THE EFFECTS OF EARLY REFLECTIONS
ON SUBJECTIVE ACOUSTICAL QUALITY
IN CONCERT HALLS

by

Michael F.E. Barron, B.A.

Thesis submitted for the degree of Doctor of Philosophy in the University of Southampton, Faculty of Engineering and Applied Science.

Institute of Sound and Vibration Research,
The University,
Southampton,
England.

January, 1974.

Frontispiece: Interior of Christchurch Town Hall, New Zealand.

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE
INSTITUTE OF SOUND AND VIBRATION RESEARCH

Doctor of Philosophy

THE EFFECTS OF EARLY REFLECTIONS ON SUBJECTIVE ACOUSTICAL QUALITY IN CONCERT HALLS

by Michael Francis Evan Barron

An investigation of current theories of subjective concert hall acoustics indicated that a multi-dimensional process is probably involved, though there is no general agreement over physical correlates of this subjective experience. Subjective experiments with simulated reflections in an anechoic chamber were conducted to gain an understanding of the subjective importance of early reflections in concert halls. The effect of 'spatial impression', which is produced by early lateral reflections, was identified as being the only desirable effect of early reflections, other than loudness effects. An extensive series of subjective experiments indicated that the physical quantity, which correlates with the subjective degree of spatial impression, is the ratio of lateral to non-lateral sound within 80 ms of the direct sound. The bass component of lateral reflections, below 400 Hz, was found to be substantially responsible for the effect of spatial impression. The evidence available also suggested that an interaural cross-correlation process is involved, and that an alternative measure for the degree of spatial impression is the height of the localisation maximum of the cross-correlation function, when only the early sound is considered.

Impulse measurements of the temporal and spatial distribution of sound in three Australasian concert halls are described. These measurements and a computer investigation, which assumed specular reflection at room surfaces, indicated that the degree of spatial impression is probably satisfactory and relatively uniform in many concert halls. Certain hall shapes, such as the fan-shape, were found to be undesirable; to design a large concert hall (about 3000 seats) with a uniform and satisfactory degree of spatial impression appears to be difficult.

Current literature suggests that spatial impression is a characteristic of the best concert halls, though its significance relative to other subjective effects remains to be determined.

ACKNOWLEDGEMENTS

Grateful thanks go to Professor P.E. Doak and Professor A.H. Marshall who both provided much assistance, encouragement and sustaining influence. I should like particularly to thank the long suffering subjects who performed the subjective experiments, especially Dr. George Dodd, who must hold the record of the highest number of sessions. Thanks are due to many others both in Southampton and Perth, Western Australia; among them Roy Caruth, who among other things managed admirably to keep temperamental apparatus in working condition. Finally I would like to thank Maureen Strickland for her most proficient typing of the text.

CONTENTS

PART II

PART I

Chapter 1

INTRODUCTION

A discussion some years ago about how human beings determine the direction of a source of sound was titled "Two ears - but one world" |1|. A more suitable title, without the implication of redundancy, might be: "Two ears - but three dimensions". The ability of our hearing system, with just two receptors, to determine direction in three dimensions, without having to rely on head movement, is certainly one of the most intriguing aspects of the human auditory system. Indeed, three functions for which we generally rely on our visual senses can also be performed by our hearing system: perception of direction, perception of distance and perception of the nature of an enclosed space. For these we rely heavily on binaural processes, which many blind people have learnt to perfect to enable them to orientate themselves.

Perception of direction, distance and space is fundamental to appreciation of the acoustics of concert halls. The physical acoustic attribute to be first associated with subjective appreciation of space was the reverberation time. It is obvious that attributes associated with reverberation time have been appreciated ever since certain forms of music were associated with particular types of building. Perhaps the most obvious example of this is the development of the Gregorian Chant, though the interaction between reverberation characteristics and style of music has been very strong all through the classical and romantic periods of musical composition. Perception of acoustic space is an essential aspect of listening to music in rooms, and one that is strongly reliant on binaural processes. The willingness of people to double or even quadruple their outlay on electronic equipment to attain the sensation of perceived space attests to this fact.

Historically, design of concert halls has been on the basis of precedent, with the inevitable occasional failures due to excessive extensions of known successes, or ignorance of acoustical behaviour. With the discovery at the turn of the century that reverberation time was not only a function of hall volume but also of the character of wall surfaces, the possibility of prediction and design for the most important acoustical characteristic of spaces has existed. However, it has since become apparent that other aspects of hall design also influence the subjective acoustic impression. Hall shape, the orchestra enclosure, and surfaces responsible for early reflections are generally considered important. However, architecture in this century has become so flexible that frequently no precedents exist. No longer is there an established style or a desire to conform to an established style, as existed in previous centuries. The desire for novelty, as much as reinterpretation of the visual function of the concert hall, has been responsible for many extreme divergences from the classical rectangular shape. This has in part been necessary due to the economic requirement to seat larger audiences, which can no longer satisfactorily be accommodated in the rectangular shape suitably scaled-up. The need for information concerning the physical acoustic requirements for good acoustics has thus become imperative if concert halls are to be built with characteristics that match the expectations. The solution of the problem, however, lies in subjective experiments, with their inherent inaccuracies, problems of definition of dimensions, etc.

The notion that acoustic experience in concert halls is not a one-dimensional situation (i.e., that reverberation time, or some variant of it, does not fully express subjective acoustic conditions in a hall) has only recently been confirmed (|2|, see section 2.6). It appears that between four and six subjective dimensions are required to express subjective impression in concert halls. The problem remains, however, of finding accepted words in which to express acoustic experience. Without an accepted vocabulary, experimental methods which make no prior assumptions become necessary with their inherent lack of precision.

It is therefore understandable that as yet no universally accepted quantity exists which correlates with subjective experience. Even the significance of reverberation time has been much criticised recently. Many other quantities have been proposed, some less substantiated than others. However, it is only recently that suitable technical facilities have been developed to make recordings in halls and suitable psychological techniques have been applied to help resolve the problem of what is significant in concert halls.

It is perhaps ironic that one hall with the best reputation in the world, the Musikvereinsaal in Vienna, was designed with no quantitative acoustic knowledge. One is tempted to cite the example of the Stradivarius violin and suggest that technology is destroying what should be the art of acoustics. This however is unjust criticism (see also an eloquent defence of the role of the acoustician in connection with a recent adaptation of an old malt house for use as a concert hall |3|). The narrow rectangular hall, as will be seen in later chapters, has, apart from other virtues, a particular uniformity in its acoustics which becomes increasingly more difficult to achieve for larger halls. Whilst the most valued violin was developed on the basis of pure empiricism, it is probable that more fortune than empiricism was responsible for the discovery of the virtues of the classical rectangular shape.

Although we may well be losing the romantic figure of the acoustical Merlin, whose ear can communicate directly with the drawing board, nevertheless the application of valid scientific techniques to the problem of subjective concert hall acoustics promises, some seventy years after the birth of physical architectural acoustics, to provide practical answers, which become more valuable the more radical are the deviations from conventional designs. The present state of the art is to a certain extent reflected in the present popularity of variable acoustic elements (particularly in the USA) which are tuned for best conditions. Both the best use of such elements, however, and the application of electronic systems (see, e.g., references |4, 5|) which are likely to gain popularity, depend ultimately on an understanding of the subjective requirements for good acoustics. One is tempted to ask how it is that a majority of halls have been successful in the absence of definite criteria. Yet it is frequently the failures which have led to the most conclusive evidence concerning the significant physical quantities (see, for example, section 2.1(b)).

One comparative failure in concert hall acoustic design has been the fan-shaped hall. This and other failures, which are not attributable to reverberation time, lead one to consider a possible relationship with the early sound. The significance of the direct sound and early reflections has in the past been much overshadowed by studies of the later arriving sound, but recently, as mentioned above, the subjective significance of reverberation time has been much criticised. One physical attribute of the early sound that has received mention by many people is the presence of early lateral reflections. In particular, Marshall |6, 7| has suggested that the presence of such reflections in classical rectangular halls is responsible for their excellent reputation, whereas such reflections are frequently not perceived in some modern halls, which includes fan-shaped halls. This thesis commences with an investigation of the subjective effects of early reflections when they are simulated in an anechoic chamber. The desirable spatial effect produced by early lateral reflections emerges as the only positive subjective effect of early reflections, other than loudness effects. This partial confirmation of Marshall's theory led to a thorough investigation of the physical requirements to produce this desirable spatial effect, which will be called "spatial impression" (S.I.). It was apparent that integrated energy in the early sound determined subjective response and that the ratio of lateral to non-lateral early sound correlates well with the subjective degree of spatial impression. Further subjective experiments indicated that the bass component of the lateral sound was particularly important for the creation of spatial impression, and answered to a certain extent the question of the relation between the subjective effects of reverberation and early lateral reflections. This subjective experimental work forms the body of Part I, which is concerned with the subjective aspects of the problem. This work confirms many qualitative comments made by other workers in the field, but provides quantitative data derived from subjective experiment. In particular, the notion that spatial impression is derived in the human auditory system by a cross-correlation process between the signals at the two ears, first proposed by Keet |8|, has been both substantially confirmed and extended.

In Part II the implications of the results of Part I for concert hall design are investigated. Both physical measurements in real halls and computer investigations of simple hall shapes were made; design procedure for spatial impression in concert halls is included in the penultimate chapter. It is concluded that spatial impression is potentially a very important requirement for excellent concert hall acoustics, but confirmation of the importance of spatial impression must await the publication of results of experiments being conducted elsewhere. Two considerations help explain the fact that only recently has the significance of early lateral reflections been widely accepted: firstly Marshall |6| suggests that the subjective effects of early lateral reflections have been traditionally attributed to reverberation, and secondly studies in Part II of this thesis suggest that, in particular in smaller halls, the degree of spatial impression is automatically satisfactory. Only recently with attempts to accommodate audiences of 3000 or more has the consideration of spatial impression become necessary.

To enable readers to experience for themselves the subjective effects described in this thesis, a recorded tape is included inside the rear cover of this thesis. This tape is suitable for playing on a stereophonic tape recorder (tape speed 19 cm/s) through headphones. Details of the contents of this tape are included in Appendix IV, as well as a summary being contained opposite the tape itself.

In the text, most chapters contain a final conclusions section. Readers unfamiliar with the content of this thesis may like to refer to this section before approaching the relevant chapter, to gain a preview of the arguments involved.

Chapter 2

A BRIEF REVIEW OF LITERATURE CONCERNING SUBJECTIVE CONCERT HALL ACOUSTICS

2.1 THE PHYSICAL SITUATION

(a) Early reflections.

If a pulse is emitted on the stage of a concert hall, a listener receives first the direct sound, then early reflections whose intensity gradually decreases whilst their density (number arriving per second) increases until a situation is reached (normally, in concert halls, about 100 ms after the direct sound) where direction of incidence is random, and the instantaneous level decays generally exponentially. The transition from early reflections to the final condition of reverberation is a gradual one; the earliest reflections are often discrete, becoming less precisely defined in time due to diffusion at reflecting surfaces. This transition is clearly demonstrated by "oscillograms" or "echograms" measured in halls. These are derived simply by using an explosive sound source (frequently from an electric spark) and displaying a microphone signal on an oscilloscope screen.

An extensive study of the statistics of early reflections was made at Göttingen during the period 1950-1960, principally by Meyer, Schodder, Thiele and Junius. This work is well summarised in reference |9|. For oscillograms measured in halls it is found that they differ widely for different receiving positions within a hall. For positions in the stalls, near the source, the direct sound predominates, whilst towards the back of a hall, the direct sound is considerably less prominent and the reflection density becomes high even in the first 50 ms after the direct sound.

Statistical analysis of reflections was performed at Göttingen in terms of their amplitude relative to the direct sound. Within the first 100 ms the number of reflections deviates substantially from the statistically calculated mean number, N (given in equation (15.1)). It has not proved possible to relate the results of such a statistical analysis to subjective impressions in different halls; one can say only that it is perhaps subjectively significant that within a given hall the number of reflections is always larger in the circle than in the stalls.

This study raises the question of whether a diffuse reflection (which, because its energy is spread temporally, would not be registered in this statistical analysis) is equivalent subjectively to a discrete reflection of equal total energy. This equivalence does appear to exist, at least in certain situations: at threshold (Seraphim |10| and Kuhl |11|), and for the subjective effect of lateral reflections (see section 10.3). This investigation may be further criticised since a pulse response contains no frequency information, and the results are principally relevant to frequencies of maximum intensity in the pulse (around 3 kHz) which seem unlikely to be the most important in concert halls.

Thiele proposed a measure related to energy received in different time periods after the direct sound; this will be further discussed in section 2.4 below. Oscillograms are still frequently made in halls since they provide information about early reflections and are readily measured. If energy arriving within certain time periods is important, however, they are difficult to interpret visually. Cremer has proposed the 'Christmas tree criterion': that for good acoustics the sequence of reflections should be more or less regular and of gradually decreasing amplitude. This is certainly relevant for the reverberation (say after 100 ms) to avoid echoes, but no evidence is available regarding what constitutes an undesirable deviation from the criterion for early sound. Oscillograms also carry no information regarding the direction of reflections.

Thiele and others have made measurements at high frequencies of the directional distribution in concert halls. It was found that the directional diffusivity in halls decreased with increasing volume. The measuring method is distinctly laborious, however, and the value of the results has been questioned by Reichardt |12| and others who suggest that whilst a high degree of diffusion is generally considered desirable for late reverberant sound, different conditions may apply to the early sound. The work in this thesis partly answers the question of what directional characteristics are desirable for the early sound. Gilford |13| cites various correlation techniques which may be more valuable for determining the degree of diffusion for the later reverberant sound, though such techniques appear limited in their usefulness to large rooms.

(b) Audience attenuation at grazing incidence.

Only in 1964 was it discovered that sound passing over audience seating suffers severe attenuation (15-20 dB) at frequencies around 180 Hz. Measurements of this have been reported by Schultz and Watters |14| and Sessler and West |15|; the typical attenuation in excess of spherical divergence is shown in Figure 2.1. It was found that this excess attenuation is established after the sound had crossed only six seating rows and that beyond that it is independent of position. The attenuation also occurs for lateral reflections, as can be seen in Figure 2.1. The same response occurs with and without an audience and is not dependent on the absorption of the seating.

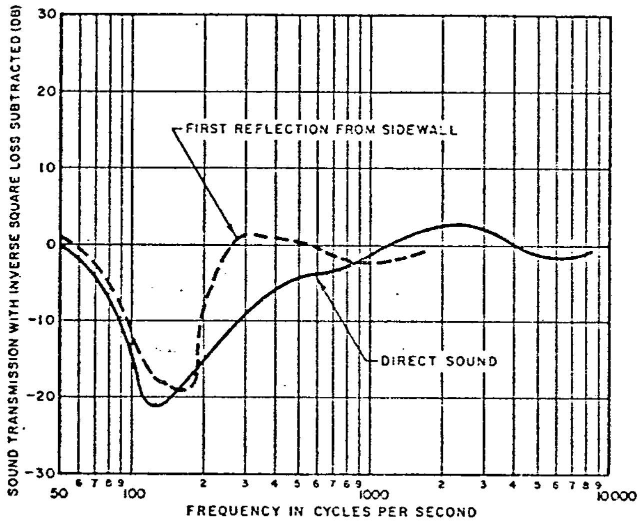


Figure 2.1. Typical transmission characteristic of direct sound and the first reflection on the main floor of a concert hall (after Schultz and Watters |14|.).

The elevation of the sound has to be 15° relative to the seating according to West and Sessler |16| and at least 30° according to Schultz and Watters for the selective low-frequency attenuation to be eliminated. Both these figures were derived from model studies; a 15° floor rake as a means of overcoming the attenuation effect is a possible solution if this figure is correct. Reflections off the ceiling will, in general, therefore not suffer audience attenuation.

The physical mechanism responsible for this attenuation appears to be a vertical resonance in the cavities between consecutive seating rows. Sessler and West |15| suggest that a diffraction effect is also responsible for the broad-band absorption. Reports that high frequency sound is attenuated as it passes over an audience (see, e.g., reference |17|) are frequently found but no quantitative measurements are available in the literature, other than those by Meyer, Kuttruff and Schultz |18|, which were found by Kunstmann |17| to be results based on a special case of symmetry. Schultz and Watters found no excess attenuation at high frequencies due to the seats alone, and speculate that it would not occur with an audience if line of sight is maintained. Furduev |19| cites measurements on open-air seating which support this criterion.

The subjective significance of audience attenuation at grazing incidence will be discussed in Chapters 11, 12 and 23.

2.2 REVERBERATION TIME AND SOUND DECAY

It is not necessary here to record the history of the concept of reverberation time and its measurement. It is now however generally accepted that precision to the nearest 1/20th of a second for reverberation time (R.T.) and examination of the fine structure with pure tones |20| are of little subjective significance. Recently the significance of R.T. measured between -5 dB and -35 dB of the stationary level has been questioned, as well as the traditional method of measurement. Schroeder |21| has demonstrated that the average sound intensity during decay (as measured when noise decays are averaged) equals the squared impulse response from time t to infinity. When this "integrated impulse method" is used replication of R.T. measurements is no longer necessary, and the details, particularly of the initial decay, are readily available. Schroeder has further shown that build-up and decay are purely complementary with such integrated impulse responses |22|.

Atal, Schroeder and Sessler |23| asked subjects to subjectively compare linear (exponential) with non-linear decays, and discovered that for simulated decays there was good correlation between the decay rates over the first 160 ms. However for real decays recorded in halls, measurement of the decay rate over the first 15 dB correlated fairly well with reverberance of linear decays. Marshall |6| has also questioned the significance of the reverberant decay as such, which is rarely perceived in musical performance.

The subjective contribution of reverberation has been well summarised by Parkin and Morgan |4|: that it bridges the gap between notes of music, that unlike the direct sound it arrives from all directions, including from behind, which is important, and also that it contributes to the total loudness of music.

2.3 THE HAAS EFFECT AND SPEECH INTELLIGIBILITY

Quantitative consideration of the subjective contribution of early reflections dates from the work of Haas in 1951 |24|. The Haas effect can be summarised as follows: for reflections in or near the horizontal plane one will localise on the direct sound (precedence effect), even if, for short delays, the reflection level somewhat exceeds that of the direct sound. Haas made measurements of echo disturbance with speech, which were extended by Bolt and Doak |25| to provide disturbance contours, which have essentially been confirmed |26|.

The contribution of early reflections to speech intelligibility has been investigated by Lochner and Burger |27, 28|, and both the latter and Niese |29| have proposed measures based on the ratio of early (useful) energy to the late (disturbing) energy. With speech one has a unique criterion, speech intelligibility, which is readily quantifiable. Music differs from speech in being generally less impulsive, but comparison has frequently been made between clarity for music and intelligibility for speech; in both cases the ratio of early to late energy has been considered.

2.4 MEASURES OF EARLY TO REVERBERANT ENERGY

Thiele (see again reference |9|) proposed the 50 ms energy fraction as a significant measure for subjective appreciation. Correlation with speech intelligibility appears to be quite good |30, 29|, at least in the absence of late disturbing echoes. Beranek and Schultz |31| have also found this measure very significant in halls, and found that the acceptable range for the ratio of reverberant to early (50 ms) energy is less than 10 dB. They also suggest that the ratio is substantially constant throughout halls. Schmidt |32|, using simulated direct sound and reverberation, found that the "Raumlichkeit" (or room impression) depended not only on the ratio of reverberant to direct sound but also on the reverberation time. On the basis of this result, the 50 ms energy fraction is not a unique measure, though comparison with the predicted value for an exponential decay would still appear to be valid. Reichardt |12, 33| has also attempted to include early sound in a measure corresponding with "Raumlichkeit", but the argument is not well substantiated by experiment (see section 12.2).

Comparison of measured values of a physical quantity and those predicted for an exponential decay has been used by Jordan |34| as a criterion for subjective suitability. Jordan has proposed the rise time, steepness, and most recently the early decay time as being significant. Schroeder |22| has shown that all these measures may be derived from the integrated impulse decay, and Jordan's early decay time differs from Schroeder's measure of the initial reverberation time only to the extent that Schroeder considers the first 15 dB of the decay, whilst Jordan considers the first 10 dB. Jordan proposes that the early decay time (measured as a reverberation time) should not be smaller than the average R.T. (measured in the conventional manner).

Kürer |30| has investigated the use and validity of a measure proposed by Cremer, called "Schwerpunktzeit" or "centre time", defined as follows:

$$t_s = \frac{\int_o^\infty (t - t_o)p^2(t).dt}{\int_o^\infty p^2(t).dt} ,$$

where t_o is the arrival time of the direct sound. This measure has the advantage that no discrete time interval is involved. There is a good correlation between this measure and speech intelligibility, and it also appears to be linearly related to Reichardt's "Raumlichkeit", but equal degrees of "Raumlichkeit" do not correspond with identical values of t_s as the reverberation time is varied.

Reichardt, Jordan and Kürer were all seeking a single number to express the acoustical sensation in a hall. At present there is insufficient evidence to consider any one as being more valid than another. Furthermore, musical acoustical experience has been seen to be multidimensional, so it would be surprising indeed if only one measure could express the experience in all situations. One particular assumption should be questioned: namely, that the sensation of "Raumlichkeit" (room impression, a sense of being surrounded by reverberant sound) is simply the inverse of the sense of clarity. Reichardt and Schmidt's results |35, 32| have been concerned with room impression in a simulation environment, but have been applied to consideration of clarity and definition by Reichardt himself and Kürer. In the author's experience spatial effects can be readily appreciated in a simulation environment, whilst in a concert hall the definition, both of the whole, and between instruments, is the most significant factor. The work of

Hawkes and Douglas |2| would suggest that the two are not the same. As yet only one subjective measurement of definition in a hall has been made (by Macfadyen |36|), but no physical measures were included.

Of particular relevance to the work reported here is that consideration of early lateral reflections offers the possibility of increasing both spatial effect and clarity.

2.5 THE SPATIAL EFFECTS OF LATERAL EARLY REFLECTIONS

An early mention (though only a casual one) of the effects of early lateral reflections was made by Meyer and Schodder |37| in 1952: "... the presence of a second loudspeaker creates an apparent enlargement of the spatial extent of the primary source and with a delay of some 10 ms also a certain 'reverberance'" (author's translation).

Only in 1966 does further comment about such reflections reappear |6, 38, 39, 40| and since that time other authors have also quoted the desirable qualities of lateral reflections |12, 34|. Marshall |7| and Reichardt |12| also claim that the presence of bass frequencies in lateral reflections is also critical.

Marshall |6| proposed that there is a desirable subjective quality present when music is played in classical rectangular halls, which is not present in fan-shaped, low-ceiling halls. He traced this quality, termed "spatial responsiveness", to the presence of unmasked lateral reflections in classical halls, which he claimed would be masked in the case of low-ceiling halls. A criterion was established based on the relative delays of lateral and ceiling reflections, which for rectangular halls leads to consideration of the cross-section ratio (the ratio of hall width to height), good halls having a small cross-section ratio. West |40| claims to have obtained a good correlation between cross-section ratio and overall acoustic quality in halls.

With the availability of threshold data for reflections with music |38|, and details of audience attenuation at grazing incidence (see above), Marshall |7| stresses the necessity for bass lateral sound to produce a sense of "warmth" of tone. He proposes that bass lateral sound can be derived from high lateral reflecting surfaces.

Keet |8| suggested that the degree of the spatial effect was related to the short-time cross-correlation between signals at the two ears. He also discovered that the degree of the effect was a function of music level.

Apart from the work of Keet, no study has been made of the conditions necessary for this effect of lateral reflections to occur. What is the significance of angle of elevation and azimuth of the reflection? Does the reflection delay have to be less than the ceiling reflection delay, as Marshall suggests, or not? Indeed, Marshall was obliged to extrapolate from thresholds of a side reflection in the presence of direct sound alone to the case of a side reflection in the presence of direct sound plus a ceiling reflection. Is this extrapolation valid? Before describing experiments aimed at answering such questions, it is appropriate to consider various techniques which have been used to deal with the multi-dimensional subjective situation in a concert hall.

2.6 THE MULTI-DIMENSIONAL APPROACH TO CONCERT HALL ACOUSTICS

Beranek |41| conducted a study of 54 halls throughout the world and on the basis of interviews with conductors and musicians was able to rate them on a scale from A+ to C+ and 'poor'. He also established a list of 18 subjective qualities important for listening (which included intimacy, liveness, warmth, clarity and brilliance). Each of these were related to a physical measure and by comparison of these physical measures with the subjective ratings, weightings for each of these physical measures as a contribution to perceived acoustical quality were derived. In particular, 'intimacy', which according to Beranek is related to the initial time delay gap, accounts for 40% of the total rating.

This work has been criticised by Hawkes and Douglas |2| who question two implicit assumptions of Beranek's work: firstly, that the positive attributes used are independent and linearly additive (for example, this would imply that a hall deficient in one feature can be compensated by excellence in another); secondly, that there is a linear relationship between the physical variable and the hall's rating on a good/bad scale. To eliminate these assumptions, a multi-dimensional assessment was undertaken in the Royal Festival Hall and other halls and the results analysed by factor analysis. They found that subjective assessment is made on the basis of between four and six subjective dimensions, rather than one dimension which has often been assumed. They discovered that "reverberance" seemed to be related to R.T., "evenness" to distance from and alignment with the orchestra, and "intimacy" to proximity to the orchestra, to a short initial time delay gap and to a high narrow cross-section. The significant dimensions do vary with the type of music and the concert hall.

However, the possibility of deriving precise relevant physical quantities from such studies is remote. In any case the precision of the method is limited to the extent to which different subjects interpret the subjective scales identically.

To evade the problem of expressing acoustical experience in words, work is now being conducted at Göttingen (as yet unpublished) in which subjects are required to rate recordings according to similarity and dissimilarity and preference. This has become possible through the development of an artificial head, both in Berlin |42| and Göttingen |43|. Recordings in different halls are replayed to subjects either through earphones or a pair of compensated loudspeakers. In this way it is hoped to determine the most significant physical requirements in concert halls. Whilst this approach is limited to existing halls and their accidental admixture of parameters, the possibility that the situation in a concert hall may be precisely simulated by using a computer should enable the required physical parameters relating to subjective appreciation to be found.

The methods used for the study reported here are more modest, though inasmuch as the physical situations employed elicited subjective responses, they are felt to be relevant to the real situation which these situations were approximating.

Chapter 3

THE EXPERIMENTAL SIMULATION SYSTEMS

3.1 INTRODUCTION

The acoustical situation for a listener in a concert hall can be simulated in the following manner |44|: the direct sound is radiated from in front of the subject in an anechoic environment, early reflections are radiated from the relevant directions by loudspeakers fed with signals from a delay machine and reverberation, derived either from a reverberation chamber or a reverberation plate, is fed to generally four loudspeakers arranged symmetrically about the subject. The limitation of this method is that there is no gradual transition from discrete reflections to diffuse (temporally irregular) reflections to statistically defined reverberation. A further criticism is that without duplication of loudspeakers and the use of a stereo signal, the finite width and resulting arrival time differences for different sections of the orchestra are not reproduced. Experiments reported in this thesis were concerned with investigation of the simplest situations of direct sound plus a few reflections; thus divergence between the simulation and reality could be estimated at each stage.

Experiments were conducted in two locations: Southampton (S) and Perth, Western Australia (P). The apparatus used was different in each case, but the systems were otherwise equivalent.

3.2 ANECHOIC FACILITIES

At Southampton, the large anechoic chamber, with internal measurements 9.15 x 9.15 x 7.32 m, was used. The cut-off frequency for this room is about 70 Hz. Subject-loudspeaker distances of 3m were used, though "non-elevated" loudspeakers were here at an elevation of -16° (a typical figure for direct sound, etc., in a hall with raked seating). The subject was seated on a chair; though he was asked to face forwards, there were no physical restrictions on head movement. Experiments were performed in subdued lighting conditions to aid concentration.

The anechoic facility in Perth was more modest, with internal dimensions 3.7 x 3.0 x 2.7m. Subjectively no intrusion from the exterior existed; the internal wall treatment consisted of a 15 cm polyurethane foam lining with, for the majority of the wall surfaces, additional triangular wedges 15 cm high glued on the plane foam surface. Minimum subject-loudspeaker distances of 1.2m had to be used for the subject at the centre of a loudspeaker array.

It is difficult to assess whether the limitations of the Perth facility influenced results; certainly no subjective situations arose which could be directly associated with its limitations, though this is not surprising. Threshold measurements and measurements of the effect of lateral reflection delay, made in both locations, were in good agreement with one another.

3.3 SIMULATION APPARATUS

Figure 3.1 is a diagram of the test set-up for a typical comparison experiment. Music motifs, recorded on tape loops, were played on a conventional tape recorder (Ferrograph Series 7(S), Nagra IV (P)). The output signal was fed into a tape delay machine, the two machines used will be described below. Output signals from the tape delay machine then passed

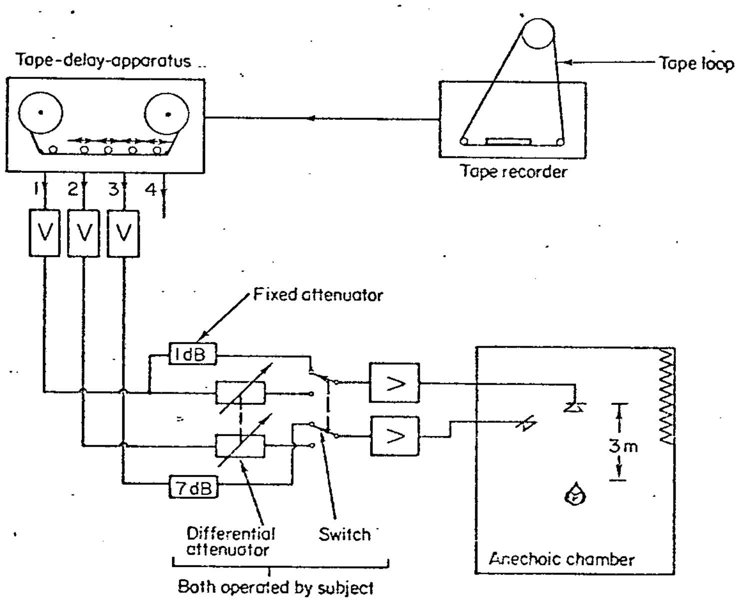


Figure 3.1. Typical test set-up (as used for the experiment described in section 7.4).

through impedance matching devices, filters, etc., to attenuators and switches in the anechoic chamber for the subject to control. After further amplification the signals were played in the anechoic chamber through loudspeakers. Further details are included below.

(a) Tape delay machines.

At Southampton a specially made Leevers-Rich system was used, which operated with 12.7 mm (½ in) wide tape running from reel-to-reel at 76 cm/s tape speed. It produced four outputs which could be delayed at will relative to one another.

In Perth two Philips delay machines (Type 6911/2) were used though considerably modified, as illustrated in Figure 3.2. To provide continuously variable delays, the tape was passed round two movable pulleys between the two record heads and their relevant replay heads. The considerable drag in the system as a whole required that durable but smooth tape be used (6.3 mm (¼ in) wide Ampex 434) and to minimise wow-and-flutter the movable pulleys were replaced with high inertia pulleys. Tape loops running at a tape speed of 76 cm/s could run for about ½ hour without serious deterioration. The electronic circuitry was rewired to eliminate interchannel crosstalk, and four additional replay channels per machine were also added. With two machines a maximum of 16 reflections with different delays could be simulated. The performance of these two tape delay systems is summarised in Table 3.1, in each case signal-to-noise-ratio was subjectively quite satisfactory.

Table 3.1

Electronic performance of tape delay machines.

Machine	Frequency Response	Wow-and-Flutter (weighted)
Leevers-Rich (S)	90 Hz - 20 kHz ± 1.5 dB	0.2%
Modified Philips (P)	60 Hz - 12 kHz ± 2.0 dB	0.1%

(b) Auxiliary apparatus and differential attenuator.

Impedance matching becomes important if a system is to be flexible; where possible high input and low output impedances were used. Various filters were introduced (one of which is described in Appendix II). Impedance dropping amplifiers were also included.

All comparison experiments were conducted at constant loudness; this was achieved by a differential attenuator which, for the case of direct

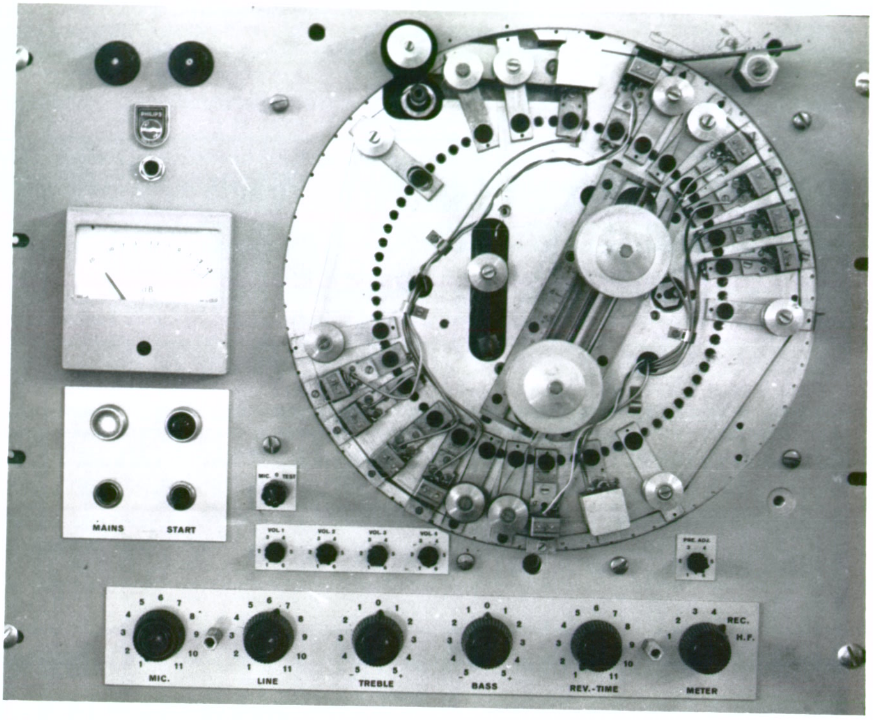


Figure 3.2. Philips tape-delay machine (modified)

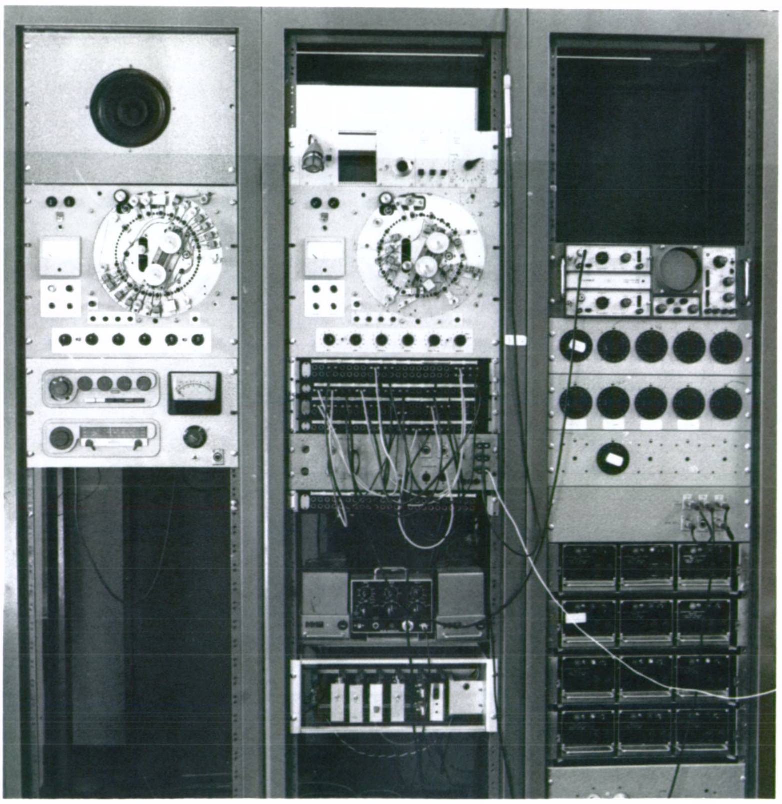


Figure 3.3. Simulation apparatus

sound and a single reflection, attenuates the direct sound for increasing reflection level such that the incoherent sum of their energies is constant. The design of the differential attenuator is outlined in Appendix I. The simulation apparatus in Perth is illustrated in Figure 3.3, and the block diagram in Figure 3.4.

(c) Loudspeakers.

At Southampton, Quad electrostatic loudspeakers were used, which proved particularly suitable due to their good frequency response and impulse response. Smaller loudspeakers were required in Perth; after subjective comparison between many commercially available small loudspeakers, Celestion Ditton 10's were chosen as being particularly free of colouration. Their small size permitted minimum relative angles of azimuth of less than 10°. To improve the bass response the original input to the delay machines was passed through a bass-boost filter. Figure 3.5 shows the frequency responses for a Quad loudspeaker and a Ditton 10 (including bass-boost).

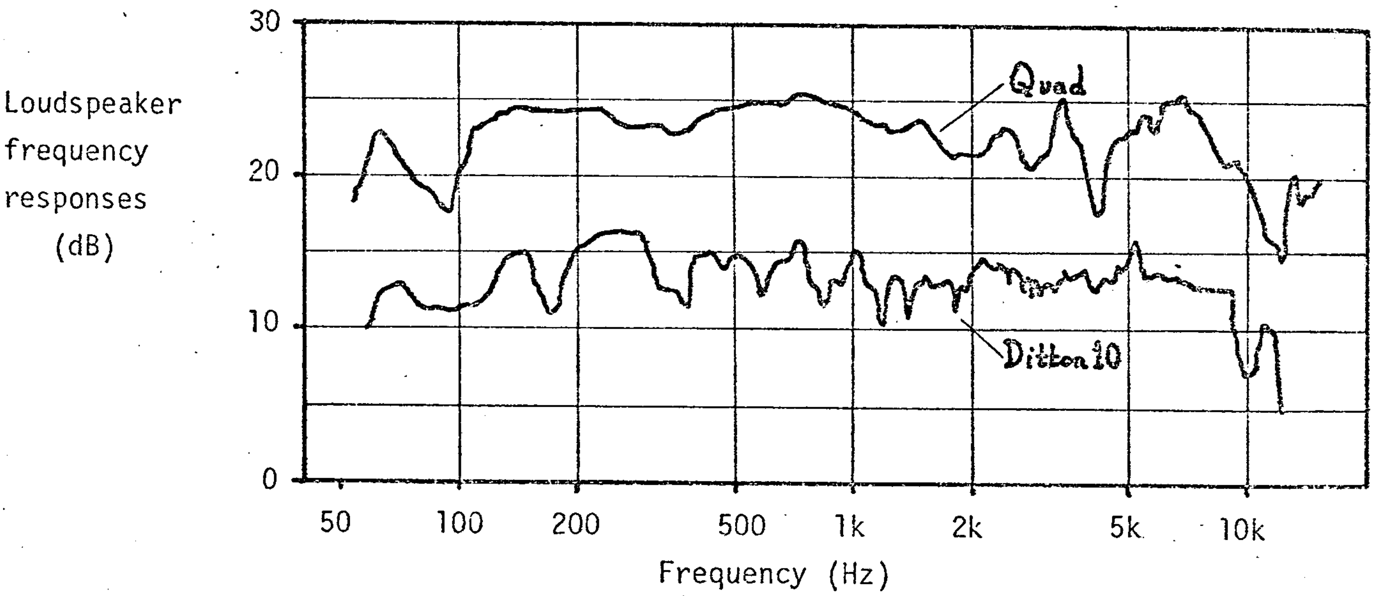


Figure 3.5. Frequency response of a Quad electrostatic loudspeaker (S) and a Celestion Ditton 10 loudspeaker, including bass-boost (P).

(d) Reverberation.

In both locations reverberation was derived from a reverberation plate (EMT 140 St). To produce a diffuse reverberant field, four loudspeakers were used placed symmetrically about the head of the subject. Opposite loudspeakers were wired in series and placed equidistant from the subject. In Perth, two incoherent outputs from the reverberation plate were available,

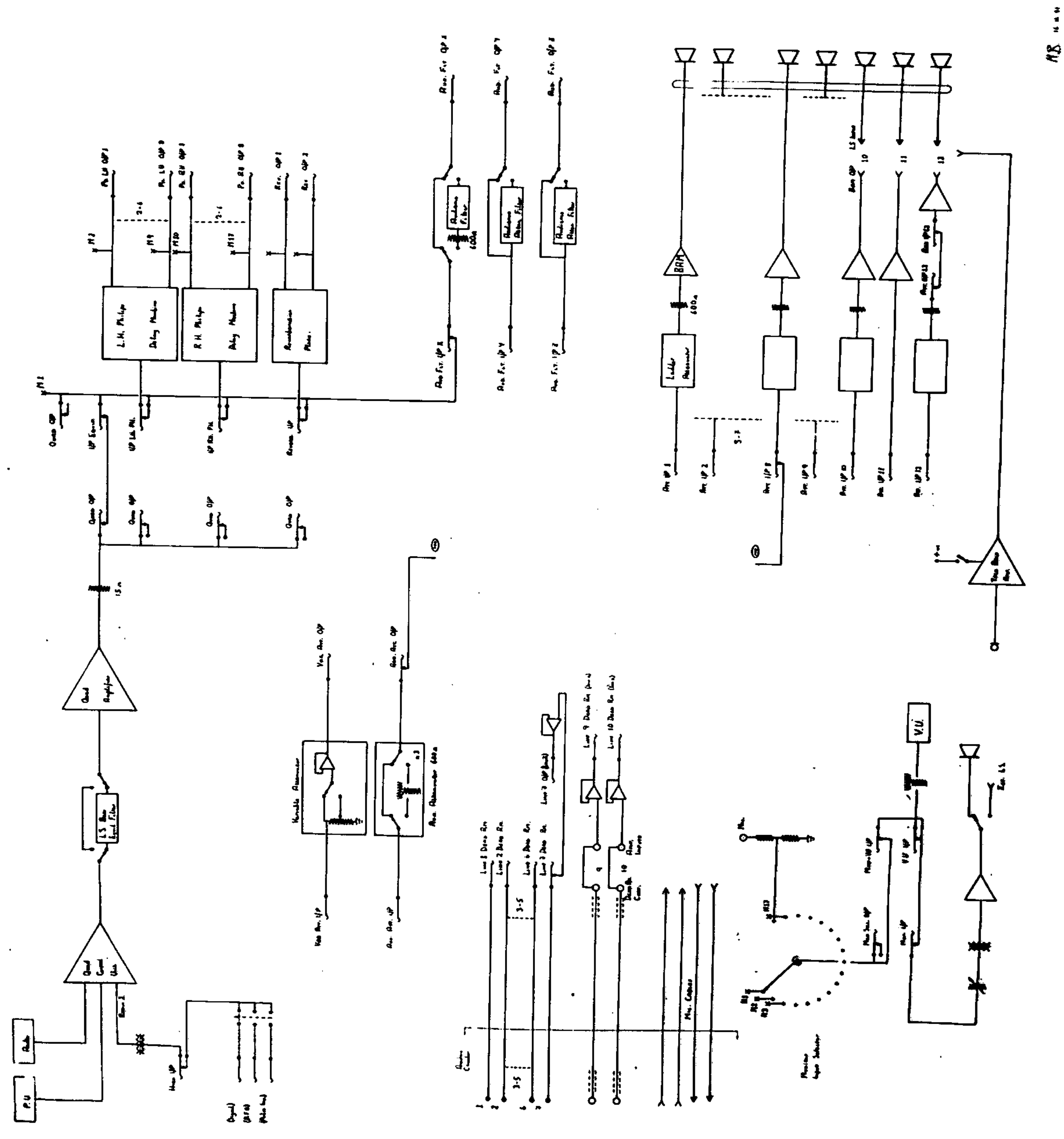


Figure 3.4. Block diagram of simulation system

whilst at Southampton only one reverberation signal was played to all four loudspeakers, though one pair, opposite one another, was placed about 40 cm further from the subject's head than the other pair to introduce a time delay difference for temporal incoherence |44|. Reverberation signals were delayed relative to the direct sound such that their onset came after the discrete reflections (between 50-100 ms after the direct sound). Conventional paper-cone loudspeakers were used to radiate reverberation.

For the experiments described in Chapter 12 and for reverberation on the tape included with this thesis a mean R.T. of 1.9 sec was used; the mean R.T. (for two output channels) measured by the integrated impulse method |21| is given in Figure 3.6.

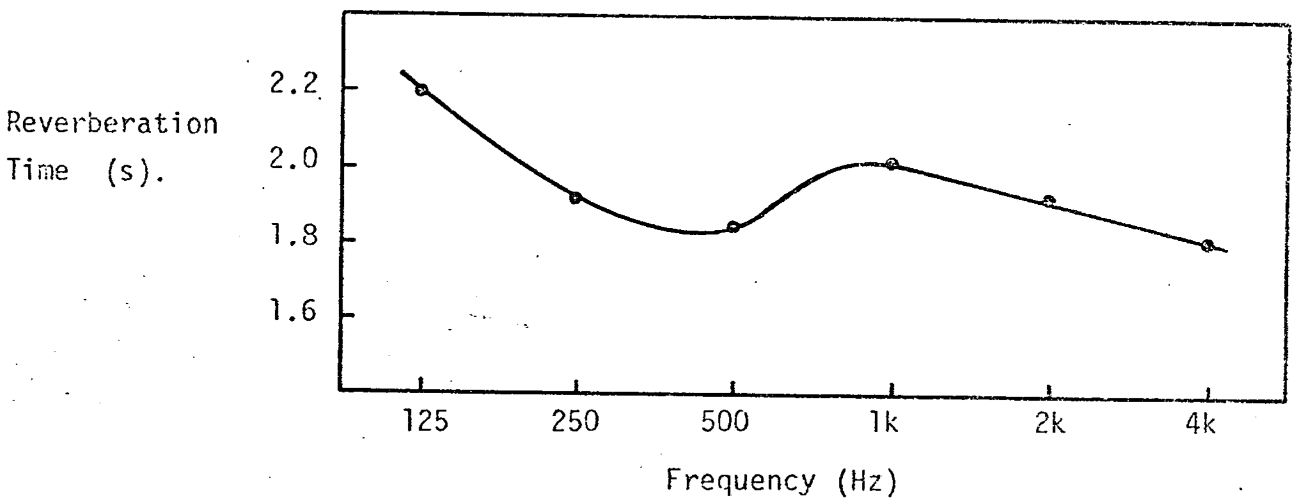


Figure 3.6. Mean reverberation time for signals from reverberation plate used in experiments reported in Chapter 12 and on the tape included with this thesis.

Readers can judge the quality of the Perth simulation system for themselves, since nearly all elements of it were used for the recording of the tape contained in this thesis.

3.4 SETTING-UP PROCEDURE

To obtain the correct delays, pure tones were fed into the tape-delay apparatus at Southampton and the phases of the output signals from it were compared at different pure tone frequencies (loudspeaker-subject distances were constant here). In Perth a microphone was placed in the position of the subject's head and with an impulse fed into the system the delay was adjusted using an oscilloscope. The accuracy of stated delays was better than 0.5 ms.

Before each experiment the level for each reflection at a microphone in the position of the subject's head was adjusted. At Southampton a pure tone at 1 kHz was used since loudspeaker responses were substantially flat in this region. This was not the situation in Perth where a warble tone, of frequency 800 Hz $\pm$ 250 Hz, was employed. The stated level can be considered as being accurate to within 0.5 dB.

The stationary level of reverberation channels was adjusted by using low-pass filtered white noise (cut-off frequency 5 kHz). The level was adjusted relative to the level for the direct sound with the same signal.

The mean level at which music motifs were played to subjects will be referred to as E dB relative to 0.0002 dyne/cm^2. Setting-up accuracy was within about 1 dB. The angle of azimuth will be represented by the symbol α, measured in degree anti-clockwise from straight ahead. The angle of elevation will be represented by β, measured in degrees above straight ahead.

3.5 MUSIC MOTIFS AND SUBJECTS

The motif used for the majority of the experiments was a 47 second section of the 4th movement of Mozart's Jupiter Symphony (No. 41), bars 94-151, recorded by the English Chamber Orchestra in the Building Research Station anechoic chamber. The section is a typical example of fast classical music, it has a wide dynamic range (about 20 dB) and contains most instruments of the orchestra. As an example of slow classical music (also used for the thesis tape) a 19 second section from Wagner's "Siegfried Idyll" (with a dynamic range of about 10 dB) was also used. In each case the sound level E was set up to be the most natural subjectively.

It was readily apparent that the most suitable subjects were critical listeners to music. Subjects were rejected if they did not give a sufficiently low threshold result for a single reflection. A degree of training was also found necessary; subjects did at least two experimental sessions before their results were used.

Chapter 4

THE SUBJECTIVE EFFECTS OF FIRST REFLECTIONS

4.1 INTRODUCTION

Haas |24| established that early reflections made a desirable contribution to the perception of speech in a room. This was a contribution in terms of intelligibility: reflections with delays up to 50 ms increase the effective loudness of the useful sound for intelligibility, |27|. The limiting time delay may be longer for music, the contribution for music being in terms of clarity. These effects will be called "loudness effects", with the implication that increasing the level of the direct sound is equivalent to adding such reflections to the direct sound. The subject of this chapter is the investigation of other subjective effects of early reflections. Since the effects also depend on reflection direction, they will be discussed under broad headings as the effects of lateral reflections and of frontal or ceiling reflections.

Schubert |38| included a table indicating the subjective effects of early reflections as a function of delay, which subjects reported using to determine the threshold. This table is duplicated below, Table 4.1.

Table 4.1

Threshold determining effects of a single reflection in the presence of direct sound (after Schubert).

Delay (ms)	Frontal Reflection	Lateral Reflection
0	Loudness	Apparent image size or
5	Tone colouration	image shift
10 - 20		Tone colouration
30 - 60		Room impression
> 80	Echo disturbance	Echo disturbance

To derive a two-dimensional version of Schubert's table, subjects were presented with different delay reflections and were able to adjust the differential attenuator for themselves in 1 dB steps (i.e., loudness changes were eliminated). First the general qualitative effects were established through consultation between subjects. Then the subjects were asked to determine for themselves the maximum and minimum delays and levels at which the various effects occurred. Because the transition from one effect to another is a gradual one, the results are not more accurate than ± 3dB (though the threshold curve and curve of equal spatial impression are exceptions to this statement, since a different technique was used for their determination). The Mozart motif was used at a mean level of E = 77 dB. Results shown are the average of two subjects.

Figure 4.1 summarises the effects observed for a lateral reflection at azimuth $\alpha = 40^o$. Each section will be discussed below.

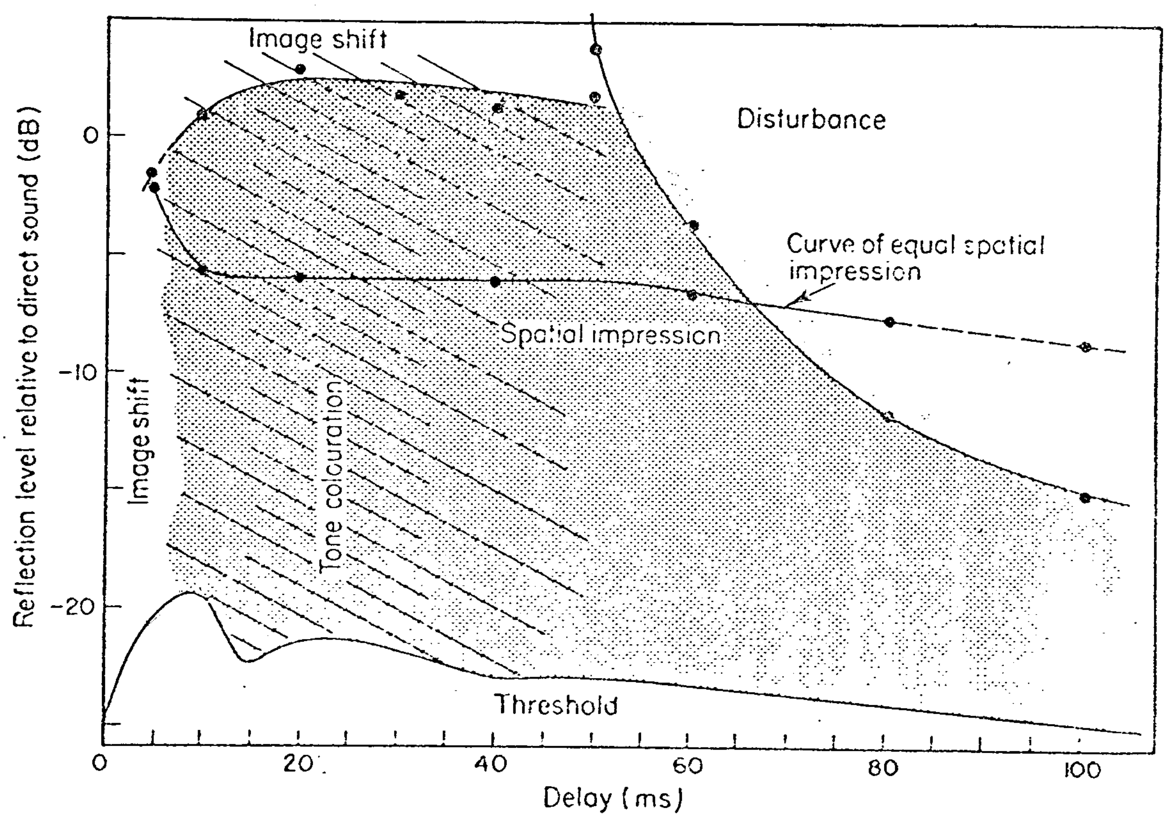


Figure 4.1. Subjective effects of a single side reflection ($\alpha = 40^o$) of variable delay and level with music.

4.2 SUBJECTIVE EFFECTS OF A LATERAL REFLECTION

(a) Threshold.

As a base line for the effects of a side reflection, the threshold from Figure 5.1 is included in Figure 4.1; reflections below threshold produce no audible effect. The threshold line and all other solid lines in Figure 4.1 represent lines of equal subjective impression.

(b) Localisation effects.

The precedence effect |24| normally operates with music and a reflection. However in the extreme situation of a small delay (< 5 ms) or a high level reflection with delay less than 50 ms, the apparent source moved from the direct sound loudspeaker towards the reflection loudspeaker. The effect is very similar to that observed when the balance control of a stereo system is adjusted. This movement of the point of localisation is indicated in Figure 4.1 by the unshaded areas marked image shift.

As will be discussed in Chapter 13, localisation in the horizontal plane for frequencies below 1500 Hz can be considered as a correlation process between the signals at the two ears. Section 13.8(f) contains an explanation of image shift for intense lateral reflections, similar arguments can also be used to explain image shift for short delays. At high frequencies localisation is due to interaural intensity differences; such image shifts can also be explained.

(c) Tone colouration.

For certain delay reflections (from about 10-50 ms, but especially around 20 ms), the tone of the music appeared to sharpen, especially the violin tone. The degree of colouration, however, appeared to be relatively independent of level for reflections more than 10 dB above threshold. This tone colouration is indicated in Figure 4.1 by diagonal-line shading, the density of the shading roughly corresponding to the degree of colouration.

Interference between a signal and its repetition is called the "comb filter effect". Figure 4.2 shows the frequency response of a signal plus a 20 ms delayed reflection. (The signal from a beat-frequency oscillator was fed into the experimental arrangement in place of the music signal. For the sound from the two loudspeakers a conventional loudspeaker response was taken with a microphone placed in the position of the subject's head.) The response in Figure 4.2 is for the signal and reflection at equal sound level. It may be expressed mathematically as $F(\omega) = 2|\cos \omega\tau/2|$, where τ is the reflection delay.

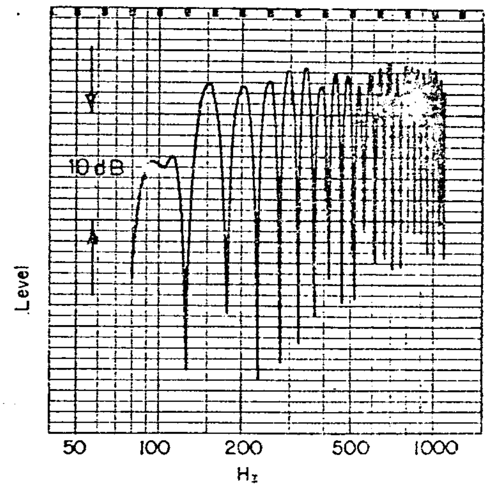


Figure 4.2. Frequency response of a signal plus 20 ms delayed reflection.

Müller |45|, among others, also reports tone colouration: "Short delay reflections with music can cause melodic as well as harmonic distortions ... (Colouration) is particularly prominent with broad band spectra, with heavy instrumentation and especially with percussion instruments" (Müller - author's translation).

Other workers in this field however have concluded that analysis by the auditory system responsible for detecting colouration occurs in the time-domain, on the model of Licklider |46|, rather than the frequency domain. Atal, Schroeder and Kuttruff |47| proposed that, since response fluctuations in the steady state response in a room can be of the order of 30 dB, the reason such fluctuations are generally not perceived is due to short-time analysis by the ear (rather than infinite-time analysis). An experiment to measure the threshold of colouration for white noise and its repetition (presented diotically through earphones) gave the result in Figure 4.3. It was suggested that the threshold is related to the maximum value of the short-time autocorrelation function $\Phi_s(\tau)$ for $\tau \neq 0$, where

$$\Phi_s(\tau) = \Phi(\tau).e(\tau),$$

$\Phi(\tau)$ being the actual autocorrelation function and $e(\tau)$ being the auto-correlation function of the weighting function.

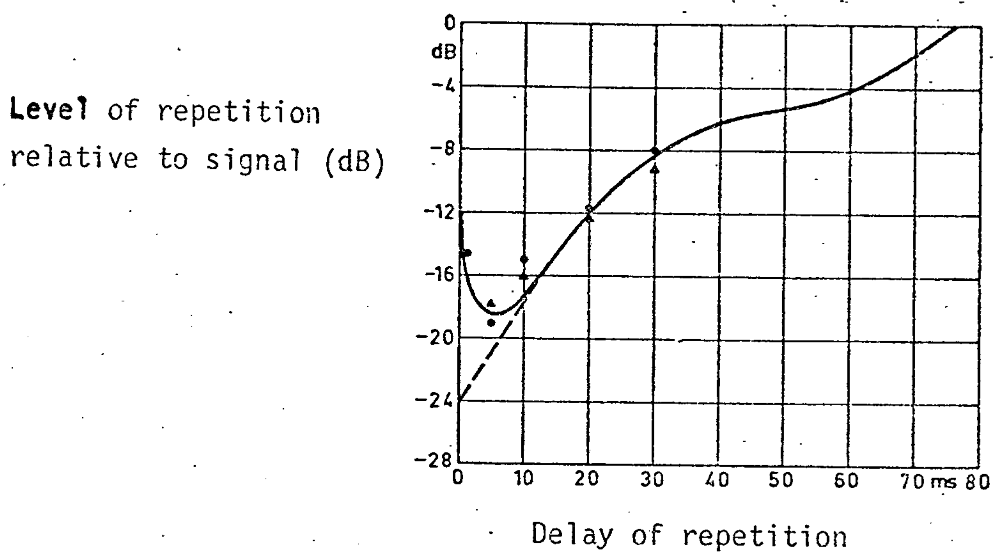


Figure 4.3. Threshold of a delayed repetition for white noise, after Atal et al |47|.

Bilsen |48| was able to extend the work of Atal et al to obtain a better agreement with subjective results. Bilsen claimed that two subjective effects are perceived with white noise: a repetition pitch ($f = 1/\tau$, for a signal and its repetition delayed τ) and a timbre change. He considers that these two effects are perceived simultaneously; in particular the threshold is the same for each. For music both effects are apparent as colouration. Bilsen was able to confirm that analysis was performed in the time-domain, and proposed a revised autocorrelation weighting function $\rho'(\tau)$, shown in Figure 4.4. Colouration with a white noise signal (plus repetitions) is

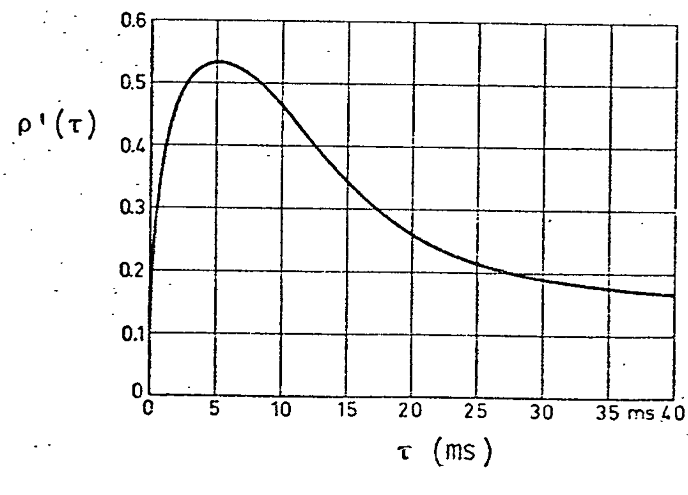


Figure 4.4. Revised autocorrelation weighting function for colouration, after Bilsen |48|.

perceived if

$$\frac{\Phi(0)}{\Phi(\tau)_{max}} < 16\rho'(\tau), \tag{4.1}$$

where $\Phi(0)$ is the autocorrelation function of the signal for $\tau = 0$, and $\Phi(\tau)_{max}$ is the maximum value of the autocorrelation function for $\tau \neq 0$. $\tau \to \infty$

Kuttruff |49| published a method of finding periodicities in a room. The method is based on the room's autocorrelogram, which may be derived from the impulse response, $h(t)$;

$$h(t) = \sum_k a_k . \delta(t - t_k) ,$$

a_k is the reflection intensity, δ is the Dirac delta function. The autocorrelation function of this response is given by

$$\phi(\tau) = \delta(\tau) . \sum_k a_k^2 + \sum_{k \neq \ell} \sum a_k a_\ell \delta(\tau + t_k - t_\ell) . \tag{4.2}$$

This is illustrated in Figure 4.5 for the simple case of a signal and its

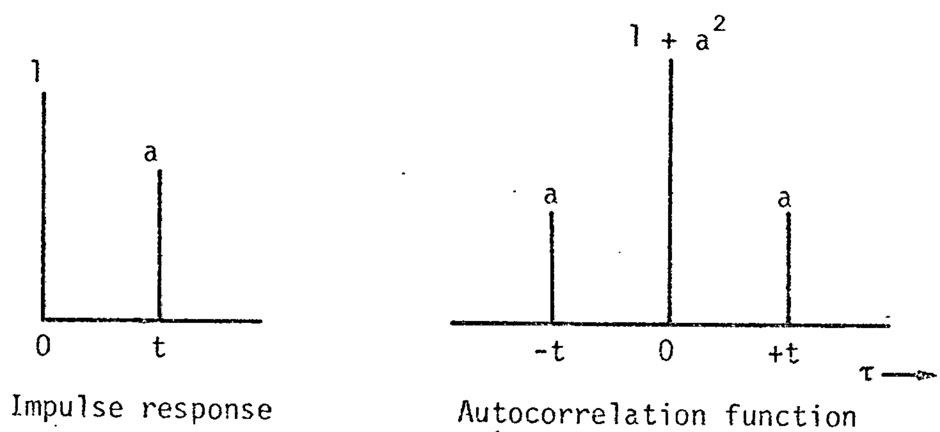


Figure 4.5. Impulse response and autocorrelation function for a signal plus its repetition.

repetition delayed by t. For periodic equally-spaced reflections (such as occur with a flutter echo), the secondary maxima of the autocorrelation function are particularly intense relative to the central maximum at $\tau = 0$. Kuttruff proposed that the ratio Δ of the peak at $\tau = 0$ to the highest peak of the autocorrelogram is a suitable quantitative measure for the number and strength of periodic reflections, which was called 'Temporal Diffusion'. However the ratio Δ is identical to the left-hand side of the inequality (4.1); there is thus a means of predicting that colouration will be perceived in a room, if

$$\Delta = \frac{\Phi(0)}{\Phi(\tau)_{max}} < 16\rho'(\tau). \qquad (4.3)$$

For the case of direct sound and a single reflection, the criterion in expression (4.3) is equivalent to the threshold in Figure 4.3. Kuttruff also describes an elegant method of deriving the autocorrelation function for a source and receiver in a room, which employs nothing more than a loudspeaker, microphone, tape recorder and storage oscilloscope.

The criterion in expression (4.3) corresponds to white noise. Threshold measurements by Somerville et al. |50| suggest that perception of colouration with speech is not as acute as with noise, so this criterion is a safe one for music. The criterion, however, is applicable to monophonic listening; colouration is frequently detectable on a microphone recording, but not detectable when listening in a studio oneself |51|. Flanagan and Lummis |52| have also investigated this phenomenon. It is very evident in a simulation of early reflections that frontal reflections produce considerably higher degrees of colouration than lateral reflections, which explains the difference between monophonic recording and binaural hearing. Bilsen's criterion is thus probably too strict for lateral reflections. The implications of perception of colouration for concert halls will be discussed in section 24.4.

(d) Echo disturbance.

High level reflections of delay > 50 ms became disturbing; as the delay increased the level at which they first became disturbing decreased. The onset of disturbance was determined by the two subjects; the average results are included in Figure 4.1 as a curved solid line. This result is in close agreement with the 20% disturbance limit measured by Muncey, Nickson and Dubout |53| for fast string music, using a large number of trained subjects. Muncey et al. found, however, that the disturbance limit also depended on music tempo. Determining the threshold of disturbance is much complicated by the fact that the preceding reflection sequence critically determines the threshold: Meyer and Kuhl |54| observed that the degree of disturbance produced by a reflection of delay, say, between 50 and 100 ms, will be much reduced by adding a preceding reflection. Fortunately in the real reverberant case the situation becomes simpler.

Nickson, Muncey and Dubout |26| measured the threshold of disturbance for reverberated speech and music; it is evident that the dependence on the signal is considerably less for the reverberant situation. Dubout |55| was able to get good agreement between measured and predicted results according to the following criterion: an echo is detected if at any time its envelope lies above the envelope of the original sound. He found, however, that for short delays account should be taken of the transient response of the ear, an approach pursued by Niese |56|. Niese postulates that the ear integrates incident energy as a ballistic instrument with a certain time constant. Loudness is proportional to pressure squared; for convenience let loudness $L = p_o^2$ for a continuous sound with pressure amplitude p_o. The perceived loudness is determined by the following equation:

$$\tau_o \frac{dL}{dt} = |p^2(t)| - L, \tag{4.4}$$

where τ_o is the auditory time constant. For sounds of duration t, the loudness is thus given by $L = p_o^2(1 - e^{-t/\tau_o})$. Subjective measurements fit this description well, giving a value of $\tau_o = 23$ ms. Following continuous excitation the instantaneous loudness has an "after-ring" of $L(t) = p_o^2 e^{-t/\tau_o}$. However, with a reverberant signal, the exciting signal has a characteristic

$$|p^2(t)| = p_o^2 e^{-t/\tau_R}, \qquad \tau_R = \text{R.T.}/13.82. \tag{4.5}$$

For this function, the solution of equation (4.4) is

$$L(t) = p_o^2 \frac{\tau_R e^{-t/\tau_R} - \tau_o e^{-t/\tau_o}}{\tau_R - \tau_o}. \tag{4.6}$$

Niese suggests that 50% of listeners will be disturbed if the echo pressure amplitude p_t^2 exceeds L by 3 dB for the particular value of t: i.e., if

$$\frac{p_E^{\,2}}{p_o^{\,2}} > 2L \qquad \text{or } 10 \log \frac{p_E^{\,2}}{p_o^{\,2}} > 10 \log L + 3.$$

This criterion fits the results of Nickson et al. |26| very well. Dubout's approach stresses the dynamic characteristics of the actual sound whilst Niese's stresses the transient behaviour of the ear. Both are certainly relevant.

(e) Spatial impression.

None of the effects of a single lateral reflection mentioned so far constitute a significant positive (i.e., desirable) contribution to the sound quality. It was found that, for the majority of reflection situations, the subjective effect of a side reflection was "spatial impression"; it occurs for all delays > 5 ms. When, for example, the level of a 40 ms delay lateral reflection is increased from threshold, the source appears to broaden, the music begins to gain body and fullness. One has the impression of being in a three-dimensional space. As the reflection level is increased, the amount of source broadening is also increased, until for high reflection levels there is an image shift. This broadening effect, or spatial impression, is easy to appreciate; subjects in fact found it relatively easy to equate two spatial impressions. Readers unfamiliar with the effect are referred to Band B on the tape at the back of this thesis.

The dominant shaded area in Figure 4.1 indicates the extent of "spatial impression". The density of shading is varied simply to indicate that the degree of "spatial impression" increases as the reflection level rises. Other authors' reports of this effect have already been listed in section 2.5, calling it "spatial responsiveness" |6|, "ambience" |57|, etc. Marshall |6| gives a good description of spatial impression from the Manager of the Concertgebouw Orchestra of Amsterdam, who described it as the difference between feeling inside the music and looking at it, as through a window. The impression is different, however, from that produced by reverberation; the latter tends to remove the starkness of anechoic music, providing a certain degree of being surrounded by the sound and giving an impression of distance from the source.

This binaural impression will here be called "spatial impression" (corresponding to the German "Raumeindruck"), though reverberation also produces a form of room or spatial impression. In the absence of any agreement between authors as to a suitable term, the term "spatial impression" (S.I.) does at least convey its binaural and subjective qualities. The bulk of the remainder of this thesis is concerned with an investigation of spatial impression; Chapter 13 discusses the probable auditory mechanism involved.

4.3 SUBJECTIVE EFFECTS OF A CEILING REFLECTION

Subjective observations on the effects produced by a ceiling reflection ($\alpha = 0^{\circ}$, $\beta = 40^{\circ}$) included level change (similar to that produced by a side reflection), image shift and tone colouration. Both the latter were more intense than the same effects with a side reflection and occurred for the majority of delay and level situations.

Somerville <u>et al</u>. |50| noticed that the Haas (or precedence) effect does not determine the point of localisation for a ceiling reflection: one localises in between the direct sound and the ceiling reflection. Localising above the orchestra in halls with a reflector has also been reported, as has tone colouration. It is possible that the localisation shift is due to the tone colouration, since localisation in the vertical plane is due to peaks in the frequency response at the eardrums |58|. Blauert |59| reports that with a fixed head interference effects determine the point of localisation in the median plane for sound from in front and behind with a small relative time delay ($\lesssim 0.5$ ms).

Chapter 5

REFLECTION THRESHOLDS

5.1 INTRODUCTION

Threshold measurement has been the basis of much audiological work and theory, but perhaps one reason for the relatively slow progress in the study of early reflections in concert halls has been that threshold results are rarely directly applicable in the multi-dimensional situation that exists for appreciation of music. The virtue of a threshold, that for its measurement no prior assumptions need be made of the subjective dimension for measurement, is also its limitation: it tells us simply that reflections below this level produce no audible effect, the threshold level being the lowest level at which the reflection can just be detected.

The threshold of reflections with speech have been measured by Seraphim |60, 10|; a typical result for a single lateral reflection masked by direct sound and a further frontal reflection are given in Figure 5.3. The gradient of the sloping sections of the threshold curves is about 0.5 dB/ms. A similar threshold is encountered with tone impulses |61|, which suggests that the impulsive nature of speech is critical at threshold. As can be seen in Figure 5.3, the presence of additional reflections considerably influences the threshold behaviour, indeed the details with only one additional reflection become relatively complex |10|. Additional reflections with speech cause not only forward but also backward masking: i.e., addition of a reflection of delay x ms affects not only the threshold of reflections arriving later than x ms but also before. The form of this masking was found to be dependent on the reflection directions. Seraphim estimated that on average only 12-15 reflections are perceptible in concert halls. As regards diffuse reflections it was found that behaviour is determined by the total reflection energy.

The threshold of a single reflection with music has been measured by Schubert |38, 62|; Figure 5.1 contains a typical result. It can be seen that the slope of the threshold curve is much less with music than with speech: for a pizzicato motif Schubert obtained a slope of 0.4 dB/ms, but in general slopes between 0.12 and 0.04 dB/ms were measured. The implication of the small slope with music in multiple reflection situations will be discussed in section 5.3. Schubert found that, for the masking of a single reflection by direct sound, that the threshold is dependent on reflection azimuth, but independent of elevation; this characteristic will be further discussed in section 8.8. The threshold determining criteria were found to vary with reflection delay, though this has already been mentioned in the previous chapter.

5.2 THRESHOLDS OF SINGLE REFLECTIONS WITH MUSIC

The threshold of a single side reflection in the presence of direct sound was measured. For this experiment conducted at Southampton an attenuator in 1 dB steps was used, whilst experiments conducted at Perth, reported in section 5.3, used an attenuator in 1.5 dB steps. A self-testing procedure was adopted, subjects as well as adjusting the attenuator could also switch off the test reflection for reference.

Schubert |38| compares the threshold results of experienced and inexperienced listeners as well as test subjects who had frequently performed such threshold measurements. Whilst threshold results of the latter were found to be lower than those of an average group of listeners, such results were highly consistent and were very suitable for comparison between different threshold situations. All experiments reported in this chapter were conducted by subjects who had been "trained" by performing at least two threshold measuring sessions.

The mean threshold for a single side reflection with azimuth $\alpha = 40^{o}$ as measured by two experienced subjects is given in Figure 5.1. The result is for the Mozart motif at a mean level E = 81 dB. The intersubject differences were never more than 3 dB. The result obtained by Schubert for an analogous situation is also included in Figure 5.1; Schubert used a "choral" motif with reflection azimuth of 30^{o}; the result is taken from Abb.8 in reference |38|. The agreement is surprisingly good. The minor differences, that the first peak occurs for a different delay and that the slope for long delays is different, correspond with differences between thresholds with different motifs as shown in Schubert's Abb. 2 |38|.

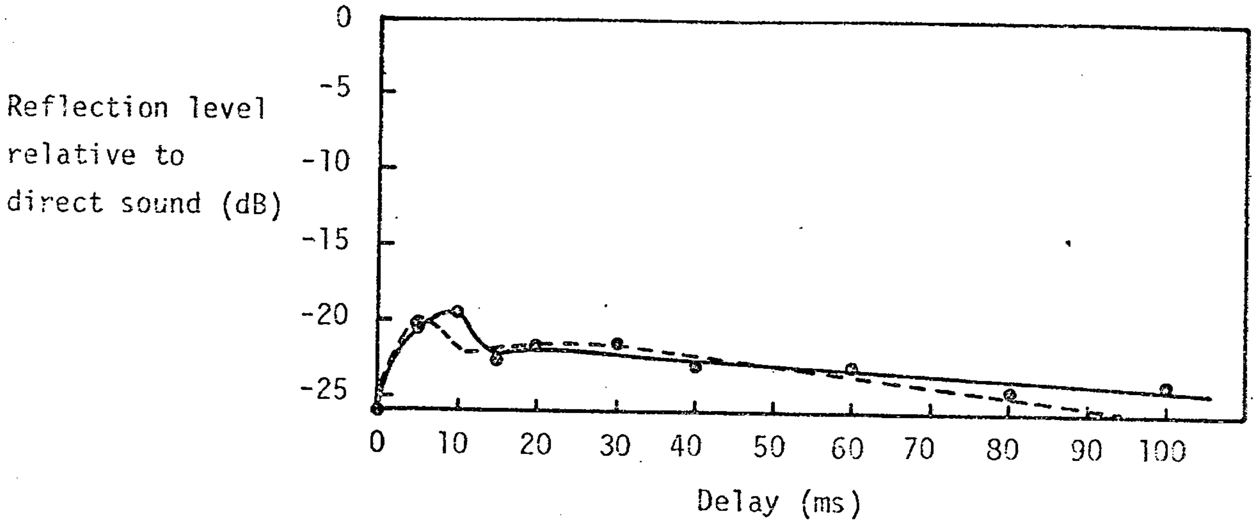


Figure 5.1. Threshold of a single side reflection for music.
o——o , Threshold for Mozart motif for reflection at $\alpha = 40^{\circ}$;
– – – , threshold (after Schubert) for "choral" motif for reflection at $\alpha = 30^{\circ}$.

Whilst the above measurement was performed at Southampton, the same experiment was repeated by four subjects in Perth; the result is the lower curve in Figure 5.2. The agreement between the two results is a reassuring subjective confirmation of the equivalence of the two simulation systems. The slope in each case of the threshold curve for long delays is 0.06 dB/ms.

5.3 THRESHOLD OF A SINGLE SIDE REFLECTION IN THE PRESENCE OF A CEILING REFLECTION

In general the first three reflections to arrive in a concert hall are from the two side walls and the ceiling. For the study of the effect of lateral reflections the possibility of masking by a ceiling reflection requires investigation. The measurement of a side reflection threshold in the presence of direct sound and a ceiling reflection has already been reported in reference |63|; however, the motif for that experiment was not purely anechoic. The experiment was repeated with the Mozart motif, though the conclusions are the same as previously reported.

The lateral reflection was at an azimuth of 40°, whilst the elevation of the ceiling reflection was also 40°. The ceiling reflection had a delay of 32 ms, and a level 2 dB below the direct sound level. Four subjects performed the experiment with the Mozart motif at a mean level E = 77 dB.

