JARO
Journal of the Association for Research in Otolaryngology

Distortion Product Otoacoustic Emissions Evoked by Tone Complexes

SEBASTIAAN W. F. MEENDERINK[1] AND MARCEL VAN DER HEIJDEN[1]

[1]*Department of Neuroscience, Erasmus MC, Dr. Molewaterplein 50, Rotterdam, 3015 GE , The Netherlands*

ABSTRACT

Distortion product otoacoustic emissions (DPOAEs) are traditionally evoked by two-tone stimuli. In this study, emission data from Mongolian gerbils are reported that were obtained with stimuli consisting of six to 10 tones. The stimuli were constructed by replacing one of the tones of a tone pair by a narrowband multitone complex. This produced rich spectra of the ear canal sound pressure in which many of the third-order DPOAEs originated from the interaction of triplets of stimulus components. A careful choice of the stimulus frequencies ensured that none of these DPOAE components coincided. Three groups of DPOAEs are reported, two of which are closely related to DPOAEs evoked by tone pairs. The third group has no two-tone equivalent and only arises when using a multitone stimulus. We analyzed the relation between multitone-evoked DPOAEs and DPOAEs evoked by tone pairs, and explored the new degrees of freedom offered by the multitone paradigm.

Keywords: DPOAEs, otoacoustic, gerbil, *Meriones unguiculatus*, tone complex, zwuis

INTRODUCTION

To cope with a dynamic range exceeding 100 dB, several stages of auditory processing employ compres-

Correspondence to: Marcel van der Heijden · Department of Neuroscience · Erasmus MC · Dr. Molewaterplein 50, Rotterdam, 3015 GE, The Netherlands. URL: http://neuro.nl/research

sive transformations. In the cochlea, for example, the growth rate of mechanical responses to pure tones can be as low as 0.2 dB per decibel increase of the stimulus intensity (Rhode 2007). This mechanical nonlinear transformation has side effects. The incoming signals are distorted and additional, nonstimulus components are produced.

When the stimulus consists of two pure tones having frequencies f_1 and f_2, their intermodulation results in multiple distortion products (DPs) of which the third-order ("cubic") difference tones (CDTs) at frequencies $2f_1-f_2$ and $2f_2-f_1$ are best known. Distortion products are often audible; reports of these "subjective tones" date back several centuries (see Plomp 1965, for a historical perspective). Their audibility suggests that they are associated with propagating components in the inner ear, which is supported by psychophysical (Smoorenburg 1972) and electrophysiological studies (Goldstein and Kiang 1968; Kim et al. 1980).

Under favorable stimulus conditions, DPs return to the ear canal, where they can be recorded with a sensitive microphone. Reports of acoustic byproducts of nonlinear auditory processing date back many decades (Wever and Lawrence, 1954; von Békésy, 1960), but their systematic study started with the pioneering work on otoacoustic emissions by Kemp and colleagues (reviewed in Kemp 2008). Distortion products measured acoustically in the ear canal are called distortion-product otoacoustic emissions.

Virtually all studies of DPOAEs employ two stimulus tones to evoke them. The few studies that did address three-tone DPOAEs merely pointed out their occurrence (Kemp and Brown 1986; Kemp 1998), or analyzed them in the narrow context of low-frequency biasing (Bian and Scherrer 2007; Marquardt et al. 2007) or time-varying suppression (Meenderink and van der

Heijden 2010). It occurred to us that there is no *a priori* reason for such constraints, and that the systematic study of multitone (>2) emissions may enlarge the scope of DPOAE research beyond what is possible with tone pairs. Mathematically, third-order distortions arise from those terms in the Fourier expansion of the distorted stimulus that involve the interaction of three stimulus components (Schroeder 1975). If f, g, and h are stimulus frequencies, such third-order terms produce distortions at frequencies $\pm f \pm g \pm h$. Since f, g, and h need not be all different, these third-order terms include harmonics $3f$, (i.e., $f = g = h$), as well as CDTs at $2f - g$ (for which $f = h$). Thus, third-order DPs evoked by tone pairs are a limited subset of the more general case of interacting triplets of tones.

This study explores the generation of third-order DPOAEs by stimuli that contain more than two frequencies. We replaced one of the primaries of a tone pair by a narrowband tone complex, resulting in a stimulus that contained multiple tone triplets at once. The number of third-order DPs grows as the third power of the number of primary components, and it can easily happen that several DPOAE components coincide. We avoided such coincidences by a careful choice of the primary frequencies. The resulting uniqueness of third-order DPOAE frequencies greatly facilitated their analysis. We observed three groups of DPOAEs evoked by the new stimulus paradigm, two of which are closely related to DPOAEs evoked by tone pairs. The third group does not have a two-tone equivalent. We analyzed the relation between multitone-evoked DPOAEs and DPOAEs evoked by tone pairs, and explored the new degrees of freedom offered by the multitone paradigm.

MATERIALS AND METHODS

Animal Preparation

Recordings were made from adult Mongolian gerbils (*Meriones unguiculatus*; three males and 12 females; 21 ears; body weight, 46–66 g). Animals were anesthetized by intraperitoneal injection of ketamine/xylazine solution (effective dosage, 80 and 12 µg/g body weight, respectively). Supplementary anesthesia was administered subcutaneously at $\sim^1/_3$ of the initial dose in 1-h intervals. A small metal rod was attached to the dorsal surface of the skull and was used to fix the head of the animal. Body temperature was maintained at 37°C using a thermocontrolled heating pad. The pinna and cartilaginous ear canal were removed and a custom-built probe was sealed over the bony ear canal using Vaseline. In eight of the animals a small hole was drilled in the wall of the inferior posterior mastoid chamber of the ipsilateral bulla to prevent pressure buildup in the middle ear cavity. We did not notice apparent differences between record-

ings with or without such venting, and no distinction between the data from these two conditions is made. Animal procedures were in accordance with guidelines provided by the animal committee of the Erasmus MC.

Stimuli

To evoke otoacoustic emissions, acoustic stimuli that consisted of multiple frequency components (each with a random starting phase) were used. For convenience of description, the components are separated into two constituents: (1) a single frequency component and (2) a tone complex consisting of M tones. In this study, M varied between 5 and 9. The frequency of the single component will be denoted by f; the frequencies of the tone complex by g_1, $g_2 \ldots g_M$. The stimulus frequencies were chosen such that all possible difference and sum frequencies were unique (Victor et al. 1977; van der Heijden and Joris 2003; 2006). Following van der Heijden and Joris (2003), we will refer to this stimulus property as "zwuis". Nonlinear interaction between any combinations of the stimulus components can give rise to third-order DPOAEs, and different groups of third-order DPOAEs will be identified and analyzed in Results.

The primary frequencies were further restricted as described in Table 1 (the terminology used in this table anticipates the naming of the DPOAE groups described in Results). This restriction served to avoid the coincidence of DP components across groups. Combined, the restrictions in the choice of stimulus frequencies ("zwuis" and "periodicity") ensure that all third-order DPOAEs of the type $g_i \pm g_j \pm f$ are unique and never equal to a primary component. Consequently, based on its frequency, each DPOAE component can be unambiguously identified with the unique triplet (f, g_i, g_j) of stimulus components whose nonlinear interaction produced it.

TABLE 1

Relation between the number of periods for the stimulus components and the different groups of DPOAEs that result from them

Frequency	Number of periods
g; components from tone complex	$8k+4$
f; single component	$8k+2$
$2f - g_i$; 'far' group	$8k$
$g_i + g_j - f$; 'near' group	$8k+6$
$f + g_i - g_j$; 'sideband' group	$8k+2$
$g_i + g_j - g_k$; within-complex OAEs	$8k+4$

All stimulus components were chosen to have an even number of periods (rows 1–2, with integer k) over an integer number of sample points (typically 97,656). This results in the third-order DPOAEs of the different groups (rows 3–5; see text) to be periodic over the same number of samples, and never coincide with any stimulus component. The choice for the number of periods also rules out any coincidences across members of the three groups and the fourth group (bottom row) of the third-order DPOAEs arising from interaction of triplets (g_i, g_j, g_k) within the tone complex

The tone complex and the single frequency component were presented over separate D/A channels (Tucker-Davis Technologies RP2.1). To minimize transients, stimuli were gated using a raised cosine window (10–90% in 10 ms). The output of each channel was fed through a stereo power amplifier (TDT SA1) and broadcast from a separate driver (2× TDT CF1 or 2× Visaton FRS-7). These drivers where connected to the recording probe by means of plastic tubes. The correct sound pressure levels were attained by calibrating the drivers *in situ* while taking the probe transfer characteristics into account. The recording probe contained a 1/2″ pressure-microphone (GRAS 40AG) that was used to record the ear canal sound pressure synchronously with stimulus delivery. The recorded signal was band-pass filtered (0.02–100 kHz, NEXUS 2690), acquired via an A/D channel (TDT RP2.1), and stored on computer disk for offline analysis. Signal generation and data acquisition were both done at a rate of 48.8 kHz.

Data Analysis

All stimulus frequencies were chosen such that an integer number of sample points held an even number of periods for these tones (see Table 1). As a consequence, these periodic segments (typically 97,656 samples, i.e., ~2 s) also held an even number of periods for all third-order DPOAEs. This allowed for averaging of a single, continuous recording by breaking it down into its periodic segments, and eliminated any spectral leakage when extracting frequency components using Fourier analysis (Papoulis 1962).

After excluding the first and last periodic segment of each recording (to exclude the ramps and transient phenomena), the remaining signal was averaged over the periodic segments and the magnitude and phase spectra were calculated via Fourier analysis. No form of artifact rejection was used. The stimulus design resulted in the even components of these spectra to be commensurate with the stimulus (and any DPOAEs evoked by it). The odd components of the spectra, which are not commensurate with the stimulus, were used to estimate the noise floor. Additional recordings were made in an artificial ear to check for distortion in the hardware. For the stimuli used in this study, these system distortions were well below the DPOAEs, often not exceeding the noise floor.

