

Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets

Citation for published version (APA):

Kohlrausch, A. G., & Sander, A. (1995). Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets. *Journal of the Acoustical Society of America*, 97(3), 1817-1829. <https://doi.org/10.1121/1.413097>

DOI:

[10.1121/1.413097](https://doi.org/10.1121/1.413097)

Document status and date:

Published: 01/01/1995

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Phase effects in masking related to dispersion in the inner ear.

II. Masking period patterns of short targets^{a)}

Armin Kohlrausch

Institute for Perception Research (IPO), P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Andres Sander^{b)}

Drittes Physikalisches Institut, Universität Göttingen, Bürgerstr. 42-44, D-37073 Göttingen, Germany

(Received 7 October 1992; revised 15 August 1994; accepted 7 November 1994)

This article investigates how the amplitude and phase characteristics of the inner ear influence the spectrotemporal representation of harmonic complex sounds. Five experiments are reported, in each of which three sets of maskers are compared that differ only in their phase spectra. The amplitude spectra of the complexes were flat and the phase choices were (a) zero phase, (b) Schroeder phases with a positive sign, and (c) Schroeder phases with a negative sign. In the first four experiments, the spectra contained all harmonics between 200 and 2000 Hz. In experiments 1 and 2, the signal frequency was fixed at 1100 Hz and the fundamental frequency of the maskers was varied. In experiments 3 and 4, the fundamental frequency of the maskers was fixed and the signal frequency varied between 200 and 2000 Hz. In experiments 1 and 3, the signal duration was long compared to the period of the maskers. In experiments 2 and 4, the signal duration was only 5 ms and thresholds were determined for different time points within the masker's period. The results show a strong correlation between the minima of the short signal's thresholds and the threshold of the long signal. In experiment 5, the spectral extent of the masker was shifted to values one octave lower (100 to 1000 Hz) or one or two octaves higher (400 to 4000 Hz and 800 to 8000 Hz, respectively). For each spectral region, masked thresholds of a long signal were obtained for three values of the fundamental frequency. In all five experiments the thresholds depended very much on the specific phase choices with differences of up to 25 dB. The masker with a negative Schroeder phase always led to the highest thresholds. The thresholds of the masker with a positive Schroeder phase, on the other hand, were for a wide range of parameters *lower* than the thresholds for the zero-phase masker. These phase effects are most likely caused by the phase characteristic of the basilar-membrane filter, which affects the flat envelopes of the two Schroeder-phase maskers in a very different way. For an appropriate choice of parameters, one of the two becomes even more strongly modulated than the zero-phase complex. This latter observation imposes some restrictions on the second derivative (curvature) of the phase-versus-frequency relation for the auditory filters.

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Mk, 43.66.Nm [HSC]

INTRODUCTION

One of the intriguing qualities of the human hearing system is its ability to follow fast temporal variations in the acoustic input signals with high accuracy. This property is a prerequisite for higher perceptual processes based on the analysis of temporal cues within the signals. One example of these central processes is the separation or perceptual grouping of simultaneous sounds on the basis of differences in envelope modulation frequency (e.g., Bregman *et al.*, 1985). The measurement of temporal effects and the modeling of temporal resolution within the human hearing system is therefore one of the main fields of interest in current psychoacoustic research.

Temporal properties are deduced from experiments, in which a specific signal parameter varies in time. An example

of this approach, closely related to the procedure used in the present study, is the masking technique. In this paradigm, a modulated sound with a certain modulation depth is used as masker. The threshold for detection of a short-duration signal is measured as a function of its temporal position within the modulation period. This experimental approach is based on the assumption that the threshold of the signal reflects in a direct way the temporal course of the masker's internal excitation (Zwicker, 1976/77; Fastl, 1977; Rodenburg, 1977).

In the present study, we investigate masking properties of periodic complex-tone maskers in relation to phase dispersion in the inner ear. The envelope modulation of the maskers is varied by manipulating their phase spectrum. Thus these experiments are on the one hand related to the topic of "monaural phase effects" (e.g., Goldstein, 1967; Patterson, 1987). On the other hand, by manipulating the phase spectrum, we generate broadband stimuli with an amplitude or a frequency modulation, two classes of signals that are usually not directly compared in terms of their spectrotemporal excitation (e.g., Zwicker, 1974 for frequency-

^{a)}Some of the data have been presented previously at the 8th international symposium on hearing in Paterswolde, The Netherlands (Kohlrausch, 1988).

^{b)}Current address: Carl-von-Ossietzky Universität Oldenburg, Graduiertenkolleg Psychoakustik, Postfach 2503, D-26129 Oldenburg, Germany.

modulated maskers and Zwicker, 1976/77 for amplitude-modulated maskers).

These experiments continue earlier work on complex-tone masking, where we compared the masking of continuous targets produced by pairs of maskers differing only in their phase spectra (Smith *et al.*, 1986). Although the envelopes of both maskers were flat and identical, differences of up to 20 dB were obtained in the masked thresholds. Through simulations with a digital, linear basilar-membrane filter we could demonstrate that differences in masked thresholds were largest for those maskers which had the largest difference in modulation depth at the output of the basilar-membrane filter. If low-masked thresholds of a stationary signal are produced by pronounced valleys in the envelope of the masker, the basilar-membrane approach can explain most of the effects observed in this previous investigation.

In the present study, this way of reasoning is tested with the technique of masking period patterns (MPPs). A signal pulse with a duration shorter than the period of the masker is placed at different time points within the period of the masker. If the previous explanation is correct we should find a masking pattern of the short tone pulse which closely resembles the envelope fluctuations of the basilar-membrane filtered masker. The signal levels at the minima in these masking patterns should be highly correlated with the thresholds obtained for stationary signals. For a zero-phase complex (or periodic pulse), such a correlation had been found by Duifhuis (1971).

In addition to the two Schroeder-phase maskers used in the first study, a sine-phase masker with all components in zero phase was included in all experimental conditions. As this masker has a constant phase value at the input to the ear, the peak factor of the input time signal—and thus the modulation depth—is maximal. Due to its simple phase spectrum, the results for a zero-phase masker should allow a better understanding of the role of the phase characteristic of the inner ear for the internal representation of complex sounds.

In a final experiment phase effects are compared in four spectral regions. Measurements are performed at signal frequencies of 550, 1100, 2200, and 4400 Hz. At each signal frequency three different fundamental frequencies for the harmonic maskers are used. The spectral parameters of the maskers at the various signal frequencies are chosen to test the hypothesis that (according to a first approximation) the amplitude and the phase characteristic of the auditory filters is invariable by shifting on a logarithmic frequency scale.

I. METHOD

A. Stimuli

All maskers used in the experiments consisted of a sum of equal-amplitude sinusoids of common fundamental frequency f_0 . In experiments 1 to 4, the spectrum included all harmonics between 200 and 2000 Hz. In experiment 5, the signal frequency and the masker bandwidth were chosen from four different spectral regions which are listed in Table I. The actual number of harmonics in a complex depended on the value of f_0 :

TABLE I. Relation between the signal frequency and the spectral range and the fundamental frequencies of the maskers.

Signal frequency (Hz)	Masker spectral range (Hz)	Fundamental f_0 (Hz)		
550	100–1000	12.5	25	50
1100	200–2000	25	50	100
2200	400–4000	50	100	200
4400	800–8000	100	200	400

$$m(t) = \sum_{n=n_1}^{n_2} A_0 \sin(2\pi n f_0 t + \theta_n). \quad (1)$$

Three different values for the starting phases θ_n were used: A constant value of zero for all θ_n , for which the time function resembles a pulse sequence with an interpulse interval given by the inverse of the fundamental frequency. This signal will be named the “sine-phase” complex or m_0 complex.

The two other phase choices are based on a formula proposed by Schroeder (1970):

$$\theta_n = -\pi n(n-1)/N, \quad (2)$$

where $N = n_2 - n_1 + 1$ gives the total number of components in the complex. This phase choice reduces the peak factor of the time function and leads to a relatively flat temporal envelope. The formula can be used both with a “-” sign before the fraction [as in Eq. (2)] and with a “+” sign. The signals generated with θ_n values given by Eq. (2) will be called m_- signals, while the term m_+ signals will be used, if the “+” sign is used.

The three time functions for a complex with fundamental frequency 100 Hz and the different phase choices are shown in Fig. 1. The upper panel shows the sine-phase complex, the middle panel a Schroeder-phase complex with a “-” sign as in Eq. (2) (m_-) and the bottom panel shows a Schroeder-phase complex with a “+” sign (m_+). An important feature of the Schroeder-phase signals is the slope of the instantaneous frequency within each period, which is linearly increasing for the m_- complex and linearly decreasing for the m_+ complex. This can be seen by comparing the “wavelength” of the fine structure at the beginning and the end of each period.

In the time function of the sine-phase signal, a low- and a high-frequency ripple is apparent between the peaks. These oscillations correspond to the lower- and upper-edge frequencies of the complex and can be perceived as separate pitches superimposed on the 100-Hz virtual pitch of the complex (Kohler and Houtsma, 1991,1992; see also Moore and Glasberg, 1989).

