

ACOUSTICS *and* AUDIO TECHNOLOGY

Third Edition



Mendel Kleiner

Room Acoustics Planning and Design

6

6.1 INTRODUCTION

The purpose of this chapter is to give examples of the kind of deliberations and concerns that are involved in the planning and design of rooms. Many of the physical and psychoacoustic aspects of room acoustics have already been presented and discussed in Chapters 4 and 5.

The purpose of room acoustics planning and design is usually to provide rooms that allow good communication between people—for example, by providing good acoustics for speech and music. This requires not only specific acoustic properties regarding the way the room guides the sound from the speaker or performer to the audience and other orchestra members, but also that competing sounds, such as noise are suitably controlled.

The requirements for the way in which sound is distributed in the room are not only expressed by sound level and sound level distributions but are also expressed with timbre and temporal characteristics. The temporal characteristics are a result of the room size and shape, as well as the way the audience and other sound-absorbing and scattering areas are distributed over the room's boundaries. It is important to note that there are substantial interindividual preference differences between people, and that the psychoacoustics data presented in Chapter 5 is not exact. Much room acoustic data still needs to be collected and compared to personal preference for improving the mapping between the physical domain of room design and the subjective domain of acoustics appreciation.

Performers and audiences may have conflicting requirements regarding suitable acoustics. Solo performers, such as speakers and singers will usually want to sense the acoustic response of the auditorium, whereas musicians playing in groups or orchestras will usually put more emphasis on the way the room allows them to interact.

Background noise levels must also be sufficiently low to allow good communication. In rooms for music, low levels of background noise are needed not only to keep the noise from masking the *running* music, but also so that people will be able to hear the full dynamic range and spatial qualities of the reverberation processes. Background noise can further influence and predispose people's reactions to the acoustic quality of the room. The sound quality of the noise will be different depending on whether the noise is due to heating, ventilation, and air conditioning noise (usually called HVAC noise), to other activities in the building, or to traffic, neighboring industry, and other exterior noise. This leads to different noise criteria (NC) being necessary depending on the noise and the use of the room.

For many types of rooms, the room acoustical conditions are so important that acoustic planning and design must be introduced in the early sketching stage. Cooperation between architect and acoustician is necessary for excellent acoustics. The possibilities of changing the general room acoustics of a room are limited once the room exists physically, without building a new interior shell or implementing other drastic changes.

Electronic architecture—that is, active sound field control, using microphones and loudspeakers in the room in conjunction with electronic signal processing—may sometimes be a way to improve acoustical conditions so that the room may be useful for some purpose other than the ones allowed by *natural* or *passive* acoustics. *Sound reinforcement systems* are often used to provide better conditions for listening to speech in acoustically difficult environments, such as those characterized by excessive distance or reverberation, or inadequate sound reflections.

It is important to stress that the quality of room acoustics is difficult to measure. Chapter 5 discussed some metrics for room acoustics quality. These metrics often require considerable experience and expertise to be applied correctly. Often, listening and subjective judgment by experts is the only way to assess the room acoustics quality. Hearing is an important measurement device.

6.2 BASIC REQUIREMENTS FOR GOOD ROOM ACOUSTICS

Quiet: Acceptable Noise Levels and Noise Sound Quality

The requirements of noise that does not damage hearing were discussed in Chapter 3. In designing the acoustics of rooms for homes, offices, industry, and entertainment, additional NC must be met. What is appropriate will depend on the use of the room and on the requirement for visual and auditory impressions to be matching. We expect a concert hall to be quiet, whereas we will not be surprised to find an industrial workplace somewhat noisy. Sometimes the noise is added intentionally as *acoustic perfume*, such as Muzak-type background music, broadband noise having an unobtrusive spectral character, or even intentional HVAC noise.

Typically, noisy environments may be found in schools, offices, industry, and restaurants. From the viewpoint of the acoustician, these environments require special consideration: limiting noise so that it does not cause hearing loss and is acceptable for the use of the room and designing sound reinforcement and other sound amplification equipment for maximum speech intelligibility and other program enjoyment.

Noise level requirements will vary between rooms and room uses. Different noise criteria typically must be applied to office equipment for use in small offices and in office landscapes. The maximum noise is often specified by the maximum allowable A-weighted sound pressure level, without considering the spatial characteristics of the sound field relative to the person being exposed to the sound. The sound pressure at the entrance of the ear canal will typically vary by at least ± 10 dB at frequencies above 2 kHz due to the angle of sound incidence.

Not only does a 10 dB sound pressure level increase in the midfrequency range typically correspond to more than a doubling of loudness, the change may also considerably increase the risk of noise-induced hearing loss and influence *task performance*. It is also important to note that an increase in A-weighted sound pressure level does not necessarily increase *annoyance*. The temporal and other char-

acteristics of the noise are important in this respect. One often uses various metrics of *Product Sound Quality* to further describe the properties of noise, see Chapter 17. The noise may also be described using special terminology usually specific to a certain engineering field; for example, in describing the noise of air handling equipment, words such as hiss, rumble, and roar are used.

It is important to stress the need for low background noise. Most noise criteria have been set with persons having normal hearing in mind. In societies with much sound-generating equipment (mechanical as well as electroacoustical), people are subject to extra noise-induced hearing impairments. Persons having hearing impairments or hearing loss generally have difficulty in understanding speech buried in noise and are often more annoyed by noise than persons having normal hearing.

Noisy Environments

Much of the negative influence of heating, ventilation, and air conditioning noise on speech is typically due to masking. Frequency domain masking is primarily active upward in frequency as discussed in Chapter 3, where metrics such as signal-to-noise (S/N) ratio, articulation index, speech transmission index, among others, were introduced. As indicated by Figure 3.34, it is reasonable to expect that the S/N ratio (assuming typical speech and HVAC noise spectra) must be at least 15 to 25 dB for speech intelligibility to be unaffected by noise in the speech frequency range.

The Speech Interference Level (SIL) can be used to estimate the seriousness of noise affecting speech. SIL is defined as the mean of the 0.5, 1, and 2 kHz octave band noise levels (see Reference 6.1). Using the data shown in Figure 6.1, one can roughly determine the interference by a particular noise on

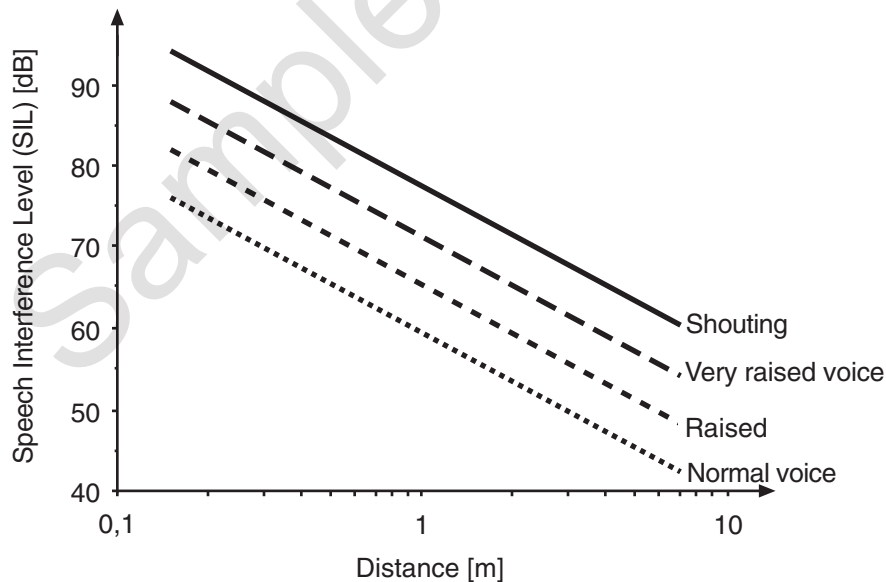


Figure 6.1 The Speech Interference Level (SIL) is defined as the mean of the sound pressure level values for noise spectrum in the 0.5, 1, and 2 kHz octave bands. The graph shows SIL for speech at various distances and for different voice modes (after Ref. 6.1).

speech, not only for offices and other workplaces but also for telephone transmission and similar cases. The data can also be used to determine the necessary sound pressure levels for sound reinforcement equipment.

Quiet Environments

Some environments are characterized by communication activities that are particularly sensitive to noise. This is particularly true for audio and video recording studios, radio and television stations, concert halls, operas, theaters, and similar venues. The background noise limits are usually set during consultation between director, fundraiser, architect, and acoustician.

