tone responses obtained at various SPLs. Blue solid circles mark the primary components; red dots mark DP3s that were significantly phase locked to the stimulus (Rayleigh test,  $p < 0.001$ ). In most cases, the DP3s outnumbered the primary components. To evaluate the relative levels of DP3s and linear response, we summed the power of all the DP3s falling in each frequency band around a primary; the geometric means between subsequent primaries served as edges of bins. The summed power of the DP3s in each bin ( $P_{DP3}$ ) is shown as triangles in Figure 8. It is to be compared with the power of the linear response components at the same frequency ( $P_{LIN}$ ).



**FIG. 7.** Phase change against sensitivity change for all frequencies of multitone recordings. Colored labels indicate the frequencies (kHz) of the closest curve with the same color. Panels contain data from different ears.

This comparison revealed a number of systematic features. At the lowest SPLs, no significant DP3s were observed. With increasing SPL, DP3s began to appear,

and their power  $P_{DP3}$  grew faster than  $P_{LIN}$ . Often, the appearance of DP3s at the lowest SPLs was a more sensitive indicator of the emergence of nonlinearity



**FIG. 8.** Linear response components and distortions. All panels show power spectra of multitone responses. Phase-locking primary components are marked by blue circles; and phase-locking DP3s by red dots. Black triangles indicate the power of DP3s summed over

the frequency band around each primary. Gray dots are the remaining components. Each panel shows data for a given SPL per component as indicated in the graphs. CF=11.2 kHz.

than the reduction of sensitivity (e.g., compare Figs. 8B and 2C). At higher SPLs, the growth of DP3s slowed down and fell behind that of  $P_{LIN}$ .

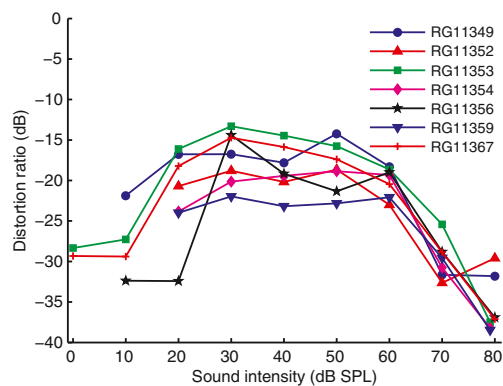
Compared to the linear response, the DP3s showed a bias towards higher frequencies, reflected by a shallower high-frequency roll-off (Fig. 8). As a result, the ratio  $P_{DP3}/P_{LIN}$  depended strongly on frequency for all SPLs. Below CF,  $P_{DP3}$  was typically 20 dB or more below  $P_{LIN}$ . Near CF,  $P_{DP3}$  came as close as 10 dB below  $P_{LIN}$  (ear RG11359 came as close as 6 dB). Above CF,  $P_{DP3}$  often even exceeded  $P_{LIN}$  (Figs. 8D–F) indicating a very strong degree of distortion at intermediate SPLs.

The overall degree of distortion (disregarding frequency) was assessed by summing the power of both the linear response components and the DP3s over all frequencies. The resulting “distortion ratio”  $\Sigma P_{DP3}/\Sigma P_{LIN}$  (shown for seven ears in Fig. 9) peaked at intermediate SPLs, but never exceeded –13 dB, consistent with a quasilinear characterization of BM responses.

## DISCUSSION

### Rayleigh test and low-SPL linearity

The Rayleigh test is routinely used to assess significance of phase locking of neural spike trains to tones. We showed (Fig. 1) that it is applicable to analog data too, providing a straightforward way to judge whether a Fourier component reflects a phase-locked response to the stimulus or is part of the noise floor. The test is generally applicable to the analysis of responses to periodic stimuli, e.g., evoked otoacoustic emissions. In cochlear mechanics, it distinguishes low-SPL floor effects from genuine compressive growth. After



**FIG. 9.** Distortion ratio versus SPL extracted from the responses to multitone stimuli for seven ears. This ratio compares the power of third-order DPs, summed over the entire frequency range, with the summed power of the linear response components. Based on the number of DPs available ( $N=21,274$ ) and the Rayleigh criterion ( $p < 0.001$ ) we required a minimum of  $2Np=43$  significant DPs for a data point to be included.

restricting the analysis to significant responses, we found linear growth at the lowest SPLs (Figs. 2 and 3) for all frequencies, in agreement with previous studies of BM responses in gerbil (Cooper 1999; Ren and Nuttall 2001).

