On the audibility of all-pass phase in electroacoustical transfer functions

Møller, Henrik; Minnaar, Pauli; Olesen, Søren Krarup; Christensen, Flemming; Plogsties, Jan

Published in:
Journal of the Audio Engineering Society

Publication date:
2007

On the Audibility of All-Pass Phase in Electroacoustical Transfer Functions*

HENRIK MØLLER, *AES Member,* **PAULI MINNAAR,**[1] *AES Member,* **SØREN KRARUP OLESEN,**
(hm@acoustics.aau.dk) (pauli@am3d.com) (sko@acoustics.aau.dk)

FLEMMING CHRISTENSEN, AND JAN PLOGSTIES,[2] *AES Associate Member*
(fc@acoustics.aau.dk) (Jan.Plogsties@iis.fraunhofer.de)

Department of Acoustics, Aalborg University, 9220 Aalborg Ø, Denmark

Audible effects of second-order all-pass sections with center frequencies in the range of 1–12 kHz were studied in headphone listening experiments. All-pass sections give rise to two effects. 1) A perception of "ringing" or "pitchiness," which is related to an exponentially decaying sinusoid in the impulse response of all-pass sections with high Q factors. The ringing is especially audible for impulsive sounds, whereas it is often masked with everyday sounds such as speech and music. With an impulse signal the ringing was found to be audible when the decay time constant for the sinusoid exceeds approximately 0.8 ms (peak group delay of 1.6 ms), independent of the center frequency within the frequency range studied. 2) A lateral shift of the auditory image, which occurs when an all-pass section is inserted in the signal path to only one ear. The shift is related to the low-frequency phase and group delays of the all-pass section, and it was found to be audible whenever these exceed approximately 35 μs, independent of the signal.

0 INTRODUCTION

A linear and time-invariant sound reproduction system can be characterized by its transfer function in the s domain. From this the magnitude and phase at a given frequency ω can be found by inserting $s = j\omega$ (ω real). This is implied throughout this paper. The transfer function can be split into a minimum-phase component and an excess-phase component (see [1]). The excess-phase component can be split further into a linear-phase component and an all-pass component. In the s domain this can be written as follows:

$$\begin{aligned} H(s) &= H(s)_{\text{minimum phase}} \cdot H(s)_{\text{excess phase}} \\ &= H(s)_{\text{minimum phase}} \cdot H(s)_{\text{linear phase}} \cdot H(s)_{\text{all pass}}. \end{aligned} \quad (1)$$

The minimum-phase component has the same magnitude as the original transfer function, but the phase is the smallest possible for the particular magnitude. It has a causal and stable impulse response and a causal and stable inverse $[1/H(s)_{\text{minimum phase}}]$. The linear-phase component represents an initial delay in the impulse response. It has a magnitude of unity and a phase that is a linear function of frequency with a negative slope. The all-pass component also has a magnitude of unity, but the phase is not a linear function of frequency. The all-pass component can be divided into a number of first- and second-order all-pass sections.

The aim of this work was to study the effects of all-pass components in transfer functions and to determine under which circumstances such effects are audible. A listening experiment was carried out in which signals that had been filtered with a second-order all-pass section were compared to the original unfiltered signals. Listeners adjusted the properties of the all-pass section until its presence was just audible. In this way a range of all-pass sections that may have an audible effect was determined. Signals were presented through headphones with the all-pass section inserted in one or both sides. The signals were chosen to make the effect of the all-pass section as audible as possible.

The work was carried out to give a background for evaluating the significance of all-pass phase in binaural synthesis and playback. However, the results may also be

*Manuscript received 2006 August 12; revised 2007 January 24.
[1]Now at AM3D A/S, 9200 Aalborg SV, Denmark.
[2]Now at Fraunhofer Institut für Integrierte Schaltungen, 91058 Erlangen, Germany.

useful for other purposes and applications, for example, evaluating all-pass phase in multipath transmission systems such as multidriver loudspeakers and headphones and their crossover filters.

0.1 Phase and Delay

It is our impression that there are a number of potential misconceptions about the phase of a transfer function and the connection between phase and delay. The phase shift is sometimes interpreted as being equivalent to a shift in time of a signal. Depending on the signal and the course of the phase along frequency, this time shift can materialize in different ways, and it is important to distinguish between these. In particular it is important to understand the concepts of delay, phase delay, and group delay. A summary of the definitions is given in the following, whereas more details and some instructive examples are given in Appendix 1.

A system performs a delay if the output is identical to the input, except that it occurs later. If the shift in time is τ_d, the phase $\phi(\omega)$ of the transfer function is given by

$$\phi(\omega) = -\omega\tau_d. \tag{2}$$

Thus the phase of a delay is linear in the (ω, ϕ) plane, and it has a value of zero at a frequency of zero.

The phase delay $\tau_p(\omega)$ is defined as

$$\tau_p(\omega) = -\frac{\phi(\omega)}{\omega}. \tag{3}$$

The phase delay indicates a shift in time that applies to a single frequency of a signal. As seen, $\tau_p(\omega)$ can vary with frequency, while τ_d cannot since it applies to the entire signal.

The group delay $\tau_g(\omega)$ is defined as

$$\tau_g(\omega) = -\frac{\partial\phi(\omega)}{\partial\omega}. \tag{4}$$

The group delay indicates a shift in time that applies to the envelope of a signal or part of a signal that has its frequency content in a narrow region around ω. Like $\tau_p(\omega)$, $\tau_g(\omega)$ can vary with frequency.

Since a delay is represented by a single value, it is not correct to say that a signal has been exposed to a frequency-dependent delay, whereas it is fine to speak of frequency-dependent phase and group delays. Neither the phase nor the group delays should be interpreted as a delay of the signal as a whole.

0.2 All-Pass Sections

This subsection gives a brief review of all-pass sections and their characteristics with focus on second-order sections and issues that are important for the understanding of the present study. More details and examples are found in Appendix 2.

In the complex s plane all-pass sections have zeros in the right half-plane and poles at corresponding mirror-image positions in the left half-plane. In the complex z plane all-pass sections have zeros outside the unit circle and poles at the reciprocal positions inside the unit circle. It is the mirror-image positions, or the reciprocal positions, that guarantee the flat magnitude. Examples of pole–zero positions are given in Fig. 1.

The transfer function in the complex s plane of a second-order all-pass section is given by

$$H(s)_{\mathrm{ap2}} = \frac{s^2 - \frac{\omega_0}{Q}s + \omega_0^{\,2}}{s^2 + \frac{\omega_0}{Q}s + \omega_0^{\,2}} \tag{5}$$

where Q is the quality factor, or Q factor, and ω_0 (or $f_0 = \omega_0/2\pi$) is the center frequency. If Q is ½ or lower, a second-order all-pass section turns into two first-order sections.

As mentioned, the magnitude is unity. Fig. 2 shows an example of phase, phase delay, and group delay. The phase is zero at a frequency of zero. As the frequency increases, the phase decreases monotonically to -2π while passing $-\pi$ at the center frequency. The higher the Q value, the more rapid is the change of phase around the center frequency.

The phase and group delays are nearly constant in a frequency range from zero to somewhere below the center frequency. The two have the same value given by

$$\tau_p(\omega)_{\mathrm{ap2}\atop \lim\,\omega\to 0} = \tau_g(\omega)_{\mathrm{ap2}\atop \lim\,\omega\to 0} = \frac{2}{\omega_0 Q}. \tag{6}$$

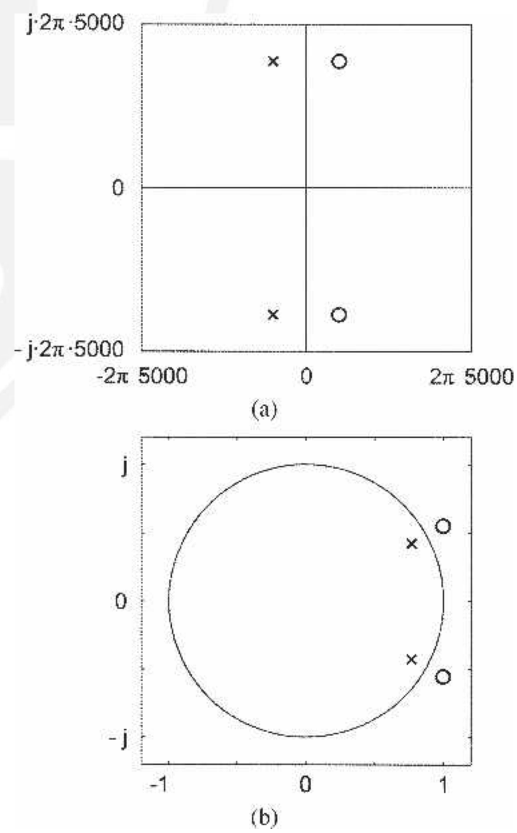

Fig. 1. Examples of pole–zero positions for second-order all-pass section. f_0 = 4 kHz; Q = 2; 48-kHz sampling frequency. (a) In complex s plane. (b) In complex z plane.

Around the center frequency the phase and group delays each have a peak. The peak is especially prominent for the group delay, and for high values of Q ($Q^2 \gg 1/4$) it appears nearly at the center frequency. The highest value, the peak group delay, is given by

$$\tau_{\text{g peak}} \approx \frac{4Q}{\omega_0}. \tag{7}$$

The group-delay bandwidth, that is, the width of the range around the center frequency, where the group delay is more than half of its peak value, is for high Q given by

$$\text{BW}_\text{g} \approx \frac{\omega_0}{Q}. \tag{8}$$

The impulse response of a second-order all-pass section consists of a single impulse (delta function) and an exponentially decaying sinusoid as given by

$$h(t)_{\text{ap2}} = \delta(t) + A\text{e}^{-t/\tau}\cos(\omega_1 t + \theta) \tag{9}$$

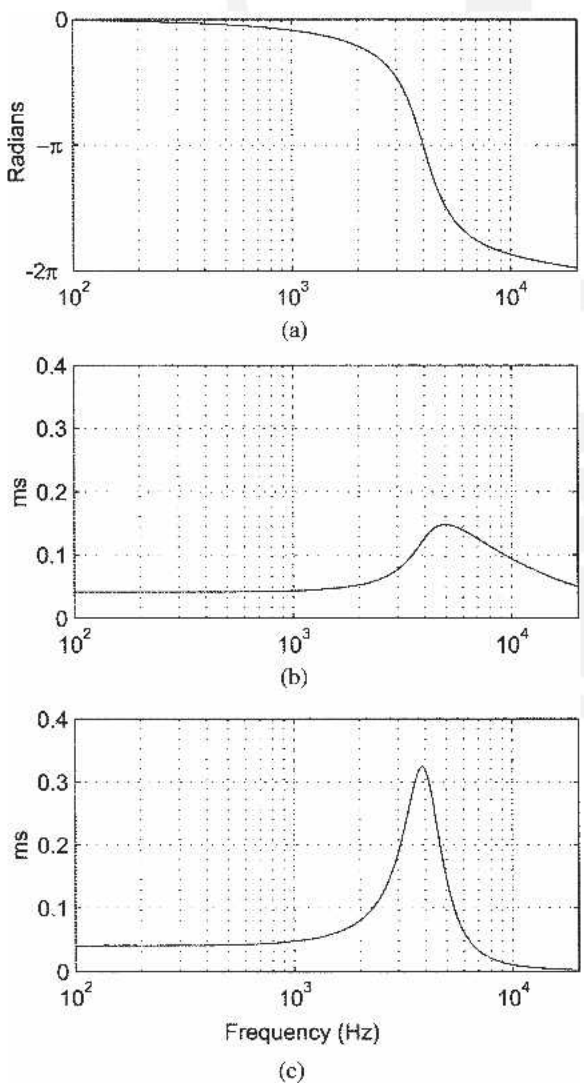

Fig. 2. Example of phase, phase delay, and group delay for second-order all-pass transfer function. $f_0 = 4$ kHz; $Q = 2$. (a) Phase. (b) Phase delay. (c) Group delay.

where the decay time constant is given by

$$\tau = \frac{2Q}{\omega_0}. \tag{10}$$

At high Q ($Q^2 \gg 1/4$) the frequency of the decaying sinusoid ω_1 is nearly equal to the center frequency ω_0, and the number N_τ of cycles within a time constant is given by

$$N_\tau \approx \frac{Q}{\pi}. \tag{11}$$

The amplitude A is given by

$$A \approx -\frac{2\omega_0}{Q}. \tag{12}$$

Hence for high values of Q the sinusoid has low amplitude and decays slowly, while it has higher amplitude and decays faster for low values of Q. The effect of the center frequency on the decaying sinusoid is the opposite. Fig. 3 shows examples of impulse responses.

