# Higher-order Finite Difference Time Domain Algorithms for Room Acoustic Modelling

Jelle van Mourik

Ph.D.

University of York
Electronics

November, 2016

# Abstract

The acoustic qualities of indoor spaces are fundamental to the intelligibility of speech, the quality of musical performances, and perceived noise levels. Computationally heavy wave-based acoustic modelling algorithms have gained momentum in the field of room acoustic modelling, as ever-increasing computational power makes their use more feasible, even for large-scale room acoustic models. Most notably the Finite Difference Time Domain (FDTD) method is often employed for rendering the low- and mid-frequency part of room impulse responses (RIRs). However, this algorithm has known disadvantages, most prominently dispersion error, which renders a large part of the simulated RIR invalid.

This thesis is concerned with the implementation and analysis of higher-order FDTD stencils as a means to improve the current state-of-art FDTD methods that solve the room acoustic wave equation. A detailed analysis of dispersive properties, stability, and required grid spacing of current and higher-order stencils is presented, and has been verified using a GPU implementation of the several different algorithms. It is argued that, of the analysed stencils, the $4^{\text{th}}$-order stencil gives the best result in terms of output quality versus computational effort. In addition, this thesis focusses on the derivation of absorbing boundaries for the $4^{\text{th}}$-order scheme, its stability analysis, and detailed analysis of absorptive properties compared to established boundary models for $2^{\text{nd}}$-order schemes.

The newly proposed $4^{\text{th}}$-order scheme and its boundaries are tested in two case studies: a large shoebox model, in order to test the validity against a common benchmark, and a complex acoustic space. For the latter study, impulse responses were measured in the National Centre for Early Music in York, UK, and computationally generated using the current state-of-the-art as well as the proposed $4^{\text{th}}$-order FDTD algorithm and boundaries. It is shown that the $4^{\text{th}}$-order stencil gives at least as good as, or better results than those achieved using the $2^{\text{nd}}$-order stencil, at lower computational costs.

# Contents

# List of Figures

# List of Tables

# List of accompanying material

- Higher-order stencils (supporting material for Chapter 5)
  - IRs for room – 1.5m x 1.25m x 1m
  - IRs for room – 5m x 3.5m x 2m
  - IRs for room – 7m x 5m x 4m
  - IRs for room – 16 x 16 x 16 nodes
- Shoebox case study (supporting material for Chapter 7):
  - IRs for 2$^{\text{nd}}$-order accurate method with LRS1 boundaries
  - IRs for 2$^{\text{nd}}$-order accurate method with LRS2 boundaries
  - IRs for 2$^{\text{nd}}$-order accurate method with SAT-SBP boundaries
  - IRs for 4$^{\text{th}}$-order accurate method with SAT-SBP boundaries
  - Matlab code to generate all the above
  - Graphics output that plots the decay graphs and FFT plots for all of the above
- NCEM case study (supporting material for Chapter 7):
  - NCEM Geometry: the Blender and the OBJ file
  - Sound files: drum and speech sounds, dry and convolved with all IRs

**Note:** See C for a more detailed description of the accompanying material.

# Acknowledgments

This thesis would not not have been possible without the help and support of a great number of people, either in advice, in technical support, in encouraging cheers along the way, or in cookies. It would require another chapter to do justice to every single one, for which unfortunately there is space nor time, so a concise summary will have to suffice.

First, I would like to thank my supervisor Dr. Damian Murphy for his continuous expert advice, his patience every time I let myself be side-tracked for the greater benefit of the research, and his guidance when guidance was needed.

A very special mention to my brother Tim van Mourik, for proving himself a great twin by being one of the few people to actually be interested in the technical details of this research, including its Matlab implementations, and for providing invaluable technical aid by means of a great many hours of his lab's supercomputer's CPU time.

Another big thanks to my colleagues at Sony Interactive Entertainment for their support, and for giving me the space to finish my PhD while working in the Creative Services Group.

But most of all, I would like to express my appreciation to my parents, my family, and every single one of my friends, in The Netherlands, in York, in London, in Spain, and wherever else in the world you are. The continuous support and understanding you have shown during my rocky road of research has been invaluable. There have been times in which my PhD had to take precedence over things I would probably have rather done, and I am extremely grateful for the understanding and patience of all of you during these five years. If I had to bake every one of you a cookie, I would be baking a lot of cookies.

# Declaration

I hereby declare that the contents of this thesis are entirely the product of my own work, and that all contributions from outside sources, such as publications, websites, and personal contact, have been explicitly stated and referenced to the best of my ability. Additionally, I declare that part of this work has been previously published and/or presented in papers and at conferences. These publications are the following:

- Jelle van Mourik and Damian T. Murphy. Geometric and wave-based acoustic modelling using Blender. In *Proceedings of the AES 49th International Conference*, Audio for Games, London, UK (February 2013).
- Jelle van Mourik and Damian T. Murphy. Explicit higher-order FDTD schemes for 3D room acoustic simulation. *IEEE Transactions on Acoustics, Speech and Language Processing*, 22(12): pp. 2003–2011, 2014.
- Jelle van Mourik, Stephen Oxnard, Aglaïa Foteinou, and Damian T. Murphy. Hybrid acoustic modelling of historic spaces using blender. In *Proceedings of Forum Acusticum 2014*, Krakow, Poland (November 2014).
- Stephen Oxnard, Dean O'Brian, Jelle van Mourik, and Damian T. Murphy. Frequency-Dependent Absorbing Boundary Implementations in 3D Finite Difference Time Domain Room Acoustics Simulations. In *Proceedings of EuroNoise 2015*, Maastricht, The Netherlands (June 2015).

# Chapter 1

# Introduction

## 1.1 Background

The quality and quantity of sound, music, and noise greatly impact people's lives all over the world, and the medium through which these are delivered plays an important role in humans' perception of these sources. There are aesthetic aspects of sound and music delivery, which are demonstrated in, for example, concert halls and other musical venues. Also intelligibility of speech is greatly influenced by the environment in which it is heard. Additionally, there are significant health aspects to exposure to continuous sound and noise: the World Health Organisation reports that 1.1 billion people are at a risk of hearing loss at some point in their lives [151]. In many risk cases, the acoustic environment heavily influences the actual and perceived noise level, and is therefore as important as a subject of study as the sound sources themselves.

Acoustics is the field of study concerned with sound wave propagation in solid, fluid, and gaseous mediums. Room acoustics studies wave propagation particularly in air in enclosed spaces, such as offices, concert halls, auditoria, classrooms, restaurants, etcetera. Computational room acoustic modelling aims at creating and/or recreating the sound field in such spaces using computer methods, and has many applications in architectural design, music creation and reproduction methods, and computer gaming. This thesis is concerned with all aforementioned applications: its goal is to develop and improve numerical techniques that benefit the modelling of arbitrarily shaped indoor acoustic spaces.

A traditional method for acoustic research in architectural design is using physical scale models of spaces. While this method is still in use by present-day acoustic consultancy companies, it is labour-intensive, and poses many new problems, both practically, in terms of building, storing, and archiving a project, and theoretically, as acoustic properties are not necessarily all scaleable to a model. Computer simulation software provides a more convenient way to run acoustic simulations without the need for expensive equipment and laborious model construction. Whereas early methods mainly focussed on finding echo related parameters (e.g. reverberation time) using statistical methods, computational methods and power have advanced sufficiently to produce full impulse responses within a reasonable time frame. Many acoustic modelling algorithms that produce such an impulse response have been described in literature (see Chapter 4 for a detailed analysis). A very common and useful distinction is the one between *geometric* algorithms and *wave-based* algorithms, which distinguishes the generally faster methods with good high-frequency performance from the typically much slower methods with accurate low-frequency performance. These two types are not mutually exclusive, and can be used in conjunction to produce a *hybrid* solution that combines two or more algorithms, exploiting their respective advantages.

The main body of this thesis studies one particular family of algorithms based on Finite Difference Time Domain (FDTD) modelling [31, 166, 32]. As a wave-based method, it aims to solve the wave equation numerically at discrete intervals in time and space. It has been demonstrated to be very suitable for the modelling of arbitrarily shaped acoustic spaces [103, 18, 164]. The algorithm has been subject to intensive study, not merely in the field of acoustics but also in electromagnetism, geophysics, fluid dynamics, material science, etcetera. Though subject to intensive wider study, the finite difference modelling of the *acoustic wave equation* poses a particular set of problems. For example, while its rigorous approach makes it appropriate for low-frequency wave modelling, the spatial and temporal sampling means that the algorithmic spatial complexity is linear in terms of room volume and modelling frequency [35]. Modelling large environments at full human hearing range (up to ∼20 kHz) is therefore not realistic with current consumer-type computers. Recent research into the FDTD for sound propagation has focussed on several aspects, such as algorithmic improvements to increase the valid bandwidth [103, 81], efficient computer implementations [218, 164], several different reactive boundary models [103, 43, 153], and directional source and receiver implementations [79, 78]. This thesis is primarily concerned with the former three.

## 1.2 Statement of hypothesis

This thesis is motivated by the notion that incremental improvements to acoustic modelling algorithms in both quality and efficiency are beneficial to the field of acoustics, which can in turn positively impact the quality of acoustic designs of concert halls, auditoria, classrooms, public spaces, offices, etcetera. The main focus of this research is to improve one particular kind of wave-based algorithm, the Finite Difference Time Domain method, in order to make it computationally lighter while at the same time increasing the validity of its output, and to verify these results using real-life case studies. The hypothesis tested in this thesis is the following:

> **Hypothesis**
>
> The 'leggy' 4th-order Finite Difference stencil in combination with the presented corresponding boundaries presents an improvement in both output quality and computational speed compared to the currently prevailing Standard Rectilinear and Interpolated Wideband 2nd-order schemes.

Quality will be measured objectively by the amount of dispersion error of the stencils, which gives a good measure for 'exactness' of the stencil. Computational speed can be analysed both on theoretical grounds and in practical simulations.

## 1.3 Contributions of this thesis

Through the research that has been conducted in the exploration of this thesis, the following novel contributions to the field have been made:

- A thorough investigation of the time and memory savings and dispersion error of higher-order (that is, greater than 2nd-order) 'leggy' stencils for the purposes of 3D room acoustic modelling;
- Recommendations based on the above for a new state-of-the-art FDTD method that outperforms current common 2nd-order implementations;
- A new way of graphically representing the quality of a stencil by means of plotting the confidence level versus the valid bandwidth;

- A novel implementation of $4^{\text{th}}$-order accurate absorbing boundaries for the acoustic wave equation;
- Case studies that prove the validity of the aforementioned new methods.

## 1.4 Thesis outline

This thesis is structured as follows:

**Chapter 2** discusses the fundamental mathematical and physical concepts underlying the relevant theories and equations. The wave equation is derived from first principles, several of its solutions are discussed, as well as common acoustic theory that follows from it, such as the inverse square law, Kirchhoff-Helmholtz integral, spherical harmonics, and the acoustic rendering equation.

**Chapter 3** aims to lay the foundation for the acoustics principles relevant to this thesis. The first part deals with physical acoustic theory such as sound propagation in air, wave-boundary interactions such as reflection, diffusion, refraction and absorption. The second part deals with the larger-scale acoustic principles, such as room impulse responses and reverberation. It also discusses perceptually relevant acoustic parameters, most of which are formally defined in ISO 3382 [66].

**Chapter 4** aims to summarise the efforts and progress in the field of room acoustic modelling over the last couple of decades. It is separated into two parts: the geometric acoustic modelling methods, which historically were the earliest to be developed as they are computationally less intensive, and the wave-based acoustic modelling methods, which are more rigorous from a physical point of view, but consequently also heavier in terms of required computational efforts.

**Chapter 5** presents the first novel results of this thesis. It first discusses higher-order stencils, which have been the subject of some prior investigations but, to the author's knowledge, have not been adopted more widely. The contributions of this chapter are a thorough theoretical stability, dispersion, and performance analysis, and a performance analysis of a GPU implementation of the aforementioned stencils.

**Chapter 6** presents the investigation into stable higher-order accurate absorbing boundaries. An introduction is given into several methods that have traditionally been used to determine stability in FDTD simulations of several kinds. It is noted that not all of these are appropriate for the particular case of the acoustic wave equation with impedance boundary. Several attempts at higher-order impedance boundaries are presented. One is presented as a good candidate, and is analysed further in terms of stability, polar eigenvalue analysis, and absorptive properties.

**Chapter 7** aims to verify the newly proposed higher-order stencil and boundary implementation by means of two case studies. The first one deals with the classic benchmark case of a shoebox room, and an extensive analysis is performed on all aspects of the simulated impulse responses. The second case study deals with the acoustics of an actual space, in which impulse responses have been measured previously. The current state-of-art $2^{\text{nd}}$-order FDTD method and the newly proposed $4^{\text{th}}$-order case are used to simulate this space and their results are compared with each other and with the measurements.

**Chapter 8** concludes the thesis by summarising the contributions of its preceding chapters, with particular attention to the chapters 5–7 that present novel work. It reviews the novel work in the light of the hypothesis stated in Sec. 1.2, and suggests several topics for further study in continuation of the research presented in this thesis.

# Chapter 2

# Essential Mathematical and Physical concepts

Wave propagation can be mathematically described by the linear second-order partial differential equation known as the *wave equation*. While many variations are known that describe e.g. electromagnetic or water waves, this thesis is particularly concerned with the *acoustic wave equation*, which describes air waves in the free field. While the wave equation is generally derived using a microscopic basis, it gives rise to many macroscopic acoustic phenomena such as sound diffraction, standing waves, and the inverse square law of sound decay.

This chapter deals with the fundamentals of the mathematics and physics relevant to the rest of this work. The wave equation is introduced and derived in detail in Sec. 2.1. Sec. 2.1.1 goes through its derivation from first principles, and the next sections discuss some of the solutions and alternative formulations of the wave equation. The subsequent sections derive the Kirchhoff-Helmholtz integral, spherical harmonics, and Kirchhoff diffraction, all fundamental constructs for acoustics and acoustic modelling and reproduction methods.

## 2.1 The wave equation

Linear acoustic modelling primarily deals with the propagation of sound in a medium - generally air. The homogeneous wave equation describes simple wave-like motion and can be applied to e.g. air, fluids, strings, electromagnetic fields, etcetera. In the context of this thesis, wave propagation in extended spaces are of interest and thus a derivation of the wave equation in air will follow, using Hooke's laws. Other derivations can be found in e.g. [23, 202]

(for a string) and [159] (for fluids).

### 2.1.1 Derivation

The following derivation of the wave equation draws from the derivations by [23, 64]. Let us consider a tube with cross-section $S$ filled with some gas. We assume that the length of the tube is large compared to its radius, such that we limit ourselves to considering only longitudinal motion. The displacement $u(x,t)$ of the gas particles at a position $x$ along the tube is the net displacement of air along the tube's longitudinal axis at time $t$. If we consider the net displacement a small distance $dx$ away, then by Taylor series expansion we get:

$$u(x+dx,t) = u(x,t) + \frac{\partial u}{\partial x}\, dx + O(x^2). \tag{2.1}$$

The rarefaction, i.e. the lowering in density of the volume between $x$ and $x+dx$ is then:

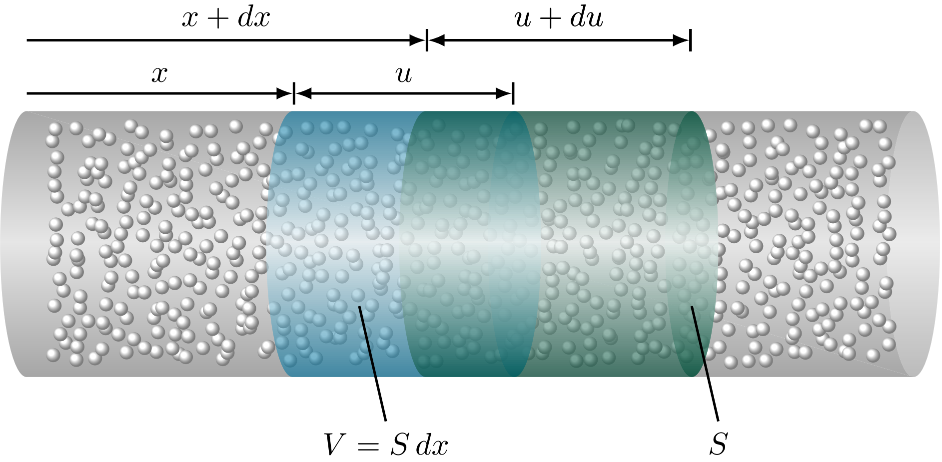


**Figure 2.1:** A tube with cross-section $S$ filled with gas particles. $x$ indicates position along the tube, and $u$ represents the particle displacement.

$$du = \left(u(x,t) + \frac{\partial u}{\partial x} dx\right) - u(x,t) = \frac{\partial u}{\partial x}\, dx, \tag{2.2}$$

which corresponds to the following change in volume of the cross section:

$$dV = S\,\frac{\partial u}{\partial x}\, dx. \tag{2.3}$$

The *bulk modulus* $K$ of an ideal gas is a measure of its resistance to compression and is formally defined as:

$$K = -V \frac{\partial p}{\partial V}, \tag{2.4}$$

where $p$ represents the pressure of the volume $V$. This can be rewritten as:

$$\partial p = -K \frac{\partial V}{V} = -K \frac{S \frac{\partial u}{\partial x} dx}{S\, dx} = -K \frac{\partial u}{\partial x}. \tag{2.5}$$

Hooke's law states that the force $F$ on the surface of a cross-section is proportional to the difference in pressure:

$$\frac{F}{S} = -\partial p = -K \frac{\partial u}{\partial x}, \tag{2.6}$$

and taking the derivative of this with respect to $x$ gives:

$$\frac{\partial p}{\partial x} = -K \frac{\partial^2 u}{\partial x^2}. \tag{2.7}$$

Newton's second law states that a force is equal to the product of the mass and the acceleration of that mass. The mass of a cross-section is its volume times its density $\rho$, and the acceleration is the second time derivative of its displacement.

$$F = \rho V \frac{d^2 u}{dt^2} = \rho\, S\, dx \frac{d^2 u}{dt^2}. \tag{2.8}$$

Combining this result with Eq. 2.6, we get:

$$\begin{aligned} \rho\, S\, dx \frac{d^2 u}{dt^2} &= -S\, \partial p \\ \rho \frac{d^2 u}{dt^2} &= -\frac{\partial p}{\partial x}. \end{aligned} \tag{2.9}$$

Substituting the right-hand side by Eq. 2.7, this results in:

$$\rho \frac{d^2 u}{dt^2} = K \frac{\partial^2 u}{\partial x^2} \tag{2.10}$$

The speed of sound $c$ is in a gas determined by the density $\rho$ and the bulk modulus $K$:

$$c = \sqrt{\frac{K}{\rho}}, \tag{2.11}$$

such that we are left with the homogeneous wave equation:

$$\frac{d^2u}{dt^2} = c^2\,\frac{\partial^2 u}{\partial x^2} \tag{2.12}$$

This equation is the fundamental equation that describes wave propagation in an acoustic medium, and is therefore the basis of many acoustic problems. It is worth noting that it is a reasonably valid approximation at standard temperature and air pressure for low pressure waves, though there are many cases where it breaks down noticeably. High-pressure waves are not modelled adequately, as higher-order terms previously ignored in the Taylor series approximation start to have non-negligible effects. Moreover, Eq. 2.12 conserves all energy in the system, which in reality is not accurate, especially at high frequencies. The *viscous wave equation* is a variation of the wave equation that takes this into account and reads:

$$\frac{d^2u}{dt^2} = c^2\,\frac{\partial^2 u}{\partial x^2} + c\,\alpha\,\frac{\partial^2 u}{\partial x^2}\frac{\partial u}{\partial t}, \tag{2.13}$$

where $\alpha$ is some viscosity constant. No derivation of this will be given but one can be found in e.g. [134]. This equation accounts for air absorption, which increases as frequency increases. Frequency-dependent air absorption will be discussed in more detail in Sec. 3.1.6.

### 2.1.2 Solutions

Solutions to the homogeneous wave equation can be formulated in several ways. The general complex harmonic solution can easily be verified to be:

$$u(x,t) = A\,e^{j(\omega\,t-k\,x)} + B\,e^{j(\omega\,t+k\,x)}, \tag{2.14}$$

where $A, B$ are arbitrary complex numbers, which represent magnitude and phase of the wave. $k = \frac{\omega}{c}$ is generally referred to as the *wave number*, and $\omega$ as *radial frequency.* $j$ is the imaginary unit given by $j^2 = -1$. We also introduce the *frequency* variable $f$, which relates to the former as $f = \frac{\omega}{2\pi}$, and the *wave length* $\lambda$, defining $\lambda = \frac{2\pi}{k}$. It should also be noted that the following is a solution of the wave equation:

$$u(x,t) = A\,e^{j(\omega\,t-k\,x)} + B\,e^{j(\omega\,t+k\,x)} + C\,t + D\,x + E, \tag{2.15}$$

where $C, D, E$ are also arbitrary constants. The interpretation of this is that any global offset in pressure does not influence wave-like motion, nor does a linear drift in time or space.

A particular formulation of a solution to the wave equation is the *d'Alembert solution*:

$$u(x, t) = u_+(x - c\,t) + u_-(x + c\,t), \tag{2.16}$$

where the solution is explicitly separated into two waves: one travelling to the right, $u_+(x - c\,t)$, and one travelling to the left, $u_-(x + c\,t)$.

Only 1-dimensional systems have been discussed so far, but acoustic problems are generally 2- and 3-dimensional. We can replace the spatial variable $x$ by a position vector $\mathbf{x} = x_1, x_2, ...$, indicating the position along all Cartesian axes. (Note that this is not restricted to just 2 or 3 dimensions.) The second spatial derivative is now denoted as $\nabla^2$ (often written as $\Delta$), also known as the *Laplacian operator*, which along three Cartesian axes resolves to $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$.

$$\frac{d^2}{dt^2} u(\mathbf{x}, t) = c^2\, \nabla^2 u(\mathbf{x}, t). \tag{2.17}$$

Henceforth in this thesis, bold-faced variables indicate vector quantities, i.e. $\mathbf{x} = \{x_1, x_2, x_3, ...\}$ describes a multi-dimensional position vector, in the particular 3-dimensional case: $\mathbf{x} = \{x, y, z\}$.

Another way of looking at spatial coordinates is using spherical coordinates. Here we define a point in 3-dimensional space by its azimuth $\phi$ (taken anti-clockwise with respect to the $x$-axis), its elevation angle $\theta$ (with respect to the $y$-axis), and its radius $r$ from the origin (see Fig. 2.2). In spherical coordinates, the Laplacian operator becomes:

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2}{\partial \phi^2}. \tag{2.18}$$

Though a more complicated formulation to work with, the spherical formulation of the wave equation gives us useful information on the behaviour of a point source radiating sound into space. It can be verified that the following is a solution to the wave equation in radial form:

$$u(r, t) = \frac{1}{r} u(r - c\,t) + \frac{1}{r} u(r + c\,t), \tag{2.19}$$

thus representing the sum of an inward and outward travelling spherical wave which decays in intensity with factor $\frac{1}{r}$.

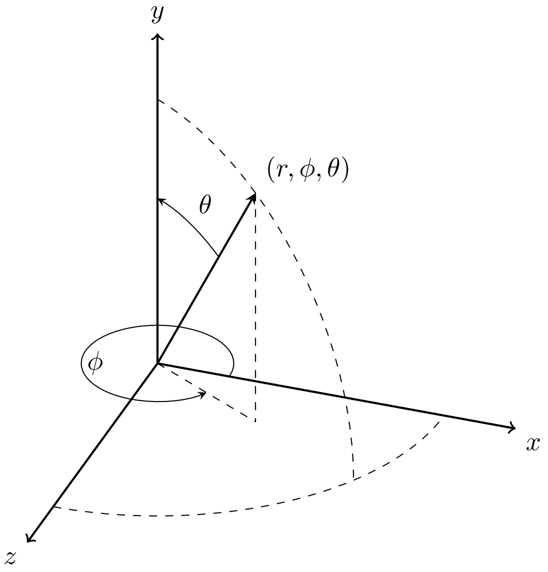


**Figure 2.2:** A polar coordinate system. In this work, $\phi$ will represent the azimuthal angle, $\theta$ the elevation, and $r$ the radius.

### 2.1.3 Inverse square law

Related to the spherical solution of the wave equation is the decay of a point source's intensity $I$ (as sound energy $E$ per unit area) with distance. Eq. 2.19 shows that the amplitude of a sound wave is inversely proportional to the distance from the source. In 3-dimensional air propagation, if the energy of a source is $E$, then the total energy on the sphere around the source at some radius $r$ is proportional to the sphere's surface area: $I = \frac{E}{4\pi r^2}$. Hence the intensity of a sound on a surface area is inversely proportional to the square of the distance to the sound source. This result is known as the *inverse square law.* (Note that this is a theoretical solution in the free field in which air absorption is ignored.)

### 2.1.4 Simple boundary conditions

At the boundaries of a domain, several conditions can be imposed that will determine the shape of the wave motion in the rest of the domain. Of particular interest are two simple boundary conditions: the *Dirichlet* and the *Von Neumann* boundary conditions. The Dirichlet boundary condition demands that at the boundary of the domain, the displacement is constant (and generally taken as 0). For simplicity, we place the boundary at $x = 0$:

$$u(0, t) = 0 \qquad \text{(Dirichlet condition).} \tag{2.20}$$

In the case of a tube, this boundary condition corresponds to a tube with an open ending. The Von Neumann condition on the other hand imposes that there be no net change in displacement at the boundary, corresponding to a tube with a closed end:

$$\frac{\partial}{\partial x} u(0,t) = 0 \qquad \text{(Von Neumann condition).} \tag{2.21}$$

We can solve the wave motion in a tube exactly using either one of these boundary conditions. Consider the transverse motion in a thin tube of length $L$. We impose the Dirichlet condition on both ends: $u(0,t) = u(L,t) = 0$. Substituting the left conditions into Eq. 2.14, we get:

$$u(0,t) = A\, e^{j(\omega\, t)} + B\, e^{j(\omega\, t)} = 0, \tag{2.22}$$

from which follows $A = -B$. Substituting the right boundary condition gives:

$$\begin{aligned} u(L,t) = A\, e^{j(\omega\, t - k\, L)} + B\, e^{j(\omega\, t + k\, L)} &= 0 \\ A\left(e^{jLk} + e^{-jLk}\right) &= 0 \\ e^{jLk} &= -e^{-jLk}. \end{aligned} \tag{2.23}$$

The above is only true for $L\, k = n\, \pi$, with $n \in \mathbb{N}$, from which it follows that $k = \frac{n\,\pi}{L}$. Substituting this again into Eq. 2.14 gives after simplification:

$$u(x,t) = A \cos\left(\frac{n\,\pi}{L} c\, t\right) \sin\left(\frac{n\,\pi}{L} x\right). \tag{2.24}$$

Any linear combination of the above solutions is also valid, and hence the general solution is:

$$u(x,t) = \sum_{n=0}^{\infty} A_n \cos\left(\frac{n\,\pi}{L} c\, t\right) \sin\left(\frac{n\,\pi}{L} x\right). \tag{2.25}$$

These are the eigensolutions of the wave equation with Dirichlet terminations. They are visualised in Fig. 2.3.

For a tube with Von Neumann boundary conditions, the same approach can be employed, though now the derivative at the boundaries needs to equal 0:

$$\frac{\partial}{\partial x} u(0,t) = -A\, \omega\, e^{j(\omega\, t)} + B\, \omega\, e^{j(\omega\, t)} = 0, \tag{2.26}$$

from which it follows that $A = B$. Omitting the steps similar to the ones above, it can be

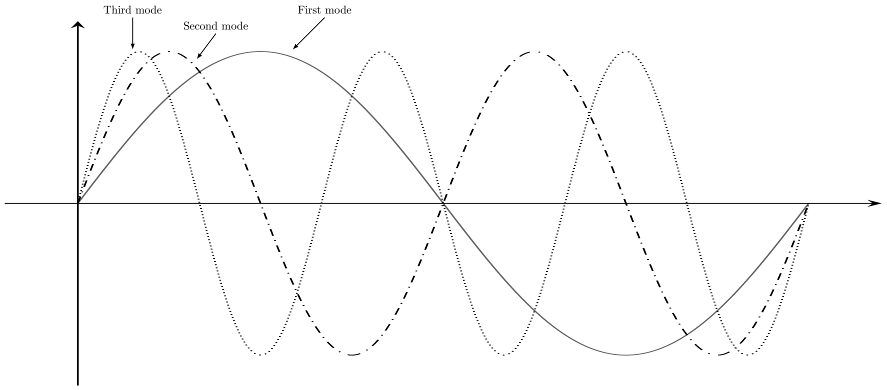


**Figure 2.3:** The first three vibrational modes for a tube with open, i.e. Dirichlet terminations at both endings.

shown that for this case, the general solution to the wave equation is:

$$u(x,t) = \sum_{n=0}^{\infty} A_n \sin\left(\frac{n\,\pi}{L}c\,t\right)\cos\left(\frac{n\,\pi}{L}x\right). \tag{2.27}$$

Fig. 2.4 visualises this result for the first 3 modes. From the above derivations and figures, it

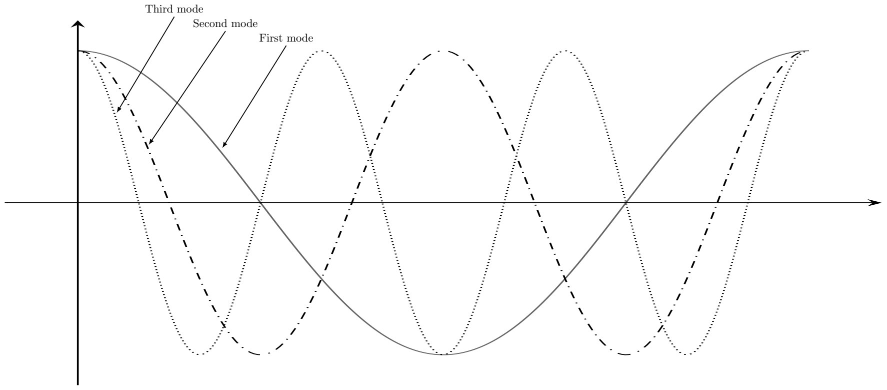


**Figure 2.4:** The first three vibrational modes for a tube with closed, i.e. Von Neumann terminations at both endings.

is clear that the boundary conditions determine the type of nodal pattern that appears along the tube: wherever the Dirichlet condition produces nodes along the tube (e.g. at the centre of the tube), the Von Neumann boundary condition produces anti-nodes, and vice versa.

### 2.1.5 Exact solutions in 3D

Of particular interest mainly on a theoretical ground is the 3D solution in a rectangular room. This provides a good test case for many algorithms because of its 3-dimensional yet simple nature, and there is an exact expression for the solution of the wave equation. Let us take a room of dimensions $L_x, L_y, L_z$, and let all the boundary conditions be of the rigid Von Neumann boundary type.

$$u(x,t) = \sum_{\substack{n_x=0\\n_y=0\\n_z=0}}^{\infty} A_{n_x,n_y,n_z} \sin\left(\sqrt{\left(\frac{n_x\,c}{L_x}\right)^2 + \left(\frac{n_y\,c}{L_y}\right)^2 + \left(\frac{n_z\,c}{L_z}\right)^2}\,\pi t\right) \cdot \\ \cos\left(\frac{n\,\pi}{L_x}x\right)\cos\left(\frac{n\,\pi}{L_y}y\right)\cos\left(\frac{n\,\pi}{L_z}z\right). \tag{2.28}$$

Evidently, this is analogous to the 1D tube example (Eq. 2.27), but the contributions of all axes are multiplied, and will form a 3-dimensional nodal and anti-nodal pattern in the shoebox-like room.

### 2.1.6 Alternative formulations of the wave equation

Even though the classical formulation of the homogeneous wave equation is presented in Eq. 2.12, there are several other possible formulations with their own specific advantages. First of all, it can be observed that Eq. 2.12 can be rewritten as:

$$\left(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right) = 0, \tag{2.29}$$

recognising it as the product of two first-order differential equations.

The formulation as a system of two separate equations can be useful as well:

$$\begin{cases} \frac{\partial u}{\partial t} - c\frac{\partial u}{\partial x} = p \\ \frac{\partial p}{\partial t} + c\frac{\partial p}{\partial x} = 0 \end{cases}, \tag{2.30}$$

where $p$ represents pressure. By differentiating the first equation over time and substituting it into the second, the correctness of this solution can easily be verified by manipulating it into the form of Eq. 2.12. A second useful two-step formulation is the following:

$$\begin{cases} \frac{\partial p}{\partial x} = -\rho\frac{\partial u}{\partial t} \\ \frac{\partial p}{\partial t} = -\rho\, c^2\frac{\partial u}{\partial x} \end{cases}. \tag{2.31}$$

This is perhaps the most common two-step formulation, as it has a clear interpretation: the first expression formulates conservation of momentum, while the second expression states conservation of mass. It also makes explicit that the pressure and displacement are each proportional to the other's derivative.

If we take $u(x,t)$ to be the d'Alembert solution to the wave equation (see Eq. 2.16), then we get:

$$\frac{\partial}{\partial x}u(x,t) = u'_{+}(x - c\,t) + u'_{-}(x + c\,t). \tag{2.32}$$

Combining this with the conservation of mass equation, we get:

$$\begin{aligned} p(x,t) &= \int -\rho\, c^2 \Big( u'_{+}(x - c\,t) + u'_{-}(x + c\,t) \Big) dt \\ &= \rho\, c \Big( u_{-}(x + c\,t) - u_{+}(x - c\,t) \Big) \\ &= p_{-}(x + c\,t) - p_{+}(x - c\,t), \end{aligned} \tag{2.33}$$

where we have set: $p_{\pm}(x \mp c\,t) = \rho\, c\, u_{\pm}(x \mp c\,t)$. From this, we can define:

$$\frac{p_{\pm}(x,t)}{u_{\pm}(x,t)} = \pm \rho\, c = \pm Z, \tag{2.34}$$

where $Z$ is called the *characteristic acoustic impedance.*

Lastly, to the formulation of the wave equation as in Eq. 2.12, a forcing/driving function $F(x,t)$ is often added to represent an external force actively exciting the domain:

$$\frac{d^2 u}{dt^2} = c^2 \frac{\partial^2 u}{\partial x^2} + F(x,t). \tag{2.35}$$

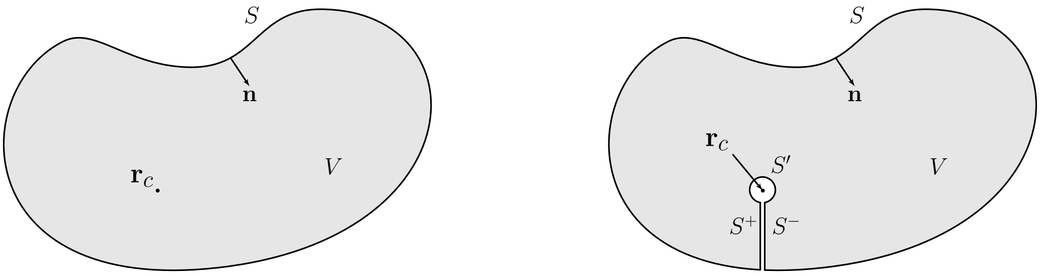
## 2.2 Kirchhoff-Helmholtz integral

The *Kirchhoff*- or *Kirchhoff-Helmholtz-integral theorem* provides a different way of describing a space that is governed by a wave equation. In particular, it reduces the problem of an $n$-dimensional wave equation to an $(n-1)$-dimensional problem. In the following derivation, only the 3-dimensional wave equation will be considered. The general form of a *Helmholtz equation* is:

$$\nabla^2 f + A^2 f = 0, \tag{2.36}$$

where $f$ is an arbitrary function, and $A$ is a constant. The homogeneous wave equation is essentially a Helmholz equation in time and space. Green's theorem relates a 3-dimensional volume integral to a (2-dimensional) surface integral as follows:

$$\int_V F\,\nabla^2 G - G\,\nabla^2 F dV = \oint_S G\frac{\partial F}{\partial \mathbf{n}} - F\frac{\partial G}{\partial \mathbf{n}} dS. \tag{2.37}$$



**(a)** A volume $V$ bounded by the surface $S$ and surface normal $\mathbf{n}$. $\mathbf{r}_c$ represents a sound source in $V$.

**(b)** The surface $S$ can be tightly folded around $\mathbf{r}_c$ in order for $V$ not to contain any singularities.

**Figure 2.5:** An illustration of the derivation of Kirchhoff-Helmholtz integral: the surface area $S$ bounding volume $V$ is folded around a source lest the volume contain any singularities, thus allowing integration over $S$.

Here $V$ represents any simply connected volume over which is integrated, $S$ is the closed surface that bounds that volume, and $F$ and $G$ are arbitrary integrable functions. $\mathbf{n}$ represents the surface normal pointing inside on $S$, and thus the term $\frac{\partial}{\partial \mathbf{n}}$ represents a derivative with respect to the surface normal. Even though Eq. 2.37 holds for a more general $F$ and $G$, we later choose $F$ and $G$ such that they satisfy the wave equation (Eq. 2.12). Using two Helmholtz functions, the inside of the volume integral reduces to:

$$F\,\nabla^2 G - G\,\nabla^2 F = A_1 F\,G - A_2 F\,G \;= (A_1 - A_2) F\,G\,. \tag{2.38}$$

If we demand that $A_1 = A_2$, then we get:

$$\oint_S G\frac{\partial F}{\partial \mathbf{n}} - F\frac{\partial G}{\partial \mathbf{n}} dS = \int_V F\,\nabla^2 G - G\,\nabla^2 F dV = 0\,. \tag{2.39}$$

If we now take $G$ to represent the Fourier transform of a radiating source at centre $\mathbf{r}_c$ within $S$ (see Eq. 2.19 for its time domain version), we get:

$$\oint_S \left(\frac{1}{|\mathbf{r}_c - \mathbf{r}|}\,e^{\pm j|\mathbf{r}_c - \mathbf{r}|k}\right)\frac{\partial F}{\partial \mathbf{n}} - F\,\frac{\partial}{\partial \mathbf{n}}\left(\frac{1}{|\mathbf{r}_c - \mathbf{r}|}\,e^{\pm j|\mathbf{r}_c - \mathbf{r}|k}\right)\,dS = 0\,. \tag{2.40}$$

Although $G$ satisfies the Helmholtz equation, it is not an integrable function in the entire domain, since the function is discontinuous at $\mathbf{r} = \mathbf{r}_c$. To circumvent this problem, we redraw the surface of $S$ by pushing it in at the surface such that it surrounds $\mathbf{r}_c$, as demonstrated in Fig. 2.5. $V$ is still simply connected, but does not include $\mathbf{r}_c$ any more. The surface area can now be written as the sum of connected surfaces $S$, $S'$, $S^+$, and $S^-$, and the surface integral becomes:

$$\oint_{S+S'+S^{+}+S^{-}} dS = \oint_{S} dS + \oint_{S'} dS + \underbrace{\oint_{S^{+}} dS + \oint_{S^{-}} dS}_{=0} = 0\,. \tag{2.41}$$

If we bring $S^{+}$ and $S^{-}$ infinitesimally close to each other, the two integrals cancel each other out because their normals point in opposite directions. We are therefore left with:

$$\oint_{S} G\frac{\partial F}{\partial \mathbf{n}} - F\frac{\partial G}{\partial \mathbf{n}} dS = -\oint_{S'} G\frac{\partial F}{\partial \mathbf{n}} - F\frac{\partial G}{\partial \mathbf{n}} dS. \tag{2.42}$$

The right-hand side is an integral over the surface that surrounds $\mathbf{r}_c$. If we set this sphere to be of size $\mathbf{r}_c - \mathbf{r} = \epsilon(\mathbf{r})$, the integral will run over the entire solid angle of the sphere, such that $dS = \epsilon^2 d\Omega$, and $\frac{\partial}{\partial \mathbf{n}} = \frac{\partial}{\partial r}$. Taking the limit for $\epsilon \to 0$, it follows that:

$$\begin{aligned}
&\lim_{\epsilon \to 0}\left(\oint_{S'} G\frac{\partial F}{\partial \mathbf{n}} - F\frac{\partial G}{\partial \mathbf{n}} dS\right) = \\
&\lim_{\epsilon \to 0}\left(\oint_{S'} \left(\frac{1}{\epsilon}\, e^{\pm j\epsilon k}\frac{\partial F}{\partial r} - F\,\frac{\partial}{\partial r}\frac{1}{\epsilon}\, e^{\pm j\epsilon k}\right)\, \epsilon^2 d\Omega\right) = \\
&\lim_{\epsilon \to 0}\left(\oint_{S'} \left(\epsilon\, e^{\pm j\epsilon k}\frac{\partial F}{\partial r} - F\left(-1 \pm j\epsilon^3\right) e^{\pm j\epsilon k}\right)\, d\Omega\right) = \\
&\oint_{S'} -F * (-1)\, d\Omega = 4\pi F
\end{aligned} \tag{2.43}$$

Substituting this result into Eq. 2.42 gives:

$$F = \frac{1}{4\pi}\oint_{S} F\frac{\partial}{\partial \mathbf{n}}\left(\frac{1}{|\mathbf{r}_s - \mathbf{r}|}\, e^{\pm j|\mathbf{r}_s - \mathbf{r}|k}\right) - \left(\frac{1}{|\mathbf{r}_c - \mathbf{r}|}\, e^{\pm j|\mathbf{r}_c - \mathbf{r}|k}\right)\frac{\partial F}{\partial \mathbf{n}}\, dS\,, \tag{2.44}$$

which is commonly know as the Kirchhoff-Helmholtz integral. Although $F$ is still a general function that satisfies the Helmholtz equation with $A^2 = k^2$, it is generally taken to represent the frequency spectrum at $\mathbf{r}$: $F = P(\mathbf{r}, \omega)$. As such, the frequency spectrum at some point in $V$ is now expressed in terms of an integral over the frequency spectra at the boundaries, multiplied with a time delay and a correction factor. The Kirchhoff-Helmholtz integral stands at the basis of several applications and theories, such as wave-field synthesis [17, 163, 96], the Boundary Element Method (see Sec. 4.4.3) and diffraction theory (see Sec. 2.5 and [97, 102, 204]).

## 2.3 The rendering equation

The rendering equation provides another means of expressing the sound field at any point in space in terms of the incoming acoustic radiance transferred from the rest of the space.

It derives from the field of computer graphics, where it was used by Kajiya [94] to describe the light transfer from objects onto each other. Siltanen *et al.* [183] argue that this equation is equally valid in room acoustics. In its time-dependent form, the rendering equation is as follows:

$$L_o(\mathbf{r}, \boldsymbol{\omega}_o, t) = L_e(\mathbf{r}, \boldsymbol{\omega}_o, t) + \int_0^{4\pi} \Phi(\mathbf{r}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o)\, L_i(\mathbf{r}, \boldsymbol{\omega}_i, t) \cos\theta_i \, d\boldsymbol{\omega}_i \, . \tag{2.45}$$

Here $L(\mathbf{x}, \boldsymbol{\omega}_i, t)$ and $L(\mathbf{r}_c, \boldsymbol{\omega}_o, t)$ are the respective incoming and outgoing radiance at $\mathbf{r}_c$ at time $t$ in direction $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_o$, and $L_e(\mathbf{r}_c, \boldsymbol{\omega}_o, t)$ is the emitted radiance. The angle $\theta_i$ is the angle that the incoming ray makes with respect to the surface normal. The distribution $\Phi(\mathbf{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o)$ represents the *bidirectional reflection distribution function* (BRDF) at $\mathbf{r}_c$, and is time-independent. The interpretation of this equation is that the emitted (acoustic) radiance at a point $\mathbf{r}_c$ in direction $\omega_o$ is equal to the radiance that the point itself emits, plus a fraction of the incoming radiance that is reflected into this direction. The integral runs over all possible incoming angles, i.e. a solid angle of $4\pi$.

The sound that arrives at a receiver position $\mathbf{r}_r$ is the sum of all the sound that radiates into its direction from all other points in space, with an additional time delay and diminution factor to account for the time it takes the sound to travel to the receiver. It is convenient to express this in the frequency domain $\hat{L}(\omega)$:

$$\hat{L}_i(\mathbf{r}_r, \omega) = \int_0^{4\pi} \int_0^{\infty} \frac{e^{-j|\mathbf{r}_r - \mathbf{r}|k}}{|\mathbf{r}_r - \mathbf{r}|} \hat{L}_o(|\mathbf{r}_r - \mathbf{r}|, \boldsymbol{\omega}_i, \omega)\, V(\mathbf{r}_r, \mathbf{r}) \, d\mathbf{r} \, d\boldsymbol{\omega}_i \, . \tag{2.46}$$

Here $V(\mathbf{r}_r, \mathbf{r})$ stands for a visibility factor, a measure for whether or not $\mathbf{r}$ and $\mathbf{r}_r$ are visible to each other. It is often taken to be 1 for visible and 0 for invisible, but it may also hold values between 0 and 1 if sound transmission through objects is allowed. Note that in both formulations, the rendering equation does not account for diffraction effects, and is therefore a purely geometric approximation.

## 2.4 Spherical harmonics

Sec. 2.1.5 presented 3-dimensional eigenfunctions to the wave equation using Cartesian coordinates. The spherical formulation of the wave equation (see Eq. 2.47) presents another set of eigenfunctions defined on the sphere. We start out by ignoring the radial part and focus on the angular part of the Laplace operator in spherical coordinates:

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}. \tag{2.47}$$

The spherical harmonic equation (see e.g. [91]) uses this to find solutions of the equation:

$$\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}+l(l+1)\right)u(\phi,\theta)=0, \tag{2.48}$$

where $l \in \mathbb{N}$. Using separation of variables, we assume that $u(\phi,\theta)$ can be written as $u(\phi,\theta) = u_\phi(\phi)\,u_\theta(\theta)$, so that Eq. 2.48 can be written as:

$$\frac{u_\phi(\phi)}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)u_\theta(\theta)+\frac{u_\theta(\theta)}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}u_\phi(\phi)+l(l+1)u_\phi(\phi)\,u_\theta(\theta)=0, \tag{2.49}$$

Multiplying with $\frac{\sin^2\theta}{u_\phi(\phi)\,u_\theta(\theta)}$ gives:

$$\frac{\sin\theta}{u_\theta(\theta)}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)u_\theta(\theta)+\frac{1}{u_\phi(\phi)}\frac{\partial^2}{\partial\phi^2}u_\phi(\phi)+l(l+1)\sin^2\theta=0 \tag{2.50}$$

$$\sin\theta\left(\frac{1}{u_\theta(\theta)}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)u_\theta(\theta)+l(l+1)\sin\theta\right)=-\frac{1}{u_\phi(\phi)}\frac{\partial^2}{\partial\phi^2}u_\phi(\phi) \tag{2.51}$$

In the latter formulation, each side is constant in terms of the other. The solution to the right-hand side can readily be seen to be:

$$u_\phi(\phi)=A\,e^{\pm j\,m\,\phi}, \tag{2.52}$$

where $m$ is some complex number. The equation in terms of $\theta$ is more involved. Substituting the previous result back into Eq. 2.53, we get:

$$\frac{\sin\theta}{u_\theta(\theta)}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)u_\theta(\theta)+l(l+1)\sin^2\theta=-m^2, \tag{2.53}$$

and multiplying with $\frac{u_\theta(\theta)}{\sin^2\theta}$ gives:

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)u_\theta(\theta)+l(l+1)\,u_\theta(\theta)+\frac{1}{\sin^2\theta}m^2\,u_\theta(\theta)=0. \tag{2.54}$$

We can now use the substitution $-\frac{1}{\sin\theta}\frac{\partial}{\partial\theta} = \frac{\partial}{\partial(\cos\theta)}$ to rewrite the above expression as:

$$\begin{aligned} -\frac{\partial}{\partial(\cos\theta)}\left(\sin^2\theta\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\right)u_\theta(\theta) + l(l+1)\,u_\theta(\theta) + \frac{1}{\sin^2\theta}m^2\,u_\theta(\theta) = 0 \\ \frac{\partial}{\partial(\cos\theta)}\left(\left(1-\cos^2\theta\right)\frac{\partial}{\partial(\cos\theta)}\right)u_\theta(\theta) + l(l+1)\,u_\theta(\theta) + \frac{m^2}{1-\cos^2\theta}\,u_\theta(\theta) = 0 \end{aligned} \tag{2.55}$$

This can be recognised as a *Legendre differential equation* [193] in terms of $\cos\theta$, to which the solution are the *associated Legendre Polynomials* $P_l^m(\cos\theta)$, provided that $m \in \mathbb{N}$ and $-l \leq m \leq l$. This gives us the angular part of the solution to the wave equation in spherical coordinates:

$$u_l^m(\phi,\theta) = A\,e^{\pm j\,m\,\phi}\,P_l^m(\cos\theta). \tag{2.56}$$

These are the eigenfunctions to Eq. 2.48, and there are $2l+1$ combinations for every order $l$.

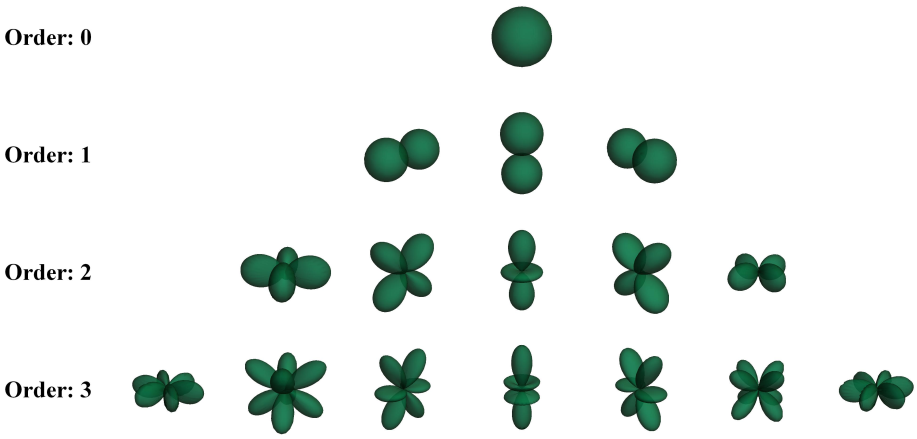


**Figure 2.6:** Spherical harmonics of order $0 \leq l \leq 3$. For each order $l$, there are $2l+1$ associated harmonical patterns.

Like the eigenfunctions in Cartesian coordinates, also the spherical harmonics form orthonormal and countable basis functions that are solutions to the Laplacian operator, on a spherical rather than cubical volume.

## 2.5 Kirchhoff-Fresnel Diffraction

When a wave encounters an obstacle, it will slightly bend around it into the shadow of the object. This phenomenon is known as *diffraction*, and can be quantified using the Kirchhoff-Helmholtz integral under some restrictions. Fig. 2.7 depicts a simple radiating point source

$P$ whose sound is incident onto a wall with a small slit into it. For simplicity, we assume that the sound produced at $P$ is a complex monopole: $u(P) = \frac{A}{r} e^{j\,k,r}$. We are interested in the sound incident on point $P'$, which represents a receiver on the other side of the wall.

We call the incident ray onto the slit $\mathbf{r}_1$, under incoming angle $\alpha$, and the outgoing ray to $P'$ $\mathbf{r}_2$, under angle $\beta$. By the Kirchhoff-Helmholtz integral (Eq. 2.44), we know that the sound at $P'$ can be expressed as a surface integral. We are free to choose any convenient surface, so we take a spherical surface $S_4$ around $P'$ that includes the slit $S_1$, and is bounded by the wall at $S_2$ and $S_3$. The integral then becomes:

$$F = \frac{1}{4\pi} \oint_{S_1+S_2+S_3+S4} F \frac{\partial}{\partial \mathbf{n}} \left( \frac{A}{|\mathbf{r}|} e^{\pm j|\mathbf{r}|k} \right) - \left( \frac{A}{|\mathbf{r}|} e^{\pm j|\mathbf{r}|k} \right) \frac{\partial F}{\partial \mathbf{n}} \, dS \,. \tag{2.57}$$

We now make the assumption that at the surface of the wall $S_2$ and $S_3$, the sound energy of $P$ and its surface normal is zero: $F_{S_2,S_3} = 0$, $\frac{\partial}{\partial \mathbf{n}} F_{S_2,S_3} = 0$. The latter can be justified by arguing that the wavefront only moves parallel to the wall, and thus the derivative with respect to the normal is zero. The former, however, is not actually correct, and is a necessary approximation, which will indeed invalidate the final formula to some extent. At $S_1$, we know $F$ and $\frac{\partial}{\partial \mathbf{n}} F$:

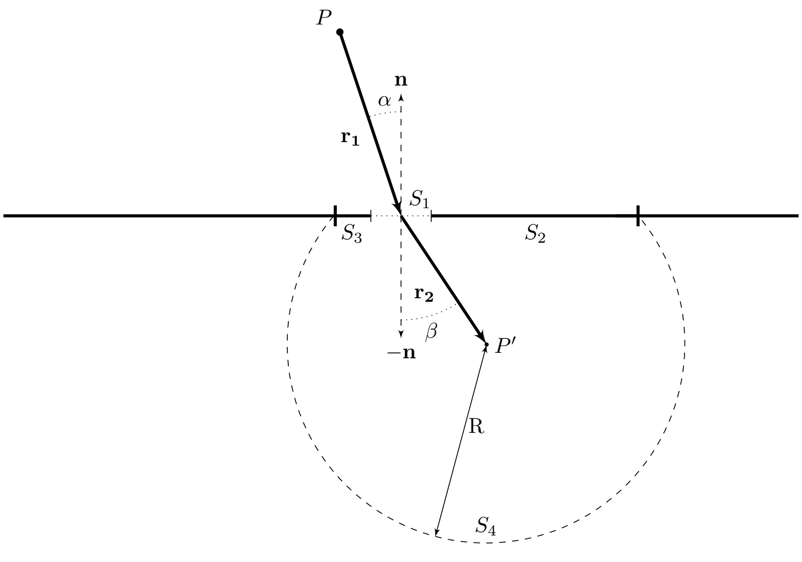


**Figure 2.7:** An incoming sound from point source $P$ enters through an opening in a wall.

$$\begin{aligned} F_{S_1}(\mathbf{r}_2) &= \frac{1}{|\mathbf{r}_2|} e^{j|\mathbf{r}_2|k} \\ \frac{\partial}{\partial \mathbf{n}} F_{S_1}(\mathbf{r}) &= \left( j\,k - \frac{1}{|\mathbf{r}_2|} \right) \frac{1}{|\mathbf{r}_2|} e^{j|\mathbf{r}_2|k} \cos\alpha. \end{aligned} \tag{2.58}$$

The final region we have to deal with is $S_4$, the part of the circle around $P'$. Clearly $F$ and its derivative are not zero here, but it was shown by Kirchhoff [100] and later e.g. Born [30] that the total contribution of $S_4$ is largely negligible. Thus we are left with the following integral over $S_1$:

$$\begin{aligned} F(P') = \frac{1}{4\pi} \int_{S_1} \Bigg( &\frac{1}{|\mathbf{r}_2|} e^{j|\mathbf{r}_2|k} \left( j\,k - \frac{1}{|\mathbf{r}_1|} \right) \frac{A}{|\mathbf{r}_1|} e^{j|\mathbf{r}_1|k} \cos\beta - \\ &\frac{A}{|\mathbf{r}_1|} e^{\pm j|\mathbf{r}_1|k} \left( j\,k - \frac{1}{|\mathbf{r}_2|} \right) \frac{1}{|\mathbf{r}_2|} e^{j|\mathbf{r}_2|k} \cos\alpha \Bigg) dS \end{aligned} \tag{2.59}$$

We may assume that $\mathbf{r}_{1,2}$ are much larger than the wavelength such that the terms $\frac{1}{\mathbf{r}_1}$ and $\frac{1}{\mathbf{r}_2}$ as subtractions from $j\,k$ vanish, and we get:

$$\begin{aligned} F(P') &= \frac{1}{4\pi} \int_{S_1} \frac{j\,k\,A}{|\mathbf{r}_1|\,|\mathbf{r}_2|} e^{j(|\mathbf{r}_1|+|\mathbf{r}_2|)k} (\cos\beta - \cos\alpha)\, dS \\ &= \frac{jA}{2\lambda} \int_{S_1} \frac{e^{j(|\mathbf{r}_1|+|\mathbf{r}_2|)k}}{|\mathbf{r}_1|\,|\mathbf{r}_2|} (\cos\beta - \cos\alpha)\, dS \end{aligned} \tag{2.60}$$

This result is commonly known as the *Kirchhoff* or *Kirchhoff-Fresnel diffraction equation.*

## 2.6 Speed of sound

The speed of sound varies greatly per medium. As a general rule, it propagates slowest through air, faster through liquids, and fastest through solids. In this work, the only medium we will consider—unless explicitly stated otherwise—will be atmospheric air (which is a composite of nitrogen ($\sim$ 79%), oxygen ($\sim$ 20%), argon ($\sim$ 1%), and small quantities of various other gases). A great many studies have been devoted to deriving the speed of sound in air mathematically and/or finding it experimentally—for a good summary see e.g. [221]. The general formula for the speed of sound is:

$$c = \sqrt{\frac{R\,T\,\gamma}{M}}, \tag{2.61}$$

where $R$ is the universal gas constant ($8.314 \frac{J}{K\,\mathrm{mol}}$), $T$ is the absolute temperature (in [K]), $\gamma$ the specific heat ratio, and $M$ is the molar mass of the air (in $[\frac{\mathrm{mol}}{m^3}]$). We define standard barometric pressure $p_0 = 101.325$ kPa. Then at a temperature of 293.15 K (i.e. 20°C), the speed of sound is $343.26 \frac{m}{s}$. This is the speed of sound that will be adhered to in the rest of this work. Other works often use slightly different choices, such as the speed at 273.15

K (or 0°C), or different pressure rates. However, the differences are relatively small, of the order of $\pm 10\frac{m}{s}$, and to the best of author's knowledge there is no evidence that, assuming a homogeneous static medium, these slight differences in speed of sound are perceptible by the human hearing system.

## 2.7 Summary

The 1-dimensional acoustic wave equation has been derived from first principles using the example of a gas-filled tube. It is recognised that this is the simplest version of the homogeneous wave equation, and higher-order terms that cover e.g. air absorption are left out. Several solutions to the 3-dimensional wave equation have been given, which give rise to some concepts fundamental to the rest of this thesis, such as Von Neumann and Dirichlet boundary conditions, spherical harmonics, and diffraction. Two more fundamental concepts that stand at the foundation of several acoustic modelling algorithms are the Kirchhoff-Helmholtz integral and the room acoustic rendering equation. The following chapter will use these concepts to present a slightly higher level view on room acoustics.

