The Differential Time-Delay Distortion and Differential Phase-Shift Distortion as Measures of Phase Linearity*

W. MARSHALL LEACH, JR.

Georgia Institute of Technology, School of Electrical Engineering, Atlanta, GA 30332-0250, USA

The relationships between phase shift, phase delay, and group delay of a system are reviewed. The differential time-delay distortion and differential phase-shift distortion are defined. A rationale is presented for the use of these as indicators of phase linearity of audio amplifier circuits. The differential time-delay and phase-shift distortions are calculated for several low-pass and high-pass filter functions. The calculations are used to illustrate the use of the differential phase-shift distortion as a criterion for calculating minimum amplifier bandwidths for a specified phase linearity.

0 INTRODUCTION

There have been many conflicting and confusing statements concerning the relationship between the phase response of an amplifier and its bandwidth. An incorrect but commonly quoted criterion for acceptable high-frequency phase response is that the bandwidth of an amplifier must be approximately 10 times the signal bandwidth. Not only have advertisements propagated this misunderstanding, but it has also been quoted in refereed journals. Apparently the confusion has been a result of the misinterpretation of phase shift as phase distortion. In reality, phase distortion can be defined only after time delay has been accounted for. A phase shift that is caused by a time delay that is uniform for all frequencies in the signal bandwidth is not phase distortion.

This paper presents an investigation of the relationships between amplifier bandwidth and phase response. A differential time-delay distortion is defined, which is the difference between the phase delay and the group delay. This is then used to define a differential phase-shift distortion. Bandwidth requirements for a specified maximum differential phase-shift distortion are established by examining the response of familiar first- and second-order low-pass and high-pass filter functions. The criterion used to establish the bandwidth requirements is that the differential phase-shift distortion through the amplifier be less than or equal to 5° at the signal bandwidth extremes of 20 Hz and 20 kHz. There is no analytical reason for the use of the 5° figure in establishing the bandwidth criteria. It is used only for purposes of illustrating the theory. It is straightforward to repeat the calculations for any other value of phase shift.

* Presented at the 61st Convention of the Audio Engineering Society, New York, 1978 November 3–6; revised 1989 July 7.

1 DIFFERENTIAL TIME-DELAY AND PHASE-SHIFT DISTORTIONS

Two different delay functions are commonly used to specify the phase linearity of systems, the phase delay and the group delay. The phase delay is defined as the apparent time delay of a steady-state sinusoidal signal as it passes through the system. Ideally, it should be a constant at every frequency in the signal bandwidth. The group delay is defined as the apparent time delay of an amplitude envelope variation modulated on a sinusoidal input signal. Ideally, it also should be a constant at every frequency in the signal bandwidth. Although the difference between these two delay functions is intimately related to the phase linearity of a system, the literature normally treats them as separate delay functions.

To establish a meaningful criterion for the phase linearity of a system that is designed to process audio signals, it is first necessary to examine the nature of the acoustic signals that it will be required to reproduce. First, the signals are band-limited. (Electronic synthesizers could be considered an exception if one takes only the electrical signal output into account. However, when the electrical signal generated by a synthesizer is reproduced acoustically, the acoustic signal will be band-limited.) Not only are all sources of sound an acoustic bandpass filter, but also all microphones used to record these sounds and the recording media used to record and play them back are bandpass filters, not to mention the bandpass filter characteristics of any loudspeaker used to reproduce the sound and of any ear used to perceive it. Response to zero frequency is meaningless. The lower limit of human hearing is normally taken to be 20 Hz. An acceptable amplitude and phase response above this frequency is sufficient for low-frequency reproduction. The upper limit of human hearing is normally taken to be 20 kHz. An acceptable amplitude and phase response below this frequency is sufficient for high-frequency reproduction. Because reproduced signals outside the 20-Hz to 20-kHz band cannot be heard by the ear, the author believes that phase and amplitude response deviations outside this band are meaningless.

The nature of acoustic signals from both nonpercussive and percussive musical instruments has been studied and reported by Saponas, Matson, and Ashley [1]. A conclusion stated in their paper is that the sounds from all musical instruments "can be studied in terms of single pulse amplitude modulation of a single sinusoidal term. The harmonic structure of the carrier waveform can be handled by individually studying the Fourier components of the carrier signal." In other words, acoustic signals consist of a superposition of "packets" of nonoverlapping frequency components that vary dynamically in amplitude and frequency. Complex acoustic signals can be handled by studying each individual packet.

