JARO
Journal of the Association for Research in Otolaryngology

Psychophysical Estimates of Cochlear Phase Response: Masking by Harmonic Complexes

JENNIFER J. LENTZ AND MARJORIE R. LEEK

Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, DC 20307, USA

ABSTRACT

Harmonic complexes with identical component frequencies and amplitudes but different phase spectra may be differentially effective as maskers. Such harmonic waveforms, constructed with positive or negative Schroeder phases, have similar envelopes and identical long-term power spectra, but the positive Schroeder-phase waveform is typically a less effective masker than the negative Schroeder-phase waveform. These masking differences have been attributed to an interaction between the masker phase spectrum and the phase characteristic of the basilar membrane. To explore this relationship, the gradient of stimulus phase change across masker bandwidth was varied by systematically altering the Schroeder-phase algorithm. Observers detected a signal tone added in-phase to a single component of a masker whose frequencies ranged from 200 to 5000 Hz, with a fundamental frequency of 100 Hz. For signal frequencies of 1000–4000 Hz, differences in masking across the harmonic complexes could be as large as 5–10 dB for phase gradients changing by only 10%. The phase gradient that resulted in a minimum amount of masking varied with signal frequency, with low frequencies masked least effectively by stimuli with rapidly changing component phases and high frequencies masked by stimuli with more shallow phase gradients. A gammachirp filter was implemented to model these results, predicting the qualitative changes in curvature of the phase-by-frequency function estimated from the empirical data.

In some cases, small modifications to the gammachirp filter produced better quantitative predictions of curvature changes across frequency, but this filter, as implemented here, was unable to accurately represent all the data.

Keywords: masking, phase, auditory filters, harmonic complexes

INTRODUCTION

Although spectral analysis has long dominated our understanding of auditory processing, in recent years, temporal waveform information has come to be recognized as contributing significantly to auditory perception. Cochlear analyzing mechanisms have been likened to a bank of overlapping bandpass filters, whose spectral and temporal characteristics interact with acoustic stimuli to produce changes in the shapes of waveforms. The amplitude characteristics of the filters have been described extensively (e.g., Patterson and Moore 1986; Glasberg and Moore 1990; Leek and Summers 1993a; Rosen and Baker 1994) but little is known about the phase response.

The lack of information concerning the phase characteristics of cochlear filtering has been especially troubling in the implementation of spectro-temporal models of hearing. Patterson et al. (1995) assumed a gammatone filter in the initial versions of the Auditory Image Model. Moore et al. (1989) made a similar assumption to support modeling of gap detection results, while acknowledging that the true phase characteristic was unknown in both normal and impaired ears. However, Kohlrausch and Sander (1995)

UNFILTERED STIMULI   **FILTERED STIMULI**

Negative Schroeder (C=-1.0)

Crest Factor = 1.83   Crest Factor = 3.51

Positive Schroeder (C=+1.0)

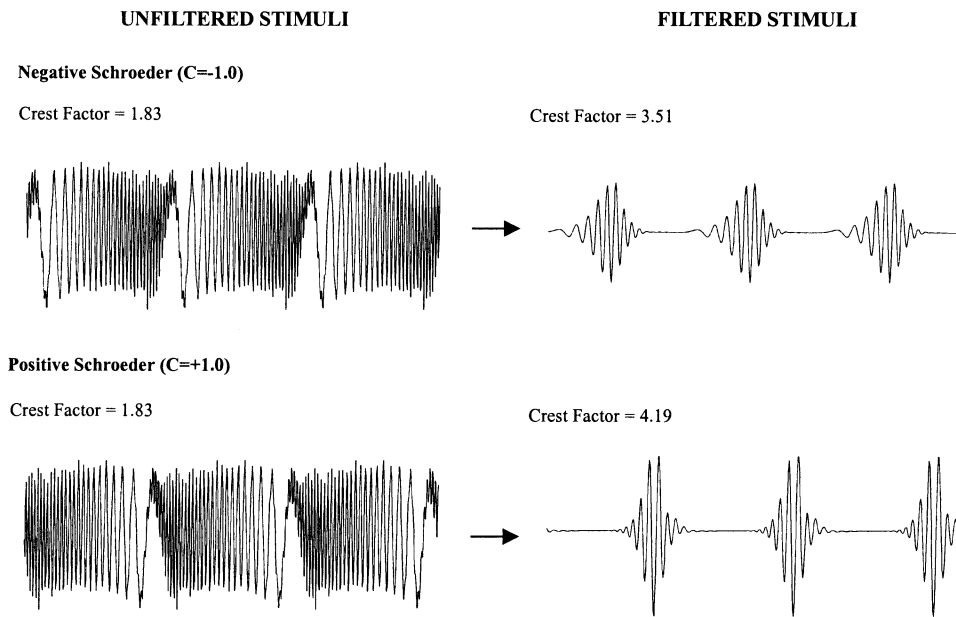Crest Factor = 1.83   Crest Factor = 4.19

FIG. 1. Two Schroeder-phase harmonic maskers are shown in the left panels. The upper-left panel shows the negative Schroeder-phase masker, and the lower-left panel shows the positive Schroeder-phase masker. In the right panels, the Schroeder-phase waveforms were passed through a computer-simulated gammachirp filter with a resonant frequency of 2000 Hz to generate the filtered waveforms shown.

reported that their masking data were not well-modeled by the gammatone filter and suggested that the phase response was almost certainly incorrect. Carlyon (1996) also pointed out that a more realistic phase response was needed in order to model results from a study using complex dynamic maskers. Questions concerning the perceptual effects of abnormal cochlear filtering in hearing-impaired listeners have been restricted primarily to the magnitude response, with only general reference to possible interactions between the stimulus phase spectrum and the phase response of the filters (Leek and Summers 1993b; Leek et al. 1996).

No direct measurements of the phase characteristics of human auditory filters have been made, but some investigators have incorporated impulse responses derived from mammalian VIIIth nerve data in descriptions of psychophysical data (Moore et al. 1989; Patterson et al. 1995). Patterson et al. (1987) discussed the relative value of several choices of auditory filter phase. They concluded that a time-domain filter described by the gammatone function, corresponding to the impulse response that reflects patterns of auditory nerve firing, was the best candidate. This function had been described earlier by de Boer and de Jongh (1978) to account for physiological data obtained using the reverse correlation technique. Irino and Patterson (1997) later modified this time-domain auditory filter to include an onset "chirp," allowing the resulting filter to have slope asymmetries about the center frequency in the spectral domain. They termed this more realistic auditory time-domain filter the gammachirp filter and incorporated it into more recent models and simulations of auditory processing along both spectral and temporal dimensions.

Smith et al. (1986) first suggested that masking by selected harmonic complexes might provide insight into the phase response of the human cochlea. They used harmonic complexes with equal-amplitude components as maskers. Component phases were chosen to be either monotonically increasing or decreasing with component frequency following an algorithm developed by Schroeder (1970). The resulting Schroeder-phase waveforms have identical long-term power spectra and equally flat temporal envelopes, but one of them is the time reverse of the other, as shown in the left panels of Figure 1. The negative Schroeder-phase waveform, shown in the top left panel of Figure 1, is generated using negative phase selections and has a monotonically increasing instantaneous frequency within each fundamental period. The negative Schroeder-phase waveform is typically a more effective masker than the positive Schroeder-phase waveform (shown in the bottom left panel), which has positive phases and a monotonically decreasing instantaneous frequency within the fundamental period. Depending on other stimulus characteristics, these two maskers can result in masked threshold differences of 5–25 dB (Smith et al. 1986; Kohlrausch and Sander 1995; Summers and Leek 1998). Because the positive and negative Schroeder-phase waveforms have identical power spectra, the measured masking differences are not consistent with an energy model of masking.

Following preliminary psychophysical and modeling work on the phase dispersive properties of the inner ear by Smith et al. (1986), Kohlrausch and Sander (1995) argued that the basis for the large differences in masking by the two Schroeder-phase waveforms lies partially in an interaction between the stimulus phase spectrum and the phase characteristic

of the basilar membrane. They measured thresholds for a 5-ms tone, added in-phase to a single masker component within a Schroeder-phase waveform having a fundamental frequency of 100 Hz. This short tone occurred at different times within the 10-ms fundamental period of the masker. When the masker was the negative Schroeder-phase waveform, thresholds depended slightly on the time placement of the signal within the period. However, for the positive Schroeder-phase masker, thresholds differed by as much as 20 dB, depending on the time of the signal onset within the masker period. The threshold variability with the positive Schroeder-phase masker is level-dependent, with lesser effects at low and high stimulus levels and more threshold variability across the masker period for moderate-level stimuli (Carlyon and Datta 1997; Summers 2000).

