



**Technische Universität Berlin**  
Fakultät V - Verkehrs- und Maschinensysteme  
Institut für Strömungsmechanik und Technische Akustik  
Fachgebiet Technische Akustik  
Sekr. TA 7 - Einsteinufer 25 - 10587 Berlin

## **Room Acoustics (lecture notes)**

Drasko Masovic, TU Berlin, 2018      (last update: April 13, 2021)

# Contents

|   |           |
|---|-----------|
| <b>List of symbols</b>                              | <b>4</b>  |
| <b>1 Introduction</b>                               | <b>7</b>  |
| 1.1 Types of rooms . . . . .                        | 9         |
| 1.2 Characteristics of sources . . . . .            | 11        |
| 1.2.1 Dynamics and ambient noise . . . . .          | 11        |
| 1.2.2 Frequency range . . . . .                     | 13        |
| 1.2.3 Directivity . . . . .                         | 15        |
| 1.3 Subjective criteria . . . . .                   | 17        |
| <b>2 Basic equations</b>                            | <b>21</b> |
| 2.1 Conservation laws . . . . .                     | 21        |
| 2.2 Viscous and thermal losses . . . . .            | 22        |
| 2.3 Wave equation . . . . .                         | 24        |
| 2.4 Sources and boundary conditions . . . . .       | 25        |
| 2.5 Sound energy and sound pressure level . . . . . | 26        |
| <b>3 Modal analysis</b>                             | <b>29</b> |
| 3.1 Complex sine waves . . . . .                    | 29        |
| 3.2 Acoustics in frequency domain . . . . .         | 35        |
| 3.3 Damping . . . . .                               | 36        |
| 3.4 Green's function in frequency domain . . . . .  | 38        |
| 3.5 Room eigenmodes . . . . .                       | 42        |
| 3.6 Rectangular room with hard walls . . . . .      | 48        |
| 3.7 Modal density . . . . .                         | 55        |
| <b>4 Analysis in time domain</b>                    | <b>59</b> |
| 4.1 Green's function in time domain . . . . .       | 59        |
| 4.2 LTI systems and impulse response . . . . .      | 63        |
| 4.3 Directivity . . . . .                           | 64        |
| 4.4 Spherical and plane waves . . . . .             | 66        |
| 4.5 Acoustic impedance . . . . .                    | 71        |
| 4.6 Sound rays . . . . .                            | 72        |
| <b>5 Geometrical and statistical theory</b>         | <b>74</b> |
| 5.1 Summation of energy . . . . .                   | 74        |
| 5.2 Statistical theory and diffuse field . . . . .  | 77        |
| 5.2.1 Stationary state . . . . .                    | 81        |
| 5.2.2 Energy decay . . . . .                        | 83        |
| 5.3 Ray tracing in a rectangular room . . . . .     | 84        |

|   |            |
|---|------------|
| <b>6 Measurement and descriptors of room acoustics</b>            | <b>89</b>  |
| 6.1 Measurement of a room impulse response . . . . .              | 89         |
| 6.2 Descriptors of room acoustics . . . . .                       | 92         |
| <b>7 Estimation and optimization of sound fields in rooms</b>     | <b>101</b> |
| 7.1 Direct sound and visibility . . . . .                         | 104        |
| 7.2 Reverberation time and sound level . . . . .                  | 105        |
| 7.3 Early reflections . . . . .                                   | 108        |
| 7.4 Acoustic design . . . . .                                     | 109        |
| 7.4.1 Noise control in rooms . . . . .                            | 109        |
| 7.4.2 Rooms for sound reproduction . . . . .                      | 109        |
| 7.4.3 Recording studios . . . . .                                 | 110        |
| 7.4.4 Lecture halls . . . . .                                     | 111        |
| 7.4.5 Theatres . . . . .  | 112        |
| 7.4.6 Opera houses . . . . .                                      | 113        |
| 7.4.7 Concert halls . . . . .                                     | 114        |
| <b>8 Basic acoustic elements</b>                                  | <b>117</b> |
| 8.1 Infinite uniform plane surface . . . . .                      | 117        |
| 8.2 Diffuse surface . . . . .                                     | 121        |
| 8.3 A layer of porous absorber . . . . .                          | 123        |
| 8.4 Porous absorber in front of a rigid motionless wall . . . . . | 126        |
| 8.5 Helmholtz resonator and perforated panels . . . . .           | 128        |
| 8.6 Membrane resonator . . . . .                                  | 132        |
| 8.7 Quarter-wavelength resonator and seat dip effect . . . . .    | 134        |
| <b>9 Small elements</b>   | <b>138</b> |
| 9.1 Finite rectangular surface . . . . .                          | 138        |
| 9.1.1 Rigid motionless plate as a reflector . . . . .             | 142        |
| 9.1.2 Schroeder diffuser . . . . .                                | 143        |
| <b>10 Modelling in room acoustics</b>                             | <b>149</b> |
| 10.1 Ray tracing . . . . .  | 149        |
| 10.2 Image sources . . . . .                                      | 155        |
| 10.3 Scale models . . . . .                                       | 157        |
| <b>Literature</b>   | <b>159</b> |

## List of frequently used symbols

|               |  |
|---------------|--|
| $A_s$         | equivalent absorption area                         |
| $AL_{cons}$   | articulation loss of consonants                    |
| $a$           | radius   |
| $BQI$         | binaural quality index                             |
| $BR$          | bass ratio   |
| $C$           | clarity  |
| $C_p$         | specific heat capacity at constant pressure        |
| $c$           | speed of sound                                     |
| $D$           | damping, definition                                |
| $D_i$         | directivity  |
| $d$           | distance, thickness, depth                         |
| $E$           | acoustic energy (density)                          |
| $EDT$         | early decay time                                   |
| $e$           | unit vector  |
| $F$           | force (one-dimensional)                            |
| $f$           | frequency, generic function                        |
| $f_{Schroed}$ | Schroeder frequency                                |
| $G$           | Green's function, strength factor                  |
| $g$           | impulse response                                   |
| $H$           | height   |
| $H_e$         | Helmholtz number                                   |
| $\mathbf{I}$  | intensity vector                                   |
| $IACC$        | interaural cross correlation coefficient           |
| $IACF$        | interaural cross correlation function              |
| $I_s$         | irradiation strength                               |
| $K$           | heat conductivity                                  |
| $k$           | wave number  |
| $\mathbf{k}$  | wave vector  |
| $L$           | length, sound pressure level                       |
| $L_W$         | sound power level                                  |
| $LEF$         | lateral energy fraction                            |
| $l$           | length, mean free path length                      |
| $M$           | mass   |
| $m$           | attenuation constant, modulation transfer function |
| $\mathbf{n}$  | unit vector normal to a surface                    |
| $P$           | sound power  |
| $p$           | pressure, received signal                          |
| $Q$           | source function, Q-factor                          |
| $q$           | source function (per unit volume)                  |
| $\mathbf{q}$  | heat flux  |
| $R_s$         | reflection coefficient                             |

|                   |  |
|-------------------|--|
| $RASTI$           | rapid speech transmission index                                |
| $RP$              | radiation pattern  |
| $\mathcal{R}$     | specific gas constant  |
| $r$               | distance, magnitude of radius vector                           |
| $r_c$             | critical distance  |
| $S$               | stiffness, surface area  |
| $SC$              | Schroeder curve  |
| $SNR$             | signal-to-noise ratio  |
| $ST$              | stage support factor   |
| $STI$             | speech transmission index                                      |
| $s$               | signal, sequence   |
| $s_s$             | scattering coefficient   |
| $T$               | temperature, period, duration, reverberation time              |
| $T_s$             | transmission coefficient                                       |
| $t$               | time   |
| $t_s$             | centre time  |
| $V$               | volume   |
| $v$               | velocity   |
| $\mathbf{x}$      | position vector  |
| $\mathbf{y}$      | position vector (for a source or surface)                      |
| $Z$               | impedance  |
| $\mathcal{Z}$     | specific impedance   |
| $z$               | random number between 0 and 1                                  |
| $\alpha_s$        | absorption coefficient   |
| $\beta$           | volume fraction  |
| $\gamma$          | heat capacity ratio, directivity factor (gain)                 |
| $\Delta t_{init}$ | initial time delay gap   |
| $\delta$          | Dirac delta function, end correction, boundary layer thickness |
| $\epsilon$        | infinitesimal  |
| $\zeta$           | damping constant   |
| $\theta$          | polar angle  |
| $\iota$           | polar angle (for a source or surface)                          |
| $\lambda$         | wavelength   |
| $\mu$             | dynamic viscosity  |
| $\nu$             | kinematic viscosity  |
| $\Xi$             | flow resistivity   |
| $\xi$             | displacement (one-dimensional)                                 |
| $\rho$            | density  |
| $\sigma$          | porosity, perforation ratio                                    |
| $\tau$            | viscous stress tensor, time (for a source or surface)          |
| $\Phi$            | phase, phase shift   |
| $\phi$            | azimuthal angle  |
| $\chi$            | azimuthal angle (for a source or surface)                      |
| $\psi$            | mode (eigenfunction)   |

|          |                                   |
|----------|-----------------------------------|
| $\Omega$ | solid angle, modulation frequency |
| $\omega$ | angular frequency                 |

Note: Bold symbols denote vectors. Subscript  $_0$  is used for steady background values and  $\hat{\cdot}$  above a symbol indicates a complex amplitude. The functions  $\mathcal{R}_e(\cdot)$  and  $\mathcal{I}_m(\cdot)$  give real and imaginary part of a complex number, respectively,  $*$  denotes a complex conjugate value, and  $j$  is used as the imaginary unit. Symbol  $\partial$  is used for partial derivatives,  $d$  for total derivatives, and  $\nabla$  is the nabla operator ( $\nabla$  is typically gradient of a scalar and  $\nabla \cdot$  divergence of a vector). Symbol  $\propto$  stands for “proportional to”,  $\sim$  for “scales as”, and  $\sim \mathcal{O}(\cdot)$  for order of magnitude. Angle brackets  $\langle \rangle$  are used for averaging (for example, over time:  $\langle \cdot \rangle_t$ ).  $||$  gives magnitude of a vector or an absolute value (modulus). If a vector, say  $\mathbf{v}$ , has complex components, the expression  $|\mathbf{v}|$  gives its magnitude  $|\mathbf{v}| = \sqrt{\mathbf{v}^* \cdot \mathbf{v}}$ , which is a scalar (not a vector of the moduli of its components).

## 1 Introduction

Room acoustics is a sub-discipline of classical acoustics in fluids dealing with sound fields in closed (or partly open) spaces (see Fig. 1, left). Although most commonly associated with relatively large rooms filled with air, it can also include small spaces (cavities) and other fluids in which sound propagates. Boundedness of the observed space implies that besides the direct sound, which in the absence of obstacles reaches a receiver directly from the source, a series of reflections usually takes place at the boundary surfaces of the room, which substantially contribute to the sound field in it.

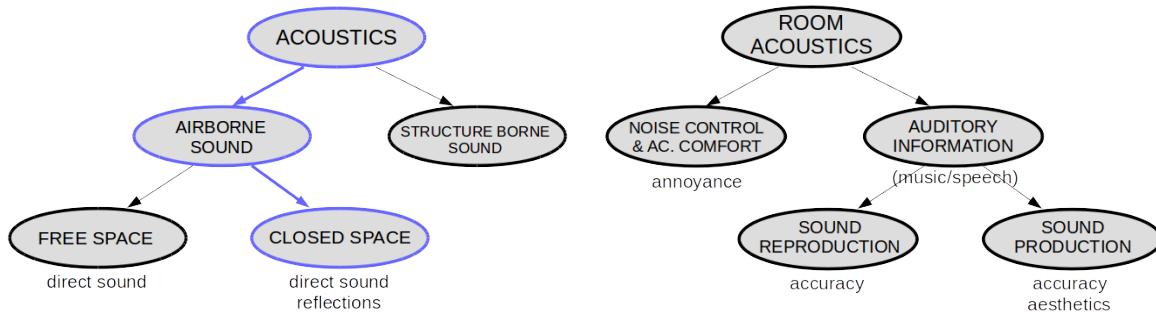


Figure 1: Room acoustics: (left) the object of study indicated with blue colour, (right) general types of applications.

Under normal circumstances, a room can be treated as a physical linear time-invariant (LTI) transmission system<sup>1</sup> between a source emitting a certain signal  $s(t)$  and a receiver which receives its possibly modified version  $p(t)$ . This is illustrated in Fig. 2. Room acoustics provides means for assessment, estimation, measurement, and optimization of the system – the direct sound and reflections – such that the resulting sound field at the receiver location satisfies certain objective or subjective criteria, which depend on the type of the room.

As shown in Fig. 1 (right), the applications of room acoustics spread in two main directions. The first group of applications involves primarily noise control of various sources of unwanted sound which may be present inside or, less commonly, outside the room. The acoustics of such rooms should ensure a sufficient acoustic comfort to people in them or protection of sensitive equipment against high noise levels. Moreover, even if a potentially disturbed listener or sensitive equipment are in a separate room, knowledge of the sound field in the source room can be used to efficiently suppress noise closer to the source.

For this type of applications, sound field is usually well described with basic acoustic quantities, such as sound pressure level or acoustic energy (often supplemented with

<sup>1</sup>As we will see in section 4.2, such systems are entirely characterized with their impulse responses.

The conditions for linearity will be discussed shortly, while the time invariance holds whenever the acoustically relevant conditions in the room do not change significantly over time, for example, there are no moving objects in the room and the thermodynamic properties of the medium are essentially stable.

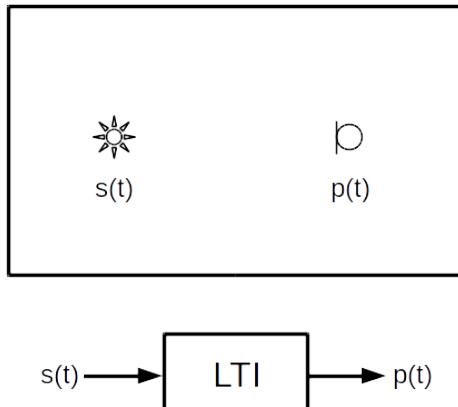


Figure 2: Room as a linear time-invariant (LTI) system.

reverberation time), averaged over the entire room volume, without a detailed analysis of the field. Furthermore, such applications are most common in engineering practice, frequently in combination with sound insulation and building acoustics (which will not be covered here<sup>2</sup>). The reason for this are various regulations concerned with environmental noise and acoustic comfort which exist in many countries and which set the upper limits for noise levels in different spaces. Very wide range of rooms in which people live or work can belong to this group of applications. Some examples are factories, machine rooms and workshops, open-plan offices, railway stations, airports, and other public places.

Quite different is the second group of applications, in which **sound fields in rooms carry a specific auditory information**, typically music or speech, which is of interest to its receivers. The listeners are often located in the same room, although the sound can also be first recorded with a microphone placed in the room. According to the type of the information, rooms can be simply but crudely divided into rooms for music (concert halls, opera houses, music studios, etc.), rooms for speech (lecture halls, theatres, conference rooms, etc.), and rooms with general purposes (including both speech and music). However, such a division becomes impractical inasmuch as speech and music appear together, often with similar significance (for example, in an opera house). A more appropriate division can be made according to the origin of sound emitted by the source in the room, which leads to two sub-cases.

In the first case, the main goal is accurate reproduction of already complete and given sound information, which has been previously stored on a medium. The reproduction is achieved with the aid of an audio system and the loudspeakers as sources of sound in the rooms. Typical examples are control rooms in studios for music or speech as well as cinemas. Especially sensitive to the influence of room acoustics are stereo and multichannel recordings, which contain not only sound but also spatial information on the location of virtual sources. In general, the reproduced audio signal should be distorted as least as possible by the room as the transmission system.

<sup>2</sup>Room acoustics and building acoustics together are commonly referred to as architectural acoustics.

In the second case, the auditory information is created (produced) directly in the room and the room is expected not only to transmit it accurately to the receiver, but also to **enhance it in a desirable way**. The usual aim is to support the source of sound by increasing sound energy at the receiver location (the room gain) and to meet more sophisticated subjective, even aesthetic criteria of the listeners. Such rooms “actively” contribute to the sound field and affect the perceived sound and, therefore, their acoustics becomes critical. The examples are concert halls and opera houses, as well as lecture halls and theatres. The acoustics of these rooms is commonly considered as part of the overall design of the room interior, thus interweaving with architectural aesthetic criteria. Architects and acoustic consultants are then expected to collaborate in order to achieve a successful design satisfying both visual and auditory criteria. The described general applications of room acoustics are briefly summarized in Table 1.1.

Table 1.1: Applications of room acoustics in engineering practice.

|                      | <b>room acoustics</b>             |                             |
|----------------------|-----------------------------------|-----------------------------|
|                      | <b>noise control, ac. comfort</b> | <b>auditory information</b> |
| main goal            | noise suppression                 | accuracy/enhancement        |
| criteria             | objective                         | objective and subjective    |
| target values        | regulations                       | guidelines/suggested values |
| acoustic treatment   | norms                             | best practice               |
| related discipline   | building acoustics                | interior architecture       |
| project collaborator | civil engineer                    | architect                   |

## 1.1 Types of rooms

Several common types of rooms in which an auditory information is produced or reproduced are listed in Table 1.2. Some of them, such as theatres, may involve both produced (by the performers on the stage) and reproduced (by the public address (PA) system) sound. The table indicates also the usual sources of sound for each type of the room, whether the receivers are mainly human listeners or microphones, as well as the expected spatial distribution of the sources and receivers. Note that a podium or stage can also be treated as a receiver location, since the speakers and performers are also listeners of the sound they produce and the room provides them with a necessary acoustic feedback. Type of the auditory information (music/speech/singing) is also relevant for the specific applications and given in the table.

The table gives only a general overview of typical scenarios and various deviations from them are certainly possible, for example, use of PA systems in lecture halls or opera houses, a single room used as a control room and recording studio, and so on. Placement of sources and microphones in recording studios can also be controlled and predetermined, so that their preferred locations are more or less fixed. All these aspects should be carefully considered when analysing acoustics of a room.

Table 1.2: Types of rooms and common sources and receivers of sound in them. Abbreviations: room type – (P) sound production, (R) sound reproduction; sound – Mu music, Si singing, Sp speech; source – MI musical instrument, Hu human, LS loudspeaker (sound reinforcement); receiver – Hu human, Mi microphone.

| room type        | sound      | source | source location  | rec. | rec. location  |
|------------------|------------|--------|------------------|------|----------------|
| concert hall (P) | Mu, Si     | MI, Hu | stage            | Hu   | auditorium     |
| opera house (P)  | Si, Mu, Sp | Hu, MI | stage, orch. pit | Hu   | auditorium     |
| theatre (P, R)   | Sp, Mu     | Hu, LS | stage            | Hu   | auditorium     |
| rec. studio (P)  | Mu, Si, Sp | MI, Hu | not fixed        | Mi   | not fixed      |
| control room (R) | Mu, Si, Sp | LS     | fixed            | Hu   | fixed, limited |
| cinema (R)       | Sp, Mu     | LS     | fixed            | Hu   | auditorium     |
| lecture hall (P) | Sp         | Hu     | podium           | Hu   | auditorium     |
| classroom (P)    | Sp         | Hu     | podium           | Hu   | auditorium     |
| club (P, R)      | Mu, Si     | MI, LS | stage, fixed     | Hu   | not fixed      |

Other types of spaces which are not included in Table 1.2 but can also be treated using the methods of room acoustics are churches, amphitheatres and stadiums (as partly open spaces), sport halls, offices (especially open-plan offices), various public spaces, small cabins (for example, in vehicles) and cavities, and many other. A special type of rooms which appears in practice quite frequently are multipurpose rooms. In general, multifunctionality of these rooms assumes different types of the auditory information (speech and music) or varying locations of sources and receivers in the room. Although certain flexibility can be achieved with changeable geometry of the room and adaptive acoustic elements, the outcome is always a compromise between the more general functionality of the room and reasonable but sub-optimal acoustic quality.

The **ultimate receiver of the auditory information is usually human**, whether directly, being present in the room, or indirectly, when listening the recordings made with the microphones located in the room. However, if the receiver in the room is a microphone, its directivity can be taken into consideration in order to relax the requirements for the room acoustics. For example, undesirable strong reflections in a recording studio can be largely avoided with a careful placement of directional microphones. Similar consideration holds for directional sources. For example, controllable directivity of line array loudspeaker systems is frequently used to circumvent acoustic deficiencies in large spaces (very large halls, stadiums, public spaces, etc.), whether they lack useful reflections which "amplify" the source without major distortion, or the undesired (late) reflections must be suppressed. In general, we can conclude that certain acoustic properties of sources and receivers (besides their location in the room) complement room acoustics and should be taken into consideration. In the rest of this section we study in more details the most important characteristics of sources which commonly appear in rooms, as well as subjective criteria of human listeners, which are all relevant for the appropriate room acoustics.

## 1.2 Characteristics of sources

Next we consider some of the **basic acoustic properties of various sources of sound** (music, singing, or speech) which affect the criteria for optimal room acoustics. These are dynamic range, frequency range, and directivity.

### 1.2.1 Dynamics and ambient noise

Dynamic range of a source of sound determines the maximum acceptable level of noise in a room, as well as the minimum gain of the room, which ensures sufficient sound energy at the receiver's location. Table 1.3 gives a very rough indication of relatively sustained maximum sound power levels of some common sources of sound. For easier comparison with everyday sounds, the values are also converted to sound pressure levels in free space, at the distance<sup>3</sup> 1 m, estimated using the equality<sup>4</sup>  $L_{free,max,1m} = L_{W,max} - 11 \text{ dB}$ . Time average values can be considerably lower, say, by 10 dB or more. Including the attenuation due to sound propagation (6 dB per doubling the distance from a point source in free space, as will be derived in eq. (4.18)), it becomes evident that most sources of interest, such as human voice, cannot provide sufficient sound pressure levels for typical distances of the receivers without additional gain of the room due to the reflections. Especially critical are relatively high frequencies, at which the sources typically radiate less efficiently, but which are essential for intelligibility of speech and clarity of music.

With the exception of percussions, the values shown in Table 1.3 correspond to relatively sustained achievable sound levels. Of course, sound pressure level is expected to vary in time. Extreme examples of time-dependence are short impacts of percussions. If the duration of such transient sounds is shorter than around 200 ms, the natural inertia of human auditory system can lead to even lower perceived sound levels compared to the actual, objective ones. For instance, a transient sound with duration 10 ms is perceived as around 10 dB weaker than a sound with equal amplitude and duration 100 ms. The perceived sound level of very short sounds increases nearly linearly per doubling the duration. However, slower variations of sound energy are also important for clear reception of sound information, especially speech (as will be discussed further in section 6.2). A good illustration of this are highly reverberant spaces, which tend to smooth out the time variations of sound level. Parts of the information which are contained in the dynamics of the signal, such as silent consonants or short quiet intervals between successive musical tones, are thus diminished. This results in poor speech intelligibility and inappropriate room response for most of the music styles (especially those which involve

<sup>3</sup>According to eq. (5.36), the distance of 1m matches quite closely the critical distance at middle frequencies in many common rooms. At larger distances from the source, sound pressure level in an ideally diffuse sound field remains approximately constant, so the given values can also be taken as very roughly estimated maximum values of sound pressure level in the room outside the zone of the direct sound dominance.

<sup>4</sup>See eq. (4.34). For simplicity, no directivity is taken into account and the relation for a small (point) source is used. Therefore, the estimated sound pressure levels are indeed only indicative, especially for large sources.

Table 1.3: Maximum sound power levels ( $L_{W,max}$ ) and approximate sound pressure levels at the distance 1m in free space ( $L_{free,max,1m}$ ) of common sources of sound.

| source                     | $L_{W,max}$ | $L_{free,max,1m}$ |
|----------------------------|-------------|-------------------|
| normal voice               | 85 dB       | 74 dB             |
| loud voice                 | 95 dB       | 84 dB             |
| loud singing               | 100 dB      | 89 dB             |
| plucked string instruments | 90 dB       | 79 dB             |
| bowed string instruments   | 95 dB       | 84 dB             |
| woodwind instruments       | 100 dB      | 89 dB             |
| piano                      | 110 dB      | 99 dB             |
| brass instruments          | 115 dB      | 104 dB            |
| percussions (peaks)        | 120 dB      | 109 dB            |
| organs                     | 130 dB      | 119 dB            |
| orchestra                  | 135 dB      | 124 dB            |

fast succession of tones). Consequently, not only sufficient sound pressure level must be ensured at the location of the receiver, but the dynamic range of the source should be preserved to a certain extent.

Dynamic range, which is the difference between maximum and minimum sound power level, varies between different sources. It is around 30 dB for woodwind instruments and somewhat larger for brass instruments. Violins can achieve even larger dynamic range, up to around 50 dB. Dynamics of human voice is around 20-30 dB. These are also the approximate ranges of sound pressure levels which a room should provide at the location of the receiver.

While the maximum sound pressure level depends on the sound power of the source and the room's gain, lower limit of the perceivable useful sound is most often determined by the background noise in the room. Its level sets a threshold for the quietest sounds which can be heard by a listener without being masked by the ambient noise. The sources of airborne and structure-borne noise can be located outside the room, such as traffic or many other sources of noise in adjacent rooms and corridors (when they are not efficiently suppressed by means of sound insulation), but also inside the room, such as ventilation systems, lighting, and other installations. The highest allowed levels of noise are often strictly defined in norms and regulations for different types of rooms. Since both the sensitivity of human hearing and ambient noise are frequency dependent, single-number values can be obtained after A-weighting and energy summation or by implementation of the noise rating curves. In the latter case, rating of broadband noise with a single number is done in the following way: octave noise spectrum is drawn on the top of the NR-curves shown in Fig. 3; it is then represented by the lowest NR-curve which has all its values (at all frequencies) larger than the values of the noise spectrum; number in the name of each NR-curve, which corresponds to its value at 1 kHz, is the rated value in dB.

Suggested maximum values of the background noise level are given in Table 1.4 for the rooms from Table 1.2 and several other types. These values vary somewhat between countries and their regulations. For easier comparison with Table 1.3 and assessment of the dynamic range, A-weighted equivalent sound pressure levels in the last column are estimated to be 10 dB above the rated value of the NR-curve, which is a reasonable estimation for broadband sounds. Table 1.4 clearly shows that the requirements with regard to ambient noise are particularly stringent for the rooms for music listening or recording. This is primarily due to the large expected dynamics of music. They are followed by the rooms for speech. The highest levels of ambient noise are allowed when sound reinforcement is present.

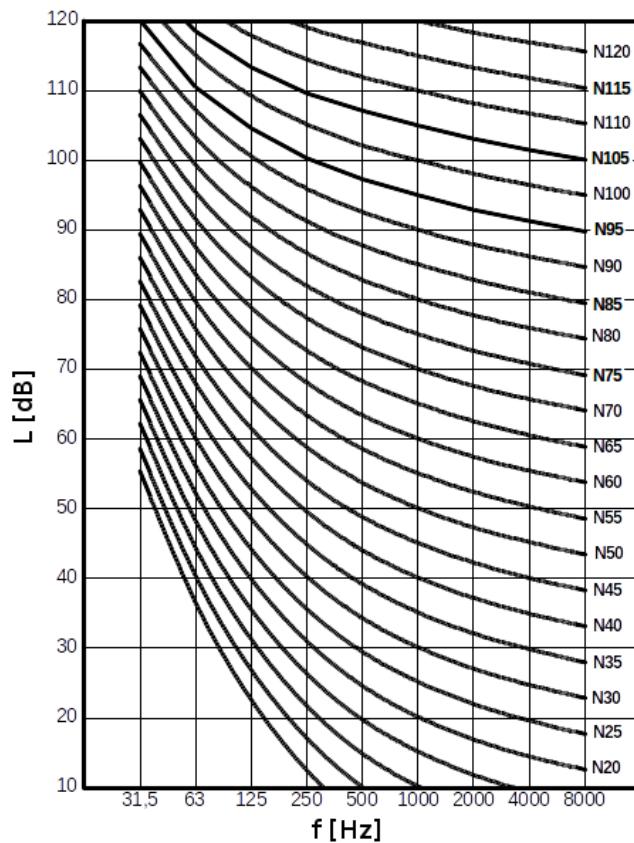


Figure 3: The NR-curves.

### 1.2.2 Frequency range

Although the audible frequency range is approximately from 20 Hz to 20 kHz, the fact that not all its parts are equally relevant for particular applications of room acoustics can substantially simplify the treatment. At moderate sound pressure levels, around 80 dB, human hearing is much less sensitive to the tones at low frequencies than at mid-to-high frequencies, 1-5 kHz, where the sensitivity is highest (for example, around

Table 1.4: Suggested maximum ambient noise levels in rooms expressed in terms of NR-curves and approximate A-weighted sound pressure levels.

| room type            | NR-curve | $L_A$ |
|----------------------|----------|-------|
| concert hall         | 15       | 25 dB |
| opera house          | 15       | 25 dB |
| theatre              | 20       | 30 dB |
| rec. studio          | 15       | 25 dB |
| control room         | 15       | 25 dB |
| cinema               | 30       | 40 dB |
| lecture hall         | 25       | 35 dB |
| classroom            | 25       | 35 dB |
| club (public spaces) | 35       | 45 dB |
| sleeping room        | 25       | 35 dB |
| living room          | 30       | 40 dB |
| church               | 30       | 40 dB |
| stadium              | 45       | 55 dB |
| library              | 30       | 40 dB |
| open-plan office     | 35       | 45 dB |

20 dB at 50 Hz). On the other hand, useful auditory information is rarely contained in the sound at very high frequencies, above this range. Moreover, most sources of sound in rooms, such as musical instruments and human voice, radiate less sound energy at high frequencies above several kilohertz and the energy is strongly dissipated in air or absorbed at room surfaces. In practice, we are rarely interested in octaves above 4kHz.

Bulk of the energy of speech is contained in the octaves from 63 Hz to 8 kHz. The fundamental frequency of an adult male voice is between 80 Hz and 180 Hz and of a female voice around 150-250 Hz. However, most of the spoken information is carried by the higher harmonics (formants). Sound energy of vowels is predominantly in the octaves 250-2000 Hz. Consonants appear at even higher frequencies, in the octaves between 2 kHz and 8 kHz. While low and middle frequencies carry most of the energy of speech (the vowels), high-frequency contribution of consonants is critical for intelligibility, even though their energy can be 10-20 dB weaker.

Frequency bands in which the fundamental frequencies of singing appear are 63-250 Hz (bass), 125-500 Hz (tenor), and 250-1000 Hz (soprano). Another frequency region of special importance is around 3000 Hz, which contains the so-called singer's formant. It allows trained singers to stand out from the orchestra, which is at lower frequencies often a challenging task. Table 1.5 summarizes the discussion above by listing the frequency ranges of fundamental frequencies of common musical instruments, or in which bulk of sound energy is contained, in the case of percussions and human voice. It should be mentioned that higher harmonics of musical tones as well as broadband, non-harmonic (stochastic) components can extend somewhat above these ranges.

Table 1.5: The most relevant frequency ranges (in octave bands) of musical instruments and human voice.

| source        | freq. range      |
|---------------|------------------|
| speech        | 125 Hz – 4 kHz   |
| bass          | 63 Hz – 250 Hz   |
| tenor         | 125 Hz – 500 Hz  |
| soprano       | 250 Hz – 1 kHz   |
| guitar        | 63 Hz – 1 kHz    |
| harp          | 31.5 Hz – 4 kHz  |
| double bass   | 31.5 Hz – 250 Hz |
| cello         | 63 Hz – 1 kHz    |
| violin        | 125 Hz – 4 kHz   |
| contrabassoon | 31.5 Hz – 250 Hz |
| clarinet      | 125 Hz – 2 kHz   |
| flute         | 250 Hz – 4 kHz   |
| piano         | 31.5 Hz – 4 kHz  |
| tuba          | 31.5 Hz – 500 Hz |
| trombone      | 63 Hz – 500 Hz   |
| trumpet       | 125 Hz – 1 kHz   |
| timpani       | 63 Hz – 250 Hz   |
| drums         | 125 Hz – 500 Hz  |
| xylophone     | 250 Hz – 4 kHz   |
| triangle      | 2 kHz – 8 kHz    |
| organs        | 16 Hz – 8 kHz    |

As a general outcome, in room acoustics we are typically interested in the frequency range of octave bands from **63 Hz to 4 kHz**. This corresponds to the third-octave bands from 50 Hz to 5 kHz. **If the room is intended only for speech, the lowest octave (63 Hz) becomes less relevant, as far as useful sound information is concerned.** In contrast to this, particularly significant are the octaves in the middle of the audible range, 500Hz, 1kHz and 2kHz, in which the human auditory system is most sensitive and the information content is high.

### 1.2.3 Directivity

Practically all real sources of sound are directional, that is, **they do not radiate sound equally in all directions**, at least not at all frequencies. While pronounced directivity can cause too low sound levels in certain directions from the source, it can also help the acoustic optimization of rooms. For example, locations and orientations of musical instruments on the stage of a concert hall with regard to the auditorium are well defined by the tradition of music performance. Positions of the instruments and singers in a

recording studio can be controlled. Location of a speaker on the podium of a lecture hall or a loudspeaker of a PA system is also fixed and known. Together with more or less known orientation and directivity of the sources, regions and directions of higher sound radiation can be identified and utilized for the optimization. This is, of course, much more difficult if the source is expected to move or change its orientation frequently, in which case its directivity becomes less relevant for the acoustic design.

Since it involves the dependence on both angle to the source and frequency, directivity is quite cumbersome to represent accurately graphically or numerically (see also the discussion in section 4.3). Table 1.6 gives very approximate values of coverage angles (here angles of principle radiation within which the far-field sound pressure level varies less than  $\pm 1.5\text{dB}$ ) of human voice and musical instruments. The values are given for the horizontal plane, which is, for example, plane of the bridge of string instruments or the bore of wind instruments. The directivity clearly increases with the size of the source. Almost all of the listed sources are omnidirectional only at frequencies below 250Hz, which means that they are directional in most of the frequency range of interest.

Table 1.6: Coverage angle (0 to -3 dB) of musical instruments and human voice in the horizontal plane. Only one main radiation lobe is considered. Note: sound generation of flute in the octave band 125 Hz is practically negligible (compare with Table 1.5).

| source        | 125 Hz | 250 Hz | 500 Hz | 1 kHz | 2 kHz | 4 kHz |
|---------------|--------|--------|--------|-------|-------|-------|
| speech/voice  | 300°   | 280°   | 160°   | 160°  | 120°  | 90°   |
| harp          | 360°   | 360°   | 180°   | 80°   | 60°   | 30°   |
| double bass   | 320°   | 270°   | 220°   | 100°  | 90°   | 80°   |
| cello         | 360°   | 150°   | 110°   | 80°   | 70°   | 80°   |
| violin        | 360°   | 360°   | 220°   | 180°  | 140°  | 90°   |
| contrabassoon | 360°   | 360°   | 120°   | 80°   | 60°   | 40°   |
| clarinet      | 360°   | 360°   | 360°   | 220°  | 100°  | 40°   |
| flute         | /      | 100°   | 130°   | 130°  | 110°  | 110°  |
| tuba          | 190°   | 160°   | 80°    | 45°   | 30°   | 20°   |
| trombone      | 360°   | 360°   | 140°   | 100°  | 45°   | 45°   |
| trumpet       | 360°   | 360°   | 360°   | 90°   | 80°   | 40°   |

Coverage angle is a simple and convenient parameter when the source has a single direction of principle radiation. However, certain musical instruments exhibit more complicated directivity patterns. For example, a flute radiates sound both from the blow hole and the tone holes or the open end and, accordingly, its radiation pattern has maxima in these different directions, which makes the values in Table 1.6 inaccurate. This is especially true at high frequencies, when complicated radiation patterns are expected in general.

Radiation patterns are also often asymmetric with respect to the direction of maximum radiation, which is not captured by the value of coverage angle. This is usually the case

when the source itself is asymmetric and large, or the player of a musical instrument is located very close to the instrument thus affecting its radiation. For instance, head of a violinist presents an obstacle for sound radiation in a particular direction. Directivity patterns of piano and organs, which are not listed in Table 1.6, are especially complicated due to their large sizes. Most of the sound energy of a grand piano is radiated from the soundboard upwards and downwards and certain fraction of the energy is reflected from the lid. Therefore, position of the lid can also affect the radiation pattern. Percussions are also omitted from the overview in the table above. In general, they radiate predominantly normal to the membrane or plate. Finally, directivity can vary even between the instruments of the same type, for instance, due to different physical properties of wooden boards of string instruments or their particular assembly.

A special type of the sources of sound are loudspeakers. Their directivity depends largely on their size and increases with frequency. Conveniently, the directivity patterns and other related parameters of loudspeakers are usually provided by the manufacturers and thus well known. Of special interest are line arrays of loudspeakers which consist of several loudspeakers clustered typically in the form of a straight or J-shaped array. Careful pre-processing of the input signals supplied to each of the loudspeakers allows a high flexibility and control over the directivity of the entire system. With larger number of loudspeakers in the array, higher directivity can be achieved and the direct sound energy can be distributed more uniformly over a large auditorium, compared to a single loudspeaker of equal sound power. In certain applications, this can be used to suppress the deficiencies of room acoustics, such as lack of useful reflected sound energy in (semi) open spaces, for example stadiums, or detrimental late reflections in highly reverberant rooms, sport halls, and churches. The pronounced and controllable directivity of line arrays is thus utilized both to avoid undesired reflections and minimize the losses of useful sound energy. Note that essentially the same applies to directional microphones and microphone arrays, when they are used as receivers.

### 1.3 Subjective criteria

Since we mostly consider human listener as the receiver of sound information, sound field in a room is eventually assessed subjectively. With regard to that, Table 1.7 lists the most relevant subjective criteria for the already discussed applications of room acoustics. It is important to note that these criteria must be satisfied only at those particular locations in rooms where the receivers are expected. This is especially critical when the locations of sources and listeners are known and fixed, for example at the stage or in the auditorium, or even spatially limited, such as in control rooms (locations of the loudspeakers and at the mixing desk). Acoustic design can then be focused on these key locations in the room, which makes the treatment somewhat easier and more efficient.

Already at the first glance the table indicates that the subjective requirements are most complicated and diverse for rooms in which the sound information is created. In addition to the acoustic-related aspects which we consider next, in many cases, such as opera

Table 1.7: The main subjective criteria for different applications of room acoustics.

| application                | subjective criteria  |
|----------------------------|--|
| noise control, ac. comfort | low loudness, appropriate reverberance   |
| sound reproduction         | suppressed effects of the room   |
| sound production           | appropriate loudness and dynamics, balance, reverberance, intelligibility (speech) or clarity (music), synchronization with the visual component, spaciousness, envelopment, apparent source width and localization, intimacy, absence of echo and coloration; <i>on the stage</i> : ease of ensemble, support; *visual criterion – visibility of the source |

houses and theatres, visibility of the source (actor or performer) presents an important additional component. It is determined at the first place by the distance between a listener (viewer) and the source, as well as by the angle at which the source is observed. The question how the listed subjective criteria can be met in practice will be considered in section 7. Here we will only give a few general remarks on each criterion and how they can be linked to certain objective physical properties of sound fields.

**Loudness** is closely related to the **energy of sound and sound pressure level**. Since **human auditory system reacts to the excitation with certain inertia**, it tends to “integrate” the sound energy over a certain finite time interval, which is of the order of 10 ms. The total (rather than instantaneous) **sound energy which is received in such short time windows determines the subjectively perceived loudness of sound events**. Consequently, the perceived sound level should be well above the (supposedly steady) ambient noise level in the room (see Table 1.4). A room should also provide an appropriate **balance** between sound energy at different frequencies, especially rooms for music performances.

For good **intelligibility** of speech or **clarity** of music, sufficient loudness of the signal is not enough. Certain minimum dynamic range of the received sound has to be ensured, since much of the information in both speech and music is contained in the temporal fluctuations of sound energy between the spoken syllables or played musical tones as well as in the quieter parts. Characteristic time scale of such fluctuations has typically the order of 100 ms or larger<sup>5</sup>. Accordingly, sound energy inside a room should decay fast enough during the quiet intervals not to mask the succeeding sounds<sup>6</sup>. On the other hand, the early decaying energy which reaches the listener shortly after the direct sound, well within the first 100 ms, does not compromise clarity but even contributes positively by increasing the loudness.

<sup>5</sup>This value can be directly compared with the modulation frequencies for speech which are given in section 6.2.

<sup>6</sup>As a quick analysis, if we suppose that the sound energy should decay at least 10 dB in 100 ms, the reverberation time in the room (which will be defined later) below  $T_{60} = 60 \text{ dB} \cdot 0.1 \text{ s} / 10 \text{ dB} = 0.6 \text{ s}$  is necessary for good speech intelligibility. For comparison, optimal reverberation time of lecture halls is typically around 1 s and even below 0.5 s in small classrooms.

**Synchronization** between the perceived auditory and visual components depends mainly on the distance between the source and receiver. Since the visual information is received practically instantaneously, time delay with which the direct sound reaches the listener (and the attenuation during propagation) should not be too large. Maximum distance of around 30 m is recommendable for concert halls, when the visual component is of secondary importance. When good **visibility** of the source is more or less equally important as the auditive component, as in theatres and opera houses, the maximum distance should be smaller. For a proper distinction of facial expressions of actors on the stage, the viewer should not be placed more than around 20 m from the stage.

The propagation time of the direct sound affects the synchronization. On the other hand, the delay (with respect to the direct sound) and energy of the earliest reflections which reach the listener determine the perceived distance from the source. If the earliest reflections arrive relatively long after the direct sound and with low energy, the source appears to be closer (both source and receiver appear to be far from the reflecting surfaces and close to each other). This quality of the perceived sound is often called **intimacy**. It decreases when strong reflections immediately follow the direct sound.

**Reverberance** makes the sound more natural and subjectively more appealing than a pure “dry” sound. However, optimal reverberance (and the associated objective physical phenomenon – reverberation) for music performance depends on the music style and has partly historical roots. For example, Baroque music is usually associated with longer reverberation times than Romantic music. Based on this, the halls for music performances (the first two rows of Table 1.2) can be further subdivided according to the type of the music which is produced in them. A compromise is often inevitable, which should make the room’s reverberance satisfactory for most of the music performed in it. While reverberation time of 2 seconds or longer is appropriate for church music, shorter reverberation times are more suitable for contemporary music genres, and even shorter (around and below 1 s) for speech.

**Spaciousness** is often stated to consist of two components: **listener envelopment** (impression of being immersed in the sound field, or being surrounded by it) and **apparent source width**. It depends mostly on the amount of sound energy which reaches the listener laterally. Hence, appropriate room acoustics, especially for music performances, should provide not only sufficient level of sound and clarity (which can in principle be achieved with localized sound reinforcement even in an anechoic environment), but also appropriate angular distribution of sound energy with respect to the listener. Uniform angular distribution of the incoming sound energy approaches the idealization of a diffuse field (see section 5.2), in which many reflections with equal energy reach the listener from all directions.

During the performance, the players and singers have to be able to hear well the sound of their own instrument or voice, as well as the other performers. This introduces additional requirements for room acoustics. In fact, in the absence of other listeners (for example, in recording studios), these requirements and associated subjective components become essential. They are often divided into ease of ensemble and support. **Ease of**

**ensemble** rates the ability of the performers to hear well each other, which allows easier common performance. When it is low, the result can be poor synchronization and balance (relative strength) of the sound produced by different performers. In contrast to this, **support** relates to the feedback which the performers obtain from the room. Lack of support means that the performers might not hear their instruments or voices loud enough. As a consequence, they will probably tend to play or sing louder than actually necessary. Both ease of ensemble and support can be low in overdamped rooms (such as in poor recording studios) or at stages lacking local early reflections (typically from the closest rear wall or ceiling).

Presence of strong (not more than around 20 dB weaker than the direct sound<sup>7</sup>) and distinct reflections or resonances of the room can cause undesired echo or coloration (change of the spectral content of the sound). Both of these phenomena are quite easily perceived and therefore the room acoustic design should prevent or at least minimize them. The major difference between echo and coloration is in the length of the delay of the reflections with respect to the direct sound. A rough border between the two is the delay time of 100 ms. The reflections which reach the receiver with shorter delays are difficult to be perceived as an echo, a replica of the direct sound, but rather as a change of timbre<sup>8</sup>.

If the reflections immediately follow the direct sound, with the delays even less than around 5 ms, they may cause an apparent **shift** of the location of the source or change of its **width**. While this may be less problematic or even desirable in rooms for music production, since it contributes to the spaciousness (see above), it can be extremely detrimental for a stereo or multichannel sound image. Especially interfering are reflections which reach a listener at relatively low angles, around 30-50°, in the horizontal plane, where spatial resolution of the human auditory system is particularly high.

It should be bear in mind that although the subjective components of a sound field can be associated with certain physical parameters, these relations are by no means direct and simple. For example, minimum energy of the reflected sound which is required for the perception of echo, coloration, or apparent shift of the source depends strongly on the character of the sound (tonal sound, random noise, transient, etc.) and its angle of incidence. Correlation between objective and subjective parameters is the object of many studies, based on which widely used numerical descriptors of room acoustics are derived, which will be presented in section 6.2.

---

<sup>7</sup>In certain cases, reflected sound can even be several decibels stronger than the direct sound, for example, when a concave surface focuses the reflections to a receiver or when the direct sound is significantly attenuated due to the grazing propagation over an absorbing surface, such as auditorium.

<sup>8</sup>The delay of 100 ms corresponds to the difference in sound path lengths of around 34 m. This is one of the reasons why loudspeakers of public address systems at large events are normally placed at shorter distances from each other.

## 2 Basic equations

In this section we consider the governing equations and derive some of the fundamental equations for sound fields in rooms.

### 2.1 Conservation laws

Sound waves in air involve oscillatory motion of its particles. As such, they obey the governing equations of fluid dynamics:

- conservation of mass

$$\frac{\partial \rho_f}{\partial t} + \nabla \cdot (\rho_f \mathbf{v}_f) = 0 \text{ and} \quad (2.1)$$

- conservation of momentum

$$\frac{\partial}{\partial t}(\rho_f \mathbf{v}_f) + \nabla \cdot (\rho_f \mathbf{v}_f \mathbf{v}_f) + \nabla p_f - \nabla \cdot \tau_f = 0. \quad (2.2)$$

The physical quantities which depend on time ( $t$ ) and location ( $\mathbf{x}$ ) are:  $\rho$  density,  $\mathbf{v}$  velocity,  $p$  pressure, and  $\tau$  viscous stress tensor (which we will neglect in most of the cases). Both  $\tau_f$  and  $\mathbf{v}_f \mathbf{v}_f$  are second-order tensors. The subscript  $f$  stands for the fluid, air in this case. In their most general form, the conservation equations are non-linear and without a closed-form analytical solution.

The calculation is greatly simplified when acoustic waves are the only perturbation of otherwise uniform and steady fluid. We can then write:  $\rho_f = \rho_0 + \rho(\mathbf{x}, t)$ ,  $p_f = p_0 + p(\mathbf{x}, t)$ , and  $\mathbf{v}_f = \mathbf{v}_0 + \mathbf{v}(\mathbf{x}, t)$ , where the subscript  $0$  denotes the steady values in the (background) fluid, which are constant in both time and space, and  $\rho$ ,  $p$ , and  $\mathbf{v}$  are purely acoustic quantities. In the following, we will consider only still air at room temperature ( $T_0 = 293K$ ) as the medium, for which  $\mathbf{v}_0 = 0$  and  $\rho_0 = 1.2\text{kg/m}^3$ . Neglecting viscosity, the two conservation equations become, respectively:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho_0 \mathbf{v} + \rho \mathbf{v}) = 0 \text{ and} \quad (2.3)$$

$$\frac{\partial}{\partial t}(\rho_0 \mathbf{v} + \rho \mathbf{v}) + \nabla \cdot (\rho_0 \mathbf{v} \mathbf{v} + \rho \mathbf{v} \mathbf{v}) + \nabla p = 0. \quad (2.4)$$

Another reasonable assumption is that the acoustic waves are only small perturbations of the steady values. By that, we mean that  $p \ll p_0$ . The static pressure is given by the equation of state of an ideal gas:

$$p_0 = \rho_0 \mathcal{R} T_0, \quad (2.5)$$

where  $\mathcal{R}$  is specific gas constant (for air,  $\mathcal{R} = 287\text{J/kgK}$ ) and  $T_0$  is temperature in kelvins. This gives approximately  $p_0 = 101\text{kPa}$  for air at room temperature. If we

take one percent of this value, 1010Pa, as the maximum amplitude of pressure for small acoustic perturbations, the inequality  $p \ll p_0$  holds whenever the sound pressure level (which will be introduced in eq. (2.29); the factor of  $\sqrt{2}$  is the ratio of amplitude and root mean square value for simple oscillations) is below

$$20 \log_{10} \left( \frac{1010 \text{Pa}/\sqrt{2}}{2 \cdot 10^{-5} \text{Pa}} \right) \approx 151 \text{dB}.$$

This will be satisfied in practically all applications of room acoustics. Moreover, as equations (2.17) and (2.18) will show for homentropic flow of a perfect gas,  $\rho/\rho_0 \sim \mathcal{O}(p/p_0)$ , which implies  $\rho \ll \rho_0$ . The conservation laws can thus be approximated for all but extremely loud sounds (such as in the vicinity of strong sources of sound) as

$$\frac{\partial \rho}{\partial t} + \rho_0 \nabla \cdot \mathbf{v} = 0 \quad \text{and} \quad (2.6)$$

$$\rho_0 \frac{\partial \mathbf{v}}{\partial t} + \rho_0 \nabla \cdot (\mathbf{v} \mathbf{v}) + \nabla p = 0, \quad (2.7)$$

without major loss of accuracy.

As we will see in section 4.4, if we are not very close to the source of sound,  $|\mathbf{v}| \sim \mathcal{O}(p/(\rho_0 c_0)) \sim \mathcal{O}(\rho c_0 / \rho_0) \ll c_0$ , where  $c_0$  is the speed of sound in air. Since the acoustic waves are the only perturbation, we can scale  $\partial/\partial t \sim \omega$  and  $\nabla \sim k = \omega/c_0$ , where  $\omega$  is angular frequency of the sound and  $k$  is wave number. Therefore, the first two terms in eq. (2.7), scale as  $\rho_0 \omega |\mathbf{v}|$  and  $\rho_0 \omega |\mathbf{v}|^2 / c_0$ , respectively, and we can neglect the second term in favour of the first one. This finally gives the two conservation laws for classical acoustics:

- conservation of mass

$$\boxed{\frac{\partial \rho}{\partial t} + \rho_0 \nabla \cdot \mathbf{v} = 0} \quad \text{and} \quad (2.8)$$

- conservation of momentum

$$\boxed{\rho_0 \frac{\partial \mathbf{v}}{\partial t} + \nabla p = 0}. \quad (2.9)$$

These equations are linear. Their linearity is a direct consequence of the assumption of weak acoustic perturbations (for sound levels below around 150dB), adding to the assumption of a uniform, motionless, and homentropic (inviscid and adiabatic) perfect background fluid. All acoustic quantities,  $p$ ,  $\rho$ , and  $\mathbf{v}$ , depend on time  $t$  and location  $\mathbf{x}$ , while the static quantities are constants.

## 2.2 Viscous and thermal losses

In the derivation above, we neglected the viscosity, which is usually acceptable for the bulk of air in a room. However, viscous effects can be significant inside thin boundary

**layers at solid surfaces.** An important example for this are porous materials, which are commonly used as sound absorbers and which, indeed, rely on viscous (and thermal) dissipation of sound energy in their fine structures as the physical mechanism for sound absorption. They will be treated in more details in section 8. Here, we will only try to estimate roughly how thick is the layer of air in which such phenomena are significant.

First we estimate the thickness of a viscous boundary layer (also called Stokes layer),  $\delta_\tau$ , next to a motionless rigid surface. We balance the rate of increase of the momentum from eq. (2.2)

$$\frac{\partial}{\partial t}(\rho_f \mathbf{v}_f) \sim \rho_0 \omega |\mathbf{v}| \quad (2.10)$$

with the viscous term, after applying Stokes' hypothesis for Newtonian fluids<sup>9</sup>, from which it follows:

$$\nabla \cdot \tau_f \sim \frac{1}{\delta_\tau} \mu \frac{1}{\delta_\tau} |\mathbf{v}| = \rho_0 \nu \frac{1}{\delta_\tau^2} |\mathbf{v}|, \quad (2.11)$$

where  $\mu = \rho_0 \nu$  is dynamic viscosity and  $\nu$  is kinematic viscosity of the fluid. The spatial derivatives are scaled as  $1/\delta_\tau$ . For air at room temperature and atmospheric pressure  $\mu \approx 1.8 \cdot 10^{-5} \text{ kg/(ms)}$  and  $\nu \approx 1.5 \cdot 10^{-5} \text{ m}^2/\text{s}$ . Equations (2.10) and (2.11) then give:

$$\delta_\tau(\omega) \sim \mathcal{O}\left(\sqrt{\frac{\nu}{\omega}}\right). \quad (2.12)$$

We see that the viscous boundary layer thickness is inversely proportional to the square root of frequency.

In order to estimate the thickness of thermal boundary layer,  $\delta_T$ , we will use Fourier's law:

$$\mathbf{q} = -K \nabla T, \quad (2.13)$$

where  $\mathbf{q}$  denotes heat flux vector (with the unit  $\text{W/m}^2$ ) due to heat conduction,  $K$  is thermal conductivity ( $K \approx 0.025 \text{ W/(Km)}$  in air at the room temperature and atmospheric pressure), and  $T$  is (unsteady) temperature in kelvins. The rate of increase of energy is given as:

$$\rho_0 C_p \frac{\partial T}{\partial t} \sim \rho_0 C_p \omega T, \quad (2.14)$$

where  $C_p$  is specific heat capacity at constant pressure, which is for air  $C_p \approx 1005 \text{ J/(kgK)}$ , and time derivative is again replaced with angular frequency. This term is balanced by the rate of energy due to the heat conduction:

$$-\nabla \cdot \mathbf{q} = K \nabla \cdot \nabla T \sim \frac{KT}{\delta_T^2}, \quad (2.15)$$

where we now substituted spatial derivatives with  $1/\delta_T$ . From the last two equations it follows that the order of magnitude of thermal boundary layer thickness is:

$$\delta_T(\omega) \sim \mathcal{O}\left(\sqrt{\frac{K}{\rho_0 C_p \omega}}\right), \quad (2.16)$$

---

<sup>9</sup> $\tau_f = \mu \nabla \mathbf{v} + \mu (\nabla \mathbf{v})^T - \frac{2}{3} \mu (\nabla \cdot \mathbf{v}) I$ , where  $I$  is the unit tensor and  $T$  means transpose.

which is also inversely proportional to the square root of frequency. In fact, the thickness of both boundary layers has the same order of magnitude. For example, it is around 0.1mm for frequencies around 1kHz. Such small length scales are irrelevant for sound propagation in rooms. However, they are characteristic length scales for the inner structures of porous materials. As a consequence, major sound attenuation in such materials takes place in the thermal and viscous boundary layers at their solid parts.

## 2.3 Wave equation

Since all static values are given, the two conservation equations (2.8) and (2.9) contain three unknowns (two scalars, acoustic pressure and density, and the velocity vector). An additional equation is necessary in order to close the system of equations. This is the thermodynamic equation of state, which in its linearized form for homentropic flows (with constant entropy, that is, inviscid and adiabatic flows) and with  $\mathbf{v}_0 = 0$  reads:

$$p = c_0^2 \rho. \quad (2.17)$$

In a perfect gas,

$$\rho_0 c_0^2 = \gamma p_0, \quad (2.18)$$

where  $\gamma$  is heat capacity ratio, which can be taken as constant ( $\gamma = 1.4$  for air). This gives the speed of sound in air at room temperature around  $c_0 = 343\text{m/s}$ .

The equations (2.8), (2.9), and (2.17) form a closed system of three equations which can be solved in general. We can combine the equations (2.8) and (2.17) to remove the density<sup>10</sup>. The two remaining equations are

$$\frac{1}{c_0^2} \frac{\partial p}{\partial t} + \rho_0 \nabla \cdot \mathbf{v} = 0 \quad (2.19)$$

and

$$\rho_0 \frac{\partial \mathbf{v}}{\partial t} + \nabla p = 0. \quad (2.20)$$

These can be further simplified to a single scalar equation with acoustic pressure as the only unknown. This is done by differentiating eq. (2.19) with respect to time and subtracting the divergence of eq. (2.20) from it. The result is:

$$\frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} + \rho_0 \frac{\partial}{\partial t} (\nabla \cdot \mathbf{v}) - \rho_0 \nabla \cdot \left( \frac{\partial \mathbf{v}}{\partial t} \right) - \nabla^2 p = 0.$$

Since the time derivative and divergence can switch their positions, the second and third terms cancel, which leaves the wave equation for sound pressure:

$$\frac{1}{c_0^2} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} - \nabla^2 p(\mathbf{x}, t) = 0.$$

(2.21)

<sup>10</sup>Alternatively, we can remove pressure by inserting eq. (2.17) into eq. (2.9). This would lead to the wave equation with acoustic density as the unknown. In general, compressible longitudinal sound waves in fluids are completely described with a single scalar quantity, such as pressure, density, or scalar potential, which all satisfy the wave equation.

## 2.4 Sources and boundary conditions

The sound field in practically whole volume of a room satisfies the derived wave equation. However, a complete solution requires additional information on the sources of sound, as well as the initial<sup>11</sup> and boundary conditions. Appropriate boundary conditions have to be specified for all surfaces in the room, that is for the bounding surfaces of the room (walls, ceiling, and floor), as well as the visible surfaces of all solid objects which are present in the room. This is, of course, one of the major challenges in room acoustic calculations.

Sources of sound are most easily added to eq. (2.21) by introducing a scalar function  $q(\mathbf{y}, \tau)$ , which is non-zero only at the locations of the sources  $\mathbf{y}$  (thus,  $\mathbf{x} = \mathbf{y}$  in the source region) and only at times  $\tau$ , when the sources are active. Alternatively, the sources can be introduced as time-varying boundary conditions. Wave equation with the source term reads

$$\boxed{\frac{1}{c_0^2} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} - \nabla_x^2 p(\mathbf{x}, t) = q(\mathbf{y}, \tau)}. \quad (2.22)$$

Note that we use the subscript  $x$  in the nabla operator (Laplacian) in order to emphasize that the differentiation is performed with respect to the receiver's location  $\mathbf{x}$ , that is,  $\nabla_x = \partial/\partial x_i$  in Cartesian coordinates, where  $i = 1, 2, 3$ . A distinction between  $\mathbf{x}$  and  $\mathbf{y}$  (although they belong to the same physical space inside the room), as well as between  $t$  and  $\tau$ , is made for mathematical clarity, which will become significant later<sup>12</sup>. In the source region, however,

$$\frac{1}{c_0^2} \frac{\partial^2 p(\mathbf{y}, \tau)}{\partial \tau^2} - \nabla_y^2 p(\mathbf{y}, \tau) = q(\mathbf{y}, \tau). \quad (2.23)$$

The source term has the unit  $\text{kg}/(\text{m}^3\text{s}^2)$  and can be associated, for example, with the physical rate of volume injection:

$$q(\mathbf{y}, \tau) = \rho_0 \frac{\partial^2 \beta(\mathbf{y}, \tau)}{\partial \tau^2}, \quad (2.24)$$

where  $\beta$  represents the injected volume fraction (a dimensionless quantity). This is especially useful if the fluid around a compact source is essentially incompressible.

Boundary condition for acoustic pressure can be formulated quite generally as<sup>13</sup>:

$$\boxed{ap(\mathbf{y}, \tau) + b\nabla_y p(\mathbf{y}, \tau) \cdot \mathbf{n}(\mathbf{y}) = c(\mathbf{y}, \tau)}, \quad (2.25)$$

<sup>11</sup>It is usually supposed that there is no sound field before some source in the room is switched on, so we will neglect the trivial initial condition most of the time.

<sup>12</sup>See for example the derivation of eq. (3.46).

<sup>13</sup>We do not consider moving objects in the room. However, the surfaces are generally allowed to oscillate around some defined location.

for any  $\mathbf{y}$  which is located at the surface<sup>14</sup>. Vector  $\mathbf{n}(\mathbf{y})$  represents a unit vector at  $\mathbf{y}$  which is normal to the boundary and pointing into the room,  $a$  is a dimensionless constant,  $b$  is a constant with the unit [m], and  $c$  is a time-dependent function (in Pa). It is non-zero if the surface is active (oscillating even when not excited by an incident sound). For simplicity, we will practically always assume that  $a$ ,  $b$ , and  $c$  do not vary over the surface and that the surface is locally reacting (for more explanations, see the remarks after the definition of impedance in eq. (4.46)).

Particularly important type of a surface is motionless rigid surface, which is also an idealization of an acoustically hard wall. At such a surface normal component of velocity is zero:

$$\mathbf{v}(\mathbf{y}, \tau) \cdot \mathbf{n}(\mathbf{y}) = 0. \quad (2.26)$$

From the scalar product of eq. (2.9) with time-independent  $\mathbf{n}(\mathbf{y})$ , it follows (after replacing  $\mathbf{x}$  with  $\mathbf{y}$  and  $t$  with  $\tau$ ):

$$\rho_0 \frac{\partial}{\partial t} (\mathbf{v}(\mathbf{y}, \tau) \cdot \mathbf{n}(\mathbf{y})) + \nabla_y p(\mathbf{y}, \tau) \cdot \mathbf{n}(\mathbf{y}) = \nabla_y p(\mathbf{y}, \tau) \cdot \mathbf{n}(\mathbf{y}) = 0. \quad (2.27)$$

Therefore, the coefficient  $a$  in eq. (2.25) equals zero and  $c(\mathbf{y}, \tau) = 0$  for any  $b \neq 0$  (we can set  $b = 1\text{m}$ ) for this particular surface.

