# Pulse Statistics Analysis of Room Acoustics*

R. H. BOLT, P. E. DOAK, AND P. J. WESTERVELT
*Acoustics Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts*
(Received February 2, 1950)

Many sounds of speech and music more nearly resemble pulsed wave trains than abruptly terminated continuous sounds as used in reverberation measurement. It is therefore not surprising to find that two rooms can differ markedly in acoustical quality even if they appear identical under reverberation analysis which ignores details of short transients.

This paper introduces a pulse statistics point of view which takes immediate account of the pulse-like nature of common sounds. Fundamentally, the method consists in examining the response of the room to a short pulse. The walls are replaced by an array of image sources (simple images if the walls are hard, or appropriately modified if there is absorption). These image arrays are then considered statistically.

From this approach one can derive such classical quantities as reverberation time and mean free path. One can also analyze the detailed nature of discrete reflections including interference effects, and thus obtain an average correlation between room geometry and the character of its pulse response.

Idealized experiments in a hard-walled rectangular room are employed to illustrate the essential features of this approach. A point source emits an exponential damped 3600-c.p.s. wave train of about 2 msec. duration. The received signals are recorded logarithmically on an oscillograph and the system is calibrated for quantitative results. Several dozen discrete reflections can be measured and correlated with calculation. The pulses merge into a more or less continuous background after a time that is calculated and confirmed experimentally. Detailed differences arise according to the positions of the source and microphone in the room.

## INTRODUCTION

IT is well known that reverberation time is not a completely adequate index of the acoustical quality of a room. Implicit in the concept of reverberation time is the assumption that the room reaches a steady-state diffuse condition before the source of sound is abruptly terminated. In practice, however, most sounds of speech and music can be generally classified as pulsed wave trains whose amplitudes and frequency components fluctuate sufficiently within time intervals shorter than the time constant of the room, so that the room seldom reaches steady state. Thus it would seem that the response of the room to transient sounds of this general type is an especially important physical problem of room acoustics. The results obtained recently by Mason and Moir[1] and others[2] who have used short tone bursts to investigate acoustics of auditoriums lend support to this point of view.

The task of describing mathematically the response of a room to an arbitrary transient, and of studying the roles of room geometry and distribution of absorbing materials in this response, is extremely complicated. The problem can be approached from a normal mode point of view,[3] or one can attempt to "follow" the sound waves around in the room as they are reflected back and forth from the walls. The latter approach has been recently investigated by Mintzer[4] using Laplace transform methods.

From the transient point of view, it is desirable to use the second approach. In essence, this method consists of replacing the effect of the boundaries of the room by an infinite array of image sources, each image corresponding to one of the multiple reflections of the original wave emitted by the source. Finding these images analytically is no simple matter, and only in very special cases will the images be "mirror images" of the source.[3] The image concept has been used extensively in earlier geometric studies[5,6] where the source is considered to be incoherent. Due mainly to mathematical difficulties, little use of images has been made as yet in wave-acoustical investigations of rooms.

However, it has become increasingly apparent in recent years that the first 20 db of decay of a sound in a room is of primary importance in differentiating between two rooms which have approximately equal overall reverberation times. Further, as the work of Mason and Moir indicates,[1] the time and amplitude distributions of reflected tone bursts can be correlated with the acoustical quality of a room. These facts indicate that the images relatively close to the source, i.e., the first few reflections, are primarily responsible for certain important features of the acoustical character of rooms, as Brillouin has observed.[6] It should therefore be worth while to study this "short term" transient response by a method of images in which all wave properties of the image sources can be considered (i.e., where the assumption of an incoherent source is not made). Further, if an image array satisfying the boundary conditions can be found, one should be able to treat this array statistically and thus obtain the long term average transient response as well.

This paper is confined for the most part to a discussion of an idealized case, a hard-walled rectangular

* This work was supported in part by the ONR, Department of Navy, under Contract NObs 25391, Task 7.

[1] C. A. Mason and J. Moir, "Acoustics of cinema auditoria," J. Elec. Eng. 88, Part III, No. 3 (September, 1941).

[2] British Broadcasting Corporation, Engineering Division, Research Department Reports B.027 and B.035.

[3] P. M. Morse and R. H. Bolt, "Sound waves in rooms," Rev. Mod. Phys. 16, 117 (1944).

[4] D. Mintzer, "Transient sounds in rooms," J. Acous. Soc. Am. 22, 341 (1950).

[5] C. F. Eyring, "Reverberation time in 'dead' rooms," J. Acous. Soc. Am. 1, 217–241 (1930).

[6] J. Brillouin, "Sur l'acoustique des salles," Rev. d'Acoustique 1 (September–November, 1932).

room containing a simple source which emits a short pulse. Statistical properties of the image array are investigated to illustrate the image method and to provide a basis for future work on more general cases.

The limited applicability of results based on a simple mirror image picture has already been pointed out.[3] However, as the discussions of both Morse and Bolt[3] and Mintzer[4] indicate, a specular approximation is allowable in cases where the walls are not too soft, and, in any case, one can assume that the image of a simple source can be described analytically by an expansion in spherical harmonics, which is essentially a multipole expansion of the image source taken about the mirror image point. Thus, if the walls are not fairly hard, neither the simple mirror image nor the specular approximation can be used, and other suitable representations of the images must be found.

## PULSE STATISTICS THEORY

### 1. Image Space and the Time Distribution of Reflected Pulses

With these restrictions in mind, let us now set up a working picture for pulse analysis. We shall consider a simple rectangular room with perfectly reflecting walls. The room has dimensions $L_x$, $L_y$, $L_z$, one corner being at the origin of a Cartesian coordinate system.

A sharp pulse of sound is emitted from a source in this room. The source is assumed to be a simple one, with spherically symmetric radiation. This point source may be located by the vector:

$$\mathbf{r}_s = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}, \tag{1}$$

as shown in Fig. 1. A point receiver of sound pressure is located at a position:

$$\mathbf{r}_R = U\mathbf{i} + V\mathbf{j} + W\mathbf{k}. \tag{2}$$

The vector displacement of the receiver from the source is:

$$\mathbf{R}_0 = \mathbf{r}_s - \mathbf{r}_R. \tag{3}$$

Associated with this sound source is an infinite array of image sources each occupying, alone, an image replica of the physical room. This image array also is illustrated in Fig. 1. Each image cell is designated by three numbers $(l, m, n)$, and these three numbers can take all integral values from minus infinity to plus infinity. The cell (0, 0, 0) is the actual room.

The pulse associated with the image (1, 0, 0) reflects once from the wall $x = L_x$. The pulse from (2, 0, 0) reflects first from the wall $x = 0$, then from the wall $x = L_x$. Obviously the absolute value of the cell number, $l$, for cells lying along the $x$ axis gives directly the number of wall reflections suffered by the pulse from the image in question. The pulse from (1, 1, 0) reflects once from the wall $x = L_x$ and once from the wall $y = L_y$, so that its number of reflections is $|l| + |m|$. In fact, by this system of designation, the total number of wall reflections suffered by the pulse from the image $(l, m, n)$ is directly:

$$N_{lmn} = |l| + |m| + |n|. \tag{4}$$

The vector position of each image source is:

$$\mathbf{r}_{lmn} = x_l\mathbf{i} + y_m\mathbf{j} + z_n\mathbf{k}. \tag{5}$$

The vector position of each image with respect to the receiver is:

