Chapter 1
Principles of Acoustics and Hearing

What is a sound wave? A visual interpretation can be obtained by imagining the membrane of a speaker that moves back and forth. Air molecules immediately next to the speaker will co-oscillate with the speaker. These air molecules will push and pull their neighbors, who in turn will push and pull their neighbors, and so on, forming a longitudinal wave of oscillating molecules. This is in essence a sound wave: tiny pockets of air, oscillating around an equilibrium position, causing small air pressure variations imposed on the static air pressure. A sound wave travels with a speed of 343 m/s in air at 20 °C whereas in solids and fluids the speed is faster and the total distance it can travel is longer, mainly due to the higher density of molecules.

The magnitude of the wave determines the amplitude of the sound given by a pressure maximum and a pressure minimum as graphically interpreted in Fig. 1.1. Because the human ear is capable of detecting sounds levels ranging from 1 unit to sounds one million times louder, the *decibel* (dB) scale has been introduced. Named after Alexander Graham Bell (1877–1922) the decibel is a logarithmic unit and overcomes thereby the handling of very large numbers. The decibel is used to measure sound levels and of course therefore also sound level differences.

The quietest sound that the human ear can detect is about 0 dB and the least noticeable difference is about 1 dB. Doubling the number of bass players in a symphonic orchestra will increase the sound level by 3 dB, however, a double brick wall will cause an isolation of the sound pressure of approximately 55 dB. The sound level in the open drops by 6 dB per doubling of distance. A symphony orchestra produces only approximately 2.5 W of acoustic power when playing *fortissimo* (very loud). The dynamic range of a symphony orchestra is as much as 70 dB, that of a pop and rock band considerably less, whereas jazz performances can span over a large dynamic interval from *ppp* to *fff*, in musical terms, as well.

The decibel is hence based on a ratio and is therefore not a unit like a meter or a watt. In order to express the absolute value of a sound pressure the ratio is taken between a given level and a reference level being 2.10^{-7} mbar or 20 μPa corresponding roughly to the lowest audible sound at 1,000 Hz.
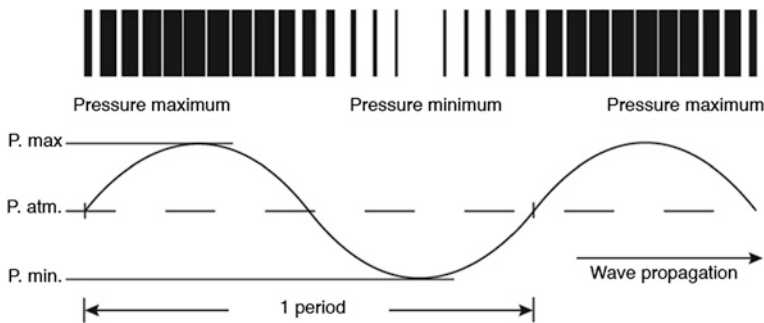
Fig. 1.1 Graphic representation of a pure tone. One period equals the wavelength of the tone

Frequency, *f,* and wavelength, λ, are two parameters expressing a wave. In, for instance, musical tones, the wavelength equals one period of the sound wave implying that the shape of the wave is periodic. The distance between pressure maxima (or minima) in the sound wave is the wavelength. The wave travels with the speed of sound *c,* and the three parameters interrelate as expressed by the equation

$$c = \lambda f \tag{1.1}$$

which is in good agreement with one's intuitive understanding: the shorter the wave, the more often will one pass a certain point in as much as all waves pass with a constant speed, the speed of sound. The unit of frequency is s^{-1} (cycles/s) which is written as Hertz, Hz, named after Heinrich Rudolf Hertz (1857–1894). The wavelength unit is meters.

A pure tone has a single frequency associated with it. Musical instruments as well as the human voice, however, produce a number of overtones creating a unique sound where the lowest tone normally determines the pitch. A doubling of frequency corresponds to an octave in musical terms. Acoustic measurements are traditionally often made over octave intervals with center frequencies being: 63, 125, 250, 500, 1,000 Hz, and so on.

Many instruments depend on *resonance* in order to create their sound. The length and diameter of a flute, the volume of a drum and the tension of its head, and the body volume of an acoustic guitar are examples of resonators and ways to achieve a particular resonance frequency. But in amplified music, electric instruments are often used and the vibration of a string on an electric bass or guitar, for example, is picked up and amplified by the use of a transducer and amplifiers. The lowest note is called the *fundamental frequency* (Fig. 1.2) and the overtones are referred to as *harmonics* of the instrument. The A above middle C on a piano is often tuned to 440 Hz and the lowest E string on an acoustic or electric bass is about 41 Hz. The shapes of such complex waves are still close to periodic. The overtones that create the spectrum of the instrument and thereby its characteristic sound together with the sound created where the hand, stick, bow, or another tool touches the instrument, in many cases extend all the way to the highest sound possible for the human ear to hear above 15–20 kHz.

Fig. 1.2 Fundamental frequencies of different instruments used in pop and rock music
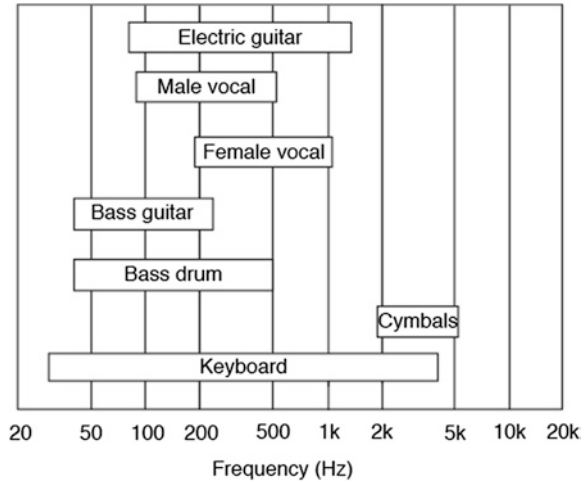
Sound Propagation

Sound rarely propagates in a uniform pattern away from the sound source. The pattern by which sound is distributed is dependent on, among other things, the frequency produced. Such patterns can be drawn for each frequency band and are called directivity patterns.

Sound propagates in its simplest form outdoor in the open, where the waves are not affected by any obstacles. Without reflections from walls or ceiling the unamplified human voice is capable of reaching approximately at least 50 m, which is also the longest distance from a performer to a listener in, for instance, the ancient Greek theatres. In these theatres, though, the proscenium walls behind the performers reflect the part of the performers' sound that propagates away from the audience to be redirected towards the audience permitting a higher sound level for the performer and audience.

Indoors, the level of the direct sound of course decreases farther away from the sound source. From pop and rock concerts, most of us have experienced being placed behind taller audience members standing in front of us blocking us visually. Actually we are also blocked from the direct sound if the placement of the sound system does not compensate for this by being placed high enough or if the floor is not sloped.

Only frequencies with wavelengths comparable to the size of the obstacle in front or smaller are blocked because large wavelengths bend around obstacles of inferior size. This phenomenon is called diffraction (Fig. 1.3). The obstacle is simply being surrounded by lower frequency sound. Likewise, it takes a very large and heavy object such as a concrete wall to reflect a low-frequency sound precisely. When the reflecting surface is large compared to the wavelength of an incoming sound wave, the sound will be reflected just like light in a mirror. This is called *specular reflection*.

