THE UNIVERSITY *of* EDINBURGH

Finite Difference and Finite Volume Methods for Wave-based Modelling of Room Acoustics

Brian Hamilton

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF EDINBURGH

2016

Abstract

Wave-based models of sound propagation can be used to predict and synthesize sounds as they would be heard naturally in room acoustic environments. The numerical simulation of such models with traditional time-stepping grid-based methods can be an expensive process, due to the sheer size of listening environments (e.g., auditoriums and concert halls) and due to the temporal resolution required by audio rates that resolve frequencies up to the limit of human hearing. Finite difference methods comprise a simple starting point for such simulations, but they are known to suffer from approximation errors that may necessitate expensive grid refinements in order to achieve sufficient levels of accuracy. As such, a significant amount of research has gone into designing finite difference methods that are highly accurate while remaining computationally efficient.

The problem of designing and using accurate finite difference schemes is compounded by the fact that room acoustics models require complex boundary conditions to model frequency-dependent wall impedances over non-trivial geometries. The implementation of such boundary conditions in a numerically stable manner has been a challenge for some time. Stable boundary conditions for finite difference room acoustics simulations have been formulated in the past, but generally they have only been useful in modelling trivial geometries (e.g., idealised shoebox halls). Finite volume methods have recently been shown to be a viable solution to the problem of complex boundary conditions over non-trivial geometries, and they also allow for the use of energy methods for numerical stability analyses. Finite volume methods lend themselves naturally to fully unstructured grids and they can simplify to the types of grids typically used in finite difference methods. This allows for room acoustics simulation models that balance the simplicity of finite difference methods for wave propagation in air with the detail of finite volume methods for the modelling of complex boundaries.

This thesis is an exploration of these two distinct, yet related, approaches to wave-based room acoustic simulations. The overarching theme in this investigation is the balance between accuracy, computational efficiency, and numerical stability. Higher-order and optimised schemes in two and three spatial dimensions are derived and compared, towards the goal of finding accurate and efficient finite difference schemes. Numerical stability is analysed using frequency-domain analyses, as well as energy techniques whenever possible, allowing for stable and frequency-dependent boundary conditions appropriate for room acoustics modelling. Along the way, the use of non-Cartesian grids is investigated, geometric relationships between certain finite difference and finite volume schemes are explored, and some problems associated to staircasing effects at boundaries are considered. Also, models of sound absorption in air are incorporated into these numerical schemes, using physical parameters that are appropriate for room acoustic scenarios.

Declaration

I declare that that this thesis was composed by myself, that the work herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Brian Hamilton)

Acknowledgements

There are many people that I must thank for help and support throughout this PhD programme.

My deepest thanks go to my main PhD supervisor: Dr. Stefan Bilbao — for guidance throughout all phases of this work; for your wisdom in all things related to acoustics, signal processing, and numerical methods; and also for your good humour.

I would like to thank Prof. Lauri Savioja for co-supervising this thesis, and for the many welcoming visits I have made to Aalto University. Thanks also to Jonathan Botts and Jukka Saarelma for collaborations and fruitful discussions regarding all aspects of numerical wave simulation.

Thanks for the support from the friends I have made in Edinburgh over the years. Thanks to my colleagues in the Acoustics and Audio Group for encouragement along the way. In particular, special thanks to Alberto and Craig for collaborations and for proofreading portions of this thesis.

Thanks to Prof. Dick Botteldooren, external examiner on this thesis, and Dr. James Hopgood, internal examiner, for thorough reviews of this thesis and for many helpful suggestions and corrections.

This work would not have been possible without the financial support of the European Research Council, under grant StG-2011-279068-NESS, and of the Natural Sciences and Engineering Research Council of Canada (NSERC) under the Postgraduate Scholarships-Doctoral (PGS-D) Program.

Lastly, thanks to my parents for a lifetime of encouragement, and sincere thanks to Laura for personal support throughout this arduous thesis-writing journey.

Contents

Chapter 1

Introduction

1.1 Introductory remarks

One of the aims of room acoustics modelling is to reproduce the behaviour of sound waves in an enclosed space (i.e., the acoustics of a room), so that one can predict how a room will sound, and eventually to understand what makes certain rooms "sound better" than others. The ability to accurately reproduce the acoustics of a room has many applications: it can aid in the architectural design of any enclosed space—in particular, spaces meant for listening and musical production, such as sound recording studios, auditoriums, and concert halls [140]; it can be used for the creation of virtual (acoustic) spaces [226]; and it can serve as a digital audio effect [315], adding artificial reverberation [284] to musical compositions and recordings.

A natural starting point for a model of room acoustics is the wave equation, a partial differential equation (PDE) that describes sound waves in air [73, 74], among other wave phenomena. By combining this PDE with a geometrical model of a room (the domain) and a description of the enclosing walls (the boundary conditions), one arrives at a basic mathematical and physical model that encapsulates the acoustics of a room [140]. Solutions to such PDE problems can be shown to exist and to be unique (for a given domain and set of boundary conditions), but they are not generally known in analytic forms. As such, these types of problems must be solved numerically. In other words, the acoustics of a room must be simulated, and typically with the aid of a computer.

For the last 50 years, room acoustics simulation techniques have relied mainly on the simplifying assumptions of geometrical acoustics—wherein sound is assumed to behave akin to light rays. From this simplified description of room acoustics, simulation techniques could be employed at relatively low computational costs, on digital computers in the late 1960s and and in the 1970s [138, 232, 3]. Such simulation techniques—henceforth, *geometric methods*—are still used today due to their simplicity [190], although they make use of high-frequency approximations, so they are, in general, not reliable in low (audible) frequencies [228]. Research in geometric methods has progressed towards the inclusion of missing wave-like properties (such as diffraction and phase effects) in order to extend their validity into low frequencies, but this requires additional complexities and compromises the inherent simplicity of these methods.

In the room acoustics community, simulation techniques (numerical methods) that aim to directly solve the wave equation, or related PDEs, without such simplifying assumptions, are classified as *wave-based* methods [298]. In contrast to geometric methods, wave-based methods can be computationally expensive, as they typically rely on a grid-based, time-stepping approach, which carries unavoidable

computational costs due to sampling considerations in time (related to the range of audible frequencies) and space (related to the spatial wavelengths of audible frequencies). Furthermore, the desire to obtain an accurate numerical solution, relatively free of approximation errors, may necessitate expensive grid refinements. On the other hand, as they are free from the simplifying assumptions of geometrical acoustics, wave-based methods allow for all wave-like phenomena to be simulated in a natural way.

At the time when geometric methods started to be used for room acoustics simulations (in the 1970s), the use of wave-based methods for room acoustics simulation would have been unimaginable. In 1973, Schroeder said it would be "utterly impossible" [233] using computers at the time. In the early 1990s, when wave-based approaches started to become computationally viable, finite difference methods—a classic family of numerical methods with a long history in simulating wave equations [60, 206, 86, 2] and related PDE systems [311, 264, 195, 295]—started to be applied to the problem of simulating 3-D room acoustics in low-frequencies, using mathematical formulations stemming from analogous equations in electromagnetics [51, 37] and, independently, stemming from developments in *digital waveguide* sound synthesis techniques [246, 286, 223]. Other well-known families of numerical methods were also applied to room acoustic problems, including finite volume (in the time-domain) [37, 38], and finite element and boundary element methods in both time- and frequency-domains (see [262] and references therein). With regards to the methods implemented for the wave equation itself (i.e., in the time-domain), finite difference methods seem to have gained the most popularity over the years (e.g., [181, 48, 218, 222, 136]), most likely due to their simplicity in formulation and ease in implementation.

While wave-based methods will most likely remain more expensive than the simplest forms of geometric methods for some time, there is no question that the geometrical acoustics model is insufficient as a description of room acoustics in low frequencies, where wave diffraction and interference effects play a major role [140]. Consequently, finite difference methods (or more generally, wave-based methods) are increasingly being seen as necessary alternatives to geometric methods for low frequencies [299]. A balance of these two distinct approaches may be appropriate, i.e., a hybridization of wave-based and geometric methods, respectively used for low and high frequencies [243, 299, 228], where the division of frequencies is justified by perceptual characteristics of large-room acoustics [299], or more strictly—adhering to the fundamental principles of geometric acoustics—by the dimensions of surfaces in the room under analysis. However, as wave-based methods are not limited to any frequency range (in theory), finite difference methods—or any other wave-based methods—can be seen as *complete* alternatives to geometric methods when accuracy is of primary concern, as they are ultimately based on a more physically correct and generally applicable description of acoustics, from which perceptual characteristics of room acoustics can be reproduced naturally. Indeed, finite difference simulations of large spaces in the full audible range carry immense computational costs (by today's standards), but they are increasingly becoming more accessible thanks to continued increases in computing power. Furthermore, there is currently an important trend towards expanding computing power into parallel architectures, and these methods benefit greatly from parallel implementations, such as on repurposed computer graphics hardware [210, 236, 303, 215].

It should also be mentioned that geometric methods are typically only used for the calculation of the various global attributes used in room acoustics (such as reverberation time and clarity) [299] or for the calculation of point-to-point impulse responses to be convolved with dry audio signals for *auralization* purposes [128]. The scope of time-domain finite difference methods is not limited to these interests. Finite difference methods also comprise essential tools in the *physical modelling sound synthesis* framework [24]—a framework that is widely used in the area of computer-based music [209]. As

they inherently solve for the entire acoustic field in a space, finite difference methods allow for the embedding and two-way acoustical coupling with other physically modelled sounding objects that occupy space and emit sound directionally according to their physical shape, such as, e.g., virtual percussion instruments [202, 33, 273]—something that would be difficult (to say the least) to accomplish from a ray-based (geometrical) description of acoustics. Not only that, there is scope for the integration of head and torso models (occupying space), in order to extract binaural audio in a natural way [310, 305], as well as the ability to have multiple sources—potentially occupying space, and with directionality being modelled naturally—and multiple listening positions, moving throughout the space in the run-time loop at no extra cost.

While finite difference methods have shown promise over the years for room acoustics modelling (and for physically modelled sound synthesis), there is much work still to be done. In particular, and despite important recent work in the application of these numerical methods to room acoustics modelling [131, 136], two fundamental properties have been somewhat lacking: *consistency* (with the underlying boundary value problem) and (guarantees of) *numerical stability*. Without both of these properties, one cannot make the claim that a particular numerical method actually solves the PDE problem in question. More generally, without these properties, one cannot reliably close the gap between the mathematical model of room acoustics and the real-world room acoustics. Indeed, consistency and stability are relatively straightforward to obtain with finite difference schemes over the interior of the domain, which is largely governed by the wave equation (e.g., [60, 86, 241, 53]). But as it turns out, it is deceptively difficult to simultaneously achieve consistency over the boundary of the domain in room acoustics simulations and maintain numerical stability, as these types of simulations require accurate modelling of non-trivial boundary surfaces (i.e., beyond a rectangular box or infinite flat wall conveniently aligned with Cartesian grid axes), along with frequency-dependent behaviours over both the interior and the boundary of the domain. Furthermore, all of this has to be achieved while keeping computational costs to a minimum, as they are already quite heavy (as dictated by sampling considerations), and are additionally subject to the needs of mitigating approximation errors, mainly in the form of numerical dispersion over the interior of the domain. By addressing some of these issues—using finite difference methods and closely-related finite volume methods—this thesis can be seen as a step towards the ultimate goal of accurate, efficient, and detailed room acoustics numerical modelling.

1.2 Thesis objectives

The objective of this thesis is to develop an efficient framework that allows for the accurate and numerically stable modelling of the wave behaviour of sound in an enclosed space, while incorporating the following details:

- Accurate geometrical modelling of non-trivial domains
- Accurate numerical modelling of general frequency-dependent boundary conditions
- Accurate numerical modelling of sound absorption in air

To achieve these goals, finite difference methods will be primarily used, operating over regular grids in order to cover most of the *interior* of the domain, which is governed, largely, by the acoustical wave equation. The accurate and efficient simulation of the wave equation in *free-space* is thus an important area of investigation for the purposes of this thesis. Finite volume methods, operating over unstructured (or structured) grids will be employed as a generalisation of finite difference methods, in

order to address concerns at the boundaries. As such, the relationship between finite volume and finite difference methods, using not only conventional Cartesian grids, is an important topic to be addressed in this thesis. Energy methods will be essential in order to ensure numerical stability, as such techniques allow for a global and flexible approach to stability analysis. Not only that, they provide numerical energy calculations that can be used to verify the numerical correctness (down to the limits of machine precision) of practical implementations of these methods. Frequency-domain stability analyses will also be used for the problem of free-space wave propagation, as well as matrix stability analyses to discuss some of the challenges associated with boundary modelling using only finite difference methods, and in discussing underlying problems in other popular approaches that can be found in the literature.

1.3 Outline of thesis

An outline for the rest of the thesis is as follows.

Chapter 2

Chapter 2 is a background chapter. It presents the model equations and boundary conditions that we aim to discretise in the subsequent chapters. It also provides a brief description of the wave-behaviour of sound, including details on sound absorption in air, simple point sources, plane-waves and reflection coefficients, and the concept of a room impulse response. This chapter also includes a review of computational methods that are commonly used for room acoustics, covering (briefly) geometric methods, and covering some other wave-based methods that have been applied to the simulation of room acoustics. A detailed review of related work in the area of finite differences for room acoustics is also provided in this chapter.

Chapter 3

Chapter 3 is an introduction to the use of finite difference methods for room acoustics modelling in three spatial dimensions, under the guarantees of energy-based stability, and considering frequency-dependent losses over the interior and boundaries. In this chapter, we discretise most of the equations and boundary conditions presented in Chapter 2, using the simplest Cartesian finite difference scheme. As such, this chapter presents a simple, yet detailed, discrete model for wave-based room acoustics, but certainly not one that completely addresses the concerns of this thesis. In another sense, this chapter is an introduction to some of the challenges associated with the ultimate goals of this thesis.

Chapter 4

Chapter 4, while admittedly long, is about tackling the computational cost of accurate finite difference modelling of wave propagation over the interior of a large enclosed space. The interest here is to identify accurate and computationally efficient schemes, from the point of view of mitigating numerical dispersion over a wide range of frequencies, and from the point of view of traditional accuracy (i.e., order of accuracy). As such, Chapter 4 is an exploration of finite difference schemes for the wave equation in free space (i.e., not considering boundary modelling), in one, two, and three spatial dimensions. We begin with the simplest form of the wave equation (in one spatial dimension) in order to illustrate and discuss, in a simple and clear manner, the important concepts that relate to the overall accuracy of these methods. This is followed by the wave equation in two and three spatial dimensions (the more

challenging cases). There should be, hopefully, a clear progression in the approaches used to derive schemes as we consider higher spatial dimensions, since in many cases, the possible schemes found in various dimensions are closely-related (such as, e.g., the well-known simple Cartesian one [60, 86]). This chapter explores computational stencils that reach out to more than just nearest-neighbours on a Cartesian grid, and it also explores the use of non-Cartesian (lattice-based) grids, including hexagonal grids in 2-D, and analogous grids in 3-D (face-centered cubic and body-centered cubic), which are somewhat unconventional in the wider use of finite difference methods. We make use of modified equation methods to derive schemes with high orders of accuracy, in both space and time, and we identify schemes that have numerical dispersion optimised over a wide range of frequencies. These schemes are also compared to conventional schemes that employ high-order differencing in space (with stencils reaching out along coordinate directions only). Many new schemes are identified, with potential for use not only in room acoustics. The development of schemes using non-Cartesian (hexagonal, face-centered cubic and body-centered cubic) grids and the analysis of these non-Cartesian schemes in the frequency domain, taking into account the appropriate non-Cartesian wavenumber domains of analysis, are important contributions of this chapter. Implicit formulations are also used to derive other (novel) high-order accurate schemes.

Chapter 5

In Chapter 5 we take a closer look at the challenges associated with finite difference boundary modelling in room acoustics, in order to identify what is missing from approaches previously seen in the literature. This chapter considers the simplest scheme in many spatial dimensions, and it considers the stability and accuracy of commonly-used finite difference boundary conditions, including those derived in Chapter 3. For stability analyses in this chapter, we make use of matrix and energy stability techniques, and we discuss the relationships between these two approaches. Simple examples (such as boxes conveniently aligned with grid axes) are used to demonstrate fundamental and informative properties of these schemes. An important contribution of this chapter is to identify the problem of complex eigenvalues arising in the underlying Laplacian matrix for some widely-used boundary models, and to provide insights on instabilities that have been observed recently in the literature.

Chapter 6

Chapter 6 is concerned with the use of finite volume methods for room acoustics modelling. These finite volume schemes, while formulated for completely unstructured grids, are considered in a complementary role to finite difference methods, in order to accomplish the goals of this thesis for which regular-grid finite difference schemes are simply ill-equipped. An introduction to finite volume methods is provided, followed by energy analyses for stability conditions and frequency-independent boundary conditions, and implicit formulations. Connections are established between the finite volume framework and finite difference approaches, providing underlying geometrical interpretations for various digital waveguide mesh topologies (and associated finite difference schemes) that have appeared in the literature over the years, as well as some discrete boundary models. Additionally, in this chapter we investigate staircasing effects with various test cases, and we present detailed and complex finite volume (and finite difference) simulations of room acoustics, on non-trivial geometries, using fitted tilings of finite volume cells, and with general impedance frequency-dependent boundary conditions, while demonstrating stability and conservation of numerical energy to machine precision.

Chapter 7

In Chapter 7 we give concluding remarks and provide insights into future developments.

A glossary is provided at the end of the thesis, comprising a list of abbreviations and acronyms and a redundant list of symbols by chapter.

1.4 Main contributions

The main contributions of this thesis are as follows:

- A systematic presentation of the simple Cartesian scheme in 3-D considering energy-based stability and incorporating frequency-dependent losses over the interior and boundary (Chapter 3)
- A comprehensive analysis and comparison of high-order Cartesian grid–based schemes for the wave equation in 2-D and 3-D, leading to novel high-order accurate schemes (Sections 4.3.5–4.3.8 and 4.5.4–4.5.6)
- A derivation of novel high-order explicit and implicit finite difference schemes using modified equation methods employing non-Cartesian grids in two and three spatial dimensions (Sections 4.4.4, 4.4.5, 4.6.4, and 4.7)
- A comprehensive investigation and explanation of instabilities that can arise in seemingly innocuous approaches to finite difference boundary modelling, relating to some recently proposed and widely-used boundary models in the room acoustics literature (Section 5.4)
- An exposition of inconsistencies in the simulation of room decay times due to staircasing effects in finite difference modelling, revealing important shortcomings of finite difference methods, and other methods that exhibit staircasing effects (Section 6.6.2)
- Various simulations using fitted (irregular) finite volume cells in 2-D and 3-D, derived from Cartesian, hexagonal, and face-centered cubic grids, maintaining numerical stability and computational efficiency, and demonstrating conservation of numerical energy to machine precision when relevant (Sections 6.3, 6.6.1 and 6.6.2–6.8.1)
- Simulations over non-trivial domains, using general frequency-dependent losses modelling real-world materials (in collaboration with others), demonstrating insensitivity to the choice of the finite volume grid (mesh), indicating convergence to the unique underlying solution of a given room acoustics model (Section 6.8.2)

There are various incremental contributions along the way, and some perhaps worth mentioning are:

- A discussion of spatial decoupling effects in some commonly-used compact explicit schemes (Sections 4.3.4 and 4.6.1)
- A general family of parametrised discrete Laplacians on hexagonal grids (Section 4.4.2)
- A systematic derivation via energy methods of centered impedance boundary conditions in finite difference schemes over d-dimensional shoebox domains (Section 5.2.2)
- Establishing some connections between stability conditions derived using energy methods and those derived using matrix-type stability analyses (Section 5.2.2)

- Two novel multi-dimensional digital waveguide mesh topologies (Section 6.4.3)
- Finite volume interpretations of centered finite difference impedance boundary conditions 6.5.1
- An analysis and review of some digital waveguide mesh–type boundary conditions, from the point of view of corresponding finite volume tilings (Section 6.5)
- Provably-stable fixes to re-entrant corner instabilities observed in some commonly used finite difference impedance boundary conditions (Section 6.5.2)
- Establishing connections between finite volume frequency-dependent boundary conditions and some previously proposed finite difference frequency-dependent boundary conditions (Appendix A)

1.5 Related Publications

Elements of this thesis have been published in the following papers. The contribution of the current author is detailed in each case.

1. B. Hamilton and S. Bilbao. Hexagonal vs. rectilinear grids for explicit finite difference schemes for the two-dimensional wave equation. In *Proceedings of the International Congress of Acoustics (ICA)*, Montréal, Canada, June 2013.

   The current author produced the entire paper, with the manuscript edited by Bilbao.

2. S. Bilbao and B. Hamilton. Construction and optimization techniques for high order schemes for the two-dimensional wave equation. In *Proceedings of the International Congress of Acoustics (ICA)*, Montréal, Canada, June 2013.

   The current author edited the manuscript and helped with mathematical derivations.

3. B. Hamilton and S. Bilbao. On finite difference schemes for the 3-D wave equation using non-Cartesian grids. In *Proceedings of the Sound and Music Computing (SMC) Conference*, pages 592–599, Stockholm, Sweden, August 2013.

   The current author produced the entire paper, with the manuscript edited by Bilbao.

4. S. Bilbao, B. Hamilton, A. Torin, C. J. Webb, P. Graham, A. Gray, K. Kavoussanakis, and J. Perry. Large scale physical modeling sound synthesis. In *Proceedings of the Stockholm Musical Acoustics Conference (SMAC)*, pages 593–600, Stockholm, Sweden, August 2013.

   The current author helped write Section 6.4.

5. B. Hamilton and S. Bilbao. Fourth-order and optimised finite difference schemes for the 2-D wave equation. In *Proceedings of the Digital Audio Effects (DAFx) Workshop*, Maynooth, Ireland, September 2013.

   The current author produced the entire paper, with the manuscript edited by Bilbao.

6. B. Hamilton and C. J. Webb. Room acoustics modelling using GPU-accelerated finite difference and finite volume methods on a face-centered cubic grid. In *Proceedings of the Digital Audio Effects (DAFx) Workshop*, pages 336–343, Maynooth, Ireland, September 2013.

   The current author produced the entire paper, with the exception of Section 7 and coding of GPU simulations.

7. B. Hamilton. Sampling on a diamond grid and the tetrahedral digital waveguide mesh. *IEEE Signal Processing Letters*, 20(10):925–928, October 2013.

   The current author produced the entire paper.

8. B. Hamilton and A. Torin. Finite difference schemes on hexagonal grids for thin linear plates with finite volume boundaries. In *Proceedings of the Digital Audio Effects (DAFx) Workshop*, Erlangen, Germany, September 2014.

   The current author produced the entire paper, except for Section 2 and parts of Section 1 and 7. Torin also helped with derivations and edited the manuscript.

9. B. Hamilton, S. Bilbao, and C. J. Webb. Revisiting implicit finite difference schemes for 3-D room acoustics simulations on GPU. In *Proceedings of the Digital Audio Effects (DAFx) Workshop*, Erlangen, Germany, September 2014.

   The current author produced the entire paper, with the manuscript edited by Bilbao, and help with simulations from Webb.

10. B. Hamilton. Finite volume perspectives on finite difference schemes and boundary formulations for wave simulation. In *Proceedings of the Digital Audio Effects (DAFx) Workshop*, Erlangen, Germany, September 2014.

    The current author produced the entire paper.

11. A. Torin, B. Hamilton, and S. Bilbao. An energy conserving finite difference scheme for the simulation of collisions in snare drums. In *Proceedings of the Digital Audio Effects (DAFx) Workshop*, Erlangen, Germany, September 2014.

    The current author wrote Section 5.3.

12. B. Hamilton, S. Bilbao, and C. J. Webb. Improved finite difference schemes for a 3-D viscothermal wave equation on a GPU. In *Proceedings of Forum Acusticum*, Krakow, Poland, September 2014.

    The current author produced the entire paper, with the exception of some parts of Section 2 (written with Bilbao). The manuscript was edited by Bilbao, and Webb helped with GPU simulations.

13. B. Hamilton, C. J. Webb, A. Gray, and S. Bilbao. Large stencil operations for GPU-based 3-D acoustics simulations. In *Proceedings of the Digital Audio Effects (DAFx) Workshop*, Trondheim, Norway, December 2015.

    The current author produced the entire paper. Webb and Gray helped with GPU simulations and editing. Bilbao edited the manuscript.

14. S. Bilbao, B. Hamilton, J. Botts, and L. Savioja, "Finite volume time domain room acoustics simulation under general impedance boundary conditions," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 1, pp. 161–173, 2016.

    The current author produced Section VI and edited Sections I-IV and VII.

15. S. Bilbao and B. Hamilton. Finite volume modeling of viscothermal losses and frequency-dependent boundaries in room acoustics simulations. In *AES 60th Conference on Dereverberation and Reverberation of Audio, Music, and Speech*, Leuven, Belgium, February 2016.

    The current author produced Section 4.2 and edited the rest of the manuscript.

16. S. Bilbao, B. Hamilton, R. Harrison, and A. Torin. *Springer Handbook of Systematic Musicology*, chapter Finite Difference Schemes in Musical Acoustics: A Tutorial. Springer-Verlag, (to appear).

    The current author produced the section regarding the wave equation in two and three spatial dimensions.

17. J. Saarelma, J. Botts, B. Hamilton, and L. Savioja. Audibility of dispersion error in room acoustic FDTD simulation as a function of simulation distance. *Journal of the Acoustical Society of America*, vol. 139, no. 4, pp. 1822–1832, 2016.

    The current author edited the manuscript and consulted on aspects relating to numerical dispersion of non-Cartesian schemes.

18. B. Hamilton and S. Bilbao. Optimised 25-point finite difference schemes for the three-dimensional wave equation. In *Proceedings of the 22nd International Congress on Acoustics*, Buenos Aires, Argentina, September 2016.

    The current author produced the entire paper, with the manuscript edited by Bilbao.

19. B. Hamilton, C. J. Webb, N. D. Fletcher, and S. Bilbao. Finite difference room acoustics simulation with general impedance boundaries and viscothermal losses in air: Parallel implementation on multiple GPUs. In *Proceedings of the International Symposium on Musical and Room Acoustics*, La Plata, Argentina, September 2016.

    The current author produced the entire paper, with the manuscript edited by Bilbao. Fletcher helped with computer modelling and Webb helped with GPU programming.

20. S. Bilbao and B. Hamilton, "Wave-based room acoustics simulation: Explicit/implicit finite volume modeling of viscothermal losses and frequency-dependent boundaries," *Journal of the Audio Engineering Society*, 2017 (accepted for publication).

    The current author produced Section 3, co-authored Sections 1 and 4, and edited the rest of the manuscript.

Chapter 2

Background

In this chapter we review some of the physical equations describing room acoustics, including the main partial differential equation of interest (the 3-D wave equation), as well as models for sound absorption in air that are appropriate for room acoustics, and models for frequency-dependent boundary conditions. We also discuss the concepts of plane waves, reflection coefficients, simple sound sources, and room impulse responses. Finally, we review some of the computational methods (geometric and wave-based) that are typically used for room acoustics simulation and modelling.

2.1 Room acoustics fundamentals

In this section we present some of the equations related to wave-based models of 3-D room acoustics. We begin with model equations of lossless sound propagation in air, namely the conservation equations of linear acoustics and the 3-D wave equation. We discuss some properties of the wave equation that are relevant to room acoustics, followed by a brief description of simple sound sources in free-space. The subject of boundary conditions is covered next, with models of locally-reacting walls. This begins with simple frequency-dependent immittance (admittance or impedance) conditions, and ends with a more general frequency-dependent boundary model inspired by passive electrical circuit networks. We then discuss viscothermal extensions of the lossless sound propagation model in free space to dissipation due to viscothermal and molecular relaxation effects. Finally, we discuss the concept of room impulse responses in the context of room acoustic PDE models.

2.1.1 Linear and lossless sound propagation in air

From a general perspective, sound propagation in air is governed by the non-linear equations of fluid dynamics, such as the Navier-Stokes equations or the Euler equations. In the context of room acoustics, several simplifying assumptions can be made in order to reduce such equations to a simpler set of linear equations. For brevity we omit a detailed derivation, as this can be found in many acoustics textbooks (e.g. [174, 127, 46]); but briefly, the simplifications are as follows. In the absence of viscosity and thermal conduction, the Navier-Stokes can be reduced to the Euler equations, which in turn can be linearised about atmospheric pressure and density and zero mean velocity. Using various other assumptions (irrotational flow, homogeneous medium, source-free), one arrives at the following system

of equations:

$$\frac{1}{\rho c^2}\partial_t p = -\nabla \cdot \mathbf{v} \tag{2.1a}$$

$$\rho \partial_t \mathbf{v} = -\nabla p \tag{2.1b}$$

Here, $p = p(\mathbf{x}, t)$ is a scalar field that represents the acoustic pressure (the deviation from ambient pressure), in units of kg/(m·s^2), $\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$ is the vector particle velocity (components in m/s), c is the speed of sound in m/s, and ρ is the ambient density of the fluid (kg/m^3). The variable t represents time, $\mathbf{x} = (x, y, z)$ is a position vector in 3-D space, the symbol ∂_t represents a partial derivative with respect to time, $\nabla\cdot$ is the 3-D divergence operator ($\nabla \cdot \mathbf{v} = \partial_x v_x + \partial_y v_y + \partial_z v_z$), and ∇ is the 3-D gradient operator ($\nabla p = (\partial_x p, \partial_y p, \partial_z p)^{\mathrm{T}}$. As these equations feature first-order time derivatives, the equations themselves are said to be first-order. This system, which could be seen as two equations (one scalar and one vector) or as four scalar equations, comprises the conservation equations of linear acoustics. In particular, (2.1b) represents conservation of momentum. Eq. (2.1a) is known as the equation of continuity and it represents conservation of mass [127, 46, 208]. The bulk modulus κ sometimes appears in (2.1a) via the substitution:

$$\kappa = \rho c^2 \quad \Longrightarrow \quad \frac{1}{\kappa}\partial_t p = -\nabla \cdot \mathbf{v} \tag{2.2}$$

Initial conditions (initial data at $t = 0$) required for this system are $p(\mathbf{x}, 0)$ and $\mathbf{v}(\mathbf{x}, 0)$. In the context of room acoustics, c is generally taken be in the range of 340–344 m/s (indoor temperatures of air ranging from 15°C–20°C), and ρ is taken to be in the range of 1.21–1.25 kg/m^3 [194].

Apart from mass and momentum, another important conserved quantity is the total energy of the system. This is shown by multiplying (2.1a) by p and integrating over the full-space ($\mathbb{R}^3$), which gives:

$$\frac{1}{\rho c^2}\int_{\mathbb{R}^3} p(p_t)\,\mathrm{d}\mathbf{x} = -\int_{\mathbb{R}^3} p(\nabla \cdot \mathbf{v})\,\mathrm{d}\mathbf{x} \tag{2.3}$$

where here we have used the shorthand notation $p_t = \partial_t p, p_{tt} = \partial_t^2 p$, etc.; this shorthand will be used throughout this thesis when convenient. After making use of the divergence theorem (applied to the product of a scalar function and a vector field), the chain rule, and (2.1b),[1] we obtain an energy balance of the form:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(t) = 0\,, \quad \mathcal{H}(t) = \mathcal{V}(t) + \mathcal{T}(t) \tag{2.4}$$

where

$$\mathcal{V}(t) = \frac{1}{2\rho c^2}\int_{\mathbb{R}^3} p^2\,\mathrm{d}\mathbf{x}\,, \quad \mathcal{T}(t) = \frac{\rho}{2}\int_{\mathbb{R}^3} |\mathbf{v}|^2\,\mathrm{d}\mathbf{x} \tag{2.5}$$

The quantity $\mathcal{H}$ represents the total energy (in Joules), also known as the Hamiltonian, and $\mathcal{V}$ and $\mathcal{T}$ represent potential and kinetic energies, respectively. Eq. (2.4) says the total energy remains constant over time (conservation of energy).

While pressure and/or particle velocity are generally the quantities of interest in the room acoustics literature, another useful representation for acoustic fields in fluids is the acoustic velocity potential $\Psi = \Psi(\mathbf{x}, t)$ with units m^2/s. In the absence of losses, this quantity can be related to the pressure and

[1] More specifically, the divergence theorem says that: $-\int_{\mathbb{R}^3} p(\nabla \cdot \mathbf{v})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^3} \mathbf{v} \cdot (\nabla p)\,\mathrm{d}\mathbf{x}$ and the chain rule gives: $(\frac{1}{2}p^2)_t = p(p_t)$. After substituting in (2.1b) and subsequent application of the chain rule to scalar components of $\mathbf{v}$, one obtains (2.4) and (2.5).

velocity fields by [174]:

$$p = \rho \Psi_t \tag{2.6a}$$

$$\mathbf{v} = -\nabla \Psi \tag{2.6b}$$

where the second relation holds when the particle velocity field is irrotational. In some texts, Ψ takes the opposite sign [127, 46], affecting the above relations, but the convention taken here follows [174, 15]. This representation, although somewhat of an abstraction, is useful because it encapsulates both the pressure and particle velocity fields.

Another useful equation—central to this thesis—results by substituting the above relations into (2.1a):

$$\Psi_{tt} = c^2 \Delta \Psi \tag{2.7}$$

where $\Delta := \partial_x^2 + \partial_y^2 + \partial_z^2$ is known as the 3-D Laplacian operator. This second-order PDE is known as the 3-D wave equation and it will be the primary focus of this work. Initial conditions required for this system are $\Psi(\mathbf{x}, 0)$ and $\Psi_t(\mathbf{x}, 0)$.

In terms of Ψ, the energy balance (2.4) can also be expressed with:

$$\mathcal{V}(t) = \frac{\rho}{2c^2} \int_{\mathbb{R}^3} \Psi_t^2 \, \mathrm{d}\mathbf{x} \,, \quad \mathcal{T}(t) = \frac{\rho}{2} \int_{\mathbb{R}^3} |\nabla \Psi|^2 \, \mathrm{d}\mathbf{x} \tag{2.8}$$

where these terms could be derived from direct substitution of (2.6) into (2.5), or by multiplying (2.7) by $(\rho/c^2)\partial_t \Psi$ and making use of the divergence theorem and chain rule.[2]

One could also recover a wave equation in terms of the pressure field by applying $\rho \partial_t$ to (2.7) and making use of (2.6a), or by applying $\rho \partial_t$ to (2.1a) and substituting in (2.1b). This *pressure* wave equation is written as:

$$p_{tt} = c^2 \Delta p \tag{2.9}$$

which requires the initial conditions $p(\mathbf{x}, 0)$ and $p_t(\mathbf{x}, 0)$.

Travelling wave solutions to the wave equation

The wave equation gets its name from the fact that it admits wave-like solutions. A classic example of a wave-like solution results if we make the following simplifying assumption. Suppose that p only varies along the x-direction. It can then be written as a one-dimensional wave equation:

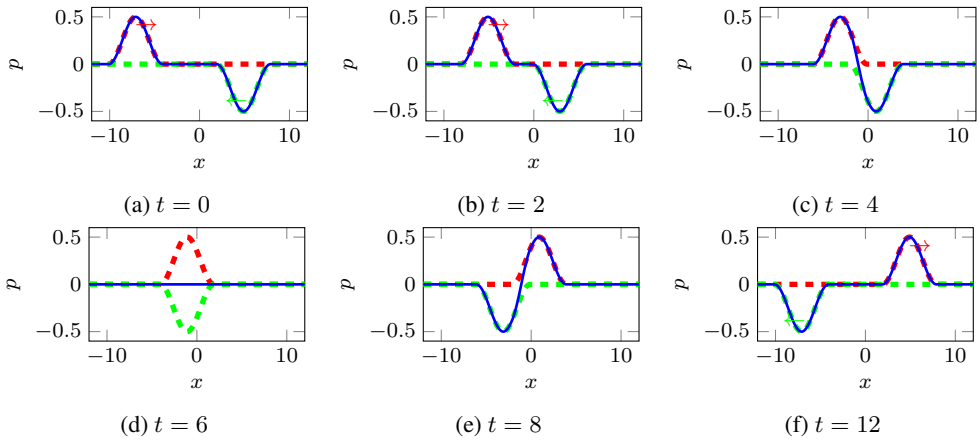
$$p_{tt} = c^2 p_{xx} \tag{2.10}$$

where here $p = p(x, t)$. Although originally derived as a model for a vibrating string [147], in which case the variable p would be replaced by another symbol (e.g., y) representing the transverse displacement of the string, the 1-D wave equation can represent acoustic wave propagation through a tube with a uniform cross-section [73]. A general solution for this 1-D wave equation can then be written as the sum of right- and left-going travelling waves:

$$p(x, t) = p^+(x - ct) + p^-(x + ct) \tag{2.11}$$

[2] In this case, the divergence theorem gives: $\int_{\mathbb{R}^3} \Psi_t (\nabla \cdot \nabla \Psi) \, \mathrm{d}\mathbf{x} = -\int_{\mathbb{R}^3} (\nabla \Psi_t) \cdot (\nabla \Psi) \, \mathrm{d}\mathbf{x}$, with $\Delta \Psi = \nabla \cdot (\nabla \Psi)$.

The above is known as d'Alembert's travelling wave solution, after [147]. An example of this solution is provided in Fig. 2.1, demonstrating the behaviour of the two components p^+ and p^- and their sum p.



(a) $t = 0$ (b) $t = 2$ (c) $t = 4$

(d) $t = 6$ (e) $t = 8$ (f) $t = 12$

Figure 2.1: Snapshots of the solution to 1-D wave equation with arbitrary left and right travelling wave components, and $c = 1$. Red: $p^+(x - ct)$, green: $p^-(x + ct)$, blue: $p = p^+(x - ct) + p^-(x + ct)$.

When p is varying along more than one direction, there is not a known general solution available that is so simply expressed as a sum of travelling waves. There is however another special case in 3-D that leads to similar expressions of travelling wave solutions. Consider a solution p that is radially symmetric about some point, such as the origin ($\mathbf{x} = 0$). The wave equation expressed in spherical coordinates, as opposed to Cartesian coordinates, then simplifies to:

$$p_{tt} - c^2 \frac{1}{r}(rp)_{rr} \tag{2.12}$$

where the term on the right is the Laplacian operator, expressed in spherical coordinates, applied to the spherically symmetric $p = p(r, t)$ with $r > 0$. This equation can also be seen as a 1-D wave equation in the variable rp, and a travelling wave solution, similar to d'Alembert's solution, results [15]:

$$p = \frac{1}{r}\left(p^+(r - ct) + p^-(r + ct)\right) \tag{2.13}$$

In this case, the travelling waves decay by a factor of $1/r$, where r is the distance from the origin. As a model of sound propagation in air, this $1/r$ scaling intuitively makes sense, since a listener at a distance from a sound source hears sounds softer than a listener close to the source does.

Another agreement with our intuition about sound arises from this solution in the case of a localised sound source (expressed as some initial data for the PDE). Consider the case where p^+ and p^- are localised around the origin at $t = 0$. The behaviour of p^+ predicts that the total pressure field p would be non-zero at some distance r after a time r/c. This relates to the experience that sound from a distance is generally heard *after* an amount of time that follows the time-instant of sound generation. The sound of thunder arriving after a flash of lighting is a classic example of this. On the other hand, the solution $p^-(r + ct)$ describes a sound that would be heard at a distance $r > 0$ for $t < 0$, or *before* the sound-generating event; this goes against intuition (thunder heard before lightning). Since p^- violates the principle of causality, it is often discarded in the context of sound propagation [211].

It is important to note that while the travelling wave component p^+ decays in amplitude over distance, the system as a whole remains lossless since the total energy of the system must obey (2.4). The

energy is spread evenly over the surface of a growing sphere, so the amplitude at any given point on the sphere must decrease appropriately. We will return to the discussion of sound sources in 3-D space (in Section 2.1.2), but for now we move on to other well-known solutions to the wave equation.

Plane-wave solutions to the wave equation

Another important wave-like solution is a *plane wave*, that is, a wave with a fixed rate of oscillation that varies, and travels, along a single direction (and is defined over the whole space). In 3-D, this means that the plane is constant along 2-D planes perpendicular to the direction of propagation, hence the name. A real-valued plane wave can be written as:

$$\Psi = A\cos(\boldsymbol{\beta}\cdot\mathbf{x} - \omega t) \tag{2.14}$$

where $A \in \mathbb{R}$ represents a maximum amplitude, $\boldsymbol{\beta} := (\beta_x, \beta_y, \beta_z) \in \mathbb{R}^3$ is a spatial angular frequency, also known as the *wave vector* (or vector wavenumber), and ω is an angular (temporal) frequency in rad/s (radians per second). The frequency in Hertz (cycles per second) is related to the angular frequency by $\omega = 2\pi f$. The magnitude of the wave vector, $\beta := |\boldsymbol{\beta}|$, is the *wavenumber* and has units rad/m. The direction of the wave vector, denoted $\hat{\boldsymbol{\beta}} := \boldsymbol{\beta}/|\boldsymbol{\beta}|$, will be called the *wave direction*.

Let us consider this plane wave as a trial solution to (2.7). After insertion into (2.7), we arrive at the following *dispersion relation* (a relation between temporal and spatial frequencies) that must be satisfied by the plane wave solution:

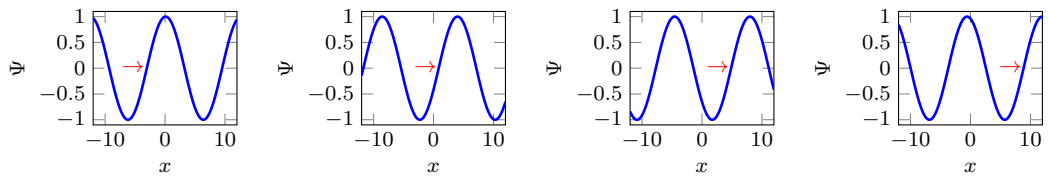
$$\omega^2 = c^2\beta^2 \quad \Longrightarrow \quad \omega = \pm c\beta \tag{2.15}$$

As this dispersion relation is linear (between ω and β) the system is said to be *non-dispersive*. That the dispersion relation features two possible plane-wave solutions for each ω is a consequence of the PDE being second-order in time.

A plane wave, while already occupying an infinite space, can also be said to be "travelling". In order to see this travelling behaviour, we rewrite the solution as:

$$\Psi = A\cos\left(\beta(w - ct)\right), \quad w = \hat{\boldsymbol{\beta}}\cdot\mathbf{x} \tag{2.16}$$

In the above form, Ψ can more readily be seen as a travelling wave that moves in the direction $\hat{\boldsymbol{\beta}}$ with a speed of propagation c, and is constant along planes where $w = \hat{\boldsymbol{\beta}}\cdot\mathbf{x}$ is constant. An example 1-D plane wave is illustrated in Fig. 2.2. Note that all such plane waves travel with the same speed, c, which is a result of the system being non-dispersive. Note also that the maximum amplitude (A) does not decay with time, which is ultimately a result of the system being lossless.



(a) $t = 0$ (b) $t = 4$ (c) $t = 8$ (d) $t = 12$

Figure 2.2: Snapshots of a 1-D plane-wave solution $\Psi = \cos(\beta(x - ct))$ with $c = 1$ and $\beta = 0.5$. The red arrow serves as a travelling frame of reference.

2.1.2 Sound sources in free space

Let us return to the subject of sound sources. For the purposes of this study, the sound source model employed will be based around a simple point source. This can be expressed in terms of the inhomogeneous (or nonhomogeneous) pressure wave equation:

$$p_{tt} - c^2 \Delta p = c^2 \rho \dot{q}_s(t) \delta(\mathbf{x} - \mathbf{x}_0) \tag{2.17}$$

where $\delta(\mathbf{x} - \mathbf{x}_0)$ is a Dirac delta[3] in 3-D (centered around $\mathbf{x}_0$), $q_s = q_s(t)$ is called the *source volume velocity* [194, 211] and $\dot{q}_s = \frac{\mathrm{d}}{\mathrm{d}t} q_s$. In terms of the acoustic velocity potential,[4] the equation reads:

$$\Psi_{tt} - c^2 \Delta \Psi = c^2 q_s(t) \delta(\mathbf{x} - \mathbf{x}_0) \tag{2.19}$$

In 3-D, the solution to the above equation for a distance $D = \|\mathbf{x} - \mathbf{x}_0\| > 0$ reads:

$$\Psi = \frac{q_s(t - D/c)}{4\pi D} \tag{2.20}$$

Thus, the acoustic field generated by the point source takes the form of a travelling spherical wave, expanding away from the source position $\mathbf{x}_0$. Note that the acoustic response at the source position goes to infinity as $D \to 0$. As such, the source is assumed to occupy a finite region of space, and typically, it is assumed that the size of the source is smaller than the smallest wavelength generated by the sound source [174].

The above solution can also be obtained from the closely-related problem:

$$\Psi'_{tt} - c^2 \Delta \Psi' = c^2 \delta(t) \delta(\mathbf{x} - \mathbf{x}_0) \tag{2.21}$$

The solution to this problem, Ψ', is known as a free-space Green's function [15, 68]. For the 3-D wave equation, this is:

$$\Psi' = \frac{\delta(t - D/c)}{4\pi D}, \quad D = |\mathbf{x} - \mathbf{x}_0| \tag{2.22}$$

Given a source term $q_s(t)$, the solution to (2.19) is then recovered by convolution (over time) with Ψ':

$$\Psi = (\Psi' * q_s)(t) = \int_0^t \Psi'(\mathbf{x}, \tau) q_s(t - \tau) \, \mathrm{d}\tau \tag{2.23}$$

One can also consider multiple sound sources with the equation:

$$\Psi_{tt} - c^2 \Delta \Psi = c^2 \sum_{m=1}^{M^{(\mathrm{s})}} q_s^{(m)}(t) \delta\left(\mathbf{x} - \mathbf{x}_0^{(m)}\right) \tag{2.24}$$

where $M^{(\mathrm{s})}$ is the number of sources, each with respective source volume velocity $q_s^{(m)}$ and source

[3] The Dirac delta (in 1-D) can be defined by [15]:

$$\int_{x_1}^{x_2} g(x) \delta(x) \, \mathrm{d}x = \begin{cases} g(0) & \text{if } 0 \in [x_1, x_2] \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

The Dirac delta in 3-D is then $\delta(\mathbf{x}) = \delta(x)\delta(y)\delta(z)$.

[4] Here we must assume that the time-derivative of the particle velocity is irrotational.

location $\mathbf{x}_0^{(m)}$. The solution follows from the previous example by linearity:

$$\Psi = \sum_{m=1}^{M^{(\mathrm{s})}} \frac{q_s^{(m)}(t - D_m/c)}{4\pi D_m}\,, \quad D_m = \|\mathbf{x} - \mathbf{x}_0^{(m)}\| \tag{2.25}$$

In the case of a sound source that is arbitrarily distributed over space, one has the equation:

$$\Psi_{tt} - c^2 \Delta \Psi = c^2 F \tag{2.26}$$

where now $F = F(\mathbf{x}, t)$ represents the more general source (forcing) term, and the solution reads [211, 175]:

$$\Psi = \int_{\mathbb{R}^3} \frac{F(\mathbf{y}, t - D/c)}{4\pi D}\,\mathrm{d}\mathbf{y}\,, \quad D = \|\mathbf{x} - \mathbf{y}\| \tag{2.27}$$

where $\mathrm{d}\mathbf{y}$ is a differential element over space, and the volumetric integral is over the entire space.

In terms of an energy analysis, the energy balance in the above case becomes:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(t) = \mathfrak{F}(t) \tag{2.28}$$

where $\mathcal{H}$ is as before ((2.4)), and $\mathfrak{F}$ is

$$\mathfrak{F}(t) = \rho \int_{\mathbb{R}^3} (\Psi_t)(F)\,\mathrm{d}\mathbf{x} \tag{2.29}$$

Thus the energy balance of the system takes into account the energy injected into the system by the source term.

2.1.3 Boundary conditions modelling locally reactive walls

For a model of room acoustics, we need to move from the unbounded (free-space) problem to a problem involving the wave equation on a closed domain. As acoustic waves travelling between the walls of a room impinge on a bounding wall they are expected to reflect, and subsequently travel in a direction pointing away from the wall. Upon reflection, a wave may lose energy, but it should not gain energy as this could lead to an unphysical growth in the system. In other words, the boundaries should be *passive* for a well-posed system.[5] For the purposes of this thesis, only boundary conditions that serve to model *locally reactive* walls will be presented. In essence, locally reactive boundary conditions mean that the wall does not act as a vibrating system of its own (such as, e.g., a wall being modelled as a plate or membrane). This is an assumption that is commonly employed in order to simplify room acoustics problems [140].

Passive boundary conditions for room acoustics modelling can be derived from an *energy analysis* of the wave equation on a closed domain (i.e., a domain that includes a boundary). Consider a closed domain $\overline{\Omega} = \Omega \cup \Gamma$, where Ω represents the interior (without the boundary) and Γ represents the boundary.[6] An energy analysis for the lossless wave equation, carried out over such a domain, then reads:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(t) = -\mathcal{B}\,, \quad \mathcal{H}(t) = \mathcal{V}(t) + \mathcal{T}(t) \tag{2.30}$$

[5] One example where waves could gain energy at walls is it the walls were moving and exerting a force on the acoustic field. In this case the walls would be *active*. Such scenarios will not be considered in this thesis.

[6] Typically, a room will also be a *bounded* domain, along with being closed, but to analyse portions of a room, it helps to consider unbounded, closed domains, such as half-spaces. As such, we do not restrict $\overline{\Omega}$ to being bounded.

where $\mathcal{V}$ and $\mathcal{T}$ are now:

$$\mathcal{V}(t) = \frac{1}{2\rho c^2} \int_{\overline{\Omega}} p^2 \, \mathrm{d}\mathbf{x} \,, \quad \mathcal{T}(t) = \frac{\rho}{2} \int_{\Omega} |\mathbf{v}|^2 \, \mathrm{d}\mathbf{x} \tag{2.31}$$

and $\mathcal{B}$ is a boundary term:[7]

$$\mathcal{B}(t) = \int_{\Gamma} p \, v_\perp \, \mathrm{d}\sigma = - \int_{\Gamma} (\rho \Psi_t)(\hat{\mathbf{n}} \cdot \nabla \Psi) \, \mathrm{d}\sigma \tag{2.32}$$

where, here, $\mathrm{d}\sigma$ is a differential surface element over the boundary with outward normal vector $\hat{\mathbf{n}}$ (i.e., normal to the wall), and $v_\perp$ is the outward normal component of $\mathbf{v}$:

$$v_\perp := \hat{\mathbf{n}} \cdot \mathbf{v} \,, \quad \mathbf{x} \in \Gamma \tag{2.33}$$

Thus, the total energy of the system will be non-increasing (a stable system) as long as $\mathcal{B}(t) \geq 0$. Three simple boundary conditions that produce a non-negative $\mathcal{B}$ are:

$$p = 0 \,, \qquad \mathbf{x} \in \Gamma \tag{2.34a}$$

$$v_\perp = 0 \,, \qquad \mathbf{x} \in \Gamma \tag{2.34b}$$

$$v_\perp = Y p \,, \quad \mathbf{x} \in \Gamma \tag{2.34c}$$

where in the third case $Y = Y(\mathbf{x})$ is the real-valued characteristic admittance of the boundary, which can also be written as:

$$Y = Y_0 \gamma \tag{2.35}$$

where $Y_0 = (\rho c)^{-1}$ is the characteristic admittance of air, and $\gamma = \gamma(\mathbf{x}) \in \mathbb{R}$ is a specific surface admittance. Note that $\gamma \geq 0$ is required for $\mathcal{B} \geq 0$.

The first condition, (2.34a), is known as a (homogeneous) Dirichlet condition and it models "soft" walls; these are not commonly used in room acoustics models.[8] The second condition, (2.34b), models rigid walls and it is the starting point of room acoustics models. Both of these Dirichlet-type boundaries are lossless, i.e., $\mathcal{B} \equiv 0$. The third condition, (2.34c), is a simple admittance (resistive) condition, and it models frequency-independent wall losses. In the case of $\gamma > 0$, energy is dissipated from the system through the walls and the term $\mathcal{B}$ represents a *power outflow*.

Clearly, $v_\perp = 0$ is a special case of the admittance condition (when $\gamma = 0$), but the soft-wall condition (2.34a) could also be seen as a special case of the admittance condition, (2.34c). We can also rewrite the admittance condition as an *impedance* condition:

$$p = Z v_\perp \,, \quad \mathbf{x} \in \Gamma \,, \quad Z = \xi Z_0 \tag{2.36}$$

where $Z = Z(\mathbf{x})$ is the real-valued characteristic impedance of the boundary, $\xi = 1/\gamma \geq 0$ is a specific surface impedance of the boundary surface and $Z_0 = 1/Y_0 = \rho c$ is the characteristic impedance of air. Then the boundary condition $p = 0$ is a special case of the above with $\xi = 0$.

It is also worth rewriting these boundary conditions in terms of the acoustic velocity potential. The

[7] In this case, the divergence theorem gives: $\int_{\overline{\Omega}} \Psi_t (\nabla \cdot \nabla \Psi) \, \mathrm{d}\mathbf{x} = - \int_{\Omega} (\nabla \Psi_t) \cdot (\nabla \Psi) \, \mathrm{d}\mathbf{x} + \int_{\Gamma} (\Psi_t)(\hat{\mathbf{n}} \cdot \nabla \Psi) \, \mathrm{d}\sigma$.

[8] A "soft wall" is simply the opposite of a rigid wall, but, as opposed to a rigid wall, the soft wall does not have a clear physical realisation in the context of room acoustics.

conditions (2.34) are then:

$$\Psi_t = 0\,, \qquad \mathbf{x} \in \Gamma \tag{2.37a}$$

$$\hat{\mathbf{n}} \cdot \nabla\Psi = 0\,, \qquad \mathbf{x} \in \Gamma \tag{2.37b}$$

$$-\hat{\mathbf{n}} \cdot \nabla\Psi = \frac{\gamma}{c}\Psi_t\,, \quad \mathbf{x} \in \Gamma \tag{2.37c}$$

In terms of the velocity potential, we can identify (2.37b) (the rigid wall condition) as a *homogeneous Neumann* condition. Although of limited physical significance to room acoustics, the "soft wall" condition (2.37a) could also be more simply expressed as $\Psi = 0$ (which implies (2.37a)), in which case it is a *homogeneous Dirichlet* condition in Ψ.

Modal solutions and room modes

Modal solutions to the wave equation are fundamental in the study of room acoustics. In many cases, the method of separation of variables can be used to assume a solution to the wave equation of the form $\Psi(\mathbf{x}, t) = \phi(\mathbf{x})\psi(t)$. Inserting this trial solution into the wave equation results in an ordinary differential equation (ODE) and a PDE, respectively:

$$\ddot{\psi}(t) = -c^2\beta^2\psi(t) \tag{2.38a}$$

$$\Delta\phi(\mathbf{x}) = -\beta^2\phi(\mathbf{x}) \tag{2.38b}$$

The first equation, (2.38a), is in the form of a simple harmonic oscillator with frequency $\omega = c\beta$. The second equation, (2.38b), is an eigenvalue problem for the Laplace operator, also known as the *Helmholtz equation*, which plays a central role in the *frequency-domain* analysis of room acoustics.[9]

In the free-space case, one obtains a continuous spectrum of eigenvalues from the Helmholtz equation, pertaining to plane-wave solutions. In the case of a closed, bounded domain, the eigenvalues appear as a discrete, yet infinite set. Provided that the solution is indeed separable in that case (such as, e.g., under rigid wall conditions), one can write a more general solution as a weighted sum of *modes*:

$$\Psi(\mathbf{x}, t) = \sum_m \phi_m(\mathbf{x})\left(a_m \cos(\beta_m c t) + b_m \sin(\beta_m c t)\right) \tag{2.39}$$

where m is a modal index, and a_m, b_m are modal coefficients that can be determined from a set of initial conditions for the problem. In this context, each frequency $f_m = \frac{c\beta_m}{2\pi}$ (in Hz) is known as a modal frequency, and $\phi_m(\mathbf{x})$ is its associated modal shape.

In order to make use of this modal solution to the wave equation, one has to first solve the Helmholtz eigenvalue problem, which is generally not solvable by simple means, apart from some special cases. One well-known such special case, particularly relevant to room acoustics, is that of a box-shaped domain (an simplified shoebox hall) with Neumann conditions (rigid walls). Considering a box of size $L_x \times L_y \times L_z$ (dimensions in metres), under rigid wall conditions the mode-shapes take the form:

$$\phi_{\boldsymbol{m}}(\mathbf{x}) = \cos\left(\frac{m_x \pi x}{L_x}\right)\cos\left(\frac{m_y \pi y}{L_y}\right)\cos\left(\frac{m_z \pi z}{L_z}\right) \tag{2.40}$$

where $\boldsymbol{m}$ is a vector modal index, and $\boldsymbol{m} = (m_x, m_y, m_z) \in \mathbb{Z}^{3+}$.[10] These modal shapes oscillate at

[9] As opposed to time-domain analysis, which is governed largely by the second-order wave equation.

[10] $\mathbb{Z}^{3+}$ is defined as $\mathbb{Z}^{3+} = \{(x, y, z) \in \mathbb{Z}^3, x \geq 0\}$.

the associated modal frequencies:

$$f_{\boldsymbol{m}} = \frac{c}{2}\sqrt{\left(\frac{m_x}{L_x}\right)^2 + \left(\frac{m_y}{L_y}\right)^2 + \left(\frac{m_z}{L_z}\right)^2} \tag{2.41}$$

From this expression for the modal frequencies, we can start to discuss the *modal density* of the room. Clearly, the modal frequencies appear in an inharmonic series that is sparse in low frequencies, and becomes increasingly more dense in high frequencies. In fact, it can be worked out that there will be, *on average*, $\frac{8\pi L_x L_y L_z}{6c^3} f^3$ modal frequencies below some f in Hz [174]. Thus, the modal density under some frequency f in a box-shaped room is proportional to f^3.

In the case of a non-rectangular room, it is not as simple to express the modal frequencies in an analytic form. Nevertheless, it can still be said that the modal density of any room (in 3-D), regardless of shape, has (on average) a modal density proportional to f^3 under some frequency f [174]. These number of these modal frequencies is directly related to the number of *degrees of freedom* of the underlying PDE problem. As such, we can say that this 3-D problem has a complexity that scales as $\mathcal{O}(f^3)$. This has natural consequences for any numerical algorithm that aims to solve to for the acoustic field Ψ (to be discussed more in Section 2.2.2.

Reflection coefficients for frequency-independent boundaries

At this point we should discuss in more detail how energy is dissipated in the simple admittance boundary model, and how acoustical waves reflect off a surface. First, consider as a domain the half-space $\mathbb{Z}^{3+}$ with a boundary along $x = 0$. Next we consider a sum of two complex-valued plane waves:

$$\Psi = A\mathrm{e}^{\mathrm{j}(\omega t + (\beta_x, \beta_y, \beta_z)\cdot \mathbf{x})} + B\mathrm{e}^{\mathrm{j}(\omega t + (-\beta_x, \beta_y, \beta_z)\cdot \mathbf{x})} \tag{2.42}$$

There are two assumptions built into this trial solution. The first assumption is that one wave is travelling towards the wall (the *incident wave*) and the other is travelling away from the wall (the *reflected wave*), and the other assumption is that the wave vectors are related by a simple change of sign in the x-component (the direction normal to the boundary). This means that the angle of incidence will be equal to the angle of reflection, so the boundary acts like a mirror does to light rays.[11]

It remains to determine the relationship between A and B and the characteristic admittance of the wall, γ. To derive this relationship, we impose the admittance boundary condition (2.37c) at $x = 0$ with $\gamma \geq 0$ left unspecified. Inserting this trial solution into (2.37c) and noting that the normal outward vector on the boundary is $\hat{\mathbf{n}} = -\hat{\mathbf{e}}_x$ we get the following relation at $x = 0$:

$$(A - B)\cos\theta = \gamma(A + B) \tag{2.43}$$

where $\theta \in [0, \pi/2)$ is the angle of incidence. This reduces to:

$$R = \frac{B}{A} = \frac{\cos\theta - \gamma}{\cos\theta + \gamma} \tag{2.44}$$

where $R = R(\theta)$ is an angle-dependent *reflection coefficient*. Note that $|R| \leq 1$, which means that the reflected wave cannot be larger in amplitude than the incident wave. This is a consequence of passivity, as one could easily see that $\gamma < 0$ also leads to $|R| > 1$, and indeed $\gamma < 0$ violates the energy stability condition.

[11] This is of course and underlying principle in the geometrical acoustics model, to be discussed in Section 2.2.1.

Some special cases of interest are: when $\gamma = 0$, $R \equiv 1$ (rigid wall, behaviour independent of angle); when $\gamma = \cos\theta$, $R = 0$ and the wall is completely absorbing for angle θ; and when $\gamma \to \infty$, $R \to -1$ (a limiting case) and the reflected wave is *phase-reversed* with respect to the incident wave. The special case of $\gamma = 1$, is known as a first-order Engquist-Majda absorbing boundary condition [69], and it is fully absorbing ($R = 0$) at normal incidence ($\theta = 0$). It is also worth mentioning that this reflection coefficient can be written in terms of the impedance $\xi = 1/\gamma$:

$$R(\theta) = \frac{\xi - \cos\theta}{\xi + \cos\theta} \tag{2.45}$$

In this case, the rigid wall is a limiting case, i.e., $\xi \to \infty$, $R \to 1$.

Positive real functions for frequency-dependent conditions

So far we have considered the real-valued immittances, and as such, reflection coefficients that are independent of frequency. However, aside from walls that are very rigid, this only provides a superficial description of walls in the real-world. For a more detailed description of walls, one must consider walls that have a *complex* immittance (admittance or impedance), leading to a frequency-dependent behaviour, with frequency-dependent losses and *storage* of energy. Of course, this complicates the formulation of a stable boundary value problem. Fortunately, stable frequency-dependent boundaries can be constructed with the help of *positive real* functions [306]. A positive-real function $F(s)$ is complex-valued function in the variable $s = \sigma + \mathrm{j}\omega$ (with $\sigma \in \mathbb{R}$) that satisfies the following conditions:

$$\begin{aligned} \Re\{F(s)\} &> 0 \quad \text{if } \sigma > 0 \\ \Im\{F(s)\} &= 0 \quad \text{if } \omega = 0 \end{aligned}$$

Frequency-dependent admittance boundary condition

A general locally-reactive admittance boundary condition can be written in the Laplace domain as:

$$v_\perp(\mathbf{x}, s) = Y(\mathbf{x}, s) p(\mathbf{x}, s)\,, \quad \mathbf{x} \in \Gamma \tag{2.46}$$

A simple example of a positive-real function is the following admittance boundary condition:

$$Y(\mathbf{x}, s) = Y_0(As + B + Cs^{-1}) \tag{2.47}$$

Here, $A = A(\mathbf{x}) \geq 0$, $B = B(\mathbf{x}) \geq 0$, and $C = C(\mathbf{x}) \geq 0$ are quantities that represent acoustic analogues of an electrical inductance (A, in units s^{-1}), resistance (B, dimensionless), and inverse-capacitance (C, in units s) in a parallel "RLC" circuit. In this electrical circuit analogy, the pressure is taken to be the analogue of voltage, and the normal velocity is the analogue of current. The coefficients A, B, C can vary over the boundary but must remain non-negative.

In the time domain, this boundary condition can be written as:

$$v_\perp = Y_0(Ap_t + Bp + Cg)\,, \quad \partial_t g = p\,, \quad \mathbf{x} \in \Gamma \tag{2.48}$$

where $g = g(\mathbf{x}, t)$ is an extra variable on the boundary, linked to the storage of energy over the boundary. For $A = C = 0$, B is simply γ from the frequency-independent admittance condition (2.34c), so this condition can be seen as a generalisation of (2.34c).

Under this boundary condition, the energy balance of the lossless wave equation system is:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_{\mathrm{i}}(t) = -\mathcal{B}(t)\,, \quad \mathcal{H}_{\mathrm{i}}(t) = \mathcal{V}(t) + \mathcal{T}(t) \tag{2.49}$$

where $\mathcal{H}_{\mathrm{i}}$ represents the stored energy on the interior, with $\mathcal{V}$ and $\mathcal{T}$ are as before (in (2.31)), and $\mathcal{B}$ is now:

$$\mathcal{B}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_{\mathrm{b}}(t) + \mathcal{Q}(t) \tag{2.50}$$

where $\mathcal{H}_{\mathrm{b}}$ is the stored energy over the boundary, and $\mathcal{Q}$ is the power dissipated over the boundary. These terms are written as:

$$\mathcal{Q}(t) = Y_0 \int_\Gamma B p^2 \mathrm{d}\sigma \geq 0\,, \quad \mathcal{H}_{\mathrm{b}}(t) = \frac{Y_0}{2}\int_\Gamma A p^2 + C g^2 \mathrm{d}\sigma \geq 0 \tag{2.51}$$

where the chain rule was applied in order to obtain $\mathcal{H}_{\mathrm{b}}$. The energy balance can then be rewritten as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{H}_{\mathrm{i}} + \mathcal{H}_{\mathrm{b}}\right) = -\mathcal{Q} \tag{2.52}$$

and the total stored energy $\mathcal{H} = \mathcal{H}_{\mathrm{i}} + \mathcal{H}_{\mathrm{b}}$ is thus non-increasing over time, and we can say that the system is stable, or passive. As a consequence of this passivity, it must be the case that $R = R(\omega, \theta)$—now complex and a function of both angle and frequencies—is bounded as $|R| \leq 1$ for $\omega \in [0, \pi]$ and $\Re\{s\} = 0$.

For an admittance-type condition of the form (2.46), the frequency-dependent reflection coefficient at normal incidence ($\theta = 0$) is expressed as:

$$R(\omega) = \frac{Y_0 - Y(\omega)}{Y_0 + Y(\omega)} \tag{2.53}$$

A plot of the magnitude and phase of $R(\omega)$ is shown in Fig. 2.3, for various choices of non-zero A, B, C; also shown in that figure are three examples of the frequency-independent admittance condition with parameter $\gamma \in \mathbb{R}$.

In each of the frequency-dependent cases in Fig. 2.3, it can be seen that there is a "dip" in the magnitude ($|R|$) when the phase ($\angle R$) crosses zero—this is due to a resonance at approximately $f = \frac{1}{2\pi}\sqrt{\frac{C}{A}}$ Hz in the equivalent circuit network. If the frequency-dependence of a wall model calls for more than one such resonance, a more general immittance circuit can be employed; this will be treated shortly. The frequency-independent examples (with $\gamma \in \mathbb{R}$) in Fig. 2.3 have no such resonances and indeed, in these cases R has no dependence on frequency. For the case $\gamma = 5$, R is negative and thus the reflection at the boundary is phase-reversing (this happens when $\gamma > 1$), while for $\gamma < 1$ the reflection is phase-preserving. The homogeneous Neumann condition $\gamma = 0$ has $R = 1$ (phase-preserving). It can also be observed that with this admittance condition, $R \to -1$ ($|R| \to 1$ and $\angle R \to \pi$) as $\omega \to 0$ when there is a frequency-dependence (A and/or C non-zero).

Frequency-dependent impedance boundary condition

A similar boundary condition can be written in terms of an impedance, generalising the simple impedance condition (2.36). In general, a locally-reactive impedance boundary condition can be written in the Laplace domain as:

$$p(\mathbf{x}, s) = Z(\mathbf{x}, s) v_\perp(\mathbf{x}, s)\,, \quad \mathbf{x} \in \Gamma \tag{2.54}$$

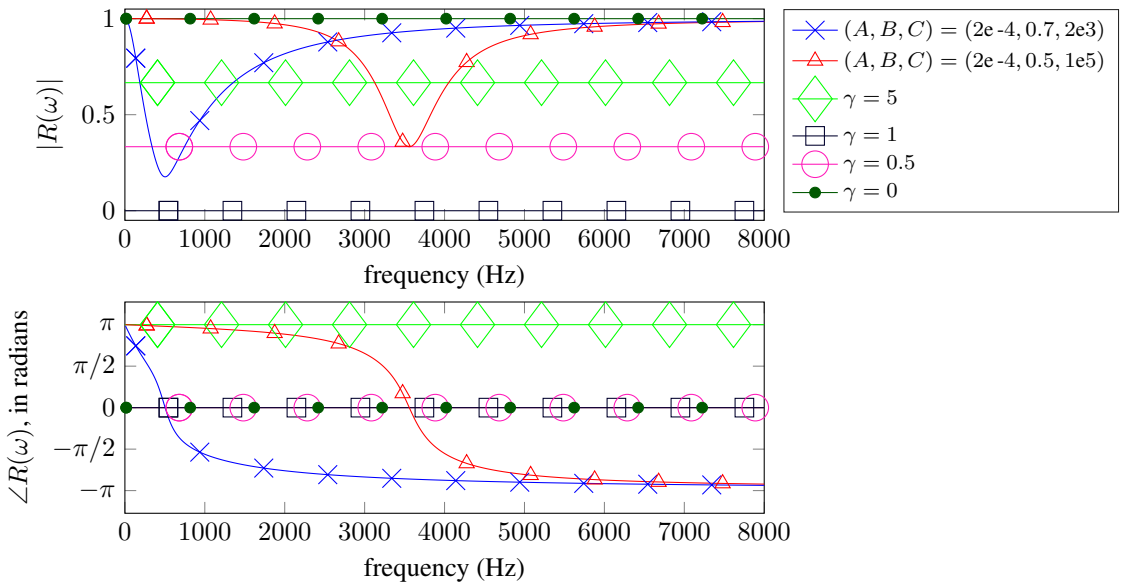


Figure 2.3: Reflection coefficients for admittance boundary conditions at normal incidence.

One possible generalisation of (2.36), in the electrical analogy, is a *series* RLC circuit. This is expressed in the Laplace domain as:

$$Z(\mathbf{x}, s) = Z_0(Ds + E + Fs^{-1}) \tag{2.55}$$

where $Z_0 = \rho c$ is the characteristic impedance of air, and $D, E, F \geq 0$ are analogous to series inductance (D, in units s^{-1}), resistance (E, dimensionless), and inverse-capacitance (F, in s), and all of these coefficients can spatially vary over the boundary ($D = D(\mathbf{x})$, etc.). In the time domain, this boundary condition can be written as:

$$p = Z_0 \left(D\partial_t v_\perp + Ev_\perp + Fg\right) , \quad \partial_t g = v_\perp , \quad \mathbf{x} \in \Gamma \tag{2.56}$$

and it is straightforward to arrive at an energy balance of the form (2.49) and (2.50), where now

$$\mathcal{Q}(t) = Z_0 \int_\Gamma E v_\perp^2 \mathrm{d}\sigma \geq 0 , \quad \mathcal{H}_\mathrm{b}(t) = \frac{Z_0}{2} \int_\Gamma D v_\perp^2 + F g^2 \mathrm{d}\sigma \geq 0 \tag{2.57}$$

As in the frequency-dependent admittance case, we have here that the total energy of the system is non-increasing. For the impedance condition, a frequency-dependent reflection coefficient at normal incidence is written as:

$$R(\omega) = \frac{Z(\omega) - Z_0}{Z(\omega) + Z_0} \tag{2.58}$$

This reflection coefficient is plotted in Fig. 2.4 for various choices of D, E, F, and also for various (frequency-independent) choices of $\xi \in \mathbb{R}$. The numerical values are chosen to be the same respective values as A, B, C in Fig. 2.3, and it can be seen that $|R|$ behaves exactly as in the admittance case, but $\angle R$ has changed by π rad. In particular, it can be noted that when $D > 0$ and/or $F > 0$, $R \to 0$ as $\omega \to 0$. Thus, as opposed to the frequency-dependent admittance conditions which approach the "soft wall" at DC, these conditions approach a rigid wall at DC.

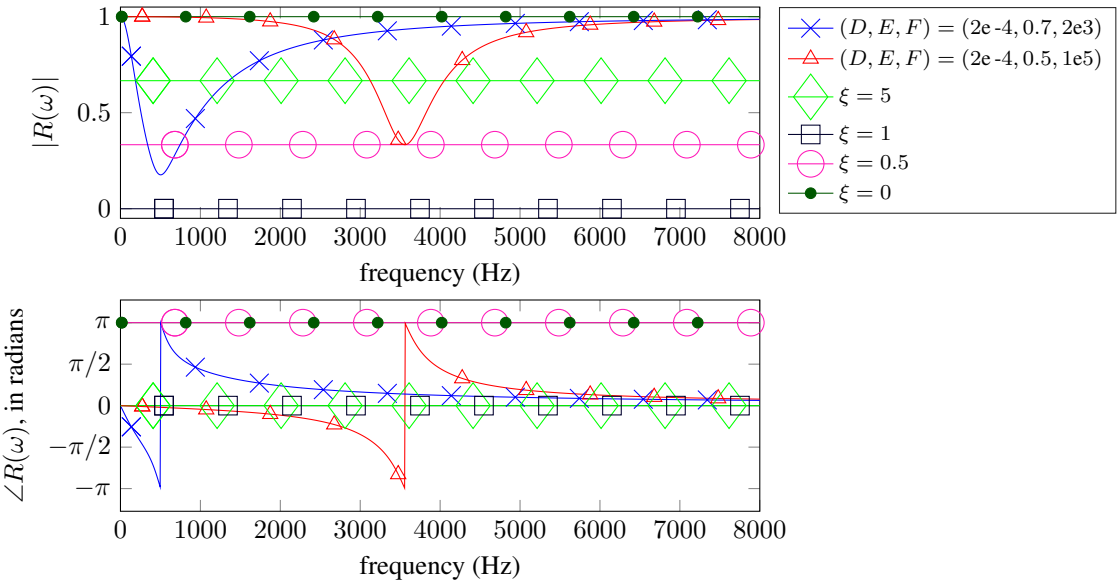


Figure 2.4: Reflection coefficients for impedance boundary conditions at normal incidence

A general impedance boundary condition

As explained earlier, each equivalent RLC circuit network in the above boundary condition is capable of producing at most one resonance. When the wall to be modelled has a more complex behaviour (though still locally reactive), a more general boundary condition is required.

Keeping with positive-real functions, one possible boundary condition that can lead to such general frequency dependences has recently been proposed in [30], and it is also based on equivalent circuit representations. This general impedance model is made up of a parallel combination of $M^{(\mathrm{b})} \geq 1$ series RLC circuit branches, where the analogues of voltage and pressure are again pressure and normal velocity. The circuit network is illustrated in Fig. 2.5.[12]

The total admittance of this network is:

$$Y(\mathbf{x}, s) = \sum_{m=1}^{M^{(\mathrm{b})}} \frac{s}{L^{(m)}(\mathbf{x})s^2 + R^{(m)}(\mathbf{x})s + \frac{1}{C^{(m)}(\mathbf{x})}} \tag{2.59}$$

where $R^{(m)}, L^{(m)}, C^{(m)}$ are the resistance, inductance, and capacitance of the mth of $M^{(\mathrm{b})}$ branches, as illustrated in Fig. 2.5, and they are permitted to vary over the boundary, yet must remain non-negative. The condition $C^{(m)} > 0$ is imposed such that zero admittance (rigid walls) is enforced at DC [30].

The boundary condition associated with this admittance is expressed by (2.54) (in the Laplace domain). In order to express this boundary condition in the time domain, the normal velocity boundary component is split into M branch components (as illustrated in the figure):

$$v_\perp = \sum_{m=1}^{M^{(\mathrm{b})}} v_\perp^{(m)} \tag{2.60}$$

The pressure (voltage) at the boundary, which is constant across each branch, is related to the branch

[12]The convention of using the symbols $R^{(m)}, L^{(m)}, C^{(m)}$ in this section follows [30], whereas previously we used the symbols D, E, F (or A, B, C) to describe simpler frequency-dependent immittance boundary conditions, which follows the notation used in [26]. In the case of a single branch ($M = 1$), we have that $(L^{(1)}, R^{(1)}, 1/C^{(1)}) = Z_0(D, E, F)$.

normal velocity (branch current) by $M^{(\mathrm{b})}$ parallel impedance relations:

$$p = L^{(m)}\partial_t v_\perp^{(m)} + R^{(m)} v_\perp^{(m)} + \frac{1}{C^{(m)}} g^{(m)}\,, \quad \partial_t g^{(m)} = v_\perp^{(m)}\,, \quad m = 1, \ldots, M^{(\mathrm{b})}\,, \quad \mathbf{x} \in \Gamma \tag{2.61}$$

A reflection coefficient $R(\omega)$ (not to be confused with a branch-resistance $R^{(m)}$) can be expressed by (2.58). The energy balance is of the form (2.49), where now (2.50) has:

$$\mathcal{Q}(t) = \sum_{m=1}^{M^{(\mathrm{b})}} \int_\Gamma R^{(m)} \left(v_\perp^{(m)}\right)^2 \mathrm{d}\sigma \geq 0\,, \quad \mathcal{H}_\mathrm{b}(t) = \sum_{m=1}^{M^{(\mathrm{b})}} \int_\Gamma \frac{L^{(m)}}{2} \left(v_\perp^{(m)}\right)^2 + \frac{1}{2C^{(m)}} \left(g^{(m)}\right)^2 \mathrm{d}\sigma \geq 0 \tag{2.62}$$

Reflection coefficients and admittances from an example three-branch circuit network are plotted in Fig. 2.6. It can be seen that the combined reflection coefficient has three resonances, each of which is close to a resonance of the individual branches.

2.1.4 Sound absorption in air

So far we have only considered lossless propagation of sound in air, but we also need a description of sound absorption (attenuation) in air for a more complete model of room acoustics. Let us then return to the case of free-space wave propagation. The classic model for sound attenuation in air is Stokes' equation [258], which can be written as:

$$\Psi_{tt} = c^2 \Delta\Psi + c\eta\Delta\Psi_t \tag{2.63}$$

where η is a constant that has dimensions of metres, and it is related to viscous and/or thermal effects in air. This constant is generally a small value, and in this work it will be assumed that $\eta \ll c/\omega$ for the temporal frequencies of interest.[13] Note that this equation reduces to the lossless wave equation when $\eta \to 0$. It will also be convenient to define the time τ_η as:

$$\tau_\eta = \eta/c \tag{2.64}$$

This is also known as a *relaxation time*, and it is related to an interpretation of viscothermal losses as a molecular relaxation effect [174].

In order to take a first look at the behaviour of this system, it is useful to conduct an energy analysis in the free-space case. Multiplying both sides of (2.63) by $(\rho/c^2)\partial_t\Psi$ and integrating over the full-space, an energy balance can be written as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(t) = -\mathcal{Q}(t) \tag{2.65}$$

where $\mathcal{H}$ is:

$$\mathcal{H} = \frac{1}{2\rho c^2}\int_{\mathbb{R}^3} p^2\,\mathrm{d}\mathbf{x} + \frac{\rho}{2}\int_{\mathbb{R}^3} |\mathbf{v}|^2\,\mathrm{d}\mathbf{x} \tag{2.66}$$

and $\mathcal{Q}$ is a power outflow:

$$\mathcal{Q}(t) = -\frac{\rho}{c^2}\int_{\mathbb{R}^3} (\Psi_t)(c\eta\Delta\Psi_t)\,\mathrm{d}\mathbf{x} = \rho\tau_\eta \int_{\mathbb{R}^3} |\nabla\Psi_t|^2\,\mathrm{d}\mathbf{x} \geq 0 \tag{2.67}$$

where the above equality follows from the application of the divergence theorem. Thus, the additional

[13] For example, with $c = 340$ m/s and $\omega = 40000\pi$ rad/s (generally taken to be the limit of human hearing), we assume that $\eta \ll$ 2.7e-3 m, whereas η should be on the order 10^{-6} m–10^{-7} m for indoor conditions, as will be seen shortly.

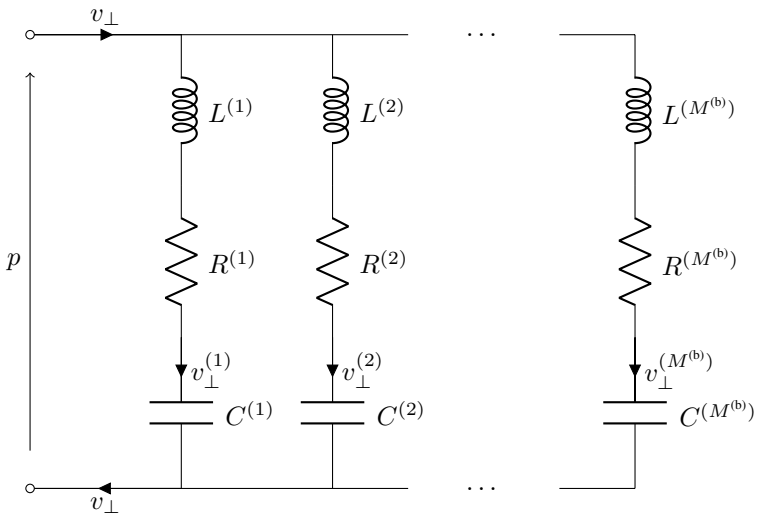


Figure 2.5: Parallel-series RLC circuit for general impedance boundary condition

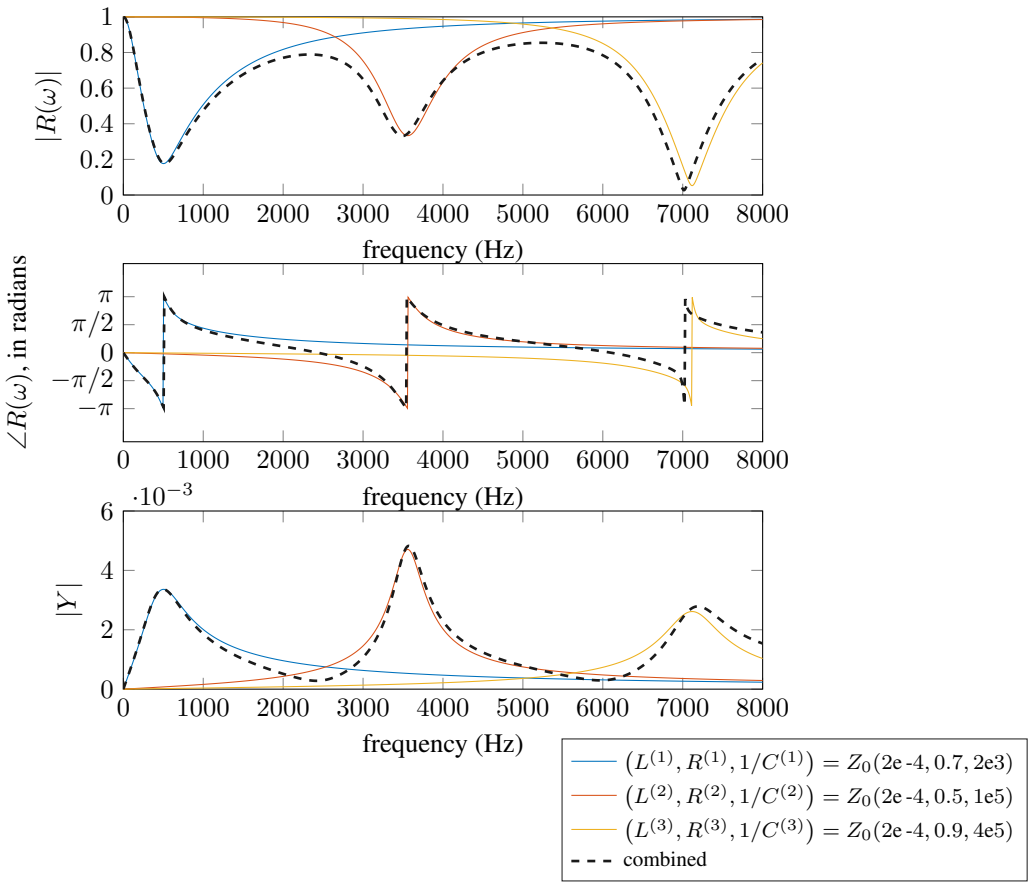


Figure 2.6: Reflection coefficients and surface admittances for three-branch general impedance boundary condition at normal incidence. The reflection coefficients and surface admittances are plotted for individual branches and for the combined network.

term in (2.63) (over that of the wave equation) leads to energy dissipation. In a sense, energy is still conserved, but in this case energy conservation is expressed using the accumulation of power dissipated:

$$\mathcal{H}(t) = \mathcal{H}(0) - \int_0^t \mathcal{Q}(\tau)\,\mathrm{d}\tau \tag{2.68}$$

where $\mathcal{H}(0)$ is the energy of the system at time zero.

