Acoustic Simulation and Visualization Algorithms

Vom Fachbereich Informatik der Technischen Universität Kaiserslautern
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
genehmigte Dissertation
von

Eduard Deines

Danksagung

An dieser Stelle möchte ich mich ganz herzlich bei allen, die mich während der Arbeit an dieser Dissertation unterstützt haben, bedanken. Mein Dank gilt meinem Doktorvater, Hans Hagen, der mir nicht nur die einzigartige Gelegenheit gab im Rahmen des internationale Graduiertenkollegs ("Visualization of Large and Unstructured Data Sets") an der TU Kaiserslautern zu promovieren, sondern während der gesamten Zeit mit wertvollen Ratschlägen zum Erfolg dieser Arbeit beigesteuert hat. Jörg Meyer möchte ich für die Unterstützung und hilfreiche Gespräche während meines Aufenthalts an der UC Irvine danken.

Auch möchte ich mich bei den MitarbeiterInnen der Arbeitsgruppe "Graphische Datenverarbeitung" der TU Kaiserslautern und des Fraunhofer ITWM Kaiserslautern bedanken. Besonderer Dank gilt Martin Hering-Bertram, der immer Zeit hatte anstehende Probleme ausführlich zu diskutieren und mit seinen Anregungen und seiner konstruktiven Kritik sehr viel zur erfolgreichen Fertigstellung dieser Arbeit beigetragen hat. Jan Mohring möchte ich dafür danken, dass er ein offenes Ohr für meine Fragen betreffend des mir neuen Gebietes "Akustik" hatte. Frank Michel und Jevgenij Jegorovs danke ich für das Mitwirken an den gemeinsamen wissenschaftlichen Projekten, Teile von denen diese Arbeit ausmachen. Gerd Reis für das sorgfältige Korrekturlesen der Arbeit und die hilfreichen Verbesserungsvorschläge. Mady Gruys und Inga Scheler für die aufmunternden Worte während den kritischen Phasen meiner Arbeit.

Nicht zuletzt möchte ich meiner Familie für ihre Unterstützung fernab des Schreibtisches danken. Besonders danke ich meiner Frau Svetlana, die immer an mich geglaubt und immer Verständnis für die langen Abende im Büro hatte. Ihr möchte ich diese Arbeit widmen.

Zusammenfassung

Computergestützte Simulation und Visualisierung der akustischen Verhältnisse eines geschlossenen Raumes kann beim Design von Konzertsälen, Vorlesungsräumen, Theratern oder aber auch Wohnzimmern helfen. Denn nicht nur die visuelle Erscheinung eines Raumes sondern auch seine "Akustik" ist für eine angenehme Atmosphäre im Raum entscheidend. In Fabrikhallen ist die Lärmbekämpfung eine wichtige Aufgabe, denn Lärm ist gesundheitsschädigend.

Abgesehen von den offensichtlichen Unterschieden zwischen dem Seh- und dem Hörsinn eines Menschen, sind viele Techniken, die bei der fotorealistischen Darstellung von Bildern und Auralisation akustischer Umgebungen eingesetzt werden, einander sehr ähnlich. Beide Probleme können mit Hilfe von geometrischen Ansätzen wie Ray-tracing oder Particle-tracing gelöst werden wenn einige, weniger wichtige Effekte, vernachlässigt werden.

Mit Hilfe der Simulation wollen wir die akustischen Eigenschaften eines virtuellen Modells vorhersagen. Für die Auralisation muss eine Impulsantwort für jede Konstellation von Schallquelle und Zuhörer bestimmt werden. Die Faltung der Impulsantwort mit einem echolosen Signal ergibt das an der Zuhörerposition empfangene Schallereignis. Die Impulsantwort enthält alle reflektierten Anteile des Schalls, die durch die unterschiedlichen Materialien des Raumes beeinflusst werden.

Betreffend der Raumakustiksimulation wurde im Rahmen dieser Dissertation ein Simulationsalgorithmus, namens Phonon Tracing, entwickelt, welcher auf Schallteilchenverfolgung basiert. Das Verfahren berechnet die Energie bzw. den Druck eines Schallteilchens, das von der Schallquelle ausgesandt wird, und verwendet diese Information zur Konstruktion der Antwort auf ein Schallereignis an einer vorgegebenen Hörposition im Raum.

Während des Verfolgens der Schallteilchen, werden deren Pfade, sowie weitere Informationen in der sogenannten Phonon Map gespeichert. Unter Verwendung dieser Informationen wurden verschiedene Visualisierungsansätze für die Schallausbreitung von der Schallquelle, sowie des beim Zuhörer ankommenden Schallanteils entwickelt. Diese erlauben eine Analyse der akustischen Verhältnisse in dem zu untersuchenden Raum.

Der Simulationsalgorithmus und die Visualisierungsmethoden wurden in ein gemeinsames System integriert, welches Virtual Reality Technologien anwendet um eine bessere Immersion in die virtuelle Welt zu gewährleisten. Das System wurde als Prototyp im Rahmen eines Projektes an der TU Kaiserslautern entwickelt und befindet sich zur Zeit in Weiterentwicklung. Es beinhaltet ein stereoskopisches Rückprojektionssystem für das visuelle Rendering und eine professionelle Audioanlage zur akustischen Ausgabe im Ziele der Auralisation.

Abstract

Computer-based simulation and visualization of acoustics of a virtual scene can aid during the design process of concert halls, lecture rooms, theaters, or living rooms. Because, not only the visual aspect of the room is important, but also its "acoustics". In factory floors noise reduction is important since noise is hazardous to health.

Despite the obvious dissimilarity between our aural and visual senses, many techniques required for the visualization of photo-realistic images and for the auralization of acoustic environments are quite similar. Both applications can be served by geometric methods such as particle- and ray tracing if we neglect a number of less important effects.

By means of the simulation of room acoustics we want to predict the acoustic properties of a virtual model. For auralization, a pulse response filter needs to be assembled for each pair of source and listener positions. The convolution of this filter with an anechoic source signal provides the signal received at the listener position. Hence, the pulse response filter must contain all reverberations (echos) of a unit pulse, including their frequency decompositions due to absorption at different surface materials.

For the room acoustic simulation a method named phonon tracing, since it is based on particles, is developed. The approach computes the energy or pressure decomposition for each particle (phonon) sent out from a sound source and uses this in a second pass (phonon collection) to construct the response filters for different listeners. This step can be performed in different precision levels.

During the tracing step particle paths and additional information are stored in a so called phonon map. Using this map several sound visualization approaches were developed. From the visualization, the effect of different materials on the spectral energy / pressure distribution can be observed. The first few reflections already show whether certain frequency bands are rapidly absorbed. The absorbing materials can be identified and replaced in the virtual model, improving the overall acoustic quality of the simulated room. Furthermore an insight into the pressure / energy received at the listener position is possible.

The phonon tracing algorithm as well as several sound visualization approaches are integrated into a common system utilizing Virtual Reality technologies in order to facilitate the immersion into the virtual scene. The system is a prototype developed within a project at the University of Kaiserslautern and is still a subject of further improvements. It consists of a stereoscopic back-projection system for visual rendering as well as professional audio equipment for auralization purposes.

Contents

List of Figures

List of Tables

1 Introduction

In our daily life, we are surrounded by different kinds of sound, by pleasant music as well as by perturbing noise. Human hearing is capable of perceiving sound in a frequency range covering about 10 octaves, with a great number of distinguishable timbres. We are able to determine the direction and the approximate distance of a sound source. Based on our impression we can judge wheter a theater, a concert hall, a lecture room, or a church, has good or poor acoustics.

Room acoustics is a sub-area of acoustic science, challenging the question about good and poor acoustics of a room. More precisely, room acoustics is concerned with researching the influence of the structural condition of a building that is the influence of the geometry and the material properties on the sound propagating inside. The acoustics at a listener position inside the room is expressed by the impulse response, which is the response of a unit pulse propagated from the sound source and received by the listener. It contains the direct sound from the source as well as the sound reflected from the room boundaries. The acoustic impression is characterized by the fraction of the early reflections in relation to the total received sound, time delay and the incident direction of the early reflections, and the amount and spreading of late reverberations. The challenge is now to influence these factors by the room design in order to achieve appropriate acoustic quality according to the room usage. Thus, the goals are different depending on the usage of the room:

- Considering a recording studio, the room should not influence the recorded sound. Thus, it should be designed as non-reflective as possible.

- A class or lecture room, or a theatre should provide a good comprehensibility of speech. The fraction of the direct sound should be high. The delay of the early reflections should be low and they should reach the listener mostly from the direction of the speaker to enhance clarity and to support localization. Consequently, the reverberation time should be kept short.

- In a concert hall a spatial impression of the music experience is desirable. Therefore, the direct sound and the reflections should be well-balanced. But, the fraction of the direct peak should be high enough to perceive the music precisely. In order to achieve spatial impression, the early reflections should be strong compared to the total sound amount and well-distributed. In contrast to the lecture room the duration of the reverberations should be greater resulting in a fuller spatial impression of the music.

As one can see, it is barely possible to design a room for different kinds of usage, unless this room can be reconfigured for different presentations.

For a fairly long time, research has been done in the area of characterizing the acoustics of rooms. One of the pioneers was W.C. Sabine, who evaluated a room by means of its reverberation time which he defined as the time period of the sound pressure level decaying about 60 dB of its initial level [CM78a]. Since then, numerous studies have been made in order to measure the subjective assessment criteria of acoustical behavior of a room with objective characterization quantities [Rei68, Ber04, Thi53, LB61, Yam72, Kür69]. Some of these metrics, which can be derived from the measured or simulated room impulse response, are incorporated into the ISO 3382 norm [ISO97] (German version [DIN00]). Most of them are based on a relation of the early and late sound energy received by the listener.

A helpful tool in the room characterization is the simulation of the acoustic behaviour of the room. In the recent years the field of virtual acoustics has become more important, because the rapid development of computer technology enables complex calculations of sound propagation during the design process, improvement of the acoustics of existing rooms, or enhancement of the representation of Virtual Reality scenarios. The simulation of the sound wave propagation from the source through the scene is called *acoustic modelling*. By means of computer-aided simulation, a priori listening to music or speech inside a new or re-designed room is possible. The auditive reproduction of simulation results is called *auralization*.

Nowadays, many computer-aided algorithms for modeling and simulating acoustics inside enclosures exist. Mostly, they base on wave-theoretical or geometric acoustics. Wave-based acoustics deals with the numerical solution of the wave equation. The numerical methods often used include *Finite Difference Time Domain* (FDTD) techniques [Yee66, Bot95, SLV02], *Boundary Element Methods* (BEM) [Wu00], and *Finite Element Methods* (FEM) [Bra03, Ihl98]. These methods subdivide the domain into small elements such that the differential wave equations can be approximated by ordinary equations simplifying the solution. Since the elements must be smaller than the smallest simulated wave length, these methods are restricted to low frequencies only.

In room acoustic design and especially in virtual and acoustic reality applications, mostly methods based on geometrical acoustics are applied. They are based on optical fundamentals, and make use of approaches developed there. The *image source method* [AB79, Bor84] models specular reflections by inserting new sound sources, which are obtained by mirroring the location of audio source at polygonal surfaces inside the scene. In the *ray tracing method* [Kro68, Kul84, Vor88] several rays are traced from the sound source to listeners, which are typically represented as spheres collecting all intersecting rays. The reflections at the surfaces in the scene take place according to Snell's law of specular reflection or diffuse reflections.

Usually, Lambert's law is used for diffuse reflection calculation, but there exist approaches for modelling scattering more realistically [EAS01, CR05, TLDD07b]. Another widely used technique in acoustic computation is the *beam tracing method*. Beam tracing overcomes the aliasing problem of classical ray tracing by recursively tracing beams (i.e. sets of rays) of polyhedral [FCE$^+$98, FMC99, FTC$^+$04], conic [Ama84, MM93, Dal96], triangular [Lew93, Far95], or rectangular (frustum tracing) [LCM07a, LCM07b] cross-sections. Ray tracing and image source methods depend on the receiver position. Changing the listener position implies a recalculation of the room impulse response. In contrast, radiosity methods [SZ93, NMI04] are view independent. Thereby ideal diffuse reflections, i.e. surfaces reflecting sound uniformly in all directions, are assumed. Incorporating also specular reflections, the acoustic rendering equation [SLS07a, SLS07b] was derived for use in room acoustic simulations applying also acoustic BRDFs. Due to the computation complexity these methods do not seem practical for large environments. Due to the shortcomings of the approaches described above, continuative approaches have been developed e.g. [Vor89, Hei93, Nay93, Lew93, AFST04b, WH05]. Mostly, they employ parts of the classical schemes or a combination of them.

Approaches extending the geometric algorithm by modeling diffraction effects are proposed in e.g. [TFNC01, TSK01, PLS02, LSS02, PL03, CSF05]. Also, methods exist exploiting programmable graphics hardware (GPU) to accelerate the simulation calculations [JK04, TG97, TG98, TLDD07b, TLDD07a].

Sonel mapping [KJM04, KJM05b, KJM05a] is a particle based Monte Carlo method for geometric acoustics inspired by photon mapping (an approach for photorealistic rendering in computer graphics [Jen96]). It is a two stage sound particle based method modeling specular and diffuse reflections, as well as diffraction of the sound wave.

As mentioned before it is important to determine the amount and the direction of the received sound energy / pressure in order to judge acoustic quality at a listener position. The visualization of the sound wave behavior inside the room is a helpful tool. It is important, not only for teaching purposes, but also for engineers designing and treating rooms. Visual representation of sound propagation as well as representation of sound received at a listener position can help understanding these complex processes. The existing visualization techniques can be classified into two groups. The methods of the first group consider the propagation of the sound waves or rays inside the room independent of the listener position. Whereas the algorithms in the second group base on the measured or simulated room impulse response for individual listeners. Yokota et. al. [YST02] visualized sound propagation of 2-dimensional room sound fields resulting from a wave-based simulation using the difference time domain method. Petrausch and Rabenstein [PR05] introduce a program for simulation and OpenGL based visualization of 2-dimensional sound wave propagation in real time. Tokita and Yamasaki [TY05] visualized particle displacements on a rigid

3-D grid in a cube shaped room. The displacements resulted from a wave-based numerical simulation. Lokki [Lok02] presented in his PhD thesis a visualization using the image-source method. In [PL03] Pulkki and Lokki present an approach visualizing the phenomenon of edge diffraction. Funkhouser et. al. [FCE⁺98] used visualization of source points, receiver points, pyramidal beams, reverberation path etc. in order to understand and evaluate their acoustic modeling method. Lauterbach et. al. [LCM07b] shown the sound propagation resulting from frustum tracing. Khoury et. al. [KFW98] represented the sound pressure levels inside the room by means of color maps. Additionally, the authors analyzed the precedence effect (or "law of the first wave front") by using isosurfaces. Lokki and Nenonen [LN06] utilized cave equipment for immersive visualization of trace paths and particle paths propagating inside closed rooms. Merimaa et. al. [MLPK01] presented a visualization of measured impulse response using 2D plots in horizontal an median plane. Omoto and Uchida [OU04] visualized the arriving direction and strength of measured sound intensity using circles. Weyna [Wey05b, Wey05a] used flow visualization methods in order to study acoustic flow fields in front of and behind a thick rectangular plate. Stettner et. al. [SG89] and Monks et al. [MOD00] visualized several acoustic parameters by use of specific icons. A couple of commercial systems [ODE, CAT, Bos] provide several tools for visualizing measured or computed acoustic quantities.

Virtual Reality technologies provide an immersive representation of computer-generated scenes. Integrating acoustic simulation, visualization, and auralization into the design process aims at interactive design and immediate exploration of virtual models. Several approaches to integrate acoustic simulation into Virtual Reality systems have been proposed in recent years. The DIVA system [SHLV99, SLH02, LSV⁺02] provides sophisticated synthesis and spatialization algorithms. In [NSG02] an audio-rendering system for use in immersive virtual environments, the blue-c, is proposed. Since the applications of this system are restricted to fictional scenes, no physical-based acoustic simulation is provided. In [LN06] the authors describe a four wall VR system for immersive visualization of sound rays propagation inside closed rooms. A real-time audio system applied in a CAVE like environment with four loudspeakers is proposed in [LSVA07].

In this thesis we propose the phonon tracing algorithm, a novel simulation algorithm allowing auralization and visualization of sound in enclosed spaces. Phonon tracing is a two stage algorithm inspired by the photon mapping for realistic image rendering [Jen96]. The idea of the algorithm is the separation of the tracing (emission step) and the impulse response calculation (collection step) processes. In contrast to the traditional ray tracing approaches, where changing the listener position implies recalculation of the particle paths, only one tracing step is needed for a given sound source, followed by multiple impulse response calculation steps for different listener positions. Of course, when the source position has been changed the particle tracing has to be re-executed. As result of the emission step we obtain

the phonon map which represents all reflections of an emitted wave front by a large set of particles (phonons). The approach is capable of treating sound energy as well as sound pressure, whereas to our knowledge all other current geometric approaches are energy based. The image source method can also be expressed in terms of sound pressure, but, compared to phonon tracing, the number of considered reflections is low, due to the computational complexity. In contrast to the sonel mapping algorithm, which is also inspired by the photon mapping and is terminated based on randomly generated numbers, phonon tracing is a deterministic approach. The phonon path is terminated if the energy / pressure of the phonon falls below a prescribed threshold or the number of the phonon's reflections exceed a given number of maximum possible reflections.

Additionally, the collection step of the phonon tracing algorithm can be executed at different precision levels, depending on user requirements of computation time and accuracy. Thus a faster computation of the impulse response can be performed if lower accuracy is required, e.g. to get a first quick impression of the acoustics of the scene. To achieve this, the phonons in the map are grouped into clusters of similar trajectories representing different wave fronts and thus decreasing memory requirements of the phonon map.

The phonon map contains the path information of the sound particles. Using this information, we develop several visualization techniques of the acoustics of the scene. First, a visual representation of the propagation of the sound wave from the source through the room such that the effect of different materials on the spectral energy / pressure distribution can be observed. The first few reflections already show whether certain frequency bands are rapidly absorbed. The absorbing materials can be identified and replaced in the virtual model, improving the overall acoustic quality of the simulated room. After a large time period, a great number of reflections have occurred, such that individual wave fronts cannot be identified, anymore. The color mix of the phonons, however, shows the overall acoustic quality. Also, we present several further approaches including the visualization of individual phonons on surfaces and reflected sound wave fronts providing insight how the room influences the propagated sound waves. Furthermore, we introduce an approach to represent the direction and amount of the sound received by the listener. Moreover we propose a visual analysis of an acoustic parameter at given listener positions using the phonon tracing providing insight which surface material needs to be changed in order to enhance the value of this metric.

We integrate the phonon tracing algorithm as well as several sound visualization approaches into a common system utilizing Virtual Reality technologies to facilitate the immersion into the virtual scene. The system is a prototype developed within a project at the University of Kaiserslautern and is still a subject of further improvements. It consists of a stereoscopic back-projection system for visual rendering as well as professional audio equipment for auralization purposes.

The remainder of the thesis is arranged as follows. In the next chapter fundamentals of the sound physics, human hearing, and room acoustics are provided. Chapter 3 summarize the existing acoustic modeling techniques in more detail. The description of the phonon tracing algorithm is given in chapter 4 followed by the discussion of the sound visualization approaches developed within this thesis (chapter 5). The Virtual Reality system is introduced in chapter 6. Chapter 7 concludes this dissertation and provides an outlook on future work. Some additional information is given in the appendix.

2 Basics on sound waves and room acoustics

Acoustics (Greek *akuein* = to hear) is the science of sound. This research area is concerned with the generation, the synthesis, the propagation, the impact as well as the analysis of sound. Sound exists due to the oscillation of a medium, which could be gaseous, fluid, or solid. Human ears can perceive sound within a frequency range between 20 Hz and 20 kHz, called audible sound. Sound below this range is called infrasonic, and above ultrasonic.

The sound fields appearing in the nature can in general not be calculated, exactly. There exist different methods and approaches for the approximation of this natural phenomenon. In *wave-based* acoustics the wave fronts are expressed as plane, spherical, cylindrical waves of diffuse sound fields, which can be calculated with satisfactory accuracy [FKS84a]. *Geometrical acoustics* examine single sound rays spreading out from a source. *Statistical acoustics* deals with the observation of several sound portions using statistical methods. In the area of *psychological acoustics* the research concentrates on the question, how humans perceive sound. [CM78a].

In the field of *room acoustics* researchers investigate the problem of improving acoustical behavior inside closed rooms. Among other areas, this implies improvement of speech intelligibility inside lecture halls, optimization of theatres or concert halls for speech and music presentations, and damping of noise inside factory halls. The major question appears, how can we define "good" or "bad" acoustics, or how can the acoustical quality of a given room be measured.

In the recent years the field of virtual acoustics became important pushed by the rapid development of computer technology. This allows complex calculations of sound propagation also during the design process.

The following chapter gives an overview of the basics of sound waves, the sound propagation in closed rooms, the theory of room acoustics, as well as the basics of the possible evaluation of the acoustics inside a room.

2.1 Phenomena of sound waves

In the following a survey of the general laws and phenomena regarding the sound wave behavior is provided.

2.1.1 Sound wave

A wave can be defined as a disturbance that propagates within a medium with some velocity, witch depends on the medium, and transfers energy [BS95a]. Depending on the direction of the wave oscillation relative to to the direction of the wave propagation, two different types of waves can be classified: *transverse* and *longitudinal*. A *transverse* wave oscillates perpendicular, or transverse, to the propagation direction. Examples of transverse waves are rope waves or electromagnetic waves, which are unique in that they require no medium in which to propagate. In longitudinal wave the motion of the particles of the medium, forming the wave, is in the same direction as the direction of the wave propagation. The sound waves belong to this type of waves.

In room acoustics the sound fields can be considered as a superposition of single simple sound waves, e.g plane and spherical waves. Furthermore, we only deal with air as the medium, therefore we can restrict our considerations to sound propagation in gases. The sound waves are described by sound pressure p and sound particle velocity v, which both are specified depending on time and position. Pressure and particle velocity, as well as material density ρ, are calculated using the three basic acoustic equations, the *Euler equation*, the *continuity equation*, and the *constitutive equation (ideal gas law)* [Sku71]. Using this basic equations the wave equation is derived as:

$$\Delta p = div \nabla p = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \tag{2.1}$$

where c is the velocity of sound that, by the assumption of a homogeneous medium being at rest, can be calculated depending on the temperature τ as [Kut00]:

$$c \approx (331.4 + 0.6\tau) \ \left[\frac{m}{s} \right] \tag{2.2}$$

In large rooms, the variation of temperature and hence of the sound velocity can not be entirely avoided. But the effects which are caused by this certainty are small and thus may be neglected.
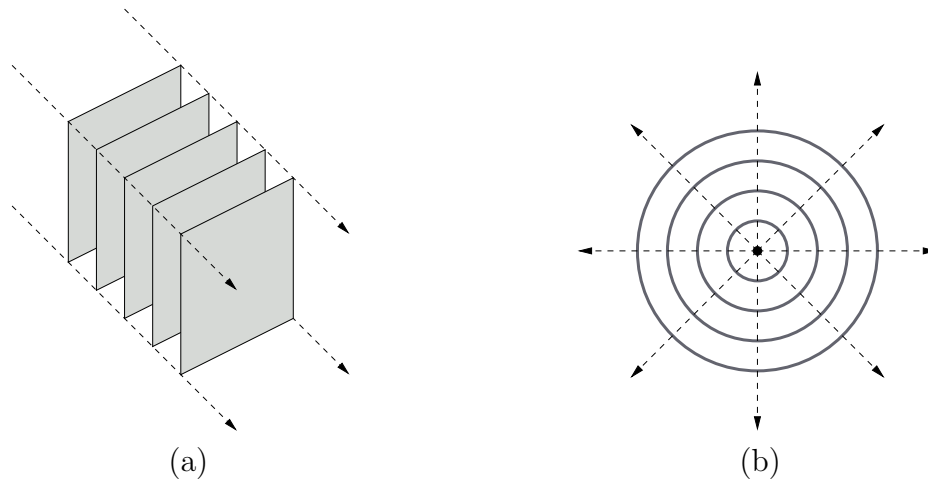
Figure 2.1: Simple wave types: (a) plane wave; (b) spherical wave

Plane waves. A plane wave is a progressive wave, where the sound pressure p is constant in any plane perpendicular to the propagation direction (see figure 2.1 (a)). These planes of constant pressure are called *wave fronts*. The time and space dependence of the acoustical quantities is described by a sine or cosine functions [Kut00]. The progressive harmonic plane wave reads:

$$p(x,t) = \hat{p}\,cos(\omega t - kx) \tag{2.3}$$

where $\hat{p}$ denotes the wave amplitude, $\omega = 2\pi f$ the angular frequency, and $k = \frac{\omega}{c}$ the wave number. The above equation describes a spatial harmonic wave with the period $\lambda = \frac{2\pi}{k}$, which is the *wave-length* of the harmonic wave. Within harmonic waves the pressure p and the particle velocity v are in phase [FKS84a].

Spherical waves. Another type of simple sound waves are the spherical waves. A spherical wave is a sound wave that is propagated uniformly in all directions as concentric wave fronts [Rie82]. Figure 2.1 (b) depicts a schematic representation of a spherical wave. Here, the surface of constant pressure are concentric spheres. In the center of this wave fronts we have a *point source*. The pressure of the spherical wave decreases proportional to the traversed distance, $p \sim \frac{1}{r}$, the energy proportional to the squared distance, $e \sim \frac{1}{r^2}$. The harmonic spherical wave reads:

$$p(x,t) = \frac{\hat{p}}{r}\,cos(\omega t - kx) \tag{2.4}$$

Another simple wave type is the *cylindrical wave*. For a description of the cylindrical wave as well as for the detailed derivation of the wave equations introduced above see [Kut00, FKS84a, Rie82]. The equations 2.3 and 2.4 are derived assuming

that the medium is free of losses. By sound propagation in air we have to take the air absorption into account (see section 2.1.2.4).

A plane wave in its pure form does not exists in the real world. However, a limited region of a spherical wave can be considered as a good approximation for a plane wave. On the other hand, a point source producing a spherical wave can be realized by a sound source which is small compared to the wave length. However, most natural sound sources do not behave like spherically symmetric point sources. The sound pressure depends also on the emission characteristics of the source.

2.1.2 Basic laws and sound propagation

2.1.2.1 Huygens's principle

Huygens's principle, due to the Dutch physicist Christian Huygens (1629–1695), says that every point on a wave front acts as a source of a spherical wave [Gri91], which have equal propagation velocity and frequency as the original wave. This individual Huygens's wavelets sum up to the total wave front at succeeding time. Figure 2.2 depicts how this principle is applied to the propagation of a plane and a spherical wave.

(a)   (b)
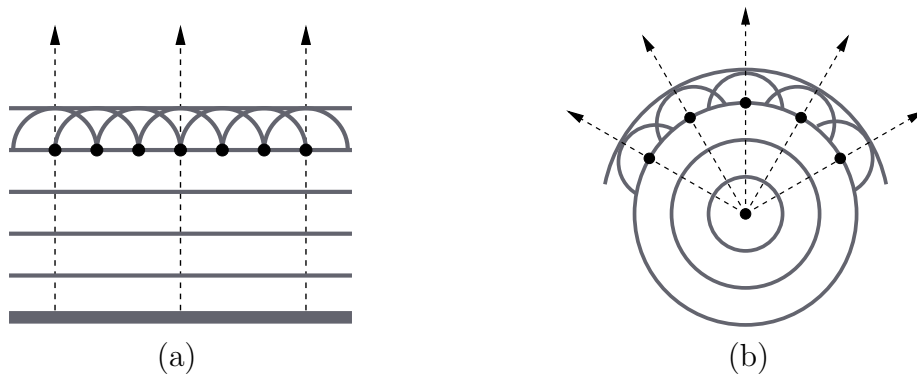
Figure 2.2: Huygens's wavelets along plane (a) and spherical (b) waves.

On the basis of the Huygens's principle the propagation of waves can be interpreted. Phenomena like refraction, diffraction or interference can be explained using Huygens's wavelets. A result of this wave behavior is that the wave propagates always perpendicular to the wave front.

2.1.2.2 Reflection

A reflection occurs if the sound wave bounce off a surface, e.g. off a wall inside a room. If the surface is "smooth" then the sound wave reflects *specularly*. Thereby is the incident angle equal to the reflected angle with respect to the surface normal. Figure 2.3 (a) depicts this phenomenon. The "smoothness" of the surface is measured depending on the wave-length. Are the dimensions of the surface variation significantly smaller compared to the wave-length, the wave is specularly reflected. Sound waves in the low frequency range, have great wave-length, and thus are specularly reflected on almost all surfaces. When the surface variations are located in the same order of magnitude as the wave-length, the wave is reflected in all directions even if the incident wave front comes from a particular direction. This type of reflection is called *diffuse* reflection, and surfaces that reflects the sound wave diffusely are called *diffuse surfaces* [Cow94, CM78a]. An example of diffuse reflection is illustrated in figure 2.3 (b).
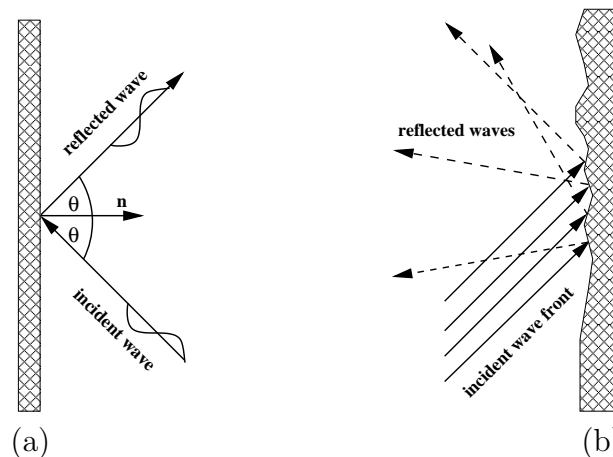
Figure 2.3: Reflections at surfaces: (a) specular; (b) diffuse.

2.1.2.3 Refraction

On the boundary of two different media the wave is bending as it passes between these two media. This phenomenon is called *refraction*. Acoustic refraction satisfies Snell's law [Cow94]:

$$\frac{c_2}{c_1} = \frac{sin(\theta_t)}{sin(\theta_i)} \tag{2.5}$$

where c_2 is the sound velocity in the transmitted medium, c_1 the speed of sound in the incident medium, θ_t the angle of transmission, and θ_i the angle of incidence, both

with respect to the line perpendicular to the boundary. Refraction is characterized by the refraction coefficient τ [Rie82], which is defined as the ratio between the transmitted sound intensity I_t and the incident sound intensity I_i:

$$\tau = \frac{I_t}{I_i} \tag{2.6}$$

The sound wave changes its propagation direction when travelling into a medium that conducts sound at a different velocity. This is the case if the sound wave travels into a different medium or in a region of the same medium where the conditions have been changed. A sound wave traveling through the atmosphere changes the velocity and direction with change of temperature, humidity and wind-velocity, since the speed of sound depends on these factors. Figure 2.4 illustrates the refraction of waves.
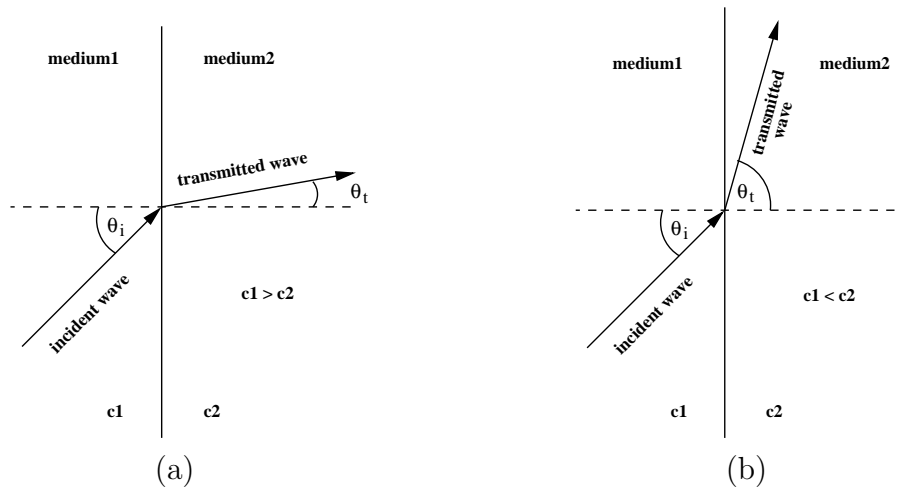
Figure 2.4: Refraction of waves.

2.1.2.4 Reflection factor and absorption coefficient

Depending on the surface material properties, only a part of the energy and pressure of the sound wave is reflected back. The residual part is refracted or absorbed by the material as a result of the transformation of sound into heat (sound dissipation) [FKS84a]. Thus, the sound wave undergoes a change in amplitude and phase. These changes which takes place during the reflection of a wave are expressed by the complex *reflection factor* [Kut00]:

$$R = |R|\, exp(i\chi) \tag{2.7}$$
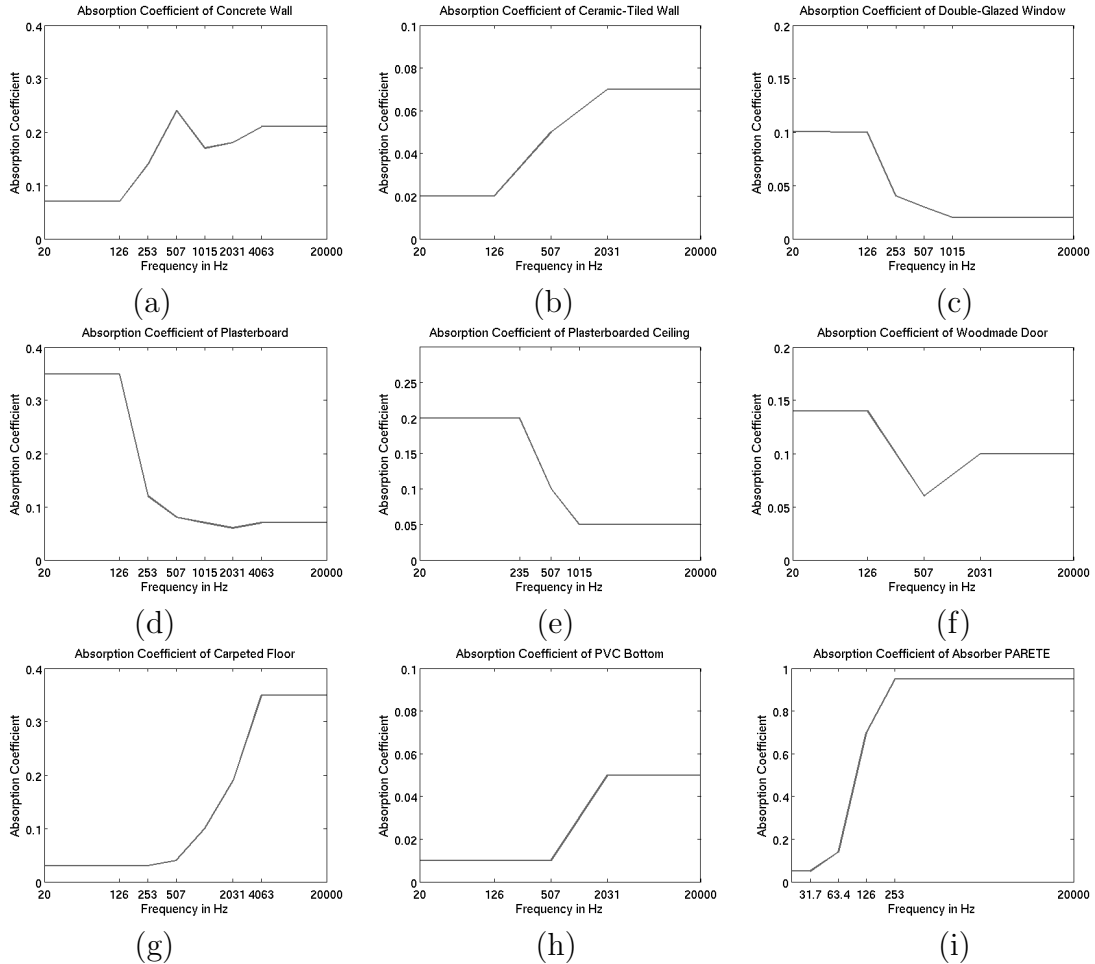
Figure 2.5: Absorption functions for different materials.

which is a property of the surface. The absolute value of R as well as its phase angle χ depends on the frequency and on the direction of the incident wave.

The intensity of a wave is proportional to the squared pressure amplitude [Kut00]. Consequently, the intensity of the reflected wave is reduced by the factor $\rho = |R|^2$ (the *reflection coefficient*) compared to the incident wave. The fraction

$$\alpha = 1 - |R|^2 \tag{2.8}$$

of the incident energy is lost. This quantity is called the *absorption coefficient*. The absorption coefficient has its maximum value $\alpha = 1$ for a surface with zero reflectivity. The surface is called *totally absorbent*. The surface is *hard* or *rigid* for $R = 1$, and *soft* for $R = -1$. In both cases there is no absorption, thus the surface is *totally reflective*. In room acoustics, the latter case is not common and occurs only in limited frequency ranges. Figure 2.5 shows examples of the absorption functions for

different materials provided by CARA[1]. The reflection factor specifies the acoustical properties of all surfaces for all frequencies and for all angles of incidence completely. For the reflection coefficient ρ, the refraction coefficient τ, and the dissipation coefficient δ we can draw the following energy balance [FKS84a]: $\rho + \tau + \delta = 1$.

Air absorption. In the field of room acoustics we deal with the sound propagation in air. Air, as a medium, absorbs sound energy. The amount of the absorption is not only frequency dependent, but varies with the relative air humidity, air temperature, as well as atmospheric pressure. The attenuation caused by the air absorption follows the exponential law [Kut00]:

$$I(r) = I_0 \, exp(-mr) \tag{2.9}$$

where I_0 is the original sound intensity of the source, r the wave traversed distance, and m the attenuation coefficient. Air attenuation coefficients at 20° and normal atmospheric pressure are listed in table 2.1. These values result from calculations presented in [BSZ90].

Relative	Frequency in kHz						
humidity in %	0.5	1	2	3	4	6	8
40	0.60	1.07	2.58	5.03	8.40	17.71	30.00
50	0.63	1.08	2.28	4.20	6.84	14.26	24.29
60	0.64	1.11	2.14	3.72	5.91	12.08	20.52
70	0.64	1.15	2.08	3.45	5.32	10.62	17.91

Table 2.1: Air attenuation coefficient at 20° and normal atmospheric pressure in $10^{-3} \, m^{-1}$ [Kut00].

2.1.2.5 Diffraction

Refraction of sound waves is caused by the change of sound velocity. Sound waves do not need a change of the medium to be bent. This bending without a change of the medium is called *diffraction* [BS95a]. Sound can be heard around corners or behind obstacles occluding the sound source. The amount of diffraction depends on the wave length and the size of the obstacle. If the obstacle is small compared to the wave length, the propagation seems to be unhindered by the obstacle. On the other hand, behind large obstacles no sound is observed [CM78a]. Consequently, low-frequency sound waves will spread over a larger angle around a corner than high-frequency waves. This is the reason why voices sound somewhat dull when heard from next room. An illustration of the diffraction effect is shown in figure 2.6.

---

[1]www.cara.de
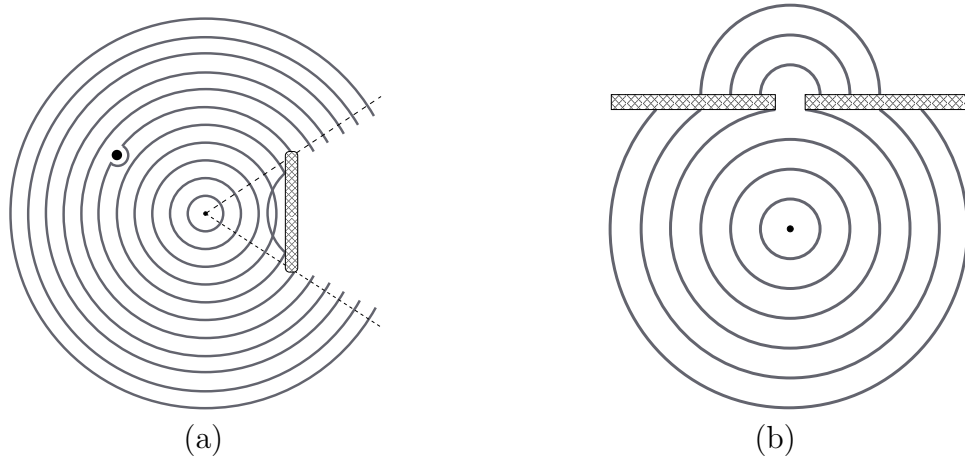
Figure 2.6: Diffraction of waves: (a) on a small and a large obstacle; (b) on a gap.

2.1.2.6 Superposition and interference

The principle of superposition denotes that two waves do not effect each other, even if they are in the same place at the same time [BS95a]. In other words, the waves add algebraically, such that the displacement of the sum of the waves is equal to the displacement of the first wave added to the displacement of the second wave at the same point and time. This law follows directly from the linearity of the wave differential equation.

This superposition leads to certain patterns when adding waves of equal frequency. Combining or adding of two or more similar waves is called *interference*, and has a special meaning in the physics of wave. Two different types of interference are defined:

1. *Constructive interference*: Superposition of waves with equal angular frequency in phase such that the waves enhance each other.

2. *Destructive interference*: Superposition of waves with equal angular frequency but of complementary phase resulting in an effective disappearance of the waves.

An example is the interference of two identical point sources, shown in figure 2.7.

2.1.2.7 Standing waves

A special type of interference phenomena are *standing waves*. A standing wave appears by superposition of two waves with equal frequency and phase propagating in opposite direction [BS95a]. The amplitudes of the waves need not necessarily match. The waves may be generated from two different sources or originate from a
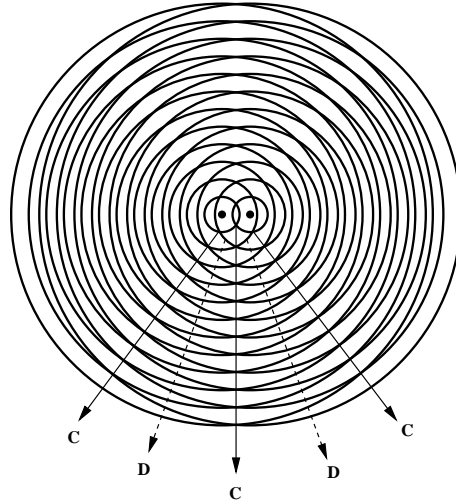
Figure 2.7: Interference between two identical point sources. C labels constructive interference and D destructive interference.

reflection at a surface. Characteristically for every standing wave patterns is that there are points that appear to be standing still (therefore the name "standing" wave). These points are called *nodes* [BS95a]. There are other points along the medium which undergo oszillations between a large positive and and large negative displacement. In a sense, these points are the opposite of nodes, and so they are called *antinodes*. Figure 2.8 depicts an example of a standing wave, where nodes and antinodes are labeled with N and A, respectively. The nodes are the result of a destructive interference of the two waves, whereas the antinodes arise from a constructive interference. The distance between two nodes (or two antinodes) corresponds to the half of the wave length.

The simplest standing wave occurs at the lowest vibration frequency and is called *fundamental* or *first harmonic* [BS95a]. Other standing waves, called *harmonics*, are identified by their harmonic number h and have the frequency h times the fundamental frequency (see figure 2.8). These different types of oscillation are called *modes*. Particularly in wave-based acoustics the modes (eigenmodes) are important, since they can be calculated as a solution of the homogeneous wave equation with boundary conditions [FKS84a]. This subject will be discussed in section 2.2.1. Inside rooms, standing waves cause changes in the timbre, since certain tones are emphasized.
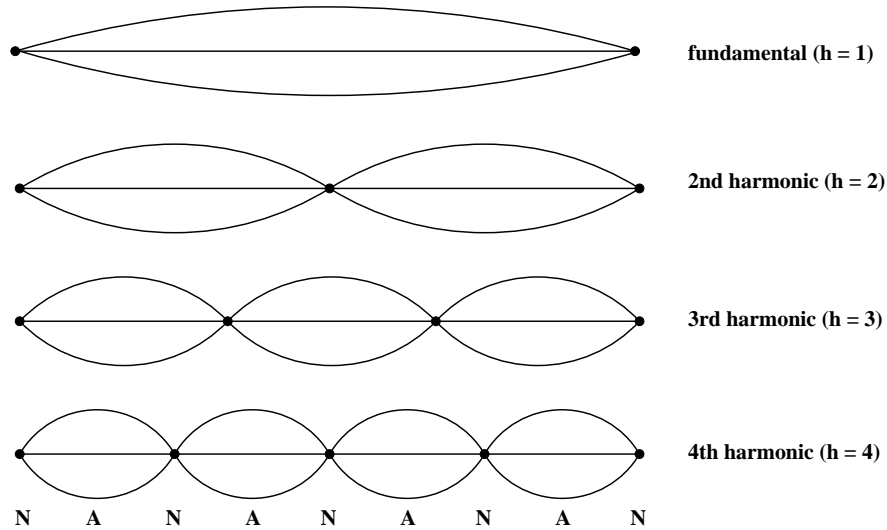
Figure 2.8: Standing waves.

2.2 Room acoustics

One may think that the loudspeakers of a home hi-fi unit are the most important reason for a good sound, pleasant to our ears. But, it's not that simple. There is by far more as what is coming out of the loudspeakers. What we hear is on the one hand the direct sound from the speakers and on the other hand the indirect, reflected, sound that is bounced on the walls, ceiling, floor and other objects inside the room, and reaches our ears at different times. This reflected part can be useful and bothersome as well. The advantages of the reflections are that music and speech appear fuller and louder as usually in the free space (dull and thin). Furthermore, the reflections can induce comfortable spaciousness. However, the frequency dependent absorptions at room walls cause certain notes sound louder while canceling out others. The result could be e.g. a boomy bass, a harsh sounding treble, or echos as well. Hence, we should consider the room as the last link in our system chain. The science of room acoustics deals with designation, evaluation and improvement of acoustic quality inside rooms.

2.2.1 Wave-based acoustics

In the field of the wave-based room acoustics the main intention is the solution of the wave equation (2.1) introduced in this section. Assuming a periodical behaviour of the pressure (harmonic wave) the wave equation then reads [Kut00]:

$$\Delta p + k^2 p = 0 \tag{2.10}$$

where $k = \frac{\omega}{c}$ is the wave number, with the angular frequency $\omega = 2\pi f$, and Δ is the Laplacian operator. This equation is known as the *Helmholtz equation.*

Locally reacting room surfaces (surfaces with impedances independant on the incident sound [Kut00]) imply the following boundary condition:

$$\zeta \frac{\partial p}{\partial n} + ikp \tag{2.11}$$

where ζ is the specific impedance of the surface (see appendix A), and $\frac{\partial}{\partial n}$ denotes the partial differentiation in the direction of the surface normal n. Given the boundary condition the wave equation has non-zero solutions only for particular values k_n of k, the *eigenvalues* (see [Kut00]). Each eignevalue k_n is assigned to a solution, the *eigenfunction* $p_n(r)$ $r \in \mathbb{R}^3$, which is a three dimensional standing wave (sec 2.1.2.7), a so-called *normal mode* or *eignenmode* of the room.

In closed form the eigenfunctions can be represented only for few room shapes. An important example of such simple shaped rooms is a rectangular room with the dimensions l_x, l_y, l_z, depicted in figure 2.9. It is important since many concert halls, lecture rooms, churches etc. are close to the rectangular shape, and therefore the results obtained for a rectangular room can be applied to many existing rooms.
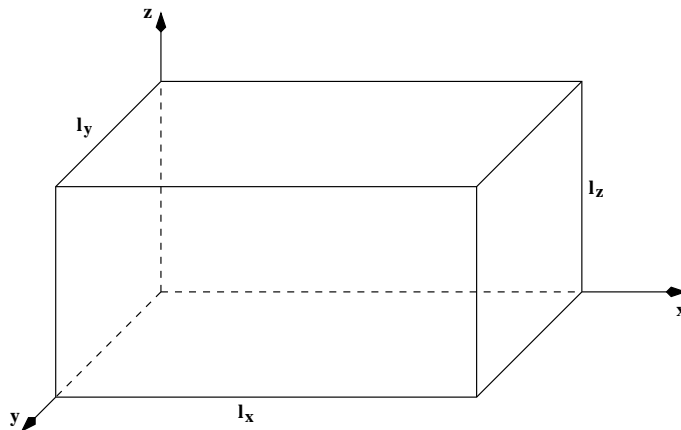
Figure 2.9: Rectangular shaped room.

The *Helmholz* equation 2.10 in Cartesian coordinates in $\mathbb{R}^3$ reads:

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} + k^2 p = 0 \tag{2.12}$$

The variables can be separated in three parts which only depend on x, y, and z, respectively. The *Helmholtz* equation is then split up into three ordinary differential equations for x, y, and z with equivalent boundary conditions. The solution of the wave equation 2.12 for discrete eigenvalues is given by [Kut00]:

$$p_{h_x,h_y,h_z}(x,y,z) = C \cos\left(\frac{h_x \pi x}{l_x}\right) \cos\left(\frac{h_y \pi y}{l_y}\right) \cos\left(\frac{h_z \pi z}{l_z}\right) exp(i\omega t) \tag{2.13}$$

where C is a constant that can be chosen. The eigenvalues of the wave equation can be calculated as follows:

$$k_{h_x,h_y,h_z} = \pi \sqrt{\left(\frac{h_x}{l_x}\right)^2 + \left(\frac{h_y}{l_y}\right)^2 + \left(\frac{h_z}{l_z}\right)^2} \tag{2.14}$$

The eigenfrequencies corresponding to the eigenvalues are given by [Kut00, CM78b]:

$$f_{h_x,h_y,h_z} = \frac{c}{2\pi} k_{h_x,h_y,h_z} = \frac{c}{2} \sqrt{\left(\frac{h_x}{l_x}\right)^2 + \left(\frac{h_y}{l_y}\right)^2 + \left(\frac{h_z}{l_z}\right)^2} \tag{2.15}$$

The formula in equation 2.13 represents the three-dimensional standing wave (see section 2.1.2.7). The numbers h_x, h_y, and h_z are the harmonic numbers in x, y, and z directions, respectively. These values indicate the number of nodal points described in section 2.1.2.7, in this case of course the *nodal planes*, where the pressure becomes zero. This pressure vanishing occurs for all values of x which are odd integers of $l_x/2h_x$ and for equivalent values of y and z. The nodal planes are equidistant and mutually orthogonal [Kut00].

The eigenmodes can be divided into three different classes depending on their propagation direction [MEG74]:

1. *Axial modes:* The propagation direction is parallel to one of the room edges. Two of the harmonic numbers h_x, h_y, h_z are equal to zero. These modes are induced by a single wall pair.