The measured threshold of the lateral reflection is plotted in Figure 5.2 as a function of reflection delay; the threshold level is relative to the sum of the direct and ceiling reflection level, incoherent addition being assumed. Mean 95% confidence limits for this threshold are ± 3 dB. For comparison, the threshold for a side reflection masked by direct sound only

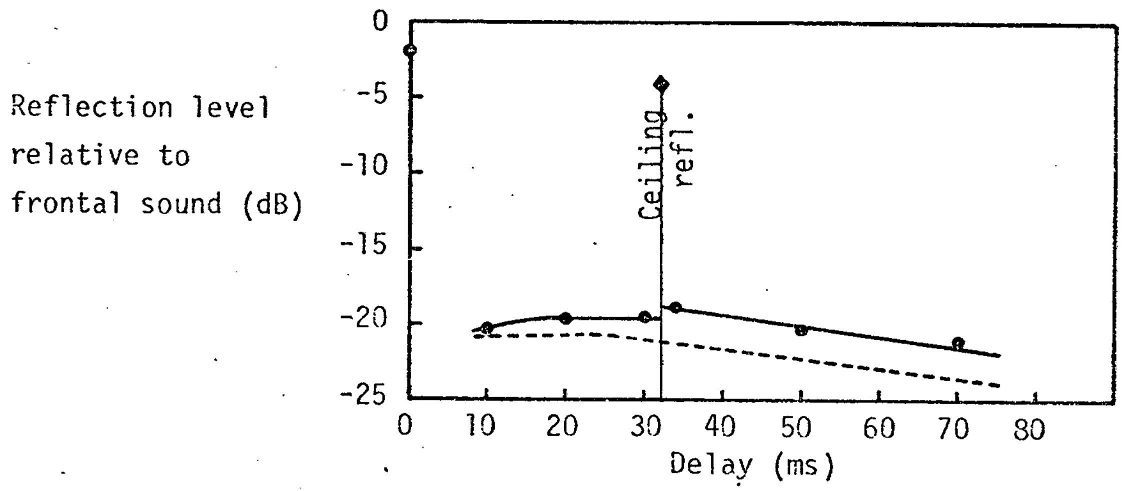


Figure 5.2. Threshold of side reflection with direct sound and a ceiling reflection for music (Mozart). – – – –, Threshold with direct sound only.

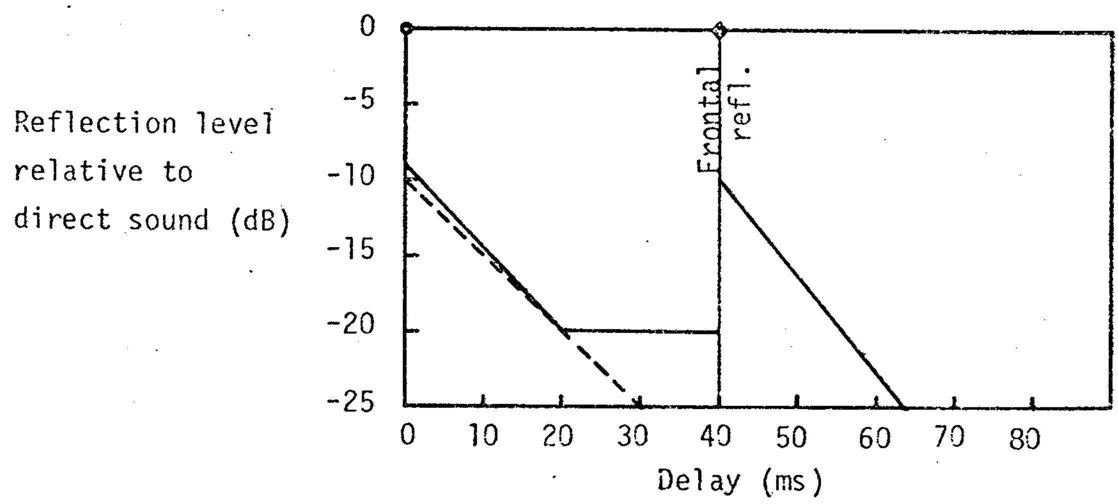


Figure 5.3. Threshold of side reflection with direct sound and a frontal reflection for speech (after Seraphim). – – – –, Threshold with direct sound only.

is also included. Figure 5.3 contains the corresponding thresholds for speech as measured by Seraphim |60|.

The effect on the threshold of a side reflection at the introduction of a ceiling reflection is best considered as two component effects: an overall threshold shift independent of delay and an additional threshold shift (d) for reflections arriving later than the ceiling reflection. For speech the former is insignificant, whilst the latter is determined by the delay of the ceiling reflection: the threshold level for delays just greater than that of the ceiling reflection is equal to that for 0 ms delay, so the threshold shift for speech (d_s) is 0.5 × x dB, where x is the ceiling reflection delay, and 0.5 is the gradient of the threshold line for speech. For the 40 ms ceiling reflection used by Seraphim, d_s = 20 dB.

With music the gradient of the threshold curve is considerably less than it is for speech; thus even if the mechanism is similar to that with speech, the threshold shift at the delay of the ceiling reflection for music (d_m) is considerably less. From Figure 5.2, d_m takes a value of the order of 1 dB, which is quite insignificant at threshold. Thus with music the relative delays of the ceiling and side reflection are not significant whereas with speech they are critically significant. The overall shift of the threshold when a ceiling reflection is introduced is about 2 dB with music. This again is a small change at threshold when viewed as a change of subjective impression, as will be seen in section 9.1. The results of the same experiment performed with a 57 ms delay ceiling reflection also supports the argument above.

5.4 CONCLUSIONS

The threshold results for a single side reflection agree well with Schubert's results |38|. Whilst with speech the introduction of a frontal or ceiling reflection is critical for the threshold of side reflections arriving after the ceiling reflection, this behaviour does not occur with music due to the much smaller gradient for threshold with delay. The maximum threshold shift at the introduction of a ceiling reflection is about 2 dB with music which is hardly significant at threshold; the threshold level in this case is determined substantially by the frontal sound level. With such a threshold level at about -20 dB lateral sound will only very rarely be masked by frontal sound with music, and most reflections will be above threshold. Just as Seraphim |10| discovered that with diffuse reflections the threshold behaviour is determined by the summed reflection level, so with music the subjective effect of three reflections at threshold is equivalent to the effect of a single reflection of level equivalent to the summed reflection level (see section 10.3). Since the threshold with music is relatively independent of delay, to view reflection masking as masking by total frontal and/or total lateral sound regardless of delay is more useful, and evades the inherent problem of using threshold curves: that two reflections below threshold constitute an above-threshold situation.

As was discovered by Schubert |62|, the threshold determining criteria depend on reflection delay. Whilst the threshold curve in Figure 4.1 appears to be related to the domains for the various subjective effects, the fact that more than one subjective effect determines the threshold precludes extrapolation from threshold to above-threshold behaviour. It is more salutary to view threshold behaviour as a limiting case of above-threshold behaviour. For this reason, thresholds will only be considered in the remainder of this thesis when below threshold situations are likely to occur with the total sound field in concert halls.

Chapter 6

THE COMPARISON TECHNIQUE FOR SUBJECTIVE EXPERIMENTS WITH SPATIAL IMPRESSION

6.1 POSSIBLE SUBJECTIVE EXPERIMENT TECHNIQUES

The principal effect of early lateral reflections was found in Chapter 4 to be a sense of envelopment, or source broadening, which will be referred to, collectively, as spatial impression. To determine the variation of degree of spatial impression (S.I.) the technique pioneered by Hawkes and Douglas |2|, as described in section 2.6, has the advantage of requiring very few assumptions, though for significant results to emerge a uniform understanding of the meaning, in subjective terms, of the descriptive scales clarity, warmth, intimacy, etc., is necessary. Alternatively one can present subjects with a lateral reflection situation and having told them that the effect of such a reflection will be called "spatial impression", ask them to indicate on a scale of spatial impression the perceived degree for varying situations.

However, if the assumption of working in a single subjective dimension is made, more precise quantitative experiments are possible. Two techniques are available: the method of constant stimuli and the self-testing comparison technique. For the method of constant stimuli the fixed test sound field, for which the degree of spatial impression is required, is presented for paired comparison with a varying comparison sound field; for each pair the subject is required to say which field produces the greatest sense of spatial impression. Since the variable comparison field is varied independent of the subject no bias can occur due to subject-variable situation interaction. For the self-testing comparison technique subjects can switch at will from the fixed test field to the variable field, which they also adjust until the best equivalence seems to exist regarding the degree of spatial impression. The advantages of the comparison technique are high precision and minimum experimental time. However, as an experimental technique it is valid only if certain requirements are fulfilled:

(a) that the test and comparison fields are sufficiently similar for a comparison to be made, and that additional subjective effects are secondary in perceived importance to that being measured;

(b) that the region of uncertainty, over which equivalence appears to exist, is small relative to the minimum change the subject can produce in the variable comparison field;

(c) that the degree of the subjective effect is monotonically related to the variable describing the variable comparison field.

6.2 THE COMPARISON TECHNIQUE FOR SPATIAL IMPRESSION

Throughout the remainder of this chapter the variation of spatial impression with delay will be discussed though the principles and techniques used were identical for variations of other physical reflection parameters. Figure 6.1 contains the echograms for a comparison experiment to determine the degree of spatial impression produced by a 10 ms pair of lateral reflections. Figure 3.1 shows the block circuit diagram for this experiment,

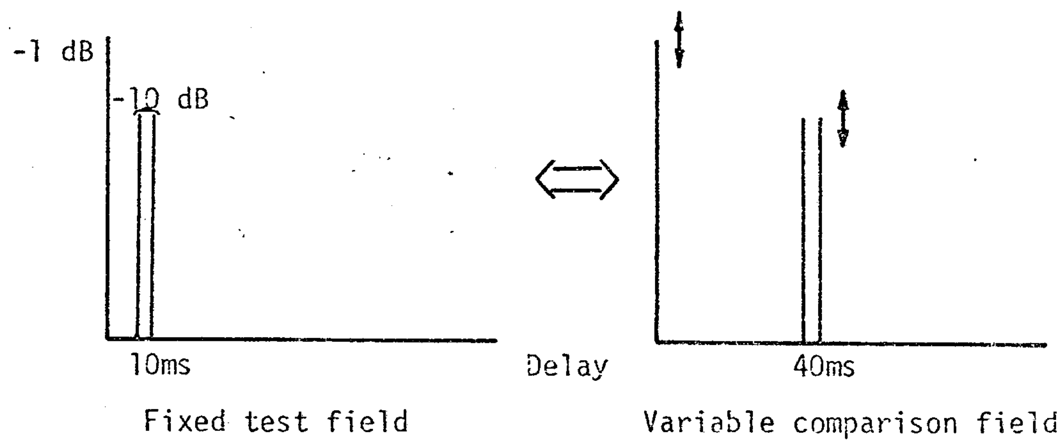


Figure 6.1. Echograms for sound fields used in the experiment to determine the variation of S.I. with reflection delay.

with one important difference: that here the delayed signal is fed to a pair of loudspeakers, wired in parallel, one on each side of the subject's head to provide a symmetrical situation more similar to the real situation. To avoid a degenerate situation, mentioned in section 13.6, one of the lateral reflection loudspeakers was placed at least 35 cm further from the subject's head than the other to introduce a minimum time delay difference of 1 ms.

To comply with requirement (a) above, loudness equality should exist between the two sound fields for comparison. To achieve this the differential attenuator, described in section 3.6, is used for the variable comparison field, so that as the lateral reflection level is increased the direct sound level is decreased to maintain a constant incoherent energy sum. The delay of 40 ms for the comparison field was chosen since it lies in the middle of the delay range for spatial impression, for which the predominant subjective effect is spatial impression (tone colouration being only just detectable, and the delay being insufficient for echo disturbance to occur). The level of this reflection relative to the direct sound is used as a measure of the degree of spatial impression.

Before reporting an analysis of variance for this experiment it is convenient at this point to discuss to what extent the requirements listed at the end of the previous section are fulfilled. In each case if a requirement is not fulfilled this would be made obvious by a large spread of subjective results. Requirement (c) is fulfilled for experiments in which the ratio of lateral to frontal sound is varied (see section 9.1). Requirement (b) was also completely fulfilled; on occasions subjects considered that equivalence would be achieved for a point intermediate between the steps on the differential attenuator, which were only 1 dB apart. Depending on the similarity or dissimilarity of delays being compared, so the requirement (a) is fulfilled better or worse. The particular additional effects that complicate subjective comparison will be mentioned in the next chapter. This is reflected in the scatter of results, though since in each case the 95% confidence limit of the mean was similar or less than the minimum step on the differential attenuator and was in each case less than one difference limen of spatial impression, requirement (a) can be considered fulfilled.

6.3 ANALYSIS OF VARIANCE FOR THE DEGREE OF SPATIAL IMPRESSION WITH DELAY EXPERIMENT

For the experiment described in the previous section, using reflection sequences similar to those in Figure 6.1, 10 subjects performed the experiment with three values of reflection delay in the test field: 10 ms, 35 ms and 70 ms. Each measurement was replicated 4 times, yielding 120 experimental results in all. A two-factor analysis of variance was conducted, for the factors subjects and delay; the results are given in Table 6.1. The interaction term is significant at the 0.1% level, so F values for subjects and delay are related to the interaction mean square. It is found that the delay effect is highly significant, whilst the subject effect is not significant.

This analysis is not completely valid since the overall variance at each delay differs significantly. This is a consequence of the experimental design; the smallest variance occurs for the two comparison fields which are most similar, namely, for the delay value 35 ms. Common transformations of the data were tried; however, only a logarithmic transformation reduced the divergence of variance, but it remained significant at the 1% level. With this transformation, the interaction remains significant

TABLE 6.1

Analysis of variance for the degree of spatial impression with delay experiment.

Source	SS	DF	MS	F	Prob.
Subject	19.4	9	2.15	1.70	N.S.
Delay	151.4	2	75.70	59.60	<0.001
Interaction	22.8	18	1.27	3.42	<0.001
Residual	33.3	90	0.37		
Total	226.8	119			

at the 1% level, the subject effect remains insignificant, whilst the delay effect becomes even more significant (F-value = 80). Given the pronounced significance of delay, deriving a special transformation to deal with these results did not seem warranted.

To determine the relative contributions of inter- and intra-subject variance to the total experimental variance, a one-factor analysis was performed for each delay value. The subject effect was not significant for delays of 35 and 70 ms, but was highly significant ($P < 0.1\%$) for a delay of 10 ms. This significant subject effect explains the interaction in the two-factor analysis. For each delay value the residual mean square was very close to the value in Table 6.1 (i.e., $\sigma^2 = 0.37$ (dB)^2).

To summarise, the experimental design leads to results determined principally by delay rather than by any subject effect. For the delay situation where the two sound fields for comparison are most different, the overall variance is largest and is attributable to a highly significant intersubject variance. This suggests an experimental procedure, which uses numerous subjects rather than replicating measurements. The general procedure for subjective experiments reported in the remainder of this thesis was to use a sufficient number of subjects to obtain an adequately small 95% confidence interval of the mean. In each case a confidence limit of the order of one subjective difference limen was considered adequate. This procedure was considered valid since the majority of experiments were of an exploratory nature, i.e., to determine the dependence of a subjective effect on a physical parameter.

Chapter 7

THE VARIATION OF DEGREE OF SPATIAL IMPRESSION WITH REFLECTION DELAY

7.1 INTRODUCTION

The great virtue of using a simulation system over the real concert hall situation is that reflection parameters can be varied independently. In particular, in a concert hall for a particular seat position the level of a reflection is, in general, a function of its delay according to inverse square law attenuation. In addition to independent variability, with a simulation the dependence of a subjective effect on physical acoustic parameters can be derived with much greater precision than is likely in the real situation. This is particularly relevant for a study of early reflections, since each reflection is defined by at least four dimensions, and many reflections should be considered.

As a basic assumption, variation of spatial impression with one reflection parameter is considered independent of any others, i.e., that no interaction terms exist. No experiments were conducted to test this hypothesis explicitly, though in the large variety of experiments conducted none suggested it to be incorrect.

To determine the variation of spatial impression with reflection delay the comparison technique described in the previous chapter was used. As a measure of the degree of spatial impression, the level of a 40 ms lateral reflection for subjective equivalence is employed; this is not in fact a linear measure of the degree of S.I. as will be discussed in Chapter 9, but, being monotonic, it is suitable for this purpose. Experiments were conducted to derive a curve of equal spatial impression, and a curve of degree of S.I. for different delays, both with unilateral and bilateral reflections. The behaviour with a slow tempo motif and with the simultaneous presence of reverberation was also investigated.

7.2 CURVE OF EQUAL SPATIAL IMPRESSION

Four subjects on average conducted the experiment to derive a curve of equal spatial impression with the Mozart motif. A single lateral reflection was used at an azimuth of 40^{o}. With a 40 ms reflection at - 6 dB relative to the direct sound used as a reference, the aim of the experiment was to determine the required level of a reflection from the same direction but with a delay of x ms to produce the same degree of S.I. The echograms of the two sound fields, for comparison, are shown in Figure 7.1, the variable field being shown on the right. The loudness level was maintained constant

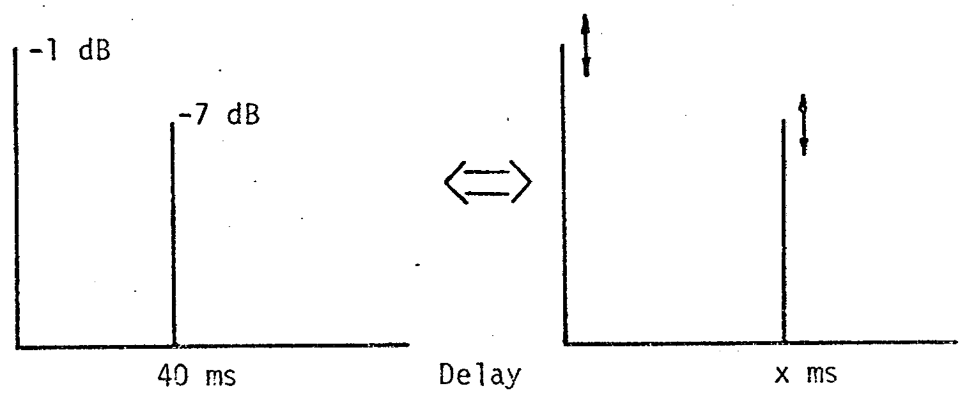


Figure 7.1. Echograms of sound fields for comparison in equal spatial impression experiment.

by the differential attenuator at a mean level of E = 77 dB.

The subjectively determined results of this experiment are included in Figure 4.1 as a solid line running across the figure. It is evident from this curve that for delays greater than 10 ms only small changes of reflection level are required to maintain the same degree of spatial impression. The gradient of the curve for reflection delays greater than 50 ms is 0.07 dB/ms, a value which compares well with a gradient of 0.2 dB/ms for the aUs measured by Reichardt and Schmidt |64| and gradients between 0.04 ... 0.4 dB/ms for the threshold with different music motifs found by Schubert |38|. As expected, the shape of the curve in Figure 4.1 proved to be the inverse of the curve of degree of S.I. against reflection delay, which will be discussed in the next section.

7.3 DEGREE OF S.I. AGAINST DELAY FOR BILATERAL REFLECTIONS

In this experiment seven subjects were used with the same motif at the same loudness level, but with a pair of lateral reflections at azimuth $\pm 40^{o}$ (again as in Chapter 6 the lateral loudspeakers were so placed as to give a minimum difference of arrival time between them of 1 ms). However, to determine the degree of spatial impression, contrary to the procedure described in section 7.1, here it is the level of the 40 ms reflection which is varied. The echograms for this experiment are shown in Figure 7.2.

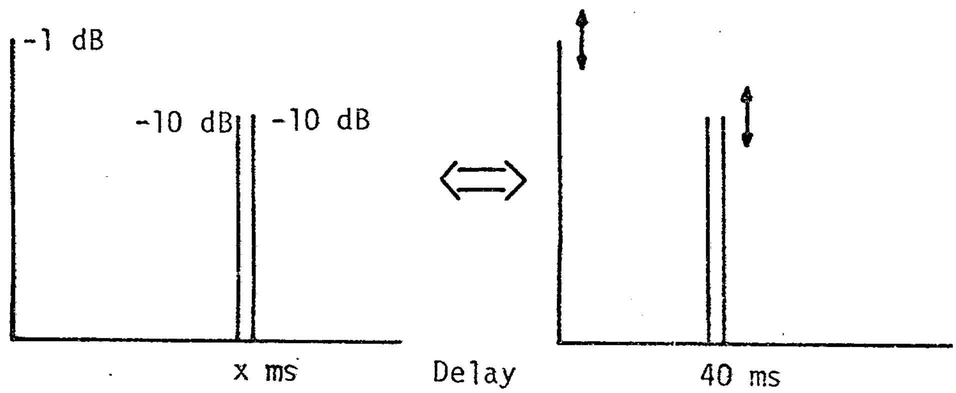


Figure 7.2. Echograms of sound fields for comparison in degree of S.I. experiment.

Reflection delays between 5 ms and 90 ms were used. Subjects appeared to have no particular difficulty performing this experiment in spite of additional subjective effects occurring, such as tone colouration and, for large delays, echo disturbance. As has been shown in Chapter 6, in which an identical experiment was described, scatter about the mean is substantially intersubject rather than intra-subject for this experiment.

The results are plotted in Figure 7.3; the ordinate here is the degree of S.I. as measured by the level of a 40 ms reflection pair for the subjectively equivalent degree of S.I. 95% confidence limits of the mean values are also included.

The level of the 40 ms reflection pair refers to the incoherent sum of the two reflection levels; the ordinate is thus the ratio of reflected to direct sound.

It can be seen that for the delays investigated, the degree of spatial impression is substantially independent of reflection delay. Considering that the difference limen for spatial impression is at least a 1.4 dB change in the ratio of lateral to frontal early energy (see section 9.1), one can say that for delays greater than 8 ms the spatial impression produced by

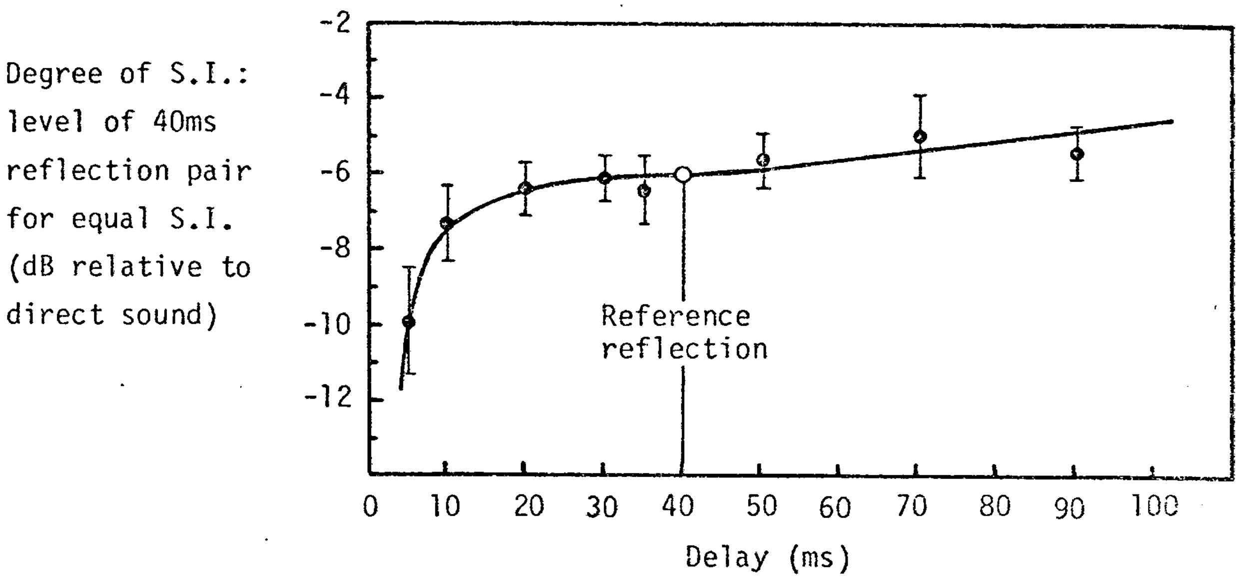


Figure 7.3. Curve of degree of S.I. against reflection delay for a lateral reflection pair ($\alpha = \pm 40^{o}$). , Mean and 95% confidence limit of the mean.

a lateral reflection is independent of delay. In concert halls reflections with delays less than 8 ms do occur, though not in the majority of seat positions in a hall.

7.4 DEGREE OF S.I. AGAINST DELAY FOR A UNILATERAL REFLECTION

Using an identical procedure to that described in the previous section, but with a single lateral reflection at a level equivalent to the incoherent sum of the levels of the lateral pair, produced the results shown in Figure 7.4. Both the mean and 95% confidence limits of the mean are shown. The dotted line is the curve for the bilateral case from Figure 7.3. In both Figures 7.3 and 7.4 the ordinate is the same quantity, the ratio of lateral reflected to direct sound. The agreement between the unilateral and bilateral case can be seen to be very good: within 0.5 dB, though statistical comparison is not possible due to significantly different variances.

Although the number of subjects in this experiment was on average only four, the scatter of results was very small: for all delays the variance in the unilateral case was less than, and for three out of five

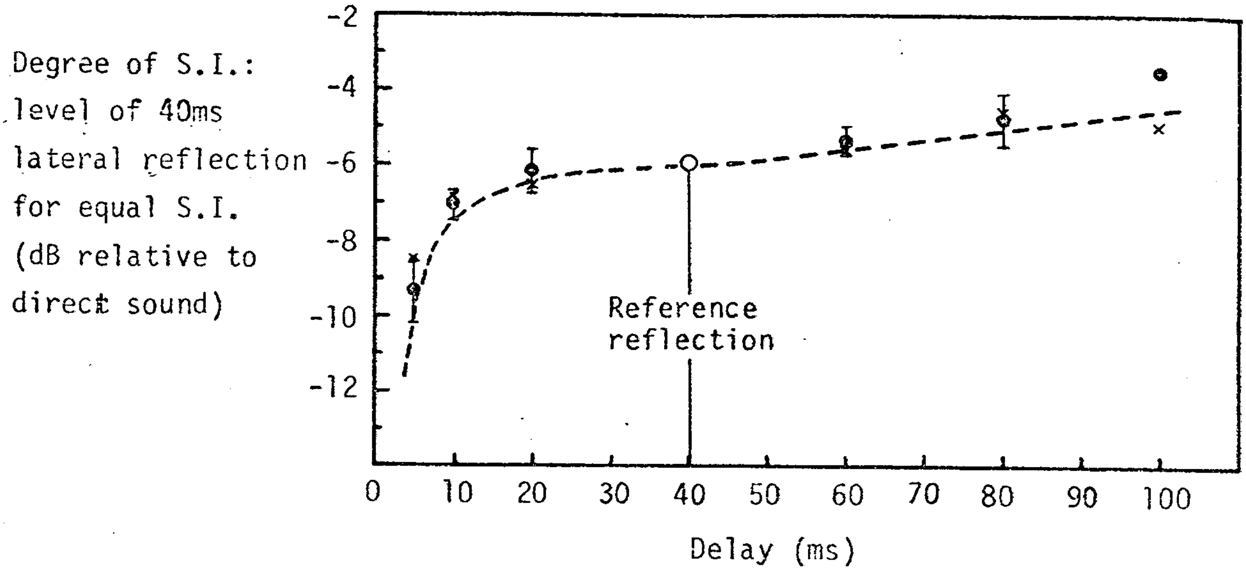


Figure 7.4. Curve of degree of S.I. against reflection delay for a single lateral reflection ($\alpha = 40^{o}$). , Mean and 95% confidence limit of mean, no reverberation; ×, mean of results with reverberation; - - - -, curve for lateral pair from Figure 7.3.

results significantly less than, the variance in the bilateral case. This is in spite of the fact that in the unilateral case there is the additional problem for short delays that the effect is more a shift of the point of localisation rather than a broadening of the source, which makes comparison difficult. This result indicates that subjects were better able to "concentrate" on a reflection from one side rather than both.

7.5 DEGREE OF S.I. AGAINST DELAY FOR A UNILATERAL REFLECTION WITH REVERBERATION PRESENT

The experiment in section 7.4 was repeated with reverberation added (at a stationary level of -8 dB relative to E and reverberation time = 1.5 s) and a further reflection (-12 dB rel. to direct sound, 83 ms delay, $\alpha = -40^{o}$, $\beta = -16^{o}$) added to produce a more natural sound. This "room sound" was identical in the two sound fields for subjective comparison. Subjects found this experiment more difficult, as was reflected in the 95% confidence limits, which were double those without reverberation added. There was, however, no significant change in the results by the addition of reverberation; the average points are included in Figure 7.4.

7.6 DEGREE OF S.I. AGAINST DELAY FOR BILATERAL REFLECTIONS USING A SLOW TEMPO MOTIF

The experiment with bilateral reflections described in section 7.3 was repeated by four subjects using the legato, slow tempo motif from Wagner's "Siegfried Idyll", at a mean level of 73 dB. The results are plotted in

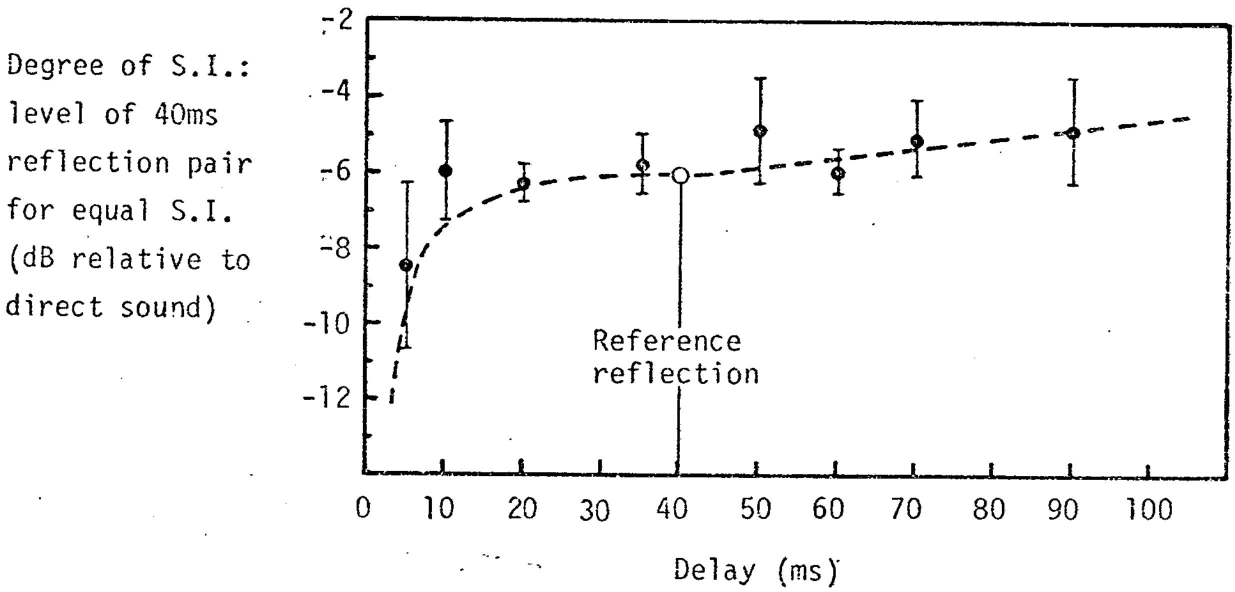


Figure 7.5. Curve of degree of S.I. against reflection delay for a lateral reflection pair for the Wagner motif. ⧲ , Mean and 95% confidence limit of the mean; ----, curve for Mozart motif (from Figure 7.3).

Figure 7.5; the dotted curve is taken from Figure 7.3, the corresponding result for the Mozart motif. The mean values, particularly at small delays, suggest a slightly different characteristic behaviour, though the agreement is within the 95% confidence limits. The contention that the degree of S.I. is independent of delay for delays greater than 8 ms remains unchallenged, however.

7.7 CONCLUSIONS

It was found that the degree of spatial impression is highly constant for lateral reflections of delays 20-60 ms. The degree of S.I. increases slightly for delays up to 100 ms, and decreases significantly for delays below 10 ms. However, since the difference limen for spatial impression is at least 1.4 dB as a change in the ratio of lateral to frontal energy, it can be said that the degree of spatial impression is independent of delay for delays greater than 8 ms. This result holds whether a pair of lateral reflections or just a single one is used, and is independent of motif and the presence or absence of reverberation.

Chapter 8

THE VARIATION OF SPATIAL IMPRESSION WITH DIRECTION OF INCIDENCE OF REFLECTIONS

8.1 INTRODUCTION

In experiments discussed in previous chapters, lateral reflections were generally at an azimuth of 40^{o} to straight ahead and an elevation near zero. This was chosen as being a typical reflection direction for sound reflected off a concert hall side-wall, being also convenient for simulation, though the choice was otherwise arbitrary. A comparison technique similar to that described in Chapter 6 is also suitable for investigating the transition from subjectively frontal to subjectively lateral sound. Subjects had to compare the degree of spatial impression produced by a reflection from one direction with that from another. Preliminary experiments soon indicated that results obtained with single reflections from just one side differed from results when a bilateral pair of reflections was used (see section 8.7). Since the bilateral situation corresponds much closer to the real concert-hall situation, the bulk of experiments were conducted with pairs of lateral reflections. Again the loudspeakers used to simulate the lateral reflections were placed so that for a pair there was a difference in arrival time of the two reflections of at least 1 ms.

Contrary to results reported in reference |63| (now considered erroneous), the maximum subjective spatial effect was observed to occur for azimuth angles of 90^{o}. Since this corresponds with what one considers physically as purely lateral, the experimental method which involved equating the spatial impression produced by a fixed level pair of reflections at azimuth α^{o} with a variable level pair of reflections at azimuth of 90^{o} seemed the obvious choice. To avoid confusion, in the remainder of this chapter, "lateral sound energy" will refer to energy from an azimuth of 90^{o}, and "subjectively lateral and non-lateral (or frontal)" refer to the subjective effect for azimuth 90^{o} and 0^{o} (or 180^{o}), respectively.

8.2 VARIATION OF DEGREE OF S.I. WITH AZIMUTH

Figure 8.1 indicates the echograms of the two sound fields to be compared in this experiment. For each a reflection delay of 40 ms was employed, and loudness level was maintained constant at a mean level of E = 77 dB by using the differential attenuator. The Mozart motif was used throughout and five subjects performed the experiments.

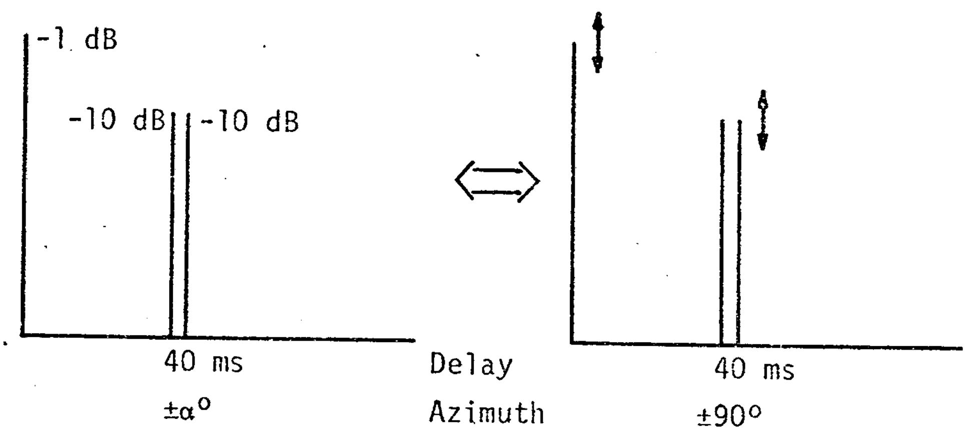


Figure 8.1. Echograms of sound fields for comparison in the variation with azimuth experiment.

In the fixed level situation (l.h.s. in Figure 8.1), the ratio of reflected to direct sound energy was -6 dB. Results of this experiment were in the form of the ratio of lateral to frontal early energy for 90° azimuth reflections which gave the same degree of spatial impression as a -6 dB ratio for that particular angle of azimuth. Results are plotted in Figure 8.2.

The results for azimuth angles of 0° and 90° are a priori $-\infty$ and -6 dB, respectively. The measured results are plotted as solid circles together with their respective 95% confidence limits of the mean. For an azimuth of 20° no confidence limit is included since all subjects gave the same experimental result. It is noticeable that results for angles of azimuth between 90° and 180° agree well with results for the complementary angles of azimuth (i.e., $180^{\circ} - \alpha^{\circ}$). Surprisingly, these reflections from behind were not perceived by subjects as coming from behind, but they simply created a pleasant sense of envelopment. A reflection at $\alpha = 180^{\circ}$ produced an apparent source in the median plane, with no sensation

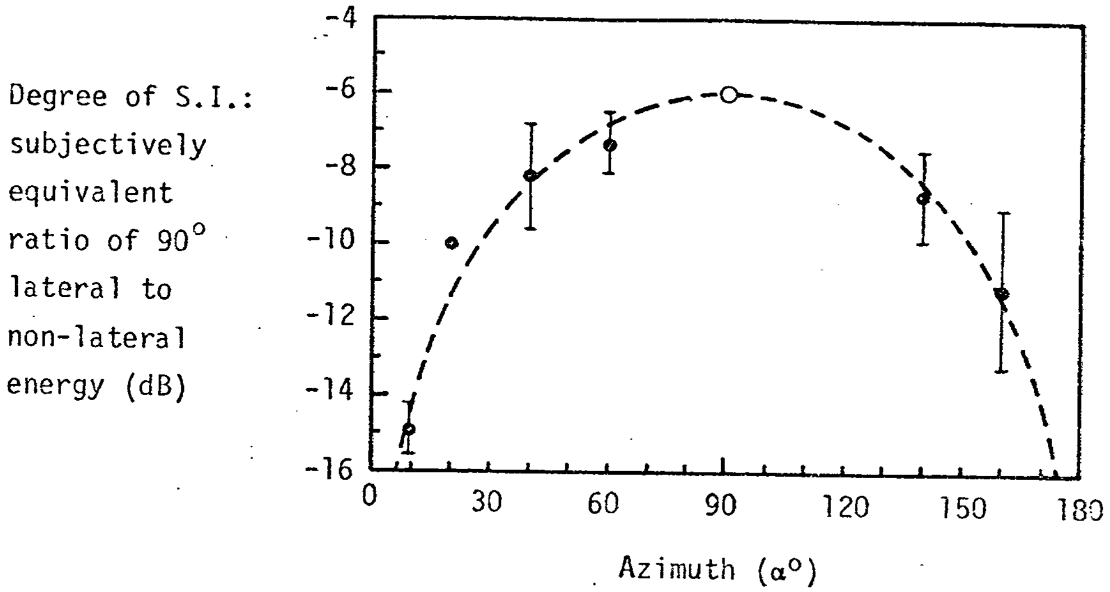


Figure 8.2. Degree of S.I. for pairs of reflections at different angles of azimuth. ɸ , Mean and 95% confidence limit of the mean of experimental results. Dotted line corresponds to predicted result for a sine relationship.

of source broadening apparent to permit a comparison of the two sound fields to be made. A value of $-\infty$ has thus been assigned to an azimuth of 180^{o}.

The experimental results can be treated in the following manner: a lateral reflection is assumed to contribute a proportion of its energy, λ, to the subjective lateral sound, and the remaining proportion $(1 - \lambda)$ to the subjective frontal sound. The parameter λ is thus a function of angle of azimuth and takes by definition the values $\lambda_{0^o} = 0$ and $\lambda_{90^o} = 1$.

In the above experiment the sound level of side reflections in the fixed situation was - 6 dB relative to the direct sound level. So if Δ is the direct sound energy, the reflected sound energy is $\Delta/4$. Thus in the fixed field used in the experiment the ratio of (90^o) lateral to non-lateral early sound was

$$\frac{\frac{\lambda\Delta}{4}}{\Delta + \frac{\Delta}{4}(1 - \lambda)} = \frac{\lambda}{5 - \lambda}$$

$$= \text{antilog} \frac{S_{90^o}}{10} , \qquad (8.1)$$

where S_{90^o} is the measured subjectively equivalent ratio of lateral to non-lateral early sound energy for 90^o azimuth reflection. From equation (8.1) and the results plotted in Figure 8.2 the parameter λ can be determined and this is tabulated in Table 8.1. In the final column the sine of the angle of azimuth is also listed, which can be seen to correspond

TABLE 8.1

Experimentally determined value of contribution to lateral sound for reflections of different angles of azimuth.

Angle of azimuth (α)	λ (measured)	sin α
0^o	0	0
10^o	0.16	0.17
20^o	0.46	0.34
40^o	0.66	0.64
60^o	0.79	0.87
90^o	1	1
140^o	0.59	0.64
160^o	0.35	0.34
180^o	0	0

pretty closely to the experimentally determined value of λ. To test the validity of the empirical relationship, $\lambda = \sin\alpha$, the experimental values corresponding to this relationship are shown by the dotted line in Figure 8.2. For all angles of azimuth, except $\alpha = 20^o$, the dotted line lies within the 95% confidence limit of the mean; the $\alpha = 20^o$ result also deviates by a typical 95% confidence limit. So one can conclude that, within experimental accuracy, $\lambda = |\sin\alpha|$, an extremely simple and convenient result.

8.3 COMPARISON OF NON-90^o LATERAL REFLECTION SOUND FIELDS

In the previous section, lateral reflection pairs at an azimuth of α^o were compared with a lateral reflection pair at 90^o azimuth. Two experiments were also conducted in which reflection pairs at α^o were compared with a reflection pair at 40^o azimuth, the chosen values of α being 10^o

and 20^o. An identical experimental procedure was used; five subjects adjusted the ratio of reflected to frontal sound at an azimuth of 40^o to obtain subjective equivalence with respect to spatial impression with a pair of reflections at α^o azimuth at a ratio of reflected to frontal sound of -6 dB. Experimentally determined results for $\alpha = 10^o$ and 20^o are given in Table 8.2. The fourth column in Table 8.2 contains the predicted equivalent ratio (S_{90^o}) of lateral to non-lateral sound energy for 90^o azimuth reflections, on the assumption that the parameter λ takes the value $|\sin\alpha|$. This is determined as follows:

$$\text{antilog}\ \frac{S_{90^o}}{10} = \frac{\lambda . P}{\Delta + (1-\lambda)P}, \qquad (8.2)$$

where the ratio of reflected to direct sound P/Δ is used, P being the reflection sound energy. The fifth column contains the predicted ratio of reflected to direct sound energy for 40^o azimuth reflections which would produce the value of S_{90^o} in column four. The fifth column thus contains the predicted experimental result according to the hypothetical relationship $\lambda = |\sin\alpha|$.

TABLE 8.2

Results of experiment with 40^o azimuth comparison field.

Angle of azimuth	Subjectively equivalent ratio of reflected to direct for 40^o reflections (dB)	95% conf. limit of the mean (dB)	S_{90^o} (dB)	Predicted ratio of reflected to direct for 40^o reflections (dB)
10^o	-11.4	±1.1	-14.4	-12.4
20^o	-8.2	±0.6	-11.3	-9.2

Whilst the agreement between experimentally determined and predicted results is within the 95% confidence limit for the case of $\alpha = 10^o$, this is not the case for $\alpha = 20^o$. As the discrepancy is in the same direction as that illustrated in Figure 8.2 for $\alpha = 20^o$, it appears that the relationship $\lambda = \sin\alpha$ does not hold exactly for $\alpha = 20^o$. This deviation will be further discussed at the end of section 8.5.

8.4 VARIATION OF DEGREE OF S.I. WITH ELEVATION

Using an identical procedure to that used in section 8.2, but varying the elevation, β, rather than the azimuth, leads to results plotted in Figure 8.3. Unfortunately, due to the limited ceiling height of the anechoic chamber used for these experiments, only two angles of elevation proved feasible. A priori results for $\beta = 0^{o}$ and 90^{o} are -6 dB and $-\infty$ dB, respectively; thus one can postulate the relationship $\lambda = |\cos \beta|$. The predicted result according to $\lambda = |\cos \beta|$ is shown as the dotted line in Figure 8.3.

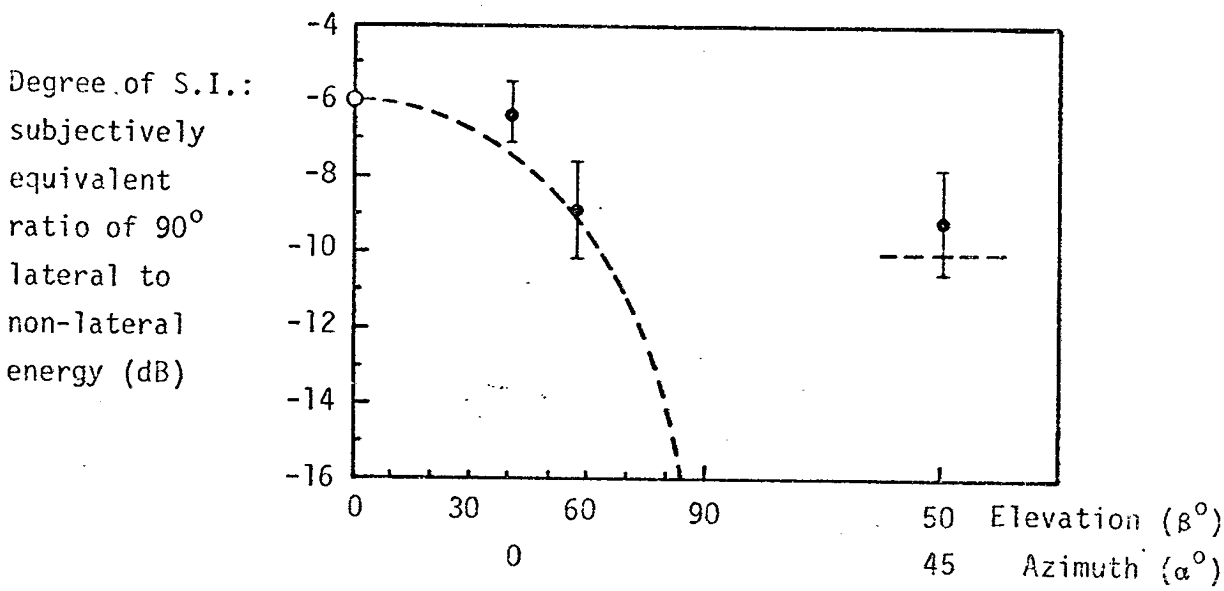


Figure 8.3. Degree of S.I. for pairs of reflections at different angles of incidence. ϕ , Mean and 95% confidence limit of the mean of experimental results. Dotted line corresponds to predicted result.

Agreement between predicted and measured results is good for $\beta = 57^{o}$, but just outside the 95% confidence limit for $\beta = 40^{o}$.

8.5 THE GENERAL CASE

Figures 8.2 and 8.3 show that the nearer the direction of incidence of a reflection approaches to lateral the larger is the spatial effect it produces. To deal with the general case a relationship is required between λ and the angle ϕ between the direction of incidence and the lateral axis which passes through the listener's ears. The relationship must be such that for $\alpha = 90^{o}$, $\lambda = \cos \beta$ and for $\beta = 0^{o}$, $\lambda = \sin \alpha$.

Figure 8.4 shows a listener's head with the lateral axis through his ears, and the frontal axis corresponding to straight ahead. The circle (in

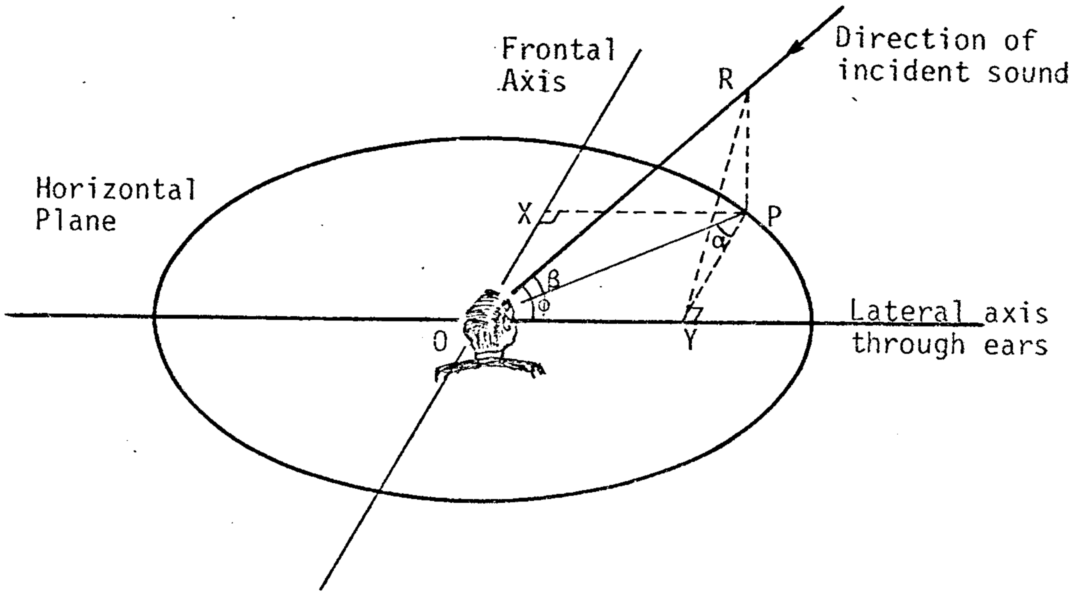


Figure 8.4. Perspective diagram for incident sound relative to a listener.

projection) lies in the horizontal plane. From the diagram it can be seen that

$$\sin \alpha . \cos \beta = \frac{OY}{OP} \cdot \frac{OP}{OR} = \frac{OY}{OR} = \cos \phi .$$

Making $\lambda = \cos \phi$ satisfies the conditions above; thus one has the empirical relationship

$$\lambda = \cos \phi = \sin \alpha . \cos \beta , \qquad (8.3)$$

which can be used for the completely general case. No modulus sign is required in the relationship $\lambda = \cos \phi$, due to the definition above of the angle ϕ. One can alternatively consider the angle of the incident sound to the median plane, Φ, but since the median plane is perpendicular to the lateral axis, one arrives at the result $\lambda = \sin \Phi$.

Figure 8.3 also includes the result of an experiment with the general case $\alpha = 45^{\circ}$ and $\beta = 50^{\circ}$. The predicted result according to equation (8.3) is again given as the dotted line; agreement is within the 95% confidence limit of the mean of the experimental results.

The statistical analysis described in Chapter 6 suggests that the variance in these experiments is substantially intersubject. Minor deviations from $\lambda = \cos\phi$ were experienced in these experiments (for $\alpha = 20^{o}$, $\beta = 0^{o}$ and $\alpha = 0^{o}$, $\beta = 40^{o}$). However considering the limited precision of these experiments, and the fact that in complex real situations the difference limen for spatial impression is about 1.4 dB (see section 9.1), these minor deviations do not warrant a refinement of the relationship $\lambda = \cos\phi$, which is highly convenient in its simplicity. Results with the relationship $\lambda = \cos^2\phi$, for instance, are far outside the 95% confidence limits.

8.6 PHYSIOLOGICAL IMPLICATIONS

Figure 8.5 shows a projection similar to Figure 8.4, but the listener's ears are here labelled E_1 and E_2, a distance d apart. It can be seen from

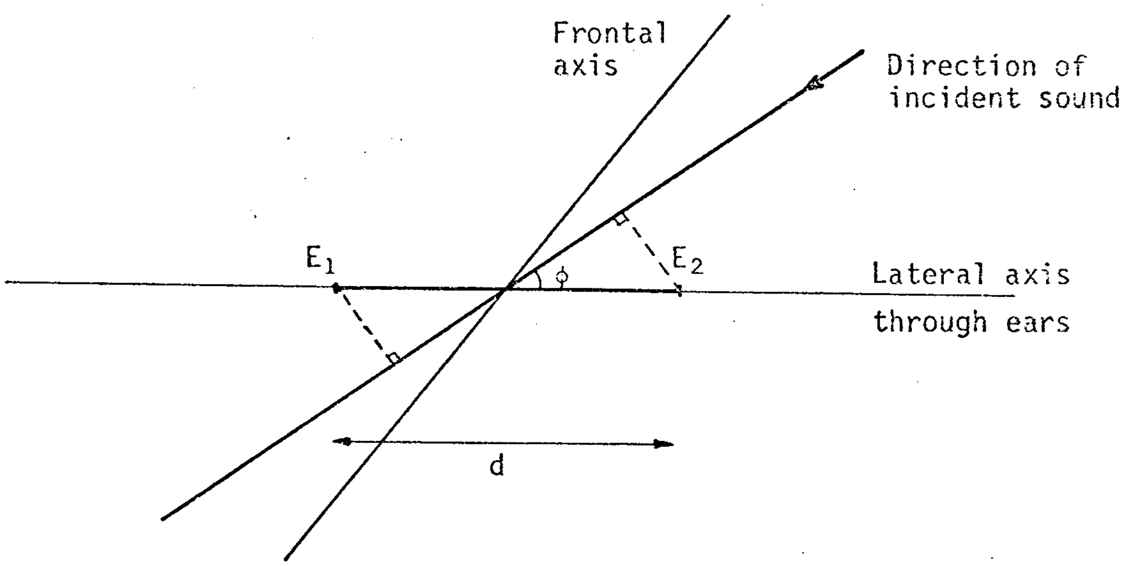


Figure 8.5. Perspective diagram to illustrate interaural time difference.

Figure 8.5, that if E_1 and E_2 are two points in space, then:

Contribution to lateral sound for a reflection (= $\cos\phi$)

$$= \frac{\text{Interaural time difference}}{\text{Maximum interaural time difference}} \qquad (8.4)$$

However the two ears are separated by the head, which produces an area of shadow. The implications of this for interaural time difference (which will be described in section 13.2) are that the right hand side of the relation (8.4) is for angles of azimuth: $(\alpha + \sin \alpha)/(1 + \frac{\pi}{2})$. The errors involved in assuming it to be $\sin \alpha$ are however small, and as suggested above such refinement is not warranted. The important implication, however, of the relation (8.4) is that it is the interaural time difference which is important. This consideration is basic to the theory, discussed in Chapter 13, that spatial impression is a cross-correlation process.

8.7 EFFECT OF DIRECTION FOR UNILATERAL REFLECTIONS

As was mentioned in reference |63|, to compare two single lateral reflections of different azimuth proved to be subjectively difficult. It involved comparison of sound from two non-identical loudspeakers, and variations in tone colouration were also apparent which made comparisons of spatial impression more difficult. Perhaps more significant was the fact that with a single reflection there is a slight shift in the point of localisation, rather than in terms of pure spatial impression. This was demonstrated in an experiment conducted with a single $\alpha = 40^{\circ}$ reflection which was compared with a single 90° reflection. This experiment gave the result of -10.7 ± 1.5 dB, significantly smaller than the bilateral result of -8.2 dB which appears in Figure 8.2.

Experiments with bilateral reflections obviate most of the subjective difficulties encountered with unilateral reflections and, as has been mentioned above, the bilateral case approximates much closer to the real situation.

8.8 THRESHOLD VARIATION WITH ANGLE OF AZIMUTH

Detection of a lateral reflection is relatively easier than detection of a frontal one. If, as suggested in Chapter 13, a cross-correlation process is involved in detection of spatial impression, the cross-correlation mechanism must then be more sensitive than alternative mechanisms which have to be employed to detect a frontal reflection. Schubert |62| has measured thresholds with a series of music motifs and discovered that the variation of threshold with azimuth is substantially independent of reflection delay and musical motif (a result which supports the general assumptions of this thesis, that spatial impression behaves very similarly for different motifs and that the effects on S.I. of varying different reflection dimensions are substantially independent of one another). He obtained an average difference in threshold between a frontal and a 90^o lateral reflection of 10 dB, (see Bild 10 in |62|), and a mean value for the threshold of a 90^o lateral reflection of -25 dB relative to the direct sound.

It can be hypothesised that a reflection is either detected by the spatial impression it produces or by the frontal component. The limiting value for spatial impression is thus a ratio of lateral to non-lateral sound of -25 dB. At threshold, from equation (8.2), with $\lambda = \sin\alpha$:

$$\text{antilog}\,\frac{-25}{10} = 3.10^{-3} = \frac{P.\sin\alpha}{\Delta + P(1-\sin\alpha)}\,, \qquad (8.5)$$

and except for α very small, this may be simplified since $P \ll \Delta$;

$$3.10^{-3} = \frac{P}{\Delta}\,.\,\sin\alpha$$

or

$$\frac{P}{\Delta} = -25 - 10\log\sin\alpha \text{ in dB.} \qquad (8.6).$$

Hence the threshold value P/Δ may be calculated for different values of azimuth. This is plotted as the dotted line in Figure 8.6.

According to Schubert the mean threshold value for a frontal reflection is -15 dB. Since the frontal component of a lateral reflection decreases with increasing azimuth, the threshold due to the frontal component is $\geqslant$ -15 dB. Only for $\alpha = 6^o$ does P/Δ from equation (8.6) equal -15 dB, thus only for $\alpha < 6^o$ is it possible that reflections are detected at threshold by the frontal component; i.e., the line plotted in Figure 8.6 is predicted to be valid for $\alpha > 6^o$. The values measured by Schubert (from Bild 10 in |62|, assuming an $\alpha = 90^o$ threshold of -25 dB) are also included as the solid circles in Figure 8.6. Agreement is well within the 95% confidence limits of Schubert's measurements.

Schubert (again in reference |62|) points out that the threshold variation with azimuth is much larger for a reflection than it is for a single sound. With a simple experiment he demonstrated that the simultaneous presence of a frontal sound (which may be totally uncorrelated in acoustical content to the "reflected" sound) causes an increased directional discrimination for lateral ("reflected") sound. A threshold measured by

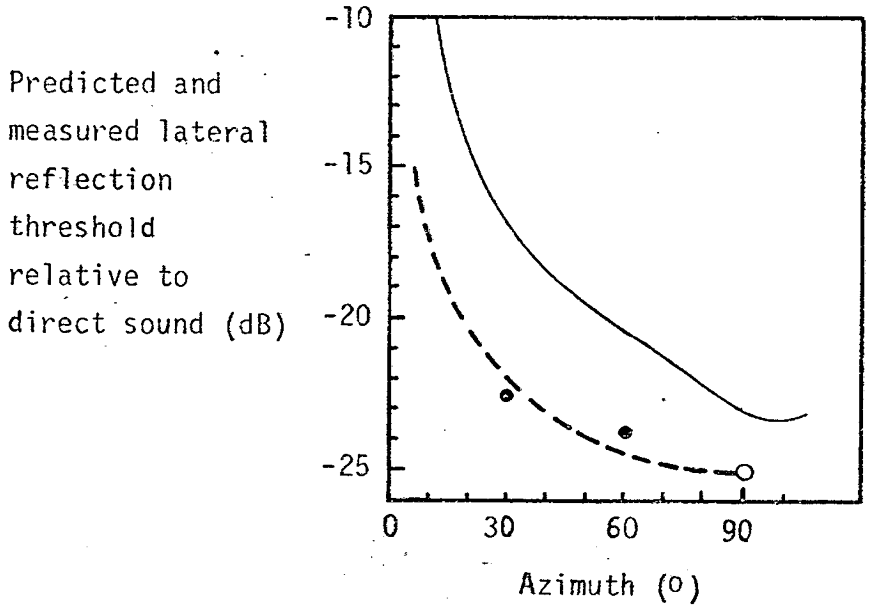


Figure 8.6. Predicted (— — —) and measured (•, after Schubert |62|) threshold for a lateral reflection as a function of azimuth. (———), Threshold value for incoherent signals after Damaske |65|.

Damaske |65| is also included in Figure 8.6; this is the threshold for an uncorrelated lateral filtered music signal (250-2000 Hz) in the presence of a frontal signal. Except for a small threshold shift the agreement between Damaske's result and the predicted result is good. These threshold results will be further discussed in section 13.3, but the agreement between them suggests a similar mechanism in each case.

8.9 CONCLUSIONS

Experiments involving comparison of the spatial impression produced by reflections from different directions showed that the greatest degree of S.I. was created for "purely lateral" reflections at an azimuth of 90^o. For the general case it was considered that a reflection contributed a proportion λ to the lateral sound and a proportion $(1 - \lambda)$ to the frontal sound. It was found that within experimental accuracy (with two minor deviations) that the experimentally determined λ equalled $\cos\phi$ ($= \sin \alpha.\cos \beta$), where ϕ is the angle of incidence to the lateral axis through the listener's ears, α is the angle of azimuth, and β is the angle of elevation. This relationship also predicted results in good agreement with Schubert's measurements of threshold variation of reflection azimuth |62|.

Chapter 9

VARIATION OF SPATIAL IMPRESSION WITH SOUND LEVEL

9.1 VARIATION OF DEGREE OF S.I. WITH LATERAL REFLECTION LEVEL

It was initially noted in section 4.2(e) that the degree of spatial impression varies with reflection level for a lateral reflection, and indeed the level of lateral reflected sound relative to the direct sound has been used as a measure of the degree of spatial impression in the preceding chapters. It is very obvious subjectively that the subjective difference limen for changes in lateral reflection level is not a constant, and the subjective degree of spatial impression is not a linear function of lateral reflection level relative to direct.

Measurements of difference limen made by Reichardt and Schmidt |64| permit the derivation of a relationship between subjective spatial impression and reflection level. The simulation field used by Reichardt and Schmidt is shown in Figure 9.1. The presence of the reverberant field need not

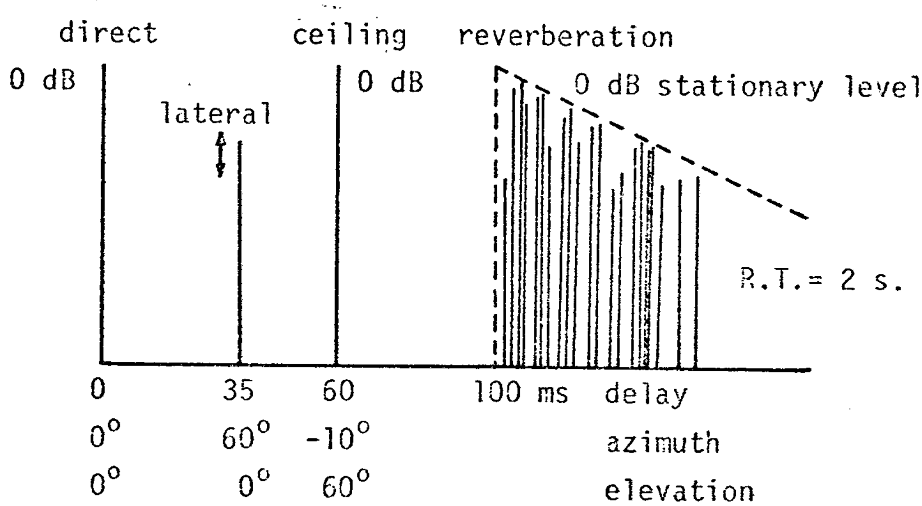


Figure 9.1. Echogram of test field used by Reichardt and Schmidt |64|.

be of concern here beyond the consideration that the results of the experiment would be applicable to the real (reverberant) situation; experiments have shown the subjective effects of reverberation to be substantially independent of those of early reflections (see Chapter 12). Subjects were presented with sound fields in pairs for comparison, the variable being the lateral reflection level, and were asked whether they could detect a change or not. The maximum sound level of the total field was 76 dB; one can estimate the mean level as being about 70 dB. From these results they derived the minimum detectable change or difference limen for different lateral reflection levels. The difference limen results of Reichardt and Schmidt are replotted in Figure 9.2 for the case of level decrease - this is taken directly from Figure 3 of reference |64|; the abscissa being the lateral reflection level relative to the direct sound.

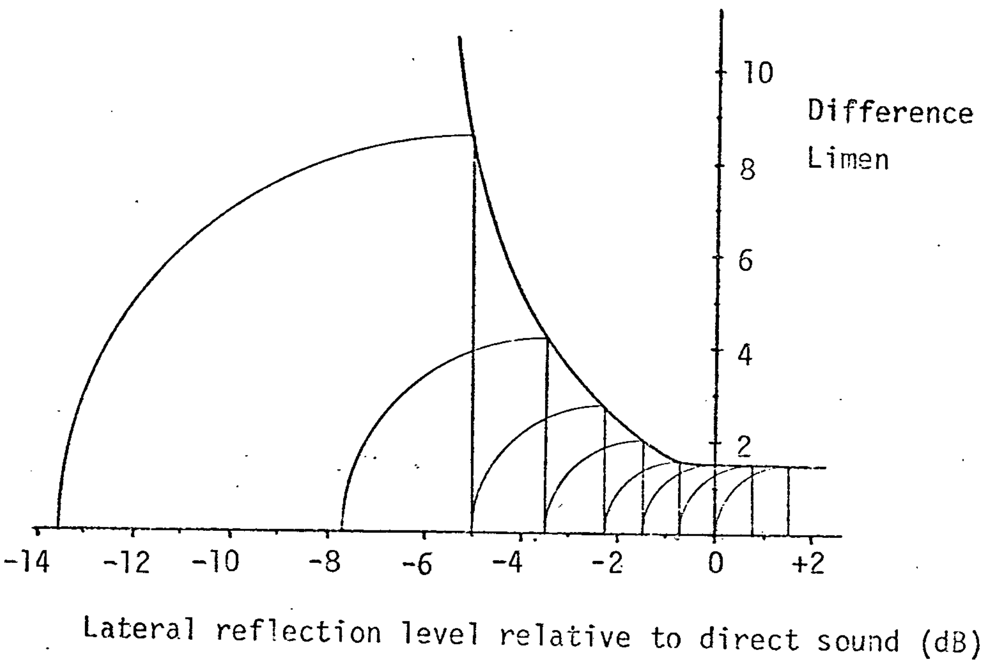


Figure 9.2. The relationship between difference limen and reflection level for a lateral reflection.

Reichardt and Schmidt in another paper |35| describe a method whereby a scale of degree of the subjective quality may be derived from difference limen results. Since, from reference to Figure 4.1, it is apparent that the principal subjective effect of a 35 ms lateral reflection is spatial impression, it is reasonable to assume that the difference limens were determined from changes in spatial impression. A change in level equivalent to the difference limen can be considered as a change of one subjective unit of S.I. Thus a scale of degree of spatial impression can be derived as a function of lateral reflection level. The procedure is as follows: placing a compass point at the +1.5 dB point on the abscissa, one draws a quarter circle, of radius equal to the difference limen of 1.5 dB. Where the circle cuts the abscissa, one places the compass point again and repeats the process. In this way five points are obtained on the degree of S.I. curve in Figure 9.3 (the points for degrees of S.I.: $\frac{1}{2}$, $1\frac{1}{2}$, $2\frac{1}{2}$, $3\frac{1}{2}$ and $4\frac{1}{2}$).

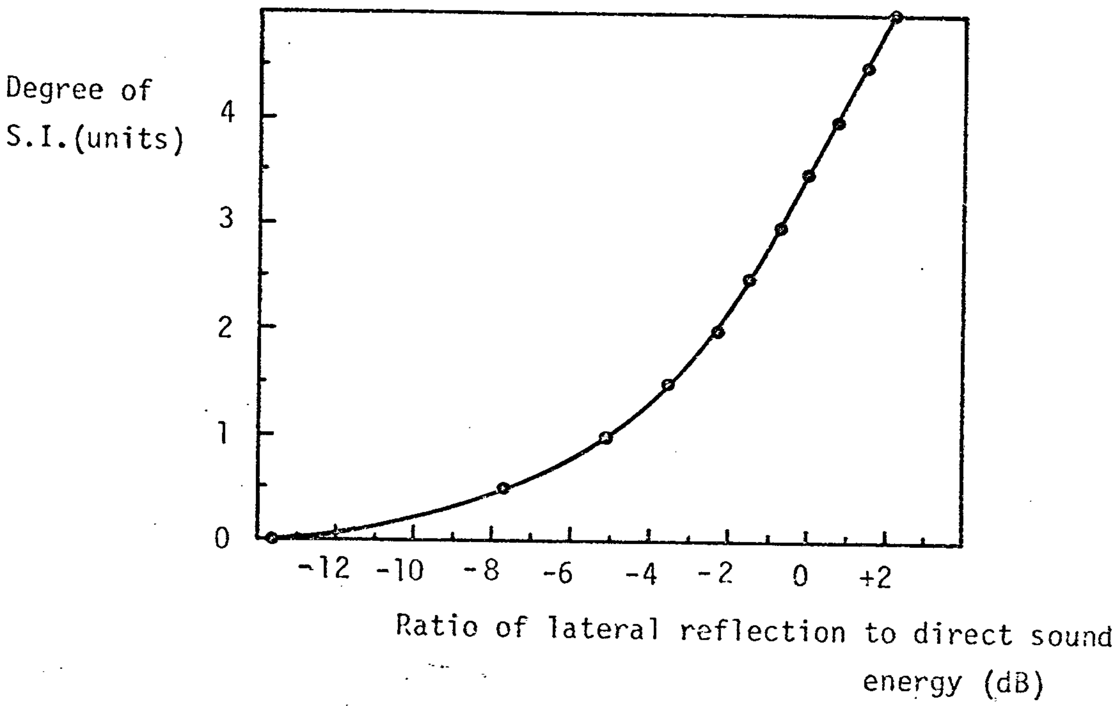


Figure 9.3. Relationship between the degree of S.I. and the lateral reflection level relative to the direct sound in the reflection sequence in Figure 9.1.

The choice of +1.5 dB is, however, arbitrary as a starting point. Since in Figure 9.2 the difference limen is constant at 1.5 dB for high level reflections, an interval of 0.75 dB is equivalent to half a subjective unit of S.I. at these high reflection levels. So one can obtain a new series of points for Figure 9.3 by using the same graphical method but starting at +0.75 dB reflection/direct level. The interval between points in Figure 9.3 is thus half a subjective unit of S.I. The designation of zero degree of S.I. for a reflection level of -13.6 dB corresponds with the threshold measured by Reichardt and Schmidt.

As a result of experiments involving ceiling reflections, reported in Chapter 10, the effect of a ceiling reflection on the degree of S.I. is such that the degree of S.I. is determined by the ratio of lateral to non-lateral early sound. This means, roughly, that for a ratio of lateral reflection to direct sound of 0 dB in Reichardt and Schmidt's experiment, the presence of the ceiling reflection causes the ratio of lateral to non-lateral early sound to be -3 dB. A further small correction is, however, also necessary: since the lateral reflection azimuth is 60^o, a correction according to the results of Chapter 8 should be made to obtain the ratio of (90^o) lateral to non-lateral early sound, S dB ($S = S_{90^o}$). If f_d, f_c and f_ℓ are the levels of the direct sound, ceiling and lateral reflections, respectively, then

$$\text{antilog}\ (\frac{S}{10}) = \frac{\sin 60^o . f_\ell}{f_d + f_c + (1 - \sin 60^o) f_\ell} . \qquad (9.1)$$

Since f_ℓ/f_d is the abscissa in Figure 9.3, and $f_d = f_c$ in Reichardt and Schmidt's experiment, a graph of degree of S.I. against ratio of lateral to non-lateral early sound, S(dB) can be derived. This is shown in Figure 9.4.

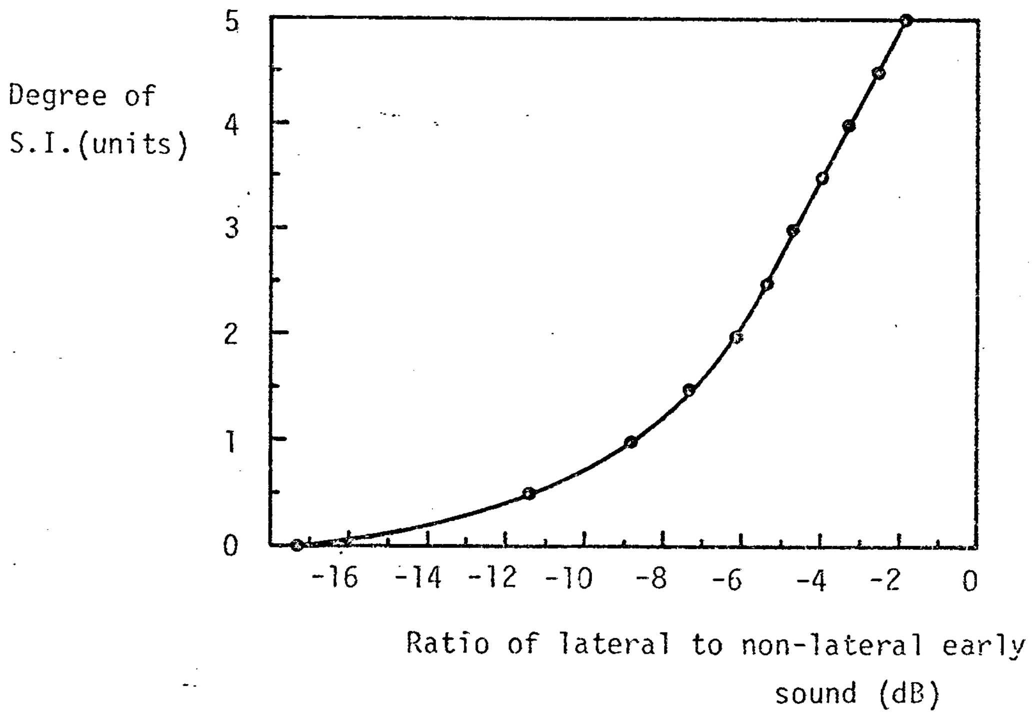


Figure 9.4. Relationship between the degree of S.I. and the ratio of lateral to non-lateral early sound, S dB.

The result is much as expected from subjective observation: the curve is monotonically increasing and the rate of change of S.I. at low levels is much less than it is for high levels of lateral sound. It will be noticed that the degree of S.I. is very nearly linear with respect to the ratio of lateral to non-lateral sound (S) for values of the latter > −6 dB. Since typical hall values for S lie in this region, this result facilitates manipulation. In this region one subjective unit of S.I. corresponds to a change in S of 1.4 dB.

Wagener |66| has also measured the variation in spatial effect with a single lateral reflection. Subjects were required to indicate the perceived area (in the upper hemispherical space) from which they heard sound. The increase in perceived area as a function of lateral reflection level is in good agreement with Reichardt and Schmidt's results, though the precision of Wagener's results is probably low. For the proportion of lateral sound generally found in concert halls, the perceived area was found to be similar both for a fast and slow motif.

Wagener also investigated the variation of perceived area with reflection delay and direction. It is very unfortunate, however, that for each of these investigations a reflection level equal to that of the direct sound was used. This is both unrealistic for concert halls and subjectively in such a simulation one is aware of two discrete sources: i.e., no longer a situation of spatial impression. Wagener's experimental method could provide an independent check on the results found by comparison experiments. However, experiments with a bilateral pair of reflections are to be preferred.

9.2 VARIATION OF DEGREE OF S.I. WITH MUSIC LEVEL

Measurements by de V. Keet (reported in reference |8|) show that the degree of spatial impression is a function of music sound level. Keet made stereo recordings of single source recordings of orchestral music in real halls. These were replayed to subjects at different listening levels via a pair of loudspeakers. The subjects were asked to assess the apparent source width in degrees, with reference to a scale placed in front of them. Measurements made with recordings taken in four positions (in three concert halls) gave the result that on average for higher levels a change of 1.4^{o} occurs in the apparent source width for a change in listening level of 1 dB.

Further analysis of Keet's results in Chapters 13 and 14 (see especially section 14.4) indicates that the difference limen for spatial impression in terms of apparent source width is about 5.3°, a change which would occur for a change in listening level of about 4 dB. As the difference limen for level for a monophonic music recording is of the order of 3 dB, this is a very small value. It means, for instance, that the change in degree of spatial impression during a piece of music, for which typically the dynamic range is at least 20 dB, is of very similar order to the maximum variation in S.I. between or within halls - in both cases about five subjective units.

The actual level of the early sound is thus a significant factor in the degree of spatial impression produced in a concert hall. The implications of this for desirable concert hall designs are discussed in Part II. This result, however, also illustrates the particular dynamic nature of spatial impression: that one's impression of being surrounded by sound is most intense in 'forte' music passages. It probably also accounts for the very exciting sound which occurs in small halls (in which the absolute level of early sound is high).

Chapter 10

THE DEGREE OF SPATIAL IMPRESSION IN MULTIPLE REFLECTION SITUATIONS

10.1 INTRODUCTION

From the unlimited variety of possible multiple reflection situations available for comparison, two situations were chosen to indicate the behaviour of subjective spatial impression: the effect of the introduction of a ceiling reflection and the effect of multiple lateral reflections. The behaviour of the subjective degree of S.I. under these two situations will now be discussed.

10.2 THE DEGREE OF S.I. WITH A CEILING REFLECTION

An experiment to determine the effects of a ceiling reflection has been reported in reference |63|. Whilst the conclusions of this experiment remain as they were reported, the experiment has been repeated with seven rather than just three subjects, and with modifications to overcome the problems that were encountered in the earlier experiment. Since a sound field with a ceiling reflection was compared with one without, the additional effects of tone colouration and vertical shift in the point of localisation, which occur only in the presence of the ceiling reflection, complicate subjective comparison. Further, with a single lateral reflection the subject's attention, if not also to a certain extent the perceived point of localisation, is drawn to one side. This was remedied by using a lateral pair of reflections as described in previous chapters, whilst the vertical shift in point of localisation was minimised by using the minimum practical angle of elevation for the ceiling reflection (which was $\beta = 12^o$). With this system tone colouration effects did not prove troublesome, the scatter of results about mean values being no larger than in the equivalent experiment without a ceiling reflection described in section 7.3. The echograms for the two fields for comparison are shown in Figure 10.1; the ceiling reflection is at a delay of 32 ms, whilst the delay of the lateral reflections in the fixed field was varied from 5 to 90 ms.