Fig. 1.3 Sound waves with
larger wavelengths bend
around smaller objects and
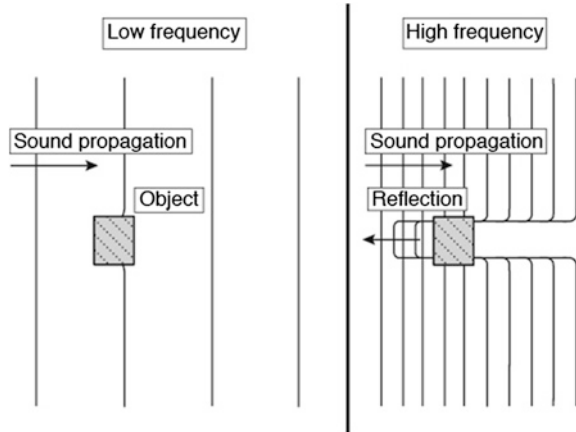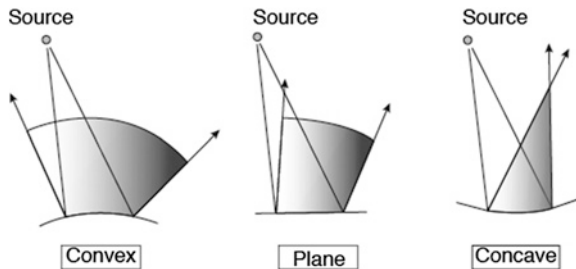higher frequency sounds are
blocked

Fig. 1.4 The reflected sound
will be affected differently
depending on the shape of the
boundary the incoming sound
wave meets

As light spreads scattered by a matte white piece of paper, sound can also be
scattered more or less universally out in all directions from the scattering surface.
The term *diffusing surface* is also used to account for the same effect and should
not be confused with the diffuse sound field discussed later. Scattering is a very
important property that is used extensively in auditorium design.

Conversely, black matte paper will not let any light get away, and in the same
way, for instance, a thick layer of mineral wool will absorb incoming sound. Any
building material, and the way it is mounted, will determine how much sound it
absorbs and reflects and at which frequencies it does so.

Sound is reflected in different ways depending on its frequency and the
shape of the boundary it meets. At mid and high frequencies sound waves can
be considered as rays or beams, and the incoming angle of the wave is equal to
the angle at which it bounces away from the boundary, very similar to a billiard
ball. This situation with a plane boundary is sketched in the middle drawing of
Fig. 1.4. If the surface is convex, however, the beam will be scattered, whereas
a concave surface will lead to a situation where the sound is reflected from sev-
eral places on the surface back to the listener's ear with small time differences.
This causes a so-called focusing effect, highly disruptive for the intelligibility
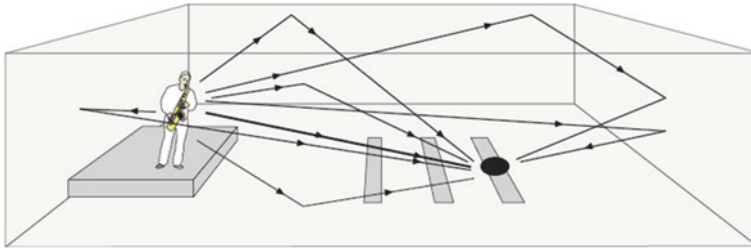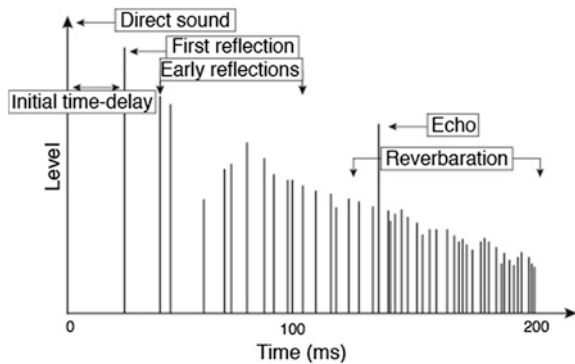of the sound.

Fig. 1.5 Direct sound and first reflections from various surfaces in a room

Fig. 1.6 Reflectogram indicating how sound reflections soon multiply into reverberation

Sound in Rooms

Soon after the direct sound arrives at the listener, a series of early reflections follows from the side walls, ceiling, and so on (Fig. 1.5). This sound evidently will arrive later at the listener because it travels longer. The initial sound wave can spread in many directions; it hits numerous surfaces in the room and soon multiplies into hundreds of reflections. The later sound arriving after about 0.1 s (depending on hall size, etc.) is called the *reverberant sound* or simply *reverberation*. The characteristic sound at a given location in any hall is thus a result of thousands of reflections and where, at which frequencies, and to what degree they are reflected, scattered, or absorbed. Luckily the human hearing system is capable of distinguishing subtle differences and acknowledging the unique quality of different halls.

A diagram showing the amplitude of reflections versus time is called the *impulse response* or a *reflectogram* (Fig. 1.6). The term "impulse response" can also mean an actual recording of the room's response to a sound.

Any enclosed space has normal modes given by its dimensions. These modes are also referred to as the *standing waves* of the room when it is exposed to sound. The room modes are responsible for there being places in the room where a given frequency is louder than in other places of the room.

For a rectangular room with dimensions l_x, l_y, l_z it can be shown that the modes, also called *natural frequencies*, are

$$f_{l,m,n} = \frac{c}{2}\left[\left(\frac{\ell}{l_x}\right)^2 + \left(\frac{m}{l_y}\right)^2 + \left(\frac{n}{l_z}\right)^2\right]^{\frac{1}{2}} \tag{1.2}$$

where l, m, and n are integers that indicate the number of nodal planes perpendicular to the x-, y-, and z-axes, the three dimensions of the room. They are, for example, [1, 0, 1] or [7, 4, 3] or any other integers. The important point here is that the natural frequencies are given by the room dimensions. This is also true in odd-sized rooms whose natural frequencies can be calculated by use of computer programs.

It can be proven that the density of natural frequencies is proportional to the frequency squared, f^2, indicating that at higher frequencies the density is high, in fact so high that the modes can be regarded as a continuum. Likewise the size of the room plays a role and in small rooms the modes are very separated indeed at low frequencies, leading to severe colorations of the sound at certain frequencies. Such colorations will be even bigger if the room dimensions are a low integer multiple of one another because the modal frequencies will then coincide. So at higher frequencies, the sound energy at a given frequency is most likely more uniformly distributed in the room. It is said that the sound field is more diffuse at these higher frequencies.

In any room with a continuous (stationary) sound source, the sound field will be formed by standing waves. The decay of sound "from" the standing waves at the time right after a sound source has stopped, is one feature of reverberation.

But reverberation can occur without standing waves; think of a hand clap in a room. There will be reverberation because of hundreds of echoes repeated soon after one another, but standing waves will not have had time to build up. The same thing happens in a forest: a sudden percussive sound creates a reverb because of reflections from the many trees, but the sound waves would propagate or "run," probably not "stand".

