Group and Phase Delay Requirements for Loudspeaker Systems (Invited)

J. Robert Ashley

Koss Corporation

Milwaukee, WI 53212

ABSTRACT

The question of true requirements regarding phase shift tolerance in audio systems is reopened. No need for phase delay linearity can be justified for music signals. The actual requirement is a tolerance of 2 ms for group delay through the parallel channels (woofer and squawker) of a speaker. Simulated drum tone experiments are reported to verify this well known number, 2 ms. Group delay in the highly desirable 3rd order Butterworth crossover network is found to be identical in high and low pass channels. Phase shift in transducers causes the phase troubles in speaker systems. An idealized two way speaker combines the response of the transducers into the crossover network and solves all of the phase shift related problems in a two way speaker.

1. The Questions

Is phase linearity needed in loudspeakers? Does time alignment make an audible difference? What is the true bound on phase shift in an audio system? These are among the many questions usually asked regarding the topic of phase shift and time delay in all kinds of audio systems. The people willing to give answers fall into two camps: 1) The phase shift is not important followers of Helmholtz, or 2) Phase linearity is needed. Since the conventional wisdom of time alignment is readily "understood" by the loudspeaker buyers, those in the second camp are making more money than those in the first camp. Still, the questions deserve scientific answers and those of us in camp 1 would like to be able to explain why time alignment is not needed in a state-of-the-art loudspeaker system.

2. What is Music?

The most important answer is that music is what I want to sound good coming out of my speaker. If using pulses, square waves, tone bursts, etc., as test signals will tell me something useful regarding music reproduction, then I will be happy to use these signals. But, I see no need to place demands on my speaker regarding the time domain oscilloscope pictures of pulse or impulse response.

The original analyses of music considered only the steady state portions of a tone--say the middle of a long bowed note on a violin. Based on some simplifying assumptions, the vibration is found to be periodic with integer relationships between the eigenfunctions; that is, we can model the sound radiated by

$$p(t) = \sum_n C_n \exp(2\pi n f_o t + \theta_n) \qquad (1)$$

where n is a precise integer. We have generated a signal consisting of fundamental and second harmonic with variable θ_2. We have yet to find a subject who can detect change in θ_2, even when his hand is on the control knob. [1]

We have also repeated several other authors' experiments of this kind and not verified the authors' contentions regarding audibility of phase shift. The usual flaw we find in the work of others is failure to consider linearity of the loudspeakers used. For example, we used a carefully chosen and tested dynamic speaker in place of an electrostatic speaker in one experiment and did not hear the phase shift as claimed by the original experimenter.

Regardless of the result of experiments based on precisely periodic signals, there is another more important reason why gradual phase shift is not audible. The above assumption that n is precisely integer is not correct. For our violin tone, the relatively soft end of the finger which sets the length of the string will set a slightly different length for the fundamental and the various harmonics; thus, n will not be precisely integer. Now, θ_n is a slow function of time--it is fortunate that we cannot hear phase shift. From another standpoint, the well tempered musical scale (based on a ratio between semitones of $\sqrt[12]{2}$) creates musical intervals which are not the naturally expected integer harmonic relationship. We prefer to listen to this slightly discordant music where there is not a precise integer relationship.

If phase requirements cannot be based on steady state music, then we must explore transient music. A stacatto note on a piano (Fig. 1)