The spectrotemporal properties of the three signals can be analyzed more clearly in a short-time-spectrum representation. Figure 2 shows the spectra of the three signals calculated using a moving 5-ms Hanning window. Both the edge frequencies of the sine-phase complex and the sawtoothlike frequency modulation of the two Schroeder-phase complexes are pronounced in this representation.¹

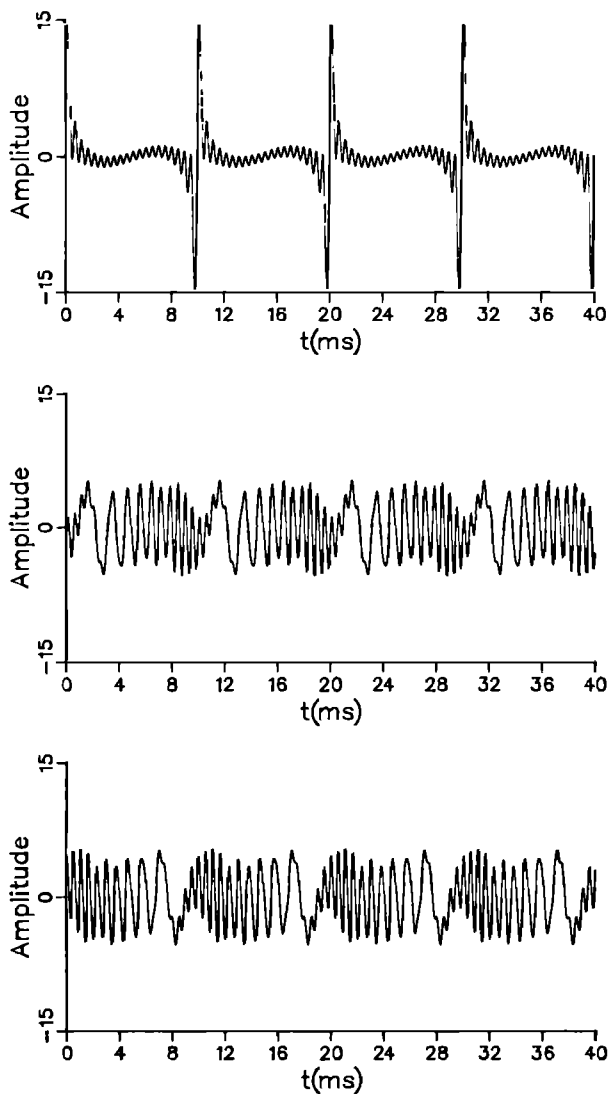


FIG. 1. Time functions of harmonic complexes for three different choices of the component's starting phases. Top: $\theta_n=0$ (sine-phase or m_0 complex), middle: $\theta_n=-\pi n(n-1)/N$ (m_- Schroeder-phase complex), bottom: $\theta_n=+\pi n(n-1)/N$ (m_+ Schroeder-phase complex). All complexes are composed of the equal-amplitude harmonics number 2 to 20 of fundamental frequency 100 Hz. For this plot, the amplitude of an individual harmonic was set at 1.

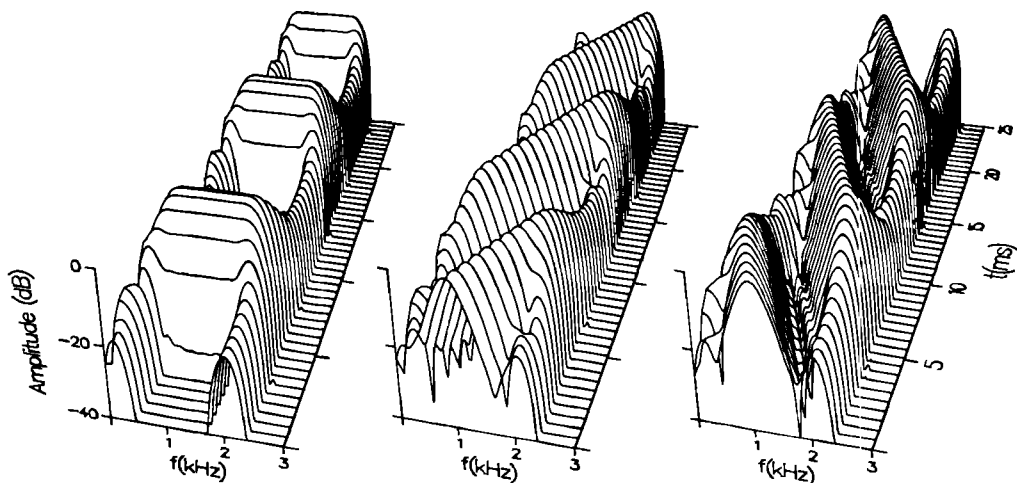


FIG. 2. Short-time spectra using 5-ms Hanning windows for the three signals from Fig. 1. Left: sine-phase complex, middle: m_- complex, right: m_+ complex.

In experiments 1 to 4, the masker level was 75 dB SPL. The level of individual harmonics therefore varied between approximately 66.5 dB SPL for a complex with only 7 components ($f_0=275$ Hz) and 55 dB SPL for a complex with 91 components ($f_0=20$ Hz). In experiment 5, the masker level was 70 dB SPL. This has to be considered when analyzing the experimental data in the following figures, as the masked thresholds of the sinusoidal signal are always expressed relative to the level of a single masker component.

The signal frequency always corresponded to the frequency of a specific masker component. In experiments 1 and 2, the center frequency of the signal was 1100 Hz. In experiments 3 and 4, it was varied in the range 200 to 2000 Hz. In experiment 5, different spectral regions for signal and masker were chosen and the signal was spectrally centered in the masker. The duration of the signal in experiments 1, 3, and 5 was 260 ms and it was gated with raised-cosine ramps of 30 ms (10 ms in experiment 5). In experiments 2 and 4, the signal was a 5-ms Hanning pulse consisting of two 2.5-ms raised-cosine ramps. The long signal was temporally centered in the masker; the short signal was added to the masker with a variable onset time difference t_v . The signal was always added in phase to the corresponding masker component.

B. Threshold estimation procedure

Masked thresholds were measured with an adaptive three-interval forced-choice procedure. The three masker intervals had a duration of 320 ms each and were separated by breaks of 200 ms. In one randomly chosen interval, the sinusoidal signal was added to the masker. The subject had to specify the interval containing the signal. The signal level was varied adaptively following a two-down one-up rule with a step size of 1 dB (Levitt, 1971). During the beginning of each measurement, a larger step size of 5 or 10 dB was used. This was progressively reduced by a factor of 2 after each second turnaround point of the signal level. This initial part of the measurement ended when the step size reached 1 dB. The median value of the levels within the following 20

trials was then taken as a single threshold value (Kollmeier *et al.*, 1988). During these 20 trials, typically 4 to 6 reversals of the signal level occurred.

In experiments 1, 2, 4, and 5, each subject performed four measurements per condition and the median was taken as the threshold value. In experiment 3 the individual data points are averages of two single measurements. If these two values differed by more than 3 dB, a third value was obtained and the two closer values were averaged. If subjects had difficulties reaching a stable threshold a larger number of measurements was performed and the final four values were used to calculate the median.

Results of experiments 1 and 3 are shown as averages of the two observers. For experiments 2, 4, and 5, median values across all observers are plotted. These medians were based on all single measurements (16 for the four observers of experiments 2 and 5, 12 for the three observers of experiment 4). For these data points, the interquartile values are plotted if they exceed 3 dB on at least one side.

Results for the three phase choices are represented in the following way: Data for the sine-phase masker are indicated by open squares. Data for the Schroeder-phase maskers are indicated by open triangles with the tip pointing either to the left or to the right. The direction of the tip is the same as the course of the instantaneous frequency, i.e., a tip pointing to the right is used for the m_- masker and a tip pointing to the left is used for the m_+ masker.

C. Apparatus and subjects

A Gould Concept 32/9705 computer generated all stimuli and recorded the responses of the subjects. The digital stimuli were converted into analog signals by means of a two-channel 12-bit D/A converter at a sampling rate of 20 kHz and were low-pass filtered at 5 kHz. In experiment 5, a 16-bit D/A converter was used. Its sample frequency was set to 30 kHz (and the low-pass filter to 10 kHz), if the audio signal contained components above 4 kHz. After an appropriate amplification, the signals were presented diotically to the listener over a headphone (Beyer DT880 monitor with diffuse field equalizer).

Seven subjects aged between 25 and 36 years participated in the experiments, but only one subject (the first author) performed all measurements. All listeners were members of the psychoacoustic and room acoustic research group at the Drittes Physikalisches Institut in Göttingen and had prior experience in psychoacoustic listening experiments.