It is easy to understand intuitively that the more the sound reflections hit some sound-absorptive surface, the faster the sound dies out in a given room. And likewise: the bigger the room is, the longer a reverberation. The time it takes for the sound in a room to attenuate 60 dB is called the *reverberation time* of the room. Despite many advances in acoustics over the latest decades this remains the single most important parameter in room acoustics. It is essential to note that the reverberation time is usually not the same in different frequency bands for a given room. It can be shown that the time T, in seconds, it takes for the sound to attenuate by 60 dB in a perfectly diffuse room proposed by Sabine (Wallace Clement Sabine 1868–1919), in a given frequency band, is approximately

$$T_{60} \approx 0.16\frac{V}{A} \tag{1.3}$$

This equation is called Sabine's formula and is rightfully the most well-known relation of parameters in room acoustics. In reality, especially smaller rooms are

Fig. 1.7 Graphic representation of reverberation time (RT) and T_{30}
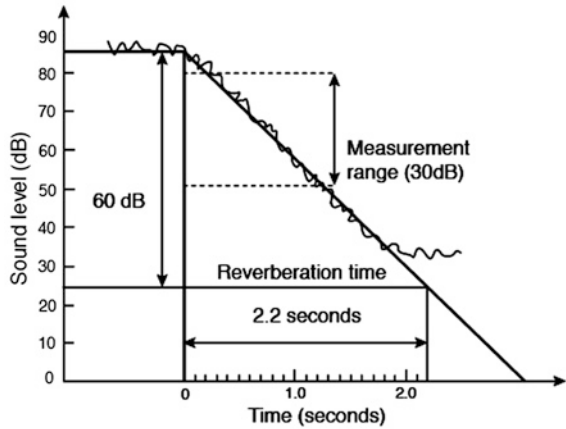
not perfectly diffuse and certainly not in all frequency bands. V is the volume of the hall in cubic meters and A is the absorption area, meaning the number of equivalent square meters of 100 % sound-absorbing material in a given frequency band. The reverberation time is sometimes, also in this book, denoted simply T or RT.

When actually measuring the reverberation time of a room usually only the sound from –5 to –35 dB or even –25 dB is measured because background noise usually makes it difficult actually to measure anything as silent as 60 dB below the sound source measurement signal. The time this 30 or 20 dB decay takes is then multiplied by 2 or 3, respectively (most often, this multiplication is performed by the measurement device itself), assuming that the decay is linear over time so that the decay of the first some 30 dB takes the same time as the latter 30 dB. This measured reverberation time is called T_{30} (Fig. 1.7) or T_{20}. One can get a rough estimate of the reverberation time in a room at approximately 2 kHz by clapping one's hands once and counting the number of seconds it takes for that sound to die out. Low-frequency reverberation time can be estimated by producing a similar low-frequency sound. Of course, such tests are not in compliance with any standards for measurement of reverberation time where an exact measurement is made in every octave, or third octave band, for instance by firing a gun, an abrupt loud signal of noise, or measuring the decay of a sine wave sweeping through frequencies.

This reverberation time based on a long decay is also called the terminal reverberation time because it denotes a sound and the decay to more or less complete inaudibility. This seldom being the case during a concert (mostly only at the last note of a song), there is also another parameter describing sound decay called the *running reverberation time*. This takes into account only the first 10 dB of decay and to make it comparable to the terminal reverberation time, this time is multiplied by 6. If the decay slope is strictly linear from 0 to –60 dB the two reverberation times are identical, but this is seldom the case. The running reverberation time is denoted EDT—early decay time—and was suggested in 1968 by Vilhelm
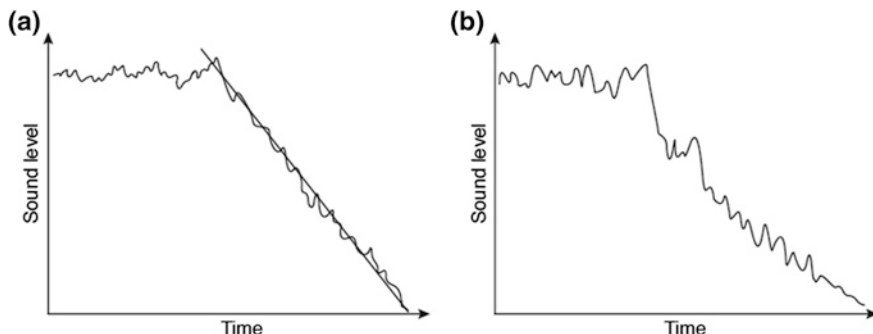
Fig. 1.8 A linear (**a**) and a nonlinear (**b**) sound decay in a room. This explains why T_{30} and EDT can be different

Lassen Jordan (1909–1982). The EDT is heavily dependent on the position in the hall where it is measured. As shown later, a given volume of hall for a given purpose has an ideal associated terminal reverberation time. Bigger halls allow, as mentioned, longer reverberation. The sound decay is often nonlinear and EDT is often a more relevant parameter to describe the reverberation actually experienced by a listener in a given position of a room. See Fig. 1.8.

The so-called Schröder frequency suggested in 1962 by Manfred R. Schröder (1926–2009) denotes the frequency below which the sound field can be regarded as being dominated by discrete standing waves

$$fs = 2{,}000\sqrt{\frac{T}{V}} \tag{1.4}$$

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

There is no sharp line defining above which frequency the sound field can be regarded as diffuse. Below this frequency, on the other hand, as discussed above, the modes will not be close to one another, thus standing waves at single frequencies will occur. In either case reverberation will occur at least as the decay of sound "from" the standing waves. As seen in Eq. (1.4), in bigger rooms the Schröder frequency is lower. In smaller rooms, measurements of low-frequency reverberation time can be very dependent on the position where it is measured because of the modes; therefore many measurements at different positions are averaged.

The halls investigated in this book are in almost all cases over some 1,000 m^3. With an ideal reverberation time for pop and rock music of 0.6 s for that size volume (see Chap. 5), this yields a Schröder frequency f_s below about 50 Hz, close to the lower extremity of the 63-Hz octave band. Although a lot of halls suffer from a long reverberation time at low frequencies moving f_s upwards, for the purposes of this book, sound even in the 63-Hz octave band is generally regarded as diffuse. This is

important to stress, for instance, for sound engineers who often think of room acoustics solely in terms of standing waves. For studios and very small venues such as clubs and cafés with, for example, a low ceiling, it is certainly correct that distinguished annoying modes exist at lower frequencies. For mid-sized or larger halls this is not an essential point; the density of the natural frequencies is high enough even for low tones to be regarded as a diffuse sound field that decays over time as reverberation.

The relation between sound-level difference and reverberation time difference under the diffuse sound field is

$$\Delta L_\mathrm{p} = 10 \log \frac{T1}{T2} \tag{1.5}$$

According to this formula doubling the reverberation time in a given frequency band compared to others, results in $10 \cdot \log 2 \approx 3$ dB extra response from that frequency band. A common misconception is the belief that when that frequency band has been lowered level wise, for instance on an equalizer, then the room acoustics have been corrected! This is completely incorrect because the reverberation time at that frequency is still the same. Level adjustments work in the level domain, not in the time domain. As a matter of fact the reverberation of a given room size has to be not too long and usually not too short either for a given purpose. Furthermore, musicians and sound engineers adjust their playing and their levels so that it conforms to the hall and the situation. Thereby the level difference that the possibly uneven reverberation times at different frequencies impose is evened out, still leaving the clarity of the sound, and thereby the overall sound quality, to be affected. As we show, a too-high reverberation value in a given frequency band will cause an undefined blurred sound in that frequency domain and often that is enough to damage the overall sound in a venue.

Acoustic consultants use this connection between reverberation and sound level as means to enhance, or acoustically amplify, lower frequency sound in halls for acoustical music, for example, or inversely, to reduce noise in a kindergarten or factory. In pop and rock venues, the amplification aspect of reverberation is close to uninteresting inasmuch as all levels are adjusted by the turn of a knob either by the musician or by the sound engineer. For amplified music the reverberation solely affects the sound by either washing away the core message in the music (when it is too long) or by making the performance seem dull and lacking liveliness (when too short) as discussed in Chap. 5.