## 2.5 Sound energy and sound pressure level

As already discussed in the previous section, we usually consider human listener as the receiver of a sound in a room. Roughly speaking, human auditory system perceives the energy of the sound averaged over certain time interval, rather than the instantaneous values of sound pressure. For this reason, the oscillating sound pressure of a sustained pure tone is not perceived as time-varying, but as a continuous sound with constant loudness. This also suggests that it is not the sound pressure which is averaged (which would result in the constant value of  $p_0$  for acoustic perturbations as the only unsteady component) but its squared value, which is a second-order quantity.

Such averaging can be mathematically expressed as:

$$\sqrt{\frac{1}{T_{avg}} \int_{-T_{avg}/2}^{T_{avg}/2} p^2 dt} = \sqrt{\langle p^2 \rangle_{T_{avg}}}, \quad (2.28)$$

where  $T_{avg}$  is the time interval over which sound pressure is averaged. The right-hand side of the equation is simply a more compact notation. This expression defines a root mean square (RMS) value. The square root in it retains pascal as the unit, but the RMS value is strictly positive and does not correspond to the instantaneous pressure value or its absolute value. For simple tones, the averaging interval can be conveniently set

---

<sup>14</sup>We use the same variable  $\mathbf{y}$  for locations inside the source region and at the boundaries. This is acceptable since we are actually interested in the sound field at the locations inside the room  $\mathbf{x} \neq \mathbf{y}$ , so the source region and boundaries will be addressed quite similarly (as, for example, in eq. (3.46)).

to one period of the oscillations, or, what is more often in practice, many periods. For transient (impulse) sounds, it is difficult to define a single universally applicable value. However, it is usually of the **order of magnitude 10ms or longer** (see also the discussion in section 1), for example, S(low) 1s, F(ast) 125ms, or I(mpulse) 35ms.

Sound pressure level is defined as the logarithm with the base 10 of the RMS value:

$$L = 20 \log_{10} \left( \frac{\sqrt{\langle p^2 \rangle_{T_{avg}}}}{2 \cdot 10^{-5} \text{Pa}} \right) = 10 \log_{10} \left( \frac{1}{T_{avg}} \int_{-T_{avg}/2}^{T_{avg}/2} \frac{p^2}{4 \cdot 10^{-10} \text{Pa}^2} dt \right) \quad (2.29)$$

The reference value  $2 \cdot 10^{-5} \text{Pa}$  (which is not equal to  $p_0$ ) is chosen such that it closely matches the threshold of human hearing at middle frequencies. The unit of sound pressure level is decibel (dB).

Squared values of the basic (first-order) acoustic quantities ( $p, \rho, \mathbf{v}$ ) are closely related to sound energy. We can multiply eq. (2.19) with  $p/\rho_0$  and eq. (2.20) with  $\mathbf{v}$  (the scalar product). Sum of the two equations gives the conservation law for acoustic energy:

$$\begin{aligned} \frac{p}{\rho_0 c_0^2} \frac{\partial p}{\partial t} + p \nabla \cdot \mathbf{v} + \mathbf{v} \cdot \rho_0 \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla p &= \frac{1}{2\rho_0 c_0^2} \frac{\partial p^2}{\partial t} + \frac{\rho_0}{2} \frac{\partial(\mathbf{v} \cdot \mathbf{v})}{\partial t} + \nabla \cdot (p\mathbf{v}) \\ &= \frac{\partial}{\partial t} \left[ \frac{p^2}{2\rho_0 c_0^2} + \frac{\rho_0 |\mathbf{v}|^2}{2} \right] + \nabla \cdot (p\mathbf{v}) = \frac{\partial E}{\partial t} + \nabla \cdot \mathbf{I} = 0. \end{aligned} \quad (2.30)$$

Thereby, we defined sound intensity (rate of energy flux) as the vector:

$$\mathbf{I} = p\mathbf{v}, \quad (2.31)$$

with the unit  $\text{W/m}^2$ . Acoustic energy<sup>15</sup> is defined as:

$$E = E_{pot} + E_{kin} = \frac{p^2}{2\rho_0 c_0^2} + \frac{\rho_0 |\mathbf{v}|^2}{2}, \quad (2.32)$$

where the first term represents potential energy and the second term is kinetic energy.

Although both sound energy and intensity are obviously second-order acoustic quantities (the relation between acoustic pressure and velocity, eq. (2.9), is linear), they do not involve any time averaging by definition<sup>16</sup>. In order to associate them with sound pressure level or RMS value, averaging over time must be performed, which will be done in the following section. We can also express instantaneous (in contrast to time-averaged) acoustic power output of a source:

$$P_q(t) = \int_V \nabla \cdot \mathbf{I}(\mathbf{x}, t) d^3x = \oint_S \mathbf{I}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) d^2x = \oint_S p(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) d^2x, \quad (2.33)$$

<sup>15</sup>Strictly speaking, it is energy density with the unit  $[\text{J/m}^3]$ . However, it is most often simply referred to as acoustic energy.

<sup>16</sup>However, a caution is necessary, because many authors do use the terms “energy” and “intensity” for their time-averaged values and this is often implicit in the quantities  $E$  and  $\mathbf{I}$ . In order to avoid confusion, we will be explicit and any averaging shall be denoted with angle brackets.

where  $S$  is any closed surface which completely encloses the source and  $\mathbf{n}$  is unit vector normal to it and pointing outwards. We used the divergence theorem to switch from the volume to the surface integral.

## 3 Modal analysis

When all openings of the room are excluded, or their effect is negligible, rooms can be treated as fully closed cavities<sup>17</sup>. Even partly open spaces are expected to have more or less pronounced resonances at certain frequencies (the eigenfrequencies). At the same time, linearity of the governing equations allows very efficient implementation of the Fourier analysis on the sound fields. This means that the fields can be decomposed into simple sine waves with different frequencies and the time dependence can be replaced with frequency. This is demonstrated in the following, as the basis for the modal analysis of closed spaces.

### 3.1 Complex sine waves

We will now try solving the homogeneous wave equation (2.21) (the effects of boundary conditions and sources will be added later). For this, we will apply the usual method of separation of variables. We suppose a solution in the form

$$p(\mathbf{x}, t) = \hat{p}(\mathbf{x})\tilde{p}(t), \quad (3.1)$$

where  $\hat{p}(\mathbf{x})$  depends only on location<sup>18</sup> and  $\tilde{p}(t)$  depends only on time. Then, if  $p \neq 0$ :

$$\frac{\hat{p}(\mathbf{x})}{c_0^2} \frac{d^2\tilde{p}(t)}{dt^2} - \tilde{p}(t)\nabla^2\hat{p}(\mathbf{x}) = 0 \Rightarrow \frac{1}{c_0^2\tilde{p}(t)} \frac{d^2\tilde{p}(t)}{dt^2} = \frac{1}{\hat{p}(\mathbf{x})}\nabla^2\hat{p}(\mathbf{x}). \quad (3.2)$$

The independent variables are separated on the two sides of the last equality, which can thus be satisfied only if they are both equal to some real constant  $C$  (with the unit  $1/\text{m}^2$ ).

The time dependence gives

$$\frac{d^2\tilde{p}(t)}{dt^2} = Cc_0^2\tilde{p}(t). \quad (3.3)$$

This equation has the solution

$$\tilde{p}(t) = A \cos(\omega t) + B \sin(\omega t), \quad (3.4)$$

with  $A$  and  $B$  some real constants and, after replacing  $\tilde{p}(t)$  in eq. (3.3),  $C = -(\omega/c_0)^2 = -k^2$ . Here,  $\omega$  is angular frequency, which is related to frequency  $f$ , period  $T = 1/f$ , and wavelength  $\lambda = c_0T$  by the equalities:

$$\omega = 2\pi f = 2\pi/T = 2\pi c_0/\lambda. \quad (3.5)$$

<sup>17</sup>We also neglect any interaction between the sound field inside the room and the exterior. Relations between the interior and exterior sound fields and sound insulation are the object of building acoustics.

<sup>18</sup>More precisely, it is sufficient to assume that it changes slowly enough over time, so that the essential time dependence is inside the exponential term of  $\tilde{p}(t) \sim e^{j\omega t}$ . See the comment after eq. (3.27) for more details. Note also that no complex quantities have been introduced yet and that the units of  $\hat{p}$  and  $\tilde{p}$  are not important as long as their product is expressed in pascals.

For future use, we also introduced the wave number:

$$k = \omega/c_0 = 2\pi f/c_0 = 2\pi/(c_0 T) = 2\pi/\lambda. \quad (3.6)$$

All the quantities which appear in eq. (3.5) and eq. (3.6) are larger than zero. Inserting the value of  $C$  in the spatial part of eq. (3.2) gives:

$$\nabla^2 \hat{p}(\mathbf{x}) + k^2 \hat{p}(\mathbf{x}) = 0, \quad (3.7)$$

which is exactly the **Helmholtz equation** (3.30) which will be derived below, only without the source term and with real  $\hat{p}(\mathbf{x})$ .

The two terms in the solution in eq. 3.4 involve two constants,  $A$  and  $B$ , which in general depend on the source and initial conditions. They can be written more compactly if we let  $p$  (and, consequently, all other first-order acoustic quantities<sup>19</sup> which describe the sound field, sources, or boundary/initial conditions in the linearized theory) take complex values. First we note that the simple oscillatory functions  $\sin(\omega t)$  and  $\cos(\omega t)$  represent the same solution with the phase difference  $\pi/2$  between them. Then we use Euler's formula,

$$e^{j\omega t} = \cos(\omega t) + j \sin(\omega t), \quad (3.8)$$

in order to write<sup>20</sup>:

$$\tilde{p}(t) = D e^{j\omega t}, \quad (3.9)$$

where  $D$  is now a complex factor. Its real and imaginary part entirely substitute the two degrees of freedom of  $A$  and  $B$ .

Since the entire solution we are interested in is the product  $\hat{p}(\mathbf{x})\tilde{p}(t)$  in eq. (3.1), where  $\hat{p}(\mathbf{x})$  is now also complex in general, we can absorb the factor  $D$  into  $\hat{p}(\mathbf{x})$  and keep only the dimensionless exponential term as the essentially time-dependent part of the solution. The solution of eq. (2.21) is thus the product of time-independent (or slowly varying) complex amplitude  $\tilde{p}(\mathbf{x})$  and the factor  $e^{j\omega t}$ :

$$p(\mathbf{x}, t) = \hat{p}(\mathbf{x}) e^{j\omega t}. \quad (3.10)$$

Consequently, the other first-order quantities take the forms:  $\mathbf{v}(\mathbf{x}, t) = \hat{\mathbf{v}}(\mathbf{x}) e^{j\omega t}$ ,  $\rho(\mathbf{x}, t) = \hat{\rho}(\mathbf{x}) e^{j\omega t}$ , and  $q(\mathbf{y}, \tau) = \hat{q}(\mathbf{y}) e^{j\omega \tau}$ , the last one being a complex source function. Note

<sup>19</sup>Since only the real part is physical, the second-order quantities such as  $E$ ,  $\mathbf{I}$ , and  $P_q$  remain real-valued (see below), as well as all the independent variables and constants:  $t$ ,  $\tau$ ,  $\mathbf{x}$ ,  $\mathbf{y}$ ,  $\omega$ ,  $f$ ,  $T$ ,  $\lambda$ ,  $k$ ,  $c_0$ ,  $\rho_0$ , etc., at least for now. The unit vector  $\mathbf{n}$  in the boundary condition, eq. (2.25), also remains real, but  $a$ ,  $b$ , and  $c$  can take complex values.

<sup>20</sup>Some authors use  $e^{-j\omega t} = \cos(\omega t) - j \sin(\omega t)$  for the time dependence, which also obeys eq. (2.21), because its double time derivative cancels the factor  $-1$  in the exponent. Such complex conjugate solutions can be included in our notation when we let  $\omega$  take negative values (which will reflect in the bounds of the Fourier integral in eq. (3.22)). However, this does not bring any new physical solutions, which are only the real part of  $p$  (see also footnote <sup>24</sup>), and we can still consider  $\omega$  to be strictly positive. Nevertheless, the adopted convention has to be made clear for accurate mathematical treatment.

that the magnitude of a complex vector is real, for example  $|\mathbf{v}| = \sqrt{\mathbf{v}^* \cdot \mathbf{v}}$ , where  $\mathbf{v}^* = \mathcal{R}_e(\mathbf{v}) - j\mathcal{I}_m(\mathbf{v})$  is complex conjugate of  $\mathbf{v}^*$ . An important property of complex exponential functions is that they are mutually orthogonal<sup>21</sup>:

$$\int_{-\infty}^{\infty} e^{j\omega_1 t} (e^{j\omega_2 t})^* dt = \int_{-\infty}^{\infty} e^{j\omega_1 t} e^{-j\omega_2 t} dt = \int_{-\infty}^{\infty} e^{j(\omega_1 - \omega_2)t} dt = 2\pi\delta(\omega_1 - \omega_2), \quad (3.11)$$

which is equal to zero whenever  $\omega_1 \neq \omega_2$ . Here we used the equality

$$\int_{-\infty}^{\infty} e^{j\omega t} dt = 2\pi\delta(\omega). \quad (3.12)$$

The complex amplitude  $\hat{p}(\mathbf{x})$  (in pascals) will be treated further below. For now, it is sufficient to say that continuing separation of the three spatial variables in eq. (3.7) is feasible only when all space-dependent variables (including the boundary conditions) can be decoupled in terms of the components of  $\mathbf{x}$ . This is the case only with simple geometries of rooms, such as the rectangular room which will be treated in section 3.6.

Following from eq. (3.10), we should mention that we use the same symbol  $p$  for complex sound pressure and actual “physical” sound pressure. This shall not cause a confusion, since it is understood that the latter one can always be expressed as the real part of the former one:

$$\begin{aligned} \mathcal{R}_e(p) &= \mathcal{R}_e(\hat{p}e^{j\omega t}) = \mathcal{R}_e[(\mathcal{R}_e(\hat{p}) + j\mathcal{I}_m(\hat{p}))(\cos(\omega t) + j\sin(\omega t))] \\ &= \mathcal{R}_e[\mathcal{R}_e(\hat{p})\cos(\omega t) + j\mathcal{R}_e(\hat{p})\sin(\omega t) + j\mathcal{I}_m(\hat{p})\cos(\omega t) - \mathcal{I}_m(\hat{p})\sin(\omega t)] \quad (3.13) \\ &= \mathcal{R}_e(\hat{p})\cos(\omega t) - \mathcal{I}_m(\hat{p})\sin(\omega t). \end{aligned}$$

Hence, no particular notation is needed in order to distinguish between real and complex  $p$ ,  $\mathbf{v}$ ,  $\rho$ , or  $q$ . In fact, we often refer to complex sound pressure simply as sound pressure, complex amplitude as amplitude, complex sine wave as sine wave, and so on. Moreover, all the linear equations which we have derived so far hold for the complex acoustic variables, as well, since they do not involve a product of two first-order quantities.

In contrast to this, the energy-based, second-order quantities require more care. For example, intensity is by definition in eq. (2.31) equal to

$$\mathbf{I} = \mathcal{R}_e(p)\mathcal{R}_e(\mathbf{v}), \quad (3.14)$$

This is in general not equal to the product  $p\mathbf{v}$ , if  $p$  and  $\mathbf{v}$  are complex. Fortunately, we are most often interested only in time-averaged energy (similarly as in the case of sound

---

<sup>21</sup>Since the exponential functions are complex, the integral has to contain a product with the complex conjugate value in order to return a real value. It can be compared with the result of time averaging in eq. (3.15), the derivation of which also involves integration over time or with eq. (3.54), which is an integral over space. Dirac delta function  $\delta$  and its properties will be introduced in eq. (3.42) and further.

pressure level in eq. (2.29)) or its flux. The latter one is for simple acoustic oscillations equal to the time average of sound intensity over one period<sup>22</sup>  $T = 2\pi/\omega$ :

$$\begin{aligned}
 \langle \mathbf{I} \rangle_T &= \frac{1}{T} \int_{-T/2}^{T/2} \mathbf{I} dt = \frac{1}{T} \int_{-T/2}^{T/2} \mathcal{R}_e(p) \mathcal{R}_e(\mathbf{v}) dt \\
 &= \frac{1}{T} \int_{-T/2}^{T/2} [\mathcal{R}_e(\hat{p}) \cos(\omega t) - \mathcal{I}_m(\hat{p}) \sin(\omega t)] [\mathcal{R}_e(\hat{\mathbf{v}}) \cos(\omega t) - \mathcal{I}_m(\hat{\mathbf{v}}) \sin(\omega t)] dt \\
 &= \frac{1}{T} \int_{-T/2}^{T/2} \{ \mathcal{R}_e(\hat{p}) \mathcal{R}_e(\hat{\mathbf{v}}) \cos^2(\omega t) - \mathcal{I}_m(\hat{p}) \mathcal{R}_e(\hat{\mathbf{v}}) \sin(\omega t) \cos(\omega t) \\
 &\quad - \mathcal{R}_e(\hat{p}) \mathcal{I}_m(\hat{\mathbf{v}}) \sin(\omega t) \cos(\omega t) + \mathcal{I}_m(\hat{p}) \mathcal{I}_m(\hat{\mathbf{v}}) \sin^2(\omega t) \} dt \\
 &= \frac{1}{T} \mathcal{R}_e(\hat{p}) \mathcal{R}_e(\hat{\mathbf{v}}) \int_{-T/2}^{T/2} \cos^2(\omega t) dt + \frac{1}{T} \mathcal{I}_m(\hat{p}) \mathcal{I}_m(\hat{\mathbf{v}}) \int_{-T/2}^{T/2} \sin^2(\omega t) dt \\
 &= \frac{1}{T} \mathcal{R}_e(\hat{p}) \mathcal{R}_e(\hat{\mathbf{v}}) \left( \frac{2\omega T/2 + \sin(2\omega T/2)}{4\omega} - \frac{-2\omega T/2 + \sin(-2\omega T/2)}{4\omega} \right) \\
 &\quad + \frac{1}{T} \mathcal{I}_m(\hat{p}) \mathcal{I}_m(\hat{\mathbf{v}}) \left( \frac{2\omega T/2 - \sin(2\omega T/2)}{4\omega} - \frac{-2\omega T/2 - \sin(-2\omega T/2)}{4\omega} \right) \\
 &= \frac{1}{T} \mathcal{R}_e(\hat{p}) \mathcal{R}_e(\hat{\mathbf{v}}) \frac{T}{2} + \frac{1}{T} \mathcal{I}_m(\hat{p}) \mathcal{I}_m(\hat{\mathbf{v}}) \frac{T}{2} = \frac{1}{2} [\mathcal{R}_e(\hat{p}) \mathcal{R}_e(\hat{\mathbf{v}}) + \mathcal{I}_m(\hat{p}) \mathcal{I}_m(\hat{\mathbf{v}})].
 \end{aligned}$$

We used the fact that  $\sin(\omega t) \cos(\omega t)$  is an odd function of  $t$ , the integral of which over the symmetric interval  $[-T/2, T/2]$  vanishes. The last result can be written shorter as

$$\boxed{\langle \mathbf{I} \rangle_T = \frac{1}{2} \mathcal{R}_e(\hat{p}^* \hat{\mathbf{v}}) = \frac{1}{4} (\hat{p}^* \hat{\mathbf{v}} + \hat{p} \hat{\mathbf{v}}^*)}. \quad (3.15)$$

Sustained acoustic power of a source can be expressed from eq. (2.33):

$$\begin{aligned}
 \langle P_q \rangle_T &= \frac{1}{T} \int_{-T/2}^{T/2} \oint_S \mathbf{I}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) d^2 \mathbf{x} dt = \oint_S \langle \mathbf{I}(\mathbf{x}, t) \rangle_T \cdot \mathbf{n}(\mathbf{x}) d^2 \mathbf{x} \\
 &= \frac{1}{2} \oint_S \mathcal{R}_e(\hat{p}^* \hat{\mathbf{v}}) \cdot \mathbf{n}(\mathbf{x}) d^2 \mathbf{x} = \frac{1}{4} \oint_S (\hat{p}^* \hat{\mathbf{v}} + \hat{p} \hat{\mathbf{v}}^*) \cdot \mathbf{n}(\mathbf{x}) d^2 \mathbf{x},
 \end{aligned} \quad (3.16)$$

for a fixed closed surface  $S$ . Analogously to the sound pressure level, sound power level is defined as

$$\boxed{L_W = 10 \log_{10} \frac{\langle P_q \rangle_T}{10^{-12} \text{W}}}. \quad (3.17)$$

The reference value  $10^{-12} \text{W}$  follows from the value  $4 \cdot 10^{-10} \text{Pa}^2$  in eq. (2.29) divided with  $\rho_0 c_0 \approx 400 \text{kg}/(\text{m}^2 \text{s})$ , which will be identified as the characteristic impedance of air in eq. (4.45).

<sup>22</sup>We assume constant amplitude of the sound during this time interval.

Sound energy can be found from eq. (2.32),

$$E = \frac{\mathcal{R}_e(p)^2}{2\rho_0 c_0^2} + \frac{\rho_0 |\mathcal{R}_e(\mathbf{v})|^2}{2}, \quad (3.18)$$

and after averaging over one period:

$$\begin{aligned} \langle E \rangle_T &= \frac{1}{2\rho_0 c_0^2 T} \frac{1}{T} \int_{-T/2}^{T/2} [\mathcal{R}_e(\hat{p}) \cos(\omega t) - \mathcal{I}_m(\hat{p}) \sin(\omega t)]^2 dt \\ &\quad + \frac{\rho_0}{2} \frac{1}{T} \int_{-T/2}^{T/2} [\mathcal{R}_e(\hat{\mathbf{v}}) \cos(\omega t) - \mathcal{I}_m(\hat{\mathbf{v}}) \sin(\omega t)]^2 dt \\ &= \frac{\mathcal{R}_e(\hat{p})^2}{2\rho_0 c_0^2 T} \int_{-T/2}^{T/2} \cos^2(\omega t) dt - \frac{2\mathcal{R}_e(\hat{p})\mathcal{I}_m(\hat{p})}{2\rho_0 c_0^2 T} \int_{-T/2}^{T/2} \sin(\omega t) \cos(\omega t) dt \\ &\quad + \frac{\mathcal{I}_m(\hat{p})^2}{2\rho_0 c_0^2 T} \int_{-T/2}^{T/2} \sin^2(\omega t) dt + \frac{\rho_0 \mathcal{R}_e(\hat{\mathbf{v}})^2}{2T} \int_{-T/2}^{T/2} \cos^2(\omega t) dt \\ &\quad - \frac{2\rho_0 \mathcal{R}_e(\hat{\mathbf{v}})\mathcal{I}_m(\hat{\mathbf{v}})}{2T} \int_{-T/2}^{T/2} \sin(\omega t) \cos(\omega t) dt + \frac{\rho_0 \mathcal{I}_m(\hat{\mathbf{v}})^2}{2T} \int_{-T/2}^{T/2} \sin^2(\omega t) dt \\ &= \frac{\mathcal{R}_e(\hat{p})^2 T}{2\rho_0 c_0^2 T} + \frac{\mathcal{I}_m(\hat{p})^2 T}{2\rho_0 c_0^2 T} + \frac{\rho_0 \mathcal{R}_e(\hat{\mathbf{v}})^2 T}{2T} + \frac{\rho_0 \mathcal{I}_m(\hat{\mathbf{v}})^2 T}{2T} \\ &= \frac{\mathcal{R}_e(\hat{p})^2 + \mathcal{I}_m(\hat{p})^2}{4\rho_0 c_0^2} + \frac{\rho_0 (\mathcal{R}_e(\hat{\mathbf{v}})^2 + \mathcal{I}_m(\hat{\mathbf{v}})^2)}{4}. \end{aligned}$$

This can also be written as

$$\langle E \rangle_T = \frac{|\hat{p}|^2}{4\rho_0 c_0^2} + \frac{\rho_0 |\hat{\mathbf{v}}|^2}{4} = \frac{\hat{p}\hat{p}^*}{4\rho_0 c_0^2} + \frac{\rho_0 \hat{\mathbf{v}} \cdot \hat{\mathbf{v}}^*}{4}. \quad (3.19)$$

Finally, sound pressure level from eq. (2.29) becomes

$$L = 10 \log_{10} \left( \frac{1}{T} \int_{-T/2}^{T/2} \frac{\mathcal{R}_e(p)^2}{4 \cdot 10^{-10} \text{Pa}^2} dt \right), \quad (3.20)$$

which gives

$$L = 10 \log_{10} \left( \frac{|\hat{p}|^2/2}{4 \cdot 10^{-10} \text{Pa}^2} \right) = 20 \log_{10} \left( \frac{|\hat{p}|/\sqrt{2}}{2 \cdot 10^{-5} \text{Pa}} \right). \quad (3.21)$$

Note that sound pressure level is usually expressed in terms of the root mean square value, which is for simple oscillations  $|\hat{p}|/\sqrt{2}$  and the division with square root of two in the last equality drops out.

The solution in eq. (3.10) is given for a fixed angular frequency. More generally, any continuous complex function  $p(\mathbf{x}, t)$  (or some other acoustic variable replacing  $p$ ) can

be written as a superposition of (in general infinite number) of complex sine functions. Mathematically this follows from Fourier's theory and it is expressed by the integral:

$$p(\mathbf{x}, t) = \int_{-\infty}^{\infty} \hat{p}(\mathbf{x}, \omega) e^{j\omega t} d\omega = \mathcal{F}_{\omega}^{-1}\{\hat{p}(\mathbf{x}, \omega)\}, \quad (3.22)$$

where  $\hat{p}(\mathbf{x}, \omega) d\omega$  takes the role of  $\hat{p}(\mathbf{x})$  in eq. (3.10)<sup>23</sup>. In other words, an acoustic quantity is entirely represented by its spectrum. The complex amplitude  $\hat{p}(\mathbf{x}, \omega)$  can thus vary with frequency but does not depend on time. It can also incorporate any phase shift of the complex sine wave  $p(\mathbf{x}, t)$ , for example,  $p = \hat{p}e^{j(\omega t + \Phi)} = (\hat{p}e^{j\Phi})e^{j\omega t}$ , where  $\Phi$  can be defined, for example, by initial conditions. In general, the spatially-dependent amplitude will depend on the sources and boundary conditions. Hence, by utilizing simple time dependence  $e^{j\omega t}$ , we leave all the effects of boundary/initial conditions and the sources to determine the complex amplitude.

Integral in eq. (3.22) is known as Fourier integral. In fact, eq. (3.22) represents inverse Fourier transform of  $\hat{p}$  with  $\omega$  as the parameter (hence the notation  $\mathcal{F}_{\omega}^{-1}\{\hat{p}\}$ ). Fourier transform reads

$$\hat{p}(\mathbf{x}, \omega) = \mathcal{F}_t\{p(\mathbf{x}, t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} p(\mathbf{x}, t) e^{-j\omega t} dt, \quad (3.23)$$

with time  $t$  as the parameter of the transform<sup>24</sup>. Additional factor  $1/(2\pi)$  is necessary in order to satisfy the equality:

$$\begin{aligned} p(\mathbf{x}, t) &= \int_{-\infty}^{\infty} \hat{p}(\mathbf{x}, \omega) e^{j\omega t} d\omega = \int_{-\infty}^{\infty} \frac{1}{2\pi} \left( \int_{-\infty}^{\infty} p(\mathbf{x}, t') e^{-j\omega t'} dt' \right) e^{j\omega t} d\omega \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi} p(\mathbf{x}, t') \left( \int_{-\infty}^{\infty} e^{j\omega(t-t')} d\omega \right) dt' = \int_{-\infty}^{\infty} \frac{1}{2\pi} p(\mathbf{x}, t') 2\pi \delta(t - t') dt' \\ &= \int_{-\infty}^{\infty} p(\mathbf{x}, t') \delta(t' - t) dt' = p(\mathbf{x}, t). \end{aligned}$$

The independent time variable  $t'$  is introduced here only to distinguish it from  $t$ . We also used eq. (3.12) with  $t$  and  $\omega$  switched:

$$\int_{-\infty}^{\infty} e^{j\omega t} d\omega = 2\pi \delta(t). \quad (3.24)$$

<sup>23</sup>Strictly speaking, the unit of  $\hat{p}(\mathbf{x}, \omega)$  is Pa·s and  $\hat{p}(\mathbf{x}, \omega)$  is density of  $\hat{p}(\mathbf{x})$  over  $\omega$ . However, this distinction does not have far-reaching implications and the use of the same symbol (here  $\hat{p}$ ) is common in literature. The integral can be observed as an extension of the sum of simple sinusoidal components from eq. (3.10) with different frequencies.

<sup>24</sup>Fourier transform introduces negative angular frequencies. However, since we are interested in real  $p(\mathbf{x}, t)$ , then  $\hat{p}(\mathbf{x}, \omega) = \mathcal{F}_t\{p(\mathbf{x}, t)\} = \hat{p}^*(\mathbf{x}, -\omega)$  and no new information on the sound field is introduced by these frequencies.

### 3.2 Acoustics in frequency domain

Since practically all equations we are dealing with are linear (Fourier transform is also a linear operation), we can observe a single component from eq. (3.10) with generic angular frequency  $\omega$  without any loss of generality. This is possible because a given angular frequency  $\omega$  does not change in linear systems:

$$A\hat{p}_1(\mathbf{x})e^{j\omega t} + B\hat{p}_2(\mathbf{x})e^{j\omega t} = (A\hat{p}_1(\mathbf{x}) + B\hat{p}_2(\mathbf{x}))e^{j\omega t}, \quad (3.25)$$

where  $A$  and  $B$  are some constant complex factors. The spectral components at different frequencies remain independent from each other and do not affect each other's frequencies. Consequently, we can work with  $\hat{p}(\mathbf{x})$  instead of  $\hat{p}(\mathbf{x}, \omega)$  from eq. (3.22) for general considerations in a linearized theory (which again justifies the use of the same symbol  $\hat{p}$ ). Dependence of the amplitude and other quantities on frequency and integration over frequency are implicit. In this way, time-independent equations are easily obtained by dividing the corresponding equations with time dependence with  $e^{j\omega t}$ , which is not affected by the linear operations. As an example, boundary condition for the pressure amplitude  $\hat{p}(\mathbf{x})$  is obtained by dividing eq. (2.25) (with complex  $p$ ,  $a$ ,  $b$ , and  $c$ ) with  $e^{j\omega\tau}$ :

$$\boxed{a\hat{p}(\mathbf{y}) + b\nabla_y\hat{p}(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y}) = \hat{c}(\mathbf{y})}, \quad (3.26)$$

where  $\mathbf{y}$  is location on the surface and  $\mathbf{n}(\mathbf{y})$  is real, as before. It is understood that  $a$ ,  $b$ , and  $c$ , as well as  $\hat{p}$  can depend on frequency.

Both  $p(\mathbf{x}, t)$  and  $\hat{p}(\mathbf{x})$  have the same physical unit, pascal. The symbol  $\hat{\phantom{p}}$  thus only implies that the variable is observed in the frequency (rather than time) domain. At a specific frequency, it depends only on the location  $\mathbf{x}$ . The form of eq. (3.10) clearly separates time and spatial dependence. In the cases when the amplitudes are allowed to (slowly) vary in time, it shall be clearly stated so.

An important property of complex sine functions, for example complex pressure, is that their time derivative is very simple:

$$\frac{\partial p}{\partial t} = \left( \frac{\partial \hat{p}}{\partial t} \right) e^{j\omega t} + \hat{p} \left( \frac{\partial}{\partial t} e^{j\omega t} \right) = \hat{p} \left( \frac{\partial}{\partial t} e^{j\omega t} \right) = j\omega \hat{p} e^{j\omega t} = j\omega p. \quad (3.27)$$

This equality of the derivation with respect to time and the product with  $j\omega$  is one of the main reasons for introducing complex quantities at the first place, in eq. (3.8). Note that we supposed that  $|e^{j\omega t} \partial \hat{p} / \partial t| \ll |\hat{p} \partial e^{j\omega t} / \partial t| = |j\omega \hat{p} e^{j\omega t}|$ . This is justified when any characteristic time scale  $2\pi\hat{p}/(\partial \hat{p} / \partial t)$  over which  $\hat{p}$  might vary appreciably (if we indeed allow  $\hat{p}$  from eq. (3.1) to vary with time to a certain extent) is much larger than the period of oscillations  $T = 1/f = 2\pi/\omega$ .

We can use the derivative from eq. (3.27) to completely remove time dependence from the wave equation. The second-order time derivative equals

$$\frac{\partial^2 p}{\partial t^2} = j\omega \frac{\partial p}{\partial t} = (j\omega)^2 p = -\omega^2 p. \quad (3.28)$$

If we now insert this in the wave equation (2.22), we obtain:

$$k^2 p(\mathbf{x}, t) + \nabla_x^2 p(\mathbf{x}, t) = -q(\mathbf{y}, \tau), \quad (3.29)$$

where  $k$  is defined in eq. (3.6). Writing the source term as  $q(\mathbf{y}, \tau) = \hat{q}(\mathbf{y})e^{j\omega\tau}$  and dividing both sides with the factor  $e^{j\omega\tau}$ , which is not a function of space, we obtain the Helmholtz equation with the source term:

$$k^2 \hat{p}(\mathbf{x}) + \nabla_x^2 \hat{p}(\mathbf{x}) = -\hat{q}(\mathbf{y}). \quad (3.30)$$

Complex amplitude  $\hat{p}(\mathbf{x})$  contains also the phase shift due to the wave propagation from the source to the receiver,  $e^{j\omega(t-\tau)} = e^{-jk|\mathbf{x}-\mathbf{y}|} = e^{-j\Delta\Phi}$ . Together with initial conditions and boundary conditions (eq. (3.26)), which all reflect in the value of complex amplitude, this closes the time-independent acoustic problem. According to eq. (2.24), the complex source function can also be sometimes modelled as

$$\hat{q}(\mathbf{y}) = -\rho_0 \omega^2 \hat{\beta}(\mathbf{y}). \quad (3.31)$$

Similarly as with the equations (2.22) and (2.23), eq. (3.30) can be formally written in the source region:

$$k^2 \hat{p}(\mathbf{y}) + \nabla_y^2 \hat{p}(\mathbf{y}) = -\hat{q}(\mathbf{y}). \quad (3.32)$$

### 3.3 Damping

In order to obtain a simple expression for time derivative in eq. (3.27), we had to suppose that the sound amplitude varies slowly. The amplitude of a sound wave in a room changes over time due to damping. If we suppose that  $\hat{p}$  decays exponentially in time<sup>25</sup>, starting with its maximum value  $\hat{p}_{max}$ :

$$\hat{p}(\mathbf{x}) = \hat{p}_{max}(\mathbf{x}) e^{-\zeta t}, \quad (3.33)$$

with  $\zeta$  denoting real and positive damping constant<sup>26</sup> with the unit 1/s, then the condition from above leads to  $|e^{j\omega t} \partial \hat{p} / \partial t| = |e^{j\omega t} (-\zeta) \hat{p}_{max} e^{-\zeta t}| \ll |j\omega \hat{p}_{max} e^{-\zeta t} e^{j\omega t}|$ . In other words:

$$\zeta \ll \omega. \quad (3.34)$$

We refer to this as a condition for low damping, for example, a weakly damped room. It can also be formulated such that the amplitude at each frequency of interest does not change significantly during one period or along one wavelength of the sine wave. The condition is normally satisfied for reasonably low attenuation of sound in air or absorbing materials, as will be shown shortly. This justifies leaving out time as an argument of the complex amplitude in eq. (3.10).

<sup>25</sup>A reasonable assumption in many cases of practical importance, as it will turn out later.

<sup>26</sup>Which can, of course, like  $\hat{p}(\mathbf{x})$ , vary with frequency.

Largely for historical reasons which date back to the works of W. C. Sabine (the founder of room acoustics), the main quantity which is used for expressing overall damping in rooms is reverberation time  $T_{60}$ . It is defined as the time interval it takes sound energy in a room to decay 60dB after a source which produces a sustained sound level in the room is switched off. Reverberation time can be related to the defined damping constant  $\zeta$ . If  $\hat{p}$  is given in eq. (3.33), then at the fixed location of the receiver,  $\mathbf{x}$ :

$$10 \log_{10} \left( \frac{|\hat{p}_{max}|^2}{|\hat{p}(t = T_{60})|^2} \right) = 20 \log_{10} \left( \frac{|\hat{p}_{max}|}{|\hat{p}_{max} e^{-\zeta T_{60}}|} \right) = 20 \log_{10} \left( \frac{1}{e^{-\zeta T_{60}}} \right) = 60 \text{dB}.$$

Therefore, the frequency-dependent reverberation time for exponentially decaying amplitudes of the sound equals

$$T_{60} = \frac{1}{\zeta} \ln (10^{60/20}) \approx \frac{6.91}{\zeta}. \quad (3.35)$$

Reverberation time in ordinary rooms is between around 0.3s and several seconds. Hence,  $\zeta < 23 \text{s}^{-1}$  for most of the rooms. Since we are interested only in audible frequencies above 20Hz, that is, for  $\omega = 2\pi f > 125 \text{rad/s}$ , this justifies our assumption of low damping in rooms ( $\zeta \ll \omega$ ), practically for the entire audible frequency range. Moreover, the shortest reverberation times, around 0.3s and below, occur usually only at relatively high frequencies and reverberation time increases at low frequencies.

Here we introduced reverberation time at particular frequencies rather than within finite frequency ranges (bands), as it is usually observed. However, it rarely varies substantially between close frequencies. A possible exception worth mentioning are coupled rooms (rooms connected by relatively small apertures which act as interfaces between them). Depending on their damping, the eigenmodes<sup>27</sup> of the coupled rooms at similar frequencies can cause larger fluctuations of energy decay, which are called beats. Such occurrences can make the estimation of reverberation time difficult.

It is of interest to note that according to equations (3.10) and (3.33):

$$p(\mathbf{x}, t) = \hat{p}(\mathbf{x}) e^{j\omega t} = \hat{p}_{max}(\mathbf{x}) e^{-\zeta t} e^{j\omega t} = \hat{p}_{max}(\mathbf{x}) e^{j(\omega + j\zeta)t}, \quad (3.36)$$

that is, if we allow the phase shift to be complex,  $\Phi = j\zeta t$ , it will correspond to the decay of sound amplitude in time. This is very often used in order to model an exponential decay by promoting the real-valued angular frequency  $\omega$  to a complex value:

$$\omega \rightarrow \omega + j\zeta, \quad (3.37)$$

while keeping the wave number  $k$  and other associated variables real (recall the footnote <sup>19</sup>). Alternatively, the same can be achieved for a forward propagating wave with a complex wave number:

$$k \rightarrow k - j\zeta/c_0, \quad (3.38)$$

<sup>27</sup>The room eigenmodes are discussed in section 3.5.

while keeping the angular frequency real. The sign is changed since the phase of a forward propagating wave (in the positive direction of, say,  $x$ -axis) has the form<sup>28</sup>  $\omega t - kx = (\omega/c_0 - k)x$ . In contrast to this,  $k \rightarrow k + j\zeta/c_0$  would be valid for a wave travelling in the opposite direction. It follows that an amplitude decay can be modelled as occurring both in time and over space. Typically, it takes place during the wave propagation through the medium, for example, due to different mechanisms of sound dissipation in air. However, energy losses after the reflections from absorbing (or transmitting) surfaces can also be taken into account in this way. The abrupt changes of the amplitude after each reflection may appear to be distributed in space, but the cumulative effect on the sound field in the room is essentially the same (if the decay is exponential), especially in low damped rooms, in which many reflections take place before the sound wave vanishes.

Attenuation in air is nearly exponential and it is commonly expressed with the **attenuation constant**  $m_{air}$ , which quantifies the decrease of sound energy per unit length (in  $1/m$ ). From eq. (3.33) it follows:

$$\hat{p}^2 \sim e^{-2\zeta t} = e^{-2\zeta x/c_0}, \quad (3.39)$$

and therefore:

$$m_{air} = \frac{2\zeta_{air}}{c_0}. \quad (3.40)$$

Attenuation in  $\text{dB}/\text{m}$  is then  $10 \log_{10} e^{m_{air}} = m_{air} \cdot 10 \log_{10} e \approx 4.34m_{air}$ . Table 3.1 shows the values of attenuation in air, expressed in  $1/\text{m}$  and  $\text{dB}/\text{m}$ , for standard room temperature, atmospheric pressure, and relative humidity 50%. We can see that it can be significant for frequencies above 1kHz when the propagation paths are very long, around 100m or more. In room acoustics, this is relevant only in very large and weakly damped spaces, such as churches or large halls. Under most circumstances in practice, especially at relatively low frequencies or in the presence of other causes of energy losses in the room (absorbing materials, transmitting walls or openings), the contribution of dissipation in air is negligible.

Table 3.1: Attenuation constant in air  $m_{air}[1/\text{m}]$  and attenuation in  $\text{dB}/\text{m}$  at room temperature, normal atmospheric pressure, and relative humidity 50%.

|               | <b>500Hz</b>        | <b>1000Hz</b>       | <b>2000Hz</b>       | <b>4000Hz</b>       | <b>8000Hz</b>        |
|---------------|---------------------|---------------------|---------------------|---------------------|----------------------|
| $m_{air}$     | $0.6 \cdot 10^{-3}$ | $1.1 \cdot 10^{-3}$ | $2.3 \cdot 10^{-3}$ | $6.8 \cdot 10^{-3}$ | $24.3 \cdot 10^{-3}$ |
| $4.34m_{air}$ | 0.003               | 0.005               | 0.010               | 0.030               | 0.105                |

### 3.4 Green's function in frequency domain

In the previous sections we reduced the acoustic problem to solving the **time-independent** Helmholtz equation (3.30) supplied with boundary conditions<sup>29</sup>. Here we will show how

<sup>28</sup>See for example the sound field from a point source in eq. (4.18)

<sup>29</sup>We ignore the initial conditions and allow the complex sine waves to have an infinite duration. Their phase is set by the source function. The boundary conditions will either determine the tailored

this can be done with the aid of Green's function,  $\hat{G}(\mathbf{x}|\mathbf{y})$ , which satisfies the equation<sup>30</sup>

$$k^2 \hat{G}(\mathbf{x}|\mathbf{y}) + \nabla_x^2 \hat{G}(\mathbf{x}|\mathbf{y}) = -\hat{\delta}(\mathbf{x} - \mathbf{y}). \quad (3.41)$$

We will see that Green's function captures sound propagation from  $\mathbf{y}$  to  $\mathbf{x}$ . The source term is replaced by Dirac (delta) function, a generalized function which satisfies (we remove the hat symbol, since we are listing general properties):

$$\int_{-\infty}^{\infty} \delta(x) dx = \int_{-\epsilon}^{\epsilon} \delta(x) dx = 1, \quad (3.42)$$

where  $x$  is a generic independent variable and  $\epsilon \rightarrow 0^+$  is a positive infinitesimal. The value of  $\delta(x)$  is zero for any  $x \neq 0$ . Thus, delta function is defined with finite values only in the integral sense<sup>31</sup>. Since the integral gives a dimensionless value, the product  $\delta(x)dx$  is also dimensionless. According to eq. (3.41), the unit of  $\hat{G}(\mathbf{x}|\mathbf{y})$  is equal to the unit of  $\hat{\delta}(\mathbf{x} - \mathbf{y})$  multiplied with  $\text{m}^2$ . Since  $\hat{\delta}(\mathbf{x} - \mathbf{y}) = \hat{\delta}(x_1 - y_1)\hat{\delta}(x_2 - y_2)\hat{\delta}(x_3 - y_3)$ , the unit of  $\hat{\delta}(\mathbf{x} - \mathbf{y})$  is  $1/\text{m}^3$  and, therefore, the unit of  $\hat{G}(\mathbf{x}|\mathbf{y})$  (in a three-dimensional case) is  $1/\text{m}$ .

The most important properties of delta function are symmetry:

$$\delta(x - y) = \delta(y - x) \quad (3.43)$$

and sampling:

$$\int_{-\infty}^{\infty} f(x)\delta(x - y) dx = f(y). \quad (3.44)$$

The latter property means that the integral with delta function extracts the value of a continuous function  $f(x)$  only at  $x$  for which the argument of  $\delta$  is zero, that is, at  $x = y$ . An important property of Green's function, which we will not prove here, is reciprocity<sup>32</sup>:  $\hat{G}(\mathbf{x}|\mathbf{y}) = \hat{G}(\mathbf{y}|\mathbf{x})$ . Since delta function is a generalized function, so is Green's function.

Due to the symmetry of delta function and the reciprocity of Green's function, we can switch  $\mathbf{x}$  and  $\mathbf{y}$  in eq. (3.41) and obtain:

$$k^2 \hat{G} + \nabla_y^2 \hat{G} = -\hat{\delta}(\mathbf{x} - \mathbf{y}). \quad (3.45)$$

---

Green's function in eq. (3.48) or the surface integral in eq. (3.46).

<sup>30</sup>Characterization of room acoustics should not depend on any particular source in the room and its properties, such as directivity. Therefore, temporary replacement of the source function with delta function also makes practical sense. Using the hat symbol above delta function is actually unnecessary. It only indicates that we are working in frequency domain and time is not its argument.

<sup>31</sup>It is often loosely defined such that it equals zero for every  $x \neq 0$  and infinity for  $x = 0$  (which is necessary in order to satisfy the integral in eq. (3.42)). However, it has a physical meaning only inside an integral, which is the reason why it is sometimes referred to as a distribution rather than a function. A pointwise definition of such generalized functions is not necessary.

<sup>32</sup>The notation  $\hat{G}(\mathbf{x}, \mathbf{y})$  is also frequently used. Similarly as with complex amplitude  $\hat{p}(\mathbf{x})$ , we often omit the arguments for brevity.

The only change is in the variable of the differential operator. If we multiply eq. (3.32) with  $\hat{G}(\mathbf{x}|\mathbf{y})$  and subtract eq. (3.45) multiplied with  $\hat{p}(\mathbf{y})$  from it, we obtain:

$$\begin{aligned}\hat{G}k^2\hat{p}(\mathbf{y}) + \hat{G}\nabla_y^2\hat{p}(\mathbf{y}) + \hat{G}\hat{q}(\mathbf{y}) - \hat{p}(\mathbf{y})k^2\hat{G} - \hat{p}(\mathbf{y})\nabla_y^2\hat{G} - \hat{p}(\mathbf{y})\hat{\delta}(\mathbf{x} - \mathbf{y}) \\ = \hat{G}\nabla_y^2\hat{p}(\mathbf{y}) + \hat{G}\hat{q}(\mathbf{y}) - \hat{p}(\mathbf{y})\nabla_y^2\hat{G} - \hat{p}(\mathbf{y})\hat{\delta}(\mathbf{x} - \mathbf{y}) = 0.\end{aligned}$$

Then we integrate the result with respect to the source location  $\mathbf{y}$  over some volume  $V$ :

$$\begin{aligned}\hat{p}(\mathbf{x}) &= \int_V \hat{p}(\mathbf{y})\hat{\delta}(\mathbf{x} - \mathbf{y})d^3\mathbf{y} \\ &= \int_V \hat{G}\hat{q}(\mathbf{y})d^3\mathbf{y} + \int_V [\hat{G}\nabla_y^2\hat{p}(\mathbf{y}) - \hat{p}(\mathbf{y})\nabla_y^2\hat{G}] d^3\mathbf{y} \\ &= \int_V \hat{G}\hat{q}(\mathbf{y})d^3\mathbf{y} + \int_V \{\nabla_y \cdot (\hat{G}\nabla_y\hat{p}(\mathbf{y})) - (\nabla_y\hat{G}) \cdot (\nabla_y\hat{p}(\mathbf{y})) \\ &\quad - \nabla_y \cdot (\hat{p}(\mathbf{y})\nabla_y\hat{G}) + (\nabla_y\hat{p}(\mathbf{y})) \cdot (\nabla_y\hat{G})\} d^3\mathbf{y} \\ &= \int_V \hat{G}\hat{q}(\mathbf{y})d^3\mathbf{y} + \int_V \nabla_y \cdot \{\hat{G}\nabla_y\hat{p}(\mathbf{y}) - \hat{p}(\mathbf{y})\nabla_y\hat{G}\} d^3\mathbf{y}.\end{aligned}$$

For the first equality we used the sampling property of delta function. Finally, we can apply the divergence theorem on the last integral:

$$\boxed{\hat{p}(\mathbf{x}) = \int_V \hat{q}(\mathbf{y})\hat{G}d^3\mathbf{y} + \oint_S \{\hat{G}\nabla_y\hat{p}(\mathbf{y}) - \hat{p}(\mathbf{y})\nabla_y\hat{G}\} \cdot \mathbf{n} d^2\mathbf{y}}, \quad (3.46)$$

where  $S$  is the surface which encloses the volume  $V$  and  $\mathbf{n}$  is the unit vector normal to it and pointing outwards. If the volume  $V$  does not contain any sources of sound, the first integral vanishes. However, the volume is usually defined such that it matches the interior of the room and thus includes all sources in it. The surface  $S$  coincides then with the boundary surfaces of the room and the locations  $\mathbf{y}$  refer both to the source region and boundaries (this is why we used the same variable  $\mathbf{y}$  for locations in the source region in the wave equation (2.22) and at the boundaries in eq. (2.25); Green's function characterizes propagation of sound to the receiver at  $\mathbf{x}$  either from the source or surface, which is thus treated in eq. (3.46) as a virtual source).

Equation (3.46) is called Kirchhoff's (integral) equation (sometimes also called Kirchhoff-Helmholtz or Rayleigh equation). It shows that for given sources, Green's function indeed provides the solution of the entire acoustic problem. Unfortunately, it is only by means of the integral equation which has the acoustic variable  $\hat{p}$  appearing on its both sides. Calculation of the integral on the right-hand side requires knowledge of the field on the surface  $S$ , which belongs to the solution we search for. Still, eq. (3.46) has important implications.

For a source located at  $\mathbf{y}$  which produces outgoing sound waves in free space (without any obstacles or boundaries present) with angular frequency  $\omega = kc_0$ , it can be shown that Green's function equals

$$\boxed{\hat{G}_{free}(\mathbf{x}|\mathbf{y}) = \frac{e^{-jk|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} = \frac{e^{-jkr}}{4\pi r}}, \quad (3.47)$$

where  $r = |\mathbf{x} - \mathbf{y}|$  is distance between the source and receiver point. The last equality clarifies the notation  $\mathbf{x}|\mathbf{y}$  which is used for the arguments of Green's function, since it refers to particular locations of the source and receiver and describes propagation of sound between them.

As expected, free space Green's function in eq. (3.47) satisfies the reciprocity condition,  $\hat{G}(\mathbf{x}|\mathbf{y}) = \hat{G}(\mathbf{y}|\mathbf{x})$ , since it depends only on the modulus  $|\mathbf{x} - \mathbf{y}|$ . It is also spherically symmetric, depending only on radial distance and not on polar and azimuthal angles. However, its value is not defined at the source location ( $r = 0$ ). This fact allows even infinitesimally small (point) sources to exhibit directivity (angularly dependent radiation), which can be mathematically shown by expanding  $r = |\mathbf{x} - \mathbf{y}|$  into a Taylor series around  $\mathbf{y} = 0$  (the multipole expansion). Rather than expanding Green's function in order to treat the sources, we will just bear in mind that any solution which is obtained with the non-expanded form, such as the spherically symmetric function in eq. (3.47), has to additionally take into account directivity of the source. For example, this is done in section 4.4 in eq. (4.20).

If Green's function  $\hat{G}(\mathbf{x}|\mathbf{y})$  satisfies both eq. (3.41) and the same boundary conditions as the initial variable  $\hat{p}$  at all surfaces  $S$  of a closed space, it is called tailored Green's function. General boundary condition is given in eq. (3.26) and, therefore, tailored Green's function has to satisfy

$$a\hat{G}_{tail} + b\nabla_y \hat{G}_{tail} \cdot \mathbf{n}(\mathbf{y}) = \hat{c}(\mathbf{y}) \quad (3.48)$$

at every point  $\mathbf{y}$  on each surface and with the same values of  $a$ ,  $b$ , and function  $\hat{c}$  (with the adapted units). In such a case, the two terms inside the curly brackets in the second integral of eq. (3.46) will cancel at the surfaces and the equation simplifies to:

$$\hat{p}(\mathbf{x}) = \int_V \hat{q}(\mathbf{y}) \hat{G}_{tail} d^3 \mathbf{y}. \quad (3.49)$$

This integral equation is evidently much easier to solve than eq. (3.46), since it does not include the pressure amplitude  $\hat{p}$  on its right-hand side. Moreover, tailored Green's function completely describes acoustic behaviour of a room (excluding the sources) for given locations of the source and receiver. Unfortunately, it is usually very difficult to obtain its explicit form for any given room geometry and boundary conditions. We will be able to calculate it analytically in section 3.6 for a simple rectangular room with hard walls. In the trivial case when the boundaries are absent,  $\hat{G}_{tail} = \hat{G}_{free}$ .

If the source is acoustically compact (point source) and located at point  $\mathbf{y}$ , we can write it as the distribution  $\hat{q}(\mathbf{y}') = \hat{Q}(\mathbf{y}')\delta(\mathbf{y}' - \mathbf{y})$ , where  $\hat{Q}$  has the unit kg/s<sup>2</sup>. As before, it can correspond to the unsteady volume injection in eq. (3.31). Notice, however, that we turn the function  $\hat{q}$  into a generalized function. This is not a problem because we will use it in the integral in eq. (3.49) and integral over the small volume which contains the source is

$$\int_V \hat{q}(\mathbf{y}') d^3 \mathbf{y}' = \int_V \hat{Q}(\mathbf{y}') \delta(\mathbf{y}' - \mathbf{y}) d^3 \mathbf{y}' = \hat{Q}(\mathbf{y}), \quad (3.50)$$

according to the sampling property of delta function. Similarly,

$$\begin{aligned}\hat{p}(\mathbf{x}) &= \int_V \hat{q}(\mathbf{y}') \hat{G}_{tail}(\mathbf{x}|\mathbf{y}') d^3\mathbf{y}' = \int_V \hat{Q}(\mathbf{y}') \delta(\mathbf{y}' - \mathbf{y}) \hat{G}_{tail}(\mathbf{x}|\mathbf{y}') d^3\mathbf{y}' \\ &= \hat{Q}(\mathbf{y}) \hat{G}_{tail}(\mathbf{x}|\mathbf{y}),\end{aligned}\quad (3.51)$$

which makes the connection between pressure amplitude, source, and tailored Green's function even more explicit. In fact, tailored Green's function entirely corresponds to the frequency response of the room. Still, the last equality makes physical sense only when  $\hat{G}_{tail}$  is a regular function at  $\mathbf{x}$  outside the source region. For example, in free space:

$$\hat{p}(\mathbf{x}) = \hat{Q}(\mathbf{y}) \hat{G}_{free}(\mathbf{x}|\mathbf{y}) = \hat{Q}(\mathbf{y}) \frac{e^{-jk|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}. \quad (3.52)$$

The pressure amplitude at  $\mathbf{x} \neq \mathbf{y}$  is merely a scaled version of  $\hat{Q}$  with changed phase.

### 3.5 Room eigenmodes

We will first treat a room completely generally as a closed cavity with volume  $V$  in which an omnidirectional point source produces a sound field. The complete solution for the sound field is obtained after integrating eq. (3.51) over all wave numbers  $k$ . However, **tailored Green's function of closed spaces can always be expanded as a sum of an infinite number of distinct eigenmodes (complex eigenfunctions, or simply called modes)  $\psi_n(\mathbf{x})$ :**

$$\hat{G}_{tail}(\mathbf{x}|\mathbf{y}) = \sum_n A_n(\mathbf{y}) \psi_n(\mathbf{x}), \quad (3.53)$$

where  $\psi_n$  denotes mode  $n$  ( $n = 1, 2, 3, \dots$ ) and  $A_n$  is its complex scaling factor, which does not depend on  $\mathbf{x}$ . Since  $\hat{G}_{tail}(\mathbf{x}|\mathbf{y})$  depends on both  $\mathbf{x}$  and  $\mathbf{y}$ , we expect the coefficient  $A_n$  to depend on  $\mathbf{y}$ , that is, on the location of the source (and, of course, frequency or, equivalently, wave number). Thus, we separate the two variables which depend on the source and receiver locations.

Note that this is very similar to the expansion in time domain, which was done in section 3.1 with the sum of complex sine waves  $e^{j\omega t}$  and separation of time and space dependent quantities. **Equation (3.53) represents a Fourier series** and can be directly compared with the Fourier transform in eq. (3.22). **Appearance of the sum instead of the integral is due to the spatial boundedness of sound field** (there is no “boundary” in time) and **more general form of eigenmodes**,  $\psi_n(\mathbf{x})$  instead of a simple exponential function such as  $e^{j\omega t}$ , is because of the variety of possible boundary conditions. While space involves in general three coordinates, which can be coupled within the boundary conditions, time is a single scalar and we assumed no restrictions by initial conditions. In fact, we can treat the terms  $e^{j\omega t}$  as eigenmodes in time with no additional constraints. Similarly, if we set  $\mathbf{y} = 0$ ,  $e^{-jkr}/r = e^{-jk|\mathbf{x}|}/|\mathbf{x}|$  in the expression for  $\hat{G}_{free}$  (eq. (3.47)) can be seen as eigenmodes with respect to the radial coordinate of a free space (without

boundary conditions), and  $\psi_n(\mathbf{x})$  as eigenmodes in a closed space with boundary conditions. While  $\omega$  and  $k$  can take any (real and positive) value in the first two cases without constraints, we will see that each  $\psi_n(\mathbf{x})$  is associated with a single discrete value of  $k_n$ , which depends entirely on the boundary conditions.

The modes are mutually orthogonal, which is written as:

$$\int_V \psi_m^*(\mathbf{x}) \psi_n(\mathbf{x}) d^3 \mathbf{x} = \begin{cases} K_n & \text{for } m = n \\ 0 & \text{otherwise.} \end{cases} \quad (3.54)$$

The integral over the entire room volume is zero for any two different modes. For  $m = n$ , the integral is equal to a real-valued constant  $K_n$ . Following the discussion from above, this can be compared with the orthogonality of complex exponential functions of time in eq. (3.11). The finite value on the right-hand side instead of delta function is due to the finite volume.

Next we would like to express the eigenmodes  $\psi_n(\mathbf{x})$  more specifically. Outside the source region, Green's function has to satisfy eq. (3.41) with the right-hand side equal to zero. Using the expansion from eq. (3.53), we obtain:

$$\begin{aligned} k^2 \hat{G}_{tail} + \nabla_x^2 \hat{G}_{tail} &= \sum_n k_n^2 A_n(\mathbf{y}) \psi_n(\mathbf{x}) + \nabla_x^2 \sum_n A_n(\mathbf{y}) \psi_n(\mathbf{x}) \\ &= \sum_n A_n(\mathbf{y}) [k_n^2 \psi_n(\mathbf{x}) + \nabla_x^2 \psi_n(\mathbf{x})] = 0, \end{aligned} \quad (3.55)$$

where  $k_n$  is now wave number of the mode  $\psi_n$ . Therefore, for every non-trivial  $A_n \neq 0$ :

$$k_n^2 \psi_n(\mathbf{x}) + \nabla_x^2 \psi_n(\mathbf{x}) = 0. \quad (3.56)$$

The modes which constitute the tailored Green's function satisfy homogeneous Helmholtz equation with appropriate wave number.

In order to evaluate the coefficients  $A_n(\mathbf{y})$ , we will use the property of orthogonality of eigenmodes. We again insert the expansion from eq. (3.53) into eq. (3.41). This time, however, we solve for the coefficients  $A_n(\mathbf{y})$  which do not *a priori* depend on the modes. Therefore, we have to allow an arbitrary  $k$ . Furthermore, the delta function on the right-hand side of eq. (3.41) brings the dependence on  $\mathbf{y}$ . Thus, we obtain:

$$\begin{aligned} k^2 \hat{G}_{tail} + \nabla_x^2 \hat{G}_{tail} &= k^2 \sum_n A_n(\mathbf{y}) \psi_n(\mathbf{x}) + \nabla_x^2 \sum_n A_n(\mathbf{y}) \psi_n(\mathbf{x}) \\ &= \sum_n A_n(\mathbf{y}) [k^2 \psi_n(\mathbf{x}) + \nabla_x^2 \psi_n(\mathbf{x})] = \sum_n A_n(\mathbf{y}) [k^2 \psi_n(\mathbf{x}) - k_n^2 \psi_n(\mathbf{x})] \\ &= \sum_n A_n(\mathbf{y}) (k^2 - k_n^2) \psi_n(\mathbf{x}) = -\hat{\delta}(\mathbf{x} - \mathbf{y}), \end{aligned}$$

where we also used the last result from eq. (3.55).

