ROBERT W. CARVER *

— SONIC — HOLOGRAPHY

The illusion presented by conventional stereo sound is, in my estimation, severely flawed by fundamental distortions of space and time. Although we accept it because we must, and love it when we can, I don't think any of us really believes this illusion approaches a real-life experience. By "conventional stereo" I mean the sort of presentation that works mostly or entirely through loudness or amplitude manipulations. For a sound supposed to come from the left, the left speaker dominates in amplitude; for a sound from the right, the right speaker plays louder. The speakers are equally loud for a sound intended to come from exactly in between, and appropriate degrees of louder or softer are employed for "somewhat left" and "somewhat right" sounds. *Interaural amplitude difference* or IAD is one common name for these manipulations we use to fool the ear and brain. Thus, if the left ear picks up a sound as being significantly louder then the right ear, the ear-brain mechanism will locate the sound source to the left. The converse is true when amplitude favors the right ear. If both ears detect no amplitude difference in what they separately hear, the ear-brain "visualizes" the sound source as being directly in front of the listener.

A classic acoustical experiment, made some years ago, vividly illustrated the power of this mechanism. Listeners were exposed to identical and simultaneous impulse "pips" coming from two spaced loudspeakers, one of which was a little closer to the listeners so that its output arrived slightly ahead in time. The listeners perceived both sound sources as one, and identified "it" as coming from the near speaker. But when the near speaker was turned down in level and the far one up, it was possible to make the pips shift gradually over to the far speaker, showing that amplitude cues could be made to override temporal cues in the ear-brain's localization process. It took quite a difference in amplitude to bring this about, however, and listeners also noted that the sonic character of the pip changed with the shift, becoming somehow fuller and more diffuse.

*President, Carver Corp., Woodinville, Wash.

