

M. R. SCHROEDER

MUSIC PERCEPTION
IN
CONCERT HALLS

Paper given at a seminar
organized by the Committee
for the Acoustics of Music
in October 1978

Publications issued by
the Royal Swedish Academy of Music 26

Stockholm 1979

M. R. SCHROEDER

MUSIC PERCEPTION

IN

CONCERT HALLS

Paper given at a seminar
organized by the Committee
for the Acoustics of Music
in October 1978

Publications issued by
the Royal Swedish Academy of Music 26

Stockholm 1979

CONTENTS

Preface	III
M.R. Schroeder: Music perception in concert halls	1
References	27
Sound examples	30

ISBN 91-85428-14-0

© 1979 by The Royal Swedish Academy of Music, Stockholm, Sweden

Ekenäs Tryckeri Aktiebolag, Ekenäs, Finland

P R E F A C E

In collaboration with the Swedish Acoustical Society and the Center for speech communication, research and musical acoustics the Committee for the Acoustics of Music, which is sponsored by the Royal Swedish Academy of Music, arranged a seminar with Dr Manfred Schroeder, Drittes Physikalisches Institut, Universität Göttingen, as the only lecturer. This seminar was held at the Royal Institute of Technology, Stockholm in October 1978. Schroeder's three lectures concerned "Fundamentals of Room Acoustics", "Music Perception in Concert Halls" and "Speech Perception in the Presence of Reverberation and Echoes". In this booklet we present a condensed version of these lectures, which Schroeder has prepared. The sound illustrations are included in an EP-grammophone record. It contains examples of orchestral music recorded with the musicians playing in an anechoic room plus this same music signal with digitally added reverberations of various kinds. Also, the record contains music combined with the actual acoustic characteristics of a number of concert halls. These quite unique examples should ideally be reproduced over loudspeakers in an anechoic room. However, the main effects will remain even if the examples are reproduced over loudspeakers in an ordinary room with shorter reverberation than the concert halls.

The editorial work has been done by Doc. Erik Jansson and the grammophone record was produced by Lennart Fahlén. We are indebted to Dipl. Phys. Herbert Alrutz for the preparation of the sound illustrations.

The Royal Institute of Technology
October, 1979

Johan Sundberg
President of the Committee for
the Acoustics of Music



The author listening to music coming from loudspeakers suspended in Bell Laboratories anechoic chamber - a room whose walls, ceiling, and floor are lined with highly sound absorbing "wedges" to eliminate echoes that would interfere with acoustic measurements and subjective evaluations. The music radiated from the loudspeakers has been processed by a digital computer to simulate sound transmission in a concert hall. In this manner "the acoustics" of future halls can be tested before they are built.

(Schroeder, Am. J. Phys. 41/4, April 1973)

MUSIC PERCEPTION IN CONCERT HALLS

M.R. SCHROEDER

Drittes Physikalisches Institut
Universität Göttingen
Göttingen, Germany

Why, precisely, is concert hall acoustics such a difficult subject? The reason, simply stated, is that concert hall acoustics is governed by three sets of interacting problems, all three of them exceedingly complex in themselves: (1) The physical problem of wave propagation and attenuation in irregularly shaped enclosures; (2) the psychological problem of how we, the human listeners, perceive the sound waves impinging on our ears; and, finally, (3) the problem of subjective preference: What do people really prefer to hear, do tastes change etc.?

If concert halls were bare rectangular cells with neither chairs, human performers nor listeners inside, the physical problem would be easy. The wave equation for sound in air (at levels below the threshold of pain) is linear and simple enough and, for the rectangular geometry, the independent variables (the three spatial coordinates and time) are separable. Thus, the wave equation for the sound pressure (a second-order partial differential equation) can be integrated explicitly leading to the well-known sinusoidal (or complex exponential) solutions in space and time. These solutions are the normal modes - much talked about by expert and layman alike, but audible as such only in small, tiled bathrooms and similar enclosures.

Normal modes are, of course, not the prerogative of bare, rectangular chambers but prevail in more complicated shapes too - with one important difference: They can no longer be calculated and specified explicitly. The magnitude of the problem is further aggravated by the fact that a large hall, such as Boston Symphony Hall, has a total of more than ten billion normal modes in the audio frequency range (20 - 20 000 Hz for young listeners), not one of which we know or can ever hope to know.

At this stage of comprehension of the magnitude of the problem it would seem to be quite reasonable to give up any attempt of shedding light on concert hall acoustics by physical theory. But that is not in human nature. The very complexity of the problem suggests other viewpoints and, in fact, implies different solutions.

The enormous number of normal modes means that the wavelengths of audible sound are small compared to the overall dimensions of the hall. Thus, at least, for smooth inner surfaces, one can, as an approximation, assume the wavelength to be zero, thereby neglecting all effects of wave diffraction and arriving at a description of sound propagation in terms of rays - just as the light rays emerge from electromagnetic theory by letting the wavelength go to zero.

Geometrical Acoustics

The discipline based on the ray approximation is commonly called "geometrical acoustics" in analogy to geometrical optics.

Geometrical acoustics has been, and continues to be, one of the mainstays of concert hall design. Because of the close analogy to light rays, the propagation of sound in an enclosure is often studied by means of optical scale models lined with mirrors (and sometimes filled with smoke to enhance the visibility of the light rays).

In recent times, ray tracing by computer has given further impetus to geometrical acoustics. In fact, one of the oldest but still unresolved problems of architectural acoustics is presently being successfully attacked by computer ray studies: the relationship between the reverberation time of an enclosure and the sound absorption on its walls. (The absorption of sound in air is in general negligible, except at high frequencies and, in any case, presents no theoretical difficulties. The reason is that the air, and therefore its absorptive effect, is spread evenly throughout the entire three-dimensional space as opposed to the wall absorption which is distributed highly non-uniformly, being confined to a two-dimensional subspace - the surface of the enclosure.)

Reverberation time (T) is defined as the time in which the sound energy density at a given point decays to one millionth of its starting value after all sound sources have been turned off.¹ Experience teaches us that reverberation time, in most halls, is not a strong function of location and initial conditions and in this sense we can speak of the reverberation time of a hall. (However, reverberation time may depend on frequency and on whether the hall is empty or filled with people. If no frequency value is stated, quoted values usually refer to frequencies between 500 - 1000 Hz. Example: Vienna Grosser Musikvereinssaal T = 2.0 sec.)

Modern concert halls are designed so that the unoccupied seats have very nearly the same absorption as that of an adult person sitting in that seat. Thus, reverberation time becomes independent of attendance. In older halls, with wooden seats and little upholstery, reverberation time of the empty hall may be 50% higher (or more) than in the fully occupied hall. This can make orchestra rehearsals extremely unpleasant unless a "rehearsal curtain" is utilized to shut out excessive reverberation from the empty hall.

At frequencies below 250 Hz, one often finds an increase in reverberation time by 20% or more (a significant increase,

Most experts agree that such an increase in reverberation time at low frequencies is desirable for music (as opposed to speech).