2. *Tangential modes:* Induced by two wall pairs. The propagation direction is perpendicular to one of the room edges. One of the harmonic numbers is equal zero.

3. *Oblique modes:* Induced by all wall surfaces of the room. The wave strikes all three wall pairs oblique. The harmonic numbers are all non-zero.

The amount of the eigenfrequencies from 0 to an upper limit f of the room can be approximated by [Kut00]:

$$N_e = \frac{4\pi}{3} V \left(\frac{f}{c}\right)^3 + \frac{\pi}{4} S \left(\frac{f}{c}\right)^2 + \frac{L}{8} \frac{f}{c} \tag{2.16}$$

where $S = 2(l_x l_y + l_x l_z + l_y l_z)$ is the area of all walls and $L = 4(l_x + l_y + l_z)$ the sum of all edge lengths. This equation is also valid for rooms of arbitrary shapes.

The eigenfrequencies can be strictly separated at relatively low frequencies only. With increasing frequency the corresponding eigenmodes interfere more and more with each other. The condition for almost complete interference is given by [MM04, SK62]:

$$\frac{V f^2}{T} > 4 \cdot 10^6 \left(\frac{m}{s}\right)^3 \tag{2.17}$$

here T is the reverberation time introduced in section 2.2.3. The corresponding limiting frequency

$$f_s = 2000 \sqrt{\frac{V}{T}} \tag{2.18}$$

is called *Schroeder frequency*. The number of the eigenfrequencies below the Schroeder frequency is calculated as [MM04]:

$$N_e(f < f_s) = 850 \sqrt{\frac{T^3}{V}} \tag{2.19}$$

For the frequencies exceeding the Schroeder frequency the description of the sound field using the superposition of the eigenfrequencies becomes complicated even for simple room shapes.

For a more complicated case of non-rectangular rooms we have to fall back on the numerical solutions of the wave equations using e.g. the Finite Element Method (FEM). This subject will be discussed in section 3.1.

2.2.2 Geometric acoustics

Waves traverse the path between two points which takes the least time (*Fermat* principle [CM78a]). In a medium with constant propagation velocity this path is also the shortest, thus at free propagation a line.

In geometric acoustics the concept of a wave described in the previous section is replaced by the examination of single sound rays. A sound ray is regarded as an evanescent part of a spherical wave emitted from a point source. The rays travel along a well defined propagation direction, and are subject to the same laws of propagation as the light rays in optics, except for their different velocities. The energy decreases proportional to the squared traversed distance, where the pressure amplitude decreases proportional to simply the distance. The rays emitted from the source do not necessarily need to carry the same amount of energy, thus modeling the directivity of the source (human voice for example).

The simple linear propagation of the sound ray is disturbed by obstacles or by changes inside the medium. In this case phenomena like reflection, diffraction, diffraction etc. occur, as discussed in section 2.1. Most notably in room acoustics is the law of reflection including the frequency dependent sound energy absorption at room walls, whereas refraction does not occur in room acoustics [Kut00]. Furthermore, diffraction and interference effects have to be modeled using different approaches.

Hence, geometrical acoustics deals with the examination of the propagation of single sound rays from a source to a particular listener position. An additional aspect that has to be taken into account is the propagation velocity of sound, since it is responsible for such important effects as reverberation, echoes, etc. Suppose the source emits a single pulse, the Dirac pulse (see appendix B). In reality this is not possible, since acoustic pressure signals are zero on average. At the listener position arrive successively different sound rays: first the Dirac itself as *direct peak* from the source, after that the sound rays which are reflected once or multiple times at the room walls. By-and-by the rays carry less and less energy as a result of the energy decay during the propagation as well as absorption at room surfaces. The first strong reflections are called *early reflections*, whereas the numerous weak reflections arriving at later time are called *late reverberations*. The role of the early reflections with respect to our subjective hearing (see section 2.2.4) is different from that of the late reverberation. However, strong reflections arriving at later time are perceived as echos (see figure 2.10 (b)). Particularly, periodic sequences of echos, the *flutter echos* [MEG74] are displeasing.

When the intensities of all the rays are plotted on the time axis the so called *echogram* [Kut00] is obtain. A sketch of an echogram is depicted in figure 2.10. Using the echogram diverse acoustic metrics can be derived in order to evaluate the acoustic behavior of the room. A couple of these metrics are introduced in section
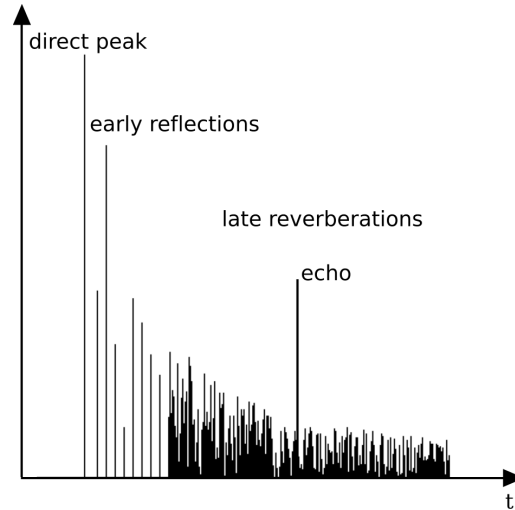
Figure 2.10: Sketch of a pulse response (echogram).

2.3. Furthermore, by means of the echogram the calculation of the *room impulse response*, i.e. the pressure response of the Dirac pulse, is feasible. The convolution of the impulse response with an anechoic signal results in a signal perceived at a given listener position.

However, the geometrical approach has its limitations. The consideration of the sound wave as a set of rays is permitted only if the dimensions of the room and its surfaces are large compared with the wave length of sound. For low frequencies the effects of diffraction and interference become more and more important. An example how complicated the relation between the wave length and room dimensions are, is shown in the following example [CM78a]. We consider the reflection of a sound wave at a structured wall (see figure 2.11). At low frequencies (wave length greater than the wall structures, figure 2.11 (a)) the wave is reflected perfectly specular. Are the wall structures at the same order of magnitude as the wave length (middle frequency, 2.11 (b)) the reflection is perfectly diffuse. Whereas, at high frequency range (small wave length compared to the wall structure, figure 2.11 (c)) each of the wall faces reflects the wave specular.

Since the computation complexity increases with increasing number of reflections and diffusion, statistical methods (section 2.2.3) have to be taken into account for examination of late reflections.

Despite the limitation of geometric acoustics, this concept is of great practical importance, particularly for architectural design of large rooms, such as concert halls, theaters or lecture rooms, where the room dimensions are large with respect to the wave length. Furthermore, geometric acoustics allows the examination of the acoustical behavior in rooms of arbitrary shape, whereas the wave-based methods
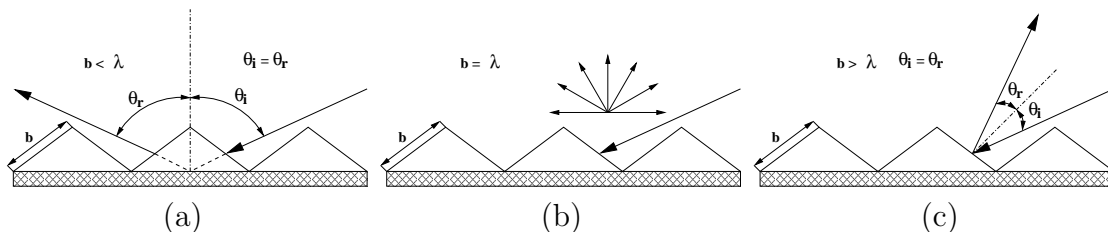
Figure 2.11: Reflection of a sound wave at a structured wall. (a) Low frequent wave, (b) Middle frequent wave, and (c) high frequent wave.

fail due to the complexity of the sound wave distribution. Many recently developed room acoustic modeling algorithms are based on geometric acoustics (see section 3.2).

2.2.3 Statistical acoustics

A sound wave emitted from the source undergoes multiple reflections on the walls of a room. Ideally, the sound field is equally distributed in all directions. The sound field becomes then diffuse i.e. the energy density is uniformly distributed in the room [MM04] such that it can be described statistically. Hence, the name *statistical acoustics* which mainly deals with the temporal characteristic of sound field quantities inside rooms. The intention is the description of the reverberation which is an important phenomenon in room acoustics. We consider here the sound wave as a set of sound particles or small energy packets traveling with constant velocity on straight lines. Forming a power balance, simple laws of the sound-field decay as well as the sound-field energy density can be derived.

In statistical room acoustics the shape of the room is irrelevant. In fact, the room volume, the total surface area of room faces, and the absorption of the surfaces are important. Different walls have different absorption coefficients. From a statistical point of view the mean absorption coefficient $\bar{\alpha}$ is used:

$$\bar{\alpha} = \frac{1}{A} \sum_{i=1}^{n} A_i \alpha_i \qquad (2.20)$$

where A_i is the surface area and α_i the absorption coefficient of room surface i. A is the total surface area given by:

$$A = \sum_{i=1}^{n} A_i \qquad (2.21)$$

The sound field decays with the increasing number of reflections at walls. We can

derive the mean number of reflections as [MEG74]

$$\bar{n} = \frac{Ac}{4V} \tag{2.22}$$

where V is the volume of the room and c the sound velocity. The total energy of the sound field after a time period t is calculated as follows [Kut00]:

$$E(t) = E_0(1 - \bar{\alpha})^{\bar{n}t} = E_0(1 - \bar{\alpha})^{\frac{Ac}{4V}t} = E_0 \exp\left(\frac{Ac\,\ln(1 - \bar{\alpha})}{4V}t\right) \tag{2.23}$$

E_0 is the initial value of the sound energy after a short impulse or after the sound source is paused promptly. The energy density and consequently the mean sound pressure in a room decay exponentially in time. This is the general law of sound decay.

The reverberation process is characterized by the *reverberation time.* According to W.C. Sabine (1868–1919) the reverberation time is the time interval where the sound energy decays to 10^{-6} of its initial value [CM78a]. In practice, the sound pressure level is measured. Therefore the reverberation is also referred to as the time period where the sound level pressure decreases by 60 dB [Rie82]. Sabine empirically found the following formula for the calculation of reverberation time based on the volume, the total surface area and the absorption properties of the room:

$$RT_{60} = 0.163\,\frac{V}{\bar{\alpha}A} = 0.163\,\frac{V}{\sum_{i=1}^{n}\alpha_i A_i}\quad[s] \tag{2.24}$$

C.F. Eyring (1889–1951) confirmed Sabine's experiments theoretically. The reverberation time can be derived from the equation 2.23 including air absorption. The Eyring formula for reverberation time reads [MM04]:

$$RT_{60} = 0.163\,\frac{V}{4mV - A\,\ln(1 - \bar{\alpha})}\quad[s] \tag{2.25}$$

where m is the air attenuation coefficient. If the absorption coefficient $\bar{\alpha}$ is small ($\bar{\alpha} \ll 1$) then it holds $\ln(1 - \bar{\alpha}) \approx -\bar{\alpha}$ and the Eyring formula is transformed into Sabine's formula. The reverberation time can be computed easily and quite accurate. For a long time period it was the primary acoustic metric.

Schroeder [Sch65, Sch79] proposed a method to measure the reverberation time. Using the MLS pseudo-random noise the room impulse response is determined. The integration of the impulse response yields a decay curve (see figure 2.12). The reverberation time is then calculated by use of the slope of the decay curve.

Another significant quantity in statistical room acoustics is the *critical distance* (or *reverberation radius*). The sound field in a room can be described by the direct
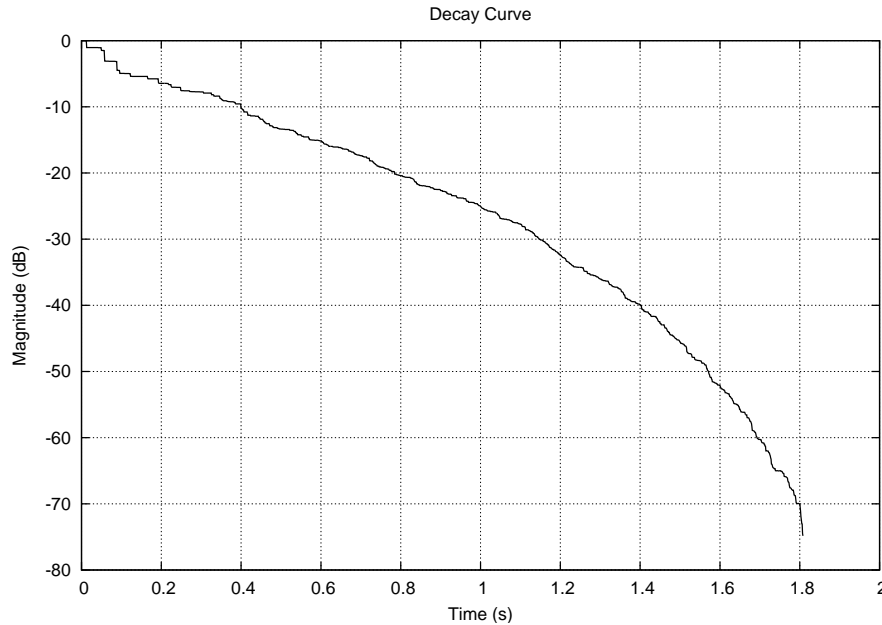
Figure 2.12: Decay curve illustrating reverberation.

field from the source superposed with the diffuse field (stationary field) induced by the room. The critical distance is the distance from the source where the energy density of the direct field is equal to that of the diffuse field. For a point source, i.e. spherical emission, the critical distance can be calculated as follows [MM04]:

$$r_h = 0.057 \sqrt{\frac{V}{RT_{60}}} \qquad (2.26)$$

In case of directional emission, the critical distance in the direction of the main emission is greater compared to the residual directions. At distances $r < r_h$ the energy density of the direct field is greater than that of the diffuse field, and less than the energy density of diffuse field at $r > r_h$. The critical distance increases with increasing room volume V and decreases with the increasing reverberation time RT_{60}.

Optimal reverberation time of a room depends on its functionality [MEG74]. Because of the comprehensibility smaller reverberation time for speech presentations is preferred as for music shows. Table 2.2 includes guide values for different rooms.

The statistical model applies to large rooms in which the stationary field dominates the properties of the room.

cabaret	spectacle or talk	chamber music	opera	concert	organ music
0.8	1.0	1.4	1.3 ... 1.6	1.7 ... 2.0	2.5

Table 2.2: Guide values of reverberation in [s] for the frequency range 500 – 1000 Hz (corresponding to [MEG74]). .

2.2.4 Psychological acoustics

In the room acoustic theories described in previous sections we did not consider a listener the room is actually designed for. Psychological acoustics treats the correlations between the human sound perception and physical parameters of sound. First we review some facts on human hearing.
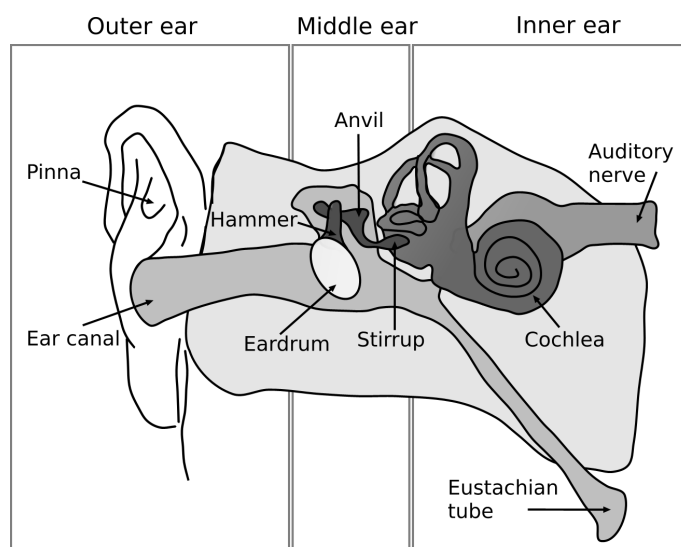
Figure 2.13: Human ear.

The sense of hearing represents the most direct contact to our environment, since it can not be shut down like vision by closing the eyes. The human ear consists of three parts (see figure 2.13): the outer ear (with pinna and ear canal), the middle ear (with eardrum and ossicles: hammer, anvil, stirrup), and the inner ear (with cochlea and auditory nerve) [Gel04]. Sound waves raise the eardrum to vibrate. The ossicles pass this oscillations on to the cochlea, the actual hearing organ. The vibrating motions of the ossicles produces a wave motion on the membrane that runs the full length of the cochlea. Super-sensitive hair cells are situated on top of the moving membrane. When a hair cell is bent because of membrane motion, the nerve fires and the message is transmitted to the brain over the auditory nerve. If the vibrations are fast (high-frequency sound), the membrane has its greatest motion near the vibrating ossicle. If the vibrations are slow (low-frequency sound),

the maximum membrane motion occurs at the tip of the cochlea. Thus hair cells near the ossicles transmit information about high-frequency, while those at the tip provide information about low-frequent sound. The chochlea performs a frequency analysis of the sound. Some more details about the frequency differentiation of human hearing can be found in [Gör06].

The lowest perceivable frequency of human ear is about 20 Hz, the highest about 20 kHz. The just noticeable pressure is $p_0 = 2 \cdot 10^{-5}$ Pa, which corresponds to a sound pressure level L_p of 0 dB:

$$L_p = 20 \lg \frac{p}{p_0} \tag{2.27}$$

However, human hearing senses tones of equal pressure level not equally loud [CM78a]. Figure 2.14 (a) depicts curves of equal loudness. Thus, the just perceivable pressure threshold holds for the frequency range between 1000 and 4000 Hz. For the frequencies below and above this range the hearing threshold is higher (see figure 2.14 (b)). The upper curve in figure 2.14 (b) depicts the threshold of pain. Sound pressure levels above this threshold can cause damage. Concerning the fact that between the hearing threshold and the threshold of pain there is a difference of about 134 dB (ratio of about $1 : 10^{13}$) we can see that our hearing system is very powerful. The perceivable frequency range of 10 octave is remarkable, too, compared to less than 1 octave of visible light. However, the speech use only a small part of this "hearing area" (see figure 2.14 (b)). Even music does not span the entire range, generally.
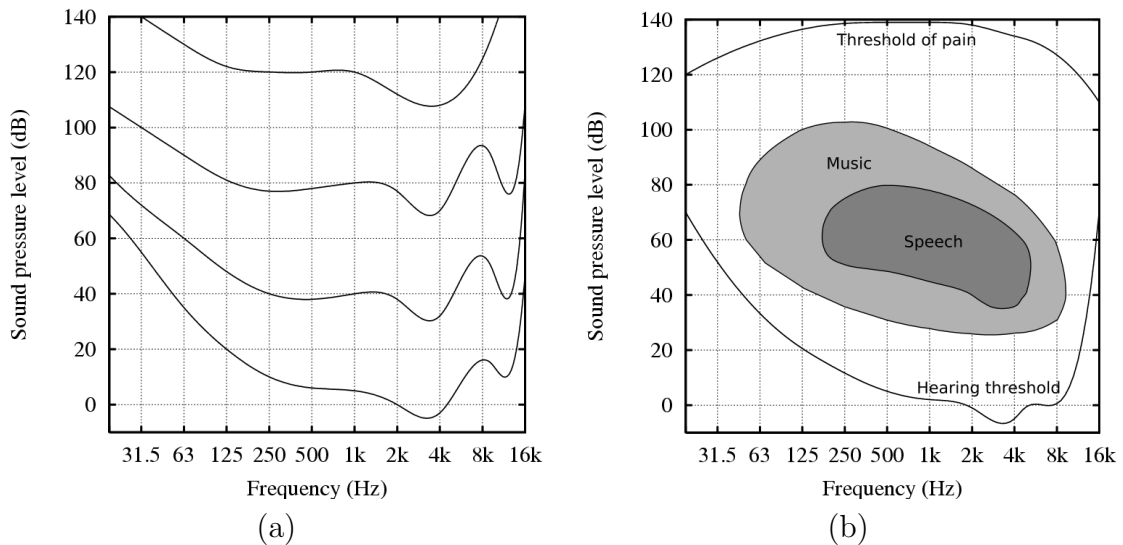
Figure 2.14: Curves of equal loudness.

Since tones of equal pressure level do not sound equally loud to us, some quanti-

ties are defined reflecting the subjective loudness sense, the *loudness level* and the *loudness*. The loudness level L_N in [phon] denotes the sound pressure level of a sine signal which is found as equally loud as the incident sound signal [CM78a]:

$$L_N = L_p(1000 \text{ Hz})/\text{dB} \quad [\text{phon}] \tag{2.28}$$

The loudness N in [sone] is defined as the perceived loudness of an incident sound of 40 [phon] loudness level. If the sound signal is felt twice as loud, L_N is doubled. This occurs by increasing L_N by 10 [phon]. The following relation arises from this fact [CM78a]:

$$N = 2^{(L_N/\text{phon}-40)/10} \quad [\text{sone}] \tag{2.29}$$

The human hearing is capable of localizing a sound source pretty precisely, even out of an amount of environmental noise. A wide variety of acoustic cues are responsible for this matter of fact [Har99], [Bla97]. J.W. Strutt (Lord Rayleigh) proposed in 1907 the first theory concerning the bilateral spacial hearing, the *duplex theory*, defined for pure-tones [SLR07]. He introduced two cues dominating the perceived location of the sound source. The first is the *inter-aural level (intensity) difference* (ILD or IID), indicating the different pressure levels received at ipsilateral (closest to the source) and contralateral (farthest from the source) ear. The second is the *inter-aural time difference* (ITD) representing the difference of arrival times of the sound wave front between the both ears.

At frequencies below about 1.5 kHz ITD is perceptually important for the localization, since no suitable level differences (ILD) can be observed due to diffraction effects of the head. Whereas at frequencies above 1.5 kHz limit, no distinct delay time differences (ITD) can be distinguish. The contralateral ear is located in the acoustical shadow of the head. The ILD becomes important for the localization process [Bla97]. The duplex theory and the frequency dependent dominance of the corresponding cues have been experimentally supported (see for example [WK92, MM02]).

However, the duplex theory does not describe the human sound localization completely. Humans who are hearing impaired in one ear are able to localize a sound origin [SIM94]. The two cues ITD/ILD are suitable for the localization of sound sources in the azimuth plane (left-right direction), only. For sources in the median plane the inter-aural differences for an hypothetical spherical head (as assumed in the duplex theory) are zero. The result is that the listener can not distinguish between sources in back, in front, or overhead [Har99]. In general, a sound source positioned anywhere on the shell of a cone centered in the axis between the ears will have identical inter-aural differences. This is the so called *cone of confusion* [Bla97]. However, humans are capable to differentiate sounds originating from the

median plane. Studies of the acoustical effects of the ear ([Bat67, ST68, Bla74]) have shown that the pinna together with the ear canal act as a direction selective filter. Depending on the direction and the distance of the source different eigen resonances in the pinna structure are stimulated. Accordingly, each direction holds different coloration, interpreted by the brain to distinguish back, front or above (see directional bands [Bla74] or [Bla97]). The localization in the median plane is less accurate as that in the azimuth plane using ITD and ILD. Furthermore, in order to localize the elevation in the median plane, the sound signals have to be broadband and contain frequencies above 7 kHz [Bla97]. It is widely believed that our auditory system uses ITD, ILD, and the spectral cues for spatial localization of sound sources.

In fact, not only the outer ear affects the sound wave, but also the head, shoulder and upper torso. This directional and distance dependent anatomical filtering of the sound before it reaches the eardrum is described by the complex response function the *head related transfer function* (HRTF) or the *anatomical transfer function* (ATF) [Har99, Beg00]. A lot of research has been done in order to physically model [KW92, Dud94, TPT04], as well as empirically measuring [WK89, GM96][2] the HRTF. HRTFs are used for localized rendering using headphone to create virtual auditory environments, see e.g. [ZDD04, Møl92]. The shape of the head, shoulder, torso and pinna differ from person to person, thus the HRTFs are individual. Headphone rendering using non-individualized HRTFs as filters can lead to localization failure (e.a. front-back or up-down confusion) [WAKW93] as well as to in-head-localization (IHL) where the virtual source is perceived to be inside the head [Bla97].

Further aspects affect the capability of sound source localization. In rooms for example, reverberations support a more precise determination of the distance to the source [SC00]. The reflections at room surfaces do not influence locating. The human auditory system has the directional hearing ability even in presence of reflections, due to the fact that much more importance is attached to the wave front which arrived first the auditory system as to the following wave fronts (the so called *law of first wave front* [CM78a]) [Bla74]. Also, sound that is familiar to us can be localized more precisely.

A vast amount of research has been done in the field of human perception of sound. Detailed description of this research area is out of the scope of this dissertation. The interested reader is referred to [Bla74, Bla97, Gol97, Beg00] for an extensive discussion. A tutorial to head related transfer functions is presented in [CW99].

---

[2]http://sound.media.mit.edu/KEMAR.html

2.3 Room acoustic parameters

In room acoustics numerous studies have been done in order to measure the subjective assessment criteria of acoustical behavior of a room with objective characterization quantities [Rei68, Ber04, Thi53, LB61, Yam72, Kür69]. Some of these metrics, which can be derived from the measured or simulated room impulse response, are incorporated into the ISO 3382 norm [ISO97] (German version [DIN00]). In the following several objective metrics are discussed. More criteria can be found in [FKS84b].

2.3.1 Sound strength

The first subjective criterion is the *sound strength* G [ISO97]. It describes the perceived loudness of a sound source inside a room. G is defined as the energy covered by the room impulse response:

$$G = 10 \; lg \frac{\int_0^\infty p^2(t) \mathrm{d}t}{\int_0^\infty p_{10}^2(t) \mathrm{d}t} \quad \mathrm{dB} \tag{2.30}$$

where $p(t)$ is the room impulse response at a given listener position and $p_{10}(t)$ is the impulse response at a distance of 10 meters from the source in free field.

2.3.2 Early and late sound energy relation

The following parameters base on the experiences that the early arrived sound (direct from the source as well as reflected wave fronts) enhances the auditory event. It facilitates e.g. the intelligibility of speech. The time limit lies approximately between 50ms and 100ms [Kut00]. Mostly, the ratio between the energy of the early sound ("useful" energy)

$$E_e = \int_0^{t_l} p^2(t) \mathrm{d}t \tag{2.31}$$

and the residual energy ("disturbing" energy)

$$E_r = \int_{t_l}^\infty p^2(t) \mathrm{d}t \tag{2.32}$$

is build [ISO97] with limit time t_l.

Thiele [Thi53] assigned the time limit of useful energy t_l to 50ms and defined the *definition* (original *Deutlichkeit*) as:

$$D_{50} = \frac{\int_0^{50} p^2(t)\mathrm{d}t}{\int_0^{\infty} p^2(t)\mathrm{d}t} \tag{2.33}$$

Definition is primary a metric for measuring the intelligibility of speech inside the room, but it can be applied to determine the quality of the recognizability of music instruments. A value of 0.5 denotes a 90 % syllable intelligibility[3] [MM04]

Ahnert [FKS84b] proposed a logarithmic metric for the definition of the *definition index* (original *Deutlichkeitsmaß*) which is the tenfold of the base 10 logarithm of the ratio between early and residual energy with $t_l = 50$ ms:

$$C_{50} = 10 \; lg \left(\frac{\int_0^{50} p^2(t)\mathrm{d}t}{\int_{50}^{\infty} p^2(t)\mathrm{d}t} \right) = 10 \; lg \left(\frac{D_{50}}{1 - D_{50}} \right) \quad \mathrm{dB} \tag{2.34}$$

$C_{50} > 0$ dB indicates good speech intelligibility [FKS84b].

Beranek and Schulz [BS65] used the reciprocal term of the definition to classify different concert halls in North America, the *ratio of reverberant to early energy*:

$$R = 10 \; lg \left(\frac{\int_{50}^{\infty} p^2(t)\mathrm{d}t}{\int_0^{50} p^2(t)\mathrm{d}t} \right) = -C_{50} \tag{2.35}$$

If R is too high the sound is muddy, though full and well-blended. If R is small, the sound is clear and well-defined, but dry. Good values are in the range between 3 dB and 8 dB [FKS84b].

Reichardt et. al. [RAAS74] have detected that the limit time t_l can (or should) be determined higher than that for speech. Therefore they set $t_l = 80$ ms and defined the *clarity index* (original *Klarheitsmaß*) to characterize the transparency of music presentations:

$$C_{80} = 10 \; lg \left(\frac{\int_0^{80} p^2(t)\mathrm{d}t}{\int_{80}^{\infty} p^2(t)\mathrm{d}t} \right) \quad \mathrm{dB} \tag{2.36}$$

Convenient values of C_{80} depend on the music style. For baroque music, Mozart for example, values of $C_{80} = 0$ dB may be acceptable, whereas music of romantic period (Tchaikovsky) may require a clarity index between -3 dB and 0 dB [RAAS74].

---

[3]The syllable intelligibility is defined as the ratio between the amount of correctly catched syllables and the amount of total spoken syllables [Rei68]. It is subjectively measured using so called logatoms (meaningless syllables)

The so far described metrics assume a strict limiting value where the late arriving
sound has a negative affect to speech or music. A continuous shift of a strong
reflection beyond this limit will cause a jump in the value of these criteria, which is
not natural. Therefore, Lochner and Burger [LB61, LB58] defined a metric where
the early energy is weighted by a function $a(t)$ depending on the delay time (and
also relative echo level). This function decreases gradually from 1 at the starting
time (arrival of direct sound) to 0 at $t_l = 95$ ms. This quantity was named the
signal-to-noise ratio by the authors and is calculated as follows:

$$\eta = \frac{\int_0^{95} p^2(t)a(t)\mathrm{d}t}{\int_{95}^{\infty} p^2(t)\mathrm{d}t} \tag{2.37}$$

The running time limit is also used for the next criterion, the *center time* (original
Schwerpunktszeit), proposed by Kürer [Kür69]

$$t_s = \frac{\int_0^{\infty} tp^2(t)\mathrm{d}t}{\int_0^{\infty} p^2(t)\mathrm{d}t} \tag{2.38}$$

Low values of t_s indicate high intelligibility of speech or high transparency of
music [MM04].

Speech intelligibility can also be characterized by the *speech transmission index*
(STI), or the simplified version the *rapid speech transmission index* (RASTI) pro-
posed by Houtgast and Steeneken [HS73, HS84]. STI and RASTI can be calculated
by use of the modulation transfer function (see [Sch81]) which, when white noise
is used as a source signal, can be derived form the impulse response [MM04]. The
interested reader is referred to the original work of the authors for more details
concerning this metric and the calculation algorithm.

2.3.3 Spatial impression

A further subjective effect in acoustics, which is mainly important for concert halls,
is the sensation of space usually experienced in a room. It is caused due to the
fact that in closed rooms sound reaches the listener from quite different directions,
and that our hearing is not able to distinguish between the different wave fronts
and treats them as one overall impression, the *spaciousness* or *spatial impression*
[Kut00].

First systematic studies of the spaciousness were accomplished by Baron [Bar71].
He detected that the space impression mainly relies on the early lateral reflections
rather than on the reverberation. According to this the spaciousness can be objec-
tively characterized by the ratio of the early lateral energy (from 5 to 80 ms after

arriving of the direct sound) and the overall early energy (from 0 to 80 ms including the direct sound), the *early lateral fraction* [BM81]:

$$LFC = \frac{\int_5^{80} p^2(t) \, cos(\theta) \mathrm{d}t}{\int_0^{80} p^2(t) \mathrm{d}t} \tag{2.39}$$

where θ indicates the angle between the incident direction of a certain energy part and the axis connecting both ears.

Due to the technical measurement possibilities alternative definition of the early lateral fraction can be found in literature [Bar00]:

$$LF = \frac{\int_5^{80} p^2(t) \, cos^2(\theta) \mathrm{d}t}{\int_0^{80} p^2(t) \mathrm{d}t} \tag{2.40}$$

Another way to characterize the subjective spatial impression is based upon the fact that later reflections produces different responses at the left and the right ear. This similarity or dissimilarity can be measured by the normalised inter-aural cross correlation function [ISO97] or [DIN00]:

$$IACF_{t_1,t_2}(\tau) = \frac{\int_{t_1}^{t_2} p_l(t) \cdot p_r(t + \tau) \mathrm{d}t}{\int_{t_1}^{t_2} p_l^2(t) \mathrm{d}t \, \int_{t_1}^{t_2} p_r^2(t) \mathrm{d}t} \tag{2.41}$$

The lower time limit t_1 is usually set to zero (arrival of first wave front), whereas the upper limit t_2 can differ, but is restricted to the integration of early reflections [Kut00]. The maximum of the absolute value of this function within the range $-1\mathrm{ms} < \tau < 1\mathrm{ms}$ is called the inter-aural cross correlation coefficient [DA72]:

$$IACC_{t1,t_2} = max|IACF_{t_1,t_2}(\tau)| \qquad -1\mathrm{ms} < \tau < 1\mathrm{ms} \tag{2.42}$$

High values of $IACC$ indicate a low degree of spaciousness.

Bradley and Soulodre [BS95b] experimentally found that at least two elements contribute to the spatial impression, the *apparent source width* (ASW) and the *listener envelopment* (LEV). The former describes the impression of the acoustical extension of the source compared to the source's visual perception. The latter depicts the impression of being surrounded by the music, being in the middle of the sound field. The above described metrics mark the suitable ASW. For the characterization of LEV, Bradley and Soulodre proposed the following criterion:

$$LG_{80}^\infty = 10 \, lg \left(\frac{\int_{80}^\infty p^2(t) \, cos^2(\theta) \mathrm{d}t}{\int_0^\infty p_{10}^2(t) \mathrm{d}t} \right) \tag{2.43}$$

where p_{10} is the impulse response measured in 10 m distance from the source in free space.

2.3.4 Initial reverberation

Historically, the first acoustical quantity which can be measured objectively is the reverberation time introduced by Sabine (see section 2.2.3). The reverberation time is still considered as the most important objective quantity of a room, even if it requires additional parameters for a complete description of acoustical conditions inside a room [Kut00]. This matter of fact has at least three reasons: (1) It can be measured or predicted with reasonable accuracy. (2) It does not depend significantly on the listener position. (3) Data on reverberation times of existing halls are available, and can be used as a benchmark.

However, it has become evident that by continuous music or speech only the initial decay rate is deciding for the subjective reverberation sensation of the listener. Therefore, several researchers proposed to limit the evaluation interval of the 60 dB level decay to the initial portion of the decay process [CM78a]. In the literature appear different initial decay time definitions depending on the evaluation interval, $RT15$ (from 0 to -15 dB), $RT20$ (from 0 to -20 dB), or $RT30$ (from -5 to -35 dB) [CM78a]. The initial reverberation time depends, in contrast to the Sabine's reverberation on the time location.

Nowadays, it has become common to characterize the reverberation by use of the *early decay time* (EDT, $RT10$) proposed by Jordan [Jor70]. It is the time in which the first 10 dB of the decay occurs multiplied by six, and can be calculated using the following equation (see [FKS84b]):

$$10 \, lg \left(\frac{\int_0^\infty p^2(t)\mathrm{d}t - \int_0^{EDT/6} p^2(t)\mathrm{d}t}{\int_0^\infty p^2(t)\mathrm{d}t} \right) \, \mathrm{dB} = -10 \, \mathrm{dB} \qquad (2.44)$$

3 Computer-aided room acoustics

Nowadays, many computer-aided algorithms for modeling and simulation of acoustics inside enclosures exist. Mostly, they relay on wave-based or geometric acoustics. This chapter provides an overview of existing acoustic simulation approaches. A survey of such methods is also given in [RSS97, Sve02].

3.1 Wave-based methods

As mentioned in section 2.2.1 wave-based acoustics deals with the numerical solution of the wave equation. Common numerical methods include *Finite Difference Time Domain* (FDTD) techniques [Yee66, Bot95, SLV02], *Boundary Element Methods* (BEM) [Wu00], and *Finite Element Methods* (FEM) [Bra03, Ihl98]. These methods subdivide the domain into small elements such that the differential wave equations can be approximated by a system of equations which can be solved by means of well known techniques. Since the elements must be smaller than the smallest simulated wave length, these methods are restricted to low frequencies. A brief overview of the applications of BEM and FEM to acoustic problems is given in [Est07].

The FDTD methods basically use a Cartesian grid for discretization. Sound pressure and particle velocity are determined for each grid point. Therefore the corresponding differential equations are approximated via finite difference functions simplifying the solution of the system. The implementation of the FDTD methods is straightforward. As the name implies, all the calculations are done in the time domain.

The boundary element methods discretize only the boundary of the considered domain but not the entire surface or volume. The basic idea is the description of the propagation of a sound wave in the interior of the domain by means of the state variables at the boundary. The degrees of freedom used for the discretization is lower than that for FEM, but a fully populated system of equations has to be solved.

For closed rooms the wave equation is preferably solved by the finite element method (FEM), which approximates the wave equation by a large system of ordinary differential equations (ODEs). The unknowns are the pressure values at grid points inside the room. In general, there are by far too many unknowns to solve these systems of ODEs in real time. Hence, a reduction of the system to a concise state-

space model with similar input-output behavior in the frequency range of interest is needed.

There are many different approaches to model reduction [AS01]. The common observation is that system dynamics can often be represented quite well by a superposition of a few (generalized) eigenmodes. The coefficients of these modes are the unknowns of the new reduced system. Finally, assuming sample-wise constant input (e.g. acceleration of the loudspeaker membrane), the continuous state-space model is transformed into a discrete one, which can be solved in real time.

In the following the steps to get from the wave equation (3.1) to a reduced discrete state-space model (3.6) describing the transient response of a room to excursion of a loudspeaker membrane are listed. This reduction approach was developed at the Fraunhofer ITWM within the scope of a collaborative work and is published in [DBM$^+$06]. The wave equation and associated boundary conditions read:

$$
\begin{aligned}
\frac{\partial^2 p}{\partial t^2} - c_0{}^2 \Delta p &= 0 \quad \text{on} \quad G \\
c_0 \frac{\partial p}{\partial n} &= -\frac{1-R}{1+R} \frac{\partial p}{\partial t} \quad \text{on} \quad \Gamma_w \\
\frac{\partial p}{\partial n} &= -\rho_0 \frac{\partial^2 x_m}{\partial t^2} \quad \text{on} \quad \Gamma_m .
\end{aligned}
\tag{3.1}
$$

$p = p(t, x)$ denotes pressure, c_0 the velocity of sound, and $\rho_0 = 1.2 \ kg/m$ the density of air at room temperature, G the interior of the room, and Γ_w and Γ_m the surfaces of walls and membrane, respectively. x_m is the excursion of the membrane and R is a reflection coefficient. It may depend on the particular wall, but is constant for all frequencies. This is a minor problem, as the model is used only for a small frequency band.

Approximating the pressure distribution by a superposition of, for instance, piecewise quadratic ansatzfunctions $p(t, x) = \sum_{i=0}^{N} p_i(t) \, \varphi_i(x)$ and integrating (3.1) with respect to the φ_i gives a FE model of the form:

$$
\begin{aligned}
M\ddot{p} + D\dot{p} + Kp &= Fu \\
y &= Pp .
\end{aligned}
\tag{3.2}
$$

The real $N \times N$ matrices M, D, K are called mass, damping, and stiffness matrix. $p = p(t)$ is a vector composed of the coefficients p_i. $u = u(t)$ is the input, e.g. the acceleration of the membrane. F transforms this input into a force. P is a projection matrix extracting certain examined pressures y_i. Setting

$$
\begin{aligned}
\hat{x} = \begin{bmatrix} p \\ \dot{p} \end{bmatrix}, \quad
\hat{E} = \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix}, \quad
\hat{A} = \begin{bmatrix} 0 & I \\ -K & -D \end{bmatrix}, \\
\hat{B} = \begin{bmatrix} 0 \\ F \end{bmatrix}, \quad and \quad
\hat{C} = \begin{bmatrix} P & 0 \end{bmatrix}
\end{aligned}
\tag{3.3}
$$

the FE model may be rewritten as a generalized state-space model:

$$\hat{E}\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}u$$
$$y = \hat{C}\hat{x} . \tag{3.4}$$

Assuming that $\hat{x}$ is essentially composed of the columns of a matrix $U \in \mathbb{R}^{N \times n}$, $n \ll N$, and projecting the equations on the columns of $V \in \mathbb{R}^{N \times n}$ yields a reduced state-space system where

$$\hat{x} = U\tilde{x}, \quad \tilde{E} = V^t \hat{E} U, \quad \tilde{A} = V^t \hat{A} U, \quad B = V^t \hat{B}, \quad \tilde{C} = \hat{C} U . \tag{3.5}$$

The columns of U and V may be found by expanding the associated transfer function $H(s) = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B}$ about some shifts $s_j = \mathrm{i}\omega_j$. Here, the *rational dual Arnoldi* algorithm described in [Ols02] is used. Finally, dividing the first equation of the reduced version of (3.4) by $\tilde{E}$ and performing a balanced truncation [ZD98] result in a state-space system of typically a few hundred unknowns rather than several 10,000 degrees of freedom. Integrating the reduced version of (3.4) over the length Δt of one sample for constant input u_n, i.e.

$$x_n = \mathrm{e}^{\tilde{A}\Delta t} x_{n-1} + \int_0^{\Delta t} \mathrm{e}^{\tilde{A}(\Delta t - \tau)} \, \mathrm{d}\tau \, \tilde{B} \, u_{n-1},$$

leads to a discrete state-space system

$$x_n = Ax_{n-1} + Bu_{n-1}$$
$$y_n = Cx_n . \tag{3.6}$$

By switching to a suitable basis $x_n = T\xi_n$, it is always possible to turn the system into *companion canonical form* [KK85], where the first $n-1$ rows of A coincide with the last $n-1$ rows of the identity matrix of order n, the last row contains the negative coefficients of the characteristic polynomial of A, and B is the n-th unit vector. Hence, updating the state vector x_n and evaluating the pressure at a certain position requires $2n$ multiplications and $2n-1$ additions.

3.2 Geometric methods

In room acoustic design and especially in Virtual Reality applications mostly, methods based on geometrical acoustics are applied. In the following different geometrical approaches are presented.

3.2.1 Image source method
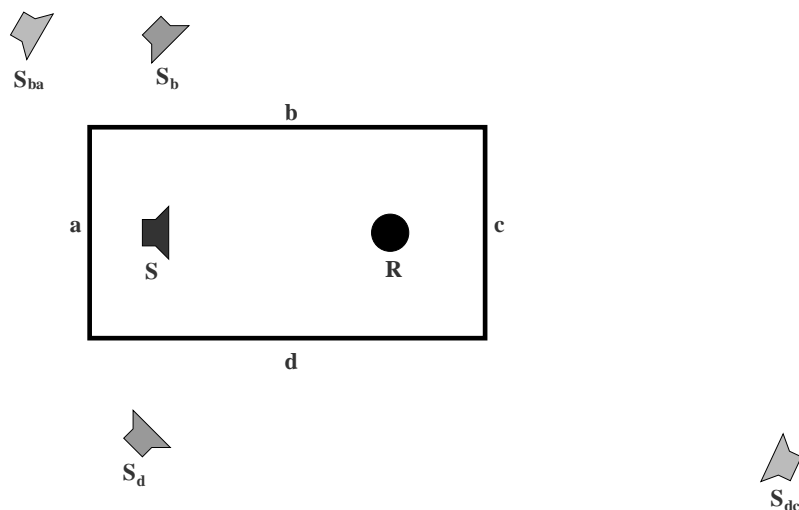
Figure 3.1: Sketch of first order (S_b and S_d) and second order (S_{ba} and S_{dc}) image sources.

The *image source method* [AB79, Bor84] models specular reflections by inserting new sound sources, which are obtained by mirroring the location of an audio source at polygonal surfaces inside the scene (see figure 3.1). The key idea here is that a sound ray from a sound source S reflected at a surface, has the same acoustical effect as one from the mirrored source S', reduced by the absorbed part. In a rectangular box-shaped room it is easy to construct image sources up to a certain order of reflections. Following the image source calculations, for a receiver position R, each of the virtual sources must be checked if it can be heard in R. The contributions of all "hearable" sources are summed to build the acoustical impression at R. The advantage of the image source method is that it is very accurate, but there is a problem if the room is not box-shaped. The number of the image sources increases exponentially with reflection order. Therefore, this approach is suitable for simple room geometry or low reflection orders, only.
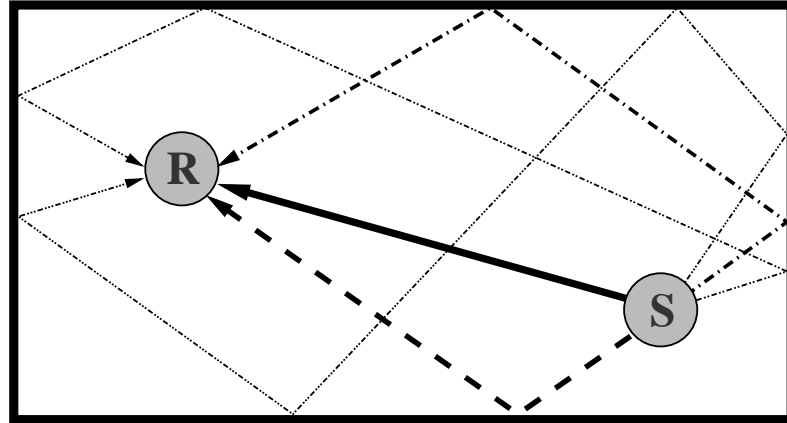
Figure 3.2: Sketch of the ray tracing method up to fourth order of reflections.

3.2.2 Ray tracing methods

In the *ray tracing method* [Kro68, Kul84, Vor88] several rays are traced from the sound source to listeners (see figure 3.2), which are typically represented as spheres collecting all intersecting rays. The reflections at the surfaces in the scene take place according to Snell's law of specular reflection or diffuse reflection. Usually, Lambert's law is used for diffuse reflection calculation, but there exist approaches to model scattering more realistically [EAS01, CR05, TLDD07b]. Typically each ray is reflected either specularly or diffusely, the choice of which is determined by a scattering coefficient. The result may be more precise if each ray generates several secondary rays but the computation time would increase exponentially. In addition to the absorption properties of considered surface time delay and air absorption can be taken into account, too. The information contained in the resulting energy histogramm can be converted into a impulse response used for auralization applying for example the method described in [Kut93]. The ray tracing approach is general and simple to implement. It can be easily scaled for accuracy and/or speed trade-offs by limiting the number of traced rays. However, when changing the receiver's position, the tracing of all rays needs to be redone.

3.2.3 Beam tracing and related methods

Another widely used technique in acoustic computation is the *beam tracing method* [WD82, HH84, DKW85]. Beam tracing overcomes the aliasing problem of classical ray tracing by recursively tracing beams (i.e. sets of rays) through the scene. Each beam is checked for intersections with scene polygons in a back-to-front order ensuring subsequent processing of partially hidden polygons. At intersection of the beam with a polygon the original beam is clipped in order to remove the

shadow region. Sub-beams are constructed representing refraction and reflection. Further sub-beams can be constructed modeling additional sound wave phenomena. In case of polyhedral beams the cross-sections of the sub-beams are the intersection polygons.

Beam tracing overcomes the problem of sampling artifacts common for traditional ray tracing approaches. The primary disadvantage is the complexity of the intersection calculations with scene polygons. Compared to image source methods beam tracing is an efficient method to determine visible image sources, since each beam represents the region for which the corresponding virtual source is visible. Thus less image sources have to be treated. Diffraction phenomena can be realized with beam tracing, whereas scattering can not easily be modeled [Sve02].

Another beam tracing approach is *cone tracing* [Ama84, MM93, Dal96]. These approaches consider, as the name implies, conical beams. To avoid complex intersection calculations mostly only the center ray is traced instead of the actual cone face. Unlike polyhedral beam tracing the drawback of the cone tracing is the overlapping of the cones due to the fact that circular faces can not cover a spherical wave exactly. In order to overcome the problem Gaussian weighting functions can be used [MM93] tracing some kind of Gaussian beams or to use pyramids with triangular cross-section closer approximating the sphere surface [Lew93, Far95].

Funkhouser et. al. [FCE⁺98, FMC99, FTC⁺04] presented a polyhedral beam tracing approach for interactive virtual reality applications. The method consists of four phases, two of them are executed within an off-line process and the residual two on the fly while the user explores the virtual scene. In the first phase, the spatial subdivision phase, the 3-D scene is subdivided into convex polyhedral regions using binary space partitioning. The result of this phase is a cell adjacency graph representing the neighborhood relations of the convex cells which enables an efficient space traversal. In the second phase, the beam tracing phase, the tracing process proceeds considering transmission, specular reflection, and diffraction [TFNC01] phenomena of sound while constructing new beams. The resulting beam tree structure is used in the next phase, the path generation phase, in order to determine beam paths from the sound source to each receiver. In the last phase the paths are then applied for auralization. The sound diffraction is modelled using the Uniform Theory of Diffraction (UTD) [KP74], an extension of the Geometrical Theory of Diffraction (GTD) [Kel62], which incorporates diffraction into the ray theory of light. In the UTD an infinite wedge is handled as a secondary source of diffracted waves that can be reflected or diffracted before arriving at the listener position. An incident ray results in a cone of diffracted rays, where the aperture angle of the cone is equal to the angle between the ray and the diffracting edge. The axis of this cone is the edge. For a given source and receiver position a single ray represents the diffracted field. The UTD assumes infinite length of the diffracting edge, and is valid only if source and receiver are located far from the edge. For each diffracting edge the algorithm

computes a UTD-based diffraction depending on the shortest path through the edge. In order to reduce computation time diffraction calculations can be limited to receivers in shadow zones. Of course, this reduces the computation accuracy. The proposed approach has several limitations [FTC+04]. The 3-D model must contain only planar polygons that are greater in diameter than the sound wave length. As mentioned above, incorporating diffuse reflections into the beam tracing algorithm is not straightforward, thus currently, no diffuse reflections are considered.

A recently proposed approach for tracing volumetric rays is the *frustum tracing algorithm* [LCM07a, LCM07b]. This algorithm deals with the intersection complexity by using four-sided convex frusta instead of polyhedral beams. Also, the clipping on the intersection polygon is not performed exactly, but discrete. Therefore each frustum is uniformly subdivided into sub-frusta which are represented by a sample ray. An intersection of a sub-frustum with a scene primitive occurs only if the sample ray intersects the primitive thereby introducing some errors in the algorithm. The authors have demonstrated that the frustum tracing algorithm can be executed at interactive rates for complex scenes allowing reflection orders up to seven. In the current version of the method transmission and specular reflection are considered, but no diffuse reflection and diffraction.

3.2.4 Radiosity related methods

Ray tracing and image source methods depend on the receiver position. Changing the listener position implies a recalculation of the room impulse response. In contrast, radiosity methods are view independent. Originally used in thermodynamics the radiosity methods are often applied in computer graphics for illumination calculations in enclosures [GTGB84, CG85].