Next we multiply the last equality with an arbitrary eigenfunction  $\psi_m(\mathbf{x})$  and integrate over the entire room volume (with respect to  $\mathbf{x}$ ):

$$\begin{aligned} \int_V \sum_n A_n(\mathbf{y})(k^2 - k_n^2) \psi_n(\mathbf{x}) \psi_m(\mathbf{x}) d^3 \mathbf{x} &= \sum_n A_n(\mathbf{y})(k^2 - k_n^2) \int_V \psi_n(\mathbf{x}) \psi_m(\mathbf{x}) d^3 \mathbf{x} \\ &= (A_n(\mathbf{y})(k^2 - k_n^2) K_n)_{n=m} = A_m(\mathbf{y})(k^2 - k_m^2) K_m \\ &= - \int_V \psi_m(\mathbf{x}) \hat{\delta}(\mathbf{x} - \mathbf{y}) d^3 \mathbf{x} = -\psi_m(\mathbf{y}). \end{aligned}$$

Orthogonality of the modes from eq. (3.54) left only one member of the sum, for  $n = m$ , and the selectivity of delta function replaced  $\mathbf{x}$  with  $\mathbf{y}$  as the argument of the eigenfunction. Since this is valid for any mode  $m$ , it also holds for any mode  $n$  and we can simply change the index:

$$A_n(\mathbf{y}) = \frac{\psi_n(\mathbf{y})}{K_n(k_n^2 - k^2)}. \quad (3.57)$$

If we can determine the eigenmodes  $\psi_n$  and their wave numbers  $k_n$ , which satisfy both eq. (3.56) and the boundary conditions (since we started with tailored Green's function in eq. (3.53)), we can also determine the constants  $K_n$  from eq. (3.54) and then the coefficients  $A_n$  from eq. (3.57). These coefficients together with the eigenmodes define the tailored Green's function in eq. (3.53), which can then be used in eq. (3.49) together with a known source function to obtain the complete solution in terms of the complex amplitude of sound pressure (the time dependence is simply attached by multiplication with  $e^{j\omega t}$ , according to eq. (3.10)). The tailored Green's function equals

$$\hat{G}_{tail}(\mathbf{x}|\mathbf{y}) = \sum_n A_n(\mathbf{y}) \psi_n(\mathbf{x}) = \sum_n \frac{\psi_n(\mathbf{x}) \psi_n(\mathbf{y})}{K_n(k_n^2 - k^2)} = c_0^2 \sum_n \frac{\psi_n(\mathbf{x}) \psi_n(\mathbf{y})}{K_n(\omega_n^2 - \omega^2)}, \quad (3.58)$$

which also satisfies the reciprocity:  $\hat{G}_{tail}(\mathbf{x}|\mathbf{y}) = \hat{G}_{tail}(\mathbf{y}|\mathbf{x})$ .

However, we are still far from a closed form solution, since the boundary conditions at all surfaces determine the eigenmodes and their associated wave numbers. In fact, since tailored Green's function is superposition of eigenmodes (eq. (3.53)), the eigenmodes have to satisfy the same boundary conditions as the tailored Green's function, which are, as already discussed, also satisfied by the original variable  $\hat{p}$ . From equations (3.48) and (3.53) it follows:

$$\begin{aligned} a \hat{G}_{tail} + b \nabla_y \hat{G}_{tail} \cdot \mathbf{n}(\mathbf{y}) &= \sum_n a_n A_n \psi_n(\mathbf{y}) + \nabla_y \left( \sum_n b_n A_n \psi_n(\mathbf{y}) \right) \cdot \mathbf{n}(\mathbf{y}) \\ &= \sum_n A_n [a_n \psi_n(\mathbf{y}) + b_n \nabla_y \psi_n(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y})] = \hat{c}(\mathbf{y}), \end{aligned}$$

where we left only  $\mathbf{y}$  at the boundary (not in the source region) as the argument, in order to avoid confusion. Consequently,

$$a_n \psi_n(\mathbf{y}) + b_n \nabla_y \psi_n(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y}) = \frac{\hat{c}_n(\mathbf{y})}{A_n}, \quad (3.59)$$

with  $\sum_n \hat{c}_n(\mathbf{y}) = \hat{c}(\mathbf{y})$ .

Even without knowing the exact forms of eigenmodes, certain important conclusions can already be made based on eq. (3.58). In addition, we can introduce damping as in equations (3.36) and (3.37), that is, we model the exponential decay of amplitude<sup>33</sup> by adding the imaginary part  $j\zeta_n$  to the real angular frequency  $\omega_n$  of mode  $n$ . Tailored Green's function becomes

$$\hat{G}_{tail}(\mathbf{x}|\mathbf{y}) = c_0^2 \sum_n \frac{\psi_n(\mathbf{x})\psi_n(\mathbf{y})}{K_n(\omega_n^2 + 2j\zeta_n\omega_n - \zeta_n^2 - \omega^2)}. \quad (3.60)$$

From inequality (3.34),  $\zeta_n \ll \omega_n$  in weakly damped rooms, so we can approximate:

$$\hat{G}_{tail}(\mathbf{x}|\mathbf{y}) = c_0^2 \sum_n \frac{\psi_n(\mathbf{x})\psi_n(\mathbf{y})}{K_n(\omega_n^2 + 2j\zeta_n\omega_n - \omega^2)}. \quad (3.61)$$

Inserting this into eq. (3.49), the formal solution of eq. (3.30) is

$$\hat{p}(\mathbf{x}) = c_0^2 \int_V \hat{q}(\mathbf{y}) \sum_n \frac{\psi_n(\mathbf{x})\psi_n(\mathbf{y})}{K_n(\omega_n^2 + 2j\zeta_n\omega_n - \omega^2)} d^3\mathbf{y}. \quad (3.62)$$

The solution for a point source, eq. (3.51), reads

$$\boxed{\hat{p}(\mathbf{x}) = \hat{Q}(\mathbf{y})\hat{G}_{tail}(\mathbf{x}|\mathbf{y}) = \hat{Q}(\mathbf{y})c_0^2 \sum_n \frac{\psi_n(\mathbf{x})\psi_n(\mathbf{y})}{K_n(\omega_n^2 + 2j\zeta_n\omega_n - \omega^2)}}. \quad (3.63)$$

Each member of the sum corresponds to a single mode of the room. Its contribution to the complex pressure amplitude is

$$\hat{p}_n(\mathbf{x}) = \hat{Q}(\mathbf{y})c_0^2 \frac{\psi_n(\mathbf{x})\psi_n(\mathbf{y})}{K_n(\omega_n^2 + 2j\zeta_n\omega_n - \omega^2)}. \quad (3.64)$$

Different (complex) terms in this expression can be interpreted as follows:

- $\hat{Q}(\mathbf{y})$  depends only on the (point) source,
- $\psi_n(\mathbf{x})\psi_n(\mathbf{y})/K_n$  indicates how strongly the mode  $n$  can be excited by the source; it depends on the geometry of the problem – shape and size of the room, boundary conditions, as well as the locations of the source and receiver, and
- $1/(\omega_n^2 + 2j\zeta_n\omega_n - \omega^2)$  is frequency dependence of the contribution of mode  $n$  to the total sound pressure; the logarithm of its modulus,  $20 \log_{10}(1/|\omega_n^2 + 2j\zeta_n\omega_n - \omega^2|)$ , is shown in Fig. 4 for several realistic values of the damping constant  $\zeta_n$  and for normalized frequency  $f/f_n$  (left) and  $f_n = 1\text{kHz}$  (right).

<sup>33</sup>Such modelled absorption obviously does not follow directly from the governing equations. However, it makes the calculation much more transparent and approximates well the realistic effects of damping in a room.

Since  $\zeta_n \ll \omega_n$  and thus  $2\zeta_n\omega_n \ll \omega_n^2$ , it follows that the damping term is important only when  $\omega_n^2$  and  $\omega^2$  largely cancel, that is, when  $\omega \approx \omega_n$ . Therefore, the damping is critical close to the frequency of the mode and the entire associated term can also be written as  $1/(\omega_n^2 + 2j\zeta_n\omega - \omega^2)$ . It determines how high the peaks in the frequency response of a room due to the modes will be. More damped rooms will have less pronounced resonance behaviour.

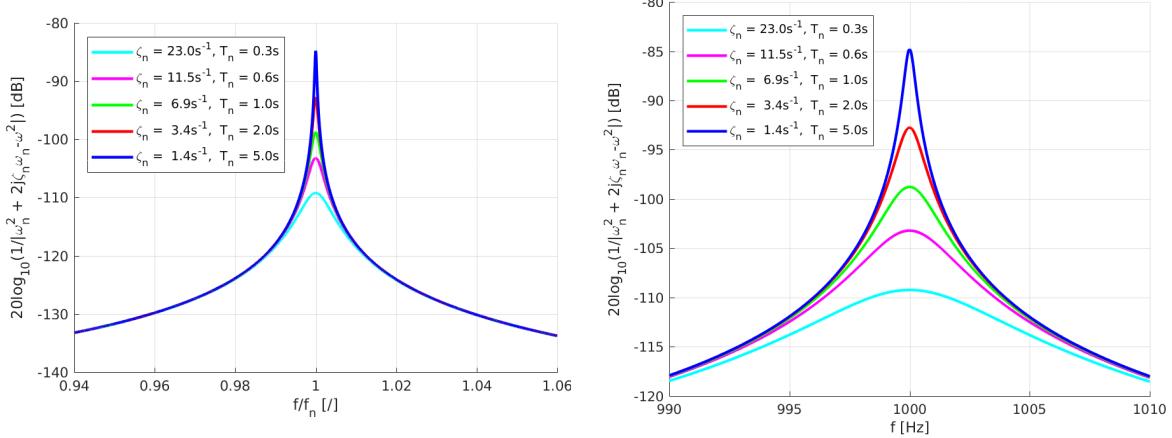


Figure 4: Amplitude of a damped eigenmode: (left) normalized frequency, (right) mode at 1000Hz.

The curves in Fig. 4 depict typical behaviour of a damped linear oscillator, such as a mass-spring system. This follows from the analogy between eq. (3.30) and the equation of an undamped<sup>34</sup> oscillator, with the major difference being in the independent variable (location instead of time). The ordinary (time is the only independent variable, so we can write  $d/dt$  instead of  $\partial/\partial t$ ) differential equation of a simple slightly damped forced oscillator reads

$$M \frac{d^2\xi(t)}{dt^2} + D \frac{d\xi(t)}{dt} + S\xi(t) = F(\tau), \quad (3.65)$$

where  $\xi$  is one-dimensional displacement (m),  $M$  is mass (kg),  $D$  is damping (kg/s),  $S$  is stiffness ( $\text{kg}/\text{s}^2$ ), and  $F$  is one-dimensional force (N) which acts on the mass  $M$ . By comparing this equation with the Helmholtz equation (3.30), we see that the second-order spatial derivative (in three dimensional space) is replaced with the second-order time derivative and the following substitutions apply:  $\hat{p}(\mathbf{x}) \rightarrow \xi(t)$ ,  $k^2 \rightarrow S/M$ , and  $-\hat{q}(\mathbf{y}) \rightarrow F(\tau)/M$ . The remaining term  $(D/M)(d\xi(t)/dt)$  is due to the damping, which is not contained in eq. (3.30), but modelled in eq. (3.60).

If we write the displacement and force as  $\xi(t) = \hat{\xi}e^{j\omega t}$  and  $F(\tau) = \hat{F}e^{j\omega\tau}$ , respectively, and divide eq. (3.65) with  $e^{j\omega t}$ , we obtain a simple algebraic equation

$$-M\omega^2\hat{\xi} + j\omega D\hat{\xi} + S\hat{\xi} = \hat{F}. \quad (3.66)$$

<sup>34</sup>We introduced damping first in eq. (3.60).

Since the force acts directly on the moving mass, there is no phase shift from  $e^{j\omega(t-\tau)}$ . The solution is

$$\hat{\xi} = \frac{\hat{F}/M}{S/M + j\omega D/M - \omega^2}. \quad (3.67)$$

For small damping, resonance of the mass-spring system occurs at the angular frequency  $\omega_0 = \sqrt{S/M}$ . Damping affects the Q-factor of the resonance, which is defined as  $Q = \omega_0 M / D = \sqrt{SM}/D$ . The Q-factor is related to the half-width of the resonance curve,  $\Delta\omega$ , which is angular frequency range in which amplitude of the oscillations drops by the factor not larger than  $\sqrt{2}$  and the energy by the factor not larger than 2, compared to the peak values. The relation is  $Q = \omega_0 / \Delta\omega$ , or  $\Delta\omega = \omega_0 / Q = D/M$ . Mechanical impedance is also defined as

$$Z = \frac{F}{v} = \frac{F}{d\xi/dt} = \frac{\hat{F}}{j\omega\hat{\xi}} = j\omega M + D + \frac{S}{j\omega}. \quad (3.68)$$

By comparing the denominator in eq. (3.67) with  $S/M = \omega_0^2$  and the term  $\omega_n^2 + 2j\zeta_n\omega - \omega^2$  from above, we can conclude that every room (so far, we have only assumed that the damping is low and that the room is excited by a point omnidirectional source, with no other restrictions on the geometry of the room or boundary conditions) behaves as an infinite series of simple harmonic oscillators, with the resonances at frequencies  $f_n = \omega_n/(2\pi)$ , with half-widths  $\Delta\omega_n = 2\zeta_n$  and Q-factors  $Q_n = \omega_n/(2\zeta_n)$ . The damping constant  $\zeta_n$  corresponds to  $D/(2M) = \Delta\omega_n/2$ . The excitation “force” of the oscillators is, of course, introduced by the source of sound  $\hat{Q}$  for each  $f_n$ . However, the amount of excitation of the room resonances (and the actual values on the vertical axes in Fig. (4)) is also spatially dependent. It depends on the source and receiver locations and the shape of the modes (the term  $\psi_n(\mathbf{x})\psi_n(\mathbf{y})/K_n$ ). Moreover, the total sound field in eq. (3.63) is a sum of the contributions of many modes with different (or even equal) frequencies. Depending on the signs of their terms  $\psi_n(\mathbf{x})$  and  $\psi_n(\mathbf{y})$ , the modes can add constructively or cancel, increasing the peaks in Fig. 4 or creating dips in the frequency response.

Summarizing the methodology we have used so far for solving the room acoustic problems, Table 3.2 lists the main quantities which appeared, the key equations (and boundary conditions) which they satisfy, and the variables on which they depend. It should demonstrate gradual decomposition of the initial problem (defined by the governing equations and the wave equation), which we achieved by separating the variables and decoupling the dependencies.

In order to obtain a more specific solution, we need to calculate the eigenmodes  $\psi_n$  of the room and their wave numbers, which satisfy eq. (3.56) and boundary conditions, eq. (3.59). All boundary conditions have to be satisfied simultaneously by each mode, since we removed any time dependence by working in the frequency domain (the solution, complex pressure amplitude, is also independent of time). This is obviously a very difficult task in general, which is the main limitation of the modal analysis. It can be achieved analytically only in very few idealized cases. These are the rooms with very

Table 3.2: Main quantities and their dependencies. Notes: BC – boundary conditions; initial condition in time domain is  $p = 0$  before the source  $q$  is switched on and in frequency domain the duration of sine waves is assumed to be unlimited; frequency dependence is always implied.

| quantity         | equation  | depends on  |
|------------------|-----------|---|
| $p$              | eq.(2.22) | $\mathbf{x}, t, q(\mathbf{y}, \tau)$ , BC in eq. (2.25) |
| $\hat{p}$        | eq.(3.30) | $\mathbf{x}, \hat{q}(\mathbf{y})$ , BC in eq. (3.26)    |
| $\hat{G}_{tail}$ | eq.(3.41) | $\mathbf{x}, \mathbf{y}$ , BC in eq. (3.48)             |
| $\psi_n$         | eq.(3.56) | $\mathbf{x}$ , BC in eq. (3.59)                         |
| $A_n$            | eq.(3.57) | $\mathbf{y}, \psi_n$                                    |

simple geometry and boundary conditions, for which it is possible to further separate the spatially-dependent variables in a suitable coordinate system. For a rectangular room with uniform boundary conditions, this is most easily done (as in the next section) with Cartesian coordinates<sup>35</sup>. In any case, we should bear in mind that the full solution of a sound field necessarily depends on the source function, which must be given.

### 3.6 Rectangular room with hard walls

In order to separate the variables in Cartesian coordinates, we suppose the following form of the solution of eq. (3.56):

$$\psi_n(\mathbf{x}) = \psi_{n1}(x_1)\psi_{n2}(x_2)\psi_{n3}(x_3) = \psi_{n1}\psi_{n2}\psi_{n3}, \quad (3.69)$$

where we placed the components of  $\mathbf{x}$  in the subscripts for shorter notation. (We will do the same with the components of  $\mathbf{y}$  below.) As a convention,  $n_i = 0, 1, 2, \dots$  for  $i = 1, 2, 3$ , but if all three  $n_i$  are zero the mode  $\psi_{n=0}$  is non-acoustical. Equation (3.56) becomes

$$\begin{aligned} \nabla_x^2(\psi_{n1}\psi_{n2}\psi_{n3}) + k_n^2\psi_{n1}\psi_{n2}\psi_{n3} &= \psi_{n2}\psi_{n3}\frac{d^2\psi_{n1}}{dx_1^2} + \psi_{n1}\psi_{n3}\frac{d^2\psi_{n2}}{dx_2^2} \\ &\quad + \psi_{n1}\psi_{n2}\frac{d^2\psi_{n3}}{dx_3^2} + k_n^2\psi_{n1}\psi_{n2}\psi_{n3} = 0, \end{aligned} \quad (3.70)$$

Dividing with  $\psi_{n1}\psi_{n2}\psi_{n3}$  for any  $\psi_n \neq 0$  gives

$$\frac{1}{\psi_{n1}}\frac{d^2\psi_{n1}}{dx_1^2} + \frac{1}{\psi_{n2}}\frac{d^2\psi_{n2}}{dx_2^2} + \frac{1}{\psi_{n3}}\frac{d^2\psi_{n3}}{dx_3^2} + k_n^2 = 0, \quad (3.71)$$

Since  $\psi_{n1}$ ,  $\psi_{n2}$ , and  $\psi_{n3}$  are functions of different components of  $\mathbf{x}$ , we can also promote  $k_n$  to wave vector  $\mathbf{k}_n$  with components  $k_{n1}$ ,  $k_{n2}$ , and  $k_{n3}$  in three-dimensional  $k$ -space, such that

$$k_n^2 = |\mathbf{k}_n|^2 = k_{n1}^2 + k_{n2}^2 + k_{n3}^2. \quad (3.72)$$

<sup>35</sup>Similarly, cylindrical coordinates would be more appropriate for cylindrical cavities and spherical coordinates for spherical cavities. Such geometries have much less practical relevance in room acoustics, so we will cover only rectangular rooms in more details.

Then eq. (3.71) gives three decoupled ordinary differential equations:

$$\boxed{\begin{aligned} \frac{d^2\psi_{n1}}{dx_1^2} + k_{n1}^2\psi_{n1} &= 0, \\ \frac{d^2\psi_{n2}}{dx_2^2} + k_{n2}^2\psi_{n2} &= 0, \\ \frac{d^2\psi_{n3}}{dx_3^2} + k_{n3}^2\psi_{n3} &= 0. \end{aligned}} \quad (3.73)$$

General solutions of equations (3.73) are, respectively:

$$\begin{aligned} \psi_{n1} &= C_{11}e^{-jk_{n1}x_1} + C_{21}e^{jk_{n1}x_1} \\ &= (C_{11} + C_{21})\cos(k_{n1}x_1) - j(C_{11} - C_{21})\sin(k_{n1}x_1), \\ \psi_{n2} &= C_{12}e^{-jk_{n2}x_2} + C_{22}e^{jk_{n2}x_2} \\ &= (C_{12} + C_{22})\cos(k_{n2}x_2) - j(C_{12} - C_{22})\sin(k_{n2}x_2), \\ \psi_{n3} &= C_{13}e^{-jk_{n3}x_3} + C_{23}e^{jk_{n3}x_3} \\ &= (C_{13} + C_{23})\cos(k_{n3}x_3) - j(C_{13} - C_{23})\sin(k_{n3}x_3), \end{aligned} \quad (3.74)$$

where  $C_{ji}$  ( $j = 1, 2$ ) are complex constants and we also applied Euler's formula. Note that we have to consider both complex conjugate solutions for the components of complex modes,  $e^{jk_{ni}x_i}$  and  $e^{-jk_{ni}x_i}$ , unlike in eq. (3.9) where only  $e^{j\omega t}$  was sufficient to cover all physically relevant cases for real sound field (see footnote <sup>20)</sup>). In particular, for our convention with  $e^{j\omega t}$ , the waves with  $e^{-jk_{ni}x_i}$  propagate forwards (in the direction of axis  $x_i$ ) and the waves with  $e^{jk_{ni}x_i}$  propagate backwards (in the opposite direction).

The last equations have 9 unknowns: 6 complex constants  $C_{ji}$  and three (still) real-valued components  $k_{ni}$  (as we will see, complex constants  $a_n$  and  $b_n$  in the boundary conditions can introduce damping and lead to complex values of the wave number). They depend on the boundary conditions at the plane surfaces of the rectangular room, which are described by eq. (3.59). If we assume that all surfaces are passive,  $\hat{c}_n(\mathbf{y}) = 0$ :

$$a_n\psi_n(\mathbf{y}) + b_n\nabla_y\psi_n(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y}) = 0 \quad (3.75)$$

for the locations  $\mathbf{y}$  at the surfaces and  $\mathbf{n}$  points into the room.

If the room extends from 0 to  $L_1$ , 0 to  $L_2$ , and 0 to  $L_3$  along the axes  $x_1$ ,  $x_2$ , and  $x_3$ , respectively, the component  $n_1$  of  $\mathbf{n}$  is different than zero only for the two surfaces which are normal to the  $x_1$ -axis, which is when  $y_1 = 0$  or  $L_1$ . Similarly,  $n_2 \neq 0$  for  $y_2 = 0$  or  $L_2$  and  $n_3 \neq 0$  for  $y_3 = 0$  or  $L_3$ . More precisely:  $n_1 = 1$  for  $y_1 = 0$  and  $n_1 = -1$  for  $y_1 = L_1$ ,  $n_2 = 1$  for  $y_2 = 0$  and  $n_2 = -1$  for  $y_2 = L_2$ ,  $n_3 = 1$  for  $y_3 = 0$  and  $n_3 = -1$  for  $y_3 = L_3$ . Uniform boundary conditions ( $a_n$  and  $b_n$  are assumed to be constant over each surface) can then be defined for each of the six surfaces with separate variables. Exactly this fact allows the entire procedure of separation of variables, which thus critically depends on the geometry and boundary conditions. In more complicated (and realistic) cases, the variables may be coupled such that the separation is much less efficient, if not impossible.

We again neglect the trivial case when  $\psi_{ni} = 0$  for any  $i = 1, 2, 3$ , so that the boundary conditions read

$$\begin{aligned}
 a_{n10}\psi_{n1} + b_{n10}\frac{d\psi_{n1}}{dy_1} &= a_{n10}C_{11}e^{-jk_{n1}y_1} + a_{n10}C_{21}e^{jk_{n1}y_1} \\
 &\quad + b_{n10}C_{11}(-jk_{n1})e^{-jk_{n1}y_1} + b_{n10}C_{21}(jk_{n1})e^{jk_{n1}y_1} = 0 \text{ for } y_1 = 0, \\
 a_{n1L}\psi_{n1} - b_{n1L}\frac{d\psi_{n1}}{dy_1} &= a_{n1L}C_{11}e^{-jk_{n1}y_1} + a_{n1L}C_{21}e^{jk_{n1}y_1} \\
 &\quad - b_{n1L}C_{11}(-jk_{n1})e^{-jk_{n1}y_1} - b_{n1L}C_{21}(jk_{n1})e^{jk_{n1}y_1} = 0 \text{ for } y_1 = L_1, \\
 a_{n20}\psi_{n2} + b_{n20}\frac{d\psi_{n2}}{dy_2} &= a_{n20}C_{12}e^{-jk_{n2}y_2} + a_{n20}C_{22}e^{jk_{n2}y_2} \\
 &\quad + b_{n20}C_{12}(-jk_{n2})e^{-jk_{n2}y_2} + b_{n20}C_{22}(jk_{n2})e^{jk_{n2}y_2} = 0 \text{ for } y_2 = 0, \\
 a_{n2L}\psi_{n2} - b_{n2L}\frac{d\psi_{n2}}{dy_2} &= a_{n2L}C_{12}e^{-jk_{n2}y_2} + a_{n2L}C_{22}e^{jk_{n2}y_2} \\
 &\quad - b_{n2L}C_{12}(-jk_{n2})e^{-jk_{n2}y_2} - b_{n2L}C_{22}(jk_{n2})e^{jk_{n2}y_2} = 0 \text{ for } y_2 = L_2, \\
 a_{n30}\psi_{n3} + b_{n30}\frac{d\psi_{n3}}{dy_3} &= a_{n30}C_{13}e^{-jk_{n3}y_3} + a_{n30}C_{23}e^{jk_{n3}y_3} \\
 &\quad + b_{n30}C_{13}(-jk_{n3})e^{-jk_{n3}y_3} + b_{n30}C_{23}(jk_{n3})e^{jk_{n3}y_3} = 0 \text{ for } y_3 = 0, \\
 a_{n3L}\psi_{n3} - b_{n3L}\frac{d\psi_{n3}}{dy_3} &= a_{n3L}C_{13}e^{-jk_{n3}y_3} + a_{n3L}C_{23}e^{jk_{n3}y_3} \\
 &\quad - b_{n3L}C_{13}(-jk_{n3})e^{-jk_{n3}y_3} - b_{n3L}C_{23}(jk_{n3})e^{jk_{n3}y_3} = 0 \text{ for } y_3 = L_3.
 \end{aligned}$$

The coefficients  $a_{n10}$  and  $b_{n10}$  are defined for the surface  $y_1 = 0$ ,  $a_{n1L}$  and  $b_{n1L}$  for the surface  $y_1 = L_1$ ,  $a_{n20}$  and  $b_{n20}$  for the surface  $y_2 = 0$ , and so on. After inserting the appropriate values of  $y_i$ :

$$\begin{aligned}
 C_{11}(a_{n10} - jb_{n10}k_{n1}) + C_{21}(a_{n10} + jb_{n10}k_{n1}) &= 0, \\
 C_{11}(a_{n1L} + jb_{n1L}k_{n1})e^{-jk_{n1}L_1} + C_{21}(a_{n1L} - jb_{n1L}k_{n1})e^{jk_{n1}L_1} &= 0, \\
 C_{12}(a_{n20} - jb_{n20}k_{n2}) + C_{22}(a_{n20} + jb_{n20}k_{n2}) &= 0, \\
 C_{12}(a_{n2L} + jb_{n2L}k_{n2})e^{-jk_{n2}L_2} + C_{22}(a_{n2L} - jb_{n2L}k_{n2})e^{jk_{n2}L_2} &= 0, \\
 C_{13}(a_{n30} - jb_{n30}k_{n3}) + C_{23}(a_{n30} + jb_{n30}k_{n3}) &= 0, \\
 C_{13}(a_{n3L} + jb_{n3L}k_{n3})e^{-jk_{n3}L_3} + C_{23}(a_{n3L} - jb_{n3L}k_{n3})e^{jk_{n3}L_3} &= 0.
 \end{aligned} \tag{3.76}$$

These are 6 equations for the 9 unknowns,  $C_{ji}$  and  $k_{ni}$ . In fact, these are 3 systems of two equations. If they have non-trivial solutions for the coefficients  $C_{ji}$ , their determinants have to be zero, which gives:

$$\begin{aligned}
 (a_{n10} - jb_{n10}k_{n1})(a_{n1L} - jb_{n1L}k_{n1})e^{jk_{n1}L_1} &= (a_{n10} + jb_{n10}k_{n1})(a_{n1L} + jb_{n1L}k_{n1})e^{-jk_{n1}L_1}, \\
 (a_{n20} - jb_{n20}k_{n2})(a_{n2L} - jb_{n2L}k_{n2})e^{jk_{n2}L_2} &= (a_{n20} + jb_{n20}k_{n2})(a_{n2L} + jb_{n2L}k_{n2})e^{-jk_{n2}L_2}, \\
 (a_{n30} - jb_{n30}k_{n3})(a_{n3L} - jb_{n3L}k_{n3})e^{jk_{n3}L_3} &= (a_{n30} + jb_{n30}k_{n3})(a_{n3L} + jb_{n3L}k_{n3})e^{-jk_{n3}L_3},
 \end{aligned}$$

or after multiplication with  $e^{jk_{ni}L_i}$ :

$$\begin{aligned} e^{2jk_{n1}L_1} &= \frac{(a_{n10} + jb_{n10}k_{n1})(a_{n1L} + jb_{n1L}k_{n1})}{(a_{n10} - jb_{n10}k_{n1})(a_{n1L} - jb_{n1L}k_{n1})}, \\ e^{2jk_{n2}L_2} &= \frac{(a_{n20} + jb_{n20}k_{n2})(a_{n2L} + jb_{n2L}k_{n2})}{(a_{n20} - jb_{n20}k_{n2})(a_{n2L} - jb_{n2L}k_{n2})}, \\ e^{2jk_{n3}L_3} &= \frac{(a_{n30} + jb_{n30}k_{n3})(a_{n3L} + jb_{n3L}k_{n3})}{(a_{n30} - jb_{n30}k_{n3})(a_{n3L} - jb_{n3L}k_{n3})}. \end{aligned}$$

These are now three equations for  $k_{ni}$  and they can be solved for given values of the coefficients  $a_n$  and  $b_n$  (although, they are transcendental equations, which prevents a closed form solution in general). The obtained values of  $k_n$  can be complex if damping is introduced by the boundary conditions, with the same physical interpretation as in eq. (3.38). In any case, the obtained values can be inserted into equations (3.76) in order to find the values of the remaining coefficients  $C_{ji}$ , which determine the eigenmodes according to eq. (3.74). For simplicity, we will assume that all surfaces are equal, that is,  $a_{n10} = a_{n1L} = a_{n20} = a_{n2L} = a_{n30} = a_{n3L} = a_n$  and analogously for  $b_n$ . The last equations then reduce to

$$\begin{aligned} e^{2jk_{n1}L_1} &= \frac{(a_n + jb_n k_{n1})^2}{(a_n - jb_n k_{n1})^2} \Rightarrow e^{jk_{n1}L_1} = \pm \frac{a_n + jb_n k_{n1}}{a_n - jb_n k_{n1}}, \\ e^{2jk_{n2}L_2} &= \frac{(a_n + jb_n k_{n2})^2}{(a_n - jb_n k_{n2})^2} \Rightarrow e^{jk_{n2}L_2} = \pm \frac{a_n + jb_n k_{n2}}{a_n - jb_n k_{n2}}, \\ e^{2jk_{n3}L_3} &= \frac{(a_n + jb_n k_{n3})^2}{(a_n - jb_n k_{n3})^2} \Rightarrow e^{jk_{n3}L_3} = \pm \frac{a_n + jb_n k_{n3}}{a_n - jb_n k_{n3}}, \end{aligned}$$

with arbitrary signs  $\pm$ . If in addition to this all surfaces are rigid and motionless (acoustically hard and passive), normal component of pressure gradient at them is zero (eq. (2.27)) and the boundary condition in eq. (3.75) reads with  $a_n = 0$  and  $b_n = 1$

$$\nabla_y \psi_n(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y}) = 0. \quad (3.77)$$

Hence:

$$\begin{aligned} e^{jk_{n1}L_1} &= \pm 1, \\ e^{jk_{n2}L_2} &= \pm 1, \\ e^{jk_{n3}L_3} &= \pm 1, \end{aligned}$$

which is satisfied only if

$$\begin{aligned} k_{n1} &= n_{L_1} \pi / L_1, \\ k_{n2} &= n_{L_2} \pi / L_2, \\ k_{n3} &= n_{L_3} \pi / L_3, \end{aligned} \quad (3.78)$$

where<sup>36</sup>  $n_{L_i} = 0, 1, 2, \dots$  for  $i = 1, 2, 3$ . The components of wave vector are thus real (no damping).

Dimensionless quantity  $kL$ , where  $L$  is in general some characteristic length scale and  $k$  is wave number, is called **Helmholtz number**,  $H_e$ . According to eq. (3.6), it can be written as

$$H_e = kL = \frac{\omega L}{c_0} = \frac{2\pi f L}{c_0} = \frac{2\pi L}{\lambda}. \quad (3.79)$$

Hence, it expresses ratio of the characteristic length scale (for example, length of a room, as in the case above, or width of a surface) and the sound wavelength. It allows us do describe something as acoustically very small (compact), when  $H_e \ll 1$ , or very large when  $H_e \gg 1$ . This has such strong implications, that whenever we say that something is small/large (a room, an object, etc.) in the context of (room) acoustics, we mean that the value of Helmholtz number (with an appropriate length scale) is small/large. Equation (3.78) suggests that room modes cannot occur for small values of Helmholtz number with the relevant room dimensions as the length scales.

Each mode  $n$  is associated with one particular combination of  $n_{L_1}$ ,  $n_{L_2}$ , and  $n_{L_3}$ . Its wave number  $k_n$  is defined by eq. (3.72). In order to determine shapes of the eigenmodes for the rectangular room with rigid walls, we notice that eq. (3.76) gives

$$\begin{aligned} -C_{11} + C_{21} &= 0, \\ C_{11}e^{-jk_{n1}L_1} - C_{21}e^{jk_{n1}L_1} &= 0, \\ -C_{12} + C_{22} &= 0, \\ C_{12}e^{-jk_{n2}L_2} - C_{22}e^{jk_{n2}L_2} &= 0, \\ -C_{13} + C_{23} &= 0, \\ C_{13}e^{-jk_{n3}L_3} - C_{23}e^{jk_{n3}L_3} &= 0. \end{aligned} \quad (3.80)$$

For  $k_{ni}$  from eq. (3.78), these are only 3 linearly independent equations:

$$\begin{aligned} C_{11} - C_{21} &= 0, \\ C_{12} - C_{22} &= 0, \\ C_{13} - C_{23} &= 0. \end{aligned} \quad (3.81)$$

We can use them in eq. (3.74) to obtain:

$$\begin{aligned} \psi_{n1} &= 2C_{11} \cos(k_{n1}x_1), \\ \psi_{n2} &= 2C_{12} \cos(k_{n2}x_2), \\ \psi_{n3} &= 2C_{13} \cos(k_{n3}x_3). \end{aligned}$$

---

<sup>36</sup>Negative values  $n_{L_i} = -1, -2, \dots$  for  $k_{ni} < 0$  would also satisfy the equalities. However, these solutions are already included, as can be seen in eq. (3.80) below. Since  $C_{11} = C_{21}$ ,  $C_{12} = C_{22}$ , and  $C_{13} = C_{23}$ , we can insist that  $k_{n1}$ ,  $k_{n2}$ , and  $k_{n3}$  are strictly non-negative and still keep all the solutions.

Equation (3.69) then reads

$$\begin{aligned}\psi_n(\mathbf{x}) &= \psi_{n1}\psi_{n2}\psi_{n3} = 8C_{11}C_{12}C_{13} \cos(k_{n1}x_1) \cos(k_{n2}x_2) \cos(k_{n3}x_3) \\ &= 8C_{11}C_{12}C_{13} \prod_{i=1}^3 \cos(k_{ni}x_i).\end{aligned}\quad (3.82)$$

As expected, three coefficients,  $C_{11}$ ,  $C_{12}$ , and  $C_{13}$ , remain unknown. Their product is a constant which can always multiply an eigenfunction to give essentially the same eigenfunction (with the same eigenvalue).

We can now write tailored Green's function with damping, which is given in eq. (3.61):

$$\hat{G}_{tail}(\mathbf{x}|\mathbf{y}) = 64c_0^2 C_{11}^2 C_{12}^2 C_{13}^2 \sum_n \frac{\prod_{i=1}^3 \cos(k_{ni}x_i) \prod_{i=1}^3 \cos(k_{ni}y_i)}{K_n(\omega_n^2 + 2j\zeta_n\omega_n - \omega^2)}.$$

The constant  $K_n$  is determined from eq. (3.54):

$$\begin{aligned}K_n &= \int_V \psi_n^*(\mathbf{x})\psi_n(\mathbf{x})d^3x \\ &= 64C_{11}^2 C_{12}^2 C_{13}^2 \int_0^{L_1} \int_0^{L_2} \int_0^{L_3} \cos^2(k_{n1}x_1) \cos^2(k_{n2}x_2) \cos^2(k_{n3}x_3) dx_1 dx_2 dx_3 \\ &= 64C_{11}^2 C_{12}^2 C_{13}^2 \int_0^{L_1} \cos^2(k_{n1}x_1) dx_1 \int_0^{L_2} \cos^2(k_{n2}x_2) dx_2 \int_0^{L_3} \cos^2(k_{n3}x_3) dx_3.\end{aligned}$$

The first integral equals

$$\begin{aligned}\int_0^{L_1} \cos^2(k_{n1}x_1) dx_1 &= \frac{1}{k_{n1}} \int_0^{k_{n1}L_1} \cos^2(k_{n1}x_1) d(k_{n1}x_1) \\ &= \frac{1}{2k_{n1}} [(k_{n1}L_1) + \sin(k_{n1}L_1) \cos(k_{n1}L_1)] = \frac{L_1}{2},\end{aligned}$$

For the last equality, we used the fact that  $\sin(k_{n1}L_1) = 0$ , which follows from eq. (3.78). In the same manner, we can calculate the remaining two integrals, which equal  $L_2/2$  and  $L_3/2$ , respectively. Therefore,

$$K_n = 8C_{11}^2 C_{12}^2 C_{13}^2 L_1 L_2 L_3,$$

where  $V = L_1 L_2 L_3$  is volume of the room, and

$$\boxed{\hat{G}_{tail}(\mathbf{x}|\mathbf{y}) = \frac{8c_0^2}{L_1 L_2 L_3} \sum_n \frac{\prod_{i=1}^3 \cos(k_{ni}x_i) \prod_{i=1}^3 \cos(k_{ni}y_i)}{(\omega_n^2 + 2j\zeta_n\omega_n - \omega^2)}} \quad (3.83)$$

The three unknown coefficients got cancelled by the normalization with  $K_n$ , which made them irrelevant for the tailored Green's function. Finally, complex sound pressure amplitude can be calculated from eq. (3.49) or, if the sound source is both compact and omnidirectional, from eq. (3.63):

$$\boxed{\hat{p}(\mathbf{x}) = \frac{8c_0^2 \hat{Q}(\mathbf{y})}{L_1 L_2 L_3} \sum_n \frac{\prod_{i=1}^3 \cos(k_{ni}x_i) \prod_{i=1}^3 \cos(k_{ni}y_i)}{(\omega_n^2 + 2j\zeta_n\omega_n - \omega^2)}} \quad (3.84)$$

Sound field is superposition of all the modes which are excited by the source located at certain point in the room. The upper part of Fig. 5 shows amplitude of the frequency response (in dB) calculated using eq. (3.83) (which is equivalent to eq. (3.84) with  $\hat{Q} = 1$ ) for two values of reverberation time. The damping constant is expressed from eq. (3.35) and frequency resolution is 0.1Hz. Figure 5 (below) shows the same frequency responses calculated numerically, with lower frequency resolution of 1Hz for shorter computation time.

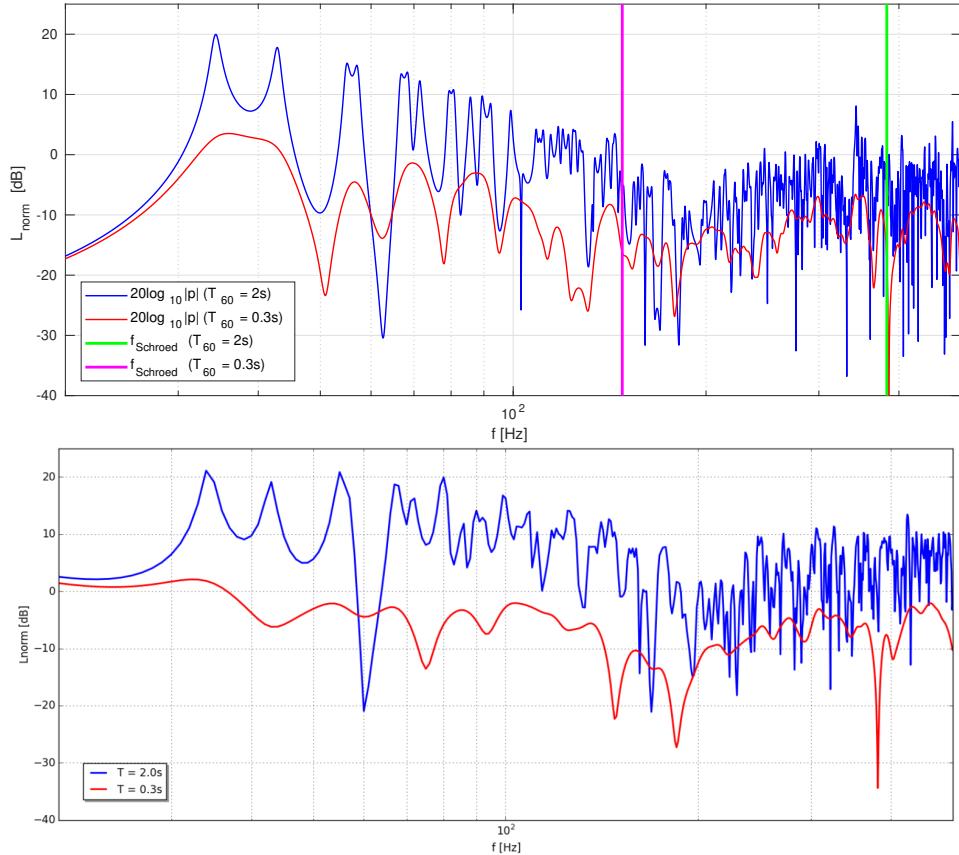


Figure 5: Frequency response of a rectangular room with rigid walls; the room dimensions are  $(L_1, L_2, L_3) = (5\text{m}, 4\text{m}, 3\text{m})$ , receiver location is  $(x_1, x_2, x_3) = (4.5\text{m}, 0.5\text{m}, 0.5\text{m})$ , and source location is  $(y_1, y_2, y_3) = (0.1\text{m}, 0.1\text{m}, 0.1\text{m})$ ; speed of sound is  $c_0 = 343\text{m/s}$  and reverberation time is  $T_{60} = 2\text{s}$  and  $T_{60} = 0.3\text{s}$ ; the frequency response is (above) based on the lowest 1000 modes (with frequencies up to 512Hz) calculated analytically, with the frequency resolution 0.1Hz (below) calculated numerically with the frequency step 1Hz.

Note that the overlapping modes can add constructively or cancel each other, depending on the signs of the cosines in eq. (3.83), which represent the modes  $\psi_n$ . This causes higher peaks or deeper drops of amplitude of the response,  $|p| = |\hat{p}|$ . The figure also demonstrates how strong variations of the response with frequency, which are often undesired (particularly when the modes are not dense enough at low frequencies to be

indistinguishable by human hearing), can be smoothed out by damping the modes, that is decreasing their Q-factor (compare with Fig. 4). Additional consequence is drop of the total energy of the sound field in the room, which can be desirable, in the case of noise control, but also undesirable, if the sound is a carrier of information and produced by a weak source.

### 3.7 Modal density

It is clear from Fig. 5 that the density of modes and their overlapping increase with frequency. This necessarily leads to higher computational requirements for calculating sound fields, also for the numerical calculations of more complicated geometries. On the other hand, human hearing has a limited frequency resolution which drops with frequency. It appears to average sound energy within certain finite frequency ranges in a similar manner as it performs averaging over time. This makes a detailed calculation of each higher mode unnecessary. Hence, it is of interest to inspect how the number of modes and their overlapping depend on frequency.

According to equations (3.72) and (3.78), resonance frequencies (frequencies of the modes) can be represented as points in  $k$ -space scaled with the factor  $f/k = c_0/(2\pi)$ . For the rectangular room:

$$f_n^2 = \left(\frac{c_0}{2\pi}\right)^2 k_n^2 = \frac{c_0^2}{4\pi^2} \sum_{i=1}^3 k_{ni}^2 = \left(\frac{c_0 n_{L_1}}{2L_1}\right)^2 + \left(\frac{c_0 n_{L_2}}{2L_2}\right)^2 + \left(\frac{c_0 n_{L_3}}{2L_3}\right)^2. \quad (3.85)$$

The number of frequencies  $f_n$  which are below frequency  $f$  is given approximately as the ratio of volume  $(4f^3\pi/3)/8$  of the spherical sector in the first octant of the scaled  $k$ -space and its smallest volume element, which equals  $(c_0/2L_1)(c_0/2L_2)(c_0/2L_3) = c_0^3/(8V)$ . This is only approximate because the points on the boundaries of the octant are not counted equally as the inner points. In particular, only half of the modes with one of the values  $n_{L_i}$  equal to zero are included and only one quarter of the modes with two out of three  $n_{L_i}$  equal to zero. Consequently, the approximation is valid for large values of  $f$ , when the number of the omitted modes on the boundaries and edges of the spherical sector is small compared to the total number of points. The ratio gives:

$$N_{f_n < f} \approx \frac{(4f^3\pi/3)/8}{c_0^3/(8V)} = \frac{4\pi V f^3}{3c_0^3}. \quad (3.86)$$

The accurate equality which includes all modal frequencies correctly can also be derived and it reads

$$N_{f_n < f} = \frac{4\pi V f^3}{3c_0^3} + \frac{\pi S f^2}{4c_0^2} + \frac{L f}{8c_0}, \quad (3.87)$$

where  $S = 2(L_1 L_2 + L_1 L_3 + L_2 L_3)$  is total area of the boundary surfaces of the room and  $L = 4(L_1 + L_2 + L_3)$  is total length of the edges. As we see, the contribution of the last two correction terms becomes irrelevant for high  $f$ . It is interesting to note that these

and the following equations hold even for an arbitrary shape of the room when  $f \rightarrow \infty$ . For later reference, Fig. 6 shows logarithms of the ratios  $L/S$  and  $S/V$  for a rectangular room scaled with the longest dimension of the room,  $L_1$ .

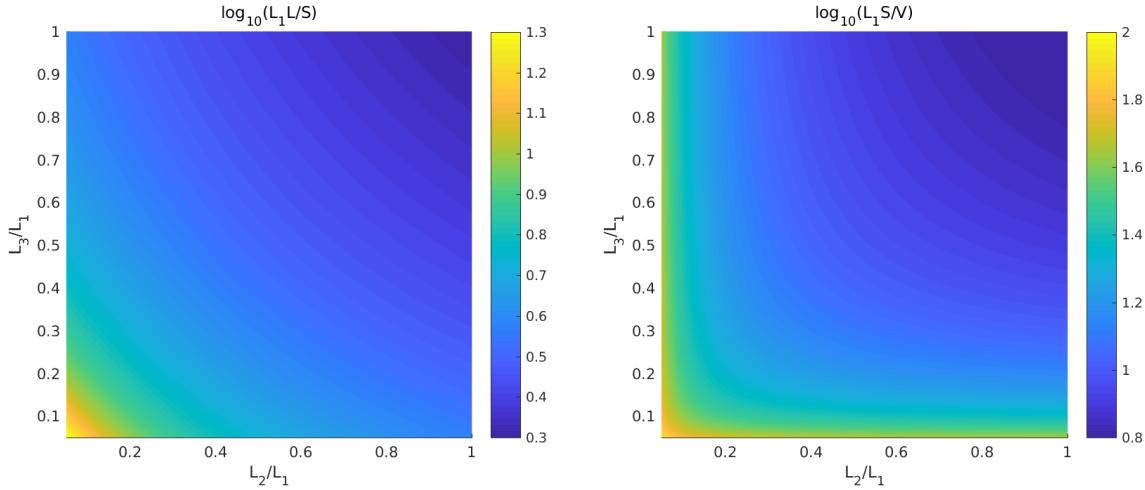


Figure 6: Logarithms of the ratios  $L/S$  (left) and  $S/V$  (right) in a rectangular room scaled with the longest dimension  $L_1$ .

The number of modes increases rapidly with frequency,  $N_{f_n < f} \sim f^3$ . Density of modes, or the number of resonance frequencies per hertz, is expressed as the derivative

$$\frac{dN_{f_n < f}}{df} = \frac{4\pi V f^2}{c_0^3} + \frac{\pi S f}{2c_0^2} + \frac{L}{8c_0} \approx \frac{4\pi V f^2}{c_0^3}, \quad (3.88)$$

where the approximation holds for large  $f$ . On the other hand, the overlapping of modes will also depend on their widths. We recall that the resonance curves, given by eq. (3.64) and with the form presented in Fig. 4, have half-widths  $\Delta f_n = \Delta\omega_n/(2\pi) = \zeta_n/\pi$ . If the number of different modes within this frequency range is relatively large, say larger than 3, the modes are substantially overlapping and no single mode is expected to dominate. This happens when  $dN_{f_n < f}/df > 3/\Delta f_n$ , that is at the frequencies which satisfy

$$f > f_{Schroed} = \sqrt{\frac{3c_0^3}{4\pi V \Delta f_n}} = \sqrt{\frac{3c_0^3}{4V\zeta_n}} \approx 5500 \sqrt{\frac{1\text{m}^3/\text{s}^3}{V\zeta_n}}, \quad (3.89)$$

where  $f_{Schroed}$  is Schroeder frequency, named after M. Schroeder who first proposed it. Expressed in terms of reverberation time in eq. (3.35), its value is

$$f_{Schroed} \approx 2100 \sqrt{\frac{T_{60,\Delta f_n} \cdot 1\text{m}^3/\text{s}^3}{V}}. \quad (3.90)$$

The overlapping modes are assumed to have similar values of the damping constant and reverberation time within the range  $\Delta f_n$ . However, broadband reverberation time

or average reverberation time in central octave bands is often used in practice for the estimation of Schroeder frequency. The inaccuracies which are introduced thereby are usually not critical, since Schroeder frequency is used anyway as an indicator of high/low modal overlapping, not as a clear limit between the two regimes.

Schroeder frequency indicates the low frequency range in which resonance behaviour of a room can be perceivable and significant. As an example, its value for the analysed rectangular room is shown in the upper part of Fig. 5. Above Schroeder frequency, the overlap of a large number of modes leads to (after averaging in relatively narrow frequency bands) smoother frequency response of the room. Figure 7 shows density of the modes of the same room as in Fig. 5 and their cumulative distribution, according to equations (3.86)-(3.88) with and without the correction terms in eq. (3.87). Schroeder frequency for  $T_{60} = 2\text{s}$  is also indicated, which is around 383Hz (this is more realistic value for the room considered here with very weak damping than the value for  $T_{60} = 0.3\text{s}$ , which is around 148Hz). The number of modes below Schroeder frequency in weakly damped rooms can be quite large for practical calculations.

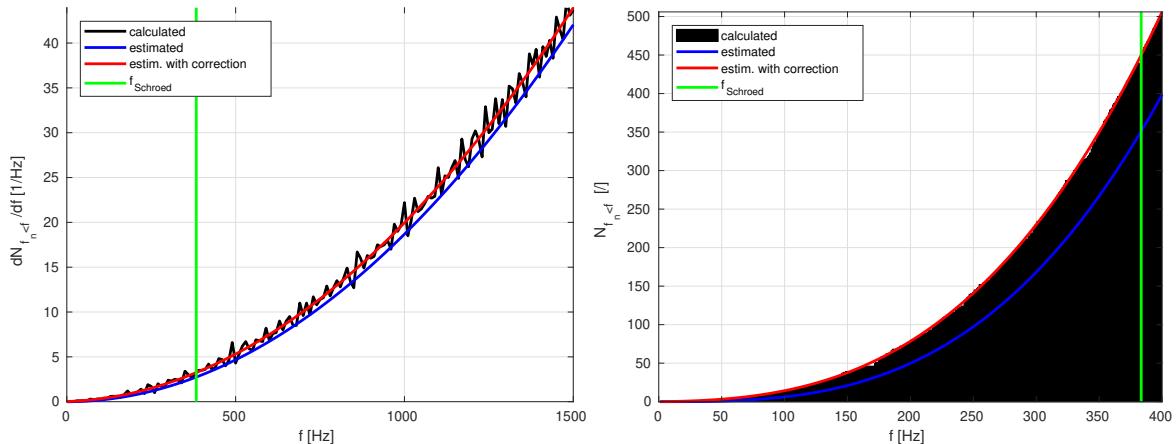


Figure 7: Density of modes (left) and cumulative distribution (right) for a rectangular room with dimensions  $(L_1, L_2, L_3) = (3\text{m}, 4\text{m}, 5\text{m})$ . The calculated values are obtained using eq. (3.85). For smoothness of the curve in the left diagram, frequency bands with the width 10Hz were considered. The estimated values (equations (3.86)-(3.88)) are given with and without the two correction terms in eq. (3.87). Schroeder frequency is shown for reverberation time 2s.

In this section we considered the modal theory and obtained an analytical solution for the sound field in a rectangular room with rigid walls. Although very idealized, the case provided valuable information on the resonant behaviour of rooms and lead to some general conclusions. In practice, many rooms do actually have a rectangular shape or can be approximated so. The equations such as eq. (3.85) can then be used for rough estimations (sometimes even with a surprising accuracy), at least at low frequencies, when the treatment of resonances is of special interest, and for low damping.

If complicated geometries and boundary conditions are to be treated more accurately, the solutions of Helmholtz equation or tailored Green's functions have to be estimated numerically, typically using Finite Element Method (FEM) or Boundary Element Method (BEM). The former method, which is based on the discretization of the three-dimensional computational domain, is often used for solving Helmholtz equation (3.30) directly, supplied with boundary conditions, such as the one in eq. (3.26). The latter technique is based on the discretization of surfaces and often used for solving the integral equation (3.46), with the well known free-space Green's function from eq. (3.47).

One of the major difficulties for accurate numerical (or analytical) solutions is appropriate representation of realistic boundary conditions. As already discussed, the modes have to satisfy the conditions at all the boundaries simultaneously, since the time dependence is suppressed from the calculations. Moreover, complicated materials and micro geometries of surfaces cannot be easily mathematically expressed. Besides this, the applicability of numerical solvers in room acoustics is still largely limited by reasonable computational requirements. In practice, a certain limited number of modes can be solved, which means that satisfactory accuracy can be achieved only at relatively low frequencies (actually, Helmholtz numbers, since modal density at certain frequency depends on the dimensions of the room, as well), when the density of modes is low. Schroeder frequency can be seen as a very rough indicator of the highest frequencies which can be covered with such calculations. The calculations for higher frequencies are usually unacceptable, especially if multiple scenarios should be assessed in some reasonable time. For these reasons, it is worthwhile to reconsider sound field analysis in time domain, in search of a more suitable methodology for high frequencies, particularly for large rooms.

## 4 Analysis in time domain

Modal analysis from the previous section gave us basic insight into acoustic behaviour of rooms (and closed spaces in general), with particular applicability of the approach for relatively low frequencies at which the resonant behaviour of rooms is dominant. However, the analysis becomes impractical when many modes have to be included which contribute to the response of the room. With the exception of small rooms, this is most often the case in practice, if considerable part of the audible frequency range should be covered. In addition to this, all boundary conditions in the calculations of sound fields or tailored Green's functions had to be satisfied at once by the eigenfunctions, since the time dependence of the acoustic quantities was omitted. This makes the calculations involved, except for the simplest room geometries. In this section, we return to the analysis in time domain, which should relax these requirements.

### 4.1 Green's function in time domain

Similarly as Helmholtz equation (3.30), wave equation (2.22) can be formally solved in terms of Green's function. Analogously to eq. (3.41), it has to satisfy the following equation:

$$\frac{1}{c_0^2} \frac{\partial^2 G(\mathbf{x}, t | \mathbf{y}, t)}{\partial t^2} - \nabla_x^2 G(\mathbf{x}, t | \mathbf{y}, \tau) = \delta(\mathbf{x} - \mathbf{y}, t - \tau) = \delta(\mathbf{x} - \mathbf{y})\delta(t - \tau), \quad (4.1)$$

and from eq. (2.23):

$$\frac{1}{c_0^2} \frac{\partial^2 G(\mathbf{x}, t | \mathbf{y}, \tau)}{\partial \tau^2} - \nabla_y^2 G(\mathbf{x}, t | \mathbf{y}, \tau) = \delta(\mathbf{x} - \mathbf{y}, t - \tau) = \delta(\mathbf{x} - \mathbf{y})\delta(t - \tau). \quad (4.2)$$

Following the procedure in section 3.4, we subtract the last equality multiplied with  $p(\mathbf{y}, \tau)$  from eq. (2.23) multiplied with  $G(\mathbf{x}, t | \mathbf{y}, \tau)$  and obtain (we leave out the arguments of  $G$  for brevity):

$$\begin{aligned} G \frac{1}{c_0^2} \frac{\partial^2 p(\mathbf{y}, \tau)}{\partial \tau^2} - G \nabla_y^2 p(\mathbf{y}, \tau) - p(\mathbf{y}, \tau) \frac{1}{c_0^2} \frac{\partial^2 G}{\partial \tau^2} + p(\mathbf{y}, \tau) \nabla_y^2 G - G q(\mathbf{y}, \tau) \\ = -p(\mathbf{y}, \tau) \delta(\mathbf{x} - \mathbf{y}) \delta(t - \tau). \end{aligned}$$

We integrate the result with respect to  $\mathbf{y}$  over some volume  $V$  and with respect to  $\tau$  over the time interval<sup>37</sup>  $(t_0^-, t)$ , where  $t_0^- < t_0$  is time just before  $t_0$ , when the source  $q$  is

---

<sup>37</sup>Due to causality in time domain, we can limit the integration up to the reception time  $t > t_0$ , since no event which happens after  $t$  can influence the field at  $t$ .

switched on. This gives:

$$\begin{aligned}
& \int_{t_0^-}^t \int_V Gq(\mathbf{y}, \tau) d^3 \mathbf{y} d\tau + \frac{1}{c_0^2} \int_{t_0^-}^t \int_V \left( p(\mathbf{y}, \tau) \frac{\partial^2 G}{\partial \tau^2} - G \frac{\partial^2 p(\mathbf{y}, \tau)}{\partial \tau^2} \right) d^3 \mathbf{y} d\tau \\
& + \int_{t_0^-}^t \int_V (G \nabla_y^2 p(\mathbf{y}, \tau) - p(\mathbf{y}, \tau) \nabla_y^2 G) d^3 \mathbf{y} d\tau \\
& = \int_{t_0^-}^t \int_V p(\mathbf{y}, \tau) \delta(\mathbf{x} - \mathbf{y}) \delta(t - \tau) d^3 \mathbf{y} d\tau \\
& = \int_{t_0^-}^t \left( \int_V p(\mathbf{y}, \tau) \delta(\mathbf{x} - \mathbf{y}) d^3 \mathbf{y} \right) \delta(t - \tau) d\tau \\
& = \int_{t_0^-}^t p(\mathbf{x}, \tau) \delta(t - \tau) d\tau = p(\mathbf{x}, t).
\end{aligned}$$

We notice that

$$p \frac{\partial^2 G}{\partial \tau^2} - G \frac{\partial^2 p}{\partial \tau^2} = \frac{\partial}{\partial \tau} \left( p \frac{\partial G}{\partial \tau} \right) - \frac{\partial p}{\partial \tau} \frac{\partial G}{\partial \tau} - \frac{\partial}{\partial \tau} \left( G \frac{\partial p}{\partial \tau} \right) + \frac{\partial G}{\partial \tau} \frac{\partial p}{\partial \tau} = \frac{\partial}{\partial \tau} \left( p \frac{\partial G}{\partial \tau} - G \frac{\partial p}{\partial \tau} \right)$$

and similarly

$$G \nabla_y^2 p - p \nabla_y^2 G = \nabla_y \cdot (G \nabla_y p - p \nabla_y G).$$

After also applying the divergence theorem, we obtain:

$$\begin{aligned}
p(\mathbf{x}, t) &= \int_{t_0^-}^t \int_V Gq(\mathbf{y}, \tau) d^3 \mathbf{y} d\tau + \int_{t_0^-}^t \oint_S (G \nabla_y p(\mathbf{y}, \tau) - p(\mathbf{y}, \tau) \nabla_y G) \cdot \mathbf{n} d^2 \mathbf{y} d\tau \\
&+ \frac{1}{c_0^2} \int_{t_0^-}^t \int_V \frac{\partial}{\partial \tau} \left( p(\mathbf{y}, \tau) \frac{\partial G}{\partial \tau} - G \frac{\partial p(\mathbf{y}, \tau)}{\partial \tau} \right) d^3 \mathbf{y} d\tau,
\end{aligned}$$

where the closed surface  $S$  around the volume  $V$  has at each point a unit vector  $\mathbf{n}$  normal to it and pointing outwards. The last integral equals

$$\begin{aligned}
& \int_{t_0^-}^t \int_V \frac{\partial}{\partial \tau} \left( p(\mathbf{y}, \tau) \frac{\partial G}{\partial \tau} - G \frac{\partial p(\mathbf{y}, \tau)}{\partial \tau} \right) d^3 \mathbf{y} d\tau \\
&= \int_V \left[ \int_{t_0^-}^t \frac{\partial}{\partial \tau} \left( p(\mathbf{y}, \tau) \frac{\partial G}{\partial \tau} - G \frac{\partial p(\mathbf{y}, \tau)}{\partial \tau} \right) d\tau \right] d^3 \mathbf{y} \\
&= \int_V \left[ \left( p(\mathbf{y}, \tau) \frac{\partial G}{\partial \tau} - G \frac{\partial p(\mathbf{y}, \tau)}{\partial \tau} \right)_{\tau=t} - \left( p(\mathbf{y}, \tau) \frac{\partial G}{\partial \tau} - G \frac{\partial p(\mathbf{y}, \tau)}{\partial \tau} \right)_{\tau=t_0^-} \right] d^3 \mathbf{y} \\
&= \int_V \left( G \frac{\partial p(\mathbf{y}, \tau)}{\partial \tau} - p(\mathbf{y}, \tau) \frac{\partial G}{\partial \tau} \right)_{\tau=t_0^-} d^3 \mathbf{y},
\end{aligned}$$

where  $(\cdot)_{\tau=t}$  indicates that the expression in the parentheses is to be evaluated at time  $t$  (that is,  $(f(\tau))_{\tau=t} = f(t)$ ) and  $G(\mathbf{x}, t | \mathbf{y}, \tau) = 0$  for  $\tau = t$  and any  $\mathbf{x} \neq \mathbf{y}$  (we are only

interested in  $p$  outside the source region and no instantaneous propagation is allowed). Therefore,

$$\begin{aligned} p(\mathbf{x}, t) &= \int_{t_0^-}^{t^-} \int_V q(\mathbf{y}, \tau) G d^3 \mathbf{y} d\tau \\ &+ \int_{t_0^-}^{t^-} \oint_S (G \nabla_y p(\mathbf{y}, \tau) - p(\mathbf{y}, \tau) \nabla_y G) \cdot \mathbf{n} d^2 \mathbf{y} d\tau \\ &+ \frac{1}{c_0^2} \int_V \left( G \frac{\partial p(\mathbf{y}, \tau)}{\partial \tau} - p(\mathbf{y}, \tau) \frac{\partial G}{\partial \tau} \right)_{\tau=t_0^-} d^3 \mathbf{y}, \end{aligned} \quad (4.3)$$

where  $t^-$  denotes time just before  $t$ . We can compare this with eq. (3.46). Besides the necessary integration over time due to recovered time dependence, the additional term is due to the initial conditions at the moment when the source is switched on. It vanishes if we assume that  $p = 0$  for all  $t < t_0$  and we can replace  $t_0^-$  with  $t_0$  in the integration intervals. The equation becomes

$$p(\mathbf{x}, t) = \int_{t_0}^{t^-} \int_V q(\mathbf{y}, \tau) G d^3 \mathbf{y} d\tau + \int_{t_0}^{t^-} \oint_S (G \nabla_y p(\mathbf{y}, \tau) - p(\mathbf{y}, \tau) \nabla_y G) \cdot \mathbf{n} d^2 \mathbf{y} d\tau \quad (4.4)$$

for every  $\mathbf{x} \neq \mathbf{y}$ . The first integral is contribution of the direct sound from the source and the second term is due to the boundaries. As before, the second term vanishes for tailored Green's function satisfying the same boundary conditions as  $p$  and the equation simplifies to

$$p(\mathbf{x}, t) = \int_{t_0}^{t^-} \int_V q(\mathbf{y}, \tau) G_{tail} d^3 \mathbf{y} d\tau = \int_{-\infty}^{t^-} \int_V q(\mathbf{y}, \tau) G_{tail} d^3 \mathbf{y} d\tau, \quad (4.5)$$

where we formally extended the time integration to  $-\infty$  ( $q = 0$  for  $-\infty < t < t_0$ ). This equation is the counterpart of the solution of Helmholtz equation in frequency domain, eq. (3.49).