It is important to note, however, that while this energy balance reduces to the physical energy of the lossless system when $\eta \to 0$, it is not, strictly speaking, the correct physical energy of the system. Here, the energy balance is in terms of energetic (energy-like), positive semi-definite quantities (similar to Lyapunov functions [161]) that are used to bound the behaviour of the system.[14]

Frequency-dependence of viscothermal losses

In order to get a more detailed view of the behaviour of Stokes' equation, let us consider a complex wave-like trial solution:

$$\Psi = A\mathrm{e}^{st - \mathrm{j}\boldsymbol{\beta}\cdot\mathbf{x}} \tag{2.69}$$

where $s = \sigma + \mathrm{j}\omega$ and $\sigma \in \mathbb{R}$ represents a decay factor when $\sigma < 0$ and a growth factor when $\sigma > 0$. When $\sigma = 0$, this trial solution satisfies the lossless wave equation (2.7), and its real part is equivalent to (2.16).[15] Inserting this trial solution into (2.63) leads to a characteristic equation in s:

$$s^2 + c\eta\beta^2 s + c^2\beta^2 = 0 \tag{2.70}$$

which has the following complex roots for $\omega \ll c/\eta$ (see [47]):

$$s \approx -\frac{c\eta\beta^2}{2} \pm \mathrm{j}\omega \tag{2.71}$$

Also, for $\omega \ll c/\eta$, we find that the system is dispersive but that the effect is negligible, and thus we can take $\omega \approx \pm c\beta$ [47]. Then we can take the solution to be (replacing '$\approx$' by '$=$'):

$$\Psi = A\mathrm{e}^{-\alpha t}\mathrm{e}^{\mathrm{j}(\omega t - \boldsymbol{\beta}\cdot\mathbf{x})} \tag{2.72}$$

where

$$\alpha = \frac{\omega^2\tau_\eta}{2} \tag{2.73}$$

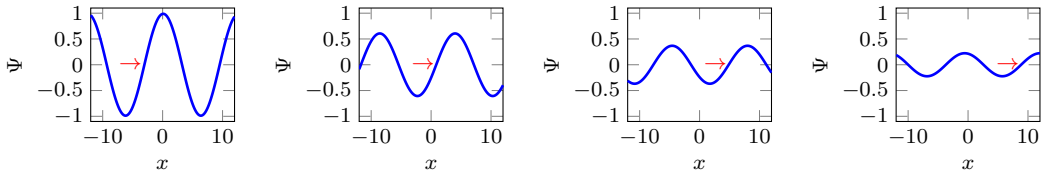
The coefficient α can be called a *temporal absorption coefficient*, and by (2.72) the amplitude $A\mathrm{e}^{-\alpha t}$ decreases by a factor of $\mathrm{e}^{-\alpha}$ every second (still frequency-dependent). Since the decay rate is proportional to the square of the frequency, the absorption is said to follow a *power law*. The absorption coefficient is also sometimes expressed in units of Nepers, in which case the decay rate is α Nepers/s. An example 1-D solution with viscothermal dissipation (using unrealistic coefficients, for illustrative purposes) is shown in Fig. 2.7, analogous to the lossless example featured in Fig. 2.1.

One can also express the decay rate as a function of space through the wave speed c:

$$\alpha' = \alpha/c \quad \Longrightarrow \quad \Psi = A\mathrm{e}^{-\alpha' ct}\mathrm{e}^{\mathrm{j}(\omega t - \boldsymbol{\beta}\cdot\mathbf{x})} \tag{2.74}$$

[14] A more correct energy follows from the more general linearised Navier-Stokes system, but for small η ($\omega \ll c/\eta$), the behaviour of this second-order equation is virtually identical to that of the more general first-order system. See [99] (published work of the current author) for more detail.

[15] While a real-valued trial solution could be employed, this complex-valued wave is preferred for its ease of mathematical manipulation.



(a) $t = 0$ (b) $t = 4$ (c) $t = 8$ (d) $t = 12$

Figure 2.7: Snapshots of a 1-D plane-wave solution $\Psi = \mathrm{e}^{-\alpha t}\cos(\beta(x - ct))$ with $c = 1$, $\beta = 0.5$, and $\alpha = 1/8$. The red arrow serves as a travelling frame of reference.

and then it can be said that any sound will decay by a (frequency-dependent) factor of $\mathrm{e}^{-\alpha'}$ for every metre of distance travelled, or α' Nepers/m.

Specific values for η

With the absorption behaviour of Stokes' equation sufficiently described, it remains to specify an appropriate value for η. It should be mentioned that while we will continue to make use of Stokes' equation, η will not necessarily represent the same quantities as originally considered by Stokes' [258].

As originally formulated by Stokes [258], the value of η was taken to be a combination of viscosity terms and ρ, c for the fluid of interest:

$$\eta = l_v := \frac{1}{\rho c}\left(\frac{4}{3}\mu + \mu_\mathrm{b}\right) \tag{2.75}$$

where μ and μ_B represent dynamic (shear) and bulk viscosities, respectively.[16] For air at standard conditions, the dynamic viscosity is known and generally taken to be a constant [174, 194, 127]. Meanwhile, the bulk viscosity μ_B remains unknown, as it is difficult to measure [65], but it is estimated to be on the order of the dynamic viscosity [194, 65]. Stokes assumed that the effect of bulk viscosity was negligible and it was set to zero in his original formulation [258]. With η defined as above, Stokes' equation can also be called the "viscous wave equation". One could also derive a viscous wave equation in terms of the pressure [127]:

$$p_{tt} = c^2\Delta p + c l_v \Delta p_t \tag{2.76}$$

Another effect that can induce a decay of sound is thermal conduction in air. For thermal effects, we define the characteristic length:

$$l_h := \frac{\kappa'}{C_\mathrm{p}\rho c} \tag{2.77}$$

where κ' is the coefficient of thermal conductivity (not to be confused with κ the bulk modulus), and C_p is the heat coefficient at constant pressure for a constant unit mass. These constants are also known for air [174, 194, 127]. Then we have for η:

$$\eta = l_v + (\gamma' - 1)l_h = \frac{1}{\rho c}\left(\frac{4}{3}\mu + \mu_\mathrm{b} + (\gamma' - 1)\frac{\kappa'}{C_\mathrm{p}}\right) \tag{2.78}$$

where γ' is the ratio of specific heats. In this case, α' is known as the *modified classical absorption coefficient* [194] (simply *classical* when the bulk viscosity is neglected). Taking the known values of $\mu, \kappa', C_\mathrm{p}$ for air into account, and with an estimated bulk viscosity on the order of the dynamic viscosity,

[16] The dynamic viscosity is associated to losses due to shear stress, whereas the bulk viscosity is typically associated to losses due to translation and rotational molecular relaxation effects [194].

the value of η is on the order of 10^{-7} m [174]. With the inclusion of heat conduction, Stokes' equation is only changed by the value of η, but at this point it is more appropriate to refer to (2.63) as the "viscothermal wave equation".

In this case, the pressure is related to the acoustic velocity potential through the following equation, which can be derived from a linearised Navier-Stokes system [174, 99]:

$$p = \rho\Psi_t - \rho c l_v \Delta\Psi \tag{2.79}$$

Viscothermal effects can also be seen as components of molecular relaxation effects, namely the translational modes of molecules [174]. Other relaxation effects include those due to vibrational and rotational modes of molecules. For the frequencies of acoustical interest, the rotational modes of molecules in air induce a power-law absorption on the order of the classical absorption [194] and the two are often lumped together [117].

As for the vibrational relaxation effects in air, mainly due to oxygen and nitrogen molecules, the associated absorption behaves as follows:

$$\frac{\alpha_{\text{vib}}}{f^2} \propto \frac{f_{\text{r}}}{f^2 + f_{\text{r}}^2} \tag{2.80}$$

where f_{r} is a *relaxation frequency*, or the inverse of a *relaxation time*: the time it takes for relaxation process to reach an equilibrium state. According to the above relation, the absorption due to vibrational relaxation processes behaves as a power-law for $f \ll f_{\text{r}}$ and as a constant for $f \gg f_{\text{r}}$. Thus, if the latter behaviour is dominant in the frequencies of interest, then the viscothermal wave equation is only a first-order approximation to more general relaxation effects.

Typically, the sound absorption coefficient is expressed as a sum of the aforementioned effects [117]:

$$\alpha_{\text{tot}} = \alpha' + \alpha_{\text{vib,O}} + \alpha_{\text{vib,N}} \tag{2.81}$$

where α' now encompasses the power-law rotational and translation relaxation effects (or from another point of view, the viscothermal effects). These absorption coefficients can be calculated as a function of frequency, temperature, ambient pressure, and relative humidity according to [117] or [16].

Fixing the ambient pressure to standard conditions, the absorption coefficients α_{tot} is a function of temperature, relative humidity, and frequency. Considering four distinct pairs of temperature and humidity, the absorption coefficient (2.81) is plotted as a function of frequency in Fig. 2.8, in terms of decibels (dB) per metre attenuation. It can be seen from the figure that the classical absorption dominates in ultrasonic frequencies ($\gg$ 20 kHz), whereas the total absorption is dominated by the vibrational relaxation effects of oxygen and nitrogen in audible and subsonic frequencies.

It can be seen in Fig. 2.8 that for 40%–60% relative humidity and 15–20 degrees Celsius, which are typical indoor conditions, the sound attenuation in 2–20 kHz is dominated by the power-law absorption due to oxygen vibrational modes. Since the bulk viscosity is ultimately unknown, we can attribute such relaxation effects to the value of μ_{b}, and thus into η, after [194, 174]. Taking a first-order approximation to the absorption coefficient $\alpha_{\text{vib,O}}$ from [117], we can then find a suitable value for η. A full expression for this temperature and humidity-dependent first-order approximation to $\alpha_{\text{vib,O}}$ is left out for brevity (this can be obtained from [117]), but as an example, at 20°C and 30% relative humidity, the value for η would be approximately 2×10^{-6} m, and this value would also work for 15°C and 40% relative humid-

ity.[17] Extracting a value for η from a first-order approximation of $\alpha_{\text{vib,O}}$ is a valid approximation (i.e., agrees with [117, 16]) for frequencies in the 2–20 kHz range for standard indoor conditions (e.g., 40–60%) (it can be assumed that wall absorption will dominate under 2 kHz). Outside of these standard indoor conditions, a more detailed wave equation taking into account full relaxation effects may be required for accurate modelling of sound absorption in air, but such models are beyond the scope of this work.

Boundary conditions with viscothermal losses

Let us now consider extensions of the frequency-dependent boundary models to the viscothermal case. We should note that viscothermal effects are accompanied by boundary layer effects [174], but these are generally only significant in tubes and ducts. It can safely be assumed that boundary layer effects are too small to be significant in the context of room acoustics [140], so we will ignore them for the purposes of this study.

For the viscothermal case, boundary conditions can be obtained in a simple manner: by redefining the normal component of the velocity field to be:

$$v_\perp = -\hat{\mathbf{n}} \cdot (\nabla\Psi + \tau_\eta \nabla\Psi_t) \tag{2.82}$$

This association leads to an energy-like balance of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{H}_\mathrm{i}(t) + \mathcal{H}_\mathrm{b}(t)\right) = -\left(\mathcal{Q}_\mathrm{i}(t) + \mathcal{Q}_\mathrm{b}(t)\right), \tag{2.83}$$

where

$$\mathcal{H}_\mathrm{i}(t) = \frac{\rho}{2c^2}\int_{\overline{\Omega}} \Psi_t^2 \,\mathrm{d}\mathbf{x} + \frac{\rho}{2}\int_\Omega |\nabla\Psi|^2 \,\mathrm{d}\mathbf{x}\,, \qquad \mathcal{Q}_\mathrm{i}(t) = \frac{\rho}{c^2}\int_\Omega |\nabla\Psi_t|^2 \,\mathrm{d}\mathbf{x} \tag{2.84}$$

and $\mathcal{H}_\mathrm{b}$, $\mathcal{Q}_\mathrm{b}$ are given by either (2.51) (for the admittance boundary condition) or (2.57) (for the impedance boundary condition). In the energetic quantities above, $v_\perp$ is of the modified form (2.82). As in the free-space viscothermal case, this is not the physical energy of the system, it is akin to a Lyapunov function [161] that reduces to the physical energy when $\eta \to 0$. Since all the individual terms are positive, the "energy" of the system is non-increasing.

2.1.5 Room impulse responses

A central topic in the study and reproduction of room acoustics is the concept of a *room impulse response* (RIR). A room impulse response represents a signal that would be recorded at some listening position in a room (using an infinitely small omnidirectional microphone) in response to an impulsive point-source located at some other source position in the same room. This serves as a *partial*[18] signature of the acoustics of the room, and can be used in a straightforward manner to produce artificial reverberation through the convolution operation (in time).

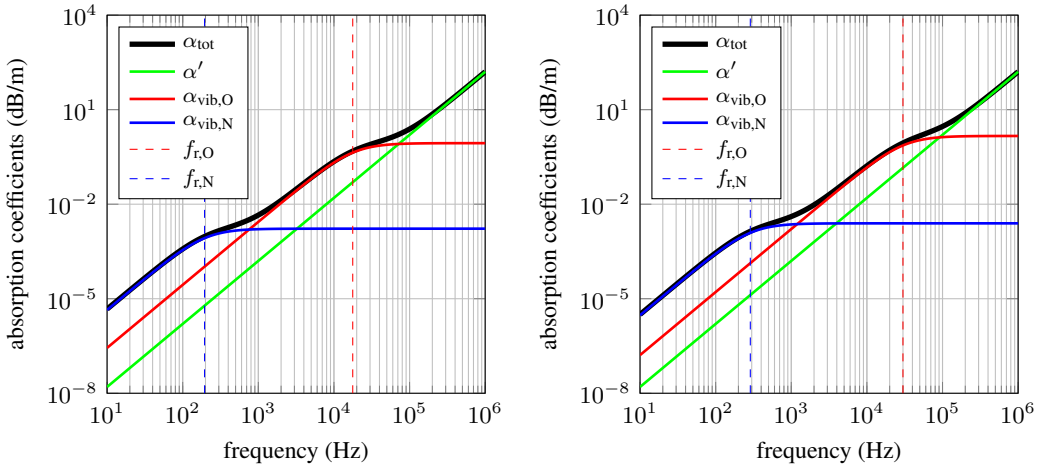
In order to illustrate this concept, we will consider the following PDE problem:

$$\Psi_{tt} - c^2\Delta\Psi - c\eta\Delta\Psi_t = c^2 F(\mathbf{x}, t)\,, \quad F(\mathbf{x}, t) = q_s(t)\delta(\mathbf{x} - \mathbf{x}_0) \tag{2.85}$$

on an enclosed domain $\overline{\Omega}$ with some associated boundary conditions over the boundary surface, and

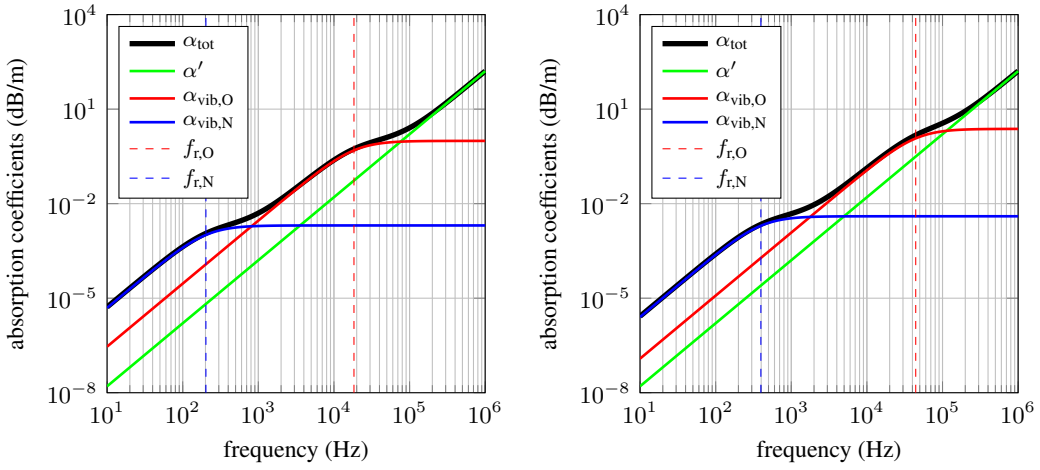
[17] 50% relative humidity and 18°C–20°C is perhaps more typical for indoor conditions, in which case $\eta \approx$ 1e-6 m.

[18] There are an infinite number of impulse responses for a given room—one for each source and listener pair of positions.



(a) $T = 15°\text{C}$, $h_{\text{r}} = 40\%$

(b) $T = 15°\text{C}$, $h_{\text{r}} = 60\%$



(c) $T = 20°\text{C}$, $h_{\text{r}} = 30\%$

(d) $T = 20°\text{C}$, $h_{\text{r}} = 60\%$

Figure 2.8: Absorption coefficients in dB/m and relaxation frequencies, for various values of temperature (T) and relative humidity (h_{r}) at standard air pressure (101.325 kPa).

with the initial conditions on the acoustic field such that $\Psi = \Psi_t = 0$ for $t \leq 0$.

The input to this linear system is considered to be $x(t) = q_s(t)$ at the source position $\mathbf{x}_0$. The output of the system is considered to be $y(t) = \Psi(\mathbf{x}_1, t)$, where $\mathbf{x}_1$ is known as the listening position.[19] With the constants c, ρ, η fixed, we can consider this input/output system in the form of a signal processing block diagram, as shown in Fig. 2.9. In this context, the block that represents the room can be seen as a filter, and the transformation from $x(t)$ to $y(t)$ is known as a filtering operation.

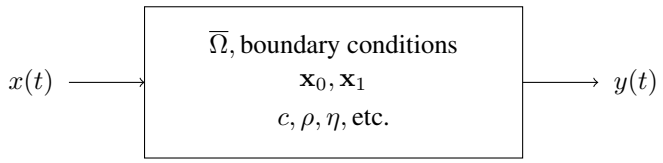


Figure 2.9: Representation of a room as a time-domain filter.

Since this PDE problem is both linear and time-invariant,[20] we can completely describe this filter by its impulse response. The impulse response is obtained by using the Dirac delta as an input to the system, and the output, denoted $h(t)$, is the impulse response of the filter, which is also a "room impulse response". Fig. 2.10 illustrates the acquisition of the impulse response.

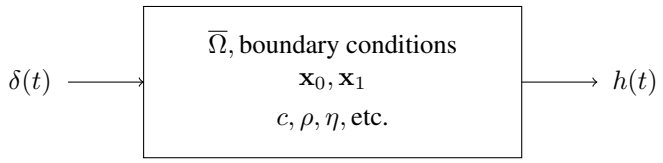


Figure 2.10: Illustration of method to obtain impulse response

Provided that one can obtain this impulse response, the output $y(t)$ to the filtering operation illustrated in Fig. 2.9 can be calculated by the convolution operation:

$$y(t) = (x * h)(t) = \int_{-\infty}^{\infty} x(\tau) h(\tau - t) \, \mathrm{d}\tau \tag{2.86}$$

The above convolution operation can also be represented in the frequency domain as a multiplication:

$$\hat{Y}(\omega) = \hat{X}(\omega)\hat{H}(\omega) \tag{2.87}$$

where

$$\hat{X}(\omega) = \int_{-\infty}^{\infty} x(t) \mathrm{e}^{-\mathrm{j}\omega t} \, \mathrm{d}t \tag{2.88}$$

and as such, $\hat{H}(\omega)$ serves as a frequency-domain signature of the room in question. Finally, we note that the limits of integration in the above convolution operation can be reduced to $[0, t]$ provided that $x(t) = 0$ for $t < 0$.

[19] One could equally start with the pressure wave equation (2.17) and then the input to the system could be $\dot{q}_s(t)$ and the output would be in terms of a pressure signal.

[20] This wave equation is time-invariant because ∂_t is a time-invariant operator, provided that the source and receiver are static. In the case of a moving source or listener, the system is no longer time-invariant. Moving sources and listeners can be incorporated into a wave-based model of room acoustics (not necessarily using impulse responses), but they will not be considered in this work.

2.2 Computational methods for room acoustics in the time-domain

In this section, we give an overview of some computational methods that can be applied to the simulation of room acoustics, starting with the standard geometric methods (ray-tracing and image source methods), followed by numerical methods for solving the wave equation and related PDEs, such as finite difference, finite element, finite volume, and boundary element methods, and as well as Fourier transform–based spectral methods.

2.2.1 Geometric methods

Geometrical acoustics

Geometrical methods are based on the *geometrical acoustics* description of sound, which is to assume the sound wavelength is small in comparison to the dimensions of the room or obstacles in the room [140], precluding wave diffraction effects. Under these assumptions, sound can be described by "sound rays" (akin to light rays in optics) which represent normals to infinitely small portions of a spherical sound wave emanating from a simple point source. It is assumed that sound rays travel in straight lines and decay in intensity (energy) proportionally to $1/r^2$ where r is the distance from the source [298].[21] Sound rays reflect off surfaces in a room, and a portion of a given reflection is taken to be specular, behaving according to a mirror-image principle,[22] while the other portion is taken to be non-specular, which means that rays scatter in a somewhat random direction, or uniform directions due to the roughness of the boundary surface [140]. The ratio of non-specular to specular reflection is defined by a *scattering coefficient* or a *diffusion coefficient* [63]. Also at each reflection, a portion of the energy of the sound ray is absorbed, as characterised by an *absorption coefficient* for the surface. All of these coefficients take on values between zero and one [228]. Absorption coefficients may be defined in terms of a specific incidence angle (e.g., normal incidence), or specific as angle-averaged coefficients, under the assumptions of a *diffuse field* [63].

From this geometrical framework, various geometric methods (simulation techniques) have been developed over the years, but we will only cover the two main approaches: the *ray-tracing* approach [232, 138] and the *image-source method* [3, 35]. See [228] for an exhaustive review of the various other geometric methods presently available. There are of course important extensions of geometric methods to include simplistic diffractions effects, using geometric interpretations of diffraction [125], but we cannot reasonably cover them here; see, e.g., [263, 278, 228]. Air absorption effects are treated in various ways: either by lumping air absorption losses into octave-band propagation decays, by applying a frequency-dependent filter to each propagation path [115], or by post-processing the impulse response with frame-based filters [253].

The main algorithm

Before describing these methods individually, it is worth explaining what they both aim to achieve; something we will refer to as the "main algorithm".

The aim of both ray-tracing and image source methods is to retrace, or reconstruct, some or all of the possible *paths* taken by a *ray* emanating from a *source position*—considered to be a point—and arriving at some *listener position*, which is considered to be a point or a small volume. Generally these methods

[21] Sound rays can be seen as plane waves, as described by (2.16), and in accordance with a plane-wave decomposition of the free-space Green's function (2.22).

[22] The plane-wave interpretation of a fully specular reflection is described in Section 2.1.3.

differ in the manner by which they reconstruct the paths that include wall reflections. Once a valid path between the source and listener is reconstructed by either the image source or ray tracing methods, the distance travelled by the ray, combined with the speed of sound, and the reflections undertaken—characterised by absorption and possibly scattering coefficients—determine the amount of energy (or pressure amplitude) that arrives from that ray during some epoch of time. This energy, along with the energy from other valid ray-paths, is collected into a *histogram*, wherein each bin corresponds to a time-sample, whose duration is chosen according to some sampling frequency. This histogram is also called an *echogram* [228], and is interpreted as the energy envelope of a room impulse response (RIR), which is also known as the "energy impulse response" [297].

This main algorithm is terminated at some point as designated by the user, e.g., based on some decision about a sufficient number of paths being reconstructed, or based on some assumption about the decay of energy over time in the system (e.g., using Sabine's formula [140]).

From the echogram, one can compute room acoustics attributes such as reverberation time and clarity [297]. If one desires an (approximate) impulse response for auralization purposes, one must convert from the energy impulse response to a pressure impulse response. Generally, the phase is not modelled by these methods (inherently), so this conversion is not uniquely determined, but various conversion methods have been proposed [297, 141], and the simplest approaches are taking the (positive) square root of the energy histogram [297], or simply reinterpreting the energy as a pressure [228]. The resulting impulse response can then be convolved with a dry audio signal in order to provide an *auralization* [128] of the room.[23]

Frequency-dependent behaviour can be incorporated in the following approaches by assuming that the energy associated with each ray pertains to some frequency band of a given width (e.g., an octave). This type of frequency-dependency requires multiple simulations, one for each frequency band of interest. The individual responses obtained can then be recombined using a bank of digital filters. An alternative way to incorporate frequency-dependence into geometric methods is to interpret the ray as a travelling Dirac delta impulse, and to interpret a wall reflection as a convolution of the Dirac delta with a digital filter representation (impulse response) of the wall's frequency response. Reflections from multiple wall materials can be treated by additional convolutions with associated wall impulse responses [10].

Ray-tracing

In the classic ray-tracing approach [232, 138], ray-paths are traced from the source to the listener (receiver) by shooting a large number of rays in various directions and checking if they eventually strike the receiver. Since one cannot be sure which firing directions will yield valid paths, it makes sense to distribute the firing directions uniformly over a sphere, or to randomise the paths according to some probability distribution (e.g., uniformly for an omnidirectional source) [298]. Since the chances of identifying a valid path between the source and *pointwise* listener position will be generally quite small, it makes sense for the receiver to occupy some volume (rather than being a point in space), in order to increase the chances of "collecting" a ray. Rays do not necessarily terminate upon hitting the listener, they can keep being traced. The amount of energy collected at a receiver can also depend on the distance travelled through the receiver volume [228].

The ray tracing algorithm allows for non-specular reflections, either by redirecting ray reflections randomly (weighted accordingly to the scattering coefficient), or splitting them into specular + diffuse

[23] This process is described by (2.86) in Section 2.1.5, but it is important to note that the obtained impulse response in geometric methods is not necessarily that of an underlying PDE model.

rays (this can be more costly as the number of rays grows). Rays eventually "die off" and stop being traced once they decay sufficiently in energy. A source directivity can be modelled by weighting the distribution of fired rays accordingly. The cost of ray-tracing, which consists mainly of ray-triangle intersection tests, scales with the number of surface triangles or polygons. As such, room models (surface meshes) are often simplified before carrying out ray-tracing, based on prior assumptions about the necessary geometrical details of a room that will be perceptually relevant to acoustic responses [299].

One disadvantage of ray-tracing is the fact that not all early reflections are necessarily found (traced), and such early reflections are perceptually important [140]. As such, the early reflections are often computed with the image source method, to be described next.

Image source method

The image source method [3, 35] has the same aim as ray-tracing—to collect energy (or pressure amplitude) emanating from a source into temporal bins—but the way it traces source-receiver paths is fundamentally different. In the image source method, source-receiver paths are calculated in a deterministic manner, with every path being found in the limit of the algorithm being run for an infinite time. Essentially, one creates a mirror image of the domain by reflecting the domain through one of its boundary faces (the domain is assumed to be polyhedral). This creates a *virtual domain* (mirror domain) with its own *virtual source* that traces a path to the listener passing through the reflecting wall. The ray energy that arrives in a given epoch of time is determined by the distance from the virtual source to the listener and the wall that the path intersects. Once this process is carried out for all boundary faces, the set of *first-order* reflections has been calculated.

Let us call each one of those virtual domains, the "first-order virtual domains". In order to calculate the *second-order* reflections, the first-order virtual domains are reflected through their own boundary faces, in order to create new (second-order) virtual domains. Second-order reflections are then traced from the second-order virtual sources to the listener, taking into account all recursive orders of reflections. This process continues until the user decides to stop, but this is typically only carried out to low orders, since the computational complexity grows exponentially with the order of reflection. Beam tracing algorithms can help in this regard [184, 143]. When the angles between adjacent walls are not acute, each path must be verified for validity (visibility) [35]; e.g., may be cases where the reconstructed path is obstructed by an concave part of the room. By its nature, the image source method cannot handle non-specular reflections.

It is worth mentioning that the image source method is essentially a discrete version of the *reflection method* [251], which is based on the well-known *image principle* in the study of PDEs [61]. When first introduced to the room acoustics world, it was shown to solve the wave equation in the special case of a 3-D box with rigid walls, in the limit of infinite-order reflections [3]. However, strictly speaking, the mathematical scope of the image method in solving PDEs is limited to special case geometries, and generally only with lossless boundaries [124]. For non-cuboid geometries with non-rigid boundaries, the obtained impulse responses are not necessarily those of the underlying PDE, even in the limit of infinite-order reflections.

The image source technique is generally regarded to accurately simulate *early reflections* (e.g., first and second orders), and it is often combined with the ray tracing algorithm for the later reflections [228].

2.2.2 Finite difference methods

Finite difference methods have seen the most development in terms of numerical methods for transient equations for room acoustics. As they are the main focus of this thesis, they are deserving of a short historical treatment, which will help to put their use in room acoustics into context with their wider use in other fields.

Finite difference methods – an early history

Finite difference methods are perhaps the oldest numerical method for solving PDEs (or ODEs) (e.g., [34, 240]), and they are closely-related to *numerical integration* and *quadrature* techniques. Their use for the wave equation, and other time-dependent PDEs, generally starts with the seminal work of Courant et al. [60] and Richardson [203, 205, 204]. Many important developments in finite difference methods took place in the 1940s and 1950s for fluid modelling [185, 296, 146, 206], and other early applications were in numerical weather modelling [50, 195, 153, 78, 9]. There is also parallel work to be found in the Russian literature [121]. Later, with the spread of computer technology, finite difference methods found early uses in geophysical acoustics applications [4], in electromagnetics [311], and as sound synthesis techniques based on musical acoustics [212]. Seminal texts on the subject of finite difference methods (and other numerical methods) include [206, 86, 57, 5, 268, 260]. See also [269] for a detailed history of finite difference methods.

On centered differences and staggered grids

In the context of the above-mentioned applications (neglecting many other PDEs, such as, e.g., the heat equation), finite difference schemes are generally employed for first-order systems of equations (conservation equations) or second-order equations (e.g., the wave equation). In either case, it has been common practice for some time to use finite difference operators that are *centered* about the point in time or space of interest (if possible), for accuracy reasons [240]. A convenient way to accomplish this task for first-order systems is to place different discrete field components on *staggered grids*, i.e., individual grids (generally Cartesian) that are staggered in space and/or time. Probably the first to apply this technique to time-dependent PDEs was Richardson, in 1922 [205]. Influenced by Richardson's early work, staggered grids were commonly employed in the modelling of fluids and in numerical weather prediction in the 1950s and 1960s [296, 145, 206, 153, 78, 9]. Though today, perhaps the most well-known (at least in the area of room and musical acoustics) use of staggered grids is in Yee's classic scheme from 1966 for Maxwell's equations [311], later popularised by Taflove and dubbed the "finite difference time domain" (FDTD) method [264].

Early uses of finite difference methods in related areas of acoustics

Before discussing the uses of finite difference methods in room acoustics, it is worth discussing briefly their (earlier) use in other areas of acoustics.

Finite difference methods had been applied to acoustical wave problems since the 1950s [206], with much development in the geophysics community in the late 1960s, and in the 1970s and 1980s [4, 2, 295, 66, 111, 120]. They continue to be popular tools in that field [53, 156, 72] often as a first-step in imaging applications for depth (oil) exploration [17]. Staggered grid formulations for the linear acoustics first-order system (2.1) have appeared in the geophysics literature [295, 120], inspired by the use of staggered

grids for 1-D linear acoustics [206], independently of developments of the same schemes adapted from Yee's scheme, such as, e.g., [166].

As mentioned previously, finite difference schemes were used for early *sound synthesis* techniques [212, 110, 12]. Musical acoustics was also studied from the point of view of finite difference (and finite element) methods in the 1990s [49, 202]. See also [24, Chapter 1] for more examples.

It is also worth mentioning a closely-related technique known as the *digital waveguide* [246, 245]—an algorithm for sound synthesis and musical acoustics modelling. The digital waveguide algorithm can be seen as a computationally efficient reformulation of Courant's classic finite difference scheme for the 1-D wave equation [60]. Its simple (and efficient) formulation led to its wide use in the area of computer music and physical modelling sound synthesis, as a simple model for strings and wind instruments [248]. Digital waveguides spurred large amounts of research in physical modelling techniques and analysis of musical acoustics in the early 1990s and 2000s (e.g., [285, 229, 292, 18]; also see [24] and [250] and references therein).

The development from the digital waveguide literature that is most relevant to this thesis is certainly the *digital waveguide mesh* (DWM) [286], a multidimensional version of the digital waveguide algorithm. The simplest DWM topology (the "rectilinear DWM" [286]) can be viewed as a reformulation of Courant's classic scheme in 2-D and 3-D [60, 86], although in this case the digital waveguide formalism did not turn out to be advantageous from a computational efficiency point of view. On the other hand, it led to many important developments related to finite difference approaches in the field of room acoustics, to be discussed next.

Early uses of finite difference methods in room acoustics

An early application of finite difference methods to room acoustics—although not a time-domain application—can be found as an exercise in the Collatz's 1960 textbook on numerical methods [56], wherein an eigenvalue problem is posed to calculate the modes of a 3-D room with rigid walls and two open doorways. The use of finite difference methods for *time-dependent* room acoustics came much later, in the early 1990s, by various authors, independently and around the same time.

The first uses of finite difference methods for transient room acoustics were adapted from Yee's scheme, to the first-order system (2.1). In Japan, Chiba et al. [51] published an adaptation of Yee's seminal electromagnetic staggered-grid scheme [311] in a room acoustics setting, but this contribution seems to have gone largely unnoticed outside of Japan (as it was published in Japanese and has yet to be translated to English). In the same year, Botteldooren (independently) adapted Yee's scheme to the same acoustics system for room acoustics [40, 37, 38], with his contribution having more of an effect in the wider room acoustics community. At the same time, Botteldooren also presented a generalisation of his finite difference scheme to the finite volume framework—another central topic in this thesis (more to be said about finite volume methods later). From another point of view, Savioja et al. applied van Duyne's digital waveguide mesh [286] to the problem of 3-D room acoustics [223], opting for an equivalent (second-order) finite difference formulation [227] rather than the preferred (and less efficient) wave-scattering formulation from the digital waveguide formalism.

As a general remark, it is worth noting that in the room acoustics community, "finite difference methods", whether expressed in first-order staggered form or second-order unstaggered form, have become synonymous with "finite difference time domain" (FDTD) methods, after the electromagnetics nomenclature [264]. Furthermore, in the room acoustics community, it is common to find a *staggered grid* referred to as a "Yee grid". However, it is worth recalling that staggered schemes for the acous-

tic system (2.1), essentially equivalent to those of Chiba et al. and Botteldooren, were independently formulated in the geophysics community [295, 120], stemming not from Yee's contribution, but from earlier use of staggered grids in acoustics [206].

Further developments in finite difference methods for room acoustics

The use of finite differences in the room acoustics community seem to have developed at first along two independent paths, respectively stemming from the contributions of Botteldooren (in first-order form), and from the work of Savioja et al. (and van Duyne [286]) in digital waveguide/finite difference mesh (second-order) form. Before listing some of these developments, it is should be mentioned—to put these two lines of development into context—that eventually it was realised that these two flavours of finite difference schemes are equivalent for the calculation of the interior acoustic field [21, 131, 43].[24] In terms of the staggered-grid form, there have been many important contributions (e.g., [312, 41, 42, 217, 218, 231, 271, 300, 118, 93]); however, as this thesis is mostly concerned with the second-order (wave equation) form, we will focus on developments related to the second-order form.

Initially, work that progressed along the lines of second-order formulations was heavily influenced by the digital waveguide wave-scattering formalism, and was largely inspired by signal processing techniques. While many made use of the more efficient (yet not the most general) equivalent finite difference formulation *in practice*, most theoretical developments remained faithful to the DWM formalism [221, 179, 144, 1, 122, 48]. Bilbao showed that the "finite difference/digital waveguide meshes" were special cases of more general finite difference schemes, which opened the door to significant improvements in accuracy and efficiency [21, 22]. The DWM formalism is not as prevalent these days, with finite difference methods (or "FDTD") taking its place in the room acoustics community (e.g., [131, 254, 36, 237]), but the DWM remains influential and a source of inspiration in the particular developments towards non-Cartesian mesh topologies[25] (summarised in, e.g., [48, 182]).

There have been many practical (applied to real-world acoustics) and theoretical contributions in the use of these schemes for room acoustics, but we focus on the developments directly related to the concerns in this thesis. Beyond the initial presentation of van Duyne's digital waveguide mesh [286], the "interpolated digital waveguide mesh" of Savioja et al. [221, 225] was an important first-step in addressing the problem of numerical dispersion, as were the many non-Cartesian digital waveguide mesh topologies that arose over the years [80, 287, 21, 81, 144, 1, 48]. Bilbao [21, 32, 23] generalised most of these digital waveguide mesh topologies (and interpolated forms) to the finite difference formalism and worked out the underlying constraints required for *stability* (which were be implicitly satisfied by the DWM formalism), as well as constraints for consistency (which were not necessarily satisfied by the DWM formalism), and constraints for *isotropy* in the approximation error (of practical and theoretical importance).

Kowalczyk and van Walstijn continued the development of the schemes initially put forth by Savioja et al. and later generalised by Bilbao, supplementing them with frequency-dependent boundary conditions in order to model wall materials with general frequency dependencies [131, 133, 137, 135, 136]. Kowalczyk and van Walstijn also introduced a particularly useful metric for computational efficiency in finite difference schemes [291],[26] allowing them to identify computationally efficient special cases of the (so-called) "compact explicit schemes" [134, 136]. Kowalczyk and van Walstijn also put forth

[24] In hindsight, this is not surprising as they comprise the simplest possible discretisations to equivalent PDE systems, and such equivalences were known for precursors to these multidimensional schemes [206].

[25] In fact, two new digital waveguide mesh topologies are presented in Section 6.4.3.

[26] This efficiency metric will be of particular use in Chapter 4.

an alternative depiction of numerical dispersion error, as a function of temporal frequencies rather than wavenumbers, through an inversion of the numerical dispersion relation,[27] (inspired by [230]), which provided important insights into the temporal bandwidth limitations of finite difference schemes [291, 131].

The boundary conditions of Kowalczyk and van Walstijn for the interpolated ("compact explicit") schemes comprised an important step towards to the adoption of finite difference methods for use in room acoustics modelling, and they were put to use by many in the same field of research (e.g., [222, 159, 279, 238, 36, 255, 19]). However, as first indicated only recently by Borrel-Jensen [36], the boundary conditions that were proposed are not necessarily stable under general room geometries (beyond rectangular boxes aligned with Cartesian grid axes); this will be investigated in more detail in Section 5.4.

Other important developments include a general construction for discrete Laplacians on Cartesian grids in 3-D put forth by Bilbao, leading to a generalisation of second-order accurate and isotropic finite difference for the wave equation in 3-D [25] which is relevant to Chapter 4. Also, Botts provided important insights into the relationships between staggered and second-order forms [43] (as will be further discussed in Chapter 3), and he identified and analysed important truncation effects that can appear in these types of finite difference schemes [44], which is relevant to some of the analyses presented in Chapter 5.

While the more practical side of finite difference modelling with *parallel computing architectures* will not be covered in this thesis, it is important to keep in mind the scope for such parallel implementations. As such, the following studies are worth noting: [177, 210, 236, 222, 36, 303, 100, 304].

Finally, we note that there of course many important developments in the use of finite difference methods for wave simulation outside of the field of room acoustics that will be referred to in this thesis (e.g., since the 1980s: [294, 275, 241, 66, 281, 282, 54, 7, 53, 90]), but we will not attempt to provide a well-deserving historical treatment here.

Finite difference methods for the wave equation: computational costs

Before moving on to other numerical methods, it is worth briefly discussing the computational costs of finite difference schemes for the wave equation. The computational costs of finite difference methods for room acoustics can seem heavy, but it is important to put these costs into context with the underlying PDE model.

We consider only the second-order form, as it is generally more computationally efficient than the first-order form under the assumption of a constant c [43, 304]. Most of the second-order finite difference methods considered in this thesis (and most of the finite difference schemes commonly employed for room acoustics) can be represented as the following matrix recursion:

$$\mathbf{A}\mathbf{u}^{n+1} = \mathbf{B}\mathbf{u}^{n} + \mathbf{C}\mathbf{u}^{n-1} \tag{2.89}$$

where $\mathbf{u}^n$ is an $N \times 1$ column vector, representing the entire spatially-discretised acoustic field (pressure or velocity potential) at time $t = nk$, where k is the time-step, and where $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are $N \times N$ matrices. Typically $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are sparse matrices and $\mathbf{A}$ and $\mathbf{C}$ are often diagonal. Furthermore, in regular-grid finite difference schemes (such as Cartesian schemes), $\mathbf{A}, \mathbf{B}, \mathbf{C}$ do not need explicit storage, as they are relatively uniform. In such cases, only one row needs to be stored, and a few variations of that row for grid points near the boundary of the domain.

[27] Although not necessarily a well-defined inversion (i.e., one-to-one), as will be seen in Section 4.5.3.

The majority of schemes used in room acoustics applications are *explicit* schemes, wherein $\mathbf{A}$ and $\mathbf{C}$ are diagonal matrices. As such, they can be written as follows:

$$\mathbf{u}^{n+1} = \mathbf{B}'\mathbf{u}^n + \mathbf{C}'\mathbf{u}^{n-1} \tag{2.90}$$

where $\mathbf{B}' = \mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{C}' = \mathbf{A}^{-1}\mathbf{B}$ are rewritten as such in order to reflect the fact that $\mathbf{A}^{-1}$ is trivially calculated, and these can be seen as new sparse matrices, and furthermore, $\mathbf{C}'$ remains diagonal. In this case, the storage requirements for the scheme are simply $2N$, since $\mathbf{u}^{n+1}$ can overwrite $\mathbf{u}^{n-1}$ in place. As the matrices are sparse, and since the main operation at each time-step is a sparse matrix multiplied by a vector, the complexity of this explicit algorithm (per time-step) is $\mathcal{O}(N)$.[28]

Let us now consider the specific case of room acoustics in three spatial dimensions. By 3-D spatial sampling considerations, and by the dispersion relation of the wave equation (relating temporal frequencies to spatial frequencies), one finds that $N \propto f^3$, where f is the maximum temporal frequency of interest.[29] Thus, the storage requires are $\mathcal{O}(f^3)$, and the computation per time-step is $\mathcal{O}(f^3)$. It is worth recalling from Section 2.1.3 that for any 3-D domain, the modal density, or the number of degrees of freedom, will be $\mathcal{O}(f^3)$. In other words, the complexity of the continuous problem scales with that of the finite difference scheme, and the computational costs of the finite difference scheme—without considering approximation errors—reflect the complexity of the underlying problem. Considering any extended duration of time, the computational cost will be $\mathcal{O}(f^4)$, as the number of time-steps required for a given duration of time will be proportional to f (simply by temporal sampling considerations).

2.2.3 Finite element and finite volume methods

Finite element methods [62, 314, 116] and finite volume methods [293, 14, 107, 152] can be seen as closely-related generalisations of finite difference methods [313, 189] (see also [269]). Applied to wave equation, these methods can also be written in the form (2.89), although the structure of these matrices will vary from the finite difference approach, and these differences are discussed below.

The main advantage of these approaches over finite difference methods is they can easily operate over *unstructured grids*, allowing for adaptation to boundary surfaces that do not align conveniently with grid axes. The disadvantage of these methods is that they are not as easily implemented in practice as regular-grid finite difference methods, as they typically require a pre-processing *meshing* step, and additional storage for the mesh data. This additional storage means that most, if not all, of the rows of the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ need to be stored. However, it is also often the case that $\mathbf{A}$ and $\mathbf{C}$ are diagonal.

Finite element methods

In finite element methods, the spatial domain is discretised into polygonal or polyhedral *elements*—typically triangles or tetrahedra—and localised basis functions are chosen to span these elements [314]. The solution is projected onto the basis elements and some chosen residual is minimised, resulting in a linear system of equations whose solution (the weighting coefficients) leads to an approximation to the unknown variable of interest (e.g., the pressure field). The construction of the matrices involved in the linear system of equations is an important pre-processing step, as is the choice of the finite element *mesh*. There are generally two matrices that result from this process, a *mass matrix*, and a *stiffness*

[28] Implicit schemes, where $\mathbf{A}$ and $\mathbf{C}$ are sparse but not diagonal, require a linear system solution, but the scheme can be chosen such that this can be accomplished with iterative methods, and the complexity remains $\mathcal{O}(N)$, whereas the storage costs are then at least $4N$. For simplicity in this discussion, we focus on the explicit case.

[29] This will be explained in more detail in Chapter 4.

matrix and both are sparse due to the localised nature of the finite element basis functions chosen. Finite element methods in room acoustics (or other enclosed spaces) are often employed in the frequency domain (e.g., [242, 64, 10]), but time-domain approaches can also be derived, using, e.g., standard leapfrog time-integration [262, 53, 186]. In the time-domain formulation, the mass matrix is essentially $\mathbf{A}$ and the stiffness matrix is part of $\mathbf{B}$. When the mass matrix is not diagonal, the time recursion is implicit (requiring a linear system solution), but *diagonal mass lumping* techniques can be used to make it explicit (not requiring a linear system solution) [53].

Finite volume methods

Finite volume methods can be seen as special cases of finite element methods, where the basis is simply a flat function over the element (rather than, e.g., a "hat" function) [313]. Typically, polygonal or polyhedral elements (*cells* in the nomenclature of finite volume) are used for finite volume methods, but curved cells are also permitted. For a derivation of finite volume methods independent from a finite element construction, one typically considers integral forms of the PDE of interest (involving the divergence operator) and one applies Gauss' (divergence) theorem, followed by the use of averaging over cell-faces and cell-volumes (one such derivation is given in Chapter 6). Finite volume methods also require some directional derivatives to be approximated, for which finite differences are used. These finite differences generally operate across the centres of adjacent cells (cell-centered finite volume) or along the edges of cells (i.e., between vertices of cells, vertex-centered finite volume) [198]. In this work, we consider mostly the cell-centered approach. Finite volume methods are often used in the time-domain for computational fluid dynamics [152].

In regards to the matrix update equation (2.89) for finite volume methods (or finite volume time domain (FVTD) methods), $\mathbf{A}$ is typically a diagonal matrix that represents the volumes of cells, and $\mathbf{B}$ has information about directional derivatives and surface areas of faces that adjoin cells. An important property of finite volume methods is that they are *flux conserving* [152], which is a property that lends itself naturally to the use of energy methods for stability analysis.

Finite volume methods were first used for room acoustics by Botteldooren [40, 37], proposed as a generalisation of the simple Cartesian finite difference approach and recommended as a complementary method for the modelling of irregular boundaries.[30] Finite volume methods were not as widely adopted in the room acoustics community as were finite difference methods, most likely due to the additional complexity required to take advantage of this generalisation. However, finite volume methods have recently re-emerged in the room acoustics literature in a more convincing way, allowing for flexible modelling of boundaries with frequency-dependent behaviour under a numerically stable framework [26]. Finite volume methods applied to simple room acoustic problems also appear in [219, Section 5.3].

Computational costs

Finite element and finite volume methods have the same computational complexities as finite difference methods: $\mathcal{O}(f^3)$ for storage and $\mathcal{O}(f^4)$ for operations (for the 3-D wave equation). However, when using fully unstructured grids, the actual memory costs may be higher than those of finite difference methods,[31] since $\mathbf{A}$ and $\mathbf{B}$ need to be stored in full (although still in a sparse matrix representation). On the other hand, when used as a complement to regular finite difference schemes, additional storage is

[30] This complementary use of finite volume methods is an essential part of the approach to room acoustics modelling that will be used in this thesis.

[31] It must be said that this is only for approximately the same grid spacing or cell size, and disregards accuracy considerations, which may allow one approach use a coarser grid or mesh than another.

only necessary near the boundary of the domain. More will be said about this in Chapter 6 in regards to finite volume methods.

2.2.4 Boundary element methods

Boundary element methods (BEM) are based on the idea that one can express the solution to an inhomogeneous wave equation boundary value problem in terms of the free-space Green's function, the initial conditions, and the boundary conditions (and the known source term) [68]. This general solution is in the form of an integral equation which is not easily evaluated, therefore it must be approximated numerically. As such, boundary element methods do not discretise the wave equation itself, rather they discretise a general solution to the wave equation (or the Helmholtz equation). These methods are not represented by the matrix recursion (2.89).

For brevity, we do not present the full integral equation here; see [68, Eq. (3.0.13)]. In the absence of sources and initial conditions, this integral equation is known as the "Kirchhoff integral equation" (KIE) [68, Eq. (3.5.24)], and for a time-harmonic solution it is known as the "Kirchhoff-Helmholtz integral equation" (KHIE) (see, e.g., [106, Eq. (6)]). BEM in the time-domain typically uses the KIE, and BEM in the frequency-domain uses the KHIE. These integral equations state that the pressure (or acoustic velocity potential), in reaction to a point source somewhere inside the boundary can be predicted *anywhere inside the domain* if the pressure and normal velocity are known *everywhere on the boundary surface*. The unknown variable in this problem is the pressure over the boundary surface.

The unknown variable is solved for using some sort of element method (e.g., finite element) over the boundary. One discretises the boundary surface into *surface elements* (boundary elements), then one chooses a set of basis functions over the elements (e.g., piecewise polynomials). Enforcing the error to be zero either at sampling points in the elements (collocation methods) or over the entire elements (Galerkin methods) leads to a linear system of equations that solves for the pressure on the boundary, or more specifically, the weighting coefficients required to reconstruct the pressure from the basis functions. Once this system is solved over the boundary, the pressure on the interior domain is obtained by propagating the approximation over the boundary back to the interior point of interest using the known free-space Green's function.

The main advantage of BEMs is that the problem is reduced by one dimension, meaning that the number of degrees of freedom in the numerical approximation is decreased significantly. However, the system of equations that arises typically features a full matrix, so the linear system solution can be expensive computationally, and memory costs can be important for the storage of this matrix. In 3-D, and for some maximum frequency of interest f, the number of degrees of freedom in the problem scales as $\mathcal{O}(f^2)$ for BEMs, which is better than the $\mathcal{O}(f^3)$ required for other grid-based approaches. However, because of the generally full matrix involved, storage costs are $\mathcal{O}(f^4)$. BEMs are typically used in the frequency domain in the room acoustics (or for other sound enclosures) literature [13, 262]. BEMs in time-domain formulations are also possible and have been used for room acoustics modelling [262, 105, 219].

2.2.5 Fourier methods

Finally, we review another class of time-stepping, grid-based methods that have been applied to room acoustics [254, 113, 199, 165], which we will call "Fourier methods" after [83], also known as "Fourier spectral collocation methods" [187, 108]. Fourier methods aim to circumvent the problem of numerical dispersion in grid-based methods by transforming the problem to the wavenumber space, such that

spatial derivatives transform to scalar multiplications that can be evaluated exactly (in theory) [187, 130, 201, 266, 70, 154, 155, 85]. In practice, this is carried out using some discrete form of the spatial Fourier transform (discrete Fourier transform (DFT), discrete cosine transform (DCT), etc.), which in turn is carried out efficiently using the fast Fourier transform (FFT) algorithm [289]. However, these transforms, at least in discrete forms, generally operate over box domains (aligned with Cartesian grid axes), with particular boundary conditions naturally enforced (e.g., rigid or periodic). In other words, when applied to box domains with rigid or periodic boundary conditions, these methods are efficient at evaluating the wave equation. On the other hand, many challenges arise when they are applied to more realistic room acoustics problems.

The main approach of adapting Fourier methods to more general domains involves the concatenation of box-shaped domains, resulting in somewhat crude staircase approximations to non-trivial domains [199, 113]. This is a form of "domain decomposition" technique, which faces the problem of interfacing (passing data) between box domains in an accurate and stable manner. Ideally, such methods would be able to employ "two points per wavelength (PPW)", but their usage beyond box-domains often requires employing higher-grid resolutions in order to mitigate reflection errors that arise at subdomain interfaces [199],[32] and/or complicated iterative data exchanges at subdomain interfaces, which themselves require fine-grid glue layers operating at higher sample rates [113].

In such domain-decomposition Fourier methods, interfacing is typically carried out in the spatial domain, whereas spatial derivatives are evaluated in the wavenumber domain [199, 113]. Consequently, these methods require two or more sets of d-D FFTs per time-step (e.g, twelve FFTs per time-step in [113]), making the computational complexity higher than for other grid-based approaches: $\mathcal{O}(N \log N)$ for a total number grid of N points[33] (rather than, e.g., $\mathcal{O}(N)$ in finite difference methods) with $N \propto f^3$ in 3-D where f is the maximum frequency of interest.

The advantage of these methods is, of course, that they can be effective in addressing the problem of numerical dispersion over the interior of large box-shaped domains. The main disadvantage—beyond the aforementioned difficulties of practical implementation for non-cuboid geometries—is their ability to provide consistent and stable approximations to the room acoustics problem at hand. This is perhaps best understood by considering the fact that Fourier methods can also be seen as finite difference schemes in the limiting case of high-order accuracy in space (and possibly, also time) [70, 108]. Consequently, the problems of boundary modelling and numerical stability in finite difference schemes carry over to Fourier methods, and are, accordingly, more challenging. It should be said that at the time of writing, the formulation of reliable numerical stability conditions for such methods remains an open problem, leaving Fourier methods somewhat unreliable in practice.[34] Furthermore, impedance boundaries conditions (i.e., relating pressure and outward-normal particle velocity) have yet to be formulated in a consistent manner. Currently, such methods rely on adaptations of PMLs (perfectly matched layers) to model frequency-independent wall reflections [199, 113], approximating frequency-dependent wall losses via multi-band techniques (requiring a separate simulation for each frequency band of interest). Finally, as Fourier methods operate over box-shaped partitions, staircasing effects are inevitable problems that arise, and as a consequence, consistency with the underlying PDE problem—in the case of non-trivial room geometries under non-rigid boundary conditions—remains an open question.

[32] It has been claimed that 2.6 PPW is sufficient for accurate interfacing the "adaptive rectangular decomposition" method [199, 165], but experiments from a more recent study suggest that 8–10 PPW may be required [45].

[33] It is assumed here that the number of subdomains is much smaller than N.

[34] Indeed, numerical instabilities have been reported in the adaptation of such methods to non-trivial geometries (e.g., [114, 173, 178]).

2.3 Summary

In this chapter, we provided an overview of PDEs that can be used for time-domain room acoustic modelling, incorporating viscothermal losses in air and general frequency-dependent boundary conditions formulated as positive-real circuit networks. Energy analyses were provided and these will be mimicked in discrete forms in the following chapter in order to derive numerically stable schemes approximating—in a simple manner—the room acoustic PDE models presented so far. Additionally in this chapter, various approaches to room acoustic simulations were reviewed, including geometric (ray-based and image source–based) methods, and some standard wave-based methods, including finite difference, finite volume/element, boundary element, and Fourier methods, along with advantages and disadvantages of each approach.

Chapter 3

Room acoustics modelling in 3-D with the simple Cartesian scheme

The purpose of this chapter is to introduce some basic concepts in finite difference–based room acoustics modelling using the simplest Cartesian scheme in three spatial dimensions. For the most part, this chapter will follow Section 2.1 in a discrete setting. We start with the simplest finite difference approximation to the lossless wave equation, followed by an equivalent staggered-grid scheme for the first-order system (2.1). This is followed by boundary conditions in frequency-independent and frequency-dependent forms. We also provide a brief overview of approximation errors in the form of numerical dispersion (a topic to be treated in more detail in Chapter 4). Finally, these boundary conditions and schemes are extended to include viscothermal losses. Energy methods are used throughout this chapter in order to derive numerical stability conditions for the interior problem, and also to derive stable boundary conditions. Numerical examples are provided to illustrate properties of the presented schemes and boundary conditions, including conservation of numerical energy and frequency-dependent losses.

3.1 The basic 3-D scheme

For convenience, we begin with a discretisation of the second-order wave equation on the full space with the dependent variable $\Psi(\mathbf{x}, t)$, after which we will consider an equivalent staggered formulation for the first-order system in $p(\mathbf{x}, t)$ and $\mathbf{v}(\mathbf{x}, t)$. The model initial value problem (IVP) is then:

$$\Psi_{tt} = c^2 \Delta \Psi \,, \qquad t > 0 \,, \quad \mathbf{x} \in \mathbb{R}^3 \tag{3.1a}$$

$$\Psi = f_0 \,, \; \Psi_t = g_0 \,, \qquad t = 0 \,, \quad \mathbf{x} \in \mathbb{R}^3 \tag{3.1b}$$

where $f_0 = f_0(\mathbf{x})$ and $g_0 = g_0(\mathbf{x})$ represent the initial conditions, defined at $t = 0$.

3.1.1 Preliminaries

For the entirety of this chapter, and unless otherwise specified, it will be assumed that $c = 340\,\text{m/s}$ and $\rho = 1.2\,\text{kg/m}^3$, and any spatial quantities (positions, lengths) will be in units of metres and all times in seconds.

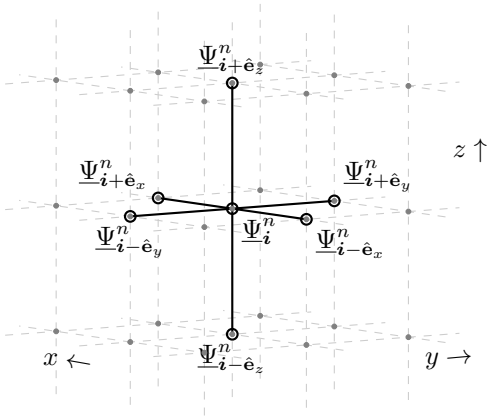


Figure 3.1: Seven-point stencil for δ_Δ

Grids and grid functions

In order to approximate a solution to above problem, we define a spatiotemporal grid function:

$$\underline{\Psi}^n_{\boldsymbol{i}} \approxeq \Psi(\boldsymbol{i}h, nk) \tag{3.2}$$

where

$$\boldsymbol{i} = (i_x, i_y, i_z) \in \mathbb{Z}^3\,, \quad n \in \mathbb{Z}^+ \tag{3.3}$$

where $\mathbb{Z}^+ := \{n \in \mathbb{Z}, n \geq 0\}$, and where $h > 0$ is the *spatial step* (or grid spacing) and $k > 0$ is the *time-step*. For convenience, we will also represent the infinite Cartesian grid by $\mathbb{Z}^3_h := \{\boldsymbol{i}h : \boldsymbol{i} \in \mathbb{Z}^3\}$.

Shift and finite difference operators

Temporal and spatial shift operators may be defined as follows:

$$e_{t\pm}\underline{\Psi}^n_{\boldsymbol{i}} := \underline{\Psi}^{n\pm1}_{\boldsymbol{i}} \tag{3.4a}$$

$$e_{x\pm}\underline{\Psi}^n_{\boldsymbol{i}} := \underline{\Psi}^n_{\boldsymbol{i}\pm\hat{\mathbf{e}}_x} = \underline{\Psi}^n_{i_x\pm1,i_y,i_z} \tag{3.4b}$$

$$e_{y\pm}\underline{\Psi}^n_{\boldsymbol{i}} := \underline{\Psi}^n_{\boldsymbol{i}\pm\hat{\mathbf{e}}_y} = \underline{\Psi}^n_{i_x,i_y\pm1,i_z} \tag{3.4c}$$

$$e_{z\pm}\underline{\Psi}^n_{\boldsymbol{i}} := \underline{\Psi}^n_{\boldsymbol{i}\pm\hat{\mathbf{e}}_z} = \underline{\Psi}^n_{i_x,i_y,i_z\pm1} \tag{3.4d}$$

where $\hat{\mathbf{e}}_x, \hat{\mathbf{e}}_y, \hat{\mathbf{e}}_z$ are the standard unit vectors in $\mathbb{R}^3$. Throughout this chapter, we will use the symbol w as a placeholder for the coordinate directions x, y, z. For example, the shift operators could be summarised by:

$$e_{w\pm}\underline{\Psi}^n_{\boldsymbol{i}} := \underline{\Psi}^n_{\boldsymbol{i}\pm\hat{\mathbf{e}}_w}\,, \quad w \in \{x, y, z\} \tag{3.5}$$

The above could also be written as:

$$e_{(-w)\mp}\underline{\Psi}^n_{\boldsymbol{i}} := \underline{\Psi}^n_{\boldsymbol{i}\pm\hat{\mathbf{e}}_w}\,, \quad w \in \{x, y, z\} \tag{3.6}$$

where $(-w)$ simply indicates the direction opposite to w.

Using the shift operators defined in (3.4), second-order difference operators along spatial dimensions

$w \in \{x, y, z\}$ can be constructed as:

$$\delta_{ww} := \frac{1}{h^2}(e_{w+} - 2 + e_{w-}), \quad w \in \{x, y, z\} \tag{3.7}$$

and this allows us to build the simplest discrete Laplacian in 3-D:

$$\delta_\Delta := \delta_{xx} + \delta_{yy} + \delta_{zz} \tag{3.8}$$

Expanding this discrete Laplacian operator,[1] applied to $\underline{\Psi}^n_{\boldsymbol{i}}$, we have:

$$\begin{aligned}\delta_\Delta \underline{\Psi}^n_{\boldsymbol{i}} = \frac{1}{h^2}\Big(&\underline{\Psi}^n_{i_x+1,i_y,i_z} + \underline{\Psi}^n_{i_x-1,i_y,i_z} + \underline{\Psi}^n_{i_x,i_y+1,i_z} \\ &+ \underline{\Psi}^n_{i_x,i_y-1,i_z} + \underline{\Psi}^n_{i_x,i_y,i_z+1} + \underline{\Psi}^n_{i_x,i_y,i_z-1} - 6\underline{\Psi}^n_{i_x,i_y,i_z}\Big)\end{aligned} \tag{3.9}$$

Note that this operator essentially reaches out over seven points in space, so it is also known as a seven-point discrete Laplacian. The spatial stencil[2] of points used in this discrete Laplacian is illustrated in Fig. 3.1. For convenience, we also write this operation as:

$$\delta_\Delta \underline{\Psi}^n_{\boldsymbol{i}} = \frac{1}{h^2}(Q^n_{i_x,i_y,i_z} - 6\underline{\Psi}^n_{i_x,i_y,i_z}) \tag{3.10}$$

where $Q^n_{i_x,i_y,i_z} = Q^n_{\boldsymbol{i}}$ is a weighted average over space, defined as:

$$Q^n_{i_x,i_y,i_z} = \underline{\Psi}^n_{i_x+1,i_y,i_z} + \underline{\Psi}^n_{i_x-1,i_y,i_z} + \underline{\Psi}^n_{i_x,i_y+1,i_z} + \underline{\Psi}^n_{i_x,i_y-1,i_z} + \underline{\Psi}^n_{i_x,i_y,i_z+1} + \underline{\Psi}^n_{i_x,i_y,i_z-1} \tag{3.11}$$

At this point we can reveal the (well-known) underlying approximations of these finite difference operators. First, we let the shift operators defined in (3.4) also be applicable to continuous functions:

$$e_{t\pm}\Psi = \Psi(\mathbf{x}, t \pm k) \tag{3.12a}$$

$$e_{x\pm}\Psi = \Psi(\mathbf{x} \pm \hat{\mathbf{e}}_x h, t) \tag{3.12b}$$

and similarly for $e_{y\pm}, e_{z\pm}$. Then for the finite difference operators we have the Taylor expansions:

$$\delta_{tt}\Psi = \partial_{tt}\Psi + \mathcal{O}(k^2) \tag{3.13a}$$

$$\delta_{ww}\Psi = \partial_{ww}\Psi + \mathcal{O}(h^2), \quad w \in \{x, y, z\} \tag{3.13b}$$

$$\delta_\Delta\Psi = \Delta\Psi + \mathcal{O}(h^2) \tag{3.13c}$$

where the terms $\mathcal{O}(h^2)$ and $\mathcal{O}(k^2)$ represent the *truncation errors* of the finite difference operators. The order of accuracy of each operator indicates how quickly the approximation approaches the continuous operation in the limit of small h or k. As such, these operators are said to be second-order accurate.

[1] It is worth noting that an early form of this operator appears in [193].

[2] The word "stencil", sometimes called a "computational molecule", can take on various meanings in the literature of numerical methods. It could be a temporal, spatial, or spatiotemporal set of points, and one could also associate "stencil weights" to each point and incorporate such weights into the meaning of "stencil". For the most part in this study, "stencil" will refer to a set of points (vectors) in space, while "update stencil" will refer to the set of spatiotemporal points involved in a local update equation.

It follows that if we combine these operators, we have an approximation to the wave equation:

$$(\delta_{tt} - c^2\delta_\Delta)\Psi = (\partial_{tt} - c^2\Delta)\Psi + \mathcal{O}(h^2) + \mathcal{O}(k^2) \tag{3.14}$$

Note that the approximation has a second-order truncation error in both the spatial and temporal grid spacings. Since the truncation error goes to zero in the limit of $h, k \to 0$, the approximation is said to be *consistent*.

Now that we have a consistent approximation to the wave equation, we replace the unknown solution Ψ with the grid function $\underline{\Psi}_{\boldsymbol{i}}^n$ to arrive at the following finite difference scheme:

$$\delta_{tt}\underline{\Psi}_{\boldsymbol{i}}^n = c^2\delta_\Delta\underline{\Psi}_{\boldsymbol{i}}^n \tag{3.15}$$

The above is known as the simplest scheme for the 3-D wave equation [60, 86]. In the room acoustics literature, this scheme is also referred to as the "rectilinear scheme" [21], "standard rectilinear" [134], or "standard leapfrog" [136].[3] The essential operation of the scheme is provided by rearranging (3.15) into the following update equation:

$$\underline{\Psi}_{\boldsymbol{i}}^{n+1} = -\underline{\Psi}_{\boldsymbol{i}}^{n-1} + (2 - 6\lambda^2)\underline{\Psi}_{\boldsymbol{i}}^n + \lambda^2 Q_{\boldsymbol{i}}^n \tag{3.16}$$

where $Q_{\boldsymbol{i}}^n$ is the summation (3.11), and where $\lambda := ck/h$ is a dimensionless quantity known as the *Courant number*. The importance of the Courant number will be discussed shortly, but for now it is worth mentioning that for the special case of $\lambda = \sqrt{1/3}$, the update equation reduces to:

$$\underline{\Psi}_{\boldsymbol{i}}^{n+1} = \frac{1}{3}Q_{\boldsymbol{i}}^n - \underline{\Psi}_{\boldsymbol{i}}^{n-1} \tag{3.17}$$

This special case is also known by its equivalent *digital waveguide mesh* formulation [286, 223, 21].[4] The significance of the choice $\lambda = \sqrt{1/3}$ is linked to convergence and stability of the scheme, as will be seen shortly.

It is also worth noting that we could rewrite this special case as:

$$\frac{\underline{\Psi}_{\boldsymbol{i}}^{n+1} + \underline{\Psi}_{\boldsymbol{i}}^{n-1}}{2} = \frac{\underline{\Psi}_{i_x+1,i_y,i_z}^n + \underline{\Psi}_{i_x-1,i_y,i_z}^n + \underline{\Psi}_{i_x,i_y+1,i_z}^n + \underline{\Psi}_{i_x,i_y-1,i_z}^n + \underline{\Psi}_{i_x,i_y,i_z+1}^n + \underline{\Psi}_{i_x,i_y,i_z-1}^n}{6} \tag{3.18}$$

which, essentially, describes the operation of the scheme in terms of temporal and spatial averages. In other words, (3.18) says that the local average of the acoustic field in time must be equal to the local average of the acoustic field in space. This is related to a classic interpretation that is used to describe the wave equation and Laplacian operator in the acoustics literature [176].

Returning to the more general update equation (3.16), we note that the time-step $n+1$ requires the storage of only two time-steps in the past, since $\underline{\Psi}_{\boldsymbol{i}}^{n+1}$ can overwrite $\underline{\Psi}_{\boldsymbol{i}}^{n-1}$ in place. Since the PDE problem starts at time $t = 0$ ($n = 0$), the first two states ($\underline{\Psi}_{\boldsymbol{i}}^0, \underline{\Psi}_{\boldsymbol{i}}^1$) must be specified. The first state is

[3] The moniker "rectilinear" could be seen as a misnomer since it might imply the use of a either a "rectangular grid", which is a regular grid based on congruent rectangular (prism) tiles, or a "rectilinear grid", which is a irregular (yet still structured) grid based on rectangular (prism) tiles that are not necessarily congruent to each other [256, 270]. The Cartesian grid used in this simple scheme is neither of these types.

[4] This was essentially the form used by Savioja et al. [223], but in terms of a pressure quantity rather than velocity potential. The pressure wave equation scheme (Eq. (3.58)) is discussed later.

simply set by the initial condition f_0:

$$\underline{\Psi}_{\boldsymbol{i}}^0 = [f_0]_h \tag{3.19}$$

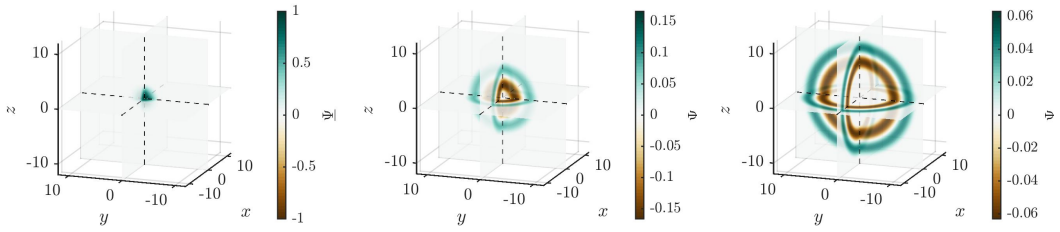
where $[f_0]_h$ represents a consistent approximation to f_0 on $\mathbb{Z}_h^3$.[5] Meanwhile, the second state can be approximated from f_0 and g_0 with the following *initialisation scheme*:

$$\underline{\Psi}_{\boldsymbol{i}}^1 = \underline{\Psi}_{\boldsymbol{i}}^0 + k[g_0]_h \tag{3.20}$$

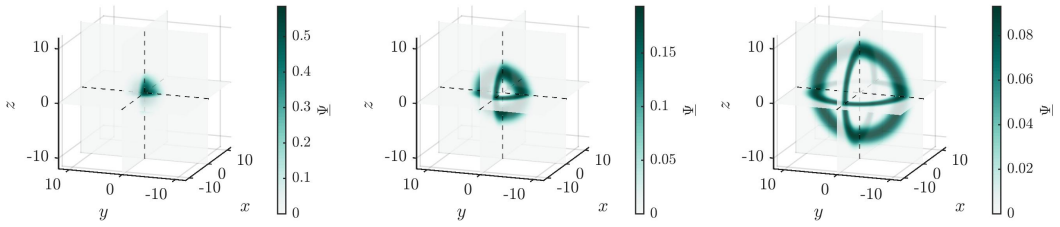
Although, for slightly more accuracy in the estimation of $\underline{\Psi}_{\boldsymbol{i}}^1$, it may also be preferable to make use of the following initialisation scheme [260, 244]:

$$\underline{\Psi}_{\boldsymbol{i}}^1 = \underline{\Psi}_{\boldsymbol{i}}^0 + k[g_0]_h + \frac{c^2k^2}{2}\delta_\Delta \underline{\Psi}_{\boldsymbol{i}}^0 \tag{3.21}$$

Some examples of 3-D wave propagation, simulated with this scheme starting from initial conditions, are illustrated in Fig. 3.2.



(a) $t = 0$ (b) $t = 4$ (c) $t = 8$



(d) $t = 2$ (e) $t = 4$ (f) $t = 8$

Figure 3.2: Snapshots of the solution to the 3-D wave equation with $c = 1$, plotted along three 2-D slices of the 3-D field (when $x = 0$, $y = 0$, and $z = 0$). The initial conditions are: in the top row $f_0 = g_{\text{rc}}, g_0 = 0$ and in the bottom row $f_0 = 0, g_0 = g_{\text{rc}}$, where $g_{\text{rc}} = g_{\text{rc}}(\mathbf{x})$ is a raised cosine defined as $g_{\text{rc}} = \frac{1+\cos(\mathbf{x})}{2}$ for $|\mathbf{x}| \leq \pi$ and zero elsewhere. Here, h and k are chosen such that any underlying approximation error is not visibly apparent.

3.1.2 An equivalent staggered grid formulation

Another commonly used time-dependent finite difference scheme in the field of room acoustics stems from the equivalent first-order system (2.1), using staggered grids for pressure and vector velocity

[5] In the usual case, this denotes a Cartesian sampling of f_0, but it is not true in general that an initial condition will be bandlimited. For example, $f_0 = \delta(\mathbf{x})$ is not bandlimited.

grid function components. The essentials of this staggered scheme, which will be presented shortly, appeared in the acoustics-related literature around the same time, independently, from various authors [51, 38, 166, 120], but it is mostly known in room acoustics literature via Botteldooren's study [38]. For convenience, we make use of the formulation appearing in the paper by Chiba et al. [51], as their choice of staggered grids coincides directly with the second-order scheme in $\underline{\Psi}_{\boldsymbol{i}}^n$, although shortly we will also relate this to the more familiar staggered formulation of Botteldooren [38].[6]

Preliminaries

The following are some preliminaries that will be necessary (or just convenient) for deriving and presenting the staggered scheme of interest.

To start, we define the following staggered grid functions:

$$\underline{p}^{n+\frac{1}{2}}_{i_x,i_y,i_z} = \underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}} \qquad \approxeq p(\boldsymbol{i}h, (n+1/2)k) \tag{3.22a}$$

$$\underline{v}^n_{x,i_x+\frac{1}{2},i_y,i_z} = \underline{v}^n_{x,\boldsymbol{i}+\hat{\mathbf{e}}_x/2} \approxeq v_x((\boldsymbol{i}+\hat{\mathbf{e}}_x/2)h, nk) \tag{3.22b}$$

$$\underline{v}^n_{y,i_x,i_y+\frac{1}{2},i_z} = \underline{v}^n_{y,\boldsymbol{i}+\hat{\mathbf{e}}_y/2} \approxeq v_y((\boldsymbol{i}+\hat{\mathbf{e}}_y/2)h, nk) \tag{3.22c}$$

$$\underline{v}^n_{z,i_x,i_y,i_z+\frac{1}{2}} = \underline{v}^n_{z,\boldsymbol{i}+\hat{\mathbf{e}}_z/2} \approxeq v_z((\boldsymbol{i}+\hat{\mathbf{e}}_z/2)h, nk) \tag{3.22d}$$

where v_w is the scalar component of $\mathbf{v}$ along the coordinate direction $w \in \{x, y, z\}$.[7] This placement of staggered grid variables will be illustrated shortly (in Fig. 3.3(a)). For convenience, we also define the vector grid function:

$$\underline{\mathbf{v}}^n_{\boldsymbol{i}+\frac{1}{2}} := (\underline{v}^n_{x,\boldsymbol{i}+\hat{\mathbf{e}}_x/2}, \underline{v}^n_{y,\boldsymbol{i}+\hat{\mathbf{e}}_y/2}, \underline{v}^n_{z,\boldsymbol{i}+\hat{\mathbf{e}}_z/2})^{\mathrm{T}} \tag{3.23}$$

where the subscript of $\underline{\mathbf{v}}^n_{\boldsymbol{i}+\frac{1}{2}}$ is a slight abuse of notation that is only meant to serve as a reminder that its vector components are individually staggered in space.

We will also make use of the following asymmetric (non-centered) finite difference operators:

$$\delta_{t\pm} := \pm\frac{e_{t\pm} - 1}{k}, \qquad \delta_{t\pm}\Psi = \partial_t\Psi + \mathcal{O}(k) \tag{3.24a}$$

$$\delta_{w\pm} := \pm\frac{e_{w\pm} - 1}{h}, \quad \delta_{w\pm}\Psi = \partial_w\Psi + \mathcal{O}(h) \tag{3.24b}$$

which, as shown above, have a first-order accuracy with respect to the spatiotemporal points at which they are applied. δ_{w+} is also known as a forward-difference, and δ_{w-} is a backward-difference (and similarly for δ_{t+} and δ_{t-}).

For convenience, we also define the discrete gradient (vector) operator:

$$\boldsymbol{\delta}_{\mathbf{x}\pm} := (\delta_{x\pm}, \delta_{y\pm}, \delta_{z\pm})^{\mathrm{T}} \tag{3.25}$$

[6] We should note that the form presented by Botteldooren is also in a more general finite volume framework, which will be treated shortly. For now we consider the "basic Cartesian FDTD" in [38] rather than the finite volume "Voronoi cell FDTD" generalisation [38].

[7] N.B.: In this case, the subscripts x, y, z denote the scalar components of $\mathbf{v}$ and not partial derivatives.

and the following identities will be of use:

$$\delta_{tt} = \delta_{t+}\delta_{t-} \tag{3.26a}$$

$$\delta_{w\pm} = e_{w\pm}\delta_{w\mp} \tag{3.26b}$$

$$\delta_{ww} = \delta_{w+}\delta_{w-} \tag{3.26c}$$

$$\delta_{\Delta} = (\boldsymbol{\delta}_{\mathbf{x}-})^{\mathrm{T}}(\boldsymbol{\delta}_{\mathbf{x}+}) \tag{3.26d}$$

Also the following property of $\delta_{w\pm}$ will be of use:

$$\delta_{w-} = -\delta_{(-w)+} \tag{3.27}$$

where the notation above is akin to that in (3.6). This property deserves a short explanation; it says that the backward-difference along some coordinate direction is the negative of the forward-difference applied along the opposite direction. More generally, this property can be written as: $\delta_{w\mp} = -\delta_{(-w)\pm}$.

Finally, we also make use of the following half-shift operators, defined for an arbitrary grid function $\underline{u}_{\boldsymbol{i}}^{n} \approxeq u(\mathbf{x}_{\boldsymbol{i}}, t_n)$:

$$e_{t\pm}^{1/2}\underline{u}_{\boldsymbol{i}}^{n} := \underline{u}_{\boldsymbol{i}}^{n\pm 1/2} \qquad e_{t\pm}^{1/2}u = u + \mathcal{O}(k) \tag{3.28a}$$

$$e_{w\pm}^{1/2}\underline{u}_{\boldsymbol{i}}^{n} := \underline{u}_{\boldsymbol{i}\pm\hat{\mathbf{e}}_w/2}^{n} \qquad e_{w\pm}^{1/2}u = u + \mathcal{O}(h)\,, \quad w \in \{x, y, z\} \tag{3.28b}$$

and clearly, $\left(e_{w\pm}^{1/2}\right)^2 = e_{w\pm}$, and $\left(e_{t\pm}^{1/2}\right)^2 = e_{t\pm}$. It is important to point out that the seemingly first-order accurate difference operators defined above are second-order accurate with respect to shifted positions:

$$\delta_{t\pm}u = \partial_t\left(e_{t\mp}^{1/2}u\right) + \mathcal{O}(k^2) \tag{3.29a}$$

$$\delta_{w\pm}u = \partial_w\left(e_{w\mp}^{1/2}u\right) + \mathcal{O}(h^2) \tag{3.29b}$$

This second-order accuracy about a half-shifted position is the underlying motivation behind the spatial and temporal staggering of vector-component grid functions in (3.22).

A staggered grid scheme for the first-order system of linear acoustics

At this point we can present a staggered-grid scheme for the first order system (2.1), after [51]. This scheme can be written compactly as follows:

$$\delta_{t-}\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}} = -\kappa(\boldsymbol{\delta}_{\mathbf{x}-})^{\mathrm{T}}\underline{\mathbf{v}}_{\boldsymbol{i}+\frac{1}{2}}^{n} \tag{3.30a}$$

$$\rho\delta_{t+}\underline{\mathbf{v}}_{\boldsymbol{i}+\frac{1}{2}}^{n} = -\boldsymbol{\delta}_{\mathbf{x}+}\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}} \tag{3.30b}$$

where $\kappa = \rho c^2$ is the bulk modulus. The operation of this discrete first-order system is given by the following rearrangement of (3.30):

$$\underline{p}^{n+\frac{1}{2}}_{i_x,i_y,i_z} = \underline{p}^{n-\frac{1}{2}}_{i_x,i_y,i_z} - \lambda Z_0 \Big(\underline{v}^n_{x,i_x+\frac{1}{2},i_y,i_z} + \underline{v}^n_{y,i_x,i_y+\frac{1}{2},i_z} + \underline{v}^n_{z,i_x,i_y,i_z+\frac{1}{2}} - \underline{v}^n_{x,i_x-\frac{1}{2},i_y,i_z} - \underline{v}^n_{y,i_x,i_y-\frac{1}{2},i_z} - \underline{v}^n_{z,i_x,i_y,i_z-\frac{1}{2}} \Big) \tag{3.31a}$$

$$\underline{v}^{n+1}_{x,i_x+\frac{1}{2},i_y,i_z} = \underline{v}^n_{x,i_x+\frac{1}{2},i_y,i_z} - \lambda Y_0 \left(\underline{p}^{n+\frac{1}{2}}_{i_x+1,i_y,i_z} - \underline{p}^{n+\frac{1}{2}}_{i_x,i_y,i_z} \right) \tag{3.31b}$$

$$\underline{v}^{n+1}_{y,i_x,i_y+\frac{1}{2},i_z} = \underline{v}^n_{y,i_x,i_y+\frac{1}{2},i_z} - \lambda Y_0 \left(\underline{p}^{n+\frac{1}{2}}_{i_x,i_y+1,i_z} - \underline{p}^{n+\frac{1}{2}}_{i_x,i_y,i_z} \right) \tag{3.31c}$$

$$\underline{v}^{n+1}_{z,i_x,i_y,i_z+\frac{1}{2}} = \underline{v}^n_{z,i_x,i_y,i_z+\frac{1}{2}} - \lambda Y_0 \left(\underline{p}^{n+\frac{1}{2}}_{i_x,i_y,i_z+1} - \underline{p}^{n+\frac{1}{2}}_{i_x,i_y,i_z} \right) \tag{3.31d}$$

where we recall that $Z_0 = \rho c$ and $Y_0 = (\rho c)^{-1}$. From these update equations, we can recover the more commonly seen formulation presented by Botteldooren [37, 38] simply by applying $e^{1/2}_{t-}$ throughout (3.31). For completeness, Botteldooren's scheme [37] can be written as follows:

$$\rho \delta_{t-} \underline{\mathbf{v}}^{n+1/2}_{\boldsymbol{i}+\frac{1}{2}} = -\boldsymbol{\delta}_{\mathbf{x}+} \underline{p}^n_{\boldsymbol{i}} \tag{3.32a}$$

$$\delta_{t+} \underline{p}^n_{\boldsymbol{i}} = -\kappa (\boldsymbol{\delta}_{\mathbf{x}-})^{\mathrm{T}} \underline{\mathbf{v}}^{n+1/2}_{\boldsymbol{i}+\frac{1}{2}} \tag{3.32b}$$

where the ordering of the above two equations reflects how they are sequentially carried out in practice, as per [38].

Relation to second-order scheme

The first-order staggered scheme (3.30) can be made equivalent to the second-order scheme in $\underline{\Psi}^n_{\boldsymbol{i}}$ through the following relations, which are simple discretisations of (2.6):

$$\underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}} = \rho \delta_{t+} \underline{\Psi}^n_{\boldsymbol{i}} \tag{3.33a}$$

$$\underline{\mathbf{v}}^n_{\boldsymbol{i}+\frac{1}{2}} = -\boldsymbol{\delta}_{\mathbf{x}+} \underline{\Psi}^n_{\boldsymbol{i}} \tag{3.33b}$$

By substituting (3.33) into (3.30a), we recover a second-order scheme:

$$\rho \delta_{t+} \delta_{t-} \underline{\Psi}^n_{\boldsymbol{i}} = \kappa (\boldsymbol{\delta}_{\mathbf{x}-})^{\mathrm{T}} (\boldsymbol{\delta}_{\mathbf{x}+} \underline{\Psi}^n_{\boldsymbol{i}}) \tag{3.34}$$

which is equivalent to (3.15) by the identities (3.26a) and (3.26d) and since $c^2 = \kappa/\rho$.

Initial conditions for this scheme can also be related, through (3.33), to those of the second-order scheme:

$$\Psi = f_0, \Psi_t = g_0 \Rightarrow \qquad p = \rho g_0, \qquad \mathbf{v} = -\nabla f_0, \qquad t = 0 \tag{3.35}$$

$$\Rightarrow \qquad \underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}} = \rho [g_0]_h, \qquad \underline{\mathbf{v}}^n_{\boldsymbol{i}+\frac{1}{2}} = -\boldsymbol{\delta}_{\mathbf{x}+} [f_0]_h, \qquad n = 0 \tag{3.36}$$

The above initialisation of the first-order scheme is followed by the velocity component update for $n = 0$ (to calculate $\underline{\mathbf{v}}^1_{\boldsymbol{i}+\frac{1}{2}}$), then one can proceed with (3.31) starting from $n = 1$.

Cartesian finite volume form

The staggered scheme also has a special relationship with the basic form of a Cartesian *finite volume* method for the first-order system [37]. To establish this, we will only show that the basic staggered

scheme approximates an integral form of the equation of continuity,[8] saving a detailed and more general finite volume derivation for later (Chapter 6). As such, we rewrite the discrete system as:

$$h^3\delta_{t+}\underline{p}^{n+\frac{1}{2}}_{i_x,i_y,i_z} = -\kappa\Big(h^2\underline{v}^n_{x,i_x+\frac{1}{2},i_y,i_z} + h^2\underline{v}^n_{y,i_x,i_y+\frac{1}{2},i_z} + h^2\underline{v}^n_{z,i_x,i_y,i_z+\frac{1}{2}} + h^2\underline{v}^n_{(-x),i_x-\frac{1}{2},i_y,i_z} + h^2\underline{v}^n_{(-y),i_x,i_y-\frac{1}{2},i_z} + h^2\underline{v}^n_{(-z),i_x,i_y,i_z-\frac{1}{2}}\Big) \tag{3.37a}$$

$$\rho\delta_{t-}\underline{v}^n_{(\pm w),\boldsymbol{i}\pm\hat{\mathbf{e}}_w/2} = -\delta_{w\pm}\underline{p}^{n+\frac{1}{2}}_{i_x,i_y,i_z}\,, \quad w\in\{x,y,z\} \tag{3.37b}$$

where we have reversed half of the vector components' directions by changing their signs, i.e.:

$$\underline{v}^n_{(-w),\boldsymbol{i}-\hat{\mathbf{e}}_w/2} = -\underline{v}^n_{w,\boldsymbol{i}-\hat{\mathbf{e}}_w/2} \tag{3.38}$$

This is essentially the same property as (3.27).