II. RESULTS

A. Fixed signal frequency, variable fundamental frequency of the masker

In the first two experiments, the signal frequency was fixed at 1100 Hz. Thus in these measurements, the influence of temporal resolution within a fixed auditory filter was analyzed.

1. Experiment 1: Stationary signal

In the first experiment, the duration of the signal (260 ms) was long compared to the duration of a single period of

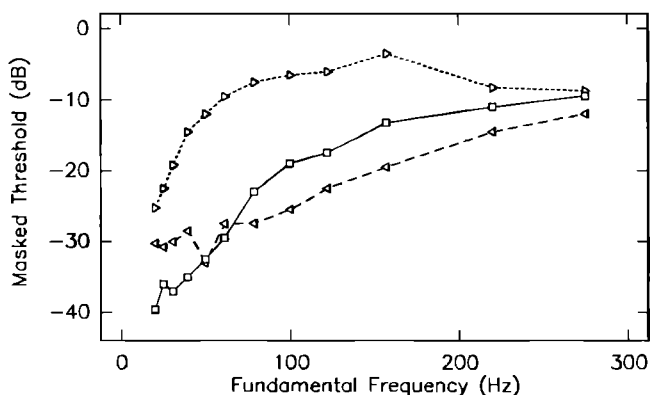


FIG. 3. Simultaneous masked thresholds of a 260-ms, 1100-Hz signal as a function of the fundamental frequency f_0 of the harmonic complex masker. Thresholds are expressed relative to the level of a single masker component. The maskers were presented at a level of 75 dB SPL. \square : m_0 masker; \triangleleft : m_+ masker; \triangleright : m_- masker. Average results of two listeners.

the masker. This experiment is closely related to experiment 1 in Smith *et al.* (1986) with the minor modification that, previously, the maskers were defined as a sum of *cosines* instead of as a sum of *sines*.

In Fig. 3, the average results for two subjects are presented. The abscissa denotes the fundamental frequency of the masker and the ordinate gives the masked threshold of the signal relative to the level of the 1100-Hz component of the masker. The results for the two Schroeder-phase signals (\triangleleft and \triangleright) agree well with the data in Smith *et al.* (1986). The differences between these two maskers can amount to more than 20 dB and they are maximal for fundamental frequencies between about 50 and 200 Hz.

A new aspect of these data is the relation between the sine-phase and the positive Schroeder-phase masker. Only at the lowest values of f_0 does the sine-phase masker lead to the lowest masked thresholds. For all fundamental frequencies above 60 Hz, however, the positive Schroeder-phase masker produces less masking than the sine-phase masker, with a maximum difference in masked thresholds of more than 6 dB.

2. Experiment 2: Short signal pulse

In the second experiment, thresholds of a 5-ms signal were measured as a function of its temporal position within the masker. Three different values for f_0 were used for the maskers: (1) A high value of 220 Hz, which is somewhat greater than the critical bandwidth at 1100 Hz, and for which there should be only a small interaction between adjacent masker components and phase effects should be small; (2) a medium value of 100 Hz, for which the results of experiment 1 indicate a large difference between the two Schroeder-phase maskers; and (3) a low value of 25 Hz, for which, in the previous experiment, the sine-phase masker led to the lowest thresholds, and the differences between the two Schroeder-phase maskers became small.

In Fig. 4, the results for the maskers with fundamental frequency 220 Hz are plotted. The abscissa describes the temporal position of the signal onset relative to the masker onset. For all three maskers, the masking patterns are only

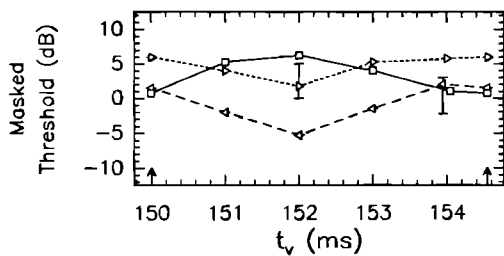


FIG. 4. Masked thresholds of a 5-ms signal (relative to the level of a single masker component) as a function of the onset time difference t_v between masker and signal. The masker's phase is indicated by the three symbols as in Fig. 3: m_0 (\square), m_+ (\triangleleft), and m_- (\triangleright). The arrows mark the duration of one period of the masker which has a fundamental frequency of 220 Hz. Data at the end of the masker's period are copies of those at the beginning of the masker's period. The data points are the medians of the 16 individual threshold values from the four observers. The interquartile ranges are indicated, if they exceed 3 dB on at least one side.

slightly modulated. The different positions of minima and maxima in the MPPs for the two Schroeder-phase maskers on the one hand and the sine-phase masker on the other hand follow from our definition of the starting phase for the masker components. This relation is described in detail in the Appendix. The modulation within the patterns for the same masker and the differences between the patterns for different masker phases indicate that there is some interaction between the 1100-Hz component and adjacent masker harmonics. The thresholds for the m_- masker (\triangleright), for instance, are always 4 to 7 dB higher than the values for the m_+ masker (\triangleleft). The relative position of the thresholds for the three maskers corresponds well with the results for the stationary signal (Fig. 3).

In Fig. 5, the corresponding results for the masker with fundamental frequency 100 Hz are presented. Whereas the pattern for the m_- masker is very flat for different values of t_v —the median value varies only by 5.5 dB—the pattern for the m_+ masker is highly modulated with a variation of 17 dB within the 10-ms period of the masker. The variation in the thresholds for the sine-phase masker is similar to that for the m_+ masker, but the pattern is shifted to values about 7 dB higher.

Finally, Fig. 6 shows the results for maskers with 25-Hz fundamental frequency. The patterns of all three maskers are highly modulated with a variation of 20 dB for the two

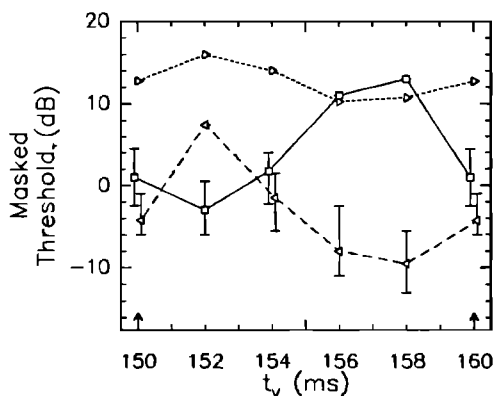


FIG. 5. Same as Fig. 4 for a fundamental frequency of 100 Hz.

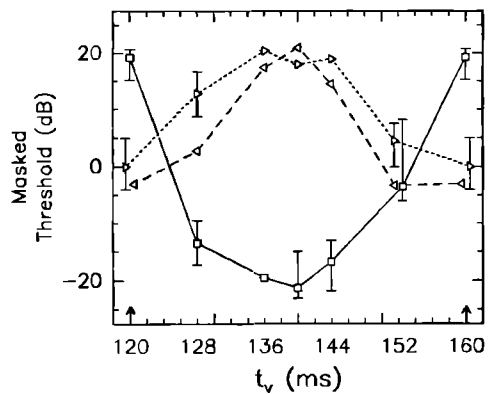


FIG. 6. Same as Fig. 4 for a fundamental frequency of 25 Hz.

Schroeder-phase maskers and 40 dB for the sine-phase masker.² The two Schroeder-phase maskers produce very similar masking patterns. The width of the minimum, however, differs between the two maskers and can be related to differences in the threshold values in Fig. 3. The wider valley in the MPP for the m_+ masker in Fig. 6 corresponds to a lower masked threshold for the stationary signal in Fig. 3. The sine-phase pattern, on the other hand, is much more modulated and closely resembles the envelope of the acoustic input signal, a series of short pulses with long periods of low envelope values in between.

B. Variable signal frequency, fixed fundamental frequency of the masker

In this second part of the study, the fundamental frequency of the maskers was fixed at 100 Hz and the signal frequency was varied over the whole spectral range of the maskers.

1. Experiment 3: Stationary signal

The temporal properties of this experiment were chosen to be parallel to those of experiment 1: The 260-ms signal was added to the temporal center of the 320-ms masker and was long compared to the period of the complex. Masked thresholds were determined at all 19 frequencies between 200 and 2000 Hz that were present in the masker spectrum.

The average results of the two listeners are shown in Fig. 7. At both spectral edges of the maskers (200 and 2000 Hz), the differences between the three phase choices disappear. However, in the spectral center, the phase values strongly influence the masked thresholds: The m_+ and m_0 curves pass through a minimum around 1000 Hz, whereas the m_- data increase below the upper edge, reach a maximum around 1600 Hz and decrease for lower frequencies. Comparing the results for the two Schroeder-phase maskers with the sine-phase masker, we can state the following: Below a signal frequency of about 1500 Hz the m_+ masker leads to threshold values lower than or equal to those for the sine-phase masker. The largest difference, of about 10 dB, is found around a signal frequency of 900 Hz. The m_- masker, on the other hand, generally leads to higher values than the sine-phase masker. The greatest difference between these two maskers, of about 20 dB, occurs around 1600 Hz.

