Efficient Hybrid Virtual Room Acoustic Modelling

Stephen Oxnard

Doctor of Philosophy

University Of York

Electronics

September 2016

Abstract

This thesis investigates approaches to virtual room acoustic modelling and auralisation in order to a develop hybrid modelling solution that is capable of efficient and accurate simulation of enclosed sound propagation. Emphasis is placed on the advantages and disadvantages of state of the art numerical and geometric acoustic modelling methods. Numerical methods have been shown to preserve important sound wave characteristics such as diffraction and room modes, and are considered more accurate for low frequency acoustic modelling than geometric techniques which fail to preserve such wave effects. However, the implementation of numerical acoustic models inherently requires large computational effort compared to more efficient geometric techniques such as ray-tracing. This is particularly problematic for simulations of large-scale 3D acoustic environments and for high spatio-temporal sampling rates. A novel acoustic modelling solution is presented, which seeks to circumvent the disadvantageous computational requirements of 3D numerical models while producing a suitable approximation to low frequency sound behaviour. This modelling technique incorporates multiple planar cross-sectional 2D Finite Difference schemes that are utilised in combination to synthesise low frequency wave propagation throughout the target acoustic field. In this way a subset of prominent low frequency sound wave characteristics may be emulated with low computational cost compared to 3D numerical schemes. These low-frequency results can then be combined with the high-frequency output from efficient geometric simulations to create a hybrid model providing accurate broadband results at a relatively low computational cost. Investigation of room impulse response rendering for a series of theoretic and real spaces demonstrates advantages of this new hybrid acoustic modelling technique over purely ray-based methods in terms of low frequency accuracy, and over 3D numerical methods in terms of computational efficiency. Conclusions are drawn from objective measurements obtained from simulation results of the virtual models produced. Results demonstrate the applicability of the novel hybrid approach to the enhancement of purely ray-based room impulse response rendering by which a more realistic representation of low frequency wave phenomena may be simulated in an efficient manner, improving the theoretical accuracy of objective and audible results. This study contributes towards research and design of high-speed, interactive virtual acoustic simulations appropriate for industrial and creative virtual reality applications.

Contents

List of Figures

List of Tables

For Compassion, Coffee and POWERMETAL...

Acknowledgements

I would like to extend my deepest thanks to Dr. Damian Murphy who has supported me the whole way through the PhD process. I am extremely grateful for his guidance and unwavering positivity. It has been a great experience to work with such a talented supervisor. My thanks also go to Dr. Helena Daffern who has been an excellent second supervisor to me over the past few years and to Dr. Gavin Kearney who, in his role as thesis advisor, has enthusiastically encouraged the development of this work.

My warmest thanks go to everyone at the Audio Lab who have made my PhD experience one to recall fondly.

Last, but not least, my heartfelt thanks to my family and my yaboh Pixie for their endless support and compassion - they are a patient bunch.

Declaration of Authorship

I, Stephen Oxnard, declare that this thesis titled, 'Efficient Hybrid Virtual Room Acoustic Modelling' is an original work and that I am the sole author. This work has not been previously presented for an award at any University. All sources, including the published work of others, are clearly acknowledged as references.

I addition, I declare that parts of this research have been presented at conferences held during the course of the research degree. The related publications are as follows:

Oxnard, S., O'Brien, D., van Mourik, J. and Murphy, D., "**Frequency-Dependent Absorbing Boundary Implementations in 3D Finite Difference Time Domain Room Acoustics Simulations**". Invited Paper in *Proceedings of the 10^{th} European Congress and Exposition on Noise Control Engineering, EURO-NOISE*, Maastricht, The Netherlands, June, 2015.

Oxnard, S. and Murphy, D., "**Achieving Realistic Auralisations using an Efficient Hybrid 2D Multi-plane FDTD Acoustic Model**". In *Proceedings of the EAA Joint Symposium on Auralization and Ambisonics*, Berlin, Germany, April, 2014.

Oxnard, S. and Murphy, D., "**Efficient Hybrid Room Impulse Response Synthesis for Room Acoustic Modelling**" [Poster Abstract and Presentation only]. In *Proceedings of the 6^{th} Annual York Doctoral Symposium*, York, UK, October, 2014.

Oxnard, S. and Murphy, D., "**Room Impulse Response Synthesis Based on a 2D Multi-plane FDTD Hybrid Acoustic Model**". In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, October, 2013.

Other Publications:

van Mourik, J., Oxnard, S., Foteinou, A. and Murphy, D., "**Hybrid Acoustic Modelling of Historic Spaces using Blender**". In *Proceedings of Forum Acusticum*, Krakow, Poland, September, 2014.

Oxnard, S. and Murphy, D., "**Achieving Convolution Based Reverberation through use of Geometric Acoustic Modelling Techniques**". In *Proceedings of the 15^{th} International Conference on Digital Audio Effects (DAFx12)*, York, UK, September, 2012.

Chapter 1

Introduction

The process of room acoustic modelling seeks to digitally simulate the propagation of sound through enclosed environments. This, in turn, facilitates the analysis and sonic reproduction of real or theoretical spaces. In recent decades, various acoustic simulation procedures have undergone significant development to provide reliable means of designing, examining and shaping the world around us. The acoustic characteristics of a given space may be predicted in an objective sense and subjectively experienced through *auralisation*: the reproduction of a soundfield over loudspeakers or headphones. Several commercial and industrial applications rely on and benefit from this outcome. For instance: in architectural consultancy, acoustic modelling processes are used to investigate and listen to the sonic properties of a new building or structure to inform important acoustic design decisions prior to construction; for virtual reality systems, where a primary aim is to convince a person they are somewhere they are not, the simulation and presentation of accurate auditory information is required to produce an immersive and realistic user experience; room acoustic modelling methods may be applied to the simulation of external environments supporting the study of noise control to objectively and perceptually evaluate the impact of day-to-day factors (e.g. transportation networks, large-scale industry projects, noise from machinery) on soundfields in which people live, work and socialise; digital heritage uses acoustic modelling to represent soundfields in areas of historic architectural/cultural significance for posterity and to recreate the sound experienced by historic communities in spaces which no longer exist.

1.1 Motivation

There are two categories of acoustic modelling techniques: numerical and geometric. While geometric methods are considered appropriate for simulating only mid-to-high frequency sound propagation, they are in general less computationally expensive than alternative numerical methods. Currently, there is a prevalence of geometric modelling methods applied to both industry (e.g. acoustic consultancy, architectural design, noise pollution control) and creative arts (e.g. game audio development, virtual reality systems, sound design and composition). The author postulates that this prevalence is influenced by established use of geometric acoustic modelling techniques, which have been in existence for a longer period than numerical methods, and the prohibitive computational costs inherent to numerical acoustic modelling. As such, areas of industry in which sound simulation is a primary concern have yet to harness the proven capabilities of numerical acoustic modelling for faithfully representing wave motion. Furthermore, acoustic modelling procedures that facilitate real-time manipulation of simulated soundfields are an attractive proposition for all applications listed previously. With dynamic control of model attributes, such as geometry and source/receiver locations, it is possible to rapidly examine and experience changes in a simulated soundfield as the model attributes are altered.

This thesis is concerned with reducing the computational requirements of numerical acoustic models while maintaining a level of accuracy in the representation of low frequency wave characteristics which cannot be achieved using alternative geometric modelling methods. To this end, numerical modelling is applied to simulate sound propagation over 2D planar cross-sections of a target acoustic field as part of a hybrid acoustic modelling approach. This reduction in dimensionality provides an approximation to 3D wave simulation while providing significant savings in memory requirements and simulation run-times compared to full 3D numerical modelling.

1.2 Statement of Hypothesis

This research examines the applicability of 2D Finite Difference Time Domain (FDTD) schemes to the simulation of low frequency sound propagation in enclosed spaces as part of a hybrid RIR synthesis solution. As discussed in the statement of motivation, the justification for this work stems from the following concepts:

- Prohibitive computational requirements of numerical acoustic modelling paradigms is significantly alleviated by reducing the spatial dimensionality of the synthesised target acoustic field.

- Low frequency sound wave phenomena in 3D space may be represented, in part, by efficient 2D numerical acoustic simulations.
- Wave characteristics, such as diffraction and resonances, may be emulated to a higher level of accuracy in 2D numerical acoustic simulations than that attained using alternative geometric acoustic modelling paradigms.

In line with the concepts above, a novel numerical acoustic modelling strategy, which synthesises low frequency enclosed sound propagation by means of multiple 2D FDTD planes, is developed in this study. This technique, termed '2D *multiplane* FDTD', is intended to be applied to virtual acoustic modelling in combination with high frequency geometric simulations forming an efficient hybrid solution.

In order to examine the extent to which this new hybrid acoustic model is suited to the task of simulating soundfields, the following hypothesis is investigated in this research:

2D multiplane FDTD acoustic modelling facilitates efficient simulation of low frequency soundfields to a level of accuracy higher than geometric models and comparable to 3D FDTD simulations as part of a hybrid modelling solution.

This hypothesis is assessed primarily through objective analysis of low frequency room impulse responses (RIRs) synthesised using the 2D multiplane FDTD approach with reference to results obtained from both 3D FDTD and geometric acoustic models. These results are drawn from a range of virtual acoustic models created during the course of this study. Simulation data pertaining to the computational cost of 2D multiplane and 3D FDTD is documented in order to demonstrate the notable reduction in memory and run-time requirements achieved by the multiplane approach compared with 3D FDTD modelling.

1.3 Novel Contributions of this Thesis

In the completion of this thesis, the following novel contributions have been identified:

1) Comparison of three implementations of frequency-dependent absorbing boundary conditions used in 3D FDTD acoustic simulations provides insight on the performance of each in terms of accuracy and computational efficiency.

2) Investigation into the numerical stability of frequency-dependent and frequency-independent locally reacting surface absorbing boundary conditions for standard rectilinear 3D FDTD acoustic simulations leads to a reformulation of frequency-dependent

boundary conditions that is shown to be robust against numerical instability issues which arise for complex acoustic model geometries.

3) The development and evaluation of a 2D multiplane hybrid acoustic modelling solution demonstrates that this novel approach is capable of simulating RIRs with increased accuracy at low frequencies compared to that produced using geometric methods only while achieving significant reductions in computational cost compared to alternative 3D numerical acoustic modelling.

1.4 Thesis Structure

This section details the contents of this thesis as presented.

Chapter 2: A range of concepts drawn from fundamental acoustic theory are discussed. Particular attention is given to the principles of sound wave propagation in enclosed environments and the means by which characteristics of room acoustics may be quantified, analysed and explained. The theory reviewed in this chapter informs the implementation and evaluation of virtual acoustic modelling paradigms developed for the purposes of this work.

Chapter 3: An extensive review of current state of the art in virtual acoustics modelling is presented in order to inform the design of the hybrid RIR synthesis method investigated later in this thesis. The theoretical basis and implementation of a range of geometric and acoustic modelling methods are examined to ascertain the scenarios in which each may be successfully applied to the problem of RIR synthesis. This involves assessing the capabilities and limitations inherent to each method. Additionally, details of numerical/geometric hybrid acoustic modelling solutions are given to demonstrate the advantages of hybrid approaches over purely geometric or numerical methods. Findings support the application of the FDTD method to 2D multiplane modelling.

Chapter 4: Having noted that the FDTD paradigm is appropriate for constructing the 2D multiplane acoustic model, important aspects of rectilinear FDTD schemes are examined in depth. Specifically, scheme isotropy, sound source excitation, RIR capture, frequency-independent and frequency-dependent absorbing boundary conditions, and numerical stability issues are detailed for both 2D and 3D FDTD simulations. With reference to RIR capture, a method for deriving spatial RIRs from FDTD models is briefly investigated. The concepts detailed in this chapter underpin and influence the implementation of the 2D multiplane and 3D FDTD models created in order to investigate the hypothesis that informs this thesis. This chapter includes an investigation into

the implementation of three frequency-dependent absorbing boundary conditions for 3D FDTD acoustic models. Findings are drawn from analysis of objective results calculated from RIRs simulated using each boundary condition implementation and details pertaining to the computational cost incurred in each modelling case. Additionally, examination of numerical stability issues arising due to the use of absorbing boundary conditions in rectilinear 3D FDTD acoustic model of complex room geometries are also presented.

Chapter 5: In this chapter, the aspects of virtual acoustic modelling and acoustic theory discussed in previous chapters are applied to the definition, implementation and analysis of the 2D multiplane FDTD hybrid RIR synthesis approach. Objective results drawn from RIRs produced for three scenarios of virtual room acoustic modelling are applied to investigate the stated research hypothesis.

Chapter 6: Key findings of this work are summarised and reviewed with regards to the guiding research hypothesis. The novel contributions of this work are reflected upon in the context of recent developments in the field. Additionally, routes for future studies that have been brought to light through completion of this work are detailed.

Chapter 2

Elements of Acoustic Theory for Sound Simulation and Analysis

Theoretical acoustics is the study of physical systems that facilitate sound transmission. In general, such systems are derived from observed events and translated into mathematical equations that seek to describe the nature of sound propagation in any given medium. The phenomenon of sound may manifest in solids, liquids and gases and exist beyond the ranges of human perception in the form of 'infrasonics' and 'ultrasonics'. For the purposes of this work, the interest lies in sound that is detectable by the human auditory system and propagates over a medium of bounded air, or more plainly, 'Room Acoustics'. When investigating the emulation of room acoustics via a computer modelling method, it is required that an understanding of sound properties be applied to inform the model's design. More crucially, an appreciation of the assumptions made to construct the governing mathematical equations provides insight into potential limitations of any given modelling approach. Key principles of theoretical acoustics must be applied in order to evaluate the correctness of an acoustic model through comparison of modelled results and those derived from theory. Furthermore, an overview of the human hearing system and psychoacoustic concerns is necessary to provide insight on the subjective evaluation of simulated audio. This chapter provides a review of particular concepts in acoustics that are central to the topics of study presented throughout this work. The fundamental theory detailed in the following may be found in various texts on the subject of acoustics, e.g. [1, 3, 23–27].

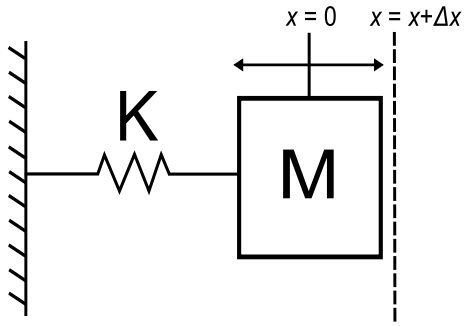

FIGURE 2.1: A free-body diagram of the Simple Harmonic Oscillator with one degree of spatial freedom x, comprised of a Mass (M) and Spring with constant (K) connected to a rigid termination.

2.1 Principles of Wave Motion

Sound exists as waves. A wave may be described as the transport or movement of a disturbance in a medium through the medium itself. In reality, the shape of a disturbance of this kind, when viewed over space or time (or both), can take infinite forms depending on the state of the medium before excitation and the nature of the excitation that caused the disturbance. In an idealistic case, sound waves travel unimpeded throughout an isotropic fluid medium. Although this is a radical simplification of what occurs in practice, analysis of such ideal cases provides important and useful expressions that are central to the study of acoustics.

2.1.1 The Simple Harmonic Oscillator

Simple harmonic motion is a concept that provides the fundamental basis for the mathematics of wave motion. The most basic mechanical system that supports this particular type of behaviour is the simple harmonic oscillator (SHO). This system, as depicted in Figure 2.1, is comprised of a mass body that is linked to a rigid termination via a spring. To begin analysis, it is first assumed that the system is lossless and, therefore, does not account for the effects of friction and so forth. Displacing the mass M in the positive x direction produces a force F acting on the mass which is governed by Hooke's law:

$$F = -Kx \tag{2.1}$$

where force F is measured in Newtons (N), K is the stiffness constant of the spring with units $\frac{\text{kg}}{\text{s}^2}$ and x is the amount by which the mass is displaced in metres (m). Employing Newton's second law of motion, it is possible to express the transient behaviour of the

system by means of the time domain differential equation for SHO:

$$\frac{d^2x}{dt^2} + \frac{K}{M}x = 0 \tag{2.2}$$

a general solution to which is:

$$x = A\cos(\omega_0 t + \phi_0) \tag{2.3}$$

where $\omega_0 = \sqrt{\frac{K}{M}}$ is angular frequency (rads/s), ϕ_0 is a phase constant (rads) and A is amplitude. As such, it is apparent that when the mass is displaced from its equilibrium position and released, the resulting motion of the mass is unimpeded sinusoidal oscillation with resonant frequency $f = \frac{\omega_0}{2\pi}$ Hz. This regular, or harmonic, mass displacement over time represents lossless temporal sound wave motion and is recognised and perceived as a single frequency 'tone'. SHO is the fundamental basis of all sound waveforms due to the fact that any complex waveform may be expressed as the weighted sum of sinusoidal components in accordance with Fourier theory. In addition, the SHO can be viewed as the building block of any acoustic system as it describes the oscillation of a point mass over time. Coupling a series of these point masses in any degree of spatial dimensionality creates a physical system that facilitates wave motion through time and space.

2.1.2 Travelling Waves

As waves propagate throughout a medium, the medium itself deforms taking on the shape of the wave across the space where the waveform is situated at some point in time. This deformation is composed of a series of compressions and rarefactions that are produced due to the transport of kinetic energy resulting from a disturbance. An apt analogy to this process is provided by Howard & Angus [1], referred to as the 'golf ball and spring' model. This model consists of a number of SHO elements interconnected by springs where the point masses are depicted as golf balls. This arrangement is depicted in Figure 2.2.

It is assumed that the 1D model, which comprises a simple representation of an air column, is lossless and that the mass and spring constants are uniform throughout. Upon excitation, the left-most golf ball is displaced in the positive x direction causing the right hand spring to compress and the left hand spring to extend. This compression acts to transfer kinetic energy to the next mass and the rarefaction imposed on the first mass restores the mass to its equilibrium position. The resulting movement of the first mass causes a rarefaction between the first two masses as the second mass is compressed against its right hand spring. An interaction of this kind is repeated

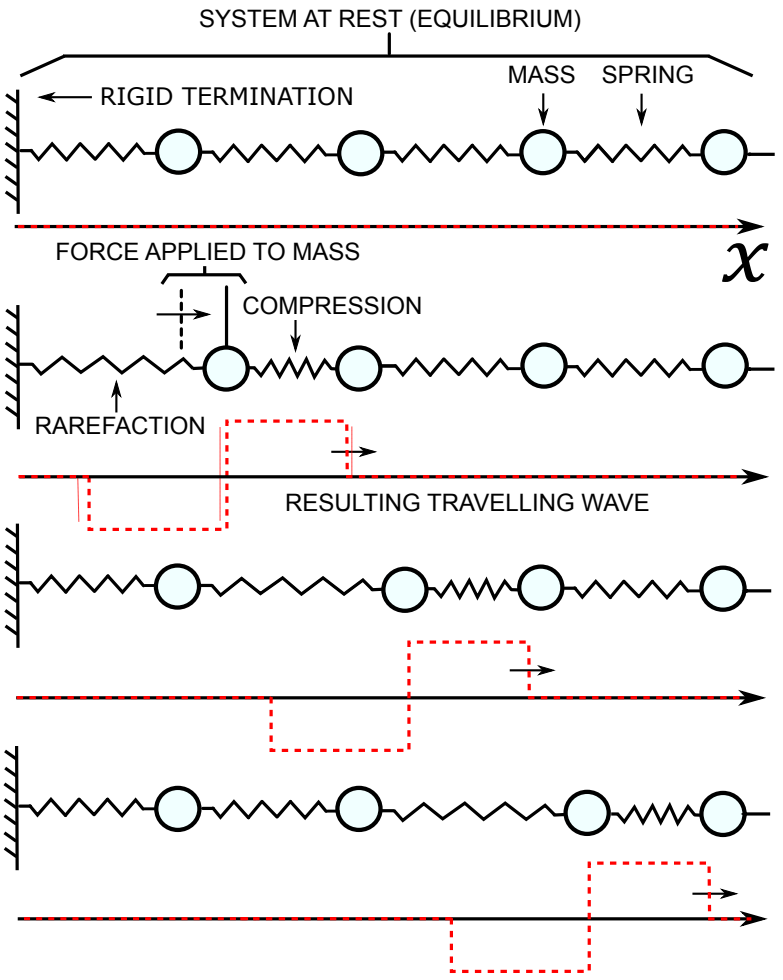

FIGURE 2.2: Depiction of a travelling wave based on the golf ball and spring model, after [1]. The waveform resulting from the series of rarefactions and compressions is shown below each stage by the dashed red line.

for every mass-spring-mass group. Hence, the transport of energy throughout the model is achieved by a series of compressions and rarefactions. Notice that the shape of the resulting waveform does not change shape as it propagates over time through the medium, hence this type of wave is called a *travelling wave*. In terms of sound propagation in air, this model represents the way in which pressure fluctuations, i.e. the displacement of air molecules, progress throughout an infinitely thin column of air. In order to further study the nature of wave movement in space and time, this concept can be translated into a single governing mathematical expression.

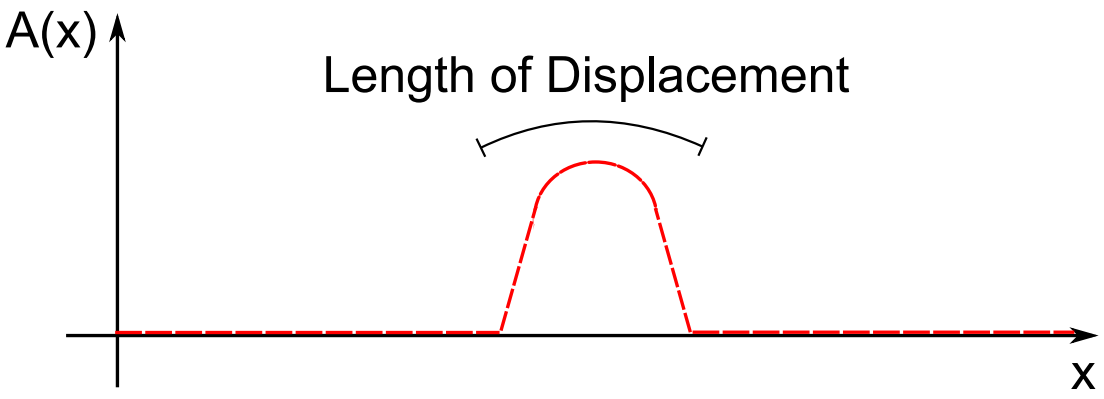

FIGURE 2.3: A simple 1D string model, represented by the red dashed line, with displacement along a section length.

2.1.3 The Wave Equation

Central to the theoretical description of any propagating acoustic waveform is a second-order partial derivative expression: the *wave equation*. This single equation is a common starting point for numerical acoustic models and can be employed to realise simulation of sound propagation in any given space. In addition to close examination of the wave equation and related properties, this section provides insight on simple boundary conditions that must be used to appropriately terminate the spatial domain in order to emulate and examine the properties of enclosed sound waves.

2.1.3.1 1D Wave Equation

A simple 1D physical system may be envisaged as a flexible string that occupies a set length in a single spatial dimension, x, giving a system similar to that used during discussion of travelling waves. At rest, the string has no deformations. Upon excitation, a portion of the string is displaced from equilibrium position as shown in Figure 2.3. Intuitively, it is expected that when the string is displaced in this manner and then released, some form of wave motion will occur over time and space. Thus, what is desired is an expression that relates the shape of the waveform, i.e. the initial deformation, to the transient behaviour of a single point on the string. This is achieved by examining a small portion of the string as it is displaced as shown in Figure 2.4.

It is assumed that the string is infinitely flexible, has uniform mass ρ_0 and is being acted upon by some tension, T. Under such conditions, the overall force acting on the string section ds in the p-direction is given by $T(\sin(\theta_2) - \sin(\theta_1))$. A further assumption that the displacement of the string is very small suggests that the angles θ_1 and θ_2 are also very small. Under this condition it is possible to state that $\sin(\theta_1) \approx \tan(\theta_1)$ with a

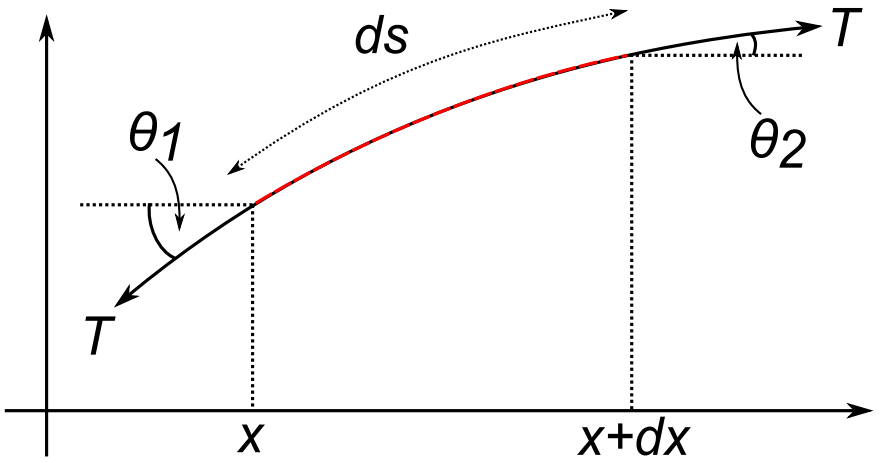

FIGURE 2.4: An analytical diagram showing uniform tension (T) acting on a subsection (ds) of the 1D string.

similar expression for θ_2. Since $\tan(\theta_2)$ and $\tan(\theta_1)$ are equal to the slope of the string at points $(x + dx)$ and x respectively, the total vertical force on this portion of the string is:

$$T\left(\left(\frac{\partial p}{\partial x}\right)_{x+dx} - \left(\frac{\partial p}{\partial x}\right)_{x}\right) = Tdx\frac{\partial}{\partial x}\left(\frac{\partial p}{\partial x}\right) = Tdx\frac{\partial^2 p}{\partial x^2} \tag{2.4}$$

after [23, 26], using the following relationship:

$$f(x + dx) = f(x) + dx\frac{\partial f}{\partial x} \tag{2.5}$$

which states that the value of some function $f(x)$ at the point $(x + dx)$ is equal to the value of $f(x)$ plus the the spatial element dx multiplied by the variance of f as x increases to the point $(x+dx)$. Hence, the displacement of the string p has been related to the spatial variable x. What remains is to relate this expression to the transient behaviour of the element of the string. Using Newton's second law of motion, the net force in the direction of displacement F_p on the portion of the string can be written as:

$$F_p = \rho_0 dx\frac{\partial^2 p}{\partial t^2} \tag{2.6}$$

where the term $\rho_0 dx$ gives the total mass of the string section. Given that an expression for the vertical force is known (2.4), the spatial and temporal terms may now be equated.

$$Tdx\frac{\partial^2 p}{\partial x^2} = \rho_0 dx\frac{\partial^2 p}{\partial t^2} \tag{2.7}$$

Finally, (2.7) may be rearranged to yield the homogeneous 1D wave equation:

$$\frac{\partial^2 p}{\partial t^2} = c^2 \frac{\partial^2 p}{\partial x^2}; c = \sqrt{\frac{T}{\rho_0}} \tag{2.8}$$

The solutions to (2.8) are recognised as waves that, under the assumptions previously defined, propagate without change in shape at a constant speed of c (m/s). In this one dimensional case, the value $p(x,t)$ simply refers to the distance (m) of a point on the string from its equilibrium position at time t (s).

In order to validate the simulation of wave motion of this kind, it is useful to gain a theoretical understanding of how a 1D system reacts to a disturbance. Green's functions provide a means of expressing the propagation of an impulsive disturbance in free field conditions. The Dirac delta function is an impulsive excitation that simplifies the derivation of the Green's function for the 1D wave equation. This impulse is defined as follows:

$$\delta(t) = \begin{cases} \infty, & \text{t} = 0 \\ 0, & \text{t} \neq 0 \end{cases} \tag{2.9}$$

with the following property:

$$\int_{-\infty}^{\infty} \delta(t) dt = 1 \tag{2.10}$$

Applying such an impulse to the homogeneous wave equation yields the following inhomogeneous expression:

$$\frac{\partial^2 p}{\partial t^2} = c^2 \frac{\partial^2 p}{\partial x^2} + A\delta(t)\delta(x) \tag{2.11}$$

where A is the magnitude of the source excitation. Following [28], the aim is to seek solutions to (2.11) by removing the differential terms through Laplace and Fourier transforms to yield a simple algebraic expression for the unknown $p(x,t)$. Firstly, the system is defined to operate under the initial time condition,

$$p(x,t)\Big|_{t=0} = \frac{\partial p(x,t)}{\partial x}\Big|_{t=0} = 0 \tag{2.12}$$

and the infinite convergence criterion,

$$p(x,t)\Big|_{x\to\pm\infty} = \frac{\partial p(x,t)}{\partial x}\Big|_{x\to\pm\infty} = 0 \tag{2.13}$$

after [28]. The time derivative may be removed by means of the Laplace transform, defined as:

$$P(s) = L[p(t)] = \int_0^{\infty} p(t) e^{-st} dt \tag{2.14}$$

where $s = \sigma + jw$ is the Laplace variable. Applying (2.14) to (2.11) yields,

$$\int_0^\infty \frac{\partial^2 p(x,t)}{\partial t^2} e^{-st} dt = c^2 \frac{\partial^2}{\partial x^2} \int_0^\infty p(x,t) e^{-st} dt + A\delta(x) \int_0^\infty \delta(t) e^{-st} dt \tag{2.15}$$

Carrying out the integrations with employment of the initial time condition and the properties of the Dirac delta function results in the following ordinary differential equation:

$$\frac{d^2 P}{dx^2} = \left(\frac{s}{c}\right)^2 P - A\delta(x) \tag{2.16}$$

of the transformed function $P = L[p(x,t)]$. A similar approach is taken to now remove the spatial derivative through use of the Fourier transform:

$$\bar{p}(\phi) = \int_{-\infty}^{\infty} p(x) e^{j\phi x} dx \tag{2.17}$$

Applying the Fourier transform to (2.16) gives,

$$\int_{-\infty}^{\infty} \frac{d^2 P}{dx^2} e^{j\phi x} dx = \left(\frac{s}{c}\right)^2 \int_{-\infty}^{\infty} P e^{j\phi x} dx - A \int_{-\infty}^{\infty} \delta(x) e^{j\phi x} dx \tag{2.18}$$

where the dual transformed unknown function is now,

$$\bar{P} = \int_{-\infty}^{\infty} P e^{j\phi x} dx \tag{2.19}$$

Calculation of the integrals in (2.18), which relies on the properties of the Dirac delta function as well as the convergence criterion, results in the sought after algebraic relation.

$$-\phi^2 \bar{P} = \left(\frac{s}{c}\right)^2 \bar{P} - A \tag{2.20}$$

Through basic manipulations, the unknown may now be solved for.

$$\bar{P} = \frac{A}{\phi^2 + \left(\frac{s}{c}\right)^2} \tag{2.21}$$

What remains at this point is to transform (2.21) back into the time and space domains to achieve an expression for the unknown relative to the original independent variables x, t. Firstly, the inverse Fourier transform (2.22) is applied.

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{p}(\phi) e^{-j\phi x} d\phi \tag{2.22}$$

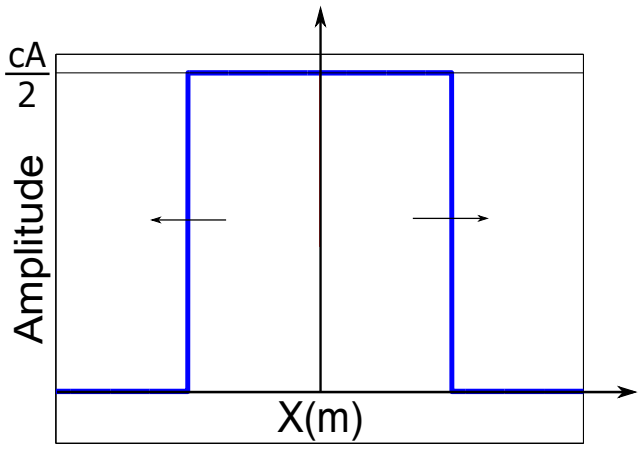

FIGURE 2.5: 1D Green's function for the Wave Equation - the arrows signify the directions of propagation away from the point of excitation.

Now, since the integrand $\bar{P}$ is symmetrical, a single sided semi-infinite integral will suffice to transform the integrand back into the spatial domain.

$$P = \frac{A}{\pi}\int_0^\infty \frac{1}{\phi^2 + \left(\frac{s}{c}\right)^2}\cos(\phi x)d\phi = \frac{A}{2(\frac{s}{c})}e^{-(\frac{s|x|}{c})} \tag{2.23}$$

This result, after [28], is a standard transform as detailed in [29]. Finally, the inverse Laplace transform is employed to regain the time variable,

$$p(x,t) = L^{-1}[P] = \frac{cA}{2}L^{-1}\left[\frac{1}{s}e^{-(\frac{s|x|}{c})}\right] \tag{2.24}$$

Fortunately, the inverse Laplace transform of the type in (2.24) is known to result in the Heaviside unit step function H [29] and hence, the Green's function solution to the 1D wave equation in response to an impulsive excitation is,

$$p(x,t) = \frac{cA}{2}H(ct - |x|) \tag{2.25}$$

This result states that the expected free-field waveform produced in 1D in response to an impulse excitation will take the form of two wavefronts that travel in opposing directions with amplitude $\frac{cA}{2}$ at a speed of c (m/s). Figure 2.5 demonstrates this type of wave propagation. In order to verify the correct implementation of a numerical system, the resulting simulated waveform may be compared to this theoretical expression to ensure that the wave equation is being numerically solved as desired. A similar approach to that described here may be taken to generating the Green's function for waves in 2D and 3D [28, 30].

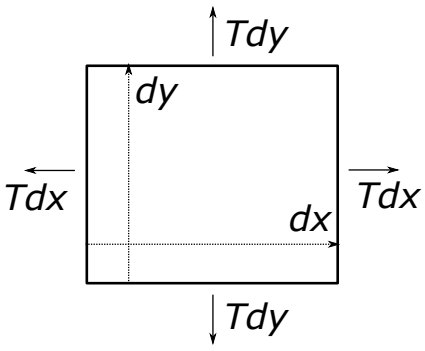

FIGURE 2.6: A square sub-element of a 2D membrane of size $dxdy$ with equal tension acting on all sides.

2.1.3.2 2D Wave Equation

The derivation of the 2D wave equation is a similar process to that of its 1D equivalent. To begin, the wave motion is defined as propagating throughout an infinitely flexible membrane with uniform mass and density. If the membrane is pulled with equal tension T around its bounding edges, it can be stated that any small area $dxdy$ of the membrane will have the same tension acting on each side. This notion is encapsulated in Figure 2.6. As with the 1D system, the objective is to relate a disturbance in the medium to the transient motion of a single point on the membrane. In this case, the disturbance will be a displacement of the membrane from its equilibrium position perpendicular to the xy plane denoted $p(x, y, t)$. In the x-direction, the overall force acting on the small area is given as,

$$Tdy\Big[\Big(\frac{\partial p}{\partial x}\Big)_{x+dx} - \Big(\frac{\partial p}{\partial x}\Big)_{x}\Big] = T\frac{\partial^2 p}{\partial x^2}dxdy \tag{2.26}$$

Similarly, the net force acting in the y-direction is,

$$Tdx\Big[\Big(\frac{\partial p}{\partial y}\Big)_{y+dy} - \Big(\frac{\partial p}{\partial y}\Big)_{y}\Big] = T\frac{\partial^2 p}{\partial y^2}dydx \tag{2.27}$$

Hence, the total net force F_p acting normal to the xy plane is simply the sum of these tensional forces:

$$F_p = T\frac{\partial^2 p}{\partial x^2}dxdy + T\frac{\partial^2 p}{\partial y^2}dydx \tag{2.28}$$

which can be related to the acceleration of the area element to form the 2D wave equation,

$$T\frac{\partial^2 p}{\partial x^2}dxdy + T\frac{\partial^2 p}{\partial y^2}dydx = \rho_0 dxdy\frac{\partial^2 p}{\partial t^2} \tag{2.29}$$

$$\frac{\partial^2 p}{\partial t^2} = c^2\nabla^2 p; \qquad \nabla^2 = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} \tag{2.30}$$

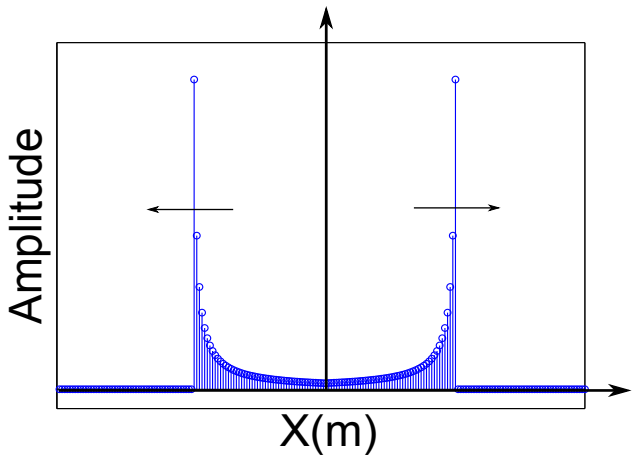

FIGURE 2.7: A cross section of the 2D Green's function for the Wave Equation - the arrows signify the directions of propagation away from the point of excitation. The non-zero pressure field within the two wavefronts is referred to as *afterglow*.

where ρ_0 is the density of the membrane material, ∇^2 is the 2D Laplacian operator and the wave propagation speed $c = \sqrt{T/\rho_0}$ (m/s). In this system, as with the 1D wave equation, it is clear that the shape of the disturbance propagating through the medium is directly proportional to the acceleration of a point on the membrane normal to the xy plane. The expression in (2.30) is the homogeneous form of the 2D wave equation for an isotropic medium. In order to examine the nature of waveform propagation in such a system, a non-homogeneous, impulsive excitation term is introduced to (2.30) and the resulting Green's function is derived. Hence, in 2D the initial equation is,

$$\frac{\partial^2 p}{\partial t^2} = c^2 \nabla^2 p + A\delta(t)\delta(x)\delta(y) \tag{2.31}$$

and the Green's function is sought under quiescent initial conditions with convergence at infinity criterion applied for both spatial dimensions [28]. Following a similar derivation as that of the 1D case, as noted in [28, 30], and references therein, the Green's function in 2D for wave motion in response to an impulsive excitation may be written,

$$p(x, y, t) = \frac{cA}{2\pi} \frac{H(ct - r)}{\sqrt{((ct)^2 - r^2)}} \tag{2.32}$$

where $r = \sqrt{x^2 + y^2}$ (m) is the magnitude distance from the excitation location. This is a very useful result as it exposes an interesting property of 2D wave motion. If the resulting function is plotted out across one dimension of the membrane over time, as in Figure 2.7, it is shown that decaying non-zero components are produced after the passing of the initial wavefront. This quality, known as *afterglow* [30–33], is common to all 2D wave-based systems and special care must be taken if the effects of afterglow are

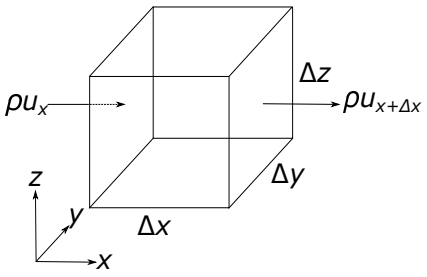

FIGURE 2.8: Diagram of a cubic volume of air ($\triangle x \triangle y \triangle z$ m^3) with the flow of mass ρ and at velocity u highlighted in the x-direction.

to be removed. Further discussion of afterglow is given in relation to dimensionality reduction for efficient numerical acoustic modelling in Chapters 4 and 5.

2.1.3.3 3D Wave Equation

The derivation of the 3D wave equation is a more involved process than that of the lower dimensionality equations. This is mainly due to the fact that the idea of tension cannot strictly be applied to describe forces acting on a volume. Instead, the derivation begins with the introduction of the law of mass conservation. The diagram given in Figure 2.8 represents a small cubic element of space with volume $\triangle V = \triangle x \triangle y \triangle z$. Considering the scenario where a fluid, for example air, is able to flow through the element from every direction, it may be stated that the rate of increase of mass within the element must equal the amount of mass flowing in and out through each element surface. Hence, in the case of air flow in the positive x-direction, the following expression must be true:

$$-\frac{\partial \rho}{\partial t} \triangle y \triangle z = (\rho u_x|_x - \rho u_x|_{x+\triangle x}) \triangle y \triangle z \tag{2.33}$$

where u_x is the fluid velocity component in the x-direction and ρ is the density of the fluid. Similar expressions exist for inflow and outflow in the remaining two spatial dimensions.

$$-\frac{\partial \rho}{\partial t} \triangle x \triangle z = (\rho u_y|_y - \rho u_y|_{y+\triangle y}) \triangle x \triangle z \tag{2.34}$$

$$-\frac{\partial \rho}{\partial t} \triangle x \triangle y = (\rho u_z|_z - \rho u_z|_{z+\triangle z}) \triangle x \triangle y \tag{2.35}$$

As such, the rate of increase of mass in element $\triangle V$ is related to the total fluid inflow by,

$$-\frac{\partial \rho}{\partial t} \triangle V = (\rho u_x|_x - \rho u_x|_{x+\triangle x}) \triangle y \triangle z + (\rho u_y|_y - \rho u_y|_{y+\triangle y}) \triangle x \triangle z + (\rho u_z|_z - \rho u_z|_{z+\triangle z}) \triangle x \triangle y \tag{2.36}$$

Dividing through by $\triangle V$ yields the *continuity equation* [26] in terms of particle velocity and instantaneous density:

$$\frac{\partial \rho}{\partial t} + \nabla \bar{u}\rho = 0 \tag{2.37}$$

where $\bar{u}(x, y, z, t)$ is the particle velocity vector in three spatial dimensions. However, for the study of acoustics the variable of interest is acoustic pressure $p(x, y, z, t)$ which can be related to fluid density using the following equations.

$$s = \frac{\rho - \rho_0}{\rho_0} \tag{2.38}$$

$$p = Bs \tag{2.39}$$

The new term s is a measure of condensation which calculates the density of a fluid with reference to the equilibrium density. For example, a value of $s = 0$ suggests that the fluid is neither compressed or in a state of rarefaction, $s > 0$ suggests that the fluid is compressed in the volume under observation and $s < 0$ implies that a rarefaction is occurring. Additionally, the value B refers to the bulk modulus of the fluid. After [26], the instantaneous density my be written as $\rho = \rho_0(1 + s)$. It is then assumed that ρ_0 can be regarded as constant over space and time, hence $s << 1$ giving:

$$\rho_0 \frac{\partial s}{\partial t} + \rho_0 \nabla \bar{u} = 0 \tag{2.40}$$

Finally, dividing through by ρ_0 allows expression of the continuity equation in terms of acoustic pressure and particle velocity.

$$\frac{1}{B}\frac{\partial p}{\partial t} + \nabla \bar{u} = 0 \tag{2.41}$$

Next, the law of conservation of momentum is considered in order to relate the particle velocity $\bar{u}$ to acoustic pressure such that it will then be possible to construct an expression in terms of pressure, space and time only. Referring again to Figure 2.8, it may be stated that to displace the element of fluid mass $\triangle m$ at an acceleration $\bar{a}$, a force $\triangle \bar{f}$ is required. In a similar fashion to the balance of mass flow, the net force acting on the element in the x-direction may be expressed as follows:

$$\triangle f_x = (P|_x - P|_{x+\triangle x})\triangle y \triangle z = -\frac{\partial P}{\partial x}\triangle V \tag{2.42}$$

where P is a measure of instantaneous pressure at a discrete point in space and time. Similar relations exist for forces acting in the remaining two spatial dimensions. Following the derivation given in [26], the acceleration of the element mass may be obtained from the particle velocity at location $[x, y, z]$ at time t and that which occurs at a new, nearby location at time $t + \triangle t$. Employing Taylor's series expansion allows expression

of the new velocity in a convenient form.

$$\bar{u}(x+u_x\triangle x, y+u_y\triangle y, z+u_z\triangle z, t+\triangle t) = \bar{u}(x,y,z,t) + \frac{\partial \bar{u}}{\partial x}u_x\triangle t + \frac{\partial \bar{u}}{\partial y}u_y\triangle y + \frac{\partial \bar{u}}{\partial z}u_z\triangle t \tag{2.43}$$

The acceleration may, consequently, be found by taking the first-order difference of velocity with respect to time. Letting the right hand side of (2.43) $= I(x,y,z,t)$,

$$\bar{a} = \lim_{\triangle t\to 0}\frac{I(x,y,z,t)-\bar{u}(x,y,z,t)}{\triangle t} = \frac{\partial \bar{u}}{\partial t} + (\bar{u}.\nabla)\bar{u} \tag{2.44}$$

where,

$$(\bar{u}.\nabla) = u_x\frac{\partial}{\partial x} + u_y\frac{\partial}{\partial y} + u_z\frac{\partial}{\partial z} \tag{2.45}$$

Now, employing Newton's second law facilitates the relating of velocity and instantaneous pressure:

$$-\triangle P\triangle V = \Big(\frac{\partial \bar{u}}{\partial t} + (\bar{u}.\nabla)\bar{u}\Big)\triangle m \tag{2.46}$$

Since the mass of the small cubic element $\triangle m = \rho\triangle V$ and assuming only small density and acoustic pressure fluctuations (such that P and ρ are respectively equal to the acoustic pressure and fluid density at equilibrium), (2.46) becomes:

$$\triangle p + \rho_0\Big(\frac{\partial \bar{u}}{\partial t} + (\bar{u}.\nabla)\bar{u}\Big) = 0 \tag{2.47}$$

This expression may now be arranged into the linear Euler's equation [3, 26] under further application of the assumption that the variance in acoustic variables are small compared to space. As such, the time derivative of $\bar{u}$ has a much greater magnitude than the spatial derivative giving the linear Euler's equation:

$$\triangle p + \rho\frac{\partial \bar{u}}{\partial t} = 0 \tag{2.48}$$

The linear 3D wave equation results from combination of the laws of mass and momentum conservation. This is achieved by eliminating the particle velocity term in (2.48) using (2.41) giving,

$$\nabla\triangle p + \rho_0\frac{\partial}{\partial t}\Big(\frac{-1}{B}\frac{\partial p}{\partial t}\Big) = 0 \tag{2.49}$$

and by rearranging,

$$\frac{\partial^2 p}{\partial t^2} = c^2\nabla^2 p; \qquad \nabla^2 = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \tag{2.50}$$

where $c = \sqrt{\frac{B}{\rho_0}}$ is the speed of sound in the medium and ∇^2 is the 3D Laplacian operator. The derivation of this equation has involved numerous assumptions that

mostly relate to the restriction that sound waves cause only very small fluctuations in a medium from the equilibrium state. Additionally, the wave equation is linear and assumes an isotropic medium where movement of the medium from one location to another is not facilitated. This is a simplification of what occurs in reality where most instances of sound propagation will involve many non-linear and chaotic processes. However, the homogeneous linear wave equation is sufficient for the purposes of this thesis.

The Green's function for the 3D wave equation in response to an impulsive sound source may be stated as follows,

$$\frac{\partial^2 p}{\partial t^2} = c^2\nabla^2 + A\delta(x)\delta(y)\delta(z)\delta(t) \quad \Rightarrow \quad p = \frac{A}{4\pi R}\delta\Big(t - \frac{R}{c}\Big) \tag{2.51}$$

where R (m) is the magnitude distance from the sound source location. Again, as per [28], this result is derived assuming quiescent conditions and the convergence criterion in all spatial dimensions. Examination of (2.51) shows that the 3D wave equation is the only case discussed so far that preserves the shape of the initial excitation function as the wave propagates over the space at speed c (m/s). As the surface area of the wave increases, the amplitude of the waveform will decrease inversely proportional to this area increase. Again, this result is very useful in verifying correct sound wave propagation in numerical models.

2.1.3.4 Simple Boundary Conditions

In order to fully describe a realistic acoustic system, the spatial domains must be terminated appropriately leading to the requirement of boundary conditions. Such conditions, which govern the nature of the interaction of a pressure wave on impact with a surface, can take many forms. This is due to the fact that any combination of basic mechanical elements can be arranged at the termination of a spatial domain to reflect and absorb acoustic energy in a manner approximating reality. This subject will be discussed at various points throughout the following chapters, however a review of basic boundary conditions is given here to introduce important characteristics of particular boundary formulations.

The most basic boundary conditions are the Dirichlet and Neumann conditions that describe ideal, lossless terminations. Firstly, the Dirichlet ('Fixed') boundary is defined as,

$$p(L_x, y, z, t) = 0 \tag{2.52}$$

for a right-hand side surface boundary at locations $[L_x, y, z]$ where $p(x, y, z, t)$ is acoustic pressure. A condition of this kind forces the pressure at such boundaries to a zero value. In 1D and 2D systems, the terminations can be thought of as physically clamped as the string or membrane will never be displaced in these locations. A further property of this condition is that waves incident upon the boundary will undergo phase reversal.

The Neumann boundary condition is given by,

$$\frac{\partial p(x, y, z, t)}{\partial n} = 0 \tag{2.53}$$

In contrast to the fixed condition, the Neumann boundary forces the gradient of the pressure field $p(x, y, z, t)$ in the direction normal to the boundary n. As such, in 1D and 2D systems this termination type facilitates variance in pressure at the boundary and phase preservation of an ingoing waveform. Application of each boundary type results in different solutions to the wave equation and in both cases no energy will be absorbed at the terminations of the domain. Ideal systems of this form, for simple spatial geometries, may be solved to produce theoretical solutions which can be employed to validate the results of an acoustic modelling approach.

For more realistic boundaries, that allow for absorption at terminations a simple condition may be imposed [34],

$$\frac{\partial p}{\partial t} = \alpha \frac{\partial p}{\partial n} \tag{2.54}$$

which relates the velocity of the pressure field p to the gradient of the pressure field normal to the boundary. Recalling that the variance in displacement of the medium from the equilibrium position is very small, this condition imposes a restriction on the rate of movement of the pressure field with a variance governed by the values of a loss coefficient α. This condition facilitates the rate at which energy is dissipated from an acoustic system to be altered as required to approximate a real physical process. Further development of this useful absorbing boundary condition is detailed in Chapter 4.

2.1.4 A Lossy Wave Equation

A simple lossy wave equation is defined in [26], as follows:

$$\frac{\partial^2 p}{\partial t^2} = c^2 \nabla^2 - \beta \frac{\partial p}{\partial t} \tag{2.55}$$

where all variables are as previously defined and the additional term β is a loss coefficient. This loss term governs the extent to which the velocity of the pressure field is

reduced as the term $-\beta\frac{\partial p}{\partial t}$ acts contrary to the direction of acceleration. As such, for values of $\beta > 0$ the loss term will cause a decay of energy within the medium itself, regardless of boundary conditions, as wavefronts propagate over space and time. In general, β can be defined as a function of frequency facilitating frequency-dependent loss consistently throughout the medium. However, for the purposes of this work, β will take a positive-real numerical value to provide a frequency-independent lossy wave equation. In a 2D system, the additional loss term can be thought of as reducing the kinetic energy in a membrane. For example, if the membrane is pulled up from its equilibrium position and released, the counteracting velocity term acts to slow its returning movement implying that kinetic energy is being dissipated at a rate proportional to β.

2.2 Sound Propagation in Air

2.2.1 The Speed of Sound

Following on from the formulation of the 3D wave equation, it can be stated that the speed of sound in air, c, is equal to the square root ratio of the bulk modulus and the equilibrium density of air $\sqrt{\frac{B}{\rho_0}}$. Calculating a numeric constant for the speed of sound requires empirical investigation of the properties of air. Furthermore, the bulk modulus of a fluid varies depending on temperature. Kuttruff [3] provides an expression for calculating the speed of sound in air as follows:

$$c = 331.4 + 0.6\theta_T \quad ms^{-1} \tag{2.56}$$

where θ_T is the temperature in degrees Celius ($^\circ$). For the purposes of this study, (2.56) is used to calculate sound wave propagation speed as constant speeds for a given room temperature (i.e. 20°) may be found.

2.2.2 Air Absorption

Air is a viscous medium. As such, it is not sufficient to treat the propagation of sound through air as a lossless process. In reality, air absorption of sound is a highly complex phenomenon dependent on an array of factors, especially in exterior environments. When disregarding the impacts of the weather, i.e. considering only interior environments, air absorption levels are found to be related to humidity, temperature and frequency [20]. The frequency dependency of air absorption is encountered in practice and can be experienced by listening to distant sound sources in large external environments. In the vicinity of the sound source, the frequency range of the source

Frequency (Hz)	125	250	500	1k	2k	4k
Attenuation (dB/100m), approx.	0.032	0.07	0.18	0.46	1.00	2.62

TABLE 2.1: Approximate attenuation (dB SPL) of sound in air over a distance of 100m measured for a range of frequencies, after [20].

is experienced within the limits of the human auditory system. At a large distance away from the source, the same sound is perceived as dull or lacking in timbre. This is the effect of frequency-dependent absorption reducing the amplitude of high frequency sound components as the sound is propagated through air.

In a study conducted by Harris [20], an experimental procedure was devised to empirically define a temperature, humidity and frequency-dependent attenuation coefficient for air. The results produced demonstrate that air absorption is reasonably low for room temperature and frequencies up to 2 kHz. For example, given a relative humidity of 50% and a temperature of 20°C (typical values for interior environments) the attenuation of sound pressure (dB) over a distance of 100m are given approximate values shown in Table 2.1 for a range of frequencies.

Hence, it may be stated that air absorption becomes an increasingly important factor as sound source frequency increases above the 2kHz range. Furthermore, pressure attenuation due to the properties of air must be considered for spaces of large dimensions and boundaries with low absorption. More recent investigation of air attenuation [35, 36] verifies these statements. In the ISO documentation [36] an atmospheric attenuation coefficient is defined as the reduction in pressure (dB) of a sound source over a distance of 1 km. For a relative humidity of 70% and temperature 20° the attenuation for a noise source in the 500 Hz band is noted as 2.8 dB/km, whereas for an 8 kHz band noise source the attenuation is markedly higher at 76.6 dB/km.

2.2.3 Monopoles and Plane Waves

As sound waves emanate from an excitation position in space, the resulting wavefronts expand travelling away from the excitation point. The most basic conceptualisation of an expanding waveform is the single frequency monopole or *point source*. In 2D, a monopole sound source is one that produces an expanding circular pressure wave. In 3D, the pressure wave is spherical. Mathematically, a monopole is defined as follows,

$$p(\bar{r}, \omega) = \frac{A_0}{|\bar{r} - \bar{r_s}|} e^{-j(\mathbf{k}|\bar{r} - \bar{r_s}|)} \tag{2.57}$$

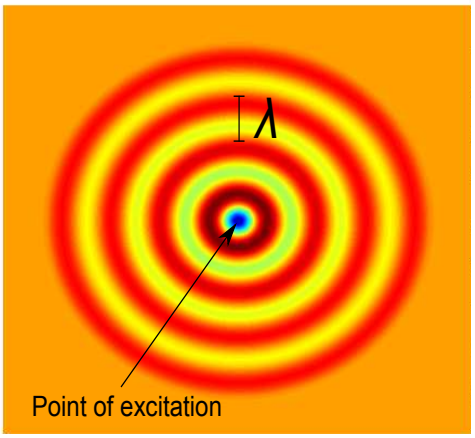

FIGURE 2.9: A 2D monopole sound source of wavelength λ.

where A_0 is the pressure amplitude of the wave and $\bar{r_s}$ and $\bar{r}$ are the source and reference positions respectively expressed as vectors in 2D or 3D space. The wavenumber, $\mathbf{k}$, in 3D space is given by,

$$\begin{aligned} \mathbf{k} &= k_x\hat{i} + k_y\hat{j} + k_z\hat{k} \\ k_x &= k\cos\theta\cos\phi \\ k_y &= k\sin\theta\cos\phi \\ k_z &= k\sin\phi \end{aligned} \tag{2.58}$$

where $k = \frac{\omega}{c}$ (the wavenumber, radians per metre) and $\omega = 2\pi f$ is the angular frequency in radians per second corresponding to a specific frequency f(Hz) with wavelength $\lambda = \frac{c}{f}$ (m). The angles of azimuth and elevation are given by θ and ϕ respectively. By forcing k_z to zero, the definition of a 2D monopole results. This expression is depicted in Figure 2.9. Notice the similarities between (2.57) and the Green's function for the 3D wave equation (2.51) derived previously. It is evident that the amplitude of the expanding wave is consistent in all directions around the solid angle and that this amplitude decreases as the wavefront travels further from the source position. Furthermore, an arbitrary source may be represented by weighted summation of exponential terms of different frequencies thus removing the restriction to a single wavenumber in (2.57).

The planewave approximation to a propagating wavefront arises from examination of a monopole at infinite distance. As the circular/spherical wavefront continues to expand, the distance over which a portion of the wavefront approximates a straight line increases. Hence, at sufficiently large distances from the source location, the monopole

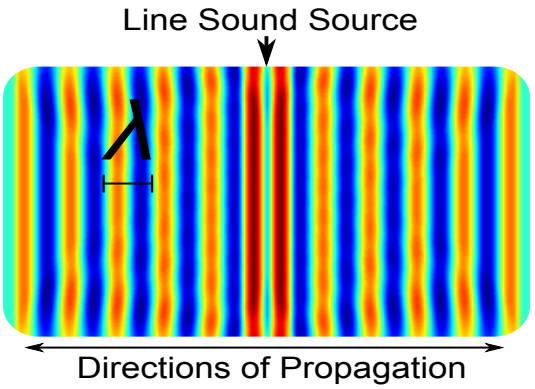

FIGURE 2.10: A 2D planewave as produced by a bi-directional line sound source producing a sinusoid with wavelength λ.

begins to emulate the type of propagation produced by a line source. This concept is exemplified in Figure 2.10.

At a reference location $\bar{r}$, a planewave may be written as,

$$p(\bar{r}, t) = A_0 e^{j(\omega t - \mathbf{k}\cdot\bar{r})} \tag{2.59}$$

Where all terms are as previously defined. Planewaves provide a useful means of investigation for many acoustic systems. For example, the reflection model detailed in section 2.3.2 relies on this approximation.

2.2.4 Sound Pressure - SPL

For the purposes of room acoustic description and modelling, the leading quantity of interest is sound pressure. A unit of pressure, or 1 Pa (Pascal), refers to a force of 1 Newton acting on an area of 1 square metre. However, acoustic pressure is commonly quantified through use of sound pressure level, or 'SPL'. This quantity is defined as the ratio of the instantaneous pressure occurring at an arbitrary position in space and the lower threshold of the human auditory system. Due to the extensive range of values in which sound pressure may be perceived as audible, sound pressure level is defined on a logarithmic scale as follows:

$$SPL = 20\log_{10}\left[\frac{p_{rec}}{p_{ref}}\right] dB \tag{2.60}$$

where p_{ref} is the lowest audible pressure value in the region of 20μPa. p_{rec} is the pressure value recorded at some instant in time. Values for recorded pressure levels

can be as high as 20 Pa [1]. As such, the choice of a logarithmic scale is justified as the pressure range of interest can vary by a factor of 10^6.

2.2.5 Sound Intensity - SIL

An alternative measure corresponds to the flow of sound energy through a unit area as opposed to the force acting on the area. This quantity is the sound intensity level, 'SIL', which corresponds to the number of Joules of energy passing through a unit area per second. For simplicity, it is useful to introduce 'sound power', P (Watts), defined as the number of Joules of energy passing a point in space per second. Hence, SIL has units W/m^2 providing insight into the power density at an arbitrary position in the soundfield under examination. Due to the fact that SIL describes a flow of energy, it must assume a direction of sound propagation. In general, the direction of propagation is assumed to be travelling away from an active sound source and perpendicular to the unit area of interest. As such, SIL holds only for free-field conditions realised, in the most part, under anechoic conditions. Mathematically, SIL is calculated using the following equation:

$$SIL = 10 \log_{10} \left[\frac{I_{rec}}{I_{ref}} \right] dB \tag{2.61}$$

where I_{rec} is the recorded intensity level and I_{ref} is a reference intensity level. As with SPL, the reference level is set to the threshold of human perception, 10^{-12} W/m^2. Again, a logarithmic scale is employed to cover the range of intensities that may be achieved by a real sound source (10^{-12} - 10 W/m^2) [1]. The SIL measure bears less relevance than SPL to the perception of sound experienced in a sonic environment due to the fact that variance in sound energy is perceived as fluctuations in pressure at a localised position. However, examination of SIL leads to an important law concerning the spread of sound energy throughout free space: the *inverse square law*.

2.2.6 The Inverse Square Law

The diagram in Figure 2.11 shows a monopole sound source projecting sound energy into a portion of the free field. As the spherical wavefront progresses away from the point of excitation, the total area, A, over which the wavefront is spread increases as per:

$$A = 4\pi r^2 \tag{2.62}$$

where r is the distance from the source location. Recalling that sound intensity is the energy flow through a unit area, it is apparent the decrease in intensity over space is

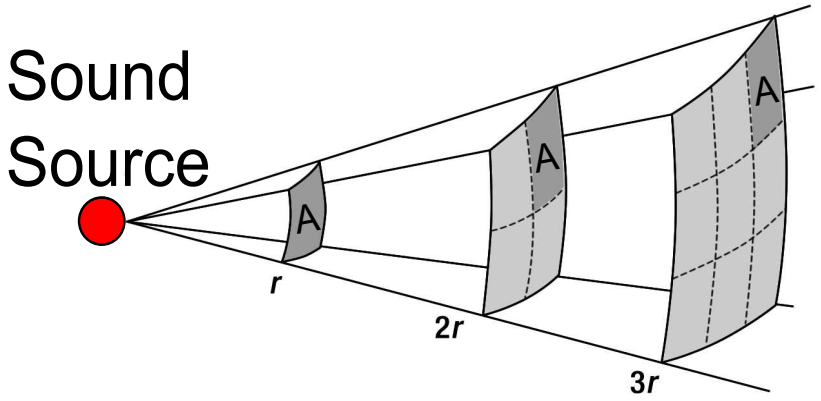

FIGURE 2.11: Depiction of the Inverse Square Law. This diagram shows how the surface area and spread of energy increases proportionally to the distance from the point of excitation. Image adapted from [2].

inversely proportional to the distance from the excitation position. This law, which holds only when assuming negligible loss due to the viscosity of air, is:

$$I = \frac{P}{4\pi r^2} \tag{2.63}$$

where P is the sound power of the source in Watts. This law is commonly encountered in practice and provides explanation as to why sound sources are perceived as loud when close by and less loud when further away with a consistent sound source power.

2.3 Elements of Room Acoustic Theory

Examination of sound propagation in free field conditions provides useful relations and concepts relating to sound characteristics. However, this idealistic system does not occur in reality as any waveform created in practice must eventually dissipate energy through interaction with resistive mediums, for example solid surface obstructions. In the specific case of room acoustics, a sound environment is generally considered to be an enclosure of bounded air. In such an environment, the propagation of sound energy throughout the space is far more complex as the waves are contained due to reflection and forced to interact with one another. This section provides insight on a number of important concepts through which the acoustics of a room may be investigated.

2.3.1 Stages of Enclosed Sound Propagation

There are minimum requirements for the creation and perception of sound in an room. These are: a bounded volume of air, a sound source and a sound receiver. A simple

model that seeks to describe sound propagation from source to receiver temporally divides the arrival of sound energy at the receiver into three stages: *Direct Sound*, *Early Reflections* and *Reverberant Sound*. Referring to Figure 2.12, the left most panel depicts the direct sound path from source to receiver in a simple 2D square geometry. Upon excitation of the soundfield, assuming that the path between source and receiver is unobstructed, the direct sound is that which first arrives at the receiver. In a real space, the reduction of sound energy occurring over the direct path is due only to the inverse square law and loss due to the viscosity of air. For this reason, the direct sound generally gives rise to the strongest pressure fluctuation experienced at the receiver location. As sound waves propagate throughout a space, they interact with bounding surfaces and undergo reflection. This interaction, which is detailed in more depth in section 2.3.2, acts to partially dissipate energy from the sound waves. Early reflections are those which travel along the shortest propagation paths between source and receiver via at least one boundary. In some cases, a high temporal density of early reflections can cause the pressure fluctuations to increase to a level greater than that of the direct sound. This can occur when the room geometry acts to focus reflection paths onto the receiver location. Examples of 1$^{\text{st}}$ and 2$^{\text{nd}}$ order reflections are shown in the center and right panels of Figure 2.12 respectively. These reflections are separated from the direct sound component, and in most cases each other, in the following respects:

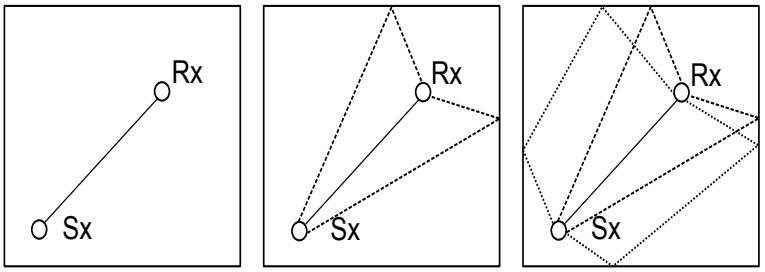

FIGURE 2.12: A simple 2D diagram of the direct sound path between sound source (Sx) and receiver (Rx) within a rectangular enclosure (left). Examples of 1^{st} order (centre) and 2^{nd} order (right) reflections are also shown.

Amplitude: Reflected sound waves travel an increasingly larger distance as the order of reflection increases. Hence, they are subject to increased reduction in pressure amplitude due to the inverse square law and the effects of air absorption. In addition, pressure reduction is also induced by the dissipation of energy upon successive boundary interactions.

Time of Arrival: Sound waves approaching a receiver position via one or more boundaries take longer to reach the position than the direct sound. This is simply due to the

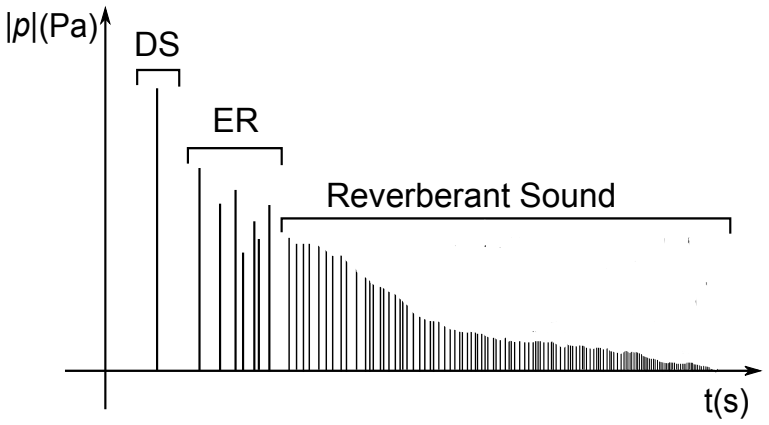

FIGURE 2.13: Magnitude of pressure variation at a receiver position in response to an impulsive sound excitation. The three prominent temporal stages of enclosed sound propagation are highlighted ('DS' - Direct Sound, 'ER' - Early Reflections).

differences between propagation path lengths of reflections and results in a temporal distribution of reflections after the time of arrival of the direct sound component.

Angle of Arrival: Reflections arrive at the receiver location as secondary sources from the direction of the boundary with respect to the receiver's orientation within a room.

The nature of early reflections, as perceived by a listener at the receiver location, will change significantly if the receiver and/or source locations or orientations are altered. As suggested in [1], these changes are of psychoacoustical importance as they provide a listener with auditory cues that describe the size of the environment and the arrangement of source/receiver locations within the soundfield. In large room geometries, temporal delays between early reflections can become very large. For delays greater than approximately 30 ms, the reflections are perceived as echoes [1].

The final temporal stage of sound in an enclosure is the reverberant sound. Reverberant sound is composed of many reflections that have interacted with several boundary surfaces and arrive at the receiver location with a high temporal density. This stage begins when the soundfield is said to have become *diffuse*, meaning that the soundfield is largely stochastic in state with indistinguishable rapidly successive reflections. These diffuse reflections possess decreasing amounts of energy as their traversal distance and reflection order increase. Due to the rapid arrival of the diffuse soundfield components at the receiver location, a listener at this position will have difficulty discerning the reflections as individual events. Nor will the angle of arrival be obvious. As such, the diffuse field is generally considered perceptually constant throughout a sonic environment. When a sound source stops emitting energy into the acoustic field, the diffuse

sound components will decay exponentially over time giving rise to a reverberation tail that acts to colour the reception of the sound generated. This is an important acoustic feature that is exploited in many audio applications to enhance the performance of speech and music. Figure 2.13 provides a graphical summary of the 3-stage temporal model of sound propagation in an enclosure in response to a sudden excitation. A more rigorous extension of this model is detailed in section 2.3.6.

2.3.2 Reflection and Absorption

In a homogeneous medium, sound waves propagate unimpeded in an outward direction from a sound source. This continues until the wavefronts encounter a change in medium density. Such a change arises at a boundary where the density of a solid surface is far greater than that of air. Upon interaction, the total wavefront splits into two portions, one of which is transmitted into the new density medium while the other is reflected back into the air. Formally, this physical interaction is described by means of an impedance quantity and relationship. Following discussion in [3], the characteristic impedance of air may be calculated as,

$$Z_{air} = \rho c = 414 \text{kgm}^{-2}\text{s}^{-1} \tag{2.64}$$

The impedance of a boundary is defined as the ratio of pressure and velocity normal to the boundary divided by the characteristic impedance of air, [3]:

$$\zeta_\omega = \frac{p}{\rho \bar{u}_n c} \tag{2.65}$$

where ζ_ω is assumed to be, in general, complex in order to facilitate frequency-dependency and phase characteristics. A simple model of reflection at normal incidence provides insight into how this ratio of impedances acts on a reflected wavefront. Consider a 1D, single frequency planewave of amplitude A_o with pressure and velocity component:

$$p(x,t) = A_o e^{j(\omega t - kx)} \tag{2.66}$$

$$\bar{u}(x,t) = \frac{A_o}{\rho c} e^{j(\omega t - kx)} \tag{2.67}$$

This simple model is depicted in Figure 2.14. Upon reflection, the reflected wavefront will have a lesser amplitude than the incident wave due to the fact that some acoustic energy will be transmitted into the boundary medium. It is assumed at this point that the transmitted acoustic wave dissipates as heat or otherwise and no longer has any impact on the acoustic field present in the adjacent air. As such, the amplitude of the reflected wave will be reduced by the factor R - the reflection coefficient. Hence, the

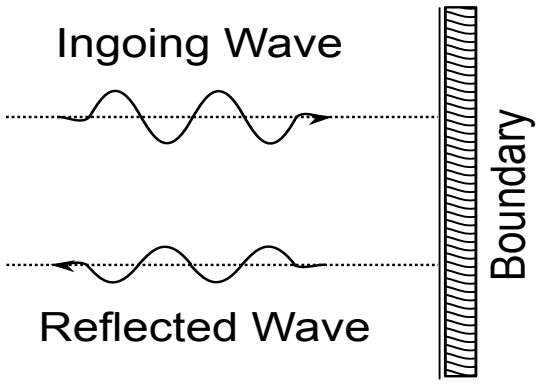

FIGURE 2.14: A simple model of wave reflection for a single frequency wave interacting with a phase-preserving absorbing boundary.

reflected wave components may now be written, [3]:

$$p_{ref}(x,t) = RA_o e^{j(\omega t - kx)} \tag{2.68}$$

$$\bar{u}_{ref}(x,t) = -R\frac{A_o}{\rho c}e^{j(\omega t - kx)} \tag{2.69}$$

noting that the change in sign of the velocity component corresponds to a reversal in propagation direction along the x-axis. By examining the pressure and velocity components at the boundary itself, assuming that the boundary is situated at $x = 0$ for convenience, is it possible to define the impedance of the boundary, Z_b, in terms of the reflection coefficient,

$$Z_b = \rho c\frac{1+R}{1-R} \tag{2.70}$$

This in turn leads to a relationship between the reflection coefficient and the characteristic impedance of the boundary:

$$R = \frac{\zeta_\omega - 1}{\zeta_\omega + 1} \tag{2.71}$$

For realistic systems, the reflection coefficient R is bounded such that $-1 < R < 1$ in order that acoustic energy is always conserved or dissipated at a boundary, never increased. This leads to three distinct extremes of impedance values which correspond to specific boundary types [3]:

- $\zeta_\omega = 0$, $R = -1$: A soft boundary type from which an incident planewave of frequency ω will be fully reflected. Additionally, the phase of the outgoing wave will be inverted due to the negative reflection coefficient.
- $\zeta_\omega = 1$, $R = 0$: A completely absorbing boundary surface that can be approached in acoustically treated anechoic environments. Reflections at frequency ω will not occur under this condition.

- $\zeta_\omega = \infty$, $R = 1$: A fully reflective, phase-preserving boundary type. The reflected wave will differ from the incident wave only in terms of a reversed velocity component.

By including a further spatial degree in equation (2.71), such that the wavenumber k becomes $\hat{k} = x\cos(\theta) + y\sin(\theta)$, the reflection coefficient for reflections at oblique incidence may be defined as follows:

$$R = \frac{\zeta_\omega \cos(\theta) - 1}{\zeta_\omega \cos(\theta) + 1} \tag{2.72}$$

for angles of incidence θ with respect to the boundary normal. From the above expression, it is possible to calculate angle and frequency dependent values for impedance given the corresponding reflection coefficient.

In the majority of the literature, the value used to quantify the absorption characteristics of different surfaces is the *absorption coefficient*. This measure relates the intensity of incident and reflected waves rather than the amplitude. Formally, the intensity of a sound wave in air is given as,

$$I = \frac{p^2}{\rho c} \tag{2.73}$$

for pressure p. Hence, the intensity of a reflected wave will be reduced by $|R^2|$ giving a measure of the incident energy that is dissipated upon interaction with a boundary as $1 - |R^2|$. This is the definition of the absorption coefficient α,

$$\alpha = 1 - |R^2| \tag{2.74}$$

which, like the related reflection coefficient and impedance values, can be frequency and angle dependent.

2.3.3 Scattering

Ideal specular reflection, as defined by Snell's law [26, p. 138], will not occur in practice at all frequencies. Instead, portions of a wavefront incident on a boundary will be reflected at an angle which deviates from the angle of incidence. This concept is called wave scattering. Scattering occurs when a wave encounters a boundary with surface irregularities, as is the case in most real environments. In brief, the relationship between the amount of scattered energy upon reflection and the frequency of a wave is dependent on the size of the surface irregularities. For example, an irregularity of length d metres will have negligible scattering effect on a wave of length $\lambda \gg d$. Conversely, a surface with many small irregularities is likely to scatter sufficiently

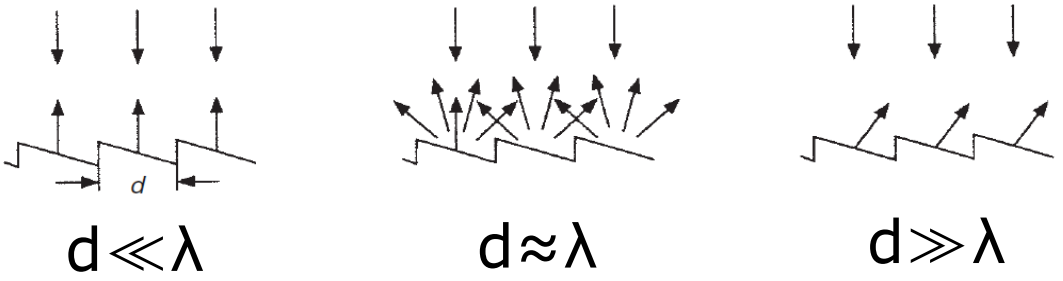

FIGURE 2.15: The general relationship between surface irregularities and scattering of different wavelengths, adapted from [3].

high frequency wave components upon reflection. Kuttruff [3] further expands on this concept through the example shown in Figure 2.15. This diagram demonstrates the low frequency scenario as previously described. Additionally, wavelengths equal to the length of surface irregularities are greatly diffused while wavelengths much less than this length are steered towards a direction normal to the surface orientation.

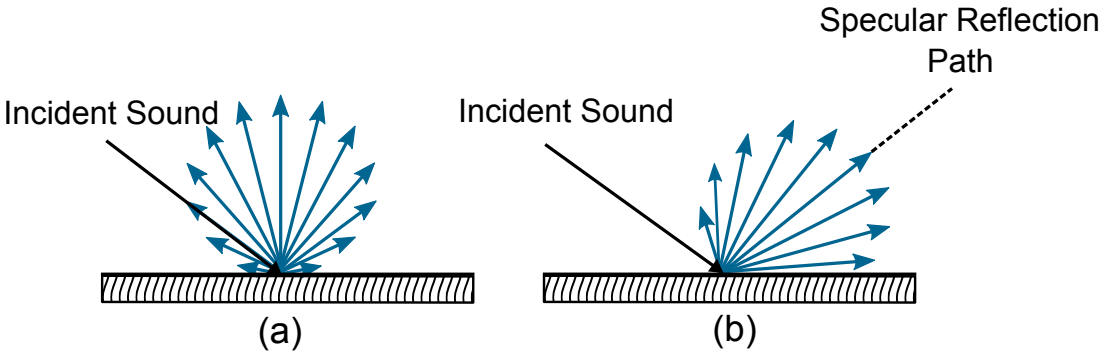

FIGURE 2.16: (From left to right) Visual representation of Lambertion scattering (diffusion) where sound is scattered in all directions (a) and oblique Lambert scattering where the distribution of scattered sound is centered on the angle of specular reflection (b), after [4].

Employing a geometric approach to the investigation of predictable scattering leads to the Lambertian scatterer, as applied in optics [37]. The underlying assumption of the geometric investigation is that sound wavefronts may be represented as 1D paths orientated in the direction of sound propagation. Hence, the frequency-dependency of scattering effects is not preserved in this approximation. As shown in Figure 2.16 (a), a lambertian scatterer distributes reflected energy in all directions proportional to $\cos(\theta)$ where $\theta = \frac{\pi}{2}$ is the surface normal. This is the case regardless of angle of incidence. To further improve this concept and facilitate variance of the amount of diffusion, the oblique Lambert law is commonly employed in commercial acoustic modelling programs such as ODEON [4, 21, 38]. The implementation of the oblique Lambert scattering requires the definition of a scattering coefficient that governs the amount by which the

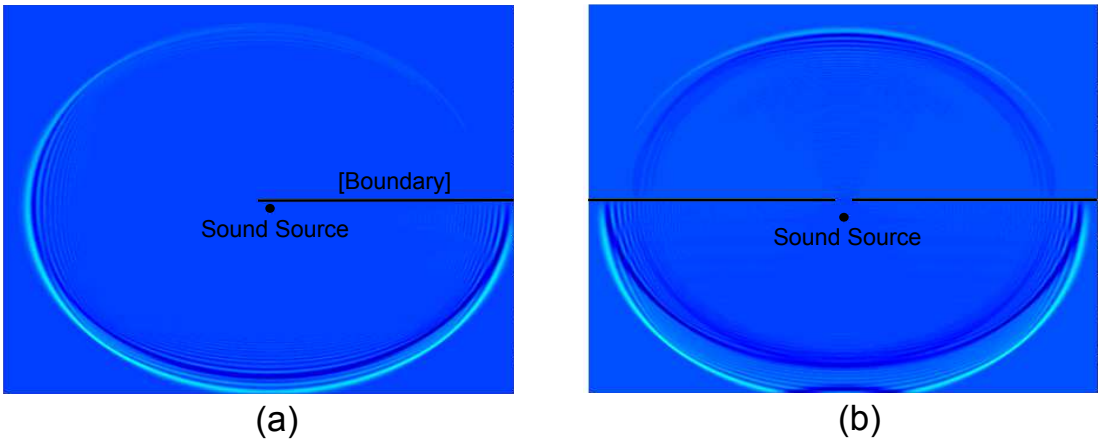

FIGURE 2.17: (From left to right) Diagram (a) displays an expanding monopole sound wave diffracting around a single boundary. Diagram (b) shows a propagating monopole sound wave upon exit from an aperture.

propagation path of a reflection deviates from the angle of incidence. The extremes of scattering coefficients, as applied in geometric acoustic models, are '0' yielding specular reflection and '1' leading to a random angle of reflection as detailed in [39].

