History and Development *of* Stereophonic Sound Recording

Reprinted by permission of Journal of the Audio Engineering Society, issue April, 1953

By Ross H. Snyder
Ampex Electric Corporation, Redwood City, California

The pioneering work on stereophonic reproduction by Dr. Fletcher and the Bell Laboratories is described. Consideration is given to the relative importance of frequency range, noise, and distortion in monaural *vs* multichannel reproduction. Distinction is made between binaural and stereophonic reproduction. The realization in practical form of the requirements laid down by early research through the use of magnetic tape recording is described. Phase relations are analyzed, and requirements for the preservation of three-dimensional effect are defined.

The search for "facsimile" reproduction of sound has gone on for more years than the recent mushrooming of the high-fidelity industry might at first indicate. It is, indeed, startling to realize that facsimile was accomplished as long ago as 1933, although this was under conditions which would be difficult if not impossible of duplication on a commercial scale, even with today's facilities. Facsimile may be said to have been achieved when subjective judgment by a significant number of persons of normal hearing gives no better than chance distribution—50%—of correct distinctions between the live, original sound and the reproduced sound.

However acceptable as a *substitute* for the real thing, reproduction which is distinguishable from the original by a significant proportion of listeners must be accorded no higher rank than high fidelity.

The requirements for crossing the sharply defined line from high fidelity to facsimile were laid down more than ten years ago by Dr. Harvey Fletcher.[1] These were, in substance (1) transmission noise significantly below the ambient noise in the receiving location, (2) dynamic range sufficient to reproduce a level upward of 100 acoustic db[2] peak at the listening chamber, in the case of symphonic music, (3) nonlinear distortion due to transmission characteristics below the level of detectability, (4) frequency range equal to or exceeding that of the human hearing apparatus, which is limited by the peak listening level, and (5) preservation of spatial orientations.

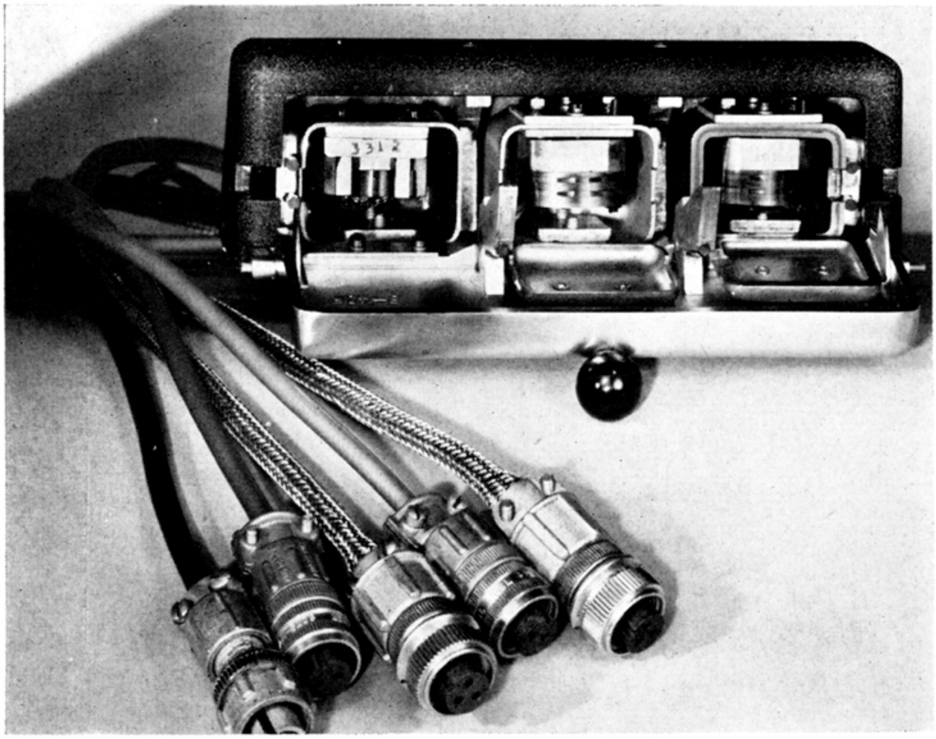
It is interesting to note in the classical studies which emerged from the Bell Laboratories during the late 1920's and 1930's that ambient noise at the point of listening so very commonly exceeds that which is routinely characteristic of the auditoriums in which music originates.[3] In the world's finer auditoriums, it is not uncommon to find ambient levels as low as 20 acoustic db, and levels of 40 to 50 are common in the living rooms in which ear is given to reproduced music, so that the quieter passages in the music may often be expected to be masked by noises entirely extraneous to the program material or to the transmission system.

