Hearing, the Determining Factor for High-Fidelity Transmission*

HARVEY FLETCHER**

In two parts—Part 1

FROM THE ARCHIVES OF BELL TELEPHONE LABORATORIES

This is perhaps the first authoritative study of the requirements for ideal systems for the transmission of speech and music. Much of our present-day knowledge and practice stems from this article, which presents conclusions derived from measurements of hearing on more than 500,000 people during the World's Fairs in 1939 and 1940.

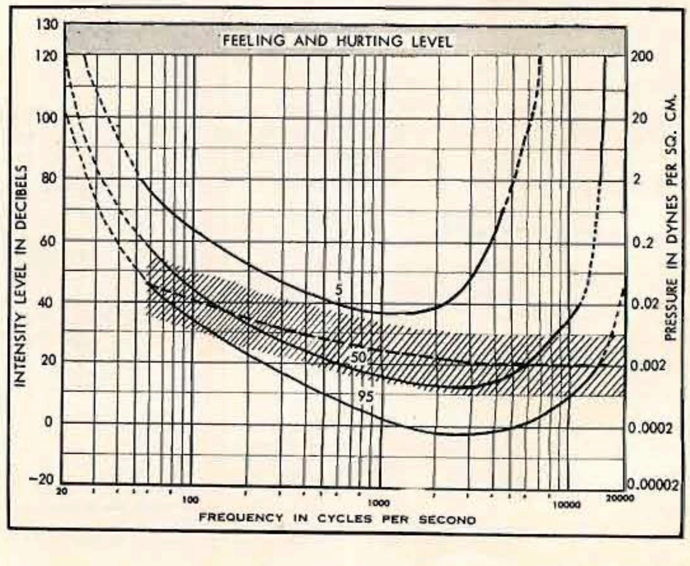
WHENEVER a sound is made by a sudden impact of one solid body upon another, a wave train is set up in the air which contains components ranging in frequency from zero to infinity. As the impact becomes more sudden, the higher-frequency components carry a greater portion of the total acoustic energy. An ideal transmission system from a physicist's standpoint might be defined as one which would transmit such sounds to a distant point and there reproduce a disturbance in the air which is a facsimile of that produced by the original source. The requirements for such a system are very severe and it is difficult, if not impossible, to attain them.

The purpose of transmitting sounds to a distant place is usually so that persons may hear them. Certainly this is true of broadcast systems, telephone systems, and sound-picture systems. Under such conditions the properties of the hearing mechanism and the characteristics of the listening location, rather than the properties of the sounds transmitted, will very largely determine the fundamental requirements of the transmission system. This will certainly be true if we wish to transmit all kinds of sounds which can be heard. However, if we are interested in only a limited number of sounds, then the characteristics of these sounds play a greater part in determining the requirements for the transmission system.

During the years 1938–1940 a survey of the hearing capabilities of persons in a typical population was made by the Bell Telephone Laboratories. This was done in connection with the exhibits at the World's Fairs at San Francisco and New York City, sponsored by the Bell Telephone companies. At these exhibits records of the hearing of more than one half million persons were analyzed. The record expressed the hearing acuity as a relative hearing loss or gain with respect to an arbitrary reference. Measurements at the Laboratories on this reference have made it possible to express these data on an absolute scale and the results have been published by Steinberg, Montgomery, and Gardner.[1] *Figure* 1 has been constructed from data taken from this paper. The lower curve labeled 95 indicates that 95 out of 100 persons in a typical group cannot hear pure tones whose frequency and intensity levels lie below this curve. The top curve indicates that 5 out of 100 cannot hear these tones until they exceed the intensity levels indicated by this curve. The middle curve indicates the levels where one half the group can hear and the other half cannot hear. The dashed portions of the curves indicate regions where no measurements have been made. Feeling and hurting levels lie somewhere above 120 decibels as indicated by the field of dots at the top of the chart. Our experience with reproduced music has taught us that it is undesirable and probably unsafe to reproduce for a general audience sounds that have greater intensity levels than 120 decibels.