Free space Green's function of wave equation (2.22), which satisfies eq. (4.1) for an outgoing wave is

$$G_{free}(\mathbf{x}, t | \mathbf{y}, \tau) = \frac{\delta(t - \tau - |\mathbf{x} - \mathbf{y}|/c_0)}{4\pi|\mathbf{x} - \mathbf{y}|} = \frac{\delta(t - \tau - r/c_0)}{4\pi r}. \quad (4.6)$$

In contrast to  $\hat{G}_{free}$  from the previous section (in eq. (3.47)),  $G_{free}$  is scaled delta function, non-zero only when  $\tau = t - r/c_0$ , which is called emission (or retarded) time. Actually, the delta function  $\delta(t - \tau)$  is an eigenfunction in time domain like the exponential function  $e^{-jkr}$  in free space in frequency domain. Every function of time can be represented as an integral of scaled  $\delta(t - \tau)$  with respect to  $\tau$  (the sampling property of symmetric delta function in eq. (3.44) with  $x$  and  $y$  replaced by  $\tau$  and  $t$ ) like any spherically symmetric function of space can be represented as an integral of scaled  $e^{-jkr}$  with respect to  $k$  (the Fourier transform in eq. (3.23) without dependence on  $\mathbf{x}$  and with

$\omega$  and  $t$  replaced by  $r$  and  $k$ ) or, indeed, any function of time as an integral of scaled  $e^{j\omega t}$  with respect to  $\omega$  (inverse Fourier transform in eq. (3.22) without dependence on  $\mathbf{x}$ ). The delta function and  $e^{j\omega t}$  are eigenfunctions (in the absence of time constraints) in time and frequency domain, respectively.

If a point source located at  $\mathbf{y}$  emits an ideal impulse at time  $t_0$ , we can write the source function as a generalized function  $q(\mathbf{y}', \tau) = Q(\mathbf{y}', \tau)\delta(\mathbf{y}' - \mathbf{y})\delta(\tau - t_0)$ , with  $Q$  in kg/s, so that

$$\int_{-\infty}^{\infty} \int_V q(\mathbf{y}', \tau) d^3 \mathbf{y}' d\tau = \int_{-\infty}^{\infty} \int_V Q(\mathbf{y}', \tau) \delta(\mathbf{y}' - \mathbf{y}) \delta(\tau - t_0) d^3 \mathbf{y}' d\tau = Q(\mathbf{y}, t_0). \quad (4.7)$$

For a tailored Green's function, from eq. (4.5):

$$\begin{aligned} p(\mathbf{x}, t) &= \int_{-\infty}^{t^-} \int_V q(\mathbf{y}', \tau) G_{tail}(\mathbf{x}, t | \mathbf{y}', \tau) d^3 \mathbf{y}' d\tau \\ &= \int_{-\infty}^{t^-} \int_V Q(\mathbf{y}', \tau) \delta(\mathbf{y}' - \mathbf{y}) \delta(\tau - t_0) G_{tail}(\mathbf{x}, t | \mathbf{y}', \tau) d^3 \mathbf{y}' d\tau \\ &= Q(\mathbf{y}, t_0) G_{tail}(\mathbf{x}, t | \mathbf{y}, t_0), \end{aligned} \quad (4.8)$$

if  $t_0 \leq t^-$  (or zero otherwise). This is very similar to the result in eq. (3.51) with added time dependence. In this case, Green's function corresponds to the impulse response (the impulse is emitted by the point source), which contains all the information about sound propagation in a room. Green's function in frequency domain, which was treated in the previous section, corresponds to the frequency response.

In free space,  $G_{tail} = G_{free}$  and

$$p_{free}(\mathbf{x}, t) = Q(\mathbf{y}, t_0) G_{free}(\mathbf{x}, t | \mathbf{y}, t_0) = Q(\mathbf{y}, t_0) \frac{\delta(t - t_0 - |\mathbf{x} - \mathbf{y}|/c_0)}{4\pi|\mathbf{x} - \mathbf{y}|} \quad (4.9)$$

is merely a scaled and delayed (due to propagation from  $\mathbf{y}$  to  $\mathbf{x}$ ) version of the radiated impulse<sup>38</sup>. As expected, the delay  $|\mathbf{x} - \mathbf{y}|/c_0$  has the same effect as the phase shift  $jk|\mathbf{x} - \mathbf{y}|$  in eq. (3.52). In general, tailored Green's function of a closed space in time domain is a series of impulses as in eq. (4.9) but additionally scaled (by different propagation path lengths, absorption at the surfaces, etc.) and taking the locations  $\mathbf{y}$  in the source region and at the surfaces (for the reflection) as the arguments. Analogously to the superposition of reflections with different delays,  $\hat{G}_{tail}$  in closed spaces in frequency domain is a series of eigenfunctions at distinct frequencies (eq. (3.53)).

The fact that  $G_{tail}$  is a generalized function, a series of delta impulses, has another consequence. Sound pressure as expressed by the last equalities in equations (4.8) and (4.9) is not physical. The reason is that the integrals in eq. (4.8) did not act on the delta

<sup>38</sup>General time dependence of  $q$  of a point source in free space will be given in eq. (4.11).

functions of  $G_{tail}$ . For example, a physical result in free space would be

$$\begin{aligned} p_{free}(\mathbf{x}, t) &= \int_{-\infty}^{t^-} \int_V q(\mathbf{y}', \tau) \frac{\delta(t - \tau - r/c_0)}{4\pi r} d^3\mathbf{y}' d\tau \\ &= \int_V \frac{1}{4\pi r} \left( \int_{-\infty}^{t^-} q(\mathbf{y}', \tau) \delta(\tau - (t - r/c_0)) d\tau \right) d^3\mathbf{y}' = \int_V \frac{q(\mathbf{y}', t - r/c_0)}{4\pi r} d^3\mathbf{y}', \end{aligned} \quad (4.10)$$

where we used the symmetry of delta function and the fact that  $r$  is not a function of time (the source and receiver are stationary). If the source is compact,  $r$  is nearly constant in the source region and we can write

$$\begin{aligned} p_{free}(\mathbf{x}, t) &= \frac{1}{4\pi r} \int_V q(\mathbf{y}', t - r/c_0) d^3\mathbf{y}' = \frac{1}{4\pi r} \int_V Q(\mathbf{y}', t - r/c_0) \delta(\mathbf{y}' - \mathbf{y}) d^3\mathbf{y}' \\ &= \frac{1}{4\pi r} Q(\mathbf{y}, t - r/c_0), \end{aligned} \quad (4.11)$$

where now, similarly as in eq. (3.50),  $q(\mathbf{y}', \tau) = Q(\mathbf{y}', \tau) \delta(\mathbf{y}' - \mathbf{y})$ ,  $Q$  has again the unit  $\text{kg/s}^2$ , and

$$Q(\mathbf{y}, \tau) = \int_V q(\mathbf{y}', \tau) d^3\mathbf{y}' = \int_V Q(\mathbf{y}', \tau) \delta(\mathbf{y}' - \mathbf{y}) d^3\mathbf{y}'. \quad (4.12)$$

Hence, eq. (4.11) holds for any function of time  $q$ , not only for an ideal impulse. It only states that for  $p_{free}$  at time  $t$ ,  $Q$  should be evaluated at the emission time  $t - r/c_0$ .

## 4.2 LTI systems and impulse response

A type of a system which is completely described with its impulse response is linear time invariant (LTI) system. If we assume that all sources and receivers in a room are stationary, the room can also be treated as an LTI system. Its impulse response  $g(t)$  is response at certain receiver location to the impulse emitted from a point source at some other location. The resulting signal at the receiver location is given by the convolution integral ( $\delta$ ,  $s$ ,  $p$ , and  $g$  are here real-valued functions of time only):

$$p(t + r/c_0) = \int_{-\infty}^{\infty} s(\tau) g(t - \tau) d\tau = \int_{-\infty}^{\infty} g(\tau) s(t - \tau) d\tau, \quad (4.13)$$

where  $s(t)$  is the input signal (emitted by the source), as in Fig. 2. Time dependence of the first equality can be compared with eq. (4.5), with  $s$  acting as  $q$  (in fact,  $Q$  for a point source, due to the volume integral) and  $g$  as  $G_{tail}$ . By inserting, for example, free space Green's function from eq. (4.6), we can see that the essential difference is that the emission time  $t - r/c_0$  is replaced with the reception time  $t \geq r/c_0$ . This explains the time shift to  $t + r/c_0$  on the left-hand side of eq. (4.13). Equation (4.13) and thus the integral with respect to time in eq. (4.5) are convolutions and Green's function is indeed an ideal impulse response. However, it must not be forgotten that in this case  $g(t)$  depends on the particular locations of the source and receiver ( $\mathbf{x}$  and  $\mathbf{y}$  in  $G_{tail}$ ).

The last integral in eq. (4.13) gives after replacing  $t + r/c_0$  with  $t$  (which does not affect the interval of the integral):

$$p(t) = \int_{-\infty}^{\infty} g(\tau) s(t - r/c_0 - \tau) d\tau. \quad (4.14)$$

Since  $\delta(t) = 0$  for  $t \neq 0$ , all causal systems have impulse responses which satisfy  $g(t) = 0$  for  $t < 0$  and start at  $t = 0$ , which is here time of arrival of the direct sound. The integral in eq. (4.14) thus becomes

$$p(t) = \int_0^{\infty} g(\tau) s_e(t - \tau) d\tau, \quad (4.15)$$

where  $s_e(t) = s(t - r/c_0)$  is given at the emission time. From the theory of LTI systems, we also know that in frequency domain convolution turns into product:

$$\hat{p}(\omega) = \hat{s}_e(\omega) \hat{g}(\omega), \quad (4.16)$$

and  $\hat{g}(\omega)$  is frequency response of the system. This has the same form as eq. (3.51).

### 4.3 Directivity

Working in time domain allows us to start our analysis with an isolated source in free space, which relates to the conditions before the emitted sound has reached any obstacle or boundary during its propagation. We will consider the boundaries later on. However, we can suppose that the source emits a complex sine wave in time with generic angular frequency  $\omega$ , that is,

$$Q(\mathbf{y}, \tau) = \hat{Q}(\mathbf{y}) e^{j\omega\tau}. \quad (4.17)$$

Complex sound pressure equals then

$$p(\mathbf{x}, t) = \frac{\hat{Q}(\mathbf{y})}{4\pi r} e^{j\omega(t-r/c_0)} = \frac{\hat{Q}(\mathbf{y})}{4\pi r} e^{j(\omega t - kr)}. \quad (4.18)$$

According to eq. (4.11),  $Q$  has emission time  $t - r/c_0$  as the argument and we left out the subscript  $_{free}$  for brevity. The complex amplitude is

$$\hat{p}(\mathbf{x}) = \frac{\hat{Q}(\mathbf{y})}{4\pi r} e^{-jkr}, \quad (4.19)$$

which is in full agreement with eq. (3.52).

Like  $\hat{G}_{free}$  in eq. (3.47), Green's function in eq. (4.6) is spherically symmetric, which leaves sound pressure amplitude dependent only on the distance from the source,  $r$ . When appropriate, directivity of a point source can be introduced explicitly with a real

and non-negative<sup>39</sup> dimensionless factor  $D_i$  (also called directivity) as<sup>40</sup>:

$$\boxed{p(\mathbf{x}, t) = \frac{\hat{Q}(\mathbf{y}) D_i(\theta, \phi)}{4\pi r} e^{j(\omega t - kr)}}, \quad (4.20)$$

for  $\mathbf{y} = 0$  and  $r = |\mathbf{x}| > 0$ . Here,  $0 \leq \theta \leq \pi$  and  $0 \leq \phi < 2\pi$  are polar and azimuthal angle, respectively, at the receiver's location  $\mathbf{x}$  given with:  $x_1 = r \sin(\theta) \cos(\phi)$ ,  $x_2 = r \sin(\theta) \sin(\phi)$ ,  $x_3 = r \cos(\theta)$ , for the source located at the origin  $\mathbf{y} = 0$  of the spherical coordinate system. The complex amplitude is therefore

$$\hat{p}(\mathbf{x}) = \frac{\hat{Q}(\mathbf{y}) D_i(\theta, \phi)}{4\pi r} e^{-jkr}. \quad (4.21)$$

Like  $\hat{p}$  and  $\hat{Q}$ ,  $D_i$  is also a function of frequency (usually increasing the directivity with frequency). Hence, the values of any descriptor of directivity should be given in the entire frequency range of interest, typically in octave frequency bands.

Directivity  $D_i$  is most often normalized with its maximum value  $D_{i,max}$  (which is typically at the axis of an axisymmetric source, for example a loudspeaker), so that its values are limited to the range between 0 and 1. Logarithm of such an angularly dependent quantity represents the radiation pattern<sup>41</sup>:

$$RP = 20 \log_{10} \left( \frac{D_i(\theta, \phi)}{D_{i,max}} \right) \leq 0 \text{dB}. \quad (4.22)$$

As an example, Fig. 8 shows different representations of the directivity of an ideal monopole and dipole as a function of  $\theta$  (the values do not depend on azimuth, so the curves in polar plots can be rotated around the horizontal axis to obtain the three-dimensional radiation patterns).

For convenience, directivity of a source is often expressed as a single number value. A simple but crude descriptor is coverage angle, which is the angular spread around the direction of maximum radiation of a source, over which the radiation pattern remains above some given value, typically -3dB or -6dB. It takes values between 0° and 360° (see, for example, Table 1.6). For a dipole, it equals 90° for -3dB and 120° for -6dB, while monopole has full coverage angle, 360°. For more complicated sources, the radiation patterns of which are not axisymmetric, it is usually given in horizontal and vertical planes.

Another approach to express directivity with a single number is with directivity factor (gain):

$$\gamma = \frac{4\pi}{\int_0^{4\pi} (D_i(\theta, \phi)/D_{i,max})^2 d\Omega} = \frac{4\pi D_{i,max}^2}{\int_0^{4\pi} D_i^2(\theta, \phi) d\Omega} \geq 1, \quad (4.23)$$

<sup>39</sup>Minus sign of a negative factor would be absorbed in  $\hat{Q}(\mathbf{y})$ .

<sup>40</sup>Directivity is by definition angular dependence of the acoustic far field (to be defined shortly) generated by the source. Therefore, the introduced representation of directivity of a source also holds only in the far field.

<sup>41</sup>Other normalizations are also possible, for example, with the power level of the source.

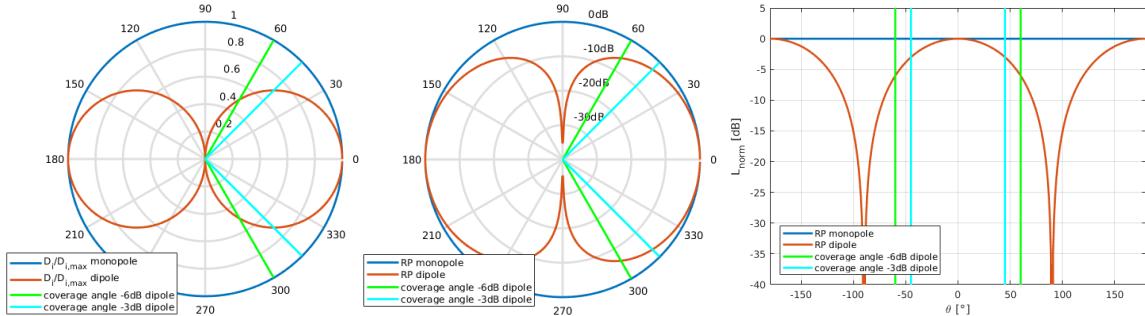


Figure 8: Directivity of an ideal point monopole and dipole.

where  $d\Omega = \sin(\theta)d\theta d\phi$  is the differential for solid angle. It essentially represents the ratio of maximum magnitude of time-averaged sound intensity produced by the source (in the direction of maximum radiation of the source) and its average value over all directions. If a source has acoustic power  $\langle P_q \rangle_T$ , it will effectively radiate  $\gamma \langle P_q \rangle_T$  in the direction of maximum radiation. Directivity factor is larger than or equal to 1. It equals 1 for a perfect monopole (hence the name gain,  $\gamma$  times more sound energy is radiated in the direction of maximum radiation than with a monopole with the same total radiated power) and for a dipole

$$\begin{aligned} \gamma &= \frac{4\pi}{\int_0^{4\pi} \cos^2(\theta) d\Omega} = \frac{4\pi}{\int_0^{2\pi} \int_0^\pi \cos^2(\theta) \sin(\theta) d\theta d\phi} \\ &= \frac{2}{\int_0^\pi \cos^2(\theta) \sin(\theta) d\theta} = -\frac{6}{[\cos^3(\pi) - \cos^3(0)]} = 3. \end{aligned} \quad (4.24)$$

Associated logarithmic quantity is

$$10 \log_{10} \gamma = 10 \log_{10} \left( \frac{4\pi D_{i,max}^2}{\int_0^{4\pi} D_i^2(\theta, \phi) d\Omega} \right) \geq 0 \text{dB}. \quad (4.25)$$

It equals 0dB for omnidirectional sources, around 4.8dB for a dipole, and generally rarely exceeds 20dB.

#### 4.4 Spherical and plane waves

Sound pressure amplitude in free space in equations (4.19) and (4.21) decays proportionally to  $1/r$  (regardless of the directivity of the source). In rooms, however, the propagating wave is expected to hit solid surfaces, either boundary surfaces of the room, or other objects inside the room. As implied by eq. (2.25), boundary conditions can be defined with respect to  $p$  and/or its gradient. The gradient is also related to the acoustic velocity by the conservation of momentum, eq. (2.9). Therefore, it is of interest to consider the gradient of acoustic pressure which we obtained above.

If we suppose that the point source is omnidirectional<sup>42</sup> with given source function  $Q = \hat{Q}e^{j\omega t}$ , pressure is a function of  $r$  and  $t$  only (eq. (4.18)) and we can replace  $\nabla_x p$  with  $\mathbf{e}_r \partial p / \partial r$  in spherical coordinates<sup>43</sup>, where  $\mathbf{e}_r$  is a unit vector parallel to and in the direction of sound propagation. Consequently:

$$\begin{aligned}\nabla_x p(\mathbf{x}, t) &= \mathbf{e}_r \frac{\partial}{\partial r} \left( \frac{\hat{Q}(\mathbf{y})}{4\pi r} e^{j(\omega t - kr)} \right) \\ &= \frac{\hat{Q}(\mathbf{y})}{4\pi} \left[ \left( \frac{1}{r} \frac{\partial}{\partial r} e^{j(\omega t - kr)} \right) + \left( e^{j(\omega t - kr)} \frac{\partial}{\partial r} \frac{1}{r} \right) \right] \mathbf{e}_r \\ &= \frac{\hat{Q}(\mathbf{y})}{4\pi} \left( -\frac{jk}{r} e^{j(\omega t - kr)} - \frac{1}{r^2} e^{j(\omega t - kr)} \right) \mathbf{e}_r \\ &= -\frac{\hat{Q}(\mathbf{y})}{4\pi} \left( \frac{jk}{r} + \frac{1}{r^2} \right) e^{j(\omega t - kr)} \mathbf{e}_r \\ &= -p(\mathbf{x}, t) \left( jk + \frac{1}{r} \right) \mathbf{e}_r.\end{aligned}\tag{4.26}$$

From eq. (2.9),

$$\rho_0 \frac{\partial \mathbf{v}}{\partial t} + \nabla_x p = j\omega \rho_0 \mathbf{v} + \nabla_x p = 0,\tag{4.27}$$

and so

$$\begin{aligned}\mathbf{v}(\mathbf{x}, t) &= -\frac{1}{j\omega \rho_0} \nabla_x p(\mathbf{x}, t) = \frac{1}{j\omega \rho_0} p(\mathbf{x}, t) \left( jk + \frac{1}{r} \right) \mathbf{e}_r \\ &= \frac{1}{\rho_0 c_0} p(\mathbf{x}, t) \left( 1 - \frac{j}{kr} \right) \mathbf{e}_r = \frac{\hat{Q}(\mathbf{y})}{4\pi r \rho_0 c_0} \left( 1 - \frac{j}{kr} \right) e^{j(\omega t - kr)} \mathbf{e}_r.\end{aligned}\tag{4.28}$$

From eq. (3.15), we can express time-averaged sound intensity of such a spherical sound wave:

$$\begin{aligned}\langle \mathbf{I} \rangle_T &= \frac{1}{4} (\hat{p}^* \hat{\mathbf{v}} + \hat{p} \hat{\mathbf{v}}^*) \\ &= \frac{1}{4} \hat{p}^* \left[ \frac{1}{\rho_0 c_0} \hat{p} \left( 1 - \frac{j}{kr} \right) \mathbf{e}_r \right] + \frac{1}{4} \hat{p} \left[ \frac{1}{\rho_0 c_0} \hat{p} \left( 1 - \frac{j}{kr} \right) \mathbf{e}_r \right]^* \\ &= \frac{1}{4} \hat{p} \hat{p}^* \left[ \frac{1}{\rho_0 c_0} \left( 1 - \frac{j}{kr} \right) \mathbf{e}_r \right] + \frac{1}{4} \hat{p} \hat{p}^* \left[ \frac{1}{\rho_0 c_0} \left( 1 - \frac{j}{kr} \right) \mathbf{e}_r \right]^* \\ &= \frac{1}{4} |\hat{p}|^2 \frac{1}{\rho_0 c_0} \left( 1 - \frac{j}{kr} + 1 + \frac{j}{kr} \right) \mathbf{e}_r = \frac{|\hat{p}|^2}{2\rho_0 c_0} \mathbf{e}_r \\ &= \frac{1}{2\rho_0 c_0} \left| \frac{\hat{Q}(\mathbf{y})}{4\pi r} e^{-jkr} \right|^2 \mathbf{e}_r = \frac{1}{2\rho_0 c_0} \left| \frac{\hat{Q}(\mathbf{y})}{4\pi r} \right|^2 |e^{-jkr}|^2 \mathbf{e}_r = \frac{|\hat{Q}(\mathbf{y})|^2}{32\rho_0 c_0 \pi^2 r^2} \mathbf{e}_r.\end{aligned}\tag{4.29}$$

<sup>42</sup>If the source is directional, the following analysis is still valid in the far field (with additional multiplication with the factor  $D_i$ ). It will be shown that sound wave acts as a plane wave in the far field and thus (locally) depends spatially essentially only on  $r$  and not  $\theta$  and  $\phi$ .

<sup>43</sup>Gradient of  $p$  is in spherical coordinates  $\nabla p = \frac{\partial p}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial p}{\partial \theta} \mathbf{e}_\theta + \frac{1}{r \sin(\theta)} \frac{\partial p}{\partial \phi} \mathbf{e}_\phi$ .

Note that the terms with  $j/(kr)$  cancelled. Similarly, time-averaged energy is according to eq. (3.19)

$$\begin{aligned}\langle E \rangle_T &= \frac{\hat{p}\hat{p}^*}{4\rho_0c_0^2} + \frac{\rho_0\hat{\mathbf{v}} \cdot \hat{\mathbf{v}}^*}{4} = \frac{\hat{p}\hat{p}^*}{4\rho_0c_0^2} + \frac{\rho_0}{4} \frac{1}{\rho_0^2c_0^2} \hat{p}\hat{p}^* \left(1 + \frac{1}{k^2r^2}\right) \mathbf{e}_r \cdot \mathbf{e}_r \\ &= \frac{|\hat{p}|^2}{2\rho_0c_0^2} \left(1 + \frac{1}{2k^2r^2}\right) = \frac{1}{c_0} \langle \mathbf{I} \rangle_T \cdot \mathbf{e}_r \left(1 + \frac{1}{2k^2r^2}\right) \\ &= \frac{|\hat{Q}(\mathbf{y})|^2}{32\rho_0c_0^2\pi^2r^2} \left(1 + \frac{1}{2k^2r^2}\right)\end{aligned}\quad (4.30)$$

and sound pressure level is given in eq. (3.21):

$$\begin{aligned}L &= 10 \log_{10} \left( \frac{|\hat{p}|^2/2}{4 \cdot 10^{-10} \text{Pa}^2} \right) = 10 \log_{10} \left( \frac{\rho_0 c_0 \langle \mathbf{I} \rangle_T \cdot \mathbf{e}_r}{4 \cdot 10^{-10} \text{Pa}^2} \right) \\ &= 10 \log_{10} \left( \frac{|\hat{Q}(\mathbf{y})|^2 / (32\pi^2 r^2)}{4 \cdot 10^{-10} \text{Pa}^2} \right) \approx 10 \log_{10} \left( \frac{|\hat{Q}(\mathbf{y})|^2 / r^2}{1.3 \cdot 10^{-7} \text{Pa}^2} \right).\end{aligned}\quad (4.31)$$

Acoustic power of an omnidirectional point source emitting the spherical wave is from eq. (3.16)

$$\begin{aligned}\langle P_q \rangle_T &= \oint_S \langle \mathbf{I}(\mathbf{x}, t) \rangle_T \cdot \mathbf{n}(\mathbf{x}) d^2\mathbf{x} = 4\pi r^2 \langle \mathbf{I} \rangle_T \cdot \mathbf{e}_r = \frac{2\pi r^2}{\rho_0 c_0} |\hat{p}|^2 \\ &= \frac{2\pi r^2}{\rho_0 c_0} \frac{|\hat{Q}(\mathbf{y})|^2}{16\pi^2 r^2} = \frac{|\hat{Q}(\mathbf{y})|^2}{8\pi\rho_0 c_0},\end{aligned}\quad (4.32)$$

where we applied the spherical control surface over which  $|\hat{p}|^2$  and  $\mathbf{I} \cdot \mathbf{e}_r$  are constant, with the centre at the source location, radius  $r$ , and unit vector  $\mathbf{n}(\mathbf{x}) = \mathbf{e}_r$  normal to it. Sound power level in eq. (3.17) equals then

$$\begin{aligned}L_W &= 10 \log_{10} \frac{\langle P_q \rangle_T}{10^{-12} \text{W}} \\ &= 10 \log_{10} \frac{|\hat{p}|^2/2}{4 \cdot 10^{-10} \text{Pa}^2} + 10 \log_{10} \frac{4 \cdot 10^{-10} \text{Pa}^2}{\rho_0 c_0 10^{-12} \text{W/m}^2} + 10 \log_{10} \left( \frac{4\pi r^2}{1 \text{m}^2} \right),\end{aligned}\quad (4.33)$$

so sound pressure level is approximated as

$L(r) \approx L_W - 10 \log_{10} \left( \frac{4\pi r^2}{1 \text{m}^2} \right).$

(4.34)

We used the first equality in eq. (4.31) and replaced  $\rho_0 c_0 \approx 400 \text{kg}/(\text{m}^2\text{s})$ . The last approximation is commonly adopted in order to relate sound power level of a source and sound pressure level at certain distance from it. Equation (4.34) holds for an omnidirectional source.

In the expressions for velocity, eq. (4.28), and energy, eq. (4.30), we recognize again Helmholtz number  $kr$ , this time with the distance from the source  $r$  as the reference length scale. If its value is much larger than 1, the second term in the brackets can be neglected compared to the first one, and we can write

$$\boxed{\mathbf{v}(\mathbf{x}, t) = \frac{1}{\rho_0 c_0} p(\mathbf{x}, t) \mathbf{e}_r = \frac{\hat{Q}(\mathbf{y})}{4\pi r \rho_0 c_0} e^{j(\omega t - kr)} \mathbf{e}_r}, \quad (4.35)$$

or after dividing with  $e^{j\omega t}$

$$\hat{\mathbf{v}}(\mathbf{x}) = \frac{1}{\rho_0 c_0} \hat{p}(\mathbf{x}) \mathbf{e}_r = \frac{\hat{Q}(\mathbf{y})}{4\pi r \rho_0 c_0} e^{-jkr} \mathbf{e}_r. \quad (4.36)$$

This is acoustic far field<sup>44</sup>, at large enough (compared to the wavelength) distances from the source, when

$$kr \gg 1 \Rightarrow r \gg \lambda/2\pi. \quad (4.37)$$

Acoustic pressure and velocity are in phase as in a plane wave and the energy from eq. (2.32) is simplified to

$$E = \frac{p^2}{2\rho_0 c_0^2} + \frac{|p\mathbf{e}_r|^2}{2\rho_0 c_0^2} = \frac{p^2}{\rho_0 c_0^2}. \quad (4.38)$$

Time-averaged energy and intensity are also simply related:

$$\boxed{\langle E \rangle_T = \frac{|\hat{p}|^2}{2\rho_0 c_0^2} = \frac{1}{c_0} \langle \mathbf{I} \rangle_T \cdot \mathbf{e}_r = \frac{1}{c_0} |\langle \mathbf{I} \rangle_T| = \frac{|\hat{Q}(\mathbf{y})|^2}{32\rho_0 c_0^2 \pi^2 r^2} = \frac{\langle P_q \rangle_T}{4\pi r^2 c_0}}, \quad (4.39)$$

where we used eq. (4.32) for the last equality. Sound pressure level from eq. (4.31) equals

$$L = 10 \log_{10} \left( \frac{\rho_0 c_0 \langle \mathbf{I} \rangle_T \cdot \mathbf{e}_r}{4 \cdot 10^{-10} \text{Pa}^2} \right) = 10 \log_{10} \left( \frac{\rho_0 c_0^2 \langle E \rangle_T}{4 \cdot 10^{-10} \text{Pa}^2} \right) \quad (4.40)$$

and if we again approximate  $\rho_0 c_0 \approx 400 \text{kg}/(\text{m}^2\text{s})$ :

$$\boxed{L = 10 \log_{10} \left( \frac{\langle \mathbf{I} \rangle_T \cdot \mathbf{e}_r}{10^{-12} \text{W}/\text{m}^2} \right) = 10 \log_{10} \left( \frac{c_0 \langle E \rangle_T}{10^{-12} \text{J}/(\text{m}^2\text{s})} \right)}. \quad (4.41)$$

It follows that in the far field time-averaged sound energy (a scalar) is sufficient for calculation of sound pressure level regardless of the source, which is on the other hand closely related to the subjective perception of sound (see eq. (2.29) and the discussion related to it). This is an important conclusion which will be used in statistical theory and

---

<sup>44</sup>We emphasize that it is acoustic far field to distinguish it from geometric far field, which uses size of the source as the reference length scale (see section 9.1). However, due to its significance, acoustic far field is often called simply far field.

geometrical (ray) acoustics. For example, rather than a wave, the latter theory treats sound as a beam of energy radiated from a source in the direction  $\mathbf{e}_r$ . Time averaging allows neglecting the phase differences between waves, which substantially simplifies the calculations. As will be shown, only the absolute values of complex amplitudes remain relevant.

In contrast to the far field, acoustic near field of a point source<sup>45</sup> can be defined for distances from the source at which  $kr \ll 1$ . Then from eq. (4.28)  $\mathbf{v}(\mathbf{x}, t) = -jp(\mathbf{x}, t)\mathbf{e}_r/(kr\rho_0c_0)$ . Pressure and velocity are out of phase by the factor  $\pi/2$  and the velocity magnitude is much larger than  $|\hat{p}|/(\rho_0c_0)$ . In addition to this, from eq. (4.30),  $\langle E \rangle_T \gg \langle \mathbf{I} \rangle_T \cdot \mathbf{e}_r/c_0$ , which shows that most of the energy in the acoustic near field does not propagate far from the source, to the far field. This energy will affect the value of sound level in eq. (4.31). However, this case is of less practical importance in room acoustics, since inequality (4.37) will be usually satisfied, except very close to the sources (or other objects which might reflect or scatter the sound and thus behave as secondary sources) of relatively low frequency sound. For example, a distance of  $r = 5\text{m}$ , satisfies far field condition for frequencies above 100Hz, while for frequencies above 500Hz, already 1m can be considered as acoustic far field. Moreover, between  $r \ll 1/k$  and  $r \gg 1/k$  is a transition zone, in which the near-field effects are still rather small, so eq. (4.37) is in practice often replaced with less restrictive  $r > \lambda$  or even  $r > \lambda/2$ .

In the far field of a compact directional source, magnitude  $|\hat{p}(r)|$  in eq. (4.33) should be additionally scaled with angularly dependent  $D_i(\theta, \phi)$ , which gives

$$L(r, \theta, \phi) \approx L_W - 10 \log_{10} \left( \frac{4\pi r^2}{1\text{m}^2} \right) + 10 \log_{10} D_i^2(\theta, \phi). \quad (4.42)$$

Similarly, power of a directional point source can be expressed with acoustic quantities in the far field by introducing factor  $D_i^2$  in the surface integral in eq. (4.32) (see also footnote<sup>42</sup>). We obtain:

$$\begin{aligned} \langle P_q \rangle_T &= \oint_S D_i^2(\theta, \phi) \langle \mathbf{I}(\mathbf{x}, t) \rangle_T \cdot \mathbf{n}(\mathbf{x}) d^2\mathbf{x} = |\langle \mathbf{I}(r, t) \rangle_T| \oint_S D_i^2(\theta, \phi) d^2\mathbf{x} \\ &= c_0 \langle E(r, t) \rangle_T \oint_S D_i^2(\theta, \phi) d^2\mathbf{x} = \frac{|\hat{p}(r)|^2}{2\rho_0c_0} \oint_S D_i^2(\theta, \phi) d^2\mathbf{x} \\ &= \frac{|\hat{Q}(\mathbf{y})|^2}{32\rho_0c_0\pi^2r^2} \oint_S D_i^2(\theta, \phi) d^2\mathbf{x}, \end{aligned} \quad (4.43)$$

where we also used eq. (4.39). For an omnidirectional source,  $D_i = 1$ , and this reduces to eq. (4.32), since

$$\oint_S 1 d^2\mathbf{x} = 4\pi r^2 \quad (4.44)$$

for the spherical control surface.

---

<sup>45</sup>If the source is not compact, the receiver can be in the far field with regard to some parts of the source region, no matter how close to the source it actually is, so that the field there is mixed, both near and far field. This is the reason why some authors introduce the term geometric near field for such locations. The condition for acoustic far field, eq. (4.37), remains unchanged.

## 4.5 Acoustic impedance

In the far field of a source of sound in free space, amplitude of pressure and magnitude of velocity are simply related in eq. (4.36) with the characteristic impedance of air,  $Z_0$ :

$$\frac{\hat{p}}{\hat{v} \cdot \mathbf{e}_r} = \frac{\hat{p}}{|\hat{v}|} = Z_0 = \rho_0 c_0 \approx 412 \text{ kg}/(\text{m}^2\text{s}). \quad (4.45)$$

In general, this ratio can be complex, such as in the near field of a point source (eq. (4.28) for  $kr \ll 1$ ). On an arbitrary surface, impedance at point  $\mathbf{y}$  is defined as

$$Z(\mathbf{y}) = \frac{\hat{p}(\mathbf{y})}{\hat{v}(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y})}, \quad (4.46)$$

with the unit vector  $\mathbf{n}(\mathbf{y})$  normal to the surface at  $\mathbf{y}$  and pointing into the surface<sup>46</sup>. It is infinite for a rigid motionless wall, since  $\hat{v} \cdot \mathbf{n} = 0$  for any  $\hat{p}$  at its surface. Hence,  $Z = \infty$  is an alternative form of the boundary condition for a rigid wall.

The impedance is defined pointwise. This is a property of locally reacting surfaces. The impedance of such surfaces at certain point  $\mathbf{y}$  does not depend on the sound field at any other point and, within linear approximation, this makes the impedance a property of the surface alone, regardless of the sound field. This may not hold for the ratio  $\hat{p}/(\hat{v} \cdot \mathbf{n})$  in general, for example due to the waves propagating through the solid structure of walls. In such cases, we may rather talk simply about the ratio of pressure and velocity component normal to the wall. When we talk about impedance, we assume that the surface is locally reacting.

Alternatively, the effect of a surface in far field can also be expressed as a relation between the amplitudes of incident sound pressure  $\hat{p}_{inc}$  (or velocity) and reflected  $\hat{p}_{refl}$  or total  $\hat{p}_{inc} + \hat{p}_{refl}$  sound pressure (or velocity). If the incoming wave reaches the surface from the direction given with polar and azimuthal angles  $(\iota_{inc}, \chi_{inc})$ , the relation can be written as

$$\hat{p}_{refl}(\theta, \phi) = f(\hat{p}_{inc}(\iota_{inc}, \chi_{inc})), \quad (4.47)$$

where  $0 \leq \iota_{inc}, \theta \leq \pi/2$ , and  $0 \leq \chi_{inc}, \phi < 2\pi$ . For example, reflection coefficient is defined at a surface as  $\hat{R}_s = \hat{p}_{refl}(\mathbf{y})/\hat{p}_{inc}(\mathbf{y})$  and will be used in section 8.

From the comparison of equations (4.45) and (4.46), it follows that if a surface in acoustic far field is normal to the direction of propagation of an incoming sound wave, so that  $\mathbf{n}(\mathbf{y}) = \mathbf{e}_r$ , and its impedance is equal to  $Z_0$ , it does not affect the amplitude or phase of pressure or velocity. It is acoustically invisible in the sense that its impedance is matched to the impedance of air and there is no reflection of the incoming sound. This

---

<sup>46</sup>This may sometimes cause a confusion, since we often set  $\mathbf{n}$  so that it points into the room, for example for defining boundary conditions (see also the derivation of eq. (4.49) below). However, we stick to the usual convention for the definition of impedance. This also agrees with the direction of an incoming sound wave (the vector  $\mathbf{e}_r$ ) for normal incidence.

is one of the reasons why impedance is often normalized with the value  $Z_0$ :

$$\mathcal{Z}(\mathbf{y}) = \frac{Z(\mathbf{y})}{Z_0} = \frac{Z(\mathbf{y})}{\rho_0 c_0}, \quad (4.48)$$

which is a dimensionless quantity called specific impedance.

Impedance presents a mixed type of boundary condition and we can relate it to the general boundary condition given in eq. (3.26). First, we notice that from the complex form of conservation of momentum, eq. (4.27), multiplied with  $\cdot \mathbf{n}(\mathbf{y})$  and divided with  $e^{j\omega\tau}$ , we have:  $j\omega\rho_0(\hat{\mathbf{v}} \cdot \mathbf{n}) = -\nabla_y \hat{p} \cdot \mathbf{n}$ . Then, from eq. (4.46):

$$Z(\mathbf{y}) \frac{\nabla_y \hat{p}(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y})}{j\omega\rho_0} - \hat{p}(\mathbf{y}) = 0. \quad (4.49)$$

Since we want to relate the impedance and boundary condition, which by definition takes  $\mathbf{n}$  pointing in the opposite direction (into the room), we also replaced  $\mathbf{n}$  with  $-\mathbf{n}$  in eq. (4.46). Multiplying both sides with  $j\omega/c_0$  gives:

$$\frac{Z(\mathbf{y})}{\rho_0 c_0} \nabla_y \hat{p}(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y}) - \frac{j\omega}{c_0} \hat{p}(\mathbf{y}) = -jk\hat{p}(\mathbf{y}) + \mathcal{Z}(\mathbf{y}) \nabla_y \hat{p}(\mathbf{y}) \cdot \mathbf{n}(\mathbf{y}) = 0. \quad (4.50)$$

Comparing this with eq. (3.26) gives  $a = -jk$  and  $b = \mathcal{Z} = Z/Z_0$  for a passive locally reacting surface. Impedance is just an alternative form of a boundary condition for a locally reacting surface.

Impedance of a surface can also be related to the flow of acoustic energy through it. According to equations (3.15) and (3.16) (with  $\mathbf{n}$  pointing outwards), the associated power is expressed as the integral over the open surface:

$$\begin{aligned} \langle P \rangle_T &= \frac{1}{2} \int_S \mathcal{R}_e(\hat{p}^* \hat{\mathbf{v}} \cdot \mathbf{n}) d^2 \mathbf{y} = \frac{1}{2} \int_S \mathcal{R}_e(Z^* (\hat{\mathbf{v}} \cdot \mathbf{n})^* (\hat{\mathbf{v}} \cdot \mathbf{n})) d^2 \mathbf{y} \\ &= \frac{1}{2} \int_S \mathcal{R}_e(Z) |\hat{\mathbf{v}} \cdot \mathbf{n}|^2 d^2 \mathbf{y}. \end{aligned} \quad (4.51)$$

Therefore, only the real part of impedance (resistance) is related to the flow of energy. When it is larger than zero, the surface is absorbing sound energy, when it is negative the surface is active (it produces sound energy).

## 4.6 Sound rays

The concept of sound ray is central in the geometrical theory of acoustics. It is an energy approximation of a plane sound wave in the far field of a (directional) source radiated within a small solid angle  $d\Omega = \sin(\theta)d\theta d\phi$ . Surface element of the control sphere centred at the source location with radius  $r \gg 1/k$  has area  $dS = r^2 d\Omega$ . Since  $r$  is large, the surface element is approximately flat and the sound field over it is nearly uniform, which justifies the plane wave approximation.

Like the energy of a plane wave, the energy of a sound ray does not decay with  $r$ , because the solid angle  $d\Omega$  is fixed and not  $dS$ . Integral of a generic function  $f(\mathbf{x})$  over the entire spherical control surface equals

$$\oint_S f(\mathbf{x}) d^2\mathbf{x} = r^2 \int_0^{4\pi} f(r, \Omega) d\Omega = r^2 \int_0^{2\pi} \int_0^\pi f(r, \theta, \phi) \sin(\theta) d\theta d\phi. \quad (4.52)$$

Using this in eq. (4.43),

$$\langle P_q \rangle_T = \oint_S c_0 \langle E(r, t) \rangle_T D_i^2(\theta, \phi) d^2\mathbf{x} = \int d\langle P_q \rangle_T, \quad (4.53)$$

implies that the time-averaged sound energy which is radiated within the angle  $d\Omega$  can be related to the acoustic power of the source as

$$d\langle P_q \rangle_T = c_0 D_i^2(\theta, \phi) \langle E(r, t) \rangle_T r^2 d\Omega.$$

Therefore, the rate of energy (power) which is transferred by a sound ray with angle  $d\Omega$  equals

$$\langle P_{q,ray}(\theta, \phi, t) \rangle_T = \frac{d\langle P_q \rangle_T}{d\Omega} = c_0 D_i^2(\theta, \phi) \langle E(r, t) \rangle_T r^2 = D_i^2(\theta, \phi) \frac{\langle P_q \rangle_T}{4\pi}, \quad (4.54)$$

which does not depend on  $r$ . For the last equality, we used eq. (4.39), which also implies

$$|\langle \mathbf{I}_{ray}(\theta, \phi, t) \rangle_T| = \langle E_{ray}(\theta, \phi, t) \rangle_T c_0 = \frac{\langle P_{q,ray}(\theta, \phi, t) \rangle_T}{1\text{m}^2} = D_i^2(\theta, \phi) \frac{\langle P_q \rangle_T}{4\pi \cdot 1\text{m}^2}, \quad (4.55)$$

since there is no energy decay due to the expansion of the wavefront ( $d\Omega$  is fixed).

Energy losses due to dissipation in air or at absorbing surfaces can be included by additional multiplication of the second-order quantities with real positive factors taking values between 0 and 1. This is the case with absorption coefficient which will be introduced later, in eq. (5.22).

## 5 Geometrical and statistical theory

After considering sound fields in time domain in the previous section, we will introduce here statistical and geometrical theory, which are together with modal analysis basic theories of room acoustics. Both of them rely on the statistical approach and averaging of acoustic quantities, which puts acoustic energy as the main descriptor of the fields. Therefore, we will first consider energy summation as the basis of the theories.

### 5.1 Summation of energy

In the previous section we have considered sound field around a point source in free space and how the acoustic properties of locally reacting surfaces can be represented. We also saw that the value of sound pressure level in acoustic far field, which is closely related to the subjective perception of loudness, is fully determined by the value of sound energy (eq. (4.40)). Now we will continue with studying the field at the location of a listener,  $\mathbf{x}$ . In section 3, the problem was treated in frequency domain with modal analysis, according to which the field consists of superimposed modes of the room. In time domain, the sound field at  $\mathbf{x}$  is expected to be a superposition of travelling waves, in particular, direct sound from the source (which is the free space component) and reflections from the surfaces and objects in the room. Therefore, our intention is to calculate total energy of the superposed waves in the far field.

We will start with simple real-valued sine waves with frequency  $f = \omega/(2\pi)$ . Two plane waves in the far field are most generally represented with sound pressures  $p_1 = A \cos(\omega t)$  and  $p_2 = B \cos(\omega(t + \Delta t))$ , where  $A$  and  $B$  are real amplitudes larger than one and  $\omega\Delta t$  is the phase shift between the two waves at certain location of the receiver. If we suppose that the source provides sound starting from  $t = 0$ , this holds for any  $t > 0$ . Without any loss of generality, we can also suppose that  $\Delta t > 0$ . When these two waves are superposed, the total energy is from eq. (4.38)

$$\begin{aligned} E_{sum} &= \frac{(p_1 + p_2)^2}{\rho_0 c_0^2} = \frac{1}{\rho_0 c_0^2} [A \cos(\omega t) + B \cos(\omega(t + \Delta t))]^2 \\ &= \frac{1}{\rho_0 c_0^2} \{(A \cos(\omega t))^2 + [B \cos(\omega(t + \Delta t))]^2 + 2AB \cos(\omega t) \cos(\omega(t + \Delta t))\} \\ &= E_1 + E_2 + \frac{AB}{\rho_0 c_0^2} \{\cos(-\omega\Delta t) + \cos(2\omega t + \omega\Delta t)\} \\ &= E_1 + E_2 + \frac{AB}{\rho_0 c_0^2} \{\cos(\omega\Delta t) + \cos(2\omega(t + \Delta t))\}, \end{aligned}$$

where  $E_1 = p_1^2/(\rho_0 c_0^2)$  and  $E_2 = p_2^2/(\rho_0 c_0^2)$  are energies of the two waves. The energy  $E_{sum}$  does not have to correspond to any particular sound wave. It is merely the total sound energy at the given receiver location.

Time-averaged energy is accordingly

$$\langle E_{sum} \rangle_T = \langle E_1 \rangle_T + \langle E_2 \rangle_T + \frac{AB}{\rho_0 c_0^2} \langle \cos(\omega \Delta t) + \cos(2\omega(t + \Delta t)) \rangle_T \quad (5.1)$$

and depends not only on the energy of the incoming sound waves, but also on the time difference of their arrival at the receiver location. Depending on the phase shift between them, sine waves can add constructively or (partly) cancel. A single fixed value of the delay gives different phase shifts,  $\Delta\Phi = \omega\Delta t$ , depending on the frequency. The value of the last term in eq. (5.1) without the amplitude-dependent factor  $AB/(\rho_0 c_0^2)$  is shown in the left graph in Fig. 9.

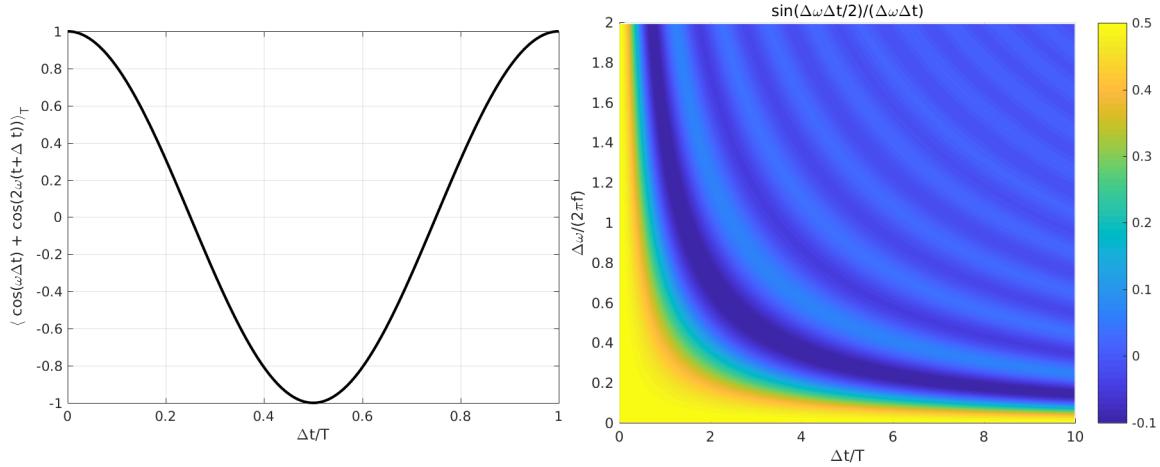


Figure 9: Correction of the sum of energy averaged over time (left) and frequency (right) for sine waves. Time delay is normalized with the period  $T = 1/f$  and  $f$  is central frequency of the range  $\Delta\omega$ .

Acoustic quantities are most commonly measured or calculated within finite frequency bands, for example, octaves or third-octaves. Human auditory system also integrates sound energy over certain frequency ranges in addition to the integration over short time intervals. This justifies averaging over frequency in addition to already introduced time averaging of sound energy. Of course, this makes sense only if the analysed sound is broadband enough within the selected frequency ranges, which means that it is a superposition of many simple sinusoidal components that we consider here.

Frequency averaging of the total energy in the range of angular frequencies  $\Delta\omega$  gives

(compare with time averaging in eq. (2.28)):

$$\begin{aligned}
\langle E_{sum} \rangle_\omega &= \langle E_1 \rangle_\omega + \langle E_2 \rangle_\omega + \frac{AB}{\rho_0 c_0^2} \langle \cos(\omega \Delta t) + \cos(2\omega(t + \Delta t)) \rangle_\omega \\
&= \langle E_1 \rangle_\omega + \langle E_2 \rangle_\omega + \frac{AB}{\rho_0 c_0^2} \frac{1}{\Delta \omega} \left[ \int_{-\Delta \omega/2}^{\Delta \omega/2} \cos(\omega \Delta t) d\omega + \int_{-\Delta \omega/2}^{\Delta \omega/2} \cos(2\omega(t + \Delta t)) d\omega \right] \\
&= \langle E_1 \rangle_\omega + \langle E_2 \rangle_\omega + \frac{AB}{\rho_0 c_0^2} \frac{1}{\Delta \omega} \frac{1}{\Delta t} \int_{-\Delta \omega \Delta t/2}^{\Delta \omega \Delta t/2} \cos(\omega \Delta t) d(\omega \Delta t) \\
&\quad + \frac{AB}{\rho_0 c_0^2} \frac{1}{\Delta \omega} \frac{1}{2(t + \Delta t)} \int_{-\Delta \omega(t + \Delta t)}^{\Delta \omega(t + \Delta t)} \cos(2\omega(t + \Delta t)) d(2\omega(t + \Delta t)) \\
&= \langle E_1 \rangle_\omega + \langle E_2 \rangle_\omega + \frac{AB}{\rho_0 c_0^2} \frac{1}{\Delta \omega} \frac{1}{\Delta t} [\sin(\Delta \omega \Delta t/2) + \sin(\Delta \omega \Delta t/2)] \\
&\quad + \frac{AB}{\rho_0 c_0^2} \frac{1}{\Delta \omega} \frac{1}{2(t + \Delta t)} [\sin(\Delta \omega(t + \Delta t)) + \sin(\Delta \omega(t + \Delta t))] \\
&= \langle E_1 \rangle_\omega + \langle E_2 \rangle_\omega + \frac{2AB}{\rho_0 c_0^2} \frac{1}{\Delta \omega} \frac{1}{\Delta t} \sin(\Delta \omega \Delta t/2) + \frac{AB}{\rho_0 c_0^2} \frac{1}{\Delta \omega} \frac{1}{(t + \Delta t)} \sin(\Delta \omega(t + \Delta t)).
\end{aligned}$$

The last two terms vanish in favour of the sum  $\langle E_1 \rangle_\omega + \langle E_2 \rangle_\omega$  when  $\Delta \omega \Delta t \gg 1$  and  $\Delta \omega(t + \Delta t) \gg 1$ , which is when  $\Delta \omega \gg 1/\Delta t$  (this automatically satisfies both conditions, since  $t, \Delta t > 0$  and thus  $t + \Delta t > \Delta t$ ). This is also demonstrated in Fig. 9 (right), which shows the value of  $\sin(\Delta \omega \Delta t/2)/(\Delta \omega \Delta t)$  as a function of  $\Delta \omega$  and  $\Delta t$  normalized with the central angular frequency of the range and associated period, respectively. As  $\Delta \omega \Delta t \rightarrow \infty$ , the value approaches zero. On the other hand, the factor  $2AB/(\rho_0 c_0^2)$  is of the same order of magnitude as  $\langle E_1 \rangle_\omega + \langle E_2 \rangle_\omega \sim (A^2 + B^2)/(\rho_0 c_0^2)$  or lower.

It follows that energy summation of sound waves (or sound rays as their representations, with which we will work later) is justified only if the condition  $\Delta \omega \Delta t \gg 1$  is satisfied. For example, if a delay between two reflections, replicas of the direct sound, is only 1ms, the condition becomes  $\Delta f \gg 1/(2\pi \Delta t) \approx 160\text{Hz}$ , so the frequency range should be at least around 1kHz wide. However, if the delay is 10ms, width of the range should be  $\Delta f \gg 16\text{Hz}$ . As a conclusion, frequency averaging may not be satisfactory for narrowband sounds (small  $\Delta \omega$ ) or broadband sounds which are coherent<sup>47</sup> and with similar time of arrival (small  $\Delta t$ ). In all other cases, the approximation

$$\boxed{\langle E_{sum} \rangle_\omega = \langle E_1 \rangle_\omega + \langle E_2 \rangle_\omega} \quad (5.2)$$

holds. In practice, short delays are quite common, so the condition is usually that the waves are broadband enough and incoherent. Fortunately, even the reflections which

<sup>47</sup>By coherent, we mean that the spectral content of the waves, including the phase differences between different spectral components of each wave, is very similar. This similarity of the spectra also implies the similarity of waveforms. In the analysis above, we obviously did not cover such dependencies, since we considered only sine waves with the same frequency. However, if broadband sounds are incoherent, random variations of phase differences between their corresponding spectral components at different frequencies will again justify frequency averaging, even if  $\Delta t$  is small.

originate from the same source can usually be treated as incoherent after a few reflections from real surfaces.

Now we can again allow complex values of sound pressure. As before, averaging over time will then remove the phase dependence and leave an expression involving only the amplitudes:

$$\begin{aligned}\langle E_{sum} \rangle_{\omega,T} &= \langle E_1 \rangle_{\omega,T} + \langle E_2 \rangle_{\omega,T} = \frac{\langle |\hat{p}_1|^2 + |\hat{p}_2|^2 \rangle_{\omega}}{2\rho_0 c_0^2} \\ &= \frac{\langle |\hat{Q}(\mathbf{y})|^2 \rangle_{\omega}}{32\rho_0 c_0^2 \pi^2} \left( \frac{1}{r_1^2} + \frac{1}{r_2^2} \right) = \frac{\langle P_q \rangle_{T,\omega}}{4\pi c_0} \left( \frac{1}{r_1^2} + \frac{1}{r_2^2} \right),\end{aligned}\quad (5.3)$$

where we used eq. (4.39) and the last two equalities hold when both waves originate from the same point source (for example, two reflections with different paths to the receiver with total lengths  $r_1$  and  $r_2$ ).

Under some circumstances, energy summation may not reflect well the subjective perception, especially in the case of narrowband and coherent sounds. For example, two simultaneous tones with slightly different frequencies may cause beating, which is time modulation of the sound. Oscillating energy of their sum is not captured by a simple sum of the two energies. Strong early reflections (with short delays with regard to the direct sound) can cause coloration of the perceived sound (recall section 1.3). Such information is also suppressed by the energy summation. Too long intervals for averaging over time (longer than several tenths of milliseconds) also disagree with the natural inertia of human hearing. Hence, energy summation is not recommendable in frequency ranges smaller than, say,  $\Delta f = 5/(2\pi \cdot 20\text{ms}) \approx 40\text{Hz}$ . In general, frequency averaging does involve a loss of information, which might be relevant in certain situations, both objectively and subjectively. However, in spite of these deficiencies and mainly due to the substantial simplification of the calculations which it introduces, frequency averaging has a crucial role in acoustic analyses at high frequencies, when the modal analysis becomes tedious. It gives rise to the statistical theory, as well as ray tracing modelling, which, in contrast to the modal theory, consider only (relatively broadband) sound energy.

## 5.2 Statistical theory and diffuse field

Application of the statistical theory based on energy summation is especially simple if the sound field is ideally diffuse, meaning that the sound waves reach any point in the room uniformly (with equal energy) from all possible directions. In terms of geometrical (ray) acoustics, any receiver location is hit by sound rays, the representations of plane waves, from all directions and the received sound energy does not depend on the angle of incidence. Under the assumption of such a sound field in the room, we can treat energy of the rays as independent of angle. According to eq. (4.55), it equals

$$\langle E_{ray,diff}(t) \rangle_T = \frac{1}{c_0} |\langle \mathbf{I}_{ray,diff}(t) \rangle_T|. \quad (5.4)$$

Next we can relate magnitude of sound intensity  $\mathbf{I}$  or energy  $E$  at  $\mathbf{x}$  with intensity of the incident rays  $\mathbf{I}_{ray}$ . Integration over all directions of the incoming rays and eq. (4.39) give

$$|\langle \mathbf{I} \rangle_T| = c_0 \langle E \rangle_T = \int_0^{4\pi} \langle \mathbf{I}_{ray}(\Omega) \rangle_T \cdot \mathbf{e}_r d\Omega = \int_0^{4\pi} |\langle \mathbf{I}_{ray}(\Omega) \rangle_T| d\Omega, \quad (5.5)$$

where  $\Omega$  is again solid angle and we left the time dependence in  $\langle \cdot \rangle_T$  implicit. Similarly as  $E_{sum}$  above,  $\mathbf{I}$  and  $E$  do not describe any particular wave. In fact, the resulting intensity vector which would be the sum of all intensity vectors of the incoming rays would be zero, according to the diffuse field assumption. The quantity  $|\mathbf{I}|$  here is simply magnitude of the intensity associated with the total energy  $E$  at  $\mathbf{x}$  and eq. (5.5) can be seen as its definition. The integration of time-averaged intensities implies simple summation of energy of the rays as in eq. (5.3) and under the same assumptions – broadband and incoherent sound waves and averaging over frequency. In a diffuse field, the last equation simplifies to

$$c_0 \langle E_{diff} \rangle_T = |\langle \mathbf{I}_{ray,diff} \rangle_T| \int_0^{4\pi} d\Omega = 4\pi |\langle \mathbf{I}_{ray,diff} \rangle_T|. \quad (5.6)$$

and therefore:

$$\langle E_{diff} \rangle_T = \frac{4\pi |\langle \mathbf{I}_{ray,diff} \rangle_T|}{c_0} = 4\pi \langle E_{ray,diff} \rangle_T. \quad (5.7)$$

It can also be shown that a diffuse field is uniform, that is,  $\langle E_{diff} \rangle_T$  is constant in space (does not depend on the location  $\mathbf{x}$  in the room). An intuitive explanation of this would be that if energy in some region of the sound field were lower than in another region, this would mean that the sound rays coming from the first region would transport less energy and the field would not be diffuse (uniform over all angles of incidence) at the locations where the rays from both regions intersect. The energy of rays,  $\langle E_{ray,diff} \rangle_T$  is also constant in space.