As previously, the Mozart motif was used with the loudness level kept constant by the differential attenuator at a mean value of 77 dB. The azimuth of the lateral reflections was $\pm 40^o$. Again subjects were told to

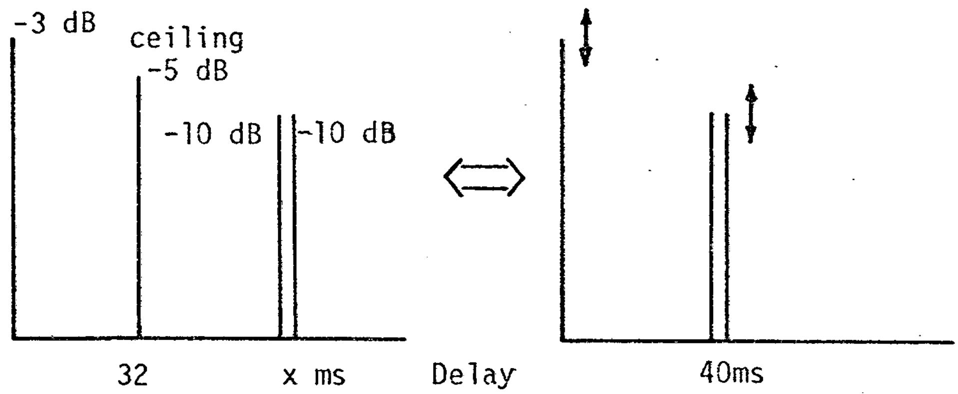


Figure 10.1. Echograms for comparison to determine effect of ceiling reflection.

equate the degree of spatial impression, and to ignore other subjective changes.

The results of this experiment are plotted in Figure 10.2; again the degree of S.I. was measured as the level of a 40 ms delay reflection pair to produce an equivalent spatial effect. The dotted line in Figure 10.2 shows the variation in S.I. with delay without a ceiling reflection present (from Figure 7.3). According to the hypothesis that the ceiling

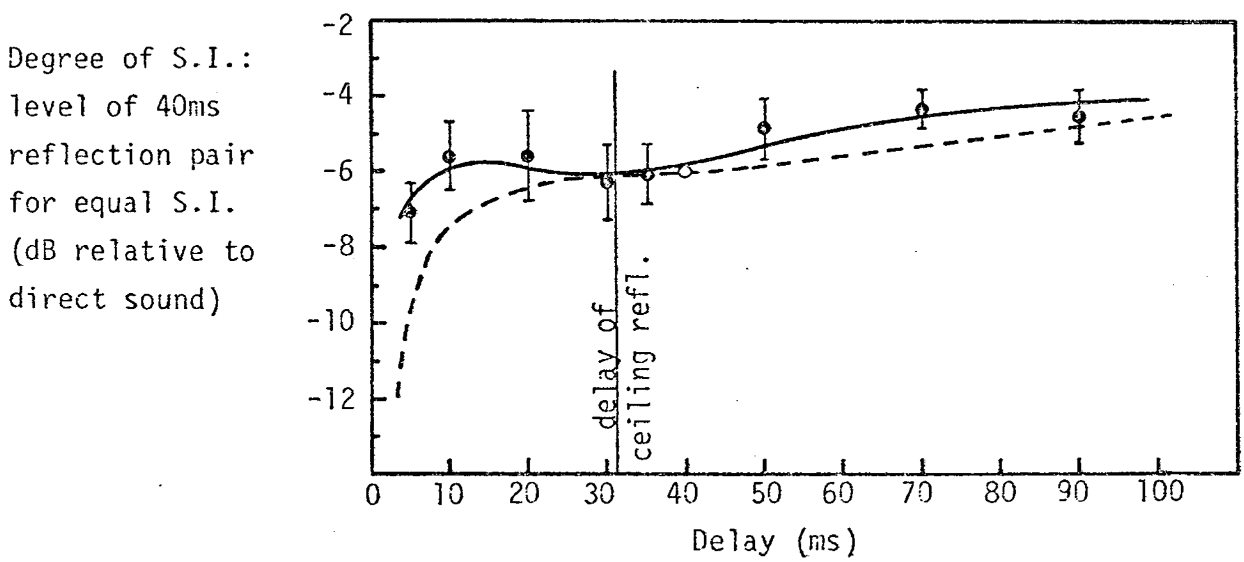


Figure 10.2. The degree of S.I. for a sound field containing a 32 ms ceiling reflection as a function of lateral reflection delay. , Mean and 95% confidence limit of the mean; (- - - -) curve of degree of S.I. without ceiling reflection (from Figure 7.3).

reflection level adds incoherently to the direct sound level for subjective spatial impression, or in other words that the degree of spatial impression is a function of the ratio of lateral to non-lateral early energy, then the two curves in Figure 10.2 would be identical. However, with the additional assumption that the degree of S.I. is independent of reflection delay within the difference limen of 1.4 dB, the hypothesis is supported for results in this experiment of between -7.4 and -4.6 dB. This is the case for all lateral reflection delays other than 5 ms, for which the degree of S.I. has already been seen to be low.

Limited time prevented investigation of the behaviour with a ceiling reflection of different delay; however, threshold results, as reported in section 5.3, suggest that behaviour is independent of ceiling delay.

10.3 DEGREE OF S.I. WITH MULTIPLE LATERAL REFLECTIONS

The effect of two lateral reflections was reported in reference |63| as being equivalent to a single lateral reflection of level equal to the sum of the incoherent lateral reflection levels. This result also conforms to the hypothesis, used in the previous section, that the degree of S.I. is a function of the ratio of lateral to non-lateral early energy. Two further short experiments also supported the hypothesis, one in which a relatively high level of lateral sound was used and in the other a low level. The subjective effect of four lateral reflections (two each side of the subject), each at a level of -9 dB relative to the direct sound, was found equivalent to a bilateral pair of reflections (summed incoherently) at -2.1 dB (±1.3 dB) relative to the direct sound. This is in good agreement with the hypothetical result of -3 dB. In the second experiment three lateral reflections (with delays of 15, 40 and 54 ms) were arranged on one side of the subject at an azimuth around $+40^{\circ}$.. Each was fed with a signal at -20 dB relative to the direct sound, a value very close to the average threshold value. Out of five subjects asked to judge the required level of a single lateral reflection to produce the same subjective effect, four gave a mean result of -15.3 dB (±0.8 dB) whilst the fifth found the situation below threshold. The hypothetical result for this experiment is -15.2 dB.

A further series of experiments was conducted in which the spatial effect of a pair of reflections at $\pm\alpha^{\circ}$ azimuth (and β° elevation) plus a pair at $\pm 90^{\circ}$ azimuth was compared with a single 90° pair. Five subjects performed the experiment, the Mozart motif being used as before. Figure 10.3

shows the echograms of the two fields used for comparison. The subjective

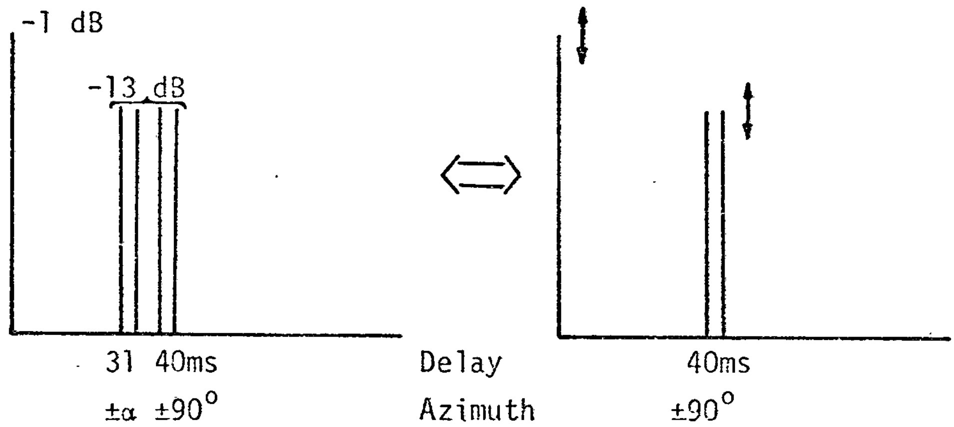


Figure 10.3. Echograms of sound fields for comparison in multiple lateral reflection experiment.

results are shown in Table 10.1. According to the hypothesis, the ratio of lateral to non-lateral early sound (S) is given by equation 10.1 in the general case:

$$\text{antilog } (S/10) = \frac{\sum \cos \phi . P}{\Delta + \sum(1 - \cos \phi)P} , \tag{10.1}$$

where, as before P and Δ are the reflection and direct sound levels. In this particular experiment P was the same for all reflections and so for the fixed sound field one has

$$\text{antilog } (S/10) = \frac{2P + 2P \cos \phi}{\Delta + 2P(1 - \cos \phi)}. \tag{10.2}$$

S is the predicted result for this experiment according to the hypothesis; it is tabulated in the last column of Table 10.1.

Measured and predicted results agree within experimental accuracy except for minor deviations for small angles of azimuth. This behaviour was not investigated further; since this minor deviation is in a positive direction (i.e., larger S.I. than anticipated), it was not considered sufficient to warrant revision of the hypothesis.

TABLE 10.1

Measured and predicted results of multiple lateral reflection experiment.

α^o	β^o	Measured mean ratio of lateral to non-lateral (dB)	95% conf. limit of mean (dB)	Predicted ratio of lateral to non-lateral (dB)
10	0	-7.6	0.7	-8.7
20	0	-7.1	0.3	-8.1
40	0	-6.8	1.0	-7.0
60	0	-6.4	0.8	-6.4
140	0	-6.3	1.0	-7.0
160	0	-7.3	0.8	-8.1
0	40	-5.9	0.7	-6.7
0	57	-7.3	1.4	-7.4
45	50	-6.8	1.0	-7.7

10.4 CONCLUSIONS

Experiments with multiple reflection sequences suggest that the subjective degree of spatial impression is related to the objective ratio of lateral to non-lateral early energy, measured in dB. A reflection contributes a proportion cos ϕ to the "lateral sound" and (1 - cos ϕ) to the "non-lateral sound".

Chapter 11

THE EFFECT ON SPATIAL IMPRESSION OF AUDIENCE ATTENUATION AT GRAZING INCIDENCE

11.1 THE FREQUENCY COMPONENTS OF SPATIAL IMPRESSION

The attenuation effect, reported in section 2.1(b), that occurs when sound passes over audience seating affects both the direct sound and reflections off vertical wall surfaces. It is likely to affect more than half the early reflections in a hall and should therefore be considered as regards its significance for spatial impression. However before describing experiments to investigate the degree of spatial impression that occurs with audience attenuated sound, it is necessary to consider the subjective effects of lateral reflections in different frequency bands.

If one listens to direct sound and lateral reflections, both octave filtered, it is apparent that different frequencies produce different subjective effects. This is summarised in Figure 11.1. At low frequencies there is a subjective impression of being surrounded by the sound, a sensa-

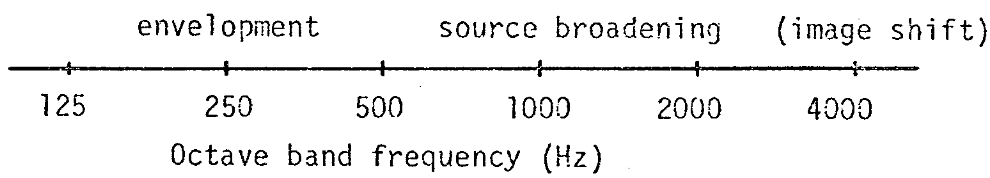


Figure 11.1. The subjective effects of lateral reflections in different octave bands.

tion well described by the word "envelopment". At these frequencies, the apparent area of the source is large, involving a degree of vertical spread as well as horizontal spread. (This observation agrees with one by Wettschureck |67|, that precision of localisation in the median plane is low with high frequencies absent.) At higher frequencies (around 1 kHz) the subjective impression is more one of source broadening; the broadening takes place very much in the plane of the loudspeakers. For frequencies in the 4 kHz octave, the broadening effect is greatly diminished, and for a single lateral reflection the effect is principally one of image shift away from straight ahead. Marshall |7| claims that this envelopment effect at low frequencies is highly desirable and Reichardt |12| also notes the importance of the bass component in lateral reflections. In the experiments reported in section 12.4 an attempt was made to discover the subjective significance of this low frequency lateral sound.

11.2 THE AUDIENCE ATTENUATION FILTER

A filter was built with a characteristic similar to the audience attenuation effect at grazing incidence. The circuit is contained in Appendix II. Figure 11.2 shows the filter response, whilst the dots are taken from measured results by Schultz and Watters |14| in three halls.

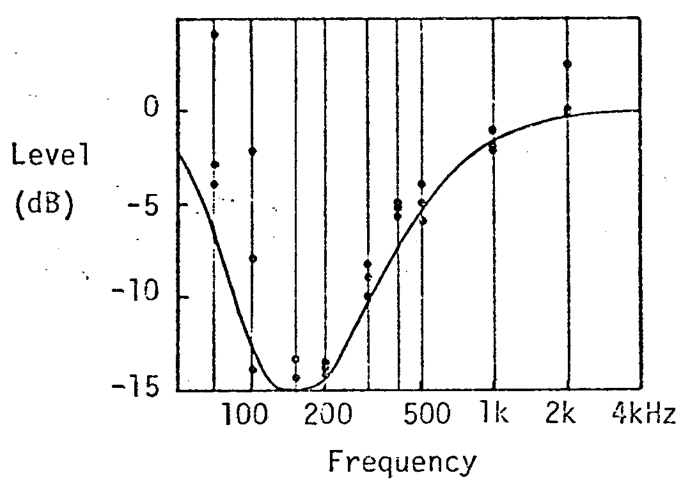


Figure 11.2. Filter response to simulate audience attenuation at grazing incidence (————); (•), measured results by Schultz and Watters |14|.

The subjective effect of this filter was similar to introducing severe bass cut, this can readily be appreciated by listening to Band A on the tape contained inside the rear cover of this thesis. The degree of filtering is certainly severe enough to be readily appreciated in a concert hall, if the total early sound had this response. In terms of loudness, however, the effect of introducing the filter was small, perhaps due to subjective compensation for the absent lower harmonics |68|. In objective terms, for the Mozart motif used for the majority of subjective experiments the effect of filtering was a mean level reduction of 5 dB (linear) or 2 dBA; the latter figure substantiates the subjectively small loudness change.

11.3 THRESHOLD EXPERIMENTS WITH AUDIENCE ATTENUATION FILTERING

To determine the effect of audience attenuation filtering the threshold experiment reported in section 5.3 was repeated with non-elevated sound audience filtered. The echogram for this experiment is shown in Figure 11.3(a). Delays for the lateral reflection ($\alpha = 40^{o}$) were varied from

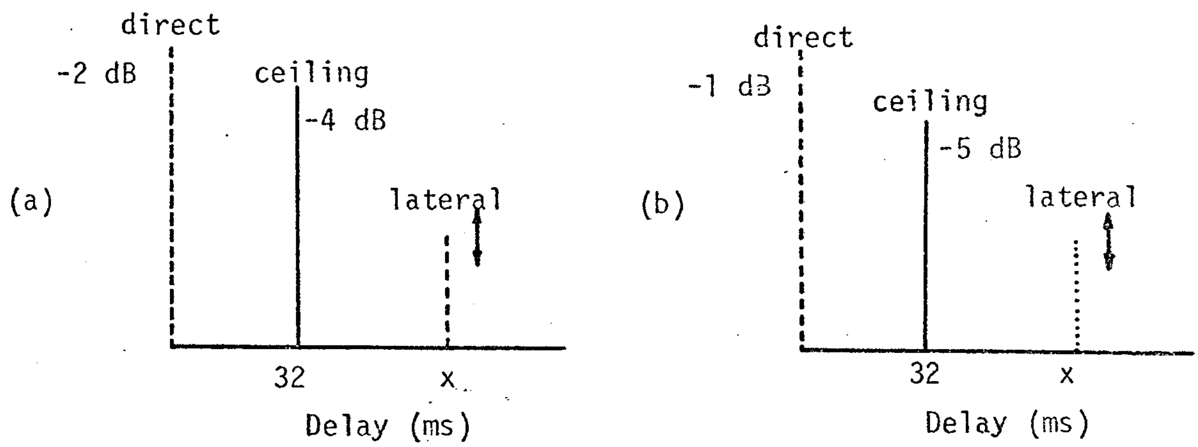


Figure 11.3. Echograms for experiments to determine the threshold of an audience filtered lateral reflection. (— — —), Audience filtered reflections; (··········), audience filtered and low pass filtered below 400 Hz.

10 to 70 ms, audience filtered sound is indicated in Figure 11.3 by dashed lines.

It soon became apparent that what the results of this experiment were indicating was the nature of what sound is masking what, rather than simply giving a figure for the threshold of audience filtered sound. There are at least four credible possibilities: the bass frontal sound masks the bass lateral, the total frontal sound masks the bass lateral, or the total frontal sound masks the total lateral or high frequency frontal sound masks high frequency lateral. The measured result is given in Figure 11.4, together with the result from Figure 5.2 for full frequency lateral and frontal sound. The threshold level for the filtered lateral reflection refers to the level the reflection <u>would</u> have <u>without</u> filtering, its actual level is about 2 dB less (dBA being taken as subjective level) whilst the level of the bass sound is about 10 dB less. It is immediately obvious that subjects were not using bass lateral sound to determine the

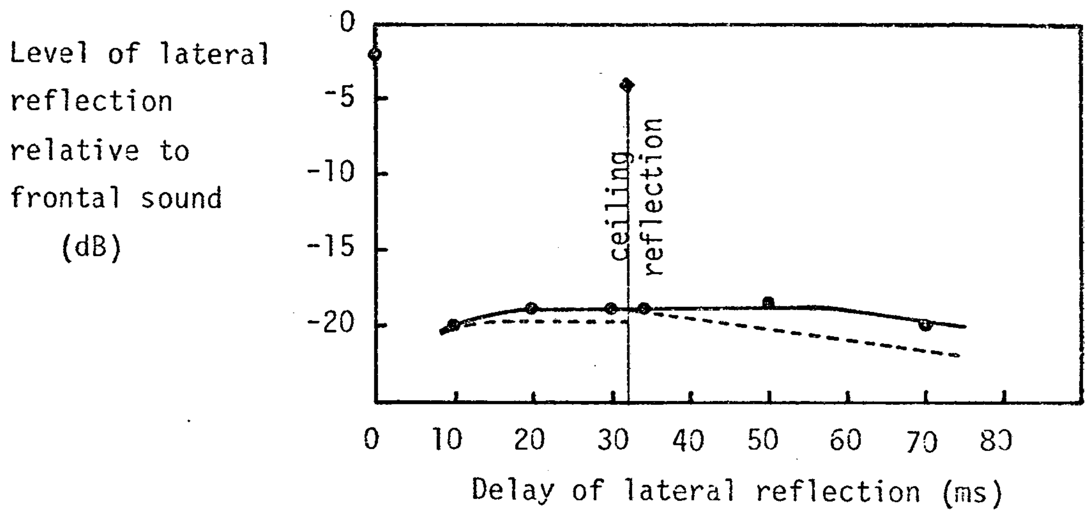


Figure 11.4. Threshold of audience filtered lateral sound in the presence of audience filtered direct sound and an unfiltered ceiling reflection, (———). (- - - -), Threshold with all sound unfiltered. Levels refer to unfiltered sound.

threshold. Not surprisingly they reported using the most intense component, the high frequency lateral sound, to determine the threshold; this leaves the question of whether the high frequency lateral sound is being masked by high frequency frontal sound or total frontal sound. Since, however, the masking of high frequency reflections is not of great interest to this thesis, it is pertinant to revert to the question of what is the threshold for the bass component of the lateral sound.

The experiment was therefore repeated with lateral sound, not only passing through an audience attenuation filter but also a low-pass filter excluding sound above 400 Hz (-3 dB at 400 Hz, 24 dB/octave roll-off). The echogram is shown in Figure 11.3(b), the change in reflection levels was for reasons immaterial to this discussion. A 5 second section of the Mozart motif was used, which contained a prominent 'cello line. The mean threshold levels for two subjects for this experiment are given in Figure 11.5. It can be seen that the variation with delay is again minimal, though the mean threshold level here is about 10 dB.

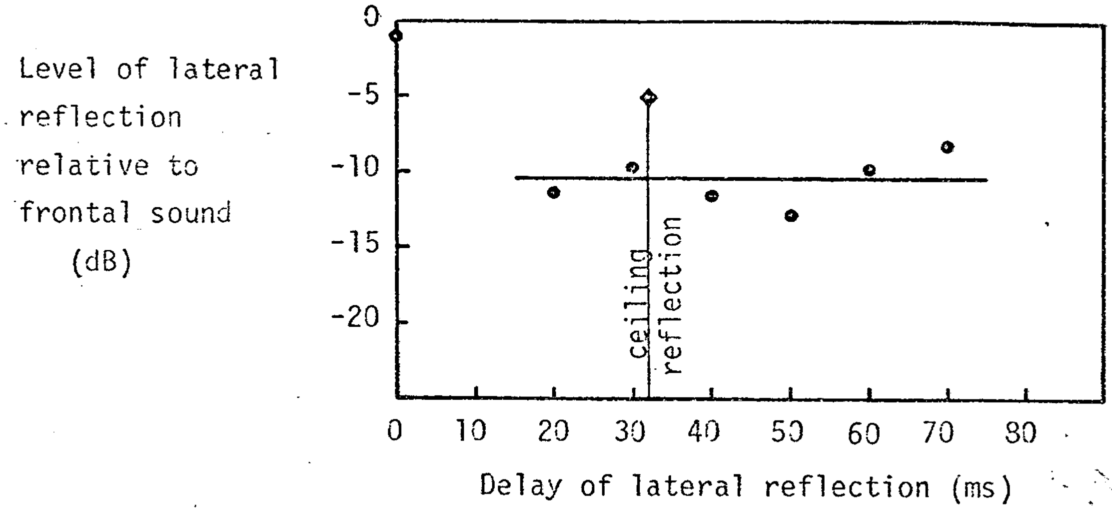


Figure 11.5. Threshold of low-pass filtered, audience filtered, lateral sound in the presence of audience filtered direct sound and an unfiltered ceiling reflection. All levels refer to equivalent levels for unfiltered sound.

If the actual reflection levels are considered, this apparently high threshold is explained. The actual direct sound level is (due to filtering) about 2 dB less than quoted, i.e., −3 dB, so that the total frontal sound level is (−3) + (−5) = −1 dB. The lateral sound level is about 10 dB less than quoted (since the mean level from Figure 11.2 below 400 Hz is −10 dB or less); thus the actual threshold level is −20 dB, which means a threshold level in terms of energy contained in reflections of −19 dB for the lateral sound relative to the frontal sound. This value is very similar to that quoted in Figure 11.4, when all sound components are unfiltered. There remain two possibilities: that the total frontal sound masks the bass lateral sound, or that the bass frontal sound masks the bass lateral sound. It is difficult from these results to determine which of these mechanisms is operative, though masking of bass sound by treble is obviously a possible mechanism, and, subjectively, whilst performing this experiment, one is very much aware of treble frontal sound.

In these experiments the direct sound and the first two reflections in a concert hall were simulated. Whilst in the second experiment subjects were measuring the desired threshold for bass lateral sound, the situation was also unrealistic since bass lateral sound does not in general occur in the absence of treble lateral sound, which in fact is normally more intense. It was pointed out in section 5.3 that not only does backward as well as forward masking occur for music, as it does for speech |60|, but also with music the delay of reflections is substantially irrelevant for thresholds. If as suggested in the previous paragraph, masking of bass lateral sound is by the total rather than just other bass sound, what is required is the threshold of a bass lateral reflection relative to the total early sound as a function of the treble frequency ratio of lateral to non-lateral early sound.

Unfortunately no measure of this threshold was taken and since this threshold may be critical for listening conditions in rooms, it is necessary to speculate on its probable value. For a treble-frequency ratio (of lateral to non-lateral early sound) of $-\infty$ dB the threshold was found above to be -19 dB. For a ratio of $+\infty$ (i.e., no frontal sound) lateral treble sound is masking lateral bass; since audience attenuation filtering does not completely mask bass sound, the threshold in this situation is less than the -10 dB that the audience filter attenuates bass sound, a probable threshold value being about -15 dB. The dependence of this bass lateral threshold on the spatial distribution of early treble sound is thus probably small; a mean value of -17 dB for the normal spatial distribution of early treble sound can be postulated, decreasing slightly for situations with little treble lateral sound.

11.4 SPATIAL IMPRESSION WITH AUDIENCE ATTENUATION FILTERING

With direct sound, one ceiling and two lateral reflections, but with all non-elevated sound audience filtered, the variation of spatial impression with lateral reflection delay was investigated. This was a repeat of the unfiltered experiment reported in section 10.2. The relevant echograms are contained in Figure 11.6. Seven subjects performed the experiment with the Mozart motif. Again levels in Figure 11.6 and also in Figure 11.7, which contains the results, are levels which the components would have had without filtering.

There are three possible ways in which subjects might have performed the experiments: to equalise the bass frequency effects, to equalise the

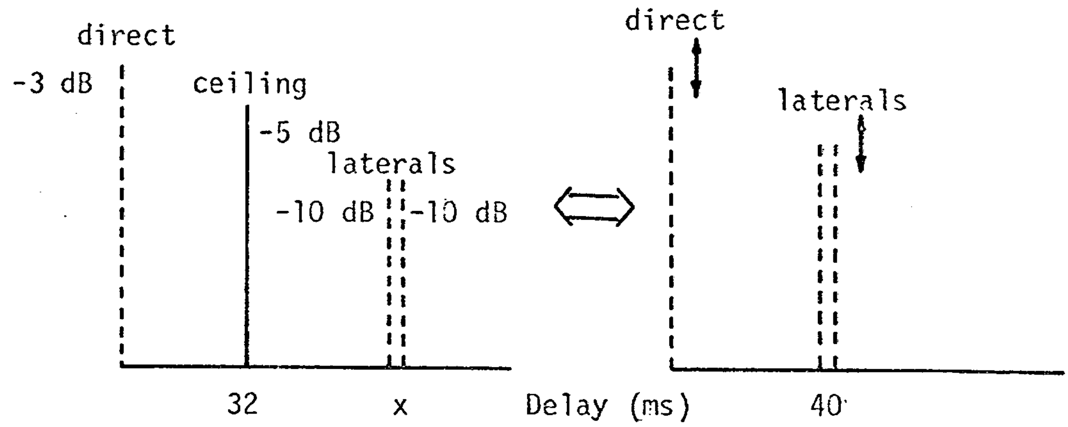


Figure 11.6. Echograms of sound fields for comparison in the spatial impression experiment with audience attenuation filtering; (----), audience attenuated reflections.

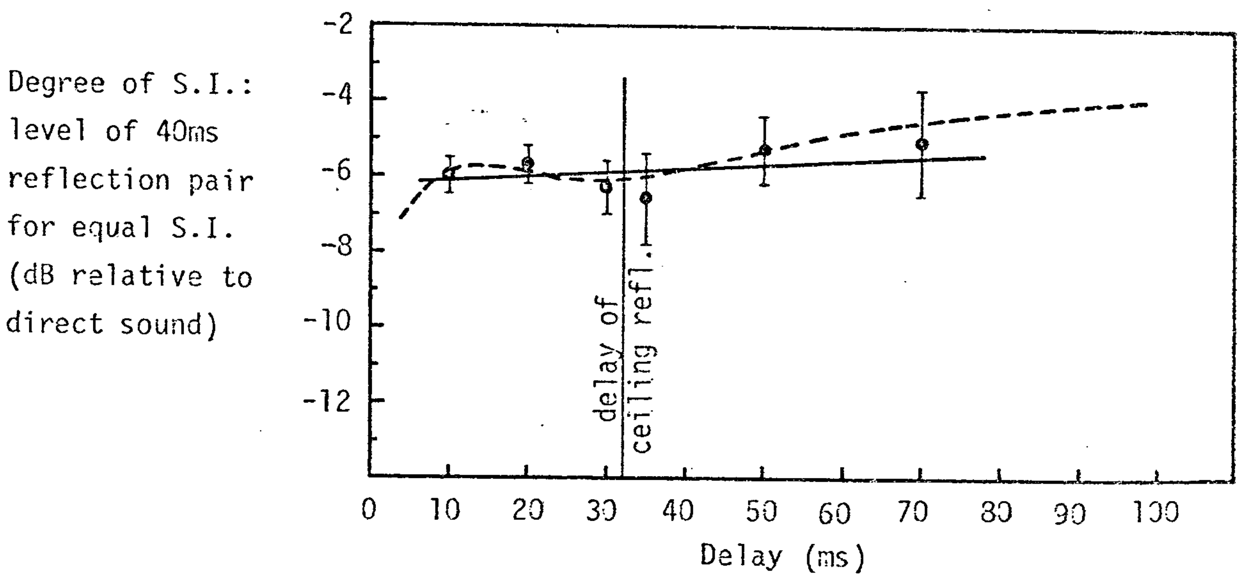


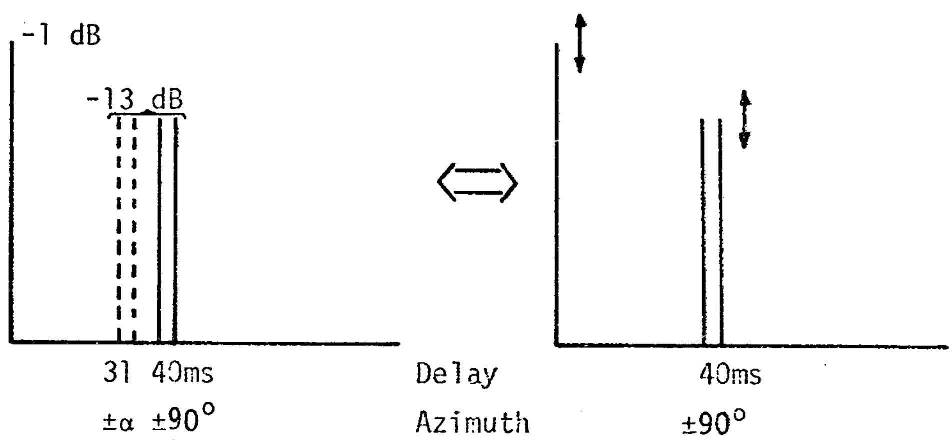
Figure 11.7. The degree of S.I. as a function of reflection delay for a lateral reflection pair of reflections with a ceiling reflection present. ɸ, Mean and 95% confidence limit of the mean with all non-elevated sound audience filtered; (----), result with no audience filtering from Figure 10.2.

spatial distribution in terms of energy irrespective of frequency, or to equalise the high frequency effects. With bass frequencies the ratio of bass lateral to non-lateral sound in the fixed sound field is about -13 dB, whilst the total energy ratio is about -8 dB. It is clearly not the bass frequencies which are being equalised here, and -8 dB is also just outside the 95% confidence limits. The similarity between results with and without filtering leads to the conclusion that subjects were equalising high frequencies, which is what subjects reported after performing the experiment. What is surprising, however, but perhaps also understandable if subjects were just concentrating on high frequencies, is the small scatter of results, considering the different spectral content in the two sound fields: bass sound was only present in the ceiling reflection in the fixed sound field. Subjects were told to ignore as far as possible spectral changes and evidently concentrating on one frequency band proved not to present much problem.

This experiment thus proved useless for elucidating the effect on the spatial impression of audience attenuation filtering, but it did illustrate the limitations on the use of the comparison technique for subjective experiments. Subjects will obviously aim for the most obvious equality between the two sound fields. There is, for instance, no point in instructing subjects to equalise bass spatial effects if in doing so they are making a more obvious effect out of balance. And what may be the most obvious effect in a limited simulation such as these may not be the most significant in the real concert hall situation. An attempt to use a bass frequency variable field, to force subjects to consider the bass frequency content of the fixed sound field, was soon abandoned since it was felt it would yield little information: not only does the comparison technique become more open to criticism the more different are the two sound fields, but also this particular experiment would yield data in a different subjective dimension to previous experiments, and thus not be comparable with previous results.

The most important question relevant to concert halls is whether there are two subjectively important effects, a bass frequency spatial effect and a treble frequency spatial effect, or one? Whether a listener in a concert hall appreciates spatial impression as one or two dimensions? The extreme situation for subjective operation in one dimension would be that a lack of bass lateral sound could be compensated by more treble lateral sound. This is obviously not realistic subjectively, at least not for an extreme lack of bass lateral sound. The conclusion of the previous experiment might be that subjects work in two subjective dimensions, but the fixed sound field in this experiment was slightly extreme in that there was significantly less lateral bass sound than is encountered in most concert halls. The experiment demonstrated that subjects are capable of operating in more than one dimension. However in concert halls Hawkes and Douglas |2| have found that listeners operate in four to six subjectively independent dimensions; that two of these are related to the proportion of lateral to non-lateral early sound would be to exaggerate the importance of lateral reflections. The compromise conclusion is that in general early lateral sound contributes to only one subjective dimension, to which different frequencies contribute to a different extent, bass frequencies being probably more important than treble frequencies. To quantify this weighting function would be difficult; a report on a simple subjective study to throw light on the significance of the different frequency components is contained in section 12.4. For such a study it is necessary to simulate as nearly as possible the real situation, with the inclusion of reverberation.

By the choice of the variable field in a comparison experiment, the subjective dimension in which subjects are to operate is defined. With a full frequency variable field and fixed sound field with almost full spectral content, subjects will treat spatial effects as a single dimension. Such an experiment was conducted as an addition to that described in Section 8.2. Figure 11.8 shows the echograms for this experiment. Equation (10.2) gives the expected result for the full frequency experiment. If one assumes that audience filtering attenuates the sound by a factor k for spatial impression, k can be calculated by comparing the measured result with the expected result. The calculated values of k are also shown in Figure 11.8 for the four angles of azimuth tested. For $\alpha = 20^{\circ}$ subjects on average measured a larger spatial impression with filtering than without, a strange result difficult to account for. For the other measurements the calculated value of k is about -2dB, which agrees with the assumed value used previously based on the dBA mean level change for the motif, when the filter is introduced. However, it should be added that the confidence limits for these values of k are large.



α^o	k (dB)
20	(+3.2)
40	-2.4
60	-2.3
140	-1.3

Figure 11.8. Echograms of sound fields for comparison and results of the experiment to determine the contribution of audience filtered lateral reflections to overall spatial impression. (- - - -), Audience filtered reflections.

11.5 CONCLUSIONS

The subjective effect of lateral reflections was found to be a function of their frequency content. However, comparison experiments presuppose the answer rather than answering the question of whether listeners in concert halls appreciate bass lateral sound as an experience distinct from treble lateral sound. Subjects in an experimental situation are definitely aware of the frequency content of the lateral sound, whilst on the basis of the results of Hawkes and Douglas |2| it appears unlikely that two dimensions related to early lateral sound are involved in one's overall acoustic judgement in a concert hall. If subjects are required to work in a single dimension, the subjective effect of audience filtering is only an attenuation of about 2 dB. This result imposes a maximum reduction on the subjectively relevant ratio of lateral to non-lateral sound of about 2 dB, when audience filtering is introduced. This value, however, seems too small to reflect such a gross filtering effect.

To clarify the role of bass frequency lateral sound, it is necessary to conduct experiments which do not involve any prior assumptions concerning the form of subjective judgement. Threshold and difference limen experiments are probably the first step in such an investigation. An exploratory experiment will be discussed in section 12.4 to determine the relative significance of the bass frequency content of lateral reflections. The threshold value for bass lateral sound was estimated as being about -17 dB relative to total other sound. In the absence of further information regarding the relative importance of bass and treble lateral sound, for the investigation of the physical situation in concert halls it appears best to treat the two as separate subjective entities, i.e., to consider both the bass frequency and mid-frequency ratio of lateral to non-lateral early sound.

Chapter 12

THE RELATION BETWEEN SPATIAL EFFECTS PRODUCED BY EARLY LATERAL REFLECTIONS AND REVERBERATION

12.1 INTRODUCTION

To satisfactorily describe in words the subjective distinction between the effect of reverberation and early lateral reflections is difficult, yet it is a distinction which can readily be appreciated by a subject alternately presented with a simulation of each. Both reverberation and lateral reflections produce a spatial effect, the combined effect of which is to provide a listener with a feeling of being in a three-dimensional space. But whilst lateral reflections produce a spread of the apparent source, reverberation has no particular direction associated with it (assuming it is diffuse); it provides an environment for the early sound, removing the starkness of basically anechoic music. Reverberation is traditionally the most obvious acoustic quality associated with a room, and a high reverberant level relative to the early sound gives the impression of an apparently distant source, whilst a low reverberant level gives an impression of source proximity. Bekesy has also reported this quality |69|. Lateral reflections provide a sense of involvement with the performance which reverberation does not. The tape contained inside the back cover of this thesis gives a good representation of the subjective effects of reverberation and early lateral reflections.

Whilst reverberation and lateral reflections produce different subjective effects, this is not to say that a listener in a concert hall associates the reverberant sound with one subjective dimension and the lateral sound with another. The reverberation is probably related to 'reverberance', and lateral reflections may also contribute to 'warmth' and 'intimacy', but a low ratio of reverberant to early sound might also contribute to 'intimacy', and both factors contribute to 'clarity'. In terms of one's subjective impression of a room (i.e., the spatial effects), both factors are likely to contribute; indeed Marshall has claimed |6| that the subjective effect of early lateral reflections has previously been assigned to reverberation. It is safe to say that there is definitely some overlap between the two effects, but to resolve the relative importance of reverberation and lateral reflections requires elaborate subjective studies.

By performing comparison experiments one can force, or attempt to force, subjects to consider the spatial effects of reverberation and lateral reflections as equivalent and in this way determine a relationship between the two. This was done by Reichardt |12|, when he asked subjects to measure the 'Raumlichkeit' (room impression) for a series of reflection sequences including reverberation. This experiment will be discussed in the next section, more for the lessons to be learnt from it than the actual results achieved. Similar experiments were also conducted in Perth, but it soon became apparent that relatively high levels of lateral sound could not be compensated by reverberation, and that the effects of lateral reflections were too dissimilar for the experiment to be satisfactory. If the effects of lateral reflections and reverberation are distinct, it is necessary to consider what additional contribution is made by lateral reflections. This is discussed in section 12.4, whilst section 12.5 contains a discussion of the possibility of a lack of bass lateral early sound being compensated by boosted bass reverberant sound.

12.2 REICHARDT'S 'RAUMLICHKEIT' EXPERIMENT

Using direct sound and a simulated reverberant field, Reichardt and Schmidt |35| have derived a relationship between the ratio of reverberant to direct sound and a subjective scale of 'Raumlichkeit' (room impression) by measuring difference limen. The ratio was called the 'Hallabstand' H (reverberant separation) though here it is more convenient to use its inverse, H' (= 1/H):

$$H' = \frac{\text{Reverberant level}}{\text{Direct sound level}}.$$

It was found that the room impression (R.I.) was not only a function of H' but also of the reverberation time: a fall of H' of 4 dB is equivalent to a fall of R.T. from 2.0 to 1.7 secs. Beranek and Schultz |31|, and others, have considered the ratio of reverberant to early sound (R) to be more significant in a real sound field:

$$R = \frac{\text{Reverberant level}}{\text{Early sound level}}.$$

The early sound is generally considered to be that which arrives within 50 ms of the direct sound, whilst the reverberant sound is all that arrives after 50 ms. There is, however, no published evidence concerning the validity of this measure.

Reichardt |12| questions which of these measures is more relevant in the real situation: "is the contribution of reflections within the first 50 ms more important for an increase of room impression or transparency (clarity)?" (Reichardt- author's translation. Note the implicit assumption that room impression and clarity are mutually exclusive.) Reichardt also reports in the same paper on an experiment in which a simulated lateral reflection was used, which indicated that lateral reflections contribute to room impression. He then proceeds to investigate which of the two measures H' or R is more valid as a measure correlatable with subjective room impression; the method for these experiments is to compare a reflection sequence containing direct sound, a ceiling and a side wall reflection and reverberation with one containing direct sound and reverberation. For the particular reflection sequences tried, it is apparent that R is more suitable for reverberant situations, whilst H' is more suitable for unreverberant situations. The validity of this result is however limited to the range of situations investigated. It is easy, however, to conceive a situation in which neither H' nor R are correlatable with room impression: if the ratio of lateral to frontal early sound (excluding direct sound) is varied whilst the total early sound level is held constant, this will (as Reichardt himself has noted) alter the degree of room impression, though it will not affect either H' or R, since neither measure contains any directional factor. What is obviously required is a measure which takes account of the spatial effects of lateral reflections as well as that of reverberation. Reichardt's results provide material which can be used to gain an impression of the form such a measure might have.

Reichardt's experiment involved comparing two types of sound field, shown in Figure 12.1, and denoted as H and R-experiments. For comparison a field with just direct sound and reverberation was used, the ratio between them will be referred to as R_c (= reverberant/direct sound level). The subjective relationship between R_c and H' for the H-experiment, and R_c and R in the R-experiment as obtained by Reichardt is given in Figure 12.2 (derived directly from Bild 4 in reference |12|). The 45° line is also included; for the regions over which the H' or R lines correspond with the 45° line, the particular measure is a correlate of subjective measurements. As can be seen, this occurs for R large and H' small. However, for R large the influence of the early reflections is minimal compared with reverberation, whilst for H' small the early reflections are nearly masked by the direct sound. For this experiment there was a particularly poor choice of test reflection sequence

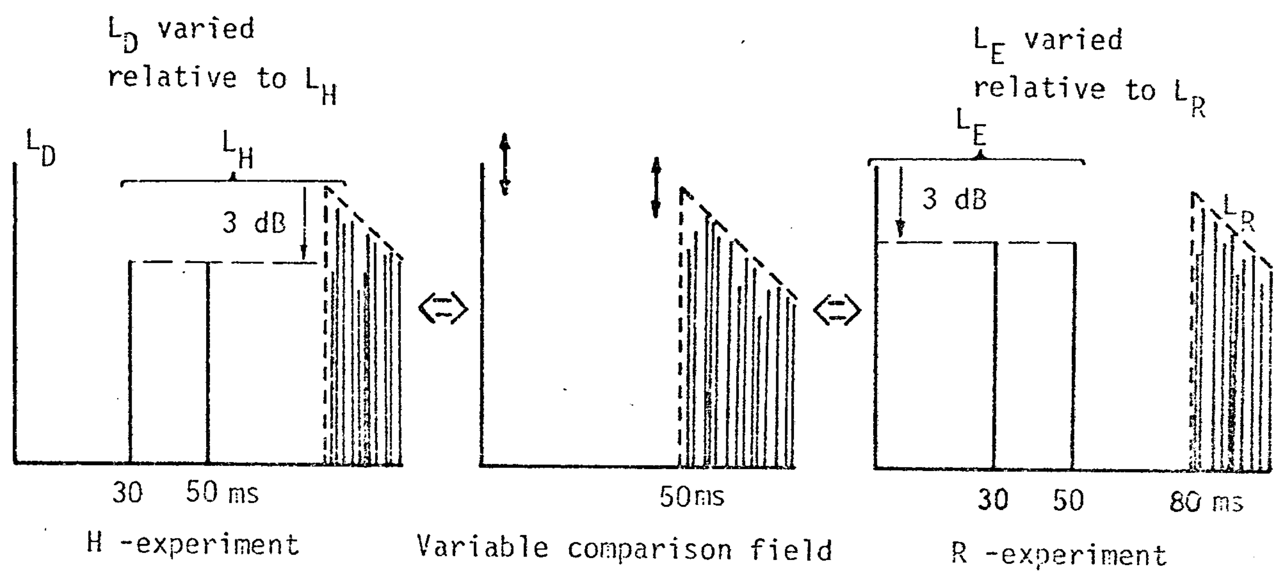


Figure 12.1. Echograms used in Reichardt's 'Raumlichkeit' experiment |12|.

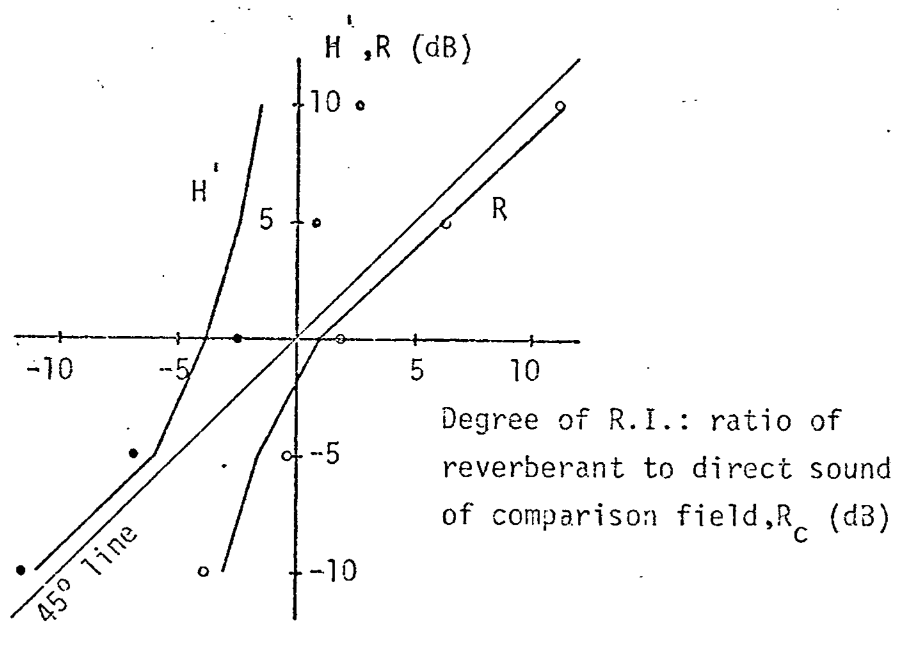


Figure 12.2 Room impression of the H' and R fields after Reichardt |12|. (———), Reichardt's subjective results. •, Predicted results for H'-experiment; ∘, predicted results for R-experiment, both according to equation (12.1).

since in each case there is an equal level side reflection and ceiling reflection. If, as Reichardt suggests, the lateral reflections contribute to room impression, adding the two early reflections is not likely to give a significant change in room impression, particularly for the R-experiment with $R > 0$ dB, which is what appears in practice, as in Figure 12.2.

In the search for a physical quantity which might predict the results of Figure 12.2, one credible possibility does not fit the results at all well. Starting from the ratio of lateral to non-lateral early sound, it is possible that the overall ratio of lateral to non-lateral sound determines the degree of room impression. As will be shown later (see section 19.7), with the auditory "directional characteristic" for spatial impression, a diffuse field has a ratio of lateral to non-lateral sound of 0 dB: i.e., the reverberant energy should be divided equally between the lateral and non-lateral sound components. As mentioned above, however, the predicted results according to this measure agree very poorly with subjectively measured results after Reichardt.

Analysis described in the next chapter on the premise that a cross-correlation process is involved suggests the following measure:

$$R_i = \frac{\text{Reverberant level + Early lateral sound}}{\text{Total sound level}} \qquad (12.1)$$

$$= \text{Reverberant energy fraction + Early lateral energy fraction.}$$

If R_i is calculated for each of the sound fields used by Reichardt, and the ratio of reverberant to direct (R_c) for the comparison field is also calculated, which produces the same value of R_i, the points included in Figure 12.2 are obtained: one series of points for each experiment. Agreement is good except for H' large. H' = +10 dB is however an unnatural condition with direct sound dominated by early reflections and reverberation. But without further details of Reichardt's results (and their precision) elaboration of the measure R_i in equation (12.1) is not warranted to attempt to fit results better.

Experiments could no doubt be validly conducted to determine first of all the validity of R, the ratio of reverberant to early sound, as a measure of room impression with only frontal early sound and reverberation. The contribution of early lateral sound to room impression could then be determined. However, attempts by the author to perform such experiments proved abortive since subjects were unable to compensate for the effects of lateral reflections with reverberation. This raises the question of how Reichardt managed to perform this experiment; one significant difference is that

Reichardt only used a single lateral reflection rather than lateral reflections on both sides. This is in fact a serious criticism of Reichardt's experiment. The author's experiments of the same nature will be reported in the following section.

12.3 COMPARISON EXPERIMENTS FOR 'ROOM IMPRESSION'

A short series of comparison experiments were undertaken to determine the contribution of reverberation and lateral reflections to room impression. For the first two experiments the variable field contained only direct sound and lateral reflections, whilst for the later experiments it contained only direct sound and reverberation (as was used by Reichardt). For the first

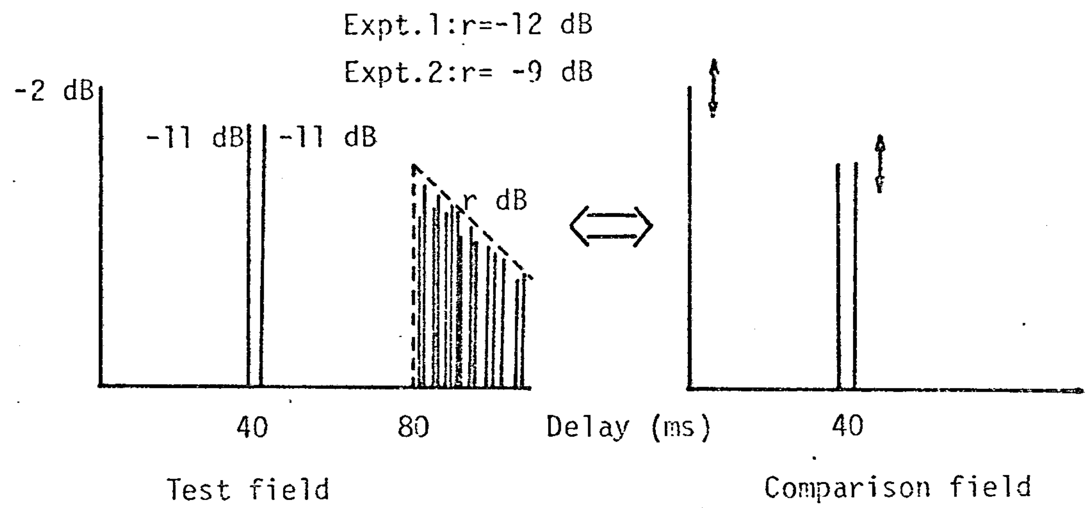


Figure 12.3. Echograms of sound field for comparison in experiments 1 and 2.

two experiments, different values of r, the stationary reverberant level, were used, as shown in Figure 12.3. The subjectively measured result is in terms of the ratio of lateral to non-lateral early sound, S. Subjects were told to equalise the sense of being surrounded by sound (as opposed to source breadth for spatial impression); if this presented difficulties, which it appeared not to do, subjects were advised to equalise their awareness of the direct sound. Again the scatter of results was lower the more similar were the two sound fields for comparison. The Mozart motif was again used, five subjects performed the experiment. The results are given in Table 12.1. The fourth and fifth columns contain the value of the quantity R_i, as defined by equation (12.1), for the test (fixed) and comparison (variable) fields. In each case the agreement is good, so for small amounts of reverberation, the quantity R_i corresponds with subjective measures.

In the third experiment the test field in Figure 12.3 (with r = -12 dB) was compared with the comparison field used by Reichardt, as shown in Figure 12.1. This comparison field contains only direct sound and reverberation, the subjectively measured result is in terms of the ratio of reverberant to early energy, R. The result for five subjects is also included in Table 12.1. Again agreement using the measure R_i is within the 95% confidence limit for the measured results.

TABLE 12.1

Comparison experiment results with lateral reflections and reverberation.

Expt. No.	Reverberant level (dB) and motif	Subjectively comparable comparison field	95% conf. limit (dB)	R_i of comparison field (dB)	R_i of test field (dB)
1	r = -12 Mozart	S = -4.1 dB	±0.7	-5.5	-5.9
2	r = -9 Mozart	S = -2.9 dB	±1.1	-4.7	-5.1
3	r = -12 Mozart	R = -6.4 dB	±2.0	-7.3	-5.9
4	(see Figure 12.4) Wagner	R > +4 dB		>-1.5	-1.7

With again the comparison field consisting of only reverberation and direct sound, the test field in Figure 12.4 was used. Subjects were not able to give a precise result for this experiment and complained that the sound fields were too different for a meaningful comparison to be made. Each subject gave a range of values over which rough equivalence existed, the results ranging from R = +4 dB to +10 dB. The measure R_i predicts a value at the bottom of this range.

The test field used in this last experiment, given in Figure 12.4, corresponds with a typical concert hall situation, at least as far as the ratio of lateral to non-lateral early sound and the ratio of reverberant to early sound are concerned. The result of this last experiment suggests that the effects of lateral reflections are distinct from those of reverberation, each contributing significantly to the total spatial effect. The subjective effects of lateral reflections in a reverberant field will be discussed in the next section.

For relatively low levels of lateral early sound and reverberant sound, subjective comparison between the effects of each is possible; the quantity in equation (12.1) appears to correlate well with results of subjective comparison experiments for "room impression" for such sound fields.

12.4 THE SUBJECTIVE EFFECTS OF LATERAL REFLECTIONS IN A REVERBERANT FIELD

In section 4.2(e) the subjective effect of adding lateral reflections to direct sound was described as source broadening, a three-dimensional sense, etc. To determine the value of lateral reflections in concert halls what is required is a description of the subjective effect of adding lateral reflections to reverberant sound. Readers can judge this for themselves by listening to Band D of the tape contained in the rear of this thesis. Five subjects were presented with the echogram in Figure 12.4 and were able to switch the lateral reflections on and off. The reflection and reverberation

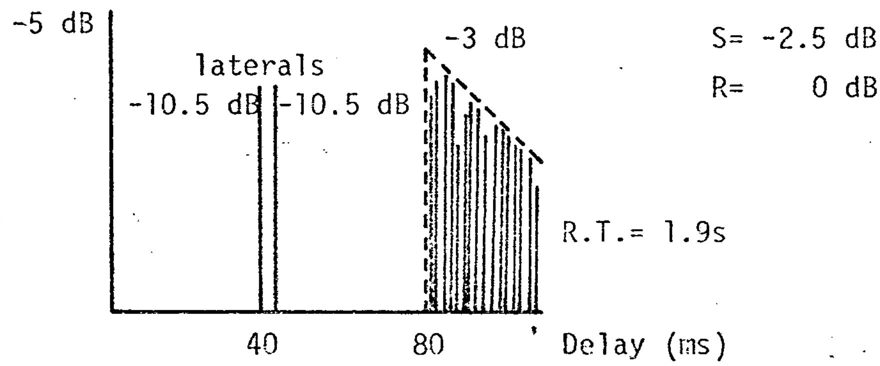


Figure 12.4. Echogram of sound field used to investigate subjective effects of lateral reflections in a reverberant field.

levels were chosen to be close to typical values in concert halls: the ratio of lateral to non-lateral sound was S = -2.5 dB, whilst the ratio of reverberant to early sound was R = 0 dB. The Wagner motif was used. As well as being asked to describe the effect of addition of lateral reflections, subjects were also asked to choose two parameters which best described the effect. To make this choice they were provided with a list of the following possible scales (since subjects were later to be asked to operate on two additional named scales: spatial effect and clarity, these were excluded from the list): warmth, source distance, envelopment, bass sound, reverberance, intimacy, bass envelopment. Choices not contained in the list were also allowed. Out of five subjects, three chose 'envelopment' whilst two chose 'bass envelopment'. As a second dimension, two chose 'bass sound', two 'source distance' and one 'intimacy'. The descriptions complemented the subjects' respective choices; all subjects, either in the choice of parameter or in their description, commented on the bass frequencies being enhanced, or the spatial effect being predominantly a bass effect. There was frequent mention of feeling closer to the source, and feeling a greater sense of involvement when the lateral reflections were present. Two subjects also commented that one could better appreciate the contribution of individual instruments with lateral reflections.

Two aspects are clear from this experiment: that lateral reflections cause a sensation of envelopment, and that the effect is predominantly associated with bass frequencies (although the reverberation and direct sound are both full frequency). Subjects were then asked to rate on line scales with a single centre mark the change that occurred when lateral reflections were introduced. They were asked to use their two chosen scales and the scales 'spatial effect' and 'clarity'. Subjects were presented with four lateral reflection situations:

(1) unfiltered lateral reflections
(2) audience filtered lateral reflections
(3) low-pass filtered lateral reflections (below 400 Hz)
(4) audience filtered lateral reflections with bass-boosted reverberation.

It is apparent from the mean results that subjects were scoring on the 'spatial effect' and 'envelopment' scales as a single dimension. The mean scores for these four situations are shown in Figure 12.5.

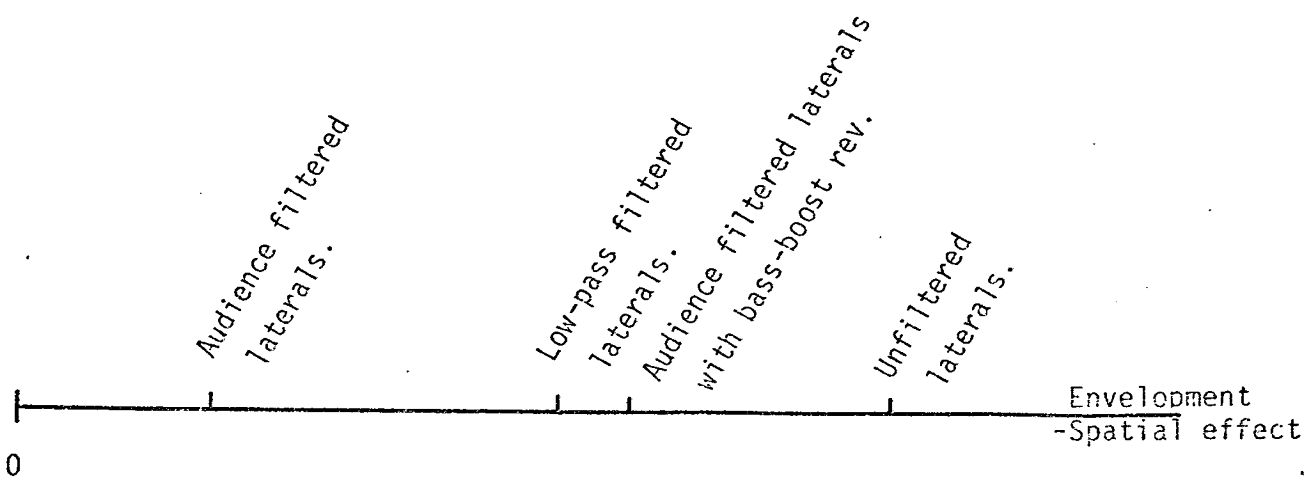


Figure 12.5. Mean subjective scores on the Envelopment/Spatial Effect scales for the addition of filtered and unfiltered lateral reflections in a reverberant field.

The superiority of unfiltered laterals is evident. Since the audience filter and low-pass filter are roughly complementary, it is not surprising that the sum of the scores for each of these situations approximates to that of the unfiltered situation. The mean score with audience filtered laterals was 23%, whilst for the low-pass filtered situation it was 63% of the score for the unfiltered situation. If subjects scores for unfiltered lateral reflections were normalised, the 95% confidence limit for the mean with filtered lateral reflections was ±20%, i.e., the mean (23% ± 20%) is statistically below 50%. The fourth situation with bass-boosted reverberation will be discussed in the next section.

Scores on the clarity scale were generally haphazard, explainable perhaps since the ratio of reverberant to early sound was maintained constant on switch over. Information on other scales was insufficient to gain a valuable picture of behaviour.

This study can be criticised for the small number of subjects used and for not allowing subjects completely unbiased judgement. However, the consistency with which subjects judged the effect of lateral reflections to be enveloping and strongly associated with bass frequencies is persuasive. The bass component of the lateral reflections accounts for the majority of the spatial effect, though the high frequency component also contributes.

12.5 BASS LATERAL SOUND AND BASS REVERBERATION

Schultz and Watters |14| suggest that the lack of bass sound which occurs on the main floor of concert halls due to audience attenuation filtering is compensated by bass sound in the reverberation. One can speculate that increased R.T. at bass frequencies, which is frequently sought in concert halls, might compensate for a lack of bass early lateral sound.

To determine whether increasing the bass reverberant energy could compensate for a lack of bass lateral sound, subjects were asked to rate on the four subjective scales, described in the previous section, the situation with audience filtered lateral reflections and bass-boosted reverberation. The reverberant signal was passed through a high-fidelity tone control circuit and the bass control adjusted to give a mean increase in level of 5 dB, in the frequency region 100-400 Hz. With this, admittedly crude, form of compensation, subjects rated the degree of envelopment/spatial effect as considerably increased relative to the situation with unboosted reverberation, as can be seen from Figure 12.5, though the degree of envelopment is nevertheless inferior to that with full frequency lateral sound.

A further experiment was also conducted by six subjects, with the Wagner motif, in which they were asked to compare the sound field in Figure 12.4 with the same field but with the lateral reflections audience filtered and the reverberant sound bass-boosted. Subjects were asked to determine at what level of bass boost the two sound fields sounded most similar. One subject considered the two sound fields too dissimilar for a meaningful balance to be made, whilst the others detected a difference in quality but on average found that a 5 dB bass-boost in the reverberation compensated for the lack of bass-lateral sound. Subjects generally preferred the situation with unfiltered lateral reflections, and found that bass-boosting the reverberation involved a deterioration of clarity.

The situation with audience filtered laterals is extreme relative to most real halls, though cases will be discussed in Part II of halls with areas with no early bass lateral sound. The experiments described in this section were of an exploratory nature, they do point, however, to two conclusions: that a lack of bass early lateral sound can be in part compensated by an elevated bass reverberant sound level, but that in terms of a higher reverberation time at low frequencies, the increase needs to be substantial. To obtain a 3 dB increase in the reverberant energy a doubling of the reverberation time is required. Thus any rise in the reverberation time at bass frequencies of less than 50% above the mid-frequency R.T. is unlikely to be significant from this point of view. Band E on the tape at the back of this thesis enables readers to experience the comparisons discussed in this section, though it should be pointed out that one's discrimination in the real simulation is higher than this recording.

12.6 CONCLUSIONS

Examination of Reichardt's experiment |12| indicated that <u>if</u> a single measure exists which correlates with the spatial effects of both lateral reflections and reverberation neither of the quantities considered by Reichardt can be expected to be useful. A measure is required which contains terms corresponding to both the reverberant energy and the lateral early sound energy. Such a measure, which predicts results in reasonable agreement with Reichardt's, is given in equation (12.1). This measure was also found to agree well with measured results using a low level of lateral and reverberant sound.

However, when using a field corresponding to the real situation in a concert hall, the effects of lateral reflections and reverberation were found to be too dissimilar for subjective comparison. Subjects, when asked to describe the effects of lateral reflections in a reverberant sound field, described the effect by the word 'envelopment' and considered the effect predominantly associated with bass frequencies (although all sound components were full frequency). Whilst it was shown that the high frequency lateral reflections do contribute to a spatial effect, the bass frequencies account for the majority of the effect.

Since much of the early sound in concert halls is lacking in bass sound due to audience attenution filtering, the possibility was investigated that a lack of bass early lateral sound can be compensated by additional bass reverberant energy. This was found to be possible though increases in the bass R.T. of at least 50% appear to be necessary. It is interesting to note that using the same reflection sequence with reverberation, given in Figure 12.4, subjects were able to compensate for a lack of bass early lateral sound with bass reverberant sound, whilst they were not able to compensate for full frequency early lateral sound with reverberant sound. This demonstrates the relative temporal insensitivity of the hearing system to bass sound. It should be noted, however, that the degree of temporal sensitivity depends on the signal used: changes are more readily detected in fast (e.g., the Mozart motif) than in slow motifs (e.g., the Wagner motif). It will be noted that certain experiments aimed at determining possible relationships between early and late sound were conducted with the Wagner motif for this reason.

A mention was made in section 12.2 of a comment by Reichardt |12| suggesting that clarity and room impression are mutually exclusive. This would seem to be erroneous. The measure, R, the ratio of reverberant to early sound, is generally considered to be a correlate of subjective clarity (though this has never been proven). Reichardt however considers it as a correlate of room impression, and discovers it not to be universally applicable. The particular virtue of lateral early reflections would seem to be that they contribute both to clarity and to producing a desirable spatial effect. This is one more example of the danger of considering subjective assessment of the acoustical situation in a concert hall as a uni-dimensional process.

Chapter 13

SPATIAL IMPRESSION AS AN AUDITORY CROSS-CORRELATION PROCESS

13.1 INTRODUCTION

It is now well accepted that localisation in the horizontal plane, and judgement of diffuseness, both involve a cross-correlation process between the signals at the two ears. The ability to concentrate on a signal from a particular direction to the exclusion of other (uncorrelated) signals (the cocktail party effect) probably also involves a cross-correlation procedure. The spatial effect of lateral reflections, which has been the subject of study in previous chapters, may also be related to a cross-correlation process. Since in each case one is interpreting a different aspect of the same cross-correlation function, it is instructive to review the present theories concerning localisation and judgement of diffuseness related to the cross-correlation between the signals at the two ears.

13.2 LOCALISATION IN THE HORIZONTAL PLANE

Cherry and Sayers have conducted a series of experiments to determine the relationship between signals presented over earphones and perceived direction. These are well summarised in reference |1|. With earphones, only lateralisation is possible (judgement of left, right or centre); this appears to be related to the delay (τ) at which a maximum occurs in the cross-correlation function between the signals at the two ears. For τ positive one lateralises on one side, for τ negative on the other side. If more than one maximum exists (as will occur for sine wave signals) the relevant maximum is the one which occurs in the interval $-630\mu s < \tau < 630\mu s$, where $630\mu s$ is the maximum interaural time delay (corresponding to a path length difference of 21.6 cm), which occurs for 90^o lateral sound.

The general expression for the cross-correlation function between two signals, $f_\ell(t)$ and $f_r(t)$, in the left and right ears, respectively, is

$$\phi_{\ell r}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} f_\ell(t) f_r(t+\tau).dt. \tag{13.1}$$

Cherry and Sayers point out that for a finite energy signal this tends to zero. Evidently the ears perform a running correlation; the following form has been suggested by Atal, Schroeder and Kuttruff |47|:

$$\phi^{s}_{\ell r}(\tau) = \phi_{\ell r}(\tau).e(\tau),$$

where $e(\tau)$ is the autocorrelation function of a weighting function $r(t)$, which is zero for positive time (future) and has decreasing values for negative time (past). The function $r(t)$ was determined by Atal et al. for the hypothesised autocorrelation process in the ear, which enables colouration to be detected (see section 4.2(c)). The weighting function for localisation is evidently related to measures of the inertia of auditory localisation by Blauert |70|, who found that signals presented concurrently at alternate ears could not be distinguished from each other if their duration was less than 200 ms. Work of Tobias and Zerlin |71| has shown that for a variable duration burst of noise, the interaural time difference threshold is a minimum for durations greater than 700 ms. The weighting function $r(t)$ for localisation is thus zero for positive time and decreases down to zero for t negative, being zero for $t < -700$ ms. This weighting function need not be considered here, since for such long time periods the effect on the correlation function for delays less than 630 μs will be insignificant.

Sound presented to a listener is located outside the head due to transformation of the sound by the pinna and head; direction in azimuth can be determined with great precision. At low frequencies it is well established (e.g., in reference |72|) that the localisation is on the basis of phase/time differences. It is generally assumed that this is performed by calculating the interaural time difference. Figure 13.1 shows measured interaural time differences (ITD) for clicks as a function of azimuth (α) reproduced from Feddersen et al. |73|. The dotted curve shows the calculated ITD according to equation (13.2):

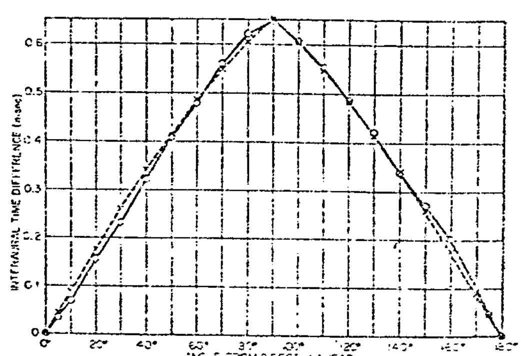


Figure 13.1. The measured (——) and calculated (- - - -, according to equation (13.2)) interaural time difference as a function of azimuth (after Feddersen et al. |73|).

$$ITD = \Delta t = \frac{r}{c}(\alpha + \sin\alpha), \tag{13.2}$$

where r is the radius of the head and α is here measured in radians. This equation can be derived by assuming that the incident wavefront travels round the surface of the head in the shadow region. Cross-correlation between the ear signals is a mechanism by which the ITD can be determined, as given by the delay at the correlation maximum. The autocorrelation function for a rectangular band of noise is given by

$$\Phi(\tau) = \frac{\sin\pi.\Delta f.\tau}{\pi.\Delta f.\tau} . \cos 2\pi f\tau,$$

where f is the centre frequency and Δf is the bandwidth. A sine wave signal thus has a "cosine wave" cross-correlation function. For frequencies below 750 Hz only one correlation maximum exists in the interval $|\tau| < 630\mu s$, irrespective of angle of azimuth, (see Figure 13.2). For 750 Hz < f < 1500 Hz,

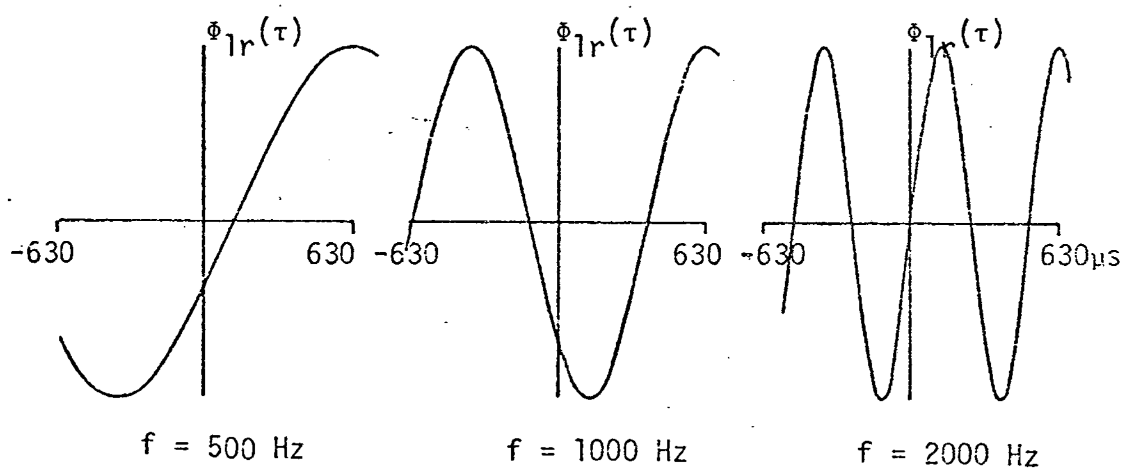


Figure 13.2. Interaural cross-correlation functions for three sine wave signals at three frequencies for directly lateral sound.

there exist two correlation maxima in the interval $|\tau| < 630\mu s$ for large angles of azimuth. One assumes that interaural amplitude differences allow one to determine whether the sound is left or right of straight ahead. For $f > 1500$ Hz more than one maximum lies on each side of $\tau = 0$, so localisation by cross-correlation no longer remains possible. It is therefore not surprising that for frequencies above about 1500 Hz localisation is determined not by phase but by interaural intensity differences. Various experiments reported in reference |72| confirm this result. Figure 13.3, taken from Mills |74|, indicates that precision of localisation is poorest around 1500 Hz, in particular for more lateral source directions; this is

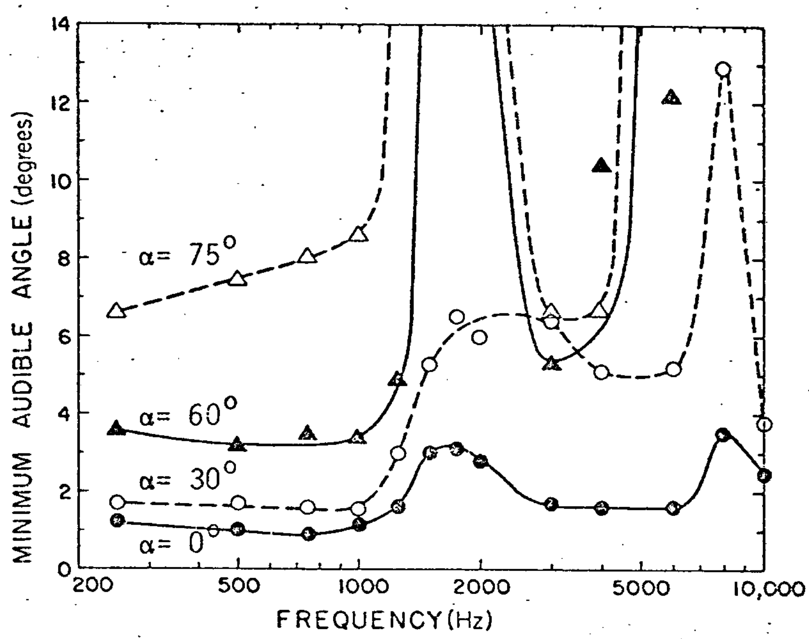


Figure 13.3. Minimum audible angle between successive pulses of tone as a function of tone frequency and direction of the source (after Mills |74|).

in complete agreement with correlation prediction. As the frequency increases above 1500 Hz so the intensity variation with azimuth increases; thus precision of localisation also exists at these frequencies. It is interesting to note that the maximum frequency at which individual fibres of the auditory nerve can fire in synchronism with the stimulating tone is also about 1500 Hz!

It can also be noted in Figure 13.3 that the precision of localisation is almost independent of frequency below 500 Hz. Such behaviour extends at least to 100 Hz (Stevens and Newman |75|). In terms of the correlation function, the "sharpness" of the maximum decreases with frequency, suggesting a decrease in precision of localisation at lower frequencies. However, correlation in the auditory system must occur at the neural level; with individual nerve fibres firing in synchronism with the stimulating signal at these frequencies (e.g. Sayers and Lynn |76|), the correlation process is made more sensitivie due to a "neural sharpening effect". Thus only the delays of maxima and minima and not the precise shape are relevant for auditory cross-correlation functions.

13.3 THE JUDGEMENT OF DIFFUSE SOUND FIELDS

A considerable amount of work has been applied to achieving simulated diffuse fields with a small number of loudspeakers. To achieve this the first requirement is two or more mutually incoherent signals; since these diffuse fields were required for reverberated sounds, two methods have been used: temporal incoherence and spatial incoherence. For temporal incoherence, the reverberated signal is subjected to a time delay and it is found for noise reverberated in a reverberation chamber that the original signal and its delayed version are mutually incoherent for delays >1 ms, |44|. This result is particularly relevant for the discussion in the next section, since what is normally called a "coherent" reflection is for the ear an incoherent signal (as long as its delay is greater than 1 ms). In general this method of achieving incoherent signals has not been used due to the undesirable tone colouration that results. For spatial incoherence microphones are placed in a reverberation chamber and for sufficient separation between microphones (> 50 cm) the coherence between signals is very low. Four microphones at the corners of a tetrahedron have frequently been used |65|, to provide four mutually incoherent signals.

Damaske |65|, using two such spatially incoherent signals, measured the threshold of a signal with an angle of azimuth, α, in the presence of a signal from straight ahead using filtered music (250-2000 Hz) as the sound signal; the experimental arrangement eliminated loudness changes. This threshold result has been included in Figure 8.6; considering the difference in signals used, agreement with results for an anechoic signal and a reflection is persuasive. Damaske discovered that the lowest threshold occurs for an angle of azimuth of $\alpha = 100^{\circ}$, whilst experience with lateral reflections suggests a figure of $\alpha = 90^{\circ}$, though this discrepancy is minimal considering the low precision of measurement with lateral reflections. Damaske also notes that the threshold determining criterion was movement of the apparent source area.

The fact noted above, that a signal and a delayed version of that signal are mutually incoherent subjectively, helps explain the similarity in the thresholds. It also helps explain why the same variation with azimuth occurs for a lateral (music) signal in the presence of a completely unrelated frontal noise signal |62|. In each case the same mechanism appears to be operating, and in the two cases with related sound signals the apparent source area is the determining criterion. An explanation in terms of cross-correlation between signals at the two ears to explain the apparently greatest sensitivity for such lateral sound is contained in section 13.07.

Damaske also notes that the apparent source area with a frontal and lateral noise signal (200-10,000 Hz) increased both with increasing level of the additional incoherent sound, and also with increasing angle of azimuth α. This again coincides with experience with early lateral reflections, though with an important difference: namely that the apparent source area in Damaske's experiment increased for $\alpha > 90^{\circ}$, whilst it decreases for early lateral reflections.

Damaske |65| also determined that, for variations between the coherence of the signals (by varying the microphone separation for spatial incoherence), the apparent source size increases with increasing incoherence. "With greater incoherence, the measurement of the interaural time difference in the hearing mechanism becomes increasingly more difficult, localisation tends to become more imprecise .... The ability to localise is limited when so many incoherent sources are placed that the cross-correlation function has no predominant maximum" (Damaske - author's translation).

Damaske |77| has measured the cross-correlation functions for noise signals for different angles of azimuth with a dummy head. Response in this case was dominated by the resonance at 3.5 kHz due to the ear canal. As expected the cross-correlation maximum coincides with previously measured interaural time delays. Damaske and Ando |78| used a lower frequency band (250 Hz-2 kHz) such that the ear-canal resonance is no longer dominant. Using a new dummy head they discovered that the cross-correlation function was dominated by a resonance at 1.5 kHz due to sound diffraction of the head and outer ear. They measured and compared with calculated values the correlation functions for different combinations of sound sources, with both coherent and incoherent signals. In particular they demonstrated how with coherent signals, cross-correlation predicts a central localisation for identical signals from each side, whilst with phase reversal localisation occurs inside the head; further, they were able to show how with a quadraphonic sound system 90° lateral localisation cannot be simulated. For incoherent signals it was observed that a predicted array of four incoherent sources produced a more highly diffuse field than the general symmetrical arrangement.