The phase of each DPOAE component was expressed in cycles *re*. the phase of the stimulus components producing it as determined from the microphone signal. Thus, the phase $\Phi_{dp}(f, g_i, g_j)$ of a DP component having frequency $f_{dp} = g_i \pm g_j \pm f$ becomes

$$\Phi_{dp}(f, g_i, g_j) = \varphi_{dp} - (\varphi_i \pm \varphi_j \pm \varphi_f) \qquad (1)$$

where $\varphi_{...}$ denotes the phase values extracted from the Fourier spectrum of the recording.

The relation between DPOAE phase and stimulus frequency was analyzed by fitting the phase data using two linear models. Any DPOAE components not exceeding the noise floor were excluded from the fitting procedure. Weight factors were assigned to the remaining DPOAE components to prevent the phases of weak components (close to the noise floor) from dominating the data fits. The weight factor W was

$$W = 1/\sigma^2 \qquad (2)$$

where σ^2 is the variance in amplitude of a tonal component caused by the addition of noise. This variance was evaluated using numerical simulations of the effect of adding random components (with phase drawn from a uniform distribution, and amplitude drawn from a Rayleigh distribution) on the phase of single spectral components. The simulations yielded a polynomial description of the variance σ^2

$$\sigma^2 = 23.6*\text{NSR}^6 - 57.1*\text{NSR}^5 - 52.4*\text{NSR}^4 \qquad (3)$$
$$-20.8*\text{NSR}^3 + 4.06*\text{NSR}^2 + 5.88*\text{NSR}$$

where NSR denotes the reciprocal of the signal-to-noise amplitude ratio.

The first model was a straight line of DPOAE phase Φ_{dp} against DPOAE frequency f_{dp}

$$\Phi_{dp} = \Phi_0 - \tau f_{dp} \qquad (4)$$

The phase offset Φ_0 and group delay τ were estimated by minimizing the weighted sum of squared differences between data and model,

$$\sum_k W_k \left(\Phi_k^{\text{model}} - \Phi_k^{\text{data}} \right)^2 \qquad (5)$$

The second model, which was used in the case of two-dimensional representations of the DPOAE phase data (see Fig. 5), is given by

$$\Phi_{dp}(f, g_i, g_j) = \Phi_0 + \tau_+(g_i + g_j) + \tau_-(g_i - g_j) \qquad (6)$$

Notice that Eq. 6 does not explicitly include a term for the single stimulus frequency f, because f was not varied within data sets that were fitted. The phase offset Φ_0 and group delays τ_+ and τ_- were estimated by minimizing the weighted sum of squared residuals (Eq. 5). Group delays will be reported by their mean ±95% confidence interval. All stimulus generation, data acquisition, and offline analysis were done via custom software in MATLAB.

RESULTS

Figure 1 shows amplitude spectra of the ear canal sound pressure when either the lower primary (A) or the upper primary (B) consisted of the tone complex. Nonlinear interaction of the single stimulus tone with one or two components from the tone complex produced three readily identified groups of third-order DPOAEs, with each DP frequency resulting from a unique set of three stimulus frequencies. The DPOAEs are of biological origin: they exceed the estimated noise floor as well as the system distortion.

The three groups of third-order DPOAEs evoked by a stimulus that consists of a single primary at frequency f and a tone complex having M components at frequencies g_1, g_2...g_M, are:

1. The "far" group: M DPOAEs having frequencies $2f-g_i$, where $i=1...M$; (*left-pointing triangles* in Fig. 1).
2. The "near" group: $^1/_2 M(M+1)$ DPOAEs having frequencies g_i+g_j-f, where $i,j=1...M$; $i \le j$ (*circles* for $i = j$, and *right-pointing triangles* when $i < j$ in Fig. 1).
3. The "sideband" group: $M(M-1)$ DPOAEs having frequencies $f + g_i-g_j$, with $i,j=1...M$; $i \ne j$ (*diamonds* in Fig. 1).
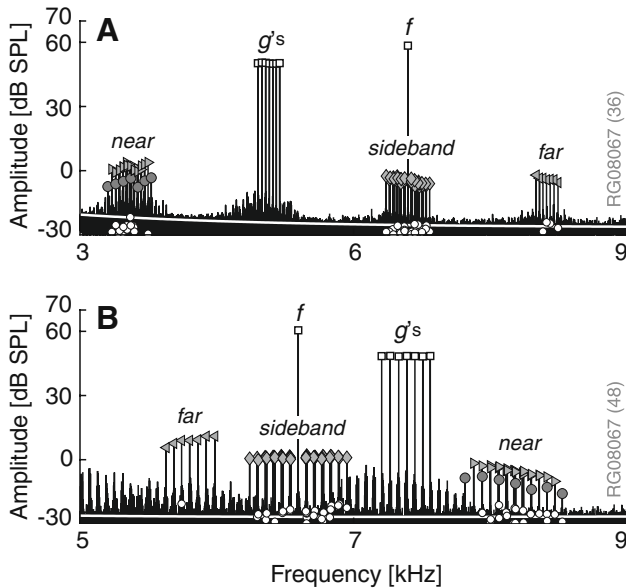

FIG. 1. DPOAEs evoked by multitone stimuli. Amplitude spectra of ear canal sound pressure are shown. Either the lower (**A**) or the upper (**B**) primary of the two-tone stimulus was replaced by a tone complex. The stimulus (*squares*) consisted of a seven-tone complex (g_1 to g_7) and a single tone (*f*). Three groups of third order DPOAEs that result from the interaction between f and g can be identified: the far group at $2f-g_i$ (*left-pointing triangles*), the near group at frequencies $g_i + g_j-f$ (*right-pointing triangles* for $i < j$ and *closed circles* for $i = j$), and the sideband group at $f+g_i-g_j$ (*diamonds*). The estimated noise floor (*line*) and the recorded system distortion (*open circles*) are also given. Stimulus parameters: $f=$ 6.6 kHz; $L_f=60$ dB SPL; $L_g=50$ dB SPL per component; seven components in tone complex; $g=5...5.2$ kHz in **A** and $g=7.2...7.6$ kHz in **B**.

Other third-order DPOAEs also arise from such a stimulus (e.g., $3f$, $g_i + g_j-g_k$), but we restrict our analysis to those components that involve both the single stimulus component f and one or two components from the tone complex g. In the next sections, each of the three groups will be considered.

The Far Group

In the frequency spectrum, these DPOAEs (*left-pointing triangles* in Fig. 1) occur at the "far side" of the tone complex (at the opposite side of the single tone as the tone complex). Using a tone complex having M components results in a far group that consists of M DPOAEs, at frequencies $f_{far}(i)=2f-g_i$, $i=1...M$. Each DPOAE thus results from the interaction of only two stimulus components, the single-tone primary and one frequency component from the tone complex. This is analogous to DPOAEs evoked by two-tone stimuli, but instead of sweeping one primary across frequencies to evoke subsequent DPOAEs, all the different stimulus frequencies are here presented simultaneously as a tone complex. More specifically, the far group resembles $2f_1-f_2$ DPOAEs when f_2 is replaced by the tone complex ($f < g$; Fig. 1B), or $2f_2-f_1$ DPOAEs when f_1 is replaced by the tone complex ($f > g$; Fig. 1A).

We tested the similarity between the far group and DPOAEs evoked by two-tone stimuli by also recording the latter emissions using either a fixed-f_2 or a fixed-f_1 paradigm (i.e., one stimulus tone was fixed in frequency, while the other was varied across recordings). DPOAEs from the two-tone recordings were then compared to the far group obtained from a single recording in which all of the stimulus components were presented at once. As an example, Figure 2A shows the superimposed amplitude spectra of 10 recordings. In nine of these recordings, two-tone stimuli (light gray) were used, in which the upper stimulus component f_2 was fixed, while the frequency of the lower component f_1 was varied across recordings (fixed-f_2 paradigm). In the 10th recording, all stimulus frequency components were presented simultaneously (dark gray). The intensity of each component in the tone complex was reduced relative to the f_1 level in the two-tone paradigm in such a way that the total power of the tone complex equaled the power of each single f_1 component in the fixed-f_2 paradigm. The amplitude and phase data for the DPOAEs from these recordings are shown in more detail in Figure 2B, C, respectively (*triangles* for far group, *circles* for DPOAEs from consecutive two-tone stimuli). Notice that, although the amplitude of each DPOAE component in the far group is smaller than that resulting from the two-tone stimulus, their total power (horizontal dashed line in Fig. 2B) is nearly identical to the power of each of the latter DPOAEs. The phase data exhibit very similar group delays, but show a small vertical offset
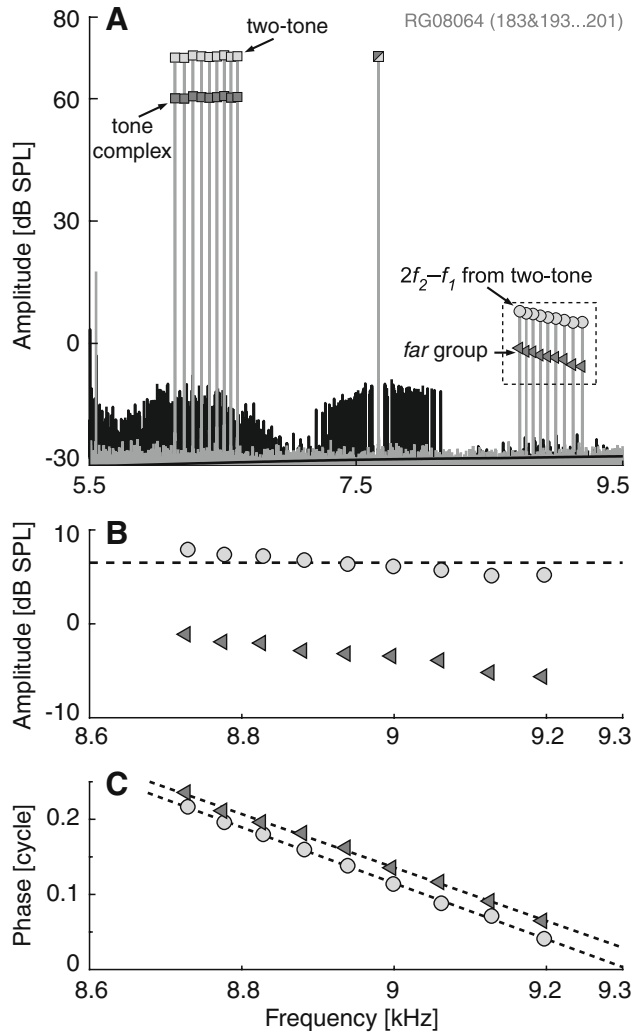
FIG. 2. Comparison between DPOAEs evoked with either a single multitone stimulus or a series of two-tone stimuli. **A** Ten superimposed amplitude spectra obtained either with nine separate two-tone stimuli (*light gray*) or with a nine-tone zwuis-complex (*dark gray*). The *dashed box* around the DPOAEs is shown again in (**B**), with the corresponding phase data in (**C**). *Circles* represent the $2f_2-f_1$ DPOAEs from the consecutive two-tone stimuli; *triangles*, the far group. The *horizontal dashed line* in **B** indicates the total power of the DPOAEs in the far group. Phase data were fitted with straight lines (*dashed lines*), yielding group delays of 372±12 μs and 356± 17 μs for the $2f_2-f_1$ DPOAEs (*circles*) and the far group (*triangles*), respectively. Stimulus parameters: f (or f_2)=7.7 kHz; g (or f_1)=6.1... 6.6 kHz (nine components); L_f (or L_2)=70 dB SPL; L_g=60 dB SPL per component; L_1=70 dB SPL.