Now let us present the equation of continuity in divergence form over a *cubic* cell $\mathcal{C}$ with volume h^3 and centered about the position $\mathbf{x} = \boldsymbol{i}h$. This can be written as:

$$\int_{\mathcal{C}}\partial_t p\,\mathrm{d}\mathbf{x} = -\kappa\Big(\int_{\mathcal{S}_{x+}} v_{x+}\,\mathrm{d}\sigma + \int_{\mathcal{S}_{y+}} v_{y+} + \int_{\mathcal{S}_{z+}} v_{z+}\,\mathrm{d}\sigma\,\mathrm{d}\sigma + \int_{\mathcal{S}_{x-}} v_{x-}\,\mathrm{d}\sigma + \int_{\mathcal{S}_{y-}} v_{y-}\,\mathrm{d}\sigma + \int_{\mathcal{S}_{z-}} v_{z-}\,\mathrm{d}\sigma + \Big) \tag{3.39}$$

Here, $\mathcal{S}_{x\pm}$, $\mathcal{S}_{y\pm}$, and $\mathcal{S}_{z\pm}$ represent the six square faces of the cubic cell (with surface areas h^2) with respective outward normals $\pm\hat{\mathbf{e}}_x$, $\pm\hat{\mathbf{e}}_y$, and $\pm\hat{\mathbf{e}}_z$; and $v_{x\pm}$, $v_{y\pm}$, and $v_{z\pm}$ represent the outward normal velocity components associated to each respective face. It can then be realised that (3.37a) is a discrete form of (3.39), where spatial averages are applied over the cell and its faces.

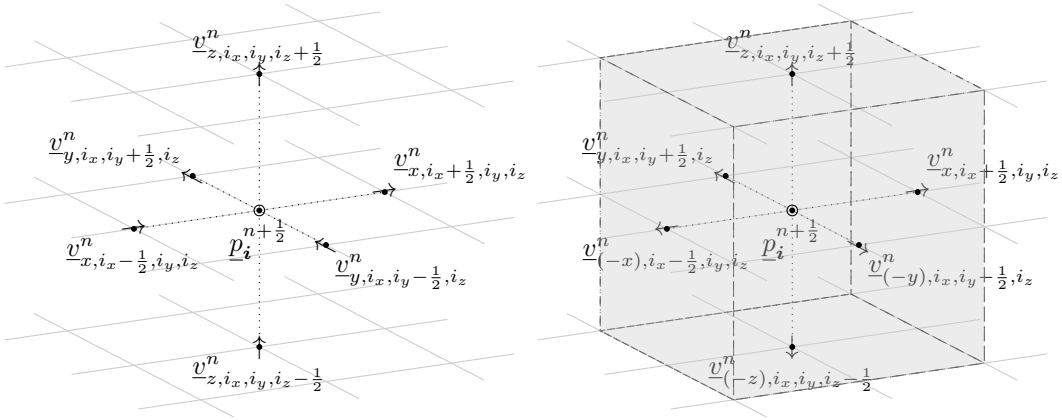
The second finite difference equation appearing above, (3.37b), is simply a discretisation of the equation of motion in terms of the same normal outward vectors, of which there are six per cell; however, only three need to be computed per index $\boldsymbol{i}$, as in the standard FDTD (finite difference time domain) scheme. The placements of discrete field components and the orientations of vector components in both the basic Cartesian "FDTD" (3.31) and "FVTD" (finite volume time domain) interpretations (3.37) are illustrated in Fig. 3.3.

Transitional note on convergence and stability

Thus far, we have not specified a value (or a range of values) for the Courant number λ. Next we will conduct an energy analysis, which will provide some insight into appropriate choices of λ, but first it is worth briefly discussing some implications of the Courant number. The choice of the Courant number is inextricably linked to the critical topics of convergence and stability in the study of finite difference schemes.

With regards to the topic of convergence, this finite difference scheme in $\underline{\Psi}^n_{\boldsymbol{i}}$ is said to be pointwise convergent if $|\underline{\Psi}^n_{\boldsymbol{i}} - \Psi(\boldsymbol{i}h, nk)| \to 0$ in the limit of $h, k \to 0$. This scheme could also be said to be globally convergent if $\|\underline{\Psi}^n - \Psi(\cdot, nk)\| \to 0$ in the limit of $h, k \to 0$, where $\|\cdot\|$ is an appropriate norm over the spatial grid. Such convergence is an important property because it provides a basic level of confidence in the numerical solution—if the scheme is convergent, it will reveal a unique (true) solution in the limit of a small grid spacing. However, convergence does not follow from consistency alone. A

[8] The update equations for the velocity components could be similarly shown to satisfy an integral form of the equation of motion [37], but it is simpler to just consider them as a finite difference approximation [26].



(a) Cartesian FDTD (b) Cartesian FVTD

Figure 3.3: Illustration of staggered pressure and velocity vector components in FDTD and FVTD forms. Note the directions of arrows denoting velocity components.

classic result in the study of numerical methods is a *necessary* condition for (pointwise) convergence derived by Courant et al., commonly known as the "Courant-Friedrichs-Lewy (CFL) condition" [60]. The CFL condition essentially says that the domain of dependence of the numerical solution must contain the domain of dependence of the underlying continuous solution (see, e.g., [268] for more detail). For this simple scheme, the CFL condition leads to a constraint on the Courant number (hence the name), and this constraint happens to be:

$$\lambda \leq \sqrt{1/3} \tag{3.40}$$

Numerical stability is another crucial feature of any numerical method. It is required to ensure that the numerical solution remains bounded as time marches forward.[9] Also, with numerical stability, it can be assured, by the Lax equivalence theorem [207, 260], that the scheme is *convergent*, as long as the scheme is consistent with the model problem. Generally, numerical stability imposes its own constraint on the Courant number, which is not necessarily the same as that imposed by the CFL condition.[10] For free-space problems, one can apply a frequency-domain technique known as von Neumann analysis [185], which will be seen later, to derive a necessary and sufficient condition for stability. In order to derive a stability condition for this scheme, we will make use of energy methods, which are, additionally, useful in deriving boundary conditions that are stable.

3.1.3 Energy stability analysis

Energy methods are useful techniques for the design of stable finite difference schemes [60, 148, 207, 91, 24]. The principle of energy methods is to derive an energetic quantity for the approximation in order to establish a discrete analogue of the conservation of energy principle. Stability for the system results from conditions that ensure that the energy is non-negative and non-increasing. The analysis follows closely the energy analysis used in the continuous case (as in Chapter 2).

[9] More specifically, the numerical solution should remain bounded *appropriately for the model equation* for any valid set of (bounded) initial conditions.

[10] The CFL condition is a necessary condition for *convergence*, but it is not necessarily sufficient for stability [268].

Preliminaries

In order to proceed with an energy analysis, following [24], we define the following finite difference and averaging operators, and we give their associated Taylor expansions after acting on continuous functions:

$$\delta_{t\cdot} := \frac{e_{t+} - e_{t-}}{2k}, \qquad \delta_{t\cdot}\Psi = \partial_t \Psi + \mathcal{O}(k^2) \tag{3.41a}$$

$$\mu_{t\pm} := \frac{e_{t\pm} + 1}{2}, \qquad \mu_{t\pm}\Psi = \Psi + \mathcal{O}(k) \tag{3.41b}$$

$$\mu_{tt} := \mu_{t+}\mu_{t-}, \qquad \mu_{tt}\Psi = \Psi + \mathcal{O}(k^2) \tag{3.41c}$$

$$\delta_{w\cdot} := \frac{e_{w+} - e_{w-}}{2h}, \qquad \delta_{w\cdot}\Psi = \partial_w \Psi + \mathcal{O}(h^2) \tag{3.41d}$$

$$\mu_{w\pm} := \frac{e_{w\pm} + 1}{2}, \qquad \mu_{w\pm}\Psi = \Psi + \mathcal{O}(h) \tag{3.41e}$$

where $w \in \{x, y, z\}$ for (3.41d) and (3.41e). We also define the following inner product on $\mathbb{Z}_h^3$:

$$\langle \underline{u}^n, \underline{v}^n \rangle_{\mathbb{Z}_h^3} := \sum_{\boldsymbol{i} \in \mathbb{Z}^3} h^3 (\underline{u}_{\boldsymbol{i}}^n)(\underline{v}_{\boldsymbol{i}}^n), \quad \|\underline{u}^n\|_{\mathbb{Z}_h^3}^2 := \langle \underline{u}^n, \underline{u}^n \rangle_{\mathbb{Z}_h^3} \tag{3.42}$$

where $\underline{u}_{\boldsymbol{i}}^n$ and $\underline{v}_{\boldsymbol{i}}^n$ are arbitrary scalar grid functions on $\mathbb{Z}_h^3$. Note that $\|\cdot\|_{\mathbb{Z}_h^3}$ is an ℓ^2-norm on $\mathbb{Z}_h^3$.

The following identities will also be useful [24]:

$$\delta_{t+}\left(\delta_{t-}\underline{u}_{\boldsymbol{i}}^n\right)^2 = 2(\delta_{t\cdot}\underline{u}_{\boldsymbol{i}}^n)(\delta_{tt}\underline{u}_{\boldsymbol{i}}^n) \tag{3.43a}$$

$$\delta_{t+}(\underline{u}_{\boldsymbol{i}}^n e_{t-}\underline{u}_{\boldsymbol{i}}^n) = 2(\underline{u}_{\boldsymbol{i}}^n)(\delta_{t\cdot}\underline{u}_{\boldsymbol{i}}^n) \tag{3.43b}$$

$$\delta_{t+}(\mu_{t-}\underline{u}_{\boldsymbol{i}}^n)^2 = 2(\delta_{t\cdot}\underline{u}_{\boldsymbol{i}}^n)(\mu_{tt}\underline{u}_{\boldsymbol{i}}^n) \tag{3.43c}$$

$$\delta_{t+}(\underline{u}_{\boldsymbol{i}}^n)^2 = 2(\mu_{t+}\underline{u}_{\boldsymbol{i}}^n)(\delta_{t+}\underline{u}_{\boldsymbol{i}}^n) \tag{3.43d}$$

$$\underline{u}_{\boldsymbol{i}}^n e_{t-}\underline{u}_{\boldsymbol{i}}^n = (\mu_{t-}\underline{u}_{\boldsymbol{i}}^n)^2 - \frac{k^2}{4}(\delta_{t-}\underline{u}_{\boldsymbol{i}}^n)^2 \tag{3.43e}$$

where the first four identities are discrete versions of the chain rule of differentiation. We also make use of the following summation-by-parts (SBP) property on the full-space:

$$\langle \underline{u}^n, \delta_{w-}\underline{v}^n \rangle_{\mathbb{Z}_h^3} = -\langle \delta_{w+}\underline{u}^n, \underline{v}^n \rangle_{\mathbb{Z}_h^3}, \quad w \in \{x, y, z\} \tag{3.44}$$

Energy analysis

The energy analysis for the discrete system follows closely that of the continuous system in Section 2.1.1. We multiply both sides of (3.15) by $\frac{\rho}{c^2}\delta_{t\cdot}\underline{\Psi}_{\boldsymbol{i}}^n$, and take the inner product (3.42) to obtain:

$$\frac{\rho}{c^2}\langle \delta_{t\cdot}\underline{\Psi}, \delta_{tt}\underline{\Psi} \rangle_{\mathbb{Z}_h^3} = \sum_{w \in \{x,y,z\}} \rho \langle \delta_{t\cdot}\underline{\Psi}, \delta_{ww}\underline{\Psi} \rangle_{\mathbb{Z}_h^3} \tag{3.45}$$

where we have dropped the superscript n for now. Making use of identity (3.26c) and the SBP property (3.44), we can rewrite (3.45) as:

$$\frac{\rho}{c^2}\langle \delta_{t\cdot}\underline{\Psi}, \delta_{tt}\underline{\Psi} \rangle_{\mathbb{Z}_h^3} + \sum_{w \in \{x,y,z\}} \rho \langle \delta_{t\cdot}\delta_{w+}\underline{\Psi}, \delta_{w+}\underline{\Psi} \rangle_{\mathbb{Z}_h^3} = 0 \tag{3.46}$$

where we made use of identity (3.26b).

Next, we use identities (3.43a) and (3.43b) to arrive at an energy balance:

$$\delta_{t+}\mathfrak{h} = 0\,, \quad \mathfrak{h} = \mathfrak{v} + \mathfrak{t} \tag{3.47}$$

where

$$\mathfrak{v} = \frac{\rho}{2c^2}\|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{Z}^3_h}\,, \qquad \mathfrak{t} = \frac{\rho}{2}\sum_{w\in\{x,y,z\}}\langle\delta_{w+}\underline{\Psi}, e_{t-}\delta_{w+}\underline{\Psi}\rangle_{\mathbb{Z}^3_h} \tag{3.48}$$

We note that $\mathfrak{h}$, $\mathfrak{v}$, and $\mathfrak{t}$ respectively represent the discrete analogues to the Hamiltonian, and potential energy and kinetic energies in (2.5). Furthermore, (3.47) shows that the total discrete energy is conserved over time. In other words, we have that:

$$\mathfrak{h}^n \equiv \mathfrak{h}^0 \tag{3.49}$$

It remains to find conditions such that $\mathfrak{h} \geq 0$ (for stability). Making use of identity (3.43e), we rewrite the quantity $\mathfrak{t}$ as:

$$\mathfrak{t} = \frac{\rho}{2}\sum_{w\in\{x,y,z\}}\left(\|\mu_{t-}\delta_{w+}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} - \frac{k^2}{4}\|\delta_{t-}\delta_{w+}\underline{\Psi}\|^2_{\mathbb{Z}^3_h}\right) \tag{3.50}$$

Then we have that:

$$\mathfrak{t} \geq -\frac{\rho k^2}{8}\sum_{w\in\{x,y,z\}}\|\delta_{t-}\delta_{w+}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \tag{3.51}$$

Next, we can make use of the following inequality [24]:[11]

$$\|\delta_{w\pm}\underline{u}\|^2_{\mathbb{Z}^3_h} \leq \frac{4}{h^2}\|\underline{u}\|^2_{\mathbb{Z}^3_h} \tag{3.52}$$

to arrive at:

$$\mathfrak{t} \geq -\frac{\rho k^2}{2h^2}\sum_{w\in\{x,y,z\}}\|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \geq -\frac{3\rho k^2}{2h^2}\|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \tag{3.53}$$

This leads to a bound on $\mathfrak{h}$:

$$\mathfrak{h} \geq \frac{\rho}{2c^2}\left(1 - 3\lambda^2\right)\|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \tag{3.54}$$

which provides a condition on the Courant number:

$$\lambda \leq \frac{1}{\sqrt{3}} \quad \implies \quad \mathfrak{h} \geq 0 \tag{3.55}$$

This condition is also, not surprisingly, the CFL condition for convergence (3.40). Provided that the above condition is satisfied, the scheme will be lossless and stable, in the sense that the energy is non-increasing. It follows that the staggered grid scheme has the same stability condition, as it is equivalent to this scheme in $\underline{p}$ and $\underline{\mathbf{v}}$.

It is important to note that while the energy remains bounded (when initial conditions are in L^2), the solution is still permitted to have some growth, which is allowed because the wave equation itself

[11] This inequality follows from the fact that for any $a, b \in \mathbb{R}$, $(a+b)^2 \geq 0 \implies (a-b)^2 \leq 2(a^2+b^2)$, and the additional factor of two follows from pairwise-symmetry in the summation, i.e., $(a-b)^2 = (b-a)^2$.

permits linearly growing solutions.[12] Using $\lambda \leq \sqrt{1/3}$ and (3.55) we have that:

$$\mathfrak{h}^0 \geq \mathfrak{h}^n \geq \frac{\rho}{2c^2}\left(1-3\lambda^2\right)\|\delta_{t-}\underline{\Psi}^n\|^2_{\mathbb{Z}^3_h} \quad \implies \quad \sqrt{\frac{2\mathfrak{h}^0}{1-3\lambda^2}} \geq \frac{\rho}{c}\|\delta_{t-}\underline{\Psi}^n\|_{\mathbb{Z}^3_h} \geq \frac{\rho}{c}\delta_{t-}\|\underline{\Psi}^n\|_{\mathbb{Z}^3_h} \tag{3.56}$$

$$\implies \quad \|\underline{\Psi}^n\|_{\mathbb{Z}^3_h} \leq \frac{cnk}{\rho}\sqrt{\frac{2\mathfrak{h}^0}{1-3\lambda^2}} \tag{3.57}$$

which shows that the norm of the numerical solution grows at most linearly over time.[13]

A second-order wave equation scheme in pressure

One could equally derive a second-order scheme in the pressure variable directly from (3.15) and (3.33):

$$\delta_{tt}\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}} = c^2\delta_{\Delta}\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}} \tag{3.58}$$

which is subsequently equivalent to the staggered scheme. It is more common to see this scheme shifted by $e_{t-}^{1/2}$:

$$\delta_{tt}\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}} = c^2\delta_{\Delta}\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}} \qquad \overset{e_{t-}^{1/2}}{\implies} \qquad \delta_{tt}\underline{p}_{\boldsymbol{i}}^{n} = c^2\delta_{\Delta}\underline{p}_{\boldsymbol{i}}^{n} \tag{3.59}$$

which is subsequently equivalent to the staggered formulation of Botteldooren,[14] since it is related to the formulation provided here through the same application of $e_{t-}^{1/2}$.

The stability condition of this second-order scheme is unchanged from (3.15). However, it is important to note that while linear growth in the acoustic velocity potential is valid from a physical point of view (total energy remains bounded), linear growth in pressure—while permitted in the wave equation itself—leads to a physical energy that is unbounded (for initial conditions in L^2) and is thus not appropriate from a physical point of view. Nevertheless, such growth can be prevented by setting the initial condition: $p_t(\mathbf{x}, 0) = 0$.

Remarks on the choice of discrete system

So far we have presented three choices of finite difference discrete models that could be (and have been) used for room acoustics simulations: the second-order scheme in velocity potential (3.15), the staggered scheme in terms of pressure and vector particle velocity (3.30) (or (3.32)), and the second-order scheme in pressure (3.58). In the room acoustics literature, it has been more common to make use of the staggered formulation derived by Botteldooren [38] or the pressure wave equation scheme [223, 182] for time-domain finite difference simulations. The use of the second-order scheme in the acoustic velocity potential, which is the primary model in this study, is rather recent in the finite difference room acoustics literature [302], although it has been used for some time in the geophysics literature [2]. In terms of a *sound pressure output* (one of the ultimate aims of such simulations), the three choices of finite difference systems will lead to equivalent outputs, provided that any boundary and initial conditions are also related through (3.33). Any of these formulations are thus sufficient for standard room acoustics simulations, but there are slight differences between the choices (3.15), (3.31), or (3.58) that are worth discussing further.

[12] One such example is given later, in Footnote 24, see also [23, Section 4.3.4] and [260] for more discussion on this.

[13] This bound is only valid for $\lambda < \sqrt{1/3}$, but it may not be possible to obtain a bound at $\lambda = \sqrt{1/3}$, see [24, Section 2.4.3].

[14] This relationship can also be established without the use of $\underline{\Psi}_{\boldsymbol{i}}^n$; see, e.g., [21, 43].

With regards to the derivation of boundary conditions, the first-order system and associated staggered scheme is the most natural choice, due to the fact that immittance (real or complex) boundary conditions are generally formulated in terms of p and $\mathbf{v}$. The staggered formulation also leads to more control over the placement of outward normal boundary velocity components, and has an underlying finite volume interpretation. After the first-order system, the second-order formulation in terms of Ψ is perhaps the next most natural choice for the specification of boundary conditions, since boundary conditions in p and $\mathbf{v}$ are equivalently expressed in terms of Ψ_t and $\nabla\Psi$, and staggered grid/finite volume formulations in these quantities are, therefore, equivalently derived. With regards to the pressure wave equation model, boundary conditions are less often expressed in terms of the pressure variable alone, but such conditions can be obtained from the underlying first-order system (or velocity potential system) and can then be re-expressed in pressure-only forms for the second-order pressure scheme [132].

In terms of computational efficiency, if both the pressure and velocity fields are desired everywhere, it is cheapest (computationally) to make use of the first-order scheme directly. If only the pressure is desired, as in standard room acoustics applications, the second-order forms are clearly more efficient in terms of memory and arithmetic operations. For example, in the first-order scheme, one must store four states per grid point $\boldsymbol{i}$, which is twice the storage requirement of the second-order scheme. Also, operation counts are higher in the case of the first-order scheme.[15] It should also be mentioned that the second-order scheme in Ψ requires an additional differentiation at the output to obtain a sound pressure, but the cost of this operation is negligible compared to the overall cost of the scheme since the number of outputs taken from the scheme is generally much less than the total number of grid points. We also note that one is not restricted to using δ_{t+} in order to obtain the pressure, and more accurate temporal difference operators could be employed to obtain a pressure output.

For more general virtual and musical acoustics applications (e.g., physical modelling sound synthesis), the first-order scheme and second-order scheme in Ψ are appropriate choices, as they provide straightforward direct or indirect access to both pressure and velocity field components, which are ultimately required as coupling conditions for the inclusion of other sound-making objects (e.g., physically-modelled musical instruments [24, 33]). In such settings, the pressure wave equation model is the least flexible because one does not have easy access to the velocity components. Velocity components can be estimated from the pressure field alone, as a cumulative sum of pressure gradients over time [43], but this is less straightforward and can only be accurate to within a constant. Another option, provided initial conditions are zero, is to carry out a hybrid scheme, incorporating both the second-order pressure scheme and staggered first-order scheme [43], but it is simpler still to use $\underline{\Psi}_{\boldsymbol{i}}^n$ directly.

Finally, we should note that while the first-order scheme carries a greater computational cost, the fact that it does not permit linear growth from a bounded set of initial conditions (as opposed to second-order schemes) is a desirable feature from a practical standpoint.

3.1.4 Numerical dispersion

Numerical dispersion is a critical issue in the design of finite difference schemes for room acoustics applications. It is essentially a warping of frequencies that can lead to a mistuning of room modes and unnatural modal densities. Additionally, it gives rise to a direction- and frequency-dependent wave speed, which in turn gives rise to audible artefacts, such as a smearing of transients into "chirp"-like sounds and other phase artefacts. In this section, we present an introductory analysis of numerical dispersion in the simplest 3-D scheme.

[15] Operation counts are left out for brevity, see, e.g., [43, 304].

As discussed in Chapter 2, the wave equation admits plane-wave solutions that adhere to a linear dispersion relation, and as such, the wave equation can be said to be non-dispersive. Finite difference schemes for the wave equation also admit plane-wave solutions, but due to approximation errors, these plane-wave solutions follow a (generally) non-linear dispersion relation. In order to present the dispersion relation of the scheme, we consider a test solution (ansatz) made up of a discrete plane wave:

$$\underline{\Psi}_{\boldsymbol{i}}^{n} = \mathrm{e}^{\mathrm{j}(\omega nk + \boldsymbol{\beta}\cdot\boldsymbol{i}h)} \tag{3.60}$$

where ω represents the temporal angular frequency (in rad/s) and $\boldsymbol{\beta} = (\beta_x, \beta_y, \beta_z)$ is the wave vector. These are the symbols as used in the continuous case, but the relation between ω and $\boldsymbol{\beta}$ (the dispersion relation) is left unspecified in this discrete setting.

Inserting this trial solution into the finite difference scheme (3.15), we obtain a dispersion relation for the 3-D scheme:[16]

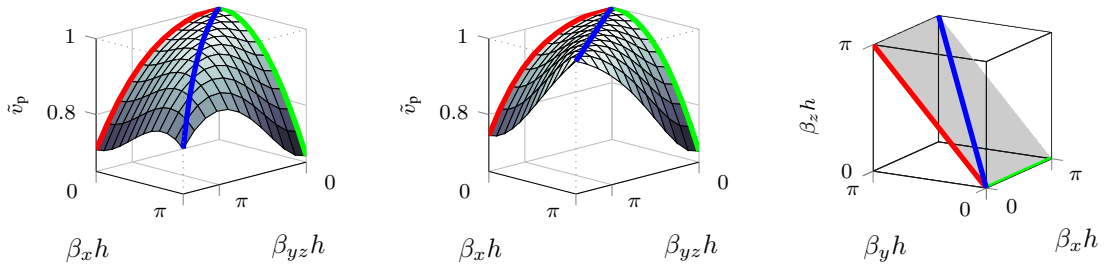
$$\omega = \frac{2}{k}\sin^{-1}\left(\pm\lambda\Big(\underbrace{\sin^2(\boldsymbol{\beta}h\cdot\hat{\mathbf{e}}_x/2) + \sin^2(\boldsymbol{\beta}h\cdot\hat{\mathbf{e}}_y/2) + \sin^2(\boldsymbol{\beta}h\cdot\hat{\mathbf{e}}_z/2)}_{G(\boldsymbol{\beta},h)} \Big)^{1/2} \right) \tag{3.61}$$

That this dispersion relation is generally nonlinear indicates that the scheme features *numerical dispersion* (or *dispersion error*). Note also that in the limit of small $h, k \to 0$ the sine function becomes linear (small-angle approximation) and the numerical dispersion relation approaches that of the wave equation, as expected by consistency.

From this dispersion relation we can point out a few more important properties of the scheme. One such detail is that the function labelled $G(\boldsymbol{\beta}, h)$ is periodic in the cubic wavenumber cell $\boldsymbol{\beta} \in [-\pi/h, \pi/h]^3$. As such, the finite difference scheme can only resolve wavenumbers in this cell, which also follows from multidimensional sampling theory [191]. It can also be noted that $G(\boldsymbol{\beta}, h)$ takes on its maxima at the corners of this cubic region ($\max_{\boldsymbol{\beta},h} G(\boldsymbol{\beta}, h) = 3$) and has a minimum value of zero. A von Neumann stability analysis can be easily conducted from this point, by constraining $\Im\{\omega\} = 0$ in (3.61), one eventually obtains to the stability condition:[17] $\lambda \leq \sqrt{1/3}$, which is the same condition derived from energy methods (and the CFL condition for convergence). Provided that this constraint is satisfied, it can be determined from the properties of the inverse sine function that $\omega \in [-\pi/k, \pi/k]$, which is the same temporal bandwidth constraint imposed by sampling considerations. However, contrary to sampling considerations, it also follows from (3.61) that the full bandwidth of temporal frequencies *cannot* be reproduced unless $\lambda = \sqrt{1/3}$; if $\lambda < \sqrt{1/3}$ then ω can only be in some subset of $[-\pi/k, \pi/k]$, and this band of frequencies shrinks as $\lambda \to 0$. This phenomenon will be investigated in more detail in Chapter 4.

[16] The first-order scheme has an equivalent dispersion relation [38], and the same dispersion relation is found in Yee's FDTD scheme [265], and staggered schemes for first-order elastodynamic systems [234].

[17] For brevity, the steps are left out here, but the same analysis is carried out in detail in Chapters 4 and 5.



(a) $\lambda = 0.5$ (b) $\lambda = \sqrt{1/3} \approx 0.577$ (c) side-diagonal 2-D slice

Figure 3.4: Surface plots of $\tilde{v}_\mathrm{p}$, for two values of λ along the 2-D plane: $\beta_y = \beta_z$, and we define: $\beta_{yz} = (\beta_y + \beta_z)/\sqrt{2}$. This 2-D plane is intersected with the 3-D cubic region: $\boldsymbol{\beta}_h \in [0, \pi]^3$ to give the 2-D slice shaded in grey and shown in (c). The green line denotes an axial direction ($\beta_y = \beta_z = 0$), the red line denotes a side-diagonal direction ($\beta_x = 0, \beta_y = \beta_z$), and the blue line denotes a space-diagonal direction ($\beta_x = \beta_y = \beta_z$).

Numerical phase velocity

In the room acoustics literature, numerical dispersion is usually analysed from the point of view of the numerical phase velocity.[18] A *relative* numerical phase velocity, ideally unity, can be defined as:

$$\tilde{v}_\mathrm{p}(\boldsymbol{\beta}, \lambda, h) = \frac{\omega(\boldsymbol{\beta}, \lambda, h, k)}{c|\boldsymbol{\beta}|} \tag{3.62}$$

where $\omega(\boldsymbol{\beta}, \lambda, h, k)$ is simply the dispersion relation (3.61).[19]

We can analyse $\tilde{v}_\mathrm{p}$ to get a better idea of how numerical dispersion is manifested in the numerical solution. However, there are many ways to visually analyse $\tilde{v}_\mathrm{p}$, which is essentially a function of five variables: h and λ and $\boldsymbol{\beta} = (\beta_x, \beta_y, \beta_z)$ in Cartesian coordinates, or $\boldsymbol{\beta} = (\beta, \phi(\boldsymbol{\beta}), \theta(\boldsymbol{\beta}))$ in spherical coordinates. The time-step k is left out because it can be fixed by choosing λ and h (c is fixed). To further simplify this parameter space, we can use the fact that $G(\boldsymbol{\beta}, h)$ features a large degree of symmetry and is periodic over a cubic wavenumber cell tiling. With that said, it will suffice to analyse $\tilde{v}_\mathrm{p}$ over a 2-D slice of the non-negative octant of the cubic wavenumber cell. This 2-D slice is oriented along the plane with normal vector $\hat{\mathbf{n}} = (0, 1/\sqrt{2}, -1/\sqrt{2})$ and passes through the origin (for which $\beta_y = \beta_z$), as shown in Fig. 3.4(c). For the other free variable, λ, we consider two choices: $\lambda = \sqrt{1/3} \approx 0.577$ and $\lambda = 0.5$; these are shown in Figs. 3.4(a) and 3.4(b).

Let us first compare the two choices of λ. From Figs. 3.4(a) and 3.4(b), we can see that $\tilde{v}_\mathrm{p}$ is generally more correct (closer to unity) over the wavenumbers displayed for $\lambda = \sqrt{1/3}$ than it is for $\lambda = 0.5$. This follows a trend that will be seen for many (but not all) schemes in Chapter 4, that $\tilde{v}_\mathrm{p}$ gets worse as λ is chosen away from the stability limit. Consequently, $\lambda = \sqrt{1/3}$ can be seen as the optimal choice of the Courant number for this scheme (keeping in mind the constraints of stability).

Next, we can investigate $\tilde{v}_\mathrm{p}$ as a function of wavenumber (magnitude). In this respect, it can be noted, in either Fig. 3.4(a) or 3.4(b), that the phase velocity is ideal near the zero wavenumber (DC) and tends to incur more error as the wavenumber increases along any given direction. This is expected by consistency with the model system.

Finally, we can compare the various wave directions for fixed wavenumbers, i.e., directions of prop-

[18] Strictly speaking, the term "phase velocity" is a bit of a misnomer, and this quantity should, more appropriately, be called a phase *speed*, since a velocity must also carry a direction. However, it is common in the literature to call this quantity the "phase velocity".

[19] Note ω is a function of four variables (one a three-vector), but $\tilde{v}_\mathrm{p}$ is relative to c, so it does not depend on k for λ and h fixed ($\lambda = ck/h$).

agation for plane-wave solutions. First we note that there is a large degree of directional dependence in $\tilde{v}_\mathrm{p}$; this is known as *isotropy error*, or *anisotropy*. The amount of anisotropy also decreases in the limit of small wavenumbers, as expected by consistency. Highlighted in the figures are axial (green), side-diagonal (red), and space-diagonal (blue) directions of the cubic wavenumber cell. It can be seen that wavenumbers along the space-diagonal feature the least amount of dispersion error, remaining exact in the case of $\lambda = \sqrt{1/3}$. Meanwhile, the axial direction demonstrates the greatest errors, reaching as high as 30% error at the spatial Nyquist ($\beta_x = \pi/h, \beta_{yz} = 0$).

According to this analysis, and as is commonly known, it is best to fix λ as close as possible to the stability limit $\lambda = \sqrt{1/3}$ in order to operate with as little dispersion error as possible for a fixed h (in this scheme). In order to further eliminate dispersion error, one can vary the last remaining free parameter, h. Choosing h small means that one is operating closer to DC in the space of normalised wavenumbers $\boldsymbol{\beta} h \in [-\pi, \pi]^3$, where dispersion approaches zero. Provided that the initial conditions are relatively bandlimited, a minimum grid spacing can be determined by multidimensional sampling considerations [191], and reducing h beyond this minimum (refining the grid or "oversampling" the initial conditions while keeping λ fixed) can provide a numerical solution relatively free of dispersion error. Of course, this comes at a great cost, to be discussed shortly.

In order to see the effects of dispersion error in a numerical solution, as well as the effects of grid refinement, consider the following example. We take an initial condition ($\Psi(\mathbf{x}, 0)$) that only varies as a Gaussian along the x-direction (and $\Psi_t(\mathbf{x}, 0) = 0$), for which the continuous solution is a sum of left- and right-going travelling waves (along the x-axis). Numerical solutions computed from the simple Cartesian scheme, with $\lambda = \sqrt{1/3}$ and various choices of h, are shown in Fig. 3.5, wherein the red curve displays the ideal solution (right-going only) and the numerical solution is in blue. For each choice of h, it can be seen that the numerical wavefront spreads apart over time (and space), eventually losing coherence as it propagates to the right. This is an effect of numerical dispersion, which can be indirectly explained by $\tilde{v}_\mathrm{p}$; the wavefront can be seen as a continuous spectrum of plane-wave components (with a Gaussian magnitude envelope).[20] The rate at which the numerical wavefront loses coherence is inversely proportional to the chosen h, which is a consequence of consistency (or convergence), and also it demonstrates the effect of grid oversampling.

According to the above discussion and the presented numerical example, oversampling the spatial grid is a reliable way to eliminate dispersion errors, but this strategy carries a heavy computational cost in three spatial dimensions and quickly becomes impractical. Storage (memory) requirements scale to the third power in the oversampling factor, and arithmetic operations scale to the fourth power (when λ is fixed), which means that the compute time will also generally scale to the fourth power in the oversampling factor. These increases in computational costs are the main motivation for the pursuit of more accurate finite difference schemes (to be pursued in Chapter 4).

Since oversampling in 3-D carries such a heavy computational cost, it should only be used as necessary to achieve a desired level of accuracy (or confidence) in the numerical solution. There are many ways to interpret "accuracy" in the context of room acoustics, but as a starting point, one could attribute a level of confidence in the numerical solution to the "percent-error" in the numerical phase (or group) velocity. Oversampling could then be performed such that dispersion error is kept below a desired threshold over a range of wavenumbers of interest. For example, if one desires a one-percent error threshold in $\tilde{v}_\mathrm{p}$, it can be found by closer examination of $\tilde{v}_\mathrm{p}$ (as in Fig. 3.4(b)) that 0.19π rad is approximately the minimum normalised wavenumber (along the axial directions) with a one-percent error in $\tilde{v}_\mathrm{p}$, provided that $\lambda = \sqrt{1/3}$. Thus, one would need approximately a 5.3$\times$ oversampling factor

[20] In such case, the numerical solution may be more aptly described by a numerical *group velocity*, see, e.g., [275, 260]

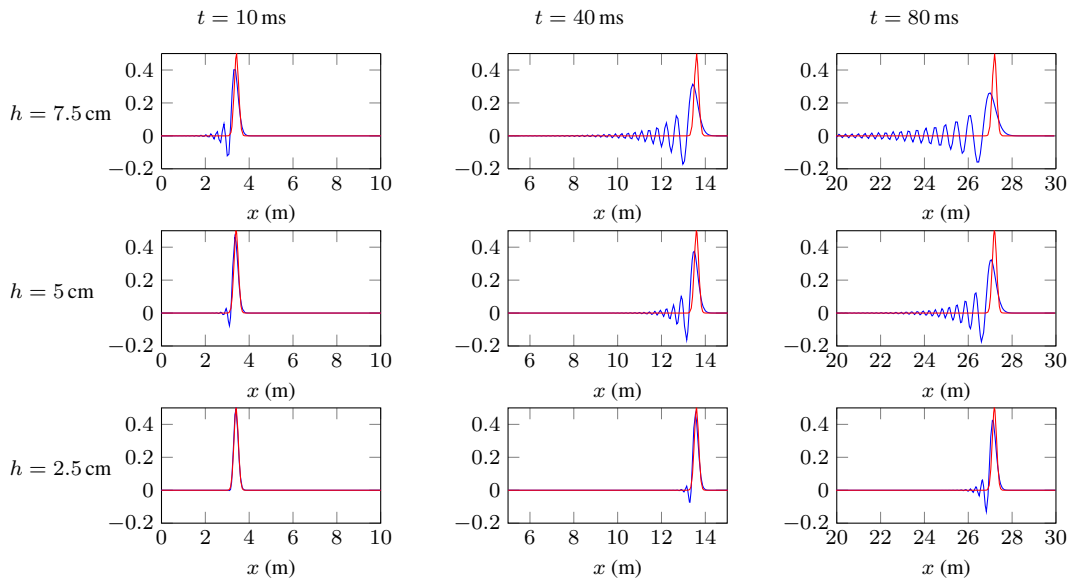


Figure 3.5: Demonstration of numerical dispersion effects for a pulse propagating and varying only along the x-direction, for various choices of h. The blue curve corresponds to the numerical solution, and the red curve corresponds to the exact solution. Snapshots are shown at times as indicated. Here $c = 340\,\text{m/s}$ and $\lambda = \sqrt{1/3}$.

$(1/0.19)$ over the minimum h dictated by sampling considerations in order to have all initial conditions bandlimited to 0.19π rad, and thus within the one-percent threshold. In other words, for a one-percent error threshold in $\tilde{v}_\text{p}$ one would need to sample the initial conditions with at least ten samples (points) per wavelength of interest (a commonly used "rule of thumb"). If a two-percent error threshold is desired, one would require approximately a $3.7\times$ oversampling factor ($\approx 1/0.27$), or 7–8 samples per wavelength, and if a 5% error is desirable, a $2.4\times$ oversampling factor would be necessary, etc.. Determining what is an appropriate threshold for a room acoustics problem from a perceptual point of view is another question, and one that cannot be reasonably addressed here; see, e.g. [214] (published work involving the current author).

3.2 Boundary conditions

In this section we derive boundary conditions for the simple Cartesian scheme through the use of energy methods. Energy methods comprise an extremely useful set of tools in this respect, since any boundary conditions can be proposed to be accordant with an energy-based stability framework.[21]

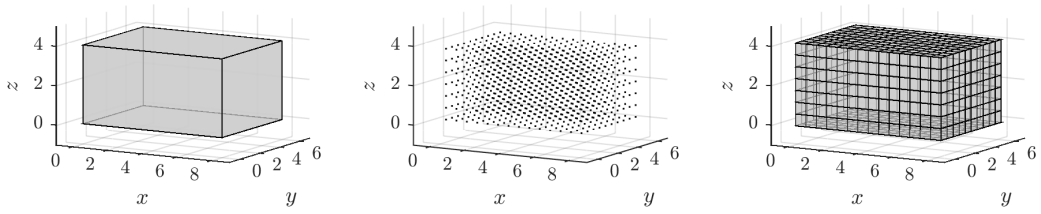
3.2.1 Preliminaries

For this section we consider a closed domain, denoted $\overline{\Omega} \subset \mathbb{R}^3$, and which can be decomposed into an open domain Ω and its closure (boundary) Γ, i.e., $\overline{\Omega} = \Omega \cup \Gamma$.

For illustrative purposes, it will be convenient, as a test case, to define an open box-shaped region:

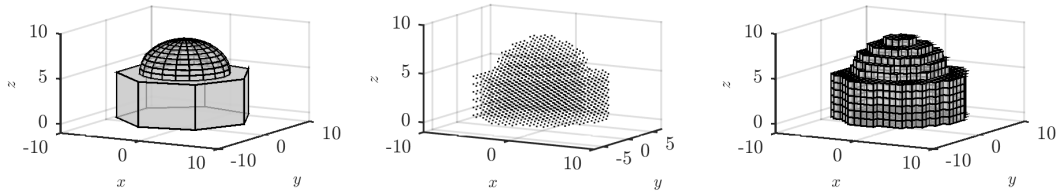
$$\mathcal{V}_\text{box} := (0, L_x) \times (0, L_y) \times (0, L_z) \tag{3.63}$$

[21] This contrasts with another often-used methodology, where one first proposes a boundary condition, followed by a stability analysis when possible (as in, e.g., [133]).



(a) $\overline{\Omega}$ (b) $\mathbb{G}_h$ (c) $\overline{\Omega}_h$

Figure 3.6: An example domain, $\mathcal{V}_{\text{box}}$ with $(L_x, L_y, L_z) = (8, 6, 4)\,\text{m}$ (left); along with a Cartesian grid, $\mathbb{G}_h$, for $h = 0.6\,\text{m}$ (middle); and the associated volume $\overline{\Omega}_h$ (right).



(a) $\overline{\Omega}$ (b) $\mathbb{G}_h$ (c) $\overline{\Omega}_h$

Figure 3.7: An example non-trivial domain made up of a heptagonal prism with an ellipsoidal dome (left); along with a Cartesian grid, $\mathbb{G}_h$, for $h = 0.8\,\text{m}$ (middle); and the associated volume $\overline{\Omega}_h$, demonstrating staircase effects (right).

where $\times$ denotes the Cartesian product.

Grids and index sets

The Cartesian grid under consideration is now defined as the intersection of a scaled integer (cubic) lattice and $\overline{\Omega}$, i.e.:

$$\mathbb{G}_h := \mathbb{Z}^3_h \cap \overline{\Omega} \tag{3.64}$$

and for convenience, we also define a vector index set for these grid points:

$$\mathcal{I} := \{\boldsymbol{i} \in \mathbb{Z}^3 : \boldsymbol{i}h \in \overline{\Omega}\} \tag{3.65}$$

For illustrative purposes it will also help to define a continuous region associated to the discrete set $\mathbb{G}_h$, which will serve as a "staircase approximation" to the continuous domain $\overline{\Omega}$. Recalling the basic Cartesian finite volume formulation discussed in Section 3.1.2, it makes sense to consider this region to be a union of cubic cells centered about points on the grid. This staircase region can formally be defined as:

$$\overline{\Omega}_h := \bigcup_{\boldsymbol{i} \in \mathcal{I}} \{\boldsymbol{i}h + \mathbf{x}' : \mathbf{x}' \in \mathbb{R}^3, \|\mathbf{x}'\|_\infty \leq h/2\} \tag{3.66}$$

where $\|\cdot\|_\infty$ is the Chebyshev norm (the "chessboard distance"). It is important to note that in general, $\overline{\Omega}_h \neq \overline{\Omega}$, i.e., these regions are not necessarily equivalent. We can decompose this staircased region into an open region and its boundary, $\overline{\Omega}_h = \Omega_h \cup \Gamma_h$, such that Γ_h represents the boundary surface of the staircased region. The sets $\mathbb{G}_h$ and $\overline{\Omega}_h$ are illustrated with example domains in Figs. 3.6 and 3.7.

For the following energy analyses, it will be useful to define an indicator function and its complement:

$$q_{\boldsymbol{i},w\pm} = \begin{cases} 1 & \text{if } \boldsymbol{i} \pm \hat{\mathbf{e}}_w \in \mathcal{I} \\ 0 & \text{otherwise} \end{cases}, \quad \overline{q}_{\boldsymbol{i},w\pm} = \begin{cases} 0 & \text{if } \boldsymbol{i} \pm \hat{\mathbf{e}}_w \in \mathcal{I} \\ 1 & \text{otherwise} \end{cases}, \quad w \in \{x, y, z\} \tag{3.67}$$

It will also be useful to recognise the property that

$$q_{\boldsymbol{i},w-} = q_{\boldsymbol{i},(-w)+} \quad \Longrightarrow \quad \sum_{w\in\{x,y,z\}} (q_{\boldsymbol{i},w+} + q_{\boldsymbol{i},w-}) = \sum_{w\in\pm\{x,y,z\}} q_{\boldsymbol{i},w+} \tag{3.68}$$

and we note that the latter summation is over six directions.

For convenience, we define index sets for "regular interior points" (points that have six interior neighbours) and for "boundary points" (points that have fewer than six interior neighbours):

$$\mathcal{I}_{\mathrm{i}} := \{\boldsymbol{i} \in \mathbb{Z}^3 : q_{\boldsymbol{i},x\pm} = q_{\boldsymbol{i},y\pm} = q_{\boldsymbol{i},z\pm} = 1\} \tag{3.69}$$

$$\mathcal{I}_{\mathrm{b}} := \mathcal{I} \setminus \mathcal{I}_{\mathrm{i}} \tag{3.70}$$

For example, in the case of $\Omega = \mathcal{V}_{\text{box}}$, we have:

$$\mathcal{I} = \{0, 1, \ldots, N_x\} \times \{0, 1, \ldots, N_y\} \times \{0, 1, \ldots, N_z\} \tag{3.71}$$

$$\mathcal{I}_{\mathrm{i}} = \{1, 2, \ldots, N_x - 1\} \times \{1, 2, \ldots, N_y - 1\} \times \{1, 2, \ldots, N_z - 1\} \tag{3.72}$$

where $N_w = \lfloor L_w/h \rfloor + 1$ for $w \in \{x, y, z\}$.

Inner products and summation-by-parts

We can also define an inner product on $\mathbb{G}_h$:

$$\langle \underline{u}^n, \underline{v}^n \rangle_{\mathbb{G}_h} := \sum_{\boldsymbol{i}\in\mathcal{I}} h^3 (\underline{u}^n_{\boldsymbol{i}})(\underline{v}^n_{\boldsymbol{i}}), \quad \|\underline{u}^n\|^2_{\mathbb{G}_h} := \langle \underline{u}^n, \underline{u}^n \rangle_{\mathbb{G}_h} \tag{3.73}$$

where $\underline{u}^n_{\boldsymbol{i}}$ and $\underline{v}^n_{\boldsymbol{i}}$ are arbitrary scalar grid functions on $\mathbb{G}_h$. We note that $\|\cdot\|_{\mathbb{G}_h}$ is an ℓ^2-norm on $\mathbb{G}_h$.

For this energy analysis, we will use a summation-by-parts (SBP) property for the closed domain. In the simple case of $\Omega = \mathcal{V}_{\text{box}}$, we have the following SBP property for δ_{x-} [24]:

$$\langle \underline{u}^n, \delta_{x-}\underline{v}^n \rangle_{\mathbb{G}_h} = -\langle \delta_{x+}\underline{u}^n, \underline{v}^n \rangle_{\mathbb{G}_h} + \sum_{i_y=0}^{N_y} \sum_{i_z=0}^{N_z} h^2 \left((e_{x+}\underline{u}^n_{N_x,i_y,i_z})(\underline{v}^n_{N_x,i_y,i_z}) - (\underline{u}^n_{0,i_y,i_z})(e_{x-}\underline{v}^n_{0,i_y,i_z}) \right) \tag{3.74}$$

which we can rewrite with the indicator function $q_{\boldsymbol{i},x+}$ as:

$$\langle \underline{u}^n, \delta_{x-}\underline{v}^n \rangle_{\mathbb{G}_h} = -\langle q_{x+}\delta_{x+}\underline{u}^n, \underline{v}^n \rangle_{\mathbb{G}_h} + \sum_{i_y=0}^{N_y} \sum_{i_z=0}^{N_z} h^2 \left((\underline{u}^n_{N_x,i_y,i_z})(\underline{v}^n_{N_x,i_y,i_z}) - (\underline{u}^n_{0,i_y,i_z})(e_{x-}\underline{v}^n_{0,i_y,i_z}) \right) \tag{3.75}$$

In the more general case of a non-trivial Ω, it is straightforward to arrive at the SBP property:

$$\langle \underline{u}^n, \delta_{w-}\underline{v}^n \rangle_{\mathbb{G}_h} \;=\; -\langle q_{w+}\delta_{w+}\underline{u}^n, \underline{v}^n \rangle_{\mathbb{G}_h} + \sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^2 \left(\overline{q}_{\boldsymbol{i},w+}(\underline{u}^n_{\boldsymbol{i}})(\underline{v}^n_{\boldsymbol{i}}) - \overline{q}_{\boldsymbol{i},w-}(\underline{u}^n_{\boldsymbol{i}})(e_{w-}\underline{v}^n_{\boldsymbol{i}}) \right) \tag{3.76}$$

3.2.2 Frequency-independent boundary conditions

Let us consider the case of frequency-independent immittance boundary conditions. The boundary value problem (BVP) under consideration here can be written as:

$$\Psi_{tt} = c^2\Delta\Psi \qquad t>0\,, \quad \mathbf{x}\in\Omega \qquad \text{(wave equation)} \tag{3.77a}$$

$$-\hat{\mathbf{n}}\cdot\nabla\Psi = \frac{\gamma}{c}\Psi_t \qquad t>0\,, \quad \mathbf{x}\in\Gamma \qquad \text{(boundary conditions)} \tag{3.77b}$$

$$\Psi = f_0\,,\; \Psi_t = g_0 \qquad t=0\,, \quad \mathbf{x}\in\overline{\Omega} \qquad \text{(initial conditions)} \tag{3.77c}$$

where $\gamma = \gamma(\mathbf{x}) \in \mathbb{R}^+$ is a (dimensionless) specific surface admittance. Two special cases are of note: for $\gamma = 0$, homogeneous Neumann conditions results, modelling rigid perfectly reflective walls ($R = 1$); and for $\gamma = 1$, one has absorbing conditions of the first-order Engquist-Majda type [69], which model perfectly absorbing walls at normal angles of incidence ($R = 0$) (as discussed in Section 2.1.3).

Energy analysis

The energy analysis for this discrete system is a continuation of the free-space case in Section 3.1.3, or it can be seen as a discrete form of the continuous analysis in Section 2.1.3. Boundary conditions will be chosen according to energy-stability considerations. Proceeding with the energy analysis, we multiply both sides of (3.15) by $\frac{\rho}{c^2}\delta_{t\cdot}\underline{\Psi}^n_{\boldsymbol{i}}$, and take the inner product (3.73) to obtain:

$$\frac{\rho}{c^2}\langle \delta_{t\cdot}\underline{\Psi}, \delta_{tt}\underline{\Psi}\rangle_{\mathbb{G}_h} = \sum_{w\in\{x,y,z\}} \rho\langle \delta_{t\cdot}\underline{\Psi}, \delta_{ww}\underline{\Psi}\rangle_{\mathbb{G}_h} \tag{3.78}$$

where we have dropped the superscript n for now. Making use of identities (3.26c) and (3.27) and the SBP property (3.76), we can rewrite (3.78) as:

$$\frac{\rho}{c^2}\langle \delta_{t\cdot}\underline{\Psi}, \delta_{tt}\underline{\Psi}\rangle_{\mathbb{G}_h} + \sum_{w\in\{x,y,z\}} \rho\langle q_{w+}\delta_{t\cdot}\delta_{w+}\underline{\Psi}, \delta_{w+}\underline{\Psi}\rangle_{\mathbb{G}_h} = -\mathfrak{b} \tag{3.79}$$

where

$$\mathfrak{b} = -\rho\sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^2 \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}(\delta_{t\cdot}\underline{\Psi}_{\boldsymbol{i}})(\delta_{w+}\underline{\Psi}_{\boldsymbol{i}}) \tag{3.80}$$

and where we have made use of identity (3.26b). $\mathfrak{b}$ is the discrete analogue of (2.32), and each term in $\mathfrak{b}$ accesses the grid function at spatial points that are outside of $\mathbb{G}_h$. Such exterior points are known as "ghost points". The choice of the boundary condition should be one that eliminates the ghost points in an explicit update for boundary nodes, and one that keeps $\mathfrak{b} \geq 0$. Some appropriate choices will be specified shortly.

Next we use identities (3.43a) and (3.43b) to arrive at an energy balance:

$$\delta_{t+}\mathfrak{h} = -\mathfrak{b}\,, \quad \mathfrak{h} = \mathfrak{v} + \mathfrak{t} \tag{3.81}$$

where

$$\mathfrak{v} = \frac{\rho}{2c^2}\|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{G}_h}\,, \qquad \mathfrak{t} = \frac{\rho}{2}\sum_{w\in\{x,y,z\}}\langle q_{w+}\delta_{w+}\underline{\Psi}, e_{t-}\delta_{w+}\underline{\Psi}\rangle_{\mathbb{G}_h} \tag{3.82}$$

In the absence of wall losses, we have $\mathfrak{b} = 0$ and the discrete energy is conserved over time. In the presence of wall losses, numerical energy is still conserved in a sense, as long as one considers the accumulation of lost energy. With that said, we define the following energetic quantities:

$$\overline{\mathfrak{h}}^n = \mathfrak{h}^n + \overline{\mathfrak{b}}^n\,, \qquad \overline{\mathfrak{b}}^n = k\sum_{m=0}^{n-1}\mathfrak{b}^m \tag{3.83}$$

Here, $\overline{\mathfrak{b}}$ represents the accumulation of dissipated power, and $\overline{\mathfrak{h}}$ represents the total energy including these losses. Then we can say that:

$$\overline{\mathfrak{h}}^n = \overline{\mathfrak{h}}^{n-1} = \cdots = \mathfrak{h}^0 \tag{3.84}$$

which should hold to numerical precision. In other words, it should be the case that:

$$1 - \frac{\overline{\mathfrak{h}}^n}{\overline{\mathfrak{h}}^{n-1}} \equiv \mathcal{O}(\text{machine epsilon}) \tag{3.85}$$

or more strictly, that:

$$1 - \frac{\overline{\mathfrak{h}}^n}{\overline{\mathfrak{h}}^{0}} \equiv \mathcal{O}(\text{machine epsilon}) \tag{3.86}$$

These energy conservation properties will be tested in subsequent numerical experiments.[22]

It remains to find conditions such that $\mathfrak{b}, \mathfrak{h} \geq 0$ for stability. Returning to the kinetic energy of indeterminate sign, we make use of identity (3.43e) to rewrite $\mathfrak{t}$ as:

$$\mathfrak{t} = \frac{\rho}{2}\sum_{w\in\{x,y,z\}}\left(\|q_{w+}\delta_{w+}\mu_{t-}\underline{\Psi}\|^2_{\mathbb{G}_h} - \frac{k^2}{4}\|q_{w+}\delta_{w+}\delta_{t-}\underline{\Psi}\|^2_{\mathbb{G}_h}\right) \tag{3.87}$$

and we have that:

$$\mathfrak{t} \geq -\frac{\rho k^2}{8}\sum_{w\in\{x,y,z\}}\|q_{w+}\delta_{w+}\delta_{t-}\underline{\Psi}\|^2_{\mathbb{G}_h} \tag{3.88}$$

Then we can make use of the following bound, which is not necessarily tight:

$$\|q_{w\pm}\delta_{w\pm}\underline{u}\|^2_{\mathbb{G}_h} \leq \frac{4}{h^2}\|\underline{u}\|^2_{\mathbb{G}_h} \tag{3.89}$$

to arrive at:

$$\mathfrak{t} \geq -\frac{\rho k^2}{2h^2}\sum_{w\in\{x,y,z\}}\|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{G}_h} \geq -\frac{3\rho k^2}{2h^2}\|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{G}_h} \tag{3.90}$$

This leads to the following bound on $\mathfrak{h}$:

$$\mathfrak{h} \geq \frac{1}{2c^2}\left(1 - 3\lambda^2\right)\|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{G}_h} \tag{3.91}$$

[22] There may be other finite precision effects to consider when verifying such numerical energy conservation properties; see, e.g., [272, Appendix A].

and the usual condition on the Courant number:

$$\lambda \leq \frac{1}{\sqrt{3}} \quad \implies \quad \mathfrak{h} \geq 0 \tag{3.92}$$

With regards to the boundary term, we require that $\mathfrak{b} \geq 0$ for the system to be stable as a whole. This leads to the following natural choice for a discretised form of (3.77b):

$$\mp \delta_{w\pm} \underline{\Psi}_{\boldsymbol{i}}^n = \frac{\gamma_{\boldsymbol{i},w\pm}}{c} \delta_{t\cdot} \underline{\Psi}_{\boldsymbol{i}}^n \,, \quad \boldsymbol{i} \in \mathcal{I}_{\mathrm{b}} \,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},w\pm} = 1 \tag{3.93}$$

where $\gamma_{\boldsymbol{i},w\pm}$ represents the real admittance on the boundary of Γ_h.

Under this choice of boundary condition, the boundary term in the energy balance is:

$$\mathfrak{b}^n = \frac{\rho}{c} \sum_{\boldsymbol{i} \in \mathcal{I}_{\mathrm{b}}} h^2 \sum_{w \in \pm\{x,y,z\}} \gamma_{\boldsymbol{i},w+} \overline{q}_{\boldsymbol{i},w+} (\delta_{t\cdot} \underline{\Psi}_{\boldsymbol{i}}^n)^2 \geq 0 \tag{3.94}$$

Thus, with (3.55), the energy of the system is strictly non-increasing (energy-stable).

In summary, we have the discrete finite difference boundary value problem, which is analogous to (3.77):

$$\delta_{tt} \underline{\Psi}_{\boldsymbol{i}}^n = c^2 \delta_\Delta \underline{\Psi}_{\boldsymbol{i}}^n \qquad n > 1 \,, \quad \boldsymbol{i} \in \mathcal{I} \tag{3.95a}$$

$$\mp \delta_{w\pm} \underline{\Psi}_{\boldsymbol{i}}^n = \frac{\gamma_{\boldsymbol{i},w\pm}}{c} \delta_{t\cdot} \underline{\Psi}_{\boldsymbol{i}}^n \qquad n > 1 \,, \quad \boldsymbol{i} \in \mathcal{I}_{\mathrm{b}}, \text{ when } \overline{q}_{\boldsymbol{i},w\pm} = 1 \tag{3.95b}$$

$$\underline{\Psi}_{\boldsymbol{i}}^n = [f_0]_h, \;\; \delta_{t+} \underline{\Psi}_{\boldsymbol{i}}^n = [g_0]_h \qquad n = 0 \,, \quad \boldsymbol{i} \in \mathcal{I} \tag{3.95c}$$

and where (3.21) could also be used instead of (3.95c) for the calculation of $\underline{\Psi}_{\boldsymbol{i}}^1$.

It remains to combine the interior update with boundary conditions for boundary nodes. As mentioned previously the boundary condition serves to eliminate the ghost points from the system. Thus, we can rearrange (3.95b) to get:

$$\underline{\Psi}_{\boldsymbol{i} \pm \hat{\mathbf{e}}_w}^n = \underline{\Psi}_{\boldsymbol{i}}^n + \frac{\gamma_{\boldsymbol{i},w\pm}}{2\lambda} (\underline{\Psi}_{\boldsymbol{i}}^{n+1} - \underline{\Psi}_{\boldsymbol{i}}^{n-1}) \,, \quad \boldsymbol{i} \in \mathcal{I}_{\mathrm{b}} \,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},w\pm} = 1 \tag{3.96}$$

and the ghost point value $\underline{\Psi}_{\boldsymbol{i} \pm \hat{\mathbf{e}}_w}^n$ will be substituted back into the update equation for $\underline{\Psi}_{\boldsymbol{i}}^n$. After making this substitution, one obtains the following boundary update equation:

$$\underline{\Psi}_{\boldsymbol{i}}^{n+1} = \frac{1}{1 + \lambda \gamma'_{\boldsymbol{i}}/2} \left((2 - \lambda^2 K_{\boldsymbol{i}}) \underline{\Psi}_{\boldsymbol{i}}^n + (\lambda \gamma'_{\boldsymbol{i}}/2 - 1) \underline{\Psi}_{\boldsymbol{i}}^{n-1} + \lambda^2 \tilde{Q}_{\boldsymbol{i}}^n \right) \tag{3.97}$$

where

$$\gamma'_{\boldsymbol{i}} = \sum_{w \in \{x,y,z\}} (\overline{q}_{\boldsymbol{i},w+} \gamma_{\boldsymbol{i},w+} + \overline{q}_{\boldsymbol{i},w-} \gamma_{\boldsymbol{i},w-}) \tag{3.98}$$

and

$$K_{\boldsymbol{i}} = \sum_{w \in \pm\{x,y,z\}} q_{\boldsymbol{i},w+} \,, \qquad \overline{K}_{\boldsymbol{i}} = 6 - K_{\boldsymbol{i}} \tag{3.99}$$

and where $\tilde{Q}_{\boldsymbol{i}}^n$ is a partial summation:

$$\tilde{Q}_{\boldsymbol{i}}^n = \sum_{w \in \pm\{x,y,z\}} q_{\boldsymbol{i},w+} \underline{\Psi}_{\boldsymbol{i}+\hat{\mathbf{e}}_w}^n \tag{3.100}$$

This partial summation is specialised at the boundary nodes, but $\tilde{Q}_{\boldsymbol{i}}^n = Q_{\boldsymbol{i}}^n$ for $\boldsymbol{i} \in \mathcal{I}_{\mathrm{i}}$. Also, $\overline{K}_{\boldsymbol{i}}$ is

simply the number of "ghost point"–neighbours, and $K_{\boldsymbol{i}}$ is the number of non-zero summands in $\tilde{Q}^n_{\boldsymbol{i}}$. Eq. (3.97) is also valid for interior points, for which $\gamma_{\boldsymbol{i},w\pm} = 0$.

For simplicity, we could also consider one admittance value per boundary node, i.e., $\gamma_{\boldsymbol{i},w\pm} \equiv \gamma_{\boldsymbol{i}}$. In this case we have $\gamma'_{\boldsymbol{i}} = \overline{K}_{\boldsymbol{i}}\gamma_{\boldsymbol{i}}$, and the update equation simplifies to:

$$\underline{\Psi}^{n+1}_{\boldsymbol{i}} = \frac{1}{1 + \lambda\overline{K}_{\boldsymbol{i}}\gamma_{\boldsymbol{i}}/2}\left((2 - \lambda^2 K_{\boldsymbol{i}})\underline{\Psi}^n_{\boldsymbol{i}} + (\lambda\overline{K}_{\boldsymbol{i}}\gamma_{\boldsymbol{i}}/2 - 1)\underline{\Psi}^{n-1}_{\boldsymbol{i}} + \lambda^2\tilde{Q}^n_{\boldsymbol{i}}\right) \tag{3.101}$$

In order to illustrate how $\tilde{Q}^n_{\boldsymbol{i}}$ varies over the boundary, consider the special case $\Omega = \mathcal{V}_{\text{box}}$. In that case, the modified partial summation for wall, edge, and corner nodes closest to the origin are:

$$\text{wall :} \qquad \tilde{Q}^n_{\boldsymbol{i}} = \underline{\Psi}^n_{1,i_y,i_z} + \underline{\Psi}^n_{0,i_y+1,i_z} + \underline{\Psi}^n_{0,i_y-1,i_z} + \underline{\Psi}^n_{0,i_y,i_z+1} + \underline{\Psi}^n_{0,i_y,i_z-1} \tag{3.102a}$$

$$\text{edge :} \qquad \tilde{Q}^n_{\boldsymbol{i}} = \underline{\Psi}^n_{1,0,i_z} + \underline{\Psi}^n_{0,1,i_z} + \underline{\Psi}^n_{0,0,i_z+1} + \underline{\Psi}^n_{0,0,i_z-1} \tag{3.102b}$$

$$\text{corner :} \qquad \tilde{Q}^n_{\boldsymbol{i}} = \underline{\Psi}^n_{1,0,0} + \underline{\Psi}^n_{0,1,0} + \underline{\Psi}^n_{0,0,1} \tag{3.102c}$$

where for the wall $i_x = 0, i_y \in [1, N_y - 1], i_z \in [1, N_z - 1]$, and $i_x = i_y = 0, , i_z \in [1, N_z - 1]$ for the edge, and $i_x = i_y = i_z = 0$ for the corner node.

Example initial conditions

In order to present numerical examples, it will help to first define some example initial conditions that will be used to excite the discrete system. One useful function that can be used as either initial condition ($f_0(\mathbf{x})$ or $g_0(\mathbf{x})$) is the Gaussian in 3-D centered about the point $\mathbf{x}_0$. This is written as:

$$\exp\left(\frac{-\|\mathbf{x} - \mathbf{x}_0\|^2}{2\sigma^2}\right) \tag{3.103}$$

where σ^2 is the variance in squared-metres. Since this function has a Fourier transform that is also a Gaussian, it is straightforward to work out the useful property that the magnitude spectrum of this function takes on its maximum at DC ($\beta = 0$) and decreases monotonically, having decayed by N dB at the wavenumber

$$\beta = \sqrt{\frac{N\ln(10)}{10\sigma^2}} \tag{3.104}$$

It is also useful to define a *normalised* Gaussian, such that it integrates to one. This is simply:

$$\frac{1}{(\sigma\sqrt{2\pi})^3}\exp\left(\frac{-\|\mathbf{x} - \mathbf{x}_0\|^2}{2\sigma^2}\right) \tag{3.105}$$

and the decay rate is unchanged. The normalised Gaussian can also be used to represent the Dirac delta $\delta(\mathbf{x} - \mathbf{x}_0)$, which it approaches in the limit of $\sigma \to 0$ [15].

Another useful initial condition, in discrete form, is the Kronecker delta in both space and time. This can be used to construct the simplest approximation to the Dirac delta, which can then be used to compute a discrete room impulse response. This set of initial conditions can be expressed by the two initial states:

$$\underline{\Psi}^0_{\boldsymbol{i}} = 0\,, \quad \underline{\Psi}^1_{\boldsymbol{i}} = \begin{cases} V_0 c^2 k/h^3 & \boldsymbol{i} = \boldsymbol{i}_0 \\ 0 & \text{otherwise} \end{cases} \tag{3.106}$$

where $\boldsymbol{i}_0 h \approxeq \mathbf{x}_0$ and V_0 is an effective source volume, which for simplicity here, is associated with the (arbitrary) value $V_0 = 1\,\text{m}^3$.

Numerical examples: wave propagation and energy conservation

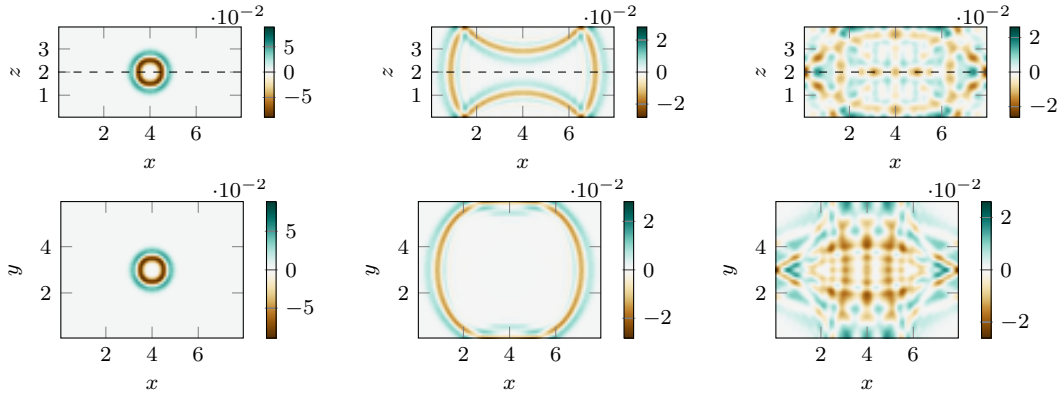
In order to demonstrate the operation of this scheme, and to demonstrate some interesting properties of numerical energy conservation, we provide a few simple numerical examples. First, we consider a box of size $(L_x, L_y, L_z) = (8, 6, 4)$ (in metres). The wall admittance coefficient is set to two values. The first chosen admittance is $\gamma = 0$ in order to demonstrate lossless behaviour. The second choice is $\gamma = 0.05$, which provides a model reflection coefficient at normal incidence of $R \approx 0.9$. The initial conditions are $g_0 = 0$ and f_0 is a Gaussian with variance $0.02\,\mathrm{m}^2$, centered about the middle of the domain. For the finite difference scheme, we set $h = 0.1\,\mathrm{m}$ and $\lambda = \sqrt{1/3}$ (and $k = h\lambda/c$). Consequently, the sample rate $1/k$ is approximately 6 kHz.

Snapshots of the simulation, through the 2-D planes $z = h\lfloor N_z/2 \rfloor$ and $y = h\lfloor N_y/2 \rfloor$, are displayed in Fig. 3.8 for the lossy case, demonstrating typical wave propagation in a shoebox room. Numerical energy quantities for both cases are shown in Figs. 3.9 and 3.10, where it can be seen that the accumulated energy is indeed conserved to numerical precision, in both forms (3.85) and (3.86). The variations in numerical energy appear in integer multiples of 2^{-53}, or machine epsilon in double precision floating-point arithmetic.

It is worth remarking that such energy calculations have an important practical significance; if energy conservation is not satisfied down to numerical precision for an energy-conserving (linear) scheme, it usually means that there is a bug in the underlying code, and consequently, these calculations provide a way to check that the finite difference algorithms are properly implemented [24]. For these reasons, they will be provided for many of the numerical examples in this thesis.

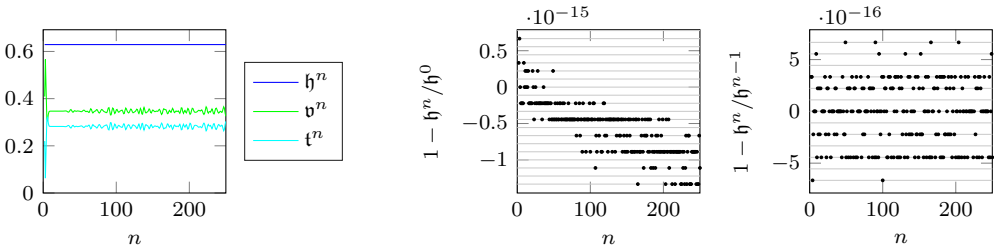
The non-trivial, domed-shaped domain is considered next. In this example, we set $h = 0.2\,\mathrm{m}$ and $\lambda = \sqrt{1/3}$ (and $k = h\lambda/c$). The initial conditions in this case are $f_0 = 0$ and g_0 is a Gaussian with variance $0.1\,\mathrm{m}^2$, centered about the middle of the domain. Two slices of the 3-D field are shown in Fig. 3.11 at a few instants in time for the case of $\gamma = 0.05$. Numerical energy conservation for this example, in lossless and lossy forms, is illustrated in Figs. 3.13 and 3.12.[23]

[23] We note the energy variation $1 - \mathfrak{h}^n/\mathfrak{h}^0$ drifts in the lossless case in Fig. 3.12, which may be due to the fact that the solution itself also drifts, degrading the relative accuracy of the solution. A detailed investigation of such round-off effects in numerical energy calculations can be found in [272, Appendix A].



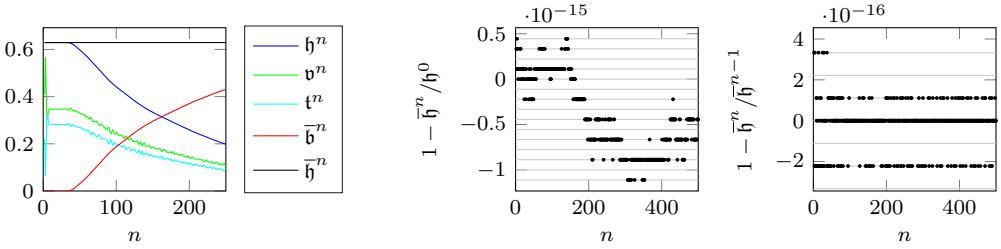
(a) $t = 2\,\mathrm{ms}$ (b) $t = 10\,\mathrm{ms}$ (c) $t = 25\,\mathrm{ms}$

Figure 3.8: Snapshots of simulation of 3-D wave equation through 2-D slices $y = L_y/2$ (top) and $z = L_z/2$ (bottom), at times (approximately) as indicated. The dotted line in the top figures demarcates the z-slice for the bottom figures.



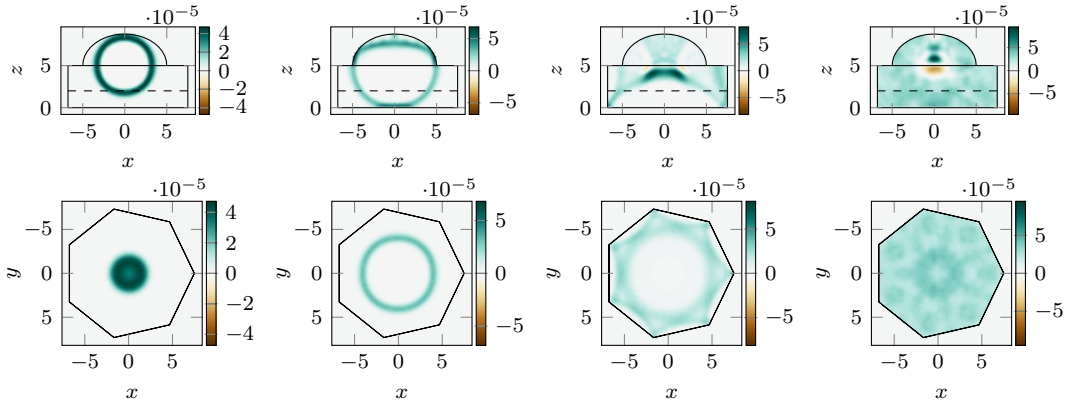
(a) Numerical energies (in Joules) (b) Relative variations in numerical energy

Figure 3.9: Plots of numerical energy and variations in numerical energy for the lossless case $\gamma = 0$. The variations in energy are integer multiples of double precision machine epsilon, as denoted by grey lines.



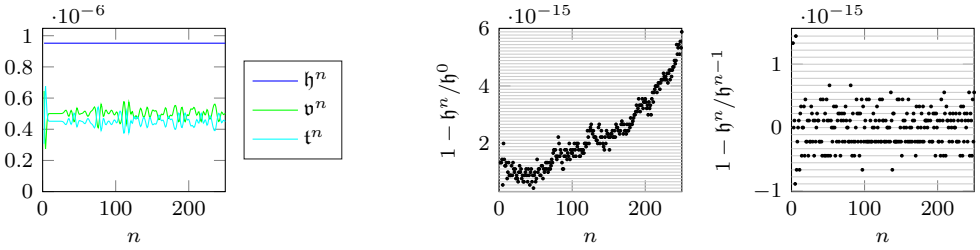
(a) Numerical energies (in Joules) (b) Relative variations in numerical energy

Figure 3.10: Plots demonstrating conservation of numerical energy including losses (compare Figs. 3.9(a) and 3.10(a)). The variations in energy are integer multiples of double precision machine epsilon, as denoted by grey lines.



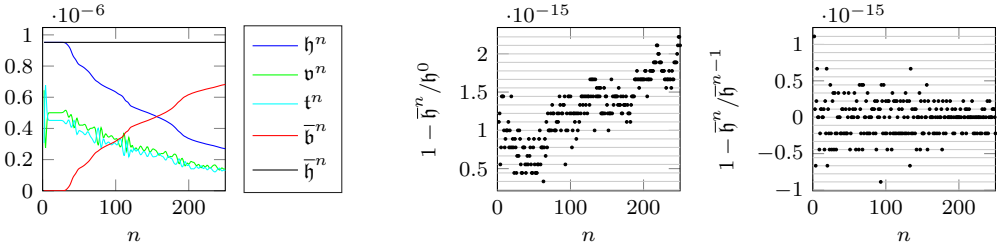
(a) $t = 10\,\text{ms}$ (b) $t = 15\,\text{ms}$ (c) $t = 25\,\text{ms}$ (d) $t = 75\,\text{ms}$

Figure 3.11: Snapshots of simulations of 3-D wave equation, at times as indicated. x- and y-axes are in units of metres.



(a) Energetic quantities (in Joules) (b) Relative variations in numerical energy

Figure 3.12: Plots of numerical energy and variations in numerical energy for the lossless case $\gamma = 0$.

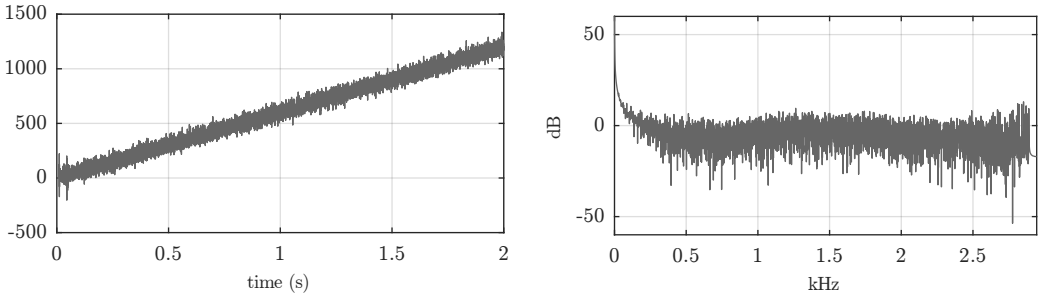


(a) Energetic quantities (in Joules) (b) Relative variations in numerical energy

Figure 3.13: Plots of numerical energy and variations in numerical energy for the lossy case $\gamma = 0.05$.

Numerical examples: room impulse responses

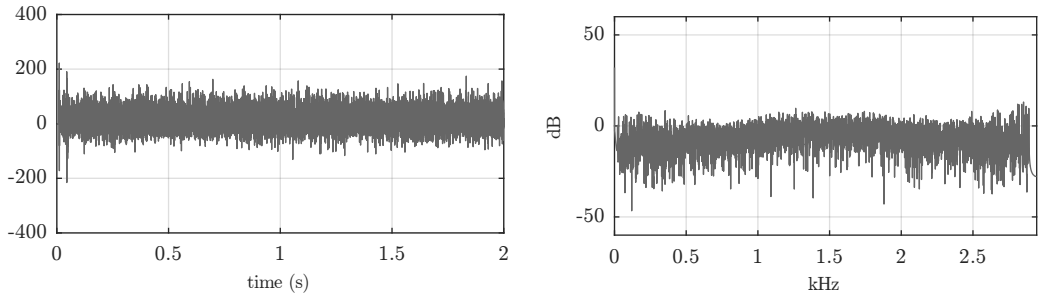
It is also useful to analyse simulation outputs at fixed locations on the grid (time-series outputs), in order to get a view into the frequency-domain behaviour of this scheme, along with some time-domain aspects. A useful starting test case is the case of a box-domain with rigid walls ($\gamma = 0$). Considering again the box-domain of size $(L_x, L_y, L_z) = (8, 6, 4)$, in metres, we compute an impulse response using a Kronecker delta initial condition with input and output locations $\mathbf{x}_0 = (5, 4, 3)$ and $\mathbf{x}_1 = (2.8, 1.7, 0.9)$ (in metres), respectively. The spatial step is set to $h = 0.1\,\text{m}$ and $\lambda = \sqrt{1/3}$ as before. The output signal (an impulse response), denoted $\underline{\Psi}^n_{\text{out}}$, is the time-series given by $\underline{\Psi}^n_{\boldsymbol{i}_1}$ where $\boldsymbol{i}_1 = \lfloor \mathbf{x}_1/h \rceil$. The computed impulse response and its magnitude spectrum are displayed in Fig. 3.14.



(a) Impulse response, $\underline{\Psi}^n_{\text{out}}$. Vertical axes in m^2/s.

(b) Magnitude spectrum of impulse response

Figure 3.14: Impulse response and its magnitude spectrum, obtained from box-domain with rigid walls.



(a) Impulse response, $\underline{\Psi}^n_{\text{out}}$. Vertical axes in m^2/s.

(b) Magnitude spectrum of impulse response

Figure 3.15: Impulse response (DC-filtered) and its magnitude spectrum, obtained from box-domain with rigid walls.

There are a few details worth pointing out in Fig. 3.14. First, we can see a linear drift in the acoustic velocity potential output, which at first seems "unphysical", but it is actually expected for the case of rigid walls and the chosen excitation,[24] and it is important to remember that the corresponding pressure, related through one temporal derivative, has no such drift. The exhibited linear drift in the acoustic velocity potential can be seen in the spectrum as a boost in signal-energy near DC. For the purpose of analysing the acoustic velocity potential, it is useful to remove this energy near DC using a simple DC blocker [249]. This impulse response after filtering is shown in Fig. 3.15. It is also important to mention that the scaling of the vertical axes, while exhibiting large amplitudes, is not a cause for alarm, since this response is still an abstraction and we must recall that the Dirac delta source has, in a sense, infinite amplitude, and thus the numerical output will tend to have large amplitudes. Subsequent time-domain

[24]This is akin to a 1-D string with free ends excited by an initial velocity or an upward force (these two excitations being related by Duhamel's principle [61, 15, 75]), which would cause the string to drift upwards linearly (as expected).

convolution with a source volume velocity $q_s(t)$ (see Section 2.1.2) renormalises such scalings, as long as the discrete convolution is appropriately scaled to the sample rate.

A useful "validation" test is to compare the resonant frequencies appearing in the computed impulse response to the known analytical frequencies for this box domain with rigid walls. In Fig. 3.16 we display the first few resonances that appear in this impulse response, along with theoretical modal frequencies for modal indices $\boldsymbol{m} = (m_x, m_y, m_z)$ (see Section 2.1.3) where $\max(m_x, m_y, m_z) \leq 3$. It can be seen that there is an agreement between the impulse response and theoretical modal frequencies in the range of frequencies displayed. In general, discrepancies appear in higher frequencies due to numerical dispersion.

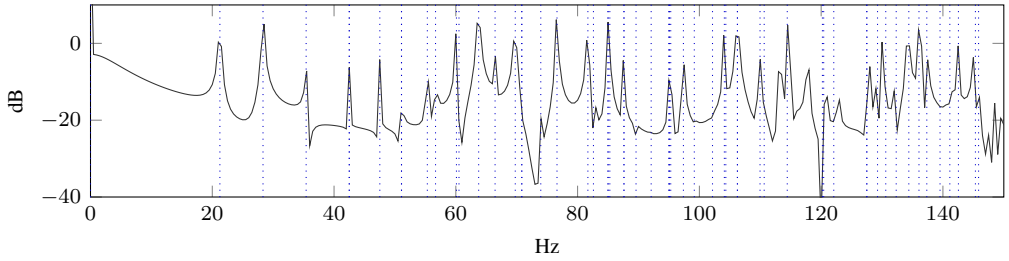
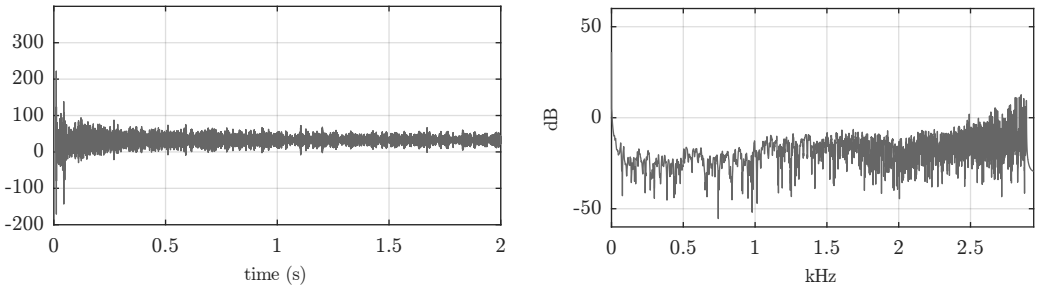


Figure 3.16: Magnitude spectrum of an impulse response obtained from box-domain with rigid walls, plotted alongside modal frequencies (blue, dotted).

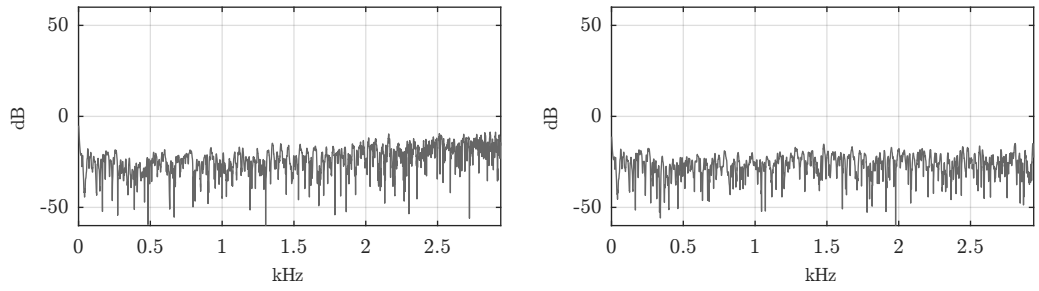
Next we consider the case of absorptive walls, with $\gamma = 0.05$. An impulse response computed from an analogous example problem (to those presented earlier) is shown in Fig. 3.17. First, we note that there is no longer a DC drift, but there is a general DC offset present; this can also be removed for analysis purposes. It can also be seen that the loss experienced by the numerical solution is not constant across the whole frequency range, despite the fact that the model boundary condition is supposedly frequency-independent. This is mostly due to the fact that the boundary condition assumes a constant wave speed c, but in practice the numerical wave speed (the numerical phase velocity) is frequency-dependent due to numerical dispersion. As a consequence, the absorption at the boundaries (at any given angle of incidence) is erroneously frequency-dependent. On the other hand, and as expected, errors in the boundary absorption caused by numerical dispersion can be mitigated by using a smaller spatial step. Fig. 3.18 shows the same band of frequencies from an impulse response computed from finer grids, with $h = 0.05\,\mathrm{m}$ and with $h = 0.025\,\mathrm{m}$. Indeed, it can be seen that as the grid is refined, the loss due to wall absorption approaches a flat spectrum (apart from variations due to resonances), in accordance with the frequency-independent boundary conditions.



(a) Impulse response, $\underline{\Psi}^{n}_{\text{out}}$. Vertical axes in m^2/s.

(b) Magnitude spectrum of impulse response

Figure 3.17: Impulse response (DC-filtered) and its magnitude spectrum, obtained from box-domain with non-rigid walls ($\gamma = 0.05$) and $h = 0.1$ m.