Radiosity methods are based on the energy balance. Thereby ideal diffuse reflections, i.e. surfaces reflecting light uniformly in all directions, are assumed. The scene is subdivided into finite sub-faces (patches) between which the energy transport is determined. The fraction of energy exchanged between two patches is expressed by the so called form factors. The calculation of the form factors is time-consuming making the radiosity approach computational expensive.

An algorithmic realization of the radiosity method simulating room acoustics is proposed in [SZ93]. Two essential differences between sound and light waves which should be incorporated into the radiosity formulation are the air absorption and the finite velocity of sound. The air damping is given by the exponential law and is modelled by the air attenuation factor between two patches analogous to the form factors. The more important modification of the radiosity equation is the consideration of the time delay of arriving sound. The acoustic radiosity algorithm described in [SZ93] can be summarized as follows. For a given scene geometry and

reflection coefficients, first the form factors and air attenuation factors are computed. These factors do not change in time, thus they have to be calculated only once. The energy balance equation is then evaluated for every time step t. The impulse response for a listener position can be calculated by placing an additional surface representing the listener.

By changing the listener position only the form factors and air attenuation factors for the listener patch have to be updated. Since every surface patch can emit energy, integration of moving sources is straightforward. Nosal et. al. [NMI04] reviewed and discussed existing radiosity approaches. Furthermore, they introduce several improved techniques for computation time reduction of the radiosity methods in arbitrary polyhedral rooms.

As mentioned before only diffuse reflections are considered. In computer graphics, there also exist methods extending the radiosity algorithm in order to consider specular reflections, for example by applying bidirectional reflection distribution functions (BRDFs) [SAWG91]. Kajiya [Kaj86] introduced a general rendering equation for global illumination calculations considering diffuse and specular light reflections. This equation has been extended to the acoustic rendering equation [SLS07a, SLS07b] for use in room acoustic simulations applying also acoustic BRDFs. Diffraction and transmission of sound are not modelled in the current version.

Due to the computation complexity these methods do not seem practical for large environments.

3.2.5 Further approaches

Due to the shortcomings of the approaches described above, continuative approaches have been developed. Mostly, they employ parts of the classical schemes or a combination of them.

One approach that makes use of the advantages of the image source method and ray tracing is introduced in [Vor89]. Here the visibility check of the image source algorithm is performed via ray tracing. This hybrid method has been further extended in [Hei93]. The direct sound and early reflections are calculated using a combined method to predict specular parts of the impulse response. To calculate the diffuse part of early reflections a second ray tracing stage with low time resolution is performed and only if the last ray reflection is non-specular, decided by a random variable, the energy is collected at the listener position. A statistical method using the frequency dependent reverberation times is applied to simulate the reverberant tail of the impulse response. An upgrade and extension of both methods mentioned above in respect of real-time capabilities is the RAVEN acoustic simulation system [LSVA07, SV07]. The system takes effects such as flutter echos into account as well.

Odeon [Nay93] is another room acoustic simulation software implementing a hybrid method. Rays are used in order to discover the location of image sources. Also cones are used to find possibly missed valid image sources.

In [KBAA01] an extended radiosity approach is combined with the image source method. Here, the image source method is applied to determine specular-specular reflected energy between surfaces (patches), whereas specular-diffuse, diffuse-specular, and diffuse-diffuse reflections are handled by use of extended radiosity.

Another hybrid simulation method combining beam tracing with radiosity is proposed in [Lew93]. The approach consists of two stages. In the first stage triangular beams are traced trough the environment in order to calculate the early specularly reflected part of the echogram. At each reflection diffuse energy is stored at the surfaces. This stored energy is then used in the second step to determine the diffuse reverberant. Therefore a radiosity method is applied.

The method has been modified in [DL00] to trace polyhedral beams. Emitting a triangular beam, the shape of reflected beams is governed by the shape of the reflecting surface. The computationally expensive polygon-polygon intersections are avoided by the algorithm. Starting at a suitable position at the reflecting surface the algorithm maps out all illuminated sections using a progressive description process that automatically takes into account all plane edges, including those of occluding planes. Each section of illumination is used to produce a new reflected beam. Also diffuse reflections are facilitated by the algorithm using radiant exchange.

A further modification of the triangular beam tracing is the approach described in [WH05]. The beams carry sound pressure and phase information. The complex reflection coefficients of the boundary surfaces are determined by use of the transfer-matrix approach.

An approach accelerating the beam tracing algorithm is presented in [AFST04b, AFST04a, AFST04c]. Therefore, all informations independent of the source and receiver positions are precalculated and stored. In fact, it is the visibility information between two reflectors stored in the so called visibility diagrams. By a given source position a beam tree can be constructed through a look-up process on the visibility diagrams. Through a look-up on the beam tree the paths of the sound wave from the source to a receiver can then be calculated. Also diffraction effects are modeled utilizing the Uniform Theory of Diffraction (UTD) [KP74]. The proposed version of the approach can be applied to 2.5-D scenes (vertical walls and horizontal floor and ceiling).

An idea to unify the advantages of the (ray, beam) tracing algorithm, the image source method, and the radiosity approach, the quantized pyramidal beam tracing is proposed in [Ste96, Ste04]. The sound energy is carried within polygonal beams, called pyrs, each representing an image source, and are splitted at reflections and diffraction. In order to reduce the calculation time the pyrs are re-unified. Thereto

a quantization of solid angles and room surfaces into small patches is performed, corresponding to the quantization of the image source space into cells. Any pyr's source point falls into such cell and is shifted to the center point of the cell. The associated energy of the pyr is interpolated and distributed to previously created quantized pyrs according to their overlap with the surface patches. The shapes of these quantized pyrs are given by the patches. In this way pyrs that arrive at patches within a certain time interval are unified. Thus, any kind of reflection and diffraction can be combined. The quantized pyr tracing corresponds to a combination of the image source method and the radiosity approach. The edge diffraction is realized by an energetic ray diffraction model inspired by Heisenberg's uncertainty principle. Therefor, reflection-angle-probability-density-functions [Ste04, SS07] are defined, whose values increase the closer the ray or beam passes an edge. Rays or beams are splitted into several new ones according to this distribution. The method reduces computation time, however the memory usage can be very high [Ste04].

Another approach often used for modeling diffraction in acoustics is the Biot-Tolstoy theory [BT57] by infinite rigid wedges, which was extended by Medwin et. al. [Med81, MCJ82] considering also finite barriers. This approach is often referred to as Biot-Tolstoy-Medwin (BTM) solution in the literature. Here, the edges are discretized into a number of secondary point sources whose contributions are added to determine the diffracted field. The resulting edge-diffraction impulse response is subject to two singularities in its first sample (onset singularities), first of which is addressed by an analytical solution in [SFV99], and the second one by an approximation [SC06]. Based on the extended BTM solution [SFV99] an approach to integrate edge-diffraction into geometrical acoustics is proposed in [CSF05] re-using intermediate values of the diffraction calculations to determine specular reflections. Furthermore, the solution described in [SFV99] is utilized in [TSK01, PLS02, PL03] as an extension of the image-source method. The diffracting edges are treated as additional diffraction-image sources. Diffractions upto second order are considered. The BTM method is computational complex, thus several approximations were developed to reduce the computation time. In [LSS02] an implementation of edge-diffraction in parametric auralization is presented. In auralization, the diffraction is approximated with a low-order warped IIR filter fitted to the analytic diffracted response based on BTM. This auralization is integrated in the dynamic rendering system, the DIVA system [SHLV99, SLH02, LSS02]. However, the diffraction calculations have to be performed in an offline process by given source and listener positions over time. In [CS07] an edge-subdivision strategy is proposed to speed up the time-domain edge-diffraction calculations.

Other approaches exploit programmable graphics hardware (GPU) to accelerate the simulation calculations. In [JK04] the GPU is used to perform acoustic ray tracing. In [TG97, TG98] graphics hardware is used for calculation of sound occlusion and diffraction effects by obstacles based on the Fresnel-Kirchoff approximation. In

[TLDD07b, TLDD07a] the GPU is utilized to model first-order sound scattering, both diffraction and reflection, off complex surfaces based on a surface-integral formulation and the Kirchhoff approximation. For audio-scattering calculations combined normal and displacement maps are used.

3.2.6 Sonel mapping

Sonel mapping [Kap06, KJM04, KJM05b, KJM05a] is another geometric acoustics approach also inspired by Photon mapping (see appendix C). It is a two stage sound particle based method modeling specular and diffuse reflections, as well as diffraction of sound waves. In the first stage sound particles, called sonels, are traced through the virtual scene in order to populate the sonel map. This map is used in the second stage, the rendering stage, in combination with distribution ray tracing to calculate the echogram at a given listener position. Only diffusely reflected sonels are stored in the sonel map.

In the sonel tracing stage a prescribed number of sound particles is emitted from the sound source. At interaction of the sonel with a scene surface, it is checked whether the intersection point is located within a diffraction zone or a non-diffraction zone. In the first case either the sonel will be transmitted through the diffraction zone unaltered or a new sonel is generated and emitted in a random direction over a hemisphere centered at the intersection point. In case of a non-diffracting zone one of the following three options occurs. The sonel can be either reflected ideal specularly, reflected ideal diffusely (Lambertian) and stored in the sonel map, or it can be absorbed. In the latter case the sonel is terminated and a new one is started from the source. In order to decide which of the five cases occurs the Russian roulette technique is used.

In the rendering stage rays are emitted from the listener constructing the echogram at this position. At intersection of the ray with a scene surface one of five possibilities occurs, similar to the rendering stage. Here Russian roulette is used as well. In case of a diffuse reflection, the sonel map is applied to determine the portion of the energy leaving the intersection position and contributing to the echogram. The ray tracing process it terminated at this point. In case of a specular reflection, the tracing is continued assuming ideal specular reflection. If the ray hits the sound source, represented as a sphere, the attenuated energy due to the medium impact is added to the echogram. If the ray hits a diffraction zone, a random variable is used to decide whether the ray transmits through the diffraction zone and the tracing is continued or diffraction terminates the ray. The portion of the diffracted energy added to the echogram is determined using the Huygens-Fresnel principle. In case of absorption the ray is terminated as well.

Sonel mapping is a particle-based Monte Carlo method which can be applied for acoustic calculations inside arbitrary complex environments.

4 Phonon tracing

Photon mapping [Jen96, JC98] is often used for rendering photo-realistic images, supplementing uni-directional ray-tracing by a variety of visual effects, like color bleeding and caustics. A similar approach is adopted to the simulation of sound, where a number of issues need to be re-considered. For example, the human ear is not capable of locating the origin of a sound source as precise, as the human eye is capable of locating the origin of light. In the case of high frequencies where diffraction is negligible, one can locate directional sound with a precision of about 30 degrees. Thus, a lower number of particles (phonons) may be sufficient for auralization.

Another difference is that the human ear can distinguish a lot more frequency bands than the eye. Since the eye has only three different receptors for red, green, and blue light, any continuous distribution within the visible spectrum between 400 and 700 μm of wavelength is mapped onto three variables. Consequently, only three wavelengths need to be considered for rendering images. The human ear, however, can distinguish frequencies in the spectrum $\Omega = [20Hz, 20000Hz]$ fairly precise. In addition, the spectrum of an audio signal is highly time-dependent in contrast to most illumination problems that require only a static solution.

In this chapter we present a novel geometric acoustics algorithm, the Phonon Tracing, for the simulation of mid- and high-frequency sound. For a given source and listener positions, our method computes sound energy as well as sound pressure response accounting for the different reflections at various surfaces with frequency-dependent absorption coefficients. We do not consider diffraction effects due to low frequencies, since these can be better computed by e.g. finite element methods.

4.1 Tracing the sound energy

The idea of the algorithm is the separation of tracing and filter calculation processes. This way, only one tracing step is needed for a given sound source, followed by multiple impulse response calculations steps for different listener positions. Of course, when the source position has been changed the particle tracing has to be re-executed. Thus the phonon tracing algorithm consists of two steps: the *phonon emission step* where the particle traces are calculated and the *phonon collection step* used to determine the impulse response at given listener position. As result of

the *emission step* we obtain the *phonon map* which represents all reflections of an emitted wave front by a large set of particles (phonons).

4.1.1 Problem specification

The Phonon Tracing algorithm requires the following input:

- position of the sound source S
- emission distribution E of the sound source
- number of phonons n_{ph} to be traced from the sound source
- one or more listener positions l_i
- a triangulated scene with tagged materials m_j
- an absorption function $\alpha_j : \Omega \mapsto (0,1]$ for each material
- an acoustic BRDF for each material (if not specular)
- an energy threshold ε and maximum number of reflections n_r for terminating the phonon paths.

The output of our method is a finite response filter (FIR) f_i for each listener position l_i corresponding to the energy pulse response with respect to the sound source.

4.1.2 Phonon emission

Every phonon p emitted from the sound source carries the following information:

- an energy spectrum $e_{ph} : \Omega \mapsto \mathbb{R}^+$
- the distance d_{ph} traversed from the source
- the phonon's current position pt_{ph}
- the outgoing direction v_{ph}

Additional information, such as the number of reflections and the material index of the last reflection may also be attached for visualization purposes.

The used absorption and energy functions α_j are represented by $n_e = 10$ coefficients associated with the frequencies $40, 80, 160, ..., 20480$ Hz. The corresponding basis functions for the energy spectrum are wavelets ω_i ($i = 1, ..., n_e$) adding up to a unit impulse (discretized Dirac pulse). The wavelets will be described later. For now, we assume that every phonon is composed of different frequencies, which is much more efficient than tracing a single phonon for each individual frequency band.

Phonons are emitted from the source S according to the emission probability distribution E. In the case of a sound source with equally distributed emission over the unit sphere, three independent random variables x, y, z are used with normal (Gaussian) distribution based on the probability function:

$$P(x) = \frac{2}{\sqrt{\pi}} exp\left(-x^2\right) \tag{4.1}$$

This is obtained by inserting a random variable ξ equally distributed on $[0, 1]$ into $x = \Phi^{-1}(\xi)$ where Φ is the normal distribution function:

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x exp\left(-t^2\right)\mathrm{d}t = erf(x) \tag{4.2}$$

Since the variable $(x, y, z)^T$ forms a rotationally symmetric Gaussian distribution, its normalized direction is equally distributed on the unit sphere [Dev86] and can be used to emit phonons at a spherical sound source (point source).

The phonons are started at the source with a (discretized) unit energy spectrum $e_{ph,i} = 1$ ($i = 1, ..., n_e$). At the first intersection of the phonon ray with the scene, the phonon direction d_{ph} is reflected with respect to the local surface normal, the absorbed energy is subtracted according to the local material m_j,

$$e_{ph,i} := (1 - \alpha_{j,i}) e_{ph,i}, \tag{4.3}$$

and the distance d_{ph} is set to the traversed distance. The phonon is fixed at the intersection point, contributing to a global phonon map (whose evaluation is discussed later). To efficiently compute the ray-surface intersections, a kd-tree based partitioning of the scene (see [PH04]) is used.

If the maximal energy of the phonon exceeds the energy threshold, i.e.

$$max\{e_{ph,i}\}_{i=1}^{n_e} > \varepsilon_e, \tag{4.4}$$

the next phonon re-uses the path and energy of the preceding one, saving computation time. It is started at the current position with respect to the outgoing direction v_{ph} and contributes to the phonon map at the next surface intersection. If the threshold ε is not exceeded and a minimum number of reflections have been computed, then a new phonon is started from the sound source S. After a prescribed number n_{ph} of phonons have been spread out from the source, the tracing is terminated. Figure 4.1 shows a flow chart of the emission step.

The approach can be extended by considering bi-directional reflection distribution functions (BRDF's), for example when modeling corrugated materials by plain surfaces. A BRDF provides the reflection factor for a pair of incoming and outgoing directions. For any fixed incoming direction, it can be used as probability
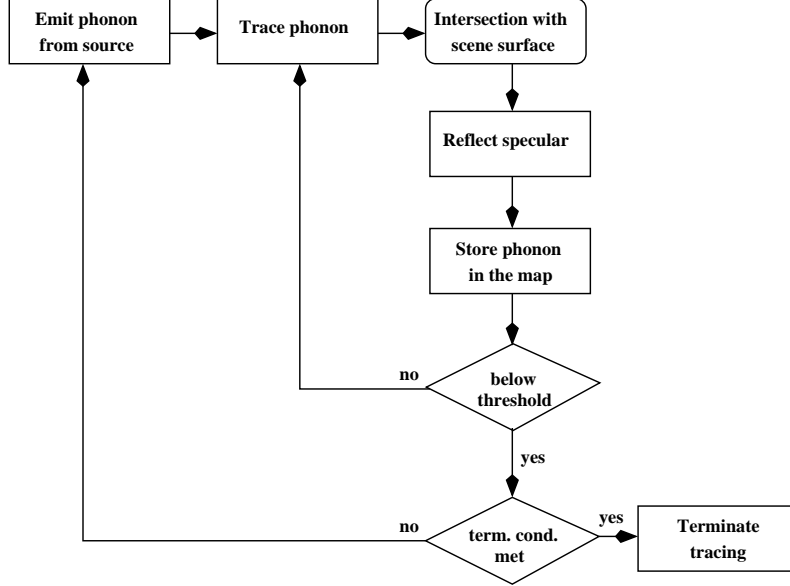
Figure 4.1: Flow chart diagram of the phonon emission step.

function for a random variable determining the reflection vector. We note that the material's corrugation must be very coarse to justify this approach, since the wave lengths of sound are much wider than those of light. Due to the strong dependency of the wavelengths and reflection distribution, individual BRDF's may be defined for several bandwidths, making it necessary to trace multiple succeeding phonons in different directions, carrying distinct parts of the energy spectrum.

4.1.3 Phonon collection and filtering

The remaining task of the method is collecting the phonons' contributions to a FIR filter f for every listener position l. This filter f corresponds to the energy pulse response from the source, recorded at l, such that convolution with an anechoic source signal g_s reproduces the perceived signal g_l, where

$$g_l = g_s * f = \int_{\mathbb{R}} g_s\left(\tau\right) f\left(\cdot - \tau\right) \mathrm{d}\tau. \tag{4.5}$$

In the case of uniform absorption for all frequencies, the contribution of a phonon visible from the listener is simply a scaled, translated Dirac impulse (unit impulse),

$$\delta_{ph,l,i}\left(t\right) = exp\left(-kh_{pl}^2\right) e_{ph,i}\delta\left(t - \left(d_{ph} + d_{pl}\right)/c\right), \tag{4.6}$$

where k is a constant, h_{pl} denotes the shortest distance between the continued

phonon path $pt_{ph} + tv_{ph}$ and the listener l, $e_{ph,i}$ denotes the phonon's energy (at frequency ω_i), d_{pl} denotes the distance between phonon's pt_{ph} and listener l, and c is the speed of sound. The term $(d_{ph} + d_{pl})/c$ corresponds to the time elapsed between emission and reception of the phonon. The Dirac pulse δ is shifted by this time interval and scaled by the phonon's energy $e_{ph,i}$, multiplied by a Gaussian, weighting the distance of the ray to the listener.

In classical acoustic ray-tracing [Kro68, Kul84], a sphere is used to collect rays at a listener's position. Using a Gaussian, however, provides much smoother filters, since more phonon rays contribute to the filter, weighted by their shortest distance. The constant k determines the radius of this Gaussian. For now, $k = 200m^{-2}$ is used, keeping in mind that this constant needs to be adapted to the phonon density. Figure 4.2 shows a sketch of the phonon emission step.
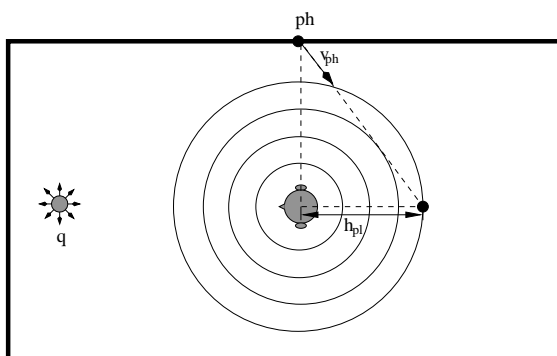
Figure 4.2: Sketch of the phonon collection step.

In the more general case of frequency-dependent absorption, the Dirac pulse is subdivided into wavelets ψ_i representing the individual frequency bands,

$$\delta(t) = \sum_{i=1}^{n_e} \psi_i(t).$$ (4.7)

With the different energy coefficients $e_{ph,i}$, the filter becomes

$$f(t) = \sum_p exp\left(-kh_{pl}^2\right) \sum_{i=1}^{n_e} e_{ph,i}\psi_i\left(t - (d_{ph} + d_{pl})/c\right)$$ (4.8)

For a large number of phonons, it is more efficient to precompute n_e different filters, one for each bandwidth, based on equation 4.6. Their convolution with the individual wavelets are then efficiently computed employing the Fast Fourier

Transform (FFT). By the convolution theorem, equation 4.8 is equivalent to

$$f(t) = \sum_{i=1}^{n_e} \psi_i * \sum_{p} \delta_{ph,l,i}(t), \qquad (4.9)$$

where the convolution translates and scales the wavelets properly for each individual phonon.

4.1.3.1 Filter design

The band-pass filters $b_i = \widehat{\psi_i}$ (whereas $\widehat{\cdot}$ denotes the Fourier transform) are constructed as follows. To obtain quickly decaying wavelets, the filters need to be smooth in the spectral domain. Therefore, the shape of a cosine functions is used,

$$c_i(x) := \begin{cases} \frac{1}{2} + \frac{1}{2}\cos(\pi(x-i)) & \text{if } x \in [i-1, i+1]; \\ 0 & \text{else.} \end{cases} \qquad (4.10)$$

With the re-definition $c_1(x) := 1$ for $x < 1$ and $c_{n_e}(x) := 1$ for $x > n_e$, we obtain the property $\sum_{i=1}^{n_e} c_i(x) = 1$, everywhere.
In the implementation, $n_e = 10$ and absorption coefficients for the frequencies $\omega_i = 20 \cdot 2^i$ Hz $(i = 1, ..., 10)$, covering the range from 40 to 20480 Hz are used. On the logarithmic scale, the filters are defined as

$$b_i(\omega) = \begin{cases} c_i(log_2(\omega/\omega_0)) & \text{if } \omega > 0; \\ \delta_{i1} & \text{if } \omega = 0; \\ c_i(log_2(\omega_0/-\omega)) & \text{if } \omega < 0, \end{cases} \qquad (4.11)$$

where $\omega_0 = 20$ Hz and δ_{i1} denotes the Kronecker delta. The wavelets ψ_i are obtained by the inverse Fourier Transform. b_1 and b_{10} are low- and high-pass filters, respectively, whereas the remaining b_i are bandpass filters with support $[\omega_{i-1}, \omega_{i+1}]$, see figure 4.3. The filters have the following properties:

- compact support in Ω
- symmetry and smoothness in both domains
- their sum is one in Ω and a Dirac in the time domain.

(a)   (b)
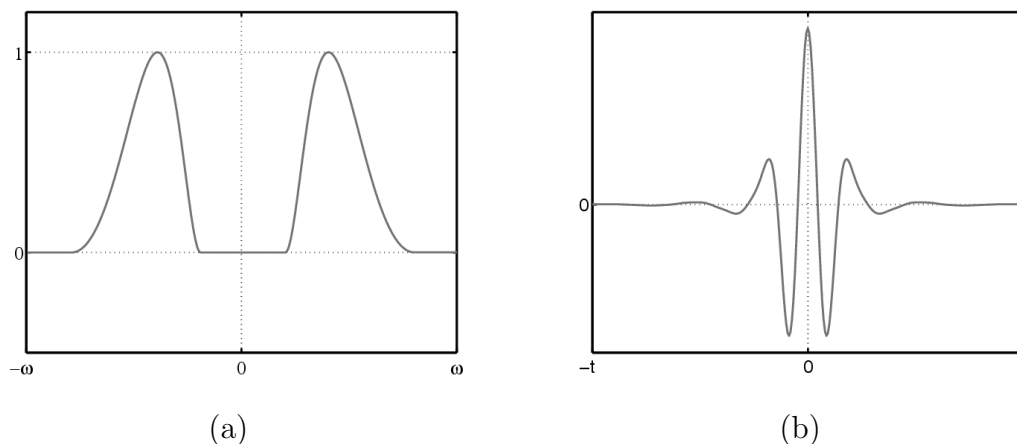
Figure 4.3: Bandpass filter. (a) in spectral domain, (b) in time domain.

4.1.3.2 Discretization

The energy response filters produced by the implementation are sampled at a rate of 48kHz. For generating the filter bank, 2^{14} samples are used, employing the inverse FFT (see appendix B) for discretized wavelets. Due to the symmetry, the impulse response filters are non-causal. To limit this effect, the widest filter for low frequencies are combined into one (decaying much faster) and all filters are cut off after 314 samples (corresponding to 6.5 msec.) in both directions. Hence, the low-frequency portion of an echo, filtered due to different absorption coefficients, may be received up to 6.5 msec. in advance of the main peek, which in fact cannot be recognized by human ear. In the case of uniform absorption, the discretized wavelets still form the partition of the unit impulse, providing perfect echos.

Once, the energy response filter f has been computed, it needs to be normalized with respect to the number of phonons used and the radius of the Gaussian located at the listener. Therefore, the filter is normalized such that the signal's intensity at 1 meter distance from the sound source S (without any reflections) corresponds to the intensity of the anechoic source signal. The normalization factor is simply one over the sum of all Gaussian terms collected from the outgoing phonon paths in 1 meter distance from the source.

An extension which is considered by the algorithm is the absorption due to air described in section 2.1.2.4.

4.2 Tracing the sound pressure

This section describes the modification of the phonon tracing approach, where sound pressure is used for calculations, instead of energy. Thus, phase changes due to scattering can be integrated in the algorithm easily, and interference phenomena can be simulated. The main idea is, analogous to the algorithm described in the previous section, to trace sound particles outgoing from the sound source through the given scene building the phonon map. Afterwards the phonons are collected in order to calculate the room impulse response at a given listener position. In contrast to the energy approach where quadratic attenuation is used because of the spatial particle density, linear pressure attenuation is facilitated. This is obtained by modeling linear pressure attenuation by Gaussian basis functions dilated proportional to the traversed distance (see figure 4.4). This way the approximate partition of unity for the basis functions is preserved. The pressure amount of a particle is scaled inverse proportional to the traversed distance due to the linear attenuation of the sound. The single particles can be considered as individual micro-sources in analogy to the image-source method.
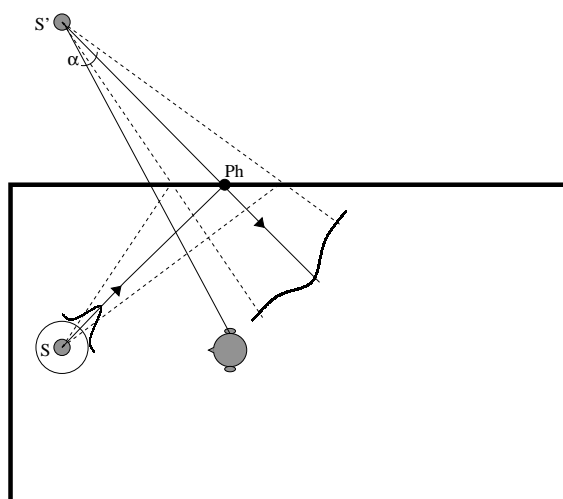
Figure 4.4: Sketch of the phonon tracing algorithm to trace sound pressure inside closed rooms.

4.2.1 Algorithm specification

We define:

- S_2 a unit sphere in $\mathbb{R}^3$, i.e. $S_2 = \{x \in \mathbb{R}^3; \|x\| = 1\}$

- $\{v_i\}_{i=1,...,N}$ a set of directions distributed on S_2 with regard to the emission distribution of the sound source S

- $\psi_i : S_2 \mapsto \mathbb{R}, x \mapsto g(arccos(\langle x, v_i \rangle)), i = 1, ..., N$ the partition of one over S_2

We consider a spherical sound source S in point $x_0 = 0$ in $\mathbb{R}^3$. Then for a unit impulse (Dirac) δ of the length $\Delta t = \frac{1}{f_s}$ the pressure $P(t, x)$ at time $t \in \mathbb{R}$ in x reads:

$$P(x, t) = P_1 \frac{1}{\|x\|} \delta \left(t - \frac{\|x\|}{c} \right) \sum_{i=1}^{N} \psi_i \left(\frac{x}{\|x\|} \right) \tag{4.12}$$

where P_1 is the pressure in 1 m distance from the sound source S, c is the sound velocity, f_s the sampling rate, and $\delta(t)$:

$$\delta(t) = \begin{cases} 1 & \text{if } t \in [0, \Delta t] \\ 0 & \text{else} \end{cases} \tag{4.13}$$

As mentioned before we choose g as Gaussian basis function of an angle $\vartheta = arccos(\langle x, v_i \rangle)$, $i = 1, ..., N$:

$$g(\vartheta) = \alpha \cdot exp \left(-\frac{\vartheta^2}{2\sigma^2} \right) \tag{4.14}$$

For evaluation of the room acoustics a uniform distribution of sound rays is used. In order to satisfy the partition of unity of the basis functions for the uniform distribution (the needed parameters can also be derived for unequally distributed direction vectors), it must hold:

$$\oint_{S_2} g \, \mathrm{d}S \;=\; \frac{4\pi}{N} \Leftrightarrow \tag{4.15}$$

$$\int_{\phi=0}^{2\pi} \int_{\vartheta=0}^{\pi} \alpha \cdot exp \left(-\frac{\vartheta^2}{2\sigma^2} \right) \cdot sin\vartheta \, \mathrm{d}\vartheta \mathrm{d}\phi \;=\; \frac{4\pi}{N} \Leftrightarrow$$

$$\alpha \;=\; \frac{2}{N} \left(\int_{\vartheta=0}^{\pi} exp \left(-\frac{\vartheta^2}{2\sigma^2} \right) \cdot sin\vartheta \, \mathrm{d}\vartheta \right)^{-1}$$

where 4π is the surface area of the unit sphere S and N the number of direction vectors. The integral in the above equation has to be solved numerically, but if we assume $\sigma \ll 1$, then it holds:

$$\int_{\vartheta=0}^{\pi} exp\left(-\frac{\vartheta^2}{2\sigma^2}\right) \cdot sin(\vartheta)\, \mathrm{d}\vartheta \approx \int_{\vartheta=0}^{\pi} exp\left(-\frac{\vartheta^2}{2\sigma^2}\right)\vartheta\, \mathrm{d}\vartheta \qquad (4.16)$$

By substituting $u = -\frac{\vartheta^2}{2\sigma^2}$ and $\mathrm{d}u = -\frac{1}{\sigma^2}\vartheta\, \mathrm{d}\vartheta$ we get:

$$\begin{aligned}
\int_{\vartheta=0}^{\pi} exp\left(-\frac{\vartheta^2}{2\sigma^2}\right)\mathrm{d}\vartheta &= -\sigma^2 \cdot \int_0^{-\frac{\pi^2}{2\sigma^2}} exp(u)\, \mathrm{d}u \\
&= \sigma^2 \left[exp(0) - exp\left(-\frac{\pi^2}{2\sigma^2}\right)\right] \\
&\approx \sigma^2
\end{aligned} \qquad (4.17)$$

This leads to:

$$\alpha = \frac{2}{N\sigma^2} \qquad (4.18)$$

Choosing σ there is a trade-off between smoothness of the partition of unity and the resolution of geometric details of the scene. For better approximation of the partition of unity the support of the Gaussian function g should contain n different direction vectors. The surface area for a direction reads:

$$a = n\frac{4\pi}{N} \qquad (4.19)$$

Furthermore to fulfill the partition of unity it should hold:

$$\begin{aligned}
exp\left(-\frac{\vartheta_0^2}{2\sigma^2}\right) &\geqslant \frac{1}{2} &\Rightarrow \\
\frac{\vartheta_0^2}{2\sigma^2} &\leqslant ln2 &\Rightarrow \\
\vartheta_0 &= \sigma\sqrt{2ln2} &\Rightarrow \\
n\frac{4\pi}{N} &= 2\pi \int_0^{\vartheta_0} sin(\vartheta)\, \mathrm{d}\vartheta & \\
&= 2\pi(1 - cos(\vartheta_0)) & \\
&= 2\pi(1 - cos(\sigma\sqrt{2ln2})) &\Leftrightarrow \\
\sigma &= \frac{arccos\left(1-\frac{2n}{N}\right)}{\sqrt{2ln2}} &
\end{aligned} \qquad (4.20)$$

Thus, the Gaussian basis function with an approximate partition of unity on the unit sphere S reads:

$$g(\vartheta) = \frac{2}{N\sigma^2} \cdot exp\left(-\frac{\vartheta^2}{2\sigma^2}\right) \tag{4.21}$$

Halton sequence. For the accuracy of the Monte-Carlo integration, the uniformity is more important than the randomness. The Halton quasi-random sequence (also applied in computer graphics [Kel96a, Kel96b]) is used for generating direction vectors. The Halton sequence is obtained by calculating the radical inverse function, which takes a number $i \in \mathbb{N}$ represented in prime base p and reflects it through the radical point:

$$\Phi_p(i) := \sum_{j=0}^{\infty} a_j(i)p^{-j-1} \in [0,1) \Leftrightarrow i = \sum_{j=0}^{\infty} a_j(i)p^j \tag{4.22}$$

The m dimensional Halton sequence reads:

$$z_i = (\Phi_{p_1}(i), ..., \Phi_{p_m}(i)) \tag{4.23}$$

The radical inverse function can be intutively understood in the following way. To get the radical inverse of i in base p, we write down the digits of i in the base p, then we invert their order and place the floating point in the beginning. The resulting number is a desired floating point number, expressed in base p. The Halton sequence is incrementally, i.e. it is possible to increase the number of samples without discarding the already calculated samples. The algorithm described in [HS64] can be used for the calculation of the Halton sequence, which is not more expensive as determination of usual pseudo-random numbers.

4.2.2 Implementation

Now, as we have defined the Gaussian basis functions we can formulate the phonon tracing algorithm to trace pressure. The algorithm requires the following input parameter:

- position of the sound source S
- reference pressure P_1 in 1m distance from S
- emission distribution E of S
- one ore more listener positions l_i
- a triangulated scene with tagged materials m_j

- a reflection function $\rho_j : \Omega \mapsto [-1, 1]$ for each material
- an acoustic BRDF for each material (if applicable)
- a number of phonons n_{ph} to be traced from the source
- a lower pressure threshold ε and a maximum number of reflections n_{refl} for terminating the phonon paths

The output of our approach is a FIR filter f_i for each listener's position l_i corresponding to the impulse response with respect to the sound source and the phonon-map. For each phonon ph the map contains the pressure spectrum p_{ph}, the phonon's position pt_{ph} at the reflection point, the image source q_{ph} from which we can calculate the phonons outgoing direction v_{ph} and the traversed distance d_{ph}, the number of reflections r_{ph}, and the material m_{ph} at the current reflection.

4.2.2.1 Phonon emission step

The phonons sent out from the source are associated with the following quantities:

- a pressure spectrum $p_{ph} : \Omega \mapsto \mathbb{R}^+$
- the virtual source q_{ph}
- the phonon's current position pt_{ph}

As described in section 4.1, $n_e = 10$ frequency bands are used for the simulation. Furthermore, the wavelets introduced in section 4.1 are the basis functions for the pressure spectrum.

Phonons are emitted from the source S according to the emission probability distribution E and have a unit pressure spectrum $p_{ph,i} = 1$ $(i = 1, ..., n_e)$ at their starting point. At the intersection of the phonon ray with the scene, the virtual source q_{ph} is calculated as follows:

$$q_{ph} \quad \leftarrow \quad q_{ph} + 2 \cdot \langle (pt_{ph} - q_{ph}), n \rangle \cdot n \tag{4.24}$$

where n is the surface normal at the intersection point pt_{ph}, and $\langle \cdot, \cdot \rangle$ denotes the scalar product. The pressure is reduced according to the reflection function ρ_j of the local material m_j. The phonon is fixed at the intersection point, and stored in the global phonon map. If the termination conditions are not satisfied, the tracing process is continued. Otherwise a new phonon is started from the source.

4.2.2.2 Phonon collection and filtering

The phonons are now collected in order to build the impulse response f_i at the listener position l_i. In the case of the uniform absorption for all frequencies the

contribution of a single phonon to the impulse response, corresponding to equation 4.12 reads:

$$p(t, x) = \frac{\rho_{tot} \, P_1}{d_{ph}} \, g\left(\angle \left(v_{ph}, l_i - q_{ph} \right) \right) \cdot \delta\left(t - \frac{d_{ph}}{c} \right) \qquad (4.25)$$

where P_1 is the reference pressure at 1m distance from the source, ρ_{tot} is the product of the reflection coefficients along the phonon path and g is a Gaussian weighting function defined in equation 4.21 with $N = n_{ph}$.

Since most absorption coefficients α provided in the literature refer to energy, we have to calculate the pressure related reflection coefficients:

$$\rho = \sqrt{1 - \alpha} \; . \qquad (4.26)$$

In the general case of a frequency-dependent absorption, the unit pulse is subdivided as presented in section 4.1 (see also [BDM+05]). The filter f_i then becomes a sum of wavelets (band pass filters) scaled and shifted as described above. Analogously to the energy phonon tracing algorithm the air absorption is considered.

4.3 Modeling diffuse reflections

In the phonon tracing algorithm described in the previous sections only specular reflections are considered. In fact for virtual acoustic applications surfaces are generally assumed to reflect the sound wave specularly. This assumption is true if the surface structure is significantly smaller compared to the wave length. Otherwise, the sound wave is reflected diffusely (see section 2.1.2.2). This section describes the diffuse reflection model used in the phonon tracing algorithm.

In order to model diffusion in room acoustics, a scattering coefficient δ_i for each scene surface is introduced. The *phonon emission step* of the phonon tracing algorithm is then modified as follows. If the sound particle hits a scene surface, a number $\xi \in [0, 1]$ is randomly chosen. If $\xi > \delta_i$ the phonon is reflected diffusely, otherwise the reflection is specular. In the latter case the outgoing direction of the phonon is calculated as described in previous sections. In case of diffuse reflection the outgoing direction v_{ph} is determined assuming ideal (Lambertian) diffuse reflection where the direction of the reflection is perfectly random over a hemisphere surrounding the phonon's position pt_{ph} (intersection point of the ray with the surface) [Kut71]. The azimuthal angle ϑ is a random number ξ_1 in the interval $(-\pi, \pi)$ and the polar angle ϕ is given by the arc-cosine of the square root of a random number ξ_2 chosen in the

interval $(0, 1)$:

$$\vartheta = \xi_1 \qquad \phi = arccos(\sqrt{\xi_2}) \tag{4.27}$$

In the pressure phonon tracing algorithm the virtual source in equation 4.24 is now calculated as follows:

$$q_{ph} = pt_{ph} - d_{ph} \cdot v_{ph} \tag{4.28}$$

where pt_{ph} is the phonon's current position and d_{ph} the traversed distance. Figure 4.5 depicts the *phonon emission step* including the diffuse reflection model.
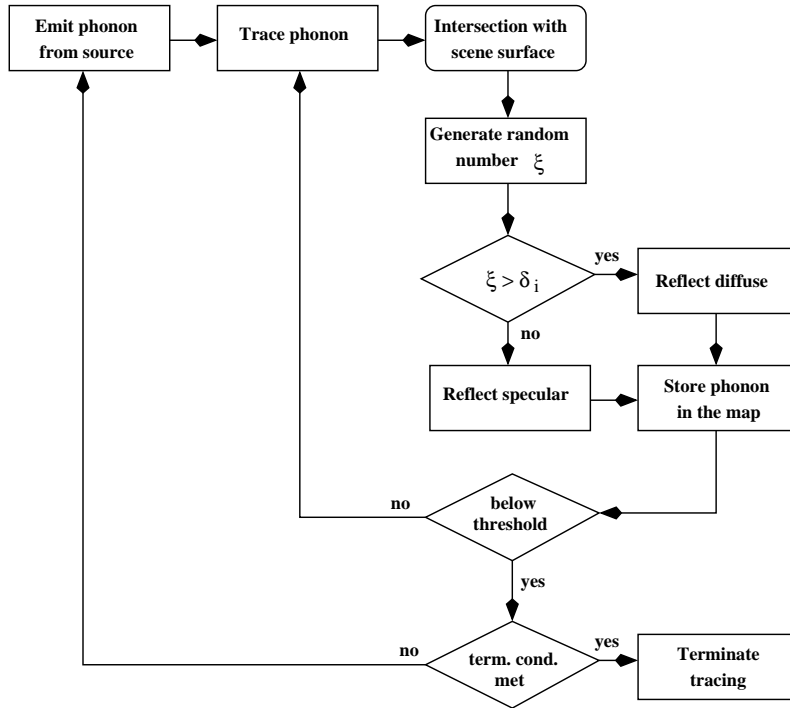
Figure 4.5: Flow chart diagram of the phonon emission step including diffuse reflections.

Since the scattering coefficients δ_i are frequency dependent we have to either perform a tracing step for each frequency band or to use an average scattering coefficient for all frequencies. A better solution for this situation is subject of future research.

4.4 Acoustic level of detail

In computer graphics, multiple techniques considering the level of detail representation of geometric objects exist [LRC+03]. In Virtual Reality, for example, the objects located in a remote distance from the viewer position are rendered using fewer triangles or patches reducing rendering time. More details are added when the viewer approaches an object providing a more accurate representation. This section describes some similar concepts for acoustic simulation, i.e. how the phonon tracing algorithm can be modified in order to provide calculation of the room impulse response at different levels of detail depending on user requirements on computation time and computation accuracy.

4.4.1 Clustering

4.4.1.1 Cluster definition

Before we discuss the different levels of detail approaches we need to partition the particles stored in the phonon map into clusters which represent different wave fronts reflecting at the walls. We can use the information stored in the phonon map to accomplish this partitioning process. For early reflections all phonons within a cluster have the same trajectory that means that they reflected at the same surface having equal absorption properties. Thus, they satisfy the following criteria:

- equal numbers of reflections n_{ph}
- and for each reflection:
    - equal material indices m_{ph} (same object of the scene)
    - equal surface normals n at the reflection position

Consequently, phonons included in an early reflection cluster have equal virtual source positions and absorbed pressure fraction.

At higher reflection orders, the clusters become smaller and smaller, until they contain only a single phonon. Hence, for cluster reconstruction at higher reflection order the constraint of equal trajectory is disregarded and all phonons having with the same last reflection material are combined to a cluster. The cluster criteria are then as follows:

- equal material index m_{ph} of last reflection
- equal surface normal n at the last reflection position

A cluster contains the residual phonons (late reflections) at different room surfaces after a prescribed number of first reflections. All the phonons within such

a residual cluster have different virtual sources and also different pressure/energy
decomposition.

4.4.1.2 Data structure

In the following the data structure of the clustered phonon map is described. Since
the construction of the map is partially performed during the emission step of the
phonon tracing algorithm, a dynamic data structure is needed. Here, a singly-linked
list structure is used. The principal benefit of a linked list over a conventional array
is that the order of the linked items may be different from the order that the data
items are stored in memory or on disk, allowing the list of items to be traversed in
a different order. A linked list is a self-referential data-type because it contains a
pointer or link to another datum of the same type. Linked lists permit insertion and
removal of nodes at any point in the list in constant time.
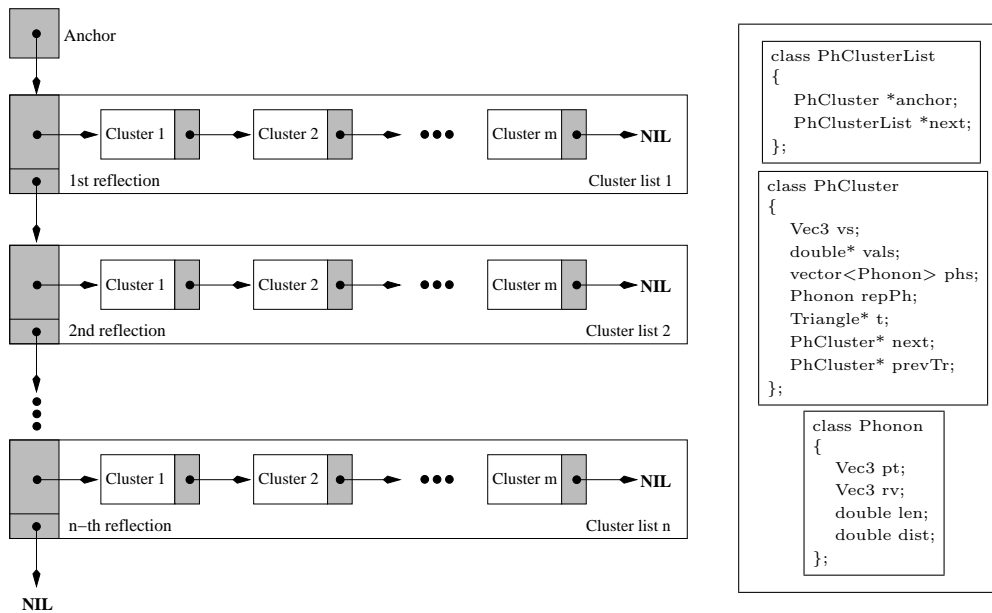
Figure 4.6: Data structure for early reflections clustered phonon map.

In figure 4.6 the data structure for the early reflection clustered map is depicted.
The clusters are grouped into a cluster list containing all clusters with the same
number of reflections. Since the phonons of each early reflection cluster have the
same trajectory, as described in the last subsection, the absorbed pressure fraction
(double* *vals*) as well as the position of the virtual source (Vec3 *vs*; see the phonon
tracing algorithm description in section 4.2) are stored for the cluster instead of stor-
ing them for each individual phonon. In addition to the phonons (vector<Phonon>

phs), every cluster contains the information about the material and the reflecting surface (covered by the Triangle* *t*), a pointer to its successor (PhCluster* *next*), a pointer to a cluster in a predecessor cluster (PhCluster* *prevTr*) marking the trajectory, and a representative phonon (Phonon *repPh*), as described in the next subsection. For a single phonon its position (Vec3 *pt*), outgoing direction (Vec3 *rv*), traversed distance (double *len*), as well as its distance to the representative phonon (double *dist*) needed for further computations is stored. For storing the representative phonon the phonon class is used, too. The number of reflections is implicitly given by the arrangement of the cluster lists.
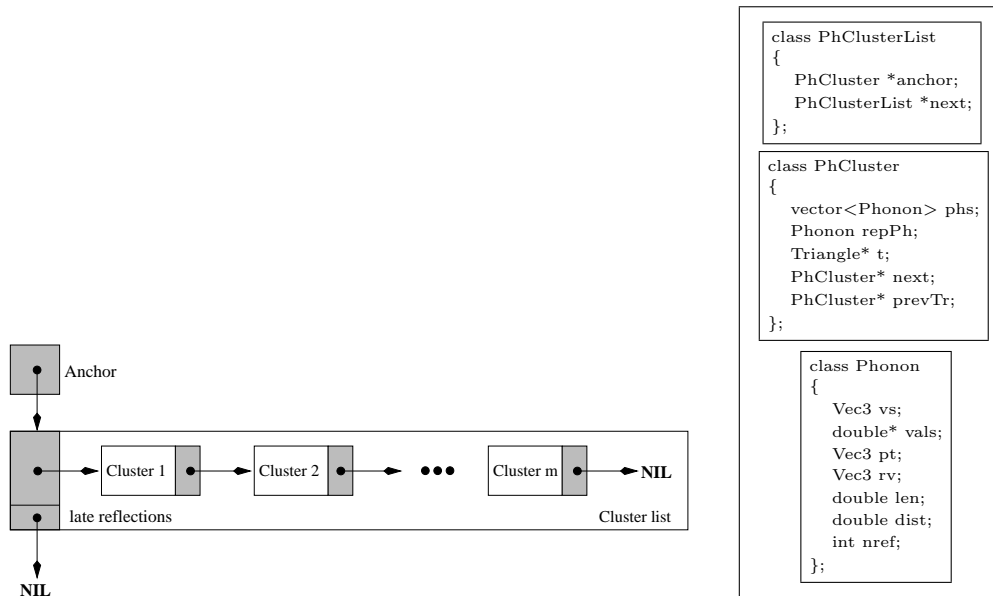
Figure 4.7: Data structure for late reflections clustered phonon map.

The late reflection clusters (residual clusters) are stored in a separate cluster list (see figure 4.7). As mentioned before, the phonons inside these clusters have different trajectories and thus different position of virtual source and absorbed energy/pressure fraction. Therefore, these values have to be stored for each individual phonon. The material and plane information are equal for common cluster and are stored per cluster.

For visualization purposes we need the exact trace for each phonon. Therefore, a separate data structure containing a list with references to the corresponding clusters and the phonons inside these clusters is used.

4.4.1.3 Calculation algorithm

Since the number of the phonons in the map as well as the number of the phonon paths can be large, the cluster calculation may be time extensive. For each phonon we need to compare the material and surface information with other phonons. In order to speed up this process, the clusters (initially containing only one phonon) inside the cluster list are sorted and afterwards equal clusters are grouped into one cluster. The sort algorithm is described below.

Before we discuss the clustering algorithm, we first describe the methods comparing two clusters which are used during the sorting and cluster building. Pseudo-code of these methods is outlined in figure 4.8. The method `compare_sort()` is used for the comparison of two clusters during the sorting task of the cluster list. Thereby the clusters are ordered by their material id, surface id, and the path traversed. Since initially each cluster contains only one phonon, it is sufficient to compare the pointer to the predecessor cluster in the previous cluster list. The methods `compare_early()` and `compare_late()` are used for the comparison of two clusters inside the early and late reflections cluster lists, respectively. These methods are applied during the second step of the clustering algorithm described below. The difference between these comparison methods is that the traversed paths of the phonons do not need to be compared for the last reflection clusters. For early reflection clusters the paths are previously compared. If the triangle pointers are not equal, comparison of material and surface (same scene object) is required.