Both the above studies can be criticised for the following reason: since cross-correlation is necessarily a neural process, why should cross-correlation behaviour be determined by the spectrum at the ear-drum? This criticism is particularly relevant to the work of Damaske and Ando. Not only does it seem unreasonable to suggest that a state of diffusion cannot be assessed for frequency bands other than 1500 Hz; but it is also curious that at this particular frequency the precision of auditory localisation appears to be poorest, and that it is at the limit of the frequency region for which the cross-correlation process can be used for localisation. This will be further discussed in section 13.9.

13.4 THE CROSS-CORRELATION FUNCTION IN COMPLEX SOUND FIELDS

Since early reflections are in general incoherent from an auditory point of view, the following derivation, taken from Damaske and Ando |78|, is also relevant to the situation of direct sound and early reflections. It is shown that the interaural cross-correlation function for incoherent signals from a number of sources in different directions is given by the sum of the cross-correlation functions for the signals from the individual sources.

For the general cross-correlation function in equation (13.1), the pressures at the ears for a single source, $f_\ell^1(t)$ and $f_r^1(t)$, can be expressed as the convolution of the pressure impulse responses, $h_{1\ell}(t)$ and $h_{1r}(t)$, for the paths from the source to the left and right ears and the signal from that source, $s_1(t)$:

$$f_\ell^1(t) = h_{1\ell} * s_1,$$

$$f_r^1(t) = h_{1r} * s_1.$$

For a source direction, in azimuth α_1, the cross-correlation function is given by

$$\phi_{\ell r}^{\alpha_1}(\tau) = \phi_{h_{1\ell}h_{1r}} * \phi_{s_1 s_1},$$

e dep. upon path
[illegible] source.

where $\phi_{h_{1\ell}h_{1r}}$ is the cross-correlation function of the impulse responses $h_{1\ell}$ and h_{1r}, and $\phi_{s_1 s_1}$ is the autocorrelation function of the signal.

For two sources, the pressures at the ears becomes

$$f_{\ell}^{II}(t) = h_{1\ell}(t) * s_1(t) + h_{2\ell}(t) * s_2(t)$$

$$f_r^{II}(t) = h_{1r}(t) * s_1(t) + h_{2r}(t) * s_2(t).$$

For incoherent signals, the cross-correlation function of the signals $s_1(t)$ and $s_2(t)$ is negligibly small, i.e. $\phi_{s_1 s_2} \approx 0$, so the interaural cross-correlation function is simply expressed by

$$\phi_{\ell r}^{II}(\tau) = \phi_{\ell r}^{\alpha_1}(\tau) + \phi_{\ell r}^{\alpha_2}(\tau). \qquad (13.3)$$

$\Phi_{s_1 s_2} = 0$

For a number of sources, N, (or reflections), the cross-correlation function is given by

$$\phi_{\ell r}^{N}(\tau) = \sum_{n=1}^{N} \phi_{\ell r}^{\alpha_n}(\tau). \qquad (13.4)$$

The normalised interaural cross-correlation function, $\Phi_{\ell r}(\tau)$, for N incoherent signals is given in the next section in equation (13.6).

13.5 THE SHORT-TERM CROSS-CORRELATION COEFFICIENT AS A MEASURE OF SPATIAL IMPRESSION

Keet |8|, using stereo recordings of single source reproductions of orchestral music in real halls replayed them to subjects who were then asked to measure the degree of spatial impression by assessing the apparent source width in degrees. He suggested that in arriving at an apparent source width the ear performs a short-term cross-correlation between the signals arriving at each ear. An objective measure was established by recording the impulse response of the hall in a pair of stereo cardioid microphones; the short-term cross-correlation coefficient is then the measure of coherence between these two impulse responses. If A and B are the impulse responses in the two microphones, the short-term cross correlation coefficient, K, is

$$K_0^{50} = \frac{\int_0^{50\text{ msec}} A \times B.dt}{\left\{\int_0^{50\text{ msec}} A^2.dt \int_0^{50\text{ msec}} B^2.dt\right\}^{\frac{1}{2}}} \qquad (13.5)$$

He found that there was a straight line relationship between the subjective apparent source width and the objective degree of incoherence $(1 - K_0^{50})$, as measured with stereo microphones. A high degree of incoherence corresponded to a large apparent source width. It is instructive to relate Keet's short-term cross-correlation coefficient to the more general cross-correlation function $\phi_{\ell r}(\tau)$.

As outlined above, the cross-correlation function for incoherent signals from different directions is given by the sum of the cross-correlation functions for each signal. This is in fact also the situation for direct sound and early reflections. The normalised cross-correlation function for N incoherent signals is

$$\Phi_{\ell r}(\tau) = \frac{\sum\limits_{n=1}^{N} \phi_{\ell r}^{\alpha_n}(\tau)}{\left\{\sum\limits_{n=1}^{N} \phi_{\ell\ell}^{\alpha_n}(0) \,.\, \sum\limits_{n=1}^{N} \phi_{rr}^{\alpha_n}(0)\right\}^{\frac{1}{2}}} \,. \tag{13.6}$$

Whilst the denominator is expressed in terms of autocorrelation functions it is simply the product of the incoherent energy sums in the two ears.

For Keet's measurement of the "short term" cross-correlation coefficient, the value of the correlation function at $\tau = 0$ was used (i.e., the height of the central maximum). Further, for his measurements the test signal was of short duration. However, the validity of equation (13.6) depends only on $\phi_{s_1 s_2} = 0$ for all pairs of signals; for a noise or music signal this occurs for delayed signals of relative delay >1 ms, whilst with Keet's test signal the limiting delay is larger (see Figure 13.~~7~~8). It is thus apparent that "short-term" refers to the proportion of sound which is to be considered, rather than the nature of the test signal. To make measurements in real halls this demands a short duration test signal, to permit one to sample only the early reflections; however, in a simulation such a signal is not necessary, or even particularly desirable. This will be further discussed below.

Considering now the correlation between the signals received at the two microphones used by Keet,

$$K_0^{50} = \Phi_{\ell r}(0) = \frac{\sum \phi_{\ell r}^{\alpha_n}(0)}{\left\{\sum \phi_{\ell\ell}^{\alpha_n}(0) \,.\, \sum \phi_{rr}^{\alpha_n}(0)\right\}^{\frac{1}{2}}} \,, \tag{13.7}$$

for all reflections arriving within 50 ms of the direct sound. It is assumed that the coherence between microphone signals for lateral sound is zero, whilst for unit intensity frontal sound it is unity: i.e.,

$$\phi^{90^\circ}_{\ell r}(0) = 0, \quad \phi^{0^\circ}_{\ell r}(0) = 1,$$

for unit intensity signals. In this discussion sound is assumed to be either left or right lateral or frontal. Sound arriving on axis in one microphone was attenuated by about 5 dB in the other microphone at mid-frequencies |79|; the effect on the lateral sound from the left or right, f_L and f_R, is indicated in Figure 13.4 (0.56 = antilog -5/20). A typical value for cardioid microphones for the attenuation of sound 45° off-axis i.e., for the frontal sound for each microphone, is 1.5 dB (0.85 = antilog -1.5/20). Hence if $\sum f_D^{\,2}$ is the total direct (or frontal sound),

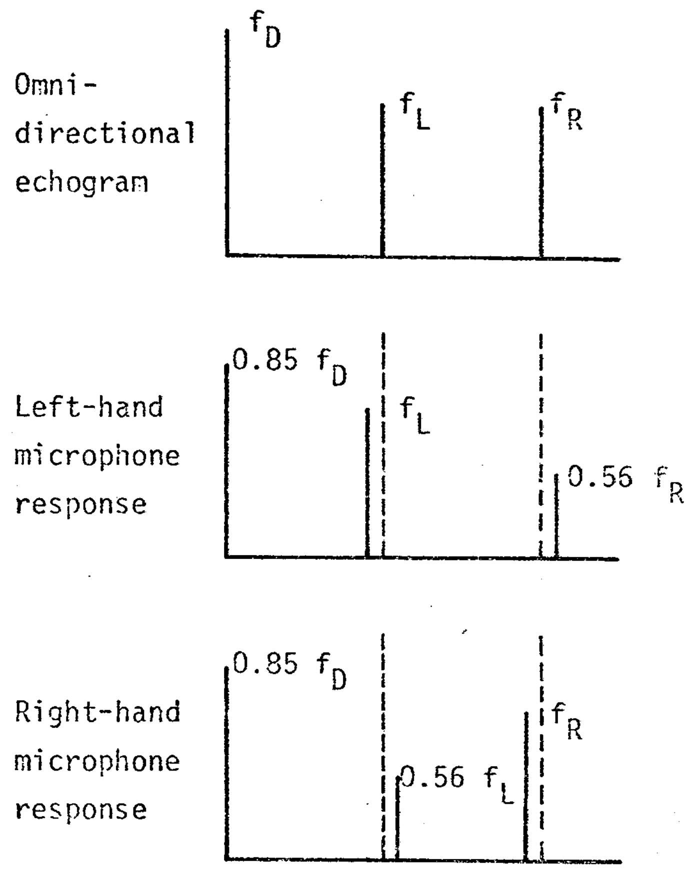


Fig. 13.4. Echograms for Keet's cross-correlation measurement.

$$\sum \phi^{\alpha_n}_{\ell r}(0) = \sum \phi^{0^\circ}_{\ell r}(0) = (0.85)^2 . \sum f_D^{\,2}$$

and

$$K_0^{50} = \Phi_{\ell r}(0) = \frac{(0.85)^2 . \sum f_D^{\,2}}{\{((0.85)^2 . \sum f_D^{\,2} + \sum f_L^{\,2} + (0.56)^2 . \sum f_R^{\,2}) \times ((0.85)^2 . \sum f_D^{\,2} + (0.56)^2 . \sum f_L^{\,2} + \sum f_R^{\,2})\}^{\frac{1}{2}}}.$$

It is convenient to assume that the left-hand sound and the right-hand sound are equally intense: i.e., $\sum f_L^{\,2} = \sum f_R^{\,2}$. Thus

$$K_0^{50} = \frac{0.7\sum f_D^{\,2}}{0.7\sum f_D^{\,2} + 1.32\sum f_L^{\,2}} = \frac{1}{1 + 1.86\,\dfrac{\sum f_L^{\,2}}{\sum f_D^{\,2}}} .$$

However, the ratio of lateral to non-lateral early sound is,

$$\text{antilog}\ \frac{S}{10} = \frac{\sum f_L^{\ 2} + \sum f_R^{\ 2}}{\sum f_D^{\ 2}} = \frac{2.\ \sum f_L^{\ 2}}{\sum f_D^{\ 2}}$$

Thus,

$$1.86\ \frac{\sum f_L^{\ 2}}{\sum f_D^{\ 2}} = \text{antilog}\ \frac{S - 0.3}{10}$$

With the 0.3 dB term ignored,

$$1 - K_0^{50} = \frac{\text{antilog}\ \frac{S}{10}}{1 + \text{antilog}\ \frac{S}{10}} . \qquad (13.8)$$

Keet's degree of incoherence is thus seen to be a particular measure of the interaural cross-correlation coefficient. From equation (13.8) it is seen that it is monotonically related to the measure derived from subjective experiments as being related to subjective impression. In fact the right hand side of equation (13.8) is simply the lateral energy fraction. It was mentioned above that Keet discovered a linear relationship between the perceived apparent source width and the degree of incoherence. By using the relationship in equation (13.8) and the relationship derived in Chapter 9 between the subjective impression and the ratio S, a relation can be derived between subjective spatial impression and the degree of incoherence.

13.6 THE DEGREE OF INCOHERENCE AS A MEASURE OF SPATIAL IMPRESSION

The relationship between $(1 - K_0^{50})$ and S, the ratio of lateral to non-lateral early sound, from equation (13.8), is plotted in Figure 13.5. The form of this curve is in fact reminiscent of that derived previously for the subjective degree of S.I. (Figure 9.4); the change in the degree of incoherence is small for small levels of lateral sound.

Figure 9.4 gives the relationship between the subjective scale of S.I. and the quantity S. If the degree of S.I. is related to $(1 - K_0^{50})$, S being used as the common parameter, the relationship in Figure 13.6 is derived. It is apparent that $(1 - K_0^{50})$ is a linear measure of the degree of spatial impression. Divergence from linearity only occurs for the lowest values

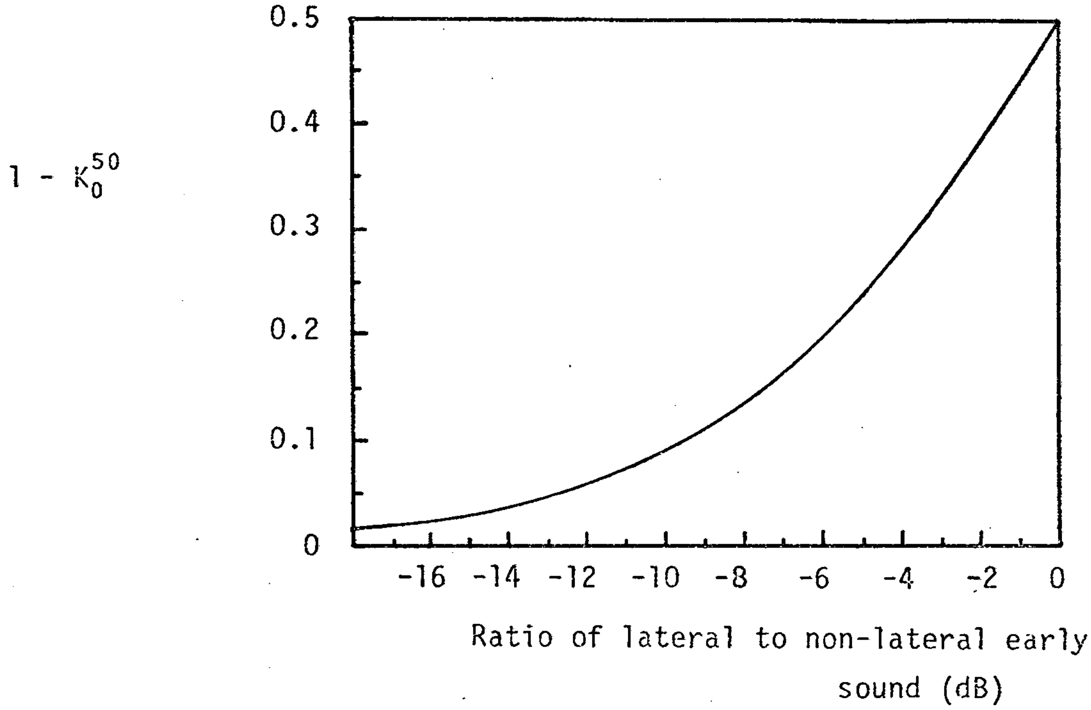


Figure 13.5. Relation between the degree of incoherence ($1 - K_0^{50}$) and the ratio of lateral to non-lateral early sound, S (dB).

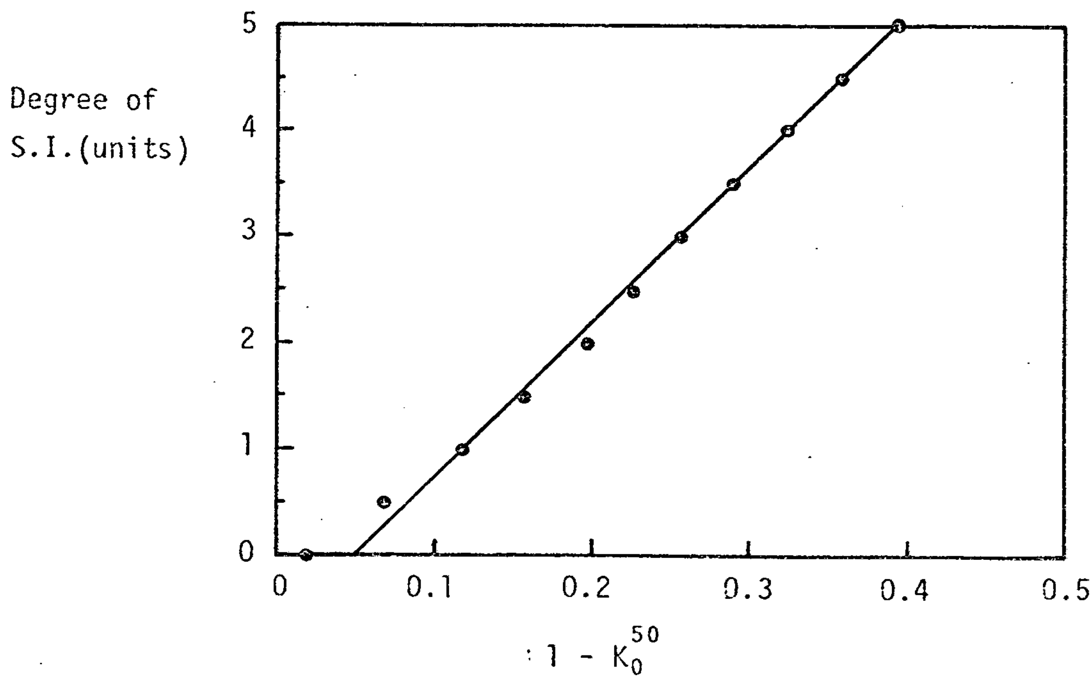


Figure 13.6. Relation between the degree of S.I. and the degree of incoherence ($1 - K_0^{20}$).

of S, for which in any case Reichardt's results (used to derive Figure 9.4) have a high margin of inaccuracy.

Keet's measure of the degree of incoherence is thus not only shown to be a unique measure of the degree of spatial impression, but also a linear one. This leads to the interesting conclusion that the apparent source width (measured in degrees) perceived by subjects is also a linear measure of the degree of spatial impression.

Damaske's measurement |65| of the perceived source size with noise signals (250 Hz - 2 kHz) with four incoherent signals as the degree of coherence is varied is also found to give a linear relationship. This is plotted in Figure 13.7, which is derived from Figure 3 and 22 in reference |65|. The degree of coherence $\bar{k}$ was also derived by Damaske from the height of the normalised maximum of the cross-correlation function between the loudspeaker signals, a measure similar to Keet's. It is intriguing that a linear relation exists both for the case of apparent source width and perceived (solid angle) source size.

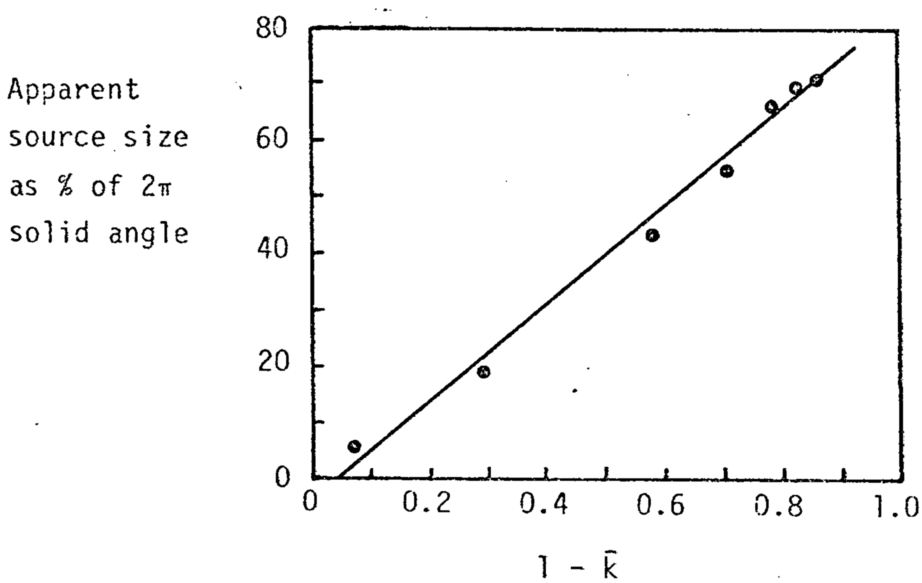


Figure 13.7. Perceived apparent source size as a function of signal incoherence (1 - $\bar{k}$) for four incoherent signals (after Damaske |65|).

The evidence that the degree of incoherence is a linear measure of the degree of spatial impression does not constitute proof that the process by which the auditory system derives a spatial impression is a correlation process. A simple experiment which demonstrates that spatial impression relies on different stimulation of the two ears (i.e., that S.I. depends on comparison between ear signals) is to present a subject with identical lateral reflections from each side. Direct sound from straight ahead and two side reflections, each at -3 dB relative to the direct sound, each delayed 40 ms at angles of azimuth of $\pm 40^{\circ}$ were presented to subjects, who were able to compare this situation with that when the reflection delays were different. Whilst, in general, varying the delay of reflections produces little change in the subjective impression, all subjects agreed that there was definitely something different about the degenerate case, and found the degree of spread of sound much less in this case. By moving one's head about 30 cm to each side, one returned to the normal impression with two side reflections. This last experience compares with the simple localisation experiment when two loudspeakers, as in a stereo system, are fed with the same signal. Upon moving one's head adequately to introduce a time delay of 630 μsec (path length difference 21.6 cm) between the signals at each ear, the point of localisation moves from between the loudspeakers to the loudspeaker nearest one's head.

This experiment elegantly demonstrates not only the fact that spatial impression depends on a correlation process between the signals at the two ears, but also that spatial impression is closely connected to the localisation process. Its significance is virtually limited, however, to experimental situations, such a degenerate case would only occur in a perfectly symmetrical concert hall with both source and listener on the axis of symmetry, and is not relevant to a broad orchestral source.

The evidence quoted so far completely supports Keet's measure of the degree of incoherence as a measure of spatial impression. Keet, to measure the quantity $(1 - K_0^{50})$, used cardioid microphones, which at least for sound from in front, corresponds roughly to the directionality of human ears. Modern measurements might be made by using an artificial head, but this would probably not alter measurements very much. However, a unique measure of spatial impression should give a maximum degree of S.I. for purely lateral sound; it remains to be seen if the measure of $(1 - K_0^{50})$ would give a maximum for this angle of azimuth. Consideration of this also throws light on the choice of test signal used by Keet, and the validity of the assumption $\phi_{\ell r}^{90^{\circ}}(0) = 0$ used in section 13.4.

13.7 THE VARIATION OF DEGREE OF INCOHERENCE WITH REFLECTION AZIMUTH

Consider a situation with direct sound and a single lateral reflection at azimuth α (a pair of reflections would in fact be more suitable from a subjective point of view, but the argument is identical and simpler to express for a single lateral reflection). The degree of coherence for $\tau = 0$ is, from equation (13.3),

$$\phi_{\ell r}(0) = \phi^{0^o}_{\ell r}(0) + \phi^{\alpha}_{\ell r}(0).$$

(Typical correlation functions for this situation are sketched in Figure 13.10 below.) If k^{α}_{ℓ} and k^{α}_{r} are the pressure levels at the left and right ears for the lateral reflection (it being assumed that $k^{0^o}_{\ell} = k^{0^o}_{r} = 1$), and τ_{α} is the interaural time delay, then

$$\phi^{\alpha}_{\ell r}(0) = k_{\ell}.k_{r}.\Phi_{ss}(\tau_{\alpha})$$

and

$$\phi^{0^o}_{\ell r}(0) = 1,$$

where $\Phi_{ss}(\tau)$ is the autocorrelation function of the test signal. Hence the degree of coherence, $\phi_{\ell r}(0)$, is given by

$$\phi_{\ell r}(0) = 1 + k_{\ell}.k_{r}.\Phi_{ss}(\tau_{\alpha}). \qquad (13.9)$$

Since k^{α}_{ℓ} and k^{α}_{r} are relatively constant for small variations of α, and are virtually unity at low frequencies, which have been shown in Chapter 12 to be dominant for spatial impression, the maximum degree of incoherence (the minimum value of $\phi_{\ell r}(0)$) occurs for the minimum value of $\Phi_{ss}(\tau_{\alpha})$ for $0 \leq |\tau_{\alpha}| \leq 630\mu s$. In other words the angle of azimuth that produces the maximum degree of incoherence is determined by the autocorrelation function of the test signal used. $\Rightarrow$ Ando's paper

Keet for his measurements used a 5 ms duration swept-frequency pulse starting at 1 kHz and ending at 1.2 kHz. The autocorrelation function for this test signal is given in Figure 13.8. This function has a minimum value for $\tau = 400\ \mu s$, which is the interaural time delay for a signal of azimuth $\alpha = 50^o$, (from Figure 13.1). Thus Keet's measure $(1 - K^{50}_{0})$ would predict a maximum degree of spatial impression for $\alpha = 50^o$. To predict a maximum degree of spatial impression for $\alpha = 90^o$, the test signal autocorrelation function should have a minimum value for $\tau = 630\mu s$, for which a test signal centre frequency 700 Hz is required.

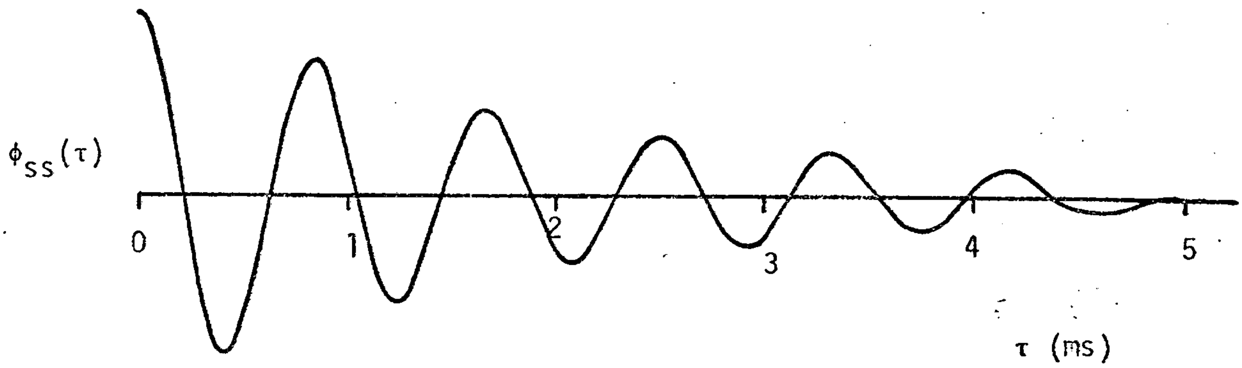


Figure 13.8. Autocorrelation function of the test signal used by Keet |8|.

However, for concert hall measurements, a further consideration governs the choice of test signal used. In a concert hall with a large number of reflections, the effects of lateral reflections were found to sum, so the duration of the test signal should be as short as possible, to obviate possible cross-correlation between one reflection signal and another. Keet provides no explanation for his choice of test signal, but a duration of 5 ms is relatively long compared with the sampling time of 50 ms. The autocorrelation function of the ideal test signal would decrease from unity to zero for τ in the range 0 to 630μs; this is sketched in Figure 13.9(a). A reasonable approximation to this is provided by a single cycle 750 Hz tone pulse, whose autocorrelation function is given in Figure 13.9(b).

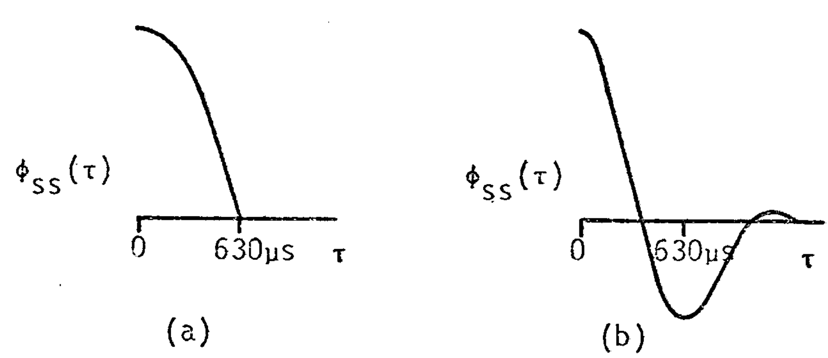


Figure 13.9. Autocorrelation functions of (a) the ideal test signal and (b) a single cycle 750 Hz tone pulse.

Such a 750 Hz tone pulse would be ideal as a test signal if all reflections arrived at the listener "unfiltered". This is not the case, however, due to audience attenuation at grazing incidence, quite apart from any frequency-selective reflection that might occur. Since audience attenuation filtering has been shown in the previous chapter to severely reduce the degree of spatial impression, this aspect cannot be ignored. The results in section 12.4 suggest a test signal in the region of 200-400 Hz as being suitable. Such a signal does in fact also fit the azimuth angle for maximum incoherence criterion, since for all frequencies below 750 Hz, $\Phi_{ss}(\tau)$ decreases from $\tau = 0$ to $\tau = 630\mu s$, and so for a single reflection maximum incoherence occurs for $\alpha = 90^{o}$ (see Figure 13.10 below). Two problems arise when low frequency signals are used: the duration becomes long, but perhaps more important the value of $\phi_{\ell r}(0)$ only changes marginally for changes in azimuth. Since the auditory system must be performing a correlation between neural signals, the "neural sharpening effect" permits the auditory system to distinguish azimuth independent of frequency below 500 Hz. Before deciding whether a realistic measure is possible by correlation techniques, it is relevant to discuss the probable auditory mechanism involved in creation of spatial impression.

13.8 A MODEL OF AUDITORY PROCESSING FOR SPATIAL IMPRESSION

The fact that it is predominantly bass frequencies which determine the degree of spatial impression indicates a filtering process on the total sound field incident at the eardrums. The results in Chapter 11 and 12 suggest that the contribution of frequencies above 1500 Hz to spatial impression is minimal; hence the processing for S.I. occurs in the frequency range over which localisation depends on interaural phase or time differences. Such behaviour is consistent with a correlation process. Further, in the previous section it was shown that for the maximum spatial impression to occur for an azimuth of 90^{o} the centre frequency for the relevant frequency band for S.I. must be below 750 Hz. This is also consistent with the findings in Chapter 11 and 12.

In the normal concert hall situation, one localises on the direct sound; hence, for a listener facing the source, localisation at frequencies below 1500 Hz corresponds with a maximum of the correlation function $\phi_{\ell r}(\tau)$ for $\tau = 0$. Keet's measure of the degree of incoherence is based on the height of this maximum, and was found in section 13.5 to be uniquely related to the subjectively determined measure S, the ratio of lateral to non-lateral sound, at least in terms of reflection level, and also by making simplifying assumptions concerning reflection direction. The following model can therefore be proposed for the auditory process involved in creating a subjective spatial impression.

The correlation process which determines the location of the source at low frequencies is also responsible for determining the degree of spatial impression. The process for S.I. is dependent on a band of frequencies probably centred in the region of 300-400 Hz, and at least having a centre frequency below 750 Hz. The degree of spatial impression is a function of the height of the correlation maximum at the interaural time delay (usually zero) corresponding to the direction of localisation.

Before discussing implications of this model, it is in order to mention the possible location of the filtering action in the neural pathway. Individual neurons leaving the cochlea respond to particular frequency bands, and thus a correlation at a synapse over a limited frequency band is a likely mechanism. In fact in mathematical terms a correlation coefficient for a limited frequency band can be constructed from a broad frequency band correlation by Fourier transformation; thus assumptions of the location of the filtering action in the auditory system are not necessary.

(a) Reflection direction and level.

That this model accounts for the maximum sensitivity for purely lateral sound has been fully discussed. The behaviour of the degree of incoherence with reflection level for the head (rather than a pair of cardioid microphones) is derived by a similar procedure to that in section 13.5. For the human head at low frequencies the sound level at the ears is independent of the angle of incidence and so the analysis is identical to that in section 13.5 with all the numerical coefficients replaced by unity, which leads in fact to the same expression as in equation (13.8).

(b) Reflection delay.

The model also explains, at least qualitatively, the dependence of spatial impression on reflection delay. Equations (13.7) and (13.9) are valid for reflections which are incoherent from an auditory point of view. In this respect reflection delay is unimportant as long as it is larger than a certain value, of the order of a few milliseconds. This corresponds with observed behaviour in Chapter 7. For very small delays, reflections are coherent and contribute ~~the~~ to increase the degree of coherence at $\tau = 0$.

(c) Spatial impression and reverberation.

The correspondence between spatial impression produced by lateral reflections and the spatial effects produced by reverberation is also explained by this model. In equation (13.6), a fully diffuse reverberant field makes no contribution to the numerator, and simply increases the denominator by a factor $(1 + k^2)$, where k = antilog R/10, R being the ratio of reverberant to early sound. As an extension of equation (13.8),

$$(1 - K) = \frac{\text{Reverberant level + Early lateral sound level}}{\text{Total sound level}} \quad . \qquad (13.10)$$

This quantity was shown in sections 12.2 and 12.3 to predict measured results for subjective 'room impression'. The expression also helps explain the significance of late reflections: that spatial impression is created by any discrete lateral reflection, but once conditions are sufficiently diffuse to provide equal energy from all directions, the degree of incoherence is determined by the reverberant level. The addition of reverberant sound thus increases the degree of incoherence. The physical measurement of the expression in equation (13.10) can be a steady state correlation between two omni-directional microphones spaced 22 cm apart, with a low frequency test signal; however, this approach involves ignoring the subjectively perceived difference between the effects of early lateral reflections and reverberation. The model nevertheless explains well the inter-relation; the fact that there is a perceived difference indicates an ability to distinguish reverberant diffuse conditions by some other analysis than that used for spatial impression.

(d) Effect of head rotation.

In considering the ratio of lateral to non-lateral sound, one assumes that the listener is facing the source. However, if in the concert hall situation the listener rotates his head, little change in the spatial impression is perceived, though the ratio is likely to change significantly. The model, not being dependent on actual direction relative to the listener's head, explains this perceived behaviour, since the degree of S.I. is given by the height of the correlation maximum, rather than the particular value for $\tau = 0$.

(e) Perceived and actual source width.

A further comment may also be made concerning the relation between the actual source width and the perceived source width in a concert hall. In a tutti passage, the actual source width also causes a reduction in the height of the correlation maximum; thus actual and perceived source width due to lateral reflections are complementary and involve the same perception mechanism.

(f) Bilateral and unilateral reflection situations.

The variation of spatial impression with reflection delay was found in Chapter 7 to be independent of whether a unilateral or bilateral reflection was used. However, in Chapter 8 it was noted that one problem with investigating the variation of S.I. with azimuth was that a small localisation shift occurs with a single lateral reflection. This is also explained by the cross-correlation model. As can be seen in Figure 13.10 due to the

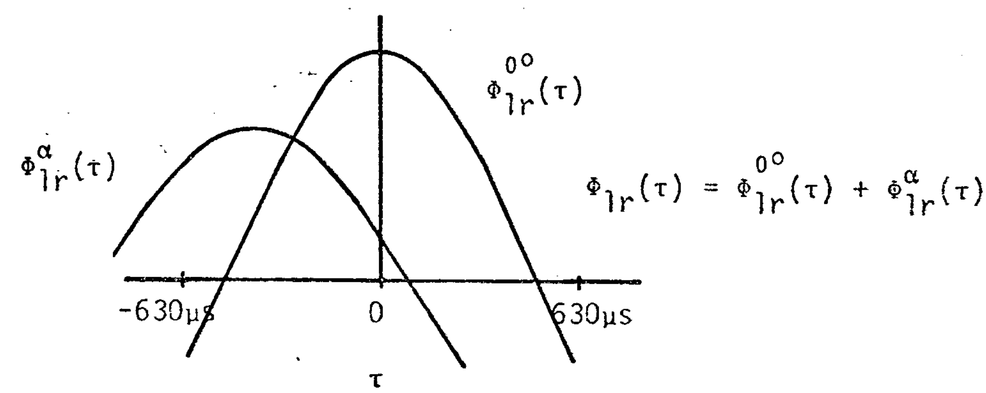


Figure 13.10. Cross-correlation functions for direct sound and a single lateral reflection. The resultant cross-correlation function is the sum of the two individual functions. Signal centre frequency, 500 Hz.

negative slope of the correlation function for the lateral reflection at $\tau = 0$, the resultant cross-correlation function has a maximum slightly displaced negatively from $\tau = 0$. Bilateral reflections do not produce such a displacement. For spatial impression, both a unilateral and bilateral reflection contribute about equally, depending on reflection level.

13.9 THE PERCEPTION OF DIFFUSE SOUND FIELDS

Consideration of the frequency behaviour also helps to clarify the results of experiments mentioned in section 13.3 on creation of a synthetic diffuse field. In a subjectively diffuse field, in the absence of direct sound, there is no prominent maximum in the cross-correlation function between the two ears. However, whilst with spatial impression the effect is predominantly low frequency, for appreciation of a diffuse field two frequency areas are relevant: below 1500 Hz and above 1500 Hz. Below 1500 Hz, the localisation mechanism performs a cross-correlation process to determine the degree of diffuseness, whilst above 1500 Hz amplitude and pinna effects determine the perceived sound directions. The dual mechanism helps explain the apparent contradiction in Damaske's results |65|: namely, that at threshold, the maximum sensitivity occurs for purely lateral sound, whilst for an incoherent lateral source in addition to a frontal one, the perceived source area increases for $\alpha > 90^{o}$. The threshold measurement was made with a signal below 2 kHz, for which a cross-correlation process dominates, whilst for the above threshold experiment high frequencies (up to 10 kHz) were present and sound was perceived from behind with a source behind the listener. Since front-back confusion is high at low frequencies (Stevens and Newman |75|), the perceived source area is probably not perceived as being frequency dependent.

For the experiments of Damaske and Ando |78| it is interesting that they chose a noise signal below 2 kHz. For this situation the auditory cross-correlation process can be considered dominant for localisation. It was noted above, however, that there is no a priori reason for correlation to be determined by the frequency spectrum as measured at the eardrums. Subjective observations included in their paper are generally concerned with coherent sources; in each case the correlation maximum is identical for any centre frequency. The only observation on subjective diffuseness is with a four-loudspeaker system arranged symmetrically around the head; in a comparison experiment with a predicted optimally diffuse situation (for correlation using a 1.5 kHz "signal") they found the predicted situation to be subjectively superior. However, for lower frequency sound the predicted situation would also be superior from a cross-correlation point of view, since the predicted situation is less symmetrical about the head, and individual correlation maxima for each signal are more random.

13.10 A REVISED METHOD OF MEASURING THE DEGREE OF INCOHERENCE FOR SPATIAL IMPRESSION

The principal elements of the proposed system for measuring the degree of incoherence for spatial impression have already been mentioned above, and so the rationale behind them will not be repeated here. It is required to measure the short-term cross-correlation function (equation 13.5) used by Keet |8| with two significant modifications: that a low frequency tone pulse of short duration should be used, and that sampling should be over a longer time period. The discovery that spatial impression is essentially a low frequency effect suggests that either a signal should be chosen with a suitable bandwidth or that the microphone response should be filtered. As suggested in section 13.7, a single cycle tone pulse at a frequency below 750 Hz is desirable, such a pulse, however, has a very broad bandwidth (the 3 dB bandwidth for a 630 Hz single cycle in Figure 16.1 was 700 Hz). Filtering therefore appears necessary, though a signal frequency outside the pass-band is quite acceptable. The signal frequency and filter require very careful selection so that the frequency "distortion" due to audience attenuation at grazing incidence in halls is accounted for correctly for the measurement to be subjectively valid. Further subjective experiments are required for this choice.

Since only low frequencies are being considered, for which no amplitude head-shadowing occurs, microphones can be omni-directional, separated by 22 cm. The line joining their centres should be perpendicular to the line from the source to the test position. Two microphones located on the sides of a sphere of radius equal to the human head (or an artificial head) is more true to reality but the errors inherent in using two microphones located to give the correct maximum interaural/inter-microphone time difference would be very small.

Keet investigated the correlation over different 50 ms duration intervals and found that only for the first did the subjective rank ordering agree with correlation measurements. More significant, however, would have been the investigation of different intervals starting from zero time delay. In anticipation of arguments to be presented in section 14.3, an 80 ms interval seems more appropriate.

The microphone signals could readily be recorded andanalysed digitally; by using digital filters in such a system no problems would arise due to filter phase response. For a direct reading analogue system there is, in addition to the problem of filter phase response, also the problem of the use of analogue multipliers followed by integration, leading to errors due to integrating drift voltages. For both systems it is necessary to maintain an adequate signal-to-noise ratio on measurement, which might present transducer problems or influence the choice of signal duration. The precision requirements for the system are high since the degree of coherence for a relatively high degree of spatial impression (S = -2.5 dB) would only be 0.75 when using a 400 Hz centre frequency. The interval $\Phi(0) = 1$ to 0.75 corresponds to about five subjective units.

For high levels of coherence, Burger and Keet |80| have recently pointed out that a good estimate of the degree of coherence may be made by simply displaying one signal on the x-plates and the other on the y-plates of an oscilloscope. In fact, this method only has quantitative value for random signals, though in qualitative terms it offers a simple solution for early sound.

The advantage of such a measuring system, if built, over the method described in Part II is that with a single measurement the degree of spatial impression is obtained for that particular receiver position. Scaling down the system for use in 1:8 or 1:10 models should also present no problems (though scaled down seating is also necessary for measurements to be valid).

13.11 CONCLUSIONS

Modern auditory theory of localisation was reviewed; it was found that substantial evidence exists to suggest that below 1500 Hz a correlation process between the signals at the two ears is performed. Reported measurements of simulated diffuse sound fields also point to an auditory cross-correlation process being used to determine the degree of diffuseness. It was suggested, however, that since correlation must take place at the neural level, there is no reason to assume that peaks in the frequency response at the eardrum determine the shape of the cross-correlation function. It was also noted that at low frequencies, a "neural sharpening effect" causes the cross-correlation function to be more "peaked" than in a corresponding measurement on the signals at the eardrums.

Keet |8| has proposed that a short-term cross correlation process is involved in determining the degree of spatial impression, and found that an objectively measured correlation coefficient (in equation 13.5) was linearly related to perceived apparent source width. It has been shown here that Keet's correlation coefficient is simply a special case of the general correlation coefficient between the signals at the two ears. It has also been found that Keet's degree of incoherence $(1 - K_0^{50})$ is uniquely related to the ratio of lateral to non-lateral sound; indeed $(1 - K_0^{50})$ equals the lateral early energy fraction. Whilst Keet's measurement probably gives good average measures, it has been found that only for test signals below 750 Hz would a maximum degree of spatial impression be predicted for directly lateral sound. In line with findings reported in Chapter 12 it has been suggested that measurements with a tone pulse with centre frequency around 400 Hz would be more suitable.

The following model for the auditory process involved in deriving a perceived spatial impression is postulated: that the degree of spatial impression is given by the height of the correlation maximum at the inter-aural time delay corresponding to the direction of localisation, and that the correlation process is dependent on a frequency band centred in the region of 400 Hz. This model, as well as explaining the variation of spatial impression with reflection delay, level and direction, also explains the relation between the effects of reverberation and early lateral reflections, and the relation between perceived and actual source width. Further quantification of the model is not considered valid due to the "neural sharpening effect" mentioned above.

Finally a possible measuring system to determine the degree of incoherence was discussed, which employs two omni-directional microphones. Analysis by digital methods could readily be achieved, though an analogue system is probably also possible.

Chapter 14

A PHYSICAL PARAMETER RELATED TO THE SUBJECTIVE DEGREE OF SPATIAL IMPRESSION

14.1 THE BASIC PHYSICAL PARAMETER

It was found in Chapter 7 that the subjective degree of spatial impression produced by lateral reflections was predominantly independent of delay for delays greater than 5 ms. Experiments reported in section 10.2 and 10.3 indicated that the degree of the spatial impression is related to the ratio of lateral to non-lateral early sound. In Chapter 8 it was shown that for reflections at a general angle of incidence, the angle to the axis through the listener's ears, ϕ, was relevant. It was found also that the contribution to the lateral sound was given by multiplying the reflection level by cos ϕ. This leads to the concept of the parameter S, the ratio of lateral to non-lateral early sound, which correlates with the subjective degree of spatial impression:

$$S = \frac{\sum\limits_{t=5\text{ms}}^{80\text{ ms}} P.\cos\phi}{\sum\limits_{t=0\text{ ms}}^{80\text{ ms}} P(1 - \cos\phi)}, \qquad (14.1)$$

where P is the "reflection" level ("reflection" here also embracing the direct sound). The choice of the upper time limit will be discussed below. S is generally expressed in dB. The threshold value for S is about -25 dB, though the "working threshold" relevant to the real situation is probably at least 5 dB higher, at about -20 dB.

In the previous chapter it was shown that the lateral energy fraction was a linear measure of spatial impression. The ratio of lateral to non-lateral early sound, measured in dB, will however be retained for the following reasons:

(1) energy ratios are generally expressed in dB,

(2) such a ratio in dB can be estimated at sight from a simple echogram expressed in dB,

(3) S dB is a linear measure over the range most commonly encountered in concert halls (section 9.1),

(4) S = 0 dB corresponds with a uniform angular distribution of early energy (section 19.7).
Thus, since the variable S was used in a previous publication |63|, there are sufficient arguments for retaining it.

Since $(1 - K_o^{50})$, or the lateral energy fraction, is a linear measure of spatial impression (Figure 13.6), a relationship between S and the subjective degree of S.I. can readily be derived. The equation of the line in Figure 13.6 is:

$$\text{Degree of S.I.} = 14.5\,(1 - K_o^{50}) - 0.7. \tag{14.2}$$

Using equation (13.8), namely,

$$(1 - K_o^{50}) = \frac{\text{antilog}\,(S/10)}{1 + \text{antilog}\,(S/10)},$$

one arrives at

$$\text{Degree of S.I. (units)} = \frac{\text{antilog}\,\frac{S + 11}{10}}{1 + \text{antilog}\,(S/10)} - 0.7, \tag{14.3}$$

where S is measured in dB. Since the degree of spatial impression is a function of total early sound level, this expression is valid for the mean level used by Reichardt and Schmidt |64| in their experiment, which was estimated previously as 70 dB. The validity of the expression (14.3) rests on the validity of the experiments discussed in Chapters 7, 8, 10 and on the validity of Reichardt and Schmidt's difference limen experiments.

14.2 FREQUENCY CONSIDERATIONS

It was suggested in Chapter 11 that spatial impression is probably a single subjective effect, but that gross distortions in the spatial spectral balance are probably detected as such. For this reason it was considered necessary in concert hall measurements to consider the value of S at about 200 Hz, the frequency of maximum audience attenuation due to grazing incidence, and the value of S at mid-frequencies, such as, say, 800 Hz. In the event of $S_{200\text{ Hz}} \simeq S_{800\text{ Hz}}$, obviously no problems arise in interpreting the measurements, but for the case of $S_{200\text{ Hz}}$ significantly less than $S_{800\ Hz}$, results summarised in Figure 12.5 suggest that $S_{200\ Hz}$ should be weighted more strongly than $S_{800\ Hz}$ to obtain a single value.

It was suggested in Chapter 13 that it is unlikely, according to the probable auditory mechanism involved, that frequencies above 1500 Hz contribute to spatial impression.

14.3 THE UPPER DELAY LIMIT FOR SPATIAL IMPRESSION

Discussions in the literature generally consider the first 50 ms as constituting the early sound, whilst all sound after that is considered reverberant. The figure 50 ms seems to be an extension of the period generally assumed for the integration time of the ear. The argument seems to run that since the integration time for tone pulses is about 23 ms |56|, and that for speech it is 30 ms |27|, then, since music is less impulsive than speech, 50 ms must be an approximate value for the integration time with this signal. The minimum duration for a musical note is about 100 ms, far in excess of the apparent integration time for tone pulses. A musical experience consists of receiving a number of incoherent signals (from different instruments) and reflections and reverberation both incoherent relative to the direct sound from an auditory point of view (section 13.3). The concept of integration time does not seem particularly relevant to the musical situation. The "integration time" for speech was measured on the basis of intelligibility, some equivalent time period no doubt exists for clarity with music. If the concept of integration time is abandoned, there is no unifying principle regarding the time period for early sound.

It was discovered in Chapter 7 that lateral reflections with delays of 80 ms or more produced a distinct spatial effect. The limiting time period evidently does not describe a discrete cut-off, but the discussion in section 13.8(c) suggests a gradual transition from the effects of lateral reflections to those of reverberation. Since in many halls seat positions frequently occur where lateral reflections and also significant cornice reflections (second order reflections off the ceiling and side walls) arrive later than 50 ms, extension of the sampling time period to 80 ms is desirable. The limit of 80 ms corresponds roughly with the time limit after which discrete reflections can become disturbing |26|, though for bass frequencies most significant for spatial impression, results of

Haas |24| for filtered speech suggest a much longer permissible time period for bass sound before disturbance sets in. This would be in line with communications theory in which temporal resolution decreases with decreasing frequency.

In the event, at least for measurements in two rectangular halls, the respective choices of an 80 ms or a 100 ms limiting time delay led to identical values for the ratio of lateral to non-lateral early sound (section 18.2).

14.4 THE EARLY SOUND LEVEL FACTOR

Virtually all attention has been paid in previous chapters to the lateral energy factor of spatial impression. One reason for this is that it is an invariate characteristic of a hall (for a certain source-receiver combination). The early sound level factor provides a further dynamic aspect to a listener's appreciation of music: the sense of envelopment is intense at 'forte' passages and may be virtually imperceptible in 'piano' lightly orchestrated sections. From Keet's results |8| it is possible to get a rough estimate of the difference limen for total early sound relevant to spatial impression.

From equation (14.2) a change of $1/14.5 = 0.07$ in $(1 - K_o^{50})$ corresponds to a change of one difference limen of spatial impression. From Keet's Figure 3, the mean change in apparent source width for a 0.1 change in $(1 - K_o^{50})$ is 7.7^o; thus the difference limen in terms of apparent source width is 5.3^o. The gradient in Keet's Figure 2, which relates apparent source width to listening level is, for higher levels, about 1.4^o/dB, which leads to a difference limen for spatial impression for the early sound level of $5.3/1.4 \simeq 4$ dB.

This value is indeed small and further experimental evidence is to be desired. It suggests, for instance, that the degree of spatial impression varies between imperceptible to five or more subjective units within the dynamic range of most classical music. The early sound level factor is therefore highly significant for a listener's appreciation of the spatial impression, though the variation between halls is of less significance from this point of view than musical dynamic range.

An additional term may be added to the equation for the degree of S.I. in subjective units (14.3) to account for the early sound level factor. If E dB is the early sound level, since Reichardt and Schmidt's experiment was conducted at an estimated mean level of 70 dB, the additional term is +(E - 70)/4.

14.5 THE RATIO OF LATERAL TO NON-LATERAL EARLY SOUND VERSUS THE DEGREE OF INCOHERENCE AS A MEASURE OF SPATIAL IMPRESSION

In the previous section the difference limen in terms of spatial impression of the degree of incoherence was found to be 0.07. This is of the same order of magnitude as the difference limen for changes in coherence measured by Pollack and Trittipoe |81, 82| for noise signals presented over earphones. 75% of subjects detected a change in interaural coherence from 1 to 0.96. For a high degree of coherence Pollack and Trittipoe also found an increase in difference limen with a decrease in sound level. Their results do differ, however, in one significant respect: whereas Keet's and Reichardt and Schmidt's results suggest a linear relationship between the degree of coherence and spatial impression, Pollack and Trittipoe found that the difference limen increased markedly with decreasing coherence. This disparity of results may indicate the limit of the comparability of results; noise signals through earphones are localised in the head and source broadening, with such signals also occurs within the head.

A measuring system for the degree of incoherence relative to spatial impression was outlined in section 13.9. Such a system has the advantage for measurement that the degree of spatial impression is derived from a single measurement. However for such a system further subjective evidence is required to determine the frequency components for spatial impression. Both the quantities S and E can be readily calculated for a given reflection sequence, and, with a small margin of error, be measured in a concert hall (see Chapter 16). Such measurements permit an assessment of the frequency variation of quantities related to the early sound, rather than the measurement itself including a frequency weighting function to provide a single measure. For this reason and reasons mentioned in section 14.1 above, the ratio of lateral to non-lateral early sound, S dB, as defined by equation (14.1) will be considered for the majority of the remainder of this thesis as the unique measure of spatial impression, together with the early sound level, E.

PART II

Chapter 15

THE THEORETICAL BEHAVIOUR OF SOUND INTENSITY IN ROOMS ACCORDING TO THE GEOMETRICAL IMAGE MODEL

15.1 INTRODUCTION

The subject matter of Part II is an analysis of the physical situation in a concert hall in the light of the findings of Part I. In this introductory chapter, the possible contributions of acoustical theory, according to a simple geometrical image model, are explored. In line with the approach in Part I, discussion is related principally to integrated sound intensity. Predictions concerning the early sound cannot be treated statistically without direct reference to the details of the hall shape and conditions; this will be investigated in Chapters 20-22. The study in this chapter concerns the acoustical situation once the receiver position is no longer critical: that is, in the time period after the direct sound when reverberant conditions apply.

The geometrical approach to room acoustics has received much criticism from theoreticians, but it does remain the standard method for deriving reverberation time (R.T.) formulae. Apart from the virtue of its obvious simplicity as an approach, it also remains the only method capable of dealing theoretically with the effect of room shape on auditorium acoustics |83| or of investigating the acoustics of rooms with specific treatments on different room surfaces (see, e.g. reference |84|).

Bolt, Doak and Westervelt |85| used the geometrical model of room acoustics to derive various well known formulae, which are normally derived on the basis of different assumptions. Their study is based on the assumption of specular reflection at the room surfaces, which they argue is valid if the walls are fairly hard: if the absorption is less than 20%, say. By making the further assumption of incoherent addition of acoustic energies from the various image sources, the image model can be used to derive a formula describing the temporal sound intensity distribution in a room. (Throughout this chapter, the term "sound intensity" has its usual room acoustics meaning: i.e., it is the sound power per unit area arriving at a point from all directions.) The predictions of this theoretical average formula for a rectangular room may then be compared with results computed on the basis of a "discrete images" model of the room.

A similar approach to the one presented here is to be found in papers by Doak |86| and by Gibbs and Jones |87|, though in each the analysis is used only to estimate the total sound intensity in a room. Results in sections 15.5 and 15.6 indicate the validity of these estimates relative to the values computed for discrete image arrays, and demonstrate the form of deviations for different receiver positions. Finally the theoretical analysis of the case of a fully absorbent floor is given, with a discussion of the role of diffusing surfaces in real halls (i.e., surfaces responsible for creating a situation of equal probability of energy flow in all directions).

15.2 INTEGRATED SOUND INTENSITY IN A RECTANGULAR ROOM WITH UNIFORM ABSORPTION

Consider, as in reference |85|, a rectangular room containing a simple point source which emits a pulse at time $t = 0$. Associated with this source is an array of image sources, each source being contained in an image cell which is a replica of the original room. Each image source is considered to emit an identical pulse simultaneous with the primary source. To calculate the sound intensity at a receiver position the intensities of the pulses arriving at that position are summed. If the precise location of the source and receiver can be ignored, then the average number of pulses, N, that arrive within an interval t of the pulse being emitted from the source is given by equation (10) of reference |85|, i.e.,

$$N = \frac{4\pi c^3 t^3}{3V} \tag{15.1}$$

and the number of pulses/second is

$$\frac{dN}{dt} = \frac{4\pi c^3 t^2}{V} \; . \tag{15.2}$$

Here c is the speed of sound and V is the room volume.

In time t the pulse travels a distance ct and the average number of reflections experienced is $ctS/4V$, where S is the effective area of the room boundaries, since $S/4V$ is the mean collision frequency (i.e., the reciprocal of the mean free path; see the paper by Hunt |88|). At each reflection the pulse is attenuated by a factor $(1 - \bar{\alpha})$, where $\bar{\alpha}$ is the mean absorption coefficient in the room. So the attenuation of a pulse due to absorption at room surfaces is $(1 - \bar{\alpha})^{ctS/4V}$.

The pulse is also attenuated due to air absorption by a factor e^{-mct}, where m is the air attenuation coefficient. Further, the pulse suffers spherical divergence by a factor $I_o r_o^2/c^2t^2$, where I_o is the intensity of the direct pulse and r_o is the source-receiver distance. Thus the net intensity of the n^{th} pulse with a travel time t is

$$I_r = \frac{I_o r_o^2}{c^2 t^2} (1 - \bar{\alpha})^{ctS/4V} . e^{-mct}. \tag{15.3}$$

The number of pulses arriving in a time interval δt is, from equation (15.2),

$$\frac{4\pi c^3 t^2}{V} \delta t.$$

Thus the total intensity arriving at the receiver during the interval t to $(t + \delta t)$ is

$$\delta I_t^{t+\delta t} = \frac{I_o r_o^2}{c^2 t^2} .(1 - \bar{\alpha})^{ctS/4V} .e^{-mct} . \frac{4\pi c^3 t^2}{V} . \delta t$$

$$= \frac{4\pi c I_o r_o^2}{V} .e^{(-ct/4V)\left[-S\ell n(1 - \bar{\alpha}) + 4mV\right]} .\delta t \tag{15.4}$$

This last expression is valid insofar as equation (15.1) is valid and insofar as the expression for the mean collision frequency expresses correctly the number of reflections experienced. Both these expressions are realistic if the precise location of the source and receiver can be ignored and, as has been shown in reference |85|, it should be possible to ignore these locations in the general rather than the degenerate case for t greater than a certain value, t_c, depending on the size and shape of the room. The relevant value of t_c is discussed in sections 15.3 and 15.4.

Equation (15.4) is thus valid for $t > t_c$ and the total ("integrated") sound intensity arriving after time t (the pulse from the source having been emitted at time t = 0) is

$$I_t^{\infty} = \frac{4\pi c I_o r_o^2}{V} \int_t^{\infty} e^{(-ct/4V)\left[-S\ell n(1-\bar{\alpha})+4mV\right]} .dt, \quad \text{for} \quad t > t_c,$$

or

$$I_t^{\infty} = \frac{16\pi I_o r_o^2}{-S\ \ell n(1 - \bar{\alpha}) + 4mV} .e^{(-ct/4V)\left[-S\ell n(1 - \bar{\alpha})+4mV\right]} \tag{15.5}$$

If Π is the power contained in the pulse emitted by the source, then $I_o = \Pi/4\pi r_o^2$ and

$$I_t^\infty = \frac{4\Pi}{-S\ \ell n(1 - \bar{\alpha}) + 4mV} . e^{(-ct/4V)\left[-S\ell n(1 - \bar{\alpha}) + 4mV\right]} \tag{15.6}$$

Extending integration down to t = 0 (and ignoring the air absorption term) leads to the following familiar expression:

$$I_o^\infty = \frac{4\Pi}{-S\ \ell n(1 - \bar{\alpha})}, \tag{15.7}$$

which is the classical formula for the total sound intensity, $I = 4\Pi/A$, where A is the total absorption in the room. That equation (15.6) is only valid for $t > t_c$ suggests that equation (15.7) may not be particularly accurate in predicting the total sound intensity. This will be further discussed in section 15.6.

The implications of equation (15.6) are thus that for all times later than t_c after the pulse is emitted the intensity throughout the room space is uniform and decays at a uniform rate throughout the room according to the Eyring formula. The base intensity value (at t = 0) for the decay is the classical value for the total sound intensity.

In practice, it is generally more useful to have sound intensity in a room expressed in terms of the reverberation time T and the intensity I_{01} of the direct sound at unit distance from the source ($I_{01} = I_o r_o^2$). In metric units, equation (15.6) becomes

$$I_t^\infty = 312 . \frac{T}{V} . e^{-13.82t/T} . I_{01} .$$

15.3 NUMBER OF REFLECTIONS RECEIVED IN A RECTANGULAR ROOM

Throughout this section a similar notation will be used to that in reference |85|, except that the coordinate system origin is here chosen as being at the centre of the junction of the front wall and floor. The x-axis runs along the length of the hall, the y-axis across it and the z-axis vertically upwards. All distance measurements are in metres. Image cells are designated by three integers (ℓ, m, n). The hall dimensions are referred to as L_x, L_y and L_z.

A computer investigation was made to compare the number of discrete reflections received at a point in a room with the "theoretical average" number predicted by equation (15.1). The computer programme listed all the computed reflections in temporal order and for each calculated the percentage error between the computed number (M) of reflections and the theoretical average number (N) received by the same time t after the pulse is emitted (percentage error = 100 x (M - N)/N%). A preliminary investigation up to t = 300 ms was made for a rectangular hall of typical concert hall dimensions (50 m x 20 m x 20 m) with an on-axis source at (10, 0, 2). It was apparent that the character of the disagreement between the computed discrete and theoretical average values depended primarily on the distance of the receiver from the front wall. Since the average theoretical number of reflections received increases with the cube of t, investigation of the three-dimensional (3-D) case involves lists of over 1000 reflections for t = 500 ms, and so initially investigations were made of the two-dimensional (2-D) case (floor and ceiling reflections being ignored) and subsequent listing of the 3-D case was used to confirm the conclusions. Plots of the percentage error between the computed discrete and theoretical average numbers of reflections received against t are shown in Figures 15.1-5 for the 3-D case. Results for the corresponding 2-D cases are also shown in Figures 15.1-3.

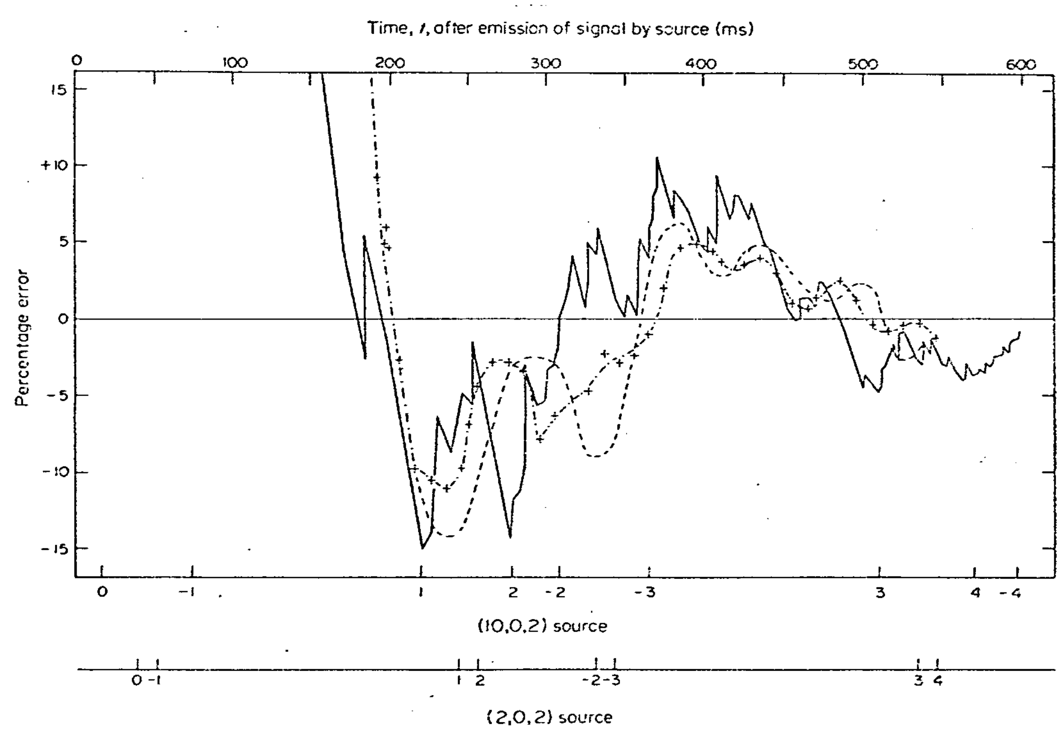


Figure 15.1. Percentage error between computed and theoretical average number of reflections received. The two lower abscissae indicate arrival times and ℓ-values of reflections from (ℓ,0,0) image cells. For source at (10,0,2) and receiver at (15,2.5,2): ———, 2-D and —·—·—,3-D. For source at (2,0,2): - - - - -, 3-D.

As in Figure 2 in reference |85| the disagreement between computed discrete and average numbers of reflections received is large for small values of t, and on a percentage basis disagreement becomes particularly large and essentially haphazard. Thus the behaviour of the curves in Figures 15.1-5 for t < 200 ms is of little interest. The remaining sections of the curves of Figures 15.1-4, however, show a damped cyclic, or almost cyclic, relationship with a definite correspondence between the 2-D and 3-D cases, and the curve of Figure 15.5 also shows a rather randomly cyclic error. This correspondence between 2-D and 3-D suggests that the behaviour of the error curve depends on some aspect not connected with the z-dimension.

Figures 15.1-4 are for seat positions (15,2.5,2),(25,7.5,2), (35,7.5,2) and (45,2.5,2), respectively; the y-values (positions across the room) are chosen at random and in fact changing them has minimal effect on the results. In Figures 15.1, 3 and 4 the cyclic behaviour has a period of roughly 300 ms, while in Figure 15.2 the period is random and the deviations from zero are smaller. It is also apparent that the sign of the error in

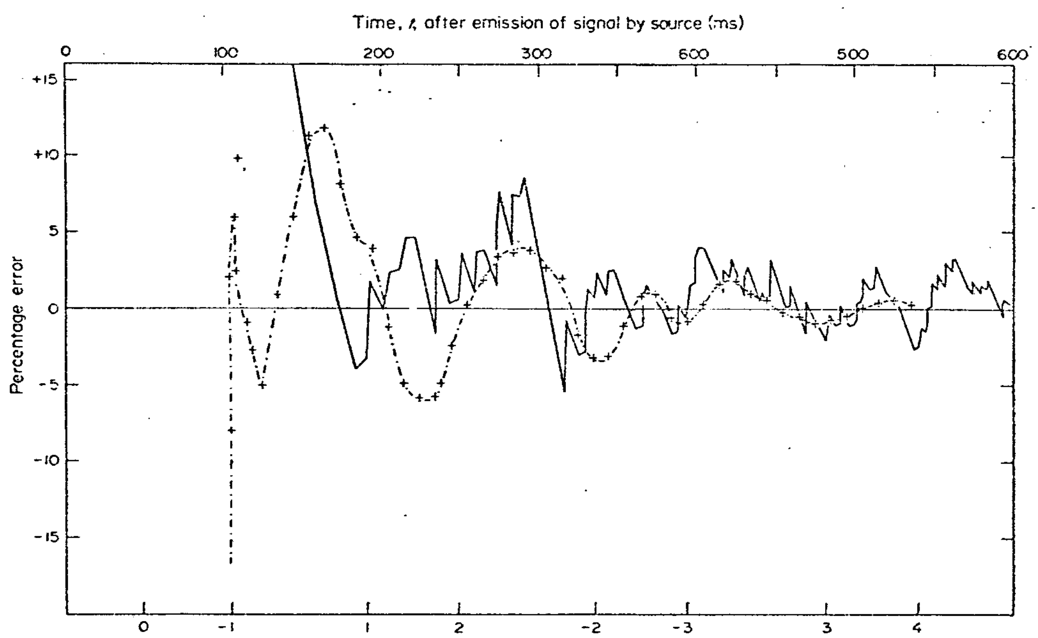


Figure 15.2. Percentage error between computed and theoretical average number of reflections received. The lower abscissa indicates arrival times and ℓ-values of reflections from (ℓ,0,0) image cells. For source at (10,0,2) and receiver at (25,7.5,2): ———, 2-D; —·—·—,3-D.

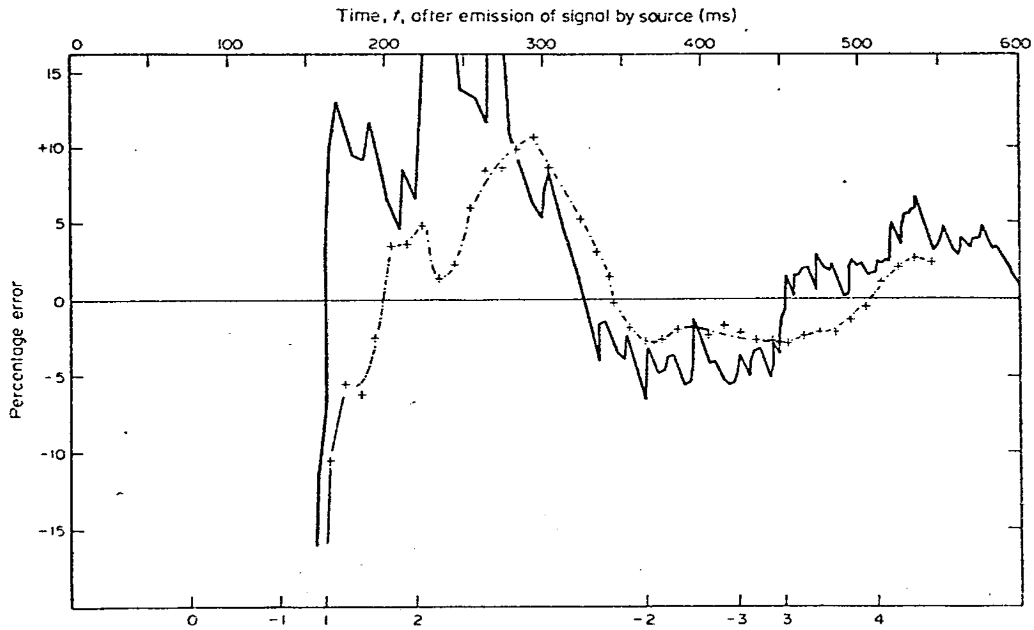


Figure 15.3. Percentage error between computed and theoretical average number of reflections received. The lower abscissa indicates arrival times and ℓ-values of reflections from (ℓ, 0,0) image cells. For source at (10,0,2) and receiver at (35,7.5,2): ———, 2-D; —·—·—, 3-D.

Figure 15.1 is opposite to that in Figures 15.3 and 15.4. These characteristics suggest that the location of the receiver in the x-direction (i.e., the direction along the length of the hall) determines the behaviour of the error, and therefore it is pertinent to examine the location of the receiver relative to the image field.

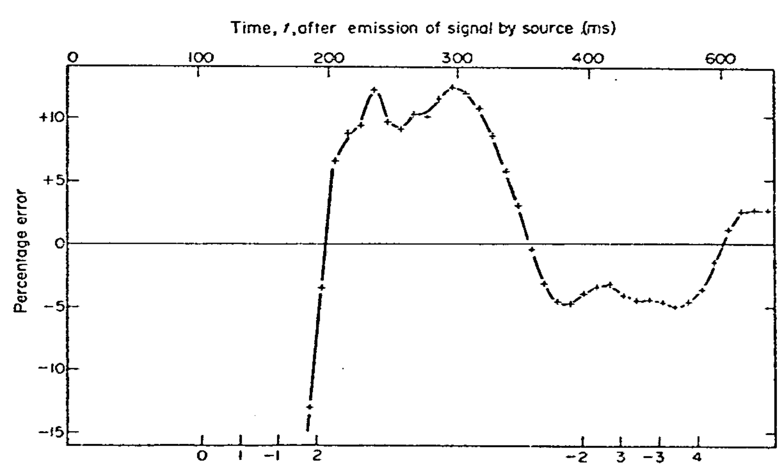


Figure 15.4. Percentage error between computed and theoretical average number of reflections received (3-D). The lower abscissa indicates arrival times and ℓ-values of reflections from (ℓ, 0,0) image cells. Source at (10,0,2) and receiver at (45,2.5,2).

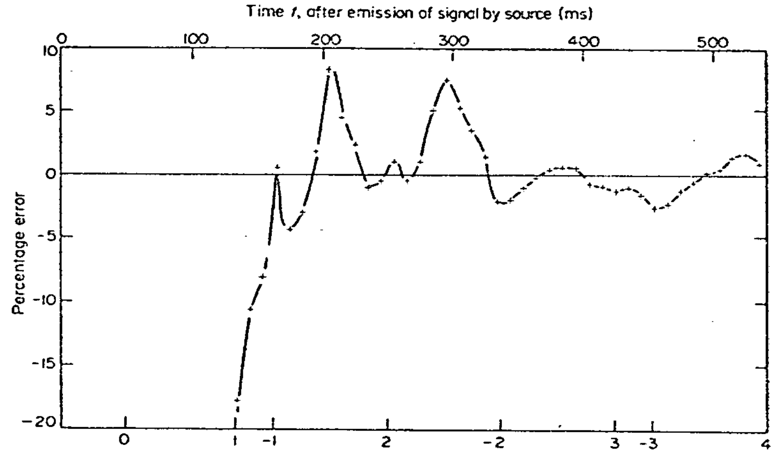


Figure 15.5. Percentage error between computed and theoretical average number of reflections received (3-D). The lower abscissa indicates arrival times and ℓ-values of reflections from (ℓ, 0,0) image cells. Source at (20,0,2) and receiver at (35,7.5,2).

Figure 15.6 shows the image array in plan on the x-y plane; images of the source, as represented by dots, are arranged in planes in 3-D, each plane corresponding to a different (integer) value of ℓ. With the source towards one end of the hall, the planes occur in pairs: ℓ = 0, -1; 1, 2; -2, -3 and 3, 4. The position of the receiver determines the relative arrival times of the pulses from these image planes. Arrival times of reflections from (ℓ, 0,0) image cells are given below each graph in Figures 15.1-5 together with the associated values of ℓ. In Figures 15.1-4 it is noticeable that, when the pulses from (ℓ, 0, 0) images arrive in clusters (as in Figures 15.1, 3 and 4), a regular cyclic error between the computed and theoretical average number of reflections results, whilst when these images are relatively regularly spaced a smaller, random error results (e.g., as in Figure 15.2). It may be readily shown that with the source towards one end of the hall, the most regular arrival of (ℓ,0,0) reflections occurs for a receiver position half-way down the length of the hall, which is the case illustrated in Figure 15.2.

The number of reflections received before time t from an image plane with a particular value of ℓ is theoretically, from simple geometrical considerations, given by

$$N_\ell = \frac{\pi c^2(t^2 - t_\ell^2)}{L_y L_z}, \quad \text{for} \quad t > t_\ell, \tag{15.9}$$

where t_ℓ is the travel time for the particular (ℓ, 0, 0) reflection. So if t_ℓ is larger than it would be for a regular spacing of reflections from (ℓ, 0, 0) images, N_ℓ (for $t > t_\ell$) is less by a constant amount than it would be with a regular spacing of these reflections. Thus comparison of Figure 15.1 with Figure 15.2 indicates that the late arrival of reflections (1,0,0) and (2,0,0) in the case in Figure 15.1 causes the computed number of reflections received to be less than the theoretical average number in the region t = 200-290 ms. However since the late arrival

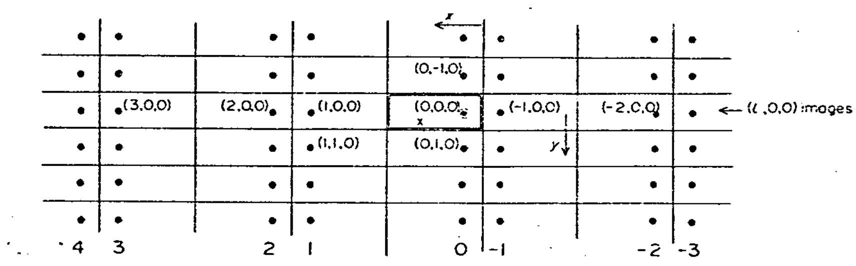


Figure 15.6. Plan view of image array in x-y plane. Cell (0,0,0) represents the actual room.

of the reflections (1,0,0) and (2,0,0) is accompanied by the early arrival of the reflections (-2,0,0) and (-3,0,0), and reflections from their corresponding image planes, so the negative error mentioned above is corrected in the region t = 290-370 ms in Figure 15.1. In this manner the error behaves cyclically and since the pattern of (ℓ,0,0) images repeats itself every $2L_x$, so the period of the cyclic behaviour of the error in Figure 15.1 is about 290 ms ($= 2L_x/c$ for $L_x = 50$m). Also since irregular spacing of image planes results in constant errors (either positive or negative) for each plane, as mentioned above, percentage errors decrease as t and N increase, hence producing the damped nature of the error curves. The behaviour of the error curves in Figures 15.3 and 4 is entirely consistent with the explanation above, which also explains why the sign of these error curves is opposite to that in Figure 15.1. It is also noteworthy that the more dense clustering of (ℓ,0,0) reflections for the case in Figure 15.4 results in larger errors than the 3-D case in Figure 15.3.

The situation with the source very near the front wall, at (2, 0, 2) was investigated for the receiver position of Figure 15.1; the percentage error curve is included in Figure 15.1 as the dashed line. The errors are slightly larger than for the more central source position, as one would expect considering the more sense clustering of (ℓ,0,0) reflections, but the differences are small. When the source no longer lies near the front wall, complete regularity of arrival of $(\ell,0,0)$ reflections occurs for particular source-receiver pairs (the behaviour being unchanged, of course, if the source and receiver are interchanged). Figure 15.5 gives the example of the source at (20,0,2) and the receiver at (35,7.5,2), for which the arrival times of $(\ell,0,0)$ reflections are fairly regular; the error curve for this source-receiver pair has a random period and small values for the error, as occurs in Figure 15.2.

In summary, it can be concluded that the computed number of discrete reflections received in a rectangular room varies cyclically about the theoretical average number given by equation (15.1). With the source towards one end of the largest room dimension (taken here as along the x-axis), as in a concert hall, this cyclic behaviour has a period roughly equal to $2L_x/c$, except at receiver positions half-way down the length of the hall, where percentage differences between computed and theoretical average values are smallest and the period random. The character of this cyclic behaviour is determined by the regularity with which reflections arrive from image planes perpendicular to the x-axis. Thus the validity of equation (15.1) is determined more by the relation of the travel time to the length of the hall than the relation of the travel time to the geometric mean of the hall dimensions as used in reference |85| in the definition of the dimensionless time parameter τ. It remains to be decided what is the minimum value of the travel time (as related to L_x) for equation (15.6) to be valid.