2.3.4 Diffraction

As waves progress throughout a medium, they encounter obstacles and boundaries containing the medium itself. While the total interaction process of a wavefront with any object is complex, specific elements are easily observed. Such an element is *diffraction* which, in the field of acoustics, refers to the ability of sound waves to 'bend' around objects that lie in the path of propagation. Diffraction is regularly encountered in practice, for instance, when a listener perceives a sound source located around a corner as arriving from the approximate direction of the corner itself (in absence of any reflections from other surfaces). Figure 2.17 provides visual examples of diffraction as modelled in a 2D wave-based acoustic simulation. Diagram A displays a 2D monopole wavefront interacting with, and diffracting around, a partial boundary. In diagram B, the same 2D wavefront interacts with two boundaries and diffracts upon exit from an aperture between the two obstacles. If diffraction did not occur, the wavefront would continue only along paths perpendicular to the aperture. A more rigorous explanation for diffraction arises from the condition that continuity must be maintained in air pressure over space. As such, the pressure field in the region of an obstacle edge (or corner) must move in sympathy with incoming and outgoing wavefronts. This is further exemplified by investigation of Huygen's principle which states that any wavefront many be represented by a series of point sources located along the wavefront itself. The relationship between this principle and diffraction is shown in Figure 2.18.

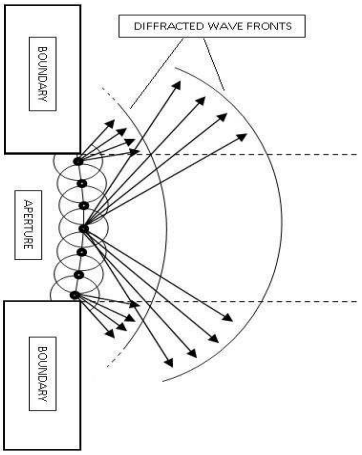

FIGURE 2.18: Graphical interpretation of sound wave diffraction as described by means of Huygen's principle. In this example, a wavefront is radiating from the exit of an aperture between two boundaries.

The wavefront passing through the aperture is approximated by a number of monopole sources. Upon exit from the aperture, the wavefront will expand over space beyond the boundaries reaching propagation paths that are not parallel to the orientation of the aperture sides. Referring back to Figure 2.17, it is apparent that the amplitude of the diffracted wavefront decreases as the diffraction angle increases. This is due to the fact that the diffraction process is both frequency-dependent and related to the size of an obstacle.

Consider, for example, the case of a wavefront encountering a small obstruction as displayed in Figure 2.19 (a). The wave diffracts around the obstruction and proceeds

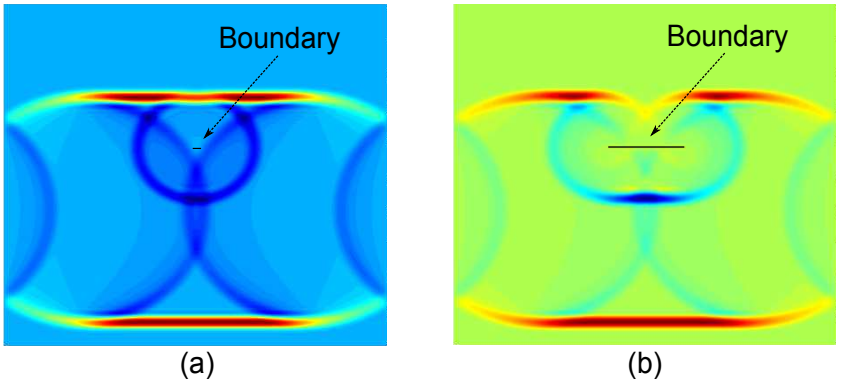

FIGURE 2.19: (From left to right) Diagram (a) shows the case of planewave diffraction around an object when the object length is smaller than the wave length. Diagram (b) displays the same planewave diffracting to a lesser extent due to the increase in obstruction size. In both diagrams, red signifies high pressure values.

relatively unaltered. Conversely, in the case of a larger obstruction (Figure 2.19 (b)), a greater amount of shading occurs leading to a decrease in pressure amplitude of the

wavefront perpendicular to the center of the obstruction. This seeks to demonstrate the relationship between wavelength (and by extension, wave frequency) and the length of the obstruction. As the length of the obstruction increases, the spectral content of the diffracted components decreases to include wavelengths longer than or equal to the obstacle length. Hence, in general, it can be stated that in the study of room acoustics, where the majority of primary boundary dimensions are the length of low frequency wavelengths, diffraction is most prevalent at low frequencies.

2.3.5 Interference

Interference effects arise when multiple waveforms interact. In the free field, this can occur when sounds emanating from two or more sound sources merge in space. In the case of fully or partially enclosed environments, sound waves produced by a single source interact with reflected waves originating at boundary surfaces. Interference may be catagorised as either *constructive* or *destructive*. Constructive interference results when sound pressure waves interact in a region of space to produce a net increase in the amount of air compression or rarefaction occurring in the region. Conversely, destructive interference arises when rarefactions and compressions interact in space producing a net decrease in the extent of compression or rarefaction. The extremes of both interference types are displayed in Figure 2.20 for the case of 1D travelling waves. This concept can be extended to 2D and 3D systems.

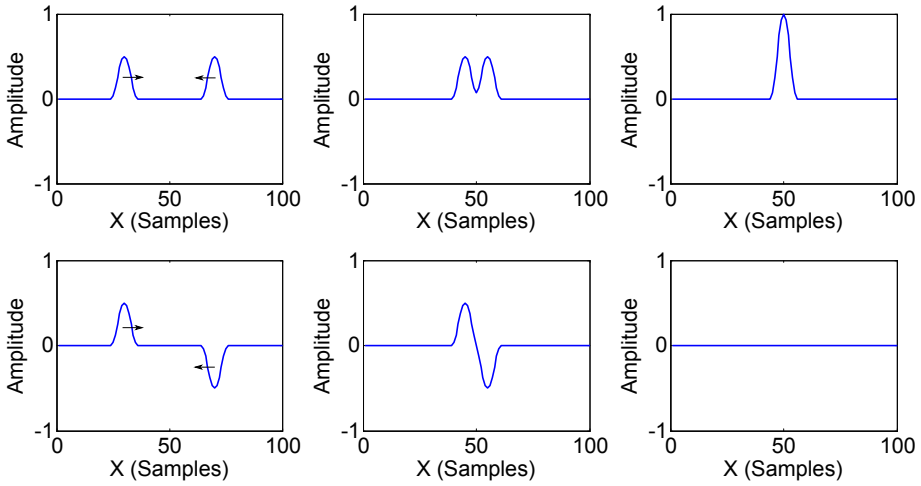

FIGURE 2.20: (First row from left to right) Interaction of positive right-going and left-going travelling waves resulting in constructive interference. (Second row from left to right) Interaction of positive left-going and negative right-going travelling waves resulting in complete waveform cancellation due to destructive interference.

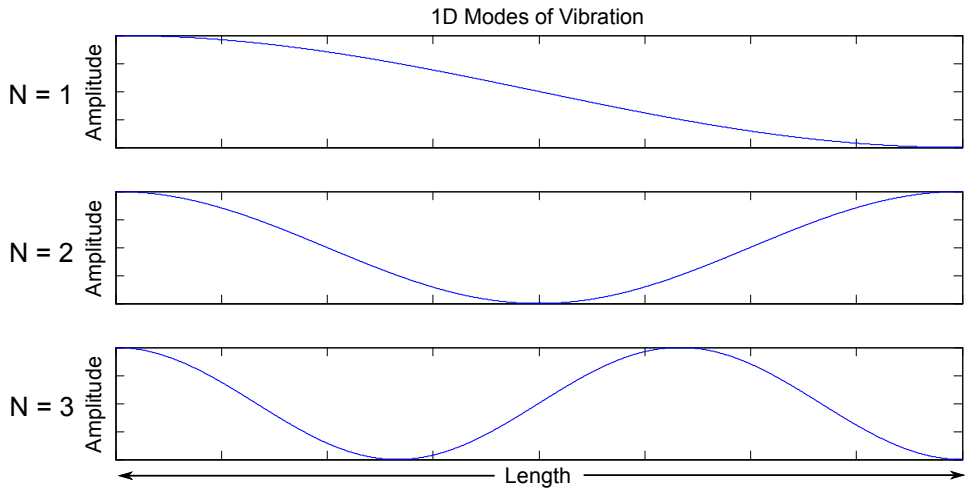

FIGURE 2.21: Modes of vibration (or 'standing waves') between two boundaries separated by a distance 'L' (length). The panels depict the first three modes resulting from cyclical propagation of pressure waves with resonant frequencies.

2.3.6 Standing Waves and Room Modes

Physical systems that exhibit vibrational characteristics resonate at particular frequencies that vary depending on the composition of the system itself. In an acoustic system, resonances are perceived as pronounced frequency components that are spaced in the frequency spectrum to the extent that the pitch of the components may be discerned. These resonances are produced when the spatial dimension(s) of an acoustic environment are directly related to the wavelength of particular sinusoidal sound wave components. For a single spatial dimension, it may be shown (as per Figure 2.21) that constructive cyclic wave propagation occurs between terminating boundaries at frequencies given by:

$$f_{res} = \frac{cN}{2L} \tag{2.75}$$

for a length L (m), wave speed c (ms^{-1}) and an integer value N. Hence, f_{res} is a resultant series of resonant frequencies. An important underlying assumption of (2.75) is that both boundaries are subject to the Neumann ('free') condition facilitating phase-preservation. Upon phase-preserving reflection at the boundaries, as is usually the case in room acoustics, the reflected waveform is exactly in phase with waveforms that exist in the space for every instance of f_{res}. Therefore, the spatial positioning of compressions and rarefactions will appear stationary yielding a *standing wave*. This is a particular type of constructive interference that facilitates a build up of sound wave energy at resonant frequencies leading to large pressure fluctuations.

An example of a 1D acoustic system that exhibits this resonant behaviour is that of

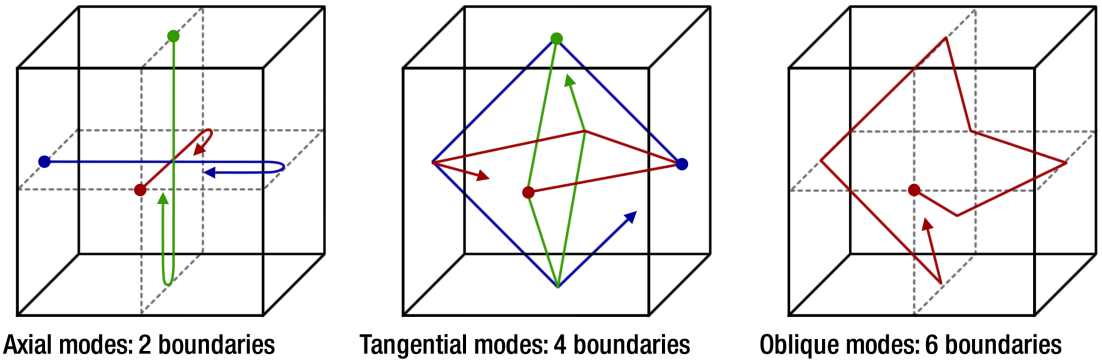

FIGURE 2.22: Graphical depiction of the cyclical modal sound wave paths throughout a simple cuboid space. Examples of Axial, Tangential and Oblique modes are given. Image sourced directly from [2].

an air column, or pipe, with two open ends. Changing the boundary condition at one end to the Dirichlet, 'fixed', type yields the case of a pipe with one closed end or *closed flue pipe*. This adaption of boundary type impacts the pattern of interference arising in the system. From (2.75) it is evident that resonant frequencies are integer multiples of half wavelengths that are equal to the length of an open end pipe. Conversely, it may be shown that resonances in a closed flue pipe are odd integer multiples of quarter wavelengths that are equal to the pipe length [1].

The notion of standing waves can be extended to 2D and 3D where resonances arise between a number of surfaces. For simplicity, examples are given for parallel boundaries in 3D space where resonances (or *room modes*) are categorised as 'Axial', 'Tangential' or 'Oblique'. Axial room modes manifest in a single dimension and result from cyclic wave propagation between two parallel surfaces. Tangential modes occur over two dimensions between four surfaces and oblique modes arise over three dimensions and six boundaries. Examples of these cyclic propagation paths are given in Figure 2.22. In such simplistic topologies, it is possible to analytically derive the frequencies for each type of room mode [1, 24]:

$$f_{n_x,n_y,n_z} = \frac{c}{2}\sqrt{\left(\frac{n_x}{L}\right)^2 + \left(\frac{n_y}{W}\right)^2 + \left(\frac{n_z}{H}\right)^2} \tag{2.76}$$

where f_{n_x,n_y,n_z} is the series of modal frequencies (Hz), L, W and H are the length, width and height (m) of the enclosure respectively and n_x, n_y and n_z are integer values corresponding to the number of half wavelengths existing between the bounding surfaces. Note that for phase-reversing boundaries, a condition exists such that n_x, n_y and n_z must be greater than zero. Such a condition does not apply in the case of phase-preserving boundaries allowing a 'DC mode' to amplify within the system. Room modes also occur in sonic environments of more geometric complexity, however deriving an

analytical means of calculating the resonant frequencies by analytical methods becomes very complicated, if not impossible. Furthermore, room modes are generally regarded as a low frequency sound phenomenon. This is due to the fact that the separation of modes in the frequency spectrum is larger at low frequencies. As frequency increases, this separation becomes less apparent as modal frequency values become increasingly coincident in the spectrum. As such, resonant sound behaviour in 3D enclosures, at higher frequencies, results as the combination of various modes of vibration which are difficult to distinguish. Schroeder and Kuttruff [40] provide a statistical means for calculating an approximate critical frequency, F_c below which a space may said to produce distinguishable modal activity as follows:

$$F_c \approx 2000\sqrt{\frac{RT_{60}}{V}} \tag{2.77}$$

where, V (m^3) is the volume of the space and RT_{60} (s) is the reverberation time (defined in section 2.3.8).

The experience of room modes will vary significantly as a listener's position within a space changes. This is due to the fact that standing waves are most prominent along the cyclical paths which they traverse. In addition, interaction between spatially distributed standing waves can lead to constructive and destructive interference causing spatially varying locations of modal frequency increase, reduction and cancellation. Hence, the nature of enclosed soundfields is inherently inconsistent throughout space in terms of the spectral content of perceived sound. This presents a problem for room acoustic design where, in most cases, the ideal outcome is to produce a soundfield which consistently compliments the reception of sound at all listening positions. For this reason, modal analysis through virtual modelling methods is extremely useful as it provides insight into how to counteract manifestation of undesired resonances by altering room topologies and surface absorption characteristics.

2.3.7 Echograms and the Room Impulse Response

An important method of analysing the acoustic characteristics of a space may be conducted by recording the pressure fluctuations at a discrete point in response to an acoustic excitation. Ideally, the excitation function should be full bandwidth, containing all audible frequency components with a consistent amplitude. This signal type is the Dirac delta function (2.9), defined previously. In terms of acoustic pressure, the Dirac delta function is an infinitely large and infinitely thin amplitude distribution located at time $t = 0$. This infinitely short burst of pressure constitutes an acoustic impulse. In reality, such a signal cannot be generated and, therefore, approximations

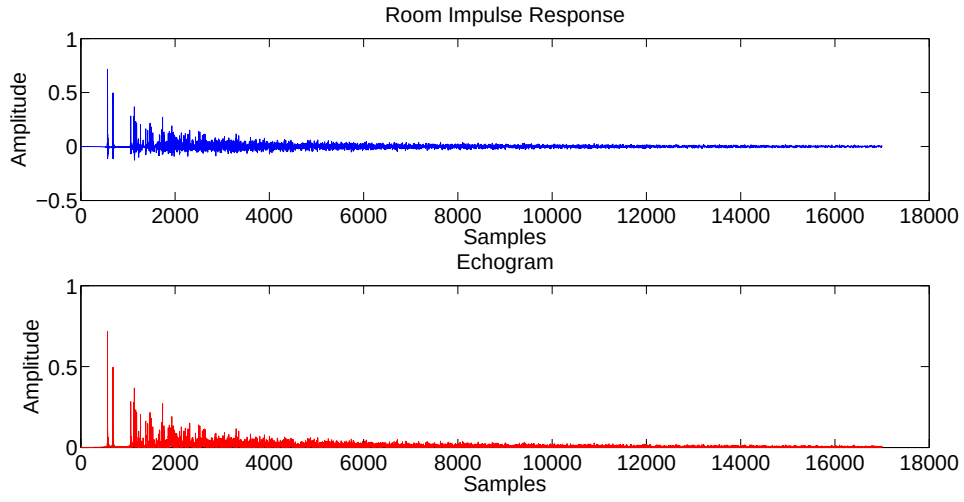

FIGURE 2.23: Examples of a sampled Room Impulse Response (RIR) (top) and an acoustic echogram (bottom).

must be made to realise useful measurements. In accordance with ISO documentation [17], there are best-practice guidelines that detail the equipment and signal types that should be employed. In particular, the sound source should be near omni-directional, meaning that sound is distributed evenly about the solid angle with, ideally, a flat frequency response. In terms of sound capture, the recording microphone should be small as to not interfere significantly with the local soundfield and also possess a flat frequency response. Suggestions pertaining to the nature of the impulsive signal required by [17] have given rise to the creation of a number of sound excitation and capture methods. Among these, the most notable are the *Maximum Length Sequence* [41] and *Exponential Swept Sine* [42, 43] approaches.

The data collected by employing these techniques facilitates the examination of *Echograms* and *Room Impulse Responses (RIRs)*. These two signal types describe the pressure variations recorded, in response to an optimal excitation signal, as functions of time. An example of each type is provided in Figure 2.23.

The plots depicted serve to highlight the difference between the two representations of pressure variation: an echogram contains information on the pressure magnitude variations over time, whereas a RIR contains both magnitude and phase information. Both signal types allow examination of the time-of-arrival and amplitude of the direct sound and significant early reflections experienced at the receiver position. Additionally, the overall nature of the reverberant soundfield and temporal decay of pressure levels may be ascertained. However, RIRs provide a more complete representation of the soundfield due to the fact that sound pressure compressions and rarefactions are distinguishable. Moreover, a room impulse response generated using an optimally flat

sound source defines the approximate transfer function between source and receiver. A RIR of this kind may, therefore, be thought of as the acoustic fingerprint of a sonic environment which, when analysed, yields a wealth of objective measurement data regarding the acoustic qualities of the space. It is for this reason that digital emulation of RIRs is often the ultimate goal in the realm of acoustic modelling and prediction.

2.3.8 Acoustic Parameters

Upon collection of a RIR (or echogram), it is possible to calculate numerous acoustic parameter values that objectively quantify important acoustic characteristics of the space measured. The ISO documentation [17] defines the means by which these parameters are derived from room impulse responses. What follows is a description of particular parameters that are deemed important in relation to the studies detailed in later sections.

Reverberation Time:
There are numerous approaches to the prediction of the reverberation time related to a soundfield. Each method is based on the same underlying principle that the reverberation time is the amount of time taken, in seconds, for sound energy to drop by 60 dB after initial excitation, assuming no further excitation. For this reason, reverberation time is commonly referred to as 'RT_{60}'. In absence of a room impulse response it is possible to determine an approximate value for RT_{60} via statistical means developed by Wallace C. Sabine [44]. According to Sabine, the reverberation time of a space is directly related to the volume of the enclosure V, the total area of bounding surfaces S and the absorption coefficient possessed by each surface α:

$$RT_{60} = \frac{0.161V}{S\alpha} \tag{2.78}$$

The term on the denominator of the above equation refers to the total absorption, given in Sabins (A) calculated as $A = \sum_{n=1}^{N} S_n \alpha_n$ where N is the total number of surfaces, of area S_n metres squared, present within the space. For realistic surface materials, the absorption value will vary with frequency, therefore the value of RT_{60} is also frequency-dependent. In order to compensate for this variance, both absorption characteristics and RT_{60} are calculated in octave bands in order to attain more accurate results.

Sabine's equation provides a quick and simple means of predicting sound decay times in an enclosure. However, its employment is subject to a number of conditions. Firstly, the formulation of (2.78) relies on the definition of a 'mean free path' (MFP) length which is the average distance between bounding surfaces. This can be problematic in the case of sonic environments that are radically longer than they are wide as the

MFP will not be representative of either the large or smaller inter-surface distances. Secondly, it is assumed that the soundfield is diffuse, i.e. the sound energy distribution and direction of travel is stochastic. In practice this will not be the case for spaces that exhibit numerous strong early reflections in response to excitation. Lastly, it may be demonstrated that the accuracy of Sabine's formula decreases for increasing values of $\alpha > 0.3$ [1]. As an example, for fully absorbing surfaces, $\alpha = 1$, a reverberation time of zero would be expected, however (2.78) will yield a non-zero value. An alternative statistical means of calculating reverberation time is that provided by the Norris-Eyring equation:

$$RT_{60} = \frac{-0.161V}{S\ln(1-\alpha)} \tag{2.79}$$

where, as before, S is the total bounding surface area and α is the average absorption coefficient. This formula is subject to the same underlying assumptions as Sabine's RT_{60} calculation, however results yielded from its use are more realistic for high values of absorption [26]. For example, the case of full absorption, $\alpha = 1$, returns a reverberation time of zero as expected due to the inclusion of the natural logarithm. As noted in section 2.3.2, surface absorption characteristics are frequency dependent and, therefore, RT_{60} will also vary with frequency. Specific values of absorption coefficients are commonly documented in octave bands to provide an approximation to the variation of absorption levels over frequency. Hence, the calculation of RT_{60} may be conducted in octave bands to provide a frequency-dependent measure of sound energy decay rates.

A more rigorous approach to RT_{60} calculation that derives reverberation time from a captured or synthesized RIR is detailed in [17]. Recorded RIR data must be manipulated into a form that better represents the decay of sound energy over time. This process culminates in the created of an *energy decay curve* (EDC), also known as the 'Schroeder Curve' after Manfred Schroeder's original work on reverberation time calculation [45]. The EDC is calculated as follows, [17, 45]:

$$EDC(t) = \int_t^\infty p^2(\tau)d\tau = \int_0^\infty p^2(\tau)d\tau - \int_0^t p^2(\tau)d\tau \tag{2.80}$$

where $p(t)$ is the recorded RIR giving a pressure p at time t and $EDC(t)$ is the energy decay curve. Interpretation of the integral limits is as follows: $t = 0$ corresponds to the temporal point of sound excitation; $t = \infty$ is the time by which the soundfield local to the capture position has settled to its equilibrium position. Upon calculation of the EDC, the decay curve is converted onto a logarithmic scale to better suit the range of values returned. Figure 2.24 shows an example logarithmic EDC.

A value of RT_{60} is obtained by constructing a linear least squares regression line that

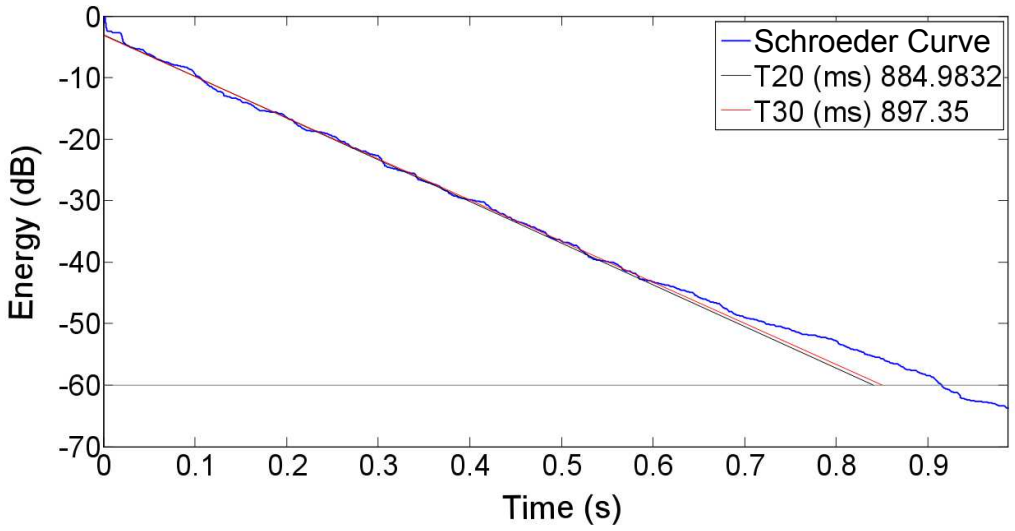

FIGURE 2.24: An example EDC or 'Schroeder' Curve, represented by the blue curve. The extrapolation for T_{20} and T_{30} regression lines are shown in black and red respectively.

approximates the decay of energy over time. Ideally, this regression line would be derived such that the temporal point at which a -60 dB decay has occurred is found. However, in practice, a pressure drop of this magnitude is rarely attained due to the presence of a noise floor or prominent background noise leading to a regression line that is not representative of the decay curve and overestimated reverberation time values. To overcome this problem, the regression line is instead extrapolated from a portion of the EDC defined between lower and upper energy level limits. In this way, the reverberation parameter 'T_{30}' provides the reverberation time calculated by extrapolating the gradient of the regression line in the range EDC $= [-5 : -35]$ dB assuming a normalised initial energy level of 0 dB. In environments with large amounts of background noise or a relatively high noise floor, an alternative reverberation parameter 'T_{20}' may be calculated in an analogous manner using the energy limits, EDC $= [-5 : -25]$ dB. RT_{60}, and indeed every other parameter described in the following, should be evaluated in specific frequency bands. Hence, it is required that RIRs are first band-pass filtered to obtain the response in either one-octave or third-octave bands.

The reverberation experienced in a sonic environment is an extremely important acoustic quality in a subjective sense. A space with large reverberation times may enhance the reception of musical performances if the nature of the space acts to colour the sound produced by instruments in a timbrally pleasing manner. Conversely, there are scenarios where low reverberation times are of great importance, for example recording studio control rooms and other acoustically dry environments. This is due to the fact that reverberation can lead to problems, such as unacceptable loss of source signal fidelity upon reception if RT_{60} values are too large. Hence, this parameter is of great

interest to persons wishing to evaluate or design ideal acoustic conditions for spaces of particular sonic purposes.

Early Decay Time:
The early decay time (EDT) measures the sound energy decay for the initial part of a RIR and is defined as the time taken, in seconds, for the sound energy to decrease to 10 dB below an initial level. EDT is derived by extrapolating the portion of the EDC in the range EDC = $[0 : -10]$ dB (assuming a normalised initial energy level) to a level of -60 dB. Hence, EDT is an alternative value of RT_{60}. The difference between EDT and T_{20} or T_{30} is that the direct sound and early reflections contribute to the EDC portion from which EDT is derived. As such, EDT can provide insight into the strength and temporal width of the non-diffuse soundfield leading to a parameter that is closely related to perceived reverberance [1, 17]. For this reason EDT values may be deemed more psychoacoustically relevant than T_{20} and T_{30}.

Early-to-Late Index:
The early-to-late index, C_{t_e}, is the logarithmic ratio of squared sound pressure arriving at a receiver from the early and late soundfields. Commonly, the early soundfield is defined as containing both the direct sound component and significant early reflections when calculating this parameter. The late sound is that which arrives when the soundfield begins to become more diffuse. C_{t_e} is formally defined as follows:

$$C_{t_e} = 10 \log_{10} \left(\frac{\int_0^{t_e} p^2(t) dt}{\int_{t_e}^{\infty} p^2(t) dt} \right) dB \tag{2.81}$$

where p is the instantaneous pressure recorded at time t and t_e is the temporal limit imposed to differentiate between early and late arriving RIR components. The lower limit of the numerator integral refers to the time at which the direct sound is recorded. The upper limit of the denominator integral, in reality, is set to a time by which sound energy has decayed to a sufficiently low level, usually by at least 30 dB [17]. In a perceptual sense, this quantity describes how intelligible the audio information emanating from a sound source will be to a listener for a given source/listener arrangement within a soundfield. For this reason, C_{t_e}, is commonly referred to as the 'clarity index'. Quantitatively, a high value of this measure calculated from a RIR suggests that the sonic features experienced at the listening position are clear and distinguishable and are without excessive colouration due to the features of the room from which the RIR was collected. In contrast, a low value corresponds to a soundfield in which the fidelity of the audio information produced by a sound source is greatly reduced en route to the listening position. In practical terms, the ratio between early and late arriving energy

provides insight into the diffusivity of the acoustic environment and is, therefore, highly related to reverberation time.

The ISO document [17] provides advice on the temporal limits to apply when attempting to monitor the clarity of a space for different types of sound source. Officially, the Clarity Index is the measure of C_{t_e} obtained when the temporal division t_e is set to 80 ms giving a quantity best suited to evaluating the reception of live music. For the purposes of investigating the clarity of speech, a value of $t_e = 50$ ms is advised. Reflecting on the fact that C_{t_e} is related to RT_{60}, these time limits are justified. Generally, reverberation is understood to support the performance of music and colour the perceived sound in a subjectively pleasing manner. Hence, the perception of clarity in this sense will differ to that of speech where reverberation can degrade understanding in communication and cause annoyance. For this reason, the division between the early and late energies should be altered depending on the nature of the sound being communicated in the space under investigation.

Definition:

Definition, D_{50}, is a similar quantity to the early-to-late index in that it provides a measure of how recognisable the sonic features of a source signal are to a listener when transferred over an acoustic field. D_{50} is defined as follows, after [17]:

$$D_{50} = \frac{\int_0^{50} p^2(t)dt}{\int_0^{\infty} p^2(t)dt} \tag{2.82}$$

In this calculation, the numerator of the above expression is a measure of the acoustic energy arriving at a receiver position in the first 50 ms of impulse response capture (direct sound and early reflections). The denominator integral calculates the total energy experienced at the listening position over the entire duration of the recorded signal. As such, Definition corresponds to the ratio of early arriving energy to the total energy present in the RIR and is presented as a percentage. Subjectively, D_{50} is another advised measure for investigating the intelligibility of speech within a sonic environment. Notice that the same temporal division between early and late arriving sound is used for both D_{50} and the early-to-late index for speech. The close relationship between D_{50} and C_{50} is demonstrated by the fact that one may be calculated from the other:

$$C_{50} = 10 \log \left(\frac{D_{50}}{1 - D_{50}} \right) dB \tag{2.83}$$

When calculated independently from a RIR, C_{50} and D_{50} values can be compared to validate the correctness of results.

Source Strength:
The difference between sound pressure fluctuations at a receiver position and those emitted by a sound source can be expressed by means of the sound strength measure, G. This quantity is a logarithmic value pertaining to the difference between the pressure signal recorded at the receiver, in response to an omni-directional source, and the sound pressure experienced at a distance of 10 m from the same source in the free field (i.e. anechoic conditions). As a function,

$$G = 10 \log \left(\frac{\int_0^\infty p^2(t)dt}{\int_0^\infty p_{10}^2(t)dt} \right) dB \tag{2.84}$$

where p and p_{10} are instantaneous pressure values recorded at time t at the receiver location and in the free field at 10 m respectively. The integral limits, as with the calculation of the parameters described previously, refer to integrating from the start of the RIR signal, $t = 0$, to a temporal position at which the soundfield has decayed sufficiently, $t = \infty$. Through use of G it is possible to discern the nature of sound distribution throughout a measured space by applying the above expression to RIRs captured in various locations. Moreover, comparison of sound levels experienced in different acoustic environments may be conducted in a similar manner.

2.4 Audio Convolution

Any Linear Time Invariant (LTI) expression that describes the nature of an acoustic system may be reduced to a simple input/output system. Figure 2.25 (a) displays such a system with input $X(z)$ and output $Y(z)$, where (z) signifies the z-domain resulting from the discrete z-transform. The transfer function $H(z) = \frac{Y(z)}{X(z)}$ governs the modification of the input to create the output. An impulse response can be found for the system by simply setting $X(z)$ equal to the z-transform of the Dirac delta function (i.e. a discrete frequency-domain signal with constant magnitude at all frequencies). The modification of an input signal by a transfer function signal is the basis of convolution which, when applied to two discrete time-varying signals, results in the combination of both signals. In the discrete-time domain, this process is defined as follows:

$$y[n] = p_1[n] \star p_2[n] = \sum_{m=-\infty}^{\infty} p_1[m] p_2[n-m] \tag{2.85}$$

with y[n] being the resulting convolution of discrete signals p_1 and p_2. Integers n and m reference to discrete time samples. Returning to the concept of the input/output system given above, the temporal process of sound propagation through a space may be described if the transfer function is defined as the RIR of that space. Hence, once the

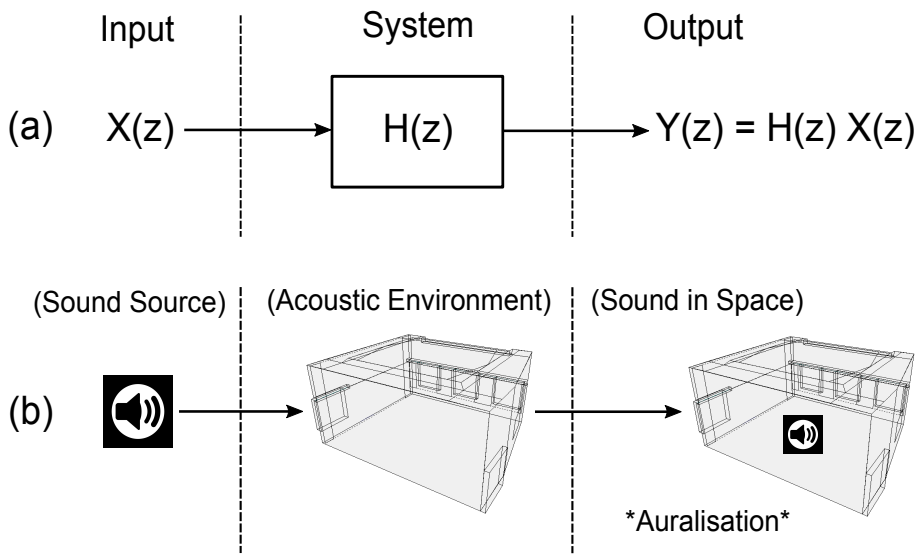

FIGURE 2.25: Diagram (a) shows a simple abstraction of an input/output transfer function system defined in the discrete z-domain. Diagram (b) displays the transfer function concept as applied to room acoustics where the system (i.e. a room) modifies input audio to produce the sound of the input audio as if it were played in the space. The transfer function ($H(z)$) representing the room is the RIR captured in the space.

RIR is known, it is possible to calculate the propagation of any sound source through the space in which the RIR was captured by setting the input equal to the choice of sound source, as per Figure 2.25 (b). This concept is central to the rendering of acoustic auralisations (see section 2.6), allowing any sound to be experienced as if played in the environment from which the RIR was extracted. Furthermore, the discrete z-domain system highlights the potential for carrying out convolution in the frequency domain as follows:

$$y[n] = F^{-1}\{P_1[w] \times P_2[w]\} \tag{2.86}$$

where P_1 and P_2 are discrete frequency-domain representations of the signals p_1 and p_2 with $[w]$ referencing discrete frequency and F^{-1} denotes the inverse Fourier Transform. This is a useful property of convolution that allows for efficient calculation by means of the Fast Fourier Transform [46].

2.5 Elements of the Human Auditory System

2.5.1 The Sound Reception Chain

The human auditory system, under usual conditions, incorporates two ears as sound receivers giving rise to the term *binaural*. This natural sound reception system, as discussed here, is composed of numerous elements which together contribute to the

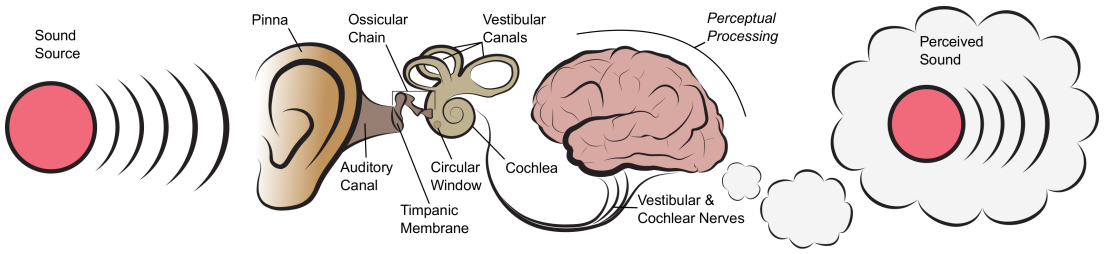

FIGURE 2.26: A simplified flow diagram (from left to right) of the sound reception chain in human hearing. Major physiological components of the auditory system are noted. Image adapted from [5].

sound reception chain. The nature of the physiological components of the reception chain vary from person to person. As such, it is broadly accepted that every person perceives sound in a different way owing to this variance. Perceptual cues, such as localisation and distance, will therefore also vary from person to person as these cues are reliant on the physical attributes of the auditory system.

A simplified overview of the sound reception chain is displayed in Figure 2.26. As shown, a sound impacting on a human ear is first manipulated by the pinna which acts as a direction-dependent filter and channels incoming sound into the auditory canal. The canal is terminated by the timpanic membrane or 'ear drum' and is approximately 0.027 m in length. Interestingly, this arrangement is analogous to a pipe with one open and one closed end and, therefore, the auditory system exhibits increased sensitivity in frequency ranges that correlate with the theoretical resonant frequencies of such a system.

The timpanic membrane defines the boundary between the outer and middle ear segments. This a thin layer of tissue that separates the air in the atmosphere from the internal mechanisms in the head. Under normal conditions, the membrane takes on a concave conical shape, as viewed from the auditory canal, due to the fact that it is pulled taught via the attachment to the ossicular chain. The tension provided facilitates vibration in response to small variations in the pressure of sound waves incident on the membrane.

The subsequent transmission of sound through the reception chain occurs via the ossicular chain due to its fixture to the timpanic membrane. The ossicles are a collection of three connected bones which, together, straddle the middle ear portion and transmit vibrations from the ear drum to the inner ear via mechanical movement. The innermost termination of the ossicular chain is attached to the cochlea via the *oval window*. In its entirety, this system acts to amplify the vibrations experienced by the ear drum due to the smaller dimensions of the oval window compared to the ear drum.

The cochlea is situated in the inner ear and is comprised of two fluid filled chambers separated by the *basilar membrane* which are, together, rolled into a spiral. Vibrations from the oval window pass into the fluid and are transmitted through the length of one of the chambers to the apex of the cochlea. At this point, the vibrations are transferred into the second chamber which terminates with the circular window. Hence, the circular window (a small membrane) moves in sympathy with the oval window as pressure fluctuations in the fluid are passed through the full length of the cochlea. The separating basilar membrane possesses some 30000 nerve endings [1] that sense the pressure variations in the surrounding fluid. Finally, these nerve endings send appropriate signals to the brain via the cochlear nerve in order that the sound may be processed and perceived.

It is apparent that the generation of auditory cues is reliant on the combined effects of each stage of the reception chain. Hence, the modulation of sound pressure waves via the mechanisms present in the system briefly described above produces and relays the characteristic auditory cues that shape the perception of sound on an individual basis.

2.5.2 Sound Localisation

Broadly speaking, sound localisation refers to the ability of the human auditory system to allow the direction of incoming sound to be determined upon reception. This, in turn, enables the perception of spatialised sound attributes whereby the positioning of incoming sound components is perceived for all directions around the solid angle. With reference to binaural hearing, sound localisation is partly characterised by the Duplex theory, initially devised and investigated by Lord Rayleigh [47]. This theory involves two distinct localisation models: the *Interaural Time Difference*, ITD and *Interaural Level Difference*, ILD. These models are briefly described here in order to provide context for the following discussion of spatial sound reproduction formats.

Interaural Time Difference: Due to the distance present between a human's ears, sound incident on a listener from any angle other than 0° and 180° around the azimuthal plane will arrive at the ears at different times. This causes an inherent time delay in the reception of sound at the ear that is located furthest away from the sound source. In turn, this time delay cue is resolved psychoacoustically by a listener producing a perception of the direction from which the sound is originating.

Interaural Level Difference: The presence of the human head between the ears acts to occlude sound from the ear that is located furthest away from a sound source. This occlusion causes a difference in the pressure amplitude received at each ear whereby the level is lower in the ear that is occluded in relation to the source location. This provides

a listener with a further perceptual cue that is resolved to indicate the direction of the incoming sound. However, as discussed in [1], low frequency wave components are capable of diffracting around the head and, therefore, the pressure amplitude difference arising at each ear becomes less perceptible as frequency decreases. As such, it is widely accepted that the human auditory system applies a combination of both ITD and ILD to localise incoming sound. For frequencies below around 1500 Hz, ITD is the dominant acoustic cue. For frequencies above around 1500 Hz ILD is dominant.

2.5.3 Just Noticeable Differences (JNDs)

The study of the perception of sound provides insight into how sensitive the human auditory system is to changes in a soundfield or in received audio. This concept of detecting variances in the attributes of received sound is made quantifiable by means of *Just Noticeable Differences* (JNDs). JNDs correspond to the minimum change in any sound attribute (e.g. pitch, loudness, localisation) that the human auditory system is capable of resolving. Hence, the notion of JNDs may be readily extended to the objective acoustic parameters derived from RIRs (see section 2.3.8) in order to ascertain whether objective differences between two or more RIRs are likely to translate into an easily perceived difference. Such quantities can be useful in the design and analysis of room acoustic properties and so the definition of reliable JND values has been the focus of extensive research (see e.g. [48, 49] and references therein). Currently, JNDs

Perceived Attribute	Acoustic Parameter	Frequency Range (Hz)	JND
Clarity	C_{80} (dB)	500 - 1000 Hz	1 dB
Definition	D_{50} (%)	500 - 1000 Hz	5%
Reverberance	EDT (s)	500 - 1000 Hz	5%

TABLE 2.2: Just Noticeable Difference (JND) values for C_{80}, D_{50} and EDT acoustic parameters as defined in [17]. These values are assumed to be valid over the frequency ranges provided.

pertinent to room acoustic parameters listed in ISO 3384 [17] are generally regarded as the benchmark values to be used when evaluating RIRs in a perceptual sense. A subset of these values are provided in Table 2.2. However, particular attention has recently been given to the derivation of a reliable JND for reverberation time, T_{30} [16]. While a JND for RT_{60} is stated in ISO 3384 [17], this value refers to EDT and is defined as 5% as derived in [50]. Hence, this EDT JND is not readily applicable to the alternative T_{30} measure of reverberation time and is considered too stringent for T_{30} [16]. This consideration is justified by studies that indicate a much larger range for T_{30} JND, for example [16] and [51] which document ranges of +/- 24.5 % and 30 % respectively. For

the purposes of this study, the T_{30} JND is defined as the former due to the fact that this range is shown to be equally applicable to band-limited and full audio bandwidth stimuli [16] and is the more conservative of the two. EDT and C_{80} JNDs are defined in this work as per Table 2.2 in order to maintain consistency with recognised standards.

Following the discussion in [48] it is noted that while JND ranges provide a basis of quantifiable perceptual comparison between audio signals, these ranges are approximate and can vary depending on two notable factors. Firstly, the derivation of JNDs for acoustic parameters documented in related literature relies on the use of different audio stimuli (e.g. noise, instruments etc.) leading to a lack of consistency between reported ranges. Secondly, the ranges of acoustic parameter measures incorporated in studies also varies. As such, there is uncertainty when considering whether the application of a JND range derived using a large parameter range in audio stimuli is equally applicable to both low and high objective parameter measures. Hence, JNDs are considered an approximate guideline for perceptual RIR analysis for the purposes of this work and are applied only to further quantify the relative agreement of modelled RIR parameters obtained from simulations (see sections 5.2-5.3).

2.6 Auralisation

Room impulse responses, captured in real environments, or simulated in their virtual representations, provide the data required to numerically quantify acoustic characteristics such as reverberation and clarity. Although these measures are significant in the realm of acoustic design and analysis, they do not provide a subject with a *sense* of the auditory nature of soundfields. This is where *auralisation* becomes central to the experiential evaluation of room acoustics. In [52], Kleiner et al. aptly define auralisation as follows:

"Auralisation is the process of rendering audible, by physical or mathematical modelling, the soundfield of a source in space, in such a way as to simulate the binaural listening experience at a given position in the modelled space."

The desired outcome of this procedure is the relation of an aural impression of the acoustic characteristics inherent to a given space in response to a sound source (or sound sources) to a listener. Approaches to auralisation may be divided into two categories: *Fully Computed Auralisation* and *Computed Multiple-Loudspeaker Auralisation*, after [52]. The first category relies on the calculation of RIRs by means of a virtual modelling approach. The RIR generated is then processed to create an auditory representation of a sound source in the modelled space suitable for playback over headphones.

Computed multiple loudspeaker and Fully Computed Auralisation approaches differ only in the means by which the modelled soundfield is presented to a listener. In this case, the sound reproduction system consists of an array of loudspeakers arranged in a manner governed by the type of sound capture method used (i.e. B-Format, see section 4.3.2). This gives the approach a considerable advantage in that a more immersive, natural listening experience is attained owing to the directivity inherent to multi-channel speaker setups and the removal of the requirement for listeners to wear headphones. However, auralisations of this kind must be undertaken in anechoic or suitably treated listening environments in order to avoid colouration of the presented audio. In addition, the loudspeakers used must also adhere to particular restrictions. As stated in [52], they must be: small as to avoid undesirable scattering effects that may give rise to increased diffusivity of the rendered auralisations; full-range such that the entire audible bandwidth may be represented.

2.6.1 Walkthrough Auralisation

This auralisation procedure involves the reproduction and presentation of a dynamic auditory scene such that a listener may experience the evolution of a changing soundfield. The progressive alterations commonly applied in walkthrough auralisations include: moving sound source and receiver positions; adjustment of room geometries; modifying model parameters such as surface absorption and scattering characteristics. Various implementations of walkthrough auralisations are described in related literature (e.g. [53–61]). These examples demonstrate rendering interactive auralisations in either *online* or *offline* processes.

The offline approach, i.e [55, 59], relies on the pre-computation of a selection of varying RIRs that differ with respect to the changing aspects of a soundfield or virtual model. For example, a moving listening location may be represented by a series of RIRs recorded at minimally spaced instances along a predetermined route throughout the measured space. The impression of movement, in relation to a static sound source, may then be achieved by interpolating between RIRs as required during the auralisation process. Alternatively, a moving source position, or a combination of moving source and receiver, can be handled by similar means. The advantage of this approach, when using virtual models to produce RIRs, is that no time constraints are imposed on simulation and rendering of results. Therefore, accurate, yet computationally expensive, wave-based or hybrid methods (see Chapter 3) may be utilised to produce a set of detailed RIRs, thus enabling the production of highly realistic auralisations.

The online approach, i.e [54, 60, 61], is by far the more difficult of the two to implement. This is largely due to the requirement that alterations of the modelling scenario, and their consequent effects on the auralisation, must be carried out in real-time. As a result, restrictions on the level of detail and accuracy of the chosen modelling technique are usually implied to seek a compromise between between simulation speed and resulting realism of auralisations. Implementations such as those documented in [53, 54, 57] operate by accurately modelling early reflections and synthesizing the reverberant stages of sound by means of estimation or delay and filtering networks. The simplification in the representation of diffuse soundfields is justified by the listener's inability to distinguish separate reflection components that occur closely in time. Reduction in levels of geometric complexity may also be considered to lower computational load [54]. This is particularly useful when employing the Image Source Method (see section 3.1.2) to compute early reflection paths where simulation times are prohibitive for large numbers of model surfaces.

There is a considerable benefit that arises from the use of real-time auralisations. The introduction of user-interaction with such systems allows a listener to specify changes to model attributes and immediately receive a means of perceptually monitoring the effects of these changes on acoustic characteristics. Recently, such systems have been developed to inform the design of buildings in terms of acoustic properties [60, 61]. In the alternative offline case, a new set of RIRs would be required for each alteration made to the model. Hence, walkthrough auralisations constitute a highly useful extension to static auditory scene reproduction in the context of all the related applications of acoustic modelling outlined in Chapter 3. This is especially true in cases where comparative testing is desired for subjective evaluation purposes.

2.7 Summary

This chapter has provided a review of acoustic theory and concepts that are related to the studies documented in the following. The nature of sound waves in air has been investigated and the governing mathematical expressions for wave propagation in 1D - 3D space have been derived and analysed. A review of important acoustic phenomena that arise in real enclosed sonic environments has been documented in order to highlight key characteristics of sound that an acoustic simulation should aim to model. The discussion of RIRs addresses the ISO standardised approach to the calculation of acoustic parameters that define the sonic properties of a space. Lastly, the concept of audio convolution has been introduced in relation to the production of auralisations which are, in turn, described alongside elements of the human auditory

system and subjective JND ranges for ISO acoustic parameters. The topics reviewed here serve as a basis for analysis and explanation of results derived from virtual acoustic models documented in later chapters.

Chapter 3

Virtual Acoustic Modelling Methods

The primary aim of virtual acoustic modelling is to digitally represent the behaviour of sound propagation within a sonic environment ensuring faithful recreation of the acoustic characteristics inherent to that environment. In most cases, the process of virtual acoustic modelling involves simulating a room impulse response. As discussed in Chapter 2, a RIR reveals the acoustic properties of the space in which it is captured. Hence, the ability to simulate RIR signals is significant as it provides a means to investigate spaces that are under development, inaccessible, no longer in existence, theoretical or in initial design/development stages. For this reason, acoustic modelling processes find application in a number of disciplines including:

- Architectural Design: acoustic prediction informs the design of spaces where acoustic qualities are of importance (e.g. concert halls, theatres and recording studio control rooms).
- Environmental Noise: the impact of sound barriers and other approaches to noise control and reduction may be investigated prior to implementation to assess and optimise performance.
- Virtual Reality: the creation of realistic reverberation effects can enhance the sense of immersion experienced in the virtual realm. This notion also extends to audio in film and game production.
- Architectural Heritage: the acoustics of a site of architectural significance may be simulated such that it is possible to provide and contribute to a complete account of the properties of the site for posterity.

There exists a wide range of different approaches to virtual acoustic modelling, all of which have their strengths and weaknesses. These approaches may be divided into two distinct categories: Geometrical and Numerical. What follows in this chapter is a review of select acoustic modelling paradigms that have been investigated during this study to inform the design and development of the novel RIR synthesis method presented in Chapter 5.

3.1 Geometrical Acoustic Models (GAMs)

Geometrical approaches include the most established acoustic modelling paradigms with a history of digital implementations that dates back to the 1960s [62]. All GA models are constructed with the assumption that sound propagates throughout space as a series of 1D paths of traversal or "rays". While this underlying principle contributes to the efficiency of simulations, it also impacts negatively on the accuracy of the results produced. This is due to the fact that wave properties are not preserved in such simulations leading to inaccuracies in low frequency ranges where wave phenomena (such as standing waves, diffraction and occlusion) are most apparent. For this reason, it is generally accepted that GA models are valid only at mid-high frequency ranges where wavelengths are much shorter than the dimensions of the surfaces that compose the modelled space.

Despite this general limitation and recent developments in more modern numerical techniques that preserve wave effects, GAMs remain a highly useful and popular means of room acoustic modelling as noted by Savioja and Svensson [63]. This claim is supported by the current dominance of GAMs in commercially available acoustic prediction programs such as CATT-Acoustic [19], ODEON [21] and EASE [64].

3.1.1 Ray Tracing

One of the earliest examples of the application of ray-tracing to the problem of acoustic prediction was produced by Krokstad et al. [62] in 1968. This work defines and examines the basic principles of ray-tracing methods that continue to inform the design of current acoustic prediction techniques. The ray-tracing algorithm begins by defining a sound source location in a given space, from which a number of 1D ray paths are projected into the space. These rays represent the progression of sound energy along the paths, the directions of which are distributed (approximately) uniformly around the solid angle. As the ray paths emanate from the sound source and traverse the space they interact with boundary surfaces and are consequently reflected according

to the laws of geometrical reflection. In [62], the ray paths terminate upon arrival at a receiver area, defined as a rectangluar audience area in the space. Upon arrival at the receiver area, the time of arrival (defined as the time elapsed since excitation) is recorded. The recorded data takes the form of an echogram that can be utilised to describe the temporal nature of the soundfield relative to a predefined listening position. Note that echograms may be further processed to include synthesised/approximated phase components in order to yield a room impulse response [63].

Since the early work by Krokstad et al., ray-tracing algorithms for acoustic modelling have undergone significant developments to increase the accuracy of simulations. In [65], Kulowski examines the implementation of sound reflection/absorption, emission and reception in ray-tracing models which together facilitate more accurate emulation of sound propagation in space. Following [65], the approach to omni-directional source modelling is refined by considering two scenarios. Firstly, if the number of rays cast is known, outgoing rays may be defined to yield a uniform distribution over the solid angle. Alternatively, a random distribution is applied when the number of rays to be cast is unknown. The second situation may arise when the amount of ray paths required to investigate a space is determined during run-time as the results produced approach and eventually attain a predefined condition. Furthermore, for the case of a directional sound source, the initial energy applied to outgoing rays can be weighted with respect to a desired source directivity function.

The attribution of initial energy to each ray cast facilitates the simulation of absorption due to the absorptive characteristics of boundary surfaces. Upon each ray reflection, the amount of energy dissipated in the surface encountered is calculated and subtracted from the incident energy level. Hence, successive reflections over time cause a decrease in the total energy of a ray. Commonly, the amount of absorption applied during ray reflection is assumed to be independent from angle of incidence. This assumption holds for diffuse soundfields and may be utilised for calculating absorption for late reflections as applied in an early implementation of ODEON [66]. Furthermore, the length of a calculated ray path may be used to calculate frequency-dependent energy loss due to the viscosity of air. If a ray dissipates enough energy to fall below a defined minimum before arriving at a receiver, the traced path will be terminated and discarded.

The reception of a ray at a receiver point occurs when the ray path intersects a volume located around the receiver location. Point-like receivers are not applied in ray-tracing due to the low probability that a suitable number of rays will intersect the receiver's position in a reasonable run-time. However, the use of receivers which occupy a defined amount of space leads to the potential for invalid or duplicated ray paths to register

as contributions to the resulting echogram. This problem is addressed in [67] where examples of invalid ray paths are given as follows:

- Multiple detections of ray paths emulating the direct sound component or early reflections.
- Ray paths between source and receiver that are obstructed by elements of the room geometry.
- Rays that are detected on the opposite side of an interior boundary with respect to the receiver due to the fact that the receiver volume encompasses both sides of the boundary.

The first example can be remedied by tracing a series of pyramidal volumes between source and receiver for direct sound and early reflection paths, as is the case with beam tracing algorithms (see section 3.1.3). This reduces the potential for casting several rays towards the receiver and instead the sound components arriving at the receiver are recognised by an intersection of the receiver volume with the volume cast.

In order to identify invalid rays that are in some way obstructed by boundaries, it is possible to reduce the volume of the receiver to a minimum allowable size. The governing equation for this minimum value is documented in [67] for the case of sphere. A further solution, which may be used in conjunction with the first, is to trace a reverse path from the receiver to the source. If a backtracked ray path fails to arrive at the source at the correct time (the time of arrival), the ray is discarded. While this is a rigorous means of ray validation, this approach does incur increased computational effort as each ray path must be traced twice.

Ray tracing is currently employed in a range of acoustic modelling programs, including CATT-Acoustic [19] and ODEON [21], as a means of simulating late reflections in diffuse soundfields [63]. The inclusion of diffuse reflection in ray tracing was examined by Hodgson [68]. A previous work by the same author, [69] demonstrates that accounting for only specular reflections in a ray-based acoustic prediction model produces results which are likely to deviate from those recorded in practical measurements. The discrepancies were hypothesized to be a consequence of the presence of diffuse sound reflections occurring in practice and their omission in the models. In [68], Lambertian scattering was applied in order to simulate the effects of diffuse reflection and, in turn, this led to the notion of controlling the amount of diffuse reflection arising in the later stages of a room impulse response. The incorporation of diffusion is shown to lead to the prediction of reverberation times that better match those found in practice [68]. Further refinement of the methods by which diffuse soundfields are simulated in

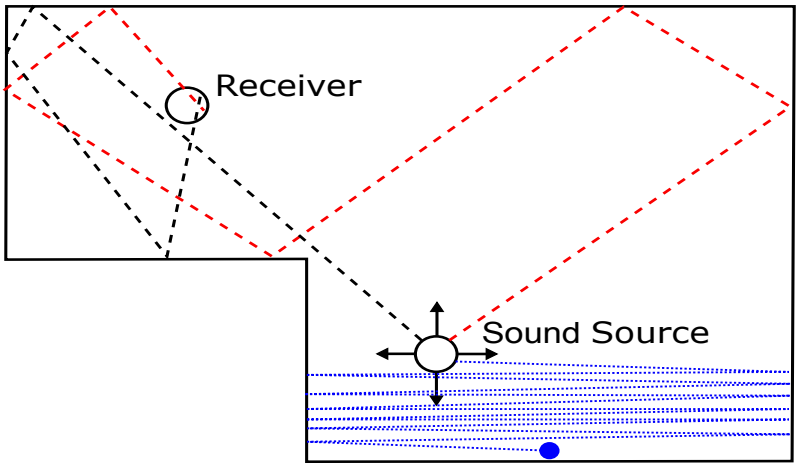

FIGURE 3.1: 2D depiction of rays traced throughout a simple domain. The ray casts signified by the red and black lines traced from the sound source are cases of received rays as they intersect the receiver domain. The blue ray cast is an example of one which does not arrive at the receiver before imposition of the termination criteria.

ray-based algorithms are given in [70, 71] and [71] in relation to ODEON specifically. While GAMs do not inherently emulate sound wave characteristics, diffraction modelling has been introduced to ray-tracing algorithms in order to increase the accuracy of auralisations. A study by Torres et al. [72] demonstrates that synthesis of diffraction in the early stages of a RIR contributes significantly to the perception of spatial and low-frequency qualities within a model. However, the inclusion of diffracted ray components in simulations can greatly increase computation time.

Having examined the ray-tracing technique, it is possible to conclude that this method is well suited to the simulation of the late temporal stages of sound propagation in mid-to-high frequency bands. However, due to the nature of ray casting, reflection and reception in ray-tracing algorithms, there is the potential for valid paths to be omitted during simulation. This, in turn, suggests that ray tracing is not reliable for simulating prominent early reflections. Hence, this technique is commonly used in combination with the image source method which directly calculates specular early reflection paths.

3.1.2 Image Source Method

The theory behind the Image Source Method (ISM) as applied to acoustics was presented by Eyring in [73] where the process was employed to predict the reverberation time of acoustically treated rooms. This approach was then digitally implemented many years later by Allen and Berkeley [74] to simulate RIRs in spaces with rectangular, parallel surfaces. The case of such cuboid topologies yields the simplest form of ISM and is best described with the aid of the diagram depicted in Figure 3.2.

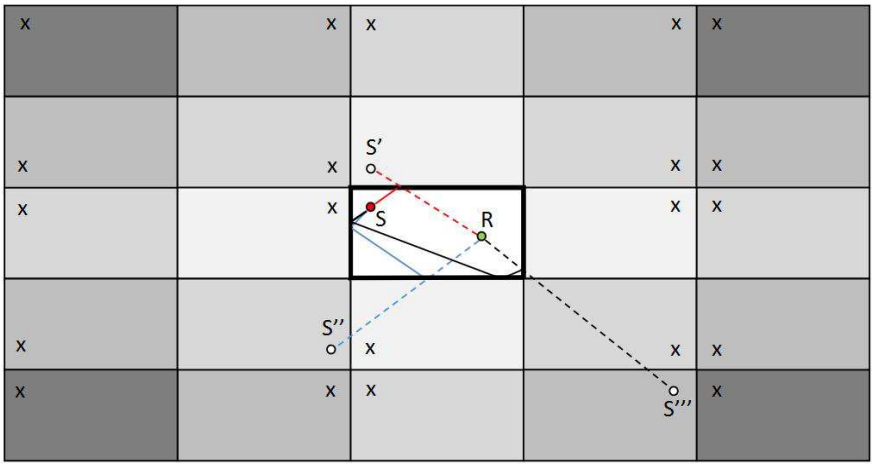

FIGURE 3.2: 2D graphical interpretation of the Image Source Method. Source (S) and receiver (R) positions are shown within the boundaries of a simple rectangular space. Examples of 1^{st} - 3^{rd} order reflections are given alongside their respective image sources. Further examples of valid image sources are denoted 'x'.

The diagram shows a 2D plan view of a cuboid space with a sound source location 'S' and receiver location 'R' situated within the boundary surfaces (shown in bold black). As with all GAMs, the propagation of sound energy is assumed to occur along a number of 1D paths originating at the sound source. However, in the ISM each reflection occurring in a path is represented using a secondary *image* source. Examples of first, second and third order reflections are displayed in the diagram with their respective image sources, S' - S''', each positioned in a fictitious space with dimensions equal to those of the original space. The path length between an image source and the receiver is equal to that of the reflected path to be represented. In addition, the angle of arrival of each reflection at the receiver is preserved and the reflection order is given by the number of boundaries intersecting the path from the image source to the original space.

It is evident from Figure 3.2 that a symmetry exists in the positioning of image sources for cuboid room topologies. This arrangement of contributing image sources is the result of the underlying assumption in the ISM that all reflections are specular. Defining a lattice of image sources in this way yields an efficient algorithm that guarantees the simulation of all specular reflection paths between source and receiver up to a given order. In terms of sound energy decay, absorption due to surface reflections and the viscosity of air may be applied in a manner similar to that found in ray-tracing methods. Furthermore, the recursion of the ISM algorithm produces an echogram from the direct sound and reflection paths recorded at the receiver over time. This recursion ceases when a prescribed condition is met, for example when a defined reflection order is attained or when reflected components have decayed below a given sound energy level.

The development of a more rigorous implementation of the ISM, which is capable of handling arbitrary enclosures, is described by Borish [75]. This work exposes two prominent challenges that arise when applying the ISM to complex room geometries. Firstly, defining the locations of image sources that reflect across surfaces of arbitrary orientation in space requires a greater number of mathematical operations in comparison to the simple case previously discussed. This is due to the fact that the repeated image source structure (lattice) will not occur for any non-cuboid topology. Naturally, this gives rise to an increase in computation time. Secondly, the recursive definition of image sources is subject to validity checking in order to ascertain whether a given reflection is visible to the receiver. In brief, this iterative operation progresses by reflecting each located image source across every boundary surface, which, in turn, generates a new set of image sources. A resulting reflection path is deemed invalid if it is either not visible to the receiver via successive specular reflection or in excess of a defined maximum path length.

For a given room topology, it may be stated that $\mathrm{O}(N^r)$ image sources are required to accommodate a reflection order 'r' in an enclosure composed of 'N' boundary surfaces [76, 77]. For arbitrary geometries, the same number of validity checks must be undertaken for each image source. It is, therefore, apparent that increases in geometric detailing and required reflection order can lead to prohibitive simulation times. This prompts model design challenges when attempting to approximate structures such as curved surfaces, which must be represented as a series of planar surfaces [75], leading to proposed methods of geometry simplification to reduce computational load [78].

The ISM is capable of predicting specular reflection paths and, as such, is well suited to the calculation of strong early reflections arriving at a receiver position. In this way, the ISM is deemed more accurate than ray-tracing techniques in which strong specular reflections may be omitted during simulation. However, as previously noted, this accuracy comes at the cost of prohibitive computational requirements for high order reflections and successive specular reflection rarely occurs in practice. Hence, the ISM is commonly applied to the calculation of the early arriving sound contributing to a RIR. Restricting the utilisation of this method allows for low-order early reflections to be computed at interactive rates [79] allowing alterations in the acoustic model to be carried out in real-time. This process may then be supplemented by applying an alternative, more efficient and representative means of acoustic modelling to the simulation of the diffuse soundfield, creating a hybrid modelling method capable of rendering temporally complete RIRs (see e.g [66, 80, 81]).

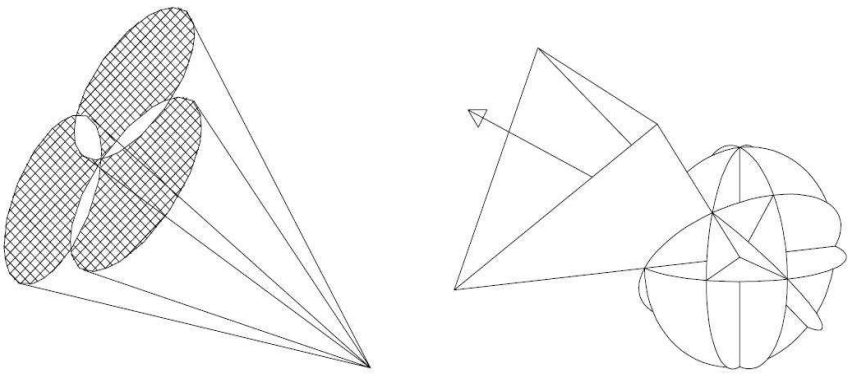

FIGURE 3.3: Visual comparison of conical and pyramidal beams projected from an omni-directional sound source. Image sourced directly from [6].

3.1.3 Beam Tracing

Beam tracing provides a means of optimising the ISM by removing a large quantity of the aforementioned visibility checks that are required to validate sound propagation paths in complex enclosure geometries. This process is conducted by determining approximate valid paths by recursively tracing beams (or rays with volume) throughout the modelled space during a pre-computation stage. Potential propagation paths are then revealed upon the intersection of a receiver location with a beam volume. In this way, a number of invalid image sources that are not visible to the receiver location are ignored entirely. The application of several adaptations of this optimisation procedure in the realm of virtual acoustics modelling are detailed in [6, 76, 77, 82, 83]. A significant difference between the various implementations discussed in the literature is the form of which the beams take. Figure 3.3 highlights a comparison between *conical* and *pyramidal* beam tracing approaches. This diagram demonstrates the advantage of tessellating an omni-directional sound source with pyramidal beam volumes over conical volumes. In the case of conical volumes, a certain degree of overlap is unavoidable when all path directions are to be represented around the solid angle of a sound source. Consequently, multiple detections of valid propagation paths can occur giving rise to imbalances in sound energy levels emitted into the modelled space [6]. This problem is overcome by projecting pyramidal beams from a source position such that the solid angle is completely represented while overlapping beams are avoided. However, even with the use of pyramidal volumes, there is the potential for spaces within an enclosure to remain uninvestigated. Studies, detailed in [82, 83], describe a technique in which beam topologies are derived from the shape of the sounding surfaces. This 'adaptive' beam tracing technique ensures a geometrically exact reproduction of the soundfield [82]. In order to provide a brief overview of a general approach to beam tracing, Figure 3.4 displays a simple modelling scenario. A single beam is shown to originate from a

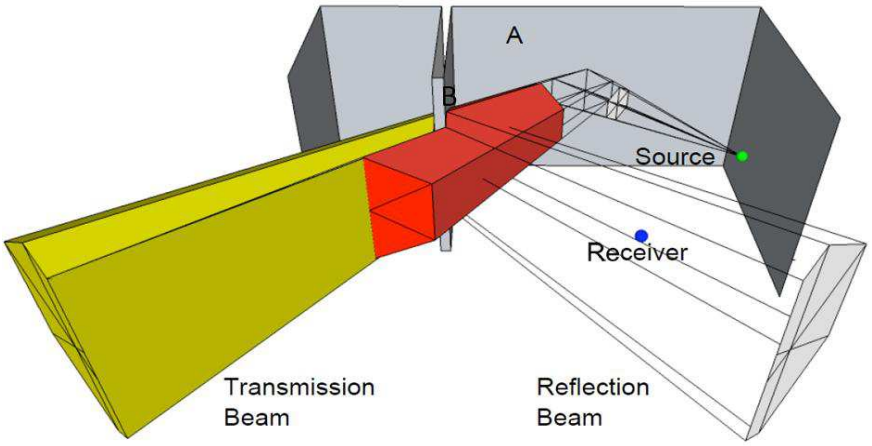

FIGURE 3.4: The traversal of a single beam throughout a 3D space with examples of entire (1^{st} order - surface A) and partial (2^{nd} order - surface B) reflections. Image sourced directly and adapted from [7].

defined source location and undergo consequent interaction with a 3D geometry. Upon interaction with an obstructing surface, the portion of the beam in contact with the surface is removed to form a shadow region. This shadow region, or *transmission beam*, is reflected in the plane of the surface to create a *reflection beam* [76]. Examples of entire (surface A) and partial (surface B) obstruction are provided. The recursion of this process is repeated until a maximum path length, or alternative termination condition, is attained.

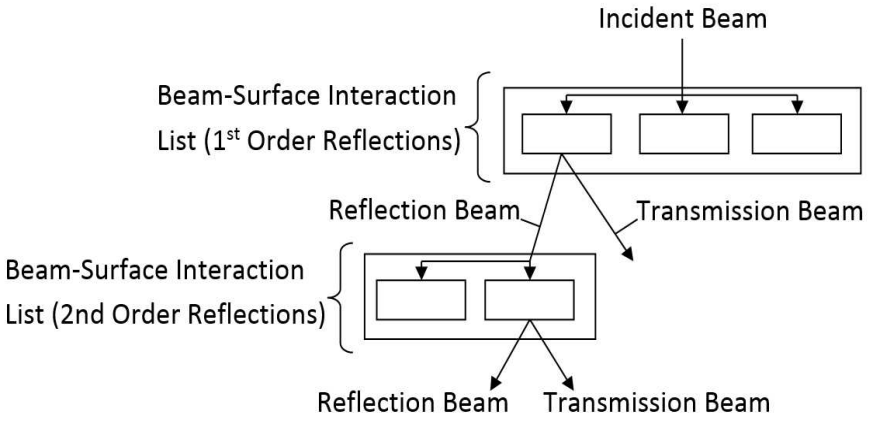

FIGURE 3.5: Example beam-tree (partial) highlighting the fragmentation of a beam incident upon multiple surfaces, after [8]. In this way, approximate propagation paths are mapped and stored for later use by the ISM.

This recursive process defines a series of potential (volumetric) propagation paths in which valid image sources are located and is applied during pre-computation. The result of this pre-computation stage is referred to as a *beam-tree*, an example of which is given

in Figure 3.5. In brief, a beam-tree is a hierarchical data-structure that stores the details of the approximate valid propagation paths between source and receiver [76]. Upon completion of the beam-tree, the ISM is then applied to detect exact specular reflection paths through investigation of the volumes mapped by the beam-tree. This acts to reduce the amount of geometric space that the ISM is required to examine and ensures omission of invalid image sources thus removing the need for validation procedures. The resulting optimisation of the ISM has been proven capable of rendering interactive auralisations for large-scale auditory scenes incorporating multiple stationary sound sources and a moving receiver location [76].

3.1.4 Acoustic Radiosity

Radiosity methods are derived from a different basis than that of the techniques previously described. As opposed to tracing discrete paths of energy propagation throughout a volume, radiosity approaches seek to simulate the transport of energy between the surfaces of a bounded enclosure. The acoustic radiosity (AR) method is a time-dependent implementation of the radiosity approach for illumination that is applied in the realm of computer graphics rendering in which the transferred energy is that of light [84]. This time-dependency is required in order to facilitate the comparatively slow wave speed inherent to sound propagation.

The transfer of energy in an AR system assumes that all bounding surfaces are Lambertian diffusers such that sound is dispersed uniformly in all directions upon boundary interaction, regardless of angle of incidence. This constraint gives rise to both an advantage and a limitation of the method. The limitation arises due to the inability of this technique to model significant, low-order, specular reflections which may arise in real scenarios. However, from a computation point of view, benefits to the underlying algorithm are attained both in terms of speed and memory as the calculation and storage of individual ray-paths (including angles of incidence and reflection) are not required [9, 63].

Building upon the work by Kuttruff [85], the AR method was extended in later work to include arbitrary enclosure topologies, air absorption and surface absorption yielding the following governing integral equation, as documented in [9, 86],

$$B(r,t) = \int_s \rho(r) B(r^{'}, t - \frac{R}{c}) e^{-mR} \left(\frac{\cos(\theta)\cos(\theta^{'})}{\pi R^2} \right) dS^{'} + B_d(r,t) \tag{3.1}$$

The spatial elements of this expression are shown in Figure 3.6. In the diagram, a sound source position is shown within the confines of an arbitrarily shaped enclosure bounded

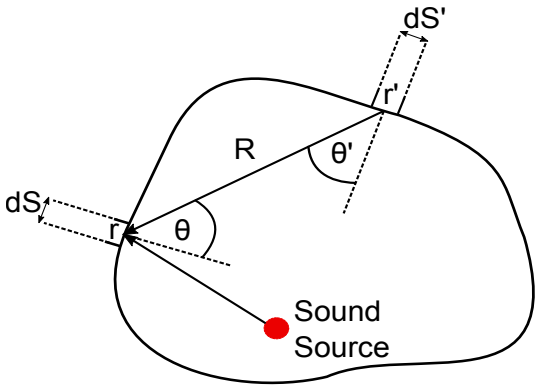

FIGURE 3.6: A 2D cross-section interpretation of the spatial terms of the AR integral equation, after [9]. A radiation path, of length R, is shown between two surface patches dS and $dS^{'}$. The angles θ and $theta^{'}$ provide the angles between the radiation path and the surface normals of patches dS and $dS^{'}$, denoted r and $r^{'}$ respectively

by the surface S. This boundary is broken down into a series of elements of equal length, known as 'patches', of which dS and $dS^{'}$ are examples. The sound energy emitted from the source is given in terms of acoustic radiation density $B_d(r,t)$, defined as the surface integral of sound intensity over the solid angle at a given distance from the source location. The radiation is transferred to all surface patches. In terms of a single patch $B(r,t)$, whose location is given by the vector r, the emitted radiation is calculated from the energy received from the source and all other surface patches (denoted collectively by $B(r^{'},t)$ for all locations $r^{'}$). Additionally, the radiation transferred from all other patches is subject to a reduction in level imposed by the surface reflection coefficient $\rho(r)$. In this way, each patch is defined as a simultaneous emitter and receiver of acoustic energy.

The distance R yields the path length between two patches and, therefore, it is possible to account for the time delay present due to the speed of sound c and energy dissipation due to air arising from the exponential term with loss calibration coefficient m. Finally, angles θ and $\theta^{'}$ are those found between the radiation path R and the surface normals of dS and $dS^{'}$ respectively, providing the cosine terms that impose Lambertian diffusion.

From the governing AR integral equation it is evident that the rendering method is independent of receiver location. An output RIR may be arrived at by calculating the total radiation transferred from all patches (and sound sources) to a prescribed receiver location and recording the value in terms of pressure for individual discrete time-steps. This leads to another advantage over the GA methods previously discussed in which simulations required predefinition of a single listening position. In contrast, once a scene has been rendered using radiosity, the soundfield may be analysed at any point in the modelled space without the requirement of repeating the simulation. AR

is, therefore, particularly suitable for the production of interactive simulations with a moving receiver, as the computational cost of calculating the radiation present at any point in space is deemed minimal and can be carried out in real-time [9, 63].

3.1.5 Geometric Acoustic Models Summary

Having examined and reviewed a selection of GA approaches to virtual acoustic modelling, it is possible to discern suitable applications for each method when seeking to render a RIR by computational means. Firstly, the most apparent limiting factor common to all GAMs is their inability to inherently preserve sound wave characteristics as previously noted. While techniques exist to incorporate diffraction modelling in ray-based simulations, implementing this phenomenon is computationally expensive and achieves only a coarse approximation to the behavior of diffracted sound that occurs in reality. Further elements of low frequency sound propagation are not represented for arbitrary environment topologies. Hence, the use of GAMs in virtual acoustic modelling must be considered on a case-specific basis. If low frequency sound propagation is deemed an important element of investigation, alternative means of providing accurate results must be sought.

Secondly, each GAM discussed here has associated strengths and weaknesses which may be summarised as follows:

- Ray-Tracing: does not guarantee the representation of specular early reflections, but can efficiently simulate diffuse soundfields.
- ISM: directly calculates specular reflections to provide good representation of the early stages of a RIR, however computational cost increases dramatically with modelled reflection order and sound diffusion is not emulated. Furthermore, the underlying algorithm design is non-trivial for complex room geometries.
- Beam-Tracing: optimises the ISM to gain a reduction in computational requirements. As with ISM, non-specular sound reflection is not modelled.
- Acoustic Radiosity: suited to the simulation of diffuse soundfields such as those arising during the later stages of a RIR. The assumption of Lambertian scattering limits this approach to the case of an idealised, stochastic soundfield implying that the early stages of a RIR (strong specular reflections) are not modelled.

Given these broad conclusions, it is reasonable to state that an efficient and accurate approach to mid-high frequency RIR simulation is possible through appropriate combination of a selection of GAMs where each method used is applied to its strength. As

previously stated, such combinations are implemented in industry standard acoustic prediction software. Furthermore, each GAM investigated has been demonstrated capable of achieving interactive simulation rates under certain conditions. This implies that temporally complete and accurate mid-high frequency RIRs may be calculated in real-time. Recent work in this area conducted by Aspöck et al. [61] corroborates this claim.

For the purposes of this work, commercial software was utilised in order to produce high frequency RIRs for the purposes of developing and examining the hybrid modelling paradigm discussed in Chapter 5. Specifically, ODEON and CATT-Acoustic were used to produce results for the cases studies and final study respectively. In brief, ODEON applies the ISM for reflections up to a user-defined order known as the *transition order*. For reflection orders above the transition order, ODEON reverts to a ray-tracing algorithm to produce the later diffuse portion of the output RIR. CATT-Acoustic provides the facility to combine specular cone (beam) tracing, the ISM and ray-tracing to render a temporally complete RIR. Further details on the employment of these programs in this work are given where relevant in Chapter 5. It is assumed, in light of the previous discussion, that the results of these methods may be approximated using real-time GA algorithms.

3.2 Numerical Acoustic Models (NAMs)

Numerical, or 'wave-based', techniques for virtual acoustic modelling share a simple unifying principle: they seek to directly solve the wave equation using a physics-based approach. For the purposes of modelling sound propagation throughout an isotropic medium, the continuous wave equation takes the form of (2.8), (2.30) and (2.50) for 1, 2 and 3 spatial dimensions as derived in section 2.1.3. The number of spatial dimensions included in a given model is determined by the system that is to be represented. For example: a 1D system approximates waves on a string; 2D systems may be applied to the modelling of vibrations in membranes or thin plates; for the purposes of emulating sound propagation in space, 3 spatial dimensions are required.

Numerical acoustic modelling paradigms possess a significant advantage over the previously described GAMs due to the nature of their implementation: wave phenomena (i.e. diffraction, interaction and standing waves/resonances) are inherently preserved in the models. Hence, these methods are deemed to be valid over the entire audible bandwidth in contrast to GAMs which are assumed to be accurate at mid-high frequencies only. However, this advantage comes at the cost of extensive computational

resources required to simulate realistic and spectrally complete soundfields. What follows is an investigation into a variety of NAMs that are currently applied to the problem of acoustic modelling and prediction. This review, while not exhaustive, informs and justifies the selection of Finite Difference Time Domain methods for use in this work over alternative NAM paradigms.

3.2.1 Finite Difference Time Domain (FDTD)

The mathematical descriptions of dynamic system behaviour, such as the nature of sound wave propagation, commonly rely on the use of partial differential expressions (PDEs). Finite Difference Time Domain (FDTD) approximations provide a means of calculating numerical solutions to PDEs, rendering possible the discrete-time simulation and analysis of physical systems through use of iterative algorithms. This approach was utilised by Yee in 1966 [87] to solve Maxwell's equations for electrodynamics and later applied to the realm of acoustic modelling by Savioja [88] and Botteldooren [89] in the 1990s. As an introduction to FDTD algorithms, also referred to as 'schemes', the numerical solution of the 1D wave equation, which provides a rudimentary string model, is detailed here.

The primary components of FDTD schemes are the *forward-*, *backward-* and *centered-difference* approximations to the derivatives of an arbitrary function $f(x)$. These are obtained from truncated forms of the Taylor's series expansion of $f(x+\triangle x)$ and $f(x-\triangle x)$ given respectively as:

$$f(x+\triangle x) = f(x) + \triangle x f'(x) + \frac{\triangle x^2}{2!} f''(x) + \frac{\triangle x^3}{3!} f'''(x) + ... \tag{3.2}$$

$$f(x-\triangle x) = f(x) - \triangle x f'(x) + \frac{\triangle x^2}{2!} f''(x) - \frac{\triangle x^3}{3!} f'''(x) + ... \tag{3.3}$$

Truncation and rearrangement of (3.2) yields the first-order forward-difference approximation to the first derivative of $f(x)$ as:

$$f'(x) = \frac{f(x+\triangle x) - f(x)}{\triangle x} + O(\triangle x^2) \tag{3.4}$$

where $O(\triangle x^2)$ refers to terms of the order $\triangle x^2$ or greater and are, in this instance, assumed negligible. Similarly, the backward-difference approximation to the first-order derivative may be derived from (3.3):

$$f'(x) = \frac{f(x) - f(x-\triangle x)}{\triangle x} + O(\triangle x^2) \tag{3.5}$$

Additionally, the second-order derivative of $f(x)$ may be approximated through summing (3.2) and (3.3) to arrive at the following relation:

$$f''(x) = \frac{f(x + \triangle x) - 2f(x) + f(x - \triangle x)}{\triangle x^2} + O(\triangle x^3) \tag{3.6}$$

giving the centered-difference approximation to $f''(x)$ where, in this case, third-order terms $O(\triangle x^3)$, and greater, are assumed to be negligible. The second-order finite difference approximation provides the basis of the numerical solutions to the wave equation discussed in this work. In the following example, the means by which the 1D wave equation may be solved by FDTD modelling is demonstrated.