(~0.02 cycles) between the two types of recordings. Such an offset was, however, not observed systematically across recordings. Comparisons between the far group and two-tone DPOAEs from similar recordings in four animals are made in Figure 3 (*triangles*) as scatter plots for the group delay (A) and DPOAE power (B). Depending on the relative primary frequencies, the far group resembled either the $2f_1-f_2$ (when $f < g$) or the $2f_2-f_1$ (when $f > g$) DPOAEs. There appears to be no need to distinguish between these two situations.

The general observation is that the far group is practically identical to DPOAEs evoked by consecutive two-tone stimuli. Thus, for this group of DPOAEs there is little difference between sweeping one primary across a set of frequencies (fixed-f_1 or fixed-f_2 paradigm) and presenting the whole set together as a tone complex.

The Near Group

Spectrally, this group of emissions occurs at the side of the tone complex opposite of the single stimulus tone. In analogy with the far group we will refer to it as the "near group." The frequencies of this group are given by $f_{near}(i,j) = g_i + g_j - f$. Figure 4 gives an example for which $f > g$ (see Fig. 1B for an example with $f < g$).

This DPOAE group contains more frequency components than there are stimulus components. In the example of Figure 4, a tone complex that consisted of nine components evoked a total of 36 unique DPOAE components in the near group. Also, the near group appears to be composed of two subsets that systematically differ in their amplitudes by ~6 dB (compare close circle and *triangles* in Fig. 4A).

Both the number of DPOAE components and the apparent existence of two subsets can be explained by realizing that the generation of each DPOAE component involves three stimulus components: the single stimulus component f, and two components (g_i, g_j) from the tone complex. The single frequency component f is constant within the near group, so all possible combinations are represented by the square matrix (g_i, g_j). Graphically, these combinations (g_i, g_j) are given by the lattice points in Figure 5A. The subset of M combinations for which $i = j$ produce DPOAE components (*circles*) involving only one frequency component from the tone complex. This subgroup is the "mirror image" of the far group of DPOAEs. The remaining $\frac{1}{2}M(M-1)$ combinations for which $i \neq j$ (*triangles* in Fig. 5A) produce the additional frequency components of the near group of DPOAEs. Their generation involves two different components from the tone complex. Notice that not all possible combinations result in a unique DPOAE component. Since $g_i + g_j = g_j + g_i$, only half of all $i \neq j$ combinations result in a unique frequency. In combinatorial terms, both $g_i + g_j$ and $g_j + g_i$ contribute to the same DPOAE component. This combinatorial effect corresponds to the coefficient 2 of the xy term in the binomial expansion of $(x+y)^2$, and causes the DPOAEs for which $i \neq j$ to have their amplitudes doubled (+6 dB) *re.* the amplitude of DPOAEs for which $i = j$. Apart from this combinatorial factor 2 in amplitude, there is no fundamental difference between the two groups, as is shown by the phase data (Fig. 4B) in which all near compo-
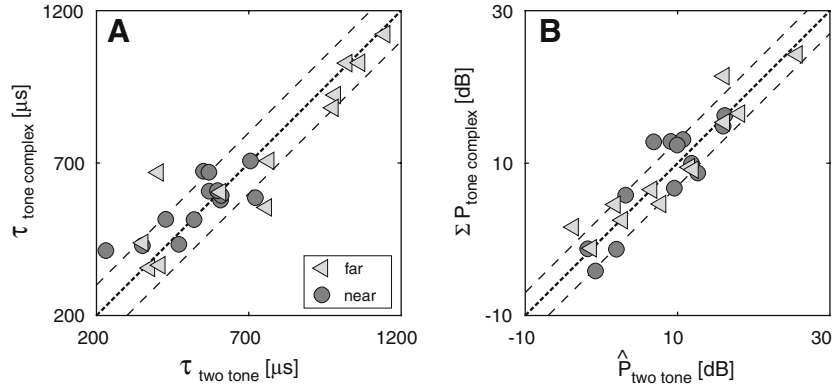
FIG. 3. Comparison of DPOAEs obtained with two-tone sweeps and multitone stimuli. DPOAEs from four animals are displayed in scatter plots of group delay (**A**) and power (**B**). Data were obtained either from a single recording using a tone complex (*ordinate*) or from several consecutive two-tone recordings that visited the frequencies of the tone complex one-by-one (*abscissa*). In **B**, the

total power ΣP of the DPOAE group is plotted against the average power of the two-tone evoked DPOAEs. See Figure 2 for an individual example. Data for the far group (*triangles*) and near group (*circles*) are shown. *Dotted* and *dashed lines* in **A** indicate equality ± 100 µs; in **B**, equality ± 3 dB.

nents are well described by a single function of DPOAE frequency.

To investigate the analogy between the near group and DPOAEs evoked by two-tone stimuli, recordings identical to those described for Figure 2 were made: in a series of two-tone recordings, DPOAEs were obtained using either a fixed-f_2 or a fixed-f_1 paradigm. These two-tone data were compared with a single recording in which all stimulus components were presented at once. Results are given in Figure 3 (*circles*) together with the data for the far group. As before, the grouping of all data around the diagonals suggests that the DPOAEs evoked by the two stimulus protocols are essentially the same.

The similarity between the DPOAEs obtained using a tone complex or a series of two-tone recordings again shows that there are no essential differences between DPOAEs recorded sequentially and simultaneously. They are generated by the same nonlinear

mechanism, and there are no indications that they propagate along different cochlear paths.

Given the close similarity between the near and far groups and consecutive two-tone stimuli, the multi-tone data may be considered as $2f_1 - f_2$ or $2f_2 - f_1$ DPOAEs as if recorded using a fixed-f_1 or a fixed-f_2 paradigm. Specifically, the near group resembles either $2f_1 - f_2$'s with f_2 fixed ($f > g$) or $2f_2 - f_1$'s with f_1 fixed ($f < g$), while the far group corresponds to either $2f_2 - f_1$'s with f_2 fixed ($f > g$) or $2f_1 - f_2$'s with f_1 fixed ($f < g$). Using this interpretation of the multitone data, DPOAE group delays were calculated for these four different {DPOAE, stimulus} configurations (Fig. 6A, B). For the "$2f_1 - f_2$" DPOAEs, the fixed-f_1 paradigm yielded systematically larger group delays than the fixed-f_2 paradigm. A comparison of the former group delays with $2f_1 - f_2$ group delays from two-tone data of the gerbil (adapted from Faulstich and Kössl 2000), shows that the absolute group delays as well
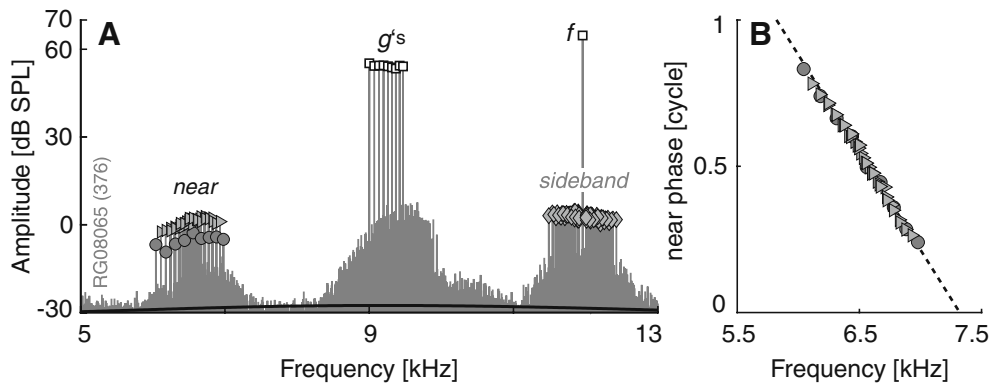

FIG. 4. Near group of DPOAEs. **A** Power spectrum of the ear canal sound pressure. DPOAE components of the near group are given by $g_i + g_j - f$ (*triangles* for $i < j$ and *circles* for $i = j$). Notice the systematic, ~6 dB, difference in magnitude between the two subgroups. **B**

