Improved Quasi-Stereophony and "Colorless" Artificial Reverberation

M. R. SCHROEDER
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey
(Received April 19, 1961)

"Quasi-stereophony" is defined as the reproduction over two or more loudspeakers (or binaural earphones) of different sound signals derived from a single audio signal. The purpose of quasi-stereophony is to create (from a single audio signal) an illusion of spatially distributed sound sources. Quasi-stereophonic reproduction does not permit correct localization but does share with true stereophony the properties of "depth" and "ambience" which are important attributes of stereophony (for the casual listener perhaps even more important than correct localization).

This paper describes a new filtering method for producing quasi-stereophony. In contrast to earlier proposals, the new filtering method leaves the amplitude spectrum of the sound intact. The same kind of filter has also been used for generating "colorless" artificial reverberation. Experimental results indicate that both quasi-stereophony and artificial reverberation can be achieved without spectral distortion.

INTRODUCTION

An early proposal for quasi-stereophonic reproduction was made by Lauridsen.[1] Recognizing that Lauridsen's proposal amounted to filtering the monophonic signal by a pair of interleaved "combfilters" the present author[2] described alternative methods for quasi-stereophonic reproduction including the filtering of sound by means of reverberant rooms.[3]

In a recent paper Lochner and Keet[4] reported subjective results of comparisons between two-channel stereophony and quasi-stereophony obtained by filtering a monophonic signal by means of a room or a reverberation plate.

In referring to my earlier work on quasi-stereophonic reproduction, Lochner and Keet say the following: "In practice there is an intensity-vs-frequency as well as an intensity-vs-time difference in the reflection pattern set up at the two ears. This gives the impression that different frequency components arrive from different directions and probably contribute substantially to better diffusion and naturalness of reproduction. A directional distribution of frequency components alone can in fact create a certain 'stereophonic effect' as was found by Schroeder when applying spectral intensity modulations between the two ears of an observer listening to a program through earphones, using two interlocking combfilter systems, one for each ear. We do not, however, agree with him that the presence of different intensity-vs-frequency signals at the two ears is the essential prerequisite for stereophonic effects to be obtained from a single input signal."

Here I should like to add that I no longer agree with the conclusions that I drew in 1957. My point had been that, in order to create a strong spatial impression, the two filters utilized to split the single-channel sound must have different amplitude-frequency responses. I concluded this from the negative results obtained in earlier experiments with a variety of all-pass filters, i.e., filters which have *flat* amplitude-frequency responses but differ in their phase-vs-frequency responses.

After reading the Lochner and Keet paper, I resumed the quasi-stereophonic experiments with a new kind of all-pass filter with a more irregular phase characteristic. In informal tests, subjects agreed that the spatial illusion is comparable or even superior to that obtained with interlocking combfilters. Thus, quasi-stereophonic reproduction can indeed be realized with filters having flat amplitude responses and differing only in their phase characteristics.

UNDISTORTED QUASI-STEREOPHONY

The filters used in the recent experiments are of more than academic interest. Because their amplitude-frequency responses are flat, they do not introduce the "hollow" or "reedy" sound quality associated with the combfilters of Lauridsen's original proposal. A block diagram of a pair of all-pass filters[5] suitable for quasi-stereophony is shown in Fig. 1. Each filter consists of

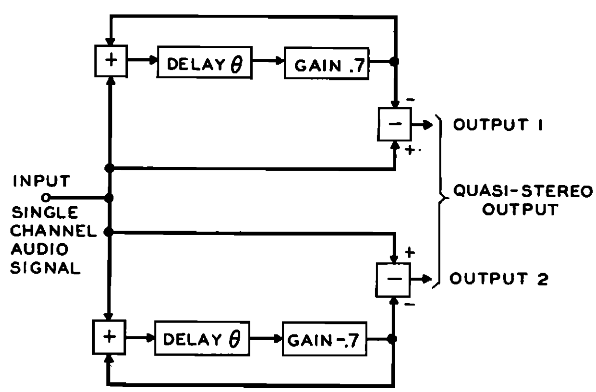

FIG. 1. Filters with flat amplitude-frequency responses, to separate single-channel sound into two signals for quasi-stereophonic reproduction.