# Chapter 3

# Introduction to room acoustics

Whereas the previous chapter deals with the mathematical constructs of computational room acoustics, this section discusses the physical interpretation of wave behaviour in some acoustic medium. Sec. 3.1 discusses the meaning of a travelling wave and the interaction between different mediums, which result in wave reflection, absorption, and diffraction. In Sec. 3.2, the aforementioned are taken to a macro-scale, thus forming the room acoustical properties of a space. Room acoustic concepts such as the ‘impulse response’ are introduced, and the relevant standardised ISO 3382 parameters are discussed. Sec. 2.6 and Sec. 3.1.6 derive the speed of sound and air absorption from first principles, and provide a reference for the values that will be used throughout this thesis.

## 3.1 Acoustic media

An acoustic environment is defined by the objects in it, the boundaries surrounding it, and the propagation medium (which in this work will always be air). The correct modelling of the acoustic properties of both the boundaries and the medium is of great importance for the final output. Since the acoustics of spaces are mostly determined by the room’s boundaries and the large objects in the space, small details in the geometry can often be ignored. The level of detail of the input data can thus be relatively low: Vorländer [214] states that as a rule of thumb, a resolution of 0.5 meter can be a guideline for a geometric acoustic model; objects smaller than this may be ignored. He also claims that a high level of detail will not only increase computation time, but will also yield wrong or biased results, although he does not further support this claim. This statement is probably incorrect, as smaller objects can often have a diffusing effect, thus making a non-negligible contribution to the overall acoustics of the space. For the purposes of this work, however, the 0.5 meter model detail shall suffice as a rule of thumb, as this work is not directly concerned with these intricacies of acoustic modelling.

Every acoustic medium has several acoustic properties, and every surface (i.e. an interface between two different acoustic mediums) its own behaviours, including their acoustic absorption, transmission, and scattering. The following sections describe the relevant acoustic interactions between two acoustic mediums. Most of following theory and derivations follow those by [23, 99, 64].

### 3.1.1 Acoustic impedance

The *specific acoustic impedance* describes the impedance at a single point in space $\mathbf{x}$, and is quantified as:

$$Z(\mathbf{x}, t) = \frac{p(\mathbf{x}, t)}{|\mathbf{u}(\mathbf{x}, t)|}. \tag{3.1}$$

$\mathbf{u}(\mathbf{x}, t)$ is a vector quantity, and its magnitude together with the pressure value define the impedance $Z$, which is a scalar function of space. Since the pressure and displacement are generally time-dependent functions, $Z$ may also be a time-varying quantity. Since this is not very convenient in practice, it is useful to express $Z$ in the frequency domain. $Z$ is in general frequency-dependent, and can also be written as the transfer function of the pressure and velocity:

$$Z(\mathbf{x}, \omega) = \frac{P(\mathbf{x}, \omega)}{U(\mathbf{x}, \omega)}, \tag{3.2}$$

where $P(\mathbf{x}, \omega)$ and $U(\mathbf{x}, \omega)$ are the Fourier transformed pressure and velocity signals. The impedance may be complex valued: $Z = R + Bj$, in which case $R$ is referred to as the *resistance* and $B$ as the *reactance*. The reciprocal of the impedance, $\Xi(\omega) = \frac{1}{Z(\omega)}$, is called the *admittance*, and is a measure for the particle velocity that is realised by a certain amount of pressure. Though less convenient a formulation, for completeness' sake let us also state the time-domain version of the impedance relation:

$$p(\mathbf{x}, t) = Z(\mathbf{x}, t) * u(\mathbf{x}, t), \tag{3.3}$$

where $*$ signifies convolution, and $Z(\mathbf{x}, t)$ is the inversely Fourier transformed signal of $Z(\mathbf{x}, \omega)$.

### 3.1.2 Wave-boundary interaction

The homogeneous wave equation is approximately valid in a continuous medium such as air. At boundaries however, which can be considered discontinuities in the medium's density, it no longer holds. According to the principle of conservation of mass, the difference in acoustic pressure over time at such a boundary is related to the density of and the speed of sound in the main medium, and the velocity gradient:

$$\frac{\partial}{\partial t} p(\mathbf{x}, t) = -\rho\, c^2\, \nabla \mathbf{u}(\mathbf{x}, t). \tag{3.4}$$

According to the conservation of momentum, the pressure gradient equals the negative density times the time derivative of the velocity field:

$$\nabla p(\mathbf{x}, t) = -\rho \frac{\partial}{\partial t} \mathbf{u}(\mathbf{x}, t). \tag{3.5}$$

The acoustic impedance $Z$, already introduced in Sec. 2.1.6, relates the pressure at a point to the particle displacement $u$ (not to be confused with the propagation speed of sound $c$). More specifically, it is a measure for the amount of pressure generated by a certain particle velocity. Acoustic modelling mostly deals with homogeneous volumes, in whose interior the impedance value is of no particular interest as it is constant throughout. Only at surfaces, the impedance plays a significant role. The *acoustic impedance* is therefore defined as the impedance normal to the boundary. It is typically formulated in the frequency domain:

$$Z(\omega) = \frac{P(\omega)}{\mathbf{n} \cdot \mathbf{U}(\omega)}\ , \tag{3.6}$$

where $Z(\omega)$ is defined at the boundary of two surfaces, and $\mathbf{n}$ is the surface normal. $\mathbf{U}$ is now a vector quantity, as the boundary orientation may have an effect on the direction of the particle velocity. Unless otherwise specified, all boundary interactions we consider are interactions between air and some other material. Therefore, it is convenient to define the *normalised characteristic impedance* $\zeta_w$, which is the impedance of the wall $Z_w$ divided by the impedance of air:

$$\zeta_w(\omega) = \frac{Z_w(\omega)}{Z_{\text{air}}(\omega)} = \frac{Z_w(\omega)}{(\rho\, c)_{\text{air}}}\ , \tag{3.7}$$

where $\rho_{\text{air}}$ and $c_{\text{air}}$ are the density of and the speed of sound in air.

We can make the (non-physical) simplification of $Z(\omega)$ being frequency-independent, and thus $Z(t)$ being constant. As such, Eq. 3.6 has a simple time-domain version (equivalent to Eq. 3.3 in multiple dimensions):

$$p(t) = Z(t) * \mathbf{n} \cdot \mathbf{u}(t) = u(t) Z \cos\theta, \tag{3.8}$$

which is valid across the entire surface of a uniform boundary. This simplification also allows us to differentiate Eq. 3.8 with respect to time, which, combined with Eq. 3.5, gives the following impedance formulation at a boundary:

$$\rho \frac{\partial}{\partial t} p(\mathbf{x}, t) = -Z \cos\theta \; \nabla p(\mathbf{x}, t)\,. \tag{3.9}$$

This formulation has as particular advantage that it is expressed only in terms of pressure.

### 3.1.3 Absorption and reflection

If we consider a wave $p^{+}$ that arrives at a boundary of two mediums under an angle $\theta_i$, it produces a reflected wave $p^{-}$ under angle $\theta_o$ and a transmitted wave $p^{tr}$ under angle $\theta_t$. This process is graphically represented in Fig. 3.1. For simplicity, we only consider a plane wave travelling in two dimensions, such that we can write:

$$p^{+} = p\left(t - \frac{x \sin\theta_{\mathrm{i}} + y \cos\theta_{\mathrm{i}}}{c_1}\right) \tag{3.10}$$

$$p^{-} = p\left(t + \frac{x \sin\theta_{\mathrm{o}} + y \cos\theta_{\mathrm{o}}}{c_1}\right) \tag{3.11}$$

$$p^{tr} = p\left(t - \frac{x \sin\theta_{\mathrm{t}} + y \cos\theta_{\mathrm{t}}}{c_2}\right)\,, \tag{3.12}$$

where $c_{1,2}$ are the speed of sound in the respective mediums. This is illustrated in Fig. 3.1. Because of continuity of pressure and velocity, we have:

$$p^{+} + p^{-} = p^{t} \tag{3.13}$$

$$u^{+} + u^{-} = u^{t}\,. \tag{3.14}$$

If we only consider the wave propagation in the $x$-direction (i.e. along the boundary), then using the previously stated equation of pressure continuity, we get:

$$p^{+}(t) + p^{-}\left(t + \frac{x\left(\sin\theta_{\mathrm{i}} - \sin\theta_{\mathrm{o}}\right)}{c_1}\right) = p^{tr}\left(t + x\left(\frac{\sin\theta_{\mathrm{i}}}{c_1} - \frac{\sin\theta_{\mathrm{t}}}{c_2}\right)\right)\,. \tag{3.15}$$

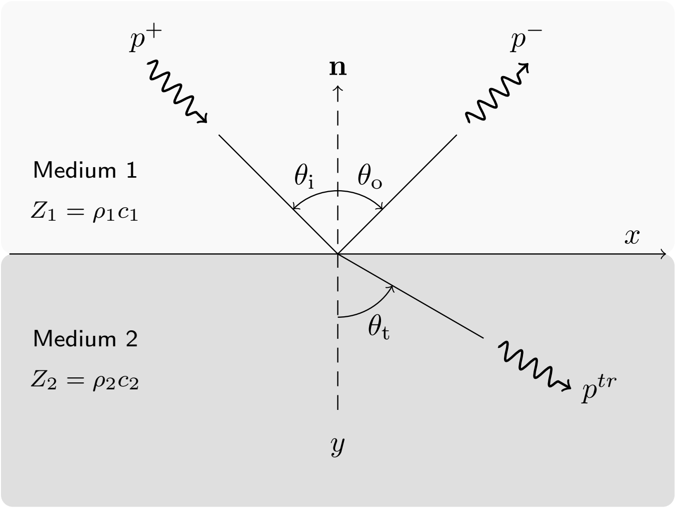


**Figure 3.1:** A wave $p^+$ arriving at a boundary under an angle $\theta_i$ produces a reflected wave $p^-$ and a transmitted wave $p^{tr}$, under angles $\theta_o$ and $\theta_t$, respectively.

This can only be satisfied for all possible values of $t$ and $\theta$ if $\sin\theta_\mathrm{i} = \sin\theta_\mathrm{o}$ (i.e. specular reflection of outgoing wave) and:

$$\frac{\sin\theta_\mathrm{i}}{c_1} = \frac{\sin\theta_\mathrm{t}}{c_2} \,. \tag{3.16}$$

This result is commonly known as Snell's law, and relates the propagation speeds in both mediums to the amount of diffraction. If we then define the reflection coefficient $R = \frac{p^-}{p^+}$ and the transmission coefficient $T = \frac{p^{tr}}{p^+}$, we find that:

$$1 + R = T. \tag{3.17}$$

Using the impedance relationship between pressure and velocity at a boundary (see Eq. 3.6), and adding a minus-sign to $u^-$ since it travels backwards, Eq. 3.14 can also be written as:

$$\frac{p^+}{Z_1\cos\theta_\mathrm{i}} - \frac{p^-}{Z_1\cos\theta_\mathrm{o}} = \frac{p^{tr}}{Z_2\cos\theta_\mathrm{t}} \,. \tag{3.18}$$

From this, and recalling that $\theta_\mathrm{i} = \theta_\mathrm{o}$, equations for the reflection and transmission coefficient can be derived in terms of $Z_1$ and $Z_2$:

$$R = \frac{Z_2\cos\theta_\mathrm{i} - Z_1\cos\theta_\mathrm{t}}{Z_2\cos\theta_\mathrm{i} + Z_1\cos\theta_\mathrm{t}} \qquad\qquad T = \frac{2Z_2\cos\theta_\mathrm{i}}{Z_1\cos\theta_\mathrm{t} + Z_2\cos\theta_\mathrm{i}} \,. \tag{3.19}$$

From these equations, it follows that the transmission coefficient can only be negative when $\cos\theta_\mathrm{i} < 0$ and/or $\cos\theta_\mathrm{t} < 0$. Since this would relate to incident and outgoing angles larger than $\pi$, this does not correspond to a physical process. $R$ on the contrary can clearly be negative. Since $p^-(t) = R\,p^+(t)$, a positive reflection coefficient relates to in-phase reflection, whereas a negative $R$ corresponds to a phase shift of $\pi$ with respect to the incoming wave. If we solve Eq. 3.19 for $Z_2$, we get:

$$Z_2 = Z_1 \cdot \frac{\cos\theta_\mathrm{t}}{\cos\theta_\mathrm{i}} \cdot \frac{1-R}{1+R}\,. \tag{3.20}$$

In acoustic problems, it can almost always be assumed that Medium 1 is air. Combining this with Eq. 3.7, we find the characteristic impedance $\zeta$ for a wall:

$$\zeta = \frac{Z_2}{(\rho\,c)_\mathrm{air}} = \frac{Z_1}{(\rho\,c)_\mathrm{air}} \cdot \frac{\cos\theta_\mathrm{t}}{\cos\theta_\mathrm{i}} \cdot \frac{1-R}{1+R} = \frac{\cos\theta_\mathrm{t}}{\cos\theta_\mathrm{i}} \cdot \frac{1-R}{1+R}\,. \tag{3.21}$$

### 3.1.4 Transmission and refraction

Sound transmission is an acoustic phenomenon not often implemented in modelling software. The primary reason is that transmission effects are not very large compared to reflection and scattering. Moreover, designers are often more interested in the acoustics within a space and assume that outside influences are negligible. Nonetheless, transmission is a real acoustic phenomenon that may be important in specific cases, and is therefore important to discuss.

Sound transmission is the phenomenon of a sound wave travelling from one medium into another. It was shown in the previous section that the transmission coefficient $T$ of a surface is related to the characteristic impedances of the two neighbouring surfaces. It was also shown that the angle of the transmitted sound need not be the same as the angle of incidence of the incoming sound, but that they are related by Snell's law (Eq. 3.16). Therefore, it follows that the outgoing angle should be:

$$\theta_\mathrm{t} = \arcsin\left(\frac{c_2}{c_1}\sin\theta_\mathrm{t}\right)\,. \tag{3.22}$$

For this result to be real, the argument of the arcsin function must be smaller than or equal to 1. Therefore:

$$\sin\theta_\mathrm{i} \leq \frac{c_1}{c_2}\,. \tag{3.23}$$

The *critical angle* $\theta_c$ is the maximum angle of incidence that allows transmission and is defined as $\theta_\mathrm{cr} = \arcsin\left(\frac{c_1}{c_2}\right)$. When $\theta_\mathrm{i} < \theta_{cr}$, transmission will occur. In geometric acoustics, the transmitted sound power is proportional to $T^2$.

### 3.1.5 Diffusion

Diffusion or scattering is the process where an incident sound wave is reflected in multiple directions, rather than only specularly. The particular distribution of scattered energy is often very complex, and often depends on parameters such as frequency, angle of incidence, surface material, and surface roughness. It is therefore hard if not impossible to find exact expressions that cover all of these parameters.

There are several ways to treat diffusion. It has been proven to be a reasonable approximation to model surfaces as Lambertian reflectors [117]. This approximation assumes that all surfaces reflect sound according to Lambert's cosine law: disregarding the angle of incidence of the wave, the wave energy is scattered in each direction proportional to the cosine of its angle to the surface normal. In other words, assuming wave energy $E_\mathrm{i}(\boldsymbol{\omega}_\mathrm{i})$ incident upon a surface $dA$ from direction $\boldsymbol{\omega}_\mathrm{i}$, then the energy $E_\mathrm{o}(\boldsymbol{\omega}_\mathrm{o})$ scattered into the solid angle $d\Omega$ in the direction of $\boldsymbol{\omega}_\mathrm{in}$ with respect to the surface normal reads as:

$$E_\mathrm{o}(\omega_\mathrm{o}) = E_\mathrm{i}(\boldsymbol{\omega}_\mathrm{i}) \cos(\theta_\mathrm{o})\, d\Omega\, dA, \tag{3.24}$$

which is demonstrated graphically in Fig. 3.2.

This approach does not take into account the angle of incidence of an acoustical ray, and is therefore a fairly crude approximation. It has often been proposed in literature [215, 45, 160] to circumvent this problem by introducing a scattering coefficient $\sigma$ between 0 and 1, which signifies the fraction of energy that is scattered in a Lambertian way. The distribution for the outgoing energy then becomes:

$$E_\mathrm{o}(\boldsymbol{\omega}_\mathrm{o}) = \begin{cases} E_\mathrm{i}(\boldsymbol{\omega}_\mathrm{i})\ \cos(\theta_\mathrm{o})\, d\Omega\, dA & \sigma \text{ of the time} \\ E_\mathrm{i}(\boldsymbol{\omega}_\mathrm{i})\, \delta(\boldsymbol{\omega}_\mathrm{i} - \boldsymbol{\omega}_\mathrm{o})\, d\Omega\, dA & (1-\sigma) \text{ of the time.} \end{cases} \tag{3.25}$$

A more precise way of storing scattering data is by means of a *bidirectional reflectance distribution function* (BRDF) [183]. This is a data structure that saves the reflected and scattered energy distribution as a function of incoming and outgoing angle. For each angle $d\theta_\mathrm{i}$, the fraction of energy reflected to $d\theta_\mathrm{o}$ and $d\phi_\mathrm{o}$ is stored. The advantage of BRDFs is that they can be measured or computed off-line and may give a very accurate approximation. The disadvantage is that they are frequency-dependent and, depending on the frequency resolution chosen, they may become relatively large data structures.

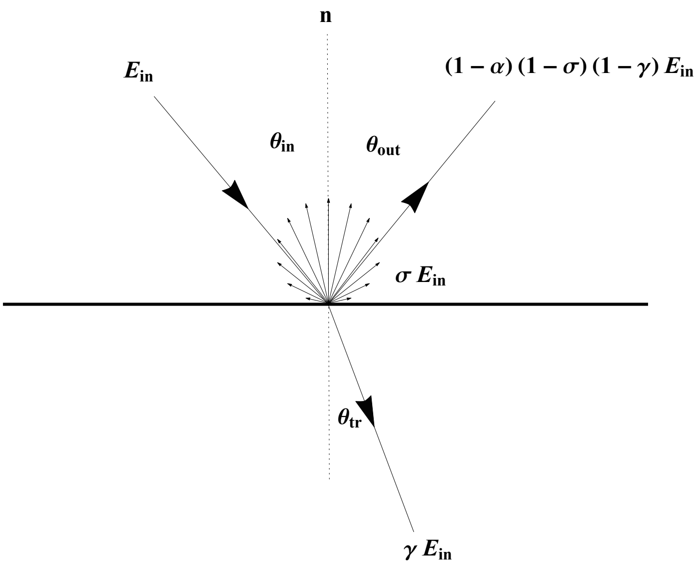


**Figure 3.2:** A schematic representation of the effect of reflection, diffusion, and scattering. An incoming wave with energy $E_\mathrm{i}$ under an angle $\theta_\mathrm{i}$ is partially absorbed (by absorption $\alpha$), partially transmitted ($\gamma E_\mathrm{i}$) under $\theta_\mathrm{t}$, partially scattered ($\sigma E_\mathrm{i}$), and the rest is reflected.

### 3.1.6 Air absorption

Though the homogeneous wave equation is lossless, air absorption is a very real phenomenon that plays an important role in the acoustics of any space. Air absorption is heavily frequency dependent and is especially prominent in high frequencies. An example of a more general wave equation that models this behaviour was already given in Eq. 2.13. The absorption of sound pressure by air is heavily dependent on its temperature and humidity, and is generally not negligible. Elaborate research has been done by, among others, Bass *et al.* [11, 12, 13] into the theoretical and experimental value of air absorption across the frequency spectrum, between 50 Hz and 1 MHz. They use the following constructs.

The *vibrational relaxation time* of a molecule is the time it takes for a molecule to return to an equilibrium from a high-energy state. Its reciprocal times $2\pi$ is the *molecular relaxation frequency* [226]. The relaxation frequency of oxygen $f_{r,O_2}$ and nitrogen $f_{N_2}$ are experimentally determined to be:

$$f_{r,O_2}(\eta) = 24 + 4.04 * 10^4\,\eta\,\frac{0.02+\eta}{0.391+\eta}, \tag{3.26}$$

$$f_{N_2}(\eta) = \sqrt{\frac{T_0}{T}}\left(9 + 280\,\eta\,\exp\left(-4.17\left[\left(\frac{T_0}{T}\right)^{\frac{1}{3}} - 1\right]\right)\right), \tag{3.27}$$

both of which are a function of the humidity $\eta$ (in [%]) and $T$, the absolute temperature (in [K]), $T_0$ is the reference temperature 273.15 K [227, 13]. The air absorption $\alpha_{\text{air}}(f)$ in $\left[\frac{\text{dB}}{\text{m}}\right]$ for a frequency $f$ is then given as:

$$\begin{aligned}\alpha_{\text{air}}(f) = \frac{20}{\ln 10} f^2 \Bigg( & 1.84 \cdot 10^{-11} \sqrt{\frac{T}{T_0}} + \\ & \left(\frac{T}{T_0}\right)^{-\frac{5}{2}} \left( \left( 0.01278 \exp\left(-\frac{2239.1}{T}\right) \frac{1}{f_{r,O_2} + \frac{f^2}{f_{r,O_2}}} \right) + \right. \\ & \left. \left( 0.1068 \exp\left(-\frac{3352.0}{T}\right) \frac{1}{f_{N_2} + \frac{f^2}{f_{N_2}}} \right) \right) \Bigg). \end{aligned} \tag{3.28}$$

Fig. 3.3 graphically represents this, for several different humidity levels, at 293.15 K (20°C). Air absorption across all frequencies increases approximately linearly on a dB scale, with two clear inflection points in them which relate to the oxygen and nitrogen relaxation frequency. At 0% humidity, air absorption is lowest at all frequencies. For humidity levels between 10% and 90%, absorption at low and middle frequencies (<~ 2500Hz) becomes less with increasing humidity. At higher frequencies, until ~ 20 000Hz, it is very humidity-dependent, and at frequencies higher than ~ 20 000Hz (beyond the human hearing range), air absorption is roughly linear for all frequencies.

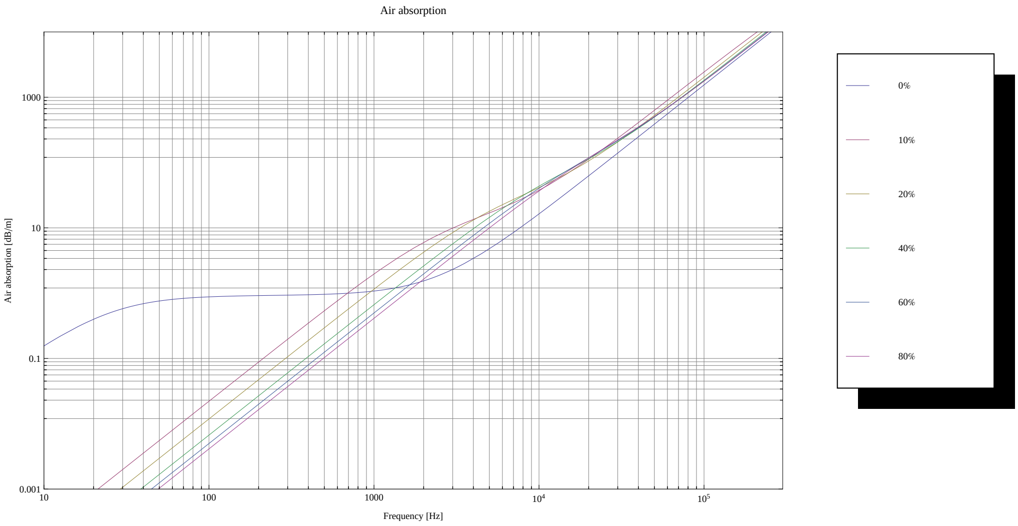


**Figure 3.3:** Air absorption (in $\left[\frac{\text{dB}}{\text{m}}\right]$) for frequencies between 10 and 20 000 Hz, at different humidity levels, at 293.15 K (20°C).

## 3.2 Room acoustics

### 3.2.1 Room impulse response

The *room impulse response* (RIR) is the time-domain transfer function from the sound source at one point in space onto a receiver in the same space. The RIR is formally defined as the *linear time-invariant* (LTI) transfer function $h(t)$ of the sound energy from source to receiver in a particular environment, with the source sending out a *Dirac delta function* $\delta(t)$. The delta function in continuous time is defined as the first time derivative of the unit step function $s(t)$:

$$\delta(t) = \frac{\partial}{\partial t} s(t) = \frac{\partial}{\partial t} \begin{cases} 0 & \mathrm{t} < 0 \\ 1 & \mathrm{t} \geq 0 \end{cases} \tag{3.29}$$

Hence, $\delta(t)$ equals 0 except at $t = 0$, where the function value is undefined but has an area under the curve of 1. (There is more to be said about the delta function, which, strictly mathematically speaking, is not a function but a distribution—see e.g. [197]. For the purpose of this work, however, this is irrelevant.) In discrete time, the delta function is defined as $d[0] = 1$ and $d[n] = 0$ for all $n \neq 0$, and is referred to as the *Kronecker delta function.* The Fourier transform of the delta function is a constant, i.e. a flat distribution. It is therefore very well suited for determining a transfer function of a room in a computational domain without frequency bias, since all frequencies are represented equally.

An RIR is generally said to consist of three parts: the direct sound, the early reflections, and the reverberant tail (or late reverberation) [90]. Fig. 3.4 shows a schematic impulse response indicating these three concepts. The direct sound is the proportion of the sound that travels from source to receiver directly, without any reflections. The early reflections are the reflections that bounce off the walls and reach the receiver relatively soon (typically $< 80ms$) afterwards. The direct sound and the early reflections together are mainly responsible for sound localisation by the human brain [85]. The late reverberation is the sound that reaches the ear after several (largely diffuse) reflections. It contains information about the size and material of the room, but less about its exact shape [85].

As an RIR is the sound energy transfer function from source to receiver taking into account all aspects of a room and as it represents an LTI system, it can be used to also determine how other sounds propagate in the space. If one takes an anechoic recording, i.e. a recording (theoretically) free of any reverberation, and convolves this with the RIR $h(t)$, this signal will represent the sound heard at the receiver position of the RIR, had it been played back

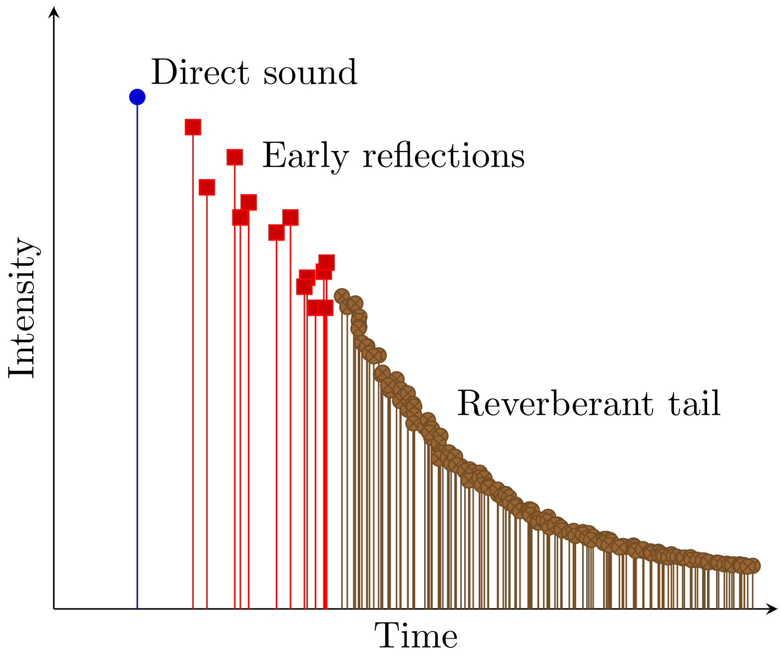


**Figure 3.4:** A schematic representation of an RIR, divided into the three most important sections. The direct sound is the sound that arrives from source to receiver without any obstruction. The early reflection are distinguishable reflections that arrive soon after. The reverberant tail comprises the later reflections that cannot be distinguished visually in the reflectogram and form a diffuse field.

at the source position of the RIR. A caveat to this observation is that it assumes both sound source and receivers had the same directivity pattern. Moreover, the fact that we restrict the RIR to be LTI means that nonlinear effects such as those produced by high intensity sounds cannot be taken into account.

### 3.2.2 Reverberation

The reverberation time is a measure for the time it takes a sound to die away. Sabine [161] was among the first to measure reverberation times in different spaces in the early $20^{\text{th}}$ century. He defined the reverberation time $RT_{60}$ as the time it took for the level of a signal to drop to a millionth of its original strength (i.e. -60 dB). He found that the most important variables were the volume of the room, the surface area of the room, and the absorption coefficients of the materials these surfaces were made of. For more about his methods and derivation, see Sec. 4.2.1. Ever since, it has been the main parameter by which the acoustics of a space can be defined. The well-known *Sabine equation* he experimentally found is:

$$RT_{60} = \frac{0.164\,V}{S\,\alpha_{av}}, \tag{3.30}$$

where $V$ is the volume of the room, $S$ the surface area, and $\alpha_{av}$ the average absorption power across all surfaces. As the last one is frequency-dependent, also the outcome of this formula is valid for just one frequency. It must be noted that this formula is merely stated here as a historical reference, but extremely crude and not actually suitable for acoustic prediction. More on this can be found in Sec. 4.2.1.

Sabine observed that it is important to note that the reverberation time is frequency-dependent, because also material and air absorption are heavily frequency-dependent. To simplify this frequency-dependency, absorption times across a range of frequencies (often an octave, fifth, or third) are averaged into a *frequency band* that represents these frequencies. Thus, the reverberation characteristics of a room can be plotted as a graph with $RT_{60}$ against frequency, or, more commonly, for $N$ (often 3, 8, 12, or 36) separate frequency bands.

### 3.2.3 ISO 3382

The International Standard ISO 3382 [66] was designed to provide a larger reference frame of how to compare acoustical properties. It contains a set of minimum requirements for good measurement procedures and defines a set of objective acoustical parameters. Depending on the parameter, they can be derived from one impulse response, or from several with different characteristics (e.g. different polar patterns). Since the acoustical parameters are based on psychoacoustically relevant phenomena, they are useful when comparing the acoustical properties of different spaces. Moreover, the fact that they can be derived from one or two impulse responses gives a valid way of comparing measured results with computer simulated data.

#### Impulse response

Exactly determining the transfer function of a room for a certain source and receiver position would require the following: a perfectly omnidirectional source and receiver, no environmental noise, and a delta function as excitation at the source position. In practice of course, all these conditions have to be approximated. The source and receiver should be as omnidirectional as possible. The noise level should be kept to a minimum, at least 45 dB below the signal level [66]. As a delta function is a theoretical concept, it needs only to be approximated. This can be done by using a signal whose power is equally distributed over all frequencies, for example a gunshot or white noise. Often more conveniently, a long signal known as a *sine sweep* can move through all frequencies to cover the full frequency range of interest. For a good reference on impulse response measurements, see e.g. [59, 61].

An impulse response $p(t)$ can then be measured: in the case of an impulse source, the recorded sound already is the transfer function and may need little processing. In the case where a sine sweep was used, deconvolution can be applied to obtain $p(t)$ [61].

From the impulse response, another important function can be derived: the *integrated squared impulse response* $L(t)$:

$$L(t) = \int_t^\infty p^2(\tau)\, d\tau. \tag{3.31}$$

$L(t)$ is a measure for the amount of sound energy that has yet to arrive at the receiver. Infinity is a theoretical concept again, which in practice signifies the end of the signal. The signal length should at least be $2\, RT_{60}$ to give reliable results [66].

### Reverberation time

The reverberation time $RT_{60}$ in [s] can easily be derived from $L(t)$. Plotting the integrated squared impulse response on a dB scale, a best-fit linear regression line $(t)$ can be drawn to approximate the curve, starting at the 5 dB decay. Typically, $L(t)$ is approximately linear on a logarithmic scale up to a decay of 60 dB or more. In this case, the linear regression line is to be constructed using the data points up to the 65 dB decay. Subsequently, the reverberation time $T_{60}$ is the time it takes $l(t)$ to drop 60 dB in strength.

If the noise level is too high to observe the 60 dB decay in $L(t)$, a shorter estimate period of time has to be used. For example if only the first 45 dB decay is accurate, $l(t)$ has to be constructed of the data points between -5 dB up to the 45 dB decay. The line can be extrapolated to find the theoretical 60 dB decay time, which is then called $T_{40}$, as it is based on a 40 rather than 60 dB decay. In the same way, the reverberation time $T_{30}$ or $T_{20}$ can be computed from the data between -5 dB and -35 dB, and -5 dB and -25 dB, respectively. Note that the index does not refer to the decay of the reverberation (which in all cases corresponds to a 60 dB decay), but rather to the way in which this result was derived. In all cases, the same measurements need to be done several times in order to average out statistical errors. Depending on the intent of the impulse response measurements, it may be required to also take multiple measurements at different positions, in order to get a more representative depiction of the entire space.

### Sound strength

The sound strength $G$ (in [dB]) of a signal is the level of the signal at the receiver point, compared to what would be heard at a 10 meter distance from the source in a free field (i.e. an anechoic chamber). It can be computed as follows:

$$G = 10 \log_{10} \frac{\displaystyle\int_0^\infty p^2(t)\,dt}{\displaystyle\int_0^\infty p_{10\mathrm{m}}^2(t)\,dt} \tag{3.32}$$

where $p_{10\mathrm{m}}(t)$ is the impulse response measured at a 10 meter distance in anechoic conditions.

### Early decay time

The *early decay time* (EDT, in [s]) is an important parameter for the subjective perception of acoustics. According to the ISO 3382, it is “related to the to perceived reverberance, while $T$ is related to the physical properties of the auditorium” [66]. The EDT can be determined in a similar way as the reverberation time. A linear least-squares fit line can be drawn through the part of $L(t)$ in which it drops 10 dB in strength. The decay time can be computed from the slope.

### Early-to-late arriving energy

The *early-to-late index* $C$ (from ‘clarity’) in [dB] represents the ratio of early versus late arriving sound energy. It is computed from the impulse response as follows:

$$C_{t_e} = 10 \log_{10} \frac{\displaystyle\int_0^{t_e} p^2(t)\,dt}{\displaystyle\int_{t_e}^\infty p^2(t)\,dt} \tag{3.33}$$

where $t_e$ is the early time limit. This is generally chosen at 50 ms for speech applications and 80 ms for music.

### Centre of gravity

The *centre of gravity* of the impulse response $T_c$ (in [s]) is defined as the weighted mean of the squared impulse response:

$$T_c = \frac{\displaystyle\int_0^\infty t\, p^2(t)\, dt}{\displaystyle\int_0^\infty p^2(t)\, dt} \tag{3.34}$$

This quantity $T_c$ relates to “the balance between clarity and reverberance” [66].

### Lateral energy fraction

The *lateral fraction* (LF) represents the fraction of sound in the early part of the response that does not come from the direct sound. Another impulse response $p_L(t)$ is therefore recorded in such a way to record as little direct sound as possible. This is done by using a figure-of-eight microphone with its null pointing at the sound source. $LF$ can then be computed in the following way:

$$LF = \frac{\displaystyle\int_{0.005}^{0.080} p_L^2(t)\, dt}{\displaystyle\int_0^{0.08} p^2(t)\, dt}. \tag{3.35}$$

The integration limit of the integral in the numerator starts only at 5 ms, to make sure no direct sound is taken into account.

## 3.2.4 Time-frequency decay graph

Pätynen *et al.* [154, 200] argue that omni-directional measurements and acoustic parameters are not sufficiently descriptive to describe the acoustic of large spaces. They propose two additional ways of qualifying and visualising acoustical properties of an impulse response. The first one is a time-frequency decay graph. At regular intervals (they suggest between 20 and 200 ms) the RIR is windowed and transformed into the Fourier domain. This allows for plotting the magnitude of the frequency spectrum at subsequent time steps, which can all be plotted in the same graph. Fig. 3.5 shows an example of such a frequency decay graph.

A second visualisation proposed by Pätynen *et al.* [154] is of the spatio-temporal kind, as they observed that the largely direction-ignorant parameters by the ISO 3382 are insufficient to describe the characteristics of concert halls. They describe a way to get a spatial directivity pattern of the impulse response using time difference of arrival between several impulse responses from microphones positioned closely to one another. Their precise algorithms

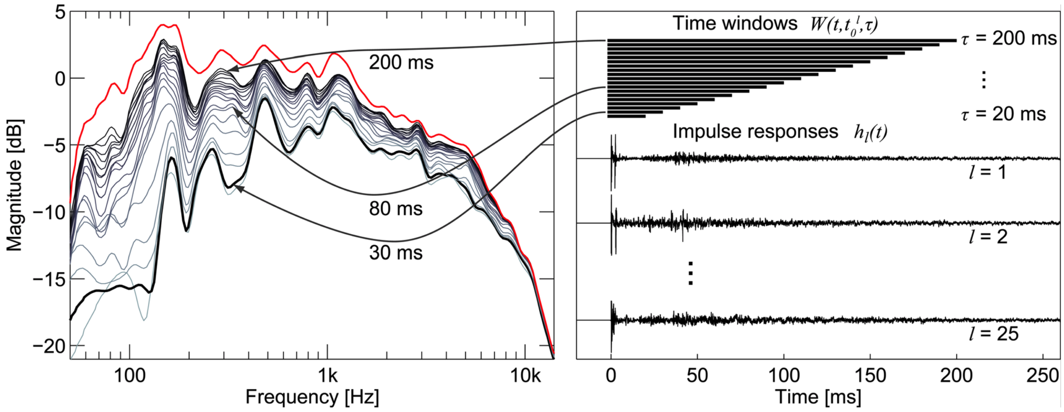


**Figure 3.5:** The time-frequency decay graph of an impulse response. Each curve in the graph on the left is the magnitude response of the time-windowed impulse response, thus forming an energy decay plot. Image from [154] with permission.

for this are beyond the interests of this thesis, as directional analysis of impulse responses is not central to this thesis.

### 3.2.5 Schröder frequency

It was shown in earlier sections (Sec. 2.5, Sec. 2.1.4) that standing waves (i.e. room modes) and diffraction effects are prominent wave effects resulting from the wave equation. However, at higher frequencies, they become more and more indistinguishable: room modes become more densely distributed within the octave, and diffraction effects are inversely proportional to the wave- length (see Eq. 2.60). There is therefore a certain frequency (though not particularly well-defined) at which waves can conceivably be modelled in a geometric rather than a wave-like fashion. This cut-off point is called the *Schröder frequency* [174, 125], first proposed by Schröder and Kuttruff, and reads as:

$$f_S = 2000\sqrt{\frac{RT_{60}}{V}}, \tag{3.36}$$

where $RT_{60}$ and $V$ are the reverberation time and the volume of a space, respectively. This formula is derived from the theoretical overlap between modes, though the precise amount of overlap is not exact. Though Schröder [173] initially proposed a more conservative formula using a factor of 4000, Schröder and Kuttruff [174] later corrected this to the above formula, arguing that “Measurements by various authors have shown that the theory is actually valid for frequencies as low as [red, 2000 $\sqrt{\frac{RT_{60}}{V}}$].”

## 3.3 Summary

This chapter discusses the fundamental physical and room acoustics concepts important to this thesis. The concept of impedance is introduced and will be used extensively in later chapters. Boundary acoustics are introduced, i.e. rather than considering merely the free field, the interactions with and through different mediums are addressed. This is important as all room acoustic problems require some way of modelling a finite enclosure. Additionally, high-level methods to concisely present room acoustic information have been discussed: the ISO 3382 and the time-frequency decay graphs. Finally, the concept of 'Schröder frequency' is introduced. This is a useful concept in the light of the next chapter, as this frequency is used to inform the decision whether to use geometric or wave-based acoustic modelling methods.

# Chapter 4

# Room acoustic modelling

**Contents**

In the scope of this thesis, the goal of room acoustic modelling is to obtain an accurate impulse response from a model of an environment. Such a model can either by physical, such as an architectural scale model, or virtual, for instance those constructed in *computer-aided design* (CAD) programmes. In this work, the emphasis lies on virtual acoustic modelling, and physical architectural models will only briefly be discussed in Sec. 4.2. The RIR may be made audible statically by convolution of anechoic material, or dynamically via real-time feedback in virtual environments.

The scope of acoustic modelling has changed over the last couple of decades: whereas early modelling methods mainly occupy themselves with finding numerical acoustical properties (such as reverberation time) of spaces, later attempts have tried to generate a full impulse response from source to receiver. In the last decade, real-time acoustic modelling has become a feasible pursuit. In real-time acoustic modelling, the 'listener' can move through a virtual environment, and the acoustic feedback adapts to the listener's position and orientation while moving through the (virtual) space. From an academic point of view, the objective is often to find the most accurate impulse response that represents reality best, whereas in some commercial applications such as game audio, the aesthetics of the (potentially real-time generated) IRs are more important. In this thesis, the emphasis of acoustic modelling lies on physical accuracy rather than aesthetically pleasing IRs—though they can of course go hand-in-hand. Standardised objective metrics have been discussed in the previous chapter.

This chapter discusses the most common acoustic modelling methods. Sec. 4.1 introduces some computer science-related terminology required for the understanding of the rest of the chapter. Sec. 4.2 deals with the earliest ways of studying room acoustics, for example through physical models or the Sabine and Eyring equations. Different virtual acoustic modelling methods are classified according to a widely used distinction [117]: geometric modelling methods, which assume ray-like properties for sound propagation, and wave-based methods, which try to reconstruct the sound field in a more rigorous way by solving (part of) the wave equation directly. The former are generally used for frequencies above the Schröder frequency, whereas wave-based methods are used for the frequency range below it. Geometric modelling methods are discussed in Sec. 4.3, followed by wave-based ones in Sec. 4.4. Finally, Sec. 4.5 deals with computational benchmarks and comparing acoustic modelling methods.

## 4.1 Asymptotic behaviour

Some essential definitions and terminology related to computer algorithms are necessary to describe, qualify and quantify the behaviour of the algorithms further on in this chapter. *Asymptotic notation* is used as in [44]: a function $f(n)$ is *asymptotically lower bound* (notation $f(n) \in \Omega(g(n))$) if:

$$f(n) > c\,g(n) \qquad \forall n > N, \tag{4.1}$$

where $c$ and $N$ are arbitrary constants. The *asymptotic upper bound* is written as $O(g(n))$, meaning that:

$$f(n) < c\,g(n) \qquad \forall n > N, \tag{4.2}$$

where again $c$ and $N$ are arbitrary constants. This 'big O' notation is often used in relation to the worst-case running time of an algorithm. The strongest form of algorithmic complexity is *asymptotically tight bound*, written $\Theta(g(n))$, meaning that a function $f(n) \in \Omega(g(n))$ and $f(n) \in O(g(n))$. Where possible, this form of asymptotic boundedness is used, as it gives the tightest bound and therefore the most instructive information about the time and memory usage of an algorithm.

## 4.2 Early modelling methods

At the beginning of the $20^{\text{th}}$ century, a scientific interest arose to study acoustics in detail, by means of simulations and scientific measurements. The history of acoustics is of course much longer, evident from the long legacy of theatre and concert venues, and even Greek and Roman theatres in ancient times[63, 41]. However, the experiments by Sabine [161] in the early decades of the $20^{\text{th}}$ century show a keen interest in the specific reasons behind acoustic qualities, and in quantifying some room acoustical parameters. Sabine built models of concert halls, filled them with gas, and took a series of pictures using the *schlieren* method (see e.g. [34]) to visualise sound travelling through his models. He also tried to relate reverberation time to the quality and quantity of the objects in the room. Schröder and Logan [175, 176] famously designed one of the first sets of filters to create artificial reverberation. However, of more interest for this thesis are the qualitative equations as found by Sabine and later derived by Eyring, which are described in the following section.

### 4.2.1 Sabine and Eyring equation

One of the earliest measurements in the field of room acoustics was performed by W. C. Sabine [161] in the late $19^{\text{th}}$ and early $20^{\text{th}}$ century. With an organ pipe and plenty of cushions from the Sanders Theatre as his main tools, he performed pioneering research on reverberation and absorption in room acoustics. He postulated that if $\alpha_i$ is the power absorption coefficient per square meter of some surface $i$ and $s_i$ is its surface area, then the total absorption power $a$ is:

$$a = \sum_{i=0}^{N} \alpha_i \cdot s_i \, , \tag{4.3}$$

where $N$ is the number of different surfaces under consideration. The average absorption $\alpha_{av}$ is then defined as the total power absorption divided by the total surface area:

$$\alpha_{av} = \frac{a}{S} = \frac{\sum_{i=0}^{N} \alpha_i \cdot s_i}{\sum_{i=0}^{N} s_i} \tag{4.4}$$

Sabine experimentally found that the reverberation time $T$ (nowadays more commonly known as $RT_{60}$), defined as the time it takes for a 60 dB decrease in the level of the sound, could be approximated as:

$$T = RT_{60} = \frac{K\,V}{S\,\alpha_{av}} = \frac{K\,V}{a} \tag{4.5}$$

where $V$ is the total volume enclosed by the room in $m^3$. According to Sabine, $K$ is a constant “depending on the initial intensity” which he experimentally found to be 0.164 [161, p. 104]. Although this equation seems to be frequency-independent, Sabine points out that the absorption coefficient can be a function of frequency.

In 1930, Eyring observed that Eq. 4.5 could not be completely correct [56]. In anechoic conditions, with close to no reverberation as the absorption coefficient $\alpha_{av}$ approaches 1, the total absorption $a$ would equal the total surface of the room, reducing the formula to:

$$RT_{60} = \frac{K\,V}{S}\,, \tag{4.6}$$

where $S$ is the total surface of room. This is clearly an incorrect result. Eyring mathematically derived a more accurate formula for the reverberation time:

$$RT_{60} = -\frac{< l > \ln(10^6)}{c\ln(1-\alpha_{av})}\,, \tag{4.7}$$

where $< l >$ is the mean free path between reflections, $c$ is the speed of sound, and ln represents the natural logarithm. The mean free path can generally be approximated as $< l >= \frac{4V}{S}$, although more accurate algebraic expressions can be derived for specific shapes [14, 98]. $V$ and $S$ are again the volume and the surface area of a room, respectively. Hence, Eq. 4.7 expands to:

$$RT_{60} = -\frac{4V\ln(10^6)}{c\,S\ln(1-\alpha_{av})}\,, \tag{4.8}$$

generally called the *Eyring equation*. Setting $c = 343.26$ m/s for the speed of sound at standard temperature and pressure, this reduces to:

$$RT_{60} = -0.16099 \cdot \frac{V}{S\ln(1-\alpha_{av})}\,. \tag{4.9}$$

For small values of $\alpha_{av}$, Eq. 4.5 and Eq. 4.9 give very similar results, because around 0, $-\ln(1-\alpha) \approx \alpha$. However, when $\alpha_{av}$ increases, the reverberation time according to Eyring's equation (correctly) approaches 0, whereas in Sabine's equation does not.

## 4.3 Geometric modelling methods

Geometric acoustic models are based on the assumption that sound propagation can be modelled as (straight) rays. This is an inherently faulty assumption, for the wave character

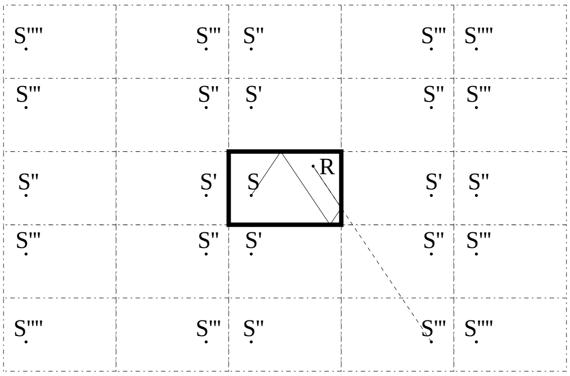


**Figure 4.1:** A visualisation of the image source method. One of the reflection paths of a third order source is shown.

of sound is being ignored. As wave phenomena become more apparent when the wavelength becomes comparable to the dimensions of the objects in the scene, this assumption can only be used for relatively large rooms and high frequencies, i.e. those above the Schröder frequency. Evidence shows, however, that reasonable acoustic predictions can be made even with this simplification. Nevertheless, it is important to keep in mind that geometric models are bound to give poor results for low frequencies. A good overview paper on geometric acoustic modelling methods was presented by Savioja and Svensson [167].

The input of these models is generally a vector scene description, in which all the objects are described by the coordinates of the corners of their boundaries. In several cases, more complex objects such as spheres or cylinder, can be described. Moreover, objects can have an associated absorption coefficient $\alpha$ and scattering coefficient $\sigma$. The following sections review the most widely used techniques in geometric acoustic modelling.

### 4.3.1 Image source method

Although the theoretical background for the "method of images" in acoustics had already been set forth by Eyring in 1930 [56], Gibbs and Jones [73] were the first to describe an algorithm and successfully implement it in a computer programme in 1972. Allen and Berkley [4] were among the first to publish the source code of such an algorithm in 1979. The image source method is based on finding all possible reflection paths from a source to a receiver, treating all walls as mirrors that specularly reflect some fraction of the sound energy. Fig. 4.1 demonstrates this: the source $S$ and the receiver $R$ are located in the original room. The image source method computes the location of all virtual sources (VSs) $S^{(i)}$ up to some order $N$, and their contribution to the impulse response at $R$.

Gibbs and Jones [73] and Allen and Berkley [4] described their algorithms for a 3-dimensional box, in which every wall could have its own absorption coefficient. They proved that their method gave better results than Eyring's equation when compared to measurements. However, they omitted to change the geometry factor related to the mean free path (see Eq. 4.7) to the box-model they were simulating. If they had, the results would have been closer to the measurements. Moreover, since their solution was formulated in terms of the dimensions of the box-model, it did not extend to different shapes. Since any non-rectangular shape will give rise to gaps or overlaps when drawing out the VSs, a more general solution had to be formulated in a different way.

Borish [26] extended the image source method to arbitrary polyhedra up to an arbitrary order of reflections. In order to keep track of all VSs, he proposed a tree structure in which nodes represent VSs, and children of a node are all valid reflections of the parent VS. Moreover, he proposed an extensive algorithm to check validity and visibility of reflections. For example, a path from a VS to a wall that first intersects another wall is considered invalid, so the new VS is considered invalid. However, although this VS might be invalid, its further reflections may be visible for the receiver and thus be valid again. Hence the tree still needs to be constructed further, since the progeny of this VS may be visible. Rather than stopping the programme after a certain order of reflections, Borish decided to stop a VS from reproducing when its distance to the receiver exceeds a certain length, thus allowing the programme to finish after a while.

Because of the necessary extra visibility checks in Borish' algorithm, it is significantly slower than Allen and Berkley's method. However, in both cases a VS produces $n$ new sources, where $n$ stands for the number of walls (and ceilings) that are to be considered. Therefore, the time complexity of both algorithms is exponential in terms of the number of reflections: $O(n^k)$, with $k$ being the number of reflections considered. In the worst case scenario in which all paths are valid, the memory needed is $O(n^k)$ in a breadth-first approach, or $O(k)$ in a depth-first approach.

Significant improvements on this algorithm were made by Lee and Lee in 1988 [122]. By writing all points in space in homogeneous coordinates and representing all reflections as a coordinate transformation [217, p. 4–10], they were able to compute reflection paths by a simple set of matrix multiplications and additions. Since the algorithm represents walls as infinite planes during its computation, however, validity and visibility checks are still necessary for all but convexly shaped rooms. By predetermining wall-to-wall visibility, they made some improvements to speed up this process. In a worst case scenario, this algorithm still has to check all reflections, hence its time complexity is $O(n^k)$. It does not need to store its precomputed data, though, such that the amount of necessary memory reduces to a constant.

In 1993, Kristiansen *et al.* tried to statistically approximate higher order reflections using the data obtained from lower orders [111]. They found that using extrapolation to estimate higher order reflections can work moderately well, provided that the differences in absorption between surfaces are not too large. Commercial software that makes use of the image source method is, for example, Odeon [146] and CATT-Acoustic [49].

An advantage of the image source method is that it can be processed almost completely in parallel by a large number of processors. It is therefore suitable as a way to compute the early reflections in a room. However, because of its exponential growth and because of processes other than reflection dominating, it is not very suitable for high-order reflections. It is a ray-based approach by nature, so its performance quality is limited to higher frequencies. Moreover, only specular reflection is being taken into account, and to the author's knowledge, no implementations exist that try to incorporate sound diffusion or transmission into this method.

### 4.3.2 Beam tracing

The foundations for beam tracing in room acoustics have been laid by Heckbert and Hanrahan [87, 86] and Dadoun *et al.* [46]. The original application of beam tracing was in graphics, as an attempt to speed up the rendering of visual images. The main alternative approach, ray tracing, was found to be too computationally intensive at the time, and beam tracing was more efficient for large, spatially coherent scenes [86]. Because of the rapid development of CPU power and more demanding scenes, however, ray tracing appeared to be more suitable in image and animation rendering. On the contrary, virtual acoustics scenes require much less detail, such that beam tracing might be the more suitable approach still.

Beam tracing exploits the fact that many rays in a 3-dimensional space with little detail will follow very similar trajectories, thus preserving their solid angle. Fig. 4.2 show an example of this sort of behaviour. The first implementations were restricted to planar polygonal models. First of all, an initial beam is defined that covers the entire image space. Secondly, all polygons in the scene are depth sorted according to visibility using a hidden surface removal algorithm [219]. The beam tracing algorithm then arranges all polygons in a tree structure,

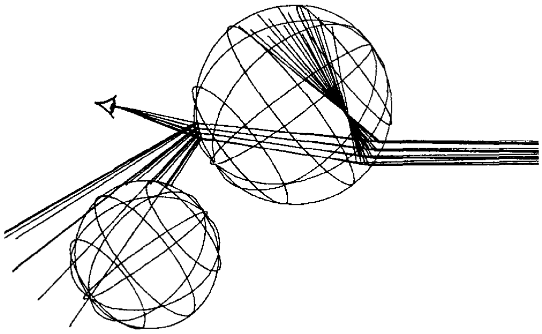


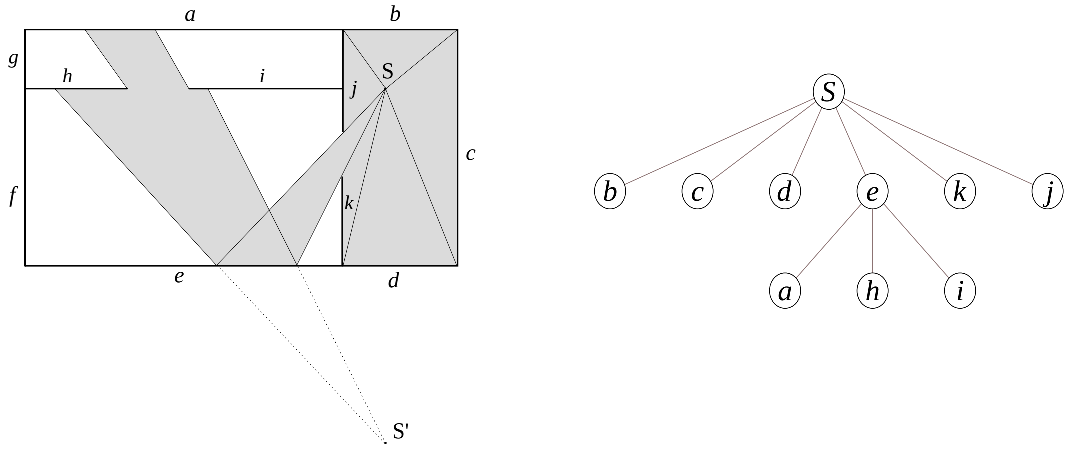
**Figure 4.2:** An incident bundle of rays passing through two spheres. Image from [86] with permission.

called a *beam tree*, according to whether, and if so when, the beam intersects the polygon. For reflective or refractive surfaces, this process may be repeated recursively.

Every time the beam intersects a polygon, the cross section is being cut out, or *clipped*, from the initial beam. The clipped beam then continues its path to the next polygon, until all polygons have been checked for visibility. This clipped beam will generally get a more complex shape after every intersection. Therefore, the beam tracing algorithm has to be equipped with high-level set operation methods that can deal with non-convex polygons with holes, such as those described by Weiler and Atherton [219]. If the intersecting polygon was reflective or translucent, then the process is repeated recursively. A virtual source will be constructed, analogous to the image source method, and the new calculations are performed in the transformed coordinate system. Fig. 4.3 shows a 2-dimensional example of beam tracing and the formation of the beam tree.

Heckbert and Hanrahan point out that the worst case running time of the construction of the beam tree is $O(n^2)$ for 0 reflections, where $n$ is the number of polygons [86]. A particular advantage of beam tracing in graphics, however, is that once the beam tree has been constructed, an image of an arbitrary resolution can be constructed at very little additional cost. A limitation of this approach is that the algorithm can only deal with planar polygons, so all round shapes need to be approximated. Moreover, since only linear transformations are possible because of the homogeneous coordinates, refraction can only be approximated, as this is not a linear operation. Lastly, an advantage of this model is that reflections up to any order are a trivial step of recursion. For each reflection, the same method is being performed in a transformed coordinate system, similar to the image source method.

There exists some confusion about the terminology of the word beam tracing, caused by several publications using the word for slightly different methods. For example, pyramid



**(a)** A beam tracer divides the space into different sections. Only the first and second part of the second iteration are shown.

**(b)** The associated beam tree.

**Figure 4.3:** An example of the first reflections of a beam tracing algorithm and its associated beam tree. Image after [189].

tracing and cone tracing are often referred to as beam tracing methods (e.g. [126, 72, 216]), despite essential differences. Both are somewhat related to beam tracing, but they show more similarities with ray tracing (see also Sec. 4.3.4). This terminological inconsistency caused Monks *et al.* [133] to reinvent beam tracing as the 'new unified beam tracing and image source approach'. They did make some improvements on the existing method, though, adding advanced methods that incorporated angle-dependent absorption, air attenuation, and multidimensional receivers into their model. Also Drumm and Lam [53] designed a very similar algorithm, called the 'adaptive beam-tracing algorithm.' The slight difference in their algorithm was that it could not handle holes in beams, so that they had to split up beams when appropriate. Moreover, they used the information gained at the beam tracing stage for a radiosity method, as will later be described in Sec. 4.3.5.

Funkhouser *et al.* [69] made significant improvements to the beam tracing method, enabling it for interactive spatial auralisation. The algorithm works in four phases: the *spatial subdivision*, *beam tracing*, *path generation*, and *auralisation*. In the first phase, the space is subdivided in planar convex polygons, which are then stored in an adjacency graph. This is then used in the next phase to speed up the process of constructing a beam tree, in a similar way as Hanrahan and Heckbert's method. Recursion accounts for transmission and reflection, and all entries in the tree contain a reference to the polygons they hit. The latter is important in the third phase. This is an interactive phase, in which the receiver can walk through the scene. Via the polygons associated with the receiver position, the beams that hit the receiver can be found. Their combined paths make up the impulse response, which can be auralised in the last phase.