Corresponding phase data. A straight line fitted to these data (*dashed line*) yielded a group delay of 666 ± 16 µs. Stimulus parameters: $f = 12$ kHz; $g = 9...9.5$ kHz (nine components); $L_f = 65$ dB SPL; $L_g = 55$ dB SPL per component.
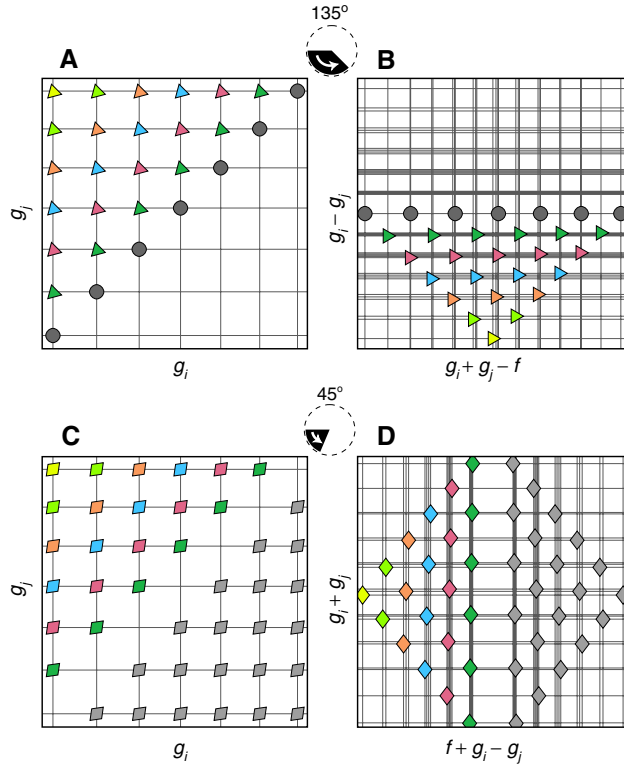
FIG. 5. (color online) Schematic diagrams illustrating how pairs of components from the tone complex combine to produce DPOAEs in the near group (*upper row*) and sideband group (*bottom row*). Each DPOAE results from a unique pair $[g_i, g_j]$ of tone complex components, represented by the two-dimensional, square matrix. For the near group (**A**), a transformation of this matrix by a 135° counterclockwise rotation (see *inset*), and by a translation over the distance $-f$, yields the equivalent matrix $[g_i+g_j-f,\ g_i-g_j]$ (**B**). In this representation the abscissa equals DPOAE frequency, and the ordinate represents stimulus frequency variations that leave DPOAE frequency fixed. For the sideband group (**C**), the translation over distance f combined with a 45° counterclockwise rotation yields matrix $[f+g_i-g_j,\ g_i+g_j]$ (**D**). This transformation has the same effect: the abscissa again equals DPOAE frequency, while the ordinate represents the variations that leave DPOAE frequency fixed. Note that the zwuis character of the stimulus ensures that none of the symbols are aligned vertically in panels **B** and **D**. The (*color*) grading of the symbols has no particular meaning, but is an aid to track the rotations.

as the trend for these delays to decrease with increasing frequency are very similar, confirming the similarity between the two-tone and multitone DPOAEs. In contrast, group delays for the "$2f_2-f_1$" DPOAEs were similar for the fixed-f_1 and fixed-f_2 paradigm, and are comparable to the {"$2f_1-f_2$"/fixed-f_2} group delays.

For the near group of DPOAEs, the zwuis character of the stimulus becomes indispensable (see Materials and Methods). It ensures that all possible combinations $g_i + g_j$ are unique, so that each DPOAE frequency $g_i + g_j-f$ can be unambiguously attributed to a unique pair (g_i, g_j) of stimulus components from the tone complex. This is illustrated in Figure 5B, in which the (g_i, g_j)-matrix of Figure 5A is transformed (rotated and translated) into the equivalent matrix

$([g_i + g_j-f],[g_i-g_j])$, where $g_i + g_j-f$ equals the DPOAE frequency. The zwuis character of the stimulus ensures that none of the symbols are vertically aligned: projecting them on the abscissa (i.e., DPOAE frequency) will never result in two coinciding points. This is necessary to disentangle the different contributions of the stimulus components to the individual DPOAEs.

The wealth of components in the near group offers an extra "degree of freedom" in the analysis of phase data. The presence of $i \neq j$ components enables an analysis of DPOAE phase that extends beyond the dependence on DPOAE frequency analyzed thus far. The additional degree of freedom is illustrated by representing the DPOAE components of the near group in the two-dimensional format of Figure 5B. In this representation, the dependence of DPOAE phase on DPOAE frequency corresponds to variation along horizontal lines (i.e., parallel to the abscissa). In contrast, variation along vertical lines indicates a
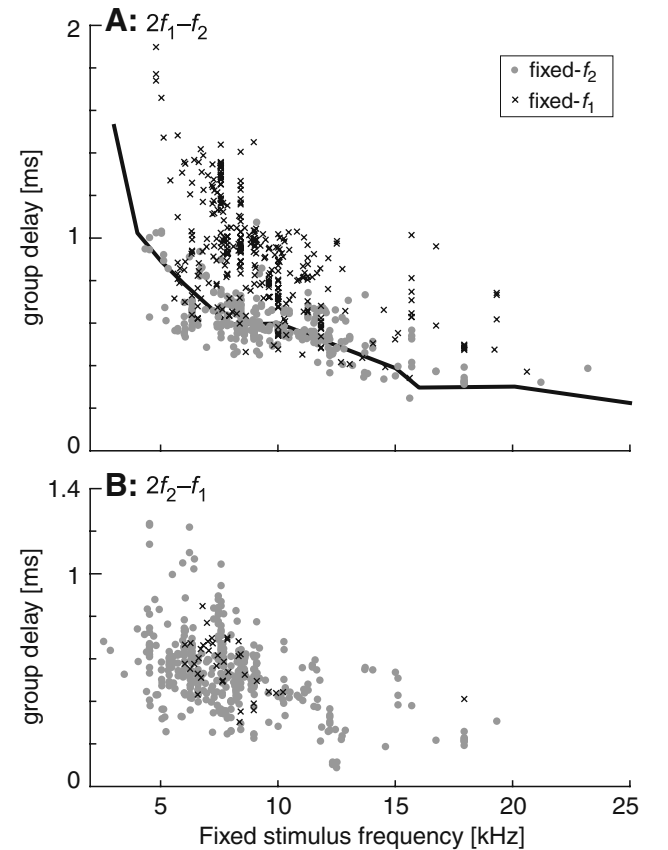

FIG. 6. Group delays for the near and far groups of DPOAEs. Based on their similarity with DPOAEs from two-tone sweeps (see text), the data were split into effective $2f_1-f_2$ group delays (**A**) and effective $2f_2-f_1$ group delays (**B**). Within each of these groups, and as indicated in the graph, the markers distinguish effective fixed-f_2 and effective fixed-f_1 group delays, corresponding to ($f>g$) and ($f<g$), respectively. For comparison, genuine two-tone data from the literature are shown in **A** (*line*). These $2f_1-f_2$ group delays (Faulstich and Kössl 2000), were obtained from gerbils using a fixed-f_2 paradigm.

dependence on stimulus frequencies in a way that does not correspond to changes in emission frequency.

Figure 7A shows the phase data from Figure 4B in the two-dimensional representation in the form of a contour plot. It is important to keep in mind that the phase data shown here were obtained from a *single* recording. In discussing them, it is hard to avoid phrases like "variations in stimulus frequencies", which might suggest the use of frequency sweeps. In reality, all stimulus components were presented together, and the "variations" refer to comparisons across DPOAE components produced by different triplets drawn from the $M+1$ stimulus components.

The contours in Figure 7A are iso-phase lines; the maximum variation of DPOAE phase occurs along the direction perpendicular to them. The vertical orientation of the contours thus indicates that DPOAE phase changes with variation along the abscissa. Thus, DPOAE phase only depends on DPOAE frequency. In contrast, variations in stimulus frequencies (g_i, g_j) that keep the DPOAE frequency fixed (variations along the ordinate) have no systematic effect on DPOAE phase. This observation is quantified by fitting the two-dimensional phase data with the model given in Eq. 6 (see Materials and Methods), which yielded $\tau_+ = 665 \pm 16$ μs, and $\tau_- = 16 \pm 26$ μs (Fig. 7B). Notice that τ_+ is the coefficient describing the phase variations with DPOAE frequency, i.e., the slope of the frequency-phase plot of Figure 4B. As explained above, τ_+ matches the group delay obtained with a two-tone stimulation paradigm. In contrast, τ_- is the coefficient for those variations in the frequencies of the tone complex g that keep the ($g_i + g_j$), and thus DPOAE frequency, fixed.

Figure 7C is a scatter plot of similarly calculated group delays for 998 recordings (830 when $f > g$; 168 when $f < g$) from 15 animals. These data show that DPOAE phase for the near group is well described by their variation with DPOAE frequency. This holds true for both stimulus configurations $f > g$ and $f < g$. Variations in stimulus

frequency that do not correspond with changes in DPOAE frequency (i.e., varying $g_i - g_j$ while fixing $g_i + g_j$) have no explanatory contribution to the observed DPOAE phase. This observation is consistent with the good fit of the straight line to the phase-versus-frequency data (Fig. 4B), which leaves little residual variance to be accounted for.

In summary, the near group results from the combination of either two or three stimulus components. The "combinatorial effects" of these components explain both the number of DPOAE components and the emergence of two subgroups whose magnitudes are 6 dB apart. There is little difference between sweeping one primary across a set of frequencies and presenting the whole set together as a tone complex. In this respect the near group is similar to the far group. The "off-diagonal" ($i \neq j$) components permit an analysis of the DPOAE phase that goes beyond their dependence on DPOAE frequency. This analysis shows that, within the range of frequencies tested, the phase of DPOAE components only varies with those variations in stimulus frequencies that correspond to changes in the DPOAE frequency.

The "Sideband Group"

The third and final group of DPOAEs does not have a two-tone equivalent; it only arises when the stimulus consists of more than two frequency components. It consists of a set of sidebands around the single stimulus component f (e.g., *diamonds* in Fig. 1A, B) at frequencies $f + g_i - g_j$, with $i \neq j$.