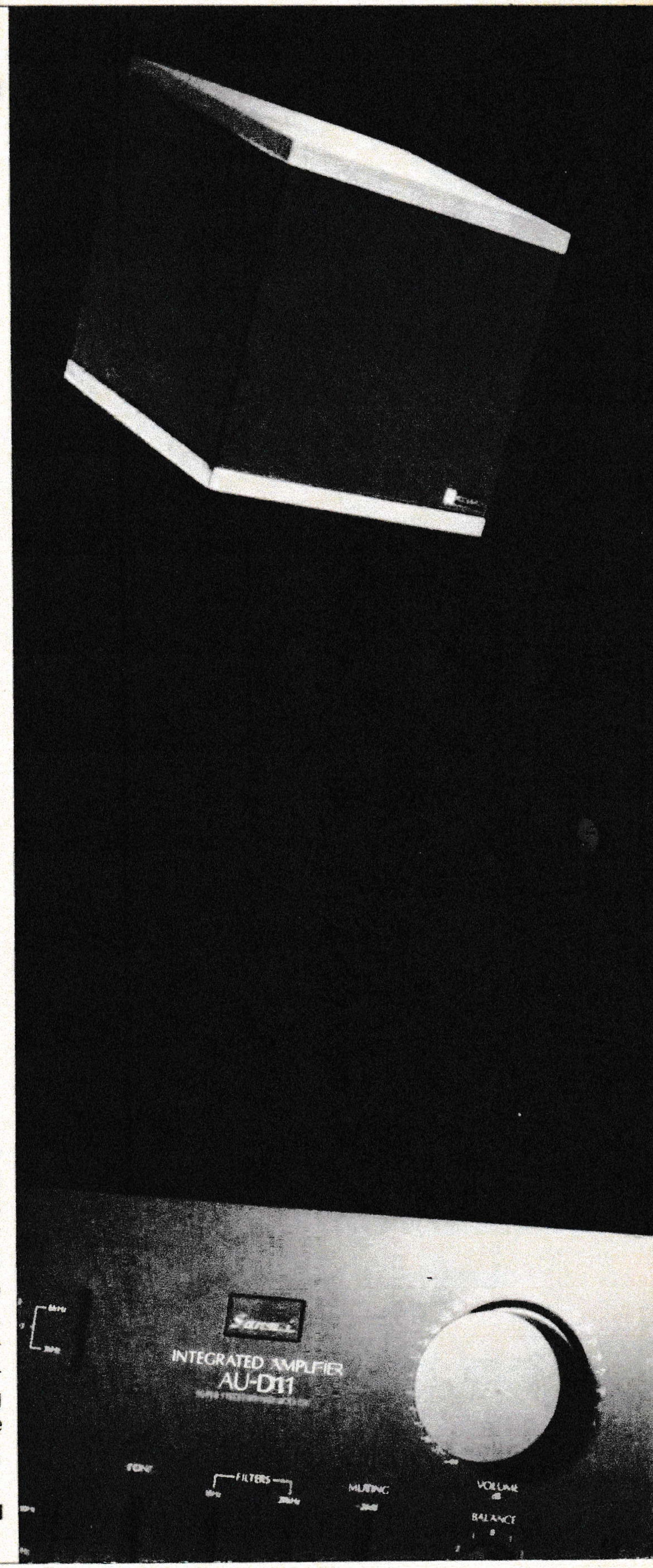


Photo: James Wojcik

Nevertheless, the mechanism can persuade the ear-brain that it heard sounds from places other than their true origins, and it is the basis for what we call stereophony. It works well enough to be the basis for most all sound recording today, and especially for the creation of records from multitrack tapes.

Just because this trick could be performed, it was easy to assume that we had unlocked the whole of nature's secret to ear-brain localization. But this was wrong because the ear-brain functions with several localization mechanisms, which can reinforce or conflict with each other. It may be true that amplitude cues such as IADs are important and even dominant for close sounds. But because of the inverse-square law, IADs reaching a listener seated back in an auditorium are negligible, even for first-arrival or direct sounds. As Fig. 1 shows, IADs can be expected to reduce to about 0.4 dB at a 16-foot distance from the sound source — an amplitude difference that would have had virtually no effect on localization of the pip in the above experiment.

On the other hand, although IADs rapidly diminish with distance, temporal cues, that is, interarual *time* differences (ITDs), remain absolutely unchanged for a given sound-source direction, no matter what its distance. We appear to have an internal clock that ticks steadily and reliably away under all circumstances, timing the interval between a sound's arrival at ear 1 and ear 2 (see Fig. 2). Therefore, it seems reasonable to conclude that for a typical audience member at a concert, ITDs have great importance in localization of sounds, but IADs mean next to nothing.

Another localization factor often cited is acoustic "shadowing" of the far ear by the head and face. Shadowing alters the frequency spectrum perceived by the far ear, presumably in a way that is informative to the ear-brain about sound-source location. Unquestionably this is an important factor, but probably more so in a nonreverberant environment such as occurs outdoors, where the head casts a truly steady and well-defined shadow. In an auditorium the far ear is not shadowed from the all-enclosing reverberant field, so the ear-brain has only the ITD period associated with the first-arrival sound to sense shadowing. Since determining spectrum shape requires a much longer sample of sound than ITD discrimination, it's by no means certain that shadowing is quite the aid to localization in concert halls it might be otherwise. But that doesn't mean shadowing won't have to be considered in the context of home listening rooms.

For purposes of source localization then, IADs are probably of very small consequence in auditoriums. Shadowing may have some consequence, but it's hard to assign an absolute or even typical rating of influence. ITDs, we know, are fixed and always present, and they seem to be of central importance.

A Matter of Time

Ironically, however, it is timing or temporal cues that get the shortest shrift in conventional stereophony. Multi-microphone techniques pretend they don't exist. There is no spaced-microphone array that rationalizes them properly for loudspeaker listening. Coincident microphones eliminate them totally, a shame because they're sorely missed by the ear-brain. Perhaps most seriously, stereo loudspeakers ultimately make a mess of the time factors by adding new and utterly inappropriate temporal cues.