The decaying sinusoid can be interpreted as the energy around the center frequency being delayed by the peak group delay. The decay time constant is proportional to the peak group delay [Eqs. (7) and (10)], thus the longer the peak group delay, the longer is the decay. Note also that the amplitude of the decaying sinusoid is proportional to the group-delay bandwidth [Eqs. (8) and (12)]. The wider the group-delay bandwidth, the more energy is delayed and, hence, the higher the amplitude.

The all-pass sections described so far have zeros in the right half of the complex s plane and poles in the left half-plane. The phase is negative, and the impulse response is causal. In principle, all-pass sections can also have zeros in the left half-plane and corresponding poles in the right half-plane. In that case the phase is positive, and the impulse response is noncausal. Such noncausal all-pass sections do not exist in physical systems, but it is relevant to consider them because they may occur as a result of calculations, such as when a transfer function that includes a causal all-pass section is inverted with the aim of equalizing it.

Whether an all-pass section is inverted or not, its impulse response is infinitely long—decaying either forward or backward in time. In order to implement the inverted version, its impulse response has to be truncated in time and delayed accordingly. In the following, such a truncated and delayed noncausal all-pass section will be referred to as an inverted all-pass section.

0.3 Audibility of Phase

The human sensitivity to phase has been the topic of many studies, and the question has generated heated discussions in the professional community and among enthusiastic audio amateurs. Helmholtz [2, pp. 124–127] is often quoted as claiming that the ear is insensitive to phase. It is true that this was his main conclusion, but he carefully pointed out limitations in his experiments, and later studies have shown that there are audible effects connected to phase aspects of signals. Some of the effects are subtle and occur only with certain signals and under certain well-controlled and ideal listening conditions. On the other hand, there is no doubt that audible effects of phase do exist—they are real and repeatable, and in some cases they are clearly audible to everyone.

One effect relates to the quality of complex tones (usually timbre and/or pitch) which, under certain conditions, change with the phase of individual tones. This has been shown for harmonic tone complexes (for example, [3]–[6], [7, sec. 70], [8]–[25]) and also for more narrow-band signals such as modulated tones (for example, [3], [7, sec. 69], [26]–[29]). Changes in timbre have also been shown as a result of a simple polarity change of the signal (180° change at all frequencies) (for example, [6], [18], [30], [31]). Another effect relates to transients, which may sound (and look) very different when the phase is changed of some or all frequency components (for example, [7, sec. 71], [30], [32] as well as several studies reported in more detail in the following).

Most of the investigations have used signals where the phase conditions were manipulated in a rather sophisticated manner, and only few have studied the effect of transmitting signals through specific transfer functions that modify the phase.

For minimum-phase transfer functions, phase and magnitude are mathematically linked, and changes of the phase only occur if the magnitude deviates from a flat response. Nevertheless minimum-phase distortion has given rise to worries, especially in connection with the high- and low-pass filters that exist at the lower and upper frequency ends of any transmission of audio signals. When the frequency range is limited in either end, the signal waveforms are obviously distorted. The natural question is, whether this distortion is audible. However, due to the link between phase and magnitude, an "undistorted reference" with the same magnitude but without the phase distortion does not exist. Lacking this reference, it is often taken for granted that the phase should preferably be zero or linear, and tests have been made of the audibility of the difference between these and minimum phase (for example, [33]). However, a transfer function with zero phase and a frequency-dependent magnitude is unnatural. Its impulse response is symmetrical around zero, so it is noncausal and does not exist in the physical world. Thus it is not a straightforward matter that such a transfer function should be a desirable reference. Making the phase linear instead of zero only means that the symmetrical impulse response of the zero-phase version is delayed. This makes an (approximate) physical implementation possible, but still it does not seem as an obvious and natural reference.

0.4 Audible Effects of All-Pass Transfer Functions

The presence of all-pass sections may be heard in two ways. 1) The decaying sinusoid in the impulse response can be heard as "ringing" or "pitchiness", and 2) a lateral shift can be observed if the all-pass phase is introduced in the signal path to only one ear. The shift goes toward the opposite side of the all-pass section (for a noninverted all-pass section).

The ringing is most obvious for impulse-like signals that concentrate their excitation at a certain time, while they leave the following period open for optimum detection of the decaying sinusoid. If an impulse-like piece of a signal is repeated, such as in a square wave of appropriate frequency, a change in timbre or tone quality may be observed. For a given center frequency the ringing is more audible with high Q than with low Q. For the lateral shift it is just the opposite; for a given center frequency, the shift is larger and more obvious with low Q than with high Q.

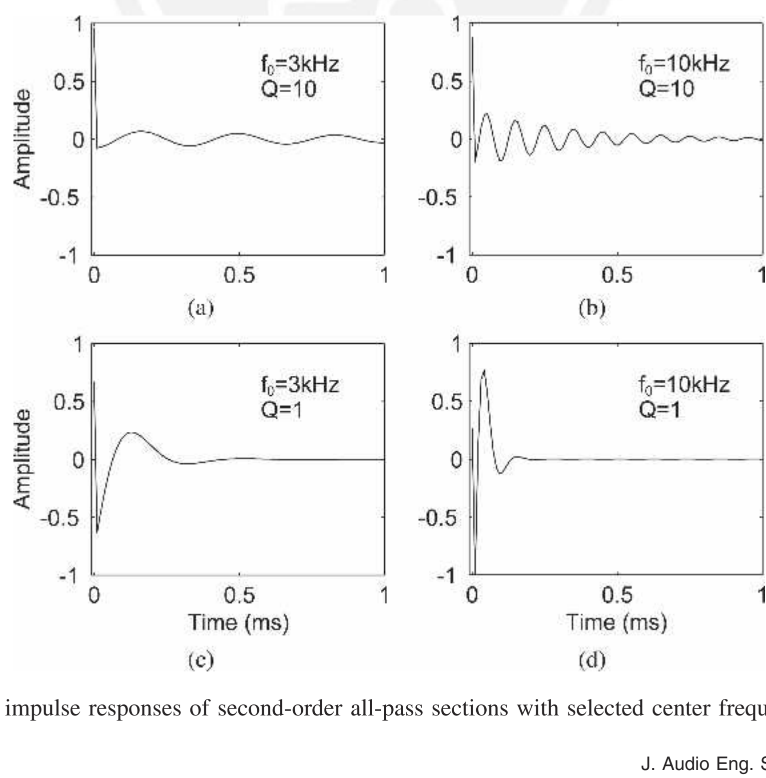

Fig. 3. Examples of impulse responses of second-order all-pass sections with selected center frequencies and Q factors.

The phenomenon of ringing is reported in several studies, but the lateral shift does not seem to have drawn the attention in previous studies of all-pass phase audibility. An exposure can be either monotic (sound in only one ear), diotic (same sound in both ears), or dichotic (different sounds in the two ears). In the previous studies, when the exposure was dichotic, it had the natural interaural differences from listening in a sound field (loudspeaker experiments). Hence all-pass sections were not inserted in one side with the original sound in the other side. The existence of the lateralization effect is easily observed in informal listening tests, and it is not unexpected since lateralization due to interaural phase differences in general is a well-known phenomenon. The effect may have significant practical implications for binaural synthesis. HRTFs at the two sides have typically different all-pass phase, and if all-pass phase is ignored or not applied properly in the synthesis, it may thus affect the localization.

0.5 Previous Investigations

A review of previous studies of the audibility of effects of all-pass sections will be given in the following. The studies differ a lot in method and terminology, but attempts have been made to use a common terminology in accordance with Sections 0.1 and 0.2. Since center frequency and Q factor are the basic parameters of all-pass sections, these have been calculated whenever possible, using the formulas of Section 0.2.

Patterson and Green [34] had two to four subjects listening to impulse responses of second-order all-pass filters that were truncated in time. Comparisons were made between two different responses, thus comparisons were not made with an original, unfiltered reference. The presentation was monotic via headphones (TDH-39). In one experiment the duration of the impulses (truncation point) was varied, and it was found that all-pass filters at 800 Hz and 1.6 kHz with peak group delays of 3.2 ms (Q values of 4 and 8, respectively) could be discriminated, provided that the impulse responses were 2.5 ms or longer. In another experiment the duration was kept constant at 10 ms, while the peak group delays were varied. It was found that all-pass sections at 400 and 800 Hz could be discriminated when the peak group delays were around 6 ms (Q values of 4 and 8, respectively), but not when the peak group delays were around 60 ms (Q values of 40 and 80, respectively). At values in-between there were large individual differences. [The authors used the term Huffman sequences (after [35]) for digital-domain versions of truncated impulse responses of all-pass sections.] Patterson and Green also varied the distance between the center frequencies of the two filters in a comparison, but they did not report the Q values.

Also Green [36] tested the audibility of the difference between truncated impulse responses of two all-pass filters. As reference he used single second-order all-pass filters with center frequencies at 650, 1900, and 4200 Hz. Peak group delays were at 1.6, 3.2, or 6.4 ms (Q values of 1.6 to 42, depending on frequency), and truncation was made at two times the peak group delay. The test filters were also all-pass filters, always at the same frequency and of the same duration as the reference, but the peak group delay was changed by varying Q and by cascading up to 16 filters. Thresholds of discrimination were found for test filters with peak group delays higher as well as lower than the reference. Thresholds were reported in terms of the difference in peak group delay between reference and test. An average value of approximately 2 ms was reported, but this value conceals a significant variation (0.69 to 4.46 ms) with other parameters of the stimuli.

Most of the later experiments compared filtered and unfiltered signals. Jensen and Møller [14] studied the audibility of effects of second-order all-pass sections with center frequencies at 250, 500, 800, 1200, and 2000 Hz. The stimulus was a 50-μs square pulse. Only two subjects participated (the authors). For each center frequency they found the lowest Q that produced an audible ringing. At frequencies above 800 Hz the ringing was found to be audible when the time constant in the exponentially decaying sinusoid of the impulse response was at least 0.8 ms (peak group delay of 1.6 ms). At frequencies below 800 Hz the ringing was found to be audible when Q was at least 1.8.

Blauert and Laws [37] did a similar experiment with center frequencies at 500 Hz, 1, 2, 4, and 8 kHz. In introductory tests with many different signals they found that the highest sensitivity was with brief impulses played back under anechoic conditions. In the formal test with seven subjects, they used 25-μs rectangular pulses. Filters with peak group delays of 1, 2, 3, and 4 ms were used (Q values of 0.8 to 50, depending on frequency), and listening was diotic via headphones. Thresholds of peak group delay varied with frequency from 3.3 ms at 500 Hz through a minimum of 0.9 ms at 2 kHz to 1.9 ms at 8 kHz. The data correspond to Q values around 3 up to 2 kHz, rising to 24 at 8 kHz. Results were not significantly different if the presentation was monotic (500 Hz, 1 kHz, and 2 kHz, four subjects). In another experiment Blauert and Laws compared signals exposed to the simulated phase of some typical transducers. One of the simulations was dominated by a 7-kHz all-pass section, and a detection threshold was found at a peak group delay of 1.1 ms (Q of 14). Blauert and Laws also showed that training could lower the threshold considerably.

Suzuki et al. [38] tested the audibility of effects of first-order all-pass sections at the frequencies of 300 Hz and 1 kHz. A variety of brief stimuli was used (single sinusoidal, square and triangular pulses, as well as short bursts of these). For eight subjects who listened to filtered and unfiltered signals, they found average detection rates of the 300-Hz filter slightly above chance for loudspeaker listening (anechoic and normal rooms), somewhat higher for headphone listening. There were large individual differences. The effect of the 1-kHz all-pass filter was not audible even in the headphone situation.

Lipshitz et al. [21] studied the audibility of effects of first- and second-order all-pass filters with center frequencies roughly in the range of 100 Hz to 3 kHz and, for the second-order filters, with Qs in the range from 0.5 to 2.8.

In search of effects they tried a large number of artificial and natural waveforms. They noted that effects were more audible the lower the center frequency and the higher the Q. With natural signals, effects were less pronounced than with artificial signals but could still be demonstrated. The study did not attempt to determine specific criteria for audibility of the effects of all-pass sections, but mainly aimed at demonstrating the existence of audible effects, and it strongly encouraged further studies.