### Comparison between wideband and tone responses

Previous studies (clicks (Recio et al. 1998), noise (De Boer and Nuttall 2000, 2002; Recio-Spinoso et al. 2009), and multitones (Rhode and Recio 2001a, b)) have shown that wideband responses share general characteristics with single-tone responses: compressive growth, widening and downward shifts of frequency tuning with SPL. The present study confirmed this similarity, analyzed it in greater quantitative detail, and explored its limits.

For each of the seven ears, the set of iso-intensity response curves obtained with tone complexes was very similar to the set of single-tone curves (Figs. 2 and 3). The individual curves matched well across the two sets when accounting for a fixed difference  $\Delta$  in effective SPL between the multitone and single-tone data (Fig. 4).  $\Delta$  varied considerably across experiments, but could be predicted in each case from the wideband responses by a simple gain control scheme without introducing any free parameters (Fig. 5).

Apart from the implications for cochlear mechanisms, our findings may inspire heuristic models of the auditory periphery. For instance, the effects of SPL on frequency resolution have long been known from behavioral data on auditory masking (Wegel and Lane 1924). Attempts to explain such data in terms of level-dependent filter shapes have raised the question how to incorporate stimulus intensity (Lutfi and Patterson 1984). The success of our gain control scheme suggests that the correct approach is to use a negative feedback configuration in which the filter is controlled by the magnitude of its own output, as previously proposed by Rosen and Baker (1994) and Lyon (2011).

### Comparison with De Boer and Nuttall (2002)

De Boer and Nuttall compared frequency selectivity of BM responses to single tones and noise bands, and analyzed their findings using a cochlear model that incorporated a saturating feedback from outer hair cells. Their amplitude–frequency curves (no phase data were presented) led them to conclude that “the response of the cochlea to [single] tones is, to a good approximation, equivalent to that of components of a noise signal” and that “insofar as frequency filtering [...] is concerned, the cochlea is not ‘more nonlinear’ or ‘less nonlinear’ for [single] tones than for noise.”

Our analyses of amplitude and phase curves (Figs. 2, 3 and 4), SPL dependence of sensitivity and phase (Figs. 6 and 7), and DP spectra (Fig. 8) confirm and generalize their conclusions. On the issue of equivalent levels, De Boer and Nuttall remarked that “stimuli of the appropriate levels have to be compared.” Interestingly, the offset between equivalent single-tone and wideband levels, which in our case varied over a ~20 dB range, appears to be less variable in their study (their Figs. 3, 4, 5, and 6 suggest a ~5-dB range). This may well be caused by the following methodological difference. The limited dynamic range of the reverse correlation analysis prevented De Boer and Nuttall from using true wideband stimuli. Instead, their amplitude–frequency curves were constructed by combining separate responses to single-octave noise bands. The authors justified using such composite curves by noting that “[in] the lower frequency ranges the cochlea is (nearly) linear.” This, however, ignores the strongly suppressive effect that low-frequency components have on the response to near-CF components. The absence of low-frequency components in De Boer and Nuttall’s CF-centered stimuli has probably weakened the effects of low-side suppression—particularly at high SPLs, because low-side suppression grows steeply with SPL. The larger bandwidth of our stimuli favored the role of low-side suppression, and (as discussed in connection with Fig. 4) this may well account for the larger variability in the offset between equivalent levels.

#### SPL-induced changes of sensitivity and response phase

In using low-SPL, linear BM responses as a reference for describing nonlinear amplitude and phase (Figs. 6 and 7) we have followed Cooper (1996), who introduced it to quantitatively compare single-tone nonlinearity (compression) with across-component nonlinearity (suppression). Cooper’s findings indicated a common origin of compression and suppression, and laid out the foundation for a unified description of cochlear nonlinearity in terms of the effects of SPL on sensitivity and phase. In several aspects, the present study is an extension of Cooper’s. The use of multitone stimuli allowed us to extend the “probe” (suppressed tone) to a much larger frequency range than the near-CF tones used previously.