Figure 3 shows how this happens. Let's take, as an arbitrary example, a "phantom" sound image that has been panned so it seems to come from directly between the two speakers. Equal-amplitude sounds travel from left speaker to left ear along the path L and from right speaker to right ear along path R. The ear-brain's IAD detector is satisfied and shouts "Sound dead ahead." For an instant the ITD sensor is satisfied too. But 100 microseconds or so later, sound traveling from the left speaker along path L reaches the right ear simultaneously with the right-speaker sound's arrival at the left ear along path R. At this point the ear-brain's ITD sensor whispers "Not a single sound dead ahead, but *two* equal sounds equally spaced to right and left," and there is confusion at the helm.

Ultimately, your conscious mind will probably accept the decision of the IAD sensor, for otherwise, stereo wouldn't work at all. But there is still the situation of conflict in which one localization mechanism is trying to "fuse" two separate sounds into a single sound while another keeps insisting that it just isn't so.

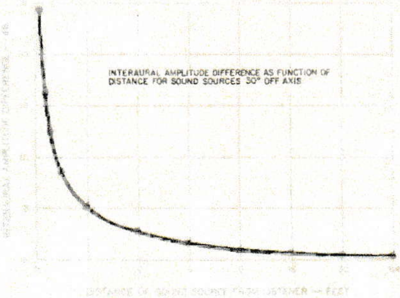


Fig. 1A—Interaural amplitude differences diminish rapidly with increases in distance from the sound source.

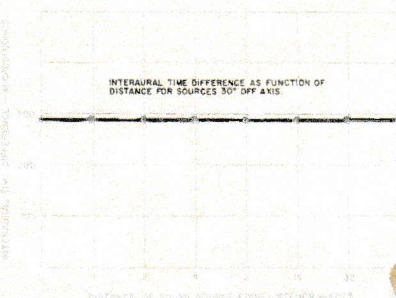


Fig. 1B—Interaural time difference remains constant whatever the subject's distance from the sound source.

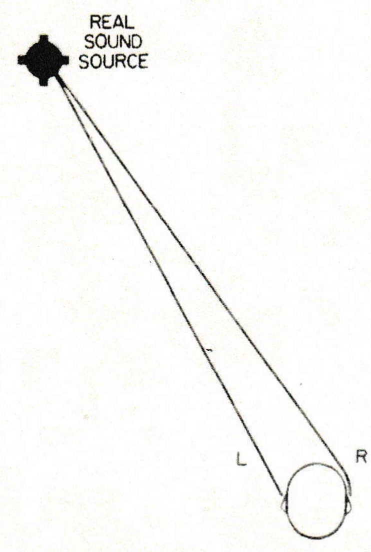


Fig. 2—How the sound from a real source, located to the listener's left, approaches the listener's ears.

Evidently, ITDs can be bullied by stereo techniques into letting IADs have the final say, but this is no victory for the integrity of our perceptions. ITDs are basic to our sound-localization ability, not secondary or supplementary. If ITD sensors are to be satisfied, they must hear two and only two arrivals from a single sonic event, and not the four that occur during conventional stereo playback. Anything else is nonrealistic and subject to deep suspicion if not outright rejection.

Is it any wonder that, if we're honest with ourselves, we never find our stereo systems sounding nearly as convincing and credible as the real thing? There is, in the end, an indistinctness, flatness, and diffuseness (if you recall the listener impressions in the experiment described earlier) that never quite go away unless certain brain centers have been lulled into a mood of acceptance. It is not so much as if something were missing, but more as if something spurious has been added. I submit that the "something" consists of two extra sound arrivals that never exist in real life and that result in a conflict of localization cues and a consequent indistinctness of time and place in the stereo image.