In 1999, more developments in interactive auralisation were made by Funkhouser *et al.* [71] They introduced priority driven and bi-directional beam-tracing to improve computation time. The former method tries to decide on a psychoacoustic basis which beams are most important for the spatialisation of sound. In the latter method, beams are being traced starting at both source and receiver. When they hit the same surface, a valid path has been found and can be auralised. They were able to speed up the process even more by making some mesh simplifications and by making some assumptions on the motion path and speed of the source and receiver.

So far, beam tracing could only handle specular reflection and transmission effects. As a result, a receiver that would pass an open door behind which a sound source was located would experience a rapid increase and decrease in volume while passing it. To smooth out this rapid change in volume, Tsingos *et al.* [204] incorporated frequency-dependent sound diffraction in the model. Sound passing a wedge would diffract according to the uniform theory of diffraction [97, 102], as shown in Fig. 4.4. In 2004, Funkhouser *et al.* [70] summarised all their research that had let to the interactive spatial auralisation of 3D environments.

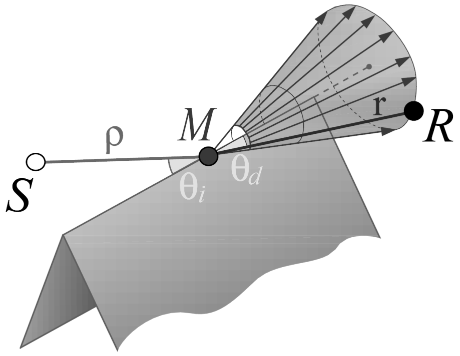


**Figure 4.4:** An incident ray $\rho$ under angle of incidence $\theta_i$ passes a wedge and forms a diffraction cone of new rays in a cone of width $\theta_d$. Image from [204] with permission.

Foco *et al.* [65] developed a way to construct a beam tree in practically linear (instead of quadratic) time in a 2-dimensional space. They mention that the method theoretically extends to 3-dimensional spaces. A few years later, they managed to expand this method to 2D×1D spaces, in other words, 3-dimensional spaces in which all walls are either completely vertical or completely horizontal [9]. In 2010, Antonacci *et al.* [8] extended the 2-dimensional case to also handle diffraction and diffusion.

The theoretical time complexity of this algorithm is debatable. One could say that for every recursive step, $O(n)$ possibilities need to be considered again, thus resulting in the same complexity as the image source method. However, Funkhouser *et al.* showed that in practice this turns out not to be the case [69]. With the help of the spatial subdivision (which runs $O(n^2)$), intersection computations can be done very quickly. For a high order of reflections (approximately $r > 5$), beams become so small they rather resemble rays, thus leading to a small number of intersections and a small branching factor. Hence, the complexity seems to be exponential, but not in terms of the number of polygons, as is the case for the image source method. Fig. 4.5 show the findings of Funkhouser regarding computational complexity.

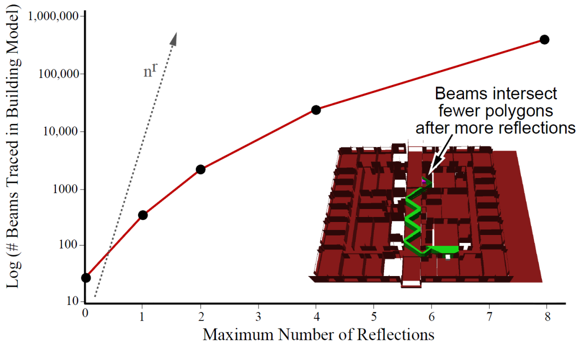


**Figure 4.5:** The computational complexity of Funkhouser's algorithm. Though theoretical analysis predicts $O(n^r)$ time-complexity, test results show that the algorithm performs much better in practice. Image from [69] with permission.

In summary, the beam tracing method provides a valuable algorithm to auralise both static and dynamic scenes. Reflection, transmission, refraction, diffusion, and diffraction can, to a large extent, be implemented. Advanced data structures enable a very quick rendering in real-time and also offer great benefits for static environments. Moreover, almost all routines can be processed in parallel. Since it is a ray-based model, its performance is still limited to higher frequencies. Another limitation is that it can only handle a scene with planar polygons, so any other shapes have to be approximated.

### 4.3.3 Ray Tracing

Ray tracing is commonly known to be the most widely used technique in graphics for visualising 3-dimensional scenes. Essentially, it is based on following rays of light through a space to find out with which object surfaces and light sources they interact. Extending this to room acoustics requires the assumption that sound has approximate ray-like behaviour. One of the first applications of ray tracing in room acoustics was put forward by Allred and

Newhouse [5]. Observing that the factor $< l >$ in Eq. 4.7 is just a geometry factor related to the mean free path, they tried to use ray tracing to more accurately measure this factor. They described a computer model of a rectangular room with a sound source, from which rays were shot in quasi-random directions. For each ray, the path length was computed.

A different approach was proposed by Krokstad *et al.* [113]. They computed the reflection path of rays that were radially shot from a source, and measured the number of the rays that hit the wall with their intersection positions. They displayed their results in a 'room response' diagram, a 2D image of the room with dots at the intersection points. Also, they placed virtual receivers with a large diameter in the area, which detected but did not influence intersecting rays [112]. In such a way, they could display some sort of an RIR.

Kulowski [114] formalised the concept of acoustic ray tracing in 1985. He stated its fundamentals and proposed several extensions to model phenomena such as frequency-dependent reflection, damping due to air, and source directionality. Wall absorption can be implemented simply by a multiplication factor at each hit. The absorption coefficient $\alpha$ can be defined in the way as demonstrated by Sabine, as the proportion of sound energy that is transformed into heat. Thus, the outgoing ray energy will be $E_{out} = (1 - \alpha)\, E_{in}$. Moreover, Kulowski points out that the ray energy should not decrease according to the inverse square law but remain constant. When the model takes into account air absorption, the ray travelling from point $\mathbf{x_i}$ to $\mathbf{x_{i+1}}$ should be $E_{\mathbf{x_{i+1}}} = E_{\mathbf{x_i}}\, e^{-m|\mathbf{x_i}-\mathbf{x_{i+1}}|}$, where $m$ is an air absorption factor. Frequency dependency can be added to this by dividing the energy spectrum into frequency bands and making the absorption factors dependent on frequency. Directional sources can simply be modelled by specifying a spatial distribution of initial rays. In 1989, Ondet and Barbry [149] thoroughly described the state of the art algorithm in acoustical ray tracing.

Vorländer [211] proposed another method to model absorption, called *controlled particle annihilation*. After a ray hits the wall, a random number is compared with the absorption coefficient $\alpha$. If the random number is lower, the ray ceases to exist. Otherwise, the ray continues to exist with the same energy it had previously. Although this method yields worse results for a small number of rays, for a large number of rays it is equivalent but faster. Vorländer [212] also used ray tracing to speed up the image source method (see Sec. 4.3.1): each time a ray hits the receiver after some number of reflections, it is certain that it completed a valid reflection path. Hence, the image source of this particular wall sequence can be computed without a need to check its validity. As a consequence, however, the image source does not search (and thus will not find) all the possible paths up to some order, so the new method does not show many advantages in practice. A variation of this method was adopted in ODEON 2.5 [140].

In 1993, Lehnert [123] published an updated overview of the state of the art of ray tracing in acoustics, and addressed systematic errors of the algorithm. He mentioned the properties of different detector shapes, their ideal size, and the aliasing errors involved. In 1996, Lam [119] discussed several improvements that had been made to incorporate diffusion into the model. Lambertian reflection has been implemented by Hodgson [88], based on the derivations of Kuttruff [115]. In this method, a part $\sigma$ of the sound is scattered in all directions according to some proportionality factor dependent on the angle, with the rest of the energy reflected specularly.

Naylor [140, 141] describes a related method, implemented in ODEON: the first $N_{TO}$ reflections are purely specular. $N_{TO}$ is a user-defined parameter which stands for the *transition order*, set equal to 2 by default. The reflections after $N_{TO}$ are purely Lambertian. For each of these reflections, the energy that goes directly to the receiver can be calculated, such that it can be added to the impulse response.

Savioja *et al.* [165] created one of the first interactive virtual acoustic environments that included ray tracing. They added a method to simulate diffusion in order to create a more realistic environment. In their model, a ray is either reflected specularly or sent in a random direction. The ratio of reflected versus random rays is determined by the scattering coefficient $\sigma$. Lentz *et al.* [124] also incorporated a ray tracing algorithm in a virtual reality system to compute the diffuse sound field. They used a precomputed binary-space-partition (BSP) tree [92] to enable real-time computation even for very complex environments.

Ray tracing has proved to be a fast method in virtual acoustic modelling. Its speed is linear with the number of rays and, depending on implementation, generally grows logarithmically or linearly with the number of objects in the scene. This is provided that all rays are being used, such as in ODEON's implementation [141]. If a large number of rays do not reach the receiver, however, many calculations will be in vain. Little computer memory is required for ray tracing. Apart from the obvious scene data, a small and constant amount of data needs to be stored for each ray, and it can be thrown away when the next ray path is computed. Some memory space is required only when all reflection points are stored, linear with number of rays and reflections.

Theoretically, not only polygons but also disks, cylinders, spheres, and ellipses can be rendered, since algebraic expressions exist for these shapes. Some research by Vercammen [210] addresses this, but to the author's knowledge no attempt has been made to implement this extensively in common acoustic modelling tools. It is questionable whether this would improve the performance significantly. Using ray tracing, transmission and refraction can be easily implemented, although to date this has not been done in virtual acoustic modelling. Reflection and diffusion can be easily modelled in a variety of ways. Disadvantages of ray tracing are that diffraction cannot accurately be modelled, and that there is a chance that important reflection paths may be overlooked.

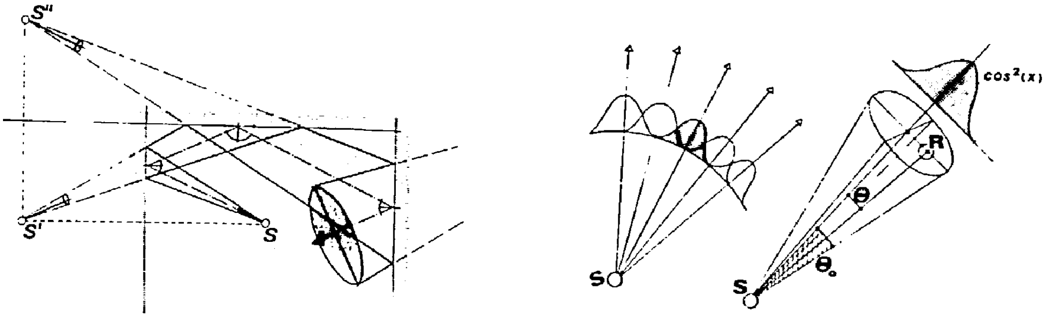
### 4.3.4 Pyramid and cone tracing

Cone and pyramid tracing combine the ideas of ray tracing and beam tracing. They use the idea of ray tracing in the sense that a finite number of rays are cast into space from the source. However, each of these rays now has a volume associated with it, modelled as either a cone or a pyramid (or more generally: beams). The idea behind this, as with beam tracing, is to cover the entire room such that no detail will be overlooked. Since only the centre of the cone is tracked, there is no need for the additional beam tree data structure.

One of the first applications was in graphics, proposed by Amanatides [7]. He showed that cone tracing could in some cases be more effective than the traditional ray tracing, especially for high resolutions. A simple example is shown in Fig. 4.6a. A difference with beam tracing is that in cone/pyramid tracing, an initial number of beams needs to be defined. Amanatides describes a method in which each cone covers an entire pixel, thus covering the entire screen with some overlap between beams. For each beam, a list of eight closest visible objects is created and for each object, the fractional coverage is computed. This process can be repeated recursively to account for reflection and refraction. A pixel is then formed by the weighted sum of all the objects in the list.

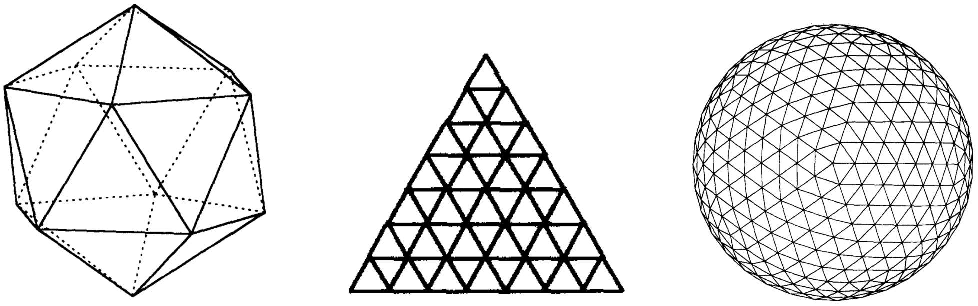
Van Maercke [205] introduced an implementation of the cone tracing mechanism in the field of acoustics. He showed that intersection computations for a cone with a point or plane are relatively simple [206]. A receiver can be modelled as a point, such that the energy transfer is related to the distance from the receiver to the source, the number of reflections (and associated absorption), and the place in the beam. Fig. 4.6b shows a possible energy distribution within the cone.

A very similar idea is proposed by Lewers [126]. The main difference lies in the subdivision of the source into pyramids instead of cones, as demonstrated in Fig. 4.7. A small advantage of this is that pyramids can cover the space exactly, without overlap. However, since neither



(a) Cone propagation and reflection.
(b) Energy distribution within a cone.

**Figure 4.6:** A cone tracer showing the first three reflections and the corresponding virtual images. Each cone has an associated energy distribution. Images from [205] with permission.

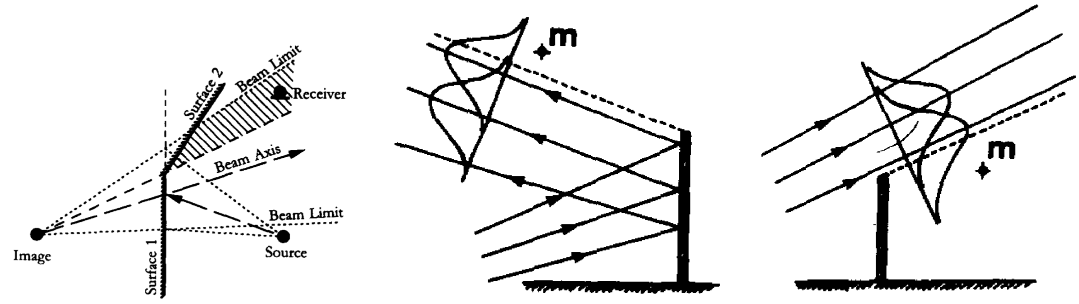


(a) An icosahedron.
(b) Subdivision of triangle.
(c) Subdivision of the whole source.

**Figure 4.7:** The subdivision of a source into different pyramids. Images from [126] with permission.

method splits up the beam after 'partial intersections' with objects, both are prone to several sorts of errors. In some cases a receiver will pick up beams that should not reach the receiver, and in some cases it does not register beams that should have reached the receiver. Naylor [141] showed that the number of valid detections is always underestimated. Farina [58, 57] wrote the programme RAMSETE that is purely based on pyramidal cone tracing and showed that its results very well matched measurement data of reverberation time [60]. Stephenson [194] introduced beam splitting for pyramid tracing, and his *Quantized Pyramidal Beam Tracing* algorithm deals with specular and diffuse reflections, and also with diffraction around corners.

Dalenbäck [47] argued that a point receiver and a cone that increases in size with time is equivalent to a (1-dimensional) ray and a receiver with a radius that increases with time. This variation makes ray tracing effectively equivalent to cone tracing. Moreover, he devised another method to model diffusion effects in the cone tracing algorithm. He accounts for diffusion by creating a secondary source every time a ray hits a wall with $\sigma > 0$. From this secondary source, $n_r$ new rays are fired in random directions according to Lambert's law, each



(a) (b) (c)

**Figure 4.8:** Several errors related to pyramid and cone tracing: in every example, the receiver registers a hit related to an invalid path. Fig. 4.8a from [126] with permission, Fig. 4.8b and Fig. 4.8c from [206] with permission.

carrying some of the energy. A specular ray is reflected as usual, carrying $E_{\text{in}}(1-\alpha)(1-\sigma)$ energy. Although he used some acceleration methods for more efficient computation, the multiple rays for each diffuse surface theoretically result in an exponential time-complexity. Therefore, in a later version [48], the reflected ray is a single one, either specular or randomised, depending on $\sigma$. This is the *randomised tail-corrected cone-tracing method*, used in the earlier versions of CATT-Acoustic. The methods used in later versions are relatively similar but employ the image source method to account for early reflections.

Both pyramid tracing and cone tracing have proven to be successful tools in acoustic modelling, as both are used in commercially available software. Despite some computational inaccuracies, their fast computation time and relatively reliable results make them appropriate prediction methods for mid- and high-frequency acoustics. As they assume a large number of initial beams which do not split up, the complexity is linear in time, instead of exponential such as in beam tracing. Depending on the precise implementation, the spatial complexity will be at most linear in number of objects (i.e. triangles) and number of rays. The algorithm allows for parallel processing, since all cones/pyramids travel independently from each other. Scene objects are not necessarily restricted to planar polygons, although in practice this generally is the case. The method takes into account reflection and diffusion considerably well. Transmission and refraction can be implemented with the same approximations as Dalenbäck made: by treating the beams as rays and the receiver as an expanding sphere. Diffraction, however, cannot be modelled in this way. Using ray tracing as a means of achieving real-time interactive acoustics is an area of active research [40]. A recent development in this field is the release of NVIDIA's *VRWorks Audio SDK* [145] (announced but yet to be released at the date of this writing), which claims to use aforedescribed methods to generate impulse responses in real-time, particularly targeting interactive Virtual Reality applications.

### 4.3.5 Radiosity method

The radiosity method (also radiant exchange or acoustic radiance transfer method) is a geometrical prediction model largely based on diffusion. The theoretical background has been laid out by, amongst others, Kuttruff [115, 117], Kayija [94], and Carroll and Miles [39]. They are all based on what has come to be called the *radiosity equation* or *rendering equation* [94, 116]. This equation relates to the energy exchange of all the surfaces with one another. Since it is an integral equation which cannot be solved exactly, the radiosity method is a way to approximate the energy exchange between finite surfaces. More specifically, when dividing the object space in a finite number of patches, the energy $B$ radiating away from some patch is, according to Shi *et al.* [181]:

$$B_i = E_i + \rho_i \sum_{j \neq i} B_j F_{i,j} \,, \tag{4.10}$$

where $B_i$, $B_j$ are the radiosity of some patch $i$ and $j$, $E_i$ is the energy patch $i$ emits itself, $\rho_i$ is the reflectivity of a patch $i$ (generally $\rho = 1 - \alpha$), and $F_{j,i}$ is the fraction of energy from a patch $j$ that reaches patch $i$. The variable $F_{i,j}$ is called the form factor, and is related to the geometry of the two patches, the distance between them, the air attenuation, etcetera. The challenge in most applications is to find an appropriate form factor for each pair of patches.

Lewers [126], as one of the first to implement the radiosity method in acoustics, obtained the form factor by beam tracing. He divided each plane into grid points, from which $N$ beams with solid angle $\Delta\Omega$ were fired into a hemisphere. All intersections with other patches were registered such that the total cross section was a measure $F$. Also, the distances from all patches to one another were averaged in order to establish a measure for the time delay. Lewers established the form factor in the following way:

$$f_{i,j} = \frac{\Delta\Omega \, \Delta A_i}{A_i \, \pi} \sum_{m}^{M} \sum_{n}^{N} \cos\theta_{i(m,n)} . \tag{4.11}$$

The two summations run over the number of grid nodes $M$, each covering an area $\Delta A$ of the total area $A$, and the number of beams $N$. Computing all the form factors should be an off-line process that can be time consuming. This is an investment, though, as now for

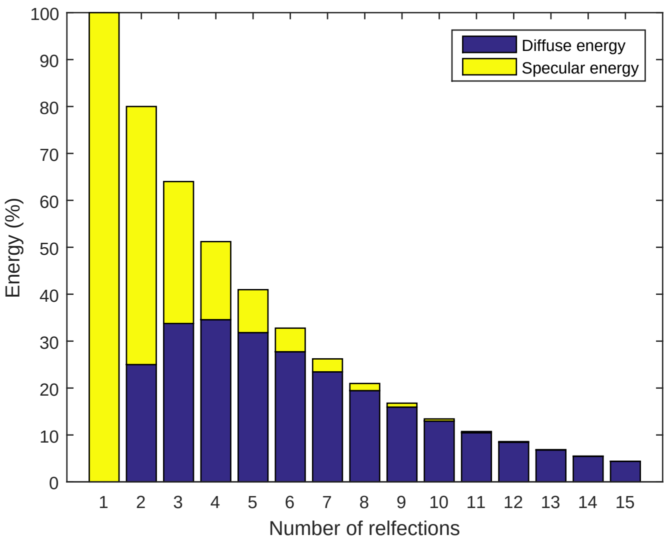


**Figure 4.9:** The total amount of scattered and specularly reflected energy in a room, in which $\alpha = 0.2$ and $\sigma = 0.25$. The horizontal axis shows the number of reflections. Image after [116].

any source and receiver an impulse response can be constructed via straightforward visibility checks and multiplications. Since the radiosity method is also a frequently used method in computer graphics, Shi *et al.* [181] devised an algorithm that could do graphical and audio rendering at the same time. They also incorporated air attenuation to account for absorption by the air.

So far, surfaces have been assumed to portray only absorption and Lambertian reflection. Kuttruff [116] derived a theoretical method to also incorporate specular reflection in the radiosity equation. Accordingly, Fig. 4.9 shows the amount of diffuse versus specular sound energy in a room after a certain number of reflections.

Tsingos [203] devised a method to more easily deal with complex scenes. By means of a hierarchical scene approach—i.e. ordering the scene objects in a 4-tree—he developed a method to more efficiently render the sound, at the expense of a more costly preprocessing. Moreover, he added the possibility of directional diffusion to the model, i.e. the preference of a surface to scatter energy in a certain direction. In 2000, Le Bot and Bocquillet [121] proved on a formal theoretical basis that the radiosity equation and Monte Carlo ray tracing should yield equivalent results under certain assumptions, and backed this up with a lot of experimental data of various test cases. The computation time of the both methods was similar, though, and they did not express a preference towards either. In a following paper, Le Bot [120] extended this method with some ideas from the image source method and ray tracing to come to a more general radiosity method, which also modelled specular reflection, and even refraction.

In 2002, Nosal [143] completed a master's thesis in which she laid out the state of the art of radiosity method algorithms. She also devised a quick method to estimate late radiation rather than compute it exactly [144], since this may otherwise be a costly process. Reboul *et al.* [157] invented a way to model diffraction. Using the Geometrical Theory of Diffraction [97], they tried to take into account the energy transfer that is diffracted into a cone by a wedge (see Fig. 4.4). They compute diffraction in frequency bands, but point out that their model may still not be valid for low frequencies.

Siltanen *et al.* [183] added minor improvements to the existing model by allowing the diffusion and specular reflection to be dependent on angle, stored in a *bidirectional reflectance distribution function* (BRDF). This allowed for a more precise definition of material properties, at the small cost of storing this data in BRDF matrices. Siltanen and Lokki [182] subsequently added diffraction to this model, based on the Biot-Tolstoy solution [22] of diffraction modelling. At all surfaces by an edge, contributions of diffraction are calculated. These are then stored in a *diffraction BRDF*, which works essentially the same as the ordinary one.

The radiosity method works in several stages. A preprocessing stage is needed to compute the influence of all patches on all others. However, since patches need to be small in order to make a reliable approximation, a subdivision algorithm is generally required first. Dividing the space into $N$ small patches, the spatial complexity is at least $O(N)$. A typical case could easily take up hundreds of megabytes, according to Siltanen [183]. Computing the contribution of all patches to each other is typically done in $O(N^2)$. This can be reduced with a hierarchical approach in which objects and patches are stored according to visibility. For $m$ objects split up into a total of $N$ patches, computation time will then be $O(mN)$ [143, p. 63]. Especially for large environments, this can be a significant reduction. Constructing the visibility tree can take up to $O(m^3)$ time [203, p. 121]. The final stage is to determine the paths from source to receiver. Without the tree, all patches have to be checked at least once, hence resulting in a $\Theta(N)$ algorithm. With the visibility tree, this can often be improved significantly, as the lookup is at best $\Omega(\log_2 m)$, but at worst $O(N)$ still.

The advantage of the radiosity method is that, once the preprocessing has been done, it is general for all receiver positions. If the method of collecting the contributions from all patches can be done fast enough, it is suitable for interactive applications. Lambertian diffusion is the main phenomenon described by this method, but successful attempts have been made to incorporate specular reflection, transmission and refraction, and even diffraction. However there is no unified approach that incorporates all of these in one algorithm. A disadvantage of the algorithm is the necessary subdivision of surfaces, which greatly increases computation time and required memory.

## 4.4 Wave-based modelling methods

Wave-based room acoustic prediction models are based on numerically solving the wave equation, or parts of it. Most of the approaches present a time domain solution, but some methods use transformations to find an RIR in a different way. Generally speaking, they require more computational power and memory than most geometrical methods. Most of the 3D applications of wave-based methods have therefore been developed relatively recently, over the last 10 years, with computer power and memory on commercially available computers ever increasing. They often intrinsically account for wave effects like interference and diffraction, but because of their rigorous nature, real-time 3-dimensional modelling is often not feasible.

### 4.4.1 Digital waveguide

Smith [185, 186, 187] was among the first to describe and use *digital waveguides* (DWGs), initially mainly for physical modelling purposes. It is based on modelling the propagation of a wave through a delay line. The link between instruments (and string instruments in particular) and DWG is often relatively intuitive, which makes the method especially suitable for the physical modelling of instruments [188]. However, its method has also been expanded to 2- and 3-dimensional models [54, 138].

For a string, the homogeneous wave equation Eq. 2.12 approximately holds, where $c = \pm\sqrt{\frac{\kappa}{\mu}}$, and $\kappa$ is the tension force along the string, and $\mu$ is the mass density. The DWG method uses a discretised version of the d'Alembert solution to the wave equation (see Eq. 2.16). In order to discretise the system in intervals of time and space, $t$ is set to $n\,T$ and $x = m\,X$, with $n, m \in \mathbb{N}$. Ensuring that $T = \frac{X}{c}$ allows for writing the displacement as:

$$\begin{aligned} y(x,t) &= y^{+}[(n-m)T] + y^{-}[(n+m)T] \\ &= y^{+}[n-m] + y^{-}[n+m] \end{aligned} \tag{4.12}$$

where $T$ is ignored for convenience. Fig. 4.10 shows how this idea can then be implemented as a forward and backward delay line. In a similar fashion, it is also possible to couple multiple

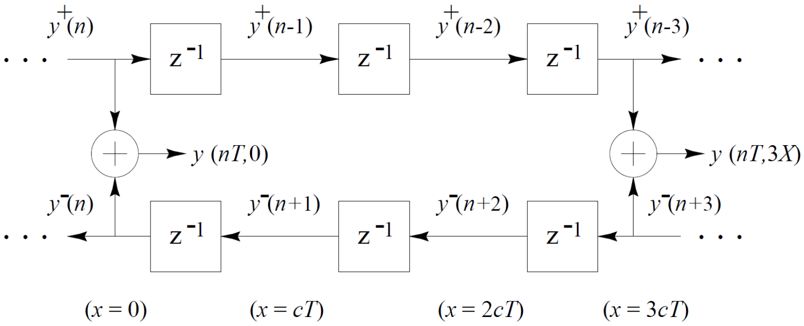


**Figure 4.10:** A digital waveguide represented as a set of delay lines. Image from [187] with permission.

strings of, for example, a violin with each other via another delay line that represents the bridge [188]. The place where several strings intersect is called a *scattering junction*, and is bound to meet some mathematical requirements as described by Van Duyne and Smith [54].

Smith [187] points out that the model is in fact an exact solution to the wave equation at all sampling instants if the travelling wave is bandlimited to the Nyquist frequency. Moreover, he showed that not only the displacement $y(x, t)$ can be computed in this way, but that also *velocity* $v(x, t)$, *acceleration* $a(x, t)$, and *force* $f(x, t)$ can be transmitted through a digital waveguide in a similar fashion. He also described ways to include (frequency dependent) air attenuation, boundary conditions, and ways of simulating a plucked string or a blown reed.

The DWG method can be extended to multiple dimensions [54]. In the two-dimensional case, the displacement at $y(x, y, t)$ is dependent on an infinite number of points in 360° around it. Since this is impossible in any computer application, the 2D space needs to be rasterised. Van Duyne and Smith [54] rasterised the space with a square grid and approximated the acceleration at $a_{x,y}[t]$ at point $(x, y)$ with a standard difference approximation as well as with a five-point stencil method:

$$\begin{aligned} a_{x,y}[t] &= \frac{\partial}{\partial t} v_{x,y}[t] \approx v_{x,y}[t+1] - 2v_{x,y}[t] + v_{x,y}[t-1] \\ &= \frac{1}{2}\left(v_{x+1,y}[t] + v_{x-1,y}[t] + v_{x,y+1}[t] + v_{x,y-1}[t] - 4v_{x,y}[t]\right). \end{aligned} \tag{4.13}$$

This equation can then further be manipulated in order to obtain $v_{x,y}[t+1]$ at all points in space. An important issue they bring forward is the problem of *dispersion error*. Because of the rasterisation, waves cannot travel equally fast in all directions. The more parallel to the grid axes the wave motion, the slower it propagates compared to actual wave motion. This effect, known as the dispersion error, becomes more prominent at higher frequencies. (A more detailed discussion of dispersion effects will be presented in Chapter 5). Fig. 4.11

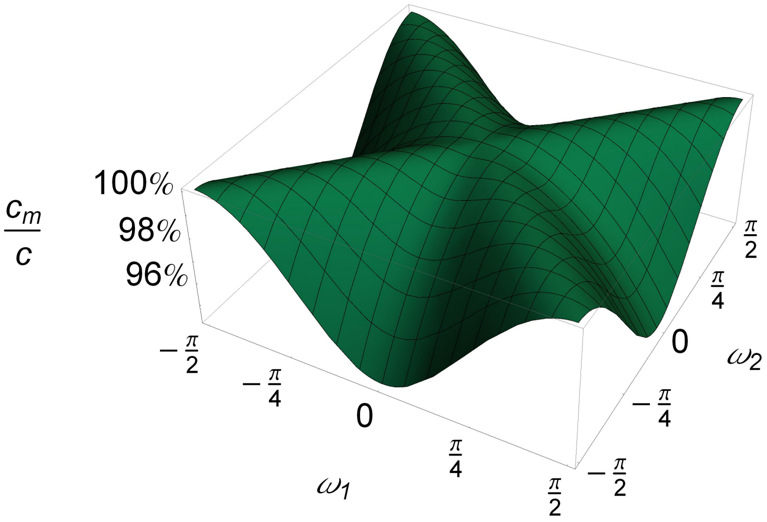


**Figure 4.11:** The dispersion error $\frac{c_m}{c}$ in percentages, in 2 dimensions. $\omega_1$ and $\omega_2$ represent the spatial frequencies in both directions. Image after [168].

plots the wave speed in the model $c_m$ versus the actual propagation speed $c$, as a function of direction and spatial frequency [168]. There are alternatives to the square grid: for example a triangular grid structure in 2D space or a tetrahedral grid in 3D space [55, 137, 138]. Dispersion error in these meshes may be less, but will always be present.

The two-dimensional DWG method has been used for a variety of physical modelling purposes, for example for modelling drums [2, 118] or the vocal tract [135, 136]. Moreover, the method can be extended to three-dimensional spaces and can thus be used in acoustic modelling. For example, the software RoomWeaver was designed by Beeson and Murphy [15] as a tool to produce an RIR from a user defined environment. Hacıhabiboğlu *et al.* [79] devised a method to implement directional sources in a 3D DWG mesh. Southern and Murphy [190] proposed a method for encoding DWG impulse responses into a second order B-format ambisonic recording, and Hacıhabiboğlu *et al.* [78] presented a way of simulating receivers with a directional frequency-dependent response. Shelley and Murphy [179] showed a way of modelling diffusely reflecting surfaces in 2D. Shelley [180] published the extension to 3-dimensional spaces.

The DWG method has proved to be a valuable means of physical modelling. The question is whether it is also fit for room acoustic simulation. In contrast to geometric modelling methods, the speed of the algorithm is not dependent on the number of objects in the room, but on the actual size they represent. The entire modelling space needs to be subdivided into grid points. The precise number of points depends on the size of the model, the waveguide size, and the required accuracy. It seems that in literature, a typical 3D DWG mesh contains a number of nodes of the order-of-magnitude $10^5$ (e.g. [15]). The algorithm is linear in terms of air nodes, which all need to be updated every sample. Hence the algorithm can be considered to be of $\Theta(N_g \cdot T)$ in time and $\Theta(N_g)$ in space, where $N_g$ represents the number of nodes in the mesh and $T$ represents the length of the impulse response in samples.

Most acoustical phenomena can be modelled by the DWG method. A major advantage is that diffraction is inherently modelled in this wave-based method, in contrast to geometric approaches. Reflection and absorption can be implemented in a relatively straight-forward way. Also transmission is not a theoretical restriction. Refraction is considerably more difficult to implement, though, and to the author's knowledge no attempts have been made to incorporate this in a state of the art method. Diffusive boundaries belong to the possibilities, as do directional sources and receivers.

### 4.4.2 Finite difference time domain method

The Finite Difference Time Domain (FDTD) method is another wave-based technique, which aims at computing the sound field in the time domain. In free air, it can be seen as a stencil method rather than as a waveguide method, although the methods have been shown to be equivalent [95]. A detailed derivation of the coefficients of the stencil have, for example, been derived by Botteldooren [31, 32]. Rather than formulating the update equations in terms of forwards and backwards travelling waves, explicit update formulas are derived for all nodes. The difference operators in the wave equation can be approximated using a centred difference scheme:

$$\begin{aligned}\frac{\partial^2}{\partial x^2}p[x,t] &= \frac{1}{X^2}\left(p[x-1,t]-2p[x,t]+p[x+1,t]\right)+\mathcal{O}(X^2)\\ \frac{\partial^2}{\partial t^2}p[x,t] &= \frac{1}{T^2}\left(p[x,t-1]-2p[x,t]+p[x,t+1]\right)+\mathcal{O}(T^2)\end{aligned} \tag{4.14}$$

Using these approximations by ignoring the higher order terms, the homogeneous 3-dimensional wave equation can be approximated as:

$$\frac{p^{t-1}_{l,m,n}-2\,p^{t}_{l,m,n}+p^{t+1}_{l,m,n}}{c^2\,T^2}=\frac{p^{t}_{l-1,m,n}+p^{t}_{l+1,m,n}+p^{t}_{l,m-1,n}+p^{t}_{l,m+1,n}+p^{t}_{l,m,n-1}+p^{t}_{l,m,n+1}-6p^{t}_{l,m,n}}{X^2}, \tag{4.15}$$

where $X$ and $T$ are the grid spacing and the time step, respectively. Solving for $p^{t+1}_{l,m,n}$ and recognising that the neighbouring points are the faces of the cube that surrounds this point, we obtain an explicit update formula for the pressure at a point:

$$p_{l,m,n}^{t+1} = \left(2 - 6\lambda^2\right) p_{l,m,n}^{t} - p_{l,m,n}^{t-1} + \lambda^2 \sum_{f \in \text{faces}} p_f^t \,, \tag{4.16}$$

where $\lambda = \frac{cT}{X}$, also known as the Courant number. A pressure update is required for every node in the modelled volume at every time step. The spatial complexity is therefore clearly $\Theta(N_g)$ and the time complexity $\Theta(N_g\, n)$, where $N_g$ is the number of nodes in the space and $n$ is the number of time steps that are modelled.

Kowalczyk and Van Walstijn [107] derived a more general 3D case in which a $3 \times 3 \times 3$ cubic stencil is used, and the update equation for the pressure $p_{x,y,z}^t$ reads as:

$$p_{x,y,z}^{t+1} = d_1 \left( \sum_{f \in \text{faces}} p_f^t \right) + d_2 \left( \sum_{e \in \text{edges}} p_e^t \right) + d_3 \left( \sum_{c \in \text{corners}} p_c^t \right) + d_4\, p_{x,y,z}^t - p_{x,y,z}^{t-1} \,,$$

where:

$$\begin{aligned} d_1 &= \lambda^2 (1 - 4a + 4b) \\ d_2 &= \lambda^2 (a - 2b) \\ d_3 &= \lambda^2\, b \\ d_4 &= 2(1 - (6a - 4b - 3)\lambda^2) \,, \end{aligned} \tag{4.17}$$

and where $p_f$ are the 6 points at the faces of the cube, $p_e$ the 12 edges of the cube, and $p_c$ the 8 corners. The variables $a$ and $b$ are free parameters that define the nature of the model [107]. They are bound to meet the following stability conditions:

$$\begin{aligned} a &\leq \frac{1}{2} \\ b &\geq \frac{1}{16}(12a - 3) \\ \lambda &\leq \min\left(1, \frac{1}{2 - 4a}, \frac{1}{3 - 12a + 16b}\right) . \end{aligned} \tag{4.18}$$

For example, $a$ can be set to $\frac{1}{4}$ and $b$ to 0 to form an octahedral grid. Of particular interest is the Standard Rectilinear (SRL) stencil ($a = 0$ and $b = 0$), in which only the 6 neighbouring nodes are considered and the update equation reduces to Eq. 4.16. Different choices for $a$ and $b$ allow for a wide range of different stencils with different properties in terms of isotropy and dispersion. Nearly all stencils are more isotropic and suffer from less dispersion than the SRL. However, this is often at the cost of an increasing stability limit for $\lambda$, which results in a higher spatial sampling for the same sample rate. It is therefore a trade-off between dispersion, computational efficiency, and memory overhead. Kowalczyk and Van Walstijn [107] argue for the Interpolated Wideband (IWB) scheme to be the most favourable in terms of isotropy and dispersion. While this is correct, this is also the stencil with the highest stability limit ($\lambda \leq 1$), which makes it less attractive in terms of memory and thus computational overhead.

Impedance boundaries and frequency dependent absorbing boundaries have been investigated by, most notably, Kowalczyk and Van Walstijn [105, 106]. They discretise the frequency-independent boundary equation Eq. 3.9 with a centred-difference scheme to obtain:

$$p_x^{t+1} - p_x^{t-1} = -\zeta \cos\theta \left(p_{x+1}^t - p_{x-1}^t\right). \tag{4.19}$$

As we only consider the FDTD method on a rectangular grid, $\cos\theta = 1$. At a boundary, the value for $p_{x+1}^t$ as in Eq. 4.17 is not defined, as this value is beyond the modelled domain. However, we can solve Eq. 4.19 for $p_{x+1}^t$, and substitute this into Eq. 4.17. This gives us an update equation that includes the impedance boundary condition, but does not include any nodes outside of the modelled domain. In a similar fashion, Kowalczyk and Van Walstijn defined boundaries for the other stencils with general $a$ and $b$.

It has been shown that these boundaries indeed absorb energy and using Von Neumann analysis they were shown to be stable. However, using spectral and pseudo-spectral analysis, Botts and Savioja [33] showed that there are still cases in which a system with the above boundaries can be unstable, particularly when employing floating point arithmetic. A different type of boundary was introduced by Webb and Bilbao [218], derived in the same way but using a forward-difference approximation for the time derivative. Botts and Savioja call this the *velocity-centred* boundary, in contrast to the former *pressure-centered* boundary, relating to the equivalent absorptive boundary approximations in a staggered velocity and pressure update method [33, 32]. Spectral analysis seems to imply that the velocity-centred boundary is more stable than the pressure-centered one.

Higher-order stencils have been discussed by e.g. Sakamoto [162] and will be discussed further in Chapter 5. Following this study, Hamilton *et al.* [82] performed an in-depth study of fourth-order stencils of various configurations. Their focus lay on the GPU implementation and memory throughput, and they conclude that, expectedly, GPU performance in terms of compute times per node roughly scales linearly with stencil size. They mention that impedance boundary implementations are a problem yet to be solved. More on this topic will be addressed in Chapter 6.

### 4.4.3 Boundary element method

The use of the boundary element method (BEM) in acoustics has been elaborately described by Kirkup [101]. Besides its use in acoustics, the BEM also finds many applications in, for example, thermo- and fluid dynamics [222, 3]. The idea behind the BEM is that the wave equation is only solved at the boundaries of a mesh, after which the effect thereof can be computed for any point within the mesh. This is done by defining a Green's function $G(\mathbf{p}, \mathbf{q})$, which describes the effect of a boundary point $\mathbf{q}$ onto another point $\mathbf{p}$. If the system equation at $\mathbf{p}$ can be described as a Laplace equation:

$$\sum_{i=1}^{N} \frac{\partial^2 \phi(\mathbf{p})}{\partial p_i^2} = 0\,, \tag{4.20}$$

then it can also be described as:

$$\int_S G(\mathbf{p}, \mathbf{q}) \frac{\partial}{\partial n} \phi(\mathbf{q})\, d\mathbf{q}\,, \tag{4.21}$$

where $S$ represents the surrounding boundary over which is integrated, and $\frac{\partial}{\partial n}\phi(\mathbf{q})$ represents the derivative of $\phi$ at $\mathbf{q}$ with respect to the surface normal. Kirkup [101] described a detailed algorithm to subdivide a boundary into discrete sections, find Green's function and compute its values.

An advantage of the BEM is that it also easily works in outside regions, whereas a 3D FDTD method might require much more memory space and time to solve the same equation [101]. Absorbing boundaries can for example be modelled by a Robin boundary condition, as described by, amongst others, Shaw [178], Ha-Duong *et al.* [77], and Hargreaves [84]. However, according to Thompson [201, p. 1315], the BEM is "limited in the ability to model complex, inhomogeneous regions." There is no theoretical limit on the complexity model, but the computation time increases rapidly with the model growing more complex. Harari and Hughes [83] thoroughly analysed the complexity of the BEM. It is not a trivial function, however, and it is dependent on the number of dimensions, boundaries and boundary conditions, as well as on the discretisation of the mesh. In summary, the BEM can be very fruitful for simple scenes and exterior problems, but not for complex environments with different materials.

### 4.4.4 Finite element method

Rather than describing one well-defined method, finite element analysis encompasses a range of different approaches for solving PDEs. Just as the BEM, finite element methods (FEMs) are not limited to acoustic modelling, but include applications in fluid dynamics [225], electromagnetics [93], and mechanics [224]. Thompson [201] wrote a detailed overview of the state of the art of several FEMs in acoustics. The FEM does not give a very intuitive approach to acoustic modelling. Nonetheless, a great number of different physical phenomena can be modelled, and a large amount of research has gone into optimising the methods for different purposes.

In a FEM simulation, a simulation domain needs to be defined, which then needs to be discretised into a finite number of elements. These elements may, but need not be, uniformly distributed in space. In some cases, it may be useful to have a high resolution of elements at places where a higher resolution is required. The essential goal of the FEM method is to express a function $p(\mathbf{x})$ (pressure in acoustic modelling) at any point $\mathbf{x}$ in space in terms of the discrete set of points for which values can easily be calculated. An elegant piece of linear algebra enables the FEM to express the physical system in terms of matrix multiplications, such that the speed of the algorithm is generally manageable. As with the BEM, the precise complexity of FEMs depends on several factors, and Harari and Hughes [83] made a detailed analysis of the spatial and time complexity.

FEMs are able to model a great number of physical phenomena, such as absorption, transmission, and diffraction. However, their resolution generally is not high enough to generate a reliable acoustic impulse response of sufficient resolution. It is therefore an appropriate method for modal analysis [201], but less appropriate for room acoustic modelling. Moreover, it suffers from dispersion error, the precise quantity depending on the design of the mesh.

### 4.4.5 Adaptive rectangular decomposition

A relatively novel method was proposed by Raghuvanshi *et al.* [156], which is an extension of the FDTD algorithm and was named “adaptive rectangular decomposition” (ARD). Making use of the fact that the eigenfunctions of the wave equation in a rectangular environment are known and exact (see Eq. 2.28), they propose to divide the space enclosed by the geometry into the smallest number of axis-aligned rectangles. (The exact algorithm for this subdivision is beyond the interest of this work.) As the eigenfunctions of a rectangular section are all the product of three cosines, a point within this section can be represented as:

$$\begin{aligned} p(x,y,z,t) = \sum_{n=\{n_x,n_y,n_z\}}^{\infty} & m_n(t) \cos\left(\frac{n_x\,\pi}{L_x}x\right) \cos\left(\frac{n_y\,\pi}{L_y}y\right) \cos\left(\frac{n_z\,\pi}{L_z}z\right) = \\ & = \mathrm{DCT}_{3\mathrm{D}}\left\{m_n(t)\right\} = M_n(\omega)\,, \end{aligned} \tag{4.22}$$

where $m(t)$ is some time-dependent modal function. This can be recognised as the (3-dimensional) Discrete Cosine Transform (DCT) of $m(t)$, which we call $M(\omega)$. This gives us the relation $p \overset{\mathbf{DCT}}{\Longleftrightarrow} M$. They further note that by transforming also the wave equation with driving function (Eq. 2.35), an expression of the wave equation in the discrete cosine domain can be formulated:

$$\frac{\partial^2 M_n(\omega)}{\partial t^2} + \left(\omega_{n_x}^2 + \omega_{n_y}^2 + \omega_{n_z}^2\right) M_n(\omega) = \mathrm{DCT}_{3\mathrm{D}}(F(t)). \tag{4.23}$$

An update equation can now be found for $M$ rather than the usual $p$, which Raghuvanshi *et al.* show to be:

$$M_n^{t+1} = 2M_n^t \cos\left(\omega_n T\right) - M_n^{t-1} + \frac{2\,\mathrm{DCT}_{3\mathrm{D}}(F(t))}{\omega_n^2}\left(1 - \cos\left(\omega_n T\right)\right). \tag{4.24}$$

Seemingly this is not an advantage, as each time update for a volume of $N$ nodes now requires a DCT, which typically runs at $O(N\log_2(N))$. However, the fast Fourier Transform (FFT) algorithm, and similarly the DCT, is an extremely common and well-understood algorithm, and as such knows many efficient implementations. Raghuvanshi *et al.* use a GPU implementation described in [74], which in practice can conceivably be faster than most traditional implementations of an $O(N)$ FDTD implementation. In addition, the advantage is that this method does not produce any dispersion error, contrary to known and heavy dispersion effects typically produced by finite difference modelling.

While the above provides an efficient means of computing the wave propagation in large rectangular spaces, the interfaces between rectangles as well as the boundaries of the domain still need to be handled. However, as this ARD lives “on top” of the FDTD solution (i.e. the space is already sampled in a FDTD-like fashion), Raghuvanshi *et al.* deal with these interfaces using a standard FDTD algorithm. In particular, they claim to use a 6th-order FDTD stencil, although they quote an unstable Courant number at which this is evaluated, and they use perfectly-matched layers at the boundaries [158]. The full process of voxelising, decomposition, excitation and interface interaction is graphically represented in Fig. 4.12.

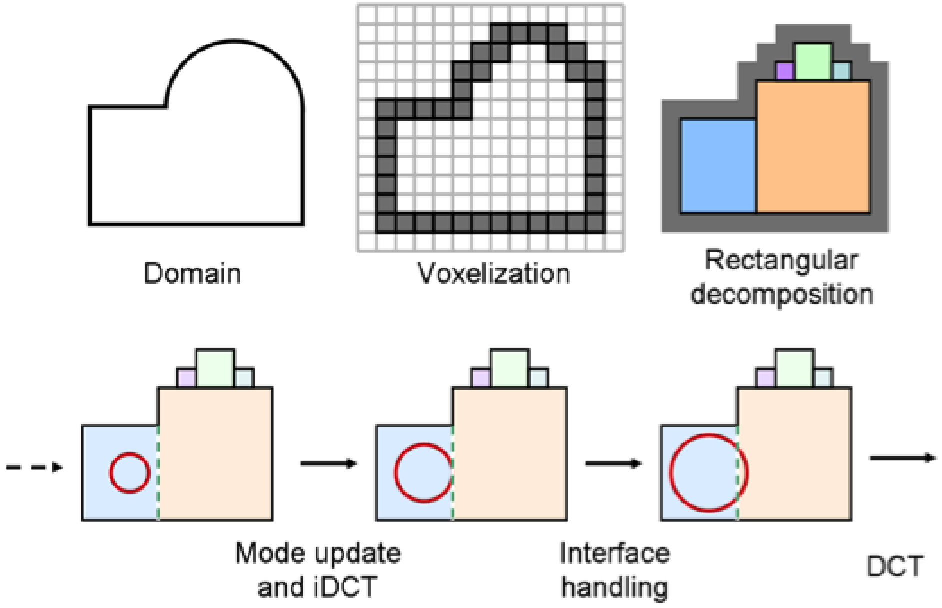


**Figure 4.12:** The preprocessing and simulation process of the Adaptive Rectangular Decomposition method. Image from [156] with permission.

The successfulness of this algorithm depends on the amount of space that can be decomposed into rectangles. Because of the small amount of FDTD solutions where this decomposition is not possible (e.g. curved boundaries), the algorithm is not entirely dispersionless, though in most cases it will outperform a traditional FDTD implementation. The performance benefits are difficult to describe. At identical sample rate, the ARD method runs with a time complexity $\Theta(N \log(N))$ versus $\Theta(N)$ for the FDTD, where $N$ is the number of nodes in the space, and both run at the same spatial complexity of $\Theta(N)$. Raghuvanshi *et al.* compare their ARD method with a 2.5x oversampled FDTD method, though, arguing that this is the fairer comparison because of the poorer quality of the FDTD algorithm in terms of dispersion. As a result, their claimed speed-up is of the order of a factor 100x and memory savings of order 10x. While the ARD method demonstrates no dispersion error, it does suffer from interface errors at the boundaries of the rectangular sections, and, as in the classic FDTD method, staircasing error remains present.

### 4.4.6 Finite volume method

Of the aforementioned wave-based methods, the DWG, FDTD, and ARD methods can produce ‘staircase errors’ at the boundaries due to regular grid discretisation. Boundaries that are not axis-aligned, including circular and elliptic boundaries, cannot be discretised very well and will need to be approximated by nodes that produce a staircase-like pattern, as depicted in Fig. 4.13. There are numerous studies into the quality and quantity of artifacts produced by this staircasing error [36, 19, 80, 21]. For low-frequency propagation the error is small, but the higher the frequency, the more error is introduced into the model, resulting in modal shifts and erroneous reflective and diffusive properties.

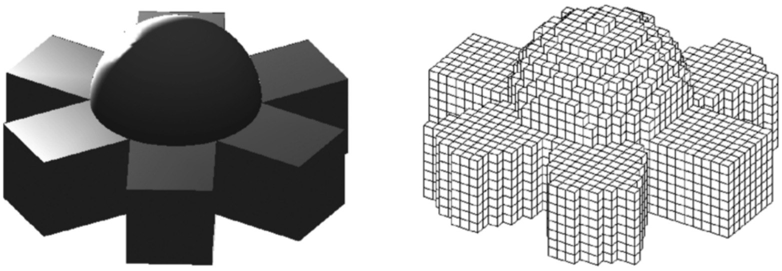


**Figure 4.13:** A cubic discretisation of a 3D shape. It is evident that in the dome and the non-axis-aligned surfaces, staircase discretisation errors are introduced. Image from [19] with permission.

The *finite volume time domain* (FVTD) method, often simply called 'finite volume' method, aims to address this error by allowing more general polygons to fill up the gaps at the boundaries. It is a common method in fluid and gas mechanics, and has a little more recently made its way into acoustic modelling. Botteldooren [31] already described non-Cartesian grids that allow this to some extent, as a generalisation of the FDTD method. Mattsson *et al.* [130] used the FVTD method to successfully model wave propagation, including rigid Von Neumann and Dirichlet boundary conditions, using an irregularly sized tetrahedral grid.

A comprehensive study of the FVTD in acoustics was published by Bilbao [19], and was elaborated upon by Bilbao *et al.* [21] to also include stable impedance boundaries and a thorough analysis on the error produced by different discretisation methods. They used the *energy method* (see Sec. 6.1.1) to prove stability. By using the same box under different rotations, they compared the effect of the staircasing error on the $\mathrm{RT}_{60}$ time. They convincingly showed that using the FDTD algorithm, the $\mathrm{RT}_{60}$ time is much less stable than when using the FVTD method. However, it must be noted that this error is frequency-dependent, and can be reduced significantly by choosing a smaller grid spacing.

## 4.5 Computational acoustic modelling and benchmarking

In the context of this thesis, the purpose of computational acoustic modelling is to computer-generate impulse responses in the most realistic way as possible. That is to say, if an impulse response was taken according the above described standards, it should be identical to one that was computed using some computer algorithm that takes the geometry of the scene and its materials as input. In practice it is impossible to obtain identical measured and virtual RIRs, but the differences may not be perceptually relevant, or if they are, the character of the RIR may still be preserved. Rather than comparing the RIRs themselves, it is therefore more useful to combine the room acoustic parameters derived from the respective impulse responses (a.o. those listed in Sec. 3.2.3).

Some efforts to standardise, benchmark, and compare measured and virtually generated RIRs were made through the *round robins* on room acoustical simulation [213, 27, 28, 29]. In these round robins, computational acoustic modelling tools from various participants, from both academic and industry backgrounds, were benchmarked against a measured RIR of a real space of relatively simple geometry. While not exactly being a competition, the round robins provide a good insight into the performance of different acoustic modelling tools, while also providing a means to standardise acoustic modelling techniques, input and output parameters.

It was observed by Hornikx *et al.* [89] that, despite the aforementioned round robins, there is a lack of common processes, techniques, benchmark test cases, and reporting methods to properly validate and compare acoustic modelling methods. They propose an elaborate and long-term framework the "...start of a long term project, about deploying benchmarks in the entire field of computational acoustics." Their benchmark platform considers four different types of acoustic problems: linear acoustics, high-frequency acoustics, acoustics in vibration, and heterogeneous and moving fluid acoustics. They propose a number of benchmark test cases, along with detailed specifications of the problem, and the requirements to which result submissions should adhere. This being a project intended to last a long time and be expanded upon, a technical committee of members of the European Acoustics Association has been installed to oversee this endeavour.

## 4.6 Summary

The most common computational acoustic modelling techniques have been described. Among the geometric methods, the most frequently used in commercial software are the image source method and ray tracing (including its variations such as cone/pyramid tracing). A major advantage is that their performance is not influenced by the physical size of the modelled space, and additionally, ray tracing-like algorithms are fast and well-studied. They do not perform well below the Schröder frequency, however, as they are inherently bad at modelling wave-like behaviour. At low frequencies, wave-based methods are more appropriate. Most of the above methods have been used in the context of acoustics, but to the best of the author's knowledge, the BEM and FEM are not generally used for room acoustic modelling problems due to their more prohibitive algorithmic time complexity. Both the DWG and FDTD method, and more recently also the FVTD and ARD method, have been used extensively to tackle room acoustic problems below the Schröder frequency. It is recognised that a set of benchmark cases would be beneficial in order to facilitate objective comparisons. The next chapter focusses on the FDTD method, as it will be the main subject of the rest of the thesis.

# Chapter 5

# Higher order FDTD stencils

**Contents**

The stencils used in the Sec. 4.4.2 all relied on dropping the higher order terms in the approximation of the second derivative (see Eq. 4.14). It is commonly assumed that a higher-order approximation gives a more accurate approximation, reduces the amount of dispersion error, and increases the valid bandwidth [42, 162, 75]. In this chapter, a family of higher-order accurate schemes called ‘large-star’ [42] or ‘leggy’ [82] stencils are studied: schemes that only take nodes along the grid axes into their update equation. They offer improvements over the SRL but without the complex stencils required for other approaches, including IWB. The goal, therefore, is to identify an optimal scheme, better than SRL, that should be at least comparable to what is offered by IWB. Both theoretical dispersion patterns and practical implementation on a GPU device are discussed.

The structure of this chapter is as follows: Sec. 5.1 is a theoretical background section that deals with a common derivation of higher-order approximations and defines the stencils used later in the chapter. Sec. 5.2 discusses the theoretical stability conditions of these stencils, and Sec. 5.3 presents a detailed dispersion analysis of the 3-dimensional stencils. In Sec. 5.4, the practical GPU implementation is discussed, and in Sec. 5.5 the experimental results are presented and discussed. This also includes more recent work by Hamilton *et al.* [82]. Sec. 5.6 presents a detailed modal analysis to verify the theoretical dispersion patterns obtained in Sec. 5.3.

## 5.1 Deriving higher order approximations

Consider first of all the derivation of the common $2^{\text{nd}}$-order accurate stencil, using the Lagrange polynomial method (following e.g. [51]). A different but equivalent method using Taylor series is presented by e.g. Sakamoto [162]. Let us define three points equally spaced over the $x$-axis: $(0,\, f_0)$, $(X,\, f_1)$, and $(2X,\, f_3)$. A Lagrange polynomial $g(x)$ is the polynomial of smallest degree that passes through each of these points, and can be written as:

$$g(x) = a_2\, x^2 + a_1 x + a_0. \tag{5.1}$$

In order to ensure that this polynomial indeed passes through all three points, the following must hold:

$$\begin{pmatrix} 0 & 0 & 1 \\ X^2 & X & 1 \\ 4X^2 & 2X & 1 \end{pmatrix} \begin{pmatrix} a_2 \\ a_1 \\ a_0 \end{pmatrix} = \begin{pmatrix} f_0 \\ f_1 \\ f_2 \end{pmatrix}. \tag{5.2}$$

This provides us with three equations and three unknowns (i.e. the polynomial coefficients), so it can be solved for $a_n$:

$$\begin{cases} a_0 = f_0 \\ a_1 = \dfrac{-3f_0 - 4f_1 + f_2}{2X} \\ a_2 = \dfrac{f_0 - 2f_1 + f_2}{2X^2} \end{cases}, \tag{5.3}$$

such that the Lagrange polynomial becomes:

$$g(x) = \frac{f_0 - 2f_1 + f_2}{2X^2} x^2 + \frac{-3f_0 - 4f_1 + f_2}{2X} x + f_0. \tag{5.4}$$

The approximation for the second derivative of $g(x)$ then becomes:

$$\frac{d^2g(x)}{dx^2} = \frac{f_0 - 2f_1 + f_2}{X^2}, \tag{5.5}$$

for all values of $x$, which is exactly equivalent to the approximation in Eq. 4.14, without the higher order terms.