$$\mathbf{R}_{lmn} = \mathbf{r}_{lmn} - \mathbf{r}_R. \tag{6}$$

To evaluate these last two equations, we note first that

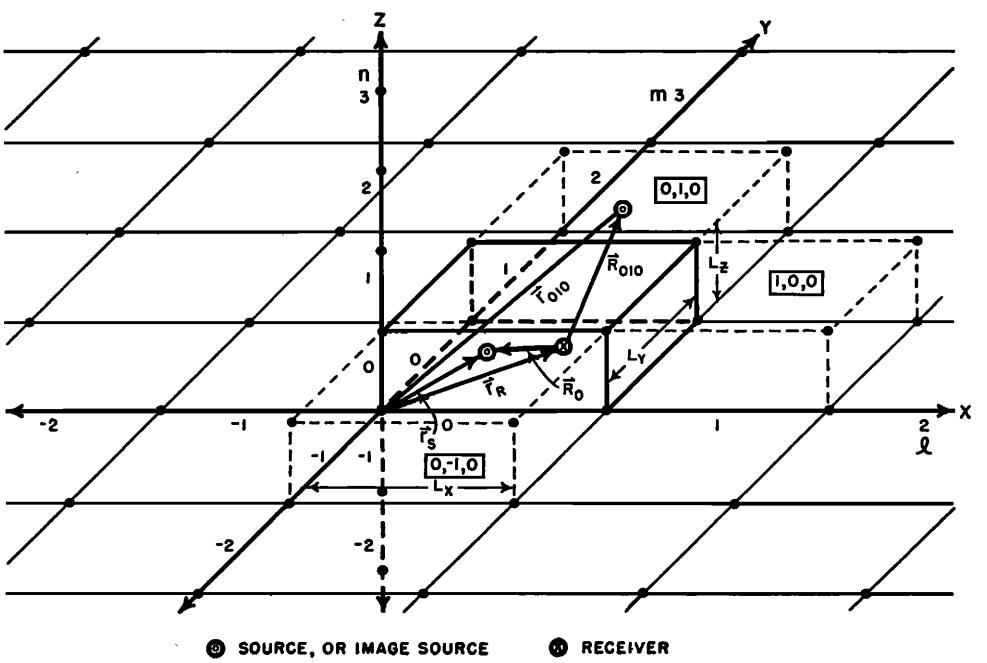



FIG. 1. An image array, showing image cells and source and receiver locations. The cell designation numbers are enclosed in the small boxes. ($L_x = 23.0$ ft., $L_y = 13.4$ ft., $L_z = 8.44$ ft.)

each image cell is a mirror reflection of the cells adjacent to it. This complicates the analysis somewhat in that we have two kinds of symmetry with respect to the basic coordinate system. Fortunately the cell designation numbers indicate directly the type of symmetry: even numbers, including 0 for the actual room, designate cells in which the source position duplicates that in the original room, while odd numbered cells have sources at a reflected position. Therefore the components for Eq. (6) are given by:

$$\left.\begin{aligned} x_l &= lL_x + X \\ y_m &= mL_y + Y \\ y_n &= nL_z + Z \end{aligned}\right\} \quad l, m, n \text{ even},$$

$$\left.\begin{aligned} x_l &= (1+l)L_x - X \\ y_m &= (1+m)L_y - Y \\ z_n &= (1+n)L_z - Z \end{aligned}\right\} \quad l, m, n \text{ odd}. \tag{7}$$

In order to generalize some illustrative calculations we select the longest dimension of the room as a scale unit and introduce the following definitions:

$$\begin{aligned} &L_x = L, \quad L_y = pL, \quad L_z = qL, \quad p, q \leq 1; \\ &(U-X)/L = \mu_{xe}, \quad (U+X)/L = \mu_{xo}, \\ &(V-Y)/L = \mu_{ye}, \quad (V+Y)/L = \mu_{yo}, \\ &(W-Z)/L = \mu_{ze}, \quad (W+Z)/L = \mu_{zo}. \end{aligned} \tag{8}$$

Thus $p$ and $q$ are dimension ratios and the $\mu$'s give the source-to-receiver displacements for both kinds of symmetry. From Eq. (6) we then get:

$$\frac{\mathbf{R}_{lmn}}{L} = \left\{\begin{matrix} [l-\mu_{xe}]\mathbf{i} \\ \text{or} \\ [l+1-\mu_{xo}]\mathbf{i} \end{matrix}\right\} + \left\{\begin{matrix} [mp-\mu_{ye}]\mathbf{j} \\ \text{or} \\ [(m+1)p-\mu_{yo}]\mathbf{j} \end{matrix}\right\} + \left\{\begin{matrix} [nq-\mu_{ze}]\mathbf{k} \\ \text{or} \\ [(n+1)q-\mu_{zo}]\mathbf{k} \end{matrix}\right\}, \quad \left\{\begin{matrix} \text{even} \\ \text{or} \\ \text{odd} \end{matrix}\right\}. \tag{9}$$

Some calculations from this equation are illustrated in Figs. 2 and 3 which will be discussed later.

Next let us find the average number of pulses received up to a specified time, $t$, after the emission of the pulse from the original source. These pulses come from all of the images within a radius $|\mathbf{R}_{lmn}| = ct$. One image is contained in each cell of volume $V = L_xL_yL_z$. The number of pulses, $N_p$, is directly given by the volume of the sphere out to $ct$ divided by the volume of one cell:

$$N_p = \frac{4\pi|\mathbf{R}_{lmn}|^3}{3V} = \frac{4\pi c^3 t^3}{3V}, \tag{10}$$

and the number of pulses received per second is:

$$\frac{dN_p}{dt} = \frac{4\pi c^3 t^2}{V}. \tag{11}$$

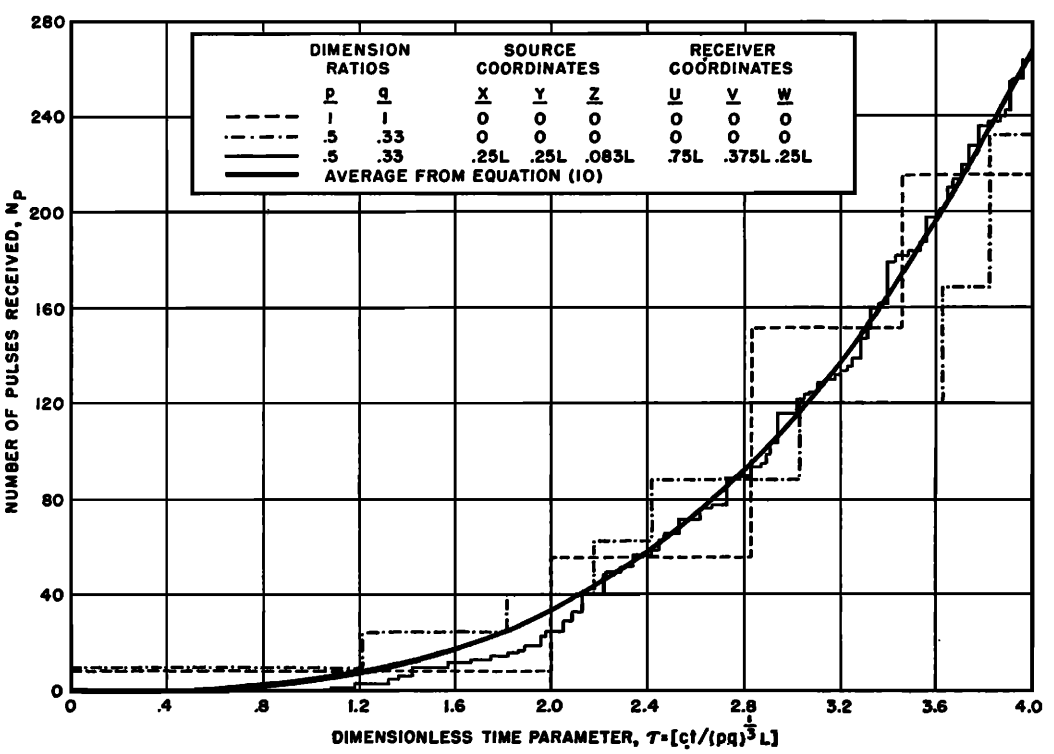



FIG. 2. Graph showing $N_p$, the number of pulses arriving at the receiver up to a time $t$ as a function of the dimensionless time parameter $\tau$. Note the extreme stepped behavior of the dashed curve for a cubical room.

These equations, (10) and (11), are valid for values of $l$, $m$, and $n$, large enough so that the details regarding positions of source and receiver in the room can be neglected. These equations also are illustrated in Figs. 2 and 3. A calculation of an actual case illustrates that a surprisingly large number of reflections per second are predicted. In a room of 10,000 cu. ft., Eq. (11) gives a rate of 180 pulses per second at 1/100 of a second. There are several conditions met in practice which greatly reduce this number. For one thing, the floor is generally quite absorptive when an audience is present, so that half of the spherical volume just assumed is essentially eliminated. Also there are usually a large number of degeneracies or coincidences which reduce the effective number of pulses, as we shall see.

Figures 2 and 3 show $N_p$, the number of pulses received up to a time $t$, for rooms of several dimension ratios and various positions of source and receiver. $N_p$ is plotted for convenience as a function of the dimenless time parameter

$$\tau=\frac{ct}{(pq)^{\frac{1}{3}}L}, \tag{12}$$

where $p$, $q$, and $L$ are defined in Eq. (8). The heavy solid line in both figures gives the average value of $N_p$ as computed from Eq. (10).