CH1559-4/80/0000-1030$00.75

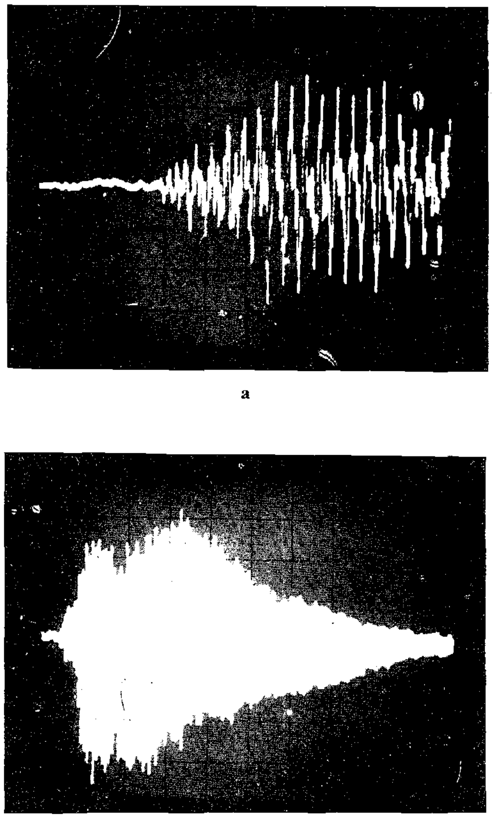

Fig. 1. Sound pressure waveform for staccato middle C on a grand piano. a. 10 ms/div. b. 50 ms/div.

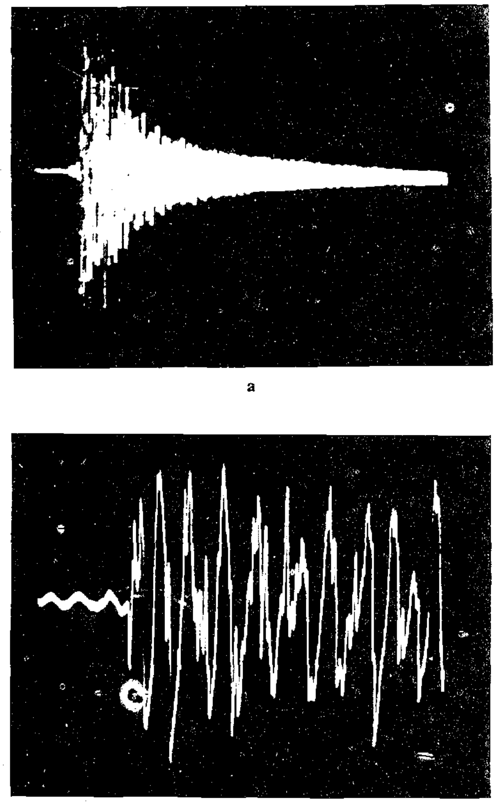

Fig. 2. Sound pressure waveform for one stroke on a 16-inch floor tom-tom. a. 100 ms/div. b. 10 ms/div.

or a single stroke on a drum (Fig. 2) are musical transients. Note that in the time domain they are not single pulses, impulses, or step functions. Instead, they are sinusoids amplitude modulated by relatively long envelope functions. In the frequency domain, the energy is clustered in narrow packets near the expected eigenfrequencies. Time delay specifications must be concerned with delay of the envelope functions.

3. What is Important about Envelope Delay?

Most music is performed by groups of musicians numbering from 2 to 200. Spacings on the stage mean that tens of milliseconds can be expected between, say, a cello playing at 110 Hz and the violin at 440 Hz. We must consider just a single instrument playing a short note and determine a tolerance on relative delay between the envelope of the fundamental and the envelopes of the harmonics.

To determine this tolerance, we have devised an experiment to simulate the drum tone on the Hitachi 505 hybrid computer. As shown in Fig. 3 two groups of integrators are used to generate damped sinusoids. In listening, the lower frequency signal by itself reminded our auditors of a heart beat. Adding in the next eigenfrequency (at the proper ratio of 2.29 times the fundamental) markedly improved the naturalness of our drum tone simulation.

Most subjects tested detected a subtle difference in tone quality with 5 to 7 ms delay. Our most acute subject could detect 3 ms. Thus, we conclude that the result by Hilliard of 2 ms delay difference is a reasonable tolerance. This corresponds to a distance tolerance of some 70 cm. This makes us somewhat skeptical of the audibility of time alignment changes where perhaps a quarter of this distance is involved.

With this magic number, 2 ms, in hand, we must consider the sources of group delay differences in loudspeaker systems.