Lee [39] tested the audibility of effects of first- and second-order all-pass filters. Up to six identical, cascaded filters were used (center frequencies in the range of 17 Hz to 17 kHz, Qs at 1.4, 5.0, or 14.3). With a variety of real-life signals (such as music, speech, key jangling) subjects were not able to detect the difference between filtered and unfiltered signals. With 1-ms square pulses repeated at 100–500 Hz the difference was audible with six first-order filters at 1 kHz (not mentioned, but most likely also with other configurations).

In an informal experiment Fincham [40] demonstrated that two second-order all-pass sections, each with a Q of 1.4, produced a distinct change in sound quality of an eight-cycle 40-Hz tone burst. The center frequency was not given, but the filters were said to be those used by Lipshitz et al. [21] (that is, in the range 100 Hz to 3 kHz).

Deer et al. [41] used 2-kHz second-order all-pass filters with 0.5-, 1-, and 2-ms peak group delay (Q values of 1.6, 3.1, and 6.3). By using two, four, or eight filters in cascade, they produced peak group delays of 1, 2, 4, 8, and 16 ms. 2, 4, and 8 ms were produced by more than one combination, that is, with different group-delay bandwidths; these were analyzed in common. Six subjects listened to the impulse responses low-pass filtered at 4.1 kHz and played back diotically via headphones (Beyer DT-100). A statistically significant detection was seen for peak group delays above roughly 2 ms. In a second experiment the three combinations with peak group delays of 4 ms (but with different bandwidths) were compared against each other. Only the difference in the extreme comparison was audible, that is, four 1-ms filters (wide-band) versus one 4-ms filter (narrow-band). For this comparison the largest difference in group delay was slightly above 2 ms (occurring slightly below and above 2 kHz).

Karjalainen et al. [42] did an informal listening test with transients and impulses as stimuli. Their description is very short, but it seems that they compared signals that were filtered with two different all-pass filters rather than filtered and unfiltered signals. They found an audible difference between all-pass filters with a difference in group delays on the order of 2 ms, when the difference occurred within a critical band at "middle frequencies" (not further specified). For more complex phase curves (also not specified) values at 1 ms or even lower were reported. Their description gives reason to believe that at least some of their all-pass filters were much more complex than single first- and second-order sections.

Flanagan et al. [43] used all-pass filters at 1, 2, and 4 kHz to obtain peak group delays of 0.5, 1, 2, 4, 8, and 16 ms. The range of group delays was achieved by varying the Q of second-order sections in the range of 0.8 to 6.3 and by cascading 1 to 32 sections. Filtered and unfiltered impulses were compared monotically and diotically in a headphone experiment with five subjects. Mean group-delay thresholds were found to be 1.4–2.0 ms. No statistically significant effect of frequency or listening condition was found in an analysis of variance. Slightly higher thresholds were found with loudspeakers in a low-reverberation room and still higher ones in a more reverberant room.

In general the studies on all-pass phase agree that the playback conditions have a significant impact on the audibility, and it is usually found that the effects are most audible with headphone playback or loudspeaker playback under anechoic conditions. Most studies used brief impulses as test signals, since these were in general found to be the signal with which the effects of the all-pass phase was most audible. Several of the studies mention specifically a perception of ringing or "pitchiness," as would be expected from the exponentially decaying sinusoid of the impulse response. There is also general agreement that the effect of all-pass phase is much more difficult to detect with normal program material, such as speech and music.

Despite the agreement on these qualitative issues, it is difficult to see a general agreement about which all-pass filters produce audible effects and which do not, for instance given in terms of thresholds of Q as a function of frequency. Some studies were mainly made to demonstrate the existence of an audible effect of all-pass sections, and of the systematic studies, some did not use an unfiltered reference but compared the output of two all-pass filters, and some used cascaded filters rather than single second-order sections. The number of studies using all-pass sections with high center frequencies is limited, and none of the experiments included inverted all-pass sections. The studies did not examine the lateral shift from an all-pass section introduced in one side.

0.6 Present Study

As mentioned in the beginning, our main concern is the audibility of effects of those all-pass sections that occur in binaural synthesis and playback. All-pass sections in head-related transfer functions (HRTFs) and in headphone transfer functions seem to be always in the high-frequency range [44] (for headphones probably except multidriver types with low crossover frequencies), so we decided to restrict the study to all-pass sections at and above 1 kHz. It was further decided to use single second-order all-pass sections in noninverted and inverted forms, to always compare filtered signals with an unfiltered reference, and, for each center frequency, find values of Q for which the all-pass section has an audible effect on a signal. We also decided to study both the phenomenon of ringing and the phenomenon of lateral shift.

1 METHOD

By means of headphones, listeners were presented with two consecutive signals separated by an interval of 1 s.

The first signal was unfiltered, whereas the second signal was filtered with a second-order all-pass section, either in both sides or in only one side. When only one side was filtered, the other side still carried the unfiltered signal. The all-pass section had a given center frequency, and the task of the listener was to adjust the Q until the difference between the signals was just audible.

1.1 Conditions and Stimuli

As reported in previous studies and confirmed by informal listening tests, the ringing that occurs at high Q is most audible with short signals. Therefore a single impulse at the sampling frequency of 48 kHz was chosen as stimulus. For a given center frequency the lowest Q that produced an audible effect was determined. Such thresholds of ringing were found for all-pass sections with center frequencies at 1, 2, 4, 8, and 12 kHz. All-pass filters were inserted in either one or both sides. At 1 kHz and 2 kHz, however, filters in only one side were omitted, since pilot experiments revealed that the presence of the all-pass section was always audible for center frequencies below about 2 kHz—if not due to ringing then due to lateralization.

Pilot tests showed that the lateralization that occurs with all-pass sections in only one side is especially pronounced with low values of Q. Therefore, for a given center frequency the highest Q that produced an audible lateral shift was determined. Such thresholds of lateralization were found for all-pass sections with center frequencies of 1, 2, 4, 8, and 12 kHz. The stimulus for these tests was a pink-noise signal (700 ms long, with 50-ms raised-cosine ramps). Pilot experiments had shown that the lateralization is relatively independent of the signal, and pink noise was chosen because it effectively masks the ringing of the all-pass section. Therefore all-pass sections with center frequencies at 1 and 2 kHz could be included in the lateralization experiment without risk of being detected due to ringing.

Stimuli were presented through headphones (Beyerdynamic DT 990 PRO) that were equalized for their minimum-phase response measured at the blocked ear canal entrance of humans (mean of 13 subjects). In all cases both noninverted and inverted all-pass sections were used. Table 1 gives an overview of stimuli and conditions.

1.2 Discrete-Time Implementation of All-Pass Sections

All processing was done on digital signals with a sampling frequency of 48 kHz. Since the all-pass sections were implemented as digital filters, a discrete-time equivalent of the analog transfer function in Eq. (5) was required. This was done by determining the roots of the polynomials in the numerator and denominator of the analog transfer function and mapping these from the s plane to the z plane by means of the matched-z transform, $z = \exp(sT)$, where T is the sampling period. These poles and zeros were then used to construct the polynomials that describe the discrete-time transfer function,

$$H(z)_{\mathrm{ap2}} = \frac{(z - z_0)(z - z_0^*)}{(z - 1/z_0)(z - 1/z_0^*)} \frac{1}{z_0 \cdot z_0^*}. \tag{13}$$

Filters were implemented as FIR filters, using impulse responses of $H(z)_{\mathrm{ap2}}$ with such long duration that the impulse had effectively died out (varies with f_0 and Q).

1.3 Listening Test Procedure

The psychometric method was the method of adjustment. For a given center frequency and condition, the listener adjusted the Q by selecting one of the points on a 12-point scale drawn on an electronic tablet interfaced to a computer that played the preprocessed sound files. The steps were logarithmic (each factor of 1.29), and the scale always spanned 1.2 decades (a factor of 16). The Q at the top of the scale was generally above the threshold, whereas that at the bottom was below. The appropriate ranges of Q were determined during pilot tests. Each time a point on the scale was pressed, a stimulus pair (unfiltered and filtered) was played, and the subjects could listen to as many pairs as they wanted to encircle the threshold. When they had found the threshold, an OK button on the tablet was pressed, and a new adjustment was initiated at another center frequency and/or condition.

The eight listeners (four male and four female) had normal hearing according to an audiometric test and had not participated in similar experiments before. The listeners were seated at a table in a sound insulated room, while the computer and other equipment were placed in an adjacent control room. Before the start of the experiment, the listener was familiarized with the task of determining the threshold, but no training took place. Each listener participated in three sessions, each with two sound conditions (A+B, C+D, and E+F in Table 1). All combinations of sound condition and center frequency were presented twice, and within a session (that is, from two sound conditions) they were randomized individually for every listener. There were no restrictions in time for carrying out the adjustments, and a session typically lasted 10–12 min.

Table 1. Conditions used in listening experiments.

Condition	Type of Threshold	Center Frequencies	All-Pass Section	Transfer Function	Signal
A	Ringing	1,2,4,8,12 kHz	Both sides	Noninverted	Impulse
B	Ringing	1,2,4,8,12 kHz	Both sides	Inverted	Impulse
C	Ringing	4,8,12 kHz	One side	Noninverted	Impulse
D	Ringing	4,8,12 kHz	One side	Inverted	Impulse
E	Lateralization	1,2,4,8,12 kHz	One side	Noninverted	Pink noise
F	Lateralization	1,2,4,8,12 kHz	One side	Inverted	Pink noise

2 RESULTS

For each of the six conditions described in Table 1, and for each center frequency, means and standard deviations of the thresholds were calculated. Since the original scale of adjustment was logarithmic, and the data seem normally distributed on a logarithmic scale, both were determined logarithmically, that is, from log Q.

2.1 Thresholds of Ringing

The thresholds of ringing are given in Fig. 4. It is seen that data lie nearly perfectly on straight lines for all conditions. The slope of all indicated lines is 1 (note the different scales of abscissa and ordinate). In the logarithmic plane this means that Q/f_0 is constant, which again means that the time constant τ of the exponential decay of the ringing is constant [see Eq. (10)]. Then the peak group delay is also constant [Eq. (7)]. This means that the thresholds of ringing can be described by a single value independent of frequency, either of the decay time constant or of the peak group delay.

For the noninverted all-pass sections (left frames, conditions A and C) the indicated straight-line approximations correspond to a decay time constant of 0.8 ms, or a peak group delay of 1.6 ms. For the inverted all-pass sections (right frames, conditions B and D) the lines correspond to a decay time constant of 0.6 ms, or a peak group delay of 1.2 ms. Differences between data for inverted and noninverted all-pass sections are statistically significant (analyses of variance, significance levels of 0.003 for comparison A–B and 0.025 for comparison C–D). There is no statistically significant difference whether the all-pass section is present in one or both sides.

2.2 Thresholds of Lateralization

The thresholds of lateralization for the two conditions E and F are given in Fig. 5. Data points can be well approximated by lines with a slope of −1. In the logarithmic plane this means that $Q \cdot f_0$ is constant, which means that the low-frequency phase and group delays are constant [Eq. (6)]. Thus the thresholds of lateralization can be described by a single value for the low-frequency phase and group delays independent of center frequency. The straight-line approximations that are included in the two frames correspond to low-frequency phase and group delays of 35 μs. There is no statistically significant difference between the results for inverted and noninverted all-pass sections but, of course, the lateral shift goes in either direction.

Although the low-frequency phase and group delays are essentially the same for the different points, the frequency range, for which the "low-frequency" phase and group delays apply, varies a lot. The upper limit varies from just below 1 kHz to roughly 10 kHz.

3 DISCUSSION

The results of the present investigation were obtained with untrained listeners and using a simple method of adjustment to obtain the thresholds. As pointed out in the literature (see [37]), training may improve the subjects' ability to detect the all-pass phase significantly. Also the choice of psychometric method may influence the results.