Sonic Holography Defined

Let's look back at Fig. 3 again. Remember that when we had just paths L and R, both the IAD and ITD sensors were satisfied. For a magic 100 microseconds or so, everything was fine. But then L_r and R_l intruded, and the "one source versus two" dilemma arose. Suppose it were possible to eliminate L_r and R_l, not just for a center-front image, but for sound images occurring anywhere on the stereo stage — to make it as if the left speaker didn't exist for the right ear and vice versa? Might this help? And if it does, how could it be accomplished?

First of all, it clearly could not be done by any electronic method alone. It would require an electronic method designed to invoke the assistance of some other mechanism, specifically, the mechanism of wave interference. We know that sound waves interfere in the air of any environment in which sound is contained, and that the interference gives rise to local reinforcements and cancellations of sound throughout the area. We're used to thinking of such interference as being, if not unpredictable, at

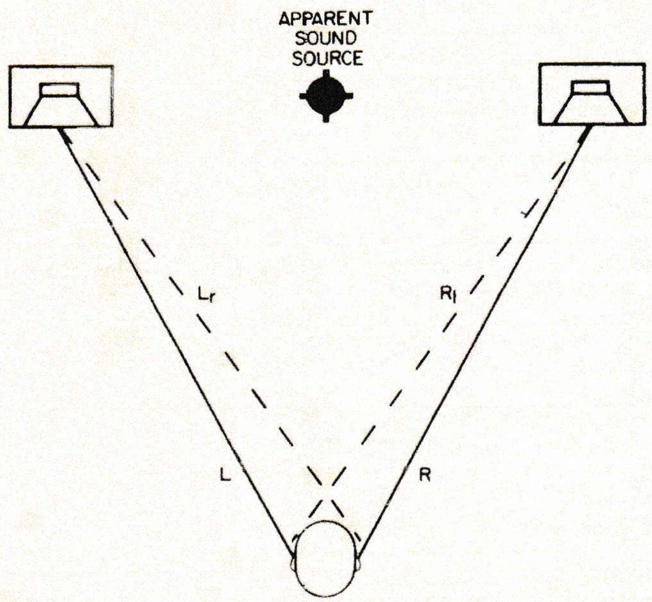


Fig. 3— How conventional stereophony uses equal amplitude signals to produce a stage-center location for the apparent sound source. But note that both speakers are heard by both ears.

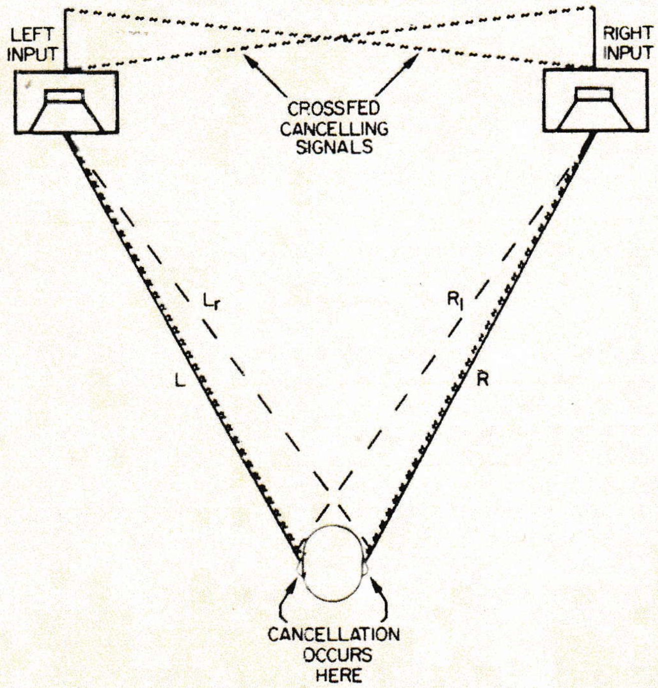


Fig. 4— Careful addition of the crossfed cancelling signals will cause the L_r and R_l signals to drop out of the sound field at a selected listening location.

least too complex to be controllable. Yet while this may be essentially true for much reflected sound, it is not true for sound proceeding directly from a sound source to a prescribed destination. In this case, we know the sound's point of origin, its speed, and how far it has to go. Therefore, it's theoretically possible to arrange for an interfering sound to meet the unwanted sound at its destination and, by acoustical interference, cancel much or all of it. If the sound's origin is the left speaker, its path L_r, and its destination the right ear, cancelling that sound upon its arrival should make the left speaker disappear for the right ear. And the same thing should be possible for the left ear vis-a-vis the right speaker.