15.4 VALIDITY OF INTEGRATED SOUND INTENSITY FORMULA

The computed integrated sound intensity at the receiver position for discrete reflections is derived by again using equation (15.3) but with k_n, the actual number of wall reflections experienced by the pulse, substituted for the number derived from the mean collision frequency. For present purposes the air absorption term may be omitted to simplify the computation. Thus the discrete integrated intensity between $t = t_1$ and $t = t_2$ is

$$I_{t_1}^{t_2} = \sum \frac{I_o r_o^2}{c^2 t^2} (1 - \bar{\alpha})^{k_n}, \qquad (15.10)$$

summed for all reflections arriving between $t = t_1$ and $t = t_2$, where $k_n = |\ell| + |m| + |n|$. This sum was computed for a range of seat positions in the hall used above (a reverberation time of 2 seconds being assumed) and the results were compared with corresponding results from equation (15.8). A large upper limit t_2 results in long computation times; t_2 = 800 ms was chosen and since the theoretical energy arriving after t = 800 ms is 15 dB less than that arriving after t = 300 ms, an upper limit of 800 ms is equivalent to infinity in the discussion below of the validity of the theory. Lower limits between t_1 = 150 ms and 300 ms were investigated and again a source at (10,0,2) was used. The differences between the computed discrete and theoretical average integrated intensities are expressed in dB in Table 15.1 (first four columns).

TABLE 15.1

Differences between computed discrete and theoretical average integrated sound intensities, for various intervals

Receiver coordinates (x, y)	Difference in dB between computed discrete and theoretical average integrated intensity			
	150-800 ms	200-800 ms	300-800 ms	100 ms delay-800 ms
15, 2.5	-0.11	0.49	1.09	0.42
25, 2.5	0.40	0.54	0.70	0.39
35, 2.5	0.65	0.53	0.46	0.80
45, 2.5	1.11	0.52	0.38	0.51
15, 7.5	0.02	0.49	1.05	0.17
25, 7.5	0.31	0.59	0.70	0.28
35, 7.5	0.91	0.53	0.46	0.68
45, 7.5	1.15	0.55	0.38	0.53

It can be seen that the difference lies between -0.1 and 1.2 dB; thus for this hall t_c = 150 ms is a useable limit for t in equation (15.6). Since the travel time for the direct sound for the receiver at (45, 7.5, 2) is 104 ms, smaller values of t_c would not be acceptable. Expressed in terms of the hall's length, this limiting value corresponds to $t_c = L_x/c$.

In general one considers reflection sequences in terms of delay-time relative to the direct sound rather than travel time. The final column of Table 15.1 gives the difference values in dB for the interval between $t_1 = T_o$ + 100 ms and t_2 = 800 ms (where T_o is the direct sound travel time). Only for receiver positions near the source does this correspond to values of t smaller than 150 ms. Results for all receiver positions are, however, within the range 0 to +1 dB. Thus using 100 ms delay as a lower limit for t in equation (15.6) appears valid within 1 dB, the limiting delay corresponding to a delay of $2L_x/3c$.

It appears odd that the computed discrete integrated intensity in general exceeds the theoretical average value. One does not suspect that it is due to equation (15.1), as Figures 15.1-4 show no distinct bias between computed discrete and theoretical average numbers of reflections received, but rather a damped cyclic behaviour tending to zero error. Nor for $t > 200$ ms is $ctS/4V$, derived from the mean collision frequency, a biased estimate of k_n, the actual number of reflections experienced by a pulse (see the work of Schroeder |89| for information on the validity of the mean collision frequency formula in two dimensions). The bias appears to originate from the fact that k_n determines the power of $(1 - \bar{\alpha})$ in equation (15.10). Due to the concave curvature of the function $y = (1 - \bar{\alpha})$, the scatter of k_n about a mean value $(\bar{k}_n)$ results in a mean value for $(1 - \bar{\alpha})^{k_n}$ larger than $(1 - \bar{\alpha})^{\bar{k}_n}$, and hence positive dB differences between computed discrete and theoretical average values in Table 15.1. Hunt, in section VI of reference |88|, lists ways in which the general inequality between the product of averages and the average of products would lead to errors in a reverberation time formula. The bias mentioned above however is distinct from those mentioned by Hunt, and occurs when α is independent of direction and uniform for all room surfaces, with a magnitude at least as large as errors mentioned by Hunt.

15.5 REVERBERANT DECAYS FOR A RECTANGULAR ROOM WITH UNIFORM ABSORPTION

Schroeder has shown |21| that the ensemble average of an infinite number of noise decays is equivalent to an integrated intensity measured in reverse time, when the room is excited by an impulse. Equation (15.6) expresses this integrated intensity, measured in reverse time, and thus predicts an exponential sound decay for sound in a rectangular room according to geometric theory. The Eyring reverberation formula can thus be derived directly from equation (15.6). It can further be said that since t_c is a limit for the validity of equation (15.6), it will also be the limit for an exponential decay in a room with uniform absorption.

Reverberant decays were computed for eight seat positions in the hall for values of t up to 800 ms. To calculate the decay it is necessary to assume a base value for I for the range $t = 800$ ms to infinity. This is derived from equation (15.8) with the relevant value for the R.T. The hall is assumed to have a uniform absorption coefficient corresponding to an

Eyring R.T. of 2 seconds; however, the bias of results mentioned in section 15.4 indicates that the mean R.T. is larger than that according to the Eyring formula. For a series of R.T. values, computed decays were compared with linear decay behaviour according to equation (15.3). Best agreement occurs for an R.T. of about 2.19 seconds, a value about mid-way between the Eyring value (2 seconds) and the Sabine value (2.35 seconds). Decays for four positions along the hall are given in Figure 15.7; details of the initial decay are also included. The straight lines drawn through the plotted points have a slope equivalent to an R.T. of 2.19 seconds. It can be seen that the computed decays are exponential to within a very short interval after the direct sound (less than 50 ms). The divergence of the computed results from an exponential decay for $t \geq 100$ ms is between -0.8 dB and +1.1 dB.

It is interesting to note that the fluctuations about the linear decay correspond to behaviour which was noted previously in section 15.3. The fluctuations are again cyclic with a period of about $2L_x/c$, and best agreement occurs again for the seat half-way along the hall at (25, 7.5, 2).

15.6 TOTAL SOUND INTENSITY

In section 15.2 it was suggested that the analysis was not valid for small values of t, but the analysis when t was taken down to zero gave the classical value for the total sound intensity, $I_o^\infty = 4\Pi/A$. It is thus not surprising to discover in Figure 15.7 that there is a considerable spread in the value for the total sound intensity (as indicated by the initial horizontal section of the decays). The classical value does indeed represent a good average but computed values differ by as much as ± 3 dB from the classical value. This is true whether the total absorption, A, is taken as $S\alpha$ or $-S\,\ell n(1 - \alpha)$.

The computed values were also compared with the alternative classical formula which takes account of the direct sound:

$$I_o^\infty = \frac{\Pi}{4\pi r_o^2} + \frac{4\Pi}{R}, \qquad (15.11)$$

where R is the room constant, such that $R = \bar{\alpha}S/(1 - \bar{\alpha})$. The inclusion of the direct sound considerably reduces the disagreement between computed and predicted values to between -1.5 dB and 1.0 dB for the eight receiver

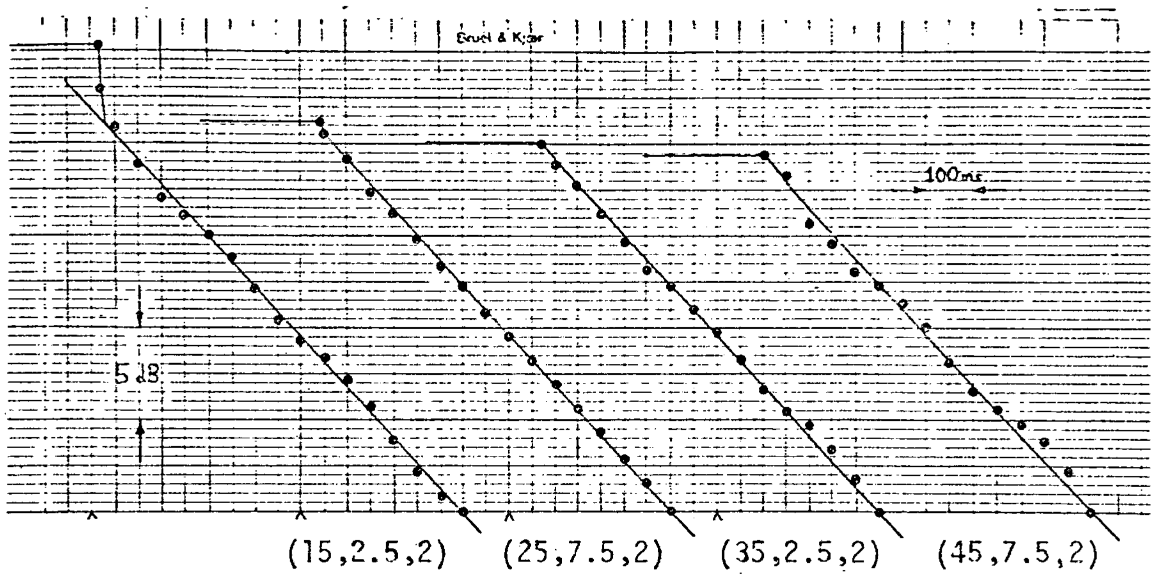


Figure 15.7. Computed reverberant decays for 4 receiver positions.

positions considered previously. Gibbs and Jones |87| have made a similar comparison for the case of a room with one wall highly absorbent, which gave similar differences between classical and computed values. Indeed in the study by Gibbs and Jones for the case of a single wall highly absorbent the agreement between the classical prediction and the computed values is much better than the agreement between either and the measured values.

To obtain an estimate of the total sound intensity, Doak |86| used an expression similar to equation (15.6) to represent the reverberant component. As the lower limit for t, he used one effective radius of the room, though in section 15.7 of reference |86| first reflections are also considered. Considering the reverberant component constant throughout the room leads to results similar to those recorded in the previous paragraph. Examinations of the decay curves in Figure 15.7 suggests the use of $t = T_o$, the direct sound travel time ($T_o = r_o/c$), as the lower limit of integration. Thus the total sound intensity can be considered as the sum of the direct sound intensity and the integrated energy from time T_o to infinity, a linear exponential decay being assumed, so that

$$I_o^{\infty} = \frac{\Pi}{4\pi r_o^{\,2}} + \frac{4\Pi}{-S\,\ell n(1-\alpha)} \cdot e^{[S\,\ell n(1-\alpha)\,.\,r_o]/4V.} \qquad (15.12)$$

The difference between the computed values and those predicted by equation (15.12) lie between +0.3 dB and +1.0 dB, a significant improvement on the classical equation (15.11). One can expect that equation (15.12) is valid, insofar as one is dealing with situations with an exponential decay.

15.7 INTEGRATED INTENSITY IN A RECTANGULAR ROOM WITH AN ABSORBENT FLOOR

Until now the discussion has been confined to situations where reflections at surfaces are such that the assumption of specular reflection is valid. Morse and Bolt |90| have shown that the reflected wave from a plane absorbent surface is no longer spherically symmetric so that to assume a simple point image is no longer valid.

In the case of a concert hall the audience provides a highly absorbent surface while the remaining surfaces are generally hard. As the absorption coefficient of the audience is about 0.9 at mid-frequencies |91|, the contribution of sound energy reflected off this surface will be minimal, and errors due to assuming a simple point image can probably be neglected in calculations of integrated intensity.

As a theoretical model the case of a completely absorbent floor will be investigated by using a procedure identical to that used in section 15.2 (though air absorption will be ignored here). With a fully absorbent floor images of the source only exist in two planes parallel to the floor, for n = 0 and 1. Using equation (15.9) to give the number of reflections received from the plane n = 1:

$$N = \frac{\pi c^2 t^2}{L_x . L_y} + \frac{\pi c^2 (t^2 - t_1^2)}{L_x . L_y}, \text{ for } t > t_1, \qquad (15.13)$$

where t_1 is here the travel time of the pulse from the ceiling image to the source. If S' is the floor surface area (= $L_x . L_y$), then

$$N = \frac{2\pi c^2 t^2}{S'} - \frac{\pi c^2 t_1^2}{S'}$$

and

$$\frac{dN}{dt} = \frac{4\pi c^2 t}{S'}. \qquad (15.14)$$

For reflections in the n = 0 plane, the relevant mean collision frequency is the two-dimensional one: $L'/\pi S'$, where L' is the perimeter of the floor surface. No simple formula will express the mean collision frequency for ceiling reflections from the n = 1 plane, though using $L'/\pi S'$ as an approximation proves to be a valid assumption. The pulse also experiences spherical divergence, so that the net intensity of the n^{th} pulse is

$$I_n = \frac{I_o r_o^2}{c^2 t^2} (1 - \alpha')^{ctL'/\pi S'}, \tag{15.15}$$

where α' is the mean absorption coefficient of the wall and ceiling surfaces. Thus

$$I_t^{t+\delta t} = \frac{4\pi c^2 t}{S'} \cdot \frac{I_o r_o^2}{c^2 t^2} .(1 - \alpha')^{ctL'/\pi S'} .\delta t \tag{15.16}$$

and

$$I = \frac{4\pi I_o r_o^2}{S'} \int \frac{1}{t} \cdot e^{(ctL' \ln(1-\alpha')\ /\pi S'} .dt, \tag{15.17}$$

or

$$I = \frac{\Pi}{S'} \int \frac{1}{t} \cdot e^{ctL' \ln(1 -\alpha')/\pi S'} .dt. \tag{15.18}$$

The solution of this integral can only be expressed in series form. For computation however it proves quicker to treat the integral as a sum over small intervals δt.

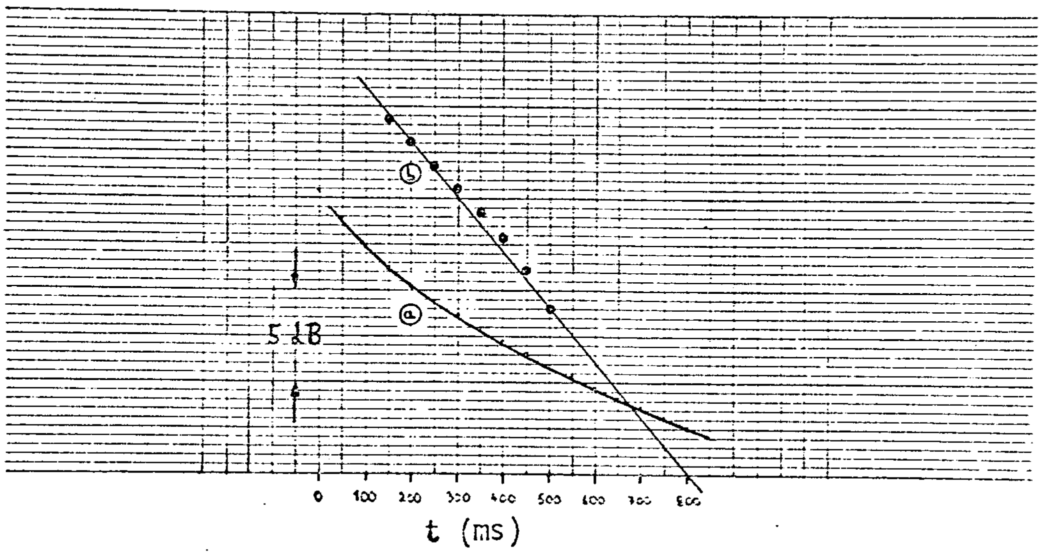


Figure 15.8. Theoretical reverberant decays for absorbent floor model. (a) Theoretical decay from equation (15.18); (b) theoretical decay assuming diffuse conditions for t > 500 ms (see text).

The theoretical behaviour of the quantity I_t^∞ was computed for the hall used previously; the value of α' used was that which would give a 2 seconds R.T. according to the Eyring formula with $\alpha = 1$ for the floor. Figure 15.8, curve (a), shows this theoretical decay for the absorbent floor model, which is predictably non-linear. The discrete computed decays, again computed by using equation (15.10), with $\alpha = 1$ for the floor surface, agree well with the average theoretical value from equation (15.18) for $t \geqq 150$ ms; agreement is at least as good as it was for the uniform absorption case (this agreement validates the mean collision frequency assumption above). Agreement within these limits is also achieved when the discrete computed decay is calculated with $\alpha = 0.9$ for the floor surface and α' is again chosen to give an Eyring R.T. of 2 seconds. Thus equation (15.16) is valid for $t \geqq 150$ ms, or L_x/c, for the case of a plane-walled rectangular hall with a highly absorbent floor but no internal diffusing surfaces.

15.8 SOUND DIFFUSION IN GEOMETRIC THEORY

In the geometric image model a diffuse sound field will exist when the images of the source are regularly distributed in the three-dimensional image space. Such a situation exists when the walls have a uniform absorption coefficient, but not when one wall surface is totally absorbent. Hunt |88| has provided an elegant demonstration of the diffuse state of the reverberant sound field for a rectangular room.

As mentioned above, if the value of the integrated energy from time t to infinity (I_t^∞) is an exponential function of t, as in equation (15.6), the reverberant decay will be exponential. Inspection of the theoretical derivations of I_t^∞ for both the uniform absorption and absorbent floor case indicates that, for an exponential decay, it is necessary that dN/dt is proportional to t^2. So to obtain an exponential decay with a highly absorbent floor it is necessary to introduce diffusing surfaces such that source images are regularly distributed in the hemispherical image space above the plane of the floor. According to this view the reverberation time will only be a function of the absorption coefficients of the surfaces other than the floor and the mean collision frequency for reflections off these surfaces. The correspondence between diffuse conditions and exponential decay behaviour is thus apparent, though, with a fully absorbent floor, diffusion only exists, of course, over half the solid angle.

Introduction of diffusing surfaces in the above theory presents considerable problems. For computation probably the only feasible approach is to consider the source to be radiating in a finite number of directions and to follow the diverging sound "rays" around the room until they strike the audience area. This approach has been used by Krokstad, Strøm and Sørsdal |83| for a series of room shapes and by Kuttruff |84| for the case of randomly diffusing walls. The conventional geometric assumption of specular reflection is, however, still required for these studies, so that results are directly applicable only over a limited frequency range.

An interesting situation arises if the decay for the absorbent floor case is calculated by using for the base value for I between 500 ms and infinity that given by equation (15.6). This is an approximation to the case of a rectangular hall with a highly absorbent floor and a limited amount of diffusing surface. Only a long time after sound is emitted from the source will the majority of reflections arriving at the receiver have experienced sufficient reflection off the diffusing surfaces to make the sound field diffuse. It may be assumed here that this state has been reached for $t > 500$ ms (after about 10 reflections on average), whilst for smaller values of t, when only the minority of reflections have experienced reflection off the diffusing surfaces, equation (15.18) is more relevant for the average behaviour. This treatment gives curve (b) in Figure 15.8 (the straight line is the decay according to equation (15.6)). This shows deviations from the exponential decay up to $t = 150$ ms of less than 1 dB, although conditions in that period are far from diffuse. In a real situation, of course, the transition to diffuse conditions is gradual; however, since in both the diffuse and non-diffuse case the values for the integrated intensity are very similar over the interval $150 \text{ ms} < t < 500$ ms, the analysis suggests a situation in which an exponential decay can occur without diffuse conditions prevailing. This behaviour is naturally characteristic only of a certain range of R.T. values for a particular size hall; that it occurs for the average size hall and typical R.T. value used in this study suggests that such behaviour may prevail in some concert halls.

15.9 CONCLUSIONS

The geometric image model predicts that, for a rectangular room with uniform absorption, the integrated sound intensity is equivalent to the sound intensity for an exponential decay (with an R.T. according to the

Eyring formula) with as base value the classical total sound intensity. Comparison of this theoretical average prediction with values computed by considering discrete images showed that the theoretical prediction is valid for values of t, the travel time, greater than L_x/c (where L_x is the length or longest dimension of the hall) or typically greater than 150 ms. Deviations between theoretical average and discrete computed values were cyclic in nature, being smallest for receiver positions near the centre of the hall. This apparent "bunching" of reflections, which results in discrepancies from average values of the number of reflections received of the order of 10%, is responsible for deviations in the reverberant decay of less than 1 dB from exponential behaviour. Since a 1 dB fluctuation is a common occurrence in a measured reverberant decay, such predicted behaviour is only likely to be detectable in an ideal acoustic room which is purely rectangular, even when interference effects occurring in the real situation are ignored.

Computed "reverberant decays" were found to be close to exponential but to have a decay rate 8.5% less than the Eyring value. This was attributed to the fact that due to the concave curvature of the function $y = (1 - \bar{\alpha})^{\lambda}$, then $\overline{(1 - \bar{\alpha})^{\lambda}} > (1 - \bar{\alpha})^{\bar{\lambda}}$. In the derivation of the Eyring formula, a mean quantity $\bar{\lambda}$ based on the mean collision frequency is used so that discrete computed decays will always have decay rates less than the Eyring value. Computed decays proved to be exponential from delays relative to the direct sound of 50 ms or less. Such an early onset of reverberant conditions is surprising, but in a real concert hall such uniform absorption conditions do not prevail.

A theoretical investigation of the situation with a fully absorbent floor surface, similar to the concert hall situation, predicted a non-exponential decay, the anticipated result. In the real concert hall diffusing surfaces are also present, which in practice generally result in an exponential decay of the reverberant sound. In terms of the geometrical image model, images of the source must be regularly distributed in the image space for an exponential decay to result, a condition which guarantees diffuse conditions. Unfortunately this theoretical method is unsuitable for dealing with the typical concert hall situation, with diffusing surfaces and a highly absorbent floor. Such situations could only be treated theoretically if some measure of the amount of diffusing surface were introduced; in any case most halls are sufficiently complex in form to warrant computer analysis via ray tracing techniques, or model studies, to determine behaviour prior to diffuse conditions. Once a state of diffusion exists in a room, geometric theory predicts no more than the generally assumed exponential decay.

Predictions based on a geometric image model for the first reflections will be frequently quoted in the remainder of this thesis. The exercise described in this chapter serves to illustrate the results of continuing the analysis for longer time periods than the early sound. It indicates the role of diffusing surfaces in providing a state of diffusion and an exponential decay in a rectangular room with an absorbent floor, whilst the assumption of specular reflection is retained. In fact examination of oscillograms in Chapter 17 indicates that reflections not long after the direct sound are in general no longer specular, though the above analysis indicates the well known result that non-specular reflections are not a necessary condition for exponential decay.

As regards quantitative results derived here which might be compared with measurement, the equations for the total sound intensity (15.12) and for the reverberant sound relative to the direct (15.8) provide a useful reference point for behaviour in real halls (for instance, in the case of the latter, in a study of the "focussing" effect due to stage enclosures affecting reverberant sound energy). Such a comparison was not in fact pursued for either result, since they did not appear to be subjectively significant quantities relevant to the main preoccupations of the thesis. Equation (15.8), however, was used as a reference prediction for measurements of temporal energy fractions, but in fact for this one is assuming no more than an exponential decay with a rate corresponding to the reverberation time.

Chapter 16

THE MEASURING SYSTEM FOR TESTING REAL HALLS

16.1 INTRODUCTION

In Part I it was established that the ratio of lateral to non-lateral early energy is related to the subjective quality, which has been called here spatial impression. The only other reported measurement of the proportion of early lateral sound was made by Schroeder et al. |39|. Schroeder compared the signal received by a directional microphone, with solid angle of pick-up of 45°, with that received by an omni-directional microphone. The directional microphone faced the nearest side wall and early sound within 50 ms of the direct sound was considered. What is surprising about this measurement is the small angle of pick-up of the directional microphone, corresponding to a semi-apex angle for a cone of only 29°. A computer investigation, assuming specular reflection, showed that, for a hall of roughly the dimensions of that investigated by Schroeder, only at 8% of the seats did sound arrive within 50 ms within a lateral 45° solid angle of pick-up! In any case correlation between Schroeder's measure and the ratio of lateral to non-lateral early energy as defined by equation (14.1) is certainly not very high.

A measuring system was therefore built to measure the ratio of lateral to non-lateral early energy. This system also enabled measurements to be made of the effect of audience attenuation filtering on the early sound, when integrated over a 50 or 80 ms period. Again, reported measurements of this effect are few.

The results reported in Part I indicate subjective behaviour related to incoherent addition (see also section 19.2). A true measure of this is only possible with a Dirac pulse as a test signal, which is of course unobtainable in practice. A finite duration pulse offers the possibility of interference between reflections off different room surfaces. To minimise this a short duration pulse should be used, which is in addition necessary for high temporal resolution.

The measuring system comprises a source which emits a sound pulse into the enclosure. The response at a particular point in the enclosure is picked up by a microphone; the microphone signal is temporally sampled with a gating circuit, and then fed into an "energy meter", which measures the energy content of the sampled response.

16.2 THE TEST SIGNAL.

For the measurement of oscillograms performed at Göttingen (described in section 2.1(a)) an electric discharge was used to produce the acoustic pulse, which had a duration of only 0.4 ms. Such a pulse contains, however, little low frequency energy, and it is therefore not suitable for analysis of behaviour at bass frequencies, due to a poor signal-to-noise ratio at these frequencies.

Another approach is to use a tone pulse radiated from a loudspeaker. By using a tone pulse of the required bandwidth maximum use can also be made of the power handling capacities of the loudspeaker. However, to obtain a particular bandwidth, a certain duration sound is required. To obtain an optimum behaviour (i.e., a minimum of the product of duration and bandwidth) a Gaussian tone pulse is used (i.e., the envelope of the tone pulse is proportional to a function of the form e^{-t^2}). This has been employed by Kürer |30|. The pulse durations required to produce suitably small bandwidths, say 1/3 octave, are relatively large.

Atal et al. |92| have used a tone pulse with a cosine-form envelope with the characteristic that the height of subsidiary maxima in the spectrum is very low (more than 40 dB below the principal maximum). Such a tone pulse with centre frequency 500 Hz and bandwidth approximately 1/3 octave has a duration of 24 ms, though 80% of the total energy of the pulse is contained in a period of 10 ms.

The conflict between bandwidth and duration can be overcome by filtering the recorded response, and whilst this would considerably reduce the temporal resolution if the filter were placed after the microphone, by placing it after the sampling gate no such reduction in resolution occurs. The signal bandwidth, however, remains significant, since to obtain good signal-to-noise ratio in the enclosure optimum use of the power handling capacity of the loudspeaker is desirable. Too short a duration tone pulse also will not provide sufficient energy for measurements to be made. The choice of a single cycle of the measuring centre frequency proved a satisfactory compromise. As the source used for these measurements employed standard loudspeakers with an unexceptional transient response the loudspeakers modified the single cycle pulse. A typical resulting test pulse from a 630 Hz single cycle is illustrated in Figure 16.1, together with its spectrum.

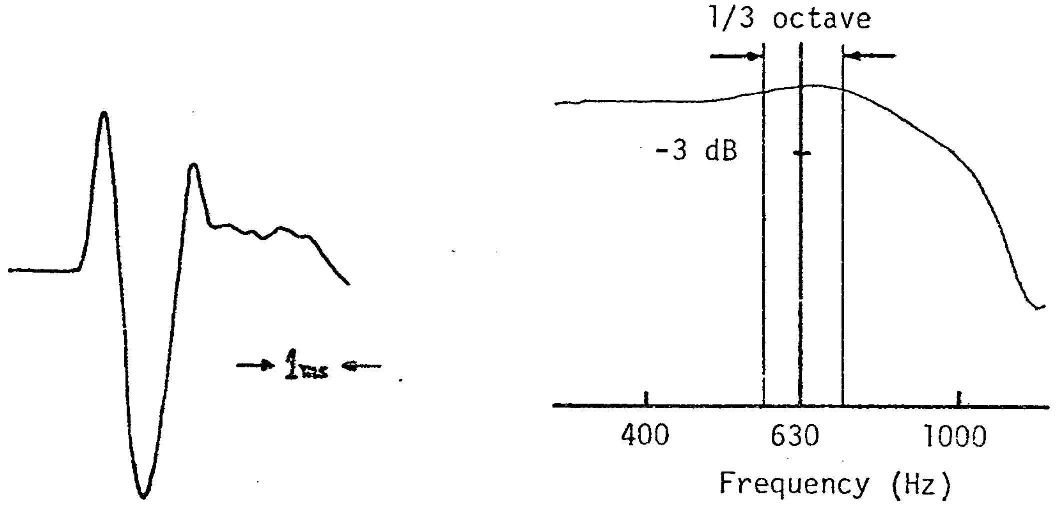


Figure 16.1. Typical test signal and its spectrum, centre frequency 630 Hz.

The degree of interference that occurs between two reflections is best illustrated by the autocorrelation function of the test signal. This is demonstrated below:

The integrated energy in two reflections which do not interfere is

$$\int p^2(t - t_1).dt + \int p^2(t - t_2).dt,$$

where t_1 and t_2 are the arrival times of the reflections. For convenience the amplitude of the two reflections is assumed the same.

When two reflections interfere, the recorded pressure is the sum of the individual pressures, so the integrated energy becomes

$$\int \{p(t - t_1) + p(t - t_2)\}^2\, dt$$

$$= \int p^2(t - t_1)dt + \int p^2(t - t_2)dt$$

$$+ 2.\int p(t - t_1).p(t - t_2).dt.$$

The final term is due to the interference of the two reflections and as a fraction of the reflection energy without interference it is identical to the normalised autocorrelation function for the test signal at delay $(t_2 - t_1)$. Figure 16.2 shows the autocorrelation functions of a single cycle sine wave and a Gaussian tone as used by Kürer |30| (Δt = 2 ms); for simplicity a centre frequency of 1 kHz has been chosen.

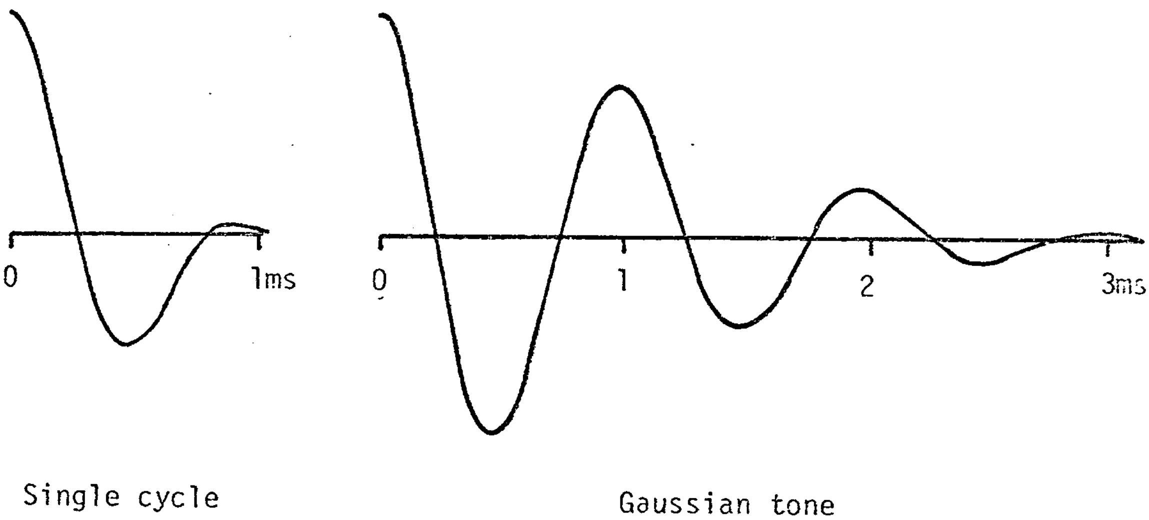
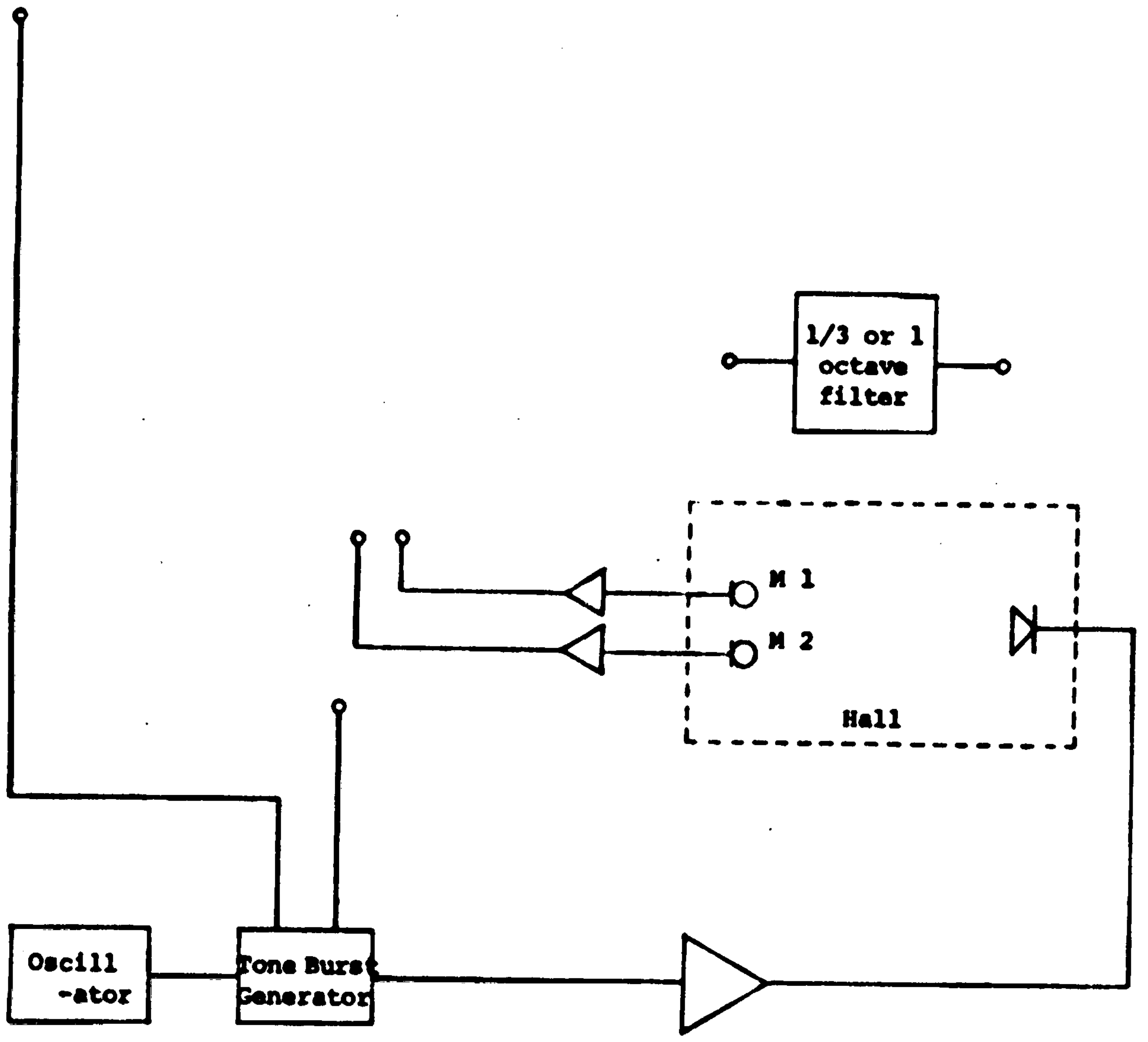


Figure 16.2. Autocorrelation functions of a single cycle (1 kHz) sine wave and a Gaussian tone pulse (centre frequency 1 kHz).

For a random temporal distribution of reflections, the interference effects will cancel out if the areas above and below the zero line in the autocorrelation function are equal. However, when working with the early sound arriving shortly after the direct sound, reflection density is not high enough for interference effects to become randomised, and a minimum value for the autocorrelation function for all delays other than zero is desirable. The superiority of the single cycle pulse in this respect is evident.

16.3 THE GATED INTEGRATED ENERGY METER

The central component of the measuring system is a gated integrated energy meter (GIEM), which was built for these measurements. The GIEM allows any desired temporal segment of the signal, received by a microphone, to be sampled and the integrated sound energy in this segment to be displayed on a meter. Since all measurements of the integrated energy were relative, provision was made to switch readily from one sampling length to another or



Auxiliary equipment used with GIEM.

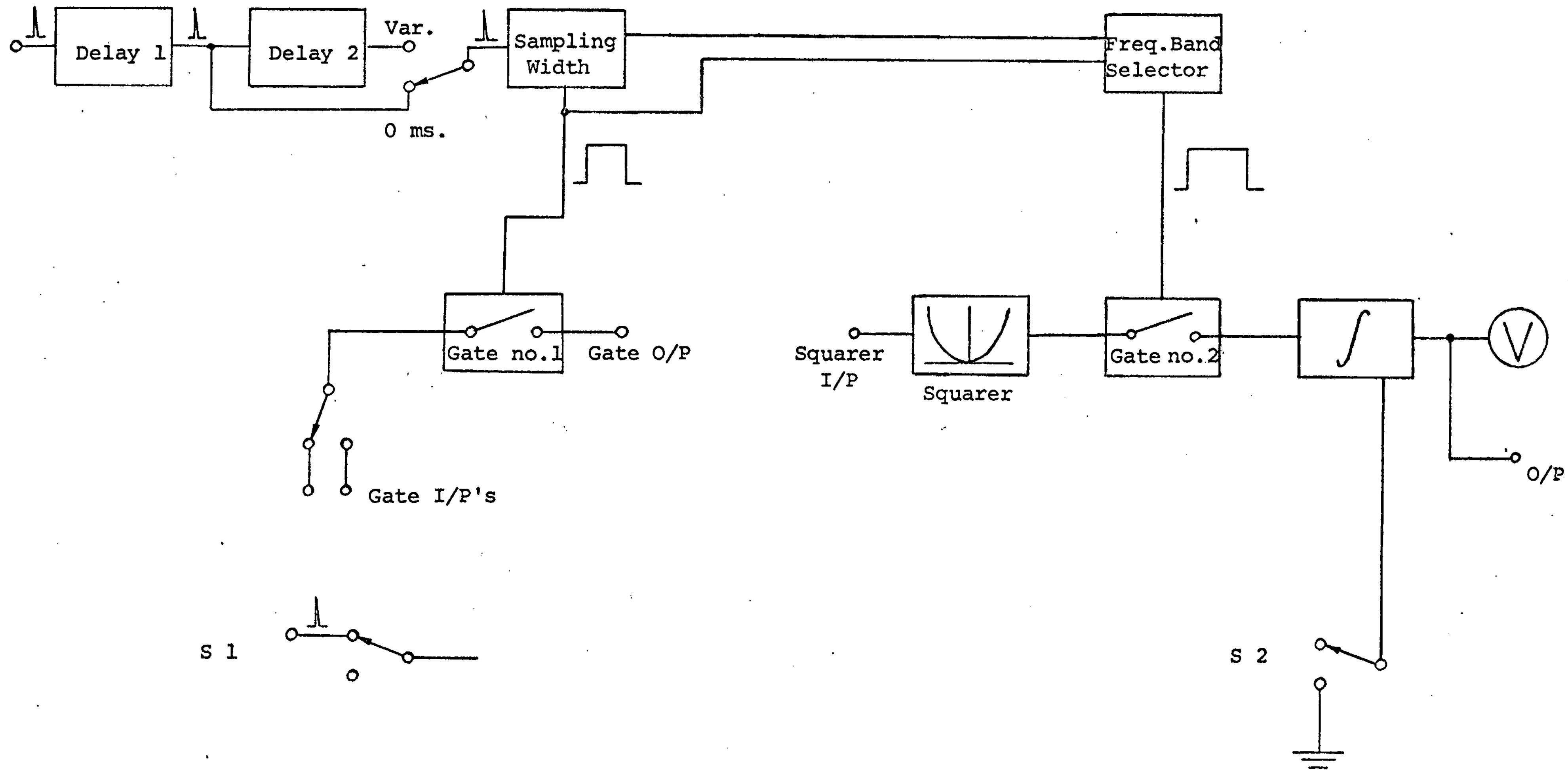


Figure 16.3. Block diagram of Gated Integrated Energy Meter

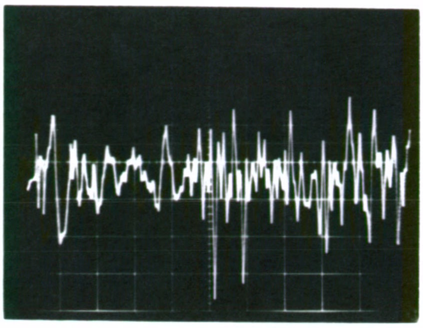
to compare the response at one microphone with another. By comparing the signal received by a directional and an omni-directional microphone, the proportion of sound from a particular direction is obtained. The block diagram of the gated integrated energy meter is given in Figure 16.3. The auxiliary equipment and the method by which this is connected to the GIEM is illustrated on the translucent sheet above Figure 16.3.

The mode of operation is as follows:

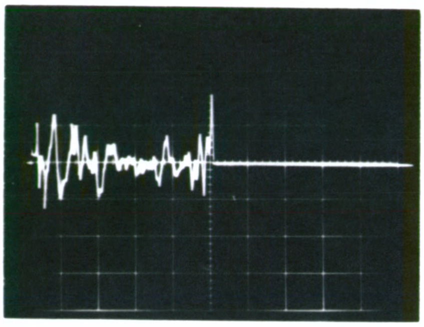
An oscillator is connected to a tone burst generator. Depressing and releasing switch S1 of the GIEM produces a trigger pulse for the tone burst generator; burst lengths of 1, 2, 4, ..., 32 cycles were available (generally a single cycle tone pulse was used). The tone pulse is radiated from an omni-directional loudspeaker at a typical source position in the hall. The signal received by the microphone is fed to gate no. 1. The initial sampling time is determined by two pulse delay circuits. Delay 1 is generally set such that the delay corresponds to the direct travel time of the pulse, so that with Delay 2 = 0 ms, the initial sampling time corresponds to zero delay at the receiver position. By altering Delay 2 the initial sampling time can be varied relative to the direct sound arrival time. The sampling width can be varied in a series of fixed steps or with a continuously variable control.

The gated microphone signal is then fed to a Brüel and Kjaer Spectrometer to provide 1/3 or single octave filtering of the signal. This has a dual function of greatly improving the signal-to-noise ratio and of limiting the measurement to the required frequency band, irrespective of the bandwidth of the radiated tone pulse. The filtered pulse is then squared to give a signal proportional to the sound energy rather than the sound pressure. The squared signal is fed to an integrator via gate no. 2. The gate is closed simultaneously with gate no. 1, but due to ringing of the filter it is necessary to leave the gate closed longer than gate no. 1 for a time depending on the frequency band being measured. The integrated energy is displayed on a moving-coil meter switchable over six ranges. Depressing switch S2 resets the integrator for the next measurement. Since S1 and S2 are coupled, depressing and releasing the switch is all that is required to repeat a measurement.

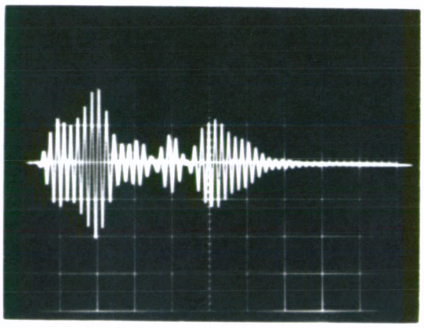
The operation of the GIEM is illustrated by a series of oscilloscope pictures in Figure 16.4 of a measurement made in a hall of the energy



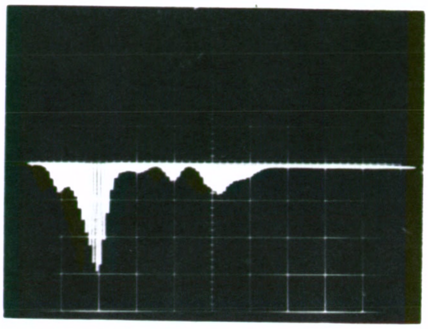
1. Microphone signal.



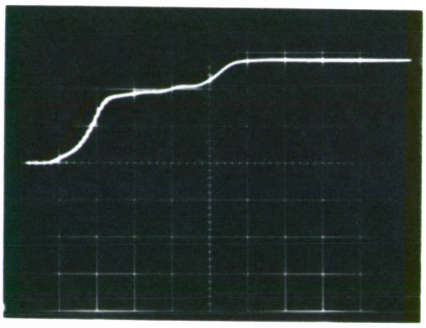
2. Gated microphone signal.



3. Gated signal after 1/3 octave filtering.



4. Squared filtered signal.



5. Integrated signal.

Oscilloscope traces illustrating operation of GIEM
- 500 Hz. signal sampled for 50ms. (starting time
identical for each trace, 10ms/div. sweep rate).

Figure 16.4.

received at a receiver position in the first 50 ms after the onset of a direct sound pulse at a frequency of 630 Hz. The meter indicates the final integrator voltage which is held indefinitely. Figure 16.6 contains a photograph of the actual instrument.

In operation the GIEM proved to be as convenient as might be expected. Problems were encountered due to voltage drift on the output of the squarer which was then integrated, giving spurious results. It is possible that a passive squaring device would be more suitable than the integrated circuit multiplier used in this instrument. However, frequent adjustment of the squarer output offset voltage and using maximum squarer input voltages reduced such errors to a negligible minimum. Linearity proved to be very good; for a 35 dB input range to the squarer, the output response was within 1 dB (with output range of 70 dB). The maximum measuring range for the meter is about 35 dB, and linearity of the squarer-integrator-meter combination over a 20 dB input range, was within 0.1 dB. Extensive modification of the instrument would be required to automate measurements. As built, it offered great flexibility in the range of measurements for which it could be used.

It will be noticed that the latter section of the circuit of the GIEM is identical to the circuitry required for measurement of the reverberation time according to the integrated impulse method after Schroeder |21|. Facility for leaving gate no. 2 closed was included, though again drift at the squarer output limited the dynamic range.

16.4 THE OMNI-DIRECTIONAL SOURCE

Musical instruments are generally directional; not only has each instrument a different directional characteristic but also the characteristic varies very significantly with frequency |93, 94|. However, for a whole orchestra, in which instruments are generally playing in groups, sound is not radiated in any preferred direction, so it is best to use an omni-directional transducer |95|.

An omni-directional source was constructed by using a rectangular box containing five 150 cm (6 in) diameter double-cone loudspeaker units mounted on five sides of the box as illustrated in Figure 16.5. Though the anechoic facility available for measurements was small, rotating an omni-directional

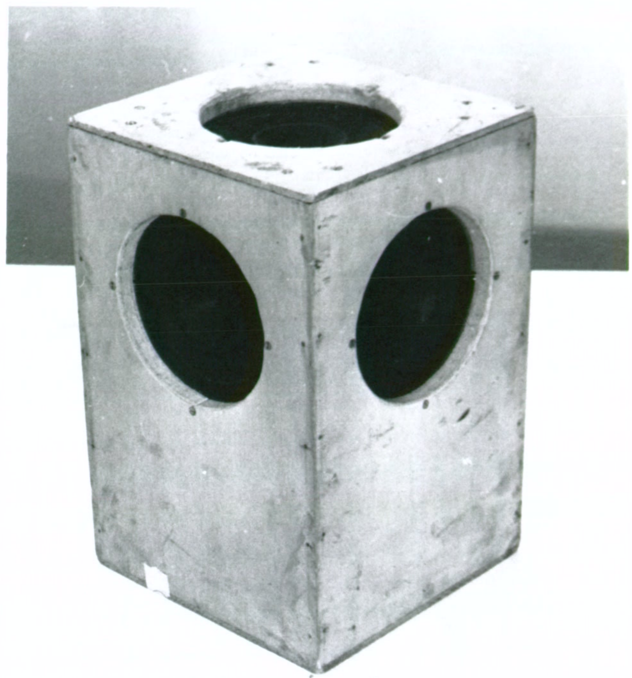


Figure 16.5. Omni-directional Loudspeaker

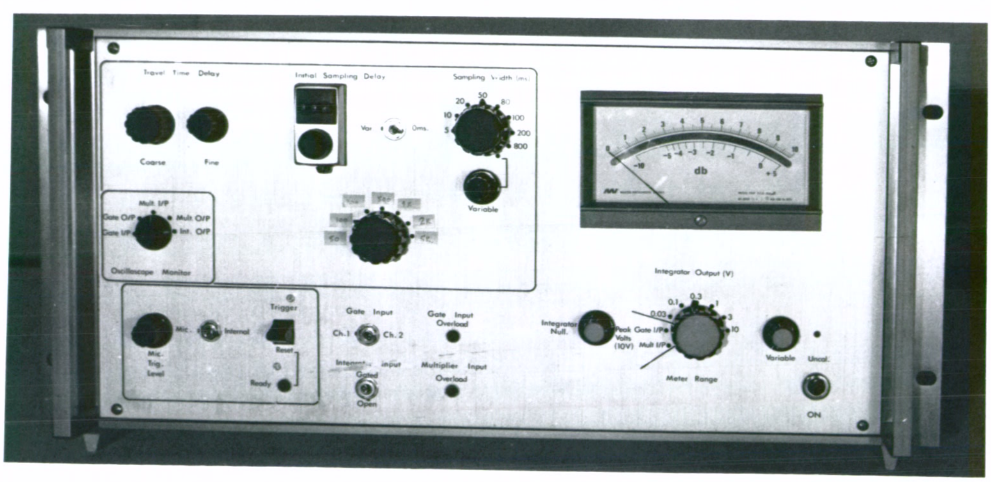


Figure 16.6. Gated Integrated Energy Meter

source about its centre causes room effects to be almost constant. The deviation from omni-directional is listed in Table 16.1, for both a horizontal scan and a vertical scan.

TABLE 16.1

Deviation from omni-directionality of loudspeaker

Frequency (Hz)	Maximum deviation from mean (dB)	
	Horizontal Plane	Vertical Plane
400	1.0	1.9
500	1.5	1.3
630	0.4	1.3
800	0.8	(3.0)
1000	1.3	(2.8)
1250	3.3	-
1600	3.5	-
2000	1.5	-

It would seem that room effects were partially responsible for the large deviations from the mean in the vertical plane at 800 and 1000 Hz, since measurements opposite the two side loudspeakers differ by more than the maximum variation in the horizontal plane. For vertical plane measurements the source was not omnidirectional since the sixth side of the loudspeaker box contained no transducer. For this reason these values have been placed in parentheses. Omni-directionality was considered suitable for 1/3 octave measurements up to 1000 kHz, though the 2kHz response was also suitable.

16.5 THE FIGURE-OF-EIGHT MICROPHONE

Whilst many measurements were made with an omni-directional microphone, the figure-of-eight response was also used to obtain a measure of the lateral sound. A stereo microphone with variable directionalities (changed by altering the polarisation voltages) proved a very flexible system; a Neumann SM69 FET microphone was used. Figure 16.7 shows the published directional response for the figure-of-eight characteristic for this microphone (the dotted lines refer to 8 and 12.5 kHz, far above the frequencies used for these measurements). The solid line is the characteristic for all frequencies below 4kHz, and illustrates the particular virtues of the figure-of-eight

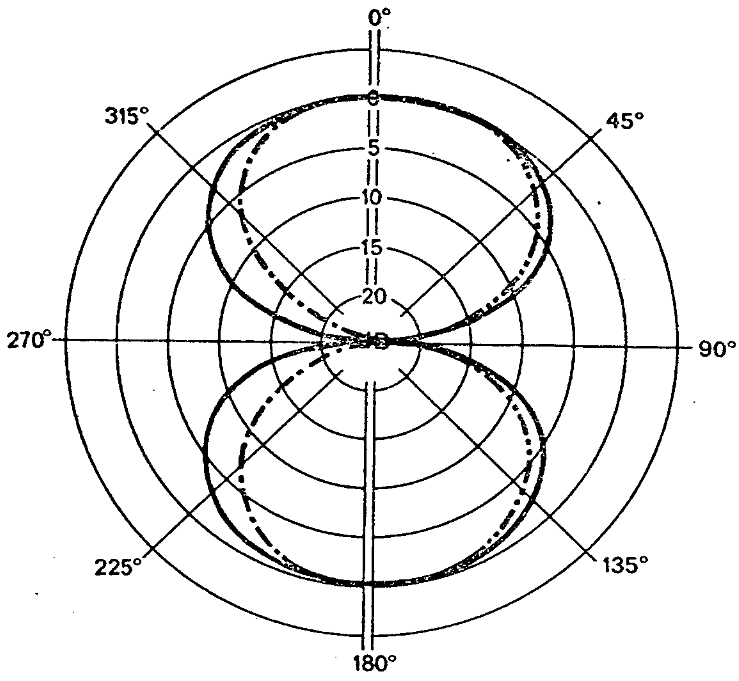
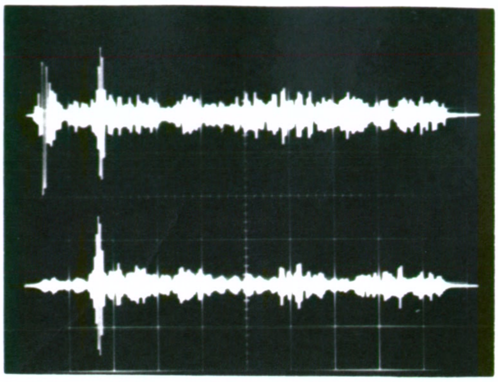


Figure 16.7. Directionality response of Neumann SM69 microphone in the figure-of-eight mode. Solid line: 4 kHz and below.

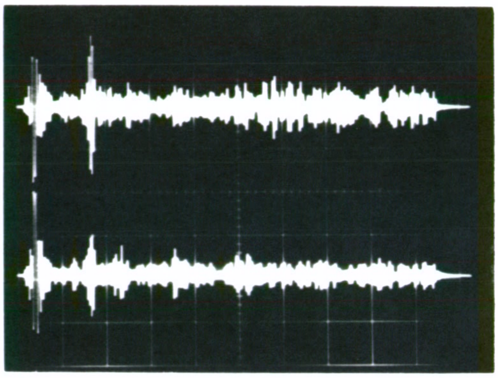
directionality response: i.e., not only is it bi-directional but also the response is independent of frequency. For computation based on this response, it was found that the response closely corresponded to two identical ellipses, the correspondence being at least as good as the variation between the two lobes.

The theoretical directionality pattern of a figure-of-eight microphone is proportional to the cosine of the angle of incidence, though the response in Figure 16.7 only approximates to this behaviour. This is, however, a cosine response in terms of voltage, whereas the response for contribution to lateral sound for spatial impression was found in Chapter 8 to be a cosine response in terms of energy. In terms of energy the figure-of-eight microphone has a cosine squared response. The implications of this unavoidable discrepancy will be discussed in section 19.7.

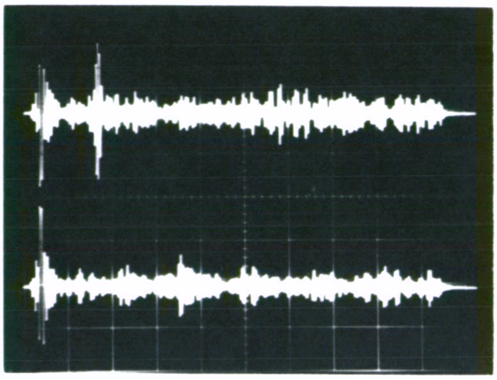
A further discrepancy from the desired response also arose due to the difference between the frequency responses of the figure-of-eight and omnidirectional characteristics. In the absence of suitable anechoic facilities, this was measured in a large room by using 20 ms pulses and a short source-receiver distance, the microphone pulses and energies being gated and



Omni-directional and directional (frontal suppressed) oscillograms.



Omni-directional and directional (90° lateral suppressed) oscillograms.



Omni-directional and directional (lateral reflection suppressed) ocsillograms.

10 ms/div. sweep rate

Figure 16.8.

measured with the GIEM. In this way reflections off neighbouring surfaces are excluded. The relative insensitivity at low frequencies for the figure-of-eight characteristic quoted in the published data was found to be as much as 5.6 dB at 80 Hz. Such a response can, however, be accounted for in measured results by applying a correction coefficient. But it was also found that the maximum responses on both sides of the figure-of-eight were not identical, differing by as much as 1.8 dB. This introduces an error term in measurements when using this microphone of up to 0.9 dB. It is likely that a single ribbon microphone would have been more suitable in this respect.

The operation of the directional microphone is illustrated in Figure 16.8 by the oscillograms taken with an omni-directional response and figure-of-eight response, with the null point set such that the direct sound, 90° to the direct sound and the first lateral reflection are suppressed. The use of this microphone in determining the direction of a particular reflection is evident from these oscillograms.

Chapter 17

THE THREE HALLS USED FOR MEASUREMENTS

17.1 INTRODUCTION

As was mentioned in the introduction of the previous chapter, the principal motive in conducting measurements in real halls was to determine the behaviour in halls of physical quantities which in Part I were found to correlate with subjective spatial impression. Since inevitably the number of halls available for tests was severely limited, it was also hoped to be able to validate computer predictions, which could be made for a much larger range of hall shapes. It was further hoped that it might be possible to correlate subjective impressions of halls with measurements made in them, though in the absence of a systematic subjective assessment, any such conclusions should be treated with circumspection.

The author had the opportunity to experience two halls, for which Dr. A.H. Marshall was the acoustical consultant, at the time of their opening. This offered the opportunity of experiencing the acoustics at a lot of different seat positions in a way that would be much more difficult during commercial operation of the hall. The halls were the Christchurch Town Hall in New Zealand and the Perth Concert Hall in Western Australia, which were opened on 30 September 1972 and 26 January 1973, respectively. The third hall in which measurements were made was Winthrop Hall, in the University of Western Australia, which was used for symphony concerts in Perth until the opening of the new hall.

17.2 CHRISTCHURCH TOWN HALL

Contrary to what the name might suggest, this hall is a large concert hall, designed primarily for musical performance; it forms part of a civic building complex containing a theatre and banqueting hall. A photograph of the interior is contained in the frontispiece and a plan of the hall is contained in Figure 17.2. The hall has a volume of 20,700 m^3 and seats an audience of 2,377, plus a choir of 400 and orchestra of 120.

The hall has a flat central floor area surrounded by raked seating underneath the gallery. The gallery, again raked, extends the total circumference of the hall, becoming the choir seating area at the orchestra end. The roof

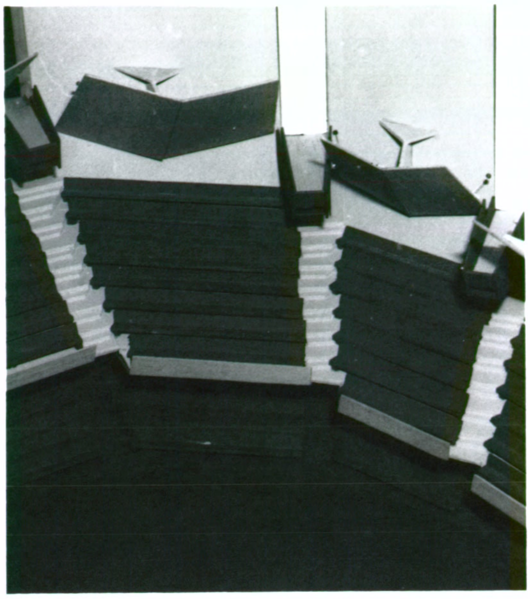


Figure 17.1. Detail of model showing dihedral reflectors

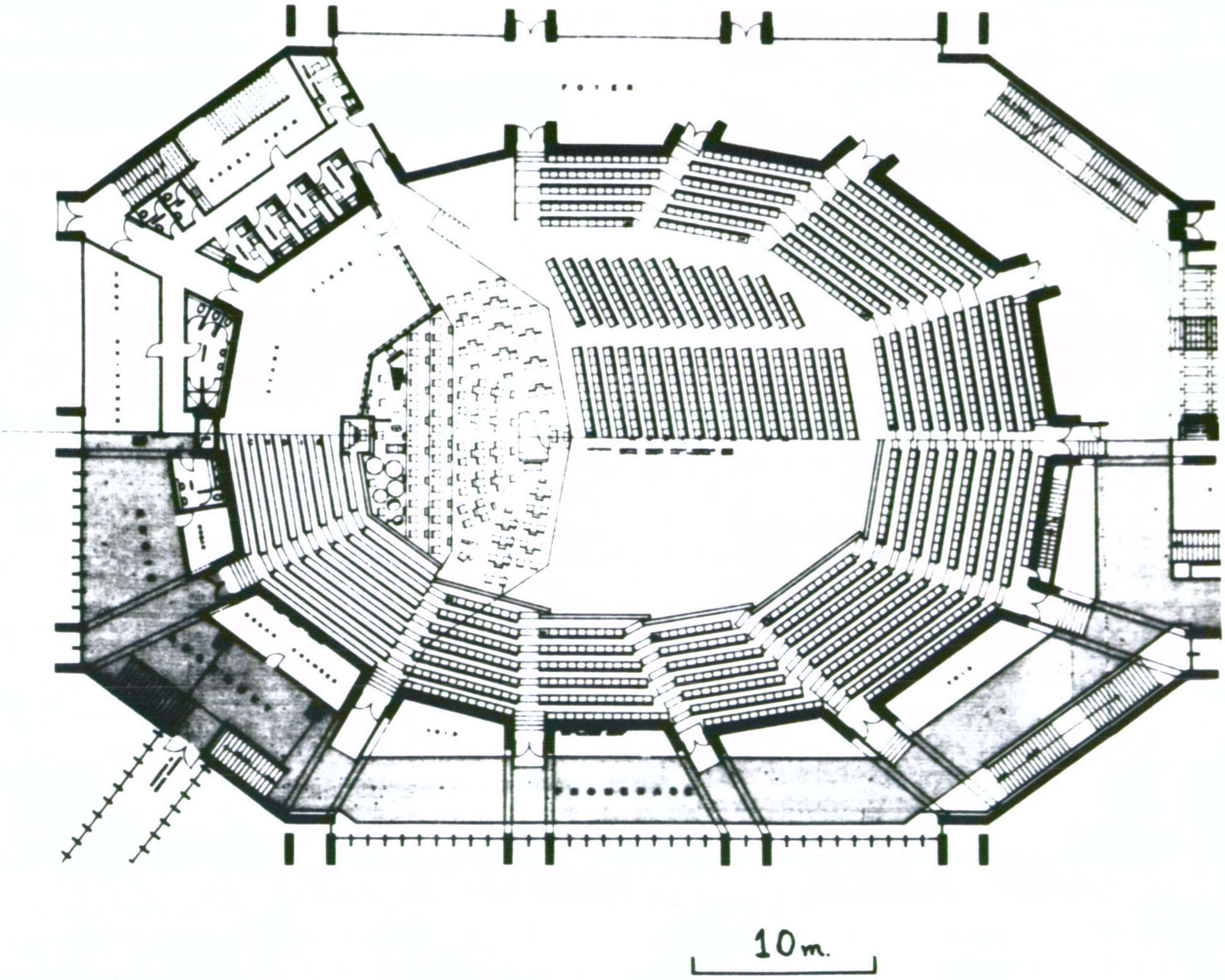


Figure 17.2. Christchurch Town Hall, New Zealand.

is made of three plane surfaces, the central one slightly sloping at about 20m above the central floor area, the surfaces at the front and back of the hall sloping more sharply.

The unique feature of this hall, however, is the elliptical plan, constructed of 14 plane sections. The virtue of this shape from a visual point of view can be appreciated to a certain extent in the photograph in the frontispiece which is taken from the gallery area at the back of the hall. Visually one tends to "convert" the shape into a circular one, so that from the gallery one feels much closer to the orchestra than one really is. This provides a much greater sense of physical involvement with the performance than occurs in many halls of equivalent length (46m). From an acoustic point of view, however, the elliptical plan is dangerous since sound from a source at one focus of the ellipse is all reflected towards the other focus. To avoid this possibility the amount of exposed vertical surface at the level of the orchestra and audience has to be minimised. This has been achieved in the following way: the area of exposed wall surface at the back of the stalls is small, since the gallery soffits slope downwards, whilst above the gallery there are huge reflecting surfaces, the wall areas between the reflectors and the gallery seating being treated with absorbent. The principal acoustic design criterion for this hall, however, was the provision of early lateral sound in all seating positions. In the stalls this is achieved in any case for most seats by reflections off the side walls, though balcony fronts were also angled to provide additional lateral sound where possible. In the gallery lateral sound is provided by the reflectors which are angled in each case to provide a lateral reflection in the relevant seating area. In the case of reflectors placed towards the back of the hall, for sound to be sufficiently "lateral" it was found necessary to use dihedral reflectors, a detail of which can be seen in a photograph of a model of this hall in Figure 17.1. The lateral reflections, provided by these reflectors, travel on paths remote from the audience seating, so one does not expect them to suffer attenuation at grazing incidence to the audience. Reflectors were also installed above the orchestra, principally to provide reflections for the orchestra, though also acting as an extension of the main reflectors.

The reverberation time of the hall was calculated for the occupied hall, two sets of absorption coefficient data being used. For both sets, the absorption of the audience according to area was taken from Beranek's results |91|. The first estimation was made by using absorption coefficient data for the individual surfaces taken from two sources: Parkin and Humphreys |96| and Beranek |97|. The second estimation employed Beranek's lumped absorption coefficients |91| for wall surfaces (with the exclusion of concrete for which results in reference |96| were used). The predicted reverberation times (R.T.'s) are given in Table 17.1, together with the measured result for the occupied hall.

TABLE 17.1

Calculated and measured octave reverberation times in seconds in Christchurch Town Hall

Frequency (Hz):	125	250	500	1000	2000
Estimate 1	3.4	2.8	2.2	1.8	1.7
Estimate 2	2.6	2.3	2.0	1.7	1.7
Measured result	2.4	2.2	2.4	2.3	2.1

At least at low frequencies, the estimate based on Beranek's lumped absorption coefficients is closer to measured values. This suggests that the absorption coefficient for wood, of which there are substantial areas in this hall, is nearer Beranek's later value in reference |91| than in his earlier publication |97|.

Reverberation time measurements were made with pistol shots at 12 receiver positions in both the gallery areas and in the stalls, both with the hall empty and occupied. It was thought that for reverberation the large reflectors might divide the volume into two coupled spaces, one above and one below the reflectors. However, in both situations, i.e., with the hall empty and occupied, no significant variation in R.T. was found with position, nor was there any consistent deviation from a linear exponential decay. The mean octave R.T.'s for the 12 positions for the hall empty and occupied are plotted in Figure 17.3. The mean R.T. for all frequencies with the hall full is 2.3 seconds, with no significant rise at bass frequencies.

The reverberant decays with the hall occupied were also measured according to Schroeder's integrated impulse method |21|, the gated integrated

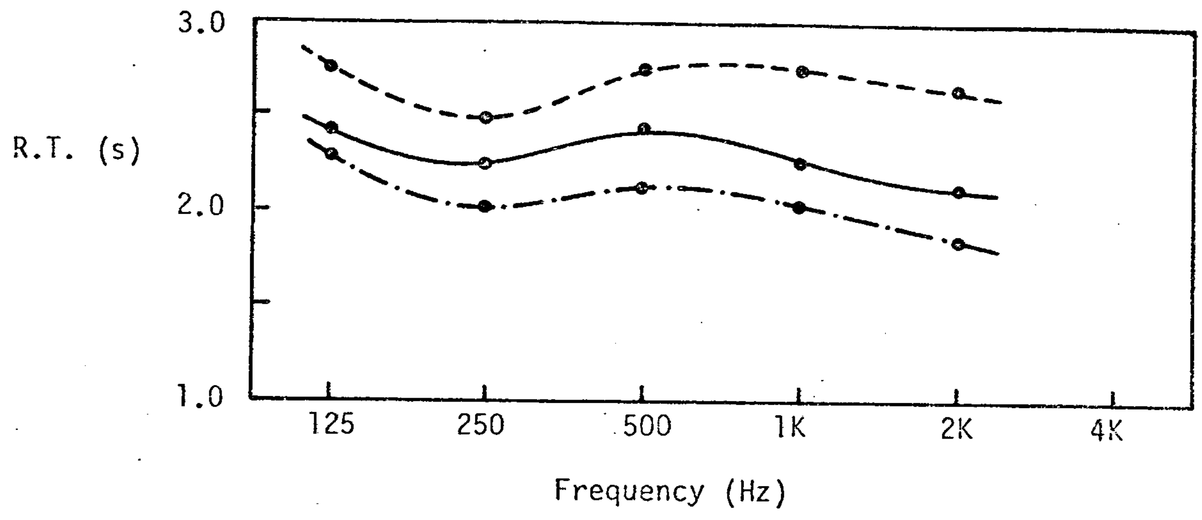


Figure 17.3. Measured reverberation times, averaged over 10-12 measuring positions, in Christchurch Town Hall. – – – –, Hall empty; ————, hall occupied; —·—·—, hall occupied, T_{15dB} by integrated impulse method.

energy meter being used as described at the end of section 16.3. Measurements were made over the first 15 dB of the decay, which, according to Atal et al. |23|, correlates better with subjective reverberance than the traditional R.T. measurement. Double sloped decays were observed in two receiver positions in the front half of the central stalls area. The results plotted in Figure 17.3 are the mean values for the remaining 10 positions. Again no consistent variation occurred with position. The mean R.T. values obtained by this method are consistently about 0.2 secs below the conventionally measured R.T.'s.

17.3 SUBJECTIVE REACTIONS TO THE CHRISTCHURCH TOWN HALL

A mean reverberation time value of 2.3 secs is slightly higher than the generally accepted optimum value of 2 secs for large concert halls. No one, however, commenting on the acoustics of the hall, considered the reverberation excessive. This may lend support to Atal et al.'s theory |23| that only the first 15 dB of the decay are subjectively significant, since the mean value for the R.T. over this range was found to be 2.1 secs. Other explanations are also possible.

The presence of early lateral sound, which this hall had been designed specifically to produce, could be appreciated throughout the hall, though the degree of this spatial effect was not as great as occurs, for instance, in the Vienna Musikvereinsaal. The most striking aspect of this hall was the remarkable degree of clarity. In the empty hall, for instance, a person on the stage could readily communicate with someone in the gallery at the back of the hall. Whilst a concert hall is not designed to have this particular quality, for music the effect is one best described by the word intimacy. This is not the almost clinical clarity occurring in some halls, which enables one to pick out individual instruments (let alone sections of the orchestra), but one gets a feeling of identification with the performer(s) which can be very exciting. The following is a comment (in private correspondence with Dr. Marshall) made by the Professor of Music, Dr. Richie, of the University of Canterbury, Christchurch, referring to a concert by a sixteen member choir:

"The clarity was really very good. One gets it both ways: the total image is clear and also the textural images are clear in all their variation. Further, intelligibility of diction is no problem. The remarkable thing is that, in spite of the large distances involved, there is an intimacy about the placing which is rewarding in this madrigal-style context". (The "placing" here refers to the choir being placed towards the back of the stage.)

There is no doubt that the best seats in this hall are those in the gallery opposite the stage, though throughout the gallery there is a surprising uniformity. The flat central floor area, a feature that was chosen for non-musical reasons, contains seats which suffer from the fact that distant instruments are obscured by near ones for full orchestra performance. Seats at the very back of the stalls right underneath the gallery inevitably suffer from the absence of reverberation from above and behind, though in all other respects the sound is good. A few comments concerning false localisation have also been made, but precise details are not available.

It is unfortunate that this hall is in a rather remote part of the world and will not receive the acclaim and publicity that a hall in Europe would. For those who have heard music in this hall and were able to compare it with experience in other halls, there was a particularly exciting quality about music heard in this hall that deserves emulation.

17.4 OSCILLOGRAMS RECORDED IN CHRISTCHURCH TOWN HALL

Oscillograms at 10 seat positions were taken for a single cycle 2 kHz pulse emitted from the omni-directional source placed centrally on the stage.

Received signals were octave analysed to improve signal to noise ratios and displayed on an oscilloscope. The recorded oscillograms are shown in Figures 17.4, 17.5 and 17.6. Contrary to all other oscillograms in this thesis (with a sweep rate of 10 ms/div), these were recorded at 20 ms/div sweep rate giving a trace of roughly 200 ms duration after the direct sound. A sketch of the hall plan is given in Figure 17.7, to indicate the location of the seat positions used for oscillograms and other measurements described

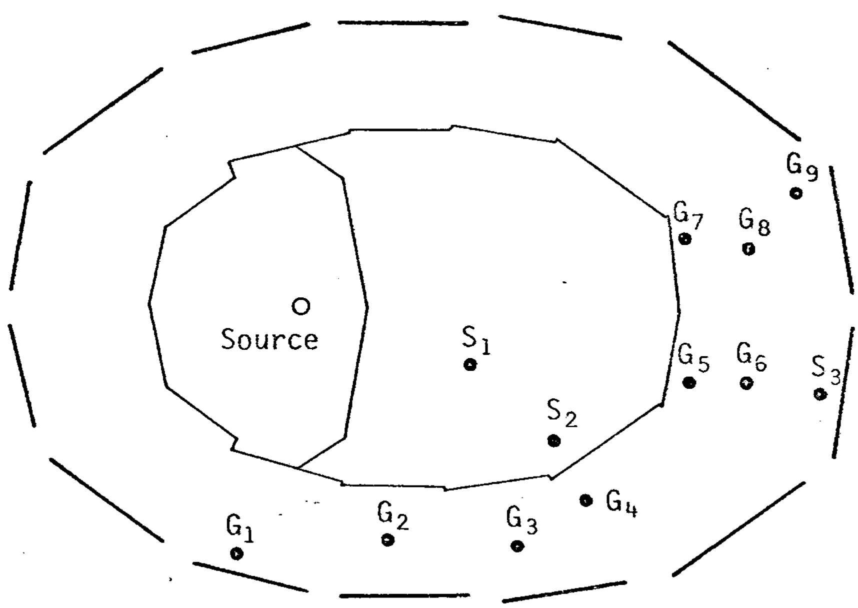
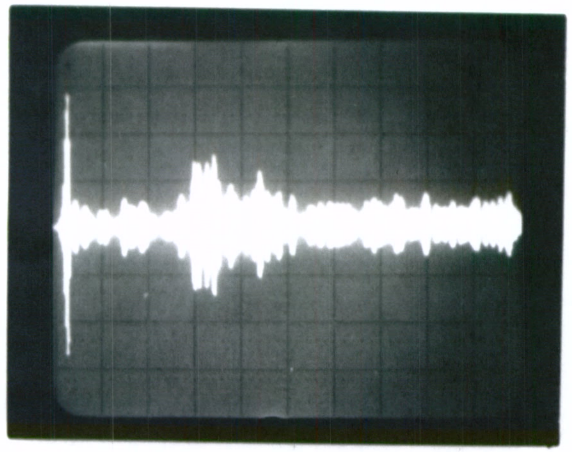


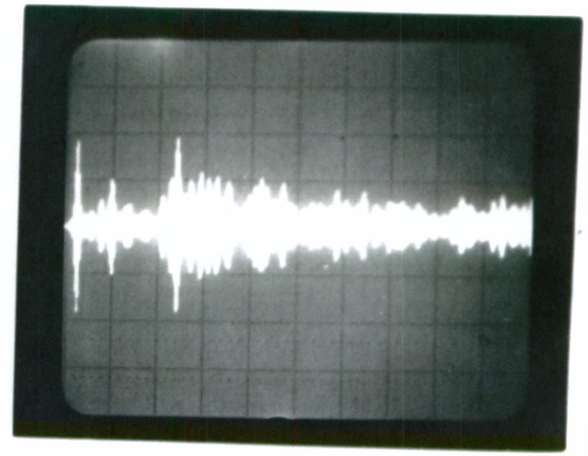
Figure 17.7. Sketch plan of Christchurch Town Hall, indicating measuring positions. S - stalls; G - gallery.

in Chapter 19. Measuring positions are labelled S for the stalls and G for the gallery, the numbers being smallest for positions nearest the source. Actual seat numbers are also given in Figures 17.4-6.

Behaviour in the stalls, illustrated in Figure 17.4, is fairly typical of responses in halls as one moves from the front to the back of the hall. The preponderence of the direct sound becomes gradually less and the reflection density greater as one moves towards the back of the hall. This behaviour however is not found in the gallery in which early reflections occur in all positions (with at least one within 20 ms), even in those closest to the source: e.g., G_1 in Figure 17.5. This can be ascribed to the reflectors primarily positioned for the gallery seating areas and may account for the uniformity of sound in the gallery. Reflection sequences



S_1 - K 25



S_2 - S 31

Inner Stalls

S_3 - K 9

Outer Stalls

Figure 17.4. 2kHz oscillograms in Christchurch Town Hall (stalls). 20ms/div. sweep rate.