In rooms for music in particular, the noise levels tolerated are very low. Curves, such as those shown in Figures 6.2 and 6.3, are often used in determining whether a noise spectrum is acceptable, or not. Such curves, specifying maximum octave band sound pressure levels from 31 Hz to 16 kHz, are useful because they, to some degree, take the character of noise into account.

In Europe, it is common to specify the maximum allowable noise level with dBA values. This is inferior compared to the U.S. way of specifying the maximum allowable noise as an NC or RC value. The noise criterion (NC) curves, shown in Figure 6.2, and the A filter curve are based on auditory threshold data. Noise shaped along the NC curves has a tendency to sound both *rumbly* and *hissy*. The NC curves were not shaped to have the best spectrum shape but rather to permit satisfactory speech communication without the noise being annoying (see Reference 6.2).

The room criterion (RC) curves shown in Figure 6.3, on the other hand, have been shaped to be perceptually neutral—they are straight lines with a slope of -5 dB/octave. The RC method involves the determination of both an RC rating and a *spectrum quality descriptor* that determines if the spectrum is rumbly or hissy (see Reference 6.2).

In some cases, one has to be content with a value for sound level L_{PA} . Concert halls usually require sound levels well below 20 dBA. The noise of microphones and HVAC equipment typically have very

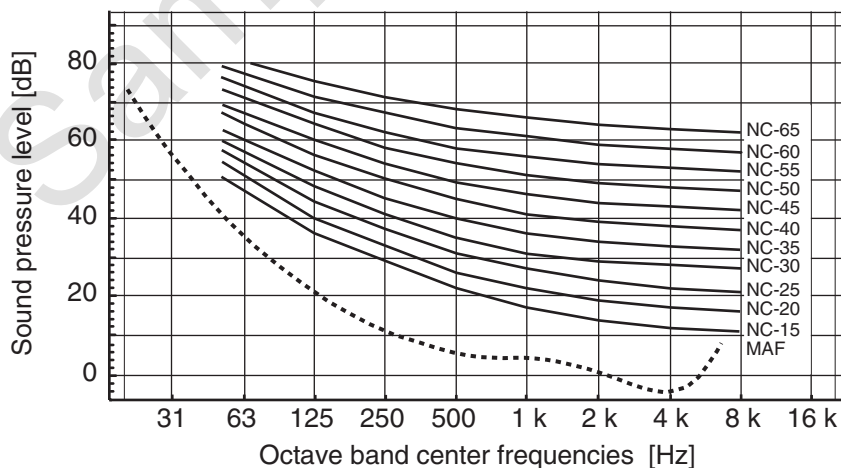


Figure 6.2 These NC curves apply to octave band-filtered noise levels and are often used in room acoustics (after Ref. 6.2).

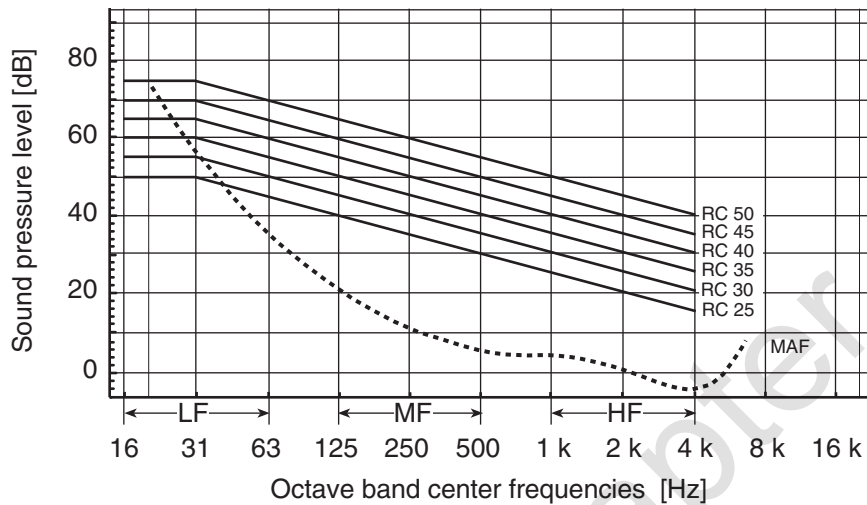


Figure 6.3 The RC curves also apply to octave band-filtered noise levels. Note that there are several editions of these curves—the ones shown are the Mark II curves (after Ref. 6.2).

different spectra. Comparing A-weighted sound levels for different spectra shapes is not relevant from the viewpoint of sound quality, as in the following example.

The acoustic background noise in a sound recording studio and the equivalent acoustic-electronic-background noise of microphones usually have different spectral shapes. Although one can say that the A-weighted sound level of microphone noise should be below the acoustic noise level in the recording room, the fact is that the noises will be perceived differently due to their different sound quality characteristics. The acoustic noise (for example, HVAC and traffic) will generally have spectra dominated by low-frequency noise, whereas microphone noise (due to its electronic generation mechanism) tends to have a *white spectrum*—that is, the octave band levels increase by 3 dB per octave. This gives the microphone noise subjectively a hissy character. Therefore, a microphone having an equivalent 20 dBA sound level internal noise level will probably sound hissy to the recording engineer when used for recording in a concert hall having an acoustic background sound level of 20 dBA. This may be a considerable problem when recording music in large concert halls, such as the one shown in Figure 6.4.

The recognizable properties of noise are usually important in the environments under discussion here. It may sometimes be easier to tolerate a noise having a random character than a noise that is transient, has a specific time pattern, or carries information content. The spatial distribution of the noise is also important. A diffuse sounding noise is often less disturbing to some people than a noise that can be localized in space (although in the latter case, one can perhaps more often find it possible to eliminate the noise source).

It is also important to avoid audible pure tones in the noise spectrum—the audibility of such tones was discussed in Chapter 3. It is reasonable to require the tones to be inaudible, which requires their sound pressure level to be below the critical band level of the other noise in the band. In practice, this results in the need for such tones to be at least 5 dB below the octave band level of the rest of the noise in the band.



Figure 6.4 In a modern concert hall, such as Symphony Hall, Birmingham, United Kingdom, the background noise sound level L_{pA} may be as low as 11 dB. (Photo by Mendel Kleiner.)

6.3 FUNDAMENTALS OF ROOM ACOUSTIC PLANNING

Rooms for Speech

The speech intelligibility properties of rooms intended for verbal communication are important, but it is essential to remember that good speech intelligibility is not the same as naturally sounding speech. One can increase the intelligibility of amplified speech, particularly in noisy surroundings, by various types of signal processing, such as spectral shaping and compression.

For our purpose here one can regard speech as a modulated signal with a complex tonal or noisy spectrum, primarily covering the frequency range 0.25 to 4 kHz that is modulated by frequencies in the 0.2 to 8 Hz range.

Reverberation and noise will decrease the *modulation strength* (also known as the *modulation depth*). A reduction of modulation depth, corresponding to the attenuation of modulation strength (sometimes called the *modulation damping*), results in a reduction of speech intelligibility.

For speech modulation at the listener to not be affected by reverberation, the reverberation time must be short, and the ratio between direct and early reflected sound to late reverberant sound must be high.

There is, however, a practical limit to shortening the reverberation time. Besides the expense of adding sound-absorbent treatment to the room, there is also another aspect. People expect rooms to have a certain reverberation time and level, determined by tradition and visual aspects (e.g., room purpose, shape, and volume).

Speech intelligibility can be shown to increase with increasing reverberation time for very short reverberation times in the range of up to 0.5 s. The reason for this speech intelligibility dependency on reverberation is that reduced sound absorption in the room leads to both an increased reverberation time and an increased sound pressure level. In addition, early sound reflections by more reflective walls

help eliminate the influence of speaker sound radiation directivity on the sound pressure. The directivity index of the human voice is approximately 6 to 10 dB in the frequency range that is most important for speech intelligibility.

These are some good rules for auditoria used primarily for the spoken word:

- Keep the travel paths of direct sound and important early reflections short. In practice, it is difficult for an untrained speaker to reach audiences with sufficient speech intelligibility at distances over 20 m without the use of amplification.
- There must be enough early reflections to provide sufficient sound level while at the same time not feeding sound energy into the late part of the reverberant sound field. It is advantageous to strive for early sound reflections to arrive close to the horizontal plane, rather than from above, to have good speech sound quality. Sometimes for practical reasons, however, it is necessary to make use of overhead sound reflectors as shown in Figure 6.5.
- Mirror-like (specular) reflections must not exceed the level of the reverberant sound if they arrive more than 30 ms after the arrival of the direct sound, such reflections may be perceived as echoes. This means that sound being focused by concave surfaces must have short propagation. The focusing effect may be remedied using alternate focal distances and sound-absorptive or scattering treatments in the travel paths as shown in Figure 6.6.