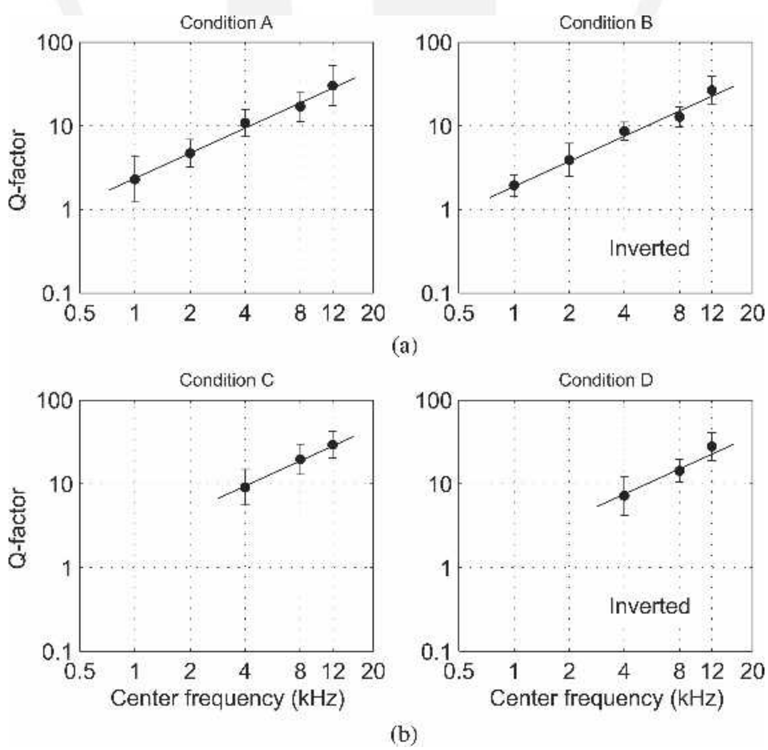

Fig. 4. Thresholds of ringing for second-order all-pass sections for four conditions A–D. Means and standard deviations of Q are given as a function of frequency (calculated from log Q). Approximations are proposed by straight lines with slopes of 1. (a) All-pass section in both sides. (b) All-pass section in one side.

For the ringing, the audibility depends much on the test signal, and the impulse stimulus was selected as especially revealing. Sound signals that are highly impulsive do occur in real life, such as from percussion, and the thresholds may apply also to such critical real-life signals. However, for most real-life signals the ringing is likely to be masked by the signal itself, and it is expected to be inaudible, even with Q values that are higher than the thresholds found. Room reflections may also mask the ringing and impair detection, whether they stem from the recorded stimulus or from playback in a room. Some authors have reported that it is more difficult to detect all-pass phase with loudspeaker reproduction than with headphone reproduction [37], [38], [43]. Since no difference was found regardless of whether the all-pass section is inserted in one or both sides, the audibility of the ringing is the same, whether it occurs in both ears or in one ear with the unfiltered signal in the other ear. This suggests that the detection takes place in connection with the individual ear and not as part of binaural processing. It was seen that the effects are slightly more audible with inverted all-pass sections than with noninverted all-pass sections. In general, backward masking is considered less effective than forward masking [45, ch. 3], and this may explain why the ringing is more audible when it arrives before the click than after the click.

The lateralization effect of an all-pass section inserted in one side depends only little on the signal, and the thresholds obtained with a pink-noise signal will also apply to everyday signals such as speech and music. It could be hypothesized that if the stimulus is restricted to a narrow frequency band around the center frequency, it would be the value of the group delay in that region, that is, close to the peak group delay, and not the low-frequency phase or group delay, which determines whether a lateral shift occurs. This was not explored further, and even if the hypothesis holds, it would only apply to very special signals, and the practical significance may be limited.

In the experiments the reproduction level was simply taken as a pleasant listening level, and means were not taken for an accurate control of this. Since the thresholds of ringing are determined by the joint presence of the impulse and the decaying sinusoid, they are not absolute thresholds for the decaying sinusoid. An influence from the reproduction level is most likely moderate but cannot be excluded. A phenomenon related to the absolute threshold of the sinusoid was observed in pilot experiments, where it was found that when Q was raised to very high values (above those used in the main experiment), the presence of the all-pass section became inaudible again. This can be explained by the fact that at very high Q the ringing becomes infinitely long while the amplitude becomes zero. Similar observations were made by Patterson and Green [34].

Only second-order all-pass sections were used in the experiments. This is not really a limitation, since a second-order section turns into two first-order sections when Q gets below ½. Such low values did occur in the stimuli if the adjustment procedure led to such values. However, all mean thresholds turned out to be above ½, and only two were close (thresholds of lateralization at 12 kHz).

3.1 Comparison with Previous Studies

Two earlier investigations found thresholds for ringing of single all-pass sections using an unfiltered reference [14], [37]. A comparison with their data is given in Fig. 6. A good agreement between investigations is seen. At frequencies where the studies overlap, data for Q are within a factor of 1.5 at the highest. It seems fair to conclude that the minimum audible Q is proportional to the center frequency in a range from 1 to 2 kHz and up, while it most likely levels out around a value of 2 to 3 below this range.

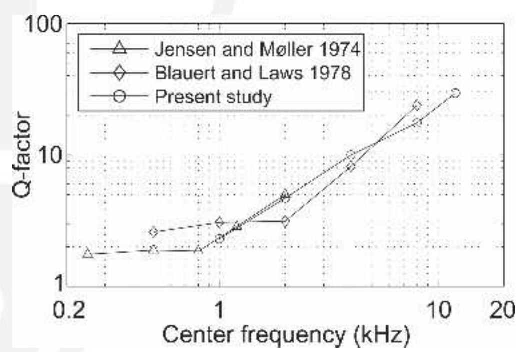

Fig. 6. Thresholds of ringing for second-order all-pass sections. Data by Jensen and Møller [14], Blauert and Laws [37] (mean of monotic and diotic presentation), and present study (noninverted, mean of data for one- and two-sided all-pass sections).

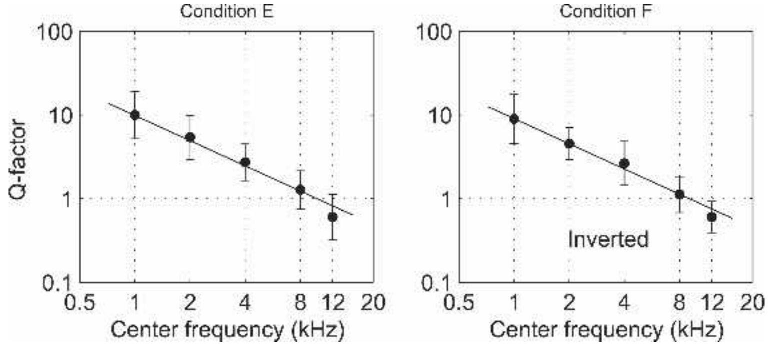

Fig. 5. Thresholds of lateralization for second-order all-pass sections (in one side) for two conditions E and F. Means and standard deviations of Q are given as a function of frequency (calculated from log Q). Approximations are proposed by straight lines with slopes of −1.

As mentioned in Section 2.1, the high-frequency threshold corresponds to a decay time constant that is independent of frequency. If *duration* of the ringing is defined as the time it takes the sinusoid to decay a certain amount, the high-frequency threshold also corresponds to a fixed duration, independent of frequency. As an example and with $\tau \approx 0.8$ ms inserted, it takes $4.6\ \tau \approx 3.7$ ms to decay to 1% of the initial value (40 dB). The constant value of Q below 1 kHz means that, for this frequency range, the number of cycles N_τ within the time constant is fixed [Eq. (11)], which also means that there is a fixed number of cycles within the duration of the ringing. As a parallel example and with $Q \approx 2$ inserted, there are $(Q/\pi)\ 4.6 \approx 2.9$ cycles within a decay to 1%. Thus the joint threshold below and above 1 kHz can be interpreted so that, in order to be audible, the ringing must have a certain duration (high frequencies) and contain a certain number of cycles (low frequencies). This interpretation was also proposed by Møller and Jensen [14].

Some researchers used all-pass transfer functions, not with the aim of studying specific all-pass configurations, but as a means to producing group delay [34], [36], [41], [43]. To obtain a certain amount of group delay, they cascaded several second-order sections rather than increasing the Q of a single section. Although then the data are not directly comparable, it is possible to compare the peak group delays across the different configurations. A comparison with the two investigations that used an unfiltered reference [41], [43] is shown in Fig. 7.

Fig. 7 also shows a good agreement between investigations with peak group delays of 1–2 ms for the frequency range above 1 kHz. In particular, data from the two investigations that used multiple all-pass sections do not seem to deviate from those of the others [41], [43]. It is thus tempting to conclude that the audibility is determined in general by the peak group delay. On the other hand this may simplify matters too much. As an example, Fig. 8 shows impulse responses of two transfer functions with the same peak group delay obtained in two different ways. In Fig. 8(a) a single second-order all-pass section with a high Q is used, whereas four second-order sections with a lower Q are used in Fig. 8(b). It is reasonable to believe that the higher level of late ringing in Fig. 8(b) is more audible. The difference can be explained by the difference in group-delay bandwidth; the low-Q sections of Fig. 8(b) have a wider bandwidth, and hence they leave more energy to be exposed to the peak group delay. It is proposed that not only the peak group delay is of importance for the audibility, and further studies are needed.

The ringing from an all-pass section occurs separated in time from the main impulse by an amount that is related to the decay time constant, and thus the peak group delay. It is interesting to note that the results for minimum detectable peak group delay in all-pass sections correspond well with values for the temporal resolution or temporal acuity of our hearing, obtained with a variety of methods. Moore gave a review [45, ch. 4] that showed values generally on the order of 2–3 ms, and in certain situations down to below 1 ms. Zwicker and Fastl [46, p. 262] gave a value of 2 ms for the temporal resolution of our hearing.

As mentioned in the Introduction, the lateralization effect of all-pass sections has not been covered in the literature. Since the effect seems closely connected to the low-frequency phase and group delays, it would be natural to compare the threshold for these with detection thresholds for an interaural time delay (ITD). For a variety of stimuli Klump and Eady [47] found values in the range of 9–75 μs, with the highest values for either pure tones at low frequencies, noise bands above 1 kHz, or very brief stimuli. For 100-μs clicks Hall [48] found thresholds in the range of 20–50 μs with large individual differences. Yost et al. [49] found thresholds in the 20–50-μs range for a 1-ms pulse. Thresholds increased significantly if frequencies below 1500 Hz were removed by high-pass fil-

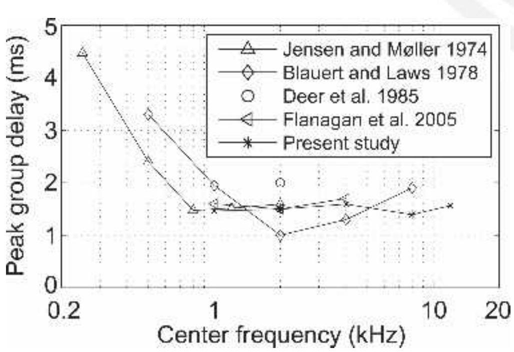

Fig. 7. Thresholds of ringing for various configurations of single and multiple all-pass sections. Data by Jensen and Møller [14], Blauert and Laws [37] (mean of monotic and diotic), Deer et al. [41] (diotic), Flanagan et al. [43] (mean of monotic and diotic), and present study (noninverted, mean of data for one- and two-sided all-pass sections).

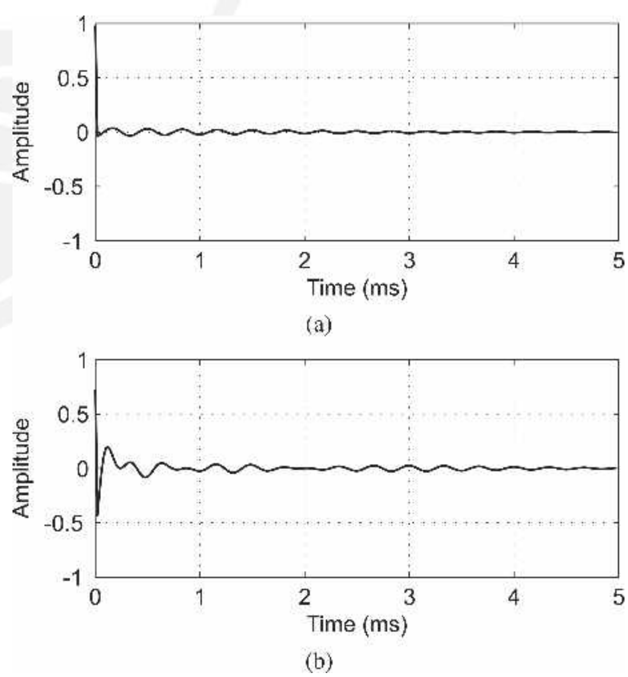

Fig. 8. Impulse responses of two all-pass transfer functions with same peak group delay of 4 ms. $f_0 = 3000$ Hz. (a) Single second-order all-pass section; $Q = 19$. (b) Four cascaded second-order all-pass sections, $Q = 4.75$.

tering of the pulse. Hafter and De Maio [50] measured ITD thresholds for band-pass-filtered clicks. Thresholds were in the range of 15–40 μs for 100-Hz to 2-kHz filtered clicks, somewhat higher for clicks filtered to the 1–2-kHz or 3–4-kHz ranges. Zwicker and Fastl [46, sec. 15.1] reported that the just noticeable interaural delay had been measured from 30 to 200 μs. Klinke [51, p. 210] gave a value of 30 μs, or half of this under optimal conditions.