[1] H. Lauridsen, Ingeniøren 47, 906 (1954).
[2] M. R. Schroeder, J. Acoust. Soc. Am. 29, 774 (A) (1957).
[3] M. R. Schroeder, J. Audio Eng. Soc. 6, 74 (1958).
[4] J. P. A. Lochner and W. de V. Keet, J. Acoust. Soc. Am. 32, 393 (1960).

[5] This type of all-pass filter was discovered in collaboration with B. F. Logan. Its first application at Bell Laboratories was in speech research in a (successful) attempt to reduce the peak factor of natural and synthetic speech.

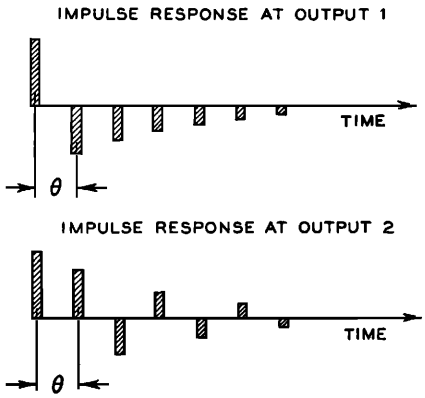

FIG. 2. Impulse responses of filters shown in Fig. 1. Decay is exponential.

a delay line with delay θ in a feedback loop with a loop gain of $\pm 1/\sqrt{2}$ and an undelayed path with unity gain. The responses of outputs 1 and 2 to an impulse supplied to the input are shown in Fig. 2.

It is not difficult to show that the transmission response from the input to either output is indeed flat. The complex frequency response from the input to output 1 is

$$H_1(\omega)=1-[(1/\sqrt{2})^1 e^{-i\omega\theta}+(1/\sqrt{2})^2 e^{-2i\omega\theta}+\cdots],$$

$$=1-\frac{(1/\sqrt{2})e^{-i\omega\theta}}{1-(1/\sqrt{2})e^{-i\omega\theta}}.$$

The following form of $H_1(\omega)$ demonstrates that its absolute value is independent of frequency:

$$H_1(\omega)=-\sqrt{2}e^{-i\omega\theta}\frac{1-(1/\sqrt{2})e^{+i\omega\theta}}{1-(1/\sqrt{2})e^{-i\omega\theta}}. \qquad (1)$$

Here both the factor $e^{-i\omega\theta}$ and the quotient have absolute value 1, the latter because it is the ratio of two conjugate complex quantities. Thus the absolute value of $H_1(\omega)$ i.e., the amplitude response at output 1, is independent of frequency.

The corresponding expression for output 2 is

$$H_2(\omega)=1-[(-1/\sqrt{2})^1 e^{-i\omega\theta}+(-1/\sqrt{2})^2 e^{-2i\omega\theta}+\cdots],$$

$$=\sqrt{2}e^{-i\omega\theta}\frac{1+(1/\sqrt{2})e^{+i\omega\theta}}{1+(1/\sqrt{2})e^{-i\omega\theta}}, \qquad (2)$$

the absolute value of which is likewise independent of frequency.