1D FDTD String Example:

The derivation of a simple 1D FDTD scheme begins with spatio-temporal sampling of the continuous space, x, and time, t, variables:

$$t = nT \tag{3.7}$$

$$x = lh \tag{3.8}$$

where T is the temporal sampling period (s), defined as the reciprocal of the temporal sampling frequency F_S, and h is the spatial sampling instance (m). The integers n and l describe the position of the system in discrete-time and space respectively. Using these discretised variables in combination with the second-order centered-difference approximation, the 1D wave equation may now be re-expressed as follows:

$$\frac{\partial^2 p}{\partial t^2} = c^2 \frac{\partial^2 p}{\partial x^2} \Rightarrow \frac{p_l^{n+1} - 2p_l^n + p_l^{n-1}}{T^2} = c^2 \frac{p_{l+1}^n - 2p_l^n + p_{l-1}^n}{h^2} \tag{3.9}$$

where p_l^n and p_l^{n-1} are regarded as the displacement of the string at spatial location lh during the current and previous time steps respectively. Hence, the single unknown term of (3.9) refers to the displacement at the next time step, p_l^{n+1}. This quantity may be solved for by re-arranging the discrete approximation of (3.9) yielding the final iterative FDTD algorithm:

$$p^{n+1} = \frac{c^2 T^2}{h^2}(p_{l+1}^n - 2p_l^n + p_{l-1}^n) + 2p_l^n - p_l^{n-1} \tag{3.10}$$

The above algorithm provides an approximate numerical solution to the 1D wave equation through recursive computation of displacement values along the length of the discrete domain as time increases. As previously shown, the FDTD approximations applied are second-order accurate and, therefore, small numerical errors between computed and theoretical values are expected, yet deemed insignificant.

The extension of the FDTD method to the numerical solution of the 2D and 3D wave equations is realised by similar means as those provided in the above example. In brief, finite difference approximations are applied to both the second-order time derivative of pressure and the 2D or 3D spatial Laplacian operator. As such, a 2D FDTD wave equation simulation may be envisaged as a grid of *pressure nodes* which facilitate the propagation of scalar pressure values throughout 2D space in discrete time steps. The extension of this notion to 3D results in the conception of a lattice of pressure nodes occupying a given volume (i.e. virtual room model).

While this numerical acoustic modelling approach is capable of inherently preserving wave phenomena, which are particularly apparent at low frequencies, the required discretisation of the simulated spatial domain leads to implementation issues. These issues include: dispersion error which arises due to the anisotropic wave propagation simulated over a discretised domain; numerical stability issues which limit the resolution of the discrete spatial domain relative to a given temporal sampling rate; the requirement of extensive computational resources. These aspects of FDTD simulations are addressed and examined in depth in Chapters 4 and 5.

3.2.2 Digital Waveguide Mesh (DWM)

Digital waveguides gained considerable attention in the 1980s and 1990s as an approach to physical modelling of musical instruments [90–92]. Techniques used previously, for example additive/spectral synthesis and frequency modulation, provide a means of approximating complex timbres. However, these synthesis techniques are abstracted from the physical mechanisms present within the modelled instrument and rely on experimentation and subjective evaluation to determine the correctness of resulting sounds. Through physical modelling it became possible to describe and simulate virtual instruments to a high level of detail leading to improved accuracy of audible results.

Digital waveguides seek a computable solution to the wave equation via a discrete form of the travelling wave solution discovered by d'Alembert in 1747,

$$p(nT, lh) = p^{+}(n - l) + p^{-}(n + l) \tag{3.11}$$

for a 1D system where all variables are as defined in the previous section. In this solution, p^{+} exists as a travelling wave of displacement in the positive x-direction and p^{-} exists likewise but travels in the opposite direction. The behaviour of (3.11) may be represented by two digital delay lines, which together form a simple bi-directional digital waveguide as demonstrated in Figure 3.7.

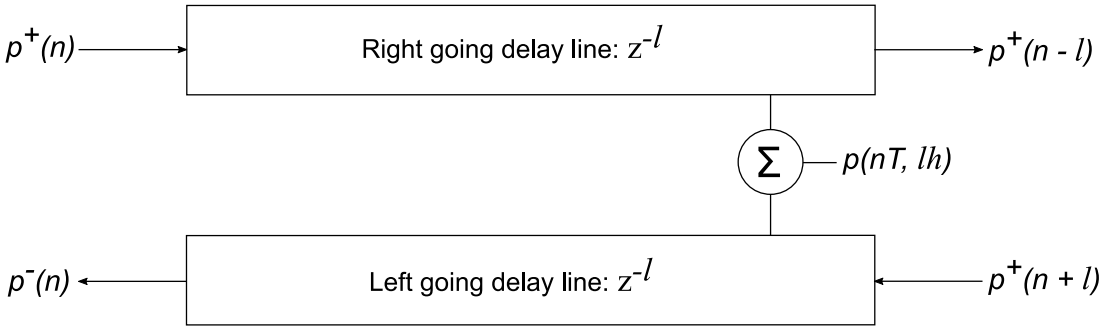

FIGURE 3.7: Diagrammatic depiction of a basic digital waveguide which provides numerical solution of the 1D wave equation via a computable travelling wave approximation, after [10]. The system is comprised of two delay lines of length l (signified by the z-transform quantity z^{-l}) that propagate right- and left-going displacement travelling waves $p^+(n)$ and $p^-(n)$ respectively. A time-varying output is gained by summing the values present on each delay line at the desired location in space lh where h is the spatial sampling instance.

Multidimensional wave propagation is simulated through combination of DWG elements to construct a Digital Waveguide Mesh (DWM) [93]. The DWM structure is created by interconnecting unit delay waveguide elements via an arrangement of *scattering junctions*. As with FDTD models, the arrangement of scattering junctions is defined to spatially discretise the target acoustic field at regularly spaced sampling instances. The role of these junctions is to apply appropriate equations to incoming wave components such that outgoing scattered wave components may be calculated in accordance with the wave equation thus facilitating wave propagation. In this way, scattering junctions may be seen as providing the same function as grid nodes used for FDTD simulations. However, the equations governing the interconnecting element are slightly more complex in DWMs. A generic arrangement of scattering junctions is depicted in Figure 3.8. As shown, the interconnection of bi-directional unit delay lines is such that incoming and outgoing signals are directed along separate delay lines. There are two conditions that enable solution of the wave equation as described in [93]. Firstly, the pressure value that occurs at a given junction is equal to the sum of all incoming and outgoing pressure values directed along interconnecting delay lines. Secondly, the net velocity arising at a given junction is zero. This second condition is met by ensuring the sum of velocities arising on incoming delay lines is equal to the sum of velocities on outgoing delay lines. Combination of these two conditions gives rise to the general equation for pressure at a scattering junction P_J with associated impedance Z_J, after [11, 94]:

$$P_J = \frac{2\sum_{i=1}^{N} \frac{p^+_{J,i}}{Z_J}}{\sum_{i=1}^{N} \frac{1}{Z_J}} \tag{3.12}$$

where $p^+_{J,i}$ is the pressure incoming from the i^{th} of N neighbouring junctions. The value P_J may then be applied to calculate the outgoing pressure values scattered to

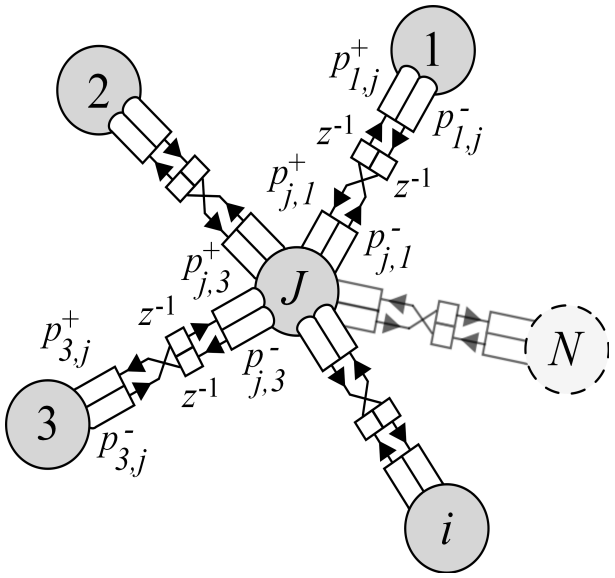

FIGURE 3.8: A generic scattering junction, J with associated pressure P_J and impedance Z_J values and N neighbouring scattering junctions interconnected by bidirectional unit delay (z^{-1}) lines. Examples of incoming ($p^+_{J,i}$) and outgoing ($p^-_{i,J}$) pressure values arising across the interconnecting delay lines are shown. Image sourced directly from [11].

neighbouring junctions, $p^-_{J,i}$ using:

$$p^-_{J,i} = P_J - p^+_{J,i} \tag{3.13}$$

Temporal recursive updating of the pressure values passed from junction to connected junction is facilitated by the unit delay line through which outgoing pressure values from one junction become the incoming pressure values to a neighbouring junction, as per:

$$p^+_{i,J} = \mathrm{z}^{-1} p^-_{J,i} \tag{3.14}$$

Hence, the iterative DWM based solution of the wave equation for a given multidimensional acoustic field is carried out through recursive calculation of (3.12) - (3.14).

DWM modelling has been applied to a range of virtual acoustic modelling scenarios such as drum membranes [95] and room acoustics [94, 96]. The DWM paradigm is considered to be a subset of a broader family of FDTD schemes [34] which, under certain conditions, produces effectively equivalent results to the FDTD counterpart. Hence, it is assumed that the DWMs have the same advantages and disadvantages as those noted for FDTD schemes in the previous section. However, DWMs possess a particular advantage in that impedance values associated with scattering junctions are readily accessible via the governing equations and, therefore, may be altered dynamically to emulate changes in the geometry of the target acoustic field. The ability to facilitate

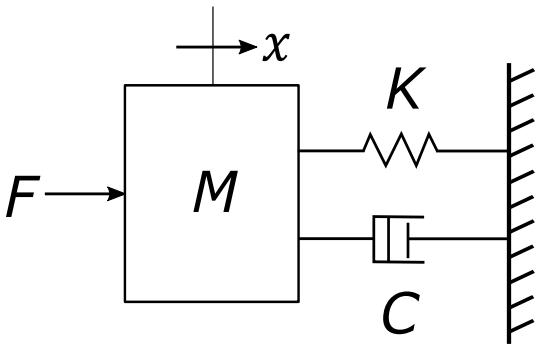

FIGURE 3.9: Graphical depiction of a mechanical Mass (M), Spring and Damper system with spring constant K, damping coefficient C and one degree of freedom, x, after [12]. This system forms a single element which, when connected to other elements of the same form, can be used to represent pressure fluctuations due to sound wave propagation.

transient changes in model topologies has placed DWMs at the center of dynamic vocal tract modelling (see e.g. [97]).

3.2.3 Finite Element Method (FEM)

Finite element analysis (FEA) is a well established numerical method that is commonly applied to the investigation of a range of mechanical systems [98, 99]. In the field of virtual acoustic modelling, the principles of FEA have been used to develop the Finite Element Method (FEM) that facilitates the numerical solution of the Helmholtz equation, also referred to as the *reduced wave equation* [100]. A straight-forward derivation of the Helmholtz equation is obtained by applying the Fourier Transform to the time-dependent wave equation. This transformation may be carried out by inserting a single frequency ansatz of the form:

$$p(x, y, z, t) = P(x, y, z)e^{-j\omega t} \tag{3.15}$$

into the wave equation (assuming 3 spatial dimensions), yielding:

$$\frac{\partial^2}{\partial t^2}[P(x, y, z, t)e^{j\omega t}] = c^2\nabla^2[P(x, y, z, t)e^{-j\omega t}] \tag{3.16}$$

which reduces to the Helmholtz equation,

$$\nabla^2 P + k^2 P = 0; \qquad k = \frac{\omega}{c} \tag{3.17}$$

where k is the wavenumber for a wave with angular frequency ω.

In [12], Wright exposes the underlying principles of FEM through use of vibrating mechanical elements. The diagram in Figure 3.9 portrays such a system consisting of

a mass M connected to a spring of stiffness K and a damping element C. If the mass is acted upon by a force F, causing displacement in the single spatial dimension x, the governing equation of motion may be stated as,

$$M\frac{\partial^2 x}{\partial t^2} + C\frac{\partial x}{\partial t} + Kx = F \tag{3.18}$$

which corresponds to the lossy wave equation given in section 2.1.4. Removing the external force and damping condition from the system yields the Helmholtz equation in terms of the physical constants for the mass and spring,

$$M\frac{\partial^2 |x| e^{-jwt}}{\partial t^2} + K|x|e^{-jwt} = 0 \tag{3.19}$$

for an arbitrary frequency ansatz $x = |x|e^{-j\omega t}$. This equation further simplifies to,

$$(K - w^2 M)x = 0 \tag{3.20}$$

from which the natural frequencies (or 'modes') of vibration can be calculated. The application of FEM to the simulation of sound propagation involves representing an enclosed volume of air with a discretized, multidimensional arrangement of mechanical elements. This process can be thought of as a concatenation of mass-spring systems in 1, 2 or 3 spatial dimensions depending on the nature of the system to be modelled. In this way, it is possible to model acoustic pressure in a soundfield as the displacement of a series of interconnect masses from their respective equilibrium positions. Solution of the Helmholtz equation for the given discretized space may then be achieved by defining a frequency-domain driving function $\mathbf{F}(\omega)$ and calculating,

$$\mathbf{x} = (\mathbf{K} - \mathbf{M}\omega^2) \setminus \mathbf{F}(\omega) \tag{3.21}$$

where '\' signifies solution through the pseudo-inverse. In the expression above, $\mathbf{x}$ is an array containing the multidimensional pressure distribution arising from the displacement of masses $\mathbf{M}$, interconnected by springs $\mathbf{K}$, in response to the excitation for a particular frequency ω. Hence, to obtain a spectrally complete response in the frequency domain, this expression must be repeated for each required frequency (i.e. discrete frequency values ranging across the audible bandwidth). In order to derive a RIR for a specific location in the pressure field $\mathbf{x}$, the frequency-domain pressure signal for a single element (or 'node') of array $\mathbf{M}$ is transformed into the time-domain via the inverse Fourier transform.

This method of solution highlights a notable disadvantage of the FEM approach in that the pressure field calculation for each frequency value requires a matrix inversion

which becomes increasingly inefficient for large discretized spatial domains [101]. Furthermore, pressure values for the complete spectrum must be gained prior to applying the inverse Fourier transform to obtain a RIR. This implies that the entire system must be solved before time-domain information is accessible for use.

As with most examples of wave-based modelling methods, the use of FEM is usually constrained to low frequency sound simulation (see e.g. [102–104]). This is mainly due to the computational challenges arising from the relationship between the spatial dimensions of elements and the upper simulation frequency. As stated in [104], approximately 6-10 elements are needed per shortest wavelength modelled, leading to vast numbers of nodes required to represent large pressure fields at high frequencies. It is for this reason, together with the fact that the approach is applied in the frequency-domain, that FEM is deemed unsuitable for rendering RIRs at interactive simulation rates [101].

3.2.4 Boundary Element Method (BEM)

As highlighted in [105] and previous discussion, FEM requires the medium to be discretised and modelled leading to high element counts for high frequency acoustic simulations and large room geometries. The same is also true for discrete time simulation paradigms such as DWMs and FDTD schemes. In contrast, the Boundary Element Method (BEM) seeks a solution to the wave equation through simulation of sound behaviour at domain boundaries only. Therefore, the dimensionality of BEM simulations is always one less than those which require discretisation of the bounded medium. This in turn leads to a significant reduction in the number of elements needed to sufficiently model sound waves in space, especially for modelling scenarios involving a large medium volume-to-surface-area ratio.

The formulation of a BEM scheme appropriate for acoustic simulation begins with the Kirchoff-Helmholtz equation (as documented in [24, 105]). This equation constitutes the boundary value integral expression for the wave equation. To reiterate briefly, this equation states that the scalar acoustic field existing in a medium can be calculated via the projection and summing of free-field Green's functions from discrete boundary surface nodes. As described in [106], there are two approaches to such a solution. The boundary integral can be evaluated for a given spatial topology by means of a linear system solve applied iteratively over a range of discrete frequency values in a manner similar to that taken in the FEM. The resulting frequency domain representation of the soundfield may then be transformed into the time domain to recover the impulse response of the system. Alternatively, the governing Kirchoff-Helmholtz equation may

be transformed into the time domain prior to applying a discrete time domain approximation. The derivation of the latter solution is addressed in [106] and [105].

There are two clear advantages of BEM over alternative approaches. Firstly, the aforementioned inherent reduction in problem dimensionality leads to reduced memory and computation time relative to the FEM. Secondly, the BEM preserves Sommerfield's radiation condition given as:

$$\lim_{\mathbf{r}\to\infty}\left(\frac{\partial P(\mathbf{r},\omega)}{\partial \mathbf{r}} - ikP(\mathbf{r},\omega)\right) = 0 \tag{3.22}$$

which is required in order that free-field propagation is preserved over infinite space with no reflected wave components. The preservation of this condition is possible due to the avoidance of spatial discretisation of the domain medium. In alternative numerical approaches, dispersion error due to anisotropic wave propagation leads to spurious reflections from theoretically fully absorbing boundaries. As such, the BEM is particularly applicable to exterior acoustic simulations such as those conducted to model the impact of human physiology on the soundfield [107, 108] and investigation of surface scattering properties [105, 109]. However, examples of the application of BEM to RIR generation are relatively sparse in the literature as the development of this approach has predominantly focused on models for sound scattering rather than enclosed wave propagation.

3.2.5 Adaptive Rectangular Decomposition and the Discrete Cosine Transform

The basis of this approach stems from the Helmholtz equation. As discussed in [110], analytical solutions to the Helmholtz equation exist under certain conditions giving rise to the so-called 'Eigenmode model' for cuboid spatial topologies. Following [110], (3.17) may be solved by separation of variables yielding the solution:

$$p(x,y,z,\omega) = p_x(x,\omega)p_y(y,\omega)p_z(z,\omega) \tag{3.23}$$

for 3 dimensions, where each term on the right hand side is an equation that satisfies the ordinary differential equation:

$$\frac{\partial^2 p_i}{\partial i^2} + k^2 p_i = 0; \qquad i = x,y,z \tag{3.24}$$

As such, the terms p_i describe wave propagation along each axis in a Cartesian coordinate system with wave number contributions k_x, k_y and k_z given,

$$k^2 = k_x^2 + k_y^2 + k_z^2 \tag{3.25}$$

For a cuboid space with dimensions L_x, L_y and L_z, where each surface is subject to Neumann boundary conditions (2.53), solutions of the following form result:

$$p_i = \cos(k_i i); \qquad i = x, y, z \tag{3.26}$$

in which the wavenumber contributions are constrained by,

$$k_i^2 = \frac{N_i^2 \pi^2}{L_i^2}; \quad i = x, y, z; \quad N_i \in \tag{3.27}$$

Given this condition, the eigenfunctions of the wave equation in a rectangular domain may be defined as,

$$p_i(x, y, z) = \cos\left(\frac{i_x \pi x}{L_x}\right) \cos\left(\frac{i_y \pi y}{L_y}\right) \cos\left(\frac{i_z \pi z}{L_z}\right) \tag{3.28}$$

after [110], which yields the frequencies of resonance (or Room Modes, see section 2.3.6) of the system.

In [111–113], the authors propose a means of applying the above result to wave-based room acoustics modelling in enclosures of non-rectangular topologies. It is noted in [111] that a time-dependent expression for sound pressure at a given point in a rectangular space is obtained through a weighted sum of calculated eigenfunctions:

$$p(x, y, z, t) = \sum_{i=i_x, i_y, i_z} m_i(t) p_i(x, y, z) \tag{3.29}$$

where $p_i(x, y, z)$ is as previously defined and $m_i(t)$ are temporal weighting coefficients. These coefficients are calculated through application of a discrete cosine transform (DCT). Following [111, 114], values for $m_i(t)$ are solutions to the forced simple harmonic oscillator equation. Hence, they take the form:

$$m_i(t) = e^{j\omega_i t} + \frac{F_i}{\omega_i^2} \tag{3.30}$$

as a solution to:

$$\frac{\partial^2 m_i(t)}{\partial t^2} + \omega_i^2 m_i(t) - F_i = 0 \tag{3.31}$$

where F_i is the forcing term, after [114]. The calculation of the weighting coefficients

may be carried out as a discrete-time iterative process. Expressing $m_i(t)$ in discrete-time samples gives $m_i(nT)$, where T is the sampling period and n is an integer. The terms for the next and previous sample values for m_i^n are denoted m_i^{n+1} and m_i^{n-1} respectively where,

$$m_i^{n+1} = e^{j\omega_i nT} e^{j\omega_i T} + \frac{F_i}{\omega_i^2} \tag{3.32}$$

$$m_i^{n-1} = e^{j\omega_i nT} e^{-j\omega_i T} + \frac{F_i}{\omega_i^2} \tag{3.33}$$

Summation of the above expressions yields,

$$m_i^{n+1} = \frac{2F_i}{\omega_i^2} + 2\cos(\omega nT) e^{j\omega_i nT} - m_i^{n-1} \tag{3.34}$$

Inserting the discrete-time form of (3.30) into (3.34) yields the update equation that results in the weighting coefficients to be applied in (3.29) at the next time sample [111],

$$m_i^{n+1} = 2\cos(\omega_i nT) m_i^n - m_i^{n-1} + \frac{2F_i}{\omega_i^2}(1 - \cos(\omega_i nT)) \tag{3.35}$$

Hence, the pressure field $p(x, y, z, t)$ is recursively calculated by means of the DCT. This approach has a significant advantage in that the solution to the wave equation is discretised in time only. The simulated spatial pressure field is continuous and, therefore, is not subject to dispersion effects that are inherent to alternative methods such as digital waveguide meshes and finite difference schemes (see sections 3.2.6 and 3.2.7 respectively). However, this advantage only holds for rectangular domains. Furthermore, the DCT has been demonstrated as an extremely efficient means of simulating a sound pressure field as the fast Fourier transform may be employed to calculate the required Cosine transform. As noted in [113], parallel computation of the required transform using GPUs results in low simulation times and memory requirements in comparison with alternative finite difference schemes.

In order to simulate arbitrary room topologies by means of the DCT, the domain must be segmented into a series of rectangular/cuboid sub-domains. A simple example is shown in Figure 3.10. This segmentation is facilitated through 'Adaptive Rectangular Decomposition' (ARD) as described in [111]. In brief, this decomposition is carried out pre-computation and seeks to fit the largest rectangular sub-domain possible in the room geometry. The limit of the rectangular domain volume is reached when one or more boundaries of the sub-domain reaches a boundary of the room topology. This process continues for the remaining volume in the space until the entire room geometry is represented by rectangular domains of the largest possible size.

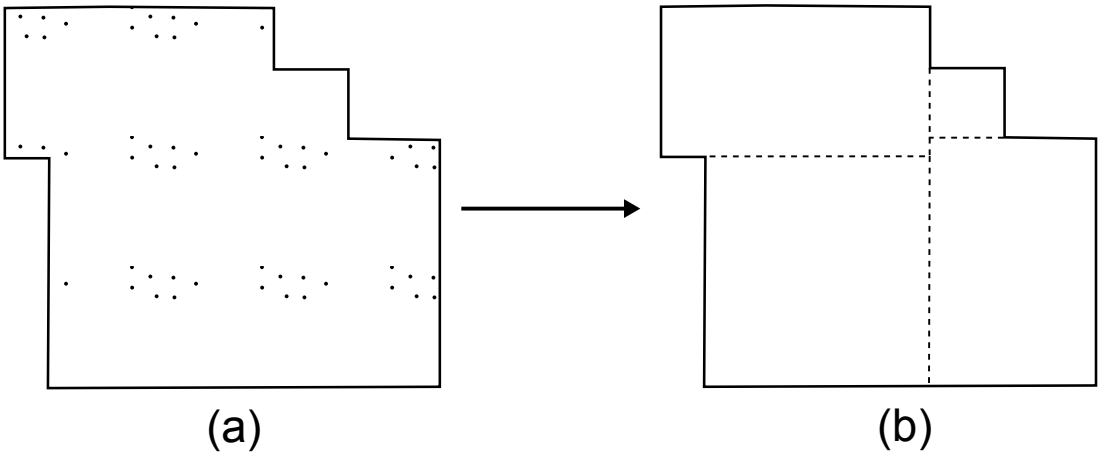

FIGURE 3.10: An example of rectangular decomposition as applied to a simple 2D domain, given in (a). The way in which the interior domain may be segmented into a series of rectangular sub-domains as per the ARD approach is given in (b).

Once the decomposition is complete, the DCT update may be applied to calculate the pressure field in each sub-domain. In order to link each sub-domain, a finite difference scheme is employed at the interface between adjacent sub-domains such that sound pressure waves are transferred from one sub-domain to another. This finite difference scheme subjects pressure waves to the Neumann boundary condition on each interface allowing phase-preserving continuation of a wavefront from one domain to another. The pressure fluctuations represented by the finite difference scheme are applied as driving functions for the adjacent sub-domain in the governing update equations (3.29) and (3.31). A similar approach is used to interface sub-domains with absorbing boundary conditions at the surfaces of the topology. It is noted in [111, 113] that this approach to interface handling gives rise to spurious wavefronts that are reflected back into the sub-domain from which they were travelling. The intensity of the erroneous reflections varies with frequency and is documented to be as high as -28 dBFs in the 1 kHz range [111]. The authors in [111, 113] state that these errors are imperceptible in resulting auralisations, however evidence to support this claim (e.g. via subjective testing) is not provided. Similar errors are also produced as a result of absorbing boundary interfacing [111].

There are a few shortcomings in the literature on the application of the DCT/ARD to room acoustic modelling:

- Spectral analysis of resulting RIRs is not provided to demonstrate correct/approximate solution to the wave equation for the spaces modelled.
- The issue of numerical stability is not addressed. Therefore, the conditions under which this simulation method operates correctly are unknown.

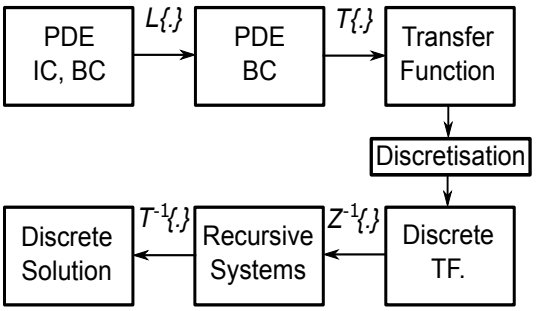

FIGURE 3.11: Diagrammatic representation of the FTM as applied to PDEs, after [13].

- The dispersive effects of the interfacing approach are not documented. While it may be assumed that there are is no dispersion error in each sub-domain, the global solution of the wave equation achieved using interfaced sub-domains will, to some extent, exhibit dispersion due to the discretisation of the pressure field on each interface.

Hence, while this method demonstrates clear advantages in terms of run-time and memory reduction comparative to alternative algorithms, the accuracy of results requires further investigation.

3.2.6 Functional Transform Method (FTM)

The Functional Transform Method (FTM), as presented in [115, 116], provides a general approach to the frequency-domain solution of partial differential equations (PDEs). In contrast to FDTD/DWM/FEM techniques, the direct discretisation of differential operators is avoided by instead producing a transfer function description of the system under examination. In relation to room acoustics modelling, the FTM shares similarities with the DCT/ARD method (described previously) in that it seeks to find the eigenfunctions and, therefore, the modes of geometrically simplistic resonant systems. However, the approach to this outcome, as shown in Figure 3.11 differs notably.

The initial step is to define a PDE describing the acoustic system to be modelled (i.e. a tube, membrane or enclosure). Associated initial conditions (ICs), such as sound source excitations, and boundary conditions (BCs) must also be defined. Following [117], the Laplace transform, $L\{.\}$ is applied to remove time derivatives terms yielding a differential equation with spatial derivatives only. The result then undergoes a further transformation to remove the remaining spatial derivatives and arrive at a purely algebraic transfer function expression for the system. This second transform, $T\{.\}$, is a problem specific Sturm-Liouville transformation which requires calculation

of a transformation kernal that is adapted to the system under investigation [117–119]. The transfer function expression for the system may be discretised via the Impulse Invariant Transformation which provides a discrete-time representation of the complex variable introduced by the initial Laplace transform. Finally, inverse z- and Strum-Liouville transformations are applied to produce a discrete-time solution of the system in the form of weighted first order resonating systems - a result that is analogous to the related DCT/ARD approach.

The FTM has been demonstrated as applicable to the physical modelling of membranes [117, 119, 120] and wave propagation throughout enclosures [121]. The latter modelling scenario highlights a key limitation of this method that arises due to the reliance on the definition of a PDE describing the overall system. For complex enclosed geometries, incorporating internal obstacles and/or apertures, the process of calculating resonant frequencies (eigenfunctions) becomes complicated due to the difficulty in defining an appropriate governing PDE. To alleviate this problem, it is possible to segment the spatial domain into an arrangement of rectangular or triangular elements, for which governing PDEs are known [13]. This process, which is similar to ARD, forms the basis of 'block-based' modelling [117, 122]. The segregation of enclosure geometry leads to the requirement of interfacing adjacent blocks via wave digital filters or DWMs [120, 122]. As with the ARD approach, the chosen methods of interconnecting adjacent sub-domains have been demonstrated to produce undesirable spurious reflections within the modelled domain [120].

The FTM has inherent advantages over alternative numerical acoustic modelling paradigms including DWMs, FDTD and FEM/BEM approaches:

- Results gained from model simulations do not suffer from the effects of numerical dispersion error due to the avoidance of direct discretisation of the spatial domain [13]. However, this is only true for models that do not rely on interconnected sub-domains.
- Source sources and receivers may be placed at liberty within virtual models and are not constrained to locations dictated by a lattice of spatial sampling points [13].
- The solution of the wave equation for simple topologies may be realised through application of the fast Fourier transform yielding efficient computation [121]. However, such an approach to implementation has only been demonstrated for a simple cuboid room topology [121].

In summary, the FTM shares similar advantages to those exhibited by the ARD/DCT approach in terms of dispersion reduction and efficiency of implementation. However, this modelling paradigm has to date only been applied to soundfield simulation in simplistic 3D environments and validation of resulting RIRs has not been carried out. Hence, it is not possible to conclude from associated literature that the FTM is appropriate for, or indeed capable of, rendering accurate representations of room acoustics characteristics.

3.2.7 Numerical Acoustic Summary

In review of the discussions put forward on numerical modelling paradigms, it is evident that each approach is highly applicable to accurate simulation of propagating sound waves. All techniques highlighted here facilitate iterative solution of the wave equation for a target acoustic field, thus preserving wave-like characteristics that are omitted in purely ray-based models. Unfortunately, even with optimised use of parallel computation architectures, these numerical solutions present computational issues in terms of lengthy run-times and large memory demands. It is noted that in the case of ARD techniques and FTM, this concern is somewhat alleviated due to the alternative, mostly spatially non-discrete, formulation of numerical solution.

Particular findings drawn from the literature are of significance in relation to this work. With regards to the methods that rely on segregation of the target acoustic field, such as the FTM and ARD technique, it may be concluded that the method of interfacing adjacent sub-domains and interior domains with surface boundaries is not optimal. Undesirable wave scattering effects that occur at these interfaces, which are required for all but the most basic cuboid model, reduce the accuracy of simulated wave properties. Therefore, for geometrically complex sonic environments, the validity of such approaches is somewhat compromised. This claim is supported by considering requirements for predictable surface absorption characteristics which cannot be ensured given the nature of erroneous frequency-dependent wave scattering effects.

Alternative frequency domain approximations to the wave equation, facilitated by the FEM and BEM, approach RIR generation through a solution for discrete frequency values. This imposes the requirement for a linear system solution for each simulated frequency in turn, prior to transforming the solution into the time domain to render the RIR. The computation times associated with such an approach for large scale acoustic environments are deemed too extensive for interactive modelling scenarios. Although time-domain formulations of BEM have been developed, the application of

this method has, to date, been limited to basic room topologies [106] and particular surface scattering problems [105, 109].

In recent years, considerably more research has been conducted on the application of discrete time domain DWM and FDTD techniques to RIR simulation than their frequency domain counterparts. In particular, the work carried out by Kowalczyk and van Walstijn [110, 123, 124] and the NESS project [125] on FDTD simulation has provided extensive research into this method culminating in an established room acoustic modelling technique. While both DWM and FDTD solutions inherently give rise to numerical dispersion error, the properties of this phenomenon are predictable and well understood. Hence, the conditions under which simulation results are deemed to be valid and accurate are known. Furthermore, issues surrounding the nature of surface boundary conditions, preservation of numerical stability and optimised parallel scheme computation have been addressed such that emulation of realistic enclosed sound wave propagation may be facilitated with low simulation run-times. Finally, it is noted that the DWM is a particular case of FDTD scheme [34, 126] and, therefore, these approaches achieve effectively equivalent results. In light of these findings, the selection of the FDTD method is justified for the purposes of this work. Specific considerations and attributes of this method, as applied to room acoustic modelling, are discussed further in Chapter 4.

3.3 Geometric/Numerical Hybrid Acoustic Modelling

Hybrid acoustic modelling (HAM) techniques are generally based on the combination of two or more modelling methods, such that each contributing method is used to its advantage. Most often, the overall aim is to create a means of virtually representing sound propagation which seeks the optimal balance between computational efficiency and accuracy of simulation results. As detailed throughout this chapter, all virtual acoustic modelling possess inherent strengths and limitations. Therefore, the combination of different methods facilitates the opportunity to devise an approach that applies each method to its strengths to tackle specific elements of RIR generation. For the purposes of this study, discussion of GAM/NAM hybrid examples is provided here.

An initial example of a GAM/NAM HAM solution is the RenderAIR system [14, 96]. The underlying motivation for the design of this modelling approach stems from the inaccuracy of GAMs at low simulation frequencies and prohibitive computation costs associated with NA solutions. The DWM paradigm is utilised to simulate wave propagation at low frequencies such that an accurate representation of prominent wave features may be achieved. By limiting the simulation bandwidth of the DWM to low

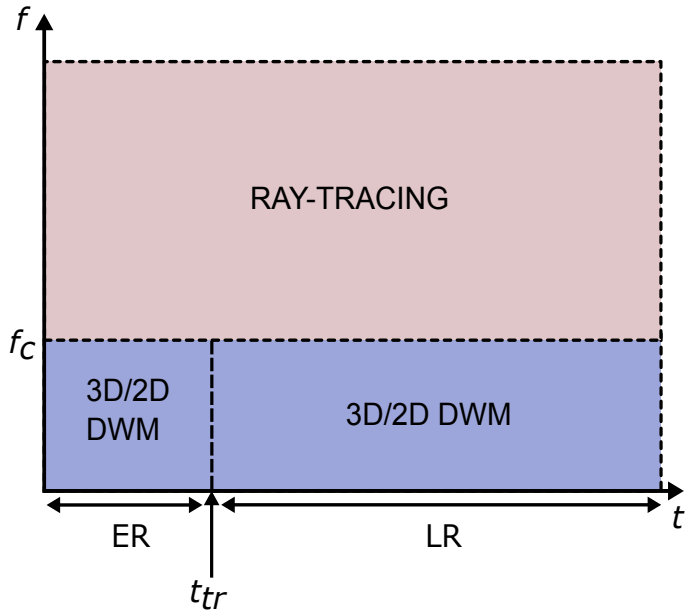

FIGURE 3.12: The time/frequency division between application of modelling methods in the RenderAIR system [14]. The frequency bandwidth f is plotted on the vertical axis with f_c denoting the frequency of transition between the use of DWM and ray-tracing paradigms during simulation. The simulation time range t is given on the horizontal axis with t_{tr} denoting the optional temporal division between early reflection (ER) and late reflection (LR) RIR stages.

frequencies only, a larger spatial sampling instance compared to that required for full audio bandwidth simulations may be applied in order to restrict the total number of pressure nodes to a manageable amount yielding reasonable simulation times. High frequency RIR content is then rendered by a comparatively efficient ray-based acoustic model. This hybrid arrangement of modelling approaches is shown in Figure 3.12. A further point of interest noted in [96] is the optional application of a single 2D DWM to the synthesis of the low frequency soundfield. The usage of the 2D DWM, as documented in [14, 96], is variable and can be applied to either the low frequency diffuse field (consisting of late reflections only) or the entire low frequency soundfield. It is noted in [96] that simulation of early reflections is more accurate using 3D DWMs compared with their 2D counterparts, however this accuracy comes at higher computational cost. Results presented in [14] demonstrate a significant reduction in run-time and memory requirements when comparing the cases in which the 2D DWM is applied. At the same time, low frequency analysis provided for the same study reveals the improvement of modelled low frequency RIR features comparative to purely ray-based simulation results. The notion of significantly reduced computational cost and improved low frequency accuracy over geometric acoustic simulations gained through use of 2D numerical modelling is central to the investigations of this work (see Chapter 5).

A more recent study, conducted by Southern et al. [80, 81], combines the FDTD NAM approach with beam tracing and radiosity methods to realise a hybrid RIR synthesis system. As with the RenderAIR project, the central motivation is to produce an accurate digital representation of a given soundfield while maintaining realistic computational requirements. Again, this results in a transition between numerical and geometric RIR simulations as frequency increases. A notable advantage of the HAM described in [81] is the use of beam tracing and radiosity RIR synthesis which seeks to provide an optimal solution to high frequency acoustic modelling. Through this combination, efficient early reflection modelling may be handled by a beam tracing algorithm while the radiosity approximation can be introduced at some transition time to calculate the diffuse field. Hence, the prohibitive computation associated with beam tracing for high reflection orders, and therefore late stages of a RIR, is avoided. This, in theory, leads to a more faithful emulation of early RIR components than that achieved by means of a ray-tracing algorithm in which prominent low order specular reflection paths may be omitted (see section 3.1.1). A significant contribution presented in [81] is the derivation of an energy matching procedure that enables appropriate combination of RIRs rendered by geometric and SRL FDTD models. This calibration process is utilised throughout the work presented in Chapter 5, facilitating the creation of hybrid RIRs as required for this study.

3.4 Summary

In review of the discussion put forward in this chapter, it is evident that approaches to acoustic modelling are many. This survey addresses the implementation and design of established geometric and numerical acoustic simulation techniques. In turn, this facilitates the evaluation and review of inherent advantages and limitations common to each approach. This review culminates in the finding that the definition of an optimal solution to RIR synthesis is most likely realised by means of HAM systems. The primary justification for this claim stems from the relative advantages/disadvantages of GAM and NAM paradigms in upper and lower regions of simulation bandwidths which are detailed in depth. For the purposes of this work, a combination of the FDTD and ray-based methods is highlighted as the most appropriate means of hybrid RIR generation.

Recent research and investigation into the use of parallel computation for NA simulation has demonstrated reductions in required run-times, particularly for the case of FDTD schemes. However, there remains a significant trade-off between model scale/simulation bandwidth and computational cost for 3D modelling scenarios, especially when

attempting to approach real-time sound synthesis. In order to alleviate this trade-off, while retaining agreeable levels of accuracy in synthesised RIRs at low frequencies, a new HAM system is devised in this study. This novel approach is influenced by both the RenderAIR system and the 2D DWM room modelling work of Kelloniemi [127]. It is proposed that the use of a series of 2D cross-sectional FDTD schemes may be applied to sufficiently represent low frequency sound propagation throughout a 3D enclosure. Hence, through combination of RIRs rendered in this way with those produced by mid-high frequency GA models, a spectrally complete sonic description of an environment may be produced in a computationally inexpensive manner. The design, implementation and validation of this *Multiplane* FDTD HAM is documented in Chapter 5, following coverage of considerations for generic FDTD scheme development presented in the next chapter.

Chapter 4

Developing 'SRL' Finite Difference Schemes for Acoustic Simulations

Based on previous discussion of approaches to acoustic modelling, the Finite Difference Time Domain (FDTD) method has been identified as an appropriate means of simulating low frequency room acoustic behaviour for the purposes of this study. This chapter addresses important properties of FDTD acoustic models and investigates the implementation of key aspects of these models such as sound source/receivers and realistic boundary surface conditions.

4.1 Derivation of Standard Rectilinear (SRL) Finite Difference Schemes

The *standard rectilinear* (SRL) FDTD scheme has been previously used for numerical acoustic modelling in a range of studies to simulate low frequency sound propagation in rooms [80, 81, 128–131]. The derivation of such a scheme begins with centered-difference approximations (see section 3.2.1) to the second-order partial derivative terms of the wave equation (2.8, 2.30 & 2.50) for the required number of spatial dimensions. This acts to discretise the continuous time (t) and Cartesian spatial dimensions ($[x, y, z]$) given:

$$t = nT, \qquad x = lh, \qquad y = mh, \qquad z = qh \tag{4.1}$$

where T and h are the temporal and spatial sampling intervals respectively. In the following, T will be referred to as the *time step* and h the *grid spacing*. The values n,

l, m and q are integers that describe the position of discrete pressure values in time and space. Each instance of pressure over the scalar pressure field is termed a *node*. Using these variables and centered difference approximations, the second order time derivative of pressure in 3D may be approximated as follows:

$$\frac{\partial^2 p}{\partial t^2} \approx \frac{p(t+\triangle t) - 2p(t) + p(t-\triangle t)}{\triangle t^2} \Rightarrow \frac{p^{n+1}_{l,m,q} - 2p^n_{l,m,q} + p^{n-1}_{l,m,q}}{T^2} \tag{4.2}$$

Extending this practice to the spatial derivatives of (2.50) yields a discrete approximation to the 3D wave equation:

$$\frac{p^{n+1}_{l,m,q} - 2p^n_{l,m,q} + p^{n-1}_{l,m,q}}{T^2} = c^2\Big(\frac{p^n_{l+1,m,q} - 2p^n_{l,m,q} + p^n_{l-1,m,q}}{h^2} + \frac{p^n_{l,m+1,q} - 2p^n_{l,m,q} + p^n_{l,m-1,q}}{h^2} + \frac{p^n_{l,m,q+1} - 2p^n_{l,m,q} + p^n_{l,m,q-1}}{h^2}\Big) \tag{4.3}$$

where $p^{n+1}_{l,m,q}$, $p^n_{l,m,q}$ and $p^{n-1}_{l,m,q}$ are, respectively, the next, current and previous pressure values for the node position [lh,mh,qh] and time $t = nT$. From this expression, the SRL FDTD iterative update scheme may be derived by solving for the single unknown term $p^{n+1}_{l,m,q}$:

$$p^{n+1}_{l,m,q} = \lambda^2 S^n_{l,m,q} + 2(1-3\lambda^2)p^n_{l,m,q} - p^{n-1}_{l,m,q} \tag{4.4}$$

where,

$$S^n_{l,m,q} = p^n_{l+1,m,q} + p^n_{l-1,m,q} + p^n_{l,m+1,q} + p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q-1} \tag{4.5}$$

and

$$\lambda = \frac{cT}{h} \tag{4.6}$$

Such a scheme may be envisaged as a rectilinear lattice of pressure nodes distributed throughout a given topology with a constant inter-nodal distance in all dimensions. This lattice represents the discretised scalar pressure field arising in a medium (air) in response to a disturbance and facilitates simulation of wave propagation by means of (4.4). Following the same method of formulation, the 2D SRL FDTD update scheme for the wave equation may be derived as:

$$p^{n+1}_{l,m} = \lambda^2 S^n_{l,m} + 2(1-2\lambda^2)p^n_{l,m,q} - p^{n-1}_{l,m} \tag{4.7}$$

where, in this case:

$$S^n_{l,m} = p^n_{l+1,m} + p^n_{l-1,m} + p^n_{l,m+1} + p^n_{l,m-1} \tag{4.8}$$

The SRL FDTD scheme is one example of many numerical approximations that may be

applied to calculate the wave equation. Alternative schemes, such as those documented in [110], apply a different numerical approximation to the spatial Laplacian operator leading to a variance in the characteristics and accuracy of wave simulation across a discretised domain. For example, 'Interpolated Wideband' [110, 123] and higher-order explicit [131, 132] schemes have been demonstrated as superior to SRL formulations in terms of run-time/memory requirements for a given simulation bandwidth. However, the SRL formulation remains one of the more simple topologies for which to implement realistic boundary conditions for the following reasons (which will be expanded upon in section 4.4):

- Development of 'Locally Reacting Surface' (LRS) boundary conditions has shown that this condition remains numerically stable for all values of reflection coefficient in SRL schemes [110]. The same cannot be stated for higher order explicit schemes such as those documented in [132].
- Imposition of the LRS boundary condition in 3D SRL topologies requires only 27 distinct pressure node update formulations to simulate an arbitrary bounded topology with rectilinear spatial sampling. The interpolated wideband approach requires a further 72 formulations in order to handle re-entrant boundary nodes sufficiently.

Hence, the SRL scheme poses the most straight-forward, reliable means of producing FDTD based sound wave simulations for arbitrary topologies. For the purposes of this work, the accuracy of resulting acoustic models is deemed sufficient (as discussed in the following section) when considering the ease of implementation greatly reduces potential for algorithmic errors in simulation code. Finally, the widespread use of SRL schemes in recent research (see e.g. [133–135]) is testament to its applicability to wave-based modelling scenarios.

4.2 Numerical Stability and Dispersion

The spatio-temporal discretisation required to derive a computable solution of the wave equation introduces numerical errors to any given FDTD scheme. These errors are due to the truncation of Taylor's series expansion used to derive the difference approximations (see section 3.2.1) and the employment of finite precision arithmetic during digital computation. The inherent FDTD algorithmic inaccuracies impact on both the numerical stability and the conditions under which wave propagation in multidimensional schemes may be regarded as sufficiently isotropic. The following provides

an overview of the mathematic means by which the issues of numerical stability and dispersion may be investigated in order to quantify the stability and accuracy of SRL simulations.

4.2.1 Numerical Stability

Numerical stability, in the context of this work, refers to the absence of continuously amplified solutions to a given discrete-time recursive update equation. For example, the wave equation is solved for all simulated frequencies by means of (4.4). If a particular solution is allowed to grow continuously over time, the output of the algorithm will tend towards an infinite value at the corresponding frequency of resonance, thus greatly overpowering and obscuring the behaviour of the scheme at all other simulated frequencies. Hence, preservation of stability is crucial in FDTD schemes for simulating realistic wave propagation and producing usable/meaningful results.

Conditions for stability in discrete-time algorithms may be derived by means of von Neumann analysis [93, 136]. In relation to FDTD wave equation schemes, this analysis places constraints on the constant values in (4.4) and (4.7) to ensure that growing solutions are not present in the algorithm. For the case of the 2D SRL scheme (4.7), a single frequency ansatz of the form:

$$p_{l,m}^{n} = e^{jwnT} e^{jk_x lh + jk_y mh} \tag{4.9}$$

is inserted to the discretised 2D wave equation as a trial solution. This results in the following expression:

$$e^{jwT} - 2 + e^{-jwT} = \lambda^2 (e^{jk_x h} - 2 + e^{-jk_x h} + e^{jk_y h} - 2 + e^{-jk_y h}) \tag{4.10}$$

Applying the discrete time relation $\mathrm{z} = e^{jwT}$ and noting that $e^{j\theta} - 2 + e^{-j\theta} = -4\sin^2(\frac{\theta}{2})$:

$$\mathrm{z} - 2 + \mathrm{z}^{-1} = \lambda^2 \left[-4\sin^2\left(\frac{k_x h}{2}\right) - 4\sin^2\left(\frac{k_y h}{2}\right) \right] \tag{4.11}$$

which may be written as:

$$\mathrm{z}^2 + 2B\mathrm{z} + 1 = 0 \tag{4.12}$$

where,

$$B = -1 + 2\lambda^2 \left[\sin^2\left(\frac{k_x h}{2}\right) + \sin^2\left(\frac{k_y h}{2}\right) \right] \tag{4.13}$$

In order to preserve numerical stability, the condition $|\mathrm{z}| \leq 1$ must be ensured. Adhering to this condition positions the solutions to (4.12) (and therefore (4.7)) within the

unit circle of the z-plane such that no growing solutions exist for all wave numbers k_x and k_y [136]. Following from (4.12), it can be stated that:

$$(\mathrm{z} + B)(\mathrm{z} + B) = 0 \tag{4.14}$$

and hence,

$$\mathrm{z} = -B \tag{4.15}$$

Therefore, in order to realise the condition $|\mathrm{z}| \leq 1$, an equivalent condition exists for real-valued B:

$$|B| = -1 + 2\lambda^2 \left[\sin^2\left(\frac{k_x h}{2}\right) + \sin^2\left(\frac{k_y h}{2}\right) \right] \leq 1 \tag{4.16}$$

The above expression may be simplified by considering the maximum unity value of the $\sin^2$ terms leading to:

$$-1 + 4\lambda^2 \leq 1 \tag{4.17}$$

which imposes the constraints on the constant values applied in (4.7):

$$\lambda = \frac{cT}{h} \leq \frac{1}{\sqrt{2}} \tag{4.18}$$

The parameter λ is referred to as the *Courant number* and the maximum value it can take while retaining stability ($\lambda = \frac{1}{\sqrt{2}}$) is termed the *Courant limit*.

Stability analysis can be conducted following the procedure above for the case of the 3D SRL approximation to the wave equation. Insertion of a trial solution $p^n_{l,m,q} = e^{jwnT} e^{jk_x lh + jk_y mh + jk_z qh}$ into (4.4) leads to a z-domain expression with the same form as (4.12), with:

$$B = -1 + 2\lambda^2 \left[\sin^2\left(\frac{k_x h}{2}\right) + \sin^2\left(\frac{k_y h}{2}\right) + \sin^2\left(\frac{k_z h}{2}\right) \right] \tag{4.19}$$

Again, noting that $\sin^2(\theta) \in [0,1]\ \forall\theta$, the maximum value of each sine term is known to be unity. Hence, the condition on B, such that $|\mathrm{z}| \leq 1$ for the 3D scheme may be written:

$$|B| = -1 + 6\lambda^2 \leq 1 \tag{4.20}$$

giving the constraint on the Courant number as:

$$\lambda \leq \frac{1}{\sqrt{3}} \tag{4.21}$$

4.2.2 Dispersion Analysis

The speed of sound in air is constant for all frequencies in real environments. However, in discretised SRL simulations the propagation of modelled sound waves is anisotropic and, therefore, the characteristics of propagation are direction-dependent. In order to quantify the extent to which this non-uniform propagation impacts on simulation results, dispersion analysis examines the phase velocity of sound waves as they travel over a discrete spatial domain. Ideally, modelled phase velocity will be constant for all frequencies and directions of travel in order to be consistent with real systems. Hence, dispersion error is measured by comparing the characteristics of phase velocity in discretised solutions to the wave equation with the ideal, realistic case. This measure is termed the 'relative phase velocity', $\hat{v}$, and is defined as:

$$\hat{v}(\mathbf{k}, \theta) = \frac{\hat{c}}{c} = \frac{\omega}{\mathbf{k}c} \tag{4.22}$$

where c is the ideal, constant wave propagation speed, $\hat{c}$ is the wave speed produced in a given simulation for angular frequency ω and wavenumber $\mathbf{k}$. From this expression, it is possible to determine the relative phase velocity for a prescribed propagation direction provided by $\mathbf{k}$ and frequency. In order to calculate a value for (4.22) for a particular scheme, ω must be derived from a trial solution of the scheme under investigation. As with the case of numerical stability analysis, a single frequency ansatz is employed for this purpose.

2D SRL Dispersion Error:

For the case of the 2D SRL wave equation scheme, insertion of the ansatz $p^n_{l,m} = e^{j\omega nT} e^{jk_x lh + jk_y mh}$ into (4.7) results in:

$$4\sin^2\left(\frac{\omega T}{2}\right) = \lambda^2 \left[4\sin^2\left(\frac{k_x h}{2}\right) + 4\sin^2\left(\frac{k_y h}{2}\right) \right] \tag{4.23}$$

for wavenumber $\mathbf{k} = \sqrt{k_x^2 + k_y^2}$. Re-arranging the above expression for ω gives:

$$\omega = \frac{2}{T}\sin^{-1}\left(\lambda \sqrt{\sin^2\left(\frac{k_x h}{2}\right) + \sin^2\left(\frac{k_y h}{2}\right)} \right) \tag{4.24}$$

which may be inserted into (4.22) to provide the relative phase velocity occurring in each direction around the azimuthal angle for a given Courant number, λ. Referring back to the stability condition imposed on the Courant number (4.18) for 2D SRL schemes, it is apparent that the smallest grid spacing h occurs when λ is set to the Courant limit giving $h = \sqrt{2}cT$. Under this condition, the resolution of the spatial

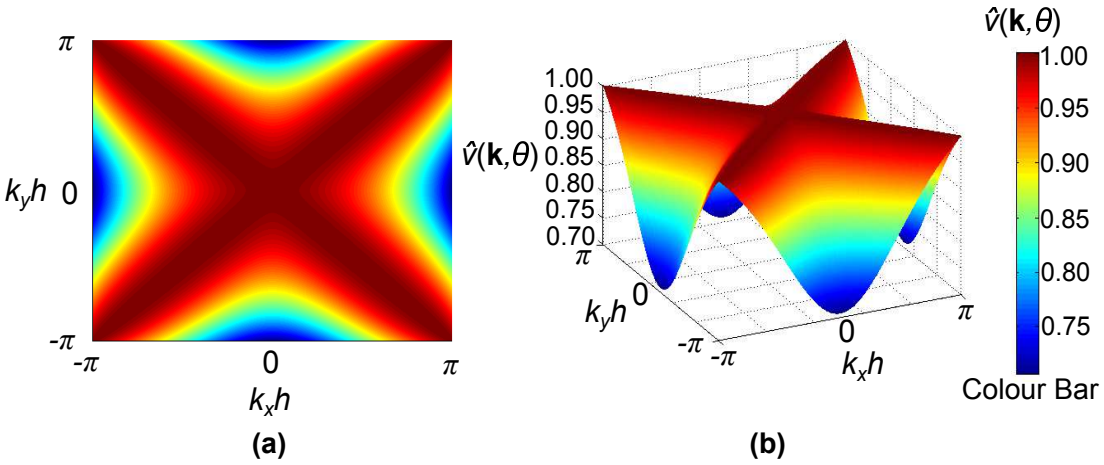

FIGURE 4.1: (a) Colour map depiction of relative phase velocity $\hat{v}(\mathbf{k},\theta)$ for the SRL FDTD scheme for wavenumber contributions k_x and k_y with grid spacing h. (b) Surface plot of SRL relative phase velocity. Colour bar scale is provided for reference.

sampling grid is as fine as possible while retaining numerical stability and, therefore, dispersion error is reduced to its minimum value for a given wave speed and time step. The relative phase velocity in the horizontal plane for a 2D SRL scheme operating at the Courant limit is depicted in Figure 4.1. As shown, the relative phase velocity is unity in directions of travel that are diagonal to the axes of the grid for all wave number contributions k_x and k_y. As such, the phase velocity is constant at all frequencies up to 0.5 F_s, where F_s is the temporal sampling rate (Hz), in diagonal directions only. Moreover, the dispersion pattern shown possesses symmetry in both x and y axes.

As the direction of travel deviates from diagonal paths, the relative phase velocity is shown to decrease markedly as wave frequency and angle of deviation from the diagonal increases. Hence, for all non-diagonal directions, wave component propagation speeds decrease as the frequency of components increases. The worst instances of dispersion error occur in axial directions aligned with the SRL grid.

By measuring the two extremes of dispersion error, it is possible to define the extent to which the SRL scheme may be considered reasonably isotropic. Following the discussion in [110], an acceptable level of anisotropy is defined for a range of frequencies in which the maximum deviation of relative phase velocity is no greater than 2% for all azimuthal angles. For the 2D SRL scheme, when using the axial directions to find the relative phase velocity, this is calculated as 0.1 F_s [110]. Therefore, to produce a simulation response accurate up to a desired upper frequency of f_{max}, a 2D SRL scheme must be run at the Courant limit with a temporal sampling rate of $F_s \geq 10\ f_{max}$. This condition is adhered to for all simulations documented in Chapter 5 for the purposes of rendering low frequency RIRs.

3D SRL Dispersion Error:

In the case of 3D SRL schemes, the calculation of relative phase velocity is dependent on the following expression for ω derived by inserting a single frequency ansatz into (4.3):

$$\omega = \frac{2}{T}\sin^{-1}\left(\lambda\sqrt{\sin^2\left(\frac{k_x h}{2}\right)+\sin^2\left(\frac{k_y h}{2}\right)+\sin^2\left(\frac{k_z h}{2}\right)}\right) \tag{4.25}$$

given for 3D wave number $\mathbf{k} = \sqrt{k_x^2 + k_y^2 + k_z^2}$. Hence, in this instance, the relative phase velocity is defined for both azimuthal and elevation angles θ and ϕ respectively:

$$\hat{v}(\mathbf{k}, \theta, \phi) = \frac{\omega}{\mathbf{k}c} \tag{4.26}$$

As with the 2D SRL scheme, it may be shown that the relative phase velocity is unity in directions diagonal to all axes for schemes operating at the Courant limit. Therefore, dispersion error does not arise in diagonal paths of propagation with respect to the rectilinear lattice of grid points. The lowest values of relative phase velocity occur in axial directions, which reveals the direction of maximum dispersion. Again, by imposing the condition for sufficient isotropy of wave propagation whereby $\hat{v}(\mathbf{k}, \theta, \phi) \geq 0.98$ for a maximum error in $\hat{v}$ of 2%, the maximum valid simulation frequency for 3D SRL simulations is 0.075 F_s [110, 137].

While this result suggests a lower valid bandwidth for the 3D scheme compared to 2D simulations previously discussed, studies on the perceptual impact of dispersion effects on auralisations imply that usable bandwidths may be as high as 0.15 F_s for 3D SRL models [138]. However, it should be noted that this bandwidth may only hold for particular source audio material and domain topologies. The perception of dispersion depending on the distance of source/listener separation and scheme topology remains an area of active research (see e.g. [137]). Hence, for the purposes of producing auralisations in this work, the objective measure of a maximum 2% error in relative phase error is adhered to in all simulations that are used to produce audible results.

4.3 Source Excitation and Virtual Sound Receivers

As discussed in section 2.3.7, the creation of a RIR involves the appropriate excitation of an acoustic environment and the measurement of resulting pressure fluctuations local to a receiver location. In FDTD schemes, this process requires appropriate initial conditions to be defined and applied to the wave equation approximation. The capture of simulated sound behaviour at a listening position (or positions) within a modelled domain is handled by simply recording the time-varying pressure value occurring at a

single grid point. However, in order to render spatial impulse responses, it is necessary to capture multiple grid point outputs and utilise post-processing techniques to create receivers with appropriate directivity functions. Details of the source and receiver implementations employed for this study are provided in the following for reference.

4.3.1 Sound Excitation Sources

The theoretical excitation function associated with RIR measurement is the continuous-time Dirac delta function as defined by (2.9) and (2.10). The discrete-time equivalent of this excitation signal is the Kronecker delta function, defined for an excitation at discrete-time sample τ:

$$d[n,\tau] = \begin{cases} 1, & n = \tau \\ 0, & n \neq \tau \end{cases} \tag{4.27}$$

where n is the discrete-time sample index. The Kronecker delta can be demonstrated as being the discrete-time equivalent of $\delta(t)$ using two properties. Firstly, given (4.27) it may be stated that for a discrete time function, $p[n]; n = [-N : N]$:

$$\sum_{n=-N}^{N} p[n]d[n,\tau] = p[\tau] \tag{4.28}$$

which is analogous to the sifting property of the Dirac delta function, stated here for a continuous-time function $f(t)$:

$$\int_{-\infty}^{\infty} f(t)\delta(t-t_0)dt = f(t_0) \tag{4.29}$$

Secondly, $\delta(t)$ is known to possess an ideally flat magnitude spectrum with equal energy at all frequencies. This property is demonstrated by:

$$F[\delta(t-t_0)] = \int_{-\infty}^{\infty} e^{-j\omega t}\delta(t-t_0) = e^{-j\omega t_0} \tag{4.30}$$

where,

$$e^{-j\omega t_0}\Big|_{t_0=0} = 1; \qquad \forall\omega \in \mathbb{R} \tag{4.31}$$

Hence, the magnitude spectrum of $\delta(t)$ is shown to be equal at all frequencies when the impulse is defined at time $t = 0$. Applying the discrete Fourier transform to $d[n,\tau]$ yields the same result:

$$\sum_{n=-N}^{N} d[n,\tau]e^{-j\omega\frac{n}{N}} = e^{-j\omega\frac{\tau}{N}} = 1; \qquad \tau = 0, \forall\omega \in \mathbb{R} \tag{4.32}$$

Hence, the Kronecker delta function is shown to be a suitable form of impulsive excitation function for the purposes of simulating RIRs. However, further consideration must be given to how the impulse is applied to a discrete finite difference pressure domain in order to produce realistic and serviceable results.

4.3.1.1 Generic Source Implementation

The various methods of exciting the discretised acoustic field of FDTD models have gained recent attention in related studies [139, 140]. There are three notable approaches to sound source implementation in FDTD simulations. These are commonly referred to as *soft*, *hard* and *transparent* sources. Referring to the discussion provided in [139], each source type is briefly described here.

Soft Source:

The soft source type is defined as an input signal that is summed with the discrete pressure field of a FDTD model at a grid point located at the given sound source location. In the case of a Kronecker delta excitation function, a soft source is implemented by means of initial conditions. To obtain a RIR, the pressure field of the model is assumed to be in an equillibrium state (i.e. all nodes are given a zero value) with an exception at the source location which is initialised to a unity value. Upon iterative updating of the FDTD wave equation, the impulse propagates away from the source location with monopole directivity characteristics. No further treatment of the source node is required. In the case of exciting the pressure field with a discrete-time varying pressure signal, each sample of the signal is summed with the pressure arising at the source node. A notable property of the soft source is that it allows preservation of resonance at DC for systems that incorporate phase-preserving boundaries. Referring to section 2.3.6, there is a solution to the wave equation at 0 Hz for such systems. Hence, it is expected that the resulting RIR will exhibit DC drift, an example of which is displayed in Figure 4.2.

Hard Source:

In a similar fashion to soft sources, a Kronecker delta excitation can be implemented as a hard source by intialising the source node to a unity value. The difference between the two source types arises in the treatment of the source node after the imposition of the excitation. For a hard source, the source node is fixed to a zero value after the initial excitation according to (4.27). This acts to clamp or 'fix' the source node such that the global solution to the wave equation rendered by the FDTD scheme does not facilitate preservation of the solution at DC. In simple terms, the incorporation of a clamped node

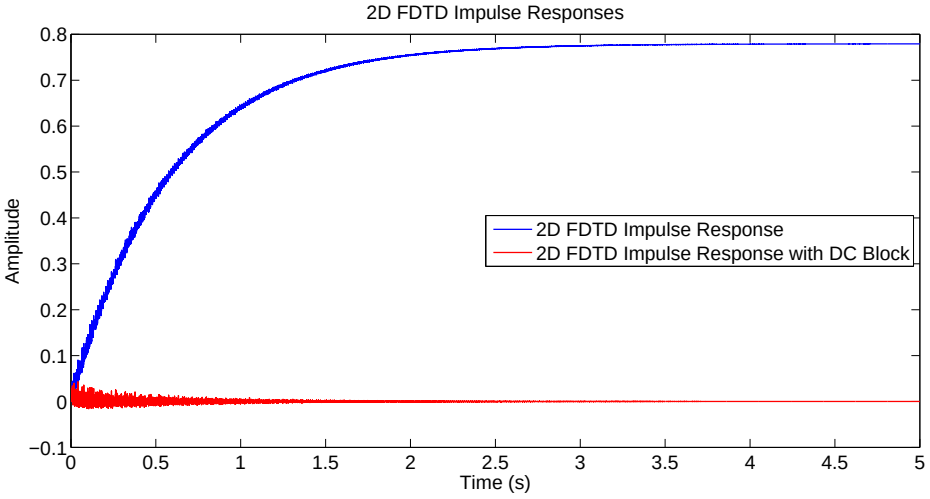

FIGURE 4.2: Comparison of 2D FDTD impulse responses generated using soft source implementation. The response given in blue exhibits the expected DC offset drift while the response given in red shows the same impulse response after filtering with a DC blocking filter.

in the pressure field creates a Dirichlet boundary node in the domain. While this has the positive effect of removing DC drift from resulting RIRs, the source node becomes an unwanted wave scatterer leading to deviations in resonant frequency values from theoretical values. In the case of a time-varying excitation signal, the source node is forced to take on the value of each signal sample regardless of the pressure fluctuations that occur at the source node.

Transparent Source:

The transparent source type seeks to provide the advantage of the hard source (suppression of DC drift) while avoiding the implications of introducing a fixed boundary node in the pressure field. This is achieved by applying a compensation signal to the source node update equation during a simulation in which the source node is forced to take on the value of the excitation signal at each sample. The compensation signal is defined as the free-field impulse response generated using a hard source excitation convolved with the required input excitation signal and is subtracted from the update expression for the source node only. Hence, the update expression for a source node for 2D transparent source implementation may be written, after [139, 140]:

$$p_{l,m}^{n+1} = \lambda^2 S_{l,m}^n + 2(1-2\lambda^2)p_{l,m}^n - p_{l,m}^{n-1} + p_{in}^{n+1} - \sum_{t=0}^{n} h_{2D}^{n-t+1} p_{in}^t \tag{4.33}$$

$$h_{2D}^n = \lambda^2 S_{l,m}^{n-1} + 2(1-2\lambda^2)p_{l,m}^{n-1} - p_{l,m}^{n-2} \tag{4.34}$$

where p_{in} is the desired excitation signal. By removing the response of the scheme to the excitation signal at the source location, the scattering effects of the hard source node

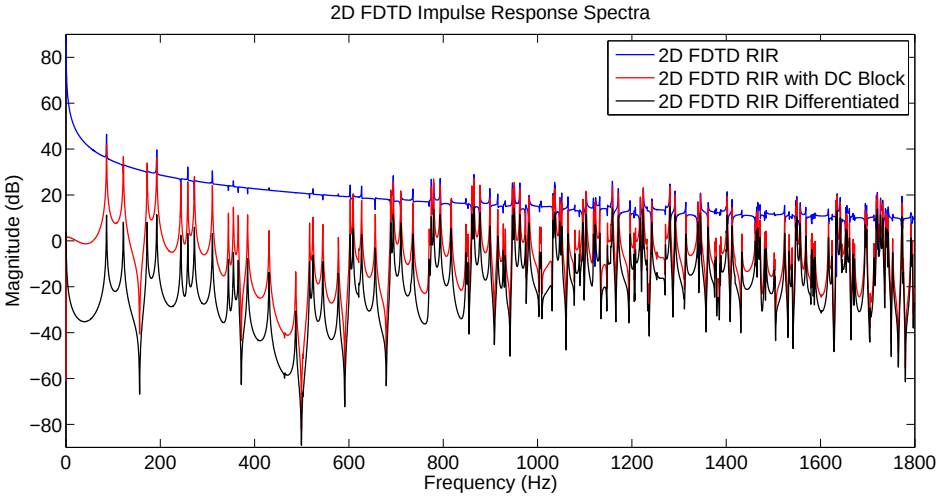

FIGURE 4.3: 2D FDTD impulse response spectrum resulting from soft source excitation (blue) showing a large DC component. The red and black spectra depict the frequency content of the same impulse response after DC blocking filtering and differentiation respectively. Note the low frequency roll-off in the case of the differentiated response.

are effectively removed. The result is a global scheme response that neither exhibits DC offset drift nor irregularities due to the constrained source node. However, rendering this source type comes at the cost of computing a suitable compensation signal either prior to simulation or as a post-processing stage. As noted in [139], this can lead to significant computational overhead, especially for high resolution 3D FDTD schemes.

For the purposes of this work, the soft source excitation type is utilised. The justification for this approach stems from the issues associated with the alternative hard and transparent sources in terms of undesirable wave scattering effects and increased computational efforts respectively. However, as noted above, impulse responses obtained from soft source excitation require pre- or post-processing in order to remove unwanted drift due to resonance at DC. An example of DC drift as rendered in a simple 2D FDTD scheme and the resulting magnitude spectrum are shown in Figures 4.2 and 4.3 respectively. In [140], it is suggested that the offset present in RIRs rendered using a Kronecker delta soft source excitation may be removed by differentiating the output signal or by exciting the pressure field with a differentiated Kronecker delta function. While this does indeed suppress the DC component, the act of differentiation creates an unwanted roll-off of low frequency components [140]. This roll-off characteristic is depicted in Figure 4.3 for the case of the impulse response shown in Figure 4.2.

A solution to the problem of DC drift which has less of an impact on the audible frequency range of the resulting output is realised through use of a DC blocking filter. A Kronecker delta function may be filtered using a 1^{st}-order high pass filter with a cut-on frequency of 20 Hz prior to simulation. However, this may only be carried

out if the impulse response length is known in advance of simulation. The resulting filtered delta function can then be summed into the pressure field as per the soft source implementation. Alternatively, the captured response to a simple soft source Kronecker delta function can be filtered using the DC blocking filter post-simulation. Owing to the linear time-invariant nature of the discrete wave equation, both approaches yield the same output signal. Hence, this approach to soft source excitation is deemed optimal as DC offset drift is suppressed while incurring minimal computational overhead and avoiding source node scattering effects. Furthermore, as displayed in Figure 4.3, the low frequency roll-off introduced by the alternative differentiation approach described previously is also circumvented.

4.3.1.2 2D Source Implementation

As with the case of 3D FDTD source implementation, a simple omnidirectional impulsive soft source can be produced by initialising the grid point at the source location to a unity value. However, application of a Kronecker delta function to a 2D FDTD simulation of the wave equation inherently gives rise to the afterglow phenomenon [30, 33] as per the associated Green's function (see section 2.1.3.2). The afterglow effects manifest as non-zero pressure values which trail the passage of the first wavefront. While this is the expected outcome of 2D wave equation simulations, there are scenarios in which these effects are undesirable. One such scenario is the emulation of 3D sound wave propagation using 2D FDTD schemes. This process is central to the studies presented in this work and, therefore, the removal of afterglow effects from resulting simulation results requires investigation.

The basic notion of afterglow removal may be understood as a means by which the 2D Green's function for the wave equation may be manipulated into a form that resembles the 3D equivalent. The aim is not to seek an equality between the two functions, but rather attain similar function characteristics. A critical concern in numerical RIR simulation is the resulting frequency content of the output. Figure 4.4 provides a visual comparison between the resulting spectral response of a 3D SRL FDTD model of a simple 2 x 2 x 2 (m) cube and a 2D cross-sectional model of the space using a 2D SRL FDTD scheme. Each simulation is run at the respective Courant limit with a temporal sampling rate, F_s= 44.1 kHz. Frequency-independent Locally Reacting Surface boundary conditions, which are fully described in section 4.4, are applied to provide an absorption coefficient of $\alpha = 0.01$ in both 3D and 2D simulations. As can be seen from the spectra displayed in Figure 4.4, it is evident that the 2D cross-sectional RIR is of greater magnitude in the low frequency range, f = [0:200] Hz. This result is consistent with those documented in [31] and is a direct consequence of the afterglow

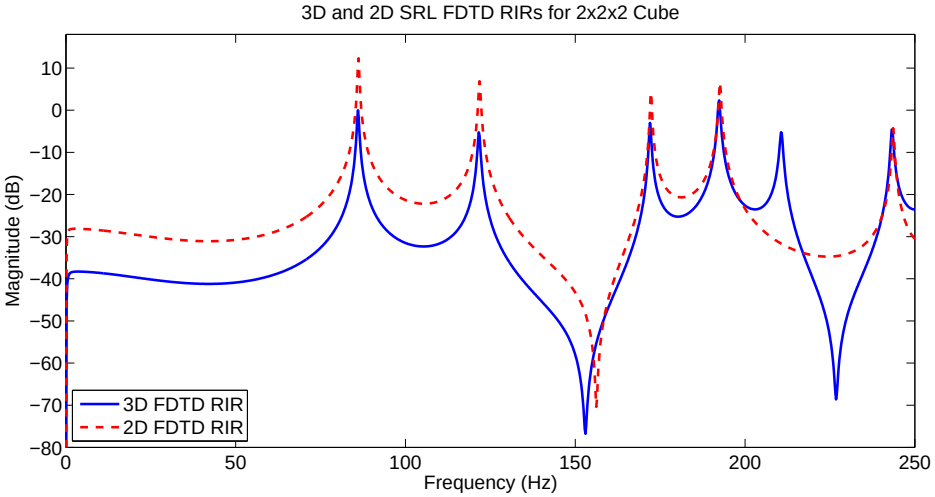

FIGURE 4.4: Low frequency spectra of 3D (blue) and 2D (red dashed) SRL FDTD simulated RIRs for a simple cubic room topology. A significant increase in the magnitude of spectral peaks in the low frequency range is observed in the 2D RIR case compared to the 3D RIR.

phenomenon. In order to better equalise the low-frequency response characteristics between the 3D and 2D cases, the 2D RIR may be treated by considering the associated Green's function (2.32), re-written in the following form:

$$p(x,y,t) = \frac{A}{2\pi}\frac{H(t-r/c)}{\sqrt{t+r/c}\sqrt{t-r/c}}; \qquad r = \sqrt{x^2+y^2} \tag{4.35}$$

as per (4) of [31]. Now, considering the case where $t \approx r/c$, (4.35) may be reduced to the following form to yield the Green's function as it tends towards a maximum value:

$$p(x,y,t) \approx \left(\frac{Ac}{8\pi^2 r}\right)^{\frac{1}{2}} \frac{H(t-r/c)}{\sqrt{t-r/c}} \tag{4.36}$$

Hence, under this condition, the 2D Green's function can be recognised as a displaced Kronecker delta function of amplitude $\frac{Ac}{8\pi^2 r}$ (where A is the amplitude of the initial excitation) multiplied by a scaled Heaviside step function $H(t-r/c)$. Following [31], this step function may be assumed to preserve minimum-phase properties. As a result, it is possible to remove the step function component through deconvolution of (4.36) using the inverse filter:

$$p_{2D_{inv}} = \frac{H(t)}{\sqrt{t}} \tag{4.37}$$

Note that the time dependence on the term r/c has been removed as this serves only to represent the decay proportional to $r^{-\frac{1}{2}}$ as the wavefront propagates outward from the source location. It therefore follows that a 2D FDTD RIR can be defined for the

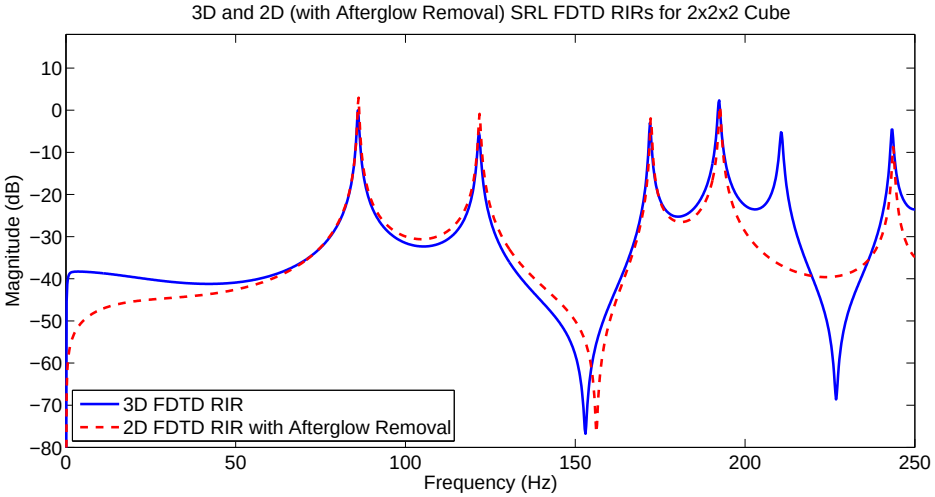

FIGURE 4.5: Low frequency spectra of 3D (blue) and 2D (red dashed) SRL FDTD simulated RIRs for a simple cubic room topology. The 2D RIR has been treated with afterglow removal processing to reduce magnitude differences between 3D and 2D cases at low frequencies.

case of equalised responses with afterglow removal applied:

$$p(x,y,t)_{2D_{RIR}} = p_{2D_{inv}} \star p(x,y,t) \tag{4.38}$$

where $\star$ signifies convolution in the time domain. The above expression serves to approximate the Kronecker delta function response characteristics common to 3D FDTD simulations. The result of applying afterglow removal to the RIR cases previously displayed previously in Figure 4.4 is shown in Figure 4.5. As depicted, the difference between the low frequency spectral components of the 3D and 2D response is markedly reduced in magnitude. Therefore, it is evident that this process is significant towards the rendering of RIRs by means of 2D FDTD modelling such that results possess similar characteristics to 3D FDTD RIRs. The application of afterglow removal, as a means of post-processing RIRs, is further documented in Chapter 5.

4.3.2 Virtual Sound Receivers

In numerous acoustic modelling scenarios, it is desirable to utilise sound receivers that model those encountered in practice. For example, virtual receiver arrays may be used to collect spatial impulse responses which, in turn, facilitate the creation of auralisations with enhanced immersive attributes. In order to achieve such an outcome, a means of producing the directional characteristics of a microphone must be attained by adapting the process through which acoustic pressure data is retrieved from a given virtual acoustic model. The most direct and commonly employed method of obtain-

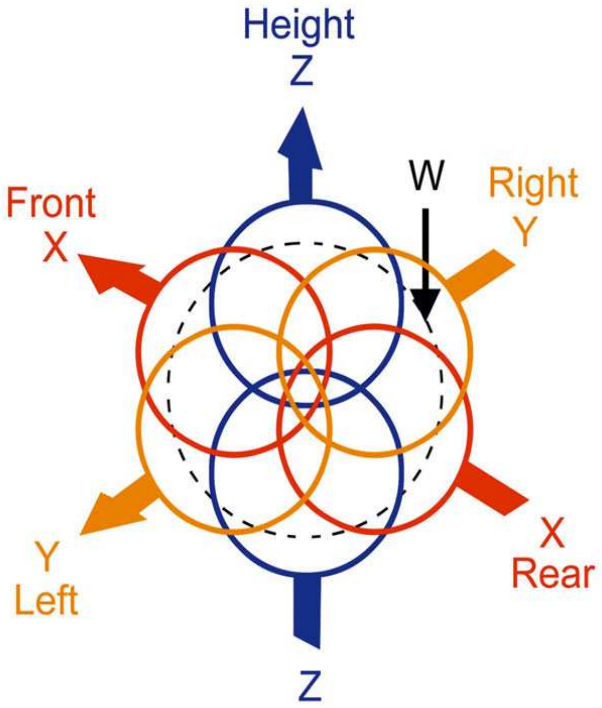

FIGURE 4.6: Directivity patterns for encoded B-Format channels W, X, Y and Z. Image sourced directly from [15].

ing impulse responses, or other such pressure signals, from FDTD simulations involves recording the numerical pressure values at a given grid point. This inherently gives rise to omnidirectional response characteristics at the chosen receiver location as equal consideration is given to sound incoming from all directions. The creation of directional sound receivers, therefore, relies on the ability to partially or fully reject sound from particular directions in order to achieve a desired response pattern.

One such directional receiver type is derived from *Ambisonics*. Ambisonics is a particular category of sound capture/rendering methods that are based on spherical harmonic decomposition of a soundfield. Initial work in this area, conducted by Gerzon [141, 142], led to the classification of the B-Format spatial audio encoding method. B-Format audio consists of four channels denoted W, X, Y and Z where the W channel refers to an omnidirectional receiver and the remaining X, Y and Z channels refer to pressure-gradient (bi-directional) receivers orientated in the x-, y- and z-axis respectively. The arrangement of directivity response patterns is depicted in Figure 4.6. As shown, the B-Format signals not only refer to pressure fluctuations local to the receiver location (W channel), but also to the directionality of the soundfield, which may be derived from the pressure-gradient signals (X, Y and Z channels), around the solid angle. The consequent ability to record the direction of the soundfield at a receiver location, in

turn, gives rise to the ability to spatialise sound components across a loudspeaker rig for the purposes of rendering accurate 3D soundscapes for auralisation purposes.

A notable advantage of Ambisonics arises from the way in which B-Format signals may be manipulated post-processing. Once the pressure-gradient characteristics of the soundfield over all axes are known, the soundfield may be rotated around the solid angle by applying simple trigonometric functions to the audio data. Thus, the directional perspective of a listening position may be altered during playback without having to repeat the measurement procedure. A method to derive the B-Format signals from RIR data simulated in a SRL FDTD model is presented here in order to detail the means by which 3D auralisations may be created using this acoustic modelling approach.