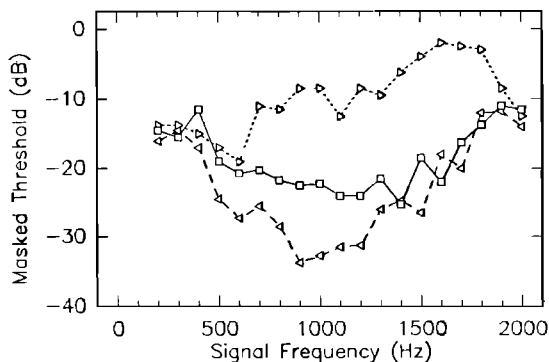


FIG. 7. Simultaneous masked thresholds of 260-ms signals in a periodic masker with fundamental frequency 100 Hz as a function of the signal frequency. All signals coincided in frequency and phase with one of the 19 masker components. The various symbols are used to indicate the three phase choices of the masker as in Fig. 3. Average results of two listeners.

2. Experiment 4: Short signal pulse

In this experiment, masking period patterns with a 5-ms signal were measured for the three maskers with fundamental frequency 100 Hz for the following signal frequencies: 200, 600, 1600, and 2000 Hz.³

In Fig. 8, the results for a signal frequency of 200 Hz are presented. The temporal variation within a period is extremely small and also the differences between the three curves are restricted to a few dB. This result agrees well with what would be expected from Fig. 7.

At a higher signal frequency of 600 Hz, the patterns are much more modulated (Fig. 9). The maxima for the two Schroeder-phase maskers no longer occur at the same signal delay. Following the considerations given in the Appendix, the maxima for the m_- masker (\triangleright) are expected at $t_v=150$ ms and those for m_+ (\triangleleft) at $t_v=155$ ms. The vertical arrangement of the three curves matches the average results for the stationary signal in Fig. 7 with m_- leading to the highest and m_+ leading to the lowest value.

In Fig. 10, the data for the 1600-Hz signal are presented. Here, the m_+ masker (\triangleleft) produces a broad minimum between $t_v=154$ and 158 ms, which is about 2 dB lower than the minimum of the sine-phase masker. The pattern of the m_- masker is very flat over the period of the masker, as it is expected from experiment 3.

The last set of data was obtained for a signal placed at

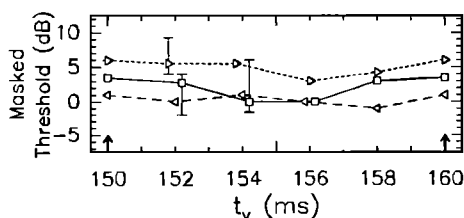


FIG. 8. Masked thresholds of a 5-ms signal as a function of the onset time difference t_v between masker and signal. The masker's phase is indicated by the same three symbols as in Fig. 3. The arrows mark the duration of one period of the masker which has a fundamental frequency of 100 Hz. Data are medians across 12 individual threshold values from three observers. The interquartile ranges are indicated, if they exceed 3 dB on at least one side. The signal frequency is 200 Hz.

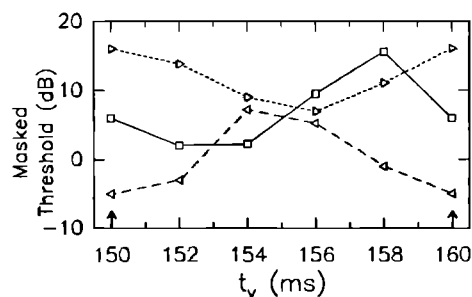


FIG. 9. As Fig. 8 for a signal frequency of 600 Hz.

the upper spectral edge at 2000 Hz (Fig. 11). Here, m_- and m_+ maskers produce nearly identical and flat masking patterns. The sine-phase values differ only in showing a relative maximum at delays 156 and 158 ms, but the value of the minimum is the same as for the two other maskers.

C. Experiment 5: Variable spectral range of masker and signal

The question for the final experiment was whether the previously observed phase effects at 1100 Hz are related in a systematic way to phase effects at other frequencies. This was investigated by measuring masked thresholds at four signal frequencies of 550, 1100, 2200, and 4400 Hz. The spectral parameters at the signal frequency 1100 Hz (line two in Table I) agreed with those used in experiment 1. The parameters at the other signal frequencies were chosen to test whether the amplitude and phase characteristics of the inner-ear filters are invariable on a logarithmic frequency scale. In this case the temporal envelope of a specific masker, transformed into the various frequency regions according to Table I and filtered at the corresponding signal frequency, always has the same degree of amplitude modulation and the relative arrangement of thresholds for the three phase choices should remain the same in all spectral regions. The spectral composition of the maskers was derived from the 1100-Hz condition by scaling each frequency including f_0 down by the factor 2 (for the signal frequency of 550 Hz, line 1 in Table I) or enlarging it by the factor 2 or 4 (signal frequencies 2200 and 4400 Hz, line 3 and 4 in Table I). With this spectral transformation, the number of masker components is the same in all spectral regions for corresponding values of f_0 (low, middle, high).

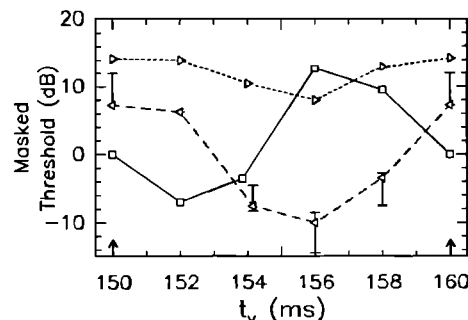


FIG. 10. As Fig. 8 for a signal frequency of 1600 Hz.

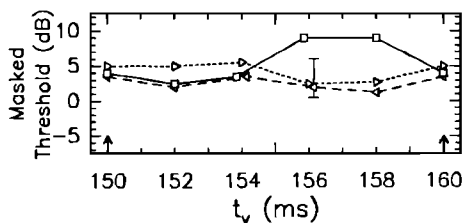


FIG. 11. As Fig. 8 for a signal frequency of 2000 Hz.

The average results of the four listeners are shown in Fig. 12 for the four signal frequencies (panel at the top left: 550 Hz, top right: 1100 Hz, bottom left: 2200 Hz, and bottom right: 4400 Hz). The three different phase choices are indicated by the same symbols as in the previous figures. The relative position of the data points is very similar at the two lowest signal frequencies of 550 and 1100 Hz (top panels). For the low value of f_0 , the m_0 masker leads to the lowest threshold, and for the high value, the m_+ masker, respectively. The m_- masker always leads to the highest threshold. Many of these relations are also found for signal frequencies of 2200 and 4400 Hz (bottom panels). The only exception is the relative position of m_0 and m_+ thresholds at the two lowest fundamentals. With increasing signal frequency, thresholds for m_0 increase more than those for m_+ . Thus at 4400 Hz, the m_+ masker leads to the lowest thresholds at all three fundamental frequencies.

Before we proceed to discuss the experimental data, we will, in the following section, first summarize the experimental results and emphasize the important details.

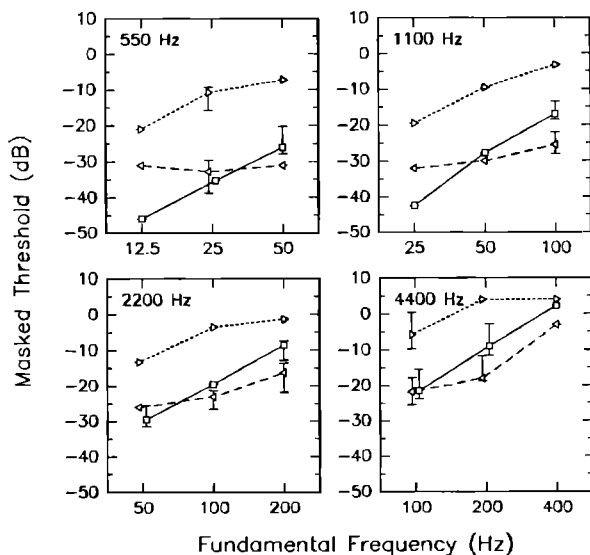


FIG. 12. Simultaneous masked thresholds of 260-ms signals as a function of the fundamental frequency of the harmonic complex masker. The signal was centered in the masker spectrum and has one of the following frequencies: 550 Hz (panel top left), 1100 Hz (top right), 2200 Hz (bottom left), and 4400 Hz (bottom right). The spectral range of the masker varied proportionally with the signal frequency. The maskers were presented at a sound pressure level of 70 dB. Data are medians across 16 individual threshold values from four observers. The interquartile ranges are indicated, if they exceed 3 dB on at least one side.