What is required of the cancelling signal? First of all, it should be appropriately derived from and related to the unwanted signal coming along L_r, but 180° out of phase with it. Second, it must arrive at the ear the instant the L_r signal does. Finally, it must come from some other source than the left speaker, or the signals would cancel at the speaker itself and nothing would come out. The obvious choice is projecting it from the right speaker. This turns out to be the way to do it. Conceptually, all we need do is tap off part of the left-channel signal at some line-level stage, invert its phase, adjust it temporally for simultaneous arrival with the L_r signal at the right ear, and then crossfeed it to the right channel for launching through the right speaker. A portion of the right channel is tapped off and processed in the same way to cause cancellation of the R signal at the left ear. Figure 4 shows the total result.

The temporal adjustment needed is about a 100-microsecond delay (125 microseconds, to be exact), which is nothing more than the ITD between the left ear, which receives sound along the shorter path L_l and the right ear, which hears the left speaker along the slightly longer path L_r. Needless to say, the listener's ears must be in the approximate spots the whole process expects if the proper cancellations are to take place right at them. However, although such interference effects are relatively local, position is not quite so critical as you might expect. So long as the listener's seat is where it should be, he can move his

SONIC HOLOGRAPHY RECORDED

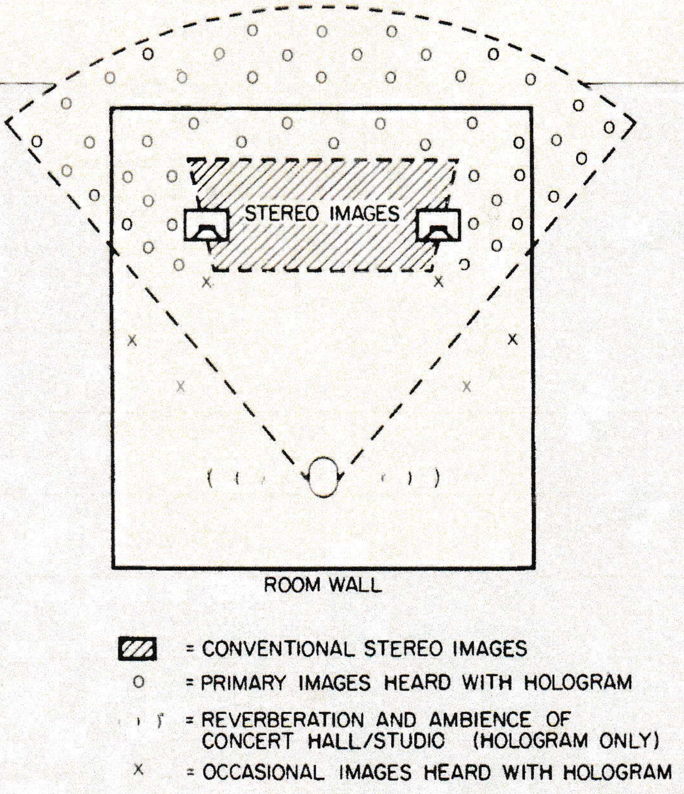
Sonic holography happens because of acoustical interference patterns set up in the air around a listener's head, but the signals that bring about holography are not so intangible. They can be detected, quantified, analyzed, and replicated. They can even be recorded — on a tape by you, if you have a Sonic Hologram Generator, or on an ordinary phonograph record such as the Soundsheet bound into this magazine. Playing the tape or record produces the full holographic effect. A Sonic Hologram Generator is needed only for material that has not been holographically encoded.