In Fig. 2 the dashed curve shows the extreme degeneracy of a cubical room when both source and receiver are in a corner. The pulses arrive in large groups because of the symmetry of the image point lattice for a cubical room and the location of source and receiver. The dash-dot curve for a rectangular room with $p=0.5$ and $q=0.33$ shows similar but smaller groups of coincident pulses. Both these curves start at $N_p=8$ since the direct pulse and the pulses from the seven source-corner images reach the receiver simultaneously at $t=\tau=0$. The light solid curve is for the same rectangular room, source and receiver now being at point such that coincidences are more or less "accidental." We note that this curve follows the average curve very closely for $\tau>2.2$. For $\tau<2.2$, the curve lies consistently below the average curve. This is partly because the separation of source and receiver are such that no pulse arrives at the receiver until $\tau=1.1$. In detail, this initial part of the curve depends in a somewhat more involved way on the relative positions of the real and image sources and the receiver.

Figure 3 shows in more detail the initial rise of the curve just discussed. The dash-dot and dashed curves show similar initial rises for a room of $p=0.8$ and $q=0.6$. The same general behavior is shown by these curves, although both reach the average curve sooner. The duration of this initial departure appears to decrease as the room becomes more nearly cubical.

## 2. Some Special Cases of Coincidences

In Eqs. (10) and (11) it was assumed that all images are distinct, that is, that no two pulses arrive at the receiver at exactly the same time. However, if source and receiver are both placed in certain locations, it is possible for the pulses to arrive in groups, as is evident from Figs. 2 and 3. It is of interest to find the

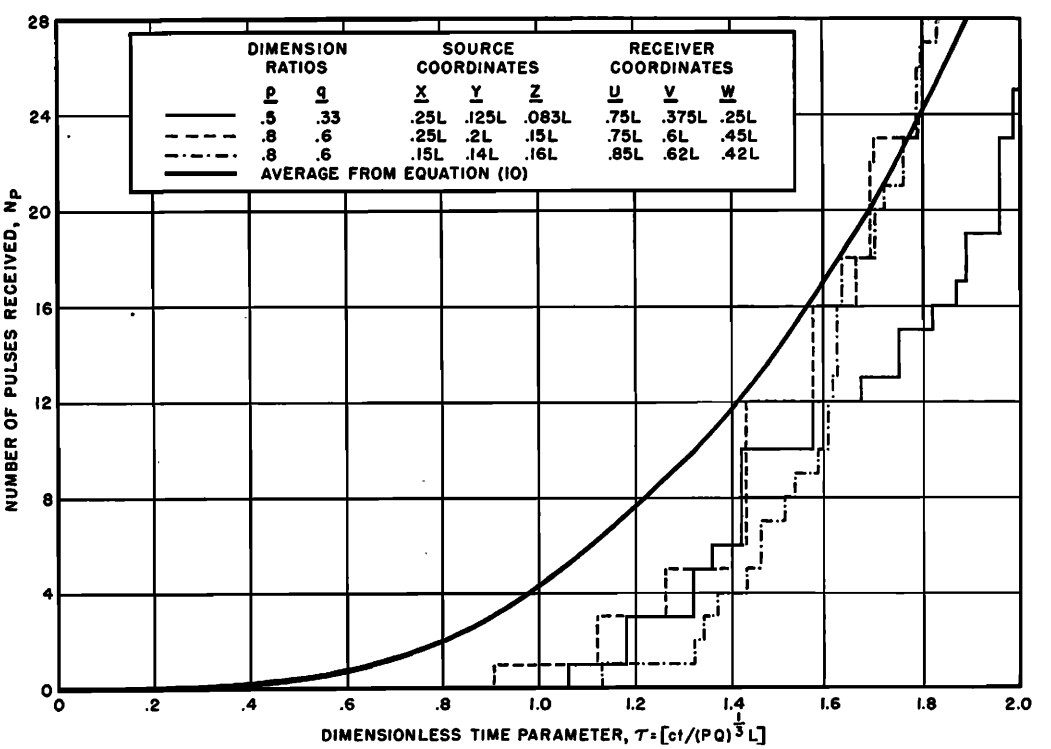



FIG. 3. Graph on an expanded scale showing the initial behavior of $N_p$ as a function of $\tau$.

number of distinct pulse groups reaching the receiver for some of these degenerate cases. In all cases it is assumed that the room is rectangular, hard-walled, and of incommensurate dimensions.

*Case 1*

Consider the source to be in a corner and the receiver in a position incommensurate with respect to the images but otherwise arbitrary. If the source is close enough to the corner, the images will clump together in groups of eight, being so close together that, effectively, each clump of eight acts as a single distinct image. Since pulses reaching the receiver will arrive in groups of eight, the *effective* number $N_p'$ of distinct pulses arriving per second is:

$$N_p'=\frac{1}{8}\frac{4\pi(ct)^3}{3V}=\frac{\pi(ct)^3}{6V}. \tag{13}$$

*Case 2*

Here we place both source and receiver in the center of the room. All the images are distinct, but because of the symmetry arising from the position of the source and receiver, a pulse from an image along a negative coordinate axis (the origin being taken at the center of the room for convenience) will arrive simultaneously with a pulse from an image along the positive coordinate axis. A pulse from a non-axial image in the first quadrant of a coordinate plane will arrive simultaneously with pulses from corresponding images in each of the other quadrants, and a pulse from an "oblique" image will arrive simultaneously with pulses from corresponding images in the other seven octants. Thus the effective distinct images are all contained in one octant of image space. Counting of the number of distinct pulses arriving can then be accomplished in exactly the same manner as the counting of the number of normal modes with frequencies less than a certain value.[3] The volume occupied in image space by all distinct image points within a radius, $ct$, divided by the volume, $V=L_xL_yL_z$, occupied by each image point, gives for $N_p'$ in this case:

$$N_p'=\frac{\pi}{6V}(ct)^3+\frac{\pi L_T}{32V}(ct)^2+\frac{S}{8V}(ct), \tag{14}$$

where:

$$V=L_xL_yL_z$$
$$S=2(L_xL_y+L_yL_z+L_zL_x)$$
$$L_T=4(L_x+L_y+L_z).$$

Similar arguments for the cases when both source and receiver are in a corner, at the center of a wall, or at the center of an edge show that all these cases (except *Case 1*, of course) can be expressed by the formula:

$$N_p'=\frac{1}{\eta_x\eta_y\eta_zV}\left[\frac{\pi(ct)^3}{6}+\frac{\pi}{8}(\eta_xL_x+\eta_yL_y+\eta_zL_z)(ct)^2+\tfrac{1}{4}(\eta_x\eta_yL_xL_y+\eta_y\eta_zL_yL_z+\eta_z\eta_xL_zL_x)(ct)\right], \tag{15}$$

where $\eta_x$, $\eta_y$, and $\eta_z$ are to be chosen as follows:

| | |
|---|---|
| $\eta_x=\eta_y=\eta_z=2$, | source and receiver in corner; |
| $\eta_x=\eta_y=2$, $\eta_z=1$ | source and receiver at center of a $z$-edge, etc.; |
| $\eta_x=2$, $\eta_y=\eta_z=1$ | source and receiver at center of a $yz$-wall, etc.; |
| $\eta_x=\eta_y=\eta_z=1$, | source and receiver at center of room. |

It is clear that for a given room the formula for $N_p'$ will depend upon the position of both the source and receiver. Further, for pulses of a finite length, for actual sources and receivers, and for most positions of source and receiver, exact coincidences will be rare, and "almost coincidences," with interference effects between the various pulses, will be the rule rather than the exception.

### 3. Pulse Spacing Statistics

The close similarity between the normal frequency lattice and the three-dimensional image source lattice has already been noted. The average statistical properties of pulses bear a close resemblance to the equivalent properties of the normal frequencies. Therefore, similarities in the fluctuation statistical properties common to both normal frequencies and pulses might be expected.

The frequency spacing index $\psi$ has been defined as the mean squared ratio of actual to average normal frequency spaces, for a specified frequency interval. This index has been evaluated for a certain class of rectangular rooms, and it can be calculated if normal frequency values are available.[7] We shall see next that the exact analog of $\psi$ for pulse spacing (defined as the mean squared ratio of actual to average time intervals for returning pulses) can be calculated and evaluated in every case for which the frequency spacing index index can be calculated or evaluated, provided that source and receiver are maintained in one corner.