If the listener is in a quiet place, these curves set the limits for the ideal transmission system. This ideal of no noise is seldom if ever realized by listeners. Measurements of room noise have been made by the Bell Telephone Laboratories and from these measurements the average noise spectrum can be deduced. In a paper by Seacord[2] it was found that 43

Fig. 1. Contours of hearing loss and room noise.



* Reprinted by permission from *Bell System Telephone Technical Monograph B-1351.* Originally presented at the Broadcast Engineering Conference, Columbus, Ohio, in February, 1941.

** *Bell Telephone Laboratories.*

[1] J. C. Steinberg, H. C. Montgomery, and M. B. Gardner, "Results of the World's Fair hearing tests," *Bell Sys. Tech. Jour.*, vol. 19, pp. 533–562; October, 1940.

[2] D. F. Seacord, "Room noise at subscriber's telephone locations," *J. Acous. Soc. Am.*, vol. 12, pp. 183–187; July, 1940.

decibels was the average sound level in residences not having radios playing. The standard deviation of levels in different residences from this figure was 5.5 decibels. The distribution about this average value indicated that about one half the residences have noise levels between 39 and 47 decibels, and 90 per cent are in the range between 33 and 52 decibels.

Hoth[3] found that the form of the noise spectrum was about the same for all types of rooms. Using his relation, the spectrum for the average room noise having a total level of 43 decibels is that shown in *Fig.* 2, lower curve. The ordinates give the spectrum level. This is obtained as follows. The intensity I in a band of frequency width W is measured. Then the ordinates y are given by

$$y = 10 \log I/WI_0$$

where I_0 is the reference level of 10^{-16} watt per square centimeter. It has been shown[4] that the masking level could be obtained directly from the spectrum level. Using this relation, the curve shown in *Fig.* 2, labeled MASKING LEVEL, was obtained. This curve then gives the level of pure tones which can just be perceived in the presence of average room noise. This masking curve is the one which is shown in *Fig.* 1 as a cross-hatched band. It shows the range of the masking levels for about 90 per cent of the residences in a typical group. The dashed curve gives the average. If we wish to include only the middle 50 per cent of the residences, the top part of this masking-level band would be lowered about 5 decibels.

Transmission Limits

Figure 1 then enables us to set several limits for an ideal transmission system. If the noise emitted by the radio set in a typical residence is not to be heard, its level should be below the average threshold of hearing in the room. For an average room this is seen to be determined by the hearing mechanism from low frequencies to 200 cps, and by the room noise from 200 to 6000 cps, and again by the hearing mechanism above 6000 cps. For example, the fundamental of a 60-cps hum should be kept below a 57-decibel level, whereas any components of this hum around 1000 cps should be kept below 25-decibel intensity level for the average room. It is seen that for the 5 per cent of the rooms which are quietest the limit is set entirely by the hearing-acuity curve. From *Figure* 1 one can also set the frequency limits for an average listener if all sounds that can be heard are to be used in the broadcast. This range is from 20 to 15,000 cps for the highest possible levels, and for any lower levels this frequency range is smaller, as indicated.

Figure 1 also gives the maximum levels such an ideal system might be called upon to transmit. This maximum level is taken as 120 decibels and the same for all frequencies. There is an uncertainty of about 10 decibels concerning the level that can be tolerated by the average ear. Our experience with reproduced sounds near this level indicates that it is very unlikely that higher levels will be used even though levels somewhat higher than 120 decibels may be tolerated by an average listener without producing permanent injury to the ear. It is probable that there will seldom if ever be a demand for such high intensity levels in a home, but if we are thinking in terms of an ideal which is set by human-hearing capabilities the upper limit must be taken at least as high as 120 decibels. The power P for producing the maximum level of 120 decibels varies from 3 watts for a typical residential room to 400 watts for a concert hall such as the Academy of Music in Philadelphia. It may be obtained from the formula

$$P = 0.00012\, V/T \text{ watts}$$

where V is the volume of the room in cubic feet and T is the reverberation time in seconds.