The relation between sound absorption and reverberation time was studied extensively by Wallace Clement Sabine at Harvard University around the turn of the century. He derived a formula (named after him) for the reverberation time

$$T = 55.2 V / cS\alpha, \quad (1)$$

where V is the total volume of the enclosure, S its surface area, c the velocity of sound and α the (average) absorption coefficient of the walls. Its derivation proceeds roughly as follows. The energy as a function of time, of one wave packet of sound energy travelling as a ray is given by

$$E(t) = E(0) \cdot (1-\alpha)^{n(t)} \quad (2)$$

where $n(t)$ is the number of "collision" of the ray during the time interval $(0,t)$ with the absorbing walls, assumed here to have a uniform absorption coefficient α .

By the definition of α , the energy of the ray is reduced by a factor $(1-\alpha)$ after each wall reflection. Thus, Eq. (2) follows immediately.

Different sound rays have, of course, different fates. Thus, $n(t)$ depends on the particular sound ray considered. The average energy of many sound rays is given by an ensemble average:

$$\langle E(t) \rangle = E(0) \cdot \langle (1-\alpha)^{n(t)} \rangle. \quad (3)$$

Unfortunately, the average on the right-hand side of Eq. (3) is unknown in all but the most trivial cases.

What is known is the average number of collisions for a spatially uniform ensemble of rays:

because a 5% difference can be perceived by an attentive listener).

$$\langle n(t) \rangle = cSt/4V \quad (4)$$

Equation (4) is a most interesting result: It asserts that $\langle n(t) \rangle$ is independent of the shape of the enclosure being a function only of its volume V and the surface area S .

Given the insolubility of Eq. (3) and the neat result of Eq. (4), physicists, throwing mathematical rigor to the wind, gave in to the temptation of transferring the ensemble average in Eq. (3) to the exponent, resulting in

$$\langle E(t) \rangle = E(0) \cdot (1 - \alpha)^{cSt/4V}. \quad (5)$$

Now with $\langle E(T) \rangle = E(0) \cdot 10^{-6}$, one easily obtains

$$T = -55.2V/cS \ln(1 - \alpha). \quad (6)$$

This is the Eyring-Schuster-Waetzmann reverberation time formula which, as $\alpha \rightarrow 0$, goes over into Sabine's formula [Eq. (1)].

For enclosures with surfaces having nonuniform absorption coefficients, the α in Eq. (6), or Eq. (1), is replaced by an average absorption coefficient:

$$\alpha = \sum_i \alpha_i S_i / S ,$$

where α_i is the absorption coefficient of a subsurface and S_i its area.

These formulas, neat, simple, and easy to apply, have but one drawback: they are incorrect. Even after allowing for wave diffraction effects, which are, of course, not included in the ray theory on which these formulas are based, there are two main inaccuracies. In applying Eqs. (1) or (6) to concert hall design, the resulting reverberation time usually turns out to

be smaller than predicted - giving a noticeably "dry" sound. Many modern concert halls are mute witnesses to this acoustical deficiency.

Another important application of the reverberation time formulas is the determination of the sound absorption coefficient α of new acoustical materials by measuring the reverberation time T in a so-called "reverberation chamber."

In this method, reverberation times of the bare chamber and the chamber partially lined with absorbing material are measured.

Values for α determined in this manner are usually too large. In fact, calculated values of the absorption coefficient often exceed 100% - a physical impossibility for a passive material!

Computer Ray Studies

Ray studies on the computer have shed considerable light on this perplexing situation. They have shown, in particular, that reverberation time depends not only on volume, surface area, and average absorption coefficient, but also on the shape of the enclosure and the distribution of the absorbing materials.

Figure 1 illustrates one of the many two-dimensional configurations with which these ray studies by B.S. Atal of Bell Laboratories and the author were begun. The "enclosure" is a quadrangle, as shown in Fig. 1, with one "surface" highly absorbing ($\alpha = 0.8$) while the others have no absorption ($\alpha = 0$). A sound source near the lower right-hand corner emits 300 rays of equal initial energy at angles 1.2° apart. The computer follows these 300 rays through many reflections until their energy has decayed to less than one millionth of the initial value. The reflection can either be specular (as shown for one ray in Fig. 1) or random ("diffuse") following, for example,

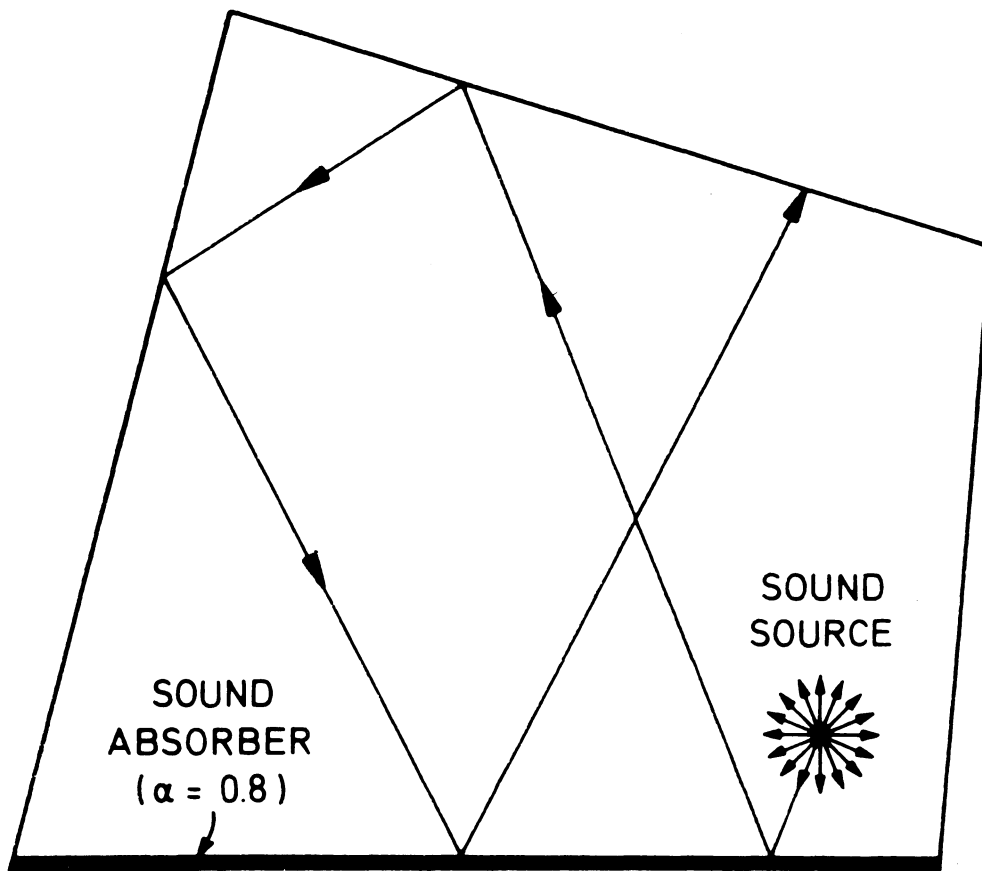


Fig. 1 Configuration for ray simulation on the digital computer.
(Schroeder, Am. J. Phys. 41/4, April 1973.)