Given the fact that acoustic signals are band-limited and can be studied by individually considering narrow-band component packets, a criterion for acceptable phase response of any system used to reproduce these signals can be developed. A narrow-band signal can be described mathematically as a sinusoidal signal with an amplitude envelope variation modulated on it. The spectrum components in the envelope variation are much lower in frequency than the frequency of the sinusoid. In the case of acoustic signals of interest, the envelope variation is in the form of a pulse which has a duration that is long compared to the period of the sinusoid. Thus each packet is modeled as a pulse amplitude variation modulated on a sine wave. Such a signal is illustrated in Fig. 1.

The signal in Fig. 1 can be reproduced exactly by a system that has a flat frequency response over the bandwidth of the signal and a phase delay that is equal to the group delay, that is, the sinusoid should be delayed by an amount equal to the envelope delay. Thus the differential time-delay distortion of a system will be defined as the difference between the phase delay and the group delay. This is given by

$$\Delta\tau = \tau_\varphi - \tau_g \tag{1}$$

where τ_φ is the phase delay and τ_g is the group delay. These are given by [2]

$$\tau_\varphi = -\frac{1}{2\pi}\frac{\varphi}{f} \tag{2}$$

$$\tau_g = -\frac{1}{2\pi}\frac{d\varphi}{df} \tag{3}$$

where φ is the system phase response function and f is the frequency. The units of τ_φ and τ_g are seconds.

An ideal system should exhibit zero or negligible differential time-delay distortion. For $\Delta\tau$ in Eq. (1) to be zero, the following differential equation must hold:

$$\frac{d\varphi}{df} = \frac{\varphi}{f} . \tag{4}$$

This equation can be solved by direct integration to obtain a solution of the form $\ln(\varphi) = \ln(f) + k$, where k is the constant of integration. This constant will be taken to be $k = \ln(2\pi T)$. It is then straightforward to rewrite the solution as

$$\varphi = 2\pi f T . \tag{5}$$

This equation represents the phase response of a system that exhibits a constant time delay of T seconds at every frequency, that is, the system is a perfect delay line. It follows, therefore, that $\Delta\tau$ can be interpreted as a measure of the deviation of the time delay of a system from that of an ideal system with constant delay.

For a signal of the type illustrated in Fig. 1, the specification of a maximum acceptable differential time-delay distortion through a system is most easily done by specifying the maximum acceptable phase shift in degrees of the sinusoidal carrier with respect to the envelope at the system output. This is obtained as the

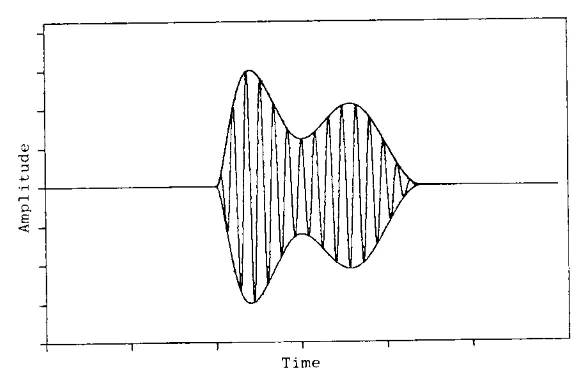

Fig. 1. Example narrow-band signal component modeled as sinusoidal signal with a amplitude envelope variation modulated on it.

product of the differential time-delay distortion in seconds times the number of cycles per second times the number of degrees per cycle. It is given by

$$\Delta\varphi = \Delta\tau \times f \times 360 = -\frac{180}{\pi}\left(\varphi - f\frac{d\varphi}{df}\right) \tag{6}$$

where the units are degrees. This represents the phase shift of the carrier with respect to its envelope for any signal component that passes through a system. The quantity $\Delta\varphi$ given by Eq. (6) will be defined as the differential phase-shift distortion. It will be used in the following to establish example bandwidth requirements for audio amplifiers.