Based on comparison of clusters, we now describe the clustering algorithm. Figure 4.9 shows the pseudo-code of the algorithm. The cluster phonon map is populated during the tracing of the sound particles through the virtual scene. For each early reflection a cluster list is created. Initially at a reflection of the phonon at a scene object a cluster containing this phonon is added to the list including the clusters at this reflection order and accurately linked. The number of the early reflections cluster lists corresponds to the maximum number of reflections. One additional list for late reflections clusters is created. Following the emission step, duplicated clusters are merged completing the cluster building process. First, each cluster list is sorted utilizing the comparison method described above. Afterwards, the phonons of equal clusters are combined and the empty cluster is deleted. Since the list is sorted, only one pass is required. This step is applied to early and late reflections cluster lists, utilizing the corresponding comparison method.

```
boolean compare_early(PhCluster* cl1, PhCluster* cl2)
```
```
#Check if the same trace (pointer comparison only)
IF cl1->prevTr != cl2->prevTr THEN
  RETURN FALSE
#Check if the same triangle (pointer comparison only)
IF cl1->tri == cl2->tri THEN
  RETURN TRUE
#Check if the same material
IF cl1->tri->imat != cl2->tri->imat) THEN
  RETURN FALSE
#Check if triangle in the same plane
IF cl1->tri->iplane != cl2->tri->iplane) THEN
  RETURN FALSE
#Else equal cluster
RETURN TRUE
```
```
boolean compare_late(PhCluster* cl1, PhCluster* cl2)
```
```
#Check if the same triangle (pointer comparison only)
IF cl1->tri == cl2->tri THEN
  RETURN TRUE
#Check if the same material
IF cl1->tri->imat != cl2->tri->imat) THEN
  RETURN FALSE
#Check if triangle in the same plane
IF cl1->tri->iplane != cl2->tri->iplane) THEN
  RETURN FALSE
#Else equal cluster
RETURN TRUE
```
```
int compare_sort(PhCluster* cl1, PhCluster* cl2)
```
```
#Compare material
IF cl1->tri->imat < cl2->tri->imat) THEN
  RETURN -1
IF cl1->tri->imat < cl2->tri->imat) THEN
  RETURN 1
#Compare plane
IF cl1->tri->iplane < cl2->tri->iplane) THEN
  RETURN -1
IF cl1->tri->iplane > cl2->tri->iplane) THEN
  RETURN 1
#Check if the same trace (memory address comparison only)
IF cl1->prevTr < cl2->prevTr THEN
  RETURN -1
IF cl1->prevTr > cl2->prevTr THEN
  RETURN 1
#Else equal cluster
RETURN 0
```

Figure 4.8: Pseudo-code of the cluster comparison methods.

During the phonon emission build cluster lists

At reflection of the phonon *ph*
Create new cluster *cl*
Set triangle pointer *t* to the triangle at the reflection of *ph*
Add *ph* to *cl*
IF *nref* $<=$ max early reflections **THEN**
  Add *cl* to the *nref*-th early reflections cluster list
  IF first reflection **THEN**
    Set *prevTr* pointer **NIL**
  ELSE
    Set *prevTr* to the cluster of previous reflection in the (*nref*-1)-th cluster list
  ENDIF
ELSE
  Add *cl* to the late reflections cluster list
  Set *prevTr* pointer **NIL**
ENDIF

After the emission step build clusters

FOR all early reflections cluster lists **DO**
  Sort phonon clusters inside the list
  Begin with the first cluster in the list *cl1 = cl(0)*
  Set *cl2 = cl(1)*
  WHILE NOT End of the cluster list **DO**
    IF compare_early(*cl1*, *cl2*) == **TRUE THEN**
      Add phonon *ph* to *cl1*
      cl1->next = cl2->next
      Delete *cl2*
    ELSE
      cl1 = cl2
    ENDIF
    cl2 = cl1->next
  ENDWHILE
ENDFOR

Sort phonon clusters inside the late clusters list
Begin with the first cluster in the late reflections list *cl1 = cl(0)*
Set *cl2 = cl(1)*
WHILE NOT End of the cluster list **DO**
  IF compare_late(*cl1*, *cl2*) == **TRUE THEN**
    Add phonon *ph* to *cl1*
    cl1->next = cl2->next
    Delete *cl2*
  ELSE
    cl1 = cl2
  ENDIF
  cl2 = cl1->next
ENDWHILE

Figure 4.9: Pseudo-code of the clustering algorithm.

4.4.2 LoD approaches

Using the cluster information the impulse response for a give source and listener position can be calculated at different levels of detail. The LoD algorithm is summarized in the following steps:

1. Find a representative phonon ph_r for each cluster

2. Calculate distance d_i of each phonon in the cluster to ph_r

3. Sort phonons according to d_i

4. Collect only phonons which are in a user prescribed distance d from ph_r or collect only a user-prescribed number of first phonons.

Next we will discuss the individual steps in detail.
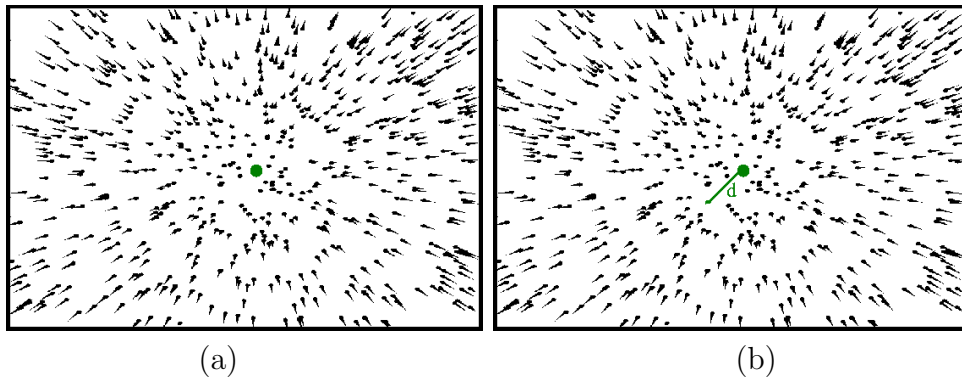
Figure 4.10: Phonon cluster. (a) Representative phonon of the cluster (green point). (b) Distance of one cluster phonon to the representative phonon.

Step 1. In the first step we determine the position pt_{ph_r} of the representative phonon ph_r for each cluster consisting of n particles (see figure 4.10(a)):

$$pt_{ph_r} = \frac{\sum_{i=1}^{n} w_i \cdot pt_{ph,i}}{\sum_{i=1}^{n} w_i} \tag{4.29}$$

The weight w_i used for a single phonon depends on the phonon's pressure (or energy) absorption factor due to reflections on the room walls. Since all phonons in an early reflection cluster have equal absorption factor the weights are all equal to 1 ($w_i = 1$). Whereas inside the late reflection clusters the weight w_i grows inversely proportional with this factor. Thus the representative phonon ph_r is shifted towards phonons having less pressure (energy) amount absorbed due to reflections at least in one frequency band.

Step 2. Now, we calculate the distance d_i (see figure 4.10(b) of each phonon inside the cluster as Euclidean distance to ph_r:

$$d_i = |pt_{ph,i} - pt_{ph_r}| \qquad (4.30)$$

The distances are used for the adaptive collection of particles in the next step.

Step 3. In order to ensure a rapid collection of the particles we sort them in ascending order of the distance d_i.

Step 4. The information precalculated in the previous steps is now used for the collection of the particles. The user can prescribe the cluster size by a distance from the representative phonon ph_r. All phonons which are not farther away than this distance are collected. Another possibility is to predefine the number m of phonons to collect. Then only the m nearest particles to ph_r in the cluster are considered for the impulse response calculation. Thus, for the lowest level of detail only the representative phonon of each cluster is collected and all particles are used for the finest detail calculation. An example of this task with resulting impulse responses is depicted in figure 4.11.

Concerning the clustered phonon map, several modifications of the level of detail algorithm may be considered. In room acoustics, early reflections are of greater interest than late reverberations, thus an approach utilizing this would be more appropriate. Additional benefits regarding the clustering should be investigated. For example greater reduction of memory use for the phonon map storage. All phonons inside early reflection clusters have equal absorbed pressure (energy) fraction, thus this information needs to be stored only once for the cluster and not for each phonon separately. Also, the position of the virtual source is equal, so we can store only the position of the virtual source and the pressure (energy) absorption factor for the cluster at the room surface but drop the actual phonons. Statistical methods can then be used to perform the collection step, for example, collection of a prescribed number of particles equally distributed in the cluster. The outgoing direction of the phonon can be calculated using the virtual source position for specular reflections or can be random in case of diffuse reflection. Therefor the regions of the clusters need also to be determined and stored. Regarding the late reflections particle clusters, more information has to be stored, due to the fact that the particles have both different pressure (energy) decomposition and different virtual source positions. Statistical approaches can be applied if the pressure (energy) as well as the virtual source positions distributions can be determined.
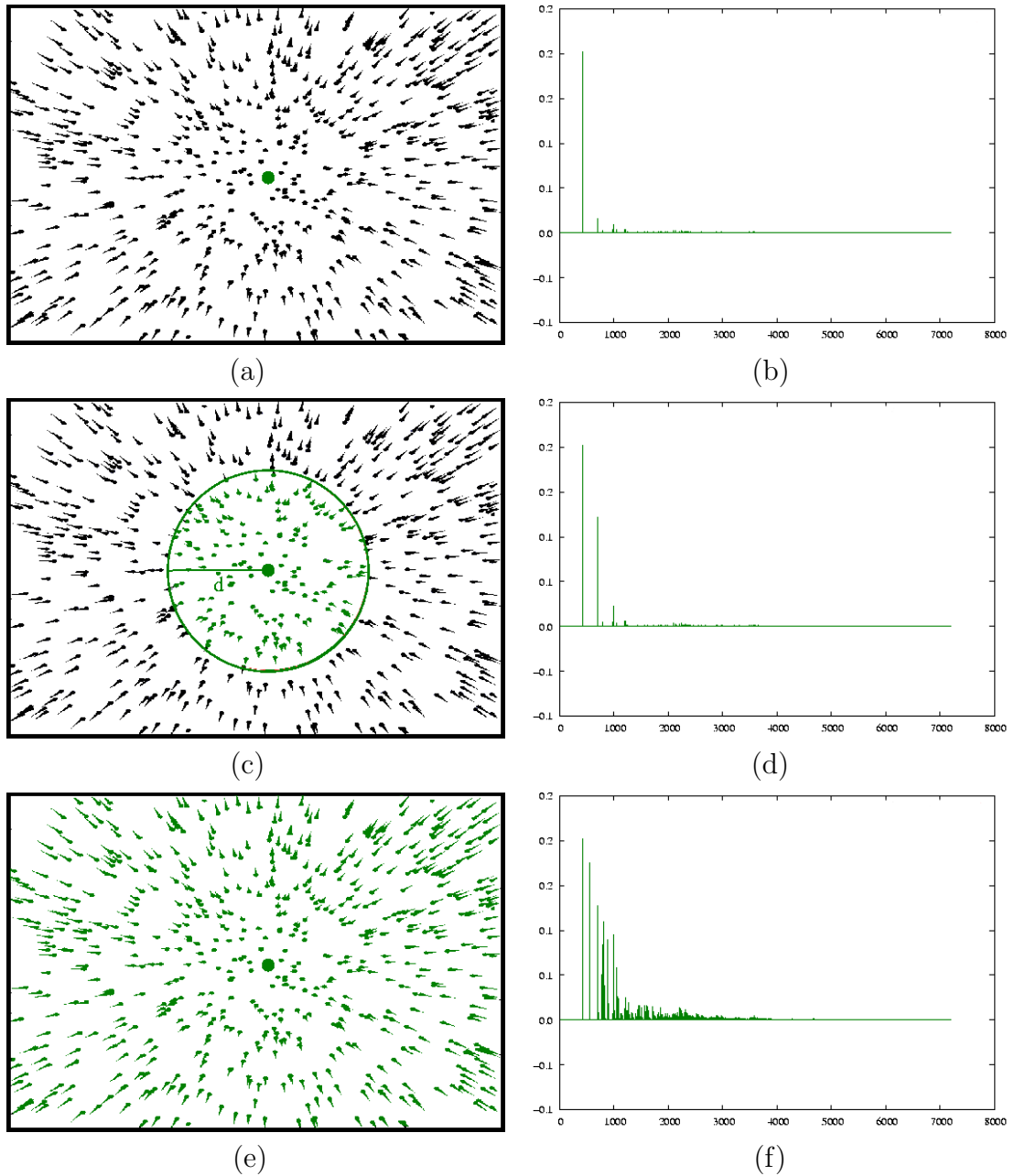
Figure 4.11: Impulse responses calculated at different levels of detail. (a+b) Only the representative phonons are collected. (c+d) Phonons collected until a prescribed distance around the representative phonon. (e+f) All phonons of the clusters collected.

Choosing the level of detail (by number of collected phonons per cluster or distance from the representative phonon) there is a trade-off between the computation time and the accuracy of the simulation results. Consider the following scenario with one sound source and two listener positions depicted in figure 4.12. One of the listeners has no direct view to the sound source.
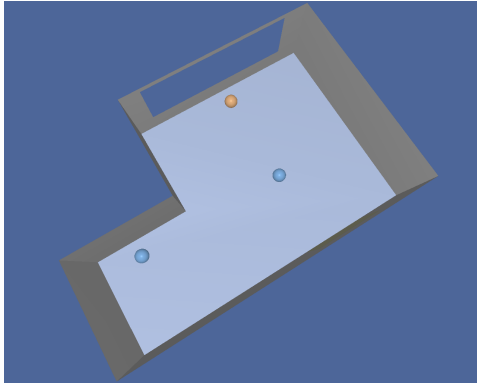
Room dimensions (max values):	$12 \times 8 \times 3$ m
Absorption:	0.1
Number phonons traced:	50000
Number phonons in the map:	1548648

Figure 4.12: Simulation setup with one source and two listeners. One of the listener has no direct view to the source.

(a)   (b)
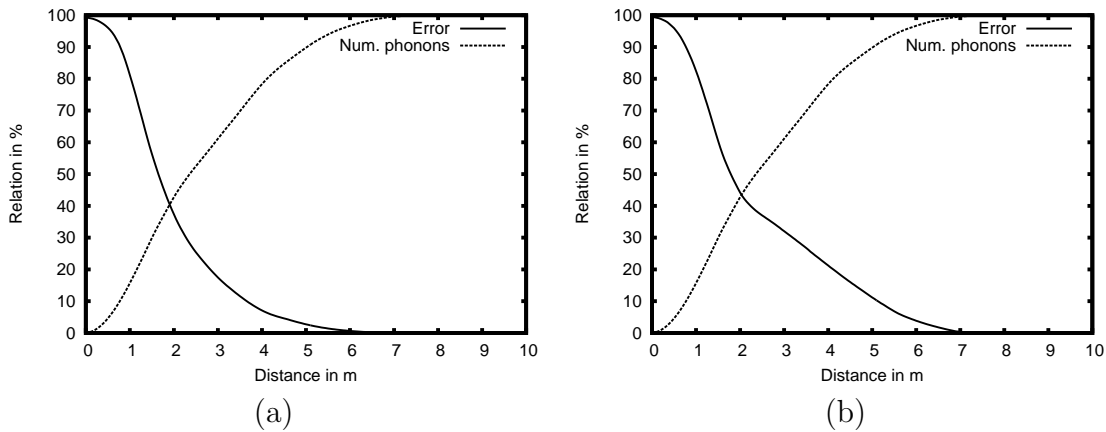
Figure 4.13: Error of the mean sound pressure level of impulse responses at different levels of detail related to the mean sound pressure level of the full detailed impulse response. (a) For the listener having the direct view to the source. (b) For the listener receiving no direct sound.

For both listeners the impulse responses are calculated at different levels of detail using the distance from the representative phonon as described above. Figure 4.13 shows the error of the mean sound pressure level of each impulse response related to the mean sound pressure level of the full detailed impulse response and the number of collected phonons related to the total number of phonons in the map, both

depending on the chosen distance from the representative phonon. As expected the error decreases with the increased distance and thus the increased number of collected phonons. However, the error at the position with no received direct sound (figure 4.13 (b)) is greater than that at the other listener position (figure 4.13 (a)) for the same number of collected phonons. The reflections for the listener with no direct view to the source are more important considering the sound pressure level of the signal.

4.5 Evaluation

4.5.1 Computation times

(a)	(b)	(c)
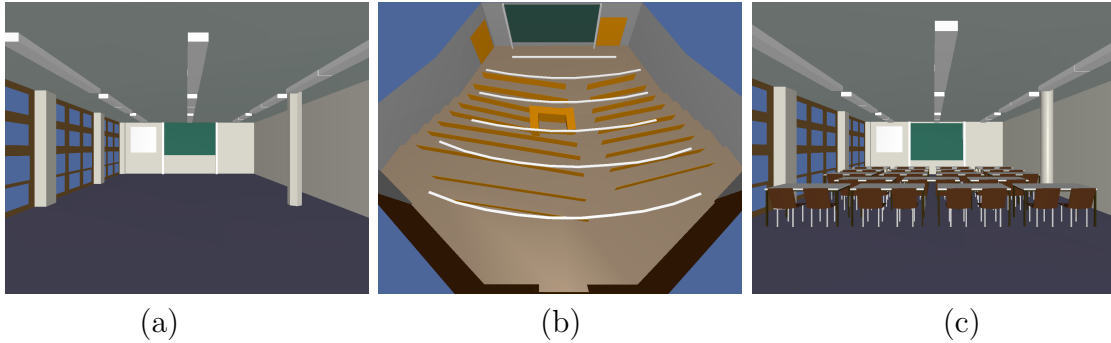
Figure 4.14: Models of lecture rooms consisting of (a) 436 triangles, (b) 1255 triangles, and (c) 16492 triangles.

In order to evaluate the performance of the phonon tracing algorithm we run it on three different room models with different conditions. The models consist of 436, 1255, and 16492 triangles, respectively, and are depicted in figure 4.14. The room surfaces are full absorptive in all three scenes. We execute the algorithm for 6, 10, 20, and 40 maximum numbers of reflections and trace 20K, 50K and 100K phonons from the source. The computation times for several stages of the algorithm on a PC with a 3GHz processor and 2GB RAM are listed in table 4.1. The table shows that the time complexity of the tracing stage of the method does not depend very much on the complexity of the scene, due to kd-tree partitioning. The times for collecting phonons and filtering are about the same as the tracing times, due to the visibility test. The cluster building task depends linearly on the number of the clusters and the number of the phonons in the phonon map (which depends on the number of reflections). Furthermore the collecting step can be scaled using the level of detail approach described in the previous subsection.

As we can see the phonon tracing algorithm is scaleable and can handle complex scenes in reasonable times on a standard PC. The algorithm can be accelerated for example utilizing CPUs with multiple cores running the simulation on several threads.

Room model	#paths/ #phonons $\times 10^3$	#reflections	#clusters	tracing [sec]	clustering [sec]	collecting [sec]
	20/120	6	9410	1.14	0.54	1.17
	20/200	10	9416	1.95	1.03	2.09
	20/400	20	9416	4.36	2.49	4.15
	20/800	40	9417	9.60	5.67	8.79
	50/300	6	13266	2.84	1.56	2.94
Lecture room	50/500	10	13267	4.89	3.02	5.26
436 triangles	50/1000	20	13267	10.90	7.38	10.44
	50/2000	40	13267	24.06	16.46	22.04
	100/600	6	16266	5.66	3.46	5.93
	100/1000	10	16266	9.66	6.83	10.59
	100/2000	20	16266	21.53	16.44	20.98
	100/4000	40	16266	48.48	39.70	44.13
	20/120	6	11772	1.45	0.54	1.58
	20/200	10	11772	2.36	1.02	2.82
	20/400	20	11773	5.28	2.41	5.35
	20/800	40	11773	11.63	5.45	11.30
	50/300	6	16977	3.53	1.54	3.93
Lecture room	50/500	10	16978	5.93	2.95	7.07
1255 triangles	50/1000	20	16978	13.00	7.09	13.42
	50/2000	40	16978	28.19	15.73	28.34
	100/600	6	21580	7.13	3.42	7.87
	100/1000	10	21580	12.04	6.70	14.16
	100/2000	20	21580	25.92	15.78	26.95
	100/4000	40	21580	55.67	34.91	56.69
	20/120	6	10759	1.38	0.55	1.26
	20/200	10	10837	2.34	1.05	2.23
	20/400	20	10866	5.30	2.53	4.34
	20/800	40	10884	11.54	5.75	9.09
	50/300	6	16778	3.51	1.57	3.12
Lecture room	50/500	10	16835	5.97	3.05	5.59
16492 triangles	50/1000	20	16848	13.03	7.45	10.85
	50/2000	40	16854	28.11	16.57	22.78
	100/600	6	22689	7.05	3.50	6.24
	100/1000	10	22731	11.97	6.82	11.19
	100/2000	20	22738	26.02	16.57	21.80
	100/4000	40	22743	56.41	38.71	45.65

Table 4.1: Computation times for several scenes.

4.5.2 Comparative visualization of wave-based and geometric acoustics

In this section the results of the phonon tracing algorithm are compared with the results calculated by the FEM based solver described in section 3.1. The comparison is done numerically and visually using the sound strength as acoustic quality metric (see section 2.3) and by visualization of the interference patterns. In the following a test scenario is discussed.

4.5.2.1 Test scenario

<table>
<tr><td>(a)</td><td>(b)</td></tr>
</table>
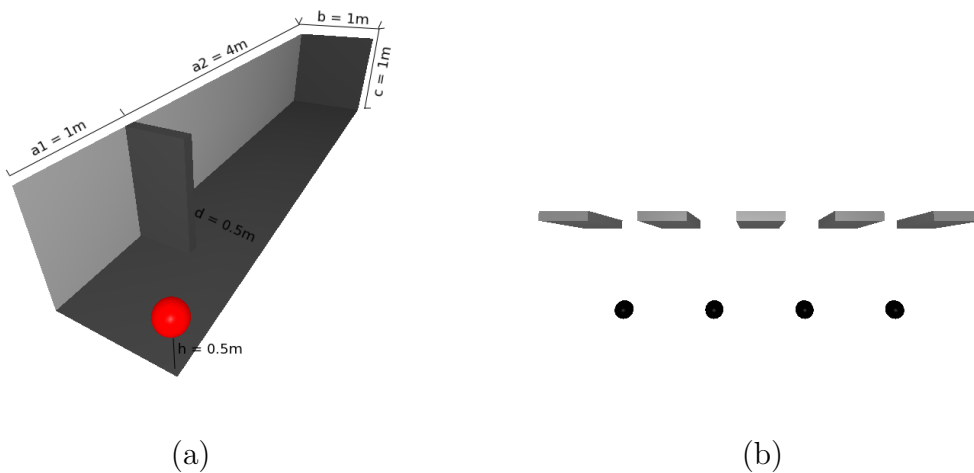
Figure 4.15: The geometry of the test room (a) and the simulation scenario (b).

In order to compare both simulation approaches the following test case is considered, designed to produce interference patterns. The FEM and phonon tracing are applied to a room of 5 meters length, 1 meter width, and 1 meter height with a gap of 0.5 meters (see figure 4.15 (a)). The sound source is located in the corner of the room. The two long walls are perfect sound reflectors, whereas the remaining room surfaces are perfect sound absorbers. Choosing the absorption characteristics of the room in this way an array of sound sources as outlined in figure 4.15 (b) is modeled. Now we observe interference effects as well as diffraction effects behind the gap.

In phonon tracing the pressure inside the room is simulated on a regular grid (561 points in total). An amount of 100000 phonons is traced from the sound source in order to calculate the room impulse responses f_i at the grid points g_i. Figure 4.16 depicts the wave propagation from the sound source. Afterwards, in order to determine the pressure at a grid point g_i for a given frequency ω a sine signal of frequency ω is convolved with f_i and the pressure at g_i is obtained at the position
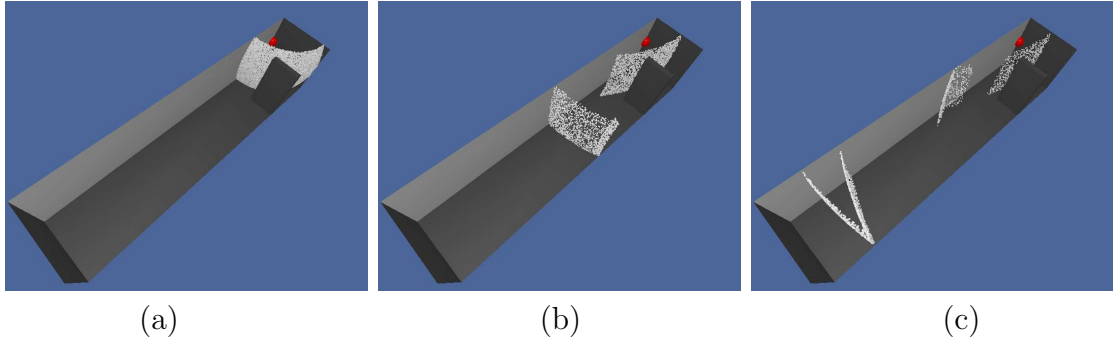
Figure 4.16: Visualization of particle (phonons) propagation from the sound source in three consecutive time steps.

$le_{f_i} + 1$ of the resulting signal. le_{f_i} is the length of the room impulse response at i-th grid point.

Comparing the results of phonon tracing and FEM, stationary sound fields originating from a source emitting only one particular frequency are considered. The corresponding FEM solution is computed in terms of the reduced continuous time system, described in section 3.1 (see also [DBM+06]) :

$$y(s) = \tilde{C} \left(s\tilde{E} - \tilde{A} \right)^{-1} \tilde{B} \qquad (4.31)$$

where $s = 2\pi i f$. The mean squared pressure at the i-th grid position reads:

$$p_i^2 = \frac{1}{T} \int_0^T p_i(t)^2 \mathrm{dt} = \frac{1}{2} |\mathrm{y_i}|^2 \qquad (4.32)$$

For visualization the results are interpolated to the same grid that is used for phonon tracing. The simulation results are compared for different wave numbers $k = \frac{2\pi}{\lambda} \in \{3, 6, 9, 12, 15\}$ where λ denotes the wave length. We consider wave lengths in the order of the gap width. Note that when choosing the shifts in the Arnoldi algorithm to match the above wave numbers, i.e. $s_j = ik_jc_0$, the reduced model is exact at these wave numbers and results coincide with those of the full FEM simulation.

4.5.2.2 Interference pattern visualization

In the first step the pressure distributions of both methods are compared in order to validate whether similar interference patterns appear. For visualization purposes a quad mesh is used, which is color coded in the following way. Positive pressure values are mapped to red color and negative pressure values to blue color, the saturation of the color is reduced, depending on the absolute pressure value at the considered

position. All pressure values with an absolute value less than the hearing threshold level $(2 \cdot 10^{-5}$ Pa) are mapped to gray color. The resolution of the displayed mesh is greater than that of the simulation mesh, the pressure values of the additional points are then bi-linearly interpolated. Figure 4.17 shows examples of the interference pattern visualization for wave number $k = 6$ and $k = 12$. The results show that both methods faithfully reproduce matching interference patterns, where the results obtained by FEM appear to be somewhat smoother. The patterns at $k = 12$ are closer to each other than that at $k = 6$.

(a)

(b)

(c)

(d)
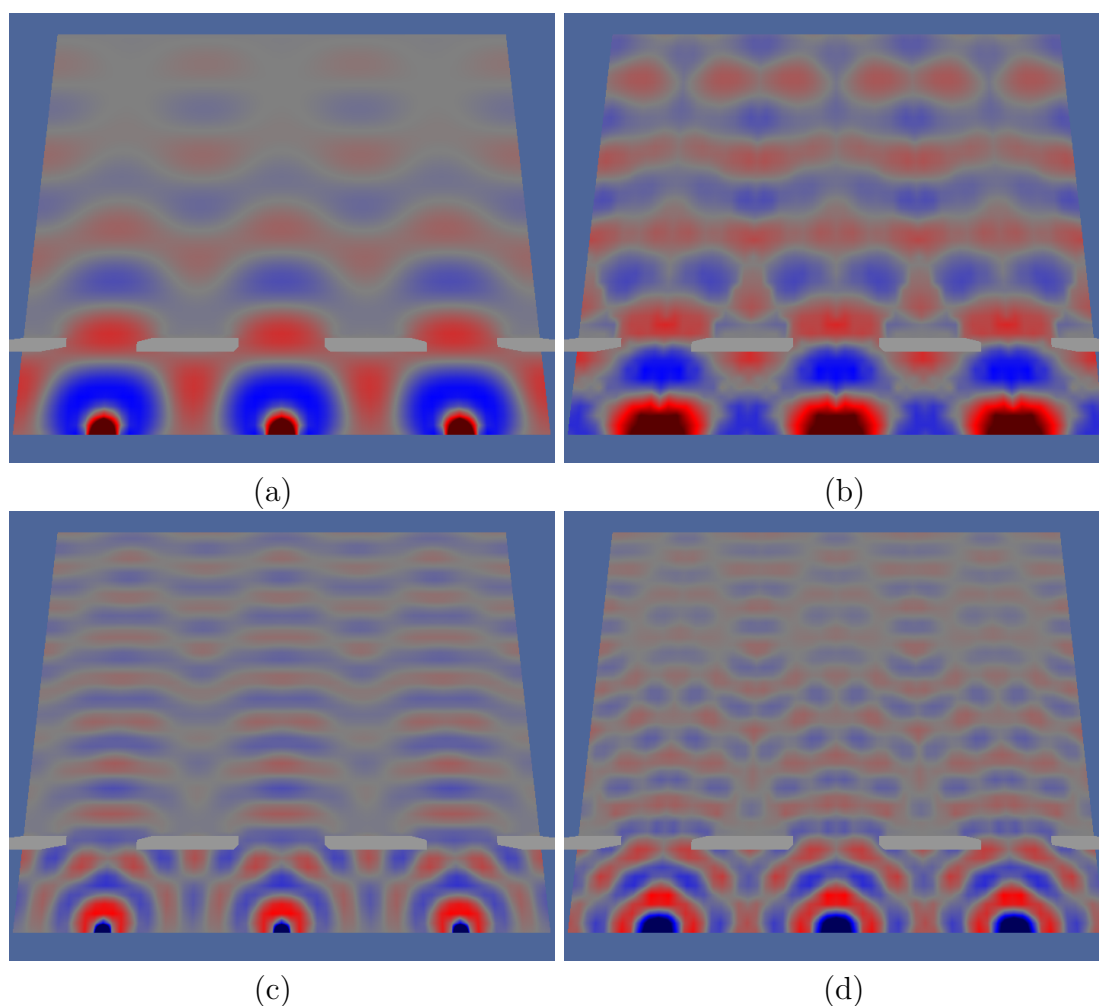
Figure 4.17: Interference pattern. FEM simulation (a+c) and phonon tracing (b+d) for the wave numbers k=6 (a+b) and k=12 (c+d).

Wave propagation is illustrated more intuitively, when pressures are mapped along the normal to the listening plane as depicted in figure 4.18.
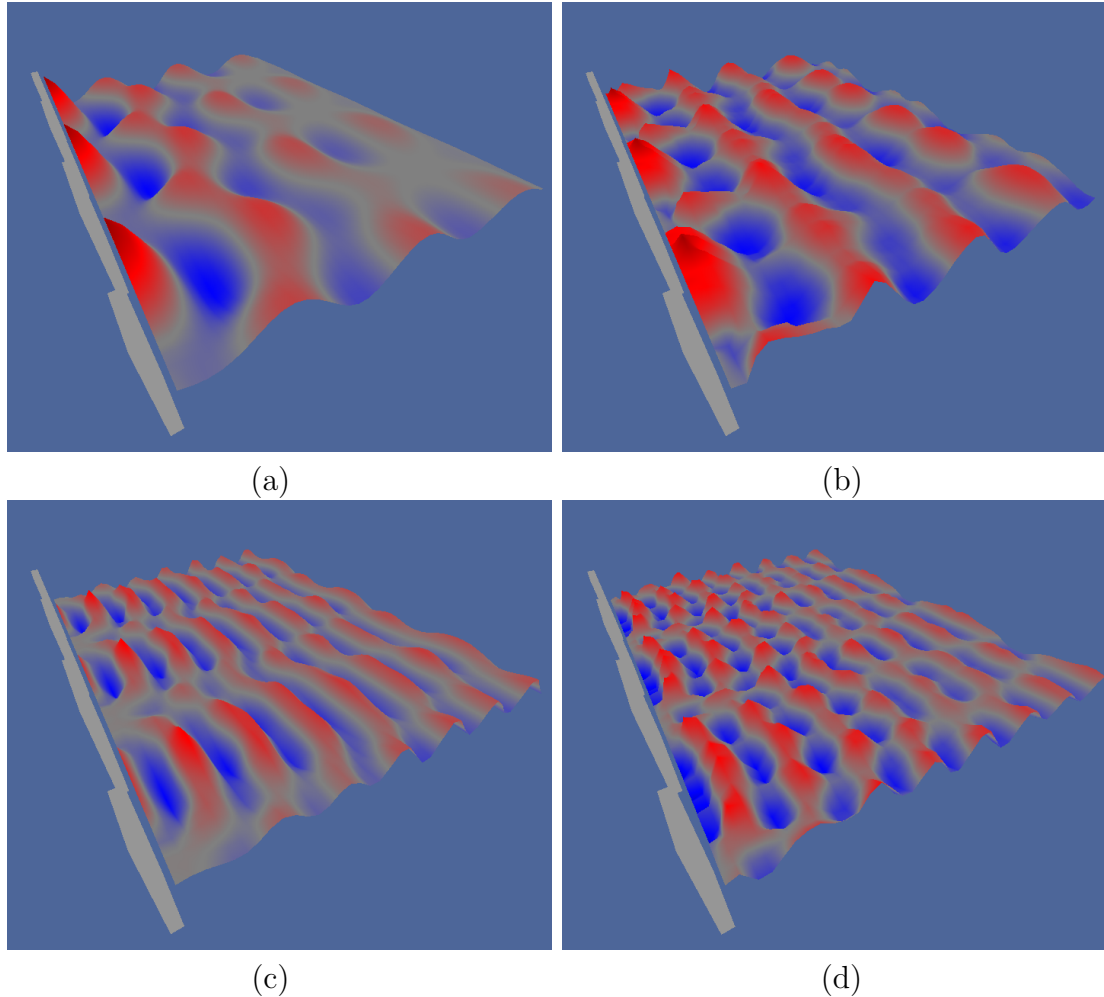
Figure 4.18: Visualization of wave propagation. FEM simulation (a+c) phonon tracing (b+d) for the wave numbers k=6 (a+b) and k=12 (c+d).

4.5.2.3 Sound strength

For a more detailed comparison we turn to an acoustic metric, called sound strength, which is essentially the logarithm of the mean squared pressure [ISO97], and for discrete values it can be calculated as follows:

$$G = 10 log_{10} \left(\frac{\sum_{i=0}^{m} p_i^2}{\sum_{i=0}^{m} p_{10_i}^2} \right) \quad dB \qquad (4.33)$$

Since the sources have been modelled differently (a point source in phonon tracing and a small membrane in FEM), the sound fields are normalized by a linear fit excluding the direct neighborhood of the sources. For visualization of the sound strength and the error a quad mesh is used color coded from red (maximum value) to

blue (minimum value). Therefore we interpolate the hue value of the HSV color space according to the sound strength and error value, respectively. Values for additional mesh points for rendering are bi-linearly interpolated as mentioned before. To ensure better comparison feasibility of the sound strength the same color range is used for G resulting from the two simulation algorithms at same frequencies. Moreover, the same color range is used for all error plots. Since the sound strength is a logarithmic measure the relative error of the pressures is proportional to the difference of the sound strength values and can be calculated as:

$$\varepsilon = \delta \frac{ln(10)}{20} \quad with \quad \delta = \|G_{ph_j} - G_{fem_j}\| \tag{4.34}$$

where G_{ph_j} and G_{fem_j} is the sound strength at the grid point j calculated using results from the phonon tracing algorithm and the FEM method, respectively. The results are summarized in table 4.2. We can observe a decrease of the error until $k = 12$, since the diffraction effects neglected by the phonon tracing approach become smaller. Due to the shortcomings of the FEM method at higher frequencies the error amount increases. Considering the results we state that both simulation methods matched best at wave number $k = 12$. In figure 4.19 (c), 4.19 (f), and 4.19(i) we observe that the blue regions on the error plots are predominant indicating overall error decrease. The sound strength plots of FEM and phonon tracing results at $k = 12$ (figure 4.19 (h+i)) are closer to eachother as those at $k = 3, 6$ (figure 4.19 (a+b), 4.19 (d+e)). The FEM approach is the mathematically correct method in the frequency spectrum that can be represented by the grid (due to the Nyquist limit). In order to keep the calculation cost of the FEM appropriate, the wave length reaches the order of the grid size and FEM becomes inaccurate. The phonon tracing is the more efficient method and provide at $k = 12$ similar results as the FEM, thus we can use it for the simulation of the acoustics at the frequencies above $k = 12$, which corresponds to a wave length of half the gap width.

k	3	6	9	12	15
δ	8.216	5.164	3.989	3.784	4.903
ε	0.946	0.595	0.459	0.436	0.565

Table 4.2: Absolute and relative error in dB between FEM and phonon tracing.
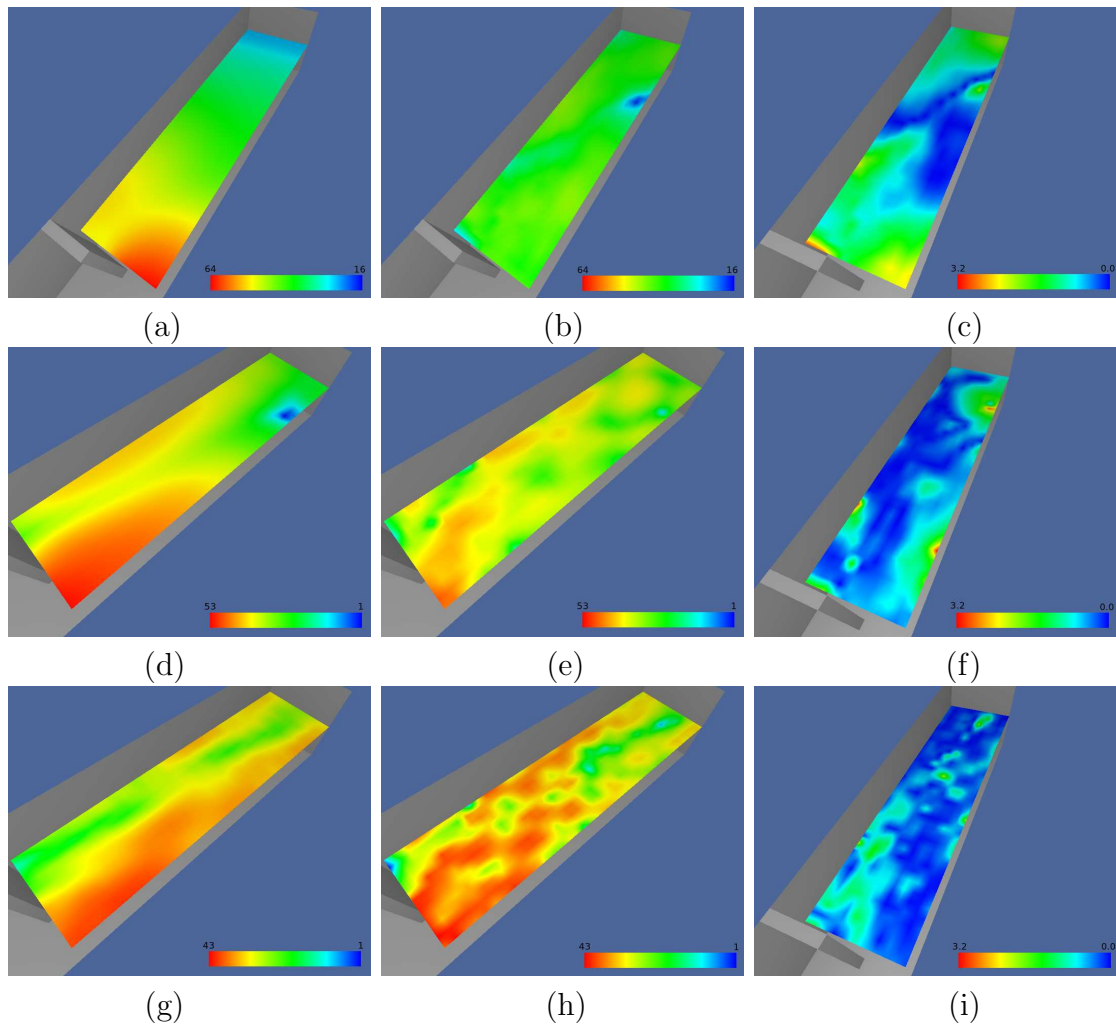
Figure 4.19: Sound strength visualization (values in dB). (a) FEM simulation, (b) phonon tracing, and (c) the relative error for wave number k=3. (d) FEM simulation, (e) phonon tracing, and (f) the relative error for wave number k=6. (h) FEM simulation, (h) phonon tracing, and (i) the relative error for wave number k=12.

4.5.3 Acoustic round robin

In the previous section the results of the phonon tracing algorithm were compared numerically and visually with the results of a FEM method. In the following the phonon tracing algorithm is evaluated using the conditions given in the 3rd round robin test [Bor02, Bor05a, Bor05b]. The geometry data, the assumed conditions as well as the results of other simulation algorithms are publicly available [1]. This way the phonon tracing can be compared with other sound prediction software.

| (a) | (b) |
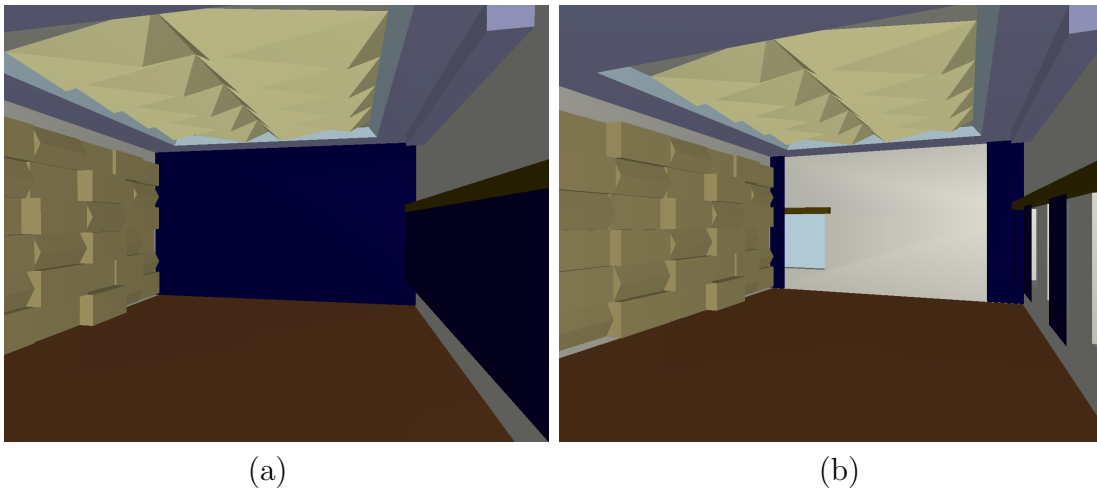
Figure 4.20:   Room   geometry   used   by   the   3rd   round   robin   at   3rd   phase (8.6×10.2×5m). (a) Curtains closed, (b) curtains opened.

The 3rd round robin on acoustical simulation software was performed in three phases with different levels of detail for the room geometry. Eight acoustic parameters, definition (D_{50}), clarity (C_{80}), sound strength (G), early decay time (EDT), center time (t_s), reverberation time (T_{30}), and lateral energy fractions (LF and LFC), were compared. In this work the geometry given in the 3rd phase is used since it has the highest geometry details and thus represents best the real room. The room geometry is depicted in the figure 4.20. Unfortunately, in the current version of the phonon tracing algorithm not all of the eight parameters can be predicted, therefore only D_{50}, C_{80}, G, EDT, and t_s are used. The absorption and scattering parameters for the room materials are given in the tables 4.3 and 4.4, respectively. The parameters are predicted using the energy responses of the phonon tracing algorithm in order to keep it comparable to other methods, which are mostly energetic. The responses are simulated for two source and three listener positions (see table 4.5) and for six octave bands (125Hz, 250Hz, 500Hz, 1kHz, 2kHz, 4kHz) with closed (figure 4.20 (a)) and open (figure 4.20 (b)) curtains.

---

[1]http://www.ptb.de/de/org/1/17/172/_raumakustik.htm

Material	Frequency in Hz					
	125	250	500	1000	2000	4000
parquet	0.04	0.04	0.07	0.06	0.06	0.07
wood	0.20	0.05	0.03	0.01	0.01	0.01
curtain(closed)	0.15	0.30	0.35	0.40	0.50	0.55
curtain(open)	0.30	0.45	0.60	0.70	0.70	0.70
wilhelmi	0.42	0.28	0.49	0.78	0.58	0.62
studio wall	0.02	0.02	0.02	0.02	0.02	0.02
windowglas	0.10	0.04	0.03	0.02	0.02	0.02
wood absorber	0.40	0.33	0.21	0.16	0.15	0.16
ceiling	0.30	0.20	0.06	0.02	0.02	0.02

Table 4.3: Absorption coefficients for the room used in the round robin.

Material	Frequency in Hz					
	125	250	500	1000	2000	4000
parquet	0.20	0.20	0.20	0.20	0.20	0.20
wood	0.20	0.20	0.20	0.20	0.20	0.20
curtain(closed)	0.21	0.26	0.32	0.39	0.47	0.56
curtain(open)	0.22	0.31	0.39	0.48	0.59	0.73
wilhelmi	0.20	0.20	0.25	0.30	0.35	0.40
studio wall	0.20	0.20	0.20	0.20	0.20	0.20
windowglas	0.10	0.10	0.10	0.10	0.10	0.10
wood absorber	0.20	0.20	0.20	0.20	0.20	0.20
ceiling	0.20	0.20	0.20	0.20	0.20	0.20

Table 4.4: Scattering coefficients for the room used in the round robin.

Sources	x	y	z
S1	1.5	3.5	1.5
S2	-1.5	5.5	1.5
Receivers	x	y	z
R1	-2.0	3.0	1.2
R2	2.0	6.0	1.2
R3	0.0	7.5	1.2

Table 4.5: Sources and listeners positions.

The simulation results are compared to the given measurements. Therefore, the mean error for the six frequency bands to emphasize the positional dependencies and the mean error of the six source listener combinations related to the reference values given in table 4.6 are determined for each of the five acoustic parameters.

Acoustic parameter	Symbol	JND
Definition	D50/%	5 %
Clarity	C80/dB	1 dB
Sound strength	G/dB	1 dB
Early decay time	EDT/s	50 msec
Center time	$t_s/msec$	10 msec

Table 4.6: Reference values (just noticeable difference JND) of used acoustic parameters for the relative evaluation.

The relative errors are depicted in the following figures for closed (figures 4.21 and 4.22) and open (figures 4.23 and 4.24) curtains. As we can see in the figures 4.21 and 4.23 the errors for all five parameters are acceptable for both closed and opened curtains. For the higher frequencies where the diffraction effects are smaller the error becomes also smaller compared to the lower frequencies. In figures 4.22 and 4.24 the results of the phonon tracing algorithm are compared to results obtained by other acoustic prediction methods with closed and opened curtains, respectively. Red dashed lines indicate the maximum error value of the according parameter, the green dashed lines the minimum error value, and the blue dashed line the average error of all participants of the 3rd round robin. The black solid lines are the relative errors of the phonon tracing algorithm. The error of the phonon tracing is mostly in the range of the minimum error and in most cases well below the average error. The maximum error is never reached. Thus, the phonon tracing algorithm compares favorably to existing methods predicting acoustical parameters of a room. The error of the early decay time (EDT) prediction is somewhat greater compared to other parameters, thus the algorithm is less suitable for prediction of the EDT.