It is also important to note that in 8 hr of monitoring the acoustic levels produced by the Philadelphia Orchestra,[1] playing a very wide variety of selections, an acoustic range of 65 db was exceeded only twice: once in the tremendous crash of sound which occurs toward the end of Stravinsky's *Fire Bird*, at a point where both the cymbals and the bass drum are sounded simultaneously at their maximum loudness, and once in the midst of an orchestral transcription of Debussy's *Clair de Lune*, when a long hush comes over the whole orchestra. Fletcher estimates that the one great crash reaches an acoustic level of 112 db peak, basing his estimate upon theoretical development of the evidence contained in the Sivian, Dunn, and White[4] papers on the power and frequency relationships existing in the output of musical instruments, in which momentary intensities of very high level were reported producible by these two instruments.

The point at which nonlinear distortion becomes perceptible has been reported in the papers of Dr. Harry Olson.[5] Although dependent upon the character of the music and of the harmonic structure of the nonlinear distortion, it may be taken as somewhat under 1% on a total rms basis. Fletcher conclusively demonstrates that the frequency range needed for facsimile duplication of sound depends firmly upon the levels of listening: when symphonic levels do not include the extreme peaks already mentioned, but do include the peaks more usually encountered, a range of 40 to something short of 15,000 cps is all that the ear can detect. This is because the extremes, below 40 and above 15,000 cps, are audible to persons of normal hearing only at levels approaching the threshold of pain. Fletcher goes so far as to state that "substantially complete" reproduction of orchestral music may be accomplished within the range of 50-8,000 cps.

The fifth consideration turned out to be crucial: the preservation of spatial orientation. However many duplicate reproducers may be provided as termination for a single transmission channel, in order to "spread out" the apparent sound source, no important number of observers were ever deceived into mistaking the copy for the original. In 1933 at the Chicago World's Fair, however, a binaural system, produced under laboratory conditions and presented under optimum circumstances, finally achieved the desired deception. Carefully fitted ear probes were used, of which the listeners could hardly be unaware, but the comparison with the live sound was unanimously declared to be nearly perfect. It remained, in that year, for the Bell Laboratories to duplicate the feat with loudspeakers, using a system of three

1 Harvey Fletcher, Hearing, The Determining Factor for High Fidelity Transmission, *Proc. Inst. Radio Engs.*, **30** 266-277 (1942).
2 Decibels above the reference level 10^{-16} w/cm^2.
3 D. F. Seacord, Room Noise at Subscribers' Telephone Locations, *J. Acoust .Soc. Amer.*, **12** 183-187 (1940); D. F. Hoth, Room Noise Spectra at Subscribers' Telephone Locations *ibid.*, **12**, 499-504 (1940)
4 L. J. Sivian, H. K. Dunn, and S. D. White, Absolute Amplitudes and Spectra of Certain Musical Instruments and Orchestras, *J. Acoust. Soc. Amer.*, **1**, 172 (1930); **2**, 330-371 (1931).
5 H. F. Olson, *Musical Engineering*, pp. 345-348, D. Van Nostrand Co., New York, 1952.
6 *Elec. Eng.*, **53**, No. 1 (1934); six articles.



The Magnetic erase, record, and playback heads of a commercial two-channel stereophonic tape recorder.

rather than two channels, and labeled "stereophonic."[6] The theoretical development for the two different systems is reviewed in a Bell System monograph of 1934.[7] Briefly, binaural and stereophonic systems attempt to accomplish somewhat the same ends by different means. Binaural transmission assumes the placing of a substitute for the human head, containing two microphones separated by a sound-absorbent material, at some position in the origination chamber at which, it is judged, the listener might most like to be. Separate channels are then provided for the output of the microphone, all the way to the ears of the listener, where the transmission is again made audible, still without intermixing the signals; this necessitates the use of earphones or, perhaps, standing wedged within a soundproof wall, one loud speaker on each side, thus effectively confined to one ear each.