As before, the origin of the DPOAE components (in terms of the stimulus components from the tone complex) is illustrated by considering the matrix (g_i, g_j) (Fig. 5C). Each $i \neq j$ combination results in a DPOAE component, and the zwuis character of the stimulus ensures that these components are all unique (see Fig. 5D). Because ($g_i - g_j$) $\neq$ ($g_j - g_i$), the combinatorial
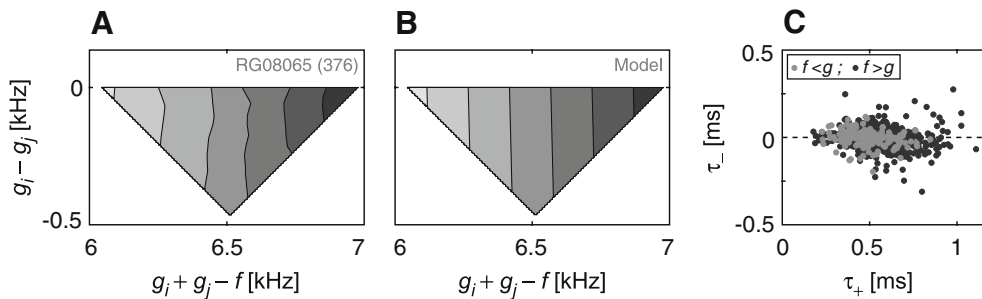

FIG. 7. **A** Two-dimensional representation of DPOAE phase for the near group. The phases of Figure 4B are replotted as a contour plot, using the two-dimensional representation of stimulus frequencies of Figure 5B. The abscissa gives DPOAE frequency f_{dp}; the ordinate, fixed-f_{dp} variations. **B** Results from the two-dimensional model (Eq. 6) fitted to these phase data ($\tau_+ = 665 \pm 16$ μs; $\tau_- = 16 \pm 26$ μs). Contours in **A** and **B** are drawn at 0.1-cycle intervals. Stimulus parameters: $f = 12$ kHz; $g = 9...9.5$ kHz (nine components); $L_f = 65$ dB SPL; $L_g = 55$ dB SPL per component. **C** Scatter plot of τ_+ and τ_- calculated for 998 recordings from 15 animals (830 when $f > g$; 168 when $f < g$). DPOAE phase for the near group is well described by DPOAE frequency, irrespective of the stimulus configuration.

effect described for the near group does not apply here: no subgroups of DPOAEs arise. DPOAE components for which $i = j$ coincide with the stimulus component f, and cannot be resolved in the spectra of the recorded signals. Thus an M-component tone complex results in $M^2 - M = M(M-1)$ different DPOAE components within the sideband group.

Since the DPOAE frequencies of the sideband group depend on the frequency *difference* $g_i - g_j$, a collective frequency shift of the entire tone complex g does not affect the frequencies of the DPOAEs. Recently, we used this property to study reverse intracochlear propagation of otoacoustic emissions in the gerbil ear (Meenderink and van der Heijden 2010). In the phase analysis of the sideband group, we distinguish the two cases $f > g$ (single tone above tone complex) and $f < g$ (single tone below tone complex).

DPOAE Phase; $f > g$

As an example, Figure 8A shows phase data for the sideband group of DPOAEs as a function of DPOAE frequency. These data correspond to the amplitude spectrum in Figure 4A, and were obtained using a stimulus for which $f > g$. A straight line (Eq. 4) fitted to
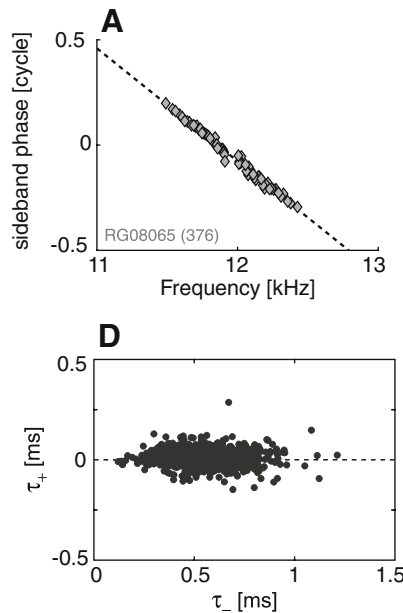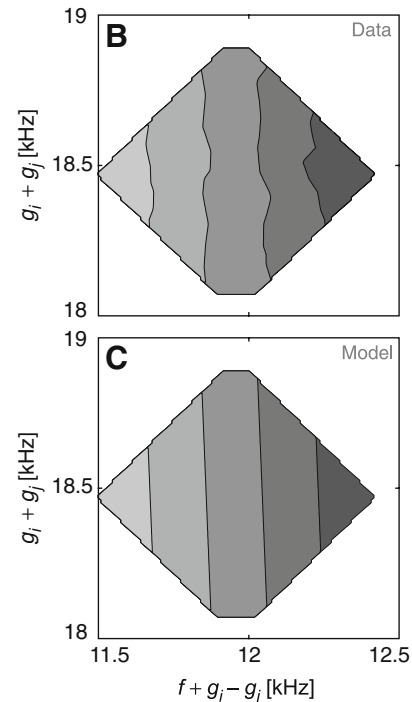
these data yields a group delay of 536 ± 13 μs. As was the case for the near group, the phase data can be rearranged in a two-dimensional format (Fig. 8B) with DPOAE frequency along the abscissa and non-DPOAE frequency variations represented along the ordinate. Similar to the near group, the contour lines are vertical, indicating that the variation in DPOAE phase only occurs with those changes in stimulus frequency that cause the DPOAE frequency to change. Fitting Eq. 6 to these data (Fig. 8C) yielded $\tau_+ = 22 \pm 14$ μs and $\tau_- = 536 \pm 13$ μs. Note that compared to the near group, the roles of τ_+ and τ_- are reversed: this time it is the τ_- coefficient that represents the group delay along the DPOAE frequency axis. A scatter plot (Fig. 8D) of group delays for phase data obtained from all 15 animals ($N = 1,065$) generalizes this observation. DPOAE phase varies systematically with DPOAE frequency (captured in τ_-), whereas non-DPOAE-related frequency variations (quantified by τ_+) do not improve the explanatory power of the linear model Eq. 6 for stimulus conditions with $f > g$.

DPOAE Phase; $f < g$

From the DPOAE amplitude spectra, no differences are apparent between the sideband DPOAEs obtained

FIG. 8. Phase data for the sideband group obtained with $f > g$. **A** DPOAE phase plotted against DPOAE frequency. A straight line was fitted to these data (*dashed line*), yielding a group delay of 536 ± 13 μs. In **B**, the same data are plotted as a contour plot, using the two-dimensional representation of Figure 5D. The abscissa gives DPOAE frequency f_{dp}; the ordinate, fixed-f_{dp} frequency variations. **C** Results from the two-dimensional model (Eq. 6) fitted to these phase data ($\tau_+ = 22 \pm 14$ μs; $\tau_- = 536 \pm 13$ μs). Contours in **B** and **C** are drawn

at 0.1-cycle intervals. The corresponding magnitude spectrum is shown in Figure 4A. Stimulus parameters: $f = 12$ kHz; $g = 9...9.5$ kHz (nine components); $L_f = 65$ dB SPL; $L_g = 55$ dB SPL per component. **D** Scatter plot of τ_- and τ_+ calculated for 1,065 recordings from 15 animals, all with $f > g$. DPOAE phase for the $f > g$ sideband group is well described by DPOAE frequency; fixed-f_{dp} frequency variations have no systematic effect on DPOAE phase.

with $f > g$ (Fig. 1A, *diamonds*) versus those evoked with $f < g$ (Fig. 1B, close *diamonds*). The phase data, however, do reveal a contrast between the two stimulus configurations. As an example, Figure 9A, B show amplitude and phase spectra obtained with $f < g$. As before, the phase data change with DPOAE frequency, and this trend is largely captured by a straight line fit to these data. On closer inspection, however, the phase also show a systematic patterning along near-vertical lines; a trend that cannot be explained by their dependence on DPOAE frequency alone. Rearranging the phase data in the two-dimensional format as described earlier (see Fig. 5D) clarifies the nature of the dependence (Fig. 9C). This time, the contours are not vertical, but tilted clockwise. The tilting indicates that, although a major portion of the phase variation is explained by the DPOAE frequency, phase is also affected by the stimulus frequencies in a different way. This is quantified by fitting Eq. 6 to these data, which yielded $\tau_+=-113\pm20$ μs and $\tau_-=642\pm19$ μs. The resulting plot (Fig. 9D) clearly shows non-vertical contours that reflect the nonzero group delay τ_+. Analysis of a total of 298 recordings obtained in 13 gerbils generalizes this observation. Unlike the corresponding analysis

for the $f > g$ case (Fig. 8D), the scatter plot of the group delays τ_+ and τ_- (Fig. 9E) is no longer distributed around the line $\tau_+=0$. Instead, it shows a systematic trend towards negative group delays for the non-DPOAE-related frequency components ($\tau_+<0$). In this respect, these DPOAEs (i.e., sideband group obtained with $f < g$) are different from all other groups, including the sideband DPOAEs obtained with $f > g$.

Theoretical Analysis of the Stimulus Dependence of DPOAE Phase: Residual Phase Effects

The multitone stimuli provided an extra degree of freedom in the analysis of DPOAE phase beyond the straightforward, one-dimensional graphs of DPOAE phase versus DPOAE frequency. In the two-dimensional phase plots (Figs. 7–9), the phase gradients in the *horizontal* direction correspond to the straightforward dependence on DPOAE frequency, which is the dominant effect in all of the phase data of the near and sideband groups. The two-dimensional phase data, however, also allowed the assessment of phase gradients in the *vertical* direction, which may be interpreted as data obtained with a "fixed-f_{dp}" recording paradigm. These gradients allowed us to assess
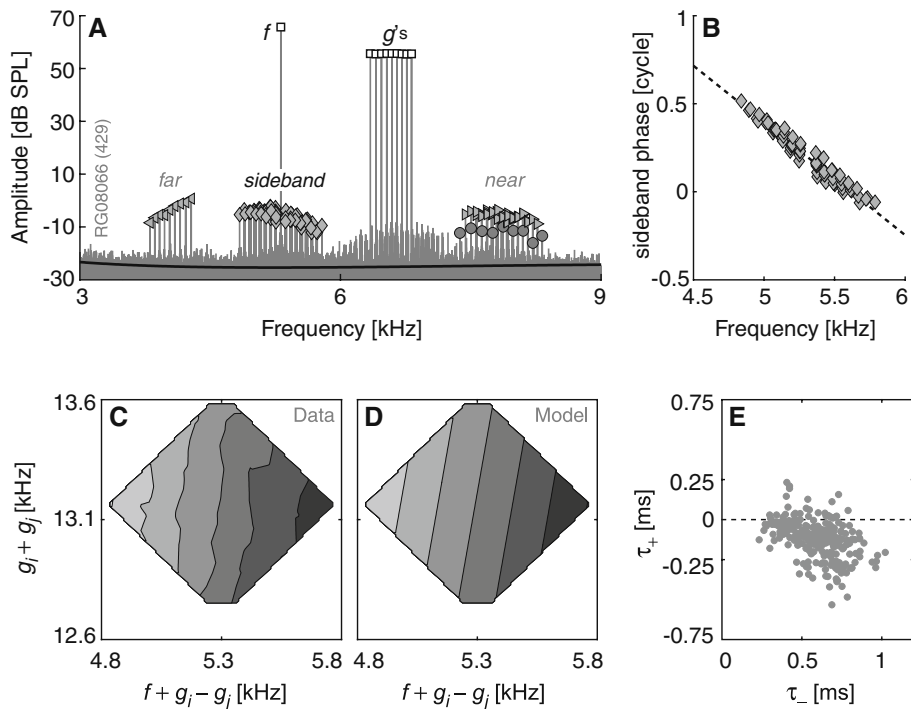

FIG. 9. Sideband DPOAEs obtained with $f<g$. **A** Amplitude spectrum of ear-canal sound pressure. **B** Corresponding phase data for the sideband group of DPOAEs. Fitting a straight line to these data yielded a group delay of 641 ± 32 μs. The line fit (*dashed line*) captures the major trend in these data, but misses the systematic near-vertical patterning in the data. **C** Two-dimensional representation of the same phase data and **D** the result from fitting the linear model of Eq. 6 to these phase data ($\tau_+=-113\pm20$ μs; $\tau_-=642\pm19$ μs).