Despite the fact that the algorithm does not model diffraction effects, since it is a pure geometric approach without adding some statistical reverberation tail, and considering the inaccuracy brought by the possible errors in the computer model of the geometry, we conclude that we have obtained reasonable results with the phonon tracing algorithm.
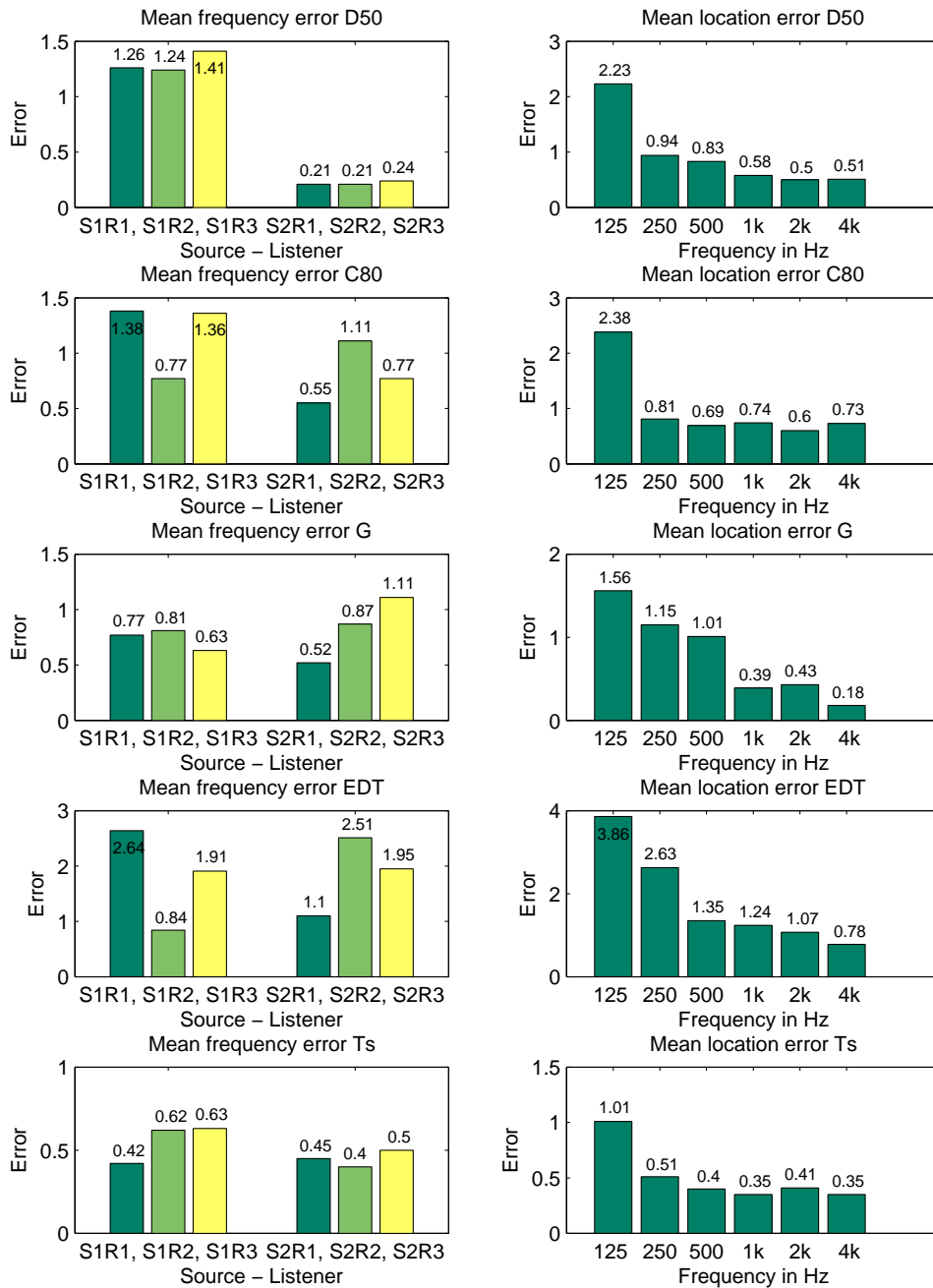
Figure 4.21: Relative error of the acoustic parameter prediction using phonon tracing algorithm at closed curtains.
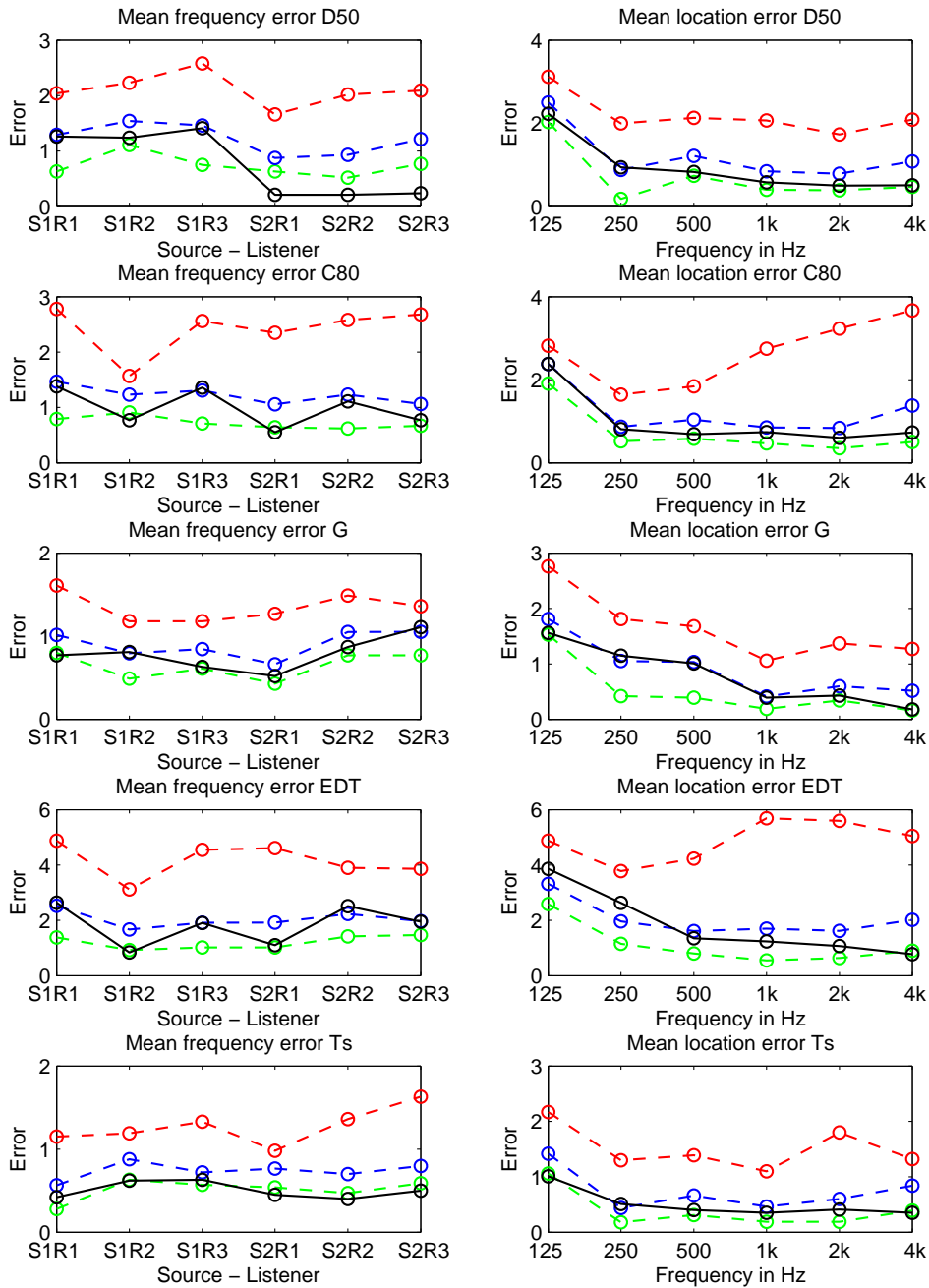
Figure 4.22: Relative error of the acoustic parameter prediction using phonon trac-
ing algorithm compared to other acoustic simulation methods at closed
curtains. Red dashed line indicates the maximum error value, the green
dashed line the minimum error and the blue dashed line the average er-
ror value of all participants of the 3rd round robin. The black line is
the error of the phonon tracing algorithm.
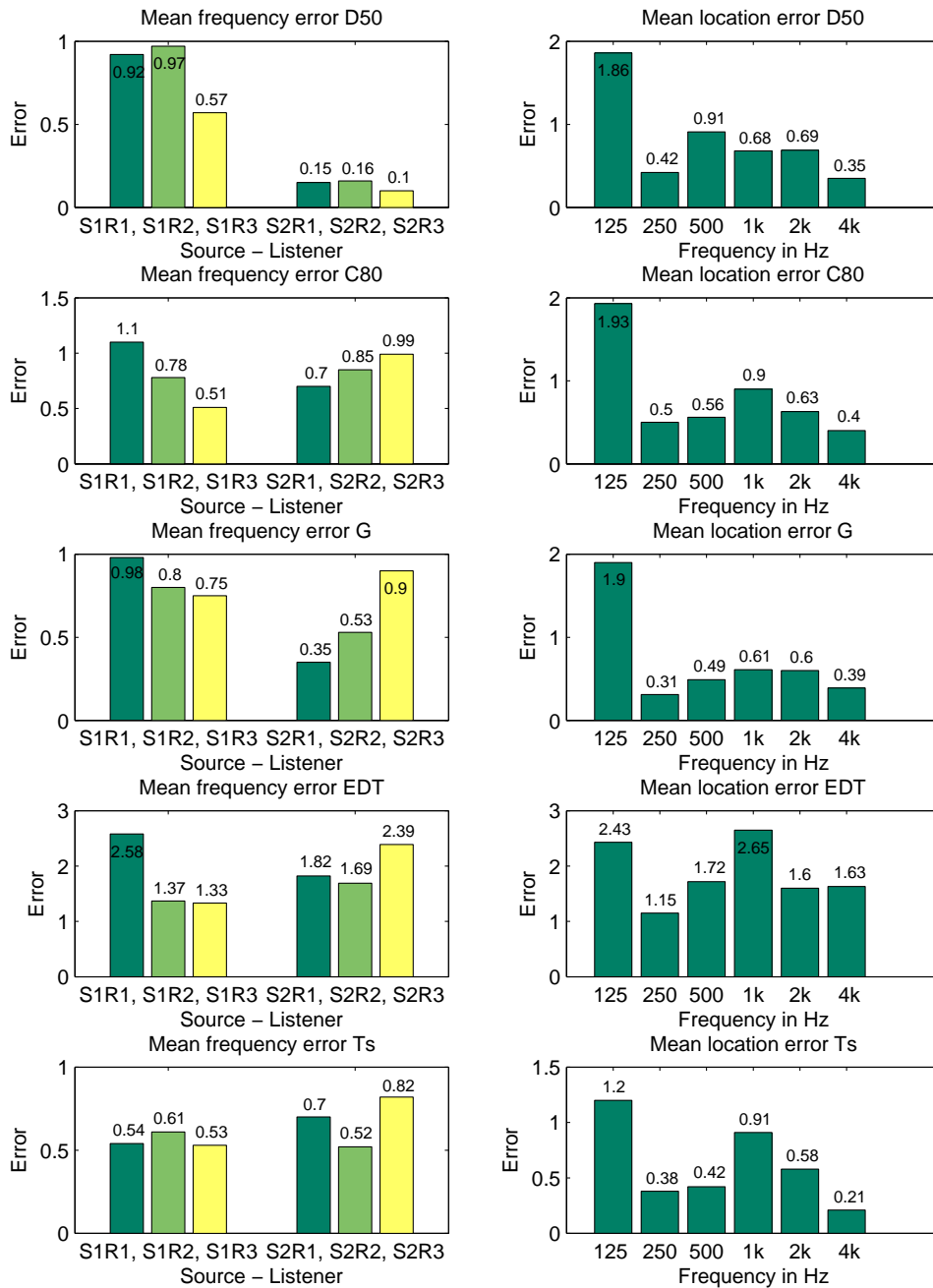
Figure 4.23: Relative error of the acoustic parameter prediction using phonon tracing algorithm at opened curtains.
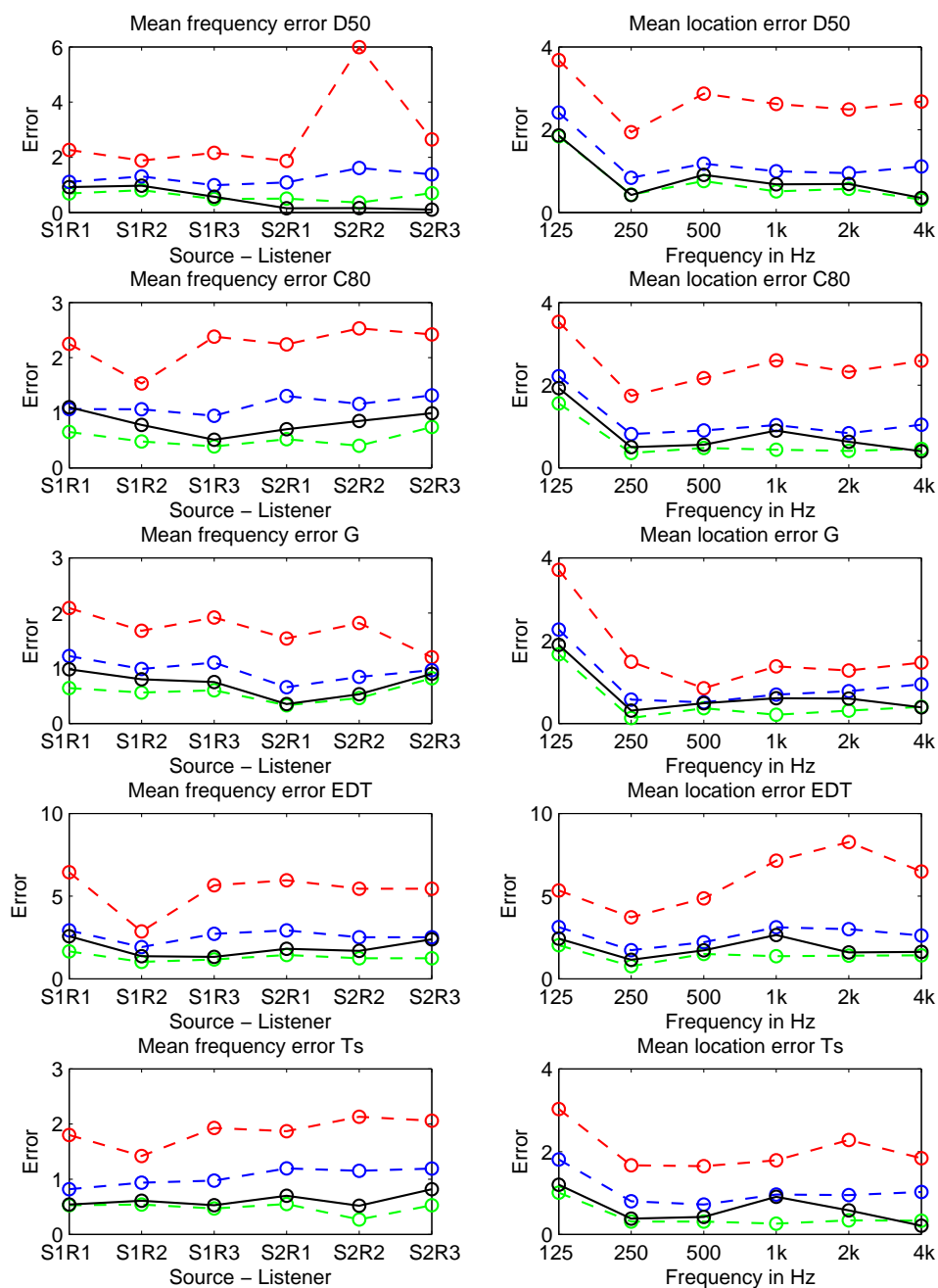
Figure 4.24: Relative error of the acoustic parameter prediction using phonon tracing algorithm compared to other acoustic simulation methods at opened curtains. Red dashed line indicates the maximum error values, the green dashed line the minimum error and the blue dashed line the average error values of all participants of the 3rd round robin. The black line is the error of the phonon tracing algorithm.

5 Sound visualization

The visualization of the sound wave behavior inside the room is important, not only for teaching purposes, but also for engineers designing and modifying rooms. Visual representation of sound propagation as well as representation of sound received at a listener position can help understanding these complex processes.

In this chapter we introduce various approaches for visualization of room acoustics utilizing the phonon map. First, a visual representation of the propagation of the sound wave from the source through the room such that the effect of different materials on the spectral energy (pressure) distribution can be observed. The first few reflections already show whether certain frequency bands are rapidly absorbed. The absorbing materials can be identified and replaced in the virtual model, improving the overall acoustic quality of the simulated room. After a great number of reflections, individual wave fronts cannot be identified, anymore. The color composition of the phonons, however, express the overall acoustic quality. Furthermore, we introduce several approaches including the visualization of individual phonons on surfaces and reflected sound wave fronts providing insight how the room influences the propagated sound waves. In addition, we present an approach to visualize the direction and amount of the sound received by the listener. Beyond, we propose a visual analysis of an acoustic parameter at given listener positions using the phonon tracing providing insight which surface material needs to be changed in order to enhance the value of this metric. We proceed to discuss several related methods.

5.1 Related work

The existing visualization techniques can be classified into two groups. The methods of the first group consider the propagation of the sound waves or rays inside the room independent of the listener position. Whereas the algorithms in the second group base on the measured or simulated room impulse response for individual listener.

Yokota et. al. [YST02] visualized sound propagation of 2-dimensional room sound fields resulting from a wave-based simulation using the difference time domain method. They consider different simple shaped rooms which are typical for concert halls. The effects of changing reflection characteristics of the walls as well as the effect of diverse suspended panel arrays were investigated. However, the

reflection coefficients are chosen equal for all rooms and frequencies.

Petrausch and Rabenstein [PR05] introduce a program for simulation and OpenGL based visualization of 2-dimensional sound wave propagation in real time. The simulation is performed by use of functional transformation method. According to the authors the program supports frequency dependent and independent damping, dispersion, multiple excitations, spatially distributed excitations, as well as flexible boundary conditions.

Tokita and Yamasaki [TY05] visualized particle displacements on a rigid 3-D grid in a cube shaped room. The displacements resulted from a wave-based numerical simulation. The visual results are presented at several time steps using VRML97[1].

Lokki [Lok02] presented in his PhD thesis a visualization using the image-source method (see section 3.2.1). The image-sources and the edge-sources (sources of the corners of room objects) of different order are depicted as boxes. Furthermore, the impulse response for each pixel is calculated and plotted for a time moment resulting in a representation of wave fronts emitting from each of those virtual sources.

The image source method has been extended in order to simulate edge diffraction (see section 3.2.5). In [PL03] Pulkki and Lokki present an approach visualizing this phenomenon. First order edge image-sources are displayed as cylinders. The radii and color of these cylinders varies depending on the pressure magnitude. The cylinders are connected to a degenerate curve represented first-order edge diffractions. Each of the first-order diffraction sources itself produce a curve diffraction on another edge. For all first-order edge-sources diffracts on a given edge the result is a large number of image-sources arranged on a planar surface. The authors rendered this surfaces color coded depending on the pressure magnitude. A blue-white-red color scheme is used to depict negative and positive polarity of the wave. The authors restricted the visualization to first- and second-order diffraction.

Funkhouser et. al. [FCE+98] used visualization of source points, receiver points, pyramidal beams, reverberation path etc. in order to understand and evaluate their acoustic modeling method. Lauterbach et. al. [LCM07b] showed the sound propagation resulting from frustum tracing. Khoury et. al. [KFW98] represented the sound pressure levels inside the room by means of color maps. Additionally, the authors analyzed the precedence effect (or "law of the first wave front") by using isosurfaces. Lokki and Nenonen [LN06] utilized cave equipment for immersive visualization of trace paths and particle paths propagating inside closed rooms.

Sound wave fronts of different frequency arrive at the listener position from different directions at different moments. Thus, the room impulse response is multidimensional. Merimaa et. al. [MLPK01] presented a visualization of measured impulse response using 2-D plots in horizontal an median plane. The directional

---

[1]http://www.web3d.org/vrml/vrml.htm

intensity is depicted with vectors on top of sound pressure related spectrogram.

Omoto and Uchida [OU04] visualized the arriving direction and strength of measured sound intensity using circles. The circles are mapped onto the surface of a rectangular wire-frame box approximating the room shape. This way the arriving direction is represented. The radii of the circles vary with the intensity. An extension of this approach using photographs of the enclosure instead of the wire-frame box is described in [FSO06].

Weyna [Wey05b, Wey05a] used flow visualization methods in order to study acoustic flow fields in front of and behind a thick rectangular plate. Measured intensity is visualized using hedgehogs and color coded streamlines. Furthermore the intensity amount is depicted by means of isosurfaces.

As discussed in section 2.3 the impulse response of the room characterizes the acoustic quality of the room. Several objective metrics are derived depending on the purpose of the room, the use for speech or musical representations. Visualizations addressing the acoustical behavior of the room also exist. Stettner et. al. [SG89] visualized clarity and definition as well as spatial impression by use of specific icons. Furthermore, the authors represent the overall strength of sound by displaying global pressure levels as color maps at the room boundaries at different time moments.

Monks et al. [MOD00] introduced an interactive optimization system for acoustic design. The system solves the inverse problem. After the user provided constraints to the surface properties and geometric variations for a range of architectural components, as well as desired acoustic performance defined by acoustic metrics, the system searches the configuration space in order to satisfy the target configuration. Simulating annealing followed by steepest descent algorithm is used for this purpose. The authors visualized the results by means of icons depicted early decay time (EDT), interaural cross-correlation coefficient (IACC), and bass ratio (BS). Furthermore, the sound strength was visualized at room surfaces using color to indicate the sound-level data at different time moments. The authors shown not only the simulated or optimized values of acoustic quantities but also the difference values between the simulation, the optimization, and specified target. Two real concert halls were investigated applying this system.

A couple of commercial systems[2] provide several tools for visualizing measured or computed acoustic quantities.

---

[2]www.odeon.dk, www.catt.se, www.bose.com

5.2 Visualization using the phonon map

The phonon map, as a result of the phonon tracing algorithm described in chapter 4, characterizes the acoustic behaviour of a scene considering the location of a specific sound source. It consists of the reverberations of a unit pulse, coming from different directions with different time delays and specific energy and pressure distributions. The amount of particles stored in the phonon map can be very high. A simple representation of the phonons at their positions in the scene depicts neither the sound propagation from the source nor the sound energy received at a listener and is often too dense to be informative as can be seen in figure 5.1. Also, the particle paths can get very complicated in complex virtual environments, and the depiction of even few paths (as line segments) is cluttered and the different paths can not be recognized clearly (see figure 5.2).

In the following sections we introduce our approaches visualizing sound propagation from the source including reflections at scene boundaries as well as sound received at different listener positions inside the scene using the information given in the phonon map.
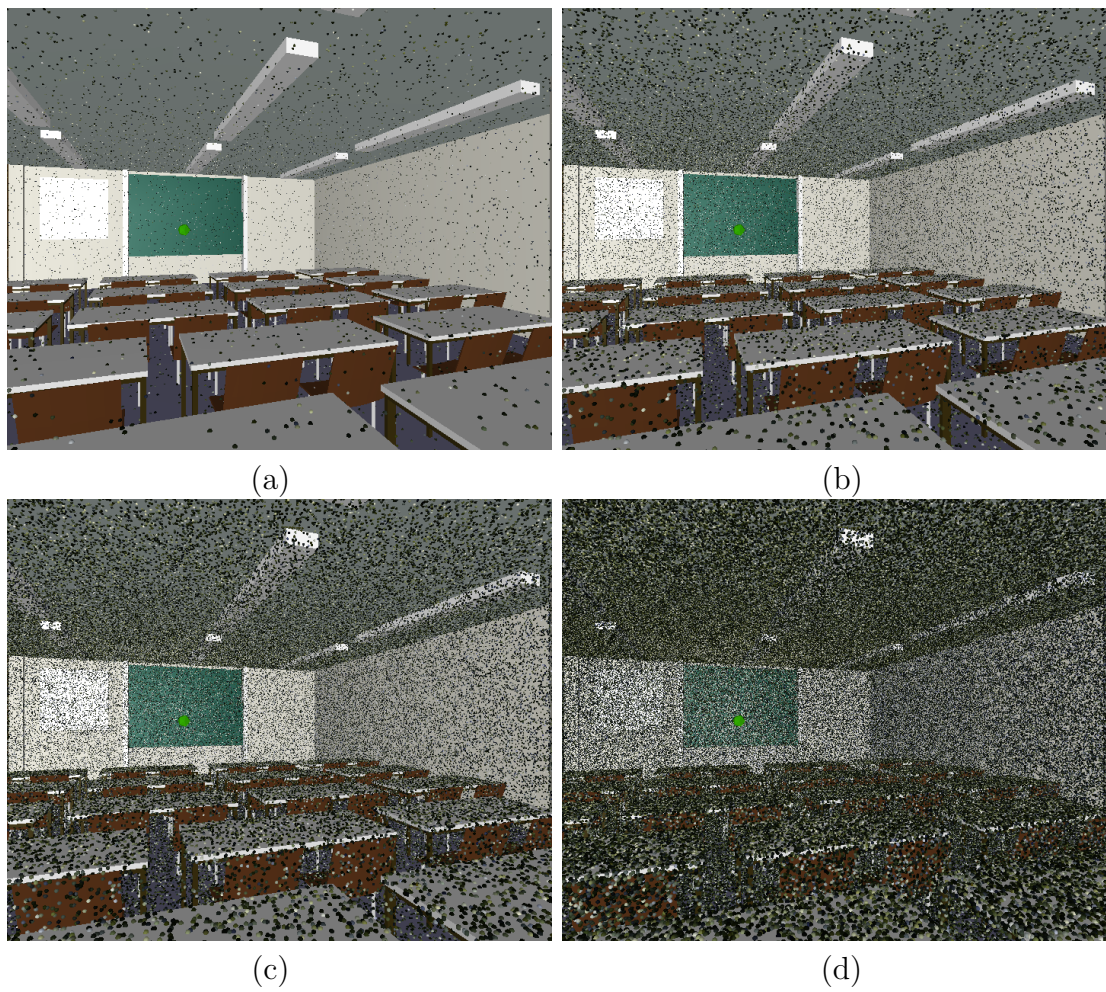
Figure 5.1: Phonon map representation with increasing number of phonons in the map from (a) to (d).
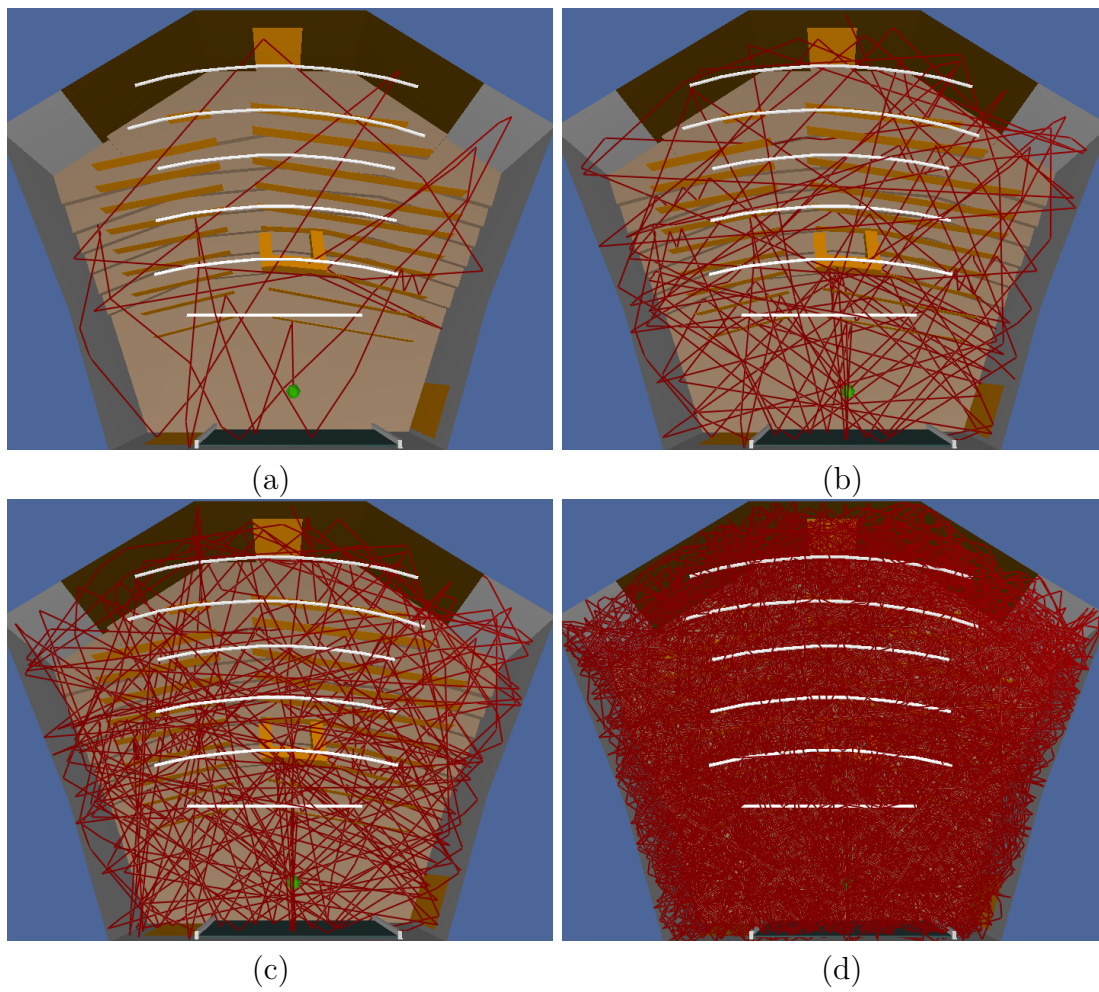
Figure 5.2: Particle paths representing by line segments. One (a) path, five (b), ten (c) and hundred (d) paths.

5.2.1 Wave front propagation

The first visualization method focuses on the spatial propagation of sound waves from the sound source inside the environment. The corresponding wave front traverses the room and is reflected at surfaces, altering its intensity and energy spectrum. In the following several alternatives of the wavefront representation are described.

5.2.1.1 Wave fronts represented using single particles

first reflections

multiple wave-fronts

spectral distribution
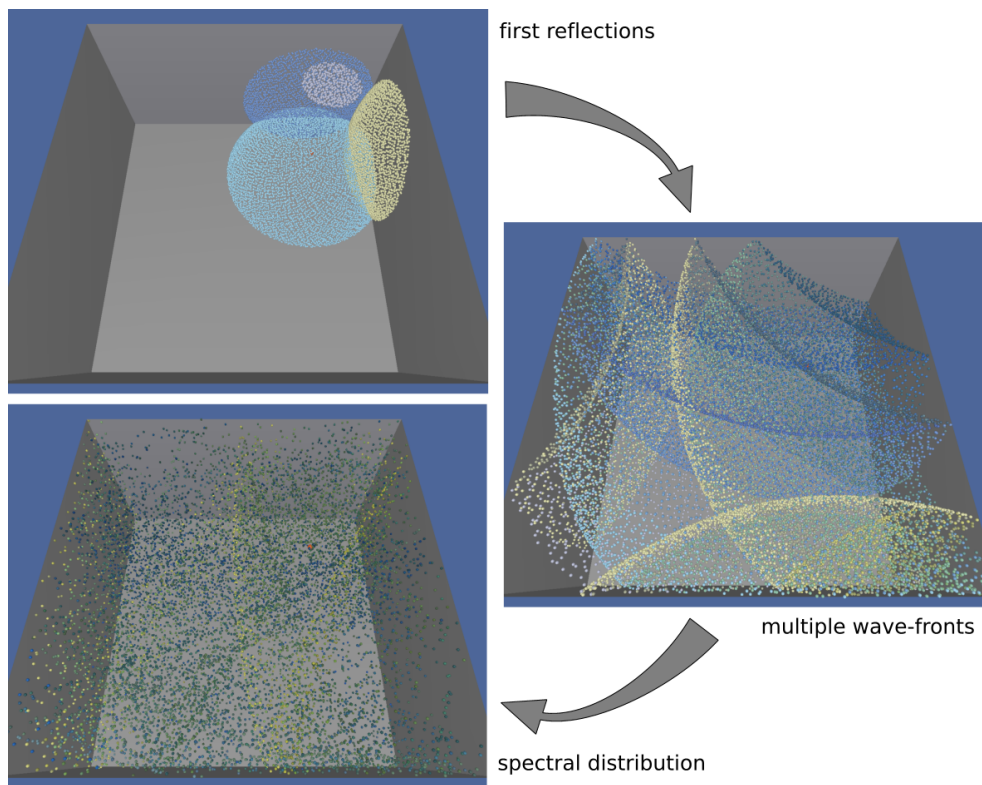
Figure 5.3: Wave front propagation from a spherical sound source at three progressive time steps. The visualization shows phonons with color-coded spectral energy due to reflection at different materials.

The first alternative is the depiction of the sound waves by rendering small spheres representing the phonons. These are color coded by means of their spectral energy or pressure. Therefore, we use the RGB components, such that blue corresponds to the average of the energy (pressure) in the frequency bands around 40, 80, 160, 320 Hz, green corresponds to the average by 640, 1280, 2560 Hz, and red to the average by 5120, 10240, 20480 Hz. When sliding through time, the spheres follow

the simulated phonon paths as depicted in figure 5.3. The following functionalities are integrated in the interactive visualization system:

- varying the percentage of phonons to be rendered

- rendering only the phonons reflected from a selected material (see figure 5.5)

- exchanging selected materials (see section 5.2.5)

- varying time / transversed distance (see figure 5.3)

Since the path information is stored in the map, and all phonons sharing the same path are known, it is simple, for example, to select all consecutive pairs ph_i, ph_{i+1} where the current time t satisfies $t_{ph_i} < t < t_{ph_{i+1}}$ and to draw a sphere on the line segment pt_{ph_i}, $pt_{ph_{i+1}}$, corresponding to a phonon location at time t. In addition, certain predicates can be used, for example selecting all phonon whose path was reflected at the first (or at any) reflection off a selected material.

The exchange of a certain material requires only the phonons' energy to be re-evaluated, where the phonon paths remain fixed. To allow the exchange, it is necessary to enforce a minimum number of reflections for every path in advance, since otherwise materials with high absorption coefficients cannot be replaced. An application scenario of this kind is provided in section 5.2.5.

5.2.1.2 Wave fronts represented using triangulated surfaces

To improve the visualization of wave fronts, in the second alternative polygonal surfaces with phonons as vertices are used. The phonons are emitted from the sound source S according to the emission distribution E over a unit sphere. We construct the surface of the wave front coming from S as a convex hull of phonons on the unit sphere around S and obtain the neighborhood relationship of these phonons. This relation does not change in time, so this initial polygonal representation of the wave front must be calculated only once. For the construction of the convex hull on the unit sphere providing a Delaunay-triangulation the CGAL library[3] is employed. The paths of the phonons as a result of the simulation are known. The position and color coded spectral energy can be calculated as described above. In a given point of time t we draw only this faces, which include phonons relevant at time t.

When the sound wave hits an object inside the room, we need to subdivided the initial wave front, since parts of it travel in different directions and have different energy (pressure) spectra. The splitting of the wave surface is decided on following criteria:

- $c_{ph_i} = c_{ph_{i+1}}$ i.e. only the faces, which vertices (phonons) have the equal color coded spectral energy (or pressure) are drawn

---

[3]http://www.cgal.org

- $\langle v_{ph_i}, v_{ph_{i+1}} \rangle \geq 0$ i.e the moving direction of the vertices differs not more than about 90°

where c_{ph_i} is the color coded spectral energy of the phonon ph_i, and $\langle \cdot, \cdot \rangle$ denotes the scalar product. The second constraint is necessary in the case if two phonons of the same triangle face hit for example different walls with the same material.

Figure 5.4 shows the wave front propagating from the source. Where it hits the ceiling, (figure 5.4 top left), the faces that now belong to the wave front reflected at the ceiling are separated from the initial triangulated surface.
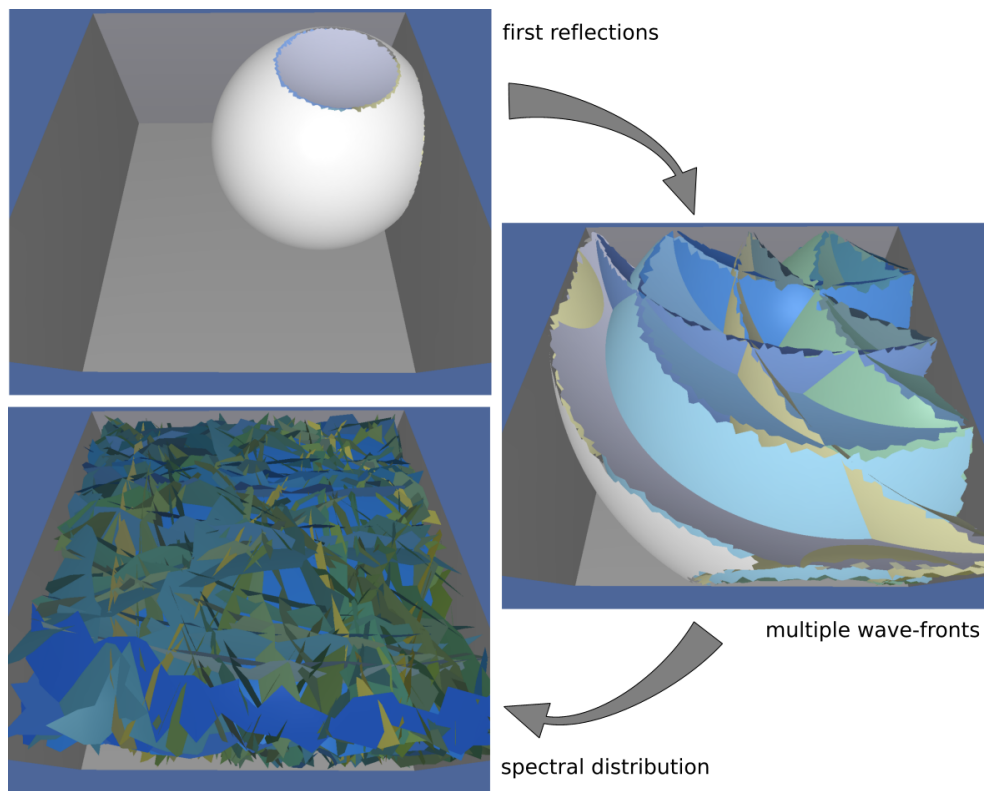
Figure 5.4: Visualizing wave front traversed from the sound source using triangulated surfaces.

An example of further visualizations of sound wave fronts reflected from different room surfaces having different materials using triangulated surfaces is shown in figure 5.5.
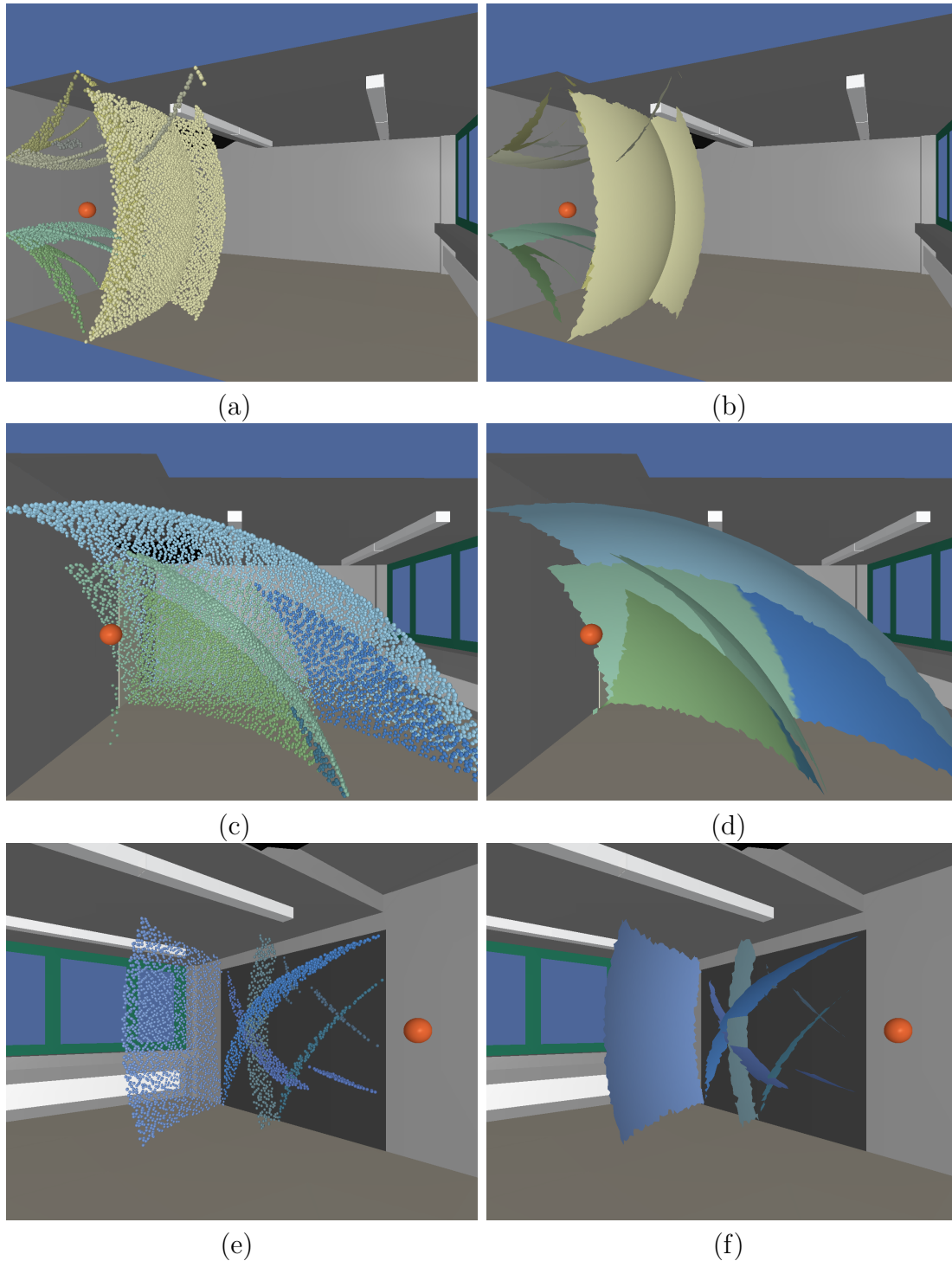
Figure 5.5: Selecting wave fronts by material reflections visualized using single
phonons as well as triangulated surfaces. Wave fronts reflected from the
walls (a+b), from the floor (c+d), and from the projection wall (e+f).
The red sphere represents the sound source.

5.2.1.3 Wave fronts represented using surface elements

The advantage of the wave front representation as triangulated surfaces in contrast to the spheres (representing single phonons) is the improved ability to distinguish between different reflected wave fronts, as shown in figure 5.5. At least, this can be observed for early reflected wave fronts. After multiple reflections the number of different wave fronts is too high to recognize each different reflection and the benefit over the sphere representation is small. The disadvantage of the triangulated surfaces visualization is the limited scalability. While we can vary the number of the phonons to be rendered in the single particle visualization, a new Delaunay triangulation for reduced number of triangles (resulting in reduced number of phonons considered) needs to be computed to allow scalability in the surface representation. A compromise between these visualizations is the usage of surface elements (surfels) [PZvBG00]. In this case disks representing a part of a sphere surface, since a spherical point source is considered during the simulation.
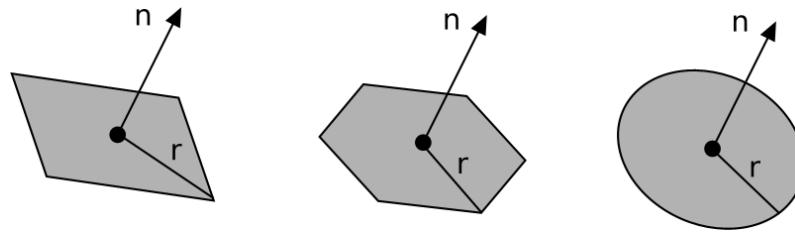
Figure 5.6: Sketch of surface elements (surfels) at different levels of detail.

In figure 5.6 a sketch of surfels is depicted. The surfel representation can be adjusted according to users needs on accuracy as well as graphics hardware capabilities. The level of detail can be changed by prescribing the number of points describing the circumcircle of the disk (from quad to circle in figure 5.6).

In order to depict the propagated sound wave front, a surfel is rendered on each particle position at given time. The size of the surfel is given by the radius r_i and the orientation by the normal n_i which corresponds the sound particle travelling direction. The radius r_i varies depending on the number n_{ph} of phonons to be rendered (given by the user) as well as on the traversed distance l of the sound wave (time step given by the user):

$$r_i = s \sqrt{\frac{l}{n_{ph}}} \tag{5.1}$$

The radius r_i can be also be expanded by the user (the scaling factor s in the above equation) in order to obtain a surface rendering of the sound wave front. A small radius r_i results in a visualization on the wave front similar to that using spheres,

whereas a large r_i leads to a triangulate surface-like presentation. An example of sound wave front visualization using surfels is depicted in figure 5.7. Like in the both previously described approaches an option to observe reflections only from one material is possible. An example is shown in figure 5.8.

We note that, the visualization approach described in this subsection can be improved by culling the surfels on the scene geometry or by use of other already published techniques [ZPvBG01, KB04] e.g. utilizing the GPU [BK03].

(a)   (b)

(c)   (d)

(e)   (f)
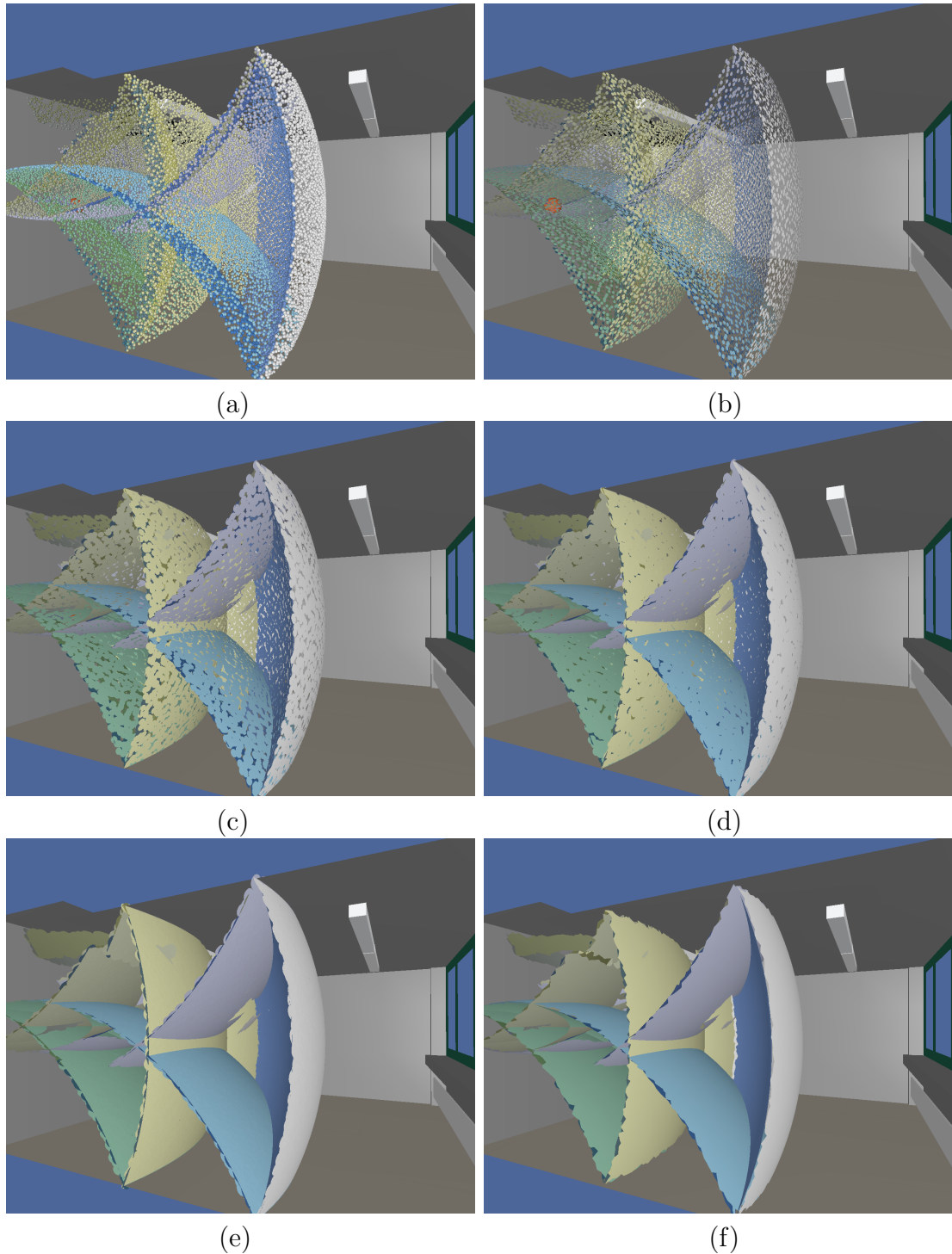
Figure 5.7: Wave front representation using surface elements (surfels) in contrast to spheres and triangulated surface visualization approaches. Spheres used for rendering of phonons (a). Surfels with increased radius from (b) to (e). Triangulated surface representing sound wave (f).
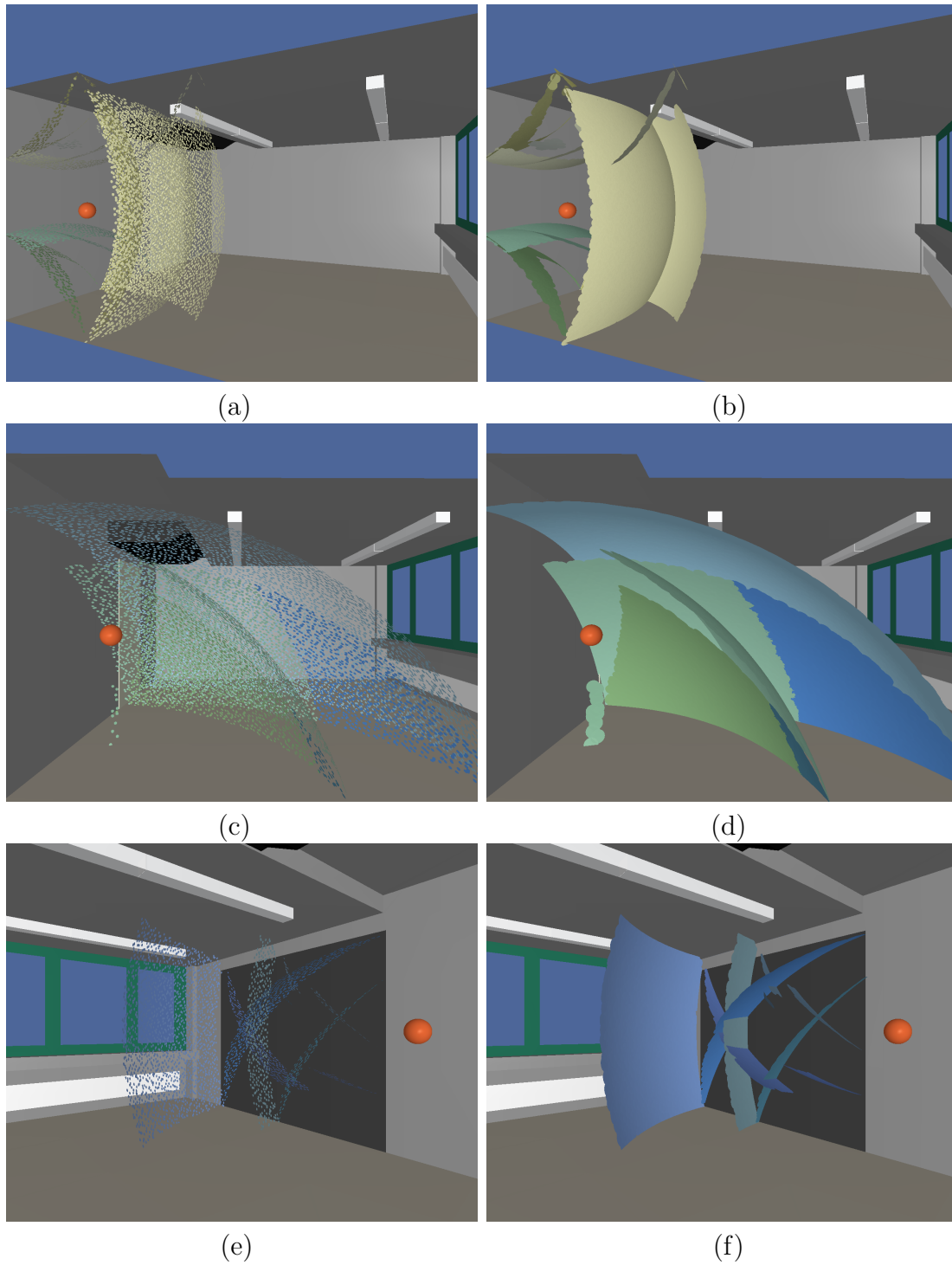
Figure 5.8: Wave fronts reflected from walls (a+b), floor (c+d), and projection wall (e+f) using smaller (a,c,e) and greater (b,d,f) radius.

5.2.2 Visualizing phonons on surfaces

The next method to examine the acoustics inside the room using the phonon map information is the visualization of certain phonons at their position inside the given scene. Each phonon is rendered as a sphere and is colored according to its spectral energy. The option to consider all frequency bands in total or each of them separately is provided, whereas 10 frequency bands associated with the frequencies $40, 80, 160, ..., 20480$ Hz are considered. In the first case the spheres are color coded by using the RGB components as described in the previous section.

In the second case, considering only one frequency band, the energy of the phonons in this frequency band is color coded using the HSV model. Therefore, we can interpolate the color of the spheres between red (full energy) and blue (energy equals zero) corresponding to the energy $e_{p,i}$ of the i-th frequency band. In order to show the phonons outgoing direction v_p a cone whose peak is rotated towards v_p is rendered. Also, the color of the cone corresponds to the phonons energy.

Since the number of phonons in the phonon map is large (as discussed earlier) we render only phonons with a given number of reflections n_p simultaneously. With this approach we can examine how the surfaces of the considered scene affect the overall acoustic of the room. The phonon map contains for each phonon the number of reflections n_p, so only those phonons with a certain number of reflections are displayed, visualizing their frequency spectrum and outgoing direction.