Stereophonic reproduction proceeds from the theory that, if a very large number of microphones were situated in a three-dimensional region where sound is originating, each with its own transmission system to another location, acoustically identical with the originating auditorium, and each terminated there in an ideal reproducer located spatially exactly in the same relation to the other reproducers as were the original microphones, then a listener at the distant location would experience the same sound sensation as a listener at the source.

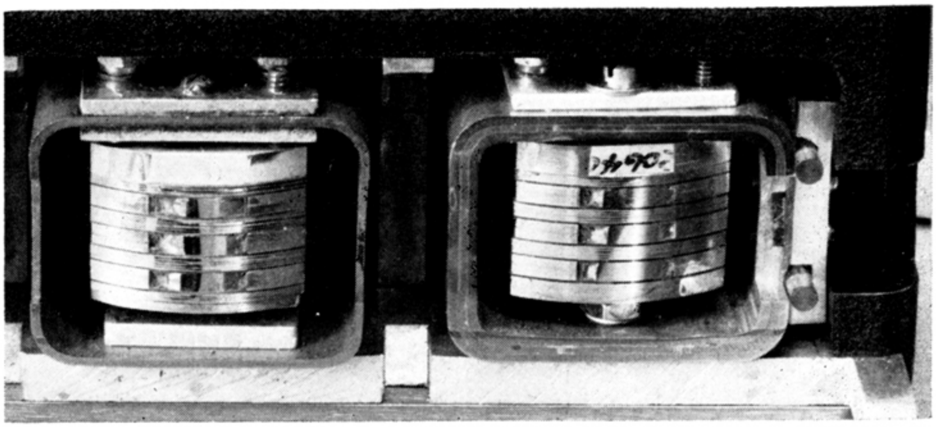
The first compromise with this ideal situation is made in reducing the number of microphones and channels to a two-dimensional plane, so that the whole transmission system samples a sort of 'curtain" of sound between the audience and the orchestra, and the reproducers then reconstitute that curtain at the point of listening. Thus all the time and space between the origination and audition of the music are made transparent for the listener. The second compromise is reduction of the two-dimensional curtain to a straight line, eliminating those microphones located vertically above the others, and eliminating the listener's ability to distinguish if any sounds are meant to originate above the orchestra, a compromise which may be considered an acceptable expedient in order to reduce the system to practicality. The final compromise is that of reducing the number of points on the line of origin from indefinitely many, to three, or two perhaps, but not to one. Earlier studies have indicated, and subsequent experiment has not denied, that the step from one channel to two is a major one; and that the step from two to three is a major one; but that beyond this point an increase in the number of channels is of only marginal significance. The distinction between binaural and stereophonic in the case of a two-channel system may appear obscure unless the 'acoustically transparent curtain" analogy is kept in mind. Experiments by R. J. Tinkham, in cooperation with the University of Illinois,[8] have resulted in a proposal that location of the microphones within the theoretical curtain holds the key to successful stereophonic duplication, in which the listener at the point of reproduction can move about freely within the sound field produced, keeping at all times a sense of spatial orientation with relation to the unseen orchestra, indeed moving, himself, with respect to it while the sound source remains apparently stationary. Since the listerner wears headphones or is otherwise confined in order to remain within the binaural concept, the sound source must inevitably appear to move as his head moves. Two-channel recordings made with the binaural artificial human head device but reproduced by loudspeakers in a comparatively free sound field, it is reported, produce an effect that, although different from the one-channel experience, is nevertheless puzzling to the human nervous system, which is not conditioned for meaningful interpretation of acoustic data presented in this manner. The distinction between binaural and stereophonic becomes more readily apparent when three channels are presented, of course.

The early experiments in stereophonic transmission were limited to simultaneous origination and audition by the character of the then-existing recording systems. Means of preserving the exact phase relations between the two or among the three channels were not readily available. Even the semistereophonic effects with which the motion picture industry experimented in the late 1930's were made dependent upon careful physical separation of the reproducers, with the dominant effect that of amplitude differences among the reproducers. The first wholly practical means of recording separate sound tracks, so that their phase inter-relations might be preserved with great precision, appears to have been offered with the arrival of magnetic tape recording systems. Camras, of the Armour Research Foundation, used this means in experiments on binaural reproduction shortly after the war[9] with notably successful results.