(a) $h = 0.05\,\text{m}$

(b) $h = 0.025\,\text{m}$

Figure 3.18: Magnitude spectra of impulse responses obtained from box-domain with non-rigid walls ($\gamma = 0.05$), with h as indicated above (DC-filtering applied).

3.2.3 Boundary conditions for the staggered scheme

In this section, we recast the boundary conditions for the second-order scheme in equivalent staggered forms, adapted from a more general finite volume framework presented in [26]. We will not need to conduct a stability analysis because stability will be ensured for these staggered forms by the previous energy analyses, provided that all boundary conditions are related through (3.33).

Making use of the identity $\delta_{t\cdot} = \mu_{t\mp}\delta_{t\pm}$, equivalent boundary conditions to (3.93) can be written as:

$$\pm \underline{v}^n_{w,\boldsymbol{i}\pm\hat{\mathbf{e}}_w/2} = Y_0\gamma_{\boldsymbol{i},w\pm}\mu_{t-}\underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}}\,, \quad \boldsymbol{i}\in\mathcal{I}_\mathrm{b}\,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},w\pm} = 1 \tag{3.107}$$

Taking into account outward normal vectors of a boundary cell, this could also be written as:

$$\underline{v}^n_{(\pm w),\boldsymbol{i}\pm\hat{\mathbf{e}}_w/2} = Y_0\gamma_{\boldsymbol{i},w\pm}\mu_{t-}\underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}}\,, \quad \boldsymbol{i}\in\mathcal{I}_\mathrm{b}\,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},w\pm} = 1 \tag{3.108}$$

These equations can be substituted back into (3.31a) to obtain a modified update equation for the pressure grid function:

$$\underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}} \;=\; \frac{1}{1+\overline{K}_{\boldsymbol{i}}\lambda\gamma_{\boldsymbol{i}}/2}\left((1\;-\;\overline{K}_{\boldsymbol{i}}\lambda\gamma_{\boldsymbol{i}}/2)\underline{p}^{n-\frac{1}{2}}_{\boldsymbol{i}}\;-\;\lambda Z_0\sum_{w\in\{\pm x,\pm y,\pm z\}} q_{\boldsymbol{i},w+}\underline{v}^n_{w,\boldsymbol{i}+\hat{\mathbf{e}}_w/2}\right) \tag{3.109}$$

where here we have taken $\gamma_{\boldsymbol{i},w\pm} \equiv \gamma_{\boldsymbol{i}}$ for simplicity. Note that here the boundary velocity component is treated like a ghost point, in the sense that it is substituted back into another update and does not need to be explicitly calculated. Boundary conditions for the staggered formulation of Botteldooren can be similarly derived by the application of $e^{1/2}_{t+}$ to the above conditions. It is important to point out, however, that these conditions (e.g., (3.107)) are different from those proposed by Botteldooren in [38] (for the frequency-independent special case). More specifically, the above conditions (e.g., (3.107)) do not place additional restrictions on the Courant number and they are formulated in terms of a pressure update, whereas the boundary conditions proposed in [38] place additional restrictions on the Courant number and are formulated as a boundary normal velocity update [38].

Summary of the first-order BVP in continuous and discrete forms

Returning to the conditions derived from the energy analysis, we can write out an equivalent BVP to (3.77)[25] for the first-order system using the relations (2.6):

$$\rho\mathbf{v}_t = -\nabla p \qquad t>0\,, \quad \mathbf{x}\in\Omega \tag{3.110a}$$

$$p_t = -\kappa\nabla\cdot\mathbf{v} \qquad t>0\,, \quad \mathbf{x}\in\Omega \tag{3.110b}$$

$$\hat{\mathbf{n}}\cdot\mathbf{v} = Y_0\gamma p \qquad t>0\,, \quad \mathbf{x}\in\Gamma\,, \quad \gamma=\gamma(\mathbf{x})\geq 0 \tag{3.110c}$$

$$p = \rho g_0 \qquad t=0\,, \quad \mathbf{x}\in\overline{\Omega} \tag{3.110d}$$

$$\mathbf{v} = -\nabla f_0 \qquad t=0\,, \quad \mathbf{x}\in\overline{\Omega} \tag{3.110e}$$

[25] Equivalent in p and $\mathbf{v}$, and to a constant in Ψ.

Similarly, we can summarise the fully discrete staggered grid scheme with the analogous discrete BVP:

$$\rho\delta_{t+}\underline{v}^n_{w,\boldsymbol{i}+\hat{\mathbf{e}}_w/2} = -\delta_{w+}\underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}} \qquad n \geq 0\,, \quad \boldsymbol{i} \in \mathcal{I}\,, \quad \text{when} \quad q_{\boldsymbol{i},w+} = 1 \tag{3.111a}$$

$$\delta_{t-}\underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}} = -\kappa\boldsymbol{\delta}^{\mathrm{T}}_{\mathbf{x}-}\underline{\mathbf{v}}^n_{\boldsymbol{i}+\frac{1}{2}} \qquad n > 0\,, \quad \boldsymbol{i} \in \mathcal{I} \tag{3.111b}$$

$$\underline{v}^n_{(\pm w),\boldsymbol{i}\pm\hat{\mathbf{e}}_w/2} = Y_0\gamma_{\boldsymbol{i},w\pm}\mu_{t+}\underline{p}^n_{\boldsymbol{i}} \qquad n > 0\,, \quad \boldsymbol{i} \in \mathcal{I}_{\mathfrak{b}}\,, \quad \text{when} \quad \overline{q}_{\boldsymbol{i},w\pm} = 1 \tag{3.111c}$$

$$\underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}} = \rho[g_0]_h \qquad n = 0\,, \quad \boldsymbol{i} \in \mathcal{I} \tag{3.111d}$$

$$\underline{\mathbf{v}}^n_{\boldsymbol{i}+\frac{1}{2}} = -\boldsymbol{\delta}_{\mathbf{x}+}[f_0]_h \qquad n = 0\,, \quad \boldsymbol{i} \in \mathcal{I} \tag{3.111e}$$

which is fully equivalent to the second-order system (3.95) in $\underline{p}^{n+\frac{1}{2}}_{\boldsymbol{i}}$ and $\underline{\mathbf{v}}^n_{\boldsymbol{i}+\frac{1}{2}}$ by the relations (3.33).

3.2.4 Frequency-dependent boundary conditions

In this section, we present frequency-dependent boundary conditions for locally reactive walls with a single local resonance. These conditions are adapted from a more general finite volume framework presented in [26] and simplified here for the basic Cartesian scheme.

A frequency-dependent admittance condition

We start with the case of a single resonance admittance boundary. The continuous boundary condition under consideration is thus (2.48), which is rewritten here in terms of Ψ:

$$\hat{\mathbf{n}} \cdot \nabla\Psi = -\frac{1}{c}(A\Psi_{tt} + B\Psi_t + C\Psi)\,, \quad \mathbf{x} \in \Gamma \tag{3.112}$$

where $A = A(\mathbf{x}), B = B(\mathbf{x}), C = C(\mathbf{x})$ (for $\mathbf{x} \in \Gamma$) and $A, B, C \geq 0$. It will be convenient to make use of the outward normal velocity component $v_\perp = -\hat{\mathbf{n}} \cdot \nabla\Psi$, in which case the boundary condition is rewritten as:

$$v_\perp = \frac{1}{c}(A\Psi_{tt} + B\Psi_t + C\Psi)\,, \quad \mathbf{x} \in \Gamma \tag{3.113}$$

Considering now the finite difference scheme, we can also define a discrete outward normal velocity component as:

$$\underline{v}^n_{w,\boldsymbol{i}} = -\delta_{w+}\underline{\Psi}^n_{\boldsymbol{i}} \tag{3.114}$$

where $w \in \pm\{x, y, z\}$ represents an outward normal vector on Γ_h. We also note the property:

$$\pm\,\underline{v}^n_{w,\boldsymbol{i}} = \underline{v}^n_{(\pm w),\boldsymbol{i}} \tag{3.115}$$

From the discrete energy analysis carried out in Section 3.2.2, we can rewrite the boundary term ($\mathfrak{b}$) in (3.80) as:

$$\mathfrak{b} = \rho \sum_{\boldsymbol{i} \in \mathcal{I}_{\mathfrak{b}}} h^2 \sum_{w \in \pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}(\delta_{t\cdot}\underline{\Psi}_{\boldsymbol{i}})(\underline{v}_{w,\boldsymbol{i}}) \tag{3.116}$$

For stability, there should be, ideally, no additional conditions on the model parameters (besides non-negativity). One possible choice of a consistent discretisation of (3.113) that achieves this is:

$$\underline{v}_{(\pm w),\boldsymbol{i}} = \frac{1}{c}(A_{\boldsymbol{i},w\pm}\delta_{tt}\underline{\Psi}_{\boldsymbol{i}} + B_{\boldsymbol{i},w\pm}\delta_{t\cdot}\underline{\Psi}_{\boldsymbol{i}} + C_{\boldsymbol{i},w\pm}\mu_{tt}\underline{\Psi}_{\boldsymbol{i}})\,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},w\pm} = 1 \tag{3.117}$$

where $w \in \{x, y, z\}$, and where $A_{\boldsymbol{i},w\pm}, B_{\boldsymbol{i},w\pm}, C_{\boldsymbol{i},w\pm}$ are A, B, C projected onto Γ_h (one triplet per

ghost point), which are non-zero only for $\boldsymbol{i} \in \mathcal{I}_\mathfrak{b}$. Then, inserting (3.117) into (3.116) and making use of the identities (3.43a) and (3.43c), we have:

$$\mathfrak{b} = \delta_{t+}\mathfrak{h}_\mathfrak{b} + \mathfrak{q}_\mathfrak{b} \tag{3.118}$$

where

$$\mathfrak{h}_\mathfrak{b} = \frac{\rho}{2c}\sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^2 \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}(A_{\boldsymbol{i},w+}(\delta_{t-}\underline{\Psi}_{\boldsymbol{i}})^2 + C_{\boldsymbol{i},w+}(\mu_{t-}\underline{\Psi}_{\boldsymbol{i}})^2) \geq 0 \tag{3.119}$$

$$\mathfrak{q}_\mathfrak{b} = \frac{\rho}{c}\sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^2 \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}B_{\boldsymbol{i},w+}(\delta_{t\cdot}\underline{\Psi}_{\boldsymbol{i}})^2 \geq 0 \tag{3.120}$$

In terms of the overall energy balance, we have now:

$$\delta_{t+}\mathfrak{h} = -\mathfrak{q}_\mathfrak{b}\,, \quad \mathfrak{h} = \mathfrak{h}_\mathfrak{i} + \mathfrak{h}_\mathfrak{b}\,, \quad \mathfrak{h}_\mathfrak{i} = \mathfrak{v} + \mathfrak{t} \tag{3.121}$$

with $\mathfrak{v}$ and $\mathfrak{t}$ unchanged from (3.82). Thus, the total stored energy is non-increasing (the scheme is stable) as long as $\lambda \leq \sqrt{1/3}$ and $A, B, C \geq 0$. Finally, after substituting ghost points back into the update equation of $\underline{\Psi}_{\boldsymbol{i}}^n$ for $\boldsymbol{i} \in \mathcal{I}_\mathfrak{b}$, we can arrive at the explicit boundary update:

$$\underline{\Psi}_{\boldsymbol{i}}^{n+1} = \frac{1}{1+a_{\boldsymbol{i}}^{(1)}}\left(\left(2-\lambda^2 K_{\boldsymbol{i}} + a_{\boldsymbol{i}}^{(2)}\right)\underline{\Psi}_{\boldsymbol{i}}^n + \left(a_{\boldsymbol{i}}^{(3)} - 1\right)\underline{\Psi}_{\boldsymbol{i}}^{n-1} + \lambda^2\tilde{Q}_{\boldsymbol{i}}^n\right) \tag{3.122}$$

where

$$a_{\boldsymbol{i}}^{(1)} = \lambda \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}\left(\frac{A_{\boldsymbol{i},w+}}{k} + \frac{B_{\boldsymbol{i},w+}}{2} + \frac{C_{\boldsymbol{i},w+}k}{4}\right) \tag{3.123a}$$

$$a_{\boldsymbol{i}}^{(2)} = \lambda \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}\left(\frac{2A_{\boldsymbol{i},w+}}{k} - \frac{C_{\boldsymbol{i},w+}k}{2}\right) \tag{3.123b}$$

$$a_{\boldsymbol{i}}^{(3)} = \lambda \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}\left(\frac{B_{\boldsymbol{i},w+}}{2} - \frac{A_{\boldsymbol{i},w+}}{k} - \frac{C_{\boldsymbol{i},w+}k}{4}\right) \tag{3.123c}$$

Note that this requires additional storage of A, B, C at boundary nodes. The storage requirements of the boundary model parameters will ultimately depend on the variations of boundary admittances desired in the model, and the additional storage requirements will generally be negligible compared to the storage of the state variables $\underline{\Psi}_{\boldsymbol{i}}^n$ and $\underline{\Psi}_{\boldsymbol{i}}^{n-1}$. We also note that (3.122) reduces to (3.97) in the case of $A = C = 0$ and with $B_{\boldsymbol{i},w\pm} = \gamma_{\boldsymbol{i},w\pm}$.

A simplified boundary update results if we assume that there is only triplet (A, B, C) per boundary node (rather than one per boundary face). In that case, we take the discretised boundary condition to be:

$$\underline{v}_{(\pm w),\boldsymbol{i}} = \frac{1}{c}(A_{\boldsymbol{i}}\delta_{tt}\underline{\Psi}_{\boldsymbol{i}} + B_{\boldsymbol{i}}\delta_{t\cdot}\underline{\Psi}_{\boldsymbol{i}} + C_{\boldsymbol{i}}\mu_{tt}\underline{\Psi}_{\boldsymbol{i}})\,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},w\pm} = 1 \tag{3.124}$$

where $A_{\boldsymbol{i}}, B_{\boldsymbol{i}}, C_{\boldsymbol{i}}$ are A, B, C projected onto $\mathbb{G}_h$ (one triplet per boundary node, i.e., $A_{\boldsymbol{i},w\pm} \equiv A_{\boldsymbol{i}}$,

$B_{\boldsymbol{i},w\pm} \equiv B_{\boldsymbol{i}}$, $C_{\boldsymbol{i},w\pm} \equiv C_{\boldsymbol{i}}$). In this form, we have (3.122) with:

$$a_{\boldsymbol{i}}^{(1)} = \lambda \overline{K}_{\boldsymbol{i}} \left(\frac{A_{\boldsymbol{i}}}{k} + \frac{B_{\boldsymbol{i}}}{2} + \frac{C_{\boldsymbol{i}} k}{4} \right) \tag{3.125a}$$

$$a_{\boldsymbol{i}}^{(2)} = \lambda \overline{K}_{\boldsymbol{i}} \left(\frac{2A_{\boldsymbol{i}}}{k} - \frac{C_{\boldsymbol{i}} k}{2} \right) \tag{3.125b}$$

$$a_{\boldsymbol{i}}^{(3)} = \lambda \overline{K}_{\boldsymbol{i}} \left(\frac{B_{\boldsymbol{i}}}{2} - \frac{A_{\boldsymbol{i}}}{k} - \frac{C_{\boldsymbol{i}} k}{4} \right) \tag{3.125c}$$

With the above parameters, (3.122) reduces to (3.101) in the case of $A = C = 0$ and $B_{\boldsymbol{i}} = \gamma_{\boldsymbol{i}}$.

Numerical examples

In order to demonstrate conservation of numerical energy with these discrete boundary conditions, we carry out a simulation using the dome-shaped room pictured in Fig. 3.7. We set $h = 0.2\,\text{m}$ and $\lambda = \sqrt{1/3}$, and the initial conditions are: $f_0 = 0$ and g_0 is a spatial Gaussian with variance $0.1\,\text{m}^2$, centered about the middle of the domain. Fig. 3.19 shows the numerical energy quantities from this simulation, along with the normalised energy variations. The presence of stored energy ($\mathfrak{h}_\text{b}$) can be seen in Fig. 3.19(a), and it can be seen in Fig. 3.19(b) that the total accumulated energy varies on the order of machine epsilon. In this case we have used an accumulated power dissipated at the boundaries, defined as:

$$\overline{\mathfrak{q}}_\text{b}^n = k \sum_{m=0}^{n-1} \mathfrak{q}_\text{b}^m \tag{3.126}$$

In order to demonstrate the frequency-dependent behaviour of this boundary condition, we compute impulse responses using a box domain with dimensions as before, and for various triplets (A, B, C). These coefficients are arbitrarily chosen for illustrative purposes, with some corresponding to the examples in Fig. 2.3. We set $\lambda = \sqrt{1/3}$ and $h = 0.025\,\text{m}$ so that the sample rate is approximately 24 kHz. Fig. 3.20 shows spectra of the impulse responses, along with the triplets (A, B, C) in the captions. The first example is lossless (Fig. 3.20(a)), and the second and third examples (Figs. 3.20(b) and 3.20(c)) use coefficients taken from Section 2.1.3. It can be seen that these cases produce frequency-dependent losses that generally agree with the reflection coefficients plotted in Fig. 2.3, and any discrepancies are due to approximation effects which can manifest as frequency-warping (e.g., numerical dispersion). For the example in Fig. 3.20(d), the coefficients are chosen to produce a heavily damped resonance at 2.25 kHz, which is indeed where a notch appears in the spectrum.

A frequency-dependent impedance condition

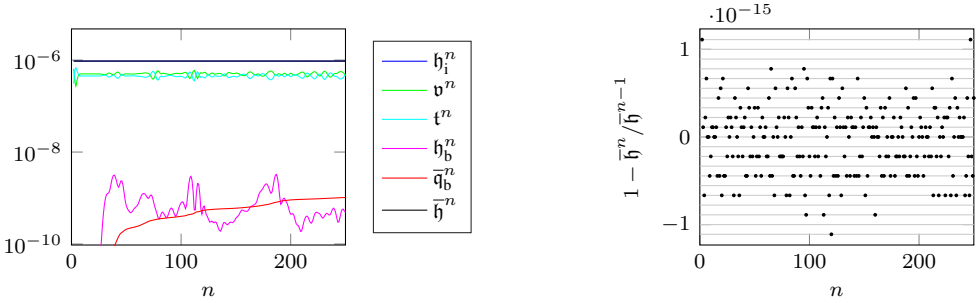
Let us now consider, as a continuous boundary condition, the frequency-dependent impedance condition (2.56). We should mention that this impedance condition has been considered previously in the literature (e.g., [51, 38, 216, 133]) and more will be said about other finite difference approaches to this boundary condition shortly.

In terms of Ψ and $v_\perp$ (2.56) is rewritten as:

$$\Psi_t = c\left(D\partial_t v_\perp + E v_\perp + F g\right), \quad \partial_t g = v_\perp, \quad \mathbf{x} \in \Gamma \tag{3.127}$$

where $D = D(\mathbf{x}), E = E(\mathbf{x}), F = F(\mathbf{x})$ (for $\mathbf{x} \in \Gamma$) and $D, E, F \geq 0$.

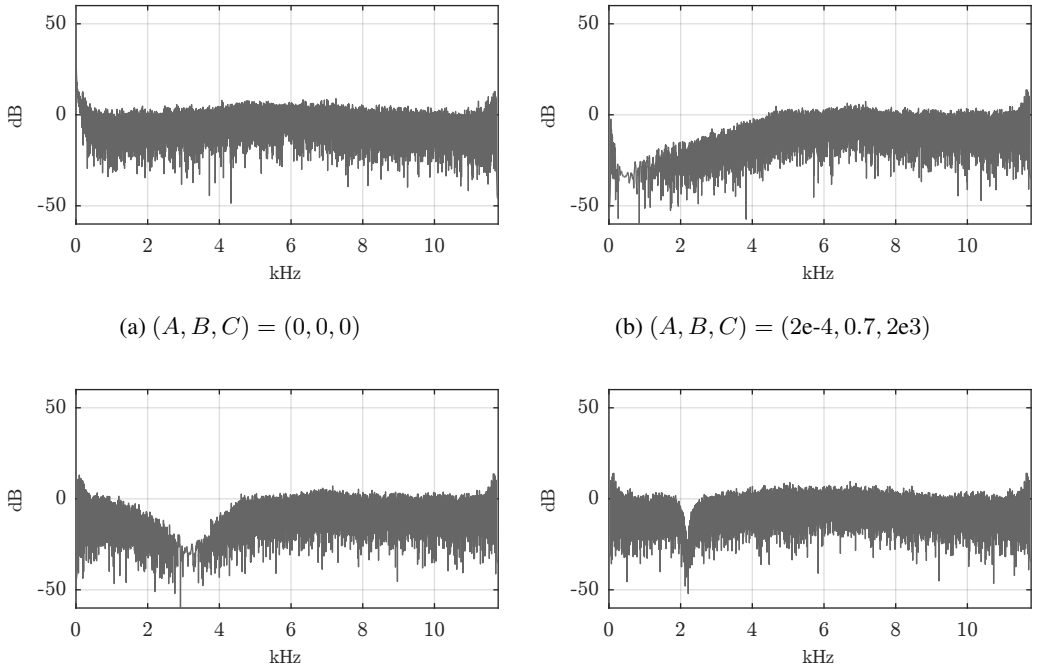
We define $D_{\boldsymbol{i},w\pm}, E_{\boldsymbol{i},w\pm}, F_{\boldsymbol{i},w\pm}$ as the projections of D, E, F onto Γ_h. The discrete form of (2.56)



(a) Energetic quantities (in Joules)

(b) Relative variations in total energy

Figure 3.19: Plots of numerical energy and variations in numerical energy for the frequency-dependent admittance boundaries with $A, B, C = (2\text{e-}4, 0.5, 1\text{e}5)$.



(a) $(A, B, C) = (0, 0, 0)$

(b) $(A, B, C) = (2\text{e-}4, 0.7, 2\text{e}3)$

(c) $(A, B, C) = (2\text{e-}4, 0.5, 1\text{e}5)$

(d) $(A, B, C) = (0.0118, 8.16, 2.35\text{e}6)$

Figure 3.20: Spectra of impulse responses taken from scheme under various choices of admittance boundary conditions.

proposed in [26] can be written in terms of $\underline{\Psi}_{\boldsymbol{i}}^n$ and $\underline{v}_{w,\boldsymbol{i}}^n$ for $w \in \pm\{x,y,z\}$ as:

$$\delta_{t\cdot}\underline{\Psi}_{\boldsymbol{i}}^n = c(D_{\boldsymbol{i},w+}\delta_{t+}\hat{\underline{v}}_{w,\boldsymbol{i}}^{n-\frac{1}{2}} + E_{\boldsymbol{i},w+}\mu_{t+}\hat{\underline{v}}_{w,\boldsymbol{i}}^{n-\frac{1}{2}} + F_{\boldsymbol{i},w+}\mu_{t+}\underline{g}_{w,\boldsymbol{i}}^{n-\frac{1}{2}})\,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},w+} = 1 \tag{3.128}$$

where $\hat{\underline{v}}_{w,\boldsymbol{i}}^{n-\frac{1}{2}}$ and $\underline{g}_{w,\boldsymbol{i}}^{n-\frac{1}{2}}$ are auxiliary variables related to $\underline{v}_{w,\boldsymbol{i}}^n$ by the following relations:

$$\delta_{t+}\underline{g}_{w,\boldsymbol{i}}^{n-\frac{1}{2}} = \mu_{t+}\hat{\underline{v}}_{w,\boldsymbol{i}}^{n-\frac{1}{2}} = \underline{v}_{w,\boldsymbol{i}}^n \tag{3.129}$$

This discrete boundary condition follows from the application of the bilinear transform (trapezoidal integration/differentiation) and the staggering in time results from the desire to have temporal derivatives be centered (in time) [26].

With this boundary condition, the boundary term in the energy balance becomes:

$$\mathfrak{b} = Z_0 \sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^2 \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+} \underbrace{\left(D_{\boldsymbol{i},w+}\delta_{t+}\hat{\underline{v}}_{w,\boldsymbol{i}} + E_{\boldsymbol{i},w+}\mu_{t+}\hat{\underline{v}}_{w,\boldsymbol{i}} + F_{\boldsymbol{i},w+}\mu_{t+}\underline{g}_{w,\boldsymbol{i}}\right)\left(\mu_{t+}\hat{\underline{v}}_{w,\boldsymbol{i}}\right)}_{W} \tag{3.130}$$

Making use of identity (3.43d) and the definition of $\underline{g}_{w,\boldsymbol{i}}^{n-\frac{1}{2}}$ ((3.129)) we have for the term W (labelled above):

$$W = \frac{1}{2}\delta_{t+}\left(D_{\boldsymbol{i},w+}(\hat{\underline{v}}_{w,\boldsymbol{i}})^2 + F_{\boldsymbol{i},w+}(\underline{g}_{w,\boldsymbol{i}})^2\right) + E_{\boldsymbol{i},w+}(\mu_{t+}\hat{\underline{v}}_{w,\boldsymbol{i}})^2 \tag{3.131}$$

Finally, we have $\mathfrak{b} = \delta_{t+}\mathfrak{h}_\mathfrak{b} + \mathfrak{q}_\mathfrak{b}$ with:

$$\mathfrak{h}_\mathfrak{b} = \frac{Z_0}{2}\sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^2 \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}\left(D_{\boldsymbol{i},w+}(\hat{\underline{v}}_{w,\boldsymbol{i}})^2 + F_{\boldsymbol{i},w+}(\underline{g}_{w,\boldsymbol{i}})^2\right) \tag{3.132}$$

$$\mathfrak{q}_\mathfrak{b} = Z_0 \sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^2 \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+} E_{\boldsymbol{i},w+}(\mu_{t+}\hat{\underline{v}}_{w,\boldsymbol{i}})^2 \tag{3.133}$$

and the energy balance is again of the form (3.121), with $\mathfrak{h}_\mathfrak{i}$ unchanged from (3.121). Thus, the scheme is stable with the above boundary conditions provided that $\lambda \leq \sqrt{1/3}$ and $D, E, F \geq 0$.

The update equation for boundary nodes can be obtained by writing out the intermediate step (from (3.129)):

$$\underline{g}_{w,\boldsymbol{i}}^{n+\frac{1}{2}} = \underline{g}_{w,\boldsymbol{i}}^{n-\frac{1}{2}} + \frac{k}{2}\left(\hat{\underline{v}}_{w,\boldsymbol{i}}^{n+\frac{1}{2}} + \hat{\underline{v}}_{w,\boldsymbol{i}}^{n-\frac{1}{2}}\right) \tag{3.134}$$

which can be substituted into (3.128) to obtain:

$$\hat{\underline{v}}_{w+,\boldsymbol{i}}^{n+\frac{1}{2}} = \alpha_{\boldsymbol{i},w+}\left(\frac{1}{2ck}(\underline{\Psi}_{\boldsymbol{i}}^{n+1} - \underline{\Psi}_{\boldsymbol{i}}^{n-1}) + \left(\frac{D_{\boldsymbol{i},w+}}{k} - \frac{E_{\boldsymbol{i},w+}}{2} - \frac{kF_{\boldsymbol{i},w+}}{4}\right)\hat{\underline{v}}_{w,\boldsymbol{i}}^{n-\frac{1}{2}} - F_{\boldsymbol{i},w+}\underline{g}_{w,\boldsymbol{i}}^{n-\frac{1}{2}}\right) \tag{3.135}$$

where

$$\alpha_{\boldsymbol{i},w+} = \left(\frac{D_{\boldsymbol{i},w+}}{k} + \frac{E_{\boldsymbol{i},w+}}{2} + \frac{kF_{\boldsymbol{i},w+}}{4}\right)^{-1} \tag{3.136}$$

Finally, (3.135) can be used to eliminate any ghost points for the update of $\underline{\Psi}_{\boldsymbol{i}}^n$ for $\boldsymbol{i} \in \mathcal{I}_\mathfrak{b}$, resulting in

the boundary update:

$$\underline{\Psi}_{\boldsymbol{i}}^{n+1} = \frac{1}{1+a_{\boldsymbol{i}}}\Bigg((2-\lambda^2 K_{\boldsymbol{i}})\underline{\Psi}_{\boldsymbol{i}}^{n} + (a_{\boldsymbol{i}}-1)\underline{\Psi}_{\boldsymbol{i}}^{n-1} + \lambda^2\tilde{Q}_{\boldsymbol{i}}^{n} \\ - c\lambda \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}\alpha_{\boldsymbol{i},w+}\left(D_{\boldsymbol{i},w+}\underline{\hat{v}}_{w,\boldsymbol{i}}^{n-\frac{1}{2}} - \frac{kF_{\boldsymbol{i},w+}}{2}\underline{g}_{w,\boldsymbol{i}}^{n-\frac{1}{2}}\right)\Bigg) \tag{3.137}$$

where

$$a_{\boldsymbol{i}} = \frac{\lambda}{4}\sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}\alpha_{\boldsymbol{i},w+} \tag{3.138a}$$

The scheme functions as an explicit update by carrying out the above steps in the following order: (3.137), (3.135), (3.134). Note that three extra variables, $\underline{\hat{v}}_{w,\boldsymbol{i}}^{n+\frac{1}{2}}$ and $\underline{\hat{v}}_{w,\boldsymbol{i}}^{n-\frac{1}{2}}$ and $\underline{g}_{w,\boldsymbol{i}}^{n-\frac{1}{2}}$, must be stored for each boundary face (the update (3.134) can overwrite in place).

A more simplified update results if we take one triplet (D,E,F) per boundary node (a simplifying assumption), i.e., where $D_{\boldsymbol{i},w\pm} \equiv D_{\boldsymbol{i}}$, $E_{\boldsymbol{i},w\pm} \equiv E_{\boldsymbol{i}}$, $F_{\boldsymbol{i},w\pm} \equiv F_{\boldsymbol{i}}$. In this simplified case, we have $\underline{\hat{v}}_{w,\boldsymbol{i}}^{n-\frac{1}{2}} = \underline{\hat{v}}_{\boldsymbol{i}}^{n-\frac{1}{2}}$ and $\underline{g}_{w,\boldsymbol{i}}^{n-\frac{1}{2}} = \underline{g}_{\boldsymbol{i}}^{n-\frac{1}{2}}$ as auxiliary variables, and the update for a boundary node becomes:

$$\underline{\Psi}_{\boldsymbol{i}}^{n+1} = \frac{1}{1+a_{\boldsymbol{i}}}\left((2-\lambda^2 K_{\boldsymbol{i}})\underline{\Psi}_{\boldsymbol{i}}^{n} + (a_{\boldsymbol{i}}-1)\underline{\Psi}_{\boldsymbol{i}}^{n-1} + \lambda^2\tilde{Q}_{\boldsymbol{i}}^{n} - c\lambda\overline{K}_{\boldsymbol{i}}\alpha_{\boldsymbol{i}}\left(D_{\boldsymbol{i}}\underline{\hat{v}}_{\boldsymbol{i}}^{n-\frac{1}{2}} - \frac{kF_{\boldsymbol{i}}}{2}\underline{g}_{\boldsymbol{i}}^{n-\frac{1}{2}}\right)\right) \tag{3.139}$$

where $\alpha_{\boldsymbol{i}} \equiv \alpha_{\boldsymbol{i},w\pm}$ and

$$a_{\boldsymbol{i}} = \frac{\lambda}{4}\overline{K}_{\boldsymbol{i}}\alpha_{\boldsymbol{i}} \tag{3.140}$$

This is followed by the updates for auxiliary variables:

$$\underline{\hat{v}}_{\boldsymbol{i}}^{n+\frac{1}{2}} = \alpha_{\boldsymbol{i}}\left(\frac{1}{2ck}(\underline{\Psi}_{\boldsymbol{i}}^{n+1} - \underline{\Psi}_{\boldsymbol{i}}^{n-1}) + \left(\frac{D_{\boldsymbol{i}}}{k} - \frac{E_{\boldsymbol{i}}}{2} - \frac{kF_{\boldsymbol{i}}}{4}\right)\underline{\hat{v}}_{\boldsymbol{i}}^{n-\frac{1}{2}} - F_{\boldsymbol{i}}\underline{g}_{\boldsymbol{i}}^{n-\frac{1}{2}}\right) \tag{3.141}$$

$$\underline{g}_{\boldsymbol{i}}^{n+\frac{1}{2}} = \underline{g}_{\boldsymbol{i}}^{n-\frac{1}{2}} + \frac{k}{2}\left(\underline{\hat{v}}_{\boldsymbol{i}}^{n+\frac{1}{2}} + \underline{\hat{v}}_{\boldsymbol{i}}^{n-\frac{1}{2}}\right) \tag{3.142}$$

All of the above frequency-dependent boundary updates can be re-expressed for a second-order scheme in pressure simply by replacing $\underline{\Psi}_{\boldsymbol{i}}^{n}$ with $\underline{p}_{\boldsymbol{i}}^{n}$, but the auxiliary variables $\underline{\hat{v}}_{\boldsymbol{i}}^{n+\frac{1}{2}}$ and $\underline{g}_{\boldsymbol{i}}^{n+\frac{1}{2}}$ then take on alternative interpretations (one time derivative higher and with units changed appropriately).

Numerical example: energy conservation

In order to confirm the property of conservation of numerical energy for these impedance conditions, we carry out a simulation using the dome-shaped room pictured in Fig. 3.7, setting $h = 0.2\,\mathrm{m}$ and $\lambda = \sqrt{1/3}$. The initial conditions here are: $f_0 = 0$ and g_0 is a spatial Gaussian with variance $0.1\,\mathrm{m}^2$, centered about the middle of the domain. The presence of stored energy at the boundaries ($\mathfrak{h}_\mathrm{b}$) and numerical energy conservation can be seen in Fig. 3.21.

Numerical example: room impulse responses

Impulse responses can also be computed in order to demonstrate the frequency-dependent behaviour of this boundary condition. First, however, it is worth comparing the behaviour of admittance conditions to impedance conditions. Two impulse responses computed with similar boundary parameters are

displayed in Fig. 3.22. It can be seen that even in the lossy case, the frequency-dependent impedance conditions permit a linear drift in the acoustic velocity potential, whereas admittance conditions do not permit the same sort of drift. This is explained by the reflection coefficient analyses in Section 2.1.3: at DC the impedance conditions become rigid walls, whereas the admittance conditions become "soft walls" at DC. For this reason, we apply DC-filtering (as carried out for the rigid wall examples previously). It is important to remember, this is not an instability (the corresponding pressure output has no such growth) since the numerical energy of the system remains bounded.

With that said, the obtained impulse response spectra (DC-filtered) are displayed in Fig. 3.23 for the same triplets used previously in the admittance case (as in Fig. 3.20). As in the case of admittance conditions, an agreement can be observed between the obtained impulse responses and the reflection coefficients in Fig. 2.4. Also for the example in Fig. 3.23(d), a notch appears at the theoretical resonant frequency of 2.25 kHz, as expected from the impedance parameters used.

Remarks on other frequency-dependent impedance finite difference boundary conditions

As mentioned previously, this simple RLC locally-reactive boundary model has been discretised by many in the literature [51, 38, 216, 133]. In the early work of Botteldooren, this boundary condition was discretised using an explicit update for the velocity boundary node, however, a stability analysis was only provided for the special case of a flat-wall aligned with grid axes. Also, the stability conditions provided were more restrictive at the boundaries than on the interior, and they were dependent on D, F in a manner not analogous to the continuous problem (i.e., other than $D, F \geq 0$).

Kowalczyk and van Walstijn presented what seemed to be a less restrictive set of finite difference boundary conditions, using centered spatial differences and the bilinear transform for the time-dependent parts [133]. The impedance boundary condition presented here uses (essentially) the same discretisation for time-dependent parts (the bilinear transform is an obvious and safe choice), but it differs in the way it treats the spatial discretisation. A detailed relationship between these boundary conditions and those appearing in [133] can be found in Appendix A.

Summarising what is detailed in Appendix A, the boundary conditions proposed in [133] can be rewritten in a form similar to the one presented here, except that the non-centered difference $\delta_{w\pm}$ is used here for the boundary conditions, and in [133], a centered difference ($\delta_{w\cdot} = \frac{1}{2}(\delta_{w+} + \delta_{w-})$) is used wherever possible. The choice of a centered difference can have important implications for numerical stability, as will be discussed in more detail in Section 5.4.

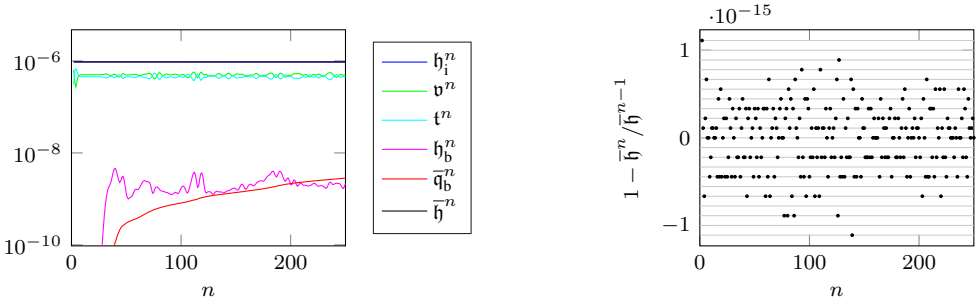
3.3 Incorporating viscothermal losses

Viscothermal loss, or more generally, sound absorption in air, is a crucial feature in any room acoustics model that aims to cover the full audible range of frequencies. A simple explicit approximation to the viscothermal wave equation (2.63) can be obtained by an additional term to the simplest scheme [302]:

$$\delta_{tt}\underline{\Psi}_{\boldsymbol{i}}^n = c^2\delta_{\Delta}\underline{\Psi}_{\boldsymbol{i}}^n + c\eta\delta_{\Delta}\delta_{t-}\underline{\Psi}_{\boldsymbol{i}}^n \tag{3.143}$$

The resulting update equation can also be seen as the explicit lossless update plus an additional term:

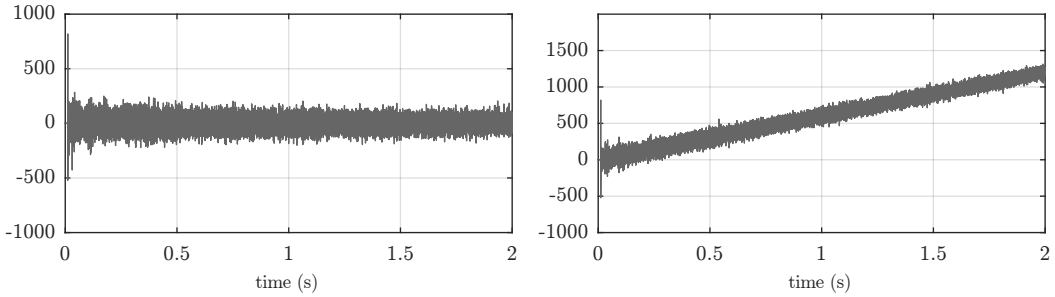
$$\underline{\Psi}_{\boldsymbol{i}}^{n+1} = (2 - 6\lambda^2)\underline{\Psi}_{\boldsymbol{i}}^n - \underline{\Psi}_{\boldsymbol{i}}^{n-1} + \lambda^2 Q_{\boldsymbol{i}}^n + W_{\boldsymbol{i}}^n \tag{3.144}$$



(a) Energetic quantities (in Joules)

(b) Relative variations in total energy

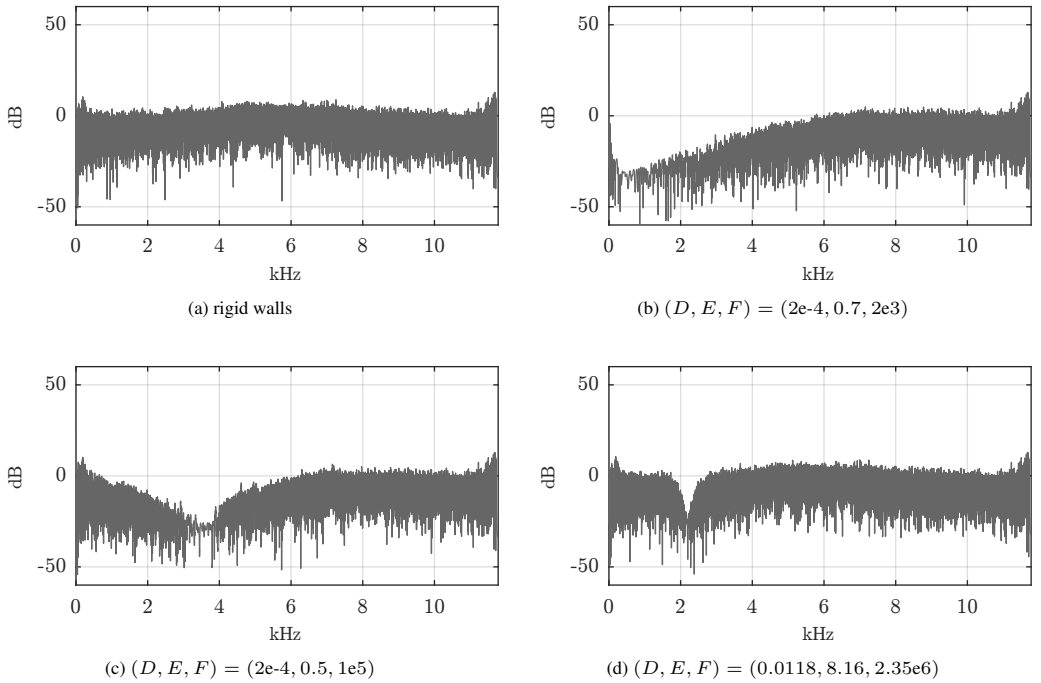
Figure 3.21: Plots of numerical energy and variations in numerical energy for the frequency impedance boundaries with $D, E, F = (2\text{e-}4, 0.5, 1\text{e}5)$.



(a) $(A, B, C) = (2\text{e-}4, 0.7, 2\text{e}3)$

(b) $(D, E, F) = (2\text{e-}4, 0.7, 2\text{e}3)$

Figure 3.22: Comparison of impulse responses ($\underline{\Psi}_{\mathrm{out}}^n$) obtained from admittance (left) and impedance (right) frequency-dependent boundary conditions. Vertical axes in m^2/s.



(a) rigid walls

(b) $(D, E, F) = (2\text{e-}4, 0.7, 2\text{e}3)$

(c) $(D, E, F) = (2\text{e-}4, 0.5, 1\text{e}5)$

(d) $(D, E, F) = (0.0118, 8.16, 2.35\text{e}6)$

Figure 3.23: Spectra of impulse responses (DC-filtered) taken from scheme under various choices of impedance boundary conditions.

where $Q_{\boldsymbol{i}}^n$ is defined as in (3.11), and where $W_{\boldsymbol{i}}$ represents the terms that are additional to the lossless update, defined as:

$$W_{\boldsymbol{i}}^n = \lambda\eta'(-6\underline{\Psi}_{\boldsymbol{i}}^n + 6\underline{\Psi}_{\boldsymbol{i}}^{n-1} + Q_{\boldsymbol{i}}^n - Q_{\boldsymbol{i}}^{n-1}) \tag{3.145}$$

and where $\eta' = \eta/h$ is a dimensionless numerical parameter related to the viscothermal losses. It is also convenient to re-express the scheme in the following form:

$$\underline{\Psi}_{\boldsymbol{i}}^{n+1} = a_1\underline{\Psi}_{\boldsymbol{i}}^n + a_2\underline{\Psi}_{\boldsymbol{i}}^{n-1} + a_3Q_{\boldsymbol{i}}^n + a_4Q_{\boldsymbol{i}}^{n-1} \tag{3.146}$$

where

$$a_1 = 2 - 6\lambda^2 - 6\lambda\eta'\,, \quad a_2 = 6\lambda\eta' - 1\,, \quad a_3 = \lambda^2 + \lambda\eta'\,, \quad a_4 = -\lambda\eta' \tag{3.147}$$

This scheme requires three states to be stored in memory, since $\underline{\Psi}_{\boldsymbol{i}}^{n+1}$ can no longer overwrite $\underline{\Psi}_{\boldsymbol{i}}^{n-1}$ in place. The two local sums, $Q_{\boldsymbol{i}}^n$ and $Q_{\boldsymbol{i}}^{n-1}$, can be computed at each time-step, or $Q_{\boldsymbol{i}}^n$ could be stored in memory for re-use as $Q_{\boldsymbol{i}}^{n-1}$ in the following time-step. In practice, this can be accomplished without the use of extra memory (beyond three full states).[26] The numerical dispersion relation of this scheme is virtually unchanged for $\eta \ll c/\omega$, so (for brevity) we do not present a dispersion analysis.[27]

To obtain a pressure output, a straightforward discretisation of (2.79) is:

$$\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}} = \rho\delta_{t-}\underline{\Psi}_{\boldsymbol{i}}^n - \rho c l_v \delta_\Delta \underline{\Psi}_{\boldsymbol{i}}^n \tag{3.148}$$

However, since this is only applied at the output, $\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}} = \rho\delta_{t-}\underline{\Psi}_{\boldsymbol{i}}^n$ is virtually identical to the above for $\eta \ll c/\omega$. The relationship between Ψ and $\mathbf{v}$ is unchanged in the viscothermal case [174] because we can assume that $\mathbf{v}$ is irrotational. We recall that in the absence of thermal effects (which are negligible in comparison to bulk viscosity or oxygen relaxation effects), one could also derive a viscous wave equation in the pressure [127], and a similar discretisation would be:

$$\delta_{tt}\underline{p}_{\boldsymbol{i}}^n = c^2\delta_\Delta\underline{p}_{\boldsymbol{i}}^n + c l_v \delta_\Delta \delta_{t-}\underline{p}_{\boldsymbol{i}}^n \tag{3.149}$$

3.3.1 Energy analysis

A stability condition for the viscothermal scheme (3.143) can be derived using energy methods. We note that this viscothermal model is a simplified form of a more general linearised Navier-Stokes system, and as such, the energetic quantity used here is just a positive semi-definite function of the state that is, strictly speaking, not the physical energy of the more general (linearised Navier-Stokes) system, but it is one that reduces to the correct energy of the lossless case when $\eta \to 0$.

The energy analysis here follows that of the lossless system. Thus, we take the inner product of (3.143) with $\frac{\rho}{c^2}\delta_{t\cdot}\underline{\Psi}_{\boldsymbol{i}}^n$, which arrives at the energy balance:

$$\delta_{t+}\mathfrak{h} = -\mathfrak{q}\,, \quad \mathfrak{h} = \mathfrak{v} + \mathfrak{t} \tag{3.150}$$

[26] This "trick" can be carried out by swapping pointers to the state arrays (in, e.g., the C programming language). We leave this as an exercise for the reader. The practical performance boosts that can be obtained from this trick are reported in [99] (published work of the current author).

[27] For a numerical dispersion analysis, see [99] (published work of the current author).

where $\mathfrak{v}$ is unchanged from (3.82), and $\mathfrak{t}$ and $\mathfrak{q}$ are:

$$\mathfrak{t} = \frac{\rho}{2} \sum_{w\in\{x,y,z\}} \langle \delta_{w+}\underline{\Psi}, e_{t-}\delta_{w+}\underline{\Psi}\rangle_{\mathbb{Z}^3_h} - \sum_{w\in\{x,y,z\}} \frac{\rho\eta k}{4c} \|\delta_{t-}\delta_{w+}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \tag{3.151a}$$

$$= \frac{\rho}{2} \sum_{w\in\{x,y,z\}} \left(\|\mu_{t-}\delta_{w+}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} - \left(\frac{k^2}{4} + \frac{\eta k}{2c}\right) \|\delta_{t-}\delta_{w+}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \right) \tag{3.151b}$$

$$\mathfrak{q} = \sum_{w\in\{x,y,z\}} \frac{\rho\eta}{c} \|\delta_{w+}\delta_{t\cdot}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \tag{3.151c}$$

where we have made use of the identity $\delta_{t-} = \delta_{t\cdot} - \frac{k}{2}\delta_{tt}$.

Continuing with a bound on the energetic term $\mathfrak{t}$, we have:

$$\mathfrak{t} \geq -\rho\left(\frac{k^2}{8} + \frac{\eta k}{4c}\right) \sum_{w\in\{x,y,z\}} \|\delta_{t-}\delta_{w+}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \geq -\frac{3\rho}{2c^2}\left(\frac{c^2k^2}{h^2} + \frac{2c\eta k}{h^2}\right) \|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \tag{3.152}$$

For a stored energy we have:

$$\mathfrak{h} \geq \frac{\rho}{2c^2}\left(1 - \frac{3c^2k^2}{h^2} - \frac{6c\eta k}{h^2}\right) \|\delta_{t-}\underline{\Psi}\|^2_{\mathbb{Z}^3_h} \tag{3.153}$$

which provides a condition on the spatial step:

$$h \geq \sqrt{3c^2k^2 + 6c\eta k} \implies \mathfrak{h} \geq 0 \tag{3.154}$$

and it is assumed that $\eta \leq ck/2$, which is satisfied by the assumption that $\eta \ll c/\omega$ for this model equation. If we define a dimensionless numerical parameter $\tau'_\eta = \tau_\eta/k$, where τ_η is a relaxation time $\tau_\eta = \eta/c$, then we can rewrite this condition as:

$$h \geq ck\sqrt{3(\tau'_\eta + 1)} \tag{3.155}$$

It is also worth rewriting the above condition in terms of the time-step. It can be worked out that an equivalent condition is:

$$k \leq \frac{1}{c}\left(\sqrt{h^2/3 + \eta^2} - \eta\right) \tag{3.156}$$

or in terms of the normalised parameter η', we have:

$$k \leq \frac{h}{c}\left(\sqrt{1/3 + \eta'^2} - \eta'\right) \tag{3.157}$$

In terms of the Courant number $\lambda = ck/h$, the above is:

$$\lambda \leq \sqrt{1/3 + \eta'^2} - \eta' \tag{3.158}$$

which reduces to the standard CFL condition for $\eta \to 0$. A von Neumann stability analysis can be used to arrive at the above conditions, but it is left out for brevity; see, e.g., [99] (by the current author).

3.3.2 A staggered first-order viscous scheme

It is worth mentioning out that a first-order staggered scheme in two spatial dimensions has appeared in the literature for sound propagation through an absorbing atmosphere [67]. It turns out that this

scheme is closely related to the above second-order scheme. The first-order system under consideration is (from [67]):

$$\frac{1}{\rho c^2}\partial_t p = -\nabla \cdot \mathbf{v} \tag{3.159a}$$

$$\rho \partial_t \mathbf{v} = -\nabla p - \mu' \nabla p_t \tag{3.159b}$$

where $\mu' = 4\mu/(3\rho c^2)$ is a viscosity term in units of seconds, and μ is the usual dynamic viscosity coefficient. Added forcing terms and terms identified as being heuristic [67, Eq. (18)] have been neglected here. These two equations can also be combined to give [67, Eq. (A17)] in the absence of heuristic terms and with constant density, which is (2.76) in the absence of the bulk viscosity.

The proposed staggered scheme, originally written in two spatial dimensions [67], but augmented here to three (in a straightforward manner) is, essentially:

$$\delta_{t+} \underline{p}_{\boldsymbol{i}}^n = -\kappa (\boldsymbol{\delta}_{\mathbf{x}-})^{\mathrm{T}} \underline{\mathbf{v}}_{\boldsymbol{i}+\frac{1}{2}}^{n+\frac{1}{2}} \tag{3.160a}$$

$$\rho \delta_{t-} \underline{\mathbf{v}}_{\boldsymbol{i}+\frac{1}{2}}^{n+\frac{1}{2}} = -(1 + \mu' \delta_{t-}) \boldsymbol{\delta}_{\mathbf{x}+} \underline{p}_{\boldsymbol{i}}^n \tag{3.160b}$$

where here, ρ and μ' are assumed to be constant, and recall that $\kappa = \rho c^2$. By applying δ_{t-} to (3.160a) and substituting in (3.160b), we have:

$$\delta_{tt} \underline{p}_{\boldsymbol{i}}^n = c^2 \delta_\Delta \underline{p}_{\boldsymbol{i}}^n + c^2 \mu' \delta_\Delta \delta_{t-} \underline{p}_{\boldsymbol{i}}^n \tag{3.161}$$

which is equivalent to (3.149) for the case of zero bulk viscosity ($\mu_{\mathfrak{b}} = 0$). This is also equivalent to (3.143) if one replaces η by $c\mu'$ and $\underline{p}_{\boldsymbol{i}}^n$ by $\underline{\Psi}_{\boldsymbol{i}}^n$.[28]

3.3.3 Boundary conditions for the viscothermal scheme

In this section we extend the boundary conditions presented for the lossless scheme to the viscothermal scheme. We consider a closed domain $\overline{\Omega}$, and the finite Cartesian grid is (again) denoted $\mathbb{G}_h$. The first steps in this energy analysis are as in the free-space case (Section 3.3.1), and we arrive at an energy balance of the form:

$$\delta_{t+} \mathfrak{h}_{\mathfrak{i}} = -\mathfrak{q}_{\mathfrak{i}} - \mathfrak{b}\,, \quad \mathfrak{h}_{\mathfrak{i}} = \mathfrak{v} + \mathfrak{t} \tag{3.162}$$

where $\mathfrak{v}$ is unchanged from (3.82), and $\mathfrak{t}$ and $\mathfrak{q}_{\mathfrak{i}}$ are now:

$$\mathfrak{t} = \frac{\rho}{2} \sum_{w \in \{x,y,z\}} \left(\|q_{w+} \delta_{w+} \mu_{t-} \underline{\Psi}\|_{\mathbb{G}_h}^2 - \left(\frac{k^2}{4} + \frac{\eta k}{2c} \right) \|q_{w+} \delta_{w+} \delta_{t-} \underline{\Psi}\|_{\mathbb{G}_h}^2 \right) \tag{3.163a}$$

$$\mathfrak{q}_{\mathfrak{i}} = \sum_{w \in \{x,y,z\}} \frac{\rho \eta}{c} \|q_{w+} \delta_{w+} \delta_{t\cdot} \underline{\Psi}\|_{\mathbb{G}_h}^2 \tag{3.163b}$$

Using the bound (3.89), the condition for $\mathfrak{h}_{\mathfrak{i}} \geq 0$ is unchanged from (3.154) or (3.156). For the boundary term $\mathfrak{b}$, we have now:

$$\mathfrak{b} = -\rho \sum_{\boldsymbol{i} \in \mathcal{I}_{\mathfrak{b}}} h^2 \sum_{w \in \pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+} (\delta_{t\cdot} \underline{\Psi}_{\boldsymbol{i}}) (\delta_{w+} \underline{\Psi}_{\boldsymbol{i}} + \tau_\eta \delta_{w+} \delta_{t-} \underline{\Psi}_{\boldsymbol{i}}) \tag{3.164}$$

[28] It follows that the stability condition derived in [67, Eq. (29)] is not the necessary and sufficient stability condition; it should be $k \leq c^{-1} \left(\sqrt{\mu'^2 + h^2/2} - \mu' \right)$ in the two-dimensional case, where in [67] μ' is written with the symbol μ.

At this point, it makes sense to redefine the outward normal velocity component $\underline{v}^n_{w,\boldsymbol{i}}$ as:

$$\underline{v}^n_{w,\boldsymbol{i}} = -\left(\delta_{w+}\underline{\Psi}^n_{\boldsymbol{i}} + \tau_\eta \delta_{w+}\delta_{t-}\underline{\Psi}^n_{\boldsymbol{i}}\right) \tag{3.165}$$

as we did in the continuous case in (2.82). With this redefinition, the boundary term is simply:

$$\mathfrak{b} = \rho \sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^2 \sum_{w\in\pm\{x,y,z\}} \overline{q}_{\boldsymbol{i},w+}(\delta_{t\cdot}\underline{\Psi}_{\boldsymbol{i}})(\underline{v}_{w,\boldsymbol{i}}) \tag{3.166}$$

Analogously to the continuous case in Section 2.1.4, all of the previously presented boundary conditions are unchanged in the viscothermal case, considering the more general $\underline{v}^n_{w,\boldsymbol{i}}$. The stability conditions are then the usual non-negativity conditions on A, B, C or D, E, F for $\mathfrak{b} \geq 0$, and for $\mathfrak{h}_\mathrm{i} \geq 0$ we require that k and h are set according to (3.154) or (3.156). The final boundary update equations invariably change, and they are provided below.

Frequency-independent boundaries

For the special case $\gamma_{\boldsymbol{i},w\pm} \equiv \gamma_{\boldsymbol{i}}$, one can work out the following update equation:

$$\underline{\Psi}^{n+1}_{\boldsymbol{i}} = \frac{1}{1+\lambda\overline{K}_{\boldsymbol{i}}\gamma_{\boldsymbol{i}}/2}\left(\left(2-\lambda^2 K_{\boldsymbol{i}}\right)\underline{\Psi}^n_{\boldsymbol{i}} + (\lambda\overline{K}_{\boldsymbol{i}}\gamma_{\boldsymbol{i}}/2-1)\underline{\Psi}^{n-1}_{\boldsymbol{i}} + \lambda^2\tilde{Q}^n_{\boldsymbol{i}} + \tilde{W}^n_{\boldsymbol{i}}\right) \tag{3.167}$$

where $\tilde{W}^n_{\boldsymbol{i}}$ is defined as:

$$\tilde{W}^n_{\boldsymbol{i}} = \lambda\eta'\left(K_{\boldsymbol{i}}(\underline{\Psi}^{n-1}_{\boldsymbol{i}} - \underline{\Psi}^n_{\boldsymbol{i}}) + \tilde{Q}^n_{\boldsymbol{i}} - \tilde{Q}^{n-1}_{\boldsymbol{i}}\right) \tag{3.168}$$

Note that this boundary update is the same (3.101) apart from the extra term $\tilde{W}^n_{\boldsymbol{i}}$. For $\gamma_{\boldsymbol{i}} \equiv \gamma$ we get essentially the boundary update proposed in [302].[29]

Frequency-dependent admittance boundaries

Admittance boundaries with a frequency-dependent behaviour, generalising the above conditions, are also formulated in a straightforward manner. After substituting the boundary condition into the update equation for boundary nodes, a boundary update results that is similar to (3.122):

$$\underline{\Psi}^{n+1}_{\boldsymbol{i}} = \frac{1}{1+a^{(1)}_{\boldsymbol{i}}}\left((2-\lambda^2 K_{\boldsymbol{i}} + a^{(2)}_{\boldsymbol{i}})\underline{\Psi}^n_{\boldsymbol{i}} + (a^{(3)}_{\boldsymbol{i}}-1)\underline{\Psi}^{n-1}_{\boldsymbol{i}} + \lambda^2\tilde{Q}^n_{\boldsymbol{i}} + \tilde{W}^n_{\boldsymbol{i}}\right) \tag{3.169}$$

where $a^{(1)}_{\boldsymbol{i}}, a^{(2)}_{\boldsymbol{i}}, a^{(3)}_{\boldsymbol{i}}$ are either given by (3.123) or (3.125).

Frequency-dependent impedance boundaries

It is also straightforward to extend the frequency-dependent impedance boundaries to the viscothermal case. In this case, the updates for auxiliary variables are unchanged (considering the more general $\underline{v}^n_{w,\boldsymbol{i}}$). In the simplified case (one triplet D, E, F per boundary point) we have $\hat{\underline{v}}_{w,\boldsymbol{i}} = \hat{\underline{v}}_{\boldsymbol{i}}$, and for the boundary

[29] We note a slight discrepancy between this update and the one presented in [302], where $K_{\boldsymbol{i}}\gamma_{\boldsymbol{i}}$ would be replaced by $\gamma_{\boldsymbol{i}}$, which leads to less absorbing at edges and corners—a simplification that was chosen for coding efficiency.

node update we have:

$$\underline{\Psi}_{\boldsymbol{i}}^{n+1} = \frac{1}{1+a_{\boldsymbol{i}}} \left((2 - \lambda^2 K_{\boldsymbol{i}}) \underline{\Psi}_{\boldsymbol{i}}^{n} + (a_{\boldsymbol{i}} - 1) \underline{\Psi}_{\boldsymbol{i}}^{n-1} + \lambda^2 \tilde{Q}_{\boldsymbol{i}}^{n} - c\lambda \overline{K}_{\boldsymbol{i}} \alpha_{\boldsymbol{i}} \left(D_{\boldsymbol{i}} \hat{\underline{v}}_{\boldsymbol{i}}^{n-\frac{1}{2}} - \frac{kF_{\boldsymbol{i}}}{2} \underline{g}_{\boldsymbol{i}}^{n-\frac{1}{2}} \right) + \tilde{W}_{\boldsymbol{i}}^{n} \right) \tag{3.170}$$

where $a_{\boldsymbol{i}}$ and $\alpha_{\boldsymbol{i}}$ are given by (3.140) and (3.136).

Numerical examples: energy conservation

The dome-shaped domain is used again to demonstrate energy conservation of the viscothermal scheme under frequency-dependent boundary conditions. For brevity, we only consider the impedance conditions. The parameters of the model and scheme are unchanged from the usual case, except here k is set according to (3.156) and $\eta = 2\text{e-}6\,\text{m}$. Plots of the energetic quantities and energy variations are shown in Fig. 3.24, demonstrating numerical energy conservation to machine accuracy, which includes the accumulated losses over the interior, represented by:

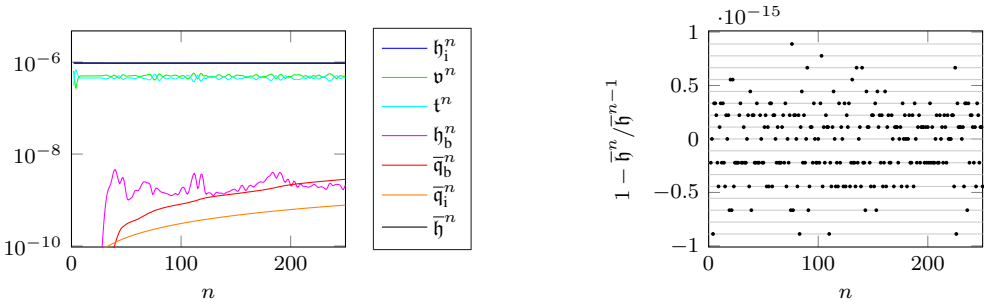
$$\overline{\mathfrak{q}}_{\mathrm{i}}^{n} = k \sum_{m=0}^{n-1} \mathfrak{q}_{\mathrm{i}}^{m} \tag{3.171}$$

Numerical examples: room impulse responses

An impulse response is computed for the box-shaped domain with $h = 0.0125\,\text{m}$, resulting in a sample rate of approximately 48 kHz. The viscosity parameter η is set to 2e-6 m, and impulse response spectra are displayed in Fig. 3.25 with rigid walls or with frequency-dependent impedance conditions at walls. Viscothermal dissipation is observed by a general low-pass characteristic in the frequency responses displayed in Figs. 3.25(b)–3.25(d). Also, one can observe frequency-dependent absorption from walls in low frequencies in Figs. 3.25(c) and 3.25(d). Finally, we note that the example in Fig. 3.25(d) reproduces an expected wall resonance at 2.25 kHz.

3.4 Summary

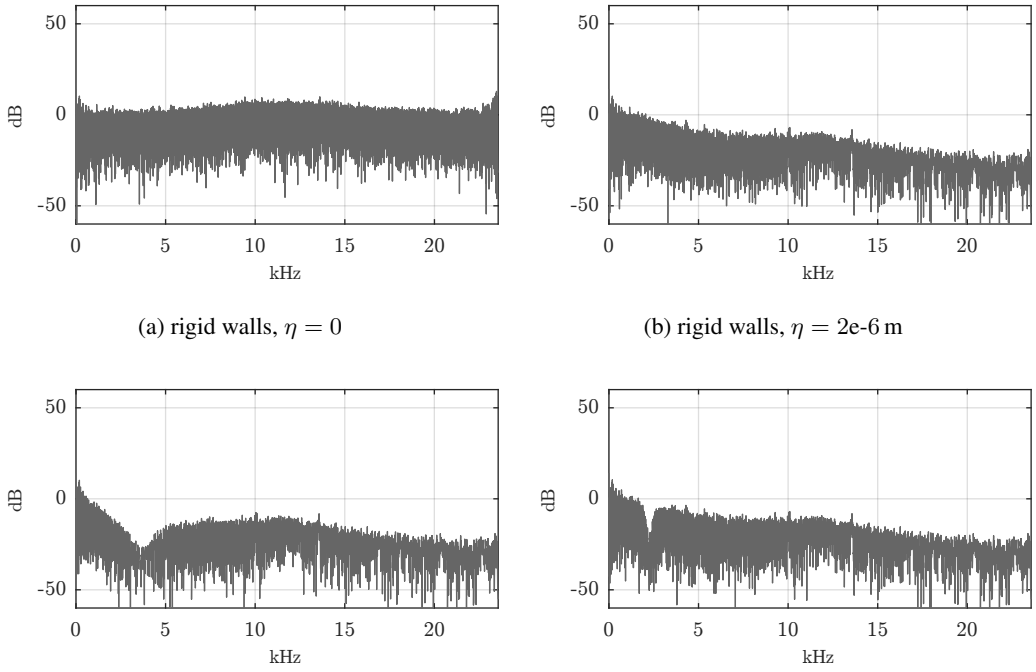
In this chapter we presented an overview of finite difference room acoustics simulation using the simplest Cartesian scheme in 3-D, including frequency-dependent boundary conditions and viscothermal losses in air. Second-order schemes and staggered, first-order schemes were discussed, and equivalences between these two formulations were established for a discrete boundary value problem that models simplified room acoustics. Boundary conditions were derived through energy-based stability analyses and accompanied by numerical examples demonstrating conservation of energy to machine precision. It is worth pointing out that the case of general impedance boundary conditions (i.e., those presented in Section 2.1.3) were not discretised in this chapter (for brevity), but they may be found in [102] (published work of the current author) adapted for the simple Cartesian finite difference scheme. General impedance boundary conditions will be treated in more detail, under a more general finite volume framework, in Chapter 6. Another topic that was discussed in this chapter is numerical dispersion in the simplest Cartesian scheme—an important problem that must be addressed for the efficient simulation of large-scale room acoustics. As such, the focus of the next chapter is on numerical dispersion, or more generally, on designing accurate finite difference schemes (beyond the simplest Cartesian scheme) for the wave equation in free-space.



(a) Energetic quantities (in Joules)

(b) Relative variations in numerical energy

Figure 3.24: Plots of numerical energy and variations in numerical energy for the frequency impedance boundaries with $D, E, F = (2\text{e-}4, 0.5, 1\text{e}5)$.



(a) rigid walls, $\eta = 0$

(b) rigid walls, $\eta = 2\text{e-}6\,\mathrm{m}$

(c) $(D, E, F) = (2\text{e-}4, 0.5, 1\text{e}5)$, $\eta = 2\text{e-}6\,\mathrm{m}$

(d) $(D, E, F) = (0.0118, 8.16, 2.35\text{e}6)$, $\eta = 2\text{e-}6\,\mathrm{m}$

Figure 3.25: Comparison of impulse response magnitude spectra, in the lossless case with rigid walls, and in the viscothermal case with rigid walls and frequency-dependent impedance conditions.

Chapter 4

Finite difference schemes for the wave equation in free space

The simple Cartesian scheme seen in Chapter 3 is but one of many possible finite difference schemes for the 3-D wave equation. In this chapter, we investigate wave equation schemes that are more complex, using stencils that reach out farther than nearest neighbouring points, and using regular (lattice-based) grids that are non-Cartesian. The goal of this chapter is to identify and compare accurate and efficient schemes for the wave equation in free space, with a particular focus on mitigating dispersion errors.

While room acoustic problems naturally call for three spatial dimensions, it will be informative to start by considering the wave equation in one and two spatial dimensions. Accordingly, in this chapter we consider the wave equation in an abstract form, where the variable of interest could take on various meanings. The variable of interest in this chapter will be denoted $u = u(\mathbf{x}, t)$, a continuous scalar function in time and d-dimensional space ($\mathbf{x} \in \mathbb{R}^d$) for $d \leq 3$. The spatial position will be $\mathbf{x} = (x, y, z)$ when $d = 3$, $\mathbf{x} = (x, y)$ when $d = 2$, and $\mathbf{x}$ is simply x when $d = 1$. The variable u will not be attributed to a particular physical quantity in this chapter, but it could represent, e.g., pressure, acoustic velocity potential, the transverse displacement of a membrane ($d = 2$) or a string ($d = 1$).

The wave equation in $d \leq 3$ can be written as:

$$\partial_t^2 u = c^2 \Delta u \tag{4.1}$$

where Δ is the Laplacian operation in $d \leq 3$ spatial dimensions:

$$(d = 1): \quad \Delta = \partial_x^2 \tag{4.2a}$$

$$(d = 2): \quad \Delta = \partial_x^2 + \partial_y^2 \tag{4.2b}$$

$$(d = 3): \quad \Delta = \partial_x^2 + \partial_y^2 + \partial_z^2 \tag{4.2c}$$

For well-posedness, we must also specify the initial conditions:

$$u(\mathbf{x}, 0) = f_0(\mathbf{x}), \qquad u_t(\mathbf{x}, 0) = g_0(\mathbf{x}) \tag{4.3}$$

It will also be useful to express the wave equation in terms of the wave operator (or the "d'Alembertian",

named after d'Alembert [147]), which can be written with the box symbol ($\Box$) as:

$$\Box := \partial_t^2 - c^2\Delta \tag{4.4}$$

and then the wave equation is simply:

$$\Box u = 0 \tag{4.5}$$

4.1 Preliminaries

Before embarking on an investigation of the many possible schemes for the wave equation in $d \leq 3$, it will be useful to discuss some preliminary topics that will be relevant to finite difference schemes in any number of spatial dimensions. Also, in this preliminary section we will introduce some concepts and notation that will be used throughout the rest of this chapter.

We start by discussing a hypothetical situation involving a solution to the wave equation that is to be sampled in both space and time, considering the constraints imposed by Cartesian sampling in space and time, linked through the ideal dispersion relation of the wave equation in a discrete setting. After this topic, we discuss the general form of explicit two-step schemes, to be seen throughout this chapter, and some general details for such schemes, including operation counts, numerical stability and dispersion, computational densities, and normalisations that will be of use to compare schemes which have different operation counts and computational densities.

4.1.1 Sampling considerations on Cartesian grids

In this section, we discuss the constraints imposed by multidimensional sampling theory [191] underlying any regular grid finite difference scheme. In particular, we discuss how such constraints relate to the dispersion relation of the wave equation in a discrete setting, and subsequently, how this relates to finite difference schemes that aim to discretise the ideal dispersion relation of the wave equation. The purpose of this discussion is to gain some insight into appropriate choices of spatial and temporal grid spacings (i.e., the Courant number) for Cartesian grid–based finite difference schemes, as this relates to the overall computational costs (keeping in mind additional constraints imposed by numerical stability).

Sampling considerations for solutions to the wave equation

We start by considering a solution to the d-D wave equation that is to be sampled using a Cartesian grid in both space and time. By Fourier theory, we can say that this solution is composed of plane-wave components with temporal angular frequencies ω in rad/s and spatial (angular) frequencies $\boldsymbol{\beta} \in \mathbb{R}^d$, where $\boldsymbol{\beta} = \beta_x$ for $d = 1$, $\boldsymbol{\beta} = (\beta_x, \beta_y)$ for $d = 2$, and $\boldsymbol{\beta} = (\beta_x, \beta_y, \beta_z)$ for $d = 3$, and $\boldsymbol{\beta}$ is also known as the vector wavenumber, or the wave vector, whose components have units rad/m. Under the framework of Fourier theory, we can also describe the solution in terms of its spatial or temporal bandwidth (its support in frequency domains).

By multidimensional sampling considerations [191], we would expect this solution to only be composed of plane waves with spatial frequencies limited to $\boldsymbol{\beta} \in [-\pi/h, \pi/h]^d$, in order to avoid *aliasing errors* in a subsequent continuous reconstruction from discrete samples. It thus makes sense to consider a normalised wave vector $\boldsymbol{\beta}_h = \boldsymbol{\beta}h \in \mathbb{B}$, where the region $\mathbb{B} := [-\pi, \pi]^d$ will be known as the Cartesian *wavenumber cell* [191]. Also, we define the normalised (scalar) wavenumber: $\beta_h = \beta h = |\boldsymbol{\beta}|h$. Let

us also define a normalised angular frequency $\omega_k \in [-\pi, \pi]$, since we would expect ω to be confined to $\omega \in [-\pi/k, \pi/k]$ by temporal sampling considerations.

Let us turn our attention to the dispersion relation of the wave equation. In terms of these normalised variables we have:

$$\omega = \pm c|\boldsymbol{\beta}| \quad \Longrightarrow \quad \omega_k(\boldsymbol{\beta}_h) = \pm\lambda|\boldsymbol{\beta}_h| \tag{4.6}$$

where $\lambda = ck/h$. We should make it clear that in this context, the Courant number, λ, is not linked to stability or convergence; it is just a dimensionless grid ratio.

It will also be useful to further manipulate this dispersion relation by introducing the normalised frequency $\omega_{h/c} = \omega h/c = \omega_k/\lambda$. We have then:

$$\omega_k(\boldsymbol{\beta}_h) = \pm\lambda|\boldsymbol{\beta}_h| \quad \Longrightarrow \quad \omega_{h/c}(\boldsymbol{\beta}_h) = \pm|\boldsymbol{\beta}_h| \tag{4.7}$$

This ideal dispersion relation is plotted in Fig. 4.1 for the two-dimensional wave equation ($d = 2$), as a set of contours inside the square wavenumber cell $\mathbb{B}$, and for $\omega_k \geq 0$.

It should be evident from Fig. 4.1 that $\beta_h = \pi$ is the limiting wavenumber that can be discretised *isotropically* (independently of the wave direction) on this 2-D Cartesian grid. Meanwhile, the maximum wavenumber supported on the 2-D Cartesian grid is $\beta_h = \sqrt{2}\pi$, which appears at the corners of the wavenumber cell $\mathbb{B} = [-\pi, \pi]^2$. Extending these properties to d spatial dimensions, the maximum wavenumber supported on a d-D Cartesian grid is $\beta_h = \sqrt{d}\pi$, but the limiting wavenumber that can be discretised isotropically on a Cartesian grid is always $\beta_h = \pi$. This is an important constraint that is imposed simply by the nature of the Cartesian grid.

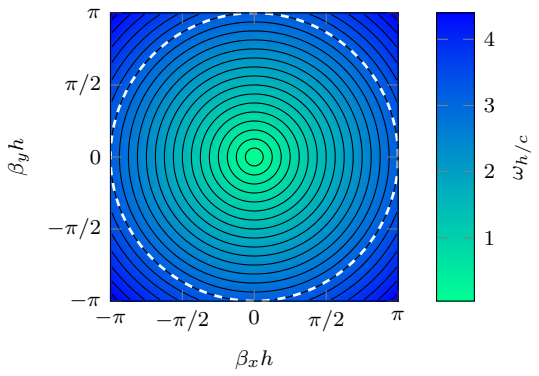


Figure 4.1: Contour plot of ideal dispersion relation $\omega_{h/c}(\boldsymbol{\beta}_h)$ for $\boldsymbol{\beta}_h \in [-\pi, \pi]^2$. Contour lines denote increments of $\pi/16$. The white dotted line denotes $|\boldsymbol{\beta}_h| = \pi$.

Let us now explore the temporal sampling constraints associated to these extreme cases, (i.e., $\beta_h \in \{\pi, \sqrt{d}\pi\}$). First, we assume that our solution has a d-dimensional spatial bandwidth with a maximum wavenumber denoted by $\beta_{\max}$ (in rad/m), and let us also define a normalised maximum wavenumber $\beta_{h,\max} = h\beta_{\max}$, which has the associated temporal frequency $f_{\max} = \frac{c\beta_{h,\max}}{2\pi h}$. We assume that $\beta_{h,\max}$ falls within the extreme cases, $\beta_{h,\max} \in [\pi, \sqrt{d}\pi]$, and we assume that c is fixed.

In order to temporally sample this hypothetical discretised-in-space solution, we need to choose an appropriate time-step k (or sample rate $1/k$). Of course, by temporal sampling considerations we would first consider the limitation $k \leq \frac{1}{2f_{\max}}$. It will be insightful to see how this temporal sampling condition relates to the underlying Courant number.

By (4.7) and temporal sampling considerations (Shannon-Nyquist sampling theory) [235], we can

derive the following temporal sampling constraint:[1]

$$k \leq \frac{1}{2f_{\max}} = \frac{\pi h}{c\beta_{h,\max}} \quad \Longrightarrow \quad \lambda \leq \frac{\pi}{\beta_{h,\max}} \tag{4.8}$$

Thus, the Shannon-Nyquist sampling criterion can be rewritten as a condition on the Courant number, and this condition depends on the value of $\beta_{h,\max}$, which is further constrained by spatial sampling considerations ($\beta_{h,\max} \leq \sqrt{d}\pi$).

Consider then the following two extreme cases, leading to constraints on the Courant number:

$$\beta_{h,\max} = \sqrt{d}\pi \quad \Longrightarrow \quad \lambda \leq \sqrt{1/d}\,, \qquad \beta_{h,\max} = \pi \quad \Longrightarrow \quad \lambda \leq 1 \tag{4.9}$$

Interestingly, with the first extreme case we have recovered the CFL condition for convergence in the simplest finite difference scheme for the d-D wave equation [60, 86]. However, for the most part, it would be desirable to have $\beta_{h,\max} \leq \pi$ such that the discretisation of the solution would be independent of the orientation of the grid. Furthermore, in a room acoustics setting (considering now a hypothetical d-D room), modal densities in the discretised solution would tend to be incorrect above the temporal frequency associated with $\beta_h = \pi$ [25]. Thus, it will be assumed henceforth that $\beta_{h,\max} \leq \pi$.

Computational costs of sampling solutions to the wave equation

Let us now investigate the computational (storage) costs of this sampling scheme. We start by defining the spatial and spatiotemporal densities for the Cartesian grid:

$$\rho_{\text{SD}} = \frac{1}{h^d}\,, \quad \rho_{\text{STD}} = \frac{1}{kh^d} \tag{4.10}$$

The spatial density, ρ_{SD}, is the number of spatial samples required per unit volume of space, while the spatiotemporal density, ρ_{STD}, is the number of spatiotemporal samples required per unit volume of *space and time* (space-time). We can proceed with this sampling scheme either by initially choosing h or initially choosing k. Let us discuss the implications for computational costs of these initial choices, starting with the spatial step.

If we initially set h according to spatial sampling considerations (keeping in mind the previous discussion about isotropic sampling), we would choose $h = \frac{\pi}{\beta_{\max}}$, which implies that:

$$\rho_{\text{SD}} = \left(\frac{\beta_{\max}}{\pi}\right)^d, \quad \rho_{\text{STD}} = \frac{c}{\lambda}\left(\frac{\beta_{\max}}{\pi}\right)^{d+1} \tag{4.11}$$

In this case it would, at first, make sense to choose λ as large as possible to keep ρ_{STD} as small as possible. However, we must also satisfy temporal sampling considerations (recast as a constraint on the Courant number). With (4.8) in mind, we find that the ideal choice is $\lambda = 1$. It will be important to keep this ideal choice (for an ideal scenario) in mind for the corresponding finite difference schemes, to be discussed shortly.

Let us now consider a scenario where k is initially chosen, rather than h, as is common in the audio-related literature. In such settings, k is generally chosen to give a sample rate $1/k$ that satisfies temporal sampling considerations. Considering $f_{\max}$, the maximum frequency of interest in Hz (e.g. 20 kHz),

[1] It could be argued that these constraints should be "strictly less than", but we ignore such details for this discussion.

one might choose $k = \frac{1}{2f_{\max}}$ (by temporal sampling arguments alone), in which case we have:

$$k = \frac{1}{2f_{\max}} \quad \Longrightarrow \quad \rho_{\text{SD}} = \left(\frac{2\lambda f_{\max}}{c}\right)^2, \quad \rho_{\text{STD}} = (2f_{\max})^{d+1}\left(\frac{\lambda}{c}\right)^d \tag{4.12}$$

In this case, it would seem that choosing λ small would be beneficial to keep both storage costs and overall computational costs down. However, this contradicts with the reasoning that led to identifying $\lambda = 1$ as being the optimal choice. The source of this contradiction lies in the fact that the naive choice of $k = \frac{1}{2f_{\max}}$ neglects the associated constraints of spatial sampling.

As mentioned previously, $\beta_h = \pi$ is the maximum wavenumber that may be discretised on a Cartesian grid regardless of the wave direction. Thus, it makes sense to align $f_{\max}$ with $\beta_h = \pi$ through the ideal dispersion relation. This leads to the following "optimal" choice for the time-step:

$$k = \frac{\lambda}{2f_{\max}} \tag{4.13}$$

This choice may seem counter-intuitive at first, but it ensures that frequencies up to $f_{\max}$ can be represented in every wave direction (isotropically). Under this choice of k, we have:

$$\rho_{\text{SD}} = \left(\frac{2f_{\max}}{c}\right)^d = \left(\frac{\beta_{\max}}{\pi}\right)^d, \quad \rho_{\text{STD}} = \frac{(2f_{\max})^{d+1}}{\lambda c^d} = \frac{c}{\lambda}\left(\frac{\beta_{\max}}{\pi}\right)^{d+1} \tag{4.14}$$

and we have recovered the values of ρ_{SD} and ρ_{STD} derived in (4.11). In other words:

$$h = \frac{\pi}{\beta_{\max}} \quad \Leftrightarrow \quad k = \frac{\lambda}{2f_{\max}} \tag{4.15}$$

Thus, whether we set h or k initially, one would ideally choose $\lambda = 1$ for such a (hypothetical) Cartesian-based sampling scheme.

Sampling considerations in Cartesian finite difference schemes

So far we have only discussed a hypothetical and idealised sampling scenario; let us now discuss how these idealised constraints relate to a finite difference scheme. First we should mention that such sampling considerations would only apply in the limit of an *infinitely accurate* finite difference scheme—to which we ultimately aim in the forthcoming sections—so the previous discussion only provides *partial* guidance into the choices of h, k and λ for a conventional finite difference scheme.

On the other hand, from the above discussion we have arrived at a theoretical upper limit to the "valid bandwidth" that could be obtained in a Cartesian grid–based scheme—this upper limit is $\frac{\lambda}{2k}$ in Hz, for a scheme with some fixed $\lambda \leq 1$. As such, rather than choosing $k = \frac{1}{2f_{\max}}$ as a starting point for a finite difference scheme (as is commonly seen in the audio-related literature), it would make sense to choose $k = \frac{\lambda}{2f_{\max}}$, provided that $\lambda \leq 1$, or simply $h = \frac{\pi}{\beta_{\max}}$. As an example, for the simplest scheme in 3-D with $\lambda = \sqrt{1/3}$, the "valid bandwidth" cannot go beyond $0.5F_s/\sqrt{3} \approx 0.29F_s$, where $F_s = 1/k$. Thus it would make sense to start with $k = (2\sqrt{3}f_{\max})^{-1}$ if the time-step is to be set initially.

Then again, for finite difference schemes there are more critical considerations to take into account: *accuracy* (e.g., dispersion error) and *numerical stability*. If we were to choose h or k solely based only on sampling considerations then we would be ignoring the presence of potentially significant dispersion

errors. Taking into account some measure of dispersion errors, one could generally choose h as:

$$h = \frac{\pi}{\sigma \beta_{\max}} \tag{4.16}$$

where $\sigma \geq 1$ can be called an "oversampling factor". The parameter σ is meant to ensure that frequencies up to $f_{\max}$ (or wavenumbers up to $\beta_{\max}$) are *accurately* reproduced, with some error criterion appropriate to the application of interest. For example, $\sigma = 1$ would result in the minimal "two points per wavelength", $\sigma = 2$ would give "four points per wavelength", and so on.

Keeping in mind the discussion about the transition frequency $\frac{\lambda}{2T}$, one could equivalently choose time-step k to be

$$k = \frac{\lambda}{2\sigma f_{\max}} \tag{4.17}$$

assuming $\lambda/\sigma \leq 1$.

Let us now consider the spatial and spatiotemporal densities of a finite difference scheme where h or k are chosen according to (4.16) or (4.17) with $\sigma > 1$. Let ρ'_{SD} and ρ'_{STD} now represent the spatial density and spatiotemporal densities in the *ideal* sampling scenario, i.e., $\rho'_{\text{SD}} = \left(\frac{2f_{\max}}{c}\right)^d$, $\rho'_{\text{STD}} = \frac{(2f_{\max})^{d+1}}{\lambda c^d}$. Then we have:

$$\rho_{\text{SD}} = \sigma^d \rho'_{\text{SD}}\,, \quad \rho_{\text{STD}} = \sigma^{d+1} \rho'_{\text{STD}} \tag{4.18}$$

where now ρ_{SD} represents the spatial density of the grid function, and ρ_{STD}, represents the number of pointwise updates per unit volume of space and time required by the finite difference scheme. Clearly by the above relations, oversampling is to be avoided as much as possible in order to keep computational costs down, especially for the case of room acoustics in three spatial dimensions ($d = 3$).

As for the Courant number, the ideal choice in a Cartesian grid–based scheme is not necessarily $\lambda = 1$, since for most explicit schemes, λ will be constrained for stability as:

$$\lambda \leq \lambda_{\max} \tag{4.19}$$

Furthermore, it is often the case that $\lambda_{\max} < 1$ (e.g., $\lambda_{\max} = \sqrt{1/3}$ for the simplest scheme in 3-D). Generally then, it would make sense to choose $\lambda = \lambda_{\max}$ to keep computational costs as low as possible. This dovetails nicely with the tendency of numerical dispersion to be best when $\lambda = \lambda_{\max}$ for a given scheme, as we saw in Section 3.1.4 for the simplest scheme in 3-D.