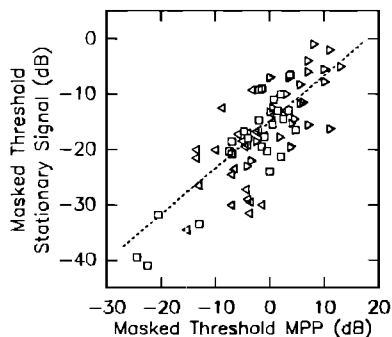


FIG. 13. Scatter plot showing the relation between minima of the masking period patterns (from experiments 2 and 4) and thresholds for stationary signals (from experiments 1 and 2) for eight combinations of signal frequency and masker fundamental frequency. The points are based on the individual threshold values of the subjects. The three symbols again indicate the three masker phase choices as in the prior figures. The linear regression curve (dotted curve) has a correlation value of 0.78.

III. SUMMARY OF EXPERIMENTAL RESULTS

(a) The masked thresholds for a given combination of fundamental frequency f_0 of the masker and frequency f_T of the signal depend very much on the phase of the masker's components. The difference between the thresholds for the three phase choices amounts to up to 25 dB and is observed both for stationary signals and in the minima of masking period patterns. The influence of masker phase disappears for signals placed at the spectral edges of the masker (cf. Figs. 7, 8, and 11) and for high fundamental frequencies above about 20% of the signal frequency (Figs. 3 and 4).

(b) There is a positive correlation between the thresholds for the long signal and the minima of the masking period patterns. This relation is illustrated in the following two figures. Figure 13 shows a scatter plot of the individual masked thresholds for the two signal durations. On the abscissa, values for the minima of the MPPs are shown. Values for stationary signals are shown on the ordinate. The three symbols indicate the three choices of the masker phase.⁴ The dotted line indicates the linear regression curve. It has a slope of 0.9 and the correlation value of the fit amounts to 0.78. An even closer correspondence is observed by analyzing the *differences* between the three phase choices for a fixed combination of f_0 and f_T . In Fig. 14, the differences between pairs of maskers ($m_+ - m_0$ shown by the triangles pointing to the left, $m_- - m_0$ shown by the triangles pointing to the right) are presented. The abscissa gives the differences at the minima in the MPPs and the ordinate gives the differences between the thresholds for stationary signals. The linear regression line is indicated by the dotted curve. It has a slope of 0.77 and the correlation value of this fit is 0.88. For comparison, the identity relation is indicated by the dashed line.

(c) With decreasing fundamental frequency the masked thresholds decrease. This result cannot be explained on the basis of the masker energy within a single critical band: For lower fundamental frequencies, the number of masker components within a critical band increases and so does the averaged masker energy relative to the level of an individual masker component. Rather than the averaged masker energy, it is the degree of modulation of the masker's envelope

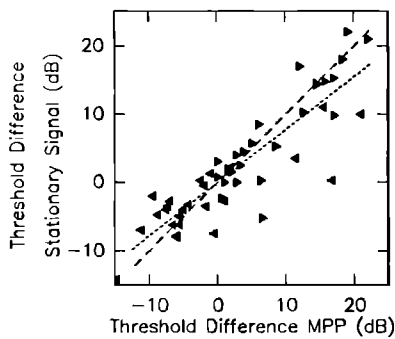


FIG. 14. Scatter plot similar to Fig. 13, but showing differences between thresholds on the two axes. The triangles pointing to the left indicate differences between m_- and m_0 , the triangles pointing to the right indicate differences between m_+ and m_0 . The linear regression curve (dotted curve) has a correlation value of 0.88. The dashed curve indicates the identity relation.

which determines the masked thresholds. The more components are summed with the appropriate phase relation, the higher the degree of modulation of the masker is and the deeper the minima in its envelope are.

(d) The slope of the steepest parts in the MPPs—for the m_0 as well as for the m_+ masker—increases with the signal frequency. For the 600-Hz signal, it amounts to about 5 dB/ms. The corresponding values for the 1100 and the 1600-Hz signals are 6.4 and 7.5 dB/ms. For a first-order bandpass filter, such slopes in the impulse response correspond to bandwidths of about 91, 117, and 137 Hz, respectively.

(e) The masking period patterns of the m_0 complex for different signal frequencies (Figs. 5, 8–11) show the following results: At both spectral edges, the patterns are relatively flat. For the three frequencies well within the spectrum of the masker, the minimum in the pattern becomes lower for higher frequencies. This result is shown in Fig. 15, where the MPPs for these three signal frequencies are combined. The 600-Hz pattern (filled triangles) is the shallowest one, whereas the 1600-Hz data (filled squares) reach the deepest minimum.

(f) The m_- and m_+ maskers have the property of a linearly rising and falling instantaneous frequency. The maxima in the MPPs of these complexes occur at those time points where the signal frequency coincides with the instantaneous frequency of the masker. This relation is demonstrated for

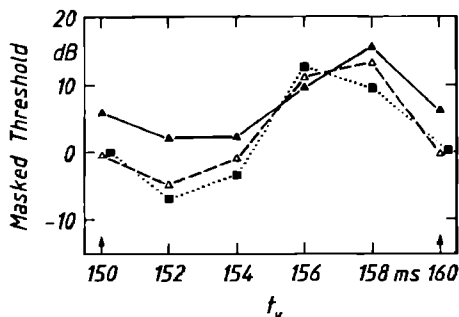


FIG. 15. Masking period patterns in a sine-phase masker of 100 Hz fundamental frequency for signal frequencies 600 Hz (filled triangles), 1100 Hz (open triangles), and 1600 Hz (filled squares).

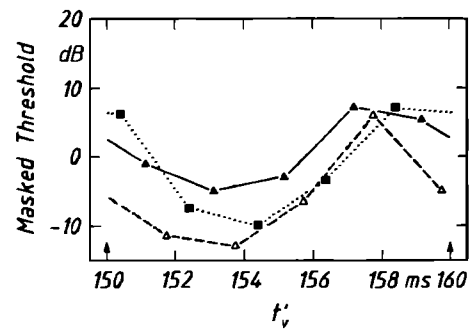


FIG. 16. Masking period patterns in an m_+ Schroeder-phase masker of 100-Hz fundamental frequency for signal frequencies 600 Hz (filled triangles), 1100 Hz (open triangles), and 1600 Hz (filled squares). The temporal position of the three curves was shifted to compensate for the frequency-dependent group delay in the m_+ masker. Shifts were 2.9 ms for 600 Hz, 5.5 ms for 1100 Hz, and 8.2 ms for 1600 Hz.

the highly modulated patterns of the m_+ complex ($f_T=600$, 1100, and 1600 Hz). If, on the basis of the frequency-dependent group delay within the complex (see Appendix), we calculate the appropriate temporal offset for each signal frequency, we can plot the MPPs on a new time scale (Fig. 16). Indeed, the maxima and minima now occur at the same time points within the masker period.

(g) In general the masker with a falling instantaneous frequency (m_+) produces lower masked thresholds than the masker with a rising instantaneous frequency (m_-). Thus we can conclude that internally the temporal excitation pattern of an m_+ complex is more modulated, i.e., the peaks in the envelope are higher and the valleys are lower than for the m_- masker.

(h) If masking conditions are transformed to other frequency regions by keeping the frequency ratios between masker fundamental, masker bandwidth and signal frequency constant, the relation between m_+ and m_- maskers remains the same: Thresholds of stationary signals in the spectral center of these maskers are always lower for the m_+ than for the m_- maskers. The relation between m_+ and m_0 maskers remains the same only for 550 and 1100 Hz. At higher signal frequencies, the range of fundamental frequencies for which m_+ thresholds are lower than those for m_0 is enlarged.

IV. DISCUSSION

A. Phase effects

The basic observation in the experiments was the strong influence of the masker's phase values on the masked thresholds. The different phase values lead of course to different time functions for the masker and it may be a question of taste to refer to these results as "monaural phase effects" or "time effects." As long as at least three sinusoidal components interact internally, i.e., excite the same auditory filter, changes in the component's phase can alter the time function at the output of this filter. Thus phase effects like differences in the just-noticeable amount of amplitude or frequency modulation can be taken as an indication of how strongly the spectral components interact internally and this interaction

can give some hints for the width and shape of the internal filter (cf. Zwicker, 1952; Goldstein, 1967; Sek and Moore, 1994).

It is, however, much more difficult to derive predictions about the *phase characteristic* of the internal filter from psychophysical data. For instance, in the study on monaural phase perception by Goldstein, it was explicitly assumed that “the unknown phase characteristic of the analyzing filter has a negligible effect upon the excitation envelope, in comparison with the effect of external AM/QFM phase transformation” (Goldstein, 1967, p. 468). As was pointed out later in the same paper, neglecting the internal phase is equivalent to neglecting a deviation from a constant plus linear phase, i.e., neglecting any curvature in the phase-versus-frequency characteristic. In the following it is shown that the present data do impose restrictions on the curvature of the internal phase of the auditory filter.

As an example, we will consider a condition with very strong differences for the three phase choices: $f_0=100$ Hz and $f_T=1100$ Hz. The flat envelope of the Schroeder-phase maskers is achieved by choosing phase values with a constant curvature (second derivative of the phase-versus-frequency characteristic). This phase choice results in a linear frequency modulation within each period (Schroeder, 1970). If a frequency-modulated signal is passed through a bandpass filter, the frequency modulation is transformed into an amplitude modulation. However, for a linear-phase filter this would not result in different degrees of modulation for these two maskers and the MPPs should be similar.