These differences in the masking period patterns as a function of signal onset time provide evidence that the "internal" waveform may be altered from the physical stimulus during auditory processing. Dau et al. (2000) also showed data that support this interpretation, measuring the magnitude of the auditory brainstem response (ABR) to a modified "chirp" stimulus. Taking into account phase dispersion along the length of the basilar membrane, Dau et al. (2000) showed that the magnitude of the ABR was greatest for this "chirp" stimulus, suggesting that a synchronous response of the basilar membrane was produced by a stimulus with a rising frequency sweep. Phase dispersion along the length of the basilar membrane may cause the internal representation of the positive Schroeder-phase waveform to be highly modulated (Kohlrausch and Sander 1995). In a manner reminiscent of masking by amplitude-modulated stimuli, the peaks within a highly modulated internal waveform would lead to high thresholds for the short-duration probe, whereas valleys in the waveform would result in lower thresholds (Zwicker 1976; Buus 1985). For long-duration signals, masking by the positive Schroeder-phase masker would be less than the negative Schroeder-phase masker, as the signal would be heard within the "valleys" of the internally modulated positive Schroeder-phase masker. According to this interpretation, the internal representation of the waveform of the negative Schroeder-phase masker is less modulated, and, therefore, there would be fewer low-amplitude portions to provide a release from masking. It should be remembered that the physical harmonic complexes do not have these characteristics; both the positive and negative Schroeder-phase waveforms have fairly flat temporal envelopes.

Kohlrausch and Sander (1995) described how alterations in the internal waveform shape of the Schroeder-phase maskers might promote an understanding of the phase characteristic of the auditory

system. They reasoned that it is the phase characteristic of a given auditory filter (or channel) that alters the shape of the internal within-channel temporal waveform. The least effective masker, reflecting the ability to hear the probe signal in a "valley" within the internal representation of that masker, might represent an interaction between the input phase spectrum and the phase characteristic of the relevant auditory channel. When the phase spectrum of the masker closely mirrors (and, therefore, counteracts) the internal phase characteristic, the resulting internal waveform at the output of the relevant auditory channel should have a very peaky temporal structure. In fact, it should be similar to an external waveform having all components in phase (e.g., a cosine- or sine-phase harmonic complex). Masking should be minimal in response to such a modulated waveform.

Figure 1 describes the effects of cochlear filtering on the Schroeder-phase stimuli. The left panels show the positive and negative Schroeder-phase stimuli. These stimuli were passed through a gammachirp filter centered around 2000 Hz to generate the waveforms in the right panels of Figure 1. The gammachirp filter counteracts the phase characteristics of the positive Schroeder-phase stimulus to greater extent than it does the negative Schroeder-phase stimulus, leading to different temporal patterns in the output waveforms. The waveform resulting from the negative Schroeder-phase stimulus (upper panels) is less modulated than the waveform resulting from the positive Schroeder-phase stimulus (lower panels).

This logic was implemented in a search for psychophysical reflections of the phase characteristic of the basilar membrane. The phase spectra of the Schroeder-phase harmonic complexes represent a constant curvature of the phase-by-frequency function. In order to estimate the phase characteristic across frequency, differences in masking effectiveness were measured for waveforms constructed with systematically modified Schroeder-phase stimuli, resulting in changes in the phase-by-frequency gradient of each masker stimulus. The psychophysical data of Kohlrausch and Sander (1995) and phase descriptions in the animal literature (e.g., Ruggero et al. 1997) suggest that the curvature of the phase-by-frequency function of the basilar membrane is not constant along its length. Therefore, the modified Schroeder-phase stimulus whose phase characteristic best mirrors the phase response of the basilar membrane must be different for different frequencies (i.e., places of maximum stimulation along the membrane). By identifying the phase gradient producing the least masking for a given signal frequency, the curvature of the phase-by-frequency characteristic of the underlying auditory system might be estimated.
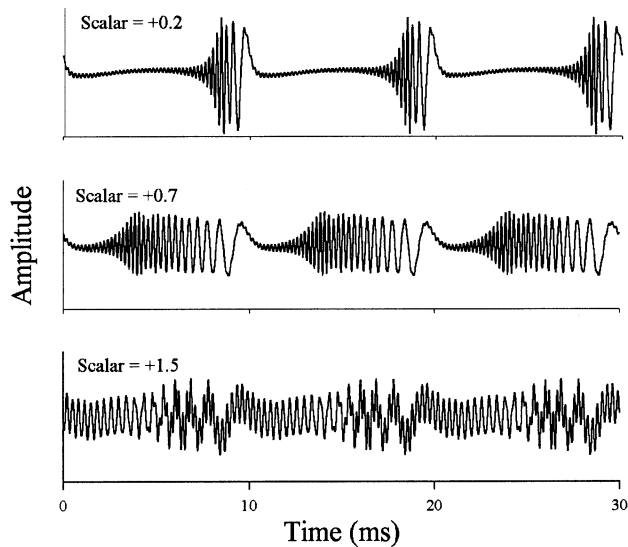
FIG. 2. Three sample stimulus waveforms are shown. Scalars (*C*) of +0.2, +0.7, and +1.5 were used to generate the stimuli in the upper, middle, and lower panels, respectively.

METHODS

Stimuli

The maskers were constructed by summing equal-amplitude cosine tones, harmonically spaced every 100 Hz from 200 to 5000 Hz. The starting phase of each component was selected according to a modification of the algorithm given by Schroeder (1970):

$$\theta_n = C\pi n(n+1)/N \qquad (1)$$

where θ_n represents the phase of the nth harmonic, N is the total number of harmonics, and C is a scalar. Maskers were generated for scalar values C selected to sample the range between -1.0 and $+2.0$, with approximately ten scalars tested for each observer. Mehrgardt and Schroeder (1983) and Schreiner et al. (1983) first presented a scaled modification of the Schroeder-phase algorithm to produce "gaps" in the temporal structure of waveforms. When $C = +1.0$ or -1.0, the original Schroeder-phase waveforms are generated (see Fig.1). Three additional examples of the masker waveforms, with scalars $C = +0.2$, $+0.7$, and $+1.5$, are presented in Figure 2. The maskers were 400 ms in duration, with 30-ms raised-cosine rise/fall times.

The phase arrangement of the positive and negative Schroeder-phase maskers is such that the frequency sweep from 200 to 5000 Hz occurs exactly within the fundamental period of the waveform (for these stimuli, 10 ms). Smaller scalars produce faster frequency sweeps and larger scalars produce slower frequency sweeps. When the frequency sweep rate is high ($-1.0 < C < +1.0$), there are low-energy portions within the masker period (see the top two panels of Fig. 2). For

$C < -1.0$ and $C > +1.0$, however, the entire frequency sweep from 200 to 5000 Hz lasts longer than the 10-ms period of the masker. The high-frequency components (for $C > +1.0$) have phase leads which cause them to wrap across more than one period and thus appear in the previous masker period (see Fig. 2, bottom panel). The phase wrapping leads to a more complex waveform, which, however, probably only minimally affects the masking results reported here. This issue will be revisited as the results are presented.

The signal to be detected was a tone added to its corresponding frequency component in the masker with the same starting phase as the masker component. Four signal frequencies were tested: 1000, 2000, 3000, and 4000 Hz. The duration of the signal was 200 ms, with 30-ms rise/fall times, and it was temporally centered within the masker. The signal was presented at a fixed level of 40 dB SPL. Threshold was defined as the masker power in dB SPL needed to mask the fixed-level signal tone at a performance level of 79% correct detections. These values will be presented as overall masker levels (masker level per component is 17 dB less, given that there are 49 components in the masker).

Informal listening to these maskers and signals suggested that, when the signal was a 4000-Hz tone, there was a chance of some confusion with the edge pitch resulting from the abrupt spectral offset of the masker at 5000 Hz. Therefore, a second set of maskers was constructed with a spectral ramp in which the five highest-frequency harmonic components were progressively attenuated by 3 dB (see Alcantara and Moore 1995). Thus, the 49th component was attenuated by 15 dB with respect to the other components of the masker. These "ramped" stimuli were used only for a signal frequency of 4000 Hz when there appeared to be some confusion between masker and signal.

The signal and masker stimuli were generated using two channels of a 16-bit digital-to-analog converter (DAC) (TDT DD1) at a sampling rate of 20.48 kHz. The outputs of the DAC were low-pass filtered at 8.5 kHz (TDT FLT3 105 dB/oct attenuation band with maximum attenuation of 65 dB), attenuated, and summed. The resulting stimulus was fed into one earphone of a TDH-49 headset.

Procedure

Following the procedure used by Summers and Leek (1998), the overall level of the masker was varied in a standard/two-alternative forced-choice task, and a 3-up, 1-down tracking procedure was used to estimate masked threshold (Levitt 1971). When the interval containing both masker and signal was chosen correctly on three consecutive trials, the level of the

masker was increased; one incorrect selection was followed by a decrease in the masker level. Observers were seated in a sound-attenuating room and indicated which interval contained the signal tone by touching a designated area on a touch-screen terminal. Correct-answer feedback was provided after each trial.