Human Hearing

A sound source emits sound in many directions and the reverberation consists of thousands of reflections from many surfaces in the room (the surfaces that are not 100 % sound absorbing at all frequencies). In a concert hall for symphonic music there are approximately 8,000 reflections from one single note during the first second. Of course in smaller and less reverberant halls there are fewer reflections, but still

Fig. 1.9 Equal loudness
contours at various levels.
Different frequencies
need different levels to be
perceived equally loud for
humans. The contours vary
with sound level. *Lowest
curve* is the hearing threshold

Equal-loudness contours (red) (from ISO 226:2003 revision)
Original ISO standard shown (blue) for 40-phons
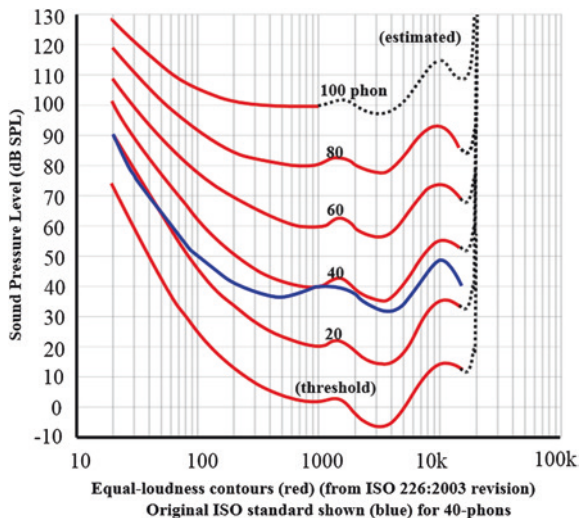
each one has a delay, a direction, and a sound level associated with it. The ear is
highly selective in interpreting this abundant information. The direct sound is pre-
cisely localized by the ear: even when the sum of later reflections is louder than
the first direct sound, the ear is capable of detecting the direction from where it
originates. The direct sound is usually not fully masked by louder later sound.

The term masking is used when the presence of a given sound *A* makes another
sound *B* inaudible (fully masking) or more or less difficult to hear (partially mask-
ing). *A* masks *B* or *B* is masked by *A*: a person needs to turn down the television in
order to understand what is being said on the phone. Generally a single note masks
more towards the upper frequencies of other tones than to the lower. In addition,
the louder the masking tone is, the broader a frequency range it masks. The term
forward masking is used to describe the phenomenon that a loud sound signal can
mask another weaker signal which is presented up to 200 ms after the first signal
stops. The opposite also is possible and is called *backwards masking* because it
goes back in time and the effect is restricted to 20 ms before the start of the strong
signal when loud enough compared to the weak one.

Perceived strength of a sound is called loudness. The unit is son or sone. A
short sound with the same sound level as a longer sound is perceived as less loud.
A 5-ms sound must be 25 dB louder for the loudness to be the same as a 60-dB
sound of 200-ms duration. The ear is most sensitive sounds at different frequencies
but at the same sound level don't occur equally loud to the ear.

The ear is most sensitive around 3.4 kHz (the approximate resonance frequency
of the typical ear canal) whereas our hearing rolls off at higher and certainly at
lower frequencies below about 100 Hz although less at higher sound levels. This is
one reason why at pop and rock concerts so much electric power is used to amplify
the low frequencies for them to be perceived equally loud as higher frequencies.
The graph showing this is referred to as the equal loudness contours (Fig. 1.9).

Fig. 1.10 The human ear can hear sounds that are louder than the "threshold in quiet" *curve*. One sound, the "masker" casts a shadow at frequencies both higher and lower than itself that prevents other sounds, "masked sounds", to be audible if they are at a lower level than the masking threshold
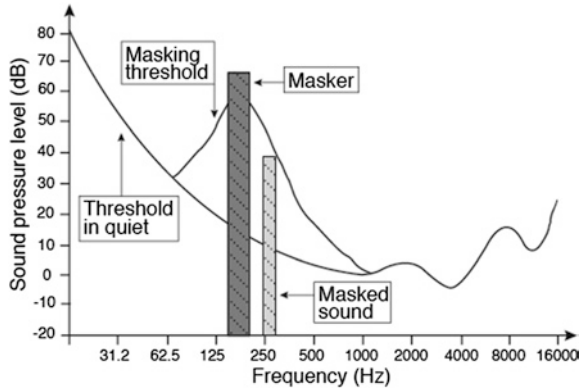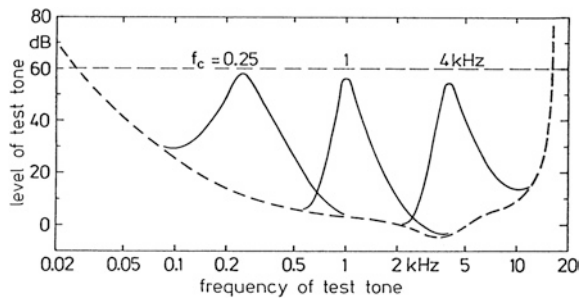
Fig. 1.11 The *curves* show the masking thresholds when the ear is exposed to about 60 dB narrowband noise at 250 Hz, 1 and 4 kHz (Zwicker and Fastl)

When regarding the figure it is seen that a sound decay from, for instance, 100 dB at 50 Hz gets inaudible quicker than a decay at, for instance, 200 Hz because the difference in dB from the perceived 100 dB curve to the hearing curve threshold is smaller at 50 Hz (given that the reverberation time of the room is close to being the same at those two frequencies). See Figs. 1.10, 1.11, 1.12 and 1.13.

Acoustic Defects: Echo

If a reflection is delayed by more than 50 ms compared to the direct sound, the ear will perceive that reflection as a distinct echo if it is louder than adjacent reflections that arrive before or after. These other reflections could otherwise mask it. An echo is always unacceptable in a music hall and must be eliminated. Echoes are more apt to occur in halls with short reverberation times because there will be few other reflections to mask them.

Flutter echoes are a series of rapidly repeated echoes mainly by parallel planar surfaces. Whereas *distinct echoes* involve long path lengths, flutter echoes are likely to occur at shorter distances such as between parallel walls on a stage, particularly if not many obstructive elements are present.
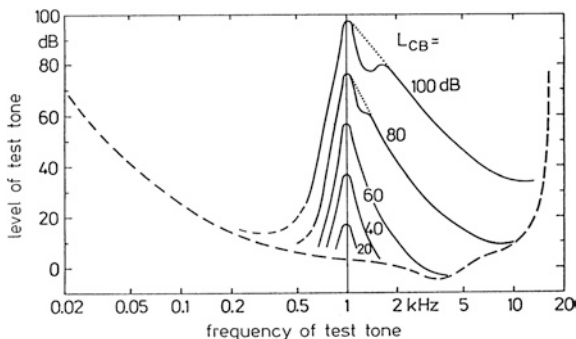
Fig. 1.12 The influence of level on the masking threshold. The masking effect gets broader with increasing level. Masking thus increases nonlinearly with level (Zwicker and Fastl)
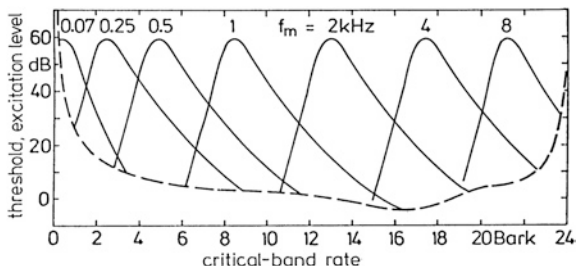

Fig. 1.13 The broadness of the masking curve changes with frequency. The slope at 70 Hz is steeper than those at higher frequencies indicating that sounds in the 63-Hz band do not mask as broadly as higher frequency sound. Here, instead of the octave band, the Bark scale is used. This scale is used in psychoacoustics instead of octave bands (Zwicker and Fastl)

Scattering

Flutter as well as distinct echoes can be avoided by applying a diffusive or angled structure on surfaces where the specular reflection originates. In (smaller) rooms with little or no obstructive elements the standing waves are not taken apart whereby the amplitude of the standing waves and thereby colorations of the sound can reach higher levels. The same goes for rooms with parallel surfaces especially, of course, if the room dimensions are low integer multiples of one another. These waves should be broken up to avoid colorations. An analogy to waves in water can be made: a pier or mole forms a breakwater that will crush the wave. But obstructive elements in a room of a size comparable to the wavelength of the wave do indeed break the wave so that a more uniform distribution of sound energy of the tone becomes present in the room. The sound field is scattered. The big peaks and dips are somewhat evened out so that one will encounter fewer level differences throughout the room. The reverberation time of a room with uneven distribution of sound absorption is greatly reduced when the room is made more diffuse.