Different degrees of modulations for the two Schroeder-phase maskers are expected at the output of a filter for which the curvature in the phase characteristic has a constant sign within its passband. If the curvature has a similar, but opposite value as the phase curvature of a Schroeder-phase complex, the phase differences between adjacent components will be flattened out by the filter and the resulting waveform will be highly peaked. For the Schroeder-phase complex with an opposite sign, the curvature in the phase characteristic will be even increased by such a filter and the waveform after filtering will have a rather flat envelope.

The optimum match between stimulus phase and filter phase can be derived from the threshold values in Figs. 3 and 7 by looking for the maximum differences between m_0 and m_+ data. In Fig. 3, differences are maximal for f_0 values between 100 and 150 Hz. The second derivative of the phase-versus-frequency relation for the positive Schroeder-phase complex is independent of the frequency and amounts to

$$\frac{d^2 \phi}{df^2} = \frac{2\pi}{Nf_0^2}. \quad (3)$$

For complexes with a constant bandwidth as in experiments 1 to 4, the total number of components N is indirectly proportional to f_0 . Thus, the phase curvature for the Schroeder-phase complexes is indirectly proportional to f_0 . For f_0 values between 100 and 150 Hz, the curvature lies between $1.05 \times 10^{-5} \pi/\text{Hz}^2$ and $0.74 \times 10^{-5} \pi/\text{Hz}^2$. We can conclude that the curvature of the phase characteristic for the basilar-membrane filter centered at 1100 Hz should be in the

range of these two values, but with a negative sign.

Another interesting point in Fig. 3 is the value of f_0 for which m_+ and m_0 thresholds agree. Thresholds should agree, if the envelope after filtering has the same degree of modulation and this is realized if the phase at the output of the filter has a similar absolute value of the curvature for the m_0 and the m_+ complex. The m_0 complex does not have any curvature in the input phase. The filter curvature must thus have half the value of the curvature of the m_+ complex. In this case, it cancels *half* of the phase curvature of the m_+ complex and it introduces the same amount of phase curvature into the m_0 complex.

With the above inverse relation between phase curvature and f_0 it is easy to see that m_+ complexes with f_0 values between 50 and 75 Hz have a curvature which is just twice as large as the curvature of the filter at 1100 Hz derived earlier. And indeed, at these values of f_0 , the experimental data for m_0 and m_+ cross each other.

In Fig. 7, the lowest thresholds for m_+ complexes are reached for signal frequencies between 900 and 1200 Hz. Since the 100-Hz m_+ complex has a constant phase curvature at all frequencies, one can conclude that the phase curvature of the auditory filters in this range of frequencies matches the phase curvature of the m_+ complex sufficiently well. The fact that m_0 and m_+ thresholds approach each other at about 1500 Hz indicates that filters in this region have a phase curvature which is about half that of the m_+ complex and thus half that of filters centered on frequencies of 900 to 1200 Hz. Thus one can conclude from these data that the phase curvature of auditory filters decreases by a factor of 2 for an increase in center frequency by about half an octave.

The results of experiment 5 allow a comparison of the phase curvature of the auditory filters over a wider range of signal frequencies. The specific spectral scaling of the stimuli scales the curvature of the phases for the Schroeder-phase complexes in a simple way. If complexes with the same number of harmonics are compared (e.g., complexes having the middle value of f_0 in the various spectral regions), the curvature decreases by a factor of four for an increase of the spectral range by a factor of two [cf. Eq. (3)].

The middle value of f_0 (50 Hz for a signal frequency of 1100 Hz) is of particular interest, since for this parameter value m_0 and m_+ thresholds are about equal at the two lowest frequencies (top panels in Fig. 12). Using the arguments applied to the data at 1100 Hz, we might conclude that the curvature of the auditory filter at 550 Hz is about a factor of 4 larger than the curvature at 1100 Hz. Such a relation of the second derivative of the phase-versus-frequency relation is in line with the assumption that the phase-versus-frequency relation remains constant on a logarithmic frequency scale. This frequency dependence agrees with the above estimate for the changes in phase curvature in the frequency range 1000 to 1500 Hz.

For the two highest signal frequencies, m_0 and m_+ thresholds do no longer cross at the middle value of f_0 . By increasing the signal frequency from 1100 to 4400 Hz, the crossover value of f_0 decreases from the middle to the low value. If this change is caused only by the change of the

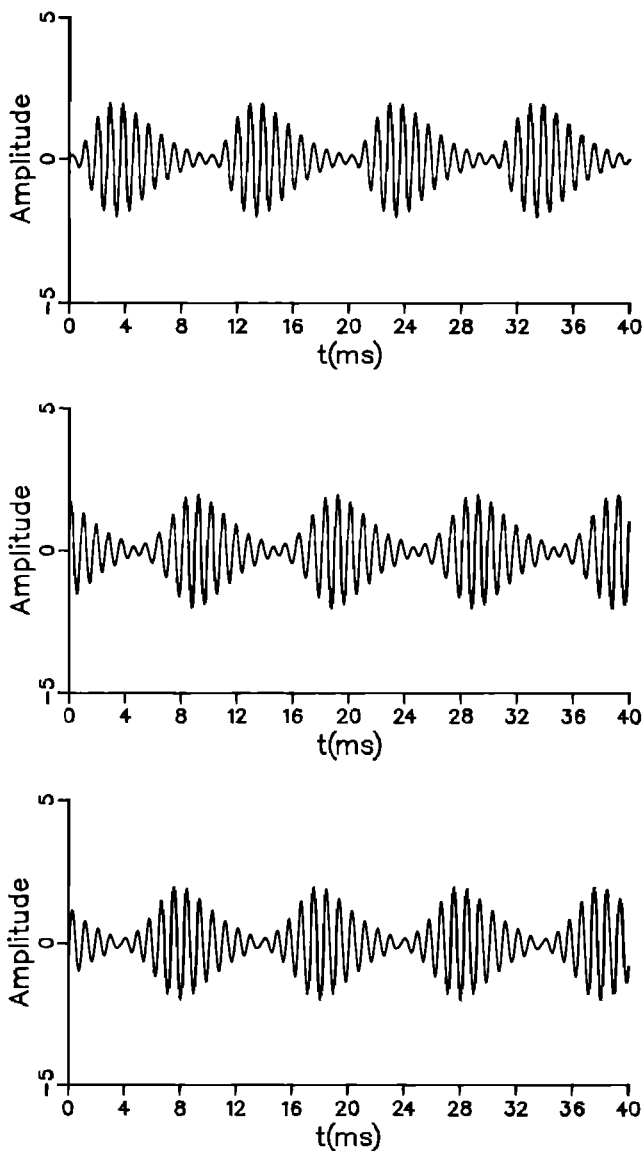


FIG. 17. Calculated waveforms for the three maskers with fundamental frequency 100 Hz using a fourth-order Gamma-tone filter centered at 1100 Hz. Top: m_0 masker, middle: m_- masker, bottom: m_+ masker. In the input signal, all masker components had the amplitude 1.

phase characteristic of the auditory filters with frequency, one has to conclude that between 1100 and 4400 Hz, the phase curvature of the filters decreases less than a factor of 16. This would imply that in this frequency range the phase-versus-frequency relation of the auditory filters is not constant on a logarithmic frequency scale.

B. Model calculations

In the following section we support the considerations about the filter phase by using two filters with different curvatures in their phase characteristics. The first filter is the popular rounded-exponential auditory filter (Patterson *et al.*, 1982). Using the implementation as a fourth-order Gamma-tone filter (Patterson *et al.*, 1987) with a center frequency of 1100 Hz, the filtered waveforms for the m_0 , the m_+ and the m_- maskers with $f_0=100$ Hz were calculated. The three panels in Fig. 17 each show four periods of the waveforms at the