On every trial, the masker alone was presented in the first interval. The other two intervals contained either masker alone or masker-plus-signal, presented in random order with equal probability. The interstimulus interval was 400 ms. At the start of each experimental session, the observer was presented with four practice trials in which the signal was clearly audible within the masker. The level of the masker on each adaptive track was set initially to 10–15 dB below the listener's estimated threshold. Masker levels were changed in steps of 4 dB at the start of each track and then reduced to 1-dB steps after three reversals in direction. The threshold search continued until the track direction had reversed 12 times. The mean of the masker levels at the last eight reversal points was taken as threshold. Three threshold tracks were collected for each experimental condition. The final reported thresholds are the means of the three track thresholds.

Observers and order of data collection

Four normal-hearing observers, ranging in age from 28 to 52 years, participated. All observers had pure tone audiometric thresholds within 15 dB of the 1989 ANSI standard at 250–6000 Hz. All listeners had previous experience with psychoacoustic tasks. Observers 1 and 2 are the first and second authors, respectively.

Thresholds were measured using a randomized block design. For each observer a signal frequency was chosen at random and thresholds were measured for scalars in random order (except that, in many instances, scalars $C = -1.0, -0.5,$ and 0 were tested last). Each threshold was measured at least three times over several days. After three replicates of each scalar had been tested for a selected signal frequency, a different frequency was chosen and the process repeated. No two replications were completed on the same day. To assess possible practice effects, thresholds were spot-checked for several scalars at each frequency. These checks resulted in replications that were within one standard deviation of thresholds obtained previously.

Following data collection at all signal frequencies, because of a possible confusion between the masker edge pitch at 5000 Hz and the 4000-Hz signal, listeners were retested at 4000 Hz using the maskers with a high-frequency spectral ramp. Thresholds reported for the 4000-Hz signal for Observers 2, 3, and 4 were measured with the spectrally ramped maskers (Observer 1

showed no difference in thresholds whether the ramp was present or not).

RESULTS

Threshold data are plotted as a function of scalar (C) in Figure 3. Each panel represents data for one signal frequency (1000, 2000, 3000, or 4000 Hz), and thresholds for individual observers are shown as different symbols within the panels. Error bars indicate standard errors of the mean across three threshold replicates. Recall that the signal level was fixed and the threshold values are the level of the harmonic-complex maskers required to mask the signal; data toward the bottom of the panels indicate more effective maskers, while higher thresholds reflect poorer maskers.

For each signal frequency, there is a nonmonotonic relationship between masker level at threshold and the scalar being tested. At the most negative scalar values, masker levels tend to be lowest, indicating highly effective maskers. As the scalar approaches zero, masker levels generally rise slightly and continue to rise beyond scalars of zero until a maximum level is reached. For scalar values larger than that associated with the highest masker level, levels fall slightly with further increases in scalar. Within a single observer and signal frequency, the difference in masking generated by changing the scalar can be quite large, often exceeding 10 dB. Summers and Leek (1998) reported similar masking differences of 8–15 dB between the positive and negative Schroeder-phase maskers at several probe frequencies. These data are also in accord with those of Smith et al. (1986) and Kohlrausch and Sander (1995) who measured a masking difference between the positive Schroeder-phase and negative Schroeder-phase maskers on the order of 10–20 dB for signal frequencies of 1100, 2200, and 4400 Hz. Their stimuli differed considerably from those used here in that the bandwidths and number of components were different and the signal level was varied to find threshold in a fixed-level masker. Nonetheless, we find the same pattern of results for these different threshold measurement techniques.

Across nearly all observers and signal frequencies there is a peak in the functions relating threshold masker level to scalar value. The scalar associated with this maximum level (the peak scalar) is generally similar across listeners. Within each panel of Figure 3, a shaded region highlights the range of peak scalars across all listeners. The larger the width of the shaded region, the greater the variability of peak scalar location across listeners. Some signal frequencies have more variability than others (e.g., 4000 Hz has a narrower shaded region than the other signal frequencies) but, overall, the range of peak scalars is rather
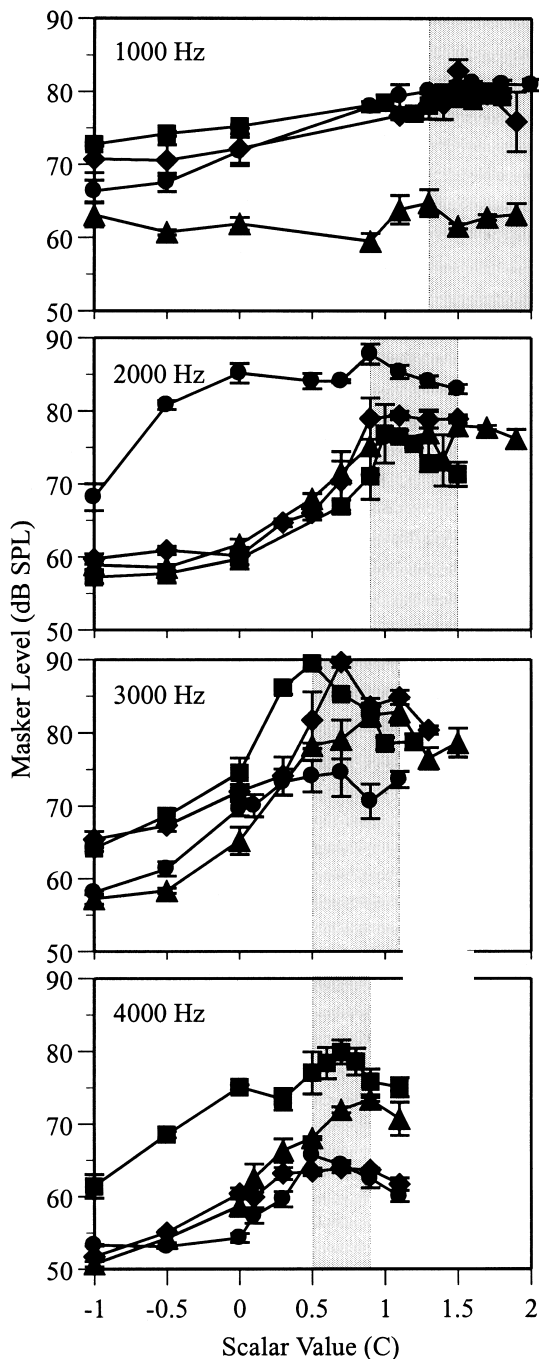
FIG. 3. Threshold (masker power in dB SPL needed to mask a 40 dB SPL signal tone) is plotted as a function of scalar value for four signal frequencies and four observers. Signal frequencies of 1000–4000 Hz are shown from the top to the bottom panel, respectively. Error bars are the standard errors of the mean across three replicate threshold estimates. The different symbols denote different observers: (−■−) Obs 1, (−♦−) Obs 2, (−▲−) Obs 3, (−●−) Obs 4.

small across listeners. Quantifying the peak scalar at 1000 Hz is difficult because the individual functions do not have a sharp peak like the functions at the other signal frequencies. The true peak scalar may be beyond the range of scalars tested here but it appears

to be greater than +1.3 for all listeners. The highlighted region at 1000 Hz includes all scalar values greater than +1.3. (To further explore this result, we tested scalar values greater than +2.0 during pilot data collection. However, for maskers generated with these scalars, we did not see a continuation of the drop in thresholds with increasing scalar; thresholds rose and fell in a manner unrelated to increases in scalar value.)

For all conditions, the peak scalar occurs at scalar values above zero, reflecting less masking by the positively scaled Schroeder-phase maskers than by the negatively scaled Schroeder-phase maskers. In addition, the location of the shaded region varies across frequency. Larger peak scalars are measured for the lower signal frequencies, and smaller peak scalars are associated with higher signal frequencies. This can be seen in the movement of the gray-shaded region toward lower scalars as signal frequency increases from the top panel to the bottom panel in Figure 3.

Peak scalars larger than +1.0 were measured for 1000- and 2000-Hz signals. Recall that, for scalars larger than +1.0, the high-frequency components "wrap" across the waveform periods occurring in the latter end of the previous masker period (see discussion in Methods section and bottom panel of Fig. 2). For the 2000-Hz signal, with a peak scalar of 1.125, the first component to wrap to the previous masker period is approximately 4500 Hz. At this signal frequency, it is unlikely that the wrapped frequencies affect masking by the peak scalar masker, as they are sufficiently distant from the signal frequency. For a signal of 1000 Hz at the peak scalar (C_{peak} > +1.3), the first component to wrap to the end of the masking period is around 3900 Hz. For cases where the C_{peak} is much greater than +1.3, masking may be influenced by the wrapping of frequencies. The C_{peak} is not well-defined at 1000 Hz and phase wrapping may adversely affect masking. Carlyon and Datta (1997) showed that masking of a 1000-Hz signal by Schroeder-phase harmonic complexes was most influenced by components at frequencies ranging from approximately 500 to 1500 Hz, with some contributions from more distant components. At 1000 Hz, masking due to the relatively remote wrapped frequencies is probably limited, though not necessarily negligible.