The depth and the overall dimensions of scattering elements, also referred to as diffusers, are decisive for which frequencies are scattered. For example, many living rooms are highly diffusive (diffuse) because of the presence of lots of small and large elements randomly placed in the room. Schröder was behind a major development in diffusive elements (1979) where he proposed a series of diffusers whose properties could be calculated in advance. The pioneering work of exploiting diffusers has since been carried forward by Marshall, Cox, d'Antonio, Konnert, and others. Vorländer and Mommertz have made progress in actually measuring scattering coefficients in the laboratory.

In mid-sized and bigger halls (which this book is mainly about) additional scattering elements are not a necessity because of the high diffusivity due to the absence of isolated standing waves in large enclosures. Also an audience scatters the sound profoundly, although not as much at lower frequencies. The stage surrounding may, however, benefit from well-designed diffusive surfaces that eliminate flutter echoes and help to distribute the sound energy on stage more evenly.

Acoustic Vocal Sound

Niels W. Adelman-Larsen and Jens Jørgen Dammerud

The vocal is often a primary instrument at pop and rock performances and without clear comprehensible lyrics a lot of the message from the band to the audience is lost. Understanding what is actually being sung is related to the term "speech intelligibility". Much research is carried out on speech intelligibility. Some of the results from this research show that the mid- to high-frequency sound is the most important for enabling understanding of what is actually being said. Consonants are important and they have most of their sound energy within this frequency range. A telephone transmits almost exclusively 300–3,000-Hz sound. And yet we understand very well what's being said. It is also within this frequency range that our hearing is most sensitive. At low levels our auditory system is not very sensitive to low-frequency sound. The low-frequency sound is not critical for getting the message. It will in most cases actually be more of the opposite: low-frequency sounds will often mask mid-frequency sound, the low-frequency sound being a masker for the other sound components, and just make communication more difficult. Figure 1.14 shows this effect. The louder the low-frequency sounds are, the broader a shadow is cast.

Of course a vocal consisting only of 300–3,000-Hz sound doesn't sound very natural. A male speaker has his fundamental sound component usually within the 125-Hz octave band, whereas female speech is usually within the 250-Hz octave band. In singing, the fundamental tones are also often within these ranges.

Within the 125- and 250-Hz octave bands a speaker or singer is close to emitting sound *omnidirectionally*, meaning that the sound propagates at close to equal levels in all directions away from the person. At higher frequencies the emitted sound is much more focused towards the front. This means that the

Fig. 1.14 Low-frequency
content of the voice can
partially mask higher
frequency content. *Vertical*:
level; *horizontal*: frequency
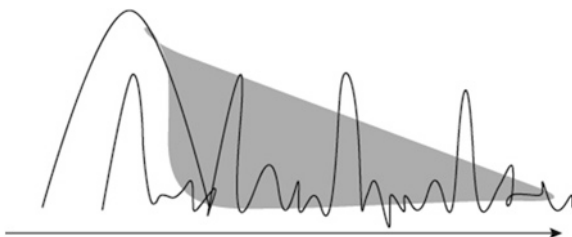
total emitted sound in all directions is usually much higher at low frequencies. Especially in an open space without reflections the directivity of the speaker is relevant with regard to the orientation of the speaker. If the person is faced away, the acoustic direct sound level within 300–3,000 Hz is significantly reduced.

What effect does the room have on speech intelligibility? For instance, in small meeting rooms the bass level due to standing waves can be so high compared to other frequencies that intelligibility is low even close to the speaker. In some cases the intelligibility is not too bad, but the low-frequency boost alone can be very disturbing. As mentioned, speech can be intelligible at a distance of 50 m in an open free field. For small rooms standing waves due to room modes will often be significant and these room modes can contribute to rather extreme acoustic gain of the speech, typically within the 125-Hz octave bands, but also within the 250-Hz octave band in very small rooms. In larger rooms where the speaker is amplified by electroacoustic means, often speech intelligibility is low if the low-frequency reverberation time is long and if the sound engineer does not lower the low-frequency content on the sound system. The long RT at low frequencies in the hall boosts the low frequency sound but it also masks the next syllable in the sentence of the speaker. The amplification is brought down to resemble the speakers' actual voice with the help of an equalizer (EQ). If the reverberant bass sound still masks the important frequencies for intelligibility (300–3,000 Hz) the sound engineer can adjust the EQ even harder and bring down the low end farther. The audience will probably not think that they did not get what they came for just because there was not a lot of low-frequency content in the voice. The important thing here is getting the message across. And that predominantly happens, for vocal sound, from 300 to 3,000 Hz. The professional sound engineer will instinctively equalize the voice so that no frequencies are masking others to the point where a transparent, "open sound" appears from the PA system in the hall. Also the sound level produced by the PA system is important for speech intelligibility; a too-high sound level introduces distortion in the listener's hearing system, which can significantly reduce intelligibility.

Reflected sound within the mid- and high-frequency range can also mask the direct sound. Furthermore, late single reflections perceived as echo are very disturbing. Such echoes can result in temporal (forward) masking where a loud

reflection can make it difficult to hear new appearing direct sound. To sum up: the accumulated reverberant response provides a general "background noise" that certainly can partially mask the direct sound. Just as the sound engineer can clean out masking frequencies from a vocal to obtain a transparent sound, so he can obtain a clear mix of a whole rock band as discussed in Chap. 5.

Absorber Types

As mentioned, a certain volume of hall used for a given purpose has a recommended reverberation time associated with it. As seen from Sabine's equation, the hall volume and the area of (100 %) absorptive materials determine the reverberation time in a perfectly diffusive room. The geometry of the interior design affects the diffusivity of the enclosed space whereby the reverberation time is affected. Once the general architectural design of the interior of a hall is chosen it becomes the job of an acoustician, in accordance with the architect, to consider which building materials and acoustical products to employ, to what extent, and how to mount them because in this way they can predict which frequency bands are absorbed and to what degree. In this way a desired reverberation time as a function of frequency can be achieved in a given hall. In Table 1.1 absorption coefficients for some building materials are shown for various frequency bands. Absorption of sound can take place by three different means: porous, vibrating panel, and resonator absorption.