Lambert's cosine law as would be expected in the case of highly irregular walls (as are found in many older concert halls).

Absorption coefficients for most materials depend, of course, on angle of incidence. Presently available reverberation theories take this into account by working with an average absorption coefficient for random incidence, i.e., for angles of incidence uniformly distributed over the solid angle. Naturally, any deviation from this angular uniformity can lead to errors.

The total energy as a function of time is shown in Fig. 2 (the "wiggly" curve). Energy is plotted in decibels (dB), a logarithmic unit for energy ratios, 10 dB corresponding to an energy ratio of 10.

The reverberation time, i.e., the time for a 60-dB decay, can be easily determined by a straight-line fit to the computed decay curve. The result obtained for the configuration shown in Fig. 2 is

$$T_{\text{simulated}} = 0.38 \text{ sec.}$$

The corresponding theoretical formulas for two-dimensional enclosures give considerably higher values:

$$T_{\text{Sabine}} = 0.63 \text{ sec,}$$

and

$$T_{\text{Eyring}} = 0.56 \text{ sec.}$$

These discrepancies have, indeed, been traced to the "sloppy" averaging in the derivation of the theoretical formulas.

In the meantime, the ray simulation studies are being extended by R. Gerlach, a student of the author at the University of Göttingen, to three-dimensional shapes.

What are the potential benefits of these ongoing studies for the acoustical consultant? For one, instead of relying on imperfect formulas, he could work out a specific reverberation time formula for his hall by ray simulation on the computer before construction starts. Naturally, for complicated hall

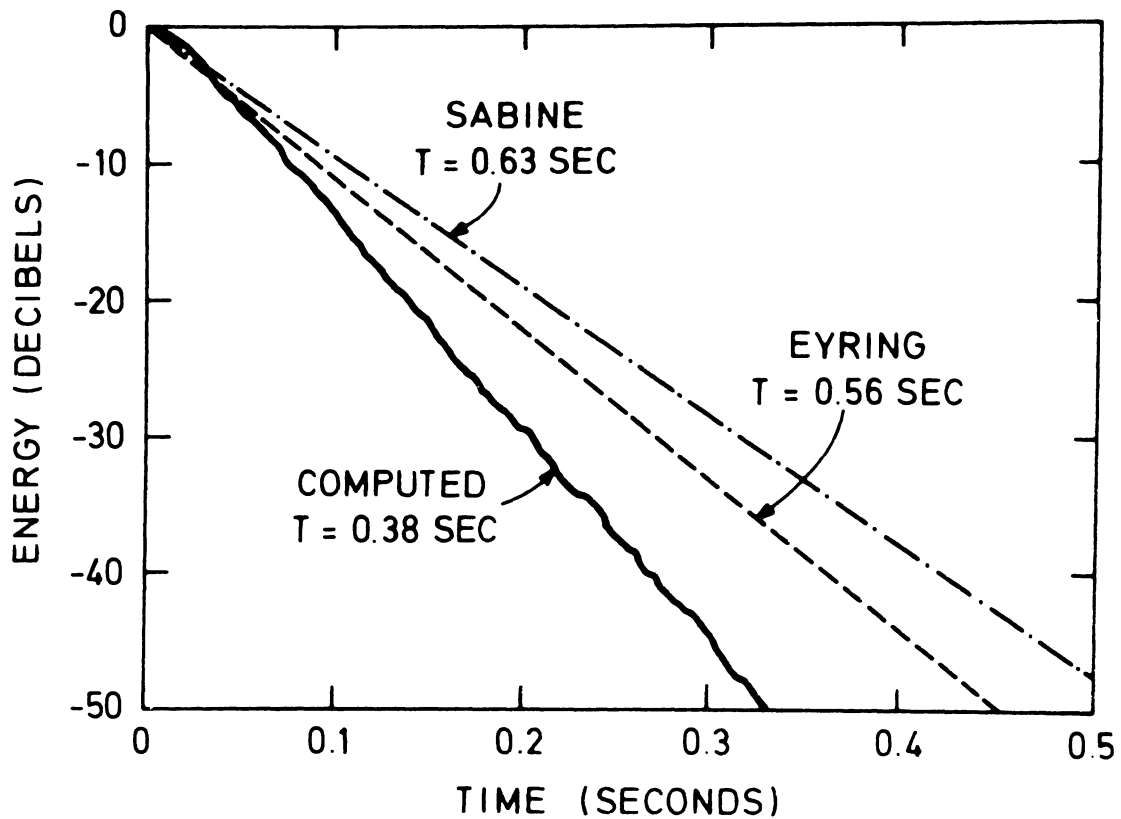


Fig. 2. Solid curve: sound decay obtained by ray simulation on the computer for the configuration shown in Fig. 1. Broken straight lines: decay predicted by Sabine and Eyring formulas, respectively. (Schroeder, Am. J. Phys. 41/4, April 1973)

shapes, the programming won't be easy. But there are no principal obstacles and the extra expense ought, no doubt, be less than the costly mistakes revealed by the finished hall.

In addition to this pragmatic use of digital ray tracing, it is hoped that continued theoretical work, aided by computer simulation, will result in new reverberation time formulas that take dependence on hall shape and absorber distribution into account - perhaps in the form of 2 or 3 additional parameters related on the geometry of the hall.

Computer Wave Studies?

If we can trace sound rays on the computer, why can't we simulate sound waves by digital machines? A simple calculation shows how utterly impossible that would be. The sampling theorem tells us that we must know the sound pressure at points closer than $\lambda/2$ apart, where λ is the shortest wavelength of interest. In the time domain, we must know the sound pressure at least every $1/2f$ seconds apart, where $f = c/\lambda$. Thus, for a space with volume V , we must perform at least $16 cV/\lambda^4$ "calculations" per second, each calculation comprising several basic computer operations lasting, at present processing speeds, perhaps 10^{-5} sec. For a volume of 10^4 m^3 , a shortest wavelength of $\lambda = 2 \cdot 10^{-2} \text{ m}$ and with $c = 340 \text{ m/sec}$, the time factor (computing time divided by real time) exceeds $3 \cdot 10^9$. While this figure may not impress the astronomer or particle physicist as excessively large, it does mean 100 years of computing time for every second of sound - not counting core storage, reliability, and other problems.

However, the digital computer is not completely out of pace on the "wave front." The exceedingly large number of normal modes, even within the bandwidth of a single mode, permits the successful application of statistical methods to the ensemble of modes. In

any auditorium, the turning-on of a single frequency in the audio frequency range (radiated, for example, from a loudspeaker) will excite a large number of normal modes which, at some distance from the source, will add with random amplitudes and phases. As the frequency is slowly increased, the complex transmission function between loudspeaker position and point of observation changes likewise in a random manner. For practical purposes, the transmission function between two points in an enclosure can be considered a complex Gaussian process with a one-sided exponential spectrum. (Note that the spectrum of a function of frequency is itself a function of a variable with the dimension of time.) The Gaussian behavior follows directly from the Central Limit theorem of probability theory applied to the case of a large number of independent normal modes. The spectrum of the transmission function is, of course, nothing but the squared impulse response between source and receiver positions.