In addition to the strong relationship between threshold and scalar, Figure 3 also demonstrates large masking differences across listeners, shown in the separations among the different symbols within each panel. For example, at 4000 Hz, Observer 1 (squares) shows little masking overall (high thresholds) while for Observer 4 (circles) the 4000-Hz tone was masked much more effectively. The threshold differences between these two observers are on the order of 15 dB, regardless of scalar value. These masking differences

among observers are not consistent across signal frequencies: A listener who shows little masking at one signal frequency may not show the same trend at a different signal frequency. Note Observer 4 as an extreme example (circles). Observer 4's thresholds exceed those of all observers at 2000 Hz, indicating the most resistance to masking by the different stimuli. However, at 3000 and 4000 Hz, the same observer shows the most masking at scalar values above zero. The source of these individual differences in level of masking is unclear. It is unlikely that they are due to the order of testing for each observer, as no patterns of particularly low or high masking levels could be related to the sequence of tested frequencies. Further, the relatively small standard errors across the three replicates for each threshold, as well as the confirmations found in the additional checks of some thresholds, do not suggest an effect of training or fatigue leading to differences in thresholds for some frequencies. The individual differences appear to be unrelated to the points of minimum masking, however, as little variability across listeners is present in the determination of peak scalar.

High thresholds (i.e., poor maskers) might indicate a highly modulated "internal" waveform in which an observer could detect a signal easily in the masker valleys. A harmonic complex constructed with the peak scalar associated with a given signal frequency is assumed to approximately mirror the internal phase characteristic in that frequency region. That is, the peak scalar generates a masker with a phase-by-frequency curvature that reflects the curvature matched to a particular auditory channel. These findings are consistent with the suggestion of Kohlrausch and Sander (1995) that the phase curvature of a single auditory filter is negative. Positive scalars counteract the phase curvature of an auditory filter, leading to a more highly peaked waveform. Their estimates also indicate that the curvature of higher-frequency auditory filters is less than lower-frequency auditory filters. The thresholds reported here provide a more detailed characterization of the change in curvature with increasing signal frequency. The masker stimuli used in this study have a constant curvature given by the equation:

$$\frac{d^2\theta}{df^2} = C\frac{2\pi}{Nf_0^2} \qquad (2)$$

where N is the number of harmonics and f_0 is the fundamental frequency of the masker (see Kohlrausch and Sander 1995). Equation (2) indicates that the curvature of the Schroeder-phase stimulus is dependent on the number of harmonics. Such a dependence of the phase curvature on number of masker components complicates comparisons between thresholds obtained by Kohlrausch and Sander (1995) for positive

and negative Schroeder-phase stimuli and those measured here for $C = \pm 1$, as the masker phase curvatures differ in the two studies. An estimate of the curvature of the phase characteristic of the auditory filter at each signal frequency is given by determining the peak scalar and using Eq. (2) to find the curvature of the corresponding stimulus.

Peak scalars averaged across the four observers are plotted as a function of frequency in Figure 4 (filled circles). The right ordinate indicates the phase curvature of the resulting stimulus. The arrow above the data point at 1000 Hz indicates that the true peak scalar is probably greater than those reported here. A repeated-measures analysis of variance indicates a significant effect of signal frequency on the peak scalar ($F_{(3,9)} = 11.70$, $p < 0.002$). Peak scalars generally decrease with increasing signal frequency. This corresponds with a magnitude of stimulus phase curvature that decreases with increasing frequency. If the curvature of auditory filters is the mirror image of the curvature of the stimulus generated by the peak scalar, the magnitude of the curvature of auditory filters also decreases with increasing center frequency, but the sign is negative. The crosses connected with a dotted line indicate estimated curvatures based on Kohlrausch and Sander's (1995) data set. The curvature at 1100 Hz was estimated explicitly by Kohlrausch and Sander. Curvatures at 2200 and 4400 were estimated using Figure 12 from Kohlrausch and Sander (1995). Following their logic, we determined the fundamental frequency at which thresholds for the positive Schroeder-phase masker and the sine-phase masker were equal, as shown in their Figure 12 (bottom panels). The phase curvature of the auditory filter centered at the signal frequency would cancel half of the phase curvature of the positive Schroeder-phase masker created with this fundamental frequency and would introduce half its phase curvature into the sine-phase masker. Therefore, the curvature of the positive Schroeder-phase masker, with the fundamental frequency at the intersection of the positive Schroeder-phase and the sine-phase maskers shown in Figure 12 of Kohlrausch and Sander (1995), may be estimated to reflect twice the curvature of the phase characteristic of the auditory filter. At 1100 Hz, Kohlrausch and Sander estimated that the phase curvature lies between -3.39×10^{-5} and -2.33×10^{-5} rads/Hz^2. The data point on Figure 4 is the average of these two curvatures. In the current experiment, the estimated average curvature of -1.86×10^{-5} rad/Hz^2 is somewhat closer to zero than Kohlrausch and Sander's estimate. However, it may be underestimated here because of the lack of a distinct peak in the functions in Figure 3 at 1000 Hz. To produce a stimulus with the same curvature as that used in the Kohlrausch and Sander experiment using frequencies ranging from 200 to
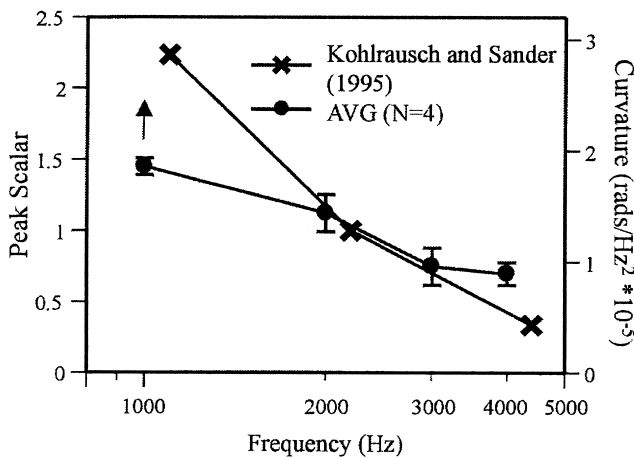
FIG. 4. Average peak scalars (shown on the left axis) and the corresponding curvature (shown on the right axis) are plotted as a function of signal frequency. The filled symbols indicate the psychophysically measured peak scalars. Error bars indicate the standard errors across four observers. The large ✕'s denote the peak scalars estimated from Kohlrausch and Sander's (1995) data set.

5000 Hz, scalars of $C = 2.6$ and $C = 1.8$ (scalars where more substantial phase wrapping may affect masking) would need to be used. The estimated curvatures for frequencies of 2200 and 4400 Hz are similar to the current estimates in those frequency regions, despite the differences in measurement techniques between these two studies.

DISCUSSION

In agreement with past studies, these data show that harmonic stimuli with different phase spectra can produce different amounts of masking (Smith et al. 1986; Kohlrausch and Sander 1995; Carlyon and Datta 1997; Summers and Leek 1998). This differential masking cannot be attributed to changes in energy with the addition of a signal because the long-term power spectra are identical, regardless of the phase spectra. Listening in the valleys of the external masker also cannot adequately explain these results because a stimulus with a flat temporal envelope often is a poorer masker than a stimulus with a scalar that produces substantial valleys in its temporal structure. The data of Kohlrausch and Sander show that, in many instances, positive Schroeder-phase maskers were less effective than maskers where all components were added in sine phase. Kohlrausch and Sander (1995) suggested that these masking differences were the result of alterations of internal waveform shape due to interactions with the phase response of the basilar membrane. Less effective maskers reflect a cancellation of the stimulus phase characteristic by phase changes imposed by cochlear processing. The

resulting internal representation of the waveform is highly modulated, and the auditory system can detect the signal in the low-amplitude portions of each masker period.