The same procedure can be applied in order to derive higher order approximations. For a $4^{\text{th}}$-order stencil we need five points instead of three. The Lagrange polynomial $g(x)$ now becomes:

$$g(x) = a_4\,x^4 + a_3\,x^3 + a_2\,x^2 + a_1\,x + a_0, \tag{5.6}$$

such that the matrix equation becomes:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ X^4 & X^3 & X^2 & X & 1 \\ 16X^4 & 8X^3 & 4X^2 & 2X & 1 \\ 81X^4 & 27X^3 & 9X^2 & 3X & 1 \\ 256X^4 & 64X^3 & 16X^2 & 4X & 1 \end{pmatrix} \begin{pmatrix} a_4 \\ a_3 \\ a_2 \\ a_1 \\ a_0 \end{pmatrix} = \begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix}. \tag{5.7}$$

This is a $5 \times 5$ matrix with 5 unknowns, so we can still solve for the polynomial coefficients, to get:

$$\begin{cases} a_0 = f_0 \\ a_1 = \dfrac{-25f_0 + 48f_1 - 36f_2 + 16f_3 - 3f_4}{12h} \\ a_2 = \dfrac{35f_0 - 104f_1 + 114f_2 - 56f_3 + 11f_4}{24h^2} \\ a_3 = \dfrac{-5f_0 + 18f_1 - 24f_2 + 14f_3 - 3f_4}{12h^3} \\ a_4 = \dfrac{f_0 - 4f_1 + 6f_2 - 4f_3 + f_4}{24h^4} \end{cases}. \tag{5.8}$$

In order to obtain a centred difference approximation, we have to evaluate $g''(x)$ at the middle point, $2X$:

$$g''(2X) = \frac{-f_0 + 16f_1 - 30f_2 + 16f_3 - f_4}{12h^2}. \tag{5.9}$$

This immediately gives the fourth order accurate update stencil: $\{-\frac{1}{12}, \frac{4}{3}, -\frac{5}{2}, \frac{4}{3}, -\frac{1}{12}\}$.

The same method can be applied to arrive at the stencils up to an arbitrary order of accuracy. In every case, this gives rise to an $n \times n$ matrix with $n$ polynomial coefficients, which always results in a determinate system. More specifically, it will be of the following form:

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 1 \\ X^n & X^{n-1} & \cdots & X^2 & X & 1 \\ (2X)^n & (2X)^{n-1} & \cdots & 4X^2 & 2X & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ ((n-1)X)^{n-1} & ((n-1)X)^{n-1} & \cdots & ((n-1)X)^2 & (n-1)X & 1 \\ (nX)^n & (nX)^{n-1} & \cdots & (nX)^2 & nX & 1 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ \vdots \\ a_1 \\ a_0 \end{pmatrix} = \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix}. \tag{5.10}$$

As we aim to find centred difference approximations, the stencils are always symmetric and have an odd number of entries. For this reason also, only even orders of accuracy are allowed. In general, an $n^{\text{th}}$ order system will have a stencil of size $n+1$. In Table 5.1 the update coefficients for all stencils up to $20^{\text{th}}$ order accuracy are displayed.

| | $p_{x-10}$ | $p_{x-9}$ | $p_{x-8}$ | $p_{x-7}$ | $p_{x-6}$ | $p_{x-5}$ | $p_{x-4}$ | $p_{x-3}$ | $p_{x-2}$ | $p_{x-1}$ | $p_x$ | $p_{x+1}$ | $p_{x+2}$ | $p_{x+3}$ | $p_{x+4}$ | $p_{x+5}$ | $p_{x+6}$ | $p_{x+7}$ | $p_{x+8}$ | $p_{x+9}$ | $p_{x+10}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2nd | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $\frac{1}{2}$ | $-1$ | $\frac{1}{2}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4th | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $-\frac{1}{12}$ | $\frac{4}{3}$ | $-\frac{5}{2}$ | $\frac{4}{3}$ | $-\frac{1}{12}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6th | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $\frac{1}{90}$ | $-\frac{3}{20}$ | $\frac{3}{2}$ | $-\frac{49}{18}$ | $\frac{3}{2}$ | $-\frac{3}{20}$ | $\frac{1}{90}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8th | 0 | 0 | 0 | 0 | 0 | 0 | $-\frac{1}{560}$ | $\frac{8}{315}$ | $-\frac{1}{5}$ | $\frac{8}{5}$ | $-\frac{205}{72}$ | $\frac{8}{5}$ | $-\frac{1}{5}$ | $\frac{8}{315}$ | $-\frac{1}{560}$ | 0 | 0 | 0 | 0 | 0 | 0 |
| 10th | 0 | 0 | 0 | 0 | 0 | $\frac{1}{3150}$ | $-\frac{5}{1008}$ | $\frac{5}{126}$ | $-\frac{5}{21}$ | $\frac{5}{3}$ | $-\frac{5269}{1800}$ | $\frac{5}{3}$ | $-\frac{5}{21}$ | $\frac{5}{126}$ | $-\frac{5}{1008}$ | $\frac{1}{3150}$ | 0 | 0 | 0 | 0 | 0 |
| 12th | 0 | 0 | 0 | 0 | $-\frac{1}{16632}$ | $\frac{2}{1925}$ | $-\frac{1}{112}$ | $\frac{10}{189}$ | $-\frac{15}{56}$ | $\frac{12}{7}$ | $-\frac{5369}{1800}$ | $\frac{12}{7}$ | $-\frac{15}{56}$ | $\frac{10}{189}$ | $-\frac{1}{112}$ | $\frac{2}{1925}$ | $-\frac{1}{16632}$ | 0 | 0 | 0 | 0 |
| 14th | 0 | 0 | 0 | $\frac{1}{84084}$ | $-\frac{7}{30888}$ | $\frac{7}{3300}$ | $-\frac{7}{528}$ | $\frac{7}{108}$ | $-\frac{7}{24}$ | $\frac{7}{4}$ | $-\frac{266681}{88200}$ | $\frac{7}{4}$ | $-\frac{7}{24}$ | $\frac{7}{108}$ | $-\frac{7}{528}$ | $\frac{7}{3300}$ | $-\frac{7}{30888}$ | $\frac{1}{84084}$ | 0 | 0 | 0 |
| 16th | 0 | 0 | $-\frac{1}{411840}$ | $\frac{16}{315315}$ | $-\frac{2}{3861}$ | $\frac{112}{32175}$ | $-\frac{7}{396}$ | $\frac{112}{1485}$ | $-\frac{14}{45}$ | $\frac{16}{9}$ | $-\frac{1077749}{352800}$ | $\frac{16}{9}$ | $-\frac{14}{45}$ | $\frac{112}{1485}$ | $-\frac{7}{396}$ | $\frac{112}{32175}$ | $-\frac{2}{3861}$ | $\frac{16}{315315}$ | $-\frac{1}{411840}$ | 0 | 0 |
| 18th | 0 | $\frac{1}{1969110}$ | $-\frac{9}{777920}$ | $\frac{9}{70070}$ | $-\frac{2}{2145}$ | $\frac{18}{3575}$ | $-\frac{63}{2860}$ | $\frac{14}{165}$ | $-\frac{18}{55}$ | $\frac{9}{5}$ | $-\frac{9778141}{3175200}$ | $\frac{9}{5}$ | $-\frac{18}{55}$ | $\frac{14}{165}$ | $-\frac{63}{2860}$ | $\frac{18}{3575}$ | $-\frac{2}{2145}$ | $\frac{9}{70070}$ | $-\frac{9}{777920}$ | $\frac{1}{1969110}$ | 0 |
| 20th | $-\frac{1}{9237800}$ | $\frac{10}{3741309}$ | $-\frac{5}{155584}$ | $\frac{30}{119119}$ | $-\frac{5}{3432}$ | $\frac{24}{3575}$ | $-\frac{15}{572}$ | $\frac{40}{429}$ | $-\frac{15}{44}$ | $\frac{20}{11}$ | $-\frac{1968329}{635040}$ | $\frac{20}{11}$ | $-\frac{15}{44}$ | $\frac{40}{429}$ | $-\frac{15}{572}$ | $\frac{24}{3575}$ | $-\frac{5}{3432}$ | $\frac{30}{119119}$ | $-\frac{5}{155584}$ | $\frac{10}{3741309}$ | $-\frac{1}{9237800}$ |

**Table 5.1:** This table shows the coefficients for a centred difference approximations up to 20th order.

| Order | 2nd | 4th | 6th | 8th | 10th | 12th | 14th | 16th | 18th | 20th | 22th | 24th |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda_{\max}$ | 1 | $\sqrt{\frac{3}{4}}$ | $\sqrt{\frac{45}{68}}$ | $\sqrt{\frac{315}{512}}$ | $\sqrt{\frac{75}{128}}$ | $\sqrt{\frac{51975}{91904}}$ | $\sqrt{\frac{4729725}{8593664}}$ | $\sqrt{\frac{4729725}{8781824}}$ | $\sqrt{\frac{241215975}{455901184}}$ | $\sqrt{\frac{4583103525}{8792178688}}$ | $\sqrt{\frac{50414138775}{97952595968}}$ | $\sqrt{\frac{1159525191825}{2277888753664}}$ |
| Memory saving | 0% | 35.1% | 46.2% | 51.7% | 55.2% | 57.5% | 59.2% | 60.5% | 61.5% | 62.4% | 63.1% | 63.7% |

**Table 5.2:** This table shows the stability limit for the Courant number for the above-mentioned centred difference approximations in one dimension, up to 24th order. It also lists the RAM savings (in percentages) in 3-dimensional grids due to the change in grid spacing for higher order stencils.

## 5.2 Stability analysis

As the new stencils have more terms, the stability condition for each of them has to be revised. First of all, we have to decide which differential operators (i.e. temporal or spatial) we are going to approximate with higher order terms. From a computational point of view, adding higher order terms to the second derivative of time means that we have to keep track of more previous time steps. This has a significant consequence for the memory usage when implemented: for each additional time step, pressure values have to be stored for the entire domain. This is probably affordable in small simulations, but may prove to be a bottleneck in large ones. On the other hand, adding higher order terms to the spatial derivative implies that for each update, not only the direct neighbour nodes but also ones further away are used to compute the new pressure value. This will most likely involve extra computational effort, but the memory usage will remain the same.

We choose to keep approximating the time derivative with a second order approximation, but to increase the order of the spatial derivative. Traditional Von Neumann stability analysis can be applied. The general update equation of an $N^{\text{th}}$ stencil can be written as:

$$p_x^{t+1} - 2p_x^t + p_x^{t-1} = \lambda^2 \sum_{n=-\frac{N}{2}}^{\frac{N}{2}} a_n \, p_{x+n}^t. \tag{5.11}$$

We define $\mathcal{F}\left\{p_x^t\right\}$ as:

$$p_x^t \xrightarrow{\mathcal{F}} e^{\omega t T} \, e^{-jkxX} = z^t \, e^{-jkxX}, \tag{5.12}$$

where $z = e^{\omega T}$. Substitution in Eq. 5.11 and simplifying the result gives:

$$\begin{aligned} z - 2 + z^{-1} &= \lambda^2 \sum_{n=-\frac{N}{2}}^{\frac{N}{2}} a_n \, e^{-jknX} \\ -4\sin^2\left(\frac{\omega T}{2}\right) &= \lambda^2 \sum_{n=-\frac{N}{2}}^{\frac{N}{2}} a_n \, e^{-jknX} \end{aligned} \tag{5.13}$$

For stability, we require that $z \leq 1$ and it follows that the right-hand side of the equation needs to satisfy:

$$0 \leq \left| -\lambda^2 \sum_{n=-\frac{N}{2}}^{\frac{N}{2}} a_n \, e^{-jknX} \right| \leq 4 \qquad \forall \, k, X. \tag{5.14}$$

For the fourth order accurate update scheme, we then get:

$$\begin{aligned}
0 &\leq \left| -\lambda^2 \left( -\frac{1}{12} e^{-2jkX} + \frac{4}{3} e^{-jkX} - \frac{5}{2} + \frac{4}{3} e^{jkX} - \frac{1}{12} e^{2jkX} \right) \right| \leq 4 \\
0 &\leq \left| -\lambda^2 \left( -\frac{1}{12} \left( e^{-2jkX} - 2 + e^{2jkX} \right) + \frac{4}{3} \left( e^{-jkX} - 2 + e^{jkX} \right) \right) \right| \leq 4 \\
0 &\leq \lambda^2 \underbrace{\left| \frac{4}{3} \sin^2 \left( \frac{kX}{2} \right) - \frac{1}{12} \sin^2 (kX) \right|}_{\text{maximum value: } \frac{4}{3}} \leq 1.
\end{aligned} \tag{5.15}$$

Therefore, the stability condition for the Courant number in a fourth order accurate centred difference scheme in 1D is: $0 \leq \lambda \leq \frac{1}{2}\sqrt{3}$. In much the same way, the stability limit for a sixth order accurate scheme is shown as follows:

$$0 \leq \quad \lambda^2 \underbrace{\left| \frac{3}{2} \sin^2 \left( \frac{kX}{2} \right) - \frac{3}{20} \sin^2 (kX) + \frac{1}{90} \sin^2 \left( \frac{3kX}{2} \right) \right|}_{\text{maximum value: } \frac{3}{2} + \frac{1}{90}} \leq 1, \tag{5.16}$$

such that the stability condition is $0 \leq \lambda \leq \sqrt{\frac{45}{68}}$. In Table 5.2, the stability limits for all stencils up to 20th order are given. Sakamoto [162] quotes similar but different stability limits for all except the 2nd-order case, though his limits are more conservative than necessary. For brevity's sake, the given stability limits only apply to the 1-dimensional case. It is trivially shown that for $D$-dimensional leggy stencils, the stability limit becomes: $\lambda_{2^{\text{nd}}} \sqrt{\frac{1}{D}}$, for example:

$$\lambda_{4^{\text{th}} \text{ order}} = \sqrt{\frac{3}{4\,D}} \quad \lambda_{6^{\text{th}} \text{ order}} = \sqrt{\frac{45}{67\,D}} \quad \lambda_{8^{\text{th}} \text{ order}} = \sqrt{\frac{315}{512\,D}} \quad \ldots \tag{5.17}$$

It it clear that the stability limit for $\lambda$ is decreasing with every increase in order of accuracy. The physical interpretation of this is that, when the sample rate and speed of sound remain constant, the grid spacing is increased. The difference in stability limit for each subsequent order becomes less and less. Fig. 5.1 graphically shows the stability limit up to 24th order. If we assume that the speed of sound and the sample rate of the simulations remain the same, a change in the Courant number implies a change in the grid spacing. For every increase in order of accuracy, the grid spacing can go up. In terms of simulating the space, this means that a smaller number of grid points is required to simulate the same space. As this paper is mostly concerned with 3D simulations, we will consider the memory savings of an $n^{\text{th}}$-order

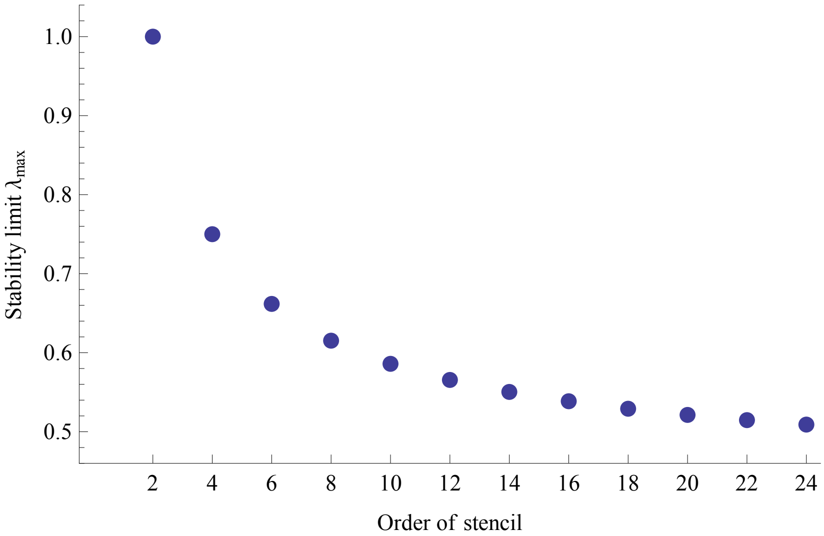


**Figure 5.1:** The value of the Courant stability limit for the 2nd up to the 24th order 1-dimensional stencils.

scheme in 3 dimensions with respect to the 3D SRL scheme:

$$\text{Relative memory usage} = \left(\frac{\lambda_{n,\text{3D}}}{\lambda_{2^{\text{nd}},\text{3D}}}\right)^3 = \left(\frac{\lambda_{n,\text{3D}}}{\sqrt{\frac{1}{3}}}\right)^3 = \lambda_{n,\text{1D}}^3. \tag{5.18}$$

In Table 5.2, the memory savings (in percentages) are listed for all quoted schemes.

## 5.3 Dispersion analysis

Dispersion error is the error in the wave speed in the computational model compared to the wave speed in the exact solution. Dispersion error is dependent on frequency, and, in higher dimensions, propagation direction. It is generally introduced as a side-effect of discretisation. The wave speed in an FDTD model can be expressed by solving Eq. 5.13 for $\omega = \omega_\phi$ to get the angular phase velocity:

$$\omega_\phi = \frac{2}{T} \arcsin\left(\lambda \sqrt{-\frac{1}{4} \sum_{n=\frac{N}{2}}^{\frac{N}{2}} a_n \, e^{-jknX}}\right). \tag{5.19}$$

The relative phase velocity $v_\phi$ is then defined as $v_\phi = \frac{\omega_\phi}{c\,k}$. In 3 dimensions, we get a similar expression:

$$\omega_\phi = \frac{2}{T}\arcsin\left(\lambda\sqrt{-\frac{1}{4}\sum_{n=\frac{N}{2}}^{\frac{N}{2}} a_n\left(e^{-jk_xnX}+e^{-jk_ynY}+e^{-jk_znZ}\right)}\right), \tag{5.20}$$

and for the relative phase velocity:

$$v_\phi = \frac{2}{c\,T\sqrt{k_x^2+k_y^2+k_z^2}}\arcsin\left(\lambda\sqrt{-\frac{1}{4}\sum_{n=\frac{N}{2}}^{\frac{N}{2}} a_n\left(e^{-jk_xnX}+e^{-jk_ynY}+e^{-jk_znZ}\right)}\right). \tag{5.21}$$

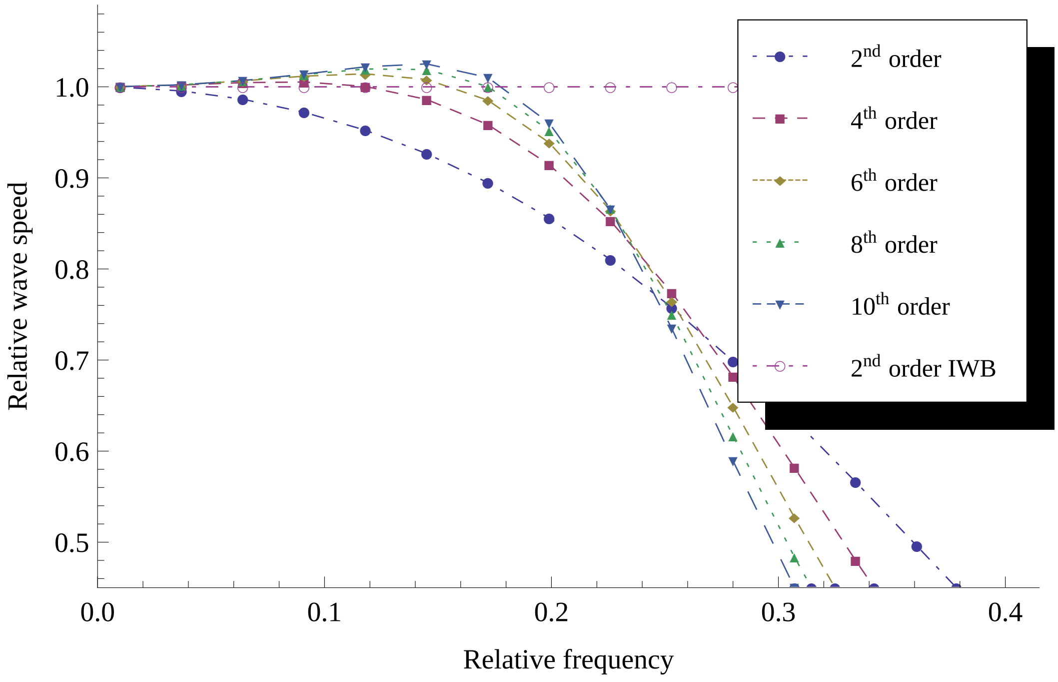
Even though the 3-dimensional dispersion is difficult to plot, we can make use of the fact that the stencil is most dispersive along either the diagonals or the grid axes, due to the minima and maxima of the goniometric functions. It is therefore sufficient to study the phase velocity along these directions only. The functions of dispersion for most of the higher order stencils are omitted as they are very lengthy. Two special cases are of special interest, however. In the 1-dimensional case, evaluated at the stability limit of $\lambda = 1$, the relative phase velocity is 1 everywhere and thus the system is dispersionless:

$$\begin{aligned} v_\phi &= \frac{2}{c\,T\,k}\arcsin\left(\lambda\sqrt{-\frac{1}{4}\left(e^{-jkX}-2+e^{jkX}\right)}\right) \\ &= \frac{2}{c\,T\,k}\arcsin\left(1\sqrt{\sin^2\left(\frac{kX}{2}\right)}\right) \\ &= \frac{2}{c\,T\,k}\frac{k\,X}{2} = 1. \end{aligned} \tag{5.22}$$

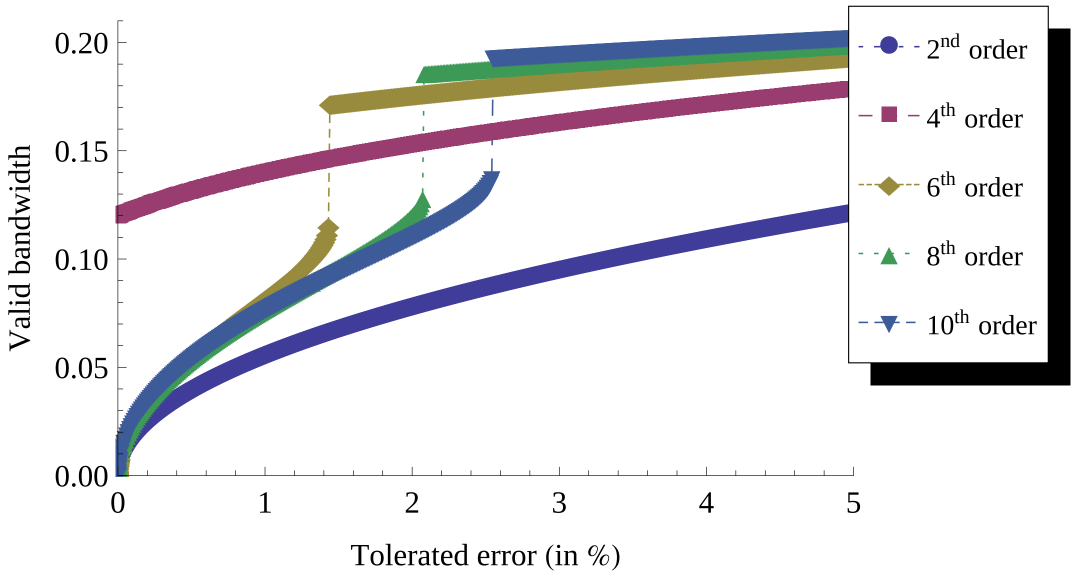
In the 3-dimensional case, wave motion along the diagonal is dispersionless at $\lambda = \frac{1}{3}\sqrt{3}$. In this direction $k_x = k_y = k_z = \frac{1}{3}\sqrt{3}$, so we get:

$$\begin{aligned} v_\phi &= \frac{2}{c\,T\sqrt{k_x^2+k_y^2+k_z^2}}\arcsin\left(\frac{1}{3}\sqrt{3}\sqrt{-\frac{3}{4}\left(e^{-j\frac{1}{3}\sqrt{3}X}-2+e^{j\frac{1}{3}\sqrt{3}X}\right)}\right) \\ &= \frac{2}{c\,T}\arcsin\frac{1}{2}\left(\sqrt{-(-4)\sin^2\left(\frac{X\sqrt{3}}{6}\right)}\right) \\ &= \frac{2}{c\,T}\frac{X\sqrt{3}}{6} = \frac{1}{\lambda}\frac{\sqrt{3}}{3} = 1. \end{aligned} \tag{5.23}$$

Dispersion patterns in all other directions for all other orders of accuracy are more involved and their analytical expressions will be omitted. Only the dispersive properties of the stencils in the axial and diagonal direction are studied as they are the most extreme. Fig. 5.2a shows the relative wave speed along the grid axis of a 3-dimensional system. It shows the $2^{\text{nd}}$- until $10^{\text{th}}$-order accurate scheme, as well as the commonly used Interpolated Wideband (IWB) scheme. More specifics on the latter can be found in e.g. [107]. Orders of accuracy higher than 10 have been studied but are omitted for the sake of clarity of the graphs. They follow the same trend as set by the lower orders. Along the grid axis, the general tendency of the higher-order stencils is that the wave velocity remains relatively stable for slightly longer than in the second-order case. It rises slightly (more so for stencils of higher order) and then drops around $0.2f_s$. The IWB scheme is dispersionless along the grid axes and is therefore represented by a horizontal line.



**(a)** The relative phase velocity in axial grid direction against the relative frequency.



**(b)** The valid bandwidth for a tolerated dispersion error. (The IWB scheme is left out, as it is disperionless in axial direction and therefore has full bandwidth validity.)

**Figure 5.2:** Fig. 5.2a shows the relative phase velocity $\frac{v_\phi}{c}$ along a grid axis for a given normalised frequency. Fig. 5.2b shows the valid bandwidth (as a fraction of the whole bandwidth) for some allowed error margin $\epsilon$ in %.

A similar analysis can be performed on the dispersion in the diagonal direction of the grid, shown in Fig. 5.3a. Along the diagonal the wave velocity becomes increasingly overestimated with increasing frequency, more so for stencils of higher order. The stencils with order of accuracy higher than 10 follow a similar trend, but have been omitted from the plot for the sake of clarity. The difference with the IWB scheme is apparent, as this shows a smoother and slower roll-off (downwards rather than upwards) than all higher-order stencils.

It is clear that the higher order stencils provide a smaller error in the lower frequencies. A convenient way of displaying the quality of a stencil is by plotting the valid bandwidth that can be reproduced for a given margin of error $\epsilon$. Fig. 5.2b plots this for wave propagation in the axial direction. The 4th-order scheme is the most favourable for high confidence simulations ($\epsilon < 1.25\%$), providing the largest valid band width of more than 10%. For less rigorous confidence levels ($\epsilon > 1.25\%$), the higher orders become increasingly more favourable, although the difference is that of a couple of percentages of the valid bandwidth. The IWB scheme is omitted in this graph as it is dispersionless and has therefore full bandwidth validity. The 6th, 8th, and 10th order stencil all show a quick jump in the confidence plot. This is because their phase velocity first goes up slightly before it drops down. The location of the corners in Fig. 5.2b therefore correspond with the location of the respective peaks in Fig. 5.2a. The 4th-order stencil does not seem to portray this behaviour in the graph, but in fact it does so around $f < 0.006 f_{\mathrm{s}}$, which is too low to be visible in this plot.

Fig. 5.3b shows a similar confidence analysis for the schemes in the axial direction. As shown previously, the 2nd-order scheme is dispersionless along the diagonal so is omitted as it has full bandwidth validity. Remarkably, of all higher-order stencils, the fourth-order accurate scheme provides the highest valid bandwidth for all $\epsilon$. This is due to the fact that the dispersion along the diagonal is now the main limiting factor, and the fourth-order scheme shows the least dispersion in this direction, save for the dispersionless second-order stencil. The IWB scheme gives a significantly better performance than all large-star systems, approximately 1.65 times more accurate than the fourth order accurate scheme.

To get an overall statement on the quality of the aforementioned stencils, the valid bandwidth for a given error margin is a good measure. For each stencil, it is enough to take the worst performing case of the axial and diagonal analysis as a measure of the best quality the stencil will produce for a given error margin. This result is presented in Fig. 5.4. We introduce the term $\mathcal{B}(\epsilon)$ to designate the valid bandwidth (on a scale from 0 to 1) for a given error tolerance $\epsilon$. In Table 5.3, several common error tolerances are shown with the corresponding valid bandwidth for the different approximation methods.

It is clear from the plot that the second-order SRL scheme performs worst of all, with only a 7.7% valid bandwidth at a 2% error tolerance. Of all large-star higher-order stencils, the

**Table 5.3:** The valid bandwidth $\mathcal{B}(\epsilon)$ for a given error tolerance $\epsilon$ for all methods and several common error tolerances.

| Order | $\mathcal{B}(1\%)$ | $\mathcal{B}(2\%)$ | $\mathcal{B}(2.5\%)$ | $\mathcal{B}(3\%)$ | $\mathcal{B}(5\%)$ |
|---|---|---|---|---|---|
| 2nd | 0.055 | 0.077 | 0.086 | 0.094 | 0.108 |
| 4th | 0.079 | 0.113 | 0.128 | 0.142 | 0.166 |
| 6th | 0.077 | 0.108 | 0.120 | 0.108 | 0.151 |
| 8th | 0.077 | 0.107 | 0.119 | 0.130 | 0.149 |
| 10th | 0.077 | 0.107 | 0.119 | 0.130 | 0.148 |
| 12th | 0.077 | 0.107 | 0.119 | 0.130 | 0.145 |
| 14th | 0.077 | 0.107 | 0.119 | 0.130 | 0.147 |
| 16th | 0.077 | 0.107 | 0.119 | 0.130 | 0.147 |
| IWB | 0.135 | 0.189 | 0.2113 | 0.231 | 0.266 |

4th-order comes out as the best one: on average it is 1.5 times better than the 2nd-order one. With each increasing order, the stencils lose a little bit of valid bandwidth. This is a remarkable result, as it is often assumed that higher orders of accuracy lead to a higher validity. The IWB outperforms all stencils: it is approximately 1.65 times more accurate than the 4th-order accurate scheme for any given error margin $< 5\%$. However, it must be noted that this increase comes with a stability limit of double that of the 4th-order scheme, and thus at the cost on an eightfold increase in memory usage.

## 5.4 GPU implementation and performance analysis

GPU implementations of the FDTD algorithm have brought the method within the scope of practical use due to the large performance gain [218, 128]. It is therefore worthwhile to discuss the implementations of higher-order stencils in this scope. A significant amount of work has been done in the area of parallelising the FDTD method. Most notably [164, 132, 128] discussed the details of GPU implementations using Nvidia's Compute Unified Device Architecture (CUDA). In this thesis, the implementation by Micikevicius [132] is closely followed, as it is easily extensible and has been designed to deal with spatial stencils of a higher order. Moreover, it aims at minimising read-redundancy by transferring blocks of nearby nodes' pressure values to shared memory, after which data can be accessed more efficiently. This is particularly useful for higher-order stencils, which require many memory reads per node update. As implementations may vary across devices and architectures, Table 5.4 shows the minimum number of arithmetic operation that is required for a single node update.

Micikevicius' was extended to suit a stencil of order $n \leq 16$. Texture memory was used

| Stencil | Add/Substract | Multiply | Memory access |
|---|---|---|---|
| 2nd order | 6 | 1 | 7 |
| 4th order | 13 | 3 | 13 |
| 6th order | 19 | 4 | 19 |
| 8th order | 25 | 5 | 25 |
| 10th order | 31 | 6 | 31 |
| IWB | 27 | 4 | 27 |

**Table 5.4:** The minimum required arithmetic operations for a single node update per stencil.

to store the pressure values in order to allow more efficient memory access. The update coefficients were placed on the GPU's constant memory. Experiments were performed on an Nvidia GPU device: a GeForce GTX 670. This device has a Kepler architecture, compute capability 3.0, and a clock rate of 980 MHz. It was found that block dimensions of $16 \times 16$ threads provided the optimal performance.

As Micikevicius' algorithm can only be applied to stencils that use on-axis neighbouring points (i.e. large-star stencils), a new algorithm had to be developed to implement the IWB in CUDA. In order to be able to make a fair comparison, an adaptation of said algorithm was therefore developed. This algorithm also tries to minimise read-redundancy, which would be particularly prevalent in the IWB scheme, as it requires 27 memory accesses per node update. As in Micikevicius' algorithm, the blocks cover the entire $xy$-plane, and iteration per thread takes place over the $z$-axis. The shared memory is a circular buffer that contains 3 instead of one slice, however, for $p_{x,y,z-1}$, $p_{x,y,z}$, $p_{x,y,z+1}$. This shared memory buffer is updated from the device memory at each iteration over the $z$-axis, after which each node updates its value with the values from shared memory. In order to maximise performance, only the value of the node at the receiver position was transferred from GPU to CPU at each time step.

Micikevicius [132] showed that his implementation has a lower performance rate for higher order stencils. This is expected behaviour, as more nodes are taken into consideration. However, we can make use of the fact that the Courant stability limit drops, which implies that a lower spatial sample rate returns similar results. As a result, the slower methods will also have less data to process, and may effectively be faster (see Table 5.2).

## 5.5 Results

### 5.5.1 Test set-up

This experiment has formerly been published in [208]. Three scenarios were defined, all shoebox type rooms of unequal dimension, one large (7.0m × 5.0m × 4.0m), one medium-sized (5.0m × 3.5m × 2.0m), one small (1.5m × 1.25m × 1.0m). All simulations were run at an audio rate of 44100 Hz, at the respective stability limit of each stencil. From this, the grid spacing and thus the number of nodes in the simulation was calculated. The boundaries were all clamped at 0, i.e. phase-reversing perfectly reflective boundaries. The grid excitation was a Dirac impulse implemented as a soft source [172]. A raw impulse response was generated by taking the value of a grid point at each time step at a fixed receiver position. Post-processing and convolution with anechoic material would be necessary to auralise the acquired impulse responses, but this is beyond the scope of this study.

### 5.5.2 Performance speed

The computation speed of the simulations are presented in Table 5.5. The dimensions (in number of nodes), number of nodes, and frames per second (fps) have been listed for all cases. It is immediately clear that the second-order scheme is not optimal: it uses too many nodes, which is a bottleneck for the performance. The $4^{\text{th}}$- and $6^{\text{th}}$-order schemes perform significantly better: they are around 40% faster and use less memory. In the medium-sized room, the eighth-order stencil has the best overall performance, although this is not the case in the other two scenarios. For orders of accuracy 8 up to 14, the computational speed decreases with every increase in order. This can be explained by the fact that the relative memory gain is smaller for each increase in order, and is outweighed by the increasing computational requirements.

Interestingly, the sixteenth order stencil shows a sudden increase in performance again. In the medium-sized room, it even matches the sixth-order stencil in speed. This can most likely be attributed to more efficient memory accessing due to memory coalescing on the GPU device.

Because of the large amount of memory required by the IWB scheme at the same audio rate, it performs fairly poorly compared to the large-star stencils. As mentioned before, the memory requirement is 5.19 times that of the SRL scheme, and 13.14 times that of the sixteenth order scheme. Also, the number of arithmetic operations per node update is high and more demanding for the GPU because of the many memory access requests. As a result,

**Table 5.5:** The computational speed of the three different scenarios for stencils of second up to sixteenth order, and the second-order IWB scheme. The room dimensions are given in number of nodes, the computational speed in frames per second.

| | Small room (1.5 × 1.25 × 1.0) | | | Medium-sized room (5.0 × 3.5 × 2.0) | | | Large room (7.0 × 5.0 × 4.0) | | |
|---|---|---|---|---|---|---|---|---|---|
| Order | Dimensions | # nodes | fps | Dimensions | # nodes | fps | Dimensions | # nodes | fps |
| 2nd | 111 × 92 × 74 | 755 688 | 3967.0 | 370 × 259 × 148 | 14 182 840 | 295.3 | 519 × 370 × 296 | 56 840 880 | 73.3 |
| 4th | 96 × 80 × 64 | 491 520 | 5750.4 | 321 × 224 × 128 | 9 203 712 | 402.4 | 449 × 321 × 256 | 36 897 024 | 99.3 |
| 6th | 90 × 75 × 60 | 405 000 | 5365.6 | 301 × 211 × 120 | 7 621 320 | 412.3 | 422 × 301 × 241 | 30 612 302 | 107.7 |
| 8th | 87 × 72 × 58 | 363 312 | 4944.3 | 290 × 203 × 116 | 6 828 920 | 422.4 | 407 × 290 × 232 | 27 382 960 | 105.0 |
| 10th | 85 × 70 × 56 | 333 200 | 4565.5 | 283 × 198 × 113 | 6 331 842 | 372.6 | 397 × 283 × 227 | 25 503 677 | 98.8 |
| 12th | 83 × 69 × 55 | 314 985 | 4260.3 | 278 × 195 × 111 | 6 017 310 | 353.2 | 390 × 278 × 223 | 24 177 660 | 95.3 |
| 14th | 82 × 68 × 55 | 306 680 | 4150.1 | 275 × 192 × 110 | 5 808 000 | 351.1 | 385 × 275 × 220 | 23 292 500 | 90.8 |
| 16th | 81 × 68 × 54 | 297 432 | 4340.9 | 272 × 190 × 108 | 5 581 440 | 416.7 | 381 × 272 × 217 | 22 488 144 | 102.4 |
| 2nd IWB | 192 × 160 × 128 | 3 932 160 | 310.1 | 642 × 449 × 256 | 73 794 048 | 18.7 | 899 × 642 × 513 | 296 082 054 | [overflow] |

the performance speed in the three test cases is more than an order of magnitude slower. In the large room, the room could not be modelled with this method because of a memory overflow.

It is important to note that the valid bandwidth that results from an IWB analysis is significantly higher than that of the higher-order stencils. Modelling the same room at a lower sample rate and upsampling to the same audio rate will therefore yield more favourable results. However, even when applying this technique, the performance does not exceed the fourth-order scheme: the IWB scheme's valid bandwidth is approximately 1.65 times higher than the fourth-order accurate stencil, so the memory usage could be reduced $1.65^3 \approx 4.49$ times. This is still larger than the memory requirements of the fourth-order scheme, and requires the slower algorithm.

### 5.5.3 Further advances

Since the above results were first published in [208], Hamilton *et al.* [82] have reviewed the experiment and repeated the experiments with a different implementation, using the slightly lower spec'd GTX 660Ti GPU device. Their approach does not use Micikevicius' [132] algorithm, and does not make use of the GPU's shared memory but rather makes efficient use of the texture cache. They express their results in 'compute time per node' (CTPN) rather than fps. Their implementation shows a slight speed-up for most cases, notably the standard SRL scheme. Only for orders of accuracy 14 and 16 is there a significant benefit of Micikevicius' algorithm. Another clear difference is in the IWB performance: it was previously mentioned that the shared-memory approach was not optimal for the IWB scheme. Hamilton *et al.* indeed show that their approach is 2.46 faster.

Though the IWB implementation of Hamilton *et al.* [82] is significantly faster than the one

**Table 5.6:** Comparison of timings for the 'medium sized room' between this work and the implementation presented by Hamilton *et al.* [82]. Table from [82].

.

| Order | Dimensions | GTX 670 CTPN (ns) | GTX 660Ti CTPN (ns) [82] | ratio |
|---|---|---|---|---|
| $2^{\text{nd}}$ | 370 × 259 × 148 | 0.239 | 0.162 | 1.47 |
| $4^{\text{th}}$ | 321 × 224 × 128 | 0.270 | 0.256 | 1.06 |
| $6^{\text{th}}$ | 301 × 211 × 120 | 0.318 | 0.271 | 1.18 |
| $8^{\text{th}}$ | 290 × 203 × 116 | 0.347 | 0.366 | 0.95 |
| $10^{\text{th}}$ | 283 × 198 × 113 | 0.424 | 0.375 | 1.13 |
| $12^{\text{th}}$ | 278 × 195 × 111 | 0.471 | 0.441 | 1.07 |
| $14^{\text{th}}$ | 275 × 192 × 110 | 0.490 | 0.574 | 0.85 |
| $16^{\text{th}}$ | 272 × 190 × 108 | 0.430 | 0.567 | 0.76 |
| $2^{\text{nd}}$ IWB | 642 × 449 × 256 | 0.725 | 0.294 | 2.46 |

stated in this work, the computations per node compared to the $4^{\text{th}}$-order scheme are still slightly lower. In return, the IWB gives approximately a 1.65 higher valid bandwidth (see Sec. 5.3) but at the cost of an eightfold memory increase. Therefore, these results also imply that the $4^{\text{th}}$-order scheme is the more favourable one.

As a means of fair comparison between computational speed and valid bandwidth, it is useful to compare the algorithms of different orders in such a way that the domains represent the same valid bandwidth for a given error margin $\epsilon$. This requires resampling of the actual outputs to get the different methods at identical sample rates, but makes for the fairest objective comparison between the algorithms and their performance. The 'medium-sized room' with SRL scheme approximation is chosen as the benchmark, and we use an error toleration $\epsilon$ of 2.5%. All methods are evaluated at their Courant stability limit, and the grid spacing is then adjusted such that their valid bandwidth matches that of the SRL approximation. For example, for the $4^{\text{th}}$-order scheme, the grid spacing $X_{4^{\text{th}}}$ compared to the grid spacing of the SRL scheme $X_{\text{SRL}}$ is:

$$X_{4^{\text{th}}} = \frac{\mathcal{B}_{4^{\text{th}}(2.5\%)}}{\mathcal{B}_{\text{SRL}}(2.5\%)} \cdot \frac{\lambda_{\text{SRL}}}{\lambda_{4^{\text{th}}}} \cdot X_{\text{SRL}} = \frac{0.128}{0.086} \cdot \frac{0.577}{0.5} \cdot X_{\text{SRL}} = 1.716\, X_{\text{SRL}} \tag{5.24}$$

The simulations have been repeated with bandwidth validity compensation, and both the grid sizes and the computation speed (in fps) are displayed in Table 5.7. The computation speed using the aforementioned algorithm by Micikevicius [132] and running on a GTX 670 are demonstrated, and show that the most favourable algorithms here are the $6^{\text{th}}$-order scheme (at 1104.3 fps) and the $4^{\text{th}}$-order scheme (at 1063.0 fps). The other higher-order accurate stencils perform well but are slower, as the computational load of the stencil outweighs the

**Table 5.7:** The 'medium-sized room' evaluated at equal valid bandwidth. The GTX 670 column shows the measurement result using the algorithm from [132]. The rightmost column scales those results with the speed improvements Hamilton *et al.* [82] achieved using a different implementation. Note that these results are theoretical and have not actually been implemented as such, so true values may differ slightly.

.

| Order | Dimensions | # of nodes | Relative grid spacing | GTX 670 (fps) | Hamilton *et al.* [82] theoretical (fps) |
|---|---|---|---|---|---|
| 2nd | $370 \times 259 \times 148$ | 14 182 840 | $X_{\text{SRL}}$ | 295.3 | 434.0 |
| 4th | $216 \times 151 \times 86$ | 2 804 976 | $1.716\,X_{\text{SRL}}$ | 1063.0 | 1126.8 |
| 6th | $216 \times 151 \times 86$ | 2 804 976 | $1.716\,X_{\text{SRL}}$ | 1104.3 | 1303.1 |
| 8th | $210 \times 147 \times 84$ | 2 593 080 | $1.765\,X_{\text{SRL}}$ | 916.0 | 870.2 |
| 10th | $205 \times 143 \times 82$ | 2 403 830 | $1.806\,X_{\text{SRL}}$ | 936.2 | 1057.9 |
| 12th | $201 \times 141 \times 80$ | 2 267 280 | $1.838\,X_{\text{SRL}}$ | 899.3 | 962.3 |
| 14th | $198 \times 139 \times 79$ | 2 174 238 | $1.863\,X_{\text{SRL}}$ | 868.5 | 738.2 |
| 16th | $196 \times 137 \times 78$ | 2 094 456 | $1.883\,X_{\text{SRL}}$ | 889.5 | 676.0 |
| 2nd IWB | $262 \times 183 \times 104$ | 4 986 384 | $1.414\,X_{\text{SRL}}$ | 273.3 | 672.3 |

gain in domain size decrease. The two 2nd-order accurate stencils (SRL and IWB) both perform poorest. In order to get a comparison that better represents the state-of-the-art IWB implementation, the results were scaled with the algorithmic improvements Hamilton *et al.* [82] reported, by multiplying our results with the performance ratios stated in Table 5.6. Note that the actual simulations have not been performed, so the scaling is theoretical and actual results may differ from speeds reported here. While the two 2nd-order algorithms perform comparatively better in these scenarios, the 6th- and 4th-order accurate scheme are still significantly more optimal.

## 5.6 Spectral analysis

In order to verify the results of the different FDTD schemes and study their dispersion characteristics, another scenario was defined: a cube of $16 \times 16 \times 16$ nodes. The cubical shape was chosen because the room modes across all axes should be equal. The small room size makes it possible to make dispersion error visible in the frequency spectrum. Impulse responses of 10 seconds were obtained by running the same scenario for all stencils at audio rate at their respective stability limit. Hence, the physical size of the room varies, and the room modes were computed accordingly for each stencil. The source and receiver were positioned in opposite corners as to accommodate excitation of as many different room modes as possible. The boundaries were perfectly reflective phase reversing surfaces.

Fig. 5.5 shows the frequency spectra up to $0.1 f_s$ for orders of accuracy 2 to 12, and the IWB scheme, with the theoretical room modes as dotted lines. It can be clearly observed that all stencils show some frequency warping higher up in the spectrum. The second-order schemes are the only ones that show a frequency drift to the left, which becomes visible around the fourth peak in the spectrum. In the other stencils, the room modes become increasingly higher than the theoretical mode values. This is in accordance with the dispersion analysis presented in Sec. 5.3: using higher-order stencils results in an overestimated wave speed for the majority of directions.

In order to compare and objectively quantify the modal shifts, a table was constructed with the measured peaks versus their theoretical values. The theoretical room modes were calculated for each model separately and matched to the peak values observed in Fig. 5.5, then the percentage error for each measured peak was computed. Fig. 5.6 plots the percentage error for each of the first 32 room modes. The room modes are indexed according to their theoretical occurrence order, and the exact configurations are shown in Table 5.8. As all room dimensions were equal, permutations of axial configurations are ignored.

The results presented in Fig. 5.6 are in close agreement with the theoretical results in dispersion predicted in Fig. 5.3a. The two second-order schemes only show a negative dispersion, and both show distinctive differences in wave speed for different modes. It can be noted that for the SRL's modal series where one axial direction dominates over any other (e.g. {1, 1, 3}, {1, 1,4}), the error is greater than with other modes due to the fact that dispersion is maximal for directions aligned axially with the FDTD grid. For the IWB's modal series, the reverse is true, which is indeed predicted by theoretical dispersion analysis [107]. Modes higher up in the spectrum are visibly rearranged as a result of this effect. The higher-order stencils show a smoother dispersion pattern and model an increased wave speed. As predicted by theoretical analysis the fourth-order stencil shows the least dispersion, though its dispersion pattern is evidently less smooth than the stencils of higher orders. With the order of the stencil increasing the stencils become more isotropic and the dispersion pattern for low frequencies becomes more even across all directions. Therefore, frequency warping techniques [169] could potentially be applied to improve the accuracy further, though this is beyond the scope of this work. Note that for the sake of clarity and space, only the graphs up to 12th order are shown. It was verified however that higher orders continue much the same pattern and have thus been omitted.

**Table 5.8:** The index of the room mode (as also used in Fig. 5.6) versus its axial configurations $\{x, y, z\}$. As the test case was a cube, all permutations of the same axial configurations are equal, so only one of each is listed.

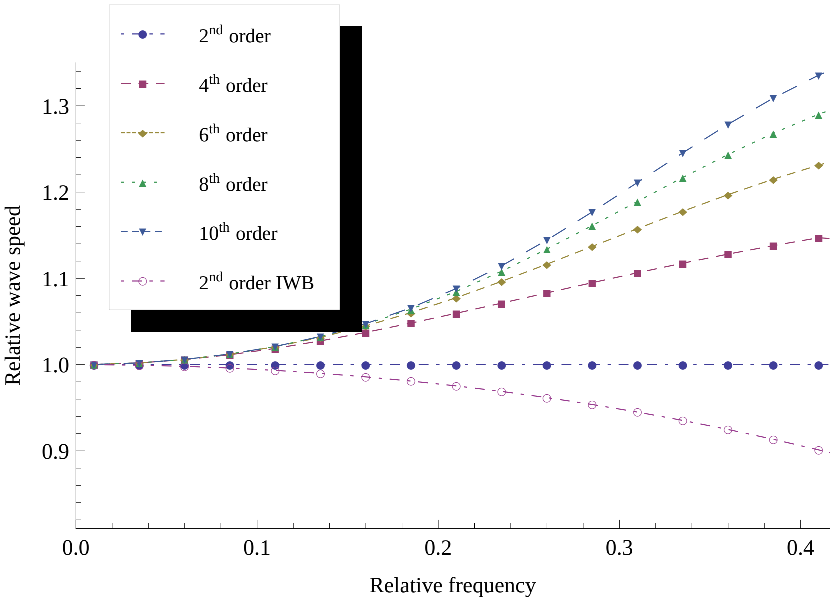
$1 \rightarrow \{ 1, 1, 1 \}$
$2 \rightarrow \{ 1, 1, 2 \}$
$3 \rightarrow \{ 1, 2, 2 \}$
$4 \rightarrow \{ 1, 1, 3 \}$
$5 \rightarrow \{ 2, 2, 2 \}$
$6 \rightarrow \{ 1, 2, 3 \}$
$7 \rightarrow \{ 2, 2, 3 \}$
$8 \rightarrow \{ 1, 1, 4 \}$
$9 \rightarrow \{ 1, 3, 3 \}$
$10 \rightarrow \{ 1, 2, 4 \}$
$11 \rightarrow \{ 2, 3, 3 \}$
$12 \rightarrow \{ 2, 2, 4 \}$
$13 \rightarrow \{ 1, 3, 4 \}$
$14 \rightarrow \{ 3, 3, 3 \}$
$15 \rightarrow \{ 1, 1, 5 \}$
$16 \rightarrow \{ 2, 3, 4 \}$
$17 \rightarrow \{ 1, 2, 5 \}$
$18 \rightarrow \{ 2, 2, 5 \}$
$19 \rightarrow \{ 1, 4, 4 \}$
$20 \rightarrow \{ 3, 3, 4 \}$
$21 \rightarrow \{ 1, 3, 5 \}$
$22 \rightarrow \{ 2, 4, 4 \}$
$23 \rightarrow \{ 2, 3, 5 \}$
$24 \rightarrow \{ 1, 1, 6 \}$
$25 \rightarrow \{ 3, 4, 4 \}$
$26 \rightarrow \{ 1, 2, 6 \}$
$27 \rightarrow \{ 1, 4, 5 \}$
$28 \rightarrow \{ 3, 3, 5 \}$
$29 \rightarrow \{ 2, 2, 6 \}$
$30 \rightarrow \{ 2, 4, 5 \}$
$31 \rightarrow \{ 2, 3, 6 \}$
$32 \rightarrow \{ 4, 4, 4 \}$

## 5.7 Conclusion

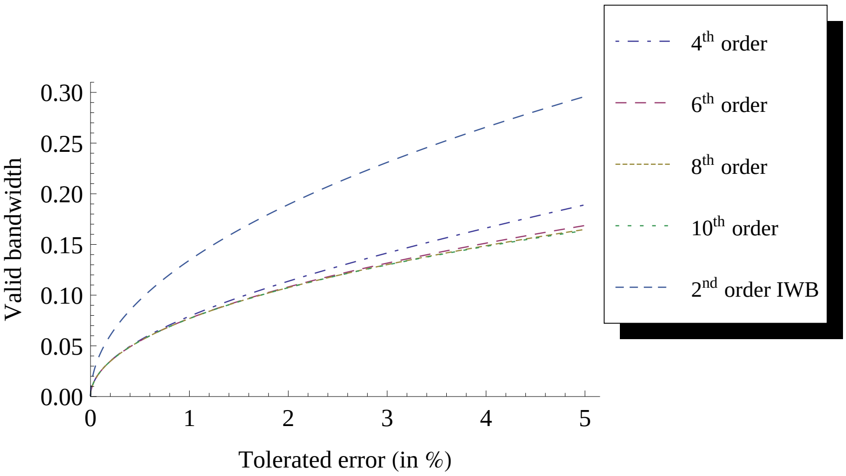
A detailed analysis of the stability and dispersive properties of higher-order stencils has been presented in this chapter. The update coefficients of stencils up to $20^{\text{th}}$ order have been presented for future reference, along with the appropriate stability limits. Higher order stencils provide a lower Courant stability limit, which implies a larger grid spacing. As a result, memory savings—and resulting computational savings—up to 64.6% are possible.

A convenient way of displaying FDTD stencils' quality is proposed, plotting valid bandwidth versus a given error margin. Dispersion analysis shows that of the 'leggy' higher-order stencils, the $4^{\text{th}}$-order scheme is optimal in terms of valid bandwidth for all accuracy errors $\epsilon < 5\%$. The current state-of-art stencil however, the IWB, is approximately 1.65 times more accurate than the fourth-order scheme over the same range. However, this number does not account for memory usage and computation time. The 3-dimensional IWB uses 8 times the amount of memory as the $4^{\text{th}}$-order scheme, which puts it at a distinct disadvantage for modelling the acoustics of large environments. The dispersion patterns of all stencils have been thoroughly examined and it was shown that they closely match the theoretical results.

Although the aforementioned stencil trivially extends to more complex geometries, impedance boundary conditions such at the Locally Reacting Surfaces method [107, 218] for the SRL scheme are necessary to make this method appropriate for practical acoustic modelling. The next chapter will address this.



**(a)** The relative phase velocity in diagonal grid direction against the relative frequency.



**(b)** The valid bandwidth for a tolerated dispersion error. (The SRL scheme is left out, as it is disperionless in diagonal direction and therefore has full bandwidth validity.)

**Figure 5.3:** Fig. 5.3a shows the relative phase velocity $\frac{v_\phi}{c}$ along a grid diagonal for a given normalised frequency. Fig. 5.3b shows the valid bandwidth (as a fraction of the whole bandwidth) for some tolerated error level $\epsilon$ in %.

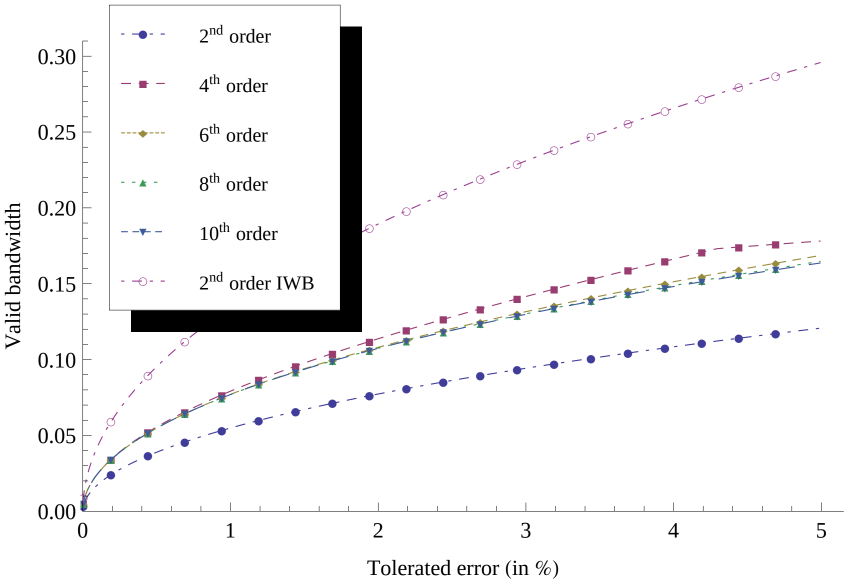


**Figure 5.4:** The valid bandwidth as a function of the allowed error margin for several stencils with different orders of accuracy.

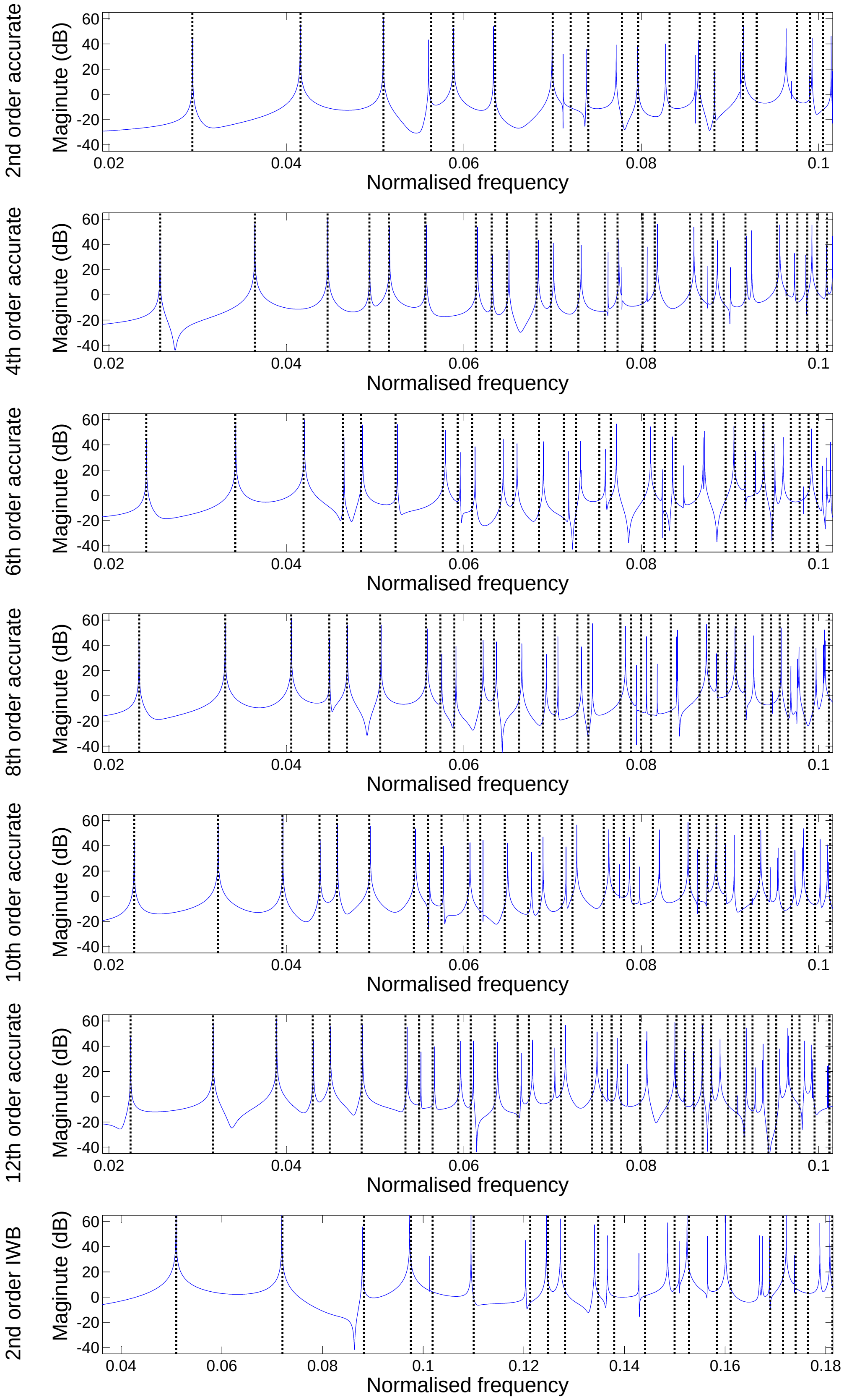


**Figure 5.5:** The frequency spectrum (from 0.02 up to $0.1f_s$) of the impulse responses of a cubic $16 \times 16 \times 16$ room. The dotted lines show the theoretical room modes for each case. Note that the results from the IWB scheme are shown on a different scale, as the room modes only show much higher in the normalised spectrum.