It is also interesting to compare the thresholds of the low-frequency phase and group delays with current knowledge of detection thresholds in localization. For frontal directions it is believed that the main cue for detecting a small azimuth angular difference is the ITD. For these directions, Blauert [52, sec. 2.1] gives values for the localization blur that are generally in the range of 1–4°. These angles correspond to ITDs of roughly 10–30 μs. The results of the current study are therefore in fair agreement with our knowledge of detection thresholds for ITD as well as of localization blur in front. In general it is interesting that our hearing seems to compare the signals from the two ears with a time resolution of around 35 μs, which is roughly 50 times better than the temporal resolution of the analysis carried out by the individual ear (1–2 ms).

Now having the data supported by observations from the literature, Fig. 9 gives a summary of the results, including an extension toward lower frequencies based on literature data.

3.2 Effect of Minimum-Phase Component

The thresholds found apply to transfer functions that are purely all-pass and thus have a flat magnitude. Such a situation occurs in practice, for example, if only the minimum-phase component of a transfer function that also includes all-pass phase is equalized. The transfer function after equalization will then be purely all-pass.

If the all-pass component is accompanied by a minimum-phase component, it may change the signal considerably. Since the ringing thresholds depend much on the signal, it is, therefore, difficult to predict in general what effect a minimum-phase component will have on the audibility of ringing. Some consideration can be given, however, to the specific and common situation where a transfer function has zeros in the right half of the complex s plane without corresponding poles in the left half-plane. Despite the "missing" poles, such zeros also represent all-pass sections. The all-pass sections can be identified by adding, for every zero in the right half-plane, a pole and a zero at the mirror-image position in the left half-plane. This does not change the transfer function, but the all-pass component can then be identified by combining the added poles in the left half-plane with the original zeros in the right half-plane.

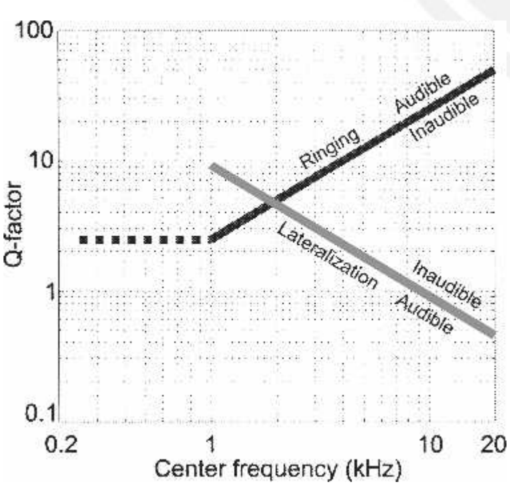

Fig. 9. Summary of thresholds. Ringing thresholds for noninverted all-pass sections. Literature data added below 1 kHz (- - -).

The added zeros in the left half-plane belong to the minimum-phase component, and they give rise to a dip in the magnitude. The dip is at the frequency where a possibly audible ringing occurs, and consequently the amplitude of the ringing is reduced. As Q gets higher, the added zeros come closer to the imaginary axis, and the dip in the magnitude becomes deeper. As a result the more obvious and audible the ringing of the pure all-pass section is, the more it is reduced in amplitude by the minimum-phase part. Unpublished inspections of HRTFs from 40 subjects (from [53]) show that zeros in the right half-plane without corresponding poles in the left half-plane are a common phenomenon in HRTFs, and informal listening tests suggest that they do not give rise to audible ringing.

The lateral shift by an all-pass section introduced in one side is nearly independent of the signal. Therefore it is reasonable to assume that the minimum-phase part of a transfer function will not affect the audibility.

3.3 Multiple All-Pass Sections

It is not straightforward to extend the thresholds for ringing to cases with multiple all-pass sections. If the center frequencies differ, ringing may occur at more frequencies at the same time, and in case of more all-pass sections at the same frequency, other decay patterns than exponential decay will be seen. Since, in general, the threshold seems closely related to the peak group delay, it is tempting to propose that the audibility can be evaluated from the peak group delay of the total all-pass component. This is to some extent supported by the similarity of peak-group-delay thresholds for multiple and single all-pass sections at the same center frequency (Fig. 7). On the other hand, other factors may also play a role (such as the group-delay bandwidth, as demonstrated in Fig. 8), and more investigations are needed to clarify this.

Transfer functions with very high numbers of all-pass sections exist in the transmission of sound in rooms. This was studied by Johansen and Rubak [54] among others, who divided room impulse responses (including a loudspeaker) into their minimum and excess-phase components. Both components turned out to have impulse responses with durations on the same order of magnitude as the reverberation time. The excess phase was not split into linear phase and all-pass phase components, but since no prominent initial delay was seen in the impulse responses, it can be concluded that the excess phase was mainly all-pass. For a room with a reverberation time of 0.6 s, the excess phase amounted to nearly 1000 rad at 10 kHz, which means that there were more than 300 pole–zero pairs in the transfer function (each pair gives up to π rad).

Listening tests indicated that the presence of the excess phase was audible with normal music signals. The authors did not show the group delay of the transfer function, and it remains to be investigated whether the peak group delay is a useful measure of audibility also with such a high number of (probably overlapping) all-pass sections.

The lateralization seems to be connected to the fact that all-pass sections work essentially as a delay at low frequencies. This suggests that such low-frequency delays of more all-pass sections in a transfer function can simply be added to get their joint effect. This means that a possible lateralization effect should be evaluated from the low-frequency phase or group delays of the total transfer function, for example, as found by $\tau_p(\omega)_{\text{all-pass}}$ or $\tau_g(\omega)_{\text{all-pass}}$, $\omega \to 0$. It is further reasonable to believe that if there are also all-pass sections on the opposite side, the delays of these should be subtracted to get the effective interaural delay. If this value exceeds the threshold of roughly 35 μs, the joint effects of all-pass sections in both sides will be audible. It is important to keep in mind, though, that the results of the current experiment were obtained with second-order all-pass sections with center frequencies at 1 kHz and above. For center frequencies below 1 kHz this cannot be assumed valid, since the frequency range where the all-pass section acts as a delay is then reduced.

4 CONCLUSIONS

The aim of this work was to study the effects of all-pass components in transfer functions, and to determine under which circumstances the effects are audible. The work was carried out with the purpose of giving a background for evaluating the all-pass phase in binaural synthesis and playback, but the results may also be useful for other purposes and applications.

All-pass sections give rise to an exponentially decaying sinusoid, which may be perceived as ringing or "pitchiness." The higher Q is, the longer is the decay, and the more audible is the ringing. It was found that the audibility of the ringing depends strongly on the signal. For the most sensitive signal, an impulse, thresholds for Q correspond to a decay time constant of the ringing around 0.8 ms for center frequencies in the range of 1–12 kHz. The peak group delay of an all-pass filter is twice the decay time constant. Therefore the threshold may also be given in terms of a peak group delay of 1.6 ms. The audibility is the same, whether the all-pass section is inserted in both sides or in only one side, which suggests that the ringing is detected in the individual ear and not as part of a binaural processing. The thresholds are slightly lower for inverted all-pass sections, which seems natural, since backward masking is usually less effective than forward masking. The thresholds found for the ringing correspond well with the few data that exist in the literature for the same conditions. In addition, the peak group delay at threshold corresponds well with existing knowledge of the temporal resolution of our hearing. For frequencies below 1 kHz, the data are well complemented by literature data, which suggest that the threshold is given by a constant Q of around 2 to 3.

The experiments were carried out with an artificial signal, an impulse, but sound signals that are highly impulsive do occur in real life, such as from percussion, and the thresholds may also apply to such critical real-life signals. However, for most real-life signals the ringing is more likely to be masked by the signal itself, and it is expected to be inaudible, even if Q is higher than the thresholds found. Reverberation in the recording or during loudspeaker playback may also impair detection.

The thresholds were obtained with isolated all-pass sections, and they may not apply to transfer functions where the all-pass phase is accompanied by a minimum-phase component. As an example, if an all-pass component stems from zeros in the right half of the complex s plane that do not have corresponding poles in the left half-plane, the magnitude of the transfer function has a local dip at the center frequency. Consequently the amplitude of the ringing is significantly reduced, and the thresholds do not apply.

If an all-pass section is applied in the signal path to only one ear, it may give rise to a lateral shift of the auditory image. The lateralization seems related to the fact that, at low frequencies, the all-pass section acts as a delay, that is, its phase and group delays are approximately constant and have the same value. The lower Q is, the longer is the delay, and the more audible is the lateralization. The effect is nearly the same for different broad-band signals. For center frequencies in the range of 1–12 kHz thresholds for Q correspond to low-frequency phase and group delays of approximately 35 μs, independent of the center frequency. The literature does not offer data for the same conditions, but the value agrees well with literature data on minimum audible interaural delay. The result also agrees with our knowledge about the localization blur for frontal sound sources in the horizontal plane, for which the interaural time difference is believed to be the main cue for discrimination.

In the experiments only one all-pass section was introduced at a time. Real-life transfer functions may include more than one all-pass section. For the ringing it is uncertain what the joint effect of several all-pass sections will be. The total peak group delay seems to play an important role, possibly together with the group-delay bandwidth, but more studies are needed. For the lateralization it seems reasonable to conclude that the joint effect of several all-pass sections can be evaluated from the difference between the two sides of the total low-frequency phase or group delays.

5 ACKNOWLEDGMENT

This work was supported in part by the Danish National Center for IT-Research, the Danish Technical Research Council, and the Danish Research Council for Technology and Production Sciences. The authors wish to thank Claus Vestergaard Skipper for his assistance in the laboratory as well as the patient listeners who participated in the listening experiment. They are also grateful to two anonymous reviewers for valuable comments and suggestions.

6 REFERENCES

[1] A. V. Oppenheim and R. W. Shafer, *Discrete Time Signal Processing* (Prentice Hall International, Englewood Cliffs, NJ, 1989).

[2] H. L. F. Helmholtz, *Die Lehre von den Tonempfindungen,* 4th German ed. (1877); A. J. Ellis, *On the Sensation of Tone,* 2nd English transl. (1885) (Dover, New York, 1954).

[3] R. C. Mathes and R. L. Miller, "Phase Effects in Monaural Perception," *J. Acoust. Soc. Am.,* vol. 19, pp. 780–797 (1947 Sept.).

[4] J. C. R. Licklider, "Effects of Changes in the Phase Pattern upon the Sound of a 16-Harmonic Tone," *J. Acoust. Soc. Am.,* vol. 29, p. 780 (Abstract) (1957 June).

[5] M. R. Schroeder, "New Results Concerning Monaural Phase," *J. Acoust. Soc. Am.,* vol. 31, p. 1579 (Abstract) (1959 Nov.).

[6] J. H. Craig and L. A. Jeffress, "Effect of Phase on the Quality of a Two-Component Tone," *J. Acoust. Soc. Am.,* vol. 34, pp. 1752–1760 (1962 Nov.).

[7] E. Zwicker and R. Feldtkeller, *Das Ohr als Nachrichtenempfänger,* 2nd rev. ed. (1967); H. Müsch, S. Buus, and M. Florentine, *The Ear as a Communication Receiver,* English transl. (American Institute of Physics, 1999).

[8] F. A. Bilsen, "Phase Sensitivity and (or?) Short Time Analysis of the Hearing Organ," *Acustica,* vol. 18, no. 3, p. 182 (1967).

[9] R. Plomp and H. J. M. Steeneken, "Effect of Phase on the Timbre of Complex Tones," *J. Acoust. Soc. Am.,* vol. 46, pp. 409–421 (1969 Aug.).

[10] J. C. Nixon, C. A. Raiford, and E. D. Schubert, "Technique for Investigating Monaural Phase Effects," *J. Acoust. Soc. Am.,* vol. 48, pp. 554–556 (1970 Aug.).

[11] H. Duifhuis, "Audibility of High Harmonics in a Periodic Pulse," *J. Acoust. Soc. Am.,* vol. 48, pp. 888–893 (1970 Oct.).

[12] C. A. Raiford and E. D. Schubert, "Recognition of Phase Changes in Octave Complexes," *J. Acoust. Soc. Am.,* vol. 50, pp. 559–567 (1971 Aug.).