Extensive coverage of approaches to 1^{st}- Ambisonics encoding of numerical acoustic models is provided in related literature [7, 143–146]. Furthermore, the emulation of directional microphone characteristics in DWM-based room acoustics models has been investigated by Hacıhabiboğlu et al. [147]. While the work in [147] demonstrates successful implementation of realistic microphone directivity patterns, the underlying method is unsuitable for the creation of bi-directional characteristics.

The principle of the Ambisonics encoding in FDTD schemes, as per [146], is largely based on the Blumlein Difference Technique (BDT) [148] which allows the synthesis of higher order spherical harmonic directivity functions through manipulation of outputs produced by adjacent receivers with lower order directivity functions. For the purposes of this work, two adjacent omnidirectional (0^{th}-order) receiver outputs are combined to produce the 1^{st}-order pressure-gradient directivity by means of the governing equation, after [143, 146]:

$$v(t) \approx \frac{1}{\rho_0 d} \int_{-\infty}^{t} [p_1(\tau) - p_2(\tau)] d\tau \tag{4.39}$$

where $p_1(t)$ and $p_2(t)$ are the pressure signals recorded from two axially aligned grid nodes in an FDTD scheme, d is the spacing between the two grid nodes (m) and ρ_0 is the density of the modelled medium. The integration limits may be interpreted as covering the range of time t from the beginning of recorded RIRs, to the time at which the pressure fluctuation amplitudes are negligible. In terms of discrete-time RIRs obtained from FDTD simulations, (4.39) may be re-expressed as follows:

$$v[n] \approx \frac{1}{\rho_0 d} \Big[\sum_{i=0}^{n} p_1[i] - p_2[i] \Big]; \qquad n = [0 : N] \tag{4.40}$$

where n is the discrete time index for a signal of N samples. The arrangement for this receiver type in a 2D SRL FDTD scheme is depicted in Figure 4.7 where the pressure-gradient directivity functions are created from axially aligned omnidirectional receivers

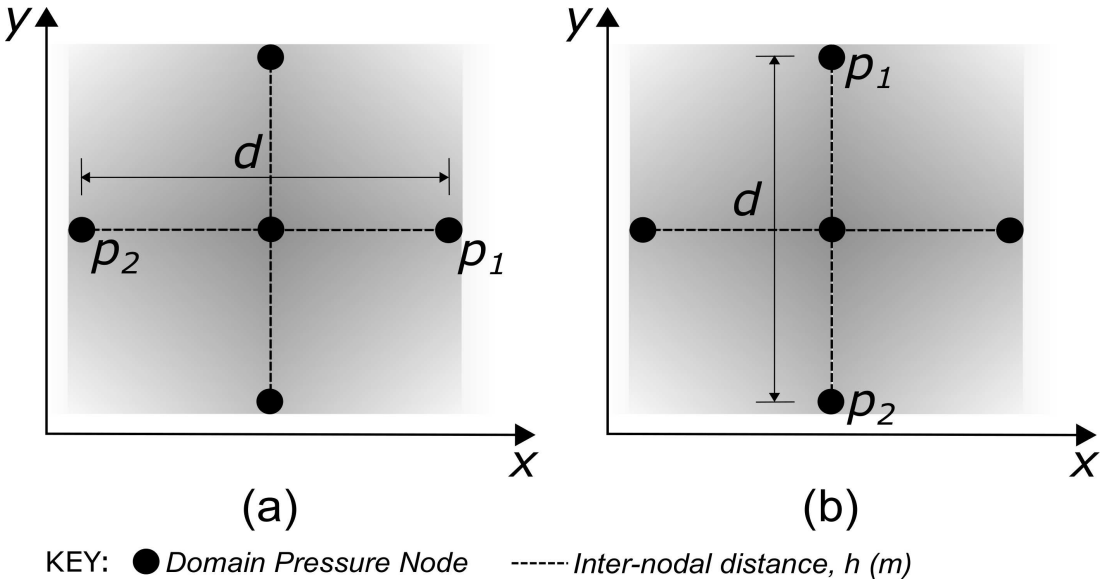

FIGURE 4.7: (a) Node arrangement for a single pressure-gradient output aligned with the x-axis in a 2D SRL FDTD scheme with contributing nodes p_1 and p_2 separated by distance $d = 2h$. (b) The same concept applied to a single pressure-gradient output aligned with the y-axis.

in the (a) x- and (b) y-directions giving the X and Y B-Format channels respectively using (4.40). The corresponding analytical directivity functions are displayed in Figure 4.8. Referring again to Figure 4.7, the arrangement of receivers is distributed around a single, central pressure node. The RIR captured from this central node provides the omnidirectional B-Format W channel. Hence, all required response patterns are coincident around a common central receiver location.

In terms of implementation, there is one further consideration that must be addressed in order to create appropriate B-Format RIRs. The scaling term and subtraction present in the discrete-time expression for pressure-gradient components, $v[n]$ (4.40), leads to an overall reduced amplitude relative to the omnidirectional W channel signal amplitude. This can be compensated for by calculating the expected first wavefront amplitude of the X and Y channel signals relative to the W channel wavefront amplitude based on the angle of incidence. Both X and Y signals may then be multiplied by an amplification factor to ensure consistency between the B-Format signal levels. These factors may be calculated from the following expressions for X and Y channels respectively:

$$G_X = \cos(\theta)\frac{p_{W_{fw}}}{p_{X_{fw}}}; \qquad G_Y = \sin(\theta)\frac{p_{W_{fw}}}{p_{Y_{fw}}} \tag{4.41}$$

where θ is defined as the azimuthal angle of incidence of incoming sound from a given source, $p_{W_{fw}}$, $p_{X_{fw}}$ and $p_{Y_{fw}}$ are the pressure amplitude value of the first wavefront captured on the W, X and Y channels respectively. Hence, G_X and G_Y yield the

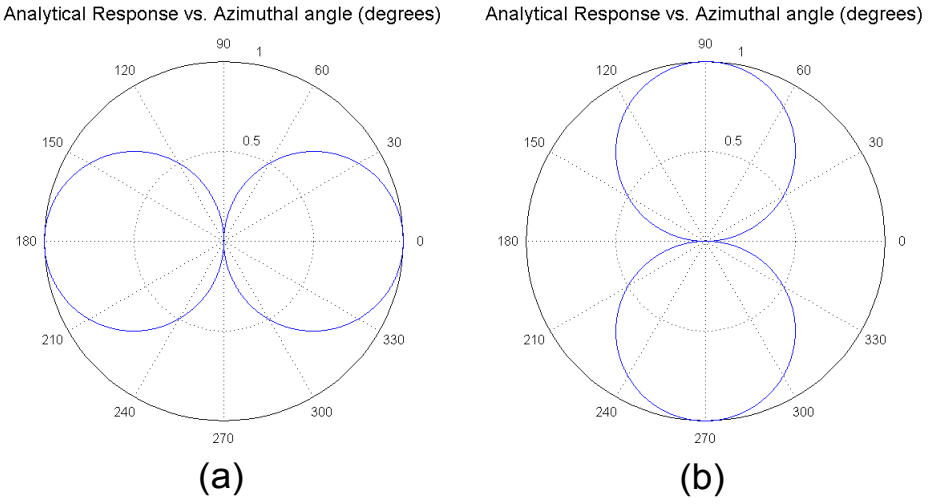

FIGURE 4.8: (a) Analytical directivity function for a pressure-gradient receiver aligned with the x-axis of a given scheme - note a negative polarity is present in the range $(90° : 270°)$. (b) Directivity function for pressure-gradient receiver aligned with the y-axis with negative polarity present in the range $(180° : 360°)$. (a) and (b) correspond to the X and Y B-Format channels respectively.

required amplification factors that equalise pressure levels between the corresponding directional B-Format channels and the omnidirectional response.

Finally, it is noted that the extension of the above process to 3D FDTD schemes is trivial as the additional vertical pressure-gradient signal may be calculated by analogous means.

4.4 Frequency-Independent Boundary Conditions

In order to accurately replicate enclosed sound propagation using FDTD simulations, it is vital that modelled boundary surfaces possess realistic characteristics. This leads to the requirement of appropriate boundary conditions that closely correspond to the reflection/absorption properties of real surface materials. Hence, the application of simple Neumann and Dirichlet conditions, while useful for validating the correctness of a given scheme through modal analysis, is not suitable for modelling real acoustic environments. Numerous examples of absorbing boundary conditions have been investigated in relation to numerical acoustic modelling paradigms (see e.g. [34, 89, 149]). In the context of this work, particular attention is given to *Locally Reacting Surface* (LRS) boundary conditions [3]. The incorporation of this boundary type into FDTD acoustic models has been examined extensively in the work of Kowalzyck and van Walstijn [110, 126, 150]. These studies demonstrate that the absorbing characteristics of the LRS condition may be accurately represented in FDTD schemes for the full

range of reflection coefficients, R = [-1:1], facilitating the simulation of surface material properties encountered in practice.

4.4.1 Locally Reacting Surface Boundaries

Following the discussion of sound wave reflection documented in [3, 110], locally reacting surface conditions may be assumed under the following conditions:

- Surface irregularities must be negated for irregularities that are significantly large in comparison to the shortest wavelength to be modelled. Hence, LRS boundaries in SRL schemes consider only planar surfaces and normal incidence reflection.
- The reflected waveform exiting a boundary is defined at a range of discrete positions over the boundary surface and sound propagation throughout the surface itself is considered negligible. Therefore, each spatially distributed reflected wave component is dependent on the local incident waveform and independent of adjacent boundary/wave interactions.

With these conditions in place, it is possible to derive the LRS condition for a discrete point on a boundary surface. Sound wave propagation, as discussed in Chapter 2, assumes conservation of mass and momentum within a given medium. However, at a boundary, conservation of mass is no longer a requirement as the boundary surface does not move. Therefore, conservation of momentum alone is sufficient to describe the behaviour of pressure waves at the boundary. From (2.42), it may be stated that the force acting on the boundary in the positive x-direction is given by:

$$F = -\frac{\partial p}{\partial x} \tag{4.42}$$

in which the force F may be calculated as the product of mass and acceleration, giving:

$$\frac{\partial p}{\partial x} = -\rho \frac{\partial u_x}{\partial t} \tag{4.43}$$

Following the derivation given in [110], the relationship between boundary impedance and particle velocity (2.65) is applied to the above expression to yield the boundary condition in terms of sound pressure only. Differentiating (2.65) with respect to time,

$$\frac{\partial u_x}{\partial t} = \frac{1}{\zeta_\omega \rho c} \frac{\partial p}{\partial t} \tag{4.44}$$

provides the velocity component in terms of pressure and the specific acoustic impedance of the boundary. Inserting this expression into the relationship for momentum conservation provides the LRS condition:

$$\frac{\partial p}{\partial t} = -c\zeta_\omega \frac{\partial p}{\partial x} \tag{4.45}$$

where, as previously described, the impedance of the boundary is related to the reflection coefficient by:

$$\zeta_\omega = \frac{1+R}{1-R} \tag{4.46}$$

Following the initial definition and use of this boundary condition, it has been applied in various FDTD acoustic modelling scenarios to simulate absorbing surface properties. However, it is noted in [110] that LRS boundaries are unsuitable for modelling purely anechoic surface properties as total absorption is only possible for waves interacting with the surface at normal incidence. Nevertheless, LRS conditions exhibit a considerable advantage over alternative boundary conditions (such as those given in [89]) as their implementation can be shown to preserve numerical stability for all realistic values of surface impedance [110].

4.4.2 LRS Implementation in SRL Schemes

At the boundary terminations of an SRL scheme, it is assumed that at least one pressure node required for the wave equation update will fall outside of the domain. This concept is depicted in Figure 4.9 where the nodes positioned beyond the interior are referred to as a *ghost nodes*. For the case of Dirichlet boundary conditions, these ghost nodes would be consistently set to a pressure of zero in order to implement physically clamped terminations producing a fixed boundary. However, for LRS boundaries the ghost nodes are treated as unknown quantities which are incorporated into the SRL wave equation update for each node adjacent to a surface by means of a discrete approximation to the boundary condition. There are two distinct approaches to achieving this outcome, as documented in [151], which are termed *pressure-centered* and *velocity-centered*.

4.4.2.1 Pressure-Centered LRS

The pressure-centered FDTD approximation to LRS conditions results from applying centered-difference finite difference terms to both the temporal and spatial derivatives in (4.45) giving (in 3D):

$$\frac{p^{n+1}_{l,m,q} - p^{n-1}_{l,m,q}}{2T} = -c\zeta_\omega \frac{p^{n}_{l+1,m,q} - p^{n}_{l-1,m,q}}{2h} \tag{4.47}$$

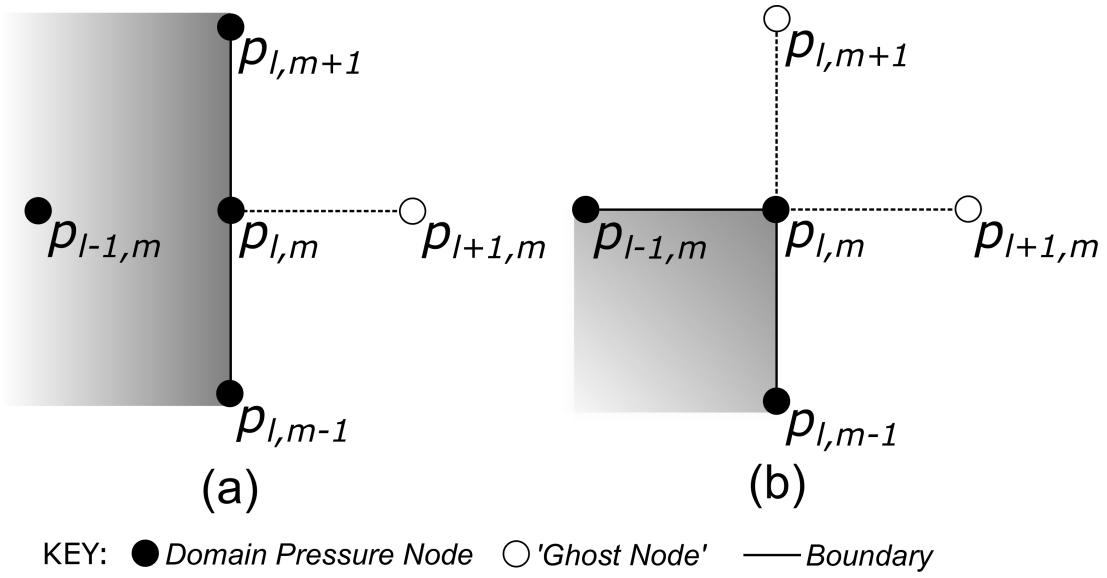

FIGURE 4.9: (a) Example of a 2D SRL, right hand boundary with known and unknown ('ghost') pressure node arrangement for the locally reacting surface condition. (b) 2D SRL corner boundary pressure node arrangement.

which yields the form investigated by Kowalczyk and van Walstijn [110, 126, 150]. The above expression corresponds to the case of a right hand, vertical boundary in the Cartesian coordinate system for which the ghost node is $p^n_{l+1,m,q}$. Rearranging for this unknown quantity gives:

$$p^n_{l+1,m,q} = -\frac{1}{\zeta_\omega \lambda}(p^{n+1}_{l,m,q} - p^{n-1}_{l,m,q}) + p^n_{l-1,m,q} \tag{4.48}$$

where λ is as previously defined. Similarly, a left hand, vertical boundary possesses the ghost node $p^n_{l-1,m,q}$ which may be solved for in the same manner:

$$p^n_{l-1,m,q} = -\frac{1}{\zeta_\omega \lambda}(p^{n+1}_{l,m,q} - p^{n-1}_{l,m,q}) + p^n_{l+1,m,q} \tag{4.49}$$

Having derived an expression for the unknown ghost node, this may now be inserted into the SRL wave equation update in order to apply the boundary condition. For a right hand, vertical surface the following update equation results for the pressure node adjacent and normal to the boundary:

$$\begin{aligned} p^{n+1}_{l,m,q} = \lambda^2 \Big[-\frac{1}{\zeta_\omega \lambda}(p^{n+1}_{l,m,q} - p^{n-1}_{l,m,q}) + 2p^n_{l-1,m,q} + p^n_{l,m+1,q} + p^n_{l,m-1,q} \\ + p^n_{l,m,q+1} + p^n_{l,m,q-1} \Big] + 2(1 - 3\lambda^2) p^n_{l,m,q} - p^{n-1}_{l,m,q} \end{aligned} \tag{4.50}$$

Rearranging the above expression to an explicit form in terms of $p^{n+1}_{l,m,q}$ results in the final LRS boundary update equation:

$$p^{n+1}_{l,m,q} = \Big[\lambda^2(2p^n_{l-1,m,q} + p^n_{l,m+1,q} + p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q-1}) + 2(1-3\lambda^2)p^n_{l,m,q} - \big(1 - \frac{\lambda}{\zeta_\omega}\big)p^{n-1}_{l,m,q}\Big]\left(\frac{1}{1+\frac{\lambda}{\zeta_\omega}}\right) \tag{4.51}$$

This derivation is equally valid for all planar boundary surface orientations which are arranged on-axis with the Cartesian coordinate system.

The imposition of the boundary condition for a vertical edge implies two ghost nodes that will lie outside the domain on two Cartesian axes. For instance, a vertical edge will produce one ghost node on the x-axis and another on the y-axis. Each ghost node is solved for independently by means of the discretised LRS expression and inserted into the update equation for the node adjacent to both boundary surfaces. Considering the case of a vertical edge where ghost nodes are present in the positive x- and y-directions, the update expression for the interior node adjacent to both boundaries becomes:

$$p^{n+1}_{l,m,q} = \Big[\lambda^2(2p^n_{l-1,m,q} + 2p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q-1}) + 2(1-3\lambda^2)p^n_{l,m,q} - \big(1 - \frac{\lambda}{\zeta_{\omega_x}} - \frac{\lambda}{\zeta_{\omega_y}}\big)p^{n-1}_{l,m,q}\Big]\left(\frac{1}{1+\frac{\lambda}{\zeta_{\omega_x}}+\frac{\lambda}{\zeta_{\omega_y}}}\right) \tag{4.52}$$

where ζ_{ω_x} and ζ_{ω_y} are the boundary impedances of the surface normal to the x- and y-directions respectively.

Lastly, the case of a corner introduces three axially aligned ghost nodes positioned normal to three joining surfaces with impedances ζ_{ω_x}, ζ_{ω_y} and ζ_{ω_z}. The update expression for the interior corner node adjacent to all three surfaces results from solving for the ghost nodes in a manner similar to that for the case of an edge, yielding:

$$p^{n+1}_{l,m,q} = \Big[\lambda^2(2p^n_{l-1,m,q} + 2p^n_{l,m-1,q} + 2p^n_{l,m,q-1}) + 2(1-3\lambda^2)p^n_{l,m,q} - \big(1 - \frac{\lambda}{\zeta_{\omega_x}} - \frac{\lambda}{\zeta_{\omega_y}} - \frac{\lambda}{\zeta_{\omega_z}}\big)p^{n-1}_{l,m,q}\Big]\left(\frac{1}{1+\frac{\lambda}{\zeta_{\omega_x}}+\frac{\lambda}{\zeta_{\omega_y}}+\frac{\lambda}{\zeta_{\omega_z}}}\right) \tag{4.53}$$

Note that, as with surface boundary update schemes, the derivation of edge and corner interior node update expressions is equally valid for all orientations that are aligned with the SRL lattice. Furthermore, similar expressions may be derived for the case of 1D and 2D boundaries by simply following the same procedure for the appropriate wave equation.

4.4.2.2 Velocity-Centered LRS

An alternative 'velocity-centered' means of applying the LRS boundary condition in SRL FDTD schemes is documented in [130] and subsequently [151]. The derivation outlined in [130] begins by applying centered-difference approximations to the time derivative of (4.45). Backward/forward-difference approximations are applied to the spatial derivative depending on the position of the bounding surface relative to the interior domain. For example, a right-hand boundary makes use of a backwards difference and a left-hand boundary uses a forward difference. For the case of a right hand boundary in 3D, (4.45) is approximated as:

$$\frac{p^{n+1}_{l,m,q} - p^{n-1}_{l,m,q}}{2T} = -c\zeta_\omega \frac{p^{n}_{l+1,m,q} - p^{n}_{l,m,q}}{h} \tag{4.54}$$

where $p^n_{l+1,m,q}$ is the unknown ghost node value which is solved for using:

$$p^n_{l+1,m,q} = \frac{-1}{2\zeta_\omega \lambda}(p^{n+1}_{l,m,q} - p^{n-1}_{l,m,q}) + p^n_{l,m,q} \tag{4.55}$$

The above expression highlights an important difference between the pressure-centered and velocity-centered LRS implementations: the impedance term is doubled due to the combination of centered and forward finite differences applied in the derivation. This feature must be included in boundary update schemes in order to approximate the correct rate of absorption. Further investigation of the absorption qualities of each implementation is documented in section 4.4.3.

Using the velocity-centered approximation, the update expressions required for every node type (air, face (one ghost node), edge (two ghost nodes) and corner (three ghost nodes)) and orientation may be condensed into a single generic form, after [130, 151]:

$$p^{n+1}_{l,m,q} = \left[\lambda^2 S^n_{l,m,q} + (2 - K\lambda^2)p^n_{l,m,q} - (1 - B_k)p^{n-1}_{l,m,q}\right]\left(\frac{1}{1 + B_k}\right) \tag{4.56}$$

$$B_k = \frac{(6 - K)\lambda}{2\zeta_\omega} \tag{4.57}$$

with the following conditions:

- The ghost nodes that appear in the term $S^n_{l,m,q} = p^n_{l+1,m,q} + p^n_{l-1,m,q} + p^n_{l,m+1,q} + p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q-1}$ for a given boundary type are assumed to have a pressure value of zero.

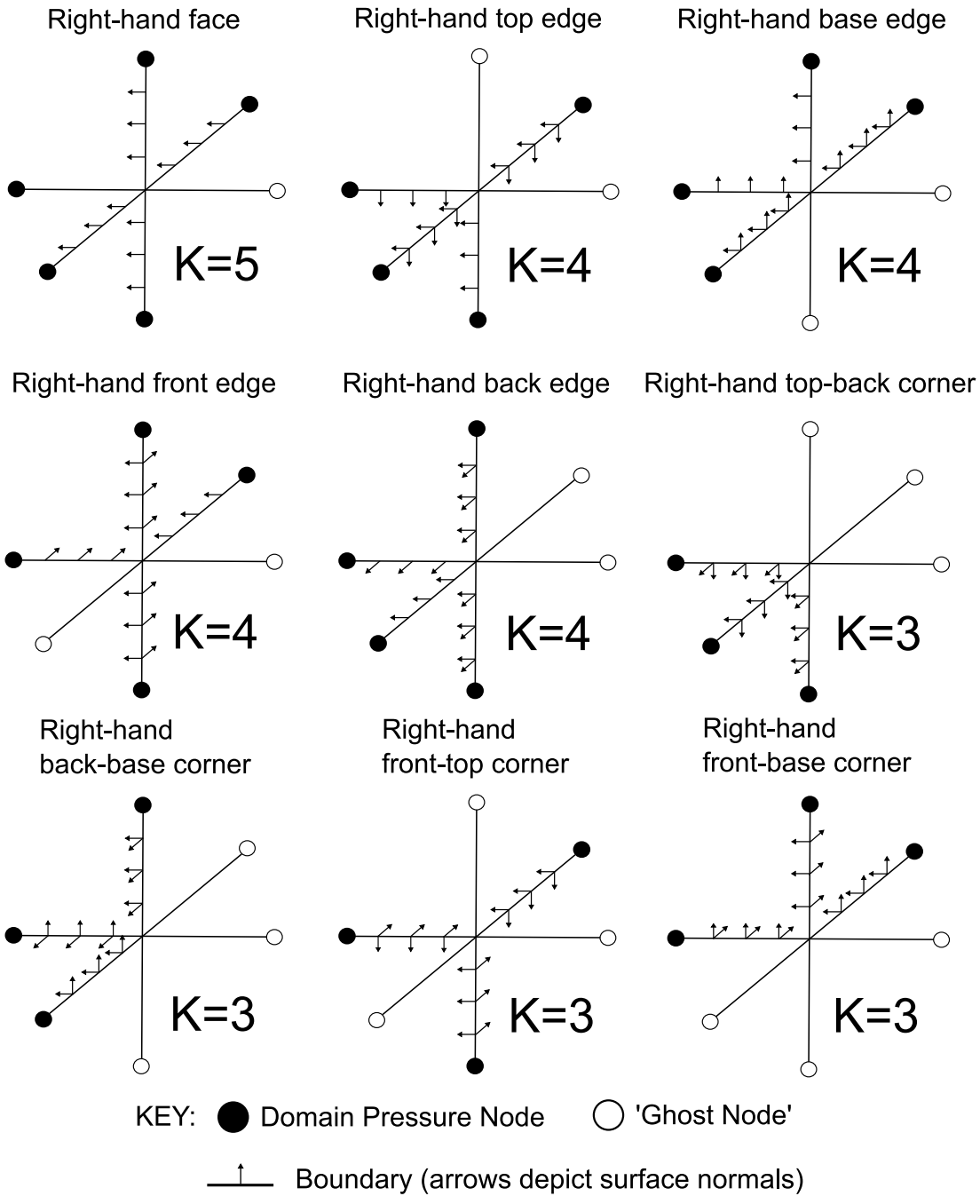

FIGURE 4.10: Example domain and 'ghost' node arrangements for all right-hand boundary types with associated K values shown for each case. Such arrangements may also be defined for left-hand boundaries.

- The constant K changes depending on the node type: $K = 6$ for interior (air); $K = 5$ for a face boundary; $K = 4$ for an edge boundary and $K = 3$ for a corner boundary.

Figure 4.10 demonstrates example node arrangements and K values for right-hand boundary types. Such a scheme facilitates an efficient means of implementing SRL/LRS schemes using accelerated parallelised algorithms on graphics processing units (GPUs). Further details of this implementation are available in [130]. As with the pressure-centered difference LRS conditions, velocity-centered LRS may also be implemented in 1D and 2D following the method of derivation detailed here for 3D.

4.4.3 Investigation of Absorbing Properties

Although pressure-centered and velocity-centered implementations of LRS boundary conditions are derived from the same governing equation, the contrasting difference approximations applied to construct the boundary node update schemes leads to differences in resulting absorption characteristics. These discrepancies are examined here using two 1D FDTD schemes, one of which incorporates pressure-centered difference LRS conditions while the other incorporates the velocity-centered approximation. The use of 1D schemes ensures that error due to numerical dispersion is not present in results.

The effect of each LRS boundary type may be observed by simply examining reflected waveforms exiting the terminations of the 1D schemes. Examples of reflected waveforms are presented in Figure 4.11 for the pressure-centered (top) and velocity-centered (bottom) boundary cases with a reflection coefficient, $R = 0.2$. The excitation function used to initialise the pressure field takes the form of a spatially distributed Hanning function of arbitrary amplitude 2.0. In accordance with the Green's function for the 1D wave equation, this excitation produces left-going and right-going travelling waves with unity amplitude. Hence, upon reflection, the travelling wave components are expected to have a peak amplitude of 0.2 as per the reflection coefficient applied. This is indeed the case for the pressure-centered LRS boundary. However, the velocity-centered condition does not reduce the amplitude of the ingoing wave as expected. Instead, the reflected travelling wave has a peak amplitude of 0.231 demonstrating that the velocity-centered LRS boundary absorbs less energy than the pressure-centered boundary which is shown to be in agreement with theory. Empirically, it may also be shown that the peak amplitude error introduced by the velocity-centered boundary increases as the reflection coefficient, R, decreases, whereas the pressure-centered boundary behaves as expected for all values of R in 1D systems.

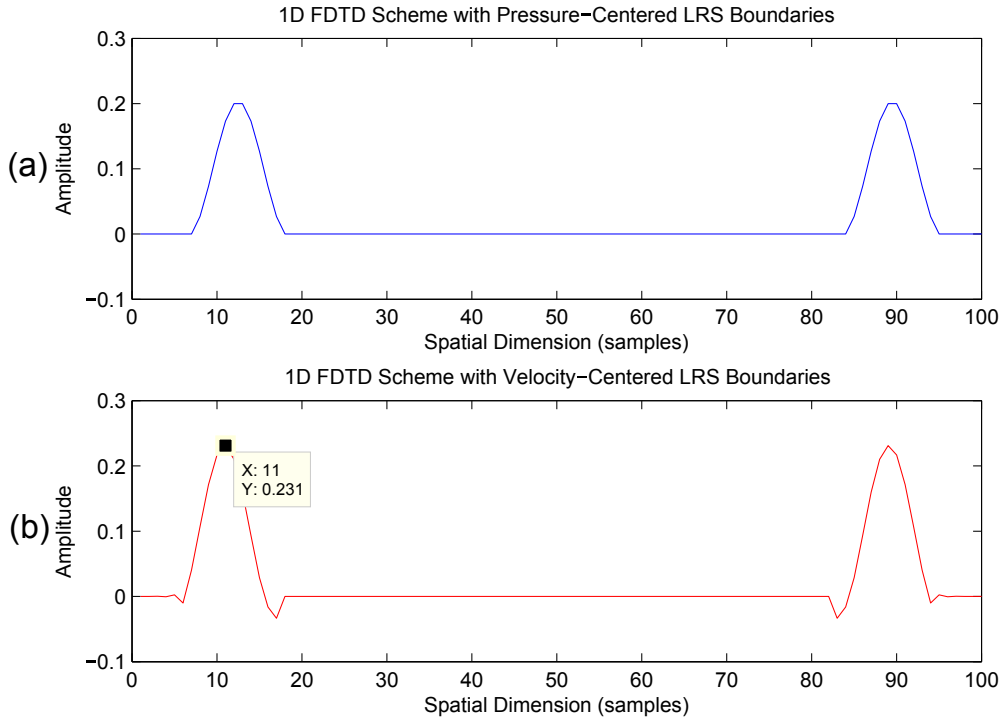

FIGURE 4.11: (a) Scalar pressure field of a 1D FDTD scheme after a single reflection of left and right travelling wave components with pressure-centered LRS boundary conditions applied to both terminations. (b) The same 1D FDTD scheme with velocity-centered LRS boundary conditions.

Another interesting feature of the velocity-centered LRS boundary is that it acts to deform the ingoing waveform producing a reflection with phase-inverted components preceding and following the main lobe. This suggests that the reflection process imposed on the ingoing wave by the velocity-centered boundary update scheme is different from that of the pressure-centered case leading to a spatial and temporal smearing of the original excitation waveform. This finding corresponds to the results documented in [151] where velocity-centered LRS conditions are demonstrated to exhibit less attenuation at high frequencies than pressure-centered conditions upon reflection. Examining the results from the 1D schemes in the frequency domain, as per Figure 4.12, it is clear that boundary absorption using the velocity-centered LRS case is indeed frequency-dependent. The discrepancy between the absorptive characteristics of pressure-centered and velocity-centered boundary formulations increases as frequency increases.

Further examination of absorption properties may be conducted by investigating the energy decay times for the two 1D schemes that result for a range of reflection coefficient values. The difference in RT_{60} values calculated by subtracting the pressure-centered scheme energy decay times from that of the velocity-centered scheme for $R = [0{:}1]$

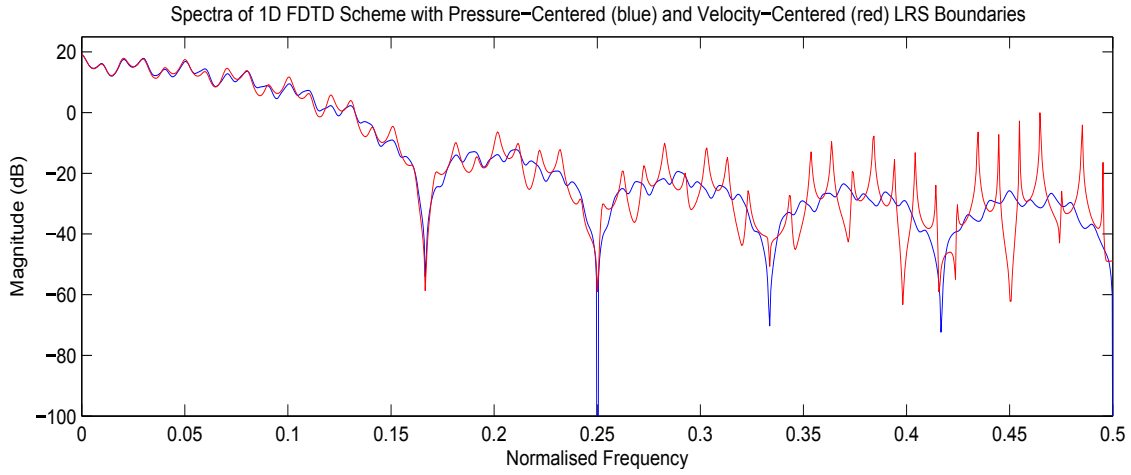

FIGURE 4.12: Comparison of magnitude spectra of impulse responses resulting from 1D finite difference schemes with pressure-centered (blue) and velocity-centered (red) LRS boundary conditions. The difference between the absorption characteristics of velocity-centered conditions compared to pressure-centered conditions becomes increasingly apparent as frequency increases.

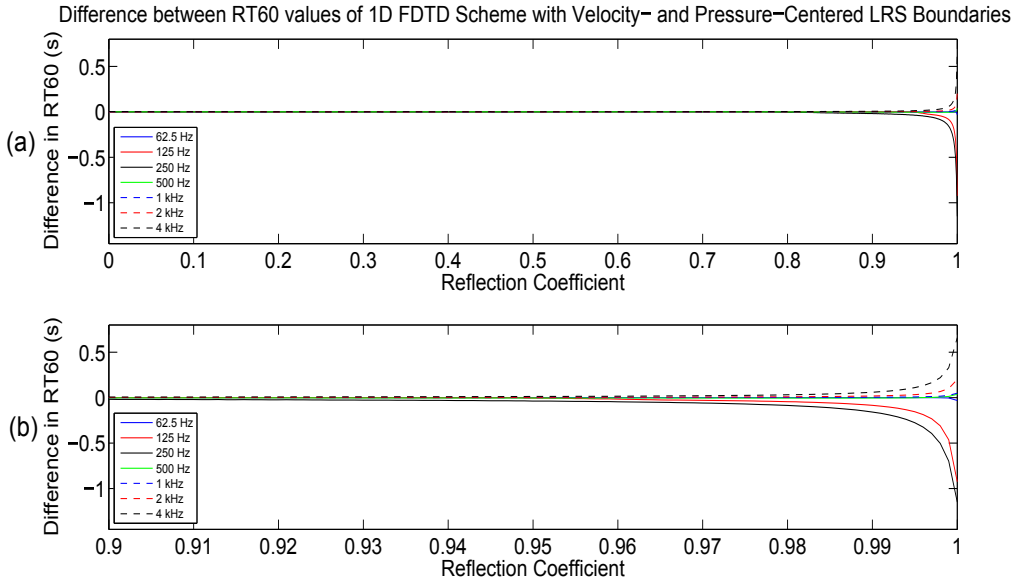

FIGURE 4.13: (a) Difference between RT_{60} values (s) for a 1D finite difference scheme using pressure-centered and velocity-centered, phase preserving boundary conditions ($0 \leq R \leq 1$). (b) Close up of RT_{60} differences given in (a) for ($0.9 \leq R \leq 1$. Values are given in octave bands up to, and including, the 4 kHz band.

are displayed in Figure 4.13 for a range of frequency octave bands. This analysis shows that while the peak amplitude discrepancy between reflected waves from each boundary implementation type increases as R decreases (as previously discussed), this does not directly translate to energy decay time results that share the same trend. Instead, the difference between RT_{60} times increases significantly for large reflection values, $R > 0.9$. Additionally, it is apparent that the deviation of energy decay times produced by the velocity-centered boundary is frequency-dependent as expected due to the high-pass filtering characteristics of the boundary. RT_{60} times in the 125 and 250 Hz octave bands are significantly less in the case of velocity-centered boundaries yielding the negative time differences obtained for high reflection values. Conversely, due to the inherent reduction in absorption at higher frequencies, the difference in RT_{60} times recorded for frequencies greater than 500 Hz increases as the reflection coefficient increases. Frequencies above the 4 kHz octave band were not considered as this higher frequency range is rarely modelled using FDTD approaches due to excessive computational requirements.

The purpose of this brief investigation is to empirically demonstrate that velocity-centered and pressure-centered LRS boundary conditions cannot be assumed to be equivalent in terms of absorptive characteristics. To the author's knowledge, the impact of this discrepancy on FDTD simulation results has not be fully examined in related literature. From the results presented here, it is possible to conclude that velocity-centered LRS implementations serve as a reasonable approximation to pressure-centered formulations for reflection coefficients $R < 0.98$. However, this conclusion is based on the assumption that the LRS boundary behaviour exhibited in 1D also arises in 2D and 3D systems. The necessity for fully quantifying and remedying the differences between the two boundary formulations stems from instability issues inherent to the pressure-centered boundary implementation as discussed in the next section.

4.4.4 Numerical Stability Issues

The pressure-centered LRS boundary conditions can be proven to preserve numerical stability for all reflection coefficient values and a given Courant limit, as per [110]. However, the analysis documented is restricted to each boundary type in isolation (i.e. only considering the boundary node and its neighbouring pressure nodes). Additionally, the stability of re-entrant nodes (those which occur when boundary surfaces connect to produce a concave structure) is examined in exactly the same way as interior/air nodes and, hence, are shown to be stable for simulations running below the Courant limit. Unfortunately, for complex room topologies which incorporate multiple re-entrant boundary node types, it may be demonstrated that instabilities can

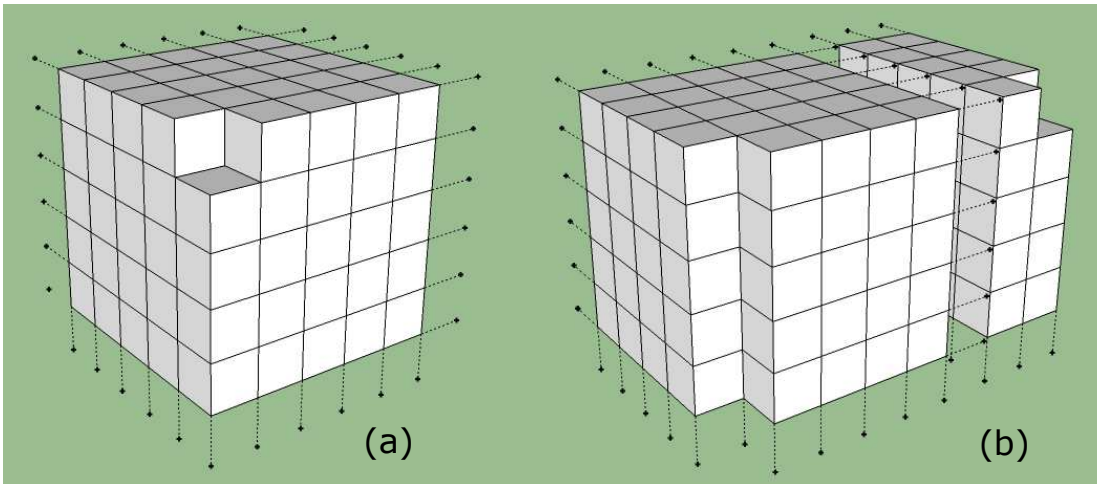

FIGURE 4.14: (a) 3D depiction of the lattice that incorporates a single re-entrant corner where vertices and lines represent pressure nodes and inter-nodal spacings respectively. (b) The lattice that incorporates both a re-entrant corner and a re-entrant edge. Dashed lines are used to represent continuation of the lattice in each dimension - these parts of the models have been excluded from the diagrams for brevity.

occur regardless of whether the stability criteria are met [151]. For the purposes of this study, it is necessary to simulate complex 3D room geometries such that resulting RIRs might be used for the purposes of comparison. Therefore, an empirical investigation into the stability of both pressure-centered and velocity-centered LRS boundary implementations is conducted to ascertain whether the simulation of 3D environments incorporating complicated boundary structures can be carried out.

Two simple 3D geometries are developed for this experiment. The first geometry is a 8 x 8 x 8 lattice of nodes with one re-entrant corner node and LRS conditions applied to face, edge and corner boundaries. The second geometry is identical, but with the addition of one vertical re-entrant edge. Examples of these two topologies are shown in Figure 4.14. Each geometry is simulated using a 3D SRL FDTD scheme running at the Courant limit with pressure-centered and velocity-centered LRS boundary implementations. The sampling rate applied is 44.1 kHz giving an inter-nodal distance of $h = \sqrt{3}ck = 0.0135$ m for wave speed $c = 344$ ms^{-1}. LRS boundaries were implemented with a reflection coefficient of $R = 0.998$, corresponding to an absorption coefficient of $\alpha = 0.004$. All scheme calculations are conducted using double precision arithmetic.

The RIRs obtained using each boundary implementation for the geometry with a single re-entrant node are displayed in Figure 4.15. As shown, both boundary implementations operate as expected preserving numerical stability to produce a valid RIR. However, the combination of both a re-entrant corner and a re-entrant edge (see Figure

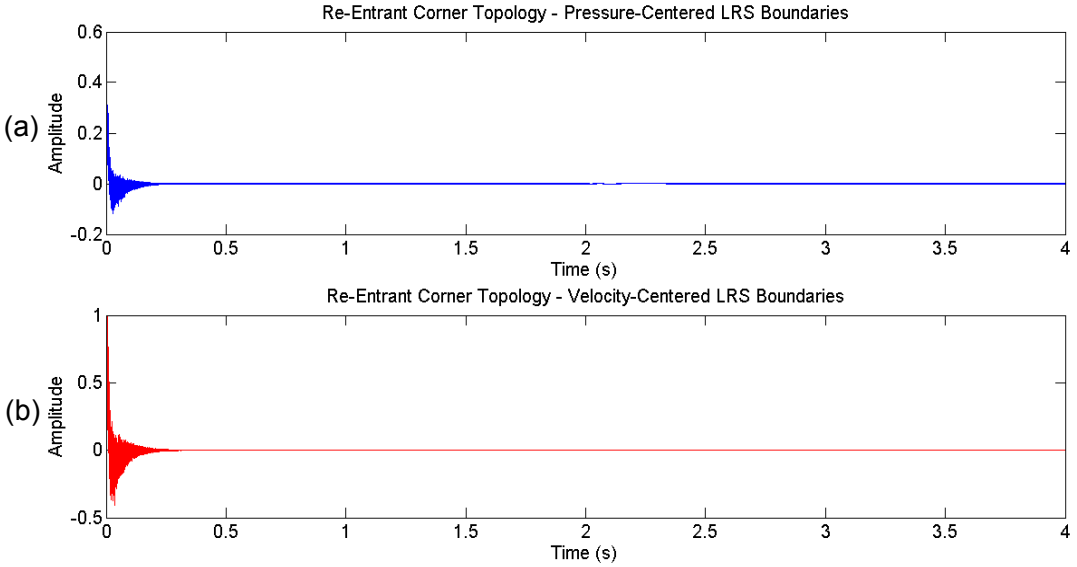

FIGURE 4.15: (a) Output RIR from a 3D FDTD SRL simulation with a single re-entrant corner node terminated with pressure-centered LRS boundary conditions. (b) Output RIR from the same scheme topology with velocity-centered LRS boundary conditions.

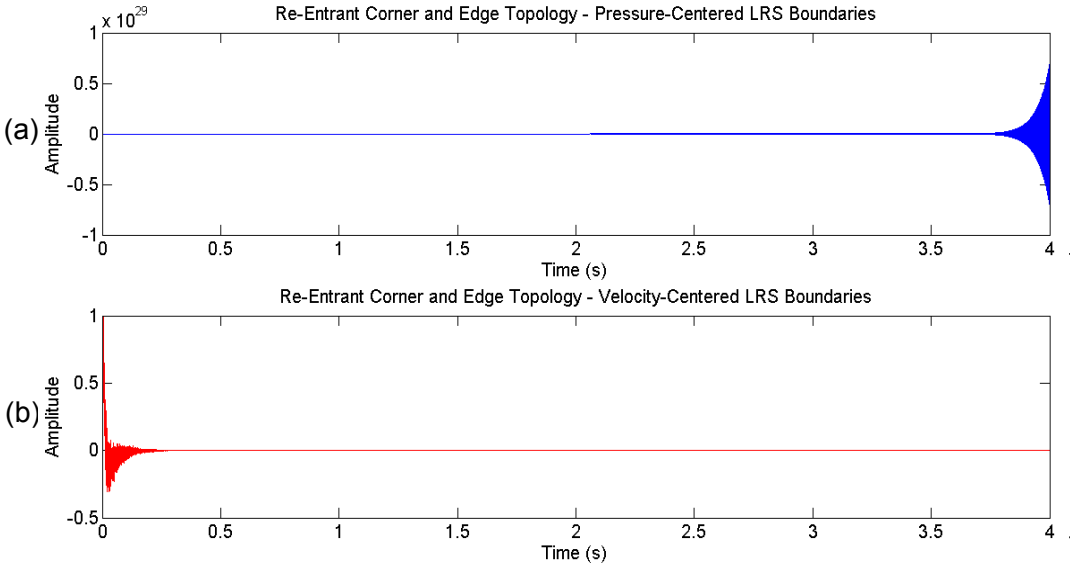

FIGURE 4.16: (a) Output RIR from a 3D FDTD SRL simulation with a re-entrant corner and edge nodes terminated with pressure-centered LRS boundary conditions. (b) Output RIR from the same scheme topology with velocity-centered LRS boundary conditions. The pressure-centered LRS scheme is shown to become numerically unstable over time.

4.14 (b)) causes the scheme with pressure-centered LRS boundaries to become unstable. Conversely, the velocity-centered boundaries do not lead to instability during simulation highlighting the finding that velocity-centered LRS implementations are more robust against instability issues. This claim is also supported by conclusions drawn in [151] which postulate that SRL schemes with velocity-centered LRS boundaries do not enable amplification of unstable resonances, particularly DC modes which occur for systems with phase-preserving terminations.

The combination of findings from the previous section and those presented above poses a challenging issue for SRL room acoustic models. In order to emulate the absorbing qualities of boundary surfaces to a sufficient level of accuracy, the pressure-centered LRS approximation is optimal. However, for complex 3D topologies the pressure-centered implementation is shown to be conditionally stable while the velocity-centered implementation demonstrates robust numerical stability. This finding is further supported by the results documented in section 5.3 in which a 3D scheme involving several re-entrant pressure node configurations is successfully simulated with preservation of numerical stability, producing usable RIR outputs. Hence, equating the absorbing properties of pressure-centered and velocity-centered LRS boundaries may prove a useful route of further investigation in order to render stable and acoustically accurate SRL finite difference room acoustic models.

4.5 Frequency-Dependent Boundary Conditions

The absorption characteristics of surface materials in real spaces are, in general, frequency-dependent as discussed in section 2.3.8. As per the convention set out in the majority of related literature (see e.g. [17, 26]), the absorption coefficient for a given material may be defined in octave bands (for example Table 4.2). This, in turn, facilitates increased virtual acoustic simulation accuracy as a more complete model of surface material properties is preserved comparative to frequency-independent boundary modelling. The work presented in this section closely follows a case study [152] in which the development and comparison of distinct approaches to frequency-dependent LRS boundary implementation were investigated. This case study was conducted to inform and support the approach to rendering RIRs for the final acoustic modelling scenario documented in section 5.3.

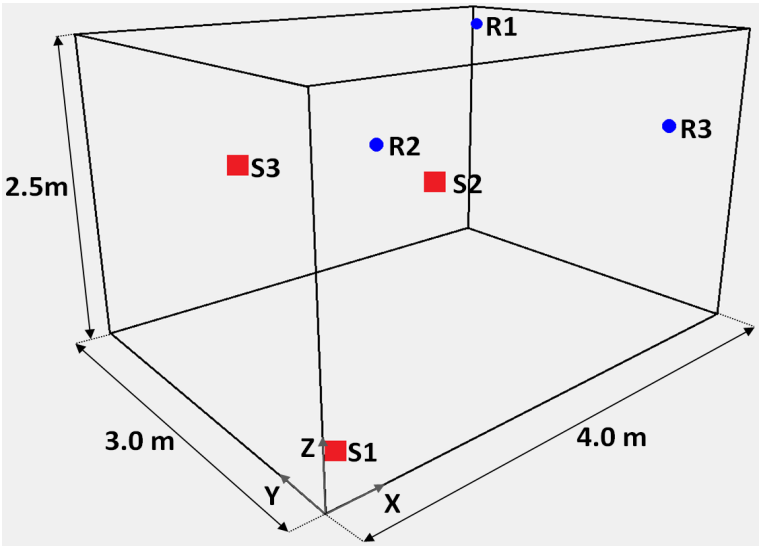

FIGURE 4.17: Overview of the room geometry used for the purposes of the case study into frequency-dependent boundary conditions. Dimensions and sound source (S1 - S3) and receiver (R1 - R3) locations are displayed.

4.5.1 Case Study Acoustic Model

The acoustic environment used to investigate frequency-dependent boundary conditions consisted of a simple, fully enclosed, cuboid topology. The selection of this room geometry allows straightforward modal analysis whereby the resonant frequencies arising in obtained RIRs can be compared with those given by known eigenvalue solutions to the wave equation in a rectilinear domain. As shown in Figure 4.17, the room geometry has dimensions 4 x 3 x 2.5 m corresponding to length, width and height respectively. Three sound source and receiver pairs are located throughout the space with positions as listed in Table 4.1. The RIRs rendered using source/receiver placements 1-3 are referred to as cases 1-3 respectively in the following.

CASE	Source (x, y, z) (m)	Receiver (x, y, z) (m)
1	(0.21, 0.21, 0.21)	(3.83, 2.77, 2.34)
2	(2.13, 1.49, 1.28)	(2.55, 2.77, 1.28)
3	(0.42, 1.50, 1.70)	(3.70, 0.42, 1.70)

TABLE 4.1: Overview of source and receiver placements defined for each RIR measurement.

Surface materials with frequency-dependent absorption coefficients, defined in octave bands, are prescribed in two different arrangements. The first arrangement consists of plasterboard material absorption characteristics applied to all surfaces. The second arrangement consists of a wooden floor, rough concrete walls and a plasterboard ceiling.

Material	f_l:Center Frequency:f_u(Hz)	α	R
Plaster (Ceiling)	44 : 63 : 88	0.08	0.9592
	88 : 125 : 177	0.08	0.9592
	177 : 250 : 355	0.2	0.8944
	355 : 500 : 710	0.5	0.7071
	710 : 1k : 1.42k	0.4	0.7746
	1.42k : 2k : 2.84k	0.4	0.7746
	2.84k : 4k : 5.68k	0.36	0.8000
Wood (Floor)	44 : 63 : 88	0.15	0.9220
	88 : 125 : 177	0.15	0.9220
	177 : 250 : 355	0.11	0.9434
	355 : 500 : 710	0.1	0.9487
	710 : 1k : 1.42k	0.07	0.9644
	1.42k : 2k : 2.84k	0.06	0.9695
	2.84k : 4k : 5.68k	0.06	0.9695
Concrete (Walls)	44 : 63 : 88	0.02	0.9899
	88 : 125 : 177	0.02	0.9899
	177 : 250 : 355	0.03	0.9849
	355 : 500 : 710	0.03	0.9849
	710 : 1k : 1.42k	0.03	0.9849
	1.42k : 2k : 2.84k	0.04	0.9798
	2.84k : 4k : 5.68k	0.07	0.9644

TABLE 4.2: Summary of absorption (α) and reflection (R) coefficients for the materials applied to model surfaces for each octave band simulated.

Absorption data applied in each model material arrangement is available from the ODEON Auditorium 10.1 software [21] and is listed in Table 4.2 alongside the frequency ranges defined for each octave band simulated.

For the purposes of this investigation, the frequency bandwidth of output RIRs is selected as the upper limit of the 4 kHz octave band (5.68 kHz) in the interests of maintaining reasonable simulation run-times and computational requirements. In order to avoid the manifestation of spatial aliasing in the required frequency range, the temporal sampling rate must follow the criterion $f_u \leq 0.196\ F_s$ where f_u is the highest required simulation frequency [123]. This criterion is derived from the dispersion relation (4.25) for the worst-case dispersion error arising in axial directions of propagation with a particular F_s. For example, for axial propagation in the x-direction, the largest valid wavenumber in x is given as $k_x h = \pi$ in accordance with the Nyquist condition. Also, it is noted that the largest wavenumber must coincide with the upper frequency of simulation, f_u. Therefore, using (4.25) for axial dispersion in the x-direction, the

following relation is derived:

$$2\pi f_u = \frac{2}{T}\sin^{-1}\left(\lambda\sqrt{\sin^2\left(\frac{\pi}{2}\right)}\right) \tag{4.58}$$

which, for a 3D SRL scheme running at the Courant limit, yields:

$$f_u = F_s\left(\frac{\sin^{-1}\left(\frac{1}{\sqrt{3}}\right)}{\pi}\right) = 0.196F_s \tag{4.59}$$

Hence, for the purposes of simulating the cuboid model up to $f_u = 5.68$ kHz, a temporal sampling rate of $F_s = 5f_u = 28.4$ kHz is selected. For a simulation wave speed of 344 ms^{-1}, the grid spacing h is set as per the Courant limit for numerical stability and minimum dispersion, giving $h = \sqrt{3}cT = 0.021$ m.

4.5.2 Octave Band Approaches

A simple approach to emulating frequency-dependent boundary conditions in FDTD simulations is presented in [153]. This method involves simulating the target acoustic field once per octave band with the associated absorption coefficients applied in each instance. The resulting full bandwidth RIRs are then filtered as appropriate to isolate the octave band for which the absorption coefficients are correctly defined. Finally, each rendered octave band RIR is then summed to create a spectrally complete RIR with frequency-dependent absorption characteristics. Following on from the work presented in [153], two different approaches to this problem are investigated here.

4.5.2.1 Constant F_s Implementation

As previously discussed, the temporal sampling rate selected for the FDTD simulations is $F_s = 28.4$ kHz. For the case of a *constant* F_s implementation, the room topology is discretised to form a SRL lattice of grid points with an inter-nodal distance of $h = 0.021$ m. This acoustic model is then simulated for each of the 7 octave bands using the SRL update scheme and LRS boundary conditions with respective absorption coefficients applied in each simulation. RIRs are generated in this way for each of the 3 source/receiver pairs using a soft source Kronecker delta excitation and an omnidirectional receiver implemented by recording the pressure fluctuations at the receiver location.

The resulting RIRs are then grouped in terms of source/receiver case and filtered by a means of a 3^{rd}-order IIR bandpass filter bank with Butterworth characteristics.

This enables the creation of a total RIR for each case by summing each octave band contribution recovered from the filter bank. Finally, a DC blocking filter is applied to removed the DC offset drift caused by the method of pressure field excitation. This process is repeated for both surface material arrangements noted in section 4.5.1.

4.5.2.2 Variable F_s Implementation

In contrast to the constant F_s approach, the *variable* F_s implementation models the topology for each octave band with the temporal sampling rate set to a value sufficiently high such that a level of dimensional consistency is maintained and spatial aliasing is avoided. It is necessary to impose a lower limit on F_s in order to realise the aforementioned geometric consistency between all octave band models. To this end, the lowest three octave bands (63, 125 and 250 Hz) simulations were defined with a sampling rate of $F_s = 2.8$ kHz in order to maintain a maximum absolute error in room geometry dimensions of $< 5\%$ when compared to the desired geometry of 4 x 3 x 2.5 m. Table 4.3 displays the sampling rates and corresponding maximum error values in all dimensions. Additionally, an error in the positioning of source and receiver locations also arises as these must be located at a pressure node in the SRL lattice. The maximum deviation from the ideal values stated in Table 4.1 is 0.11 m and arises for RIR case 2 in the 500 Hz octave band model. The different sampling rates applied across

Octave Band (Hz)	F_s (Hz)	Max. Dimension Error (%)
63, 125, 250	2800	2.14
500	3700	3.06
1000	7100	0.70
2000	14200	0.70
4000	28400	0.18

TABLE 4.3: Sampling rates applied to each octave band model for the variable F_s approach. The maximum errors in model dimensions are also provided.

all 7 octave band simulations gives rise to a large difference in the number of pressure nodes required to represent the cuboid model domain in each instance. As such, the amount of energy applied during excitation must be compensated for as the nature of pressure distribution over the propagating wavefront is inconsistent between adjacent octave bands. This topic is the subject of [154] in which a procedure to equalise the energy of each contributing octave band is formulated. Following [154], each rendered RIR is multiplied by a weighting coefficient, K_{SX}, defined as:

$$K_{SX} = \frac{\lambda F_s}{p_{DC} c} \tag{4.60}$$

where P_{DC} is the DC component of the RIR to which the weighting is being applied and λ is the Courant number associated with the FDTD scheme used. Furthermore, the resulting weighted RIRs for the 63 - 2000 Hz octave bands must be upsampled to the sampling rate of the highest octave band for consistency. Upon completion of these procedures, the variable F_s RIRs are then subject to the filtering processes applied in the constant F_s implementation to render the total RIRs.

4.5.3 LRS Boundaries using Digital Impedance Filters (DIFs)

A notable disadvantage of the octave band based approach to frequency-dependent boundary modelling is the requirement for multiple simulations of the target acoustic field. As initially documented in [124], an alternative approach to achieving frequency-dependent absorption characteristics in LRS boundary conditions is possible through use of digital filtering techniques. To begin, following the discussion in [110], frequency-dependent reflection coefficients, R_ω, are defined in terms of a discrete-time filter transfer function:

$$R_\omega(\mathrm{z}) = \frac{\psi_0 + \sum_{i=1}^{N} \psi_i \mathrm{z}^i}{\kappa_0 + \sum_{i=1}^{N} \kappa_i \mathrm{z}^i} \tag{4.61}$$

for the case of an N^{th}-order infinite impulse response (IIR) digital filter defined by a z-transformed transfer function by polynomials of z with coefficients ψ_i and κ_i; i = [0:N]. The magnitude response of a given filter is designed such that the required reflection coefficient values are defined over each octave band. This, in turn, leads to the definition of a digital impedance filter with a transfer function, $\zeta_\omega(\mathrm{z})$, given by:

$$\zeta_\omega(\mathrm{z}) = \frac{1 + R_\omega(\mathrm{z})}{1 - R_\omega(\mathrm{z})} \tag{4.62}$$

yielding,

$$\zeta_\omega(\mathrm{z}) = \frac{b_0 + B(\mathrm{z})}{a_0 + A(\mathrm{z})} \tag{4.63}$$

where,

$$b_0 = \psi_0 + \kappa_0 \tag{4.64}$$

$$B(\mathrm{z}) = \sum_{i=1}^{N} (\psi_i + \kappa_i)\mathrm{z}^i; \qquad b_i = \psi_i + \kappa_i \tag{4.65}$$

$$a_0 = \kappa_0 - \psi_0 \tag{4.66}$$

$$A(\mathrm{z}) = \sum_{i=1}^{N} (\kappa_i - \psi_i)\mathrm{z}^i; \qquad a_i = \kappa_i - \psi_i \tag{4.67}$$

A full description of the incorporation of $\zeta_\omega(\mathrm{z})$ into pressure-centered LRS boundary conditions documented in section 4.4 is detailed in [110, 124]. For brevity, the

expressions central to the FDTD algorithm development are given here. Firstly, an intermediate filter term g^n is defined for each ghost node present at a given boundary at time sample n, as per [110]:

$$g^n = \sum_{i=1}^{N} (b_i x^{n-i} - a_i y^{n-i}) \tag{4.68}$$

where x^n and y_n are the input and output digital impedance filter values respectively. It is assumed that the target acoustic field is zero for time, $t \leq 0$, throughout and, therefore, both filter input and output values are also zero. Hence, the intermediate value may be defined as zero value during the first scheme iteration prior to application in the following boundary update expressions. These are constructed for face (at the limit of the spatial domain in the x-direction), edge (at the limits of the spatial domain in the x- and y-directions) and corner (at the limits of the spatial domain in all directions) LRS boundaries respectively, after [110], for a 3D domain:

$$\begin{aligned} p^{n+1}_{l,m,q} = \Bigg(& \lambda^2 (2p^n_{l-1,m,q} + p^n_{l,m+1,q} + p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q-1}) \\ & + 2(1 - 3\lambda^2) p^n_{l,m,q} + \Big(\frac{\lambda a_{0_x}}{b_{0_x}} - 1 \Big) p^{n-1}_{l,m,q} + \frac{\lambda^2}{b_{0_x}} g^n_x \Bigg) \Bigg(\frac{1}{1 + \frac{\lambda a_{0_x}}{b_{0_x}}} \Bigg) \end{aligned} \tag{4.69}$$

$$\begin{aligned} p^{n+1}_{l,m,q} = \Bigg(& \lambda^2 (2p^n_{l-1,m,q} + 2p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q-1}) + 2(1 - 3\lambda^2) p^n_{l,m,q} \\ & + \Big(\frac{\lambda a_{0_x}}{b_{0_x}} + \frac{\lambda a_{0_y}}{b_{0_y}} - 1 \Big) p^{n-1}_{l,m,q} + \frac{\lambda^2}{b_{0_x}} g^n_x + \frac{\lambda^2}{b_{0_y}} g^n_y \Bigg) \Bigg(\frac{1}{1 + \frac{\lambda a_{0_x}}{b_{0_x}} + \frac{\lambda a_{0_y}}{b_{0_y}}} \Bigg) \end{aligned} \tag{4.70}$$

$$\begin{aligned} p^{n+1}_{l,m,q} = \Bigg(& \lambda^2 (2p^n_{l-1,m,q} + 2p^n_{l,m-1,q} + 2p^n_{l,m,q-1}) + 2(1 - 3\lambda^2) p^n_{l,m,q} \\ & + \Big(\frac{\lambda a_{0_x}}{b_{0_x}} + \frac{\lambda a_{0_y}}{b_{0_y}} + \frac{\lambda a_{0_z}}{b_{0_z}} - 1 \Big) p^{n-1}_{l,m,q} + \frac{\lambda^2}{b_{0_x}} g^n_x + \frac{\lambda^2}{b_{0_y}} g^n_y \\ & + \frac{\lambda^2}{b_{0_z}} g^n_z \Bigg) \Bigg(\frac{1}{1 + \frac{\lambda a_{0_x}}{b_{0_x}} + \frac{\lambda a_{0_y}}{b_{0_y}} + \frac{\lambda a_{0_z}}{b_{0_z}}} \Bigg) \end{aligned} \tag{4.71}$$

where coefficients with subscript x, y and z refer to the DIF terms applied while solving for the boundary ghost nodes positioned beyond the domain in the corresponding axial directions. Note that, as with the case of frequency-independent LRS boundaries, appropriate update equations for other boundary orientations may be expressed in a similar fashion. Each boundary node, therefore, requires contributions from at least

one DIF with the input and output values calculated using:

$$x^n = \frac{a_0}{\lambda b_0}(p_{l,m,q}^{n+1} - p_{l,m,q}^{n-1}) - \frac{g^n}{b_0} \tag{4.72}$$

$$y^n = \frac{1}{a_0}(b_0 x^n + g^n) \tag{4.73}$$

which are defined for a boundary node at location $[lh, mh, qh]$ and a given DIF, as per [110]. Hence, the update scheme operates as follows. The intermediate value g^n for all boundary filters is initialised to zero for the first iteration, during which all boundary nodes are updated and, consequently, the filter input and output values are calculated. Upon the subsequent iteration, g^n is recalculated for all DIFs and applied in each boundary node update. The simulation progresses in this manner with air nodes being treated as they would in a frequency-independent case.

Digital Impedance Filter Design:

As suggested in [110], the required DIFs for the case study model are designed by first defining the reflection filter $R_\omega(\mathrm{z})$. This is achieved by constructing an interpolated curve through the prescribed reflection coefficient values defined at the center frequency of each octave band. The MATLAB digital filter design tool (fdesign.arbmag [155]) is then utilised to derive suitable filter coefficients which provide an approximation to the desired magnitude response. Note that phase response characteristics are not considered in this study as data corresponding to the phase response of the surface materials used is not readily available.

Figure 4.18 depicts $|R_\omega(\mathrm{z})|$ for the plasterboard surface material alongside the interpolated reflection coefficient curve and ideal octave band reflection values. This filter response curve is generated using a 12^{th}-order digital filter. As shown, the filter magnitude response gives a good approximation to the interpolated reflection curve which, in turn, exhibits a close resemblance to the octave band values. The reflection filter is then re-formulated as a DIF by means of (4.63). The resulting impedance magnitude response is given in Figure 4.19. It can be seen that small errors in the response of $R_\omega(\mathrm{z})$ are magnified when transferring from reflection values to impedance values highlighting the sensitive dependence of the DIF response on $|R_\omega(\mathrm{z})|$ when R is large. The largest error in impedance values are shown to be in the region of the 125 and 250 Hz octave bands. This discrepancy between desired and actual DIF responses impacts the results documented in section 4.5.4.

For the cases of the wood and concrete reflection filters, the same design procedure is implemented utilising 14^{th}- and 18^{th}-order filters respectively. Filters of higher order are

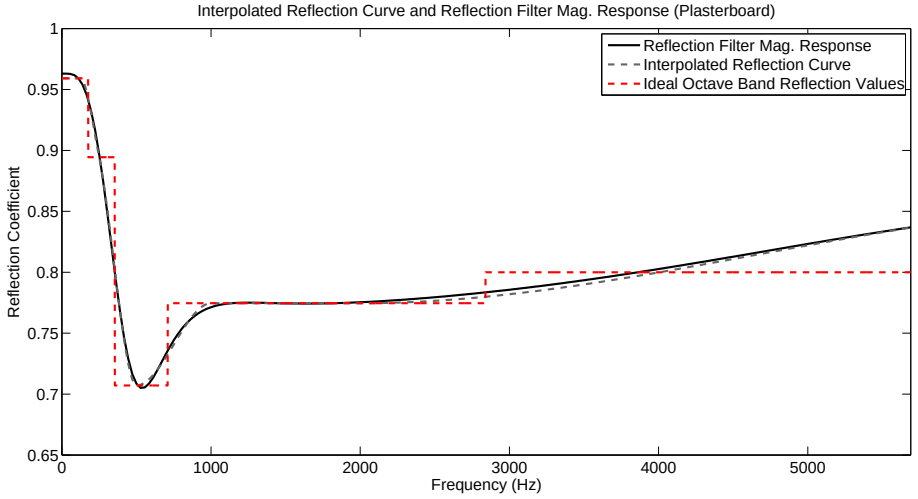

FIGURE 4.18: Reflection filter magnitude response, $|R_\omega(z)|$ (solid black line), for the case of plasterboard material depicted in terms of reflection coefficient vs. frequency. Interpolated frequency-dependent reflection curve (dashed grey line) and ideal octave band reflection values (dashed red line) are also provided.

found to produce negligible improvement in terms of response accuracy. Figures 4.20 and 4.21 show the characteristics of the wood surface reflection and impedance filter magnitude responses respectively. A high level of agreement between the interpolated reflection curve and the designed filter magnitude response is apparent. This, in turn translates to a low error between desired frequency-dependent impedance values and those resulting in the magnitude response of the corresponding DIF (as shown in Figure 4.21). The concrete reflection filter and DIF characteristics are displayed in Figures 4.22 and 4.23 respectively. Again, as with the case of the plasterboard material, the small deviations in the reflection filter magnitude response from the interpolated reflection curve cause notable errors in the DIF response (see Figure 4.23). These errors are particularly apparent for frequencies greater than 250 Hz, suggesting that a faster rate of absorption will occur for the second material arrangement in the upper 4 simulated octave bands.

As previously noted, the phase response of the DIFs was unconstrained during the computation of suitable filter coefficients. Thus, the phase response of each filter does not correspond to the physical properties of the surface materials modelled and is simply a consequence of creating the required frequency-dependent impedance characteristics. The phase response of each DIF is depicted in Figure 4.24. It is clear from this diagram that all DIFs exhibit highly non-linear phase responses in the frequency range of $[0 : 1.5]$ kHz. As such, it is expected that RIRs generated using the DIF approach will possess resonances at low frequencies that deviate from those present in RIRs simulated using the octave band implementation. This is due to the fact that the frequency-independent

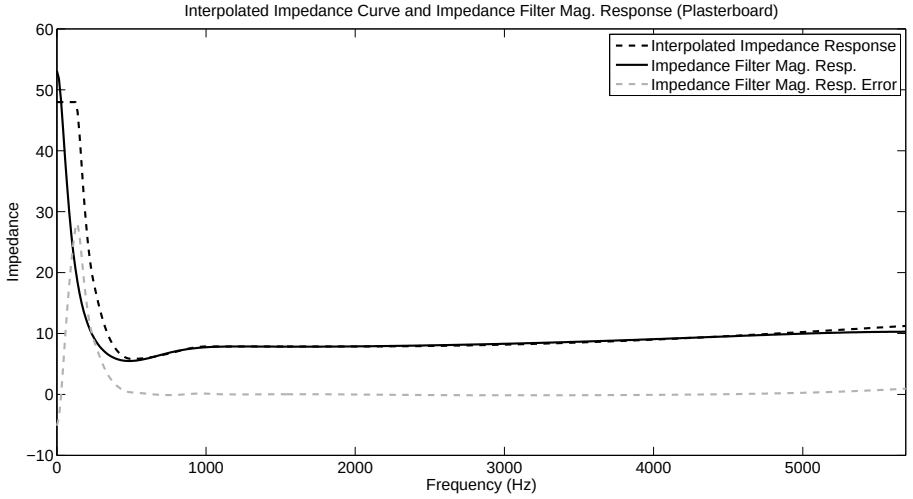

FIGURE 4.19: Magnitude response of the resulting impedance filter for the plasterboard material, $|\zeta_\omega(z)|$ (solid black line), as a function of of impedance vs. frequency. The interpolated impedance curve (dashed dark grey line) is derived from the corresponding reflection curve. The error between desired and resulting impedance values is shown by the dashed light-grey line.

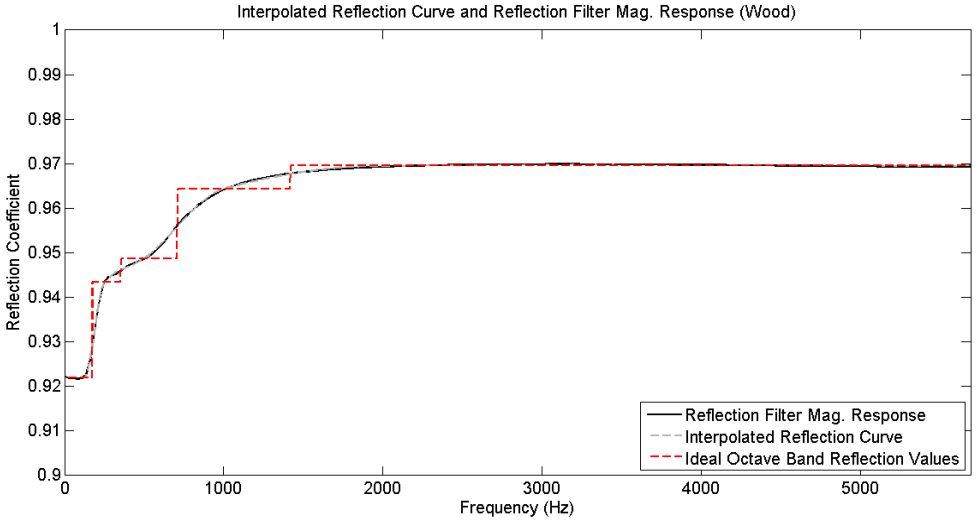

FIGURE 4.20: Reflection filter magnitude response, $|R_\omega(z)|$ (solid black line), for the case of the wood material depicted in terms of reflection coefficient vs. frequency. Interpolated frequency-dependent reflection curve (dashed grey line) and ideal octave band reflection values (dashed red line) are also provided.

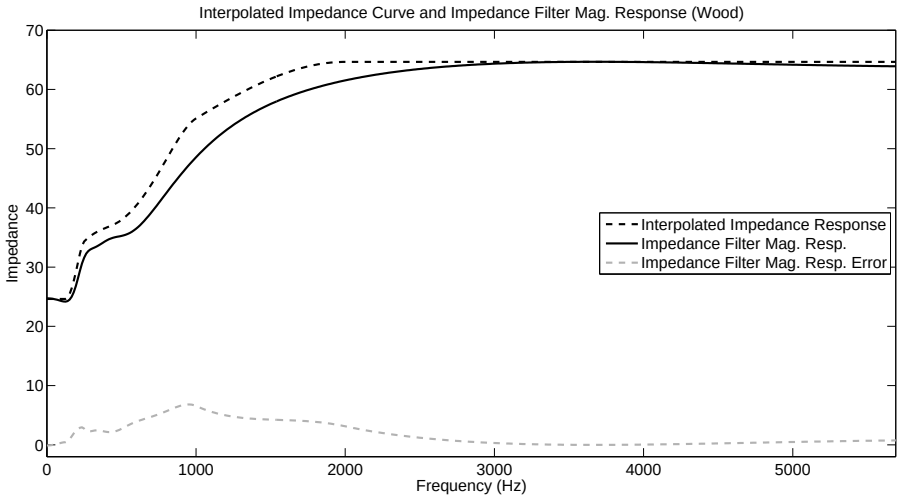

FIGURE 4.21: Magnitude response of the resulting impedance filter for the wood material, $|\zeta_\omega(z)|$ (solid black line), as a function of of impedance vs. frequency. The interpolated impedance curve (dashed dark grey line) is derived from the corresponding reflection curve. The error between desired and resulting impedance values is shown by the dashed light-grey line.

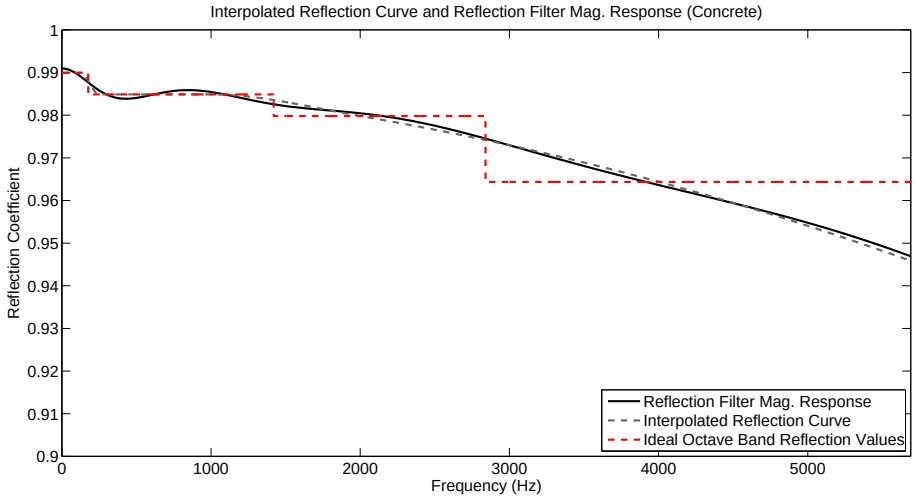

FIGURE 4.22: Reflection filter magnitude response, $|R_\omega(z)|$ (solid black line), for the case of concrete material depicted in terms of reflection coefficient vs. frequency. Interpolated frequency-dependent reflection curve (dashed grey line) and ideal octave band reflection values (dashed red line) are also provided.

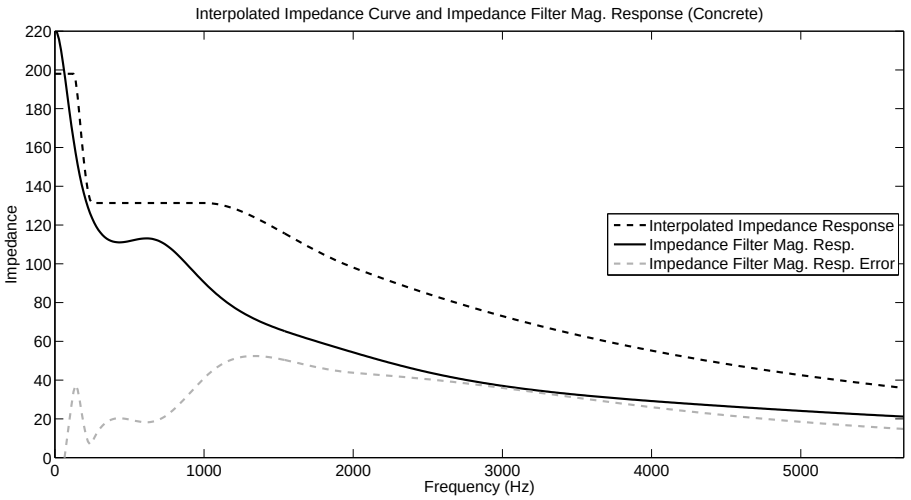

FIGURE 4.23: Magnitude response of the resulting impedance filter for the concrete material, $|\zeta_\omega(z)|$ (solid black line), as a function of of impedance vs. frequency. The interpolated impedance curve (dashed dark grey line) is derived from the corresponding reflection curve. The error between desired and resulting impedance values is shown by the dashed light-grey line.

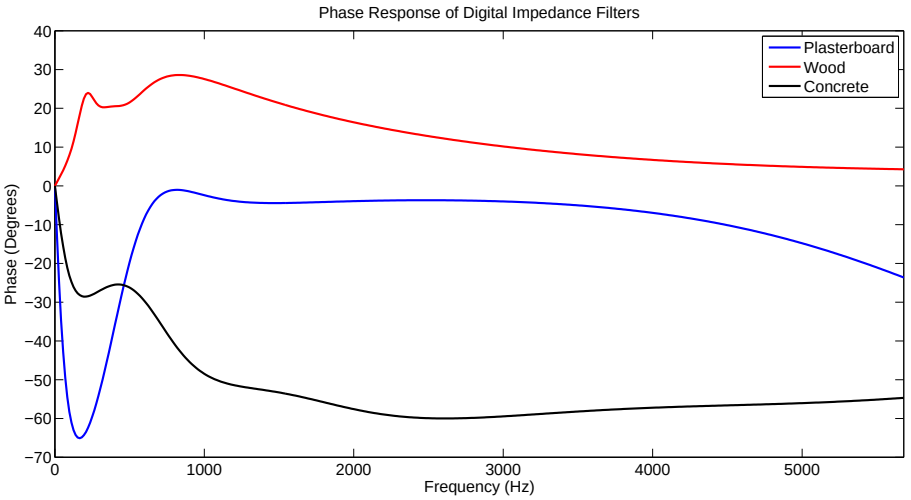

FIGURE 4.24: The phase response of the designed DIFs for each surface material type applied in the cuboid case study acoustic model, displayed over the simulation bandwidth.

LRS boundary conditions applied in the octave band approach are phase-preserving. Therefore, the process of combining each octave band contribution is analogous to applying an ideal zero-phase DIFs to boundary nodes. Given the low phase delay and advance characteristics of the plasterboard and wood DIFs respectively in the region of $[2 : 4]$ kHz, it is expected that good agreement between the nature of reflections from the octave band and DIF-based boundaries will result for the upper 2 octave bands. However, the concrete material DIF possesses significant phase delay values for the majority of the simulated bandwidth. This material is applied to the majority of the surface area present in the second surface material arrangement and, as such, deviations from resonant frequency components present in the octave band RIRs are expected across the simulated spectra.

4.5.4 Simulation Results and Discussion

A total of 6 RIRs were rendered using each of the three frequency-dependent boundary implementations. These RIRs correspond to the 3 source/receiver cases for the two model material arrangements. Findings from simulation data, presented here, impact on the design of the hybrid modelling procedure applied in the final study presented in section 5.3. Additionally, the following results provide pertinent insight into the comparison of frequency-dependent numerical acoustic modelling methods on which future studies in this area may be based (see section 4.5.4.4).

4.5.4.1 Spectral Analysis

Frequency domain analysis of the simulated RIRs facilitates the evaluation and verfication of model results in two ways. Firstly, modal analysis may be applied to compare simulated resonant frequency values to those calculated from theory. Secondly, visual inspection of relative energy levels across each modelled frequency band demonstrates the extent to which results agree with the frequency-dependent absorption characteristics defined for surface materials in each arrangement. RIR case 1 is selected for the purposes of this analysis as the positioning of source and receiver locations (in diagonally opposing corners) gives rise to the highest presence of resonant low frequency activity. Both material arrangements are considered in the following.