2 EFFECT OF UPPER CUTOFF FREQUENCY

An often-quoted criterion for acceptable high-frequency phase distortion in audio amplifiers is a maximum phase shift at 20 kHz of 5°. If a first-order, low-pass rolloff is assumed, it can be shown that a minimum bandwidth of approximately 230 kHz would be required to meet this criterion. However, the interpretation of phase shift as phase distortion is meaningless if the time delay through the amplifier is not accounted for. It is proposed here that the differential phase-shift distortion be used as a criterion for phase response. For example, a criterion for high-frequency phase distortion might be that the maximum acceptable differential phase-shift distortion at 20 kHz be 5°. This is illustrated in the following for several example low-pass transfer functions.

The high-frequency response of an amplifier is modeled with a low-pass transfer function. For many power amplifiers, a first-order transfer function is probably adequate. In general, however, amplifers have high-frequency rolloff characteristics that require higher order transfer function models. Indeed, it is not uncommon for amplifier circuits to be preceded by a passive or active low-pass filter to protect the circuit from unintentional ultrasonic or RF signals, which could be present at the input. It would be impossible here to examine the phase response of arbitrary low-pass filter transfer functions. Therefore the analysis will be limited to first-order and several second-order functions to illustrate the theory. It is straightforward to extend the analysis to higher order cases.

The transfer function of a first-order or second-order low-pass filter can be written in the form

$$F(s) = K\frac{1}{1 + a_1(s/\omega_0) + a_2(s/\omega_0)^2} \tag{7}$$

where K is the gain constant, s is the complex frequency, and $\omega_0 = 2\pi f_0$ is a normalization frequency. For a first-order transfer function, $a_1 = 1$ and $a_2 = 0$. For a second-order transfer function, $a_1 > 0$ and $a_2 = 1$.

For the first-order case ($a_1 = 1$ and $a_2 = 0$) it is straightforward to show that the upper half-power cutoff frequency is f_0. The phase and group delays are given by

$$\tau_\varphi = \frac{1}{2\pi f}\tan^{-1}\left(\frac{f}{f_0}\right) \tag{8}$$

$$\tau_g = \frac{1}{2\pi f}\frac{f/f_0}{1 + (f/f_0)^2}\,. \tag{9}$$

The differential time-delay and phase-shift distortions are obtained from Eqs. (1) and (6).

For the second-order case ($a_1 > 0$ and $a_2 = 1$), the parameter a_1 determines the alignment of the transfer function. For $a_1 = 2$, it is critically damped. For $a_1 = \sqrt{3}$, it is Bessel. For $a_1 = \sqrt{2}$, it is Butterworth. For $a_1 < \sqrt{2}$, it is Chebyshev. The 0.5-dB and 1.0-dB ripple Chebyshev alignments will be used as examples in the following. For the 0.5-dB ripple case, $a_1 = 1.1578$. For the 1.0-dB ripple case, $a_1 = 1.0455$. The upper half-power cutoff frequency for the second-order function is given by

$$f_u = f_0\left[1 - 0.5a_1^2 + \sqrt{(1 - 0.5a_1^2)^2 + 1}\right]^{1/2}\,. \tag{10}$$

The phase and group delays are given by

$$\tau_\varphi = \frac{1}{2\pi f}\tan^{-1}\left[\frac{a_1(f/f_0)}{1 - (f/f_0)^2}\right] \tag{11}$$

$$\tau_g = \frac{1}{2\pi f}\frac{a_1(f/f_0)[1 + (f/f_0)^2]}{[1 - (f/f_0)^2]^2 + a_1^2(f/f_0)^2}\,. \tag{12}$$

The differential time-delay and phase-shift distortions are obtained from Eqs. (1) and (6).