We start with the expression for the dimensionless normal frequency of the $l$, $m$, $n$th mode:[7]

$$\mu_{lmn}=\tfrac{1}{2}(pq)^{\frac{1}{2}}[l^2+(m/p)^2+(n/q)^2]^{\frac{1}{2}}. \qquad l, m, n=0, 1, 2\cdots. \tag{16}$$

We obtain the time of arrival of distinct pulses from the image sources by using Eq. (15) for the case where source and receiver are both in a corner. The resulting equation for time of arrival can be put into the dimensionless form,

$$\tau/4=\tfrac{1}{2}(pq)^{-\frac{1}{2}}[l^2+(mp)^2+(nq)^2]^{\frac{1}{2}}, \qquad l, m, n=0, 1, 2\cdots, \tag{17}$$

where $\tau$ is as defined in Eq. (12).

Notice that Eq. (17), which applies only to the case where source and receiver lie in the same corner, differs from Eq. (16) for $\mu_{lmn}$ only through the appearance of $(p)^{-1}$ and $(q)^{-1}$ in place of $p$ and $q$. Since $p$ and $q$ enter

[7] R. H. Bolt, "Normal frequency spacing statistics," J. Acous. Soc. Am. 19, 79 (1947).

symmetrically in both Eqs. (16) and (17), they can be interchanged. Thus in order to convert any equation giving the frequency spacing index into a form which will yield the pulse spacing index, one need only substitute $p^{-1}$ for $q$, $q^{-1}$ for $p$, and $\tau/4$ for $\mu$. Of course, if the room proportions are such that the required weighting factors cannot be evaluated, but actual pulse spacing values are available, for example from experiment, $\psi$ for pulse spacing is easily evaluated from its defining equation:

$$\psi_{AB}=\frac{1}{\tau_B-\tau_A}\sum_A^B\left(\frac{\delta\tau_{lmn}}{\langle\delta\tau\rangle_{Av}}\right)^2\langle\delta\tau\rangle_{Av}, \tag{18}$$

where $\langle\delta\tau\rangle_{Av}/4$ is obtained by making the substitutions $p\rightarrow(q)^{-1}$, $q\rightarrow(p)^{-1}$, $\mu\rightarrow\tau/4$ in the equation giving the average spacing between adjacent normal frequencies.[7]

### 4. A Derivation of the Mean Free Path

Pulse analysis yields a straightforward derivation of the classical mean free path. We assume that the order numbers are so large that the coordinates of the source and receiver can be neglected (that is, source and receiver can be considered to be at the origin). Using polar coordinates $(r, \theta, \phi)$ we can express the number of reflections associated with each cell as:

$$N_{lmn}=r\left[\frac{\cos\phi\,\cos\theta}{L_x}+\frac{\sin\phi}{L_y}+\frac{\cos\phi\,\sin\theta}{L_z}\right]. \tag{19}$$

We next obtain an average value of the number of reflections out to a given radius $r=ct$, by averaging Eq. (19) over all angles:

$$\bar{N}_{lmn}=\frac{\int_\sigma N_{lmn}d\sigma}{S_0}=\frac{rS}{4V}, \tag{20}$$

where $d\sigma=r^2\cos\phi d\phi d\theta$, $S_0=(4\pi r^2)/8$, $S=2(L_yL_z+L_zL_x+L_xL_y)$, and the integration is taken over one octant of space. By definition, the mean free path is equal to the total distance traveled by an average pulse in a given time $t$, divided by the number of reflections of that pulse during the same time. Therefore:

$$\text{m.f.p.}=r/\bar{N}_{lmn}=4V/S. \tag{21}$$

### 5. Energy in the Pulse

We now consider the energy contained in a sound pulse and follow the course of that energy as the sound becomes dispersed throughout the room and absorbed at its boundaries. The total power radiated by a simple source is:[8]

$$\Pi=\frac{\pi\rho\nu^2Q_0^2}{2c}=\frac{4\pi}{\rho c}(\langle p_0^2\rangle r_0^2)\text{ ergs/sec.}, \tag{22}$$

where $Q_0$ is the source strength, and $\langle p_0^2\rangle$ is the mean square sound pressure at a distance $r_0$ from the source in a free field. This equation applies to steady-state radiation. It is also valid for a single pulse wave train which has a sufficiently narrow spectral distribution. This requirement is fairly well satisfied, for example, if there are ten or more waves in the train, and if the wave amplitude is fairly constant throughout the duration of the pulse. We designate the length of this pulse $\tau_p$ and write the total energy contained in the pulse as:

$$E_p=\Pi\tau_p=(4\pi/\rho c)(\langle p_0^2\rangle r_0^2)\tau_p\text{ ergs.} \tag{23}$$

The energy density in the pulse is continually diminishing as the pulse radiates outward. At any instant of time, $t$, after emission, the volume of space containing the pulse is:

$$V_p=(4\pi/3)c^3[t^3-(t-\tau_p)^3]\simeq 4\pi c^3t^2\tau_p. \tag{24}$$

We next consider the multiplicity of pulses arriving at the receiver from all directions as time progresses (still assuming that the room is lossless). We take the number of pulses arriving per second, as given by Eq. (11), and multiply this by the energy density during the duration of a pulse passage, obtaining the energy density per second (provided the pulses add incoherently):

$$W_p=\frac{E_p}{V_p}\cdot\frac{dN_p}{dt}=\frac{4\pi}{\rho c}\frac{\langle p_0^2\rangle r_0^2}{V}\text{ ergs/cc-sec.} \tag{25}$$

If we now multiply $W_p$ by the fraction of a second occupied by the individual pulse, we obtain the average energy density in a room of volume $V$:

$$W_R=\frac{4\pi}{\rho c}\frac{\langle p_0^2\rangle r_0^2}{V}\tau_p\text{ ergs/cc.} \tag{26}$$

A more convenient quantity experimentally is the mean square pressure in the room:

$$\langle p_R^2\rangle=\rho c^2W_R=\frac{4\pi c(\langle p_0^2\rangle r_0^2)\tau_p}{V}. \tag{27}$$

Suppose we take a pulse of $\tau_p=20$ msec. duration which generates a mean square pressure of one dyne (74 db pressure level) at a distance of one meter from the source in a free field. If this pulse is emitted in a room of 10,000 cu. ft. volume, we find that the average mean square pressure throughout the room after the pulse is dispersed, as calculated from Eq. (27), is about 69 db, or only 5 db less than the sound pressure level in the original pulse as measured at one meter. In practice, the sound is rapidly dissipated by absorption and may be canceled out by interference effects.

[8] P. M. Morse, *Vibration and Sound* (McGraw-Hill Book Company, Inc., New York, 1948), second edition.

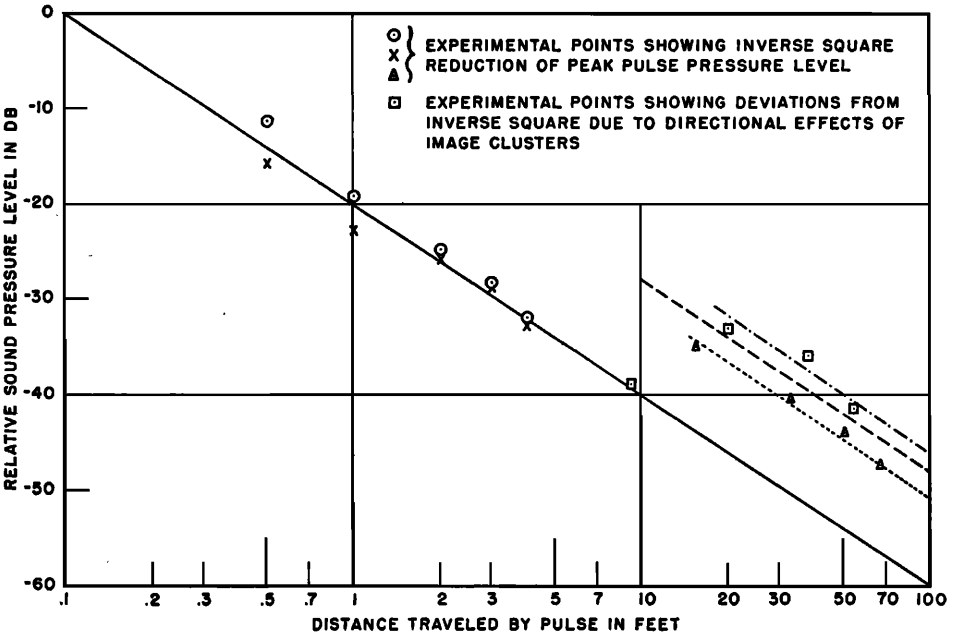



FIG. 4. Graphs of peak pulse pressure level as functions of distance from the source for several sequences of direct and reflected pulses.