It is also worth noting that in many schemes to be seen in this chapter, $\lambda = \lambda_{\max}$ is not necessarily an optimal choice and $\lambda < \lambda_{\max}$ can allow for a smaller oversampling factor σ, which will have a greater impact on reducing computational costs (for $d > 1$) than choosing λ as large as possible. In such cases, the "optimal" Courant number will depend on the error tolerance desired for a given application.

4.1.2 Explicit two-step finite difference schemes for the wave equation

In this section we present a general form (notation) for the many schemes to be investigated in this chapter.

Grid functions

Unless otherwise specified, the finite difference schemes in this chapter will act on a grid function $\underline{u}^n_{\boldsymbol{i}}$, which is approximation to the continuous solution at spatiotemporal grid points $(\boldsymbol{i}h, nk)$, i.e., $\underline{u}^n_{\boldsymbol{i}} \approxeq u(\boldsymbol{i}h, nk)$ with $\boldsymbol{i} \in \mathbb{Z}^d$ and $n \in \mathbb{Z}^+$. Additionally in this chapter, under a slight abuse of notation, we

will make use of a *continuous approximation* $\underline{u} = \underline{u}(\mathbf{x}, t)$, which can be seen as an extension of the grid function $\underline{u}_{\boldsymbol{i}}^n$ to the continuous domain. The approximation $\underline{u}(\mathbf{x}, t)$ is defined such that $\underline{u}_{\boldsymbol{i}}^n = \underline{u}(\mathbf{x}, t)$ at $\mathbf{x} = \boldsymbol{i}h, t = nk$. In other words, $\underline{u}_{\boldsymbol{i}}^n$ can be seen as the spatiotemporally sampled version of $\underline{u}(\mathbf{x}, t)$. Generally, $\underline{u}(\mathbf{x}, t)$ will only be known on the grid, but in theory it can be fully and uniquely determined from the discrete grid function $\underline{u}_{\boldsymbol{i}}^n$ using Fourier interpolation.

Two-step explicit schemes: general form

The majority of the finite difference schemes to be seen in this chapter will be explicit schemes of the following general form:

$$\delta_{\square} \underline{u}_{\boldsymbol{i}}^n = 0 \,, \quad \delta_{\square} = \delta_{tt} - c^2 \delta_{\Delta} \tag{4.20}$$

where $\delta_{\square}$ represents a discrete approximation to the d'Alembertian operator defined by (4.4), and where δ_{Δ} represents a discrete approximation to the d-D Laplacian (i.e., a discrete Laplacian). The second-order accurate time-difference operator δ_{tt} (defined in Section 3.1.1) will be used throughout this chapter, as this keeps memory to a minimum (two states per grid point). The particular form of the finite difference operator δ_{Δ} will change in each of the schemes to be seem later, and this operator will have an accuracy that is at least second-order and may contain free parameters. Such free parameters may be used for optimisation purposes, or to achieve higher orders of accuracy in space—and possibly, also time.[2] Also, the underlying grid may change; e.g., hexagonal grids are employed in Section 4.4, and face- and body-centered cubic grids are employed in Section 4.6.

Practical details

Let us define the normalised discrete Laplacian $\delta_{\Delta}^h = h^2 \delta_{\Delta}$. The update recursion for (4.20) can then be expressed as:

$$\underline{u}_{\boldsymbol{i}}^{n+1} = \lambda^2 \delta_{\Delta}^h \underline{u}_{\boldsymbol{i}}^n + 2\underline{u}_{\boldsymbol{i}}^n - \underline{u}_{\boldsymbol{i}}^{n-1} \tag{4.21}$$

which requires two states of memory ($\underline{u}_{\boldsymbol{i}}^{n+1}$ can overwrite $\underline{u}_{\boldsymbol{i}}^{n-1}$). It is worth noting that such schemes can always be rewritten as:

$$u_{\boldsymbol{i}}^{n+1} = \sum_{l=1}^{K} \xi_l u_{\boldsymbol{i}+\boldsymbol{v}_l}^n + \xi_0 u_{\boldsymbol{i}}^n - u_{\boldsymbol{i}}^{n-1} \tag{4.22}$$

Here, the underlying $(K+1)$-point stencil in the chosen discrete Laplacian δ_{Δ} is represented by K vectors in $\mathbb{Z}^d$: $\boldsymbol{v}_1 \ldots, \boldsymbol{v}_K$ and the origin ($\boldsymbol{v}_0 = \mathbf{0}$). The associated stencil weights are $\xi_0, \xi_1, \ldots, \xi_K$. For example, the seven-point stencil (seen in Fig. 3.1) used in the simplest Cartesian scheme for the 3-D wave equation has $K = 6$ and:

$$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_6\} = \{\pm\hat{\mathbf{e}}_x, \pm\hat{\mathbf{e}}_y, \pm\hat{\mathbf{e}}_z\} \,, \quad \xi_0 = (2 - 6\lambda^2) \,, \quad \xi_l = \lambda^2 \,, \quad l = 1, \ldots, 6 \tag{4.23}$$

On a computer, each of these individual pointwise update writes to one location in memory (for $\underline{u}_{\boldsymbol{i}}^{n+1}$) and accesses $K+2$ locations in memory (pertaining to the grid function). The number of arithmetic operations for each update can vary depending on the particular implementation and the choice of the Courant number. When fused-multiply add (FMA) or multiply-accumulate (MAC) operations are available (such as, e.g., on GPU architectures [307]), (4.22) is the minimal implementation, and it requires $K+2$ floating-point instructions for an arbitrary Courant number. One FMA instruction can be saved

[2] It may seem counter-intuitive at first to claim that we can raise the order of accuracy of a scheme in time by using a *spatial* finite difference operator, but this is indeed possible, through the use of modified equation methods [241], as will be seen later.

if $\xi_0 = 0$ (e.g., when $\lambda = \sqrt{1/3}$ for the simplest 3-D Cartesian scheme).

It is important to note that this explicit update recursion is parallelisable over the spatial grid; this allows one to make use of parallel computing architectures, such as GPUs [222, 304], multi-core CPUs [304], and FPGAs [177]. Another important feature is that there is an overlap between the memory reads required by the pointwise updates for neighbouring grid points, which makes the scheme amenable to memory caching; see, e.g., [103] (published work involving the current author). For more details on practical implementation of such schemes on a large scale, see [304].

Numerical dispersion and numerical stability

This general two-step explicit scheme also has a general form for its numerical dispersion and its numerical stability constraints. Considering a plane-wave trial solution, the finite difference operators δ_{tt} and δ_Δ transform to their respective Fourier symbols, acting as the following multipliers:

$$\delta_{tt}\underline{u}_{\boldsymbol{i}}^n = -\frac{4}{k^2}\sin^2(\omega_k/2)\underline{u}_{\boldsymbol{i}}^n\,, \quad \delta_\Delta \underline{u}_{\boldsymbol{i}}^n = -\frac{4}{h^2}G(\boldsymbol{\beta}_h)\underline{u}_{\boldsymbol{i}}^n\,, \quad \text{when } \underline{u}_{\boldsymbol{i}}^n = \mathrm{e}^{\mathrm{j}(n\omega_k + \boldsymbol{i}\cdot\boldsymbol{\beta}_h)} \tag{4.24}$$

for some $G(\boldsymbol{\beta}_h)$ specific to each discrete Laplacian, which may also be a function of additional free parameters.[3] This leads to a numerical dispersion relation of the form:

$$\sin^2(\omega_k/2) = \lambda^2 G(\boldsymbol{\beta}_h) \tag{4.25}$$

which is applicable to all explicit schemes of the form (4.20). The above dispersion relation defines a relationship $\omega_k = \omega_k(\boldsymbol{\beta}_h)$, which, by rearranging (4.25), can be written as:

$$\omega_k(\boldsymbol{\beta}_h) = 2\sin^{-1}\left(\pm\lambda\sqrt{G(\boldsymbol{\beta}_h)}\right) \tag{4.26}$$

Alternatively, we will also express the above in the following rescaled form:

$$\omega_{h/c}(\boldsymbol{\beta}_h) = \frac{1}{\lambda}\omega_k(\boldsymbol{\beta}_h) \tag{4.27}$$

This form allows for straightforward comparisons with the ideal dispersion relation: $\omega_{h/c}(\boldsymbol{\beta}_h) = \pm|\boldsymbol{\beta}_h|$, regardless of the chosen Courant number.

For numerical stability, a von Neumann analysis essentially requires that $\omega_k \in \mathbb{R}$.

$$0 \leq \lambda^2 G(\boldsymbol{\beta}_h) \leq 1 \tag{4.28}$$

If $G(\boldsymbol{\beta}_h)$ is itself independent of the Courant number, the above simplifies to the two following conditions:

$$G(\boldsymbol{\beta}_h) \geq 0 \tag{4.29a}$$

$$\lambda \leq \lambda_{\max} = \left(\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)\right)^{-1/2} \tag{4.29b}$$

where $\lambda_{\max}$ is specific to each scheme (since $G(\boldsymbol{\beta}_h)$ is specific to each scheme). Since $G(\boldsymbol{\beta}_h)$ may contain free parameters (assuming λ is not one of them), the first condition, which enforces negative semi-definiteness in the discrete Laplacian, may impose constraints on those free parameters. Further-

[3] For a simplified notation, we leave out subscripts relating to additional free parameters.

more, $\lambda_{\max}$ may also be a function of these free parameters.[4]

If the scheme is written in the form (4.22) and one has that the stencil weights are all non-negative, i.e., when $\xi_l \geq 0$ for $l = 0, \ldots, K$, the scheme can be said to be "concretely passive" [21] (or here, "passive" for short). For such schemes, stability is easily determined, and is in fact guaranteed provided that the scheme is consistent with the wave equation, but passive schemes are not necessarily the best choices in terms of minimising numerical dispersion [21]. More will be said about this in Section 4.3.4.

We also note that temporal frequencies produced by these schemes do not necessarily follow sampling considerations ($-\pi \leq \omega_k \leq \pi$). By (4.26), we have that:

$$\max_{\boldsymbol{\beta}_h} \omega_k(\boldsymbol{\beta}_h) \leq \pi \tag{4.30}$$

provided that $\lambda \leq \lambda_{\max}$. By (4.26), the above can only be satisfied with equality for $\lambda = \lambda_{\max}$. The quantity $\max_{\boldsymbol{\beta}_h} \omega_k(\boldsymbol{\beta}_h)$ is sometimes known as the *cutoff frequency* of the scheme, and it represents the angular frequency above which the dispersion relation can only be satisfied for *complex* wavenumbers pertaining to *evanescent waves* [294, 276]. Such evanescent solutions to the wave equation generally do not play a role in the lossless operation of these finite difference schemes, and as such will not be considered in this study.[5] However, the cutoff frequencies themselves have been of interest to the audio and room acoustics community [131], and thus will be calculated for some of the schemes to follow.

Numerical phase velocity and dispersion errors

An important quantity of interest in this chapter will be the numerical phase velocity. Considering only $\omega \geq 0$, we can define the (relative) numerical phase velocity as:

$$\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) = \frac{\omega_{h/c}(\boldsymbol{\beta}_h)}{|\boldsymbol{\beta}_h|} \tag{4.31}$$

It is important to note that the above calculation is only meaningful for $\boldsymbol{\beta}_h$ within the appropriate wavenumber cell (in most cases, the usual $\mathbb{B} = [-\pi, \pi]^d$). The appropriate wavenumber cell will be identified for each scheme, as it is not always the usual Cartesian one.

A relative error in the dispersion relation, and in the numerical phase velocity, is defined as:

$$\varepsilon_{\text{rel. err.}}(\boldsymbol{\beta}_h) = \tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1 \tag{4.32}$$

Thus, if $\varepsilon_{\text{rel. err.}} > 0$, the wave travels too quickly, and if $\varepsilon_{\text{rel. err.}} < 0$ the wave travels too slowly, and by consistency (and stability) $\varepsilon_{\text{rel. err.}} \to 0$ as $|\boldsymbol{\beta}_h| \to 0$.

Additionally, it will be of interest to define a wavenumber quantity under which the *absolute* relative dispersion error is below a certain threshold: $|\varepsilon_{\text{rel. err.}}| \leq X\%$ for some $X \geq 0$. As such, we define the quantity $\beta_{h,X\%} \leq \pi$ as follows:

$$\beta_{h,X\%} = \beta'_h, \quad \text{such that} \quad |\varepsilon_{\text{rel. err.}}(\boldsymbol{\beta}_h)| \leq \frac{X}{100} \quad \text{for all} \quad |\boldsymbol{\beta}_h| \leq \beta'_h \tag{4.33}$$

Henceforth, this quantity will be known as a *critical wavenumber* and $X\%$ will be known as the *error*

[4] Generally, when the number of free parameters is greater than two, it becomes rather difficult to rewrite (4.29) as a general function of free parameters, at which point it is helpful to make use of a computer and mathematical software.

[5] This type of evanescent behaviour can arise next to a so-called "hard source" [230], but we will not consider hard-source excitations in this thesis.

threshold.[6] In particular, the critical wavenumbers $\beta_{h,1\%}$ and $\beta_{h,2\%}$ will be identified for the schemes to come, as a simple means to make comparisons between schemes. These critical wavenumbers can be used as a starting point to measure a "valid bandwidth" in such schemes, but later we will consider additional normalisations for computational costs. The term "wideband accuracy" will be used to refer to the accuracy of a schemes in terms of critical wavenumbers, since with these metrics in mind one can start to talk about how accurately a scheme reproduces accurate wave propagation over a wide bandwidth of frequencies.

Normalising for computational costs

In this chapter we inevitably have to face the problem of comparing schemes that have different optimal Courant numbers, and thus operate at different sample rates $(1/k)$ for the same grid spacing h. To this end, it will help to introduce normalisations for the spatiotemporal density of a scheme, otherwise known as the "computational density" [21, 291, 96]. While these quantities have already been defined for regular Cartesian grids, we will redefine them here, introducing some additional parameters to represent the use of non-Cartesian grids. As such, let us rewrite the spatial and spatiotemporal densities of a finite difference scheme for the d-D wave equation as:

$$\rho_{\text{SD}} = \frac{\mu}{h^d}\,, \quad \rho_{\text{STD}} = \frac{\mu}{kh^d} \tag{4.34}$$

Here we have included the additional parameter μ, which represents the spatial density of the grid when $h = 1$. For Cartesian grids, $\mu = 1$, but for some non-Cartesian regular grids to be seen in this chapter, we will have $\mu \neq 1$. In order to normalise the spatial density (and thus memory costs) across schemes on Cartesian grids, it suffices to fix the grid spacing h. In order to also include non-Cartesian grid–based schemes into such comparisons, we can choose h to be:

$$h = \mu^{1/d} h^\star \quad \Longrightarrow \quad \rho_{\text{SD}} = \frac{1}{(h^\star)^d} \tag{4.35}$$

where $h^\star$ is fixed across schemes, and thus spatial densities can be held constant across schemes on Cartesian and non-Cartesian grids.

As for the spatiotemporal density, which depends on both the chosen Courant number and the spatial density of the grid, we can rewrite it as:

$$\rho_{\text{STD}} = \frac{c\mu}{\lambda h^{d+1}} \tag{4.36}$$

In order to normalise for spatiotemporal densities, we can fix a length h^* across schemes, and choose h for each scheme as

$$h = (\mu/\lambda)^{1/(d+1)} h^* \quad \Longrightarrow \quad \rho_{\text{STD}} = \frac{c}{(h^*)^{d+1}} \tag{4.37}$$

Keep in mind, when two schemes with the same K (stencil size $K+1$) have the same spatiotemporal density, they require the same number of FLOPs or FLOP instructions per unit volume of space-time.

In order to incorporate these normalisations into the critical wavenumbers that will be used to compare schemes, it suffices to define additional, rescaled critical wavenumbers:

$$\beta_{h^\star,X\%} = \mu^{-1/d}\beta_{h,X\%}\,, \quad \beta_{h^*,X\%} = (\mu/\lambda)^{-1/(d+1)}\beta_{h,X\%} \tag{4.38}$$

[6] Of course, it is not necessarily the case that this critical wavenumber exists for a given $X\%$.

These critical wavenumbers, in particular $\beta_{h^*,1\%}$ ($\beta_{h,1\%}$ for Cartesian schemes) and $\beta_{h^*,1\%}$ will be used to provide comparisons for the various schemes to be seen for $d > 1$.

Relative computational efficiencies

Another useful method of comparison for schemes operating at different sample rates (or grid densities) is to use the "relative computational efficiency" (RCE), as defined by van Walstijn and Kowalczyk [291]. Before defining this RCE measure here, it is worth briefly explaining what it represents. The RCE is a measure of the efficiency of a scheme in terms of the computational density (spatiotemporal density) required to maintain some error threshold up to some frequency, relative to that of some reference scheme. An obvious choice for the reference scheme is the simplest Cartesian scheme.

With that said, we define the RCE as follows:

$$\text{RCE-}X\% = \left(\frac{\beta_{h^*,X\%}}{\beta_{h^*,X\%,\text{ref.}}}\right)^{d+1} \tag{4.39}$$

where $\beta_{h^*,X\%,\text{ref.}}$ is the critical wavenumber $\beta_{h^*,X\%}$ for the chosen reference scheme, and it is assumed that $\beta_{h^*,X\%} \leq \pi$ and $\beta_{h^*,X\%,\text{ref.}} \leq \pi$. That these quantities increase to the $(d+1)$th power in $(\beta_{h^*,X\%}/\beta_{h^*,X\%,\text{ref}})$ reflects the fact that the spatiotemporal density of the scheme scales with h^{d+1}. In another sense, this RCE represents the increase (or decrease) in computational density that would be required by the reference scheme in order to attain the same level of accuracy as another scheme (in terms of the absolute relative dispersion error) up to some fixed frequency. It is important to note that as the RCE is only a function of the chosen error threshold (in percent-error), and consequently is independent of the frequency. As such, it can be defined using the normalised critical wavenumber (as done here) or using analogous critical temporal frequencies (as in [291]).

The RCE measure does not take into account the number of operations (FLOPs, FLOP instructions, and memory reads) required by the pointwise updates of varying schemes, so it also will be informative to define a relative computational efficiency that takes into account some measure of operation counts in order to compare schemes with different stencil sizes. Since $K + 2$ is the number of memory reads required for a stencil that reaches out to K spatial neighbours, as well as being the number of FLOP instructions required by each pointwise update (when FMA instructions are available), we can use $K+2$ as an additional weighting to the RCE measure. We should note, however, that compute times will not necessarily scale with $K + 2$ in a direct manner (this will also depend on the particular computing architecture being used), so this is simply a starting point to incorporate a measure of operations into the RCE.

With that said, we define the measure "RCE-$X\%$-opts" as:

$$\text{RCE-}X\%\text{-opts} = (\text{RCE-}X\%) \cdot \left(\frac{K_{\text{ref.}} + 2}{K + 2}\right) \tag{4.40}$$

Since we will use the simplest Cartesian scheme as a reference, $K_{\text{ref.}} = 2d$.

Having sufficiently discussed these preliminary topics that will apply to $d \in \{1, 2, 3\}$, we move on to finite difference schemes for the 1-D wave equation.

4.2 Finite difference schemes for the 1-D wave equation

The wave equation in 1-D is a simple starting point for the derivation of more advanced schemes in higher dimensions, and it will be useful to introduce concepts relating to higher-order accuracy, targeting errors in space and/or time, and various other effects that should be taken into consideration when deriving finite difference schemes in higher dimensions. In this section, the variable of interest will be written simply as $u = u(x,t)$, and the 1-D wave equation as:

$$\Box u = 0\,, \quad \Box = \partial_t^2 - c^2\partial_x^2 \tag{4.41}$$

Here we use the grid function $\underline{u}_i^n \approxeq u(ih, nk)$, where k, h are respectively the time-step and grid spacing, as before, and $i \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$. All of the preliminary discussion presented earlier applies to these schemes by setting $d = 1$, replacing $\boldsymbol{i}$ with i, $\mathbf{x}$ with x, and $\boldsymbol{\beta}$ with β. For the most part, the focus is on accuracy and numerical dispersion, so we will ignore initial conditions.

4.2.1 The simplest scheme

The simplest scheme for the 1-D wave equation [60] can be written as:

$$\delta_\Box \underline{u}_i^n = 0\,, \quad \delta_\Box = \delta_{tt} - c^2\delta_{xx} \tag{4.42}$$

where $\delta_\Box$ represents a discrete d'Alembertian operator, and δ_{tt} and δ_{xx} were defined in Section 3.1.1. Applying $\delta_\Box$ to $\underline{u}$ reveals an order of accuracy in space and time that is *at least* second-order:

$$\delta_\Box \underline{u} = \Box \underline{u} + \mathcal{O}(h^2) + \mathcal{O}(k^2) \tag{4.43}$$

When the Courant number $\lambda = ck/h$ is fixed, as is commonly the case, the above can be equivalently written as:

$$\delta_\Box \underline{u} = \Box \underline{u} + \mathcal{O}(h^2) = \Box \underline{u} + \mathcal{O}(k^2) \tag{4.44}$$

showing that the *global* accuracy of the scheme follows the lowest order accuracy in either time or space; thus the scheme is said to be second-order accurate.

Expanding out the finite difference operators and solving for $\underline{u}_i^{n+1}$ leads to the update equation:

$$\underline{u}_i^{n+1} = (2 - 2\lambda^2)\underline{u}_i^n + \lambda^2(\underline{u}_{i-1}^n + \underline{u}_{i+1}^n) - \underline{u}_i^{n-1} \tag{4.45}$$

and it is well-known that this scheme has the CFL condition $\lambda \leq 1$, which is also the stability condition obtained from a von Neumann analysis [185] and from an energy analysis [24].

In this section we make use of the normalised frequencies:

$$\omega_k := \omega k\,, \quad \beta_h := \beta h \tag{4.46}$$

where ω is the temporal angular frequency (in rad/s) and β is the wavenumber (in rad/m). Considering a plane-wave trial solution: $\underline{u}_i^n = \mathrm{e}^{\mathrm{j}(n\omega_k + i\beta_h)}$, the numerical dispersion relation (of the form (4.26)) is:

$$\sin^2(\omega_k/2) = \lambda^2 \sin^2(\beta_h/2) \tag{4.47}$$

or rewritten for ω_k:

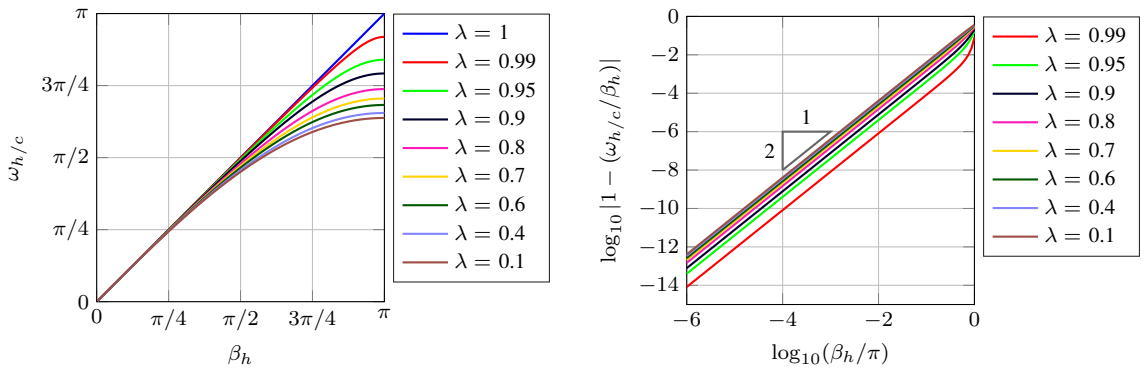
$$\omega_k = 2\sin^{-1}\left(\pm\lambda\sin(\beta_h/2)\right) \tag{4.48}$$

In terms of the normalised temporal frequency $\omega_{h/c}$, the numerical dispersion relation is:

$$\omega_{h/c}(\beta_h) = \frac{2}{\lambda}\sin^{-1}(\pm\lambda\sin(\beta_h/2)) \tag{4.49}$$

which can then be compared to the ideal dispersion relation $\omega_{h/c} = \pm\beta_h$.

The numerical dispersion relation, $\omega_{h/c}(\beta_h)$, for the simplest scheme is shown in Fig. 4.2(a) for various choices of the Courant number λ. It can be seen in the figure that the dispersion relation approaches the ideal (a unity slope) in the limit of $\beta_h \to 0$ for each $\lambda \leq \lambda_{\max}$. This reflects the consistency (and convergence) of the scheme. It can also be seen that the special case of $\lambda = 1$ results in a unity slope, matching that of the wave equation. This is a well-known special case that can provide an *exact* approximation for bandlimited solutions to the wave equation. This special case will be treated in more detail shortly; for now we focus on cases where $\lambda < 1$.



(a) $\omega_{h/c}(\beta_h)$ for various $\lambda \leq \lambda_{\max}$

(b) $|1 - \omega_{h/c}(\beta_h)/\beta_h|$ for various $\lambda \leq \lambda_{\max}$

Figure 4.2: Plots of numerical dispersion and numerical dispersion errors for simplest scheme in 1-D. (a) Normalised numerical dispersion relation, for various $\lambda \leq \lambda_{\max}$. (b) Relative dispersion error of 1-D scheme for $\lambda < \lambda_{\max}$, on logarithmic scale.

As mentioned previously, this scheme has an overall (global) accuracy (for $\lambda < 1$) that is second-order. If the scheme is indeed second-order accurate, it should be reflected in the dispersion error as $h, k \to 0$ (or equivalently, as $\beta_h, \omega_{h/c} \to 0$ for λ fixed). In order to see this property, we plot the relative error in the dispersion relation $|1 - \omega_{h/c}(\beta_h)/\beta_h|$ (which is also the error in the numerical phase velocity) on a logarithmic (log-log) scale in Fig. 4.2(b). Indeed, it can be seen that the dispersion errors, for various $\lambda < \lambda_{\max}$ have slopes of two for $\beta_h \ll \pi$, reflecting the second-order accuracy of the scheme.

An exact scheme

Now we return to the special case of $\lambda = 1$. For this choice of the Courant number, the simplest scheme reduces to:

$$\delta_\square = e_{t+} + e_{t-} - (e_{x+} + e_{x-}) \tag{4.50}$$

where $e_{t\pm}$ and $e_{x\pm}$ are temporal and spatial shifts (as defined in Section 3.1.1). With $\delta_\square \underline{u}_i^n = 0$, we get the update equation:

$$\underline{u}_i^{n+1} = \underline{u}_{i-1}^n + \underline{u}_{i+1}^n - \underline{u}_i^{n-1} \tag{4.51}$$

For some times, this scheme has been known to provide an exact approximation to the underlying continuous solution [5, 244], provided that the solution is sufficiently bandlimited in time and space. This special property of the scheme was seen in Fig. 4.2(a), and indeed for $\lambda = 1$, the dispersion relation of the scheme, (4.48), reduces to the ideal dispersion relation $\omega = \pm c\beta$.

Another way to show that the simplest scheme is exact for $\lambda = 1$ is to consider d'Alembert's travelling wave solution:

$$u(x,t) = \phi(x - ct) + \psi(x + ct) \tag{4.52}$$

Applying the discrete d'Alembertian to the above gives:

$$\begin{aligned}\delta_{\Box} u = {} & \phi(x - c(t+k)) + \psi(x + c(t+k)) + \phi(x - c(t-k)) + \psi(x + c(t-k)) \\ & + \phi((x+h) - ct) + \psi((x+h) + ct) + \phi((x-h) - ct) + \psi((x-h) + ct)\end{aligned} \tag{4.53}$$

and for $\lambda = 1$ we have $h = ck$ and all the terms on the right-hand side cancel to give $\delta_{\Box} u = 0$. Thus, d'Alembert's solution satisfies the difference equation $\delta_{\Box} u = 0$, and it follows that any solution to the exact scheme is also a solution to the 1-D wave equation. On the other hand, the inverse relationship does not necessarily follow; that is, not every solution to the 1-D wave equation is also a solution to the exact finite difference scheme for a fixed h or k. For this to be true, the solution has to be bandlimited and h, k have to be sufficiently small, which follows from Shannon's sampling theorem [235]. One can also express the exact scheme in an equivalent *digital waveguide* form, as a direct discretisation (sampling) of d'Alembert's travelling wave solution [245].

Transitional note on high-order accurate schemes in 1-D

Despite the existence of this simple and exact 1-D scheme, in the coming sections we will consider finite difference schemes beyond the simplest one. Such schemes must ultimately make use of more points (in time or space) than the minimum for such a second-order equation (three in space, three in time). The use of more than three steps in time will not be employed here, as it severely complicates stability analyses [7] (as will be discussed briefly in Section 4.2.3) and goes beyond the minimal memory requirements for such a second-order system (memory cost considerations become paramount in higher dimensions). Instead, we will focus on explicit two-step schemes that reach out to more than the usual two nearest neighbours on a regular grid of points. To this end, we will start by defining a general family of parametrised approximations to the 1-D Laplacian.

4.2.2 Parametrised approximations to the 1-D Laplacian

A second-order accurate discrete Laplacian that makes use of two neighbouring points p spatial steps away can be defined as:

$$\delta_{xx}^{(p)} = \frac{1}{p^2 h^2}(e_{t-}^p - 2 + e_{t+}^p) \quad \Longrightarrow \quad \delta_{xx}^{(p)} \underline{u}_i^n = \frac{1}{p^2 h^2}(\underline{u}_{i-p}^n - 2\underline{u}_i^n + \underline{u}_{i+p}^n) \tag{4.54}$$

where here p is a positive integer (not to be confused with pressure), and we have the special case $\delta_{xx} = \delta_{xx}^{(1)}$. It is straightforward to check that:

$$\delta_{xx}^{(p)} \underline{u} = \partial_x^2 \underline{u} + \mathcal{O}(h^2) \tag{4.55}$$

It follows that linear combinations of these operators can be used to construct other discrete Lapla-

cians with at least second-order accuracy. For example, we can consider $P > 1$ successive $(2P+1)$-point discrete Laplacians:

$$\sum_{p=1}^{P} \alpha_p \delta_{xx}^{(p)} \underline{u} = \partial_x^2 \underline{u} + \mathcal{O}(h^2)\,, \quad \text{provided that } \sum_{p=1}^{P} \alpha_p = 1 \tag{4.56}$$

where the constraint that the coefficients α_p add to unity is necessary for consistency. Henceforth, this consistency constraint will be implicitly assumed.

High-order spatial differences

It has been known for some time (since at least the work of Richardson [203]) that linear combinations of the $\delta_{xx}^{(p)}$ operators can be used to build approximations to the second-order spatial derivative with orders of accuracy higher than two. These second-order differences with high-order accuracy are ultimately special cases of (4.56). They will be more specifically denoted as:

$$\delta_{xx}^{[M]} = \sum_{m=1}^{M} \alpha_{M,m} \delta_{xx}^{(m)} \tag{4.57}$$

where $\alpha_{M,m}$ for $m = 1, \ldots M$ represent the coefficients that (uniquely) establish the following equality

$$\delta_{xx}^{[M]} \underline{u} = \partial_x^2 \underline{u} + \mathcal{O}(h^{2M}) \tag{4.58}$$

These coefficients can be worked out through Taylor expansions—by hand or with a computer program [84]—or using polynomial (Lagrange) interpolation [151]. In either case, one arrives at a general linear system for which an analytic solution is well-known [56, 7, 53]. Following [7], the coefficients $\alpha_{M,m}$ are given by the expression:

$$\alpha_{M,m} = (-1)^{m-1} \frac{2(M!)^2}{(M-m)!(M+m)!}\,, \quad m = 1, \ldots, M \tag{4.59}$$

For example, $\alpha_{1,1} = 1$, which implies that $\delta_{xx}^{[1]} = \delta_{xx}$. For $M = 2, 3, 4$ we have:

$$\delta_{xx}^{[2]} \underline{u} = \left(\frac{4}{3} \delta_{xx}^{(1)} - \frac{1}{3} \delta_{xx}^{(2)} \right) \underline{u} \qquad = \partial_x^2 \underline{u} + \mathcal{O}(h^4) \tag{4.60a}$$

$$\delta_{xx}^{[3]} \underline{u} = \left(\frac{3}{2} \delta_{xx}^{(1)} - \frac{3}{5} \delta_{xx}^{(2)} + \frac{1}{10} \delta_{xx}^{(3)} \right) \underline{u} \qquad = \partial_x^2 \underline{u} + \mathcal{O}(h^6) \tag{4.60b}$$

$$\delta_{xx}^{[4]} \underline{u} = \left(\frac{8}{5} \delta_{xx}^{(1)} - \frac{4}{5} \delta_{xx}^{(2)} + \frac{8}{35} \delta_{xx}^{(3)} - \frac{1}{35} \delta_{xx}^{(4)} \right) \underline{u} = \partial_x^2 \underline{u} + \mathcal{O}(h^8) \tag{4.60c}$$

These operators can also be written in the form:

$$\delta_{xx}^{[M]} \underline{u}_i^n = \sum_{m=-M}^{M} a_{M,m} \underline{u}_{i+m}^n \tag{4.61}$$

where $a_{M,m} = a_{-m,M}$, and

$$a_{M,0} = -2 \sum_{m=1}^{M} a_{M,m}\,, \quad a_{M,m} = \frac{\alpha_{M,m}}{m^2}\,, \quad m = 1, \ldots, M \tag{4.62}$$

It can also be shown [7] that in the limit as $M \to \infty$, one has:

$$\alpha_{m,\infty} = 2(-1)^{m-1} \quad \Longrightarrow \quad a_{m,\infty} = (-1)^{m-1}\frac{2}{m^2}\,, \quad m \geq 1\,, \quad a_{0,\infty} = -\frac{\pi^2}{3} \tag{4.63}$$

and the resulting operator $\delta_{xx}^{[\infty]}$ has an infinite-order accuracy, i.e.,

$$\delta_{xx}^{[\infty]}\underline{u} = \partial_x^2\underline{u} \tag{4.64}$$

provided that $\underline{u}$ is sufficiently bandlimited [7]. In practice this operator cannot be employed as it is non-local,[7] however it can be applied in the frequency-domain to a discrete approximation that is periodic in space. Essentially, one proceeds by applying the discrete Fourier transform (DFT) to one period of the discrete approximation, followed by a multiplication by $-\beta^2$ at regularly spaced discrete spatial frequencies, followed by an inverse DFT. This technique goes by various names; it is known as the "Fourier method" [130], a "pseudospectral method" [241], and as a Fourier-based spectral method [277]. If the solution is not periodic—which is often the case—and one applies this technique, the periodicity constraint is *imposed* on the resulting solution. This comprises a major drawback and challenge with such methods (as discussed in Section 2.2.5).

A general family of explicit two-step schemes

Now that we have defined a general family of discrete 1-D Laplacians, we can easily write a general family of two-step explicit schemes for the 1-D wave equation:

$$\delta_{tt}\underline{u}_i^n = c^2\sum_{p=1}^{P}\alpha_p\delta_{xx}^{(p)}\underline{u}_i^n \tag{4.65}$$

and for simplicity in this section it will be assumed that $\alpha_p \neq 0$ for $p = 1, \ldots, P$. We will also express these parameters by the vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_P)$.

Since this general family is of the form (4.20) with $\delta_\Delta = \sum_{p=1}^{P}\alpha_p\delta_{xx}^{(p)}$, numerical dispersion and stability follow from (4.26) and (4.29), where G, which depends on $\boldsymbol{\alpha}$, is now of the form:[8]

$$G(\beta_h) = \sum_{p=1}^{P}\frac{\alpha_p}{p^2}\sin^2(p\beta_h/2) \tag{4.66}$$

4.2.3 Schemes with high-order spatial differencing in 1-D

In this section we consider two-step schemes that make use of the aforementioned high-order difference operators (a family of well-known schemes [4, 2, 7]). These schemes are expressed by:

$$\delta_{tt}\underline{u}_i^n = c^2\delta_{xx}^{[M]}\underline{u}_i^n \tag{4.67}$$

and for this family of schemes we have:

$$(\delta_{tt} - c^2\delta_{xx}^{[M]})\underline{u} = \Box\underline{u} + \mathcal{O}(k^2) + \mathcal{O}(h^{2M}) \tag{4.68}$$

[7] From a signal processing point of view, one may recognise $a_{m,\infty}$ as the coefficients of an ideal second-order differentiator filter. This can be worked out from the definition of bandlimited sinc interpolation, following, e.g., [163].

[8] We note that while $G(\beta_h)$ is parametrised by the parameters in $\boldsymbol{\alpha}$, out of convenience we omit additional notation (such as, e.g., $G = G_{\boldsymbol{\alpha}}(\beta_h)$) to denote this dependence.

Because of the above truncation errors, these schemes can be said to be "$(2, 2M)$-accurate". If the Courant number is fixed to some value, we can make the substitution $h = ck/\lambda$ and the above becomes:

$$(\delta_{tt} - c^2\delta_{xx}^{[M]})\underline{u} = \Box\underline{u} + \mathcal{O}(k^2) + \mathcal{O}(k^{2M}) = \Box\underline{u} + \mathcal{O}(k^2) \tag{4.69}$$

showing that despite the high-order accuracy in space, the global accuracy of this family of two-step schemes remains second-order for all $M \geq 1$ (as is well-known [7]).

Numerical stability

For the investigation of numerical stability using a standard von Neumann analysis, the Fourier symbols of these high-order spatial difference operators can be written as:

$$\delta_{xx}^{[M]} \to -\frac{4}{h^2}G_M(\beta_h)\,, \qquad G_M(\beta_h) = \sum_{m=1}^{M} \frac{\alpha_{M,m}}{m^2}\sin^2(m\beta_h/2) \tag{4.70}$$

where $G_M(\beta_h)$ represents (4.66) with the particular choices of $\boldsymbol{\alpha} = (\alpha_{M,1}, \ldots, \alpha_{M,M})$ defined by (4.59). Since the coefficients $\alpha_{M,m}$ are alternating in sign (by (4.59)), it is not obvious that the stability condition $G_M(\beta_h) \geq 0$ is necessarily satisfied for $M > 1$. To address this problem, an important result was derived in [7] which re-expresses $G_M(\beta_h)$ in a non-negative form:

$$G_M(\beta_h) = \sum_{m=1}^{M} \gamma_m \sin^{2m}(\beta_h/2)\,, \qquad \gamma_m = \frac{2^{2m-1}((m-1)!)^2}{(2m)!} \tag{4.71}$$

where γ_m are strictly positive coefficients that are independent of M. Under this representation, it is clear that $G_M(\beta_h) \geq 0$. Furthermore, since each coefficient is positive, the maxima of G_M must occur at $\beta_h = \pi$, and these maxima can be written as:

$$\max_{\beta_h} G_M(\beta_h) = \Lambda_M\,, \qquad \Lambda_M := \sum_{m=1}^{M} \gamma_m \tag{4.72}$$

For example, we have $\Lambda_1 = 1, \Lambda_2 = \frac{4}{3}, \Lambda_3 = \frac{68}{45}, \Lambda_4 = \frac{512}{315}$, and it has been shown that as $M \to \infty$ [7]:

$$\Lambda_\infty = \lim_{M\to\infty} \Lambda_M = \frac{\pi^2}{4} \tag{4.73}$$

The stability condition for these schemes can then be written as:

$$\lambda \leq \lambda_{\max,M} = \frac{1}{\sqrt{\Lambda_M}} \tag{4.74}$$

and the first few $\lambda_{\max,M}$ are:

$$\lambda_{\max,1} = 1\,, \quad \lambda_{\max,2} = \sqrt{3/4}\,, \quad \lambda_{\max,3} = \sqrt{45/68}\,, \quad \lambda_{\max,4} = \sqrt{315/512} \tag{4.75}$$

By (4.73), $\lambda \leq \lambda_{\max,\infty} = \frac{2}{\pi}$ is a sufficient stability condition for all $M \geq 1$. It is worth mentioning that a corollary of (4.71) and (4.72) is:

$$\lambda_{\max,M} = \left(\sum_{p=0}^{\lceil M/2 \rceil} a_{M,2p+1} \right)^{-1/2} \tag{4.76}$$

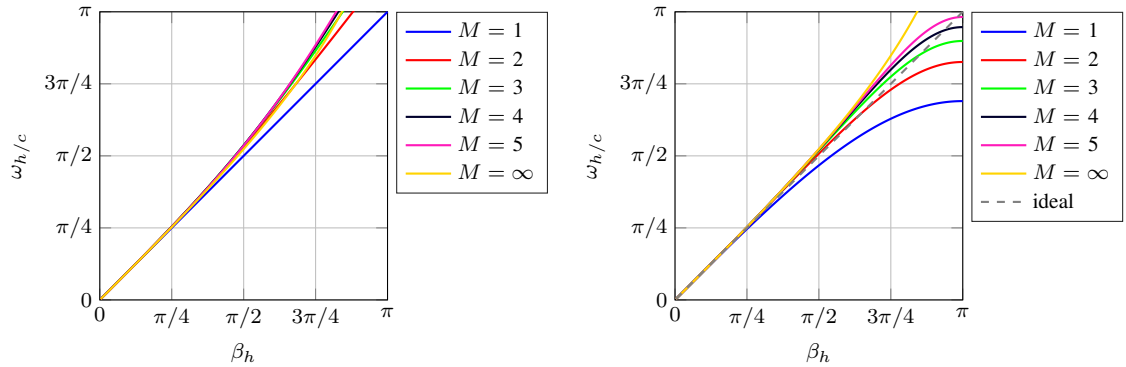
Numerical dispersion

A normalised numerical dispersion relation (corresponding to the form (4.49)) for these $(2, 2M)$-accurate schemes is:

$$\omega_{h/c}(\beta_h) = \frac{2}{\lambda} \sin^{-1}\left(\pm\lambda \sqrt{G_M(\beta_h)} \right) \tag{4.77}$$

This dispersion relation is plotted for various M with $\lambda = \lambda_{\max,M}$ in Fig. 4.3(a). As can be seen in the figure, these schemes are consistent with the ideal dispersion relation (in the limit of small h). However, nothing is gained by increasing M since $M = 1$ already provides an infinite-order accuracy in the special case of $\lambda = 1$ (this will not be the case for $d > 1$).

Ignoring the special infinite-order scheme, it is worth comparing these schemes for a fixed λ, and it is appropriate to choose $\lambda = \lambda_{\max,\infty} = 2/\pi$ for this purpose. The numerical dispersion relations with this choice of λ are shown in Fig. 4.3(b). In this figure one can see that each successive order of spatial accuracy results in a different numerical dispersion relation, but it is not necessarily clear that each successive order provides a better fit to the ideal. Consequently, the optimal choice of M and λ in these schemes will generally depend on the model problem (ignoring the special case $M = 1$ and $\lambda = 1$). This idea will be revisited in more detail when higher-dimensional forms of these schemes are considered (in Sections 4.3.3 and 4.5.2).



(a) $\omega_{h/c}(\beta_h)$ for various M with $\lambda = \lambda_{\max,M}$

(b) $\omega_{h/c}(\beta_h)$ for various M with $\lambda = 2/\pi$

Figure 4.3: Numerical dispersion for $(2, 2M)$-accurate schemes in 1-D.

Before moving on from this family of schemes, it is worth also investigating how the relative dispersion errors decrease in the limit of $h \to 0$ for λ fixed; this is illustrated in Fig. 4.4. It can be seen from this figure that regardless of the choice of M, the schemes display a rate of second-order rate convergence (in the dispersion error), reflecting the second-order global accuracy of this family of schemes. Indeed, this holds true for any $\lambda \leq \lambda_{\max,M}$ (apart from the exact case with $M = 1$).[9] We also note that there are *notches* appearing in Fig. 4.4 for finite $M > 1$; these are due to the fact that the dispersion error changes sign at certain wavenumbers, as can be deduced from Fig. 4.3(b).

[9]Consequently, it is a bit of a misnomer to call these schemes "high-order accurate" for the wave equation, as is sometimes done in the literature.

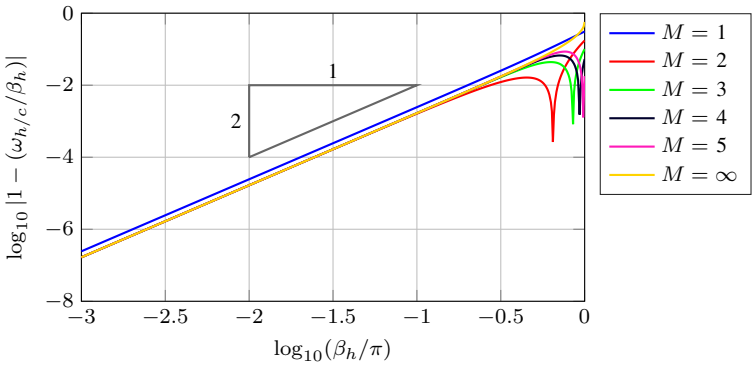


Figure 4.4: Relative dispersion error, $|1-\omega_{h/c}(\beta_h)/\beta_h|$, of $(2,2M)$-accurate schemes for $\lambda = 2/\pi$ and for various M.

On increasing the order of accuracy in time

Let us now discuss briefly some of the issues related to increasing the order of accuracy in time using high-order generalisations of standard leapfrog integrators.

In order to modify the $(2,2M)$-accurate schemes (employing $\delta_{xx}^{[M]}$) to have $(2M,2M)$-accuracy, one has to replace δ_{tt} with a time-difference operator that approximates ∂_t^2 to a higher order of accuracy. In order to derive such operators, one can simply replace x by t and h by k in $\delta_{xx}^{[M]}$ to arrive at high-order time difference operators $\delta_{tt}^{[M]}$. For example, a fourth-order accurate operator is:

$$\delta_{tt}^{[2]} = \frac{1}{12k^2}(-e_{t-}^2 + 16e_{t-} - 30 + 16e_{t+} - e_{t+}^2) \tag{4.78}$$

A seemingly fourth-order accurate explicit scheme (requiring four states in memory) could then be written as:

$$(\delta_{tt}^{[2]} - c^2\delta_{xx}^{[2]})\underline{u} = \Box\underline{u} + \mathcal{O}(k^4) + \mathcal{O}(h^4) \tag{4.79}$$

However, it is well-known that using $\delta_{tt}^{[2]}$ in this form leads to a scheme that is *unconditionally unstable* [55, 7]. As such, this scheme cannot produce an approximate solution $\underline{u}$ that converges to a true solution u. Other multi-stage time integrators could be employed (e.g., Runge-Kutta) by recasting the system in first-order form, but investigating such options will not be considered for brevity. Instead, we will make use of the same set of spatial stencils, but we will choose the stencil weights α_p in a manner that provides high orders of accuracy in both space and time by the manipulation of so-called "modified equations" [301].

4.2.4 High-order accurate schemes using modified equations

In this section we consider schemes with high-orders of accuracy in *both* space and time, derived through the manipulation of modified equations. A modified equation is essentially the Taylor expansion of a finite difference scheme about $h = k = 0$, and it represents the PDE that is solved *exactly* by the finite difference scheme [241]. In order to derive a modified equation for the general family of two-step schemes (4.65), we start with Taylor expansions of the various component operators. Assuming a sufficiently smooth $\underline{u}$, the operators δ_{tt} and δ_{xx} can be expanded in Taylor series to $2M$th-order in k

and h respectively:

$$\delta_{tt}\underline{u} = \partial_t^2\underline{u} + \sum_{m=2}^{M} \frac{2k^{2m-2}}{(2m)!}\partial_t^{2m}\underline{u} + \mathcal{O}(k^{2M}), \quad \delta_{xx}^{(p)}\underline{u} = \partial_x^2\underline{u} + \sum_{m=2}^{M} \frac{2(ph)^{2m-2}}{(2m)!}\partial_x^{2m}\underline{u} + \mathcal{O}(h^{2M}) \tag{4.80}$$

where the terms in the summations comprise the local truncation error. It follows from the second expansion above that we have for a linear combination of discrete Laplacians:

$$\sum_{p=1}^{P} \alpha_p \delta_{xx}^{(p)}\underline{u} = \partial_x^2\underline{u} + \sum_{m=2}^{M}\sum_{p=1}^{P} \alpha_p \frac{2(ph)^{2m-2}}{(2m)!}\partial_x^{2m}\underline{u} + \mathcal{O}(h^{2M}) \tag{4.81}$$

and it is assumed here that $\sum_{p=1}^{P} \alpha_p = 1$, which provides a base second-order accuracy. Combining these expansions for a finite difference scheme, we have:

$$\delta_{\square}\underline{u} = \left(\delta_{tt} - c^2\sum_{p=1}^{P}\alpha_p\delta_{xx}^{(p)}\right)\underline{u} = \square\underline{u} + \sum_{m=2}^{M}\left(\frac{2k^{2m-2}}{(2m)!}\partial_t^{2m} - c^2\sum_{p=1}^{P}\alpha_p\frac{2(ph)^{2m-2}}{(2m)!}\partial_x^{2m}\right)\underline{u} + \mathcal{O}(k^{2M}) + \mathcal{O}(h^{2M}) \tag{4.82}$$

Now if we take λ to be a constant, we can make the substitution $k = h\lambda/c$ and we have:

$$\delta_{\square}\underline{u} = \square\underline{u} + \sum_{m=2}^{M}\frac{h^{2m-2}}{c^{2m-2}}\underbrace{\left(\frac{2\lambda^{2m-2}}{(2m)!}\partial_t^{2m} - c^{2m-2}\sum_{p=1}^{P}\alpha_p\frac{2p^{2m-2}}{(2m)!}\partial_x^{2m}\right)}_{W_m}\underline{u} + \mathcal{O}(h^{2M}) = 0 \tag{4.83}$$

This modified equation represents the PDE that is exactly solved by the finite difference scheme.

The goal is now to derive a high-order scheme by manipulating this modified equation to agree with the wave equation to $2M$th-order ($\mathcal{O}(h^{2M})$). Often in the literature, this is accomplished by "trading off" time derivatives for spatial derivatives, substituting powers of ∂_t^2 for powers of $c^2\partial_x^2$ in the above expression, i.e., to assume that $\partial_t^{2m} = c^{2m}\partial_x^{2m}$ [53].[10] However, this is often accompanied by the assumption that the dependent variable on which finite difference operators act (here $\underline{u}$) is also a solution to the model PDE [53]. This assumption will not be made here, and this is the main reason for the use of a distinct symbol ($\underline{u}$) to represent the approximation, as recommended in [301]. As such, what follows is a closely-related, yet slightly different approach to modified equations than the conventional one [53].[11] We also note that the following approach follows directly from [301] and was recently used in an equivalent frequency-domain form for the 2-D wave equation [27] (published work involving the current author).

With that said, rather than substituting ∂_t^{2m} for $c^{2m}\partial_x^{2m}$, the idea is to divide out factors of $(\partial_t^2 - c^2\partial_x^2)$ from each term in brackets that is inside the summation (labelled W_m), which leads to:

$$W_m = \frac{c^{2m}\partial_x^{2m}}{(2m)!}\left(\lambda^{2m-2} - \sum_{p=1}^{P}\alpha_p p^{2m-2}\right) + (\partial_t^2 - c^2\partial_x^2)\mathcal{P}_{m-1} \tag{4.84}$$

where the first term on the right-hand side represents a remainder term, and $\mathcal{P}_{m-1} = \mathcal{P}_{m-1}\left(\partial_t^2, c^2\partial_x^2\right)$ is a multinomial of order $m-1$ in ∂_t^2 and $c^2\partial_x^2$, whose precise form is not important. The remainder

[10] The frequency domain equivalent is to substitute the ideal dispersion relation into the frequency domain version of (4.83), as done in [157].

[11] Ultimately, many of the same schemes to come have been derived by the conventional approach.

term in (4.84) cancels out provided that the following conditions can be satisfied up to order M:

$$\lambda^{2m-2} = \sum_{p=1}^{P} \alpha_p p^{2m-2}, \quad m = 2, \dots, M \tag{4.85}$$

Then we can write:

$$\delta_{\Box} \underline{u} = \Box \underline{u} + \left(\sum_{m=2}^{M} \frac{h^{2m-2}}{c^{2m-2}} \mathcal{P}_{m-1} \right) \Box \underline{u} + \mathcal{O}(h^{2M}) \tag{4.86}$$

or, more simply:

$$\delta_{\Box} \underline{u} = \Box \underline{u} + \mathcal{O}(h^2) \Box \underline{u} + \mathcal{O}(h^{2M}) \tag{4.87}$$

and since the scheme is defined by $\delta_{\Box} \underline{u} = 0$, the scheme is also represented by the modified equation:

$$\left(1 + \mathcal{O}(h^2)\right) \Box \underline{u} + \mathcal{O}(h^{2M}) = 0 \tag{4.88}$$

Consider now the inverse of the term $(1 + \mathcal{O}(h^2))$, which we can write as:[12]

$$\left(1 + \mathcal{O}(h^2)\right)^{-1} = 1 - \mathcal{O}(h^2) \tag{4.89}$$

We apply the above to (4.88), which means our modified equation is now:

$$\left(1 + \mathcal{O}(h^2)\right)^{-1} \left(1 + \mathcal{O}(h^2)\right) \Box \underline{u} = \left(1 - \mathcal{O}(h^2)\right) \mathcal{O}(h^{2M}) \implies \Box \underline{u} = \mathcal{O}(h^{2M}) \tag{4.90}$$

which means that our scheme satisfies the wave equation to $2M$th-order accuracy. Put another way, we have a local truncation error of order $2M$ for the scheme:

$$\delta_{\Box} \underline{u} = \Box \underline{u} + \mathcal{O}(h^{2M}) \tag{4.91}$$

With regards to the conditions (4.85), we note that they can be satisfied for $P = 1$ up to $M = \infty$ with $\lambda = 1$, which leads to the exact (simplest) scheme. Apart from that special case, it would make sense to choose $P \geq M$ in order to achieve $2M$th-order accuracy for any stable λ. This leads to many possible schemes (and families of schemes) of arbitrary orders of accuracy. Next, we consider what could be seen as the *minimal* schemes needed to achieve $2M$th-order accuracy, that is with $P = M$.

In the case of $P = M$, the conditions for high-order accuracy comprise a Vandermonde matrix system, for which a unique solution exists [162]. Although, considering the connection between a Vandermonde matrix and Lagrange polynomial interpolation, we can express the coefficients α_p for $M > 1$ in a simple form as a Lagrange polynomial in the variable λ^2:

$$\alpha_p = \prod_{\substack{1 \leq r \leq M \\ r \neq p}} \frac{\lambda^2 - r^2}{p^2 - r^2} \tag{4.92}$$

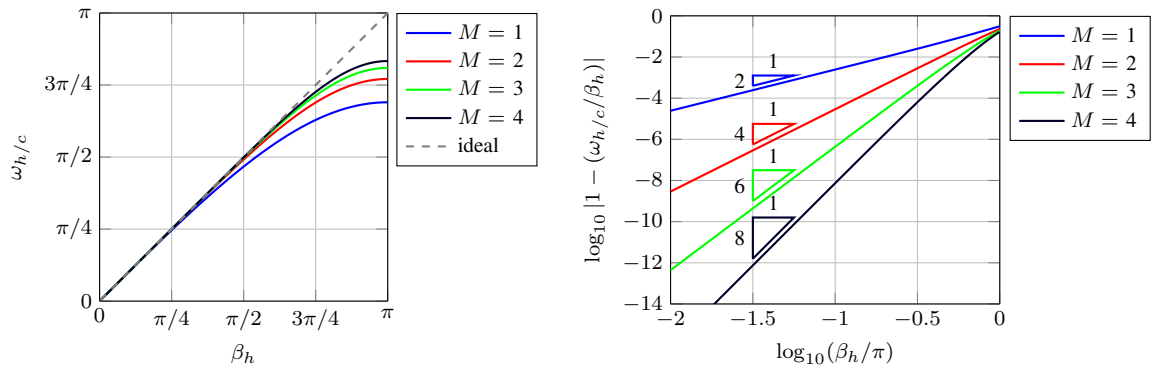
Notice that for $\lambda = 1$ we have that $\alpha_p = 0$ for $p \geq 1$, which reduces to the exact scheme. As such, this scheme is certainly stable for $\lambda = 1$. Otherwise, α_p depends on λ, so the stability condition cannot be expressed as (4.29), and we must satisfy the more general conditions (4.28), which involves the difficult problem of bounding a polynomial in λ^2. Fortunately, this problem has already been solved and the

[12] Here we use the standard Taylor series $(1 + x)^{-1} = 1 - x + x^2 + \dots$, under the assumption that $|x| < 1$, which holds here since we are considering $h \to 0$.

stability condition for this entire family has been shown to be $\lambda \leq 1$ [200, 283, 7]. This family of schemes was also derived in [157] using a modified equation approach in the frequency-domain, and in [7] using a markedly different frequency-domain approach, from which another general expression for α_p can be written as:[13]

$$\alpha_p = 2p^2(-1)^{p-1} \sum_{m=p}^{M} \binom{2m}{m-p} \frac{(m-1)!^2}{(2m)!} \prod_{r=1}^{m-1} \left(1 - \frac{\lambda^2}{r^2}\right) \tag{4.93}$$

In order to verify the order of accuracy of these schemes, we plot and compare the numerical dispersion relations for these schemes for $M \in \{1, 2, 3, 4\}$ with a fixed Courant number ($\lambda = 2/\pi$). This is shown in Fig. 4.5(a), wherein improvements in the numerical dispersion relations can be seen for increasing orders $2M$. We can also investigate these dispersion errors on a log-scale, which is shown in Fig. 4.5(b). It can be seen that the dispersion errors decrease with rates that are in accordance with the associated orders of accuracy.



(a) Numerical dispersion relations

(b) Relative dispersion errors

Figure 4.5: Plots of numerical dispersion and numerical dispersion errors on log-scale for $2M$-order accurate schemes in 1-D with $\lambda = 2/\pi$. The slopes of the curves in the right figure demonstrate high-orders of accuracy in the numerical dispersion.

[13] It has been checked with a computer that (4.92) and (4.93) give equivalent polynomials in λ^2 for each α_p, up to a very high order M. A full equivalence between these two expressions is left as an exercise for the reader.

4.3 Cartesian schemes for the 2-D wave equation

Now that some concepts relating to order of accuracy in finite difference schemes have been illustrated in one spatial dimension, we move on to the wave equation in two spatial dimensions.

The 2-D wave equation can be represented by:

$$\Box u = 0\,, \quad \Box = \partial_t^2 - c^2\Delta\,, \quad \Delta = \partial_x^2 + \partial_y^2 \tag{4.94}$$

with $u = u(\mathbf{x}, t)$ where $\mathbf{x} \in \mathbb{R}^2$, $t \in \mathbb{R}^+$. Required initial conditions for this second-order equation are $u(\mathbf{x}, 0)$ and $u_t(\mathbf{x}, 0)$, but for the most part we ignore the initial conditions in order to focus on accuracy in terms of numerical dispersion.

With regards to the grid function that will represent the fully-discrete approximation to u, we note that there is more than one option for a spatial grid with a regular grid spacing in two spatial dimensions. We will start by considering the simplest and most widely used grid, the 2-D Cartesian grid ($h\mathbb{Z}^2$), and later we will consider regular hexagonal grids (in Section 4.4). For now our grid function of interest is $\underline{u}_{\boldsymbol{i}}^n \approxeq u(\boldsymbol{i}h, nk)$, with $\boldsymbol{i} \in \mathbb{Z}^2$ and $n \geq 0$. We will also make use of the continuous approximation $\underline{u}(\mathbf{x}, t)$ where $\mathbf{x} \in \mathbb{R}^2$, $t \in \mathbb{R}^+$, and $\underline{u}(\boldsymbol{i}h, nk) = \underline{u}_{\boldsymbol{i}}^n$.

4.3.1 The simplest scheme in 2-D

We start with the simplest possible scheme on a 2-D Cartesian grid, which happens to be a straightforward extension of the simplest 1-D scheme [60]. This scheme can be represented by (4.20) with the discrete Laplacian:

$$\delta_\Delta = \delta_{xx} + \delta_{yy} \tag{4.95}$$

and δ_Δ is the standard 2-D discrete Laplacian that uses a five-point stencil. In terms of an update recursion, this scheme can be written as:

$$\underline{u}_{i_x,i_y}^{n+1} = -\underline{u}_{i_x,i_y}^{n-1} + (2 - 4\lambda^2)\underline{u}_{i_x,i_y}^n + \lambda^2(\underline{u}_{i_x+1,i_y}^n + \underline{u}_{i_x-1,i_y}^n + \underline{u}_{i_x,i_y+1}^n + \underline{u}_{i_x,i_y-1}^n) \tag{4.96}$$

This scheme is second-order accurate, and has the well-known stability condition $\lambda \leq \lambda_{\max} = \sqrt{1/2}$, which is also the CFL condition for convergence [60]. Under the choice of $\lambda = \lambda_{\max}$, we have the simplified update recursion:

$$\underline{u}_{i_x,i_y}^{n+1} = -\underline{u}_{i_x,i_y}^{n-1} + \frac{1}{2}(\underline{u}_{i_x+1,i_y}^n + \underline{u}_{i_x-1,i_y}^n + \underline{u}_{i_x,i_y+1}^n + \underline{u}_{i_x,i_y-1}^n) \tag{4.97}$$

which is also equivalent to a 2-D wave-scattering network known as a "transmission line matrix" method [119, 220] or as the "rectilinear digital waveguide mesh" [286]. The above equation also has the interpretation as a balance of temporal and spatial averages:

$$\frac{\underline{u}_{i_x,i_y}^{n+1} + \underline{u}_{i_x,i_y}^{n-1}}{2} = \frac{\underline{u}_{i_x+1,i_y}^n + \underline{u}_{i_x-1,i_y}^n + \underline{u}_{i_x,i_y+1}^n + \underline{u}_{i_x,i_y-1}^n}{4} \tag{4.98}$$

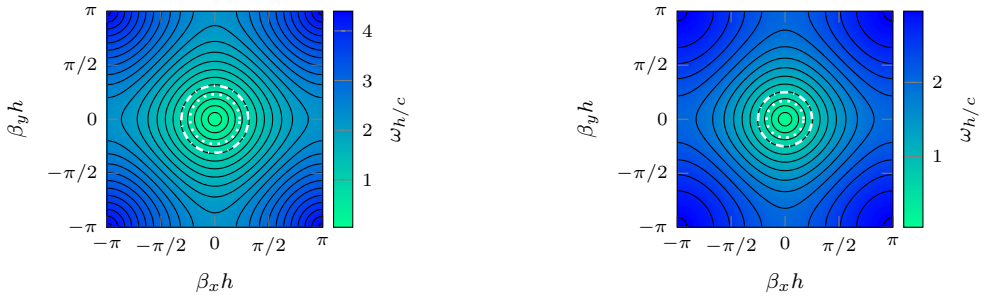
One can also work out an equivalence to 2-D forms of Yee's classic FDTD method [311]; see, e.g., [21].

The numerical dispersion relation of this scheme is expressed by (4.26) where $G(\boldsymbol{\beta}_h)$ takes the form:

$$G(\boldsymbol{\beta}_h) = \sin^2(\beta_x h/2) + \sin^2(\beta_y h/2) \tag{4.99}$$

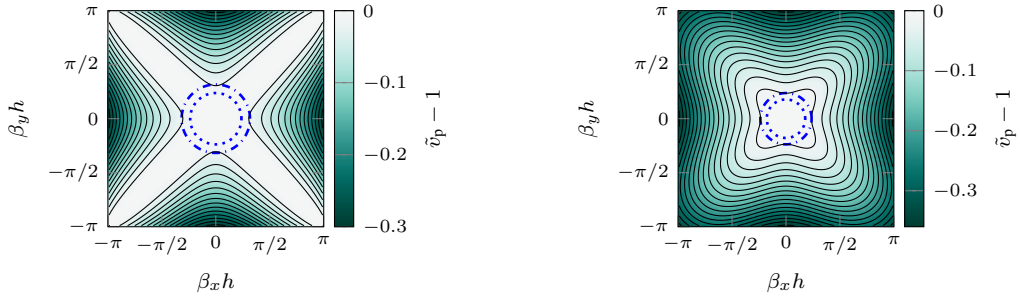
Note that by the periodicity in $G(\boldsymbol{\beta}_h)$, we only need to consider $\boldsymbol{\beta}_h \in \mathbb{B} := [-\pi, \pi]^2$, which also follows sampling considerations. As for numerical stability, (4.29) is satisfied for $\lambda \leq \sqrt{1/2}$ since $\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h) = 2$.

The numerical dispersion relation of this scheme is plotted in Fig. 4.6 for $\lambda \in \{\lambda_{\max}, 0.5\lambda_{\max}\}$, which can be compared to the ideal case in Fig. 4.1. It is also instructive to plot the relative dispersion error, and this is shown in Fig. 4.7. In both figures, a white dotted line is drawn for $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$ (the one-percent critical wavenumber, defined by (4.33)), and a white dashed-dotted line for $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ (the two-percent critical wavenumber, defined by (4.33)).



(a) $\lambda = \lambda_{\max}$ (b) $\lambda = 0.5\lambda_{\max}$

Figure 4.6: Contour plots of numerical dispersion relation $\omega_{h/c}(\boldsymbol{\beta}_h)$ for $\boldsymbol{\beta}_h \in [-\pi, \pi]^2$ with various $\lambda \leq \lambda_{\max}$. Contour lines denote increments of $\pi/16$. The white dotted line denotes $|\boldsymbol{\beta}_h| = \pi$. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.



(a) $\lambda = \lambda_{\max}$ (b) $\lambda = 0.5\lambda_{\max}$

Figure 4.7: Contour plots of $\tilde{v}_p(\boldsymbol{\beta}_h) - 1$. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

In contrast to Fig. 4.1, it can be seen from Fig. 4.6 that the numerical dispersion relations are generally anisotropic (not independent of wave direction). However, by consistency these numerical dispersion relations become increasingly isotropic (and correct) in the limit of small wavenumbers. It can also be seen from Fig. 4.7 that the numerical dispersion is closer to the ideal with $\lambda = \lambda_{\max}$ than it is with $\lambda < \lambda_{\max}$, as was the case for the simplest scheme in 1-D. In fact (and as is well-known), for the case $\lambda = \lambda_{\max}$ the numerical dispersion relation is *exact* along diagonal directions, i.e., when $\beta_x = \pm\beta_y$.[14] For both choices of λ shown here, the wave directions that feature the worst-case dispersion are the axial directions.

[14]This follows from the fact that for $\beta_x = \pm\beta_y$, the 2-D dispersion relation becomes $\omega_k = 2\sin^{-1}(\lambda' \sin(\beta h'/2))$, where $h' = h/\sqrt{2}$ and $\lambda' = ck/h'$. This is in the form of the 1-D scheme, and as such, for $\lambda' = 1 \implies \lambda = 1/\sqrt{2}$ the scheme must be exact for $\beta_x = \pm\beta_y$. Also, for $\beta_x = 0$ or $\beta_y = 0$, the dispersion relation of the 2-D scheme reduces to that of the 1-D scheme.

Unless the application of interest can tolerate significant dispersion errors ($\approx 30\%$ error), the range of wavenumbers that are accurately reproduced in this scheme (such as, e.g., to within 1% or 2% error in $\tilde{v}_\mathrm{p}$, as illustrated by Fig. 4.7) is only a small portion of the wavenumbers cell $\mathbb{B}$. In the hopes of avoiding expensive grid refinements (choosing h smaller than what is dictated by sampling constraints) we will subsequently investigate more accurate schemes, which must invariably use more points in space and/or time for each pointwise update than this simple scheme.

Before moving on, we remark that the dispersion relations shown here feature a mirror symmetry with respect to the coordinate axes (and a mirror symmetry with respect to the diagonals) and this will also be the case for the Cartesian schemes to follow. As such, we will (for the most part) present dispersion errors for the schemes to follow only in the quarter-region $\mathbb{B} \in [0, \pi]^2$. This is exemplified in Fig. 4.8 for the simplest scheme.

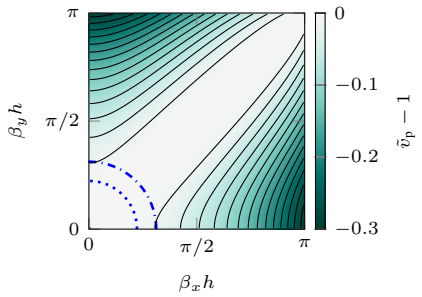


Figure 4.8: Contour plot of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for $\boldsymbol{\beta}_h \in [0, \pi]^2$, for simplest scheme with $\lambda = \lambda_{\max,M}$. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

4.3.2 Parametrised approximations to the 2-D Laplacian

In order to derive schemes beyond the simplest one, we can make use of a general framework for discrete Laplacians on the 2-D Cartesian grid, after [27] (published work involving the current author).

Consider the following spatial averaging operators:

$$\mu_{b,x} = \frac{1}{2}(e^b_{x+} + e^b_{x-}), \quad \mu_{b,y} = \frac{1}{2}(e^b_{y+} + e^b_{y-}) \tag{4.100}$$

where b denotes a positive integer–power of the shift operator (e.g., $e^2_{x+} = e_{x+}e_{x+}$). These operators have the Taylor expansions:

$$\mu_{b,x}\underline{u} = \underline{u} + \frac{b^2h^2}{2}\partial_x^2\underline{u} + \mathcal{O}(h^4), \quad \mu_{b,y}\underline{u} = \underline{u} + \frac{b^2h^2}{2}\partial_y^2\underline{u} + \mathcal{O}(h^4) \tag{4.101}$$

For convenience, let us define the following set $\mathbb{Q}$ of integer pairs $\boldsymbol{q} = (q_1, q_2)$:

$$\mathbb{Q} := \{\boldsymbol{q} \in \mathbb{Z}^2 : q_1 \geq q_2 \geq 0, q_1 \geq 1\} \tag{4.102}$$

For $\boldsymbol{q} \in \mathbb{Q}$, it follows that:

$$\mu_{q_1,x}\mu_{q_2,y}\underline{u} = 1 + \frac{q_1^2h^2}{2}\partial_x^2\underline{u} + \frac{q_2^2h^2}{2}\partial_y^2\underline{u} + \mathcal{O}(h^4), \qquad \mu_{q_2,x}\mu_{q_1,y}\underline{u} = 1 + \frac{q_2^2h^2}{2}\partial_x^2\underline{u} + \frac{q_1^2h^2}{2}\partial_y^2\underline{u} + \mathcal{O}(h^4) \tag{4.103}$$

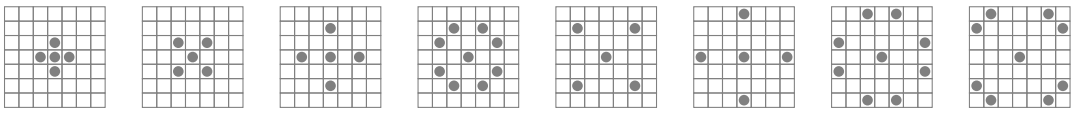
We can then define a discrete Laplacian $\delta_{\Delta,\boldsymbol{q}} = \delta_{\Delta,(q_1,q_2)}$as:

$$\delta_{\Delta,\boldsymbol{q}} = \frac{2}{(q_1^2+q_2^2)h^2}\left(\mu_{q_1,x}\mu_{q_2,y} + \mu_{q_2,x}\mu_{q_1,y} - 2\right) \tag{4.104}$$

and from the above Taylor expansions we have:

$$\delta_{\Delta,\boldsymbol{q}}\underline{u} = \Delta\underline{u} + \mathcal{O}(h^2) \tag{4.105}$$

Some of the stencils of points that are used by these discrete operators are shown in Fig. 4.9. Under this notation, the standard five-point discrete Laplacian is $\delta_{\Delta,(1,0)}$, using the stencil of points shown in Fig. 4.11(a). The next stencil in a lexicographic ordering of (q_1, q_2) is the rotated five-point stencil shown in Fig. 4.11(b), denoted as $\delta_{\Delta,(1,1)}$.



(a) $(1,0)$ (b) $(1,1)$ (c) $(2,0)$ (d) $(2,1)$ (e) $(2,2)$ (f) $(3,0)$ (g) $(3,1)$ (h) $(3,2)$

Figure 4.9: Stencils of points for operators $\delta_{\Delta,(q_1,q_2)}$ with (q_1, q_2) as indicated.

Let K represent the number of neighbouring points used in the stencil. It is straightforward to realise that $K = 8$ when $q_1 \neq q_2$ and $q_2 > 0$, otherwise $K = 4$. For the determination of stencil coefficients, it also helps to rewrite these operators as:

$$\delta_{\Delta,\boldsymbol{q}}\underline{u}_{\boldsymbol{i}}^n = \frac{4/K}{\|\boldsymbol{q}\|^2h^2}(Q_{\boldsymbol{i},\boldsymbol{q}}^n - K\underline{u}_{\boldsymbol{i}}^n) \tag{4.106}$$

where $Q_{\boldsymbol{i},\boldsymbol{q}}^n = Q_{\boldsymbol{i},(q_1,q_2)}^n$ represents the sum over the K neighbouring grid points to $\boldsymbol{i}h \in h\mathbb{Z}^2$. We can define this sum as:

$$Q_{\boldsymbol{i},\boldsymbol{q}}^n = \sum_{\boldsymbol{q}'\in\mathcal{P}(\boldsymbol{q})} \underline{u}_{i_x+q_x',i_y+q_y'}^n \tag{4.107}$$

where $\boldsymbol{q}' = (q_x', q_y')$ and $\mathcal{P}(\boldsymbol{q})$ is the set of all *unique* permutations of $\{\pm q_1, \pm q_2\}$. Clearly, $K = |\mathcal{P}(\boldsymbol{q})|$.

Consider now a set of P distinct integer pairs $\boldsymbol{q} = (q_1, q_2) \in \mathbb{Q}$ defined as $\Upsilon = \{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_P\}$. Following [27], we can define a general family of parametrised discrete Laplacians by:

$$\delta_{\Delta,\Upsilon} = \sum_{(q_1,q_2)\in\Upsilon} \alpha_{(q_1,q_2)}\delta_{\Delta,(q_1,q_2)} \tag{4.108}$$

where $\alpha_{(q_1,q_2)}$ is a real-valued parameter, and for consistency we require that:

$$\sum_{(q_1,q_2)\in\Upsilon} \alpha_{(q_1,q_2)} = 1 \tag{4.109}$$

which will be implicitly assumed. A general family of two-step explicit schemes for the 2-D wave equation can then be written as:

$$\delta_{tt}\underline{u}_{\boldsymbol{i}}^n = c^2\delta_{\Delta,\Upsilon}\underline{u}_{\boldsymbol{i}}^n \tag{4.110}$$

which is subject to the numerical stability constraints (4.28) (or (4.29)), where $G(\boldsymbol{\beta}_h)$ takes on a form given below.

Considering a plane-wave solution for $\underline{u}_{\boldsymbol{i}}^n$, we can derive a Fourier symbol for these operators. It is

straightforward to show then that these operators transform to:

$$\delta_{\Delta,\Upsilon} \to -\frac{4}{h^2} G(\boldsymbol{\beta}_h) \tag{4.111}$$

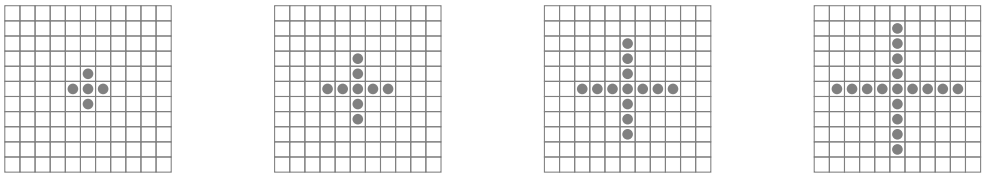
where

$$G(\boldsymbol{\beta}_h) = \sum_{(q_1,q_2)\in\Upsilon} \frac{\alpha_{(q_1,q_2)}}{2(q_1^2+q_2^2)} \left(2 - \cos(q_1\beta_x h)\cos(q_2\beta_y h) - \cos(q_2\beta_x h)\cos(q_1\beta_y h)\right) \tag{4.112}$$

Some well-known special cases of this general family of discrete Laplacians are presented next.

Standard $2M$th-order accurate discrete Laplacians

It is straightforward to extend the high-order second-order differences seen in the 1-D case to the 2-D Laplacian.[15] Under this general framework, these conventional $2M$th-order accurate discrete Laplacians arise when $q_2 = 0$, leading to the $(4M+1)$-point stencils shown in Fig. 4.10. We will use the term "leggy", after [309], to describe this family of stencils (and later, for subsequent 3-D generalisations), and it will be implicitly assumed that such leggy stencils are always used with the stencil weights that provide $2M$th-order accuracy in space.



(a) $M = 1$ (b) $M = 2$ (c) $M = 3$ (d) $M = 4$

Figure 4.10: $(4M+1)$-point leggy stencils used for $2M$th-order spatial differencing in 2-D.

In order to represent these $2M$th-order accurate discrete Laplacians under this general framework, let us define the sets Υ_M^+, for which $q_2 = 0$, as:

$$\Upsilon_M^+ = \{(1,0), \ldots, (M,0)\} \tag{4.113}$$

We can then write a family of discrete Laplacians using leggy stencils as:

$$\delta_{\Delta,\Upsilon_M^+} = \sum_{m=1}^{M} \alpha_{(m,0)} \delta_{\Delta,(m,0)} \tag{4.114}$$

Finally, for high-order accuracy in space, we set:

$$\alpha_{(m,0)} = \alpha_{M,m} \quad \Longrightarrow \quad \delta_{\Delta,\Upsilon_M^+}\underline{u} = \Delta\underline{u} + \mathcal{O}(h^{2M}) \tag{4.115}$$

where $\alpha_{M,m}$ is defined by (4.59). It should be noted that the $\mathcal{O}(h^{2M})$ term in the local truncation error for these schemes is always anisotropic.

[15] We note that this is the standard (and simplest) approach to achieving high-order accurate discrete Laplacians, but many high-order discrete Laplacians, and families of high-order discrete Laplacians, are possible under this general framework.

A nine-point family and an isotropic discrete Laplacian

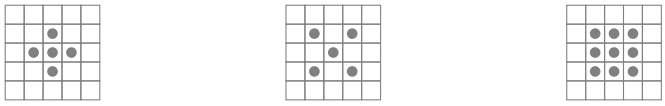
Another well-known special case of this general family of parametrised schemes is the following nine-point discrete Laplacian:

$$\delta_{\Delta,\Upsilon_9} = \alpha_{(1,0)}\delta_{\Delta,(1,0)} + \alpha_{(1,1)}\delta_{\Delta,(1,1)}\,, \quad \alpha_{(1,0)} + \alpha_{(1,1)} = 1 \tag{4.116}$$

for which we have defined the set $\Upsilon_9 = \{(1,0),(1,1)\}$. There is essentially one free parameter in this nine-point Laplacian, so we can also rewrite it as:

$$\delta_{\Delta,\Upsilon_9} = \alpha\delta_{\Delta,(1,0)} + (1-\alpha)\delta_{\Delta,(1,1)} \tag{4.117}$$

The stencils of points used by this operator, which can reduce to the standard five-point Laplacian for $\alpha = 1$ or to a rotated five-point Laplacian for $\alpha = 0$, are illustrated in Fig. 4.11. The nine-point stencil in Fig. 4.11(c) can be said to be "compact" in space, in contrast to the leggy nine-point stencil seen earlier (Fig. 4.10(b)). A finite difference scheme that uses this compact discrete Laplacian can also be said to be a "compact scheme" [187, 52, 291].[16]



(a) $\alpha = 1$ (b) $\alpha = 0$ (c) $\alpha \notin \{0,1\}$

Figure 4.11: Stencils of points used for $\delta_{\Delta,\Upsilon_9}$, as a function of α.

A useful and well-known choice for the free parameter α is $\alpha = 2/3$ [121, 56], as this leads to a particular simplification of the truncation error:

$$\alpha = 2/3 \quad \Longrightarrow \quad \delta_{\Delta,\Upsilon_9}\underline{u} = \Delta\underline{u} + \frac{h^2}{16}\Delta^2\underline{u} + \mathcal{O}(h^4) \tag{4.118}$$

The above Taylor expansion features the biharmonic (bilaplacian) operator (Δ^2) in the leading error term, and since the biharmonic is an isotropic operator, the above discrete Laplacian is said to be isotropic (to fourth order in h). For $d > 1$, isotropic error terms are crucial in achieving high-order accuracy through the manipulation of modified equations.

4.3.3 Schemes with high-order spatial differences in 2-D

The first family of scheme to consider beyond the simplest scheme will be the two-dimensional generalisations of the $(2, 2M)$th-order accurate schemes seen in Section 4.2.2, making use of the (herein-called) leggy stencils. Since for leggy stencils we only consider the stencil weights that give high-order accuracy in space, these $(2, 2M)$th-order accurate schemes will also be referred to as "leggy schemes". Leggy schemes are standard tools in acoustics simulations for geophysical applications [2, 66].

Considering the operators defined previously, this family of schemes can be written in the form of (4.20) with $\delta_\Delta = \delta_{\Delta,\Upsilon_M^+}$ (defined by (4.114), and with $\alpha_{(m,0)} = \alpha_{M,m}$). It is also useful to write

[16] However, we note that there does not seem to be an agreed upon definition of the word "compact" in the finite difference literature; e.g., "compact" has also been used to describe leggy stencils in 2-D schemes [150], and one could also apply the term "compact" to the spatiotemporal stencil used.

these schemes as the following update recursion:

$$\underline{u}^{n+1}_{i_x,i_y} = 2\underline{u}^{n}_{i_x,i_y} - \underline{u}^{n-1}_{i_x,i_y} + \lambda^2 \sum_{m=-M}^{M} a_{M,m} \left(\underline{u}^{n}_{i_x+m,i_y} + \underline{u}^{n}_{i_x,i_y+m} \right) \tag{4.119}$$

where $a_{M,m}$ are defined by (4.59). The numerical dispersion relations of these schemes are of the form (4.26), where now $G(\boldsymbol{\beta}_h)$ is:

$$G(\boldsymbol{\beta}_h) = \sum_{m=1}^{M} \gamma_m \left(\sin^{2m}(\beta_x h/2) + \sin^{2m}(\beta_y h/2) \right) \tag{4.120}$$

where γ_m are given by (4.71). Since $G(\boldsymbol{\beta}_h)$ is clearly positive, for stability one requires that:

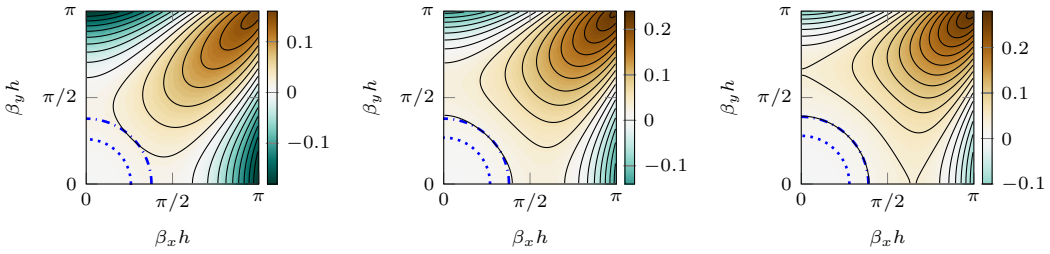
$$\lambda \leq \lambda_{\max,M} = \frac{1}{\sqrt{2\Lambda_M}} \tag{4.121}$$

where Λ_M is given by (4.72). A sufficient condition for all $M \geq 1$ is $\lambda \leq \lambda_{\max,\infty} = \sqrt{2}/\pi$. An interesting corollary of the above and (4.120) is that for $\lambda = \lambda_{\max,M}$ we have for all $M \geq 1$:

$$\omega_k((\pm\beta_x, 0)) = \omega_k((0, \pm\beta_y)) = 2\sin^{-1}\left(\frac{1}{\sqrt{2}}\right) = 0.5\pi \tag{4.122}$$

Since the dispersion relation increases monotonically along the axial directions (by (4.120)), these 2-D schemes can only reproduce at most 50% of the temporal Nyquist along axial directions. In the regard, there is no improvement over the simplest scheme ($M = 1$).

In order to analyse the dispersion errors in these schemes, we start by plotting the relative error $\varepsilon_{\text{rel. err.}}$ for $2 \leq M \leq 4$ with the Courant number set to respective stability limits; this appears in Fig. 4.12. It can be seen that the dispersion errors for $M \geq 2$ are only slightly improved over the case of $M = 1$ (which also appears in Fig. 4.7(a)), and that the improvement from $M = 2$ to $M = 3$ is marginal. The dispersion errors in these schemes, following the discrete Laplacians themselves, are always anisotropic in this family of schemes.