Fig. 2 gives the amplitude response versus normalized frequency f/f_u for six cases: first-order and second-order critical, Bessel, Butterworth, 0.5-dB ripple Chebyshev, and 1.0-dB ripple Chebyshev. The nor-

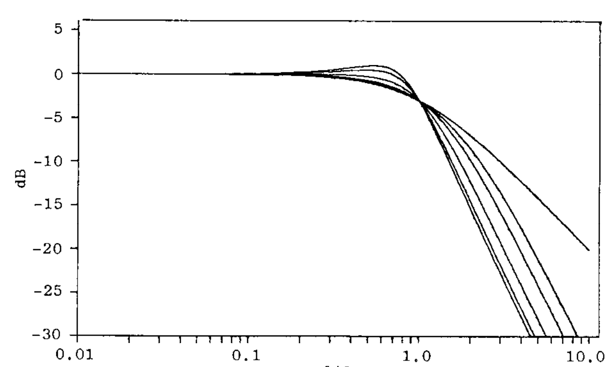

Fig. 2. Magnitude response in decibels versus normalized frequency f/f_u for example low-pass filter transfer functions. Upper curve for $f/f_u \geqslant 1$ is first order, lower curves are second order in order of decreasing a_1.

malization frequency has been chosen to be the upper half-power cutoff frequency f_u. This makes the frequency f_0 different for each case. Exceptions are the first-order and Butterworth second-order cases for which $f_0 = f_u$. Fig. 3 gives the phase responses for the six cases of Fig. 2. The phase-delay and group-delay responses are given in Figs. 4 and Fig. 5, respectively. The differential time-delay distortions are given in Fig. 6 and the differential phase-shift distortions in Fig. 7.

Examination of the figures shows that the second-order Bessel alignment exhibits the flattest differential time-delay and phase-delay distortions up to the cutoff frequency. This would be expected because the Bessel alignment corresponds to a maximally flat delay. It is interesting to note that the second-order Bessel alignment exhibits better response in these respects than the first-order alignment. This was found surprising because it is often said that higher order filters have worse phase response. The results imply that phase response can be improved if the order of a filter is increased, provided the filter alignment is chosen carefully.

To compare the phase linearity of the transfer functions numerically, let 20 kHz be the frequency at which the differential phase-shift distortion is 5°. The upper half-power cutoff frequency f_u for the different filter alignments which will give this amount of phase-shift distortion can be calculated. It is summarized in Table 1. The second-order Bessel alignment requires the least

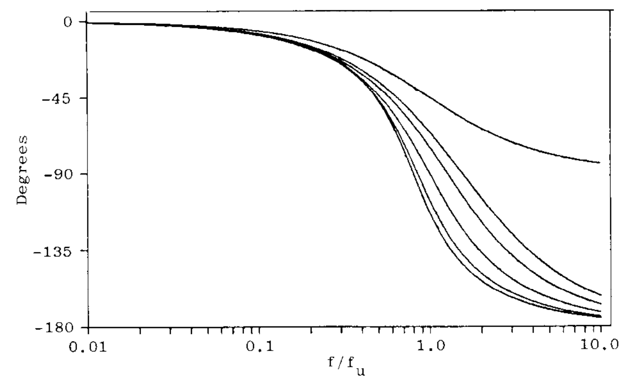

Fig. 3. Phase response in degrees versus normalized frequency f/f_u for example low-pass filter transfer functions. Upper curve for $f/f_u \geq 1$ is first order, lower curves are second order in order of decreasing a_1.

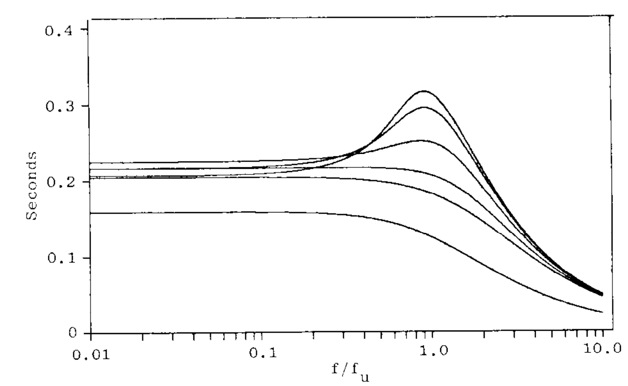

Fig. 4. Phase delay τ_φ in seconds versus normalized frequency f/f_u for example low-pass filter transfer functions. Lower curve at $f/f_u = 1$ is first order, upper curves are second order in order of decreasing a_1.