### 6. A Derivation of the Eyring Reverberation Equation

Let us assume that all of the boundaries of the room are equally absorptive, and that their absorption is less than 20 percent, so that the specular reflection approximation is valid. Further, let us suppose that the energy in a pulse is diminished upon reflection by a factor $(1-\alpha)$, where $\alpha$ is the average over all angles of the free wave absorption coefficient. Then the pulse energy associated with the $N$th image is:

$$E_{pN}=E_p(1-\alpha)^{\bar{N}}=E_p(1-\alpha)^{(ctS/4V)}, \qquad (28)$$

which leads at once to the Eyring reverberation equation:

$$T_{60}=KV/S\ \ln[1/(1-\alpha)].$$

This is the average long term reverberation time of the room. To study the short term response for a fairly hard-walled room, one could utilize the specular reflection approximation[3,4] for each of the first few reflections and thus obtain in detail the approximate short term response of the room to a particular transient. Further refinements of the long term response might be obtained by treating groups of images statistically, just as modified decay equations have been obtained by grouping of normal modes.[3]

## PULSE MEASUREMENTS IN A HARD-WALLED RECTANGULAR ROOM

### 1. Experimental Procedure

In order to obtain some simple results which could be interpreted from a pulse statistics point of view, short sound pulses were produced experimentally in a hard-plaster-walled rectangular room, and photographs of the sound arriving at a microphone were made using a cathode-ray oscillograph. The source used was a W.E. 713A receiver unit feeding into a $\frac{3}{4}$-in. diameter brass tube 8 in. long packed with steel wool to present a high acoustic impedance to the diaphragm. The effective source was the end of the tube, which was small and could be considered approximately a simple source. The microphone was a W.E. 633A dynamic type, the output of which was passed through an ERPI RA-363 octave filter and through the amplifier section of a ERPI RA 277-F sound analyzer.

The output of the analyzer was attenuated logarithmically by a Kay Labs Type 510-A Logaten and then put across the vertical deflection plates of a DuMont 247 cathode-ray oscillograph. A pulsed carrier of about 2 msec. duration was produced by mechanical switching of the speaker input which consisted of a 3600-c.p.s. signal from a Hewlett Packard 200 D audio oscillator amplified through a Fairchild audio amplifier. A slow (approximately 42 cm/sec. on the screen) external single sweep on the oscillograph was activated mechanically shortly before the beginning of the pulse. Thus the logarithm of the acoustic signal, picked up by the microphone during the 0.25-sec. interval after the pulse was emitted, was recorded linearly on the oscillograph screen as a linear function of time, and was photographed.

### 2. Calibration and Auxiliary Data

The horizontal sweep speed was calibrated by a 30-c.p.s. signal direct to the vertical plates from the oscillator. The vertical deflection was calibrated in 5-db steps, using the attenuator pad on the ERPI analyzer. The horizontal sweep was found to be linear within experimental error, while the vertical deflection was

found to be linear in db with a maximum deviation of $\pm 1$ db over the main working range of 50 db. With somewhat larger variations from linearity at low amplitudes, the vertical deflection is approximately linear over a range of 70 db, which includes background level in the pictures.

The following assumptions and conventions were adopted in interpretation of the data: (a) All pictures were calibrated in db *vs.* ft., since time in the room corresponds to distance in image space; (b) it was assumed that the acoustic pulse emitted by the source in all pictures was the same in both shape and height; (c) a decibel reference level was established in terms of the peak amplitude and was taken as zero db at 0.1 ft. from the source, all other levels thus coming out negative; (d) the velocity of sound was taken as 1120 ft./sec.

In order to get a rough check on the validity of the assumption that the actual source was a simple source, two sets of pictures were taken in which source and receiver were placed near the center of the room (so that direct pulses would arrive without interference from images) and were spaced several different distances apart from 0.50 to 4.0 ft. The peak levels of these direct pulses plotted against distance between source and receiver on semi-log paper are shown in Fig. 4. The best straight line fit to the experimental points is a line with inverse square slope. This in turn indicates that interpreting the actual source as a simple point source is a fair approximation. Another check on this approximation was made by measuring peak pulse amplitude in db at a constant radius from the source, but at six different angles with respect to the tube, from 0° to 180°. These measurements showed a maximum variation of $\pm 2$ db in the peak values.

### 3. Pulse Shape and Energy

Pulse length and carrier frequency were selected empirically. The pulse length was chosen short enough compared to room dimensions to give at least several clearly separated "echoes," but long enough to include at least five or ten cycles of the carrier. In turn the carrier frequency was selected as high as possible

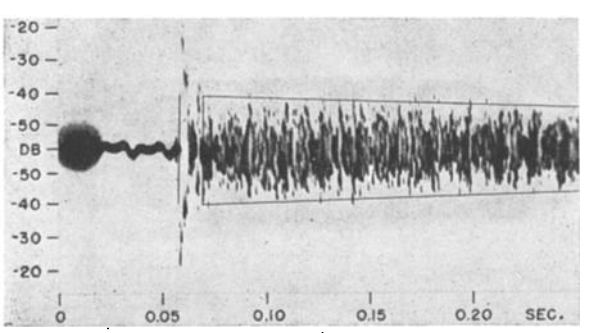



FIG. 5. Photograph for a non-degenerate case in which the distance between source and receiver is 1′. This photograph shows a typical direct pulse without interference from image pulses.

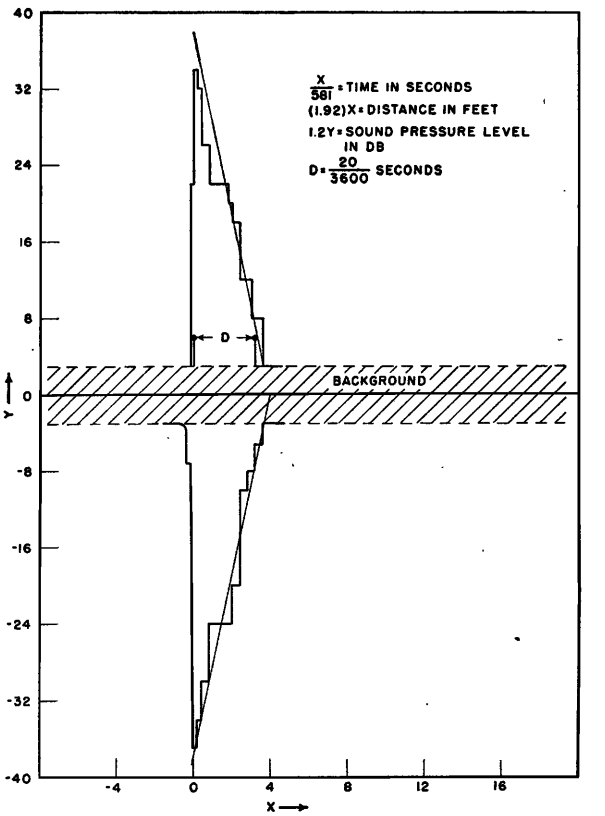



FIG. 6. Graphical enlargement of the direct pulse shown in Fig. 5, illustrating the general shape of the envelope.

(without encountering large air absorption) in order that the speaker would operate in a region of high efficiency and good transient response. The sweep speed and carrier frequency were such that carrier details are barely unresolvable, although individual oscillations of the carrier can be distinguished in some places in the pictures.

The exact nature and shape of the pulse was subsequently determined by analysis of the pictures (e.g., Fig. 5). The average pulse dimensions as determined from the photographs are illustrated in Fig. 6, which is a graphical enlargement of the direct pulse shown in Fig. 5. The detailed shape of the pulse is represented by a stepped curve in which the width of most of the steps is equivalent to several cycles of the carrier. At the onset of the pulse the first two or three swings of the alternating and exponentially increasing carrier are just discernible and the pulse builds up to maximum amplitude in about two cycles of the 3600-c.p.s. carrier. The pulse is almost symmetric but its peak amplitude in the positive direction is measurably greater than in the negative direction for this particular picture. (Symmetry of the pulse, of course, depends upon the phase of the carrier at the instants of switching it on and off.) Following the peak, the pulse decays in a rigorously exponential manner within experimental error as indicated in Fig. 6 by the two straight lines drawn through

the stepped curves. The pulse decay rate is about $2.5\times10^3$ db/sec., and depends principally upon the loudspeaker characteristics. The background noise showing on Fig. 5 is mainly a.c. power ripple. Maximum background level for all pictures was $-65$ db.