If we utilize the entire intensity level range from the average threshold in a room having average noise to 120 decibels level, it will be seen that from 2000 to 6000 cps this range is approximately 100 decibels. From 500 to 2000 cps it is about 5 decibels less than this figure, while for 100 cps the range is only 75 decibels. It should be emphasized that these figures refer to the level range of single-frequency tones. When talking about program material where complex sounds are used which are rapidly varying in intensity and frequency, the matter of measuring intensity level range is not simple as will be evident from later discussions.

A summary of these conclusions is given by the curves in *Fig.* 3, which give the limits imposed by the hearing of an average person in a room having average room noise. In using this curve it must be remembered that the lower limiting curve may be anywhere in the shaded area, depending upon the room noise condition. This shaded area covers 90 per cent of noise conditions in residences. For an average business office the lower curve will be raised about 15 decibels, and for a factory location the lower curve will be raised nearly 35 decibels, leaving a range of only 60 decibels even if the highest levels that can be tolerated by the ear are used.

(*Continued on page 49*)

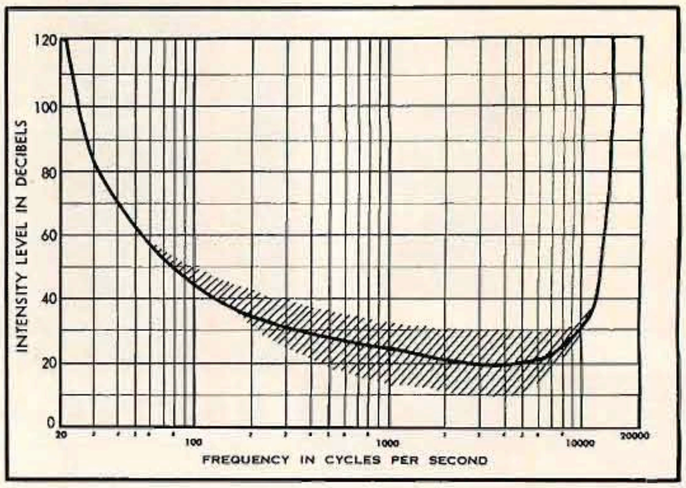


Fig. 3. Hearing limits for pure tones—typical listener in typical residential room noise.

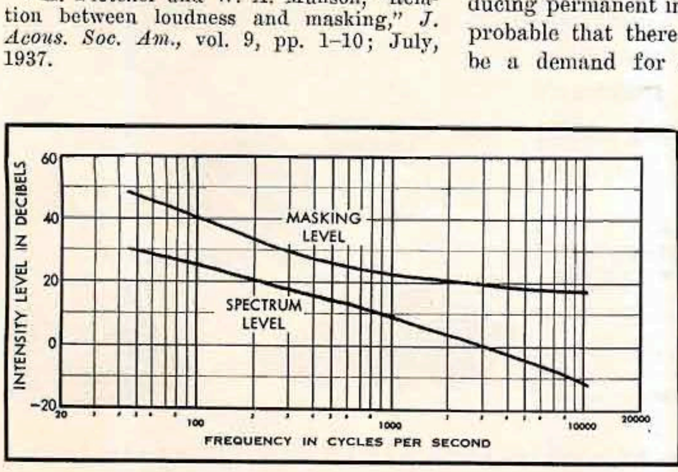


Fig. 2. Average room noise spectrum.

[3] D. F. Hoth, "Room noise spectra at subscribers' telephone locations," *J. Acous. Soc. Am.*, vol. 12, pp. 499–504; April, 1941.