Figure 6.5 A modern lecture hall equipped with sound-reflecting panels over the sending end of the room in Chalmers University of Technology, Gothenburg, Sweden. (Photo by Mendel Kleiner.)



Figure 6.6 When the sound absorption of the walls is high, a room having high concave ceilings is susceptible to echoes. The ceiling in this church has been made extremely irregular to scatter sound and avoid echo. (Photo by Mendel Kleiner.)

- Take the directivity of the human voice into account. In small rooms, such as classrooms and small auditoria having relatively hard walls, the speaker will usually be quite close to some sound-reflecting surfaces. This leads to strong early reflections that will add to the direct sound in such a way as to increase the speech intelligibility. In large auditoria, such as theaters or assembly and lecture halls, it is important to place the seating area within an approximately 120 degree arc in front of the speaker. The effective speech intelligibility will increase when the listeners can see the mouth movements of the talker.
- The reverberation time should be in the range of 0.6 to 1.0 seconds, increasing with the size of the room, see Figure 6.9. The reverberation time should also be constant to within 10% over the 125 Hz to 8 kHz octave bands, for small- and medium-sized rooms (see Figure 5.17). It is advantageous if added sound-absorbent material is placed so that it does not interfere with the propagation of important early sound reflections. This means that one should leave the central part of the ceiling nonabsorptive in auditoria, classrooms, and other verbal communication rooms.
- For small studios and control rooms, such as those used at radio stations (see Figure 6.7), having volumes smaller than 25 m^3 , it is common to reduce the reverberation time to the



Figure 6.7 Control rooms usually need to have short reverberation times. It is often necessary to remove undesirable reflections by sound absorption or sound scattering. School of Music, Gothenburg, Sweden. (Photo by Mendel Kleiner.)

range 0.2 to 0.4 s. It is also usually necessary to control the resonant modes at low frequencies using various types of porous or resonant sound absorbers.

Rooms for Music

Over the ages, the acoustic properties of the performance space have influenced the way music is written and performed. Composers of music have adapted their works to the reverberation characteristics of the spaces that were familiar to them as likely places for the performance of the particular music. Classical secular music from the baroque period was generally intended for performance in small spaces, having short reflection paths and highly reflective surfaces. Much of the classical music written during the nineteenth century, however, was written for the much larger rooms built in that era that were associated with considerably more reverberant characteristics and nearly twice as long reverberation times (see Reference 6.3).

Today it is common for musical performance spaces to be built acoustically flexible, so that many types of music can be performed successfully—one such flexible venue is shown in Figure 6.8.

Much of the popular music of today does not require a dedicated acoustically reflective environment. Using loudspeakers on and off stage, one can simulate or create the desired acoustical characteristics for the space. *Sound reinforcement systems* and *electronic reverberation enhancements systems*—particularly those outdoors—can be used effectively to simulate many types of spaces.

Some general recommendations for rooms intended for the performance and enjoyment of classical music are (see Reference 6.4):

- The *background noise levels* must be low.
- The *sound level* must be optimized—it should be neither too loud nor too low. It is defined by the *room gain*.



Figure 6.8 In this computer-controlled auditorium at IRCAM, Paris, the room acoustic conditions could be changed drastically using rotating wall sections. The sections were designed to be used as reflectors, scatterers, or absorbers (photo by Leif Rydén).

- There must be suitable *clarity*, and the *spatial* and *temporal distribution of the early reflections* must be good.
- The *reverberation time* must be appropriate for the music expected to be played.

The sound level in a real room will generally not behave exactly in the way described by the simple room acoustic theory presented in Chapter 4. The reason for this is the nondiffuseness of the sound field, sound absorption is not evenly distributed over the room surfaces. In the quasidiffuse sound field in a real room, the level will stay reasonably constant but will drop off with increasing distance. This is particularly noticeable in the audience area of, for instance, a concert hall. Far back on the main floor the sound level will also drop off due to negative interference from the sound reflected off of the seating in front, as discussed in Chapter 1.

The optimal *reverberation time* depends on the size of the room and on the type of performance that is expected to dominate the use of the room (see Figure 6.9). It is common to provide for several auditoria within one venue, which allows for the design of rooms that can complement one another regarding acoustical characteristics.

If it turns out to be necessary to build only one main auditorium, the auditorium can possibly be designed to provide for *variable acoustics* by passive or by active means. *Passive acoustic variation* can be achieved using variable absorption or volume. *Active acoustic variation* can be achieved by electronic sound field simulation or reverberation enhancement using microphones, digital signal processing, amplifiers, and loudspeakers (see Figure 6.10).

The reverberation process can often be considered as subdivided into several parts from the viewpoint of hearing, at least regarding large rooms having reverberation times between 1 and 3 seconds. Usually one considers *direct sound*, *early*, and *late* reverberation.

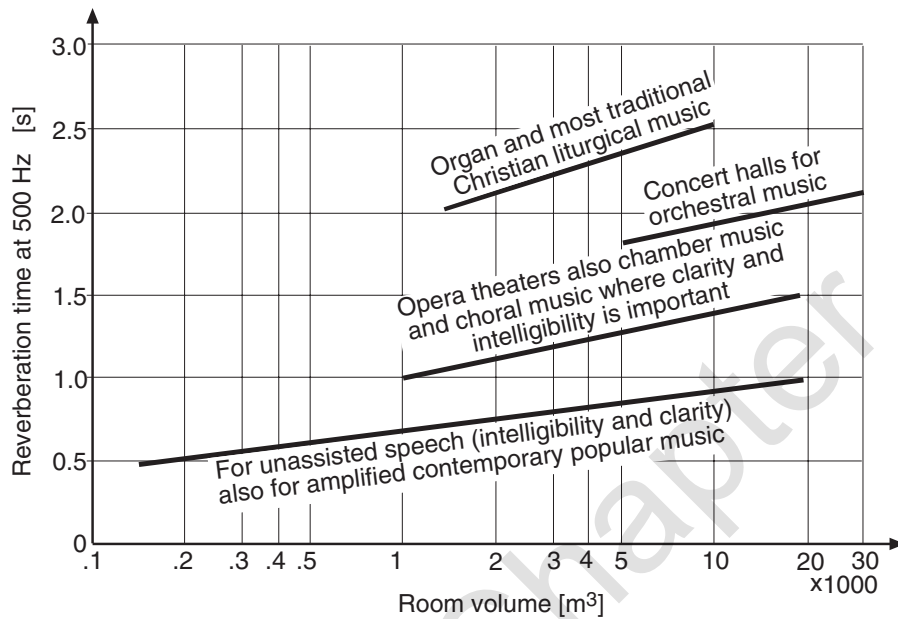


Figure 6.9 Typical recommended 0.5 kHz reverberation times for some common types of auditoria.



Figure 6.10 Using an electronic reverberation enhancement system one can achieve high flexibility in room acoustic conditions. The photo shows an active system for control of acoustics on a stage using loudspeakers integrated into a stage shell to achieve a proper blend of early reflections and local reverberation. Installation in Motala Community Center, Sweden. (Photo by Mendel Kleiner.)

The direct sound is the sound that arrives to the listener by way of the shortest path. The path may, at times, be interrupted by corners and balcony fronts, for example. It is also common to regard some reflections that follow immediately after the direct sound (having a time delay relative to the direct sound of less than approximately 5 ms) as a part of the direct sound. This is particularly the case if they are in the same vertical plane of the listener, such as from the audience area or the stage. From the listener's perceptual viewpoint, these early reflections are *fused* with the direct sound.

The early part of the reverberation process that follows the direct sound in the time range of 5 to 80 ms (5 to 50 ms for speech) is often called *early reflections*. Figure 6.11 shows reflectors suspended over the stage at the Concert Hall, Singletary Center for the Arts, University of Kentucky to enhance the early reflections. The hall was designed using results from sound field simulation, “Auditorium Synthesis”, and by using physical scale modeling (see Reference 6.17). The acoustics of the hall combines high clarity with long reverberation time.