The energy in a pulsed carrier having an envelope that rises instantaneously and then decays exponentially is

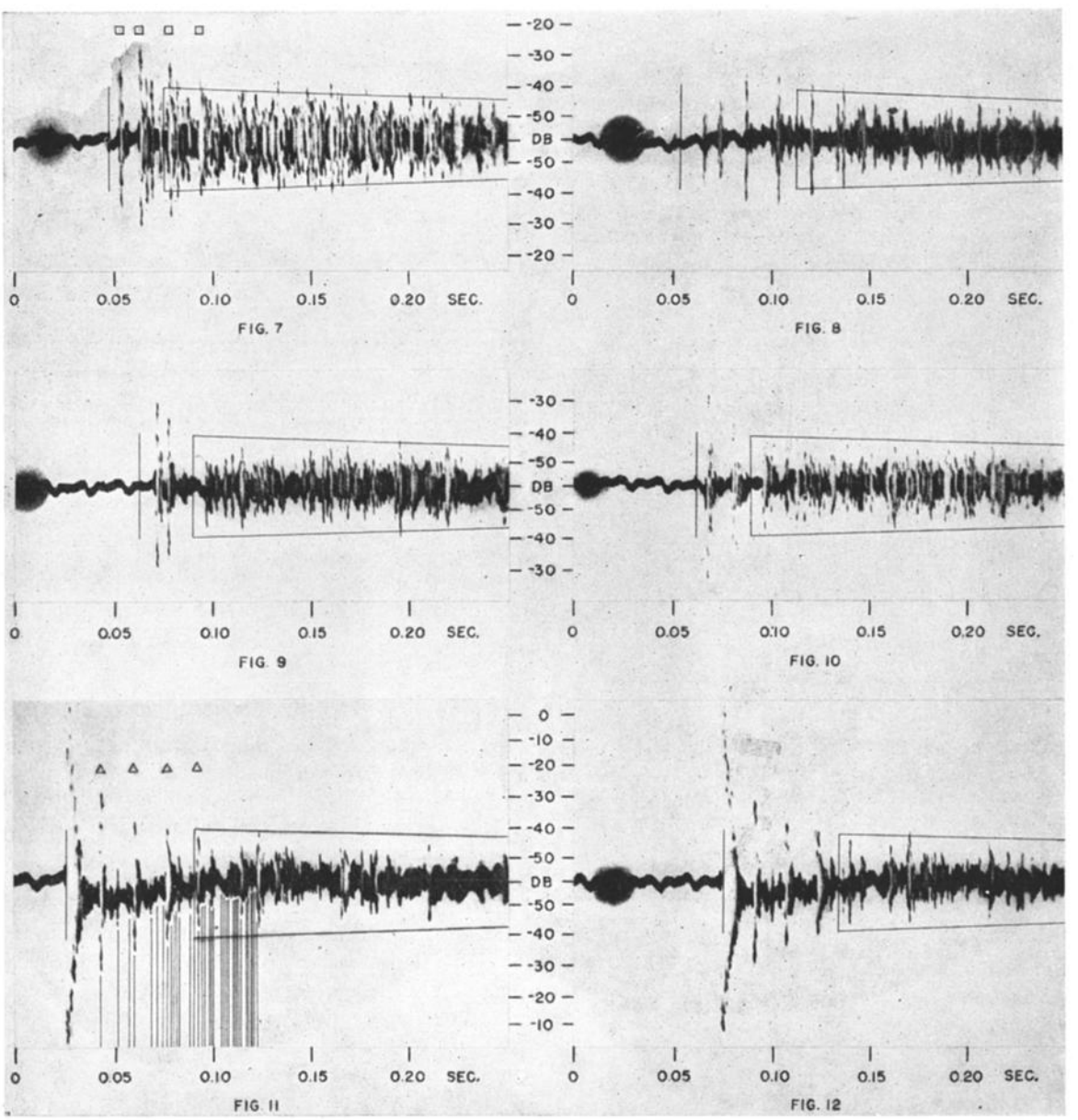
$$E_p=\frac{\pi p_0^2 r_0^2}{\rho c}\,\frac{\omega^2+2\gamma^2}{\gamma(\omega^2+\gamma^2)}\ \text{ergs}, \tag{29}$$

where $p_0$ is the peak pressure in dynes at $r_0$ cm, $\omega$ is the angular frequency, and $\gamma$ is the exponential decay constant. For the pulse shown here $\gamma=580$ and $\omega=7200\pi$ so that $(\gamma/\omega)^2\ll1$. Hence Eq. (29) reduces to

$$E_p=\frac{\pi p_0^2 r_0^2}{\rho c\gamma}. \tag{30}$$

If we let $p_s$ be the r.m.s. pressure in the room a relatively long time after emission of the pulse (i.e., when the sound is fairly uniformly distributed), then the ratio $p_s/p_0$ is, using Eq. (27):

$$\frac{p_s}{p_0}=\frac{(\langle p_R^2\rangle)^{\frac{1}{2}}}{p_0}=\frac{1}{p_0}\left(\frac{\rho c^2 E_p}{V}\right)^{\frac{1}{2}},$$




FIGS. 7–12. Series of pulse photographs showing response in a hard-walled rectangular room when the source is placed in a corner and the receiver is located at various points around the room. Note the extreme number of coincidences when the receiver is also in a corner (e.g., Figs. 8 and 11).

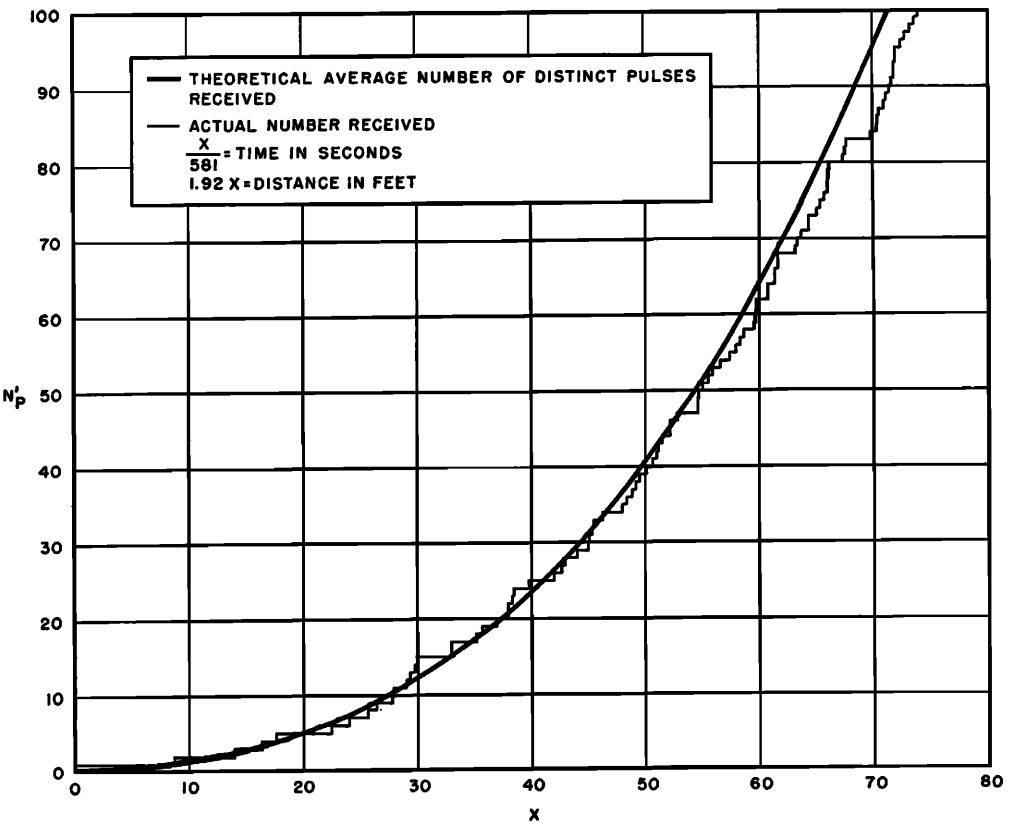



FIG. 13. The total number of distinct pulses $N_p'$ arriving at the receiver, plotted as a function of time, when the source and receiver are both in the corner of the room, as in Fig. 11.

or, using Eq. (30) and letting $\tau_p - \frac{1}{2}\gamma$:

$$\frac{p_s}{p_0} = \left(\frac{2\pi r_0^2 c \tau_p}{V}\right)^{\frac{1}{2}}. \tag{31}$$

This expression neglects dissipation of energy in the room which is easily taken into account if the reverberation time of the room is known.