This hypothetical effect of cochlear processing was explored here by measuring masking effectiveness of harmonic complexes with systematic changes in their phase spectra. The masking differences observed among these masker complexes most likely reflect the stimulus-phase and internal-phase interactions occurring within a single auditory channel. The peak scalar, determined by masking within a limited spectral region around the signal frequency, varies as a function of frequency, implying that the magnitude of the curvature of auditory filter phase decreases as frequency is increased. Thus, component phases within the masker stimuli at higher frequencies are altered less from the external stimulus than phases at lower frequencies. Fewer phase alterations also suggest that the temporal waveform shape at the level of the cochlea is better preserved at high frequencies than low. The variation of local phase curvatures within auditory channels provides evidence of more global changes in the curvature of the phase characteristic along the length of the cochlea. Thus, at higher frequencies within these masker waveforms, the phase dispersive properties of the basilar membrane have less effect on the temporal structure of the waveform, and the external shape of the waveform is more accurately preserved internally.

Data obtained from Schroeder-phase masking experiments have been modeled to some extent with basilar membrane models (Smith et al. 1986; Kohlrausch and Sander 1995) as well as with the gammatone filter described by Patterson et al. (1987). Basilar membrane models, such as developed by Strube (1985), have limited applicability to psychophysical data because the filters employed in those models have much broader filter characteristics than have been measured in humans (see comment by Moore in Kohlrausch 1988). These models, however, do predict that the output waveform for the positive Schroeder-phase stimulus is peakier than the negative Schroeder-phase stimulus. Kohlrausch and Sander (1995) used the gammatone model with a more realistic bandwidth for psychophysical data but were unable to account for the large masking difference between the positive and negative Schroeder-phase stimuli with that model. These data may be better described using the gammachirp filter developed by Irino and Patterson (1997). In fact, Micheyl et al. (1999) qualitatively described changes in Schroeder-phase stimuli using a gammachirp filter to explain their otoacoustic emission data. The gammachirp filter is implemented here to determine whether realistic auditory filtering can account for the current results.

Description of the gammachirp filter model and crest factor analysis

The gammachirp filter has a well-developed impulse response in which passband asymmetries present in human psychophysical auditory filters can be realized. The impulse response of the filter is given by

$$g(t) = t^{n-1} e^{-2\pi_r \text{ERB}(f_r)t} \cos(2\pi f_r + c\ln(t) + \phi)$$

where n, b, f_r, and c are free parameters. The order of the filter is defined by n. The bandwidth is given by ERB (f_r), the equivalent rectangular bandwidth of the filter at its resonant frequency f_r in Hz, as defined by Glasberg and Moore (1990), scaled by a bandwidth parameter b. The gammachirp filter takes into account the frequency dependence of frequency selectivity, in that filters centered at higher frequencies are broader than those at lower frequencies (Glasberg and Moore 1990). To account for nonlinear changes in bandwidth and filter asymmetry with level, Irino and Patterson (1997) defined the impulse response of the filter to be level-dependent through a chirp parameter c, which is based on the power of a signal. Depending on the time duration of the window used to estimate the power passing through the filter, the properties of the filter may change for stimuli which have large level changes across their duration. For most of the stimuli used in the current experiment, the gammachirp filter does not have a constant passband or phase characteristic throughout the stimulus duration. The filter's bandwidth is most affected by changes in the bandwidth parameter b. Changes in the chirp parameter have a smaller effect. The chirp parameter has control over the degree of filter asymmetry and changes in phase across the filter's passband. Figure 5 plots the magnitude and phase characteristics of example gammachirp filters with wide and narrow bandwidths. The filters shown in Figure 5 reflect changes imposed by altering the bandwidth parameter b. Note that slower phase changes are associated with gammachirp filters having larger bandwidths. The relationship between the chirp parameter and phase and filter bandwidth is more complex.

The effects of passing waveform with different masking effectiveness through a gammachirp filter centered at 2000 Hz are demonstrated in the right-hand panels of Figure 1. The simulated gammachirp filter was based on a stimulus level of 80 dB SPL. The upper-right panel shows a stimulus that is less peaky than that in the lower-right panel. Recall that, in masking a 2000-Hz tone, the negative Schroeder-phase waveform in the upper-left panel is a more effective masker than the positive Schroeder-phase waveform in the lower-left panel. The parameters defining this gammachirp filter are based on a 3-parameter fit described by Irino and Patterson (1997). Parameters n and b are
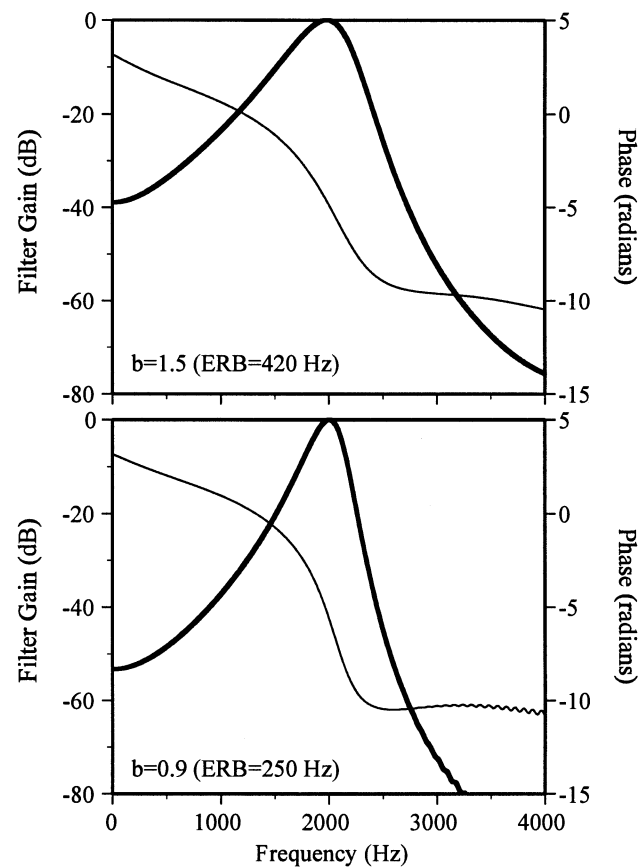

FIG. 5. Examples of two gammachirp filters that show the effects of changing the bandwidth parameter b. The heavy line represents the filter magnitude, keyed to the left ordinate, while the lighter line shows the filter phase, referred to the right ordinate.

4 and 1.51, respectively. The chirp parameter c is 3.88 $-0.109 P_s$, where P_s is the power of the stimulus passing through the filter at a given instant. Further discussions regarding the definition of P_s appear later. For many of the stimuli used in the current experiment, the power differs depending on the duration of the assumed time window of the ear. For Figure 1 and the simulations reported here, the power of the stimulus is computed over a 3-ms time window. This value is based on gap detection experiments where it has been shown that listeners can discern a temporal gap of 2–3 ms (e.g., Plomp 1964), and on estimates of the time constant of temporal modulation transfer functions (e.g., Viemeister 1979; see Viemeister and Plack 1993 for review).

The "peakiness" of the masker waveforms may be quantified by the crest factor, which is the ratio of the peak amplitude in a waveform to the root-mean-square amplitude of the waveform (Hartmann and Pumplin 1988). The internal crest factor is determined by the interaction between the external waveform and the filtering and phase dispersive properties of the basilar membrane. Internal crest factors of the

stimuli used here may be computed by passing the stimuli through a computer-simulated auditory filter. Referring again to Figure 1, after filtering, the crest factor of the negative Schroeder-phase waveform (3.51, upper-right panel) is lower than the crest factor of the positive Schroeder-phase waveform (4.19, lower-right panel), even though prior to filtering these two stimuli had the same crest factor (1.83). Assuming that maskers with high crest factors are less effective, the gammachirp filter predicts the empirical finding that the positive Schroeder-phase waveform is a poorer masker than the negative Schroeder-phase waveform.

Internal crest factors were calculated based on a bank of gammachirp auditory filters for comparison with the empirically measured peak scalars for each signal frequency. The parameters chosen for the gammachirp filter are the same as those used to generate Figure 1; however, the large range of overall stimulus levels in the data set pose an interesting problem for the implementation of the gammachirp filter. In the threshold measurements, the masker level is varied to estimate threshold. Because the signal is added in-phase to a component of the masker, the actual level of that component depends on the threshold level of that masker. Recall that the gammachirp filter is level-dependent—the chirp parameter depends on the level of the signal passing through the filter. As applied here, it is unclear whether this means the level of the signal itself (40 dB SPL), the level of the signal added to its corresponding masking component (ranging between 40 and 73 dB SPL), or the level of the stimulus passing through the filter itself. Irino and Patterson (1997) defined the gammachirp filter using notched-noise data, where the level of the masker was varied and the signal level was fixed. In those experiments, the signal and masker did not overlap in frequency, and, therefore, the level of the signal component was never in question. In addition, notched-noise stimuli usually do not have large changes in the power of the stimuli throughout the waveforms, and, therefore, a relatively static representation of the gammachirp filter can be used. In contrast, in the current experiment, the overall levels of the stimuli differ at masked threshold, and the instantaneous power of the stimuli varies within a masker period. The choice regarding how to define the power passing through the filter has great impact on the results of a gammachirp simulation as the chirp parameter is explicitly defined by this value. Greater stimulus levels will lead to more filter asymmetry and alterations in the phase characteristics of the filter.