Even when uniform in space, sound energy in the room can vary over time. If we neglect dissipation in air, the main loss of sound energy occurs at the surfaces, such as boundaries of the room. In order to include them in the analysis, we first consider a small surface element  $dS_s$  at  $\mathbf{x}$ , which is hit by a sound ray at the angle of incidence  $\theta$ , which is angle to the normal to the surface. Relation between intensity of the ray,  $\mathbf{I}_{ray}$ , and the power received by the surface element,  $dP_{s,ray}$ , is

$$d\langle P_{s,ray} \rangle_T = \langle \mathbf{I}_{ray}(\Omega) \rangle_T \cdot (-\mathbf{n}(\mathbf{x})) dS_s = |\langle \mathbf{I}_{ray}(\Omega) \rangle_T| \cos(\theta) dS_s. \quad (5.8)$$

The change of sign of  $\mathbf{n}$  (the unit vector normal to the surface and pointing into the room) compared to, for example, eq. (4.32) is because we are interested in received and not emitted power. Total power received from all the rays which hit the element  $dS$  is an integral over all angles  $\Omega$  in the half space in front of the surface:

$$d\langle P_s \rangle_T = \int_0^{2\pi} d\langle P_{s,ray} \rangle_T d\Omega = dS_s \int_0^{2\pi} |\langle \mathbf{I}_{ray}(\Omega) \rangle_T| \cos(\theta) d\Omega, \quad (5.9)$$

The last integral is called irradiation strength:

$$\langle I_s \rangle_T = \frac{d\langle P_s \rangle_T}{dS_s} = \int_0^{2\pi} |\langle \mathbf{I}_{ray}(\Omega) \rangle_T| \cos(\theta) d\Omega. \quad (5.10)$$

It is a scalar quantity with the same physical unit as intensity, W/m<sup>2</sup>, which quantifies total power incident to the surface per unit area.

In a diffuse field, the rays arriving from all angles have equal energy, so eq. (5.9) gives

$$\begin{aligned} d\langle P_{s,diff} \rangle_T &= |\langle \mathbf{I}_{ray,diff} \rangle_T| dS_s \int_0^{2\pi} \cos(\theta) d\Omega \\ &= |\langle \mathbf{I}_{ray,diff} \rangle_T| dS_s \int_0^{\pi/2} \int_0^{2\pi} \cos(\theta) \sin(\theta) d\phi d\theta \\ &= 2\pi |\langle \mathbf{I}_{ray,diff} \rangle_T| dS_s \int_0^{\pi/2} \sin(\theta) \cos(\theta) d\theta \\ &= 2\pi |\langle \mathbf{I}_{ray,diff} \rangle_T| dS_s \left[ -\frac{1}{2} (\cos^2(\pi/2) - \cos^2(0)) \right] \\ &= \pi |\langle \mathbf{I}_{ray,diff} \rangle_T| dS_s \end{aligned} \quad (5.11)$$

and the irradiation strength is

$$\boxed{\langle I_{s,diff} \rangle_T = \pi |\langle \mathbf{I}_{ray,diff} \rangle_T| = \frac{\langle E_{diff} \rangle_T c_0}{4}}, \quad (5.12)$$

where the last equality follows from eq. (5.7). Like the energy  $\langle E_{diff} \rangle_T$  in the bulk of the room, irradiation strength  $\langle I_{s,diff} \rangle_T$  is also uniform over all surfaces.

In order to take into account energy losses at the surface, we can multiply eq. (5.8) with a coefficient  $\alpha_{s,\Omega}(\Omega)$ , which takes values between 0 (no energy losses, for example at fully reflecting walls) and 1 (total energy losses) and which in general depends on both angles  $\theta$  and  $\phi$  in the spherical coordinate system with the origin at the surface. The quantity

$$d\langle P_{s,ray,loss}(\Omega) \rangle_T = \alpha_{s,\Omega}(\Omega) d\langle P_{s,ray}(\Omega) \rangle_T. \quad (5.13)$$

represents fraction of the power of sound ray with the direction of arrival  $(\theta, \phi)$  which is absorbed by the surface element (or transmitted through it outside the room). Total energy losses at the surface element from all the incident rays are then

$$d\langle P_{s,loss} \rangle_T = \int_0^{2\pi} \alpha_{s,\Omega}(\Omega) d\langle P_{s,ray}(\Omega) \rangle_T d\Omega. \quad (5.14)$$

In a diffuse field, eq. (5.8) reads

$$d\langle P_{s,ray,diff}(\theta) \rangle_T = |\langle \mathbf{I}_{ray,diff} \rangle_T| \cos(\theta) dS_s \quad (5.15)$$

The power  $dP_{s,ray,diff}$  does not depend on  $\phi$  any more, but still depends on  $\theta$  due to the factor  $\cos(\theta)$ . After inserting this in eq. (5.14), we obtain

$$\begin{aligned} d\langle P_{s,loss,diff} \rangle_T &= |\langle I_{ray,diff} \rangle_T| dS_s \int_0^{2\pi} \alpha_{s,\Omega}(\Omega) \cos(\theta) d\Omega \\ &= \frac{d\langle P_{s,diff} \rangle_T}{\pi} \int_0^{2\pi} \alpha_{s,\Omega}(\Omega) \cos(\theta) d\Omega. \end{aligned} \quad (5.16)$$

We also used the result in eq. (5.11) for the last equality. The product  $\alpha_{s,\Omega}(\Omega) dS_s$  is called equivalent absorption area:

$$dA_s(\Omega) = \alpha_{s,\Omega}(\Omega) dS_s \quad (5.17)$$

and has the unit  $\text{m}^2$ . It is equal to the area of a fully absorbing surface ( $\alpha_{s,\Omega} = 1$ ) which would give equal total absorption as the surface under consideration. The last integral in eq. (5.16) together with the factor  $1/\pi$  represents absorption coefficient in a diffuse field<sup>48</sup>:

$$\alpha_{s,diff} = \frac{1}{\pi} \int_0^{2\pi} \alpha_{s,\Omega}(\Omega) \cos(\theta) d\Omega = 2 \int_0^{\pi/2} \alpha_{s,\Omega}(\theta) \sin(\theta) \cos(\theta) d\theta, \quad (5.18)$$

where the last equality holds in the most common case when  $\alpha_{s,\Omega}$  does not depend on  $\phi$  but only on the angle of incidence. If  $\alpha_{s,\Omega}(\Omega) = \alpha_s$  does not depend on the angle of incidence either, its value is equal to the absorption coefficient in a diffuse field:

$$\alpha_{s,diff} = 2\alpha_s \int_0^{\pi/2} \sin(\theta) \cos(\theta) d\theta = \alpha_s. \quad (5.19)$$

Averaging over time<sup>49</sup> the equation of conservation of acoustic energy, eq. (2.30), with the losses at surfaces replacing the flux term and added sources, and integration over the entire inner volume of the room  $V$  give:

$$\begin{aligned} &\int_V \frac{\partial \langle E \rangle_T}{\partial t} d^3x + \int_S d\langle P_{s,loss} \rangle_T \\ &= \int_V \frac{\partial \langle E \rangle_T}{\partial t} d^3x + \int_S \int_0^{2\pi} \alpha_{s,\Omega}(\Omega) d\langle P_{s,ray}(\Omega) \rangle_T d\Omega = \int_V d\langle P_q \rangle_T. \end{aligned} \quad (5.20)$$

We replaced  $d\langle P_{s,loss} \rangle_T$  from eq. (5.14). Energy rate in the first term is balanced by the losses at the boundaries with total surface  $S$  and total power of the sources of sound in

<sup>48</sup>The absorption coefficient  $\alpha_{s,diff}$  of absorbing materials is commonly measured in a reverberation chamber, while  $\alpha_{s,\Omega}(\theta = 0)$  is measured in Kundt's tube. However, this does not mean that the absorption coefficient for some arbitrary angle will match any these values.

<sup>49</sup>Characteristic time scale with which the energy quantities in the integrals in this formulation of the conservation law change over time is obviously much larger than the period of averaging  $T$ . Otherwise, the time derivative would couple with the integral which is implicit in averaging  $\langle \cdot \rangle_T$ . This is satisfied in weakly damped rooms, when the condition (3.34) holds.

the room,  $P_q$ . In a diffuse field, from equations (5.16), (5.18), (5.11), and (5.12):

$$\begin{aligned}
 & \int_V \frac{\partial \langle E_{diff} \rangle_T}{\partial t} d^3 \mathbf{x} + \int_S d \langle P_{s,loss,diff} \rangle_T \\
 &= \int_V \frac{\partial \langle E_{diff} \rangle_T}{\partial t} d^3 \mathbf{x} + \int_S d \langle P_{s,diff} \rangle_T \alpha_{s,diff} \\
 &= \int_V \frac{\partial \langle E_{diff} \rangle_T}{\partial t} d^3 \mathbf{x} + \int_S |\langle I_{ray,diff} \rangle_T| \pi \alpha_{s,diff} dS_s \\
 &= \int_V \frac{\partial \langle E_{diff} \rangle_T}{\partial t} d^3 \mathbf{x} + \int_S \alpha_{s,diff} \langle I_{s,diff} \rangle_T dS_s = \int_V d \langle P_q \rangle_T.
 \end{aligned} \tag{5.21}$$

Since  $\langle E_{diff} \rangle_T$  is uniform in  $V$  and  $\langle I_{s,diff} \rangle_T$  is uniform over  $S$ , they can leave the integrals and the equation simplifies to

$$V \frac{d \langle E_{diff} \rangle_T}{dt} + \langle \alpha_{s,diff} \rangle_S S \langle I_{s,diff} \rangle_T = \langle P_q \rangle_T, \tag{5.22}$$

where we introduced average absorption coefficient of the surfaces,  $\langle \alpha_{s,diff} \rangle_S$ .

### 5.2.1 Stationary state

In a stationary state, while a source of sustained sound is switched on, the energy is constant and the equation becomes

$$\langle \alpha_{s,diff} \rangle_S S \langle I_{s,diff} \rangle_T = \langle P_q \rangle_T. \tag{5.23}$$

This and eq. (5.12) give

$$\langle \alpha_{s,diff} \rangle_S S \frac{\langle E_{diff} \rangle_T c_0}{4} = \langle P_q \rangle_T,$$

or

$$\boxed{\langle E_{diff} \rangle_T = \frac{4 \langle P_q \rangle_T}{\langle \alpha_{s,diff} \rangle_S S c_0}}. \tag{5.24}$$

It is an important expression which relates uniform sound energy in far diffuse field and acoustic power of the source of stationary sound. Sound pressure level can be calculated from equations (4.41) and (5.24). It is constant in space and equals

$$L = 10 \log_{10} \left( \frac{c_0 \langle E_{diff} \rangle_T}{10^{-12} \text{J}/(\text{m}^2 \text{s})} \right) \approx 10 \log_{10} \left( \frac{\langle P_q \rangle_T / 10^{-12} \text{W}}{\langle \alpha_{s,diff} \rangle_S S / 1 \text{m}^2} \right) + 6 \text{dB}, \tag{5.25}$$

that is,

$$\boxed{L = L_W - 10 \log_{10} \left( \frac{\langle \alpha_{s,diff} \rangle_S S}{1 \text{m}^2} \right) + 6 \text{dB}}, \tag{5.26}$$

where sound power level  $L_W$  was introduced in eq. (4.34).

In contrast to this, close to an omnidirectional point source, energy of the dominating direct sound will decay with the distance from the source. This suggests that there is a certain distance  $r_{c,diff}$  (called critical distance or reverberation distance) in the room, at which energies of the direct and reflected sound are equal. By using the relation between pressure and energy in the far field and power of the source, eq. (4.39), in combination with eq. (5.24), we can express the critical distance from

$$\langle E_{diff} \rangle_T = \frac{|\hat{p}_{diff}|^2}{2\rho_0 c_0^2} = \frac{4\langle P_q \rangle_T}{\langle \alpha_{s,diff} \rangle_S S c_0} = \frac{4}{\langle \alpha_{s,diff} \rangle_S S c_0} 2\pi r_{c,diff}^2 \frac{|\hat{p}_{diff}|^2}{\rho_0 c_0}. \quad (5.27)$$

It equals

$$r_{c,diff} = \sqrt{\frac{\langle \alpha_{s,diff} \rangle_S S}{16\pi}}. \quad (5.28)$$

Since it is derived under the assumption of far field, the value of critical distance holds for the frequencies which satisfy the condition  $kr_{c,diff} \gg 1$ . Only at the distances from the source which are larger than the critical distance,  $r > r_{c,diff}$ , sound field can be diffuse (a necessary but not sufficient condition). Figure 10 (left) demonstrates the decay of sound energy with the distance in free space according to eq. (4.39) and its constant value in diffuse field, according to equations (5.24) and (5.28). The values in decibels are normalized so that the level in the diffuse field equals 0dB and the two lines actually represent the sound pressure levels from eq. (4.34) and eq. (5.26) (with  $r_{c,diff}$  from eq. (5.28)).

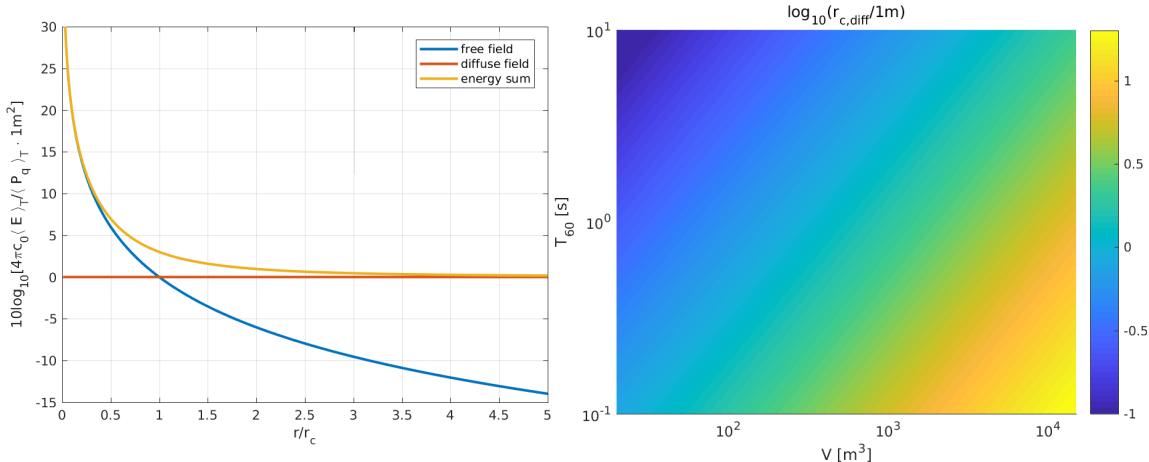


Figure 10: Normalized sound energy in a free field and diffuse field as function of distance from an omnidirectional point source normalized with critical distance (left) and the value of critical distance in a diffuse field from to eq. (5.36) (right).

If the source is directional, critical distance in the direction of its maximum radiation can be estimated with the expression

$$r_{c,diff} = \sqrt{\gamma \frac{\langle \alpha_{s,diff} \rangle_S S}{16\pi}}, \quad (5.29)$$

with  $\gamma$  directivity factor defined in eq. (4.23). However, the field is then less likely to be diffuse.

### 5.2.2 Energy decay

When the source is switched off (say, at time  $t = 0$ ), a transient regime begins during which the sound energy decays until it is completely lost. Following from eq. (5.22), this is expressed by the equality

$$V \frac{d\langle E_{diff} \rangle_T}{dt} + \langle \alpha_{s,diff} \rangle_S S \langle I_{s,diff} \rangle_T = 0, \quad (5.30)$$

which together with eq. (5.12) gives

$$V \frac{d\langle E_{diff} \rangle_T}{dt} + \frac{\langle \alpha_{s,diff} \rangle_S S c_0}{4} \langle E_{diff} \rangle_T = 0. \quad (5.31)$$

This ordinary differential equation has the solution

$$\langle E_{diff} \rangle_T = \langle E_{diff} \rangle_T^0 e^{-\frac{\langle \alpha_{s,diff} \rangle_S S c_0}{4V} t} \quad (5.32)$$

for any  $t \geq 0$ , where  $\langle E_{diff} \rangle_T^0$  is energy at the moment when the source is switched off, that is,  $\langle E_{diff} \rangle_T^0 = \langle E_{diff} \rangle_T$  for  $t = 0$ .

Comparing the last equation with the square of pressure amplitude in eq. (3.33), we find the damping constant

$$\zeta = \frac{\langle \alpha_{s,diff} \rangle_S S c_0}{8V} \quad (5.33)$$

due to the absorption at the surfaces and from eq. (3.35), the reverberation time equals

$$T_{60} \approx \frac{6.91}{\zeta} = 55.28 \frac{V}{\langle \alpha_{s,diff} \rangle_S S c_0} \approx 0.16 \frac{V \cdot 1\text{s}/\text{m}}{\langle \alpha_{s,diff} \rangle_S S} = 0.16 \frac{V \cdot 1\text{s}/\text{m}}{A_{tot,diff}}, \quad (5.34)$$

where  $A_{tot,diff} = \langle \alpha_{s,diff} \rangle_S S$  is total equivalent absorption area in the room, as introduced in eq. (5.17). This is famous Sabine's formula. Note that it is valid in the far field for any shape of the room, as long as the sound field is diffuse. In addition to this, the conditions for energy summation have to be fulfilled, which means that the superimposed sound waves are broadband and incoherent (see the comments with regard to eq. (5.2)). The ratio  $S/V$  for a rectangular room is shown in Fig. 6 (right). We can even include dissipation in air by adding the damping constant  $\zeta_{air}$  from eq. (3.40) to the one in eq. (5.33). This gives

$$T_{60} \approx \frac{6.91}{\zeta + \zeta_{air}} = \frac{6.91}{\frac{\langle \alpha_{s,diff} \rangle_S S c_0}{8V} + \frac{m_{air} c_0}{2}} \approx 0.16 \frac{V \cdot 1\text{s}/\text{m}}{\langle \alpha_{s,diff} \rangle_S S + 4m_{air} V}. \quad (5.35)$$

Neglecting the dissipation in air, we can estimate the critical distance from eq. (5.28):

$$r_{c,diff} = \sqrt{\frac{\langle \alpha_{s,diff} \rangle_S S}{16\pi}} = \sqrt{\frac{0.16V \cdot 1\text{s}/\text{m}}{16\pi T_{60}}} \approx 0.056 \sqrt{\frac{V \cdot 1\text{s}/\text{m}}{T_{60}}}. \quad (5.36)$$

Logarithm of its value is shown in Fig. 10 (right) for different values of  $V$  and  $T_{60}$ . For  $V$  around  $100 \text{ m}^3$  and  $T$  around  $1\text{s}$ , we see that the critical distance is of the order of magnitude  $r_{c,diff} \sim \mathcal{O}(1)\text{m}$ . This indicates that in most of the ordinary rooms listeners are located in the zone where the direct sound is much weaker than the reflected sound. However, this does not hold for microphones (in recording studios or on stages) or listeners at mixing desks in control rooms, which are all usually located close to the sources of sound in order to receive strong direct sound. For a directional source, eq. (5.29) gives in the direction of maximum radiation of the source

$$r_{c,diff} \approx 0.056 \sqrt{\gamma \frac{V \cdot 1\text{s}/\text{m}}{T_{60}}}. \quad (5.37)$$

It should be mentioned that truly diffuse field is very difficult to achieve in practice, even in the far field of omnidirectional sources. This can be due to geometry of the surfaces, which might reflect sound energy in different directions non-uniformly (for example, concave curved surfaces) or an unevenly distributed absorption inside the room, which leads to some surfaces reflecting more than other. Even in reverberation chambers with non-parallel and very low absorbing walls, additional diffusers are commonly used to make the sound field more diffuse. Nevertheless, the simple model of a diffuse field remains indispensable for quick estimations of reverberation time and sound energy in a room in relatively wide frequency bands far enough from the source ( $r > r_{c,diff}$  and  $kr \gg 1$ ) – equations (5.24) (or eq. (5.26) in dB) and (5.34).

### 5.3 Ray tracing in a rectangular room

Although Sabine's formula is very simple and accurate in many cases, more insight is desirable, especially concerning the assumption of diffuseness of the field, which is probably the biggest restriction of the theory. We shall go back and rather than treating only overall sound energy in the room (as in eq. (5.20) and further), we will analyse sound rays, their propagation through the room, and reflections at the surfaces. This is known as geometrical (or ray) acoustics and we study its application in room acoustics.

Similarly as with modal analysis, in order to come up with an analytical solution, we have to restrict ourselves to certain idealized conditions. We again consider a rectangular room. Furthermore, we assume that all six surfaces of the room are acoustically large. In terms of the Helmholtz number from eq. (3.79), this means that  $kL_i \gg 1$ , for  $i = 1, 2, 3$ , where  $L_1, L_2$  and  $L_3$  are, as before, dimensions of the room. Sound rays are then

supposed to specularly reflect from all surfaces of the room<sup>50</sup>. In a way, the analysis here complements the one in section 3.6 for high frequencies, when the modal analysis becomes too demanding.

If sound dissipation in air is neglected, a sound ray loses energy only when it reflects from absorbing (or transmitting) surfaces, according to eq. (5.13). The remaining power which is reflected from the surface element equals

$$d\langle P_{ray,refl}(\Omega)\rangle_T = (1 - \alpha_{s,\Omega}(\Omega))d\langle P_{s,ray}(\Omega)\rangle_T. \quad (5.38)$$

Next we will again assume a diffuse field and replace  $\alpha_{s,\Omega}(\Omega)$  with the average value  $\langle\alpha_{s,diff}\rangle_S$  for each surface and each sound ray in the room, regardless of its angle of incidence<sup>51</sup>. Therefore, the average value of absorption coefficient,  $\langle\alpha_{s,diff}\rangle_S$ , applies here not only to the room as a whole, but to each separate surface as well, which is a bit more stringent requirement than in the previous subsection and Sabine's equation. In such a case, power of the reflected ray becomes

$$d\langle P_{ray,diff,refl}(\Omega)\rangle_T = (1 - \langle\alpha_{s,diff}\rangle_S)d\langle P_{s,ray,diff}(\Omega)\rangle_T. \quad (5.39)$$

After  $N_{ray}(\Omega)$  reflections of the ray given by the angle  $\Omega$ , the remaining power of the ray is

$$d\langle P_{ray,diff,Nrefl}(\Omega)\rangle_T = (1 - \langle\alpha_{s,diff}\rangle_S)^{N_{ray}(\Omega)}d\langle P_{s,ray,diff}(\Omega)\rangle_T. \quad (5.40)$$

For a rectangular room, we can express  $N_{ray}(\Omega)$  for all rays explicitly. First we find that

$$N_{ray,i}(\Omega) = N_{ray,i}(\theta_i) = \frac{tc_0 \cos(\theta_i)}{L_i} \quad (5.41)$$

is the number of reflections in time interval  $(0, t)$  from the two surfaces which are normal to the axis  $x_i$  and  $\theta_i$  is the angle of incidence to those surfaces. Since the surfaces are parallel to each other and all reflections are specular,  $\theta_i$  remains constant for a single ray. Furthermore, the ray travels the distance  $L_i/\cos(\theta)$  between two successive reflections from the two surfaces, which leads to eq. (5.41). In a diffuse field, we can estimate the average number of reflections of all rays which uniformly arrive from all incident angles  $\theta_i$  to the two parallel surfaces:

$$\begin{aligned} \langle N_i(\Omega) \rangle_\Omega &= \frac{1}{2\pi} \int_0^{2\pi} N_{ray,i}(\Omega) d\Omega = \frac{1}{2\pi} \int_0^{2\pi} \frac{tc_0 \cos(\theta_i)}{L_i} d\Omega \\ &= \frac{tc_0}{2\pi L_i} 2\pi \int_0^{\pi/2} \cos(\theta_i) \sin(\theta_i) d\theta_i = \frac{tc_0}{2L_i}, \end{aligned} \quad (5.42)$$

<sup>50</sup>This does not apply to the areas close to the edges or corners of the room. However, we neglect these in favour of much larger regions in which specular reflections do take place.

<sup>51</sup>In reality, certain degree of diffuseness and stochasticity of the angles of incidence would be achieved even for a single ray, due to the scattering from non-flat surfaces. However, we cannot use this argument for the rectangular room discussed here, so we assume that the diffuseness is achieved by many rays, which could originate from an omnidirectional source and hit the room surfaces at different angles of incidence (indicated here with  $\Omega$  as the argument), and which at the receiver location carry approximately equal energies.

where the factor  $2\pi$  in front of the last integral results from the integral with respect to  $\phi$ . Therefore, the average number of reflections of all rays from all six surfaces equals

$$\langle N(\Omega) \rangle_\Omega = \frac{tc_0}{2} \sum_{i=1}^3 \frac{1}{L_i} = \frac{tc_0}{2} \frac{L_1 L_2 + L_1 L_3 + L_2 L_3}{L_1 L_2 L_3} = \frac{tc_0}{2} \frac{S/2}{V} = \frac{tc_0 S}{4V}. \quad (5.43)$$

As before,  $S$  denotes total area of the surfaces and  $V$  is volume of the room. The ratio  $S/V$  in a rectangular room is shown in Fig. 6 (right). Although we will not prove it, it should be mentioned that this expression for the average number of reflections remains valid even for rooms with arbitrary shapes, if the sound field is diffuse. Hence, the conclusions which follow can be generalized to non-rectangular rooms. Nevertheless, details of the distributions of ray path lengths (or time intervals between successive reflections) around the average values do depend on the shape of the room.

From equations (5.40) and (5.43), it follows that sound energy from all rays in a diffuse field decays proportionally to  $(1 - \langle \alpha_{s,diff} \rangle_S)^{\langle N(\Omega) \rangle_\Omega} = (1 - \langle \alpha_{s,diff} \rangle_S)^{tc_0 S / (4V)}$ . We can also add the dissipation in air by multiplying this factor with  $e^{-2\zeta_{air}t}$ , which follows from equations (3.39) and (3.40). Switching off the source at time  $t = 0$ , reverberation time  $T_{60}$  can then be estimated from the equality

$$\begin{aligned} 10 \log_{10} \left[ (1 - \langle \alpha_{s,diff} \rangle_S)^{\frac{T_{60} c_0 S}{4V}} e^{-2\zeta_{air} T_{60}} \right] \\ = 10 \log_{10} \left[ e^{\frac{T_{60} c_0 S}{4V} \ln(1 - \langle \alpha_{s,diff} \rangle_S) - 2\zeta_{air} T_{60}} \right] = -60 \text{dB}. \end{aligned}$$

It equals

$$T_{60} = \frac{\ln(10^{-6})}{\frac{c_0 S \ln(1 - \langle \alpha_{s,diff} \rangle_S)}{4V} - 2\zeta_{air}} = \frac{-24V \ln(10)}{c_0 (S \ln(1 - \langle \alpha_{s,diff} \rangle_S) - 4Vm_{air})} \quad (5.44)$$

or

$$T_{60} \approx 0.16 \frac{V \cdot 1\text{s}/\text{m}}{4m_{air}V - S \ln(1 - \langle \alpha_{s,diff} \rangle_S)}. \quad (5.45)$$

This is Eyring's reverberation formula. If we expand the natural logarithm into series:

$$\ln(1 - \langle \alpha_{s,diff} \rangle_S) = -\langle \alpha_{s,diff} \rangle_S - \frac{1}{2}\langle \alpha_{s,diff} \rangle_S^2 - \frac{1}{3}\langle \alpha_{s,diff} \rangle_S^3 - \dots,$$

we see that for  $\langle \alpha_{s,diff} \rangle_S \ll 1$ ,  $\ln(1 - \langle \alpha_{s,diff} \rangle_S) \approx -\langle \alpha_{s,diff} \rangle_S$  and Eyring's formula simplifies to eq. (5.35) or, for  $m_{air} = 0$  to Sabine's formula (5.34).

Since it simplifies to Sabine's formula for small  $\langle \alpha_{s,diff} \rangle_S$ , Eyring's formula appears to be more general and more accurate when the room is not very weakly damped, say for  $\langle \alpha_{s,diff} \rangle_S > 0.15$ . However, this is not necessarily true, since the derivations of both equations included several assumptions regarding the sound field and the level to which these assumptions are satisfied in an actual room determines the accuracy of the two equations. We derived Eyring's equation only for a rectangular room (although, as

mentioned above, eq. (5.43) and the results which follow from it are valid for arbitrary shapes), diffuse field (as Sabine's equation), all specular reflections in the room, and equal absorption coefficient  $\alpha_{s,diff} = \langle \alpha_{s,diff} \rangle_S$  of all surfaces. The last two are additional requirements compared with the derivation of Sabine's equation (in which  $\langle \alpha_{s,diff} \rangle_S$  is only averaged over all surfaces), which might limit the accuracy of Eyring's formula in certain cases (for example, many relatively small surfaces with different values of absorption coefficient). In any case, an unevenly distributed absorption of the surfaces can cause a non-diffuse sound field, which is a prerequisite for both theories, besides the energy summation and far-field assumption. The introduced assumptions for both formulas are summarized in Table 5.1.

Table 5.1: Assumptions in the derivations of Sabine's and Eyring's formulas for reverberation time.

| formula                              | assumptions  |
|--------------------------------------|--|
| Sabine, eq. (5.34)<br>and eq. (5.35) | energy summation (broadband and incoherent reflections), far field, diffuse field, average absorption coefficient of the surfaces                      |
| Eyring, eq. (5.45)                   | energy summation (broadband and incoherent reflections), far field, diffuse field, specular reflections, equal absorption coefficients of the surfaces |

The ray tracing analysis in this subsection points to the relation between energy decay (reverberation) and the rate of reflections (number of reflections per second), besides the absorption coefficient of the surfaces. This relation could not be made explicit by the analysis of total sound energy in the room in section 5.2. By definition, rate of reflections is

$$\dot{N} = \frac{N}{t} \quad (5.46)$$

and so, from eq. (5.43), average rate of reflections in a room of arbitrary shape in the case of a diffuse field in it is

$$\dot{N} = \frac{\langle N(\Omega) \rangle_\Omega}{t} = \frac{c_0 S}{4V}. \quad (5.47)$$

Accordingly, the mean free path length of the rays between two successive reflections is

$$l = \frac{c_0}{\dot{N}} = \frac{4V}{S}. \quad (5.48)$$

The last two parameters and their distributions for all the rays can be used to assess diffuseness of the field, when their values are available (typically from ray-tracing simulations). For example, occurrence of local peaks in the distribution of free path lengths at low values (below the mean value in the last equation) may indicate a partly decoupled space in the room, such as niche, in which many local reflections take place.

In general, reverberation time and sound field in a room do depend on the shape of the room and the distribution of absorbing surfaces. Irregular room shapes or placement

of absorbing surfaces may cause unequal mean free path lengths of different rays or uneven distribution of energy between them, resulting in a non-diffuse field. In such cases, certain absorbing surfaces may become less irradiated than other and therefore less efficient (absorbing less energy) than in a diffuse field. In other words, their effective absorption coefficient in the room can be lower than the value in a diffuse field, which is measured in a reverberation chamber and usually used for the calculations. As a consequence, reverberation time is often underestimated by eq. (5.45), even when all surfaces have similar absorption coefficient values.

In real rooms, sound absorption is often non-uniformly distributed. Typical examples are auditorium, which is often the only substantially absorbing surface in the room, or absorbing suspended ceilings or walls in offices or at public places. If all other surfaces are much less absorbing, sound field in these rooms is not diffuse in general. Sound waves (rays) which are reflected from the absorbing surfaces carry much less energy than other reflected waves and the sound energy at certain locations in the room is angularly dependent. This also means that the surfaces are not irradiated equally or uniformly over different angles of incidence, as they should be in a diffuse field. Consequently, the absorbed power (the contributions to the integral in eq. (5.20) involving  $d\langle P_{s,ray}(\Omega) \rangle_T$  of each surface) varies depending on the locations of particular surfaces with respect to the highly absorbing surface(s) and the averaging which was introduced in eq. (5.39) and the substitution of  $\alpha_{s,\Omega}$  with  $\alpha_{s,diff}$  are not justified. The entire sound field in the room is not uniform either and the statistical theory becomes inaccurate.

In particular, reverberation time can be overestimated by Eyring's and Sabine's formulas when large surface of the floor (compared with the surfaces of the walls) is absorbing. The reason is that the reflections which reach the floor from other, more reflecting surfaces get mostly absorbed. This leaves other surfaces less irradiated by the reflections from the floor, which results in a comparatively larger effective absorption of the floor than suggested by the diffuse field theory, according to which all surfaces are equally irradiated. If the reflecting ceiling is very high, reverberation time can also be longer in the upper, more reflecting part of the room, than in the lower, absorbing part. This can be sometimes even perceived in the auditorium as an additional longer reverberation coupled to the shorter local one. The opposite case is when an absorbing surface is located in a part of the room with lower sound energy (for example, at two shorter walls of a long rectangular room). It will absorb less efficiently than in a diffuse field and longer reverberation time should be expected than predicted based on the diffuse field theory. Of course, such deviations are case-dependent. More details on the sound fields can be obtained by means of ray tracing simulations, which also rely on the assumption of energy summation and acoustic far field. Some details on the numerical calculations using ray tracing will be discussed in section 10.1.

## 6 Measurement and descriptors of room acoustics

In section 4 we concluded that for given locations of the source and receiver, impulse response of the room contains all information about its acoustics. Therefore, it is natural to assess room acoustics of a room from measured (or calculated) impulse responses at the locations of interest. Nevertheless, certain properties of the acoustics of the room or sound field in it can be determined without acquisition of impulse responses. For example, spatial distribution of sound level can be measured using a source of stationary noise (usually pink noise), broadband or in octave or third-octave frequency bands. Similarly, reverberation time alone can be measured by switching off the source of noise after the stationary state in the room is achieved (as in section 5.2.1 for the case of a diffuse field; it usually suffices to leave the source on for at least the time interval equal to the expected reverberation time of the room). From the recording of the decaying sound, one can estimate the reverberation time, for example, with the least squares linear fitting of the sound level decay curve.

Such shortened procedures are often sufficient for noise control (see the right-hand side of Fig. 1 and the second column of Table 1.1). Modal analysis from section 3 (analytical or numerical) is appropriate when distinct modes of the room should be investigated, typically in small rooms at relatively low frequencies. If more details on the (broadband) room acoustics are necessary, impulse responses can be obtained quite routinely by using standardized procedures. In this section we explain how impulse responses can be measured and which descriptors of room acoustics can be calculated from them. More details on how they can be predicted will be given in section 10.

### 6.1 Measurement of a room impulse response

Like tailored Green's function  $G_{tail}(\mathbf{x}, t | \mathbf{y}, \tau)$ , room impulse response does not depend only on the properties of the room such as boundary conditions, but also on the particular locations of the source and receiver. On the other hand, it should not depend on any other properties of the source or receiver (section 1.2), such as directivity. For specific applications, these aspects can also be relevant and taken into account. However, impulse responses are by definition independent of the source and, therefore, measured with omnidirectional source and receiver.

The source should be omnidirectional (see Table 6.1) and have sufficient power over the entire frequency range of interest. As we will see, most of the descriptors of room acoustics which are derived from the impulse responses do not depend on the absolute sound levels. Hence, it is not necessary that the source radiates with equal power over different frequency bands, as long as sufficient signal-to-noise ratio (typically more than 30dB) is achieved in them for calculation of the descriptors. The source should be placed at an appropriate location in the room, where the actual sources of sound are expected (see Table 1.2). In the case of large stages, 3 or more source locations can be used. If the actual receiver is distributed over space (an auditorium), microphone locations should

cover appropriately its entire surface, with usually at least one microphone location per 100-200 seats. An advantage can be taken if the room is symmetric and the measurements can be performed only in one part (one half) of the auditorium or for fewer source locations.

Table 6.1: Maximum allowed angular variations of the far-field sound pressure level of a source which is used for measurement of room impulse response.

| octave    | 125Hz | 250Hz | 500Hz | 1kHz | 2kHz | 4kHz |
|-----------|-------|-------|-------|------|------|------|
| deviation | ±1dB  | ±1dB  | ±1dB  | ±3dB | ±5dB | ±6dB |

The source can produce a short and strong sound impulse, thus approximating the ideal delta function in time, for example, a starting pistol, firecracker, balloon, clapper, or electric spark generator. However, their radiation pattern should not vary over different angles more than several decibels (Table 6.1) and they must provide sufficient signal-to-noise ratio in the entire frequency range of interest. While the directivity of starting pistols, popped balloons, and clappers can be an issue (especially with the person conducting the measurements in their vicinity), sources of too short or weak impulses, such as small firecrackers or spark generators, may not provide sufficient sound energy at low frequencies, especially at large distances in large rooms. These properties of the sources should be verified prior to the measurements.

As an alternative to the listed sources of impulse sound, loudspeakers can be used as more controllable sources. In particular, omnidirectional radiation in wider frequency range, which is difficult to achieve with a single loudspeaker, is obtained with multiple loudspeakers commonly arranged in the form of a dodecahedron or sphere. However, insufficient signal-to-noise ratio can still be a problem in large spaces, as well as non-linearity of loudspeaker's response at high levels.

Although possible, radiation of a short impulse from loudspeakers is usually impractical, due to the limited dynamic range. The main advantage of loudspeakers over the sources of impulse sound is that they can emit other well defined deterministic signals, which have certain properties suitable for acquisition of impulse responses. These are maximum length sequence (MLS) and swept-sine (chirp, sweep). These deterministic sequences have the property of a random signal that their autocorrelation approaches ideal delta function for sufficiently long duration. This can be expressed as

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t)s(t+\tau)dt = C\delta(\tau), \quad (6.1)$$

where  $C$  is a real constant related to the mean energy of the signal  $s(t)$  representing the sequence. Of course, duration  $T$  of real signals is finite, which turns the equality in eq. (6.1) into approximation.

From eq. (4.15) it is clear that the impulse response  $g(t)$  can be acquired after deconvolution of signal  $p(t)$ , recorded with an omnidirectional microphone, with the known

excitation signal  $s(t)$  (emission time is implicit). Alternatively, convolution of  $p(t)$  (or sequence  $s(t)$ ) with time-inverted sequence  $s(-t)$  (or  $p(-t)$ ) can be performed. This is then essentially the same as cross-correlation of  $p(t)$  and  $s(t)$  (compare the integrals in eq. (4.13) with  $p$  replacing  $g$  and the integral in eq. (6.1) with one  $s$  replaced with  $p$ ,  $+$  replaced with  $-$ , and  $t$  and  $\tau$  switched). Since the deterministic sequences discussed here have autocorrelation approximating ideal impulse, deconvolution of  $p(t)$  (which is essentially a series of delayed and scaled versions of  $s(t)$ ) with them indeed gives an impulse response. In particular, cross correlation with MLS is efficiently performed using the fast Hadamard transformation. On the other hand, deconvolution with swept-sine is much simpler in frequency domain, as it comes down to the division of two spectra, according to eq. (4.16):

$$\hat{g}(\omega) = \frac{\hat{p}(\omega)}{\hat{s}(\omega)}, \quad (6.2)$$

where  $\hat{p}(\omega)$  is (fast) Fourier transform of  $p(t)$ , as in eq. (3.23), and similarly for  $\hat{g}(\omega)$  and  $\hat{s}(\omega)$ . The impulse response  $g(t)$  is then obtained as the inverse Fourier transform of  $\hat{g}(\omega)$  (eq. (3.22)).

The recorded signal  $p(t)$  must include the complete sequence emitted by the loudspeaker, as well as the entire sound decay process in the room after the emission is over. The extraneous noise outside this time window can be cut out. Signal-to-noise ratio can be further improved either by repetition of the sequence (followed by averaging out the background noise, which is assumed to be random; this results in around 3dB higher signal-to-noise ratio per doubling the number of repetitions), or by increasing its duration (and thus the total sound energy per measurement, which results in around 3dB higher signal-to-noise ratio per doubling the duration). The latter approach is preferred since sequence averaging can be sensitive to jitter. If the sequence is repeated, the duration of each repetition should be longer than the expected reverberation time of the room, in order to prevent significant overlapping of impulse responses acquired by different repetitions.

Swept-sine is generally defined as

$$s(t) = A \sin(2\pi f(t)t), \quad (6.3)$$

where  $A$  is its amplitude and the frequency  $f(t)$  depends on time. For a linear swept-sine, it is a linear function,

$$f(t) = f_1 + (f_2 - f_1) \frac{t}{T}, \quad (6.4)$$

with  $f_1$  the lowest (start) frequency,  $f_2$  the highest (end) frequency, and  $T$ , as above, total duration of the sequence. The energy of linear swept-sine is uniformly distributed over all frequencies between  $f_1$  and  $f_2$  and in this range it has the same amplitude spectrum as white noise. In contrast to this, amplitude spectrum of background noise usually decays with frequency similarly to pink noise, which means that the signal-to-noise ratio is lower at low frequencies. One of the advantages of swept-sine over MLS is that its spectrum can be easily adapted to approach the shape of the spectrum of background noise by

modifying the function  $f(t)$ , which keeps the signal-to-noise ratio relatively constant over the entire frequency range. Of course, similar can be achieved by making the amplitude  $A$  frequency (and thus time) dependent, but this is less favourable since keeping  $A$  constant (and low enough to avoid non-linear distortions) allows using maximum dynamic range of the loudspeaker. Logarithmic swept-sine is most frequently used, since it replicates the amplitude spectrum of pink noise. Its frequency increases faster than linearly with time, which gives more energy at lower than at higher frequencies. More specifically, the frequency increases exponentially from  $f_1$  to  $f_2$  over time  $T$ :

$$f(t) = f_1 \left( \frac{f_2}{f_1} \right)^{t/T}. \quad (6.5)$$

Another advantage of swept-sine is that harmonic components due to non-linear distortion can often be detected and removed. If the frequency of the swept-sine gradually increases, harmonic components will appear in  $p(t)$  before the component with equal frequency appears in the emitted sequence. As a consequence, the obtained impulse response will contain spurious reflections before the direct sound (or at the end of it), which can be easily identified and windowed out. In this way, swept-sine technique is more robust to non-linear distortions than other techniques. Nevertheless, non-linear distortion should be avoided, since it generally increases the noise level in impulse responses.

When appropriate (usually in larger projects), room impulse responses should be acquired even before the room is built, in the design phase. This allows pre-estimation of the room's acoustics, values of the descriptors described below, as well as prevention of large failures of the design. It can be done numerically, using software tools, or in scale models, which will be discussed in section 10. When the acoustics is critical, such as in concert halls and opera houses, impulse response measurements can also be performed at different stages of the construction of the hall, in order to further control the building process and avoid mistakes. As a general rule, correction of the mistakes in the design is much more difficult and costly at later stages. Finally, measurements of impulse responses in the finished room allow the final assessment of the room and comparison with the requirements.

## 6.2 Descriptors of room acoustics

Next we will introduce numerical descriptors which are commonly used for assessment of room acoustics. They can all be calculated from impulse responses of a room, which, as already discussed, contain all the information about room acoustics for the given locations of source and receiver. As such, the descriptors are objective. On the other hand, impulse responses are not readily related to the subjective criteria which were described in section 1.3. Therefore, main purpose of the descriptors is to allow evaluation of different subjectively relevant aspects of the room acoustics based on impulse responses of the room.

An important question is then how much of the subjectively relevant information which is contained in an impulse response can be extracted with the descriptors and how accurately (to which extent they correlate to the subjective assessments). Expressing many subjectively relevant components of room acoustics necessarily requires a set of such descriptors. Another important question is whether different descriptors point to the same components, that is, whether they are mutually correlated. In the search for a comprehensive set of mutually uncorrelated descriptors, new descriptors are still being studied and suggested. We will list the most important ones which are well established in literature and norms for room acoustics and thus frequently used in practice.

Reverberation time is defined as the time interval during which sound energy level drops 60dB after a source of stationary noise in the room is switched off. Note that this is different from the decay of energy of an impulse response  $g(t)$  and reverberation time cannot be assessed directly from the curve  $10 \log_{10}(|g(t)|^2)$ . If the source emits a noise signal  $s(t)$  and the room impulse response is  $g(t)$ , signal at the microphone is given by the convolution in eq. (4.15). If the source is switched off at the reception time  $t = 0$ , when the transient regime starts,  $s(t) = 0$  for  $t > 0$ , and

$$p(t > 0) = \int_t^\infty g(\tau)s(t - \tau)d\tau, \quad (6.6)$$

since  $s(t - \tau) = 0$  for  $\tau < t$ .

Reverberation time is a room property which should not depend on any single member  $s(t)$  from the ensemble of possible random noise signals. However, rather than averaging over many measurements with different  $s(t)$  (which should actually be done when reverberation time is measured using interrupted noise technique), we can use the fact that ensemble averaging of  $N$  random signals,  $\langle \cdot \rangle_N$ , gives the same result as time averaging,  $\langle \cdot \rangle_T$ . Ensemble-averaged energy of  $p(t > 0)$  is proportional to (we are interested in the rate of decay only, not in the absolute values of energy)

$$\begin{aligned} \langle p^2(t > 0) \rangle_N &= \left\langle \int_t^\infty g(\tau)s(t - \tau)d\tau \int_t^\infty g(\tau')s(t - \tau')d\tau' \right\rangle_N \\ &= \left\langle \int_t^\infty \int_t^\infty g(\tau)g(\tau')s(t - \tau)s(t - \tau')d\tau d\tau' \right\rangle_N \\ &= \int_t^\infty \int_t^\infty g(\tau)g(\tau') \langle s(t - \tau)s(t - \tau') \rangle_N d\tau d\tau', \end{aligned} \quad (6.7)$$

where we used the fact that the impulse response does not change between different measurements. Averaging the term in angle brackets over infinite time interval rather than over ensemble gives

$$\begin{aligned} \langle s(t - \tau)s(t - \tau') \rangle_T &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t - \tau)s(t - \tau')dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t)s(t + \tau - \tau')dt \approx C\delta(\tau - \tau'), \end{aligned} \quad (6.8)$$

since the last integral is equal to the autocorrelation in eq. (6.1) with the argument  $\tau - \tau'$  and  $s(t)$  is a random signal. Hence,

$$\begin{aligned}\langle p^2(t > 0) \rangle_N &\approx C \int_t^\infty \int_t^\infty g(\tau)g(\tau')\delta(\tau - \tau')d\tau d\tau' \\ &= C \int_t^\infty g(\tau) \left( \int_t^\infty g(\tau')\delta(\tau - \tau')d\tau' \right) d\tau = C \int_t^\infty g^2(\tau)d\tau \\ &= C \left( \int_0^\infty g^2(\tau)d\tau - \int_0^t g^2(\tau)d\tau \right),\end{aligned}\quad (6.9)$$

according to the selectivity of delta function, eq. (3.44). It follows that the ensemble-averaged energy decays over time  $t > 0$  with the same rate at which the cumulative energy of the impulse response increases (the second integral in the last equality). The initial (maximum) energy at time  $t = 0$  is proportional to the total energy of the impulse response (the first integral). Due to the difference of the two integrals, the entire method is called backward integration.

We can now estimate reverberation time from the logarithm of the integrals in eq. (6.9). Moreover, the energy can be normalized with the maximum value at  $t = 0$  (which also removes the constant  $C$ ), so that the logarithm reads

$$SC(t) = 10 \log_{10} \left( \frac{\int_t^\infty g^2(\tau)d\tau}{\int_0^\infty g^2(\tau)d\tau} \right) = 10 \log_{10} \left( 1 - \frac{\int_0^t g^2(\tau)d\tau}{\int_0^\infty g^2(\tau)d\tau} \right). \quad (6.10)$$

This monotonically decreasing function of time is called Schroeder curve. The curve decays starting from the value 0dB. Its minimum value is in practice determined by the noise level, which prevents the recorded impulse response from decaying to zero.

From the definition of **reverberation time** we have

$$T_{60} = t(SC = -60\text{dB}). \quad (6.11)$$

However, signal-to-noise ratio of 60dB is quite difficult to achieve in practice, especially at very low and very high frequencies. Moreover, strong early reflections in an impulse response often cause different decay of  $SC$  in its initial part compared to the rest of its decay. For these reasons, reverberation time is usually assessed from smaller dynamic range, such as

$$T_{30} = 2 [t(SC = -35\text{dB}) - t(SC = -5\text{dB})]. \quad (6.12)$$

Factor 2 is necessary since time interval for the decay of 30dB is used (as indicated by the subscript  $_{30}$ ) rather than 60dB and exponential decay of energy (linear decay of energy level) in time is assumed. If the signal-to-noise ratio of 35dB is not achievable,  $T_{20}$  can also be used:

$$T_{20} = 3 [t(SC = -25\text{dB}) - t(SC = -5\text{dB})], \quad (6.13)$$

or even  $T_{10}$ :

$$T_{10} = 6 [t(SC = -15\text{dB}) - t(SC = -5\text{dB})]. \quad (6.14)$$

Starting point at the curve at -5dB is used for all three descriptors  $T_{10}$ ,  $T_{20}$ , and  $T_{30}$  in order to avoid early decay process. In contrast to this, **early decay time** is defined as

$$EDT = 6t(SC = -10\text{dB}). \quad (6.15)$$

Unlike usual reverberation time, it should indeed quantify the decay of early reflected sound energy, which is considered to be related to the subjective perception of reverberation. This is due to the fact that usual sounds produced in rooms, such as music or speech, rarely allow large drops of sound energy, so the reverberation time is perceived based only on the early decay. Still, the factor 6 is introduced, which allows comparison of the values with reverberation time.

**Bass ratio**, which is used mostly for rooms for music performances, is defined as the ratio of reverberation times at low frequencies (in octaves 125Hz and 250Hz) and middle frequencies (in octaves 500Hz and 1000Hz):

$$BR = \frac{T_{60,125\text{Hz}} + T_{60,250\text{Hz}}}{T_{60,500\text{Hz}} + T_{60,1000\text{Hz}}}. \quad (6.16)$$

This descriptor should point to the frequency balance of the room's response. In practice,  $T_{60}$  is usually replaced with  $T_{30}$ .

While the descriptors based on reverberation time quantify energy decay, other descriptors involve total energy within a certain interval of time. **Definition**  $D_{50}$  is one such parameter which is most often associated with intelligibility of speech. It is the fraction of energy of the impulse response which is contained in its first 50ms. For the time of arrival of direct sound component  $t = 0$ , it equals

$$D_{50} = \frac{\int_0^{50\text{ms}} g^2(\tau)d\tau}{\int_0^\infty g^2(\tau)d\tau} \leq 1. \quad (6.17)$$

A very similar quantity which is used for rating transparency of music is **clarity**  $C_{80}$ . It is the ratio of energy in the first 80ms and the rest of the impulse response expressed in decibels:

$$C_{80} = 10 \log_{10} \left( \frac{\int_0^{80\text{ms}} g^2(\tau)d\tau}{\int_{80\text{ms}}^\infty g^2(\tau)d\tau} \right). \quad (6.18)$$

By changing the corresponding time intervals, variations of these parameters are possible, for example,  $C_{50} = 10 \log_{10}(D_{50}/(1 - D_{50}))$ .

Conceptually similar parameter which is used for assessment of room acoustics from the perspective of the performers is **stage support factor**. It is defined as

$$ST = 10 \log_{10} \left( \frac{\int_{t_{min}}^{t_{max}} g^2(\tau)d\tau}{\int_{0\text{ms}}^{10\text{ms}} g^2(\tau)d\tau} \right). \quad (6.19)$$

where several different pairs of values  $(t_{min}, t_{max})$  are used: (20ms, 100ms) for  $ST1 = ST_{early}$ , (20ms, 200ms) for  $ST2$ , (100ms, 1000ms) for  $ST_{late}$ , and (20ms, 1000ms) for

$ST_{total}$ . However, impulse response  $g(t)$  is in this case measured at the locations of the performers (on the stage or in a pit) at the distance 1m from an omnidirectional source. The parameter  $ST_{early}$  is supposed to correspond to the audibility between the performers (ease of ensemble) and  $ST_2$  to their perception of reverberance of the room (support). According to some authors, only  $ST_{early}$  relates well to both support and ease of ensemble. There is still no strong agreement about the preferred values of these parameters. They usually take negative values between -17dB and -10dB and some authors recommend the values  $-15\text{dB} < ST_{early} < -10\text{dB}$ .

Another approach to estimate the ratio of early and late energy in a room is with **centre time**:

$$t_s = \frac{\int_0^\infty \tau g^2(\tau) d\tau}{\int_0^\infty g^2(\tau) d\tau}. \quad (6.20)$$

It represents the first moment of energy of the impulse response (analogously to the centre of mass, for example). Lower values of centre time are usually preferred, as they indicate that most of the energy is contained in the early part of the impulse response.

A very simple and useful quantity is **initial time delay gap**  $\Delta t_{init}$ . It is the time interval between arrival of the direct sound and the first reflection which reaches the receiver. Energy of the direct sound which is scattered from the body of the source or receiver and thus immediately follows the direct sound is not taken into account for  $\Delta t_{init}$ . The quantity is related to the subjective perception of the distance from the source (also called intimacy). The larger its value is (longer delay of the reflections with regard to the direct sound), the closer the source appears to be and the feedback of the room is perceived as weaker.

**Strength factor**  $G$  quantifies overall energy gain introduced by a room compared to free space. By definition, it is the difference between sound level measured at some location in the room due to an omnidirectional source acting in it and sound level which would be produced by the same source in free space at the distance 10m. The distance of 10m should represent typical distances of the listeners in larger rooms, such as concert halls. Strength factor can also be estimated from an impulse response:

$$G = 10 \log_{10} \left( \frac{\int_0^\infty g^2(\tau) d\tau}{\int_0^{t_{dir}} g_{10m}^2(\tau) d\tau} \right), \quad (6.21)$$

where  $g_{10m}(t)$  is a reference impulse response of the room measured at the distance 10m from the source and  $(0, t_{dir})$  is time interval in which only the direct sound occurs. The interval should be long enough to include entire energy of the direct impulse, but without any succeeding reflections and with as least as possible diffracted energy. In this way, the time windowing leaves only the direct component, which is the only one in free space. Theoretically, the value of strength factor should be constant in a diffuse field at the distances larger than the critical distance. In real rooms, it usually decays with the distance from the source, which confirms the lack of diffuseness. If the two impulse responses in eq. (6.21) are recorded simultaneously, the two microphones should have

equal sensitivities and the signals should be equally amplified. If a single microphone is used for both recordings, than the source signal should have high reproducibility.

Spaciousness of sound field is often quantified by **lateral energy fraction**:

$$LEF = \frac{\int_{5\text{ms}}^{80\text{ms}} g_{fig8}^2(\tau) d\tau}{\int_0^{80\text{ms}} g^2(\tau) d\tau} = \frac{\int_{5\text{ms}}^{80\text{ms}} g^2(\tau) \cos^2(\theta) d\tau}{\int_0^{80\text{ms}} g^2(\tau) d\tau}, \quad (6.22)$$

where  $g_{fig8}(t)$  denotes an impulse response recorded with a bi-directional (figure-of-eight pattern) microphone with the minimum sensitivity (for  $\theta = \pi/2$ ) towards the source and the two maxima in the horizontal plane. This together with the integration interval starting at  $t = 5\text{ms}$  ensures that the direct sound is excluded in the numerator in eq. (6.22).

Spaciousness is perceived owing to the binaural hearing. A descriptor which takes this into account is interaural cross correlation coefficient. It is a measure of similarity between sounds received by the two ears. Normalized interaural cross correlation function equals (compare the numerator with the autocorrelation on the left-hand side of eq. (6.1))

$$IACF(t) = \frac{\int_{t_{min}}^{t_{max}} g_r(\tau) g_l(\tau + t) d\tau}{\sqrt{\int_{t_{min}}^{t_{max}} g_l^2(\tau) d\tau \int_{t_{min}}^{t_{max}} g_r^2(\tau) d\tau}}. \quad (6.23)$$

It is a function of time within the given interval  $-1\text{ms} < t < 1\text{ms}$ , which is approximately the time within which sound can travel the distance between two ears (width of the head). The function is normalized by the denominator, so that the values of  $IACF(t)$  are in the interval  $[-1, 1]$ . Impulse responses  $g_l(t)$  and  $g_r(t)$  should be recorded with the left and right microphone of an artificial (dummy) head or left and right in-ear binaural microphones.

**Interaural cross correlation coefficient** is defined as the maximum absolute value of the interaural cross correlation function:

$$IACC = \max\{|IACF(t)|\}. \quad (6.24)$$

Therefore, it rates maximum similarity between the two signals with a single value in the interval  $[0, 1]$ . Lower values of  $IACC$  indicate less correlation between the sounds at the two ears and, consequently, more spaciousness. Alternatively, **binaural quality index** is also used, which is simply

$$BQI = 1 - IACC, \quad (6.25)$$

so that its higher values indicate more spaciousness.

The integration interval  $(t_{min}, t_{max})$  is commonly  $(0\text{s}, 1\text{s})$  for  $IACC_A$ ,  $(0\text{ms}, 80\text{ms})$  for early interaural cross correlation coefficient  $IACC_E$ , or  $(80\text{ms}, 1\text{s})$  for late interaural cross correlation coefficient  $IACC_L$ . Early interaural cross correlation coefficient,  $IACC_E$ , is associated with apparent source width and  $IACC_L$  with listener envelopment. Interaural

cross correlation coefficient can also be limited to a certain frequency range. For example,  $IACC_{E,3oct}$  can denote early interaural cross correlation coefficient in the three middle octaves: 500Hz, 1000Hz, and 2000Hz. In this way, poor spaciousness and high correlation which are expected at low frequencies are excluded from the assessment.

Speech intelligibility is largely affected by masking of spoken syllables either by ambient noise or by the decaying sound energy of the preceding syllables. Temporal variations of time-averaged sound energy  $\langle E \rangle_T$  of speech can be roughly modelled as simple oscillations around the constant value  $\langle E_0 \rangle_T$ , that is:

$$\langle E \rangle_T = \langle E_0 \rangle_T [1 + m(\Omega) \cos(\Omega(t - t_0))], \quad (6.26)$$

where  $\Omega$  is modulation frequency, which approximates frequency of syllables in the speech, and  $t_0$  is some delay, for generality. In this expression,  $m(\Omega)$  satisfies  $0 \leq m(\Omega) \leq 1$  and can be seen as the amplitude of a complex modulation transfer function  $\hat{m}(\Omega)e^{j\Omega t}$ , that is,  $m(\Omega) = |\hat{m}(\Omega)|$ .

It can be shown that if the modulated signal has flat spectrum in the frequency range of interest and the ambient noise is negligible, complex modulation transfer function is equal to the normalized Fourier transform of the squared impulse response, with modulation frequency in the argument of the transform:

$$\hat{m}(\Omega) = \frac{\int_0^\infty g^2(\tau) e^{-j\Omega\tau} d\tau}{\int_0^\infty g^2(\tau) d\tau}. \quad (6.27)$$

If the ambient noise cannot be neglected,  $m$  can be estimated as  $|\hat{m}|/(1 + 10^{-SNR/10})$ , where  $SNR$  is signal-to-noise ratio. The same reasoning applies both to a broadband sound or any frequency sub-range. Therefore, for example, modulation transfer function in an octave can be estimated with the same procedure after filtering a broadband impulse response with an octave filter.

Apparent signal-to-noise ratio equals

$$SNR_a(\Omega) = 10 \log_{10} \left( \frac{m(\Omega)}{1 - m(\Omega)} \right). \quad (6.28)$$

For **rapid speech transmission index**<sup>52</sup> (*RASTI*), it is calculated in the octave 500Hz for the modulation frequencies 1Hz, 2Hz, 4Hz, and 8Hz and in the octave 2kHz for the modulation frequencies 0.7Hz, 1.4Hz, 2.8Hz, 5.6Hz, and 11.2Hz. The value of *RASTI* is according to the definition

$$RASTI = \frac{\text{avg}\{SNR_a(\Omega)\} + 15}{30}, \quad (6.29)$$

---

<sup>52</sup>*RASTI* is often used instead of the basic **speech transmission index** (*STI*), which includes more modulation frequencies and all octaves between 125Hz and 8kHz. In addition, apparent signal-to-noise ratios in the octave bands are weighted with factors 0.13, 0.14, 0.11, 0.12, 0.19, 0.17, and 0.14, respectively. It has been shown that difference between the values of the two parameters, *RASTI* and *STI*, is not large. Yet another alternative, *STIPA* (speech transmission index for public address systems), uses the same set of octaves as *STI* with only two modulation frequencies per octave.

where averaging  $\text{avg}\{ \}$  is performed over all (maximum nine) apparent signal-to-noise ratios which are in the interval  $[-15, 15]$ dB. With this condition, the values of *RASTI* take the values in the range  $[0, 1]$ . Speech intelligibility is assessed as bad for the values below 0.3, poor between 0.3 and 0.45, fair between 0.45 and 0.6, good between 0.6 and 0.75, and excellent above 0.75.