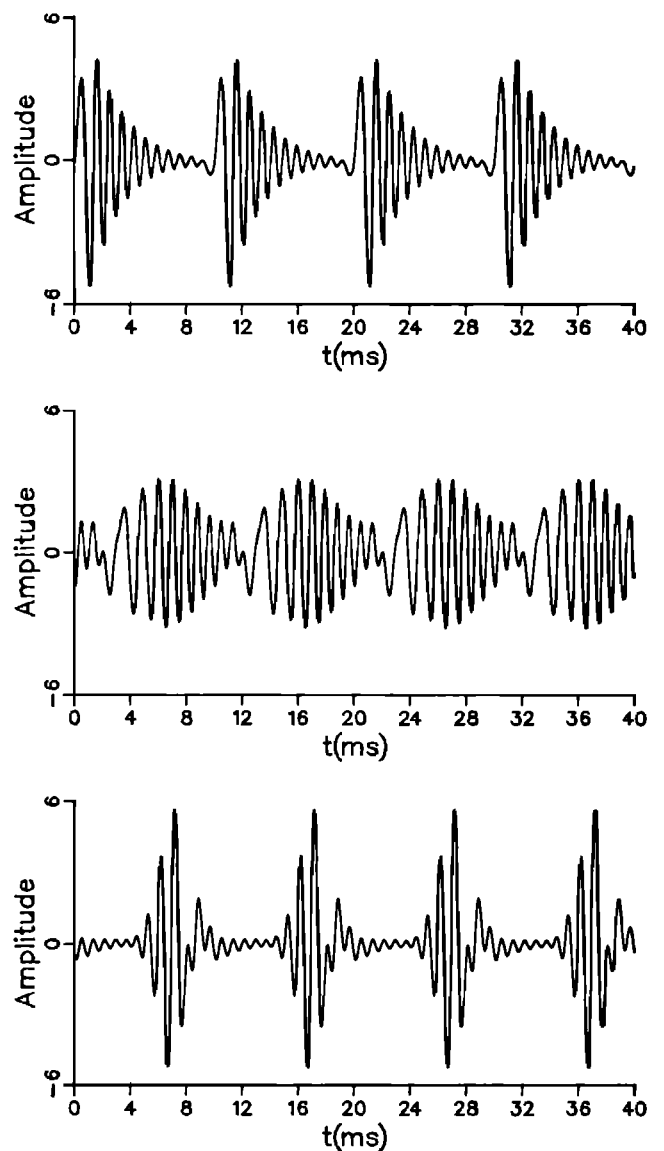


FIG. 18. Responses of a linear basilar-membrane model at resonance frequency 1100 Hz to the three harmonic maskers with fundamental frequency 100 Hz. Top: m_0 masker, middle: m_- masker, bottom: m_+ masker. In the input signal, all masker components had the amplitude 1.

output of the Gamma-tone filter. All three time functions have a similar degree of modulation and, especially, the two lower traces showing the waveforms for the two Schroeder-phase maskers are hardly distinguishable from each other. The major reason for this similar treatment of the two Schroeder-phase maskers by the Gamma-tone filter is its antisymmetric phase characteristic relative to its resonance frequency. The curvature of the filter phase changes its sign at the resonance frequency from negative to positive. A filter with such a phase characteristic can never flatten out the phase of a Schroeder-phase complex over the full range of its passband.

In contrast, Fig. 18 shows the waveforms for the same three complexes calculated with a basilar-membrane filter. For these simulations, we used a digital implementation of a linear, one-dimensional cochlear model consisting of 120 segments (Strube, 1985). This filter simulates only the passive properties of the basilar membrane and the frequency

selectivity is relatively poor compared to psychophysical estimates (cf. comments by Moore in Kohlrausch, 1988). This disadvantage is, however, more than compensated for by the phase characteristic of the model, which has a negative curvature over the range of its passband.⁵

As discussed earlier (Smith *et al.*, 1986) such a filter does show differences in the filtered waveforms for the two Schroeder-phase maskers for those values of f_0 and f_T for which the masked thresholds of stationary signals differ. With the result (b) of Sec. III, we can conclude that the same must be true for the minima in the MPPs. As can be seen by comparing the top and the bottom panel, the filter also shows differences between the m_0 and the m_+ masker, with the latter having a slightly wider minimum in its temporal waveform. Thus, at least qualitatively, the modulation of the filtered waveforms does also agree with the masked threshold differences between m_0 and m_+ maskers.⁶

This similarity between data and model simulation should, however, not be taken as a support that the basilar membrane behaves linearly. The only conclusion we can draw from the present study is that for the rather high masker level of 75 dB, the phase characteristic of the passive basilar-membrane model is suited to explain the observed phase effects. Since the amplitude and phase characteristics of the basilar membrane depend on the stimulus level (e.g., Ruggero *et al.*, 1992), one can expect that phase effects of the type described in the present paper vary with the masker level. In order to model such data, one certainly has to use a nonlinear model for the basilar membrane (e.g., Giguère and Woodland, 1994). Such modeling is, however, beyond the scope of this paper.

C. Relation to just-noticeable phase changes in harmonic complexes

Moore and Glasberg (1989) studied the sensitivity to phase changes of a single component in a broadband harmonic complex. The authors explained the results in terms of changes in the waveforms at the outputs of auditory filters introduced by the phase changes. Since the same explanation is used in the present paper for the value of the masked threshold, the two studies should give comparable results.

The phase change by an amount α of a sinusoidal component with amplitude 1 is mathematically equivalent to the addition of a component with amplitude $2 \sin(\alpha/2)$ and relative phase $90 + \alpha/2$ (and, of course, the same frequency as the original component). For psychoacoustic experiments, this consideration is only meaningful if the reference phase of the original component is somehow retained in the test interval. This is certainly the case if many harmonics in zero phase interact within a single auditory filter, since they sum up to a highly modulated waveform. In low-frequency filters, where, in the extreme case, the harmonics are perfectly resolved, the waveform (more accurately, the envelope) is not changed at all by the phase change. Thus the above formula does not apply to the low-order harmonics, but may well be used to relate phase changes for higher-order harmonics to masked thresholds.⁷

For a complex consisting of the first 20 harmonics of the fundamental 100 Hz in cosine phase the jnd was lowest for

harmonic numbers between 12 and 18 and increased for the two highest harmonics. For the lowest harmonics, no phase jnd could be determined at all. Using the above formulas, the phase jnd values transform to the following “masked thresholds”: The minimum values for harmonic numbers between 12 and 18 correspond to relative levels between -23 and -17 dB. The higher values at the upper spectral edge are equivalent to a relative level of about -6 dB.

Our data in Fig. 7 for the sine-phase complex also show the lowest thresholds for targets in the spectral center and an increase toward the upper edge. The range of thresholds (-25 to -15 dB) comes very close to the above estimate derived from the data of Moore and Glasberg. The increase toward the upper edge is somewhat smaller in our data. This might indicate that, especially at the upper edge, additional cues created by the interaction between edge pitch and target play a role.

Moore and Glasberg also performed measurements with random-phase complexes and could hardly measure a jnd. A random-phase complex does not have the strong modulation in the temporal envelope and is thus best comparable to the m_- data of our measurements. For this masker, thresholds were between -11 and -1 dB for harmonic numbers above 10. That is, in the worst case, the amplitude of the added component had to be of the same order as the amplitude of the masker component. This value comes very close to the theoretical level for a 90-deg shift ($+3$ dB). Thus the inability of subjects to detect a phase change in a random-phase complex is in line with findings from our masking experiments for a masker without a pronounced amplitude modulation.

D. Application of Schroeder-phase signals in speech research

The results of the present study are also relevant for speech manipulation techniques used to reduce the peak-to-rms ratio. Such techniques become interesting, if, e.g., the transmitter imposes some constraints on the maximal peak power. By enhancing the rms value without affecting the peak value, the loudness and the intelligibility of the speech waveform can be increased (Lynch, 1988; Quatieri and McAulay, 1988). One way to achieve a lower peak factor is, of course, to disperse the phase of the waveform. This leads exactly to the problem and the solution discussed by Schroeder (1970) and applied in the present study to generate the m_+ and the m_- complexes with a quadratic phase.

From a technical point of view, the quadratic phase—or, more generally, the phase dispersion—can be used with a positive as well as with a negative sign for the phase values. However, since the original publication by Schroeder used a negative sign, it seems as if most later studies implemented only this one of the two possible solutions. Quatieri and McAulay (1988), for instance, speak of a “chirp response with linearly increasing frequency” (p. 2559).

From the perceptual aspect, however, these two phase choices are not equivalent, since they lead to different amounts of the *internal* crest factor of the waveforms. Therefore, they might also lead to differences in, e.g., speech intelligibility. If listeners make use of temporal cues in the

speech signal, they might be more effective in listening to a speech signal with *decreasing* instantaneous frequency, since for this signal the internal crest factor is increased due to the phase dispersion in the inner ear.

ACKNOWLEDGMENTS

The experiments and most of the model simulations were performed while the authors were working at the Drittes Physikalisches Institut of the University of Göttingen in Germany. We would like to thank all colleagues who served as subjects in the experiments and the whole psychoacoustic research group in Göttingen for the many fruitful discussions concerning the topics of this paper. Technical support in generating the data and waveform plots was provided by M. v.d. Heijden and Ch. Ma in Eindhoven. For their helpful comments on earlier drafts of this paper, we wish to thank B.C.J. Moore, R. Patterson, D. Püschel, and J. Goldstein and two anonymous reviewers.