In porous absorption (such as mineral wool, drapes, porous concrete, etc.) the sound energy is dissipated into heat because the propagation of the sound wave is impeded. The porosity of a given material determines its flow resistance which is an important measure when calculating absorption properties of porous absorbers. The sound wave is impeded the most at the distance from the surface where the particle velocity of the wave is at its highest. This occurs 1, 3, or 5 or more times a quarter of a wavelength away from the surface; placing the porous absorber at a distance from the wall will increase the absorption effect at lower frequencies. But for the porous absorber to work at low frequencies it usually has to be of a certain thickness too in order to be obstructive for the relatively big wave. This is one reason why curtains, drapes, banners, and the like traditionally do not absorb any significant amount of sound energy below some 200 Hz. One can easily make a quick test if something is working as porous absorber: when blowing at the specimen, does the air seem to go through or is it blocked? If it goes through, and with some resistance, it will work by the porous absorption principle. See Fig. 1.15.

Vibrating panel absorption, also known as membrane absorption, occurs when the sound pressure on one side of a stiff or limb plate is significantly different from that on the other side. So if there is an airtight cavity behind the plate it will absorb sound because the plate will be forced to vibrate forth and back being pressed by the higher sound pressure on the one side and forced back by

Table 1.1 Possible acoustic absorption coefficients of different building materials (from Barron, Kuttruff, etc.)

Material	Center frequency of octave band (Hz)					
	125	250	500	1,000	2,000	4,000
Hard surfaces (brick walls, plaster, hard floors, etc.)	0.02	0.02	0.03	0.03	0.04	0.05
Slightly vibrating walls, suspended ceiling, glazing, etc.	0.10	0.07	0.05	0.04	0.04	0.05
Vibrating surfaces (single-layer gypsum board, etc.)	0.25	0.20	0.10	0.05	0.05	0.05
Carpet, 5-mm thick	0.02	0.03	0.05	0.10	0.30	0.50
Curtain (velour, draped)	0.06	0.31	0.44	0.80	0.75	0.65
Air absorption coefficient, 4 m (m^{-1})	0.00	0.00	0.00	0.00	0.01	0.03
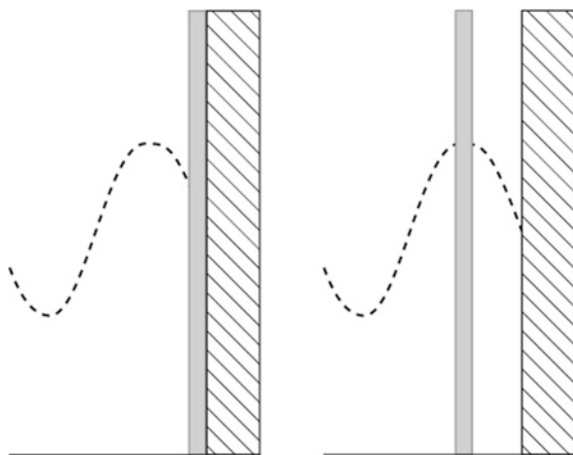
Fig. 1.15 Porous absorption works better at larger wave lengths when at a distance from the reflecting surface. The *dashed curve* represents the particle velocity in the reflected sound wave

the elasticity of the plate like a spring. Hence the system acts as does any mass-spring–damping system where the damping takes place inside the membrane as well as in the enclosed air volume of the cavity behind the plate. In order to increase the damping properties the cavity is often partly filled with porous absorption. A gypsum wall, a wooden floor on a cavity, and a window are all building elements that function by the membrane absorber principle and are certainly taken into account when designing spaces acoustically. The depth of the cavity and the mass per area of the membrane are two key factors when calculating at what frequency the absorber will achieve its maximum absorption as well as the frequency range of the absorber. Also the elasticity module of the plate material and the plate size are important attributes.

Resonating absorbers consist of an enclosed volume of air which is in open contact with the outside air through one or more openings. They are divided into three types depending on shape and number of openings: Helmholz resonators, slit

resonators, and resonating panels. They possess very different absorption properties. The sound absorption that occurs in these devices is connected to the fact that the incoming sound waves meet the reflected, phase-shifted sound waves whereby they to some degree cancel each other. Where Helmholtz resonators are used to absorb single standing waves, slit resonators and resonating panels are used to absorb broader frequency ranges. Such panels are well known, such as perforated gypsum panels mounted in the ceiling.

Absorbers are used either to achieve a suitable reverberation time for a given purpose, to eliminate unwanted reflections such as echoes (diffusers or angled surfaces are maybe more obvious for this purpose), or to lower noise levels in rooms such as kindergartens, factories, and offices.

One must always remember that in order to make a significant lowering of reverberation time across frequencies in a given hall very significant areas must be occupied with absorption of a relatively high absorption coefficient across frequencies. Often the entire ceiling area or more must be used to achieve a desired effect. This can be an expensive venture and that is why it is highly recommended to get an acoustical consultant to do exact calculations. It is simply too expensive not to get it right in the first stroke.

Audience Absorption

A common comment among live sound engineers to the musicians at sound check is, "Don't worry; it will be OK once the audience is in place," usually said with an ironic smile. Well, the audience in fact does absorb a lot of sound but almost exclusively at middle and higher frequencies. A tightly packed standing audience has an absorption coefficient of above 1.0 at frequencies higher than 1 kHz. The reason why a coefficient greater than 1 is possible is both that the surface of the people is greater than the surface at which they are standing, but also because they scatter the sound whereby the probability of the sound being absorbed elsewhere heightens, leading to a greater absorption coefficient. In Fig. 1.16 absorption coefficients of a standing and seated audience are shown as a function of frequency. If the seats are heavily upholstered, higher absorption coefficients at low frequencies can be obtained.

The absorption effect of an audience is very similar to that of heavy drapery in front of a reflective surface. Also, quite few people, diversely spread over the audience area, have an impact on the reverberation time of a hall. On the other hand, the author and his colleagues have encountered at least one hall that had to be completely filled with an audience before the sound would be acceptable on stage. In Chap. 4 20 Danish halls for pop and rock are presented. Each of the T_{30} diagrams includes a calculated T_{30} curve with a packed audience. The presence of an audience does not have a big influence on the low-frequency reverberation time of a hall. This is one reason why halls for pop and rock must be designed with a low RT at low frequencies as discussed later in detail.

Fig. 1.16 Absorption
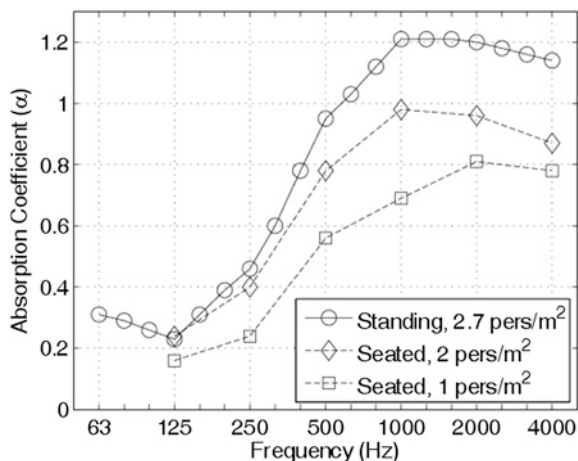coefficients of seated and
standing audiences

Table 1.2 Attenuation
constant of air at 20 °C and
normal atmospheric pressure,
in $10^{-3}\,\mathrm{m}^{-1}$ (after Kuttruff)

Relative humidity (%)	Frequency (kHz)				
	0.5	1	2	4	8
40	0.6	1.1	2.6	8.4	30.0
50	0.6	1.1	2.3	6.8	24.3
60	0.6	1.1	2.2	5.9	20.1
70	0.6	1.2	2.1	5.3	17.9

Air Absorption

A shorter reverberation above 2–4 kHz is normal in halls above a certain size as in
almost all halls presented in this book. This is due to the absorption of sound by
air. Air absorbs a significant amount of sound at very high frequencies as seen in
Table 1.2. Dry air absorbs much more sound than more humid air, so part of the high-
frequency sound above 2–4 kHz which is absorbed by the audience is "given back"
once they start to dance and sweat if the hall is not effectively dehumidified or the air
was humid to begin with. The sound engineer will usually make up for the sound-
level part of this effect on EQs. Also, at pop and rock concerts, artificial reverberation
is usually added to many instruments at higher frequencies anyway, and the amount
of this can be adjusted by the engineer according to the humidity changes. This effect
can also be part of the reason why the above-mentioned hall only sounded good with
a packed audience: the ventilation system in that hall was known to be unable to keep
up with a full house. The high-frequency sound was therefore not absorbed as much,
thus the low-frequency reverberation would not stand out as much but be somewhat
masked by higher frequency reverberant sound, making the hall just bearable.