Material Arrangement 1:

As previously noted, material arrangement 1 incorporated the plasterboard material absorption properties applied to all surfaces. The magnitude spectra of RIRs rendered using each frequency-dependent LRS boundary implementation are displayed in Figure

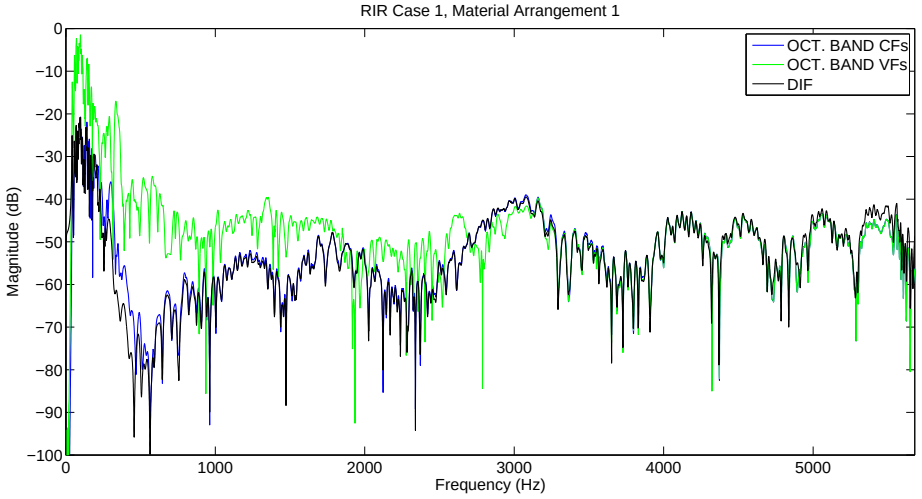

FIGURE 4.25: Magnitude spectra of constant F_s octave band (OCT. BAND CFs - blue line), variable F_s octave band (OCT. BAND VFs - green line) and DIF (black line) boundary simulation RIRs for source/receiver case and material arrangement 1.

4.25. From this diagram, it is immediately apparent that the energy equalisation process used in the creation of the variable F_s simulation RIR has not been successful. There are significant differences in energy levels for all octave bands up to the 4 kHz band which is consistent with the constant F_s RIR. While this erroneous result has not yet been fully examined, the cause may stem from the following issues:

- The sensitivity of response characteristics on source/receiver locations and model dimensions. The variable F_s approach renders contributing octave band RIRs with small errors in geometry. These errors may translate to significant changes in RIR spectral properties comparative to the more geometrically accurate constant F_s simulation results.
- In this study, the application of the energy-matching constant, (4.60), is carried out after filtering the contributing RIRs with a DC blocking filter. This causes a reduction in the value of p_{DC} in (4.60), in turn leading to an increase in the equalisation constant K_{SX}.

Both of the above points are possible areas of future research in order to investigate this comparative inaccuracy which is common to all RIRs rendered using the variable F_s approach in this study.

The level of agreement between overall spectral properties of the constant F_s and DIF simulation RIRs (as per Figure 4.25) is significantly high for the majority of the simulated bandwidth. However, there are exceptions in the lower and upper frequency

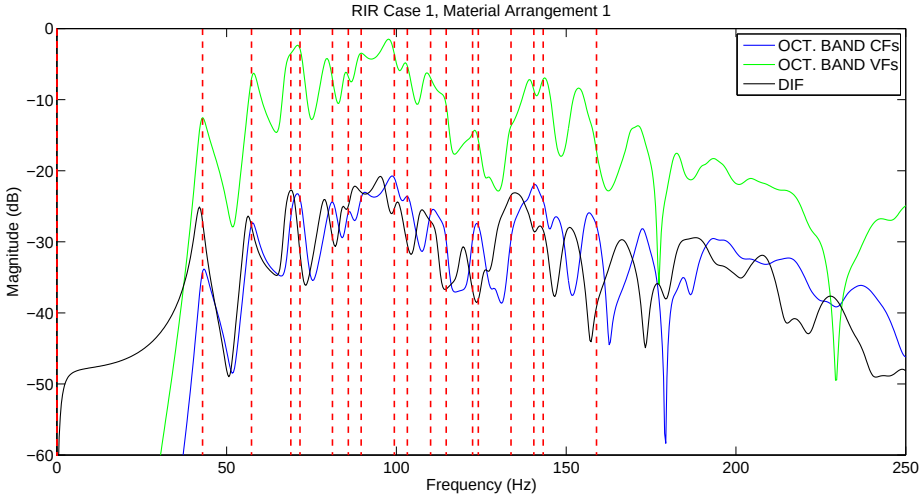

FIGURE 4.26: Low frequency magnitude spectra of constant F_s octave band (OCT. BAND CFs - blue line), variable F_s octave band (OCT. BAND VFs - green line) and DIF (black line) boundary simulations. Theoretical mode values, up to second-order, are provided by the vertical dashed red lines.

ranges. Of particular note are the overall energy levels present in the 125 and 250 Hz octave bands for these RIRs. The differences arising in these frequency ranges is a direct result of the error present in the magnitude response of the DIF developed for the plasterboard material. Compared to the octave band approach, the impedance applied in the case of the DIF boundary model is consistently less than required, leading to faster rates of absorption and less resonance of low frequency components.

Figure 4.26 provides a closer examination of the comparison between low frequency components for the Case 1 RIR as produced by each simulation method. From this diagram, it is clear that both the constant and variable F_s models share good agreement with theoretical mode values which are provided up to second-order. Slight deviations, which are more pronounced in the case of the variable F_s response, from the true modal values are due to numerical dispersion error. Hence, it may be stated that the variable F_s RIR would serve as a good approximation to the constant F_s RIR at low frequencies, provided the aforementioned error in overall energy levels was reduced. Finally, the DIF-based simulation RIR spectrum exhibits the impact of the phase response characteristics of the plasterboard material filter. The phase delay (see Figure 4.24) of the DIF causes a progressively worsening skew of resonant peak values compared to the theoretical values and those present in the constant F_s RIR in the 125 and 250 Hz octave bands. Note, as previously highlighted, that the low phase delay present at higher frequency octave bands results in a much higher correlation between constant F_s and DIF RIRs in the upper simulation bands (as shown in Figure 4.25).

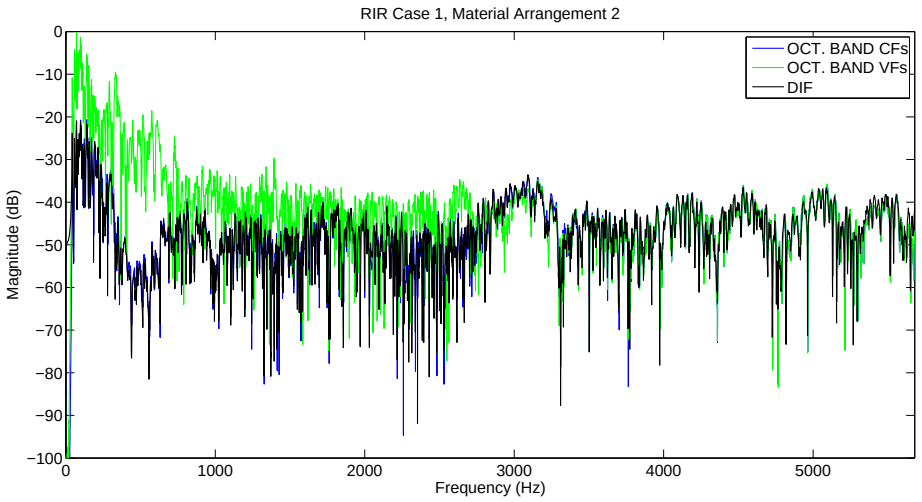

FIGURE 4.27: Magnitude spectra of constant F_s octave band (OCT. BAND CFs - blue line), variable F_s octave band (OCT. BAND VFs - green line) and DIF (black line) boundary simulation RIRs for source/receiver case and material arrangement 2.

Material Arrangement 2:

Following the same analysis procedure as used in the material arrangement 1 case, the resulting spectra of RIRs simulated for material arrangement 2 are displayed in Figure 4.27. Again, the variable F_s RIR exhibits the previously discussed energy imbalance across adjacent octave bands. By way of contrast, the agreement between DIF and constant F_s spectra is consistently high across the full simulation bandwidth with only small differences in magnitude due to the errors associated with contributing DIFs. Figure 4.28 further highlights the agreement of the DIF and constant F_s RIR spectra. As shown, the positioning and magnitude of resonant peaks is much more consistent compared to those resulting for material arrangement 1. This is due to the application of the concrete material DIF to the majority of the model surface area. As previously described, the concrete material DIF maintains a lower phase delay than the plasterboard material DIF for the lower 3 octave bands. Therefore, the deviation of resonances present in the DIF model RIR is less in this modelling case.

4.5.4.2 Acoustic Parameters

The simulated RIRs are analysed here in terms of the EDT, T_{30} and C_{80} ISO parameters [17] described in section 2.3.8. This analysis is conducted to ascertain the agreement between the acoustic characteristics rendered using each boundary modelling approach applied to the case study room for both material arrangements. Due to the fact that each RIR is bandpass filtered into frequency octaves in order to calculate parameter

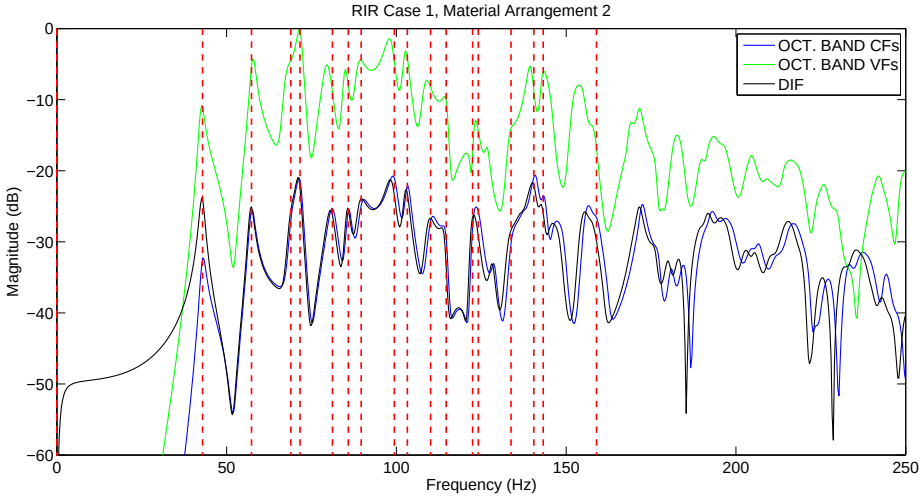

FIGURE 4.28: Low frequency magnitude spectra of constant F_s octave band (OCT. BAND CFs - blue line), variable F_s octave band (OCT. BAND VFs - green line) and DIF (black line) boundary simulations for material arrangement 2. Theoretical mode values, up to second-order, are provided by the vertical dashed red lines.

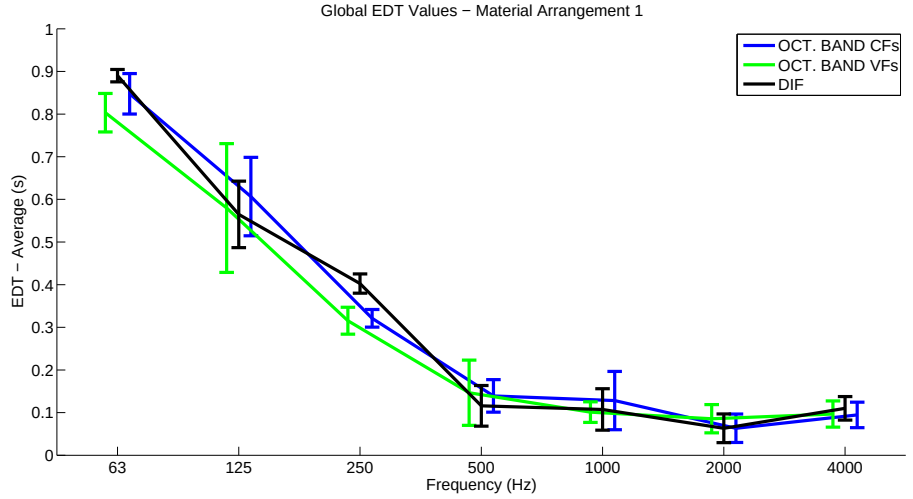

FIGURE 4.29: Global early decay time (EDT) (s) values of constant F_s octave band (OCT. BAND CFs), variable F_s octave band (OCT. BAND VFs) and DIF based simulation RIRs. Results shown are calculated by averaging values derived from the three RIR cases for model material arrangement 1 for each octave band. Standard derivations are presented by the error bars.

values for each octave band, the erroneous energy levels present in the variable F_s RIRs is expected to have minimal impact on the results documented here. Furthermore, the parameter values returned by each simulated RIR are combined and averaged across all three source and receiver cases. Hence, one *global average* value is given for each parameter (EDT, T_{30}, C_{80}) per modelling approach (variable F_s, constant F_s, DIF based) per octave band (63, 125, 250, 500, 1k, 2k, 4k (Hz)).

Material Arrangement 1:

In general, it is expected that differences will exist between the ISO parameters returned from each RIR rendering method due to the inherent variations in model characteristics. However, despite these variations, the isolated octave-band acoustic characteristics exhibited by each model are encouragingly similar. Figure 4.29 depicts the global average EDT calculated for the case study domain RIRs using each frequency-dependent boundary modelling approach. As can be seen, the general trend of all model results are in agreement with the EDT progressively decreasing across the four lower octave-bands and settling on comparatively low decay times in the 1 - 4 kHz range. The standard deviation ranges presented by the error bars give useful insight into the variance that arises between EDT values for each source/receiver case and each modelling approach. It is observed that larger deviations between values occur in the lower three octave bands. In this frequency range, the constant and variable F_s global EDT and related variance values share good agreement with the exception of a larger deviation in the variable F_s RIRs in the 125 Hz band. This result may be attributed to the differences in model geometry and source/receiver placements.

The differences arising between the octave band and DIF-based global EDTs are shown to be consistently low with the largest discrepancies arising in the 125 and 250 Hz octave bands. It is observed that this discrepancy is inconsistent with the error associated with the impedance values given by the magnitude response of the filter applied which is consistently lower than that required in this frequency range. Hence, a lower EDT is expected in the 250 Hz band for the DIF-based model than for the case of the constant and variable F_s implementations. However, referring to Figure 4.30, it is shown that the T_{30} measure of RT_{60} better demonstrates the impact of the variance in surface impedance values between octave band and DIF-based models. The lower reverberation times recorded in the DIF-based model for the 250 and 500 Hz octave bands are representative of the lower impedance values applied in this frequency range. Additionally, the higher T_{30} observed in the lower 63 Hz octave band for the DIF-based model is also consistent with error between the ideal and derived DIF magnitude response. The similarity between decay time characteristics in each octave band model is consistently high, with the single exception of the 2 kHz octave band. This result constitutes an anomaly given the agreement between the trends in T_{30} for both models and may be the result of sensitive dependence of decay times on energy calibration and/or dimensional differences inherent to the variable F_s modelling approach applied in this study.

As displayed in Figure 4.31, the Clarity (C_{80}) measures returned for both DIF-based and constant F_s models are in strong agreement with one exception arising in the

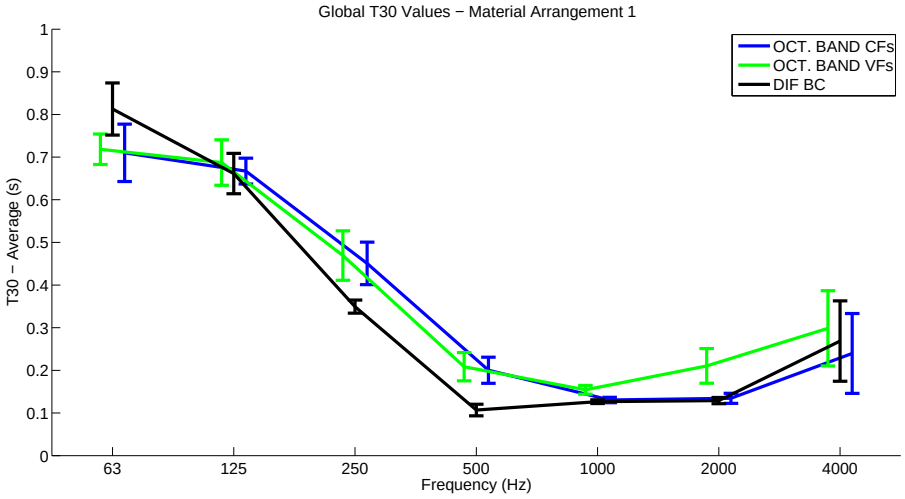

FIGURE 4.30: Global T_{30} (RT_{60}) (s) values of constant F_s octave band (OCT. BAND CFs), variable F_s octave band (OCT. BAND VFs) and DIF based simulation RIRs. Results shown are calculated by averaging values derived from the three RIR cases for model material arrangement 1 for each octave band. Standard derivations are presented by the error bars.

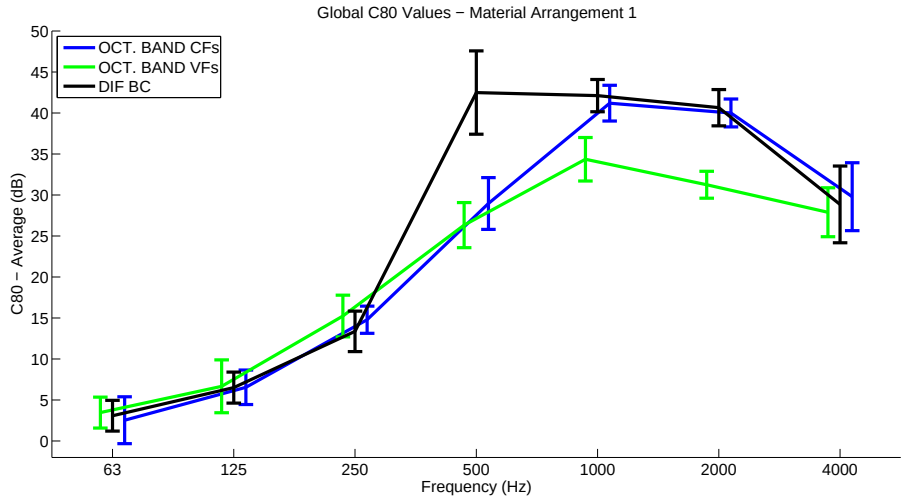

FIGURE 4.31: Global C_{80} (Clarity Index) (dB) values of constant F_s octave band (OCT. BAND CFs), variable F_s octave band (OCT. BAND VFs) and DIF based simulation RIRs. Results shown are calculated by averaging values derived from the three RIR cases for model material arrangement 1 for each octave band. Standard derivations are presented by the error bars.

500 Hz octave band. The higher Clarity level exhibited in the frequency region is a direct result of the impedance error introduced by the DIF and is consistent with the comparatively low T_{30} value previously demonstrated. The overlapping of variance and similarity of trends exhibited by the constant and variable F_s C_{80} values at low frequencies adds credence to the claim that lower resolution schemes are capable of emulating acoustic behaviour to a good level of accuracy. Unfortunately, this agreement of results diminishes in the 1 - 2 kHz octave bands. Given that T_{30} has been shown to be longer in this frequency range, this result may be related to the assumed sensitive dependence on model characteristics as previously postulated.

Finally, the 4 kHz octave band models rendered for the constant and variable F_s cases are exactly the same as they incorporate the same inter-nodal distance corresponding to $F_s = 28.4$ kHz. Hence, it follows that the results returned should be exactly the same. However, for all parameters analysed in this frequency range, this is shown to not be the case. It is expected that the bandpass filtering stage applied prior to RIR analysis, as per [17], should minimise the impact of adjacent band characteristics on the calculated results. However, the application of finite order filters allows for crosstalk between adjacent frequency bands. Hence, properties of the 2 kHz variable F_s octave band model, which has been shown to produce significantly higher RIR energy levels than that of the constant F_s model, are introducing variance between the values returned for the upper most simulation frequency range. In turn, it is noted that the method of equating energy levels across the variable F_s octave band RIRs requires further attention, not only to remedy the spectral discrepancies previously discussed, but also to reduce the differences in reverberation and Clarity measurements.

Material Arrangement 2:

Applying the same method of analysis to the second material arrangement model yields results that further demonstrate the agreement of acoustic characteristics across each boundary modelling method. Observing the nature of frequency-dependent EDT, displayed in Figure 4.32, it is apparent that all models are in agreement in terms of overall trends. For this material arrangement, three DIFs are used to model the frequency-dependency of the LRS boundary condition. Hence, ascertaining the cause of differences between recorded parameter values for constant F_s and DIF-based RIRs becomes more involved. Referring back to Figures 4.19, 4.21 and 4.22, which depict the magnitude response of the plaster board, wood and concrete DIFs respectively, the following DIF properties may be noted:

- Both the concrete and plasterboard surface impedance values are higher than those applied in the constant F_s model over the range of the lowest 63 Hz octave

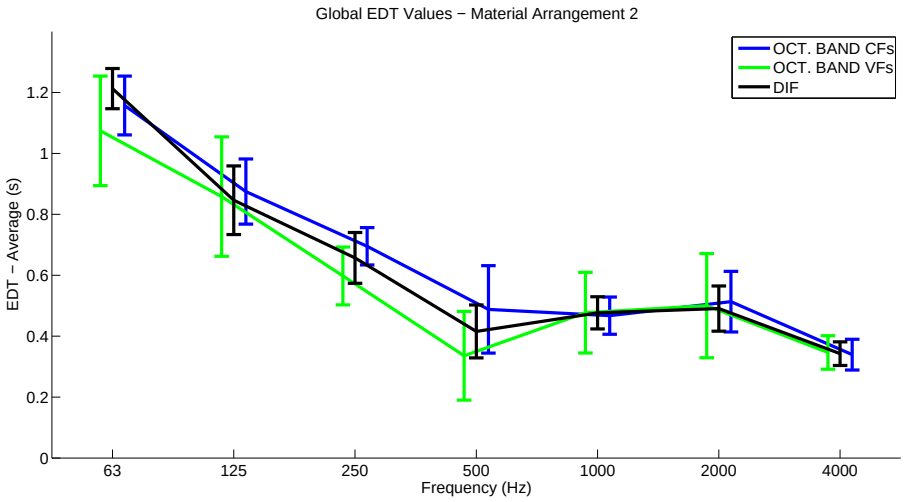

FIGURE 4.32: Global early decay time (EDT) (s) values of constant F_s octave band (OCT. BAND CFs), variable F_s octave band (OCT. BAND VFs) and DIF based simulation RIRs. Results shown are calculated by averaging values derived from the three RIR cases for model material arrangement 2 for each octave band. Standard derivations are presented by the error bars.

band. For the same octave band, the wood surface impedance values closely match those defined in the octave band model.

- The concrete material, which is applied to the four vertical walls of the model, possesses impedance values that are significantly lower than those defined in the octave band models for frequencies above $\sim$ 200 Hz.

The differences between recorded EDT for DIF-based and constant F_s RIRs may, in part, be attributed to the properties listed above. Firstly, the higher impedance values for plasterboard and concrete materials in the 63 Hz band cause longer decay times in the DIF-based RIRs compared to the constant F_s RIRs in this frequency region. This is shown in Figure 4.32. Secondly, in the 125, 250 and 500 Hz octave bands, the DIF-based RIRs produces EDTs that are shorter than those recorded in the constant F_s model. This is most likely due to the lower impedance values of the concrete DIF applied at higher frequencies. However, these lower EDT values are not recorded for the upper 1, 2 and 4 kHz octave bands in which the concrete impedance values are also lower than those of the constant F_s model. This suggests that the overall higher impedance characteristics of the constant F_s model is acting to retain a higher level of low frequency resonances which, in turn, increases the difference between the EDTs for the two models in the 125 - 500 Hz octave bands.

In terms of the variable F_s EDT values, it is apparent that a larger variance between recorded decay times for each RIR case compared to those produced by the other two

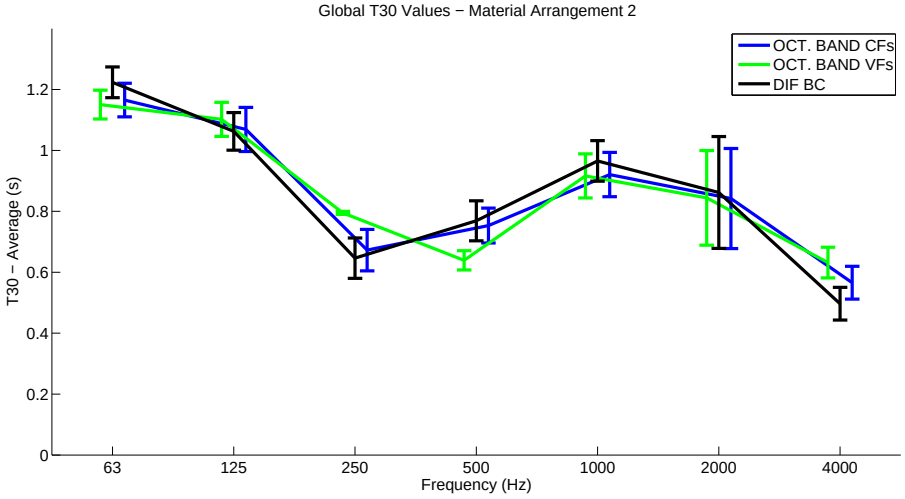

FIGURE 4.33: Global T_{30} (RT_{60}) (s) values of constant F_s octave band (OCT. BAND CFs), variable F_s octave band (OCT. BAND VFs) and DIF based simulation RIRs. Results shown are calculated by averaging values derived from the three RIR cases for model material arrangement 2 for each octave band. Standard derivations are presented by the error bars.

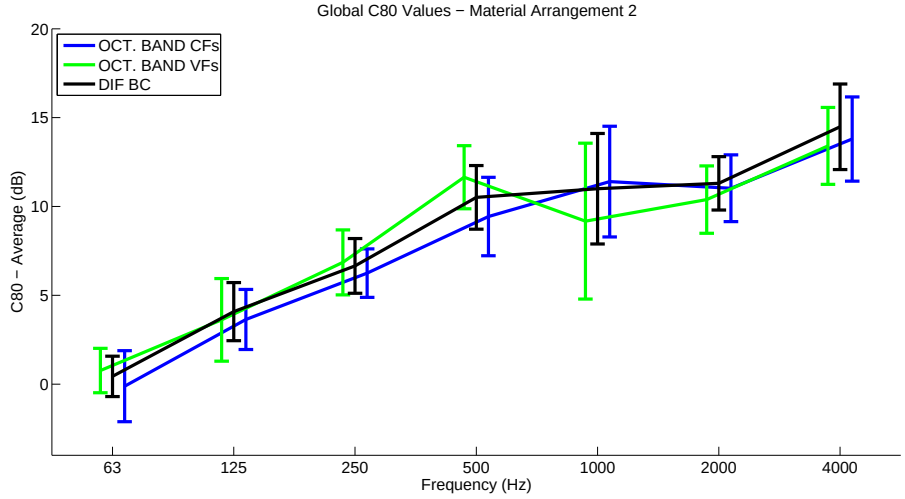

FIGURE 4.34: Global C_{80} (Clarity Index) (dB) values of constant F_s octave band (OCT. BAND CFs), variable F_s octave band (OCT. BAND VFs) and DIF based simulation RIRs. Results shown are calculated by averaging values derived from the three RIR cases for model material arrangement 2 for each octave band. Standard derivations are presented by the error bars.

models. However, the overall trend of frequency-dependent decay times is similar to that of the closely matched DIF-based and constant F_s RIRs. Furthermore, a strong agreement between average EDT values for the 1 - 4 kHz octave bands is demonstrated. Again, the variable F_s model is shown to be capable of simulating acoustic characteristics similar to those produced by the constant F_s model. Differences in calculated values are assumed to arise for the reasons put forward in the previous sections.

Encouraging results are calculated for T_{30} and C_{80} measures as presented in Figures 4.33 and 4.34 respectively. As shown in Figure 4.33, it is observed that the level of agreement between reverberation times recorded for DIF-based and constant F_s model RIRs is consistently high across the simulation bandwidth. In addition, the trend of variable F_s model T_{30} values is consistent with those of the DIF-based and constant F_s models with only slight differences occurring in the 250 Hz and 500 Hz octave bands. This overall agreement of T_{30} observed between all modelling approaches, in turn, contributes to a level of consistency across C_{80} values (see Figure 4.34). Again, highly similar C_{80} trends are observed for each modelling approach across the simulation bandwidth. These values act to support the claim that the simulation of equivalent acoustic characteristics using the three implementations of frequency-dependent LRS boundaries is an attainable target with reference to suggested model refinements provided in section 4.5.4.4.

4.5.4.3 Simulation Data

In order to demonstrate, in a broad sense, the computational requirements of each modelling approach investigated in this case study, the run-time (per second of audio output) and number of required pressure nodes are provided in Table 4.4. As shown, the variable F_s octave band model requires the smallest run-time and memory allocation in terms of number of pressure nodes. Sequentially, this approach can be simulated in approximately 5x less time than a sequential simulation of the constant F_s model. Additionally, a parallel implementation, whereby each contributing octave band model is simulated simultaneously by means of multiple processors or otherwise, achieves a reduction in memory requirement of 310 MB compared to a constant F_s parallel simulation approach. However, as demonstrated in the previous discussion of simulation results, these computational savings come at the cost of differences between rendered variable and constant F_s RIR properties.

The DIF model requires over 1.5x the number of pressure nodes required by each constant F_s model. This demonstrates that, in cases of large-scale FDTD models with high spatial sampling resolution, the constant F_s approach may be favorable as an equivalent DIF-based approach may exceed computational resources. The same consideration may also arise for implementations that require high order DIFs to achieve the required frequency-dependent surface absorption characteristics. This is demonstrated in the models created for this case study, the additional boundary filters cause an increase of approximately 48 and 66 MB for material arrangements 1 and 2 respectively. However, recorded run-times suggest that the DIF implementation is a more

Model	Band (Hz)	F_s	Nodes	Run time (s)	Mem. (MB)
OB VF_s	63, 125, 250	2800	11700	0.03	0.09
OB VF_s	500	3700	26520	0.12	0.20
OB VF_s	1000	7100	168609	1.33	1.29
OB VF_s	2000	14200	1264896	26.93	9.65
OB C/VF_s	4000	28400	9953280	429.23	79.54
DIF MA1	[ALL]	28400	16669440	1670.00	127.18
DIF MA2	[ALL]	28400	17691072	2014.00	135.00

TABLE 4.4: Run time and number of grid points (nodes) used for each model simulation. Memory requirements for double floating-point precision simulations are also provided. Key: OB VF_s - Octave Band Variable F_s, OB C/VF_s - Octave Band Constant and/or Variable F_s , DIF MA1 - DIF model, material arrangement 1, DIF MA2 - DIF model, material arrangement 2.

efficient approach than a sequential constant F_s simulation, achieving a comparative ∼4.1x reduction in required simulation time.

4.5.4.4 Case Study Discussion

This comparative investigation into implementation of frequency-dependent LRS boundaries yields useful insights towards future incorporation of realistic surface conditions in FDTD simulations. Findings drawn from rendered RIRs suggest that all three approaches are capable of producing similar results both in terms of spectral characteristics and acoustic parameters while noting the following summarised observations:

- The octave band RIR energy calibration process applied to the variable F_s model results produces energy levels that are consistently higher than those recorded in the DIF-based and constant F_s RIRs. As such, it is proposed that the application of the calibration constant be revisited in order to ascertain the cause of this erroneous result.

- Derivation of appropriate DIF coefficients for accurate representation of frequency-dependent impedance characteristics is problematic when relying on initial formulation of reflection filter magnitude curves. Direct calculation of filter coefficients for an interpolated impedance curve will avoid the observed magnification of errors introduced when transferring from approximate reflection curves to impedance values.

- Phase response characteristics inherent to DIFs, if not controlled as in this case, may cause unacceptable deviation of modal frequency values from those expected

from theory. In the absence of surface phase response data, it is therefore recommended that the phase delay/advance of each DIF applied are constrained to low values as far as possible.

- The nature of, and values returned for, acoustic parameters may be directly attributed to the accuracy of the DIFs applied to boundary surfaces in terms of impedance values. This claim is valid when assuming the constant F_s approach provides accurate results, thus constituting the base case for comparison. However, there does not appear to be a straightforward means of identifying exact causes of differences between constant and variable F_s octave band simulation results prior to remedying the issue of energy calibration noted above. It is evident that variance in parameter values do occur due to differences in model geometry and source/receiver placement.

To conclude, this study informs the design and implementation of hybrid modelling approaches documented in section 5.3 by highlighting the constant F_s octave band approach as the optimal means of frequency-dependent LRS boundary modelling. This selection is justified as both DIF-based and variable F_s approaches have been demonstrated as requiring further refinement in order to improve the accuracy of results. However, areas for further research towards this outcome are identified here that seek to improve the accuracy of DIF-based and variable F_s methods.

4.5.5 Numerical Stability Issues

As shown previously in section 4.4.4, the pressure-centered LRS approximation is prone to instability for 3D topologies that involve combinations of re-entrant node types (i.e. edges and corners). In order to study this further an example is presented here to highlight the merits of applying velocity-centered approximations to frequency-dependent boundaries modelled with DIFs. It is also necessary to reformulate the DIF LRS boundary conditions such that both pressure-centered and velocity-centered approximations may be derived in a similar fashion in the interests of consistency. This is not possible following the derivation in [110] which is applicable only to pressure-centered LRS approximations.

RIRs are rendered using both pressure-centered and velocity-centered frequency-dependent DIF formulations for the test topologies presented in section 4.4.4. Each scheme is run at the Courant limit with a temporal sampling rate of $F_s = 28.4$ kHz yielding a grid spacing of $h = 0.021$ m for wave speed $c = 344$ ms^{-1}. All scheme calculations are conducted using double point precision and each boundary node utilises DIF coefficients

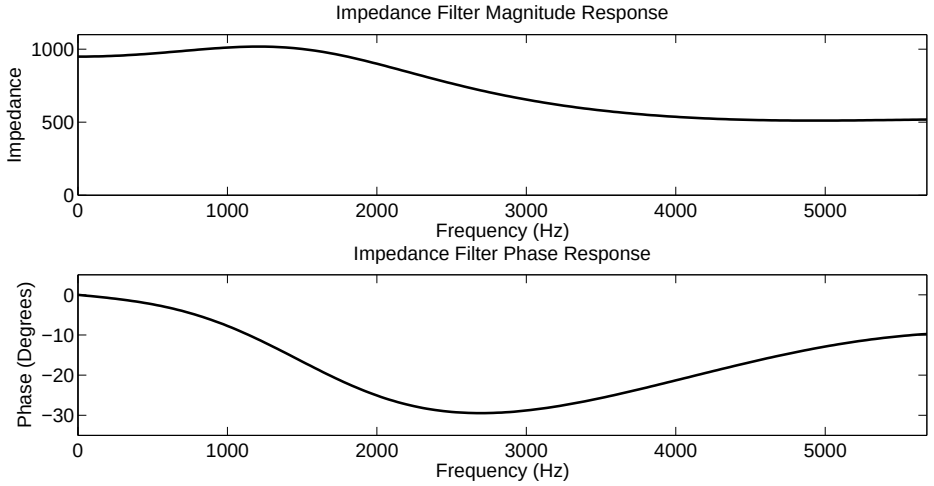

FIGURE 4.35: The magnitude and phase response characteristics of the DIF applied to investigate the stability of pressure-centered and velocity-centered frequency-dependent LRS boundary formulations as utilised in 3D SRL FDTD acoustic models.

derived to produce the filter characteristics displayed in Figure 4.35. As shown, the resulting DIF maintains high impedance values across the valid simulated bandwidth of 0 - 0.196 F_s Hz. The impedance characteristics translate to high reflection coefficients in the range $R = [0.996{:}0.998]$. This enables simulations to be run with very low absorption rates as are applied in the case of the frequency-independent LRS stability case study. Phase response characteristics are unconstrained during the calculation of required filter coefficients using the MATLAB fdesign.arbmag function.

4.5.5.1 Reformulation of DIF Boundary Model

To begin the reformulation of the pressure-centered LRS boundary, a right-hand surface termination is considered:

$$p^n_{l+1,m,q} = -\frac{1}{\lambda\zeta_\omega}(p^{n+1}_{l,m,q} - p^{n-1}_{l,m,q}) + p^n_{l-1,m,q} \tag{4.74}$$

which, when inserted into the SRL wave equation update for the pressure node at the boundary, gives:

$$p^{n+1}_{l,m,q} = \lambda^2(-\frac{1}{\lambda\zeta_\omega}(p^{n+1}_{l,m,q} - p^{n-1}_{l,m,q}) + 2p^n_{l-1,m,q} + p^n_{l,m+1,q} + p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q+1}) + 2(1-3\lambda^2)p^n_{l,m,q} - p^{n-1}_{l,m,q} \tag{4.75}$$

This expression can be re-arranged as follows considering $\zeta_\omega = \frac{b_0+B(\mathrm{z})}{a_0+A(\mathrm{z})}$:

$$p^{n+1}_{l,m,q}(\lambda a_0 + \lambda A(\mathrm{z}) + b_0 + B(\mathrm{z})) = \lambda^2(b_0 + B(\mathrm{z}))(2p^n_{l-1,m,q} + p^n_{l,m+1,q} + p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q-1}) + 2(b_0 + B(\mathrm{z}))(1 - 3\lambda^2)p^n_{l,m,q} + (\lambda a_0 + \lambda A(\mathrm{z}) - b_0 - B(\mathrm{z}))p^{n-1}_{l,m,q} \quad (4.76)$$

Finally, by expressing $C(\mathrm{z}) = \lambda A(\mathrm{z}) + B(\mathrm{z})$, $D(\mathrm{z}) = \lambda A(\mathrm{z}) - B(\mathrm{z})$ and re-arranging for $p^{n+1}_{l,m,q}$ gives the final update expression for a right hand boundary node:

$$p^{n+1}_{l,m,q} = \Big(\lambda^2(b_0 + \ B(\mathrm{z}))(2p^n_{l-1,m,q} + p^n_{l,m+1,q} + p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q-1}) + 2(b_0 + B(\mathrm{z}))(1 - 3\lambda^2)p^n_{l,m,q} + (\lambda a_0 - b_0 + D(\mathrm{z}))p^{n-1}_{l,m,q} - C(\mathrm{z})p^{n+1}_{l,m,q}\Big)\Big(\frac{1}{\lambda a_0 + b_0}\Big) \quad (4.77)$$

While the above expression appears implicit due to the unknown term, $p^{n+1}_{l,m,q}$, resulting on the right hand side, this is not the case. The update equation is explicit as required due to the fact that $p^{n+1}_{l,m,q}$ is multiplied by $C(\mathrm{z})$ referring to delayed pressure quantities for the boundary node at $[l, m, q]$ which are known. Update expressions for edge and corner boundaries may be derived by similar means. It is noted in [110] that this approach to formulation of the DIF boundary update can lead to instabilities. For this reason, the experiments conducted with combinations of re-entrant boundary nodes also include simulations that utilise the pressure-centered formulation described in section 4.5.3.

For the case of the velocity-centered approximation, the right-hand boundary node expression is:

$$p^n_{l+1,m,q} = -\frac{1}{2\lambda\zeta_\omega}(p^{n+1}_{l,m,q} - p^{n-1}_{l,m,q}) + p^n_{l,m,q} \quad (4.78)$$

which, when inserted into the discrete wave equation (4.4), gives:

$$p^{n+1}_{l,m,q} = \Big(\lambda^2(b_0 + B(\mathrm{z}))(p^n_{l-1,m,q} + p^n_{l,m+1,q} + p^n_{l,m-1,q} + p^n_{l,m,q+1} + p^n_{l,m,q-1}) + (b_0 + B(\mathrm{z}))(2 - 5\lambda^2)p^n_{l,m,q} + (\lambda a_0 - b_0 + D(\mathrm{z}))p^{n-1}_{l,m,q} - C(\mathrm{z})p^{n+1}_{l,m,q}\Big)\Big(\frac{1}{\lambda a_0 + b_0}\Big) \quad (4.79)$$

where it is assumed that all DIF terms account for the doubling of the impedance term given in (4.55). Derivations for edge and corner update expressions are developed by solving for additional unknown ghost node values as required.

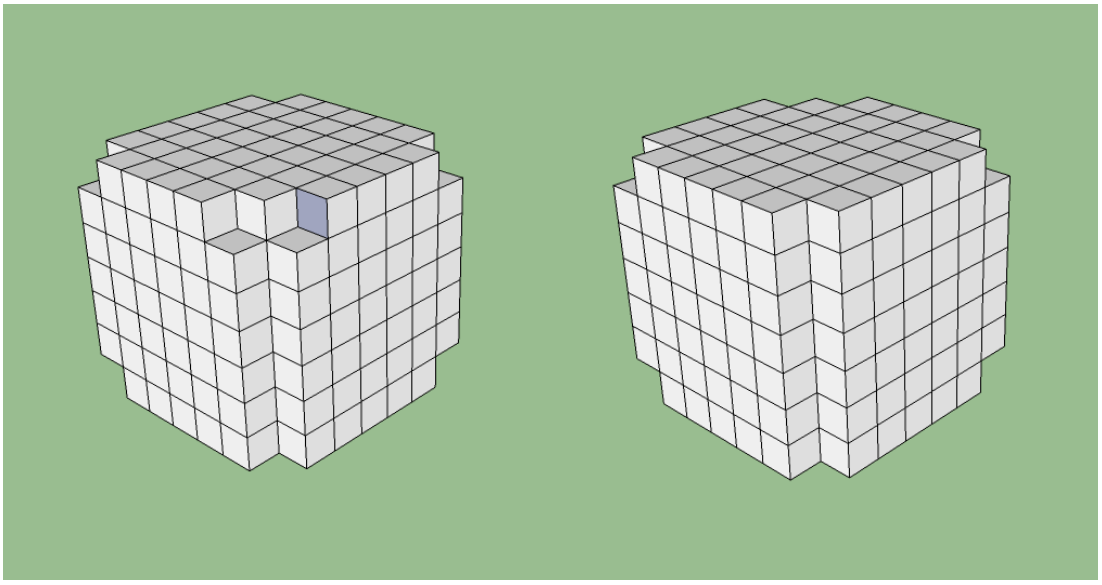

FIGURE 4.36: Arrangement of nodes in a 8 x 8 x 8 SRL lattice used for the purposes of investigating numerical stability in simulations with DIF-based frequency-dependent LRS boundaries. The left and right diagrams display the domain in diagonally opposing orientations to highlight the inclusion of multiple re-entrant node configurations.

4.5.5.2 Experimental Procedure and Results

In total, three small spatial domains were simulated for the purposes of investigation. The two topologies previously described for the frequency-independent LRS stability investigation (see Figure 4.14) were used along with a further, more complex, third domain. Figure 4.36 displays the arrangement of grid cells in this additional model. As shown, this additional model topology incorporates two diagonally opposing re-entrant edges. All corners of the 8 x 8 x 8 grid point cube are defined as re-entrant nodes with one corner possessing two re-entrant corners.

Each model domain is simulated using three implementations of the DIF-based frequency-dependent LRS boundary conditions:

- The pressure-centered approximation which makes use of an intermediate filter term as documented in [110, 124] and detailed in section 4.5.3.
- The pressure-centered approximation derived by direct manipulation of the boundary update scheme equations as per (4.77).
- A velocity-centered approximation derived by direct manipulation of resulting boundary update scheme equations as per (4.79).

Each simulation is run to produce 5 seconds of audio output with $F_s = 28.4$ kHz. Resulting RIRs are plotted in Figures 4.37, 4.39 and 4.40 for each of the three model

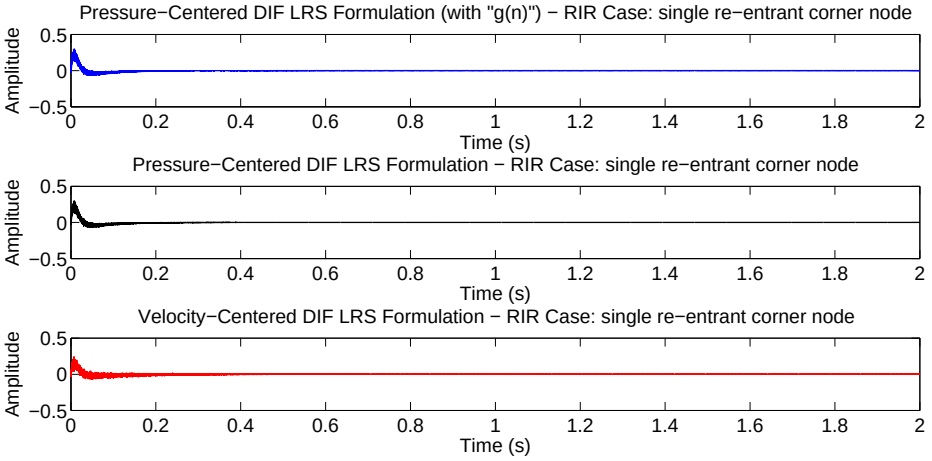

FIGURE 4.37: Simulated RIRs for the domain geometry that includes a single re-entrant corner node with frequency-dependent boundaries implemented using: (top) the pressure-centered formulation as derived in section 4.5.3; (middle) the pressure-centered formulation (4.77); (bottom) the velocity-centered formulation (4.79). The first 2 seconds of audio output (F_s = 28.4 kHz) are shown.

domains. Visual examination of the RIRs simulated in the model domain with a single re-entrant corner (see Figure 4.37) demonstrates that all LRS boundary implementations are preserving numerical stability. The initial 2 seconds of RIR output are displayed to increase clarity in the plots provided. Furthermore, the two different approaches to formulating the pressure-centered DIF-based LRS conditions are shown to produce very similar results, as per Figure 4.38. Referring to the bottom panel of this figure, the numerical difference between the two simulated RIRs (for the case of the single re-entrant node domain) may be assumed as the result of arithmetic inaccuracies due to finite precision.

A notable result is observed for the case of the simulated domain that incorporates diagonally opposing re-entrant corner and edge nodes. It is demonstrated in section 4.4.4, that this particular node arrangement gives rise to instabilities when applying frequency-independent pressure-centered LRS conditions to boundary nodes. However, in the case of frequency-dependent DIF-based LRS boundaries, it is demonstrated (as per Figure 4.39) that simulations may be run while maintaining numerical stability. This result is of particular interest as further examination may reveal the cause or conditions that encourage the amplification of solutions within a scheme with pressure-centered LRS boundaries. For example, an upper limit on the impedance value applied to boundaries may exist for a given frequency of resonance such that below this value, the scheme will maintain stability. This initial rationalisation is justified when considering the difference of impedance values applied through use of the DIF as previously

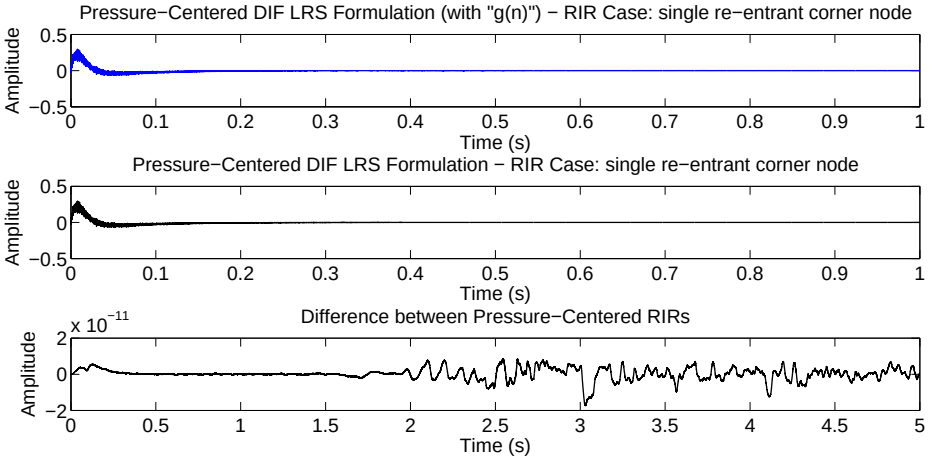

FIGURE 4.38: A comparison between the two formulations of pressure-centered DIF-based LRS boundary implementations for the re-entrant corner domain (shown top and middle for the first 2 seconds of audio output with $F_s = 28.4$ kHz). The small, non-increasing, level of numerical error between the results of the two schemes is shown for the total duration of the simulations (bottom).

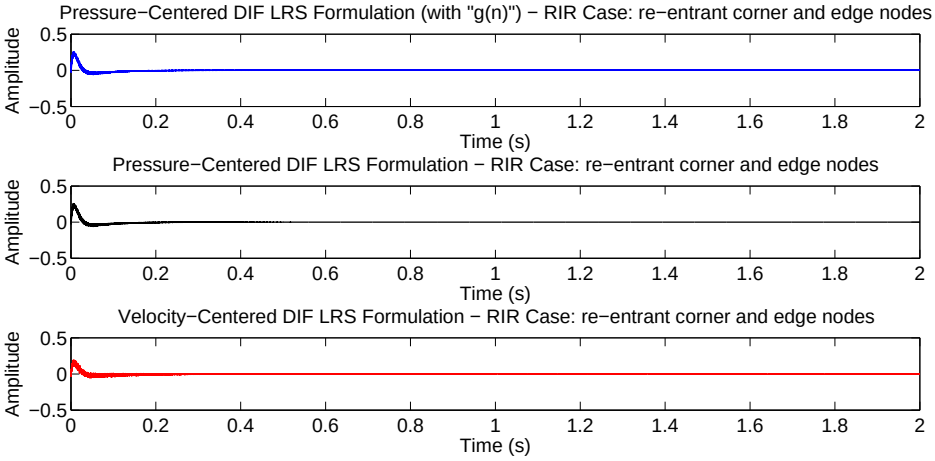

FIGURE 4.39: Simulated RIRs for the domain geometry that includes diagonally opposing re-entrant corner and edge nodes with frequency-dependent boundaries implemented using: (top) the pressure-centered formulation as derived in section 4.5.3; (middle) the pressure-centered formulation (4.77); (bottom) the velocity-centered formulation (4.79). The first 2 seconds of audio output are shown ($F_s = 28.4$ kHz).

discussed. However, seeking a generalised condition of this kind for all possibilities of domain geometries poses a significant problem when relying on empirical investigation such as is documented here.

Nonetheless, the simulation results depicted in Figure 4.40 demonstrate the potential for instability for the pressure-centered LRS formulation as the geometry of the

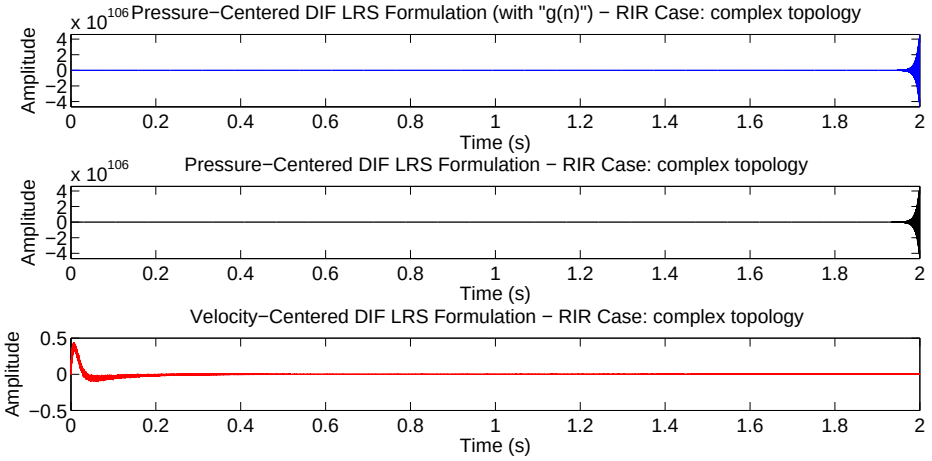

FIGURE 4.40: Simulated RIRs for the complex domain geometry with frequency-dependent boundaries implemented using: (top) the pressure-centered formulation as derived in section 4.5.3; (middle) the pressure-centered formulation (4.77); (bottom) the velocity-centered formulation (4.79). The first 2 seconds of audio output are shown ($F_s = 28.4$ kHz).

modelled domain increases in complexity. In this instance, the boundary surfaces are arranged such that numerous re-entrant nodes are present within the space. As displayed in the top two graphs, the RIRs generated for both pressure-centered DIF based boundary conditions exhibit the exponential increase in amplitude that is the result of numerical instability. Conversely, the velocity-centered formulation preserves stability with no evidence of amplified solutions present in the resulting RIR. Hence, as with the case of the frequency-independent boundary conditions discussed previously, it may be concluded from this short empirical investigation that the robustness of velocity-centered LRS boundary formulations exists also for frequency-dependent cases. This result is of significant value when considering the application of such modelling approaches to realistic room topologies, such as those documented in Chapter 5.

4.6 Summary

This chapter has highlighted several important considerations relating to the design and implementation of FDTD acoustic models. Investigations into modelling elements, including dispersion characteristics, boundary conditions and sound source excitation/receiver methods have been conducted in order to provide a basis for the construction and evaluation of the hybrid modelling method proposed in this work. Key findings that inform the design choices applied in this novel hybrid approach are given below.

- In order to ensure that numerical dispersion effects have minimum impact on audible results rendered using SRL FDTD schemes, the upper frequency limit on the usable bandwidth, f_u, must be defined in accordance with the relations $f_u \leq 0.075\ F_s$ for 3D and $f_u \leq 0.1\ F_s$ for 2D. These conditions ensure that the isotropy of the schemes remain within a 2% error bound for all directions of wave propagation.

- Soft sound source implementations form the most simple and, arguably, inherently correct means of exciting the virtual acoustic field in a FDTD model. When applied in conjunction with a simple DC filtering process, this source type produces results which can be post-processed to synthesise RIRs with minimal computational overhead.

- To imitate the sound source characteristics resulting in a 3D FDTD system, a 2D excitation signal must be manipulated by means of a simple filtering process to remove the afterglow effects that are inherent to the Green's function for the wave equation in 2D.

- For complex spatial domain topologies where SRL LRS boundary conditions are to be applied to model surfaces, the velocity-centered formulation provides the preferred approach as it has been demonstrated to be more robust against numerical stability issues compared to the pressure-centered formulation.

- For acoustic models that require frequency-dependent boundary conditions, the constant F_s octave band approach has been demonstrated to produce the most accurate results in terms of theoretical low frequency response characteristics with less computational requirements than that needed for the alternative DIF-based approach.

It has also been shown that while the velocity-centered boundary formulation maintains numerical stability to a greater extent than the pressure-centered case, the nature of absorption characteristics is not as expected from the LRS model. Initial findings from the comparative investigation of energy decay in 1D FDTD systems with velocity-centered and pressure-centered LRS boundaries (see section 4.4.3) demonstrate a difference between RT_{60} calculated from each system which increases as the reflection coefficient increases.

The variable F_s octave band approach to frequency-dependent boundary implementations has been demonstrated as able to give acoustic parameters to a level of accuracy that is, on the most part, consistent with more computationally intensive DIF-based and constant F_s octave band approaches. However, these results will, in part, be affected by differences in model geometries and source/receiver placement and the energy

matching procedure applied to RIR contributions for each octave band which has not produced expected results in terms of overall magnitude levels in resulting RIR spectra.

Finally, experimentation with different approaches to DIF-based frequency-dependent LRS boundary implementations has revealed the advantage for formulating the update scheme using velocity-centered approximations. By applying this formulation, as with the frequency-independent case, FDTD acoustic models may be run for complex geometries with increased robustness against numerical instability.

Chapter 5

Multiplane Hybrid RIR Synthesis

The work documented in this chapter investigates the design, implementation and analysis of a 2D multiplane FDTD method as part of an efficient hybrid RIR synthesis system. In brief, the multiplane method seeks to model low frequency sound propagation through the use of two or more independent 2D schemes which represent 2D planar cross sections of the target acoustic environment.

Three studies, which seek to demonstrate the applicability of the multiplane FDTD paradigm to low frequency acoustic modelling, are described. A pilot study concerning the emulation of enclosed sound propagation in a simple cuboid environment provides insight into the accuracy that may be achieved by the proposed hybrid model. RIR contributions are produced by FDTD and geometric models in the low and mid-high frequency regions (as per Figure 5.1) to render spectrally complete hybrid RIRs. Results are discussed with reference to those produced by 3D FDTD simulations in order to assess the extent to which the multiplane model is capable of emulating low frequency soundfields.

The remaining two studies examine the application of the multiplane hybrid model to the virtual representation of real acoustic environments of increasing complexity. For these modelling scenarios, data recovered from practical measurements of the real spaces are used in order to evaluate the accuracy of the multiplane method. Additionally, 3D FDTD models are rendered for each simulation case to provide further reference data for the purposes of comparison and evaluation.

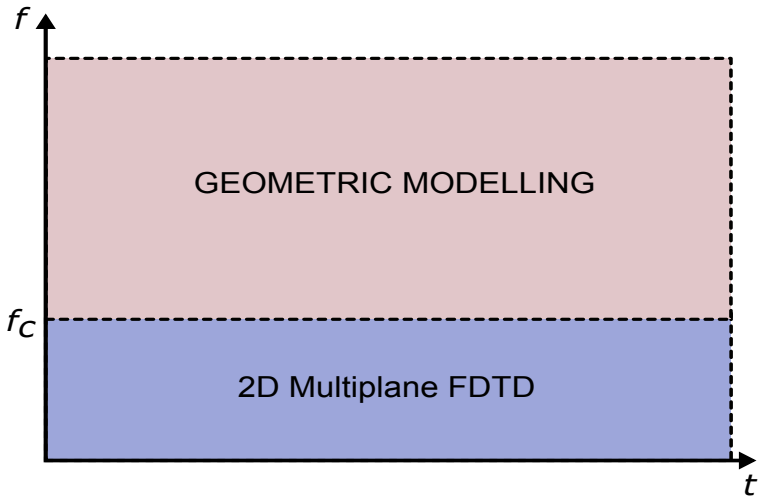

FIGURE 5.1: Diagram depicting the contribution of geometric and 2D multiplane FDTD models to the creation of hybrid RIRs. As shown, the 2D multiplane model renders the low frequency component of the full temporal range of hybrid RIRs while the geometric approach yields RIR contributions above a defined transition frequency f_c.

5.1 Pilot Study: A Simple Cuboid

The initial investigation of 2D multiplane FDTD acoustic modelling, presented here, constitutes a feasibility test in which the applicability of this new approach to low frequency sound simulation is evaluated. In terms of scope, this pilot study seeks to serve as a proof of concept in support of the claim that the accuracy of 3D FDTD modelling methods may be approached by their 2D counterparts. Low frequency analysis of resulting RIR spectra is used as the basis of analysis to verify this claim. Simulation run-times and memory requirements are presented to demonstrate the comparative efficiency of the multiplane HAM. Additionally, examination of early reflections recorded in each RIR reveals the agreement of temporal aspects simulated in each model. A subset of this work is presented in [156].

5.1.1 The Acoustic Models

The conceptual target acoustic environment modelled in this work consists of a simple cuboid geometry with uniform reflection characteristics applied to each bounding surface. A reflection coefficient of $R = 0.97$ is selected to ensure that absorption is sufficiently low to allow build up of defined resonant frequency peaks, which in turn facilitates comparative spectral analysis. While this choice of room topology and surface characteristics is rarely encountered in practice, the simplicity of the space is justified by the requirement to compare the low frequency spectra of synthesised RIRs under ideal conditions. Hence, although this modelling scenario is useful for the validation of

low frequency multiplane as an initial proof of concept, the applicability of this HAM cannot be assumed for more complex room topologies. Three scenarios are created for the purposes of this pilot study, as described in the following.

5.1.1.1 3D FDTD Model

The 3D FDTD model consists of a SRL lattice (see Chapter 4) which populates a cuboid geometry of approximate dimensions 3 x 4 x 2.5 m. In order to preserve numerical stability and minimise dispersion error, the spatial sampling instance h_{3D} is set equal to the lower limit, given as:

$$h_{3D} = \sqrt{3}ck \tag{5.1}$$

which yields a usable bandwidth of 0.196 F_s. As such, for the selected sampling rate of $F_s = 44100$ Hz and wave speed $c = 344$ ms^{-1}, an inter-nodal distance of $h_{3D} = 0.0135$ m results. It follows that a 222 x 296 x 185 node lattice gives an interior cuboid air cavity of dimensions 2.9859 x 3.9857 x 2.4860 m on which the geometric and 2D multiplane models are based.

CASE	Source (x, y, z) (m)	Receiver (x, y, z) (m)
1	(1.1889, 1.0133, 1.6213)	(1.1889, 2.9994, 1.6213)
2	(0.4999, 0.4999, 0.4999)	(0.4999, 3.4993, 0.4999)
3	(0.2027, 0.2027, 0.2027)	(0.2027, 3.7965, 0.2027)

TABLE 5.1: Source and receiver placements defined for three cases of RIR measurement in a simple cuboid model.

Frequency-independent LRS boundary conditions are defined for all surfaces by means of the pressure-centered formulation which has previously been demonstrated to retain numerical stability for simple topologies. An impedance coefficient of $\zeta = 65.667$, corresponding to the required reflection coefficient of 0.97, is applied to all nodes situated on the boundaries. The interior target acoustic field is excited by a soft-source Kronecker Delta function inserted into the lattice at a node corresponding to the given sound source location. RIRs are then rendered by recording calculated pressure fluctuations at defined receiver locations. Three source and receiver combinations, as listed in Table 5.1, provide three different RIR cases. The arrangement of source/receiver positions, displayed in Figure 5.2, ensures that each RIR case consists of a source and receiver location that share a common vector parallel to the y-axis. This particular arrangement is informed by the way in which the multiplane model is constructed, as described in section 5.1.1.3.

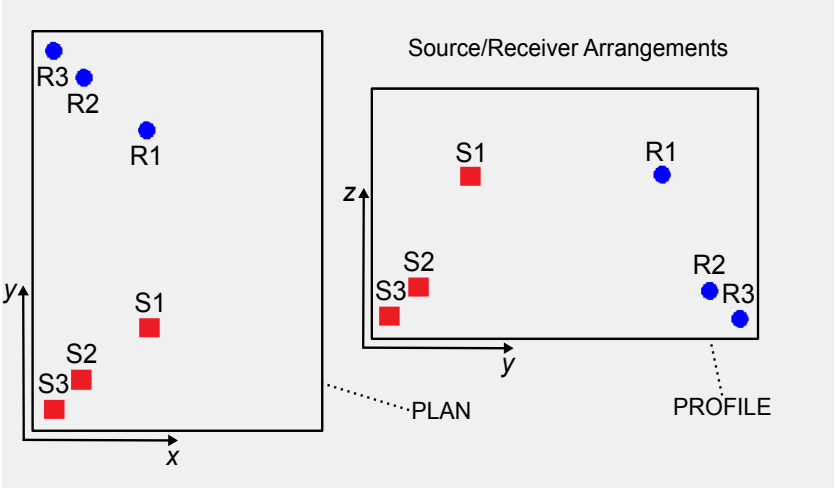

FIGURE 5.2: Plan (left) and profile (right) views of the source and receiver arrangements applied for RIR Cases 1-3. Cartesian axes are shown.

Finally, the 3D FDTD model requires approximately 12.2 million nodes to store one instance of the discretised pressure field, of which two are required for simulation purposes. This corresponds to a memory requirement of ~ 0.182 GB. Simulation run-times for CPU-based implementations are noted to be, on average, 1600 s per second of audio output excluding post-processing required for RIR calibration and hybridisation.

5.1.1.2 Geometric Model

ODEON 10.1 Auditorium [21] is utilised to calculate the mid-high frequency portion of the RIR for each source and receiver combination case. It is widely acknowledged that this acoustic prediction program is of industry standard and, therefore, this software is deemed suitable for the purposes of this work. Briefly, ODEON renders RIRs by means of the ISM for early reflections up to a user-defined reflection order. Above this reflection order, late reflections and diffuse soundfield components are calculated via a ray-tracing algorithm. Control of this variable simulation parameter is provided in ODEON via selection of a required *Transition Order* (T.O.). To accurately emulate the early stages of sound propagation within the virtual model, the T.O. is set to 5^{th}-order. This selection follows the discussion in [48] based on [39] which recommend T.O. ≥ 2 for simulation.

The room geometry is compiled in ODEON via a parameter (.par) file consisting of a list of vertices, given by co-ordinates in 3D space, and surfaces defined by a series of selected vertices. Surface dimensions are set such that they define the exact room topology of dimensions 2.9859 x 3.9857 x 2.4860 m represented by the 3D FDTD lattice

previously described. For consistency with the single node Kronecker delta excitation and RIR capture method applied in the FDTD schemes, both source and receiver directionalities are set to be omni-directional with frequency-independent amplitude. Material properties are assigned to all surfaces allowing the application of an appropriate absorption coefficient of $\alpha = 1 - R^2 = 0.591$ to all boundaries, which corresponds to the constant reflection coefficient applied in the FDTD model.

A total of 50000 rays are used to investigate the space for each RIR case. Mono RIRs are obtained by outputting 24-bit, B-Format .wav files and extracting the W-channel post-simulation.

5.1.1.3 2D FDTD Multiplane Model

The 3D cuboid space is represented by two cross-sectional planes orientated in the x-y and y-z planes. The resulting multiplane model may be envisaged as two intersecting, non-interacting rectilinear grids of pressure sampling instances with a common vector of intersection on which the source and receiver locations are defined. This arrangement, depicted in Figure 5.3, allows RIR recordings to be combined in a straight-forward manner through summation. The equal distance between source and receiver on each of the two planes ensures that the time of arrival and amplitude of the direct sound components are consistent in each scheme. Alternative arrangements involving arbitrary source/receiver positioning and/or an additional cross-sectional plane lead to the requirement of extrapolated sound excitation methods as examined in the studies documented later in this chapter. The pressure field in each planar scheme is excited using a soft-source Kronecker delta function and the two contributing responses are produced by recording and summing pressure fluctuations arising at the common receiver location. Additionally, LRS boundary conditions, consistent with those applied in the 3D FDTD model are implemented at the terminating edges of each scheme. Details of RIR post-processing and generation of the complete hybrid RIRs are given in section 5.2.1.

For 2D SRL FDTD schemes, the inter-nodal distance, h_{2D}, is subject to the following condition in order to achieve maximum spatial sampling resolution while maintaining numerical stability:

$$h_{2D} = \sqrt{2}ck \tag{5.2}$$

Therefore, by matching the simulation parameters c and k with those applied in the 3D FDTD model, a value of $h_{2D} = 0.011$ m is defined. Due to the difference that exists between h_{2D} and h_{3D}, small errors arise between model dimensions and source/receiver locations in the 3D and 2D multiplane model. The values of these errors, which are

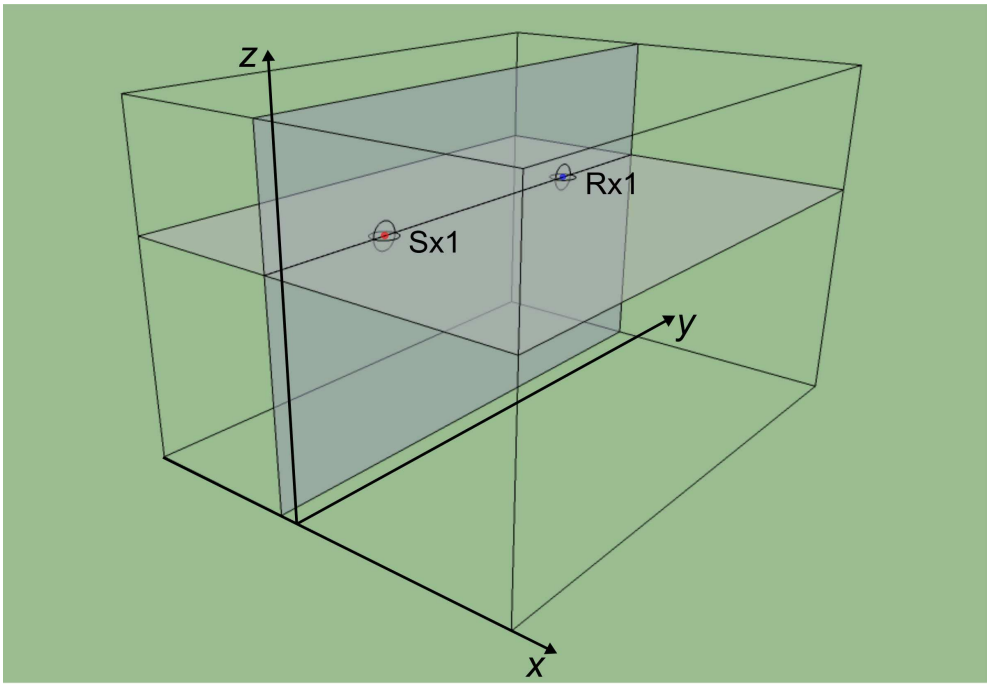

FIGURE 5.3: Arrangement of 2D planes in the multiplane model for source (Sx)/receiver (Rx) Case 1. Locations of the source and receiver, which share a common vector parallel to the y-axis, are shown. The cross-sectional planes occupy the x-y and y-z planes, intersecting on the common vector between source and receiver, providing a representation of the 3D space (given by the wire-diagram outline). This approach is replicated for the remaining 2 Sx/Rx cases.

considered to have insignificant impact on model accuracy, are detailed in Table 5.2. Also shown in the table is the difference between the number of pressure nodes required to represent the space in each FDTD model. The multiplane simulations require a total of only 179056 nodes per state of the modelled pressure field, of which two are required. This corresponds to a reduction of 11935744 sampling instances against the 3D FDTD model and, in turn, a memory saving of $\sim 98.5\%$. The impacts of this large reduction in the number of pressure nodes is reflected in simulation run-times for the multiplane model which are, on average, 15.5 s per second of RIR output. Again, the value of simulation time excludes time taken to post-process RIRs into their hybrid form. Compared to the 3D model, this translates to a run-time reduction of $\sim 99\%$. Hence, the multiplane modelling approach demonstrates significant computational savings as expected.

5.1.2 Hybridisation

A series of post-processing procedures must be applied to the simulated FDTD and ODEON RIRs in order to render them compatible for combination into final hybrid

MODEL/**CASE**	Source (x, y, z) (m)	Receiver (x, y, z) (m)
CASE 1		
3D FDTD/ODEON	(1.1889, 1.0133, 1.6213)	(1.1889, 2.9994, 1.6213)
2D Multiplane	(1.1914, 1.1049, 1.6216)	(1.1914, 3.0006, 1.6216)
Difference	(0.0025, 0.0016, 0.0003)	(0.0025, 0.0012, 0.0003)
CASE 2		
3D FDTD/ODEON	(0.4999, 0.4999, 0.4999)	(0.4999, 3.4993, 0.4999)
2D Multiplane	(0.4964, 0.4964, 0.4964)	(0.4964, 3.4970, 0.4964)
Difference	(0.0035, 0.0035, 0.0035)	(0.0035, 0.0023, 0.0035)
CASE 3		
3D FDTD/ODEON	(0.2027, 0.2027, 0.2027)	(0.2027, 3.7965, 0.2027)
2D Multiplane	(0.1986, 0.1986, 0.1986)	(0.1986, 3.7948, 0.1986)
Difference	(0.0041, 0.0041, 0.0041)	(0.0041, 0.0017, 0.0041)
	Dimensions (x, y, z) (m)	No. of Grid Points (FDTD)
ALL		
3D FDTD/ODEON	(2.9859, 3.9857, 2.4860)	12211480
2D Multiplane	(2.9895, 3.9824, 2.4821)	179056
Difference	(0.0036, 0.0033, 0.0039)	11935744

TABLE 5.2: Overview of the comparison between model dimensions and source/receiver locations. The difference between the required number of pressure nodes in the 3D and multiplane FDTD models is also provided.

RIR outputs. This collection of processes is detailed here with regards to matching both 3D and 2D multiplane FDTD model RIRs with the ODEON simulation outputs. The methods of hybridisation are applied consistently to each RIR case.

5.1.2.1 Calibrating 3D FDTD and Geometric RIRs

The initial concern with the 3D FDTD model RIRs is the impact of combining a soft-source excitation with a positive reflection coefficient applied to all boundaries. As discussed in section 4.3.1.1, these model characteristics together result in DC offset amplitude drift due to resonances at the 0 Hz solution of the discrete wave equation. Removal of the drifting effects may be achieved by use of a DC blocking filter. In this work, the DC blocking filter is defined as a 1^{st}-order high pass IIR filter with a cut-on frequency of 20 Hz with Butterworth coefficients. As such, the filter is effective in reducing the magnitude of low frequency components that contribute to amplitude drifting while having minimal impact on the audible bandwidth. In addition, it may be

demonstrated empirically that a single pass of the DC blocking filter removes drifting effects as observed in the time domain. However, a modal component centered on 0 Hz remains. This component is suppressed sufficiently by applying a second pass of the DC blocking filter, yielding a filtering process that is effectively 2^{nd}-order.

A second filtering process is required to reduce the FDTD RIR to the valid bandwidth. For an SRL scheme, the valid passband is defined as 0.196 F_s. However, dispersion effects can reduce the accuracy of the RIR in this valid passband. For this reason, a transition frequency of 1 kHz (0.023 F_s) is used. Hence, the upper frequency limit of the FDTD simulation, f_u is in accordance with the condition $f_u \leq 0.075\ F_s$ (derived in section 4.2) which ensures a dispersion error of $< 2\%$ for 3D SRL schemes. An 11^{th}-order Butterworth IIR low pass filter is applied to each RIR providing a sufficiently steep roll-off to suppress mid-high frequency components.

The final stage of processing the FDTD RIR into a state compatible with the ODEON results involves energy calibration. A method for ensuring consistent energy in the source excitation across both SRL FDTD and geometric models is discussed extensively in [81]. In brief, this method of energy calibration scales the amplitude of RIRs such that the acoustic energy of the direct sound component emitted by the sound source, E, at a distance of 1 m from the source location is unity for both models, giving:

$$E = \sum_{n=1}^{N} (p[n])^2 = 1 \tag{5.3}$$

where N is the number of discrete-time pressure samples $p[n]$ that contribute to the direct sound component. In geometric models, such as ODEON, it is trivial to scale RIRs such that this criterion is met. For a given geometric RIR, the direct sound component occupies a single time sample with amplitude p_{dir} at an arbitrary distance r from the sound source. An appropriate scaling factor may be derived from the following relationship between pressure amplitudes at two different distances from the source location:

$$p_{dir} = \frac{r_1}{r} p_1 \tag{5.4}$$

For the case of unity pressure, $p_1 = 1$, at a distance of 1m from the source, $r_1 = 1$, it can be stated that the pressure p_{dir} arising at an arbitrary distance from the source, r is:

$$p_{dir} = \frac{1}{r} \tag{5.5}$$

Therefore, a scaling constant K_{geo} that is calculated using the amplitude of the direct

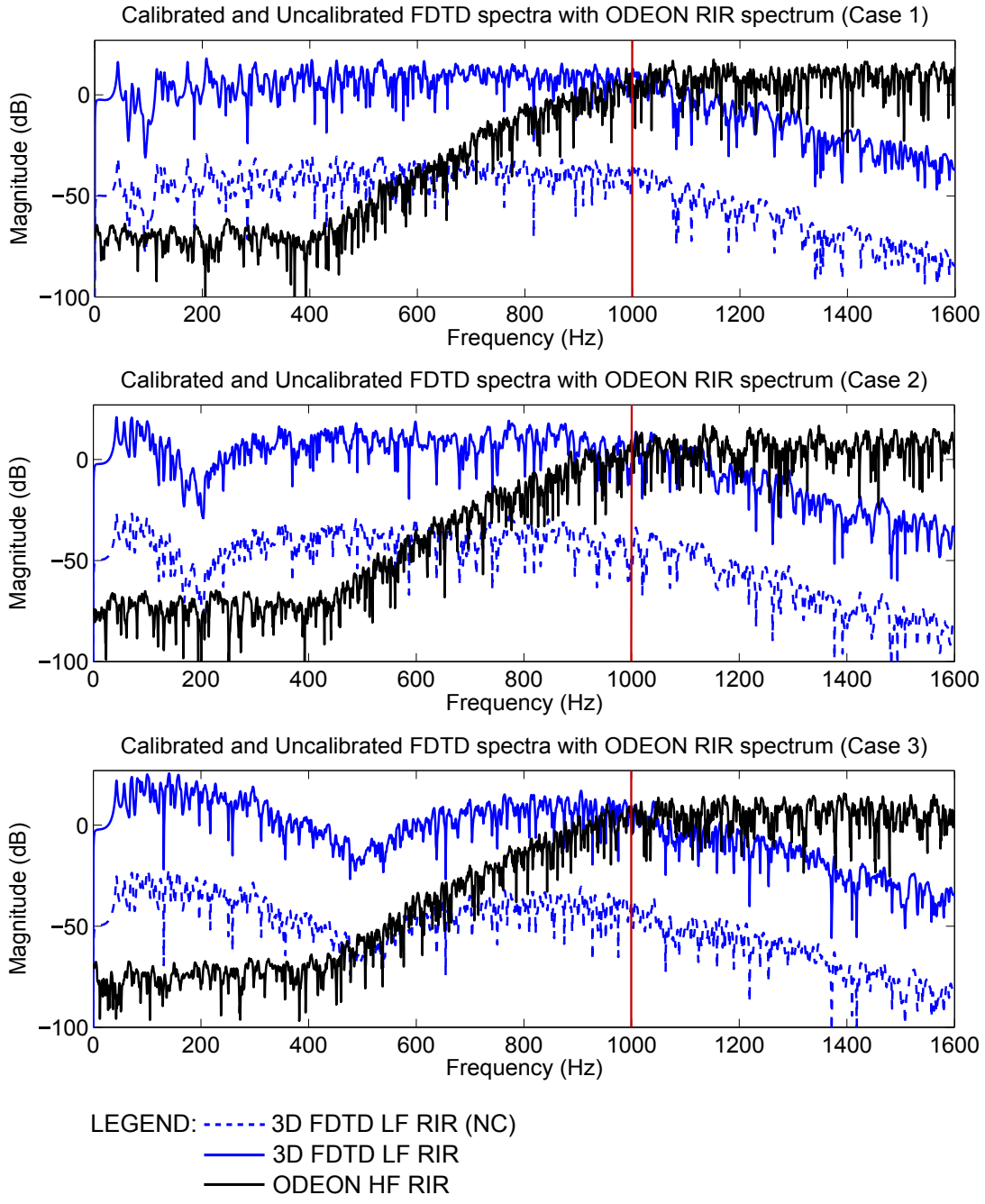

FIGURE 5.4: Magnitude spectra of calibrated (blue solid line) and non-calibrated, 'NC', (blue dashed line) 3D FDTD low frequency RIRs and respective ODEON RIRs (black solid line). The transition frequency of 1 kHz is shown in each RIR case by the vertical red solid line for reference.

sound component of each RIR may be defined as follows:

$$K_{geo} = \frac{1}{p_{dir} r} \tag{5.6}$$

Multiplication of output RIRs by the above constant preserves the unity energy condition required for the energy matching process. Each ODEON RIR is also subject to a high pass filtering process to remove the inaccurate low frequency results present below the selected transition frequency of 1 kHz. For consistency with the FDTD filtering process, the filter applied is an 11^{th}-order IIR Butterworth high pass.

The approach to calculating the total energy of the direct sound component in SRL FDTD schemes given in [81] is more involved. Briefly, the spread of energy throughout a SRL lattice is dependent on the inter-nodal distance and, therefore, the temporal sampling rate. As such, the total energy of the direct sound varies as F_s varies. The following expression, derived in [81], may be applied to calculate a scaling constant which ensures that $E = 1$ at a distance of 1m from the sound source in a 3D FDTD SRL scheme:

$$K_{3D,FDTD} = 5.437(F_s 10^{-3}) - 3.6347 \tag{5.7}$$

The use of this constant is considered accurate for $F_s > 11$ kHz and is, therefore, valid for application here. Each FDTD RIR is multiplied by $K_{3D,FDTD}$ to match the energy of the sound source to that of the ODEON model outputs. The impact of the energy calibration procedure are displayed in Figure 5.4. For each RIR case, the spectra of filtered ODEON and FDTD RIRs is shown. Note, prior to calibration, the FDTD RIRs consistently exhibit lower energy levels than those present in the ODEON counterparts highlighting the importance of the energy matching process. Post-calibration the agreement between the overall energy level of both contributing RIRs is shown to be significantly improved for all RIR cases indicating the successful implementation of the calibration process.

Once the energy calibration has been conducted, the two filtered and scaled RIRs may be combined through summation to render the final spectrally complete hybrid RIRs. Prior to summation, it is necessary to align both ODEON and FDTD simulation outputs such that the direct sound component in each occurs at the same discrete-time sample. An overview of the major hybridisation process components is presented as a flow diagram in Figure 5.5.