If the characteristics of the filter were selected based on the signal level alone (40 dB), a difference in masking between the positive and negative Schroeder-phase stimuli would not be predicted. The phase characteristics of that filter would be very similar to that of a symmetric gammatone filter ($c = 3.88 - 0.109 \times 40 = -0.48$ compared with $c = 0$ for the gammatone). Further, because the signal tone is added to a component of the masker, the level of that component is never truly 40 dB. Thus, in the current application, it seems inappropriate to define P_s in this way. Rather, P_s is defined as the power of the stimulus passing through a single gammachirp filter. This definition was chosen because it is more realistic to assume that the auditory system can determine power only within some frequency region rather than the power of a single spectral component.

The gammachirp filter and crest factor analysis were implemented here by filtering both masker and signal-plus-masker stimuli. To account for off-frequency listening, a bank of 21 gammachirp filters with center frequencies between 80% and 120% of the signal frequency was used. The filter with the largest change in power between the masker-alone stimulus and signal-plus-masker stimulus was taken as the analysis filter. The crest factor of the masker-alone stimulus was computed at the output of the selected analysis filter.

It was assumed that at the output of the gammachirp filter, stimuli with higher crest factors would be poorer maskers than stimuli with lower crest factors. Masked thresholds are assumed to be monotonically related to crest factor. Because stimulus level is critical to the implementation of the gammachirp filter, the effect of stimulus level on crest factors is discussed first. Then, results of the crest factor analysis are compared with the experimental data in two different ways. First, predicted peak scalars are compared with psychophysically measured peak scalars to determine whether the phase changes across the gammachirp filter's passband are sufficient to describe the data. Second, the results of the modeling are examined to determine whether the filter model can explain the large differences in threshold that are measured for different scalars.

Behavior of the gammachirp filter model with level

To evaluate the effects of stimulus level on crest factor, the maskers used in the current experiment were generated at three different overall levels (57, 70, and 90 dB). Maskers and signal-plus-masker waveforms (signals were 40 dB) were filtered using the same parameters as in Figure 1 (taken from Irino and Patterson 1997). Recall that the filtering process attenuates the majority of the stimulus components, resulting in reduced levels at the output of a single filter. Figure
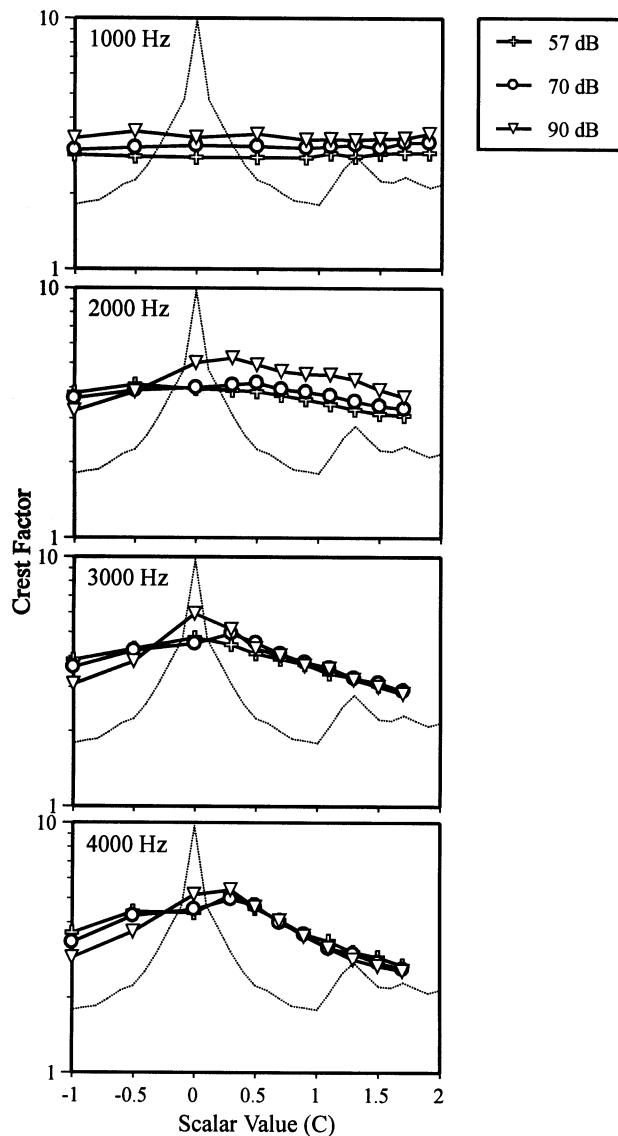
FIG. 6. Crest factor is plotted as a function of scalar value for the experimental stimuli (dotted line) and for the stimuli passed through gammachirp filters centered at 1000, 2000, 3000, and 4000 Hz, at three different stimulus levels.

6 shows crest factors of the masker-alone stimuli, after they have been passed through gammachirp filters, plotted as a function of scalar. Each panel represents a different signal frequency. Three different stimulus levels are shown by the different symbols. The crest factors for the unfiltered stimuli are indicated by the dotted line in each panel. Stimulus level does not affect the crest factors of the unfiltered stimuli.

The largest crest factor is found for the unfiltered stimuli at a scalar of zero, as all the component phases are the same and the waveform is the "peakiest." Regarding the unfiltered stimuli, large changes in crest factor are associated with small changes in scalar, especially for scalars near zero. Large changes in crest factor with scalar also occur for the filtered stimuli,

although these changes are smaller than for the unfiltered stimuli. The maximum crest factor for the unfiltered stimuli is approximately 10, whereas the maximum crest factor for the filtered stimuli is less than 6. Filtering attenuates many of the stimulus components, leading to fewer harmonic interactions and, therefore, less change in crest factor with scalar value.

At 1000 Hz, the function relating crest factor to scalar value is even flatter than the pattern in the data (Fig. 3). This result is due to the relatively narrow filter bandwidth at 1000 Hz, where very few components interact with one another. For the other three frequencies, the modeling results show smooth functions relating crest factor to scalar value. As with the psychophysical data, at negative scalar values, crest factor gradually increases with increasing scalar value. Once a peak in the function is reached, crest factor decreases with increasing scalar. Within a single frequency, the maximum crest factor increases with increasing stimulus level. Yet, crest factors are not elevated with increasing stimulus level at all scalars. For example, for signal frequencies of 2000–4000 Hz at $C = -1.0$, crest factors tend to decrease with increasing level. Thus, there is an interaction between increasing filter bandwidth and the cancellation of the phase characteristics of the stimuli.

As with the experimental data, there is a peak scalar associated with the maximum crest factor for most stimulus levels and signal frequencies. The peak scalar changes depending on the overall level of the stimulus. For example, at 2000 Hz, for a stimulus level of 57 dB, the maximum crest factor is measured at a scalar of -0.5. However, the peak scalar at a stimulus level of 70 dB is +0.5, and at 90 dB it is +0.3. A similar dependence of peak scalar on stimulus level is found at 3000 Hz but not at 4000 Hz. These shifts in peak scalar are partially a result of changes in the phase characteristic of the gammachirp filter with increasing stimulus level. The effects of off-frequency listening may cause shifts in the peak scalar as well. Filters with higher resonant frequencies may be selected as the analysis filters at greater stimulus levels. Regardless of the stimulus level, the peak scalar is always measured to be less than +0.5. These results contrast with the experimental data in which peak scalars ranged between +0.7 and +1.3. While Figure 6 simply serves to illustrate the effects of level on filtering Schroeder-phase stimuli, it also indicates that crest factors extracted from the gammachirp filter model, as implemented here, will not easily account for the psychophysical data.

The gammachirp filter parameters suggested by Irino and Patterson (1997) are based on notched-noise data of Lutfi and Patterson (1984), Rosen and Baker (1994), and Moore et al. (1990). The parameters were generated using assumptions that take into account only the spectral characteristics of sound rather than

temporal characteristics. The power spectrum model of masking has been shown to be inadequate in a number of studies, especially those tasks which implicate temporal cues (cf. Richards 1988; Kidd et al. 1989). Lentz et al. (1999) suggested that the auditory filter bandwidths recommended from the notched-noise task may be too wide to use for all psychophysical tasks. Expanding on this finding, recall from Figure 5 that narrower filter bandwidths are associated with faster phase changes across a filter's passband. Faster phase changes would be associated with a greater peak scalar. Therefore, one way to increase the predicted peak scalar would be to decrease the bandwidth of the filter. The effects of narrower filter bandwidths is discussed in the next section.

