Requirements for Loudspeaker Crossover Networks

J. Robert Ashley

Sperry Corporation
Electronic Systems
Clearwater, FL 33518

ABSTRACT

After defining terms to be used, the need for a crossover filter in a loudspeaker is explained. Justification of crossover filter functions is based on a listening test first proposed in 1962. The requirements for a crossover filter are: 1. partitioning the audio spectrum, 2. a tolerance on power response, 3. control of intermodulation distortion, 4. a tolerance on phasor magnitude response, 5. a tolerance on system group delay, and 6. a tolerance on adjacent channel group delay difference. The audio fad known as time alignment and several crossover networks will be debunked.

I. TERMINOLOGY

In deference to accepted vocabulary, a box which contains electroacoustic transducers, electrical networks, and other parts will be simply termed a "speaker". A single transducer, such as a woofer, will be called a loudspeaker driver. Combination of a loudspeaker driver with an intelligently engineered box or enclosure yields a loudspeaker system.

The first multi-driver speakers contained two loudspeaker drivers and an electrical filter to separate the spectrum into lower and higher regions. The loudspeaker systems came to be called woofers and tweeters. As speakers improved to 3-way systems, the transducer for the middle spectrum region came to be called a mid-range. To make a more consistent terminology, Paul Klipsch coined the name "squawker". If one disables the woofer and tweeter in a 3-way speaker, the resulting sound does remind one of a squawking TV set.

My marketing friends object to the term squawker. Furthermore, a four way speaker is the real state of the art. Another name for the fourth transducer is needed. Therefore, I hereby introduce the following names for the loudspeaker systems needed in a 4 way system. Starting from the lowest band and moving to the highest, we have:

- Bass system
- Baritone system
- Alto system
- Soprano system

Since we can also refer to tenors and several kinds of sopranos, there is modest room for expansion of this terminology.

II. NEED FOR A SPECTRUM PARTITIONING FILTER

It is possible to get most of the music spectrum out of a one-way speaker---until you try to get more than very low level background music. Controlling intermodulation distortion at even modest listening levels requires a 3 way speaker. Why?

First, if one watches a real time analyzer connected to an electrical source of symphonic music, the long term eyeball average is a flat line. In a crude sort of way, the spectrum envelope of most music is approximately pink.

Second, the acoustical radiation from a direct radiator loudspeaker is proportional to acceleration of the cone. To maintain uniform frequency response, the cone diameter must be reduced as the expected response frequency goes up. A 25 mm dome soprano system must accelerate at thousands of G's if several volts are applied at 3 kHz. To maintain constant acceleration as frequency is decreased, the cone displacement increases at 40 dB per decade.

Third, Klipsch [1,2,3] has shown the effects of intermodulation distortion and the causes for what is often called Doppler distortion. This angle modulation effect is not present in amplifiers and is the reason why amplifier distortion theory is not adequate for speakers.

Fourth, when all the design constraints are placed on a given loudspeaker system, either horn or direct radiator loudspeaker systems can effectively work for about 1 decade per system. As examples of the current state of the art, we tabulate:

50 liter bass system	35 to 800 Hz
1 liter baritone system	200 to 2500 Hz
5 cm cone type alto system	1500 to 7000 Hz
2 cm dome type soprano system	4000 to 18000 Hz

Consideration of these four factors leads to the need to partition the spectrum at 250, 2000, and 5000 Hz to radiate significant acoustical power (tens of milliwatts at single frequencies) with modest distortion.

III DETERMINATION OF REQUIREMENTS

What are the requirements on the electrical filters needed for frequency spectrum partitioning and how do we determine these requirements? The ultimate answer is that the crossover filters must make the speaker work well reproducing speech and

CH1945-5/84/0000-0215 $1.00 © 1984 IEEE

music. The final judge is the human hearing apparatus. The answer must be obtained by listening tests using speech and music as test signals.