Our findings confirm and generalize those of Cooper: compression and suppression are well described by the same gradual reduction of sensitivity with SPL; the nonlinear phase effects are also consistent between the two cases. In this context, our gain control scheme (Fig. 5) may be viewed as a generalization of Cooper’s analysis (Fig. 2 of Cooper 1996) of the dependence of sensitivity on the total

magnitude of BM displacement. The weakness of the scheme, i.e., the underestimation of suppression (*c.g.* sensitivity reduction) of above-CF tones, is a common finding of both studies.

The gain control scheme is based on the simplification that frequency tuning of excitation and suppression is identical. Unsurprisingly, the scheme performs worst exactly there where this assumption fails: just above CF, where suppression is stronger than expected from the steep high-frequency flanks of excitation (Schmiedt 1982). The mismatch is probably caused by the spatially distributed character of cochlear nonlinearity. Traveling waves of (above-)CF probes suffer most of their reduction at locations basal to the recording site, where any above-CF “suppressor” (including the tone itself in the single tone case) has a larger magnitude than at the recording site. Following the same line of reasoning, De Boer and Nuttall (2002) modified their cochlear model to allow the nonlinear feedback of outer hair cells to vary with cochlear location, which enabled them to account for their single-tone responses.

#### Expansive nonlinearities below CF

High-SPL multitone responses showed a small but consistent expansive nonlinearity well below CF (Figs. 6 and 7). The absence of this effect in the single-tone data is likely due to their lower effective SPL (Fig. 4C). The expansive growth reported here differs from the expansion found in the apex, which occurred above CF (Zinn et al. 2000), but appears to match the below-CF, high-SPL expansive nonlinearities reported for gerbil (Cooper 1999) and chinchilla (Rhode 2007a), which have received little attention. The joint analysis of SPL-induced sensitivity and phase effects (Figs. 6 and 7) revealed that it is not an isolated effect. In fact, with increasing frequency there is a gradual transition from expansion to compression. (The smooth transition contrasts sharply with the violent return to linearity *above* CF that has all the marks of interfering response components; see Rhode (2007a).) An interesting type of nonlinearity occurs at the transition ( $\sim 0.7 \times \text{CF}$ ) from expansion to compression: the response amplitude grows linearly with stimulus intensity, while the phase acquires a lag.

Invoking a qualitative version of “scaling invariance” (a postulated tradeoff between varying stimulus frequency and moving along the BM), the families of phase-sensitivity curves (Figs. 6E, F and 7) depict the progression of nonlinearity of a traveling wave as it propagates from base to apex. The lowest frequencies are identified with the most basal portion of the wave. Surprisingly, the SPL dependence in the initial nonlinear stage (basal portion of the wave) is expansive rather than compressive. As the wave moves on,



the expansive nonlinearity is gradually reduced and turned into compression. Over that same trajectory, larger SPLs cause larger phase lags. This implies that the basal portion of the wave slows down when the SPL is increased, in agreement with experimental data (Ren et al. 2011). The combination of these three effects of increasing SPL: (1) initial increase in gain; (2) accelerated progressive decay of amplitude; (3) reduction of wave speed; suggests that higher SPLs cause the cochlear partition to become more compliant and more resistive at the same time, as proposed in Allen (1980).

### Distortion and the limits of quasilinearity

Wideband stimulation does not appear to alter the nonlinear character of BM responses. Wideband responses grew as compressively with stimulus amplitude as did single-tone responses, and showed the same SPL-induced changes in frequency tuning and response phase; the relative power of DPs (Fig. 8) was comparable to that observed with two-tone stimuli (Robles et al. 1997; Ren 2004). For moderate SPLs, the DP power density even exceeded that of linear response components in a limited band above CF (Figs. 8D–F). The occurrence of odd-order DPs at low SPLs was a very sensitive probe of the onset of compression (Fig. 8B). Despite all of these strongly nonlinear aspects, the linear components always dominated the *total power* of the wideband responses (Fig. 9). This global dominance of the linear response explains the success of quasilinear predictions of BM responses to white noise (De Boer and Nuttall 2000; Recio-Spinoso et al. 2009). Even in the mid-intensity range, the distortion ratio never exceeded  $-13$  dB (Fig. 9), so at most 5 % of the response power originated from DPs. Thus, at least 95 % of the variance of the waveform was explained by the best linear model. The lower fractions ( $\sim 90$  %) of explained variance reported in Recio-Spinoso et al. (2009) may reflect a higher degree of nonlinearity in chinchilla than in gerbil and/or a suboptimal linear model.