[13] T. J. F. Buunen, J. M. Festen, F. A. Bilsen, and G. van den Brink, "Phase Effects in a Three-Component Signal," *J. Acoust. Soc. Am.,* vol. 55, pp. 297–303 (1974 Feb.).

[14] E. B. Jensen and H. Møller, "On the Audibility of Phase Distortion in Audio Systems," presented at the 47th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts),* vol. 22, p. 270 (1974 May), preprint A-2.

[15] V. Hansen and E. R. Madsen, "On Aural Phase Detection: Part II," *J. Audio Eng. Soc.,* vol. 22, pp. 783–788 (1974 Dec.).

[16] P. J. J. Lamoré, "Perception of Two-Tone Octave Complexes," *Acustica,* vol. 34, no. 1, pp. 3–14 (1975).

[17] R. C. Cabot, M. G. Mino, D. A. Dorans, I. S. Tackel, and H. E. Breed, "Detection of Phase Shifts in Harmonically Related Tones," *J. Audio Eng. Soc.,* vol. 24, pp. 568–571 (1976 Sept.).

[18] R. Berkovitz and B. E. Edvardsen, "Phase Sensitivity in Music Reproduction," presented at the 58th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts),* vol. 25, p. 1070 (1977 Dec.), preprint 1294.

[19] M. Kubovy and R. Jordan, "Tone-Segregation by Phase: On the Phase Sensitivity of the Single Ear," *J. Acoust. Soc. Am.,* vol. 66, pp. 100–106 (1979 July).

[20] R. Carlson, B. Granström, and D. Klatt, "Vowel Perception: The Relative Perceptual Salience of Selected Acoustic Manipulations," Speech Trans. Lab. Q. Prog. Stat. Rep. (Trita-TLF-79-4), Stockholm, Sweden (1979), pp. 73–83.

[21] S. P. Lipshitz, M. Pocock, and J. Vanderkooy, "On the Audibility of Midrange Phase Distortion in Audio Systems," *J. Audio Eng. Soc.,* vol. 30, pp. 580–595 (1982 Sept.).

[22] J. E. Daniel and M. Kubovy, "Detecting Spectral and Temporal Target Differences in the Harmonic Complex," *J. Acoust. Soc. Am.,* vol. 79, suppl. 1, p. S79 (Abstract) (1986 Spring).

[23] R. D. Patterson, "A Pulse Ribbon Model of the Monaural Phase Perception," *J. Acoust. Soc. Am.,* vol. 82, pp. 1560–1586 (1987 Nov.).

[24] B. C. J. Moore and B. R. Glasberg, "Difference Limens for Phase in Normal and Hearing-Impaired Subjects," *J. Acoust. Soc. Am.,* vol. 86, pp. 1351–1365 (1989 Oct.).

[25] A. Galembo, A. Askenfelt, L. L. Cuddy, and F. A. Russo, "Effects of Relative Phases on Pitch and Timbre in the Piano Bass Range," *J. Acoust. Soc. Am.,* vol. 110, pp. 1649–1666 (2001 Sept.).

[26] J. L. Goldstein, "Auditory Spectral Filtering and Monaural Phase Perception," *J. Acoust. Soc. Am.,* vol. 41, pp. 458–479 (1967 Feb.).

[27] T. J. F. Buunen, "Two Hypotheses on Monaural Phase Effects," *Acustica,* vol. 34, no. 2, pp. 98–105 (1975 Dec.).

[28] H. Fleischer, "Gerade wahrnehmbare Phasenänderungen bei Drei-Ton-Komplexen," *Acustica,* vol. 32, no. 1, pp. 44–50 (1975 Jan.).

[29] H. Fleischer, "Hörbarkeit von Phasenunterschieden bei verschiedenen Arten der Schalldarbietung," *Acustica,* vol. 36, no. 2, pp. 90–99 (1976).

[30] V. Hansen and E. R. Madsen, "On Aural Phase Detection," *J. Audio Eng. Soc.,* vol. 22, pp. 10–14 (1974 Jan./Feb.).

[31] R. A. Greiner and D. E. Melton, "Observations on the Audibility of Acoustic Polarity," *J. Audio Eng. Soc.,* vol. 42, pp. 245–253 (1994 Apr.).

[32] D. A. Ronken, "Monaural Detection of a Phase Difference between Clicks," *J. Acoust. Soc. Am.,* vol. 47, pp. 1091–1099 (1969 Oct.).

[33] D. Preis and P. J. Bloom, "Perception of Phase Distortion in Anti-Alias Filters," *J. Audio Eng. Soc.,* vol. 32, pp. 842–848 (1984 Nov.).

[34] J. H. Patterson and D. M. Green, "Discrimination of Transient Signals Having Identical Energy Spectra," *J. Acoust. Soc. Am.,* vol. 48, pp. 894–905 (1970 Oct.).

[35] D. A. Huffman, "The Generation of Impulse-

Equivalent Pulse Trains," *IRE Trans. Inform. Technol.,* vol. IT-8, pp. S10-S16 (1962).

[36] D. M. Green, "Temporal Acuity as a Function of Frequency," *J. Acoust. Soc. Am.,* vol. 54, pp. 373–379 (1973 Aug.).

[37] J. Blauert and P. Laws, "Group Delay Distortions in Electroacoustical Systems," *J. Acoust. Soc. Am.,* vol. 63, pp. 1478–1483 (1978 May).

[38] H. Suzuki, S. Morita, and T. Shindo, "On the Perception of Phase Distortion," *J. Audio Eng. Soc.,* vol. 28, pp. 570–574 (1980 Sept.).

[39] R. Lee, "Is Linear Phase Worthwhile?," presented at the 68th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts),* vol. 29, p. 364 (1981 May), preprint 1732.

[40] L. R. Fincham, "The Subjective Importance of Uniform Group Delay at Low Frequencies," *J. Audio Eng. Soc.,* vol. 33, pp. 436–439 (1985 June).

[41] J. A. Deer, P. J. Bloom, and D. Preis, "Perception of Phase Distortion in All-Pass Filters," *J. Audio Eng. Soc.,* vol. 33, pp. 782–786 (1985 Oct.).

[42] M. Karjalainen, E. Piirilä, A. Järvinen, and J. Huopaniemi, "Comparison of Loudspeaker Equalization Methods Based on DSP Techniques," *J. Audio Eng. Soc.,* vol. 47, pp. 14–31 (1999 Jan./Feb.).

[43] S. Flanagan, B. C. J. Moore, and M. A. Stone, "Discrimination of Group Delay in Clicklike Signals Presented via Headphones and Loudspeakers," *J. Audio Eng. Soc.,* vol. 53, pp. 593–611 (2005 July/Aug.).

[44] P. Minnaar, J. Plogsties, S. K. Olesen, F. Christensen, and H. Moller, "The Audibility of All-Pass Components with Application to Binaural Synthesis," in P. Minnaar, "Simulating an Acoustical Environment with Binaural Technology," Ph. D. thesis, Aalborg University, Aalborg, Denmark (2001).

[45] B. C. J. Moore, *Introduction to the Psychology of Hearing,* 3rd ed. (Academic Press, London, 1997).

[46] E. Zwicker and H. Fastl, *Psychoacoustics—Facts and Models* (Springer, Berlin, 1990).

[47] R. G. Klump and H. R. Eady, "Some Measurements of Interaural Time Difference Thresholds," *J. Acoust. Soc. Am.,* vol. 28, pp. 859–860 (1956).

[48] J. L. Hall, "Minimum Detectable Change in Interaural Time and Intensity Difference for Brief Impulsive Stimuli," *J. Acoust. Soc. Am.,* vol. 36, pp. 2411–2413 (1964).

[49] W. A. Yost, F. L. Wightman, and D. M. Green, "Lateralization of Filtered Clicks," *J. Acoust. Soc. Am.,* vol. 50, pp. 1526–1531 (1971).

[50] E. R. Hafter and J. De Maio, "Difference Thresholds for Interaural Delay," *J. Acoust. Soc. Am.,* vol. 57, pp. 181–187 (1975).

[51] R. Klinke, "Physiology of Hearing," in *Fundamentals of Sensory Physiology,* 3rd ed., R. F. Schmidt, ed. (Springer, Berlin, Heidelberg, New York, Tokyo, 1986).

[52] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization,* original ed. *Räumliches Hören* (S. Hirzel Verlag, Stuttgart, 1974). New material by author and English transl. by J. S. Allen (MIT Press, Cambridge, MA, 1983).

[53] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen, "Head-Related Transfer Functions of Human Subjects," *J. Audio Eng. Soc.,* vol. 43, pp. 300–321 (1995 May).

[54] L. G. Johansen and P. Rubak, "The Excess Phase in Loudspeaker/Room Transfer Functions: Can It Be Ignored in Equalization Tasks?," presented at the 100th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts),* vol. 44, p. 636 (1996 July/Aug.), preprint 4181.

[55] M. E. Valkenburg, *Analog Filter Design* (CBS College Pub., New York, 1982).

APPENDIX 1
ON PHASE AND DELAY

The phase shift of a transfer function is sometimes interpreted as being equivalent to a shift in time of a signal. Depending on the signal and the course of the phase along frequency, this time shift can materialize in different ways, and it is important to distinguish between these. Specifically, the concepts of delay, phase delay, and group delay are reviewed in this appendix.

A1.1 Delay

If the output of a system is identical to the input, except that it occurs later, the system performs a delay. The transfer function of a delay is

$$H(s) = e^{-s\tau_{\mathrm{d}}} \tag{14}$$

where τ_{d} is the delay. For a given frequency ω, $s = \mathrm{j}\omega$, and the phase is

$$\phi(\omega) = -\omega\tau_{\mathrm{d}}. \tag{15}$$

It is seen that the phase is a linear function of ω. Therefore the term "linear phase" is often used to characterize the transfer function of a delay. Note, however, that the phase of a delay is not only linear; it is zero at a frequency of zero. There are transfer functions where the phase is linear but not 0° at 0 Hz; for example, the phase of a polarity inversion is a straight line at −180°, and this is clearly not equivalent to a delay.

A1.2 Phase Delay

Consider a sinusoid with frequency ω,

$$e_{\mathrm{in}}(t) = B\sin(\omega t). \tag{16}$$

If this signal is sent through a system with a transfer function with magnitude $A(\omega)$ and phase $\phi(\omega)$, the output is given by

$$e_{\mathrm{out}}(t) = A(\omega)B\sin[\omega t + \phi(\omega)]. \tag{17}$$

If

$$\tau'(\omega) = \frac{\phi(\omega)}{\omega} \tag{18}$$

Eq. (17) can be rewritten as

$$e_{\text{out}}(t) = A(\omega)B \sin\{\omega[t + \tau'(\omega)]\}. \tag{19}$$

This shows that the phase shift is equivalent to a shift in time given by Eq. (18). $\tau'(\omega)$ can take both positive and negative values, where a positive value indicates an increase in the argument of the sinusoidal function. A positive value thus means that a particular point in the signal is reached earlier, implying that the output is ahead of the input. In real life, however, the output always comes at the same time as, or later than the input. Therefore the time shift is rather considered as a delay equal to the negative of $\tau'(\omega)$,

$$\tau_{\text{p}}(\omega) = -\tau'(\omega) = -\frac{\phi(\omega)}{\omega}. \tag{20}$$

The term $\tau_{\text{p}}(\omega)$ is referred to as the *phase delay* of the transfer function. Note that it describes the time shift to which each individual frequency is subjected, and it should not be interpreted as a delay of the signal as a whole.

A1.3 Group Delay

The term *group delay* has its origin in high-frequency radio transmission. Here an information signal is used to modulate, for example, the amplitude of a carrier signal that has a much higher frequency. The modulated signal has only frequency components in a narrow region around the carrier frequency. In such a narrow frequency band the magnitude of most transfer functions is nearly constant, and the phase can be approximated by a straight line. When the modulated signal is sent through the system, its envelope is shifted in time by an amount that is referred to as the group delay.