Another quantity which is used for the assessment of speech intelligibility (also for optimisation of public addressed systems) is **articulation loss of consonants**,  $AL_{cons}$ :

$$AL_{cons} = 0.65 \left( \frac{r}{r_c} \right)^2 T_{60} \cdot 1\%, \quad (6.30)$$

where  $r_c$  is critical distance in the room and  $r$  is distance between the source and receiver. This expression is given for signal-to-noise ratio of minimum 30dB. Higher percentage of the lost consonants is expected if the signal-to-noise ratio is lower. The value of  $AL_{cons}$  should be below 15%, preferably below 7%, and ideally below 3%. The descriptor should not be used for very large distances from the source, when its value can become even larger than 100%. If the value of speech transmission index is known, the value of  $AL_{cons}$  can be estimated as

$$AL_{cons} = 170.5405\% \cdot e^{-5.419 \cdot STI}. \quad (6.31)$$

Table 6.2 associates the introduced descriptors of room acoustics with subjective criteria discussed in section 1.3. Reliable and efficient numerical descriptors for detection and evaluation of strong distinct reflections (which are relevant for the occurrence of echo, coloration, as well as for source localization) and assessment of the texture of impulse response (time distribution of reflections and diffuseness of the field) are still sought for.

Table 6.2: Relations between subjective criteria and objective descriptors of a sound field in room.  $IACC$  can also be replaced with  $BQI = 1 - IACC$ .

| subjective property     | descriptor                                   |
|-------------------------|--|
| loudness                | $G$  |
| intelligibility/clarity | $D_{50}, C_{80}, t_s, STI, RASTI, AL_{cons}$ |
| reverberance            | $T_{10}, T_{20}, T_{30}, EDT$                |
| listener envelopment    | $LEF, IACC_L$                                |
| apparent source width   | $IACC_E$                                     |
| intimacy                | $\Delta t_{init}$                            |
| frequency balance       | $BR$   |
| ease of ensemble        | $ST1 (ST_{early})$                           |
| support                 | $ST2, ST_{late}$                             |

Table 6.3 lists appropriate values of the descriptors for different purposes of the rooms for music production (including singing) and speech. It also lists the most important types of reflections which affect values of the descriptors. The values are given for typical

conditions in the room, for example, half occupied auditorium<sup>53</sup>. However, they should be taken only as a rough indication of the optimum values at middle frequencies. For example, reverberation time value should be corrected with regard to the volume of the room. Some more details for particular types of rooms will be discussed in section 7.4 and further application-dependent values can be found elsewhere in literature as well as relevant norms and guidelines.

Table 6.3: Suggested values of numerical descriptors for rooms for speech and music production and the most relevant types of reflections.

| descriptor                                      | reflections | speech   | music   |
|---|-------------|----------|---|
| reverberation time ( $T_{30}$ )                 | late        | 0.7-1.2s | 1.5-2.2s (concert halls)<br>1.2-1.8s (opera houses) |
| early decay time (EDT)                          | early       | n/a      | 2-2.3s  |
| centre time ( $t_s$ )                           | early       | <100ms   | 100-150ms   |
| initial time delay gap ( $\Delta t_{init}$ )    | the first   | n/a      | 12-20ms   |
| strength factor ( $G$ )                         | early       | n/a      | 4-6dB   |
| definition ( $D_{50}$ )                         | early       | >50%     | n/a   |
| clarity ( $C_{80}$ )                            | early       | n/a      | -2dB to 3dB   |
| speech transmission index (STI)                 | early, late | >60%     | n/a   |
| articulation loss of consonants ( $AL_{cons}$ ) | early, late | <7%      | n/a   |
| lateral energy fraction (LEF)                   | lateral     | n/a      | >0.15   |
| interaural cross corr. ( $IACC_{E,3oct}$ )      | lateral     | n/a      | <0.4  |
| bass ratio (BR)                                 | late        | n/a      | 1.1-1.4   |

As discussed in the beginning of this subsection, the significance of different descriptors and their actual correlation with subjective criteria for room acoustics are a permanent object of research. There is no clear consensus on which set of (ideally mutually non-correlated) descriptors should be used for a complete rating. New descriptors are continuously being developed and proposed, which should provide more accurate and comprehensive assessment. One example is surface diffusivity index (SDI) which relies on the visual evaluation of scattering capacity of large surfaces, mainly ceiling and side walls in rooms for music performances. Properties of the surfaces are compared with the descriptions given in a look-up table and their rates are read out from it. High average rate of all surfaces in the room (high value of SDI) should indicate diffuse field in it. Diffuseness is also associated with texture of an impulse response. However, there is no generally adopted descriptor for it.

<sup>53</sup>Impulse responses are usually measured in empty rooms. For this reason, if the seats in an auditorium provide little absorption, they are sometimes covered with sheets of porous fabric, in order to compensate for the absorption of the absent audience.

## 7 Estimation and optimization of sound fields in rooms

An appropriate strategy for achieving good acoustics in a room is always determined by its purpose. As already discussed in section 1, the main goal can be suppression of noise, accurate sound reproduction, or subjectively appealing enhancement of sound field established by a source of sound in the room. More specifically, the task of room acoustics is treatment of the direct sound and reflections inside the room. With regard to that, Table 7.1 provides an overview of very general strategies for optimization of rooms with different purposes and means to achieve them. The distinction is made between early (fewer and stronger) reflections and late reverberation (many weaker reflections during the decay process), which are typically distinguishable in room impulse responses. The table can also be seen as an extension of Table 1.7 relating the subjective criteria with certain acoustic phenomena.

Table 7.1: General acoustic treatment of rooms depending on the purpose (in Table 1.7).

| <b>room purpose</b> | <b>direct sound</b>                      | <b>early reflections</b>   | <b>late reflections</b>                             |
|---------------------|--|--|---|
| noise control       | /  | absorption   | absorption  |
| sound reproduction  | appropriate strength and time of arrival | absorption or scattering/reflection away from the receiver                                   | absorption  |
| sound production    | appropriate strength and time of arrival | appropriate number, distribution over time, strength, spectral content, direction of arrival | scattering or absorption, appropriate reverberation |

Direct sound alone is irrelevant for noise control if the receiver is outside the zone of the direct sound dominance. In a diffuse field, this means at a distance from the source of noise larger than the critical distance in eq. (5.36) or eq. (5.37). If this is not the case or the direct sound at the receiver location dominates over the reflected sound, noise control cannot be achieved by means of room acoustics, but, for example, by direct interventions on the source or placing it in a closed casing (enclosure). In contrast to this, listener in a room for sound reproduction, such as a control room, should be close enough to the source (loudspeaker) in order to receive accurate information contained in the direct sound without effects of the room. Appropriate strength of the direct sound is somewhat easier to achieve, at least for a certain range of frequencies, if the source is directional and the receiver is localized. However, it becomes much more difficult if the sources and listeners are distributed, such as in cinemas with multichannel reproduction, where obviously not all the listeners can be in the zone of direct sound of all the sources. Moreover, different times of arrival of the direct sound coming from different loudspeakers can also affect the spatial information contained in a multichannel recording. In rooms for sound production, propagation time of the direct sound to a listener should not be too long, in order to keep the synchronization with the visual component. This often

limits the distances between sources and listeners to maximum 20-30m, as discussed in section 1.3. Additional attenuation of the direct sound (typically due to an obstacle on the propagation path or grazing propagation over the auditorium) should be avoided, if possible.

General strategy for rooms for sound reproduction and rooms in which noise should be suppressed involves absorption of both early and late reflected sound. However, it should be emphasized that excessive absorption in a room is normally not advisable. This is particularly true if human listeners are expected to spend longer time in the room, since lack of reflections does not appear natural and anechoic conditions are associated with a feel of unease. This is one of the reasons why musicians cannot perform well in too damped recording studios. Furthermore, suppression of the reflected sound allows higher intelligibility of speech and clarity of background sounds, which can be disturbing or spoil privacy in shared spaces, such as open plan offices. Indistinct weak late reflections can thus help masking the unwanted sounds and make the sound in room more natural. Fully anechoic conditions should be approached only in anechoic chambers for acoustic measurements.

As an alternative to absorption, strong undesired reflections can be scattered or reflected away from the listener. This is particularly useful in rooms for sound production, since the acoustic power of natural sources of sound (human voice and musical instruments) is quite limited. Additional damping in such rooms, besides the unavoidable absorption of the auditorium, can leave insufficient gain of the room and poor coverage with sound energy. Rather than absorbed, reflected sound energy is distributed in a controlled way, evenly over the room. In fact, after appropriate value of reverberation time has been achieved by means of absorption (including the auditorium), careful direction of early reflections and scattering late reflections lead to far better results than additional damping. Reflectors can also be efficient replacement for thick absorbers of early reflections in control rooms. The reflections can be redirected towards a diffuser or absorber in the rear part of the room.

Acoustic treatment is most delicate in rooms for sound production. In particular, more sophisticated subjective criteria for the sound in such rooms give special importance to the early reflections (compare with Table 6.3). Room acoustics of concert halls, opera houses, theatres, lecture halls, and similar rooms is largely management of early reflections. Although their desired properties (strength, time and direction of arrival, spectrum) depend on the type of the room, as will be discussed later, several universal remarks on how they can be controlled can also be given.

Very generally, a treatment from Table 7.1 can be achieved with an appropriate choice of the following:

- room geometry (macrogeometry)
  - volume
  - shape
- basic elements (large surfaces – walls, ceiling, floor)

- size
- position and orientation
- material and microgeometry
- additional (small) elements
  - reflectors
  - diffusers
  - absorbers
- source and receiver (in combination with room acoustics)
  - locations and area
  - directivity.

Unfortunately, not everything listed can be applied or optimized in each particular case in practice and the possibilities for acoustic treatment are therefore limited. If an already existing room should be acoustically treated, its size and shape, as well as the size, position, and orientation of the boundary surfaces are usually already fixed. Number of seats, size, and shape of the auditorium in concert halls, opera houses, and theatres are often set in advance. The locations of musical instruments on the stage and their radiation patterns are often predefined, at least for symphonic orchestra. Even the use of common acoustic elements – reflectors, absorbers, and diffusers – can be limited for visual aesthetic reasons, architectural design, or (especially in the case of rooms of historical value) need to preserve the visual identity of the space. Of course, available budget can also present a constraint for the acoustic treatment.

Which of the possible measures shall be applied should be decided by the acoustic consultant based on the purpose of the room and other available data, such as technical drawings of the room (horizontal and vertical cross sections are usually provided). The measures have to be discussed with and eventually approved by other persons engaged, architect or civil engineer (see Table 1.1) and the investor. Suggestions for the design or its modifications should be made according to the stage of planning or (re)construction of the room, bearing in mind that late changes are usually associated with considerably higher costs than changes in early design stages. This especially holds for rooms of greater importance for the community and large projects. In terms of room acoustics, the suggestions have to rely on the predictions of the sound field done by the acoustic consultant. Different types of tools and analyses can be used:

- analytical/empirical formulas (for a rough estimation of general properties),
- numerical simulations (typically ray tracing in combination with image sources and less frequently Finite Element Analysis or Boundary Element Analysis), and
- experimental methods (laboratory tests, *in-situ* measurements, and measurements on scale models).

Some of the most frequently used formulas which we already derived will be recapitulated here in the context of sound field predictions. Ray tracing, image source technique, and scale models will be considered briefly in section 10. *In-situ* measurements of impulse

responses and the associated descriptors were discussed in section 6.

In larger projects, the analyses should be conducted at different stages of the project, including the *in-situ* measurements of impulse response when appropriate, and particularly when new acoustically relevant data become available. As the details of the room and materials become more precisely defined, the predictions based on these data should become more detailed and accurate, but there are also less possibilities for larger modifications. According to Table 7.1, the analyses should, depending on the application, provide the estimation of

- ambient (background) noise,
- direct sound and visibility of the source,
- reverberation time (late reflections) and sound level distribution, and
- temporal and angular distribution of early reflections and their energy.

Although too high background noise can degrade the overall acoustic performances of a room, it is primarily an issue of sound insulation<sup>54</sup>. Therefore, we will consider the latter three points in the following. While the first two of them can be fairly accurately and easily assessed even with analytical or semi-empirical procedures, early reflections can be estimated only roughly, which inevitably carries certain amount of inaccuracy and risk. Acquired experience is especially helpful for such analyses.

## 7.1 Direct sound and visibility

In the absence of obstacles between them, a line of sight exists between a source and listener and the time of arrival of the direct sound,  $t_0$ , depends simply on the distance  $r$  between them. It can be calculated as  $t_0 = r/c_0$ , for the emission time  $t = 0$ . Amplitude of the direct sound decays according to eq. (4.11) proportionally to  $1/r$ , or 6dB with doubling the distance from the source. With a known power level of a directional point source, the associated sound level can be calculated using eq. (4.42). However, if the sound wave propagates just above an absorbing surface, for example an auditorium, and parallel to it (at the grazing angle), somewhat larger decay with distance should be expected. This means that the remote listeners in the back rows of the auditorium may receive even weaker direct sound than in free space, in addition to the poor visibility of the source. This problem can be addressed in two ways: by minimizing the distance between the source and receiver and by increasing the angle at which different listeners in the audience see the source (angle between the line of sight and the horizontal plane).

In terms of the average distance between listeners and a source for a fixed total area of the auditorium, round auditoriums (full circle or its segment) with the source located in the centre outperform all other room shapes. They are followed by horse shoe, fan-shaped,

---

<sup>54</sup>Sometimes also due to the installations inside the room (ventilation, light, etc.). An additional care should be taken about such issues in practice, even though they might not be directly in the scope of the work of the acoustic consultant. Moreover, sound insulation is also very often part of the same project.

and diamond-shaped auditoriums, with the source placed in front of the shortest walls of the rooms. The least favourable rooms according to this criterion are narrow rectangular rooms with a source in front of one of the two shorter walls. However, this criterion, which allows in average stronger direct sound and better visibility of the source (or larger number of seats), turns out to be in contrast to the requirement for strong lateral reflections (early reflections will be discussed in subsection 7.3). Namely, relatively narrow rectangular rooms tend to provide more lateral sound energy (contributing to spaciousness) in larger part of the auditorium than rooms with distant side walls or large angles between the front wall (behind the source) and side walls, as in fan-shaped halls. This explains why rectangular rooms are still preferred for concert halls, where spaciousness presents a very important subjective criterion besides visibility of the source and the direct sound strength, while horse shoe and fan shapes are more common for opera houses and theatres, where the visibility and strong direct sound are more critical.

Angle between the line of sight and horizontal plane can be increased either by introducing an ascent (gradual increase of height) of the auditorium, or by lifting the source above the plane of the auditorium (for example, hanging a loudspeaker, or placing the orchestra on a stage with downward slope). This avoids grazing propagation of both direct sound and early lateral and front wall reflections thus improving spaciousness and overall early-to-late energy ratio as well. In order to keep the same visibility angle (or approximately constant distance between the lines of sight for listeners far from the source) of each two listeners one behind the other<sup>55</sup> throughout the room, slope of the auditorium should gradually increase with the distance from the stage. For practical reasons, such increase can be approximated by subdividing the auditorium into several sections with different but constant slope angles, which increase towards the back of the room.

One way to increase the capacity of an auditorium but keep satisfactory distances of the listeners from the source is with balconies or galleries. However, too deep balconies usually lead to poor coverage with sound in the rear rows of seats, both on the balconies and below them. On the other hand, relatively short galleries (with several rows of seats) can improve diffuseness of the field even at low frequencies, which would be difficult to achieve with scattering surfaces of walls, and provide more early reflections to the auditorium below them, for example, from the bottom of side galleries. As a rule of thumb, height-to-depth ratio of the balconies should not be smaller than one.

## 7.2 Reverberation time and sound level

Reverberation time is usually estimated using Sabine's formula (5.35). Unless the room is very large and weakly damped, dissipation in air can be ignored in most of the cases. Eyring's equation (5.45) can also be used, especially if average absorption coefficient in

---

<sup>55</sup>Value of this distance should generally be at least 6cm and preferably around 10cm or larger. Angle between the line of sight and surface of the auditorium at the point of the listener should ideally be at least 15°.

the room is not very low. On the other hand, very damped rooms do not satisfy the assumption of a diffuse field and both equations can be inaccurate. As an alternative, reverberation time can be assessed numerically from the energy responses obtained with ray tracing simulations. In practice, the accuracy is also limited by the availability and reliability of data on the absorption coefficients of various materials in the room. Such data are usually gathered from data sheets or literature, or measured in a laboratory in a diffuse field (reverberation chamber) or at normal incidence (Kundt's tube). Deviation of the actual sound field in a room from the ideally diffuse field or plane wave at normal incidence affects the actual absorption properties of the materials.

Rooms for music performance are usually weakly damped and the main absorbing surface in them is the auditorium. For rough initial estimations of reverberation time, the absorption coefficient of a fully occupied auditorium can be set around 0.8 at middle and high frequencies, while a small non-zero value, say 0.05, can be adopted for all other reflecting surfaces. Effective seating area is often approximated as the net surface area of the auditorium extended by 0.5m in all directions without reaching any wall or boundary in the room. This extension takes into account additional absorption at the edges of the auditorium. If the auditorium is split into several separate sections, such extensions should be added to each section.

Absorption coefficient of unoccupied seats is relatively difficult to assess in general, as it depends on the applied materials and size of the upholstery. If more specific data or measured values of the absorption coefficient are not given, they can be estimated based on the data for similar types of seats in literature. In general, it is often desirable that unoccupied chairs provide similar absorption as when they are occupied, so that the total absorption and therefore reverberation time in the room do not vary significantly with the occupancy.

When present, other significant causes of absorption, such as organs, diffusers with non-negligible absorption, porous or membrane absorbers, wall linings, or dissipation in air in large reverberant spaces (such as churches and sport halls) should also be included in the estimation of reverberation time. It is up to the experience of the acoustic engineer to assess how relevant and reliable the available data on the absorption are and if they can be applied in the specific context (location of the materials, orientation, mounting, etc.). If especially high risk for an accurate estimation is associated with particular elements, usually those covering large surfaces of the room (for example, seats in the auditorium), their absorption should be measured in a laboratory.

Assessment of reverberation time is more complicated if the main room (where the listeners are) is coupled with another adjacent room (or more rooms) through an open surface (interface) which is much smaller than the total area of both main and adjacent room. Typical example are stage houses in opera houses, which are coupled with the main hall by the proscenium as the interface, as well as boxes in theatres (when they are almost closed), or aisles and naves in churches. Different sound decay processes in the coupled rooms can in general lead to non-linear level decay. Reverberation times of the coupled rooms can then be assessed independently, by treating the interface as

a boundary surface of each room which it couples with a certain absorption coefficient value. If the reverberation time in the coupled side room is much shorter than in the main room, it will not affect the perception of reverberance in the main room. The interface can then be treated as an absorbing boundary of the main room, with absorption coefficient above 0.5. In the opposite case, when reverberation time in the coupled room is longer than in the main room, the coupled reverberation can be perceived in the main room, especially if the source and/or the listener are close to (or even inside) the coupled room. This is most often undesired and the coupled room should be additionally damped.

As indicated in Table 6.3, reverberation time should be around 1s in rooms for speech and around 2s in rooms for music performance. Opera performances involve both types of sound, so the recommended values in opera houses are around 1.5s. These values hold at middle frequencies, in octaves 500Hz-2kHz. It is common that reverberation time increases towards low frequencies and decreases at high frequencies. However, it should not be much longer at low frequencies, in order to maintain the intelligibility of speech (to avoid masking by the low-frequency background noise). Increase at low frequencies up to 20% at 63Hz and drop at high frequencies up to 40% at 8kHz compared to the values at the middle frequencies are usually acceptable, with somewhat higher tolerance for music. Suggested values also depend on the volume of the room. For a room with volume  $1000\text{m}^3$ , reverberation time for speech should be around 1s and for music around 1.5s. The suggested values increase roughly 0.1-0.15s per doubling the volume around these values. Hence, the value of reverberation time around 2s in larger concert halls.

Suggested reverberation time values are approached by ensuring appropriate amount of absorption for the given volume of the room. If the room is intended for various types of sources of both speech and music (a multipurpose room), adaptive reverberation time is sometimes achieved with movable walls (for varying the volume) or absorbing surfaces, for example, curtains, or rotating elements with absorbing and reflecting (or scattering) sides. The achieved range of reverberation time values depends, according to eq. (5.35), on the absorption coefficient of the moving surfaces, fraction of the total area of the room which they occupy, and the change of room volume.

Reverberation time in a room is related with the sound level produced by a source in it. For example, constant sound pressure level in a diffuse field of an omnidirectional point source at distances larger than the critical distance can be estimated from eq. (5.26). In reality, lack of diffusion usually causes decay of sound level with the distance from the source (except close to reflecting surfaces). This is more pronounced in rooms with irregular shapes, such as long or flat rooms, as well as when absorption is not uniformly distributed inside the room.

In order to appreciably decrease sound level due to a source of noise by means of room acoustics, sufficiently large area of the room should be covered with materials with sufficiently high value of absorption coefficient. Although this is not always an easy (or inexpensive) task, even several decibels lower noise level can significantly improve the

acoustic comfort. However, it is worth mentioning once again that the room should never be overdamped.

### 7.3 Early reflections

Accurate prediction of early reflections in a room is in general more difficult than estimation of time-averaged energy (sound level) or decay of late reflections (reverberation time). Number, rate of occurrence, strength, spectral content, time of arrival after the direct sound, path length, and direction or arrival of early reflections all depend not only on the absorption coefficients of surfaces and their areas, but also their geometry, shapes, orientation, roughness, as well as locations of the sources and listeners relative to them, their directivity, and dissipation in air. It takes certain experience to recognize the most critical reflections inside the room, track down their propagation paths between a source and listener, and assess their properties. Besides technical drawings of the room, certain tools can make this task easier, such as ray tracing simulations and (partial) scale models. Of particular interest are strong distinct reflections or groups of reflections (such as those coming from concave reflecting surfaces), which may cause increase of sound energy at particular locations in the room. Such lack of diffusion and uniformity of the sound field is usually undesirable.

In general, difference between sound pressure level of the direct sound and an  $N^{\text{th}}$ -order ( $N$  times reflected) specular reflection from an omnidirectional point source can be estimated at the location of the listener as

$$L_d - L_r \approx 20 \log_{10} \left( \frac{r_r}{r_d} \right) - \sum_{n=1}^N 10 \log_{10} [1 - \alpha_{sn}(\theta_{in})] - \sum_{n=1}^N 10 \log_{10} [1 - s_{sn}(\theta_{in})]. \quad (7.1)$$

The first term follows from eq. (4.34) and involves ratio of the total path length of the reflection,  $r_r$ , and path of the direct sound,  $r_d$ . It is a consequence of the expansion of the wave front and the 6dB decay per doubling the distance. Besides the absorption coefficient of each reflecting surface on the path,  $\alpha_s(\theta_i)$ , we have also introduced the scattering coefficient,  $s_s(\theta_i)$ , which represents the ratio of non-specularly reflected (scattered) energy and total reflected energy. Both quantities can depend on the angle of incidence  $\theta_i$  (we neglect possible dependence on angle  $\phi$ ) as well as frequency. Similarly as the direct sound, early reflections which propagate over highly absorbing surfaces such as auditorium can be additionally attenuated. As already discussed, this can be avoided by introducing a slope of the auditorium or by elevating the source. Time difference of arrival of the reflection and direct sound can be estimated as  $(r_r - r_d)/c_0$ .

It is usually reasonable to expect that the lowest order reflections (for  $N = 1$  and  $N = 2$ ) will be dominant and therefore the most relevant for room acoustics. For example, a performer or a speaker on the stage should not be located too close to the front edge of the stage, if they are to benefit from the immediate reflection from the stage floor. This sound energy can be redirected towards the auditorium as a second-order reflection from the ceiling or a hanging reflector above the stage. Unfortunately, accuracy of the

estimation of early reflected energy according to eq. (7.1) is largely limited by difficult estimation of the scattering coefficients, the values of which are frequently unavailable, with the exception of large flat surfaces with  $s_s \approx 0$ .

Optimal properties of early reflections are application dependent. Therefore, they will be discussed in the next subsection for particular types of rooms, as well as certain approaches how they can be optimized.

## 7.4 Acoustic design

In the rest of this section we briefly cover the main issues and basic strategies for room acoustics of some of the most common room types. Many more specific details can be found in literature.

### 7.4.1 Noise control in rooms

Room acoustics is relevant for noise control when acoustic interventions on the source of noise inside the room (typically enclosing it with a casing or affecting directly the mechanism of noise generation) are not sufficient or feasible. This is often the case with large or distributed sources, humans as the sources of unwanted sounds, cooling equipment, etc. Diffuse field equation (5.26) indicates that sound level in the room can be decreased by increasing the average absorption in it. However, this holds only for the locations outside the zone of direct sound of the source, that is for  $r > r_c$ , where  $r$  is distance from the source and  $r_c$  is critical distance (see Fig. 10 (left) and equations (5.36) and (5.37)). Equation (5.26) can be used for estimation of noise level in octave bands. A single-number value can then be obtained using NR-curves (Fig. 3 in section 1.2.1) or A-weighting.

It should be noted that the noise suppression depends on the relative increase of the average absorption coefficient compared with the initial state. Hence, such interventions are efficient in practice only in otherwise low-damped rooms, when the average absorption coefficient can be appreciably increased with a reasonable amount of added absorbing material, for example, with absorbing suspended ceiling. If the room is already damped, sound insulation measures are much more efficient. As a side effect, increase of absorption also leads to a larger zone of direct sound, which may limit the efficiency of room acoustic interventions if the listeners are relatively close to the source.

### 7.4.2 Rooms for sound reproduction

In rooms for critical sound listening and monitoring the focus is on particular locations of the sources (loudspeakers) and receivers (typically at the height 1.2m from the floor). These are commonly defined by the audio format. Optimal response of the room must be achieved only for those locations. In control rooms it is recommendable that all

reflections are at least 20dB weaker than the direct sound from the loudspeakers, in order to avoid coloration or change of the apparent location of the source (see section 1.3). The requirement for the earliest reflections is sometimes relaxed to 10dB below the direct sound.

A frequent measure to avoid strong early reflections is mounting the loudspeakers into the walls rather than closely in front of them, in order to avoid immediate first-order reflections from the walls. They should also be placed so to avoid the immediate reflection to the listener from the mixing desk. Parts of the walls and ceiling which are closest to the loudspeakers should absorb or reflect the sound away from the listener's location. Strong reflections from the rear wall (behind the listener) are usually avoided by placing a diffuser or absorber. The combination of highly absorbing front part of the room and diffuse rear part is also known as "Live End, Dead End (LEDE)". In cinemas, broadband absorption is usually introduced at both front (behind the screen) and rear walls, as well as on the floor (carpet) and the seats, due to the larger volumes. Front parts of the side walls are sometimes left reflecting to achieve uniform coverage in the back rows (the receiver is distributed in contrast to control rooms).

Late reflections should be as diffuse as possible and reverberation time should be low, around 0.3s for the room volume  $100\text{m}^3$ . A rough dependence of the optimum reverberation time on the room volume can be given as  $T_{60} = 0.3(V/100\text{m}^3)^{1/3}\text{s}$  (sometimes  $T_{60} = 0.25(V/100\text{m}^3)^{1/3}\text{s}$ ) for  $V > 35\text{m}^3$ . Tolerance is usually  $\pm 20\%$  and somewhat larger below 250Hz and above 2kHz, even up to around +50% at 63Hz and -50% at 8kHz. Optimal volume of a cinema can be estimated with around  $4\text{m}^3$  per seat.

Additional issue associated with control rooms is due to their size which is very often relatively small. As a consequence, Schroeder frequency given in eq. (3.90) can be high, well in the audible range, and the dominance of distinct modes (as in eq. (3.63)) is to be expected at low frequencies. As a result, frequency response is not flat as it should be. This is the main reason for high damping requirements at low frequencies in small rooms for music listening. However, lack of space also limits the possibilities for installation of large absorbing elements efficient at low frequencies. Low-frequency membrane absorbers, bass traps, and careful choice of the source and listener locations (usually avoiding the central zone of the room) can all help treating this problem. Room volumes below  $100\text{m}^3$  should be avoided and room proportions should not be simple (for example, integer values) in order to avoid the overlap of the lowest modes. Simple room shapes such as rectangular should also not be used, although left-to-right symmetry with respect to the listener is often desirable, since it gives more equal frequency responses of the room for symmetrically located loudspeakers.

### 7.4.3 Recording studios

These are primarily music studios, but also broadcasting (radio and TV) studios. As in rooms for sound reproduction, optimal room impulse response should be achieved for specific locations of sources and receivers, if they are well defined. In addition to

this, proximity and directivity of the microphones used for recording can be utilized to suppress possible defects of the room acoustics. By placing a microphone in the zone of dominant direct sound and directing the maximum of its sensitivity towards the source, a response with suppressed reflections can be achieved without excessive damping inside the room. This is even more efficient if the source is directional, too. More pleasant ambient is thus achieved with less damping and possibly more scattering surfaces, in order to avoid potentially strong specular reflections. The use of microphones thus allows much greater flexibility in the design of recording studios than in the case of control rooms.

If the room is acoustically small (occasionally even smaller than  $30\text{m}^3$ ), similar remarks regarding the room modes at low frequencies apply as for control rooms. Very small rooms should have correspondingly short reverberation time, perhaps down to around 0.3s. However, if the room is used for recording only voice (for example, studios at radio stations), its frequency response below 100Hz becomes less critical than in music studios. Furthermore, sources and receivers are usually movable, so preferred locations in the room can be determined by inspection.

#### 7.4.4 Lecture halls

We consider lecture halls here as relatively large (in comparison to classrooms) rooms for speech. According to section 7.2, appropriate reverberation time is around 1s ( $\pm 20\%$ ) at middle frequencies for room volume around  $1000\text{m}^3$ . Dependence of optimum reverberation time on volume can be expressed roughly as  $T_{60} \approx 1\text{s} + 0.1\text{s} \cdot \log_2(V/1000\text{m}^3)$  or similarly  $T_{60} \approx 0.37\text{s} \cdot \log_{10}(V/1\text{m}^3) - 0.14\text{s}$ . The allowed tolerance is usually linearly increased below 250Hz and above 2kHz up to around (-50%,+70%) at 63Hz and (-50%,+20%) at 8kHz. In small classrooms (below  $500\text{m}^3$ ), the recommended reverberation time is about 0.2s shorter for the given value of  $V$  ( $T_{60} \approx 0.26\text{s} \cdot \log_{10}(V/1\text{m}^3) - 0.14\text{s}$ ). Ideally, reverberation time in rooms for speech should drop at the lowest frequencies, especially below the frequency range of speech (below the octave 125Hz). This is often achieved due to considerable transmission of low-frequency sound through the windows.

The parameter which is also frequently used is the ratio of total volume of the room and number of seats. For larger lecture halls it should be between  $4\text{m}^3/\text{seat}$  and  $6\text{m}^3/\text{seat}$ , with the maximum volume around  $5000\text{m}^3$ , and in classrooms between  $3\text{m}^3/\text{seat}$  and  $5\text{m}^3/\text{seat}$ . Very large lecture halls, with more than 500 seats, usually require the use of sound reinforcement and consequently somewhat lower reverberation time.

As indicated in Table 6.3, value of definition  $D_{50}$  should be above 0.5 and speech transmission index above 0.6, or optimally above 0.75. For useful early reflections, the ceiling should not be higher than 8-10m, otherwise hanging reflectors are recommended. In the absence of reflectors, part of the ceiling above the podium is usually inclined towards the auditorium.

Subjective impression of spaciousness in rooms for speech is less relevant than in rooms for music. Therefore, fan-shaped and round auditoriums are common, which can accom-

modate relatively large number of listeners at acceptable distances from the speaker, thus ensuring good visibility and coverage with the direct sound. Galleries can also be introduced to further increase the number of seats. Slope of the auditorium (both on the floor and the galleries) avoids attenuation of the direct sound due to grazing propagation over the auditorium and increases the overall intelligibility. Similarly, when sound reinforcement is used, the loudspeakers should be placed well above the stage and the front rows of the auditorium and directed towards the auditorium.

In the absence of sound reinforcement, early sound energy (which reaches a listener in the first 50ms after the direct sound and therefore constructively contributes to the intelligibility without masking effects) is increased by ensuring enough early reflections. Since source location is usually less distributed than the auditorium, this is most efficiently achieved by careful positioning and orientation of highly reflecting surfaces (walls or reflectors) close to the source. If the ceiling is not too high to provide desirable early reflections, its front and central parts can be used for this purpose and should be left reflecting. Since short time of arrival of such reflections is critical, the reflecting surfaces should not substantially scatter the sound away from the auditorium (which might be preferred in rooms for music performances). If the ceiling is too high so that its first-order reflections reach the listeners too late, hanging reflectors can be introduced.

In contrast to the surfaces close to the source, rear walls and remote parts of the ceiling in larger rooms are often diffuse or inclined, in order to prevent late and strong low-order reflections from reaching the front parts of the auditorium (or sometimes flutter echo between the front and back wall). The redirected energy is in such a way delivered mostly to the rear parts of the auditorium, where the delay with respect to the direct sound is shorter. When lower values of reverberation time should be achieved, these remote surfaces can also be absorbing. In general, diffuse reflections and direction of arrival of sound energy to the listeners play much less important role than in rooms for music performances.

#### 7.4.5 Theatres

Similarly as in lecture halls, fan-shaped auditoriums are quite common in drama theatres, mainly for visual reasons and higher speech intelligibility due to stronger direct sound. Galleries are also often introduced to increase the capacity of the hall. Although larger room volumes than in lecture halls are common (up to around 1000 seats without sound reinforcement, due to the trained voice of the actors), the ratio of volume and number of seats should be similar,  $4\text{-}7\text{m}^3/\text{seat}$ , as well as the reverberation time. Apart from the slope of the auditorium (on the floor and the galleries), slope of the stage (when the stage design allows it) can further improve the visibility and direct sound. If sound reinforcement is used, it should be placed above the audience, often above the proscenium.

Probably acoustically the most significant difference between large lecture halls and drama theatres is the stage. It is usually significantly larger in theatres. In addition

to this, stage houses in theatres can be much higher than proscenium, which can thus present a coupled space to the main room with the audience or provide late reflections, for example, from the back wall of the stage. The undesired late energy can be treated with absorbers inside the stage house. To a certain extent, its acoustic effects depend also on a particular scenery on the stage.

Speech intelligibility in theatres should be supported by a sufficient number of strong early reflections reaching the listeners not later than approximately 50ms after the direct sound. Their direction of arrival and subjective impression of spaciousness are not as critical as in concert halls. Since the auditorium is distributed, it is reasonable to place strongly reflecting surfaces close to the stage. Unfortunately, unlike in lecture halls, in which they can usually be placed with a greater flexibility, additional reflectors on the stage or just above it conflict with the stage design (unless they happen to be part of the scenery). This limits the possibilities for acoustic interventions and makes the reflections from parts of the side walls and ceiling closest to the stage crucial. Ideally, these surfaces should also reflect part of the sound energy back to the performers, to give them an acoustic feedback (support). The ceiling should not be higher than around 10m, in order to provide useful early reflections to the auditorium. In average, it should be somewhat lower than in the rooms for music (with height up to around 15m), which can be directly associated with the difference of time intervals for definition  $D_{50}$  and clarity  $C_{80}$  (30ms corresponds to the path length difference of around 10m).

Almost closed and highly damped boxes should be avoided in general, since they often lead to excessive damping in the room (adding to the unavoidable absorption of the auditorium) or coupling effects (when the damping inside them is low). Another cause of additional, often undesired absorption are sometimes empty slits in the ceiling, which are used for stage lighting. In such cases, the slits should be closed. The inclined surfaces of the ceiling below the lighting are typically used for controlling the coverage of the auditorium with the sound energy reflected from the ceiling.

#### 7.4.6 Opera houses

Surface area and shape of the auditorium in opera houses are determined primarily by the visual criteria. Similarly as in theatres, actions on the stage, including the details such as facial expressions, mimics, and gesticulation of the performers, should be visible to the audience. This largely limits the maximum distance between a listener and the stage and practically forces the use of galleries in order to accommodate larger number of spectators. Ascent of the auditorium on the floor and galleries improves the visibility as well as strength of the direct sound.

Optimal reverberation time is somewhat higher than in rooms for speech and lower than in rooms for music performances, normally between 1.2s and 1.8s (roughly 1.2s for room volume  $1000\text{m}^3$  and increasing 0.12s per doubling the volume). A small increase of the value at low frequency is usually acceptable to support the music performance. The volume of  $5\text{-}8\text{m}^3$  per seat is considered to be appropriate and the total volume should

not exceed  $15000\text{m}^3$ , corresponding to the maximum of around 2000 seats.

Additional reverberation in the auditorium due to the coupling with the stage house should be avoided with sound absorption inside the stage house. Proscenium thus acts practically as a highly absorbing surface with respect to the room with the auditorium. In crude predictions of the sound field, it is often approximated as a fully absorbing surface, or a high value of absorption coefficient is attributed to it (above 0.6). However, certain parts of a large stage (for example, a reflecting wall parallel to the proscenium) can be a cause of strong late reflections to the audience, which should be avoided. As in theatres, the most useful reflecting surfaces are side walls and ceiling close to the stage. They should ensure sufficient reflected energy to the auditorium and performers on the stage, and orchestra in the pit.

Orchestra pit should have floor surface around  $1.5\text{m}^2$  per member of the orchestra. Even if it does not affect the visibility of the stage from the auditorium, orchestra pit should not be too shallow (less than 2-2.5m deep). Strong reflections inside the pit allow the players to hear each other well, which improves their common performance (ease of ensemble). Moreover, height of the orchestra pit is usually adjustable. Low-frequency absorption inside the pit is often recommendable to suppress the resonances, which might also affect the balance in the rest of the hall.

#### 7.4.7 Concert halls

Compared with the rooms for speech, in rooms for music performances much larger importance must be given to the diffuseness of the sound field and, therefore, scattering of the reflected sound energy. The condition for short time of arrival of early reflections is more relaxed for the clarity of music than for speech intelligibility, with the most relevant time interval after the arrival of the direct sound about 80ms. The initial time gap should be kept below around 20ms, which can be critical in large halls. On the other hand, diffuse field achieved with many reflections reaching the listeners from different directions contributes positively to the feeling of spaciousness and increases the values of binaural quality index *BQI*. Therefore, even early reflections from the surfaces close to the stage should scatter the sound, at least partially. Too focused early reflections can also sometimes mask the overall perception of reverberation at certain places in the room.

Late reverberation contributes to the impression of the space in which the music is performed. Hence, late reflections are more significant than in rooms for speech resulting in somewhat larger values of reverberation time. Optimal reverberation time for symphonic music is around 1.4s for room volume  $1000\text{m}^3$  and increases around 0.15s per doubling the volume. A slight increase towards the lower frequencies is also common. However, it also depends on the type of music, so for organ music it is around 0.3s longer for the same room volume. Small concert halls (for small orchestra and chamber music) with volumes below  $10000\text{m}^3$  should provide  $6-10\text{m}^3$  per seat, while large halls, usually up to  $25000\text{m}^3$ , should have the ratio of volume and number of seats between

$8\text{m}^3/\text{seat}$  and  $12\text{m}^3/\text{seat}$ . Even larger volumes and ratio values up to  $14\text{m}^3/\text{seat}$  can be appropriate for organ music. Total number of seats should be between 1500 and 2000 for symphonic orchestra, in order to achieve a good coverage with sound of the entire auditorium, or somewhat smaller, less than around 1000, for small orchestra and chamber music. When necessary, additional absorption at low frequencies is usually achieved by means of membrane absorbers on the walls or ceiling and less commonly with Helmholtz resonators.

As commonly, side walls and ceiling are critical for the control of early sound energy, especially their position and orientation with respect to the stage. Conventional rectangular shape of the hall with the stage in front of a shorter wall provides sufficient lateral sound energy even in the back rows of the auditorium, especially closer to the side walls, which is not the case in round or fan-shaped halls. This is paid by a larger average distance of the listeners from the stage (meaning weaker direct sound and worse visibility) for the same total number of seats. The number of seats can be increased with galleries, which should not be too deep (preferably up to several rows only). Slopes of the auditorium on the floor and the galleries and/or the stage improve the visibility, direct sound strength, as well as the strength of early lateral reflections. Larger number of seats can be achieved in vineyard-shaped rooms. The auditorium is then usually subdivided into several sections at different heights. This introduces additional side walls for parts of the auditorium as well as important lateral reflections.

When present, bottom of the galleries can together with the closest wall introduce strong second-order reflections, which provide more lateral and back sound energy not attenuated by the grazing propagation over the audience. In the absence of galleries, similar holds for the ceiling if it is low enough (optimally between 5m and 10m high), or sufficiently large reflectors mounted on the side walls or the ceiling. Ceilings in rooms for music performances have an additional importance, since certain musical instruments, such as violin and piano, radiate large portions of sound energy in vertical direction, particularly at high frequencies. This energy can be captured and directed by the parts of the ceiling above and in front of the stage, or with hanging reflectors. When flat, the reflectors are usually not too large (with characteristic dimensions  $\sim 1\text{m}$ ; compare with section 9.1.1), in order to scatter the sound at middle frequencies.

Apart from supplying the audience with enough diffuse sound energy, walls and ceiling, and other surfaces close to the stage should provide appropriate feedback to the performers, as well, as discussed in section 1.3. Ease of ensemble is particularly important for large orchestras. The performers should hear each other well to keep the intonation and synchronisation. This can be partly achieved by decreasing the distance between the musicians on the stage. However, strong local early reflections on the stage are usually necessary for a good synchronization between the musicians at the opposite sides of the stage. If the side walls and ceiling cannot provide them (which is also often the problem with outdoor stages), the reflections can be achieved with an appropriate stage enclosure. For the same reasons, stage is sometimes placed in a recess in the hall, surrounded by reflecting surfaces. In such cases, the surrounding surfaces take over part of

the role and importance of the side walls and ceiling in the rest of the room. The most important geometrical aspects remain their distances (they should not be too close to the performers either) and orientations with regard to the stage. Parallel surfaces are not recommended in order to avoid flutter echo. Area of the stage should be around  $1.8\text{-}2\text{m}^2$  per musician, which is in practice around  $200\text{m}^2$  for large halls and around 4 times less in halls for chamber music. The floor of the stage should be massive enough and should not resonate noticeably.

Surfaces which are far from the stage are a potential cause of strong late reflections, especially in the front part of the auditorium. In extreme case, these can be perceived as echo. This can be avoided by applying absorbing materials on such remote surfaces. However, this results most often in too short reverberation time in the room (since the absorption adds to the absorption of the auditorium). Therefore, redirection of the reflections with inclined surfaces or scattering with non-flat surfaces is usually preferred. An important aspect in concert halls are the seats, which are often made lightly to moderately upholstered with porous materials (at the parts of the seats covered by the audience, when present), so that the reverberation time does not largely depend on the occupancy of the hall.

## 8 Basic acoustic elements

In this section, we study acoustic behaviour of large plane surfaces in a room, such as walls or ceiling, either as simple rigid boundaries, or treated with common types of sound absorbers.

### 8.1 Infinite uniform plane surface

We observe a flat surface normal to the  $x_1$ -axis and extending to infinity in all directions (in practice, its dimensions should be much larger than the wavelength). The surface is assumed to be uniform and in the far field of any source of sound (or other reflecting surfaces as secondary sources), so the incident sound wave denoted with sound pressure  $p_i$  can be approximated as a plane wave. According to eq. (4.36):

$$\hat{p}_i(\mathbf{x}) = \hat{p}_Q e^{-jkr} = \rho_0 c_0 \hat{\mathbf{v}}_i(\mathbf{x}) \cdot \mathbf{e}_i, \quad (8.1)$$

where  $\mathbf{e}_i$  is unit vector in the direction of wave propagation and  $\hat{Q}(\mathbf{y})/(4\pi r)$  is replaced with  $\hat{p}_Q$  which thus depends on the source and includes the initial phase of the wave. Since we are interested only in the local effects close to the surface, the value of  $\hat{p}_Q$  does not change significantly with propagation path length ( $r$ ) and we can take the value of  $\hat{p}_Q$  to be constant in both space and time, as for a plane wave in the far field. Spatial dependence is therefore contained only in the phase term  $e^{-jkr}$ .

Due to the axial symmetry of the geometry (the surface is infinite, flat, and uniform), we can treat the problem as two-dimensional, in the  $x_1 x_2$ -plane, with the surface placed at  $x_1 = 0$  and the incident sound wave reaching the surface at  $x_2 = 0$  from the quadrant  $x_1, x_2 < 0$ , without any loss of generality. If  $0 \leq \theta \leq \pi/2$  denotes the angle of incidence (which is equal to the angle between  $\mathbf{e}_i$  and the unit vector  $\mathbf{e}_1$  in the direction of the  $x_1$ -axis), then  $\mathbf{e}_i$  has components  $(\cos(\theta), \sin(\theta))$ . Since we work with Cartesian coordinates  $x_1$  and  $x_2$  attached to the surface, we should replace  $kr$  (with the radial coordinate  $r$ ) with  $\mathbf{k}_i \cdot \mathbf{x}$ , where  $\mathbf{k}_i = k\mathbf{e}_i = \omega\mathbf{e}_i/c_0$  is wave vector of the incident wave. Thus we obtain

$$\hat{p}_i(\mathbf{x}) = \hat{p}_Q e^{-jk_i \cdot \mathbf{x}} = \hat{p}_Q e^{-j(k_{i1}x_1 + k_{i2}x_2)} = \hat{p}_Q e^{-j[kx_1 \cos(\theta) + kx_2 \sin(\theta)]}. \quad (8.2)$$

Since the surface is flat and uniform, we expect only specular reflection at the same angle  $\theta$ , that is,  $\mathbf{e}_r = (-\cos(\theta), \sin(\theta))$  and  $\mathbf{k}_r = k\mathbf{e}_r$  are direction and wave vector of the reflected plane wave, respectively. Therefore, the reflected wave can be written most generally as

$$\hat{p}_r(\mathbf{x}) = \hat{p}_Q \hat{R}_s(\theta) e^{-jk_r \cdot \mathbf{x}} = \hat{p}_Q \hat{R}_s(\theta) e^{-j(k_{r1}x_1 + k_{r2}x_2)} = \hat{p}_Q \hat{R}_s(\theta) e^{-j[-kx_1 \cos(\theta) + kx_2 \sin(\theta)]}, \quad (8.3)$$

where  $\hat{R}_s(\theta) = \hat{p}_r(x_1 = 0)/\hat{p}_i(x_1 = 0)$  is by definition reflection coefficient of the surface (since the surface is uniform, it is not a function of  $x_2$ , but it can depend on the angle of incidence). It quantifies both amplitude and phase change of the sound wave after the

reflection from the surface<sup>56</sup>. Complex sound pressure amplitude in front of the surface ( $x_1 < 0$ ) equals

$$\hat{p}(\mathbf{x}) = \hat{p}_i(\mathbf{x}) + \hat{p}_r(\mathbf{x}) = \hat{p}_Q e^{-j k x_2 \sin(\theta)} \left( e^{-j k x_1 \cos(\theta)} + \hat{R}_s(\theta) e^{j k x_1 \cos(\theta)} \right). \quad (8.4)$$

It is of interest to note that the same result would be obtained if the wall were replaced with a copy of the incident plane wave scaled with  $\hat{R}_s(\theta)$  and reaching the plane of the wall from the side opposite to the incident sound (as a copy in a mirror), with the angle  $\pi - \theta$  to the vector  $e_1$ . In fact, an infinite, flat, and reflecting wall can be equivalently replaced with an image source located behind its surface, at the straight line perpendicular to it which also includes the location of the original source, and at the same distance from the wall as the original source. This is used in the image source technique for calculations of sound fields in rooms as an alternative or (more frequently) in combination with ray tracing. Image source approximation holds even if the incident wave is not plane, for example, when the source of sound is located close to the wall. The original source can also be a lower-order image source from some other reflecting surface of the room.

In the special case of motionless fully reflecting rigid wall, which reflects both specularly and without energy losses,  $\hat{R}_s = 1$  and the pressure amplitude becomes

$$\hat{p}(\mathbf{x}) = \hat{p}_Q e^{-j k x_2 \sin(\theta)} \left( e^{-j k x_1 \cos(\theta)} + e^{j k x_1 \cos(\theta)} \right) = 2\hat{p}_Q \cos(k x_1 \cos(\theta)) e^{-j k x_2 \sin(\theta)} \quad (8.5)$$

with the modulus

$$|\hat{p}(\mathbf{x})| = 2|\hat{p}_Q \cos(k x_1 \cos(\theta))|. \quad (8.6)$$

Modulus of the pressure amplitude has the maximum  $2|\hat{p}_Q|$  at the planes parallel to the surface of the rigid wall given by  $x_1 = -n\pi/(k \cos(\theta)) = -n\lambda/(2 \cos(\theta))$ , with  $n = 0, 1, 2, \dots$ . From eq. (4.39), time-averaged energy equals

$$\langle E \rangle_T = \frac{|\hat{p}|^2}{2\rho_0 c_0^2} = \frac{2|\hat{p}_Q|^2}{\rho_0 c_0^2} \cos^2(k x_1 \cos(\theta)) = \frac{2|\hat{p}_Q|^2}{\rho_0 c_0^2} \frac{1 + \cos(2k x_1 \cos(\theta))}{2}. \quad (8.7)$$

In a diffuse field, we suppose that many plane waves with equal energy reach the wall at all angles of incidence, that is  $|\hat{p}_{diff}|^2 = 4\pi|\hat{p}_Q|^2$  (the factor  $4\pi$  implies integration over the full solid angle, similarly as in the ray approximation in section 5.2, for example, in

---

<sup>56</sup>Similarly as with absorption coefficient, change of amplitude may be due to the transmission of sound through the surface.

eq. (5.7)). Therefore, total energy in front of the wall is

$$\begin{aligned}
\langle E_{diff} \rangle_T &= \frac{|\hat{p}_{diff}|^2}{2\pi\rho_0c_0^2} \int_0^{2\pi} \frac{1 + \cos(2kx_1 \cos(\theta))}{2} d\Omega \\
&= \frac{|\hat{p}_{diff}|^2}{2\pi\rho_0c_0^2} \frac{1}{2} \int_0^{\pi/2} \int_0^{2\pi} [1 + \cos(2kx_1 \cos(\theta))] \sin(\theta) d\phi d\theta \\
&= \frac{|\hat{p}_{diff}|^2}{2\rho_0c_0^2} \left( \int_0^{\pi/2} \sin(\theta) d\theta + \int_0^{\pi/2} \cos(2kx_1 \cos(\theta)) \sin(\theta) d\theta \right) \\
&= \frac{|\hat{p}_{diff}|^2}{2\rho_0c_0^2} \left( 1 + \frac{\sin(2kx_1)}{2kx_1} \right).
\end{aligned} \tag{8.8}$$

The second term in the brackets is the sinc function, which equals 1 for zero argument. Hence, at the surface of the wall ( $x_1 = 0$ ),  $\langle E_{diff} \rangle_T = |\hat{p}_{diff}|^2 / (\rho_0c_0^2)$ . This is exactly double the value of time-averaged energy in a diffuse field far from the wall, eq. (5.27). As the distance from the wall increases, the energy oscillates according to the sinc function and converges to the value  $\langle E_{diff} \rangle_T = |\hat{p}_{diff}|^2 / (2\rho_0c_0^2)$ , when the sinc function vanishes. This explains the increase of sound level close to the reflecting boundaries of a room.

According to eq. (4.46), boundary condition at a locally reacting surface can be expressed with its impedance. However, we can also define the impedance for any plane control surface parallel to the actual boundary surface. For the geometry considered here, it equals

$$Z(\mathbf{x}) = \frac{\hat{p}(\mathbf{x})}{\hat{\mathbf{v}}(\mathbf{x}) \cdot \mathbf{e}_1} = \frac{\hat{p}(\mathbf{x})}{\hat{v}_1(\mathbf{x})}. \tag{8.9}$$

From the conservation of momentum in eq. (4.27) multiplied with  $\mathbf{e}_1$  and divided with  $e^{j\omega t}$  and eq. (8.4), we find

$$\begin{aligned}
j\omega\rho_0\hat{\mathbf{v}} \cdot \mathbf{e}_1 &= j\omega\rho_0\hat{v}_1 = -\nabla_x \hat{p} \cdot \mathbf{e}_1 = -\frac{\partial \hat{p}}{\partial x_1} \\
&= -\hat{p}_Q e^{-j k x_2 \sin(\theta)} \left( -jk \cos(\theta) e^{-j k x_1 \cos(\theta)} + jk \cos(\theta) \hat{R}_s(\theta) e^{j k x_1 \cos(\theta)} \right) \\
&= jk \cos(\theta) \hat{p}_Q e^{-j k x_2 \sin(\theta)} \left( e^{-j k x_1 \cos(\theta)} - \hat{R}_s(\theta) e^{j k x_1 \cos(\theta)} \right).
\end{aligned} \tag{8.10}$$

Therefore, the impedance for arbitrary  $\mathbf{x}$  with  $x_1 < 0$  is

$$Z(\mathbf{x}) = \frac{\hat{p}(\mathbf{x})}{\hat{v}_1(\mathbf{x})} = \frac{Z_0 \left( e^{-j k x_1 \cos(\theta)} + \hat{R}_s(\theta) e^{j k x_1 \cos(\theta)} \right)}{\cos(\theta) \left( e^{-j k x_1 \cos(\theta)} - \hat{R}_s(\theta) e^{j k x_1 \cos(\theta)} \right)}, \tag{8.11}$$

where  $Z_0$  is given in eq. (4.45), and at the uniform locally reacting surface is

$$Z_s(\theta) = \frac{\hat{p}(x_1 = 0)}{\hat{v}_1(x_1 = 0)} = \frac{Z_0 \left( 1 + \hat{R}_s(\theta) \right)}{\cos(\theta) \left( 1 - \hat{R}_s(\theta) \right)}. \tag{8.12}$$

In particular, for a rigid wall  $\hat{R}_s = 1$  and

$$Z(\mathbf{x}) = \frac{Z_0 (e^{-j k x_1 \cos(\theta)} + e^{j k x_1 \cos(\theta)})}{\cos(\theta) (e^{-j k x_1 \cos(\theta)} - e^{j k x_1 \cos(\theta)})} = \frac{j Z_0 \cot(k x_1 \cos(\theta))}{\cos(\theta)}, \quad (8.13)$$

while the surface impedance  $Z_s$  is infinite (normal velocity component at the wall is zero for any pressure). Acoustically soft wall has  $\hat{R} = -1$  and  $Z_s = 0$  for any  $\theta \neq \pi/2$ . From eq. (8.13), we also see that for  $\theta < \pi/2$  the impedance is zero at the distances  $x_1 = -(2n+1)\pi/(2k \cos(\theta)) = -(2n+1)\lambda/(4 \cos(\theta))$  in front of the rigid wall. This implies that the particle velocity component  $v_1$  at such planes parallel to the wall is very high even for moderate values of sound pressure, unlike at the surface of the wall or the planes  $x_1 = -n\pi/(k \cos(\theta)) = -n\lambda/(2 \cos(\theta))$ , where  $v_1 = 0$ . For the important special case of normal incidence,  $\theta = 0$ , and small distance from the wall,  $|kx_1| \ll 1 \Rightarrow \cot(kx_1) \approx 1/(kx_1)$ ,

$$Z(\mathbf{x}) = \frac{j \rho_0 c_0}{k x_1} = \frac{j \rho_0 c_0^2}{\omega x_1}. \quad (8.14)$$

Comparing this with the mechanical impedance in eq. (3.68), we see that acoustically thin layer of air (with thickness  $d = -x_1$  and small Helmholtz number value  $kd \ll 1$ ) in front of a rigid motionless wall behaves analogously to a spring with stiffness  $S = \rho_0 c_0^2/d$  for normal incidence of the incoming plane sound wave. Notice that the mechanical analogy based on the lumped elements holds only for acoustically small length scales (thickness  $d$  in this case) and normal incidence (one-dimensional propagation).

When energy description is sufficient, reflection coefficient and impedance can be related to the absorption coefficient of the surface introduced in eq. (5.13). For example,

$$\alpha_s(\theta) = \frac{d \langle P_{s,ray,loss}(\theta, t) \rangle_T}{d \langle P_{s,ray}(\theta, t) \rangle_T} = \frac{|\hat{p}_i(\theta, x_1 = 0)|^2 - |\hat{p}_r(\theta, x_1 = 0)|^2}{|\hat{p}_i(\theta, x_1 = 0)|^2} = 1 - |\hat{R}_s(\theta)|^2. \quad (8.15)$$

Again, such absorption coefficient also involves transmitted sound energy. Es expected, it equals 0 for fully reflecting surfaces and 1 for fully absorbing surfaces. Furthermore, from eq. (8.12):

$$\hat{R}_s(\theta) = \frac{Z_s(\theta) \cos(\theta) - Z_0}{Z_s(\theta) \cos(\theta) + Z_0}. \quad (8.16)$$

Finally, since

$$|\hat{R}_s(\theta)|^2 = \frac{|Z_s(\theta) \cos(\theta) - Z_0|^2}{|Z_s(\theta) \cos(\theta) + Z_0|^2} = \frac{(\mathcal{R}_e(Z_s(\theta)) \cos(\theta) - Z_0)^2 + (\mathcal{I}_m(Z_s(\theta)) \cos(\theta))^2}{(\mathcal{R}_e(Z_s(\theta)) \cos(\theta) + Z_0)^2 + (\mathcal{I}_m(Z_s(\theta)) \cos(\theta))^2},$$

we can express absorption coefficient over impedance as

$$\begin{aligned}
 \alpha_s(\theta) &= 1 - \frac{(\mathcal{R}_e(Z_s(\theta)) \cos(\theta) - Z_0)^2 + (\mathcal{I}_m(Z_s(\theta)) \cos(\theta))^2}{(\mathcal{R}_e(Z_s(\theta)) \cos(\theta) + Z_0)^2 + (\mathcal{I}_m(Z_s(\theta)) \cos(\theta))^2} \\
 &= \frac{4\mathcal{R}_e(Z_s(\theta)) \cos(\theta) Z_0}{(\mathcal{R}_e(Z_s(\theta)) \cos(\theta) + Z_0)^2 + (\mathcal{I}_m(Z_s(\theta)) \cos(\theta))^2} \\
 &= \frac{4\mathcal{R}_e(Z_s(\theta)) \cos(\theta) Z_0}{\mathcal{R}_e(Z_s(\theta))^2 \cos^2(\theta) + 2\mathcal{R}_e(Z_s(\theta)) \cos(\theta) Z_0 + Z_0^2 + \mathcal{I}_m(Z_s(\theta))^2 \cos(\theta)^2} \\
 &= \frac{4\mathcal{R}_e(Z_s(\theta)) \cos(\theta) Z_0}{(|Z_s(\theta)| \cos(\theta))^2 + 2\mathcal{R}_e(Z_s(\theta)) \cos(\theta) Z_0 + Z_0^2}.
 \end{aligned} \tag{8.17}$$

In many practical cases, impedance of a locally reacting surface is angularly independent. Equations (8.16) and (8.17) then become, respectively:

$$\hat{R}_s(\theta) = \frac{Z_s \cos(\theta) - Z_0}{Z_s \cos(\theta) + Z_0} \tag{8.18}$$

and

$$\boxed{\alpha_s(\theta) = \frac{4\mathcal{R}_e(Z_s) \cos(\theta) Z_0}{(|Z_s| \cos(\theta))^2 + 2\mathcal{R}_e(Z_s) \cos(\theta) Z_0 + Z_0^2}}. \tag{8.19}$$

Note that the reflection and absorption coefficients are still functions of the angle of incidence. Figure 11 shows the value of absorption coefficient as a function of real and imaginary part of the normalized impedance  $Z_s/Z_0$  for two angles of incidence. Only passive surfaces (with  $\mathcal{R}_e(Z_s) \geq 0$ , as followed from eq. (4.51)) are considered. Total absorption with  $\alpha_s = 1$  is achievable for  $Z_s = Z_0$  and normal incidence. Impedance of the wall is then fully matched with the impedance of air in front of it. However, for a different angle of incidence, surface impedance which gives maximum absorption is different (to around  $1.4Z_0$  for  $\theta = \pi/4$ ). Hence, the impedance boundary condition  $Z_s = \rho_0 c_0$  does not guarantee full absorption ( $\alpha_s = 1$ ) for general angles of incidence. In the case of grazing incidence ( $\theta = \pi/2$ ), no reflection or absorption take place. In reality, however, if the surface is absorbing, sound wave refracts into it and part of the sound energy is absorbed. An example of this is sound attenuation due to propagation over audience, which was discussed earlier. The amplitude can under such circumstances decay faster than 6dB per doubling the distance. In the opposite case, when the angle of incidence approaches 0, the surface is expected to react more locally and angularly independent, which additionally justifies equations (8.18) and (8.19) even for non-locally reacting materials.