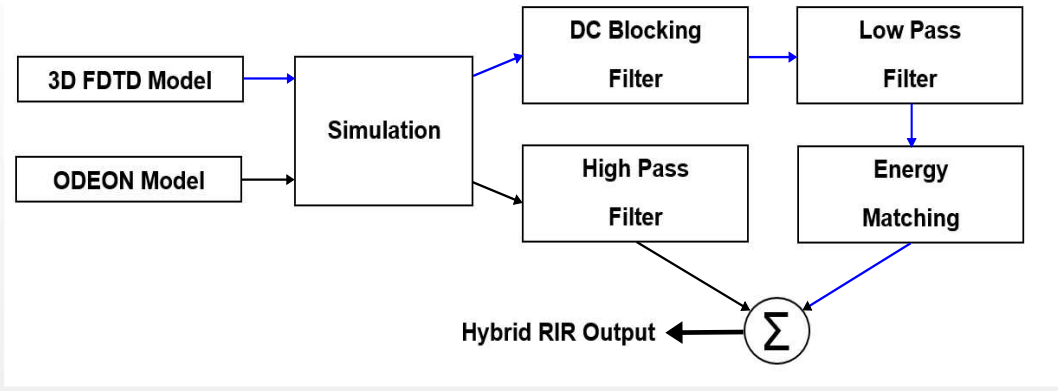

FIGURE 5.5: Summary flow diagram of the hybridisation process as conducted in the case of 3D FDTD/ODEON hybrid RIR rendering.

5.1.2.2 Calibrating 2D FDTD Multiplane and Geometric RIRs

As discussed in [31], the nature of energy decay in 2D FDTD simulations of the discrete wave equation is different from that of the 3D case. In turn, characteristics such as reverberation time and resonant frequency amplification are also expected to be different. As such, application of the same reflection coefficients to the 2D multiplane model as those used in the 3D models can lead to significant differences in the emulation of resonant behavior at low frequencies. In order to overcome this issue, for the purposes of this study, a simple coefficient matching procedure is devised. Firstly, it is assumed that the Norris-Eyring equation for 2D space, as noted in [31], offers a reasonable prediction of RT_{60}. This expression may be re-arranged for a given absorption coefficient α giving:

$$\alpha = 1 - e^{\frac{-\pi S ln(10^6)}{cLRT_{60}}} \tag{5.8}$$

where S is the area of the 2D scheme with total side length l and c is wave speed. Secondly, the RT_{60} of the 3D model may be measured to provide a numerical value that can be inserted into the above expression to yield a more appropriate absorption and, therefore, reflection coefficient. This coefficient can be calculated for each contributing cross-sectional scheme of the multiplane model. Although this is a rudimentary approach, it is suitable for the purposes of this pilot study as the reflection coefficients applied in each model are arbitrarily high in order to encourage resonant behavior.

Having simulated each 2D multiplane model yielding the 3 RIR case outputs, the results are then subject to similar filtering processes as those applied to the 3D FDTD RIRs. Each RIR is passed through the 2^{nd}-order DC blocking filter to remove DC amplitude drift and the low pass filter removing spectral components above 1 kHz. Additionally,

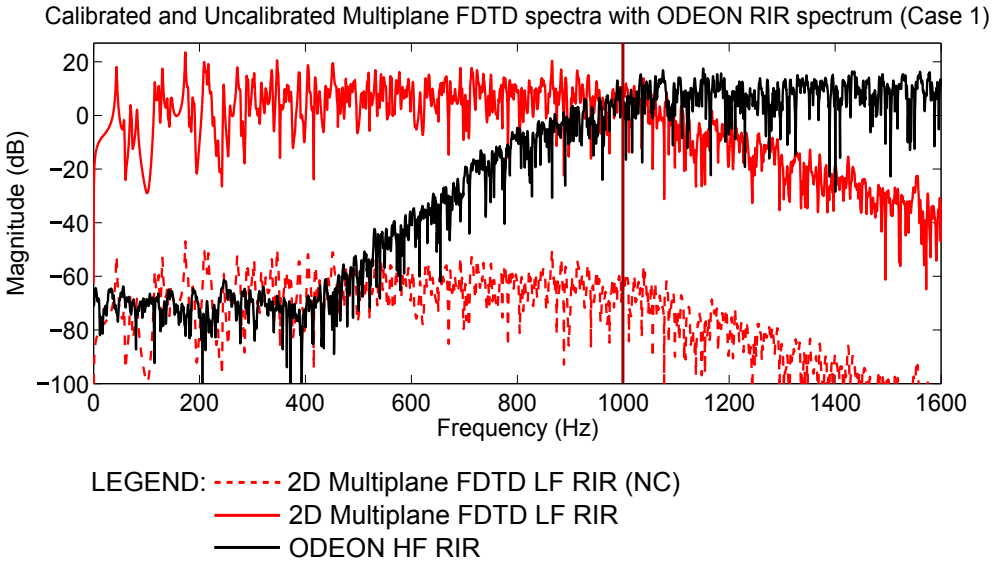

FIGURE 5.6: Magnitude spectra of calibrated (red solid line) and non-calibrated, 'NC', (red dashed line) 2D multiplane FDTD low frequency RIRs and respective ODEON RIRs (black solid line). The transition frequency of 1 kHz is denoted by the vertical red solid line for reference.

an afterglow removal filter (see section 4.3.1.2) is applied in order to reduce the effects of the afterglow phenomenon on RIR outputs.

The final calibration step involves the energy calibration between multiplane RIRs and those produced in ODEON. This procedure is conducted by matching the total energy present in the filtered multiplane RIR to that of the calibrated 3D RIR for each source/receiver case. The total energy $E_{3D,i}$ of each 3D FDTD model RIR $p_{3D,i}[n]$ can be calculated using:

$$E_{3D,i} = \sum_{n=1}^{N} (p_{3D,i}[n])^2 \tag{5.9}$$

where, N is the total number of discrete time samples n recorded in the RIR and subscript i refers to the source/receiver case (i.e i = 1,2,3). The total energy of the corresponding multiplane RIR $E_{MP,i}$ may be calculated by similar means. This leads to the definition of an energy calibration constant K_{MP}:

$$K_{MP,i} = \sqrt{\frac{E_{3D,i}}{E_{MP,i}}} \tag{5.10}$$

where subscript MP denotes the multiplane RIR energy for source receiver case i. Multiplying each multiplane RIR by $K_{MP,i}$ equates the total energy present in each case with that of the corresponding 3D RIR. By extension, the low frequency multiplane RIRs are considered to be compatible with the ODEON mid-high frequency contributions in terms of energy balance either side of the transition frequency. Figure 5.6

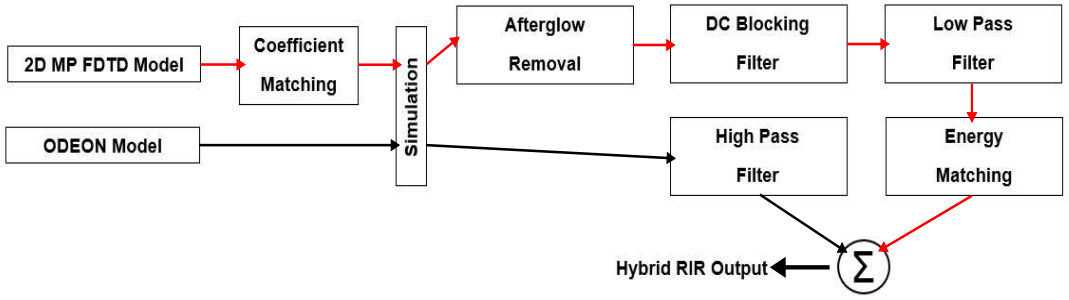

FIGURE 5.7: Summary flow diagram of the hybridisation process as conducted in the case of 2D Multiplane FDTD/ODEON hybrid RIR rendering.

demonstrates the significant improvement observed between energy level agreement of uncalibrated and calibrated multiplane RIR with the ODEON RIR for source/receiver Case 1. Such a result is also observed for RIR Cases 2 and 3.

Having processed each multiplane RIR into a form compatible with the corresponding ODEON signals, outputs from both models are time-aligned to match the temporal position of the direct sound component. The total hybrid RIR is then rendered by summing both low frequency and high frequency RIRs together. A summary flow diagram of the multiplane calibration/hybridisation processes is displayed in Figure 5.7 for reference.

5.1.3 Results

Objective results pertinent to the scope of this pilot study, presented here, are recovered from the RIRs produced by each of the modelling methods described previously. Particular attention is given to the nature of low frequency sound simulation to demonstrate the level of agreement between multiplane FDTD and 3D FDTD RIR spectra. Also observed are the representations of the low frequency soundfield as rendered by ODEON. This is required in order to compare the accuracy of geometric and multiplane models in the lower frequency region. Note that it is assumed that the 3D FDTD model produces the most accurate results given the inherent preservation of 3D wave characteristics. As such, the aforementioned levels of accuracy of the remaining two models is a comparative measure against the ideal 3D FDTD RIR cases.

5.1.3.1 Low Frequency Analysis

The low frequency magnitude spectra of each model RIR is calculated via the Fourier transform and examined in three ways. Firstly, visual inspection and comparison of low frequency spectra with reference to the 3D FDTD model results and theoretic modal values provides initial insight into the agreement of simulated low frequency characteristics. This is conducted for two frequency regions $f_{r1} = [0 : 380]$ Hz and $f_{r2} = [0 : 1000]$ Hz. The first frequency region accounts for simulation frequencies up to an approximate value for the Schroeder frequency F_c, below which resonant behaviour is most prominent. The second frequency region provides an overall description of the low frequency magnitude trends up to the transition frequency.

Secondly, the frequency response assurance criterion (FRAC), as documented in [157, 158], is used to provide a single numeric value that represents the agreement between magnitude spectra. For two frequency domain signals $X(f)$ and $Y(f)$, the FRAC calculation seeks to establish the level to which both signals are linearly related by means of the following equation, after [157]:

$$FRAC = \frac{\left| \sum_{f=f_1}^{f_2} X(f)Y^*(f) \right|^2}{\sum_{f=f_1}^{f_2} X(f)X^*(f) \sum_{f=f_1}^{f_2} Y(f)Y^*(f)} \tag{5.11}$$

assuming,

$$\frac{\left| \sum_{f=f_1}^{f_2} X(f)X^*(f) \right|^2}{\sum_{f=f_1}^{f_2} X(f)X^*(f) \sum_{f=f_1}^{f_2} X(f)X^*(f)} = \frac{\left| \sum_{f=f_1}^{f_2} Y(f)Y^*(f) \right|^2}{\sum_{f=f_1}^{f_2} Y(f)Y^*(f) \sum_{f=f_1}^{f_2} Y(f)Y^*(f)} = 1 \tag{5.12}$$

where * denotes the complex conjugate and f_1 and f_2 denote lower and upper frequency bounds respectively, thus setting the FRAC analysis bandwidth. If the two frequency domain signals, $X(f)$ and $Y(f)$ are related in an exact linear manner (i.e. they may differ only in amplitude in the time domain), then the FRAC will return a unity value. However, as the correlation between the two signals degrades, the FRAC value decreases with a minimum value of 0 signifying complete absence of any linear relationship. This measurement is applied in order to compare the extent to which the ODEON and multiplane FDTD RIRs represent the 3D FDTD RIRs in the frequency regions f_{r1} and f_{r2}.

Finally, the magnitude spectra resulting in each RIR case may be compared by means of the magnitude similarity measures as proposed in [158]. Briefly, a new signal $H(f)$ is defined to express the relationship between two arbitrary frequency domain signals,

$X(f)$ and $Y(f)$, such that:

$$\big|H(f)\big| = \left|\frac{Y(f)}{X(f)}\right| \tag{5.13}$$

In this way, $|H(f)|$ yields the ratio of magnitude spectra for the signals. Therefore, if $X(f)$ and $Y(f)$ are the same, $H(f) = 1$. Following [158], the logarithmic ratio, $|20\log_{10}(H(f))|$ dB, is deemed more straightforward to interpret. Use of logarithm ensures that the magnitude similarity value resides between 0 and -∞. This, in turn, may be translated into a magnitude similarity value, $MS(f)$, via the following equation, after [158]:

$$MS(f) = 1 - \tanh\left(\frac{\ln 3}{2}.\frac{|20\log_{10}|H(f)||}{20}\right) \tag{5.14}$$

Hence, for a given frequency f, $MS(f)$ takes on a value in the range $[0:1]$ where 0 represents a large difference in magnitude values $|X(f)|$ and $|Y(f)|$ and 1 signifies no difference. Furthermore, for a given frequency range $[f_1 : f_2]$ Hz an overall representation of the agreement between magnitude spectra can be calculated by integrating the function $MS(f)$ over frequency:

$$MI = \frac{1}{f_2 - f_1}\int_{f_1}^{f_2} MS(f)df \tag{5.15}$$

where the coefficient outside the integral serves to scale MI to the range of $[0:1]$. In combination with the FRAC measurement, MS and MI offer useful insight into the agreement of spectral components present in simulated RIRs.

Note that the ODEON RIRs analysed are calibrated in terms of energy, but not subject to the high pass filtering stage for the purposes of all comparisons conducted. For clarity, each RIR case is examined in turn in the following.

RIR Case 1:

Low frequency spectra of the modal region f_{r1} for each model simulation are shown in Figure 5.8. Modal values calculated using (2.76), signified by the vertical dashed lines, are provided for reference. As an initial analysis step, the correctness of the FDTD models may be assessed by examining the position of resonant peaks relative to the theoretic modal frequency values. It is observed, in the case of the 3D FDTD RIR, that a strong relationship exists between modal values and simulated resonances, thus verifying the correctness of the model.

The level of agreement exhibited between the 2D multiplane and 3D FDTD RIR spectra is significantly high for the first five modal values up to a frequency of $\sim$ 80 Hz. However, the multiplane model is not capable of capturing tangential modes in the x-z plane or oblique modes (denoted by the green dashed lines) due to the nature of

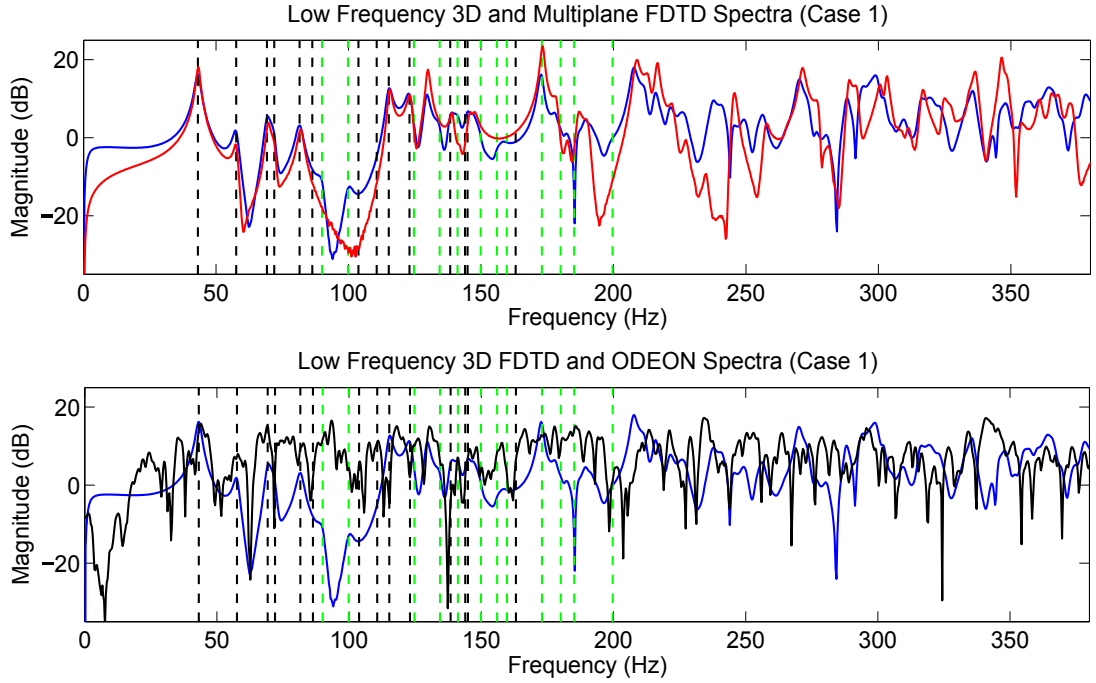

FIGURE 5.8: Low frequency spectra of 3D FDTD (blue solid line), 2D multiplane FDTD (red solid line) and ODEON (black solid line) RIRs (Case 1). Theoretic modal values are shown by the vertical dashed lines. The modal values plotted in green highlight modes that are either tangential in the x-z plane or oblique.

its construction. This is particularly evident upon comparison of resonant components simulated in the 3D model to those rendered in the multiplane model over $f \approx [90 : 100]$ Hz. This frequency region contains the first x-z tangential and oblique modes which are present in the 3D FDTD RIR, but absent in the multiplane response. Such results also exist for higher order modes of this kind. The presence of a resonant peak at oblique or x-z tangential mode values in the multiplane spectrum is the result of a coincident mode occurring across either of the two cross-sectional planes. Furthermore, it is apparent that resonances that are not active in the 3D model, due to the location of sound source and receiver, are also suppressed in the multiplane spectrum. This further highlights the agreement between both RIR spectra and demonstrates the ability of the multiplane model to produce a reasonable representation of the low frequency sound behaviour occurring in the space. In the upper half of the modal region, $f > 200$ Hz, the level of agreement between 3D and multiplane RIR spectra decreases somewhat. As frequency increases, the coincidence of multiple modal values increases in the 3D model. The same is true for the multiplane model, however the absence of higher order x-z tangential and oblique modes acts to reduce the similarity of 3D and multiplane spectra towards the upper limit of the modal region. Nonetheless, the position of several resonant peaks coincide across both modelling approaches as frequency increases.

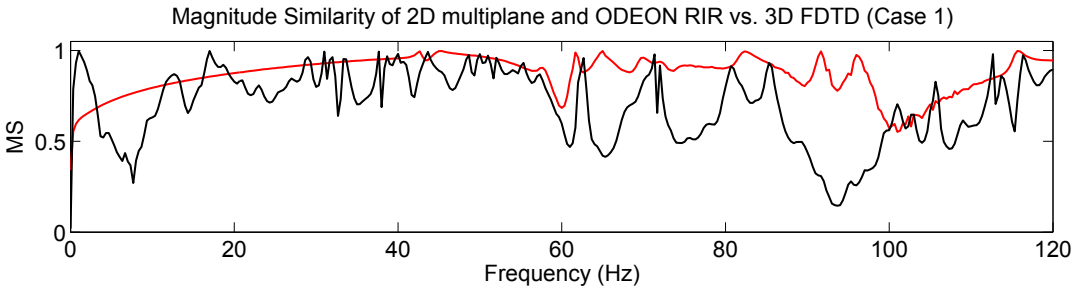

FIGURE 5.9: Magnitude similarity for 2D multiplane FDTD (red) and ODEON (black) spectra compared to 3D FDTD spectrum resulting for RIR Case 1 in the modal region f_{r1}. Values are calculated using (5.14).

Also depicted in Figure 5.8 (bottom) is the comparison of ODEON and 3D FDTD low frequency RIR spectra. It may be observed that the overall energy level trends are similar for both models with the exception of the low frequency region $f \approx$ [85:115] Hz. However, the presence of distinct modal peak values is lacking in the ODEON RIR spectrum as expected due to the inability of the modelling approach to emulate resonances. Hence, it may be determined by visual inspection alone that the 2D multiplane RIR exhibits a higher level of agreement with the 3D FDTD model in terms of low frequency spectral characteristics. This finding is further supported through analysis of magnitude similarity, MS, which compare the ODEON and 2D multiplane spectra to that of the 3D FDTD model. The MS functions for RIR Case 1 are provided in Figure 5.9. As shown, the 2D multiplane spectrum possesses a more consistent agreement with the 3D results than the ODEON RIR spectrum. This is observed through comparison of the MS values calculated for the multiplane (red curve) and ODEON (black curve) RIRs over the frequency range $f =$ [0:120] Hz. Examination of the MS curve for the ODEON RIR reveals that the similarity between geometric and 3D FDTD spectra is highly variable across the low frequency bandwidth. Conversely, the 2D multiplane MS curve maintains a higher and more consistent agreement with the 3D RIR spectrum due to the preservation of resonant characteristics. The reduction in multiplane MS values around 100 Hz is a notable result as this directly correlates with the absence of the x-z tangential and oblique modes from the multiplane RIR previously described. Furthermore, the significant difference between ODEON and 3D FDTD spectra in the region $f \approx$ [85:115] Hz is also represented by the low MS values returned over this bandwidth.

Integrating each MS curve over the modal frequency range yields the magnitude similarity indexes MI as per (5.15) provided in Table 5.3. FRAC values are also given as a measure of the correlation of ODEON and multiplane RIR magnitude spectra to the 3D FDTD RIR spectrum. Both FRAC and MI provide measurement of similarity between time varying signals through analysis of their frequency domain representations.

Case 1 (vs. 3D FDTD), $f = f_{r1}$	FRAC	MI
2D Multiplane FDTD	0.488	0.853
ODEON	0.006	0.785

TABLE 5.3: FRAC and MI values calculated to compare the agreement of multiplane/ODEON RIRs with the 3D FDTD RIR (Case 1) in the modal region f_{r1}.

However, the FRAC quantity assesses the covariance of the signals (thus incorporating phase properties), whereas the MI value provides a measure of the extent to which a linear relationship exists between magnitude spectra only. Measurement of these parameters show that the multiplane RIR is more comparable to the 3D counterpart than the ODEON RIR. This is expected given previous analyses of resulting spectra and MS properties and further supports the fact that the multiplane model is capable of emulating prominent modal components where the geometric approach is not.

Case 1 (vs. 3D FDTD), $f = f_{r2}$	FRAC	MI
2D Multiplane FDTD	0.113	0.844
ODEON	0.001	0.830

TABLE 5.4: FRAC and MI values calculated to compare the agreement of multiplane/ODEON RIRs with the 3D FDTD RIR (Case 1) in the low frequency simulation bandwidth f_{r2}.

FRAC and MI measures for the full low frequency simulation bandwidth f_{r2} up to the transition frequency of 1 kHz are noted in Table 5.4. These results provide an important comparison against those returned for the modal frequency region. Firstly, it is noted that FRAC values for both multiplane and ODEON RIRs decrease significantly compared to those given for f_{r1}. Hence as the frequency range under examination increases, the variance between simulated RIR signals also increases. This is due to the accurate simulation of wave motion provided by the 3D FDTD model which preserves a more complete representation of enclosed sound propagation compared to the alternative modelling techniques. In terms of the multiplane model, only a subset of resonant frequencies are successfully emulated. Therefore, as the number of modal components occurring in 3D space increases with frequency, the representation provided by the subset decreases in accuracy leading to a lower FRAC value. The same notion may be extended to rationalise the decrease in FRAC value calculated for the ODEON RIR in which modal components are largely omitted.

A different trend is observed between the MI returned for each frequency region. The overall magnitude similarity is shown to increase significantly for the ODEON RIR when considering the wider frequency region, whereas the similarity decreases slightly

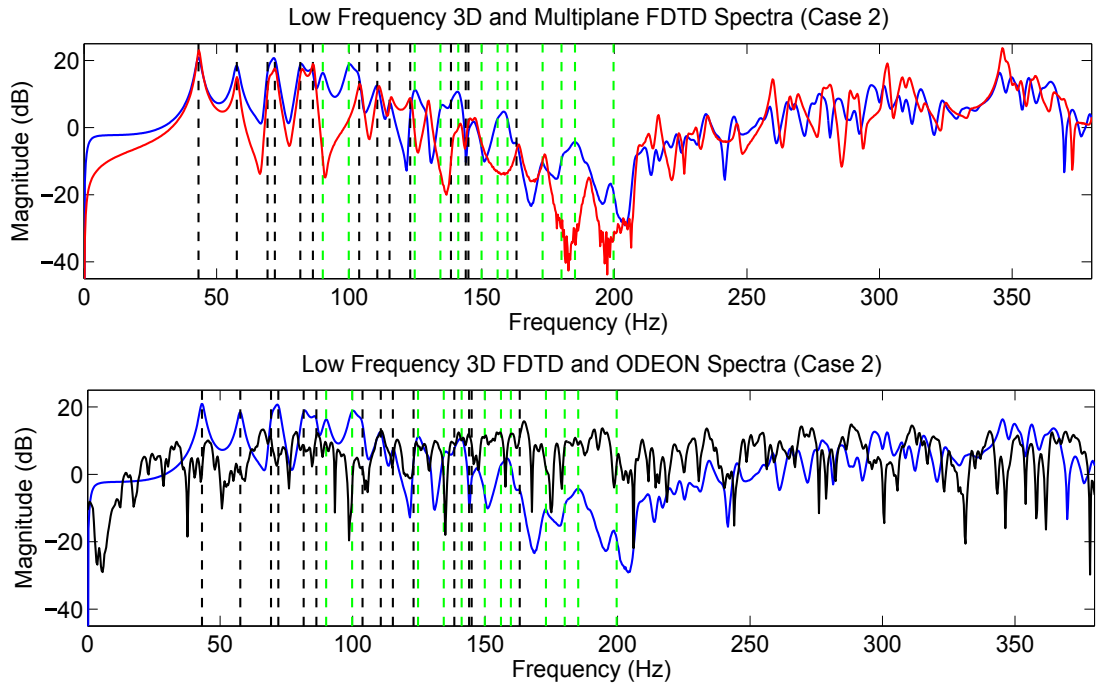

FIGURE 5.10: Low frequency spectra of 3D FDTD (blue solid line), 2D multiplane FDTD (red solid line) and ODEON (black solid line) RIRs (Case 2). Theoretic modal values are shown by the vertical dashed lines. The modal values plotted in green highlight modes that are either tangential in the x-z plane or oblique.

for the multiplane RIR. This exposes the fact that the MI measure becomes less sensitive to differences between spectra as the frequency region under analysis widens. Hence, the level of similarity between spectra local to a sub-region of the bandwidth (i.e. the modal region) will no longer contribute significantly to the MI value. This concept explains the converse trends noted in FRAC and MI values across the examination of small (f_{r1}) and larger (f_{r2}) bandwidths. However, the use of MI over the region $f = f_{r2}$ is justified for cases where notable deviations between overall RIR spectra magnitude levels are present over a large frequency region (see RIR Case 3).

Overall, for this RIR case, it may be stated that the multiplane RIR shares a higher level of similarity to the 3D RIR reference compared to the ODEON RIR. This has been demonstrated by all means of analysis applied and is particularly apparent in the modal frequency region where the multiplane model is shown to produce a good approximation to the 3D low frequency soundfield.

RIR Case 2:

The second RIR case considers sound source and receiver locations which are positioned towards the corners of the cuboid model. As such, the magnitude of low frequency components is expected to increase compared to RIR Case 1 as resonant frequency waves collect and increase in amplitude towards edges and corners. Figure 5.10 displays

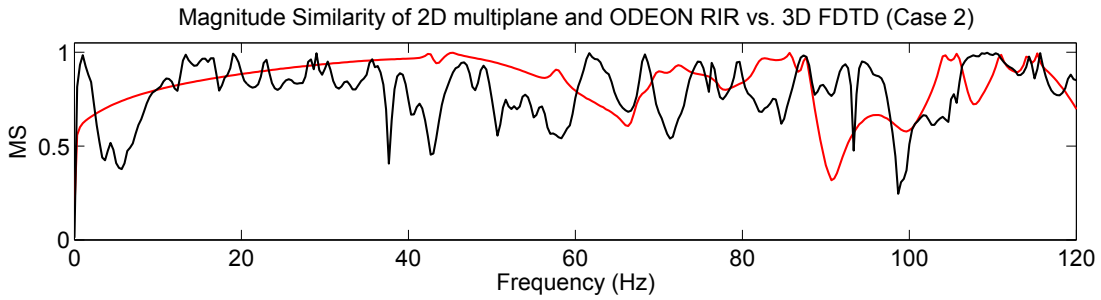

FIGURE 5.11: Magnitude similarity for 2D multiplane FDTD (red) and ODEON (black) spectra compared to 3D FDTD spectrum resulting for RIR Case 2 in the modal region f_{r1}.

the 3D FDTD (blue) and 2D multiplane (red) Case 2 RIR spectra (top). As shown, the expected increase in low frequency modal component magnitudes, with those simulated for RIR Case 1, is demonstrated by both modelling approaches. The agreement between modal response up to $f \approx 80$ Hz is shown to be high with the multiplane model correctly emulating low order resonance characteristics. This agreement degrades somewhat in frequency regions containing x-z tangential and oblique modes. In particular, the omission of these modal components in the multiplane RIR for $f = [180{:}200]$ Hz results in a significant drop in overall magnitude compared to the 3D RIR spectrum. These results are consistent with those produced in RIR Case 1. As frequency increases towards the upper limit of the modal region f_{r1}, the general trend of magnitude values in each spectra are very similar.

Conversely, the low frequency spectrum recovered from the ODEON RIR (see Figure 5.10 (bottom)), denoted by the black curve, displays little obvious agreement with the 3D RIR spectrum. Specifically, the prominent low frequency modes and increase in overall magnitude in the region $f = [40{:}100]$ Hz are not represented. This result is demonstrated in the MS curve displayed in Figure 5.11. As shown, the magnitude similarity between ODEON and 3D RIRs (black curve) is highly variable in this frequency range and is generally lower than the more consistent multiplane spectrum similarity (red curve). However, the multiplane RIR MS curve reveals the extent to which the similarity reduces for the frequency region where low order x-z tangential and oblique modes are expected ($f = [90{:}100]$ Hz). Over this frequency range, the MS values returned for the multiplane RIR reduce significantly reaching values that are notably less than those given from the ODEON RIR. This, in turn, impacts on the FRAC and MI measures calculated for this RIR case which are provided in Table 5.5.

Both FRAC and MI values for the modal region of multiplane RIR Case 2 are lower than those calculated for Case 1. This suggests that the increase in energy in modal components that are not simulated in the multiplane model has an impact on the similarity measure and correlation between 3D and multiplane RIRs and causes the

Case 2 (vs. 3D FDTD), $f = f_{r1}$	FRAC	MI
2D Multiplane FDTD	0.397	0.834
ODEON	0.005	0.761

TABLE 5.5: FRAC and MI values calculated to compare the agreement of multiplane/ODEON RIRs with the 3D FDTD RIR (Case 2) in the modal region f_{r1}.

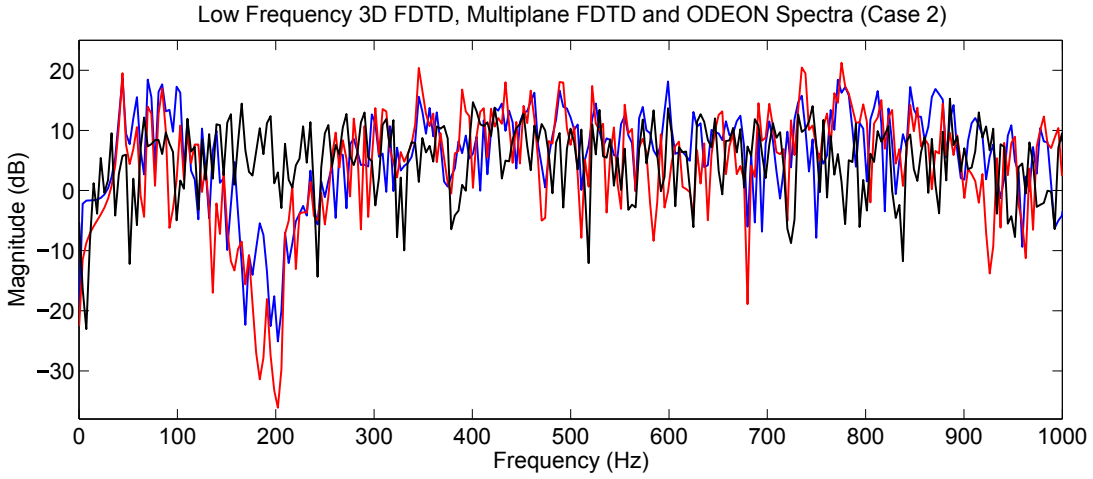

FIGURE 5.12: RIR spectra for 3D FDTD (blue), multiplane FDTD (red) and ODEON (black) for RIR Case 2 over the low frequency simulation bandwidth f_{r2} up to 1 kHz.

decrease in MI and FRAC measures. However, both measures remain higher than those calculated for the ODEON RIR demonstrating the positive effect of the simulation of a subset of modal activity in the multiplane model. Compared to a geometric approach with inherently low accuracy in the modal region, the multiplane model exhibits a notable improvement. Visual inspection of the simulated RIR spectra in the wider bandwidth f_{r2} up to the transition frequency, displayed in Figure 5.12, provides further insight on how overall magnitude levels vary between models. As shown, the ODEON RIR spectra exhibits a lower range of magnitude levels, within the range of ∼[-15:12] dB, across this bandwidth. Conversely, the FDTD RIRs both demonstrate more variable magnitude level trends. The agreement of these trends for 3D and multiplane RIRs has been discussed previously for the modal frequency region. Additionally, the 3D RIR possesses several significant spectral peaks at higher frequencies in the range of ∼[350:750] Hz. A number of these peaks, which are not present in the ODEON RIR spectrum, are shown to be successfully represented in the multiplane model. Of particular interest are the coincident spectral peaks arising in both 3D and multiplane RIRs located at approximately 350, 470, 720 and 760 Hz. However, it is noted that the same level of agreement between RIRs for both FDTD models observed in the modal region is not maintained as frequency increases. As with the findings reported

for RIR Case 1, this is due to the greater increase in the number of modal components modelled by the 3D scheme compared to the multiplane model as frequency increases. As displayed in Table 5.6, the FRAC and MI values calculated for the multiplane RIR

Case 2 (vs. 3D FDTD), $f = f_{r2}$	FRAC	MI
2D Multiplane FDTD	0.306	0.839
ODEON	0.001	0.803

TABLE 5.6: FRAC and MI values calculated to compare the agreement of multiplane/ODEON RIRs with the 3D FDTD RIR (Case 2) in the low frequency simulation bandwidth f_{r2}.

are greater than those calculated for the ODEON RIR for f_{r2}. This again reinforces the claim that the multiplane model shares a higher level of similarity with the 3D FDTD model than the ODEON model across the full low frequency simulation bandwidth. Interestingly, the MI value for the multiplane RIR increases slightly compared to the value calculated from the modal region. The same trend in MI is observed for the ODEON RIRs when comparing values from each analysis range. As with the ODEON Case 1 RIR MI value trend, this may be attributed to the tendency of MI to increase as the frequency range used for calculation increases. Hence, the levels of agreement and disagreement between resulting spectra in the modal region contribute less to the MI value calculated for the wider bandwidth f_{r2}.

The results presented here, in conjunction with those detailed for RIR Case 1, demonstrate that the multiplane model is capable of simulating low frequency sound behaviour to a greater extent than the geometric model as the sound source and receiver locations are varied. This result is exemplified in RIR Case 2 by the level of agreement between significant 3D and multiplane RIR spectral attributes in both f_{r1} and f_{r2} which are largely omitted in the corresponding ODEON RIR.

RIR Case 3:

The final RIR case in this pilot study concerns an arrangement in which the target acoustic field is excited and recorded in close proximity to opposite corners of the cuboid in one axial direction. Hence, the trend in the increase of low frequency energy as the source and receiver approach the corners (as demonstrated by comparing the spectra for RIR Cases 1 and 2) is expected to be maintained. This is observed in Figure 5.13 which displays the magnitude spectra of 3D and 2D multiplane FDTD and ODEON RIRs across the modal region f_{r1}. As with RIR Cases 1 and 2, a high level of agreement is observed between the 3D (blue) and 2D multiplane (red) RIR spectra for low order modes in the range $f = [0{:}\sim 80]$ Hz. Additionally, the 2D multiplane RIR is demonstrated to provide a good representation of the overall magnitude trends

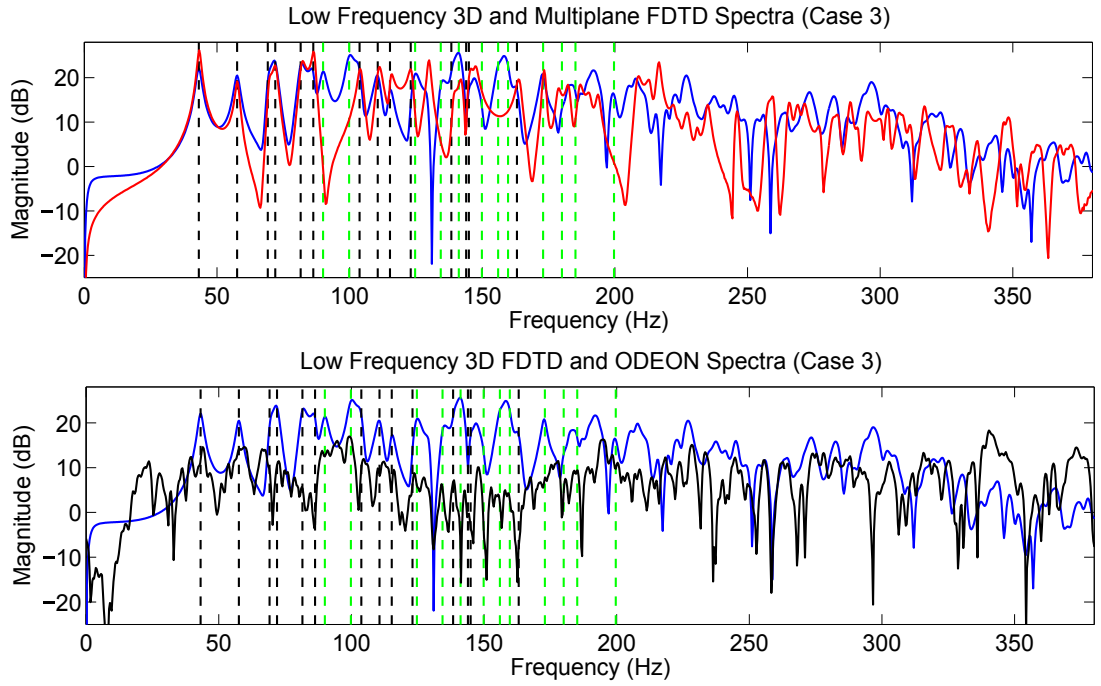

FIGURE 5.13: Low frequency spectra of 3D FDTD (blue solid line), 2D multiplane FDTD (red solid line) and ODEON (black solid line) RIRs (Case 3). Theoretic modal values are shown by the vertical dashed lines. The modal values plotted in green highlight modes that are either tangential in the x-z plane or oblique.

possessed by the 3D RIR spectra. This is most apparent when considering the level of spectral peaks across the range $f \approx [45{:}175]$ Hz and the gradual reduction in magnitude levels towards the upper limit of the modal region resulting from both models. However, the omission of prominent x-z tangential and oblique modes from the 2D multiplane RIR again acts to emphasise the inherent limitations of the two plane arrangement implemented in this study. The bottom graph of Figure 5.13 provides insight into the similarity of ODEON (black) and 3D FDTD (blue) RIR spectra. It is demonstrated that the overall magnitude trends common to both 3D and 2D multiplane FDTD RIRs are not present in the ODEON RIR spectra. This is shown by the notably lower magnitude levels of the ODEON spectral across $f \approx [60{:}200]$ Hz. Additionally, the decrease in magnitude levels towards the upper limit of f_{r1} observed in the 3D and 2D multiplane RIRs is not evident in the ODEON RIR spectrum. However, in this instance of source and receiver positioning, the geometric approach implemented in ODEON is shown to produce a higher level of similarity with the 3D FDTD RIR in terms of the coincidence of a subset of modal frequencies. This is contrast to results produced for the previous two measurement cases. For example, the ODEON RIR contains resonant peaks that coincide with theoretical modal frequency values, particularly in the range $f \approx [45{:}150]$ Hz. While this result constitutes an improvement in the correctness of spectral characteristics of the ODEON RIR compared to previous cases, the same

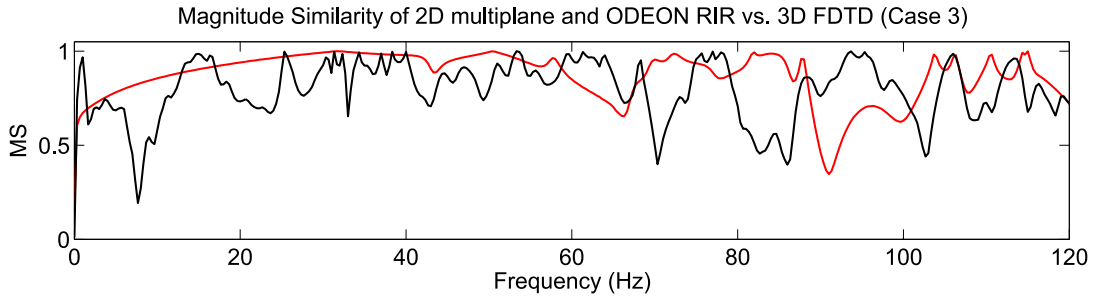

FIGURE 5.14: Magnitude similarity for 2D multiplane FDTD (red) and ODEON (black) spectra compared to 3D FDTD spectrum resulting for RIR Case 3 in the modal region f_{r1}.

improvement is not apparent elsewhere in the modal frequency range. The observations drawn from visual comparison of Case 3 RIRs are reflected in the calculation of MS, FRAC and MI values. Figure 5.14 displays the MS curves comparing the similarity of 2D multiplane (red) and ODEON RIRs (black) to the 3D FDTD RIR. As with results documented for RIR Cases 1 and 2, the multiplane RIR demonstrates a more consistent agreement with the 3D RIR in terms of low frequency spectral characteristics (f = [0:60] Hz). Again, this agreement reduces for the frequency region containing the first x-z tangential and oblique modes represented in the 3D FDTD model (f = [85:100] Hz). The MI and $FRAC$ values values calculated over the modal region are noted in Table 5.7.

Case 3 (vs. 3D FDTD), $f = f_{r1}$	FRAC	MI
2D Multiplane FDTD	0.260	0.830
ODEON	0.035	0.785

TABLE 5.7: FRAC and MI values calculated to compare the agreement of multiplane/ODEON RIRs with the 3D FDTD RIR (Case 3) in the modal region f_{r1}.

A notable improvement is observed in the FRAC value calculated for the ODEON RIR compared with Cases 1 and 2. This may be attributed to the increase in the similarity of ODEON and 3D FDTD RIR spectra for a subset of modal frequencies discussed previously. However, this claim is not supported by the corresponding MI value which is shown to be consistent with results for Cases 1 and 2. Recalling the definition of FRAC (5.11), it is apparent that this quantity accounts for the correlation of phase. Hence, the increase in ODEON RIR FRAC value may be arising due to a higher agreement in the phase of spectral components present within 3D FDTD and ODEON RIRs. However, seeking conclusive assumptions that relate this result to attributes of the geometric modelling approach implemented by ODEON is beyond the scope of this work. For the purposes of this pilot study it is sufficient to demonstrate

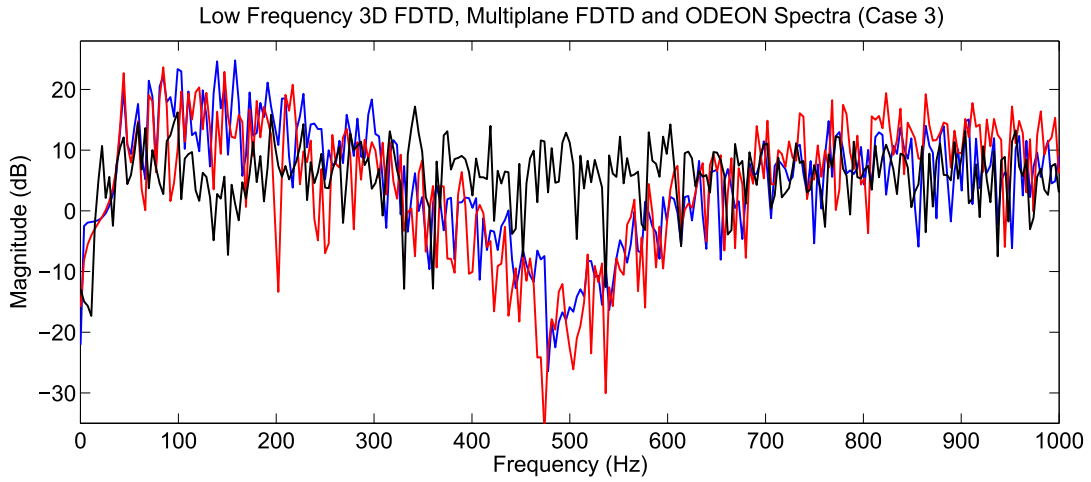

FIGURE 5.15: RIR spectra for 3D FDTD (blue), multiplane FDTD (red) and ODEON (black) for RIR Case 2 over the low frequency simulation bandwidth f_{r2} up to 1 kHz.

that the FRAC and MI values calculated from 2D multiplane RIRs are consistently greater than those calculated using the geometric approach. This finding is further supported by the values shown in Table 5.7.

Visual inspection of the low frequency simulation region, f_{r2}, further demonstrates the difference between magnitude level trends present in FDTD and ODEON spectra initially observed in the modal region. As shown in Figure 5.15, there is a notable reduction in magnitude level in the 3D FDTD RIR in the region $f \approx [400{:}600]$ Hz which also results in the multiplane RIR. Conversely, the ODEON RIR does not emulate this significant spectral feature. This, in turn, leads to FRAC and MI values for the ODEON RIR (Table 5.8) which are less than those calculated for the modal region f_{r1}. Furthermore, while the 2D multiplane RIR FRAC and MI values continue to remain

Case 3 (vs. 3D FDTD), $f = f_{r2}$	FRAC	MI
2D Multiplane FDTD	0.237	0.839
ODEON	0.014	0.782

TABLE 5.8: FRAC and MI values calculated to compare the agreement of multiplane/ODEON RIRs with the 3D FDTD RIR (Case 3) in the low frequency simulation bandwidth f_{r2}.

higher than those of the ODEON RIR, the FRAC values are lower than those recorded in RIR Cases 1 and 2. Indeed, there appears to be a trend by which the FRAC value decreases as the source and receiver locations approach the corners of the modelled space while the MI values are mostly consistent for all modelling scenarios. This final finding further demonstrates the sensitivity of the FRAC value to the correlation of RIR characteristics beyond magnitude similarity alone. Hence, the correlation between

the phase of recorded frequency responses affects the value calculated for each case. Due to the fact that the 2D multiplane FDTD model represents only a subset of valid reflection paths within the modelled domain, the difference between temporal attributes of 3D and 2D multiplane RIRs will directly impact the correlation of phase characteristics. Considering the FRAC values recorded in all three RIR cases, this issue has been shown to be exacerbated as the source and receiver approach the corners of the space. Hence, it may be concluded that the number of reflection paths represented by the multiplane model is not sufficient to fully represent more complex interference effects occurring in 3D which become increasingly significant as source and/or receiver approach a location in which resonances coincide and accumulate. Regardless, this limitation of the multiplane model, which is investigated in more depth in the following sections, is shown to have little impact on the similarity of 3D and multiplane RIR magnitude spectra characteristics.

5.1.3.2 Early Reflection Characteristics

Given the arrangement of sound source and receiver in each RIR case, in which both source and receiver are coplanar in the x-y and y-z planes, a subset of the early reflection components observed in the 3D model are expected to be represented in the multiplane equivalent. Figure 5.16 demonstrates the agreement between temporal positioning of early reflections, for the time region $t = [0 : 50]$ ms, as rendered by the 3D and 2D multiplane FDTD simulations for each RIR case. In each case, squared pressure amplitude values are shown to enable clear comparison between simulation results. Considering the RIRs rendered for RIR Case 1 (top), it is observed that the temporal positioning of early reflections is similar in both models for the first 20 ms of simulation. However, for the majority of these reflections, there is a significant difference between the squared pressure amplitude values calculated in each model. The primary reason for this result is the nature of the Green's functions for 2D and 3D wave equations which inherently give rise to different energy spreading characteristics. Although afterglow removal filtering procedure described previously acts to increase the similarity between 2D and 3D waveforms, it does not consider the amplitude of pressure fluctuations. Hence, the variance between individual reflection amplitudes between the two modelling approaches is unavoidable. As time increases ($t > 20$ ms) a greater number of strong reflections are present in the 3D FDTD model RIR as reflected wave components are free to travel in all directions around the solid angle. These later reflections are not preserved in the 2D multiplane model as wave components are constrained to travel in horizontal or vertical directions only. However, a series of later reflections occurring in

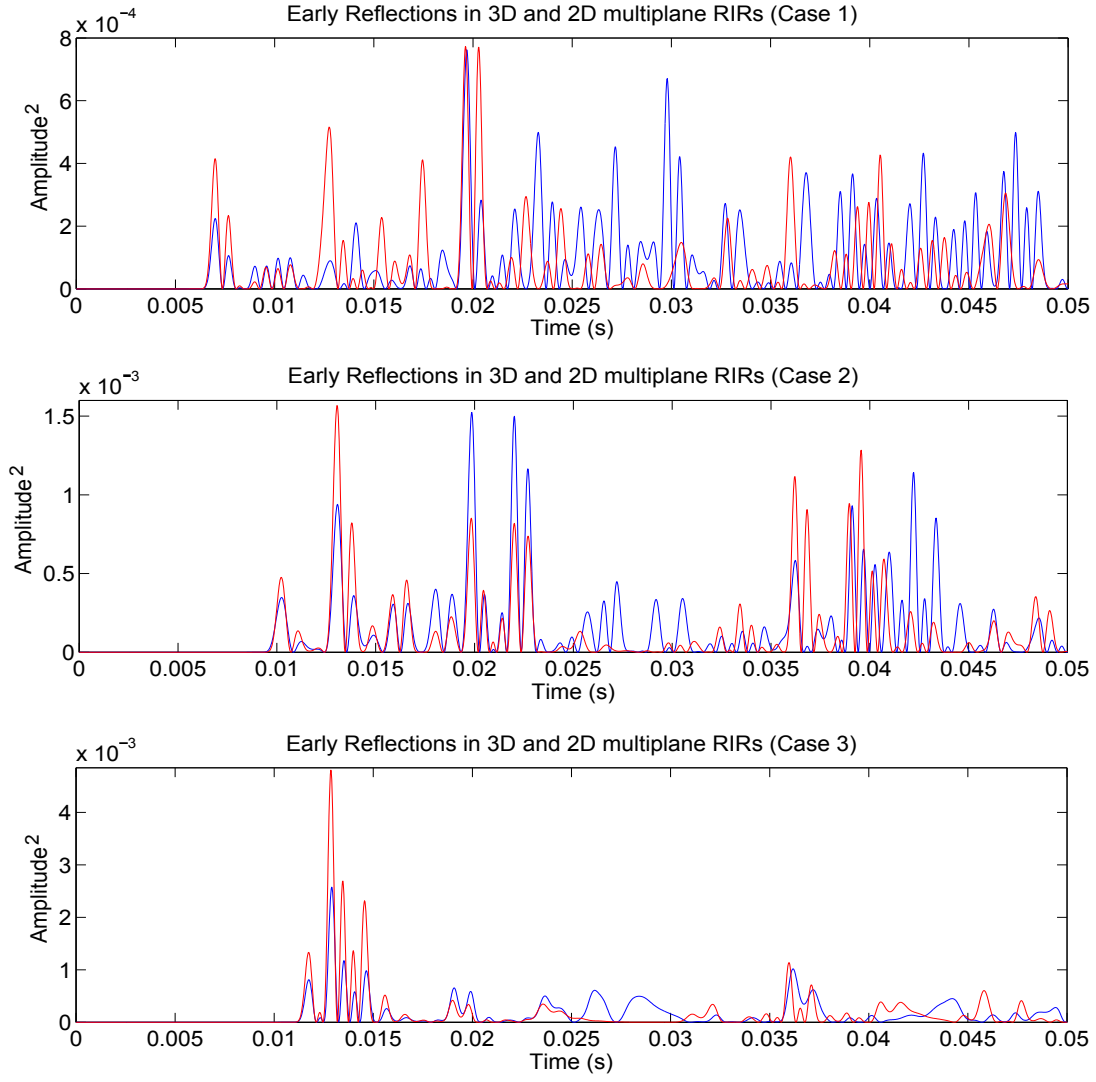

FIGURE 5.16: Early time 2D multiplane (red) and 3D (blue) FDTD RIRs. Square pressure amplitude curves are shown to highlight the temporal positioning and pressure values of early reflection components in each model RIR case. The initial 50 ms of RIR output is shown for each.

the time region $t = [40 : 50]$ ms are shown to be represented in part by the multiplane model.

Similar results are also exhibited for RIR Cases 2 and 3 (middle and bottom graphs respectively). For both cases the receiver is located equidistant from three bounding surfaces with Case 3 closer to the corner of the model. Again, good agreement between the temporal positioning of strong early reflections ($t < 20$ ms) simulated in each model is evident. Due to the location of the receiver in both cases, these early reflections occur in rapid succession shortly after the arrival of the direct sound component. This demonstrates the ability of the multiplane model to emulate low order reflection characteristics arising across the cross-sectional planes. As with RIR Case 1, a number of

later reflections simulated in the 3D model are not represented on the cross-sectional planes and overall amplitude values are markedly different between models. However, the agreement between the temporal positioning of later reflections produced by each model is shown to be reasonable in the time region $t = [30 : 40]$ ms for both RIR Cases 2 and 3.

This examination of early reflections simulated in each FDTD model serves to demonstrate that the time of arrival of these RIR components is consistent for a subset of low order reflections which exist across the cross-sectional planes included in the multiplane model. However, such results will only arise for the specific source and receiver location arrangements applied in this pilot study. For source/receiver combinations that are not located on a common vector parallel to one Cartesian axis, the agreement between the timing of early reflections will decrease. This notion is discussed in more depth with regards to the case study documented in section 5.2.

5.1.4 Concluding Remarks and Considerations

This pilot study has investigated the performance of the 2D multiplane FDTD model in terms of low frequency acoustic simulation for a cuboid environment. Three RIR cases corresponding to three different combinations source/receiver locations have been considered. Results gathered from analysis of resulting low frequency spectra suggest that the multiplane model is suitable for simulating a subset of low frequency behaviour rendered in 3D FDTD models. Moreover, the multiplane model has been demonstrated as capable of emulating the difference in low frequency behaviour that occurs when altering the source and receiver locations. These claims are supported by the following findings:

- Visual comparison of multiplane and 3D RIR spectra reveals that both models preserve theoretic modal characteristics with the multiplane model RIRs exhibiting notable similarity to the 3D counterparts in frequency regions containing prominent low order modes.
- Examination of the magnitude similarity and magnitude similarity index has revealed that the agreement between multiplane and 3D FDTD RIR spectra is consistently higher ($MI \geq 0.83$) than that observed from the results of the ODEON model ($MI = [0.761{:}0.83]$). This is particularly apparent when examining the spectra in the modal region.
- Results of FRAC analysis show that the level of correlation between RIR signals is considerably higher when comparing multiplane and 3D FDTD simulations

(FRAC = [0.260:0.445]) in the modal region than when comparing 3D FDTD and ODEON RIRs (FRAC = [0.001:0.035]). This suggests that the increased level of resonant behaviour rendered in the multiplane model, compared to the ODEON model, is culminating in a stronger resemblance to the 3D model than that achieved using ray-based techniques.

- Comparison of spectral characteristics over the low frequency bandwidth f_{r2} = [0:1000] Hz demonstrates that the overall trends in magnitude levels observed in 3D FDTD RIR cases are better represented by the multiplane model than the ODEON model. This can be confirmed, to some extent, through both FRAC and MI values. This result is of particular importance given the recorded variations in spectral attributes that occur as source and receiver locations change.

Additionally, analysis of early reflection patterns in both FDTD models have shown that the time of arrival for a subset of early reflections is correctly simulated in the multiplane model. However, the amplitudes of these reflections are inconsistent between 3D and multiplane models due to the inherent difference in acoustic energy spreading characteristics. Moreover, the correct timing of early reflections in the multiplane model is highly dependent on the specific arrangements of source/receiver locations applied in each RIR case. It is expected that deviations from this arrangement will result in a decrease in the agreement of reflection patterns yielded in each FDTD model.

Overall, the results of this pilot study are encouraging given that the level of accuracy demonstrated by the 2D multiplane model may be achieved while reducing simulation memory requirements and run-times by $\sim 98.5\%$ and $\sim 99\%$ respectively compared to 3D FDTD modelling. In summary, it may be concluded that the multiplane technique is a viable alternative acoustic modelling paradigm to hybrid ISM/ray-tracing in terms of low frequency accuracy and 3D FDTD numerical schemes in terms of computational efficiency.

Finally, specific considerations regarding the refinement of the multiplane model have been demonstrated through this pilot study. These may be summarised as follows:

- The arrangement of the multiplane model, as applied here, constrained the model to two planes only leading to the omission of a subset of modal resonances. An additional plane is required in order to simulate resonances occurring in the omitted x-z cross-sectional plane.

- A restrictive condition is imposed on the source and receiver locations. This is not ideal as realistic reproduction of particular sonic events may require free placement of both sound source and receiver.

- Investigation of reverberation times recorded in each modelling approach is required in order to assess the level to which this acoustic parameter is correctly represented in the multiplane model.

Each issue listed above is examined in more depth as part of the following case study.

5.2 Case Study: A Recording Studio Live Room

The pilot study demonstrates the applicability of the multiplane FDTD hybrid paradigm to the problem of low frequency simulation in a simple, theoretical acoustic environment. The following case study seeks to extend the investigation of multiplane low frequency RIR synthesis by assessing the accuracy of results produced when simulating the acoustic field of an existing space. To this end, RIRs captured in a recording studio live room are approximated by means of both 3D and 2D multiplane FDTD modelling approaches. Low frequency RIR spectral characteristics and acoustic parameters are examined to compare the agreement between both virtual models and practical measurements. This case study begins by defining and examining the performance of an extended 2D multiplane model as presented in the previous section. The extended multiplane model is then applied to simulate RIRs for the live room.

5.2.1 Extension of Multiplane Model to 3 Planes

The multiplane model used in the pilot study is composed of two cross-sectional planes to allow for straightforward combination of RIR outputs from each plane. Such an arrangement preserves time alignment of direct sound components recorded on each plane and the amplitude of these components are equal, therefore, the output of each plane may be summed during simulation. However, there are two considerable drawbacks to this multiplane modelling approach. Firstly, as exhibited when comparing resulting RIR spectra, restricting the model to two planes leads to the omission of prominent tangential modes occurring in the remaining cross-sectional plane which is not accounted for. Secondly, source and receiver locations are constrained in their location and are defined to lie on a common vector parallel to one Cartesian axis. This criterion limits the application of the model to specific cases of source/receiver positioning. Therefore, arbitrary positioning of sound excitation and reception within modelled environments, which is desirable for the purposes of rendering realistic scenarios, is not possible with this approach. In order to overcome both issues, the multiplane model is extended to three cross-sectional planes. This enables the simulation of additional resonances occurring across the additional cross-sectional plane facilitating the simulation of a larger

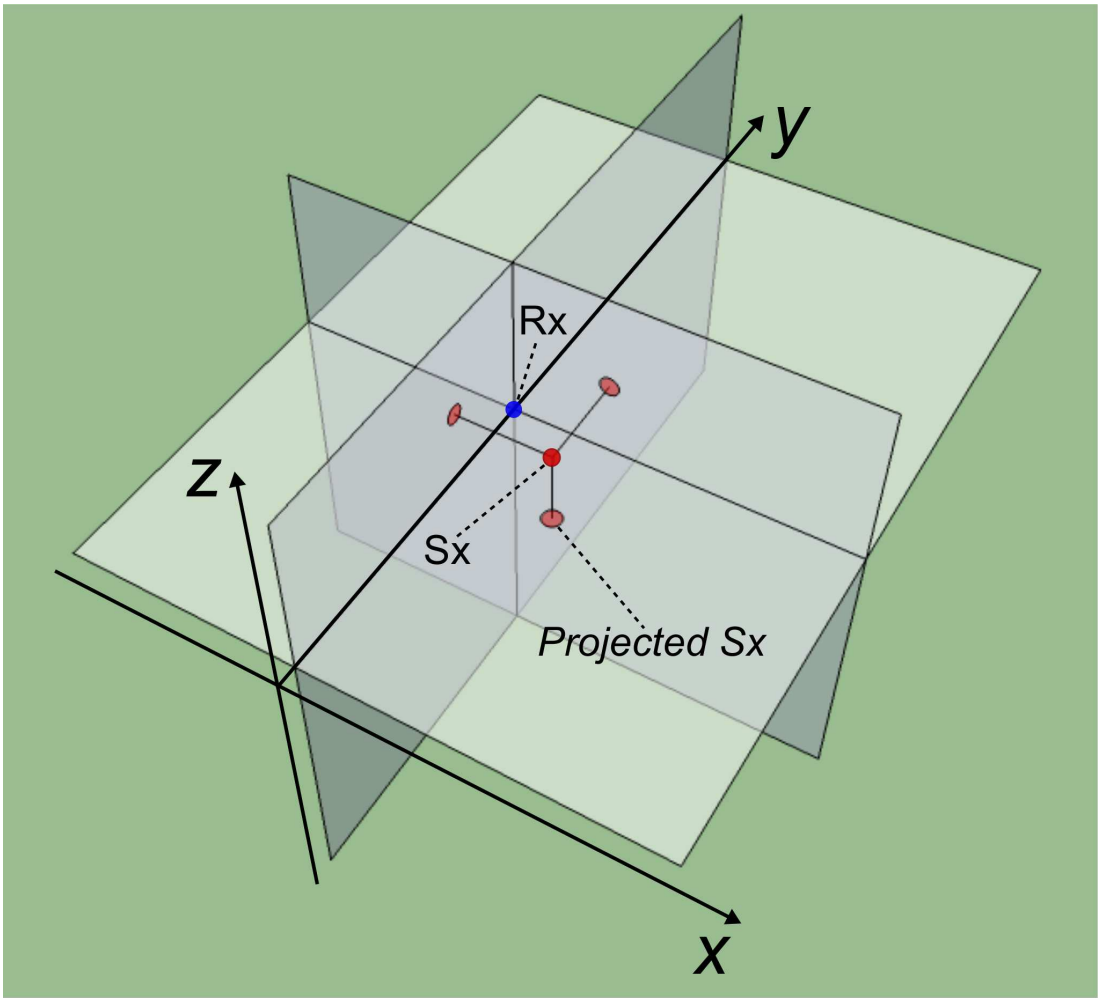

FIGURE 5.17: An example arrangement of the three plane 2D multiplane FDTD model for the case of a cuboid room topology. Source (Sx), projected sources (red) and receiver (Rx, blue) are shown.

subset of room modes that arise in the 3D environment modelled. The three planes are arranged to coincide at a common point defined by the required listening position. In order to excite the acoustic field via the cross-sectional planes, a sound source is projected perpendicularly onto each plane. As such, the source may be positioned arbitrarily within the modelled space. Consequently, both sound source and receiver may be positioned at will removing the constraint arising in the two plane multiplane model.

The overall arrangement is displayed in Figure 5.17. This diagram shows an example of the extended multiplane model arrangement for the case of a cuboid topology. Each plane is orientated with respect to the Cartesian co-ordinate system such that cross sections of the space in the x-y, y-z and x-z are each occupied by an independent 2D FDTD SRL plane. All cross-sectional planes are arranged to coincide at the receiver

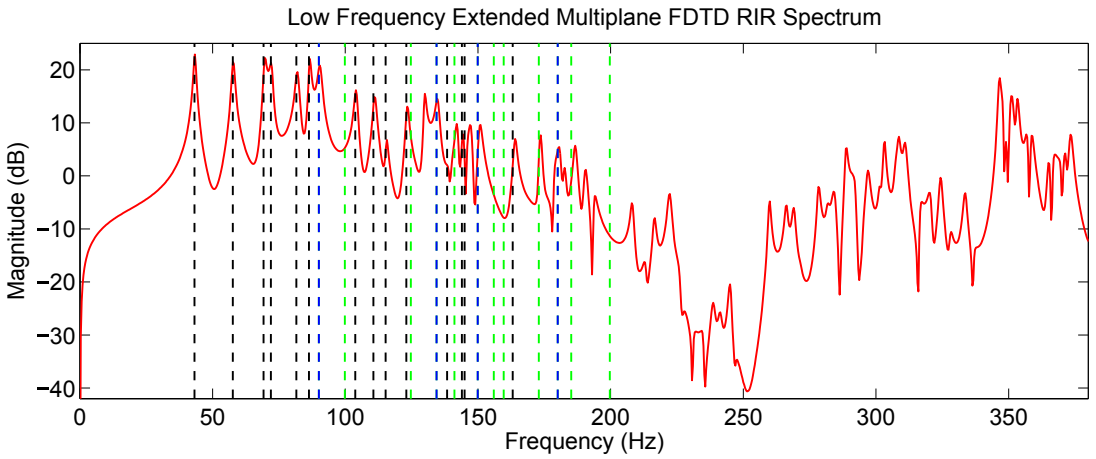

FIGURE 5.18: The magnitude spectrum of an RIR rendered using the three plane multiplane model for the cuboid topology simulated in the pilot study. Theoretic mode values are denoted by the dashed vertical lines. Tangential modes arising in the additional x-z plane are highlighted in blue and oblique modes are shown in green.

location (Rx). The source is injected into a plane at a grid node defined by the two common co-ordinates to both the plane and source location. For example, a source location with 3D Cartesian co-ordinates (1, 2, 3) would be projected onto point (1, 2) on the x-y cross-sectional plane, while the y-z and x-z plane schemes would have point source locations (2, 3) and (1, 3) respectively.

For the purposes of initial testing of the three plane multiplane model, the pilot study cuboid environment is simulated using this extended approach. The three cross-sectional planes are defined to coincide at a common receiver location of (2.59, 2.58, 2.08) m. The sound source is projected onto each plane from a position of (0.38, 0.38, 0.38) m. As such, the source and receiver are placed towards diagonally opposing corners in all dimensions. Figure 5.18 displays the spectrum of the RIR simulated for this case. For the purposes of comparison with previous implementation results, tangential modes arising in the additional x-z plane are denoted by the blue dashed vertical lines. Oblique modal values, which are not preserved by the multiplane model, are given by the green dashed vertical lines. As shown, the extended multiplane model demonstrates the capability to simulate the additional modal resonances that were omitted in the pilot study simulations. In particular, the 1^{st}-order x-z tangential mode occurring at $f \approx 90$ Hz is now present in the resulting RIR spectrum. This result demonstrates the advantage of the use of an additional plane as a larger subset of low frequency spectral components are captured in this model. Furthermore, comparison of the extended multiplane RIR with a corresponding 3D FDTD RIR, through calculation of FRAC and MI values, provides an additional means of evaluation. The MI value calculated for the modal frequency range ([0:380] Hz, as defined for the pilot study) is 0.837. This is

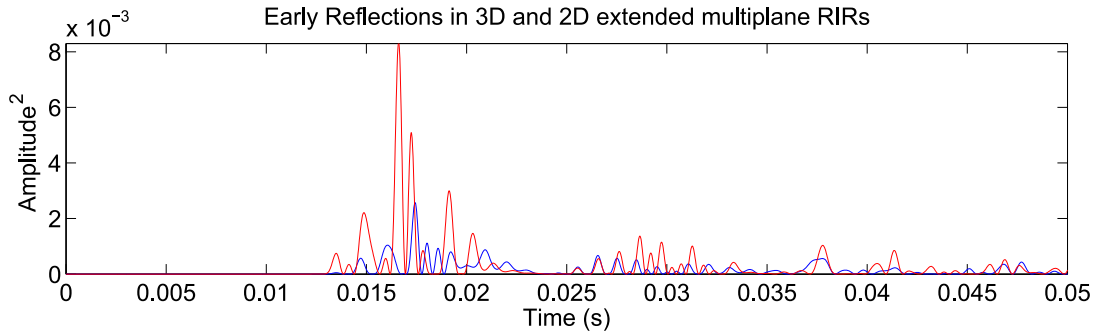

FIGURE 5.19: Early time extended 2D multiplane (red) and 3D (blue) FDTD RIRs. Square pressure amplitude curves are shown to highlight the temporal positioning and pressure values of early reflection components. The initial 50 ms of RIR output is shown.

consistent with the range of values documented for the two plane pilot study model in the modal region. However, a significant improvement is evident from FRAC analysis which, for the three plane multiplane model, is calculated as 0.557. This is notably higher than all FRAC values detailed in the pilot study, calculated as 0.488, 0.397 and 0.260 for RIR Cases 1-3 respectively ($f = f_{r1}$). Hence, the MI and FRAC values calculated for the extended multiplane model, in combination with the findings from inspection of the resulting RIR spectrum, suggest that this 3 plane approach provides a more accurate representation of low frequency sound propagation than the two plane model.

However, there is an inherent drawback associated with this extended form of multiplane modelling which arises due to the free positioning of source and receiver locations. As demonstrated in the pilot study, the temporal positioning of early reflections matched a subset of those recorded in the 3D FDTD model. This is due to the coplanar locations of source and receiver in both modelled cross-sections of the multiplane scheme. In terms of the extended multiplane model, the correlation between the timing of early reflections occurring in 3D and those rendered in the cross-sectional planes is not guaranteed for arbitrary source/receiver locations. This is demonstrated by Figure 5.19 which displays the squared amplitude of low order early reflections in the 3D and 2D extended multiplane FDTD RIRs simulated for the low frequency analysis discussed previously. As shown, the agreement between the temporal alignment of early reflections is reduced compared to that observed for the pilot study RIRs. This is most apparent for the first cluster of reflections ($t = [0.01{:}0.025]$ s) in which reflections in the extended multiplane model RIR are seen to be mostly misaligned against the 3D RIR. Hence, the additional accuracy in terms of modal response demonstrated by the extended multiplane model and free placement of source and receiver is realised at the expense of reduced accuracy in early reflection characteristics.

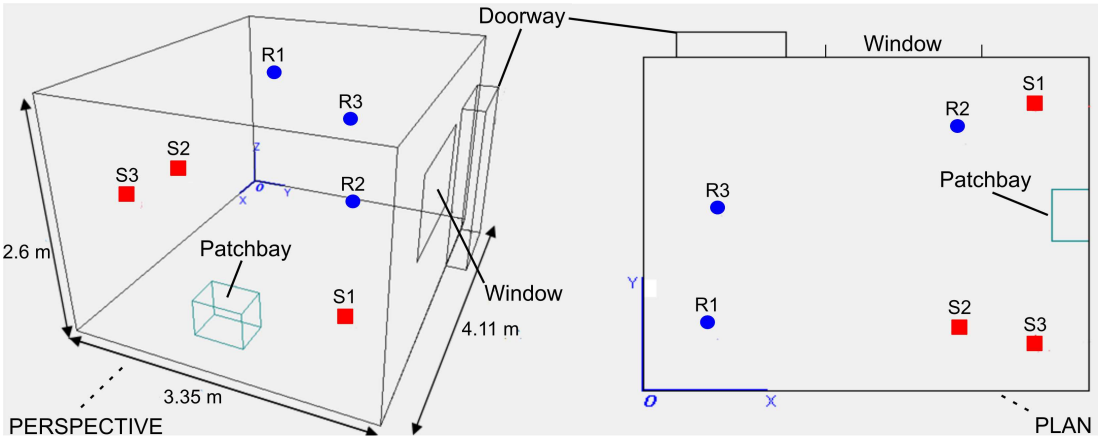

FIGURE 5.20: Perspective (left) and plan (right) views of the recording studio live room geometry with major dimensions and structural components highlighted. Source (S1-S3) and receiver (R1-R3) locations for RIR Cases 1-3 are also shown.

5.2.2 RIR Capture in the Live Room

The acoustic environment selected for this study is the recording studio live room in the Audio Lab at the University of York. This space is a (4.11 x 3.35 x 2.6) m cuboid shape with an extruded doorway area and a window positioned on one side of the room. Additionally, a small cuboid patch bay is located on one of the walls. Figure 5.20 provides an overview of the major geometric dimensions and components of the space that are included in the acoustic models created for this case study. Also shown in the diagrams are the three source and receiver locations used to capture each RIR simulated using the acoustic models.

The arrangement of the source and receiver location for RIR Case 1 is defined such that the acoustic field is excited and recorded in opposite corners of the space in all dimensions. Therefore, it is expected that the RIR recorded for this source/receiver case will contain a higher level of resonant sound energy than that of the remaining two cases due to the build up of low frequency energy that occurs towards the corners of a cuboid space. This concept is demonstrated in the pilot study (comparing RIR Cases 1 and 3 for the simple cuboid). RIR Case 2 incorporates a source/receiver combination in which both excitation and recording locations are defined on a common vector positioned on the x-y plane perpendicular to the z-axis. Finally, RIR Case 3 considers a sound source location in the proximity of a corner and a receiver positioned near to the wall opposite. This selection of source/receiver arrangements provides a means of assessing the ability of the multiplane model to simulate the soundfield of the space as source, receiver and planar cross-section are altered. Table 5.9 provides the coordinates of source and receiver locations as applied for each RIR case.

CASE	Source (x, y, z) (m)	Receiver (x, y, z) (m)
1	(3.61, 2.85, 0.68)	(0.60, 0.69, 2.00)
2	(2.91, 0.65, 1.49)	(2.91, 2.65, 1.49)
3	(3.61, 0.50, 1.45)	(0.70, 1.85, 1.50)

TABLE 5.9: Source and receiver placements defined for three cases of RIR measurement in the recording studio live room.

Practical RIR measurements are taken in the live room using the exponential swept sine technique, as developed and defined by Farina [43]. To this end, a Genelec 8130A loudspeaker is used as the sound source producing 30 s exponential sine sweeps over the approximate range of human hearing f = [20:20000] Hz. In order to approximate an omnidirectional sound source, as applied in the FDTD and geometric acoustic models, the loudspeaker is rotated around the azimuthal plane at instances of 90°. The response produced for each of the four loudspeaker orientations is recorded using a ST450 Soundfield B-Format microphone [159] and 24-bit bit-depth (F_s = 48 kHz). This results in four RIRs per source/receiver case. Mono RIRs are produced for each RIR case by summing the B-Format W-channel of the captured responses for the four loudspeaker orientations. As such, the three RIRs recorded in the live room possess approximately omnidirectional source and receiver directivity for consistency with the FDTD models created for this study. However, it must be noted the low frequency directivity of the loudspeaker is inherently omni-directional at low frequencies. Therefore, the summation of multiple responses required to approximate an omnidirectional sound source for mid-high frequencies also has the effect of enhancing the levels of low frequency energy in the responses recorded. This is shown to have little impact on the results of low frequency analysis as documented later in section 5.2.4.2. Examples of the practical RIR recording arrangement are provided in Figure 5.21.

FIGURE 5.21: (Left and Right) Examples of RIR capture setup in the live room showing the Genelec 8130A and Soundfield microphone. Note, the above images are not indicative of source and receiver locations used to record the RIRs documented in the following.

5.2.3 The Acoustic Models

The major geometric attributes of the recording studio live room (see Figure 5.20) provide the approximate representation of the space used in each acoustic model. The creation of the geometric and FDTD models closely follows the procedure set out in the pilot study. Additionally, the absorption characteristics of surfaces in the live room are represented in each acoustic model described in the following.

5.2.3.1 Geometric Model

The geometric model is rendered using ODEON 10.1 Auditorium [21] for the three RIR cases listed previously. Room dimensions and source/receiver placements are replicated exactly in this model. For consistency with the FDTD models, which incorporate frequency-independent boundary conditions, a single absorption coefficient is defined per surface material with reference to the ODEON material library. This library contains random-angle-of-incidence absorption coefficients measured for a range of material types (see e.g. [39]). The closest material types to those present in the live room were selected and the absorption coefficient defined for each material over the 500 Hz octave band is applied for all simulation frequencies. Table 5.10 provides a summary of the absorption properties applied to each surface incorporated in the model, where α_{500} refers to the absorption coefficient α for the 500 Hz octave band.

Surface	Material Type	α_{500}
Floor	Pile Carpet	0.25
Ceiling	Wood Fibre	0.20
Walls	Plasterboard	0.08
Window	Double Glazed Glass	0.05
Door	Wood (with air gap)	0.09
Patchbay (Housing)	Wood (16 mm thick)	0.10
Patchbay (Front)	Aluminium Plate	0.45

TABLE 5.10: Absorption coefficients of material types as applied to surfaces in the ODEON model of the live room. Values are sourced directly from the ODEON Auditorium 10.1 material library [21]

The geometric model is simulated in ODEON using 50000 rays and a transition order of 5 for each RIR case detailed previously. Sound source directivity is defined as omnidirectional and mono RIRs are recorded by extracting the W-channel of a B-Format 24-bit .wav file. The resulting mono RIRs are then subject to the energy calibration process described in section 5.1.2.1 to ensure consistency with the FDTD model RIRs in terms of excitation energy levels.

5.2.3.2 3D FDTD Model and RIR Synthesis

For the purposes of SRL FDTD simulation, the live room geometry is voxelised into a cubic lattice of pressure nodes with a spatial sampling instance corresponding to F_s = 44.1 kHz and wave speed c = ms^{-1}. This yields h_{3D} = 0.0135 m as defined for the 3D FDTD pilot study model. As such, a maximum deviation between the dimensions of the live room and, therefore, the geometric model, may be derived as $\sqrt{3h_{3D}^2}/2$ = 0.0117 (m) corresponding to the maximum error implied by rounding dimensions to the nearest spatial sampling instance. This error, which is also valid for the discrete location of source and receiver positions, is considered negligible.

Pressure-centered, frequency-independent LRS boundary conditions are applied to pressure nodes occupying the surfaces of the model. This constitutes an important distinction between the geometric model and the actual space in that only a single absorption coefficient is applied to each surface. Therefore, the frequency-dependent absorption characteristics possessed by the surface materials of the live room are not represented in the FDTD model. However, absorption coefficients defined in the ODEON model for the 500 Hz octave band (see Table 5.10) are utilised to provide an approximation to the properties of each surface. Table 5.11, provides an overview of the corresponding reflection and impedance values applied for each material type. With

Surface	Material Type	R_{3D}	ζ_{3D}
Floor	Pile Carpet	0.866	13.925
Ceiling	Wood Fibre	0.894	17.868
Walls	Plasterboard	0.959	47.781
Window	Double Glazed Glass	0.975	79.000
Door	Wood Fibre	0.954	42.478
Patchbay Housing	Wood (16 mm thick)	0.949	38.216
Patchbay Front	Aluminium Plate	0.742	6.752

TABLE 5.11: Reflection R_{3D} and impedance ζ_{3D} coefficients for each surface type as applied in the 3D FDTD model of the live room. Coefficients are derived from absorption data for each material type as defined in the ODEON Auditorium 10.1 material library [21] for the 500 Hz octave band.

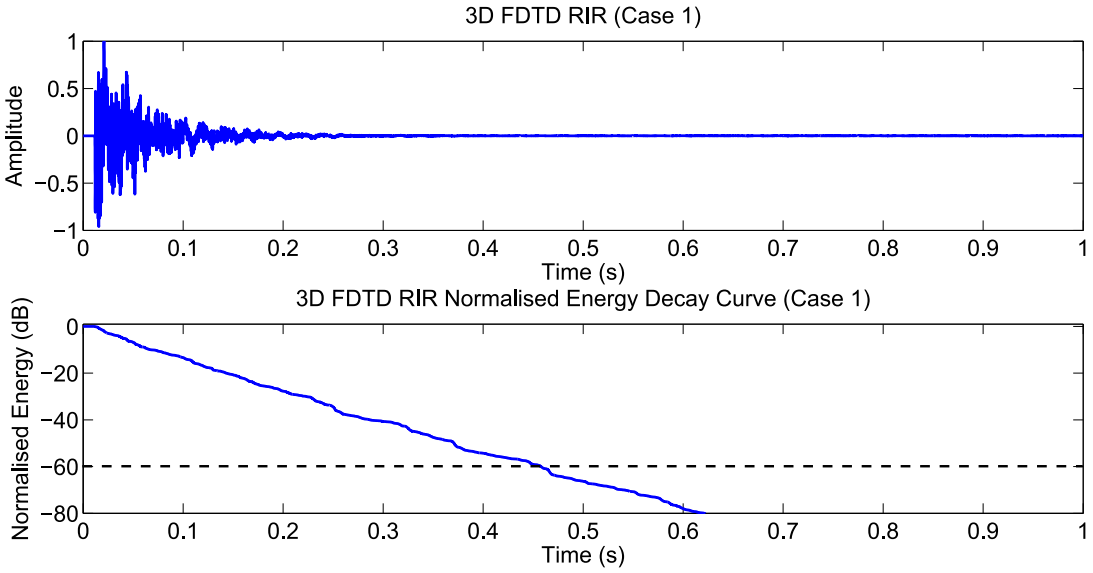

FIGURE 5.22: (Top to bottom) Normalised pressure amplitude and energy decay curves for 3D FDTD RIR Case 1. The first second of simulated output is shown and the -60dB level is displayed on the bottom graph by the dashed black line.

reference to the investigation of numerical stability of pressure-centered LRS formulations (see section 4.4.4), it is noted that this model maintains stability for a run-time sufficient to allow the RIR to decay to inaudible levels relative to the direct sound component. This is demonstrated in Figure 5.22 which displays 1 second of audio output for RIR Case 1. As shown, the normalised amplitude and corresponding energy decay curves for this RIR case decrease to a sufficient extent without any numerical stability issues. The justification for this consideration stems from the fact that the normalised energy decay curve drops below -60dB of the initial energy level produced by the source excitation.