4. Group Delay in Crossover Networks

We use as an example the third order Butterworth low pass and mirror image high pass which we have found to be the most useful crossover network. The Butterworth polynomial is

$$B_3(ju) = (1 - u^2) = j(2u - u^3) \qquad (2)$$

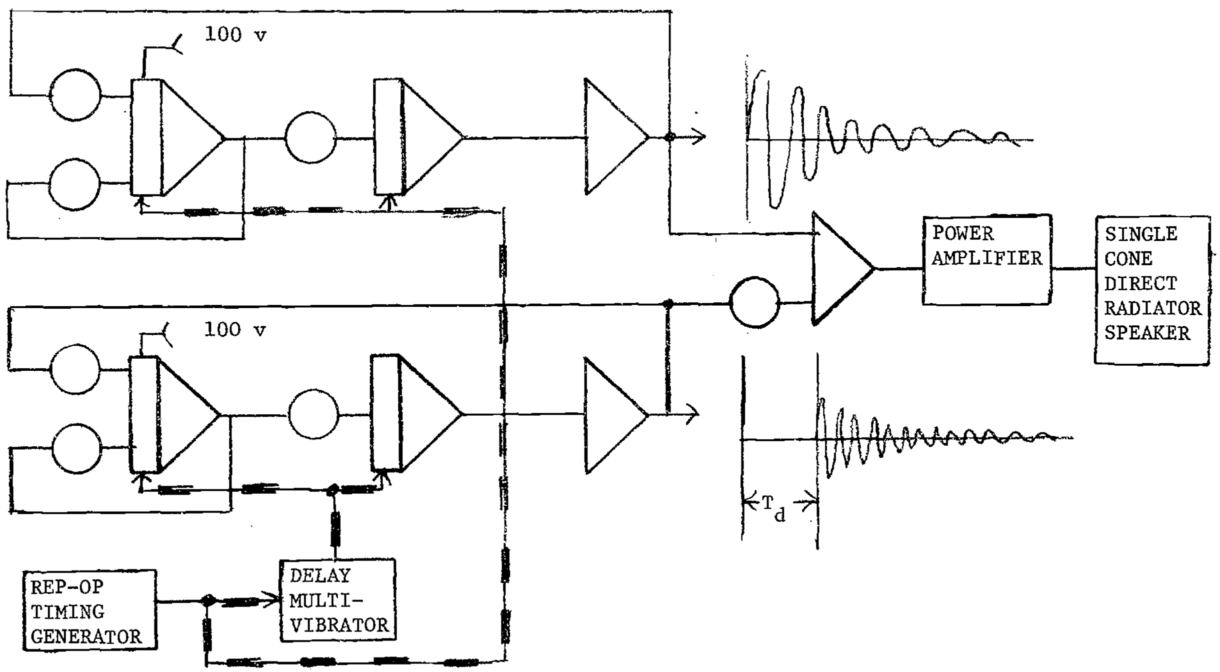

FIG. 3. A SIMULATED DRUM TONE EXPERIMENT. T_d IS VARIABLE FROM ZERO TO SEVERAL SECONDS BY THE MULTIVIBRATOR.

where u = f/fc, normalized frequency.

It is easy to show that the group delay for either the high pass or the low pass filter is

$$T_G = -\frac{d\theta}{d\omega} = \frac{1}{\omega_c}\left(\frac{2 + u^2 + 2u^4}{1 + u^6}\right) \tag{3}$$

It is a general result that any order Butterworth filter crossover network will have equal group delay in both channels. For a 318 Hz (2,000 rad/sec) crossover network, the maximum delay is 1 ms. Since there is no difference in delay, the crossover network causes no phase problems.

5. Group Delay in Woofers and Squawkers

Since any electroacoustic transducer behaves as a band-pass network, it will have group delay. For a two way speaker, we must consider the upper end of the woofer pass-band and the lower end of the squawker pass-band. For a 318 Hz crossover frequency, the woofer voice coil inductance will have to be controlled so that

$$T_{HW}(ju) = \frac{1}{1 + ju} \tag{4}$$

where u = f/600 Hz

which leads to a group delay of

$$T_{DW} = \frac{1}{(2\pi)(600)(1 + u^2)} \tag{5}$$

A 10 cm cone diameter closed box squawker might have a transfer function

$$T_{LS}(ju) = \frac{-u^2}{1 - u^2 + j\frac{u}{Q}} \tag{6}$$

where u = f/150 and Q = 1

with the corresponding group delay

$$T_{DS} = \left(\frac{1}{Q\omega_o}\right)\left(\frac{1 + u^2}{1 + \left(\frac{1}{Q^2} - 2\right)u^2 + u^4}\right) \tag{7}$$