The *late reverberation* is the reverberation that follows the early reflections—that is, more than 80 ms after the arrival of the direct sound. The *very late reverberation* that follows more than about 160 ms after

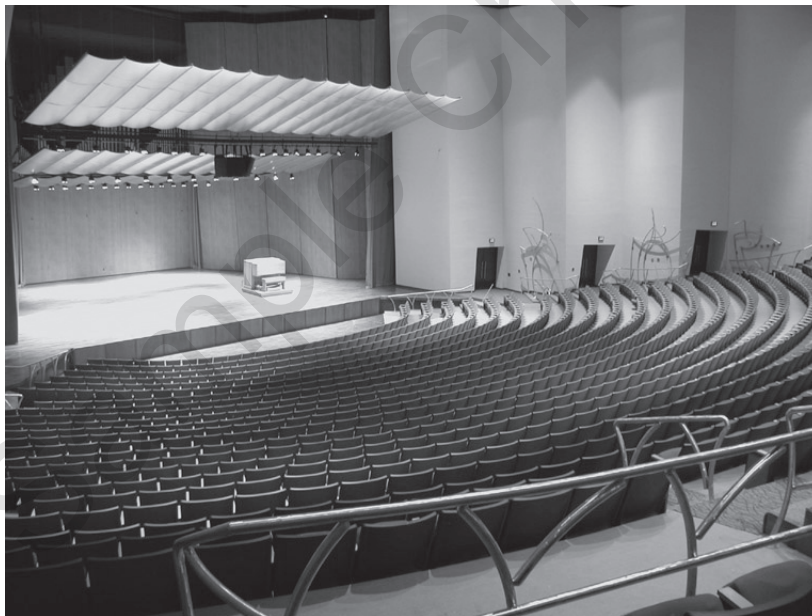


Figure 6.11 This photo shows the Concert Hall, Singletary Center for the Arts, University of Kentucky, Lexington, KY. Using sound reflectors suspended over the stage, one can add early reflections at the listener. In wide auditoria, overhead sound reflectors are usually necessary to provide both good stage acoustics and sufficient early reflected sound for the audience. It is important that the reflectors are large, have suitable scattering properties, and have enough mass so that they will provide strong and even sound coverage. The shape, number, placement, and adjustment of the reflectors are all factors that will influence the results. “Continental seating” is used in this hall. Note the sound reinforcement system’s large central cluster loudspeaker array that is used to amplify speech (photo by Dwight Newton, courtesy of University of Kentucky College of Fine Arts).

the direct sound is usually heard only at the end of a musical piece. It could be argued that this very late reverberation is extremely important since it is the last part of any piece of music heard in a concert hall.

When measuring the reverberation time we usually obtain a decay curve that shows both the early and late parts of the reverberation process. A physical reverberation process, at least when one looks at a wide frequency range response, is not characterized by the exponential decay of the theoretical process discussed in Chapter 4. The real wide bandwidth decay in the frequency range above the Schroeder frequency has many irregularities caused by the approximately discrete reflection sequence of sound in the room. The process will usually approximate that of one or more exponential processes. The part of the reverberation process that is from the start of the process out to a time limit of 0.16 s after the start is particularly important for the subjective determination of reverberation time with running music.

Note that the reverberation process will depend on the location of the sound source and the microphone position in any given room. The ISO 3382 standard gives advice on good practice for the measurement of reverberation time of performance venues.

Some target values for the reverberation time T for some different types of rooms can be found in Figure 6.9. Large rooms for classical music will typically require a volume of around 12 m³ per square meter of audience area to obtain a sufficiently long reverberation time at 1 kHz where the sound absorption of the audience and the chairs is the highest. One often strives for a frequency-independent reverberation time within $\pm 5\%$ in the octave bands from 250 Hz to 4 kHz.

In the case of large rooms, such as concert halls (volumes more than approximately 5,000 m³), the reverberation time will drop off in the 4 kHz octave band and above due to the sound absorption of air, as discussed in Chapter 1. It is common practice to leave surfaces for final adjustment of the reverberation characteristics of a room after completion.

Generally, the sound absorption by common building constructions and surfaces is fairly low in the octave bands at 125 and below. It is common to allow an increase of reverberation time down to the 63 Hz octave band of approximately 30% over the value at 0.5 kHz. This practice results in good bass response and a warm sounding reverberation. Sometimes surfaces, such as the stage floor, may act as secondary sound radiators at low frequencies for cellos and double basses that further enhance the *warmth*.

A room having a short reverberation time is usually considered *dry* and acoustically uninteresting. The desire for high clarity however makes reverberation times in the range 0.2 to 0.4 s desirable for control rooms, such as the one shown in Figure 6.7.

Another main factor influencing room acoustic quality is the spatial and temporal distribution of the early reflections.

The *time gaps* between the direct sound and major early reflections, are important and should be irregular and not arrive later than approximately 30 ms. The early reflections should be wide-band (not have a limited frequency range). One should avoid strong overhead reflections appearing in the same vertical plane as the direct sound. Since hearing cannot differentiate between signals having the same interaural delay time, such overhead reflections will cause comb filter effects unless masked by other reflections coming from the sides. A diffuse or curved ceiling will redirect these reflections to the walls.

The spatial distribution of the early reflected sound is important, particularly for creating the impression of *auditory source width*. One should strive for sound reflection angles close to the horizontal plane to maximize the advantages of our binaural hearing. It is advantageous if the directional distribution of the early reflections are such that the sound field appears to be *symmetrical* when listening. One should also strive for an even angular distribution of the reverberant sound to enhance the feeling of *envelopment* and diffuseness in the sound field.

It is important to realize that subjective diffusivity and physical diffusivity are not well correlated. One can have good subjective sound field diffusivity without the sound field being diffuse in a physical sense. Good physical diffusivity does not guarantee good listening conditions in a room.

The side walls should be somewhat scattering; for example, by using shelves, balconies, unevenly placed sound-absorbing patches or objects having sizes in the range of 0.5 to 0.05 m. Figures 6.12–6.14 show various ways of achieving high scattering by wall and ceiling surfaces. *Flutter echo* is the term given to repetitive short path sound reflections, usually appearing when reflections between two close and parallel walls dominate the sound field. It can be removed by splaying the walls, absorption or scattering.

One phenomenon of particular importance in long concert halls is that of cancellation on the main floor of direct sound by the first reflection of sound off the audience in the frequency range between 125 Hz and 1 kHz. This reflection will be at a small angle leading to phase reversal of the sound in the frequency range mentioned (that is, a reflection coefficient of $R = -1$). Since the sounds cannot be separated because of the short time difference between the direct and reflected sounds (typically less than 5 ms), there will be cancellation. The cancellation will result in *coloration*—timbral change in the sound quality of the *direct* sound.

According to the *law of the first wave front*, the direct sound determines our perception of the timbre of the sound source. Consequently, it is important to design the seating area in such a way that the coloration is minimized. This requires that:

- The chairs are designed appropriately.
- The seating area has a sufficient slope.
- The stage is sufficiently high.
- There are good sight lines to all seats.

A room having good properties regarding background noise, early reflections, reverberation time, and scattered sound will be close to ideal from the audience's acoustical point of view.

In addition, it is important that the acoustical conditions on stage give musicians good possibilities for hearing one another and give support for their own playing. The stage and its surrounding walls and ceiling must be designed so that the sound is redirected appropriately to all parts of the orchestra as well as to the audience. The stage and the auditorium must, in addition, be free of severe acoustical defects, such as echo and nonlinear distortion (for instance, a rattle).

Recording Studios

The acoustical requirements for music recording rooms are similar to those that apply to rooms for listening. Because of the different ways in which a listener interprets the sound by direct listening and by listening to sound picked up by microphones, there will be both difficulties (absence of binaural signal



Figure 6.12 One can scatter the reflections from surfaces using diffusers to soften the early reflections. This photo shows testing of diffusers to create good ensemble characteristics for a concert hall stage in Baltimore's Joseph Meyerhoff Symphony Hall. (Photo by Mendel Kleiner.)



Figure 6.13 Formerly, high sound field diffusivity was achieved by uneven surfaces, niches, and rich ornamentation, as in the Wiener Musikverein Concert Hall, Vienna shown here. (Photo by Mendel Kleiner.)

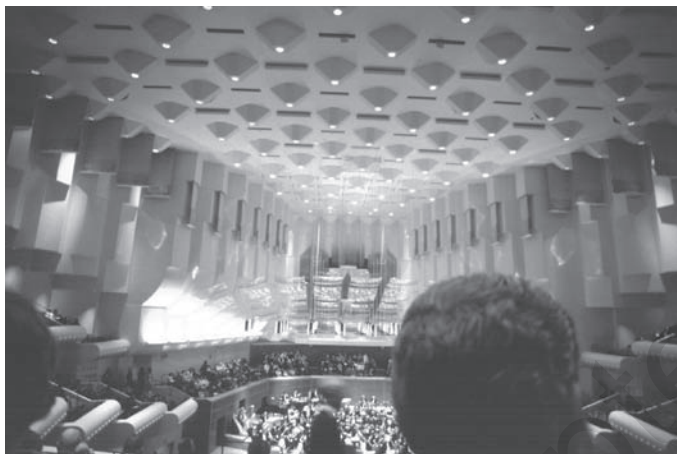


Figure 6.14 Currently, the surface unevenness is achieved by simple means and is primarily introduced to provide desired diffuse sound reflection as in the San Francisco Symphony, San Francisco, CA. (Photo by Mendel Kleiner.)

processing) and advantages (close-up recording, directional microphones, and various forms of signal enhancement). (See Reference 6.5.)