Thus, although we know almost nothing about the normal modes of an irregularly-shaped enclosure, we know everything that is knowable about its sound transmission function if the shape details of the enclosure are not specified. The situation is analogous to statistical mechanics where, without knowing the exact motions of the individual molecules, we can still make very precise statements about average quantities.

Some of the average quantities in the case of the acoustics of large halls are the average level difference between maxima and minima of the sound transmission function (roughly 10 dB) and the average distance between response maxima ($4/T$).

In 1935, E.C. Wente (see reference in footnote 2) of Bell Laboratories first reported these wild fluctuations of room response curves and, as a telephone engineer, was understandably worried by them. He wondered how man could hear as well as he did in

reverberant enclosures in the face of such response irregularities.

The question was not settled until 25 years later when it was shown³ that the human ear, in listening to speech or music, has a time resolution of about 1/50 sec and a correspondingly poor frequency resolving power that "irons out" most of the spectral irregularities of a room.⁴ However, in the intervening decades, a large number of frequency response curves were measured in auditoriums around the world in a vain effort to distill criteria for acoustic excellence from what must be considered essentially noise (albeit in the frequency domain).

Nevertheless, room response curves are a relevant measure when electroacoustic systems (such as public address systems) are to be installed. Application of Nyquist's criterion for stability (freedom from "howling") of such systems leads to a problem in extremal statistics for which only approximate analytic solutions exist.⁵ Here the digital computer has closed a significant gap by its ability to generate, Monte Carlo fashion, complex transmission functions en gros and using a simple threshold logic to implement the Nyquist criterion.^{6,7} The result of such Monte Carlo computations on the digital computer has been that public address systems become unstable for an open-loop gain of about -12 dB relative to the average power gain. Some of this loss in stability can be recovered by inserting an electronic "frequency shifter" (single-side-band modulator with a 5-Hz carrier) into the acoustic feedback loop somewhere between microphone and loudspeaker.⁵

Several other problems in wave acoustics have been solved by a proper statistical formulation of the situation and a Monte Carlo simulation on the computer based on that formulation.

Digital Simulation of Sound Transmission

So far we have mentioned the simulation by digital computer of sound rays and random wave fields. The purpose of these simulations is to elucidate some of the physical problems in room

acoustics. What about the subjective questions of sound quality and individual preference? Can they be attacked in a systematic manner also?

If we could describe human hearing and the functioning of our brains by mathematical formulas - as we describe sound transmission outside the ear by the wave equation - there would at least be hope to tackle the subjective problems in the accustomed manner. It is a truism that we cannot do this. There is, at present, only one "measuring process" that can determine the subtleties of acoustic quality relevant to the enjoyment of music and that measuring process is a human being listening to music.

How we satisfy this requirement for actual listening is determined by economic and other factors. Thus, we could build a concert hall, based on what we believe to be the best available information, and have it listened to for a few months. Then we tear it down and build a new hall etc. If the economic constraints should be unreasonably stringent, one could perhaps compromise occasionally and only rip out the interior of the hall. At any rate, practical experience with attempts along these lines have proved to be rather unprofitable.

A more sensible approach is the building of scale models of proposed halls. In such models, sound, translated in frequency by the scale factor, can be radiated from tiny loudspeakers and recorded on high-speed tape recorders. On playback, the tape speed is slowed down by the appropriate factor and the result can be listened to and evaluated subjectively.

This method, as one might expect, has a number of difficulties. While frequencies scale easily, in a manner reciprocal to geometrical dimensions, absorption factors are more cumbersome to translate. In modelling the audience area, for instance, what material has the proper absorptive properties between 2000 Hz and 400 000 Hz (assuming the scale factor to be 20)? Obviously, compromises must be made.

In another approach, one records, on an oscilloscope, acoustic impulse responses from small electrical sparks located, for example, on the "stage" of the scale model.

These impulse responses, lengthened by the scale factor, are then simulated by electrical networks to be used as filters for the selected program material.

In another possible variant of this method, the impulse response is simulated digitally, i.e., as a digital filter on a general purpose computer, thereby allowing greater precision and flexibility. In fact, the microphone output from the model could be fed directly into the computer, without prior recording, and converted automatically into a digital filter response sample by sample. To the author's knowledge, this attractive "hybrid" (part analog, part digital) possibility has not yet been exploited.

But a refinement of this method has been applied to a study of the acoustic qualities of existing concert halls by the author and two of his students (D. Gottlob and K.F. Siebrasse) at the University of Göttingen. Impulse responses from a powerful spark gap have been recorded in 20 major European concert halls including Royal Festival Hall (London), the Berlin Philharmonie, Amsterdam Concertgebouw, and the famous Grosser Musikvereinsaal in Vienna. The recordings were made with a carefully designed dummy head "seated" in various locations in these halls. Two high-quality microphones form the dummy's eardrums and their outputs are recorded on a two-track tape machine.

After these stereo-impulse responses have been transferred into the computer's memory, digital tape recordings of reverberation-free music (recorded by the Orchestra of the British Broadcasting Corporation in a "dead" room and kindly made available to us) are "convolved" (i.e., filtered) with these impulse responses resulting in a pair of signals corresponding precisely to the acoustic signals at the dummy's ear drums had the B.B.C. Orchestra in fact produced the same sound in one of the halls under study.

The remaining problem is to transfer these two signals to the eardrums of a human listener for subjective evaluation. Simply converting these digital signals into analog form and applying them to stereo earphones will not do because the auditory impression when listening over earphones is generally not the same as listening in a "free" sound field. One shortcoming of the earphone presentation is fairly obvious: in turning one's head the acoustic image (and therefore the entire hall) would turn with the head. Also, in listening over earphones, one has difficulty in "externalizing" the sound. It is as if the music originated in one's head. Such an unnatural acoustic condition would, of course, interfere seriously with any subtle quality distinctions.

How then do we get the proper acoustic signals to the listener's ear drums in a free-field listening condition? The answer was demonstrated by B.S. Atal and the author at Bell Laboratories in 1962.^{7,8}

Implementation of the method requires the prior measurement of the complex transmission functions $S(f)$ and $A(f)$ (i.e., amplitude and phase as functions of frequency) between a loudspeaker in an anechoic chamber and the right and left eardrums of a human listener (or a dummy). In the original experiment, the loudspeaker was located in front of the listener at a distance of about 3 m and 22.5° to the right.

If $A(f)$, the transmission function to the left ear, were zero, i.e., if there no "crosstalk" from the loudspeaker on the right to the left ear, the task would be simple: filtering the loudspeaker signal by the inverse of the transmission function to the right ear, $S^{-1}(f)$, would be all that is necessary. However, there is cross talk, due to sound diffraction around the human head, and it must be cancelled.

This can be done by a second loudspeaker, 22.5° to the left, radiating an appropriately filtered signal.