The value of the group delay can be determined by considering the ideal case of two sinusoidal signals of equal amplitude B and frequencies ω_1 and ω_2 that are close to each other. If the two tones are in phase at $t = 0$,

$$e_{\text{in}}(t) = B \sin(\omega_1 t) + B \sin(\omega_2 t). \tag{21}$$

This can be rewritten as

$$e_{\text{in}}(t) = 2B \cos(\omega_{\text{i}} t) \sin(\omega_{\text{c}} t) \tag{22}$$

where $\omega_{\text{i}} = (\omega_1 - \omega_2)/2$ and $\omega_{\text{c}} = (\omega_1 + \omega_2)/2$. In high-frequency radio transmission this is considered as a high-frequency carrier signal with frequency ω_{c} that is amplitude modulated with a low-frequency information signal with a frequency ω_{i}. The amplitude modulation given by $2B \cos(\omega_i t)$ determines the envelope of the carrier signal.

Now consider the signal given in Eq. (21) being sent through a system that has a transfer function with magnitude A and phase $\phi(\omega_1)$ and $\phi(\omega_2)$ at the frequencies ω_1 and ω_2, respectively. Then the output of the system can be calculated as

$$\begin{aligned} e_{\text{out}} &= AB \sin[\omega_1 t + \phi(\omega_1)] + AB \sin[\omega_2 t + \phi(\omega_2)] \\ &= 2AB \cos\left[\frac{\omega_1 - \omega_2}{2} t + \frac{\phi(\omega_1) - \phi(\omega_2)}{2}\right] \\ &\quad \times \sin\left[\frac{\omega_1 + \omega_2}{2} t + \frac{\phi(\omega_1) + \phi(\omega_2)}{2}\right] \\ &= 2AB \cos\left\{\omega_{\text{i}}\left[t + \frac{\phi(\omega_1) - \phi(\omega_2)}{\omega_1 - \omega_2}\right]\right\} \\ &\quad \times \sin\left\{\omega_{\text{c}}\left[t + \frac{\phi(\omega_1) + \phi(\omega_2)}{\omega_1 + \omega_2}\right]\right\}. \end{aligned} \tag{23}$$

Notice that the mathematical form of the output is the same as that of the input [Eq. (22)], but that the envelope is now given by

$$2AB \cos\left\{\omega_{\text{i}}\left[t + \frac{\phi(\omega_1) - \phi(\omega_2)}{\omega_1 - \omega_2}\right]\right\}. \tag{24}$$

Thus the envelope has shifted in time by an amount equal to

$$\tau'' = \frac{\phi(\omega_1) - \phi(\omega_2)}{\omega_1 - \omega_2}. \tag{25}$$

Since a positive value of this indicates that the output is earlier than the input (as was the case with τ'), it is preferable to consider the time shift as a delay equal to the negative of τ''. Furthermore it is seen that τ'' goes toward the derivative of the phase with respect to the frequency in the limit where the two frequencies ω_1 and ω_2 approach each other. The group delay at any frequency ω is therefore defined as

$$\tau_{\text{g}}(\omega) = -\frac{\partial\phi(\omega)}{\partial\omega}. \tag{26}$$

This definition can be used with any transfer function and is not restricted to high-frequency transmission. However, it should be noted that the interpretation of group delay as a delay only applies to the envelope of signals with frequency content in a narrow frequency band. As with the phase delay, the group delay should not be considered as a delay of the signal as a whole.

Other terms that are sometimes used for the group delay are *envelope delay* and *energy delay*. In addition, the term *signal delay* has been observed in the literature. While the first two alternatives are rational, the latter might seem misleading.

It follows from applying Eqs. (20) and (26) to Eq. (15) that the phase and group delays of a delay are independent of frequency and have the same value equal to the value of the delay, that is, $\tau_{\text{p}} = \tau_{\text{g}} = \tau_{\text{d}}$.

It should be noted that while the phase and group delays of a transfer function can vary with frequency, there is only one single value of a delay. Therefore it is not correct to say that a signal has been exposed to a frequency-dependent delay, but fine to speak of frequency-dependent phase and group delays.

A1.4 Examples

In order to illustrate some of the concepts discussed in Sections A1.1 to A1.3, some small examples are presented, where different signals are transmitted through a system. The phase, phase delay, and group delay of the system's transfer function are shown in Fig. 2. (These data belong to a certain second-order all-pass section. At this stage this is not important, and it is possible to follow the examples by only referring to the figure. A detailed description of second-order all-pass sections and their properties follows in Appendix 2.)

In Fig 2(a) it is seen that the phase changes from 0 to -2π as the frequency increases from zero to above 20 kHz. Although it is not clearly seen with the logarithmic frequency scale of the figure, the phase is an approximately linear function of frequency at low frequencies (up to about 1–2 kHz). The phase goes through zero at a frequency of zero. In Fig. 2(b) it is seen that the phase delay is nearly constant at low frequencies. It has a maximum around 5 kHz. The group delay shown in Fig. 2(c) is also constant at low frequencies. It has a more pronounced maximum around 4 kHz. Furthermore it can be seen that at low frequencies the values of the phase delay and the group delay are virtually the same.

Now consider three different signals to be transmitted through the system. Signal A is a pure tone with a frequency of 1 kHz, weighted by a Hanning window (see, for example, [1]). Signal B also consists of a pure tone weighted by a Hanning window, but the frequency of the tone is 4 kHz. Signal C is the sum of signals A and B. Fig. 10 shows these three input signals (grey lines) as well as their corresponding output signals (black lines).

In Fig. 10(a) it is seen that for signal A the input and output signals are practically identical, but the output is slightly later than the input. A zoom at the signals will reveal that the delay is around 45 μs. Signal A has its frequency content in a narrow band centered around 1 kHz, thus the whole signal lies in the frequency range where the phase is linear and goes through zero at a frequency of zero. Consequently, to this signal the transfer function acts like a delay. In Fig. 2(b) and (c) it is seen that both the phase delay and the group delay have values of about 45 μs.

For signal B [Fig. 10(b)] the envelope of the signal has been delayed by slightly more than 300 μs. The frequency content of this signal is confined to a narrow range around 4 kHz, and as described in Section A1.3, such signals will have their envelope delayed by the group delay. As seen in Fig. 2(c), the group delay at 4 kHz is just 320 μs. From Fig. 2(a) it is seen that the phase at 4 kHz is equal to $-\pi$, which implies that the signal component at this frequency should be exactly out of phase. This is also seen to be the case in Fig. 10(b). It can also be expressed as a 125-μs delay of the carrier, corresponding to the phase delay at 4 kHz.

For signal C, shown in Fig. 10(c), the input and output differ a lot, and it is not possible to identify the output as a delayed version of the input. This is often explained by saying that the different frequency components have been subjected to different delays, or that the signal has been subjected to a frequency-dependent delay. As mentioned, it would be more correct to say that the different frequency components have been subjected to different phase delays and group delays.

APPENDIX 2
ALL-PASS SECTIONS

As is understood from Eq. (1), the all-pass component carries the phase information, which neither is connected to the magnitude of the transfer function, nor is a simple delay. The all-pass component can be divided into first- and second-order sections. In the complex s plane, all-pass sections have zeros in the right half-plane and poles at corresponding mirror-image positions in the left half-plane. In the complex z plane all-pass sections have zeros outside the unit circle and poles at the reciprocal positions inside the unit circle. It is the mirror-image positions, or the reciprocal positions, that guarantee the flat magnitude. For first-order sections, poles and zeros are on the real

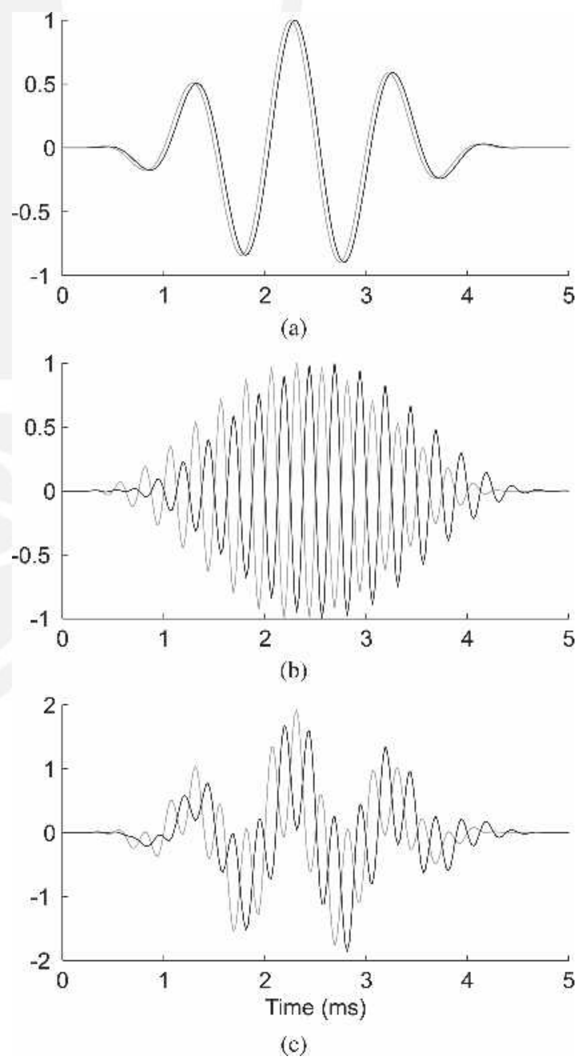

Fig. 10. Examples of three signals filtered by transfer function of Fig. 2. Original signals are shown in gray, those after filtering in black. (a) Signal A; $f_c = 1$ kHz. (b) Signal B; $f_c = 4$ kHz. (c) Signal C; signal A + signal B.

axis, while for second-order sections poles and zeros appear in complex conjugate pairs. Examples are given in Fig. 1.

A2.1 First-Order All-Pass Section

A first-order all-pass section is represented in the s plane by a single zero on the positive real axis and a pole at the same position on the negative real axis.

The transfer function can be written as (see [55])

$$H(s)_{\mathrm{ap1}} = \frac{\omega_0 - s}{\omega_0 + s}. \tag{27}$$

A2.1.1 Phase, Phase Delay, and Group Delay

The phase is found to be given by

$$\phi(\omega)_{\mathrm{ap1}} = -2\tan^{-1}\left(\frac{\omega}{\omega_0}\right). \tag{28}$$

This function is zero at a frequency of zero, and as the frequency increases, it changes monotonically from zero to $-\pi$.

By using the definition of phase delay [Eq. (20)], this is seen to be given by

$$\tau_{\mathrm{p}}(\omega)_{\mathrm{ap1}} = \frac{2\tan^{-1}(\omega/\omega_0)}{\omega}. \tag{29}$$

From the definition of group delay [Eq. (26)] this is found as

$$\tau_{\mathrm{g}}(\omega)_{\mathrm{ap1}} = \frac{2/\omega_0}{1 + (\omega/\omega_0)^2}. \tag{30}$$

A2.1.2 Impulse Response

The transfer function given in Eq. (27) can be rewritten as

$$H(s)_{\mathrm{ap1}} = -1 + \frac{2\omega_0}{s + \omega_0}. \tag{31}$$

From this equation the impulse response can be determined by means of the inverse Laplace transform,

$$h(t)_{\mathrm{ap1}} = -\delta(t) + 2\omega_0 \mathrm{e}^{-\omega_0 t}. \tag{32}$$

It is seen that the impulse response of a first-order all-pass section consists of an impulse (delta function) and an exponential decay. The time constant in the decay is $1/\omega_0$.

A2.2 Second-Order All-Pass Section

In the s plane a second-order all-pass section is represented by a complex conjugated pair of zeros with positive real parts and poles with the same values but with negative real parts.

The transfer function is given by (see [55])

$$H(s)_{\mathrm{ap2}} = \frac{s^2 - (\omega_0/Q)\,s + \omega_0^{\,2}}{s^2 + (\omega_0/Q)\,s + \omega_0^{\,2}} \tag{33}$$

where Q is the quality factor, or Q factor, and ω_0 (or $f_0 = \omega_0/2\pi$) the center frequency.

A2.2.1 Phase, Phase Delay, and Group Delay

The phase is given by

$$\phi(\omega)_{\mathrm{ap2}} = -2\tan^{-1}\left\{\frac{\omega/\omega_0}{Q[1-(\omega/\omega_0)^2]}\right\}. \tag{34}$$

Like for a first-order all-pass section the phase is zero at a frequency of zero. As the frequency increases, the phase changes monotonically to -2π. It passes $-\pi$ at the center frequency ($f_0 = \omega_0/2\pi$). An example of the phase of a second-order all-pass section is seen in Fig. 2(a).