Nevertheless, just playing the enclosed Soundsheet record may not do a thing for you unless you take steps to ensure that your listening chair (and you should use a chair) and your loudspeakers are properly positioned relative to each other, and sufficiently far from walls and other reflecting surfaces that might impair the propagation of the holographic image. Remember that you're trying to achieve the intersection of your head and two relatively small regions of acoustical interference. Special pains taken in positioning — at least at first — will not go unrewarded.

The Soundsheet record has been especially recorded in sonic holography and will play back on a normal, conventional stereo system to produce the full sonic hologram sound image. To set up your system for sonic holography, follow these step-by-step instructions. You will first need to obtain a steel tape measure, length of nonstretchable rope or speaker wire for measuring distance, and a temporary listening chair that may be easily moved about. Inspect Fig. B1, which represents a top view of a room, two stereo speakers, and a listening chair. The first step is to temporarily move your loudspeakers out and away from all walls and reflecting surfaces, as shown in the diagram. Next, space the speakers relatively close together (three to five feet, center to center, is good). Using the nonstretchable wire, rope or tape measure, carefully measure the distance from the left speaker to the center of the listening chair. Repeat the measurement for the right speaker, and make both distances (d1 and d2) *exactly* the same. Accuracy to within ⅛ inch is desired; the actual distance themselves are not too critical as long as d1 and d2 are equal.

The Soundsheet record is too flexible to be supported by a ribbed or contoured turntable mat, so support it with a conventional LP. Play the record, and listen for the following: (1) A performance extending in an arc in front of you, with a spread ranging from 45° to 90°; (2) sound images that normally spread beyond the boundaries of the left and right speakers, and occasionally turn up at full-left and full-right positions, but without instabilities or elusiveness; (3) an augmented sense of front-to-back depth, with sound images occurring well behind the speaker plane and occasionally in front of it, and (4) a palpable spatial dimension and clarity to the performance, suggestive of flesh-and-blood musicians occupying actual space.

Turning your head or otherwise moving it should not materially affect your perception of the above, so long as you remain seated in the listening chair. Getting up and leaving the chair should diminish the experience to no more than what you're used to from conventional stereo. If you're not certain you're hearing the full holographic effect, or that you notice a drastic diminution of it when you leave the chair, review the accuracy of the chair and speaker positionings and try again.



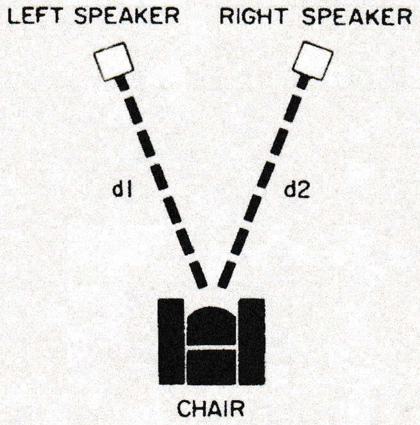


Fig. B1—Setup for sonic holography. Make d1 and d2 as close to equal as possible.

Once you've got a "fix" on the holographic experience, begin gingerly inching your speakers back toward their original locations to find out just how far you can go before unacceptable loss occurs. Be sure to keep d1 and d2 equal at all times.

One final note. The Soundsheet record is plainly not the sort of vehicle to raise expectations of ultimate sonic excellence, and indeed nothing very spectacular has been attempted or achieved on it. Thus you may find that fidelity limitations require a few mental adjustments on your part before you can appreciate the substance of what you're hearing, but you'll probably be surprised at how few will be necessary.