Since a crossover filter is needed for a completed speaker, we have difficulties developing crossover filter requirements simultaneously with developing a speaker. It is a chicken and egg problem. I gave one way to get around this in 1962 [4] although an illustration might have kept my test from being widely ignored. There are many fairly good to good speakers available for home entertainment systems. Thus, using these as shown in fig. 1 can separate the driver development process from the study of filter functions. A photograph of the speakers used for this talk as postitioned in our living room is given in fig. 2. Please notice that the spectrum above the crossover frequency always comes out of the top speaker. The radiation of the frequencies above 300 Hz will be different for the speakers as shown; thus, to keep the radiated response constant above the crossover frequency, the arrangement of fig. 1 must be carefully duplicated.

The idea behind this test is to have a reference sound which has the same coloration as the sound from the complete system. Even though transducers and filter functions may not be optimum in the speakers used, the fact that good production lines will turn out speakers within a fraction of a dB of a production standard means that the identical response assumption implied here is well met.

IV CONSTANT POWER RESPONSE

After the music spectrum has been partition and radiated, we must ask how the radiated sound is recombined as far as human hearing is concerned. In Fig. 3, we show two possible ways for making a measurement on how the sound recombines. The phasor summation is what a microphone located 50 cm from the speaker will measure. This is usually thought to be the most important response of a speaker. If we determine the power radiated by each transducer by itself and then add the power, we have eliminated phase cancellation effects. This quantity is approximately measured by putting 2 to 3 microphones at scattered locations in the listening room [5]. The intensity at each microphone is determined and the intensities added.

As the result of 30 years of serious study of speakers and room acoustics, I believe the power response is far more important than the phasor response. The listening room does seem to integrate the power radiated. For stereo listening, we are usually several meters from each speaker and the near field response is not important. We note this conclusion is not in agreement with the conventional wisdom of speaker "theory."

The requirement is a tolerance on the flatness of the power response. In a complete speaker, A vs B testing of similar speakers can detect a 1 dB peak or valley which is different between the speakers. I believe this is the tolerance on power response--that the response be flat within + or - 1 dB.

V. CONTROL OF DISTORTION

In my discussion of the need for partitioning the spectrum, I pointed out that radiating sufficient acoustical power without excessive intermodulation distortion requires several different size transducers. After a decision is made to engineer a 3-way or 4-way speaker, then the most crucial work is the choice of crossover frequencies and slopes.

I cannot state a specific tolerance on this requirement. In the evolution of speakers I have designed, the engineering of the crossover between the bass and baritone systems is most important to the musicality of the speaker. Since most vocal

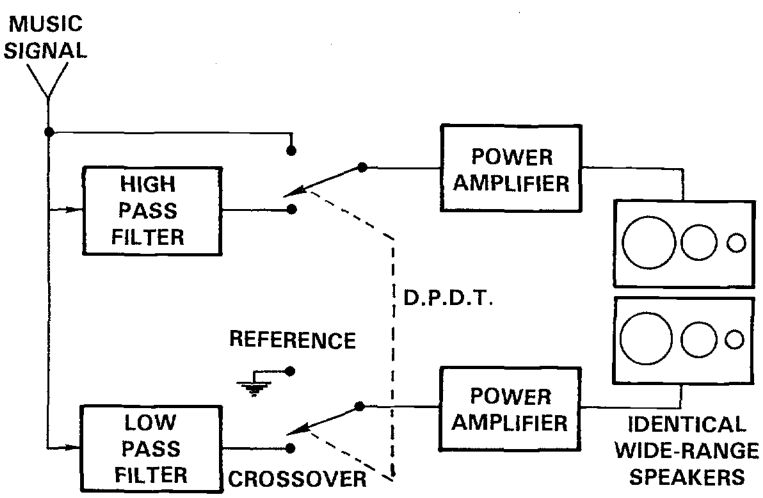

Fig. 1. A listening test for crossover filter functions.