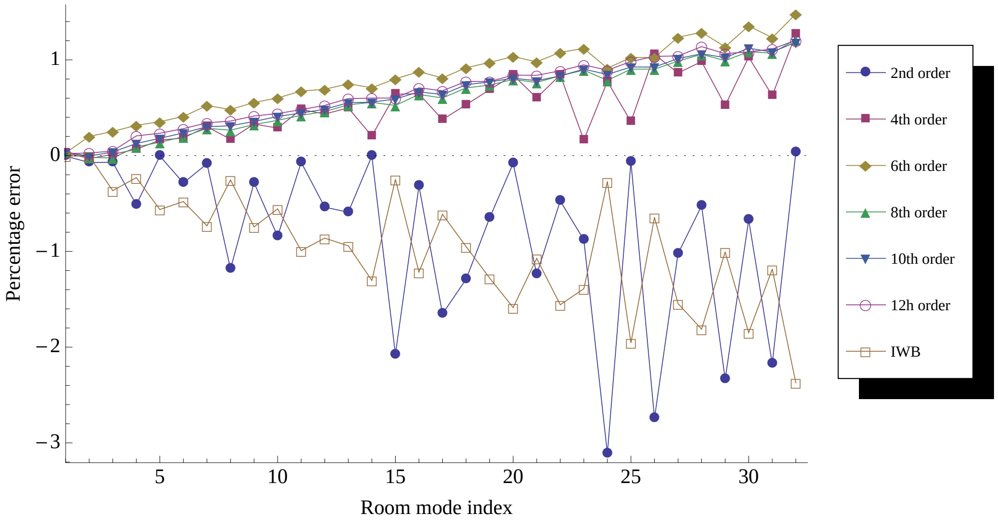


**Figure 5.6:** The percentage error of each measured room mode with respect to its theoretical value, for each mode index $i$. The mode indices are ordered, and their axial configuration can be found in Table 5.8.

# Chapter 6

# Higher order FDTD boundaries

**Contents**

The previous section describes the derivation and implementation of the wave equation in the free field, and proposes the $4^{\text{th}}$-order accurate ‘leggy’ scheme as an improvement over its $2^{\text{nd}}$-order counterparts. In order to be of practical use in room acoustic modelling, however, it is necessary to also be able to deal with wave propagation near boundaries. Of particular interest is the impedance equation, which expresses the behaviour of an absorbing boundary.

The basic substitution method as in Eq. 4.19 doesn't work for higher-order accurate stencils, and thus a different way of deriving stencils with the same properties is necessary. Developing stable time-dependent boundaries is not an easy process. While stable boundaries have been subject to extensive research in Finite Difference modelling, the particular combination of a second-order-in-time equation and a time-dependent boundary poses a particular problem. This chapter discusses several of these methods, and attempts to apply them to the $4^{\text{th}}$-order accurate wave and impedance equation. It shows several approaches and puts forward one candidate, which is investigated further for stability and absorptive properties.

The outline of this chapter is as follows. In Sec. 6.1.1. the energy method is briefly explained and exemplified. The discrete use case of it, summation-by-parts operators, is discussed in Sec. 6.1.2. Boundary interactions using the simultaneous approximation terms method are then covered in Sec. 6.1.3. Both are applied to the wave equation and Sec. 6.1.4 presents an analytical stability analysis of rigid boundaries. Sec. 6.2.1 discusses extending this approach to impedance boundaries. Using all this, in Sec. 6.2.3 boundary stencils are derived and analysed that model the impedance boundary condition in 1 up to 3 dimensions. Finally, a qualitative analysis of these stencils and their absorptive properties is discussed in Sec. 6.5.

## 6.1 Background

In order to arrive at a stable formulation of higher-order boundary stencils, some theoretical background of discrete PDE modelling is required. This is necessary to derive (stable) formulations of higher-order finite difference approximations of the impedance boundary equation. Essential to the stability analysis of many bounded and unbounded systems is the 'energy method', discussed in Sec. 6.1.1. The energy method is a general method for determining the stability of PDEs. The field of GKS theory (named after its main contributors Gustafsson, Kreiss, and Sundström) deals with the stability analysis of discrete difference problems. The essentials of GKS theory are discussed in Sec. 6.1.2. Dealing with boundaries and boundary conditions on a finite grid is a challenge that is addressed using *simultaneous approximation terms*, a method devised in, among others, [68] and [38]. Sec. 6.1.3 discusses the principles of this method. Using these constructs, Sec. 6.1.4 shows a way of obtaining an analytical expression for the error introduced by discrete wave equation modelling.

### 6.1.1 The energy method

What is now commonly known as 'the energy method' is an approach in which an exact expression of the energy in a system is derived from the differential equation that governs it. It has been described numerous times, (e.g. [75, 18, 196, 50]), and has an advantage in that it works for both the continuous and, under certain conditions, in the discrete case.

Without loss of generality, we consider the function $p(x,t)$ on the semi-bounded domain $[0,\infty]$, where we assume $p(x,t)$ goes to 0 for $x \to \infty$. This is convenient as it allows us to study one boundary at the time. The following definitions are required. Let the inner product of two real functions, continuous in time and space, $p(x,t)$ and $q(x,t)$ be defined as:

$$(p(x,t), q(x,t))_{w(x)} = (p,q)_w \overset{\text{def}}{=} \int_0^\infty p(x,t)\, w(x)\, q(x,t)\, dx, \tag{6.1}$$

where $w(x) > 0$. The norm of $p(x,t)$ on $w(x)$, written $||p(x,t)||^2_{w(x)}$, is then defined as:

$$||p(x,t)||^2_{w(x)} = ||p||^2_w \overset{\text{def}}{=} (p,p)_w = \int_0^\infty p(x,t)\, w(x)\, p(x,t)\, dx. \tag{6.2}$$

The common case where $w(x) = 1$ is referred to as simply 'the norm of $p(x,t)$', and $w$ is omitted in the notation: $||p||^2$. When denoted without subscript $w$, $w$ is assumed to be 1. For brevity's sake, we define $\frac{\partial}{\partial t}\, p(x,t) \overset{\text{def}}{=} p_t$, $\frac{\partial^2}{\partial x^2}\, p(x,t) \overset{\text{def}}{=} p_{xx}$, etcetera. Several useful scalar product relations can be derived using the chain rule and integration by parts relations:

$$\begin{aligned}
(p, p_t) &= \frac{1}{2}\frac{d}{dt}||p||^2 \\
(p_t, p_{tt}) &= \frac{1}{2}\frac{d}{dt}||p_t||^2 \\
(p, p_x) &= \int_0^\infty p\, p_x\, dx = \left[p^2(x,t)\right]_{x=0}^{x=\infty} - \int_0^\infty p\, p_x\, dx \\
&= -\frac{1}{2}\, p^2(0,t) \\
(p, p_{xx}) &= -||p_x||^2 + [p(x,t)\, p_x(x,t)]_{x=0}^{x=\infty} \\
&= -||p_x||^2 - p(0,t)\, p_x(0,t)
\end{aligned} \tag{6.3}$$

As an example, consider now the following PDE, first-order in time and $2^{\text{nd}}$-order space and commonly known as the diffusion equation, on the finite domain [0, 1]:

$$\begin{cases} p_t = \alpha\, p_{xx} \\ p(0,t) = 0 \\ p(x,0) = f(x), \end{cases} \tag{6.4}$$

with $\alpha > 0$.

Taking the scalar product of both sides of Eq. 6.4 with $p(x,t)$ and using integration by parts yields:

$$\begin{aligned} (p, p_t) &= \alpha(p, p_{xx}) \\ \frac{1}{2}\frac{d}{dt}||p||^2 &= \alpha(p, p_{xx}) \\ &= \alpha\,[p, p_x]_0^\infty - \alpha(p_x, p_x) \\ &= -\alpha\, p(0,t)\, p_x(0,t) - \alpha(p_x, p_x) \\ &= -\alpha||p_x||^2 \end{aligned} \tag{6.5}$$

The interpretation of this is that the energy in the whole system (which is proportional to $||p||^2$) is strictly non-increasing (since $||p_x||^2$ is non-negative), and thus stable.

### 6.1.2 Summation-by-parts operators

Gustafsson, Kreiss, and Sundström [108, 76, 110] were among the first to apply the energy method to discrete problems—this field has thence come to be known as GKS theory. These initial works considered the first-order-in-time differential equation of the form:

$$\begin{aligned} \frac{\partial}{\partial t}\, p(x,t) &= \sum_{i=0}^{N} \alpha_i\, \frac{\partial^i}{\partial x^i}\, p(x,t) + F(x,t) \\ &= A\, p(x,t) + F(x,t), \end{aligned} \tag{6.6}$$

where $F(x,t)$ is some driving function. Strikwerda [198] demonstrated the usefulness of the semi-discretisation approach of this system, commonly referred to as the *method of lines.* In this approach, time remains continuous and space is discretised. Kreiss and Wu [110] rigorously showed that if a semi-discrete system is stable, then, under mild restrictions, a fully discretised version of that system using Runge-Kutta or multistep techniques is also stable. This is an extremely useful result. Fully discrete systems can be much more complicated than semi-discrete systems, especially for multistep approximations. Moreover, this result allows the stability analysis to focus purely on $A$.

Let us discretise the domain $p(x)$ as $\mathbf{p} = \{p(x_0), p(x_1), ...\}^T = \{p_0, p_1, ...\}^T$, where $x_i = x\,i\,h$, and $h$ is the grid spacing, and $i = \{0, 1, 2, ...\}$. In order to be able to perform energy analysis on the discrete domain $\mathbf{p}$, we need (1) a definition of the scalar product; (2) a definition of the norm of $\mathbf{p}$; (3) difference operators that can be constructed on the entire domain. As utility notation, we will define: $\mathbf{e}_0 = \{1, 0, 0, ...\}$, and $\mathbf{E}_0 = \mathrm{diag}(\mathbf{e}_0)$. Let us define the scalar product on a grid as:

$$\langle \mathbf{p}, \mathbf{q} \rangle = \sum_{i=0}^{\infty} p(x_i) H_i \, q(x_i) h, \tag{6.7}$$

where $H_i > 0$. This can also be represented in matrix form:

$$\langle \mathbf{p}, \mathbf{q} \rangle_H = \mathbf{p}^T H \mathbf{q}, \tag{6.8}$$

where $H$ is a diagonal matrix containing all entries $H_i$. The norm is then defined as:

$$||\mathbf{p}||_H^2 = \langle \mathbf{p}, \mathbf{p} \rangle_H. \tag{6.9}$$

In fact, this is just a special case of the more general form, in which $H$ can take non-diagonal form near the boundaries [75]. (When the norm is denoted without subscript, $H$ is assumed to be $I$.) This however makes stable multi-dimensional schemes difficult if not impossible to construct (see e.g. [75]), and we will thus not digress into considering these. Summation-by-parts (SBP) operators represent the discrete equivalent of integration by parts. By definition, a matrix $D$ is a SBP operator if its behaviour mimics the properties of its continuous counterparts (see Eq. 6.5). A first-order spatial derivative for example is a SBP if for all $\mathbf{p}, \mathbf{q}$:

$$\langle \mathbf{p}, HD\,\mathbf{q} \rangle = -\langle D\,\mathbf{p}, H\,\mathbf{q} \rangle - p_0\, q_0, \tag{6.10}$$

and therefore also:

$$\langle \mathbf{p}, D\,\mathbf{p} \rangle_H = -\frac{1}{2} p_0\, q_0. \tag{6.11}$$

This restriction means that the energy in the system allows itself to be expressed purely in terms of boundary nodes, which is convenient for stability analysis.

Consider again the diffusion equation, where $A = \frac{\partial^2}{\partial x^2}$. Turning this into a semi-discrete equation, we turn $A$ into a yet-to-be-found discrete operator $D^{(2)}$, which approaches the second derivative on $\mathbf{p}$. The semi-discrete system then reads as:

$$\mathbf{p}_t = \alpha\, D^{(2)}\, \mathbf{p}. \tag{6.12}$$

Taking the scalar product with $\mathbf{p}$, we get:

$$\begin{aligned} \langle \mathbf{p}, \mathbf{p}_t \rangle &= \alpha\, \langle \mathbf{p}, D^{(2)}\mathbf{p} \rangle_H \\ \frac{1}{2}\frac{d}{dt}||\mathbf{p}||^2 &= \alpha\, \langle \mathbf{p}, HD^{(2)}\mathbf{p} \rangle \end{aligned} \tag{6.13}$$

For $D^{(2)}$ to be a SBP operator, by definition this has to equate to:

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}||\mathbf{p}||^2 &= \alpha\, \langle \mathbf{p}, HD^{(2)}\mathbf{p} \rangle \\ &= \alpha\, ||D^{(1)}\mathbf{p}||_H^2 - p_0\, D^{(1)}\, \mathbf{e}_0\, \mathbf{p}, \end{aligned} \tag{6.14}$$

where $D^{(1)}$ is an approximation of the first derivative on $\mathbf{p}$. A lot of study has gone into constructing SBP operators for first and second derivatives of several different orders of accuracy (see e.g. [109, 37, 195, 147, 148, 131, 1]). It turns out that SBP operators for some derivative and order of accuracy are not unique. This is as expected, as there are many possible approximations of higher-order derivatives. A desirable property of an SBP operator is that its matrix $H$ be diagonal. If this is not the case, then extensions to multi-dimensional systems are much less trivial, if possible at all [75].

An excellent summary and also the full SBP matrices $D^{(1)}$ and $D^{(2)}$ with diagonal $H$ for orders of accuracy 2, 4, 6, and 8, can be found in [75]. All matrixes $H$, $D^{(2)}$, $M^{(2)}$, and $S$ for orders 2, 4, and 6 are also given in A. SBP operators with diagonal $H$ for orders of accuracy higher than 8 have not been found in literature, and to the best of the author's knowledge, it is an unsolved problem whether they exist or not. However, due to practical limitations of implementation and the smaller relative gain in terms of accuracy, they are not extremely relevant to this study. Henceforward, the notation $D^{(1)}$ and $D^{(2)}$ will refer to the aforementioned matrices in literature, approaching respectively the first and second derivative.

### 6.1.3 Simultaneous approximation terms (SAT)

An interesting and common problem is how to impose a non-constant boundary condition onto the system. Let us impose the boundary condition:

$$p(0,t) = g(t). \tag{6.15}$$

We need a way to incorporate this boundary condition in the semi-discrete update equation, and thus the matrix that approximates the spatial derivative(s). The Simultaneous Approximation Terms (SAT) method [37, 38, 52, 50] provides a means of expressing the system including its above boundary equation, as:

$$\mathbf{p}_t = D^{(m)}\,\mathbf{p} - \tau\,H^{-1}\,\mathbf{e}_0\,(p_0 - g(t))\,. \tag{6.16}$$

The 'penalty parameter' $\tau \in \mathbb{R}$ here is a free parameter, which can be chosen in a way to ensure a stable system. Applying the energy method to Eq. 6.16 with $D^{(m)} = D^{(1)}$ we get:

$$\begin{aligned}
\langle \mathbf{p}, \mathbf{p}_t \rangle &= \left\langle \mathbf{p}, D^{(1)}\,\mathbf{p} - \tau\,H^{-1}\,\mathbf{e}_0\,(p_0 - g(t)) \right\rangle_H \\
\frac{1}{2}\frac{d}{dt}||\mathbf{p}||^2 &= -\frac{1}{2}\,p_0^2 - \langle \mathbf{p}, \tau\,\mathbf{e}_0\,(p_0 - g(t)) \rangle \\
&= -\frac{1}{2}\,p_0^2 + \tau\left(p_0^2 - p_0\,g(t)\right) \\
&= -\left(\tau - \frac{1}{2}\right)\,p_0^2 - \tau\,p_0\,g(t)
\end{aligned} \tag{6.17}$$

For this to be stable, it is required that:

$$\tau \geq \frac{1}{2}. \tag{6.18}$$

In this example, it seems superfluous to introduce the extra parameter $\tau$, as the most intuitive choice $\tau = 1$ is within the range of stable possibilities. This is not in general the case, however, and the SAT method can be a powerful tool to stabilise systems that otherwise would not be stable. On the other hand, it also summons the new question of what value for $\tau$ to choose, when a range of values is valid.

### 6.1.4 Analytical error analysis of the wave equation

In this section we will consider the formulation of the wave equation in direct form:

$$\frac{1}{c^2} p_{tt} = p_{xx} \tag{6.19}$$

Applying the energy method to this and using integration by parts, we get:

$$\begin{aligned}\frac{1}{c^2}\left(p_t, p_{tt}\right) &= \left(p_t, p_{xx}\right)\\ \frac{1}{2c^2}\frac{d}{dt}||p_t||^2 &= \left[p_t\, p_x\right]_0^\infty - \int_0^\infty p_{tx}\, p_x dx\\ &= \left[p_t\, p_x\right]_0^\infty - \frac{1}{2}\frac{d}{dt}||p_x||^2\\ \frac{1}{2}\frac{d}{dt}\left(\frac{1}{c^2}||p_t||^2 + ||p_x||^2\right) &= -\, p_t(0)\, p_x(0)\end{aligned} \tag{6.20}$$

Thus the wave equation is stable when the product of $p_t$ and $p_x$ at the boundary is nonnegative. Note that stability immediately follows for rigid boundary conditions, as then $p_t = 0$, and free boundary conditions, for which $p_x = 0$.

Semi-discretisation of second-order-in-time equations does not necessarily work as it does in the the first-order case. The rigorous proof in [110] is only valid for first-order-in-time systems, so we cannot readily assume that a semi-discrete second-order-in-time system will also be stable. Another means of energy analysis is necessary in order to prove stability. Ditkowski [52] demonstrated how an error growth function can be derived directly from a discretisation of the wave equation. Let $D^{(2)}$ again be the approximation to $\frac{\partial^2}{\partial x^2}$ and the approximation error be $\mathbf{T}_e$, then:

$$\begin{cases}\frac{1}{c^2}\,\mathbf{p}_{tt} = \left(D^{(2)}\,\mathbf{p} + \mathbf{T}_e\right) + \mathbf{F}(x,t)\\ p(0,t) = g(t)\end{cases} \tag{6.21}$$

describes our full system with boundary condition. If $\mathbf{p}$ represents the exact solution to Eq. 6.19 at discrete points, then let $\mathbf{q}$ be the numerical solution:

$$\frac{1}{c^2}\,\mathbf{q}_{tt} = D^{(2)}\,\mathbf{q} + \mathbf{F}(x,t). \tag{6.22}$$

We will ensure that at $t = 0$, $\mathbf{p} = \mathbf{q}$. If we add the boundary condition $p(0,t) = g(t)$ to this discrete approximation, we can formulate the SAT boundary condition. Let us make the boundary condition approximation slightly more general than in the previous section by allowing $B_0$ to be a matrix approximating boundary conditions at $p(0,t)$, and $\mathbf{g}(t) = \{g(t), g(t), ...\}$:

$$\frac{1}{c^2}\,\mathbf{q}_{tt} = D^{(2)}\,\mathbf{q} - \tau\,\left(B_0\,\mathbf{q} - \mathbf{g}(t) + \mathbf{T}_L\right) + \mathbf{F}(x,t). \tag{6.23}$$

where $\mathbf{T}_L$ is a vector containing the error at the left boundary introduced by the SAT method. In the majority of cases, $B_0$ will be 0 everywhere except for the first (few) row(s) and column(s). If we define the error $\epsilon = \mathbf{p} - \mathbf{q}$ and subtract Eq. 6.23 from Eq. 6.21, we get a formulation on the error function:

$$\frac{1}{c^2}\,\epsilon_{tt} = D^{(2)}\,\epsilon - \tau\,B_0\,\epsilon + \tau\,\mathbf{T}_L + \mathbf{T}_e. \tag{6.24}$$

By substituting:

$$M = D^{(2)} - \tau\,B_0 \tag{6.25}$$

we get:

$$\begin{aligned}\frac{1}{c^2}\,\epsilon_{tt} &= M\epsilon + \mathbf{T}_e + \tau\,\mathbf{T}_L\\ &= M\epsilon + \mathbf{T}_T,\end{aligned} \tag{6.26}$$

with $\mathbf{T}_T = \mathbf{T}_e + \tau\,\mathbf{T}_L$. This is an ordinary $2^{\text{nd}}$-order differential equation in $\epsilon$. If $M$ can be diagonalised, e.g. expressed in the form $M = Q^{-1}\,\Lambda\,Q$ with diagonal matrix $\Lambda$, then writing $\mu = Q\,\epsilon$, we get:

$$\frac{1}{c^2}\,\mu_{tt} = \Lambda\,\mu + Q\,\mathbf{T}_T. \tag{6.27}$$

$\Lambda$ is a diagonal matrix with entries $\{\lambda_0, \lambda_1, ...\}$, which are the eigenvalues of $M$. Ditkowski goes on to show that, by solving this differential equation for the $m^{\text{th}}$ equation (Eq. 6.27 is an uncoupled system), this gives:

$$\mu_m(t) = c_{m,1}\,e^{\sqrt{\lambda_m}t} + c_{m,2}\,e^{\sqrt{\lambda_m}t} + \frac{1}{\sqrt{\lambda_m}}\int_0^t \sinh\left(\sqrt{\lambda_m}\,(t-s)\right)\,Q\,\mathbf{T}_T(s)\,ds, \tag{6.28}$$

where $c_{m,1}$ and $c_{m,2}$ are some integration constants. In order to avoid exponentially growing solutions, we need to demand that the argument of the hyperbolic sine have no real part. This can only be satisfied if $\lambda_m$ is real and nonpositive. This results in at most linear growth over time proportional to $Q\,\mathbf{T}_T$. In order to guarantee this, $M^{(2)} = D^{(2)} - \tau\,B_0$ needs to be a negative semi-definite matrix by construction.

Ditkowski [52] also showed that a similar result can be obtained using the energy method. If we take the scalar product of Eq. 6.26 and $\epsilon_t$, we get:

$$\begin{aligned} \left\langle \epsilon_t, \frac{1}{c^2}\,\epsilon_{tt} \right\rangle &= \left\langle \epsilon_t, M^{(2)}\,\epsilon + \mathbf{T}_T \right\rangle \\ \frac{1}{2c^2}\frac{d}{dt}||\epsilon_t||^2 &= \frac{1}{2}\frac{d}{dt}\left\langle \epsilon, M^{(2)}\,\epsilon \right\rangle + \left\langle \epsilon_t, \mathbf{T}_T \right\rangle . \end{aligned} \tag{6.29}$$

Integrating with respect to time gives:

$$\begin{aligned} \frac{1}{2}\left( \frac{1}{c^2}||\epsilon_t||^2 - \left\langle \epsilon, M^{(2)}\,\epsilon \right\rangle \right) &= \int_0^t \left\langle \epsilon_s, \mathbf{T}_T \right\rangle ds \\ &= \int_0^t \left\langle \epsilon, \mathbf{T}_T \right\rangle_s - \left\langle \epsilon, \left(\mathbf{T}_T\right)_s \right\rangle ds \\ &= \left\langle \epsilon, \mathbf{T}_T \right\rangle - \int_0^t \left\langle \epsilon, \left(\mathbf{T}_T\right)_s \right\rangle ds \end{aligned} \tag{6.30}$$

Using the Cauchy-Schwarz inequality, which states that $-\left\langle \mathbf{p}, \mathbf{q} \right\rangle^2 \le ||\mathbf{p}||^2\,||\mathbf{q}||^2$, we get that $\left\langle \epsilon, M^{(2)}\,\epsilon \right\rangle \le \lambda_{\min}\,||\epsilon||^2$. Here $\lambda_{\min} = \min(|\lambda_{M^{(2)}}|)$ represents the eigenvalue with the smallest absolute value in $M^{(2)}$. Note that as we demanded $M^{(2)}$ to be negative semi-definite, $\min(|\lambda_M|)$ will be nonnegative. As a result:

$$\begin{aligned} \frac{1}{2}\left( \frac{1}{c^2}||\epsilon_t||^2 + \lambda_{\min}\,||\epsilon||^2 \right) &\le \left\langle \epsilon, \mathbf{T}_T \right\rangle - \int_0^t \left\langle \epsilon, \left(\mathbf{T}_T\right)_s \right\rangle ds \\ \frac{1}{2}\lambda_{\min}\,||\epsilon||^2 &\le ||\epsilon||\,||\mathbf{T}_T|| - \int_0^t \left\langle \epsilon, \left(\mathbf{T}_T\right)_s \right\rangle ds \\ &\le ||\epsilon||\,||\mathbf{T}_T|| + \int_0^t ||\epsilon||\,||\left(\mathbf{T}_T\right)_s||ds \end{aligned} \tag{6.31}$$

Ditkowski finally argues that this simplifies to $\frac{1}{2}\lambda_{\min}||\epsilon|| \le ||\mathbf{T}_T + ||(\mathbf{T}_T)_s||t$, i.e. maximum linear growth of error. He documented the construction of the full matrix $M^{(2)}$ and shows several examples with constant boundary conditions.

The above derivation shows that a semi-discrete system for the wave equation with non-constant time-independent boundary conditions shows at most linear growth. Mattsson *et al.* [130] extended the wave equation model with rigid Dirichlet and Von Neumann boundary conditions, and implemented a stable Finite Volume method using this technique.

## 6.2 Possible approaches for time-dependent boundary conditions

Thus far the only boundary conditions under consideration were time-independent, i.e. of the form:

$$\sum_{n=0}^{N} \beta_i \frac{d^n}{dx^n} p(0,t) = g(t). \tag{6.32}$$

For example Mattsson *et al.* [129, 130] have shown several stable implementation of these in wave-equation modelling. However, an interesting and common boundary condition in acoustic modelling is the impedance condition:

$$\beta_1 \frac{d}{dx} p(0,t) + \xi \frac{d}{dt} p(0,t) = B_0(t)\, p(x,t) = g(t) = 0, \tag{6.33}$$

which also contains a time derivative. Mattsson *et al.* [130] describe this scenario, but only give examples with $\xi = 0$, thus never exploring the computational difficulties this scenario poses. The objective of this chapter is to find a general formulation of these time-dependent boundaries that applies to higher orders or accuracy—at least including $4^{\text{th}}$-order. It is not trivial that boundaries of this kind will be stable, and the following sections describe the mathematical problem with regards to stability analysis that arise from including a time-dependency in the boundary condition(s). To the best of the author's knowledge, no stable higher-order boundary that models the impedance condition has been proposed in literature.

### 6.2.1 Time-dependent boundary condition for the wave equation

Let us reconsider the analytical error analysis in Sec. 6.1.4. This analysis works under the assumption that $B_0$ and therefore also $M^{(2)}$ are constant. Assuming that $B_0$ is an approximation of the impedance condition and thus $B_0 = B_0(t)$ introduces the time-dependeny, and also $M = M(t)$. In this case, Ditkowski's deduction in Eq. 6.29 is incomplete and should instead be:

$$\begin{aligned}
\left\langle \epsilon_t, \frac{1}{c^2}\epsilon_{tt} \right\rangle &= \left\langle \epsilon_t, M^{(2)}(t)\,\epsilon + \mathbf{T}_T \right\rangle \\
\frac{1}{2c^2}\frac{d}{dt}||\epsilon_t||^2 &= \frac{1}{2}\frac{d}{dt}\left\langle \epsilon, M^{(2)}(t)\,\epsilon \right\rangle - \frac{1}{2}\left\langle \epsilon, M_t^{(2)}(t)\,\epsilon \right\rangle + \langle \epsilon_t, \mathbf{T}_T \rangle \\
\frac{1}{2c^2}||\epsilon_t||^2 - \frac{1}{2}\left\langle \epsilon, M^{(2)}(t)\,\epsilon \right\rangle &= \int_0^t -\frac{1}{2}\left\langle \epsilon, M_s^{(2)}(s)\,\epsilon \right\rangle + \langle \epsilon_s, \mathbf{T}_T \rangle\, ds
\end{aligned} \tag{6.34}$$

Whereas previously we could assume a smallest eigenvalue $\lambda_{\text{min}}$ of $M$, this is no longer possible because of the time dependency. We cannot guarantee any value restriction on the eigenvalues of $M^{(2)}(t)$, and therefore this analysis will not lead to a useful stability formulation. Using this method, it is therefore impossible to say whether the impedance boundary condition using the SBP-SAT method results in a stable system.

### 6.2.2 Lower-order approximation at boundaries

As a possible solution, Sakamoto [162] proposed to use lower-order approximations near the boundaries, as the higher-order stencils need to reference invalid nodes in the system. A visual example of this is shown in Fig. 6.1. While to the author's knowledge no experiments have been conducted combining this with the locally reacting surfaces by Kowalczyk and Van Walstijn [104] (published around the same time as [162]), it is likely that the two methods can be combined and result in a stable system. However, the resulting system is not truly a higher-order system anymore: formally, the accuracy of the simulation as a whole is determined by the accuracy of the lowest approximation in the system [38, 75], thus reducing this attempt to a 2nd-order accurate approximation. Indeed, one can see how the boundary nodes will introduce the amount of dispersion error of their kind into the system. In addition to that, the lower-order stencils will be evaluated at a Courant number lower than their stability limit in order to maintain a regular grid spacing, which introduces more dispersion error yet. Hence, this is not an approach we shall pursue.

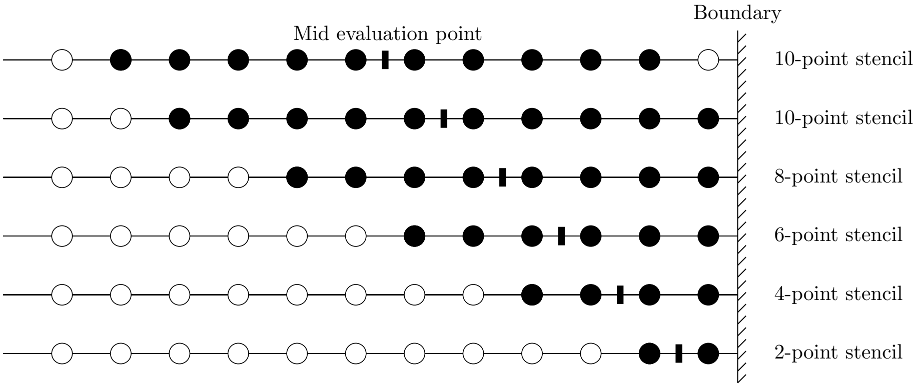


**Figure 6.1:** The boundary stencil updates for a 10th-order accurate stencil near the boundary according to Sakamoto [162]. Near the boundary, lower-order stencils are employed lest an invalid range of nodes beyond the boundary be covered. Image after [162].

### 6.2.3 SBP-SAT method for impedance boundaries

Despite the fact that a stable semi-discrete expression of the wave equation is not guaranteed to give a stable discrete version, there is evidence that several formulations are stable. Mattsson *et al.* [129, 130] performed an elaborate investigation into SBP-SAT methods for the wave equation, with several types of boundary conditions. They describe theory and implementation of fully free (Von Neumann) and rigid (Dirichlet) boundaries. Although they formulate the problem in a way that includes impedance boundaries, they did not implement these, nor is it trivial that these will result in a stable discrete system (for aforementioned reasons).

In Eq. 6.20, the analytical energy method was applied to the wave equation. Let $D^{(2)}$ be a SBP approximation of the second spatial derivative (as in Eq. 6.14). $D^{(2)}$ can also be written as:

$$D^{(2)} = H^{-1}(M^{(2)} + S), \tag{6.35}$$

where $H$ is the diagonal norm, $M^{(2)}$ is a symmetric negative definite matrix approximating the second derivative, and $S$ approximates the first spatial derivative at the boundary and is 0 everywhere else. Definitions for these matrices can be found in e.g. [131, 129, 75]. (Note that slightly different but equivalent notations are used in these works.) The energy method applied to the semi-discrete formulation of the wave equation then becomes:

$$\begin{aligned}
\left\langle H\,\mathbf{p}_t, \frac{1}{c^2}\mathbf{p}_{tt}\right\rangle &= \left\langle \mathbf{p}_t, D^{(2)}\,\mathbf{p}\right\rangle \\
\frac{1}{2c^2}\,\frac{d}{dt}||\mathbf{p}_t||_H &= \frac{1}{2}\frac{d}{dt}M^{(2)}\,\mathbf{p}^2 + (p_0)_t\,S\,\mathbf{p} \\
\frac{1}{2c^2}\,\frac{d}{dt}\left(||\mathbf{p}_t||_H^2 - c^2\,M^{(2)}\,||\mathbf{p}||^2\right) &= (p_0)_t\,S\,\mathbf{p}.
\end{aligned} \tag{6.36}$$

$E = ||\mathbf{p}_t||_H^2 - c^2\,M^{(2)}\,||\mathbf{p}||^2$ represents the energy (note that $-M^{(2)}$ is positive definite), and thus this semi-discrete version mimics Eq. 6.14. In [130], the SAT formulation of the system is then described. First of all, general boundary conditions can be written as:

$$L_0\,\mathbf{p} = \beta_1\,p_0 - \beta_2(p_0)_x + \beta_3(p_0)_t = g(t). \tag{6.37}$$

This includes rigid Von Neumann ($\beta_1 = 0$, $\beta_2 = 1$, $\beta_3 = 0$), Dirichlet ($\beta_1 = 1$, $\beta_2 = 0$, $\beta_3 = 0$) and impedance (defined later) boundary conditions. The SAT formulation of the system is then:

$$\frac{1}{c^2}\,H\,p_{tt} = (M^{(2)} + S)\,p + \tau\,\mathbf{e}_0\,(L_0\,\mathbf{p} - g(t)) \tag{6.38}$$

The energy method applied to this formulation (setting $g(t) = 0$) yields:

$$\begin{aligned}\frac{1}{2c^2}\,\frac{d}{dt}\left(||\mathbf{p}_t||_H^2 - c^2\,M^{(2)}\,||\mathbf{p}||^2\right) &= (p_0)_t\,S\,\mathbf{p} + \tau\,\left(\beta_1\,p_0^2 + \beta_2\,(p_0)_t\,S\,\mathbf{p} + \beta_3(p_0)_t^2\right)\\ &= (1+\tau\,\beta_2)(p_0)_t\,S\,\mathbf{p} + \tau\,\beta_1\,p_0^2 + \tau\,\beta_3(p_0)_t^2\end{aligned} \tag{6.39}$$

It immediately follows that for stability we at least require:

$$\tau = -\frac{1}{\beta_2}, \qquad \frac{\beta_1}{\beta_2} \geq 0, \qquad \frac{\beta_3}{\beta_2} > 0. \tag{6.40}$$

### 6.2.4 Matrix method

While the previous section poses a good starting point for stability analysis, and gives conditions for the continuous system, it does not necessarily follow that the discrete system is also stable. A more specific method to perform stability analysis is to analyse the entire fully discrete system directly. An example of this type of analysis known as *the matrix method* is presented by Strikwerda [199] and was also used by Botts and Savioja [33]. The latter describes the use of this the method to analyse two particular sets of boundary equations by [105] and [218], but the method is more generally applicable to higher-order and SAT-SBP derived update schemes. A disadvantage of the method is that, unlike the energy method or Von Neumann analysis, it has to be applied to one particular domain but has no extrapolative powers. The matrix method is therefore most useful to show the instability of systems, but cannot be used as a proof for stability of a method.

Let us write $\mathbf{p}^n = \mathbf{p}(n\,T)$, where $T$ is the time step variable. The commonly known update equation for the wave equation is:

$$\mathbf{p}^{n+1} = 2\mathbf{p}^n - \mathbf{p}^{n-1} + c^2\,T^2\,D^{(2)}\,\mathbf{p}^n \tag{6.41}$$

If we now make a new vector that contains the state of the system at two time steps: $\widetilde{\mathbf{p}}_n = \left(\mathbf{p}^n, \mathbf{p}^{n-1}\right)$, then the whole system can be expressed as a single matrix update:

$$\widetilde{\mathbf{p}}_n = A\,\widetilde{\mathbf{p}}_{n-1}, \tag{6.42}$$

where $A$ is of the form:

$$A = \begin{pmatrix} c^2\,T^2\,D^{(2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} 2\,\mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}, \tag{6.43}$$

where $\mathbf{I}$ and $\mathbf{0}$ are the identity and zero matrix, respectively, of the same dimensions as $D^{(2)}$. This has the particular advantage that the entire stability analysis can be focussed on $A$, as now:

$$\widetilde{\mathbf{p}}_{n+m} = A^m \, \widetilde{\mathbf{p}}_n \tag{6.44}$$

From this formulation, it immediately follows that the complex eigenvalues of $A$ need to lie within or on the unit circle for this system to be stable, or exponential growth will occur. Thus the problem is reduced to a matrix eigenvalue problem. Note that for the above formulation, this maps the eigenvalues of $c^2\, T^2\, D^{(2)}$, $\lambda_D$, onto the eigenvalues of $A$, $\lambda_A$, according to:

$$\lambda_A = 1 + \frac{1}{2}\lambda_D \pm \sqrt{\frac{1}{4}\lambda_D^2 + \lambda_D}. \tag{6.45}$$

For any $\lambda_D > 0$, this gives a value greater than 1, which confirms the necessity of $D^{(2)}$ being negative semi-definite. Values $-4 \leq \lambda_D \leq 0$ are mapped onto the unit circle uniformly, and $-4 \geq \lambda_D > -\infty$ are mapped on the real axis $[-1, 0)$, therefore also preserving stability.

We would like to incorporate the SBP-SAT method into the matrix method. The SAT formulation Eq. 6.38 can be separated into a time-independent part $Q$ and time-dependent part $\Xi$:

$$\begin{aligned} Q^{(2)} &= H^{-1}\left(M^{(2)} + S + \tau\,(\beta_1\,\mathbf{E}_0 + \beta_2\,S)\right) \\ \Xi &= \tau\,\beta_3\,H^{-1}\,\mathbf{E}_0, \end{aligned} \tag{6.46}$$

such that a first-order approximation to $\frac{d}{dt}\,p_0^n$ is given by:

$$\mathbf{e}_0\,\frac{d}{dt}\,p_0^n = \frac{1}{T}\left(\Xi\,\mathbf{p}^n - \Xi\,\mathbf{p}^{n-1}\right) + O(T^2), \tag{6.47}$$

where $T$ is a discrete small time step. We can then write the entire update in matrix form similar to Eq. 6.43:

$$A = \begin{pmatrix} c^2\,T^2\,Q^{(2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} 2\,\mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} + c^2\,T \begin{pmatrix} \Xi & -\Xi \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{6.48}$$

## 6.3 Methods

In the previous section, a stable formulation for the impedance boundary was sought, and a candidate was found using the SAT-SBP method. This section will investigate the stability of its discrete formulation using the matrix method, and will compare the findings against the known algorithms for absorbing boundaries.

### 6.3.1 $2^{\text{nd}}$-order accuracy

We will consider the absorptive impedance boundary condition:

$$L_0\,\mathbf{p} = \frac{d}{dx}\,p_0 + \frac{\xi}{c}\,\frac{d}{dt}p_0 = 0, \tag{6.49}$$

i.e Eq. 6.36 where $\{\beta_1 = 0, \beta_2 = 1, \beta_3 = \frac{\xi}{c}\}$. $\xi$ represents the admittance factor, as a value between 0 (no absorption) and 1 (full absorption). The aim is to get a fully discrete and stable matrix $A$ for some order of accuracy. Let us start out with the $2^{\text{nd}}$-order approximation of the spatial derivative. We take the symmetric approximation to the spatial derivative $M^{(2)}$ and boundary approximation $S$ from [75]:

$$M^{(2)} = \frac{1}{X^2}\begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \qquad S = \frac{1}{X^2}\begin{pmatrix} 1 & -1 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{6.50}$$

where $X$ is a spatial step. For simplicity, we only write down the left boundary, represented in the top left corner of the matrix. When dealing with a finite domain, also the bottom right corner will represent a boundary, and will be equal to the top left one rotated by a 180°. At this point, we will start out using $H$ taken from [75], which is $H = \text{diag}(\frac{1}{2}, 1, 1, ...)$. This gives us all information to construct $A$, which, setting $\frac{cT}{X} = \lambda$, becomes:

$$A = \begin{pmatrix} 2-2\lambda^2-2\lambda\,\xi & 2\lambda^2 & & & & -1+2\lambda\,\xi & & \\ \lambda^2 & 2-2\lambda^2 & \lambda^2 & & & & 1 & \\ 0 & \lambda^2 & 2-2\lambda^2 & \lambda^2 & & & & 1 \\ & & \ddots & \ddots & \ddots & & & \ddots \\ & \mathbf{I} & & & & & \mathbf{0} & \end{pmatrix} \tag{6.51}$$

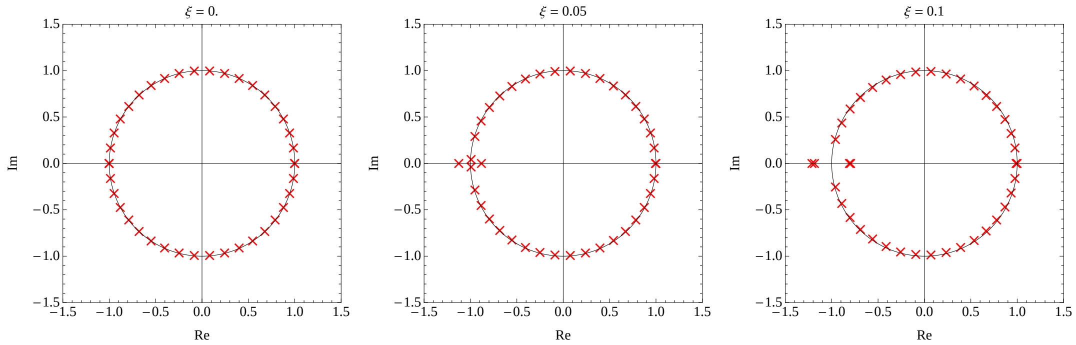


**Figure 6.2:** Polar plot of eigenvalues of a 1D $2^{\text{nd}}$-order accurate $A$ using $H = \text{diag}(\frac{1}{2}, 1, 1, ...)$, for several $\xi$. Though the lower modes indeed move inside the unit circle, thus absorbing these frequencies, the Nyquist mode immediately moves outside, thus causing a globally unstable (exponentially growingly oscillating) system.

Evaluating at $\lambda = 1$ (the Courant stability limit), this is simplified to:

$$A = \begin{pmatrix} -2\xi & 2 & & & & -1+2\xi & & & \\ 1 & 0 & 1 & & & & -1 & & \\ & 1 & 0 & 1 & & & & -1 & \\ & & \ddots & \ddots & \ddots & & & & \ddots \\ & & \mathbf{I} & & & & \mathbf{0} & & \end{pmatrix} \tag{6.52}$$

We can now perform a numerical eigenvalue analysis on $A$, with $0 \leq \xi \leq 1$. We construct a domain of size 20, such that $A$ is of dimensions $40 \times 40$, with absorbing boundary on the one side and a rigid boundary on the other, and slowly increase $\xi$. Fig. 6.2 shows the results of the eigenvalue analysis. It can be readily observed that the eigenvalues for the highest frequency modes are rapidly moving outside of the unit circle along the negative real axis. This therefore does not result in a stable approach.

In a heuristic approach to find a stable boundary model, several different methods have been tried, but for the sake of brevity the unsuccessful ones will not be discussed. Noticing that the scaling of the boundary node seems off, we try an alternative approach, by which we set $H = \mathbf{I}$. Though this is not in line with standard SBP-SAT theory, it seems like a logical choice as a means to scale the node that introduces the instability. $A$ (evaluated at the Courant

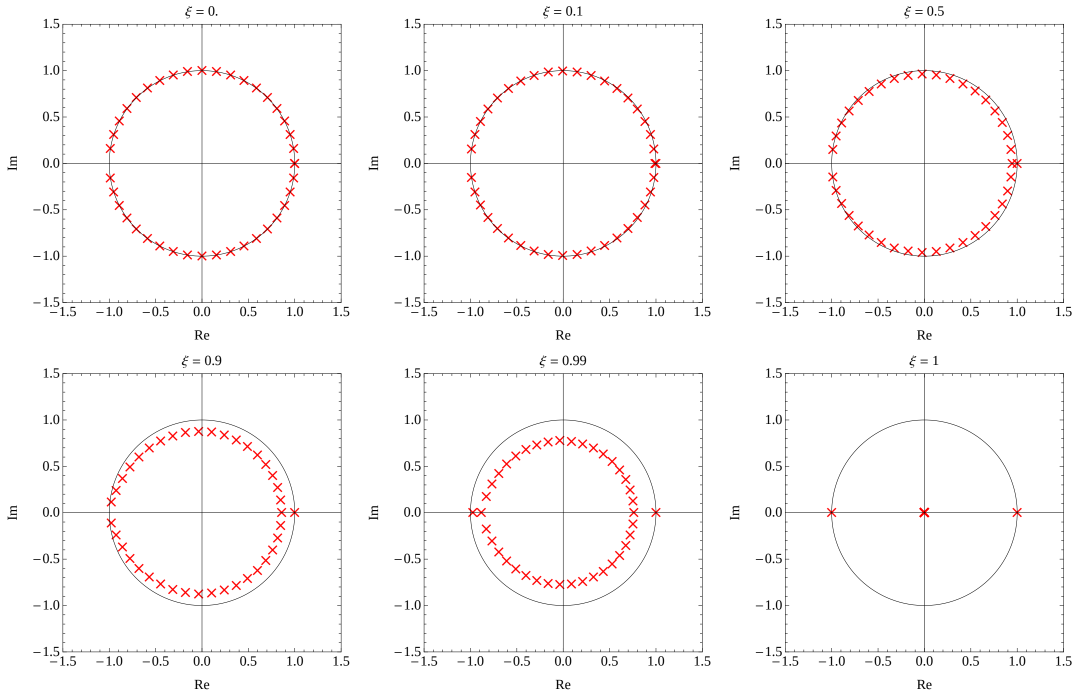


**Figure 6.3:** Polar plot of eigenvalues of 1D 2$^{\text{nd}}$-order accurate $A$ using $H = h\,\mathbf{I}$, for several $\xi$. The system is stable for all $0 \leq \xi \leq 1$, and becomes fully absorbent for $\xi = 1$ for all except the DC component and the Nyquist frequency mode.

limit) then becomes:

$$A = \begin{pmatrix} 1-\xi & 1 & & & & -1+\xi & & & \\ 1 & 0 & 1 & & & & -1 & & \\ & 1 & 0 & 1 & & & & -1 & \\ & & \ddots & \ddots & \ddots & & & & \ddots \\ & & \mathbf{I} & & & & \mathbf{0} & & \end{pmatrix} \tag{6.53}$$

The results of this experiment with several different values of $\xi$ are shown in Fig. 6.3. This boundary indeed behaves in the desired fashion: the higher the $\xi$, the more the system absorbs. The modal spectrum is correct too: for $\xi = 0$ it is perfectly equally spaced—this can be proven analytically, though this is not of current interest—and the modal pattern is preserved very well for higher absorption values. However, it can be noted that the absorption pattern isn't entirely even across all frequencies: low frequencies are absorbed more than high frequencies. For $\xi = 1$, all modal eigenvalues move towards 0, which correctly implies full absorption. The only two other eigenvalues are at ±1, which means that the DC component as well as the Nyquist frequency mode are always maintained by the system.

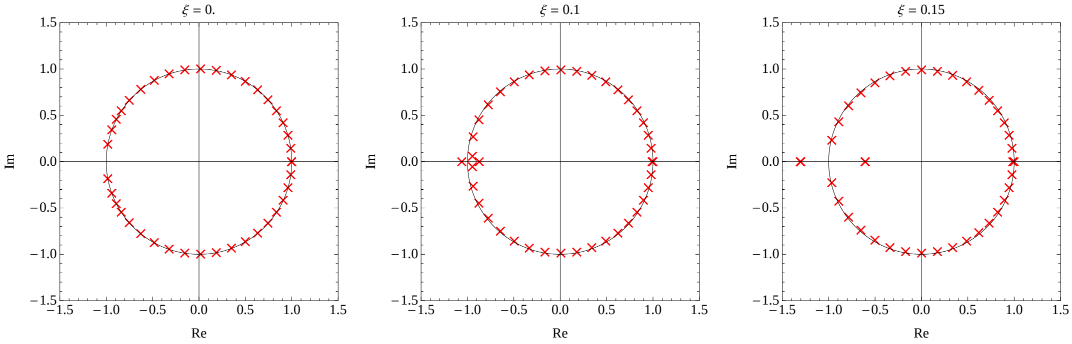


**Figure 6.4:** Polar plot of eigenvalues of 4th-order accurate $A$ using $H = \mathrm{diag}(\frac{17}{48}, \frac{59}{48}, \frac{43}{48}, \frac{49}{48}, 1, ...)$, for several different $\xi$. The system immediately becomes unstable for $\xi > 0$, demonstrating exponential growth around the Nyquist frequency mode.

### 6.3.2 4th-order accuracy

Now that we have a working solution for an impedance boundary in place for a 2nd-order accurate system, higher orders of accuracy should be considered. A requirement for this approach to work is that we have the symmetric formulation of $M$, which are known to exist until 8th order [75]. We use the 4th-order accurate matrices $M$ and $S$ (see A), with $H = \mathrm{diag}(\frac{17}{48}, \frac{59}{48}, \frac{43}{48}, \frac{49}{48}, 1, ...)$, This, evaluated at the stability limit $\lambda = \sqrt{\frac{3}{4}}$, gives for $A$:

$$A = \begin{pmatrix} 2 - \frac{54}{17}\lambda^2 - \frac{48}{17}\lambda\xi & \frac{59}{17}\lambda^2 & -\frac{2}{17}\lambda^2 & -\frac{1}{17}\lambda^2 & & & & & -1 + \frac{48}{17}\lambda\xi & & \\ \lambda^2 & 2 - 2\lambda^2 & \lambda^2 & & & & & & & -1 & \\ -\frac{4}{43}\lambda^2 & \frac{59}{43}\lambda^2 & 2 - \frac{110}{43}\lambda^2 & \frac{59}{43}\lambda^2 & -\frac{4}{43}\lambda^2 & & & & & & \ddots \\ -\frac{1}{49}\lambda^2 & 0 & \frac{59}{49}\lambda^2 & 2 - \frac{118}{49}\lambda^2 & \frac{64}{49}\lambda^2 & -\frac{4}{49}\lambda^2 & & & & & \\ & & -\frac{1}{12}\lambda^2 & \frac{4}{3}\lambda^2 & 2 - \frac{5}{2}\lambda^2 & \frac{4}{3}\lambda^2 & -\frac{1}{12}\lambda^2 & & & & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots & & & \\ & & & \mathbf{I} & & & & & & \mathbf{0} & \end{pmatrix} \tag{6.54}$$

We can evaluate this matrix again for different values of $\xi$ and perform numerical eigenvalue analysis. Fig. 6.4 show these results. It becomes immediately clear that when increasing $\xi$, the Nyquist frequency mode moves outside the unit circle along the negative real axis, thus causing exponentially growing oscillations and an unstable system.

As with the 2nd-order case, we may try the same approach with a different $H$ in order to investigate whether this solves the problem. Setting $H = h\,\mathbf{I}$, the update matrix becomes:

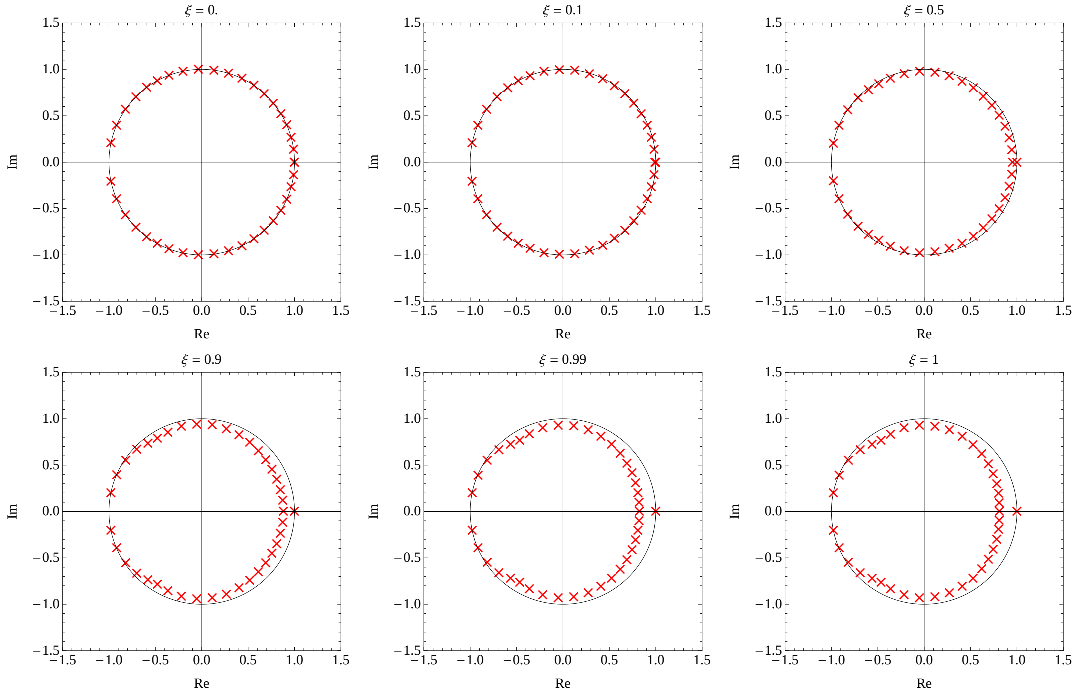


**Figure 6.5:** Polar plot of eigenvalues of 4th-order accurate $A$ using $H = h\,\mathbf{I}$, for several different $\xi$. The system is stable for all $0 \leq \xi \leq 1$.

$$A = \begin{pmatrix}
2-\frac{9}{8}\lambda^2-\lambda\xi & \frac{59}{48}\lambda^2 & -\frac{1}{12}\lambda^2 & -\frac{1}{48}\lambda^2 & & & & -1+\lambda\xi & & \\
\frac{59}{48}\lambda^2 & 2-\frac{59}{24}\lambda^2 & \frac{59}{48}\lambda^2 & & & & & & -1 & \\
-\frac{1}{12}\lambda^2 & \frac{59}{48}\lambda^2 & 2-\frac{55}{24}\lambda^2 & \frac{59}{48}\lambda^2 & -\frac{1}{12}\lambda^2 & & & & & \ddots \\
-\frac{1}{48}\lambda^2 & 0 & \frac{59}{48}\lambda^2 & 2-\frac{59}{24}\lambda^2 & \frac{4}{3}\lambda^2 & -\frac{1}{12}\lambda^2 & & & & \\
 & & -\frac{1}{12}\lambda^2 & \frac{4}{3}\lambda^2 & 2-\frac{5}{2}\lambda^2 & \frac{4}{3}\lambda^2 & -\frac{1}{12}\lambda^2 & & & \\
 & & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\
 & & & \mathbf{I} & & & & & \mathbf{0} &
\end{pmatrix} \tag{6.55}$$

We perform the analysis again, to obtain the results demonstrated in Fig. 6.5. It can be observed that this indeed yields a stable system for all $0 \leq \xi \leq 1$. Low frequency modes are absorbed more than high frequency modes, and the DC component is always maintained. Note that in contrast to the 2nd-order case, there is no Nyquist frequency mode that may cause spurious oscillations. As a source of potential concern, at $\xi = 1$ the system is not fully absorbing, unlike the 2nd-order case. Practical simulations are necessary to determine how much the 4th-order system absorbs compared to the 2nd-order one.

### 6.3.3 Von Neumann stability analysis $2^{\text{nd}}$-order stencil

As the energy method fails here for reasons formerly stated, classic Von Neumann analysis on the boundaries is employed to shed light on the stability of the boundary conditions. Starting with the $2^{\text{nd}}$-order boundary, we can express the update equation as:

$$z - 2 + z^{-1} = \lambda^2 \left(1 - e^{jkX}\right) + \lambda\,\xi \left(1 - z^{-1}\right), \tag{6.56}$$

which we rewrite as:

$$z^2 + \left(-2 + \left(1 - e^{jkX}\right)\lambda^2 - \lambda\,\xi\right) z + (1 - \lambda\,\xi) = 0. \tag{6.57}$$

This is an equation that can be solved for $z$, and for stability we require that $|z| \leq 1$:

$$|z| = \left|\frac{1}{2}\left(2 + \left(-1 + e^{iT}\right)\lambda^2 - \lambda\xi \pm \sqrt{4(\lambda\xi - 1) + \left(-2 + \left(1 - e^{jkX}\right)\lambda^2 + \lambda\xi\right)^2}\right)\right| \leq 1. \tag{6.58}$$

Evaluating this at the stability limit $\lambda = 1$ allows us to write this is a simpler way:

$$\left|\frac{1}{2}\left(1 + e^{jkX} - \xi \pm \sqrt{(\xi - 1 - e^{jkX})^2 + 4(\xi - 1)}\right)\right| \leq 1. \tag{6.59}$$

This expression has its maximum value at $e^{jkX} = -1$, using the negative discriminant:

$$\left|\frac{1}{2}\left(-\xi - \sqrt{\xi^2 + 4(\xi - 1)}\right)\right| \leq 1 \tag{6.60}$$

which gives us the stability condition:

$$\xi \leq 1. \tag{6.61}$$

This is a fortunate result, as this is indeed the maximum value of $\xi$, hence all physical phase preserving admittance values are allowed.