### 4. The Period of Resolution

A period of resolution $t_r$ can be defined as the time at which the expected interval between successive pulses is just equal to the effective pulse duration $\tau_p$. At a time $t_r$ after the first pulse, one would no longer expect to "see" the individual echoes in the clear, but rather a smear characterized by an envelope above which the peaks of the pulses would occasionally appear.

An expression for the period of resolution $t_r$ in terms of the pulse width $\tau_p$ can be obtained from the correct (i.e., appropriate for the existing number of degeneracies) expression for $N_p$ or $N_p'$ given in Eqs. (10), (13)–(15) by solving the difference equation for $t_r$:

$$N_p(t_r) - N_p(t_r - \tau_p) = 1. \tag{32}$$

This has been done for the conditions of the experiments performed and leads to the following values for the period of resolution:

| | |
|---|---|
| Source and receiver in arbitrary positions: | $t_r = 9.98 \times 10^{-3}$ sec. |
| Source in corner, receiver in arbitrary position: | $t_r = 26.3 \times 10^{-3}$ sec. |
| Source and receiver in corners: | $t_r = 56.8 \times 10^{-3}$ sec. |

These values correspond to an effective pulse duration of $\tau_p = 1.74 \times 10^{-3}$ sec., which is the pulse width at an amplitude equal to $1/e$ times the peak amplitude. The resolving time as defined is depicted in each of the pulse pictures by a vertical line appearing, in every case but one, to the right of the initial pulse. The individual pictures will be discussed in detail later.

### 5. Envelope and Decay of Unresolved Sound

The height of the vertical line that indicates the resolving time has been adjusted to equal the r.m.s. sound level relative to the peak pulse level as obtained from Eq. (31). The subsequent decay of this more or less continuous sound is portrayed by the two slightly sloping horizontal lines that form the envelope of the unresolved sound. The slopes of the envelope were determined from the measured reverberation time of the room at 3600 c.p.s.

It will be noted that the size of the envelope is, in a number of cases, considerably too large (e.g., Figs. 8, 10, 11). In every case where the theoretical envelope does not fit the photograph, the source, or both the

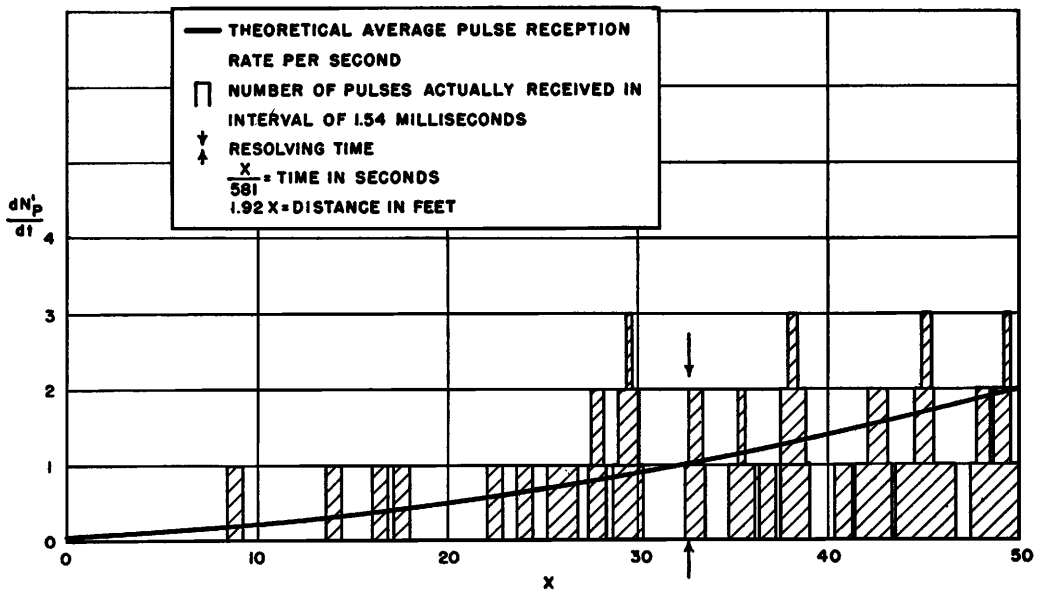
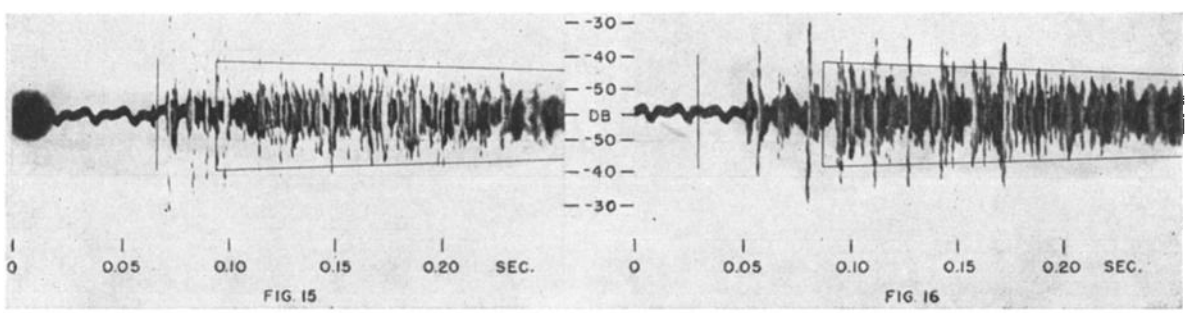



FIG. 14. The average and actual rates of arrival of pulses at the receiver when both source and receiver are in the corner of the room as in Fig. 11.




FIGS. 15 AND 16. Pulse responses for the receiver in other locations.

source and receiver, was in a corner position. Since the source was of finite dimensions it was never exactly in a corner but usually from 0.2 to 0.5 of a wave-length from one or more of the walls. Hence there was interference between the source and its images, resulting in a strongly directional *effective* source, and, consequently, in directional image clusters. Thus for the source in a corner and a particular location of the receiver, certain image sources will contribute very little effective energy to the receiver. This amounts to saying that the source effectively puts less energy into the room. This point will be made more evident from the discussion of the photographs.

## 6. Discussion of Pulse Photographs

The photographs were taken in a room of dimensions $L_x=23.0$, $L_y=13.4$ and $L_z=8.44$ ft. To specify the location of the source and receiver, we shall use $(X, Y, Z)$ for the source position and $(U, V, W)$ for the receiver position, distances being measured in feet.

In Fig. 5, the source coordinates are (11, 7.2, 3.9 ft.) and the receiver was located at (12, 7.2, 3.9 ft.). This position is non-degenerate and so the resolving time shown is $t_r=9.98$ msec. The envelope is seen to be slightly too large. This can perhaps be accounted for by the fact that both receiver and source are near the center of the room so that images from opposite pairs of walls interfere with each other consistently.

For Fig. 7, $(XYZ)$ are (1, 1.5, 0.5 in.) and $(UVW)$ are (5.75, 3.35, 2.11 ft.). The source is thus effectively eightfold degenerate, being in the corner, so the resolving time is $t_r=26.3$ msec. It will be noted that the direct pulse is exceeded in height by two succeeding pulses. The peak amplitudes of the pulses indicated by the squares are plotted on Fig. 4. This sequence corresponds to pulses from the successive images at $(0, 0, 0)$, $(0, 0, 2L_z)$, $(0, 0, 4L_z)$, and $(0, 0, 6L_z)$, each of which is a cluster of eight image sources. The differences in amplitude show clearly that each image cluster has distinct directionality, with especially great variations in the $X$, $Y$ plane. The long-dashed line on Fig. 4 is the peak level one would expect for incoherent addition of 16 sources (i.e., 12 db above peak for a single source). It is seen that the direct sound is 12 db below this line. However, the sound from the image sources lies along a line 14 db above the peak for a single source, which is,

in turn, 4 db below the level expected if the eight sources in each cluster all added coherently.