A soft source Kronecker delta excitation function is applied for each source and receiver combination. Resulting RIRs are then passed through a 2^{nd}-order DC blocking filter and low pass filtered to remove frequencies above 2 kHz. This cut-off frequency is selected to allow analysis of acoustic parameters up to and including the 1 kHz octave band while avoiding the effects of dispersion error. Noting that the 1 kHz octave band has an upper frequency of 1.42 kHz, this analysis bandwidth is well within the advised range of $0.075F_s$ (as derived in section 4.2) to maintain a maximum isotropy error of $<2\%$ in all directions around the solid angle. For the purposes of comparative low frequency analysis, detailed later in section 5.2.4, each RIR produced using the 3D FDTD is calibrated to the corresponding ODEON RIR through multiplication by the energy matching coefficient, $K_{3D,FDTD}$ (5.7).

Approximately 17 000 000 pressure nodes are required to voxelise the live room geometry, requiring a total of $\sim$ 0.253 GB of memory to simulate the target acoustic field at double point precision. An average simulation time of 3800 s per 1 s of RIR output synthesised at $F_s = 44.1$ kHz.

5.2.3.3 2D FDTD Model and RIR Synthesis

Following the procedure for constructing the extended multiplane model set out in section 5.2.1, the live room geometry is simulated using three cross-sectional, planar 2D FDTD schemes. For each RIR case, the cross-sectional schemes orientated in the x-y, x-z and y-z planes are defined to coincide at the receiver location. The cross-sections are voxelised using a rectilinear grid corresponding to a temporal sampling rate of $F_s = 44.1$ kHz and wave speed $c = 344$ ms^{-1}. Using (5.2), the internodal sampling instance is calculated as $h_{2D} = 0.011$ m providing the highest grid resolution while maintaining numerical stability. This value of h_{2D} results in a maximum deviation of cross sectional dimensions, from those present in the real space, of $\sqrt{2h_{2D}^2}/2 = 0.0078$ m provided dimensions are rounded to the nearest sampling instance. As with the case of the 3D FDTD model, this small error is also applicable to source and receiver placements and are considered negligible. Additionally, this grid resolution maintains minimal dispersion error for frequencies up to and including the 1 kHz octave band as is consistent with the 3D FDTD model. This follows from the discussion of section 4.2 which details an upper frequency limit of the usable simulation bandwidth of $f_{max} < 0.1\ F_s$ for 2D SRL FDTD schemes operating at the Courant limit. The multiplane models for each RIR case require approximately 300 000 pressure nodes for a constant internodal distance h_{2D} as defined previously. This corresponds to a memory requirement of $\sim$0.0045 GB to simulate the target acoustic field. Simulation run-times for 1 s of RIR output with $F_s = 44.1$ kHz are recorded as $\sim$60s on average. Therefore, the multiplane model used on this modelling scenario achieves a $\sim$98% reduction in terms of both memory requirements and processing time compared to the 3D FDTD model.

As noted in the pilot study, the laws governing energy decay in 3D and 2D acoustic systems are different (see e.g. [31]). This leads to the requirement of controlling the absorption rates inherent to the multiplane model such that recorded reverberation times are consistent with those recorded in the 3D FDTD models. Converse to the simple modelling scenario investigated in the pilot study where a single absorption coefficient is applied to all surfaces, the live room model incorporates a range of material types, each with a specific absorption coefficient. Therefore, a more rigorous method of calibrating 3D and 2D model absorption characteristics is warranted in this case study.

In order to approximately match 3D and 2D absorption coefficients, a 3D SRL FDTD 2 x 2 x 2 m cube model with pressure-centered LRS boundary conditions is simulated (F_s = 44.1 kHz, h_{3D} = 0.0135 m, c = 344 ms^{-1}) once per material type. In each simulation of the cube model, an absorption coefficient corresponding to one of the material types is applied consistently to all surfaces. Upon simulation of an RIR for each material type (M), the frequency-independent $RT_{60,3D_M}$ is calculated. The values for $RT_{60,3D_M}$ are then inserted into the Norris-Eyring equation for a 2D acoustic system, after [31], rearranged to make the absorption coefficient the subject:

$$\alpha_{2D,M} = 1 - e^{\frac{-\pi S ln(10^6)}{cLRT_{60,3D_M}}} \tag{5.16}$$

where S (m) and L (m) are defined as the surface area and perimeter of the 2 x 2 m cross-section of the cube respectively and c is the simulated wave speed. Using (5.16), it is possible to calculate equivalent 2D absorption coefficients α_{2D} for each material type M. Table 5.12 details the corresponding reflection R_{2D} and impedance values ζ_{2D} applied to simulate the absorption characteristics of the surface materials incorporated on each cross-sectional plane as appropriate. 3D reflection values are also listed for reference.

Surface	Material Type	R_{3D}	R_{2D}	ζ_{2D}
Floor	Pile Carpet	0.866	0.798	8.901
Ceiling	Wood Fibre	0.894	0.854	12.699
Walls	Plasterboard	0.959	0.906	20.277
Window	Double Glazed Glass	0.975	0.944	34.714
Door	Wood Fibre	0.954	0.903	19.619
Patchbay Housing	Wood (16 mm thick)	0.949	0.900	19.000
Patchbay Front	Aluminium Plate	0.742	0.674	5.135

TABLE 5.12: Approximated equivalent reflection R_{2D} and impedance ζ_{2D} coefficients for each material type as applied in the 2D multiplane model. Reflection coefficients applied in the 3D FDTD model are provided for reference.

Multiplane Model Sound Source Considerations:

The excitation of each plane is facilitated by projecting the source from the source location defined for each RIR case as detailed in section 5.2.1 (see Figure 5.17 for example). A soft source Kronecker Delta function is applied to the projected source location on each plane in order to excite the target acoustic field. Upon recording RIRs simulated within each plane, each RIR is then subject to afterglow removal, DC and lowpass filtering processes following the procedure developed for the pilot study.

An additional post-simulation process involves time-aligning and normalising the contributing RIRs from each plane prior to the summation that provides the complete multiplane RIR. Due to the differences in path lengths between the projected source locations defined on each plane, the time-of-arrival of the direct sound component recorded on cross-section must be altered such that each one possesses the same direct sound onset time. This is achieved by shifting each planar response in time as appropriate in each RIR case. Additionally, contributing RIRs recorded on each plane are normalised by equating the maximum pressure amplitude of each recorded direct sound component. This is implemented by finding the maximum amplitude of the direct sound received on each plane and scaling each planar response through multiplication with the following constant:

$$K_{MP} = \frac{DS_{RIR_{max}}}{DS_{RIR}} \tag{5.17}$$

where $DS_{RIR_{max}}$ is the largest maximum direct sound pressure amplitude of the three planar responses and DS_{RIR} is either one of the remaining direct sound amplitudes recorded on the remaining cross-sections. As such, K_{MP} provides two different values that are applied in order to equate the direct sound components of the two planar responses that possess the lowest direct sound peak amplitudes.

This process of direct sound amplitude matching is largely influenced by the characteristics of the multiplane model constructed for the pilot study. In the two plane approach, both the time-of-arrival and direct sound amplitudes are equal due to the length of the direct sound path being equal over each contributing plane. Hence, through the processes of direct sound component normalisation and time alignment, the extended three plane multiplane model replicates the attributes of the two plane model prior to combining each contributing planar response into the overall multiplane RIR. While this is considered to be a reasonable means of calibrating the output from each cross-sectional scheme, this pilot study demonstrates a scenario in which this approach does not produce reasonable results.

Considering RIR Case 2, the source and receiver locations are defined as (2.91, 0.65, 1.49) m and (2.91, 2.65, 1.49) m respectively. Noting that the x- and z-coordinates of both source and receiver are equal, it follows that the projected source location on the x-z plane is the same as the receiver location. As such, the contributing response recorded on the x-z plane begins with a large direct sound component as the source signal is inserted directly into the receiver node. Therefore, by imposing the method of direct sound matching, the two remaining planar responses with comparatively much lower direct sound amplitudes are scaled by a large K_{MP} value. The result of applying direct sound matching process to the x-y and x-z planar responses for RIR Case 2 are shown in Figure 5.23. It may be observed that the amplitude of the x-z planar

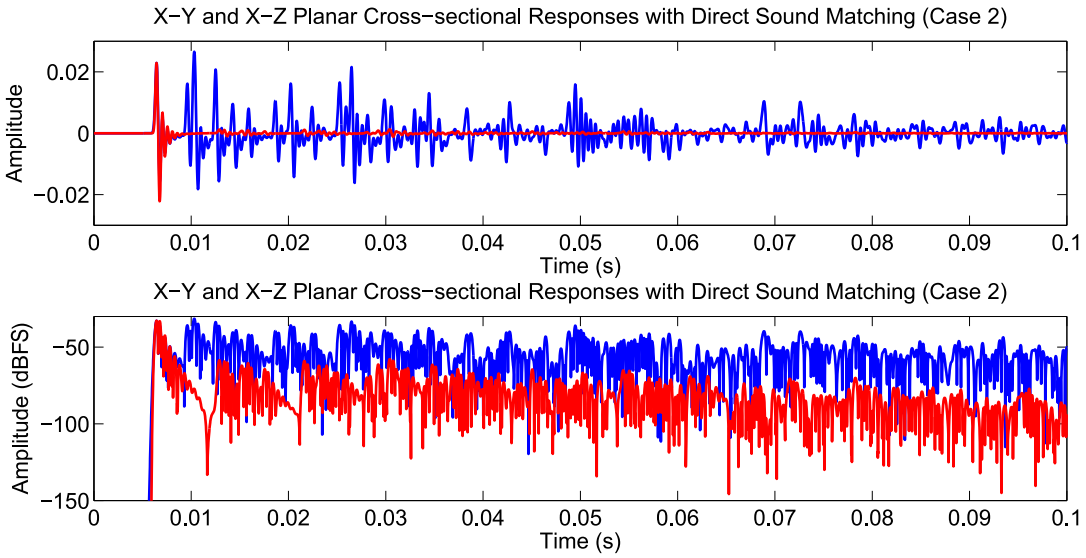

FIGURE 5.23: (Top) Cross-sectional responses recorded from the x-y (blue) and x-z (red) planes for RIR Case 2. As shown, the application of the direct sound component matching procedure leads to a significant difference in response amplitudes after the direct sound ($t > 0.01$ s). (Bottom) Logarithmic pressure response (dBFS) of the responses displayed in the top graph emphasising the difference in response amplitudes as time increases.

response is significantly lower than that of the x-y planar response after the direct sound component ($t > 0.01$ s). Therefore, upon summation of the responses from each plane for RIR Case 2 it is evident that the x-z planar response will contribute much less to the overall multiplane RIR than the other two planar responses. As a result, the resonant frequency components occurring in the x-z plane of the model will not be sufficiently represented relative to those occurring across the other two planes.

In order to correct this prior to producing the Case 2 multiplane RIR, the x-z planar response undergoes a direct sound component scaling operation. This process involves isolating the direct sound component of the x-z planar response and reducing the amplitude such that it matches the amplitude of the direct sound component recorded on the x-y plane. The result of this procedure is displayed in Figure 5.24. As observed, the amplitude of the direct sound components are equal while the response amplitudes for $t > 0.01$ are in much higher agreement. Therefore, through applying the direct sound scaling operation in this RIR case, the x-z planar response sufficiently contributes to the overall multiplane RIR.

For completeness, the contributing x-y (blue), y-z (black) and x-z (red) planar responses for RIR Cases 1 and 3 are provided in Figure 5.25. As shown, the overall amplitudes of the planar responses in these cases are of a consistent level in the early stages of the RIRs ($t < 0.02$ s). This demonstrates that the method of direct sound matching is suitable for these RIR cases as the sound source projected onto each plane

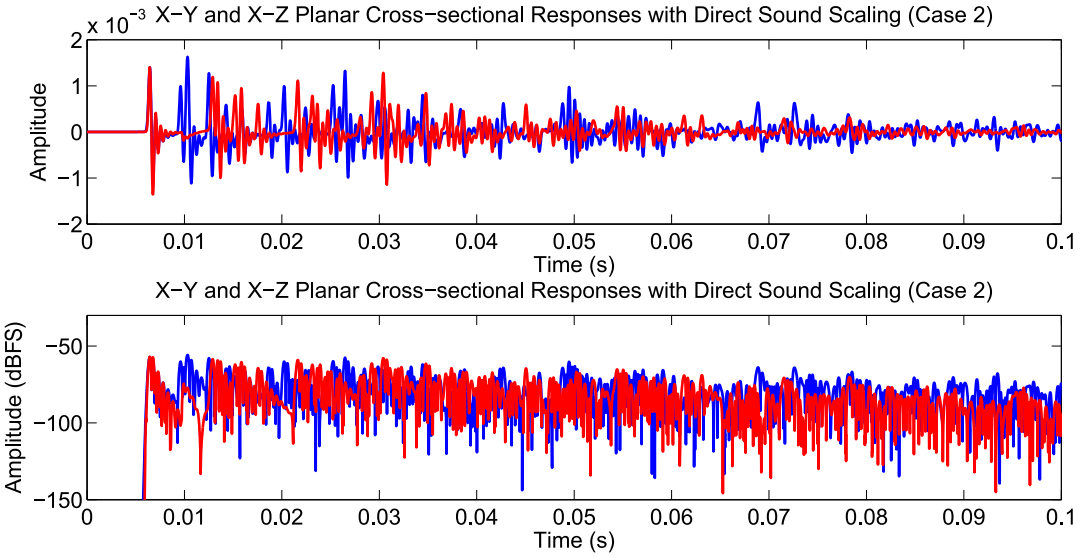

FIGURE 5.24: (Top) Cross-sectional responses recorded from the x-y (blue) and x-z (red) planes for RIR Case 2. As shown, the application of the direct sound scaling procedure leads to improved agreement between response amplitude ($t > 0.01$ s). (Bottom) Logarithmic pressure response (dBFS) of the responses displayed in the top graph highlighting the improvement in response level agreement.

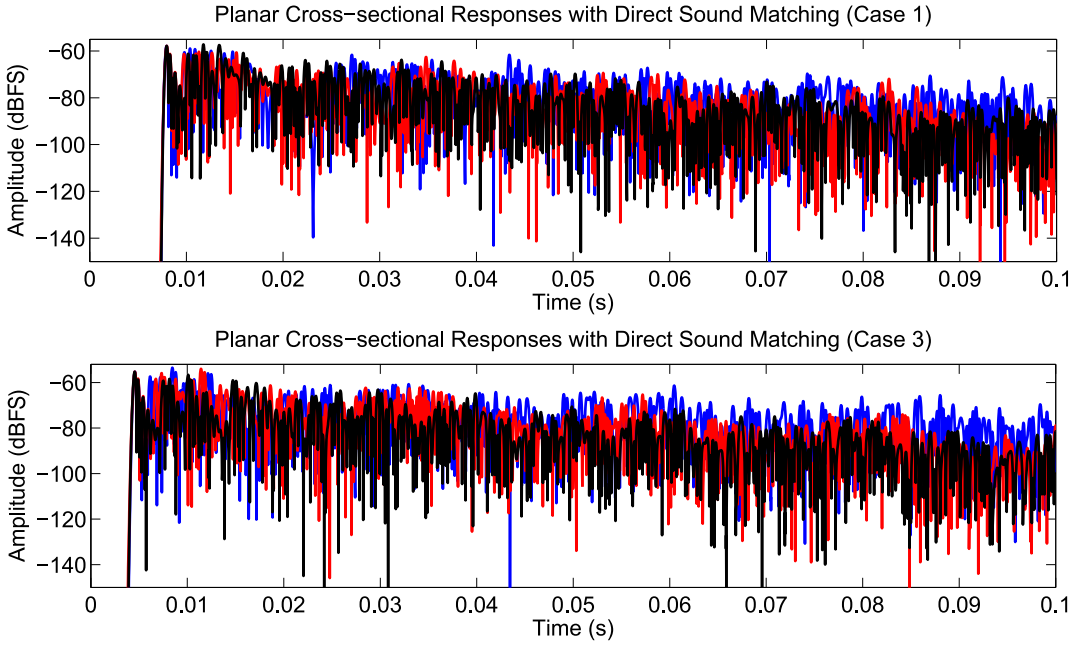

FIGURE 5.25: Planar response contributions for RIR Cases 1 (top) and 3 (bottom). Individual x-y (blue), x-z (red) and y-z (black) cross-sectional logarithmic pressure responses are shown for both cases.

is not in close proximity to the receiver. However, as observed for RIR Case 2, this approach to combining planar responses to produce the multiplane RIR cannot be considered valid for scenarios in which the projected sound source is defined on or near to a given receiver position. This issue is addressed further in section 5.3.3.2 where the multiplane source model is refined to facilitate improved matching of planar responses for arbitrary source and receiver locations in terms of overall response amplitudes.

Upon simulation and summation of the contributing planar responses for each RIR case to render the overall multiplane RIRs, the energy matching coefficient $K_{MP,i}$ (5.10) is applied. This ensures that the energy levels present in each multiplane RIR are consistent with the energy levels of the corresponding 3D FDTD RIRs for the purposes of low frequency analysis as detailed in the following section.

5.2.4 Results and Discussion

Results discussed in the following are derived from the RIRs simulated during the course of this case study. Both acoustic parameters and low frequency spectral characteristics recovered from all acoustic models for RIR Cases 1-3 are presented and examined in turn. As per the analysis procedure followed for the pilot study, the examination of objective results serves to demonstrate the capabilities of 2D multiplane acoustic modelling as an efficient means of hybrid RIR synthesis.

5.2.4.1 Acoustic Parameters

The RIR outputs from the 3D FDTD, 2D multiplane and ODEON acoustic models of the live room are examined here in terms of the reverberation time (EDT and T_{30} (s)) and Clarity (C_{80} (dB)) measures detailed in section 2.3.8. Each parameter value is derived from the RIRs for the 63, 125, 250, 500, 1k Hz octave bands using the MATLAB Acoustic Parameter Toolbox [160] in accordance with ISO documentation [17]. Additionally, JND values for each acoustic parameter, as detailed in section 2.5.3, are provided throughout in order to analyse the agreement of RIRs in a perceptual sense.

Figure 5.26 displays the T_{30} values calculated from the RIRs simulated by the 3D FDTD (blue), 2D multiplane FDTD (red) and ODEON (black) acoustic models for all source/receiver arrangements over the 63 - 1k Hz octave bands. The JND range provided by the blue error bar for each octave band is calculated as +/- 24.5% of the T_{30} obtained from the 3D FDTD RIRs corresponding to one JND after [16]. It is immediately apparent that the RIRs produced using the geometric ODEON model possess

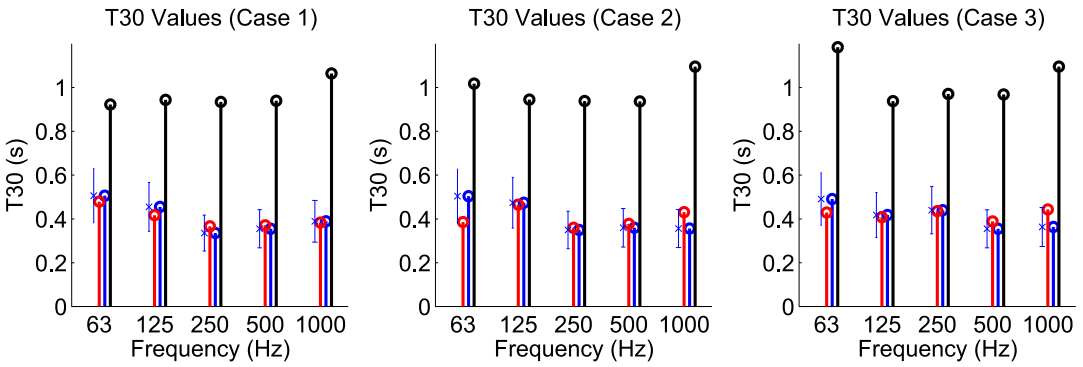

FIGURE 5.26: T_{30} values calculated from 3D FDTD (blue), 2D multiplane FDTD (red) and ODEON (black) RIRs (Cases 1 - 3) for the 63, 125, 250, 500, 1000 Hz octave bands. The error bars display the range of one JND (24.5%, after [16]) above and below the T_{30} value calculated from the 3D FDTD RIRs in each octave band.

much longer energy decay times that those recorded from both 2D multiplane and 3D FDTD RIRs. T_{30} times derived from the ODEON RIRs are consistently at least twice those arising in the FDTD models across all RIR cases and octave bands. This result demonstrates the differences in energy absorption rates inherent to geometric and SRL FDTD numerical acoustic models that arise when random-angle-of-incidence absorption coefficients are applied in both modelling methods. While this is not a primary concern in this case study, it does impact on the extent to which the comparison of geometric and FDTD RIR spectra may be considered appropriate. This is due to the fact that large differences in RIR energy decay times affect the level of agreement between spectra magnitude values, as discussed further in the following section. For this reason, normal-angle-of-incidence absorption coefficients are used in order to alleviate these differences in later work (see section 5.3) and, therefore, provide a more appropriate comparison between geometric and FDTD RIR spectra. The incorporation of normal-incidence absorption coefficients in SRL FDTD acoustic models is justified by the fact that all reflections in such models occur in a direction normal to boundary surfaces due to rectilinear spatial discretisation.

Considering T_{30} as derived from the 3D FDTD and 2D multiplane RIRs, it is observed that both models are in good agreement in terms of this parameter. The 2D multiplane RIR T_{30} values are calculated to be within one JND of the 3D RIR T_{30} for all RIR cases and all analysed octave bands. This suggests that the absorption coefficient matching procedure applied to the 2D multiplane model has been successful in emulating frequency-dependent energy decay characteristics of a 3D FDTD model with a range of different surface material types. However, in some instances (e.g. RIR Case 2, 63 and 1k Hz octave bands and RIR Case 3, 1 kHz octave band) the 2D multiplane RIR

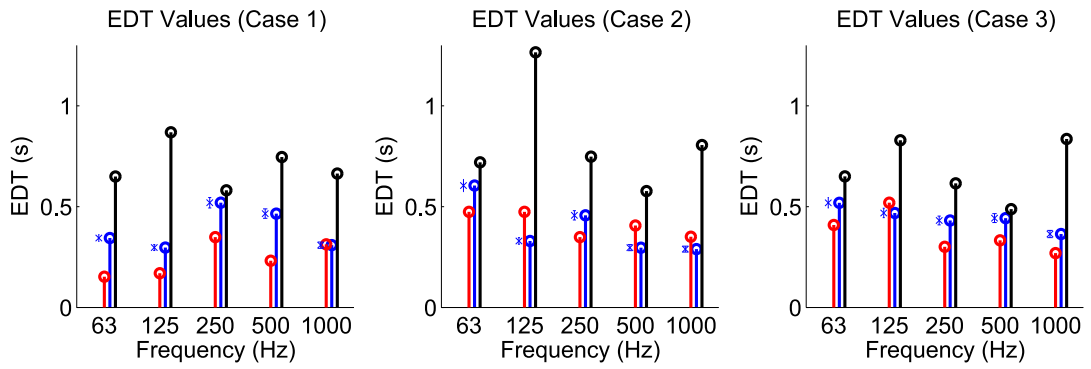

FIGURE 5.27: EDT values calculated from 3D FDTD (blue), 2D multiplane FDTD (red) and ODEON (black) RIRs (Cases 1 - 3) for the 63, 125, 250, 500, 1000 Hz octave bands. The error bars display the range of one JND (5%, after [17]) above and below the EDT value calculated from the 3D FDTD RIRs in each octave band.

T_{30} values are approaching the limit of the T_{30} JND ranges. Hence, the level of agreement between the reverberation times recorded in each FDTD model is not consistent for all octave bands. This is due to the fact that the 2D multiplane absorption coefficients are derived by matching frequency-independent RT_{60} calculated using (5.16) which does not account for frequency-dependent energy decay characteristics common to both 2D and 3D SRL FDTD acoustic simulations. As such, this finding influences the decision to implement a more rigorous RT_{60} matching procedure in octave band multiplane models as part of the final acoustic modelling study documented in section 5.3.

The EDT measurements recorded for the ODEON RIRs, displayed in Figure 5.27 share a similar trend to the T_{30} values in that they are consistently higher than those calculated from the multiplane and 3D FDTD models. This is reasonable given that EDT and T_{30} are measures of energy decay time and the FDTD models incorporate random-incidence absorption coefficients as previously discussed. It is also observed in Figure 5.27 that the EDTs calculated from the 2D multiplane model are, in general, lower than those derived from the 3D FDTD RIRs. Moreover, only one instance (RIR Case 1, 1 kHz octave band) of EDT measurement shows agreement between multiplane and 3D FDTD values whereby the multiplane EDT is within the JND range of 5%. This finding may be rationalised by considering the derivation of EDT values and the differences between RIR early reflection characteristics simulated by the multiplane and 3D FDTD models. As detailed in section 2.3.8, EDT is derived by extrapolating a best line of fit of the integrated energy decay curve in the region EDC = [0:-10] dB to an energy level of -60 dB. Hence, this measure is sensitive to the nature of both the direct sound component and early reflections recorded in an RIR. Previous time domain analysis comparing multiplane and 3D FDTD demonstrates the inherent difference between the

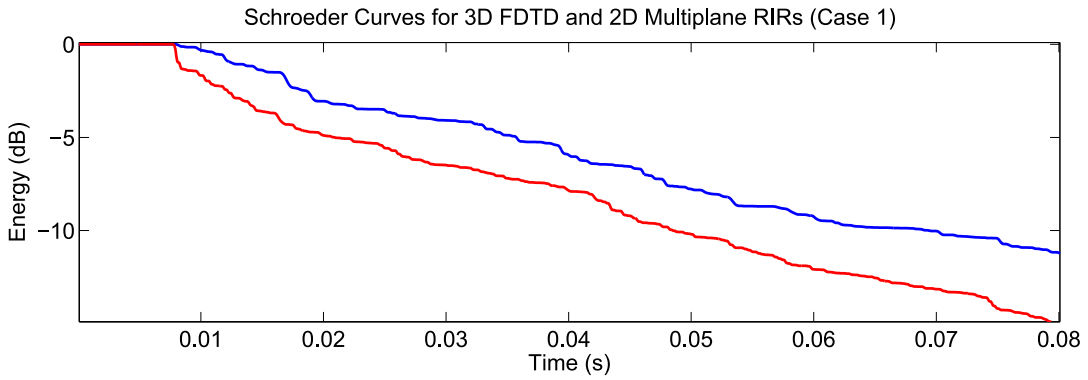

FIGURE 5.28: EDC (dB) for 3D FDTD (blue) and 2D multiplane (red) RIRs simulated for RIR Case 1. The initial stages of energy decay characteristics in the time region t = (0:80) ms is shown.

timing and pressure amplitude of the direct sound component and early reflections simulated by each modelling method. As such, it follows that the initial stages of the EDC will also differ as a result. Considering RIR Case 1, the early-time normalised EDC of 3D FDTD (blue) and 2D multiplane (red) RIRs is displayed in Figure 5.28. These EDCs are frequency-independent and are derived using the RIR analysis bandwidth f = [0:1420] Hz. From these curves, it is observed that the multiplane RIR possesses a quicker decrease in energy immediately after the direct sound component ($t \approx 8$ ms) compared to that recorded in the 3D RIR. This is the direct result of the difference between the pressure amplitudes of the direct sound component and following early reflections simulated by each model. Additionally, the multiplane RIR EDC shows a faster rate of energy decay over the range of EDC = [0:-10] dB. This results from the omission of early reflections that occur in 3D from the multiplane RIR as the 2D cross-sectional planes represent only a subset of reflection paths which are simulated in the 3D FDTD model. Hence, it is assumed that the 3D FDTD model RIR energy decay rate is comparatively slower as a greater temporal density of early reflections are inherently simulated. Consideration of these RIR characteristics explains, to good extent, the reason behind the on average lower EDTs calculated from the multiplane RIRs compared to the 3D FDTD RIRs which is particularly apparent for RIR Case 1.

Given the observed differences between T_{30} and EDT values calculated from each acoustic model, which are larger when comparing ODEON and FDTD RIRs, it is expected that Clarity, C_{80}, measurements will also differ. Figure 5.29 displays the C_{80} values calculated from the ODEON (black), 3D FDTD (blue) and multiplane (red) RIRs for the three source/receiver arrangements. In terms of the ODEON C_{80} values, it is observed that the Clarity of each RIR is consistently less than that of the FDTD RIRs. This result is due to the greater level of late arriving energy inherent to the ODEON model demonstrated by the comparatively longer T_{30} values as previously discussed.

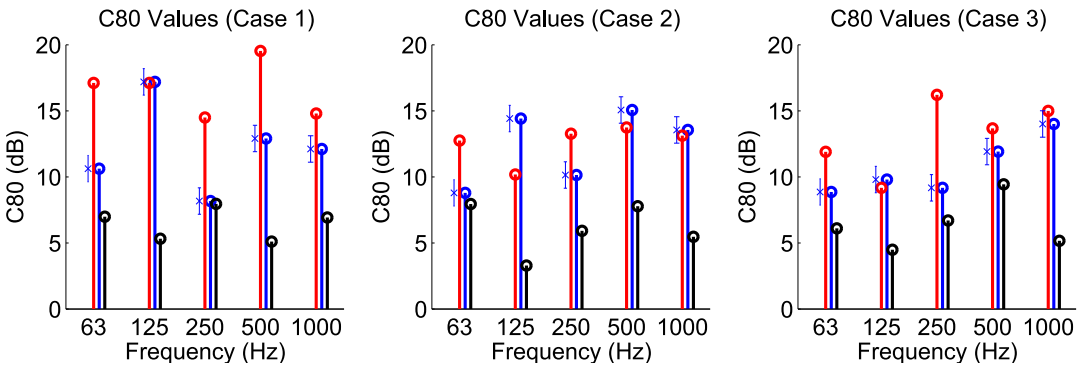

FIGURE 5.29: C_{80} (dB) values calculated from 3D FDTD (blue), 2D multiplane FDTD (red) and ODEON (black) RIRs (Cases 1 - 3) for the 63, 125, 250, 500, 1000 Hz octave bands. The error bars display the range of one JND (1 dB, after [17]) above and below the C_{80} value calculated from the 3D FDTD RIRs in each octave band.

Again, this difference in acoustic parameter measurement results from the application of random-incidence absorption coefficients to both geometric and SRL FDTD models. Referring to the Clarity measures calculated from the 2D multiplane FDTD RIRs, there is no discernible trend to the agreement of these values and those recorded from the 3D FDTD RIRs across all RIR cases. However, the *EDC*s previously presented for RIR Case 1 (Figure 5.28) demonstrate that more energy is concentrated in the direct sound component of the 2D multiplane RIR compared to the 3D RIR. This is observed by the large decrease in the multiplane *EDC* that occurs just after the passage of the direct sound component suggesting that a notable portion of the overall RIR energy is contained in this initial pressure fluctuation. The consequent effect on C_{80} measures is clear for RIR Case 1 in which the values for the multiplane RIR are, in general, greater than those calculated from the 3D RIR owing to a comparatively larger direct-to-diffuse sound energy ratio (as per (2.81)). Hence, the agreement of C_{80} measures recorded from multiplane and 3D FDTD RIRs is, as with EDT, shown to be sensitive to the inherent differences in early stage RIR characteristics previously discussed. This claim is further supported by the fact that the T_{30} values calculated from multiplane and 3D FDTD RIRs are in reasonable agreement and, therefore, the cause of differences in the early-to-late energy index C_{80} cannot be attributed to the nature of diffuse field energy decay characteristics.

It is apparent from this comparative examination of multiplane and 3D FDTD RIR acoustic parameters that agreement in objective T_{30} measures does not guarantee similar agreement in EDT and C_{80} values. Furthermore, variances in EDT and C_{80} result from differences in the simulation of early RIR characteristics inherent to the multiplane and 3D FDTD modelling methods. These findings suggest that a solution for

resolving these parametric differences will involve manipulation of the energy decay characteristics of the early stages of multiplane RIRs.

5.2.4.2 Low Frequency Analysis

As per the low frequency spectral analysis process applied in the pilot study, the live room acoustic model RIRs are examined here in terms of agreement of overall spectra magnitude levels, FRAC values (5.11) and magnitude similarity index, MI, (5.15). Analysis is performed over two different frequency regions. The first is defined as the modal region, f_{r1}, with the upper frequency limit set equal to the approximate Schroeder frequency $F_c \approx 280$ Hz, as per (2.77), giving $f_{r1} = [0{:}280]$ Hz. The second frequency analysis region is defined as $f_{r2} = [0{:}1000]$ Hz which is consistent with the value used in the pilot study in order to provide a basis of comparison between the results produced for the simple cuboid modelling scenario and the live room simulations.

To provide initial validation of the 3D FDTD model, the spectra of RIR Cases 1-3 are examined using Figure 5.30 (blue line) to investigate low order modal values. Referring to Figure 5.20, the major dimensions of the live room model are given as 4.11 x 3.35 x 2.6 m corresponding to length x width x height. Using these measurements, the theoretical 1st-order axial modal frequencies occurring in the x-, y- and z-directions may be calculated using (2.77) for simulated wave speed $c = 344$ ms^{-1} yielding 41.85, 51.34 and 66.15 Hz respectively (as denoted in Figure 5.30 by the green vertical lines). Distinct resonant frequency peaks are observed at these frequencies in the spectra of the RIRs simulated for Cases 1 and 2. The agreement between theoretical and simulated resonant frequency values for RIR Case 3 is less pronounced due to the locations of sound source and receiver. The agreement between theoretical modal values and those simulated using the 3D FDTD model demonstrates that the simulation of the live room is operating as required.

Also shown in Figure 5.30 are the magnitude spectra of RIRs simulated using ODEON which are denoted by the black line. As detailed in previous analysis of acoustic parameters, the ODEON RIRs possess longer energy decay times than those recorded for the corresponding 3D FDTD RIRs. This is reflected in the resulting magnitude spectra by the fact that the ODEON low frequency RIR energy levels are higher (post-energy calibration) than those observed in the pilot study (see e.g. Figure 5.13) relative to the 3D FDTD RIRs. Hence, it is demonstrated that differences in energy decay rates impact on the agreement of resulting RIR spectra magnitude levels. This finding is given further consideration in the final modelling scenario documented in section 5.3. Nonetheless, for the purposes of visual comparison of geometric and 3D FDTD

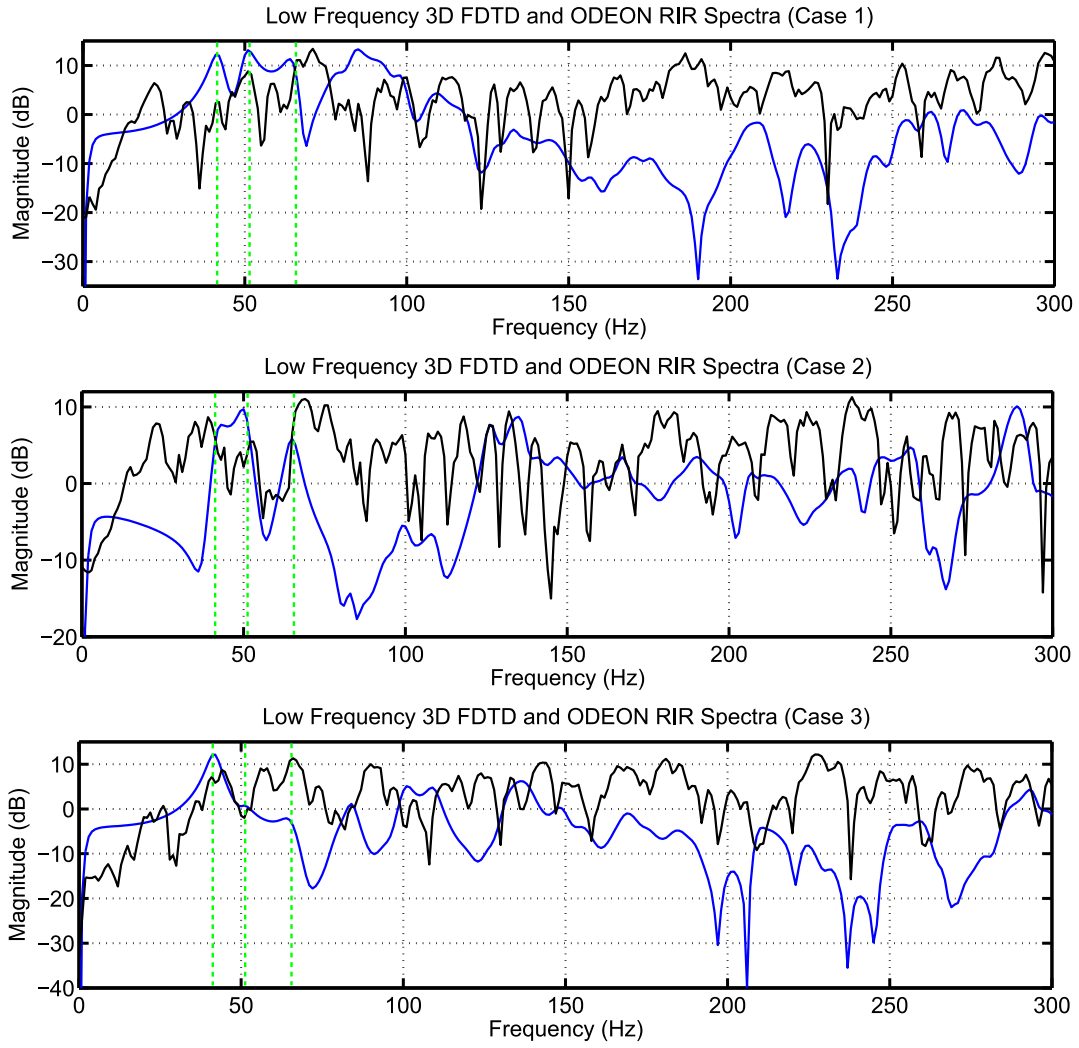

FIGURE 5.30: (Top to Bottom) Low frequency spectra of 3D FDTD (blue line) and ODEON RIR (black line) Cases 1-3. The frequency range shown includes the modal region of f_{r1} = (0:280) Hz.

RIR spectra, it is sufficient to conclude from Figure 5.30 that little agreement exists between the low frequency characteristics inherent to each model. In particular, overall magnitude level trends observed in the 3D FDTD RIRs are not represented in the ODEON spectra. This is most apparent when considering the notable dips in 3D FDTD spectra magnitude in the following regions: Case 1 - $f \approx$ [100:180] Hz; Case 2 - $f \approx$ [70:110] Hz; Case 3 - $f \approx$ [190:240] Hz. In each RIR case, these magnitude trends are not represented in the ODEON spectra. Furthermore, the positioning of magnitude peaks observed in ODEON RIR spectra for all RIR cases is observed as largely sporadic when compared to the 3D FDTD spectra. This is particularly apparent in the lower half of the modal region (f = [0:140] Hz) in which the modal frequencies simulated by the 3D FDTD model are not distinctly represented by the geometric modelling method due to the inherent limitations of ray-based simulation techniques.

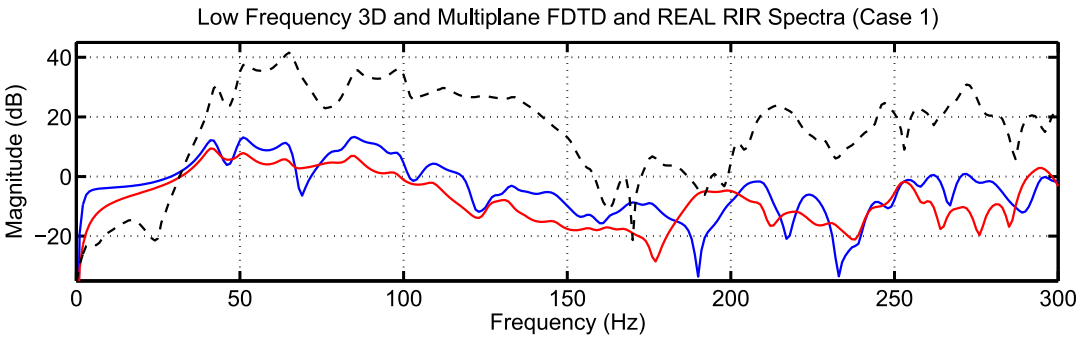

FIGURE 5.31: Low frequency spectra of recorded ('REAL' - black dashed line), 3D FDTD (blue line) and 2D multiplane (red line) Case 1 RIRs. The frequency range shown includes the modal region of $f_{r1} = (0{:}280)$ Hz.

The examination of low frequency analysis continues here by considering the modal region of the RIRs generated for each RIR case in turn. FRAC and MI values relating to the ODEON model are provided throughout for completeness while noting that these values are affected by the differences in energy decay characteristics observed between ODEON and FDTD RIRs, as previously discussed. Figure 5.31 displays the measured ('Real' - black dashed line) and simulated 3D (blue) and 2D multiplane (red) FDTD spectra for RIR Case 1. As shown, the 'Real' RIR obtained through practical measurement of the live room exhibits greater magnitude levels than the simulated RIRs. This is due to the differences between the sound source strength applied in practice and that used to excite the virtual acoustic field of each acoustic model. The low frequency energy enhancement of the practical omnidirectional sound source approximation also contributes to the magnitude offset. However, the purpose of this analysis is to display and examine the agreement between positioning of resonant peaks and overall magnitude level trends from both measurements and simulations. Hence, the spectra are presented as they are to make clear these aspects of comparison.

Comparison of both 3D and multiplane FDTD RIR spectra with the measured response spectrum demonstrates that both modelling approaches accurately simulate the first three axial modes ($f \approx [40{:}70]$ Hz). However, the magnitude levels of the corresponding peaks are lower in the multiplane spectrum compared to the 3D RIR. This level difference is also apparent over the range $f \approx [80{:}150]$ Hz. Nonetheless, the multiplane spectra is shown to contain modal peaks that are coincident with those present in the 3D model and measured RIRs in this range. This further demonstrates the capability of the multiplane model to emulate low frequency wave phenomena occurring in the 3D geometry. Furthermore, the decreasing magnitude level trend observed in the measured RIR for $f = [100{:}160]$ Hz is followed to a similar extent by both 3D and 2D multiplane models. In the region $f = [160{:}200]$ Hz, there is a notable reduction in the

agreement between simulated and measured RIRs whereby spectral notches observed in the measured RIR are not sufficiently represented by the FDTD models. Also, the positioning of spectral notches in the multiplane and 3D RIRs in this region are not coincident indicating that the cross-sectional representation of 3D wave propagation is not accurately emulating wave cancellation effects that are arising in the 3D FDTD model. However, as frequency increases towards the upper limit of the modal range, the coincidence of resonant peaks observed in the measured and 3D FDTD simulation RIRs regains stronger agreement. This is also the case for the 2D multiplane RIR spectra which contains magnitude peaks that are coincident with those present in the 3D RIR spectra ($f \approx$ [240:300] Hz), although for a subset of these peaks a notable difference in magnitude level is observed.

Overall, it is shown that both 3D and 2D multiplane FDTD share similar magnitude level trends to those present in the measured RIR spectra over the modal region. Low order modal frequencies are well represented in both simulated RIRs and magnitude peaks present in the 3D RIR are, on the most part, observed in the multiplane RIR spectra.

Model (vs. Real)	Case	FRAC	MI
3D FDTD Model	1	0.145	0.753
2D Multiplane FDTD	1	0.069	0.706
ODEON	1	0.061	0.659
Model (vs. 3D FDTD)	**Case**	**FRAC**	MI
2D Multiplane FDTD	1	0.744	0.839
ODEON	1	0.050	0.748

TABLE 5.13: (Top) FRAC and MI values calculated to compare the agreement of 3D FDTD, 2D multiplane FDTD and ODEON RIR spectra with the measured ('Real') RIR spectrum (Case 1) in the modal region f_{r1}. (Bottom) FRAC and MI values calculated to compare the agreement of 2D multiplane and ODEON RIRs with the 3D FDTD RIR (Case 1) in the modal region f_{r1}.

In order to quantify the agreement of RIR spectral characteristics, both FRAC and MI values are calculated using the measured and simulated RIRs. Table 5.13 displays the FRAC and MI values derived when comparing each simulated RIR spectra to the measured response ('vs. Real') spectra for the modal region f_{r1}. Note that prior to FRAC and MI analysis, the measured RIRs are calibrated to the corresponding 3D FDTD RIR in terms of energy by means of the energy matching procedure applied to the multiplane RIR. This is to ensure that the large difference in recorded energy levels observed through visual inspection of RIR spectra has less impact on calculated

results. It is observed that the FRAC values calculated for each simulated RIR are significantly lower than the MI values which, by way of contrast, suggest reasonable levels of agreement between real and modelled RIR spectra.

As previously discussed in relation to the RIRs examined in the pilot study, this result suggests a disagreement in the phase and temporal characteristics of real and modelled RIRs, to which the MI calculation is less sensitive than FRAC analysis. However, for the purposes of this analysis it is sufficient to compare the resulting FRAC and MI values calculated from each modelling approach. In doing so, it is observed that the 3D FDTD model produces the most accurate representation of the measured RIR given that both FRAC and MI values are higher than those calculated from multiplane and ODEON RIRs. This is as expected given the higher level of accuracy inherent to the 3D numerical model at low simulation frequencies. Additionally, the multiplane MI value, although notably less than that calculated for the 3D RIR, constitutes an encouraging result in light of the reduction in computation achieved by this modelling approach compared to full 3D FDTD simulations.

Also provided in Table 5.13 (bottom half) are the FRAC and MI values calculated by comparing ODEON and multiplane Case 1 RIRs to the 3D RIR in the modal region. Note that the MI recorded for the multiplane RIR is consistent with the high values presented in the pilot study ($MI \geq 0.83$) for the case of a cuboid room geometry. This indicates that the extended three plane model applied in this case study is capable of simulating low frequency sound propagation in a more complex 3D space, providing good agreement with 3D FDTD RIR results. This claim is further supported by the associated FRAC value which, at a level of 0.744, is notably greater than those produced using the two plane modelling approach in the pilot study (FRAC = [0.260:0.445]). Finally, the 2D multiplane RIR is shown to produce a higher level of agreement with the measured and 3D FDTD RIRs than the ODEON RIR. This is particularly apparent when comparing MI values and the FRAC values resulting from comparison with the 3D FDTD RIR.

The low frequency magnitude spectra of measured and simulated RIRs for RIR Case 2 are provided in Figure 5.32. From visual inspection of the magnitude curves it is again observed that low order modal resonances are correctly simulated by both FDTD models ($f \approx$ [40:85] Hz). However, differences between the magnitude levels of the corresponding peaks in the 3D and 2D multiplane RIR spectra are also evident. As with the analysis of RIR Case 1, the multiplane model is shown to simulate a subset of resonant frequencies that are observed in the 3D FDTD RIR spectrum across the modal frequency range. It is also noted that, due to the change in source and receiver locations compared to RIR Case 1, the overall magnitude level trends in the presented

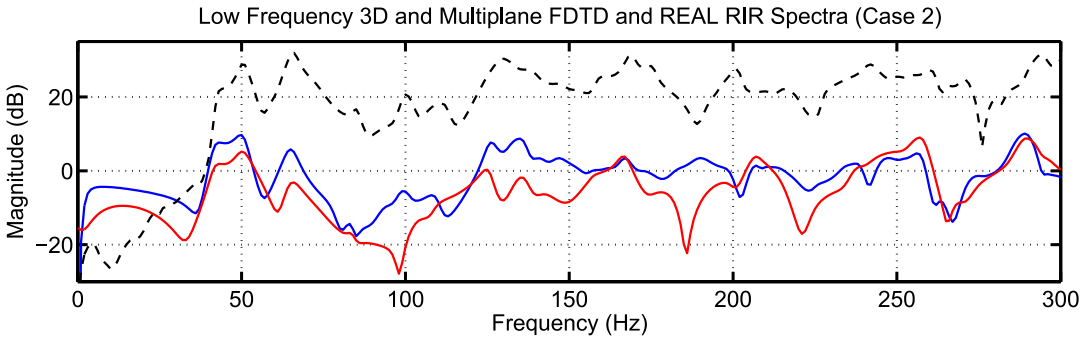

FIGURE 5.32: Low frequency spectra of recorded ('REAL' - black dashed line), 3D FDTD (blue line) and 2D multiplane (red line) Case 2 RIRs. The frequency range shown includes the modal region of f_{r1} = (0:280) Hz.

spectra differ from those displayed in Figure 5.31. Notable differences include dips in the magnitude levels of all RIR spectra in the regions $f \approx$ [75:120] Hz and $f \approx$ [210:240] Hz. Examination of these spectral features supports the claim made previously (see section 5.1.4) that the multiplane model is capable of representing the changes in the nature of the low frequency soundfield as the source and receiver positions are varied. This result is consistent with the findings drawn from analysis of multiplane RIRs presented in the pilot study (see section 5.1.4).

FRAC and MI values derived through comparison of each simulated RIR with the measured response are provided in Table 5.14 (top - 'vs. Real'). These results are not entirely consistent with those recorded for RIR Case 1. For example, the 2D multiplane RIR achieves a higher FRAC value than the 3D RIR. This result is unexpected since the 3D FDTD model is assumed to provide higher accuracy than the multiplane model as it preserves a complete representation of 3D wave motion. However, given that both 3D and multiplane FRAC values are reasonably low (FRAC < 0.26) it is not possible to conclude that the multiplane RIR is significantly more accurate than the 3D RIR in terms of low frequency sound simulation. This reasoning is further supported by the MI values which demonstrate that the 3D FDTD RIR spectrum shares greater similarity with the measured response than the multiplane RIR spectrum.

A further interesting result is observed from the MI value calculated when comparing the ODEON and measured RIRs. As shown, this value is higher than the MI derived from the multiplane RIR which suggests that the ODEON RIR spectrum shares a higher level of similarity with the measured response spectrum. However, this result is not supported by visual inspection of the ODEON RIR spectra (see Figure 5.30) which demonstrates a lack of consistency with simulated 3D FDTD RIR magnitude trends. Additionally, the associated ODEON FRAC value (FRAC = 0.089) is notably less than that given for the multiplane RIR. Hence, the ODEON MI value alone is

Model (vs. Real)	Case	FRAC	MI
3D FDTD Model	2	0.181	0.775
2D Multiplane FDTD	2	0.256	0.714
ODEON	2	0.089	0.766

Model (vs. 3D FDTD)	Case	FRAC	MI
2D Multiplane FDTD	2	0.531	0.848
ODEON	2	0.024	0.748

TABLE 5.14: (Top) FRAC and MI values calculated to compare the agreement of 3D FDTD, 2D multiplane FDTD and ODEON RIR spectra with the measured ('Real') RIR spectrum (Case 2) in the modal region f_{r1}. (Bottom) FRAC and MI values calculated to compare the agreement of 2D multiplane and ODEON RIRs with the 3D FDTD RIR (Case 2) in the modal region f_{r1}.

not evidence enough to support the claim that the representation of the low frequency measured response provided by the geometric model is, in this instance, more accurate than that simulated in the multiplane model.

FRAC and MI values pertaining to the comparison of multiplane and ODEON RIRs with the 3D FDTD RIR in the modal region are provided in the lower half of Table 5.14. As shown, the MI value calculated for the multiplane RIR demonstrates good agreement between 3D and multiplane RIR spectra ($MI = 0.848$). Furthermore, the FRAC value obtained by comparing multiplane and 3D RIRs, while less than that recorded for RIR Case 1, is still higher than those presented in the pilot study. This again demonstrates the improvement in multiplane model accuracy provided by the incorporation of an additional cross-sectional plane.

Referring to Figure 5.33, the agreement between modal values simulated by 3D (blue) and multiplane (red) FDTD models for RIR Case 3 is shown to be high up to $f \approx 110$ Hz. In turn, the positioning of the corresponding magnitude peaks are is in reasonable agreement with those present in the measured RIR spectrum in this frequency range. However, the level of similarity between modelled and measured RIR spectra decrease as frequency increases towards the upper limit of the modal region. In particular, the the multiplane model simulates a notable magnitude notch at $f \approx 175$ Hz which is not represented in 3D FDTD or measured RIRs. This suggests that wave cancellation effects, which are not consistent with the simulated and measured 3D soundfield, are occurring in the multiplane model at this frequency. Additionally, the multiplane RIR spectrum demonstrates omission of the defined magnitude peaks observed in the 3D FDTD and measured RIRs in the range $f = [200{:}250]$ Hz. However, the overall magnitude trends of both multiplane and 3D FDTD RIRs are shown to be in reasonable

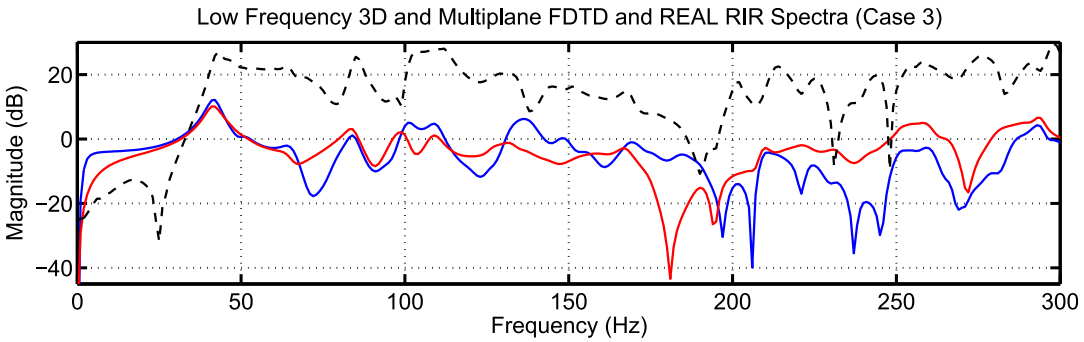

FIGURE 5.33: Low frequency spectra of recorded ('REAL' - black dashed line), 3D FDTD (blue line) and 2D multiplane (red line) Case 3 RIRs. The frequency range shown includes the modal region of f_{r1} = (0:280) Hz.

agreement for $f > 250$ Hz. Hence, for the majority of the modal region, the multiplane model is again demonstrated to be capable of emulating 3D low frequency wave phenomena. This claim is supported by the associated FRAC and MI values listed in Table 5.15.

Model (vs. Real)	Case	FRAC	MI
3D FDTD Model	3	0.062	0.737
2D Multiplane FDTD	3	0.044	0.800
ODEON	3	0.007	0.679

Model (vs. 3D FDTD)	Case	FRAC	MI
2D Multiplane FDTD	3	0.509	0.834
ODEON	3	0.030	0.748

TABLE 5.15: (Top) FRAC and MI values calculated to compare the agreement of 3D FDTD, 2D multiplane FDTD and ODEON RIR spectra with the measured ('Real') RIR spectrum (Case 3) in the modal region f_{r1}. (Bottom) FRAC and MI values calculated to compare the agreement of 2D multiplane and ODEON RIRs with the 3D FDTD RIR (Case 3) in the modal region f_{r1}.

The MI values calculated by comparing simulated Case 3 RIRs with the measured RIR show that the multiplane model, in this instance, achieves the highest similarity with the measured RIR in terms of spectral characteristics in the modal region. FRAC values calculated for all simulated RIRs demonstrate poor correlation between model results and those obtained from practical measurements. Hence, in light of these findings, in combination with those presented for RIR Cases 1 and 2, it may be concluded that MI and FRAC analysis comparing measured and simulated RIRs has yielded variable results throughout this case study and this is considered further in the concluding remarks in section 5.2.5. Referring to the MI and FRAC values calculated

from comparison of 2D multiplane and 3D FDTD RIRs, indicate an encouraging agreement between the spectra simulated by each model in the modal region. This result is consistent with those observed for RIR Cases 1 and 2 ($MI > 0.83$, FRAC > 0.5).

Model (vs. 3D FDTD)	Case	FRAC	MI
2D Multiplane FDTD	1	0.204	0.832
ODEON	1	0.013	0.739
2D Multiplane FDTD	2	0.288	0.848
ODEON	2	0.022	0.823
2D Multiplane FDTD	3	0.064	0.827
ODEON	3	0.005	0.750

TABLE 5.16: FRAC and MI values calculated to compare the agreement of 2D multiplane and ODEON RIRs with the 3D FDTD RIRs (Cases 1 - 3) in the simulated low frequency region f_{r2}.

As a final stage of low frequency analysis, the MI and FRAC values calculated by comparing 2D multiplane and ODEON RIRs with the 3D FDTD RIRs (Cases 1 - 3) over the frequency range f_{r2} are provided in Table 5.16. These results indicate that the multiplane RIRs provide a higher level of agreement with 3D FDTD RIRs than the ODEON RIRs in the low frequency simulation region.

5.2.5 Concluding Remarks and Considerations

The results and discussion presented in this case study provide insight into the capability of the extended three plane multiplane FDTD model as a means of simulating low-to-mid frequency sound propagation in a realistic acoustic environment. Three practical RIR measurements taken from a recording studio live room have been simulated by means of ODEON, 3D FDTD and 2D multiplane acoustic models. Acoustic parameters and spectral characteristics derived from each simulated RIR have been examined by comparative means leading to the following findings and considerations:

- The application of random-angle-of-incidence absorption coefficients to both geometric and FDTD acoustic models gives rise to large differences in the energy decay characteristics simulated by each modelling approach. Incorporation of normal-incidence absorption coefficients in FDTD models is investigated in the next section in order to remedy the consequent differences observed in T_{30} measures.

- T_{30} measures derived from 3D and 2D multiplane models indicate that the approach to multiplane reflection coefficient calibration applied in this study has been successful. This claim is supported by the fact that multiplane T_{30} times are within one JND of those derived from the 3D FDTD model for all RIR cases.

- Examination of EDT values derived from multiplane and 3D FDTD RIRs demonstrate that there are notable differences in the early-time energy decay characteristics inherent to each model. This, in turn, leads to notable differences in C_{80} values calculated for each model. Hence, it is postulated that manipulation of the early-time multiplane RIR amplitude may provide a means of reducing the observed differences in energy decay characteristics in the initial stages of simulated 3D and multiplane FDTD RIRs.

- Analysis of FRAC and MI measures calculated by comparing 2D multiplane and 3D FDTD RIRs in the modal region f_{r1} show that the extended three plane multiplane model is capable of producing a higher correlation to the 3D model, in terms of spectral characteristics, than the two plane implementation applied in the pilot study.

- Visual inspection of simulated and measured RIR spectra provides further evidence that the 2D multiplane FDTD acoustic model is capable of representing a subset of the spectral characteristics observed in 3D FDTD RIRs.

The FRAC and MI values which compare geometric and multiplane RIRs against 3D FDTD RIRs show that the multiplane model consistently provides a closer approximation to the 3D model than the geometric model. However, low frequency analysis conducted through examination of FRAC and MI values obtained by comparing simulated and measured RIRs does not lead to such a conclusion given the varied nature of the results presented. It is assumed that this variation in these simulation vs. measured RIR FRAC and MI values is due to the fact that the case study acoustic models are not equipped to represent the following aspects of real enclosed sound propagation:

- The frequency response and, therefore, frequency-dependent directivity of the loudspeaker and microphone used to record practical RIRs.

- The phase response of boundary surfaces and the consequent effects on the phase of reflected sound components.

- Strong vibrations of boundary surfaces due to strong modal resonances at low frequencies.

- The frequency-dependent absorption characteristics associated with real surface materials.

In light of this, the assessment of 2D multiplane FDTD acoustic modelling through comparison of measured and simulated RIR data is not extended to the final study presented in this chapter. However, a range of considerations revealed through completion of this case study are taken forward to the final study in order to inform the refinement of multiplane modelling aspects including RT_{60} calibration and appropriate sound source modelling.

5.3 Final Study: The PTB Room

The acoustic environment selected for this third and final study is a complex recording studio live room located in the Physikalisch-Technische Bundesanstalt in Braunschweig, Germany. For brevity this live room is referred to as the *PTB room* in the following. This space, which presents a complex room acoustic modelling scenario due to its geometric features, has been used in related research as a means of benchmarking virtual acoustic modelling methods [161]. As such, it is considered appropriate for the purposes of further testing the capabilities of 2D FDTD multiplane modelling as a means of low frequency RIR synthesis. By way of introducing the geometry of the PTB room, graphical depictions of the major boundary surfaces are provided in Figure 5.34. As shown, this space incorporates diffusing surfaces located on one of the walls and also on the ceiling. Further features include windows, skylights, flat fibreboard roof sections and indented wooden panelling. Hence, in comparison to the previous two acoustic modelling scenarios presented in this research, the PTB room is by far the most complex in terms of geometry. Additional details of overall room dimensions are provided in the following in relation to the geometric, 3D FDTD and 2D FDTD multiplane models created to synthesise RIRs for this space.

5.3.1 Geometric Model

CATT-Acoustic v. 8.0 acoustic prediction and auralisation software [19] is used to simulate the soundfield of the PTB room by means of geometric modelling techniques. This software is widely regarded to be of industry standard and is incorporated into this work to ensure that comparative analysis of simulated RIR data is not limited to a only one geometric acoustic modelling program. As with ODEON [21], CATT uses a hybrid geometric acoustic modelling technique that applies the ISM to early reflections and a ray-tracing algorithm to high-order reflections for diffuse soundfield simulation.

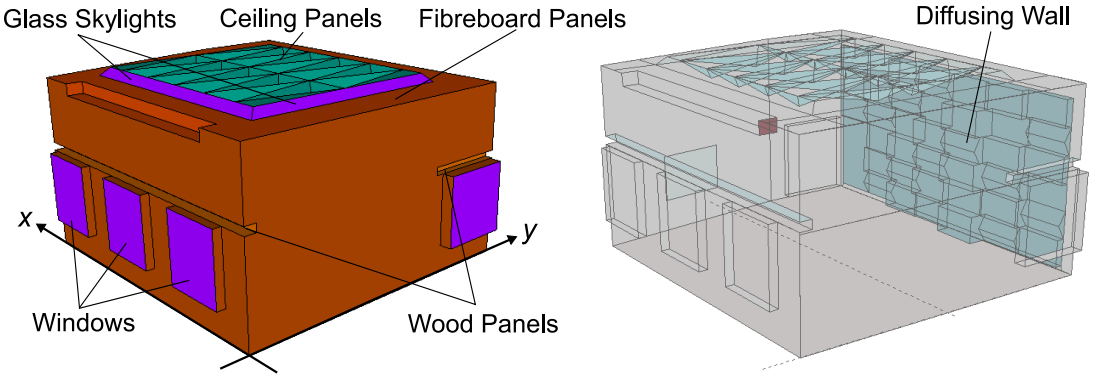

FIGURE 5.34: The 3D PTB Room geometry as viewed in Sketchup Make [18] (left) and CATT-Acoustic v. 8.0 [19] (right). Major features of the room geometry are highlighted.

However, the "Full Detailed Calculation" option (see e.g. [162]) applied in CATT-Acoustic to simulate the PTB geometric RIRs incorporates a *randomised tail-corrected cone-tracing* (RTC) modelling algorithm.

RTC uses the ISM to trace 1st- and 2st-order specular reflection paths that exist between source and receiver via each surface incorporated in the room geometry. Simulation of diffusely reflecting components is then handled by combining two different ray-based approaches. Initial (or "1st-order") diffuse reflections from diffusing surfaces are modelled by representing the radiation of scattered rays as a number of secondary sources positioned on each surface. In this way, a large number of low energy diffuse reflections are cast into the space in order to emulate attributes of the transition from early-to late-time RIRs as occurs in practice [163]. High-order reflections are simulated using a randomised cone tracing technique. As with beam-tracing approaches, cone-tracing seeks to establish valid reflection paths between source and receiver by tracing a volume rather than a 1D ray-path. However, as opposed to beam-tracing in which a single specular reflection path is established, cone-tracing allows rays to propagate freely within the traced volume providing a similar result to that generated by means of standard ray-tracing techniques.

The geometry of the PTB room is sourced from an existing ODEON Auditorium 10.1 model [21] which includes a .par file defining the surfaces of the detailed room model, after [161]. This .par file is exported to the .geo format required by CATT-Acoustic. The .geo is then updated to include the frequency-dependent absorption characteristics of the surface materials incorporated in the PTB model. These coefficients, which are provided in Table 5.17, are also sourced from the ODEON model as per [161]. Note, α_i

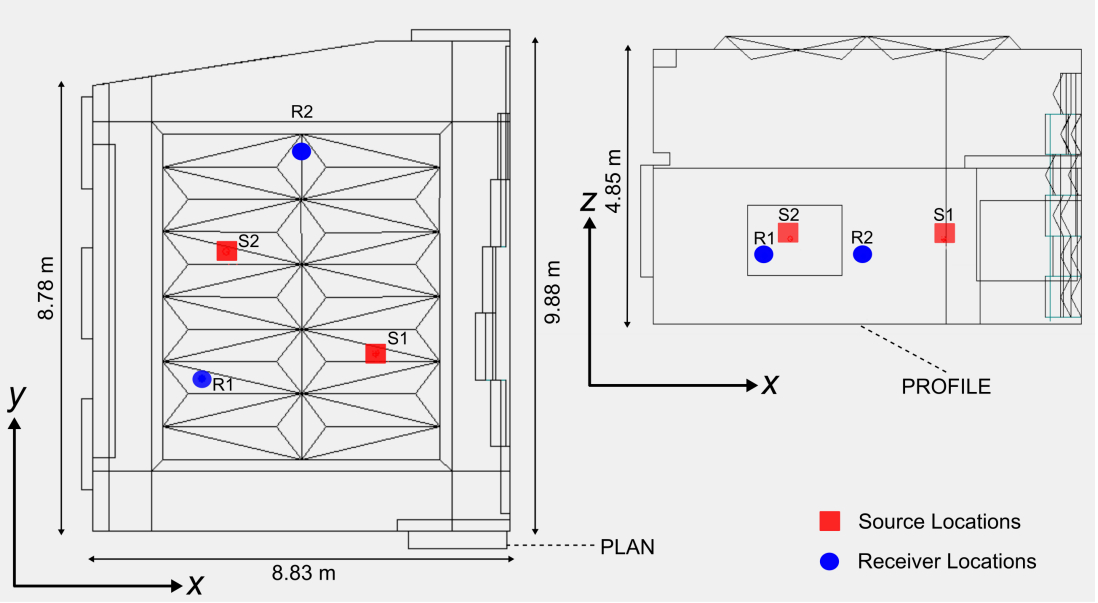

FIGURE 5.35: Plan (left) and profile (right) view of source (S1, S2) and receiver (R1, R2) locations defined in the PTB room acoustic model (see also, Table 5.18). Major dimensions of model geometry are also shown.

Surface	α_{63}	α_{125}	α_{250}	α_{500}	α_{1k}	α_{2k}	α_{4k}
Floor	0.04	0.04	0.04	0.07	0.06	0.06	0.07
Walls	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Windows	0.10	0.10	0.04	0.03	0.02	0.02	0.02
Wooden Panels	0.20	0.20	0.05	0.03	0.01	0.01	0.01
Ceiling	0.30	0.30	0.20	0.06	0.02	0.02	0.02
Fibreboard Panels	0.42	0.42	0.28	0.49	0.78	0.58	0.62
Diffusing Wall Panels	0.40	0.40	0.33	0.21	0.16	0.15	0.16

TABLE 5.17: Octave band absorption coefficients for each material type incorporated in the CATT-Acoustic PTB room model, as sourced from ODEON 10.1 Auditorium [21].

corresponds to the random-angle-of-incidence absorption coefficient of a material for the i Hz octave band.

Three RIR cases, corresponding to source/receiver locations given in [161], are defined for this study. These are listed in Table 5.18 and displayed in Figure 5.35. As shown, two sources (S1 and S2) and two receivers (R1 and R2) are used to simulate the three RIR cases with the following source/receiver combinations: S1R1; S1R2; S2R2. For consistency with the FDTD simulations, source and receiver directivity characteristics are defined as omnidirectional. Each RIR is simulated using ~16000 ray paths as set by CATT-Acoustic for the detailed calculation option previously described. Simulation run-times are recorded as approximately 1500 s per RIR, noting that the entire simulation bandwidth (all octave bands) is rendered in this time with $F_s = 44.1$ kHz.

Case	Source	Source Location (x,y,z) m	Receiver	Receiver Location (x,y,z) m
1	S1	(3.85, 5.93, 1.5) m	R1	(3.35, 2.43, 1.2) m
2	S1	(3.85, 5.93, 1.5) m	R2	(7.85, 4.43, 1.2) m
3	S2	(5.85, 2.93, 1.2) m	R2	(7.85, 4.43, 1.2) m

TABLE 5.18: Sound source (S) and receiver (R) locations defined for three RIR measurement cases in the PTB room model.

5.3.2 3D FDTD Model and Hybrid RIR Synthesis

The three RIR cases detailed in the previous section are simulated using a 3D SRL FDTD model created by voxelising the PTB room geometry as defined in ODEON [21]. This process is conducted in two stages. Firstly, a Sketchup Make [18] 3D CAD model of the PTB room geometry, created by Dr. Alex Southern [164] as part of the research conducted on hybrid acoustic modelling [81] (see section 3.3), is checked for consistency with the ODEON .par geometry and exported as a .vtk file using [165]. This .vtk geometry file is then voxelised using the WaveModeller spatial discretiser [166] to produce an SRL lattice of pressure nodes with a spatial sampling index of h_{3D} = 0.0105 m, corresponding to wave speed $c = 343.42$ ms^{-1} and temporal sampling rate $F_s = 56800$ Hz. The selected temporal sampling rate ensures that the 3D FDTD model simulations possess a usable bandwidth up to and including the 2 kHz octave band with an upper cut-off frequency of $f_u = 2.84$ kHz while maintaining isotropy error of <2% and so minimising dispersion error. This follows from the condition $f_u \leq 0.075F_s$ as detailed in section 4.2. Additionally, the internodal distance between pressure nodes in the SRL lattice imposes a maximum deviation from the geometric acoustic model dimensions of $\sqrt{3h_{3D}^2}/2 = 0.0091$ (m) provided dimensions are rounded to the nearest spatial sampling instance. The same error is also applicable to source and receiver locations compared to the geometric model and is considered negligible.

The resulting SRL lattice for the PTB room is very complex compared with the pilot and case study models and includes a large number of re-entrant edges and corners. Hence, in accordance with the investigation into the numerical stability of 3D SRL FDTD models with LRS boundaries (see section 4.4.4), the PTB room simulations are unstable when applying pressure-centered LRS boundaries to model surfaces. In contrast, implementation of velocity-centered boundaries, using the formulation documented in [130], yields stable simulations for this modelling scenario. Furthermore, in light of the large differences between energy decay characteristics of the geometric and FDTD models documented in the case study, the reflection coefficients applied to each surface of the 3D FDTD model are defined by converting the random-angle-of-incidence

absorption coefficients listed in Table 5.17 to normal-incidence coefficients. This process is carried out using the following relationship between random- and normal-angle-of-incidence reflection coefficients, developed by Siltanen [22] as used in [81]:

$$R_{Norm} = -0.22897R_{Rand}^2 + 0.99298R_{Rand} + 0.23769 \tag{5.18}$$

which is derived by interpolating the range of normal-incidence coefficients R_{Norm} = [0:1] resulting from the normal- to random-incidence absorption coefficient conversion equation given in [3] for R_{Rand} = [0:1]. Table 5.19 provides the boundary impedance ζ_i values applied for each surface material type as derived from R_{Norm}:

$$\zeta_i = \frac{1 + R_{Norm}}{1 - R_{Norm}} \tag{5.19}$$

where subscript i refers to the i Hz octave band.

Surface	ζ_{63}	ζ_{125}	ζ_{250}	ζ_{500}	ζ_{1k}	ζ_{2k}
Floor	216	216	216	112	134	134
Walls	539	539	539	539	539	539
Windows	75	75	216	310	539	539
Wooden Panels	34	34	166	310	2023	2023
Ceiling	21	21	24	134	539	539
Fibreboard Panels	13	13	23	11	5	8
Diffusing Wall Panels	14	14	18	32	44	47

TABLE 5.19: Octave band normal-incidence impedance coefficients for each material type incorporated in the 3D FDTD PTB room model (rounded to the nearest whole number). These coefficients are derived from absorption data defined in ODEON 10.1 Auditorium [21] by means of random- to normal-angle-of-incidence absorption coefficient conversion [22].

The 3D FDTD RIRs are simulated by means of the constant F_s octave band procedure detailed in section 4.5.2.1 to facilitate frequency-dependent LRS boundary conditions. This approach is selected to avoid both the prohibitive memory requirements of DIF-based LRS implementations and the energy matching process as needed for the variable F_s method (see section 4.5.4.4). As such, one full bandwidth simulation is required per octave band per RIR case with the impedance values given in Table 5.19 applied to surface pressure nodes as appropriate. For each simulation, two instances of a 977 x 824 x 488 lattice of pressure nodes is needed to store the target acoustic field for F_s = 56.8 kHz. This corresponds to a memory requirement of ∼5.86 GB which leads to prohibitive simulation run-times for CPU-based computation. For this reason, the 3D FDTD RTB room model is simulated by Dr. Craig J. Webb of the NESS Project [125] at Edinburgh University (European Research Council Grant: StG-2011-279068-NESS) using a parallel GPU. Note that the code required to parallelise the 3D SRL FDTD

update scheme such that it could be run on this GPU is authored by Dr. Craig J. Webb as listed on the accompanying CD. This process facilitates simulation of the PTB room at reduced run-time, compared to that which would be required when using a single CPU only. Recorded simulation run-times are in the region of 2800 s per second of RIR output ($F_s = 56.8$ kHz).

Upon simulation of each octave band for the three RIR cases, the octave band RIRs are isolated using a 3$^{\text{rd}}$-order zero-phase bandpass IIR filter bank with Butterworth characteristics. Cut-on and cut-off frequencies of the filters applied are defined using the lower and upper frequencies of the octave bands (see Table 4.2). The filtered octave band RIR contributions for each RIR case are then summed to produce the spectrally complete FDTD RIR. Finally, these complete RIRs are calibrated and combined with the CATT-Acoustic RIRs following the procedure defined in section 5.1.2.1 (see Figure 5.5 for reference).

5.3.3 2D Multiplane FDTD Model and Hybrid RIR Synthesis

Having voxelised the PTB room geometry into a 3D SRL lattice of pressure nodes, the multiplane model is defined by selecting the cross-sections of the lattice which contain the receiver location and are orientated in the x-y, x-z and y-z Cartesian planes. This is carried out for both receiver locations detailed in Table 5.18 resulting in two different multiplane models for simulating the acoustics of the PTB room. In order to maintain consistency between FDTD model geometries, the 2D multiplane model is simulated at a lower F_s to compensate for the inter-nodal distance h_{3D} used in the 3D lattice. To this end, the relationship $h_{2D} = \sqrt{2}ck$ is rearranged to yield:

$$F_s = \frac{\sqrt{2}c}{h_{2D}} \tag{5.20}$$

giving $F_s = 46377$ Hz for $h_{2D} = h_{3D}$ and $c = 343.42$ ms^{-1}. As such, the multiplane model is simulated at the Courant limit for a wave speed and domain geometry that are exactly equal to those of the 3D FDTD model. Note that the calculated F_s is in accordance with the condition $f_u = 2840 \leq 0.1\ F_s$, as defined in section 4.2, that ensures minimal isotropy and dispersion error in the simulation bandwidth used in this study. Moreover, the maximum difference between the PTB room dimensions and those of the multiplane model is calculated as $\sqrt{2h_{2D}^2}/2 = 0.0074$ m which also applies to discrete source/receiver locations and is considered negligible.

The target acoustic field of each 2D FDTD cross section is excited using a Kronecker delta function applied via the soft source excitation method. As with the multiplane

simulations previously discussed, RIR outputs produced by recording the pressure fluctuations at the receiver node are filtered to remove both DC offset and afterglow effects. Each RIR simulation requires 3367872 pressure nodes to store the cross-sectional acoustic field which corresponds to a memory requirement of $\sim$ 0.025 GB. As such, the multiplane simulations achieve a reduction in memory requirement of $\sim$ 99.5% compared to the 3D FDTD model. However, this statistic must account for the fact that the multiplane model is simulated at a lower temporal sampling rate than the 3D model. To better quantify the difference in memory requirements, the number of sampling instances required by the 2D multiplane model for $F_s = 56.8$ kHz may be calculated for simulations operating at the Courant limit. In this case, $h_{2D} = 0.0086$ m for $c = 343.42$ ms^{-1} giving three cross sections of 1197 x 1009, 1197 x 598 and 1009 x 598 pressure nodes. This gives a total of 5050332 nodes required to store two instances of the target acoustic field as needed. In turn, this corresponds to a memory requirement of $\sim$ 0.038 GB which still provides a substantial memory reduction of 99.36% against the 3D FDTD model. Additionally, CPU-based simulations of the PTB multiplane model record run-times of approximately 230 seconds per 1 second of RIR output. Hence, even without the benefits of accelerated simulations realised through use of GPUs, the multiplane model attains significant reductions in run-times compared to parallelised 3D FDTD simulations. The C source code file scripted to simulate the 2D multiplane FDTD PTB room model is provided on the accompanying CD.

In accordance with the findings of the case study, further refinements to the multiplane modelling approach are incorporated into the PTB multiplane model. Specifically, approaches to calibrating reverberation times and achieving an appropriate means of source excitation for arbitrary sound source/receiver locations are discussed in the following.

5.3.3.1 Control of RT_{60} in the Multiplane Model

The investigation of 2D multiplane FDTD modelling presented in the case study considers the evaluation of an approach to absorption coefficient calibration that facilitates the emulation of 3D FDTD temporal energy decay characteristics in resulting RIRs. This approach is shown to be successful in producing multiplane RIR T_{30} values that are consistent with those derived from corresponding 3D FDTD RIRs simulated for the case study acoustic modelling scenario. However, in order to enable this outcome, the calibration of each multiplane absorption coefficient involves the simulation of a cubic 3D FDTD acoustic model (see section 5.2.3.3 for reference). For the purposes of the case study, in which only seven frequency-independent surface material types are

incorporated into the modelled environment, this approach is considered feasible. With regards to the PTB room, which incorporates a much greater number of coefficients in order to represent frequency-dependent surface material absorption properties, such an approach is deemed cumbersome and inefficient. Instead, a more direct and effective means of matching the RT_{60} of multiplane and 3D FDTD RIRs is used in this final study as described here. In order to provide a concise overview of this RT_{60} matching process, it is assumed that the following description refers to the process as applied to each octave band RIR simulated for each RIR case.

Upon simulation of the contributing cross-sectional planar responses in the PTB multiplane model, the RT_{60} of each planar response is calculated. Each contributing response is then manipulated to produce the RT_{60} measured from the corresponding 3D FDTD model RIR through application of the following scaling factor, following [167]:

$$\triangle G = \frac{60}{F_s}\left(\frac{1}{RT_{60,2D}} - \frac{1}{RT_{60,3D}}\right) \tag{5.21}$$

where $RT_{60,2D}$ is the reverberation time calculated for a given contributing planar response and $RT_{60,3D}$ is the reverberation time of the corresponding 3D FDTD RIR. Following [167], the term $\triangle G$ yields the required change in planar response amplitude (dB per sample) that must be applied in order to scale the response over time to attain an amplitude value of -60 dB below the initial direct sound component amplitude at $t = RT_{60,3D}$. Having scaled all contributing planar responses using (5.21), it is then assumed that the combination of these responses will produce a total multiplane RIR that possesses an RT_{60} consistent with that of the corresponding 3D FDTD RIR. However, it is essential to note that the results referring to T_{30} documented in section 5.3.4.1 are realised by applying this procedure twice to each contributing cross-sectional response.

5.3.3.2 Refinement of the Multiplane Sound Source Model

The case study, as documented previously, reveals that the method of projected sound sources applied in the 2D multiplane FDTD model leads to an imbalance of energy contributed by each cross-sectional response under certain conditions. As demonstrated in section 5.2.3.3, a projected sound source that excites a cross-sectional plane in close proximity to the defined receiver location may give rise to an imbalance in the level to which each planar response contributes to the final multiplane RIR. This is shown to be the case when applying the direct sound matching approach (5.17) in such circumstances. Therefore, the use of this planar response matching procedure cannot be considered valid for arbitrary source and receiver combinations.