(a) $M = 2, \lambda = \lambda_{\max,2} \approx 0.61$ (b) $M = 3, \lambda = \lambda_{\max,3} \approx 0.58$ (c) $M = 4, \lambda = \lambda_{\max,4} \approx 0.56$

Figure 4.12: Contour plots of $\tilde{v}_{\text{p}}(\boldsymbol{\beta}_h) - 1$ for $(2, 2M)$-accurate schemes using $(4M + 1)$-point leggy stencils, with $\lambda = \lambda_{\max,M}$. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

On the optimal choice of the Courant number

So far we have analysed these schemes with $\lambda = \lambda_{\max}$, but for $M > 1$, this is not necessarily the best choice for the Courant number. For these $(2, 2M)$-accurate schemes, where temporal and spatial

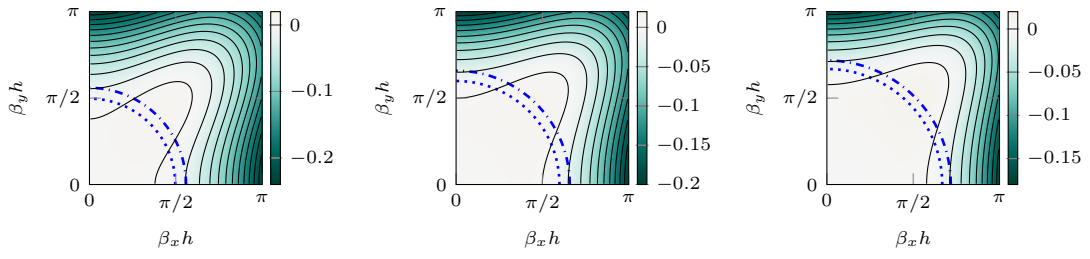
errors are unbalanced in orders, it can be beneficial to choose the time-step to be much smaller than the spatial step (choose λ small for a fixed h) in order to balance $\mathcal{O}(k^2)$ temporal errors with $\mathcal{O}(h^{2M})$ spatial errors. This is a well-known strategy [66, 53], and is also explained by the tendency of schemes that are dominated by temporal errors to have numerical wave speeds that are too fast, and visa-versa for schemes dominated by spatial errors [66]. It is thus worth reconsidering these leggy schemes under more optimal choices of the Courant number.

The interest here will mostly be a one-percent error criterion, so let us define $\lambda_{1\%,M}$ to be the Courant number $\lambda \in [0.01\lambda_{\max,M}, 0.02\lambda_{\max,M}, \ldots, 0.99\lambda_{\max,M}, \lambda_{\max,M}]$ that maximises the critical wavenumber $\beta_{h,1\%}$ in a given $(2, 2M)$-accurate leggy scheme, provided that the absolute error reaches one-percent for some $\beta_h \leq \pi$, which is generally the case for the orders M that will be considered here. We have chosen relative steps of 0.01 for simplicity. Searching through these possible Courant numbers, we find that a unique $\lambda_{1\%,M}$ exists (a global maximum in $\beta_{h,1\%}$) for the cases considered here. These optimised Courant numbers for $M \leq 4$ are found in Table 4.1.[17]

Table 4.1: Optimised Courant numbers for $(4M + 1)$-point leggy schemes in 2-D

M	$\lambda_{1\%,M}/\lambda_{\max,M}$	$\lambda_{1\%,M}$ (approx.)
1	1.00	0.707
2	0.67	0.410
3	0.52	0.299
4	0.46	0.255

Fig. 4.13 shows dispersion errors for $M \in \{2, 3, 4\}$ where now the Courant number is chosen to be $\lambda = \lambda_{1\%,M}$, which is significantly lower than respective stability limits of the schemes. It can be seen that for these choices of λ (as indicated in the figure) the dispersion error is improved over a wider range of wavenumbers. This improvement comes at the cost of a smaller time-step than at the stability limit, and thus more samples to compute for a given duration of time. We should also note that while contour lines denote 2% increments in these plots, the first contour line appears *before* the critical wavenumbers $\beta_{h,1\%}$ and $\beta_{h,2\%}$; this is simply because the $\tilde{v}_\mathrm{p}$ rises slightly above unity between the zero wavenumber and the first contour, and as such, the first contour represents 0% error, similar to the notches seen in Fig. 4.4. We will return to these $(2, 2M)$-accurate leggy schemes in order to make comparisons with other schemes to follow; next we consider schemes that make use of compact stencils.



(a) $M = 2, \lambda = \lambda_{1\%,2} \approx 0.41$ (b) $M = 3, \lambda = \lambda_{1\%,3} \approx 0.30$ (c) $M = 4, \lambda = \lambda_{1\%,4} \approx 0.26$

Figure 4.13: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for $(2, 2M)$-accurate schemes using $(4M + 1)$-point leggy stencils, with Courant numbers set as indicated. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

[17] Although not shown here, it is also worth mentioning that $\lambda_{1\%,M}$ appears to maximise $\beta_{h^*,1\%}$, which also takes into account spatiotemporal densities.

4.3.4 Compact $(K+1)$-point schemes in 2-D with $K \leq 8$

In this section we consider a family of two-step nine-point explicit schemes [89, 275, 224, 21, 131], of the form (4.20) using the discrete Laplacian $\delta_\Delta = \delta_{\Delta,\Upsilon_9}$ defined by (4.117). For this section we rewrite this discrete Laplacian in the alternative, equivalent form [24]:

$$\delta_{\Delta,\Upsilon_9} = \delta_{xx} + \delta_{yy} + (1-\alpha)\frac{h^2}{2}\delta_{xx}\delta_{yy} \tag{4.123}$$

where $\alpha = \alpha_{(1,0)}$ is the usual free parameter (consistency has been imposed). In terms of an update recursion for this scheme, we have:

$$\begin{aligned}\underline{u}^{n+1}_{i_x,i_y} = -\underline{u}^{n-1}_{i_x,i_y} + (2 - 2\lambda^2(1+\alpha))\underline{u}^n_{i_x,i_y} + \alpha\lambda^2(\underline{u}^n_{i_x+1,i_y} + \underline{u}^n_{i_x-1,i_y} + \underline{u}^n_{i_x,i_y+1} + \underline{u}^n_{i_x,i_y-1}) \\ + \frac{(1-\alpha)\lambda^2}{2}(\underline{u}^n_{i_x+1,i_y+1} + \underline{u}^n_{i_x-1,i_y+1} + \underline{u}^n_{i_x-1,i_y+1} + \underline{u}^n_{i_x-1,i_y-1})\end{aligned} \tag{4.124}$$

or more compactly:

$$\underline{u}^{n+1}_{i_x,i_y} = -\underline{u}^{n-1}_{i_x,i_y} + (2 - 2\lambda^2(1+\alpha))\underline{u}^n_{i_x,i_y} + \lambda^2\left(\alpha Q^n_{\boldsymbol{i},(1,0)} + \frac{(1-\alpha)}{2}Q^n_{\boldsymbol{i},(1,1)}\right) \tag{4.125}$$

with $Q^n_{\boldsymbol{i},(1,0)}$ and $Q^n_{\boldsymbol{i},(1,1)}$ defined by (4.107). The numerical dispersion relation of this scheme is of the form (4.26) with $G(\boldsymbol{\beta}_h)$ given by:

$$G(\boldsymbol{\beta}_h) = \sin^2(\beta_x h/2) + \sin^2(\beta_y h/2) + 2(\alpha-1)\sin^2(\beta_x h/2)\sin^2(\beta_y h/2) \tag{4.126}$$

which follows directly from (4.123). For stability conditions, it is straightforward to show that [21, 32]:

$$\alpha \geq 0\,, \quad \lambda \leq \lambda_{\max} = \min\left(1, \frac{1}{\sqrt{2\alpha}}\right) \tag{4.127}$$

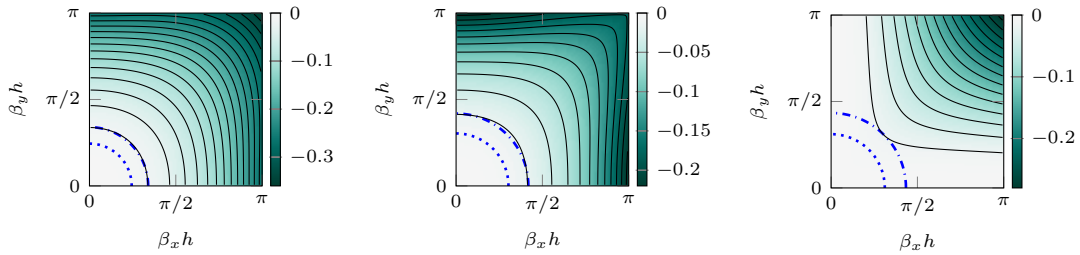
It is worth making a few remarks about the above stability conditions:

- We recover the usual $\lambda \leq \sqrt{1/2}$ for the simplest 2-D scheme ($\alpha = 1$), and for $\alpha \in [0, 0.5]$ we have $\lambda_{\max} = 1$.
- Through a modified equation analysis, it can be shown that $\lambda = \sqrt{4/3}$ would result in a global fourth-order truncation error, but this would render the scheme unstable; as such, this scheme is at most second-order accurate [32].
- Under the choices of $0 \leq \alpha \leq 1$ and $\lambda \leq \frac{1}{\sqrt{1+\alpha}}$, one has a set of schemes that permit concretely passive realisations as wave scattering networks [21], generally in the form of "digital waveguide meshes" (DWMs) [286], which are stable by construction [245]. In particular, for $\lambda = \frac{1}{\sqrt{1+\alpha}}$, one has the so-called "interpolated DWMs" in 2-D [224, 21].[18]

Let us now consider the dispersion errors for two special cases, $\alpha = 2/3$ and $\alpha = 1/2$. The first of these is the well-known isotropic scheme [275], and the second special case was identified in [22] as being optimally efficient in this family of explicit schemes. It has since been renamed the "interpolated wideband scheme (IWB)" [135], due to a special property related to its *cutoff frequency*, to be discussed shortly.

[18] Due to the close relationship between digital waveguide structures and finite difference schemes, this family of schemes could be said to be a family of "interpolated schemes" [21].

In Fig. 4.14 we display the relative dispersion errors for $\alpha = 2/3$ with the passive choice $\lambda = \sqrt{3/5}$, as well as the relative dispersion errors for $\alpha \in \{2/3, 1/2\}$ with $\lambda = \lambda_{\max}$. It can be seen that dispersion errors are generally improved over the simplest case seen earlier ($\alpha = 1$, Fig. 4.8). Comparing Figs. 4.14(a) and 4.14(b), we notice that $\lambda = \lambda_{\max}$ results in improved dispersion over $\lambda = \sqrt{1/2}$ for this isotropic scheme. Indeed, that $\lambda = \lambda_{\max}$ provided optimal dispersion was also observed for the simple scheme ($\alpha = 1$), and this is also the case for $\alpha = 1/2$. As for the case $\alpha = 1/2$ (in Fig. 4.14(b)), it can be seen that this case features slightly improved dispersion error over the cases of $\alpha = 2/3$, also at a higher Courant number (permitting a larger time-step for a fixed h). It is also worth pointing out that there is no dispersion error along axial directions for this scheme.[19]



(a) $\alpha = 2/3$, $\lambda = \sqrt{3/5}$ (b) $\alpha = 2/3$, $\lambda = \sqrt{3/4}$ (c) $\alpha = 1/2$, $\lambda = 1$

Figure 4.14: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for nine-point explicit schemes, with Courant numbers set as indicated. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

On cutoff frequencies and the interpolated wideband scheme

Cutoff frequencies [275, 294] have been a topic of interest in the room acoustics literature [131], and it is worth discussing them in the context of this family of schemes.

Recalling that $\boldsymbol{\beta}_h$ can be written as $\boldsymbol{\beta}_h = \beta_h \hat{\boldsymbol{\beta}}$, let us start by defining a *directional cutoff frequency*:

$$\omega_{k,\mathrm{d}}(\hat{\boldsymbol{\beta}}) = \max_{\beta_h} \omega_k(\beta_h \hat{\boldsymbol{\beta}}), \quad \boldsymbol{\beta}_h \in \mathbb{B} \tag{4.128}$$

This represents the maximum normalised frequency that is reproduced by the scheme for a given wave direction, and above which only evanescent solutions can satisfy the dispersion relation along the fixed wave direction.

The minimum and maximum of $\omega_{k,\mathrm{d}}(\hat{\boldsymbol{\beta}})$ over all wave directions provide two more quantities of interest, which we call the *minimum cutoff frequency* and *global cutoff frequency* of the scheme, and they are defined respectively as:

$$\omega_{k,\mathrm{m}} = \min_{\hat{\boldsymbol{\beta}}} \omega_{k,\mathrm{d}}(\hat{\boldsymbol{\beta}}), \qquad \omega_{k,\mathrm{g}} = \max_{\hat{\boldsymbol{\beta}}} \omega_{k,\mathrm{d}}(\hat{\boldsymbol{\beta}}) = \max_{\boldsymbol{\beta}_h} \omega_k(\boldsymbol{\beta}_h), \qquad \boldsymbol{\beta}_h \in \mathbb{B} \tag{4.129}$$

The minimum cutoff frequency $\omega_{k,\mathrm{m}}$ has been used as a starting measure of "valid bandwidth" in finite difference simulations [131, 136], since above this frequency one can expect modal densities to be incorrect. However, dispersion errors may be significant at such frequencies, in which cases accuracy considerations will be more critical in delimiting a "valid bandwidth".

[19] This follows from the fact that for $\alpha = 1/2$ the dispersion relation to that of the exact 1-D scheme when either $\beta_x = 0$ or $\beta_y = 0$.

For many of the Cartesian schemes to appear, one finds that the numerical dispersion relation increases monotonically with the wavenumber along any fixed wave direction. In those cases, these minimum and global cutoff frequencies are found at the extremities of $\mathbb{B}$. Furthermore, in such cases, the cutoff frequencies appear along axial and diagonal directions, at $\beta_h = \pi$ and $\beta_h = \sqrt{2}\pi$ respectively. However, it cannot be assumed that all schemes satisfy this property, as will be seen with one example scheme ($\alpha = 0$) shortly.

We should note that by (4.26) and (4.129), it is immediately the case that:

$$\omega_{k,\mathrm{m}} \leq \omega_{k,\mathrm{g}} \leq \pi \tag{4.130}$$

provided that $\lambda \leq \lambda_{\max}$. The above follows from the numerical dispersion relations we have seen, but also from temporal sampling considerations. As mentioned previously, the second inequality is satisfied with equality for $\lambda = \lambda_{\max}$.

Now that we have defined these cutoff frequencies, it is worth tabulating them for some of the schemes we have seen so far; these are listed in Tables 4.2 and 4.3. In Table 4.2 we find schemes for which $\lambda = \lambda_{\max}$. For these schemes $\omega_{k,\mathrm{g}} = \pi$, as expected. For the leggy schemes, $\omega_{k,\mathrm{m}}$ does not move from 0.5π for all M when $\lambda = \lambda_{\max,M}$, due to (4.122). In Table 4.3 we list cutoff frequencies for schemes with $\lambda < \lambda_{\max}$, including the optimised leggy schemes. As expected, choosing $\lambda < \lambda_{\max}$ results in $\omega_{k,\mathrm{m}}$ and $\omega_{k,\mathrm{g}}$ significantly lower than the temporal Nyquist.

Table 4.2: Cutoff frequencies for various $(K+1)$-point 2-D schemes with $\lambda = \lambda_{\max}$

K	stencil family	parameters $\{\alpha\}$	λ (approx.)	$\omega_{k,\mathrm{m}}$	$\omega_{k,\mathrm{g}}$
4	leggy/compact		0.707	0.500π	1.000π
8	leggy	$\{\alpha_{m,2}\}$	0.612	0.500π	1.000π
12	leggy	$\{\alpha_{m,3}\}$	0.575	0.500π	1.000π
16	leggy	$\{\alpha_{m,4}\}$	0.555	0.500π	1.000π
8	compact	$\alpha = 2/3$	0.866	0.667π	1.000π
8	compact	$\alpha = 1/2$	1.000	1.000π	1.000π

Table 4.3: Cutoff frequencies for various $(K+1)$-point 2-D schemes with $\lambda < \lambda_{\max}$

K	stencil family	parameters $\{\alpha\}$	λ (approx.)	$\omega_{k,\mathrm{m}}$	$\omega_{k,\mathrm{g}}$
8	leggy	$\{\alpha_{m,2}\}$	0.410	0.314π	0.467π
12	leggy	$\{\alpha_{m,3}\}$	0.299	0.240π	0.348π
16	leggy	$\{\alpha_{m,4}\}$	0.255	0.211π	0.304π
8	compact	$\alpha = 2/3$	0.707	0.500π	0.608π
8	compact	$\alpha = 1/2$	0.707	0.500π	0.500π

It can also be seen from these tables that the nine-point scheme with $\alpha = 1/2$ and $\lambda = 1$ (the so-called IWB scheme) is the only scheme seen so far that has the property that $\omega_{k,\mathrm{m}} = \omega_{k,\mathrm{g}} = \pi$. In other words, this is the only scheme seen so far in which frequencies up to the temporal Nyquist are reproduced (although not necessarily with sufficient accuracy) for every wave direction. This property, from which the IWB scheme derives its name, has been generally accepted as a desirable feature for room acoustics applications [131, 135, 279].

However, it is important to remember that in case of the IWB scheme, dispersion errors are quite significant along diagonal directions (nearly 30% error). In fact, somewhat ironically, the IWB scheme

only has this special property *by virtue of dispersion errors*. Indeed, one can calculate that for $\lambda = 1$ and a numerical dispersion relation that is monotonically increasing with wavenumber for every wave direction, it can only be true that $\omega_{k,\mathrm{m}} = \pi$ when the dispersion error is $1 - \sqrt{1/2} \approx 30\%$ for wavenumbers with $\beta_h = \sqrt{2}\pi$ (along the diagonal directions), which is precisely what is found in the IWB scheme. With that said, only a fraction of the bandwidth produced by the IWB scheme can be regarded as "accurate", unless a given application can tolerate such significant errors.

Accuracy considerations such as these are arguably more critical than cutoff frequencies in defining a "valid bandwidth" for a given application. To this end, one could define *critical temporal frequencies* (as opposed to critical wavenumbers) above which dispersion errors are beyond some threshold as:

$$\omega_{k,X\%} = \min_{\hat{\boldsymbol{\beta}}} \omega_k(\beta_{h,X\%}\hat{\boldsymbol{\beta}}) \tag{4.131}$$

where $X\%$ could be chosen as, e.g., $X\% \in (0, 5\%]$ to give some degree of confidence in the approximate solution.[20] Provided that the relative dispersion error reaches $X\%$, it must be the case that:

$$\omega_{k,X\%} \leq \pi \tag{4.132}$$

For the error thresholds $X\% \in \{1\%, 2\%\}$, which are targeted in this section, it is also safe to say that:

$$\omega_{k,1\%} \approx \lambda\beta_{h,1\%}\,, \quad \omega_{k,2\%} \approx \lambda\beta_{h,2\%} \tag{4.133}$$

and furthermore that:

$$\omega_{k,1\%} < \omega_{k,\mathrm{m}}\,, \quad \omega_{k,2\%} < \omega_{k,\mathrm{m}} \tag{4.134}$$

For this reason and for brevity, $\omega_{k,\mathrm{m}}$ and $\omega_{k,\mathrm{g}}$ will mostly be ignored in the rest of this chapter, and we will focus on critical wavenumbers $\beta_{h,X\%}$ for $X\% \leq 2\%$.

The rotated five-point scheme and spatial decoupling

The rotated five-point scheme ($\alpha = 0$) is another special case of interest, and in this short section we point out some important properties of this 2-D scheme. Some of these properties have important analogues in 3-D (which are inherently more difficult to visualise and explain) that have previously gone unnoticed in the literature (e.g., in [136]).

The stencil used in this scheme is clearly a rotated and scaled version of the standard five-point stencil, as can be seen from Fig. 4.11(b). Consider then the checkerboard colouring of a Cartesian grid shown in Fig. 4.15. It should be clear that the red points make up a 45°-rotated and scaled Cartesian grid, and the black points another. Let us denote these two subsets of points as "subgrids". It is not difficult to realise that if only one subgrid is initialised to non-zero values, the use of the rotated five-point scheme will never excite the other subgrid. As such, one could take the point of view that the numerical solution given by this scheme is actually two decoupled (and independent) numerical solutions, each on respective rotated Cartesian grids [21, 135]. This phenomenon will be called "spatial decoupling".

If we take the point of view that there is spatial decoupling and two independent solutions, then making use of this rotated operator is clearly wasteful, since one of two decoupled solutions will suffice.[21]

[20] It should also be the case that $\min_{\hat{\boldsymbol{\beta}}} \omega_k(\beta_{h,X\%}\hat{\boldsymbol{\beta}}) \equiv \omega_k(\beta_{h,X\%}\hat{\boldsymbol{\beta}})$.

[21] That one of two solutions suffices is related to the bandwidth of the approximate solutions. We will not go into great detail here, but essentially in this case one can recover values of $\underline{u}_{\boldsymbol{i}}^n$ on one subset of points (red or black) from the other subset of points using ideal sinc (Fourier) interpolation.

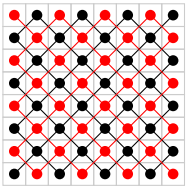


Figure 4.15: Checkerboard colouring of a Cartesian grid and connectivity of subgrids

One could also interpret the increased stability limit $\lambda \leq 1$ as an advantage over the standard scheme (as has been done in the past for closely-related staggered forms [20]), but one would be mistaken, as will be seen shortly.

Looking at the numerical dispersion of this rotated scheme, it is straightforward to rewrite the term $G(\boldsymbol{\beta}_h)$ in (4.126) with $\alpha = 0$ as:

$$G(\boldsymbol{\beta}_h) = \frac{1}{2}\left(\sin^2(\beta_x h/2 + \beta_y h/2) + \sin^2(\beta_x h/2 - \beta_y h/2)\right) \tag{4.135}$$

Let us introduce the auxiliary variables:

$$x' = \frac{x+y}{\sqrt{2}}\,, \quad y' = \frac{x-y}{\sqrt{2}}\,, \quad h' = \sqrt{2}h\,, \quad \lambda' = ck/h' \tag{4.136}$$

Then we have:

$$\frac{1}{h^2}G(\boldsymbol{\beta}_h) = \frac{1}{h'^2}\left(\sin^2(\beta_{x'} h'/2) + \sin^2(\beta_{y'} h'/2)\right) \tag{4.137}$$

and thus, we have a dispersion relation for the scheme:

$$\omega_k(\beta_h) = 2\sin^{-1}\left(\pm\lambda'\sqrt{\sin^2(\beta_{x'} h'/2) + \sin^2(\beta_{y'} h'/2)}\right) \tag{4.138}$$

In other words, and as expected, the dispersion relation is exactly that of the standard simplest scheme, only using rotated axes x' and y' and the larger grid spacing $h' = \sqrt{2}$ for the rotated grid. For numerical stability then we have $\lambda' \leq \sqrt{1/2}$, and there is no real advantage over the simplest scheme in the ability to choose a larger time-step (since the effective grid spacing is also larger). It is perhaps more appropriate to simply state that this scheme is of no practical use, as it is encapsulated by the simplest scheme.

Before moving on, it is worth discussing the appropriate wavenumber cell for this scheme. By the above manipulations it makes sense to rotate and also scale the conventional square wavenumber cell, leading to a rotated square (diamond) wavenumber cell, which we denote as $\mathbb{B}'$, with vertices at $(\pm\pi, 0)$ and $(0, \pm\pi)$ in the normalised wavenumber domain. The numerical dispersion relation of the rotated scheme is shown in Fig. 4.16 calculated over the entire wavenumber cell $\mathbb{B}$, with the boundary of $\mathbb{B}'$ drawn in white. It can be seen that the dispersion relation is periodic in $\mathbb{B}$, but more fundamentally it is periodic in $\mathbb{B}'$.[22] It can also be seen that what appears inside the white diamond in Fig. 4.16 is indeed the dispersion relation shown in Fig. 4.6(a), only shrunk by a factor of $\sqrt{2}$ (to account for the different Courant number) and rotated by 45-degrees, as explained by the above analysis. Furthermore, if $\mathbb{B}$ is taken to be the wavenumber cell, then one does not have the aforementioned property that the numerical dispersion increases *monotonically* with the wavenumber along any given wave direction (within the limits of the wavenumber cell).[23]

[22] It is important to note that such periodicity is only the case for $\alpha = 0$ in this nine-point family.

[23] Consequently, inversion of the numerical dispersion for $\boldsymbol{\beta}_h \in \mathbb{B}$, as carried out in [135], is not well-defined for this scheme.

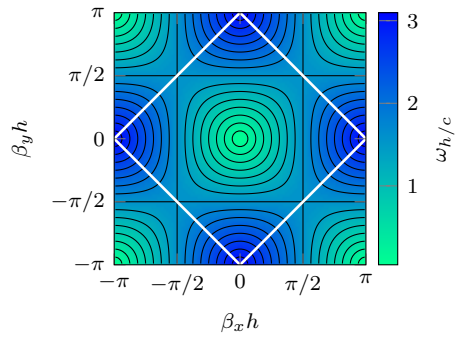


Figure 4.16: Numerical dispersion relation $\omega_{h/c}(\boldsymbol{\beta}_h)$ of rotated five-point scheme with $\lambda = 1$, for $\boldsymbol{\beta}_h \in \mathbb{B}$. Contour lines denote increments of $\pi/16$ rad/s. The boundary of $\mathbb{B}'$ is drawn in white. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

It is important to take this spatial decoupling into consideration when analysing the numerical phase velocity of this scheme, or the relative dispersion error; otherwise, one could arrive at unnecessary conclusions about observed wave propagation in this scheme. For illustrative purposes, Fig. 4.17 shows the relative dispersion error, calculated over the entire wavenumber cell $\mathbb{B}$. As expected, what appears inside the blue diamond is clearly a rotated and scaled version of what appears in Fig. 4.7(a). What appears outside of the inner blue diamond pertains to wavenumbers that would have aliased on the rotated subgrids, so under the spatial decoupling point of view, the associated values of $\tilde{v}_\mathrm{p}$ deceptive. Of course, one could extract some meaning for these relative phase velocities outside of $\mathbb{B}'$, under the point of view that this scheme produces a single numerical solution (rather than two decoupled solutions). This would lead to the conclusion that this scheme features a maximum dispersion error of 100% for normalised wave vectors with magnitude $\sqrt{2}\pi$, but this can also be interpreted as the zero wavenumber relative to another period of this dispersion relation, or as a description of the case where only one subgrid is excited.[24] We will return to the topic of spatial decoupling in Section 4.6.1.

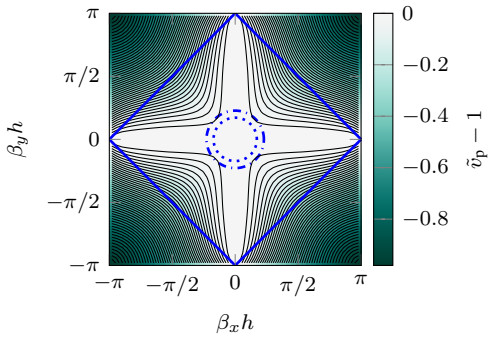


Figure 4.17: Relative dispersion error of rotated five-point scheme with $\lambda = 1$ calculated for $\boldsymbol{\beta}_h \in \mathbb{B}$. Contour lines denote 2% increments. The boundary of $\mathbb{B}'$ is drawn in blue (solid line).

4.3.5 High-order $(K+1)$-point schemes in 2-D with $K \leq 24$

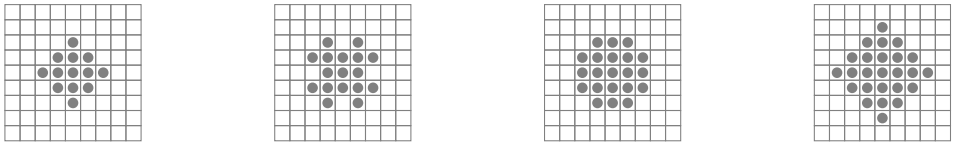
At this point we move on to schemes that employ stencils larger in size than the conventional compact nine-point stencil, yet can still be said to be "compact in space". For lack of a better term, we will call

[24] If the non-excited subgrid remains zero, diagonally-directed plane waves with wavenumbers $\sqrt{2}\pi$ persist on the full Cartesian grid for all time.

these schemes "wide-stencil" compact schemes.

Using additional free parameters provided by combining multiple discrete Laplacians, it will become possible to achieve high orders of accuracy (in both space and time) through the manipulation of modified equations, as in the 1-D case. Although, for the most part, we omit lengthy calculations required to arrive at high orders of accuracy, and we point the reader to [27] (published work involving the current author). We will also leave out the derivation of stability conditions, carried out via von Neumann analyses, and sometimes with the aid of a computer (and MATLAB [164] or Mathematica [308]). We should also note that since the schemes to be presented feature balanced orders of accuracy (in space and time), optimal numerical dispersion within these schemes are generally found for $\lambda = \lambda_{\max}$.

Among the infinite number of possible stencils afforded by (4.108), for brevity in this section, we focus on the following stencils of points shown in Fig. 4.18. The stencils in Fig. 4.18(a) and Fig. 4.18(d) are maximally compact in terms of the ℓ_1-norm (the "taxicab norm" or "Manhattan distance"), and the stencil in Fig. 4.18(c) is maximally compact in terms of a Euclidean norm. The stencil in Fig. 4.18(b) is only compact in a loose sense, but it leads to an interesting set of schemes.



(a) $K = 12$ (b) $K = 16$ (c) $K = 20$ (d) $K = 24$

Figure 4.18: $(K+1)$-point stencils of interest on 2-D Cartesian grid.

Using only these stencils, a family of discrete Laplacians could be written as:

$$\delta_{\Delta,\Upsilon} = \alpha_{(1,0)}\delta_{\Delta,(1,0)} + \alpha_{(1,1)}\delta_{\Delta,(1,1)} + \alpha_{(2,0)}\delta_{\Delta,(2,0)} + \alpha_{(2,1)}\delta_{\Delta,(2,1)} + \alpha_{(3,0)}\delta_{\Delta,(3,0)} \tag{4.139}$$

with the consistency constraint:

$$\alpha_{(1,0)} + \alpha_{(1,1)} + \alpha_{(2,0)} + \alpha_{(2,1)} + \alpha_{(3,0)} = 1 \tag{4.140}$$

The family of two-step schemes is then of the form (4.20) with $\delta_\Delta = \delta_{\Delta,\Upsilon}$ (defined above).[25] In terms of an update recursion for this scheme, we have:

$$\begin{aligned}\underline{u}^{n+1}_{i_x,i_y} = &-\underline{u}^{n-1}_{i_x,i_y} + \left(2 - \lambda^2(4\alpha_{(1,0)} + 2\alpha_{(1,1)} + \alpha_{(2,0)} + \frac{4}{5}\alpha_{(2,1)} + \frac{4}{9}\alpha_{(3,0)})\right)\underline{u}^{n}_{i_x,i_y} \\ &+ \lambda^2\left(\alpha_{(1,0)}Q^n_{\boldsymbol{i},(1,0)} + \frac{\alpha_{(1,1)}}{2}Q^n_{\boldsymbol{i},(1,1)} + \frac{\alpha_{(2,0)}}{4}Q^n_{\boldsymbol{i},(2,0)} + \frac{\alpha_{(2,1)}}{10}Q^n_{\boldsymbol{i},(2,1)} + \frac{\alpha_{(3,0)}}{9}Q^n_{\boldsymbol{i},(3,0)}\right)\end{aligned} \tag{4.141}$$

with $Q^n_{\boldsymbol{i},(q_1,q_2)}$ defined by (4.107).

In order to achieve a high-order of accuracy using modified equation methods, we first need to impose constraints on the free parameters for high orders of isotropy in the local truncation error. Let us denote the set of free parameters by a vector $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha} = \left(\alpha_{(1,0)}, \alpha_{(1,1)}, \alpha_{(2,0)}, \alpha_{(2,1)}, \alpha_{(3,0)}\right) \tag{4.142}$$

[25] It is worth mentioning that the compact nine-point and $2M$th-order accurate leggy stencils up to $M = 3$ are special cases of this family of discrete Laplacians.

From a Taylor expansion of $\delta_{\Delta,\Upsilon}$, one can find that the condition for fourth-order isotropy is:

$$(5, -10, 20, -7, 45) \cdot \boldsymbol{\alpha} = 0 \tag{4.143}$$

and a further constraint for sixth-order isotropy is:

$$(1, -4, 16, -7, 81) \cdot \boldsymbol{\alpha} = 0 \tag{4.144}$$

Under these constraints (and the consistency constraint), a Taylor expansion can be written for the scheme as:

$$\delta_{\square}\underline{u} = \square\underline{u} + \frac{1}{12}(k^2\partial_t^4 - c^2h^2F_1(\boldsymbol{\alpha})\Delta^2)\underline{u} + \frac{1}{360}(k^4\partial_t^6 - c^2h^4F_2(\boldsymbol{\alpha})\Delta^3)\underline{u} + \mathcal{O}(h^6) + \mathcal{O}(k^6) \tag{4.145}$$

where $F_1(\boldsymbol{\alpha})$ and $F_2(\boldsymbol{\alpha})$ are affine functions in terms of the remaining free α parameters. A condition for fourth-order accuracy is then:

$$\lambda^2 = F_1(\boldsymbol{\alpha}) \tag{4.146}$$

which, assuming λ constant, leaves a fourth-order truncation error, since then:

$$\delta_{\square}\underline{u} = \square\underline{u} + \frac{k^2}{12}(\partial_t^2 + c^2\Delta)\square\underline{u} + \mathcal{O}(k^4) \tag{4.147}$$

$$= \square(1 + \mathcal{O}(k^2))\underline{u} + \mathcal{O}(k^4) \tag{4.148}$$

$$= \square\underline{u} + \mathcal{O}(k^4) \tag{4.149}$$

where we have essentially used the same procedure as in (4.90). By similar calculations, we find that for sixth-order accuracy we additionally need:

$$\lambda^4 = F_2(\boldsymbol{\alpha}) \tag{4.150}$$

with which we arrive at:

$$\delta_{\square}\underline{u} = \square\underline{u} + \frac{k^2}{12}(\partial_t^2 + c^2\Delta)\square\underline{u} + \frac{k^4}{360}(\partial_t^4 + \partial_t^2c^2\Delta + c^4\Delta^2)\square\underline{u} + \mathcal{O}(h^6) + \mathcal{O}(k^6) \tag{4.151}$$

$$= \square(1 + \mathcal{O}(k^2) + \mathcal{O}(k^4))\underline{u} + \mathcal{O}(k^6) \tag{4.152}$$

$$= \square\underline{u} + \mathcal{O}(k^6) \tag{4.153}$$

It is worth briefly discussing the number of free parameters available and the constraints required to achieve fourth- and sixth-order accuracy. There are five free parameters in $\boldsymbol{\alpha}$ available and the Courant number is another free parameter. Consistency for the discrete Laplacian removes one degree of freedom in $\boldsymbol{\alpha}$, and fourth- and sixth-order isotropy each require one degree of freedom. So far we have not found it to be the case that high-orders of accuracy can be achieved unless λ is able to vary freely, so we must leave this as a free parameter.[26] Thus, we require three degrees of freedom in $\boldsymbol{\alpha}$ in order to possibly achieve a fourth-order accuracy, and we need *five* degrees of freedom in $\boldsymbol{\alpha}$ in order to possibly achieve a sixth-order accuracy. Thus, it will only be possible to achieve sixth-order accuracy with the 25-point scheme (while leaving λ as a free parameter). Similarly, for the 13-, 17-, and 21-point schemes, there are a sufficient number of free parameters for a possible fourth-order accuracy, and in the case of the

[26] For example, in the 2-D nine-point compact scheme, $\lambda = \sqrt{4/3}$ would have resulted in a fourth-order scheme but this would also be unstable [32], as mentioned previously.

21-point scheme we can additionally achieve a degree of isotropy in the error to sixth-order.

We leave out detailed derivations of solving for $\boldsymbol{\alpha}$ and λ (or $\lambda_{\max}$), but the resulting parameters of the high-order schemes are listed in Table 4.4.[27] It is worth noting that all of these schemes have α parameters of both signs (as they must for high-order accuracy), but the stability constraint $G(\boldsymbol{\beta}_h) \geq 0$ is satisfied. It is also notable that the 13- and 25-point schemes have $\lambda_{\max} = \sqrt{1/2}$, like the simplest scheme; more will be said about this in Section 4.5.4.

The relative dispersion errors for these schemes are plotted in Fig. 4.19. As expected by the higher-order accuracy, dispersion errors are improved over the second-order accuracy compact nine-point schemes seen previously. We note also that the diamond-shaped (13-point and 25-point) high-order schemes keep the property of having exact dispersion along grid diagonals. In order to confirm that these schemes are indeed high-order accurate, we plot the relative dispersion errors along grid axes for each respective scheme in Fig. 4.20, including the simplest five-point scheme and the isotropic nine-point scheme. It can be seen that the dispersion errors decrease with rates that agree with the orders of accuracy derived here.

Table 4.4: Parameters for high-order accurate schemes on 2-D Cartesian grid (OoA = order of accuracy)

K	OoA	$\alpha_{(1,0)}$	$\alpha_{(1,1)}$	$\alpha_{(2,0)}$	$\alpha_{(2,1)}$	$\alpha_{(3,0)}$	λ^2	λ (approx.)
12	$(4,4)$	1	$\frac{1}{6}$	$-\frac{1}{6}$	0	0	$\frac{1}{2}$	0.707
16	$(4,4)$	$\frac{11}{16}$	$\frac{5}{12}$	0	$-\frac{5}{48}$	0	$\frac{3}{4}$	0.866
20	$(4,4)$	$\frac{22+5\sqrt{11}}{60}$	$\frac{10+\sqrt{11}}{30}$	$\frac{4-\sqrt{11}}{30}$	$\frac{2-\sqrt{11}}{12}$	0	$\frac{3(6-\sqrt{11})}{10}$	0.897
24	$(6,6)$	$\frac{1447}{1440}$	$\frac{11}{45}$	$-\frac{11}{45}$	$-\frac{7}{144}$	$\frac{7}{160}$	$\frac{1}{2}$	0.707

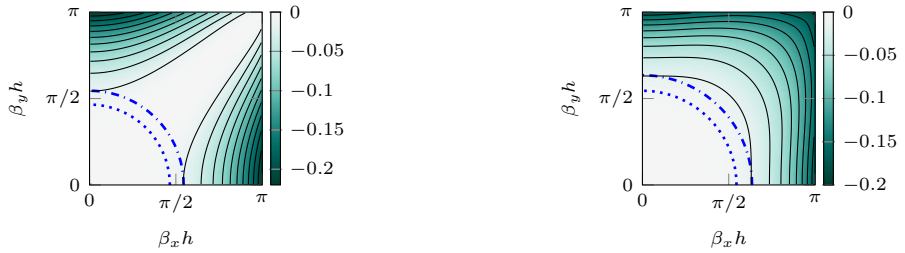
4.3.6 Optimised $(K+1)$-point schemes in 2-D with $K \leq 24$

In this section we consider the same family of schemes but we optimise the free parameters (in $\boldsymbol{\alpha}$) for wideband accuracy in numerical dispersion (i.e., accuracy over a wide range of wavenumbers).

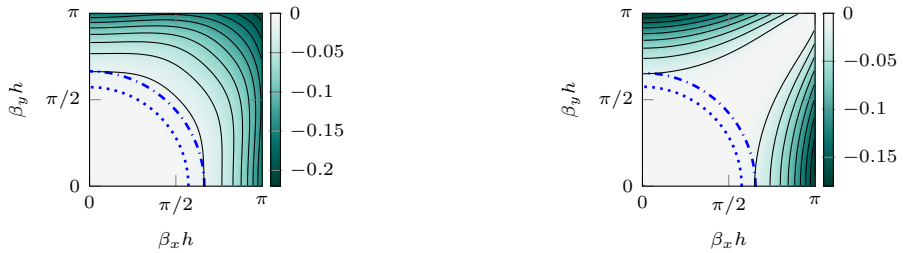
By definition, improvements in dispersion errors obtained from high-orders of accuracy are generally seen near the zero wavenumber (DC), but such improvements tend to stretch out to higher wavenumbers, leading to some wideband accuracy. However, if wideband accuracy, according to some percent-error threshold, is the ultimate interest, then it can be beneficial to abandon high orders of accuracy in order to optimise free parameters such that dispersion errors are minimised over a wider range of wavenumbers. Additionally, one could weight error criteria by some measure of computational cost that is relevant and critical to the computing hardware on hand.

There are many avenues to explore under this generally non-linear optimisation problem, and it will be treated only briefly here. This optimisation problem has been studied in a preliminary way in [27] (work involving the present author), making use of an L^2-norm in the dispersion error. The parameter space that arises for the 13-, 17-, and 21-point stencils has been explored to some degree in [96] (work by the present author), making use of a more heuristic approach that aimed to maximise quantities like $\beta_{h,1\%}$ while taking into account computational costs associated with the choice of λ. For brevity, we

[27] It should be mentioned that the 13-point and 21-point schemes are also found in [54], and the 13- and 25-point schemes are also derived in [282, 158], wherein an entire family of diamond-shaped stencils are identified with high even-order accuracy and stability limits that appear to remain at $\lambda_{\max} = \sqrt{1/2}$ for higher orders of accuracy. The 17-point discrete Laplacian also appears in [309], but we are unaware of any prior use in the context of the second-order wave equation, apart from published work involving the current author [27, 96].



(a) $K = 12$ (b) $K = 16$



(c) $K = 20$ (d) $K = 24$

Figure 4.19: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for $(K+1)$-point wide-stencil high-order schemes, with $\boldsymbol{\alpha}$ and λ as indicated in Table 4.4. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

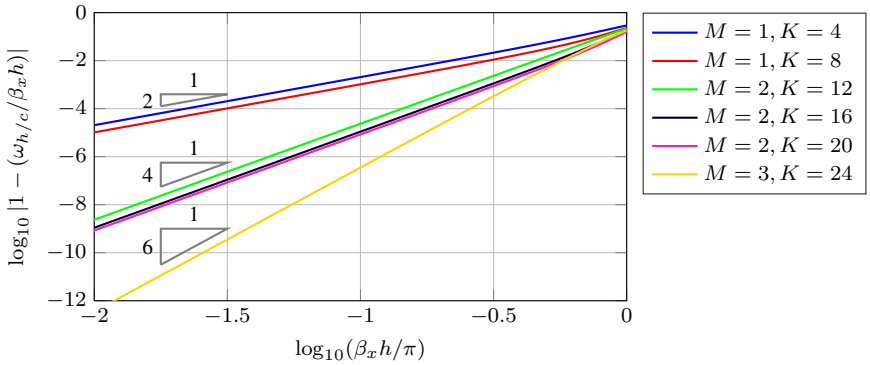


Figure 4.20: Relative dispersion errors on a log-log scale, along x-axis for $2M$th-order accurate $(K+1)$-point schemes in 2-D with $\boldsymbol{\alpha}$ and λ as indicated in Table 4.4.

give some results from that study, and subsequently we will compare such optimised schemes to the various 2-D schemes covered so far, in order to demonstrate the possible benefits that can be obtained from such optimisations.

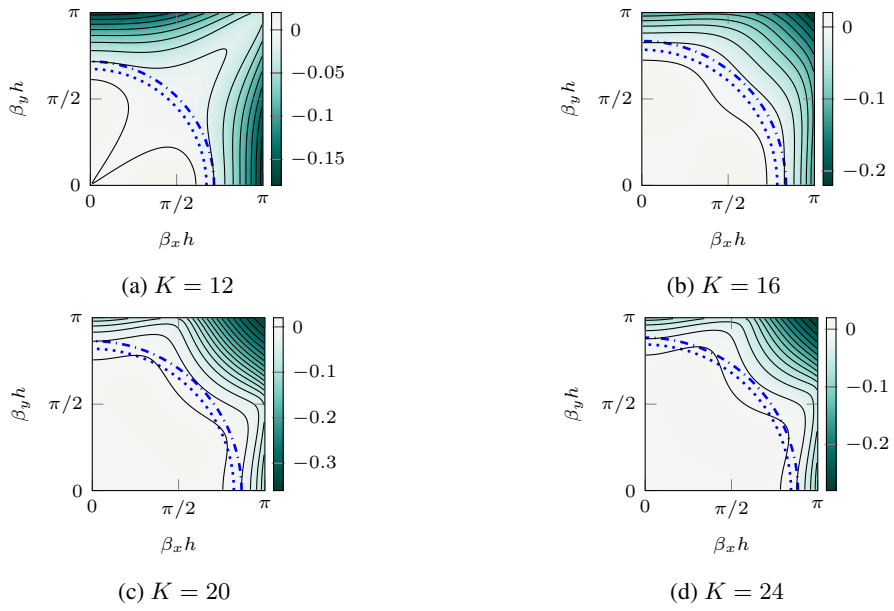
Parameters for some optimised schemes are given in Table 4.5, and the relative dispersion errors are plotted in Fig. 4.21. As can be seen from the figure, improvements are obtained for wideband accuracy considering one- and two-percent error thresholds, and from the table it can be seen that Courant numbers are higher than the optimised leggy schemes seen earlier.

4.3.7 Comparing schemes for wideband accuracy

Let us now attempt to compare these various finite difference schemes for the 2-D wave equation. The focus of this comparison will be the metric $\beta_{h,1\%}$, since optimised parameters were selected with this metric in mind.

Table 4.5: Parameters for second-order accurate optimised schemes

K	$\alpha_{(1,0)}$	$\alpha_{(1,1)}$	$\alpha_{(2,0)}$	$\alpha_{(2,1)}$	$\alpha_{(3,0)}$	λ
12	0.948	0.298	−0.246	0	0	0.726
16	0.660	0.494	0	−0.154	0	0.891
20	0.591	0.532	0.077	−0.200	0	0.931
24	0.639	0.525	0.024	−0.220	0.032	0.907



(a) $K = 12$ (b) $K = 16$

(c) $K = 20$ (d) $K = 24$

Figure 4.21: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for $(K+1)$-point wide-stencil optimised schemes, with α parameters and λ set as indicated in Table 4.5. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

We sort these schemes into four families, which are named and described as follows:

- "leggy, $\lambda = \lambda_{\max,M}$": $(2, 2M)$-order accurate leggy schemes with $\lambda = \lambda_{\max,M}$.
- "leggy, $\lambda = \lambda_{1\%,M}$": $(2, 2M)$-order accurate leggy schemes with $\lambda = \lambda_{1\%,M}$ (optimised for 1% error threshold).
- "compact, constrained": compact schemes with constraints on $\boldsymbol{\alpha}$ for high-order and/or isotropy, with $\lambda = \lambda_{\max}$. This includes the isotropic 9-point scheme ($\alpha = 2/3$).
- "compact, optimised": compact schemes with $\boldsymbol{\alpha}$ optimised with respect to 1% error threshold, with $\lambda = \lambda_{\max}$. This includes the 9-point scheme with $\alpha = 1/2$ (the IWB scheme).

The five-point scheme belongs to all four families. The associated Courant numbers and values for $\beta_{h,1\%}$ and $\beta_{h,2\%}$ for these schemes can be found in Table 4.6 under the above organisation.

In Fig. 4.22 we plot the critical wavenumbers $\beta_{h,1\%}$ for schemes in these four families, as a function of stencil size $K + 1$. With regards to this figure, it can be seen that the compact optimised schemes provide higher critical wavenumbers than the leggy schemes, apart from the case of the nine-point leggy scheme with λ optimised for $\beta_{h,1\%}$. As we are comparing schemes for fixed memory costs (by considering h to be fixed), we can conclude that if it is critical that memory storage requirements remain low (by using a large spatial step), it will generally be beneficial to make use of the optimised compact schemes

Table 4.6: Critical wavenumbers, $\beta_{h,1\%}$ and $\beta_{h,2\%}$, for various schemes with $(K+1)$-point stencils

K	family	OoA	λ (approx.)	$\beta_{h,1\%}$	$\beta_{h,2\%}$
4	—	$(2,2)$	0.707	0.219π	0.307π
8	leggy, $\lambda=\lambda_{\max}$	$(2,4)$	0.612	0.259π	0.372π
12	leggy, $\lambda=\lambda_{\max}$	$(2,6)$	0.575	0.268π	0.376π
16	leggy, $\lambda=\lambda_{\max}$	$(2,8)$	0.555	0.278π	0.388π
20	leggy, $\lambda=\lambda_{\max}$	$(2,10)$	0.541	0.284π	0.397π
24	leggy, $\lambda=\lambda_{\max}$	$(2,12)$	0.532	0.289π	0.404π
8	leggy, $\lambda=\lambda_{1\%,M}$	$(2,4)$	0.410	0.491π	0.556π
12	leggy, $\lambda=\lambda_{1\%,M}$	$(2,6)$	0.299	0.597π	0.651π
16	leggy, $\lambda=\lambda_{1\%,M}$	$(2,8)$	0.255	0.664π	0.713π
20	leggy, $\lambda=\lambda_{1\%,M}$	$(2,10)$	0.227	0.708π	0.754π
24	leggy, $\lambda=\lambda_{1\%,M}$	$(2,12)$	0.207	0.740π	0.783π
8	compact, high-order/isotropic	$(2,2)$	0.866	0.302π	0.414π
12	compact, high-order/isotropic	$(4,4)$	0.707	0.457π	0.544π
16	compact, high-order/isotropic	$(4,4)$	0.866	0.538π	0.631π
20	compact, high-order/isotropic	$(4,4)$	0.897	0.568π	0.663π
24	compact, high-order/isotropic	$(6,6)$	0.707	0.575π	0.651π
8	compact, optimised	$(2,2)$	1.000	0.309π	0.433π
12	compact, optimised	$(2,2)$	0.726	0.672π	0.716π
16	compact, optimised	$(2,2)$	0.891	0.783π	0.827π
20	compact, optimised	$(2,2)$	0.931	0.820π	0.865π
24	compact, optimised	$(2,2)$	0.907	0.843π	0.879π

to satisfy a one-percent error criterion, rather than leggy schemes or high-order schemes. One possible exception to this general trend is the nine-point leggy scheme with an optimised Courant number.

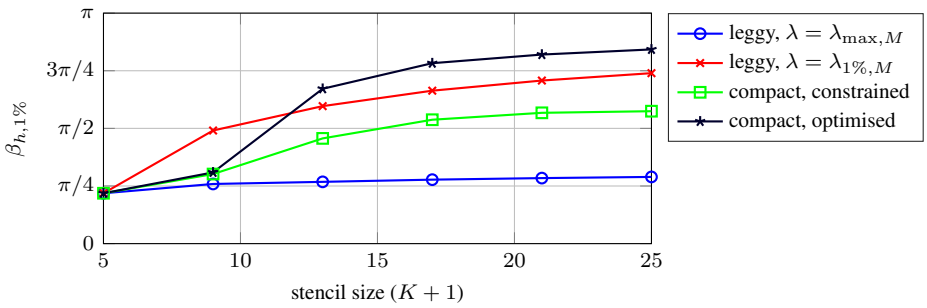


Figure 4.22: Critical wavenumber $\beta_{h,1\%}$ as a function of stencil size, for various families of schemes.

Let us now take into account the varying costs associated with the variation in Courant numbers, which, for fixed h, changes the number of samples to be computed for a given duration of solution time. We recall that the optimised leggy schemes feature Courant numbers far lower than unity, and thus require higher sample rates than the compact schemes, which tend to optimally use Courant numbers closer to unity. In order to normalise for the number of pointwise updates required for a unit volume of space-time (spatiotemporal densities) we can use the critical wavenumber $\beta_{h^*,X\%}$, defined by (4.38) (with $\mu=1$), rewritten here as:

$$\beta_{h^*,X\%} = \lambda^{1/3}\beta_{h,X\%} \tag{4.154}$$

Considering this normalisation for computational costs, we plot the critical wavenumbers $\beta_{h^*,1\%}$ as a function of stencil size for these schemes in Fig. 4.23. It can be seen that in terms of this rescaled metric, the leggy stencil schemes suffer from having small Courant numbers, and subsequently, the compact stencil schemes (both families) become more efficient choices than the leggy stencil $(2,2M)$-accurate schemes, by a large margin. However, the exception to this rule is (again) the nine-point leggy scheme with λ optimised, which is slightly superior than the compact nine-point family, according to

this metric.

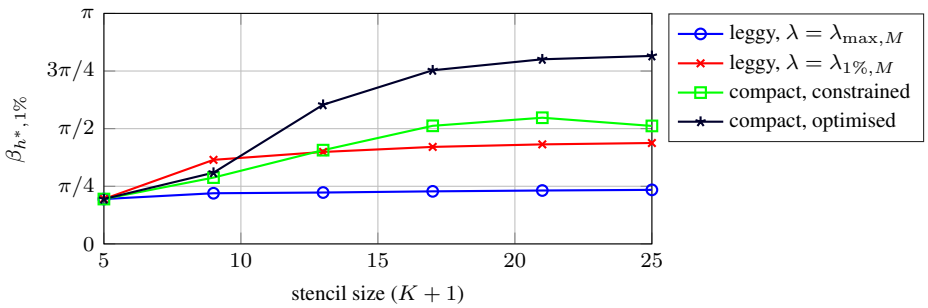


Figure 4.23: Critical wavenumbers $\beta_{h^*,1\%}$ (normalised for spatiotemporal densities) as a function of stencil size, for various families of schemes.

For a related, yet slightly different point of view, we show the *relative computational efficiencies* of these schemes for the one-percent error threshold in Fig. 4.24, where the simplest scheme is used as the reference scheme. Finally, we show the alternative relative computational efficiency metric, RCE-$X\%$-opts, in Fig. 4.25, which, interestingly, shows that the 17-point scheme is most efficient with regard to operation counts (rather, FLOP instructions), and that the leggy schemes with $\lambda = \lambda_{\max,M}$ become less efficient (in this regard) with increasing M.[28]

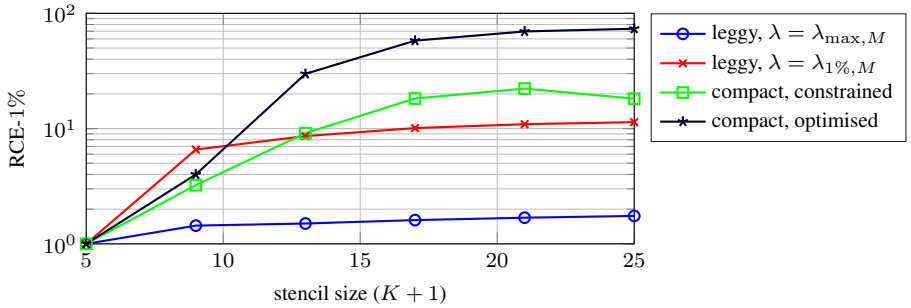


Figure 4.24: Relative computational efficiencies for one-percent error threshold (with simplest scheme as reference), as a function of stencil size, for various families of schemes.

It is important to remark that the comparisons and optimisations presented in this section are specific to the choice of $\beta_{h,1\%}$ (or $\beta_{h^*,1\%}$), which may not be appropriate for a given application, as this metric could be too strict or too relaxed. We have not considered other metrics for brevity. Nevertheless, this analysis can be used for any $\beta_{h,X\%}$ or $\omega_k(\beta_{h,X\%})$ for these schemes, but for best results, schemes to be optimised should take the desired metric into account.

4.3.8 Numerical example

At this point it is worth considering a numerical example of free-space wave propagation, in order to better view the behaviour of some of these schemes.[29] For this problem we consider $c = 1$ m/s, and for

[28] However, we must note that often in such schemes, computation times are not dictated by floating-point instructions, but rather by memory transfers, which in turn are highly dependent on the memory architecture of a given computing hardware. We cannot reasonably explore this topic further without going into more details about such computing hardware, as in [304]; see, e.g., [103].

[29] We must use a closed, bounded domain in practice, but the simulations are only run for a short duration, such that the boundaries have no influence.

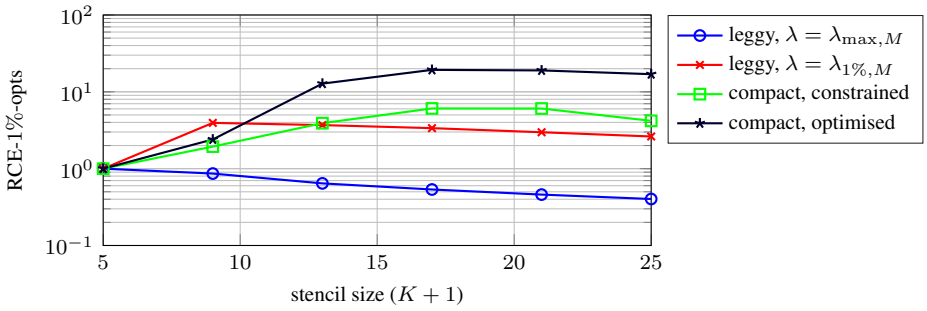


Figure 4.25: Relative computational efficiencies for one-percent error threshold, scaled by $(6/(K+2))$ (representing required FLOPS instructions or memory reads, relative to simplest scheme), as a function of stencil size $(K+1)$, for various families of schemes.

the initial conditions we make use of a Ricker wavelet:

$$u(\mathbf{x},0) = \left(1 - \|\mathbf{x}\|^2\sigma^{-2}\right)\exp\left(-\frac{\|\mathbf{x}\|^2}{2\sigma^2}\right), \quad u_t(\mathbf{x},0) = 0 \tag{4.155}$$

where σ^2 is set to $0.08\,\mathrm{m}^2$. The Ricker wavelet is the negative of the Laplacian of a Gaussian function, and as such it has a spectrum that is concentrated in higher frequencies than a Gaussian function with the same variance.

We will consider two scenarios in order to compute approximate solutions to this problem with the various schemes seen so far. In the first scenario we fix memory costs with $h = 0.2\,\mathrm{m}$. In the second scenario we normalise for spatiotemporal densities, by fixing $h^* = 0.2\,\mathrm{m}$ and setting $h = (\mu/\lambda)^{1/3}h^*$ for each scheme according to its respective λ, and here $\mu = 1$. In order to compare to some reference solution, an "exact" solution is computed with the sixth-order scheme on a refined grid ($h^* = 0.05\,\mathrm{m}$), for which there is virtually no dispersion error within the effective bandwidth of the chosen initial conditions. The initial two states for each scheme are computed with the following initialisation scheme [260]:

$$\underline{u}_{\boldsymbol{i}}^0 = [u(\mathbf{x},0)]_h\,, \quad \underline{u}_{\boldsymbol{i}}^1 = \underline{u}_{\boldsymbol{i}}^0 + k[u_t(\mathbf{x},0)]_h + \frac{c^2k^2}{2}\delta_\Delta \underline{u}_{\boldsymbol{i}}^0 \tag{4.156}$$

where $[u(\mathbf{x},0)]_h$ represents a Cartesian spatial sampling of $u(\mathbf{x},0)$.

For the first scenario, snapshots of the numerical solutions are shown in Fig. 4.26 at time $t \approx 10\,\mathrm{s}$, and Fig. 4.27 shows the numerical solutions from the various schemes as a function of time at two different read-out positions, each at a distance of $3\sqrt{2}\,\mathrm{m}$ from the origin, along the axial and diagonal directions (the extreme cases). Superimposed in the plots in Fig. 4.27 is the "exact" solution at those read-out positions. In Fig. 4.28 are the error signals of those outputs.

For the second scenario, snapshots of the computed approximate solutions are shown in Fig. 4.29 at time $t \approx 10\,\mathrm{s}$, and Fig. 4.30 shows the numerical solutions from the various schemes as a function of time at the two different read-out positions. The corresponding error signals are shown in Fig. 4.31.

Discussion

We start by discussing the results of the first scenario (fixed memory costs), as displayed by Figs. 4.26–4.28. In these figures it can be seen that the simple five-point scheme provides relatively poor results,[30]

[30]We should also note that the diagonally directed output in Fig. 4.26(b) demonstrates errors, although a frequency-domain analysis predicts no dispersion errors along the diagonal direction for the simplest scheme (for plane-wave solutions). These

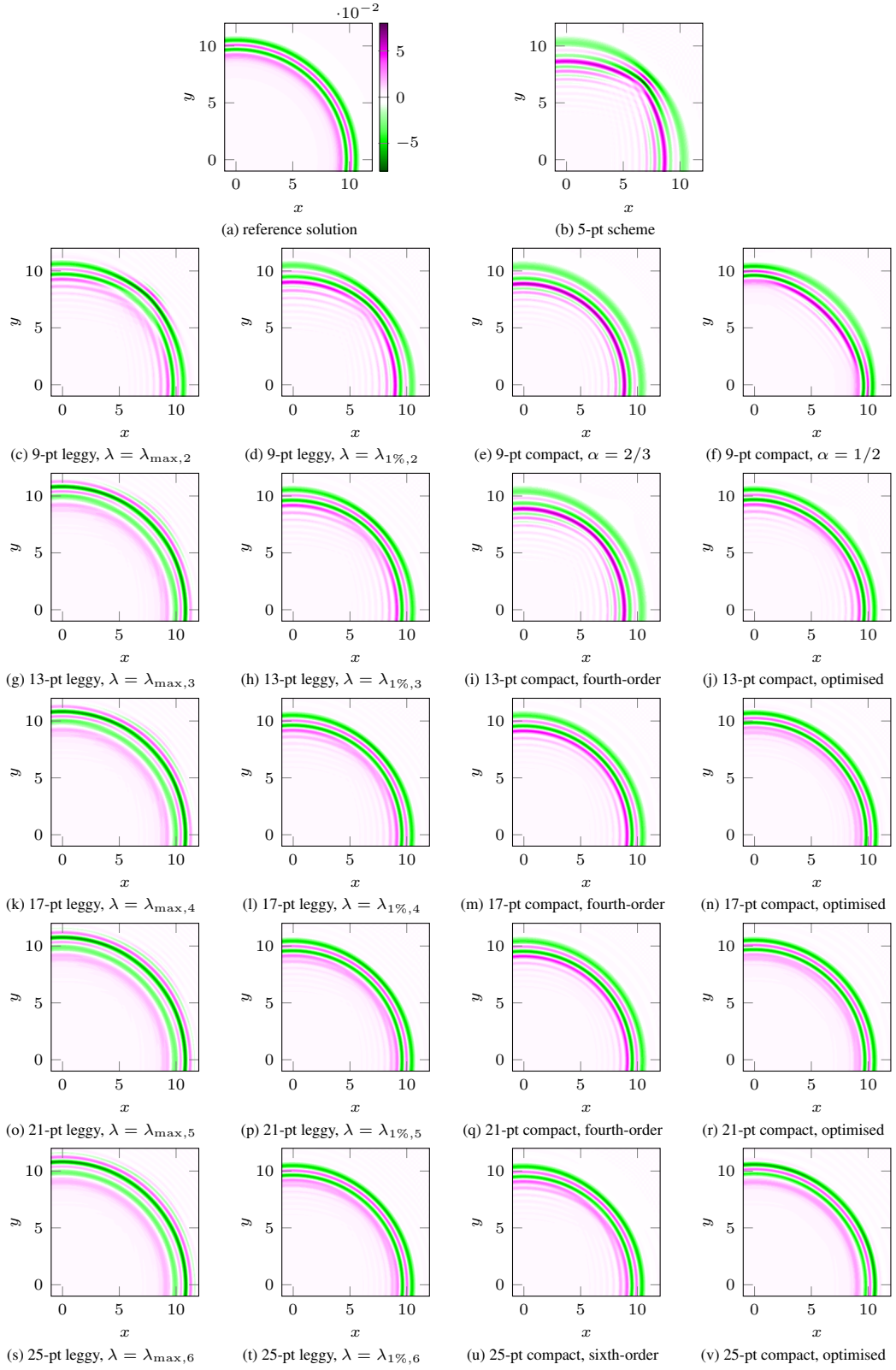
and generally, results are improved with increasing stencil size. The numerical approximations obtained with the optimised compact and leggy schemes for $K \geq 16$ are excellent, and the high-order 17–25-point schemes demonstrate agreements with the "exact" solution. Within the 9- and 13-point schemes, the compact and optimised scheme appear to give equally good results.

Next, we discuss the results of the second aforementioned scenario: fixed spatiotemporal densities, and thus fixed operation counts for equal stencil sizes. These results are displayed in Figs. 4.29–4.31. We can see that the results obtained from the compact schemes do not change significantly from the first scenario; this is due to the fact that they have Courant numbers close to unity, so they are not significantly affected by the spatiotemporal normalisations. On the other hand, in this scenario the leggy schemes give notably worse results than in the case of fixed memory costs; this is due to the fact that operate at higher sampling rates than the compact schemes (for the same h), and thus after normalising for spatiotemporal densities, they must use coarser spatial grids than in the first scenario.

According to the problem tested here, it would appear that compact schemes (optimised and second-order accurate, or high-order) are generally better choices than leggy schemes, since they also perform well when the total operation costs of a scheme (and not just memory costs) are critical.[31]

errors are due to the non-exact initialisation (4.156).

[31] We should note that the Courant numbers chosen here for the optimised leggy schemes may not have been optimal for this particular problem. However, various other Courant numbers for these leggy schemes were tested and results did not significantly improve to the point of being comparable to the compact schemes (for the same spatiotemporal densities).



(a) reference solution (b) 5-pt scheme

(c) 9-pt leggy, $\lambda = \lambda_{\max,2}$ (d) 9-pt leggy, $\lambda = \lambda_{1\%,2}$ (e) 9-pt compact, $\alpha = 2/3$ (f) 9-pt compact, $\alpha = 1/2$

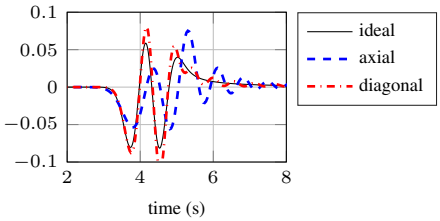
(g) 13-pt leggy, $\lambda = \lambda_{\max,3}$ (h) 13-pt leggy, $\lambda = \lambda_{1\%,3}$ (i) 13-pt compact, fourth-order (j) 13-pt compact, optimised

(k) 17-pt leggy, $\lambda = \lambda_{\max,4}$ (l) 17-pt leggy, $\lambda = \lambda_{1\%,4}$ (m) 17-pt compact, fourth-order (n) 17-pt compact, optimised

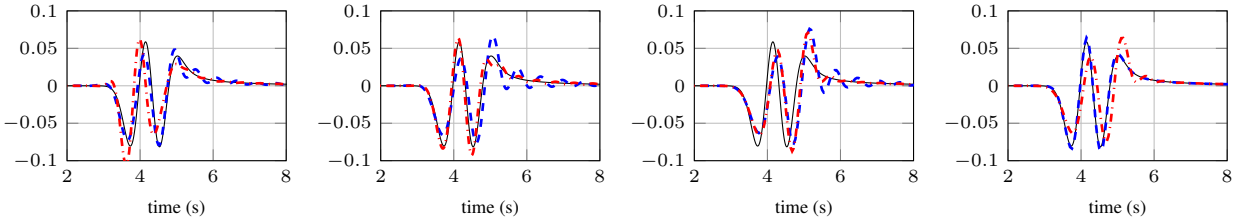
(o) 21-pt leggy, $\lambda = \lambda_{\max,5}$ (p) 21-pt leggy, $\lambda = \lambda_{1\%,5}$ (q) 21-pt compact, fourth-order (r) 21-pt compact, optimised

(s) 25-pt leggy, $\lambda = \lambda_{\max,6}$ (t) 25-pt leggy, $\lambda = \lambda_{1\%,6}$ (u) 25-pt compact, sixth-order (v) 25-pt compact, optimised

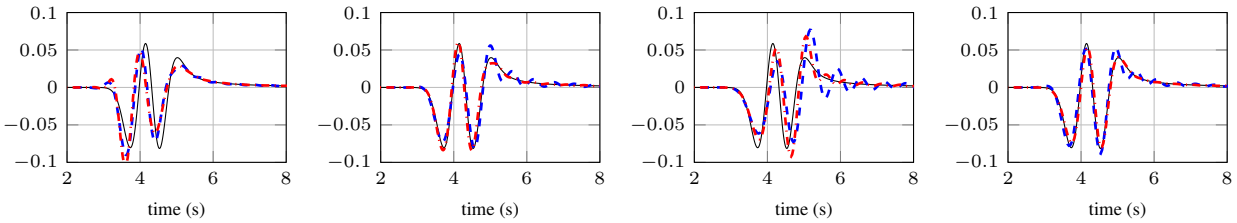
Figure 4.26: Snapshots of simulations from various schemes normalised for fixed memory costs. All plots use the same colour mapping, shown in (a). Unless otherwise specified, all schemes are second-order accurate (globally). Cubic spline interpolation is used for additional detail in the approximate solutions, and only a quarter-space region is shown (by symmetry).



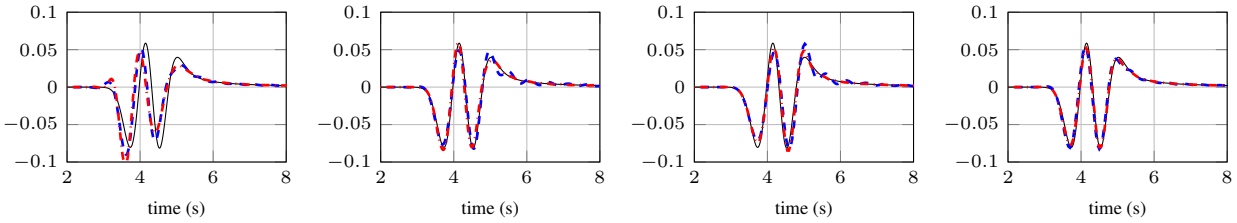
(a) 5-pt scheme



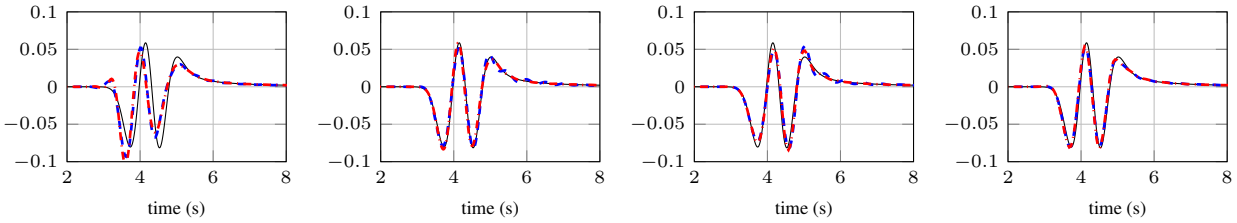
(b) 9-pt leggy, $\lambda = \lambda_{\max,2}$ (c) 9-pt leggy, $\lambda_{1\%,2}$ (d) 9-pt compact, $\alpha = 2/3$ (e) 9-pt compact, $\alpha = 1/2$



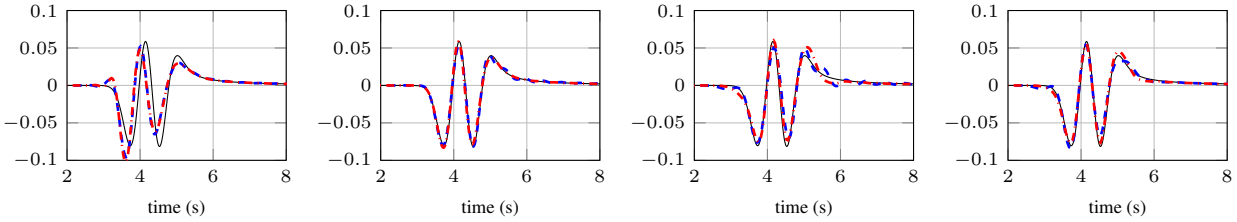
(f) 13-pt leggy, $\lambda = \lambda_{\max,3}$ (g) 13-pt leggy, $\lambda_{1\%,3}$ (h) 13-pt compact, fourth-order (i) 13-pt compact, optimised



(j) 17-pt leggy, $\lambda = \lambda_{\max,4}$ (k) 17-pt leggy, $\lambda_{1\%,M}$ (l) 17-pt compact, fourth-order (m) 17-pt compact, optimised

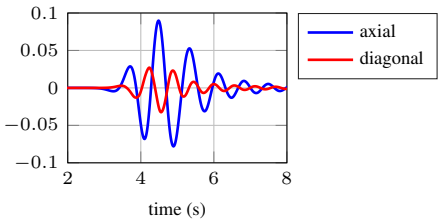


(n) 21-pt leggy, $\lambda = \lambda_{\max,5}$ (o) 21-pt leggy, $\lambda_{1\%,5}$ (p) 21-pt compact, fourth-order (q) 21-pt compact, optimised

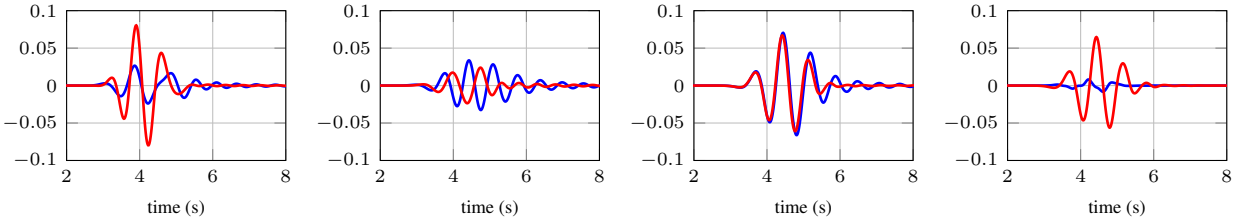


(r) 25-pt leggy, $\lambda = \lambda_{\max,6}$ (s) 25-pt leggy, $\lambda_{1\%,6}$ (t) 25-pt compact, sixth-order (u) 25-pt compact, optimised

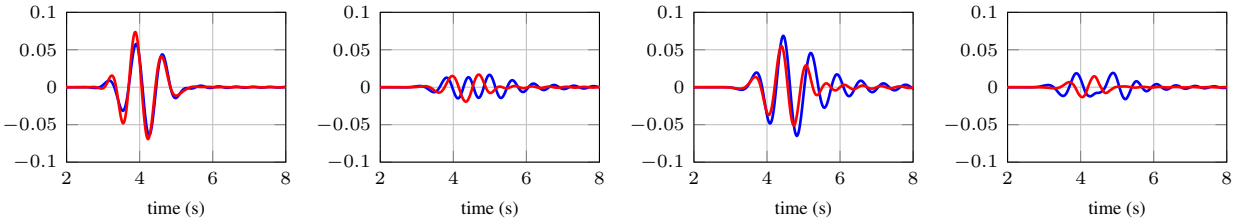
Figure 4.27: Plots of temporal outputs from various schemes, at distance $3\sqrt{2}$ m from origin, along axial direction and along diagonal direction. Here, the various schemes have the same memory costs. Cubic spline interpolation is used for additional detail in the approximate solutions.



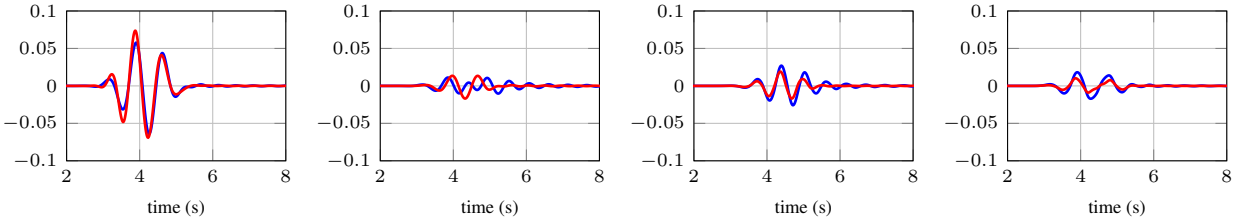
(a) 5-pt scheme



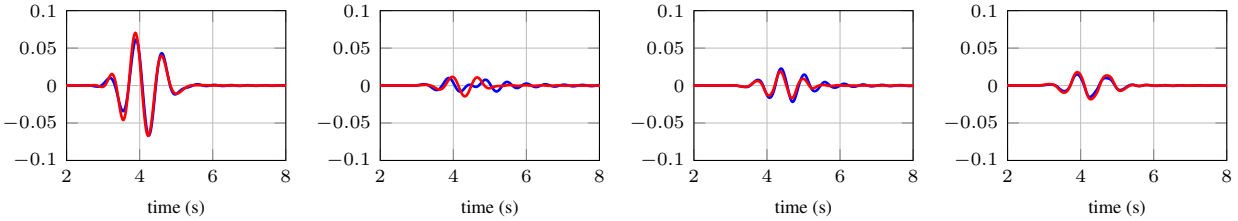
(b) 9-pt leggy, $\lambda = \lambda_{\max,2}$ (c) 9-pt leggy, $\lambda_{1\%,2}$ (d) 9-pt compact, $\alpha = 2/3$ (e) 9-pt compact, $\alpha = 1/2$



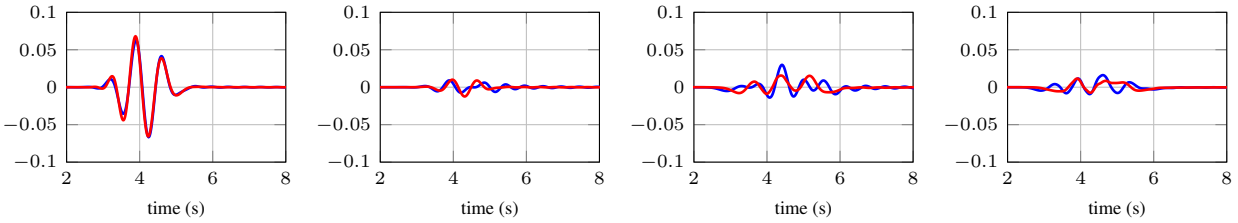
(f) 13-pt leggy, $\lambda = \lambda_{\max,3}$ (g) 13-pt leggy, $\lambda_{1\%,3}$ (h) 13-pt compact, fourth-order (i) 13-pt compact, optimised



(j) 17-pt leggy, $\lambda = \lambda_{\max,4}$ (k) 17-pt leggy, $\lambda_{1\%,4}$ (l) 17-pt compact, fourth-order (m) 17-pt compact, optimised

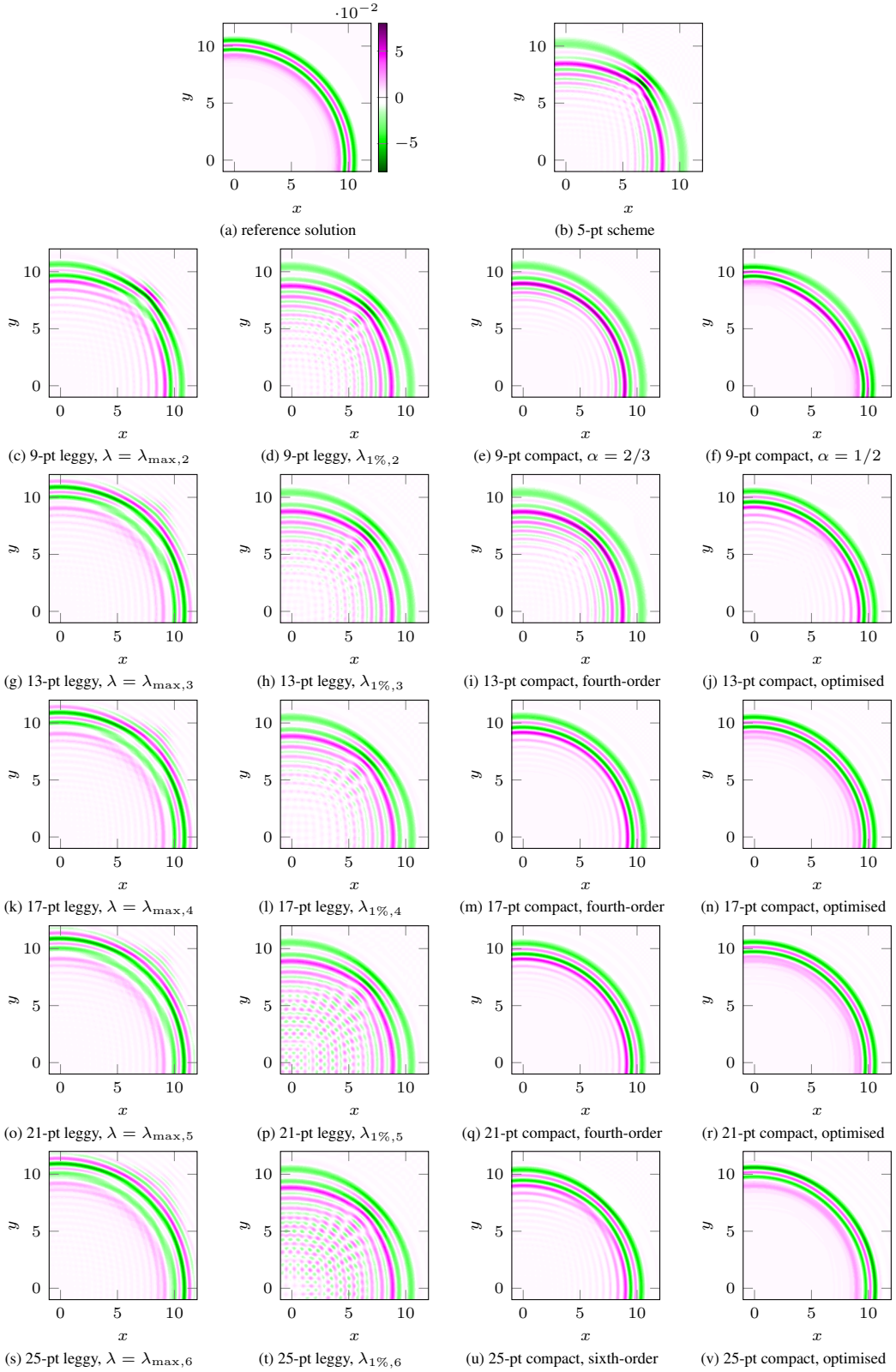


(n) 21-pt leggy, $\lambda = \lambda_{\max,5}$ (o) 21-pt leggy, $\lambda_{1\%,5}$ (p) 21-pt compact, fourth-order (q) 21-pt compact, optimised



(r) 25-pt leggy, $\lambda = \lambda_{\max,6}$ (s) 25-pt leggy, $\lambda_{1\%,6}$ (t) 25-pt compact, sixth-order (u) 25-pt compact, optimised

Figure 4.28: Plots of error signals from Fig. 4.27 (fixed memory costs).



(a) reference solution (b) 5-pt scheme

(c) 9-pt leggy, $\lambda = \lambda_{\max,2}$ (d) 9-pt leggy, $\lambda_{1\%,2}$ (e) 9-pt compact, $\alpha = 2/3$ (f) 9-pt compact, $\alpha = 1/2$

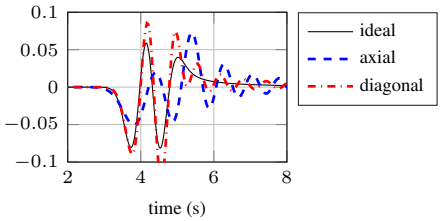
(g) 13-pt leggy, $\lambda = \lambda_{\max,3}$ (h) 13-pt leggy, $\lambda_{1\%,3}$ (i) 13-pt compact, fourth-order (j) 13-pt compact, optimised

(k) 17-pt leggy, $\lambda = \lambda_{\max,4}$ (l) 17-pt leggy, $\lambda_{1\%,4}$ (m) 17-pt compact, fourth-order (n) 17-pt compact, optimised

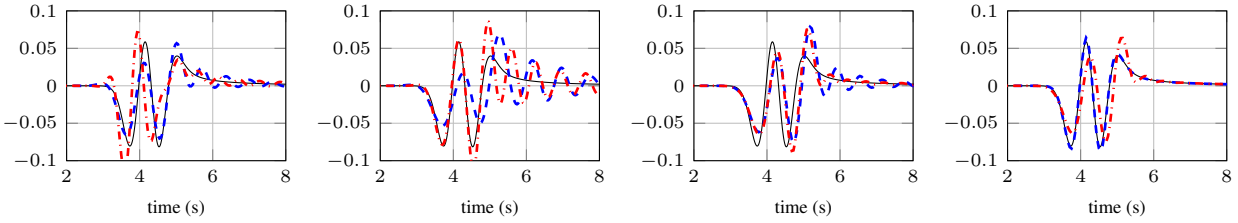
(o) 21-pt leggy, $\lambda = \lambda_{\max,5}$ (p) 21-pt leggy, $\lambda_{1\%,5}$ (q) 21-pt compact, fourth-order (r) 21-pt compact, optimised

(s) 25-pt leggy, $\lambda = \lambda_{\max,6}$ (t) 25-pt leggy, $\lambda_{1\%,6}$ (u) 25-pt compact, sixth-order (v) 25-pt compact, optimised

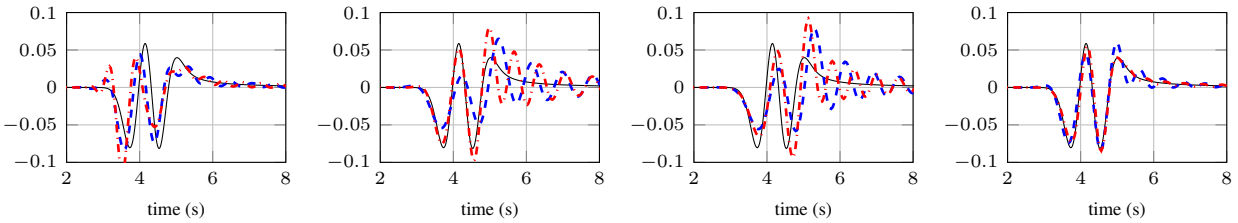
Figure 4.29: Snapshots of simulations from various schemes normalised for spatiotemporal density. All plots use the same colour mapping, shown in (a). Unless otherwise specified, all schemes are second-order accurate (globally). Cubic spline interpolation is used for additional detail in the approximate solutions, and only a quarter-space region is shown (by symmetry).