APPENDIX

In this Appendix, values are derived for the temporal positions of maxima and minima in masking period patterns for the maskers used in the experiments. The sine-phase maskers were always generated in such a way that at time instant zero the envelope of the summed waveform had a peak. This peak occurs after each integer multiple of the masker's period. Under the assumption that the maximal masked threshold is obtained for a signal pulse centered on the peak, the temporal position of the maximum in the masked thresholds will occur for those temporal positions where the signal *onset* precedes a peak by half the signal duration, in our case 2.5 ms. Those time points are $t_v = 152$ ms (for a fundamental frequency of 220 Hz), $t_v = 157.5$ ms (100 Hz) and $t_v = 117.5$ and 157.5 ms (25 Hz). Since the waveform peaks occur simultaneously over the whole bandwidth of the sine-phase masker, the above consideration is valid for all signal frequencies.

The maxima for the Schroeder-phase maskers are obtained for those signal positions at which the instantaneous frequency of the masker coincides with the signal frequency. For signal frequencies in the spectral center of the complex—as in experiment 2—this time point occurs approximately in the temporal center of the period. Compared to the maxima in the sine-phase complex, the maxima for the Schroeder-phase complexes are temporally shifted by half a period.

If the signal has a different frequency—as in experiment 4—the time point can be best calculated from the group delay of that frequency. The group delay of a system with a given phase characteristic $\phi(\omega)$ is defined as $-d\phi/d\omega$, where ω is the angular frequency. Using Eq. (2) from Sec. I, the group delay for a Schroeder-phase complex with fundamental frequency f_0 amounts at the frequency $f = n \cdot f_0$ to

$$\tau_g(f) = \pm(2n-1)/2Nf_0. \quad (4)$$

The group delay is positive for the m_- masker and negative for the m_+ masker.

For example, the instantaneous frequency of 600 Hz for the 100-Hz m_- complex is reached about 2.9 ms after the beginning of the period. The corresponding time point for the m_+ complex is about 7.1 ms. The three curves in Fig. 16 were plotted in such a way that the group delay of the corresponding signal frequencies was subtracted from the actual temporal position. The delays used were 2.9 ms ($f_T=600$ Hz), 5.5 ms ($f_T=1100$ Hz) and 8.2 ms ($f_T=1600$ Hz). After this temporal shift, the masker's instantaneous frequency and the signal frequency should coincide at 160 ms and the maxima are expected for $t_{v'} = 157.5$ ms.

¹In a preliminary report on some of these experiments (Kohlrusch, 1988), it was stated that the edge pitch for the sine-phase masker can be heard for fundamental frequencies up to about 200 Hz and that this pitch is not apparent for the Schroeder-phase masker. These statements were based on our own listening experience during the masking experiments. A systematic investigation of this phenomenon revealed that the edge pitch can also be perceived for the Schroeder-phase complexes. In addition we observed that the edge pitch can be perceived for fundamental frequencies up to about 300 Hz, if the upper-edge frequency is at least 3 to 4 kHz (Kohlrusch and Houtsma, 1991, 1992).

²During the measurements we observed that the results of the subjects varied substantially around the minimum of the sine-phase masker (t_v between 128 and 152 ms). Two of the listeners needed a longer practice period and performed approximately 20 single measurements for each of these data points, until a (relatively) stable threshold was reached. Nevertheless, at these values of t_v the averaged values for one subject remained significantly higher than for the other three subjects.

³Data were also obtained for a signal frequency of 1100 Hz. Since the median values of the three observers of this experiment agreed well with the median values of the four observers of experiment 2 for the same condition (cf. Fig. 5), these data are not shown.

⁴In order to generate these scatter plots, those subjects who originally participated only in the measurements with the short target performed afterwards the corresponding measurements for the long-duration target.

⁵For the basilar-membrane filter, both the amplitude and the phase characteristic are asymmetric around the resonance frequency. It is difficult to say how much each of these two properties contributes to the differences in the filtered waveforms of the m_+ and m_- maskers. We have some hints from own calculations (performed by Ch. Ma at the IPO in Eindhoven) that the asymmetric phase probably plays a more important role. If a Gamma-tone filter is combined with an allpass filter that has a phase characteristic similar to that of the basilar-membrane filter, the two Schroeder-phase maskers are transformed into signals with a different degree of modulation.

⁶Another example for differences in the filtered waveforms at the output of the Gamma-tone and the basilar-membrane filter can be found in Kohlrusch (1995).

⁷One remaining problem for a direct comparison between the phase jnd data and the masked threshold data lies in the relative phase of the "added component" which corresponds to the phase between the signal and the masker component in the masking experiments. This value was calculated to be $90 + \alpha/2$ for the phase jnd experiments, while it was zero throughout the present study. However, own unpublished data show that for a zero-phase complex of 100 Hz fundamental and a target frequency of 1100 Hz, masked thresholds vary by no more than 3 dB for various relative phase values of the signal. Thus the influence of the relative phase might be neglected for a first and somewhat rough comparison.

Bregman, A. S., Abranson, J., Doehring, P., and Darwin, C. J. (1985). "Spectral integration based on common amplitude modulation," *Percept. Psychophys.* **37**, 483–493.

Duifhuis, H. (1971). "Audibility of high harmonics in a periodic pulse. II. Time effect," *J. Acoust. Soc. Am.* **49**, 1155–1162.

Fastl, H. (1977). "Roughness and temporal masking patterns of sinusoidally amplitude modulated broadband noise," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 403–414.

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

- Goldstein, J. L. (1967). "Auditory spectral filtering and monaural phase perception," *J. Acoust. Soc. Am.* **41**, 458–479.
- Kohlrausch, A. (1988). "Masking patterns of harmonic complex tone maskers and the role of the inner ear transfer function," in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London), pp. 339–346.
- Kohlrausch, A. (1995). "Comment on 'The effect of temporal waveform shape on spectral discrimination by normal-hearing and hearing-impaired listeners' [*J. Acoust. Soc. Am.* **94**, 2074–2082 (1993)]," to be published in *J. Acoust. Soc. Am.*
- Kohlrausch, A., and Houtsma, A. J. M. (1991). "Edge pitch of harmonic complex tones," *IPO Ann. Progr. Rep.* **26**, 39–49.
- Kohlrausch, A., and Houtsma, A. J. M. (1992). "Pitch related to spectral edges of broadband signals," *Philos. Trans. R. Soc. London Ser. B* **336**, 375–382.
- Kollmeier, B., Gilkey, R. H., and Sieben, U. K. (1988). "Adaptive staircase techniques in psychoacoustics: A comparison of human data and a mathematical model," *J. Acoust. Soc. Am.* **83**, 1852–1862.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Lynch, J. T. (1988). "Reduction of peak/rms ratio of speech by amplitude compression and quadratic phase dispersion," *J. Audio Eng. Soc.* **36**, 147–152.
- Moore, B. C. J., and Glasberg, B. R. (1989). "Difference limens for phase in normal and hearing-impaired subjects," *J. Acoust. Soc. Am.* **86**, 1351–1365.
- Patterson, R. D. (1987). "A pulse ribbon model of monaural phase perception," *J. Acoust. Soc. Am.* **82**, 1560–1586.
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982). "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram and speech threshold," *J. Acoust. Soc. Am.* **72**, 1788–1803.
- Patterson, R. D., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (1987). "An efficient auditory filterbank based on the gammatone function," Appendix B of SVOS Final Report: The Auditory Filterbank, APU report 2341.
- Quatieri, T. F., and McAulay, R. J. (1988). "Sinewave-based phase dispersion for audio preprocessing," *Proceedings ICASSP 1988*, pp. 2558–2561.
- Rodenburg, M. (1977). "Investigation of temporal effects with amplitude modulated signals," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 429–437.
- Ruggero, M. A., Rich, N. C., and Recio, A. (1992). "Basilar membrane responses to clicks," in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany and K. Horner (Pergamon, Oxford), pp. 85–92.
- Scherer, A. (1988). "Erklärung der spektralen Verdeckung mit Hilfe von Mithörschwellen und Suppressionsmustern," *Acustica* **67**, 1–18.
- Schroeder, M. R. (1970). "Synthesis of low-peak-factor signals and binary sequences with low autocorrelation," *IEEE Trans. Inf. Theory* **16**, 85–89.
- Sek, A., and Moore, B. C. J. (1994). "The critical modulation frequency and its relationship to auditory filtering at low frequencies," *J. Acoust. Soc. Am.* **95**, 2606–2615.
- Smith, B. K., Sieben, U. K., Kohlrausch, A., and Schroeder, M. R. (1986). "Phase effects in masking related to dispersion in the inner ear," *J. Acoust. Soc. Am.* **80**, 1631–1637.
- Strube, H. W. (1985). "A computationally efficient basilar-membrane model," *Acustica* **58**, 207–214.
- Zwicker, E. (1952). "Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones," *Acustica* **2**, 125–133.
- Zwicker, E. (1974). "Mithörschwellen und Erregungsmuster stark frequenzmodulierter Töne," *Acustica* **31**, 243–256.
- Zwicker, E. (1976). "Masking period patterns of harmonic complex tones," *J. Acoust. Soc. Am.* **60**, 429–439.
- Zwicker, E. (1976/77). "Mithörschwellenperiodenmuster amplitudenmodulierter Töne," *Acustica* **36**, 113–120.