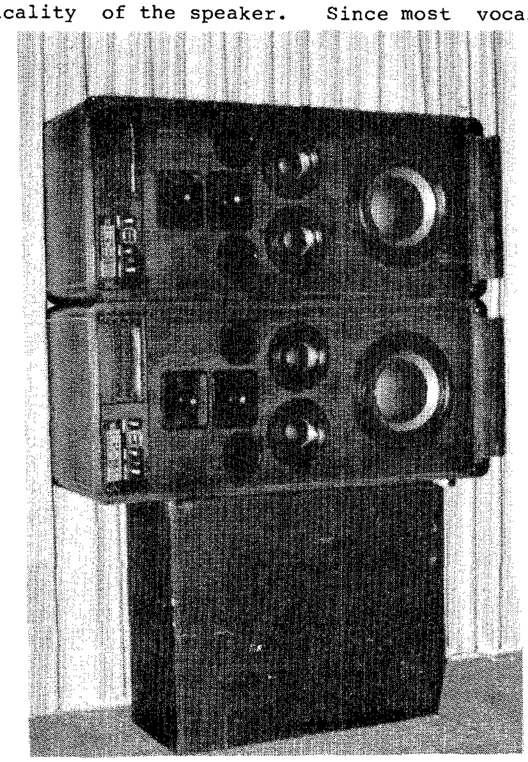

Fig. 2. Koss CM 1030 speakers as used in the listening test of fig. 1.

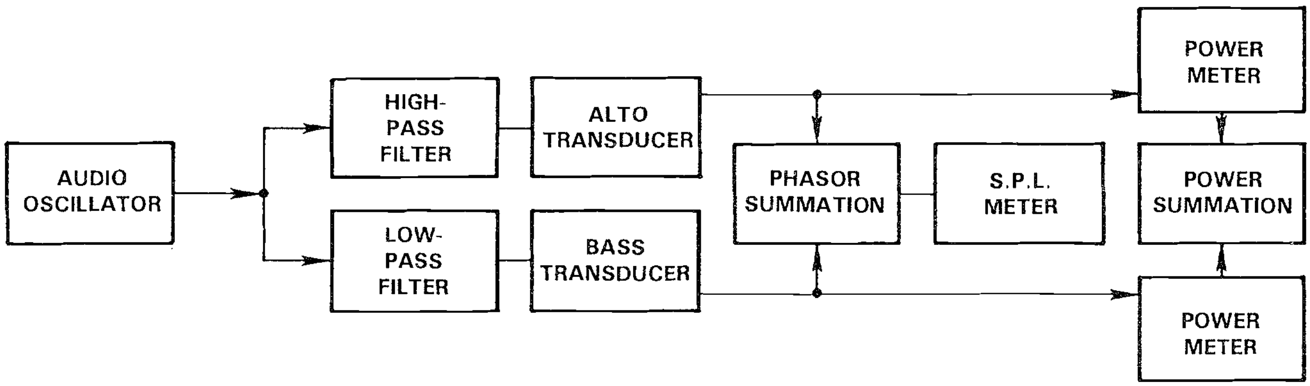

Fig. 3. Phasor and power summation of the output from a two-way speaker.

music is written in the treble clef, I find the crossover frequency should be less than 260 Hz. To control distortion, a 3rd order characteristic function is needed for the baritone high pass filter.

VI. PHASOR MAGNITUDE RESPONSE

In defining power respose, we also defined the phasor magnitude response for comparison. In study of crossover filter functions, one computes the phasor sum (including phase angles for both channels) and then considers the magnitude.

The conventional wisdom is that phasor magnitude response should be flat. I used to believe this the most important measure of speaker response. Now, I do not believe flatness is important. Augspurger [6] discusses common systems that have peaks or dips in phasor response but flat power response.

Fortunately, it is possible to have both flat phasor magnitude response and flat power response. As I taught in 1971, odd order Butterworth crossover filters have both flat phasor magnitude and power response.

VII CROSSOVER FILTER GROUP DELAY

In 1971, we first published the true phase shift requirement for any kind of an audio system [7]. This requirement is in terms of the group delay or envelope delay rather than phase delay. If we use a test signal which has some frequency domain similarity with a single musical tone or note, the spectrum will not be continuous as for a pulse. Applying the results of [7] for a sine squared envelope tone burst, putting 12 cycles of a sinusoid under a sine squared pulse as shown in fig. 4 will give a main spectral lobe which covers - 10 % to + 10 % of the carrier frequency.