Contours in **C** and **D** are drawn at 0.1-cycle intervals. Stimulus parameters: $f=5.3$ kHz; $g=6.3...6.8$ kHz (nine components); $L_f=65$ dB SPL; $L_g=55$ dB SPL per component. **E** Scatter plot of τ_- and τ_+ calculated for 298 recordings from 13 animals, all with $f < g$. DPOAE phase for the sideband group is not exhaustively described by DPOAE frequency f_{dp}; the data show a systematic effect of fixed-f_{dp} frequency variations.

whether DPOAE phase changes systematically when the stimulus frequencies (g_i, g_j) are "varied" in a way that keeps the DPOAE frequency fixed. (They are not really varied, because the various frequencies $g_1...g_M$ are all presented simultaneously. For ease of explanation, however, we will discuss the data as if collected while varying (g_i, g_j) according to a fixed-f_{dp} sweep.)

Because the dependence on DPOAE frequency (i.e., the phase gradient along the horizontal direction) is the main effect in our data (often it is also the only dependence observed), we will term the second type of DPOAE phase dependency the *residual phase dependency*. It is an interesting type of dependence, because travel times common to all components (stimulus and DPOAEs alike) do not contribute to it; they are absorbed into the main effect. As explained below, this makes the residual phase dependence a sensitive probe of the details of DPOAE generation.

Consider a generic model of DPOAE generation, in which the phase of a stimulus component f at a certain location x the cochlea is described by a function $\Phi(f,x)$, and the generation site of a third-order DP component f_{dp} is $X_G(f_1,f_2,f_3; f_{dp})$. No specific assumptions are made on the functions Φ and X_G. In the Appendix it is shown that, to first-order approximation, the phase of the DP at its place of generation is given by

$$\Phi_{dp} = \Phi_0 + \kappa \sum_{k=1}^{3} f_k \frac{\partial X_G}{\partial f_k} + \kappa f_{dp} \frac{\partial X_G}{\partial f_{dp}} - \tau f_{dp} \quad (7)$$

where κ and τ, and Φ_0 are constants. Consider now a particular joint variation of the stimulus frequencies

$$(f_1, f_2, f_3) \rightarrow (f_1 + \Delta f_1, f_2 + \Delta f_2, f_3 + \Delta f_3)$$

that leaves f_{dp} fixed, i.e. $\Delta f_{dp}=0$. For such fixed-f_{dp} variations, the only term contributing to variations in Φ_{dp} is the second term of Eq. 7, resulting in

$$\Delta \Phi_{dp} = \kappa \sum_{k=1}^{3} \Delta f_k \frac{\partial X_G}{\partial f_k} = \kappa \Delta X_G \quad (8)$$

Therefore, what we have called the residual phase dependence, directly corresponds to shifts ΔX_G of the cochlear location X_G at which the DP is generated. If the generation site is unchanged ($\Delta X_G=0$), fixed-f_{dp} "sweeps" will leave DPOAE phase unchanged. Conversely, when a residual phase effect *is* observed, it signals a shift of the generation site of the DPOAE.

With our particular stimulus design, two-dimensional phase representations were possible for two DPOAE types (near and sideband groups) in two stimulus configurations ($f > g$ and $f < g$). The analysis of DPOAE phase in the resulting four situations is summarized in Table 2. The only case showing a

TABLE 2

Population data of the residual phase effect for the four different {stimulus, DPOAE}-configurations

	$f > g$	$f < g$
Near (τ_-)	-6 ± 51 μs ($N=830$)	-3 ± 45 μs ($N=168$)
Sideband (τ_+)	8 ± 40 μs ($N=1,065$)	-115 ± 114 μs ($N=298$)

Each cell displays mean (±standard deviation) of the fixed-f_{dp} group delay (see text). The numbers between parentheses give the number of recordings. Only for the sideband group, recorded with $f < g$, does the residual phase contribute to the observed DPOAE phase variation

systematic residual phase effect was the sideband DPOAE when $f > g$ (lower right entry of the table). In order to relate this observation to the theoretical analysis of the residual phase effect, it is necessary to consider how, for each of the four situations listed in Table 2, the site of DPOAE generation is affected by the fixed-f_{dp} "sweeps". To this end, we make the following assumption: When the nonlinear interaction of three stimulus frequencies (f_1,f_2,f_3), produces a DP at frequency f_{dp}, the site of DP generation corresponds to the region with characteristic frequency $f_{max} = \max(f_1,f_2,f_3,f_{dp})$. In other words, DP generation is at the most basal best site among the four frequency components. We will refer to this as *the assumption of the most basal generation site.*

The assumption of the most basal generation site is a generalization of the evidence that, for two-tone stimuli, the $2f_1-f_2$ DP is generated near the f_2 site, whereas the $2f_2-f_1$ DP is generated near its own best site (Brown and Kemp 1984; Martin et al. 1987). The assumption is further motivated by the extremely steep high-frequency flank of mechanical tuning (Robles and Ruggero 2001), implying that tones (whether presented acoustically or generated inside the cochlea) barely propagate in cochlear regions more apical than their own best site.

The schematic diagrams in Figure 10 depict the overlapping cochlear excitation patterns for each of the four different situations of Table 2. The generation site of DPOAEs (as predicted from the assumption of most basal generation) is labeled X_G in each case. The fixed-f_{dp} "sweeps" that underlie the analysis of the residual phase effect are indicated by the arrows placed over the peaks. These arrows indicate the joint shift in primaries g_i and g_j that keep f_{dp} fixed. Notice that the requirement of a fixed f_{dp} calls for different types of joint shifts. For the near group (upper two panels), where $f_{dp} = g_i + g_j - f$, primaries g_i and g_j must move in opposite directions; for the sideband group (lower two panels), where $f_{dp} = f + g_i - g_j$, the primaries g_i and g_j must move in the same direction. A straightforward application of the above assumption of the most basal generation site now immediately leads to the result that the only situation in which a fixed-f_{dp} "sweep" causes a shift of the intra-cochlear location of DP generation X_G, is the sideband
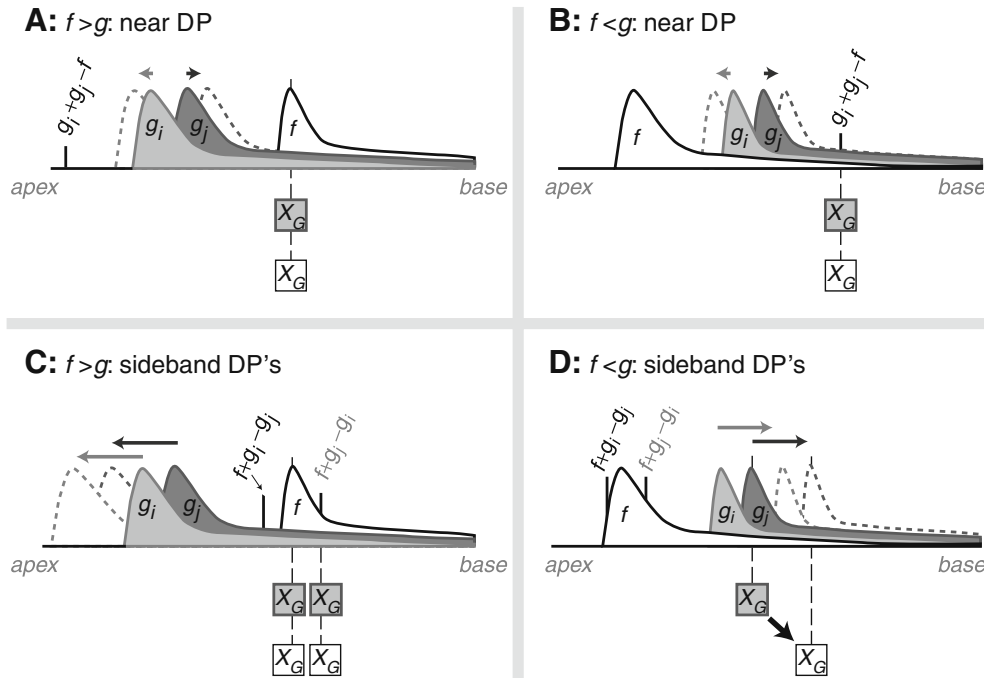
FIG. 10. Schematic diagrams of the cochlear excitation patterns for different three-tone stimuli and the surmised locations of DP generation. **A** $f > g$; near group; **B** $f < g$, near group; **C** $f > g$, sideband group; **D** $f < g$, sideband group. For a given stimulus (*solid lines*; *shaded areas*), DP generation is postulated to occur at the best site of the highest frequency component (f, g_i, g_j, or f_{dp}) involved, as indicated in the lower part of each panel by X_G (*gray*). A shift of the stimulus frequencies (g_i, g_j) that leaves the DP frequency unchanged (*dashed lines*) results in a shifted DP generation site for the sideband group when $f < g$ (**D**): in this example, it shifts to a more basal location. No such shift of the generation site X_G occurs for any of the other {stimulus, DP} configurations (**A**, **B**, **C**).

group with $f < g$ (lower right panel of Fig. 10). This is consistent with our observation that the sideband group with $f < g$ is the only one showing a residual phase effect (Table 2).