A solution to this filtering and cancellation problem is illustrated in Fig. 3, where $C(f) = -A(f)S^{-1}(f)$ is the crosstalk

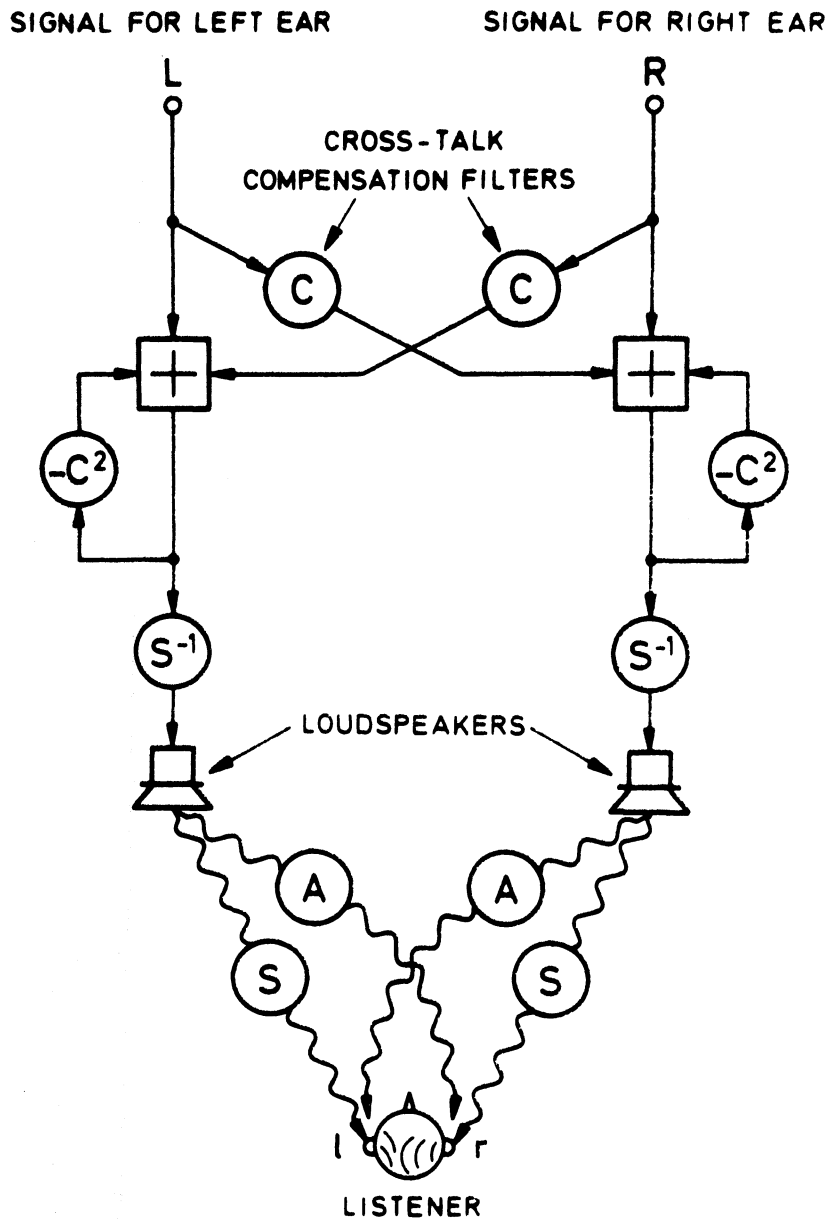


Fig. 3. Sound reproduction method for digital simulation of concert hall acoustics. (Schroeder, Am. J. Phys. 41/4, April 1973)

compensation filter. The overall transmission function from the right input (R) to the right ear (r) is

$$Rr(f) = (1 - C^2)^{-1}S^{-1}S + C(1 - C^2)^{-1}S^{-1}A . \quad (7)$$

With $C = -AS^{-1}$, $Rr = 1$ as required.

The overall response from the right input to the left ear (l) is

$$Rl(f) = (1 - C^2)^{-1}S^{-1}A + C(1 - C^2)^{-1}S^{-1}S, \quad (8)$$

which for $C = -AS^{-1}$ vanishes, also as required.

For the method to work, $A(f)S^{-1}(f)$ must be a realizable filter response, i.e., its inverse Fourier Transform must vanish for negative times. Measurements have shown that this is indeed the case within measurement accuracy. (The author wonders under what general conditions this is true. In other words, what are the necessary conditions that the ratio of the transmission functions from a point in space containing diffracting surfaces to two other points be a realizable filter response.)

In contrast, the filter $S^{-1}(f)$, which also appears in Fig. 3, does not have to be realizable because it does not occur inside a feedback loop. In fact, $S^{-1}(f)$, containing a substantial negative delay, is never realizable. However, for a suitable large positive delay, r , $S^{-1}(f) \exp(-j2\pi fr)$ is always realizable within arbitrarily small error bounds.

The practical experience with the schema of Fig. 3 has been nothing less than amazing. Although the two loudspeakers are the only sound sources, virtual sound images can be created far off to the sides and even behind the listener. In fact, even the elevation angle of a sound source is properly perceived. Since the entire system is linear, many sound sources and their echos can be reproduced simultaneously, without mutual interference, provided the listener is sitting in the proper position between

the loudspeakers and does not turn his head away from the front direction by more than about $\pm 10\%$. The spatial illusion is, indeed, so convincing that one is tempted to "look around" for the invisible sound sources. The moment one gives in to this temptation, the realistic illusion, of course, vanishes - frequently changing into an "inside-the-head" sensation.

The sound reproduction method illustrated in Fig. 3 has opened up completely new possibilities in the study of concert hall acoustics. Before, in comparing two halls, one had to base one's judgement on listening to pieces of music, played at different times, often by different orchestras under different conductors. Even if all other factors were equal, the fact that two musical experiences are separated by days, weeks, or even months makes any subtle quality assessments exceedingly unreliable if not impossible.

So far, 20 European halls have been examined by this method, using paired comparison for subjective similarity. The resulting psychological preference and similarity spaces, obtained by multidimensional scaling, are highly correlated showing that common factors underlie these judgements of acoustical quality.

With the new reproduction method instantaneous comparisons of identical program material has become possible. The author will never forget the moment he first switched himself from a seat in the Berlin Philharmonie to one in the Vienna Musikvereinssaal listening to the B.B.C. Orchestra playing Mozart's Jupiter Symphony. All he believed about the difference between these two halls based on previous visits (but wasn't too sure about) suddenly became a matter of easy distinction.

An important aspect of this psychological testing method based on multidimensional scaling of preference and similarity is its avoidance of any undefined subjective descriptors such as "warmth", "brilliance", "clarity", "fullness", "immersion", etc. (several dozen of these more or less picturesque terms are known in the art) which may mean different things to different people (assuming they mean anything).

Correlation with objective parameters

Among the objective parameters we have considered are the following:

- V the volume of the hall
- W the width of the hall (at the first sidewall reflection in the case of nonrectangular halls)
- G the time delay ("gap") between the direct sound and the first reflection at the listener's seat
- T the reverberation time obtained by a straight-line fit to the first 15 dB of the decay obtained by the integrated tone-burst method
- D the "definition", i.e., the energy in the first 50 msec of the impulse response divided by the total energy
- C the interaural "coherence" (the maximum of the cross-correlation function between the impulse responses at the two ears).

The geometric parameters V and W are obtained from drawings of the halls. The acoustic parameters G, T, D, C are obtained from measured impulse response at the dummy's ears with a powerful (5 kV) spark source on the stage.