The phase lag of output 1 follows from Eq. (1):

$$\varphi_1(\omega)=\pi+\omega\theta+2\arctan\frac{(1/\sqrt{2})\sin\omega\theta}{1-(1/\sqrt{2})\cos\omega\theta}. \qquad (3)$$

Similarly, we obtain the phase lag of output 2 from Eq. (2),

$$\varphi_2(\omega)=\omega\theta-2\arctan\frac{(1/\sqrt{2})\sin\omega\theta}{1+(1/\sqrt{2})\cos\omega\theta}. \qquad (4)$$

From Eqs. (3) and (4) we obtain the important envelope delay difference $\Delta\tau$ between outputs 1 and 2.

$$\Delta\tau\equiv\frac{d\varphi_1}{d\omega}-\frac{d\varphi_2}{d\omega}=\frac{4\sqrt{2}\cos\omega\theta}{1+8\sin^2\omega\theta}\cdot\theta, \qquad (5)$$

which is plotted in Fig. 3 as a function of frequency, f. It should be noted that $\Delta\tau$ is not the delay difference for any steady single sinusoid but rather the delay difference for the envelopes of transients, narrow groups of sinusoids, or narrow bands of noise.

The envelope delay difference oscillates with a "period" of $1/\theta$ about zero. Maxima of $\Delta\tau$ occur for $f=n/\theta$, where $n=0, 1, 2, \cdots$. For these frequencies $\Delta\tau$ equals $4\sqrt{2}\theta$. For $\theta=5$ msec, the maximum delay difference is 28 msec which is sufficient to make a narrow band of frequency components in the neighborhood of a frequency $f=n/\theta$ appear to be coming from the earphone or loudspeaker[6] connected to output 2. Conversely, for frequency bands around a frequency $f=(n+\frac{1}{2})/\theta$, the sound seems to emerge from output 1. For frequency bands around $f=(n+\frac{1}{4})/\theta$ or $f=(n+\frac{3}{4})/\theta$, the delay difference is zero and the sound image is central.

Whether the directional perception of complex sounds whose frequency components cover a wide band is liable to the above analysis in terms of narrow frequency bands is, of course, not assured *a priori*. In fact, it would be quite unreasonable to assume, for example, that for a 2-cps spacing of adjacent peaks of the $\Delta\tau$ response, the perceived directions when listening to a complex

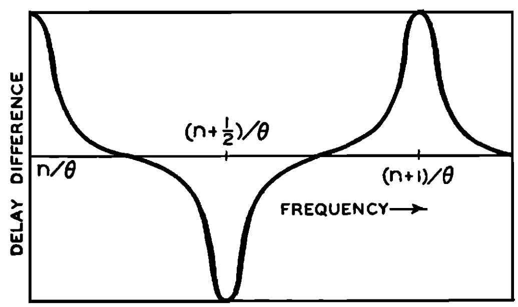

FIG. 3. Envelope delay difference between outputs *1* and *2* as a function of frequency. The envelope delay difference oscillates around zero with a "period" equal to the reciprocal of the loop delay θ.

[6] When reproducing quasi-stereophony over loudspeakers, the sound pressure at each ear will be a mixture of the two quasi-stereophonic signals. The result is a somewhat diluted spatial effect. However, this disadvantage can be overcome by radiating the *sum* of the two quasi-stereophonic signals from a central loudspeaker facing the listener, and the *difference* from a loudspeaker rotated by 90° and open at the back. In this manner, a good approximation of the proper quasi-stereophonic signals can be generated at the listener's ears within a limited listening area (see also reference 1).