Figure 5.9 shows an example of this visualization approach. The phonon map consists of one million phonons. In figures 5.9 (a) and (b) the overall frequency spectrum of the phonons after one and four reflections, respectively, is depicted using the RGB components. As we can see in figure 5.9 (a) the walls, the bottom and the canvas absorb high frequencies better than low frequencies (blue color), whereas the door, for example, reflects all frequencies equally. After four reflections at the scene surfaces we observe a shift towards lower frequencies. The representation using the RGB model shows an average of the energy spectrum at low, middle and high frequencies. Sliding through the frequency bands we observe the absorption for each individual frequency band. Figure 5.9 (c, d) shows the energy at 160 Hz and figure 5.9 (e, f) at 10240 Hz after one and four reflections, respectively. As we see, after four reflections there are about 75% of the energy of the phonons at 160 Hz, whereas the energy at 10240 Hz is almost completely absorbed by the room. By depicting the outgoing direction we can guess which object the phonon will hit next.
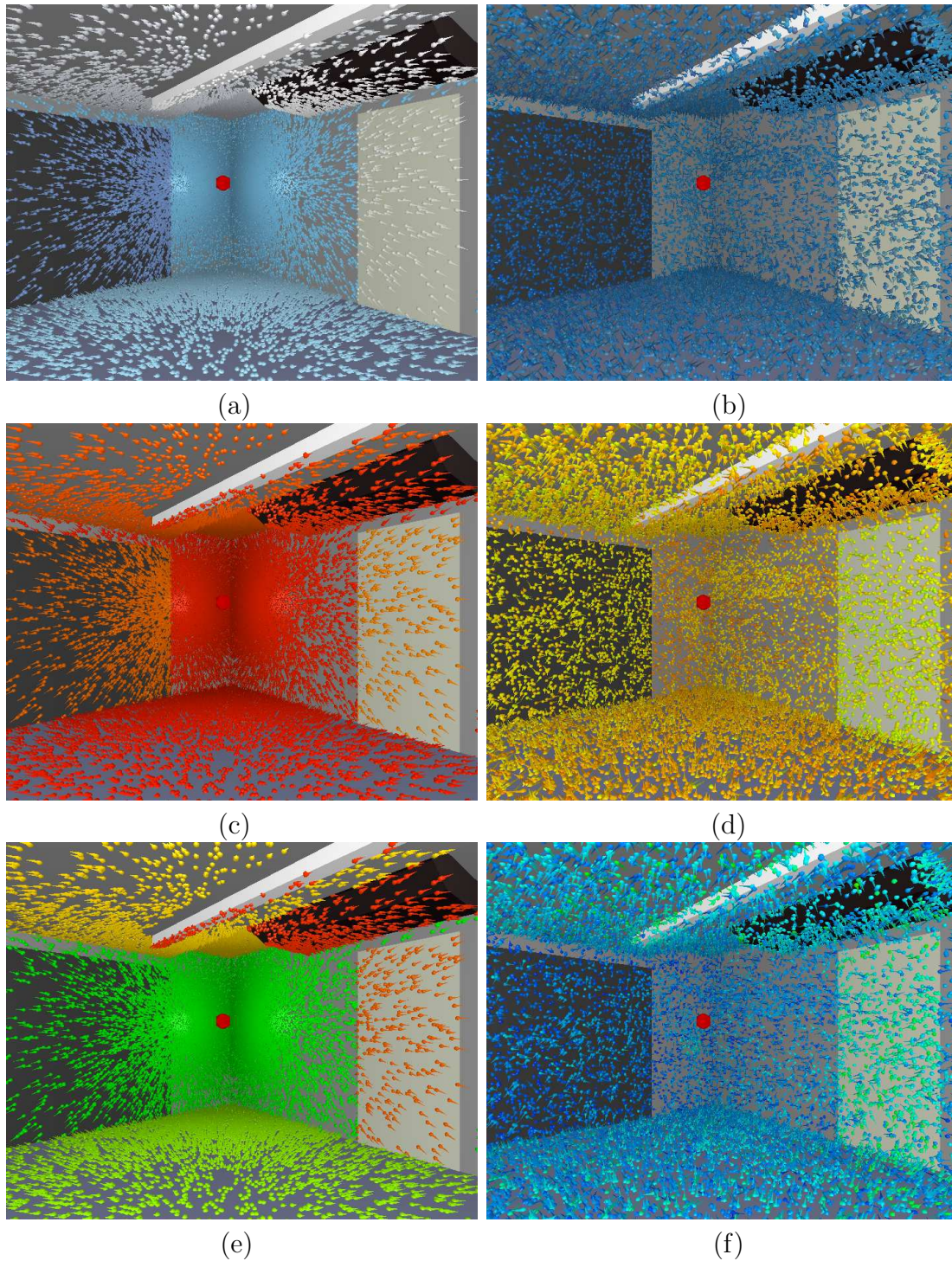
Figure 5.9: Visualization of phonons on surfaces. First (a) and fourth (b) reflection
color coded using the overall frequency spectrum. First (c, e) and fourth
(d, f) reflection color coded using the frequency band at 160 Hz and at
10240 Hz, respectively.

5.2.3 Reflected wave front visualization

(a)	(b)
(c)	(d)
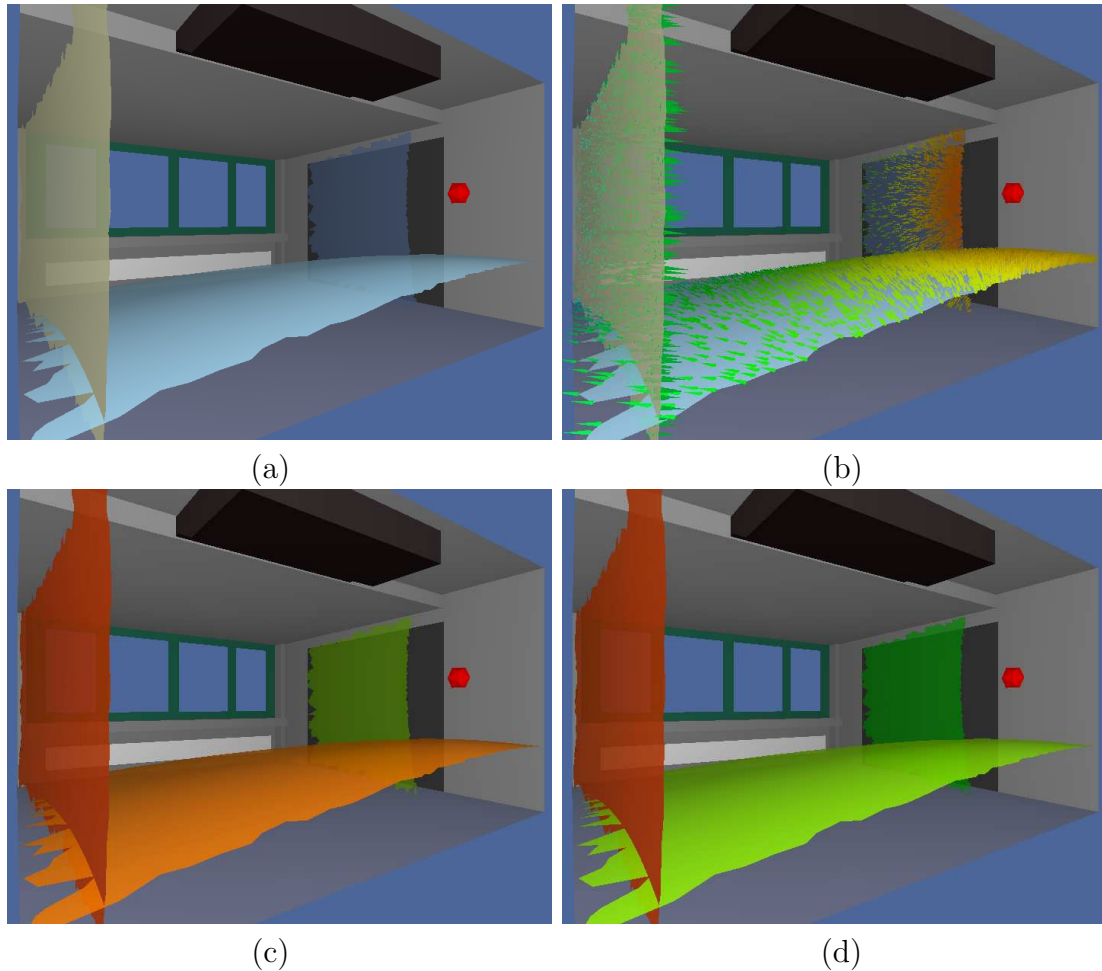
Figure 5.10: Clustered wave fronts. (a) first reflection color coded using the overall frequency spectrum. (b) Traversed distance represented by use of cones. (c) and (d) first reflection color coded using the frequency band at 640 Hz and 10240 Hz respectively.

In this section another approach visualizing wave fronts reflected at the room surfaces is described. Triangulated surfaces are used instead of rendering single particles. In order to build these surfaces we need to know which phonons belong to a common wave front. Therefore we subdivide the phonons in clusters of equal history. This subdivision is described in section 4.4, where level of detail approach using phonon tracing algorithm was discussed in more detail. This approach is used for visualization of early reflections of sound waves at the scene objects, thus all phonons inside a cluster have equal energy spectra.

The wave front surfaces reflected at the objects inside the considered room are

(a)   (b)
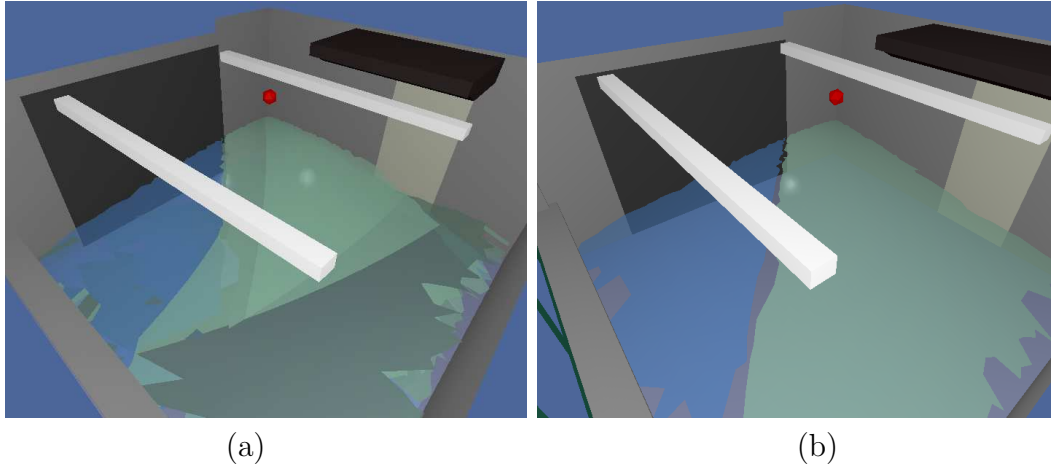
Figure 5.11: Clustered wave fronts. Second reflections from the floor. First reflected
from all other room faces (a) and from the wall and the powerwall,
respectively (b).

built by keeping only triangles of the initial wave front (the construction of which
was described in section 5.2.1.2) whose vertices (phonons) reside in the same cluster.
The other triangles are dropped since they partially belong to wave fronts reflected
from other surfaces. Now we can render the resulting surfaces for visualization of
the wave fronts reflected from the scene objects. To illustrate where the wave front
hits the object first and in which direction it propagates we can use the phonons
traversed distances d_p to deform the surface. First, the maximum traversed distance
d_{max} of all phonons inside the cluster is determined. Then, the phonons (which are
the triangle vertices of the wave front surface) are rendered in an offset distance
from the corresponding scene surface:

$$p_p \leftarrow p_p + (1 - \frac{d_p}{d_{max}}) \cdot n \tag{5.2}$$

where p_p is the phonon's position on the surface and n the scene surface normal.
This way, the offset between the original surface and the visualized wave front is
at most one unit. The wave front surfaces are color coded according to the energy
spectra of corresponding phonons. Depending on whether we want to examine all
frequency bands in common or each of them separately we can use the RGB model
or the HSV model as described in the previous section. Furthermore, we set the
surface transparency according to the average energy of the wave front.

For better clarifying the propagation direction of the wave fronts and the traversed
distance, the option to draw colored cones at the position of the phonons is provided.
The peaks of these cones are rotated towards the phonons outgoing directions. Their
color corresponds to the traversed distance and is calculated by use of the HSV model

as follows:

$$(1 - \alpha) \cdot H_{red} \; + \; \alpha \cdot H_{blue} \text{ with } \alpha = \frac{d_p}{l_{max} \cdot n_p} \tag{5.3}$$

where l_{max} is the maximum distance of the scene. In this case red corresponds to $d_p = 0$ and blue to the maximal distance depending on the number of reflections n_p. Using the above equation for color mapping the cones we can compare the traversed distance of wave fronts reflecting at different objects after a given number of reflections. Since displaying all wave front surfaces becomes complex, the implementation provides the alternative to select wave fronts by teh number of reflections, latest reflected material (object) or reflection history.

Figures 5.10 and 5.11 show examples of the described visualization approach. The colored sphere in the corner of the room represents the sound source. Figure 5.10 depicts the wave fronts of first reflections coming from the floor, the wall and the canvas. In image 5.10 (a) and (b) we see the color coded wave fronts using the RGB model and overall frequency spectrum as described above. We can see that the floor and the canvas predominantly absorb more high frequency, whereas the wall absorbs more low frequencies. Figure 5.10 (c) and (d) shows the color coded wave front surfaces by use of HSV model and the frequency band at 640 Hz and 10240 Hz, respectively. The sound energy at 10240 Hz frequency band is absorbed by the canvas and floor, but is reflected almost completely by the wall. At 640 Hz the canvas acts absorbing, whereas the floor and the wall reflects the most part of the energy.

The surfaces are deformed according to the traversed distance of the phonons belonging to the clusters. Here, we can observe where the wave fronts hit the room surfaces first and in which direction they will further spread. A representation of the traversed distance by use of colored cones (figure 5.10 (b)) shows that the wave front propagated from the sound source hit the canvas and the floor before the wall.

Figure 5.11 presents the wave front reflecting from the floor the second time. Since reflections coming from all objects inside the scene can overlap (fig. 5.11 (a)), the option to select different wave fronts is provided. In figure 5.11 (b) we see the second reflection from the floor reflected first from the wall and the canvas, respectively. In this picture we observe how the energy spectrum of the wave front changes and how it is split during the reflections at the room surfaces. This way we can examine the interaction of the single objects inside the room.

Since the wave front traversed from the sound source is split after each reflection, the visualization approach described in this section is applicable only to the first few reflections. The late reflections are handled by an visualization approach utilizing scattered data interpolation methods as described in [DMB+06].

5.2.4 Listener-based visualization

The approaches described above visualize the information stored in the phonon map considering the surfaces of the room and their acoustic properties. In this section a visualization method is presented depicting the energy received at a given listener position. With this approach we can detect from which direction the energy reaches the listener and visualize the energy spectrum.

For this purpose a triangulated sphere deformed according to the weighted phonons received at the listener position is rendered. The phonons are collected using the collection step of the phonon tracing algorithm (see chapter 4). For each phonon which contributes to the total energy at the listener position, first the intersection point $p_{intersec}$ of the ray from the center c_s of the sphere to the phonon's position ph_{pt} with the sphere is calculated. The energy e_{sp} is then increased and the displacement $disp_{sp}$ of the point sp of the intersected triangle with minimal distance to $p_{intersec}$ is calculated as follows:

$$e_{sp} \quad \leftarrow \quad e_{sp} + fac \cdot \begin{pmatrix} (e_{p,8} + e_{p,9} + e_{p,10})/3 \\ (e_{p,5} + e_{p,6} + e_{p,7})/3 \\ (e_{p,1} + e_{p,2} + e_{p,3} + e_{p,4})/4 \end{pmatrix} \tag{5.4}$$

$$disp_{sp} \quad \leftarrow \quad disp_{sp} + fac \cdot \tfrac{1}{10} \sum_{i=1}^{10} e_{p,i}$$

$e_{sp} \in \mathbb{R}^3$, $disp_{sp} \in \mathbb{R}$ if all frequency bands are considered and

$$e_{sp} \quad \leftarrow \quad e_{sp} + fac \cdot e_{p,i}$$
$$disp_{sp} \quad \leftarrow \quad disp_{sp} + fac \cdot e_{p,i} \tag{5.5}$$

$e_{sp} \in \mathbb{R}$, $disp_{sp} \in \mathbb{R}$ looking only at a certain frequency band. Thereby, fac is the weight determined using the Gaussian function described in chapter 4. In order to smooth the sphere representation the energy of the points sp_j around sp is increased if $arccos(\langle c_s - sp_j, c_s - sp \rangle) < \alpha$ for a given angle α as follows:

$$\begin{aligned} e_{sp_j} &= w \cdot e_{sp} \text{ where} \\ w &= 1 - (arccos(\langle c_s - sp_j, c_s - sp \rangle)/\alpha) \end{aligned} \tag{5.6}$$

Afterwards, the energy and the displacement of the sphere points are normalized. Therefore the maximum displacement $disp_{max}$ and energy e_{max} (in all three components) of all points is determined. Then these values are normalized:

$$e_{sp_j} \quad \leftarrow \quad \frac{e_{sp_j}}{e_{max}}$$
$$disp_{sp_j} \quad \leftarrow \quad \frac{disp_{sp_j}}{disp_{max}} \tag{5.7}$$

In order to consider occlusion of the listener position, the energy values of the sphere points sp_j are scaled as follows:

$$
\begin{aligned}
e_{sp_j} &\leftarrow e_{sp_j} \cdot \frac{n'_{ph}}{n_{ph}} \\[2mm]
disp_{sp_j} &\leftarrow disp_{sp_j} \cdot \frac{n'_{ph}}{n_{ph}}
\end{aligned}
\tag{5.8}
$$

where n_{ph} is the number of phonons in the phonon map and n'_{ph} the number of phonons which are seen from the listener.

Now, after the calculation of the displacement factors and energy at each point of the sphere we deform and display the spheres at given listener positions. The new position of the point sp_j results in:

$$
sp_j \leftarrow sp_j + disp_{sp_j} \cdot \frac{c_s - sp_j}{\|c_s - sp_j\|}
\tag{5.9}
$$

The color of the sphere points is calculated by use of the RGB (for overall frequency spectrum) or the HSV (certain frequency band) model as mentioned in the previous sections. For assessing the acoustic quality at a listener position it is important to know at which time the reflections arrive at the listener. For this reason we add the feasibility to determine the lower and upper time limits for phonon selection. In order to better illustrate the amount of the incoming energy the spheres are additionally scaled by the factor of the currently received energy (until given time) divided by total received energy.

The following figures present the results of the introduced visualization approach. Figure 5.12 (a) show spheres at four positions in the considered room. The energy is collected over the entire time interval. The spheres are deformed and color mapped by using the overall energy spectrum. As we can see the most energy of low frequencies is received at the listeners from the floor and the ceiling, since they do not absorb low frequencies. Whereas middle and high frequencies reach the listener from the walls. Considering the energy at 80 Hz (figure 5.12 (b)), at 1280 Hz (figure 5.12 (c)), and 10240 Hz (figure 5.12 (d)) we observe that most part of the energy at 80 Hz reflects from the floor and ceiling, and at 1280 Hz and 10240 Hz the most part is reflected off the walls.

The spheres depicted in figure 5.13 (a) and (b) are deformed and color coded using the energy at 80 Hz of phonons that reached the listener position before and after 50 msec, respectively. We can observe that the whole energy of the reflections from the room surfaces is received at the listener before 50 msec. This way we can compare the amount of early to late reflections.

Figure 5.13 (c) shows a room with a separating wall with total absorption. This wall prevents most reflections from the surfaces on the left side of the room to the

right side and vice versa. One listener is placed at a position behind the wall and two listeners at the corners of the room. The spheres are visualized for the 5120 Hz frequency band. As we can see, no reflections are received from the separating wall and only few reflections reach the listener from the left and right room surfaces.

(a)   (b)

(c)   (d)
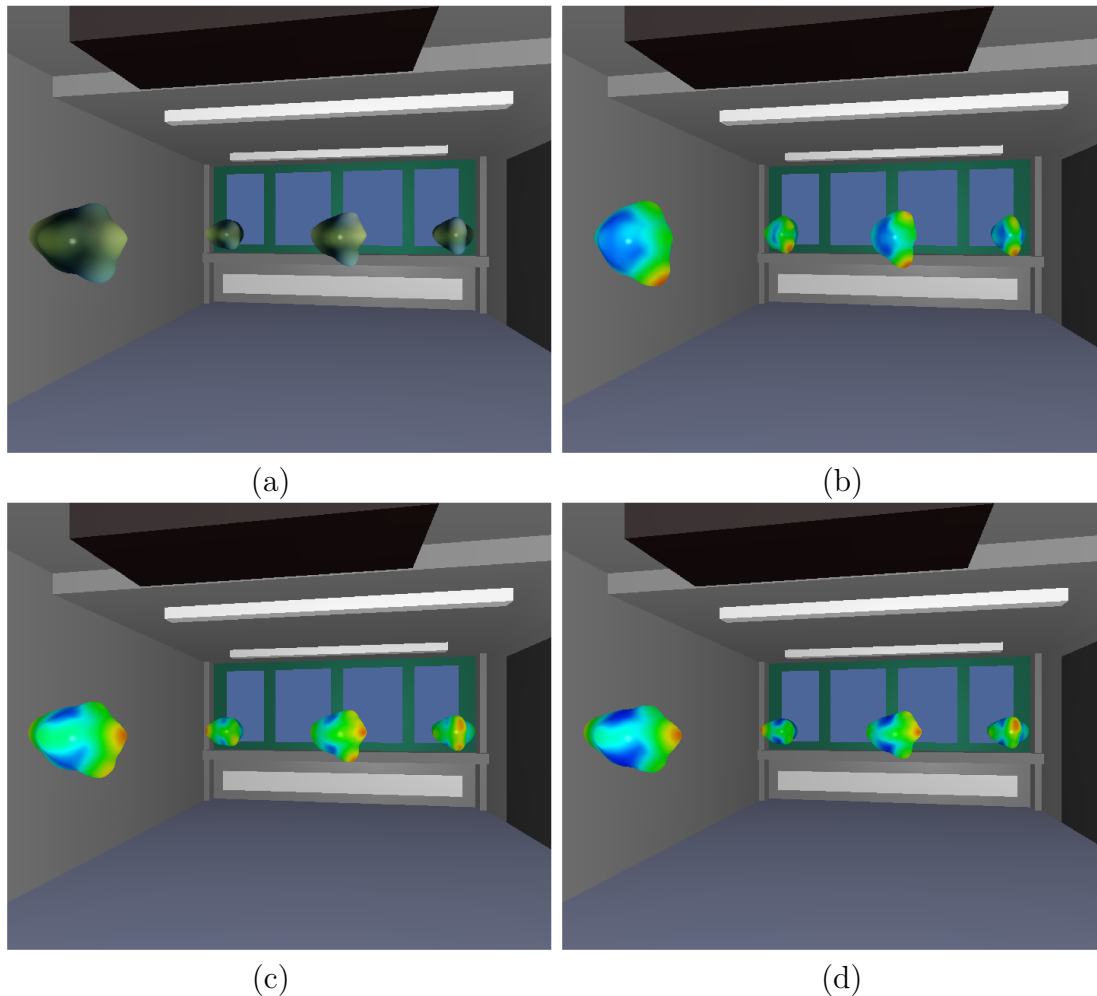
Figure 5.12: Deformed spheres representation at four listener positions. (a) color coded using the overall frequency spectrum, (b) by 80 Hz, (c) by 1280 Hz, and (d) by 10240 Hz.
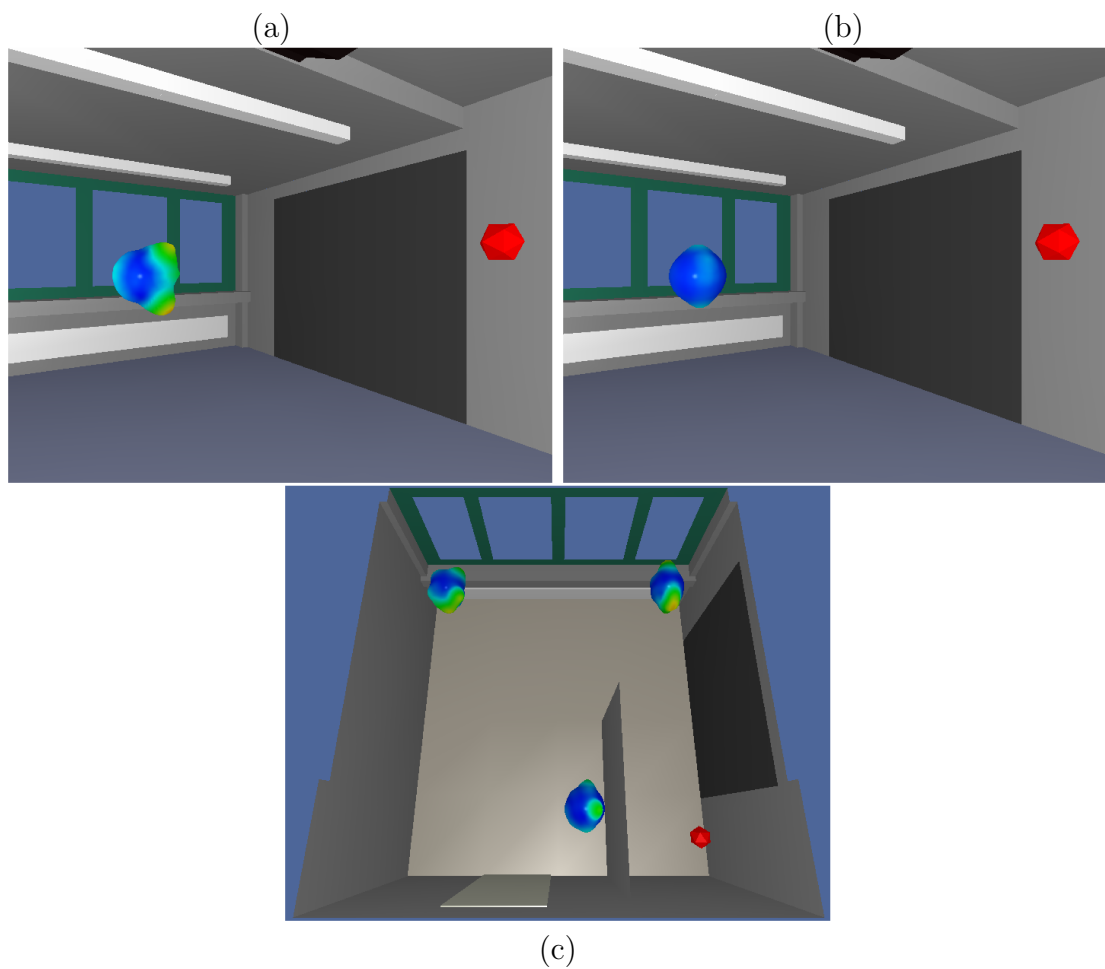
Figure 5.13: Deformed spheres representation at 80 Hz collected before (a) and after (b) 50 msec at one listener position. Deformed spheres representation at 5120 Hz at three position in a room with a separating wall with total absorption (c).

5.2.5 Visualization applications

In this section two application cases of the introduced visualization approaches are described, a VR laboratory and a lecture room.

5.2.5.1 VR laboratory

| (a) | (b) | (c) |
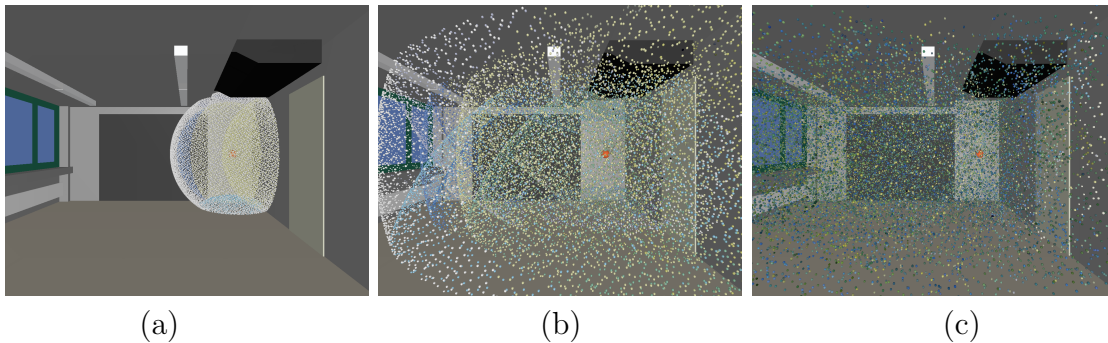
Figure 5.14: Wave front propagation from a spherical sound source. The visualization shows phonons with RGB color-coded spectral energy due to reflection at different materials at 1.5m (a), at 4.5m (b), and at 10m (c).

The visualization approaches described above were used to examine the acoustic properties of a virtual reality (VR) laboratory that is also used for auralization purposes. Figure 5.14 shows the wave front propagation based on 30000 paths traced from the sound source resulting in 910610 phonons in the phonon map, at a traversed distance of 1.5, 4.5, and 10 m. At small distances / traversal times, the individual wave fronts can be recognized, whereas large distances provide insight into the frequency decomposition of the various reflections. We observe a shift towards lower frequencies, since the phonons' color is dominated by the blue component after a number of reflections.

In order to identify the reason for the frequency shift, we look at phonons reflected from a selected material (at least once) in their earlier paths. Figure 5.15 shows these phonons for reflections from walls, floor, and canvas, respectively, for several reflection numbers. We observe that phonons reflected from walls carry mostly a yellowish color, despite of their potential reflection from additional materials. Hence, the energy of these phonons is shifted towards the mid and high frequencies. The floor and the canvas reduce high frequencies. Reflections from the canvas affect mostly the right side of the room, whereas the impact of the carpet is much greater. While a potential frequency shift can already be seen in material absorption coefficients, their impact on room acoustics can be studied much better with the aid of our visualization approach.
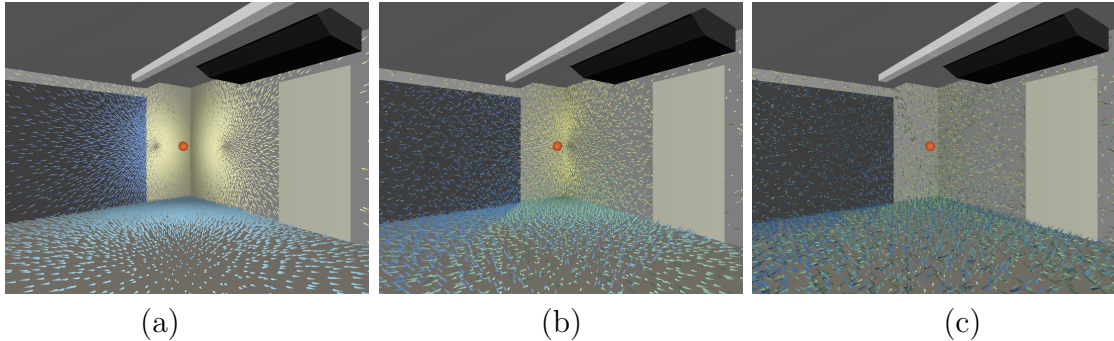
Figure 5.15: Visualization of the phonons at walls, floor, and canvas, after first (a), second (b), and third (c) reflection.

In figure 5.16 the carpet is replaced by a material with similar absorption as the walls. When comparing figures 5.16(a) and 5.16(b), it becomes evident that this change is sufficient for increasing the intensities of mid and high frequencies. The acoustic properties of this room are significantly improved by this modification, since low frequencies are also propagated by diffraction and thus do not depend on linear reflections that much. While the acoustics of the laboratory are not much of importance, it may have a greater impact on the design of larger classrooms. Optimizing the acoustics of such larger rooms may, for example, eliminate the need of using a microphone or improve the auditive quality of concerts.
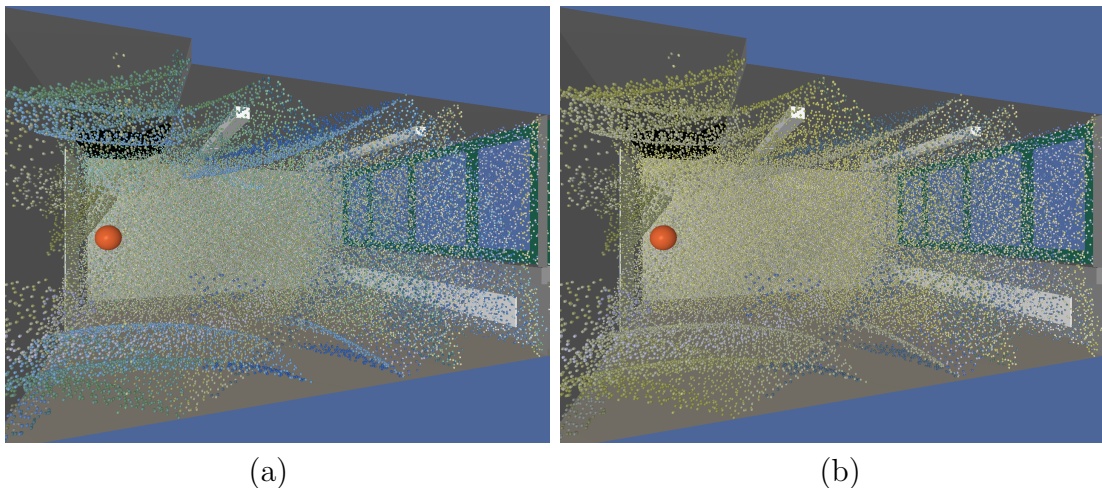
Figure 5.16: Changing material of the floor. (a) all phonons at d = 4.5m; (b) same as (a) with new material.

5.2.5.2 Lecture room

In the following application example we examine the acoustics of a lecture room concerning the speech comprehensibility. First we look at the room in a structurally complete state, where the floor, walls and ceiling are less absorbent (made of concrete for example). Now we trace the sound particles from a source located near the blackboard (red sphere in the figure 5.17) populating the phonon map. Afterwards, we look at the energy received at two different listener positions inside the room using the listener-based visualization approach described in subsection 5.2.4. We examine the sound energy at the 1280 Hz frequency band, since it is the medial frequency range of human speech. The speech comprehensibility can be judged objectively utilizing the acoustic parameter definition D_{50} (definition index C_{50}) which relates the early arrived sound energy (until 50 msec) and total energy (late energy after 50 msec). The sound definition and definition index are described in section 2.3. So we look at the energy received by the listener before and after 50 msec time limit.

As we observe in the figures 5.17 (a) and (b) the late reflected energy portion is greater compared to the early reflected sound indicating bad comprehensibility of the speaker. If we take a look at the direction of the received energy we see that the greatest amount of the late reflected sound arrives at the first listener ($Listener1$, far from the source) from the side wall and at the second listener ($Listener2$, near to the source) from the back wall. Changing the material of the walls to more absorptive material (about 50% absorption at 1280 Hz) results in the situation depicted in figures 5.17 (c) and (d). At the listener position near the source ($Listener2$) we obtained a satisfactory value of early to late energy ratio. At the position of $Listener1$ we can see an improvement but the late reflected energy part is still higher that the early sound energy part. In figure 5.17 (d) we observe that now a great amount of enegy reaches the $Listener1$ after 50 msec from the floor and also from the ceiling. Figure 5.17 (c) shows that the floor is important for early reflections at the $Listene2$, thus changing the floor material to a high absorbant material would reduce the early energy at this position. Hence, we change the floor material to a less absorbant material (about 10%) and the ceiling material to more absorbant material (about 50%). As we now see in the figures 5.17 (e) and (f), we have achieved our goal. The portion of the early energy is greater than that of the late reflected sound energy at both analyzed listener positions.

It should be noted that when changing the surface materials, the phonon map needs not to be recomputed, provided that enough reflections are incorporated. Only the collection of the phonons in order to calculate the spheres color and deformation is re-calculated what makes this examination quite interactive (depending on the number of phonons in the map). Here, each listener position needs to be analyzed more or less individually. An improved approach indicating the importance of the room surfaces for one or multiple listeners regarding an acoustic property is introduced in the next section.
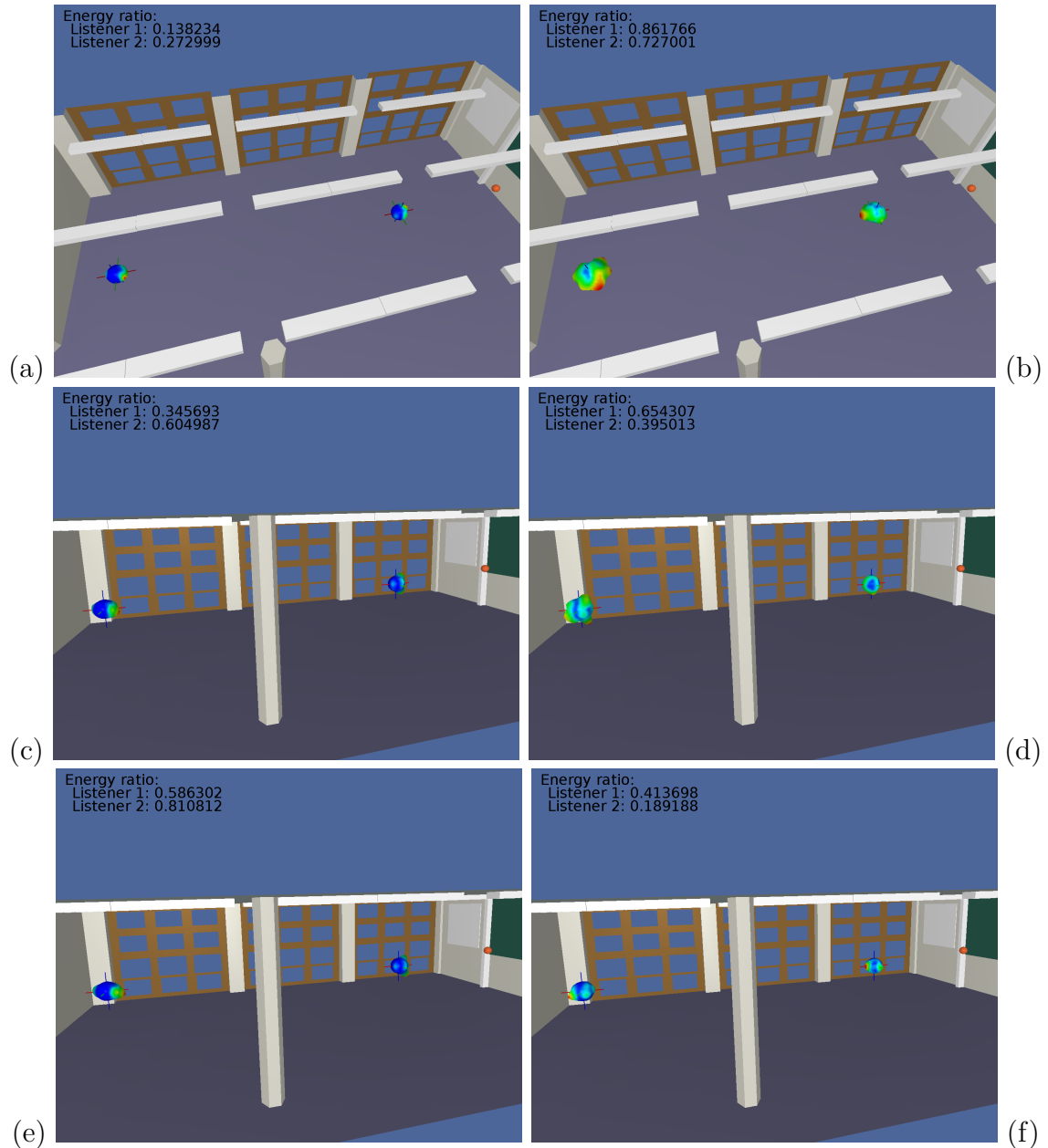
Figure 5.17: Visualization of the energy received at two listener positions before (a,c,e) and after (b,d,e) the time limit of 50 msec. In (a) and (b) the floor, the walls and the ceiling are highly reflective. In (c) and (d) the walls material is changed to a more absorbent material. Additionally, in (e) and (f) the ceiling material is changed to be more absorbent and the floor material is is changed to less absorbent one. The energy ratio is calculated relative to the total energy amount. The red sphere indicates the sound source.

5.3 Listener-based analysis utilizing acoustic metrics

This section describes a visualization application using the phonon tracing method to calculate the contribution of the reflection surfaces to the impulse response for different listener positions ([MDHB$^+$07]). This data is used to compute importance values for the geometry. Here importance denotes how much a certain scene surface contributes to a quality measure (see section 2.3). To get a visual insight into the directional aspect, the importance is mapped to the reflecting surfaces of the geometry. This visualization indicates which parts of the surfaces need to be changed to enhance a certain acoustic quality measure at different listener positions or the complete audience.

5.3.1 Importance values

The intention is to visualize which parts of the scene surface are most important for a certain acoustic metric at every listener position l_i. To calculate the importance value for a phonon the pressure contribution of the phonon ph_j to the impulse response at a given listener position l_i is used.

Because the phonon tracing method re-uses particle paths for multiple phonons (a phonon is stored at every reflection, tracing a large number of reflections until the energy contribution drops below a threshold) we have to take subsequent phonons on the same path into account when calculating the pressure contribution for a certain phonon ph_j (see figure 5.18).
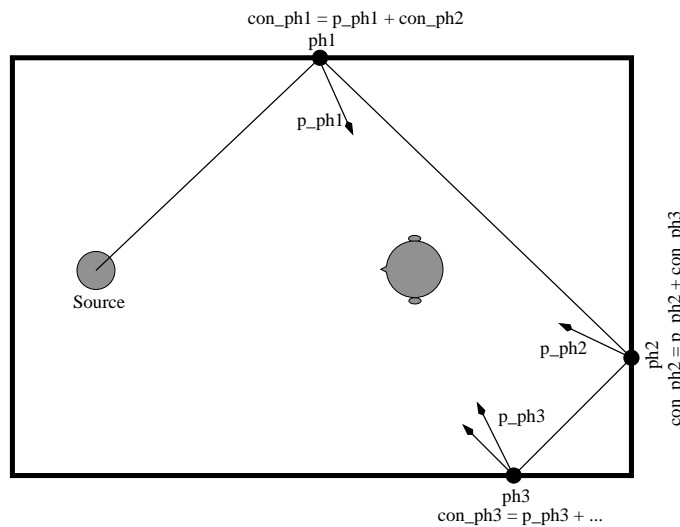
Figure 5.18: Depiction of phonons sharing the same path.

All phonons corresponding to one reflection path are listed in consecutive order within the phonon map, say $ph_{k_0}, ph_{k_0+1}, \cdots, ph_{k_n}$. The pressure and energy contribution of a phonon ph_j to the impulse response at listener position l_i is then the accumulated contribution of itself and its successors as stated in equation (5.10) for pressure con_p and energy con_e.

$$
\begin{aligned}
con_p(ph_j, l_i) &= \sum_{k=j}^{k_n} p_{ph_k}(t, l_i) \\
con_e(ph_j, l_i) &= \sum_{k=j}^{k_n} e_{ph_k}(t, l_i),
\end{aligned}
\tag{5.10}
$$

where ph_j is the current phonon, l_i the listener position and p_{ph_k}, e_{ph_k} are the pressure and energy contribution of a phonon as described in sections 4.1 and 4.2.

Due to the fact that all pathes are known we can easily compute the contribution of a phonon by forward summation of the contribution of subsequent phonons. In addition to the importance of the phonon to one listener position l_i also the contribution of the phonon to all listener positions in the scene (or to a selected subset) are calculate by:

$$
\begin{aligned}
con_{p\ all}(ph_j) &= \sum_{i=0}^{n_l} con_p(ph_j, l_i) \\
con_{e\ all}(ph_j) &= \sum_{i=0}^{n_l} con_e(ph_j, l_i)
\end{aligned}
\tag{5.11}
$$

where n_l is the number of listeners, $con_p(ph_j, l_i)$ and $con_e(ph_j, l_i)$ are the contribution of a phonon to a single listener position l_i.

To calculate an importance value $imp_e(ph, l_i)$ and $imp_p(ph, l_i)$ for a phonon ph and listener position l_i the fraction between the contribution of the phonon ph to the pressure or energy at the listener position l_i and the total received pressure or energy at the listener position (i.e. percentage of pressure received at position l_i

from phonon ph) are calculated.

$$
\begin{aligned}
imp_p(ph_j, l_i) &= \frac{con_p(ph_j, l_i)}{p_{tot,l_i}} \\
p_{tot,l_i} &= \sum_{k=0}^{n_{ph}} p_{ph_k}(t, l_i) \\
imp_e(ph_j, l_i) &= \frac{con_e(ph_j, l_i)}{e_{tot,l_i}} \\
e_{tot,l_i} &= \sum_{k=0}^{n_{ph}} e_{ph_k}(t, l_i)
\end{aligned}
\tag{5.12}
$$

where l_i is the listener position, $con_p(ph_j, l_i)$ and $con_e(ph_j, l_i)$ are the contributions of phonon ph_j to the listener position l_i, n_{ph} is the total number of phonons, p_{tot,l_i} and e_{tot,l_i} are the total pressure and energy received at l_i.

To evaluate speech comprehensibility in room acoustics, for example, it is important to relate the early reflections to the later reverberations. Therefore, an option to choose a time limit t_{end} such that the contribution of each phonon that reaches the listener after t_{end} is set to zero is provided. By means of the importance values for each phonon in the map a visualization of the impact of reflection surfaces in the scene for a chosen range of listeners is proposed. The visualization approach is described in the following section.

5.3.2 Visualization

The goal of the visualization is to highlight the surfaces of the scene that reflect the most energy of a sound wave to a given listener position and therefore are most important for the evaluation of acoustic behavior regarding acoustic measures such as definition ("Deutlichkeit", defined in section 2.3) and definition index.

First, some preprocessing steps in order to prepare the scene geometry for the visualization need to be performed. The scene consists of triangulated objects. After the tracing process phonons' positions are known and each phonon can be connected to the corresponding reflecting scene object. The triangulation of the scene used for the acoustic simulation is low in resolution to keep the computation moderate. For a detailed visualization more sample points are needed in order to provide the correct approximation of the surface importance. Therefore the triangles are refined by use of a subdivision algorithm as outlined in Fig. 5.19. The subdivision is terminated if all edges of a triangle have a smaller length than a given threshold λ.

In order to obtain the desired level of detail for the visualization color values are mapped to the vertices depending on the contributions of the phonons associated
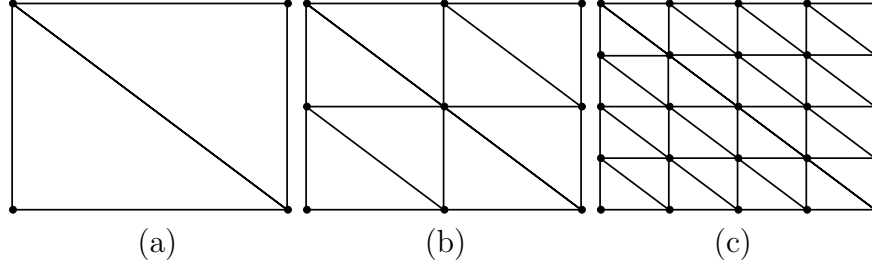
Figure 5.19: Surface subdivision. (a) initial triangulated surface. (b) triangulation after first subdivision step. (c) triangulation after second subdivision step.

with adjacent triangles. Since the number of phonons is much greater than the number of vertices a scattered data approximation approach is used:

$$
\begin{aligned}
imp_{v_i} &= \frac{\sum_{k=0}^{m} \omega \cdot imp(ph_k)}{\sum_{i=0}^{m} \omega} \\
\omega &= \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{d^2}{2\sigma^2}}
\end{aligned}
\tag{5.13}
$$

where σ is proportional to the edge length and d is the distance between the phonon's position and the vertex v_i. Furthermore only phonons ph_k which satisfy the condition shown in the equation below are considered:

$$
\langle n_t, n_{v_i} \rangle > 0
\tag{5.14}
$$

n_t is the normal of the triangle containing the phonon ph_k and n_{v_i} denotes the normal at v_i. The latter can be calculated as interpolation normal of the corresponding triangle normals.

The second approach which is utilized to calculate the importance of the vertices takes only phonons ph_k into account which are located in a tight neighborhood of v_i This means that the distance of the phonon's position to the considered vertex is smaller than a given threshold. The importance value of v_i is then the average of the importance values $imp(ph_k)$ of the phonons ph_k:

$$
imp_{v_i} = \frac{\sum_{k=0}^{m} imp(ph_k)}{m}
\tag{5.15}
$$

Additionally, the condition in 5.14 have to be satisfied for two-sided scene elements.

Now, after the calculation of the importance values for each vertex in the scene the color of each v_i is assigned according to these values. The color value for v_i is

calculated as follows:

$$
\begin{aligned}
c_{v_i} &= \alpha \cdot hue_{max} + (1.0 - \alpha) \cdot hue_{min} \\
\alpha &= \frac{imp_{v_i} - imp_{min}}{imp_{max} - imp_{min}}
\end{aligned}
\tag{5.16}
$$

where imp_{min} and imp_{max} are the minimum and maximum importance values in the scene and hue_{min} and hue_{max} denote the minimum and maximum hue values, respectively. Red color is chosen for hue_{max} and blue color for hue_{min}. The color for the vertices v_i is then interpolated from red to blue in HSV color space. Having the refined triangulated geometry of considered scene, and color values for each v_i corresponding to the importance values the entire scene can now be displayed using OpenGL texture mapping techniques or simple primitives rendering.

Furthermore, the visualization system provides some additional options for user interaction:

- selection of break-off time t_{end} for early reflections
- selection of start time t_{start} for late reflections
- selection of one listener position l_i
- selection of one object o_j inside the scene

5.3.3 Results

The following section presents the results of the listener-based analysis of surface importance for acoustic metrics described above. Furthermore, the method is applied to the acoustic improvement of a lecture hall. An important acoustic measure for a lecture hall is the comprehensibility of speech which is expressed with the acoustic metric definition index (C_{50}). The C_{50} values alone show only the current acoustic state but give no insight how it can be improved (i.e. where to put absorbing material). The listener-based analysis of surface importance overcomes this drawback by visualizing the directional aspect by means of mapping importance values to the room surfaces. In the case of definition index the importance is based on energy received in the first $50ms$ after the primary wave front and energy received after the first $50ms$, respectively. The visualization of these importance values give a clear advice which parts of the room geometry need to be changed (by means of exchanging material) to improve definition index.

To assess the value of the visualization glyphs which represent the C_{50} value at different listener positions are used. The C_{50} values are derived from simulated impulse responses for each position.