### CONCLUSIONS

Expansive nonlinearities in the “low-frequency tail” of iso-intensity curves are not an isolated phenomenon, but represent one end of a systematic, gradual transformation of cochlear nonlinearity with frequency and place. The other, more familiar end, is the compression in the peak region of the traveling wave.

Single-tone and wideband BM responses are very similar in terms of frequency tuning and its dependence on SPL, compressive nonlinearity, and SPL-induced phase changes. The difference in their effective levels can be accounted for by local negative

feedback without introducing any additional parameters. The largest mismatch between narrowband and wideband responses occurred just above CF, and our analysis suggests that this mismatch is caused by the same nonlocal aspects of cochlear nonlinearity that underlie differences in frequency tuning between excitation and suppression. The magnitude of third-order DPs evoked by wideband stimulation is comparable to DP magnitude observed in customary two-tone paradigms. The DP ratio (total power of DPs *re* total power of linear response) peaked at moderate ( $\sim 40$  dB SPL) levels, but never exceeded  $-13$  dB, which explains the success of quasilinear descriptions of peripheral frequency selectivity. In a restricted band above CF, however, the power of DPs easily exceeded that of the linear response components. Stimulus bandwidth by itself has little effect on the strongly nonlinear character of the cochlear response.

### ACKNOWLEDGMENTS

The authors thank Nigel Cooper, Wei Dong, Lisa Olson, and Alberto Recio for their generous help with setting up the measurements. This work is supported by Netherlands Organization for Scientific Research (NWO) grants 818.02.007 and 834.10.005 (Earth and Life Sciences).

#### *Open Access*

This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

### REFERENCES

- ALLEN JB (1980) Cochlear micromechanics—a physical model of transduction. *J Acoust Soc Am* 68:1660–1670
- BLOOM GS, GOLOMB SW (1977) Applications of numbered undirected graphs. *Proc IEEE* 65:562–570
- COOPER NP (1996) Two-tone suppression in cochlear mechanics. *J Acoust Soc Am* 99:3087–3098
- COOPER NP (1999) Vibration of beads placed on the basilar membrane in the basal turn of the cochlea. *J Acoust Soc Am* 106:L59–L64
- COOPER NP (2004) Compression in the peripheral auditory system. In: Bacon SP, Fay RR, Popper AN (eds) *Compression: from cochlea to cochlear implants*. Springer, New York, pp 18–61
- COOPER NP, RHODE WS (1992) Basilar membrane mechanics in the hook region of cat and guinea-pig cochleae: sharp tuning and nonlinearity in the absence of baseline position shifts. *Hear Res* 63:163–190
- DE BOER E (1997) Connecting frequency selectivity and nonlinearity for models of the cochlea. *Audit Neurosci* 3:377–388
- DE BOER E, NUTTALL AL (2000) The mechanical waveform of the basilar membrane. III. Intensity effects. *J Acoust Soc Am* 107:1497–1507