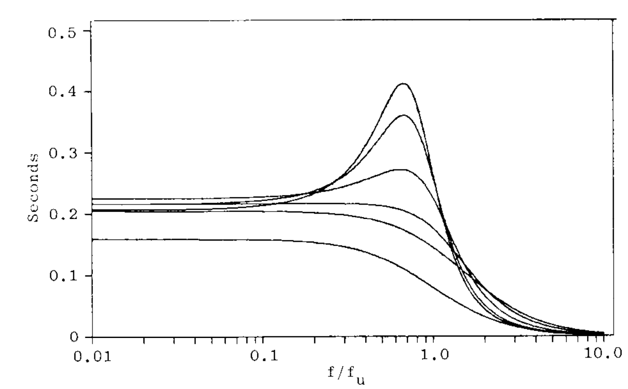

Fig. 5. Group delay τ_g in seconds versus normalized frequency f/f_u for example low-pass filter transfer functions. Lower curve at $f/f_u = 1$ is first order, upper curves are second order in order of decreasing a_1.

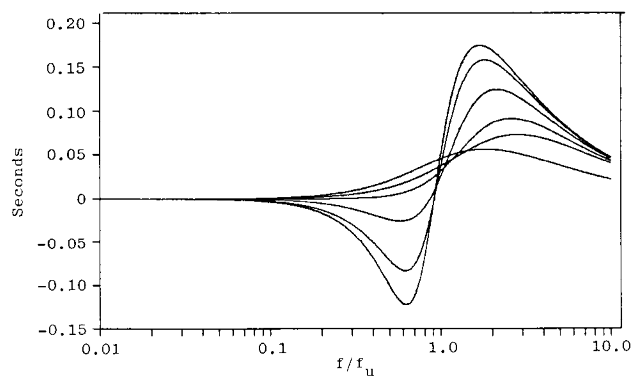

Fig. 6. Differential time-delay distortion $\Delta\tau$ in seconds versus normalized frequency f/f_u for example low-pass filter transfer functions. Upper curve for $0.1 \leq f/f_u \leq 1$ is first order, lower curves are second order in order of decreasing a_1.

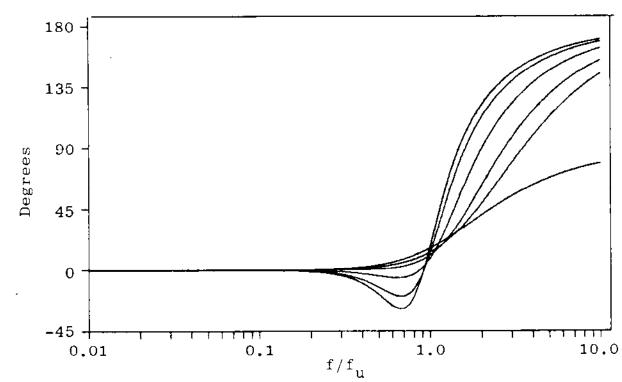

Fig. 7. Differential phase-shift distortion $\Delta\varphi$ in degrees versus normalized frequency f/f_u for example low-pass filter transfer functions. Upper curve for $0.1 \leq f/f_u \leq 1$ is first order, lower curves are second order in order of decreasing a_1.

Table 1. Upper cutoff frequency for less than 5° of differential phase-shift distortion below 20 kHz.

1st order	35 kHz
2nd order, critical	30 kHz
2nd order, Bessel	25 kHz
2nd order, Butterworth	37 kHz
2nd order, 0.5-dB Chebyshev	57 kHz
2nd order, 1.0-dB Chebyshev	63 kHz

bandwidth of 25 kHz, which is only 0.3 octave greater than 20 kHz. The first-order alignment requires 35 kHz, which is a factor of 6.5 lower than the 230 kHz required for 5° or less of phase shift at 20 kHz for the same alignment. The familiar Butterworth alignment requires 48% more bandwidth than the Bessel. Indeed, the superior phase response of the Bessel alignment over the Butterworth can easily be demonstrated by square-wave tests. The Butterworth alignment exhibits ringing, whereas the Bessel does not. Comparisons for transfer functions greater than second order have not been made. However, because the Bessel approximation to a maximally flat delay improves as the order of the alignment increases, it is stated without proof that a flatter differential phase-shift distortion should be obtained as the order of the Bessel filter is increased.