Sonic Holography in Detail

The block diagram, Fig. B2, shows the basic configuration of the Sonic Hologram Generator, with the various operators indicated. The injection-ratio block is actually switchable between nominal values of -3.5 dB relative to the main signal, which is the theoretically correct value for the interference signals, and -5 dB, which has occasionally been found subjectively preferable for certain installations and recordings — particularly those employing "manufactured" stereo sound from multitrack masters.

The delay introduced in the crossfeed lines is essentially 125 microseconds, but two designs incorporating the Hologram Generator permit two additional delays of 115 and 95 microseconds to be switched in as well. Working together, the three delays make the position of the ideal "stereo seat" somewhat less critical so as to accommodate other listeners.

The spectral-shaping circuitry has not been mentioned and certainly deserves an explanation. Shaping is performed both on the cancellation-signal channels and on the main channels. The contours employed are shown in Fig. B3. They were determined empirically, but once established they clearly fulfilled certain requirements that might have been predicted beforehand. At least three factors are involved:

Factor One. One of the great flaws of conventional stereo is that signals L and R_I, and R and L_I as well, as shown in Fig. 3, unavoidably interfere at the point where they are listened to, and one inevitable result is a change in perceived spectrum, which becomes different from what is actually embodied in the left and right channels of the recording. Often the recording engineer hears this when he listens to the master tape (in stereo, of course) in the process of mixing it down to final form, and he'll often attempt to "fix" it with equalization. If so, the holographic generator must "unfix" his fix, because sonic holography does not suffer such interference effects and needs no compensation for them.

Factor Two. A number of complex things happen when sound wavelengths shorten to approach the human-head dimension that produces TDs — things that conventional

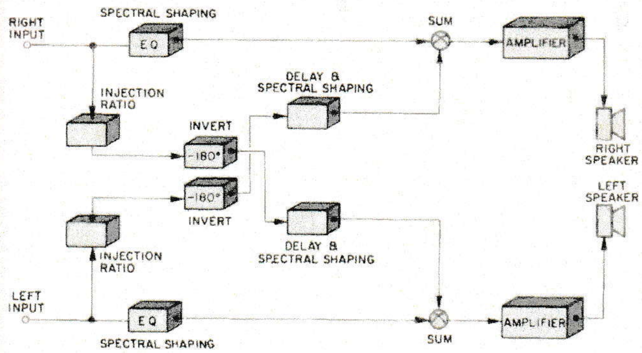


Fig. B2—Basic block diagram of the Sonic Hologram Generator.

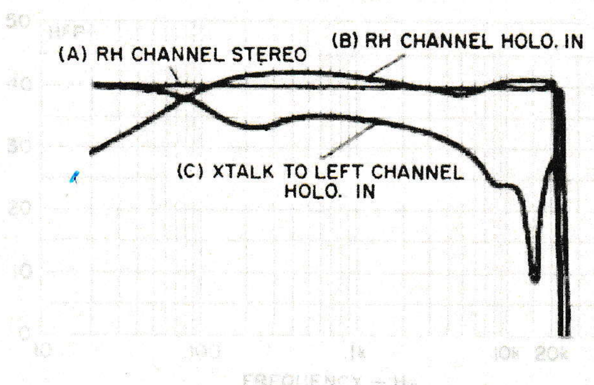


Fig. B3—Frequency shaping in the Hologram Generator: (A) Response in standard stereo, (B) contouring on main channel through Generator and (C) relative level and contouring on cross-fed signal. Chart recorder causes roll-off above 20 kHz.

stereo has never properly taken into account. First, the L and R_l (and R and L_r) interferences tend to become additive rather than subtractive, and an appreciable peak forms. Second, acoustic shadowing of the far ear by the head starts coming into its own. (You'll remember we underplayed the importance of shadowing in the reverberant-field conditions of an auditorium. However, in the direct field of nearby loudspeakers, shadowing is certain to assume a more significant role.) And third, the ear rises to its point of maximum sensitivity. The chosen contours deal with these matters in combination, in a way that properly corresponds to the spectra generated by real-life ("nonphantom") sound sources.