By using various microphone techniques, one can adapt the recording to the room acoustical conditions. Since the visual impression is not as important, more effort can be put into optimization of the room acoustics.

It is common for studios to have various types of sound reflecting, scattering, and absorbing panels, both freestanding and wall mounted. A recording studio venue will also usually contain many small rooms, or subspaces, for particular instruments. These will have splayed surfaces to avoid flutter echo.

Any recording studio must be quiet, since the listener will likely have to listen to the same recording several times. The background noise sound level in a studio for recording of acoustical instruments should preferably be below 15 dBA, as mentioned previously. It is important to observe that the noises from traffic and HVAC have noise spectra that are different from that of typical microphones.

Television studios generally pose acoustical difficulties, from the viewpoint of both noise and room acoustics. Close-up microphone placement (a wearable broadcasting microphone, for example) is often necessary to allow visually unobtrusive sound pickup (see Reference 6.6).

Rooms for Sound Reproduction

Rooms for sound reproduction are typically much smaller than those for concerts or classical music recording—typically having volumes in the range of 25 to 75 m³. This leads to problems at both low and high frequencies.

In sound reproduction the acoustical characteristics of the transfer paths between loudspeaker and listener are overlaid on the recorded signals.

It is necessary to differentiate between listening rooms in studios (control rooms of various kinds) and rooms in dwellings (living rooms, for instance). The studio environment is usually subject to a more critical evaluation and listening than the dwelling environment. Also, there is likely to be more interest and money available for creating a good listening environment in a studio than in a home. The freedom to design the room with acoustics in mind will also be larger since fewer compromises have to be made because of other uses.

Both environments require fairly short reverberation times. As a rough approximation, one should strive for a listening room reverberation time no longer than approximately half that of the recording venue if the recording has been made so that the reverberation is clearly audible and dominant (for example, using omnidirectional microphones far beyond the room's reverberation radius). Typically, a reverberation time in the interval between 0.3–0.7 s may be desired over the entire audio frequency range.

An important issue in playback rooms is to avoid early reflected sound incident in the median plane, since such sound components will cause comb filter effects as discussed previously. The components can be eliminated by using absorptive or scattering treatment (or simple redirection) of the surfaces responsible for the undesired reflections. Sometimes it is also useful to diffuse the wall patches that are responsible for the first side wall reflections. A specification of listening room characteristics is made in the IEC 268–13 standard that strives to formulate the acoustical characteristics of a typical listening room. A room that fulfills the standard is shown in Figure 6.16.



Figure 6.15 Wenger's small commercial portable cabin for music rehearsal featuring electronically variable reverberation. (Photo by Mendel Kleiner.)

In one technique often found in control rooms one half of the room (where the loudspeakers are) is treated to be highly sound absorptive, and the other half (where one listens) is “live” and treated to be diffusely reflecting. This classic approach allows a distinct and clear sound around the listener in combination with a feeling of space. More modern approaches are creating a “reflection free” zone around the listener or making walls and the ceiling diffusely reflecting.

Small rooms usually have problems in accurate music sound reproduction because of the low modal density and poor damping of modes at low frequencies as mentioned previously. This makes it virtually impossible to reproduce music accurately at low frequencies unless these low frequency modes are well damped. Using various types of signal correction by means of digital signal processing, it is possible to circumvent this problem to some extent. The use of various types of low-frequency sound absorption is in any case a must. Slit and Helmholtz resonators, as well as membrane-type sound absorbers, are suitable for this task. Figure 6.17 shows a combination of a diffuser and low-frequency sound absorber.

Correct positioning of the loudspeakers, relative to the room modes and the listening position, will ensure the least variation in frequency response. It is practical to use the reciprocity technique for finding the optimal positions.



Figure 6.16 The IEC listening room at the Section of Acoustics at the Aalborg University is a good example of an advanced IEC listening room (photo courtesy of AAU, Denmark).



Figure 6.17 An example of a combination of slit absorber and diffuser for broadband control of the acoustical conditions in a recording studio in the Gothenburg Opera, Sweden. (Photo by Mendel Kleiner.)

6.4 TOOLS FOR PREDICTION OF ROOM ACOUSTIC RESPONSE

The acoustic properties of a room can be predicted in many different ways using *computer aided design* (CAD) with varying results.

For *low frequencies*, calculations using software based on the *finite element method*, *boundary element method*, or *finite-difference time-domain method* are, in principle, quite exact, but comparatively slow. For the modeling, one needs the impedance data for the boundaries—without which the results may be quite erroneous, particularly regarding the location of resonance frequencies, mode shapes, and damping. The necessary impedance data however is seldom available. The generation of a suitable mesh is also a problem since few current mesh generators are suited to quickly and effectively meshing the geometries common in room acoustics tasks (see References 6.11 to 6.17).

For problems involving *medium and high frequencies*, geometrical acoustics is usually used. Initial design considerations usually include manual ray tracing and reverberation time calculation using Sabine's or *Fitzroy's* formula (see Reference 4.1). Sabine's formula essentially requires the sound field in the room to be diffuse, which is seldom the case. Fitzroy's formula is particularly useful in predicting the reverberation time for those rectangular rooms where sound absorption is unevenly distributed between the walls. For building acoustics and some uses in architectural acoustics software based on statistical energy analysis

(SEA), modeling is also used (see Chapter 10). This approach is based on the study of energy flow between resonant systems.

Ray tracing and mirror image method software is now common in most acoustics consultants' offices. Using CAD, it is possible to predict not only reverberation time, clarity, lateral fraction, and other room acoustics metrics but also to predict the impulse response. Many other modeling methods are, however, also possible (see Reference 6.8).

The following methods are (or have been) used in room acoustics modeling:

- Optical ray tracing using laser
- Optical intensity modeling using lamps
- Ultrasonic scale modeling by the impulse method
- Ultrasonic scale modeling by the intensity method
- Ultrasonic scale modeling by the transmission method
- Physical acoustics modeling using Finite Element (FEM) or Finite Difference Time Domain (FDTD) methods software for low frequency, small room modeling not discussed in this book
- Geometrical acoustics modeling using ray tracing or mirror image modeling software
- SEA modeling using specialized software

The optical methods must be considered obsolete except for some special uses. Optical ray tracing is based on the use of physical scale models (typical scale values are 1:10 to 1:100). The sound source is simulated using a laser or a small light source surrounded by a grill so that light is emitted in sectors. By using a scale model where reflecting surfaces are modeled by mirrors, it is possible to study the distribution of early reflected sound—sound that has been reflected only once or twice against surfaces. The advantage of the method is its simplicity and low cost. It can often be used directly in the architect's scale model. A major disadvantage of the method is that it does not take sound diffraction by edges or scattering by surface irregularities into account.

Optical intensity modeling is based on the same type of physical scale models as those used in optical ray tracing, but in this case, the static light distribution is studied as an approximation to the diffuse field sound intensity.

Ultrasonic scale modeling typically uses models in scales from 1:4 to 1:16, although some researchers have shown useful results even with scales as small as 1:50. (Outdoor sound propagation is typically modeled ultrasonically using a scale of 1:100.) Using spark source excitation, one can study the impulse response of the room. Such a scale model is shown in Figure 6.18.

The static sound distribution can be calculated from the impulse response measurements or directly studied using, for example, jet noise sources.

The following rules apply to ultrasonic scale modeling at 1:m scale:

- The frequency scale is transposed by a factor of $m:1$.
- The time scale is transposed by a factor of $1:m$.
- The sound attenuation by air is reduced by a factor of $m:1$.
- The acoustic impedances of the materials used in the scale model must be similar to those in the real room (after appropriate frequency scaling).