For the 2D case, the Z-transformed update equation of a corner node (assuming both edges to be of admittance $\xi$) gives:

$$z - 2 + z^{-1} = \lambda^2 \left(1 - e^{jk_xX} + 1 - e^{jk_yX}\right) + 2\lambda\,\xi \left(1 - z^{-1}\right), \tag{6.62}$$

Solving for $z$ then gives:

$$\begin{aligned}|z| = \Bigg| \frac{1}{2}\Big( 2 + \left(e^{jk_xX} + e^{jk_yX} - 2\right)\lambda^2 - 2\lambda\,\xi \pm \\ \sqrt{8\lambda\xi - 4 + \left(2 + \left(e^{jk_xX} + e^{jk_yX} - 2\right)\lambda^2 - 2\lambda\xi\right)^2}\Big)\Bigg| \leq 1.\end{aligned} \tag{6.63}$$

Evaluated at the stability limit $\lambda = \sqrt{\frac{1}{2}}$, this expression reaches its extreme value at $e^{jk_xX} = e^{jk_yX} = -1$:

$$|z| = \left| \frac{1}{2}\left( -\sqrt{2}\,\xi - \sqrt{2\xi^2 + 4\sqrt{2}\,\xi - 4}\right)\right| \leq 1 \tag{6.64}$$

This expression results in the stability requirement: $\xi \leq \sqrt{\frac{1}{2}} \approx 0.707$. This is not ideal, as this limits the admittance value and indicates that the system will be unstable for admittance values higher than $\sqrt{\frac{1}{2}}$.

For the 3-dimensional case, the analysis of the corner node is much the same we we will omit the steps that are similar to the ones above. We end up with the expression:

$$|z| = \left| \frac{1}{2}\left( -\sqrt{3}\,\xi - \sqrt{3\xi^2 + 4\sqrt{3}\,\xi - 4}\right)\right| \leq 1 \tag{6.65}$$

which yields the stability conditions: $\xi \leq \sqrt{\frac{1}{3}} \approx 0.577$. Note that this corresponds to a reflection value $R \approx 0.268$, which is equivalent to an absorption value of $\alpha \approx 0.928$ and higher.

### 6.3.4 Von Neumann stability analysis $4^{\text{th}}$-order stencils

The stability analysis of the $4^{\text{th}}$-order scheme follows a similar pattern as set out in the previous section. The Z-transformed update equation of the 1-dimensional boundary node is:

$$z - 2 + z^{-1} = \lambda^2\left(-\frac{8}{9} + \frac{59}{48}e^{jkX} - \frac{1}{12}e^{2jkX} - \frac{1}{48}e^{3jkX}\right) + \lambda\,\xi\left(1 - z^{-1}\right). \tag{6.66}$$

We will omit the full solution of this equation in terms of $z$ as it gets extremely lengthy and is trivial to derive using the quadratic formula. Evaluating at the stability limit $\lambda = \sqrt{\frac{3}{4}}$ and evaluating at the extreme value (when $kX = \pi$), we are left with the following expression in terms of $\xi$:

$$|z| = \left| \frac{1}{192}\left( 35 - 48\sqrt{3}\,\xi - \sqrt{6\,912\,\xi^2 + 15\,072\,\sqrt{3}\,\xi - 35\,639}\right)\right| \leq 1. \tag{6.67}$$

This is satisfied for all: $\xi \leq \frac{227}{96\sqrt{3}} \approx 1.365$, which includes the entire range of absorption from 0 to 1.

We will not go over the full analysis of the 2D and 3D schemes but the results follow here. In the 2-dimensional case, evaluated at $\lambda = \sqrt{\frac{3}{8}}$, the solution for $z$ is:

$$|z| = \left| \frac{1}{192} \left( 35 - 48\sqrt{6}\,\xi - \sqrt{13\,824\,\xi^2 + 15\,072\sqrt{6}\,\xi - 35\,639} \right) \right| \leq 1, \tag{6.68}$$

which is true for: $\xi \leq \frac{227}{96\sqrt{6}} \approx 0.965$. This covers nearly the entire absorption range, but for extremely high absorption values stability cannot be guaranteed. In the 3-dimensional case, with $\lambda = \frac{1}{2}$, we get:

$$|z| = \left| \frac{1}{192} \left( 35 - 144\,\xi - \sqrt{20\,736\,\xi^2 + 45\,216\,\xi - 35\,639} \right) \right| \leq 1, \tag{6.69}$$

This gives the stability limit $\xi \leq \frac{227}{288} \approx 0.788$, corresponding to a reflection value $R \approx 0.118$ and an absorption of $\alpha \approx 0.986$.

In summary, for all dimensions, the $4^{\text{th}}$-order stencil has less restrictive stability conditions on $\xi$ than the $2^{\text{nd}}$-order case and is therefore favourable.

### 6.3.5 Stability of multi-dimensional $4^{\text{th}}$-order stencils

For stable time-independent stencils stencils with symmetric negative definite $Q$, it is known that extensions to higher dimensions retain their stability [52], assuming operations on different axes are linearly independent. Von Neumann analysis in the previous section has posed stability requirements on $\zeta$ to avoid exponentially growing solutions. In this section, the 2D and 3D absorbing boundaries are analysed in order to verify these.

We will go by the common concept that spatial stencils for each dimension can be superimposed to obtain a higher-dimensional stencil. (The matrix representation of this is omitted as it becomes too unwieldy to write down explicitly.) We construct a domain of $11 \times 13$ nodes and apply the same boundary conditions to all boundaries, and evaluate the system at the Courant stability limit ($\lambda = \sqrt{\frac{3}{8}}$). Note that at the 4 corners of the domain, this doubles the absorption term proportional to $\xi$. Fig. 6.6 shows the result of this experiment.

First of all, it can be observed that the 2D system is now also stable for $0 \leq \xi \leq 1$ (compared to the theoretical limit 0.965), and as expected becomes more absorbing for higher values of $\xi$. For high absorption values, there is a clear difference between the axial and tangential

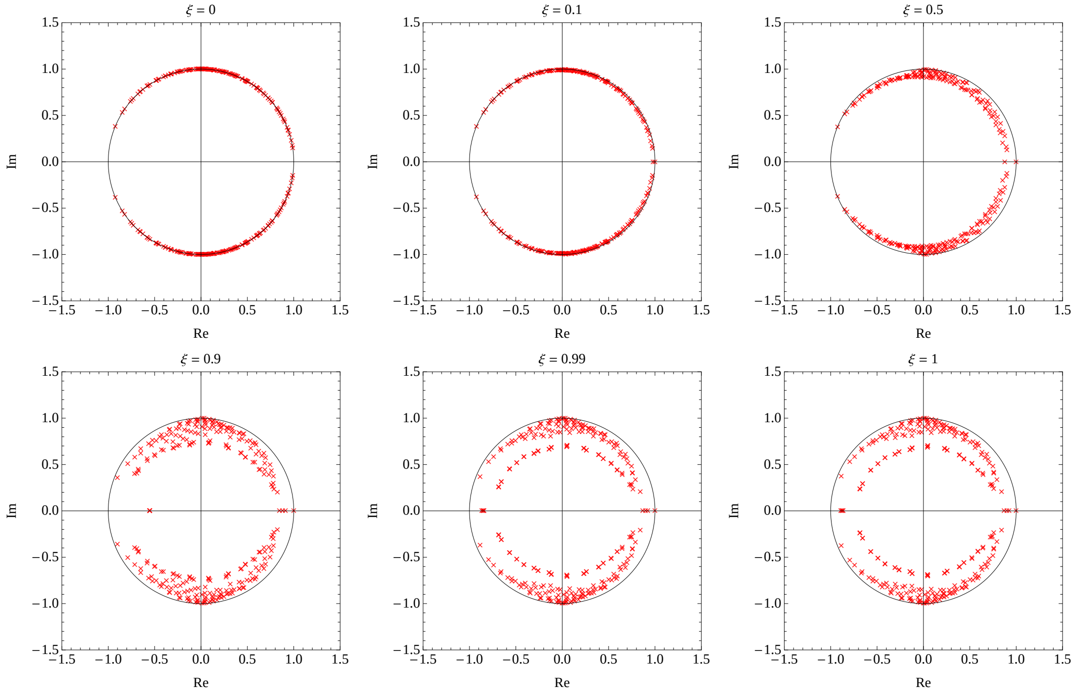


**Figure 6.6:** Polar plot of eigenvalues of 4th-order accurate update matrix for a 2D system. The system is stable for all valid $\xi$.

modes, visible as two separate circles of different radius. This is another indication that the system behaves as expected. As in the 1D case however, full absorption is not reached, so this method only serves as an approximation to the absorbing boundary condition. As with all absorbing boundaries seen before, high frequencies are absorbed less than low frequencies. A more accurate modal analysis is necessary to study the potential effects of the absorbing boundary on dispersion, as this is not clearly visible from the crowded graph of eigenvalues.

Next, we examine the properties of the 3D absorbing boundaries. The same method as for the 2D example is employed: the same stencil is superimposed for each axis, and the system is evaluated at the Courant stability limit ($\lambda = \frac{1}{2}$). An $11 \times 12 \times 13$ domain is defined and all boundaries have the absorption factor $\xi$. The results are demonstrated in Fig. 6.7.

The 3D system also absorbs energy as expected, in the same way in the previously described systems, and is stable within the theoretical stability range. Higher values of $\xi$ result in higher absorption, and oblique and tangential modes manifest themselves as smaller circles in the polar plot. However, when $\xi$ exceeds $\sim 0.86$, poles are being pushed outside of the unit circle on the negative real axis, and the system becomes unstable.

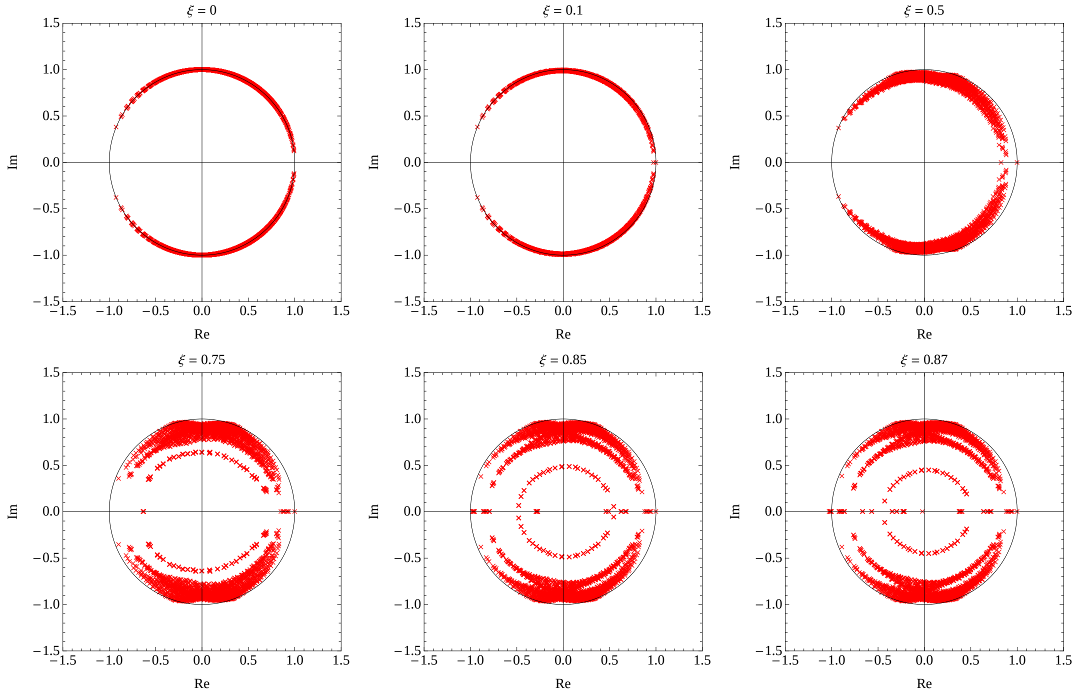


**Figure 6.7:** Polar plot of eigenvalues of $4^{\text{th}}$-order accurate update matrix for a 3D system. The system is stable for all valid $0 \leq \xi \leq 0.85$. For higher $\xi$, exponentially growing oscillations make the system unstable.

### 6.3.6 Note on $4^{\text{th}}$ order stability

As shown in the previous section, the theoretical stability limit on $\xi$ in 1D, 2D, and 3D is, respectively: $\xi \leq \frac{227}{96\sqrt{3}} \approx 1.365$, $\xi \leq \frac{227}{96\sqrt{6}} \approx 0.965$, $\xi \leq \frac{227}{288} \approx 0.788$. A stability dependency on $\xi$ is of course undesirable, but the author has not been able to find a way to mitigate this. From experiments it seems that the practical stability limit for the latter two seems to be even higher. For the 2D case, no instability is reported in the entire range of $0 \leq \xi \leq 1$. In the 3D case, only for $\xi > 0.855$ does the system seem to become unstable. We cannot be sure whether this is a more general property of these boundaries or if this is caused by the particular examples employed. Rather erring on the side of caution, the author recommends the theoretical (and more restrictive) stability requirements to be used. It should also be noted that the theoretical stability limit of the 3D system corresponds to an absorption value of 0.986, which in practice is an extremely high value, and few real-world architectural materials are likely to exceed this range.

## 6.4 Test set-up

We have two new formulations of absorbing boundaries: one for the $2^{\text{nd}}$ and one for the $4^{\text{th}}$-order boundary. Firstly, it should be verified that they absorb the same amount of energy with respect to one another. As they differ only in their order of approximation, they should absorb nearly the same amount of energy. Secondly, they are to be compared with the current common approach to absorbing boundaries, represented by Kowalczyk and Van Walstijn's locally reacting surfaces (LRS) [104, 103]. For higher-dimensional modelling, we will go by the common concept that spatial stencils for each dimension can be superimposed to obtain a higher-dimensional stencil. (The matrix representation of this is omitted as it becomes too unwieldy to write down explicitly.)

The absorptive properties were determined in 1D, 2D, and 3D simulations, using a large range of $\xi$ values to test the validity of the approach. The sizes of the domain for the $2^{\text{nd}}$-order simulations were, respectively, 40 nodes, $21 \times 28$ nodes, and $23 \times 24 \times 25$ nodes. The corresponding sizes for the $4^{\text{th}}$-order simulations were: 35 nodes, $18 \times 24$ nodes, and $20 \times 21 \times 22$ nodes. The systems were excited with a 7-point Hann window and an impulse response was obtained by running the simulation until well past the 60 dB level decay. All impulse responses were low-pass filtered with a cut-off frequency of a third of the Nyquist frequency, and on all simulations that displayed a drift, standard DC-offset filtering was performed to remove this. The time-integrated squared impulse response was then obtained, resulting in a all-frequency energy decay graph.

Oxnard [152] notes that the LRS is ambiguous and that there are two variants implemented in literature. In higher dimensions, an update scheme for the edge and corner nodes, which require absorption over more than one axis, can be derived in two different ways. The classical version by Kowalczyk and Van Walstijn (henceforth referred to as LRS1) superimposes stencils along every direction, and thus computes the gradient term for each axis separately. A different implementation was employed by Webb and Bilbao [218], who compute the gradient term once and use this across all axis. (This will be referred to as LRS2.) These implementations are not equivalent, and to the best of the author's knowledge, it is an unsolved problem as to which should be employed. Therefore, in the following section, both are used for comparison against the newly proposed SAT-SPB-inspired boundaries.

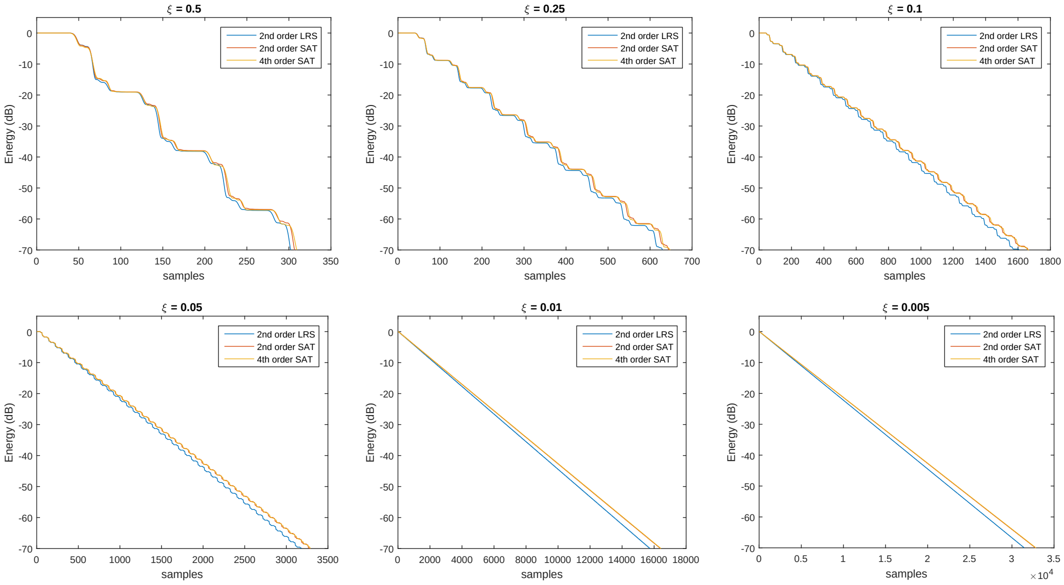


**Figure 6.8:** The decay graphs of a 1D simulation with various absorption values $\xi$. The decay graphs match extremely well and indicate identical absorptive properties for the 2<sup>nd</sup> and 4<sup>th</sup> order boundary described in this chapter.

## 6.5 Results

The results of the 1D simulations' decay graphs for various values of $\xi$ are displayed in Fig. 6.8. The decay of the $2^{\text{nd}}$- and fourth-order stencil match extremely well for the entire decay of 60 dB. The main differences are visible in the graphs that display a high absorption ($\xi = 0.5, \xi = 0.25$), but the differences are slight and can attributed to dispersion effects present in the $4^{\text{th}}$ order scheme. Both show a small but consistent difference compared to the LRS boundary type: the LRS boundary absorbs show a small but consistent $\sim 5\%$ lower reverberation time.

Fig. 6.9 shows the decay graphs of the impulse responses from 2D simulations. The modelled domain was a $21 \times 28$ rectangular domain in the $2^{\text{nd}}$-order case, equivalent to a $18 \times 24$ domain in the $4^{\text{th}}$-order case. As in the 1-dimensional case, the decay times of the $2^{\text{nd}}$ and $4^{\text{th}}$ order scheme line up extremely well, giving every indication that the taken approach is correct and that the absorptive properties are identical. As in the 1D case, compared to the LRS schemes these boundaries are slightly less. As mentioned in the previous section, two different variations of the LRS boundary have been used for comparison. The LRS1 is the most absorptive and has a $T_{60}$ that is on average $\sim 6.5\%$ lower compared to the new SAT-SBP ones. The LRS2 absorbs less energy in the corners, resulting in a reverberation time $\sim 5.3\%$ lower in the current test set-up. Note that the difference in absorption between LRS1 and

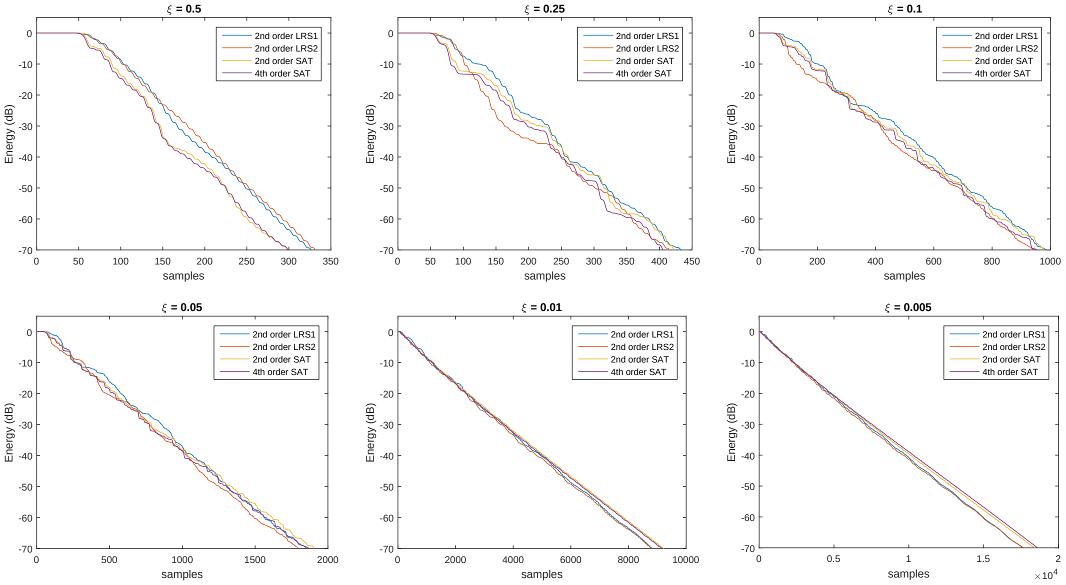


**Figure 6.9:** The decay graphs of a 2D simulation with various absorption values $\xi$.

LRS2 is only in the corner nodes, so the domain size will influence the percentage-difference of absorption.

The results of the 3-dimensional simulations in Fig. 6.10 are again similar to the lower-dimensional ones. Due to the 6 boundaries, the energy absorption is more rapid and thus the decay graph shorter, so that the slight discrepancies between the $2^{\text{nd}}$ and $4^{\text{th}}$ order up to -60 dB decay are more visible for all $\zeta \geq 0.05$. The differences can again be attributed to the different order of accuracy in modelling, though, and not to different absorptive properties. The decay graphs that display the decay of the lower impedance models show that the absorptive properties of the $2^{\text{nd}}$- and $4^{\text{th}}$-order scheme are identical for all practical purposes. The LRS1 and LRS2 follow the same pattern as in the 2-dimensional case: The LRS1 absorbs most: the average reverberation time is $\sim 9.6\%$ lower than in the SAT boundaries. The LRS2 lies in between, its reverberation time being $\sim 4.9\%$ lower.

## 6.6 Conclusion

A new method to obtain impedance boundary stencils combined with the homogenous wave-equation has been demonstrated. Their construction is based on the SBP-SAT methods, commonly known in the fields of e.g. fluid dynamics and geophysics. Mattsson and Nordström's work [129] formulated the impedance boundary condition and a similar implementation, but had not implemented them nor proven their stability in a fully discrete system. It was shown

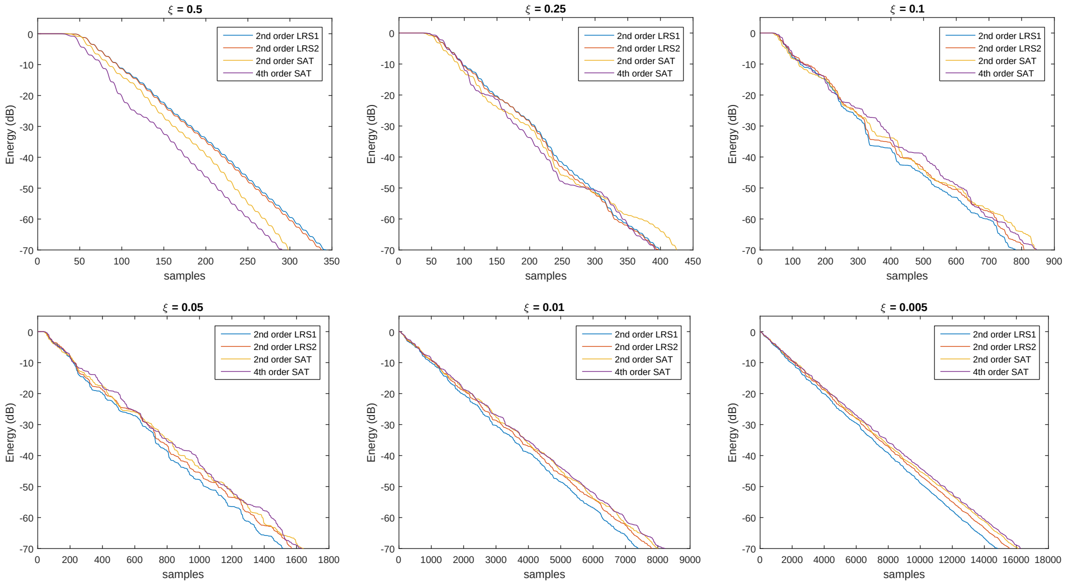


**Figure 6.10:** The decay graphs of a 3D simulation with various absorption values $\xi$.

that a variation of their proposal results in a conditionally stable boundary stencil, with a stability condition on $\lambda$ and $\xi$.

Stable impedance boundary stencils have been formulated for the second- and 4$^{\text{th}}$-order case. The method can likely be extended to 6$^{\text{th}}$ and 8$^{\text{th}}$ order, but this is matter of further investigation. The stability condition on $\xi$ implies that high absorption values (for the 4$^{\text{th}}$-order 3D scheme: $\xi > 0.788$, i.e. $\alpha > 0.986$) poses a threat of instability. This is unlikely to be a big problem for the practical purposes of acoustic modelling, as there are extremely few physical materials that would actually possess these properties. However, it should be noted that this poses a problem when trying to build in fully absorbing boundaries, such as open windows or doors.

The absorptive properties of the 2$^{\text{nd}}$- and 4$^{\text{th}}$-order accurate stencil have been investigated, and all experiments indicate that they show practically the same amount of absorption in 2D and 3D. Compared to the currently common LRS boundary, they are slightly less absorptive, depending on the LRS implementation, the latter result in reverberation times between 9.6% and 4.9% lower than the newly proposed stencils.

With the newly-proposed boundaries, a full higher-order acoustic model that includes absorbing boundaries becomes a possibility. Former implementations had to rely on lower-order approximations at the boundary, thus dragging down the accuracy of the system as a whole. These boundaries only apply to the subclass of ‘large-star’ or ‘leggy’ stencils. Future work may focus on deriving absorbing boundaries that are more generally applicable to higher-order stencils. Moreover, verification as to the correct amount of boundary absorption is necessary, as the difference between the LRS boundary can merely be observed but not attested to reality.

# Chapter 7

# Two case studies

Contents

In the previous chapters, the theoretical foundation for higher-order room acoustic modelling has been laid. In this chapter, two more realistic case studies will be presented. Firstly, the case study of a large shoebox room, much analogous to the study conducted by Southern *et al.* [191], will be presented to further assess the absorptive properties and frequency response of the higher-order boundaries. Furthermore, the methods will be tested against reality: IRs obtained from simulations using several different computational methods are compared to measured IRs. The selected building for modelling and comparison was the National Centre for Early Music (NCEM) in York [139], known for its adjustable acoustics by means of moveable panels and drapes. The space was modelled in the CAD programme Blender [24] to an appropriate level of detail, and the different materials were applied to the model. The results of both the measured IRs, a traditional $2^{\text{nd}}$-order accurate simulation, and the newly introduced $4^{\text{th}}$-order accurate simulation are compared for reverberation time and EDT. A hybrid acoustic modelling method was used so that a full bandwidth IR could be obtained, suitable for auralisation.

## 7.1 Shoebox experiment

Southern *et al.* [191] performed an extensive investigation into the comparison of the traditional FDTD algorithm, the ISM, beam tracing, and the Acoustic Radiance Transfer method. Among other things, this research contains an experiment that compares frequency response and absorptive properties of the aforementioned algorithms. They proved that the ISM and FDTD produce very similar reverberation times on shoebox rooms, well within the just-noticeable difference (JND) from one another. As these algorithms are well-established methods known to produce accurate IRs of shoebox spaces, they therefore provide a good benchmark to measure the proposed $4^{\text{th}}$-order scheme and boundaries against.

### 7.1.1 Methods

In order to stay as close as possible to the original experiment, a shoebox of the same dimensions (5.56m × 3.97m × 2.81m) was chosen, with also sources and receivers at the same locations (see Table 7.1). The sample rate was 18 126 Hz (using a speed of sound $c = 343.26$ m/s). In the $2^{\text{nd}}$-order cases, this resulted in a domain of $169 \times 121 \times 86$ nodes. In the $4^{\text{th}}$-order case, this was $146 \times 105 \times 74$. The experiment was run with all walls having the same positive real reflection value. The following reflection values were tested: $\{0.1, ..., 0.9, 0.91, ....0.99\}$. The excitation was implemented as a transparent source, in all cases a 7-point Hann window. For the second order cases, we again have 3 different boundary versions, the classic LRS scheme, the modified LRS boundaries (see Chapter 6) referred to as LRS2, and the SAT-SBP derived one (see previous chapter). Note that because of the stability dependency on $\lambda$ and $\xi$, the lowest reflection values cause an unstable system for the $2^{\text{nd}}$ order SAT-SBP simulation. All IRs underwent the exact same post-processing step: DC-offset filtering and subsequent low-pass filtering with a cut-off at 1/6 of the sample rate.

**Table 7.1:** The source and receiver positions of the shoebox experiments—values are in meters. After [191].

| Source | x | y | z | Receiver | x | y | z |
|---|---|---|---|---|---|---|---|
| S1 | 2.71 | 1.89 | 1.69 | R1 | 3.81 | 1.13 | 1.92 |
| S2 | 4.10 | 1.39 | 1.82 | R2 | 1.95 | 0.60 | 2.28 |
| S3 | 2.09 | 2.12 | 2.12 | R3 | 2.09 | 3.08 | 0.96 |
| S4 | 3.01 | 2.68 | 0.93 | R4 | 3.91 | 1.89 | 1.69 |
| S5 | 4.80 | 2.18 | 2.12 | R5 | 2.09 | 0.99 | 1.62 |

### 7.1.2 Results

Three main characteristics of the IRs have been analysed: Time-domain comparison and direct sound arrival time, frequency response, and $T_{45}$ reverberation time. The simulation time has been analysed to give an indication of IR quality versus performance speed. All output IRs have been appended to this thesis' accompanying material, including also the Matlab code to generate it, and all of the decay and frequency comparison graphs, of which an excerpt is presented further in this section.

Firstly, we consider time-domain analysis of time of arrival of the direct sound and early reflections to ensure that the models have the same basic propagation and reflection properties. Fig. 7.1 show the incoming direct sound and early reflection of 6 source/receiver pairs and several different reflection values. Note that nothing is special about the choice of the graphs presented here, they have simply been chosen to represent a wide range of cases. Note also that the graph of $R = 0.2$ doesn't show the 2$^{\mathrm{nd}}$-order SAT-SBP method, as this is unstable at this reflection amount.

The results match nearly perfectly in terms of time of arrival. In all scenarios, the 4$^{\mathrm{th}}$-order case show more pronounced peaks in the direct sound as a result of the more favourable dispersion pattern. Expectedly, the differences between LRS1 and LRS2 are so small that they can hardly be observed in time-domain plots of this short duration. Up to and including the direct sound wave, the LRS1/LRS2 and the 2$^{\mathrm{nd}}$-order SAT-SBP case are identical - as their air propagation model functions are the same. Slight differences can be observed as soon as reflections arrive at the receiver, which can be attributed to different boundary interaction behaviour. From this time-domain comparison we can conclude that the different models have roughly the same behaviour. However, it does not give any information on which one is more accurate or correct.

Secondly, we present a frequency response of the simulated IRs. Several source-receiver pairs

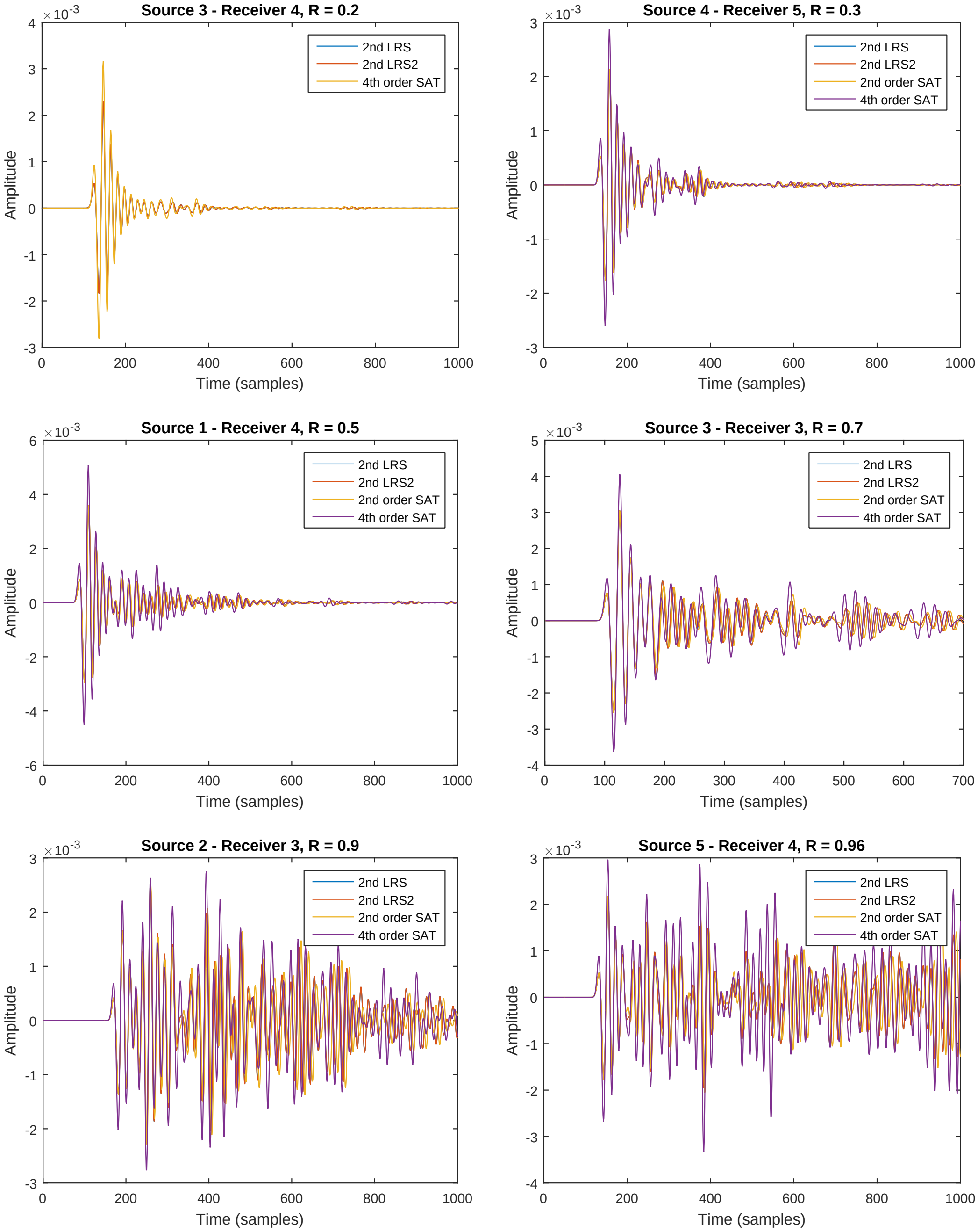


**Figure 7.1:** Time-domain comparisons between the 4 different simulation types, for several different source/receiver pairs and reflection values as indicated in the graph titles.

at different reflection values have been plotted in Fig. 7.2. Again, the choice of source-receiver pairs was arbitrary and selected to show a wide range of results. In the cases of reflection values higher than and equal to 0.95, room modes are clearly distinguishable. In these plots, grey vertical lines indicate the theoretical location of the room modes. For lower reflection values, too much energy is absorbed for room modes to be clearly distinguished.

It is well known that the classic $2^{\text{nd}}$-order FDTD algorithm produces clear room modal patterns in a shoebox room as theory prescribes. Indeed in the plots, especially those with high reflection coefficient, this is clearly the case. It can also be observed that both the $2^{\text{nd}}$- and $4^{\text{th}}$-order SAT-SBP graphs line up extremely well with the traditional LRS boundary ones. There are only small differences, which can be attributed to slight differences in domain size due to round-off error, differences in dispersion pattern, and minor differences introduced by the different boundary types. However, these plots give another clear indication that the newly introduced boundary types portray a correct behaviour in line with the currently prevalent algorithms.

As a final verification of the new SAT-SBP boundaries, the decay curves and $T_{45}$ times are analysed. Southern *et al.* [191] showed that the classical FDTD implementation and the Image Source method, which solves the wave equation exactly in a shoebox room, produce very similar $T_{30}$ results well within the commonly accepted JND [29, 223]. The LRS1 results will therefore be used as a benchmark. Although Southern *et al.* used the $T_{30}$ metric, we found that the $T_{45}$ metric gave cleaner results, while being equally valid as the simulations do not have an environmental noise floor. (The only noise floor the simulations have is introduced by numerical precision, which is far below the region of interest and can therefore easily be ignored.)

Fig. 7.3 displays the decay graphs of 12 source-receiver pairs at different reflection coefficients. It has been verified that these graphs are representative for the full set of results, and that also the pairs omitted here for conciseness' sake show very similar decay patterns. On qualitative assessment, it appears that the decay curves are extremely similar. The LRS1 and LRS2 are, expectedly, most alike, with the LRS2 boundary absorbing slightly less energy in all cases. The $2^{\text{nd}}$-order SAT-SBP method yields largest decay times, and the $4^{\text{th}}$-order one is similar but tends to decay ever so slightly faster.

In order to get the most representative quantitative $T_{45}$ comparisons, the times of all 25 source-receiver pairs were averaged for each case and each reflection value. Fig. 7.4 shows the results of this experiment, as the percentage reverberation time difference compared to

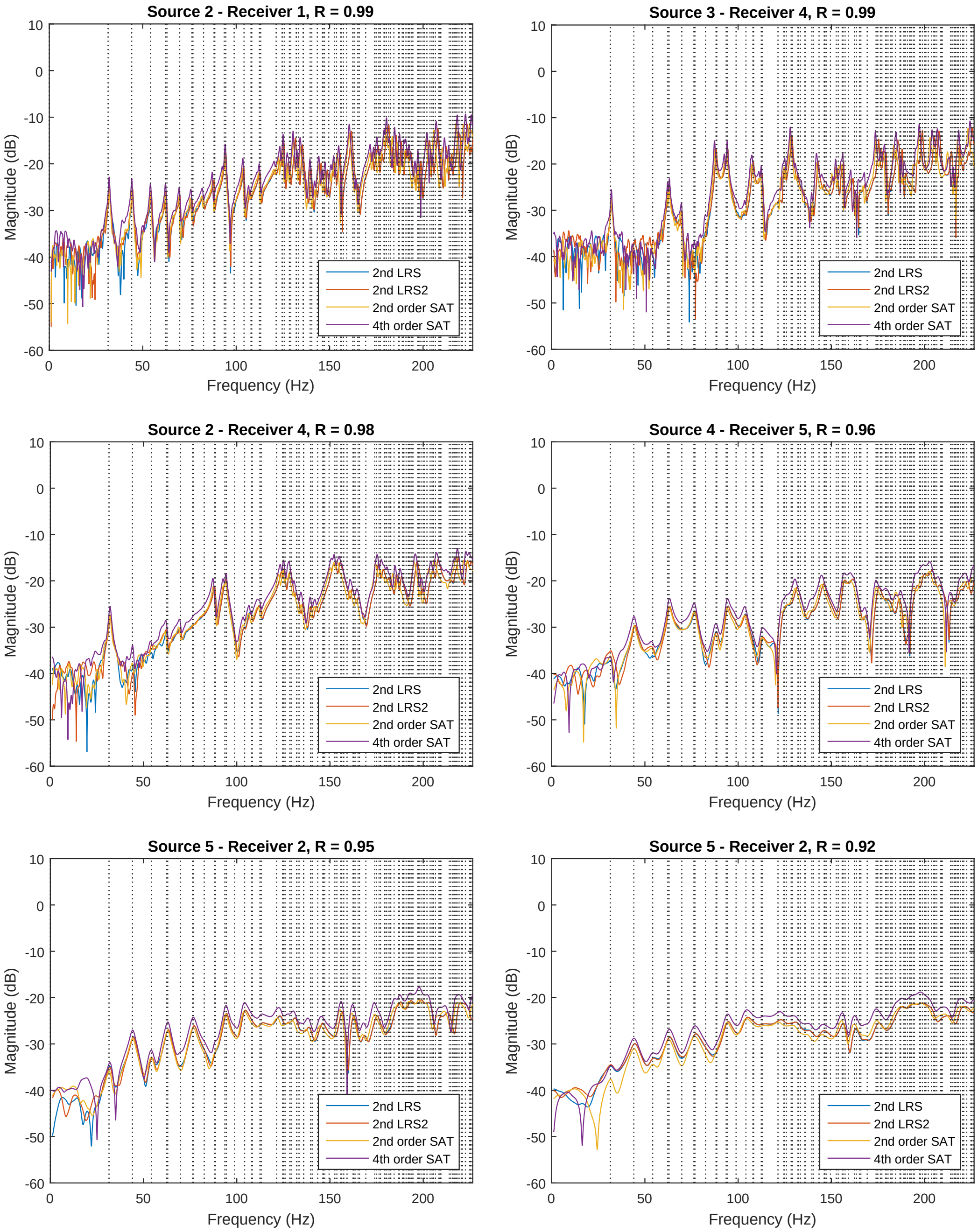


**Figure 7.2:** (1/2) Frequency domain comparisons between the 4 different simulation types, for several different source/receiver pairs and reflection values as indicated in the graph titles. Theoretical room mode values have been plotted as vertical grey lines in those graphs in which room modes can be distinguished.

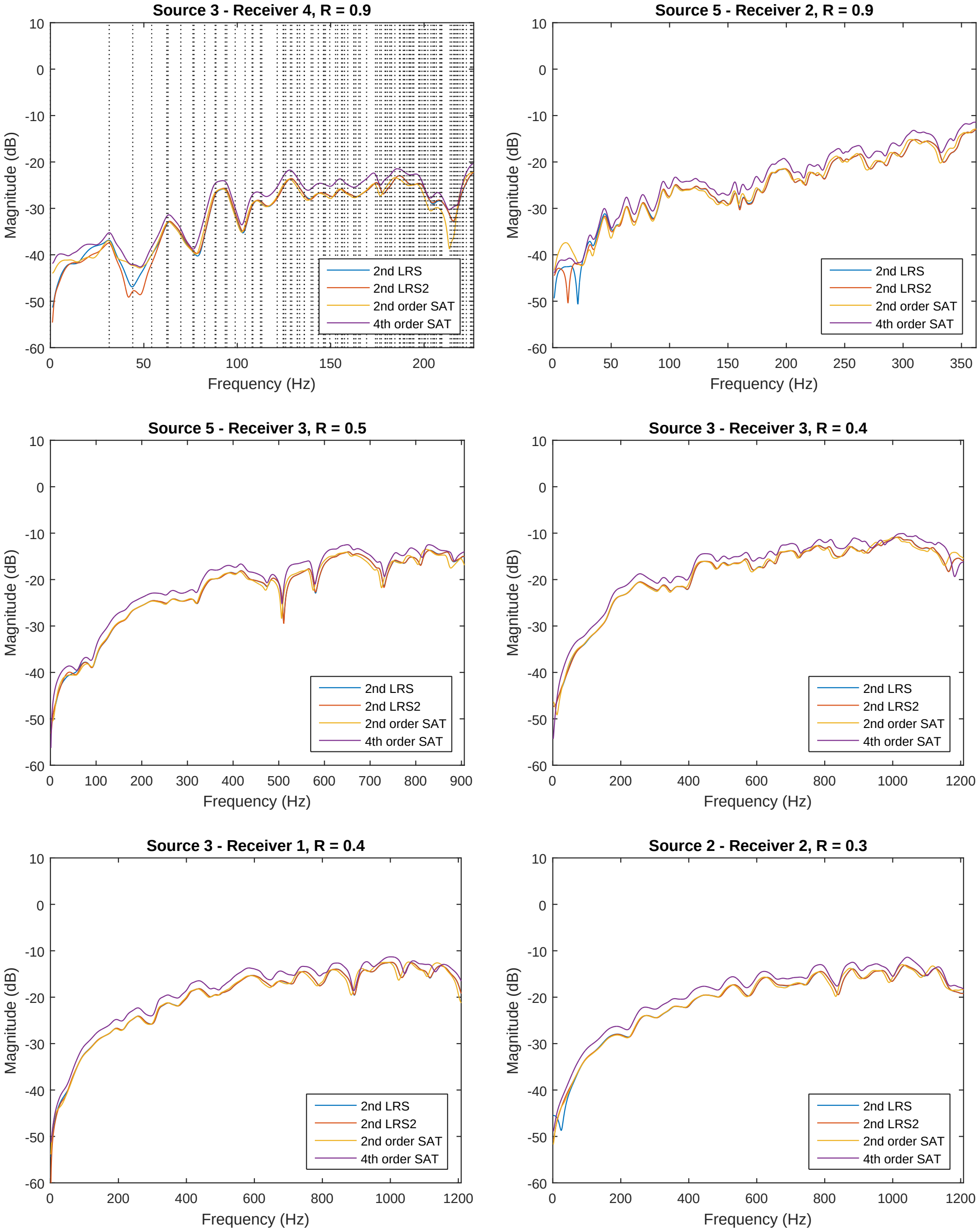
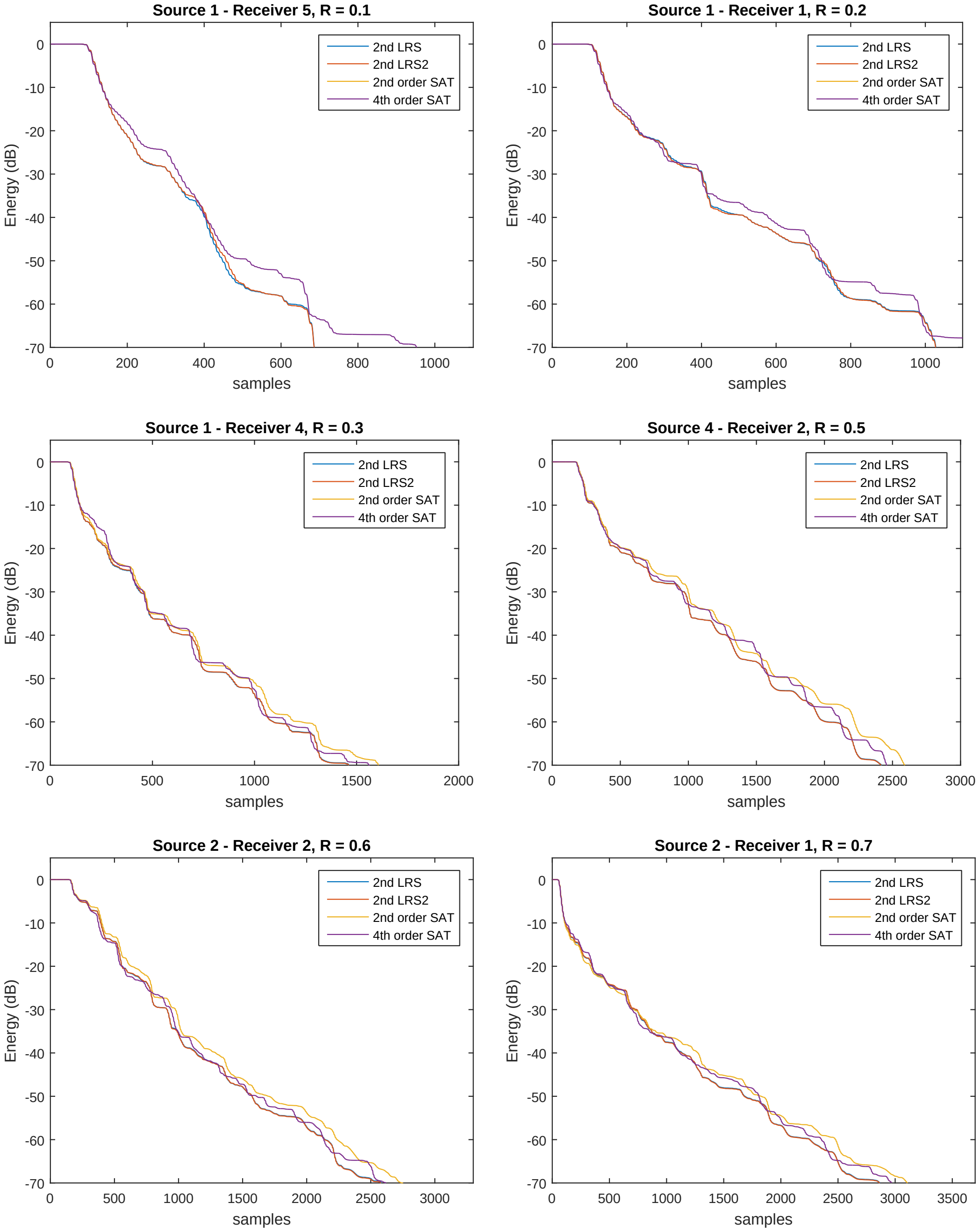


**Figure 7.2:** (2/2) Frequency domain comparisons between the 4 different simulation types, for several different source/receiver pairs and reflection values as indicated in the graph titles. Theoretical room mode values have been plotted as vertical grey lines in those graphs in which room modes can be distinguished.



Source 1 - Receiver 5, R = 0.1

Source 1 - Receiver 1, R = 0.2

Source 1 - Receiver 4, R = 0.3

Source 4 - Receiver 2, R = 0.5

Source 2 - Receiver 2, R = 0.6

Source 2 - Receiver 1, R = 0.7

**Figure 7.3:** (1/2) Energy decay comparisons between the 4 different simulation types, for several different source/receiver pairs and reflection values as indicated in the graph titles.

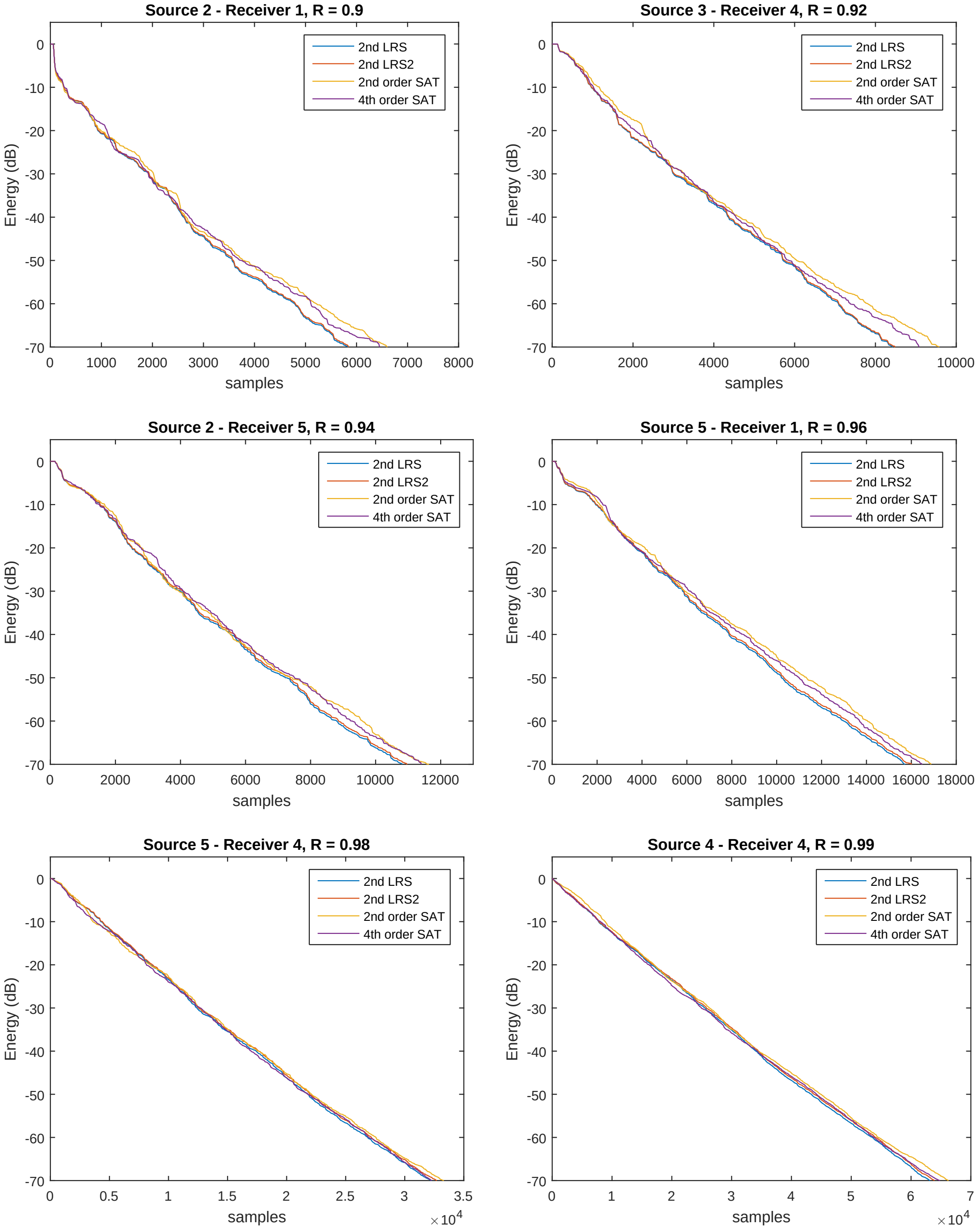


**Figure 7.3:** (2/2) Energy decay comparisons between the 4 different simulation types, for several different source/receiver pairs and reflection values as indicated in the graph titles.

the classic LRS boundary. First off, it can be noted that the LRS2 simulations show only very small deviations from the LRS1 boundary, more prominently so at high reflection values. The highest differences are reported in the range of $0.93 \leq R \leq 0.99$, where the $T_{45}$ values are up to 1.65%. The $2^{\text{nd}}$-order SAT-SBP type boundary shows results that deviate the most from the classic LRS boundary. It is consistently around 10% too high, with peak at a deviation around 14.5% for $0.7 \leq R \leq 0.8$. For reflection values higher than 0.9, however, the difference rapidly decreases. Note that there are no values for $R < 0.3$ because of instability. The $4^{\text{th}}$-order scheme portrays a smaller and more stable difference. It shows a 6–7% higher reverberation time, and also this difference decreases again for $R > 0.9$.

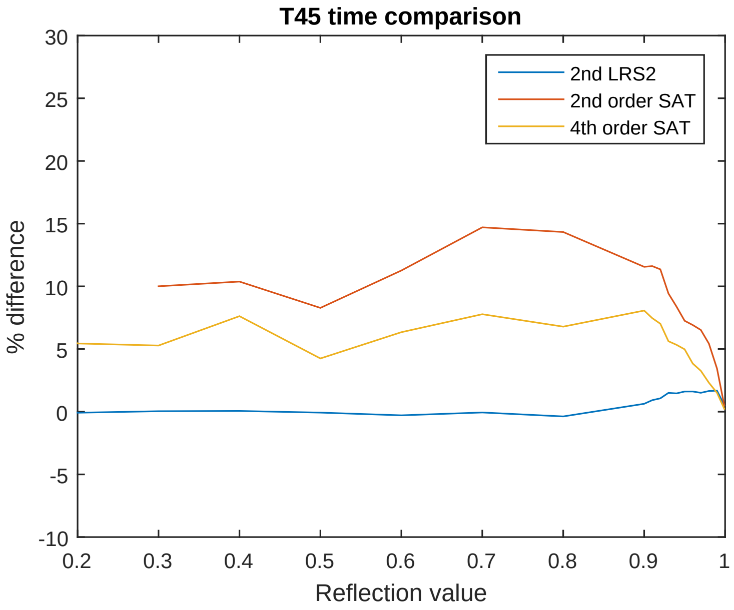


**Figure 7.4:** $T_{45}$ time comparisons of the LRS2, and $2^{\text{nd}}$- and $4^{\text{th}}$-order SAT-SBP boundaries, compared to the classic LRS boundary.

The above experiments were performed using Matlab. As implementation of the FDTD algorithm, the sparse matrix multiplication method was chosen, as this is the fairest method to compare the difference in performance because of the number of multiplications and domain size, while the underlying implementation detail remains exactly the same. The computation time of 25 simulations was averaged to obtain reliable performance data, which is displayed in Table 7.2. The values include only simulation time, not matrix construction. The LRS1 and LRS2 implementations take virtually the same amount of time, which is expected, as they are identical in everything except the value of their coefficients. The $2^{\text{nd}}$-order SAT-SBP method is on average 6.4% faster. We are unable to explain this small performance increase. The

4th-order SAT-SBP method is 13.1% faster. This is explained by the fact that the domain is smaller due to higher grid spacing, which is a more important factor than the slightly larger update equation.

**Table 7.2:** The performance comparison of the sparse-matrix implementation for 20.000 iterations for the 4 different boundary implementations, averaged over 25 simulations.

| Boundary type | Computation time (s) |
|---|---|
| LRS1 | 4478.52 |
| LRS2 | 4473.84 |
| 2nd-order SAT-SBP | 4188.81 |
| 4th-order SAT-SBP | 3889.03 |

### 7.1.3 Conclusion

In this section a more realistically sized shoebox shaped room has been analysed using 4 different types of boundaries: the classical LRS boundary implementation (LRS1), a modified LRS boundary implementation with different updates for edge and corner nodes (LRS2), and the newly presented SAT-SBP 2nd- and 4th-order algorithm. Time and frequency domain analysis show that the wave and modal characteristics are preserved by all algorithms. Decay time analysis show that there is a consistent small but significant difference between the SAT-SPB boundaries and the LRS ones. The 4th-order boundary is shown to be the most reliable, with a 5-7% difference versus the LRS1 benchmark. The 2nd-order SAT-SBP boundary shows larger and less stable deviations, up to 16.5%. From a performance point of view, the 4th-order scheme is the most favourable, as it yields the fastest results (13.1% performance increase) because of the higher allowed grid spacing. On top of this, it yields a more favourable dispersion pattern, as outlined in Chapter 5. The higher reverberation time compared to the current state-of-art may be deemed a problem. However, as the reverberation time differences are now known quantities, they could be compensated *a priori* in order to achieve a behaviour more like the LRS1, if this were deemed necessary.

## 7.2 NCEM Experiment

The second experiment concerns the case study of a real acoustic space, the IR measurements, its computer modelling, and its acoustic simulation. The goal was to compare the 2nd- and 4th- order accurate method against one another, but also to compare the simulation results against the measurements taken in the physical space.

### 7.2.1 Methods

The studied space was St. Margaret's Church in York, United Kingdom, currently known as the National Centre for Early Music. The church has been acoustically treated for concerts and conference use, with reversible acoustic panels and drapes arranged throughout the space to easily change the physical acoustic characteristics [139]. A detailed study of the acoustics of this space was presented by Foteinou [67]. For the purpose of this study, the acoustic configuration, referred to as 'musical/opera performances', was used. For this configuration, drapes and 75% of the panels were in use (open). The remaining folded panels were the ones on the north wall. During the impulse response measurements in the actual space, the temperature was measured at a constant 21.5°C and the relative humidity at 44.5-45%. The space was largely empty, without any audience or seating. It contained several furnishings and instruments, such as a piano, several tables, and a harpsichord. The Schröder frequency of the space is approximately 40 Hz.