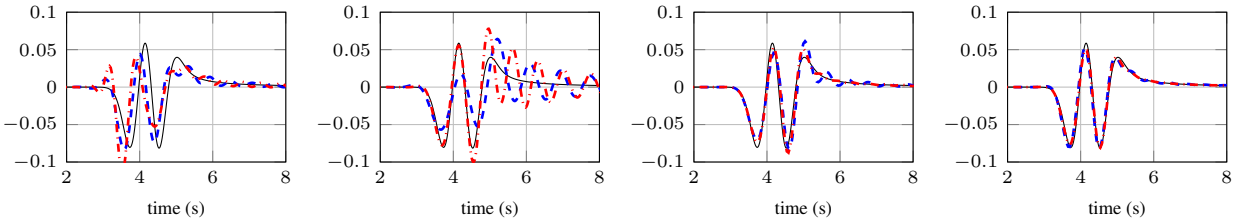
(a) 5-pt scheme



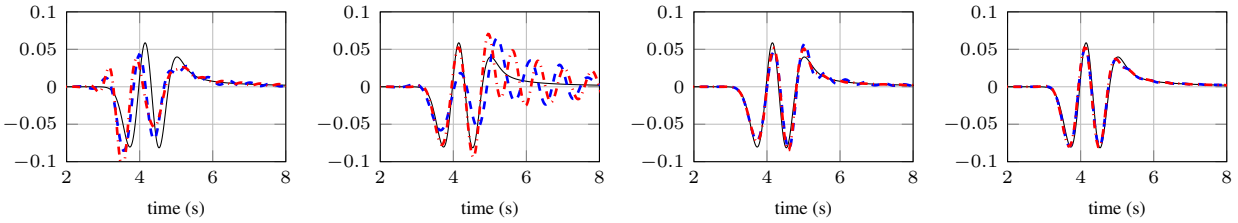
(b) 9-pt leggy, $\lambda = \lambda_{\max,2}$ (c) 9-pt leggy, $\lambda_{1\%,2}$ (d) 9-pt compact, $\alpha = 2/3$ (e) 9-pt compact, $\alpha = 1/2$



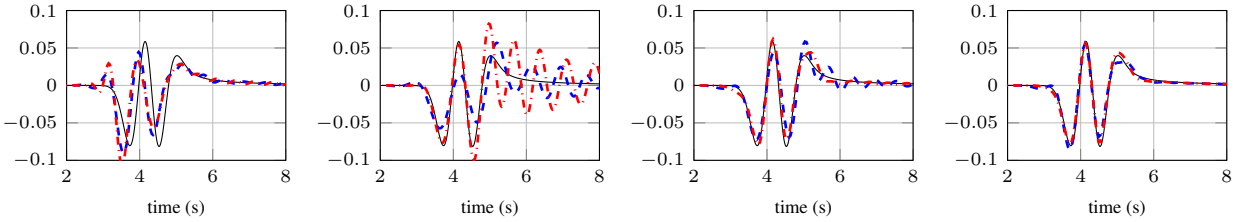
(f) 13-pt leggy, $\lambda = \lambda_{\max,3}$ (g) 13-pt leggy, $\lambda_{1\%,3}$ (h) 13-pt compact, fourth-order (i) 13-pt compact, optimised



(j) 17-pt leggy, $\lambda = \lambda_{\max,4}$ (k) 17-pt leggy, $\lambda_{1\%,4}$ (l) 17-pt compact, fourth-order (m) 17-pt compact, optimised

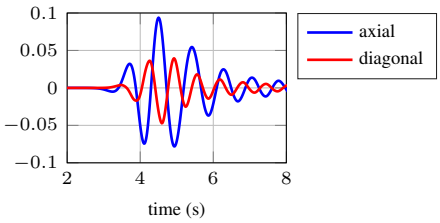


(n) 21-pt leggy, $\lambda = \lambda_{\max,5}$ (o) 21-pt leggy, $\lambda_{1\%,5}$ (p) 21-pt compact, fourth-order (q) 21-pt compact, optimised

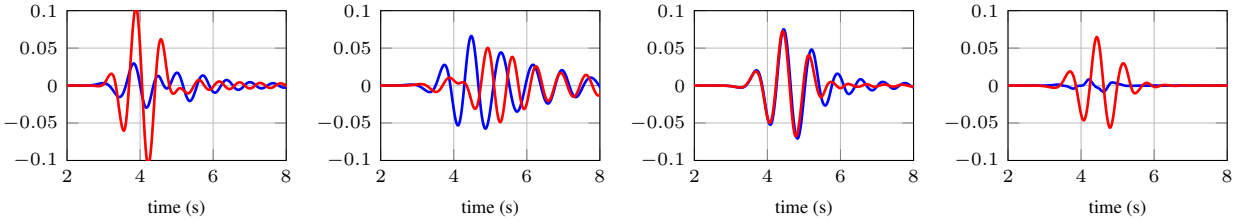


(r) 25-pt leggy, $\lambda = \lambda_{\max,6}$ (s) 25-pt leggy, $\lambda_{1\%,6}$ (t) 25-pt compact, sixth-order (u) 25-pt compact, optimised

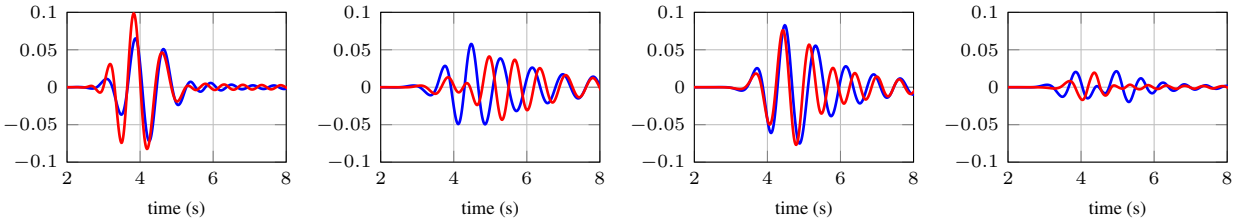
Figure 4.30: Plots of temporal outputs from various schemes, at distance $3\sqrt{2}$ m from origin, along axial direction and along diagonal direction. Here, the various schemes are normalised for spatiotemporal densities. Cubic spline interpolation is used for additional detail in the approximate solutions.



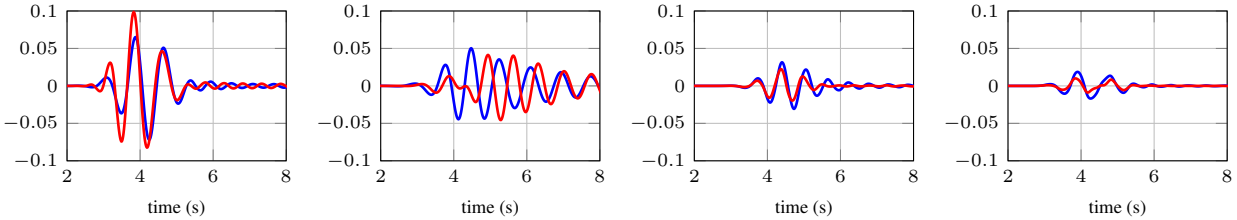
(a) 5-pt scheme



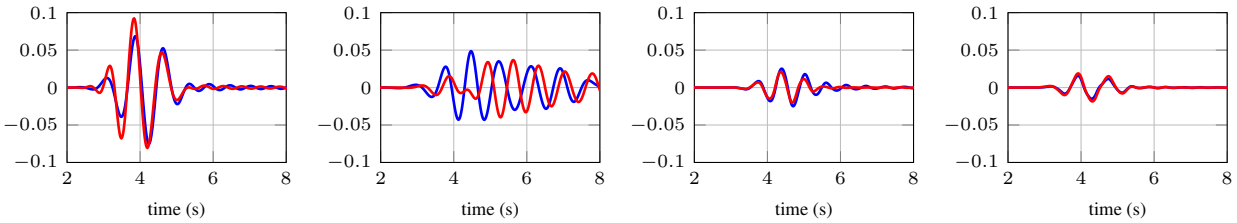
(b) 9-pt leggy, $\lambda = \lambda_{\max,2}$ (c) 9-pt leggy, $\lambda_{1\%,2}$ (d) 9-pt compact, $\alpha = 2/3$ (e) 9-pt compact, $\alpha = 1/2$



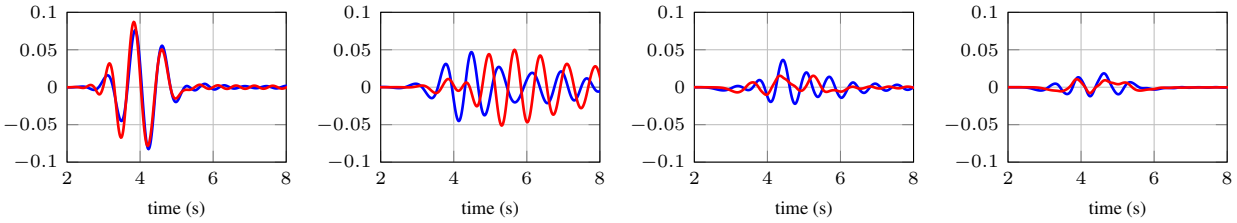
(f) 13-pt leggy, $\lambda = \lambda_{\max,3}$ (g) 13-pt leggy, $\lambda_{1\%,3}$ (h) 13-pt compact, fourth-order (i) 13-pt compact, optimised



(j) 17-pt leggy, $\lambda = \lambda_{\max,4}$ (k) 17-pt leggy, $\lambda_{1\%,4}$ (l) 17-pt compact, fourth-order (m) 17-pt compact, optimised



(n) 21-pt leggy, $\lambda = \lambda_{\max,5}$ (o) 21-pt leggy, $\lambda_{1\%,5}$ (p) 21-pt compact, fourth-order (q) 21-pt compact, optimised



(r) 25-pt leggy, $\lambda = \lambda_{\max,6}$ (s) 25-pt leggy, $\lambda_{1\%,6}$ (t) 25-pt compact, sixth-order (u) 25-pt compact, optimised

Figure 4.31: Plots of error signals from Fig. 4.30 (normalised for spatiotemporal densities).

4.4 Hexagonal schemes for the 2-D wave equation

In this section we investigate the use of regular hexagonal (non-Cartesian) grids for wave equation finite difference schemes. Hexagonal grids have been considered for finite difference methods for some time [121, 206, 56, 309], and the simplest hexagonal scheme is well-known [280, 80, 76, 21, 97]. High-order schemes on hexagonal grids have only been considered briefly (in [280]), so we will treat them in more detail here.

4.4.1 The hexagonal grid and sampling considerations

The hexagonal grid is defined here as:

$$\mathbb{G}_h^{(\mathrm{H})} = \{h\mathbf{V}_{\mathrm{H}}\boldsymbol{m} \in \mathbb{R}^2, \boldsymbol{m} \in \mathbb{Z}^2\}\,, \quad \mathbf{V}_{\mathrm{H}} = \begin{bmatrix} 1 & -1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix} \tag{4.157}$$

Under this construction, the hexagonal grid can be viewed as a skewed version of the Cartesian grid, and as such can be indexed by a column-vector $\boldsymbol{m} \in \mathbb{Z}^2$. This indexing is illustrated in Fig. 4.32.[32] $\mathbf{V}_{\mathrm{H}}$ is also known as a *generator matrix* for this lattice [58].

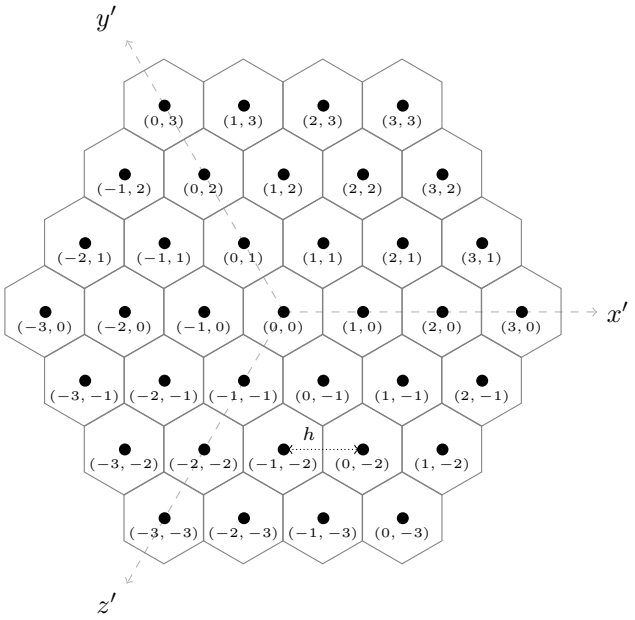


Figure 4.32: Hexagonal grid, indexed by $\boldsymbol{m} \in \mathbb{Z}^2$ according to (4.157)

In the figure we have also illustrated three axes, denoted x', y', z'. These axes are defined in relation to the standard coordinate axes (x and y) as:

$$x' = x\,, \qquad y' = -\frac{1}{2}x + \frac{\sqrt{3}}{2}y\,, \qquad z' = -\frac{1}{2}x - \frac{\sqrt{3}}{2}y \tag{4.158}$$

[32] The hexagonal grid is sometimes called the triangular lattice (for $h = 1$), since the grid points are defined at the vertices of a tiling of regular triangles. However, the name "hexagonal grid" comes from the fact that the Voronoi cell of the grid (the space occupied by each point on the grid) is a hexagon, as illustrated here, and from the fact that each point has six nearest neighbours, comprising the vertices of another hexagon.

These axes can also be described by the associated unit vectors:

$$\hat{\mathbf{e}}_{x'} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \hat{\mathbf{e}}_{y'} = \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix}, \quad \hat{\mathbf{e}}_{z'} = \begin{pmatrix} -1/2 \\ -\sqrt{3}/2 \end{pmatrix} = -(\hat{\mathbf{e}}_{x'} + \hat{\mathbf{e}}_{y'}) \tag{4.159}$$

While only two independent axes are needed in two spatial dimensions (e.g., x and y), the symmetry introduced by three axes (one redundant) is useful to navigate the hexagonal grid. Using these virtual axes we can also redefine the 2-D Laplacian operator as:

$$\Delta = \frac{2}{3}(\partial_{x'}^2 + \partial_{y'}^2 + \partial_{z'}^2) \tag{4.160}$$

Accordingly with the x'- and y'-axes, we can write the grid index-vector as $\boldsymbol{m} = (m_{x'}, m_{y'})^{\mathrm{T}}$, and it is worth noting that under this indexing system, we can determine the distance from the origin of any hexagonal point by the following:

$$\|h\mathbf{V}_{\mathrm{H}}\boldsymbol{m}\| = h\sqrt{m_{x'}^2 + m_{y'}^2 - m_{x'}m_{y'}} \tag{4.161}$$

Sampling considerations and spatial frequency domain analyses

In order to analyse hexagonal schemes in the wavenumber domain, we need to provide a short discussion about hexagonal sampling and how this leads to the appropriate wavenumber cell for the hexagonal grid, which in this case is not the usual Cartesian square cell [191, 168]. The wavenumber cell of the hexagonal grid is the Voronoi cell of its *reciprocal* or *dual* lattice, and the reciprocal of a lattice is generated by the inverse of its generator matrix, with a further scaled appropriately for the frequency quantity of interest. Here, the quantity of interest is $\boldsymbol{\beta}$ in rad/m, so we further scale by $2\pi/h$.

The dual of the hexagonal grid is defined as:

$$\widehat{\mathbb{G}_h^{(\mathrm{H})}} = \left\{(2\pi/h)\mathbf{V}_{\mathrm{H}}^{-1}\boldsymbol{m} \in \mathbb{R}^2 : \boldsymbol{m} \in \mathbb{Z}^2\right\} \tag{4.162}$$

and the Voronoi cell of this grid is another hexagon, with inradius $\frac{4\pi}{\sqrt{3}h}$; this will be illustrated shortly. We can formally define the *normalised* hexagonal wavenumber cell, which we denote as $\mathbb{B}_{\mathrm{H}}$, by the following system of linear inequalities:[33]

$$\mathbb{B}_{\mathrm{H}} = \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{A}_{\mathrm{H}}\mathbf{x}| \leq \pi\mathbf{b}_{\mathrm{H}}\} \tag{4.163}$$

where $\mathbf{x}$ is taken to be a column vector, and where:

$$\mathbf{A}_{\mathrm{H}} = \begin{bmatrix} 0 & 0 & -\sqrt{3} & -\sqrt{3} & \sqrt{3} & \sqrt{3} \\ 1 & -1 & -1 & 1 & 1 & -1 \end{bmatrix}^{\mathrm{T}}, \quad \mathbf{b}_{\mathrm{H}} = \frac{2}{\sqrt{3}}\begin{bmatrix} 1 & 1 & 2 & 2 & 2 & 2 \end{bmatrix}^{\mathrm{T}} \tag{4.164}$$

Let us briefly now consider the ideal dispersion relation in a hexagonal sampling scenario, analogous to the Cartesian sampling scheme illustrated in Fig. 4.1. The ideal (normalised) dispersion relation in the hexagonal wavenumber cell is illustrated in Fig. 4.33. It can be seen that the maximum wavenumber that can be isotropically sampled is $\beta_{h,\max} = \frac{2}{\sqrt{3}}\pi$, the inradius of the hexagonal wavenumber cell. In other words, the hexagonal grid can accommodate the isotropic discretisation of a higher density of spatial modes than the Cartesian grid for the same grid spacing h. On the other hand, the hexagonal grid has a

[33] Note that here, $\leq$ denotes a vector inequality (element-by-element).

density of $\mu = 2/\sqrt{3}$, approximately 15% higher than that of the Cartesian grid. Nevertheless, taking into account normalisations for grid densities, one arrives at the well-known result that the hexagonal grid is 15% more efficient for isotropic sampling than the Cartesian grid [191]. We also note that since $\beta_{h,\max} = \frac{2}{\sqrt{3}}$, the ideal Courant number for spatiotemporal sampling for an ideal solution to the 2-D wave equation is $\lambda = \sqrt{3}/2$ when using a hexagonal spatial grid.

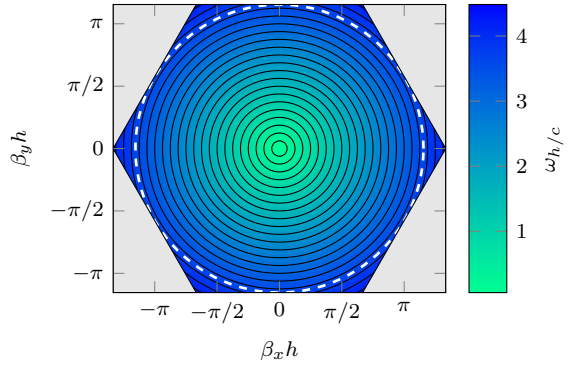


Figure 4.33: Contour plot of ideal dispersion relation $\omega_{h/c}(\boldsymbol{\beta}_h)$ for $\boldsymbol{\beta}_h \in \mathbb{B}_{\mathrm{H}}$. Contour lines denote increments of $\pi/16$. The white dotted line denotes $|\boldsymbol{\beta}_h| = \frac{2\pi}{\sqrt{3}}$.

Hexagonal grid functions and finite difference operators

For the hexagonal grid system we define a hexagonal grid function $\underline{u}^n_{\boldsymbol{m}} = \underline{u}_{m_{x'},m_{y'}}$ as $\underline{u}^n_{\boldsymbol{m}} \approxeq u(h\mathbf{V}_{\mathrm{H}}\boldsymbol{m}, nk)$. We will also use the continuous approximation $\underline{u}$, such that $\underline{u}^n_{\boldsymbol{m}} = \underline{u}(h\mathbf{V}_{\mathrm{H}}\boldsymbol{m}, nk)$, and such that $\underline{u}^n_{\boldsymbol{m}}$ can be seen as the (hexagonally) sampled version of $\underline{u}(\mathbf{x}, t)$.

In order to build finite difference operators on the hexagonal grid, we start by defining the following spatial shift operators:

$$s_{x'\pm}\underline{u}^n_{\boldsymbol{m}} = \underline{u}^n_{m_{x'}\pm 1, m_{y'}} \quad \Longrightarrow \quad s_{x'\pm}\underline{u} = \underline{u}(\mathbf{x} \pm h\hat{\mathbf{e}}_{x'}, t) \tag{4.165a}$$

$$s_{y'\pm}\underline{u}^n_{\boldsymbol{m}} = \underline{u}^n_{m_{x'}, m_{y'}\pm 1} \quad \Longrightarrow \quad s_{y'\pm}\underline{u} = \underline{u}(\mathbf{x} \pm h\hat{\mathbf{e}}_{y'}, t) \tag{4.165b}$$

$$s_{z'\pm} = s_{x'\mp}s_{y'\mp} \quad \Longrightarrow \quad s_{z'\pm}\underline{u} = \underline{u}(\mathbf{x} \pm h\hat{\mathbf{e}}_{z'}, t) \tag{4.165c}$$

Using these shift operators, we can define second-order difference operators:

$$\delta_{x'x'} = \frac{1}{h^2}(s_{x'+} - 2 + s_{x'-}) \quad \Longrightarrow \quad \delta_{x'x'}\underline{u} = \partial^2_{x'}\underline{u} + \mathcal{O}(h^2) \tag{4.166a}$$

$$\delta_{y'y'} = \frac{1}{h^2}(s_{y'+} - 2 + s_{y'-}) \quad \Longrightarrow \quad \delta_{y'y'}\underline{u} = \partial^2_{y'}\underline{u} + \mathcal{O}(h^2) \tag{4.166b}$$

$$\delta_{z'z'} = \frac{1}{h^2}(s_{z'+} - 2 + s_{z'-}) \quad \Longrightarrow \quad \delta_{z'z'}\underline{u} = \partial^2_{z'}\underline{u} + \mathcal{O}(h^2) \tag{4.166c}$$

It follows from (4.160) that the following is a discrete Laplacian on the hexagonal grid [121, 206]:

$$\delta^{(\mathrm{H})}_{\Delta} = \frac{2}{3}(\delta_{x'x'} + \delta_{y'y'} + \delta_{z'z'}) \quad \Longrightarrow \quad \delta^{(\mathrm{H})}_{\Delta}\underline{u} = \Delta\underline{u} + \mathcal{O}(h^2) \tag{4.167}$$

It also turns out that this seven-point discrete Laplacian is isotropic to fourth-order:

$$\delta^{(\mathrm{H})}_{\Delta}\underline{u} = \Delta\underline{u} + \frac{h^2}{16}\Delta^2\underline{u} + \mathcal{O}(h^4) \tag{4.168}$$

It is worth recalling that an isotropic discrete Laplacian on the Cartesian grid requires a stencil of nine points; this is a first indication that hexagonal grids can be advantageous to Cartesian grids.

4.4.2 Parametrised approximations to the Laplacian

Similar to the 2-D parametrised discrete Laplacians presented in Section 4.3.2, we can define parametrised approximations to the Laplacian on the hexagonal grid. First, let us define the averaging operators:

$$\mu_{(a,b),(x',y')} = \frac{1}{2}(s^a_{x'+}s^b_{y'-} + s^a_{x'-}s^b_{y'+}) \tag{4.169a}$$

$$\mu_{(a,b),(x',z')} = \frac{1}{2}(s^a_{x'+}s^b_{z'-} + s^a_{x'-}s^b_{z'+}) \tag{4.169b}$$

$$\mu_{(a,b),(y',z')} = \frac{1}{2}(s^a_{y'+}s^b_{z'-} + s^a_{y'-}s^b_{z'+}) \tag{4.169c}$$

where $a \geq b \geq 0$ and $a \geq 1$. Taylor expansions of these operators are, e.g.:

$$\mu_{(a,b),(x',y')}\underline{u} = \underline{u} + \frac{h^2}{2}\left(a^2\partial^2_{x'} + b^2\partial^2_{y'} - 2ab\partial_{x'}\partial_{y'}\right)\underline{u} + \mathcal{O}(h^4) \tag{4.170}$$

It is straightforward through such Taylor expansions to arrive at:

$$\begin{aligned}(\mu_{(a,b),(x',y')} + \mu_{(b,a),(x',y')} + \mu_{(a,b),(x',z')} + \mu_{(b,a),(x',z')} + \mu_{(a,b),(y',z')} + \mu_{(b,a),(y',z')})\underline{u}& \\ = 6\underline{u} + \frac{3h^2}{2}(a^2 + b^2 + ab)\Delta\underline{u} + \mathcal{O}(h^4)&\end{aligned} \tag{4.171}$$

Then for $\boldsymbol{q} \in \mathbb{Q}$, with $\mathbb{Q}$ defined in (4.102), we can define the following discrete Laplacian:

$$\begin{aligned}\delta^{(\mathrm{H})}_{\Delta,\boldsymbol{q}} = \frac{2}{3(q_1^2 + q_2^2 + q_1q_2)h^2}(&\mu_{(q_1,q_2),(x',y')} + \mu_{(q_2,q_1),(x',y')} + \mu_{(q_1,q_2),(x',z')} \\ &+ \mu_{(q_2,q_1),(x',z')} + \mu_{(q_1,q_2),(y',z')} + \mu_{(q_2,q_1),(y',z')} - 6)\end{aligned} \tag{4.172}$$

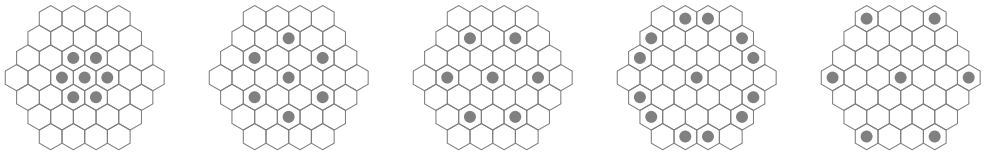
and we have

$$\delta^{(\mathrm{H})}_{\Delta,\boldsymbol{q}}\underline{u} = \Delta\underline{u} + \mathcal{O}(h^2) \tag{4.173}$$

It is worth noting that each such discrete Laplacian is isotropic to fourth-order:

$$\delta^{(\mathrm{H})}_{\Delta,\boldsymbol{q}}\underline{u} = \Delta\underline{u} + Ch^2\Delta^2\underline{u} + \mathcal{O}(h^4) \tag{4.174}$$

where C is some constant. This discrete Laplacian uses a $(K+1)$-point stencil, where $K = 6$ when $q_2 = 0$ or $q_1 = q_2$, and $K = 12$ otherwise. Some examples of the underlying hexagonal stencils are shown in Fig. 4.34.



(a) $(q_1, q_2) = (1, 0)$ (b) $(q_1, q_2) = (1, 1)$ (c) $(q_1, q_2) = (2, 0)$ (d) $(q_1, q_2) = (2, 1)$ (e) $(q_1, q_2) = (3, 0)$

Figure 4.34: Stencils of points for operators $\delta^{(\mathrm{H})}_{\Delta,(q_1,q_2)}$ with (q_1, q_2) as indicated.

For convenience, we can also rewrite this operator as:

$$\delta^{(\mathrm{H})}_{\Delta,\boldsymbol{q}}\underline{u}^n_{\boldsymbol{m}} = \frac{4/K}{(q_1^2+q_2^2+q_1q_2)h^2}(Q^n_{\boldsymbol{m},\boldsymbol{q}} - K\underline{u}^n_{\boldsymbol{m}}) \tag{4.175}$$

where $Q^n_{\boldsymbol{m},\boldsymbol{q}} = Q^n_{\boldsymbol{m},(q_1,q_2)}$ represents a sum over the neighbouring K grid points, which is slightly harder to define than the analogous summation on a Cartesian grid (i.e., (4.107)). In order to construct $Q^n_{\boldsymbol{m},\boldsymbol{q}}$, we let $s^{-a}_{x'+} = s^a_{x'-}$, and similarly for the shift operators along y'- and z'-axes. Then we can define this sum as:

$$Q^n_{\boldsymbol{m},\boldsymbol{q}} = \sum_{\boldsymbol{q}'\in\mathcal{P}(\boldsymbol{q})} s^{q_1'}_{x'+}s^{q_2'}_{y'+}s^{q_3'}_{z'+}\underline{u}^n_{m_{x'},m_{y'}} \tag{4.176}$$

where here $\boldsymbol{q}' = (q_1', q_2', q_3')$ and $\mathcal{P}(\boldsymbol{q})$ is now the set of all *unique* permutations of $\{q_1, -q_2, 0\}$ and $\{-q_1, q_2, 0\}$, and $K = |\mathcal{P}(\boldsymbol{q})|$.

Let us now consider a set of P distinct integer pairs $\boldsymbol{q} = (q_1, q_2) \in \mathbb{Q}$ denoted as $\Upsilon = \{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_P\}$. We can then construct a general family of parametrised discrete Laplacians as:

$$\delta^{(\mathrm{H})}_{\Delta,\Upsilon} = \sum_{(q_1,q_2)\in\Upsilon} \alpha_{(q_1,q_2)}\delta^{(\mathrm{H})}_{\Delta,(q_1,q_2)} \tag{4.177}$$

where $\alpha_{(q_1,q_2)}$ is a real-valued parameter, and for consistency we require that:

$$\sum_{(q_1,q_2)\in\Upsilon} \alpha_{(q_1,q_2)} = 1 \tag{4.178}$$

which will be implicitly assumed.

A general family of two-step explicit hexagonal schemes is then:

$$\delta_{tt}\underline{u}^n_{\boldsymbol{m}} = c^2\delta^{(\mathrm{H})}_{\Delta,\Upsilon}\underline{u}^n_{\boldsymbol{m}} \tag{4.179}$$

which is subject to the numerical stability constraints (4.28), where $G(\boldsymbol{\beta}_h)$ takes on a form given below:

$$G(\boldsymbol{\beta}_h) = \sum_{(q_1,q_2)\in\Upsilon} \frac{\alpha_{(q_1,q_2)}}{6(q_1^2+q_2^2+q_1q_2)}\Big(6 - c_{x,q_1}c_{y',q_2} - c_{x',q_2}c_{y',q_1} - c_{x',q_1}c_{z',q_2} \\ - c_{x',q_2}c_{z',q_1} - c_{y',q_1}c_{z',q_2} - c_{y',q_2}c_{z',q_1}\Big) \tag{4.180}$$

Here, $c_{w,b} = \cos(b\boldsymbol{\beta}_h \cdot \hat{\mathbf{e}}_w)$, and $w \in \{x', y', z'\}$ and $b \in \{q_1, q_2\}$. In many cases, $\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)$ will be found at the vertices or edge-centers of $\mathbb{B}_{\mathrm{H}}$, but not always. Also, it is important to check that $G(\boldsymbol{\beta}_h) \geq 0$ for a given discrete Laplacian, such that $\delta^{(\mathrm{H})}_{\Delta,\Upsilon}$ remains negative semi-definite.

A family of $(K+1)$-point compact discrete Laplacians with $K \leq 12$

Analogous to the compact nine-point discrete Laplacian, we have the following 13-point compact discrete Laplacian using the first two shells of points on the hexagonal grid:

$$\delta^{(\mathrm{H})}_{\Delta,\Upsilon_{13}} = \alpha_{(1,0)}\delta^{(\mathrm{H})}_{\Delta,(1,0)} + \alpha_{(1,1)}\delta^{(\mathrm{H})}_{\Delta,(1,1)} \tag{4.181}$$

where $\Upsilon_{13} = \{(1,0),(1,1)\}$. For convenience, let us redefine this discrete Laplacian operator as:

$$\delta^{(\mathrm{H})}_{\Delta,\Upsilon_{13}} = \alpha\delta^{(\mathrm{H})}_{\Delta,(1,0)} + (1-\alpha)\delta^{(\mathrm{H})}_{\Delta,(1,1)} \tag{4.182}$$

The stencil of points for this discrete Laplacian uses either seven or thirteen points, depending on the choice of α, as illustrated in Fig. 4.35.

(a) $\alpha = 1$ (b) $\alpha = 0$ (c) $\alpha \notin \{0, 1\}$

Figure 4.35: Stencils of points for operators $\delta^{(\mathrm{H})}_{\Delta,\Upsilon_{13}}$ with α as indicated.

In order to identify some special cases of interest, we can expand this operator about $h = 0$:

$$\delta^{(\mathrm{H})}_{\Delta,\Upsilon_{13}}\underline{u} = \Delta\underline{u} + \frac{h^2}{16}(2\alpha - 3)\Delta^2\underline{u} + \mathcal{O}(h^4) \tag{4.183}$$

and for $\alpha = 3/2$ we obtain a fourth-order accuracy in space [309, 56]. Upon investigation of the $\mathcal{O}(h^4)$ term in the above expansion (omitted for brevity), we arrive at another interesting choice: $\alpha = 9/10$, which gives an isotropic error to sixth-order (in h):

$$\delta^{(\mathrm{H})}_{\Delta,\Upsilon_{13}}\underline{u} = \Delta\underline{u} + \frac{3h^2}{40}\Delta^2\underline{u} + \frac{h^4}{320}\Delta^3\underline{u} + \mathcal{O}(h^6) \tag{4.184}$$

Finally, we note that $\alpha \geq 0$ is the necessary condition for this operator to remain negative semi-definite.

4.4.3 The simple seven-point hexagonal scheme

Our investigation of hexagonal schemes will naturally start with the simplest, seven-point scheme:

$$\delta_{tt}\underline{u}^n_{\boldsymbol{m}} = c^2\delta^{(\mathrm{H})}_{\Delta,(1,0)}\underline{u}^n_{\boldsymbol{m}} \tag{4.185}$$

When $\lambda = \sqrt{1/2}$ this scheme which is equivalent to the "triangular digital waveguide mesh" [80]. However, it is well-known by now that the necessary and sufficient von Neumann stability for this scheme is [280, 76, 21]:

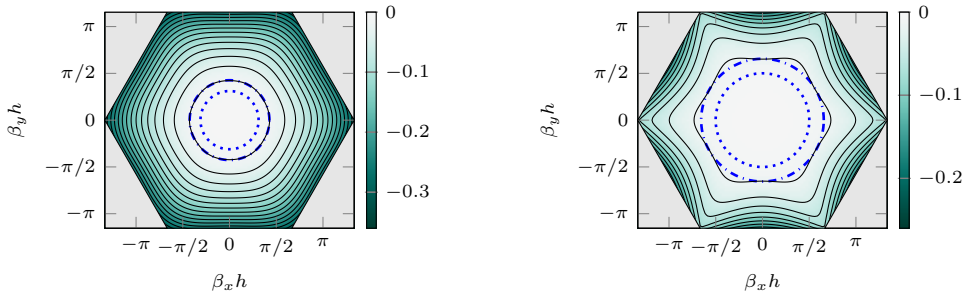
$$\lambda \leq \sqrt{2/3} \tag{4.186}$$

where $\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)$ is found at the vertices of $\mathbb{B}_{\mathrm{H}}$.

Let us now analyse the numerical dispersion of this scheme. Considering the appropriate wavenumber cell $\mathbb{B}_{\mathrm{H}}$, the relative dispersion error of this scheme is shown in Fig. 4.36 for two choices of the Courant number: $\lambda = \sqrt{1/2}$ and $\lambda = \sqrt{2/3}$.[34] As can be seen in the figure, and as has been demonstrated in the past [21], the hexagonal scheme with $\lambda = \lambda_{\max} = \sqrt{2/3}$ provides better numerical dispersion than the choice $\lambda = \sqrt{1/2}$ which is employed in the triangular DWM [80].

Due to the symmetry in these dispersion relations, it will be sufficient to consider only the positive quadrant of $\mathbb{B}_{\mathrm{H}}$; this will also serve to provide more detail. The relative dispersion errors of these schemes are shown again in Fig. 4.37. At this point we can compare the hexagonal scheme in Fig. 4.37(b) to the nine-point "IWB" Cartesian scheme in Fig. 4.14(c). One can observe that the hexagonal scheme has one- and two-percent critical wavenumbers that extend farther than those for the IWB scheme, suggesting that the seven-point scheme may provide better dispersion for less computation.

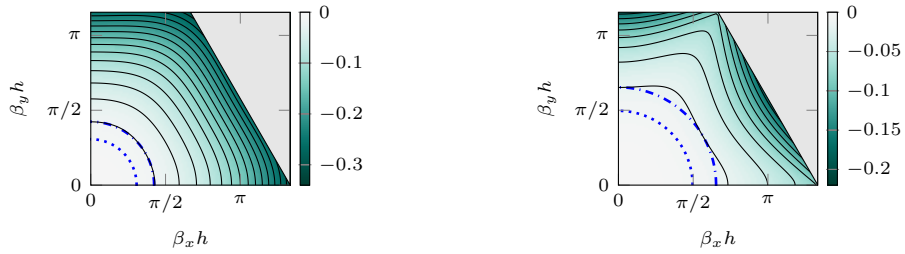
[34] It is worth noting that up until recently ([97], work by the current author), this scheme had not been analysed in the appropriate (hexagonal) wavenumber cell (e.g., [80, 81, 21, 92]). However, hexagonal schemes for other systems have been analysed in the appropriate wavenumber cell for some time (e.g., [183]).



(a) $\lambda = \sqrt{1/2}$ (b) $\lambda = \sqrt{2/3}$

Figure 4.36: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for the simplest hexagonal scheme with $\lambda = \sqrt{1/2}$, for $\boldsymbol{\beta}_h \in \mathbb{B}_{\mathrm{H}}$. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

However, we must remember that the grid density is 15% higher for the hexagonal grid (with the same h), and for $\lambda = 1$, the IWB scheme uses a larger time-step for the same h. Nevertheless, taking spatiotemporal normalisations into account, it has been shown that the *relative computational efficiency* of this scheme is more than three times that of the IWB scheme for less than one-percent dispersion error [97] (by the current author, left out for brevity).



(a) $\lambda = \sqrt{1/2}$ (b) $\lambda = \sqrt{2/3}$

Figure 4.37: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for the simplest hexagonal scheme with $\lambda = \sqrt{1/2}$, for $\boldsymbol{\beta}_h$ in the positive quadrant of $\mathbb{B}_{\mathrm{H}}$. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

4.4.4 A family of $(K+1)$-point compact schemes with $K \leq 12$

At this point we consider a more general, and previously unexplored, family of schemes:

$$\delta_{tt}\underline{u}^n_{\boldsymbol{m}} = c^2 \delta^{(\mathrm{H})}_{\Delta,\Upsilon_{13}} \underline{u}^n_{\boldsymbol{m}} \tag{4.187}$$

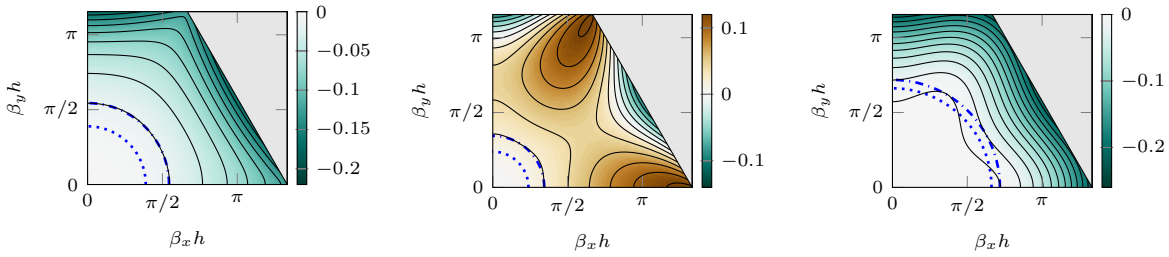
This family has one free parameter in the discrete Laplacian, α, and as mentioned previously, we need that $\alpha \geq 0$ for $G(\boldsymbol{\beta}_h) \geq 0$. Using a von Neumann stability analysis (in this case, evaluated with the help of Mathematica [308]), one finds that $\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)$ appears at the edge-centers of $\mathbb{B}_{\mathrm{H}}$ for $1/4 \leq \alpha \leq 8/11$ and at the vertices of $\mathbb{B}_{\mathrm{H}}$ for $\alpha \geq 8/11$. As such, we obtain the following stability

conditions on the Courant number for $\alpha \geq 1/4$:

$$\lambda \leq \begin{cases} \sqrt{\dfrac{9}{4+8\alpha}}\,, & 1/4 \leq \alpha \leq 8/11 \\ \sqrt{\dfrac{2}{3\alpha}}\,, & \alpha > 8/11 \end{cases} \tag{4.188}$$

The behaviour of $\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)$ is not simple to describe for $0 \leq \alpha \leq 1/4$, so it is omitted for brevity.[35]

Let us consider the special cases: $\alpha \in \{9/10, 3/2\}$, pertaining to, respectively, a sixth-order isotropic scheme and $(2,4)$-accurate scheme.[36] The $(2,4)$-accurate scheme is essentially a "leggy" scheme on the hexagonal grid. The relative dispersion errors for these schemes are illustrated in Fig. 4.38. We have displayed the $(2,4)$-accurate hexagonal leggy scheme with two choices of the Courant number: $\lambda = \lambda_{\max}$; and an optimised λ, chosen to balance spatial errors with temporal errors (i.e., $\lambda < \lambda_{\max}$), which, as can be seen, provides improved dispersion. As expected, the scheme with $\alpha = 9/10$ (in Fig. 4.38(a)) demonstrates a higher degree of isotropy than the other two illustrated schemes.



(a) $\alpha = 9/10$, $\lambda = \sqrt{20/27}$ (b) $\alpha = 3/2$, $\lambda = \lambda_{\max} = 2/3$ (c) $\alpha = 3/2$, $\lambda = 0.67\lambda_{\max} \approx 0.45$

Figure 4.38: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for compact hexagonal schemes with α and λ as indicated. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

A fourth-order accurate scheme and an optimised scheme

Apart from a 19-point fourth-order scheme on the hexagonal grid that has been presented in [280], hexagonal schemes with high order accuracy in both space and time, derived using modified equation methods have yet to be explored (to the knowledge of the current author).

We will present some high-order schemes on hexagonal grids, starting with one that can be found from this 13-point family. First, we write this scheme as a discrete d'Alembertian:

$$\delta_{\Box} = \delta_{tt} - c^2 \delta^{(\mathrm{H})}_{\Delta,\Upsilon_{13}} \tag{4.189}$$

and we have the Taylor expansion:

$$\delta_{\Box}\underline{u} = \Box\underline{u} + \frac{h^2}{48c^2}\left(4\lambda^2\partial_t^4 - 3(3-2\alpha)c^4\Delta^2\right)\underline{u} + O(h^4) = 0 \tag{4.190}$$

[35] It also does not produce interesting special cases.

[36] It is worth mentioning that the choice $\alpha = 8/9$ provides the "optimal" Courant number $\lambda_{\max} = \sqrt{3}/2$ (in terms of spatiotemporal sampling considerations), so it is another special case of interest, however it is only slightly different from the sixth-order isotropic scheme, so we omit it for brevity.

Thus, for fourth-order accuracy in both space and time, we require that:

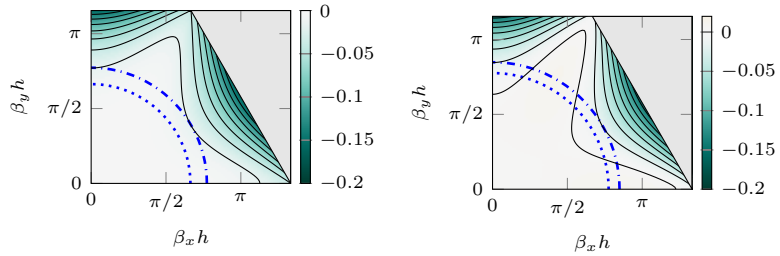
$$\lambda^2 = \frac{3(3-2\alpha)}{4} \tag{4.191}$$

It can be noticed that for a fourth-order accuracy with the seven-point scheme ($\alpha = 1$), the Courant number would have to be $\lambda = \sqrt{3/4}$ which would be unstable (as mentioned in [280]). A fourth-order scheme is thus only permitted with $\alpha > 1$. More specifically, the conditions for fourth-order accuracy are:

$$\frac{9+\sqrt{17}}{12} \leq \alpha < \frac{3}{2}\,, \quad \lambda = \sqrt{\frac{9-6\alpha}{4}} \tag{4.192}$$

The optimal choice under these constraints is $\alpha = \frac{9+\sqrt{17}}{12} \approx 1.094$, as this permits the highest possible Courant number, $\lambda = \sqrt{\frac{8}{9+\sqrt{17}}} \approx 0.78$.

The relative phase velocity for the fourth-order accurate scheme with $\lambda = \lambda_{\max}$ is shown in Fig. 4.39(a). Alongside this fourth-order scheme, we present a *second-order accurate* 13-point scheme that has been optimised for a 1% error criterion (in Fig. 4.39(b)). This optimised scheme has a value of α close to that of the fourth-order scheme, but the optimised scheme achieves a higher critical wavenumber $\beta_{h,1\%}$, as seen in the figures below. These figures could also be compared to the analogous Cartesian 13-point schemes in Figs 4.19(a) and 4.21(a).



(a) fourth-order, $\alpha = \frac{9+\sqrt{17}}{12} \approx 1.094$ (b) optimised, $\alpha = 1.165$

Figure 4.39: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for fourth-order accurate and optimised (for $\beta_{h,1\%}$) 13-point hexagonal schemes with α as indicated and $\lambda = \sqrt{2/(3\alpha)}$. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

4.4.5 High-order $(K+1)$-point schemes with $K \leq 24$

Going further in the general family of hexagonal schemes, we can consider "wide-stencil" hexagonal schemes, analogous to the 2-D Cartesian schemes considered in Section 4.3.5.

We consider the following family of schemes using stencils reaching out to as many as 24 neighbouring points:

$$\delta_{\square} = \delta_{tt} - c^2 \delta^{(\mathrm{H})}_{\Delta,\Upsilon_{24}} \tag{4.193}$$

where $\Upsilon_{24} = \{(1,0),(1,1),(2,0),(3,0)\}$. While there are many possibilities within this family (having three free parameters), we will only consider two special cases: a fourth-order accurate, sixth-order isotropic 19-point scheme (with $\alpha_{(3,0)} = 0$), and a sixth-order accurate 25-point scheme. The underlying 19- and 25-point spatial stencils are illustrated in Fig. 4.40.

To start the derivation of these high-order schemes, let us denote $\boldsymbol{\alpha} = (\alpha_{(1,0)}, \alpha_{(1,1)}, \alpha_{(2,0)}, \alpha_{(3,0)})$ for this discrete Laplacian $\delta^{(\mathrm{H})}_{\Delta,\Upsilon_{24}}$. We can expand this discrete Laplacian in a Taylor series in order to

(a) 19-point hexagonal stencil (b) 25-point hexagonal stencil

Figure 4.40: Hexagonal stencils of points involved in $\delta^{(\mathrm{H})}_{\Delta,\Upsilon_{24}}$

find the following condition for sixth-order isotropy:

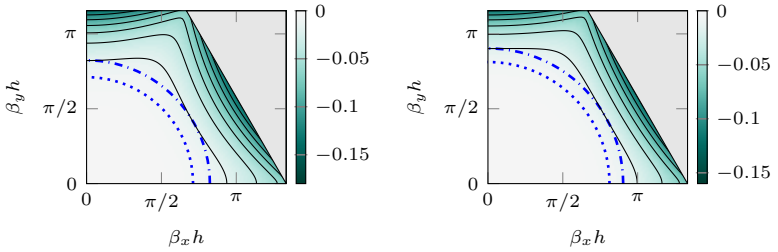
$$\boldsymbol{\alpha} \cdot (1, -9, 16, 81) = 0 \tag{4.194}$$

and clearly we need $\alpha_{(1,0)} + \alpha_{(1,1)} + \alpha_{(2,0)} + \alpha_{(3,0)} = 1$ for consistency. After imposing the sixth-order isotropy constraint and the consistency constraint, the approach to deriving fourth- and sixth-order accuracy is completely described by Eqs. (4.145)–(4.153). As such, with the 19-point stencil ($\alpha_{(3,0)} = 1$) we can possibly derive a fourth-order scheme, and a sixth-order scheme can be possibly derived using the 25-point stencil. It remains to check for stability, which can be accomplished from the expression for $G(\boldsymbol{\beta}_h)$ (Eq. (4.180)). In analysing $G(\boldsymbol{\beta}_h)$ under the above constraints, we find that $\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)$ for these schemes are again found at the vertices of the wavenumber cell, allowing us to obtain expressions for $\lambda_{\max}$, and subsequently for $\boldsymbol{\alpha}$.

The coefficients for these schemes with $\lambda = \lambda_{\max}$ are given in Table 4.7,[37] and the relative numerical phase velocities for these schemes are shown in Fig. 4.41. Numerical dispersion relations for these schemes are improved over the Cartesian counterparts with similar stencil sizes (as shown in Figs. 4.19(c) and 4.19(d)).

Table 4.7: Parameters for high-order schemes on hexagonal grid.

K	OoA	$\alpha_{(1,0)}$	$\alpha_{(1,1)}$	$\alpha_{(2,0)}$	$\alpha_{(3,0)}$	λ^2	λ (approx.)
12	$(4,4)$	$\frac{9+\sqrt{17}}{12}$	$\frac{3-\sqrt{17}}{12}$	—	—	$\frac{8}{9+\sqrt{17}}$	0.780
18	$(4,4)$	$\frac{351+5\sqrt{1761}}{540}$	$\frac{15-\sqrt{587/3}}{60}$	$\frac{27-\sqrt{1761}}{270}$	—	$\frac{81-\sqrt{1761}}{60}$	0.806
24	$(6,6)$	$\frac{16949+271\sqrt{1761}}{27000}$	$\frac{(\sqrt{1761}-81)^2}{40500}$	$\frac{638-23\sqrt{1761}}{3375}$	$\frac{21\sqrt{1761}-601}{27000}$	$\frac{81-\sqrt{1761}}{60}$	0.806



(a) 19-point, fourth-order accurate (b) 25-point, sixth-order accurate

Figure 4.41: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for fourth- and sixth-order accurate 19- and 25-point schemes. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

[37] Remarkably, the fourth-order 19-point and sixth-order 25-point schemes have identical $\lambda_{\max}$.

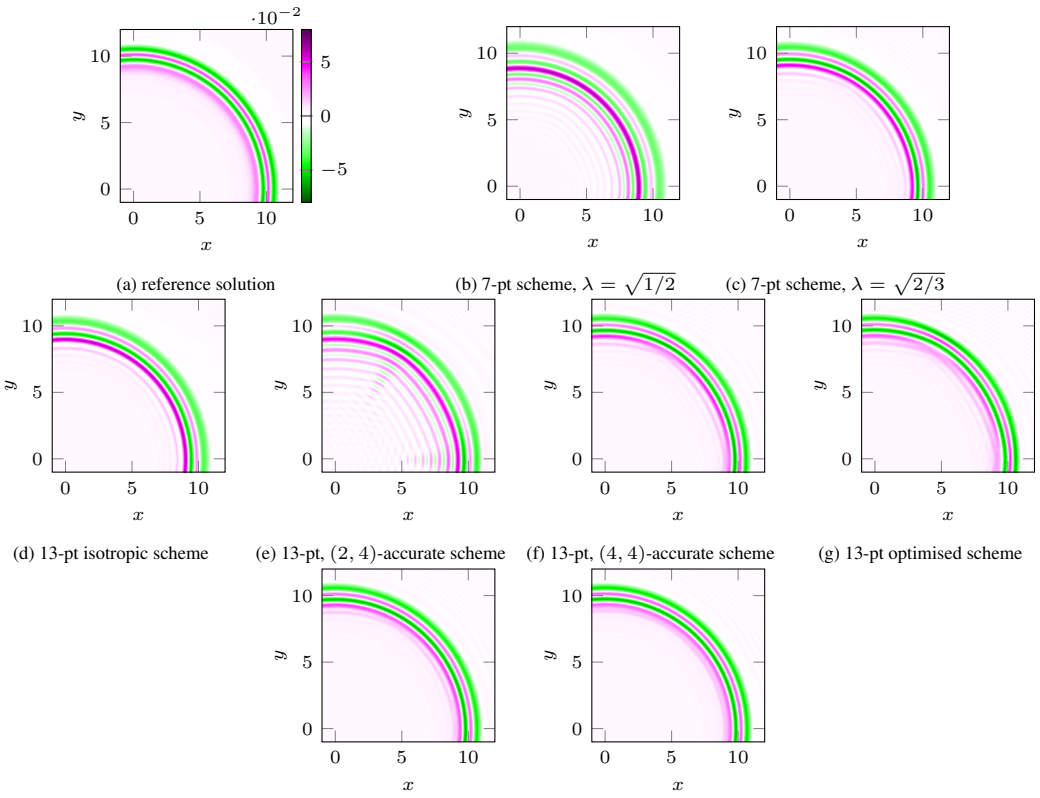
4.4.6 Numerical example

In order to see the kinds of numerical approximations obtained from these schemes, and to compare them to the Cartesian schemes, we carry out a numerical example, using the same initial conditions and spatiotemporal grid densities that were used in Section 4.3.8 (in the second scenario only, fixed spatiotemporal costs). In other words, we use a Ricker wavelet defined by (4.155) with σ^2 set to $0.08\,\mathrm{m}^2$, and $c = 1\,\mathrm{m/s}$. To fix spatiotemporal densities as in the Cartesian case, we set $h = (\mu/\lambda)^{1/3}$ with $\mu = 2/\sqrt{3}$ representing the density of the hexagonal grid (for $h = 1$) [58] and $h^* = 0.2\,\mathrm{m}$ (as in Section 4.3.8). Thus, the examples presented here have the same computational costs as their Cartesian counterparts with same stencil sizes K. The 19-point stencil ($K = 18$) has no similar-sized counterpart in the Cartesian family of schemes presented earlier, so it may be compared to the fourth-order accurate, sixth-order isotropic 21-point Cartesian scheme.

Snapshots of these simulations are shown in Fig. 4.42, along with the reference solution. Output time-responses are shown in Fig. 4.43, with residuals (error signals) in Fig. 4.44. In this case, the "diagonal direction" is along the normal $\hat{\mathbf{n}} = (\sqrt{3}/2, 1/2)$.

Discussion

For a brief discussion of these results, we compare them to the analogous Cartesian counterparts in Section 4.3.8. First we can compare the seven-point hexagonal scheme with ($\lambda = \lambda_{\max}$) to the nine-point IWB scheme in Section 4.3.8, in which we see that in terms of worst-case directional errors (along diagonals), the hexagonal scheme (Fig. 4.44(b)) outperforms the IWB scheme (Fig 4.31(e)). We can also compare the 13-point fourth-order hexagonal scheme (Fig. 4.44(e)) to the analogous fourth-order Cartesian scheme (Fig. 4.31(h)), and we find that the hexagonal scheme provides a significant improvement (with the same stencil size). Comparing the 19- and 25-point hexagonal schemes respectively to the 21- and 25-point Cartesian schemes, we find that the hexagonal schemes demonstrate smaller errors than the respective Cartesian counterparts. We can thus conclude, at least for this numerical example, that hexagonal schemes generally outperform Cartesian schemes with similar stencil sizes.

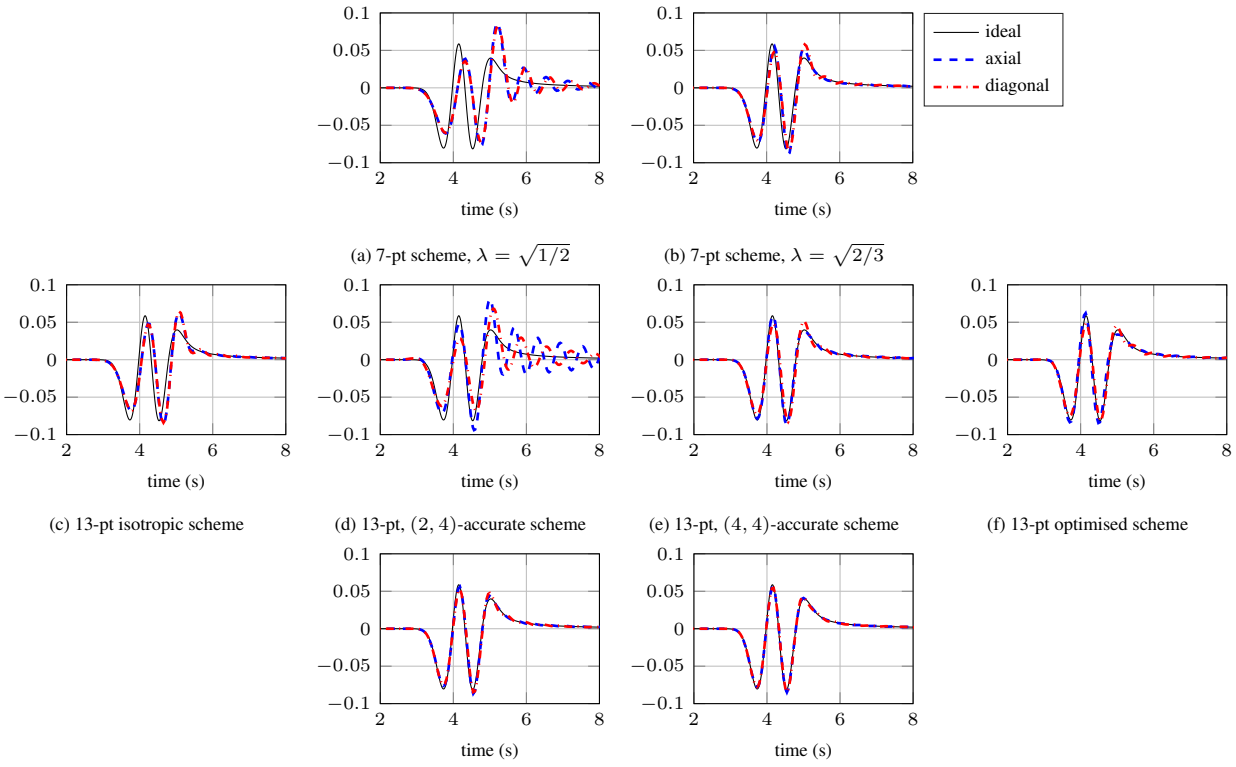


(a) reference solution (b) 7-pt scheme, $\lambda = \sqrt{1/2}$ (c) 7-pt scheme, $\lambda = \sqrt{2/3}$

(d) 13-pt isotropic scheme (e) 13-pt, $(2, 4)$-accurate scheme (f) 13-pt, $(4, 4)$-accurate scheme (g) 13-pt optimised scheme

(h) 19-pt, $(4, 4)$-accurate scheme (i) 25-pt, $(6, 6)$-accurate scheme

Figure 4.42: Snapshots of simulations from various schemes normalised for spatiotemporal densities. All plots use the same colour mapping, shown in (a). Cubic spline interpolation is used for additional detail in the approximate solutions, and only a quarter-space region is shown (by symmetry).



(a) 7-pt scheme, $\lambda = \sqrt{1/2}$

(b) 7-pt scheme, $\lambda = \sqrt{2/3}$

(c) 13-pt isotropic scheme

(d) 13-pt, $(2, 4)$-accurate scheme

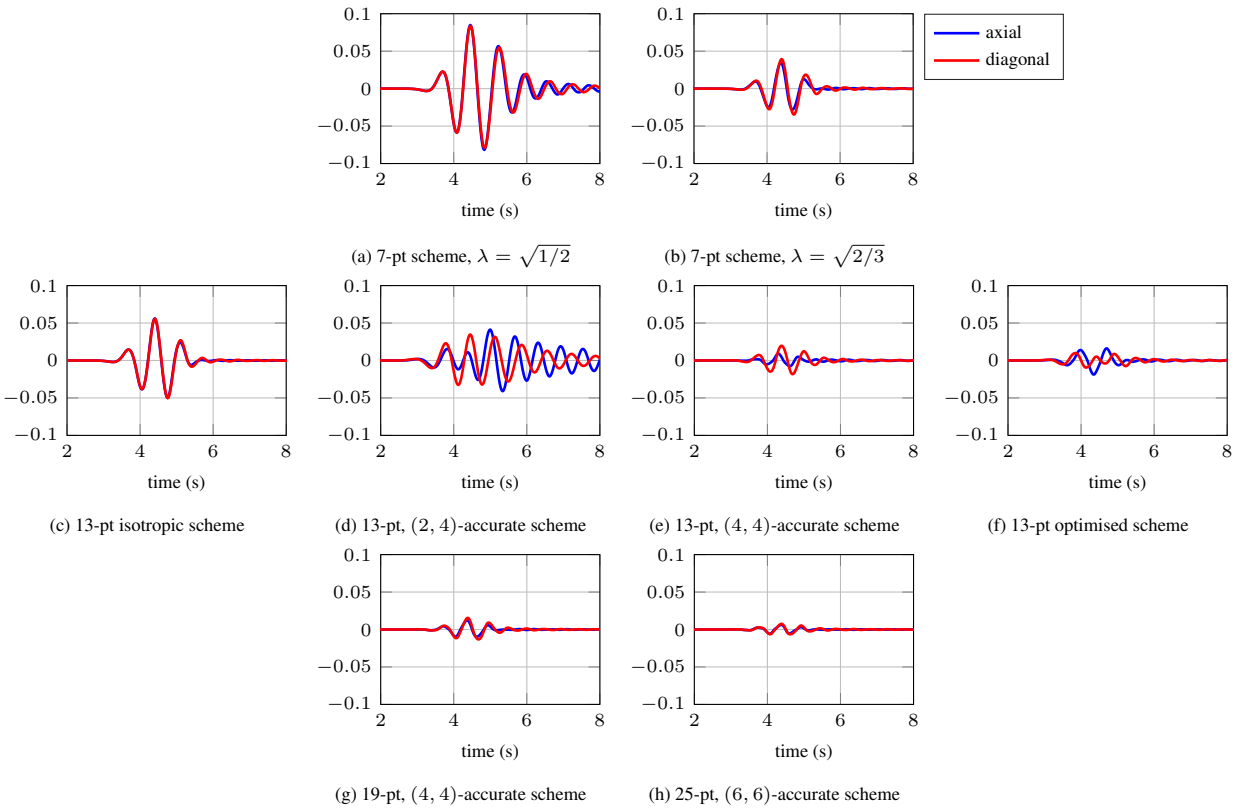
(e) 13-pt, $(4, 4)$-accurate scheme

(f) 13-pt optimised scheme

(g) 19-pt, $(4, 4)$-accurate scheme

(h) 25-pt, $(6, 6)$-accurate scheme

Figure 4.43: Plots of temporal outputs from various schemes, at distance $3\sqrt{2}$ m from origin, along axial direction and along diagonal direction. Here, the various schemes are normalised for spatiotemporal densities. Cubic spline interpolation is used for additional detail in the approximate solutions.



(a) 7-pt scheme, $\lambda = \sqrt{1/2}$

(b) 7-pt scheme, $\lambda = \sqrt{2/3}$

(c) 13-pt isotropic scheme

(d) 13-pt, $(2, 4)$-accurate scheme

(e) 13-pt, $(4, 4)$-accurate scheme

(f) 13-pt optimised scheme

(g) 19-pt, $(4, 4)$-accurate scheme

(h) 25-pt, $(6, 6)$-accurate scheme

Figure 4.44: Plots of error signals from Fig. 4.43 (normalised for spatiotemporal densities).

4.5 Cartesian schemes for the 3-D wave equation

In this section we return to the wave equation in three dimensions, which is defined by:

$$\Box u = 0\,, \quad \Box = \partial_t^2 - c^2\Delta\,, \quad \Delta = \partial_x^2 + \partial_y^2 + \partial_z^2 \tag{4.195}$$

with $u = u(\mathbf{x}, t)$ where $\mathbf{x} \in \mathbb{R}^3$, $t \in \mathbb{R}^+$. Required initial conditions for this second-order equation are $u(\mathbf{x}, 0)$ and $u_t(\mathbf{x}, 0)$. The grid function for this section is $\underline{u}_{\boldsymbol{i}}^n \approxeq u(\boldsymbol{i}h, nk)$, where $\boldsymbol{i} \in \mathbb{Z}^3$ and $n \geq 0$. We will also make use of the continuous approximation $\underline{u}(\mathbf{x}, t)$ where $\mathbf{x} \in \mathbb{R}^3$, $t \in \mathbb{R}^+$, and $\underline{u}(\boldsymbol{i}h, nk) = \underline{u}_{\boldsymbol{i}}^n$.

The goal of this section is to make use of the insights obtained from the previous sections about 2-D schemes, in order to derive schemes that are more efficient than the simple seven-point scheme discussed at length in Chapter 3. In this section we will investigate analogous "leggy" schemes with high orders of accuracy in space [8, 7, 53], a well-known family of compact explicit schemes [221, 21, 131] (analogous to the compact nine-point family discussed in Section 4.3.4), new wide-stencil schemes with high orders of accuracy in both space and time, and finally optimised wide-stencil schemes. Non-Cartesian schemes in 3-D (which also appear in closely-related Cartesian forms) will be discussed subsequently in Section 4.6.

4.5.1 Parametrised approximations to the 3-D Laplacian

As in the 2-D case, the finite difference schemes to follow will be of the form (4.20), where all that will change are the choices of the discrete Laplacian δ_Δ (and the chosen Courant numbers). It will thus be informative to present a general family of parametrised discrete Laplacians, as in the 2-D case.

Following [25], we consider averaging operations $\mu_{b,y}, \mu_{b,y}, \mu_{b,z}$ in three spatial dimensions, where $\mu_{b,x}$ and $\mu_{b,y}$ are defined by (4.100), and $\mu_{b,z}$ is similarly defined. From the Taylor expansions of these averaging operators, it follows that:

$$\mu_{q_1,x}\mu_{q_2,y}\mu_{q_3,z}\underline{u} = 1 + \frac{q_1^2h^2}{2}\partial_x^2\underline{u} + \frac{q_2^2h^2}{2}\partial_y^2\underline{u} + \frac{q_3^2h^2}{2}\partial_z^2\underline{u} \tag{4.196}$$

where $\boldsymbol{q} = (q_1, q_2, q_3)$ is a triplet of integers, and $\boldsymbol{q} \in \mathbb{Q}$, where $\mathbb{Q}$ is redefined as:

$$\mathbb{Q} := \{\boldsymbol{q} \in \mathbb{Z}^3 : q_1 \geq q_2 \geq q_3 \geq 0, q_1 \geq 1\} \tag{4.197}$$

We can then define a discrete Laplacian $\delta_{\Delta,\boldsymbol{q}} = \delta_{\Delta,(q_1,q_2,q_3)}$ as [25]:

$$\begin{aligned}\delta_{\Delta,\boldsymbol{q}} = \frac{1}{(q_1^2+q_2^2+q_3^2)h^2}\Big(&\mu_{q_1,x}\mu_{q_2,y}\mu_{q_3,y} + \mu_{q_1,x}\mu_{q_3,y}\mu_{q_2,y} + \mu_{q_2,x}\mu_{q_1,y}\mu_{q_3,y} \\ &+ \mu_{q_2,x}\mu_{q_3,y}\mu_{q_1,y} + \mu_{q_3,x}\mu_{q_1,y}\mu_{q_2,y} + \mu_{q_3,x}\mu_{q_2,y}\mu_{q_1,y} - 6\Big)\end{aligned} \tag{4.198}$$

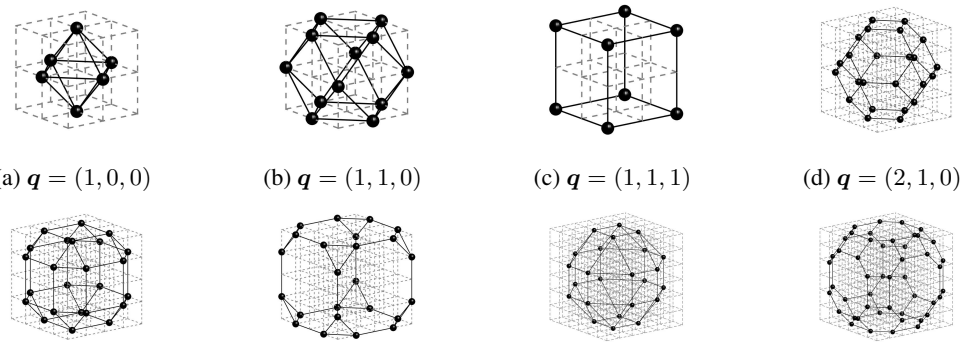
and from the above Taylor expansions we have:

$$\delta_{\Delta,\boldsymbol{q}}\underline{u} = \Delta\underline{u} + \mathcal{O}(h^2) \tag{4.199}$$

This discrete Laplacian reaches out to neighbouring points that are q_1, q_2, q_3 spatial steps away in respective x, y, z coordinate axes (and permutations thereof). Let us define $\boldsymbol{q}' = (q_x', q_y', q_z')$ and $\mathcal{P}(\boldsymbol{q})$ as the set of all *unique* permutations of $\{\pm q_1, \pm q_2, \pm q_3\}$, with $K = |\mathcal{P}(\boldsymbol{q})|$. As such, these operators use

$(K+1)$-point stencils where $K \in \{6, 8, 12, 24, 48\}$. More specifically, $K = 6$ when $q_2 = q_3 = 0$, $K = 12$ when $q_1 = q_2, q_3 = 0$, $K = 24$ when $q_1 \neq q_2 = q_3 > 0$ or when $q_1 \neq q_2 \neq q_3 = 0$, and $K = 48$ when $q_1 \neq q_2 \neq q_3 \neq 0$. The simple 7-point discrete Laplacian is denoted $\delta_{\Delta,(1,0,0)}$ using this notation.

For each such stencil, the K neighbouring grid points which are selected by the discrete Laplacian could be said to comprise a (partial) *shell of points*[38] on the cubic lattice with a squared norm of $q_1^2 + q_2^2 + q_3^2$. Some of the shells of points arising from choices of $\boldsymbol{q}$ are shown in Fig. 4.45, along with their convex hulls for illustrative purposes.[39]



(a) $\boldsymbol{q} = (1, 0, 0)$ (b) $\boldsymbol{q} = (1, 1, 0)$ (c) $\boldsymbol{q} = (1, 1, 1)$ (d) $\boldsymbol{q} = (2, 1, 0)$

(e) $\boldsymbol{q} = (2, 1, 1)$ (f) $\boldsymbol{q} = (2, 2, 1)$ (g) $\boldsymbol{q} = (3, 1, 1)$ (h) $\boldsymbol{q} = (3, 2, 1)$

Figure 4.45: Various shells of points on the cubic lattice and associated convex hulls. Bounding boxes rescaled for clarity.

For convenience, it helps to rewrite this operator as:

$$\delta_{\Delta,\boldsymbol{q}} \underline{u}_{\boldsymbol{i}}^n = \frac{6/K}{(q_1^2 + q_2^2 + q_3^2)h^2} (Q_{\boldsymbol{i},\boldsymbol{q}}^n - K \underline{u}_{\boldsymbol{i}}^n) \tag{4.200}$$

where $Q_{\boldsymbol{i},\boldsymbol{q}}^n = Q_{\boldsymbol{i},(q_1,q_2,q_3)}^n$ represents a sum over the neighbouring K grid points, defined by:

$$Q_{\boldsymbol{i},\boldsymbol{q}}^n = \sum_{\boldsymbol{q}' \in \mathcal{P}(\boldsymbol{q})} \underline{u}_{i_x+q_x', i_y+q_y', i_z+q_z'}^n \tag{4.201}$$

and we recall that $\mathcal{P}(\boldsymbol{q})$ is the set of all *unique* permutations of $\{\pm q_1, \pm q_2, \pm q_3\}$. Next we define a family of parametrised Laplacians as linear combinations of these operators. This will closely follow the 2-D case.

Consider a set of P distinct integer triplets $\boldsymbol{q} = (q_1, q_2, q_3)$ defined as $\Upsilon = \{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_P\}$, and let us define a general family of parametrised discrete Laplacians by:

$$\delta_{\Delta,\Upsilon} = \sum_{(q_1,q_2,q_3) \in \Upsilon} \alpha_{(q_1,q_2,q_3)} \delta_{\Delta,(q_1,q_2,q_3)} \tag{4.202}$$

[38] We note that some triplets $\boldsymbol{q}$ have the same Euclidean norm as other distinct triplets, such as $\boldsymbol{q} = (3, 0, 0)$ and $\boldsymbol{q} = (2, 2, 1)$, and as such one could consider these two triplets to form one shell of points, as in [58, Table 4.3]. In cases such as these, each $\boldsymbol{q}$ results in a *partial* shell of points, yet each triplet leads to a distinct discrete Laplacian, and it is useful to treat them as such.

[39] It is worth pointing out some commonly known regular and semi-regular polyhedrons displayed in these figures. An octahedron is featured in Fig. 4.45(a), a cuboctahedron in Fig. 4.45(b), a cube in Fig. 4.45(c), and a truncated octahedron in Fig. 4.45(d).

where $\alpha_{(q_1,q_2,q_3)}$ is a real-valued parameter, and for consistency we require that:

$$\sum_{(q_1,q_2,q_3)\in\Upsilon} \alpha_{(q_1,q_2,q_3)} = 1 \tag{4.203}$$

which will be implicitly assumed. The Fourier symbols of these operators are:

$$\delta_{\Delta,\Upsilon} \to -\frac{4}{h^2} G(\boldsymbol{\beta}_h) \tag{4.204}$$

where

$$G(\boldsymbol{\beta}_h) = \sum_{(q_1,q_2,q_3)\in\Upsilon} \frac{\alpha_{(q_1,q_2,q_3)}}{4(q_1^2+q_2^2+q_3^2)} \Big(6 - c_{x,q_1} c_{y,q_2} c_{z,q_3} - c_{x,q_1} c_{y,q_3} c_{z,q_2} - c_{x,q_2} c_{y,q_1} c_{z,q_3} \\ - c_{x,q_2} c_{y,q_3} c_{z,q_1} - c_{x,q_3} c_{y,q_1} c_{z,q_2} - c_{x,q_3} c_{y,q_2} c_{z,q_1} \Big) \tag{4.205}$$

Here, $c_{w,b} = \cos(b\beta_w h)$, and $w \in \{x,y,z\}, b \in \{q_1,q_2,q_3\}$. It follows that the set of free parameters α must be chosen such that $G(\boldsymbol{\beta}_h) \geq 0$, i.e., such that $\delta_{\Delta,\Upsilon}$ remains negative semi-definite.

Standard (leggy) $2M$th-order accurate discrete Laplacians

A family of $2M$th-order accurate approximations to the Laplacians on the Cartesian grid results when conventional leggy stencils are used in 3-D. Under this framework, leggy stencils result for $q_2 = q_3 = 0$. These stencils are described by the sets:[40]

$$\Upsilon_M^+ = \{(1,0,0), \ldots, (M,0,0)\} \tag{4.206}$$

We define these $2M$th-accurate 3-D discrete Laplacians as:

$$\delta_{\Delta,\Upsilon_M^+} = \sum_{m=1}^{M} \alpha_{(m,0,0)} \delta_{\Delta,(m,0,0)} \tag{4.207}$$

and it will be assumed that:

$$\alpha_{(m,0,0)} = \alpha_{M,m} \quad \Longrightarrow \quad \delta_{\Delta,\Upsilon_M^+} \underline{u} = \Delta \underline{u} + \mathcal{O}(h^{2M}) \tag{4.208}$$

where $\alpha_{M,m}$ is defined by (4.59). The $(6M+1)$-point stencils used by these operators are displayed in Fig. 4.46.

(a) $M = 1$ (b) $M = 2$ (c) $M = 3$ (d) $M = 4$

Figure 4.46: $(6M+1)$-point leggy stencils for $M = 1, 2, 3, 4$. Bounding boxes are rescaled for clarity.

[40] We use the same symbols as in the 2-D case for convenience.

A family of compact $(K+1)$-point discrete Laplacians with $K \leq 26$

Another commonly-used family of discrete Laplacians [257, 188] is given by:

$$\delta_{\Delta,\Upsilon_{27}} = \alpha_{(1,0,0)}\delta_{\Delta,(1,0,0)} + \alpha_{(1,1,0)}\delta_{\Delta,(1,1,0)} + \alpha_{(1,1,1)}\delta_{\Delta,(1,1,1)} \tag{4.209}$$

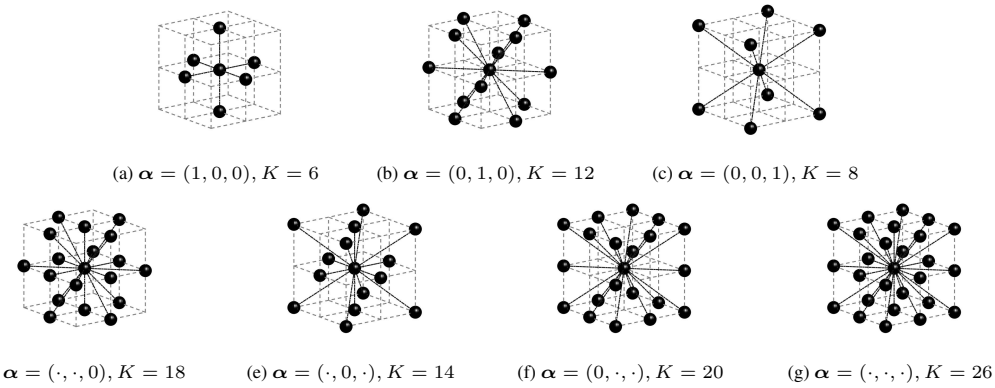
with the associated set:

$$\Upsilon_{27} = \{(1,0,0),(1,1,0),(1,1,1)\} \tag{4.210}$$

For consistency, we need:

$$\alpha_{(1,0,0)} + \alpha_{(1,1,0)} + \alpha_{(1,1,1)} = 1 \tag{4.211}$$

The $(K+1)$-point stencils used by this operator have $K \in \{6, 12, 14, 18, 20, 26\}$ depending on the choice of $\boldsymbol{\alpha} = (\alpha_{(1,0,0)}, \alpha_{(1,1,0)}, \alpha_{(1,1,1)})$. Some example choices of $\boldsymbol{\alpha}$ are illustrated in Fig. 4.47.



(a) $\boldsymbol{\alpha} = (1,0,0)$, $K = 6$ (b) $\boldsymbol{\alpha} = (0,1,0)$, $K = 12$ (c) $\boldsymbol{\alpha} = (0,0,1)$, $K = 8$

(d) $\boldsymbol{\alpha} = (\cdot,\cdot,0)$, $K = 18$ (e) $\boldsymbol{\alpha} = (\cdot,0,\cdot)$, $K = 14$ (f) $\boldsymbol{\alpha} = (0,\cdot,\cdot)$, $K = 20$ (g) $\boldsymbol{\alpha} = (\cdot,\cdot,\cdot)$, $K = 26$

Figure 4.47: $(K+1)$-point stencils for various choices of $\boldsymbol{\alpha} = (\alpha_{(1,0,0)}, \alpha_{(1,1,0)}, \alpha_{(1,1,1)})$. In the captions, "$\cdot$" is a placeholder for a non-zero value.

Through a Taylor expansion, it is straightforward to check that an isotropic discrete Laplacian results under the following constraint:

$$2\alpha_{(1,0,0)} - \alpha_{(1,1,0)} - 4\alpha_{(1,1,1)} = 0 \tag{4.212}$$

which, under the consistency condition, reduces to: $\alpha_{(1,1,0)} = 4/3 - 2\alpha_{(1,0,0)}$. Under these constraints we have:

$$\delta_{\Delta,\Upsilon_{27}}\underline{u} = \Delta\underline{u} + \frac{h^2}{12}\Delta^2\underline{u} + \mathcal{O}(h^4) \tag{4.213}$$

Analogous to the isotropic nine-point 2-D discrete Laplacian ((4.118)) is the isotropic 19-point 3-D discrete Laplacian that results from the choice $\boldsymbol{\alpha} = (\frac{1}{3}, \frac{2}{3}, 0)$ [257]. This discrete Laplacian is also related to the "D3Q19" model in the lattice Boltzmann method (LBM) literature [267]. Other discrete Laplacians related to the LBM models are the so-called "D3Q15" ($\boldsymbol{\alpha} = (\frac{2}{3}, 0, \frac{1}{3})$), and "D3Q27" ($\boldsymbol{\alpha} = (\frac{4}{9}, \frac{4}{9}, \frac{1}{9})$) models [267]. Another interesting choice is $\boldsymbol{\alpha} = (\frac{5}{12}, \frac{1}{2}, \frac{1}{12})$ [139].

4.5.2 Schemes with high-order spatial differencing in 3-D

Having already discussed the simplest Cartesian scheme in Chapter 3, we begin our investigation of finite difference schemes for the 3-D wave equation with the family of leggy $(2, 2M)$-accurate schemes (which includes the simplest scheme, for $M = 1$). Such two-step schemes, which are standard tools for

geophysical applications of the acoustic wave equation [66, 53, 71, 72], have also been investigated to some extent for room acoustics [199, 290].[41] This section will closely follow the presentation given for the analogous 2-D leggy schemes.

These $(2, 2M)$-accurate leggy schemes are represented by (4.20) with the discrete Laplacian $\delta_\Delta = \delta_{\Delta,\Upsilon_M^+}$ (in 3-D) and where $\alpha_{(m,0,0)} = \alpha_{M,m}$. We can also write these schemes as the update recursion:

$$\underline{u}^{n+1}_{i_x,i_y,i_z} = 2\underline{u}^{n}_{i_x,i_y,i_z} - \underline{u}^{n-1}_{i_x,i_y,i_z} + \lambda^2 \sum_{m=-M}^{M} a_{M,m} \left(\underline{u}^{n}_{i_x+m,i_y,i_z} + \underline{u}^{n}_{i_x,i_y+m,i_z} + \underline{u}^{n}_{i_x,i_y,i_z+m} \right) \quad (4.214)$$

where $a_{M,m}$ are defined by (4.59). The numerical dispersion relations of these schemes are of the form (4.26) with:

$$G(\boldsymbol{\beta}_h) = \sum_{m=1}^{M} \gamma_m \left(\sin^{2m}(\beta_x h/2) + \sin^{2m}(\beta_y h/2) + \sin^{2m}(\beta_z h/2) \right) \quad (4.215)$$

where γ_m are given by (4.71), and for stability one requires that:

$$\lambda \leq \lambda_{\max,M} = \frac{1}{\sqrt{3\Lambda_M}} \quad (4.216)$$

where Λ_M is given by (4.72). A sufficient condition for all $M \geq 1$ is $\lambda \leq \lambda_{\max,\infty} = \frac{2}{\sqrt{3}\pi}$. It is worth recalling that, despite the high-order accuracy in space, for a fixed Courant number these schemes are globally only second-order accurate.

As in the 2-D case, we can show from (4.215) and (4.216) that for $\lambda = \lambda_{\max,M}$ we have:

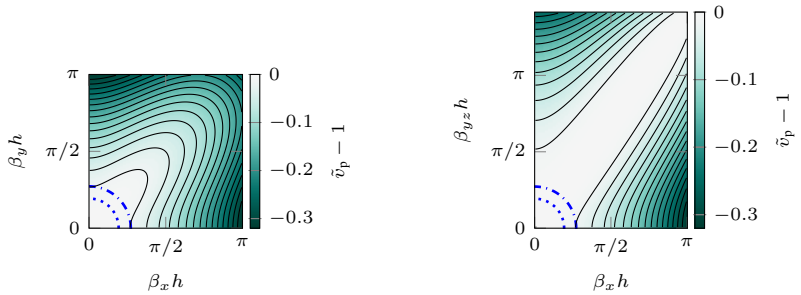
$$\omega_k((\pm\beta_x, 0)) = \omega_k((0, \pm\beta_y)) = 2\sin^{-1}\left(\frac{1}{\sqrt{3}}\right) \approx 0.392\pi \quad (4.217)$$

for all $M \geq 1$. As such, these 3-D schemes can reproduce no more than 39.2% of the temporal Nyquist along axial directions, which serves as a limit on their "valid bandwidth".

Let us now analyse the numerical dispersion, or more specifically, the relative error in the dispersion relation (and in the numerical phase velocity). Since $\boldsymbol{\beta}_h$ is now a vector in three dimensions, we will only be able to partially display the numerical dispersion properties of each scheme. By symmetry, it will suffice to consider only the positive octant of $\mathbb{B}$, as was carried out in Chapter 3 (Fig. 3.4), and we will investigate a 2-D through the 3-D wavenumber space along the plane that passes through the origin and has a normal vector $\hat{\mathbf{n}} = (0, \sqrt{1/2}, -\sqrt{1/2})$; i.e, when $\beta_y = \beta_z$. Since this normal vector is a *side-diagonal* direction, we will call this 2-D slice the "side-diagonal slice". For illustrative purposes we will also feature a 2-D slice of the relative error along the plane for which $\beta_z = 0$, which we will call "the axial slice". Following Fig. 3.4(b), let us also define $\beta_{yz} = (\beta_y + \beta_z)/\sqrt{2}$, which extends to $\sqrt{2}\pi$ along the aforementioned side-diagonal slice.

Contour plots of relative dispersion errors along these two slices, as shown for the simplest scheme in Fig. 4.48, will be used throughout this section on 3-D schemes. We note that Fig. 4.48(b) is essentially an elevated view of Fig. 3.4(b). The side-diagonal direction appears along the y-axis of Fig. 4.48(b), and the *diagonal direction* extends from the origin to the top right-corner of Fig. 4.48(b) where the wavenumber reaches $\beta_h = \sqrt{3}\pi$.