DISCUSSION

We explored the spectrum of third-order DPOAEs evoked by multitone stimuli in the gerbil ear. Multitone (>2 primaries) stimuli were used in previous work on DPOAEs, but in the majority of those studies the DPOAEs considered were still generated by two primaries; the third tone was only used to interact with the generation of the two-tone DPOAEs (Brown and Kemp 1984; Meenderink and Narins 2007). Genuine three-tone DPs have been recognized and observed in their own right, both as "additional DPs" (in widely separated primaries) and as "sidebands" around one stimulus tone (when the other two primaries were close; e.g., Kemp and Brown 1986; Kemp 1998). They have also occurred as the spectral side-effect of low-frequency biasing of two-tone DPOAEs (Bian and Scherrer 2007; Marquardt et al. 2007) or dynamic suppression of stimulus frequency OAEs (Meenderink and van der Heijden 2010). Unlike previous work, the aim of the present study

was to integrate three-tone DPOAEs into a unifying framework alongside conventional DPOAEs.

Our stimuli were constructed by replacing one of the primaries of a customary two-tone stimulus by a narrowband tone complex. This produced rich spectra of the ear canal sound pressure that contained a large number of third-order DPOAEs. The analysis of this spectrum was greatly facilitated by a particular stimulus design in which an irregular spacing of the components prevents multiple DPOAEs from coinciding (Victor et al. 1977; van der Heijden and Joris 2003; 2006).

The richness of the DPOAE spectrum renders the conventional names ($2f_1-f_2$; $2f_2-f_1$;...) insufficient. We identified and analyzed three groups of third-order DPOAEs, two of which (the "near" and "far" groups) are closely related to the well-known CDTs evoked by tone pairs; and one of which (the "sideband" group) manifested itself as sidebands around the single-tone primary. The sideband group has no two-tone evoked counterpart. We showed that a consideration of the basic properties of third-order distortions is sufficient to explain the occurrence of these three groups as well as their basic properties, such as the number of components of each group and the emergence of two sets of magnitudes in the near group.

We explored the connection between the near and far groups evoked by tone complexes on the one hand, and the customary DPOAEs evoked by tone

pairs on the other. Their frequency-phase curves, group delays, and magnitude spectra were found to be very similar (Figs. 2, 3, and 6). We conclude that the general features of these DPOAEs are the same, regardless of whether one of the tones of a pair is swept in frequency (the customary approach), or whether the different "swept" frequencies are presented simultaneously in the form of a tone complex (the approach introduced in this study). Such an equivalence of subsequent and simultaneous stimulation modes is not *a priori* obvious for a nonlinear system like the inner ear. Incidentally, simultaneous presentation does not necessarily speed up the measurements as the same response energy is now subdivided among the multiple DP components giving a lower signal to noise ratio.

Overall, the third-order DPOAEs evoked by the interaction of triplets of primary components appear to be a natural extension of those evoked by tone pairs. The use of tone complexes, however, also opens up modes of analysis that go beyond two-tone paradigms. The sideband group of DPOAEs realizes a relation between stimulus frequencies and DPOAE frequencies that is very different from the two-tone paradigm. It allows for much wider frequency spacing between primaries than is possible for the generation of two-tone DPOAEs. In a recent study (Meenderink and van der Heijden 2010), we have employed a particular realization of the sideband group, in which the tone complex was more than an octave below the single-tone primary, to analyze cochlear travel times.

More generally, the analysis of three independent interacting primaries creates "extra degrees of freedom" compared to two-tone paradigms. By representing DPOAE phase (near and sideband groups) in a two-dimensional format (Figs. 7A, 8B, and 9C), we tested whether variation of DPOAE phase was exhaustively described by its dependence on DPOAE frequency (which is given by a fixed relation between the stimulus frequencies), or whether the other degrees of freedom co-determined DPOAE phase. For the latter phenomenon, in which DPOAE phase changes while DPOAE frequency is constant, we introduced the term "residual phase effects".

The two-dimensional phase analysis (Figs. 7B, 8D, and 9E) showed that, in all except one situation, DPOAE phase exclusively varies along the DPOAE-frequency direction. The one exception that showed residual phase dependence was the sideband group emerging when the tone complex was above the single primary (Table 2). A preliminary theoretical analysis showed that, under general assumptions of DPOAE generation, this is precisely the only situation in which the intracochlear generation site of the DPOAEs is variable while DPOAE frequency is fixed (Fig. 10).

Further quantitative modeling of the residual phase effect is beyond the scope of this explorative report. It seems reasonable, however, to link our observations with the known contrast between f_1- and f_2-sweeps on the phase of $2f_1 - f_2$ emissions (e.g., Shera et al. 2000). Several studies (Moulin and Kemp 1996a, b; O Mahoney and Kemp, 1995; Schneider et al. 1999; Whitehead et al. 1996), including the present one (Fig. 6), have shown that for the $2f_1 - f_2$ CDT, a fixed-f_1 recording paradigm yields larger group delays than a fixed-f_2 paradigm. Their different group delays (evaluated from phase versus $2f_1 - f_2$ curves) imply that fixed-$(2f_1 - f_2)$ sweeps, in which f_1 and f_2 are jointly swept while keeping $2f_1 - f_2$ fixed, will still produce a phase variation. In the terminology introduced in this study, this is a case of a residual phase effect. It has been explained by the assumption that $2f_1 - f_2$
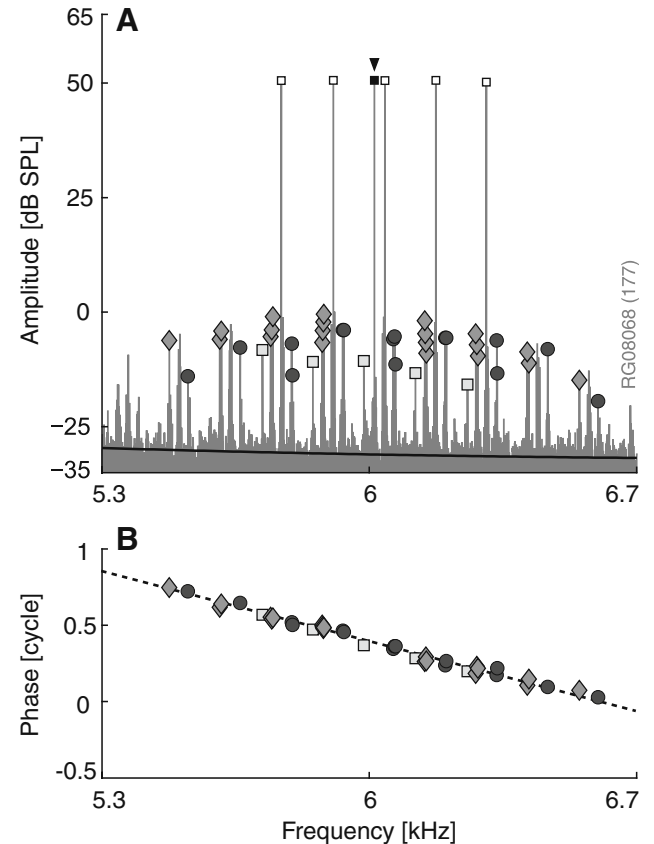

FIG. 11. Merging of DPOAE groups for $f \cong g$. **A** DPOAE spectrum evoked by a special stimulus in which the single stimulus tone f (*closed square*; *vertical arrowhead*) was placed within the frequency range of the tone complex g (*open squares*), resulting in the overlap of the three DPOAE groups (*diamonds* sideband group; *circles* near group; *open squares* far group). The stimulus design (see Table 1) ensures that DPOAE components never coincide, despite their overlapping frequency ranges. Phase data (**B**) were analyzed both separately per group and collectively, yielding group delays of 652 ± 37 μs (sideband); 667 ± 44 μs (near); 699 ± 62 μs (far); and 655 ± 29 μs (all groups pooled). Stimulus parameters: $f = 6.0$ kHz; $g = 5.8...6.3$ kHz (five components); $L_f = 50$ dB SPL; $L_g = 50$ dB SPL per component.

emissions are generated near the f_2 region (Brown and Kemp 1984; Martin et al. 1987), which moves during the fixed-$(2f_1-f_2)$ sweep. The complementary observation that $2f_2-f_1$ group delays are independent of sweep paradigm is similarly explained by assuming that it is generated near its own peak region (Martin et al. 1987). With this assumption, a joint sweep of f_1 and f_2 that leaves $2f_2-f_1$ unchanged will have no effect on DPOAE phase: the residual phase effect is absent. We have generalized these observations on the location of DPOAE generation in the "assumption of the most basal generation site" introduced in our theoretical analysis (last section of Results). This explanation parallels the description of DP generation in terms of "wave-fixed" and "place-fixed" sources (reviewed in Kemp 2008). Apparently, these concepts, which were formulated to analyze two-tone DPOAEs, are also applicable to the sideband DPOAEs reported in the present study.

Rather than introducing new names and types of emissions (see also Shera and Guinan 1999), we would like to emphasize the potential synthesizing power of multitone DPOAEs. By definition, third-order distortions emerge by the interaction of *three* primaries. It is perfectly valid for two of the three interacting tones to coincide, but there is no compelling reason to restrict the study of DPOAEs to this rather limited subclass of distortion products. In fact, such a restriction to two-tone DPOAEs is potentially biasing and may obscure the underlying unity of the family of distortion products. Any realistic nonlinear mechanism that generates a $2f_1-f_2$ component will also generate a $2f_2-f_1$ component. In this respect, lower and upper CDTs are really two sides of the same coin. Unfortunately, the typical primary ratio of ~1.2 results in a sizeable frequency separation between the lower and upper CDTs (ratio of ~1.75). This frequency difference creates a marked contrast in their propagation toward the ear canal owing to the tonotopic organization the cochlea. Therefore, this type of two-tone data emphasizes the contrast between the lower and upper CDTs, and underexposes the underlying unity. This unity is evident from the fact that extensive (f_1, f_2) area sweeps show a continuum of DPOAEs rather than two disjoint regions (Knight and Kemp 2000; Schneider et al. 2003).