### 7.2.2 Recorded impulse responses

Impulse response measurements in the church were made using the Exponential Swept Sine (ESS) Method [59], with the Aurora plug-in [62] for analysis of the acoustical parameters. The frequency range of the sine sweep was from 22 Hz to 22 kHz, and lasted 15 seconds, using a Genelec S30D as the source transducer, and a Soundfield SPS422B as the receiver microphone. The source was placed as a performer would be in the space, facing towards the north wall, while the microphone was aligned toward the south wall for each location. Although during the measurements process 26 receiver positions were used for an appropriate acoustic coverage of the space, seven receiver positions are used in this research. These receivers are selected according to the nature of their position so as to give a representative sample of the room. Fig. 7.5 shows a floor plan of the space, with source (S) and receivers (R1–R7) marked.

### 7.2.3 Simulated impulse responses

The aim of this research was to obtain high-accuracy IRs of the above described space using a combination of several different algorithms. To this end, we endeavored to model the space

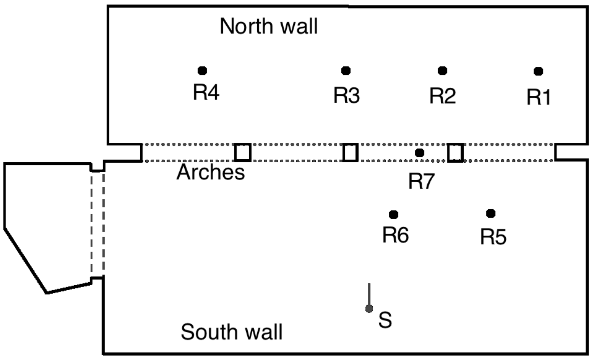


**Figure 7.5:** The source (S) and receiver (R1–R7) positions on the floor plan of the National Centre for Early Music.

as accurately as possible. The following sections describe the design stage and the algorithms employed for the room acoustic simulations.

### Blender design stage

The design stage was done using the open source 3D modelling software Blender 2.69 [24]. Using Blender's user interface allows for a quick creation of the geometry and material assignment of the scene. A plugin for this programme, first introduced in [207], was used to add acoustical data to the materials of each surface in octave bands from 62.5 Hz to 8 kHz. The plugin exports the geometry and material data to an intermediate Wavefront (.obj) geometry file, which is subsequently read by a ray tracer and FDTD solver.

Information regarding the acoustic characteristics of the surfaces, absorption and scattering coefficients, was gathered from existing libraries and literature of previous modeling work, such as [27, 127, 6, 16], and the most appropriate values were chosen for each surface in the space. Table 7.3 shows the materials used and their absorption coefficients across the eight frequency bands.

### FDTD simulation

The lower three octave bands (mid-frequencies: $\{62.5\text{ Hz}, 125\text{ Hz}, 250\text{ Hz}\}$) have been simulated in two different ways: using a $2^{\text{nd}}$-order and a $4^{\text{th}}$-order FDTD method, as described in previous sections. In the former case, the boundary implementation was that of Kowalczyk and Van Walstijn [105] (referred to as LRS1 in the previous sections). In the latter case, the newly described boundaries from Chapter 6 are used. The internal domain of the acoustic

**Table 7.3:** The absorption coefficients used in the NCEM model.

| Material | Absorption coefficients {62.5Hz, 125Hz, 250Hz, 500Hz, 1kHz, kHz, 4kHz, 8kHz} |
|---|---|
| Main wall | { 0.02, 0.02, 0.02, 0.03, 0.04, 0.05, 0.05, 0.05 } |
| Main floor | { 0.01, 0.01, 0.02, 0.03, 0.07, 0.09, 0.10, 0.10 } |
| Wood | { 0.10, 0.10, 0.07, 0.05, 0.04, 0.04, 0.10, 0.10 } |
| Stone | { 0.04, 0.05, 0.06, 0.06, 0.05, 0.05, 0.06, 0.05 } |
| Windows | { 0.10, 0.10, 0.07, 0.05, 0.05, 0.02, 0.02, 0.02 } |
| Plastic | { 0.10, 0.10, 0.25, 0.45, 0.58, 0.65, 0.70, 0.70 } |
| Reflectors | { 0.001, 0.15, 0.05, 0.05, 0.04, 0.05, 0.14, 0.14 } |
| Marble | { 0.001, 0.01, 0.01, 0.01, 0.01, 0.02, 0.02, 0.02 } |
| Fabric | { 0.03, 0.03, 0.04, 0.11, 0.17, 0.24, 0.35, 0.35 } |
| Drapes | { 0.14, 0.14, 0.35, 0.55, 0.72, 0.70, 0.65, 0.65 } |

space was voxelised into a rectilinear grid and divided into:

1. material nodes: nodes that represent a solid material and do not have a pressure value associated with them, and hence do not need to be updated,
2. boundary nodes, which have a special update equation based on their orientation and distance from the boundary (in the 4th-order case, there are several different types), and
3. air nodes, which border only boundary and/or other air nodes.

In the 2nd-order accurate simulation, the Courant number $\lambda = \frac{cT}{X}$ was chosen at its stability limit $\frac{1}{\sqrt{3}}$, with a spatial step of $X = 7.5$ cm, such that the full sample rate $f_s = \frac{1}{T}$ of the model was 7927 Hz. This was the same for each frequency band, as to ensure the highest accuracy also in the lowest frequency bands. By means of fair comparison, it was decided to model the 4th-order simulation with the same grid spacing, and a resulting sample rate of 9154 Hz. The rendering statistics are outlined in more detail in Table 7.4.

As the grid is rectilinear, incidence at a grazing angle doesn't need to be considered. The conversion from absorption value $\alpha$ to impedance value $\zeta$ (or alternatively: admittance $\xi = \frac{1}{\zeta}$), is as follows:

$$\frac{1+\sqrt{1-\alpha}}{1-\sqrt{1-\alpha}}, \tag{7.1}$$

for a phase-preserving boundary [99].

**Table 7.4:** The render statistics of the two FDTD simulations of the NCEM.

| FDTD render statistics | $2^{\text{nd}}$-order | $4^{\text{th}}$-order |
|---|---|---|
| Dimensions | $182 \times 150 \times 330$ | $182 \times 150 \times 330$ |
| Nodes | 9 009 000 | 9 009 000 |
| Grid spacing | 7.5 cm | 7.5 cm |
| Sample rate | 7927 Hz | 9154 Hz |
| Update rate | 2.830 fps | 2.708 fps |

### Ray tracer

The ray tracing algorithm finds its roots in the field of graphics [10, 220], and has since made its way to the field of acoustics. It inherently assumes ray-like behaviour, which is only a valid approximation for high frequencies. At low frequencies, diffraction and standing wave effects are more prominent, such that a ray tracer does not produce reliable results. Therefore, only the upper 5 octave bands $\{500\text{ Hz}, 1\text{ kHz}, 2\text{ kHz}, 4\text{ kHz}, 8\text{ kHz}\}$) were modelled by our ray tracer.

The ray tracing method used is forward ray tracing, i.e. tracing rays from the source to the listeners. Though backward ray tracing has advantages particularly in real-time acoustic ray tracing [171], forward ray tracing is more optimal in this case, as it can exploit the fact that there is one sound source and multiple listeners.

The omnidirectional source was modelled by casting $10^7$ rays in pseudo-random directions by sampling a unit sphere. The high number of rays provides for a representative sample of the unit sphere. Every material in the scene has a Bidirectional Reflectance Distribution Function (BRDF) $f(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o)$ associated with it [142]. For a given incoming and outgoing angle $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_o$ with respect to the surface normal, $f(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o)$ is the reflected sound energy. It is bidirectional, which means that $f(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o) = f(\boldsymbol{\omega}_o, \boldsymbol{\omega}_i)$. The rendering equation (see Sec. 2.3) states that the outgoing ray energy $L_o(\mathbf{x}, \boldsymbol{\omega}_o, t)$ can be expressed in the following terms:

$$\begin{aligned} L_o(\mathbf{x}, \boldsymbol{\omega}_o, t) = L_e(\mathbf{x}, \boldsymbol{\omega}_o, t) \; + \; ... \\ \int f(\mathbf{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o)\, L_i(\mathbf{x}, \boldsymbol{\omega}_i, t) \cos\theta_i \; d\boldsymbol{\omega}_i \; , \end{aligned} \tag{7.2}$$

where $L_e(\mathbf{x}, \boldsymbol{\omega}_o, t)$ is the emitted sound energy and $L_i(\mathbf{x}, \boldsymbol{\omega}_i, t)$ is the incoming sound energy under angle $\boldsymbol{\omega}_i$ at time $t$. As the emitted sound is 0 everywhere except for at the sound source, this term can be ignored and the sound source can be treated as a special case.

The BRDF used is loosely based around the Phong scattering model [155]. This is a simplified model of a scatterer that allows for defining a scattering function that is anything between a specular reflector and a Lambertian reflector, based on the exponent factor $k$. The formal description of the BRDF of the Phong scatterer is:

$$f_{\text{Phong}}(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o) = (1-\alpha)\frac{k+1}{2\pi}\cos^k(\boldsymbol{\omega}_o \cdot \boldsymbol{\omega}_{\text{spec}}), \tag{7.3}$$

where $k$ is some factor relating to the diffusion coefficient $\mu$, and $\boldsymbol{\omega}_{\text{spec}}$ is the specular reflection of $\boldsymbol{\omega}_i$. Fig. 7.6 shows the scattering distribution for different values of $k$. When $k = 1$, the surface behaves as a Lambertian reflector. For high values of $k$, it behaves in a predominantly specular fashion.

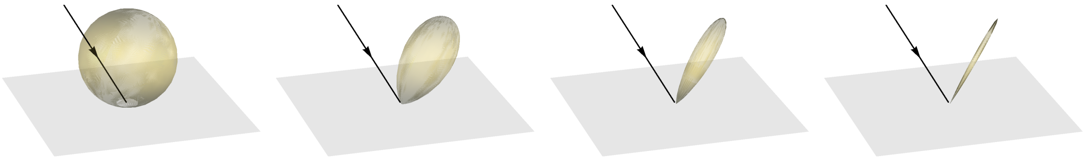


**Figure 7.6:** Phong scattering lobes with different coefficients for, from left to right $k = 1$, $k = 5$, $k = 30$, and $k = 500$.

The sound receiver used was a point receiver. Instead of registering physical ray hits as is common for a spherical receiver, the point receiver registers an impulse based on the reflected sound energy at each reflection, which can be computed exactly using Eq. 7.3. The advantage of this is that the contribution of a large number of (often small) hits can be registered, with no potential error related to the size or shape of the receiver chosen. Air absorption for each octave band was modelled using the air absorption coefficient at mid-frequency of each octave band. The air absorption coefficients were computed using the formula by Bass *et al.* [12] using the atmospheric conditions as described earlier in this section.

### 7.2.4 Combining octave bands

As the simulated IRs are computed separately for each frequency band, they need to be combined appropriately into a single RIR. To this end, a simple octave band approach was utilised (see e.g. [170]). This method combines the valid pass band, defined by the frequency range over which the absorption coefficients are applicable, of each impulse response by first band-pass filtering the responses around suitable cut-off frequency values and then summing the resulting signals. In this work, a bank of first order Butterworth filters was utilised, giving a 3 dB/octave reduction in magnitude above and below the defined cut-off frequencies provided in Table 7.5. After the band-pass filtering, the FDTD and ray-traced IRs were summed to produce the total IR.

**Table 7.5:** Overview of cut-off frequencies applied in the octave band filter bank.

| Octave Band (Hz) | Lower Cut-off (Hz) | Upper Cut-off (Hz) |
|---|---|---|
| 62.5 | 0 | 82 |
| 125 | 82 | 177 |
| 250 | 177 | 355 |
| 500 | 355 | 710 |
| 1000 | 710 | 1420 |
| 2000 | 1420 | 2840 |
| 4000 | 2840 | 5680 |
| 8000 | 5680 | 11360 |

As the FDTD and ray traced IRs are computed separately, they need to be combined in such a way that their respective energy levels are calibrated correctly. A number of energy calibration algorithms are described in related literature, e.g. [192, 184, 191]. For the purposes of this work, a simple energy calibration procedure, as discussed in [192], was deemed most suitable. More rigorous calibration techniques, such as described in [191], are valid only for high resolution FDTD schemes with temporal sampling frequencies greater than 18kHz, which was not feasible in this study. Following [192], a calibration parameter can be expressed as:

$$\eta = \frac{(f_2 - f_1)\sum_{i=g_1}^{g_2} |G_{IR}[i]|}{(g_2 - g_1)\sum_{i=f_1}^{f_2} |N_{IR}[i]|} \tag{7.4}$$

The above expression calculates the ratio of average magnitudes in the frequency spectrum of the high frequency ray-tracer IRs, $G_{IR}$, and low frequency FDTD IRs, $N_{IR}$, over a finite series of discrete frequency ranges $[f_1; f_2]$ and $[g_1; g_2]$ with index $i$. The FDTD IRs were calibrated with $f_2$ and $g_1$ set equal to the crossover frequency between low and high frequency IRs: $f_2 = g_1 = 355$ Hz. The upper and lower bounds of the frequency ranges were defined as $f_1 = 100$ Hz and $g_2 = 610$ Hz, hence the ratio of average magnitudes over a range of 255Hz above and below the crossover frequency was calculated and applied to each FDTD IR by multiplication for each IR.

Although this energy matching procedure is prone to several types of errors [192], it was deemed sufficient for the purposes of informal listening tests and octave band RIR analysis. Having calibrated the low frequency portion of each IR, the total IRs were created by summing the calibrated signals to the corresponding high frequency portions.

### 7.2.5 Results

The IRs have been successfully obtained for the $2^{\mathrm{nd}}$-order scheme with LRS boundaries, the $4^{\mathrm{th}}$-order scheme with SAT-SBP boundaries, and from the actual measurements in the NCEM. For the purposes of this chapter and this thesis, we are interested in the three lowest octave bands, which have been modelled using the FDTD solutions. A more detailed analysis of the full simulated $2^{\mathrm{nd}}$-order IR can be found in [209]. The parameters used for comparative analysis are $T_{30}$ reverberation time, EDT, and $T_c$. These parameters were chosen as fair comparison measures, as they depend least on source directivity and the receiver's polar pattern [67], neither of which are modelled by the FDTD algorithms as currently implemented.

Table 7.6 presents the aforementioned parameters for the three lowest octave bands for each receiver. Fig. 7.7 displays this information as bar plots. Additionally, in the bar plots the JND for the $T_{30}$ reverberation time measurements has been indicated using error bars. The value of the JND for reverberation time is a source of debate in literature, and although ISO-3382-1 [66] quotes a value of ±5% based on research by Seraphim [177] dated 1958, it is commonly accepted that the JND is considerably higher. Southern *et al.* [191] used an error margin of 30% in their plot, and a recent study by Blevins *et al.* [25] suggests an average JND of 24.5%. Rather erring on the side of caution, we use the latter value of 24.5% for $T_{30}$ time.

There is no consistent pattern for the circumstances under which the $4^{\mathrm{th}}$-order scheme has a higher or lower $T_{30}$ value. It does not seem to clearly correlate to position in the room, and even within the same receiver, different frequency bands show different behaviour. Most of the differences can likely be explained by the aforementioned difference in geometry voxelisation due to different stencil size. Moreover, as the analysed bands are a combination of different band-pass filtered simulations, sound energy from the ray-traced IRs bleeds into them, which may explain differences in frequency bands of the same receiver. All $T_{30}$ differences between the $2^{\mathrm{nd}}$- and $4^{\mathrm{th}}$-order accurate bands are within the JND, however, though this is not the case for either simulation compared to the measurements. EDT and $T_c$ values also demonstrate very little consistent difference between $2^{\mathrm{nd}}$- and $4^{\mathrm{th}}$-order accurate scheme, and though the values seem plausible in general, from the comparison between $2^{\mathrm{nd}}$- and $4^{\mathrm{th}}$-order IRs both with each other and with the recorded IR, there is no saying which algorithm is more 'correct.'

A particular note should be made about several clear outliers: the R2 125 Hz band, the R3 125Hz band, the R6 125 band, and the R7 250 Hz band. The former three have a much lower $T_{30}$ time compared to both simulations, while in the latter case, the 4th-order accurate scheme has significantly different values for both $T_{30}$ and $T_c$. These receivers all being relatively close to one another in the space, it may well be that some material or geometry may have been overlooked in the simulations, or that some material close to these receiver positions had particular acoustic properties that were not modelled in the simulation. The deviations in the R7 receiver are harder to explain: it seems like the calibration with the higher frequency bands has been incorrect and leakage from the band-pass filtering process with the ray-traced response is influencing the $T_{30}$ times of the lower bands too much. It is unknown, however, why this occurs in only one out of 7 receivers.

**Table 7.6:** Comparison of the 2nd-order schemes using LRS boundaries, the 4th-order scheme using SAT-SPB boundaries, and the measured IRs. The analysed acoustic parameters are $T_{30}$, EDT, and $T_c$ for the receivers R1–R7. For the $T_{30}$ measurements, the simulation values within the 24.5% JND of the measurements are printed in **bold**.

| | | 62.5 Hz | | | 125 Hz | | | 250 Hz | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | FD 2nd | FD 4th | Meas. | FD 2nd | FD 4th | Meas. | FD 2nd | FD 4th | Meas. |
| R1 | $T_{30}$ | 2.52 | 2.22 | 1.76 | **2.05** | **1.96** | 1.80 | **1.93** | **1.90** | 1.75 |
| | EDT | 2.32 | 1.99 | 2.04 | 2.04 | 1.67 | 1.99 | 1.89 | 1.70 | 1.74 |
| | $T_c$ | 0.18 | 0.15 | 0.15 | 0.13 | 0.13 | 0.09 | 0.14 | 0.11 | 0.17 |
| R2 | $T_{30}$ | **2.36** | **2.21** | 2.24 | **2.03** | **1.96** | 1.75 | **2.02** | **1.89** | 1.80 |
| | EDT | 2.40 | 2.37 | 2.03 | 2.02 | 1.92 | 1.46 | 2.05 | 1.71 | 1.66 |
| | $T_c$ | 0.16 | 0.11 | 0.13 | 0.12 | 0.12 | 0.12 | 0.13 | 0.12 | 0.12 |
| R3 | $T_{30}$ | 2.50 | **2.29** | 1.92 | **1.92** | **2.13** | 1.89 | **2.01** | **1.94** | 1.73 |
| | EDT | 1.87 | 2.27 | 1.82 | 1.60 | 2.01 | 1.43 | 2.22 | 1.89 | 1.96 |
| | $T_c$ | 0.14 | 0.13 | 0.14 | 0.11 | 0.10 | 0.11 | 0.14 | 0.15 | 0.13 |
| R4 | $T_{30}$ | 2.15 | 2.31 | 1.68 | **1.98** | **2.07** | 1.87 | **1.89** | **2.04** | 1.82 |
| | EDT | 2.06 | 2.46 | 2.18 | 1.85 | 2.13 | 1.65 | 1.76 | 1.96 | 1.58 |
| | $T_c$ | 0.15 | 0.16 | 0.15 | 0.12 | 0.12 | 0.09 | 0.14 | 0.13 | 0.13 |
| R5 | $T_{30}$ | **2.38** | **2.43** | 2.16 | **1.99** | **1.99** | 1.86 | **1.93** | **2.03** | 1.67 |
| | EDT | 1.98 | 1.92 | 1.95 | 1.57 | 1.60 | 1.67 | 1.78 | 2.06 | 1.78 |
| | $T_c$ | 0.15 | 0.14 | 0.13 | 0.13 | 0.11 | 0.10 | 0.12 | 0.13 | 0.13 |
| R6 | $T_{30}$ | **2.27** | **2.14** | 2.26 | **1.98** | **2.05** | 1.88 | **1.98** | **1.87** | 1.79 |
| | EDT | 2.33 | 2.08 | 1.92 | 1.83 | 1.90 | 1.39 | 1.65 | 1.83 | 1.70 |
| | $T_c$ | 0.10 | 0.16 | 0.10 | 0.12 | 0.13 | 0.08 | 0.12 | 0.14 | 0.08 |
| R7 | $T_{30}$ | **2.30** | **2.16** | 1.95 | **2.09** | **1.87** | 1.78 | **2.02** | **1.70** | 1.67 |
| | EDT | 2.25 | 1.99 | 2.00 | 1.92 | 1.45 | 1.67 | 2.00 | 0.97 | 1.76 |
| | $T_c$ | 0.13 | 0.07 | 0.13 | 0.10 | 0.05 | 0.10 | 0.14 | 0.04 | 0.11 |

A more concise form of the results is presented in Table 7.7, which shows the average parameter values, their standard deviations, and the mean differences in absolute values and percentages. We define the *mean difference* between the parameter values of two simulations as the mean of the absolute value of the differences across all receivers and frequency bands. It can be observed that while the mean difference between the two simulations is smaller than between the simulations and measurements, there are still clear and noticeable differences between the individual responses for each receiver and octave band.

The simulated IRs present a higher reverberation time than the measured IRs, though the $4^{\text{th}}$-order scheme's $T_{30}$ is slightly lower. This is unexpected, as the previous shoebox model showed it to have a consistently higher reverberation time. The lower reverberation time may be explained by the voxelisation process: as the $4^{\text{th}}$-order stencil is double the width of the $2^{\text{nd}}$-order scheme, it will shape rougher forms around small geometry and in corners, and will thus decrease the volume of the modelled space, as such decreasing reverberation time.

Several more observations can be made from Table 7.7. The $T_{30}$ estimations are the best, and both simulations show, on average a difference of around 11% with respect to the measurement values. The $2^{\text{nd}}$- and $4^{\text{th}}$-order are expectedly very close, with an average of 5.8% mean difference, which is well within the JND for this parameter. As mentioned earlier, comparing the two simulations, all individual bands are within the JND of one another. EDT and $T_c$ values have a much higher average error, of 14.6% and 19.1%, respectively. However, it should be noted the outliers mentioned earlier in this section have a disproportionate adverse effect on these results. Although reliable JND values for these values are absent, these large differences will most likely result in all equivalent $2^{\text{nd}}$- and $4^{\text{th}}$-order accurate simulated IRs being audibly different from one another.

### 7.2.6 Audio examples

In order to get a more subjective evaluation of the quality of these results and an assessment of audible differences, the hybrid IRs described in previous sections have been convolved with two different anechoic sources: a sample of adult female speech (retrieved from the Open AIR database [150]) and a synthesised sample drum loop. All audio files have been appended to this thesis. The previous sections have already concluded that the differences between the different methods particularly for EDT and $T_c$ are large, so it is argued that extensive listening tests such as ABX tests would be of little value, as it is very likely that differences between simulated and recorded responses would always be observed. Nevertheless, a small-scale informal listening test was performed in order to add context and weight

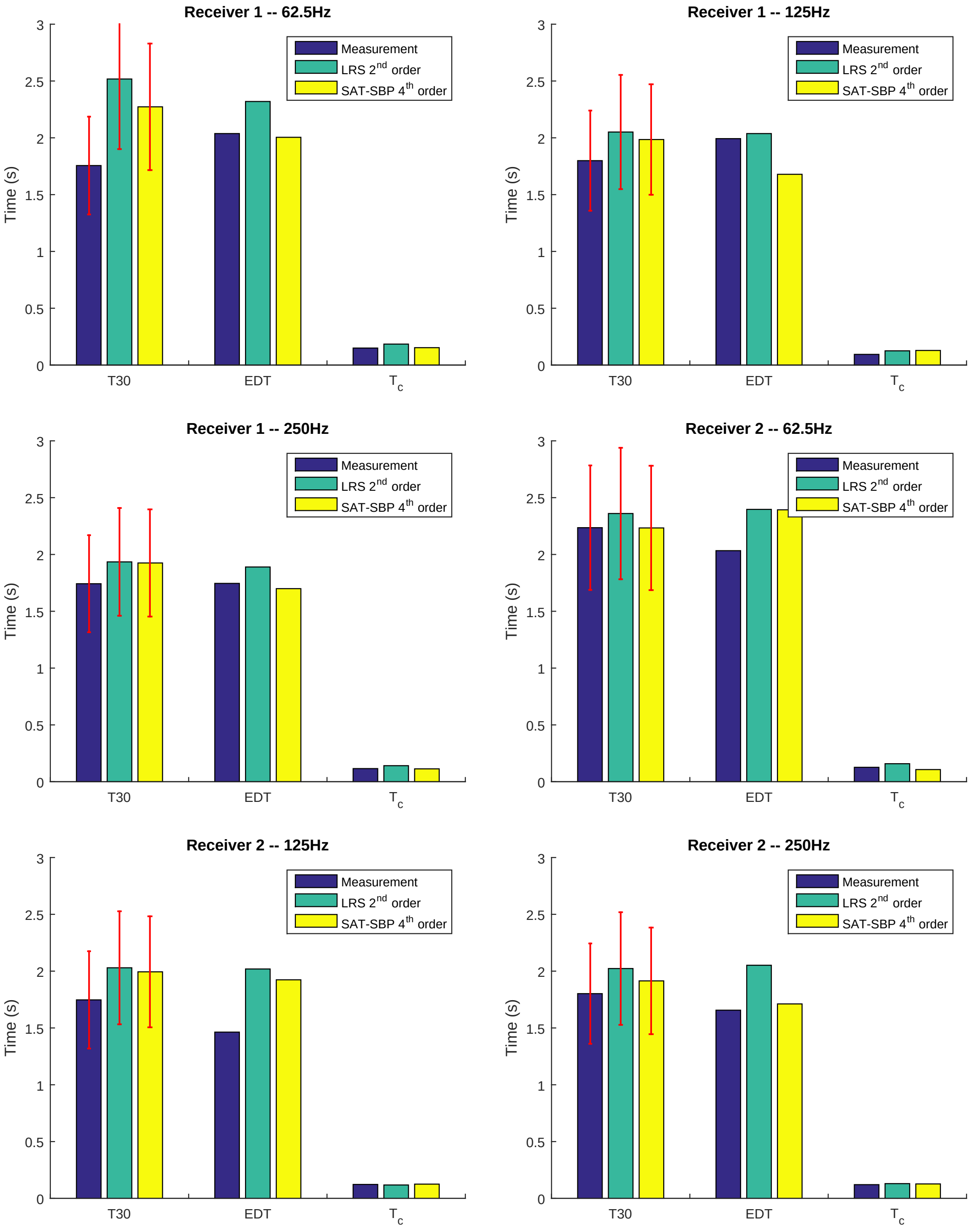


**Figure 7.7:** (1/4) Bar graphs showing the analysed acoustic parameters for each receiver and octave band. The red error bar indicates the 24.5% JND for $T_{30}$.

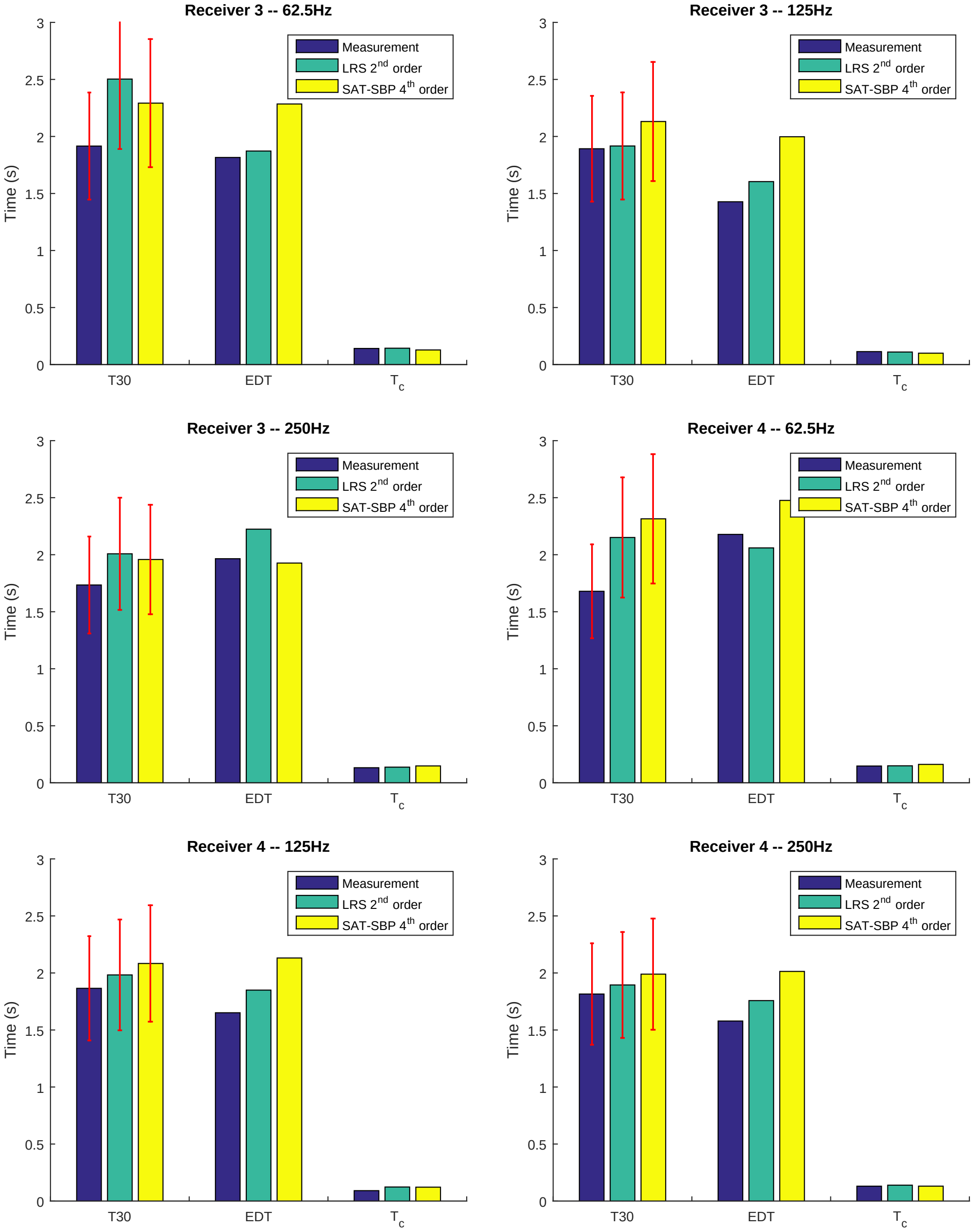
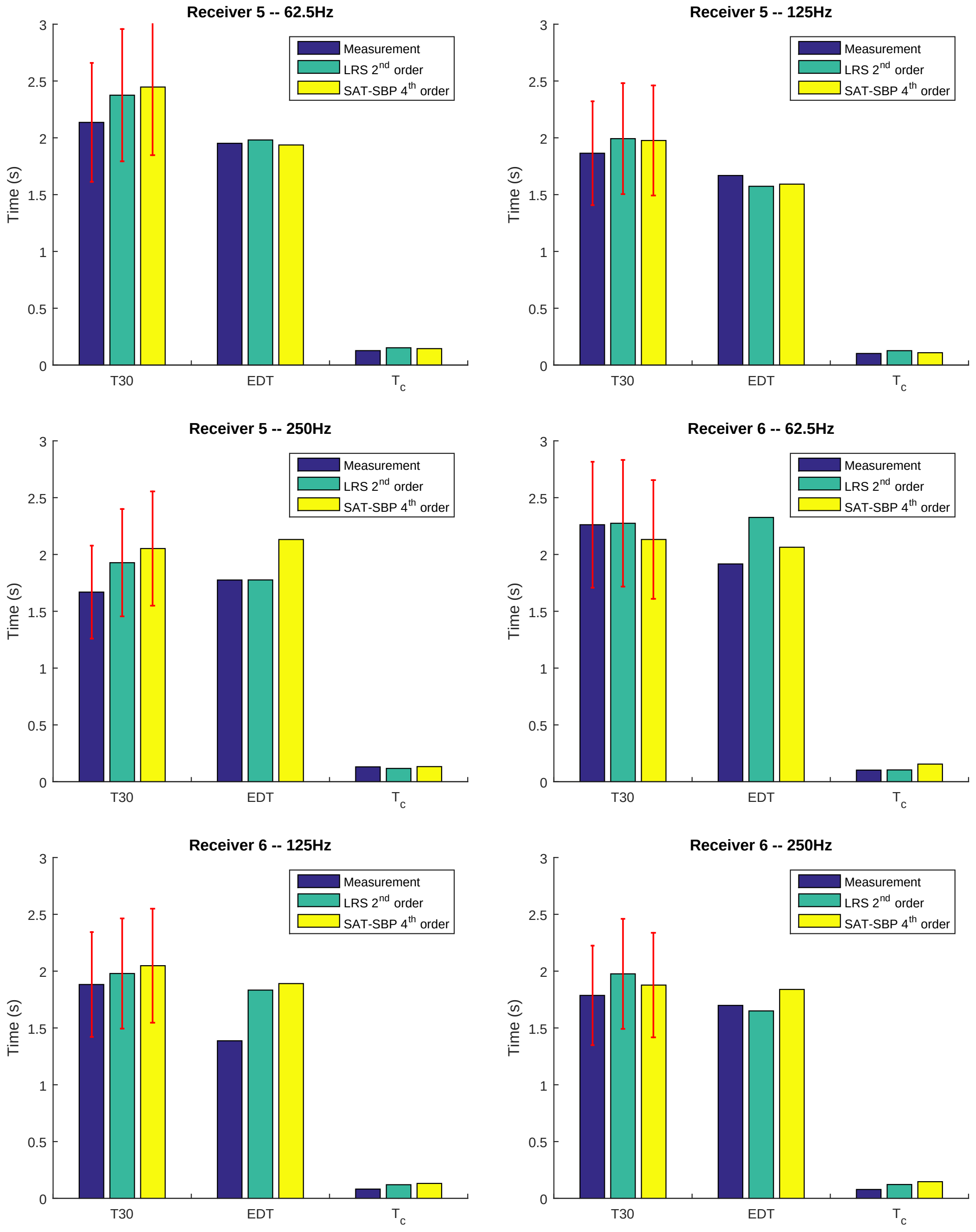


**Figure 7.7:** (2/4) Bar graphs showing the analysed acoustic parameters for each receiver and octave band. The red error bar indicates the 24.5% JND for $T_{30}$.


Receiver 5 -- 62.5Hz

Receiver 5 -- 125Hz

Receiver 5 -- 250Hz

Receiver 6 -- 62.5Hz

Receiver 6 -- 125Hz

Receiver 6 -- 250Hz

**Figure 7.7:** (3/4) Bar graphs showing the analysed acoustic parameters for each receiver and octave band. The red error bar indicates the 24.5% JND for $T_{30}$.

**Table 7.7:** The mean and standard deviation of reverberation time across all receivers and frequency bands, for each analysed acoustic parameter. The parameter differences are measured versus the recorded IR.

| | Measurement | FDTD 2<sup>nd</sup> | FDTD 4<sup>th</sup> |
|---|---|---|---|
| $T_{30}$ | $1.86 \pm 0.17$ | $2.11 \pm 0.19$ | $2.05 \pm 0.17$ |
| $T_{30}$ difference vs. Meas. | - | $0.25 \pm 0.10$ | $0.19 \pm 0.06$ |
| Mean $T_{30}$ error vs. Meas. | - | $11.7\% \pm 7.4\%$ | $10.4\% \pm 6.5\%$ |
| Mean $T_{30}$ error vs. FD2 | - | - | $5.8\% \pm 4.9\%$ |
| EDT | $1.78 \pm 0.22$ | $1.97 \pm 0.23$ | $1.90 \pm 0.32$ |
| EDT difference vs. Meas. | - | $0.19 \pm 0.03$ | $0.12 \pm 0.11$ |
| Mean EDT error vs. Meas. | - | $10.7\% \pm 7.3\%$ | $15.8\% \pm 17.9\%$ |
| Mean EDT error vs. FD2 | - | - | $14.6\% \pm 11.2\%$ |
| $T_c$ | $0.116 \pm 0.020$ | $0.132 \pm 0.020$ | $0.120 \pm 0.001$ |
| $T_c$ difference vs. Meas. | - | $0.015 \pm 0.001$ | $0.004 \pm 0.006$ |
| Mean $T_c$ error vs. Meas. | - | $13.3\% \pm 11.0\%$ | $30.0\% \pm 43.4\%$ |
| Mean $T_c$ error vs. FD2 | - | - | $19.1\% \pm 19.4\%$ |

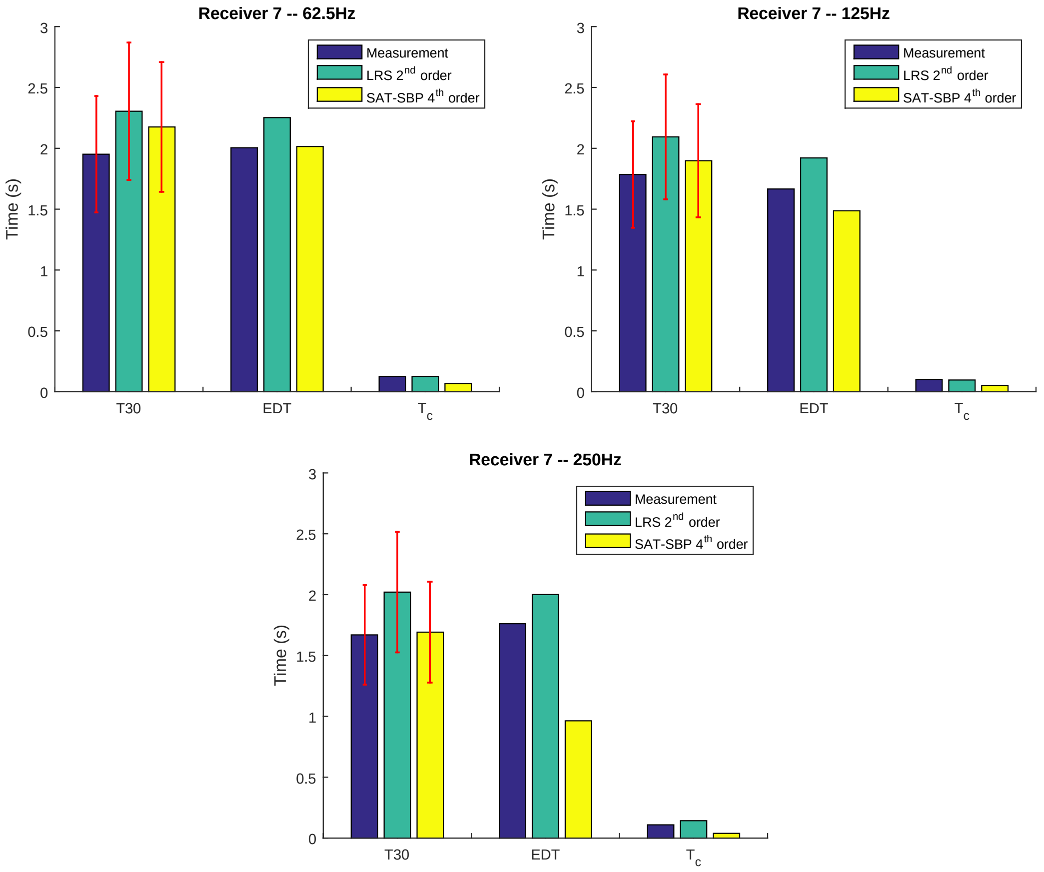


**Figure 7.7:** (4/4) Bar graphs showing the analysed acoustic parameters for each receiver and octave band. The red error bar indicates the 24.5% JND for $T_{30}$.

to the analytical results. It is informal in the sense that no specific quantitate or qualitative constructs are being examined or compared, but that the answers are merely used to get a broader understanding of the qualitative differences between the different IRs, as assessed by audio experts.

Seven test subjects were asked to evaluate the aforementioned speech and drum sample for each receiver, and to objectively and subjectively assess their differences. The subjects were all audio professionals with jobs in music and sound design, all with an extensive experience in critical listening and no known hearing disabilities. Industry standard quality headphones were used to render the sound. The samples from the three different methods were placed in random order for each receiver and marked 'A', 'B', and 'C' as to not bias the test subjects. The subjects were allowed to listen to the samples as many times as they liked, and were asked to assess the differences. The subjects were given an introduction as to the purpose of the tests, and assignment on the questionnaire was phrased as follows:

> "For each impulse response, there is a speech and a drum sample. For all 7 microphone positions, please describe the objective and subjective difference between the three impulse responses based on both samples."

Two subjects did not complete the full test and only replied to R1–R4. The full answers can be found in B. A summary of their assessments for all seven receiver positions follows here.

**R1 (total 7 answers)** The 2<sup>nd</sup>-order impulse response (FD2) drum sample is said to be heavier in low-frequency content ('bassier' and 'boomier') than the measurement IR by all subjects. It is also noted that the kick drum is much more prominent in the FD2 IR. However, three subjects indicate they hear little to no perceivable difference between these two IRs for the speech samples. The FD4 is noted to have much less low end than FD2 and to be 'clearer' and 'brighter', and the speech sample is found to be more intelligible. The reverberant tail of the FD4 is described to sound 'boxier' and with more resonances than the measurement sample.

**R2 (total 7 answers)** Three out of seven subjects express that the measurement has the longest reverberant tail, and also to be the brightest. The difference is most noticeable in the drum sample, and the snare drum is said to be less prominent in the FD2 IR, which itself is also said to be 'woolier'. One subject notes they cannot tell the difference between the measurement IR and the FD2 in the speech sample. The FD4 is generally felt to be 'closer', mainly because it is perceived to have less reverb, though one subject claims to perceive a longer tail than in FD2. It is also described to be more ‘metallic’, but ‘cleaner’ in the speech sample.

**R3 (total 7 answers)** The FD2 IR is perceived by all but one subjects to be extremely bass-heavy, particularly in the drum sample, though the length of the tail is considered to be the same. FD4 is called ‘nicer’ and ‘cleaner’ than FD2, though “not as nice to listen to as [the measurement IR] because of pronounced cluttered [mid frequencies].”

**R4 (total 7 answers)** The measurement IR and the FD2 are perceived as being very similar, and particularly the speech samples are considered close to identical. The FD2 drum sample is said by three participants to have some more resonance in the low mid frequency, and thus be slightly more ‘boomy’. The FD4 IR is generally considered to be more ‘occluded’, ‘wooly’, and ‘with lack of high end’.

**R5 (total 5 answers)** The FD2 IR is said to be the “most metallic of the three,” but also to have “more clarity in [the mid frequency region]” compared to the measurement IR. The measurement and FD2 speech samples are said to be ‘sound the same’ and ‘very similar’ and ‘[with] negligible difference’ by three subjects, whereas two subjects perceive the FD4 to be ‘closer’ and ‘with shorter reverb’ than the other two. One subject says that ‘[The FD4 IR] sounds the most pleasing of the three’, but another says that it ‘is overly compressed’.

**R6 (total 5 answers)** The measurement samples are considered much ‘bassier’ and ‘more muffled’ than the FD2 ones by all, and one subject says the former “sound unnatural.” FD4 has varying assessments, from ‘really nice [...] maybe a touch too much low end but otherwise sounds excellent’ to ‘very electronic, and close in comparison to [the FD2 IRs].’ It is also again perceived to be ‘dryer’ and therefore also ‘closer.’

**R7 (total 5 answers)** The FD2 samples are called ‘duller’, ‘drier’, and also ‘more muffled.’ The FD4 is found to be more reverberant than the former two, but still heavy on the low end (‘wooly’). The measurement IRs are called ‘wetter [than FD2]’ and to have ‘more high frequencies.’ It is perceived to be closer to the measurement samples, but ‘seems to lack lower mids’, and is called ‘boomy’ twice.

In conclusion, the simulations indeed sound different from their recorded counterparts, and also the $2^{\text{nd}}$- and $4^{\text{th}}$-order accurate stencils are audibly different in nearly all cases except some speech sampels. As was concluded earlier in [209], the state-of-art of virtual acoustic modelling, including aforementioned hybrid methods, are not yet advanced enough to produce indistinguishably similar results compared to recorded ones. However, all test subjects agreed that several of the simulated IRs sounded extremely realistic and, while they might not accurately represent the space, they would use them as part of their sound design work.

### 7.2.7 Summary

The classic $2^{\text{nd}}$-order scheme with LRS boundaries and the $4^{\text{th}}$-order scheme with SAT-SBP boundaries have been compared against the recorded IRs of a complex acoustic space. While both algorithms produce similar IRs, their differences are more complex than the initial shoebox tests suggest. Compared to the recorded IRs, all simulated IRs are objectively and subjectively different, though the extent of this varies for each receiver. Note, however, that the model that was run, including the geometry and chosen materials were all fitted *a priori*. Due to the complex nature of acoustic interactions, differences can easily be explained by slight differences in geometry or chosen material types [67]. It would be possible to edit the model parameters further to get a closer match *a posteriori*, but this is subject for future work.

## 7.3 Conclusion and discussion

Two acoustic simulation experiments have been performed as comparison and benchmark for the $4^{\text{th}}$-order scheme with SAT-SBP boundaries: a large-scale theoretical one using a shoebox model, and a real-world example of a complex acoustic environment. From the shoebox experiment, it appears that the time and frequency domain characteristics are preserved and compare well to the current state-of-art LRS boundaries. In terms of decay time, $T_{45}$ analysis suggests that the absorptive properties of the SAT-SBP boundaries are slightly lower than that of the LRS boundaries for all reflection values.

Using York's National Centre for Early Music as a case study, it was shown that the $4^{\text{th}}$-order scheme with SAT-SBP boundaries holds up well in a real acoustic modelling scenario. More study is required to create simulated FDTD impulse responses that compare even better to their real-world equivalents. As the $4^{\text{th}}$-order method has a more favourable dispersion pattern and can use a coarser grid spacing, it is an improvement over the state-of-art. However, neither method achieves to deliver IRs that are indistinguishable from measurements.

# Chapter 8

# Summary and Conclusions

**Contents**

This thesis has focussed on the FDTD method in the context of room acoustic modelling, and on methods to increase its performance and accuracy by devising more precise air and boundary stencils. The main motivation is to improve the overall state of art of room acoustic modelling and produce more accurate simulated IRs in a feasible time frame. Chapter 2 and Chapter 3 have laid the foundation in terms of concepts related to acoustic wave propagation and physical boundary interaction. Chapter 4 uses these concepts to describe the workings of a number of common room acoustic modelling methods. While not all of these methods are directly relevant to this work, they can all be part of a hybrid room acoustic modelling method, e.g. one where different octave bands are produced by different methods. Particular attention is paid to the family of ray tracing methods as they are part of the hybrid approach used in Chapter 7, and to the $2^{\text{nd}}$-order accurate FDTD method, which stands as the basis for comparisons with the higher-order accurate FDTD methods derived in later chapters.

## 8.1 Higher-order air stencils

The first contribution of this thesis is a thorough investigation of discrete-time wave propagation in free space. A particular family of higher-order accurate stencils has been investigated. Sakamoto [162] took a similar approach in a smaller-scale study on dispersion error on these stencils, and concluded that "the unavoidable phase error in the ordinary FDTD algorithm is reduced by adopting the scheme using multiple reference points." The research presented in Chapter 5 details an analysis of the dispersive properties of FDTD stencils along the axial and diagonal axes for orders of accuracy 2 up to 16, and arrives at the same conclusion. Additionally, several other conclusions were reached. First of all, the Courant stability limits for several stencils quoted by Sakamoto were slightly too limiting and it was shown they can be less restrictive. The update values for all stencils up to $20^{\text{th}}$ order and their respective stability limits were also presented for future reference. A major conclusion of this work is the higher grid spacing that is implied by a lower stability limit, which suggests that at equal sample rates, the higher-order accurate stencils provide large time and memory savings over their $2^{\text{nd}}$-order counterpart. A particular note should be made of the IWB scheme, which in terms of dispersion is two times better than the best higher-order accurate scheme, but also requires at least 8 times the amount of memory. It is therefore not the most optimal stencil from a computational point of view.

In order to verify the theoretical results, a GPU implementation was analysed, performing acoustic simulations of several different stencils on three different shoebox-type rooms. The timings clearly shows that the $4^{\text{th}}$- and $6^{\text{th}}$-order scheme are among the most efficient. Despite the lower dispersion error of the IWB, this does not compensate for the latter's high memory usage, which is clearly a bottleneck in the largest model explored as part of this investigation. It should be noted that since these results were first published in [208], Hamilton *et al.* [82] made significant improvements to the IWB GPU implementation and repeated the full experiment. While their IWB results are better, the $4^{\text{th}}$- and $6^{\text{th}}$-order scheme are still the most efficient options.

Finally, room mode plots for a small shoebox-type room have been analysed, and the shifts in frequencies have been quantified. The results strongly confirm all theory presented earlier in the chapter: room mode shifts can be observed and are particularly significant in the $2^{\text{nd}}$- and $4^{\text{th}}$-order accurate simulations, and occur most prominently for the axial and oblique modes, respectively.

## 8.2 Higher-order boundary stencils

A major shortcoming of higher-order accurate stencils used to be the absence of boundary implementations to model the impedance boundary condition. Sakamoto [162] used lower-order stencils at the boundaries, which reduces the order of accuracy of the whole simulation, rendering this approach not a true higher-order accurate method. While fully reflective hard and soft boundaries have been presented in [130], the impedance boundary condition was only formulated, but not implemented or even investigated in terms of stability.

Several approaches to obtain stable boundary implementations of higher-order accuracy have been explored. The so-called energy method has been used to formulate the problem. SBP operators for the wave equation were introduced according to GKS theory. Because of the non-constant boundary condition, the SAT method was introduced, as it has been used in similar cases [37, 38, 52] to solve boundary condition problems. Ditkowski [52] presented an analytical way of putting a bound on the error growth of the wave equation with boundary conditions, but it was shown in this thesis that Ditkowski's approach does not apply to the impedance boundary.

A novel solution inspired by the aforementioned methods was used to devise stable $4^{\text{th}}$-order accurate boundaries. It uses some adjustment of SBP-SAT operators to generate a stable solution with absorbing boundaries. The matrix method is employed to show its stability in the 1D, 2D, and 3D case. Subsequently, more rigorous Von Neumann analysis was used to prove stability, though the stability limit depends on both the Courant number $\lambda$ and the admittance value $\xi$. In the 3-dimensional case, however, the majority of absorption cases is covered by this range. Absorption analysis was done for the 1D, 2D, and 3D case on small domains, and it was shown that the absorption characteristics closely match the $2^{\text{nd}}$-order cases.

The results presented in Chapter 6 are very valuable in the context of the present-day research and use of the FDTD method. While higher-order accurate methods have been subject to moderate investigation for over a decade, to the best of the author's knowledge they have never been used for practical acoustic modelling problems, largely because of the absence of absorbing boundaries. This work helps to overcome this limitation and makes higher-order accurate FDTD modelling a real possibility, as demonstrated in Chapter 7.

## 8.3 Acoustic modelling case studies

Finally, Chapter 7 has a twofold purpose: to demonstrate the validity of the newly-derived 4th-order boundaries compared to the current most commonly implemented 2nd-order case, and to demonstrate the validity and feasibility of the approach in a real-world scenario. Time and frequency domain analyses show that the behaviour of the 4th-order accurate boundaries is nearly identical to the 2nd-order LRS boundaries, with some differences attributed to the difference in accuracy and dispersion. The absorptive properties have been analysed using a range of typical reflection values. It was shown that the $T_{45}$ times for the 4th-order accurate stencil are around 5% higher than the equivalent 2nd-order models with LRS boundaries from [103]. This difference is commonly assumed to be within the JND for reverberation time and is unlikely to be noticeable. Additionally, the differences have been quantified in this work, so the reflection values can be compensated *a priori* in order to obtain a response even closer to the 2nd-order implementation.

The final case study of the National Centre for Early Music uses the formerly presented 4th-order scheme and boundary in a hybrid acoustic model that combines FDTD modelling for the lower three octave bands and ray tracing for the higher five. IRs measured from one sound source position and several receiver positions were generated using this new method. Equivalent RIRs were measured *in situ*, and the perceptual parameters $T_{30}$, $T_c$, and EDT were compared with the IR measurements and with IRs obtained using the 2nd-order accurate simulation method. Both the 2nd- and 4th-order accurate hybrid models were shown to produce IRs that are suitable for auralisation. Their $T_{30}$ values are in close agreement and on all within the JND, though the values for EDT and $T_c$ show larger variations. Both, however, are significantly different from results obtained from the measurements, but due to the many uncertainties and inaccuracies potentially introduced by the acoustic modelling process, this experiment does not draw a conclusion as to which approach might be more accurate.

## 8.4 Restatement of hypothesis

Central to this thesis has been the improvement of the FDTD algorithm using a more accurate stencil, with corresponding boundaries, for room acoustic simulation problems. The hypothesis stated at the beginning of this thesis was:

> The ‘leggy’ 4th-order Finite Difference stencil in combination with the presented

corresponding boundaries presents an improvement in both output quality and computational speed compared to the currently prevailing Standard Rectilinear and Interpolated Wideband $2^{\mathrm{nd}}$-order schemes.

This has been tested and verified in the following ways:

- A mathematical analysis of dispersion error has been presented that has thoroughly compared all 'leggy' stencils up to $16^{\mathrm{th}}$ order, and the $2^{\mathrm{nd}}$-order IWB one. It was shown that the IWB is the least dispersive stencil, and the quality of the leggy stencils improves with increasing order of accuracy, but with diminishing returns.
- Stability limits for all 'leggy' stencils up to $24^{\mathrm{th}}$ have been derived that in some case are more leniant than in previous literature [162] and the implications for grid spacing and therefore time and memory are clearly indicative of better performance for higher-order stencils. Practical tests show that the $4^{\mathrm{th}}$- and $6^{\mathrm{th}}$-order accurate ones are the most optimal. This includes comparison to the IWB scheme.
- Boundary implementations for the $4^{\mathrm{th}}$-order accurate scheme were devised. The absorptive properties have been analysed in depth and are shown to give reverberation time results that are within the JND compared to the $2^{\mathrm{nd}}$-order scheme. The differences in absorption have been quantified and can be compensated for by *a priori* by adjusting the impedance value in the $4^{\mathrm{th}}$-order scheme.
- Two case studies have been used to demonstrate the feasibility and practicality of the $4^{\mathrm{th}}$-order scheme for real-life room acoustic modelling problems. All $T_{30}$ times of the $2^{\mathrm{nd}}$- and $4^{\mathrm{th}}$-order accurate scheme are within the JND of one another, though a small-scale listening experiment show that there are audible differences between the two.

In summary, it has been shown that the $4^{\mathrm{th}}$-order accurate leggy stencil has many advantages over the $2^{\mathrm{nd}}$-order schemes, both the SRL and IWB. What differences there are have been quantified and can be compensated for. Examples have shown that it is equally suitable for IR modelling for auralisation purposes, with the benefit of increased computational speed and a smaller memory requirement.

The results presented in this thesis make wave-based room-acoustic modelling yet a more attractive means of acoustic modelling. The presented solution can solve existing acoustic problems in a more optimal way, and allows for larger spaces to be modelled. Both are attractive features in room acoustic consultancy, and it is therefore expected that wave-based methods, including the one presented in this thesis, will make their way into commercial acoustic modelling applications.

## 8.5 Limitations and further work

New 4th-order accurate stencils that approximate the acoustic wave equation and related impedance boundary conditions have been derived. While there are many more 4th-order accurate stencils, those focussed on as part of this thesis provide the easiest solution in terms of implementation and boundary implementation. Though a method for deriving boundary solutions has been presented, only the 2nd- and 4th-order cases have been derived and implemented in this work. It remains a topic of further investigation as to whether the same approach is valid for orders of accuracy higher than 4. The stability limit being dependent on the impedance value $\xi$ is undesirable, and therefore it would be worthwhile finding alternative boundaries without such a restriction. It is unknown, however, whether these boundaries do in fact exist.

Several recent advances have been made in combining the FDTD and FVTD method [19, 21] to more accurately represent non-rectilinear geometry and to avoid the staircasing boundary problem. These results showed that the staircasing problem can have a large negative impact on simulation accuracy, and used the energy method to prove the stability of the proposed FVTD method. A more recent study by Bilbao and Hamilton [20] extends this method to the wave equation with viscothermal losses and impedance boundaries. The usefulness of the FVTD was clearly demonstrated by modal analysis of several shoebox models in different alignments. A combination of higher-order methods presented in this work and higher-order accurate FV boundaries would therefore be a significant improvement to the methods put forward in this thesis. It is an open question whether these 4th-order accurate formulations of FVTD boundaries exist, and this is recommended as a topic of further study.

While only a small subset of higher-order accurate stencils has been examined in this work, there are infinitely many more. Hamilton and Bilbao [81] classified the 2D 4th-order accurate stencils into families, and used numerical optimisation to find the optimal stencil with the lowest dispersion error. This technique is extensible to 3D schemes, and may result in more optimal stencils than those presented here. The trade-off between small dispersion error and grid spacing might mean that, as with the IWB scheme, the more optimal schemes require a smaller grid spacing, but this is also a subject for further study. It is also a non-trivial task to extend the boundary model presented in this thesis to different, ‘non-leggy’ stencils.

It should be noted that the methods described in this thesis are not only suitable to room-acoustics, but also to other problems in the field, such as the acoustics of instruments and the voice. Indeed, preliminary experiments with vocal tract modelling using this method have yielded promising results. More perceptual research is needed in order to get a better understanding and validation of the perceptual qualities of IRs produced by hybrid acoustic modelling methods. Comparisons of computer-generated IRs and measurements will potentially give more insight into the strengths and weakness of computational models, and may help close the gap between the two. The efforts by Hornikx *et al.* [89] provide an excellent way to benchmark acoustic modelling methods in a standardised way. It would therefore be very desirable to take the algorithm presented in this thesis and use it to produce room acoustic simulations of the standardised models they propose. This is also subject for further research, and will give another objective judgement on the performance of the presented method.

# Appendix A

# SBP matrices

In this appendix, the summation-by-parts operators for the second derivative up to sixth-order accuracy are listed. The notation used is $D_p^{(2)}$ for the $p$-th order accurate approximation of the 2nd derivative. We use $D_p^{(2)} = H^{-1}\left(S_p - M_p^{(2)}\right)$, where $H$ is a positive definite diagonal norm, $M_p^{(2)}$ is a positive definite symmetric matrix and $S_p$ is the approximation of the first derivative at the boundary.