3 EFFECT OF LOWER CUTOFF FREQUENCY

For 5° or less of high-frequency differential phase-shift distortion, it has been demonstrated that the upper cutoff frequency is typically less than twice the highest frequency to be reproduced. On the low-frequency side of the spectrum, however, things are different. For a high-pass filter transfer function it follows that the group delay is positive while the phase delay is negative. When these are subtracted to obtain the differential time-delay distortion, they do not tend to cancel as with the low-pass transfer functions. Instead, the difference is greater in magnitude than the magnitude of either the group delay or the phase delay alone. Thus the low-frequency bandwidth requirements are much more strenuous if phase linearity is to be maintained.

The transfer functions for the first- and second-order high-pass cases can be obtained from Eq. (7) by multiplying by $(s/\omega_0)^n$, where $n = 1$ for the first-order case and $n = 2$ for the second-order case. For the six filter alignments used for the low-pass cases, the numerical values given for coefficients a_1 and a_2 will give the same alignments for the high-pass cases. For the first-order transfer function, the lower half-power cutoff frequency is equal to the normalization frequency f_0. The phase response of the first-order transfer function is given by

$$\varphi = \frac{\pi}{2} - \tan^{-1}\left(\frac{f}{f_0}\right) . \qquad (13)$$

The phase and group delays are given by

$$\tau_\varphi = \frac{1}{2\pi f}\left[\tan^{-1}\left(\frac{f}{f_0}\right) - \frac{\pi}{2}\right] \qquad (14)$$

$$\tau_g = \frac{1}{2\pi f}\,\frac{f/f_0}{1 + (f/f_0)^2} . \qquad (15)$$

The differential time-delay and phase-shift distortions are obtained from Eqs. (1) and (6).

For the second-order high-pass transfer functions, the lower half-power cutoff frequency is given by

$$f_l = f_0\left[1 - 0.5a_1^2 + \sqrt{(1 - 0.5a_1^2)^2 + 1}\right]^{1/2} . \qquad (16)$$

The phase and group delays are given by

$$\tau_\varphi = \frac{1}{2\pi f}\left\{\tan^{-1}\left[\frac{a_1(f/f_0)}{1 - (f/f_0)^2}\right] - \pi\right\} \qquad (17)$$

$$\tau_g = \frac{1}{2\pi f}\,\frac{a_1(f/f_0)[1 + (f/f_0)^2]}{[1 - (f/f_0)^2]^2 + a_1^2(f/f_0)^2} . \qquad (18)$$

The differential time-delay and phase-shift distortions are obtained from Eqs. (1) and (6).

Fig. 8 gives the amplitude response versus normalized frequency f/f_l for six cases: first-order and second-order critical, Bessel, Butterworth, 0.5-dB ripple Chebyshev, and 1.0-dB ripple Chebyshev. The normalization frequency has been chosen to be the lower half-power cutoff frequency f_l. This makes the frequency f_0 different for each case. Exceptions are the first-order and Butterworth second-order cases for which $f_0 = f_l$. Fig. 9 gives the phase responses for the six cases of Fig. 8. The phase- and group-delay responses are given in Figs. 10 and 11, respectively. The differential time-delay distortions are given in Fig. 12 and the differential phase-shift distortions in Fig. 13.

To compare the phase linearity of the transfer functions numerically, let 20 Hz be the frequency at which the differential phase-shift distortion is 5°. The lower half-power cutoff frequency f_l for the different filter alignments that will give this amount of phase-shift distortion can be calculated. It is summarized in Table 2. The highest value of f_l occurs with the first-order transfer function. For the second-order transfer functions the values are approximately the same. It can be concluded that the first-order high-pass filter has the least phase nonlinearity. It requires a lower cutoff frequency, which is lower than 20 Hz by a factor of 23, or slightly