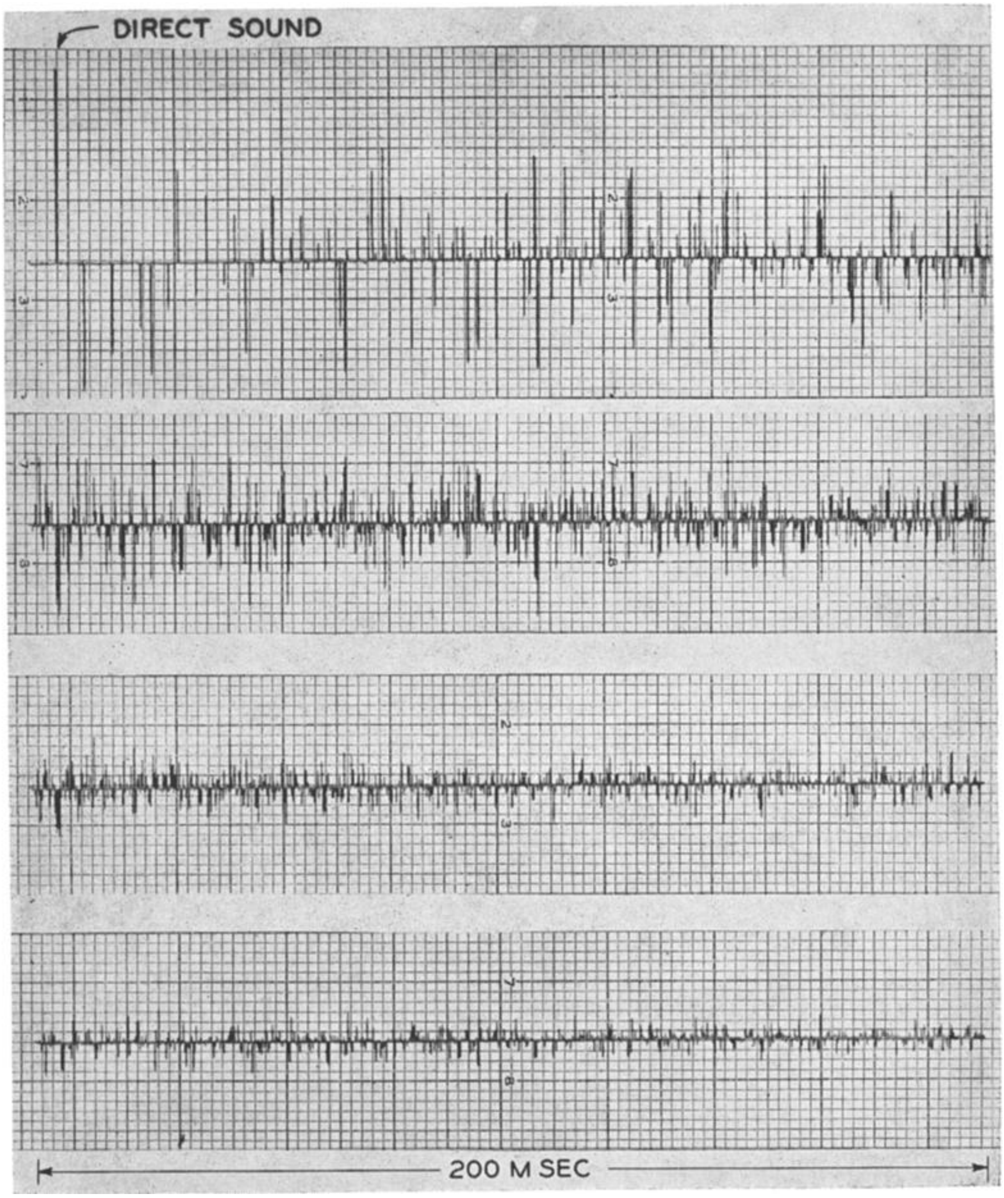

FIG. 4. Impulse response of all-pass reverberator consisting of five all-pass filters connected in tandem.

sound would alternate every 1 cps. Rather, the dominant percept in this case, corresponding to a delay $\theta=500$ msec, is a succession of distinct echoes which are alternately in and out of phase at the two ears (see Fig. 2). In other words, for very long delays θ the subjective response is based not on the frequency domain description of the filters ($\Delta\tau$– response) but on their time domain description (impulse responses). However, the fractional frequency band analysis appears to describe the perceived sound image quite well for relatively broad peaks of the $\Delta\tau$ response corresponding to delays $\theta\leq 5$ msec.