Figure 4 shows the six objective parameters as vectors in the two-dimensional preference space for 11 halls. The component of an objective parameter vector in the direction of D1 or D2 is the correlation coefficient of that parameter with the projection of the different halls on D1 or D2. The direction of an objective parameter vector is the direction of maximum correlation with the different halls. The vector's length indicates the magnitude of that correlation. (A correlation of 1 corresponds to the distance from origin to one of the endpoints of the abscissa or ordinate.)

According to Fig. 4, reverberation time T is almost coincident with D1, the consensus preference factor. Thus, the greater the reverberation time, the greater the consensus preference from these halls.

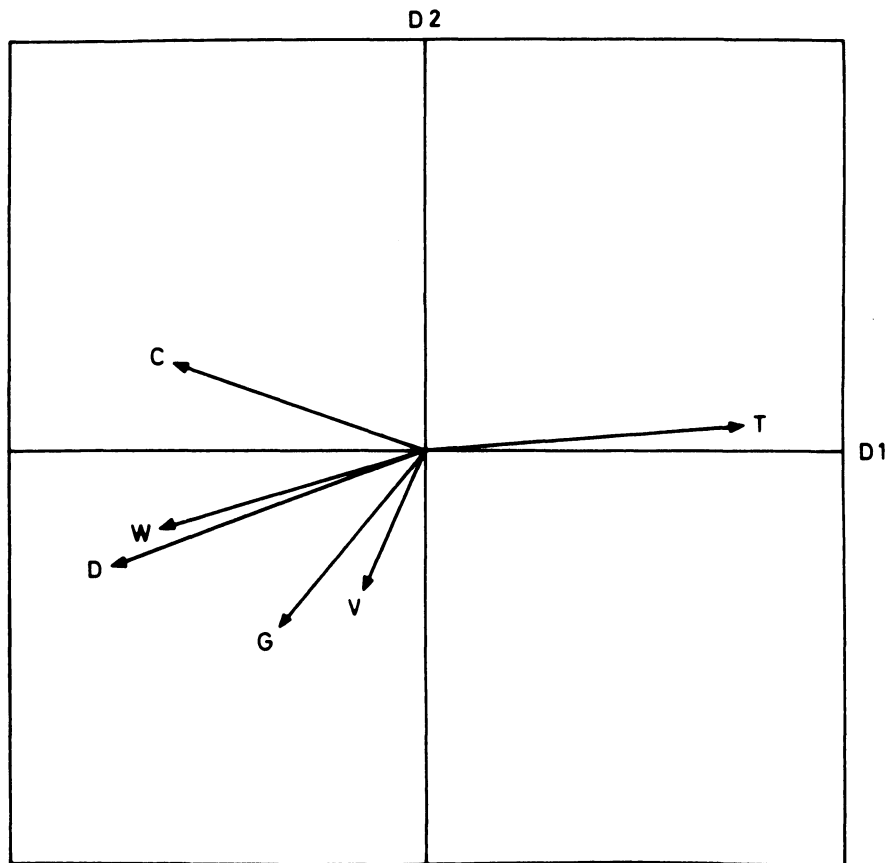


Fig. 4. Correlations of objective parameters of 11 halls with their subjective positions (V=volume, W=width at first side-wall reflection, T=initial reverberation time, D="definition", G=initial time delay gap, and C=inter-aural coherence.
(Schroeder et al., J. Acoust. Soc. Am. 56/4, 1974)

Conversely, the definition D (the relative energy in the first 50 msec of the impulse response) has a highly significant negative correlation with preference. However, this is not an independent datum because for these (and all other tested) halls we found a high negative correlation between reverberation time T and definition D: the greater the T, the smaller the D, and vice versa, as expected.

Another objective factor showing high (negative) correlation with D1 is the width W of the halls (for non-rectangular halls, W is the width at the first side-wall reflection). However, W also shows a substantial correlation (-0.62) with T for our 11 halls. Thus, while width may be a significant parameter (in the sense that wider halls are less preferred), we do not consider our present data as sufficient to draw a firm conclusion on hall width. But there is enough of a suggestion here of a possible subjective significance of a hall's width to tempt one to take a closer look.

The interaural coherence C is again almost uncorrelated with T and shows a strong (negative) correlation with dimension D1, so that we can again tentatively conclude that interaural coherence is a subjectively significant parameter. As before, smaller coherence means greater preference. (This effect might be mediated by a more pronounced feeling - of being immersed in the sound - that presumably occurs for less coherent ear signals. Whether coherence is related to sound "diffusion" and whether it can be influenced by sound diffusors - on the side walls, on the ceiling, or suspended from the latter - remains to be seen.)

Interestingly enough, the volume V of the halls also becomes a significant parameter for these (generally larger) halls. The correlation with consensus preference is, in fact, the most negative of all objective parameters. The implication of this finding is fairly obvious (if economically disastrous): once a hall has reached a certain size, don't make it any bigger - or be prepared to suffer acoustically.

These results imply that sound arriving in the median plane of a listener is detrimental to subjective preference because such

median-plane sounds lead to identical sound pressure waves at the two ears thereby decreasing binaural dissimilarity. (An explanation for the detrimental effect of median-plane sound is that it leads to a more "monophonic" - as opposed to "stereophonic" - listening experience.

This conclusion is confirmed by another result of the investigation: The highest correlation between subjective preference and a geometrical parameter of the hall was with the width of the hall: The narrower the hall, the higher the preference and vice versa. Narrow halls, of course, have more numerous and more powerful early lateral reflections that arrive at the ears near the horizontal plane as opposed to the (vertical) median plane. Such early lateral reflections lead to a greater binaural dissimilarity and therefore higher listener preference.

More recently, even more direct evidence for the importance of early lateral reflections has been obtained by the author and his collaborators at Göttingen.⁹⁻¹³ While in the study of real halls different physical parameters come in a "predetermined mix", which can only be disentangled statistically by appropriate subjective scaling methods, the technique of digital simulation¹⁴ can be employed to alter one physical parameter at a time and to study its effect on musical preference independently of all other potentially confusing factors.

Accordingly, binaural impulse responses from real halls were fed into a computer and digitally modified by adding (or deleting) lateral reflections. When the effect of these modifications is expressed as binaural similarity (as defined above), it was found that subjective preference reached its maximum value for zero binaural similarity⁹. The lowest subjective preference score was obtained for the highest value of binaural similarity. Interestingly, low preference was also found for negatively correlated earsignals. Thus, within the scope of these investigations, it can be concluded that zero binaural similarity is optimum even if in actual enclosures this goal cannot be completely realized.

Ceiling shapes

How then can we design concert hall ceilings that avoid direct (median-plane) reflections into the audience area? Absorption is verboten because we need the energy from the ceiling for reverberation. However, we can shape the ceiling to reflect most of the sound energy to the side walls, whence it will reach the listener, directly or indirectly, from lateral directions.