## $2^{\text{nd}}$ order

$$H = \begin{pmatrix} \frac{1}{2} & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \end{pmatrix} \tag{A.1}$$

$$S_2 = \frac{1}{X^2}\begin{pmatrix} \frac{3}{2} & -2 & \frac{1}{2} & \\ & 0 & & \\ & & \ddots & \end{pmatrix} \tag{A.2}$$

$$M_2^{(2)} = \frac{1}{X^2}\begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \tag{A.3}$$

$$D_2^{(2)} = \frac{1}{X^2}\begin{pmatrix} 1 & -2 & 1 & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \tag{A.4}$$

## $4^{\text{th}}$ order

$$H = \begin{pmatrix} \frac{17}{48} & & & & \\ & \frac{59}{48} & & & \\ & & \frac{43}{48} & & \\ & & & \frac{49}{48} & \\ & & & & \ddots \end{pmatrix} \tag{A.5}$$

$$S_4 = \frac{1}{X^2}\begin{pmatrix} \frac{11}{6} & -3 & \frac{3}{2} & -\frac{1}{3} \\ & 0 & & \\ & & \ddots & \end{pmatrix} \tag{A.6}$$

$$M_4^{(2)} = \frac{1}{X^2}\begin{pmatrix} \frac{9}{8} & -\frac{59}{48} & \frac{1}{12} & \frac{1}{48} & & & & \\ -\frac{59}{48} & \frac{59}{24} & -\frac{59}{48} & & & & & \\ \frac{1}{12} & -\frac{59}{48} & \frac{55}{24} & -\frac{59}{48} & \frac{1}{12} & & & \\ \frac{1}{48} & 0 & -\frac{59}{48} & \frac{59}{24} & -\frac{4}{3} & \frac{1}{12} & & \\ & & \frac{1}{12} & -\frac{4}{3} & \frac{5}{2} & -\frac{4}{3} & \frac{1}{12} & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \tag{A.7}$$

$$D_4^{(2)} = \frac{1}{X^2}\begin{pmatrix} 2 & -5 & 4 & -1 & & & & \\ 1 & -2 & 1 & & & & & \\ -\frac{4}{43} & \frac{59}{43} & -\frac{110}{43} & \frac{59}{43} & -\frac{4}{43} & & & \\ -\frac{1}{49} & 0 & \frac{59}{49} & -\frac{118}{49} & \frac{64}{49} & -\frac{4}{49} & & \\ & & -\frac{1}{12} & \frac{4}{3} & -\frac{5}{2} & \frac{4}{3} & -\frac{1}{12} & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \tag{A.8}$$

## $6^{\text{th}}$ order

$$H = \begin{pmatrix} \frac{13649}{43200} & & & & & & & \\ & \frac{12013}{8640} & & & & & & \\ & & \frac{2711}{4320} & & & & & \\ & & & \frac{5359}{4320} & & & & \\ & & & & \frac{7877}{8640} & & & \\ & & & & & \frac{43801}{43200} & & \\ & & & & & & 1 & \\ & & & & & & & \ddots \end{pmatrix} \tag{A.9}$$

$$S_6 = \frac{1}{X^2} \begin{pmatrix} \frac{25}{12} & -4 & 3 & -\frac{4}{3} & \frac{1}{4} & \\ & 0 & & & & \\ & & \ddots & & & \end{pmatrix} \tag{A.10}$$

$$M_6^{(2)} = \frac{1}{X^2} \begin{pmatrix} -\frac{15583}{12960} & \frac{253093}{172800} & -\frac{52391}{129600} & \frac{68603}{259200} & -\frac{2351}{14400} & \frac{4207}{103680} & & & & & \\ \frac{253093}{172800} & -\frac{42353}{12960} & \frac{134603}{51840} & -\frac{4141}{2880} & \frac{86551}{103680} & -\frac{24641}{129600} & & & & & \\ -\frac{52391}{129600} & \frac{134603}{51840} & -\frac{10991}{2160} & \frac{22583}{5184} & -\frac{46969}{25920} & \frac{30409}{86400} & & & & & \\ \frac{68603}{259200} & -\frac{4141}{2880} & \frac{22583}{5184} & -\frac{37967}{6480} & \frac{53369}{17280} & -\frac{54899}{129600} & \frac{1}{90} & & & & \\ -\frac{2351}{14400} & \frac{86551}{103680} & -\frac{46969}{25920} & \frac{53369}{17280} & -\frac{2747}{810} & \frac{820271}{518400} & -\frac{3}{20} & \frac{1}{90} & & & \\ \frac{4207}{103680} & -\frac{24641}{129600} & \frac{30409}{86400} & -\frac{54899}{129600} & \frac{820271}{518400} & -\frac{49}{18} & \frac{3}{2} & -\frac{3}{20} & \frac{1}{90} & & \\ & & & \frac{1}{90} & -\frac{3}{20} & \frac{3}{2} & -\frac{49}{18} & \frac{3}{2} & \frac{3}{20} & \frac{1}{90} & \\ & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \tag{A.11}$$

$$D_6^{(2)} = \frac{1}{X^2}
\begin{pmatrix}
\frac{114170}{40947} & -\frac{438107}{54596} & \frac{336409}{40947} & -\frac{276997}{81894} & \frac{3747}{13649} & \frac{21035}{163788} & & & & & \\
\frac{6173}{5860} & -\frac{2066}{879} & \frac{3283}{1758} & -\frac{303}{293} & \frac{2111}{3516} & -\frac{601}{4395} & & & & & \\
-\frac{52391}{81330} & \frac{134603}{32532} & -\frac{21982}{2711} & \frac{112915}{16266} & -\frac{46969}{16266} & \frac{30409}{54220} & & & & & \\
\frac{68603}{321540} & -\frac{12423}{10718} & \frac{112915}{32154} & -\frac{75934}{16077} & \frac{53369}{21436} & -\frac{54899}{160770} & \frac{48}{5359} & & & & \\
-\frac{7053}{39385} & \frac{86551}{94524} & -\frac{46969}{23631} & \frac{53369}{15754} & -\frac{87904}{23631} & \frac{820271}{472620} & -\frac{1296}{7877} & \frac{96}{7877} & & & \\
\frac{21035}{525612} & -\frac{24641}{131403} & \frac{30409}{87602} & -\frac{54899}{131403} & \frac{820271}{525612} & -\frac{117600}{43801} & \frac{64800}{43801} & -\frac{6480}{43801} & \frac{480}{43801} & & \\
 & & & \frac{1}{90} & -\frac{3}{20} & \frac{3}{2} & -\frac{49}{18} & \frac{3}{2} & \frac{3}{20} & \frac{1}{90} & \\
 & & & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots
\end{pmatrix}
\tag{A.12}$$

# Appendix B

# Questionnaire results

Sec. 7.2.6 discusses the subjective assessment of the measured and simulated IRs convolved with two anechoic recordings: a drum loop and a speech sample. Though summarised in Sec. 7.2.6, in this appendix the full responses have been listed.

**Note:** though in the listening test the order of the IRs had been randomised and had been given letters 'A', 'B', and 'C', for readability of the results in this appendix, all A, B, and Cs have been replaced in such a way that:

- A is the Measurement IR;
- B is the FD2 IR;
- C is the FD4 IR.

The order of the answers for each receiver has been randomised in order to avoid subject identification.

## R1

- B sounds boomier and has a resonant peak in it which you can hear in the kick drum. Also with the drum files I felt that A sounded a little less present/weaker, but hard to quantify, perhaps because it sounds slightly bright to my ears. A and B speech sound very similar to me.
- A feels like it has more air around it. Sounds nicer than B. Feels like it has more definition in the tail. B s more boomy in the lowmid. C sounds dryer than the other 2.

Much cleaner sound. Also lot less low end.

- A sounds boomy, similar length tail to B, B is more boxy low end. C is thinner with less low end and cluttered mids and shorter tail than B.
- Drums in B is bassier than A, like hollow resonant bass. C is much thinner and more mid-frequencies based than B. In the speech, cant hear much noticeable diff between A and B. C sounds kind of further away and more boxed in than B, also shorter in reverb tail, like were in a different space.
- The drum in A has a longer reverb tail than B. A has a really low frequency boost (like 100Hz) over B. B has a low mids frequency resonance (like 500 Hz?) over A. C sounds like a smaller room. B has much more low end (500Hz shelf?). C sounds tighter, though I can pinpoint why. Maybe more early reflections? C sounds brighter (more HF). C has more air/more space. Speech: C is more legible than B. C sounds dryer. Both sound like similar length reverbs C is maybe slightly shorter. A very similar to B.
- A sounds brighter to me in comparison to B, slightly noticeable on the more sibilant sounds in the speech files. C sounds like there's many resonant peaks in it. Room modes more prevalent? C has more early reflections or are more obvious? C sounds like it has less low end in the tail compared with B.
- There is some very low frequency resonance on the kick drum kits in B which isnt there in the other 2 samples. C feels like the drums are further away and in a slightly bigger room. C is missing the high and high-mid details that are in A, particularly noticeable on the cymbal crashes. For the speech, A and B are both very wet. I cant tell the difference

## R2

- C sounds almost distorted on the transients, or maybe like a bit crusher. Something weird going on here. C the cymbals sound fizzy. Maybe distorted? C has much less low end relative to the rest of the drums. Snare is less prominent in B. Both have a lot of bottom end (low frequencies). The low frequencies on B are more wooly. A sounds more compressed. Ie the drums are both wetter reverb wise, and more present (closer).
- C sound the most metallic of the three. Sounds closer than the other two. A sounds

the biggest of the three. Seems to have a much longer tail. B Sounds the most muffled of the three and sounds like its from next door rather than in the same room as me.

- C sounds closer than in B. B sounds duller than C. Tail feels longer in C. Reverb on A sounds brighter. Tail feels slightly longer on A. B has more consistent sounding resonants in it, in that I mean they feel more spread out evenly over the spectrum while in A you can clearly hear that one resnance sticks out more

- C has much more high end than B and a lot less low end, sounds more natural or real not quite as nice as A due to lack of low end. A has similar low end but more high freq B has higher absorption in highs and mids and sounds dull and boomy, A is nice

- Drums: C is much closer than B. B has much less top end than A, seems further away, duller. Speech: C sounds cleaner than B. B has more noticeable low freq rumble than A, though A also has it too.

- C sounds least obstructed. B sounds like someones taken a low pass filter to it. Missing a lot of mid/high frequencies. The speech in C has less reverb (B is wetter). C is closer. B has more low frequencies its more wooly and less intelligible. I cant really tell the difference between A and B. B sounds slightly purer, but I cant say why? More artefacts in A?

- C sounds like the drums are closer, but also sound a bit scooped out in the high mids. Speech A and B sound very similar to me, although here these two sound a bit obstructed.

## R3

- B drums is more boomy. C sounds thinner. I cant tell the difference between A and B speech. They are very very close. Not quite the same, but very close. - Reverb on A sounds brighter. Tail feels slightly longer on A. C sounds closer than in B. B sounds duller than C. Tail feels longer in C. Speech in B sounds less present in the low end than in C

- B is more boomy. More low frequencies relative to the rest. A has more presence (3-5kHz?). B sounds dryer (less reverb) than C, which sounds like a bathroom with a long corridor

- B is super subby in comparison to A. C sounds low fi but has more mids than B, which is mega sub Gonzales. Cant hear much discernible difference between A and B though. C sounds a bit more projected than B but a little nastily resonant in the mid frequencies.
- A seems to be the biggest of the three. It has more of a slap delay than the others. B sounds very dry compared to the other two. Sounds the warmest. C seems maybe as big as A but with a cleaner tail than the others.
- A is boomy but not nearly as much as B which is just nearly all low end B is slightly longer Rt60? C is way more middy than B and sounds like a hallway more natural but again not as nice to listen to as A because of pronounced cluttered mids
- Drums: B sounds like its much further away, quite low passed. A and C sound very similar to me, although A feels like It has more definition in the low end, feels cleaner. Speech: C feels like its in a bigger space and further away that the others. A and B sound very similar to me.

## R4

- Drums: C sounds distant and lower than A, sub stands out even more. Low-mid frequencies are a little more pronounced in A than B. Speech: C sounds occluded compared to B. A and B again sound pretty similar to me
- The main kick drum is more present in B. C sounds more wooly. The higher pitched kick drum is more resonant in B than in A. It catches the reverb more. C speech sounds like it has less reverb and more mid-high freq/is closer than B. A and B sound the same pretty much
- C sounds like the mid has been scooped out of it in comparison. A-B sound very close. Slightly more mid in A
- C is nasal bandpassed sound with lack of high end and feels quite resonant at certain frequencies in the mids does not sound as nice to listen to as B. A and B feel very similar to me here not much in it both sound nice IMO, B a bit more boomy so not quite as nice as A.
- C mostly sounds like reverb tail rather than much early reflection. Feels cleaner on the Speech sample. But perhaps not as accurate/real If that make sense. A sounds

the biggest of the three. B Sounds like it has much more early reflection. With speech samples A and B sounds very similar. But with drum samples the difference it much more noticeable.

## R5

- A sounds the biggest of the three. B sounds the most metallic of the three. C Sounds mostly like reverb tail and no to none early reflection. Sounds the most pleasing of the three.
- The snare drum on B drums sounds thinner. Theyre close. C is much closer/has less reverb/is drier. A and B speech sound the same. A has slightly more mids? C sounds closer/is a dryer mix.
- Drums in B sound further away than in A, wetter reverb, and B feels richer in hi freqs. C feels a bit more balanced than A and B. B speech sounds nice, A sounds a bit boxy. B and C speech are pretty similar.
- A and B again are very similar to me here I prefer B as it has slight more clarity in mids. C has more low end than B which is a little overpowering compared to B
- Drums: A has a little more hi freq in the tails. C is overly compressed, distant, but shorter reverb somehow than B. Speech: Theres probably differences between A and B but they once again seem negligible to me. C is more present than B.

## R6

- A is waaay muffled with complete lack of high end compared to B and sounds unnatural. C here is really nice IMO maybe a touch to much low end but otherwise sounds excellent really nice clarity across the spectrum with nice high absorption roll off. B is boomy and muffled in comparison
- A sounds more muffled than the others. B sounds the most metallic of the three. Has more early reflection. Medium distance to the source. C sounds like the source is much closer in this and makes the reverb tail sound and feel better.

- Drums: A is bassier than B, B has a nasty streak. C sounds very electronic, tight, and close in comparison to B. Speech: B is more pronounced/present than A. C is much less reverberant and much closer than B.
- A has almost no hi freqs at all, super noticeable in drums. C is much more broadband, B still a bit low-passed. Speech is most intelligible in A though cause its driest. C again more balanced, B less highs.
- A drums is tighter, but looses the upper mids. B looses the kick drum, loses the upper mids, but gains high frequencies. C is massively drier, you can actually hear the dry kick(!). A speech is slightly more muffled

## R7

- Drums: B sounds dull and dry in comparison to A. C sounds overly reverberant in comparison to B, but has a better broadband freq spread. Speech: B sounds dull in comparison to A, muffled. C sounds brighter and less muffled than B.
- B drums is much drier and muffled. A has more high frequencies. C is more wooly. B is drier. Speech, B has less high frequencies/is more muffled, A is wetter. C is much wetter. B has less high frequencies
- A is really boomy with no top end/mids but somehow not as boomy/muffled as B which is crazy boomy/muffled. C is nicer and has mire top end but seems to lack lower mids and sounds a bit scoopy
- B sounds the most muffled of the three. C sounds the most balanced of the three. A sounds the biggest of the three. Sounds mostly like reverb tail. More prominent in the mids.
- B sounds very muffled. C is super low-passed/wooly. A is a still a bit heavy on high-lows+low-mids but definitely nicest. Speech on C also very low-passed. B is better. A speech sounds a bit boomy.

# Appendix C

# Guide to accompanying material

The following is a guide through the appended material accompanying this thesis.

**PhD Thesis:** The PDF version of this thesis.

**Higher-order stencils:** This folder contains the IRs generated as part of the study presented in Chapter 5. For each of the three rooms (small, medium, and large), the IRs for the $2^{\mathrm{nd}}$- up to $16^{\mathrm{th}}$-order accurate stencils as well as the IWB stencil are included. The same counts for the test case of 16x16x16 nodes. All IRs have been saved as 32-bit floating-point WAV files.

**Shoebox room:** This folder contains all IRs presented in Sec. 7.1, that is to say, one IR for every source-receiver pair and method, for 10 different reflection values. Automatically generated decay graphs and FFT plots have also been included for each source-receiver pair and reflection value. Lastly, the Matlab code has been included that generated the aforementioned IRs. Again, all IRs have been saved as 32-bit floating-point WAV files.

**NCEM case study:** This folder contains the files related to the case study about the National Centre for Early Music in Sec. 7.2. The folder **NCEM Geometry** contains the Blender file of the location's geometry, which has been developped with Blender version 2.76. It also contains the same geometry as OBJ and MTL file, a standard 3D graphics extension that most CAD programmes are able to open. In **Sound files**, one can find the anechoic drum and speech sample used in aforementioned study, and the IRs of all seven receiver positions convolved with these samples, for each method (measurements, $2^{\mathrm{nd}}$-order accurate

FDTD, and 4th-order accurate FDTD).

# Symbols and abbreviations

| | |
|---|---|
| $\mathcal{B}$ | Valid bandwidth |
| $c$ | Speed of sound |
| $C$ | Clarity (ISO parameter) |
| $E$ | Energy |
| $f$ | Frequency |
| $F$ | Force |
| $G$ | Sound strength (ISO parameter) |
| $h(t)$ | Transfer function (time domain) |
| $H(\omega)$ | Transfer function (frequency domain) |
| $I$ | Sound intensity (power per unit area) |
| $j$ | Imaginary unit ($j^2 = -1$) |
| $k$ | Wave number |
| $K$ | Bulk modulus |
| $L(t)$ | Time-reversed integrated squared impulse response |
| $\mathbf{n}$ | Surface normal |
| $O(g(n))$ | Asymptotically upper bound to $g(n)$ |
| $p$ | Pressure |
| $p_0$ | Standard air pressure (101.325 kPa) |
| $RT_{60}$ | Reverberation time |
| $T_{30,45,60}$ | Reverberation time, estimated by the decay of the index in dB |
| $\mathbf{r}$ | Coordinate vector of space in higher dimension (in spherical coordinates) |
| $s(t)$ | Unit step function |
| $t$ | Time |
| $T$ | Absolute temperature |
| $T_c$ | Centre time/Centre of gravity |

| | |
|---|---|
| $u$ | Particle displacement |
| $V$ | Volume |
| $\mathbf{x}$ | Coordinate vector of space in higher dimension (in Cartesian coordinates) |
| $x$, $y$, $z$ | Cartesian coordinates of space (right/left, up/down, front/back) |
| $Z$ | Acoustic impedance |
| $\alpha$ | Absorption coefficient |
| $\delta(t)$ | Dirac delta function |
| $\zeta$ | Normalised acoustic impedance |
| $\eta$ | Humidity |
| $\Theta(g(n))$ | Asymptotically tight bound to $g(n)$ |
| $\kappa$ | Tension |
| $\lambda$ | Wave length |
| $L_o(\mathbf{r}, \boldsymbol{\omega}_o, t)$ | Radiosity intensity for outgoing angle, at some position and time |
| $\mu$ | Mass density |
| $\rho$ | Density |
| $\sigma$ | Scattering coefficient |
| $\phi$, $\theta$, $r$ | Spherical coordinates of space (azimuth, elevation, radius) |
| $\Phi(\mathbf{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o)$ | Bidirectional reflection distribution function (BRDF) |
| $\xi$ | Normalised acoustic admittance |
| $\Xi$ | Acoustic admittance |
| $\boldsymbol{\omega}_{i,o}$ | Angle of incidence and reflection, respectively |
| $\omega$ | Radial frequency |
| $\Omega(g(n))$ | Asymptotically lower bound to $g(n)$ |
| $\nabla$ | Nabla operator (spatial derivative) |
| $\nabla^2$ | Laplacian operator (second spatial derivative) |
| ARD | Boundary element method |
| BEM | Boundary element method |
| BRDF | Bidirectional reflection distribution function |
| CAD | Computer-aided design |
| CTPN | Compute time per node |
| CUDA | Compute unified device architecture |

| | |
|---|---|
| DCT | Discrete cosine transform |
| DWG | Digital waveguide |
| EDT | Early decay time (ISO parameter) |
| FDTD | Finite difference time domain |
| FFT | Fast Fourier transform |
| fps | Frames per second |
| FV | Finite volume |
| FVTD | Finite volume time domain |
| GKS | Gustafsson, Kreiss, and Sundström (theory) |
| IACC | Interaural cross correlation |
| ISO | International Organisation for Standardisation |
| IWB | Interpolated wideband |
| LF | Lateral fraction (ISO parameter) |
| LTI | Linear time-invariant |
| RIR | Room impulse response |
| SAT | Simultaneous Approximation Terms |
| SBP | Summation-by-parts |
| SRL | Standard rectilinear |
| VS | Virtual source |

# Bibliography

[1] Saul Abarbanel, Adi Ditkowski, and Amir Yefet. Bounded error schemes for the wave equation on complex domains. *Journal of Scientific Computing*, 26(1):67–81, January 2006.

[2] Marc Aird, Joel M. Laird, and John ffitch. Modelling a drum by interfacing 2-D and 3-D waveguide meshes. In *Proceedings of the International Computer Music Association Conference*, 2000.

[3] Ashraf Ali and Charles Rajakumar. *The Boundary Element Method: Applications in Sound and Vibration*. Taylor and Francis, 2004.

[4] Jont B. Allen and David A. Berkley. Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65(4):943–950, 1979.

[5] John C. Allred and Albert Newhouse. Applications of the Monte Carlo method to architectural acoustics. *Journal of the Acoustical Society of America*, 30:1–3, 1958.

[6] A. L. Álvarez, T. A. Zamarreño, S. T. Giròn, and M. S. Galindo. Virtual acoustics of the cathedral of Malaga (Spain). In *Proceedings of Forum Acusticum 2011*, 2011.

[7] John Amanatides. Ray tracing with cones. *SIGGRAPH Computer Graphics*, 18(3):129–135, January 1984.

[8] F. Antonacci, A. Sarti, and S. Tubaro. Two-dimensional beam tracing from visibility diagrams for real-time acoustic rendering. *EURASIP Journal on Advanced Signal Processing*, 2010:8.1–8.18, February 2010.

[9] Fabio Antonacci, M. Foco, Augusto Sarti, and Stefano Tubaro. Fast tracing of acoustic beams and paths through visibility lookup. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(4):812–824, 2008.

[10] Arthur Appel. Some techniques for shading machine renderings of solids. In *Proceedings of the AFIPS Spring Joint Computer Conference*, pages 37–45, 1968.

[11] H. E. Bass, L. C. Sutherland, J. Piercy, and L. Evans. Absorption of sound by the atmosphere. *Physical acoustics: Principles and methods*, 17:145–232, 1984.

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

[13] H. E. Bass, L. C. Sutherland, A. J. Zuckerwar, D.T. Blackstock, and D. M. Hester. Atmospheric absorption of sound: Further developments. *Journal of the Acoustical Society of America*, 97(1):680, 1995.

[14] A. E. Bate and M. E. Pillow. Mean free path of sound in an auditorium. *Proceedings of the Physical Society*, 59:535–541, 1947.

[15] Mark J. Beeson and Damian T. Murphy. RoomWeaver: a digital waveguide mesh based room acoustics research tool. In *Proceedings of the 7th International Conference on Digital Audio Effects (DAFX-04)*, October 2004.

[16] U. Berardi. Simulation of acoustical parameters in rectangular churches. In *Journal of Building Performance Simulation*, volume 6, 2013.

[17] A. J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *Journal of the Acoustical Society of America*, 93(5):2764–2778, 1993.

[18] Stefan Bilbao. *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics.* John Wiley & Sons, October 2009.

[19] Stefan Bilbao. Modeling of complex geometries and boundary conditions in finite difference/finite volume time domain room acoustics simulation. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(7):1524–1533, 2013.

[20] Stefan Bilbao and Brian Hamilton. Finite volume modeling of viscothermal losses and frequency-dependent boundaries in room acoustics simulations. In *Proceedings of the 60th International AES Conference on Dereverberation and Reverberation of Audio, Music, and Speech*, 2016.

[21] Stefan Bilbao, Brian Hamilton, Jonathan Botts, and Lauri Savioja. Finite volume time domain room acoustics simulation under general impedance boundary conditions. *IEEE Transactions on Audio, Speech, and Language Processing*, 2016.

[22] 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, 1957.

[23] D.T. Blackstock. *Fundamentals of physical acoustics.* Wiley-Interscience. John Wiley & Sons, 2000.

[24] Blender Foundation. Blender 2.69. [Computer programme] Retrieved from: `www.blender.org`, 2014. Last accessed: May 24, 2014.

[25] Matthew G. Blevins, Adam T. Buck, Zhao Peng, and Lily M. Wang. Quantifying the just noticeable difference of reverberation time with band-limited noise centered around 1000 hz using a transformed up-down adaptive method. In *Proceedings of the International Symposium on Room Acoustics*, 2013.

[26] Jeffrey Borish. Extension of the image model to arbitrary polyhedra. *Journal of the Acoustical Society of America*, 75(6):1827–1836, 1984.

[27] Ingolf Bork. A comparison of room simulation software - the 2nd round robin on room acoustical computer simulation. *Acta Acustica*, 86:943–956, 2000.

[28] Ingolf Bork. Simulation and measurement of auditorium acoustics – the round robins on room acoustical simulation. In *Proceedings of the Institute of Acoustics*, 2002.

[29] Ingolf Bork. Report on the 3rd round robin on room acoustical computer simulation – part ii: Calculations. *Acta Acustica united with Acustica*, 91(4), 2005.

[30] Max Born. *Optik: Ein Lehrbuch der elektromagnetischen Lichttheorie*. Springer, 1933.

[31] Dick Botteldooren. Acoustical finite-difference time-domain simulation in a quasi-Cartesian grid. *Journal of the Acoustical Society of America*, 95(5):2313–2319, 1994.

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

[33] Jonathan Botts and Lauri Savioja. Spectral and pseudospectral properties of finite difference models used in audio and room acoustics. *IEEE Transactions on Audio, Speech, and Language Processing*, 22(9):1403–1412, September 2014.

[34] Mack Breazeale and Michael McPherson. *Springer Handbook of Acoustics*, chapter Physical and Nonlinear Acoustics, pages 236–238. Springer New York, 2007.

[35] Guilherme R. Campos and David M. Howard. On the computational efficiency of different waveguide mesh topologies for room acoustic simulation. *IEEE Transactions on Speech and Audio Processing*, 13(5), 2005.

[36] Andreas C. Cangellaris and Diana B. Wright. Analysis of the numerical error caused by the stair-stepped approximation of a conducting boundary in FDTD simulations of electromagnetic phenomena. *IEEE Transactions on Antennas and Propagation*, 39(10):1518–1525, 1991.

[37] Mark H. Carpenter, David Gottlieb, and Saul Abarbanel. The stability of numerical boundary treatments for compact high-order finite-difference schemes. *Journal of Computational Physics*, 108(2):272–295, 10 1993.

[38] Mark H. Carpenter, David Gottlieb, and Saul Abarbanel. Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes. *Journal of Computational Physics*, 111(2):220–236, April 1994.

[39] M. M. Carroll and R. N. Miles. Steady-state sound in an enclosure with diffusely reflecting boundary. *Journal of the Acoustical Society of America*, 64(5):1424–1428, 1978.

[40] Zuofu Cheng. *Design of a real-time GPU accelerated acoustic simulation engine for interactive applications*. PhD thesis, University of Illinois at Urbana-Champaign, 2014.

[41] K. Chourmouziadou and J. Kang. Acoustic evolution of ancient greek and roman theatres. *Applied Acoustics*, 69(6):514–529, 2008.

[42] Melvyn Ciment and Stephen H. Leventhal. Higher order compact implicit schemes for the wave equation. *Mathematics of Computation*, 29(132):985–994, 1975.

[43] J.D. Cooper, A. Valavanis, Z. Ikonić, P. Harrison, and J.E. Cunningham. Stable perfectly-matched-layer boundary conditions for finite-difference time-domain simulation of acoustic waves in piezoelectric crystals. *Journal of Computational Physics*, 253:239–246, 2013.

[44] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to algorithms*. MIT Press, Second edition, 2005.

[45] Trevor J. Cox and Peter d'Antonio. *Acoustic Absorbers and Diffusers: Theory, design and application*. Spon Press, 2004.

[46] Norm Dadoun, David G. Kirkpatrick, and John P. Walsh. The geometry of beam tracing. In *Proceedings of the first annual symposium on computational geometry*, SCG '85, pages 55–61, New York, NY, USA, 1985. ACM.

[47] Bengt-Inge 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, August 1996.

[48] Bengt-Inge Dalenbäck. Verification of prediction based on randomized tail-corrected cone-tracing and array modeling. In *137th Convention of the Acoustical Society of America*, 1999.

[49] Bengt-Inge Dalenbäck, Adrian James, and Amber Naqvi. Computer modelling with CATT-Acoustic theory and practice of diffuse reflection and array modelling. In *Proceedings of the Institute of Acoustics 17th Conference on Reproduced Sound*, 2001.

[50] David C. del Rey Fernandez, Jason Hicken, and David W. Zingga. Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations. *Computers & Fluids*, 95:171–196, May 2014.

[51] Peter Deuflhard and Andreas Hohmann. *Numerical Analysis in Modern Scientific Computing*, volume 43. Springer-Verlag New York, Second edition, 2003.

[52] Adi Ditkowski. *Bounded-error finite difference schemes for initial boundary value problems on complex domains.* PhD thesis, Tel-Aviv University, 1997.

[53] I. A. Drumm and Y. W. Lam. The adaptive beam-tracing algorithm. *Journal of the Acoustical Society of America*, 107(3):1405–1412, 2000.

[54] Scott A. Van Duyne and Julius O. Smith. Physical modeling with the 2D digital waveguide mesh. In *Proceedings of the International Computer Music Conference*, pages 40–47, 1993.

[55] Scott A. Van Duyne and Julius O. Smith. The 3D tetrahedral digital waveguide mesh with musical applications. In *Proceedings of the International Computer Music Conference*, pages 234 –237, oct 1996.

[56] Carl F. Eyring. Reverberation time in "dead" rooms. *Journal of the Acoustical Society of America*, 1(2A):217–241, 1930.

[57] Angelo Farina. Pyramid tracing vs. ray tracing for the simulation of sound propagation in large rooms. In *Proceedings of the International Conference on Computer Acoustics and its Environmental Applications*, 1995.

[58] Angelo Farina. RAMSETE – a new pyramid tracer for medium and large scale acoustic problems. In *Proceedings of Euro-Noise 1995*, March 1995.

[59] Angelo Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. *Proceedings of the 108th AES Convention*, 2000.

[60] Angelo Farina. Validation of the pyramid tracing algorithm for sound propagation outdoors: comparison with experimental measurements and with the iso-dis 9613 standards. *Advances in Engineering Software*, 31(4):241 – 250, 2000.

[61] Angelo Farina. Advancements in impulse response measurements by sine sweeps. In *Proceedings of the 122nd Convention of the Audio Engineering Society*, 2007.

[62] Angelo Farina. Aurora plug-ins. Online source: `http://www.aurora-plugins.com/`, 2007. Last accessed: May 27, 2014.

[63] Andrea Farnetani, Nicola Prodi, and Roberto Pompoli. On the acoustics of ancient greek and roman theaters. *Journal of the Acoustical Society of America*, 124(3):1557–1567, 2008.

[64] Neville H. Fletcher and Thomas D. Rossing. *The physics of musical instruments*. Springer, Second edition, 2008.

[65] Marco Foco, Pietro Polotti, Augusto Sarti, and Stefano Tubaro. Sound spatialization based on fast beam tracing in the dual space. In *Proceedings of the Sixth International Conference on Digital Audio Effects (DAFx-03)*, 2003.

[66] International Organization for Standardization. Acoustics - measurement of the reverberation time of rooms with reference to other acoustical parameters, 1997.

[67] Aglaïa Foteinou. *Perception of Objective Parameter Variations in Virtual Acoustic Spaces*. PhD thesis, University of York, 2013.

[68] D. Funaro and D. Gottlieb. A new method of imposing boundary conditions in pseudospectral approximations of hyperbolic equations*. *Mathematics of Computation*, 51(184):599–613, 1988.

[69] Thomas Funkhouser, Ingrid Carlbom, Gary Elko, Gopal Pingali, Mohan Sondhi, and Jim West. A beam tracing approach to acoustic modeling for interactive virtual environments. In *Proceedings of SIGGRAPH 98*, pages 21–32, July 1998.

[70] Thomas Funkhouser, Nicolas Tsingos, Ingrid Carlbom, Gary Elko, Mohan Sondhi, James E. West, Gopal Pingali, Patrick Min, and Addy Ngan. A beam tracing method for interactive architectural acoustics. *Journal of the Acoustical Society of America*, 115(2):739–756, February 2004.

[71] Thomas A. Funkhouser, Patrick Min, and Ingrid Carlbom. Real-time acoustic modeling for distributed virtual environments. In *Proceedings of SIGGRAPH 1999*, pages 365–374, August 1999.

[72] Djamchid Ghazanfarpour and Jean-Marc Hasenfratz. A beam tracing with precise antialiasing for polyhedral scenes. *Computer Graphics*, 22(1):103–115, 1998.

[73] B. M. Gibbs and D. K. Jones. A simple image method for calculating the distribution of sound pressure levels within an enclosure. *Acustica*, 26:24–32, 1972.

[74] Naga K. Govindaraju, Brandon Lloyd, Yuri Dotsenko, Burton Smith, and John Manferdelli. High performance discrete fourier transforms on graphics processors. In *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*, 2008.

[75] Bertil Gustafsson. *High Order Difference Methods for Time Dependent PDE*. Springer Berlin Heidelberg, 2008.

[76] Bertil Gustafsson, Heinz-Otto Kreiss, and Arne Sundström. Stability theory of difference approximations for mixed initial boundary value problems. ii. *Mathematics of Computation*, 26(119):649–686, 07 1972.

[77] T. Ha-Duong, B. Ludwig, and I. Terrasse. A Galerkin BEM for transient acoustic scattering by an absorbing obstacle. *International Journal for Numerical Methods in Engineering*, 57(13):1845–1882, 2003.

[78] Hüyesin Hacıhabiboğlu, Banu Günel, and Zoran Cvetković. Simulation of directional microphones in digital waveguide mesh-based models of room acoustics. *IEEE Transactions on Acoustics, Speech, and Language Processing*, 16(2):213–223, February 2010.

[79] Hüyesin Hacıhabiboğlu, Banu Günel, and Ahmet M. Kondoz. Time-domain simulation of directive sources in 3-d digital waveguide mesh-based acoustical models. *IEEE Transactions on Acoustics, Speech, and Language Processing*, 16:934–946, 2008.

[80] Jon Häggblad and Olof Runborg. Accuracy of staircase approximations in finite-difference methods for wave propagation. *Numerische Mathematik*, 128(4):741–771, 2014.

[81] Brian Hamilton and Stefan Bilbao. Fourth-order and optimised finite difference schemes for the 2-D wave equation. In *Proceedings of the 16th Conference on Digital Audio Effects (DAFx-13)*, 2013. To be published.

[82] Brian Hamilton, Craig J. Webb, Alan Gray, and Stefan Bilbao. Large stencil operators for GPU-based 3-D acoustic simulations. In *Proceedings of the 18th international Conference on Digital Audio Effects*, 2015.

[83] Isaac Harari and Thomas J. R. Hughes. A cost comparison of boundary element and finite element methods for problems of time-harmonic acoustics. *Computer Methods in Applied Mechanics and Engineering*, 97:77–102, 1992.

[84] Jonathan A. Hargreaves. *Time Domain Boundary Element Method for Room Acoustics*. PhD thesis, University of Salford, 2007.

[85] W. M. Hartmann. Localization of sound in rooms. *Journal of the Acoustical Society of America*, 74(5):1380–1391, 1983.

[86] Paul S. Heckbert and Pat Hanrahan. Beam tracing polygonal objects. In *Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '84, pages 119–127, New York, NY, USA, 1984. ACM.

[87] Paul S. Heckbert and Pat Hanrahan. Introduction to beam tracing. In *International Conference on Engineering and Computer Graphics*, pages 286–289, August 1984.

[88] Murray Hodgson. Evidence of diffuse surface reflections in rooms. *Journal of the Acoustical Society of America*, 89(2):765–771, 1990.

[89] Maarten Hornikx, Manfred Kaltenbacher, and Steffen Marburg. A platform for benchmark cases in computational acoustics. *Acta Acustica united with Acustica*, 101:811–820, 2015.

[90] David M. Howard and Jamie A. S. Angus. *Acoustics and Phychoacoustics.* Focal Press, Fourth edition, 2009.

[91] Mayer Humi. Factorisation of separable partial differential equations. *Journal of Physics A: Mathematical and General*, 20(14):4577–4585, 1987.

[92] Thiago Ize, Ingo Wald, and Steven G. Parker. Ray tracing with the bsp tree. In *IEEE Symposium on Interactive Ray Tracing 2008*, pages 159 –166, aug. 2008.

[93] Jian-Ming Jin. *The Finite Element Method in Electromagnetics.* John Wiley & Sons, second edition, 2002.

[94] James T. Kajiya. The rendering equation. *SIGGRAPH Computer Graphics*, 20(4):143–150, August 1986.

[95] Matti Karjalainen and Cumhur Erkut. Digital waveguides versus finite difference structures: Equivalence and mixed modeling. *Journal on Applied Signal Processing*, 7:978–989, 2004.

[96] Thomas Sporer Karlheinz Brandenburg, Sandra Brix. Wave field synthesis. In *Proceedings of the 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video*, 2009.

[97] Joseph B. Keller. Geometrical theory of diffraction. *Journal of the Optical Society of America*, 52(2):116–130, February 1962.

[98] J. F. C. Kingman. Mean free paths in a convex reflecting region. *Journal of Applied Probability*, 2(1):162–168, 1965.

[99] Lawrence E. Kinsler, Austin E. Frey, Alal B. Coppens, and James V. Sanders. *Fundamentals of Room Acoustics.* John Wiley & Sons, Fourth edition, 2000.

[100] Gustav R. Kirchhoff. Zur theorie der lichtstrahlen. *Annalen der Physik*, 254(4):663–695, 1883.

[101] Stephen Kirkup. *The boundary element method in acoustics: a development in Fortran.* Integral equation methods in engineering. Integrated Sound Software, 1998.

[102] 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.

[103] Konrad Kowalczyk. *Boundary and medium modelling using compact finite difference schemes in simulations of room acoustics for audio and architectural design applications.* PhD thesis, Queen's University Belfast, November 2008.

[104] Konrad Kowalczyk and Maarten van Walstijn. Formulation of locally reacting wall in finite difference modelling of acoustic spaces. In *Proceedings of the International Symposium on Room Acoustics*, September 2007.

[105] Konrad Kowalczyk and Maarten van Walstijn. Formulation of locally reacting surfaces in FDTD/K-DWM modelling of acoustic spaces. *Acta Acustica united with Acustica*, 94:891–906, 2008.

[106] Konrad Kowalczyk and Maarten van Walstijn. Modeling frequency-dependent boundaries as digital impedance filters in FDTD and K-DWM room acoustics simulations. *Journal of the Audio Engineering Society*, 56(7/8):569–583, 2008.

[107] Konrad Kowalczyk and Maarten van Walstijn. Room acoustics simulation using 3-D compact explicit FDTD schemes. *IEEE Transactions on Acoustics, Speech, and Language Processing*, 19(1):34–46, 2011.

[108] Heinz-Otto Kreiss. Stability theory for difference approximations of mixed initial boundary value problems. i. *Mathematics of Computation*, 22(104):703–714, October 1968.

[109] Heinz-Otto Kreiss and Godela Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. In *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 195–212, 1974.

[110] Heinz-Otto Kreiss and Lixin Wu. On the stability definition of difference approximations for the initial boundary value problem. *Applied Numerical Mathematics*, 12(1–3):213–227, 1993.

[111] U.R. Kristiansen, A. Krokstad, and T. Follestad. Extending the image method to higher-order reflections. *Applied Acoustics*, 38(2–4):195–206, 1993.

[112] A. Krokstad and S. Strøm. Acoustical design of the multi-purpose 'hjertnes' hall in sandefjord. *Applied Acoustics*, 12(1):45 – 63, 1979.

[113] A. Krokstad, S. Strøm, and S. Sørsdal. Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibration*, 8(1):118–125, 1968.

[114] A. Kulowski. Algorithmic representation of the ray tracing technique. *Applied Acoustics*, 18(6):449–469, 1985.

[115] Heinrich Kuttruff. Simulierte nachhallkurven in rechteckräumen mit diffusen schallfeld. *Acustica*, 25:333–342, 1971.

[116] Heinrich Kuttruff. A simple iteration scheme for the computation of decay constants in enclosures with diffusely reflecting boundaries. *Journal of the Acoustical Society of America*, 98(1):288–293, 1995.

[117] Heinrich Kuttruff. *Room Acoustics*. Spon Press, 2009.

[118] Joel M. Laird. *The Physical Modelling of Drums Using Digital Waveguides*. PhD thesis, University of Bristol, 2001.

[119] Y. W. Lam. A comparison of three diffuse reflection modeling methods used in room acoustics computer models. *Journal of the Acoustical Society of America*, 100(4):2181–2192, October 1996.

[120] Alain le Bot. A functional equation for the specular reflection of rays. *Journal of the Acoustical Society of America*, 112(4):1276–1297, 2002.

[121] Alain le Bot and Arnaud Bocquillet. Comparison of an integral equation on energy and the ray-tracing technique in room acoustics. *Journal of the Acoustical Society of America*, 108(4):1732–1740, 2000.

[122] Heewon Lee and Byung-Ho Lee. An efficient algorithm for the image model technique. *Applied Acoustics*, 24(2):87–115, 1988.

[123] Hilmar Lehnert. Systematic errors of the ray-tracing algorithm. *Applied Acoustics*, 38(2–4):207–221, 1993.

[124] Tobias Lentz, Dirk Schröder, Michael Vorländer, and Ingo Assenmacher. Virtual reality system with integrated sound field simulation and reproduction. *EURASIP Journal on Advances in Signal Processing*, 2007(1):1–20, 2007.

[125] Magne Skålevik. Schroeder frequency revisited. In *Proceedings of Forum Acusticum 2011*, 2011.

[126] T. Lewers. A combined beam tracing and radiant exchange computer model of room acoustics. *Applied Acoustics*, 38(2–4):161–178, 1993.

[127] Martin Lisa, Jens H. Rindel, and Claus L. Christensen. Predicting the acoustics of ancient open-air theaters: The importance of calculation methods and geometrical details. In *Joint Baltic-Nordic Acoustics Meeting*, 2004.

[128] José J. López, Diego Carnicero, Néstor Ferrando, and José Escolano. Parallelization of the finite-difference time-domain method for room acoustics modelling based on CUDA. *Mathematical and Computer Modelling*, 57:1822–1831, April 2013.

[129] Ken Mattsson, Frank Ham, and Gianluca Iaccarino. Stable and accurate wave-propagation in discontinuous media. *Journal of Computational Physics*, 227(19):8753–8767, 2008.

[130] Ken Mattsson, Frank Ham, and Gianluca Iaccarino. Stable boundary treatment for the wave equation on second-order form. *Journal of Scientific Computing*, 41(3):366–383, December 2009.

[131] Ken Mattsson and Jan Nordström. Summation by parts operators for finite difference approximations of second derivatives. *Journal of Computational Physics*, 199(2):503–540, 9 2004.

[132] Paulius Micikevicius. 3D finite difference computation on GPUs using CUDA. In *Proceedings of the Second Workshop on General Purpose Processing on Graphics Processing Units*, pages 79–84, 2009.

[133] Michael Monks, Byong Mok Oh, and Julie Dorsey. Acoustic simulation and visualization using a new unified beam tracing and image source approach. In *Proceedings of the Convention of the Audio Engineering Society*, November 1996.

[134] Philip M. Morse and K. Uno Ingard. *Theoretical Acoustics*. Princeton University Press, 1987.

[135] Jack Mullen, David M. Howard, and Damian T. Murphy. Waveguide physical modeling of vocal tract acoustics: flexible formant bandwidth control from increased model dimensionality. *IEEE Transactions on Acoustics, Speech, and Language Processing*, 14:964–971, 2006.

[136] Jack Mullen, David M. Howard, and Damian T. Murphy. Real-time dynamic articulations in the 2-d waveguide mesh vocal tract model. *IEEE Transactions on Audio, Speech, and Language Processing*, 15(2):577–585, February 2007.

[137] Damian T. Murphy. *Digital Waveguide Mesh Topologies in Room Acoustics Modelling*. PhD thesis, University of York, 2000.

[138] Damian T. Murphy, Antti Kelloniemi, Jack Mullen, and Simon Shelley. Acoustic modeling using the digital waveguide mesh. *IEEE Signal Processing Magazine*, 24(2):55–66, March 2007.

[139] National Centre For Early Music. NCEM History. Online source: `http://www.ncem.co.uk/`, 2014. Last accessed: May 27, 2014.

[140] Graham M. Naylor. Treatment of early and late reflections in a hybrid computer model for room acoustics. *Journal of the Acoustical Society of America*, 92(4):2345, 1992.

[141] Graham M. Naylor. Odeon – another hybrid room acoustical model. *Applied Acoustics*, 38:131–143, 1993.

[142] Fred E. Nicodemus. Directional reflectance and emissivity of an opaque surface. *Applied Optics*, 4(7):767–775, 1965.

[143] Eva-Marie Nosal. Room sound field prediction by acoustical radiosity. Master's thesis, University of Calgary, December 2002.

[144] Eva-Marie Nosal, Murray Hodgson, and Ian Ashdown. 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.

[145] NVIDIA. NVIDIA VRWorks$^{\text{TM}}$. Online source: `https://developer.nvidia.com/vrworks/`. Last accessed: August 9, 2016.

[146] Odeon. Ray Tracing — Odeon. Online source: `http://www.odeon.dk/ray-tracing`. Last accessed: July 16, 2013.

[147] Pelle Olsson. Summation by parts, projections, and stability. I. *Mathematics of Computation*, 64(221):1035–1065, 1995.

[148] Pelle Olsson. Summation by parts, projections, and stability II. *Mathematics of Computation*, 64(212):1473–1493, 1996.

[149] A. M. Ondet and J. L. Barbry. Model of sound propagation in fitted workshops using ray tracing. *Journal of the Acoustical Society of America*, 85(2):787–796, 1989.

[150] Open AIR—Anechoic audio database. Online source: `http://www.openairlib.net/anechoicdb`. Last accessed: July 29, 2016.

[151] World Health Organization. Hearing loss due to recreational exposure to loud sounds - a review, 2015. NLM classification: WV 270.

[152] Stephen Oxnard. Personal communication, December 2015.

[153] Stephen Oxnard, Dean O'Brian, Jelle van Mourik, and Damian T. Murphy. Frequency-dependent absorbing boundary implementations in 3d finite difference time domain room acoustics simulations. In *Proceedings of EuroNoise 2015*, 2015.

[154] Jukka Pätynen, Sakari Tervo, and Tapio Lokki. Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses. *Journal of the Acoustical Society of America*, 133(2):842–857, 2013.

[155] Bui Tuong Phong. Illumination for computer generated pictures. *Communications of the Association for Computing Machinery*, 18(6):311–317, 1975.

[156] Nikunj Raghuvanshi, Rahul Narain, and Ming C. Lin. Efficient and accurate sound propagation using adaptive rectangular decomposition. *IEEE Transactions on Visualization and Computer Graphics*, 15(5):789–801, 2009.

[157] Emeline Reboul, Alain le Bot, and Joël Perret-Liaudet. Radiative transfer equation for multiple diffraction. *Journal of the Acoustical Society of America*, 118:1326–1334, 2005.

[158] Y.S. Rickard, N.K. Georgieva, and Wei-Ping Huang. Application and optimization of PML ABC for the 3-D wave equation in the time domain. *IEEE Transactions on Antennas and Propagation*, 51(2):286–295, 2003.

[159] Sjoerd W. Rienstra and A. Hirschberg. *An Introduction to Acoustics*. Eindhoven University of Technology, November 2012. Available online at: `www.win.tue.nl/~sjoerdr/papers/boek.pdf` [Last accessed: October 7, 2015].

[160] Jens H. Rindel. Room acoustic prediction modelling. In *Encontro da Sociedade Brasileira de Acústica XXIII*, 2010.

[161] Wallace C. Sabine. *Collected papers on acoustics*. Harvard University Press, 1922.

[162] Shinichi Sakamoto. Phase-error analysis of high-order finite difference time domain scheme and its influence on calculation results of impulse response in closed sound field. *Acoustical Science and Technology*, 28(5):295–309, 2007.

[163] Rudolf Rabenstein Sascha Spors. Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis. In *Proceedings of the 120th Audio Engineering Society Convention*, 2006.

[164] Lauri Savioja. Real-time 3D finite-difference time-domain simulation of low- and mid-frequency room acoustics. In *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx-10)*, 2010.

[165] Lauri Savioja, Tapio Lokki, Jyri Huopaniemi, and Riitta Vaananen. Creating interactive virtual acoustic environments. *Journal of the Audio Engineering Society*, 47(9):675–705, 1999.

[166] Lauri Savioja, Timo J. Rinne, and Tapio Takala. Simulation of room acoustics with a 3-D finite difference mesh. In *Proceedings of the International Computer Music Conference*, pages 463–466, 1994.

[167] Lauri Savioja and Peter Svensson. Overview of geometrical room acoustic modeling techniques. *Journal of the Acoustical Society of America*, 138(2):708–730, 2015.

[168] Lauri Savioja and Vesa Välimäki. Reducing the dispersion error in the digital waveguide mesh using interpolation and frequency-warping techniques. *IEEE Transactions on Speech and Audio Processing*, 8(2):184 –194, mar 2000.

[169] Lauri Savioja and Vesa Välimäki. Interpolated 3-D digital waveguide mesh with frequency warping. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2001.

[170] Jonathan D. Schäffer, Bruno M. Fazenda, Jamie A. S. Angus, and Damian T. Murphy. A simple multiband approach for solving frequency dependent problems in numerical time domain methods. In *Proceedings of Forum Acusticum 2011*, pages 269–274, 2011.

[171] Carl Schissler and Dinesh Manocha. GSound: Interactive sound propagation for games. In *Proceedings of the 41st AES Conference: Audio for Games*, 2011.

[172] John B. Schneider, Christopher L. Wagner, and Omar M. Ramahi. Implementation of transparent sources in FDTD simulations. *IEEE Transactions on Antennas and Propagation*, 46(8):1159–1168, 1998.

[173] M. R. Schröder. Die statistischen parameter der frequenzkurven von groen räumen. *Acta Acustica united with Acustica*, 4(7):594–600, 1954.

[174] M. R. Schröder and K. K. Kuttruff. On frequency response curves 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.

[175] Manfred R. Schroeder. Natural sounding artificial reverberation. In *Proceedings of the 13th Audio Engineering Society Convention*, November 1961.

[176] Manfred R. Schroeder and Benjamin F. Logan. "colorless" artificial reverberation. *Journal of the Acoustical Society of America*, 9(3):192–197, 1961.

[177] H. P. Seraphim. Untersuchungen über die Unterschiedsschwelle exponentiellen Abklingens von Rauschbandimpulsen. *Acta Acustica united with Acustica*, 8(5):280–284, 1958.

[178] Richard P. Shaw. Diffraction of acoustic pulses by obstacles of arbitrary shape with a robin boundary condition. *Journal of the Acoustical Society of America*, 41(4):855–859, 1967.

[179] Simon Shelley and Damian T. Murphy. The modeling of diffuse boundaries in the 2-d digital waveguide mesh. *IEEE Transactions on Acoustics, Speech, and Language Processing*, 16(3):651–665, March 2008.

[180] Simon B. Shelley. *Diffuse Boundary Modelling in the Digital Waveguide Mesh*. PhD thesis, University of York, 2007.

[181] Jiaoying Shi, Aidong Zhang, Jose Encarnação, and Martin Göbel. A modified radiosity algorithm for integrated visual and auditory rendering. *Computers & Graphics*, 17(6):633–642, 1993.

[182] Samuel Siltanen and Tapio Lokki. Diffraction modelling in acoustic radiance transfer method. In *Proceedings of Acoustics'08 Paris*, 2008.

[183] Samuel Siltanen, Tapio Lokki, Sami Kiminki, and Lauri Savioja. The room acoustic rendering equation. *Journal of the Acoustical Society of America*, 122(3):1624–1635, 2007.

[184] Samuel Siltanen, Alexander P. Southern, and Lauri Savioja. Finite-difference time domain method source calibration for hybrid acoustics modeling. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 166–170, 2013.

[185] Julius O. Smith. A new approach to digital reverberation using closed waveguide networks. In *Proceedings of the International Computer Music Conference*, pages 47–53, Burnaby, B.C., Canada, 1985.

[186] Julius O. Smith. Music applications of digital waveguides. Technical report, Stanford University, 1987.

[187] Julius O. Smith. Physical modeling using digital waveguides. *Computer Music Journal*, 16(4):74–91, 1992.

[188] Julius O. Smith. *Handbook of signal processing in acoustics*, chapter Digital Waveguide Architectures for Virtual Musical Instruments, pages 399–417. Springer, 2009.

[189] Alexander P. Southern. *The Synthesis and Auralisation of Physically Modelled Soundfields*. PhD thesis, University of York, 2010.

[190] Alexander P. Southern and Damian T. Murphy. Methods for 2nd order spherical harmonic spatial encoding in digital waveguide mesh virtual acoustic simulations. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 203 –206, oct. 2007.

[191] Alexander P. Southern, Samuel Siltanen, Damian T. Murphy, and Lauri Savioja. Room impulse response synthesis and validation using a hybrid acoustic model. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(9):1940–1952, 2013.

[192] Alexander P. Southern, Samuel Siltanen, and Lauri Savioja. Spatial room impulse responses with a hybrid modeling method. In *Proceedings of the 130th Audio Engineering Society Convention*, 2011.

[193] Irene A. Stegun. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, chapter Legendre Functions, pages 332–341. National Bureau of Standards, 1972.

[194] Uwe M. Stephenson. Quantized pyramidal beam tracing - a new algorithm for room acoustics and noise immission prognosis. *Acta Acustica united with Acustica*, 82(3), 1996.

[195] Bo Strand. High-order finite difference methods for hyperbolic IBVP. In *Proceedings of the Third International Conference on Spectral and High Order Methods*, 1995.

[196] Brian Straughan. *The Energy Method, Stability, and Nonlinear Convection.* Springer Science & Business Media, 2013.

[197] Robert Strichartz. *A Guide to Distribution Theory and Fourier Transforms.* CRC Press, 1994.

[198] John C Strikwerda. Initial boundary value problems for the method of lines. *Journal of Computational Physics*, 34(1):94–107, 1 1980.

[199] John C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations.* Society for Industrial and Applied Mathematics, Second edition, 2004.

[200] Sakari Tervo, Jukka Pätynen, and Tapio Lokki. Spatial analysis of concert hall impulse responses. In *Proceedings of the International Symposium on Room Acoustics (ISRA) 2013*, 2013.

[201] Lonny L. Thompson. A review of finite-element methods for time-harmonic acoustics. *Journal of the Acoustical Society of America*, 119(3):1315–1330, 2006.

[202] Paul A. Tippler and Gene Mosca. *Physics for Scientists and Engineers.* W. H. Freeman and Company, 2008.

[203] Nicolas Tsingos. *Simulation de champs sonores de haute qualité pour des applications graphiques interactives.* PhD thesis, Université Joseph Fourier, Grenoble, 1998.

[204] Nicolas Tsingos, Thomas Funkhouser, Addy Ngan, and Ingrid Carlbom. Modeling acoustics in virtual environments using the uniform theory of diffraction. In *Proceedings of ACM SIGGRAPH 2001*, pages 545–552, August 2001.

[205] Dirk van Maercke. Simulation of sound fields in time and frequency domain using a geometrical model. In *Proceedings of the 12th International Congress on Acoustics*, volume 7, 1986.

[206] Dirk van Maercke and Jacques Martin. The prediction of echograms and impulse responses within the epidaure software. *Applied Acoustics*, 38(2-4):93–114, 1993.

[207] Jelle van Mourik and Damian T. Murphy. Geometric and wave-based acoustic modelling using Blender. In *Proceedings of the AES 49th International Conference, Audio for Games*, 2013.

[208] Jelle van Mourik and Damian T. Murphy. Explicit higher-order FDTD schemes for 3D room acoustic simulation. *IEEE Transactions on Acoustics, Speech and Language processing*, 22(12):2003–2011, 2014.

[209] Jelle van Mourik, Stephen Oxnard, Aglaïa Foteinou, and Damian T. Murphy. Hybrid acoustic modelling of historic spaces using blender. In *Proceedings of Forum Acusticum 2014*, 2014.

[210] Martijn Vercammen. *Sound concentration caused by curved surfaces*. PhD thesis, Technische Universiteit Eindhoven, 2012.

[211] Michael Vorländer. Ein strahlverfolgungsverfahren zur berechnung von schallfeldern in räumen. *Acustica*, 65:138–148, 1988.

[212] Michael Vorländer. Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *Journal of the Acoustical Society of America*, 86:172–178, 1989.

[213] Michael Vorländer. International round robin on room acoustical computer simulations. In *5th International Congress on Acoustics*, 1995.

[214] Michael Vorländer. *Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*. RWTH edition (Berlin. Print). Springer, 2008.

[215] Michael Vorländer and Eckard Mommertz. Definition and measurement of random-incidence scattering coefficients. *Applied Acoustics*, 60:187–199, 2000.

[216] Andrew Wareing and Murray 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, 2005.

[217] Alan Watt. *3D Computer Graphics*. Pearson, 2000.

[218] Craig J. Webb and Stefan Bilbao. Computing room acoustics with CUDA - 3D FDTD schemes with boundary losses and viscosity. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 317–320, 2011.

[219] Kevin Weiler and Peter Atherton. Hidden surface removal using polygon area sorting. In *Computer Graphics*, pages 214–222, 1977.

[220] Turner Whitted. An Improved Illumination Model for Shaded Display. In J. D. Foley, editor, *Graphics and Image Processing*, volume 23, pages 343–349. Bell Laboratories, 1980.

[221] George S. K. Wong. Speed of sound in standard air. *Journal of the Acoustical Society of America*, 79(5):1359–1366, 1986.

[222] Luiz C. Wrobel. *The Boundary Element Method—Applications in Thermo-Fluids & Acoustics*, volume 1. John Wiley & Sons, 2002.

[223] H. He Z. Meng, F. Zhao. The just noticeable difference of noise length and reverberation perception. In *Proceedings of the International Symposium on Communications and Information Technologies*, pages 418–421, 2006.

[224] O. C. Zienkiewicz and R. L. Taylor. *The Finite Element Method for Solid and Structural Mechanics*. Elsevier Butterworth-Heinemann, 2005.

[225] O. C. Zienkiewicz, R. L. Taylor, and P Nithiarasu. *The Finite Element Method for Fluid Dynamics*. Elsevier Butterworth-Heinemann, sixth edition, 2006.

[226] Allan J. Zuckerwar. *Handbook of the Speed of Sound in Real Gases*. Elsevier Science, 2002.

[227] Allan J. Zuckerwar and Roger W. Meredith. Low-frequency absorption of sound in air. *Journal of the Acoustical Society of America*, 78(3):946–955, 1985.