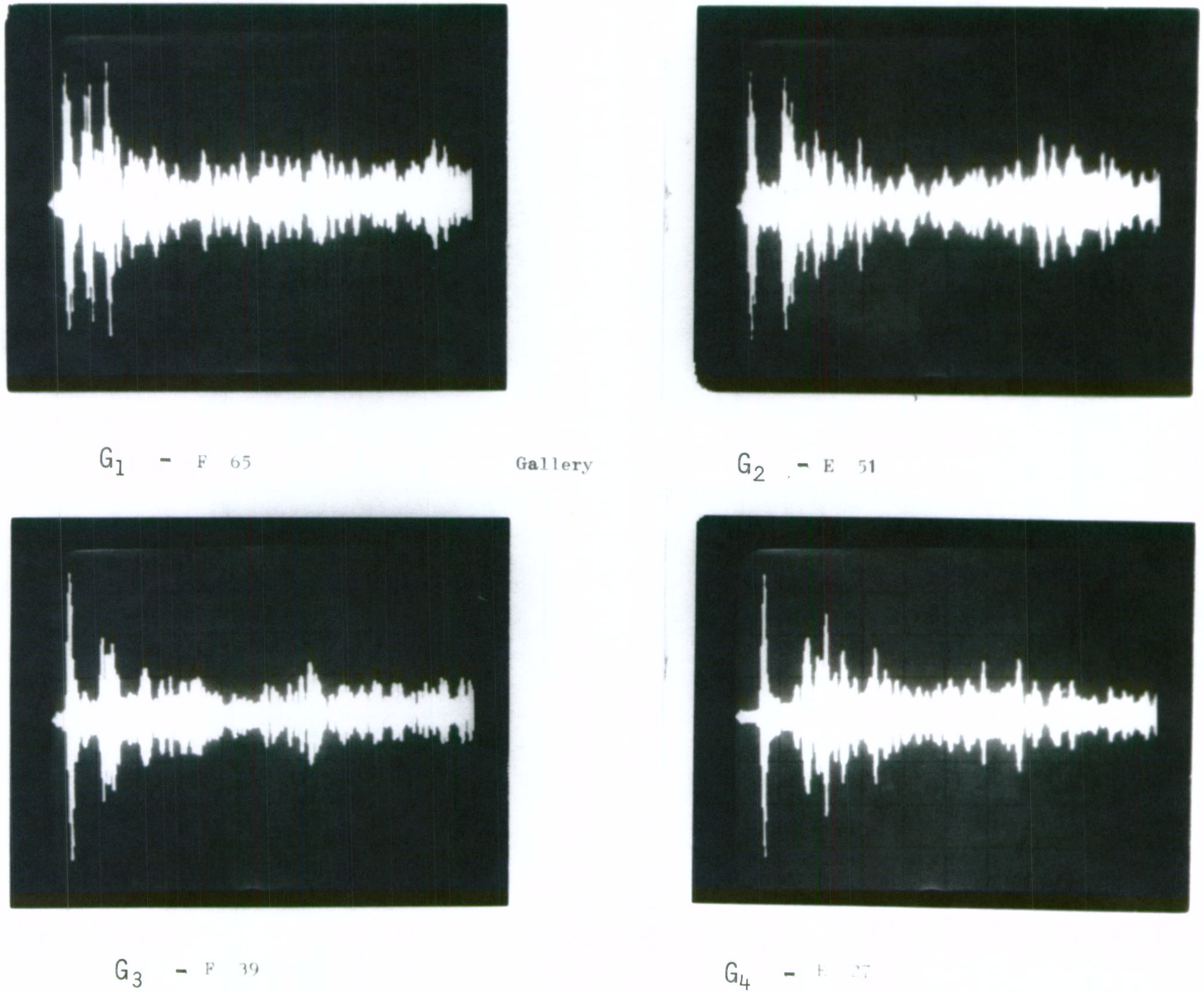
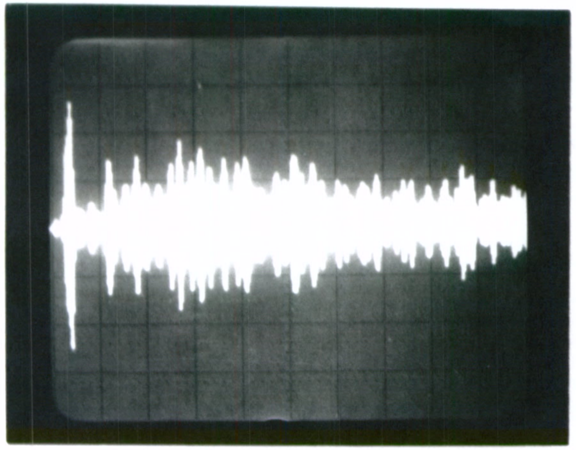
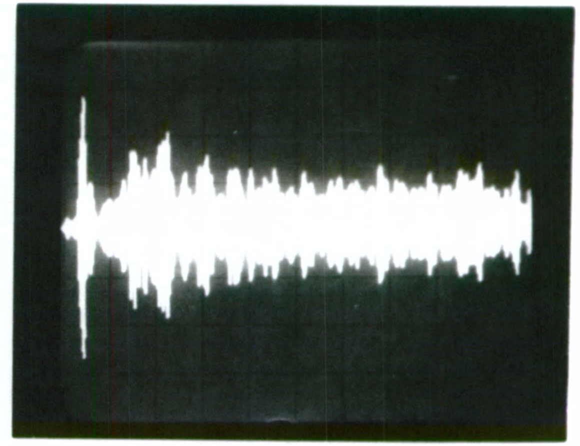


Figure 17.5. 2kHz oscillograms in Christchurch Town Hall (gallery). 20ms/div. sweep rate.

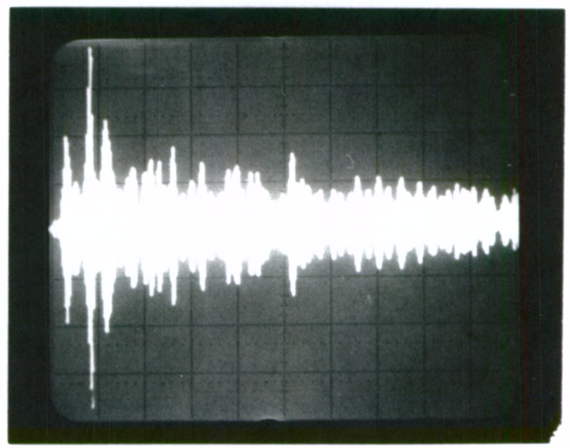


G_7 - E 84

Gallery



G_8 - H 86



G_9 - M 83

Figure 17.6. 2kHz oscillograms in Christchurch Town Hall (gallery -rear). 20ms/div. sweep rate.

in seats at the rear of the gallery, see Figure 17.6, are, as expected, particularly dense. The peculiar situation at seat G_9, where the first reflection is considerably more intense than the direct sound, can probably be ascribed to constructive interference.

A computer programme has been developed by Marshall |98| to calculate the arrival times, etc., of all early reflections up to second order, for a hall consisting of plane surfaces, according to geometrical ray behaviour. Predictions for the source and receiver positions used for oscillogram measurements were compared with measured oscillograms. The degree of agreement, in fact, was not particularly good. The measured reflection density was higher than predicted, which suggests a degree of diffusion at nominally plane surfaces which was not accounted for in the computer prediction. Lack of time unfortunately prevented further investigation of this phenomenon.

17.5 PERTH CONCERT HALL

This hall has the classical rectangular shape enclosing a volume of 16,850m^3. It seats an audience of 1,727 plus a choir of 180. The plan and long section can be seen in Figure 17.8. As well as the stalls seating there are two balconies which "cascade" down each side wall, though each is relatively shallow, containing at most 5 rows of seating. An interesting feature of this hall is the high stage enclosure, rising 4.2m above the stage, with a corrugated surface as shown in the plan. Choir seating is situated behind and above the stage. The ceiling is highly coffered with square elements, and the side walls also are made of vertical panels displaced by small amounts laterally from the mean position (this is only shown on the long section).

The mean reverberation times in the hall, averaged over four positions, are given in Figure 17.9, for the hall occupied and empty. The R.T.'s for the occupied hall are near what is considered optimum.

It is difficult to generalise about the acoustical quality in this hall, and again sufficient time has not elapsed for it to establish a reputation. The orchestra was particularly pleased with the hall, a response which can be ascribed to the enclosure used here, which is probably more elaborate than in the average hall. The value of the enclosure could also readily be appreciated at seats near the orchestra, contributing a sense of breadth to the sound. The best seats are probably at the back of the stalls, where the

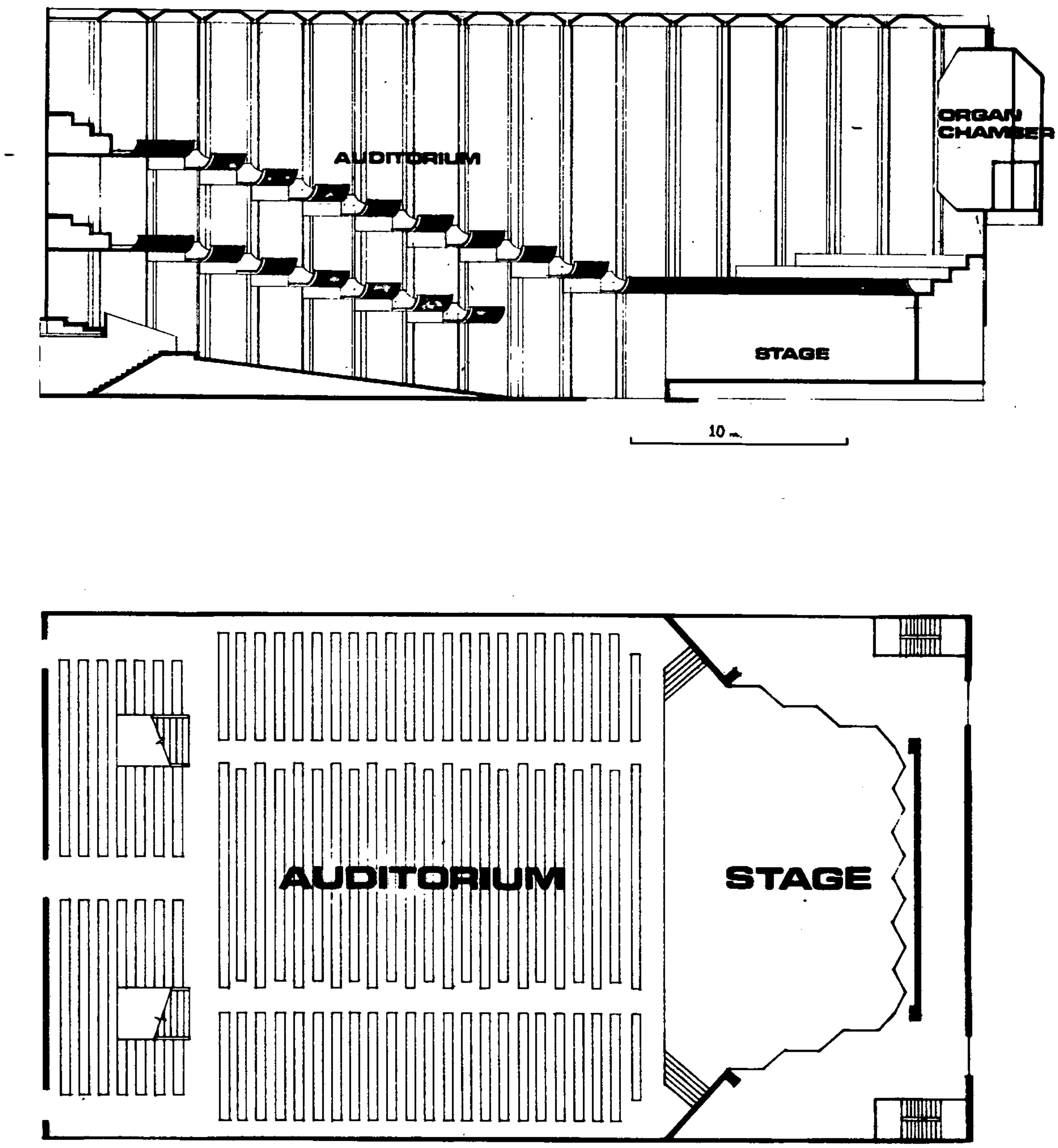


Figure 17.8. Perth Concert Hall, Western Australia.

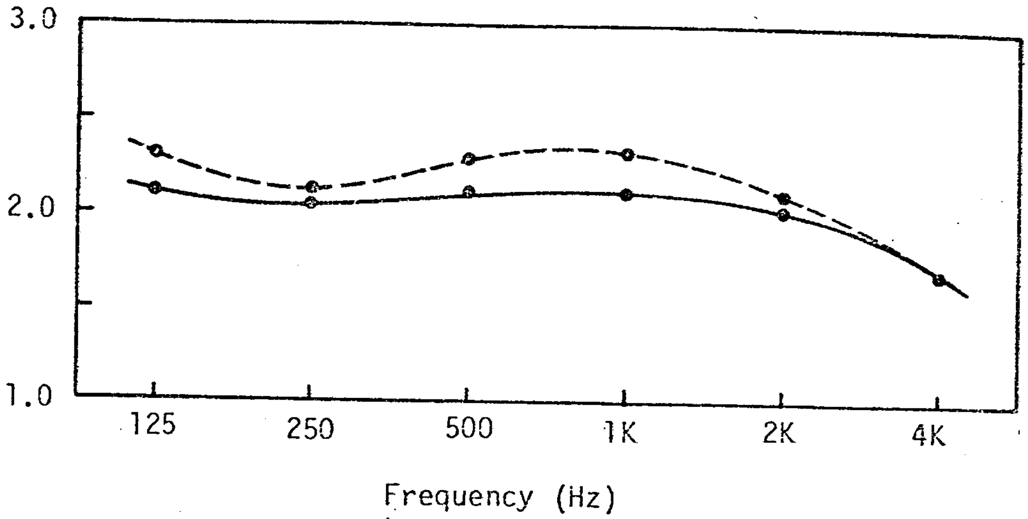
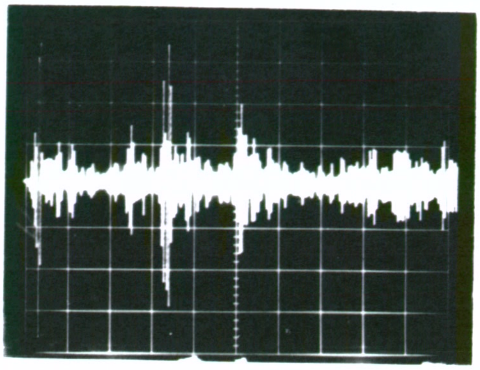


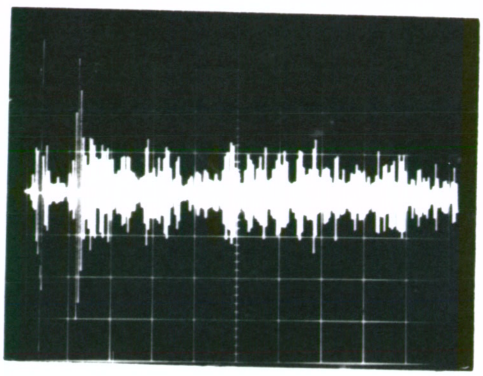
Figure 17.9. Measured reverberation times, in Perth Concert Hall.
– – – –, Hall empty; ————, hall occupied.

acoustics are well balanced and typical of the rectangular shaped hall: the clarity is good, and there is a pleasant sense of involvement in and envelopment by the music. Curiously, the seats at the centre of the first balcony proved rather disappointing, with a rather muddy sound lacking brilliance. This may be due to side reflections being obscured by the balcony tiers, but other factors may well contribute. The upper gallery proved more exciting, with a very pleasant reverberation and sense of brilliance.

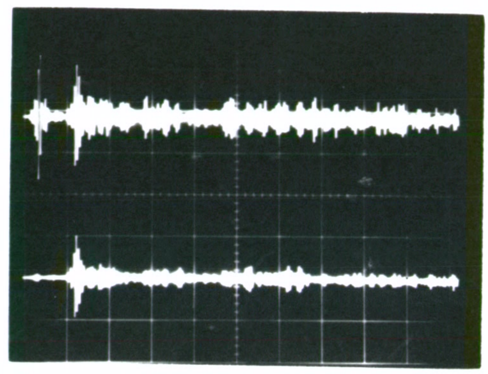
Oscillograms made in the Perth Concert Hall are shown in Figure 17.10. Measuring positions are labelled P, small numbers again are nearest the source. Seat coordinates (referred to an origin at the centre of the front wall) are also included in Figure 17.10. The location of the measuring positions is shown in the sketch plan in Figure 17.11. Unfortunately, on reproduction it is difficult in some of these oscillogram pictures to distinguish precisely the direct sound, since for the direct sound the sweep rate of the trace was greatest. However, for all oscillograms both in Perth Concert Hall and Winthrop Hall close examination reveals that the direct sound amplitude is larger than that of any of the reflections. The oscillograms show the conventional behaviour with the reflection density increasing towards the back of the hall, but the number of discrete reflections which can be distinguished is generally small. The measured oscillograms will be compared with predictions in section 18.2.



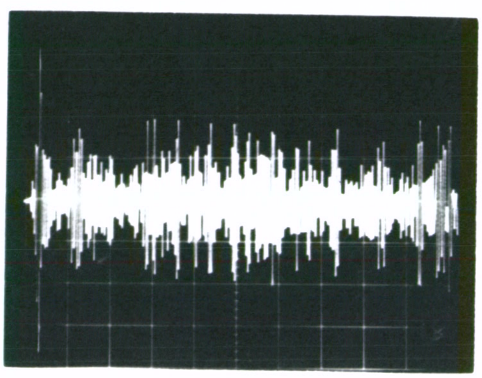
P_1 - (25.1,3.7)



P_2 - (34.0,8.7)



(34.0,8.7) - omni- and directional (~~lateral~~ frontal suppressed)



P_3 - (39.0,8.1) - 1st balcony

Figure 17.10. 2kHz oscillograms in Perth Concert Hall (receiver coordinates in m).
10 ms/div. sweep rate.

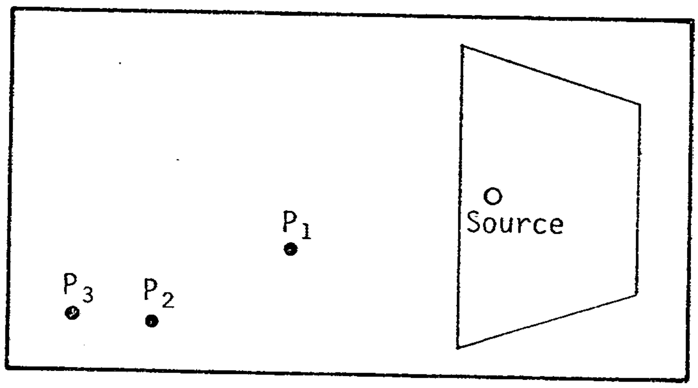
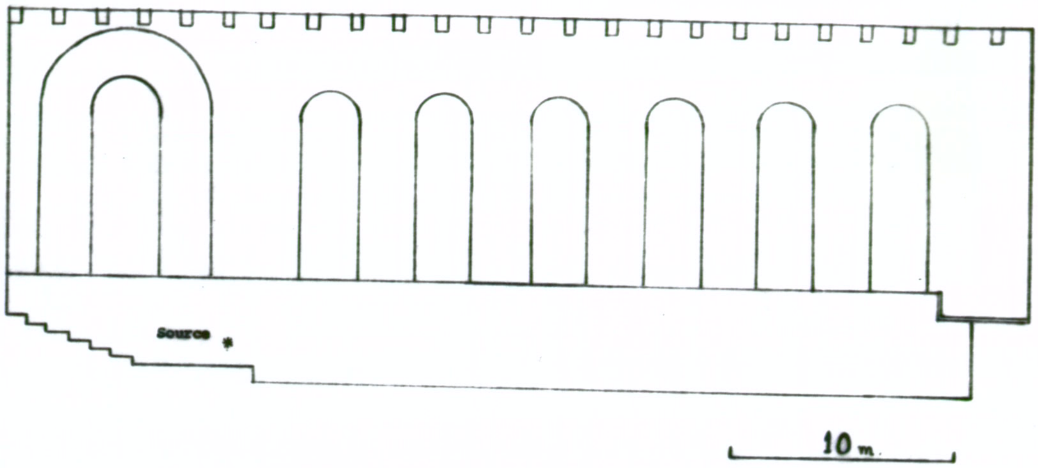


Figure 17.11. Sketch plan of Perth Concert Hall, indicating measuring positions.

17.6 WINTHROP HALL

It is unfortunate that the hall most readily available for measurements also has certain curious and undesirable acoustic features. Winthrop Hall is situated on the campus of the University of Western Australia in Perth, and has an almost purely rectangular shape, which is very convenient for comparison with predictions. The significant deviation from the pure rectangular plan is a series of deep window arches which run the length of each side wall. These can be clearly seen in the plan and long section in Figure 17.12. The volume of the hall is 12,540m^3 and it seats an audience of 980. The seating on the main floor is portable and thus does not exhibit the same degree of absorption as fully upholstered seating. A remarkable feature in this hall is that the side walls below the window arches consist of wooden panels mounted over acoustical absorbent; the rationale behind this curious treatment was never discovered. The ceiling is coffered with cross "beams", made of wooden panels.

The acoustics of Winthrop Hall have generally been criticised, though criticisms have generally been unspecific. In particular the clarity in the rear half of the hall is poor. The hall is particularly bad for solo performances, especially for instruments like the guitar. For orchestral performance this lack of clarity is not so noticeable, but in seats at the rear of the hall one feels remote from the performance. The acoustics of the hall have been described as "patchy" and this has been the author's experience in the balcony, where the sound was much less "muddy" at the



Winthrop Hall : Long section

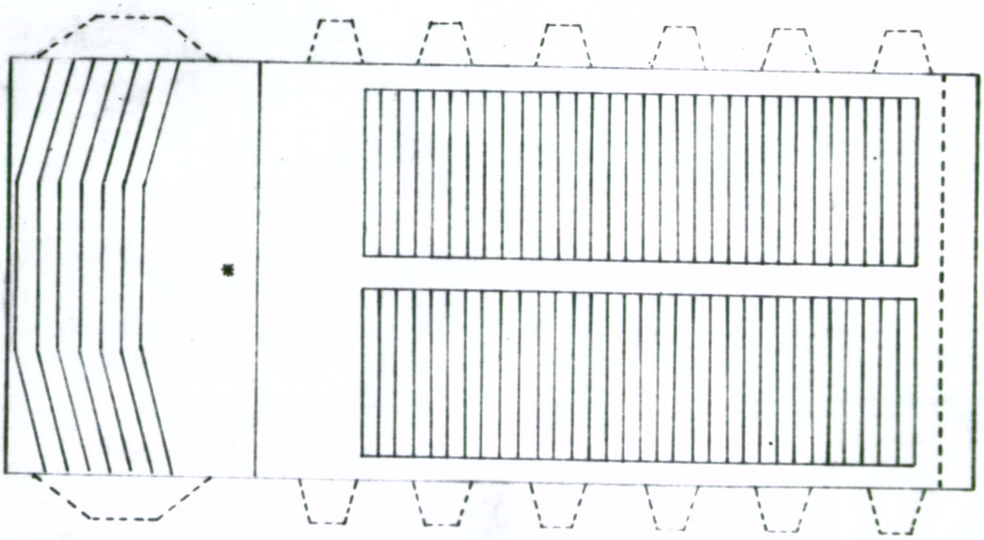


Figure 17.12. Winthrop Hall : Ground Floor Plan

side than at the centre. Whether the window arches, by obscuring reflections for certain areas of seating, produce this "patchy" quality is not known - it proved particularly difficult to isolate reflections arriving with the predicted delay of reflections off this area of the side wall. The local orchestra was generally critical of this hall; this may be attributed to the lack of reflecting surfaces around the stage.

The measured reverberation times for the empty hall are given in Figure 17.13. Unfortunately, no R.T. values are available for the occupied hall,

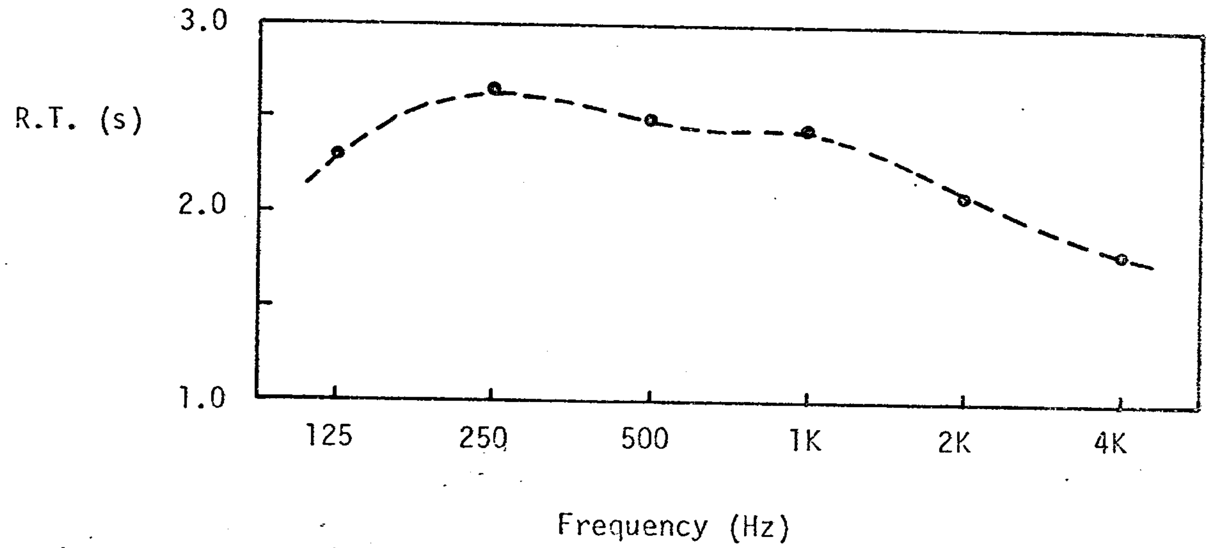


Figure 17.13. Measured reverberation of Winthrop Hall, hall empty.

though one can expect the values to be much less than those for the empty hall since the audience seating is not well upholstered. The R.T. for the occupied hall is thought to be about 1.7 secs.

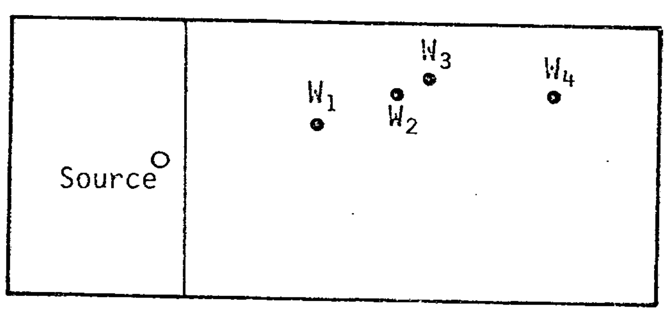
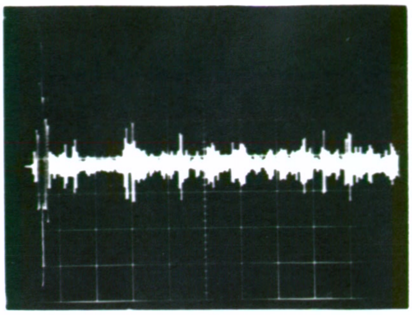
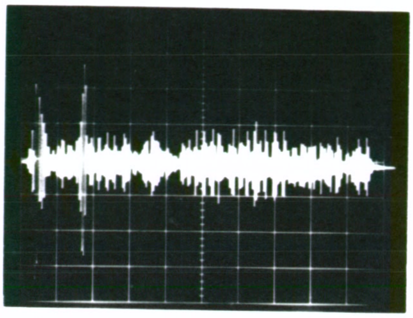


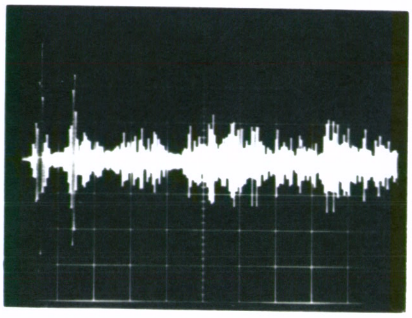
Figure 17.14. Sketch plan of Winthrop Hall, indicating measuring positions.



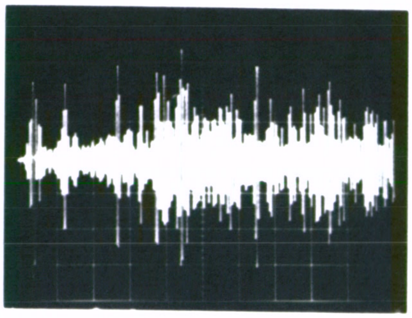
W_1 - (19.6,2.5)



W_2 - (24.6,4.5)



W_3 - (26.6,5.5)



W_4 - (34.6,4.5)

Figure 17.15. 2 kHz oscillograms in Winthrop Hall (receiver coordinates in m). 10 ms/div. sweep rate.

Figure 17.14 shows the measuring positions used for measurements described in later chapters and for the oscillograms in Figure 17.15. Again the number of discrete discernible reflections is small, in spite of this hall approximating closely to one with large, essentially plane, surfaces. A more detailed analysis of the oscillograms is described in the following chapters, which revealed that the intensity of first reflections differed from expected values.

17.7 THE INTERPRETATION OF OSCILLOGRAMS

Oscillograms are a convenient way of presenting details of the early sound in a concert hall, which may, by recording with a directional microphone, be used to illustrate directional behaviour as well. They are, however, only suitable for presenting high frequency information. A further relevant consideration is the extent to which an oscillogram illustrates the subjective response at that receiver position. An oscillogram indicates clearly the physical situation which may be compared with predictions according to the geometrical ray acoustics model. Reichardt claims |12| that oscillograms may be interpreted by the experienced acoustician, though he omits to indicate what conclusions may be derived from them. Recognition of a discrete (disturbing) echo would obviously be relatively easy, as would recognition of a flutter echo. Experience in the Perth Concert Hall, at P_2, and Winthrop Hall, at W_3, in Figures 17.10 and 17.15 respectively, two situations with very similar oscillograms, would require a very refined degree of discernment of the oscillograms to distinguish between what would probably be rated as good and mediocre, or poor, acoustics.

The above two examples have similar oscillograms but different acoustical qualities; examination of oscillograms measured in a small area is also instructive. Figure 17.16 shows oscillograms measured at seat W_2 in Winthrop Hall, and at adjacent positions 1m from the centre position. It is unlikely that the subjective impression would vary significantly within this small area, and yet, apart from the first reflection, the variation between oscillograms is considerable.

The results of experiments described in Chapter 19 would seem to support what has been the tendency in recent years, that is to consider physical quantities based on integrated energy for correlation with subjective effects:

Figure 17.16. 2 kHz. oscillograms of 5 neighbouring receiver positions (receiver coordinates in m.). 10 ms/div. sweep rate.

e.g., rate of rise and decay |23, 34|, ratio of early to late energy |31|, centre time |30|. The exhaustive studies undertaken by Schodder and Junius (summarised in reference |9|), and others, of numbers of reflections within a particular range of energies lead, one presumes, to no apparent correlation between physical and subjective measures; at least no such comparisons have been published. If integrated energy is the relevant quantity in subjective terms, the problem with an oscillogram is to estimate the area below the squared version of the oscillogram, or, in more general terms, to suitably weight the relative contributions of discrete reflections and diffuse ones. It would seem that an oscillogram is a useful addition to other physical data, but for subjective prediction its value is severely limited.

Chapter 18

DIRECT SOUND AND INTEGRATED ENERGY MEASUREMENTS IN HALLS

18.1 DIRECT SOUND TRANSMISSION OVER AUDIENCE SEATING

Measurements similar to those conducted by Watters and Schultz, and others |14, 15| were made in the three halls. The test procedure was as follows: one omnidirectional microphone was placed near the source, while a second was placed about nine rows back, or at least not so far that a reflection arrived within 20 ms of the direct sound. At each 1/3 octave frequency a tone pulse was emitted with a duration of just more than 20 ms, and the signal at the two microphones was sampled for 20 ms after the arrival of the pulse. The energy contained in the two microphone signals was compared with the predicted level difference according to the inverse square law. The excess attenuation over inverse square law is plotted against frequency for the three halls in Figure 18.1.

For the Christchurch and Perth Halls the agreement with results measured by Schultz and Watters is good (for comparison see Figure 11.2 with identical scales). For both, there is a maximum attenuation in the region 160-200 Hz of about 14 dB but the breadth of the attenuation band is narrower than measured by Schultz and Watters, though this may be attributed to the relatively short source-receiver distances used for these measurements. Both of these halls contained well-upholstered seating. For Winthrop Hall, however, whilst the maximum attenuation is characteristic, 12.5 dB, the frequency is 315 Hz, a response at significantly higher frequencies than is general (e.g., at 160 Hz there is an "amplification" of +3 dB rather than the severe attenuation occurring in general at this frequency). This behaviour can be accredited to the seating in this hall being light-weight for portability, though it is particularly interesting to see that the effect is only a frequency shift rather than a modification of the shape of the response. Sessler and West |15| measured identical frequency shifts on model seating with "underpass": there was similar open space below the seats in Winthrop Hall. Their work, however, predicts the traditional response when an audience is present.

Oscilloscope pictures of the near and far (relative to the source) microphone responses in Winthrop Hall at 315 Hz (the frequency of maximum

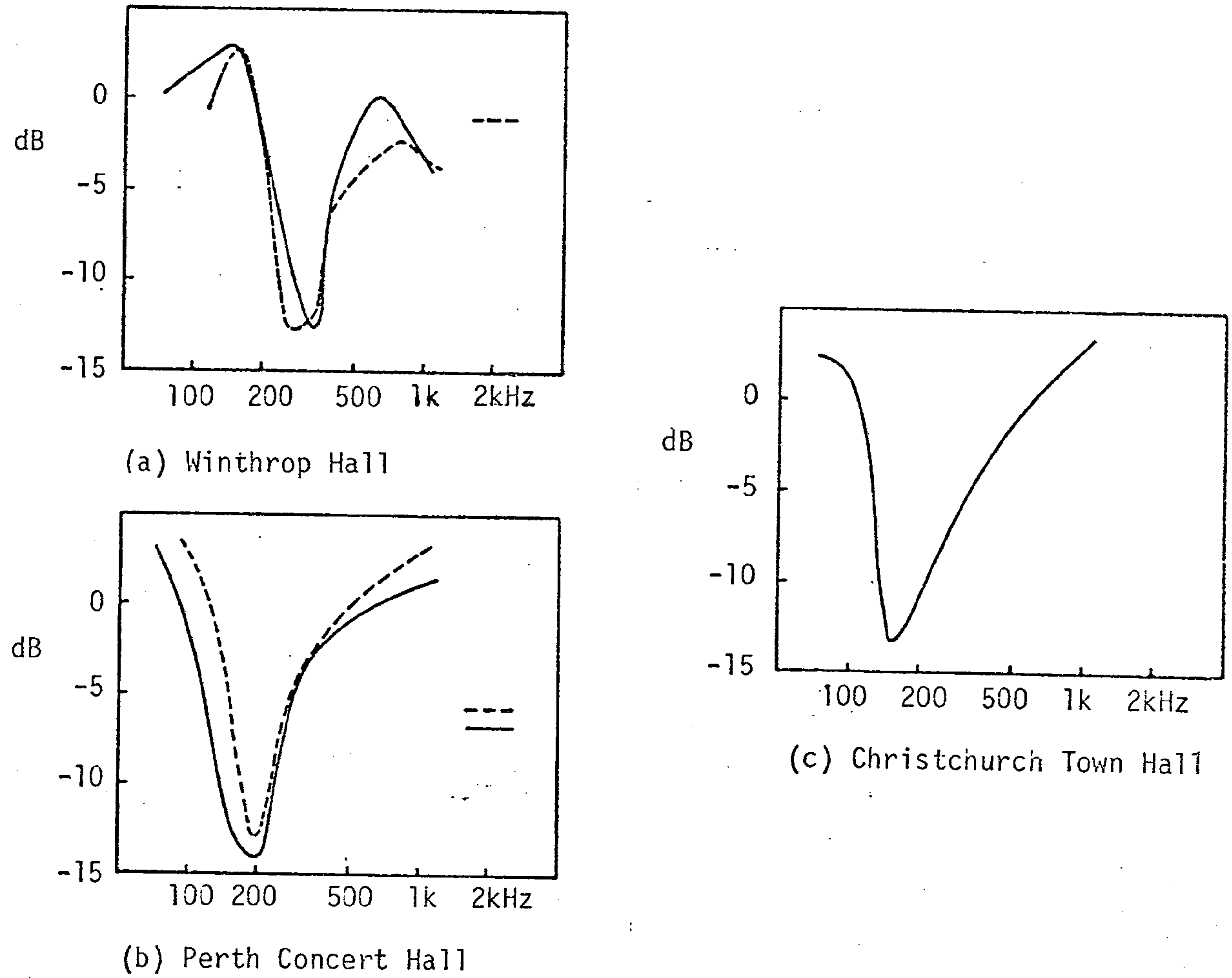
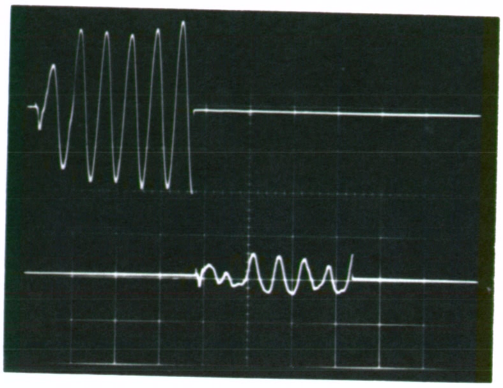


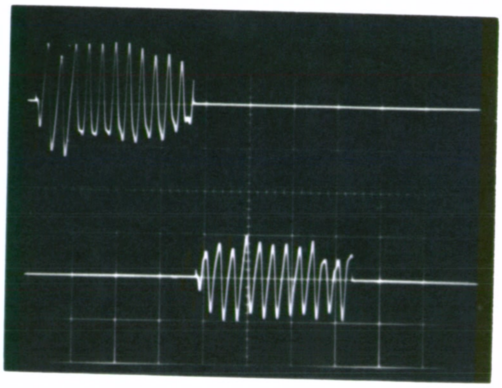
Figure 18.1. Measured excess attenuation over inverse square law as a function of frequency for direct sound passing over audience seating in the three halls. ———, attenuation for 20 ms duration sound pulses; - - - -, attenuation for single cycle sound pulses.

attenuation) and 630 Hz are shown in Figure 18.2. The relevant comparison is between the near and far response at each frequency; from Figure 18.1 excess attenuation (after corrections) at 315 Hz was measured as -12.5 dB compared with +0.8 dB at 630 Hz.

Since most measurements in halls were made with single cycle test signals, the direct sound transmission over audience seating was also measured for single cycle excitation. The sampling time for the microphone signals was varied from 20 ms to 2.5 ms, depending on frequency, to sample the total pulse energy. Measured results are included in Figure 18.1 for Winthrop Hall and Perth Concert Hall, again plotted relative to inverse square law behaviour.



315 Hz



630 Hz

Attenuation of sound at grazing incidence over audience seating at 315 Hz and 630 Hz in Winthrop Hall.
Upper trace - microphone close to source; lower trace - microphone eight rows back.
Relative sensitivity of the two microphones held constant.

Figure 18.2.

Measurements at the single frequency 2 kHz are also included. Behaviour can be seen, for both halls, to be very similar to the 20 ms pulse measurements, a surprising result perhaps in that the "interference effect" is set up with no latency. The fact that audience attenuation filtering occurs for single cycle excitation validates single cycle measurements of the effects of audience attenuation filtering on the early sound, etc., reported in the next chapter.

18.2 INTEGRATED ENERGY MEASUREMENTS IN HALLS

It is obviously advantageous to be able to make predictions of the behaviour of a particular physical quantity from a computer model. The value of oscillograms, as such for assessing concert hall acoustics has been questioned in the previous chapter. An investigation of integrated acoustic energy seemed to be an approach more similar to the likely auditory processes involved. This section contains an investigation of the integrated energy arriving within 100 ms of the direct sound, and relates results to computer predictions. From this investigation it was hoped to be able to determine the validity of the predictions and indicate situations where predicted results require qualification.

The measuring process involved using single cycle 2 kHz pulses and integrating the energy at the receiver position over intervals of 5, 10, 20, 50, 80 and 100 ms after the onset of the direct sound. Results were related to the direct sound energy. Since the nominal duration of a single cycle 2 kHz pulse is only 0.5 ms, interference effects were minimal. As oscillograms of the same receiver positions were also made, also with a single cycle 2 kHz pulse, details of the integrated responses could readily be related to individual reflections, where relevant. The measurements were made at the four receiver positions in Winthrop Hall and the three receiver positions in the Perth Concert Hall. Both omni-directional and lateral sound measurements were made.

Computer prediction was based on the assumption that these two halls were basically flat-walled rectangular halls with absorbent floor surfaces. The walls were assumed to be acoustically "hard", providing a series of specular reflections, and the integrated energy was calculated on the assumption of incoherent energy addition. For Winthrop Hall a modification of the basic rectangular hall computer programme was introduced to include the effect of the window arches, which for the purpose of computation were assumed to be absorbent. For the Perth Concert Hall additional reflections of the orchestra enclosure were calculated graphically and included in the predictions. Further details of the computer prediction procedure will be given in Chapter 20.

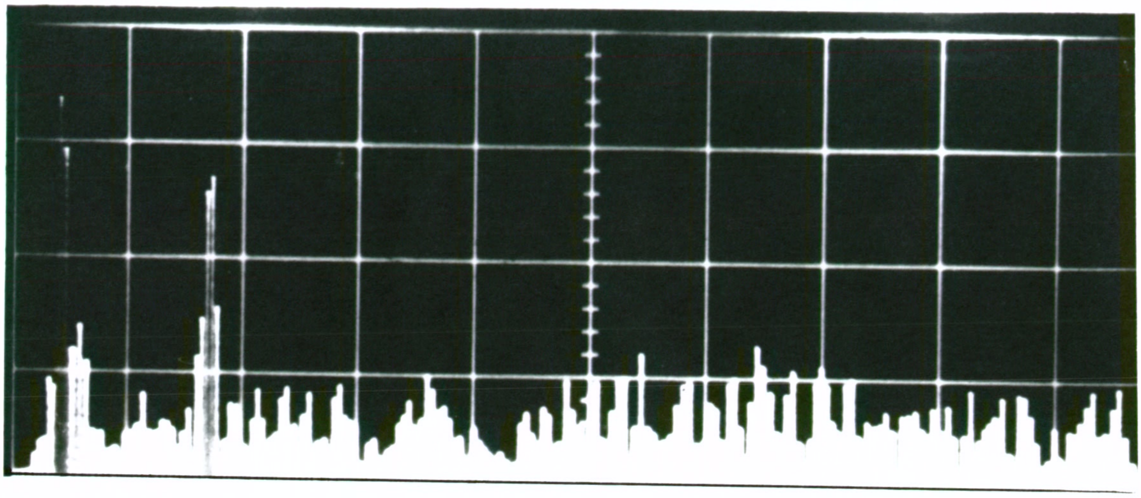
Figures 18.3 and 18.4 show a comparison between the measured and computed oscillogram for the seat at W_2 in Winthrop Hall. The time scales are identical, though the intensity values are uniformly larger in the computed oscillogram. It is immediately evident from this comparison that correspondence between measured and predicted results only extends to the first reflection in this case (for the first two or three in other cases) and that reflection density in reality is much greater than predicted after 20-40 ms, but the levels are generally lower than predicted. In other words, only for the first few reflections does the specular assumption hold and thereafter reflections are generally diffuse. The likelihood that integrated energy predictions on the basis of a specular model correspond with measured results appears remote.

Figure 18.5 contains the measured and computer predicted values for the four Winthrop Hall positions, both for the total sound and the lateral sound. The measured values are indicated by the solid lines, a straight line being drawn between each measured value; computed values are indicated by circles or crosses. The characteristic behaviour, recognizable in oscillograms, that the amount of early energy relative to the direct sound is much greater towards the back of the hall, is also evident in the integrated energy curves. Agreement between measured and predicted results, however, is within about one decibel, with the following notable exceptions.

(a) For seat W_1 the measured lateral energy is consistently less than the predicted value, whilst for seats W_3 and W_4 the 20 ms and 10 ms values, respectively, are less than predicted.

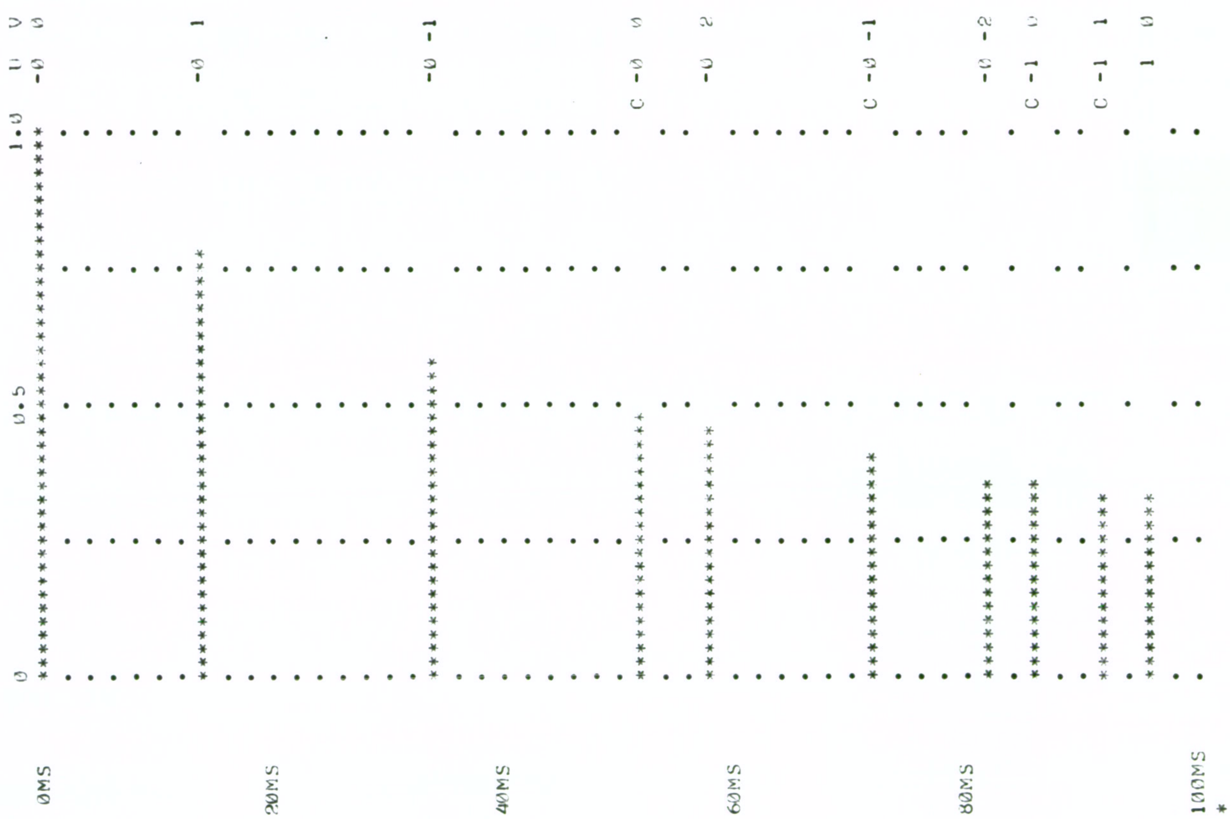
(b) Integrated energy values at receiver position W_4 are significantly larger than predicted for periods longer than 50 ms, though the relevant discrepancy of total and lateral energy from the predicted values is about the same for each.

Examination of the echograms for these receiver positions in Figure 17.15 provides an explanation of the relatively low values of the lateral sound measurements. For seat W_1 there is a double lateral reflection at 25 ms, rather than the predicted single lateral reflection; the incoherent sum of their energies is however -2.6 dB relative to the predicted level.



Measured oscillogram in Winthrop Hall (10ms/div. sweep rate).

Figure 18.3.



Computed oscillogram for the same receiver position in Winthrop Hall.

Figure 18.4.

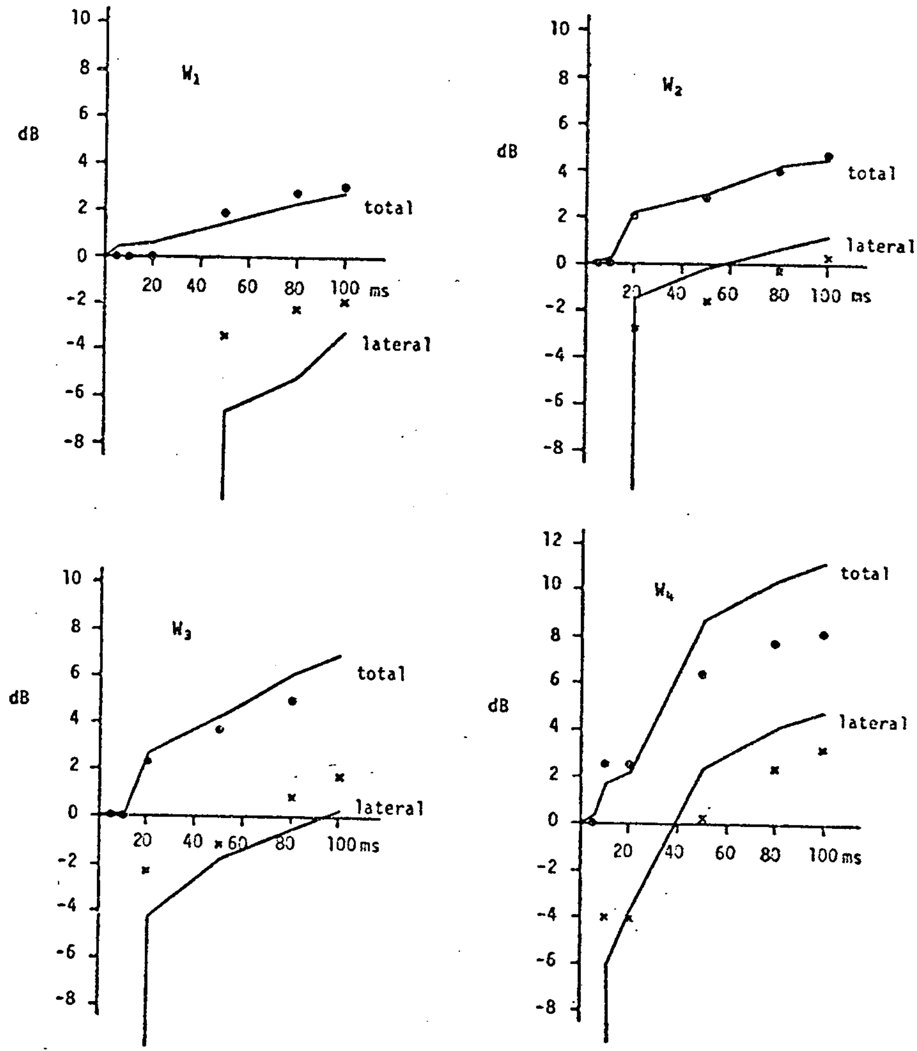


Figure 18.5. Comparison of measured and predicted integrated energy in four Winthrop Hall receiver positions. Ordinate is integrated energy relative to direct sound energy; abscissa is time period after the direct sound. ———, measured integrated energy with 2 kHz single cycle pulse for both total and lateral sound; •, computer predicted energy level of total sound; ×, computer predicted energy level of lateral sound.

The second reflection at 39 ms, another lateral reflection, is also 3.8 dB below the predicted level. These discrepancies explain the relative low values of the lateral sound energy measurements. The unpredicted low level early reflections within the first 10 ms provide compensation such that the total level measurements agree well with predictions.

For both seats W_3 and W_4 the level of the first reflection, in each case a lateral one, is found on the oscillograms to be below the predicted level by 1 dB and 2.5 dB, respectively, which explains the discrepancy in the measured lateral sound in short time periods after the direct sound. It would seem that the wall section producing lateral reflections exhibits some attenuation at 2 kHz. This section consists of wooden panels backed with absorbent!

Comparison of the oscillogram for the seat at W_4 with the predicted reflection sequence shows that for later delays the actual reflection density is very much higher than predicted. This is evidence that the number of diffuse reflections contributing to the integrated energy becomes significant towards the back of the hall. The proportion of lateral sound, however, remains close to that of specular prediction. A similar behaviour was also found in the Perth Concert Hall.

The Perth Concert Hall differs more radically from the simplified computer model than Winthrop Hall, so one would expect larger discrepancies between measured and predicted results: the ceiling was highly coffered, the side walls were lightly diffusing, and no account was taken of the balconies in the predictions. Examination of the oscillograms taken at the three seat positions in Figure 17.10 again reveals only a few discernible discrete reflections. The measured and predicted integrated energy values are shown in Figure 18.6. The results for the various seat positions will be discussed in turn.

At the seat P_1 both the total and lateral integrated energies significantly exceed predicted values, whilst examination of the oscillogram indicates a first reflection 4.1 dB more intense than predicted. This suggests that the direct sound was attenuated at the measuring frequency. The measurement of the direct sound transmission for single pulses at 2 kHz is, from Figure 18.1(b), -5.8 dB relative to the inverse square law value. Since the source and receiver positions used in the direct sound transmission measurement were close to the positions for this particular integrated energy experiment, one can conclude a similar behaviour in both situations. This 2 kHz attenuation is probably a coincidental result peculiar to the direct sound for this particular source and receiver position and for the particular seating in this hall. It does not appear to occur for the other seat positions investigated and, being distinct from the low frequency attenuation generally encountered, it cannot be presumed to exist for different directions of

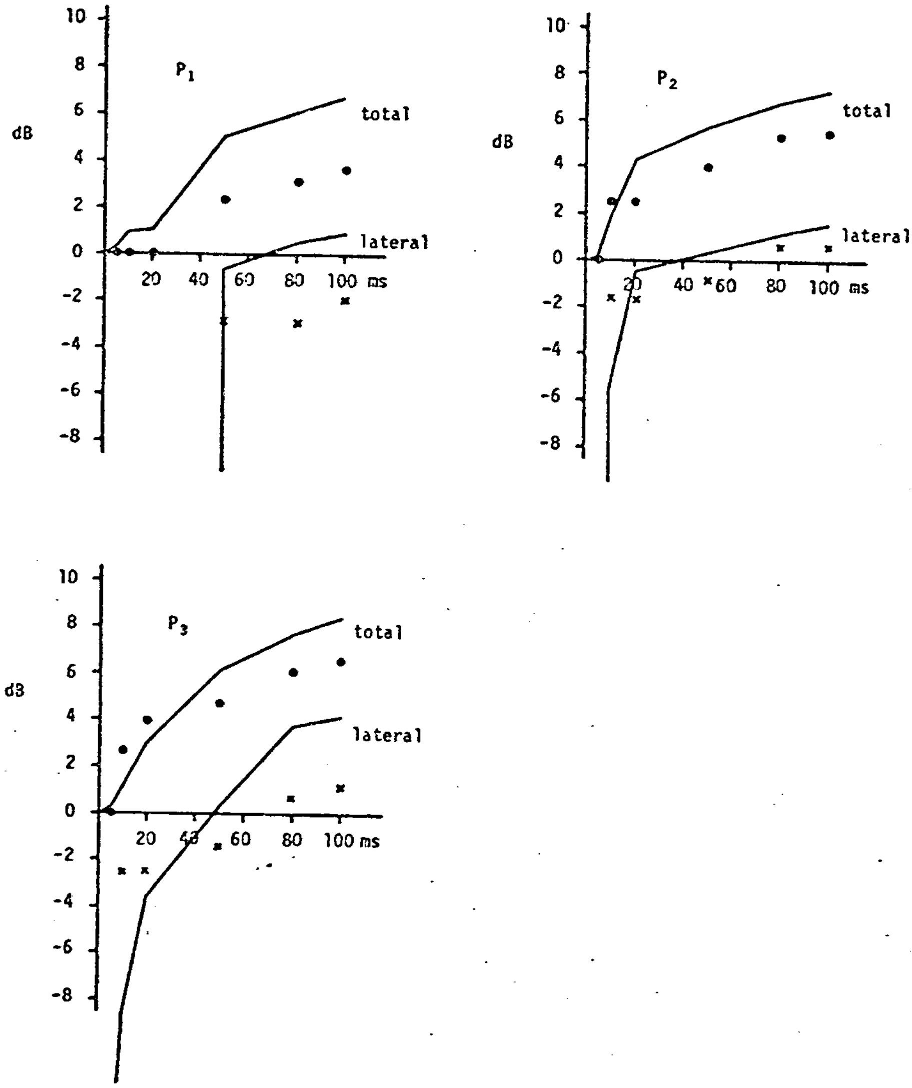


Figure 18.6. Comparison of measured and predicted integrated energy in three Perth Concert Hall receiver positions. Ordinate is integrated energy relative to direct sound energy; abscissa is time period after the direct sound. ———, measured integrated energy with 2 kHz single cycle pulse for both total and lateral sound; •, computer predicted energy level of total sound; × computer predicted energy level of lateral sound.

incidence relative to the seating, etc. Given that the direct sound energy is lower than one expects, the behaviour of the measured energies, both total and lateral, corresponds well with predictions, with an almost constant 3 dB difference.

For the seat P_2, the measured intensity from the oscillogram of the first reflection agrees, within 0.1 dB, with the predicted value. This suggests that there is either no excess attenuation of direct sound at 2 kHz or that the excess attenuation applies both to the direct sound and the first reflection. If the latter were true, the ceiling reflection would emerge as particularly prominent in the oscillogram, which it does not (the predicted delay of the ceiling reflection is 39 ms). One can therefore consider that no excess attenuation of the direct sound occurs for this seat. This being the case, the behaviour at this seat corresponds to behaviour at the seat in Winthrop hall at a similar location. The incident energy (from diffuse reflections) exceeds the predicted specular reflected energy, though the proportion of lateral sound remains close to that of prediction.

The same conclusion may be made for the seat at P_3 in the first balcony, though for this particular seat the intensity of the first (lateral) reflection, according to the oscillogram in Figure 17.10, is 6 dB below that predicted, perhaps due to obstruction by the balcony fronts "cascading" down the side walls. The discrepancy is reflected in the low lateral energy figure for the 10 ms interval.

18.3 THE VALUE OF INTEGRATED ENERGY PREDICTIONS

Analysis of the oscillograms and integrated energy measurements in these two roughly rectangular halls has shown that although a predicted oscillogram differs wildly from a measured one, due to reflections with a delay greater than 50 ms (or even less) being generally diffuse, rather than specular, a good degree of correspondence exists between integrated energy measurements and predictions. Unfortunately the study exposed a peculiarity in each of these halls: that lateral reflections were being attenuated on reflection for some seats in Winthrop Hall, and that, for a seat in the middle of the stalls in Perth Concert Hall, the direct sound at 2 kHz was suffering attenuation. Since circumstances prevented further study of these idiosyncracies, some results have had to be qualified and unfortunately precise corrections were not generally possible. Nevertheless, with these qualifications, agreement between measured and predicted values was good for seat positions in the front half of the hall. Towards the back of the halls, diffuse reflection energy exceeded the predicted specular values, but there was generally good agreement in the ratio of lateral to total energy.

Since this latter quantity is the one of interest in this thesis, these measurements validate consideration of computer predictions of this ratio for different shaped halls (see Chapters 21 and 22). The following chapter describes measurements of this quantity at frequencies at which interference effects between reflections invariably occur.

Chapter 19

EARLY ENERGY AND LATERAL ENERGY MEASUREMENTS IN HALLS

19.1 INTRODUCTION

Measurements of two physical quantities were made in the three halls over the frequency range 80-1000 Hz: the early energy fraction and the lateral early energy fraction. These measurements were aimed at answering the following questions.

(a) To what extent does audience attenuation at grazing incidence affect the energy content of the early sound?

(b) Do measurements of the early energy fraction correlate with subjective impressions of clarity?

(c) What is the measured value and variation with frequency of the ratio of lateral to non-lateral early sound which was found in Part I to correlate with subjective spatial impression?

19.2 THE PROBLEM OF MULTIPATH INTERFERENCE

When simulating a sound field, or calculating the contributions of a series of reflections, it is natural to consider incoherent addition of reflection energies. But what evidence exists to suggest that the ears react in a manner equivalent to incoherent energy addition? As regards spatial impression, comparison experiments with reflections from different directions, as reported in Chapter 10, indicated a response according to incoherent energy addition. It remains to be established whether this is a typical behaviour or a special case. It can readily be appreciated in a steady state reverberant field excited by a pure tone that the ear responds to constructive and destructive interference responsible for the modal pattern. It is obvious in this situation that the ear has no means of distinguishing the individual components responsible for maxima and minima. In situations of transient excitation some evidence is available that the auditory system can interpret the sound field in terms of its individual components, but such an ability can evidently not extend to long sustained musical chords. Whether or not the quasi-steady state situation is interpreted with reference to a short-term memory derived from the transient situation is purely a matter of speculation.

The work of Lochner and Burger |27| was one of the first in which incoherent energy addition of reflection energies was considered. They discovered that for speech the delay of a reflection has no significance if it is within a critical time period, interpreted as the integration time of the ear.

That the intelligibility is independent of reflection delay is, in fact, not surprising, since speech intelligibility is insensitive to narrow band filtering and the energy content of a broad band sound is unlikely to be affected significantly by interference effects. With broad band sound, it is, however, generally as tone colouration that the ear detects interference between delayed signals. For narrower band musical sounds the ear appears to be able to compensate for changes in sound intensity and interpret the loudness on the basis of harmonics as well as the fundamental. Two specific records of this ability are reported in the literature. Saunders |99| recounts how a group of professional violinists listened to a slow scale of semitones, played by one of their number, covering a frequency range in which there was a measured change of sound level of 10 dB due to interference effects, whilst both the player and the listeners said that the scale was uniformly loud. Schultz and Watters |68| conducted an experiment in which the fundamental of a viola theme was almost completely suppressed due to interfering paths, but this was not perceived as diminished loudness or clarity by subjects making a comparison with the "unfiltered" situation. This ability to interpret loudness in a filtered situation due to the presence of harmonics appears to be limited to solo or lightly scored passages. Schultz and Watters found that, when the viola harmonics were masked by noise simulating fundamentals and harmonics of higher instruments in a fully scored passage, the viola line was not audible when it was effectively suppressed due to interfering paths.

The experiments of Schultz and Watters were with interfering waves from similar directions. Similarly, most measurements of tone colouration with a signal and its repetition have involved monophonic reproduction (see section 4.2(c)). In a simulation, however, it is readily appreciable that the tone colouration that occurs with a frontal (e.g., ceiling) reflection is considerably more intense than that which occurs with a lateral reflection of, say, 40° angle of azimuth (see also reference |11|, Section 2). This is again an effect of great audiological interest which has yet to be investigated. Flanagan and Lummis |52| attempted to construct an instrument to permit recordings in situations in which multiple-path interference causes tone colouration in a monophonic recording. By recording with two microphones and dividing the frequency range into a number of frequency bands, the final signal was constructed by taking the signal for each frequency band from the channel in which the response was greatest in that frequency band. The auditory system evidently has an analagous processing system, since with speech, for which tone colouration is even more readily appreciable than with music, interference by multiple paths which occurs regularly in rooms is rarely perceived.

It appears from the slender evidence available that the ear analyses a sound field in terms of its individual components whenever clues are available to permit their differentiation. This facility is greater whenever the sound components arrive from different lateral directions. Situations similar to that reported by Schultz and Watters, in which multiple-path interference results in a particular instrument being inaudible over a narrow frequency range, are likely to be limited to particular source-receiver locations. For information relevant to seating areas and source regions, the average behaviour is better determined if such local variations due to multiple-path interference can be eliminated.

To obtain measurements as far as possible independent of multiple-path interference involved the use of a test signal of short duration (see section 16.2) and presentation of results as averages of measurements in three adjacent third octaves. Since the comb filter response due to two interfering waves has, in general, the characteristic that both a maximum and a minimum are contained in an octave (see Figure 4.2), averaging over an octave minimises the influence of interference effects on measured values. A similar presentation of results was also used by Schroeder *et al.* |39|.

19.3 SPATIAL AVERAGING

Rather than by averaging over an octave interval, results corresponding to incoherent addition of reflection energies may also be obtained by taking a spatial average of results in neighbouring seating positions. The long measurement time this procedure involves would be justified if the values of the quantities being measured were required at individual frequencies. However, for the quantities discussed in this chapter, this is not the case: for example, the differing sensations of spatial impression at different frequencies are associated with broad frequency bands rather than discrete frequencies (see section 11.1).

By comparing results, averaged over octave intervals, measured in neighbouring seating positions, an assessment can be made of the degree to which multiple path interference is influencing results. Such measurements also give information relevant to the suitability of a particular physical measure; since, except in special (and undesirable) cases, subjective acoustic quality does not vary from seat to seat, but a particular acoustic experience is associated with a group of seats. It follows that if a physical acoustic measure correlates with subjective experience it should not vary significantly from seat to seat but should vary between seating areas with known acoustical differences. (Indeed, such a criterion has been used by Preizer |100| as the sole determinant of parameter values for a subjectively relevant physical quantity in music halls.) It was decided to measure quantities at four points roughly 1m spaced from a central position. Unfortunately, limited time in the halls tested prevented an elaborate series of such measurements, though to make them poses the question of what is a typical seat in an area of uniform acoustics.

A measurement of the 80 ms energy fraction was made at four points at the corners of a square with diagonal of 2m in Christchurch Town Hall, in the

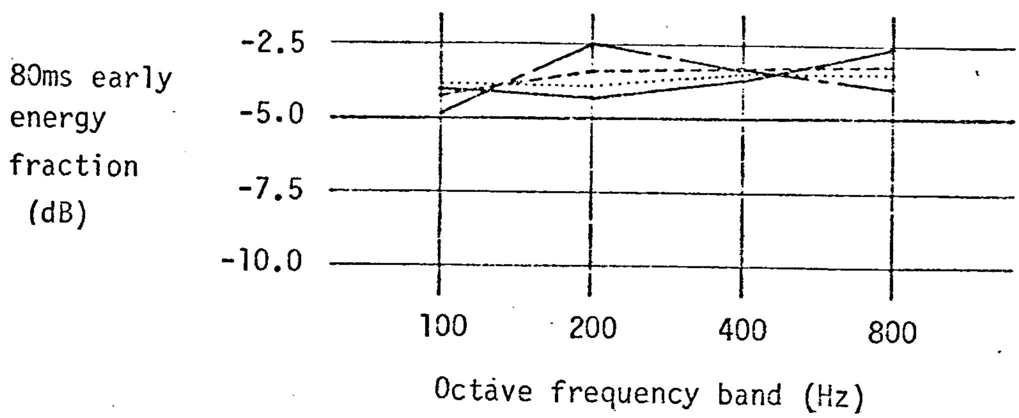


Figure 19.1. The 80 ms energy fraction measured at four neighbouring receiver positions in Christchurch Town Hall (near seat S_2).

frequency range 80-1000 Hz. Mean octave values are plotted in Figure 19.1. Deviation between the various results is small; to say that measured values of the 80 ms energy fraction are valid within 1 dB seems a reasonable conclusion. Regrettably, no similar measurement was made over a wide frequency range for the 80 ms lateral fraction, but variation of this measured quantity within a small area appeared to be slightly larger.

19.4 THE EARLY ENERGY FRACTION

The history of the measurement of the 50 ms energy fraction was considered in section 2.4; it has been generally suggested that the clarity in a hall is related to this quantity, though no experimental evidence is reported in the literature to substantiate this claim. Evidently the ear does not have a sharp cut-off period of 50 ms. It is also possible that the relevant time period is frequency dependent, being longer at low frequencies; at least low frequency echoes require a longer delay to be as disturbing as full frequency echoes |24|. Such frequency dependent integration behaviour would be in line with communication theory.

Both the 50 ms energy fraction and the 80 ms energy fraction were measured in the three halls. In the light of the discussion above both may be significant as regards subjective clarity, though the latter may be relevant only to low frequencies. The results can be compared with the value corresponding to a linear exponential decay. Such a comparison indicates whether the proportion of early energy, and therefore the clarity, is greater or less than average, a consideration particularly relevant in the Christchurch Town Hall where the early reflections were "tailored" for a particular subjective effect.

Equation (15.8) expresses the integrated energy corresponding to an exponential decay. The integrated energy in the period between t_1 and t_2 is given by

$$I_{t_1}^{t_2} = I_o^{\infty} (e^{-kt_1} - e^{-kt_2}),$$

where $k = 13.82/T$. To obtain the τ ms energy fraction, the initial integration starts at time T_o, the arrival time of the direct pulse after its emission at $t = 0$. Then

$$\tau \text{ ms energy fraction} = I_{T_o}^{T_0+\tau} / I_{T_0}^{\infty}$$

$$= I_0^{\infty}(e^{-kT_0} - e^{-kT_0-k\tau})/I_0^{\infty}e^{-kT_0}$$

$$= (1 - e^{-k\tau}) = 1 - e^{-13.82\tau/T}. \qquad (19.1)$$

Thus, if a linear exponential decay is assumed, the 50 and 80 ms energy fractions are independent of source and receiver position, and are only a function of the reverberation time. A plot of the theoretical 50 ms and 80 ms energy fractions according to equation (19.1) is given in Figure 19.2 as a function of reverberation time.

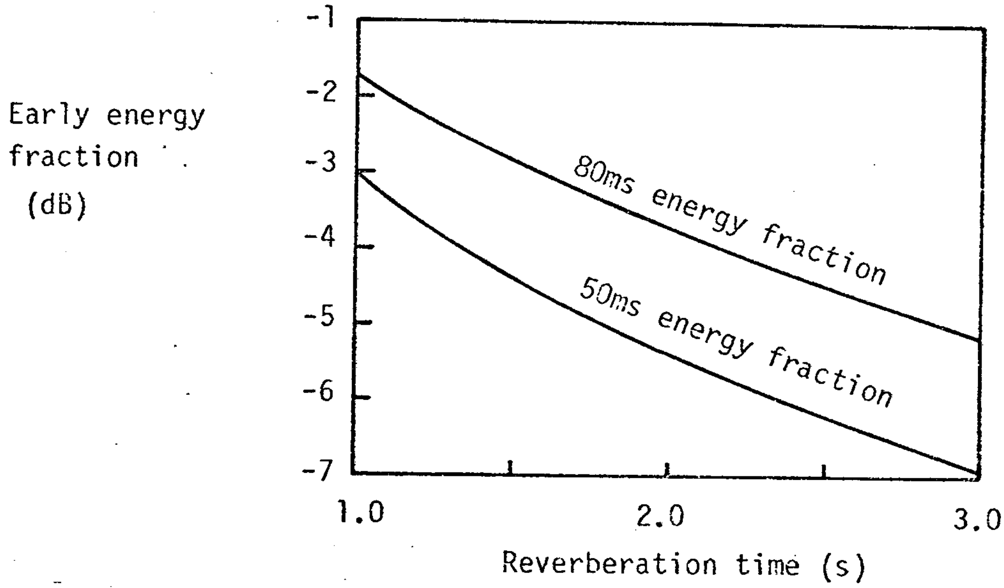


Figure 19.2. Theoretical values of the 50 ms and 80 ms energy fractions as a function of reverberation time, assuming a linear exponential decay.

The measuring system including the Gated Integrated Energy Meter is capable of measuring the energy content of the early sound at different frequencies. It was decided at the time, however, that this information could be obtained by making measurements of the early energy fraction, since to perform the latter measurements is much less time consuming. The two measures are equivalent if estimates can be made of the energy content of the reverberant field.

Statistical reverberation theory, as expressed in equation (15.8) predicts that the reverberant energy is proportional to T/V. In a hall, therefore, the variation with frequency of the reverberant energy is thus predicted as being proportional to the reverberation time. However in this theory, randomly placed absorbent surfaces and diffuse conditions are assumed, and such conditions lead to linear exponential decays (as measured in practice, and predicted, as in Figure 15.7). For a linear decay there is a predicted value for the very quantity it was hoped to be able to interpret: the early energy fraction. Thus one can only conclude that if the measured early energy fraction deviates from the statistical reverberation theory value, this theory cannot predict the reverberant energy level.

In halls with the audience area as the only area of significant absorption, four possible reasons can be suggested for deviation of the early energy fraction from that predicted by classical theory:

(a) that the source-receiver distance is small, such that the direct sound predominates (as in the computed decay in Figure 15.7 for the receiver position (15, 2.5, 2);

(b) that the hall surfaces are located to direct more (or perhaps less) early energy at the area of high absorption, the audience area;

(c) that the audience attenuation due to grazing incidence reduces the proportion of early energy especially at frequencies around 200 Hz;

(d) that surfaces exist in the hall whose reflective behaviour is distinctly frequency dependent.

The effects on the early energy fraction at different frequencies of these four situations are:

(a) a higher value than predicted at all frequencies,

(b) a higher (or lower) value than predicted, particularly at the frequencies of greatest absorption of the audience, 500-1000 Hz,

(c) a lower value than predicted in the region of 200 Hz, but the predicted value at higher frequencies,

(d) unpredictable unless precise information is available about the behaviour of the relevant surfaces.

None of the measurements made in the three halls gave a result for the position with the shortest source-receiver distance consistent with situation (a). Nor did any of the halls have obvious surfaces corresponding to those in situation (d). Situation (c) is the one under investigation; therefore situation (b) appears to be the only one which would modify interpretation of the early energy <u>fraction</u> measurements in order to determine the effect of audience attenuation due to grazing incidence on the early energy <u>content</u>. Further, it can be claimed that if the effect of situation (b) is to produce a deviation in the early energy fraction of x dB from the classical prediction at middle frequencies, the deviation around 200 Hz from this case will be less than x dB, since the absorption of upholstered seating |91| (rather than the audience here, since no audience was present for the measurements) is significantly lower at this frequency than at mid-frequencies.

In conclusion, the early energy fraction may be compared with predicted values according to a linear exponential decay. Deviation from the predicted value is likely to be significant in terms of the subjective clarity of the sound. If the measured value agrees with the predicted at middle frequencies, a variation from predicted around 200 Hz can probably be attributed to audience attenuation at grazing incidence. If the measured value exceeds that predicted at middle frequencies this may be attributed to the location of the room surfaces concentrating early sound on the area of high absorption, the seating area. The effect of audience attenuation at grazing incidence on the early energy content can be estimated, with the qualification that this concentrating effect influences the early energy fraction more at middle frequencies than at around 200 Hz.

Due to absorbent surfaces other than the seating areas and the unique behaviour of the seating for sound at grazing incidence, measurement of the early energy fraction in Winthrop Hall were not suitable for determining the energy content of the early sound, relevant to a typical hall. Measurements in the other two halls were suitable, however, for this purpose, and furthermore in each measurements were made at a seat at the front of a balcony. These provide a useful comparison since early sound at these seats does not suffer attenuation at grazing incidence.

The measuring technique was as follows: three successive single cycle pulses were radiated from the omni-directional source at each third-octave centre frequency. The gated energy in the periods 50 ms, 80 ms and 800 ms after the direct sound was measured for the signal picked up by an omni-directional microphone. Integrating up to 800 ms rather than infinity results in a theoretical error of 0.02 dB for a linear exponential decay with an R.T. of 2 seconds. By dividing the early integrated energy by the total and converting to decibels, the required result is obtained.

19.5 MEASUREMENTS OF THE EARLY ENERGY FRACTION IN HALLS

(i) 50 ms energy fraction

Measured values of the 50 ms energy fraction in the three halls are given in Figure 19.3. A key to the receiver positions is to be found in Table 19.1. As well as octave mean values for the measuring range 80-1000 Hz

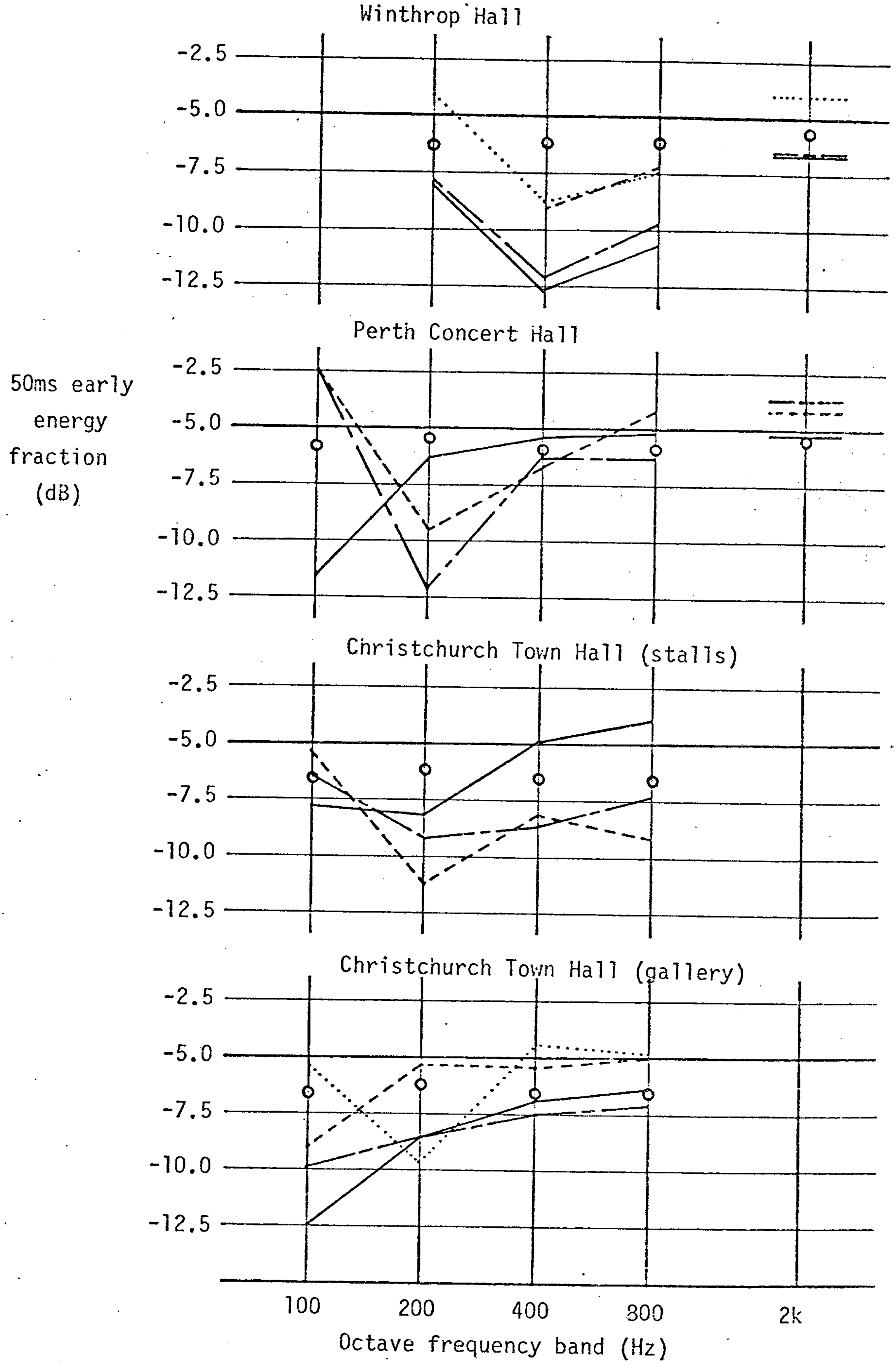


Figure 19.3. Measured values of the 50 ms energy fraction in the three halls. o, Predicted result from Figure 19.2.