[4] H. Fletcher and W. A. Munson, "Relation between loudness and masking," *J. Acous. Soc. Am.*, vol. 9, pp. 1–10; July, 1937.

HEARING

(*from page 26*)

Noise in Theaters

Measurements of noise have also been made in motion-picture theaters. Mueller found the average level in theaters without an audience to be 25 decibels and with an audience present it was 42 decibels. This last figure is within 1 decibel of that given above for the average noise level in residences. So all the limits mentioned above for residence room noise can also be applied to motion-picture theaters, that is, *Fig.* 3 gives the limits if no limitation is to be placed upon the sounds to be recorded.

Noise measurements made in the Academy of Music in Philadelphia and Constitution Hall, in Washington, D. C., during a quiet listening period indicated levels about 10 decibels lower than that given in *Fig.* 3 for residences. So at least for these concert halls the lower part of the shaded area should be used for determining minimum levels. It will be seen from *Fig.* 1 that for the average person in such quiet intervals in these concert halls the lower limit is set by the acuity of hearing rather than by the audience noise.

The foregoing gives the ideal limits of frequency and intensity for high fidelity. It is well known that within these limits the system must have a sufficiently uniform response with different frequencies so that the ear will not detect it from one having a perfectly uniform response. Due no doubt to the fact that persons usually listen in rooms which have resonances, it is difficult to detect departures of 3 or 4 decibels from uniformity —in fact, it is very difficult to measure them. Here again it would be helpful if we had some precise measurements on this point. Also, it is well known that the system must be linear with intensity, that is, the acoustic output must be proportional to the acoustic input. Also there must be no asymmetry in the vibration during transmission. Any departures from these ideals must not be larger than can be detected by the average ear.

Facsimile Requirements

When all of these requirements are met, a facsimile of the original source cannot be produced unless another factor is considered which is sometimes overlooked, namely, the spatial or auditory perspective character of the sound. If we are reproducing a moving sound source it must appear to move, and if the sound source is broad it must appear that way when reproduced. In nearly all systems now used this factor is neglected. It can be preserved in two ways, namely, by a binaural system or by a stereophonic system. In the former, two channels only are required while in the latter, theoretically an infinite number is required but practically three give a good illusion.

In the binaural system two microphones are placed in the ears of a dummy who sits in the position where the listener would like to sit if he were listening to the original sound. Two transmission lines meeting the requirements above connect these microphones to two head receivers respectively, one being placed on each ear of the observer. With such a system a complete facsimile of the original sound at the dummy is reproduced to the listener. The Oscar system presented at the World's Fair in Chicago by the Bell System was one meeting the stringent requirements outlined above.

In the stereophonic system an attempt is made to produce this spatial effect by using loudspeakers instead of head receivers in the reproduction. Suppose there were interposed between the source of sound and the audience which would normally listen to it, a sound-transparent curtain. A large number of microphones are mounted all over this curtain. An ideal line connects each microphone to a recording unit of a recording system, or to a loudspeaker if a simple transmission system is used. The loudspeakers are spaced over a similar curtain when the sound is reproduced. If the microphones and loudspeakers are close together a curtain of sound will be reproduced similar in all respects to the original sound. Again three such channels for most stages will give a sufficiently close approximation that, due to the limitations of hearing most observers cannot detect the reproduced from the original when all the other requirements discussed above are met. Indeed two channels go a long way toward this ideal. However, if the sounds were not confined essentially to the level of the stage but were permitted to go up and down as far as they went right and left, then nine channels instead of three would be required. Or if we wished the sounds to appear to come from all directions around the listeners, then channels sufficient to cover a sphere surrounding them would be necessary.

The question then arises, how much loss in quality of the reproduced sound is experienced as we depart from such an ideal? To answer this question we must know the kind of material that is used in the transmission. Without such an ideal transmission system, there is always placed some limitation on the type of material used. In general, this material may be classified as either music, speech, or noise.

TO BE CONTINUED