Direct sound component scaling is shown to provide a solution to this problem facilitating appropriate combination of planar responses such that each contributes a similar amount of energy to the total multiplane RIR (see Figure 5.24 for reference). This process is revisited here in order to establish an approach to projected source excitation that is suitable for any source/receiver arrangement in multiplane acoustic models.

Upon simulation and filtering of the contributing cross-sectional responses, the direct sound components present in each are time-aligned. The cross-sectional plane on which the direct path distance between projected sound source and receiver locations is most consistent with the 3D coordinate positions of source/receiver is then identified. It is assumed that the direct sound component peak pressure amplitude simulated in this plane best matches that which would arise for 3D sound propagation. The direct sound components of the remaining two planes are then isolated and scaled up or down to this peak amplitude. This step ensures that direct sound components captured in cross-sectional planes on which the projected sound source is located at, or in close proximity to, the receiver location are reduced in amplitude. Upon application of this direct sound scaling procedure, the initial peak amplitude of the response from each contributing plane is equal.

The final stage of the refined projected source model involves calibrating the response from each cross-section in terms of energy levels. This is carried out by calculating the total acoustic energy present in one of the contributing responses and then matching the energy of the remaining two planar responses to this value. Hence, summation of the calibrated planar responses ensures that the resulting multiplane RIR possesses equal contributions from each cross-section in terms of direct sound pressure and response energy levels. This process is conducted as described for all octave band multiplane simulations. Resulting octave band RIRs are combined to produce the final simulation output for each source/receiver case by means of the IIR bandpass filter bank operation described in relation to the 3D FDTD model (see section 5.3.2).

This development of sound source implementation in the multiplane model is guaranteed to circumvent the issue of imbalanced contributions each cross-sectional response. Furthermore, the application of this projected sound source model to the PTB room multiplane model RIRs is shown to produce simulation results that are in good agreement with those obtained from the 3D FDTD PTB model in terms of spectral characteristics. These results are presented in section 5.3.4.2.

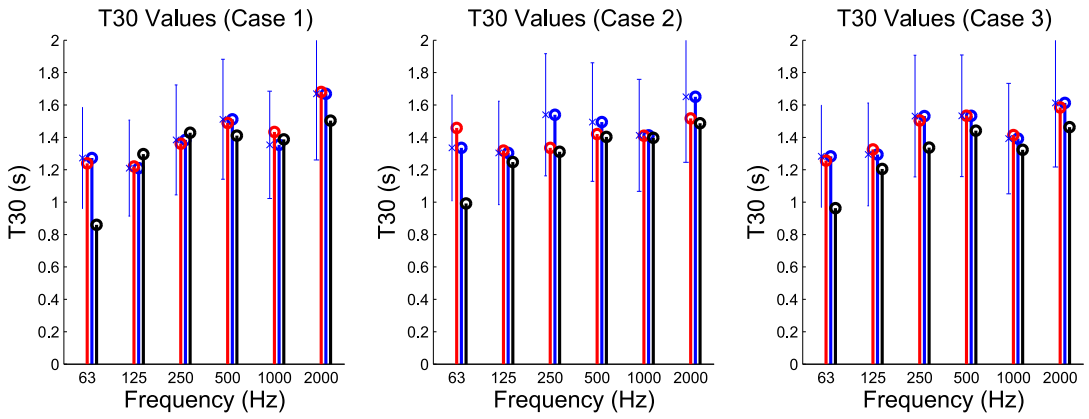

FIGURE 5.36: T_{30} values calculated from 3D FDTD (blue), 2D multiplane FDTD (red) and ODEON (black) RIRs (Cases 1 - 3) for the 63, 125, 250, 500, 1000, 2000 Hz octave bands. The error bars display the range of one JND (24.5%, after [16]) above and below the T_{30} value calculated from the 3D FDTD RIRs in each octave band.

5.3.4 Results and Discussion

The results presented in this section pertain to the objective analysis of RIRs simulated by means of the PTB room acoustic models previously described. Findings are drawn from comparative examination of the geometric, 3D FDTD and multiplane FDTD RIRs synthesised for each source/receiver arrangement defined for this final study. To this end, the means of analysis detailed in the case study is again applied here. As such, both acoustic parameter values and spectral properties obtained from simulated RIRs are considered in order to evaluate the capabilities of 2D multiplane FDTD acoustic modelling. Note that the multiplane and 3D FDTD RIRs are down-sampled to F_s = 44.1 kHz prior to analysis.

5.3.4.1 Acoustic Parameters

Frequency-dependent analysis of the T_{30}, EDT and C_{80} acoustic parameters derived from all RIRs simulated in this study is provided here. Parameter values are calculated for the 63, 125, 250, 500, 1000 and 2000 Hz octave bands using the MATLAB Acoustic Parameter Toolbox [160]. JND ranges for each acoustic parameter are also provided in order to quantify the perceptual agreement of simulated RIRs.

Figure 5.36 presents the T_{30} as simulated by each PTB room acoustic model for RIR Cases 1-3. In the charts displayed the 3D FDTD, 2D multiplane and geometric model T_{30} values are denoted by the blue, red and black bars respectively. A single JND range of 24.5% of the 3D FDTD model T_{30} is provided by the blue bar in each octave band for all RIR cases. As expected, the T_{30} values calculated for all 3D FDTD and 2D

multiplane RIRs are shown to be in consistent agreement due to the RT_{60} calibration process applied to each multiplane octave band RIR (see section 5.3.3.1). This demonstrates that it is possible to simulate predefined frequency-dependent reverberation times in the multiplane model to a high degree of accuracy (i.e. within 1 JND of the prescribed T_{30}). As such, the multiplane model T_{30} can be matched to those calculated through use of alternative acoustic modelling techniques (i.e. efficient ray-tracing) or related mathematical expressions (i.e. the Norris-Eyring equation) for any given sonic environment.

It is also shown in Figure 5.36 that the incorporation of normal-incidence reflection coefficients in the 3D FDTD model has led to a significantly higher level of agreement between the T_{30} values obtained from the FDTD and geometric models than that observed in the case study. This claim is supported by the fact that T_{30} measures obtained from the geometric model RIRs are within 1 JND of those simulated in the 3D FDTD model for all analysis frequency ranges with the exception of the 63 Hz octave band. With reference to the case study findings, it is shown that large differences in RIR energy decay characteristics simulated by geometric and FDTD models can adversely affect the process of spectral analysis. Due to the level of agreement of energy decay times observed here, it is assumed that a suitable basis for subsequent comparative low frequency analysis has been established in this study.

Comparing RIR EDT values calculated for Case 1 RIRs simulated by each acoustic model, as presented in Figure 5.37, it is apparent that the initial energy decay characteristics of multiplane and 3D FDTD RIRs are, on the most part, well matched. While there is a clear difference between EDT values calculated in the 250 Hz octave band, the multiplane and 3D FDTD RIR EDTs are consistent to within 1 JND range of +/- 5% (as denoted by the blue error bar) in the 63, 125, 500 and 2000 Hz octave bands. This, in turn suggests a high level of perceptual similarity of these RIRs in terms of perceived reverberation. In relation to the EDT values calculated for RIR Cases 2 and 3 (also shown in Figure 5.37) it is evident that reasonable agreement exists between multiplane and 3D FDTD models in terms of early energy decay properties over the simulated bandwidth. These findings suggest that the refinements implemented in the multiplane model in terms of projected sound source and RT_{60} calibration procedures have had the effect of increasing the extent to which this acoustic modelling approach is able to represent 3D acoustic soundfields. Moreover, as with the T_{30} values previously detailed, the EDTs derived from the geometric RIRs are indicative of improved agreement between FDTD and geometric model energy decay characteristics. This results from the incorporation of normal-incidence absorption coefficients in the FDTD models.

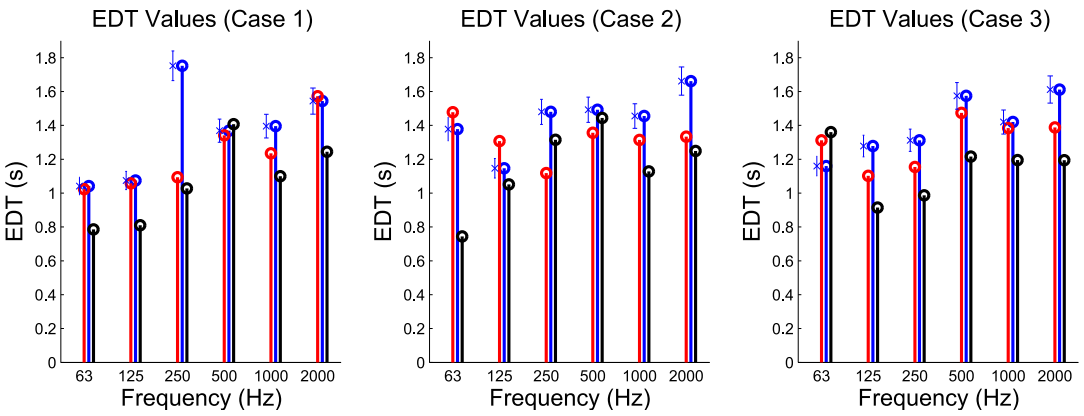

FIGURE 5.37: EDT values calculated from 3D FDTD (blue), 2D multiplane FDTD (red) and ODEON (black) RIRs (Cases 1 - 3) for the 63, 125, 250, 500, 1000, 2000 Hz octave bands. The error bars display the range of one JND (5%, after [17]) above and below the EDT value calculated from the 3D FDTD RIRs in each octave band.

Results pertaining to the Clarity (C_{80}) measures derived from each simulated RIR are displayed in Figure 5.38. From visual inspection of these results, it is readily apparent that multiplane RIR Case 3 matches well with the corresponding 3D FDTD RIR for C_{80} across the simulation bandwidth. This claim is supported by the fact that, with the noted exception of the 125 Hz octave band, the C_{80} measures derived from the multiplane Case 3 RIR are within the JND range of +/-1 dB of the 3D FDTD model results. This high level of agreement for C_{80} is not maintained for the remaining two RIR Cases. For example, the multiplane Case 2 RIR C_{80} values are observed to be consistently outside the JND range centered on the C_{80} measures derived from the 3D FDTD model for all octave bands. Furthermore, referring back to the trends of similarity observed between 3D and multiplane FDTD RIRs in terms of T_{30} and EDT measures over the simulation bandwidth, it is not possible to detect an apparent trend between the measurement of energy decay time and early-to-late energy for this RIR case. By way of contrast, the large difference previously detailed between the EDT derived from multiplane and 3D FDTD RIRs (RIR Case 1, 250 Hz octave band) is also demonstrated by the corresponding C_{80} measure. In this single instance, the longer EDT derived from the 3D FDTD model is shown to result in a C_{80} value that is less than that measured for the multiplane model. Hence, in review of the comparative examination of Clarity index measures presented here (and previously), it is concluded that the matching of early-to-late energy ratios arising in multiplane and 3D FDTD models constitutes a challenge in terms of further refining the multiplane RIR synthesis solution.

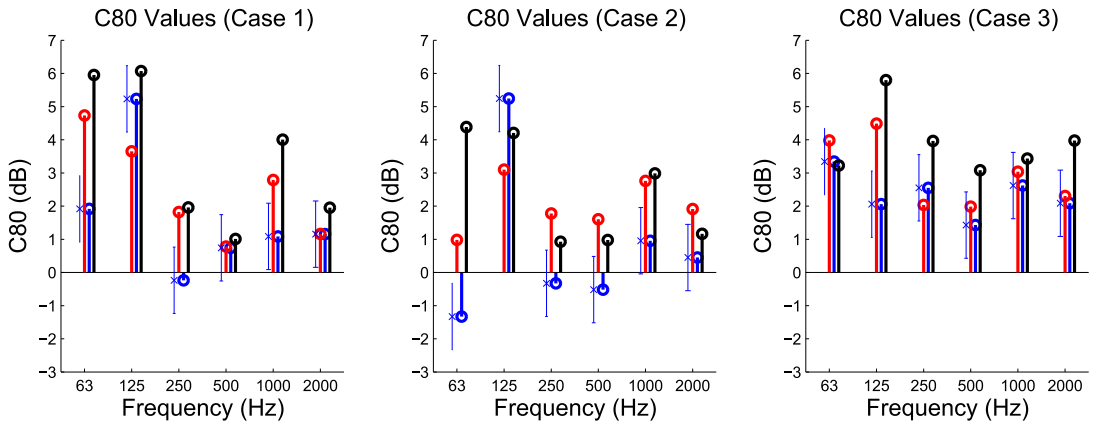

FIGURE 5.38: C_{80} (dB) values calculated from 3D FDTD (blue), 2D multiplane FDTD (red) and ODEON (black) RIRs (Cases 1 - 3) for the 63, 125, 250, 500, 1000, 2000 Hz octave bands. The error bars display the range of one JND (1 dB, after [17]) above and below the C_{80} value calculated from the 3D FDTD RIRs in each octave band.

5.3.4.2 Low Frequency Analysis

The following results pertain to the low frequency spectral analysis of RIRs obtained from CATT-Acoustic, 2D multiplane FDTD and 3D FDTD room acoustic simulations of the PTB room. This analysis follows the same format as that used for the purposes of the pilot and case studies. As such, the agreement between overall spectra magnitude characteristics derived from each simulated RIR is inspected firstly through visual comparison. This examination is conducted for frequencies in the region f_{r1} = [0:150] Hz. This is followed by examination of FRAC (5.11) and MI (5.15) values calculated through comparison of both geometric and multiplane RIRs against 3D FDTD RIRs for f_{r1}, which includes the modal region, and a wider analysis bandwidth f_{r2}. The second frequency range, f_{r2} = [0:1000] Hz, is consistent with that used for the pilot and case study low frequency analysis.

The low frequency spectra of RIRs simulated by the 3D FDTD (blue), 2D multiplane (red) and geometric (black) PTB room acoustic models for RIR Case 1 are presented in Figure 5.39. Referring to the top most diagram, it is observed that the multiplane model is representing low-order modal resonances that are indicated in the 3D FDTD RIR ($f \approx$ [20:40] Hz). However, the first two modal peaks in the multiplane spectrum are not completely coincident with those simulated in the 3D model in terms of frequency value. This is in contrast to the results presented for the more simple pilot and case study modelling in which the agreement between 1st-order mode values simulated by both FDTD models is indicated. Hence, this finding is assumed to result from applying the multiplane model to a more complex target acoustic field. As frequency increases, the agreement between multiplane and 3D RIR spectra becomes increasingly

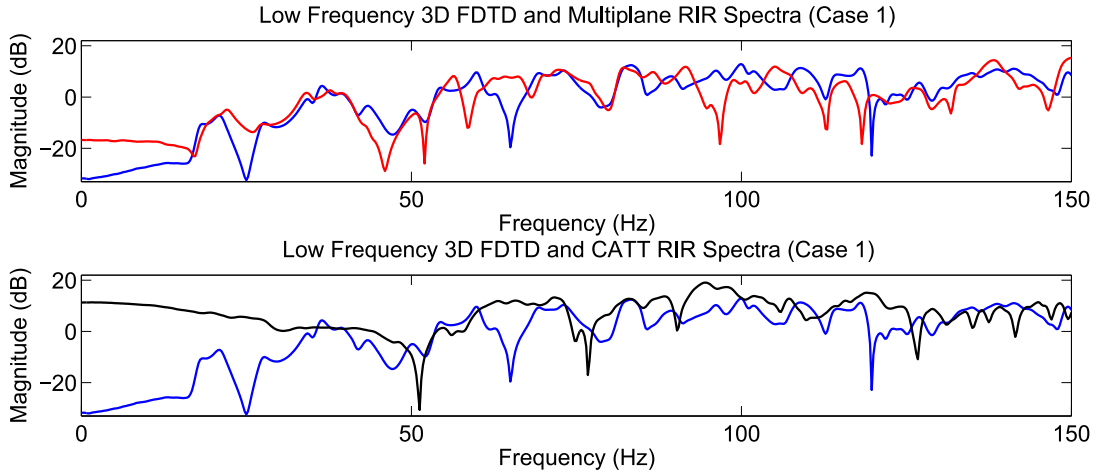

FIGURE 5.39: Low frequency spectra of 3D (blue), 2D multiplane (red) and CATT-Acoustic (black) RIRs simulated for the PTB room (Case 1). The frequency range displayed includes the modal region f_{r1}.

variable. For example, coincident magnitude peaks are indicated in the regions of f = 50 Hz and $f \approx$ [70:80] Hz. Conversely, the agreement of magnitude level trends in 3D and multiplane RIR spectra for $f \approx$ [55:70], [80:120] Hz is relatively low in terms of coincident magnitude peaks and notches. However, towards the upper limit of f_{r1} ($f >$ 120 Hz), the correlation of both spectra is shown to improve in terms of overall magnitude trends.

The level of similarity between geometric and 3D FDTD Case 1 RIR spectra is negligible in the range f = [0:50] Hz, indicating that the geometric model is incapable of correctly simulating low-order modal resonances as expected. Moreover, it may be observed from visual comparison of these two spectra in the range $f >$ 50 Hz, that only a few instances of coincident resonant frequencies exist. The most apparent of these are indicated at frequencies $f \approx$ [55:60] Hz and $f \approx$ 80 Hz. However, reasonable agreement between overall magnitude levels present in geometric and 3D FDTD RIR spectra is shown across the majority of the displayed frequency range ($f >$ 55 Hz). This is representative of the successful implementation of the energy matching procedure, as described in section 5.3.2. As displayed in Figure 5.40 (top diagram), the similarity of 3D and multiplane Case 2 RIR spectra is notably greater than that observed for RIR Case 1 for frequencies up to f = 60 Hz. However, again it is indicated that the multiplane model does not fully represent the low-order modal resonances simulated in the 3D FDTD model. This is particularly apparent for the second and third modal peaks observed in the 3D RIR spectrum. Moreover, variable trends of magnitude level correlation, similar to those previously detailed for RIR Case 1, are also displayed in this case ($f \approx$ [55:120] Hz). Towards the upper frequency limit of f_{r1} (f = [100:130]

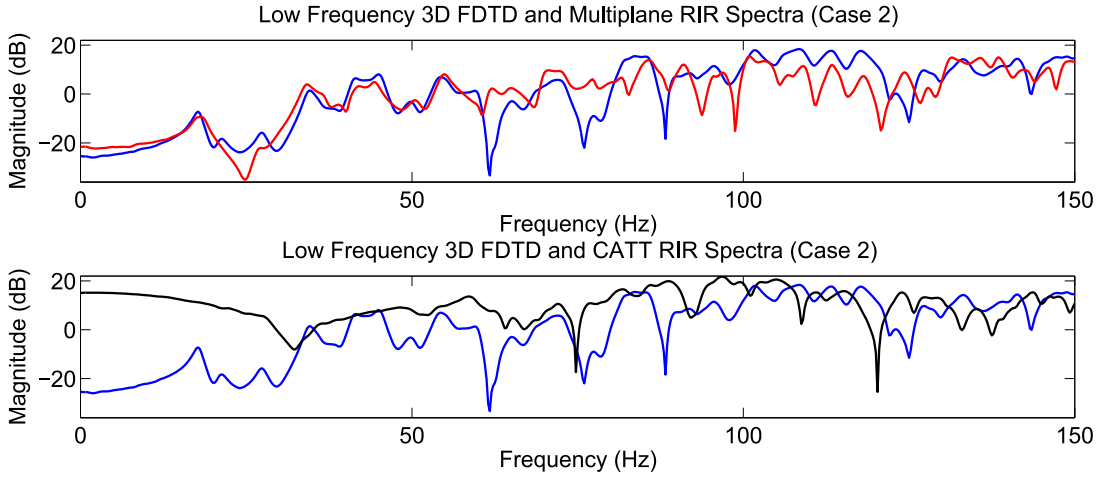

FIGURE 5.40: Low frequency spectra of 3D (blue), 2D multiplane (red) and CATT-Acoustic (black) RIRs simulated for the PTB room (Case 2). The frequency range displayed includes the modal region f_{r1}.

Hz), the coincidence of magnitude peaks in each spectrum improves while, at the same time, the agreement of overall magnitude level trends is low when compared to that shown for $f < 55$ Hz.

In terms of the comparison of geometric and 3D FDTD Case 2 RIR spectra, similar findings to those described for RIR Case 1 may be derived through visual inspection. Firstly, it is again shown that the geometric model does not represent significant low frequency spectral peaks observed in the 3D FDTD RIR spectrum ($f < 60$ Hz). Secondly, only a small number of resonant frequencies present in the 3D FDTD RIR spectrum are simulated by the geometric model for $f > 60$ Hz. The most apparent of these are observed in the regions of $f \approx$ [60:70] Hz and $f = 125$ Hz. Finally, good agreement between overall magnitude levels present in geometric and 3D FDTD RIR spectra over the region f_{r1} is again observed for this RIR case.

Visual examination of the spectra of RIRs simulated for RIR Case 3, as presented in Figure 5.41, yields findings that are, in the most part, consistent with those documented for analysis of RIR Cases 1 and 2. The geometric Case 3 RIR again exhibits a lack of low-order modal resonances for $f < 55$ Hz which are represented in the 3D FDTD model. Additionally, the geometric RIR spectrum magnitude levels are shown to be in agreement with those of the 3D FDTD RIR spectrum in the region of $f =$ [50:110] Hz. Towards the upper limit of f_{r1} this agreement diminishes. In terms of the 2D multiplane Case 3 RIR spectrum, as displayed in the top diagram of Figure 5.41, a strong agreement with the magnitude levels observed in the 3D FDTD RIR spectrum is apparent over the region f_{r1}. As with the findings drawn from analysis of RIR Cases 1 and 2, the multiplane model is demonstrated as capable of representing resonant

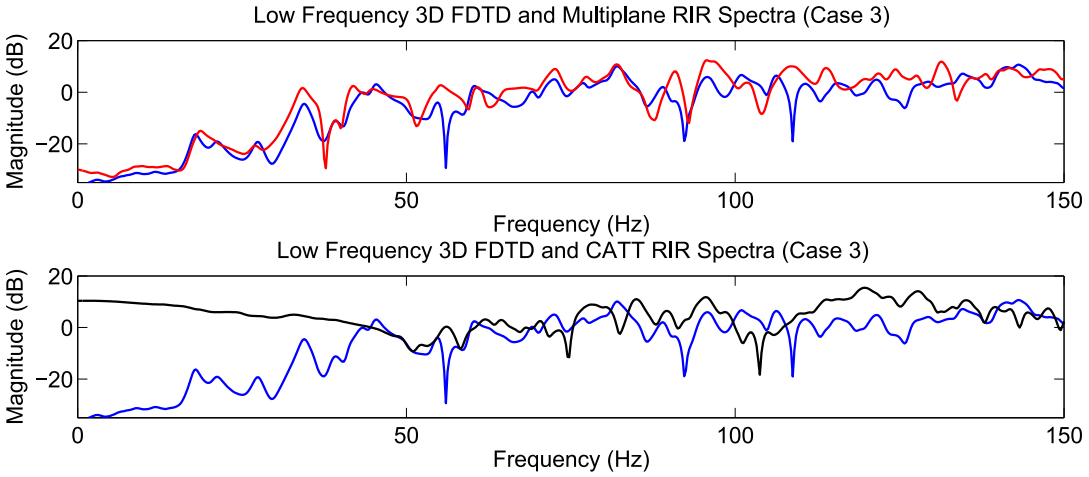

FIGURE 5.41: Low frequency spectra of 3D (blue), 2D multiplane (red) and CATT-Acoustic (black) RIRs simulated for the PTB room (Case 3). The frequency range displayed includes the modal region f_{r1}.

sound behaviour resulting in the simulated 3D soundfield to good extent. However, the omission of notable low-order modes simulated by the 3D FDTD model is again noted from inspection of the 2D multiplane RIR spectrum. Additionally, in review of the visual inspection conducted for all PTB RIR cases, it may be concluded that the 2D multiplane model is emulating the variation in magnitude trends that arise in the 3D FDTD RIR spectra as the source and receiver locations are altered. This result is consistent with the findings of the pilot study as detailed in section 5.1.4.

Model	Case	FRAC	MI
2D Multiplane FDTD	1	0.719	0.857
CATT-Acoustic	1	0.612	0.748
2D Multiplane FDTD	2	0.701	0.859
CATT-Acoustic	2	0.490	0.700
2D Multiplane FDTD	3	0.790	0.897
CATT-Acoustic	3	0.459	0.712

TABLE 5.20: FRAC and MI values calculated to compare the agreement of 2D multiplane and CATT-Acoustic RIRs with the 3D FDTD RIRs (Cases 1 - 3) in the modal region f_{r1}.

The final stage of objective analysis conducted in this study is the examination of FRAC and MI values as calculated through comparison of geometric and multiplane RIRs with 3D FDTD RIRs. These values are displayed in Tables 5.20 and 5.21 for the analysis of frequency ranges f_{r1} and f_{r2} respectively. The results pertaining to

the spectral similarity of 2D multiplane and 3D FDTD RIRs in the frequency region f_{r1} indicate a consistently high correlation between the soundfields simulated in each model. Furthermore, the FRAC and MI values calculated through comparison of geometric and 3D FDTD RIRs are notably lower than those relating to the multiplane model for all RIR cases examined. This finding is particularly apparent for RIR Cases 2 and 3. Hence, in accordance with the findings of the comparative visual examination of RIR spectra presented previously, it has been determined that the multiplane model is capable of providing a better approximation to low frequency soundfield simulation than that achieved by alternative geometric approaches. Moreover, in terms of this final study, this conclusion has been reached through examination of objective results derived from multiplane simulations of a complex 3D room geometry.

Model	Case	FRAC	MI
2D Multiplane FDTD	1	0.626	0.848
CATT-Acoustic	1	0.646	0.836
2D Multiplane FDTD	2	0.624	0.840
CATT-Acoustic	2	0.602	0.819
-			
2D Multiplane FDTD	3	0.623	0.845
CATT-Acoustic	3	0.627	0.823

TABLE 5.21: FRAC and MI values calculated to compare the agreement of 2D multiplane and CATT-Acoustic RIRs with the 3D FDTD RIRs (Cases 1 - 3) in frequency range f_{r2}.

With regards to FRAC and MI values calculated using the larger analysis frequency range f_{r2} = [0:1000] Hz, the results presented in Table 5.21 demonstrate that both geometric and multiplane models are comparable in terms of their representation of the soundfield simulated by 3D FDTD modelling. This interesting finding provides support to the claim that the multiplane model is capable of representing complex 3D soundfields up to the lower-mid frequency range to a level of accuracy that is consistent with industry standard geometric acoustic modelling software.

5.4 Review of Multiplane RIR Synthesis

Throughout the studies presented in this chapter, elements of the low frequency multiplane RIR synthesis strategy are refined to accommodate increasingly complex acoustic modelling scenarios. These refinements also have the effect of improving the accuracy of synthesised RIRs. Of particular note are the energy calibration processes devised to

model impulsive sound source excitations and match the energy levels present in the low frequency multiplane and mid-high frequency geometric model RIRs. Additionally, several approaches to the control of frequency-dependent reverberation characteristics in the multiplane acoustic models have been used. This section provides an overview of the current recommended approaches to energy calibration and RT_{60} control strategies in multiplane RIR synthesis.

Multiplane Sound Source Model and Energy Calibration:

The initial pilot study applies a two-plane multiplane model to simulate the low frequency soundfield in a cuboid. In this particular model, the source and receiver are defined to be co-planar in the x-y and y-z planes sharing a common vector parallel to the y-axis. To render the multiplane RIR, this arrangement allows a straightforward summation of impulse responses recorded on each contributing plane as the impulsive source on each plane is received with the same direct sound component amplitude and time of arrival. However, this approach to combining the contributions from each plane cannot be extended to cases where source and receiver are not coplanar and/or when there are more than two planes used in the model. Furthermore, the energy calibration between multiplane and geometric RIR contributions relies on the derivation of an appropriate energy level for an equivalent 3D FDTD model following [81].

The case study focuses on developing the former issue with the initial sound source implementation through experimentation with projected sound sources for multiplane models consisting of three contributing cross-sectional planes. For such cases, the three cross-sectional planes are defined to have a common point of intersection given by the receiver location. Each plane is then put into motion by defining a projected source point on each with reference to the 3D Cartesian position of the sound source location. In this way, the restrictive source/receiver positioning criterion imposed on the multiplane model in the pilot study is removed and both source and receiver may be positioned at will. The method of combining responses recorded on each plane is based on that used for the pilot study. The amplitudes of the direct sound components of contributing responses are made equal by weighting each response as necessary. The resulting weighted contributions are then summed to derive the overall multiplane RIR.

A particular source/receiver combination used in this study reveals the inability of this source excitation method to correctly combine responses when the sound source is projected on or in the region of the receiver position on one plane. This is shown to create a large imbalance between the contribution of each planar response to the total multiplane RIR. In order to overcome this imbalance, the direct sound component of the planar response for which the projected source is defined on the receiver location is scaled to match the amplitude of the direct sound component recorded on another

plane. This process begins the generalised approach to sound source modelling in the multiplane model used in the final study. As with the pilot study, the method of calibrating multiplane and geometric RIR energy levels requires the calculation of the energy levels recorded in an equivalent 3D FDTD model.

During the final study, a generalised method of sound source modelling and energy calibration is defined. This method provides the facility of arbitrary sound source and receiver placements by applying a standard post-processing procedure to the planar response contributions generated in a three plane multiplane model. The procedure may be defined in stages as follows:

- Having simulated the three planar responses, the direct sound component of each is isolated and scaled (see e.g. 5.2.3.3) to the direct sound amplitude value occurring on the plane whose receiver is positioned at a distance that is most representative of the 3D Cartesian source/receiver distance.
- The planar responses are then time aligned such that the time-of-arrival of the direct sound component is consistent in each.
- Appropriate low-pass, DC blocking and afterglow removal filtering is applied to each planar response.
- The total energy of a filtered planar response is calculated and matched to the low frequency energy present in the calculated geometric RIR (in the low frequency bandwidth defined for the multiplane model).
- Each planar response is calibrated to possess the same low frequency energy level using a matching coefficient derived in the previous stage.
- The planar responses are then summed to produce the total multiplane RIR. The resulting signal must be divided by three in order to compensate for the fact that each separate planar response is calibrated to match the low frequency energy levels of the geometric RIR.

A further point to be made on this sound source model and energy calibration process is that it inherently maintains an equal energy contribution from each planar response. As such, the wave behaviour modelled in each cross-section of the domain is equally represented in the final result. However, it is evident that the energy calibration stage still requires a predetermined reference energy level (in this case calculated from the geometric RIR). Given the successful studies on 3D FDTD / Geometric RIR calibration presented in [81], it is reasonable to assume that a generic calculation for the sound

source energy level in each cross-sectional plane (2D FDTD scheme) may also be defined. In doing so, the multiplane model would then comprise a self sufficient approach to low frequency RIR synthesis.

Control of RT_{60} in Multiplane Models:

As noted throughout this chapter, the nature of energy decay characteristics in 2D domains are incompatible with those of 3D environments in absence of appropriate calibration. The multiplane model implementations devised for the pilot and case studies take into account the energy decay times of 2D FDTD wave equation schemes which are, generally, longer than those of equivalent 3D schemes. In particular, the case study multiplane RIR synthesis strategy includes a means of equating absorption coefficients applied in both 3D and 2D domains by means of the Norris-Eyring equation for RT_{60}. Each absorption coefficient is applied consistently to the surfaces of a small cuboid 3D FDTD simulation. The reverberation times are then derived from the simulated RIRs for each case of absorption coefficient. These RT_{60} values are then used to calculate an equivalent absorption coefficient using the Norris-Eyring equation for energy decay in 2D domains. These calibrated coefficients are then applied as appropriate to the boundaries of the mutliplane model. Results returned from this approximate approach have been demonstrated as varying in accuracy and, moreover, only frequency-independent absorption values are investigated in these studies.

The final study improves upon the accuracy of the multiplane FDTD model RT_{60} values, with respect to those produced in the equivalent 3D FDTD models, using the frequency-dependent constant F_s absorbing boundary implementation in conjunction with a filtering process documented in [167] (see section 5.3.3.1). This procedure ensures that the decay characteristics of the 2D cross-sectional responses match that of the 3D FDTD model through direct use of the RT_{60} derived from simulated 3D FDTD RIRs. While this approach requires knowledge of the 3D model energy decay characteristics in its application in this study, the procedure may equally be applied to control the RT_{60} of the multiplane RIRs in absence of a 3D FDTD model. From the results documented in section 5.3.4.1, it is evident that the post-filtering approach facilitates accurate representation of RT_{60} regardless of how the prescribed RT_{60} value is derived. For instance, a value for reverberation time may be derived from theory, practical RIR measurement or a geometric approach thus removing the reliance of this element of multiplane modelling on computationally expensive 3D FDTD modelling. However, it is noted that such an approach will only be capable of providing approximate results when applying the multiplane RIR sythesis strategy without reference to otherwise measured or simulated RIRs (i.e. relies on theoretical calculation of energy decay times).

In summary, this section provides clarification of critical developments that the multiplane model has undergone during the course of the three studies presented in this chapter. Furthermore, insight on how the multiplane model may be utilised in absence of data derived from 3D FDTD acoustic modelling is given. The refined procedures noted above, which enable the multiplane modelling system to better simulate complex room topologies, form the basis of the current implementation of the novel RIR synthesis strategy. The final conception of this modelling approach remains to be evaluated in further acoustic modelling scenarios.

5.5 Summary

This chapter documents the processes followed in order to develop and evaluate the 2D multiplane FDTD RIR synthesis solution as a means of representing 3D sound wave propagation in enclosed acoustic environments. Findings pertaining to the accuracy and efficiency of this novel approach to virtual acoustic simulations are presented in relation to three room acoustic modelling scenarios of increasing complexity. These studies have led to the definition of an efficient means of simulating low frequency RIRs which may be utilised as part of a hybrid acoustic modelling strategy. On review of the multiplane FDTD acoustic model, it is apparent that this novel RIR synthesis strategy may be implemented as a standalone system in future implementations. An overview of the conclusions drawn from these studies is provided in the next chapter in which the support of the research hypothesis investigated in this work is discussed.

Chapter 6

Conclusion

This thesis begins with presentation and discussion of aspects of acoustic theory in order to provide a basis of understanding and evaluating the virtual acoustic models reviewed and implemented in this work. The mathematical principles underlying sound wave motion have been presented and described in the context of room acoustics. Particular attention has been given to the characteristics of enclosed sound propagation which include sound reflection, absorption, scattering, diffraction, interference and resonance. It is demonstrated that all of these characteristics are represented by a measured or simulated RIR. In turn, a primary means of examining real or modelled room acoustic properties results from the examination of acoustic parameters derived from RIRs. Examples of such parameters, including reverberation time, early decay time and early-to-late energy index have been presented and discussed.

The properties of human hearing have been detailed giving insight on perceptual auditory cues and providing context for the discussion of auralisation and approaches to spatial sound reproduction. The auralisation process has been introduced and detailed in relation to both static and dynamic (walkthrough) soundfield reproduction methods. Additionally, the concept of just noticeable differences have been introduced and discussed in relation to acoustic parameters to provide a means of objectively quantifying and evaluating the perceptual similarity of RIRs synthesised by different virtual acoustic modelling methods.

A range of virtual acoustic modelling approaches have been reviewed and evaluated in terms of the strengths and weaknesses inherent to each. Particular attention has been given to the comparative efficiency and accuracy of geometric and numerical acoustic simulation methods. It is demonstrated that while FDTD wave-based acoustic modelling is well suited to the problem of low frequency RIR synthesis, prohibitive

computational costs are incurred for 3D modelling scenarios in which the target acoustic field is large or discretised with high resolution. Additionally, ray-based geometric algorithms have been shown to a provide more efficient solution to acoustic modelling in mid-to-high frequency ranges compared to alternative numerical techniques. However, such geometric models are inaccurate when seeking to represent low frequency sound attributes. These findings serve to highlight the advantages of using numerical/geometric hybrid acoustic modelling strategies in order to achieve efficient, yet accurate RIR simulation. In such hybrid solutions, numerical and geometric approaches are applied to low and high frequency RIR simulation respectively.

In light of the evaluation of numerical acoustic modelling methods, the FDTD approach has been selected for use in this work. As such, thorough investigation of this modelling approach, as applied to 2D and 3D sound simulation in discrete rectilinear spatial domains, has been conducted. Aspects including sound source excitation strategies, RIR capture and the implementation of absorbing boundary conditions as applied in FDTD acoustic models have been examined and evaluated. Investigation of velocity-centered and pressure-centered implementations of locally reacting surface boundary conditions has confirmed that the pressure-centered formulation is prone to numerical instability issues when applied to complex model geometries. Further examination of the pressure- and velocity-centered boundaries has been documented comparing the absorption characteristics of each implementation in 1D FDTD discrete wave equation simulations. The findings of this examination demonstrate that the pressure-centered formulation provides closer agreement with theoretical results in terms of simulated absorption rates. Frequency-dependent boundary conditions have been shown to better represent the absorption characteristics of surface materials encountered in practice. A case study has been presented which examines three different approaches to implementing such boundaries in 3D FDTD acoustic models in terms of resulting RIR accuracy and computational efficiency. Moreover, a reformulated velocity-centered implementation of frequency-dependent LRS boundary conditions has been developed to provide a means of realistic surface absorption characteristics within complex 3D SRL FDTD models that is robust against numerical instability issues.

This study culminates in the development and evaluation of a computationally efficient 2D multiplane FDTD hybrid RIR synthesis solution. The motivation for this research stems from both the computational cost associated with 3D numerical acoustic modelling strategies and the inherent inaccuracy of geometric algorithms when applied to low frequency sound simulation. Prior examination of aspects of FDTD SRL acoustic modelling influences the design and implementation of the multiplane models and the 3D FDTD models created to provide a basis of comparative analysis. The capability of the multiplane model to emulate 3D low frequency soundfields is assessed through

the application of the model to three acoustic modelling scenarios of increasing levels of complexity. Throughout this work, refinements have been made to the multiplane acoustic modelling solution, each of which are demonstrated as successful in enhancing the resulting accuracy of low frequency RIRs. Such refinements include: the extension of the multiplane model to three cross-sectional planes; the definition of a projected sound source model that facilitates arbitrary source and receiver positioning within the model; the implementation of methods to control sound energy absorption rates. Major findings drawn from comparative analysis of objective simulation results obtained from 2D multiplane FDTD, 3D FDTD and geometric acoustic models created for each modelling scenario investigated are summarised below:

- The multiplane model is capable of representing resonant low frequency wave behaviour to a level of accuracy that is unachievable through use of industry standard geometric modelling methods.
- Multiplane FDTD RIR synthesis provides a solution for analysis of low frequency enclosed sound propagation that maintains a comparable level of agreement with 3D FDTD RIR synthesis in terms of spectral response characteristics.
- Significant reductions in computational cost are achieved by the multiplane model when compared to alternative 3D FDTD simulations. Through comparison of CPU based implementations of both 3D and multiplane FDTD models, this reduction is in the region of 99% in terms of both memory and run-time requirements.
- Multiplane models are demonstrated to maintain a higher level of agreement with 3D FDTD models than geometric simulations in terms of simulated RIR spectral attributes. This has been confirmed using two industry standard geometric acoustic modelling programs.
- Analysis of RIR acoustic parameters derived from simulated multiplane RIRs demonstrate that the multiplane approach is capable of modelling reverberation times which are consistent with those simulated using 3D FDTD models.

6.1 Concluding Remarks on Multiplane FDTD RIR Synthesis

The main objective of this study is to define, evaluate and develop a means of virtual acoustic modelling that overcomes the known shortcomings of geometric algorithms at

low frequencies, while reducing required computational effort compared to 3D numerical RIR simulation. Multiplane FDTD RIR synthesis provides an alternative strategy to virtual acoustic modelling that represents low frequency sound propagation to a level of accuracy that surpasses geometric modelling methods such as ray-tracing and image source method. Moreover, a high level of agreement between the results calculated through use of 3D FDTD and multiplane modelling approaches is possible. This demonstrates that multiplane acoustic models are capable of approaching the level of accuracy commonly associated with 3D numerical acoustic models.

However, from the studies and discussion thereof presented in the previous chapter, the following drawbacks of the multiplane acoustic modelling strategy are apparent:

- The use of cross-sectional planes allows for the omission of internal structures/geometric features of a given enclosure topology. This would occur if such structures do not coincide with the cross-sectional planes which are defined to possess a common point of intersection at the receiver location. As such, consideration must be given to the modelling scenarios to which the multiplane model is applied. For instance, the multiplane model can be considered appropriate for the acoustics modelling of large, mostly concave enclosures and less appropriate for RIR synthesis for spatial domains that feature significant re-entrant surfaces.
- The time-of-arrival of prominent early reflections synthesised using the multiplane model do not match those that would occur in reality for arbitrary source/receiver placement. This is due to the mapping of reflection paths to 2D cross-sections rather than a 3D volume. This issue is somewhat somewhat compensated by the fact that early reflection paths lengths are correct in mid-high frequency ISM with which the multiplane model may be combined to produce a hybrid acoustic model.
- The agreement of acoustic parameter values (e.g. EDT and C_{80}) derived from the 2D multiplane and 3D FDTD RIRs is demonstrated as variable. From the analysis of results, this has been shown to be partly due to the difference in the nature of the early decay characteristics simulated in each numerical modelling approach. However, accurate control of frequency-dependent RT_{60} in the multiplane model is shown to be possible given a prescribed expected value for this parameter.

The above drawbacks must be considered prior to applying the multiplane modelling strategy to the simulation of RIRs and auralisations. This consideration should be weighed against the considerable, proven improvement of the accuracy of low frequency sound characteristics compared to that provided by purely geometric acoustic models.

Furthermore, the significant computational savings in terms of both memory requirements and run-time achieved by the multiplane model compared to 3D FDTD modelling make the multiplane model an attractive means of emulating sound propagation as part of a hybrid RIR synthesis solution.

In light of the drawbacks and notable improvements on alternative virtual acoustic modelling approaches highlighted above, the application of the devised hybrid multiplane RIR synthesis strategy is intended to find application in industry and creative technologies. The industry applications include rapid prototyping and rendering of auralisations for the purposes of acoustic consultancy. Auralisations rendered using this approach will benefit from increased accuracy in the low frequency portions of resulting reproduced sound (against geometric approaches). Furthermore, given the efficiency of this strategy documented throughout the previous chapter, it is assumed that such a strategy would achieve real-time simulation when utilising GPU architectures for simulation. In creative technologies, the acoustic modelling approach may be applied to efficiently render virtual soundscapes for use in virtual/augmented reality systems providing a realistic and accurate representation of low frequency soundfields for immersive sound reproduction. Ultimately, the use of the multiplane FDTD model is well suited to the definition of a quick, computationally inexpensive augmentation of simulated and reproduced soundfields as part of a hybrid modelling/auralisation solution.

6.2 Restatement of Hypothesis

The primary concern of this thesis has been to investigate the following research hypothesis:

2D multiplane FDTD acoustic modelling facilitates efficient simulation of low frequency soundfields to a level of accuracy higher than geometric models and comparable to 3D FDTD simulations as part of a hybrid modelling solution.

The following research has been conducted in order to prove this hypothesis:

1) A range of virtual room acoustic modelling techniques have been reviewed and assessed in terms of advantages and disadvantages with emphasis placed on the efficiency and accuracy of RIR simulation inherent to each method.

2) Analysis of SRL FDTD acoustic modelling approaches has been conducted in order to inform the implementation of 2D multiplane and 3D FDTD acoustic models.

This included consideration of numerical stability issues and realistic boundary surface modelling methods.

3) A 2D multiplane FDTD hybrid acoustic modelling approach, which applies 2D cross-sectional FDTD discrete wave equation simulations to the modelling of low frequency enclosed sound propagation in combination with mid-to-high frequency geometric acoustic simulations, has been developed and refined.

4) The 2D multiplane model has been used to simulate a range of acoustic environments which possess increasing levels of complexity in terms of model geometry and boundary surface absorption characteristics.

5) Comparative analysis of objective RIR measures simulated by 2D multiplane FDTD with those simulated by 3D FDTD and geometric acoustic models was conducted in order to examine the agreement between simulation results. Particular attention was given to observed characteristics of resulting low frequency RIR spectra.

6) For each model/case study, a broad assessment of the computational resources (in terms of computation time and memory use) required to simulate 3D and 2D multiplane FDTD acoustic models has been provided.

6.3 Contributions

Through the completion of the research outlined above that has confirmed the stated hypothesis, the following novel contributions to the field have been identified:

1) Comparative investigation of three frequency-dependent absorbing boundary condition implementations as applied in 3D SRL FDTD acoustic simulations providing results pertaining to the performance of each in terms of accuracy and computational efficiency.

2) An investigation into the numerical stability of frequency-independent and frequency-dependent absorbing LRS boundary modelling techniques for 3D SRL FDTD acoustic modelling simulations concludes with a reformulation of frequency-dependent boundary conditions that is demonstrated as robust against numerical instability issues which arise for complex acoustic model geometries.

3) The introduction, development and evaluation of a novel 2D multiplane FDTD hybrid acoustic modelling solution which is demonstrated as capable of simulating RIRs with increased accuracy at low frequencies compared to that achieved using geometric

methods only, while attaining significant reductions in computational cost compared to alternative 3D numerical acoustic modelling.

6.4 Future Research

Reflection on the novel research presented in this thesis provides a basis for informing future work towards the development of efficient virtual acoustic modelling and auralisation procedures. The following considerations for further research are based on elements of the work detailed and discussed throughout this thesis.

6.4.1 Refining Variable F_s Octave Band LRS Boundary Modelling

Investigation of frequency-dependent LRS boundary implementations demonstrated that the variable F_s octave band modelling approach is computationally less expensive than alternative constant F_s and DIF-based approaches both in terms of simulation run-times and memory requirements. At the same time, the more efficient variable F_s octave band model has been shown as capable of simulating frequency-dependent acoustic parameters to a level of accuracy that is comparable to that produced using the alternative boundary models. However, the energy calibration method applied in this work to combine contributing octave band RIRs simulated using the variable F_s approach is shown to produce frequency-dependent RIR energy levels that are inconsistent with those simulated by the alternative modelling methods. As such, it is proposed that this aspect of the variable F_s octave band modelling solution be revisited and improved. In doing so, it will then be possible to capitalise on the low computational cost inherent to this method while achieving accurate RIR simulation.

6.4.2 FDTD Boundary Modelling using Measured Surface Phase Response Data

An area of numerical room acoustic modelling that has yet to be examined in detail is the evaluation of surface boundary modelling which incorporates material phase response characteristics. While processes that facilitate this outcome have been developed, there is little evidence to support or refute the claim that the incorporation of surface phase properties in room acoustic models leads to a higher level of objective accuracy in simulated RIRs. This potential route of research would involve comparison of measured RIRs with those simulated using numerical acoustic models with and without the application material phase response properties to boundary surfaces. Objective

comparative analysis could then be conducted on collected RIRs to ascertain whether it is appropriate to include such properties in acoustic models in order to obtain an accurate representation of a target acoustic environment.

6.4.3 Comparison of Octave Band 2D Multiplane and Measured RIRs

This study has focused primarily on the comparison of objective results pertaining to 2D multiplane and 3D FDTD RIR synthesis. Further investigation of the multiplane acoustic model could seek to extend the evaluation of this approach through comparison of simulated and measured RIRs collected for a range of real soundfields. In turn, the applicability of the multiplane approach as an acoustic prediction and analysis procedure would be further assessed.

6.4.4 Acceleration of 2D Multiplane FDTD Hybrid Acoustic Modelling Simulations

It has been demonstrated in this work and related literature that parallelised GPU implementations of 3D FDTD acoustic models provides a means of significantly reducing simulation run-times. As such, it is assumed that the 2D multiplane FDTD model developed in this work will also benefit from the acceleration in processing speeds gained through GPU use. It is expected that low frequency soundfield simulation will be possible for large scale acoustic environments at interactive rates by GPU based implementations of multiplane acoustic models. Furthermore, parallel computation procedures and refined variable F_s octave band boundary condition modelling may be combined and applied to the multiplane model in order to efficiently synthesise realistic RIRs.

Abbreviations

AR	**A**coustic **R**adiosity
ARD	**A**daptive **R**ectangular **D**ecomposition
BEM	**B**oundary **E**lement **M**ethod
DCT	**D**iscrete **C**osine **T**ransform
DIF	**D**igital **I**mpedance **F**ilter
DWG	**D**igital **W**ave**g**uide
DWM	**D**igital **W**aveguide **M**esh
FDTD	**F**inite **D**ifference **T**ime **D**omain
FEM	**F**inite **E**lement **M**ethod
FTM	**F**inite **T**ransform **M**ethod
HAM	**H**ybrid **A**coustic **M**odel
GPGPU	**G**eneral **P**urpose **G**raphics **P**rocessing **U**nit
ILD	**I**nteraural **L**evel **D**ifference
ITD	**I**nteraural **T**ime **D**ifference
ISM	**I**mage **S**ource **M**ethod
JND	**J**ust **N**oticeable **D**ifference
LRS	**L**ocally **R**eacting **S**urface
RIR	**R**oom **I**mpulse **R**esponse

Symbols

c	wave speed	ms^{-1}
d	distance	m
EDC	energy decay curve	dB
F	Force	N
F_c	Schroeder frequency	Hz
f_l	lower cut-on frequency	Hz
f_{r1}	modal frequency region	Hz
f_{r2}	low-mid simulation frequency region	Hz
F_s	temporal sampling frequency	Hz
f_u	upper cut-off frequency	Hz
$FRAC$	Frequency Response Assurance Criterion	
h	discrete inter-nodal distance	m
H	Heaviside unit step function	
I	sound intensity	$\frac{\mathrm{W}}{\mathrm{m}^2}$
K	spring constant	$\frac{\mathrm{kg}}{\mathrm{s}^2}$
$\mathbf{k}$	wavenumber	$\frac{\mathrm{rad}}{\mathrm{m}}$
l	discrete spatial index (x-axis)	
m	discrete spatial index (y-axis)	
MI	Magnitude Similarity Index	
MS	Magnitude Similarity	
n	discrete time index	
p	pressure	Pa
P	power	W (Js^{-1})
q	discrete spatial index (z-axis)	
R	Reflection Coefficient	

SIL	sound intensity level	dB
SPL	sound pressure level	dB
t	time	s
T	discrete time step	s
u	velocity	ms^{-1}
Z_{air}	characteristic impedance of air	$\frac{\mathrm{kg}}{\mathrm{m}^2\mathrm{s}}$
α	absorption coefficient	
$\delta(t)$	Dirac delta function	
ζ_ω	surface impedance	
λ	Courant number	
ω	angular frequency	rads^{-1}
ρ	density	$\frac{\mathrm{kg}}{\mathrm{m}^2}$

References

[1] David Howard and Jamie Angus. *Acoustics and Psychoacoustics*. Wiley & Sons, New York, NY, 3^{rd} edition, 2006.

[2] Sorrel Harriet. *Application of Auralisation and Soundscape Methodologies to Environmental Noise*. PhD thesis, University of York, 2013.

[3] Heinrich Kuttruff. *Room Acoustics*. Spon Press, Oxon, UK, 5^{th} edition, 2009.

[4] ODEON. Oblique Lambert, *Accessed [online]*, March 2016. URL `http://www.odeon.dk/pdf/Oblique%20Lambert.pdf`.

[5] Andrew Chadwick. Towards naturalness in sound synthesis. In *PhD Academic Research Poster, University of York*, June 2012.

[6] Angelo Farina. RAMSETE - a new pyramid tracer for medium and large scale acoustic problems. In *Proceedings of EURO-NOISE*, Lyon, France, 1995.

[7] Alex Southern. *The Synthesis and Auralisation of Physically Modelled Soundfields*. PhD thesis, University of York, 2011.

[8] P. S. Heckbert and P. Hanrahan. Beam Tracing Polygonal Objects. In *11^{th} annual conference on computer graphics and interactive techniques, SIGGRAPH '84*, volume 18, pages 119–127, 1984.

[9] 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.

[10] Davide Rocchesso and Julius O. Smith III. Generalized digital waveguide networks. *IEEE Transactions on Audio, Speech and Language Processing*, 11(3): 1063–6676, 2003.

[11] Damian Murphy. *Digital Waveguide Mesh Topologies in Room Acoustics Modelling*. PhD thesis, University of York, 2001.

[12] Julian R. Wright. An Exact Model of Acoustic Radiation in Enclosed Spaces. *Journal of the Audio Engineering Society (JAES)*, 43(10):813–820, 1995.

[13] Stefan Petrausch and Rudolf Rabenstein. Wave field simulation with the functional transformation method. In *Int. Conf. on Acoustics, Speech and Signal Processing*, pages 317–320, May 2006.

[14] Damian Murphy, Mark Beeson, Simon Shelley, Alex Southern, and Alastair Moore. *RenderAIR* - room acoustics simulation using a hybrid digital waveguide mesh approach. In *Audio Engineering Society Convention 124*, May 2008. URL `http://www.aes.org/e-lib/browse.cfm?elib=14559`.

[15] Sound on Sound. SOS Forum Post, *Accessed [online]*, March 2016. URL `http://www.soundonsound.com/sos/dec05/articles/qa1205_4.htm`.

[16] 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 *International Symposium on Room Acoustics*, June 2013.

[17] International Standards Office 3382-1:2009. Acoustics – Measurement of room parameters – Part 1: Performance spaces. *ISO*, 3382(1), 2009.

[18] SketchUp. Sketchup make, *Accessed [online]*, March 2016. URL `http://www.sketchup.com/download?sketchup=make`.

[19] CATT-Acoustic. Software for room acoustic consulting and audio virtual reality, *Accessed [online]*, March 2016. URL `www.catt.se/`.

[20] Cyril Harris. Absorption of sound in air versus humidity and temperature. *J. Acoust. Soc. Am.*, 40(1):148–159, 1966.

[21] ODEON. ODEON Room Acoustics Software website, *Accessed [online]*, March 2016. URL `http://www.odeon.dk/`.

[22] Samuel Siltanen. random2normalincidence.m - MATLAB Function File. Personal Communication, May 2015.

[23] Philip Morse. *Vibration and Sound*. McGraw-Hill Book Company Inc., York, PA, 1948.

[24] David T. Blackstock. *Fundamentals of physical acoustics*. Wiley & Sons, New York, NY, 2000.

[25] Thomas Rossing, Richard Moore, and Paul Wheeler. *The Science of Sound*. Addison-Wesley, Boston, MA, 3^{rd} edition, 2001.

[26] Lawrence Kinsler, Austin Frey, Alan Coppens, and James Sanders. *Fundamentals of Acoustics*. 1999.

[27] F. Alton Everest and Ken Pohlmann. *Master Handbook of Acoustics*. McGraw-Hill, 5^{th} edition, 2009.

[28] Kazumi Watanabe. *The Science of Sound*. Springer International, Lecture Notes in Applied and Computational Mechanics edition, 2014.

[29] Arthur Erdelyi. *Tables of Integrals*. McGraw-Hill, New York, NY, 1954.

[30] Gabriel Barton. *Elements of Green's Functions and Propagation: Potentials, Diffusion and Waves*. Claredon Press, Oxford, UK, 1989.

[31] Jose Escolano, Carlos Spa, Adan Garriga, and Toni Mateos. Removal of afterglow effects in 2-D discrete-time acoustics simulations. *Applied Acoustics*, 74(6):818–822, 2013.

[32] Carlos Spa, Jose Escolano, Adan Garriga, and Toni Mateos. An overview of the afterglow phenomena compensation. *Journal of the Acoustical Society of America*, 130(4):2364–2364, 2011.

[33] Jose Escolano, Adan Garriga, and Toni Mateos. Compensation of the Afterglow Phenomenon in 2-D Discrete-Time Simulations. *Signal Processing Letters, IEEE*, 17(8):758–761, 2010.

[34] Stefan Bilbao. *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics*. Wiley Publishing, 2009.

[35] International Standards Office 9613-1:1993. Acoustics – Attenuation of sound during propagation outdoors – Part 1: Calculation of the absorption of sound by the atmosphere. *ISO*, 9613(1), 1993.

[36] International Standards Office 9613-2:1996. Acoustics – Attenuation of sound during propagation outdoors – Part 2: Calculation of the absorption of sound by the atmosphere. *ISO*, 9613(2), 1996.

[37] Johann Lambert. *Lamberts Photometrie: Photometria, sive De mensura et gradibus luminus, colorum et umbrae*. Leipzig: W. Engelmann, Germany, 1760.

[38] Jens Holger Rindel. Modelling the directional characteristics of sound reflections. In *Joint Baltic-Nordic Acoustics Meeting, BNAM*, June 2004.

[39] ODEON. ODEON Room Acoustics Software, Version 12, User Manual, *Accessed [online]*, 2016. URL `http://www.odeon.dk/pdf/ODEONManual12.pdf`.

[40] M. R. Schroeder and K. H. 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*, 34:76–80, 1962.

[41] Ole Bjor. Maximum Length Sequence, *Accessed [online]*, March 2016. URL `http://www.gracey.co.uk/downloads/nor_mls_theory.pdf`.

[42] Angelo Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *Audio Engineering Society Convention 108*, Feb 2000. URL `http://www.aes.org/e-lib/browse.cfm?elib=10211`.

[43] Angelo Farina. Advancements in impulse response measurements by sine sweeps. In *Audio Engineering Society Convention 122*, May 2007. URL `http://www.aes.org/e-lib/browse.cfm?elib=14106`.

[44] Wallace Sabine. *Collected papers on acoustics*. Harvard University Press, Cambridge, MA, 1992.

[45] M. R. Schroeder. New Method of Measuring Reverberation Time. *Journal of the Acoustical Society of America*, 37:409–412, 1965.

[46] James Cooley, Peter Lewis, and Peter Welch. The Fast Fourier Transform and its Applications. *IEEE Transactions on Education*, 12(1):27–34, 1969.

[47] Lord Rayleigh. On our perception of sound direction. *Phlisophical Magazine*, 13, 19.

[48] Aglaia Foteinou. *Perception of Objective Parameter Variations in Virtual Acoustic Spaces*. PhD thesis, University of York, 2013.

[49] Micheal Vorländer. International round robin on room acoustical computer simulations. In *15th International Congress on Acoustics*, June 1995.

[50] H. P. Seraphim. Untersuchugen u ber die unterschiedsschwelle exponentiellen abklingens von rauschbandimpulsen. *Acta Acustica united with Acustica*, 8:280–284, 1958.

[51] Zihou Meng, Fengjie Zhao, and Mu He. The Just Noticeable Difference of Noise Length and Reverberation Perception. In *International Conference on Communications and Information Technologies (ISCIT)*, April 2007.

[52] M. Kleiner, B. Dalenbäck, and P. Svensson. Auralization - An Overview. *Journal of the Audio Engineering Society (JAES)*, 41(11):861–875, 1993.

[53] M. Noisternig, B. Katz, S. Siltanen, and L. Savioja. Framework for Real-Time Auralization in Architectural Acoustics. *Acta Acustica united with Acustica*, 94 (16):1000–1015, 2008.

[54] R. Kajastila, S. Siltanen, P. Lundén, T. Lokki, and L. Savioja. A distributed real-time virtual acoustic rendering system for dynamic geometries. In *Audio Engineering Society Convention 122*, May 2007. URL `http://www.aes.org/e-lib/browse.cfm?elib=14145`.

[55] A. Southern, J. Wells, and D. Murphy. Rendering walk-through auralizations using wave-based acoustic models. In *17th European Signal Processing Conference (EUSIPCO 2009)*, Aug. 2009. URL `http://www.eurasip.org/Proceedings/Eusipco/Eusipco2009/contents/papers/1569192174.pdf`.

[56] K. Bouatouch, O. Dielle, J. Maillard, J. Martin, and N. Noé. Real time acoustic rendering of complex environments including diffraction and curved surfaces. In *Audio Engineering Society Convention 120*, May 2006. URL `http://www.aes.org/e-lib/browse.cfm?elib=13547`.

[57] J. Houpanienmi, L. Savioja, and T. Takala. DIVA Virtual Audio Reality System. In *Int. Conf. on Auditory Display (ICAD‘ 96)*, pages 111–116, 1996.

[58] L. Savioja, J. Houpanienmi, T. Lokki, and R. Väänänen. Virtual environment simulation - advances in the DIVA project. In *Int. Conf. on Auditory Display (ICAD‘ 97)*, pages 43–46, 1997.

[59] B. N. J. Postma, A. Tallon, and B. F. G. Katz. Calibrated Auralization Simulation of the Abbey of Saint-Germain-des-Prés for Historical Study. In *Intl. Conf. on Auditorium Acoustics, Institute of Acoustics, IoA*, Oct. 2015.

[60] S. Pelzer, L. Aspöck, D. Schröder, and M. Vorländer. Integrating Real-Time Room Acoustics Simulation into a CAD Modeling Software to Enhance the Architectural Design Process. *Buildings*, 2:113–138, 2014.

[61] L. Aspöck, S. Pelzer, F. Wefers, and M. Vorländer. A real-time auralization plugin for architectural design and education. In *EAA Joint Symposium on Auralization and Ambisonics*, April 2014.

[62] A. Krokstad, S. Strom, and S. Sorsdal. Calculating the Acoustical Room Response by use of a Ray Tracing Technique. *Journal of Sound and Vibration*, 8 (1):118–125, 1968.

[63] Lauri Savioja and U. Peter Svensson. Overview of geometrical room acoustics modeling techniques. *Journal of the Acoustical Society of America*, 138(2):708–730, 2015.

[64] EASE Acoustic Simulation. Acoustic Simulation Software for Integrators, Engineers & Acoustical Consultants, *Accessed [online]*, March 2016. URL `ease.afmg.eu`.

[65] Andrzej Kulowski. Algorithmic Representation of the Ray Tracing Technique. *Applied Acoustics*, 18(6):449–469, 1985.

[66] G. M. Naylor. ODEON - Another Hybrid Room Acoustical Model. *Applied Acoustics*, 38(2-4):131–143, 1993.

[67] Hilmar Lehnert. Systematic Errors of the Ray-Tracing Algorithm. *Applied Acoustics*, 38(2-4):207–221, 1993.

[68] Murray Hodgson. Evidence of Diffuse Surface Reflections in Rooms. *Journal of the Acoustical Society of America*, 89(2):765–771, 1991.

[69] Murray Hodgson. On the prediction of sound fields in large empty rooms. *Journal of the Acoustical Society of America*, 84:253–261, 1988.

[70] Jens Holger Rindel. Computer Simulation Techniques for Acoustical Design of Rooms. *Acoustics Australia*, 23(3):81–86, 1995.

[71] 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, 1996.

[72] Rendell R. Torres, U. Peter Svensson, and Mendel Kleiner. Computation of edge diffraction for more accurate room acoustics auralization. *Journal of the Acoustical Society of America*, 109(2):600–610, 2001.

[73] Carl F. Eyring. Reverberation times in 'dead' rooms. *Journal of the Acoustical Society of America*, 1:217–241, 1930.

[74] Jont B. Allen and David A. Berkley. Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65:943–950, 1979.

[75] Jeffrey Borish. Extension of the image model to arbitrary polyhedra. *Journal of the Acoustical Society of America*, 75:1827–1836, 1984.

[76] 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 *25^{th} annual conference on computer graphics and interactive techniques, SIGGRAPH '98, pages = 21–32, year = 1998.*

[77] Thomas Funkhouser, Nicolas Tsingos, Ingrid Carlbom, Gary Elko, Mohan Sondhi, and Jim West. A beam tracing method for interactive architectural acoustics. *Journal of the Acoustical Society of America*, 115(2):739–756, 2004.

[78] S. Siltanen, T. Lokki, and L. Savioja. Geometry reduction in room acoustics modeling. *Acta Acustica united with Acustica*, 94(3):410–418, 2008.

[79] Dirk Schröder and Tobias Lentz. Real-Time Processing of Image Sources Using Binary Space Partitioning. *Journal of the Audio Engineering Society (JAES)*, 54 (7–8):604–619, 2006.

[80] Alex Southern, Samuel Siltanen, and Lauri Savioja. Spatial room impulse responses with a hybrid modelling method. In *Audio Engineering Society Convention 130*, May 2011. URL `http://www.aes.org/e-lib/browse.cfm?elib=15852`.

[81] Alex 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.

[82] I. A. Drumm. The Application of Adaptive Beam Tracing and Managed DirectX for the Visualisation and Auralisation of Virtual Environments. In *9th International Conference on Information Visualisation (IV '05)*, pages 961–965, 2005.

[83] I. A. Drumm and Y. W. Lam. The adaptive beam-tracing algorithm. *Journal of the Acoustical Society of America*, 107(3):1405–1412, 2000.

[84] Cindy M. Goral, Kenneth E. Torrance, Donal P. Greenberg, and Bennett Battaile. Modeling the Interaction of Light Between Diffuse Surfaces. *ACM SIGGRAPH Computer Graphics*, 18(3):213–222, 1984.

[85] 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*, 98:288–293, 1995.

[86] Eva-Marie Nosal. Room sound field prediction by acoustical radiosity. Master's thesis, University of British Columbia, Canada, 2003.

[87] Kane Yee. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14:463–466, 1966.

[88] Lauri Savioja, Timo Rinne, and Tapio Takala. Simulation of room acoustics with a 3-D finite difference mesh. In *Proc. of the Int. Computer Music Conference*, pages 463–466, 1999.

[89] D. Botteldooren. Acoustic finite-difference time-domain simulation of low-frequency room acoustic problems. *J. Acoust. Soc. Am.*, 95(5):2313–2319, 1994.

[90] Vesa Välimäki, Matti Karjalainen, Zoltàn Jànosy, and Unto K. Laine. A real-time DSP implmentation of a flute model. In *Int. Conference on Acoustics, Speech and Signal Processing, ICASSP*, April 1992.

[91] Perry R. Cook. A meta-wind-instrument physical model and a meta-controller for real-time performance control. In *Int. Computer Music Conference and Festival*, Oct. 1992.

[92] Rami Hänninen and Vesa Välimäki. An improved digital waveguide model of a flute with fractional delay filters. In *Nordic Acoustical Meeting, NAM*, June 1996.

[93] Scott van Duyne and Julius O. Smith III. Physical modeling with the 2-d digital waveguide mesh. In *Proc. of the International Computer Music Conference (ICMC)*, 1993.

[94] Mark J. Beeson and Damian T. Murphy. Roomweaver: A digital waveguide mesh based room acoustics research tool. In *7th Int. Conf. on Digital Audio Effects (DAFx)*, Oct. 2004.

[95] Marc-Laurent Aird. *Musical Instrument Modelling Using Digital Waveguides.* PhD thesis, University of Bath, 2002.

[96] Damian Murphy, Mark Beeson, Simon Shelley, and Alastair Moore. Hybrid room impulse response synthesis in digital waveguide mesh based room acoustics simulation. In *11th Int. Conf. on Digital Audio Effects (DAFx)*, Sept. 2008.

[97] Jack Mullen, David M. Howard, and Damian T. Murphy. Real-time dynamic articulations in the 2d waveguide mesh vocal tract model. *IEEE Transactions on Audio, Speech and Language Processing*, 15(2):577–585, 2007.

[98] Harold C. Martin and Graham F. Carey. *Introduction to Finite Element Analysis: Theory and Application.* Tata McGraw-Hill Publishing Company Ltd., New Delhi, 1973.

[99] K. C. Rockey, H. R. Evans, D. W. Griffiths, and D. A. Nethercot. *The Finite Element Method: A Basic Introduction for Engineers*. Granada Publishing Ltd., London, 1975.

[100] 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.

[101] Lauri Savioja, Dinesh Manocha, and Ming C. Lin. Use of gpus in room acoustic modelling and auralization. In *International Symposium on Room Acoustics (ISRA)*, Aug. 2010.

[102] George Kalliris, George Papanikolaou, Anastasia Papastefanou, and Christos Sevastiadis. An amphitheatric hall modal analysis using the finite element method compared to in situ measurements. In *Audio Engineering Society Convention 120*, May 2006. URL `http://www.aes.org/e-lib/browse.cfm?elib=13543`.

[103] Christos Sevastiadis, Geoge Kalliris, and George Papanikolaou. Analysis tool development for the investigation of low frequency room acoustics by means of finite element method. In *Audio Engineering Society Convention 128*, May 2010. URL `http://www.aes.org/e-lib/browse.cfm?elib=15266`.

[104] Andrzej Pietrzyk and Mendel Kleiner. The application of the finite-element method to the prediction of soundfields of small rooms at low frequencies. In *Audio Engineering Society Convention 102*, March 1997. URL `http://www.aes.org/e-lib/browse.cfm?elib=7356`.

[105] Jonathan Hargreaves. *Time Domain Boundary Element Method for Room Acoustics*. PhD thesis, University of Salford, 2007.

[106] Ernest Granados. Computation of the impulse response of small rooms with the time domain boundary element method. Master's thesis, Berlin Technical University, Germany, 2011.

[107] B. Katz. Boundary element method calculation of individual head-related transfer function. i. rigid model calculation. *J. Acoust. Soc. Am.*, 110(5):2440–2448, 2001.

[108] Y. Kahana and P. Nelson. Boundary element simulations of the transfer function of human heads and baffled pinnae using accurate geometric functions. *J. Sound Vib.*, 300:552–579, 2007.

[109] Jonathan Hargreaves and Yiu Lam. A high-frequency bem for 3d acoustic scattering. In *Int. Cong. on Sound and Vibration*, July 2015.

[110] 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, 2008.

[111] Nikunj Raghuvanshi, Rahul Narain, Rahul Narain, and Ming C. Lin. Efficient and Accurate Sound Propagation Using Adaptive Rectangular Decomposition. *IEEE Transations on Visualization and Computer Graphics*, 15(9), 2009.

[112] Nikunj Raghuvanshi, Brandon Lloyd, Naga K. Govindaraju, and Ming C. Lin. Efficient numerical acoustic simulation on graphics processors using adaptive rectangular decomposition. In *EAA Symposium on Auralization*, June. 2009.

[113] Ravish Mehra, Nikunj Raghuvanshi, Ming C. Lin, and Dinesh Manocha. Efficient GPU-Based Solver for the Acoustic Wave Equation. *Acta Acustica united with Acustica*, 73(2):83–94, 2012.

[114] Jonathan Botts and Lauri Savioja. Extension of a spectral time-stepping domain decomposition method for dispersive and dissipative wave propagation. *Journal of the Acoustical Society of America*, 137(4), 2015.

[115] Rudolf Rabenstein. Discrete Simulation Models for Multidimensional Systems Based on Functional Transformations. *Mathematics in Signal Processing, Oxford University Press (Editor: J. G. McWhirter)*, 4, 1998.

[116] Rudolf Rabenstein. Transfer function models for continuous and discrete multidimensional systems. In *IX European Signal Processing Conference (EUSIPCO-98)*, Sept. 1998.

[117] Stefan Petrausch and Rudolf Rabenstein. Two-dimensional block based physical modeling with the functional transformation method. In 4^{th}.

[118] R. Rabenstein and L. Trautmann. Solution of vector partial differential equations by transfer function models. In *IEEE Int. Symposium on Circuits and Systems (ISCAS)*, May/June 1999.

[119] L. Trautmann, S. Petrausch, and R. Rabenstein. Physical modelling of drums by transfer function methods. In *Int. Conf. on Acoustics, Speech and Signal Processing*, May 2001.

[120] Stefan Petrausch and Rudolf Rabenstein. Performance Analysis of Wave Field Simulation with the Functional Transformation Method. In *Audio Engineering Society Convention 120*, May 2006. URL `http://www.aes.org/e-lib/browse.cfm?elib=13596`.

[121] Stefan Petrausch and Rudolf Rabenstein. Efficient 3D simulation of wave propagation with the functional transformation method. In *18^{th} Symposium of Simulation Technique*, Sept. 2005.

[122] Stefan Petrausch. *Block Based Physical Modeling.* PhD thesis, University of Erlangen-Nuremberg, 2007.

[123] Konrad Kowalczyk and Maarten van Walstijn. Room acoustics simulation using 3-D compact explicit FDTD schemes. *IEEE Transactions on Audio, Speech and Language Processing*, 19(1):34–46, 2011.

[124] Konrad Kowalczyk and Maarten van Walstijn. Modelling frequency-dependent boundaries as digital impedance filters in FDTD and K-DWM room acoustic simulations. *Journal of the Audio Engineering Society (JAES)*, 56(7/8):596–583, 2008.

[125] NESS Project. Next generation sound synthesis, *Accessed [online]*, March 2016. URL `http://www.ness-music.eu/`.

[126] 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(6):891–906, 2008.

[127] Antti Kelloniemi, Lauri Savioja, and Vesa Valimaki. Simulation of room acoustics using 2-D digital waveguide meshes. In *Int. Conference on Acoustics, Speech and Signal Processing, ICASSP*, May 2006.

[128] Craig J. Webb and Alan Gray. Large-scale virtual acoustics simulation at audio rates using three dimensional finite difference time domain and multiple gpus. In *21^{st} Int. Congr. on Acoustics*, June 2013.

[129] Jonathan Sheaffer, Craig Webb, and Bruno Fazenda. Modelling binaural receivers in finite difference simulation of room acoustics. In *21st Int. Congr. on Acoustics*, June 2013.

[130] Craig J. Webb and Stefan Bilbao. Computing room acoustics with cuda - 3d fdtd schemes with boundary losses and viscosity. In *Int. Conference on Acoustics, Speech and Signal Processing, ICASSP*, May 2011.

[131] Brian Hamilton, Craig Webb, Alan Gray, and Stefan Bilbao. Large stencil operations for gpu-based 3-d acoustics simulations. In *18th Int. Conf. on Digital Audio Effects (DAFx)*, Nov. 2015.

[132] Jelle van Mourik and Damian Murphy. Explicit Higher-Order FDTD Schemes for 3D Room Acoustic Simulation. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 22(12):2003–2011, 2014.

[133] Stefan Bilbao and Craig Webb. Physical modling of timpani drums in 3D on GPGPUs. *Journal of the Audio Engineering Society (JAES)*, 61(10):737–748, 2013.

[134] Jelle van Mourik and Damian Murphy. Hybrid acoustic modelling historic spaces using blender. In *Proc. of Forum Acusticum*, 2014.

[135] Damian Murphy and Sorrel Harriet. Auralisation of finite difference time domain simulations of sonic crystal noise barriers in an urban environment. In *10th European Congress and Exposition on Noise Control Engineering, EuroNoise*, June 2015.

[136] A. Taflove and S.C. Hagness. *Computational Electrodynamics: The Finite Difference Time-Domain Method, 2nd ed.* Artech House, Norwood, MA, 2000.

[137] Jukka Saarelma, Jonathan Botts, Brian Hamilton, and Lauri Savioja. Audibility of dispersion error in room acoustic finite difference time-domain simulation as a function of simulation distance. *J. Acoust. Soc. Am.*, 139(4):1822–1832, 2015.

[138] Alex Southern, Tapio Lokki, Lauri Savioja, and Damian Murphy. The perceptual effects of dispersion error on room acoustic model auralization. In *Proc. of Forum Acusticum*, 2011.

[139] Damian T. Murphy, Alex Southern, and Lauri Savioja. Source excitation strategies for obtaining impulse responses in finite difference time domain room acoustics simulations. *Applied Acoustics*, 82:6–14, 2014.

[140] Jonathan Sheaffer, Maarten van Walstijn, and Bruno Fazenda. Physical and numerical constraints in source modeling for finite difference simulation of room acoustics. *Journal of the Acoustical Society of America*, 135(1):251–261, 2014.

[141] Michael Gerzon. Periphony: With-Height Sound Reproduction. *Journal of the Audio Engineering Society (JAES)*, 21(1):2–10, 1973.

[142] Michael Gerzon. The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound. In *Audio Engineering Society Convention 50*, March 1975. URL `http://www.aes.org/e-lib/browse.cfm?elib=2466`.

[143] Alex Southern and Damian Murphy. 2^{nd} order spherical harmonic spatial encoding of digital waveguide mes room acoustic models. In *10^{th} Int. Conf. on Digital Audio Effects (DAFx)*, Sept. 2007.

[144] Alex Southern and Damian Murphy. Methods for 2^{nd} order spherical harmonic spatial encoding in digital waveguide mesh virtual acoustic simulations. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, (WASPAA)*, Oct. 2007.

[145] Alex Southern and Damian Murphy. A second order differential microphone technique for spatially encoding virtual room acoustics. In *Audio Engineering Society Convention 124*, May 2008. URL `http://www.aes.org/e-lib/browse.cfm?elib=14462`.

[146] Alex Southern, Damian Murphy, and Lauri Savioja. Spatial Encoding of Finite Difference Time Domain Acoustic Models for Auralization. *IEEE Transactions on Audio, Speech and Language Processing*, 20(9):2420–2432, 2012.

[147] Hüseyin Hacıhabiboğlu, Banu Günel, and Zoran Cvetković. Simulation of Directional Microphones in Digital Waveguide Mesh-Based Models of Room Acoustics. *IEEE Transactions on Audio, Speech and Language Processing*, 18(2):213–223, 2010.

[148] Alan Blumlein. U.K. patent 394,325, 1931. Reprinted in *Stereophonic Techniques* (Audio Engineering Society, New York, 1986).

[149] Stefan Bilbao, Brian Hamilton, and Lauri Savioja. Finite Volume Time Domain Room Acoustics Simulation under General Impedance Boundary Conditions. *IEEE Transactions on Audio, Speech and Language Processing*, 24(1): 161–173, 2016.

[150] Konrad Kowalczyk and Maarten van Walstijn. Formulation of a locally reacting wall in finite difference modelling of acoustic spaces. In *Proc. of Int. Symp. Room Acoustics (ISRA)*, Sept. 2007.

[151] 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, 2014.

[152] Stephen Oxnard, Dean O'Brien, Jelle van Mourik, and Damian Murphy. Frequency-Dependent Absorbing Boundary Implementation in 3D Finite Difference Time Domain Room Acoustics Simulations. In *Proceedings of EURONOISE*, Maastricht, The Netherlands, 2015.

[153] Jonathan Sheaffer, Bruno Fazenda, Damian Murphy, and Jamie Angus. A simple multiband approach for solving frequency dependent problems in numerical time domain methods. In *Proc. of Forum Acusticum*, June 2011.

[154] Samuel Siltanen, Alex Southern, and Lauri Savioja. Finite-difference time domain method source calibration for hybrid acoustics modeling. In *Int. Conference on Acoustics, Speech and Signal Processing, ICASSP*, May 2013.

[155] MathWorks. Documentation: fdesign.arbmag, *Accessed [online]*, July 2016. URL `http://uk.mathworks.com/help/dsp/ref/fdesign.arbmag.html`.

[156] Stephen Oxnard and Damian Murphy. Room Impulse Response Synthesis Based on a 2D Multi-Plane FDTD Hybrid Acoustic Model. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, (WASPAA)*, Oct. 2013.

[157] R. J. Allemang. The modal assurance criterion-Twenty years of use and abuse. *Journal of Sound and Vibration*, 37(8):14–21, 2003.

[158] Kihong Shin. An alternative approach to measure similarity between two deterministic transient signals. *Journal of Sound and Vibration*, 37(1):434–445, 2016.

[159] TSL Products. ST450 MKII Soundfield Portable, *Accessed [online]*, March 2016. URL `http://www.tslproducts.com/soundfield/st450-soundfield-portable//`.

[160] Alex Southern. *Acoustic Parameter Toolbox - MATLAB Function Package*. Aalto University, Aalto, Finland, 2011.

[161] Ingolf Bork. Report on the 3rd Round Robin on Room Acoustical Computer Simulation - Part 1: Measurements. *Acta Acustica united wth Acustica*, 91:740–752, 2005.

[162] Adrian James, Bengt-Inge Dalenbäck, and Amber Naqvi. Computer Modelling with CATT-Acoustic - Theory and Practice of Diffuse Reflection and Array Modeling - Technical Paper, *Accessed [online]*, Sept. 2016. URL `http://www.adrianjamesacoustics.co.uk/technicalstuff/cattacoustic.pdf`.

[163] CATT. *CATT-Acoustic v8g - Room Acoustics Prediction and Walkthrough Auralization, User's Manual*. Gothenburg, Sweden, 2007.

[164] Alex Southern. PTB Room Geometry - Sketchup Make (.skp) File. Personal Communication, May 2015.

[165] Alex Southern. *VTK Export Plugin for Sketchup Make*. Aalto University, Aalto, Finland, 2011.

[166] Alex Southern. *WaveModeller .vtk Voxeliser*. Aalto University, Aalto, Finland, 2011.

[167] Jeremy Wells, Damian T. Murphy, and Mark J. Beeson. Temporal matching of 2d and 3d wave-based acoustic modeling for efficient and realistic simulation of rooms. In *Audio Engineering Society Convention 126*, May 2009. URL `http://www.aes.org/e-lib/browse.cfm?elib=14893`.