Predictions of peak scalars from gammachirp filter simulations

Can the gammachirp filter accurately predict the peak scalars and, therefore, the estimated phase curvature of cochlear processing? A crest factor analysis of simulated filter outputs was carried out for each observer and frequency, using the masker waveforms corresponding to each scalar value. However, the level dependence of the gammachirp filter, while realistic, introduces some difficulty in attempting to model the masking data and predict a peak scalar. Figure 6 shows that the scalar resulting in the highest crest factor at the output of a gammachirp filter varies depending on the input level of the stimulus. The psychophysical data of Figure 3 reflect a large distribution of masker levels at threshold, with distinct patterns relating threshold to scalar value and, at most frequencies, a clear peak in the functions. The peak scalars provide estimates of the curvature of the phase-by-frequency function of cochlear processing at the peak scalar threshold level, but, at other stimulus levels, it is likely that a different curvature would be estimated. If the masker threshold level for each scalar is used in a gammachirp filter simulation, the determination of peak scalar would be biased toward the highest masker level, i.e., the actual peak scalar in the data, which is to be predicted. To reduce that bias, a crest factor analysis was undertaken for filter outputs for each scaled waveform, input at the stimulus level of the peak scalar rather than at its true threshold level. It is thought that this compromise focuses the analysis on the stimulus level of the curvature estimate in each data set and diminishes the unwanted effects of different scalar threshold levels. Three filter bandwidth parameters ($b = 1.5$, $b = 1.1$, and $b = 0.9$) were used because the previous analysis suggested that a narrower bandwidth might lead to increases in the predicted peak scalar.
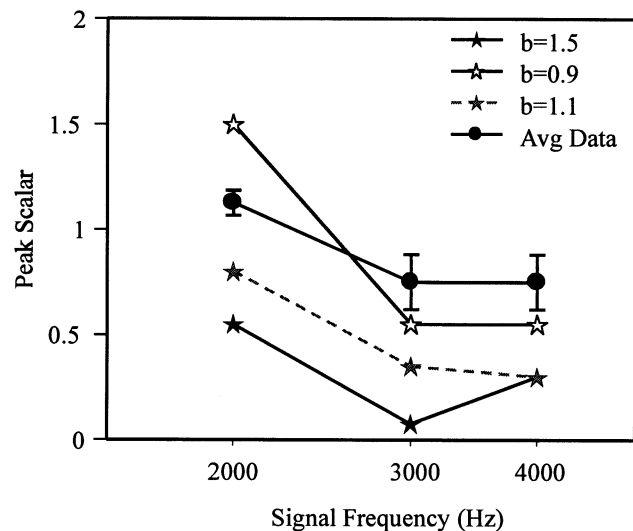

FIG. 7. Average peak scalars at three signal frequencies predicted by gammachirp filters with three different bandwidth parameters. The solid circle is the mean of the measured data, with error bars indicating the standard errors across observers.

Stimuli were generated for all scalars at the maximum threshold level for each signal frequency and observer. Figure 7 shows predicted peak scalar values averaged across observers as a function of signal frequency for the three different bandwidth parameters tested. The filled circles reflect the average of the psychophysically measured peak scalars. Error bars are the standard errors of the mean across the four observers' measurements. The filled stars indicate predicted peak scalars for $b = 1.5$; gray stars connected with a dashed line indicate predicted peak scalars for $b = 1.1$; and the unfilled stars show predicted peak scalars when $b = 0.9$. Peak scalars are not shown for 1000 Hz because of the inability to determine a true peak in the functions relating crest factor to scalar value.

Figure 7 shows that predictions of peak scalars using the gammachirp filter differ from the external stimulus (which would be $C = 0$) and that the predicted peak scalars tend to decrease with increasing signal frequency. This indicates that the gammachirp filter phase characteristic counteracts (at least to some extent) the external phase spectrum. As expected from the results shown in Figure 6, when the bandwidth parameter is 1.5, the peak scalar estimates are much lower than estimates from the data. The gray and unfilled symbols in Figure 7 reflect peak scalar predictions for narrower bandwidths, indicating that narrowing the filter has the effect of increasing the predicted peak scalar. However, for $b = 1.1$, the narrowing of the filter does not raise the peak scalars up to the experimental values. A narrower filter ($b = 0.9$) raises the peak scalar at 2000 Hz to a level that is too great, and yet the peak scalars at 3000 and 4000 Hz

still are not in accord with the data. These values come closer to the experimental data than those obtained using $b = 1.5$, however.

Recall that slower phase changes across the filter's bandwidth are associated with smaller peak scalars. So, the phase changes across the gammachirp filters, especially at 3000 and 4000 Hz, are slower than indicated by the psychophysical data. With these particular parameter specifications, the gammachirp filter does not capture the large phase changes found in the experimental data. However, when the narrower filter bandwidth is included, the gammachirp filter comes closer to predicting the measured peak scalars.

In summary, the gammachirp filter generally can account for masking differences between the positive and negative Schroeder-phase maskers (scalars of ± 1.0), but it may be unable to represent fully the psychophysical data obtained using harmonic maskers with the systematic phase changes used here. For the higher frequencies (i.e., 3000 and 4000 Hz), there would need to be a greater change of phase across the passband in order for the gammachirp filter to fully account for these Schroeder-phase masking data, at least under the assumptions of this analysis. Greater phase changes are associated with a narrower filter. Therefore, at these frequencies, narrower bandwidths in the modeled filters would produce more accurate peak scalar predictions. However, at the lower measured frequencies of 1000 and 2000 Hz, somewhat broader filters might produce better estimates, resulting in a lower peak scalar and a slower phase change across the passband. This may be especially true for the 1000-Hz signal because, referring to Figure 6, when $b = 1.5$, the filter analysis shows no change in crest factor with changes in scalar. At that frequency, it is possible that the gammachirp filter is narrow enough to resolve the 100-Hz harmonics, thus making within-channel phase interactions impossible. In that case, a broader filter description would be indicated.

Prediction of maximum–minimum threshold

Like many filter constructs, the gammachirp filter is described by a function in which the phase response of the filter is changed with changing filter bandwidth. Filters with shallow skirts impose a slower phase change than filters with steep skirts. The filter applied here has a wider passband and shallower skirts at high frequencies than at low frequencies. These two properties of the filtering process interact to produce the masking effects in Figure 3. Thus, it is expected that as signal frequency is increased, greater masking differences with changes in scalar should be a result of widening of auditory filter bandwidths. This expectation is consistent with the data of Summers and Leek (1998) who showed that, at 4000 Hz, the difference in masking

between the positive and negative Schroeder-phase waveforms was about 15 dB, whereas, at 1000 Hz, the difference in masking was around 8 dB. Second, because increasing filter center frequency is associated with a slowing of the phase change across a filter's passband, it is also expected that the scalar at the masking peak is lower at high frequencies. These qualitative predictions are consistent with the data of Figure 3.

The crest factor analysis can also be used to determine whether these predictions agree with masking differences seen across frequency. It is assumed that greater crest factors lead to less effective masking and, therefore, greater psychophysically obtained thresholds. Because it is assumed that this relationship is monotonic, a comparison between the threshold masking differences and the maximum and minimum crest factors in the simulations can be made. To compare the modeled and obtained masking differences, the obtained threshold (or modeled crest factor) at $C = -1.0$ was subtracted from the maximum threshold (or maximum crest factor). (For simplicity, the threshold for the $C = -1.0$ masker will be taken as the minimum, although in some isolated cases one of the other scalars produced slightly more masking.) Modeling results are presented using $b = 1.5$, as a similar pattern of results emerged for the narrower filter bandwidths. The top panel of Figure 8 plots the measured differences (maximum − minimum threshold) as the solid bars, and the bottom panel of Figure 8 plots the predicted differences in dB [maximum crest factor in dB − minimum crest factor in dB; 10 log(crest factor)] as the hatched bars.