One of the major problems in ultrasonic scale modeling is the correct scaling of the attenuation of sound by air. Another major problem can be found in the difficulties in designing sound sources and



Figure 6.18 Model of a church built for ultrasonic scale modeling. Note scale model of manikin for binaural sound pickup next to the hand. (Photo by Mendel Kleiner.)

microphones that can cover the frequency range of interest (typically 200 Hz to 200 kHz) with adequate directivity and sensitivity. One can approximate the correct scaling of sound attenuation by air, using dry air (typically having a relative humidity of less than 2%) or using a pure gas, such as nitrogen in the scale model. These methods have been shown to give good results not only regarding the simulation of reverberation time and levels but also regarding impulse response. Research has shown that it is possible to digitally compensate impulse response measurements made using scale models without these precautions, and still obtain quite good simulations.

A large advantage of physical ultrasonic scale modeling is that it is easy to simulate the acoustic conditions in irregularly shaped rooms, and in rooms with uneven sound absorption. In addition, ultrasonic scale modeling allows for inclusion of diffraction and scattering as well as the frequency-dependent sound absorption by various surfaces.

Using computer software, it is possible to predict the acoustic properties of a room in much more detail than by manual methods. Once the geometrical and acoustical data for the room have been entered into the computational model, one can calculate the way sound from one or several sources is incident at various listening positions. Usually one is interested in the spatial impulse response expressed as $p(t)$ relative to the sound pressure at 1 m distance from an omnidirectional source. If one is interested in auralization of the room, it is necessary to know the angles of incidence (φ , δ) for the various sound components contributing to the room impulse response. Using the room impulse response, it is possible to calculate the various room acoustic quality indices, as previously mentioned.

Most software for the prediction of room acoustics is based on sound propagation modeling using geometrical acoustics and, therefore, can only take diffraction, scattering, and real acoustic surface impedance into account in an approximate way. Some software will approximate the influence of these acoustic

phenomena on the propagation of sound and thus yield improved modeling. Most such software is based on ray tracing or mirror image modeling or possibly a hybrid between the two.

Modeling using the *ray tracing method* typically assumes a large number of rays (1000 to 1,000,000) being sent out from a sound source. The sound source can be assumed to be omnidirectional or directional. In the latter case, one can either assign different intensities to the rays or adjust the ray density to simulate the source directivity. One then follows each ray on its path and studies how the ray is reflected by the surfaces of the room and the objects it contains. Curved surfaces are generally simulated by piecewise flat surfaces. By studying the number of rays passing through a test surface or test volume (such as a sphere) at the listening position of interest, one can calculate the room impulse response at that position.

In most ray tracing models, the sound absorption is modeled by adjusting the ray intensity appropriately after calculation of the room impulse response. Scattering can be simulated by assigning various randomization rules to the angle of reflection. *Pyramid tracing* and *cone tracing* are variations on simple ray tracing.

Modeling of sound propagation using the *mirror image method* assumes that sound reflected by a flat, hard surface originates from a mirror source behind the surface. One needs to calculate the various mirror images created by the (flat) surfaces of the room and its objects. For a room having a complex

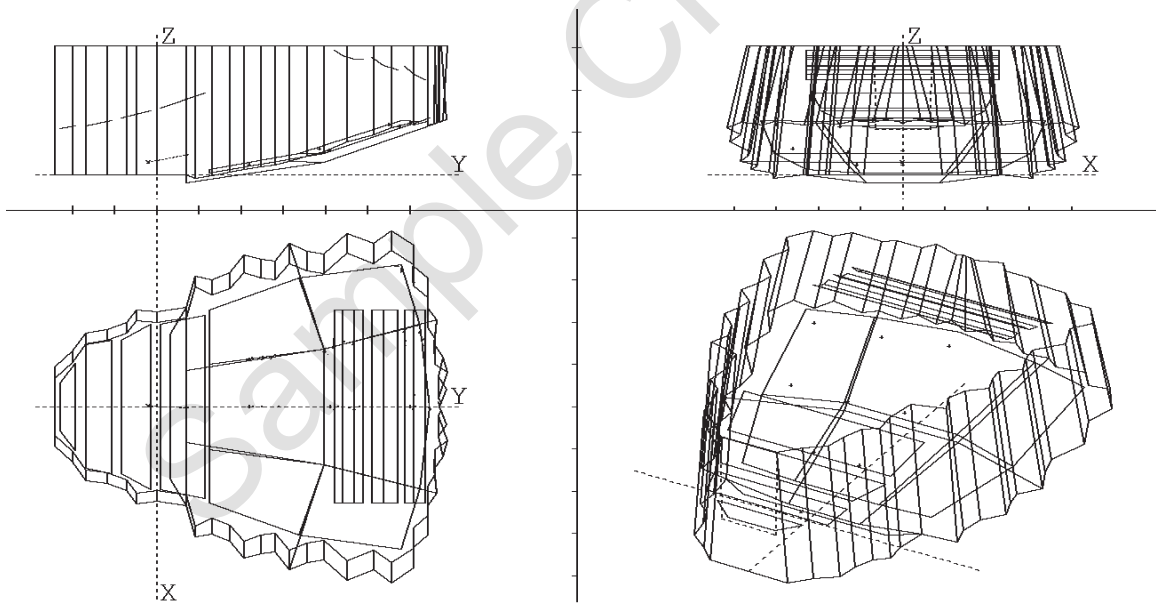


Figure 6.19 CAD using ray tracing or mirror image method-based software can be used to predict the acoustics (see also Figure 6.20). This model of the room shown in Figure 6.11 was done by Julie Byrne, MSc.

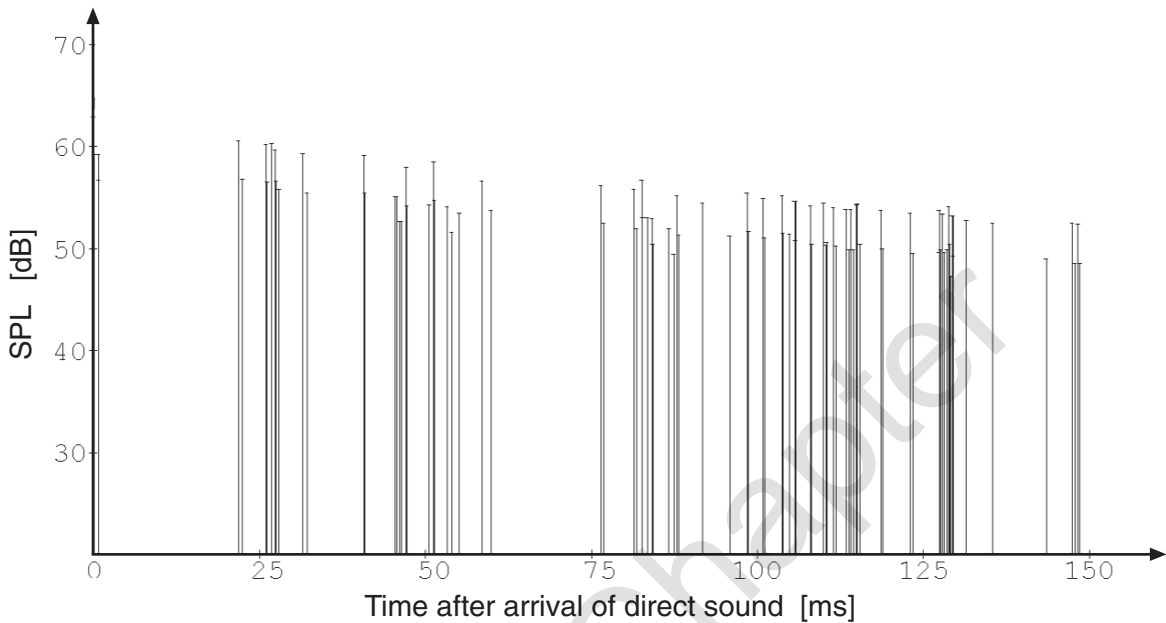


Figure 6.20 A calculated echogram for mid-stage to last row on center in the room model shown in Figure 6.19 (arbitrary reference level).

geometry, it may be necessary to calculate many millions of mirror sources, as sound is reflected multiple times by the various surfaces. The relevant sources must then be sorted out using rules for visibility. A disadvantage of the method is that it does not allow for easy inclusion of scattering and diffraction. Figure 6.20 shows the calculated echogram for the room shown in Figure 6.19 (set for not taking frequency-dependent sound absorption, scattering, or diffraction into account). One can see that the response consists of *discrete* reflections.

Hybrid models are typically based on the use of both ray tracing and mirror image modeling. Ray tracing is used initially to find the possible sound paths and then mirror image modeling is applied to find the exact sound paths. As with ray tracing, a disadvantage is the need for modeling the propagation of a large number of rays.