- DE BOER E, NUTTALL AL (2002) The mechanical waveform of the basilar membrane. IV. Tone and noise stimuli. *J Acoust Soc Am* 111:979–989
- DE LA ROCHEFOUCAULD O, DEGRAEMER WF, KHANNA SM, OLSON ES (2008) Simultaneous measurements of ossicular velocity and intracochlear pressure leading to the cochlear input impedance in gerbil. *J Assoc Res Otolaryngol* 9:161–177
- FISHER NI (1993) Statistical analysis of circular data. Cambridge University Press, Cambridge
- JOHNSON DH (1980) Applicability of white-noise nonlinear system analysis to the peripheral auditory system. *J Acoust Soc Am* 68:876–884
- LUTFI RA, PATTERSON RD (1984) On the growth of masking asymmetry with stimulus intensity. *J Acoust Soc Am* 76:739–745
- LYON RF (2011) A pole-zero filter cascade provides good fits to human masking data and to basilar membrane and neural data. In: Shera CA, Olson ES (eds) *What fire is in mine ears: progress in auditory biomechanics*. American Institute of Physics, Melville, pp 224–230
- MEENDERINK SWF, VAN DER HEIJDEN M (2011) Distortion product otoacoustic emissions evoked by tone complexes. *J Assoc Res Otolaryngol* 12:29–44
- MØLLER AR (1965) An experimental study of the acoustic impedance of the middle ear and its transmission properties. *Acta Otolaryngol (Stockh)* 60:129–149
- RAVICZ ME, COOPER NP, ROSOWSKI JJ (2008) Gerbil middle-ear sound transmission from 100 Hz to 60 kHz. *J Acoust Soc Am* 124:363–380
- RECIO A, RHODE WS (2000) Basilar membrane responses to broadband stimuli. *J Acoust Soc Am* 108:2281–2298
- RECIO A, RICH NC, NARAYAN SS, RUGGERO MA (1998) Basilar-membrane responses to clicks at the base of the chinchilla cochlea. *J Acoust Soc Am* 103:1972–1989
- RECIO-SPINOSO A, NARAYAN SS, RUGGERO MA (2009) Basilar membrane responses to noise at a basal site of the chinchilla cochlea: quasi-linear filtering. *J Assoc Res Otolaryngol* 10:471–484
- REN T (2004) Reverse propagation of sound in the gerbil cochlea. *Nat Neurosci* 7:333–334
- REN T, NUTTALL AL (2001) Basilar membrane vibration in the basal turn of the sensitive gerbil cochlea. *Hear Res* 151:48–60
- REN T, HE W, PORSOV E (2011) Localization of the cochlear amplifier in living sensitive ears. *PLoS One* 6:e20149
- RHODE WS (2007A) Basilar membrane mechanics in the 6–9 kHz region of sensitive chinchilla cochleae. *J Acoust Soc Am* 121:2792–2804
- RHODE WS (2007B) Mutual suppression in the 6 kHz region of sensitive chinchilla cochleae. *J Acoust Soc Am* 121:2805–2818
- RHODE WS, COOPER NP (1993) Two-tone suppression and distortion production on the basilar membrane in the hook region of cat and guinea pig cochleae. *Hear Res* 66:31–45
- RHODE WS, RECIO A (2001A) Basilar-membrane response to multi-component stimuli in chinchilla. *J Acoust Soc Am* 110:981–994
- RHODE WS, RECIO A (2001B) Multicomponent stimulus interactions observed in basilar-membrane vibration in the basal region of the chinchilla cochlea. *J Acoust Soc Am* 110:3140–3154
- ROBLES L, RUGGERO MA (2001) Mechanics of the mammalian cochlea. *Physiol Rev* 81:1305–1352
- ROBLES L, RUGGERO MA, RICH NC (1997) Two-tone distortion on the basilar membrane of the chinchilla cochlea. *J Neurophysiol* 77:2385–2399
- ROSEN S, BAKER RJ (1994) Characterising auditory filter nonlinearity. *Hear Res* 73:231–243
- ROSOWSKI JJ, RAVICZ ME, SONGER JE (2006) Structures that contribute to middle-ear admittance in chinchilla. *J Comp Physiol A Sens Neural Behav Physiol* 192:1287–1311
- SCHMIEDT RA (1982) Boundaries of two-tone rate suppression of cochlear-nerve activity. *Hear Res* 7:335–351
- VAN DER HEIJDEN M, JORIS PX (2006) Panoramic measurements of the apex of the cochlea. *J Neurosci* 26:11462–11473
- VICTOR JD (1979) Nonlinear systems analysis: comparison of white noise and sum of sinusoids in a biological system. *Proc Natl Acad Sci U S A* 76:996–998
- WEGEL RL, LANE CE (1924) The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. *Phys Rev* 23:266–285
- ZINN C, MAIER H, ZENNER H, GUMMER AW (2000) Evidence for active, nonlinear, negative feedback in the vibration response of the apical region of the in vivo guinea pig cochlea. *Hear Res* 142:159–183