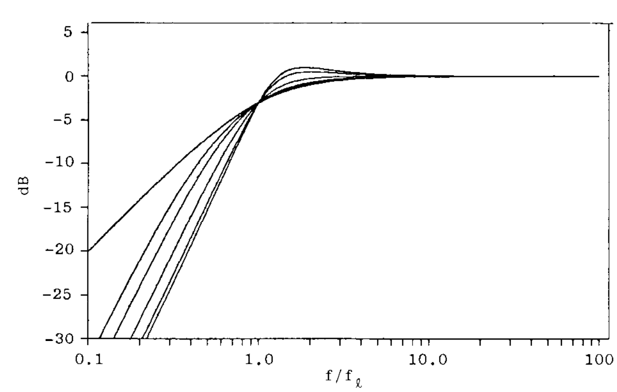

Fig. 8. Magnitude response in decibels versus normalized frequency f/f_l for example high-pass filter transfer functions. Upper curve for $f/f_l \leqslant 1$ is first order, lower curves are second order in order of decreasing a_1.

more than one decade plus an octave. This is to be contrasted to the first-order low-pass transfer function, which requires slightly less than an octave of bandwidth beyond 20 kHz for the same phase linearity.

The results imply that the lower the order for high-pass filters, the less the phase distortion. The Bessel-aligned high-pass filter does not share the linear-phase characteristic that its low-pass counterpart does. Indeed, a high-pass filter cannot be said to have a Bessel alignment because a maximally flat delay approximation is

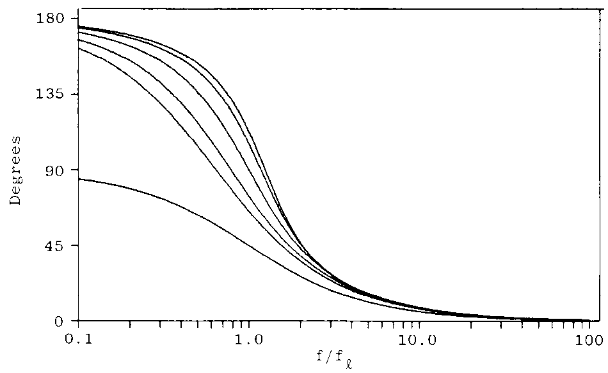

Fig. 9. Phase response in degrees versus normalized frequency f/f_l for example high-pass filter transfer functions. Lower curve for $f/f_l \leq 1$ is first order, upper curves are second order in order of decreasing a_1.

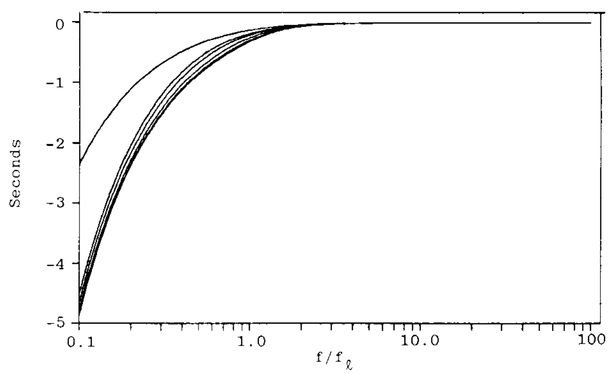

Fig. 10. Phase delay τ_φ in seconds versus normalized frequency f/f_l for example high-pass filter transfer functions. Upper curve for $f/f_l \leq 1$ is first order, lower curves are second order in order of decreasing a_1.

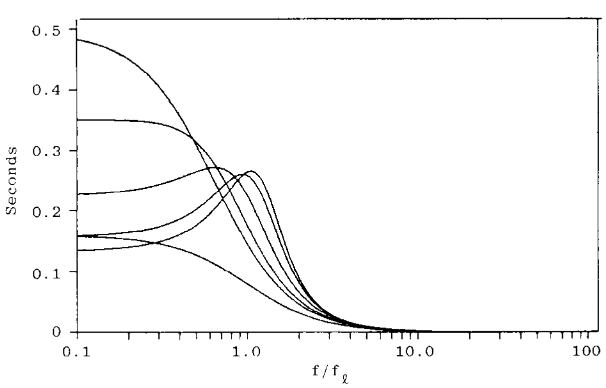

Fig. 11. Group delay τ_g in seconds versus normalized frequency f/f_l for example high-pass filter transfer functions. Lowest curve for $1 \leq f/f_l \leq 10$ is first order, upper curves are second order in order of decreasing a_1.

impossible for high-pass filters. This result can be applied to loudspeaker alignments as well as to amplifier circuits. These alignments are second-order high pass for closed-box systems and fourth-order high pass for vented-box systems. For these systems it cannot be said that a Bessel alignment has desirable phase response compared to other alignments. This means that the low-frequency phase response of a loudspeaker cannot be optimized by varying the alignment. Therefore it can be concluded that the loudspeaker alignment should be chosen for best frequency response.