Using its definition [Eq. (20)], the phase delay can be found as

$$\tau_{\mathrm{p}}(\omega)_{\mathrm{ap2}} = \frac{2\tan^{-1}\left\{\dfrac{\omega/\omega_0}{Q[1-(\omega/\omega_0)^2]}\right\}}{\omega}. \tag{35}$$

The group delay is found as [using its definition in Eq. (26)]

$$\tau_{\mathrm{g}}(\omega)_{ap2} = \frac{2}{\omega_0 Q}\left\{\frac{1 + (\omega/\omega_0)^2}{[1-(\omega/\omega_0)^2]^2 + (\omega/\omega_0 Q)^2}\right\}. \tag{36}$$

Examples of phase delay and group delay of a second-order all-pass section are given in Fig. 2(b) and (c). It is seen from the figure that both phase and group delays have peaks around the center frequency. As already noted in Section A1.4 the phase and group delays are approximately constant at low frequencies, and they seem to have about the same value. In fact, as seen in the following, they approach the same mathematical term when the frequency approaches zero.

For the phase delay the asymptotic value is found as [using Eq. (35)]

$$\underset{\lim \omega\to 0}{\tau_{\mathrm{p}}(\omega)_{\mathrm{ap2}}} = \frac{2\left\{\dfrac{\omega/\omega_0}{Q[1-(\omega/\omega_0)^2]}\right\}}{\omega} = \frac{2}{\omega_0 Q}, \tag{37}$$

and the asymptotic value for the group delay is [using Eq. (36)]

$$\underset{\lim \omega\to 0}{\tau_{\mathrm{g}}(\omega)_{\mathrm{ap2}}} = \frac{2}{\omega_0 Q}. \tag{38}$$

Since the phase and group delays approach the same value and are nearly constant at low frequencies, the all-pass section acts as a delay in this frequency region. The value of the delay is given by any one of Eqs. (37) and (38).

It can be shown that if Q is high ($Q^2 \gg 1/4$), the peak of the group delay occurs at approximately the center frequency. Thus the peak group delay can be found by evaluating Eq. (36) at the center frequency, which gives

$$\tau_{\mathrm{g\ peak}} \approx \tau_{\mathrm{g}}(\omega_0)_{\mathrm{ap2}} = \frac{4Q}{\omega_0}. \tag{39}$$

The width of the frequency range around the center frequency, where the group delay is more than half of its peak value, can be defined as the group-delay bandwidth $\mathrm{BW_g}$. It can be shown that for high Q, this is given by

$$\mathrm{BW_g} \approx \frac{\omega_0}{Q}. \tag{40}$$

In order to illustrate the influence of the variables ω_0 and Q on the phase and group delays, examples of all-pass sections with different combinations of low and high center frequencies (3 and 10 kHz) and of low and high Q (1 and 10) are shown in Fig. 11. It is seen that when Q is high [Fig. 11(a) and (b)] the peak in the group delay is very high. The peak in the phase delay is also high, although this is less pronounced. The fact that the phase and group delays are the same at low frequencies can be seen in all frames. It is also seen that the low-frequency phase and group delays are relatively large when Q is low [Fig. 11(c) and (d)]. The phase and group delays are also larger at low center frequencies [Fig. 11(a) and (c)] than at high center frequencies [Fig. 11(b) and (d)]. These observations agree with Eqs. (37), (38), and (39).

A2.2.2 Impulse Response

The impulse response of a second-order all-pass section can be derived by rewriting Eq. (33) as

$$H(s)_{\mathrm{ap2}} = 1 - \frac{2\omega_0}{Q}\left[\frac{s+1/\tau}{(s+1/\tau)^2+\omega_1{}^2} - \frac{1}{\omega_1\tau}\frac{\omega_1}{(s+1/\tau)^2+\omega_1{}^2}\right] \tag{41}$$

where

$$\omega_1 = \omega_0\sqrt{1-\left(\frac{1}{2Q}\right)^2} \tag{42}$$

and

$$\tau = \frac{2Q}{\omega_0}. \tag{43}$$

By means of the inverse Laplace transform, the impulse response is found as

$$h(t)_{\mathrm{ap2}} = \delta(t) - \frac{2\omega_0}{Q}\,\mathrm{e}^{-t/\tau}\left[\cos(\omega_1 t) - \frac{1}{\omega_1\tau}\sin(\omega_1 t)\right]. \tag{44}$$

This can be rewritten as

$$h(t)_{\mathrm{ap2}} = \delta(t) + A\mathrm{e}^{-t/\tau}\cos(\omega_1 t + \theta) \tag{45}$$

where

$$A = -\frac{2\omega_0}{Q}\frac{1}{\sqrt{1-(1/2Q)^2}} \tag{46}$$

and

$$\theta = \tan^{-1}\left[\frac{1}{\sqrt{(2Q)^2-1}}\right]. \tag{47}$$

From Eq. (45) it is seen that the impulse response consists of a single impulse (delta function) and an exponentially decaying sinusoid. The frequency ω_1 of the sinusoid is given by Eq. (42), and the decay is controlled by the time constant τ, as given in Eq. (43). It is seen that at high Q ($Q^2 \gg 1/4$) the frequency of the decaying sinusoid is nearly equal to the center frequency, and the number N_τ of cycles within a time constant can be found as

$$N_\tau \approx \frac{Q}{\pi}. \tag{48}$$

Also at high Q the initial amplitude A of the decaying sinusoid is given by

$$A \approx -\frac{2\omega_0}{Q}. \tag{49}$$

Fig. 3 shows impulse responses for the four examples from Fig. 11. In the high-Q examples [Fig. 3(a) and (b)] it is seen that the impulse response "rings" after the initial impulse. The ringing is not as obvious in the low-Q examples [Fig. 3(c) and (d)] since the decay is shorter and

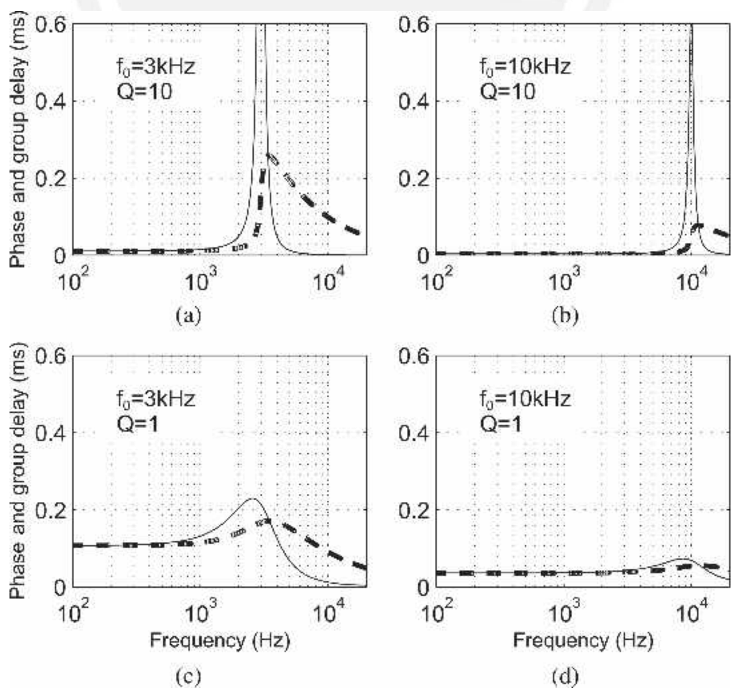

Fig. 11. Examples of phase delay (- - -) and group delay (——) for second-order all-pass sections for selected center frequencies and Q factors.

takes place over just a few oscillations of the sinusoid. On the other hand, the initial amplitude of the sinusoid is higher for the low-Q examples than for the high-Q. The lower Q is, the higher the amplitude and the shorter the decay of the sinusoid.

A comparison between the left and right frames of Fig. 3 reveals that the initial amplitude of the sinusoid is higher and the decay shorter with the high center frequencies than with the low. All these observations are in agreement with Eqs. (43) and (46) or (49). At very high Q values the ringing will be very long [Eq. (43)], but its amplitude very low [Eq. (49)]. Thus the impulse response approaches a delta function.

A comparison of Eqs. (43) and (39) shows that for high Q there is a simple relation between the time constant for the decaying sinusoid and the peak group delay,

$$\tau_{\mathrm{g\ peak}} = 2\tau. \tag{50}$$

THE AUTHORS

H. Møller

P. Minnaar

S. K. Olesen

F. Christensen

J. Plogsties

Henrik Møller was born in Århus, Denmark, in 1951. He studied electrical engineering and received a B.Sc. degree from the Danish Engineering Academy in 1974. He received a Ph.D. degree from Aalborg University, Denmark, in 1984.

From 1974 to 1976 he worked as a development engineer for Brüel & Kjær. Since then he has been working at Aalborg University. He became associate professor in 1980 and was appointed reader (docent) in 1988 and professor in 1996. From 1991 to 1994 he was partly on leave from the University to work as director for Perceptive Acoustics A/S, a research subsidiary company of Brüel & Kjær.

Dr. Møller's previous and current research reflect his long-time experience with sound, its influence on humans, acoustical measurement techniques, signal processing, hearing, and psychometric methods. His research areas include effects of infrasound and low-frequency noise on humans, investigations of hearing thresholds and loudness assessment, and exploitation of binaural techniques. He was responsible for the design and control of the work for the high-quality acoustical laboratories built at Aalborg University in 1987. Since then the activities of this section have increased and he is now manager of research and education in a wide range of areas such as human sound perception, audiology, psychometry, electroacoustics, recording and playback techniques, auralization in acoustic room modeling and virtual reality, acoustical measurement techniques, electronics, and signal processing.

Dr. Møller has organized conferences on Low Frequency Noise and Hearing (Aalborg 1980), general acoustics (Nordic Acoustical Meeting, Aalborg 1986), and Low Frequency Noise and Vibration (Aalborg 2000). He has also been deeply involved with standardization (ISO Technical Committee 43: *Acoustics*, and several of its working groups as member/convener). He is a member of the editorial board of the *Journal of Low Frequency Noise & Vibration* and serves as reviewer for a number of international journals and national and international research foundations. He is the author of numerous scientific publications and has written invited papers and conference papers. He holds several patents. Recent business relationships comprise the establishment of AM3D A/S and membership of its Business Board.

He is a member of the Danish Engineering Society, Audio Engineering Society, Acoustical Society of

America, IEEE, Danish Acoustical Society, Danish Technical-Audiological Society, Danish Society for Applied Signal Processing, and Danish Standards (Board of Acoustics and Working Group of Audiometry and Hearing).

●

Pauli Minnaar received a B.Eng. degree in mechanical engineering from the University of Pretoria in South Africa. He studied at the Institute of Sound and Vibration Research (ISVR), University of Southampton, in England, where he obtained an M.Sc. degree in acoustics. He received a Ph.D. degree in binaural technology from the Department of Acoustics, Aalborg University, Denmark, in 2001. He then worked as assistant professor doing research in binaural technology and sound quality at Aalborg University. He is currently working as a senior research engineer at AM3D in Denmark.

●

Søren Krarup Olesen was born in Denmark in 1966. He received an M.Sc.E.E. degree with a specialization in acoustics from Aalborg University, Denmark, in 1991. He began his Ph.D. study in the Department of Acoustics, Aalborg University, in 1994, then worked as technical amanuensis and research assistant in a project concerning 3-D audio in virtual rooms, and received a Ph.D. degree in 2001.

He is currently an associate professor at Aalborg University. His areas of interest are mainly binaural synthesis, design of computer-controlled psychoacoustical listening experiments, and acoustical room simulation and modeling systems for virtual reality purposes. He is a keen programmer and has held external courses in digital signal processing and C++ programming.

Mr. Krarup Olesen is a member of the Danish Acoustical Society and the Danish Engineering Society.

●

Flemming Christensen was born in Horsens, Denmark, in 1969. From 1988 to 1993 he studied electrical engineering and acoustics at Aalborg University in Denmark. He received an M.Sc.E.E. degree with a specialization in acoustics in 1993 and a Ph.D. degree in 2001. He is currently an associate professor. Since receiving his master's degree, he has been working in different research fields such as audiological measurements, directivity of hearing aids, binaural recording and playback, binaural synthesis, and sound quality.

●

Jan Plogsties was born and lives in Germany. He studied electrical engineering at the Dresden University of Technology. He received an M.Sc. degree from the Aalborg University in Denmark in 1998. There he worked as a research assistant in the Acoustics department on binaural sound perception and reproduction. Since 2001 he has been employed in the audio and multimedia department of the Fraunhofer Institute for Integrated Circuits IIS in Erlangen, Germany. He has worked on wave-field synthesis and MPEG-4 systems. His current research areas are audio coding and binaural technology. He also works as an acoustic consultant for theaters.