Figures 8–12 are a series in which $(XYZ)$ are still (1, 1.5, 0.5 in.) and in which $U=1$ in., $W=1$ in. for all photographs. In Fig. 11, $V=L_y-1$; in Fig. 12, $V=3L_y/4$; in Fig. 13, $V=L_y/2$; and in Figs. 14 and 15, $V=1$ in. The directionality of the image clusters is clearly evident in Fig. 8. Here there is no distinct "direct" pulse since the pulse from the image at $(0, 2L_y, 0)$ arrives simultaneously with the pulse from the actual source. Evidently these two pulses interfere since the height of the first pulse is small. The series of tall pulses is again from the images at $(0, 0, 2L_z)$, $(0, 0, 4L_z)$ etc., each of which arrives simultaneously with corresponding pulses from the images at $(0, 2L_y, 2L_z)$, $(0, 2L_y, 4L_z)$, and their reflections in the $X$, $Y$ plane. Each of these pulses then represents the simultaneous contributions of 32 image sources. As is evident by comparison of the peak heights with those in Fig. 7, the sources do not add coherently, however. The marked effect of interference is thus responsible for the low value of the actual envelope in this picture. The symmetry of the images corresponds closely to the symmetry when both source and microphone are in the same corner so the resolving time indicated is $t_r=56.8$ msec.

Figures 9 and 10 show effects similar to those just discussed for the other positions of the microphone mentioned previously. The resolving time for these is again $t_r=26.3$ msec. since the microphone position is no longer "degenerate." The influence of image-source directionality continues to be apparent.

In Figs. 11 and 12 both source and receiver are in the corner. Figure 12 serves to indicate the reproducibility of Fig. 11 which was possible with the experimental arrangements employed. The resolving time here is $t_r=56.8$ msec. The triangles refer to the points shown on Fig. 4. These pulses each represent the simultaneous contributions of the images at $(0, 0, \pm 2L_z)$, $(0, 0, \pm 4L_z)$ etc. As is evident from Fig. 4, the peaks very nearly follow the inverse square law. The peak amplitude, however, corresponds to 9 db above peak for a single source, whereas each pulse comes from 16 sources. The resulting peak in this case is less than would be expected for incoherent addition. Consequently it is again not surprising that the actual envelope is below the predicted value. The lines drawn on the lower half of the photograph indicate the calculated times of arrival of distinct pulses. Figure 13 shows the total number of distinct pulses $N_p'$ arriving up to a time $t$, as calculated from Eq. (15) for this location of source and receiver; and Fig. 14 shows the average rate of arrival, and a graphical portrayal of the actual pulses arriving. Each line on Fig. 11 corresponds to a pulse on Fig. 14. It is seen that agreement is quite good.

In Fig. 15 the source is still in the corner but the receiver is at $(\frac{1}{4}L_x, \frac{1}{4}L_y, 1$ in.). This picture should be compared with Fig. 10.

In Fig. 16 the source remains at (1, 1.5, 0.5 in.) while the receiver is placed in the opposite corner (nearly) at $(L_x, L_y, 0)$. Here the pulse from the source arrives almost simultaneously with the pulses from the three image clusters at $(0, 2L_y, 0)$ $(2L_x, 2L_y, 0)$, and $(2L_x, 0, 0)$. The resulting interference is clearly evident from the appearance of the first pulse. In this case the symmetry is again almost the same as that for both source and receiver in the same corner. Therefore $t_r=56.8$ msec. is used. Agreement of actual and calculated envelopes is seen to be fairly good.

Although the present experiments are restricted to an idealized room, a close connection can already be seen between these results in a small room and pulse studies in large auditoriums. For example, in Fig. 17 the source is at (1.5, 6.7, 1 ft.) and the receiver at $(\frac{3}{4}L_x, \frac{1}{4}L_y, 9$ in.). These locations correspond to typical locations of a speaker and listener in a large rectangular hall. If we consider a scaling factor of about 5, the dimensions of this hall would be $115\times67\times42.2$ ft., and our pulse would have a corresponding width (at $1/e$ peak amplitude) of 8 msec. The scaled carrier frequency would be about 700 c.p.s. The resolving time is seen to be shorter than the time actually required for the direct pulse to reach the receiver, and the character of the response shows that a "smear" sets in immediately. The actual and calculated envelopes are seen to agree well in this case, since no degeneracies are present.

Figures 18 and 19 can be compared qualitatively with Fig. 17. These were obtained in a motion picture theater of about 250,000 cu. ft. volume. Pulses were produced by applying a short pulsed carrier signal to the theater loudspeaker. In Fig. 18 the microphone was at a seat in the side section of the main floor about two-thirds of the way back from the stage, under the balcony. The pulse length was 5 msec., and the carrier frequency 2800 c.p.s.

In Fig. 19 the microphone was in the center section of the main floor, well in front of the balcony and somewhat off the center line. The pulse length was again 5 msec. and the carrier frequency was 1500 c.p.s. Qualitative observations indicated that the "hour glass" bulge in Fig. 19 can be correlated subjectively with a "slap echo" from the hard but irregular rear wall. This bulge

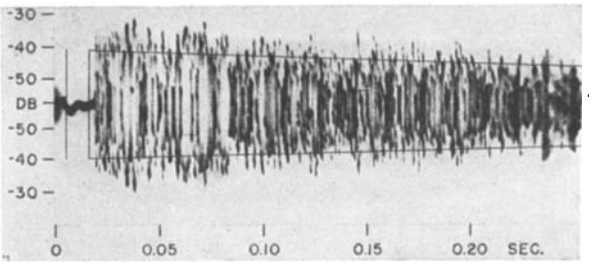
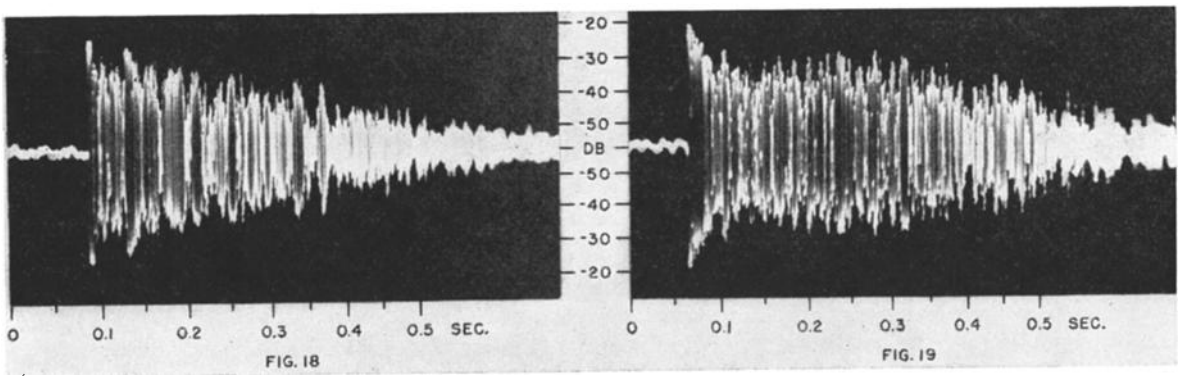



FIG. 17. Pulse response for non-degenerate locations of source and receiver which were chosen to correspond respectively to the positions of a speaker and listener in a large rectangular hall.




FIGS. 18 AND 19. Typical pulse responses in a motion picture theater of about 250,000 cu. ft. volume. Note qualitative similarity with Fig. 17.

was characteristic of pulse responses at this location for nearly all carrier frequencies. A more detailed analysis of these and similar photographs is in progress in an attempt to correlate various features of the pulse response with the results of subjective listening tests.

## CONCLUSIONS

Most of the theoretical and experimental work presented in this paper is directly applicable only to hard-walled rectangular rooms. However, as has been pointed out, appropriate modifications of the images may be made when the walls are not hard and in some cases modifications for the first few images do not involve too great computational difficulties.

It appears that fairly detailed experimental and theoretical investigations of the short term transient response (e.g., the first 100–200 msec.) for rooms having various shapes and absorptive treatments should prove highly instructive. At the same time, the statistical methods outlined in this paper can be applied to the analysis of the fluctuation characteristics of the long term response.

The experimental results presented show that appropriate transients can be easily produced and observed, and, in simple cases at least, correlated with boundary conditions. Also it appears that pulse methods are readily adaptable to scale model experiments.

## ACKNOWLEDGMENT

The authors wish to thank Dr. David Mintzer for permission to read his paper prior to publication and for participating in several discussions.