Absolute numbers have little meaning here as the scales are different and not necessarily linearly related, but common trends across frequency are of interest. For the experimental data, the difference between the maximum and the minimum threshold increases with increasing signal frequency (filled bars), except the value at 4000 Hz which is lower than the threshold differences at 2000 and 3000 Hz. The maximum − minimum threshold difference at 4000 Hz is 8 dB greater than the difference at 1000 Hz. These results are similar to those of Summers and Leek (1998), who reported masking differences between positive and negative Schroeder-phase maskers to be 7 dB greater at 4000 Hz than at 1000 Hz, although they used a 60 dB SPL signal tone rather than the 40 dB tone used here. Even though the increasing filter bandwidth would lead to the expectation of a larger maximum − minimum value at 4000 Hz than at 3000 Hz, this was not observed in the threshold masking data. The difference actually gets smaller rather than larger. However, given the lower overall masking levels for the 4000-Hz signal, shown in Figure 3, the modeled filter is not actually broader than that for 3000 Hz
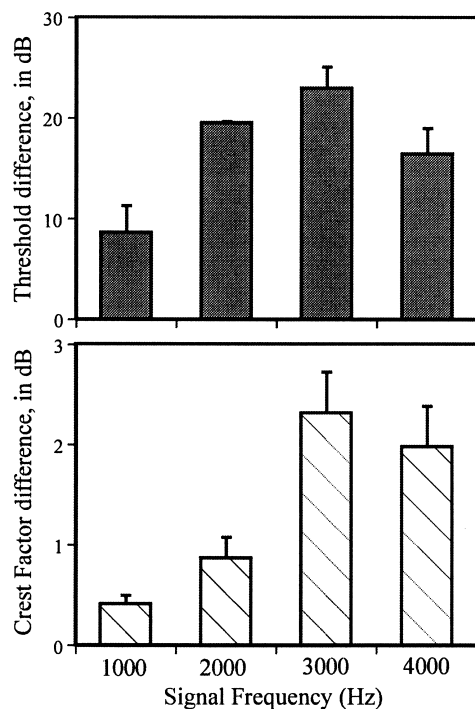
FIG. 8. The maximum threshold–minimum threshold is plotted for different signal frequencies as the solid bars in the top panel. Error bars denote standard errors across four observers. The maximum crest factor (dB) − minimum crest factor (dB) is plotted for different filter center frequencies as the hatched bars in the bottom panel.

(where overall masker levels are higher), which may account for the reduction in threshold difference both in the data and the predictions. This is a reminder that filter bandwidths likely reflect an interaction of both frequency and level.

SUMMARY AND CONCLUSIONS

This experiment measured masking produced by equal-amplitude harmonic complexes with systematically changing phase spectra. Replicating the results of Kohlrausch and Sander (1995), these data suggest that the curvature of different regions of the basilar membrane decreases with increasing frequency. Thus, the internal representation of stimuli at higher frequencies is better preserved with respect to the physical waveform.

The scalar providing the least masking at a given frequency may produce the most highly modulated internal waveform. The gammachirp filter was able to model qualitatively the changes in cochlear curvature with increasing frequency. However, this model (as implemented) does not produce appropriate changes in phase across its passband to account for the psychophysical results at all frequencies tested in this masking experiment.

REFERENCES

ALCANTARA JI, MOORE BCJ. The identification of vowel-like harmonic complexes: Effects of component phase, level, and fundamental frequency. J. Acoust. Soc. Am. 97:3813–3824, 1995.

DE BOER E, DE JONGH HR. On cochlear encoding: Potentialities and limitations of the reverse-correlation technique. J. Acoust. Soc. Am. 63:115–135, 1978.

BUUS S. Release from masking caused by envelope fluctuations. J. Acoust. Soc. Am. 78:1958–1965, 1985.

CARLYON RP. Spread of excitation produced by maskers with damped and ramped envelopes. J. Acoust. Soc. Am. 99:3647–3655, 1996.

CARLYON RP, DATTA AJ. Masking period patterns of Schroeder-phase complexes: Effects of level, number of components, and phase of flanking components. J. Acoust. Soc. Am. 101:3648–3657, 1997.

DAU T, WEGNER O, MELLERT V, KOLLMEIER B. Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. J. Acoust. Soc. Am. 107:1530–1540, 2000.

GLASBERG BR, MOORE BCJ. Derivation of auditory filter shapes from notched-noise data. Hear. Res. 47:103–138, 1990.

HARTMANN WM, PUMPLIN J. Noise power fluctuations and the masking of sine signals. J. Acoust. Soc. Am. 83:2277–2289, 1988.

IRINO T, PATTERSON RD. A time-domain, level-dependent auditory filter: The gammachirp. J. Acoust. Soc. Am. 101:412–419, 1997.

KIDD G JR., MASON CR, BRANTLEY MA, OWEN GA. Roving-level tone-in-noise detection. J. Acoust. Soc. Am. 86:1310–1317, 1989.

KOHLRAUSCH A. Masking patterns of harmonic complex tone maskers and the role of the inner ear transfer function. In: Duifhuis H, Horst JW, Wit HP, (eds.) Basic Issues in Hearing. Academic Press London, 1988, 339–350.

KOHLRAUSCH A, SANDER A. Phase effects in masking related to dispersion in the inner ear II. Masking period patterns of short targets. J. Acoust. Soc. Am. 97:1817–1829, 1995.

LEEK MR, PATTERSON RD, SUMMERS V. Phase discrimination and stimulus level in normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 99:2517A, 1996.

LEEK MR, SUMMERS V. Auditory filter shapes of normal-hearing and hearing-impaired listeners in continuous broadband noise. J. Acoust. Soc. Am. 93:3127–3137, 1993a.

LEEK MR, SUMMERS V. The effect of temporal waveform shape on spectral discrimination by normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 94:2074–2082, 1993b.

LENTZ JJ, RICHARDS VM, MATIASEK MR. Different auditory filter bandwidth estimates based on profile-analysis, notched-noise and hybrid tasks. J. Acoust. Soc. Am. 106:2779–2792, 1999.

Levitt H. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49:467–477, 1971.

Lutfi RA, Patterson RD. On the growth of masking asymmetry with stimulus intensity. J. Acoust. Soc. Am. 76:739–745, 1984.

Mehrgardt S, Schroeder MR. Monaural phase effects in masking with multicomponent signals. In: Klinke R and Hartmann R, (eds) Hearing—Physiological Bases and Psychophysics. Springer-Verlag Stuttgart, 1983, pp. 289–295.

Micheyl C, Maison S, Carlyon RP, Andéol G, Collet L. Contralateral suppression of transiently evoked otoacoustic emissions by harmonic complex tones in humans. J. Acoust. Soc. Am. 105:293–305, 1999.

Moore BCJ, Glasberg BR, Donaldson E, McPherson T, Plack CJ. Detection of temporal gaps in sinusoids by normally hearing and hearing-impaired subjects. J. Acoust. Soc. Am. 85:1266–1275, 1989.

Moore BCJ, Peters RW, Glasberg BR. Auditory filter shapes at low center frequencies. J. Acoust. Soc. Am. 88:132–140, 1990.

Patterson RD, Allerhand MH, Giguère C. Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform. J. Acoust. Soc. Am. 98:1890–1894, 1995.

Patterson RD, Moore BCJ. Auditory filters and excitation patterns as representations of frequency resolution. In: Moore BCJ, (ed.) Frequency Selectivity in Hearing. Academic Press London, 1986, 123–177.

Patterson RD, Nimmo-Smith I, Holdsworth J, Rice P. (1987) An efficient auditory filterbank based on the gammatone function. Paper presented at a meeting of the IOC Speech Group on Auditory Modelling at RSRE, December 14–15, 1987, Malvern, United Kingdom.

Plomp R. Rate of decay of auditory sensation. J. Acoust. Soc. Am. 36:277–282, 1964.

Richards VM. Components of monaural envelope correlation perception. Hear. Res. 35:47–58, 1988.

Rosen S, Baker RJ. Characterising auditory filter nonlinearity. Hear. Res. 73:231–243, 1994.

Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L. Basilar-membrane responses to tones at the base of the chinchilla cochlea. J. Acoust. Soc. Am. 101:2151–2163, 1997.

Schreiner C, Urbas JV, Mehrgardt S. (1983) Temporal resolution of amplitude modulation and complex signals in the auditory cortex of the cat. In: Klinke R and Hartmann R, (eds) Hearing—Physiological Bases and Psychophysics. Springer-Verlag Stuttgart, pp 169–175.

Schroeder MR. Synthesis of low peak-factor signals and binary sequences with low autocorrelation. IEEE Trans Inf. Theory IT 16:85–89, 1970.

Smith BK, Sieben UK, Kohlrausch A, Schroeder MR. Phase effects in masking related to dispersion in the inner ear. J. Acoust. Soc. Am. 80:1631–1637, 1986.

Strube HW. A computationally efficient basilar-membrane model. Acustica 58:207–214, 1985.

Summers V, Leek MR. Masking of tones and speech by Schroeder-phase harmonic complexes in normally hearing and hearing-impaired listeners. Hear. Res. 118:139–150, 1998.

Summers V. Effects of hearing impairment and presentation level on masking period patterns for Schroeder-phase harmonic complexes. J. Acoust. Soc. Am. 108:2307–2317, 2000.

Viemeister NF. Temporal modulation transfer functions based upon modulation thresholds. J. Acoust. Soc. Am. 66:1364–1380, 1979.

Viemeister NF, Plack CJ. Human Psychophysics. In: Yost WA, Popper AN, and Fay RR. (eds) Time analysis. Springer-Verlag New York, 1993, 116–154.

Zwicker E. Masking period patterns of harmonic complex tones. J. Acoust. Soc. Am. 60:429–439, 1976.