(160-1000 Hz in Winthrop Hall), results at 2 kHz are also included for Winthrop Hall and Perth Concert Hall. Predicted values from Figure 19.2, according to an exponential decay with an R.T. for the empty halls, are indicated by circles.

TABLE 19.1

Key to receiver positions for Figures 19.3, 19.4 and 19.6.

Winthrop Hall	Perth Concert Hall		Christchurch Town Hall Stalls	Christchurch Town Hall Gallery
W_1		....		G_1
W_2	P_1	- - - - - - -	S_1	G_3
W_3	P_2	——— — ———	S_2	G_5
W_4	P_3	—————	S_3	G_6

The influence of audience attenuation at grazing incidence is very noticeable for both Perth Concert Hall and Winthrop Hall (the latter being affected most in the 400 Hz octave which agrees with the direct sound transmission response in Figure 18.1(a)). The influence of audience attenuation in Christchurch Town Hall is less marked and only occurs at some seats, which confirms the presence of early reflections on paths remote from the audience seating. As expected, no dip in the measurements occurs in the 200 Hz octave for the two seats near the balcony fronts: seat G_5 in Christchurch Town Hall and the seat P_3 in Perth Concert Hall.

Agreement with the predicted value at 400-800 Hz is good in Perth Concert Hall, and scattered uniformly about the predicted value in Christchurch Town Hall. However, in Winthrop Hall values at 800 Hz are significantly lower than predicted, especially for the two seats towards the back of the hall; indeed the low values for these two seats (W_3 and W_4) extend over all frequencies. Although the presence of an audience in this hall is likely to shift the frequency of maximum attenuation towards 200 Hz, such consistent deficiency of early sound, especially towards the back of the hall, agrees with the perceived lack of clarity in this hall. Interestingly, for the seat below the gallery, S_3 in Christchurch Town Hall, the energy fraction significantly exceeds prediction indicating shielding by the gallery of reverberant sound.

(ii) <u>80 ms energy fraction</u>

The measured values of the 80 ms energy fraction are given in Figure 19.4. Predicted values are again represented by circles.

In each hall the influence of audience attenuation at grazing incidence is less on the 80 ms than on the 50 ms energy fraction; evidently, in general, reflections arriving within the time period 50-80 ms after the direct sound travel on paths remote from the audience seating. Responses show less variation with position for the 80 ms energy fraction: in particular, variation with position in the 800 Hz octave is very small in each hall.

Agreement with the predicted values is again good at middle frequencies in Perth Concert Hall and below predicted values in Winthrop Hall. In Christchurch Town Hall, the predicted value for the 80 ms energy fraction at 500 and 1000 Hz is -4.8 dB, whilst the mean measured values for the 400 and 800 Hz octaves were -3.9 and -3.8 dB, respectively. In terms of the <u>ratio</u> of early-to-reverberant energy, the predicted value is -3.1 dB and the measured values -1.6 dB. The subjective degree of clarity in this hall was considered to be higher than expected for the measured R.T.; however, if the difference limen of 2 dB for the case of only direct sound plus reverberation as measured by Reichardt and Schmidt |35| is relevant, this deviation of 1.5 dB between the measured and predicted ratio of early-to-reverberant energy is only likely to be marginally significant in subjective terms.

In quantitative terms, the effect on the early energy of audience attenuation at grazing incidence for the two seats in the stalls in Perth Concert Hall is given by the deviation from the predicted value at 200 Hz of the measured early energy <u>fraction</u> (since no deviation exists at mid-frequencies), which was measured as about -3 dB. When the deviation between measured and predicted values at middle frequencies in Christchurch Town Hall is taken into account, the effect on the early energy of audience attenuation is less than -3 dB in this hall, except at seat G_1. The behaviour at this seat illustrates a situation where little of the early sound travels in paths remote from the audience seating. This is more likely to occur at seats close to the source, such as this one. The effect of audience attenuation on the early energy is -5dB here.

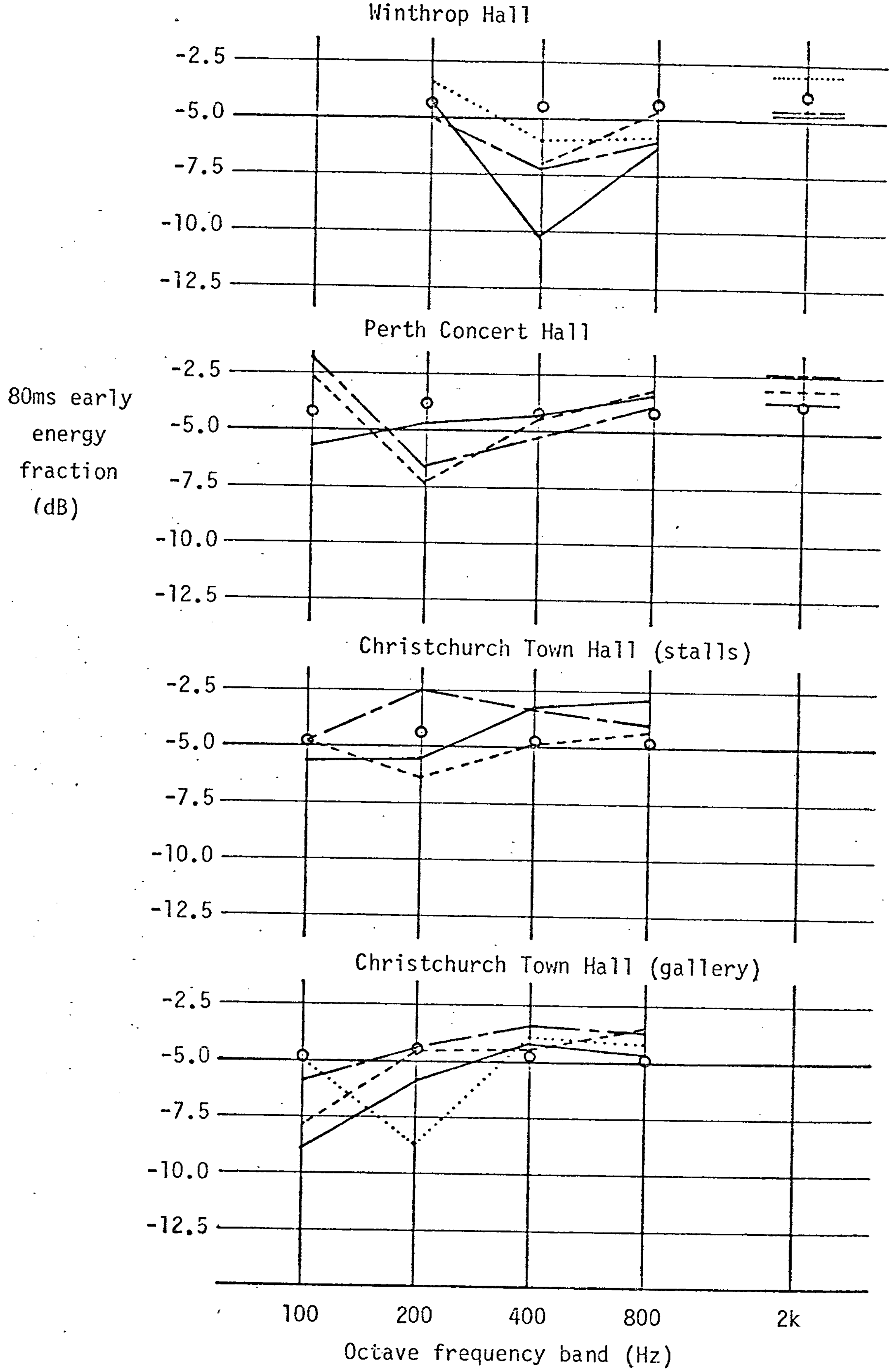


Figure 19.4. Measured values of the 80 ms energy fraction in the three halls. o, Predicted result from Figure 19.2.

19.6 THE LATERAL EARLY ENERGY FRACTION

It proves more convenient to calculate the lateral early energy fraction from measurements, than the ratio of lateral to non-lateral early energy. This has the further advantage that it avoids confusion between the measured lateral energy fraction and the subjectively relevant ratio of lateral to non-lateral sound since, due to the directionality of the figure-of-eight microphone differing from the "directionality" of the ear for spatial impression, there is not a monotonic relationship between the measured and subjectively relevant quantity. The main discussion in this section is concerned with the nature of this relationship.

The subjective response to a reflection was found, as described in chapters 8 and 10, to be proportional to the cosine of the angle ϕ of the reflection path to the lateral axis through the listener's ears. The response of the figure-of-eight microphone, orientated to receive the lateral sound, is approximately proportional to $\cos^2 \phi$; the published response was given in Figure 16.7. The consequence of this for measurements in a fully diffuse sound field may be considered first. It can be shown that for an (energy) directional response of $\lambda(\phi)$, the lateral fraction in a diffuse field (i.e., the energy received by the directional microphone as a fraction of the omni-directional energy) is given by

$$\int_0^{\pi/2} \lambda(\phi).\sin \phi.d\phi. \qquad (19.2)$$

For the subjectively relevant directional response, $\lambda(\phi) = \cos\phi$, the lateral fraction given by the integral is $\frac{1}{2}$ or −3 dB; thus the subjectively relevant ratio of lateral to non-lateral sound in a diffuse field is 0 dB, which is a convenient reference point for this measure. However, the lateral fraction in a diffuse field, for a response corresponding to the published directionality of the Neumann microphone, is −4.2 dB, whilst that for a $\lambda(\phi) = \cos^2\phi$ response is 1/3 or −4.8 dB. The error due to the different directionality responses for a diffuse field is thus 1.2 dB; consideration of typical early reflection sequences leads to even larger possible errors.

For the purpose of a comparison of the measured lateral energy fraction and the subjectively relevant ratio of lateral to non-lateral energy, the reflection sequences in two typically sized rectangular halls were chosen: 45 x 20 x 17m and 45 x 32 x 17m. At 33 and 55 seat positions, respectively, the lateral energy fraction (for the Neumann microphone) and ratio of lateral to non-lateral energy over the interval 80 ms after the direct sound were computed (see section 20.5). The resulting values for the 88 seat positions are plotted in Figure 19.5. The line corresponds to the result

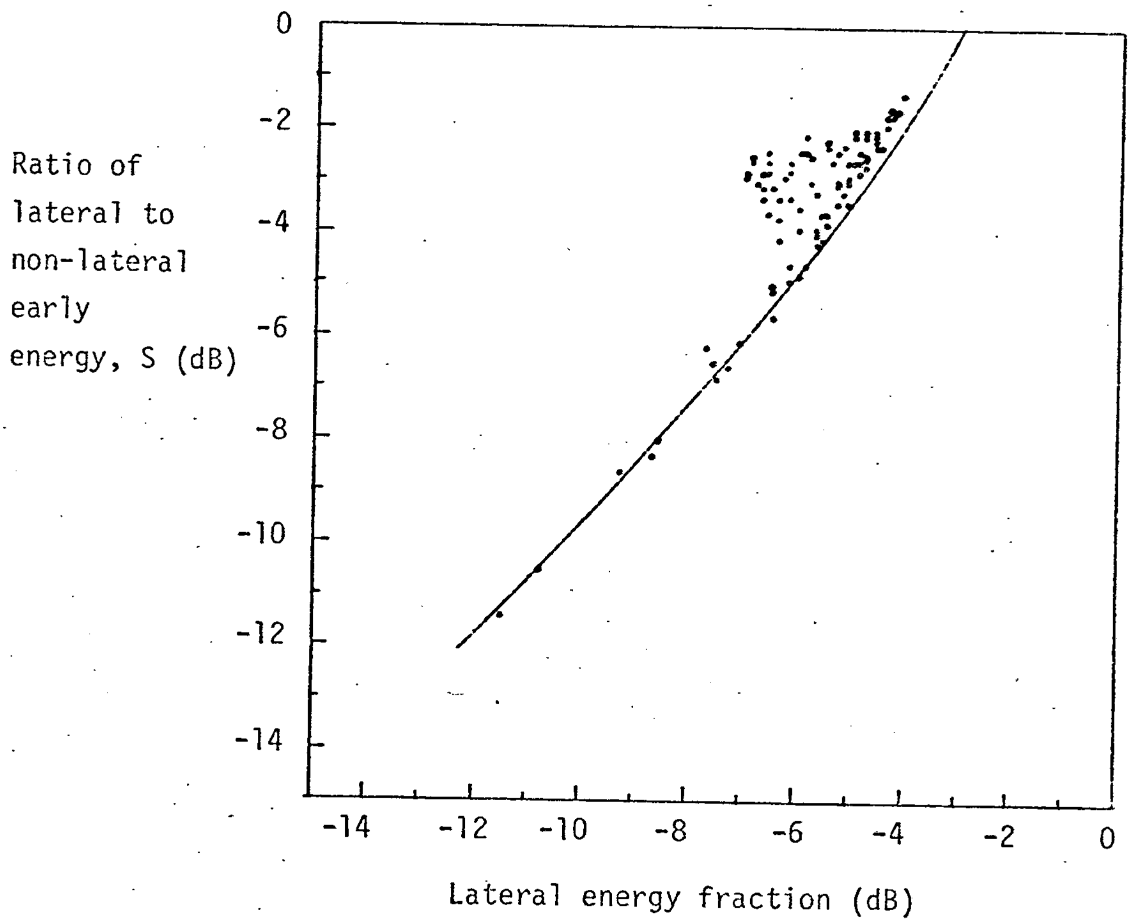


Figure 19.5. Relationship between the subjectively relevant ratio of lateral to non-lateral early sound and the physical measure of the early lateral energy fraction. Points are derived from theoretical reflection sequences in two rectangular halls.

for a lateral fraction measured with a directionality identical to the ratio of lateral to non-lateral sound. In other words, the deviation of a point from the line is a measure of the influence of the directionality of the measuring microphone used.

At seats near the source, at which the proportion of lateral sound is small, and the lateral sound comes from directions far from straight ahead, the influence of the Neumann microphone response is small. With quite insignificant exceptions, measurement with the Neumann microphone underestimates the lateral sound (which is to be expected for a $\cos^2\phi$ rather than a $\cos\phi$ response). The error can be as much as 2 dB: e.g., the subjective effect at a position with a measured lateral fraction of -7 dB may be the same as that of one with a lateral fraction of -5 dB. Within the validity of the computer model used, the points in Figure 19.5 correspond with typical situations in rectangular halls. Other shaped halls could well give points outside the main dot field.

In summary, the directionality of a figure-of-eight microphone when used to measure the lateral sound differs from the directionality of the ear for spatial impression. The result of this discrepancy is that measurements with the microphone underestimate the relevant proportion of lateral sound by as much as 2 dB. This error limits the precision with which measurements of the lateral fraction can be interpreted in subjective terms.

The measurement technique involved comparison of the energy picked up by the figure-of-eight microphone with that by an omni-directional microphone at the same position; in each case the energy is gated over 80 ms after the onset of the direct sound. The figure-of-eight microphone was positioned so that the direct sound coincided with the null position. Prior to measurement, the sensitivities of the microphone channels were adjusted to be identical at 500 Hz; the correction factors mentioned in section 16.5 were applied to the measured results.

19.7 MEASUREMENTS OF THE LATERAL EARLY ENERGY FRACTION IN HALLS

The purpose of these measurements was to determine what actual values of the lateral energy fraction occur in real halls, and whether measurements of the energy fraction at the frequency of maximum attenuation at grazing incidence differ significantly from measurements at other frequencies. Measured values of the 80 ms lateral energy fraction in the three halls are given in Figure 19.6. Results at 2 kHz are also included for Winthrop Hall and Perth Concert Hall.

In Perth Concert Hall, which corresponds perhaps closest with a typical rectangular concert hall, the measured values in the stalls were on average about -6.5 dB, which corresponds to a subjectively relevant ratio of lateral to non-lateral early sound of greater than -5.5 dB. The measured values in the first balcony (P_3) are generally lower, whilst at 2 kHz the

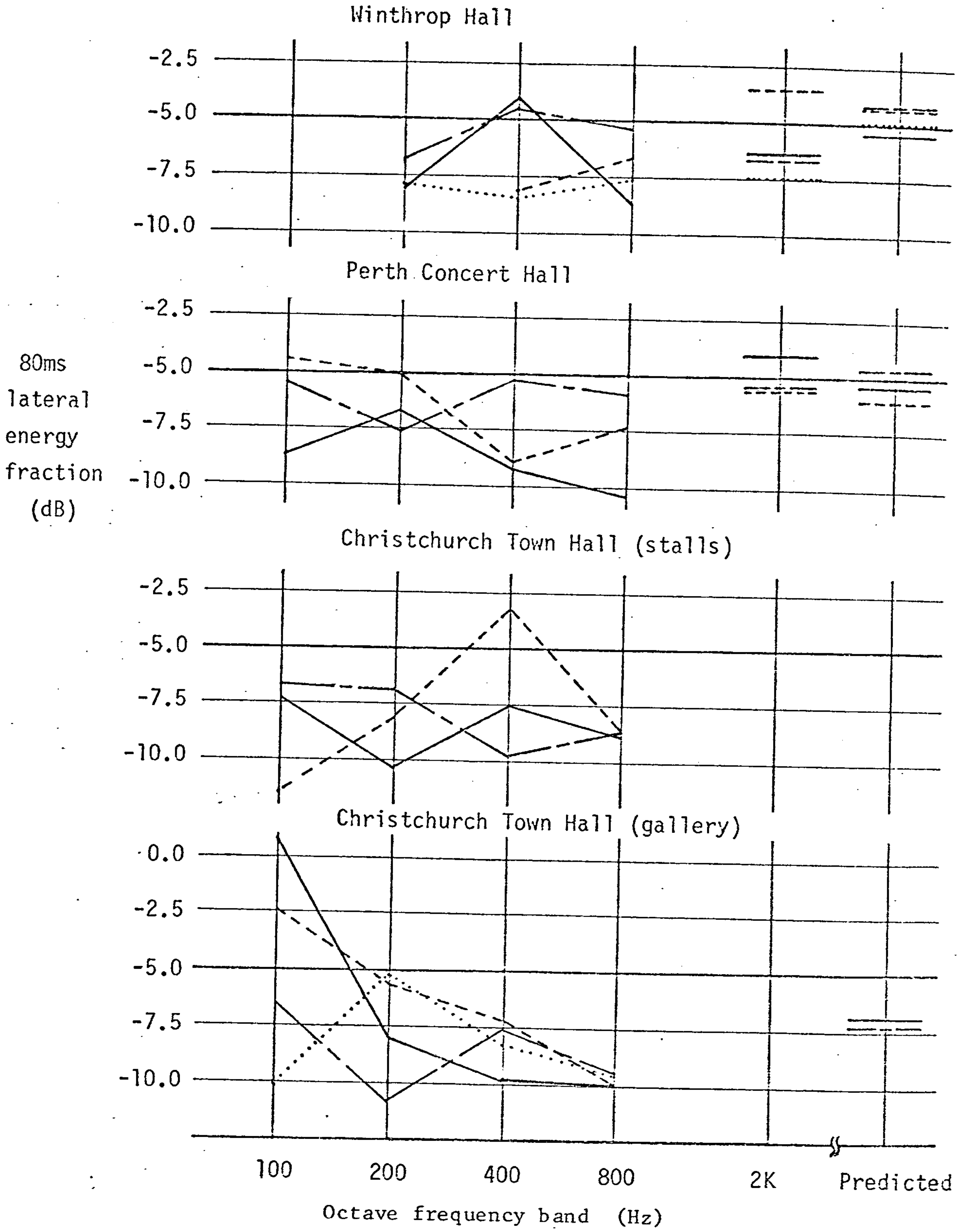


Figure 19.6. Measured values of the 80 ms lateral energy fraction in the three halls.

value, in fact, is higher than that measured in the stalls. This behaviour is difficult to explain. Measurements in Winthrop Hall gave similar values to those measured in Perth Concert Hall, which is reassuring since the two halls were very similar in shape.

Except at 800 Hz, measurements in Christchurch Town Hall showed a wide degree of spread, and included one obviously erroneous result of an energy fraction greater than 0 dB. (This was for a seat at a position near the line of symmetry down the centre of the hall; at low frequencies interference in such a position is very likely to influence results.) No significant differences exist between results in the stalls and the gallery. The mean value at middle frequencies is about -9 dB, which is significantly lower than in the rectangular halls but nevertheless above threshold.

No marked reduction in the measured lateral fraction at the frequency of maximum audience attenuation was measured in any of the halls; in fact increases in the fraction were measured as frequently as reductions. Whilst this behaviour confirms a design feature in Christchurch Town Hall, in the rectangular halls it indicates the presence of lateral sound in paths remote from the audience seating, such as cornice reflections.

19.8 CORRESPONDENCE BETWEEN THE PREDICTED AND MEASURED 80 ms LATERAL ENERGY FRACTION IN HALLS

The predicted values for the 80 ms lateral energy are also included in Figure 19.6 for the three halls; in each case the prediction is frequency independent. For Winthrop Hall and Perth Concert Hall the predicted value is simply the difference between the predicted value of the total and lateral integrated energy at 80 ms in Figures 18.5 and 18.6. Section 18.2 contains a discussion of the validity of these predictions. Agreement is relatively good in both halls at 2 kHz but measured values at lower frequencies are generally lower. Without further detailed examination in these halls it is difficult to explain why this should be.

The lateral fraction was calculated from the computed reflection sequences in the gallery of Christchurch Town Hall. By chance, the two seats towards the front of the hall, G_1 and G_3, contain no lateral sound as predicted by geometrical ray acoustics, and the predicted value for the lateral fraction is about -20 dB. The fact that the measured values of the lateral energy fraction are similar to those in other seats indicates the presence of significant non-specular reflections within 80 ms of the direct sound. The predicted values for the other two seats in the gallery exceed measured values by about 2 dB.

19.9 CONCLUSIONS

It was argued that acoustic quality relevant to seating areas is best determined by measuring quantities corresponding to incoherent addition of reflection energies. These were obtained by using a short duration test signal and with measurements over third octaves, the results being presented as averages over an octave.

Comparison of measured values of the 50 ms and 80 ms energy fraction with those predicted for a linear exponential decay with the relevant R.T. was made in the three halls. For both time fractions, agreement was good in Perth Concert Hall, whilst measured values were below those predicted in Winthrop Hall, especially towards the back of the hall, a result which agrees with perceived lack of clarity in this hall. The measured value of the 80 ms energy fraction differed by just less than a subjectively significant amount from the predicted in Christchurch Town Hall, whilst no consistent behaviour emerged for the 50 ms energy fraction. The perceived high level of clarity in this hall is thus not fully explained by these measurements.

Extrapolation from measurements of the 80 ms early energy fraction to determine the effect of audience attenuation at grazing incidence on the 80 ms total early energy indicated a fall in the early energy around 200 Hz of about -3 dB.

Measurements of the lateral early energy fraction were made with a figure-of-eight microphone. Since the microphone directionality differs from the relevant directionality of the ear for spatial impression, no precise monotonic relationship exists between the measured quantity and the subjectively relevant one. Measurements of the 80 ms lateral energy fraction in both Perth Concert Hall and Winthrop Hall were on average -6.5 dB and in Christchurch Town Hall -9 dB on average. No marked reduction in the lateral energy fraction at 200 Hz was measured.

The following chapters contain the results of computer investigations to determine the effect of hall shape on the subjectively relevant ratio of lateral to non-lateral early sound. Further discussion of the subjective significance of measurements reported in this chapter will be delayed until Chapter 23, in which the subjective significance of the computer investigations will also be discussed.

Chapter 20

PRINCIPLES OF THE COMPUTER INVESTIGATION OF DIFFERENT HALL SHAPES

20.1 INTRODUCTION

Given that it is quantities corresponding to incoherent addition of reflection energies which are of interest (see section 19.2), the problem for a computer investigation remains that only a few reflections are purely specular in real halls (see sections 17.4-7). However, no satisfactory model exists to define non-specular reflection. In the absence of such a model, two approaches are possible: (i) that in energy terms the nature of the reflection (specular or not) is not significant; (ii) that reflection off a diffusing surface is such as to randomise reflection direction. The second approach has been used by Kuttruff |84|, but it is only suitable for dealing with an acoustical situation in a room which can be treated statistically: i.e., the reverberant decay. Kuttruff used this approach to derive reverberation formulae for the case of a highly absorbent floor with diffusing room surfaces. The former approach has been used by Krokstad, Strøm and Sørsdal |83| and by Marshall |98|. In its favour as a method is the fact that when the reflection density is low, and the response cannot be averaged, reflections are most clearly specular. Later reflections, which in reality are generally more diffuse, can be considered as being equivalent to the average behaviour of the predicted specular reflections during these later time periods. For two rectangular halls, however, this last assumption was found in section 18.2 not to be the case; it was found that the integrated energy in the back half of the halls exceeded values predicted on a basis of specular reflection. It was found, however, that the measured proportion of lateral sound agreed with the predicted value. This is a result which may well not extend to different shaped halls.

Such are the limitations on the validity of computer predictions on the basis of specular reflection. It can be hoped that predicted <u>variations</u> of the quantity with a physical hall dimension will at least be in the same direction as would occur in reality.

When the assumption of specular reflection is used, two methods are available to calculate the arrival time of reflections. In that used by Krokstad <u>et al</u>. |83| the source is considered to emit a finite number of rays, each contained in a fraction of the total solid angle. The paths of these diverging rays are then calculated, specular reflection is assumed whenever the ray strikes a wall surface; the ray is considered totally absorbed when it strikes a region of audience seating. The advantage of this procedure is that it can deal with relatively complex room shapes, and computing time is roughly linearly proportional to the time period over which it is desired to "count" reflections.

The second method, used by Marshall |98|, does not require the assumption of a spherically divergent pulse being divided into a finite number of rays. The specular images of the source in each room surface are calculated, and if the path between the image and receiver cuts the surface at a real (not imaginary) point and is not obscured by any other surface, it constitutes a reflection. The method, however, becomes progressively more time consuming the higher is the order of reflections to be investigated: for second order reflections every possible pair of surfaces has to be tested, etc. The method is only suitable for lower order reflections, and, if periods up to, say, 100 ms after the direct sound are to be investigated, this limits the application of the method to large halls. Marshall used the programme to investigate reflections up to second order in Christchurch Town Hall.

The method used to calculate reflections in this study was essentially a slight modification of the second method above and is relevant to simple geometrical shapes of hall. All the results reported in the next chapter refer to rectangular halls; the unique characteristic of two surfaces at right angles is that there is a single second order image of a source in the two surfaces, independent of the receiver position. Thus, for a rectangular hall the images of a source are readily calculated since they are independent of the receiver position. In Chapter 22 three halls were investigated in which in plan or in cross-section the form was not rectangular; images of the source in this plane were calculated according to the second method above, whilst replication of the image field in orthogonal surfaces could be further readily calculated.

20.2 THE BASIC COMPUTER PROGRAMME FOR RECTANGULAR HALLS

The following assumptions were made for this study: that all wall and ceiling surfaces were perfectly reflective and produced specular reflections; that the floor area was totally absorptive. With the above conditions, in the general case all images were considered as contributing a reflection except reflections at source-receiver height off the front wall (i.e., the end wall nearest the orchestra); such a reflection was considered obscured due to the presence of other members of the orchestra. No stage enclosure was assumed in this study. The image array in plan for an off-axis source is illustrated in Figure 20.1. With the same notation as in

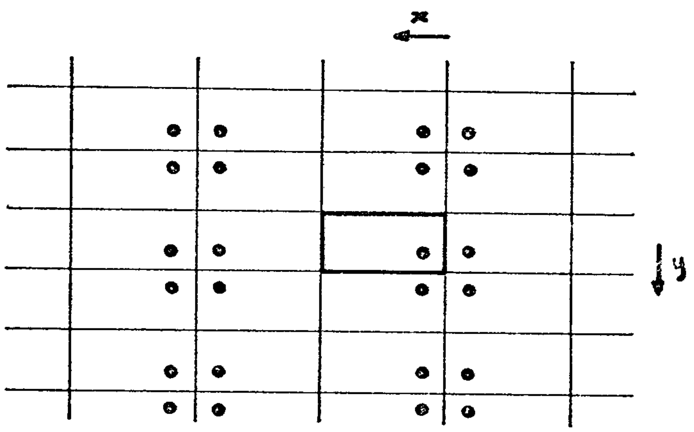


Figure 20.1. Plan view of the image array of an off-axis source in a rectangular room.

Chapter 15, for halls longer than 17 m only images in cells with $2 \geqslant \ell \geqslant -1$ need be considered for delays up to 100 ms. With a fully absorbent floor, ceiling reflections are contained in an identical array directly above that illustrated in Figure 20.1. An omni-directional source is assumed throughout.

The computer programme "asked" the hall dimensions, the source and receiver positions. The direct sound travel time was then calculated. Starting with $m = 0$, the image positions for the range of ℓ-values were calculated. Subsequently the modulus value of m was incremented by unity, the corresponding y-values for the images calculated, and the relevant image coordinates for the range of ℓ-values were determined. This latter procedure was continued until both $(0, \pm m, 0)$ images gave reflections with delays greater than 100 ms.

For the resulting image array, the required images were then eliminated (e.g., $(-1, m, 0)$ images as mentioned above), and the travel time for ceiling reflections was also calculated. The intensity of reflections relative to the direct sound was calculated according to spherical divergence, from the reflection travel time, T, and the direct sound travel time, T_o:

$$\text{Reflection intensity} = I_r = \frac{T_o^2}{T^2} \; . \qquad (20.1)$$

The speed of sound was taken as 343 m/s. The basic programme for rectangular halls in FOCAL is included in Appendix III. The various forms of print-out - as a listing of reflection details or as an echogram, as the ratio of lateral to non-lateral early sound, etc. - will be described later.

20.3 THE BASIC PROGRAMME FOR NON-RECTANGULAR HALLS

For the case of a hall with a non-rectangular cross-section, but orthogonal front and back walls, as used for the theoretical study of the "Maltings", Snape in section 22.2, the following procedure was used. Two subroutines were devised: one to calculate the image of a point in a line, and the second to calculate the point of intersection of a "ray" with a wall surface. For a second order reflection, for example, an ordered pair of wall surfaces (parallel to the x-axis) was chosen, the image of the source in the first, and the image of the primary image in the second surface were calculated. The points of intersection of the ray on a path between the final image and the receiver and the respective wall surfaces were calculated. If in each case the ray strikes a real point on the wall surface, the final image contributes a real reflection. The coordinates of the relevant images for the range of ℓ-values was then calculated, before the next ordered pair of surfaces was chosen. All possible ordered pairs of surfaces were tried.

In other respects a similar procedure to that used for rectangular halls was employed. Reflections up to third order in cross-section were calculated, which is adequate for reflection delays up to 100 ms.

20.4 SUBSIDIARY PROGRAMMES FOR REFLECTION DETAILS AND ECHOGRAM PRINT-OUT

Reflection details could be presented as a list ordered by reflection delay starting with the direct sound. As well as the reflection level in dB relative to the direct sound, the angle of azimuth of the reflection to the line of sight between the receiver and a central (i.e., y = 0) source was calculated, as was the angle of elevation of the reflection path to the horizontal. The coordinates of the image cell corresponding to the particular reflection was also listed, (U, V) being the coordinates (ℓ, m) and the prefix "C" being used to designate a ceiling reflection.

For a more graphical representation, a programme was written to present the details in the form of an echogram, as is shown in Figure 18.4. The delay is indicated along the abscissa, whilst in this case the intensity along the ordinate is that which would occur on an oscilloscope screen (intensity relative to direct sound = T_o/T). Against each reflection the coordinates of the respective image cell were also listed. Both the reflection detail and spatial impression programme, as described below, are included in Appendix III.

20.5 SUBSIDIARY PROGRAMME TO CALCULATE THE RATIO OF LATERAL TO NON-LATERAL EARLY ENERGY

To obtain the integrated intensity of reflections arriving within a certain time period as a multiple of the direct sound intensity, reflection intensity values according to equation (20.1) are summed. According to the findings in Part I, the subjectively relevant early lateral sound intensity was calculated, in addition to the total early sound intensity. The ratio of lateral to non-lateral sound energy is then given by the lateral/(total - lateral) early sound intensity.

From Chapter 14, the subjectively relevant early lateral sound intensity is given by

$$\text{Lateral sound intensity} = \sum_{\tau\,=\,5\text{ ms}}^{80\text{ ms}} I_r \sin\alpha_r \cdot \cos\beta_r, \qquad (20.2)$$

where I_r is the reflection intensity, and α_r and β_r are the reflection angles of azimuth and elevation. Rather than using discrete limits for τ, the delay time, for a reflection to contribute to the early lateral sound intensity, the intensity was linearly weighted in the regions τ = 3-8 ms and τ = 70-90 ms, a procedure closer to the subjective behaviour. The angles of azimuth and elevation were calculated as described in the previous section. The ratio of lateral to non-lateral early sound was printed in dB.

To determine the influence of audience attenuation at grazing incidence, various methods of presentation of results are possible. It was decided to calculate the ratio of lateral to non-lateral early sound at the frequency of maximum audience attenuation, and to calculate the total early sound energy at this frequency relative to sound energy at other frequencies unaffected by audience attenuation at grazing incidence. These quantities completely define the physical situation and both can be appreciated in subjective terms; the former has been the predominant subject of study in Part I, whilst the latter indicates the degree of filtering of the total early sound. For results compatible with measured values, the attenuation at grazing incidence was taken as 10 dB for the 200 Hz octave. After the results of Schultz and Watters |14|, the sound path had to be at an elevation of more than 30° to the seating area for the attenuation not to occur.

To investigate the significance of the early energy level factor on the subjective degree of spatial impression a separate study was made of the early energy levels in the first 80 ms after the direct sound in halls. Results are purely relative; the direct sound level at 10 m from the source was used as a standard, which gave mean results of the order of 0 dB.

Chapter 21

THE DEGREE OF SPATIAL IMPRESSION IN RECTANGULAR HALLS ACCORDING TO COMPUTER PREDICTIONS

21.1 INTRODUCTION

The subjective degree of spatial impression has been seen in Part I, to be determined by two factors: the proportion of lateral early sound and the total early sound level. Both these quantities were investigated in this study, though they will be dealt with separately. For comparison between halls, the important subjective quantities are the difference limens for the two factors. The difference limen for the ratio of lateral to non-lateral sound was found to be 1.4 dB (see section 9.1), whilst that for the early sound level, as it affects spatial impression, was found to be 4 dB (see section 9.2).

The first section of this study concerns the ratio of lateral to non-lateral early sound and the influence of audience attenuation due to grazing incidence. Interpreting the virtues of any particular shape hall is considerably complicated by the necessity to consider the behaviour of three quantities: the middle and low frequency ratios of lateral to non-lateral sound and the low frequency early energy relative to that at middle frequencies. The three quantities, whose derivation has been discussed in the previous chapter, will be called the mid-frequency ratio (of lateral to non-lateral early sound), the 200 Hz ratio and the 200 Hz band level (relative to mid-frequency). No information is available regarding the difference limens for the two latter quantities, but they are unlikely to be smaller than the full frequency values, 1.4 dB and 4 dB, respectively.

From a list of 19 approximately rectangular halls taken mainly from Beranek |41|, the median hall dimensions were found to be length 45 m, width 22m and height 17m. In fact the spread of values for the length and height are both small, though in each case extreme values are encountered, whilst values of the width vary widely between 18m and 39m. A hall with dimensions (45m x 20m x 17m), very similar to the Musikvereinsaal, Vienna, will be called the "average" hall for this study and a hall of dimensions (45m x 32m x 17m), similar to those of the Royal Festival Hall, will be called the "wide" hall. A characteristic source distance from the front wall was found to be 6.5m, and, unless otherwise mentioned, an on-axis source was used. In the average hall, the seating area was assumed to extend from 12m to 45m longitudinally from the front wall, and to 9m on each side of the centre line along the hall. This area was divided into 3m x 3m squares and values of the relevant quantities were calculated for points at the centres of the squares. Corresponding seating areas were used for halls of different shape.

21.2 VARIATION OF THE RATIO OF LATERAL TO NON-LATERAL EARLY ENERGY WITHIN TWO RECTANGULAR HALLS

The calculated values for the mid-frequency ratio in the average and wide halls for an on-axis source are shown in Figure 21.1, with contours for whole number values in dB's. For clarity areas rather than contour lines are labelled: -1 refers to values between -1.0 and -1.9 etc. Due to symmetry only half of the halls are shown. As one would expect, for positions near the source the values of the ratio are small due to the predominance of the direct sound. What is surprising, however, is that for the back two-thirds of the average hall the value of the ratio is very uniform,

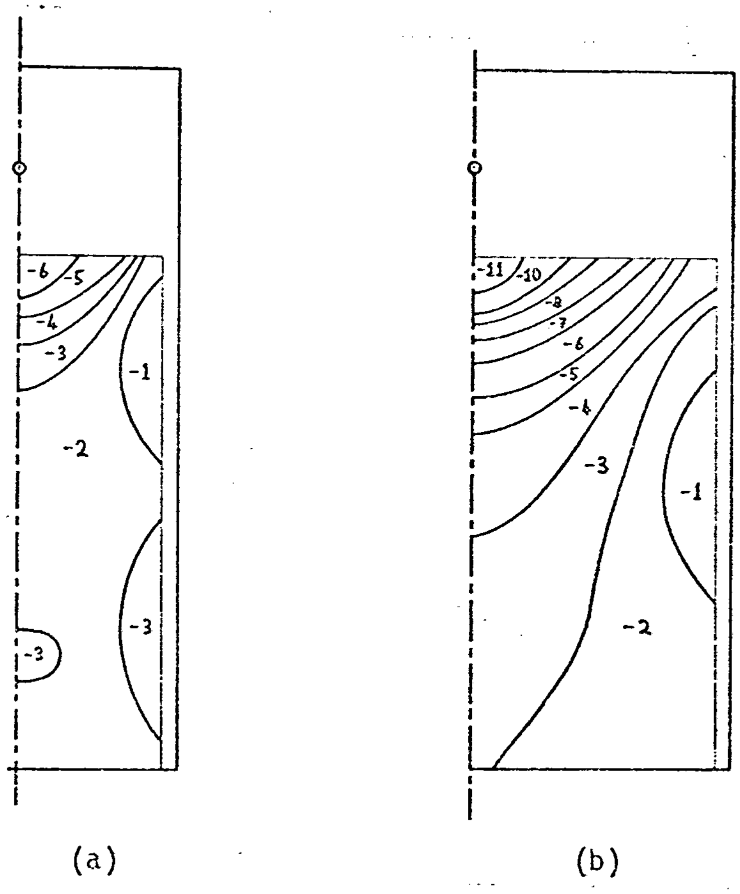


Figure 21.1. Computed values of the mid frequency ratio of lateral to non-lateral early energy in (a) the average hall and (b) the wide hall. Source position is on the axis of symmetry.

variation being within the difference limen of 1.4 dB: i.e., in this region the subjective degree of spatial impression is the same. The mean and median values of the ratio in the average hall are -2.8 and -2.6 dB, respectively.

In the wide hall, the variation with position is larger, though the characteristic behaviour is similar. In this case, seats nearer the side walls are to be preferred. The mean and median values in the wide hall are lower: -4.0 dB and -3.3 dB, respectively.

With an off-axis source 6m from the centre line in the average hall, the mean value of the mid-frequency ratio is -2.8 dB and the median value is -2.5 dB. Although the average values are essentially identical, the variation in the ratio with position is larger than with the central source, as is shown in Figure 21.2. Again low values are found in the region of the

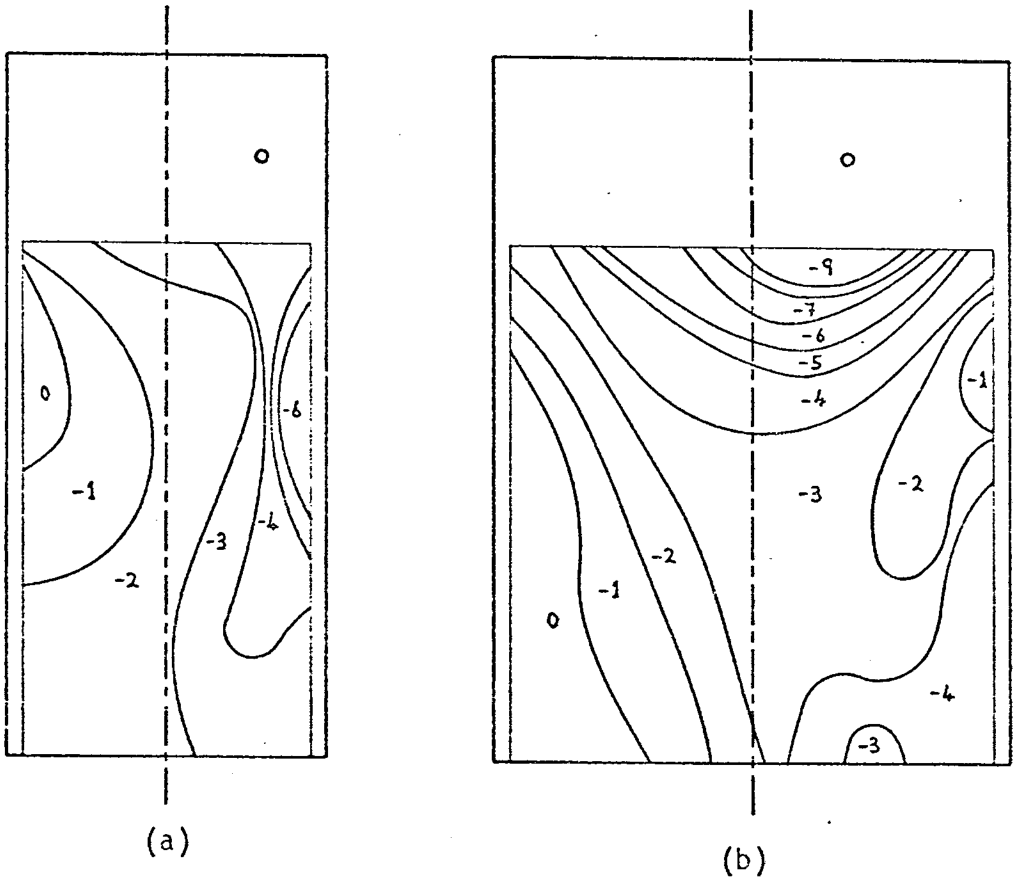


Fig.21.2. Computed values of the mid-frequency ratio of lateral to non-lateral early energy in (a) the average hall and (b) the wide hall with a lateral source position.

source, but in the back half of the hall there is a variation of the order of one subjective unit from one side to the other. Thus, if bass lateral sound is particularly desirable, as has been suggested in Chapter 11, positions on the opposite side of the hall to the bass instruments are preferable. With conventional orchestral placing this would mean that seats on the left side of a rectangular hall are to be preferred - subjective confirmation of this behaviour would lend support to the suggestion that bass-frequency lateral sound is important. (In fact it is the 200 Hz ratio which is important here, but the variation with position in the hall of the 200 Hz ratio is very similar to that of the mid-frequency ratio.)

Also shown in Figure 21.2 is the situation of an identically placed off-axis source in the wide hall. For this hall the difference between seats on opposite sides of the hall exceeds two subjective units in general. The median value for this case is again similar to that for a central source, namely -3.4 dB.

The computed values for these two halls with a central source were used in Figure 19.5 to compare the subjective relevant quantity with that measured by microphone.

21.3 VARIATION OF SPATIAL IMPRESSION WITH HALL WIDTH AND HEIGHT

For this study mean values and the standard deviation of values for a central source position will be considered. The physical situation at positions near the source are of little interest since the actual angle subtended at the listener by the orchestra is similar or even larger than the maximum perceived apparent source widths in halls due to lateral reflections. In these seats the perceived spatial impression, which is generally low, is unlikely to be subjectively significant in an overall judgement of the acoustical situation. Therefore in the remainder of this chapter only the seating area extending from 18 to 45m longitudinally will be considered. Whilst the median value is probably more relevant in subjective terms, the mean value for the reduced seating area, which is more readily computed, is almost identical to the median value. Both the mid-frequency and 200 Hz ratio of lateral to non-lateral energy will be considered as well as the 200 Hz band total level. For each quantity the spread of the results is expressed by the standard deviation in the following form: (mean) ± (standard deviation). For a normal distribution this range contains 68% of the results.

With the average hall dimension of length 45m and height 17m, the effect of varying the width from between 14m and 38m was investigated. The variation with hall width of the mid-frequency ratio is plotted in Figure 21.3. The mean value for the average hall is -2.5dB, whilst for the other halls it ranges from -2.4 to -3.9 dB. Whilst a definite relationship appears

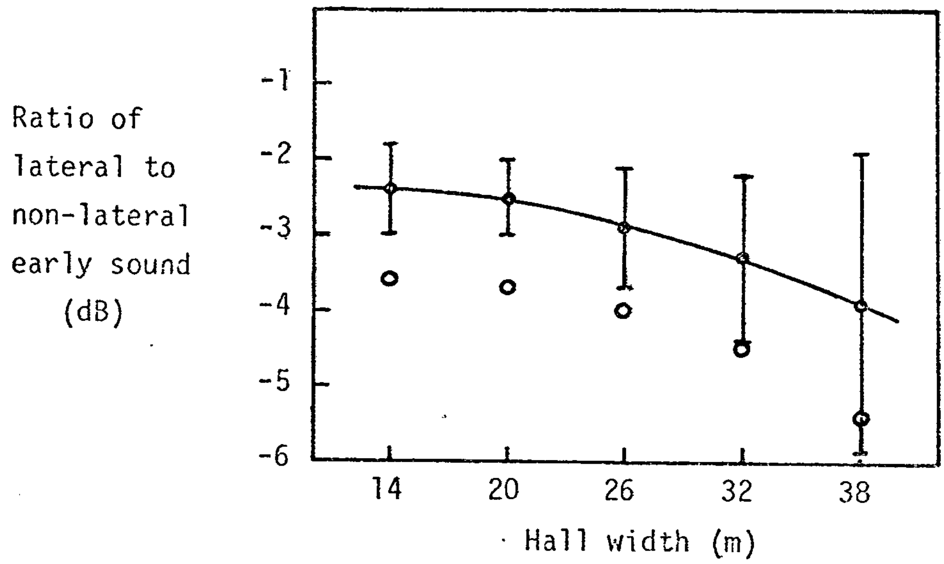


Figure 21.3. Computed values of the ratio of lateral to non-lateral early sound as a function of hall width for rectangular halls. ↨, Mean and standard deviation at mid-frequencies; o, mean for the 200 Hz octave band.

to exist, the variation between a very narrow hall and a very wide one is only one subjective unit of spatial impression. More significant, perhaps, is that the range of values is larger in the wider halls, which means that the proportion of seats with low degrees of spatial impression is larger in the wider halls, as can also be seen in Figures 21.1 and 21.2.

At first sight the small variation in the ratio of lateral to non-lateral early sound is surprising for such a wide range of hall widths, for which the delay of the first lateral reflection for a particular seat position may vary by more than a factor of two, and hence its level by more than 6 dB. Further, in the narrow halls not only are first order lateral reflections more intense but also second order. There are various reasons why variation in the ratio is small.

Firstly, the value is an average over a successively larger seating area. If a central receiver position is considered, for instance 31.5m from the front wall, the variation in the ratio is from -1.9 dB to -5.2 dB between the narrowest and widest hall. However, in the wide hall the delay of the first lateral reflection is small for a receiver position close to the wall, and at the same longitudinal distance the ratio is -1.4 dB near the wall.

A second reason for the small variation between narrow and wide halls is that whilst the intensity of lateral reflections is high in a narrow hall, towards the rear of the hall they arrive too early to contribute to spatial impression and the angle of azimuth for these reflections is also small, such that they also contribute to the non-lateral sound. As the hall becomes wider, so the angle of azimuth of lateral reflections increases such that the decrease in reflection level is partially compensated by increased spatial effect due to the larger angle of azimuth.

The mean values for the 200 Hz ratio of lateral to non-lateral are also plotted in Figure 21.3. The scatter in the results is very similar to that for the corresponding mid-frequency ratio, whilst the mean value is uniformly about 1.3 dB less than the mid-frequency mean value. This is again perhaps surprising if one considers the magnitude of the audience filtering effect; however, since the ceiling reflection image array replicates the source-receiver plane array, both contribute a similar proportion of lateral sound. As will be seen in section 21.5 cornice reflections are critical in this respect.

The more significant effect of audience attenuation filtering is on the 200 Hz band total level. The mean and standard deviation of this quantity for the different width halls is plotted in Figure 21.4. The mean

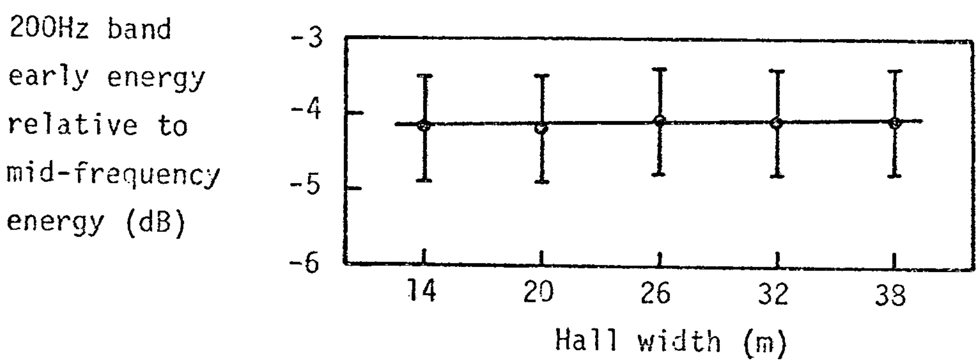


Figure 21.4. Computed values of the 200 Hz octave band early energy level relative to the mid-frequency energy level as a function of hall width for rectangular halls. ⸸ , Mean and standard deviation.

value is independent of hall width at -4.1 dB, and the scatter is small and constant for the different hall widths.

In subjective terms the effect of audience attenuation filtering is to produce a subjective degree of spatial impression at 200 Hz of two units below that at mid-frequencies. There is one unit difference due to the lower proportion of bass lateral sound, and a further unit difference due to the lower level of total early bass sound. This behaviour is independent of hall width.

Graphs of the same quantities for a variation of hall height from 12m to 22m, for a hall of length 45m and width 20m, are given in Figures 21.5 and 21.6. Due to the source-receiver plane and ceiling reflection plane containing identical image arrays, the variation in the mid-frequency ratio of lateral to non-lateral early sound is virtually independent of ceiling

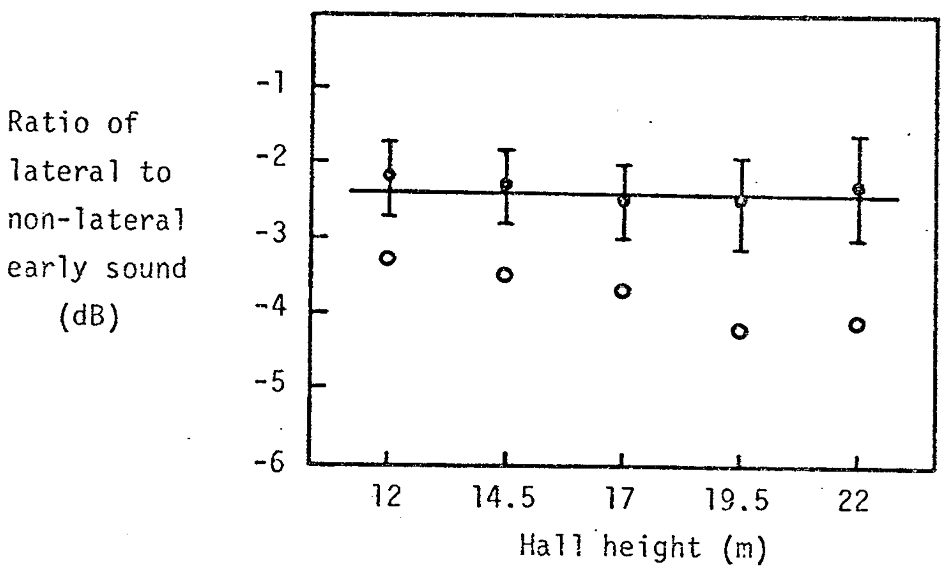


Figure 21.5. Computed values of the ratio of lateral to non-lateral early sound as a function of hall height for rectangular halls. ɸ, Mean and standard deviation at mid-frequencies; o, mean for the 200 Hz octave band.

height. Also, as one would expect, the 200 Hz ratio decreases as the hall becomes higher, but the variation is again small. The most marked variation with hall height occurs for the 200 Hz band total level, which curiously has the highest mean and smallest deviation value at the median height of contemporary halls. For low-ceilinged halls, the ceiling reflections at the

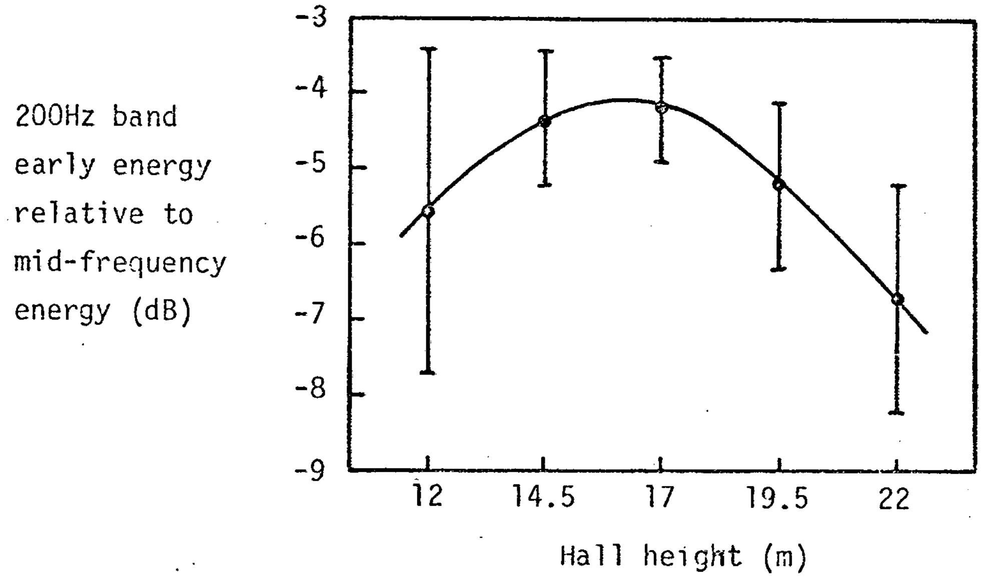


Figure 21.6. Computed values of the 200 Hz octave band early energy level relative to the mid-frequency energy level as a function of hall height for rectangular halls. ɸ, Mean and standard deviation.

back of the hall also suffer audience attenuation filtering due to their low angle of elevation, whilst for high-ceilinged halls, the ceiling reflection energy is low, especially towards the front of the hall. Thus from the point of view of good bass sound in rectangular halls a ceiling height of about 17m is to be preferred.

Two further halls were studied, with dimensions of length, width and height of (45 x 32 x 12m) and (45 x 26 x 22m). The former is a low, wide hall more suited for speech than musical performance. Together with the values for the average hall and the wide hall, the mean and standard deviation values for the three quantities are given in Table 21.1. It can

TABLE 21.1

Computed spatial impression in four rectangular halls.

Hall dimensions (m)	Mid-frequency ratio (dB)	200 Hz ratio (dB)	200 Hz band level (dB)
45 x 20 x 17 (average hall)	-2.5 ± 0.5	-3.7 ± 0.4	-4.2 ± 0.7
45 x 32 x 17 (wide hall)	-3.3 ± 1.1	-4.5 ± 1.7	-4.1 ± 0.7
45 x 32 x 12	-2.8 ± 1.0	-4.6 ± 2.2	-5.7 ± 2.2
45 x 26 x 22	-2.8 ± 0.8	-4.7 ± 0.8	-6.6 ± 1.4

be seen that the mid-frequency and 200 Hz ratios are typical for the width of the hall, whilst the 200 Hz band level is determined by the height of the hall. Again the 200 Hz band level is low at the back of the low hall and the front of the high hall.

21.4 VARIATION OF SPATIAL IMPRESSION WITH SOURCE POSITION IN THE AVERAGE HALL

The conclusions concerning variation of spatial impression with hall shape discussed in the previous section are obviously of little value if they only apply to the particular chosen source position. The mean and standard deviation values for the three quantities are given in Table 21.2 for four source positions in the average hall. The source position coordinates (longitudinal distance, lateral distance) are given relative to an origin at the centre of the front wall.

TABLE 21.2

Computed spatial impression for four source positions in average hall.

Source coordinates (m)	Mid-frequency ratio (dB)	200 Hz ratio (dB)	200 Hz band level (dB)
(6.5, 0)	-2.5 ± 0.5	-3.7 ± 0.4	-4.2 ± 0.7
(3, 0)	-2.6 ± 0.4	-3.7 ± 0.5	-3.6 ± 0.4
(10, 0)	-2.3 ± 0.7	-3.4 ± 0.6	-5.1 ± 1.0
(6.5, 6)	-2.7 ± 1.3	-3.3 ± 0.8	-4.1 ± 0.7

For each quantity the mean values and deviations vary little with source position. Since the value for the chosen central source position corresponds to an average value for a range of source positions, it was considered adequate to limit investigation to a single central source position.

21.5 VARIATION OF SPATIAL IMPRESSION WITH HALL CHARACTERISTICS

A series of "alterations" to the basic average hall were investigated to determine what was important for a high degree of spatial impression. For some "alterations" the effect on other hall shapes was also investigated. Table 21.3 contains the mean values of the three quantities for "alterations" to the average hall; deviations are only given when they differ significantly from the value for the basic hall. The "alterations" will be discussed in turn.

If instead of a flat floor seating area, the floor is raked with a typical angle of rake of 10°, the variation in spatial impression is found to be quite minimal. In many halls there are peculiarities in the front or

TABLE 21.3

Computed spatial impression for "alterations" to average hall.

	Mid-frequency ratio (dB)	200 Hz ratio (dB)	200 Hz band level (dB)
Average hall	-2.5 ± 0.5	-3.7 ± 0.4	-4.2 ± 0.7
with raked seating	-2.4	-3.7	-4.0
with absorbent front and back walls	-2.1	-3.1	-4.6
with no second or higher order lateral reflections	-3.8	-4.5	-3.9
with no cornice reflections	-2.7	-5.1 ± 1.0	-5.2
with no lateral ceiling reflections	-2.9	-7.2 ± 0.8	-6.1

back walls which cause reflections off these surfaces to be either absorbed or diffused. An extreme situation was investigated in which the front and back walls were treated as being totally absorbent. In spite of this extreme "alteration" the values of the three quantities are only slightly affected in subjective terms. The same "alteration" to a 14m and 32m wide hall also only produced very small changes in the three quantities, in each case less than a 1/3 of a subjective unit. Less extreme "alterations", which might well occur in practice, causing a transverse row of images to be non-existent, produce even less change in the three quantities.

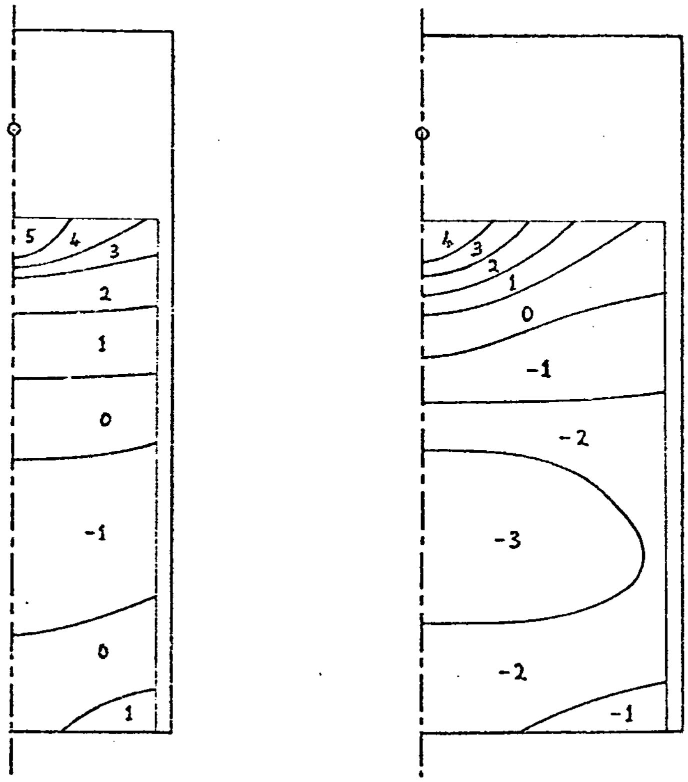
More out of interest to determine their contribution to overall spatial impression than that it might be of possible physical significance, the situation was investigated where no second or higher order lateral reflections occur. One can speculate that this might correspond with some situation with directionally diffusing side walls! Removing second and higher order laterals reduces the proportion of lateral sound, particularly towards the back of the hall, where typically the reduction is of the order of one subjective unit for both the mid-frequency and 200 Hz ratios of lateral to non-lateral sound. The mean values in Table 21.3 are also reduced by almost this amount. For the narrow hall of 14m width, the effect of removal of second and higher order laterals is even more marked: the mean mid-frequency ratio drops from -2.4 to -5.4 dB; whilst for the 32 m wide hall the mean is unchanged. Thus for narrow halls the presence of second order laterals is particularly important.

Finally the effect of the removal of cornice reflections was investigated. Both the effect of removing the pure cornice reflections (reflected off one side wall and the ceiling) and of removing all lateral ceiling reflections were calculated. In both cases the effect on the mid-frequency ratio is small, but as one would expect the 200 Hz ratio is considerably reduced as is the 200 Hz band level. It is evident from this that cornice reflections are particularly critical for bass spatial impression or envelopment, a not surprising result. Behaviour with halls of different widths was very similar.

21.6 THE EARLY SOUND LEVEL IN RECTANGULAR HALLS

This second determinant of the degree of spatial impression was investigated, the early energy being represented as a multiple of the direct sound at 10m. The absolute value is thus of little concern, but rather the difference between different halls is significant. The relevant difference limen for this quantity is 4 dB. Early energy was summed up to 80 ms.

Figure 21.7 shows the variation of this quantity in the average and wide halls. The perceived lack of intensity variation within halls is reflected in the small variation in this quantity. The mean and standard deviation for the early sound level in the 11 different shaped halls, considered in section 21.3, were calculated. Again receiver positions near the source were ignored. Theory suggests a relationship between early energy and hall volume, and if the computed mean early sound level is plotted against volume, an almost linear relationship is found to exist, as can be seen in Figure 21.8. However, the variation is barely more than one subjective unit between the largest and smallest halls. Given that variation in hall height is small in reality, and that variation is principally in hall width, this factor as a determinant of degree of spatial impression, complements the variation with hall width of the mid-frequency ratio of lateral to non-lateral, such that a narrow hall is superior to a wide one by as much as two subjective units of spatial impression.



(a) (b)

Figure 21.7. Computed values of the early energy level in (a) the average hall and (b) the wide hall. Source position is on the axis of symmetry.

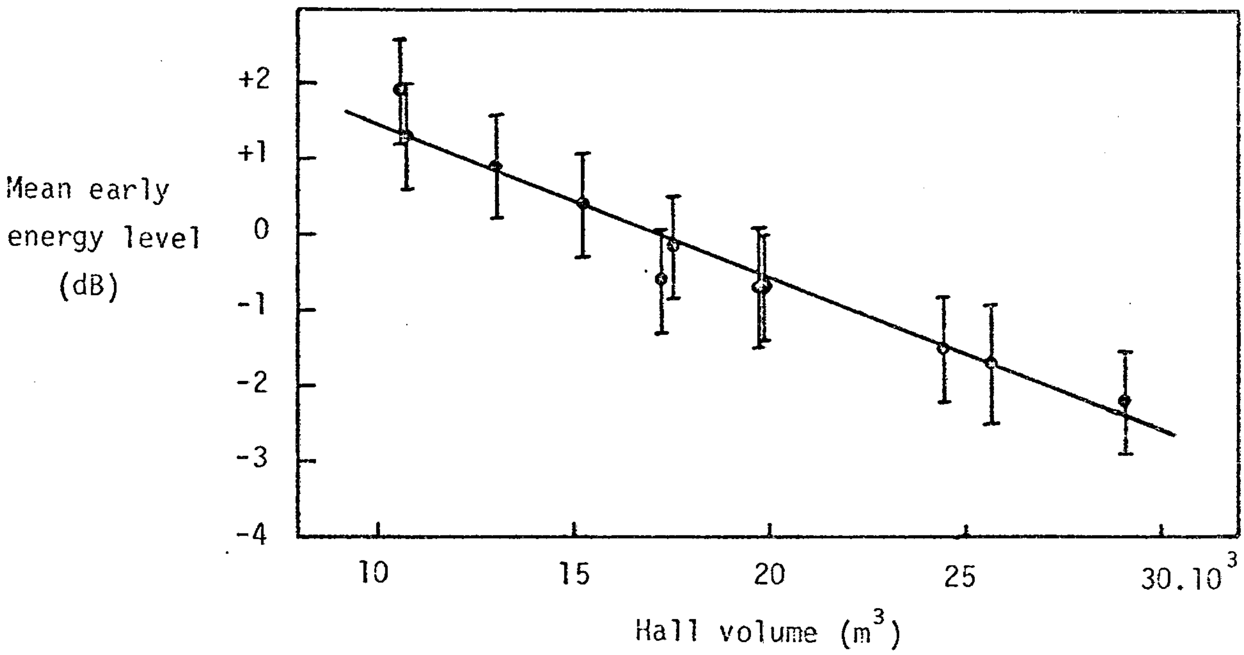


Figure 21.8. Computed values of the early energy level in 11 rectangular halls as a function of hall volume. ⌶, Mean and standard deviation.

To investigate the effect of removal of a group of reflections, the early sound level was also calculated for three hall widths with no front or back wall reflections. This relatively severe reduction in the number of reflections only produced a reduction of 1/3 of a subjective unit for spatial impression in all three halls.

21.7 CONCLUSIONS

Investigation of the mid-frequency ratio of lateral to non-lateral early sound in halls showed that the variation for central source positions was small, especially in the rear half of the hall. For a lateral source position, seats on the opposite side to the source are predicted as being preferable, by as much as two subjective units in a 32m wide hall relative to seats on the same side as the source.

For mid-frequency spatial impression, hall width variation causes a change of two subjective units, whilst hall height variation has little effect. Reflection off the front and back walls have little effect on the degree of spatial impression. Obviously the presence of lateral reflections is critical; however, particularly in narrow halls, the additional presence of second and higher order lateral reflections is essential for a high degree of spatial impression. Neither source position nor the elevation of rear seating areas influenced mean values. -2.7 dB was a typical mean value for the ratio of lateral to non-lateral early sound.

The situation at the frequency of maximum audience attenuation at grazing incidence was little affected by hall variations with a few exceptions. The 200 Hz ratio of lateral to non-lateral early sound was found to be on average -1.3 dB below mid-frequency value, whilst the total 200 Hz band early level was -4.2 dB below the mid-frequency level. The latter quantity was severely reduced at the back of low ceilinged halls and the front of high ceilinged halls. The 200 Hz ratio was, as anticipated, heavily dependent on the presence of cornice reflections. Cornice reflections can be considered critical for the sense of envelopment that is produced by bass lateral sound.

Chapter 22

THE DEGREE OF SPATIAL IMPRESSION IN FOUR NON-RECTANGULAR HALLS ACCORDING TO COMPUTER PREDICTIONS

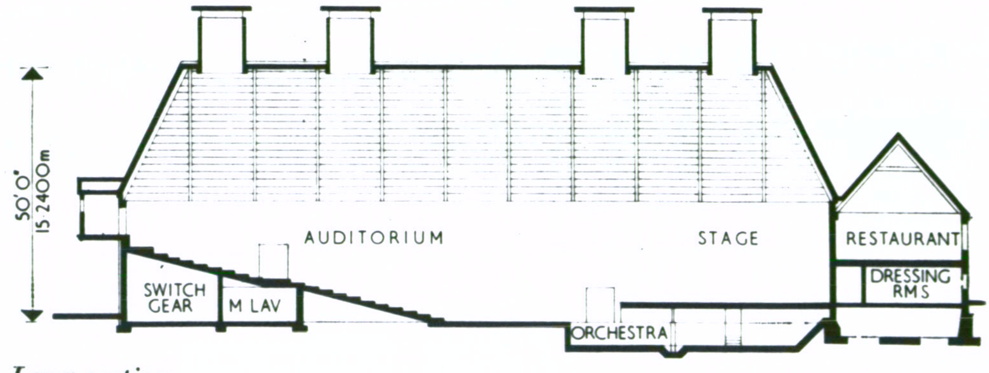
22.1 INTRODUCTION

This chapter contains an account of an investigation of four variations from the rectangular shape concert hall. Each of these shapes was considered as being likely to be significant in terms of mid- or bass-frequency spatial impression. Both changes in hall cross-section and hall plan were investigated, as well as the special case of the Royal Festival Hall, for which the basic procedure employed in the previous chapter was used. For each study the behaviour of the three quantities was again investigated: the mid-frequency and 200 Hz ratio of lateral to non-lateral early sound and the 200 Hz band early sound level. The procedure for dealing with non-rectangular hall shapes was outlined in section 20.3.

22.2 THE MALTINGS, SNAPE, SUFFOLK

This hall is relatively small by concert hall standards; the maximum audience is 824 persons. Figure 22.1 shows a long section and a photograph taken along the hall. The unusual aspect of this hall is its cross-section, which is shown in Figure 22.2 below; instead of a flat roof, the roof is gabled. The purpose of this study was to see whether this shape of roof is to be preferred for spatial impression relative to a conventional flat roof.

Ever since its opening in 1967 this hall has had a good reputation. One reason for this may well be its small volume, about 8,200 m^3, which would result in exciting climaxes, coupled with a close to ideal reverberation time of 2.0 secs at mid-frequencies, increasing to 2.9 secs at 125 Hz, |3|. Being narrow (18 m) the situation for the performers is also good. Detailed comment on the acoustics in the above reference is limited to surprise at the "distinctness" of the orchestra even when playing underneath the stage for opera, and the good balance and blend in this hall. The relatively long R.T. for such a small hall does not appear to produce excessive reverberance.



Long section

Figure 22.1. The Maltings, Snape, Suffolk

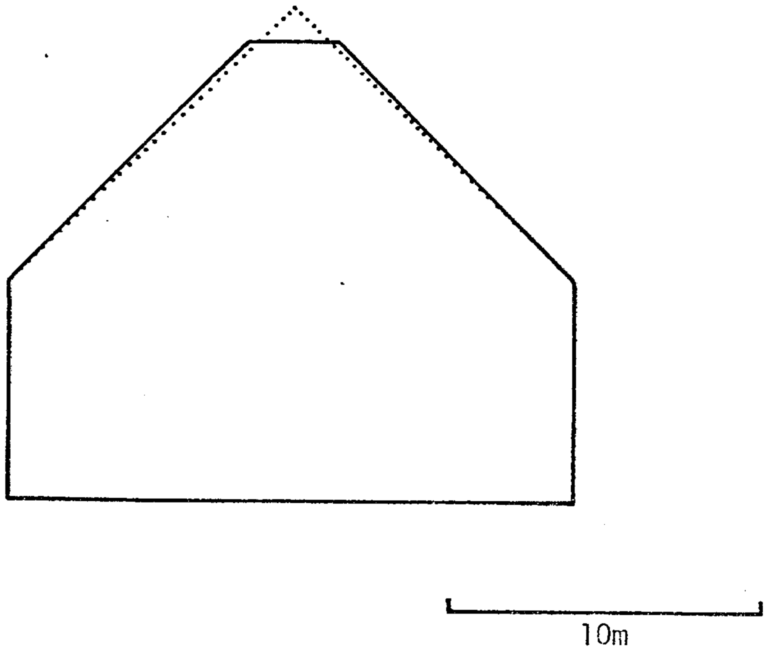


Figure 22.2. Cross-section of the "Maltings", Snape. Dotted line shows form assumed for computer study.

For the computer study, the small horizontal section at the centre of the roof was omitted and a 45^{o} inclination for the roof sections was used. This form is shown dotted in Figure 22.2. The sloping sections in the roof at each end of the hall were considered as vertical extensions of the front and back walls. The seating rake was included in the study. For comparison a flat-ceilinged hall was used, identical in all respects except for the flat-roof at a height half way up the gabled section, such that the respective volumes were identical (vertical side walls extended up to the height of the flat roof).

Calculated reflection sequences for the Maltings showed a similar number of reflections to that of the flat-ceilinged version throughout the hall. It remains to be seen whether the proportion of lateral sound is the same in both halls. The three computed quantities are listed in Table 22.1, with the relevant standard deviations; both central and lateral source positions (5m off axis) were investigated.

At mid-frequencies differences between the two halls are quite minimal. In detail the Maltings is slightly superior towards the front, whilst the flat-ceilinged equivalent is slightly superior towards the back. Curiously the 200 Hz ratio of lateral to non-lateral is not higher with the gabled

TABLE 22.1

Computed spatial impression in "Maltings" and flat-ceilinged equivalent

	Mid-frequency ratio	200 Hz ratio	200 Hz band level
Maltings:			
Central source	-2.3 ± 1.0	-3.9 ± 1.4	-4.8 ± 0.5
Lateral source	-2.2 ± 1.0	-3.1 ± 1.2	-4.6 ± 0.6
Flat-ceilinged hall:			
Central source	-2.1 ± 0.6	-3.1 ± 1.0	-6.0 ± 2.3
Lateral source	-2.0 ± 0.9	-2.4 ± 1.0	-6.1 ± 2.3

roof, since with this roof the angled roof surfaces produce lateral reflections on paths remote from the audience. Examination of the reflection sequences shows that whilst reflections off the gabled roof frequently occurred earlier, cornice reflections produced the equivalent effect. Further, not all seats received reflections off both angled roof sections, whereas all seats receive two cornice reflections in a rectangular hall. Also, since the two angled roof sections are at right angles, a double reflection off each surface is equivalent to a ceiling reflection.

The most significant difference occurs for the 200 Hz band level which is superior in the Maltings; particularly low values occur in the flat-ceilinged hall towards the rear due to the low ceiling, exaggerated further by the seating rake, which causes ceiling reflections to be of insufficient elevation to evade audience attenuation. Evidently reflections from the gabled roof have higher angles of elevation. The behaviour in the flat ceilinged hall is similar to that for the 12m high hall as shown in Figures 21.6 and 7.

In conclusion, the proportion of lateral sound at mid- and bass frequencies with a gabled roof is very similar to that of an equal volume, flat ceilinged, equivalent hall. In the particular case studied the low flat ceiling caused a lack of bass early sound, but using West and Sessler's data |16| rather than that of Schultz and Watters |14|, namely that a 15^{o} elevation (rather than 30^{o} assumed here) is adequate for audience attenuation to be eliminated, would significantly reduce the difference. Whilst in other respects no benefit is to be gained from replacing a flat ceiling by a gabled one, the addition of suspended angled reflecting surfaces within the volume of a wide hall may well increase the proportion of lateral bass sound.

22.3 THE ROYAL FESTIVAL HALL, LONDON

As can be seen in Figure 22.3, the plan of the Royal Festival Hall is based on a rectangle, whilst over the majority of its length the ceiling is basically horizontal. Apart from frequent use of diffusing surfaces in this hall, the orchestral reflector and the presence of boxes covering a substantial area of the side walls constitute a significant deviation from the rectangular halls studied in the previous chapter. The purpose of this study was to determine whether these two aspects would give a predicted change in spatial impression.

The acoustics of the Royal Festival Hall are reputed for their high degree of clarity, a clarity labelled by some as 'clinical' whilst by others as 'exciting'. There must certainly be many regular listeners in this hall who regard the ability to listen to individual instruments, as one can in this hall, as being synonymous with good acoustics. Initial criticisms of lack of fullness of tone, of "dry" acoustics have virtually ceased since the installation of the "Assisted Resonance" system |4|. There is, however, in the author's view a lack of involvement of the listener with the music being played on the stage, a sense of remoteness which might well be attributed to lack of lateral sound. This is not predicted for a pure rectangular hall of 32m width.