Both ray tracing and mirror image modeling are similarly computationally intensive. Neither method allows for exact room acoustic modeling since geometrical acoustics is used. Good modeling of absorption, diffraction, and scattering remains a problem. In any case the prediction will be limited by the experience and skill of the user. The late part of the room impulse response corresponding to the late part of the reverberation is usually too computationally intensive to be handled by a direct approach, and various approximations are generally used. Some of these may be based on the use of various statistical assumptions and random generated numbers. Such methods may lead to unevenness in the calculated full room impulse responses.

6.5 ELECTRONIC ARCHITECTURE

It is advantageous to be able to change the room acoustic properties of performance spaces to suit the type of performance, for example, various types of orchestras, music, and voice. This can be done using two approaches, *reverberation enhancement* and *sound field synthesis*, sometimes called *electronic architecture* (see References 6.4, 6.9, and 6.10).

Reverberation enhancement is typically used when the room acoustic conditions are reasonably good, but where reverberation time or level may need to be increased. It is often possible to increase the reverberation time by up to 50% with good results.

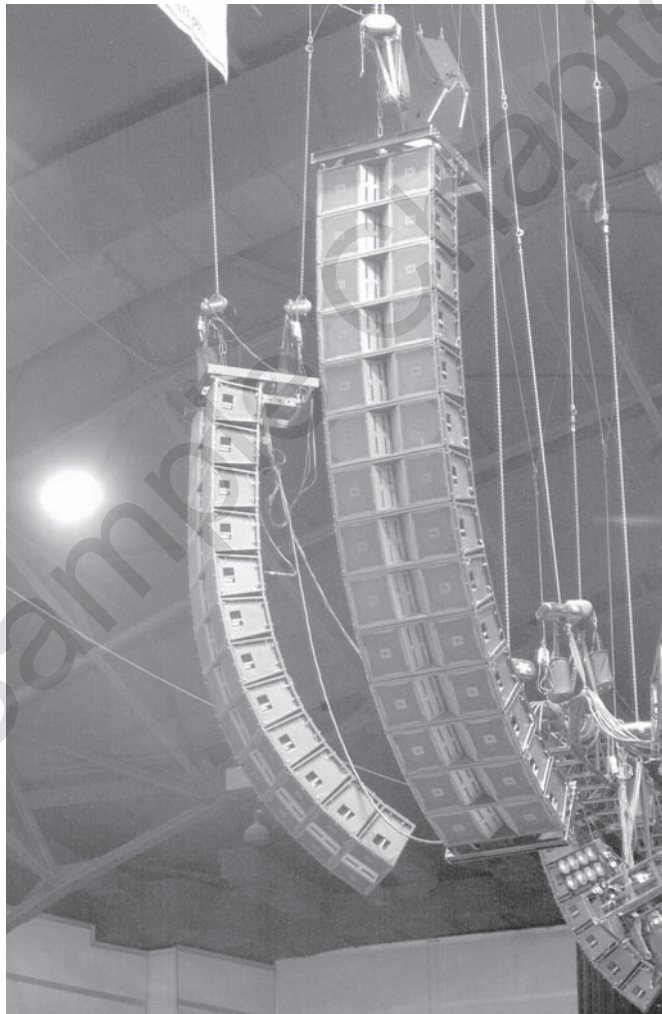


Figure 6.21 Sound reinforcement systems may use conventional direct-radiating or horn-coupled loudspeaker array elements to make up an array-type loudspeaker column to achieve the desired directivity (image courtesy JBL/Harman).

Sound field synthesis may include reverberation enhancement, but will usually include also the generation of specific early sound field components corresponding to early reflections by (imaginary) surfaces. Sound field synthesis is sometimes used to reduce the subjectively perceived reverberation time T_{subj} by reducing the relative level of the late reverberant sound. The conventional *public address system* is an example of primitive use of *sound reinforcement* to provide this effect, for example, in churches and other public meeting places (see Figure 6.21).

6.6 AURALIZATION

Audible sound field simulation is usually called *auralization*. Auralization of an environment (e.g., for speech and music) can be considered as analogous to visualization of that environment. The ambition is to make it possible to listen to the sounds in the environment without being in the actual environment (see Reference 6.8).

Computer auralization of an environment will start by entering data for the geometry of the room and the objects in the room into a computer database to be used by a mathematical model based on ray tracing or mirror image modeling. Once the software has calculated the response, it will generate a file containing a list of information about which sound components are incident at the desired listening positions:

- Levels relative to the direct sound for reflections
- Time delays relative to the direct sound for reflections
- Propagation path lengths

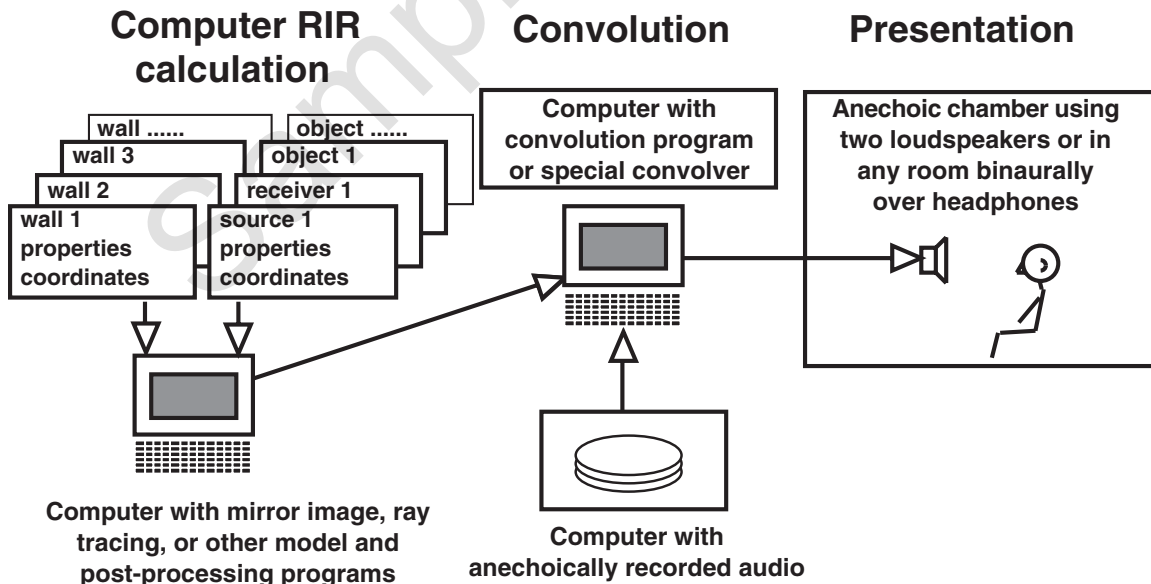


Figure 6.22 Functional diagram for fully computed auralization (After Ref. 6.8).

- Angles of radiation
- Angles of incidence

Using knowledge about the sound reflecting and scattering properties of the various surfaces, one can then add this information to the initially generated *generic* room impulse response (see Figure 6.22).

There are two main presentation methods used for auralization—binaural playback and playback by surround sound systems, such as *Ambisonics*. Binaural playback can be done either directly using headphones or by a stereo sound system using *crosstalk cancellation* (XTC). Surround sound systems need at least four loudspeakers—seven to ten loudspeakers are typically used.

In practice, XTC requires a quiet room in which the reverberation time is short, such as a semi-anechoic room. The aim is to provide the listener or listeners with sound signals at the ears corresponding to those that would be obtained when listening directly using headphones. This requires that the leakage of sound between the ears due to the loudspeaker presentation is cancelled using cancellation signals generated using knowledge about the transfer paths between the loudspeakers and the ears. Presentation using crosstalk cancellation generally gives better frontal localization than ordinary headphone presentation.

For auralization to be presented binaurally, information about the way the head influences the incoming sound waves must be added. Because of the way sound is reflected off the human body, the sound pressures at the ears are going to depend on the angles of incidence and on the frequency. At low frequencies, below 100 Hz, the effects will be negligible but already at approximately 250 Hz, the torso will contribute to the sound pressure. At frequencies above 5 kHz, there will be major effects as shown in Chapter 3 because of the reflection and diffraction of sound by the head. The frequency responses of the sound pressure at the left and right ears relative to the pressure in absence of the body are called the *head-related transfer functions* (HRTFs). Once the head-related impulse responses (HRIR) have been obtained, they can be converted into time domain signals as *head-related impulse responses*. Free libraries of HRTFs and HRIRs may be found on the Internet for various angular resolutions.