The squawker has a maximum group delay of 1.5 ms compared to a quarter ms for the woofer. As a logical consequence, proper "time alignment" for this system would place the woofer 40 cm behind the squawker. Have you ever seen one built like this? (Yes, the Klipsch Horn.)

6. An Idealized Two-way Speaker

We have just shown that the evil in loudspeaker systems is the phase shift in the transducers. We have done numerous computer simulations to verify our contention that the best one can do using the conventional ideas of overlap of bandpass of the woofer and squawker and 1st or 3rd order crossover networks is to live with the problem. In our simulations, we have added pure time delay (phase delay equal to group delay) and the usual result is to help in one frequency band at the expense of harm in another frequency band. The usual ideas of time alignment (moving the squawker back w.r.t. the woofer) do not solve the true phase shift problems in a two-way speaker.

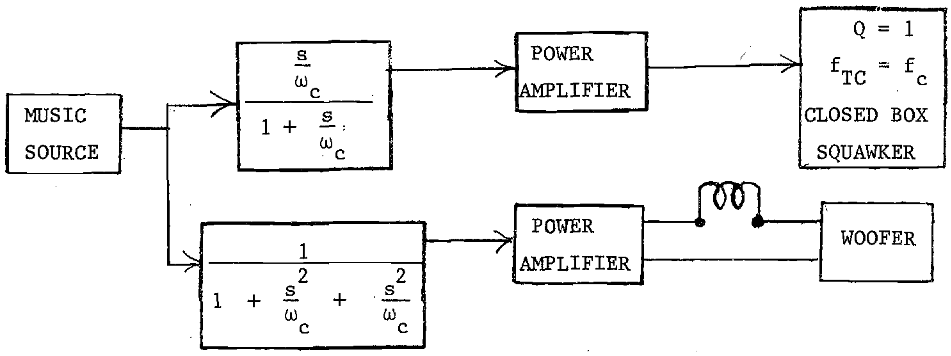

FIG. 4. BLOCK DIAGRAM OF A NOVEL LOUDSPEAKER SYSTEM WHICH ELIMINATES GROUP DELAY DIFFERENTIAL CAUSED BY THE TRANSDUCER PHASE SHIFTS.

We therefore deem it necessary and desirable to use the transfer characteristic of the woofer and of the squawker as part of the crossover characteristic. Recall that the 3rd order Butterworth polynomial can be readily factored.

$$\begin{aligned} B_3(s) &= s^3 + 2s^2 + 2s + 1 \\ &= (s+1)(s^2 + \frac{s}{Q} + 1) \end{aligned} \tag{8}$$

where $Q = 1$

Therefore, we achieve the low pass characteristic as the product

$$T_{Lo}(s) = \frac{1}{1 + \frac{s}{\omega_c}} \cdot \frac{1}{1 + \frac{s}{Q\omega_c} + \frac{s^2}{\omega_c^2}} \tag{9}$$

and the high pass characteristics as the product

$$T_{Hi}(s) = \frac{\frac{s}{c}}{1 + \frac{s}{\omega_c}} \cdot \frac{\frac{s^2}{\omega_c^2}}{1 + \frac{s}{Q\omega_c} + \frac{s^2}{\omega_c^2}} \tag{10}$$

It is obvious to the most casual observed (and, we pray, the patent examiners) that the speaker system shown in Fig. 4 overcomes all the group delay difficulties in a two-way loudspeaker system.

REFERENCE

1. T. A. Saponas, R. C. Matson, & J. Robert Ashley, "Plain and Fancy Test Signals for Music Reproduction Systems," J. Audio Eng. Soc., V 19, No. 4, Apr. 1971