5.3.3.1 Glyphs

In order to visualize definition index glyphs analog to the approach described in [SG89] are used. The representation of the glyphs is as follows. The early reflections and late reverberations are represented as semicircles where the semicircle for the early reflections is placed above that for late reverberations. The radii of the semicircles correspond to the early and late received energy, respectively. Thus, the direct comparison of early and late reflections is possible. Additionally to this ratio the glyphs also show whether the definition index is above a certain threshhold (i.e. $C_{50} > 3dB$). This is visualized by coloring the lower semicircle green for above or red for below the chosen threshhold, respectively. This is depicted in figure 5.20.

Nevertheless, this metric is independent from the direction of the arriving sound wave at a listener position and gives only insight in the current acoustic state. The definition index visualization by use of the glyphs is applied for evaluation of the listener-based importance visualization.
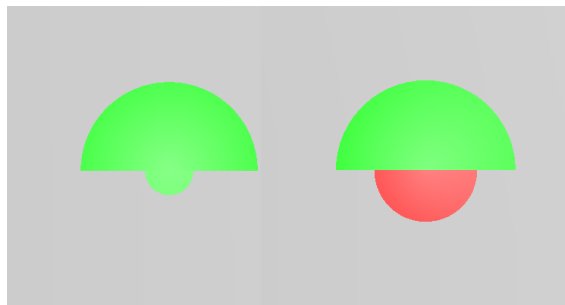
Figure 5.20: Example for definition index (C_{50}) visualization using glyphs. Left: C_{50} above given threshhold (green lower semicircle). Right: C_{50} below given threshhold (red lower semicircle).

5.3.3.2 Lecture hall

The university lecture hall subject to acoustic optimization is presented in figure 5.21. The materials used in this hall are as follows:

- concrete (floor, ceiling, gray walls)

- metal (doors, blackboard frame)

- wood (desk, benches, brown walls)
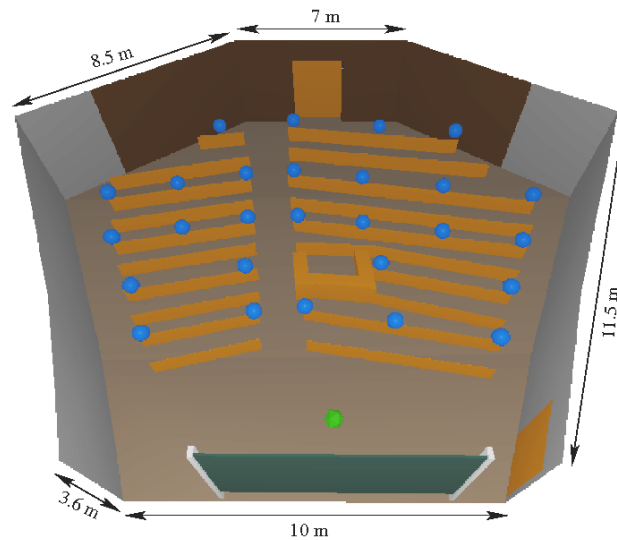
Figure 5.21: Lecture Room setup. Green: speaker, blue: audience

Speaker and audience are depicted by the green and blue spheres, respectively. The phonon tracing algorithm is used to calculate the impulse response and thus the definition and definition index at the listener positions. In order to achieve an overview of the current acoustic state of the room glyphs are visualized for the definition index (C_{50}) measure at all listener positions. Since comprehensibility of speech is the subject of improvement, the (C_{50}) values for the 1kHz frequency band are shown in figure 5.22 (a). These values are derived from the impulse responses at the different listener positions. For the simulation of the impulse responses the actual material/absorption coefficients were used. The threshold for C_{50} was set to $3dB$.

Figure 5.22 (a) clearly shows that the desired definition index is only reached at the three positions closest to the speaker (green lower semicircles of the glyphs). Depending on the individual hearing ability of the listeners the second and third row could be acceptable (red lower semicircle but smaller than upper green semicircle) whereas the definition index in the last two rows is unacceptable (red semicircle is

(a) (b) (c)
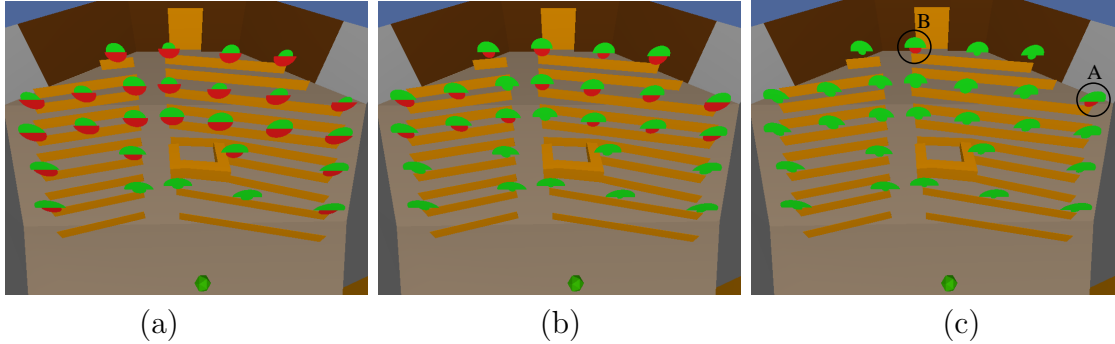
Figure 5.22: C_{50} Glyphs at listener positions in the lecture hall. (a) Actual materials/absorption coefficients. (b) Glyphs at listener positions for changed ceiling (above speaker) material with highly absorbing material (sound absorbing foam). (c) Glyphs at listener positions for changed backwall material with sound absorbing foam.

bigger than green semicircle). This clearly indicates the need of acoustic improvement, but the glyphs cannot give any cue how to alter the room so that the acoustic situation is acceptable.

5.3.3.3 Analysis of surface importance

To get insight into the directional aspect of the current acoustic situation the importance values are calculated based on the chosen acoustic metric (definition index C_{50}). As definition index describes the ratio between early ($< 50ms$) and late ($> 50ms$) received energy the break-off time is set to $t_{off} = 50ms$ for early and the start time to $t_{start} = 50ms$ for late reflections. All values were calculated for 1kHz frequency band with all listener positions taken into account. The importance values are then mapped to the room surfaces as shown in figures 5.23 (a) + (b) (topview) and 5.24 (a) + (b) (sideview). In addition to the importance for early and late reflections a difference image is calculated by subtracting the importance for early reflections from the importance for late reflections (figure 5.23 (c) + 5.24 (c)).

As described in section 5.3.2 the highest importance value imp_{max} is mapped to hue_{max} (red) and the lowest importance value imp_{min} to hue_{min} (blue). Thus red areas in figures 5.23 (a) + (b) and 5.24 (a) + (b) depict the parts of the room surface which contribute most to early (a) and late (b) reflections.

In figures 5.23 and 5.24 (c) red areas depict parts of the geometry which are important for late reflections but unimportant for early reflections and blue vice versa. Light green areas are equally important for early and late reflections. Thus the absorption coefficients of the surfaces in the red areas can be changed to alter the

<center>(a)   (b)   (c)</center>
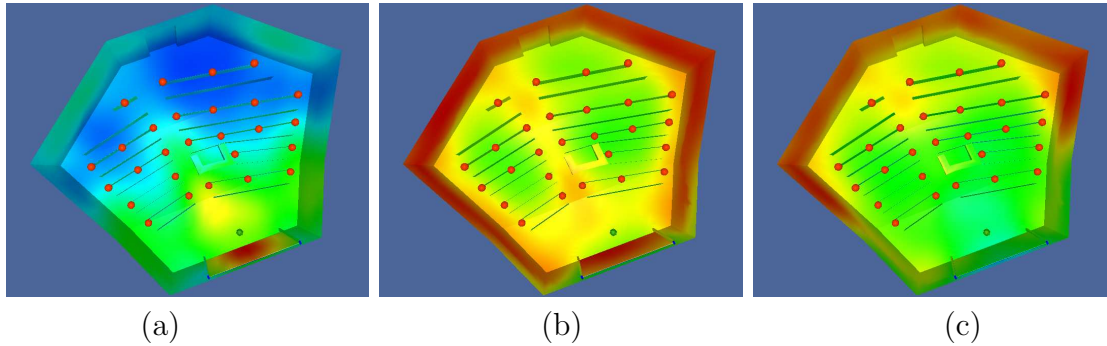
Figure 5.23: Importance values for whole audience and 1kHz mapped to the room surfaces (topview). (a) Importance for early reflections ($< 50ms$), (b) Importance for late reflections ($> 50ms$) red: highest importance, blue: unimportant. (c) Difference (importance for late reflections - importance for early reflections) red: only important for late reflections blue: only important for early reflections.

<center>(a)   (b)   (c)</center>
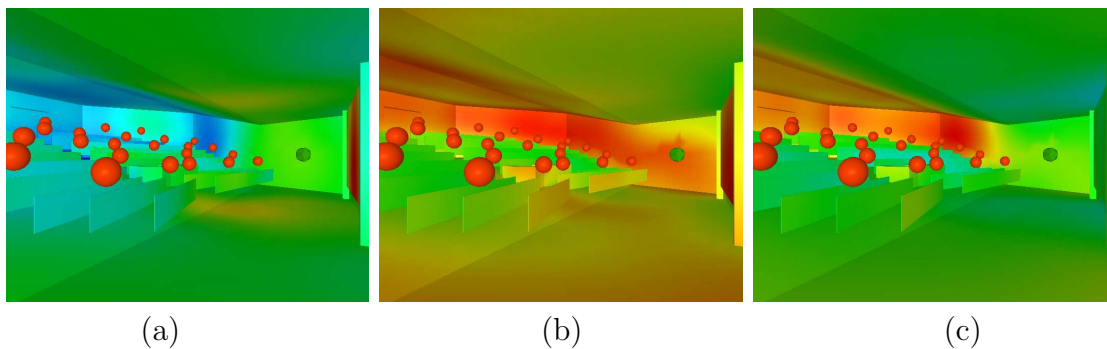
Figure 5.24: Same as figure 5.23 from side view.

amount of late received energy without large effect on the amount of early received energy, and vice versa.

To enhance the C_{50} values throughout the lecture hall, it is necessary to reduce the amount of late received energy at the listener positions. The visualization of the surface importance for early and late reflections gives a direct cue how to improve the overall definition index in the given room by simple measures. When looking at figures 5.23 (b) and 5.24 (b) one can directly locate the most important surfaces for late reflections which are the sidewalls of the room (red). An inspection of the difference images (figures 5.23 + 5.24 (c)) indicates that the change of the back wall material will not influence the early reflections too much (also red color of the back wall). Therefore the material of the back wall (wood) are exchanged with highly absorbing material (sound absorbing foam) to reduce the late reflections throughout the audience without cutting out the early reflections.

To evaluate the changes the C_{50} values are recalculated. Figure 5.22 (c) shows the improvement in definition index at the listener positions throughout the audience. The C_{50} values for almost all listener positions exceed the chosen threshold (green lower semicircles). A closer look has to be taken at the two positions which did not satisfy the desired threshold.

| (a) | (b) | (c) |
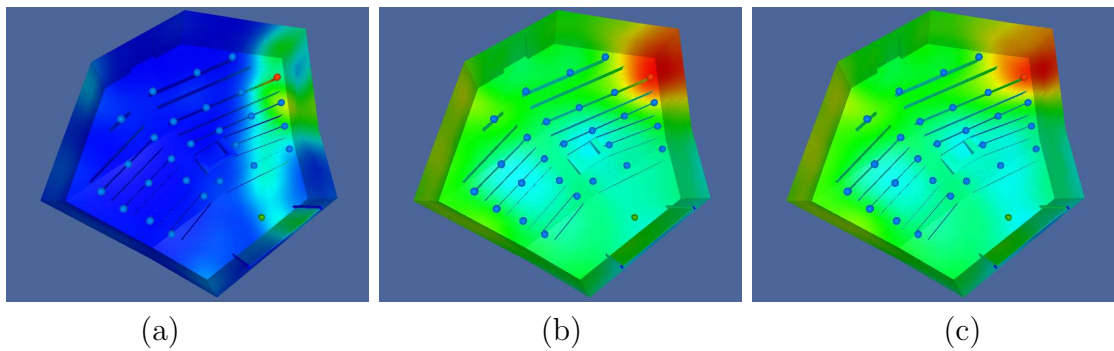
Figure 5.25: Importance values for listener position 20 and 1kHz mapped to the room surfaces (a) Importance for early reflections. (b) Importance for late reflections. (c) Difference. (Same as figure 5.23 for single listener).

| (a) | (b) | (c) |
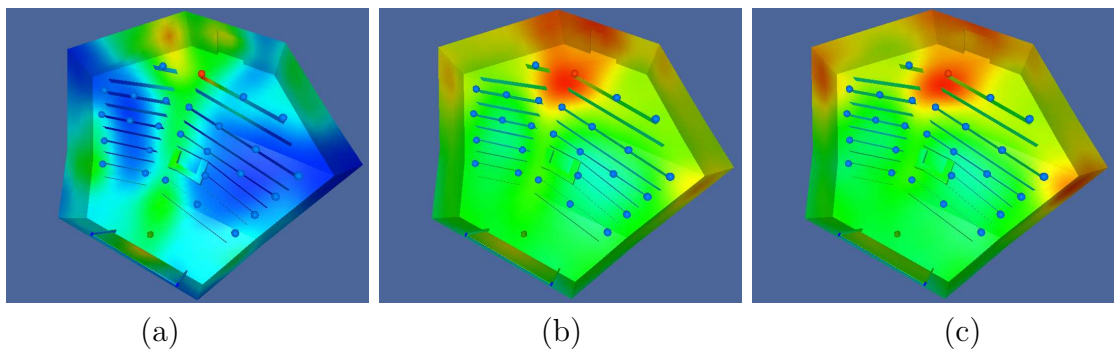
Figure 5.26: Same as figure 5.25 for listener position 26.

Figures 5.25 and 5.26 (a)-(c) show the importance values when only the single listener positions and not the whole audience is taken into account (importance for (a) early and (b) late reflections, difference image (c)). From figure 5.25 it is directly apparent that the change of the back wall material only has a medium impact on the definition index for this listener position. The main contribution for the late received energy is received from the rightmost corner of the room (which was not changed) rather than of the back wall. Thus if we want to enhance the definition index for this listener position only and not for the whole audience we need to change the material in the red areas in 5.25 (c). When looking at figures 5.26 (a)-(c) we can see a problem which occurs when we take the whole audience into account. While the back wall is important for late reflections but nearly unimportant for early reflections to the

whole audience, it is important for early and late reflections when taking only the topmost position into account. The exchange of the back wall material resulted in a decrease of late reflections to this position (which was intended) but also reduced the amount of early reflections. Thus the definition index (depending on the early to late reflection ratio) only improves slightly compared to the rest of the audience. To improve the definition index at this single position it would be most effective to alter the absorption of the floor material (red area in figure 5.26 (c)) for example exchanging the flooring material but the impact on the whole audience would not be as good as the conducted changes to the back wall.

For evaluation purposes the material of a surface (ceiling above speaker) which is on one hand of marginal importantance for late reflections and on the other hand also has some influence on the early reflections were also exchanged. The resulting C_{50} glyphs are visualized in figure 5.22 (b). As expected the improvement in definition index is only partial because the chosen surface is only medium important for late reflections taking the whole audience into account, whereas it is more important for the listeners in the first two rows (C_{50} reaches the desired threshold, i.e. green lower semicircles).

Overall, the listener-based analysis of surface importance proves to be a usefull tool to understand application specific-data and to get insight into the directional aspect of the acoustic situation in a room for a certain acoustic metric which is not apparent when looking at the resulting values only. Through the visualization of surface importances for a certain metric one could easily locate the parts of the room surface which affects this metric the most. Thus it is possible to improve the current acoustic situation with selective changes to the most important surfaces of the room.

6 Acoustic simulated reality system

Virtual Reality environments provide an immersive representation of computer-generated scenes. Integrating acoustic simulation, visualization, and auralization into the design process aims at interactive design and immediate exploration of virtual models. Several approaches to integrate acoustic simulation into Virtual Reality systems have been proposed in recent years. The DIVA system [SHLV99, SLH02, LSV+02] provides sophisticated synthesis and spatialization algorithms. Here sets of either perception-based or physics-based parameters for auralization is defined. In [NSG02] an audio-rendering system for use in immersive virtual environments, the blue-c, is proposed. It is optimized for efficient rendering of moving sound sources. Since the applications of this system are restricted to fictional scenes, no physical-based acoustic simulation is provided. In [LN06] the authors describe a four wall VR system for immersive visualization of sound rays propagation inside closed rooms. A real-time audio system applied in a CAVE like environment with four loudspeakers is proposed in [LSVA07]. A real-time simulation and rendering of moving sources and listeners is possible.
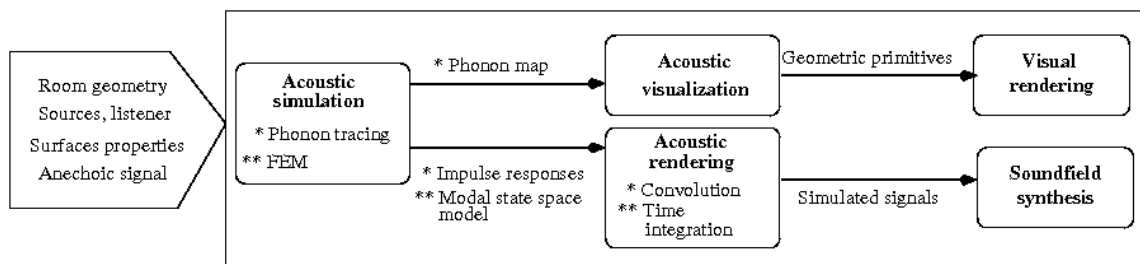
Figure 6.1: Modules of our audio-visual Virtual Reality system.

In this chapter we present a visual and auditive virtual reality display system. This system is a result of a startup project at the University of Kaiserslautern in collaboration with Fraunhofer institute of mathematics (ITWM) at Kaiserslautern. We integrate the phonon tracing algorithm as well as our visualization approaches into this system. A schematic overview of the VR system is given in figure 6.1. As input the system requires the geometry model of the room, the absorption or reflection properties of the room surfaces, sound sources characteristics, and listener positions. The module "Acoustic simulation" computes the modal state space model

(low frequencies) by use of the Finite Element Method (FEM) described in section 3.1, and the room impulse responses (RIR) and the phonon map (middle and high frequencies) by means of the phonon tracing algorithm introduced in chapter 4. On the basis of this information the anechoic source signal is modified for sound synthesis in the "Acoustic rendering" module. Thereafter, the soundfield synthesis is performed utilizing the acoustic hardware. The "Acoustic visualization" module provides the visualization approaches using the phonon map (see chapter 5). The graphical rendering ("Visual rendering" module) is implemented utilizing a stereoscopic back projection system. By use of the system, a walkthrough of the listener positions is possible, such that a visual and auditive impression of the scene can be provided.

In the following we introduce the hardware and software parts of the VR system, subsequent to a short overview of the basics of stereoscopic rendering.

6.1 Stereoscopic rendering

Stereoscopic display systems are usually classified into two basic categories: time parallel and time multiplexed stereo systems [Hod92]. Time multiplexed stereoscopic systems (see figure 6.2 (a)) alternate the images for the right and for the left eye on one display. In order to provide a stereoscopic view of the scene to an observer, a shutter system for synchronization of the left and of the right views is needed. The shutter ensures that the right eye is occluded if the left view is shown and vice versa. Usually the observer wears active shutter-glasses.
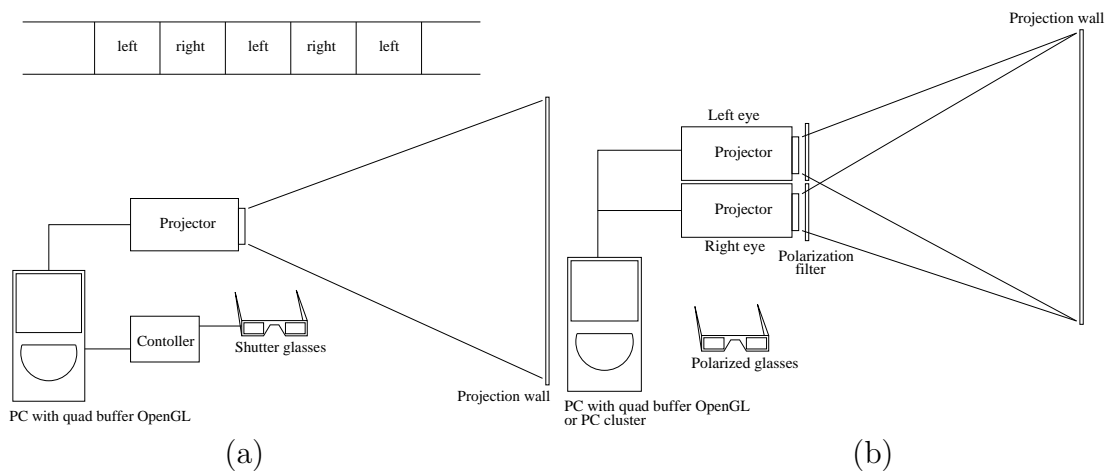
Figure 6.2: Stereoscopic displays. (a) Time multiplexed stereoscopic system and (b) time parallel stereoscopic system.

Time parallel stereoscopic systems present the views of the left eye and of the right eye simultaneously on one or two displays. In order to observe the scene one needs special glasses, so that each eye sees only the corresponding perspective view. Examples for such systems are anaglyph displays, viewer displays, head-mounted stereoscopic displays and polarized time parallel displays. The latter is the most commonly used system.

One of the possible configuration of a polarized time parallel display is as follows (see figure 6.2 (b)). Polarization filters arranged at right angles to each other are positioned before two projectors for right and left views. The observer wears eye-glasses with polarization filters in order to separate the views. A passive VR-System based on circular polarization filters (see section 6.2) has been chosen since more users can observe the display simultaneously, compared to active system, and the representation of the scene is still colorfast in contrast to anaglyph displays, for example.
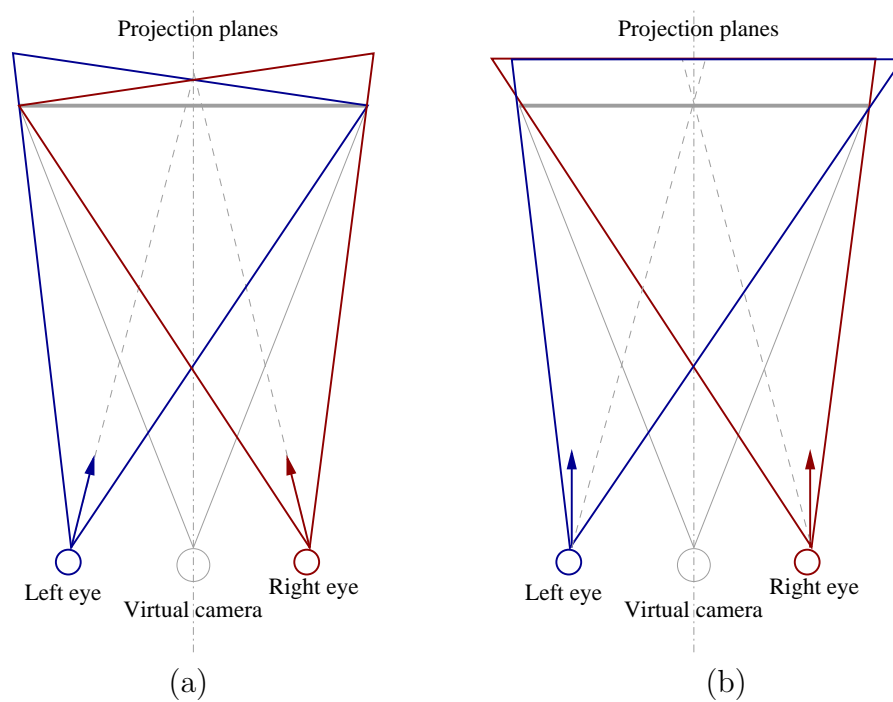
Figure 6.3: Stereoscopic projections. On-axis approach (a) and off-axis approach (b).

Basically, there are two possibilities to compute the projections of the both views for stereoscopic rendering [Hod92, Bou99]. The first approach is the "on-axis" method or the "toe-in" method. Here the cameras for the right and for the left eye have one center of projection (see figure 6.3 (a)). This technique is simple to

implement with OpenGL by using *gluPerspective*, but it is geometrically incorrect, and leads to vertical parallaxes between two homologous points [HMD97].

The correct approach, the "off-axis" method or "parallel axis asymmetric frustum perspective projection" (see figure 6.3 (b)). In this method the right and the left view directions are parallel to each other, i.e. there are two different focal points. Here we have a wider field of view than is obtained by a single perspective image computed with the same parameters, because it contains three regions (see figure 6.3 (b)): one stereoscopic seen by both eyes, and two monoscopic seen only by one eye.

In order to acquire the stereoscopic view the "off-axis" approach is used [Bou99]. This method is based on OpenGL using *glFrustum(left, right, bottom, top, near, far)* perspective projection. The parameters *left*, *right*, *bottom*, *top*, *near* and *far* for the right eye view and accordingly for the left eye view are calculated as follows:

$$
\begin{aligned}
near &= camera.near \\
far &= camera.far \\
top &= near \cdot \tan(\frac{camera.aperture}{2}) \\
bottom &= -top \\
left &= -ratio \cdot top \pm 0.5 \cdot eyesep \cdot \frac{near}{camera.focallength} \\
right &= ratio \cdot top \pm 0.5 \cdot eyesep \cdot \frac{near}{camera.focallength}
\end{aligned}
$$

where $ratio = width/height$ is the ratio of the width and the height of the window (mostly 4/3) and *eyesep* is the eye separation. Our hardware consists of two render nodes, as described in the next section. The first render node applies the parameter for the right eye and the second for the left eye.

6.2 Hardware

The hardware of the Virtual Reality system includes the 3D back projection system and the acoustic system (see figure 6.4). The 3D back projection system is composed of two high resolution digital D-ILA projectors for displaying mono and passive stereo signals, two circular mechanical shiftable polarization filters, and one projection wall, suitable for polarization filters. By means of this system a stereoscopic rendering of virtual scenes is possible, observed by wearing polarizing eyeglasses. The acoustic system contains a 7.1 soundblaster and a professional surround sound
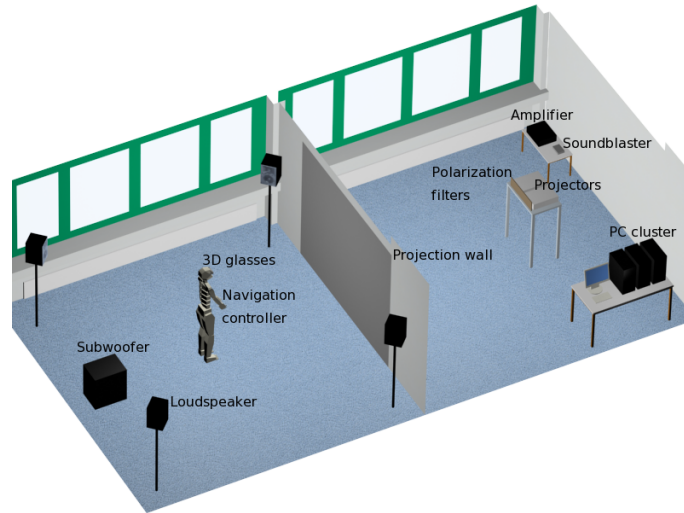
Figure 6.4: Visual and audio Virtual Reality display system.

equipment. This surround sound system on its part includes four two-way loud-speakers and one subwoofer facilitating the localization of the virtual sound source. The system is driven by the PC-cluster composed of three connected DELL computers. Two of them are the render nodes, and one acts as master controlling the navigation and audio-visual output. Technical specifications of the hardware of the Virtual Reality system are summarized in table 6.1. To improve sound synthesis acoustic absorption plates are attached inside of the Virtual Reality labor reducing the intensity of sound reflections off the walls, ceiling, and windows.

3D Powerwall	
2 projectors	1400x1050, 1500 ANSI Lumen
1 screen	plastic film suitable for polarization filters (2.88m x 2.30m)
2 filters	circular polarization filters
glasses	plastic glasses with circular polarization inserts
4.1 Acoustic system	
1 sound card	7.1 USB sound blaster
4 loudspeakers	two-way universal loudspeakers
1 subwoofer	active subwoofer with integrated amplifier
1 amplifier	four-channel amplifier with integrated programmable FIR Filters
PC cluster	
1 master	Dual Xeon 3.0 GHz, 2GB RAM, PCI-Express, NVidia Quadro FX1300
2 render nodes	Single Xeon 3.0 GHz, 2GB RAM, PCI-Express, NVidia Quadro FX5400

Table 6.1: Technical Details of the VR System

6.3 Implementation

The software application for the visual stereoscopic and acoustic rendering of the simulation results on the hardware described above is implemented in C++ using QT[1] and OpenGL[2] APIs for GUI programming and rendering. Particularly, we used QT's client-server concept for controlling the graphical representation. As noted in the previous section, both render nodes are used for drawing the view for the left and right eye view, respectively. To ensure that both rendered images are displayed simultaneously for a flicker-free representation of the scene, synchronization of the render nodes is required. Some APIs for parallel rendering already exist, for example Chromium[3], which is a system for interactive rendering on clusters.
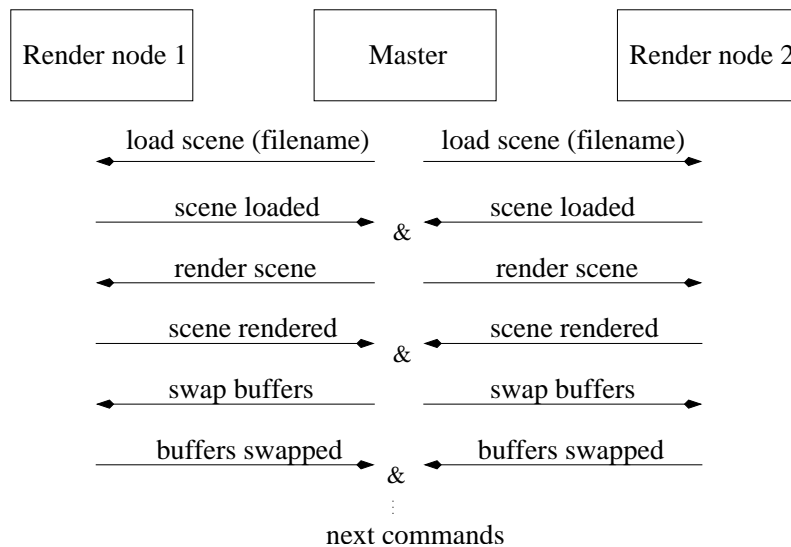
Figure 6.5: Full hand-shake synchronization for rendering

To avoid restrictions imposed by existing synchronization software packages, a simple method based on sockets and full hand-shake synchronization is implemented. Nevertheless, since pure OpenGL is used for graphic rendering the application of Chromium or another VR API's (e.g. VR Juggler[4], DIVERSE[5]) in order to be flexible to use another VR hardware configuration is straightforward. The full hand-shake synchronization is realized as follows. The master sends command messages to the render nodes. When both renderer replied on a successful execution of the

---

[1]http://trolltech.com/

[2]http://www.opengl.org/

[3]http://chromium.sourceforge.net/

[4]http://www.vrjuggler.org

[5]http://diverse-vr.org/

instruction, a new command can be send. In figure 6.5 the loading process of a room geometry from a file is conceptually shown. The master sends the command for loading the scene from file (*load scene*). After the affirmative reply from both render nodes, instruction for rendering follows (*render scene*). Is this successfully executed and confirmed, the buffers can be swapped now for the final representation of the left eye and right eye views on the display. Before the master can send new instructions to render nodes, it needs to wait for the response to the previous message, even in the case of single operations like swapping buffers. Otherwise, two consecutive messages sent by the master may collide causing a serious time delay.

Due to additional overhead for synchronization of the rendering hardware the following approach for acceleration of the visualization of sound propagation described in section 5.2.1 is proposed. First, the traversed time is subdivided into n discrete intervals, where each part holds all phonons to display at this point in time. Thus a look-up table for each time point t_i, which holds the entries with informations to the corresponding phonons, is constructed. One such entry contains the position ph_{pt} of the phonon and its color coded spectral energy c_{ph} at the current time point. While rendering the scene at t_i we have then only to show at the right position in the look-up table and to draw colored spheres, triangles, or surfels, corresponding to phonons positions and color code.

For auralization purposes the synthesis method developed at the Fraunhofer ITWM institute is applied. The algorithm is described in the next section. In order to drive the loudspeaker of the audio system separately with the corresponding signals the PortAudio library[6] for acoustical output is utilized. The audio playback is administrated by the master computer and is started in a new process (*thread*).

6.4 Acoustic synthesis

There are two major sound synthesis approaches: wave field synthesis [BdVV93] and binaural synthesis [ALKV04]. The latter approach uses either headphones or loudspeakers utilizing the cross-talk cancellation [LB04] to reproduce the sound field simulated at the ears of the virtual listener. The advantages are low cost equipment and low number of channels to be computed. However, the listener needs to wear either headphones or a tracking device. The sound field is only reconstructed for two positions inside the ears, filtered by head-related transfer functions (HRTFs) obtained from psycho-acoustic measurements. For a greater range of accuracy wave field synthesis is applied. The underlying theory is based on Huygens' principle implying that solutions of the Helmholtz equation correspond to distributions of sound sources on the surface of the domain (Kirchhoff-Helmholtz integral, [Hac89]).

---

[6]www.portaudio.com
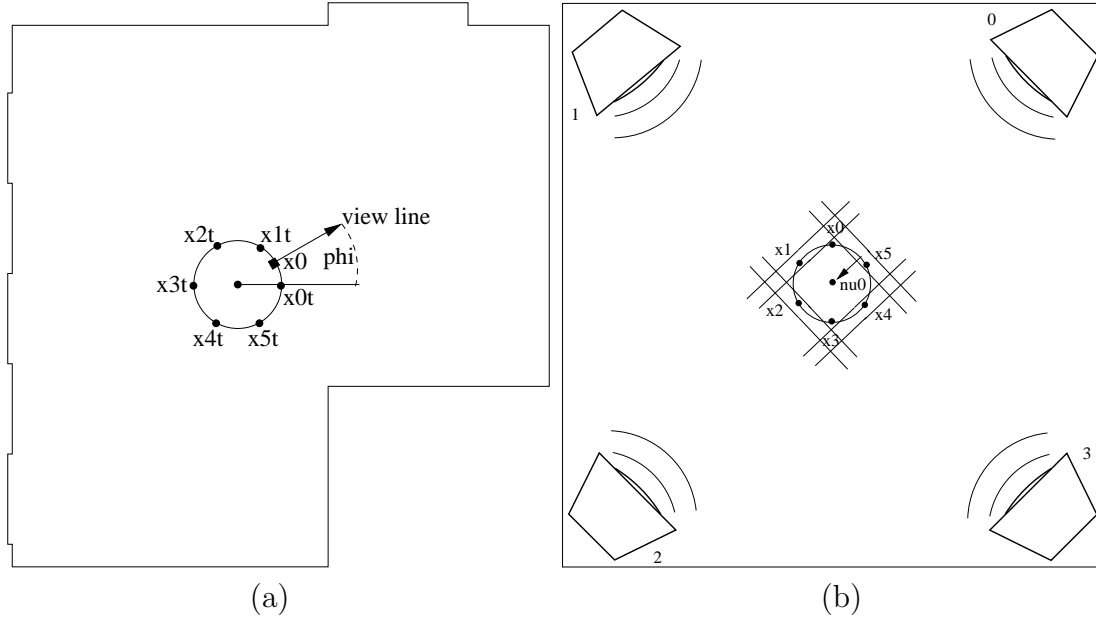
Figure 6.6: Virtual (a) and projection room (b).

This continuous distribution is approximated by a finite number of loudspeakers. The higher the number of speakers and the lower the frequency the greater is the sweet spot where the approximation is accurate. Usually, only the sound field of a 2D listening plane is synthesized as this requires already quite large number of speakers. For instance, Fraunhofer IDMT furnished a cinema with 192 speakers to achieve good reproduction at all seats [KS03]. For our purpose it is enough to have a good reproduction close to the proband's head the position of which is, except for small movements, assumed to be fixed. The present system comprises currently 4 full-band speakers and a bass. In a future system additional speakers will be added.

Let $\tilde{p}_i \in \mathbb{R}^{2n_{\mathrm{b}}}$ a section of the digital signal simulated in point $\tilde{x}_i \in \mathbb{R}^2$ of the horizontal measuring plane cutting the virtual room. n_{b} is the block length and is choosen as a power of 2 to speed up later Fourier transforms by means of the Fast Fourier transform. The n_{p} points are distributed equally on a small circle of radius r about the virtual head position. The radius has to be of the same order as the wave length corresponding to the highest frequency to be reproduced:

$$r \approx \lambda_{\min} = \frac{c_0}{f_{\max}} \; . \tag{6.1}$$

For $f_{\max}$= 20,000 Hz and a sound velocity of c_0= 343 m/s the radius is only 12 mm. The exact radius depends on the number of loudspeakers n_{l}. To get an idea, imagine that an arbitrary sound field of frequencies up to $f_{\max}$ has to be approximated by a polynomial with n_{l} coefficients which, for small n_{l}, will be valid only within a small

radius. Choosing r this small, we avoid artefacts in reproducing the high frequency part of the sound field while at lower frequencies the approximation will still be good also for larger radii.

The system for sound field synthesis is made to reproduce sound in n_p fixed positions on a ring of radius r about a point in the projection room. As the proband may change the line of vision in the virtual room these points cannot be directly associated with the n_p points from the fixed grid where sound is simulated. Hence, the simulated samples $\tilde{p}_{i,n}$ of time step n are interpolated and evaluated at the rotated real positions x_i using cubic splines in the angle.

The resulting signals $p_i \in \mathbb{R}^{2n_\mathrm{b}}$ are first windowed by a lifted cosine and then subjected to a fast Fourier transform:

$$\hat{p}_i = \mathrm{fft}(p_i^w), \quad p_{i,n}^w = \tfrac{1}{2}\left[1 - \cos\tfrac{\pi(n+\frac{1}{2})}{n_\mathrm{b}}\right]p_{i,n} \ . \tag{6.2}$$

For synthesizing the sound field, we assume that the loudspeakers induce plane waves in the central listening position and that time delay and amplification are the same for all loudspeakers. These conditions are met fairly well as we use an anechoic projection room and the speakers are furnished with FIR filters and placed on a circle. However, in a later version the plane-wave assumption will be replaced by using measured impulse responses. So far, the Fourier coefficients $\hat{q}_{j,k}$ of the signal of the j-th speaker are computed to satisfy:

$$\hat{p}_{i,k} = \sum_{j=1}^{n_\mathrm{l}} \hat{q}_{j,k} \exp\left(-\mathrm{i}\frac{\omega_k}{c_0}\langle \nu_j, x_i\rangle\right) \ . \tag{6.3}$$

$\omega_k = \frac{\pi f_\mathrm{s}}{n_\mathrm{b}}k$ is the k-th angular frequency, f_s the sampling rate, and ν_j the radiation direction of the j-th speaker. This leads to a linear system for each frequency:

$$A_k\,\hat{q}_{.k} = \hat{p}_{.k} \ . \tag{6.4}$$

For stability reasons the number n_p of reproduction points is chosen greater than the number n_l of speakers. Thus, equation (6.4) becomes overdetermined and is approximated by a least squares fit. Moreover, in order to keep speaker signals bounded we introduce a small regularization parameter α:

$$\hat{q}_{.k} = T_k\hat{p}_{.k}, \quad T_k = [A_k^*A_k + \alpha\mathrm{I}]^{-1}A_k^* \ . \tag{6.5}$$

The matrices T_k are computed once for each frequency when initializing the system. Since the speaker signals are real we have $\hat{q}_{.,k} = \hat{q}_{.,n_\mathrm{b}-k}^*$ and only half of the systems need to be solved. Next, the $\hat{q}_j$ are retransformed into time domain by the inverse Fourier transform to get q_j^w. The superscript w indicates that these speaker

signals will reproduce only the windowed measurements. To get the final speaker signals, the first half of $q_j^{w,\mathrm{m}}$ is added to the last half of the $q_j^{w,\mathrm{m}-1}$ computed for the last block:

$$q_j = q_{j,1\ldots n_\mathrm{b}}^{w,\mathrm{m}} + q_{j,n_\mathrm{b}+1\ldots 2n_\mathrm{b}}^{w,\mathrm{m}-1} \ . \tag{6.6}$$

Finally, $q_{j,n_\mathrm{b}+1\ldots 2n_\mathrm{b}}^{w,\mathrm{m}}$ is stored for adding in the next step. The following table summarizes the algorithm:

1. *Get next block of simulated measurements.*

2. *Per time step: interpolate simulated measurements at reproduction points of projection room.*

3. *Scale measurements in projection room by lifted cosine-function.*

4. *Fast Fourier transform of scaled measurements in projection room.*

5. *For each frequency: compute loudspeaker signals to optimally reproduce measurements in a regularized least squares sense.*

6. *Inverse fast Fourier transform of loudspeaker signals.*

7. *Add first half of loudspeaker signals to last half stored for last block.*

8. *Store last half of loudspeaker signals to add in the next step.*

9. *Goto 1.*

For a fast convolution of the impulse responses simulated using the phonon tracing method the overlap-add algorithm is used (*see* apendix B.4). Currently, six simulation positions around the virtual listener are calculated in order to drive the four loudspeakers (see figure 6.6). The synthesis algorithm and the six time convolution of the RIRs can be performed in real time only separately, thus the convolution is done in a preprocessing step, by the current implementation.

7 Summary and future work

In this thesis a novel geometric acoustic simulation approach, phonon tracing, was presented for the auralization and visualization of room acoustics. The algorithm facilitates both linear and quadratic attenuation for the simulation of energy and pressure, respectively. Considering sound energy, quadratic attenuation was used because of the spatial particle density. In case of sound pressure simulation, Gaussian basis functions were used with approximate partition of unity to represent wave fronts. This is obtained by constructing the basis functions on the unit sphere around a sound source and dilating them with the traversed distance of the associated phonons. The results obtained by the presented simulation method compare favorably to existing methods. Phonon tracing is a two stage method. In the first stage, the phonon emission stage, sound particles (phonons) are emitted from the sound source and traced through the room populating the so called phonon map. Thereby specular as well as diffuse reflections are considered. In the second stage, the phonon collection stage, particles stored in the phonon map are collected in order to calculate the impulse response for a given listener position. Using the impulse response acoustic parameters can be derived objectively judging the room acoustics. Furthermore, convolution of the impulse response with an anechoic signal results in a audible prediction (auralization) of the sound received by the listener.

Additionally, the collection step of the phonon tracing algorithm can be executed at different precision levels, depending on user requirements of computation time and accuracy. Thus a faster computation of the impulse response can be performed if lower accuracy is required, e.g. to get a first quick impression of the acoustics of the scene. To achieve this, the phonons in the map are grouped into clusters of similar trajectory representing different wave fronts and decreasing memory requirements of the phonon map. Since the phonon tracing algorithm is separated into to stages, changing the listener position does not require the re-execution of the tracing process. However, if the source is changed the emission step needs to be performed, again.

The information stored in the phonon is also used for sound visualization providing a method for acoustic analysis inside the room. Various visualization approaches containing the visualization of sound wave propagation inside the room as well as the visualization of sound arriving at the listener were proposed. Also, several applications were discussed, including a visual analysis method which examine the acoustics of the room according to an acoustic parameter.

The phonon tracing algorithm and several visualization approaches were integrated in a Virtual Reality frame work allowing immersive exploration of the acoustic of a virtual scene. This frame work results as a prototype within a project at the University of Kaiserslautern in collaboration with the Fraunhofer ITWM Kaiserslautern.

Some interesting and challenging tasks are still open. In the algorithm introduced in this thesis, only static sound sources are considered. Integration of moving sources would be an improvement of the method. Furthermore, integration of the diffraction effects would enhance the computation accuracy of the resulting impulse responses. While these effects are small at higher frequencies, they can be simulated at low frequencies by a wave-based method (FEM). More comparisons between phonon tracing and FEM need to be performed in order to find out when to use the individual methods. Another challenging task is the modeling of a room with given acoustic conditions. How would the room be shaped satisfying "good" acoustics for a speech presentation or a music performance. The challenge for such an inversion problem is to find an appropriate formal description how the room shape influences the given acoustic parameters. This enables optimization algorithms for automatic room improvement. As mentioned before, the Virtual Reality frame work is a prototype, thus further research needs to be done to improve it, regarding for example, the acoustic synthesis and the integration of innovative human machine interfaces for a better and more intuitive navigation in the virtual environment.

A Acoustic quantities and their relations

The following definitions and equations are taken from [Rie82].

A.1 Sound

Sound is defined as the mechanical oscillations of an elastic medium. The oscillations occur if the medium particles (molecules) are disequilibrated due to an external force. The particles swing periodically around their equilibrium position. The incidence of sound is associated with the existence of a medium. Sound can not propagate in vacuum. There are three types of sound depending on the transmitting medium:

1. *Air-borne sound*: sound in air and other gases as medium

2. *Solid-borne sound*: sound in solids as medium

3. *Fluid-borne sound*: sound in fluids as medium

Sound propagates in a wave shape (longitudinal wave) with a certain velocity the *sound velocity*:

$$c = \frac{\omega}{k} = \frac{2\pi f}{k} = \lambda f \quad \left[\frac{\mathrm{m}}{\mathrm{s}}\right] \tag{A.1}$$

where f is the *frequency*, λ is the *wave length*, $\omega = 2\pi f$ the *angular frequency*, and $k = 2\pi/\lambda$ the *wave number*. Sound velocity depends on the medium properties. In gases for example the density varies with the temperature and thus the sound velocity differs due to temperature changes. Human ears can perceive sound within a frequency range between 20 Hz and 20 kHz, called *audible sound*. Sound below this range is called *infrasonic*, and above *ultrasound*. A surface holding the wave points with equal delay time from the sound source is called the *wave front*. In case of a sine-shaped wave it is the surface of constant phase. The space area covered by the sound waves is called the *wave field*. It is described by use of physical sound field parameters.

A.2 Sound pressure and pressure level

The basic quantity for sound field description is the *sound pressure p*. The unit of measurement of p is Pascal ($1\text{Pa} = 1\text{N/m}^2$). The sound pressure is the pressure alteration of the local stationary pressure of the medium. Thus, the sound pressure is a local time dependent alternating quantity. The pressure which can just be noticed by human hearing is called *hearing threshold* and is in air determined to $p_0 = 20\mu\text{Pa}$. Sound pressure of about 20 Pa is displeasing and can causes pain, this is the so called *threshold of pain*. The region in-between of this extrema is the *hearing area*. Since the loudness sensation of human being is nearly logarithmic (Weber-Fecher-law), a logarithmic metric is chosen to measure the sound pressure, the *sound pressure level*:

$$L_p = 20 \lg \frac{p}{p_0} \quad [\text{dB}] \tag{A.2}$$

such that $L_p = 0$ dB corresponds to the hearing threshold.

A.3 Sound particle velocity and velocity level

The *sound particle velocity v* in [m/s] is the alternating oscillation velocity of a particle of the sound transmitting medium. Likewise the sound pressure, the particle velocity is an alternating quantity describing the sound field. In a plane wave the sound pressure and the particle velocity are in phase leading to sound energy propagation. In near field of a spherical wave are the pressure and particle velocity dephased (there is no energy propagation), whereas in the far field the both quantities are again in phase. The particle velocity is like other velocities a vector. Instead of indication of all three direction only one direction, the "important" one, is used. The range of the particle velocity occurring in practice spans from $v_0 = 5 \cdot 10^{-6}$ [cm/s] at hearing threshold to $v = 25$ [cm/s] at threshold of pain. v is measured on a logarithmic scale by the *particle velocity level*, as:

$$L_v = 20 \lg \frac{v}{v_0} \quad [\text{dB}] \tag{A.3}$$

A.4 Specific acoustic impedance

The *specific acoustic impedance* is the ratio between the sound pressure and the particle velocity:

$$Z_s = \frac{p}{v} \quad \left[\frac{\text{Ns}}{\text{m}^3}\right] \tag{A.4}$$

In the sound field of a planar wave this ratio is constant at every point in time and at every position of the field, well this fact approximately holds in the far field of an arbitrary sound emitter. The acoustic impedance can then be calculated as a product of medium density and the speed of sound. It is called the *characteristic acoustic impedance*:

$$Z_0 = \rho \cdot c \quad \left[\frac{\text{Ns}}{\text{m}^3}\right] \tag{A.5}$$

In a sound field of a spherical wave this ratio is not constant. The specific acoustical impedance depends on the wave length λ and the distance r from the source. Already in the distance $r = \lambda$, Z_s differs only slightly form Z_0. At $r > \lambda$ Z_s becomes real and is identical to Z_0. Otherwise Z_s is complex. The specific impedance describes the acoustical behavior of a medium completely.

A.5 Sound energy quantities

The sum of the potential and kinetic energy of a sound field is the *sound energy*:

$$W = W_{pot} + W_{kin} \quad [\text{J}] \tag{A.6}$$

Mostly, the *sound energy density w*, which is defined as the energy per volume unit, is used instead. Sound energy density is an appropriate metric describing the amount of sound energy at a certain position of the field. In a progressive wave it is associated to the sound velocity and to the sound intensity:

$$w = \frac{I}{c} \quad \left[\frac{\text{J}}{\text{m}^3}\right] \tag{A.7}$$

Sound intensity I describes the portion of sound energy per time passing through a surface perpendicular to the sound propagation direction. In the field of a plane

wave it is the product of the pressure and particle velocity:

$$I = p \cdot v \quad \left[\frac{\mathrm{W}}{\mathrm{m}^2}\right] \tag{A.8}$$

In the filed of a spherical wave we have to take the phase shift φ between the pressure and the particle velocity. The intensity then reads:

$$I = p \cdot v \cdot \cos(\varphi) \quad \left[\frac{\mathrm{W}}{\mathrm{m}^2}\right] \tag{A.9}$$

The sound intensity indicates the energy content of a sound field and is given by the *sound intensity level*:

$$L_I = 10 \, \lg \frac{I}{I_0} \quad [\mathrm{dB}] \tag{A.10}$$

where $I_0 = 10^{-12}$ W/m^2 is the reference intensity at hearing threshold.

The total energy per time unit emitted by a sound source in all directions is called the *sound power* P. The sound power passing through a imaginary surface S can be calculated as follows:

$$P = \int_S I_n \mathrm{d}S \quad [\mathrm{W}] \tag{A.11}$$

where I_n is the normal component of the sound intensity I perpendicular to S. The sound power is an important parameter for noise measuring and is given by the *sound power level*:

$$L_P = 10 \, \lg \frac{P}{P_0} \quad [\mathrm{dB}] \tag{A.12}$$

where $P_0 = 10^{-12}$ W is the reference power.

B Some basics on digital signal processing

This is a short collection of definitions, principles, and algorithms used in this dissertation. Fore more details on digital signal processing see [PM07, OS99, Hof98].