One possible ceiling shape would be a convex (curved downward over the "center aisle") hyperbolic cylinder whose axis runs parallel to the front-back axis of the hall. A hyperbolic surface, as is well known, reflects rays from one of its focal lines as if they were coming from the other focal line. Thus, if the outer focal line of the ceiling's hyperbolic cross-section is located in the stage area at the height of the musical instruments, the reflected sound rays will seem to come from a point only slightly above the ceiling and will spread out accordingly with much more energy going laterally than for a plane ceiling.

More generally and less mathematically speaking, the ceiling should droop above the center aisle with a U-shaped cross-section to minimize median plane reflections.

Highly sound-diffusing ceilings

Another and possibly more effective solution would be to design a ceiling with extremely diffuse sound reflection, scattering a simple ray into ten or more "ray-lets" with roughly equal energies over a wide angular distribution. Surfaces with such highly diffuse reflections including experimental results were described previously¹⁵⁻¹⁸. These surfaces were designed to have local reflection coefficients for normally incident sound alternating between +1 and -1 according to a mathematical sequence (binary "maximum-length sequence", also called binary "pseudorandom noise"¹⁹). Reflection coefficients of -1 for hard surfaces are easily realized by a quarter-wavelength "wells" in the wall.

However, because of the quarter-wavelength requirement, the sound diffusing properties of the surface depend upon the wavelength of the incident sound. In practice it is found that good diffusion is obtained in a range of wavelengths half an octave below and above the "design wavelength". For an incident wave of half the design wavelength, the wells are, of course, half a wavelength deep, resulting in near specular reflection by the surface. (However, at one-third the design wavelength, the surface is a good diffuser again.)

Quadratic - residue diffusers

In the meantime, we have looked hard for surface structures that give excellent sound diffusion over larger wavelength ranges. In this search, we discovered, by computer simulation, that surfaces based on m -ary maximum-length sequences¹⁹ are capable of good diffusion over larger bandwidth - presumably because such surfaces have wells of several different depths. This led us to still another kind of mathematical sequences, the quadratic-residue sequences of elementary number theory, investigated by A.M. Legendre and C.F. Gauss. These sequences are defined as follows

$$s_n = n^2 \quad (\text{where } n^2 \text{ is the least non-negative remainder after division with } N, \text{ and } N \text{ is an odd prime.})$$

As an example, for $N = 17$, the sequence reads (starting with $n = 0$):

0, 1, 4, 9, 16, 8, 2, 15, 13, 13, 15, 2, 8, 16, 9, 4, 1; 0, 1, 4 etc.

These sequences are symmetric (around $n \equiv (N-1)/2$) and periodic with period N . In addition, they have the following astounding property: The discrete Fourier Transform R_m of the exponentiated sequence

$$r_n = e^{i2\pi s_n/N}$$

has constant magnitude. In fact,

$$|R_m|^2 = \left| \left(\frac{1}{N} \right) \sum_{n=1}^N r_n e^{-i2\pi nm/N} \right|^2 = 1/N.$$

Now suppose we construct a hard wall and give it local reflection coefficients for normal incidence according to the exponentiated quadratic-residue sequence r_n . What kind of sound reflection properties might it have? This is a complicated diffraction problem. Nevertheless, by making several, rather bold, simplifications, we can show, that such surfaces reflect sound very diffusely.

Figure 5 shows the result of a measurement on a scale model of the surface structure at a frequency of 11.5 kHz (corresponding to the design wavelength). Both the observed scattering angles and the energy fluxes are in good agreement with the theoretical values.

What are the scattering properties of the quadratic-residue surfaces at wavelengths other than the design wavelength? For longer wavelengths, the ceiling will look more and more like a smooth mirror, and we may not expect good scattering for frequencies half an octave or more below the frequency corresponding to λ_0 .

However, for wavelengths shorter than λ_0 the ceiling remains a good scatterer up to a limit given by prime number N : $\lambda > \lambda_0/N$ and, more stringently, the width w : $\lambda > 2w$.

The amount of flexibility inherent in the proposals made here should be sufficient to achieve the preponderance of lateral reflections necessary for the high listener preference found in concert halls with large binaural dissimilarity. A first application of quadratic-residue diffusors to the design of a large hall is described by Mashall and Hyde²⁰.

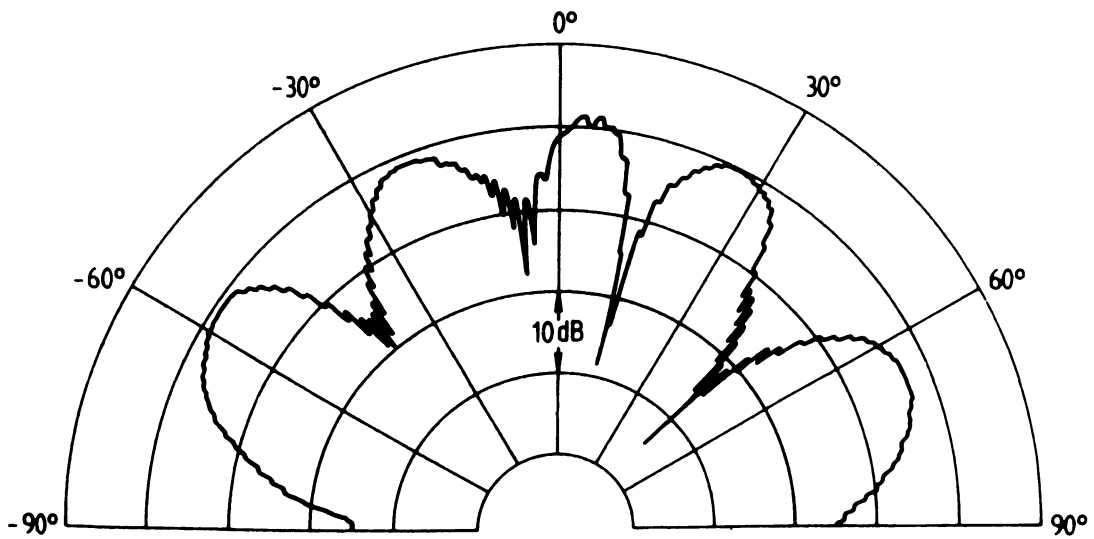


Fig. 5. Scatter diagram of a scale model diffusor for the design frequency.
(Schroeder, J. Acoust. Soc. Am. 65/4, 1979)

References:

- ¹ This definition of reverberation time was introduced by W.C. Sabine because in his early experiments, using a stopwatch and his ears to time reverberation, his hearing covered a dynamic range of one million to one in sound energy.
- ² Wente is particularly well known as the inventor of the condenser microphone. For his paper see E.C. Wente, J. Acoust. Soc. Amer. 7, 123 (1935).
- ³ B.S. Atal and M.R. Schroeder, "Perception of Coloration in Filteres Gaussian Noise - Short-Time Spectral Analysis by the Ear", Congr. Rept. Fourth Int. Congr. Acoust., Copenhagen, (1962), Vol. I, Paper H31.
- ⁴ This frequency resolution (ability to detect spectral irregularities in a passive transmission system) for speech or music should be distinguished from the frequency discrimination of steady tones. For pure tones, this frequency discrimination can attain 3 Hz at 1 kHz for trained subjects participating in paired comparison ("AB") tests.
- ⁵ M.R. Schroeder, "Improvement of Acoustic Feedback Stability in Public Address Systems", Congr. Rept. Third Int. Congr. Acoust., Stuttgart (1959), Vol. II, pp. 771-775.
- ⁶ M.R. Schroeder, J. Acoust. Soc. Amer. 45, 1077 (1969).
- ⁷ M.R. Schroeder, J. Acoust. Soc. Amer. 47, 424 (1970).
- ⁸ M.R. Schroeder and B.S. Atal, "Computer Simulation of Sound Transmission in Rooms", IEEE Int. Conv. Record (1963), Pt. 7, pp. 150-155.
- ⁹ U. Eysholdt, Doctoral Dissertation, University of Göttingen (1977).