## 8.2 Diffuse surface

If a wall is not flat (or non-uniform, for example, with spatially dependant material properties), scattering of sound in non-specular directions is expected to occur at wavelengths which are comparable to or smaller than the characteristic length scale of the

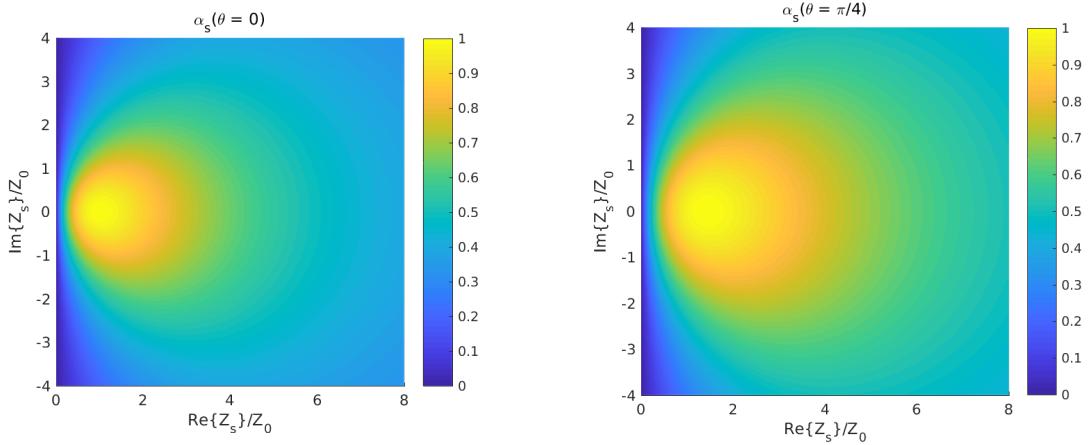


Figure 11: Absorption coefficient according to eq. (8.19) as a function of the real and imaginary parts of normalized impedance for (left) normal incidence and (right) angle of incidence  $\theta = \pi/4$ .

geometrical irregularities<sup>57</sup>. This is expressed with scattering coefficient  $s_s(\theta_i)$  introduced in eq. (7.1), where  $\theta_i$  is the angle of incidence. Energy of the specularly reflected sound,  $E_{spec}$ , can be estimated by multiplying the reflected energy,  $E_r$ , with the factor  $1 - s_s(\theta_i)$ :

$$E_{spec} = E_r - E_{scatt} = E_r(1 - s_s(\theta_i)), \quad (8.20)$$

where  $E_{scatt} = s_s(\theta_i)E_r$  is total non-specularly reflected energy. A model which is frequently used for angular dependence of the scattered energy is Lambert's cosine law:

$$E_{scatt}(\theta_r) = s_s(\theta_i)E_r \frac{\cos(\theta_r)}{\pi}, \quad (8.21)$$

where  $\theta_r$  is angle of the scattered sound wave to the normal of the surface. The factor of  $1/\pi$  is necessary for the equality

$$\begin{aligned} \int_0^{2\pi} E_{scatt}(\theta_r) d\Omega_r &= \int_0^{\pi/2} \int_0^{2\pi} s_s(\theta_i) E_r \frac{\cos(\theta_r)}{\pi} \sin(\theta_r) d\phi_r d\theta_r \\ &= 2s_s(\theta_i) E_r \int_0^{\pi/2} \cos(\theta_r) \sin(\theta_r) d\theta_r = s_s(\theta_i) E_r. \end{aligned} \quad (8.22)$$

The value of  $E_{scatt}(\theta_r)/E_r$  as a function of angle  $\theta_r$  and scattering coefficient  $s_s(\theta_i)$  is shown<sup>58</sup> in Fig. 12.

<sup>57</sup>If the characteristic length scale of the deviations from the flat surface is much larger than the wavelength, the surface may behave much like an inclined wall, with a similar treatment as above, only with adjusted coordinate system.

<sup>58</sup>For the graphical representation we let  $-\pi/2 \leq \theta_r \leq \pi/2$ . Values for the negative angles are equal to the values for the corresponding positive angles and the graph can be rotated around the axis  $\theta_r = 0$  ( $E_{scatt}$  does not depend on  $\phi_r$ ).

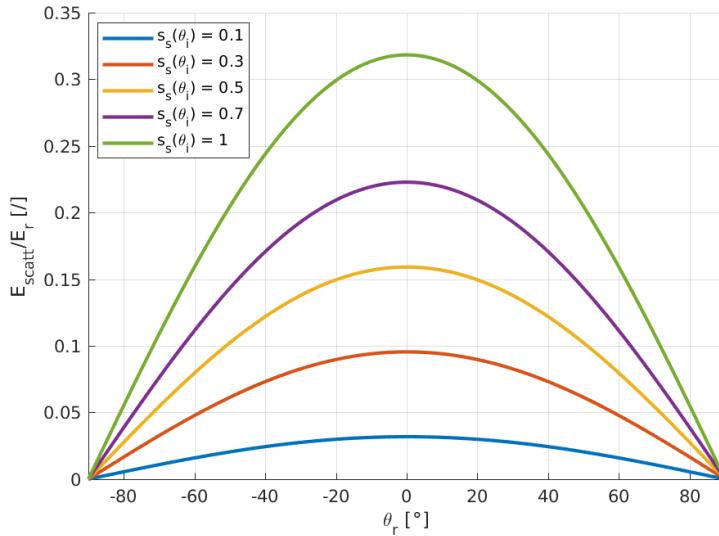


Figure 12: Ratio of scattered and specularly reflected energy according to Lambert's cosine law as a function of angle of reflection and for different values of scattering coefficient.

### 8.3 A layer of porous absorber

In section 8.1 we described an infinite plane surface with impedance, reflection coefficient, and absorption coefficient. Their values can be expressed simply for certain idealized surfaces, such as acoustically hard and soft walls. On the other hand, porous materials (and many other) have a complicated structure which makes the exact analytical treatment difficult. Values of the relevant acoustic quantities can only be estimated, for example, by using simplified models or empirical equations, or measured. The basis for modelling porous absorbers was, in fact, already introduced in eq. (3.38). By adding the imaginary part  $-j\zeta/c$  to the real wave number  $k$  of a forward propagating wave, where  $\zeta$  is real and positive damping constant and  $c$  is speed of sound in the porous material which is in general different from  $c_0$ , exponential decay of sound pressure amplitude inside the material can be modelled.

The required damping constant should be estimated from the physical properties of the material, such as its geometry, inner structure, etc. A very simple and crude method is Rayleigh model, which treats a porous material as consisting of many mutually parallel narrow channels filled with air and separated by rigid solid walls (solid structure of the material). The only required property of the material is its specific flow resistance (flow resistance per unit length, also called flow resistivity),  $\Xi'$ . By definition, it represents

the ratio of pressure drop and velocity of the flow through the narrow channels<sup>59</sup>:

$$\Xi' = -\frac{\partial p / \partial x_1}{v_1}, \quad (8.23)$$

where we suppose that the flow and the channels are both parallel to the  $x_1$ -axis. The unit is  $\text{kg}/(\text{m}^3\text{s}) = \text{Pa s}/\text{m}^2 = \text{Rayl}/\text{m}$ .

Flow resistivity is usually measured on a sample of the material with certain porosity  $\sigma$ , which is fraction of the volume of channels (air) from the total volume of the material. Flow resistivity is accordingly

$$\Xi' = \Xi\sigma, \quad (8.24)$$

where  $\Xi$  denotes such measured flow resistivity. Table 8.1 indicates typical ranges of values of density and flow resistivity of various porous materials. Flow resistivity of materials such as mineral wool, which have porosity  $\sigma \approx 1$  (hence,  $\Xi' \approx \Xi$ ), is commonly between  $5000 \text{ kg}/(\text{m}^3\text{s})$  and  $100000 \text{ kg}/(\text{m}^3\text{s})$ . The value can depend on the actual direction of the flow (angle of incidence in the case of an incoming sound wave). However, we will assume that the structure of the material is stochastic enough so that it is essentially isotropic. If flow resistivity is measured on a sample with thickness  $d$  as  $\Xi = \Delta p/(v_1 d)$ , where the pressure drop  $\Delta p$  divided with  $d$  approximates the negative gradient in eq. (8.23), flow resistance in  $\text{kg}/(\text{m}^2\text{s})$  is calculated as  $\Xi'd = \sigma\Xi d = \sigma\Delta p/v_1$ . Therefore, for normal incidence of an incoming sound wave and  $kd \ll 1$  ( $k$  is wave number), flow resistance quantifies pressure drop of the wave propagating through the porous material, which thus behaves as a lumped resistance  $\Delta p/v_1 = \Xi'd/\sigma$ .

Table 8.1: Density and flow resistivity of porous materials.

| material           | $\rho$ [ $\text{kg}/\text{m}^3$ ] | $\Xi \cdot 10^3 \text{ kg}/(\text{m}^3\text{s})$ |
|--------------------|-----------------------------------|--|
| mineral/glass wool | 15-50                             | 5-15   |
| mineral/glass wool | 50-100                            | 15-40  |
| mineral/glass wool | 100-200                           | 40-80  |
| wood wool          | 350-500                           | 0.5-2  |
| cotton             | 25-100                            | 10-150   |

As implied in section 2, momentum equation 2.9 has to be modified in order to include acoustic effects of porous materials. By subtracting the pressure gradient term from eq. (8.23), we can obtain a one-dimensional form (assuming an isotropic material) of the conservation of momentum inside the narrow channels:

$$\rho_0 \frac{\partial v_1}{\partial t} + \frac{\partial p}{\partial x_1} + \Xi' v_1 = 0. \quad (8.25)$$

<sup>59</sup>The value of flow resistivity is practically the same for mean and unsteady flow, as long as the fluid dynamics inside the channels is dominated by the viscous effects. This means that the width of the narrow channels formed by the solid skeleton of the material should not be larger than the boundary layer thickness given in eq. (2.12).

In frequency domain (linearized theory) this gives

$$j\omega\rho_0\hat{v}_1 + \frac{\partial\hat{p}}{\partial x_1} + \Xi'\hat{v}_1 = (j\omega\rho_0 + \Xi')\hat{v}_1 + \frac{\partial\hat{p}}{\partial x_1} = 0. \quad (8.26)$$

In this way we model the attenuation inside the boundary layer at the supposedly rigid solid structure of the porous material. For simplicity, we neglect energy losses due to heat exchange between the fluid and the structure.

We can now observe a plane wave inside the porous material with the complex wave number

$$k_{porous} = \frac{\omega}{c} - j\frac{\zeta}{c}, \quad (8.27)$$

where  $c$  is speed of sound in the material. Similarly as in eq. (8.2), complex amplitude of the sound wave can be written as

$$\hat{p}(x_1) = \hat{p}_Q e^{-jk_{porous}x_1} = \hat{p}_Q e^{-\zeta x_1/c} e^{-j\omega x_1/c}. \quad (8.28)$$

In order to express  $k_{porous}$  in terms of  $\Xi'$ , we analyse the perturbations in fluid inside the material (air with the steady density  $\rho_0$  and speed of sound  $c_0$ ) and refer to the conservation laws. After inserting

$$\frac{\partial\hat{p}}{\partial x_1} = -jk_{porous}\hat{p}.$$

into eq. (8.26), we obtain

$$(j\omega\rho_0 + \Xi')\hat{v}_1 - jk_{porous}\hat{p} = 0. \quad (8.29)$$

The second relation between  $\hat{p}$  and  $\hat{v}_1$  follows from the conservation of mass, eq. (2.8), and the equation of state (2.17), which in one-dimensional form and frequency domain give

$$\frac{j\omega}{c_0^2}\hat{p} + \rho_0\frac{\partial\hat{v}_1}{\partial x_1} = \frac{j\omega}{c_0^2}\hat{p} - jk_{porous}\rho_0\hat{v}_1 = 0. \quad (8.30)$$

We can multiply this equation with  $k_{porous}c_0^2/\omega$  and use the result in eq. (8.29) to obtain

$$(j\omega\rho_0 + \Xi')\hat{v}_1 - \frac{jk_{porous}^2\rho_0 c_0^2}{\omega}\hat{v}_1 = 0 \quad (8.31)$$

and therefore

$$k_{porous}^2 = \frac{\omega^2}{c_0^2} \left( 1 - j\frac{\Xi'}{\rho_0\omega} \right). \quad (8.32)$$

Since  $\mathcal{R}_e(k_{porous}) = k = \omega/c_0 > 0$  (we assume that the plane wave propagates in the direction of  $x_1$ -axis),

$$k_{porous} = \frac{\omega}{c_0} \sqrt{1 - j\frac{\Xi'}{\rho_0\omega}} \quad (8.33)$$

and we obtained the necessary relation between the complex wave number and flow resistivity as a physical property of the material, which can be used for calculations of damping inside the porous material.

It should be noted that Rayleigh model is very crude. Real porous materials do not consist of straight narrow channels, but rather small open cavities with irregular shapes and sizes. They can also be non-isotropic. The solid parts are not perfectly rigid and they do conduct heat. In order to overcome such inaccuracies, other more sophisticated models (involving more physical properties of the materials) have been suggested in literature (Delany and Bazley, Miki...), many of them being empirical. In practice, absorption coefficient of porous absorbers is very often provided by the manufacturers, estimated from the data in literature for similar materials, or, when appropriate, measured in a reverberation chamber (diffuse field). In addition to this, actual performance of porous absorbers in a room depends also on the specific arrangement, distribution of the material (see also the comments at the end of section 5.3), and mounting, especially when shapes other than simple flat layers of the material with constant thickness are applied. One example are absorbing wedges inside anechoic chambers, the geometry of which allows gradual change of surface impedance (by gradually increasing the absorbing area) and thus further prevents possible reflections, at least for wavelengths which are comparable to the size of the wedges.

## 8.4 Porous absorber in front of a rigid motionless wall

Absorbing materials are typically placed in front of acoustically hard surfaces, which increases their efficiency. In the following we analyse absorption properties of the layer of porous material from the previous section with thickness  $d$  backed by a flat, rigid, and motionless wall. From eq. (8.13), impedance at the free surface (at  $x_1 = -d$ ) of such a layer mounted directly on the rigid wall (between the planes  $x_1 = -d$  and  $x_1 = 0$ ) equals

$$Z_{\text{porous},\text{wall}}(\mathbf{x}) = \frac{\hat{p}}{\sigma \hat{v}_1} = -\frac{j Z_0 \cot(k_{\text{porous}} d \cos(\theta))}{\sigma \cos(\theta)}, \quad (8.34)$$

where we included an additional factor  $\sigma$ , the perforation ratio, which is ratio of the surface area of all the openings in the material at its free surface and the total free surface area of the material. Since the material is assumed to be isotropic and statistically uniform, this ratio is constant over the entire surface of the material. Moreover, Rayleigh's model approximates the material with straight parallel channels which extend through the entire thickness of the material and have constant cross sections. Therefore, porosity, which we used in the previous section, and perforation ratio are equal, which is why we use the same symbol ( $\sigma$ ) for them. Including  $\sigma$  in eq. (8.34) follows from the conservation of mass. Since  $v_1 = 0$  at the solid surfaces of the material at  $x_1 = -d$  and the mass is conserved, velocity component parallel to the  $x_1$ -axis immediately in front of the material, which is to be used for the estimation of impedance, is equal to  $v_1$  just inside the openings of the material multiplied with the factor  $\sigma$ .

After inserting  $k_{porous}$  from eq. (8.33):

$$Z_{porous,wall}(\mathbf{x}) = -\frac{jZ_0}{\sigma \cos(\theta)} \cot\left(\frac{\omega}{c_0} \sqrt{1 - j \frac{\Xi'}{\rho_0 \omega}} d \cos(\theta)\right). \quad (8.35)$$

If the Helmholtz number  $kd = \omega d/c_0 \rightarrow 0$ , so that the cotangent can be expanded into series to give  $\cot(k_{porous}d \cos(\theta)) \approx 1/(k_{porous}d \cos(\theta))$  and thus  $Z(\mathbf{x}) \rightarrow \infty$ , the large impedance approaches that of a hard wall, which results in low absorption. Consequently, in order to achieve appreciable absorption, thickness of the porous layer must not be too small compared with the sound wavelength. For normal incidence,  $\theta = 0$  and

$$Z_{porous,wall}(\mathbf{x}) = -\frac{jZ_0}{\sigma} \cot(k_{porous}d) = -\frac{jZ_0}{\sigma} \cot\left(\frac{\omega}{c_0} \sqrt{1 - j \frac{\Xi'}{\rho_0 \omega}} d\right). \quad (8.36)$$

Thin porous absorbers mounted directly on a hard surface are not very efficient. The reason for this can be found in eq. (8.23), which indicates that for a fixed flow resistivity, pressure drop is proportional to the velocity inside the porous material. Therefore, porous absorbers are particularly efficient at the locations in the sound field where particle velocity is high. For example, from eq. (8.13) it followed that such locations are at the distances from a rigid wall which are odd multiples of  $\lambda/(4 \cos(\theta))$ . In contrast to this, normal velocity component is zero at hard surfaces and porous absorbers are inefficient there. This means that absorption in a relatively thin layer of porous absorber (still thick enough relative to the wavelength, say, at least  $d \approx \lambda/10$ ) can be increased by placing it certain distance  $d_0$  away from the wall. If low angles of incidence dominate, the highest efficiency is achieved at the distances around  $\lambda/4 = c_0/(4f)$ , where  $f$  is frequency of the sound. However, the optimal distance depends in general both on the angles of incidence and frequency of the incoming sound waves which should be absorbed. The resulting angularly dependent absorption coefficient can be estimated by inserting the impedance from eq. (8.35) into eq. (8.17) and the diffuse field value from eq. (5.18).

If we restrict ourselves to normal incidence and assume that the absorber is acoustically thin, that is  $kd \ll 1$ , we can treat it as a lumped resistance  $\Xi' d/\sigma$  (see section 8.3). The total impedance analogous to eq. (3.68) is then sum of the resistance and the impedance of the air gap with thickness  $d_0$  in front of the rigid wall and behind the porous layer, which was given in eq. (8.13). After inserting  $\theta = 0$  and  $d_0 = -x_1$ , we obtain

$$Z_{porous,wall} = \Xi' d/\sigma - jZ_0 \cot(kd_0). \quad (8.37)$$

We can now estimate absorption coefficient from eq. (8.19) with  $\theta = 0$ :

$$\begin{aligned} \alpha_{s,porous,wall} &= \frac{4\Xi' d Z_0 / \sigma}{\{( \Xi' d / \sigma )^2 + [Z_0 \cot(kd_0)]^2\} + 2(\Xi' d / \sigma) Z_0 + Z_0^2} \\ &= \frac{4\Xi' d Z_0 / \sigma}{( \Xi' d / \sigma + Z_0 )^2 + [Z_0 \cot(kd_0)]^2}. \end{aligned} \quad (8.38)$$

It approaches zero when  $\cot(kd_0) \rightarrow \pm\infty$ , which is for  $kd_0 = n\pi$  and  $n = 0, 1, 2, \dots$ , and reaches maxima for  $kd_0 = (2n + 1)\pi/2$ , when  $\cot(kd_0) = 0$ . Maximum absorption is  $\alpha_{s,porous,wall} = 1$  only when  $\Xi'd/\sigma = Z_0$ . When  $\Xi'd/\sigma$  is larger/smaller than  $Z_0$ , the absorption coefficient will drop with increasing/decreasing value of  $\Xi'd/\sigma$ . In practice, the variations between minima and maxima are less pronounced when averaging over frequency is performed, for example in octave bands. Furthermore, the dips are not so strong in a diffuse field, when different angles of incidence appear, so the frequency at which the cotangent is infinite varies for a fixed  $d_0$ .

If the distance between the absorber and the wall is acoustically small,  $kd_0 \ll 1$ , the air gap contributes as the lumped stiffness  $\rho_0 c_0^2/d_0$ , since  $\cot(kd_0) \approx 1/(kd_0)$ , as shown in section 8.1 (eq. (8.14)). The impedance and absorption coefficient become respectively:

$$Z_{porous,wall} = \Xi'd/\sigma - j \frac{\rho_0 c_0^2}{\omega d_0} \quad (8.39)$$

and

$$\alpha_{s,porous,wall} = \frac{4\Xi'dZ_0/\sigma}{(\Xi'd/\sigma + Z_0)^2 + [\rho_0 c_0^2/(\omega d_0)]^2}. \quad (8.40)$$

At sufficiently high frequencies, when  $\rho_0 c_0^2/(\omega d_0)$  is negligible compared to  $\Xi'd/\sigma + Z_0$ , maximum absorption is achieved, up to  $\alpha_{s,porous,wall} = 1$  when  $\Xi'd/\sigma = Z_0$ . Evidently, “sufficiently high” frequency will be lower for larger distances from the wall  $d_0$ . Figure 13 shows absorption coefficient values obtained with equations (8.38) and (8.40), and demonstrates the effects of varying flow resistivity, thickness of the porous material, and width of the air gap between the material and the wall.

From the analysis above, it follows that low frequencies and associated large wavelengths require relatively large distances between the porous layer and the wall, which might take considerable volume of the room. With regard to that, distances larger than 20-30 cm are rarely used in practice. In order to achieve a more compact absorber (with even thinner layer of porous material), local particle velocity in the porous material can be increased by means of acoustic resonators, at least for a certain range of frequencies close to the resonance. Especially compact are Helmholtz resonators, which will be covered in the next section.

## 8.5 Helmholtz resonator and perforated panels

We treat a Helmholtz resonator as an acoustically compact cavity with volume  $V_c$  and arbitrary shape, bounded with rigid walls. The only opening of the cavity extends into a narrow neck with the length  $l_n$  much smaller than the wavelength ( $kl_n \ll 1$ ) and comparable to or smaller than the characteristic length scale of the cavity. For convenience, we can place the straight neck along the horizontal right-oriented  $x$ -axis (instead of the usual  $x_1$ -axis), with the cavity located at the left end of the neck. Since the entire resonator is compact, sound waves do not propagate inside it and, in terms of the mechanical analogy, both cavity and neck can be treated as lumped elements.

Limiting ourselves to the plane sound waves parallel to the  $x$ -axis and one-dimensional sound field, we can consider only two distinct velocities in the resonator which are functions of time only – inside the cavity,  $v_c(t)$ , and inside the neck,  $v_n(t)$ .

The resonance behaviour of a Helmholtz resonator is due to the change of its geometry from the narrow neck<sup>60</sup> with cross section area  $S_n$  to the wider cavity with cross section area  $S_c$ . Since the two parts of the geometry are finite and compact, we can write (linearized and inviscid) mass and momentum equations (2.8) and (2.9) in the integral form. Since  $S_c \gg S_n$ , conservation of mass at the interface between the neck and the cavity<sup>61</sup> implies  $|v_c| \ll |v_n|$  and in the linear approximation:

$$\left| \left( \frac{\partial v}{\partial t} \right)_c \right| \ll \left| \left( \frac{\partial v}{\partial t} \right)_n \right| \text{ and } \left| \left( \frac{\partial v}{\partial x} \right)_c \right| \ll \left| \left( \frac{\partial v}{\partial x} \right)_n \right|, \quad (8.41)$$

where  $( )_c$  means inside the cavity and  $( )_n$  inside the neck. On the other hand, the momentum equation (2.9) implies:

$$\left| \left( \frac{\partial v}{\partial t} \right)_c \right| \ll \left| \left( \frac{\partial v}{\partial t} \right)_n \right| \Rightarrow \left| \left( \frac{\partial p}{\partial x} \right)_c \right| \ll \left| \left( \frac{\partial p}{\partial x} \right)_n \right|. \quad (8.42)$$

Hence, pressure in the bulk of the cavity can be considered to be spatially uniform in comparison to the neck. Furthermore, the acoustically compact volumes  $V_n \ll V_c$  imply that much smaller neck behaves like a plug of incompressible air relative to the cavity, so we can suppose that the density perturbation satisfies  $|\rho_n| \ll |\rho_c|$  and therefore  $|(\partial \rho / \partial t)_n| \ll |(\partial \rho / \partial t)_c|$ . With this we can finally write the conservation of mass,

$$\begin{aligned} \int_V \frac{\partial \rho}{\partial t} d^3x + \rho_0 \int_V \frac{\partial v}{\partial x} d^3x &\approx \int_{V_c} \frac{\partial \rho}{\partial t} d^3x + \rho_0 \int_{V_n} \frac{\partial v}{\partial x} S_n dx \\ &= V_c \frac{\partial \rho_c}{\partial t} + \rho_0 S_n (v_n - v_c) \approx V_c \frac{\partial \rho_c}{\partial t} + \rho_0 S_n v_n = 0, \end{aligned} \quad (8.43)$$

and the conservation of momentum,

$$\begin{aligned} \rho_0 \int_V \frac{\partial v}{\partial t} d^3x + \int_V \frac{\partial p}{\partial x} d^3x &\approx \rho_0 \int_{V_n} \frac{\partial v}{\partial t} d^3x + \int_{V_n} \frac{\partial p}{\partial x} d^3x \\ &= \rho_0 V_n \frac{\partial v_n}{\partial t} + S_n (p_n - p_c) = 0, \end{aligned} \quad (8.44)$$

where  $v_n$  and  $p_n$  are also velocity and pressure at the opening of the neck. Compressibility of practically motionless volume of air in the cavity is balanced by the motion of essentially incompressible plug of air in the neck. The latter is determined by the difference of pressure at the opening of the neck and inside the cavity.

<sup>60</sup>Although narrow, the neck is assumed to be wider than the thickness of the boundary layer. We neglect viscous and thermal effects on the walls of the neck.

<sup>61</sup>We centre a very thin control volume at the interface with all the surfaces coinciding with the rigid boundaries of the resonator except the two flat surfaces normal to the  $x$ -axis. The volume flux at the left (cavity) side of the control volume,  $S_c v_c$ , has to be equal to the flux at the right (neck) side,  $S_n v_n$ .

After replacing  $\rho_c = p_c/c_0^2$  from eq. (2.17) as well as the ansatz  $p = \hat{p}e^{j\omega t}$  and  $v = \hat{v}e^{j\omega t}$ , we obtain two equations for the complex amplitudes:

$$j\omega V_c \frac{\hat{p}_c}{c_0^2} + \rho_0 S_n \hat{v}_n = 0 \quad (8.45)$$

and

$$j\omega \rho_0 V_n \hat{v}_n + S_n (\hat{p}_n - \hat{p}_c) = 0. \quad (8.46)$$

We can now express  $\hat{v}_n$  from eq. (8.45) in terms of  $\hat{p}_c$  and insert it in eq. (8.46), which gives

$$\frac{\hat{p}_n}{\hat{p}_c} = 1 - \frac{\omega^2 V_n V_c}{c_0^2 S_n^2}. \quad (8.47)$$

At angular frequencies close to

$$\omega_0 = c_0 S_n \sqrt{\frac{1}{V_n V_c}}, \quad (8.48)$$

sound pressure inside the cavity becomes very large in comparison to the pressure at the opening of the resonator. This is the resonance frequency. In reality, the maximum amplitude is always limited by viscous dissipation or non-linear damping (at high sound levels), which are both omitted from the analysis here.

Geometric volume of the neck equals  $S_n l_n$ . However, it turns out that the effective length of the neck is somewhat larger than  $l_n$ , due to the inertia of air at the ends of the neck, which increases the effective volume  $V_n$  of the air plug associated with the neck. Therefore, we express the effective length of the neck as

$$l_{n,eff} = \frac{V_n}{S_n} = l_n + 2\delta, \quad (8.49)$$

where  $\delta$  is end correction of the length (at one end of the neck, which explains the factor of 2 in eq. (8.49)). It generally depends on the geometry of the neck and for a circular perforation in a rigid baffle it is approximately  $0.85a$ , where  $a$  is radius of the perforation. It is of the same order of magnitude as  $a$ , which makes the end correction especially important for thin perforated screens, the thickness of which can be much smaller than  $a$ . The resonance frequency becomes

$$f_0 = \frac{c_0}{2\pi} \sqrt{\frac{S_n}{l_{n,eff} V_c}} = \frac{c_0}{2\pi} \sqrt{\frac{\sigma S_c}{l_{n,eff} V_c}} = \frac{c_0}{2\pi} \sqrt{\frac{\sigma}{l_{n,eff} d}}, \quad (8.50)$$

where  $d = V_c/S_c$  is characteristic length of the cavity and  $\sigma = S_n/S_c$ .

A common implementation of Helmholtz resonators in room acoustics is in absorbing perforated panels. A perforated screen backed by a rigid wall can be modelled as a raster of Helmholtz resonators. We will assume that all perforations are equal and equidistant. In such a case,  $\sigma$  represents the perforation ratio of the panel, which should be much

smaller than 1, according to the assumption  $S_c \gg S_n$ . Impedance at the outer surface of the perforated panel can be estimated in the following way. First, we express  $\hat{p}_c$  from eq. (8.45) and insert it in eq. (8.46), which gives

$$\frac{\hat{p}_n}{\hat{v}_n} = -\frac{j\omega\rho_0V_n}{S_n} - \frac{\rho_0S_nc_0^2}{j\omega V_c} = -j\omega\rho_0l_{n,eff} + \frac{j\rho_0\omega_0^2l_{n,eff}}{\omega} = j\rho_0l_{n,eff}\left(\frac{\omega_0^2}{\omega} - \omega\right). \quad (8.51)$$

In the definition of impedance, eq. (4.46), the unit vector  $\mathbf{n}$  is pointing into the surface. Therefore,  $\hat{\mathbf{v}}_i \cdot \mathbf{n}$  equals here  $-\hat{v}_i$  (the cavity is left from the neck) and  $\hat{\mathbf{v}}_i$  is velocity due to the incident sound wave. Furthermore, conservation of mass at the outer surface of the screen (interface between the perforation and the room) gives  $v_n = v_i/\sigma$  (compare with footnote <sup>61</sup>), while  $\hat{p}_n = \hat{p}_i$ . Hence, the impedance equals

$$Z_{perf} = -\frac{\hat{p}_i}{\hat{v}_i} = -\frac{\hat{p}_n}{\sigma\hat{v}_n} = \frac{j\rho_0l_{n,eff}}{\sigma}\left(\omega - \frac{\omega_0^2}{\omega}\right). \quad (8.52)$$

By comparing this with the mechanical impedance in eq. (3.68), we see that the backed perforated screen behaves as a combination of two lumped elements – mass  $\rho_0l_{n,eff}/\sigma$ , which is the air in the neck, and stiffness  $\rho_0l_{n,eff}\omega_0^2/\sigma = \rho_0c_0^2S_n/(\sigma V_c) = \rho_0c_0^2S_c/V_c = \rho_0c_0^2/d$ . This stiffness is equal to the stiffness of a thin layer of air for normal sound incidence, which we derived in section 8.1 (eq. (8.14)) and it is due to the air in the cavity of the resonator.

In order to introduce absorption, porous material is usually placed inside the cavity, close to the neck. From section 8.3 (see eq. (8.24) and the discussion below it) we know that a thin layer of porous material acts as a lumped resistance  $\Xi'd_{porous}/\sigma_{porous}$ , with  $\Xi'$  flow resistivity,  $\sigma_{porous}$  porosity, and  $d_{porous} = d - d_0$  thickness of the porous material ( $d_0$  is thickness of the remaining air gap inside the cavity). As a result, the impedance of the entire assembly equals

$$Z_{perf,absorb} = \frac{j\rho_0l_{n,eff}}{\sigma}\left(\omega - \frac{\omega_0^2}{\omega}\right) + \frac{\Xi'd_{porous}}{\sigma_{porous}}. \quad (8.53)$$

Absorption coefficient can be calculated from eq. (8.19):

$$\begin{aligned} \alpha_{s,perf,absorb} &= \frac{4Z_0\Xi'd_{porous}/\sigma_{porous}}{(\Xi'd_{porous}/\sigma_{porous} + Z_0)^2 + [\rho_0l_{n,eff}(\omega^2 - \omega_0^2)/(\omega\sigma)]^2} \\ &= \frac{4Z_0\Xi'd_{porous}/\sigma_{porous}}{(\Xi'd_{porous}/\sigma_{porous} + Z_0)^2 + [\rho_0c_0^2(\omega^2/\omega_0^2 - 1)/(\omega d)]^2}. \end{aligned} \quad (8.54)$$

It has the maximum value for  $\omega = \omega_0$  when it equals

$$\alpha_{s,perf,absorb} = \frac{4\Xi'd_{porous}/(\sigma_{porous}Z_0)}{[\Xi'd_{porous}/(\sigma_{porous}Z_0) + 1]^2}. \quad (8.55)$$

As before, maximum absorption  $\alpha_{s,perf,absorb} = 1$  (at the resonance frequency) is achieved when  $\Xi'd_{porous}/\sigma_{porous} = Z_0$ .

Figure 14 gives estimation of the absorption coefficient from eq. (8.54) for a panel with circular perforations and different values of the relevant parameters. The absorption increases with flow resistivity and thickness of the porous material. However, the presented values of  $\Xi' d_{\text{porous}} / \sigma_{\text{porous}}$  are all smaller than  $Z_0$ . For  $\Xi' d_{\text{porous}} / \sigma_{\text{porous}} > Z_0$ , absorption drops with further increase of  $\Xi'$  or  $d_{\text{porous}}$ . This can be compared with the last graph in Fig. 15 below, where the maximum absorption drops with the increase of damping  $D$  (which is there given larger than  $Z_0$ ). Besides that, resonance frequency is lower for a larger cavity, longer neck, larger radius of perforation (which leads to both larger effective length of the neck and width of the cavity, if perforation ratio is fixed), or lower perforation ratio (which leads to a larger width of the cavity, when radius of the perforation is fixed). This is in agreement with eq. (8.50). As pointed out in section 8.4, porous absorber is most efficient in the regions of high particle velocity. In Helmholtz resonators this is close to its neck, which is why the absorber is often applied as a thin layer of felt or fabric, glued on the back side of the perforated screen.

In the analysis above, we considered only normal incidence of the incoming sound. This is justified if the cavities of the resonators behind the perforations are physically separated with rigid walls, so that the sound cannot propagate parallel to the backing wall. This is usually not the case in practice. Still, if sound absorption inside the cavity is substantial, incident sound at larger angles of incidence will tend to refract normal to the wall. Indeed, measurements show that the impedance of perforated panels changes only slightly with the angle of incidence, at least for relatively low angles to the normal of the surface (up to around  $60^\circ$ ) and not too low frequencies (above 300Hz for typical geometries of the perforated panels).

## 8.6 Membrane resonator

In the previous section we considered how efficiency of a porous absorber backed by a rigid wall can be increased by placing a perforated screen in front of it. Addition of the mass of air plug inside the perforations to the stiffness of air between the screen and the wall sets the conditions for the resonance to occur, which increases magnitude of the particle velocity inside the porous material close to the perforation at the frequencies around the resonance. The physical mechanism behind membrane absorbers is essentially the same, except that the added mass is due to a thin solid material (membrane or a thin plate) placed in front of a rigid wall and close to an absorbing material. Damping can also be caused by the viscous losses inside the material of the membrane itself or its connections at the edges. With regard to that, we will use a general symbol  $D$  to denote the total damping regardless of the mechanism. If additional porous absorber is used, it should not be fixed directly to the membrane, so that they both move together. This would decrease the effective particle velocity inside the material and therefore its efficiency.

Pressure drop on the two sides of the solid membrane can be estimated from the con-

servation of momentum as

$$\Delta p = M_m \frac{dv_1}{dt}, \quad (8.56)$$

where  $M_m$  is surface mass of the membrane in  $\text{kg/m}^2$ . We again assumed normal incidence and one-dimensional dynamics. After switching to frequency domain,  $\Delta\hat{p}/\hat{v}_1 = j\omega M_m$  can be directly compared with eq. (3.68), confirming that an acoustically thin membrane acts as a lumped mass for normal incidence of sound. Impedance of the whole assembly is therefore

$$Z_{memb,wall} = j\omega M_m + D + \frac{\rho_0 c_0^2}{j\omega d}, \quad (8.57)$$

where the stiffness of air between the membrane and wall is, as before, equal to  $S = \rho_0 c_0^2/d$ . Neglecting the damping, resonance frequency is

$$f_0 = \frac{\omega_0}{2\pi} = \frac{\sqrt{S/M_m}}{2\pi} = \frac{c_0}{2\pi} \sqrt{\frac{\rho_0}{dM_m}}. \quad (8.58)$$

Absorption coefficient can also be estimated from eq. (8.19), with  $\theta = 0$ :

$$\begin{aligned} \alpha_{s,memb,wall} &= \frac{4DZ_0}{\{D^2 + [\omega M_m - \rho_0 c_0^2/(\omega d)]^2\} + 2DZ_0 + Z_0^2} \\ &= \frac{4DZ_0}{(D + Z_0)^2 + [\omega M_m - \omega_0^2 M_m/\omega]^2} \\ &= \frac{4DZ_0}{(D + Z_0)^2 + [(\omega^2 - \omega_0^2)M_m/\omega]^2}. \end{aligned} \quad (8.59)$$

It has the maximum value

$$\alpha_{s,memb,wall} = \frac{4D/Z_0}{[D/Z_0 + 1]^2} \quad (8.60)$$

at the resonance frequency ( $\omega = \omega_0$ ). For  $D = Z_0$ ,  $\alpha_{s,memb,wall} = 1$  at the resonance frequency. Absorption coefficient from eq. (8.59) is shown in Fig. 15. Increasing mass of the membrane or its distance from the wall lowers the resonance frequency, according to eq. (8.58). For  $D > Z_0$ , absorption coefficient decays with further increase of damping.

We can also estimate impedance of a thin membrane (without internal damping) not backed by a rigid wall as

$$Z_{memb} = j\omega M_m + Z_0, \quad (8.61)$$

where  $Z_0 = \rho_0 c_0$  is simply characteristic impedance of air behind the membrane. Absorption coefficient at normal incidence equals then (eq. (8.19)):

$$\alpha_{s,memb} = \frac{4Z_0^2}{4Z_0^2 + \omega^2 M_m^2}. \quad (8.62)$$

It decreases with the increase of membrane's surface mass and approaches 1 for a very light membrane or low frequencies, when the incident sound energy is transmitted through the membrane.

Unfortunately, modelling a membrane absorber with lumped elements gives very often unsatisfactory accuracy of the predicted values of both resonance frequency and sound absorption (more often than in the case of perforated panels). The main reason is that the membrane is not a locally reacting surface. It supports wave propagation along its surface (bending waves) and formation of its own eigenmodes. Furthermore, sound waves can propagate parallel to the wall behind the membrane, which is also neglected by the one-dimensional approximation for normal incidence. Adding to this, connections at the edges of the membrane can hinder its motion, while eq. (8.56) holds for a freely moving membrane. As a consequence, the connections can decrease the effective absorbing surface of the membrane absorber. In practice, overdamped membrane resonators with broader frequency characteristics carry less risk of the inaccuracies of the predictions.

By comparing Figures 13, 14, and 15, we can generally conclude that absorbers based only on layers of porous materials provide high absorption at middle to high frequencies. Their frequency range can be expended at low end most efficiently by increasing their distance from a hard wall. Absorbers based on the resonators are efficient in a limited frequency range around the resonance frequency, which is normally higher for perforated panels than for membrane absorbers. These absorbers are in general more expensive and difficult for prediction of performance and, particularly in the case of membrane absorbers, often provide lower values of absorption coefficient. Therefore, they are mostly used in the critical parts of the frequency range where porous absorbers are inefficient, which is at low to middle frequencies.

## 8.7 Quarter-wavelength resonator and seat dip effect

Another type of resonators which are occasionally used for sound absorbers is quarter-wavelength resonator. Its mechanism can be understood in terms of eq. (8.13) with  $\theta = 0$ , which is forced by rigid side-walls of the resonator. High values of particle velocity are achieved at the opening of the resonator, at the distance  $\lambda/4$  from the rigid bottom, and the resonance frequency can be estimated fairly accurately. However, the length of  $\lambda/4$  also implies much larger dimensions of the cavity than in the case of Helmholtz absorbers.

An interesting phenomenon associated with quarter-wavelength resonance is so-called seat dip effect. It is an occurrence of excessive attenuation of sound by empty seats of an auditorium, which is usually observed in a narrow frequency range between 100Hz and 200Hz (and to less extent at the odd multiples of the lowest frequency). The effect can be explained by the formation of vertical quarter-wavelength resonators between the adjacent rows of the seats. The damping is introduced by the upholstery of the seats. Less commonly, the effect can also be observed when the seats are occupied.

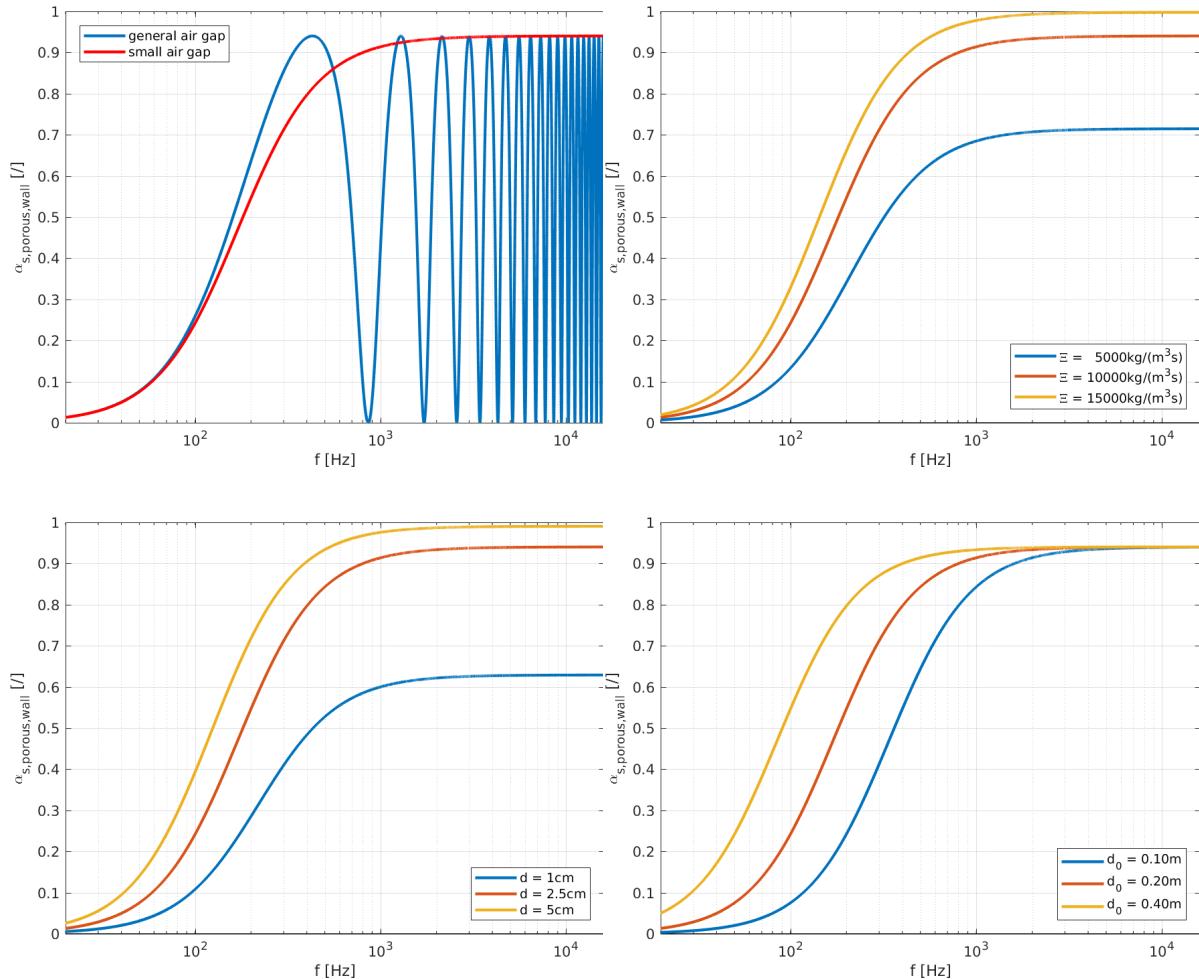


Figure 13: Absorption coefficient of a thin layer of porous absorber with porosity  $\sigma = 1$  ( $\Xi = \Xi'$ ) in front of a hard wall for normal incidence: (above left) calculated using eq. (8.38) and eq. (8.40) for  $\Xi = 10000\text{kg}/(\text{m}^3\text{s})$ ,  $d = 2.5\text{cm}$ ,  $d_0 = 0.20\text{m}$ ; calculated using eq. (8.40) for (above right)  $d = 2.5\text{cm}$ ,  $d_0 = 0.20\text{m}$ , (below left)  $\Xi = 10000\text{kg}/(\text{m}^3\text{s})$ ,  $d_0 = 0.20\text{m}$ , (below right)  $\Xi = 10000\text{kg}/(\text{m}^3\text{s})$ ,  $d = 2.5\text{cm}$ .

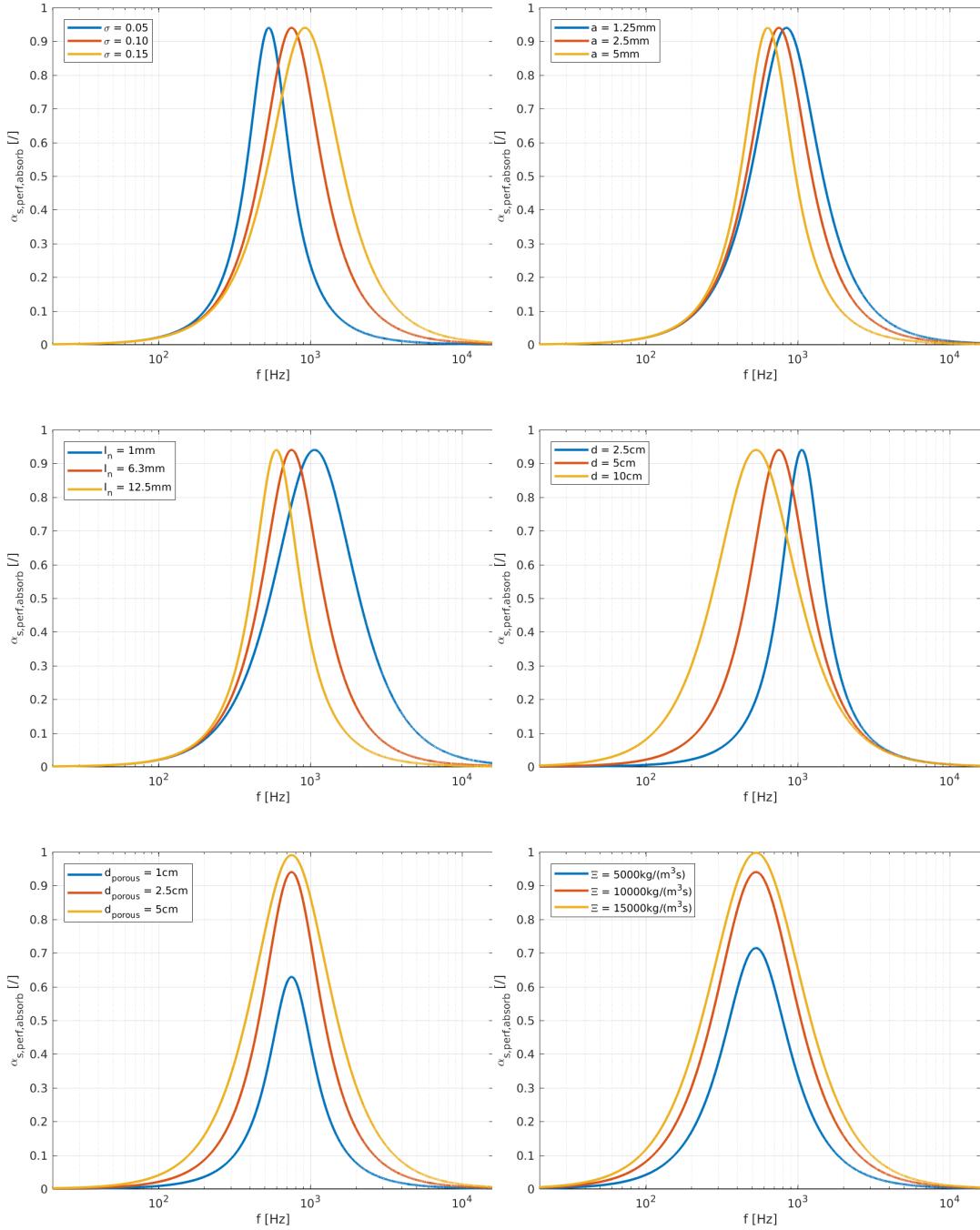


Figure 14: Absorption coefficient of a perforated panel with circular perforations backed by a layer of porous absorber ( $\sigma_{porous} = 1 \Rightarrow \Xi = \Xi'$ ) in front of a hard wall for normal incidence: (above left)  $a = 2.5\text{mm}$ ,  $l_n = 6.3\text{mm}$ ,  $d = 5\text{cm}$ ,  $d_{porous} = 2.5\text{cm}$ ,  $\Xi = 10000\text{kg}/(\text{m}^3\text{s})$ , (above right)  $\sigma = 0.1$ ,  $l_n = 6.3\text{mm}$ ,  $d = 5\text{cm}$ ,  $d_{porous} = 2.5\text{cm}$ ,  $\Xi = 10000\text{kg}/(\text{m}^3\text{s})$ , (middle left)  $\sigma = 0.1$ ,  $a = 2.5\text{mm}$ ,  $d = 5\text{cm}$ ,  $d_{porous} = 2.5\text{cm}$ ,  $\Xi = 10000\text{kg}/(\text{m}^3\text{s})$ , (middle right)  $\sigma = 0.1$ ,  $a = 2.5\text{mm}$ ,  $l_n = 6.3\text{mm}$ ,  $d_{porous} = 2.5\text{cm}$ ,  $\Xi = 10000\text{kg}/(\text{m}^3\text{s})$ , (below left)  $\sigma = 0.1$ ,  $a = 2.5\text{mm}$ ,  $l_n = 6.3\text{mm}$ ,  $d = 5\text{cm}$ ,  $\Xi = 10000\text{kg}/(\text{m}^3\text{s})$ , (below right)  $\sigma = 0.1$ ,  $a = 2.5\text{mm}$ ,  $l_n = 6.3\text{mm}$ ,  $d = 5\text{cm}$ ,  $d_{porous} = 2.5\text{cm}$ .

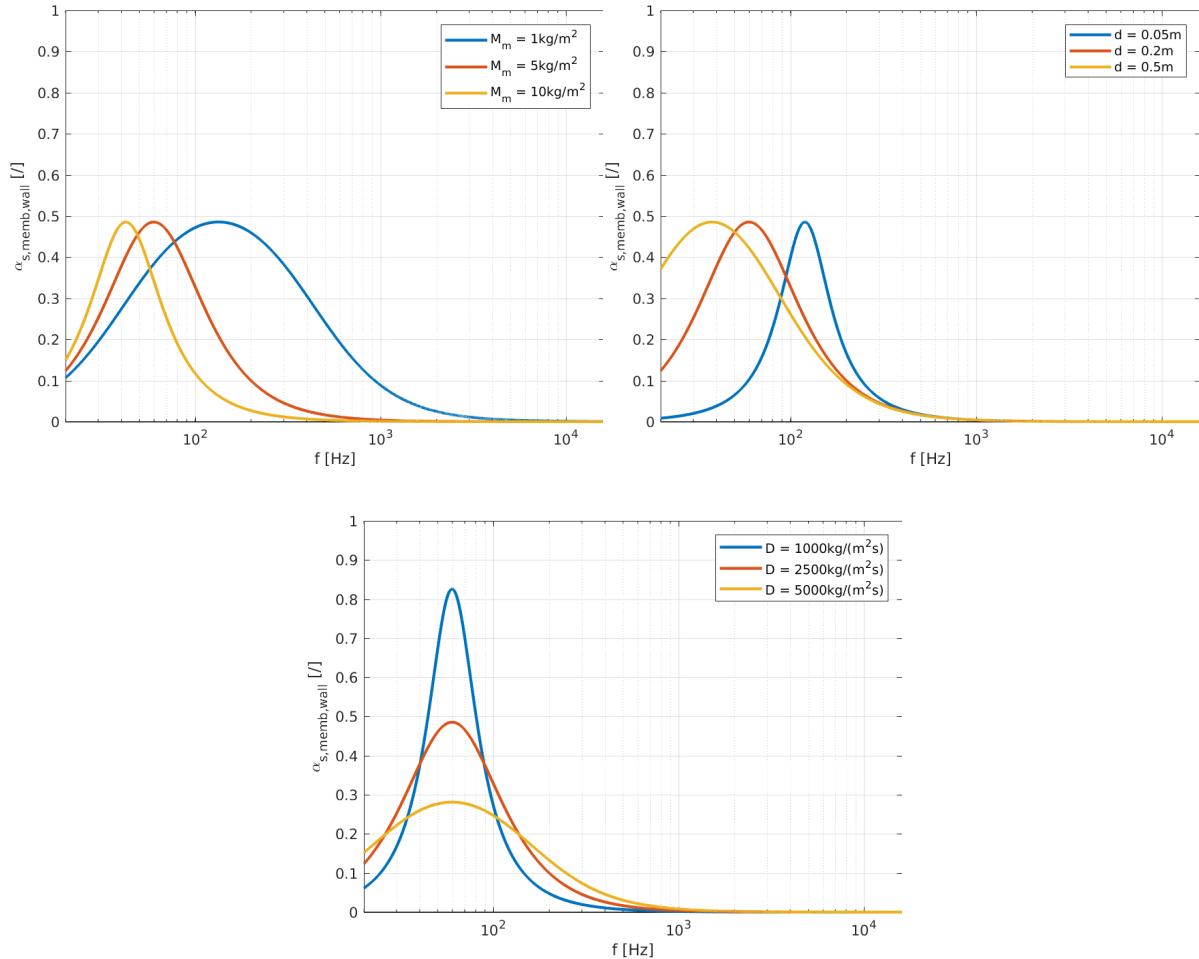


Figure 15: Absorption coefficient of a damped membrane in front of a hard wall for normal incidence: (above left)  $d = 0.2\text{m}$ ,  $D = 2500\text{kg}/(\text{m}^2\text{s})$ , (above right)  $M_m = 5\text{kg}/\text{m}^2$ ,  $D = 2500\text{kg}/(\text{m}^2\text{s})$ , (below)  $M_m = 5\text{kg}/\text{m}^2$ ,  $d = 0.2\text{m}$ .

## 9 Small elements

Room acoustics is often optimized by means of relatively small elements, with much smaller area than boundary surfaces of the room which were treated in the previous section. Such elements are reflectors and diffusers and they will be considered in the following.

### 9.1 Finite rectangular surface

Reflecting surfaces in a room often have dimensions which are not much larger than the sound wavelength, at least at relatively low frequencies. In such cases and in contrast to the analysis in section 8.1, less energy of the incident sound will be reflected specularly and greater part of it will diffract or scatter from the reflecting body at different angles. For this reason, rigid plates or more generally plane rectangular surfaces with finite dimensions, which we will consider here, are frequently used both as (specular) reflectors and diffusers. We will first estimate the sound reflected from a finite rectangular surface in the far field.

As before, we assume a plane incoming wave with pressure  $\hat{p}_i$  originating from a distant source of sound. The effect of the body is captured by eq. (3.46) in frequency domain. We replace  $\hat{p}(\mathbf{y})$  at the surface of the body with the total (incident plus reflected sound) complex pressure amplitude  $(1 + \hat{R}_s(\theta_i))\hat{p}_i(\mathbf{y})$  and  $\nabla_y \hat{p}(\mathbf{y})$  with  $-j\omega\rho_0\hat{\mathbf{v}}(\mathbf{y})$  from the conservation of momentum, eq. (4.27). The volume integral represents the free field component ( $\hat{G}$  is free space Green's function), which is the incident sound  $\hat{p}_i(\mathbf{x})$ , so the entire equation reads

$$\hat{p}(\mathbf{x}) = \hat{p}_i(\mathbf{x}) - \oint_S \{j\omega\rho_0\hat{G}\hat{\mathbf{v}}(\mathbf{y}) \cdot \mathbf{n} + (1 + \hat{R}_s(\theta_i))\hat{p}_i(\mathbf{y})\nabla_y \hat{G} \cdot \mathbf{n}\} d^2\mathbf{y}. \quad (9.1)$$

For convenience, we can choose the coordinates such that the plate lies in the plane  $x_3 = 0$  and centred at the origin, extending from  $x_1 = -L_1$  to  $x_1 = L_1$  and from  $x_2 = -L_2$  to  $x_2 = L_2$ . Moreover, due to the geometrical symmetry of the problem, we can suppose without a loss of generality that the incident sound reaches the plate from the octant<sup>62</sup>  $x_1, x_2, x_3 \geq 0$ , with the angle of incidence equal to the angle to the  $x_3$ -axis (inclination)  $0 \leq \theta_i \leq \pi/2$  and azimuthal angle of the projection of the wave path in the  $x_1x_2$ -plane to the  $x_1$ -axis  $0 \leq \phi_i \leq \pi/2$ . Only the irradiated side of the plate reflects the sound back into the half space  $x_3 \geq 0$ . Hence, we do not consider diffraction around the

<sup>62</sup>This is in a way the opposite direction of arrival of the incoming sound wave to the one in section 8.1. The incident wave travels “backwards” with  $k_1, k_2, k_3 \leq 0$ , which explains the opposite signs in the exponents in eq. (9.17) and eq. (8.2). Alternatively, we could set  $k_1, k_2, k_3 \geq 0$ , but the two angles must be then replaced as  $\theta_i \rightarrow \pi - \theta_i$  and  $\phi_i \rightarrow \phi_i + \pi$  with the same outcome. We adopt the first configuration in order to keep the angle of incidence  $\theta_i \leq \pi/2$ . Another difference compared with section 8.1 is that the plate lies in the plane  $x_3 = 0$  and not  $x_1 = 0$  and that we have to solve the problem in three dimensions.

plate and we can integrate in eq. (9.1) only over the irradiated plane rectangular surface  $S$  with area  $(2L_1)(2L_2) = 4L_1L_2$  and with  $\mathbf{n} = \mathbf{e}_3$  pointing outwards from it.

Total sound pressure in front of the plate equals

$$\hat{p}(\mathbf{x}) = \hat{p}_i(\mathbf{x}) - \int_S \left\{ j\omega\rho_0\hat{G}\hat{v}_3(\mathbf{y}) + (1 + \hat{R}_s(\theta_i))\hat{p}_i(\mathbf{y})\frac{\partial\hat{G}}{\partial y_3} \right\} d^2\mathbf{y}. \quad (9.2)$$

From the definition of impedance in eq. (4.46),

$$\begin{aligned} \hat{v}_3(\mathbf{y}) &= -\frac{\hat{p}(\mathbf{y})}{Z_s(\mathbf{y})} = -\frac{(1 + \hat{R}_s(\theta_i))\hat{p}_i(\mathbf{y})}{Z_s(\mathbf{y})} \\ &= -\frac{(1 + \hat{R}_s(\theta_i))\hat{p}_i(\mathbf{y})\cos(\theta_i)(1 - \hat{R}_s(\theta_i))}{Z_0(1 + \hat{R}_s(\theta_i))} \\ &= -\frac{\hat{p}_i(\mathbf{y})\cos(\theta_i)(1 - \hat{R}_s(\theta_i))}{\rho_0 c_0}, \end{aligned} \quad (9.3)$$

where eq. (8.12) was also used. Appearance of the minus sign in the first equality is because the surface normal  $\mathbf{n} = \mathbf{e}_3$  points outwards from the surface, while in the definition of impedance it points into the surface. Therefore, from eq. (9.2):

$$\hat{p}(\mathbf{x}) = \hat{p}_i(\mathbf{x}) + \int_S \left\{ jk\hat{G}\hat{p}_i(\mathbf{y})\cos(\theta_i)(1 - \hat{R}_s(\theta_i)) - (1 + \hat{R}_s(\theta_i))\hat{p}_i(\mathbf{y})\frac{\partial\hat{G}}{\partial y_3} \right\} d^2\mathbf{y}. \quad (9.4)$$

Free space Green's function is given in eq. (3.47):

$$\hat{G}(\mathbf{x}|\mathbf{y}) = \frac{e^{-jk|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} = \frac{e^{-jkr}}{4\pi r}, \quad (9.5)$$

where  $r = |\mathbf{x} - \mathbf{y}|$ . Although distinct parts of the surface radiate effectively into half space (being surrounded by the rest of the plate), we take free space Green's function because the reflection from the surface is taken into account by the factor  $(1 + \hat{R}_s)$  in eq. (9.1). Green's function depends on the location at the surface  $\mathbf{y}$ , which makes the integration over the surface  $S$  non-trivial. However, if we are interested in the sound field far from the plate, at distances much larger than both  $L_1$  and  $L_2$ , the denominator can be approximated as  $4\pi|\mathbf{x} - \mathbf{y}| \approx 4\pi|\mathbf{x}|$ , which is independent of  $\mathbf{y}$ . This is geometric far field,  $|\mathbf{x}| \gg |\mathbf{y}|$ . Unlike acoustic far field, in which  $|\mathbf{x}| \gg 1/k = \lambda/(2\pi)$  (and which will also be assumed below in eq. (9.15)), it is defined with respect to the dimensions of the body, not the sound wavelength.