For the computer study many details present in the real hall were ignored: the orchestral enclosure, the front stalls side walls, and the side sections of the front wall. Again all surfaces were considered as providing specular reflections, whilst in reality the ceiling form is far from flat. The orchestral reflector is assumed to be a single plane surface. The rectangular area of the side walls containing the boxes (including box fronts, etc) was assumed to be fully absorbant. Only cornice reflections strike this area, and very few of these reflections travelling on a path inclined to horizontal will be reflected by the vertical wall surface adjacent to the boxes and no other. The box fronts would provide diffuse reflection, of which one can assume only a small proportion would reach the audience area. Thus the assumption of complete absorption for this area is probably realistic. The seating area from the middle of the

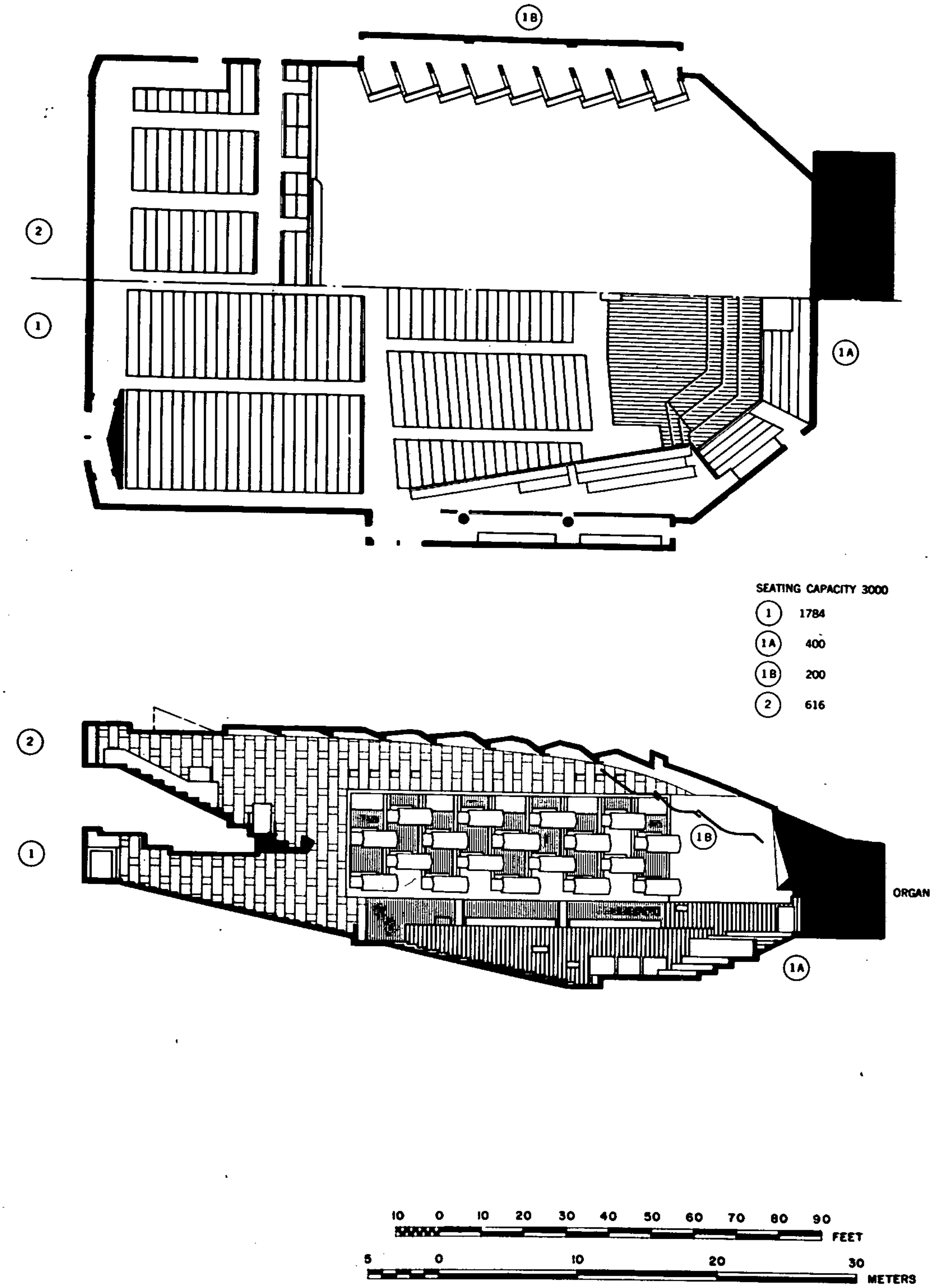


Figure 22.3. Royal Festival Hall, London.

lower stalls to the middle of the terrace stalls, from 21.5m to 39.5m from the front, was investigated. Seats behind this area are excessively screened by the grand tier to be suitable for comparison.

The computed values for the three quantities are listed in Table 22.2 for three situations: (a) the basic rectangular hall, (b) with the addition of the orchestral reflector, (c) with reflector and "boxes" on side walls. The results for situation (a) are very similar to those for a 32m wide hall recorded in Section 21.3 as one would expect. The addition of the reflector causes changes in the expected directions but only for the 200 Hz ratio is it approximately one subjective unit of change on average. The introduction of the "boxes" as absorbent surface on the side walls, whilst

TABLE 22.2

Computed spatial impression in three "versions" of the Royal Festival Hall

	Mid-frequency ratio	200 Hz ratio	200 Hz band level
(a) Basic rectangular hall	−2.9 ± 1.3	−3.5 ± 1.3	−4.3 ± 0.8
(b) Basic + orchestral reflector	−3.6 ± 1.3	−4.9 ± 1.3	−3.4 ± 0.6
(c) Basic + orchestral reflector and "boxes"			
Central source:	−4.2 ± 1.0	−6.9 ± 2.0	−3.8 ± 0.6
Lateral source:	−3.6 ± 1.3	−6.3 ± 2.1	−5.0 ± 0.9

having only a small effect on the mid-frequency ratio and the bass frequency level, severely reduces the 200 Hz ratio. This is the direct result of on average 1.1 cornice reflections being absent for the central source position, and 1.4 for the lateral source position (the 200 Hz ratio is nevertheless higher with the lateral source since cornice reflections are absent where the ratio is normally highest, on the side opposite the source). At an average seat, only one pair of cornice reflections will arrive from three sources on the stage, left-hand side, centre and right-hand side, out of a possible six cornice reflections. For bass instruments located on the right side of the stage, seats on the same side as the source are to be preferred. This presence or absence of cornice reflections causes the spread of the 200 Hz ratio to be larger for situation (c).

To anticipate arguments contained in section 23.3: according to the model used, only 40-50% of seats in the seating area studied receive unmasked bass lateral sound. This would explain the perceived lack of subjective involvement in this hall. Ross (on p.31 in |101|) also comments about "the (frequency) range which is missing in the R.F.H. which is in the middle of the musical scale up to 500 Hz, from about 150-500 Hz". Measurements of the proportion of early lateral sound in the Royal Festival Hall would be of great interest for the study of spatial impression.

22.4 FAN-SHAPED HALLS

The fan-shape has become a popular form for a hall due to its economic virtue: that it contains the largest number of people for a given maximum source-receiver distance for a given radiating angle for the source. Few useful reports are available about such halls: the Aula Magna at Caracas, Venezuela, which has an apex angle of 77^{o}, is quoted as having a high degree of clarity, with brilliant string tone. This may be due to the extensive suspended reflectors above and beyond the orchestral level. The Alberta Jubilee Auditorium at Edmonton and Calgary, Canada (apex angle 39^{o}) has been criticised for sound that "was not vital, not live enough .... low frequencies were seriously lacking, definition and clarity good to excellent". Both the above comments are taken from Beranek's book |41|. His own comment about the second hall is that "one can get pleasure out of a hall that does not take such an active part in the performance". The bass frequency R.T. for this hall is low, however, which may explain some of the comments. No fan-shaped hall, however, has gained a reputation for excellent acoustics.

Among shapes investigated by Krokstad et al. |83| in their computer study, it was found that a fan-shaped hall was deficient in energy in the 40-180 ms time period after the direct sound which is considered undesirable. The early energy level is also unsatisfactory for this shape; for the Aula Magna the computed value is -3.9 dB ± 1.5, the extremely low value occurring at the rear of the hall, whilst in an equivalent volume rectangular hall the spread is much less: viz. -3.2 dB ± 0.7. In terms of spatial impression Marshall |7| has pointed out the disadvantage that with a fan-shape the angle of azimuth of lateral reflections is small, though there is also the problem that few second order lateral reflections occur with such a hall shape. It was decided to investigate the degree of spatial impression in two fan-shaped halls, the first of dimensions equivalent to the Aula Magna and the second with an apex angle half that of the Aula Magna, but with greater length to give about the same volume, which will be called the Wedge-shaped Hall. Figure 22.4 contains a sketch of the plans of the two

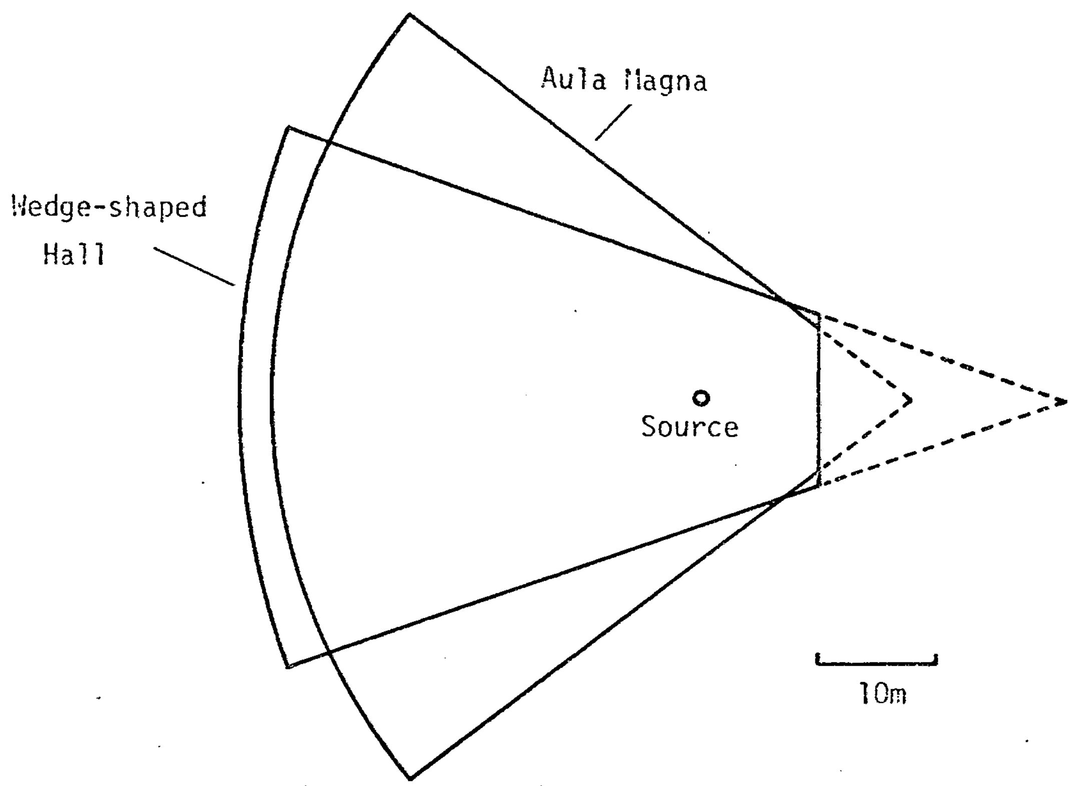


Figure 22.4. Plan of the Aula Magna and Wedge-shaped Hall.

halls. Both halls had the same height, 16m.

For the computer study the floor and ceiling were considered horizontal and the curved back wall was considered totally absorbent (a frequent feature to avoid embarrassing focussing). Reflections up to third order in plan were computed, though if these did occur they were very rare. The mean and deviations of the three quantities are listed in Table 22.3 for the two fan shapes and the rectangular hall of dimensions 45 x 38 x 17m, which has a similar volume, about $30{,}000 m^3$.

The extremely low values for the ratios of lateral to non-lateral sound are evident here, the situation being worse for the wider splay. The equivalent volume rectangular hall also has areas of poor lateral sound; though all these halls are large by concert hall standards, significant improvement could be made on the rectangular shape for this volume.

Results for basically the same fan-shaped halls but with a shorter

TABLE 22.3

Computed spatial impression in fan-shaped and a reverse-splayed hall compared with rectangular halls of similar volume.

	Apex Angle	Mid-frequency ratio (dB)	200 Hz ratio (dB)	200 Hz band level (dB)
Aula Magna	77°	-6.3 ± 1.2	-6.6 ± 1.4	-4.4 ± 0.6
Wedge-shaped hall	38.5°	-4.9 ± 0.8	-5.8 ± 1.4	-4.5 ± 0.7
Equiv. volume rectangular hall (45 x 38 x 17)	0°	-3.9 ± 2.0	-5.4 ± 3.0	-4.1 ± 0.7
"Aula Parva"	77°	-6.0 ± 1.4	-6.3 ± 1.6	-4.6 ± 0.7
Short wedge-shaped hall	38.5°	-4.5 ± 0.8	-5.1 ± 0.5	-4.3 ± 0.7
Equiv. volume rectangular hall (45 x 26 x 17)	0°	-2.9 ± 0.8	-4.0 ± 0.8	-4.1 ± 0.7
Reverse-splay Hall	(-40°)	-1.5 ± 1.2	-3.1 ± 1.6	-4.2 ± 1.2

front to back wall distance, to give a hall volume in each case of about 20,000 m^3, are also included in Table 22.3. The halls are called 'Aula Parva' and Short Wedge-shaped Hall, and can be compared with the rectangular hall of equivalent volume: 45 x 26 x 17 m. The relative inferiority of the fan-shape at this volume is even more marked relative to the rectangular shape. Again the total early energy level is lower with the fan-shaped halls, mean values being 1 dB lower than for the rectangular hall, but values at the back of the fan-shaped hall are again very low, typically -4 dB.

22.5 A REVERSE-SPLAYED HALL

If the vices of the fan-shape are converted into virtues by reversing the splay of the walls one achieves the expected high degree of S.I. The plan shape that was tested is shown in Figure 22.5. The chosen width at the source end was 32m, whilst the length and height were 45m and 17m. The side walls were each angled for two-thirds of the length of the hall at 20°. The volume of this hall is again about 20,000 m^3. The virtues of this shape are as follows: the angled walls increase the angle of azimuth of lateral sound as well as reducing its travel time (relative to rectangular); second order

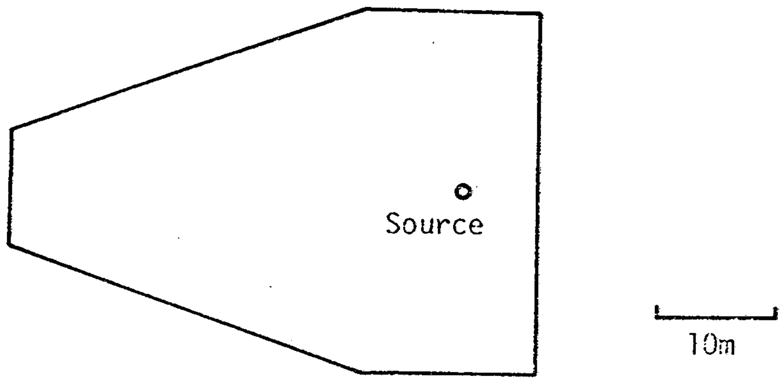


Figure 22.5. Plan of the reverse-splayed hall.

lateral reflections are also affected in the same way, and certain seats receive sound from both side wall surfaces.

The computed values listed in the last line of Table 22.3 show the superiority of this shape relative to the equivalent volume rectangular hall 45 x 26 x 17m. The improvement in the ratio of lateral to non-lateral sound is of the order of one subjective unit; this makes the hall superior to a 14m wide hall of half the volume. Seats at the rear of the reverse-splayed hall show particularly high degrees of lateral sound, and relatively high levels of bass early sound.

In architectural terms the square orchestral end is probably unsatisfactory; a shape corresponding to an elongated regular hexagon is likely to be preferred. This is unlikely, however, to be inferior to the shape tested for lateral sound. This elongated hexagonal plan form has also been espoused by Gabler |102| as the 'ideal' room shape for speech. The aim in Gabler's study was to provide a large number of early reflections for a large audience for speech. It is interesting to note that it also has virtues in terms of early lateral sound. A further possibility also exists, which evades the problem that a reverse-splay reduces the audience area: to use a concertina form side wall with useable side wall sections for reflections angled to give a large angle of azimuth for lateral reflections.

Chapter 23

A SYNTHESIS OF MEASURED AND COMPUTED RESULTS

23.1 THE EFFECT OF THE AUDIENCE ATTENUATION DIP ON BASS EARLY SOUND

Few references appear in the literature of measures of the frequency response of early sound in halls. Schultz and Watters |14| include the frequency response for sound arriving within 50 ms of the direct sound for a seat in the centre of the main floor of Boston Symphony Hall. The response was -7 dB on average in the 200 Hz octave compared with mid-frequencies. Schroeder et al. |39| also measured this same quality in the Philharmonic Hall, N.Y.; after the final alterations to the hall the response was essentially flat for the average over five main-floor positions, though with strong reflections from the suspended reflectors this is not surprising.

In section 19.5 a method was described of interpreting measurements of the early energy fraction to obtain an estimate of the early energy content at 200 Hz. (Incidentally measurements of both quantities in the Philharmonic Hall |39| are consistent with this procedure.) For the measurements in the stalls in Perth Concert Hall the estimated energy in the 200 Hz octave band was -6 dB for within 50 ms and -3 dB for within 80 ms of the direct sound in each case relative to mid-frequency. In Christchurch Town Hall the attenuation was always less than this, no doubt due to the use of reflecting surfaces to provide reflections on paths remote from the audience. A measurement by Beranek and Schultz of the ratio of reverberant to early sound energy in La Grande Salle, Montreal |31| suggests a 5 dB deficiency in the 200 Hz band, though without details of the reverberation time, it is necessary to assume a flat R.T. response here.

Further, computed levels of the 200 Hz band for the first 80 ms in rectangular halls were found to be -4.1 ± 0.7 dB independent of width, but decreasing significantly for low-ceiling or high-ceiling halls. The Perth Concert Hall result is in good agreement with this.

On the basis of this evidence it can be concluded that a paucity of bass early energy occurs in rectangular halls in the stalls, being about -6 dB or -3 dB within the first 50 ms or 80 ms relative to mid-frequency early energy. A virtually flat response occurs at the front of balconies, and less attenuation occurs when sound is specifically directed at the audience on paths remote from the audience seating. In qualitative terms this is as expected, but one interesting feature is that within the first 100 ms most of the bass deficiency appears to be compensated in the stalls of rectangular halls.

Schultz and Watters |14| suggest that "perhaps the absence of energy in any particular frequency band in the early sound can be compensated if sufficient energy is present in this band in the reverberant field". They cite the example of two seats, one at the front of and the other below a balcony for which "there is found to be little difference in the quality of sound heard from these two positions", although the latter is significantly lacking in bass early sound.

Sessler and West |15|, citing the same example, take a more pessimistic view: "In addition, the feeling of envelopment, which depends on pronounced lateral reflections, is also diminished (on the main floor). Both effects may contribute to the inferior acoustic quality experienced on the main floor of many concert halls as compared to balcony-seating areas." The conclusions above suggest a compromise solution. At least for simulations the character of reverberant and early sound are distinctly different even for low frequencies. However, the ability to discriminate delay is less at these frequencies |24|. A deficiency of about 3 dB for sound in the first 100 ms is on the border line of perception, and may well be compensated by positioning of bass instruments on the stage or by the number of such instruments.

The implication of this for concert hall design is that rather than increasing the bass frequency reverberation time, <u>a significant proportion of reflections should arrive within about 100 ms of the direct sound on paths remote from audience seating</u>. Such reflections will generally exist in rectangular halls except for two significant situations: at the front of halls with high ceilings (> 20m above the floor) or at the back of low ceiling halls (with ceiling height < 14m). In particular deficiency of bass sound is likely to occur for seating below balconies at positions far behind the balcony front. Audience attenuation at grazing incidence will not occur with a sufficient seating rake; West and Sessler |16| consider a 15^{o} rake adequate, whilst Schultz and Watters |14| quote the figure of 30^{o}.

23.2 THE MID-FREQUENCY DEGREE OF SPATIAL IMPRESSION IN HALLS

From the experiments described in Part I it was found that, for a high proportion of subjectively relevant lateral sound, reflections had to arrive later than 5 ms after the direct sound, and from directions not too close to that of the direct sound. It is not immediately obvious what implications this has even for rectangular hall shapes. Marshall |6| and West |40| have both suggested that hall cross-section is significant in this respect.

By making the assumption of specular reflection, it was possible to predict the relevant physical quantity in halls. It was found that the ratio of lateral to non-lateral early sound at mid-frequencies was relatively constant in a rectangular hall, though less so in wide halls, in which, for a central source, lower values occurred for central receiver positions and, for a lateral source, the ratio was lowest for receiving positions on the same side as the source. Maximum variation within wide halls (> 30m) corresponded to about 2 to 3 subjective units across the hall.

On average it was found that the degree of spatial impression at mid-frequencies was unaffected by hall height, but varied with hall width. If the total early sound level is also included as a factor influencing the subjective degree of S.I., the mean variation in degree between narrow (14m) and wide (38m) halls is about two subjective units, though the spread is larger for the wider halls. In no case did a situation arise at any seat where the degree of S.I. at mid-frequencies was predicted as being below threshold.

For halls which are not rectangular in plan, the suggestion of Marshall |7| that fan-shaped halls would be inferior whilst a reverse-splay hall would be superior to an equivalent volume rectangular hall was found to be realistic. For the variation of cross-section that was investigated, from rectangular to a gabled roof section, there was a quite minimal change in the ratio of lateral to non-lateral sound.

The assumption of specular reflection appeared from measurements in two rectangular halls (in section 18.3) to lead to realistic results. A more serious criticism of the computer investigation is the omission of any stage enclosure which may account for the fact that measurements in real halls were on average less than predicted, though the discrepancy is only of the order of 2 dB. In the two rectangular halls the mean mid-frequency lateral fraction was measured as about -6 dB, which corresponds to a ratio of lateral to non-lateral sound of about -4 dB. The small variation with position was in agreement with prediction. In Christchurch Town Hall the mean mid-frequency lateral fraction was -9 dB, which corresponds to a ratio of lateral to non-lateral of about -8 dB. This relatively low value was attributed to the small angle of azimuth of most lateral reflections.

However if the results of Chapter 12 are to be accepted, the low-frequency ratio of lateral to non-lateral sound is even more significant than the mid-frequency ratio, and this leads to different priorities for hall design in certain cases.

23.3 SPATIAL IMPRESSION AT BASS FREQUENCIES

The predicted ratio of lateral to non-lateral early sound for the 200 Hz octave band was found to be about 1.5 dB less than the mid-frequency ratio irrespective of hall width and height: i.e., about one subjective unit. For variation of hall width the 200 Hz band early energy was predicted as being about 4 dB below the mid-frequency level, again about one subjective unit. Measurements in halls confirmed that these figures were realistic, or if anything, that in the real situation the difference between the 200 Hz band and mid-frequency behaviour was less significant than predicted.

However, in certain situations the predicted behaviour at low frequencies was significantly inferior. It is tempting when viewing subjective spatial impression in terms of difference limens to ignore the possibility that masking may indeed take place. To view the situation in terms of a masking threshold also has the advantage that one is dealing with a single criterion rather than having to discuss the behaviour of two physical quantities.

It was suggested in section 11.3 that bass lateral sound is likely to be masked by the total early mid-frequency sound. A threshold value of -17 dB was speculated. It is usual to consider that the effective threshold for listening in the real situation may be higher than this, say by as much as 5 dB. This effective threshold of -12 dB may be applied to the computed results. The ratio of bass lateral sound to total early mid-frequency sound can be readily calculated from the 200 Hz ratio of lateral to non-lateral sound and the 200 Hz band early energy relative to mid-frequency. In certain of the hall shapes investigated a significant proportion of seats exist for which the bass lateral sound is below threshold.

This approach raises the question of what criteria does one apply for design of a concert hall. Does one aim at a good average behaviour, and consider that certain poor seats can be compensated by certain very good seats? This would seem to be an erroneous approach, since such good and bad seats may exist in the same row, and "patchy" acoustics are not popular. The approach that all seats should comply with certain minimum standards seems more realistic; that bass lateral sound be above threshold would appear to be one such minimum standard.

The predicted percentage of seats in rectangular halls with masked bass lateral sound is given in Table 23.1. The significance of values for the 200 Hz ratio and band levels is immediately obvious here (see Figures 21.4-7). For a very wide hall (38m) masking occurs towards the front and

TABLE 23.1

The percentage of seats in rectangular halls with masked bass lateral sound.

Hall width (m) L = 45m; H = 17m	% of seats below threshold	Hall height (m) L = 45; W = 20	% of seats below threshold
14	0	12	26
20	0	14.5	0
26	6	17	0
32	7	19.5	18
32 (lateral source)	2	22	48
38	20		
Hall dimensions (m):			
45 x 32 x 12	42		
45 x 26 x 22	44		

along the axis of symmetry. For a low ceiling hall (12m) masking occurs at the back, whilst with a high ceiling masking occurs over the whole front half. As expected, the absence of any lateral reflections off the ceiling is devastating; for the hall (45 x 20 x 17m) 100% masking occurs in this situation, whilst for a hall (45 x 32 x 17) with no such reflections the figure is 85%.

For rectangular halls a -12 dB threshold leads to the following conclusions:

(a) width should be less than 35m

(b) height should be between 13 and 20m

(c) cornice reflections should not be suppressed.

In a hall with balconies, balcony fronts can be designed to provide lateral reflections containing bass sound, but nevertheless avoiding dead spots would take careful design.

The masking situation for the halls discussed in Chapter 22 is given in Table 23.2. The poor performance of the large fan-shaped halls is again apparent, as is the situation in the modelled Royal Festival Hall. The figure for the reverse-splay hall is probably over-pessimistic; masking

TABLE 23.2

The percentage of seats with masked lateral sound in various non-rectangular halls

	% of seats below threshold
Aula Magna	33
"Aula Parva"	37
Wedged shaped hall	21
Short wedge-shaped hall	5
Reverse-splay hall	11
Royal Festival Hall including "boxes":	
Central source	50
Lateral source	60

occurred for positions close to the orchestra (since here the hall is widest). The proximity of the orchestra makes spatial impression less important here, and a stage enclosure would also greatly improve the situation.

For halls in which bass lateral sound is above threshold, the conclusions for mid-frequency spatial impression also apply: that narrow halls are superior. In the three halls used for measurement, no below-threshold situations occurred.

23.4 CONCLUSIONS

For adequate bass early sound it was concluded that a significant proportion of reflections should arrive within about 100 ms of the direct sound on paths remote from audience seating. This may be achieved by an adequate seating rake (15^{o}-30^{o}), but in rectangular halls with a flat floor in general a 3 dB deficiency at bass frequencies is probably not significant. Both low ceiling (< 14m) and high ceiling (> 20m) halls would have areas of insufficient bass early sound.

For spatial impression a minimum requirement of above threshold bass lateral sound was proposed. This is in agreement with Marshall's suggestion |7| that significant lateral reflections should be present arriving on paths remote from audience seating. The threshold value used was on the basis of speculation, but the qualitative implications were that rectangular halls should not be too wide, and again as above neither low ceilings or high ceilings are desirable. In particular, cornice reflections should be preserved.

For above threshold situations, the results pointed to a conclusion slightly at variance with that proposed by West |40| and Marshall |6|, who suggested hall cross-section ratio as being significant. Within the limits for ceiling height mentioned above, ceiling height was not found to be significant, whereas small hall widths were found to be superior. Since for 19 actual rectangular halls (derived mainly from Beranek |41|) the standard deviation of hall height is only 3m, the correlation between cross-section ratio and acoustic quality found by West |40| is understandable. For simple variations in hall shape, the fan-shaped hall was found to be significantly inferior whilst the reverse-splayed hall was found to be superior to the rectangular form.

Chapter 24

SPATIAL IMPRESSION AND CONCERT HALL ACOUSTICS

24.1 INTRODUCTION

In this penultimate chapter the opportunity will be taken to discuss (sometimes in less quantitative terms) the more general concert hall situation in the light of the work reported in this thesis. Firstly an assessment of the validity of Marshall's and Beranek's theories will be discussed, then the question of diffuse versus discrete reflections. Finally there is a consideration of how significant spatial impression is for concert hall acoustics and a discussion of design criteria for spatial impression.

24.2 A REASSESSMENT OF MARSHALL'S THEORY

This theory has already been summarised in section 2.5. Marshall's theory was, of necessity, based on the published experimental results at the time. These were limited to threshold results, but in section 5.3 the measured threshold of a lateral reflection was found to be independent of its delay relative to direct sound and a ceiling reflection. This removes the importance of precedence of lateral reflections relative to the ceiling reflection, though since an earlier reflection is more intense, such a reflection will produce a larger degree of spatial impression. The conclusion of the precedence approach, that for rectangular halls the cross-section ratio is important, requires modification; this has been fully discussed in the previous chapter.

Marshall |6| also proposed that discrete reflections are to be preferred to diffuse ones, since he claims the latter are more likely to be masked. Experiments described in Chapter 10 and references in the literature conflict with this view however. The question of diffuse reflections will be discussed further below.

In the light of evidence concerning audience attenuation at grazing incidence, Marshall |7| also considered the possibility of bass frontal sound masking bass lateral sound. In section 11.3 it was suggested that bass lateral sound is probably masked by the total mid-frequency early sound, but that only the relative level and not delay are important. However, each approach leads to the same conclusion: that lateral reflections on paths remote from audience seating are highly significant for the creation of spatial impression. The importance of bass lateral sound was confirmed in Chapter 12.

The virtues and vices of a reverse-splay and a fan-shaped hall, respectively, as proposed in reference |7|, were also confirmed. However, the design exercise discussed in reference |103|, in which a subdivided audience is used, can be criticised for having many seats in which lateral reflections arrive too early or with insufficient angle of azimuth to create adequate spatial impression. Furthermore, cornice reflections would be difficult to preserve in this form of design. The use of a reverse-splay shape appears from section 22.5 to be limited to dimensions larger than those proposed in reference |103| for the individual open "boxes". This is not to invalidate the concept of a subdivided audience, though provision of uniform acoustics for such a hall is a complex design exercise. Use of suspended reflecting surfaces, especially to provide lateral sound from above, as used in Christchurch Town Hall, probably offers a simpler solution.

It is convenient at this point to discuss the significance of early lateral reflections in terms of labels conventionally applied to subjective concert hall experience. Marshall |6| has suggested that the effect of early lateral reflections has frequently been ascribed to reverberation: in other words spatial impression contributes to the subjective scale of "reverberance". Later, Marshall |7| associates a sensation of "warmth" with spatial impression, though it may be questioned to what extent this is a function of total bass early sound, bass early lateral sound and bass reverberant sound. Two out of five subjects in section 12.4 considered source distance or "intimacy" to be affected by the presence of lateral reflections. It would thus seem that lateral reflections contribute to more than one subjective attribute, rather than occupying a subjective scale of their own, though with training it is relatively easy to detect the presence of early lateral reflections. This suggests that investigations which use scales of "clarity", "intimacy", etc., similar to those of Beranek |41|, are unlikely to reveal the significance of lateral reflections. The approach involving judgement of similarity and dissimilarity offers more promise in this respect.

24.3 BERANEK'S CONCEPT OF ACOUSTICAL "INTIMACY"

Beranek in his investigation of concert hall acoustics |41|, proposed that the initial time delay gap (the delay of the first reflection relative to the direct sound) is the most significant determinant of acoustical quality in a concert hall (i.e., more significant than reverberation time). He relates the initial-time-delay gap to the subjective quality of "intimacy"; for a hall to have the right degree of intimacy it must sound as if it were the appropriate size for the music being played there. For optimum intimacy, he claims that the initial-time-delay gap must be less than 20 ms.

For halls in which lateral reflections precede ceiling reflections, the criteria for spatial impression and "intimacy" are equivalent, though for spatial impression it is the level rather than the delay which is considered significant. In an experiment, reported previously |63|, it was found that the delay of a lateral reflection did not substantially influence the subjective effect; in particular the apparent room size was much more a function of lateral reflection level than delay. However, if the ceiling reflection has precedence, a delay of 20 ms would seem to be unsuitable on account of tone colouration and the image shift effect associated with such reflections.

There are no doubt many physical acoustic quantities which have a reasonable correlation with the initial-time-delay gap for existing halls. The ratio of lateral to non-lateral early sound may be one of them which also has the advantage that its subjective implications coincide to a certain extent with those claimed by Beranek for "intimacy".

24.4 THE CASE FOR DIFFUSE REFLECTIONS

The physical quantity derived in Part I as being related to the subjective degree of spatial impression contained only the level and direction of a reflection. Thus a diffuse reflection is equivalent to a discrete one if the total energy is the same. This was confirmed above threshold (Chapter 10) and at threshold (section 10.3) for spatial impression. Both Seraphim |10| and Kuhl |11| also found that a discrete and a diffuse reflection were both detected at threshold by their energy content.

It is frequently quoted that diffusing surfaces are desirable in concert halls (see, e.g., the book by Parkin and Humphreys |96|, p. 95) though many different reasons are given. They certainly provide more uniform conditions throughout the hall, and improve the degree of diffusion of the reverberant sound. Reports have been heard of the superiority of a diffuse lateral reflection over a discrete one, though nothing appears to have been published. Certainly in a simulation with just direct sound, early reflections and reverberation, it is difficult to appreciate in what way the diffuse lateral reflection might be superior, though without a more elaborate simulation system this is not sufficient evidence.

The only respect in which a diffuse lateral reflection might be preferred, according to results in this thesis, is in terms of localisation. With a discrete reflection a slight lateral image shift occurs, which has been discussed in section 13.8 (f); with a diffuse reflection this shift is less likely. It is of course only significant where the lateral sound is one-sided, which occurs near walls in a concert hall. Meyer and Schodder |37| performed equal loudness experiments with a single discrete reflection and a series of temporally diffuse reflections and found that for equal loudness the diffuse reflection energy had to be larger. This can also be interpreted as indicating the superiority of diffuse reflections if correct localisation is to be maintained.

The case for diffuse ceiling reflections is, however, much stronger, since a diffuse reflection is much less likely to cause vertical image shift, |50|, and will also produce considerably less severe tone colouration.

24.5 HOW IMPORTANT IS SPATIAL IMPRESSION IN HALLS?

It is probably fair to say that sufficient comments by eminent acousticians have been made (as listed in section 2.5) to establish that lateral reflections are significant in halls. The question remains of the degree of significance. This thesis has been concerned with determining what requirements lateral reflections should fulfil for spatial impression to occur. However, the designer requires to know the priority he should place on providing spatial impression. The necessary evidence is not very extensive.

The two halls with the best reputations for good acoustics in the world are the Boston Symphony Hall, U.S.A., and the Vienna Musikvereinsaal. There is little doubt, on the basis of investigations reported in previous chapters, that the degree of spatial impression in both these halls is uniformly high, due to their narrow rectangular shape. In the author's personal experience the sensation of spatial impression in the rear gallery of the Musikvereinsaal is immediately very obvious. The ratio of lateral to non-lateral sound has also been found above to be correlated roughly with two measures recently proposed as correlating with acoustical quality: the initial-time-delay gap and the cross-section ratio. However, apart from measurements reported in Chapter 19 which, in the circumstances cannot be considered to prove or even indicate a relation between degree of spatial impression and general acoustical quality, the only other reported measurements of the proportion of lateral sound were made in the Philharmonic Hall, N.Y. |39|: a higher proportion of lateral sound in the Second Terrace of this hall "may in part have been responsible for the difference in subjective quality when listening to music in these locations".

Spatial impression thus has the potential to be one of the most significant requirements for excellent acoustics, but what priority it takes has yet to be established. Perhaps aspects related to clarity are more significant, though frequently with a uniform exponential decay, including the initial decay, clarity is quite satisfactory. The provision of above threshold bass lateral sound offers a minimum requirement which is essentially soluble for most basic hall shapes. Regarding the desirable degree of spatial impression this becomes a matter of taste, though the Musikvereinsaal must be somewhere near the upper limit.

24.6 DESIGN CRITERIA FOR SPATIAL IMPRESSION IN HALLS

It is convenient to discuss first of all the requirements to satisfy the criterion that bass lateral sound is above threshold. For rectangular halls, these have already been extensively discussed in the previous chapter:

(a) that hall width is less than about 35 m

(b) that hall height is more than 13 m but less than 20 m

(c) that cornice reflections are not obscured.

For wide halls inclusion of an orchestral reflector to reflect sound onto the audience is likely to be detrimental to spatial impression, whilst a seating rake of more than 15^{o} removes the constraint of maximum hall height.

For small auditoria, the only relevant requirement is that the ceiling be sufficiently high for ceiling reflections at the back of the hall not to be affected by audience attenuation at grazing incidence. For large halls of non-rectangular shape the requirement for lateral reflections travelling on paths remote from the audience becomes more difficult to satisfy. It is probably realistic to consider just first order reflections, and "cornice" reflections for large halls (an investigation of eliminating all higher order reflections in rectangular halls showed errors to be small except for narrow halls, which in any case are known to satisfy the threshold criterion). The sum of the "cornice" reflection levels, when each is multiplied by $\sin\alpha . \cos\beta$ (α = angle of azimuth, β = angle of elevation) should be within 12 dB of the total early sound level. The latter can be approximately calculated from the inverse square law and Figure 21.9, which is a plot of early energy level against volume, assuming the shape approximates very roughly with rectangular. This procedure, for instance, should be suitable for assessing the situation in an elongated hexagonal shaped hall. For a fan-shape and reverse-splay shape, the results of sections 22.4, 22.5 and 23.3 can serve as a guide. In each case a single central source can be used, and about 5 seat positions investigated over the seating area, the front 8 m or so of audience seating being ignored. For more extreme shapes, a computer investigation on the lines of that described in Chapter 20 can be used.

Simple models are also readily employed: the lateral sound is detected by a figure-of-eight microphone; pulse responses can be employed. To examine overhead reflections, non-elevated sound from the source can be readily eliminated with a screen. To obtain a figure for the total early sound a gated energy meter, similar to that described in Chapter 16, is most suitable, though estimates are no doubt possible from oscillograms.

To estimate the degree of spatial impression for above threshold halls, the results reported in previous chapters can be used as a guide. For large halls, again considering just first order reflections and cornice reflections, and ignoring audience attenuation effects, is probably adequate, though a full computer investigation is obviously to be preferred. The computer investigation by Marshall |98| of Christchurch Town Hall indicated that diffuse reflections occur even from nominally plane surfaces so that results should be averaged over seating areas, whenever small reflecting surfaces, and in particular suspended surfaces, are employed. The measuring technique described in previous chapters can be readily used in large scale models with again a figure-of-eight microphone, though seating areas should be modelled to produce a suitable attenuation dip.

The problem of design of a large hall (seating 3000 people) with a relatively high level of spatial impression remains at present unresolved.

Many halls of this size can be criticised for placing sections of the audience deep under balconies or in areas unsuitably related to the orchestra (for example, behind the orchestra). Subdividing the audience or using extensive suspended reflecting surfaces to provide overhead lateral sound are possible solutions, though it remains a very challenging design exercise to provide uniformly excellent acoustics!

Chapter 25

CONCLUSION

It is now well established that subjective appreciation of concert hall acoustics is a multi-dimensional process, though at present there is no general agreement over physical correlates of this subjective experience. Reverberation time has long been considered a physical correlate of subjective impression, which can be called "reverberance". Recently two substantiated criticisms of this conventional view have been made. Firstly subjective experiments have shown that it is the initial decay rate, rather than the later decay rate, which correlates best with subjective impression. Secondly the suggestion has been made that the subjective effects of early lateral reflections also contribute to the sense of "reverberance". It was the primary aim of this project to determine the nature and behaviour of the subjective effects of early lateral reflections, called here "spatial impression", with the eventual aim of determining what implications this has for concert hall design. Subjective experiments were conducted in an anechoic environment, using synthesised direct sound, early reflections and reverberation. Once a physical correlate of spatial impression had been derived, a measuring system was built and measurements were conducted in three concert halls in Australasia. As an extension of these measurements, a computer investigation was conducted to determine the influence of hall shape and size on the degree of spatial impression.

In a simple subjective experiment, subjects were asked to describe the subjective effects of a single lateral reflection. Apart from loudness effects, the following effects were noted: image shift, tone colouration, spatial impression and echo disturbance. Of these, only spatial impression can be considered as a positive contribution to acoustical quality. Both tone colouration and echo disturbance have, at least at threshold, been the subject of elaborate study elsewhere. However, measurements of the threshold for spatial impression indicated that virtually all reflections in a concert hall are above threshold, so that the threshold situation has little relevance, except as will be mentioned below in particular situations.

Subjective experiments were therefore conducted with above threshold situations. A comparison technique was used to determine the variation of spatial impression (S.I.), with different reflection parameters. It was found

(a) that S.I. was substantially independent of delay for delays greater than about 8 ms;

(b) that the degree of S.I. was a function of the direction of the reflection relative to an axis through the listener's ears; maximum S.I. occurs for purely lateral reflections;

(c) that as a function of level, extrapolation of work by other authors indicated a relationship between reflection level and S.I., which was monotonic increasing; the degree of S.I. appears also to be a function of music level;

(d) that in multiple reflection situations the ratio of lateral to non-lateral early sound correlates with subjective spatial impression. This work led to a physical measure (expressed in equation (14.1)) as a unique correlate of spatial impression .

Direct and early sound in a concert hall suffer audience attenuation at grazing incidence, frequently being as much as 15 dB at 160 Hz and extending up to about 800 Hz. With a filter to simulate this attenuation effect, subjective experiments were conducted to determine the influence it has on spatial impression. The comparison technique, however, proved to be unsuitable, though it was apparent that bass lateral sound could be masked by other early sound. However, an experiment, in which subjects were required to mark on a scale of spatial impression, indicated that with reverberation also present the degree of spatial impression is severely reduced when lateral reflections are audience filtered. Further experiments with reverberation indicated a certain area of overlap between the subjective effects of reverberation and spatial impression.

It has been proposed elsewhere that the auditory mechanism responsible for spatial impression involves a short-term cross-correlation process. Since spatial impression is so strongly associated with bass frequencies, it can be asserted that it occurs at frequencies for which the ear relies on a cross-correlation process for localisation. A model was proposed that the degree of spatial impression is related to the height of the localisation maximum in the inter-aural cross-correlation function for frequencies below 750 Hz. This model explains nearly all observed characteristics associated with spatial impression, and suggests a simple measuring method for use in halls or models.

An investigation of the implications for sound decays in rooms according to geometric acoustic theory led to the not surprising conclusion that average behaviour cannot be predicted for an environment with diffusing surfaces. The use of geometric theory for prediction of the proportion of early lateral sound appeared, however, to be valid according to early energy measurements in halls. These measurements were performed with impulses, which were analysed by gating and measuring the integrated energy content. Measurements in halls indicated that the ratio of lateral to non-lateral early energy is relatively constant both with frequency and within halls. The effect of audience attenuation on the early energy was found to be more marked, though for longer time periods the early energy level was only about 3 dB down at the frequency of maximum attenuation.

The computer investigation of the ratio of lateral to non-lateral early sound indicated that at mid-frequencies only hall width and not hall height is significant. However, since spatial impression appears to be strongly associated with bass frequencies, it is necessary to consider the possibility that lateral bass sound is masked. This leads to the conclusion that

(a) the large majority of existing halls exhibit perceptible spatial impression over their entire seating area;

(b) rectangular halls wider than about 35m are not desirable;

(c) halls which are too low (<13m) or too high (>20m) also possess areas at which bass lateral sound is masked;

(d) cornice reflections should not be suppressed.

Fan-shaped halls were also found to be poor for spatial impression, whilst the use of a reverse-splay proved to be particularly good. There are, however, considerable problems involved in providing a relatively high level of spatial impression in a large hall.

Many of these conclusions may be verified in existing halls, though the problem of subjective assessment remains. Many references to the desirable effects of early lateral reflections have been made by eminent acousticians. There seems little doubt that early lateral reflections are subjectively significant. It is hoped that this work has determined the physical requirements for such reflections to be significant. The question, however, remains of how significant is the spatial impression produced by early lateral reflections. It is hoped that the tests of similarity and dissimilarity on artificial head recordings made in real halls (such as those being conducted at present in Göttingen) may resolve this question of significance for a listener's assessment of concert hall acoustics.

APPENDIX I

DIFFERENTIAL ATTENUATOR

The function of the differential attenuator is to vary the relative levels of two channels, whilst maintaining the sum of the two channel levels constant (assuming incoherent addition). Typically it was used to vary the level of the reflection (L_R) relative to the direct sound level (L_D) at constant total loudness. The circuit is given below; a multi-position switch was used, with specially selected resistors for the resistance

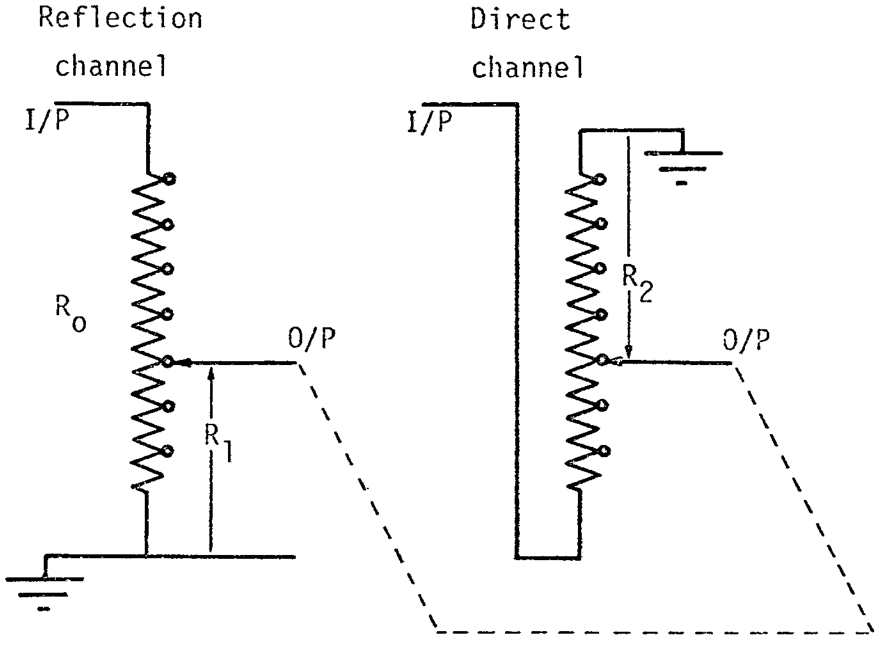


Figure I.1. Circuit of the differential attenuator.

chain.

If the ratio of reflected to direct sound (= L_R/L_D) is r, and the total resistance of each chain is R_o, then

$$R_1 = R_o\sqrt{\frac{r}{1+r}}\,; \qquad R_2 = R_o\sqrt{\frac{1}{1+r}}.$$

A total resistance of R_o =30KΩ was used; the output was fed to a high input impedance device (> 1 MΩ).

APPENDIX II

AUDIENCE ATTENUATION FILTER

The circuit of the audience attenuation filter is given below in Figure II.1. The frequency response of this filter is given in Figure 11.2. The integrated circuit amplifiers used were National LM 741C (or equivalent). The filter has unity gain and zero phase shift in the pass-band (> 2 kHz).

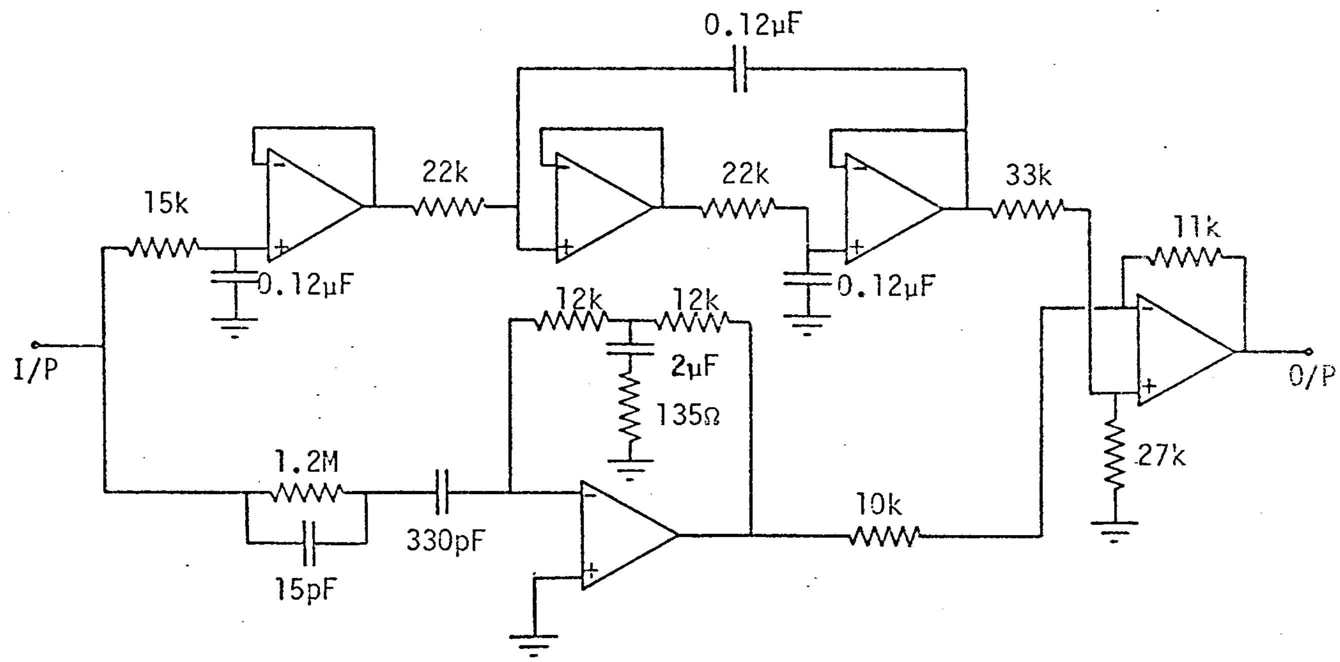


Figure II.1. Circuit of the audience attenuation filter.

APPENDIX III

COMPUTER PROGRAMMES FOR THE INVESTIGATION OF EARLY REFLECTIONS IN RECTANGULAR HALLS

The following computer programmes are written in FOCAL-11, suitable for use on a Digital Equipment Corp. PDP-11. The version of FOCAL-11 used in these programmes is a preliminary version, which differs from the official version in one respect: subroutines in this version are labelled as functions FNEW(,), whilst in the final version they are labelled FSBR(,). Subroutine 15 in these programmes computes the natural logarithm.

The computing procedure is outlined in sections 20.2 and 20.4. In Figure III.1 the basic programme is given, which contains the input instructions and calculates the relevant images, and the travel times for sound from these images. The basic programme may be used with either of the supplementary programmes: the reflection details programme and the spatial impression programme, which are given in Figures III.2 and 3 respectively.

In the reflection details programme, listing under T is the delay time (in ms) relative to the direct sound, DB is the reflection level relative to the direct sound, AZ and EL are the reflection azimuth and elevation, and U and V refer to the cell coordinates in which the particular image is contained, the prefix C before the cell coordinates indicating a ceiling reflection.

In the spatial impression programme, the output S(80) is the mid-frequency ratio of lateral to non-lateral early sound, SB(80) is the bass-frequency ratio of lateral to non-lateral early sound and "150 Hz level" is the bass frequency relative to mid-frequency early energy level.

```
1.10 A "HALL LENGTH"L," HALL WIDTH"W," HALL HEIGHT"H,!
1.20 A "SOURCE COORDINATES"D,DL,DH,!!
1.30 A "RECEIVER COORDINATES"A,B,C;S A=A-D;S C=C-DH
1.50 S TO=2.915*FSQT[A*A+(B-DL)↑2+C*C]
1.55 T "             T(0)"%4.01,TO,"MS" !!
1.80 S KC=0;S I=1;S M=0;S TT=200

2.10 D 3;S N=I-1;T !!,"     T              M          N              M/N %",!!
2.20 S TM=10000;F I=1,1,N;D 7
2.30 I (TT-TM)2.9;S M=M+1;S NM=(0.16903*T(K)↑3)/(L*W*H)
2.40 T %4.01,T(K),"     ",%3,M,"     ",%4.01,NM,"          "
2.42 T 100*M/NM-100,!
2.50 S T(K)=1000;G 2.2
2.90 Q

3.30 S Z=2*H*[FITR((21+KC)/2)-10]+2*DH*[2*FITR(10+KC/2)-20-KC]
3.40 S X=0;S Y=0;D 6.1;I (TT-T(I))3.6;D 4;S KC=KC+1;G 3.3
3.60 S KC=-1
3.70 D 3.3;S X=0;S Y=0;D 6.1;I (TT-T(I))3.9;D 4;S KC=KC-1;G 3.7
3.90 R

4.10 S KA=0
4.20 S X=2*L*[FITR((21+KA)/2)-10]+2*D*[2*FITR(10+KA/2)-20-KA]
4.30 S Y=0;D 6.1;I (TT-T(I))4.7;D 5;S KA=KA+1;G 4.2
4.70 S KA=-1
4.80 D 4.2;S Y=0;D 6.1;I (TT-T(I))4.9;D 5;S KA=KA-1;G 4.8
4.90 R

5.10 S KB=0
5.20 S DX=0;S Y=KB*W+DL*[1+4*FITR(KB/2)-2*KB];D 6
5.25 I (KB-0.5)5.3;S Y=Y-2*KB*W;D 6
5.30 I (3-DX)5.8;S KB=KB+1;G 5.2
5.80 R

6.10 S T(I)=2.915*FSQT[(X-A)*(X-A)+(Y-B)*(Y-B)+(Z-C)*(Z-C)]
6.20 I (TT-T(I))6.9;S I=I+1;R
6.90 S DX=DX+2;R

7.10 I (TM-T(I))7.2;S TM=T(I);S K=I;R
7.20 R
```

Figure III.1. Basic programme to compute early reflections in rectangular halls.

```
 6.10 T !!;D 30.1
 6.20 S TM=10000;F I=1,1,N;D 7;S I=-I;D 7;S I=-I
 6.40 S INT=4.343*FNEW(TO*TO/T(K)*T(K),15);D 11

11.30 S AN=(Y(FABS(K))-B)/(X(FABS(K))-A)
11.31 S AN=FABS[(AN-B/A)/(1+AN*B/A)];S AN=AN/FSQT(1+AN*AN)
11.50 I (K)11.52;S EL=FSQT[1-(2.915*C/T(K))↑2];G 11.6
11.52 S EL=FSQT[1-(2.915*(H-C)/T(K))↑2]
11.60 T %3.01,T(K)-TO,"     "INT,"     "%3.02,AN,"     "EL,"          "
11.61 D 32;T !

29.34 Q

30.10 T "    T              DB          AZ           EL                    U   V" !

32.12 I (RR-1.5)32.2;T "          ,"
32.20 I (K)32.3;T " ";G 32.5
32.30 T "C"
32.50 T %1.(X(FABS(K))-L/2+D)/L,Y(FABS(K))/W;R
```

Figure III.2. Supplementary computer programme for reflection details.

```
 6.10 C
 6.20 F I=1,1,N;D 11;S I=-I;D 11;S I=-I
 6.40 G 29.34

11.10 S K=I;S T=T(K)
11.20 I (T-TO-90)11.3;R
11.30 S AN=(Y(FABS(K))-B)/(X(FABS(K))-A)
11.31 S AN=FABS[(AN-B/A)/(1+AN*B/A)];S AN=AN/FSQT(1+AN*AN)
11.40 S INT=TO*TO/T*T;I (T-TO-70)11.5;S INT=INT*(90-T+TO)/20
11.50 I (K)11.52;S EL=FSQT[1-(2.915*C/T(K))↑2];G 11.6
11.52 S EL=FSQT[1-(2.915*(H-C)/T(K))↑2]
11.60 S LA=INT*EL*AN
11.80 I (T-TO-3)11.81;I (8-T+TO)11.9;S LA=LA*(T-TO-3)/5;G 11.9
11.81 S LA=0
11.90 S L8=L8+LA;S T8=T8+INT
11.92 I (0.866-EL)11.94;S TB=INT;S BL=LA;G 11.96
11.94 S TB=INT*0.1;S BL=LA*0.1
11.96 S B8=B8+TB;S E8=E8+BL

29.34 T "S(80)="%3.01,4.343*FNEW(L8/(T8-L8),15),"DB"
29.40 T "                         SB(80)="4.343*FNEW(E8/(B8-E8),15),"DB",!
29.50 F I=1,1,26;T " "
29.60 T "150HZ LEVEL="4.343*FNEW(B8/T8,15),"DB",!
29.90 Q
```

Figure III.3. Supplementary computer programme for the degree of spatial impression.

APPENDIX IV

DETAILS OF THE RECORDED TAPE CONTAINED WITHIN THE REAR COVER

This tape is suitable for playing on a twin track stereophonic tape recorder at a speed of 19 cm/s ($7\frac{1}{2}$ in/s). The tape is designed to be listened to over stereo headphones. The duration of the tape is $7\frac{3}{4}$ minutes.

The recording was made with an artificial head in the position of a subject's head at the centre of the loudspeaker array used for simulation experiments. I am indebted to Mrs. Celia Campbell for allowing me the use of this artificial head, which she constructed. Apart from slight front-back ambiguity (which is a frequent short-coming with headphone reproduction), the reproduction in the horizontal plane with this head was good. For example, comparison experiments for spatial impression, as described in Chapters 7-10, could also be conducted via the artificial head with relative ease. Localisation in the median plane via the artificial head was poor, but no recordings were made for this tape with elevated reflections. Whilst there is inevitably a certain degradation of quality with such a recording relative to the real situation for a subject, the reproduction of the subjective effects produced by early reflections and reverberation is good.

The tape consists of 18 repetitions of a short motif from Wagner's "Siegfried Idyll". The repetitions are divided into five bands which deal with the following aspects.

- Band A: The subjective effect of audience attenuation filtering (see section 11.2)
- Band B: The subjective effect of early lateral reflections (see section 4.2(e))
- Band C: The subjective effect of reverberation (see section 4.2(e) and Chapter 12)
- Band D: The subjective effect of early lateral reflections in a reverberant field (see Chapter 12)
- Band E: The importance of bass early lateral sound and a comparison of the effects of early and late bass sound (see Chapter 12).

Each band is separated by an interval of about 10s, whilst individual repetitions are separated by intervals of 5s. The differences between individual repetitions are very obvious in the earlier bands, becoming less obvious in Band D. In Band E differences are particularly difficult to distinguish; high quality headphones and a high degree of concentration by the listener are required here. This band in particular, though this also applies to the whole tape recording, is included purely for illustration. Since the recording procedure is imperfect, an inability to perceive a difference described in the text should be viewed circumspectly.

The recordings were made with direct sound, a pair of lateral reflections (delay 40 and 41.5 ms), one on each side of the head, and reverberation, derived from a reverberation plate and played through four symmetrically placed loudspeakers (see section 3.3(d)). In Band A, D and E the audience attenuation filter was also used (see section 11.2); "filtered" refers to this audience filter being in circuit. In Band B and C a gradual transition from one situation to another is made; switching from one situation to another at 2.5 or 4 second intervals was also frequently used. The symbol S indicates the ratio of lateral to non-lateral early sound (at mid-frequencies) and R indicates the ratio of reverberant to early energy. The contents of the individual tracks are listed below.

Band A

1. Direct sound alone.
2. Direct sound audience filtered.
3. Direct sound switching filtering in and out.

Band B

1. Direct sound plus lateral reflections. Gradual transition from $S = -20$ to $+4$ dB.

2. Direct sound plus lateral reflections. Gradual transition from $S = +4$ to -20 dB.

3. Direct sound with lateral reflections ($S = -6$ dB) switched in and out.

Band C

1. Direct sound plus reverberation. ($R = -2.5$dB)

2. Direct sound plus reverberation. Gradual transition from $R = -20$ to $+4$ dB.

3. Direct sound plus reverberation. Gradual transition from R = +4 to -20 dB.

Band D

1. Direct sound plus lateral reflections plus reverberation (R = -4.5 dB; S = -2.5 dB).

2. Direct sound plus reverberation. Lateral reflections switched in and out (R = -4.5 dB; S = -2.5 dB).

3. Ditto.

4. Direct sound plus reverberation. Filtered lateral reflections switched in and out (R = -4.5 dB; S = -2.5 dB).

5. Ditto.

Band E

1. Direct sound plus reverberation. Switching from filtered to unfiltered lateral reflections (settings as before).

2. Ditto.

3. Direct sound plus lateral reflections plus reverberation. Switching from unfiltered lateral reflections to filtered lateral reflections with reverberation bass boosted.

4. Ditto.

BIBLIOGRAPHY

1. E.C. CHERRY 1959 in "Sensory Communication" Ed. W.A. Rosenblith. Massachusetts: M.I.T. Press. Chapter 6: Two ears - but one world.

2. R.J. HAWKES and H. DOUGLAS 1971 Acustica 24, 235. Subjective acoustic experience in concert auditoria.

3. H.R. HUMPHREYS et al. 1967 Architects' Journal (13 Sept.), p.687. Concert hall for the Aldeburgh Festival of Music and the Arts.

4. P.H. PARKIN and K. MORGAN 1970 J. Acoust. Soc. Am. 48, 1025. "Assisted Resonance" in The Royal Festival Hall, London 1965-1969.

5. G. DODD 1972 Ph.D. Thesis, University of Southampton. Assisted Resonance and room acoustics in small auditoria.

6. A.H. MARSHALL 1967 J. Sound Vib. 5, 100. A note on the importance of room cross-section in concert halls.

7. A.H. MARSHALL 1968 J. Sound Vib. 7, 116. Levels of reflection masking in concert halls.

8. W. DE V. KEET 1968 6th Int. Congr. on Acoustics, Tokyo, E-2-4. The influence of early reflections on the Spatial Impression.

9. E.G. RICHARDSON and E. MEYER 1962 Technical aspects of sound. Vol. III. New York: Elsevier Publishing Company.

10. H.P. SERAPHIM 1963 Acustica 13, 75. Raumakustische Nachbildungen mit elektroakustischen Hilfsmitteln.

11. W. KUHL 1969 Rundfunktechn. Mitteilungen 13, 205. Unterschiedliche Bedingungen beim Hören in einem Raum und bei elektroakustischen Übertragungen.

12. W. REICHARDT 1968 6th Int. Congr. on Acoustics, Tokyo, GP-2-2. Der Impuls -Schalltest und seine raumakustische Beurteilung.

13. C.L.S. GILFORD 1972 "Acoustics for radio and television studios". I.E.E. Monograph Series 11. London: Peter Peregrinus Ltd.

14. T.J. SCHULTZ and B.G. WATTERS 1964 J. Acoust. Soc. Am. 36, 885. Propagation of sound across audience seating.

15. G.M. SESSLER and J.E. WEST 1964 J. Acoust. Soc. Am. 36, 1725. Sound transmission over theatre seats.

16. J.E. WEST and G.M. SESSLER 1966 J. Acoust. Soc. Am. 40, 1246. Model study of the sound transmission over raked theatre seats.

17. D. KUNSTMANN 1967 Acustica 18, 259. Modelluntersuchungen zur Schallausbreitung über Publikum (Schallstreuung am Kugelzeilen).

18. E. MEYER, H. KUTTRUFF and F. SCHULTE 1965 Acustica 15, 175. Versuche zur Schallausbreitung über Publikum.

19. V.V. FURDUEV 1971 Sov. Phys. Acoustics 16, 282. Audience sound absorption. Research methods and results (review).

20. T. SOMERVILLE 1953 Acustica 3, 365. An empirical acoustic criterion.

21. M.R. SCHROEDER 1965 J. acoust. Soc. Am. 37, 409. New method of measuring reverberation time.

22. M.R. SCHROEDER 1966 J. acoust. Soc. Am. 40, 549. Complementarity of sound buildup and decay.

23. B.S. ATAL, M.R. SCHROEDER and G.M. SESSLER 1965 5th Int. Congr. on Acoustics, Liège. G32. Subjective reverberation time and its relation to sound decay.

24. H. HAAS 1951 Acustica 1, 49. Uber den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache.

25. R.H. BOLT and P.E. DOAK 1950 J. acoust Soc. Am. 22, 507. A tentative criterion for the short-term transient response of auditoriums.

26. A.F.B. NICKSON, R.W. MUNCEY and P. DUBOUT 1954 Acustica 4, 515. The acceptability of artificial echoes with reverberant speech and music.

27. J.P.A. LOCHNER and J.F. BURGER 1958 Acustica 8, 1. The subjective masking of short time delayed echoes by their primary sounds and their contribution to the intelligibility of speech.

28. J.P.A. LOCHNER and J.F. BURGER 1964 J. Sound Vib. 1, 426. The influence of reflections on auditorium acoustics.

29. H. NIESE 1957 Hochfreq Tech. Elektroakust. 66, 70. Die Prufung des raumakustischen "Echogradkriteriums" mit Hilfe von Silbenverständlichbeitsmessungen.

30. R. KURER 1972 Ph.D. Thesis, Berlin. Untersuchungen zur Auswertung von Impulsmessungen in der Raumakustik.

31. L.L. BERANEK and T.J. SCHULTZ 1965 Acustica, 15, 307. Some recent experiences in the design and testing of concert halls with suspended panel arrays.

32. W. SCHMIDT 1968 Hochfreq Tech. Elektroakust. 77, 37. Zusammenhang zwischen Hallabstand und Nachhallzeit für den Raumeindruck (Halligkeit und Raumlichkeit bei Musik).

33. W. REICHARDT 1970 Hochfreq Tech. Elektroakust. 79, 121. Vergleich der objectiven raumakustischen Kriterien für Musik.

34. V.L. JORDAN 1969 Applied Acoustics 2, 59. Room acoustics and architectural acoustics development in recent years.

35. W. REICHARDT and W. SCHMIDT 1966 Acustica 17, 175. Die horbaren Stufen des Raumeindruckes bei Musik.

36. K.A. MACFADYEN 1970 Applied Acoustics 3, 181. A method of assessing musical definition in an auditorium.

37. E. MEYER and G.R. SCHODDER 1952 Nachr. Akad. Wiss. Göttingen Math.-Phys. Klasse No. 6, 31. Uber den Einfluss von Schallrückwürfen auf Richtungslokalisation und Lautstärke bei Sprache.

38. P. SCHUBERT 1966 Tech. Mitt. RFZ, Heft 3, 124. Untersuchungen über die Wahrnehmbarkeit von Einzelrückwürfen bei Musik.

39. M.R. SCHROEDER, B.S. ATAL, G.N. SESSLER and J.E. WEST 1966 J. acoust. Soc. Am. 40, 434. Acoustical measurements in the Philharmonic Hall (New York).

40. J.E. WEST 1966 J. acoust. Soc. Am. 40, 1245. Possible subjective significance of the ratio of height to width in concert halls.

41. L.L. BERANEK 1962 "Music Acoustics and Architecture". New York: John Wiley and Sons Inc.

42. H. WILKENS 1972 Acustica 26, 213. Kopfbezügliche Stereophonie - ein Hilfsmittel für Vergleich und Beurteilung verschiedener Raumeindrücke.

43. V. MELLERT 1972 J. acoust. Soc. Am. 51, 1359. Construction of a dummy head after new measurements of thresholds of hearing.

44. E. MEYER, W. BURGTORF and P. DAMASKE 1965 Acustica 15, 339. Eine Apparatur zur elektroakustischen Nachbildung von Schallfeldern. Subjective Hörwirkungen beim Ubergang Kohärenz-Inkohärenz.

45. L. MULLER 1968 6th Int. Congr. on Acoustics, Tokyo. E-2-6. Zur Klangfärbung durch Kurzzeitreflectionen bei Rauschen, Sprache und Musik.

46. J.C.R. LICKLIDER 1956 in Third Symposium on Information Theory, Ed. C. Cherry. New York: Academic Press Inc.

47. B.S. ATAL, M.R. SCHROEDER and H. KUTTRUFF 1962 4th Int. Congr. on Acoustics, H31. Perception of colouration in filtered gaussian noise - short-time spectral analysis by the ear.

48. F.A. BILSEN 1968 Ph.D. Thesis, Technische Hogeschool, Delft. On the interaction of sound with its repetitions.

49. H. KUTTRUFF 1965 Acustica 16, 166. Uber Autokorrelationsmessungen in der Raumakustik.

50. T. SOMERVILLE, C.L.S. GILFORD, N.F. SPRING and R.D.M. NEGUS 1966 J. Sound Vib. 3, 127. Recent work on the effects of reflectors in concert halls and music studios.

51. W. KUHL 1965 Rundf. Techn. Mitt. 9, 170. Das Zusammenwirken von direktem Schall, ersten Reflexionen und Nachhall bei der Hörsamkeit von Räumen und bei Schallaufnahmen.

52. J.L. FLANAGAN and R.C. LUMMIS 1970 J. acoust. Soc. Am. 47, 1475. Signal processing to reduce multipath distortion in small rooms.

53. R.W. MUNCEY, A.F.B. NICKSON and P. DUBOUT 1953 Acustica 3, 168. The acceptability of speech and music with a single artificial echo.

54. E. MEYER and W. KUHL 1952 Acustica 2, 77. Bemerkungen zur geometrischen Raumakustik.

55. P. DUBOUT 1958 Acustica 8, 371. Perception of artificial echoes of medium delay.

56. H. NIESE 1956 Hochfreq Tech. Elektroakust. 65, 4. Vorschlag für die Definition und Messung der Deutlichkeit nach subjectiven Grundlagen.

57. J.P.A. LOCHNER and W. DE V. KEET 1960 J. acoust. Soc. Am. 32, 393. Stereophonic and quasi-stereophonic reproduction.

58. J. BLAUERT 1970 Acustica 22, 205. Sound localisation in the median plane.

59. J. BLAUERT 1971 J. acoust. Soc. Am. 50, 466. Localisation and the law of the first wavefront in the median plane.

60. H.P. SERAPHIM 1961 Acustica 11, 80. Uber die Wahruehmbarkeit mehrerer Ruckwurfe von Sprachschall.

61. W. BURGTORF 1961 Acustica 11, 97. Untersuchungen zur Wahrnehmbarkeit verzögerter Schallsignale.

62. P. SCHUBERT 1969 Hochfreq Tech. Elektroakust. 78, 230. Die Wahrnehmbarkeit von Rückwürfen bei Musik.

63. M. BARRON 1971 J. Sound Vib. 15, 475. The subjective effects of first reflections in concert halls - the need for lateral reflections.

64. W. REICHARDT and W. SCHMIDT 1967 Acustica 18, 274. Die Wahrnehmbarkeit der Veränderung von Schallfeldparametern bei der Darbietung von Musik.

65. P. DAMASKE 1967 Acustica 19, 199. Subjective Untersuchung von Schallfeldern.

66. B. WAGENER 1971 Acustica 25, 203. Räumliche Verteilungen der Hörrichtungen in synthetischen Schallfeldern.

67. R.G. WETTSCHURECK 1973 Acustica 28, 197. Die absoluten Unterschiedsschwellen der Richtungswahrnehung in der Medianebene beim natürlichen Hören, sowie beim Hören über ein Kunstkopf - Ubertragungssystem.

68. T.J. SCHULTZ and B.G. WATTERS 1964 J. acoust. Soc. Am. 36, 897. Perception of music heard via interfering paths.

69. G. von BEKESY 1960 "Experiments in hearing". New York: McGraw-Hill.

70. J. BLAUERT 1970 Acustica 23, 287. Zur Trägheit des Richtungshörens bei Laufzeit- und Intensitätsstereophonie.

71. J.V. TOBIAS and S. ZERLIN 1959 J. acoust. Soc. Am. 31, 1591. Lateralisation threshold as a function of stimulus duration.

72. A.W. MILLS 1972 in "Foundations of Modern Auditory Theory". Vol. II. Ed. J.V. Tobias. New York: Academic Press. Chapter 8: Auditory Localization.

73. W.E. FEDDERSEN, T.T. SANDEL, D.C. TEAS and L.A. JEFFRESS 1957 J. acoust. Soc. Am. 29, 988. Localization of high-frequency tones.

74. A.W. MILLS 1963 in "Proceedings of the international congress on technology and blindness" Vol. 2. New York: American Foundation for the Blind. p. 111: Auditory perception of spatial relations.

75. S.S. STEVENS and E.B. NEWMAN 1936 Am. J. Psychol. 48, 297. The localisation of actual sources of sound.

76. B. McA. SAYERS and P.A. LYNN 1968 J. acoust. Soc. Am. 44, 973. Interaural amplitude effects in binaural hearing.

77. P. DAMASKE 1970 Acustica 22, 191. Richtungsabhängigkeit von Spectrum und Korrelationsfunktionen der an den Ohren empfangenen Signale.

78. P. DAMASKE and Y. ANDO 1972 Acustica 27, 232. Interaural cross-correlation for multichannel loudspeaker reproduction.

79. W. de V. KEET Personal correspondence.

80. J.F. BURGER and W. de V. KEET 1973 Applied Acoustics 6, 243. Rapid assessment of interchannel coherence.

81. I. POLLACK and W.J. TRITTIPOE 1959 J. acoust. Soc. Am. 31, 1250. Binaural listening and interaural noise cross-correlation.

82. I. POLLACK and W.J. TRITTIPOE 1959 J. acoust. Soc. Am. 31, 1616. Interaural noise-correlations: examination of variables.

83. A. KROKSTAD, S. STRØM and S. SØRSDAL 1968 J. Sound Vib. 8, 118. Calculating the acoustical room response by the use of a ray tracing technique.

84. H. KUTTRUFF 1971 Acustica 25, 333. Simulierte Nachhallkurven in Rechteckräumen mit diffusem Schallfeld.

85. R.H. BOLT, P.E. DOAK and P.J. WESTERVELT 1950 J. acoust. Soc. Am. 22, 328. Pulse statistics analysis of room acoustics.

86. P.E. DOAK 1959 Acustica 9, 1. Fluctuations of the sound pressure level in rooms when the receiver position is varied.

87. B.M. GIBBS and D.K. JONES 1972 Acustica 26, 24. A simple image method for calculating the distribution of sound pressure levels within an enclosure.

88. F.V. HUNT 1964 J. acoust. Soc. Am. 36, 556. Remarks on the mean free path problem.

89. M.R. SCHROEDER 1970 J. acoust. Soc. Am. 47, 424. Digital simulation of sound transmission in reverberant spaces.

90. P.M. MORSE and R.H. BOLT 1944 Rev. Mod. Physics 16, 69. Sound waves in rooms.

91. L.L. BERANEK 1969 J. acoust. Soc. Am. 45, 13. Audience and chair absorption in large halls, II.

92. B.S. ATAL, M.R. SCHROEDER, G.M. SESSLER and J.E. WEST 1966 J. acoust. Soc. Am. 40, 428. Evaluation of acoustic properties of enclosures by means of digital computers.

93. J. MEYER 1971 7th Int. Congr. on Acoustics, Budapest, 19A1. Uber die Hallradien einiger Musikinstrumente in verschiedenen Konzertsälen.

94. J. MEYER 1972 J. acoust. Soc. Am. 51, 1994. Directivity of the bowed string instruments and its effect on orchestral sound in concert halls.

95. W. KUHL 1969 Rundfunktechn. Mitteilungen 13, 210. Zusammenfassung der Diskussionen des Symposiums über die Planung von Konzertsälen und Studios mit Hilfe raumakustische Modele.

96. P.H. PARKIN and H.R. HUMPHREYS 1958 "Acoustics, Noise and Buildings". London: Faber and Faber Ltd.

97. L.L. BERANEK 1954 "Acoustics". New York: Mcgraw Hill Book Co. Inc.

98. A.H. MARSHALL to be published.

99. F.A. SAUNDERS 1962 Sound - Its Uses and Control 1, 10. Violins old and new - an experimental study.

100. L.B. PREIZER 1966 Sov. Phys. Ac. 11, 407. Statistics of high-level reflections in auditoriums.

101. M. ROSS 1972 Castle Terrace Theatre Project - Acoustics Symposium. Sandy Brown Assoc. and Edinburgh City Council.

102. W. GABLER, R. BUCKLEIN, E. KRAUTH and F. SPANDOCK 1968 Acustica 19, 264. Untersuchungen an einer akustisch 'idealen' Raumform.

103. A.H. MARSHALL 1968 6th Int. Congr. on Acoustics, Tokyo, E-2-3. Concert hall shapes for minimum masking of lateral reflections.

CONTENTS OF THESIS TAPE

Stereophonic recording at 19 cm/s tape speed (CCIR recording characteristic). To be listened to over stereo headphones. Tape duration: $7\frac{3}{4}$ min. For further details see Appendix IV.

Band A: The subjective effect of audience attenuation filtering

1. Direct sound alone.

2. Direct sound audience filtered.

3. Direct sound switching filtering in and out.

Band B: The subjective effect of early lateral reflections

1. Direct sound plus lateral reflections. Gradual transition from S = -20 to +4 dB.

2. Direct sound plus lateral reflections. Gradual transition from S = +4 to -20 dB.

3. Direct sound with lateral reflections (S = -6 dB) switched in and ou

Band C: The subjective effect of reverberation

1. Direct sound plus reverberation (R = -2.5 dB).

2. Direct sound plus reverberation. Gradual transition from R = -20 to +4 dB.

3. Direct sound plus reverberation. Gradual transition from R = +4 to -20 dB.

Band D: The subjective effect of early lateral reflections in a reverberant field

1. Direct sound plus lateral reflections plus reverberation (R = -4.5 dB; S = -2.5 dB).

2. Direct sound plus reverberation. Lateral reflections switched in and out (R = -4.5 dB; S = -2.5 dB).

3. Ditto.

4. Direct sound plus reverberation. Filtered lateral reflections switched in and out (R = -4.5 dB; S = -2.5 dB).

5. Ditto.

Band E: The importance of bass early lateral sound and a comparison of the effects of early and late bass sound

1. Direct sound plus reverberation. Switching from filtered to unfiltered lateral reflections (settings as before).

2. Ditto.

3. Direct sound plus lateral reflections plus reverberation. Switching from unfiltered lateral reflections to filtered lateral reflections with reverberation bass boosted.

4. Ditto.