- 10 U. Eysholdt, D. Gottlob, "Digital Synthesis of Sound Fields: A New Method to Study the Influence of Single Acoustical Parameters on Preference Judgements". Proc. 9th Intern. Congr. Acoustics, 4, (Madrid 1978) (Paper A2).
- 11 Y. Ando, U. Eysholdt, and D. Gottlob, "Subjective Preference Tests of Sound Fields with Multi-Echoes", p. 44 (Paper B8).
- 12 Y. Ando, "Subjective Preference in Relation to Objective Parameters of Music Sound Fields with a Single Echo", J. Acoust. Soc. Am. 62, 1436-1441 (1977).
- 13 D. Gottlob, A. Kohlrausch, M.R. Schroeder, "Einfluß raumakustischer Parameter auf die subjektive Beurteilung von Schallfeldern", Fortschritte der Akustik, DAGA '78, Bochum (1978).
- 14 M.R. Schroeder, "Computer Models for Concert Hall Acoustics", Am. J. Phys. 41/4, 461-71 (April 1973).
- 15 M.R. Schroeder, "Diffuse Sound Reflections by Maximum-Length Sequences", J. Acoust. Soc. Am. 57, 149-150.
- 16 M.R. Schroeder and R.E. Gerlach, Response to "Comments on 'Diffuse sound reflection by maximum-length sequences'", J. Acoust. Soc. Am. 60, 954 (1976).
- 17 R.E. Gerlach and M.R. Schroeder, "Die Anwendung von Maximalfolgendiffusoren in einem Modellhallraum," Fortschritte der Akustik, DAGA '76, Heidelberg, VDI-Verlag, Düsseldorf, 1976, 255-258.
- 18 W. v. Heesen, "Diffuse Schallreflexion durch räumliche Maximalfolgen", Fortschritte der Akustik, DAGA '76, Heidelberg (1976).

- 19 S.W. Golomb, Shift-Register Sequences (Holden-Day San Francisco, 1967).
- 20 A.H. Marshall and J.R. Hyde, "Evolution of a Concert Hall: Lateral Reflection and the Acoustical Design for Wellington Town Hall. J. Acoust. Soc. Am. 63, S36 (1978). See also Proc. 9th Intern. Congr. Acoustics, 46-47 (Madrid, 1978) (Papers B10 and B11).

Sound Examples

Drittes Physikalisches Institut, Universität Göttingen

The music examples in this record were originally devised for the reproduction system illustrated in Fig. 3. The two stereo channels are reproduced via specially designed filters by means of two equal high-quality loudspeakers in an acoustically highly damped (anechoic) room. The filters are adjusted for the case that the two loudspeakers are placed symmetrically in front of the listener. The distance listener loudspeaker should be 2.4 m and the distance between the loudspeakers 2.8 m. In the music examples of the record this filtering is included.

The best possible reproduction is obtained with a sound reproduction system according to the description above. However, a substantial part of the effects can be obtained by playing the examples over a common stereophonic equipment in a room with sufficiently low reverberation, such as a reasonably normal livingroom. The balancing of the two channels and the best listening position can be found out with the aid of the initial test examples.

Side A:

I. Test signals (Track I)

- a) Test signal for equalizing the two channels. The sound should be equally loud in both ears. (20")
- b) Test signal for selecting listening position. Find the position in the room where the signal is, if possible, only heard in the left ear (or as weak as possible in the right ear). (20")
- c) Same as b) but for the opposite ear. (20")

II. Demonstration of sound fidelity in the computer record. and reprod. system (Track II) (Analog-to-digital and digital-to-analog converters with a resolution of 12 bit and a sampling frequency of 23 kHz).

- a) Ordinary recording of an excerpt of J.S. Bach: Brandenburg Concerto No.3 (BWV 1048) (20")
- b) The same signal rerecorded and replayed by the computer (20")

III. Digital synthesis of a sound in a given listening position of a hall (Track III) (Convolution of reverberationfree music with impulse responses in original concert halls)

- a) Reverberationfree recording of excerpt of W.A Mozart, Symphony No. 41 C "Jupiter" K. 551 (15")
- b) An impulse response (5")
- c) Synthesis of the music in a listening position in a hall with short, medium and long reverberation time (15")

IV. The same reverberationfree music (excerpt from W.A. Mozart, Symphony No. 41 C "Jupiter K. 551) replayed and rerecorded in a listening position of 11 concert halls.

The preference for the listening positions of the 11 halls are given below.

Track IV	Hall 1	Preference:	Average	(46")
" V	" 2	"	Poor	"
" VI	" 3	"	Good	"

Side B:

IV. The same reverberationfree music replayed and rerecorded in a listening position of 11 different concert halls (cont.)

Track I	Hall 4	Preference:	Good	(46")
" II	" 5	"	Poor	"
" III	" 6	"	"	"
" IV	" 7	"	Good	"
" V	" 8	"	Average	"
" VI	" 9	"	"	"
" VII	" 10	"	Good	"
" VIII	" 11	"	Average	"

References

- 1) Eysholdt, U: Subjektive Untersuchungen an digitalen Nachbildungen von Schallfeldern aus Konzertsälen.
Dissertation Göttingen 1976
- 2) Eysholdt, U. & D. Gottlob: Synthese von Schallfeldern auf einem Digitalrechner.
in: Fortschritte der Akustik, p. 263-266, VDI-Verlag Düsseldorf 1976
- 3) Eysholdt, U. & D. Gottlob: Der Einfluss raumakustischer Parameter auf die subjektive Beurteilung von Schallfeldern
in: Fortschritte der Akustik, p. 251-254, VDI-Verlag Düsseldorf 1976
- 4) Eysholdt, U. & D. Gottlob: Digital synthesis of sound fields: a new method to study the influence of single acoustical parameter on preference judgements
ICA 9, Paper A2, Vol. I, p. 4, Madrid 1977.
- 5) Gottlob, D: Vergleich objektiver akustischer Parameter m. Ergebnissen subjektiver Untersuchungen in Konzertsälen, Dissertation Göttingen 1973

- 6) Siebrasse, K F: Vergleichende subjektive Untersuchungen zur Akustik von Konzertsälen, Dissertation, Göttingen 1973
- 7) Schroeder, M R., Gottlob, D. & Siebrasse, K F.: Comparative study of European concert halls, JASA (1974) 53, No 4, p. 1195.

(See also the reference list on pp. 27-29.)