B.1 Signals and systems

A *signal* is defined as a function representing the information about state or behavior of any physical quantity and is mathematically described as a function of one or more variables. The variables can be either continuous or discrete. Audio signals are expressed as functions of time (*time domain*) or frequency (*frequency domain*). *Time-continuous signals* are defined in a continuous domain, whereas *time-discrete signals* are defined at uniform positions. *Analog signals* are signals which have continuous domain and amplitude, whereas *digital signal* have both discrete domain and amplitude. In computer-aided signal processing we treat digital signals. Digital signals are represented as a sequence of digits $x = \{x[n]\}$, $-\infty < n < \infty$. The most important digital signals are:

- *Unit impulse*:

$$\delta[n] = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n \neq 0 \end{cases} \tag{B.1}$$

  It corresponds to the analog *Dirac-delta impulse*:

$$\delta(t) = \begin{cases} \infty & \text{if } t = 0 \\ 0 & \text{else} \end{cases} \tag{B.2}$$

  with $\int_{-\infty}^{\infty} \delta(t) = 1$

- *Unit step*

$$\sigma[n] = \begin{cases} 0 & \text{if } n < 0 \\ 1 & \text{if } n \geq 0 \end{cases} \tag{B.3}$$

- *Digital sine signal*

$$x[n] = \hat{x} \, \sin(\omega n + \varphi) \tag{B.4}$$

with amplitude $\hat{x}$, angular frequency ω and phase shift φ .

Digital signals can originate from sampling of analog signals or by accumulating a variable over a period of time. For a lossless reconstruction of the analog signal the following must hold.

Nyquist–Shannon sampling theorem:

> If the analog signal $x_a(t)$ is band-limited, highest frequency $f = f_{max}$, and the signal is sampled at a rate $f_s > 2f_{max}$, then $x_a(t)$ can be exactly reconstructed using the sample values. $f_N = 2f_{max}$ is called the *Nyquist rate*.

A system can be defined as a physical module performing operations on signals, e.g. filters used for noise reduction. In signal processing an input signal is transferred (processed) from source to sink via a transfer system (see figure B.1). $H(\omega)$ is the *transfer function* in frequency domain and $f(t)$ is the *impulse response* in time domain of the system. The output signal is the result of the convolution (defined in section B.4) of the input signal and the impulse response, or the frequency spectrum of the output signal is the product of the frequency spectrum of the input signal and the transfer function (which represents the frequency spectrum of the impulse response).
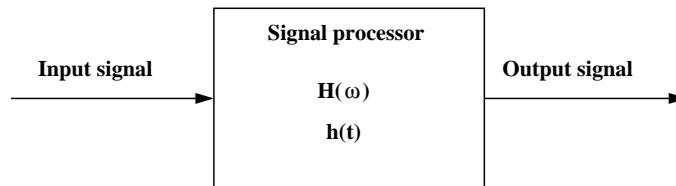
Figure B.1: Block diagram of a signal processing system.

Time-discrete systems can be either *static* or *dynamic*, *time-variant* or *time-invariant*, *linear* or *non-linear*, *causal* or *non-causal*, *stable* or *non-stable*. For definition of the different system properties see [PM07, OS99].

B.2 Fourier analysis

The spectrum of a continuous-time periodic function $x(t)$ with the period time $T = 1/f$ can be described using the *Fourier-series*. The Fourier series in complex representation reads:

$$x(t) = \sum_{k=-\infty}^{\infty} X_k e^{jk\omega_0 t} \tag{B.5}$$

where $\omega_0 = 2\pi/T$ is the angular frequency depending on the period time T, and $j = \sqrt{-1}$ the imaginary unit. The *Fourier-coefficients* X_k can be calculated as follows:

$$X_k = \frac{1}{T} \int_T x(t) e^{-jk\omega_0 t} \mathrm{d}t \tag{B.6}$$

The spectrum of a continuous-time aperiodic function $x(t)$ is called the *Fourier-integral* and reads:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) e^{j\omega t} \mathrm{d}\omega \tag{B.7}$$

with

$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} \mathrm{d}t \tag{B.8}$$

$X(\omega)$ is the *Fourier transform* of $x(t)$. The Fourier integral in equation B.7 is called the *inverse Fourier transform*.

Of course, the spectrum exists only if some conditions, the *Dirichlet conditions* are fulfilled (see [PM07]).

The Fourier series of a discrete-time periodic sequence reads:

$$x(n) = \sum_{k=0}^{N-1} c_k e^{j2\pi kn/N} \quad n = 0, ..., N-1 \tag{B.9}$$

with Fourier-coefficients

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N} \quad k = 0, ..., N-1 \tag{B.10}$$

For a discrete-time aperiodic sequence the Fourier transform pair, the inverse (equation. B.11) and the direct (equation B.12) transform, is expressed as follows:

$$x(n) = \frac{1}{2\pi} \int_0^{2\pi} X(\omega)e^{j\omega n}\mathrm{d}\omega \tag{B.11}$$

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \tag{B.12}$$

where for the existence of the Fourier series must hold:

$$\sum_{n=-\infty}^{\infty} |x(n)| < \infty \tag{B.13}$$

Using the continuous transformation $X(\omega)$ of the sequence $\{x(n)\}$ of the length N in equation B.12 the *discrete Fourier transformation* (DFT) is defined as a finite sequence $\{X(k)\}$ of sampling values of $X(\omega)$. The DFT and the IDFT (*inverse discrete Fourier transform*) reads:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} \quad k = 0, ..., N-1 \tag{B.14}$$

$$x(n) = \frac{1}{N} \sum_{n=0}^{N-1} X(k)e^{j2\pi kn/N} \quad n = 0, ..., N-1 \tag{B.15}$$

The DFT is an important tool in digital signal processing and can be calculated efficiently using the *Fast Fourier Transformation* (FFT) defined by Cooley and Tukey (see [PM07]).

B.3 Digital filter

In audio-engineering, filters are used to lock certain frequency parts, whereas to amplify other frequency parts of the sound signal. The room impulse response in room acoustics can also be considered as a filter changing the structure of an anechoic source signal. There are analog and digital filters, where in computer-aided acoustic design we deal with the latter one.

There are two classes of filter, the *recursive* or infinite impulse response filter (IIR-Filter) and the *finite response filter* (FIR-Filter). The output value of the IIR-Filter

depends not only on the input values, but also on the preceding output values:

$$y(n) = F[y(n-1), y(n-2), ..., x(n), x(n-1), ...] \tag{B.16}$$

Because of this back coupling the impulse response of the filter becomes infinite.

In contrast, the FIR-filter is non-recursive and therefor has a finite impulse response:

$$y(n) = F[x(n), x(n-1), ...] \tag{B.17}$$

Filters are also classified according to their frequency-domain characteristics as *low-pass*, *high-pass*, *band-pass*, *band-stop*, or *all-pass* filters.

B.4 Fast convolution: the overlap-add-algorithm

The behavior of a system (e.g. a filter relating source and listener considering the room specific properties) can be applied using the convolution. The convolution of the input signal with the systems impulse response results the output signal. The continuous convolution of two functions f and g is defined as follows:

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)\mathrm{d}\tau \tag{B.18}$$

The discrete convolution reads:

$$(f * g)(n\Delta t) = \sum_{k=-\infty}^{\infty} f(k\Delta t)g((k-n)\Delta t) \tag{B.19}$$

The convolution can be computed efficiently applying the Fourier transform, since the convolution in time domain corresponds to the multiplication in frequency domain, and vice versa. It holds for two impulse responses f and g and their transfer functions F and G:

$$\begin{aligned} f * g &= F \cdot G \\ f \cdot g &= F * G \end{aligned} \tag{B.20}$$

In practical application the input sequence $x(n)$ can often be very long. One method for efficient computation of the convolution of $x(n)$ with an impulse response $h(n)$ is the *overlap-add* method on a block-by-block basis using the DFT [PM07]. The input sequence $x(n)$ is segmented into blocks and each block is processed via

the DFT and IDFT to produce the output signal sequence $y(n)$. In the following the steps of the overlap-add algorithm are described for $h(n)$ of the duration M, input data block of the size L, with the DFTs and IDFTs for size $N = L + M - 1$ (in order to speed up the Fourier transformation N is chosen to the power of 2):

1. *Get input data block of size L from the input sequence $x(n)$*

2. *Append $M-1$ zeros to the block:* $x_i(n) = \{x((i-1)L, ..., x(iL-1), \underbrace{0, ..., 0}_{M-1\ zeros} \} i = 1, ..., K$(number of blocks)

3. *Compute the N-point DFT of the block $x_i(n) \Longrightarrow X_i(k)$*

4. *Multiply X_i with the transfer function $H(k)$ of the impulse response $h(n)$:* $Y_i(k) = H(k) \cdot X_i(k)$

5. *Perform the IDFT of $Y_i(k) \Longrightarrow y_i(n)$*

6. *Overlap and add the first $M - 1$ points to the last $M - 1$ points of y_{i-1}*

7. *Goto 1.*

The algorithm results in the following output sequence:

$$y(n) = \{y_1(0), y_1(1), \quad ..., y_1(L-1), y_1(L) + y_2(0), y_1(L+1)+ \\ y_2(1), ..., y_1(N-1) + y_2(M-1), y_2(M), ...\} \tag{B.21}$$

A modification of the overlap-add algorithm in order to consider long impulse responses (filters) $h(n)$, where $h(n)$ is also segmented into blocks is described in [Zöl97].

C Photon mapping

Photon mapping is a two pass method for computing global illumination in computer generated scenes proposed by Jensen [Jen96, JC98, Jen01]. The first pass is the constructing of the photon map by emitting photons from the light source and store them in the scene as they hit surfaces. This information can be seen as a rough representation of the light within the model. in the second step the photon map is used to render the final image.

C.1 Pass 1: Photon emission

Each photon emitted from the source is trace through the scene. Every time a photon interacts with a surface inside the scene it is stored in the photon map. Jensen uses Russian roulette technique to determine whether the photon is absorbed or reflected. The new direction of the photon is computed by use of the BRDF of the surface. Jensen uses two photon maps, one for caustics computation (caustics photon map) and one as the approximation of the light within the scene (global photon map). Compact and efficient storing of the photons is done by use of a balanced kd-tree. The rendering task is perform more efficiently due to the kd-tree structure.

C.2 Pass 2: Rendering

The final image is rendered by use of Monte Carlo ray tracing exploiting the photon map. The surface radiance $L_s(x, \Psi_r)$ leaving the surface at the intersection point x in the direction Ψ_r can be calculated using the rendering equation:

$$L_s(x, \Psi_r) = L_e(x, \Psi_r) + \int_\Omega f_r(x, \Psi_i, \Psi_r) L_i(x, \Psi_i) \cos\theta_i \ \mathrm{d}\omega_i \qquad \text{(C.1)}$$

where L_e is the radiance emitted by the surface itself, L_i is the incident radiance in the direction Ψ_i, f_r is the BRDF of the surface, and Ω is the sphere of incident directions. L_e is given by the surface definition, the residual of the equation can be

expressed as a sum of direct, specular, and diffuse illumination:

$$
\begin{aligned}
L_r \;=\;& \int_\Omega f_r L_{i,l}\cos\theta_i \; \mathrm{d}\omega_i + \\[6pt]
& \int_\Omega f_{r,s}(L_{i,c} + L_{i,d})\cos\theta_i \mathrm{d}\omega_i + \\[6pt]
& \int_\Omega f_{r,d} L_{i,c}\cos\theta_i \mathrm{d}\omega_i + \\[6pt]
& \int_\Omega f_{r,d} L_{i,d}\cos\theta_i \mathrm{d}\omega_i
\end{aligned}
\tag{C.2}
$$

where

$$
f_r = f_{r,s} + f_{r,d} \quad \text{and} \quad L_i = L_{i,l} + L_{i,c} + L_{i,d}
$$

Thereby, $L_{i,l}$ is the contribution from the light source, $L_{i,c}$ contributions via specular reflections, and $L_{i,d}$ that via diffuse reflections. The BRDF f_r is separated into a specular $f_{r,s}$ and a diffuse $f_{r,d}$ part.

Direct illumination. The first term in equation C.2 express the direct illumination by the light sources. This term is computed, depending on the application, either approximately or accurate. For accurate evaluation extended photon maps are used as described in [JC95]. The approximated solution is the radiance estimate determined using the global photon map.

Specular reflection. The second term in C.2 corresponds to the radiance reflected of specular and highly glossy surfaces. This term is calculated by used of standard Monte Carlo ray tracing.

Caustics. Caustics on diffuse and slightly glossy surfaces (third term in C.2) are computed using the caustics photon map as describe further in the text. Since the caustics photon map is visualized directly the number of photons in it must be high.

Diffuse reflections. The diffuse reflections represented by the fourth term in C.2 are estimated using the global photon map as described above.

Estimating radiance using the photon map. The radiance L_r leaving the surface at the intersection point x can be estimated using the photon map as follows. First N photons with the shortest distance to x are located. The radiance can then be

approximated as:

$$
\begin{aligned}
L_r(x, \Psi_r) &= \int_\Omega f_r(x, \Psi_i, \Psi_r)\frac{\mathrm{d}^2\Phi_i(x,\Psi_i)}{\mathrm{d}A\omega_\mathrm{i}}\mathrm{d}\omega_i \\
&\approx \sum_{p=1}^{N} f_r(x, \Psi_r, \Psi i, p)\frac{\Delta\Phi_p(x,\Psi_{i,p})}{\pi r^2}
\end{aligned}
\tag{C.3}
$$

The above approximation is possible under the assumption that each photon p represents flux $\Delta\Phi_p$ arriving at x from the direction $\Psi_{i,p}$. The surface ΔA around x is estimated using sphere centered in x. The radius r of the sphere is expanded until it contains N photons. ΔA is then assigned to the area πr^2.

Bibliography

[AB79]   J.B. Allen and A. Berkeley. Image method for efficiently simulating small-room acoustics. *J. Acoust. So. Amer.*, 65(4):943–950, Apr. 1979.

[AFST04a]   F. Antonacci, M. Foco, A. Sarti, and S. Tubaro. Accurate and fast audio-realistic rendering of sounds in virtual environments. In *Proceedings of the 6th IEEE Workshop on Multimedia Signal Processing*, pages 271–274, Siena, Italy, September–October 2004.

[AFST04b]   F. Antonacci, M. Foco, A. Sarti, and S. Tubaro. Real time modeling of acoustic propagation in complex environments. In *Proceedings of the 12th European Signal Processing Conference (EUSIPCO '04)*, pages 1773–1776, Vienna, Austria, September 2004.

[AFST04c]   F. Antonacci, M. Foco, A. Sarti, and S. Tubaro. Real time modeling of acoustic propagation in complex environments. In *Proceedings of the 7th International Conference on Digital Audio Effects (DAFX-04)*, pages 274–279, Naples, Italy, October 5–8 2004.

[ALKV04]   I. Assenmacher, T. Lentz, T. Kuhlen, and M. Vorländer. Integrating real-time binaural acoustics into VR applications. In *Eurographics Symposium on Virtual Environments*, pages 129–136, Grenoble, June 8–9 2004.

[Ama84]   J. Amanatides. Ray tracing with cones. *Computer Graphics*, 18(3):129–135, July 1984.

[AS01]   A.C. Antoulas and D.C. Sorensen. Approximation of large-scale dynamical systems: An overview. Technical report, Rice University, 2001.

[Bar71]   M. Baron. The subjective effects of first reflections in concert halls-the need for lateral reflections. *Journal of Sound and Vibrations*, 15(4):475–494, 1971.

[Bar00]   M. Baron. Measured early lateral energy fractions in concert halls and opera houses. *Journal of Sound and Vibrations*, 232(1):79–100, 2000.

[Bat67]   D.W. Batteau. The role of the pinna in human localization. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 168(1011):158–180, 1967.

[BDM+05] M. Bertram, E. Deines, J. Mohring, J. Jegorovs, and H. Hagen. Phonon tracing for auralization and visualization of sound. In *IEEE Visualization 2005*, pages 151–158, Minneapolis, MN, USA, October 23–28 2005.

[BdVV93] A.J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *J. Acoust. Soc. Am.*, 93(5):2764–2778, 1993.

[Beg00]   D.R. Begault. *3-D Sound for Virtual Reality and Multimedia*. National Aeronautics and Space Administration, Ames Research Center, Moffett Field, CA, 2000.

[Ber04]   L.L. Beranek. *Concert halls and opera houses : music, acoustics, and architecture*. Springer, New York, 2004. 2. ed., rev. ed.

[BK03]   M. Botsch and L. Kobbelt. High-quality point-based rendering on modern gpus. In *Proceedings of Pacific Graphics*, pages 335–343, Canmore, Alberta, Canada, October 8–10 2003.

[Bla74]   J. Blauert. *Räumliches Hören*. S. Hirzel Verlag, Stuttgart, 1974.

[Bla97]   J. Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, Mass., 1997.

[BM81]   M. Baron and A.H. Marsha. Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure. *Journal of Sound and Vibrations*, 77(2):211–232, 1981.

[Bor84]   J. Borish. Extension of the image model to arbitrary polyhedra. *J. Acoust. So. Amer.*, 75(6):1827–1836, 1984.

[Bor02]   I. Bork. Simulation and measurement of auditorium acoustics - the round robins on room acoustical simulation. *Proceedings of the Institute of Acoustics*, 24(4), Juli 2002. Presented at the Auditorium Acoustics Conference, London.

[Bor05a] I. Bork. Report on the 3rd round robin on room acoustical computer simulation - part i: Meassurements. *Acta Acustica united with Acustica*, 91:740–752, 2005.

[Bor05b] I. Bork. Report on the 3rd round robin on room acoustical computer simulation - part ii: Calculations. *Acta Acustica united with Acustica*, 91:753–763, 2005.

[Bos]   Bose Corporation. Bose modeler. *http://www.bose.com*.

[Bot95]   D. Botteldooren. Finite-difference time-domain simulation of low-frequency room acoustic problems. *Journal of the Acoustical Society of America*, 98(6):3302–3308, December 1995.

[Bou99]   P. Bourke. 3d stereo rendering using opengl (and glut). *http://astronomy.swin.edu.au/ pbourke/opengl/stereogl/*, 1999.

[Bra03] D. Braess. *Finite Elmente: Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie.* Springer, Berlin, 2003. 3., korrigierte und ergänzte Auflage.

[BS65] L.L. Beranek and Th.J. Schultz. Some recent experiences in the design and testing of concert halls with suspended panel arrays. *Acustica*, 15:307–316, 1965.

[BS95a] R.E. Berg and D.G. Stork, editors. *The physics of sound.* Prentice Hall, Englewood Cliffs, New Jersey, 1995. Second edition.

[BS95b] J.S. Bradley and G.A. Soulodre. The influence of late arriving energy on spatial impression. *Journal of the Acoustical Society of America*, 97(4):2263–2271, 1995.

[BSZ90] H.E. Bass, L.C. Sutherland, and A.J. Zuckerwar. Athmospheric absorption of sound: Update. *Journal of Acoustic Society of America*, 88(4):2019–2021, 1990.

[BT57] M. A. Biot and I. Tolstoy. Formulation of wave propagation in infinite media by normal coordinates with an application to diffraction. *Journal of the Acoustical Society of America*, 29(3):381–391, March 1957.

[CAT] CATT. Catt-acoustic. *http://www.catt.se.*

[CG85] M. F. Cohen and D. P. Greenberg. The semi-cube a radiosity solution for complex environments. *Computer Graphics*, 19(3):31–40, July 1985.

[CM78a] L. Cremer and H.A. Müller. *Die wissenschaftlichen Grundlagen der Raumakustik, Band I.* S. Hirzel Verlag Stuttgart, 1978. 2. völlig neubearbeitete Auflage.

[CM78b] L. Cremer and H.A. Müller. *Die wissenschaftlichen Grundlagen der Raumakustik, Band II.* S. Hirzel Verlag Stuttgart, 1978. 2. völlig neubearbeitete Auflage.

[Cow94] J. P. Cowan. *Handbook of environmental acoustics.* John Wiley & sons, Inc., New York, 1994.

[CR05] C. L. Christensen and J. H. Rindel. A new scattering method that combines roughness and diffraction effects. In *Proceedings of Forum Acusticum*, pages 1173–1179, Budapest, Hungary, August 29 – September 2 2005.

[CS07] P. T. Calamia and U. P. Svensson. Fast time-domain edge-diffraction calculations for interactive acoustic simulations. *EURASIP Journal on Advances in Signal Processing*, 2007:Article ID 63650, 10 pages, 2007. doi:10.1155/2007/63650.

[CSF05] P. T. Calamia, U. P. Svensson, and T. A. Funkhouser. Integration of edge-diffraction calculations and geometrical-acoustics modeling. In

Proceedings of Forum Acusticum, Budapest, Hungary, August 29 – September 2 2005.

[CW99] C.I. Cheng and G.H. Wakefield. Introduction to head-related transfer functions (hrtf's): Representations of hrtf's in time, frequency, and space. In *Proceedings of the 107th Convention of the Audio Engineering Society*, New York, USA, September 1999.

[DA72] P. Damaske and Y. Ando. Interaural crosscorrelation for multichannel loudspeaker reproduction. *Acustica*, 27:232–238, 1972.

[Dal96] B.-I. Dalenbäck. Room acoustic prediction based on a unified treatment of diffuse and specular reflection. *Journal of the Acoustical Society of America*, 100(2):899–909, 1996.

[DBM⁺06] E. Deines, M. Bertram, J. Mohring, J. Jegorovs, F. Michel, H. Hagen, and G.M. Nielson. Comparative visualization for wave-based and geometric acoustics. In *IEEE Visualization 2006*, pages 1173–1179, Baltimore, Maryland, USA, October 29 – November 3 2006.

[Dev86] M Devroye. *Non uniform random variate generation*. Springer, New York, 1986.

[DIN00] Deutsches Institut für Normung e.V. DIN. Messung der nachhallzeit von räumen mit hinweis auf andere akustische parameter. *DIN EN ISO 3382:2000*, 2000.

[DKW85] N. Dadoun, D. G. Kirkpatrick, and J. P. Walsh. The geometry of beam tracing. In *Proceedings of the first annual symposium on Computational geometry*, pages 55–61, 1985.

[DL00] I. A. Drumm and Y. W. Lam. The adaptive beam-tracing algorithm. *Journal of the Acoustical Society of America*, 107(3):1405–1412, March 2000.

[DMB⁺06] E. Deines, F. Michel, M. Bertram, H. Hagen, and G. Nielson. Visualizing the phonon map. In *Eurographics / IEEE-VGTC Symposium on Visualization*, Lisbon, Portugal, Mai 8-10 2006.

[Dud94] R.O. Duda. Modeling head related transfer functions. In *Proceedings of the Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, pages 457–461, Asilomar, CA, October 1994.

[EAS01] J. J. Embrechts, D. Archambeau, and G.B. Stan. Determination of the scattering coefficient of random rough diffusing surfaces from room acoustics applications. *Acusica*, 87:482–494, 2001.

[Est07] O. von Estorff. Numerical methods in acoustics: Facts, fears, future. In *19th International Congress on Acoustics (ICA2007)*, Madrid, Spain, September 2–7 2007.

[Far95] A. Farina. Ramsete - a new pyramid tracer for medium and large scale acoustic problems. In *Proceedings of EURO-NOISE 95 Conference*, March 1995.

[FCE⁺98] T. A. Funkhouser, I. Carlbom, G. Elko, G. P. M. Sondhi, and J. West. A beam tracing approach to acoustic modeling for interactive virtual environments. In *Computer Graphics (SIGGRAPH 98)*, pages 21–32, Orlando, FL, July 1998.

[FKS84a] W. Fasold, W. Kraak, and W. Schirmer, editors. *Taschenbuch Akustik. Teil 1*. VEB Verlag Technik, Berlin, 1984.

[FKS84b] W. Fasold, W. Kraak, and W. Schirmer, editors. *Taschenbuch Akustik. Teil 2*. VEB Verlag Technik, Berlin, 1984.

[FMC99] T. A. Funkhouser, P. Min, and I. Carlbom. Real-time acoustic modeling for distributed virtual environments. In *Computer Graphics (SIGGRAPH 99)*, pages 365–374, Los Angeles, August 1999.

[FSO06] Y. Fukushima, H. Suzuki, and A. Omoto. Visualization of reflected sound in enclosed space by sound intensity measurement. *Acoust. Sci. & Tech.*, 27(3):187–189, 2006.

[FTC⁺04] T. Funkhouser, N. Tsingos, I. Carlbom, G. Elko, M. Sondhi, J. E. West, G. Pingali, P. Min, and A. Ngan. A beam tracing method for interactive architectural acoustics. *Journal of the Acoustical Society of America*, 115(2):739–756, February 2004.

[Gel04] S.A. Gelfand. *Hearing: An Introduction to Psychological and Physiological Acoustics*. Marcel Dekker, New York, 2004. Fourth edition, revised and expand.

[GM96] W. G. Gardner and K. D. Martin. Hrtf measurements of a kemar dummy head microphone. *IEEE CS Technical Committee on Computer Generated Music*, 1996.

[Gol97] E.B. Goldstein. *Wahrnehmungssysteme: Eine Einführung*. Spektrum, Akad. Verl., Heidelberg, 1997.

[Gör06] T. Görne. *Tontechnik*. Fachbuchverlag Leipzig im Carl Hanser Verlag, 2006.

[Gri91] E. Grimsehl, editor. *Lehrbuch der Physik. Band 1. Mechanik - Akustik - Wärmelehre*. B.G. Teubner Verlagsgesellschaft KG, Leipzig, 1991. 27. Auflage.

[GTGB84] C. M. Goral, K. E. Torrance, D. P. Greenberg, and B. Battaile. Modeling the interaction of light between diffuse surfaces. *Computer Graphics*, 18(3):213–222, July 1984.

[Hac89] W. Hackbusch. *Integralgleichungen, Theorie und Numerik*. Teubner Verlag, 1989.

[Har99] W.M. Hartmann. How we localize sound. *Physics today*, pages 24–29, 1999.

[Hei93] R. Heinz. Binaural room simulation based on an image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the raverberant tail. *Applied Acoustics*, 38:145–159, 1993.

[HH84] P. S. Heckbert and P. Hanrahan. Beam tracing polygonal objects. *Computer Graphics*, 18(3):119–127, July 1984.

[HMD97] L. Harrison, D. McAllister, and M. Dulberg. Stereo computer graphics for virtual reality. In *ACM SIGGRAPH '97*, 1997. Course Notes 6.

[Hod92] L. F. Hodges. Tutorial: Time-multiplexes stereoscopic computer graphics. In *IEEE Computer Graphics & Applications*, pages 20–30, March 1992.

[Hof98] R. Hoffmann. *Signalanalyse und -erkennung. Eine Einfürung für Informationstechniker*. Springer, 1998.

[HS64] J.H. Halton and G.B. Smith. Algorithm 247: Radical-inverse quasi-random point sequnce. *Communications of the ACM*, 7(12):701–702, December 1964.

[HS73] T. Houtgast and H.J.M. Steeneken. The modulation transfer function in room acoustics as a predictor of speech intelligibility. *Acustica*, 28:66–73, 1973.

[HS84] T. Houtgast and H.J.M. Steeneken. A multi-language evaluation of the rasti-method for estimation speech intelligibility in auditoria. *Acustica*, 54:185–199, 1984.

[Ihl98] F. Ihlenburg. *Finite element analysis of acoustic scattering*. Springer, New York, 1998.

[ISO97] International Organisation of Standardization ISO. Acoustics – measurement of the reverberation time of rooms with reference to other acoustical parameters. *ISO 3382:1997*, 1997.

[JC95] H.W. Jensen and N.J. Christensen. Efficiently rendering shadows using the photon map. In *Proceedings of Compugraphics 95'*, pages 285–291, 1995.

[JC98] H.W. Jensen and P.H. Christensen. Efficient simulation of light transport in scene with participating media using photon maps. In *Computer Graphics (SIGGRAPH 98)*, pages 311–320, July 1998.

[Jen96] H.W. Jensen. Global illumination using photon maps. In *Rendering Techniques '96 (Proceedings of the 7th Eurographics Workshop on Rendering)*, pages 21–30, 1996.

[Jen01] H.W. Jensen. *Realistic Image Synthesis Using Photon Mapping*. AK Peters, 2001.

[JK04] M. Jedrzejewski and M. Krzysztof. Computation of room acoustics using programable video hardware. In *International Conference on Computer Vision and Graphics ICCVG'2004*, Warsaw, Poland, September 22-24 2004.

[Jor70] V.L. Jordan. Acoustical criteria for auditoriums and their relation to model techniques. *Journal of the Acoustical Society of America*, 2(1):408–412, 1970.

[Kaj86] J. T. Kajiya. The rendering equation. *Computer Graphics*, 20(4):143–150, August 1986.

[Kap06] B. Kapralos. *The Sonel Mapping. Acoustical Modeling Method*. PhD thesis, York University, Toronto, Ontario, Canada, 2006.

[KB04] L. Kobbelt and M. Botsch. A survey of pointbased techniques in computer graphics. *Computers & Graphics*, 28(6), 2004.

[KBAA01] N. Korany, J. Blauer, and O. Abdel Alim. Acoustic simulation of rooms with boundaries of partially specular reflectivity. *Applied Acoustics*, 62:875–887, 2001.

[Kel62] J. B. Keller. Geometrical theory of diffraction. *Journal of the Optical Society of America*, 52(2):116–130, February 1962.

[Kel96a] A. Keller. Quasi monte carlo methods in computer graphics: The global illumination problem. *Lecture in App. Math.*, 32:455–469, 1996.

[Kel96b] A. Keller. Quasi monte carlo radiosity. *X. Pueyo and P. Schröder, editors, Rendering Techniques '96 (Proc. 7th Eurographics Workshop on Rendering)*, pages 101–110, 1996.

[KFW98] S. Khoury, A. Freed, and D. Wessel. Volumetric visualization of acoustic fields in cnmat's sound spatialization theatre. In *Visualization '98*, pages 439–442 & 562. IEEE, 1998.

[KJM04] B. Kapralos, M. Jenkin, and E. Millios. Sonel mapping: Acoustic modeling utilizing an acoustic version of photon mapping. In *IEEE International Workshop on Haptics Audio Visual Environments and their Applications (HAVE 2004)*, Ottawa, Canada, October 2–3 2004.

[KJM05a] B. Kapralos, M. Jenkin, and E. Millios. Acoustical diffraction modeling utilizing the huygens-fresnel principle. In *IEEE International Work-*

shop on Haptics Audio Visual Environments and their Applications (HAVE 2005), Ottawa, Canada, October 1–2 2005.

[KJM05b] B. Kapralos, M. Jenkin, and E. Millios. Acoustical modeling using a russian roulette strategy. In *118th Convention of Audio Engineering Society*, Barcelona, Spain, May 28–31 2005.

[KK85] H.W. Knobloch and H. Kwakernaak. *Lineare Kontrolltheorie.* Springer-Verlag, 1985.

[KP74] R. G. Kouyoumjian and P. H. Pathak. A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface. *Proceedings of the IEEE*, 62(11):1448–1461, November 1974.

[Kro68] U. Krockstadt. Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibrations*, 8(18):118–125, 1968.

[KS03] B. Klehs and T. Sporer. Wave field synthesis in the real world part 1: In the living room. In *Proceedings of the 114th Audio Engineering Society (AES) Convention*, Amsterdam, March 22 - 25 2003.

[Kul84] U. Kulowski. Algorithmic representation of the ray tracing technique. *Applied Acoustics*, 18:449–469, 1984.

[Kür69] R. Kürer. Zur gewinnung von einzahlkriterien bei impulsantworten in der raumakustik. *Acustica*, 21:370–372, 1969.

[Kut71] H. Kuttruff. Simulierte nachhallkurven in rechtecksräumen mit diffusem schallfeld. *Acustica*, 25:333–342, 1971.

[Kut93] H. Kuttruff. Auralization of impulse responses modeled on the basis of ray-tracing results. *Journal of Audio Engineering Society*, 41(11):876–879, November 1993.

[Kut00] H. Kuttruff. *Room acoustics.* Spon Press, London, 2000. 4th. edition.

[KW92] D.J. Kistler and F.L. Wightman. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *Journal of the Acoustical Society of America*, 91(3):1637–1647, 1992.

[LB58] J.P.A. Lochner and J.F. Burger. The subjective masking of short time delayed echoes by their primary sounds and their contribution to the intelligibility of speech. *Acustica*, 8:1–10, 1958.

[LB61] J.P.A. Lochner and J.F. Burger. The intelligibility of speech under reverberant conditions. *Acustica*, 11:195–200, 1961.

[LB04] T. Lentz and G. Behler. Dynamic cross-talk cancelation for binaural synthesis in virtual environments. In *Proceedings of the 117th Au-*

dio Engineering Society (AES) Convention, San Francisco, CA, USA, 2004.

[LCM07a] C. Lauterbach, A. Chandak, and D. Manocha. Adaptive sampling for frustum-based sound propagation in complex and dynamic environments. In *19th International Congress on Acoustics (ICA2007)*, Madrid, Spain, September 2–7 2007.

[LCM07b] C. Lauterbach, A. Chandak, and D. Manocha. Interactive sound propagation in dynamic scenes using frustum tracing. In *In Proceeding of IEEE Visualization*, Sacramento, CA, USA, October 28 – November 1 2007.

[Lew93] T. Lewers. A combined beam tracing and radiant exchange computer model of room acoustics. *Applied Acoustics*, 38:161–178, 1993.

[LN06] T. Lokki and V. Nenonen. Imersive visualization of room acoustics. In *Joint Baltic-Nordic Acoustics Meeting*, Gothenburg, Sweden, November 8–10 2006.

[Lok02] T. Lokki. *Physically-based Auralization*. PhD thesis, Helsinki University of Technology, 2002.

[LRC⁺03] D. Luebke, M. Reddy, J. Cohen, A. Varshney, B. Watson, and R. Huebner. *Level of Detail for 3D Graphics*. The Morgan Kaufmann Series in Computer Graphics, Elsevier Science, San Frncisco, CA, USA, 2003.

[LSS02] T. Lokki, P. Svensson, and L. Savioja. An efficient auralization of edge diffraction. In *Audio Engineering Society, 21th International Conference*, pages 166–172, St. Petersburg, Russia, June 2002.

[LSV⁺02] T. Lokki, L. Savioja, R. Väänänen, J. Huopaniemi, and T. Takala. Creating interactive virtual auditory environments. *IEEE Computer Graphics and Applications*, 22(4):49–57, July 2002.

[LSVA07] T. Lentz, D. Schröder, M. Vorländer, and I. Assenmacher. Virtual reality system with integrated sound field simulation and reproduction. *EURASIP Journal on Advances in Signal Processing*, 2007:Article ID 70540, 19 pages, 2007. doi:10.1155/2007/70540.

[MCJ82] H. Medwin, E. Childs, and G. M. Jebsen. Impulse studies of double diffraction: A discrete huygens interpretation. *Journal of the Acoustical Society of America*, 72(3):1005–1013, September 1982.

[MDHB⁺07] F. Michel, E. Deines, M. Hering-Bertram, C. Garth, and H. Hagen. Listener-based analysis of surface importance for acoustic metrics. In *IEEE Visualization 2007*, pages 1680–1687, Sacramento, CA, USA, October 28 – November 1 2007.

[Med81] H. Medwin. Shadowing by finite noise barriers. *Journal of the Acoustical Society of America*, 69(4):1060–1064, April 1981.

[MEG74] E. Meyer and Neumann E-G. *Physikalische und Technische Akustik*. Fiedr. Vieweg + Sohn, Braunschweig, 1974. 2., berichtigte und erweiterte Auflage.

[MLPK01] J. Merimaa, T. Lokki, T. Peltonen, and M. Karjalainen. Measurements, analysis, and visualization of directional room responses. In *Proceedings of the 111th Audio Engineering Society (AES) Convention*, New York, NY, USA, September 21–24 2001.

[MM93] D. van Maercke and J. Martin. The prediction of echograms and impulse responses within the epidaure software. *Applied Acoustics*, 38:93–114, 1993.

[MM02] E.A. Macpherson and J.C. Middlebrooks. Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *Journal of the Acoustical Society of America*, 111(5):2219–2236, 2002.

[MM04] G. Müller and M. Möser, editors. *Taschenbuch der Technischen Akustik*. Springer, 2004. 3., erweiterte und überarbeitete Auflage.

[MOD00] M. Monks, B.M. Oh, and J. Dorsey. Audioptimization: Goal-based acoustic design. *IEEE Computer Graphics and Applications*, 20(3):76–91, 2000.

[Møl92] H. Møller. Fundamentals of binaural technology. *Applied Acoustics*, 36:171–218, 1992.

[Nay93] G. M. Naylor. Odeon – another hybrid room acoustical model. *Applied Acoustics*, 38:131–143, 1993.

[NMI04] E.M. Nosal, Hodgson M., and Ashdown I. Improved algorithms and methods for room sound-field prediction by acoustical radiosity in arbitrary polyhedral rooms. *Journal of the Acoustical Society of America*, 116(2):970–980, 2004.

[NSG02] M. Naef, O. Staadt, and M. Gross. Spatialized audio rendering for immersive virtual environments. In *Symposium on Virtual Reality Software and Technology*, pages 65–72. ACM, November 2002.

[ODE] ODEON. Room acoustic software. *http://www.odeon.dk*.

[Ols02] K.H.A. Olsson. Model order reduction in FEMLAB by dual rational Arnoldi. Master's thesis, Department of Mathematics, Chalmers University of Technology and Göteborg University, 2002.

[OS99] A.V. Oppenheim and R.W. Schafer. *Zeitdiskrete Signalverarbeitung*. R. Oldenburg Verlag, München, 1999. 3., durchgesehene Auflage.

[OU04] A. Omoto and H. Uchida. Evaluation method of artificial acoustical environment: Visualization of sound intensity. *Journal of Physiological Anthropology and Applied Human Science*, 23:249–253, 2004.

[PH04] M. Pharr and G. Humphreys. *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann, 2004.

[PL03] V. Pulkki and T. Lokki. Visualization of edge diffraction. *Acoustics Research Letters Online (ARLO)*, 4(4):118–123, 2003.

[PLS02] V. Pulkki, T. Lokki, and L. Savioja. Implementation and visualization of edge diffraction with image-source method. In *Proceedings of the 112th Audio Engineering Society (AES) Convention*, Munich, Germany, May 2002.

[PM07] J.G. Proakis and D.G. Manolakis. *Digital Signal Processing. Principles, Algorithms, and Applications*. Prentice Hall, New Jersey, fourth edition, 2007.

[PR05] S. Petrausch and R. Rabenstein. Highly efficient simulation and visualization of acoustic wave fields with the functional transformation method. In *Simulation and Visualization*, pages 279–290, Otto von Guericke Universität, Magdeburg, March 2005.

[PZvBG00] H. Pfister, M. Zwicker, J. van Baar, and M. Gross. Surfels: Surface elements as rendering primitives. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 335–342, New Orleans, LA, USA, 2000.

[RAAS74] W. Reichardt, O. Abdel Alim, and W. Schmidt. Abhängigkeit der grenzen zwischen brauchbarer und unbrauchbarer durchsichtigkeit von der art des musikmotives, der nachhallzeit und der nachhalleinsatzzeit. *Applied Acoustics*, 7:243–264, 1974.

[Rei68] W. Reichardt. *Grundlagen der technischen Akustik*. Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig, 1968.

[Rie82] M.M. Rieländer, editor. *Reallexikon der Akustik*. Verlag Erwin Bochinsky, Frankfurt, Main, 1982.

[RSS97] R. Rabenstein, O. Schips, and A. Stenger. Acoustic rendering of buildings. In *5th Int. IBPSA Conference Building Simulation*, volume 2, pages 181–188, Prag, Czech Republic, September 1997.

[SAWG91] F. X. Sillion, J. R. Arvo, S. H. Westin, and D. P. Greenberg. The semi-cube a radiosity solution for complex environments. *Computer Graphics*, 25(4):187–196, July 1991.

[SC00] B.G. Shinn-Cunningham. Distance cues for virtual auditory space. In *Proceedings of First IEEE Pacific-Rim Conference on Multimedia*, pages 227–230, Sydney, Australia, December 13–15 2000.

[SC06]   U. P. Svensson and P. T. Calamia. Edge-diffraction impulse responses near specular-zone and shadow-zone boundaries. *Acta Acustica united with Acustica*, 92:501–512, 2006.

[Sch65]   M.R. Schroeder. New method of measuring reverberation time. *Journal of the Acoustical Society of America*, 37(3):409–412, 1965.

[Sch79]   M.R. Schroeder. Integrated-impulse method measuring sound decay without using impulses. *Journal of the Acoustical Society of America*, 66(2):497–500, 1979.

[Sch81]   M.R. Schroeder. Modulation transfer functions: Definition and measurment. *Acustica*, 49:179–182, 1981.

[SFV99]   U. P. Svensson, R. I. Fred, and J. Vanderkooy. An analytic secondary source model of edge diffraction impulse responses. *Journal of the Acoustical Society of America*, 106(5):2331–2344, November 1999.

[SG89]   A Stettner and D. P. Greenberg. Computer graphics visualization for acoustic simulation. In *International Conference on Computer Graphics and Interactive Techniques*, pages 195–206. ACM, 1989.

[SHLV99]   L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen. Creating interactive virtual acoustic environments. *Journal of Audio Engineering Society*, 47(9):675–704, September 1999.

[SIM94]   W.H. Slattery III and J.C. Middlebrooks. Monaural sound localization: Acute versus chronic unilateral impairment. *Hearing Research*, 75:38–46, 1994.

[SK62]   M.R. Schroeder and K.H. Kuttruff. On frequency responsecurves in rooms. comparison of experimental, theoretical, and monte carlo results for the average frequency spacing between maxima. *Journal of the Acoustical Society of America*, 34(1):76–80, 1962.

[Sku71]   E. Skudrzyk. *The Foundations of Acoustics. Basic Mathematics and Basic Acoustics*. Springer–Verlag, Wien – New York, 1971.

[SLH02]   L. Savioja, T. Lokki, and J. Huopaniemi. Auralization applying the parametric room acoustic modeling technique - the diva auralization system. In *International Conference on Computer Auditory Display*, pages 219–224, Kyoto, Japan, July 2-5 2002.

[SLR07]   J.W. Strutt (Lord Rayleigh). On our perception of sound direction. *Philosophical Magazine*, 13:214–232, 1907.

[SLS07a]   S. Siltanen, T. Lokki, and L. Savioja. Acoustic radiance transfer method. In *19th International Congress on Acoustics (ICA2007)*, Madrid, Spain, September 2–7 2007.

[SLS07b] S. Siltanen, T. Lokki, and L. Savioja. The room acoustic rendering equation. *Journal of the Acoustical Society of America*, 2007. Accepted for publication.

[SLV02] L. Savioja, T. Lokki, and V. Välimäki. The interpolated 3-d digital waveguide mesh method for room acoustic simulation and auralization. In *Proceedings of the Audio Engineering Society (AES) 104th Convention*, Copenhagen, Denmark, August 26–28 2002.

[SS07] U. M. Stephenson and U. P. Svensson. An improved energetic approach to diffraction based on the uncertainty principle. In *19th International Congress on Acoustics (ICA2007)*, Madrid, Spain, September 2–7 2007.

[ST68] E.A.G. Shaw and R. Teranishi. Sound pressure, generated in an external ear replica and real human ears by a nearby sound source. *Journal of the Acoustical Society of America*, 44(1):240–249, 1968.

[Ste96] U. M. Stephenson. Quantized pyramidal beam tracing - a new algorithm for room acoustics and noise immission prognosis. *Acustica - acta acustica*, 82:517–525, 1996.

[Ste04] U. M. Stephenson. *Beugungssimulation ohne Rechenzeitexplosion: Die Methode der quantisierten Pyramidenstrahlen – ein neues Berechnungsverfahren für Raumakustik und Lärmimmissionsprognose. Vergleiche, Ansätze, Lösungen.* PhD thesis, RWTH Aachen University, 2004.

[SV07] D. Schröder and M. Vorländer. Hybrid method for room acoustic simulation in real-time. In *19th International Congress on Acoustics (ICA2007)*, Madrid, Spain, September 2–7 2007.

[Sve02] U. P. Svensson. Modelling acoustic spaces for audio virtual reality. In *Proceedings of the 1st IEEE Benelux Workshop on Model based Processing and Coding of Audio (MPCA-2002)*, November 15 2002.

[SZ93] J. Shi and A. Zhang. A modified radiosity algorithm for integrated visual and auditory rendering. *Computer & Graphics*, 17(6):633–642, 1993.

[TFNC01] N. Tsingos, T. Funkhouser, A. Ngan, and I. Carlbom. Modeling acoustics in virtual environments using the uniform theory of diffraction. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH 2001)*, pages 545–552, 2001.

[TG97] N. Tsingos and J.-D. Gascuel. Soundtracks for computer animation: Sound rendering in dynamic environments with occlusions. In *Proceedings of the Conference on Graphics Interface*, pages 9–16, Kelowa, British Columbia, Canada, May 1997.

[TG98]   N. Tsingos and J.-D. Gascuel. Fast rendering of sound occlusion and
           diffraction effects for virtual acoustic environment. In *Proceedings of
           the Audio Engineering Society (AES) 104th Convention*, Amsterdam,
           Netherlands, May 1998. preprint no. 4699.

[Thi53]   R. Thiele. Richtungsverteilung und zeitfolge der schallrückwürfe in
           räumen. *Acustica*, 3:291–302, 1953.

[TLDD07a] N. Tsingos, S. Lefebvre, C. Dachsbacher, and M. Dellepiane. Extending
           geometrical acoustics to highly detailed architectural environments. In
           19th International Congress on Acoustics (ICA2007), Madrid, Spain,
           September 2–7 2007.

[TLDD07b] N. Tsingos, S. Lefebvre, C. Dachsbacher, and M. Dellepiane. Instant
           sound scattering. In *Proceedings of the 18th Eurographics Symposium
           on Rendering (EGSR)*, Juli 2007.

[TPT04]   J.C.B. Torres, M.R. Petraglia, and R.A. Tenenbaum. Low-order mod-
           eling of head-related transfer functions using wavelet transforms. In
           Proceedings of the International Symposium on Curcuits and Systems,
           volume 3, pages 513–516, May 23–26 2004.

[TSK01]   R. R. Torres, U. P. Svensson, and M. Kleiner. Computation of edge
           diffraction for more accurate room acoustics auralization. *Journal of
           the Acoustical Society of America*, 109(2):600–610, February 2001.

[TY05]   Y. Tokita and Y. Yamasaki. Visualization of 3-dimensional sound fields
           by numerical solutions of particle displacement. *Acoust. Sci. & Tech.*,
           26(2):215–217, 2005.

[Vor88]   M. Vorländer. Ein strahlenverfolgungs-verfahren zur berechnung von
           schallfeldern in räumen. *Acusica*, 65:138–148, 1988.

[Vor89]   M. Vorländer. Simulation of the transient and steady-state sound prop-
           agation in rooms using a new combined ray-tracing/image-source al-
           gorithm. *J. Acoust. So. Amer.*, 86(1):172–178, 1989.

[WAKW93]   E.M. Wenzel, M. Arruda, D.J. Kistler, and F.L. Wightman. Localiza-
           tion using nonindividualized head-related transfer functions. *Journal
           of the Acoustical Society of America*, 94(1):111–123, 1993.

[WD82]   J.P. Walsh and N. Dadoun. What are we waiting for? the development
           of godot, ii. In *103rd Meeting of the Acoustical Society of America*,
           Chicago, USA, April 1982.

[Wey05a]   S. Weyna. Application of "microflown" probe to visualization of acous-
           tic power flow. In *Polish-Scandinavian Structured Conference on
           Acoustic*, Poznan-Wagrowiec, Poland, September 11–15 2005.

[Wey05b] S. Weyna. Microflown based identification of vortex shadding in the space of real acoustic flow fields. In *Twelfth International Congress on Sound and Vibration*, Lisbon, Portugal, July 11–14 2005.

[WH05] A. Wareing and M. Hodgson. Beam-tracing model for predicting sound fields in rooms with multilayer bounding surfaces. *Journal of the Acoustical Society of America*, 118(4):2321–2331, October 2005.

[WK89] F.L. Wightman and D.J. Kistler. Headphone simulation of free-field listening. i: Stimulus synthesis. *Journal of the Acoustical Society of America*, 85(2):858–867, 1989.

[WK92] F.L. Wightman and D.J. Kistler. The dominant role of low-frequency interaural time differences in sound localization. *Journal of the Acoustical Society of America*, 91(3):1648–1661, 1992.

[Wu00] T. W. Wu, editor. *Boundary element acoustics: Fundamentals and computer codes*. WIT Press, Southhampton, 2000.

[Yam72] K. Yamagushi. Multivariate analysis of subjective and physical measures of hall acoustics. *Journal of the Acoustical Society of America*, 52(5):1271–1279, 1972.

[Yee66] K. S. Yee. Numerical solution of initial boundary value problems involving maxwellâs equations in isotropic media. *IEEE transactions on Antennas and Propagation*, 14(3):302–307, May 1966.

[YST02] T. Yokota, S. Sakamoto, and H. Tachibana. Visualization of sound propagation and scattering in rooms. *Acoust. Sci. & Tech.*, 23(1):40–46, 2002.

[ZD98] K. Zhou and J.C. Doyle. *Essentials of robust control*. Prentice-Hall inc., 1998.

[ZDD04] D. Zotkin, R. Duraiswami, and L. Davis. Rendering localized spatial audio in a virtual auditory space. *IEEE Transactions on Multimedia*, 6(4):553–564, 2004.

[Zöl97] U. Zölzer. *Digitale Audioverarbeitung*. B.G. Teubner, Stuttgart, 1997. 2., durchgesehene Auflage.

[ZPvBG01] M. Zwicker, H. Pfister, J. van Baar, and M. Gross. Surfel splatting. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 371–378, Los Angeles, CA, USA, August 2001.

CURRICULUM VITAE

PERSONAL DETAILS

Name:	Eduard Deines
Day and place of birth:	14.09.1976, Novosibirsk / Russia
Marital status:	Married

EDUCATION

10/2005 - to date	Graduate student at the International Research Training Group (IRTG) "Visualization of Large and Unstructured Data Sets - Applications in Geospatial Planning, Modeling, and Engineering"
10/1997 - 03/2004	Study of Computer Science at the Technical University of Kaiserslautern **Degree:** Diplom-Informatiker (Dipl.-Inform.) **Master Thesis**: "Exploration von 3D Strömungsstrukturen mit Hilfe von Schnittflächentopologie" (Exploration of 3D Flow Structure by Means of Cutting Plane Topology)
08/1993 - 06/1996	Technical Grammar School Kaiserslautern **Certificate:** General Qualification for University Entrance (Abitur)
08/1990 - 06/1993	Secondary School HS Goethe Kaiserslautern **Certificate:** Secondary School Level I
09/1983 - 06/1990	Secondary School No. 88 Barnaul / Russia

WORK EXPERIENCE

05/2004 - 09/2005	Research associate at the Computer Graphics and Visualization Group at the Technical University of Kaiserslautern.