As we showed in 1971 [7] and as Cox [8] will illustrate at this ICASSP, music is not short pulses or impulses; instead, even the most transient stacatto note is of the form

$$M(t) = A \exp[-\alpha t] \sin[\omega t]$$

and will have a spectrum shape of a main lobe that is a few percent of the carrier frequency wide. It is not continuous across the whole audio spectrum. This spectral limitation means that we do not consider the phase delay but the group or envelope delay

$$T_g = -\frac{d\phi}{d\omega}$$

There is an obvious local requirement that Tg be constant over the major spectrum content of a note. This means that over a 10 % increment of the spectrum, the phase shift versus frequency characteristic must be linear. When considered in terms of total speaker performance, this means that the phasor magnitude response should not have peaks and valleys higher or deeper than 1 dB of the mean magnitude.

When considered over a 1 to 2 octave range, a change in Tg with frequency implies that the attack or buildup of harmonics in notes may lag or lead the buildup of the fundamental. A demonstration of this effect was developed in 1970 for an AES convention and repeated at ICASSP 80 [9]. We believe that over an octave band, a 2 ms maximum differential group delay is the only significant phase shift requirement in a speaker or any other audio component.

VIII ADJACENT CHANNEL GROUP DELAY DIFFERENCE

If we visualize the tone burst of fig. 4 traveling through the adjacent channels in fig. 3, we can imagine one lagging the other. In the crossover region, both transducers are radiating and a differential delay could cause trouble.

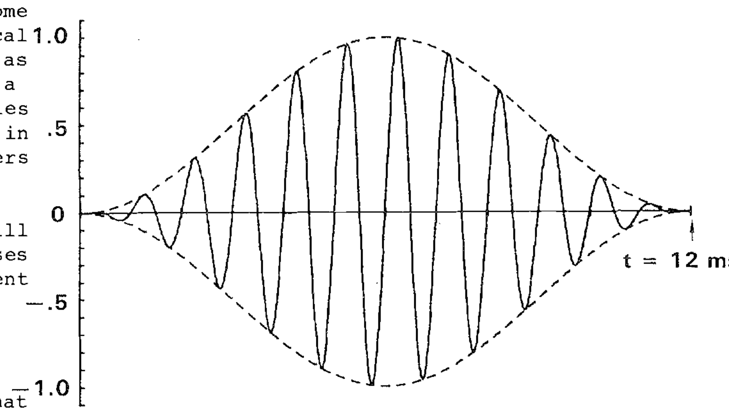

Fig. 4. A confined spectrum test pulse for audio system testing. The dashed curves are the sine squared envelope.

Discovery of this trouble dates back to 1935. As related later by Hilliard [10], two way horn speakers were being evaluated for motion picture sound. A tap dancer seemed to have an echo delayed by about 10 ms. The cause of the delay was the mouths of the horns being in the same vertical plane. Moving the baritone horn back so that the horn drivers were in the same vertical plane cured the difficulty. Hilliard quoted a tolerance of 2 ms on this experiment. Later experiments by Gillum [11] indicate that 5 ms may be a more reasonable tolerance. Since the 2 ms difference is easy to live with, we will quote it as the tolerance.

The tolerances on group or envelope delay are the only phase shift requirements on a speaker. I reiterate my statement of 1971, there are no constraints on the phase velocity of an audio system beyond those implied in the group velocity or envelope delay requirements just stated.

IX DEBUNKING

The topic known as "Audio" is 90 % gadgeteering and 10 % engineering. Audio is an avocation for many well educated people who sometimes publish results before they completely understand the problem [4,12]. In justifying the linear phase and time alignment claims, many refer to a communications theory idea related to distortionless transmission of a pulse. The fallacy of applying this theory to audio is that music is not a short pulse with a continuous spectrum. Since the music spectrum is totally different, there is no justification for applying this video theory to audio.

In light of this fallacy and the true requirements on phase shift as taught above, there is no justifiable need for linear phase or time aligned speakers. Usually, the measures taken to attain linear phase or time alignment screw up some other speaker performance factor and actually degrade the musicality. In properly conducted listening tests, the Koss CM 1030's have blitzkrieged the time aligned competitors.