The air absorption is also the reason why acoustic parameters often are not given
a value at the 8-kHz octave band for auditorium acoustics. In smaller size rooms
such as sound studios it is in that sense a relevant frequency band to consider.

Critical Distance and Level of Reverberation

As earlier stated, the sound level of a sound source in a free field decreases by 6 dB per doubling of distance. The sound in a room under the diffuse sound field assumption consists of two parts: the direct sound from the sound source and the reverberant sound which is the sum of reflections described as the diffuse sound field. The sound level of the diffuse sound field is approximately the same in any location of the hall. In rooms with a low RT this is not completely the case though, as discussed later.

$$L_p = L_w + 10 \log \left(\frac{Q}{4\pi r^2} + \frac{4(1 - \alpha')}{S\alpha'} \right) \text{dB(metric)} \tag{1.6}$$

In this equation L_p is the sound pressure level at a distance r, L_w is the sound power level of the sound source, Q is the directivity of the source, S is the surface area of the room, and α' is the area weighted average of the absorption coefficient.

The first term of the equation containing r^2 refers to the 6-dB attenuation of the direct sound field per doubling of distance, and the second term with $S\alpha'$ refers to the reverberant field. This equation is for empty rooms with diffuse sound fields with uniformly distributed sound-absorbing material.

This implies that close to the sound source, the direct sound is relatively loud and the reverberant sound has a given level which is the same anywhere in the room. Farther away from the sound source the direct sound level has decreased and the reverberant sound level is the same as closer to the source. Thereby the level of reverberant sound relative to direct sound is higher far away from the sound source compared to close to the source.

The distance from the sound source where the level of the direct sound is equal to that of the diffuse sound is called the *critical distance*, *reverberation radius*, or *room radius*.

From the above equation it can be deduced that

$$r, cr = \sqrt{\frac{QV}{100\pi T(1 - \alpha')}} \text{(metric)} \tag{1.7}$$

where Q is the directivity of the loudspeaker in a given frequency band, V is the volume of the hall, T is the reverberation time of the hall, and α' is the area weighted average of the absorption coefficient. For systems with several loud-speakers the equation yields

$$r, cr = \sqrt{\frac{QV}{100\pi T(1 - \alpha')N}} \text{(metric)} \tag{1.8}$$

where N is the number of loudspeakers or rather discrete clusters of loudspeakers. It is seen that the critical distance increases with higher values of Q and V and

with smaller values of T, α', and N. This means that a larger share of the audience will enjoy a clear sound when Q, V, and α' are increased and when T and N are decreased. This is easy to understand intuitively.

This has some consequences when designing acoustics for amplified music and also in the design of loudspeakers and loudspeaker systems for halls.

The most striking information in this equation is perhaps the effect of the low Q value of any loudspeaker at low frequencies. In the 63- and 125-Hz octave band the directivity of a sound source is low due to the omnidirectional nature of sound waves emitted from loudspeakers at these frequencies. An omnidirectional source has a Q of 1. The Q of, for example, a loudspeaker with a dispersion pattern of 90° wide by 40° high has a Q of 10. This is one reason why, as we show, the reverberation time T at low frequencies must be low.

The equation that includes the number of different sound sources implies, for instance, that one should only employ extra clusters of loudspeakers such as delay speakers when no other option is possible. On the other hand, one could be led to think that as low a reverberation time as possible would be the answer for correct acoustics for amplified music. As shown later this is not correct; in fact there is a lower limit to recommended reverberation time for pop and rock music halls just as for halls for other musical genres. Achieving enough level is usually not a problem because the speaker system normally can be turned up (or extra amplification power can be assigned) so reverberation to increase sound level is not needed (although the very low tones from up to perhaps around 70 Hz might in fact benefit from a certain amount of acoustical amplification).

Reverberation Time Design

When designing halls for classical music the challenge is often to get a high enough value of reverberation time. The reverberation time is proportional to the volume of the hall, therefore the challenge is typically met by building halls with a relatively high ceiling. One strategy used is that of counting the volume of empty space in the hall per audience. For amplified music, as we show, the aim is to obtain a lower RT. Different prediction tools such as computer models are of help as well as calculations including Sabine's formula mentioned earlier. From data sheets containing the absorption coefficient of building materials in at least the octave bands 125 Hz–2 kHz, and maybe some specific absorption coefficient measurements on certain special materials used in the hall, as well as experience, the trained acoustic consultant is capable of making good estimates of RT of a hall before it is built or refurbished.

When designing halls with seats, a type of seat is often chosen so that its absorption matches that of a person. This is a way to ensure that the acoustics don't change much from rehearsal to concert or if the hall is not completely full. This is of advantage for the orchestra and thereby also beneficial to the audience. Some typical values of absorption coefficients for different building materials are

found in Table 1.1. When using Sabine's formula it should be remembered that it is based on a perfectly diffuse room. This is seldom the case, especially when a large amount of absorption is present on, for instance, one entire surface. A packed audience on the floor represents such a surface.

Background Noise

Because the sound level is high at pop and rock concerts there are no real recommendations as to level of background noise within the hall. The audience is sometimes almost as loud between songs as the band playing their songs. Halls for classical performances have very strict background sound levels and this also indicates that clubs for dynamic jazz can benefit from a somewhat limited background noise level. In smaller clubs the bartender usually does not brew espresso during ballads. Ventilation and moving lights are among possible noise sources but also railways and highways can be too loud for a jazz club.

Sound Levels and Amplified Events

Niels W. Adelman-Larsen and Jacob Navne

Decibel denotes, as mentioned, the level of acoustic quantities relative to their reference values. In the case of sound pressure, the reference sound pressure, P_{ref}, is 20 μPa. One often used descriptor for absolute sound level is the sound pressure level (SPL) which is defined as

$$Lp = SPL = 10 \log \frac{p_{rms}^2}{p_{ref}^2}$$ (1.9)

where p_{rms} is the root mean square value of the sound pressure. RMS is the effective pressure of the time-varying sound pressure. The total SPL value is the summation of each SPL value in every octave or third octave band measured by the measurement device.

Because the human ear is less sensitive to low frequencies than to middle frequencies particularly at low levels, standardized frequency weighting filters are used to give the sound pressure a value that corresponds to the perceived hearing impression. The so-called A-weighted value, for instance, is 19.1 dB lower at 100 Hz and 10.9 dB lower at 200 Hz (third octave bands) than the sound pressure level with no filter used. The A-weighted sound pressure level is denoted L_A or can be specified by writing dB(A).