What portions of a complex sound signal appear to be coming from or near the center depends on the flatness of the $\Delta\tau$ response at its zero crossings, the degree of flatness being determined by the delay θ, and the loop gain[7] of the feedback loops around the delay lines (see Fig. 1): The higher the loop gain, the more peaked the $\Delta\tau$ response. For a given maximum value of $\Delta\tau$, this means *more* center image because $|\Delta\tau|$ is relatively small for more frequencies. By adjusting both the delay θ and the loop gain, a fairly uniform spatial distribution of sound images can be achieved.

The optimal magnitude of the delay θ may vary with the kind of program material. We have found delays as short as $\theta=5$ msec to give very pronounced spatial effects. For delays in excess of 50 msec, considerable

[7] For open loop gains g other than $\pm 1/\sqrt{2}$, the gain of the undelayed path has to be modified to $g^2/(1-g^2)$ to retain a flat response. I shall leave the proof to the interested reader that this gain for the undelayed signal results in a flat frequency response. A possible proof follows the same simple pattern as the one for $g=\pm 1/\sqrt{2}$ given in the text.

reverberation will be added to the input signal. In fact, a delay of θ corresponds to a reverberation time of $T_{60}=20\cdot\theta$. (For each delay θ the signal is attenuated by 3 db. Thus, the delay for 60-db attenuation is $20\cdot\theta$.) This observation leads to another interesting electro-acoustic application of all-pass filters.

"COLORLESS" ARTIFICIAL REVERBERATION

In addition to being useful for quasi-stereophony, all-pass filters of the kind shown in Fig. 1 are ideal for generating artificial reverberation.[8] They exhibit the required exponential decay with time, and because their frequency responses are flat, they do not add the unpleasant hollow, reedy, or metallic quality to the reverberated sound associated with other artificial reverberators employing delay and feedback.

B. F. Logan and the writer[9] have built artificial reverberators by connecting several all-pass filters of the kind shown in Fig. 1 in tandem using incommensurate delays θ to insure a sufficiently random time response. Figure 4 shows the impulse response of the series connection of five all-pass filters with delays between 100 and 6 msec. In tests with white Gaussian noise, no difference could be detected between the input and the output of the reverberator. The high-echo density resulting from the tandem connection of several filters assures a smooth and flutter-free response even to sharp transients. Further details of "colorless" artificial reverberators are described in a forthcoming paper.[10]

[8] According to a communication by W. Kuhl (Rundfunktechnisches Institut, Hamburg), the application of all-pass filters to artificial reverberation was independently suggested by J. J. Geluk of the Nederlandse Radio Unie.

[9] M. R. Schroeder and B. F. Logan, J. Acoust. Soc. Am. **32**, 1520 (A) (1960).

AMBIOPHONY

By sending single-channel sound through many parallel all-pass filters with separate outputs and feeding the filter outputs to loudspeakers distributed over the walls and ceiling of an auditorium or living room, a highly diffuse "stereo reverberation"[11,12] or "ambiophony"[13] can be achieved. If, in addition, acoustic feedback is absent or minimized,[14] the reverberated sound can be given the same spectrum as the original, or any desired timbre, thereby making ambiophonic installations compatible with the exacting quality requirements encountered in the electroacoustic conversion of auditoriums and theaters designed primarily for speech into concert halls.

ACKNOWLEDGMENTS

I am indebted to E. E. David, Jr. and B. F. Logan for fruitful discussions and to F. K. Harvey and A. J. Prestigiacomo for experimental assistance.

[10] M. R. Schroeder and B. F. Logan, J. Audio Eng. Soc. **9**, 192 (1961).

[11] R. Vermeulen, Philips Tech. Rev. **17**, 258 (1956).

[12] R. Vermeulen, J. Audio Eng. Soc. **6**, 124 (1958).

[13] D. Kleis, Philips Tech. Rev. **20**, 309 (1959).

[14] M. R. Schroeder, Radio Electronics **31**, 40 (1960); see also "Improvement of acoustic feedback stability in public address systems," in *Proceedings of the Third International Congress on Acoustics* (Elsevier Publishing Company, Amsterdam, 1961) (to be published).