4 CONCLUSIONS

The differential time-delay distortion and phase-shift distortion have been defined. A rationale has been proposed for the specification of the phase linearity of a system that is based on these two functions. It has been

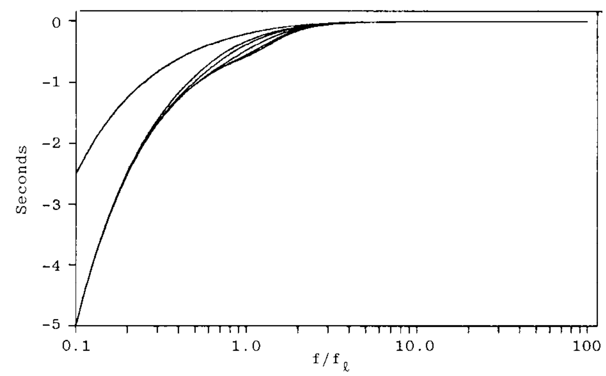

Fig. 12. Differential time-delay distortion $\Delta\tau$ in seconds versus normalized frequency f/f_l for example high-pass filter transfer functions. Upper curve at $f/f_l = 1$ is first order, lower curves are second order in order of decreasing a_1.

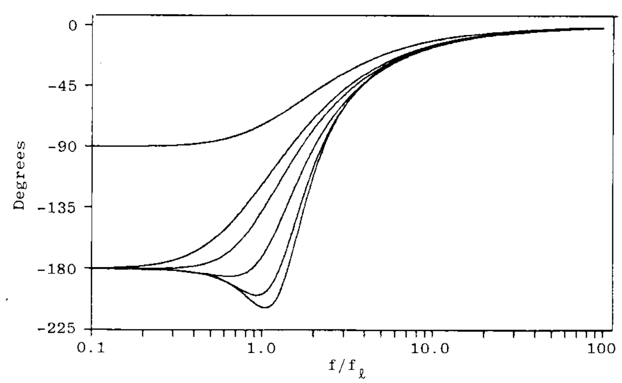

Fig. 13. Differential phase-shift distortion $\Delta\varphi$ in degrees versus normalized frequency f/f_l for example high-pass filter transfer functions. Upper curve at $f/f_l = 1$ is first order, lower curves are second order in order of decreasing a_1.

Table 2. Lower cutoff frequency for less than 5° of differential phase-shift distortion above 20 Hz.

1st order	0.87 Hz
2nd order, critical	0.68 Hz
2nd order, Bessel	0.64 Hz
2nd order, Butterworth	0.62 Hz
2nd order, 0.5-dB Chebyshev	0.64 Hz
2nd order, 1.0-dB Chebyshev	0.67 Hz

shown that low-pass transfer functions exhibit less phase-shift distortion than high-pass transfer functions of the same order. For first-order low-pass functions, a bandwidth slightly less than one octave higher than the highest frequency to be reproduced is sufficient for negligible phase distortion. For first-order high-pass functions, a bandwidth of approximately one decade plus another octave lower than the lowest frequency to be reproduced is necessary. For low-pass functions, the second-order Bessel alignment gives the optimum phase distortion characteristics of all first- and second-order alignments. Because the Bessel alignment exhibits a maximally flat delay, where the "flatness" improves as the order of the filter is increased, it should follow that the phase distortion of Bessel low-pass functions decreases as the order of the filter is increased. For high-pass functions, the phase distortion is not sensitive to the alignment but increases as the order of the filter is increased.

5 REFERENCES

[1] T. A. Saponas, R. C. Matson, and J. R. Ashley, "Plain and Fancy Test Signals for Music Reproduction Signals," *J. Audio Eng. Soc.*, vol. 19, pp. 294–305 (1971 Apr.).

[2] D. G. Fink, Ed., *Electronics Engineers Handbook*. (McGraw Hill, New York, 1975).

Dr. Leach's biography was published in the July/August issue.