Another source of misunderstanding is improperly conducted listening tests. In audio folklore, the electrostatic speaker is endowed with magical properties which I have never believed. Their real faults are horrible harmonic distortion generation and terrible intermodulation distortion. We repeated a phase shift test proposed by Madsen with a magnetic motor direct radiator and did not hear phase shift effects. Madsen's auditors actually heard a change in the magnitude of the spectrum caused by distortion in the electrostatic softspeaker.

One result regarding crossover networks which I dispute is the 2nd order filter functions proposed by Thiele and later by Linkovitz [12].

$$T_{lo}(ju) = \frac{1}{(1 + ju)^2} \qquad T_{hi}(ju) = \frac{-(ju)^2}{(1 + ju)^2}$$

A sophomore calculation shows this crossover idea to have a 3 dB hole in the power response at u = 1. Except for the minus sign in the high frequency function, it is the network I rejected in 1962 because it failed the listing test [4].

Recently, Lipschitz and Vanderkooy [13] proposed a variation of the Ashley 1962 idea [4]. They obtain flat phasor magnitude response by requiring

$$T_{hi}(s) = \exp[-\tau s] - T_{lo}(s) \qquad \tau = -\lim_{\omega \to 0} \frac{\phi_L(\omega)}{\omega}$$

Any senior equipped with a VIC 20 can find that these filters have about a 2 dB power hole at the crossover frequency. They also fail several of the group delay tolerances above. The most serious fault is that the inventors failed to understand Ashley 1962 [4] which would suggest

$$T_{lo}(s) = \exp[-Ts] - T_{hi}(s)$$

as a better way to control distortion. It is obvious to the most casual graduate student that $T \neq \tau$. The crossover filter functions of [13] were not invented here.

X CONCLUDING REMARKS

The requirements given herein do not agree with much of the conventional wisdom of audio lore. I am scientifically correct.

Time alignment has greatly aided the sales of certain speakers but it actually does more harm than good to the sound of a speaker.

The methods I taught at ICASSP 80 for designing crossover networks are correct and the state of the art [9].

XI REFERENCES

1. P. W. Klipsch, "Modulation Distortion in Loudspeakers," J. Audio Eng. Soc. 1969 April.
2. P. W. Klipsch, "Modulation Distortion in Loudspeakers: Part II," J. Audio Eng. Soc. 1979 Feb.
3. P. W. Klipsch, "Modulation Distortion in Loudspeakers: Part III," J. Audio Eng. Soc. 1972 Dec.
4. J. R. Ashley, "On the Transient Response of Ideal Crossover Networks," J. Audio Eng. Soc. 1962 July.
5. P. W. Klipsch, "Corner Speaker Placement," J. Audio Eng. Soc. 1959 July.
6. G. L. Augspurger, "Electrical versus Acoustical Parameters in the Design of Loudspeaker Crossover Networks," J. Audio Eng. Soc. 1971 June.
7. T. A. Saponas, R. C. Matson, & J. R. Ashley, "Plain and Fancy Test Signals for Music Reproduction Systems," J. Audio Eng. Soc., 1971 April.
8. J. C. Cox, "The Minimum Detectable Delay of Speech and Music," Proc. ICASSP 84.
9. J. R. Ashley, "Group and Phase Delay Requirements for Loudspeaker Systems," Proc. ICASSP 80.
10. J. K. Hilliard, "Notes on How Phase and Delay Distortions Affect the Quality of Speech, Music & Sound Effects," IEEE Trans Audio, 1964 Mar/Apr.
11. G. C. Gillum, "A Shuffleboard Experiment of Time Delay Effects in Loudspeakers," J. Audio Eng. Soc. 1978 May.
12. S. H. Linkwitz, "Active Crossover Networks for Noncoincident Drivers," J. Audio Eng. Soc. 1976 Jan/Feb.
13. S. P. Lipshitz & J. Vanderkooy, "A Family of Linear-Phase Crossover Networks of High Slope Derived by Time Delay," J. Audio Eng. Soc. 1983 Jan/Feb.