Although the approximation  $|\mathbf{x} - \mathbf{y}| \approx |\mathbf{x}|$  holds for the negligible decay of amplitude with  $\mathbf{y}$  (compared to the decay with  $\mathbf{x}$ ), it cannot be applied for the exponent of Green's function  $-jk|\mathbf{x} - \mathbf{y}|$ , since phase of the wave can vary significantly with location at the surface  $\mathbf{y}$ , depending on its size and the frequency. Here we use the law of cosines instead, according to which

$$r = |\mathbf{x} - \mathbf{y}| = \sqrt{|\mathbf{x}|^2 - 2\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2} = |\mathbf{x}| \sqrt{1 - \frac{2\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}|^2} + \frac{|\mathbf{y}|^2}{|\mathbf{x}|^2}} \quad (9.6)$$

for  $|\mathbf{x}| \neq 0$ . Taylor expansion  $\sqrt{1-a} = 1 - a/2 + \mathcal{O}(a^2)$  for  $a \rightarrow 0$  gives

$$\begin{aligned} |\mathbf{x} - \mathbf{y}| &\approx |\mathbf{x}| \left[ 1 - \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}|^2} + \frac{|\mathbf{y}|^2}{2|\mathbf{x}|^2} + \mathcal{O}\left(\left(\frac{2\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}|^2} - \frac{|\mathbf{y}|^2}{|\mathbf{x}|^2}\right)^2\right) \right] \\ &= |\mathbf{x}| \left[ 1 - \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}|^2} + \mathcal{O}\left(\frac{|\mathbf{y}|^2}{|\mathbf{x}|^2}\right) \right], \end{aligned} \quad (9.7)$$

where we used  $|\mathbf{y}| \ll |\mathbf{x}|$ ,  $\mathbf{x} \cdot \mathbf{y} \ll |\mathbf{x}|^2$ , and  $\mathcal{O}(a^2)$  denotes order of  $a^2$  or higher. Therefore,

$$r = |\mathbf{x} - \mathbf{y}| \approx |\mathbf{x}| - \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}|}. \quad (9.8)$$

If we write  $\mathbf{y}$  in spherical coordinates as  $(r_y, \theta_y, \phi_y)$  with  $\theta_y = \pi/2$ , the scalar product becomes

$$\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}| |\mathbf{y}| \cos(\theta - \pi/2) \cos(\phi - \phi_y) = |\mathbf{x}| |\mathbf{y}| \sin(\theta) \cos(\phi - \phi_y) \quad (9.9)$$

and so

$$\begin{aligned} r &= |\mathbf{x} - \mathbf{y}| \approx |\mathbf{x}| - |\mathbf{y}| \sin(\theta) \cos(\phi - \phi_y) \\ &= |\mathbf{x}| - |\mathbf{y}| \sin(\theta) [\cos(\phi) \cos(\phi_y) + \sin(\phi) \sin(\phi_y)] \\ &= |\mathbf{x}| - \sin(\theta) [y_1 \cos(\phi) + y_2 \sin(\phi)]. \end{aligned} \quad (9.10)$$

We can insert this into eq. (9.5) to obtain an approximate expression for free space Green's function in the geometric far field:

$$\hat{G}_{far}(\mathbf{x} | \mathbf{y}) \approx \frac{1}{4\pi |\mathbf{x}|} e^{-jk(|\mathbf{x}| - y_1 \sin(\theta) \cos(\phi) - y_2 \sin(\theta) \sin(\phi))}. \quad (9.11)$$

In order to express the spatial derivative of Green's function appearing in eq. (9.4), we first notice that

$$\frac{\partial \hat{G}}{\partial y_3} = \frac{\partial \hat{G}}{\partial r} \frac{\partial r}{\partial y_3}, \quad (9.12)$$

where (compare with eq. (4.26))

$$\frac{\partial \hat{G}}{\partial r} = \frac{\partial}{\partial r} \frac{e^{-jkr}}{4\pi r} = \frac{1}{4\pi r} (-jk) e^{-jkr} + \frac{e^{-jkr}}{4\pi} \left( -\frac{1}{r^2} \right) = \frac{e^{-jkr}}{4\pi r} \left( -jk - \frac{1}{r} \right). \quad (9.13)$$

Since  $r = |\mathbf{x} - \mathbf{y}|$ ,

$$\begin{aligned} \frac{\partial r}{\partial y_3} &= \nabla_y r \cdot \mathbf{n} = -\nabla_x r \cdot \mathbf{n} = -\frac{\partial r}{\partial x_3} \\ &= -\frac{\partial}{\partial x_3} (|\mathbf{x}| - \sin(\theta) [y_1 \cos(\phi) + y_2 \sin(\phi)]) \\ &= -\frac{\partial}{\partial x_3} (x_1 \sin(\theta) \cos(\phi) + x_2 \sin(\theta) \sin(\phi) + x_3 \cos(\theta) \\ &\quad - \sin(\theta) [y_1 \cos(\phi) + y_2 \sin(\phi)]) = -\cos(\theta). \end{aligned} \quad (9.14)$$

In the acoustic far field  $r \gg 1/k$ . Therefore, we can neglect the second term in the last parentheses in eq. (9.13) in favour of the first one (as in eq. (4.35)). The derivative becomes

$$\begin{aligned} \frac{\partial \hat{G}}{\partial y_3} &= \frac{\partial \hat{G}}{\partial r} \frac{\partial r}{\partial y_3} = jk \cos(\theta) \frac{e^{-jkr}}{4\pi r} = jk \cos(\theta) \hat{G} \\ &\approx jk \cos(\theta) \frac{1}{4\pi |\mathbf{x}|} e^{-jk(|\mathbf{x}|-y_1 \sin(\theta) \cos(\phi)-y_2 \sin(\theta) \sin(\phi))}. \end{aligned} \quad (9.15)$$

Now we can insert the last equality into eq. (9.4) and obtain

$$\begin{aligned} \hat{p}(\mathbf{x}) &= \hat{p}_i(\mathbf{x}) + jk \int_S \hat{G} \hat{p}_i(\mathbf{y}) \{ \cos(\theta_i) (1 - \hat{R}_s(\theta_i)) - (1 + \hat{R}_s(\theta_i)) \cos(\theta) \} d^2 \mathbf{y} \\ &= \hat{p}_i(\mathbf{x}) - jk \int_S \hat{G} \hat{p}_i(\mathbf{y}) \{ \hat{R}_s(\theta_i) (\cos(\theta) + \cos(\theta_i)) + (\cos(\theta) - \cos(\theta_i)) \} d^2 \mathbf{y}. \end{aligned} \quad (9.16)$$

The incident plane sound wave can be written as (see also footnote <sup>62</sup>)

$$\begin{aligned} \hat{p}_i(\mathbf{x}) &= \hat{p}_Q e^{-jk_i \cdot \mathbf{x}} = \hat{p}_Q e^{-j(k_{i1}x_1 + k_{i2}x_2 + k_{i3}x_3)} \\ &= \hat{p}_Q e^{jk[x_1 \sin(\theta_i) \cos(\phi_i) + x_2 \sin(\theta_i) \sin(\phi_i) + x_3 \cos(\theta_i)]}, \end{aligned} \quad (9.17)$$

from which its value at the irradiated surface of the plate equals

$$\hat{p}_i(y_3 = 0) = \hat{p}_Q e^{jk[y_1 \sin(\theta_i) \cos(\phi_i) + y_2 \sin(\theta_i) \sin(\phi_i)]}. \quad (9.18)$$

According to eq. (9.16), angular distribution of the reflected sound amplitude depends on the variation of reflection coefficient at the surface  $S$ . If  $\hat{R}_s$  does not depend on the angle of incidence, the reflected sound pressure amplitude  $\hat{p} - \hat{p}_i$  equals from equations (9.16), (9.18), and (9.11)

$$\begin{aligned} \hat{p}_r(\mathbf{x}) &= -jk \int_S \hat{G} \hat{p}_i(\mathbf{y}) \{ \hat{R}_s(\mathbf{y}) (\cos(\theta) + \cos(\theta_i)) + (\cos(\theta) - \cos(\theta_i)) \} d^2 \mathbf{y} \\ &= -jk \int_{-L_1}^{L_1} \int_{-L_2}^{L_2} \hat{G} \hat{p}_i(y_3 = 0) \{ \hat{R}_s(y_1, y_2) (\cos(\theta) + \cos(\theta_i)) \\ &\quad + (\cos(\theta) - \cos(\theta_i)) \} dy_1 dy_2 = -\frac{jk \hat{p}_Q}{4\pi |\mathbf{x}|} e^{-jk|\mathbf{x}|} \\ &\quad \int_{-L_1}^{L_1} \{ \hat{R}_s(y_1) (\cos(\theta) + \cos(\theta_i)) + (\cos(\theta) - \cos(\theta_i)) \} e^{jAy_1} dy_1 \int_{-L_2}^{L_2} e^{jBy_2} dy_2. \end{aligned} \quad (9.19)$$

For simplicity, we restrict in the last equality  $\hat{R}_s(y_1)$  to be a function of  $y_1$  only. We also introduced  $A = k[\sin(\theta_i) \cos(\phi_i) + \sin(\theta) \cos(\phi)]$  and  $B = k[\sin(\theta_i) \sin(\phi_i) + \sin(\theta) \sin(\phi)]$  for brevity. We should note that the last two integrals represent two (or a single two-dimensional) inverse spatial Fourier transforms (as in eq. (3.22)), namely  $\mathcal{F}_{y_1}^{-1}\{\hat{R}_s(y_1)(\cos(\theta)+\cos(\theta_i))+(\cos(\theta)-\cos(\theta_i))\}$  and  $\mathcal{F}_{y_2}^{-1}\{1\}$ , with the parameters  $A$  and  $B$  instead of time and over the finite intervals  $(-L_1, L_1)$  and  $(-L_2, L_2)$ , respectively<sup>63</sup>.

<sup>63</sup>With positive  $k_1$ ,  $k_2$ , and  $k_3$ ,  $A$  and  $B$  would receive the opposite sign and the transforms could be compared directly to the Fourier transform in eq. (3.23) with spatial coordinates instead of time.

### 9.1.1 Rigid motionless plate as a reflector

If the plate is rigid and motionless ( $\hat{R}_s = 1$ ), the first term in the curly brackets in the first integral in eq. (9.16) vanishes, so the reflected sound equals

$$\hat{p}_r(\mathbf{x}) = -2jk \cos(\theta) \int_S \hat{G} \hat{p}_i(\mathbf{y}) d^2\mathbf{y} = -2jk \cos(\theta) \int_{-L_1}^{L_1} \int_{-L_2}^{L_2} \hat{G} \hat{p}_i(y_3 = 0) dy_1 dy_2, \quad (9.20)$$

or from eq. (9.19)

$$\hat{p}_r(\mathbf{x}) = -\frac{jk \cos(\theta) \hat{p}_Q}{2\pi |\mathbf{x}|} e^{-jk|\mathbf{x}|} \int_{-L_1}^{L_1} e^{jAy_1} dy_1 \int_{-L_2}^{L_2} e^{jBy_2} dy_2. \quad (9.21)$$

Both integrals in eq. (9.21) can be solved to give

$$\begin{aligned} \hat{p}_r(\mathbf{x}) &= -\frac{jk \cos(\theta) \hat{p}_Q}{2\pi |\mathbf{x}|} e^{-jk|\mathbf{x}|} \frac{e^{jAL_1} - e^{-jAL_1}}{jA} \frac{e^{jBL_2} - e^{-jBL_2}}{jB} \\ &= \frac{jk \cos(\theta) \hat{p}_Q}{2\pi |\mathbf{x}|} e^{-jk|\mathbf{x}|} \frac{2j \sin(AL_1)}{A} \frac{2j \sin(BL_2)}{B} \\ &= -\frac{2jk L_1 L_2 \cos(\theta) \hat{p}_Q}{\pi |\mathbf{x}|} e^{-jk|\mathbf{x}|} \frac{\sin(AL_1)}{AL_1} \frac{\sin(BL_2)}{BL_2}. \end{aligned} \quad (9.22)$$

The last two fractions are sinc functions (or spherical Bessel functions of the first kind and zero order,  $j_0$ ), with the form  $\text{sinc}(x) = j_0(x) = \sin(x)/x$ . Therefore, we can write

$$\begin{aligned} \hat{p}_r(\mathbf{x}) &= -\frac{2jk L_1 L_2 \cos(\theta) \hat{p}_Q}{\pi |\mathbf{x}|} e^{-jk|\mathbf{x}|} \text{sinc}(AL_1) \text{sinc}(BL_2) \\ &= -\frac{2jk L_1 L_2 \cos(\theta) \hat{p}_Q}{\pi |\mathbf{x}|} e^{-jk|\mathbf{x}|} \text{sinc}(k L_1 [\sin(\theta_i) \cos(\phi_i) + \sin(\theta) \cos(\phi)]) \\ &\quad \text{sinc}(k L_2 [\sin(\theta_i) \sin(\phi_i) + \sin(\theta) \sin(\phi)]). \end{aligned} \quad (9.23)$$

Normalized value of  $\hat{p}_r$  is shown in Fig. 16 for different angles of incidence and sizes of a rectangular plate (values of the Helmholtz number  $kL_1$ ).

Note that Helmholtz number appears in the forms  $kL_1$  and  $kL_2$  with respect to the dimensions of the plate. Its low values imply fewer lobes in a less directional scattering pattern (with less dominant specular reflection), which is determined by the sinc functions, as well as lower reflected energy in total, which is due to  $kL_1 L_2$  multiplying the sinc functions (and therefore not visible in Fig. 16 due to normalization). As expected, a compact plate is acoustically transparent (does not reflect). As the value of Helmholtz number increases, less energy of the incident sound is diffracted around the plate and more energy is reflected specularly. However, it should also be noted that eq. (9.23) does not converge to give a single specular reflection from an infinite hard wall when  $kL_1, kL_2 \rightarrow \infty$ , which should (from eq. (9.17) with  $k_{r3} = -k_{i3}$ ) have the form

$$\hat{p}_r(\mathbf{x}) = \hat{p}_Q e^{jk[x_1 \sin(\theta_i) \cos(\phi_i) + x_2 \sin(\theta_i) \sin(\phi_i) - x_3 \cos(\theta_i)]}. \quad (9.24)$$

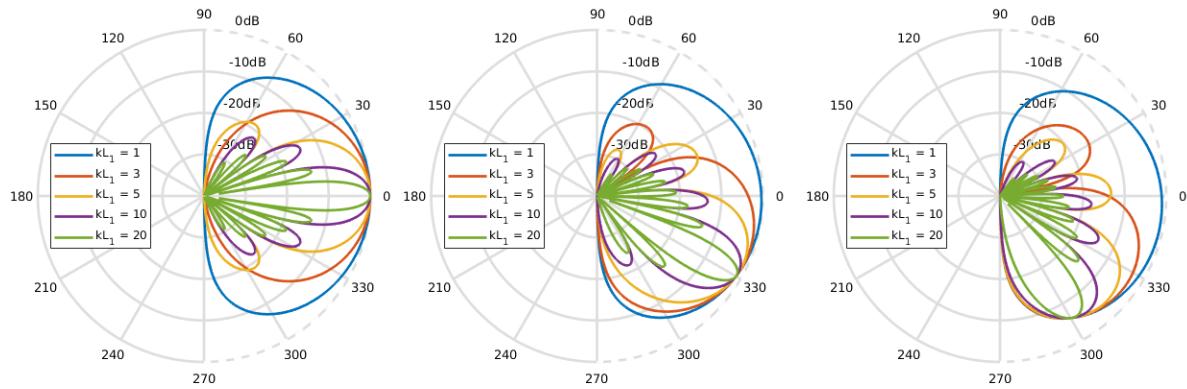


Figure 16: Scattering of a rigid rectangular plate:  $20 \log_{10}(|\hat{p}_r|\pi|\mathbf{x}|/(2kL_1L_2|\hat{p}_Q|))$  calculated in the  $x_1x_3$ -plane ( $-\pi/2 \leq \theta \leq \pi/2$  in the shown polar coordinates, where  $\theta < 0$  implies  $\phi + \pi$  in the spherical coordinates used in the text) with  $kL_2 = 5$ ,  $\phi_i = 0$ , and (left)  $\theta_i = 0$ , (middle)  $\theta_i = \pi/6$ , (right)  $\theta_i = \pi/3$ .

This is because we assumed geometric far field  $|\mathbf{x}| \gg |\mathbf{y}|$  in eq. (9.8), which cannot be satisfied for every  $\mathbf{y}$  at the plate when it is infinitely large.

Plane reflectors can be an economic and simple alternative to diffusers (which will be addressed below) or absorbers, when reflections from a particular surface and with a well-defined direction of arrival to the listener should be suppressed (for example, in a control room with localized source and listener). They can redirect the reflected energy away from the listener to an absorbing or scattering surface. Reflectors are also preferred when strong early reflections should be achieved from relatively remote surfaces, when scattered reflections might be too weak or reach the listeners too late, for example, from high ceilings in concert halls, opera houses, or lecture halls. Apart from being made of sufficiently rigid material, in order to support reflection, their sizes and shapes can be chosen such that they reflect only relatively high frequencies, for example voice of a singer (or singer's formant mentioned in section 1.2.2) which is usually much weaker than orchestra. However, certain curvature of the reflectors is often desired, especially when the locations of sources and listeners are distributed. Pronounced resonances of the reflectors should be avoided in the frequency range of interest, which sometimes lead to unexpected narrowband absorption around the resonance frequency.

### 9.1.2 Schroeder diffuser

If we limit ourselves to relatively low incident and scattering angles to the surface, say  $\theta, \theta_i < \pi/4$ , then in eq. (9.19)  $|\cos(\theta) - \cos(\theta_i)| \ll |\cos(\theta) + \cos(\theta_i)|$ . Moreover, if  $|\hat{R}_s| \rightarrow 1$  (we allow only phase changes at the surface, without energy losses), the

equation can be simplified to

$$\begin{aligned}\hat{p}_r(\mathbf{x}) &= -\frac{jk\hat{p}_Q}{4\pi|\mathbf{x}|}e^{-jk|\mathbf{x}|}(\cos(\theta) + \cos(\theta_i)) \int_{-L_1}^{L_1} \hat{R}_s(y_1)e^{jAy_1}dy_1 \int_{-L_2}^{L_2} e^{jBy_2}dy_2 \\ &= -\frac{jkL_2\hat{p}_Q}{2\pi|\mathbf{x}|}e^{-jk|\mathbf{x}|}\text{sinc}(BL_2)(\cos(\theta) + \cos(\theta_i)) \int_{-L_1}^{L_1} \hat{R}_s(y_1)e^{jAy_1}dy_1.\end{aligned}\quad (9.25)$$

Apart from the factor in front of the integral<sup>64</sup>, the angularly dependent scattered sound depends only on the inverse spatial Fourier transform of the reflection coefficient,  $\mathcal{F}_{y_1}^{-1}\{\hat{R}_s(y_1)\}$ , with the parameter  $A = k[\sin(\theta_i)\cos(\phi_i) + \sin(\theta)\cos(\phi)]$ . If for  $\phi_i = 0$  we observe the field only in the  $x_1x_3$ -plane ( $\hat{R}_s(y_1)$  also does not change along the  $x_2$ -axis), switch to polar coordinates (letting  $\theta < 0$  represent  $\phi = \pi$ ), the parameter  $A$  becomes  $k[\sin(\theta_i) + \sin(\theta)]$  and the scattered field is proportional to

$$\hat{p}_r(\mathbf{x}) \propto \int_{-L_1}^{L_1} \hat{R}_s(y_1)e^{jk[\sin(\theta_i)+\sin(\theta)]y_1}dy_1. \quad (9.26)$$

The last expression reveals the idea behind (one-dimensional) Schroeder diffusers, which is to select such surface distribution of the reflection coefficient which gives approximately constant (independent of  $\theta$ ) value of the Fourier transform integral, at least in a certain frequency range. However, angular distribution of the reflected sound amplitude cannot be uniform because the parameter of the transform in the exponent in eq. (9.26) depends not only on  $\theta$ , but also on  $\theta_i$  and  $k$ . In fact, the scattering pattern exhibits peaks (lobes) separated by dips, the number and width of which depend on frequency and size of the diffuser in a similar manner as for a rectangular plate (which is indeed a special case for  $\hat{R}_s = 1$ ).

In practical realizations, one-dimensional Schroeder diffusers are subdivided over the length  $2L_1$  into  $N$  equally wide sections with the width  $l_1 = 2L_1/N$  and different values of the discrete function  $\hat{R}_s(y_1) = \hat{R}_n$ , for  $n = 1, 2, \dots, N$ . For practical reasons,  $N$  rarely exceeds 20. This means that the integral in eq. (9.26) can be replaced with a sum over  $n$ . If  $kl_1 \ll 1$ , we can replace  $y_1$  with  $nl_1 - L_1$  (the phase does not change significantly over one section width) and write

$$\hat{p}_r(\mathbf{x}) \propto \sum_{n=1}^N \hat{R}_n e^{jk[\sin(\theta_i)+\sin(\theta)](nl_1-L_1)}. \quad (9.27)$$

Next, we assume that the sequence of reflection coefficients  $\hat{R}_n$  repeats along  $x_1$ -axis with the period  $2L_1$ . As a consequence, the scattering pattern exhibits angular periodicity. Scattering maxima<sup>65</sup> are at the angles  $\theta$  which satisfy

$$\sin(\theta) + \sin(\theta_i) = \frac{m\pi}{kL_1} = \frac{m\lambda}{2L_1}, \quad (9.28)$$

<sup>64</sup>According to the assumption  $0 \leq \theta, \theta_i < \pi/4$ , so  $\sqrt{2} < (\cos(\theta) + \cos(\theta_i)) \leq 2$  does not substantially contribute to the directivity in the scattering pattern.

<sup>65</sup>These peaks and lobes are a consequence of the periodicity of the surface and should be distinguished from the lobes discussed above, with regard to eq. (9.26).

where  $m$  is an integer called diffraction order. In particular, for  $m = 0$ , the lobe is specular, with the angle  $\theta = -\theta_i$ . The sum becomes

$$\begin{aligned}\hat{p}_r(\mathbf{x}) &\propto \sum_{n=1}^N \hat{R}_n e^{jkm\lambda(nl_1-L_1)/(2L_1)} = \sum_{n=1}^N \hat{R}_n e^{j2\pi m(nl_1-Nl_1/2)/(Nl_1)} \\ &= e^{-j\pi m} \sum_{n=1}^N \hat{R}_n e^{j2\pi mn/N}.\end{aligned}\tag{9.29}$$

In common Schroeder diffusers, different reflection coefficients are achieved with different phase shifts  $\Phi$  in the sectors, that is,  $\hat{R}_n = |\hat{R}|e^{j\Phi_n}$ , with  $|\hat{R}| \rightarrow 1$ . The phase shifts are due to different depths  $d_n$  of the wells (with equal widths  $l_1$ ), which are separated by thin rigid plates. At the outer surface of the diffuser<sup>66</sup> (control surface which is its interface to the room), the phase shift of a plane wave propagating normal to the surface which is reflected back from the supposedly fully reflecting bottom of the well  $n$  is  $\Phi_n = 2\pi(2d_n/\lambda) = 2kd_n$ . Modulus of the complex amplitude of the reflected sound is then proportional to

$$\begin{aligned}|\hat{p}_r(\mathbf{x})| &\propto \left| \sum_{n=1}^N e^{j(2\pi mn/N + \Phi_n)} \right| = \left| \sum_{n=1}^N e^{j(2\pi mn/N + 2kd_n)} \right| \\ &= \left| \sum_{n=1}^N e^{(j/N)(2kL_1n(\sin(\theta) + \sin(\theta_i)) + 2\pi s_n k/k_d)} \right|.\end{aligned}\tag{9.30}$$

Since the phase shift depends on frequency (wave number), the diffuser is expected to be efficient within a certain frequency range around the design frequency  $f_d$  (and generally its multiples). With regard to that, depths of the wells are commonly expressed with integer numbers  $s_n = 2Nd_n/\lambda_d = Nk_d d_n/\pi$ , where  $\lambda_d = c_0/f_d$  and  $k_d = 2\pi/\lambda_d$ . Similarly,  $\Phi_d = 2k_d d_n = 2\pi s_n/N$ . Scattering pattern can be defined as

$$SP_{Schroed} = 20 \log_{10} \left( \left| \sum_{n=1}^N e^{(j/N)(2kL_1n(\sin(\theta) + \sin(\theta_i)) + 2\pi s_n k/k_d)} \right| \right)$$

(9.31)

down to a constant value in dB. Normalization can be done by subtracting the maximum value of the expression on the right-hand side of eq. (9.31).

Wavelength which corresponds to the design frequency should be much larger than the well width  $l_1$  ( $k_d l_1 < 1$ ). In practice, a rule of thumb is often used according to which the

---

<sup>66</sup>The model of a plane rectangular surface which has been used here starting with eq. (9.1) does not consider sound propagation behind the control surface. Consequently, all effects of the surface and possible contributions behind it are contained only in the reflection coefficient at the plane surface. The details of sound propagation inside the wells of the diffuser, such as reflections from the side walls of the wells, cannot be captured with this model, although they sometimes lead to physically relevant effects, such as additional absorption of the diffuser.

design frequency should satisfy  $f_d < c_0/(2l_1)$ . Some scattering will take place at higher frequencies as well, even though the model presented here might not be valid any more. On the other hand, the wavelength should not be much larger than the width of one repetition of the diffuser,  $2L_1 = Nl_1$ , for the diffuser to be efficient, that is  $k_d Nl_1 > 1$ , which gives an approximate lower limit for the design frequency:  $f_d > c_0/(2Nl_1)$ . In reality, certain diffusion is achieved even 1-2 octaves below this frequency. The two limits for the design frequency can be expressed simultaneously:  $c_0/(2Nl_1) < f_d < c_0/(2l_1)$  or shorter:  $\pi < 2k_d L_1 < N\pi$ .

Several types of sequences  $s_n$  are used for achieving efficient diffusers. For a quadratic residue sequence,  $s_n = (n - 1)^2 \bmod N$ , where  $N$  is a prime number, and mod gives the least non-negative remainder (in order to avoid unnecessarily deep wells which provide the phase shifts larger than  $2\pi$  when  $s_n > N$ ). Such a sequence gives equal peaks of the lobes described by eq. (9.28). Figure 17 shows normalized scattering patterns of a Schroeder diffuser based on the quadratic residue sequence with  $N = 7$  and different values of some of the relevant parameters, estimated using eq. (9.31). Other possibilities for the sequences include:

- maximum length sequences (with only two depths of the wells, 0 and  $d$ ); such diffusers are usually efficient only in one octave but have advantage of the small depth
- primitive root sequences ( $s_n = r_{N+1}^n \bmod (N + 1)$ , with  $N + 1$  a prime number and  $r_{N+1}$  its primitive root (resulting in all unique values of  $s_n$  for  $n = 1, 2, \dots, N$ ); additional suppression of the specular reflection is achieved compared to the diffusers based on a quadratic residue sequence (compare with the bottom right graph in Fig. 17), especially for large  $N$
- Legendre sequences; such diffusers behave similarly as the diffusers based on primitive root sequences; however, some of the wells are filled with absorber which leads to higher energy losses

For simpler mathematical treatment, we have considered only one-dimensional Schroeder diffusers, the wells of which are treated as narrow rectangular surfaces with different reflection coefficients (the phase parts), which periodically extend along a single axis. Two-dimensional diffusers can be constructed in the same manner over two perpendicular axes. Furthermore, limited frequency range in which Schroeder diffusers are efficient can be expanded for example with fractal diffusers, when the bottom of each relatively wide well contains another high-frequency diffuser. The overall depth of the diffusers, which is dictated by the deepest wells, is sometimes decreased by bending the deepest wells behind the shorter ones, parallel to the surface of the diffuser. However, this requires more complicated and less robust designs.

Adding to the analysis, it should be mentioned that the scattering patterns of Schroeder diffusers and their overall acoustic behaviour are affected in reality by their finite extension (finite number of repetitions with period  $2L_1$ ), additional absorption, and three-dimensional sound propagation inside the wells at high frequencies. All these phenomena

can also lower the accuracy of predictions. Higher accuracy can be achieved with the aid of numerical simulations.

Of course, Schroeder diffusers are not the only acoustic elements which can contribute to the diffuseness of the sound field in a room. Any non-planar rigid surface is expected to scatter the sound at least for a certain range of Helmholtz number values around  $kL \approx 1$ , where  $L$  is the characteristic size (depth/width) of the geometric irregularities. Scattering pattern and scattering coefficient are in general difficult to estimate and there are accordingly less data and models available in literature compared with, for example, absorption of porous absorbers. Diffusers are also sometimes used in combination with absorbing materials, which results in hybrid absorbing-scattering surfaces. Many variations and products exist on the market.

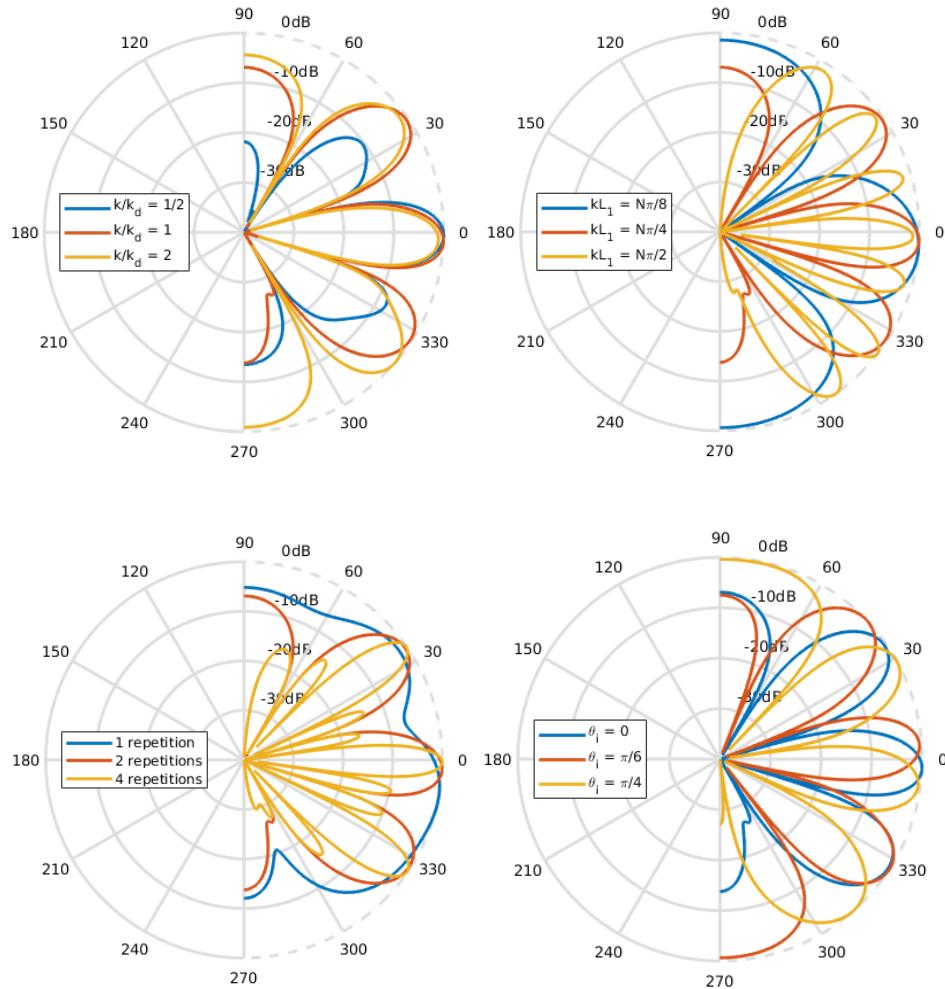


Figure 17: Normalized scattering pattern ( $SP_{Schroed} - \max(SP_{Schroed})$ ) of a one-dimensional Schroeder diffuser with quadratic residue sequence, with  $N = 7$  ( $s_n = 0, 1, 4, 2, 2, 4, 1$ ) and: (above left)  $kL_1 = N\pi/4$ , 2 repetitions,  $\theta_i = 0$ , (above right)  $k/k_d = 1$ , 2 repetitions,  $\theta_i = 0$ , (below left)  $k/k_d = 1$ ,  $kL_1 = N\pi/4$ ,  $\theta_i = 0$ , (below right)  $k/k_d = 1$ ,  $kL_1 = N\pi/4$ , 2 repetitions.

## 10 Modelling in room acoustics

In the preceding sections we have mostly considered how acoustics in a room and associated physical phenomena can be understood, modelled, and to some extent predicted theoretically. When a more detailed and accurate treatment of complicated realistic scenarios is necessary, the relevant acoustic phenomena can be measured, tested experimentally, or predicted by means of numerical simulations. While on-site measurements generally give the most reliable results, numerical calculations and experiments on scale models can offer greater flexibility in testing different possible scenarios and modifications on a model of a room and can be used for predictions of sound field before the actual room is built. Therefore, this section is concerned with how room acoustics can be modelled and analysed with ray tracing and image source technique, as the most frequently used methods in computational room acoustics, as well as physical scale models.

### 10.1 Ray tracing

Ray tracing technique relies on the high-frequency approximation of sound waves with rays, which was described in section 4.6, and energy summation, which was discussed in section 5.1. We concluded that energy summation is appropriate for sufficiently broadband incoherent waves. On the other hand, far acoustic field, in which plane wave and sound ray approximations are valid, assumes distances from any source of sound which are larger than the wavelength (eq. (4.37)). Consequently, ray tracing is reasonable to use only in acoustically large rooms, the characteristic dimension  $L$  of which is significantly larger than the wavelengths of interest. The corresponding Helmholtz number value must satisfy  $kL \gg 1$ . Another frequently used condition is that the observed frequency range is above Schroeder frequency given in eq. (3.90), so that no single mode is likely to dominate in the frequency response (or any of the frequency bands in which the energy is summed). For example, if we suppose the room volume  $V = L^2H$ , where  $L$  is characteristic length/width of the room and  $H$  its characteristic height, we can write Schroeder frequency as

$$f_{Schroed} \approx 2100 \sqrt{\frac{T_{60,\Delta f_n} \cdot 1\text{m}^3/\text{s}^3}{V}} = \frac{2100}{L} \sqrt{\frac{T_{60,\Delta f_n} \cdot 1\text{m}^3/\text{s}^3}{H}}. \quad (10.1)$$

For common values of height  $H$  and reverberation time  $T_{60}$ ,  $f > f_{Schroed}$  usually appears as a more stringent condition than  $f \gg c_0/(2\pi L) \approx 55\text{m} \cdot \text{s}^{-1}/L$ , which follows from  $kL \gg 1$ .

In order to study energy response of a room to an excitation by a source of sound in it, sound rays can be represented with virtual **particles** which transport sound energy along the ray paths at the speed of sound  $c_0$ . The particles start from the source (or the point of last reflection from a surface), which we set at the origin of a coordinate system, so that each direction of propagation is given with the angle  $(\theta, \phi)$  in usual spherical coordinates. Calculations in each frequency band of interest are performed

separately. The number of traced particles should be large enough to ensure statistically well covered calculation, but small enough to keep the computational costs (memory and processing power) reasonably low (see also footnote<sup>67</sup> below). Stochastics of the calculations, which is necessary for statistical treatment of the sound field (for example, scattering), is achieved by means of random numbers. Hence, we introduce a random real number  $0 \leq z \leq 1$  which can take values in the range between 0 and 1. When we need more than one random number simultaneously, we add indices for distinction:  $z_1, z_2$ , etc.

An omnidirectional **source** in the origin of its local coordinate system radiates particles at random angles  $\theta = \text{round}(z_1) \arccos(z_2) + (1 - \text{round}(z_1))(\pi - \arccos(z_3))$  and  $\phi = 2\pi z_4$ . The function `round()` outputs 0 or 1 thus randomly selecting one of the two hemispheres around the source. The distribution of  $\theta$  is not uniform as in the case of  $\phi$ . This is because all particles are supposed to carry the same energy. Since the energy radiated from an omnidirectional source should be spread equally over a control sphere surrounding the source, distribution of  $\theta$  should be uniform after scaling with  $1/\sin(\theta)$ . This follows from the integrals in eq. (4.52) and the equality  $d\Omega/\sin(\theta) = d\theta d\phi$ .

From eq. (4.55), energy of a single particle emitted by the omnidirectional source equals

$$E_p = \frac{|\langle \mathbf{I}_p \rangle_T|}{c_0} = \frac{\langle P_q \rangle_T / 1\text{m}^2}{N_p c_0}, \quad (10.2)$$

where  $\langle P_q \rangle_T$  is time-averaged power of the source. The full solid angle is covered with  $N_p$  particles, so  $4\pi$  (which is in the denominator for the rays with solid angle  $d\Omega$ ) is replaced with  $N_p$ . However, if ray tracing is used only for estimation of energy impulse response of a room and calculation of the descriptors (section 6.2) for which absolute sound levels are irrelevant, such as reverberation time, definition, clarity, etc., multiplication of  $E_p$  with a constant value equal for each particle does not affect the results.

Directivity of the source can be included by multiplying the energy of each emitted particle with  $D_i^2(\theta, \phi)$ . Alternatively, rather than scaling energy of the particles, their angular density can be adjusted by annihilating (removing) particles in the directions where the radiation pattern of the source drops. This decreases the total energy in these directions and thus simulates directivity of the source. The annihilation can be performed stochastically when the condition  $z > D_i(\theta, \phi)/D_{i,max}$  is met. The particles which are not annihilated keep their equal energies regardless of the direction. For a large number of particles  $N_p$ , energy scaling and particle annihilation have the same effect on the final results. While the latter technique can be less accurate for smaller numbers of particles (since fewer particles remain for a good statistical coverage), it is computationally more efficient, as some particles are left out from the computation and no additional multiplication with the scaling factor is necessary.

Acoustic impulse generated by a point source is achieved when all the particles start from the location of the source, say  $\mathbf{x}_0$ , at the same time  $t_0$ . Starting from  $t_0$ , each particle is followed (traced) and a check is made whether it hits a surface or a receiver in the room. The particles are traced until (a) they are annihilated, (b) their energy drops below

some predefined value  $E_{p,min}$  (which suffices for the extraction of the output parameter values), or (c) some specified time  $t > t_0$  (which forces the end of the calculation if the other two conditions are not met).

All the **surfaces** are assumed to consist of plane polygons. In practice, it is usual to approximate curved surfaces with a set of plane surfaces. In the following, we suppose that all the surfaces are triangles (a plane polygon can always be subdivided into triangles). For a triangle with the vertices at points  $\mathbf{x}_A$ ,  $\mathbf{x}_B$ , and  $\mathbf{x}_C$  located in the supporting plane determined by the three points, we define direction of the normal to the plane according to the right-hand rule, that is, if a right-handed screw is placed normal to the plane and rotated in the direction from  $\mathbf{x}_A$  over  $\mathbf{x}_B$  to  $\mathbf{x}_C$ , it moves in the direction of the normal. Unit vector normal to the supporting plane then equals

$$\begin{aligned}\mathbf{n} &= \frac{(\mathbf{x}_B - \mathbf{x}_A) \times (\mathbf{x}_C - \mathbf{x}_A)}{|(\mathbf{x}_B - \mathbf{x}_A) \times (\mathbf{x}_C - \mathbf{x}_A)|} = \frac{(\mathbf{x}_C - \mathbf{x}_B) \times (\mathbf{x}_A - \mathbf{x}_B)}{|(\mathbf{x}_C - \mathbf{x}_B) \times (\mathbf{x}_A - \mathbf{x}_B)|} \\ &= \frac{(\mathbf{x}_A - \mathbf{x}_C) \times (\mathbf{x}_B - \mathbf{x}_C)}{|(\mathbf{x}_A - \mathbf{x}_C) \times (\mathbf{x}_B - \mathbf{x}_C)|}.\end{aligned}\quad (10.3)$$

Equation of the supporting plane which holds for any point  $\mathbf{y}$  at the plane reads

$$\mathbf{n} \cdot \mathbf{y} = P_p, \quad (10.4)$$

where  $P_p$  is a constant. It can be found using any vertex of the triangle, for example, from  $P_p = \mathbf{n} \cdot \mathbf{x}_A$ .

Next, we need to find the intersection point  $\mathbf{x}_s$  of a particle emitted (or reflected) from the point  $\mathbf{x}_0$  in the direction given by a unit vector  $\mathbf{e}_r$  fixed with two spherical coordinates,  $(\theta, \phi)$ , with the supporting plane given by eq. (10.4). Locus of the particle's path (the ray) consists only of the points  $\mathbf{x}$  which satisfy

$$\mathbf{x} = \mathbf{x}_0 + R\mathbf{e}_r, \quad (10.5)$$

where the vector  $\mathbf{e}_r$  has Cartesian components  $e_{r1} = \sin(\theta) \cos(\phi)$ ,  $e_{r2} = \sin(\theta) \sin(\phi)$ , and  $e_{r3} = \cos(\theta)$ , and  $R$  is any real number (only  $R > 0$  implies forward travelling particle along the ray line). If  $\mathbf{n} \cdot \mathbf{e}_r = 0$ , the particle travels parallel to the supporting plane and there is no intersection point. In all other cases, the value of  $R$  for the intersection point is found after replacing  $\mathbf{y}$  from eq. (10.4) with  $\mathbf{x}$  from eq. (10.5):

$$\mathbf{n} \cdot (\mathbf{x}_0 + R\mathbf{e}_r) = P_p, \quad (10.6)$$

from which it follows:

$$R = \frac{P_p - \mathbf{n} \cdot \mathbf{x}_0}{\mathbf{n} \cdot \mathbf{e}_r}. \quad (10.7)$$

Hence, when  $R > 0$ , the intersection point is

$$\mathbf{x}_s = \mathbf{x}_0 + \frac{P_p - \mathbf{n} \cdot \mathbf{x}_0}{\mathbf{n} \cdot \mathbf{e}_r} \mathbf{e}_r. \quad (10.8)$$

Now it needs to be checked whether the intersection point  $\mathbf{x}_s$  lies inside or outside the triangle with the vertices  $\mathbf{x}_A$ ,  $\mathbf{x}_B$ , and  $\mathbf{x}_C$ . If it is outside the triangle, then at least one cross product  $(\mathbf{x}_B - \mathbf{x}_A) \times (\mathbf{x}_s - \mathbf{x}_A)$ ,  $(\mathbf{x}_C - \mathbf{x}_B) \times (\mathbf{x}_s - \mathbf{x}_B)$ , or  $(\mathbf{x}_A - \mathbf{x}_C) \times (\mathbf{x}_s - \mathbf{x}_C)$  must point into the direction opposite from  $\mathbf{n}$ . Therefore, for an intersection point which is inside the triangle or at its edges, all the following conditions have to be satisfied:

$$\begin{aligned} [(\mathbf{x}_B - \mathbf{x}_A) \times (\mathbf{x}_s - \mathbf{x}_A)] \cdot \mathbf{n} &\geq 0, \\ [(\mathbf{x}_C - \mathbf{x}_B) \times (\mathbf{x}_s - \mathbf{x}_B)] \cdot \mathbf{n} &\geq 0, \\ [(\mathbf{x}_A - \mathbf{x}_C) \times (\mathbf{x}_s - \mathbf{x}_C)] \cdot \mathbf{n} &\geq 0. \end{aligned} \quad (10.9)$$

The equalities are satisfied when the intersection point is at the corresponding edge or vertex appearing in the expression. For efficient computation, the algorithm can stop the check as soon as one of the conditions is not satisfied. If the intersection point belongs to the surface of the triangle, length of the particle's path to the surface is

$$r = |\mathbf{x}_s - \mathbf{x}_0| \quad (10.10)$$

and the elapsed travel time is  $t - t_0 = r/c_0$ . If necessary, **dissipation** in air can be taken into account at this stage according to equations (3.39) and (3.40) by scaling the particle energy as

$$E_p e^{-m_{air}r} \quad (10.11)$$

with  $m_{air}$  attenuation constant in air and  $4.34m_{air}$  attenuation expressed in dB/m. If the path intersects multiple surfaces, the particle is reflected (assuming that no transmission takes place) from the closest surface, which apart from  $R > 0$  gives the shortest distance  $r$ .

Energy losses due to reflection from an **absorbing** surface are accounted for either by particle energy scaling,  $(1 - \alpha)E_p$ , or by annihilation of the particle when  $z < \alpha$ , where  $z$  is a random number. **Scattering** is commonly modelled using Lambert's law from eq. (8.21) with the scattering coefficient  $s_s$  independent from the angle of incidence  $\theta_i$ . Rather than creating many new scattered particles, which would dramatically increase the computational efforts, only one particle can be traced with the same energy as the incident particle (neglecting additional absorption). Angle of the reflected particle relative to the normal of the surface is calculated as

$$\begin{aligned} \theta_r &= |2 \arccos(\sqrt{z_2}) - \pi/2|, & \phi_r &= 2\pi z_3, & \text{for } z_1 < s_s, \\ \theta_r &= \theta_i, & \phi_r &= \phi_i + \pi, & \text{for } z_1 \geq s_s, \end{aligned} \quad (10.12)$$

which on a large sample gives approximately the same distribution of scattering angles as Lambert's cosine law (with added uniform distribution of  $\phi_r$ ). This is demonstrated in Fig. 18 on the sample of 10000 incident particles. Local polar and azimuthal angles,  $\theta_i$  and  $\phi_i$ , are calculated at the point of reflection from  $\mathbf{e}_r(\theta, \phi)$  of the incident particle and surface normal  $\mathbf{n}$ . Similarly, new values of  $\theta$  and  $\phi$  of the reflected particle are calculated from  $\theta_r$ ,  $\phi_r$ , and  $\mathbf{n}$ . Values of the cumulated elapsed time  $t$  and particle energy  $E_p$  (if energy is scaled) are updated after each reflection. Tracing of the reflected

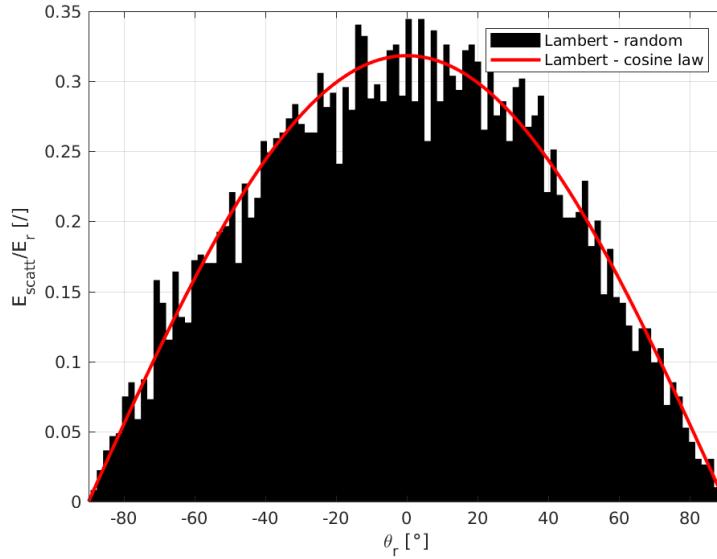


Figure 18: Normalized scattered energy (for  $z_1 < s_s$  only) according to Lambert's cosine law approximated with random variables and plotted as the histogram of  $\theta_r = 2 \arccos(\sqrt{z_2}) - \pi/2$  normalized with  $N_p \pi^2 / (2N_{bins})$ , where  $N_p = 10000$  is the number of incident particles and  $N_{bins} = 100$  is the number of bins in the histogram ( $-\pi/2 \leq \theta_r \leq \pi/2$  for easier comparison with the curve for  $s_s(\theta_i) = 1$  in Fig. 12).

particle with the calculated direction  $(\theta, \phi)$  can be continued with  $\mathbf{x}_s$  becoming new  $\mathbf{x}_0$  and  $t$  becoming  $t_0$ .

Together with the inspection of possible next reflections of the particle, a check should be made whether it hits one or more receiving spheres on its path. An omnidirectional **receiver** is represented as a virtual spherical volume with radius  $a_{rec}$ . The radius should not be too small in order to ensure that statistically relevant number of particles hit the receiver and thus contribute to the reliability of the results<sup>67</sup>. For a receiver located at  $\mathbf{y}_{rec}$ , the check can be done<sup>68</sup> by measuring the distance  $d$  between  $\mathbf{y}_{rec}$  and the path line at the closest point  $\mathbf{x}_0 + R_{rec}\mathbf{e}_r$ ,

$$d = |\mathbf{x}_0 + R_{rec}\mathbf{e}_r - \mathbf{y}_{rec}|, \quad (10.13)$$

where

$$R_{rec} = (\mathbf{y}_{rec} - \mathbf{x}_0) \cdot \mathbf{e}_r. \quad (10.14)$$

<sup>67</sup>The number of particles which hit the receiving sphere depends on its volume  $V_{rec} = 4a_{rec}^3\pi/3$  and the total number of particles  $N_p$ , as well as the size of the room and absorption in it (when particles get annihilated or fully absorbed). If  $V_{rec}$  is small,  $N_p$  should accordingly be large to maintain the statistical validity, especially for large and damped rooms. Reasonable values are, for example,  $N_p \sim \mathcal{O}(10^5)$  and  $2a \sim \mathcal{O}(1\text{m})$  for large concert halls and  $N_p \sim \mathcal{O}(10^3)$  for smaller lecture rooms, although the number of particles can be increased as more powerful computational resources become available.

<sup>68</sup>Certain complications and inaccuracies can appear if some surface in the room partly covers or intersects the receiving volume. We neglect such cases.

If  $d < a_{rec}$ ,  $R_{rec} > 0$ , and  $R_{rec}$  is smaller than  $R$  of the next reflecting surface (which has the smallest positive  $R$  of all the surfaces), then the particle crosses the receiving volume. Its current energy (scaled according to eq. (10.11) with  $R_{rec}$  replacing  $r$ , if dissipation in air is included) and time of arrival ( $t_0 + R_{rec}/c_0$ ) are stored. Angle of arrival can also be stored (both for receivers and surfaces), as it gives valuable information on spatial properties of the field and diffuseness.

For each receiver, energies of the received particles are summed within certain time frames<sup>69</sup>, based on the assumption of broadband and incoherent incoming waves. If air dissipation is neglected and absorption is implemented as annihilation of particles rather than energy scaling, this comes down to simple counting the received particles in each time frame and multiplying the result with the initial energy of the particles from eq. (10.2). No additional storage of the energies of traced particles needs to be done, which is a great advantage of this method. The obtained distribution of the received energy over time,  $E(t_i)$ , where  $i = 1, 2, \dots$  is the number of each successive time frame, is an estimation of **echogram** (reflectogram) or energy impulse response averaged within the time frames.

Echograms can be used for the assessment of energy-based descriptors of room acoustics, as well as for auralization. In the latter case, a set of echograms for different frequency bands (for example, octaves) is treated as a series of power spectra in different time frames. After interpolation of the values in the entire considered frequency range (in order to increase the frequency resolution), energy distribution  $E(f, t_i)$  is obtained from which frequency response in each time frame can be estimated as  $\sqrt{2E(f, t_i)\rho_0 c_0^2 e^{j\Phi(f, t_i)}}$ . The exact values of phase  $\Phi(f, t_i)$  cannot be extracted from energy responses, so they are often assumed to be random and therefore generated as random values for each frequency<sup>70</sup>. Inverse Fourier transform and merging the results from different time frames give an estimated impulse response of the room, which can be used for auralization, for example by convolving it with some “dry” input signal. Furthermore, impulse responses obtained at two or more close locations can be used for rendering binaural or ambisonic recordings with additional spatial information.

It should be mentioned that the algorithm sketched above is not computationally optimized. Many improvements can be made to make the computations faster with less processing or memory requirements. For example, the normal vector  $\mathbf{n}$  in eq. (10.3) does not have to be normalized to a unit vector by the denominators for the rest of the algorithm to work. Conditions (10.9) can be checked more efficiently after the transformation to two-dimensional coordinates, since all points belong to the same supporting plane. However, the focus here is not on the algorithmic optimization.

---

<sup>69</sup>While too long time frames lower the temporal resolution of the results, too short frames may increase the statistical error due to small numbers of particles hitting the receiver within different frames. Frame interval of 5-10ms can be chosen as a reasonable compromise which also fits well to the inertia of human hearing, as discussed in section 1.2.1.

<sup>70</sup>Still, phase must be an odd function of frequency for real systems:  $\Phi(f) = -\Phi(-f)$ . Minimum phase function is often used which has the properties of causality and stability.

Ray tracing simulations consist in practice of several steps. The first step is creating the three-dimensional geometry of the room with all relevant objects in it. The model of the room should not be too detailed (which is often in contrast to the architectural drawings and hinders the use of the latter ones for acoustic simulations; adaptation of the available three-dimensional CAD drawings is usually necessary, which can be time consuming). Large number of small surfaces does not only increase the computational load, but can also decrease the accuracy, for example, due to unrealistic specular reflections from acoustically small surfaces hit by the particles. When a larger curved surface has to be subdivided into smaller plane surfaces, diffusion should be approximated with higher values of the scattering coefficient of relatively large surfaces (crude subdivision), rather than a large number of small surfaces (fine subdivision).

The second step is specifying acoustically relevant properties of all the surfaces. These are primarily absorption and scattering coefficient. It is suggested that the latter one is always larger than, say, 0.05, even for large flat surfaces (finite surfaces will always scatter the sound close to their edges). Additional parameter which is sometimes introduced is transparency coefficient. It represents a fraction of incident energy which is transmitted through a surface. This coefficient is normally set to zero for boundary surfaces and the entire non-reflected energy is captured by the absorption coefficient. However, it can have a value between 0 and 1 for a partly transparent surface in the room, for example hanging absorber and reflector, curtain, panel, etc. In such a case, absorption coefficient of the same surface should not include the transmitted energy (eq. (8.15) becomes  $\alpha_s = 1 - |\hat{R}_s|^2 - |\hat{T}_s|^2$ , where  $\hat{T}_s$  is transmission coefficient and  $|\hat{T}_s|^2$  is equal to the transparency coefficient.).

In the last step of the preparation of calculations, the sources and receivers have to be defined, in particular their locations and directivity (which requires also the definition of local coordinate system). In general, different values of sound power of the sources can be specified for different frequency bands. Receivers are usually not directional. After the sources and receivers have been defined, the calculations can be performed followed by the post-processing of the results and extraction of output quantities. The major advantage of numerical simulations over measurements is that different possibilities for room designs can be tested efficiently, such as various geometries (for example, angle of reflectors, orientation of boundary surfaces) or values of the input parameters (absorption and scattering coefficient). When measurement results are available, the simulated model can be adjusted (particularly in terms of less known input parameters), so that the numerical results match well the measured values, and thus validated before further modifications are introduced.

## 10.2 Image sources

As discussed in section 7.3, strong low-order reflections are often critical for room acoustics and difficult for accurate predictions. For example, some important reflections may be completely left out in a ray tracing simulation due to a low number of traced particles,

which miss either the surface or the receiver after reflection, or estimated inaccurately due to the intrinsic assumptions of the method (energy summation instead of wave superposition). For improved accuracy, the earliest (mostly 1<sup>st</sup>- and 2<sup>nd</sup>-order) reflections can be modelled with image sources.

As pointed out in section 8.1, acoustically large plane and uniform surfaces which reflect specularly can be replaced with image sources. Each  $n^{\text{th}}$ -order image source corresponds to one  $(n-1)^{\text{th}}$ -order image source, with the actual sources in the room having zero order. Since the theory is not based on sound rays, sound decay with the distance,  $p \sim 1/r$ , has to be taken into account for all point sources, including the image sources, and sound pressure (not energy) at a receiving point  $\mathbf{x}$  (not a sphere) can be estimated from eq. (4.19) for compact omnidirectional sources, starting with the lowest order sources. If eq. (4.21) is used for directional sources, an image source has to be oriented symmetrically to the corresponding lower-order source with respect to the reflecting surface (as an image in a mirror).

If the reflecting surface is uniform, motionless, with angularly independent absorption coefficient, absorption can be included by scaling the image source strength as  $\sqrt{1 - \alpha_s} \hat{Q}(\mathbf{y})$ , which corresponds to replacing  $\hat{R}_s(\theta)$  in eq. (8.4) with  $\sqrt{1 - \alpha_s}$  following from eq. (8.15)<sup>71</sup>. Similarly as in eq. (7.1), fraction of non-specularly reflected energy can be excluded by additional multiplication with  $\sqrt{1 - s_s}$ , while image sources provide only specularly reflected components.

Calculation of sound pressure rather than energy captures more accurately superposition of waves in comparison to the ray tracing method. There is no need for simplified models of sound propagation associated with energy summation or plane waves approximation. However, the accuracy is limited by the range of validity of the image source approach, which is only for large flat surfaces reflecting specularly. As demonstrated in section 9.1, relatively small or non-planar surfaces reflect significant fraction of incident energy non-specularly.

All room surfaces have finite dimensions. Consequently, symmetry planes behind which image sources are positioned can extend outside the actual surfaces, if the sound waves from these image sources reach some receiving point through the parts of the symmetry planes where the actual surfaces are located. For the receivers for which this is not the case, the image source is not audible. Unfortunately, this means that the contribution of an image source depends on particular locations of both sources and receivers in the room. These dependencies are not easily determined in general, especially for higher-order image sources. Each image source has to be checked for audibility for each receiver independently, for example, by backtracing along the path from the receiver to the image source in question. If the path does not cross actual mirror surface but its extension in the supporting plane, the image source is not audible. If it does, another path is constructed from the intersection point at the actual surface to the lower order image source of the considered source and checked whether it crosses actual surface. The

---

<sup>71</sup>Minimum phase function can again be used for the phase.

procedure is continued until a zero-order (real) source is reached. The initial image source is audible for the selected receiver only if all the paths cross actual surfaces. When the image source technique is combined with ray tracing, in the hybrid schemes, the audibility can be extracted from the calculated ray paths. If a ray reaches a receiver after  $n$  reflections, the corresponding  $n^{\text{th}}$ -order image source is audible for the receiver.

Another disadvantage of the method is that the number of image sources increases rapidly with their order, which makes the image source technique impractical for calculation of late reflections. For a single zero order source and  $N$  surfaces in the room, there can be  $N$  image sources of the first order. If each image source is mirrored with respect to the other  $N - 1$  surfaces,  $N(N - 1)$  2<sup>nd</sup>-order image sources can be obtained. Continuing this process for higher orders gives the total of

$$N_q = \sum_{i=1}^n N(N - 1)^{i-1} = N \sum_{i=1}^n (N - 1)^{i-1} = N \frac{(N - 1)^n - 1}{N - 2} \quad (10.15)$$

image sources up to the order  $n$ . For the last equality, the sum was expressed from

$$\begin{aligned} \Sigma &= \sum_{i=1}^n (N - 1)^{i-1} = 1 + (N - 1)^1 + \dots + (N - 1)^{n-1} \\ &= \frac{1}{N - 1} [(N - 1)^1 + (N - 1)^2 + \dots + (N - 1)^n] = \frac{\Sigma - 1 + (N - 1)^n}{N - 1}. \end{aligned} \quad (10.16)$$

Such a large growth with  $n$  limits the application of image sources in practice mostly to the first few orders, unless the room geometry is very simple with small number of surfaces. Moreover, as the order of image sources and their distances from the actual room interior increase, the fraction of inaudible image sources increases. For these reasons, image source technique is usually used in combination with ray tracing, for calculating the earliest reflections with more accuracy. Late energy, starting with the reflections of the 4<sup>th</sup>- or 5<sup>th</sup>-order, is estimated more accurately and efficiently with ray tracing.

### 10.3 Scale models

Scale models in room acoustics are built physical models of rooms with scaled geometry. Typical scaling ratio for relatively large halls is 1:10 (scale factor 10) or somewhat higher. Accordingly, all distances and propagation times measured in a scale model should be multiplied with the scale factor when compared with the actual room, while the analysed frequencies should be divided by the same factor in order to keep the value of relevant Helmholtz numbers unchanged (sound wavelength should be adapted to the smaller geometry)<sup>72</sup>. While higher values of the scale factor are desirable in order to make the model less cumbersome and expensive, too large values are restricted by the acoustic

<sup>72</sup>We assume that the model is filled with air with the same speed of sound as in the original room. Other media have been used occasionally.

conditions at the scaled frequencies and expected similarity between the model and the actual room. The main reason for this is that the considered frequencies are by the scale factor higher than actual frequencies. In practice, this is associated with several complications:

- absorption and other acoustic properties of the surfaces should match those of the actual surfaces but shifted to higher frequencies, which can be achieved only approximately,
- if the medium (air) is the same as in the original room, attenuation constant  $m_{air}$  at the scaled frequencies should be higher by the scale factor than at the original frequencies (since the sound paths are shorter by the same factor), which is generally not the case,
- sound sources and receivers should be omnidirectional even at very high analysed frequencies (often in the ultrasound range),
- sound sources should provide correspondingly very short impulses for impulse response measurements.

These limit the reliability of measurements in scale models and their applicability mainly to studies of early sound energy and low-order reflections from weakly absorbing surfaces. For this reason, as well as substantially lower costs of production, partial scale models can be used for the analyses of particularly important early reflections in certain parts of the room, for example, close to a stage. For basic analyses, hard, fully reflecting materials can be used to represent all acoustically hard surfaces in the room and highly absorbing porous materials for all surfaces which essentially absorb. Dissipation in air is often neglected.

Omnidirectional radiation of a source in a scale model should be verified, as well as its dynamics over the entire (scaled) frequency range of interest. Spark generators are frequently used, since they provide very short impulses and omnidirectional radiation pattern due to their small size. This is more difficult to achieve with small electro-acoustic transducers, for example, in the form of a dodecahedron. However, the advantage of the latter ones is that the deterministic excitation signals can be used, which allows more robust and flexible methods for acquisition of impulse responses, as discussed in section 6.1. Small, usually 1/4-inch or 1/8-inch microphones are used to ensure omnidirectional characteristics of the receivers.

Due to their price and demanding process of production, scale models are used relatively rarely, mostly in large projects in which room acoustics is critical, such as for concert halls and opera houses. Even then, they are used only in combination with more flexible and efficient numerical simulations. However, one of the greatest advantages of scaled models over ray tracing and image source techniques remain more realistic diffraction and scattering from surfaces.

## Literature

L. Beranek: "Concert Halls and Opera Houses: Music, Acoustics, and Architecture" (2<sup>nd</sup> edition), Springer, 2004

T. J. Cox, P. D'Antonio: "Acoustic Absorbers and Diffusers" (2<sup>nd</sup> edition), Taylor & Francis, 2009

W. Fasold, E. Veres: "Schallschutz und Raumakustik in der Praxis", Verlag für Bauwesen Berlin, 1998

H. Kuttruff: "Room Acoustics" (6<sup>th</sup> edition), CRC Press, 2017

J. Meyer: "Acoustics and the Performance of Music" (5<sup>th</sup> edition), Springer, 2009

S.W. Rienstra, A. Hirschberg: "An Introduction to Acoustics", Eindhoven University of Technology, 2017

M. Vorländer: "Auralization – Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality", Springer, 2008