Factor Three. We all know and despair of what happens when deep bass is reproduced in a home listening room of typical size. The long wavelengths, confined within close room boundaries, set up large-scale interference patterns ("standing waves") with sizable fixed regions of reinforcement and cancellation. Since sonic holography works by controlling interference patterns in the vicinity of the listener at *all* wavelengths, it represents a golden opportunity to attack this problem at its roots. The low-frequency contours chosen for the main and cancellation signals are particularly effective in subduing the speaker-to-rear wall room mode — the most troublesome one in the majority of stereo installations. Note carefully that, despite casual appearances, the contours as shown do actually sum vectorially to flat response.

head around freely without risking serious anomalies. Two listeners cannot occupy the same seat, of course, but even a listener a bit removed from the prime stereo seat will experience some of the benefits, while listeners well removed will experience nothing worse than conventional stereo heard under those conditions.

I call the results of this acoustical mathematics "sonic holography" because of its obvious similarity to the sort of wave interference that creates an optical holographic image. In terms of what is perceived, there are other parallels between the sonic and optical images as well.

Experience of Holography

Those who have heard sonic holography properly presented report an impression of realism and palpability, with a good sense of depth, three-dimensionality and solidity imparted to the performance. Sound localizations occur behind, in front of, and beyond the plane between the two loudspeakers, and seem entirely natural when they do so. The sound becomes dissociated from the listening space, seemingly transported back to its original environment. Bass seems less troubled by room boundaries, and it seems to be spreading itself throughout a much larger volume of space. The experience is beyond the usual stereo but is clearly related to the best that conventional stereo can do.

Here I should anticipate a question that has probably occurred to many readers: Does sonic holography help with the many temporal problems that inevitably are built into the average recording? I don't think it does. They remain intact, just as undesirable as ever, awaiting some other kind of solution. But does that mean sonic holography will not make these undesirable records more listenable? Emphatically not.

It's my impression that many phase or time difficulties encountered in recordings are not fixed disfigurements that persist throughout both sides. Rather, I think they shift around or come and go with a sort of randomness that makes them less than blatantly obtrusive — or if they were blatant, they probably got tidied up a bit in the mix. But the problems that sonic holography aims to cure (and they have both temporal and spectral aspects) are fixed by the geometry of the home listening situation, and as such can be thought of as extreme and permanent colorations that influence everything coming through the sound system. Alleviating them can only improve the sound of everything that passes through the system, no matter how good or bad it intrinsically is.

Work on sonic holography was begun several years ago because the intent of conventional stereo to dissociate the stereo image from the loudspeakers that were the actual sound sources seemed fundamentally unrealizable. With some recordings, played under some circumstances, you could *convince* yourself that there was genuine depth, spread and space conveyed by the image and that the loudspeakers were no more propagators of the sound than were the microphones used to pick it up originally. But a certain amount of convincing was always necessary.

Sonic holography grew as a research project when we realized that the underlying theory of stereo reproduction needed amendment if a convincing and consistent sense of space, with integrity, was to take place. The sense of space came fairly easily; there are a number of signal processors now on the market that afford it. But consistency, integrity and credibility were harder to achieve. In the end they could be gotten only with a linear device, behaving in a theoretically correct way, that addresses itself to the several crucial faults in conventional stereo and leaves everything else pretty much alone.

Conventional stereo is a deliberate illusion, and sonic holography is a deliberate illusion. Neither one is a replica of what we experience in the real world. But by all psychoacoustical considerations, I believe sonic holography comes closer to reality and I hope you can appreciate that from the theory. But try listening to it yourself.

Editor's Note: We wish to thank Bob Carver and the Carver Corp. for the technical and financial assistance which made possible the inclusion in this issue of an Eva-Tone Soundsheet demonstrating Sonic Holography. Additional thanks to RealTime Records and the Roger Wagner Chorale for the music "Encore," RT-110, and to Boston Lettershop for the music "Chords."—E.P.