Figure 6.23 The Bose Auditioner® system is a commercial system for auralization using CAD. The binaural room impulse responses are convolved with music or speech so that it is possible to listen to a simulation of what it would be like to listen in the actual room being simulated. Crosstalk cancellation is used to render the binaural signals in this setup (photo courtesy of Bose Corporation).



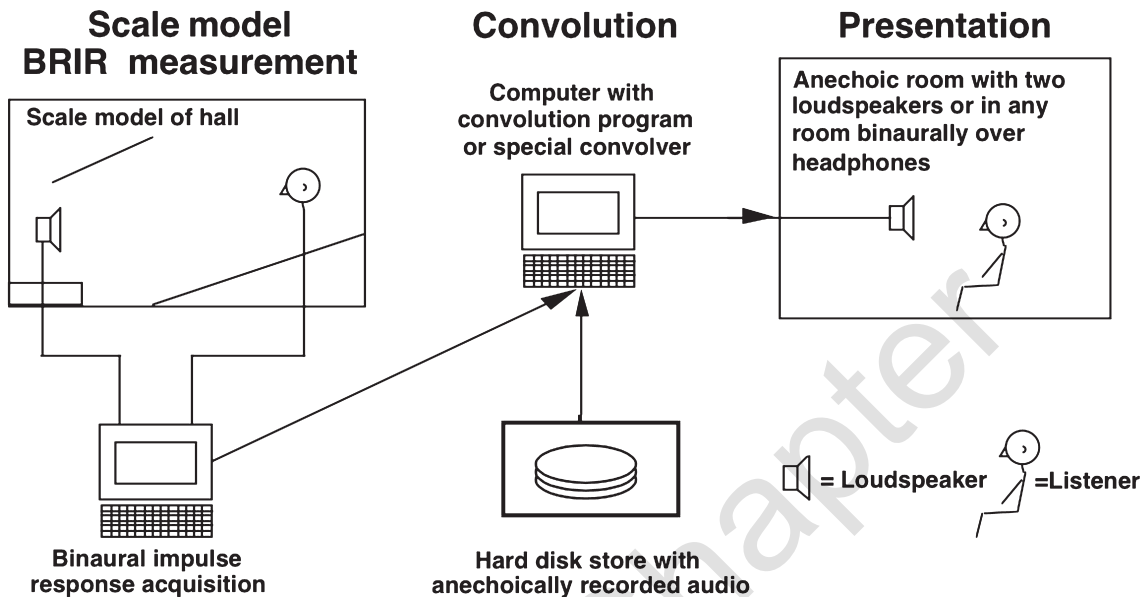


Figure 6.24 Functional diagram for auralization using the indirect ultrasonic scale modeling method (After Ref. 6.8).

Once the *generic spatial impulse responses* for the positions of the source and receiver in the room have been found, these can be appropriately *convolved* with the HRIRs and with the impulse responses of the sound-reflecting surfaces—the impulse responses of the various reflection coefficients described in Chapter 2. This yields the final binaural room impulse responses.

The binaural room impulse responses can then be convolved with *anechoically recorded* speech and music so that one can listen to what the sound of the room would be. The convolution can be done using software in a general purpose computer or by a dedicated *hardware convolver* (digital filter).

The convolved binaural signals can then be played back over headphones or further processed by a crosstalk cancellation network to allow for loudspeaker presentation. (See Figure 6.23.)

Auralization can also be done using a scale modeling approach. Direct modeling where the audio signals are played back as well as recorded in the model is quite problematic and an indirect approach where the binaural room impulse response is measured digitally and is subject to suitable digital signal processing is favored. The principle of this approach is shown in Figure 6.24. The main reasons for this is the possibility for a better S/N ratio and lower nonlinear distortion than using the direct method. In the indirect approach, the main difficulties lie in obtaining a suitably scaled manikin with torso, head, and ears, all correctly modeled to scale. Once the binaural impulse responses have been obtained, these are then convolved with the desired anechoic audio signal as in the case of fully computerized auralization described initially.

The advantages of auralization are many. By listening to auralization of rooms, costly mistakes in planning and building can be avoided, and different wall treatments and seating arrangements can be investigated. Auralization can also be used to study the properties of sound reinforcement systems, to reduce

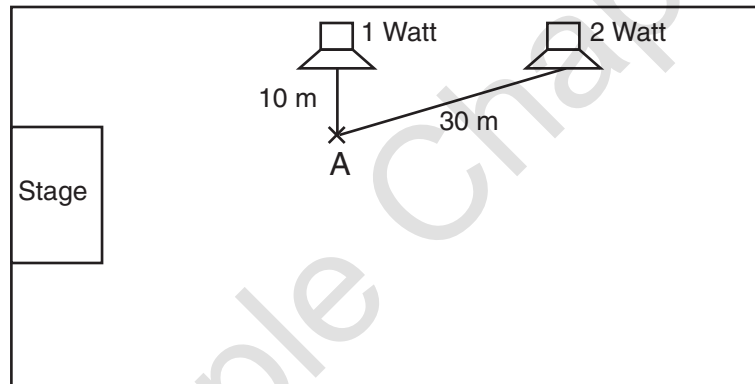
tuning time at building completion. Along with its use in architecture, auralization is used in various forms of virtual reality rendering, games, information systems, among others.

6.7 PROBLEMS

- 6.1 Two loudspeakers in an auditorium are connected to the same signal source but are radiating different sound power.

Task: Determine, using the figure on page 162 and Figure 5.11, if the sound at listening position A will be considered as good or poor, since one of the sounds will arrive delayed relative to the other.

Hint: Only consider the direct sound components from the loudspeakers and assume that the loudspeakers are omnidirectional.

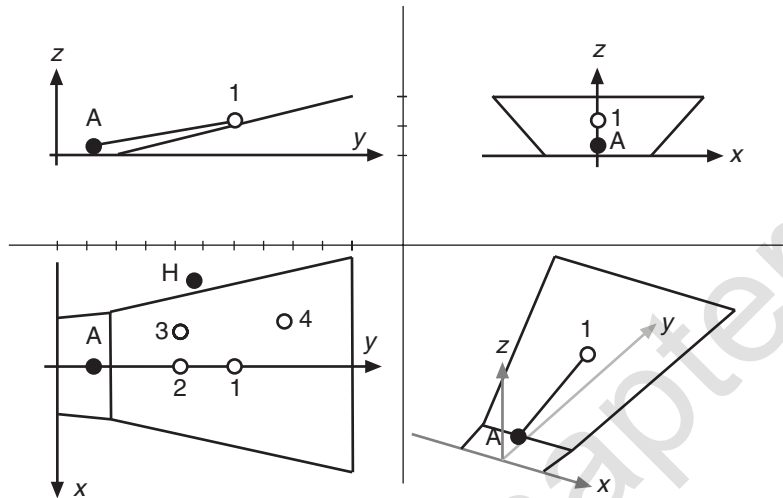


- 6.2 One wants to use a loudspeaker to improve the acoustical conditions in an outdoor theatre. Assume that the sound source is omnidirectional and all surfaces are sound absorptive. Use the figure to determine the required loudspeaker playback levels. The distances from loudspeaker H and actor A to the respective listening positions are:

	Pos. 1	Pos. 2	Pos. 3	Pos. 4	
Distance to actor	24	14	16	33	[m]
Distance to loudspeaker	15	14	9	16	[m]

The distance between the markers in the figure is 5 m.

Task: Calculate the necessary delay times for the four positions in the figure to avoid echoes.



- 6.3 One wants to use a loudspeaker to improve the listening conditions in a concert hall and mounts a loudspeaker above the stage. The direct sound level of the loudspeaker is adjusted to be 6 dB above that of the speaker in the direction of interest. The volume of the hall is $40,000 \text{ m}^3$ and the reverberation time T is 1.7 s. Assume the sound field in the reverberant field to be diffuse and that sound field components can be added on an intensity basis. Further, assume that the delay between the direct sound components from the speaker and loudspeaker is less than 0.010 s and that the directivity index of the loudspeaker in the direction of interest is 6 dB.

The directivity D is defined as the sound level at a certain distance from a directional sound source relative to the sound level that would be obtained from an omnidirectional sound source radiating the same power. The directivity index (DI) is defined as $DI = 10 \log(D)$.

Task: Calculate the level difference between direct sound and reverberant sound $H = L_{p, \text{direct}} - L_{p, \text{reverb}}$. Find the approximate subjective reverberation time T_{subj} for various seats at distances of 10, 20, and 30 m respectively from the speaker or loudspeaker with and without the loudspeaker active. Use Figures 5.10, 5.11 and 5.21 in making your decision.