Although the magnetic tape recording medium offers an elegant and simple solution to the problem of preserving these phase relations from track to track so long as playback is always on the machine on which the recording was made, it is probably well for us to appreciate

(Continued Next Page)

7 Bell Telephone System, Monograph B-784, *Auditory Perspective*, Symposium of Six Papers, 1934.



The degree of precision which is required to line up three stacked heads can be judgd from this view of a three-channel stereophonic head assembly. The stack on the left is for recording; that on the right for playback.

the degree of exactness to which the lateral displacement between the two recording or the two reproducing heads must be held, from machine to machine, and from which the heads on any one machine must not be permitted to depart through their service life. If good engineering practices are observed, it has been found that the heads may be constructed so that respective gaps are directly above one another, while still preserving adequate crosstalk rejection. By this means, the twin heads may be constructed as a single unit, and their alignment with respect to each other indefinitely assured. The difference in lateral displacement between the heads, from machine to machine, must be held to limits which are determined by the degree of exactness with which spatial orientation is to be preserved. It appears that a maximum error of 20° may be an acceptable limit to place upon this effect; that is to say, if, in playing back on one machine a tape made on another, the soloist or the conductor may appear to have shifted 20° to the left or right of his true position, this error may be regarded as tolerable. The human hearing apparatus determines directionality, at least in the middle of the audible range, primarily by measuring the phase difference between tones due to the difference in their time of arrival at the two ears. This is to say, the ear measures a time interval. This interval, in the case of tones originating a relatively few degrees off dead-ahead, is given by the expression

$$t = d \tan \lambda / v$$

where t is the time interval in seconds, d is the distance between eardrums (which may be taken as 6 in.). λ is the angular error, and v is the velocity of sound in inches per second. This is approximately 0.165 msec, during which time the tape, assuming recording at 15 ips. moves about 2.5 thousandths of an inch. If, then, the error is not to exceed 20° in the case of the most extreme combination of differences among machines, lateral displacement between each of the two heads in both recording and reproducing stacks must be held to ± 1.25 thousandths. Tolerances of this magnitude can be maintained in manufacturing operations if extraordinary precautions are observed, so it may be assumed that the problem is surmountable.

Among the important recent experiments with three-channel stereophonic recording, those of Lorin D. Grignon, of Twentieth Century-Fox Studios,[10] are perhaps the best documented. Mr. Grignon concludes that stereophonic recording eliminates the "boominess" which monaural recording unnaturally represents some studios to have. This reverberation effect, upon which quantitative data are largely lacking, is also reported greatly reduced with binaural recording by David C. Apps of General Motors.[11] Grignon further concludes that phase relation is the largest factor in the listeners' determination of sound directionality, with the gain through each channel a secondary factor. Our experiments in large measure confirm this, although we find indications that phase relations are most important in the middle range, while amplitude effects becomes more and more important in determining the source of hf sounds. Another of Grignon's conclusions is that one microphone must be placed near any soloist, if clear and definite apparent location is to be preserved, but that there must at all times be some pickup on all three channels from both the soloist and the accompaniment.

It may be said, in conclusion, that the ready preservation of stereophonic effects have been made economically and physically possible by the advent of modern multitrack magnetic tape recorders, that the realization of these effects may be expected to reach the conditions of facsimile reproduction, if known and attainable considerations are met, and, finally, that the applications of stereophonic recording are by no means limited to the preservation of music, but may be expected to have profound effect upon industrial sound analysis, and even upon physical investigations into the nature of sound itself.

8 R. J. Tinkham, Binaural or Stereophonic?, *Audio Eng.*, **37**, 22 (January 1953).

9 Marvin Camras, A Stereophonic Magnetic Recorder, *Proc. Inst. Radio Engrs.*, **37**, 440-447 (1949).

10 Lorin D. Grignon, Experiment in Stereophonic Sound, *J. Soc. Motion Picture Engrs.*, **52**, 280-292 (1949).

11David C. Apps, Use of Binaural Tape Recording in Automotive Noise Problems, *J. Acoust. Soc. Amer.*, **24**, 660 (1952).