The data of the present study demonstrate that it is well possible to generalize beyond two-tone stimuli, provided some caution is observed in the stimulus design. The capability of a multitone paradigm to reveal the unity of third-order DPOAEs is perhaps best illustrated by a recording in which the single tone was placed in the middle of the multitone primary (Fig. 11). The stimulus design still makes it possible to distinguish the near, far and sideband groups based on the unique triplets of primaries producing each

DPOAE component, and the phases of the groups are distinguished by the different symbols in Figure 11B. It is clear, however, that the distinction between the groups has become purely formal. Owing to their overlapping frequency range, the groups merge into a single family of third-order DPOAEs spanning a bandwidth of ~1.1 kHz over which they show a systematic phase variation. This DPOAE spectrum bridges the gap between lower and upper CDTs, perhaps illustrating that the nomenclature of DPOAEs ($2f_1-f_2$, near group...) is often more a reflection of the experimental methodology than of the mechanisms that shape them.

ACKNOWLEDGMENTS

Open Access

APPENDIX

Denote the phase of the cochlear vibration evoked by a single tone by

$$\Phi(x;f), \tag{A1}$$

where f is the frequency of the tone and x is the cochlear location. No specific assumptions are made regarding this function, except sufficient smoothness. Consider a triplet of stimulus tones at frequencies f_1, f_2, and f_3, which need not be all different. The frequency f_{dp} of a third-order DP produced by the nonlinear interaction of the three tones is

$$f_{dp} = \sum_{k=1}^{3} w_k f_k, \tag{A2}$$

where $w_k=\pm 1$ are weight factors. For the three DPOAE groups considered in this study (e.g., far, near, and sideband groups), the weight factors always add up to unity:

$$\sum_{k=1}^{3} w_k = 1. \tag{A3}$$

With the assumption that a DP component is generated at a single cochlear location X_G, its phase Φ_{dp} at this location becomes

$$\Phi_{dp} = \sum_{k=1}^{3} w_k \Phi(X_G; f_k). \qquad (A4)$$

In general, X_G will be a function of all three primary frequencies and the DP frequency, f_{dp}:

$$X_G = X_G(f_1, f_2, f_3; f_{dp}).$$

For sufficiently narrowband signals, the phase of a frequency component f close to f_0 at a place x close to x_0 is well described by the first-order approximation

$$\Phi(x, f) \cong \Phi(x_0; f_0) + \kappa(x - x_0) - \tau(f - f_0), \qquad (A5)$$

where the constants $\kappa = \partial\Phi/\partial x(x_0, f_0)$ and $\tau = -\partial\Phi/\partial f(x_0, f_0)$ are the average wave number (inverse wavelength) and group delay, respectively. Similarly, if the primary frequencies and DP frequency are sufficiently close together, one may approximate

$$X_G(f_1, f_2, f_3; f_{dp}) \cong X_G(f_0) + \sum_{k=1}^{3}(f_k - f_0)\frac{\partial X_G}{\partial f_k}(f_0)$$
$$+ (f_{dp} - f_0)\frac{\partial X_G}{\partial f_{dp}}(f_0), \qquad (A6)$$

where $X_G(f_0)$ is shorthand for $X_G(f_0, f_0, f_0; f_0)$.

The combination of A4, A5, and A6 yields

$$\Phi_{dp} = \Phi_0 + \kappa\sum_{k=1}^{3} f_k\frac{\partial X_G}{\partial f_k} + \kappa f_{dp}\frac{\partial X_G}{\partial f_{dp}} - \tau f_{dp}, \qquad (A7)$$

where various constant terms were absorbed into Φ_0.

Equation A7 shows that variations in primary frequency f_k affect DP phase in two possible ways. Variations in f_k that cause f_{dp} to change result in a "direct" effect on Φ_{dp} via the last term. This simply reflects the average group delay τ of the collective incoming traveling waves toward the generation site. It is the main effect of changing the primary frequencies, which incorporates the travel times toward the generation site. The main effect does not require the generation site X_G to shift with the variation in stimulus frequencies. The other two contributions to changes in Φ_{dp}, the second and third terms of A7, occur through shifts of the generation site X_G. Both terms contain the factor κ, the wave

number of the traveling waves at the generation site. The third term describes the changes of X_G with f_{dp}; this contribution merges into the main effect of the fourth term. The second term is the only term that can cause phase changes mediated by fixed-f_{dp} variations of the stimulus frequencies. This establishes the link between residual phase effects and shifts of the location of DP generation.

REFERENCES

BIAN L, SCHERRER NM (2007) Low-frequency modulation of distortion product otoacoustic emissions in humans. J Acoust Soc Am 122:1681–1692

BROWN AM, KEMP DT (1984) Suppressibility of the $2f_1$-f_2 stimulated acoustic emissions in gerbil and man. Hear Res 13:29–37

FAULSTICH M, KÖSSL M (2000) Evidence for multiple DPOAE components based upon group delay of the $2f_1$-f_2 distortion in the gerbil. Hear Res 140:99–110

GOLDSTEIN JL, KIANG NYS (1968) Neural correlates of the aural combination tone $2f_1$-f_2. Proc IEEE 56:981–992

KEMP DT, BROWN AM (1986) Wideband analysis of otoacoustic intermodulation. In: Allen JB, Hall JL, Hubbard A, Neely ST, Tubis A (eds) Peripheral auditory mechanisms. Springer, New York, pp 306–313

KEMP DT (1998) Otoacoustic emissions: distorted echoes of the cochlea's traveling wave. In: Berlin CI (ed) Otoacoustic Emissions: basic science and clinical applications. Singular, San Diego, pp 1–60

KEMP DT (2008) Otoacoustic emissions: concepts and origins. In: Manley GA, Fay RR, Popper AN (eds) Active processes and otoacoustic emissions. Springer, New York, pp 1–39

KIM DO, MOLNAR CE, MATTHEWS JW (1980) Cochlear mechanics: nonlinear behavior in two-tone responses as reflected in cochlear-nerve-fiber responses and in ear-canal sound pressure. J Acoust Soc Am 67:1704–1721

KNIGHT RD, KEMP DT (2000) Indications of different distortion product otoacoustic emission mechanisms from a detailed f1, f2 area study. J Acoust Soc Am 107:457–473

O'MAHONEY CF, KEMP DT (1995) Distortion product otoacoustic emission delay measurement in human ears. J Acoust Soc Am 97:3731–3735

MARQUARDT T, HENSEL J, MROWINSKI D, SCHOLZ G (2007) Low-frequency characteristics of human and guinea pig cochleae. J Acoust Soc Am 121:3628–3638

MARTIN GK, LONSBURY-MARTIN BL, PROBST R, SCHEININ SA, COATS AC (1987) Acoustic distortion products in rabbit ear canal. II. Sites of origin revealed by suppression contours and pure tone exposures. Hear Res 28:191–208

MEENDERINK SWF, NARINS PM (2007) Suppression of distortion product otoacoustic emissions in the anuran ear. J Acoust Soc Am 121:344–351

MEENDERINK SWF, VAN DER HEIJDEN M (2010) Reverse cochlear propagation in the intact cochlea of the gerbil: evidence for slow traveling waves. J Neurophysiol 103:1448–1455

MOULIN A, KEMP DT (1996A) Multicomponent acoustic distortion product otoacoustic emission phase in humans. I. General characteristics. J Acoust Soc Am 100:1617–1639

MOULIN A, KEMP DT (1996B) Multicomponent acoustic distortion product otoacoustic emission phase in humans. II. Implications for distortion product otoacoustic emissions generation. J Acoust Soc Am 100:1640–1662

PAPOULIS A (1962) Fourier integral and its applications. McGraw-Hill, New York

PLOMP R (1965) Detectability of threshold for combination tones. J Acoust Soc Am 37:1110–1123

RHODE WS (2007) Basilar membrane mechanics in the 6–9 kHz region of sensitive chinchilla cochleae. J Acoust Soc Am 121:2792–2804

ROBLES L, RUGGERO MA (2001) Mechanics of the mammalian cochlea. Physiol Rev 81:1305–1352

SCHNEIDER S, PRIJS VF, SCHOONHOVEN R (1999) Group delays of distortion product otoacoustic emissions in the guinea pig. J Acoust Soc Am 105:2722–2730

SCHNEIDER S, PRIJS VF, SCHOONHOVEN R (2003) Amplitude and phase of distortion product otoacoustic emissions in the guinea pig in an (f1, f2) area study. J Acoust Soc Am 113:3285–3296

SCHROEDER MR (1975) Amplitude behavior of the cubic difference tone. J Acoust Soc Am 58:728–732

SHERA CA, GUINAN JJ JR (1999) Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. J Acoust Soc Am 105:782–798

SHERA CA, TALMADGE CL, TUBIS A (2000) Interrelations among distortion-product phase-gradient delays: their connection to scaling symmetry and its breaking. J Acoust Soc Am 108:2933–2948

SMOORENBURG GF (1972) Audibility region of combination tones. J Acoust Soc Am 52:603–614

VAN DER HEIJDEN M, JORIS PX (2003) Cochlear phase and amplitude retrieved from the auditory nerve at arbitrary frequencies. J Neurosci 23:2124–2128

VAN DER HEIJDEN M, JORIS PX (2006) Panoramic measurements of the apex of the cochlea. J Neurosci 26:11462–11473

VICTOR JD, SHAPLEY RM, KNIGHT BW (1977) Nonlinear analysis of cat retinal ganglion cells in the frequency domain. Proc Natl Acad Sci U S A 74:3068–3072

VON BÉKÉSY G (1960) Experiments in hearing. McGraw-Hill, New York

WEVER EG, LAWRENCE M (1954) Physiological Acoustics. Princeton University Press, Princeton

WHITEHEAD ML, STAGNER BB, MARTIN GK, LONSBURY-MARTIN BL (1996) Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency. J Acoust Soc Am 100:1663–1679