As a measure for characterizing the sound pressure level of a fluctuating noise averaged over time the equivalent sound pressure level L_{eq} is used. This measure can also be A-weighted, denoted $L_{A,eq}$. This is the basis of calculating how big a dose of sound a person is exposed to, such as during a concert or a working

Fig. 1.17 Sound level
readout on a laptop computer
at the mixing console of
Ancienne Belgique, Bruxelles

day in a manufacturing facility. For instance, in Denmark the maximum doses dur-
ing a eight- hour working day is an $L_{A,eq}$ of 85 dB. Mathematically, a doubling in
terms of dB is approximately 3 dB but humans need an extra 10 dB to experience
a sound level increase as a doubling.

The $L_{A,eq}$ can be measured directly on most sound-level meters. Calculations
can also be made and as an example, at a concert, three tunes of respectively 3, 4,
and 5 min are played; the averaged $L_{A,eq}$ level of the three songs is, respectively,
100, 95, and 105 dB at a given location. The total $L_{A,eq}$ during the 12 min of those
three songs at that location is found as

$$L_{A,eq} = 10 \log \left(\frac{3}{12} 10^{10,0} + \frac{4}{12} 10^{9,5} + \frac{5}{12} 10^{10,5} \right) = 102 \, \text{dB} \qquad (1.10)$$

In most countries legislation sets a maximum averaged sound pressure level for any
employee to receive during a workday. These pieces of legislation are in place to
ensure the health and safety of workers in noisy environments. It is the responsibility
of the employer to make certain that the noise in the work environment is as little as
possible. This is quite a dilemma when looking at live reinforced music, given that
the noise you need to protect the workers from is the actual product that this "fac-
tory" is selling to its audience, that is, the music being performed. Two very conflict-
ing interests are at play in this scenario. The customers (i.e., the audience) want to
experience an event they cannot reproduce at home, with punchy bass and loud and
clear sound and ambiance. At the same time the employees need to be exposed to as
little noise as possible. Moving workplaces such as the bar, coat check, and so on
outside the main concert area helps reduce the exposure by architectural means.

It has become very common to measure the actual sound level in venues dur-
ing amplified music concerts. The microphone is most often placed approximately
in the center of the audience area, at the sound engineer's desk, referred to as the
"front of house" or FOH. The sound engineer monitors the level, for instance, on a
laptop, throughout the concert (Fig. 1.17).

When measuring the sound level at a concert, the preferred measurement
method is L_{eq}. As music is dynamic in nature, the L_{eq} value is a good way to

monitor the averaged level such as during a whole concert. Doing SPL measurements at live events is always a compromise between measurement accuracy and realistic possibilities at the actual event. In an ideal world, a number of measurement positions would be looked at, as the SPL varies with distance to the stage and the main loudspeaker array. In reality it is, however, difficult to place expensive measurement microphones among the audience, and therefore one single fixed, protected position at the FOH is the common standard.

In some countries, such as Germany with the DIN 15905-5 standard, the measurement needs to be compensated for the difference between the loudest point in the audience, and the actual measurement point. In this norm, a test signal is played prior to allowing the audience into the venue, and using this signal and a compensation algorithm in the measurement equipment, the difference is measured and used during the concert. In this way, the values displayed by the measurement equipment are not the actual SPL at the measurement position, but a compensated calculated SPL at the loudest point in the venue. This approach is also used in Sweden and Belgium. It makes good sense inasmuch as actual hearing damage is often a result of a few minutes of extremely loud levels rather than a couple of dB higher SPL over a couple of hours or even a large number of concerts.

A number of countries have sound-level limits for live shows. These are not directly related to health and safety but mainly based on an overall compromise between many stakeholders—the audience who wants a physical experience, the organizer who wants to have a satisfied audience, neighbors who want their sleep—and an overall concern towards the comfort of the listening audience. These limits are often based on $L_{A,eq}$ values, and typically range from 99 to 103 dB and timespan averages from 15 to 60 min.

It is important to know that for a live show of any modern genre, an average $L_{A,eq}$ of 100 dB is generally needed in order to fulfill the requirement from the audience to both hear and feel the show. As mentioned, the typical maximum doses for employees in factories is $L_{A,eq} = 85$ dB over 8 h. If this same limit were to be followed at rock shows, a show at 100 dB on average could last no more than 15 min, then the audience would have received a full day of exposure due to the relationship between dB and time mentioned above (or simply by deducing that a 3-dB increase in average SPL equals a doubling of noise exposure to the ear, and thus halves the time it takes to obtain the same dose).

In that perspective, a live show will never be a "safe" event in terms of normal health and safety regulations, however, very few people attend more than 5–10 concerts a year, and when compared to the everyday exposure of MP3 music players and headphones mounted directly into the ear canal, with a much higher SPL, the risk that a live music event is the reason for a person developing a hearing impairment is low. The risk is there, though, and is biggest in small venues with a low ceiling or in other cases where the PA speakers are not elevated high above the audience. In these cases severe differences in SPL are present between audiences close to the stage and those in the rear.

In some countries, sound level limits at live shows are fairly gentle. Denmark, Norway, and the Netherlands, for instance, are using an $L_{A,eq}$ of 103 dB over

15 min. This makes for loud shows, and only limits very few acts in being as loud as they want to: in fact even with the right to be at 103 dB, most shows are played at 99–101 dB on average, as this often proves to be an adequate SPL. Of course if this measurement is made at the sound engineer's position, without being corrected with regard to the potentially shorter distance from the speakers to some audience members it may, depending on the layout of the hall, mean that other audience members receive considerable larger doses. Other countries including Sweden and Switzerland have a stricter approach, where shows are to be at 99 dB(A), and for Sweden, the limit is lowered if any audience member is below the age of 13. Then it is set at $L_{A,eq}$ 93 dB over 60 min. A number of exceptions exist.

Especially for outdoor events in closely populated areas, the limits are very strict, and sometimes even based on instant values instead of an average. This makes it very difficult to produce a concert at a sound level expected by the audience, and complaints from ticke tholders are often the end result.

One of the main limitations with the current legislation is that the maximum levels set forth are almost always based on A-weighted values. Part of the reason for that is also that the majority of noise regulations and guidance literature was written in the mid-twentieth century. At that time loudspeaker design was in its infancy, and not comparable to today's powerful line arrays and huge sub bass cabinets. The development in loudspeaker design has enabled sound engineers to play with full range systems, frequencies from around 30 Hz all the way to 20 kHz.

Bass frequencies are the most difficult to control and sound insulate. This fact often results in neighbors mainly being bothered by bass frequency sound, and not the sound stemming from other instruments. As mentioned, A-weighting removes a lot of bass content from the measurement to better mimic the behavior of the ear, and thus there is very little correlation between the measured SPL at the event and the nuisance experienced by the neighbor. An example that describes the inadequacy of A-weighting for this purpose could be to look at the difference between a British guitar rock band, and an electronic dance act. Both orchestras may play at the same $L_{A,eq}$ of 100 dB measured inside the venue, but due to the very heavy bass that is often associated with the electronic music genre, the neighbor of such an event will be bothered much more by the electronic act than by the rock band, as the bass content is much, much louder, but not reflected on the A-weighted measurement. Switching to more frequency flat C-weighted measurements would provide a much better correlation between the value measured inside the venue and the nuisance to the neighbors. Switching to C-weighted measurements would at the same time require a significant increase in the value of the maximum allowed L_{eq}. When looking at measurements of A- and C-weighted values performed simultaneously, there is often a 10–20 dB difference between the two quantities, so a limit of $L_{A,eq}$ 15 min = 100 dB, would have to be $L_{c,eq}$ 15 min = 110 dB or maybe even more.

Regardless of the important safety issue of sound-level control at amplified concerts, it should be remembered that our hearing system distorts more at higher levels. The masking curves in Fig. 1.12 show that at higher levels the sound engineer has a more difficult job creating a clear sound.