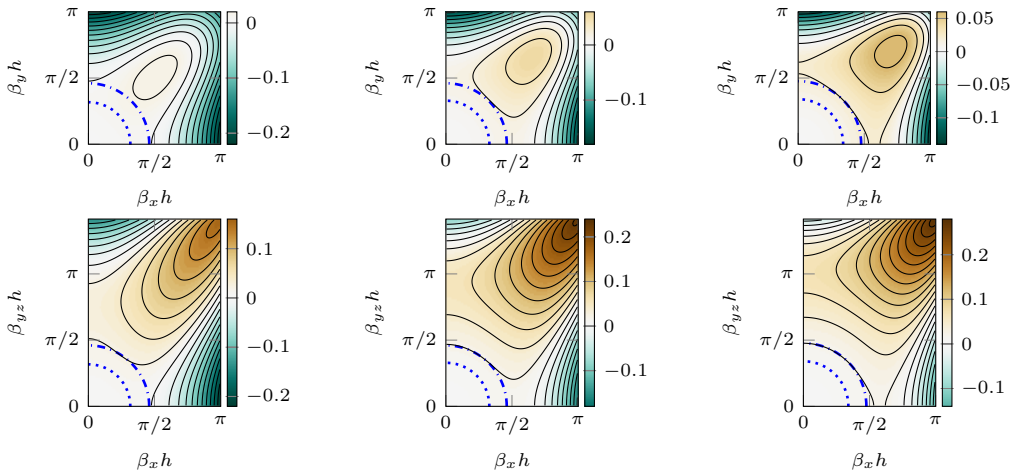
[41] In [199], the $(2, 6)$-accurate leggy scheme was employed, but the Courant number used was not specified. In [290], the leggy schemes for $M \leq 8$ were analysed, but only for $\lambda = \lambda_{\max}$, which is not necessarily optimal in terms of dispersion errors, as is well-known [66, 53], and as will be investigated in this section.



(a) $\beta_z = 0$ axial slice (b) $\beta_y = \beta_z$ side-diagonal slice

Figure 4.48: Contour plot of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for the simplest seven-point 3-D scheme with $\lambda = \sqrt{1/3} \approx 0.58$ along a 2-D slice through the cubic wavenumber oriented along (a): $\beta_z = 0$ plane, and (b): $\beta_y = \beta_z$ "side-diagonal" plane (for which we define $\beta_{yz} = (\beta_y + \beta_z)/\sqrt{2}$). Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

Analogous representations of the relative dispersion errors for $(2, 2M)$-order schemes are shown in Fig. 4.49. As in the 2-D versions of these schemes, we find that there is a noticeable improvement from $M = 1$ (the simplest scheme, in Fig. 4.48) to $M = 2$ (Fig. 4.49(a)) in terms of critical wavenumbers $\beta_{h,1\%}$ and $\beta_{h,2\%}$, and only marginal improvements thereafter with higher M (Figs. 4.49(b)-(c)).



(a) $M = 2, \lambda = \lambda_\mathrm{max} = 0.5$ (b) $M = 3, \lambda = \lambda_\mathrm{max} \approx 0.47$ (c) $M = 4, \lambda = \lambda_\mathrm{max} \approx 0.45$

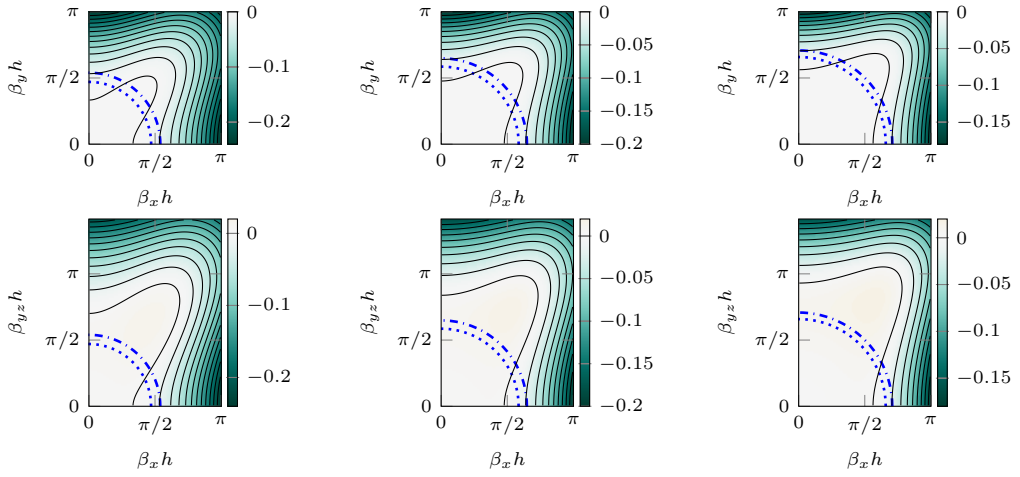
Figure 4.49: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for $(2, 2M)$-accurate schemes using $(6M + 1)$-point leggy stencils in 3-D, with $\lambda = \lambda_{\mathrm{max},M}$, for $\boldsymbol{\beta}_h \in [0, \pi]^3$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

As in the 2-D case of leggy schemes, there are better choices for the Courant number than $\lambda = \lambda_{\mathrm{max},M}$ for $M > 1$, which provide better balances of $\mathcal{O}(T^2)$ errors with $\mathcal{O}(X^{2M})$ errors. As in the 2-D case, we define $\lambda_{1\%,M}$ to be the Courant number $\lambda \in [\frac{1}{100}\lambda_{\mathrm{max},M}, \frac{2}{100}\lambda_{\mathrm{max},M}, \ldots, \lambda_{\mathrm{max},M}]$ that maximises the critical wavenumber $\beta_{h,1\%}$ in a given $(2, 2M)$-accurate leggy scheme. These optimised Courant numbers are given in Table 4.8 for $M \leq 4$. The relative dispersion errors for these choices of λ are shown in Fig. 4.50. As in the 2-D case, there are significant improvements in the critical wavenumbers $\beta_{h,1\%}$ and $\beta_{h,2\%}$ over the respective choices of $\lambda = \lambda_{\mathrm{max},M}$, but at the cost of a lower time-step k for a fixed h. Later we will compare with other schemes under normalisations for computational costs.

It is also worth tabulating the minimal and global cutoff frequencies ($\omega_{k,\mathrm{m}}$ and $\omega_{k,\mathrm{g}}$, defined by (4.129))

Table 4.8: Optimised Courant numbers for $(6M+1)$-point leggy schemes in 3-D

M	$\lambda_{1\%,M}/\lambda_{\max,M}$	$\lambda_{1\%,M}$ (approx.)
1	1.00	0.577
2	0.73	0.365
3	0.59	0.277
4	0.53	0.240



(a) $M=2$, $\lambda=\lambda_{1\%,2}\approx 0.37$ (b) $M=3$, $\lambda=\lambda_{1\%,3}\approx 0.28$ (c) $M=4$, $\lambda=\lambda_{1\%,4}\approx 0.24$

Figure 4.50: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h)-1$ for $(2,2M)$-accurate schemes using $(6M+1)$-point leggy stencils in 3-D, with λ as indicated, and for $\boldsymbol{\beta}_h\in[0,\pi]^3$. Top row: 2-D slice through $\beta_z=0$ plane. Bottom row: 2-D slice through $\beta_y=\beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h|=\beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h|=\beta_{h,1\%}$.

for these schemes. These are listed in Table 4.9, wherein the top half pertains to the choices $\lambda=\lambda_{\max,M}$, and in the bottom half are the optimised choices where $\lambda<\lambda_{\max,M}$. It can be seen that a significant portion of the overall temporal bandwidth is sacrificed to improve the dispersion errors near DC under the choices of $\lambda=\lambda_{1\%,M}<\lambda_{\max,M}$. As in the 2-D case, we find that $\omega_{k,\mathrm{m}}$ appears along the axial direction for the leggy schemes with $\lambda=\lambda_{\max,M}$, and it remains constant for all $M\geq 1$ by (4.217).

Table 4.9: Cutoff frequencies for various $(6M+1)$-point $(2,2M)$-order accurate 3-D schemes

M	λ (approx.)	$\omega_{k,\mathrm{m}}$	$\omega_{k,\mathrm{g}}$
1	0.577	0.392π	1.000π
2	0.500	0.392π	1.000π
3	0.470	0.392π	1.000π
4	0.453	0.392π	1.000π
2	0.365	0.277π	0.521π
3	0.277	0.221π	0.402π
4	0.240	0.198π	0.356π

4.5.3 Compact $(K+1)$-point schemes in 3-D with $K\leq 26$

At this point we will investigate a well-known family of compact explicit schemes, employing up to 26 nearest neighbouring points on a Cartesian grid (analogous to the 2-D family in Section 4.3.4).

This family of schemes first appeared in the room acoustics literature as the 3-D "interpolated digital waveguide mesh" (DWM) [221] formulated by Savioja, which was later presented in a more general finite difference setting by Bilbao [21, 32]. Kowalczyk and van Walstijn analysed these schemes further, and re-dubbed them the "compact explicit schemes" [136]. We should also note that this family of schemes was also considered by Tuomela [280, 281], with a focus on the isotropic special cases.

We can represent this family of schemes as:

$$\delta_{tt}\underline{u}_{\boldsymbol{i}}^n = c^2\delta_{\Delta,\Upsilon_{27}}\underline{u}_{\boldsymbol{i}}^n \tag{4.218}$$

which is in the form of (4.20) with the 27-point discrete Laplacian $\delta_\Delta = \delta_{\Delta,\Upsilon_{27}}$ (defined by (4.209)). We also redefine $\boldsymbol{\alpha} = (\alpha_{(1,0,0)}, \alpha_{(1,1,0)}, \alpha_{(1,1,1)})$ for this scheme. The update recursion for this scheme can be written as:

$$\begin{aligned}\underline{u}_{i_x,i_y,i_z}^{n+1} = &\left(2 - \lambda^2(6\alpha_{(1,0,0)} + 3\alpha_{(1,1,0)} + 2\alpha_{(1,1,1)})\right)\underline{u}_{i_x,i_y,i_z}^n \\ &+ \lambda^2\left(\alpha_{(1,0,0)}Q_{\boldsymbol{i},(1,0,0)}^n + \frac{\alpha_{(1,1,0)}}{4}Q_{\boldsymbol{i},(1,1,0)}^n + \frac{\alpha_{(1,1,1)}}{4}Q_{\boldsymbol{i},(1,1,1)}^n\right) - \underline{u}_{i_x,i_y,i_z}^{n-1}\end{aligned} \tag{4.219}$$

and (4.211) is required for consistency, so one can rewrite $\alpha_{(1,1,1)} = 1 - \alpha_{(1,0,0)} - \alpha_{(1,1,0)}$. Henceforth, this consistency constraint will be implicitly assumed.

It is also worth noting that this scheme can be rewritten in the following form:

$$\delta_t^2\underline{u}_{\boldsymbol{i}}^n = \lambda^2\Big(\delta_x^2 + \delta_y^2 + \delta_z^2 + a(\delta_x^2\delta_y^2 + \delta_x^2\delta_z^2 + \delta_y^2\delta_z^2) + b\delta_x^2\delta_y^2\delta_z^2\Big)\underline{u}_{\boldsymbol{i}}^n \tag{4.220}$$

where $a = (\alpha_{(1,1,0)} + 2\alpha_{(1,1,1)})/4$ and $b = \alpha_{(1,1,1)}/4$, and where $\delta_t^2 := k^2\delta_{tt}$ and $\delta_w^2 := h^2\delta_{ww}$ for $w \in \{x, y, z\}$. This expression (and notation) is derived from the explicit special case of compact implicit schemes in [89], and it is the preferred form in [136].

Numerical stability

The numerical stability conditions for these schemes are well-known, but it will be informative to re-derive them here.

The schemes within this family are essentially linear combinations of the three special cases for which $\boldsymbol{\alpha} = (\alpha_{(1,0,0)}, \alpha_{(1,1,0)}, \alpha_{(1,1,1)})$ has only one non-zero coefficient. Let us consider these three special cases in isolation. We will use the respective monikers "SLF" (standard leapfrog) ($\boldsymbol{\alpha} = (1, 0, 0)$, the simplest scheme), "CCP" (cubic close-packed) ($\boldsymbol{\alpha} = (0, 1, 0)$), and "OCTA" (octahedral) ($\boldsymbol{\alpha} = (0, 0, 1)$), after [136]. It is worth noting here that the CCP and OCTA special cases are the starting points for the non-Cartesian schemes to be seen in Section 4.6.3, but in this section we treat the CCP and OCTA schemes under the assumption that they are operated on Cartesian grids (i.e., $\boldsymbol{i} \in \mathbb{Z}^3$), following [136].[42]

For these schemes, the numerical dispersion relation is again of the form (4.26), so our stability condition for the Courant number depends on $\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)$ and we need that $G(\boldsymbol{\beta}_h) \geq 0$. For these three special cases (SLF, CCP, and OCTA), let us rewrite $G(\boldsymbol{\beta}_h)$ as:

$$G(\boldsymbol{\beta}_h) = \sum_{\boldsymbol{v}\in\Upsilon}\frac{3/K}{\|\boldsymbol{v}\|^2}\sin^2(\boldsymbol{\beta}_h \cdot \boldsymbol{v}h/2) \tag{4.221}$$

[42] In Section 4.6.3 we consider them on face-centered cubic (FCC) and body-centered cubic (BCC) grids, respectively (with the associated non-Cartesian wavenumber cells).

where here Υ is the set of K vectors $\boldsymbol{v}$, each with Euclidean norm $\|\boldsymbol{v}\|$. These vectors represent the neighbouring points to the origin in the underlying stencils. More specifically:

- when $\boldsymbol{\alpha} = (1, 0, 0)$ (SLF), $K = 6$ and $\Upsilon = \pm\{\hat{\mathbf{e}}_x, \hat{\mathbf{e}}_y, \hat{\mathbf{e}}_z\}$ and $\|\boldsymbol{v}\| = 1$
- when $\boldsymbol{\alpha} = (0, 1, 0)$ (CCP), $K = 12$ and $\Upsilon = \pm\{\hat{\mathbf{e}}_x \pm \hat{\mathbf{e}}_y, \hat{\mathbf{e}}_x \pm \hat{\mathbf{e}}_z, \hat{\mathbf{e}}_y \pm \hat{\mathbf{e}}_z\}$ and $\|\boldsymbol{v}\| = \sqrt{2}$
- when $\boldsymbol{\alpha} = (0, 0, 1)$ (OCTA), $K = 8$ and $\Upsilon = \pm\{\hat{\mathbf{e}}_x \pm \hat{\mathbf{e}}_y \pm \hat{\mathbf{e}}_z, \hat{\mathbf{e}}_x \mp \hat{\mathbf{e}}_y \pm \hat{\mathbf{e}}_z\}$ and $\|\boldsymbol{v}\| = \sqrt{3}$

From (4.221), it is clear that $G(\boldsymbol{\beta}_h) \geq 0$. Furthermore, an upper bound on $G(\boldsymbol{\beta}_h)$ may be obtained in a straightforward manner:

$$G(\boldsymbol{\beta}_h) \leq \frac{3}{\|\boldsymbol{v}\|^2} \quad \Longrightarrow \quad \lambda \leq \frac{\|\boldsymbol{v}\|}{\sqrt{3}} \tag{4.222}$$

More specifically, we arrive at the following *sufficient* conditions for stability for the three special cases:

$$\text{SLF:} \quad \lambda \leq \sqrt{1/3}\,, \qquad \text{CCP:} \quad \lambda \leq \sqrt{2/3}\,, \qquad \text{OCTA:} \quad \lambda \leq 1 \tag{4.223}$$

Provided that $\alpha_{(1,0,0)}, \alpha_{(1,1,0)}\alpha_{(1,1,1)} \geq 0$, which guarantees that $G(\boldsymbol{\beta}_h) \geq 0$, we can arrive at a sufficient condition for the family of schemes:

$$\lambda \leq \sqrt{\frac{2}{4\alpha_{(1,0,0)} + \alpha_{(1,1,0)} + 2}} \tag{4.224}$$

When the above condition is satisfied with equality, we have the 3-D "interpolated DWMs" first presented by Savioja [221, 225].

As first pointed out in [21], the above condition, while sufficient, is not necessary in all cases. The necessary and sufficient condition can be obtained from the form (4.205), which is *multilinear* in the individual cosine terms $(c_{x,1}, c_{y,1}, c_{z,1})$ for this family of schemes. As such, the extrema of $G(\boldsymbol{\beta}_h)$ must occur at the vertices of the cubic region $\boldsymbol{\beta}_h \in [0, \pi]^3$, as shown in [21, 32]. For the condition $G(\boldsymbol{\beta}_h) \geq 0$, this leads to the following constraints on $\boldsymbol{\alpha}$:

$$-2\alpha_{(1,0,0)} \leq \alpha_{(1,1,0)} \leq 2\alpha_{(1,0,0)} + 1 \tag{4.225}$$

This constraint immediately rules out the possibility of a stable 21-point isotropic scheme with $\boldsymbol{\alpha} = (0, \frac{4}{3}, -\frac{1}{3})$. The stability condition on the Courant number as a function of $\boldsymbol{\alpha}$ (or simply $\alpha_{(1,0,0)}$ and $\alpha_{(1,1,0)}$) can be split into three convex regions [21], as illustrated in Fig. 4.51. This can be summarised by [32]:

$$\lambda \leq \lambda_{\max} = \min\left(1, \frac{1}{\sqrt{2\alpha_{(1,0,0)} + \alpha_{(1,1,0)}}}, \frac{1}{\sqrt{2\alpha_{(1,0,0)} - \alpha_{(1,1,0)} + 1}}\right) \tag{4.226}$$

Returning to the three special cases (SLF, CCP, and OCTA), we find that (4.223) is necessary and sufficient for the SLF and OCTA schemes[43], but that (4.223) is sufficient *but not necessary* for the CCP scheme. The necessary and sufficient condition for the CCP scheme is $\lambda \leq 1$.

It is also interesting to note that at the intersection of the three stability regions lies the special case $\boldsymbol{\alpha} = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ with $\lambda_{\max} = 1$, which has become commonly known as the 3-D "interpolated wideband (IWB)" scheme [136]. The IWB special case and other special cases investigated in [136] are listed in Table 4.10 under the monikers used in [136]. The so-called IISO and IISO2 ("IISO" stands for

[43] In fact, these are the only two cases for which this is true [21].

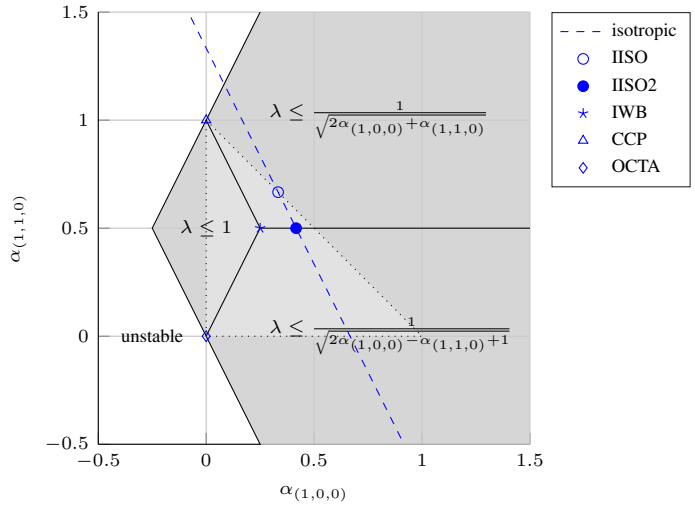


Figure 4.51: Stability regions for 27-point compact schemes (after [21]) as a function of $\alpha_{(1,0,0)}$ and $\alpha_{(1,1,0)}$. Special case schemes are labelled and isotropic schemes appear along the dashed line. The concave region to the left (not coloured) has $\min_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h) \leq 0$ and is therefore unstable. Concretely passive schemes (following (4.224)) are possible within the triangular region demarcated by the dotted lines.

"interpolated isotropic" [131, 136]) schemes are two special cases that have the highest possible $\lambda_{\max}$ for isotropic schemes within this family.[44]

Table 4.10: Special cases of 27-point compact explicit family of schemes

scheme	K	$\alpha_{(1,0,0)}$	$\alpha_{(1,1,0)}$	$\alpha_{(1,1,1)}$	$\lambda_{\max}$
SLF	6	1	0	0	$\sqrt{1/3}$
CCP	12	0	1	0	1
OCTA	8	0	0	1	1
IWB	26	$1/4$	$1/2$	$1/4$	1
IISO	18	$1/3$	$2/3$	0	$\sqrt{3/4}$
IISO2	26	$5/12$	$1/2$	$1/12$	$\sqrt{3/4}$

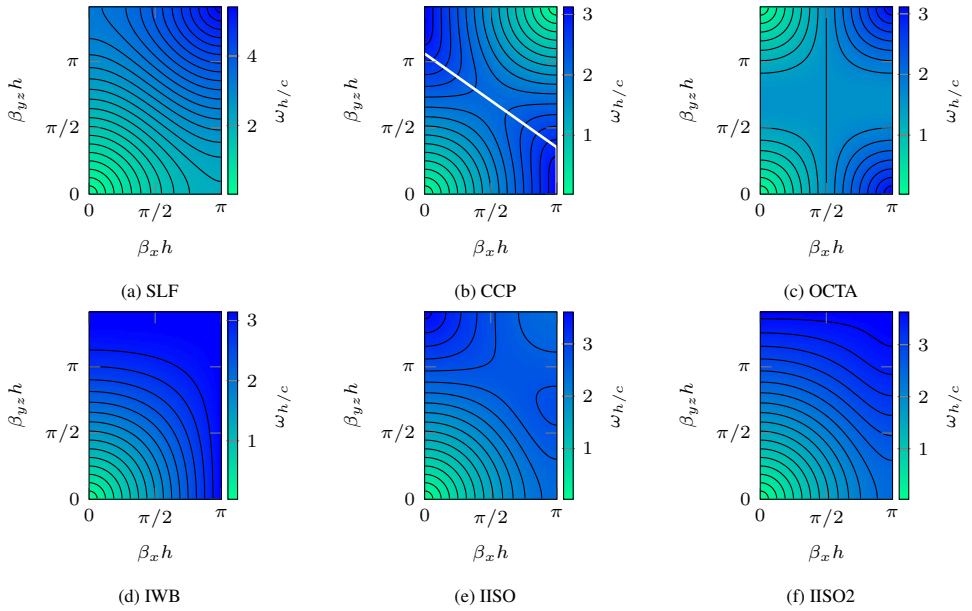
Numerical dispersion

Analyses of the numerical dispersion for this family of schemes have been presented in previous studies [21, 32, 135, 136], but we will reconsider them here in order to identify some important features which have been missing from previous analyses. Generally, the numerical dispersion of these schemes has been analysed in terms of the *numerical phase velocity*, either as a function of wavenumber [21, 32], or indirectly as function of temporal frequency and fixed wave direction via an "inversion" of the dispersion relation [136]. It will be useful here to begin with the numerical dispersion relations for this family of schemes, before analysing errors in relative phase velocities.

For the following discussion, we will say that a scheme has a "monotonic dispersion relation" when the dispersion relation increases monotonically with the wavenumber along *all* wave directions for which $\boldsymbol{\beta}_h \in [0, \pi]^3$. The significance of this property is that directional cutoff frequencies will be

[44] It is straightforward to check that any isotropic scheme in this family (i.e., with $\alpha_{1,1,0} = 4/3 - 2\alpha_{1,0,0}$) with $\alpha_{1,0,0} \geq 5/12$ has $\lambda_{\max} = \sqrt{3/4}$, as noted in [280].

found on the boundary of $[0,\pi]^3$, and global and minimal cutoff frequencies will be found at the vertices of $[0,\pi]^3$. Furthermore, a scheme has this property one can immediately "invert" the dispersion relation along any fixed wave direction in order to express β_h as a function of ω_k and $\hat{\boldsymbol{\beta}}$. It is under these assumptions that numerical dispersion of these schemes were analysed in [136], but it turns out that these assumptions do not hold in general for these schemes.



(a) SLF (b) CCP (c) OCTA

(d) IWB (e) IISO (f) IISO2

Figure 4.52: Contour plots of numerical dispersion relations of special cases of 27-point family, with $\lambda = \lambda_{\max}$, for side-diagonal slice through $[0,\pi]^3$. Contour lines denote increments of $\pi/16$.

The numerical dispersion relations $\omega_{h/c}(\boldsymbol{\beta}_h)$ for the aforementioned special case schemes are shown in Fig. 4.52, for $\boldsymbol{\beta}_h$ along a side-diagonal slice through $[0,\pi]^3$. From Fig. 4.52, we can immediately see the CCP, OCTA, and IISO schemes do not have this aforementioned monotonicity property.[45] Specifically, the CCP and IISO schemes have, respectively, local maxima at $\beta_h = \frac{\sqrt{3}\pi}{2}$ and $\beta_h = \frac{2\pi}{\sqrt{3}}$ along the diagonal direction, and the OCTA scheme has a local maximum along the side-diagonal direction at $\beta_h = \frac{\pi}{\sqrt{2}}$.

It can also be seen from Fig. 4.52 that the CCP and OCTA schemes feature multiple DC frequencies ($\omega_{h/c} = 0$) along this slice. This indicates that the dispersion relations are periodic with a fundamental tile that is non-cubic, similar to the rotated five-point scheme in 2-D (in Section 4.3.4).[46]

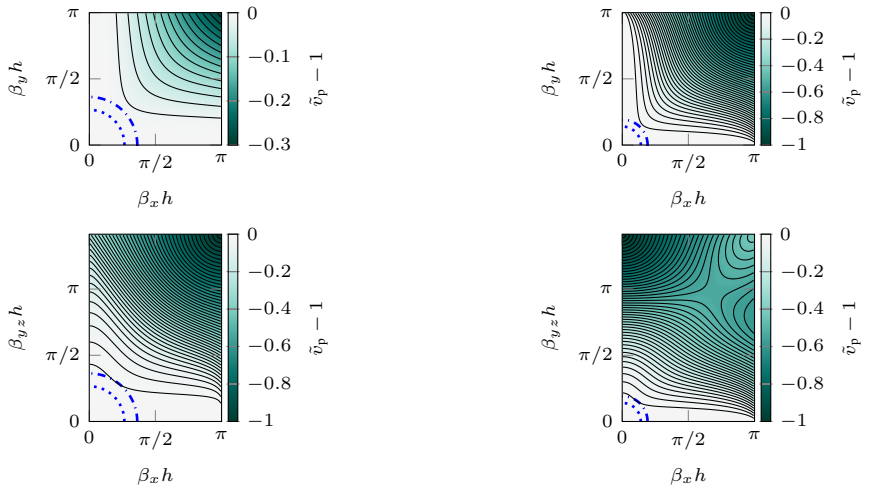
Numerical phase velocity

Having already investigated the numerical phase velocity of the simple 3-D Cartesian scheme (SLF) in Chapter 3 (Section 3.1.4), we start by investigating $\tilde{v}_{\mathrm{p}}$ for the CCP and OCTA schemes. Since we assume here that $\boldsymbol{i} \in \mathbb{Z}^3$, we will make use of the positive octant of the Cartesian wavenumber cell, i.e., $\boldsymbol{\beta}_h \in [0,\pi]^3$, and later we will consider these schemes in (more appropriate) non-Cartesian wavenumber cells. The relative dispersion error for these schemes with $\lambda = 1$ is plotted in Fig. 4.53, along the axial and side-diagonal slices. We see that the numerical relation is exact along axial directions for both of these schemes. The relative dispersion error also goes to 100% at some $\boldsymbol{\beta}_h \in [0,\pi]^3$ for both schemes;

[45] Since the associated $G(\boldsymbol{\beta}_h)$ is multilinear in the axial directions, the numerical dispersion relation must be monotonically increasing along the *axial* direction, but multilinearity does not necessarily extend to non-axial directions.

[46] Indeed, these schemes are more appropriately formulated on non-Cartesian grids; this will be treated in Section 4.6.

this is due to the presence of multiple DC frequencies within $\mathbb{B}$ (as mentioned previously), hinting at some form of spatial decoupling, to be discussed in more detail in Section 4.3.4.



(a) CCP, $\lambda = 1$

(b) OCTA, $\lambda = 1$

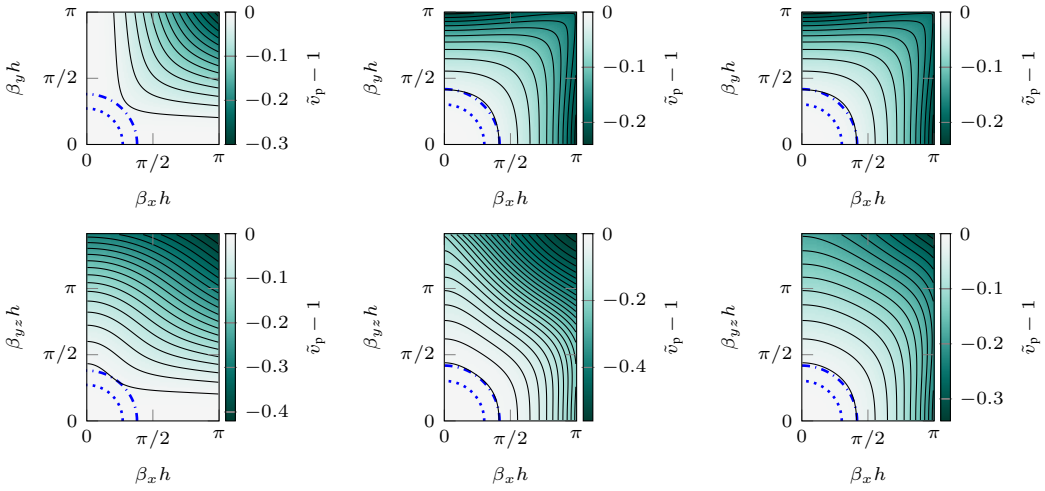
Figure 4.53: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for CCP and OCTA scheme, for $\boldsymbol{\beta}_h \in [0, \pi]^3$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

Next we consider three other special cases: the IISO, IISO2, and IWB schemes (with $\lambda = \lambda_{\max}$); these are shown in Fig. 4.54. It can be seen that, as expected, the IISO and IISO2 schemes exhibit isotropic behaviour near DC. We can also see that the dispersion errors along worst-case directions for these isotropic schemes appear to be identical (axial directions). As such, the 19-point scheme (the IISO scheme) is generally a better choice than the 27-point IISO2 scheme, as it requires fewer computations (nearly 30% fewer) per pointwise update for the same (worst-case) dispersion errors; this was also noted in [280].[47] The isotropic schemes also have slightly larger critical wavenumbers $\beta_{h,1\%}$ and $\beta_{h,2\%}$ than the IWB scheme, but they also operate at higher sample rates for the same h. As for the IWB scheme, it demonstrates exact dispersion along the axial directions, as in the analogous 2-D scheme (Fig. 4.14(c)).

Other special cases

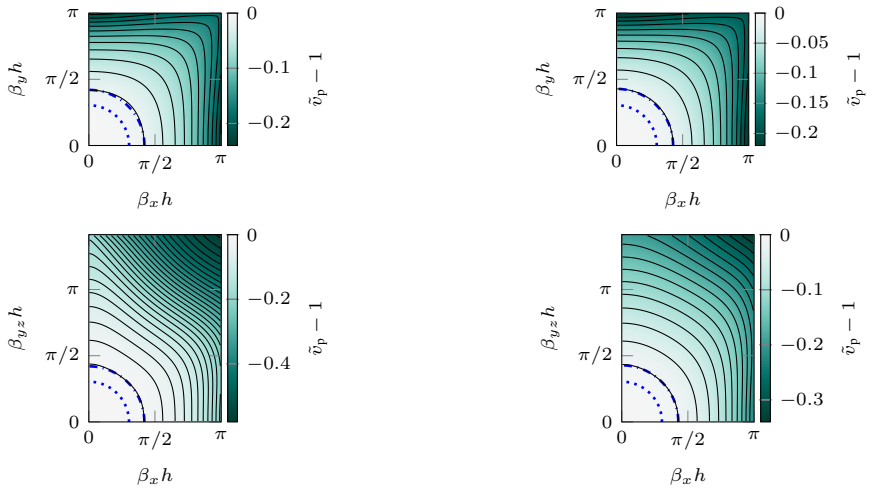
While the above three examples are essentially the optimal cases for which more than one element of $\boldsymbol{\alpha}$ is non-zero, there are alternatives not far from these cases that offer slight improvements in the one- and two-percent critical wavenumbers. For example, two schemes that are very close to the IISO and IISO2 schemes in their parameters are shown in Figs. 4.55(a) and 4.55(b). Strictly speaking, the schemes featured in Figs. 4.55(a) and 4.55(b) are *not* isotropic near DC, but specifically at the 2% error isocontour, these schemes appear to be more isotropic than the IISO and IISO2 schemes. As will be seen shortly, this leads to slightly better computational efficiency for that error threshold.

[47] We should that while the IISO scheme does not have a monotonic dispersion relation along the diagonal directions, this should not be seen as a deterrent to the use of this scheme, as this phenomenon only appears at wavenumbers for which dispersion errors are already quite significant in all schemes within this family (for at least one wave direction).



(a) $K = 26$, $\boldsymbol{\alpha} = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$, $\lambda = 1$ (b) $K = 18$, $\boldsymbol{\alpha} = (\frac{1}{3}, \frac{2}{3}, 0)$, $\lambda \approx 0.87$ (c) $K = 26$, $\boldsymbol{\alpha} = (\frac{5}{12}, \frac{1}{2}, \frac{1}{12})$, $\lambda \approx 0.87$

Figure 4.54: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for $(K+1)$-point compact schemes in 3-D, for $\boldsymbol{\beta}_h \in [0, \pi]^3$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.



(a) $K = 18$, $\boldsymbol{\alpha} = (\frac{129}{400}, \frac{271}{400}, 0)$, $\lambda \approx 0.87$ (b) $K = 26$, $\boldsymbol{\alpha} = (\frac{2}{5}, \frac{1}{2}, \frac{1}{10})$, $\lambda \approx 0.88$

Figure 4.55: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for $(K+1)$-point compact schemes in 3-D, for $\boldsymbol{\beta}_h \in [0, \pi]^3$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

Related 2-D schemes

It is worth pointing out some special relationships between this family of 27-point schemes and the 2-D compact nine-point scheme. For example, it can be seen that for the IWB scheme in Fig. 4.54(a), the $\beta_z = 0$ slice is identical to that of Fig. 4.53(a), and to Fig. 4.14(c). We also note that the $\beta_z = 0$ slices for the IISO and IISO2 schemes in Figs. 4.54(b) and 4.54(c) are identical to Fig. 4.14(b).

In fact, it is straightforward to work out when the solution is constant along one coordinate direction

(e.g., the z-direction), that analogous (related) 2-D schemes will satisfy the following relations:

$$\alpha_{(1,0)} = \alpha_{(1,0,0)} + \alpha_{(1,1,0)}/2\,, \quad \alpha_{(1,1)} = \alpha_{(1,1,0)}/2 + \alpha_{(1,1,1)} \tag{4.227}$$

As such, the 3-D IWB and the CCP schemes would reduce to the 2-D IWB scheme for a solution that is constant along one coordinate direction. Similarly, the 3-D IISO and IISO2 schemes reduce to the isotropic nine-point 2-D scheme. In fact, all isotropic schemes in this 3-D family reduce to the single 2-D isotropic scheme, with $\alpha_{(1,0)} = 2/3$ (as one might expect).[48] It is also worth noting that the OCTA scheme reduces to the rotated five-point scheme in 2-D along the $\beta_z = 0$ plane.

On cutoff frequencies

Next we will discuss the cutoff frequencies found in this family of schemes. The *directional*, *minimum*, and *global* cutoff frequencies are defined as in the 2-D case (Eqs. (4.128) and (4.129)), and the inequalities (4.130) also hold for these 3-D schemes.

The minimum cutoff frequencies for the 3-D schemes within this family of 27-point schemes are given in Table 4.11.[49] It can be seen from the table that the 3-D IWB scheme ($\boldsymbol{\alpha} = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$) has the property that $\omega_{k,\mathrm{m}} = \pi$, provided that $\lambda = \lambda_{\max}$. However, as with the 2-D IWB scheme, it is important to remember that not all of the bandwidth from the IWB scheme is accurately reproduced; dispersion errors reach as much as 43% error ($(1 - 1/\sqrt{3})\%$) in the IWB scheme.[50]

It is worth recalling that for the CCP and OCTA special cases, the respective cutoff frequencies $\omega_{k,\mathrm{m}}$ are not found at the extremities of $\mathbb{B}$, as assumed in [136]. In fact, they are found at the wavenumber $\beta_h = \sqrt{3}\pi/2$ along the diagonal directions.[51]

Table 4.11: Cutoff frequencies for various $(K+1)$-point compact schemes up to $K \leq 26$

name	K	$\alpha_{(1,0,0)}$	$\alpha_{(1,1,0)}$	$\alpha_{(1,1,1)}$	λ (approx.)	$\omega_{k,\mathrm{m}}$	$\beta_{h,1\%}$	$\beta_{h,2\%}$
SLF	6	1	0	0	0.577	0.392π	0.190π	0.268π
CCP	12	0	1	0	1.000	0.667π	0.261π	0.358π
OCTA	8	0	0	1	1.000	0.500π	0.134π	0.189π
IWB	26	$1/4$	$1/2$	$1/4$	1.000	1.000π	0.269π	0.379π
IISO	18	$1/3$	$2/3$	0	0.866	0.667π	0.302π	0.414π
IISO2	26	$5/12$	$1/2$	$1/12$	0.866	0.667π	0.302π	0.414π
—	18	$129/400$	$271/400$	0	0.870	0.671π	0.303π	0.418π
—	26	$2/5$	$1/2$	$1/10$	0.877	0.681π	0.303π	0.427π

Comparing for wideband accuracy and relative computational efficiency

Here we compare special cases of interest within this family in terms of critical wavenumbers and relative computational efficiency. The one- and two-percent critical wavenumbers for these schemes are listed in Table 4.11; it can be seen that the isotropic schemes (and close alternatives) provide the highest (spatial) bandwidth under which one- and two-percent error criteria are satisfied.

In order to take into account the different optimal Courant numbers (and thus different operating sample rates for the same h)[52], we can look at the relative computational efficiencies (RCEs) defined

[48] However, it is not always the case that the 3-D schemes permit the same Courant numbers as the 2-D analogues, for there are many isotropic schemes for which $\lambda_{\max} \leq \sqrt{3/4}$.

[49] We leave out the global cutoff because $\omega_{k,\mathrm{g}} = \pi$ when $\lambda = \lambda_{\max}$.

[50] As in the 2-D IWB scheme, this special property is only possible by virtue of dispersion errors along non-axial directions.

[51] Consequently, the closed-form expressions for cutoff frequencies provided in [136, Eq. (25)] are not sufficient for the CCP, OCTA and IISO schemes.

[52] Equivalently, the following comparisons account for different grid densities when these schemes have the same sample rate.

by (4.39), using the simplest seven-point scheme as a reference. These RCEs were given in [136, Table II] considering 2%, 4% and 8% error criteria. Here, we consider a continuous range of error criteria from 0.1%–10%, as shown in Fig. 4.56. As reported in [136], we see that the IISO (or IISO2) scheme generally has the highest RCE for less than 4% dispersion error. The alternative choices, $\boldsymbol{\alpha} = (\frac{129}{400}, \frac{271}{400}, 0)$ and $\boldsymbol{\alpha} = (\frac{2}{5}, \frac{1}{2}, \frac{1}{10})$, are slightly more efficient than the IISO and IISO2 schemes for error thresholds greater than 1%. For greater than 6% error, the IWB scheme has the highest RCE, and generally becomes the most efficient choice up to 10% error. We see that the OCTA scheme is a poor choice relative to all other schemes. It is also important to point out that the displayed RCEs tend to be constant in the limit of small errors, which is a consequence of these schemes having the same (global) order of accuracy. We will revisit these comparisons when we reconsider the CCP and OCTA schemes on non-Cartesian grids (in Section 4.6), at which point their RCEs can increase by factors of two and four, respectively.

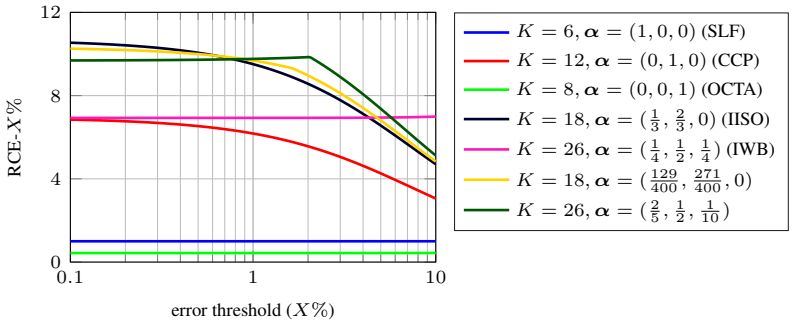
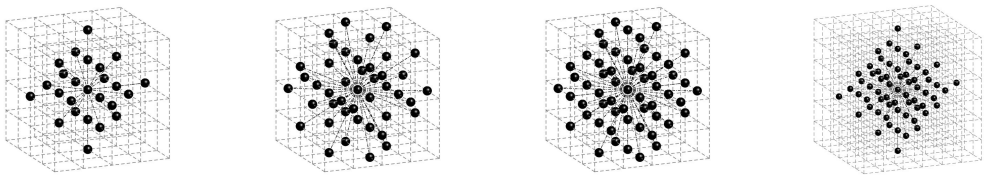


Figure 4.56: Relative computational efficiencies for various choices of $\boldsymbol{\alpha} = (\alpha_{(1,0,0)}, \alpha_{(1,1,0)}, \alpha_{(1,1,1)})$ for the family of 27-point compact explicit schemes on the Cartesian grid.

4.5.4 High-order and optimised $(K+1)$-point schemes with $K \leq 62$

In this section we investigate some high-order explicit schemes, employing stencils that reach out beyond the conventional first three compact shells, analogous to the wide stencils and schemes considered in 2-D. Among the many choices of possible stencils, the focus here will be on the four 3-D stencils shown in Fig. 4.57. These stencils involve neighbouring points that can be reached by at most three spatial steps along any combination of coordinate axes.



(a) $K = 24$ (b) $K = 42$ (c) $K = 56$ (d) $K = 62$

Figure 4.57: $(K+1)$-point 3-D stencils.

A family of discrete Laplacians using these stencils can be written as:

$$\delta_{\Delta,\Upsilon} = \alpha_{(1,0,0)}\delta_{\Delta,(1,0,0)} + \alpha_{(1,1,0)}\delta_{\Delta,(1,1,0)} + \alpha_{(1,1,1)}\delta_{\Delta,(1,1,1)} + \alpha_{(2,0,0)}\delta_{\Delta,(2,0,0)} + \alpha_{(2,1,0)}\delta_{\Delta,(2,1,0)} + \alpha_{(3,0,0)}\delta_{\Delta,(3,0,0)} \quad (4.228)$$

for which we have the 27-point family of discrete Laplacians as special cases, as well as the $2M$th-order accurate leggy discrete Laplacians for $M \leq 3$. For convenience we define $\boldsymbol{\alpha}$ here as:

$$\boldsymbol{\alpha} = (\alpha_{(1,0,0)}, \alpha_{(1,1,0)}, \alpha_{(1,1,1)}, \alpha_{(2,0,0)}, \alpha_{(2,1,0)}, \alpha_{(3,0,0)}) \tag{4.229}$$

and clearly, for consistency we require that $\boldsymbol{\alpha} \cdot (1,1,1,1,1,1) = 1$.

A two-step scheme for the 3-D wave equation follows from (4.20), with the discrete Laplacian $\delta_\Delta = \delta_{\Delta,\Upsilon}$ defined above (we have used the same notation as in the 2-D case, for brevity). The update equation for this scheme is:

$$\begin{aligned}\underline{u}^{n+1}_{i_x,i_y,i_z} = &\left(2 - \lambda^2 \left(6\alpha_{(1,0,0)} + 3\alpha_{(1,1,0)} + 2\alpha_{(1,1,1)} + \frac{3}{2}\alpha_{(2,0,0)} + \frac{6}{5}\alpha_{(2,1,0)} + \frac{2}{3}\alpha_{(3,0,0)}\right)\right) \underline{u}^{n}_{i_x,i_y,i_z} \\ &+ \lambda^2 \Big(\alpha_{(1,0,0)} Q^n_{\boldsymbol{i},(1,0,0)} + \frac{\alpha_{(1,1,0)}}{4} Q^n_{\boldsymbol{i},(1,1,0)} + \frac{\alpha_{(1,1,1)}}{4} Q^n_{\boldsymbol{i},(1,1,1)} \\ &+ \frac{\alpha_{(2,0,0)}}{4} Q^n_{\boldsymbol{i},(2,0,0)} + \frac{\alpha_{(2,1,0)}}{20} Q^n_{\boldsymbol{i},(2,1,0)} + \frac{\alpha_{(3,0,0)}}{9} Q^n_{\boldsymbol{i},(3,0,0)} \Big) - \underline{u}^{n-1}_{i_x,i_y,i_z}\end{aligned} \tag{4.230}$$

Fourth-order and sixth-order schemes

From a Taylor expansion of this parametrised discrete Laplacian, it is straightforward to arrive at the following constraint for fourth-order isotropy:

$$(2, -1, -4, 8, 2, 18) \cdot \boldsymbol{\alpha} = 0 \tag{4.231}$$

as well as the following *two* conditions required for sixth-order isotropy:

$$(-2, 3, 8, -32, -6, -162) \cdot \boldsymbol{\alpha} = 0\,, \qquad (0, 1, -4, 0, 4, 0) \cdot \boldsymbol{\alpha} = 0 \tag{4.232}$$

After imposing these constraints, the derivation of high orders of global accuracy using modified equation methods follows the two-dimensional case, so we can simply refer to Eqs. (4.145)–(4.153). Leaving λ as a free parameter, we need three degrees of freedom in $\boldsymbol{\alpha}$ to achieve a fourth-order accuracy, and *six* to achieve a sixth-order accuracy.[53] In principle, it is possible to achieve fourth-order accuracy with $K \in \{24, 42, 56\}$ and $K = 62$ could result in a sixth-order accuracy, but we must also check that stability can be ensured. Additionally, in the 57-point scheme we can impose a sixth-order isotropy.

Table 4.12: Parameters for high-order accurate schemes on 3-D Cartesian grid.

K	OoA	$\alpha_{(1,0,0)}$	$\alpha_{(1,1,0)}$	$\alpha_{(1,1,1)}$	$\alpha_{(2,0,0)}$	$\alpha_{(2,1,0)}$	$\alpha_{(3,0,0)}$	λ^2	λ (approx.)
24	$(4,4)$	1	$\frac{2}{9}$	0	$-\frac{2}{9}$	0	0	$\frac{1}{3}$	0.577
42	$(4,4)$	$\frac{3+5\sqrt{3}}{24}$	$\frac{2}{3}$	0	0	$\frac{5(1-\sqrt{3})}{24}$	0	$\frac{3-\sqrt{3}}{2}$	0.796
56	$(4,4)$	$\frac{10-\sqrt{11}}{30}$	$\frac{4\sqrt{11}}{15}$	$\frac{10-3\sqrt{11}}{30}$	$\frac{\sqrt{11}}{30}$	$\frac{2-\sqrt{11}}{6}$	0	$\frac{3(6-\sqrt{11})}{10}$	0.897
62	$(6,6)$	$\frac{181}{180}$	$\frac{44}{135}$	$\frac{1}{135}$	$-\frac{44}{135}$	$-\frac{2}{27}$	$\frac{11}{180}$	$\frac{1}{3}$	0.577

Table 4.12 lists the parameters found to provide high orders of accuracy, at maximal Courant num-

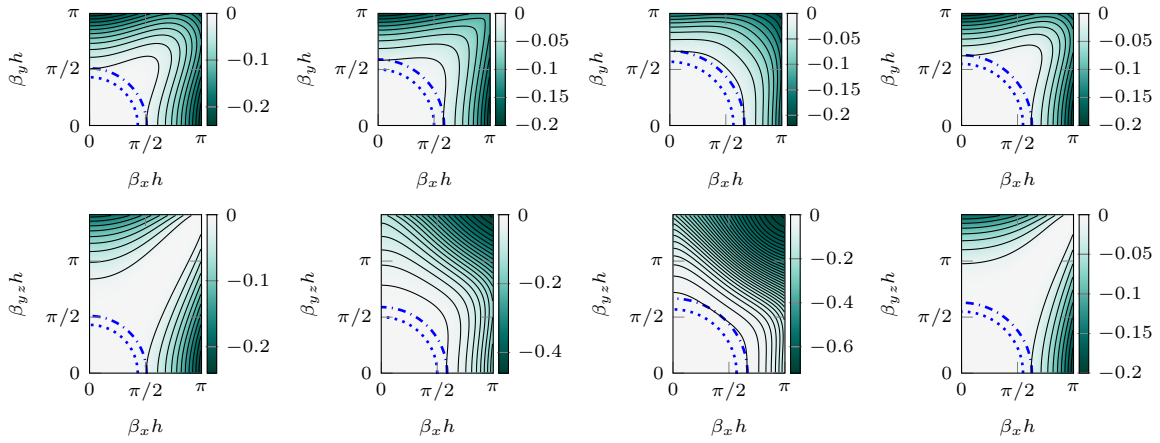
[53] As noted in [280], the isotropic special cases of 27-point compact explicit schemes have a sufficient number of free parameters to impose a fourth-order local truncation error, but λ would have to be fixed to an unstable value of $\lambda = 1$ ($\lambda_{\max} \leq \sqrt{3/4}$ for the isotropic schemes in that 27-point family).

bers for stability, which also happen to be optimal for numerical dispersion. The relative dispersion errors for these schemes are then displayed in Fig. 4.58. For $K = 24$ and $K = 62$, the stability condition is $\lambda \leq \sqrt{1/3}$, like the simplest 3-D scheme.[54] Also, like the simplest scheme in 3-D, those schemes demonstrate an exact wave speed along diagonal wave directions.

Remarkably, the 57-point scheme has the same Courant number as the 21-point scheme featured previously in 2-D, and the $\beta_z = 0$ slice of its dispersion relation appears identical to Fig. 4.19(c). In fact, it can be shown that this 57-point 3-D scheme reduces to the 21-point 2-D scheme for a solution that is constant along one coordinate direction, since this leads to the relations:

$$\alpha_{(1,0)} = \alpha_{(1,0,0)} + \alpha_{(1,1,0)}/2 + \alpha_{(2,1,0)}/10 \qquad \alpha_{(1,1)} = \alpha_{(1,1,0)}/2 + \alpha_{(1,1,1)}$$
$$\alpha_{(2,0)} = \alpha_{(2,0,0)}\,, \quad \alpha_{(2,1)} = \alpha_{(2,1,0)}/2 \qquad \alpha_{(3,0)} = \alpha_{(3,0,0)}$$

From the above relations, we also discover that the 25- and 63-point 3-D schemes reduce to the 13- and 25-point 2-D schemes, but for different $\lambda_{\max}$ in 2-D and 3-D.



(a) $K = 24$ (b) $K = 42$ (c) $K = 56$ (d) $K = 62$

Figure 4.58: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for $(K+1)$-point high-order accurate schemes in 3-D, for $\boldsymbol{\beta}_h \in [0, \pi]^3$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

Optimised schemes

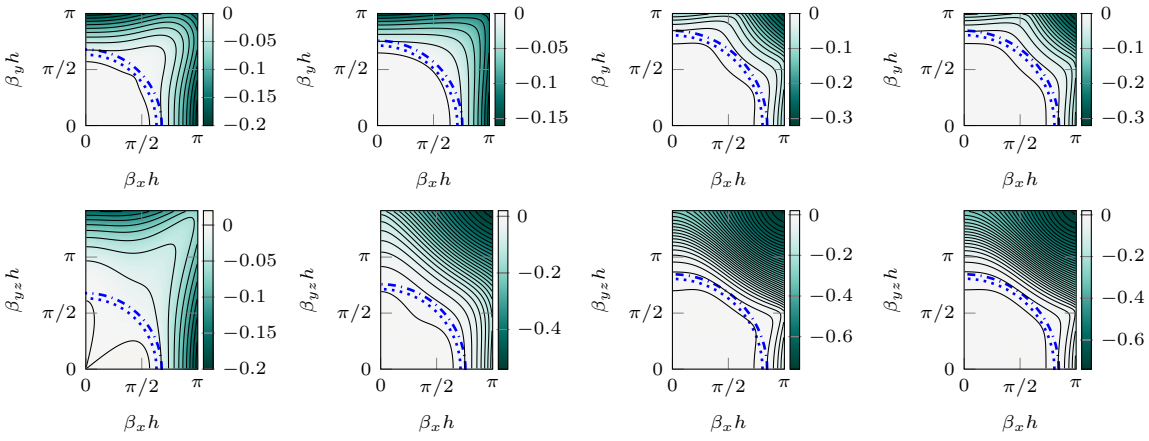
In this section, we trade off high-order accuracy in favour of a wideband accuracy. Disregarding conditions for high-order accuracy leaves us with as many as five free parameters to be optimised, keeping only the constraint for consistency (for a base second-order accuracy). As in [96] (work of the present author), these parameters have been optimised mainly using a standard implementation of the simplex method (`fminsearch` in MATLAB [164]) in order to maximise a one-percent critical wavenumber. The results of these optimisations are shown in Table 4.13, where all values have been rounded to three decimal places, for brevity (the Courant numbers are truncated). The relative dispersion errors for these

[54] This may seem remarkable at first, but it follows the 2-D case in Table 4.4, and there is indeed a pattern. A general expression for these $2M$th-order accurate schemes, using stencils that are compact in the ℓ_1-norm (in any number of dimensions) was found by Tuomela [282]. Such $(K+1)$-point stencils have $K = (2M^2 + 3M + 4)(2M/3)$ in 3-D, $K = 2M^2 + 2M$ in 2-D, and $K = 2M$ in 1-D. The 3-D cases with $K = 42$ and $K = 56$, which are not compact in the ℓ_1 sense, have not appeared in the literature (to our knowledge).

optimised schemes are displayed in Fig. 4.59. In comparison to Fig. 4.58, we observe improvements in the one- and two-percent critical wavenumbers for each respective stencil size K.

Table 4.13: Parameters for second-order accurate optimised 3-D Cartesian schemes

K	$\alpha_{(1,0,0)}$	$\alpha_{(1,1,0)}$	$\alpha_{(1,1,1)}$	$\alpha_{(2,0,0)}$	$\alpha_{(2,1,0)}$	$\alpha_{(3,0,0)}$	λ (approx.)
24	0.899	0.415	0	−0.314	0	0	0.608
42	0.423	0.824	0	0	−0.247	0	0.797
56	0.190	0.903	0.063	0.194	−0.350	0	0.925
62	0.194	0.859	0.083	0.196	−0.312	−0.020	0.933



(a) $K = 24$ (b) $K = 42$ (c) $K = 56$ (d) $K = 62$

Figure 4.59: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for $(K+1)$-point optimised wide-stencil schemes in 3-D, for $\boldsymbol{\beta}_h \in [0, \pi]^3$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$.

4.5.5 Comparing schemes for wideband accuracy

At this point it is worth comparing the various 3-D schemes we have seen so far, in terms of their critical wavenumbers, and also considering normalisations for computational costs, as we did for 2-D Cartesian schemes in Section 4.3.7. In Table 4.14, we list the one- and two-percent critical wavenumbers for a selection of the schemes presented thus far, along with their associated Courant numbers. For these comparisons, schemes are classified into four families, analogous to those defined for the 2-D case (in Table 4.6).[55]

The one-percent critical wavenumbers $\beta_{h,1\%}$ (normalised for memory costs) are plotted in Fig. 4.60 as a function of stencil size. Similarly, the critical wavenumbers $\beta_{h^*,1\%}$ (normalised for spatiotemporal densities) are plotted in Fig. 4.61 as a function of stencil size. We see from Fig. 4.60 that the optimised leggy schemes tend to give the highest one-percent critical wavenumbers when memory costs are fixed, followed closely by the optimised compact schemes, for $K \geq 24$.[56] In Fig. 4.61, we see that when

[55] As there are no leggy schemes with $K \in \{56, 62\}$, we have chosen $M \in \{9, 10\} \Rightarrow K \in \{54, 60\}$ for the two largest-stencil leggy schemes in this comparison.

[56] This contrasts with what was seen for 2-D schemes in Section 4.3.7, since the compact schemes tended to produce better results than the leggy schemes for larger stencil sizes.

Table 4.14: Critical wavenumbers, $\beta_{h,1\%}$ and $\beta_{h,2\%}$, for various 3-D Cartesian grid–based schemes with $(K+1)$-point stencils

K	family	OoA	λ (approx.)	$\beta_{h,1\%}$	$\beta_{h,2\%}$
6	—	$(2,2)$	0.577	0.191π	0.268π
12	leggy, $\lambda=\lambda_{\max,M}$	$(2,4)$	0.500	0.317π	0.456π
18	leggy, $\lambda=\lambda_{\max,M}$	$(2,6)$	0.470	0.328π	0.461π
24	leggy, $\lambda=\lambda_{\max,M}$	$(2,8)$	0.453	0.340π	0.475π
42	leggy, $\lambda=\lambda_{\max,M}$	$(2,14)$	0.428	0.359π	0.501π
54	leggy, $\lambda=\lambda_{\max,M}$	$(2,18)$	0.420	0.366π	0.512π
60	leggy, $\lambda=\lambda_{\max,M}$	$(2,20)$	0.417	0.369π	0.515π
12	leggy, $\lambda=\lambda_{1\%,M}$	$(2,4)$	0.365	0.468π	0.533π
18	leggy, $\lambda=\lambda_{1\%,M}$	$(2,6)$	0.277	0.584π	0.640π
24	leggy, $\lambda=\lambda_{1\%,M}$	$(2,8)$	0.240	0.655π	0.705π
42	leggy, $\lambda=\lambda_{1\%,M}$	$(2,14)$	0.201	0.766π	0.808π
54	leggy, $\lambda=\lambda_{1\%,M}$	$(2,18)$	0.179	0.819π	0.854π
60	leggy, $\lambda=\lambda_{1\%,M}$	$(2,20)$	0.174	0.831π	0.864π
12	compact, constrained	$(2,2)$	1.000	0.262π	0.359π
18	compact, constrained	$(2,2)$	0.866	0.302π	0.414π
24	compact, constrained	$(4,4)$	0.577	0.424π	0.508π
42	compact, constrained	$(4,4)$	0.796	0.492π	0.582π
56	compact, constrained	$(4,4)$	0.897	0.566π	0.662π
62	compact, constrained	$(6,6)$	0.577	0.546π	0.622π
18	compact, optimised	$(2,2)$	0.870	0.303π	0.418π
24	compact, optimised	$(2,2)$	0.608	0.627π	0.674π
42	compact, optimised	$(2,2)$	0.797	0.710π	0.752π
56	compact, optimised	$(2,2)$	0.925	0.803π	0.846π
62	compact, optimised	$(2,2)$	0.933	0.797π	0.840π

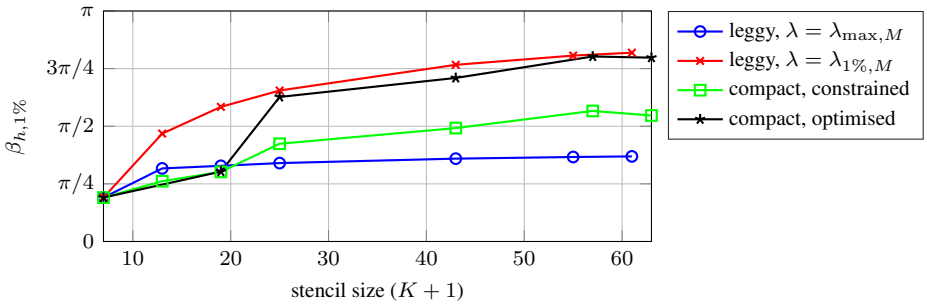


Figure 4.60: Critical wavenumber $\beta_{h,1\%}$ (same memory costs) as a function of stencil size, for various families of 3-D Cartesian grid–based schemes.

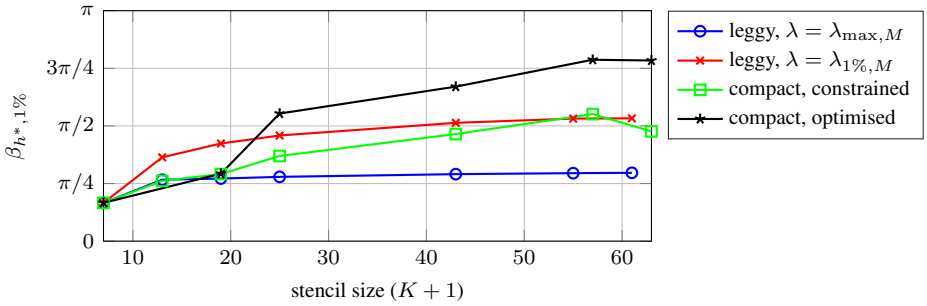


Figure 4.61: Critical wavenumber $\beta_{h^\star,1\%}$ (same spatiotemporal densities) as a function of stencil size, for various families of 3-D Cartesian grid–based schemes.

spatiotemporal densities are normalised, the compact schemes offer better wideband accuracy than the optimised leggy schemes when $K \geq 24$, since the leggy schemes must compensate for using higher sample rates for the same h.

Next we compare these schemes for their one-percent relative computational efficiencies, taking into account the number of operations (FLOP instructions) required by each pointwise update. These comparisons are illustrated in in Figs. 4.62 and 4.63, from which we find that some of the optimised schemes are nearly 500× more efficient than the simplest schemes, in terms of pointwise updates, or nearly 70× more efficient in total operation counts. In terms of the RCE-1%-opts metric (defined by (4.40)), the 57-point optimised compact scheme appears to be the most efficient choice. It is important to remember, however, that actual compute times (in practice) will not necessarily scale with these relative efficiencies, as memory caching effects must also be considered (see, e.g., related work from the current author [103]).

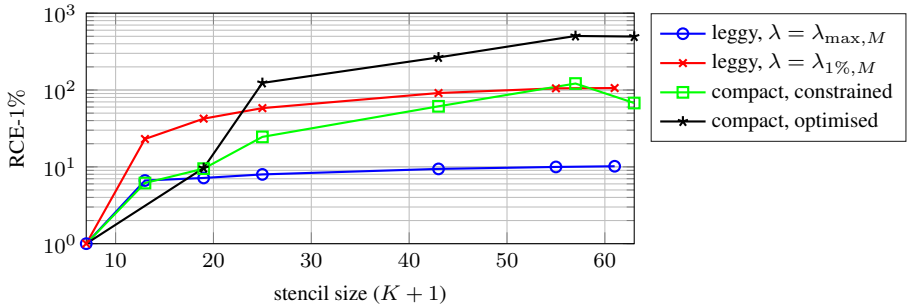


Figure 4.62: Relative computational efficiencies for one-percent error threshold (with simplest scheme as reference), as a function of stencil size, for various families of schemes.

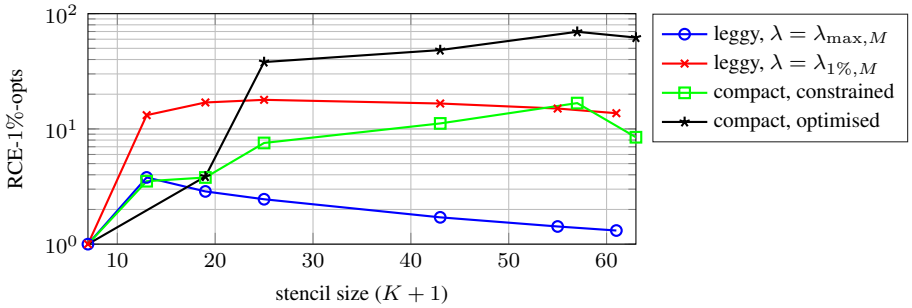


Figure 4.63: Relative computational efficiencies for one-percent error threshold, scaled by $(8/(K+2))$ (representing required FLOPS instructions or memory reads, relative to simplest scheme), as a function of stencil size $(K+1)$, for various families of schemes.

4.5.6 Numerical example

In this section we compare the various 3-D schemes using a numerical example, analogous to what was considered in Section 4.3.8. The initial conditions considered here are expressed by (4.155) (with $\mathbf{x} \in \mathbb{R}^3$), but in this case we use $c = 340\,\text{m/s}$ and $\sigma^2 = 0.03\,\text{m}^2$.

As in the 2-D case, we consider two scenarios: using $h = \mu^{1/3} h^\star$ with $h^\star = 14\,\text{cm}$ for all schemes (normalised for memory costs), and using $h^* = 14\,\text{cm}$ fixed across schemes and $h = (\mu/\lambda)^{1/4} h^*$ (i.e., normalised for spatiotemporal densities), and we recall that $\mu = 1$ for the Cartesian grid. The initial two states for each scheme are computed with the initialisation scheme (4.156). The simulations are run for a short duration (5 ms of simulation time) with outputs taken at a distance of $\sqrt{3}\,\text{m}$, along the axial,

side-diagonal, and diagonal directions (the extreme cases).[57] In order to obtain an "exact" solution to compare against, we use the sixth-order scheme with $h = 5\,\text{cm}$, such that dispersion errors are virtually absent (for this set of initial conditions). The resulting outputs from the various schemes are displayed in Figs. 4.64 and 4.66, along with the errors in Figs. 4.65 and 4.67.

Discussion

We start by discussing the case of fixed memory costs ($h = \mu^{1/3} h^\star$), as shown in Fig. 4.64 (error signals in Fig. 4.65). The simple 7-point scheme (SLF) provides relatively poor results for this choice of grid spacing and initial conditions. The OCTA scheme is worse than the SLF scheme (as expected by the analysis of relative computational efficiencies), but this can be improved by considering a more appropriate setting on the BCC grid, as will be seen in the next section. Apart from the leggy schemes with $\lambda = \lambda_{\max,M}$, there is an improvement with increasing stencil size within each family. The best results for fixed memory costs are generally obtained with $K \geq 24$ with the *optimised* leggy schemes ($\lambda < \lambda_{\max,M}$), and with the compact schemes (both high-order and optimised).

Next we discuss the scenario of fixed spatiotemporal densities (the grid spacing chosen as $h = (\mu/\lambda)^{1/4} h^*$), as shown in Figs. 4.66 and 4.65. Recall, in this case, schemes of the same stencil size can be computed using the same number of FLOP instructions. In comparison to the previous scenario, the leggy schemes suffer in performance as they must compensate for using small Courant numbers (and thus high sample rates for a fixed grid spacing) by using coarser grids; this is similar what was seen for the 2-D leggy schemes. Between the compact and the leggy schemes, we start to see a difference in results for $K \geq 24$, which is also when we can achieve higher-order accuracy in both space and time for these explicit schemes. In comparison to the scenario of fixed memory costs, results are relatively unchanged for the compact schemes. The second-order accurate optimised compact schemes tend to give better results than the high-order accurate compact schemes in this case.

[57] For the outputs, cubic spline interpolation is employed to show detail in the temporal plots, but this non-ideal interpolation does not add noticeable artefacts.

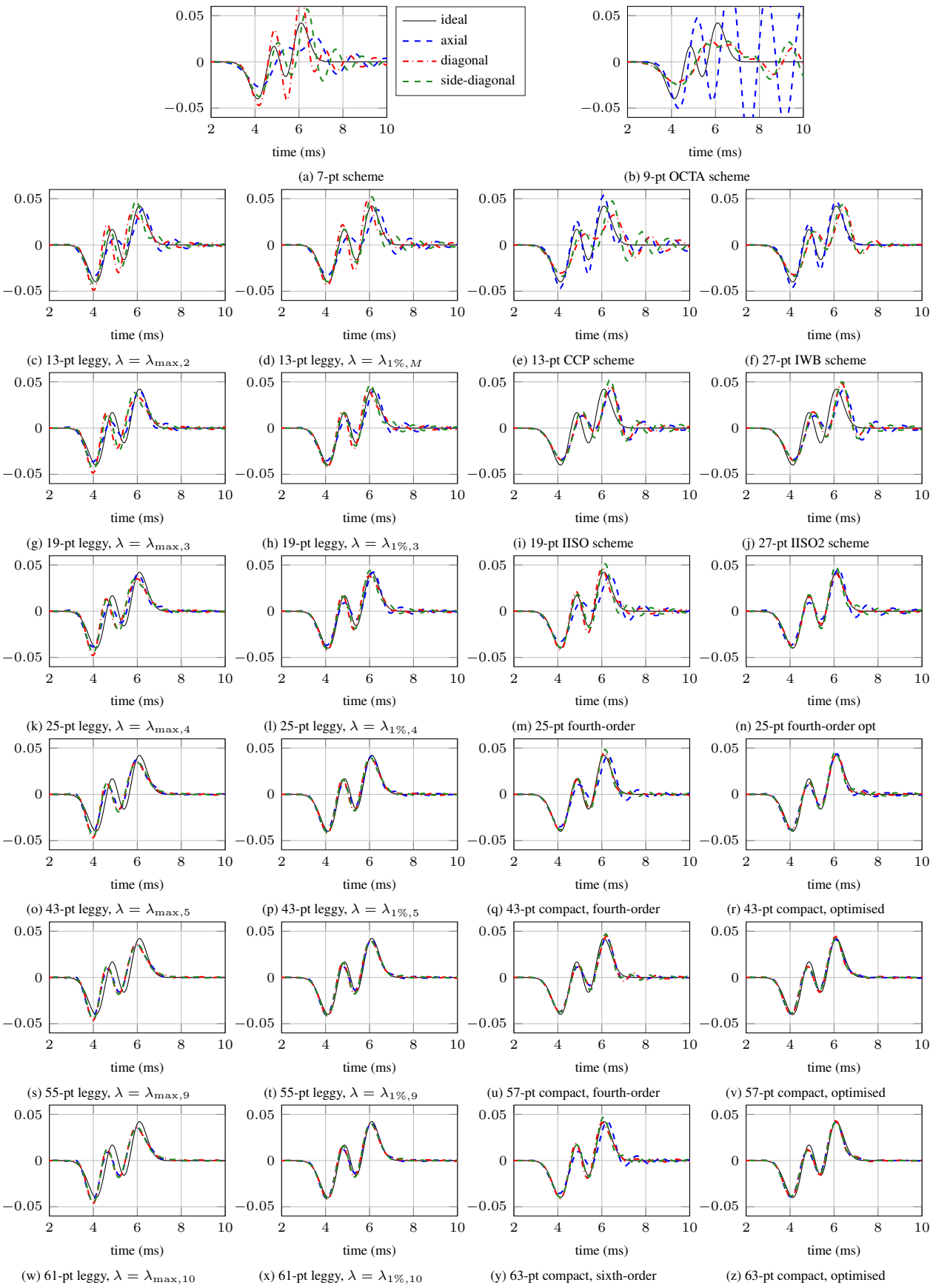


Figure 4.64: Time-domain plots of outputs at distance $\sqrt{3}$ m from origin, along axial, side-diagonal, and diagonal directions. Here, the various schemes have the same grid spacing (fixed memory costs).

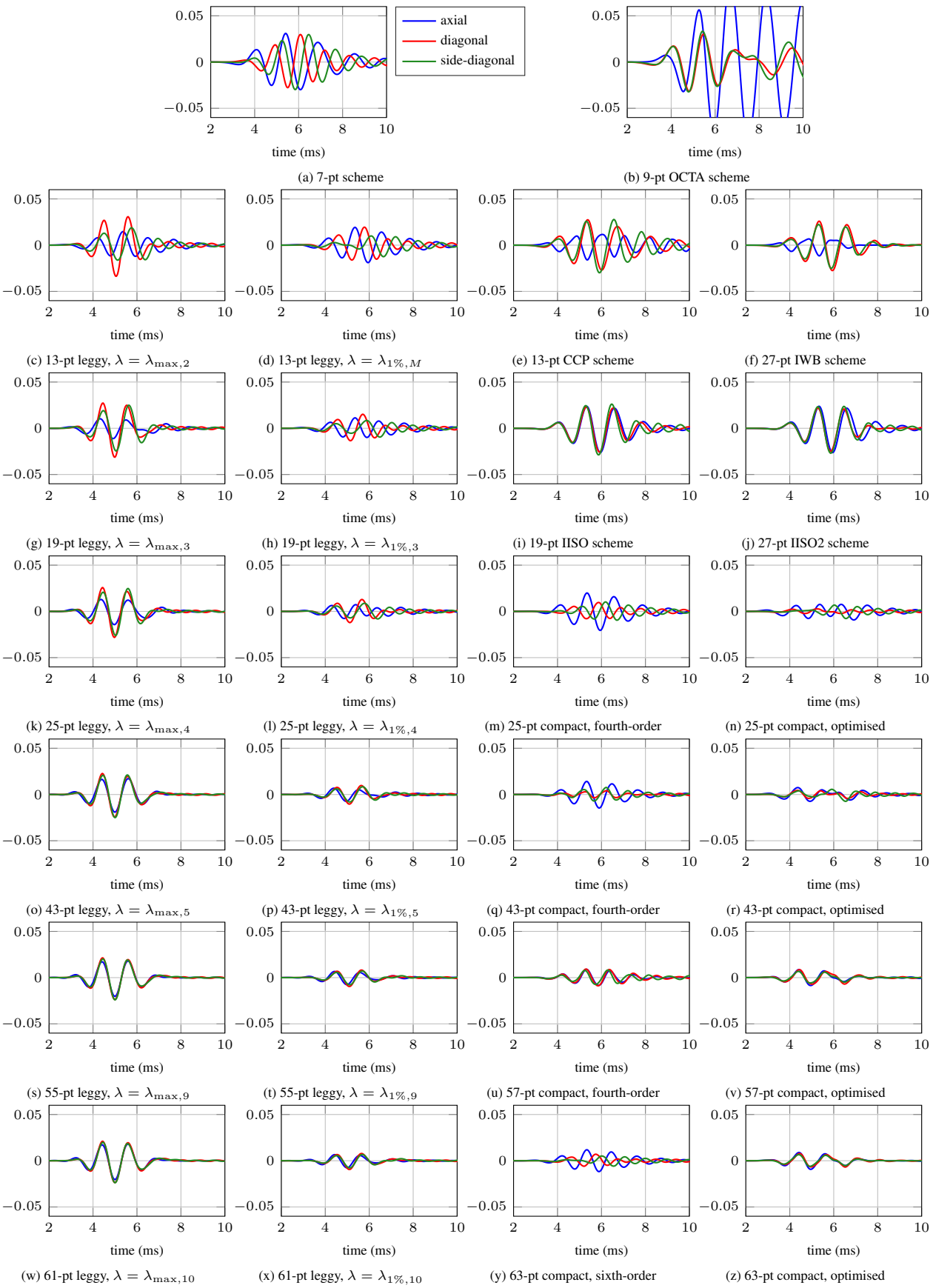


Figure 4.65: Plots of error signals from Fig. 4.64 (fixed memory costs).

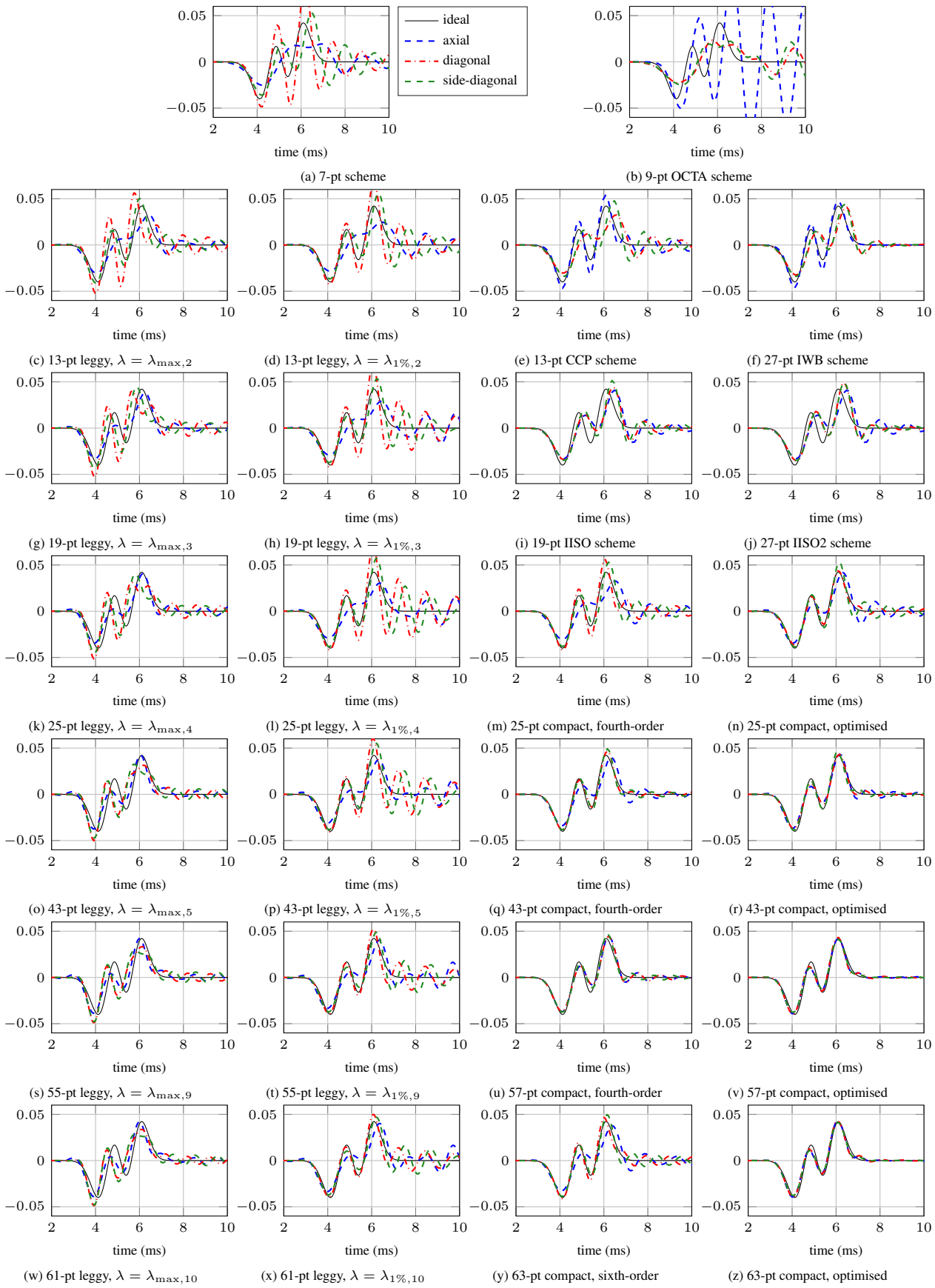


Figure 4.66: Time-domain plots of outputs at distance $\sqrt{3}$ m from origin, along axial, side-diagonal, and diagonal directions. Here, the various schemes are normalised for spatiotemporal densities, which also means similar operation counts for similar stencil sizes.

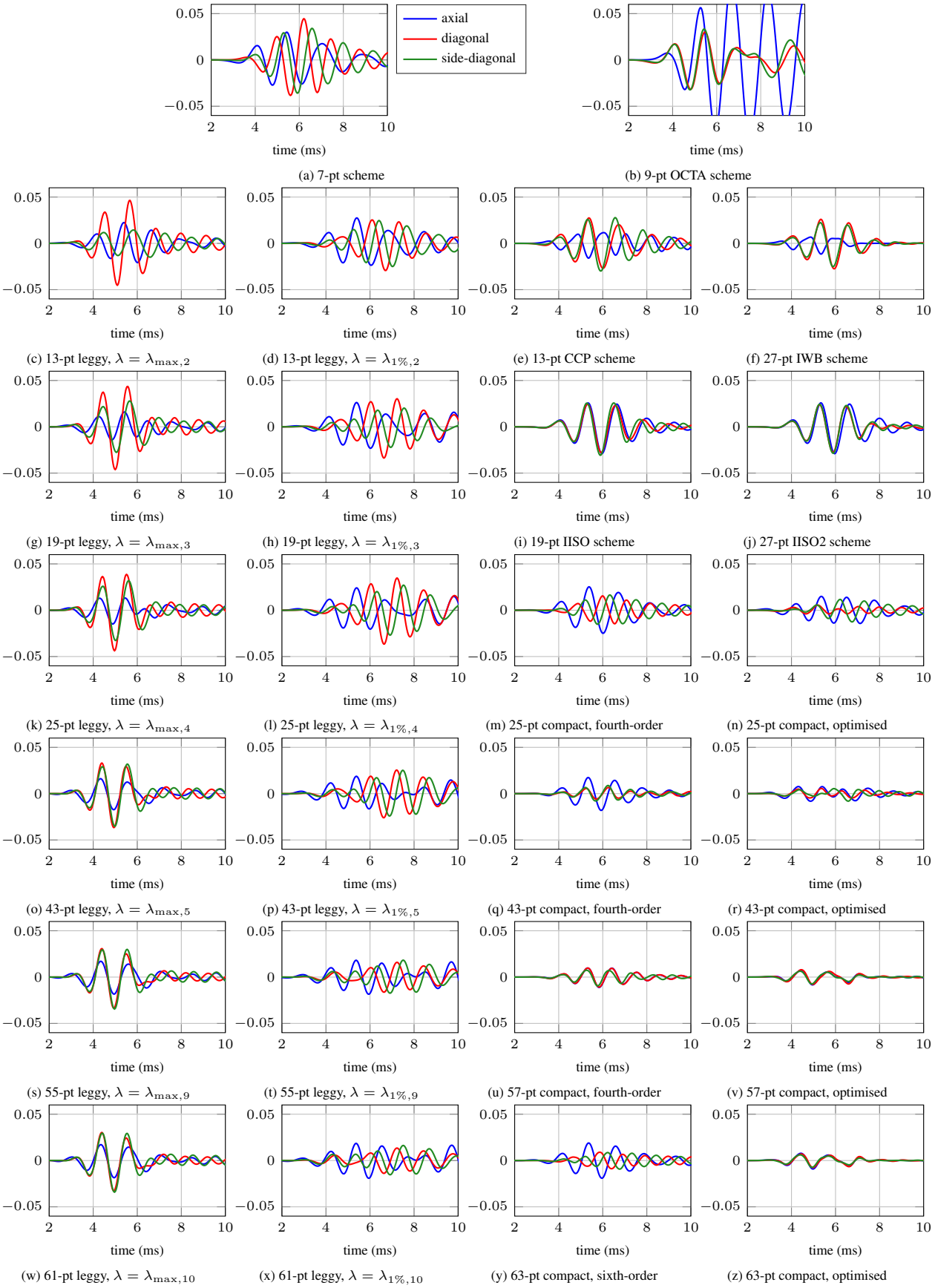


Figure 4.67: Plots of error signals from Fig. 4.66 (normalised for spatiotemporal densities).

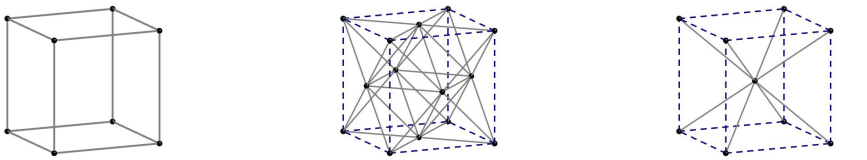
4.6 Non-Cartesian schemes for the 3-D wave equation

In this section we consider finite difference schemes on non-Cartesian (lattice-based) grids in 3-D, namely the face-centered cubic (FCC) and body-centered cubic (BCC) grids. Both of these lattices can be seen as 3-D analogues to the 2-D hexagonal grid, as will be explained shortly, and thus we might expect them to perform better than Cartesian schemes of the same stencil size (as was seen in 2-D with the hexagonal schemes). We should note that apart from the simplest schemes on these grids [144, 1, 21, 196], the non-Cartesian schemes to be considered in this section have not (to our knowledge) appeared in the literature.

FCC and BCC grids can be found as subgrids on the Cartesian grid (as will be illustrated shortly), and as such it will be convenient to use the same grid function $\underline{u}^n_{\boldsymbol{i}}$, but $\boldsymbol{i}$ will be restricted to a subset of $\mathbb{Z}^3$ that makes up the non-Cartesian grid of interest. It should be noted that as a consequence of this choice of notation, h will *not* represent the respective grid spacings on the FCC and BCC grids (i.e., the *inter-nodal* distance), but rather the grid spacing h on the overarching Cartesian grid; this will be important to remember (as such, it will be repeated later). The grid spacings for the FCC and BCC grids will be represented instead by h', and the relationship between h' and h will be specified shortly.

4.6.1 Non-Cartesian grids and sampling considerations

Apart from the cubic *lattice* ($\mathbb{Z}^3$), from which we derive the Cartesian *grid*,[58] there are two other lattices to consider in 3-D: the face-centered cubic (FCC) lattice and the body-centered cubic (BCC) lattice. In the crystallography literature [104], these lattices are often illustrated by their respective cubic "unit cells", which can be used to tile space and populate the respective lattices. The traditional unit cells for these lattices are shown in Fig. 4.68. The names "FCC" and "BCC" derive from the fact that there are lattice points at cubic *face centers* in the FCC lattice's unit cell, while there is a lattice point at the center of the cube (the *body center*) for the unit cell of the BCC lattice, as seen below.

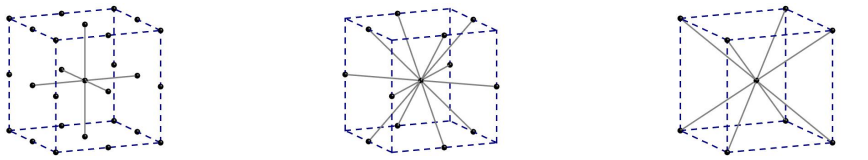


(a) simple cubic (b) face-centered cubic (c) body-centered cubic

Figure 4.68: Unit cells of simple cubic, face-centered cubic (FCC), and body-centered cubic (BCC) lattices. The grey lines show adjacencies between nearest-neighbouring lattice points.

An important characteristic of each of these lattices is its "kissing number", which is the number of non-overlapping spheres that can touch, or "kiss", a sphere of the same size centered at the origin on the lattice [58]. The cubic lattice has a kissing number of six, and the FCC and BCC lattices have kissing numbers of twelve and eight respectively [58]. Consequently, the kissing number is also the number of nearest-neighbouring points of any given point on the lattice. These nearest-neighbour adjacencies are illustrated in the unit cells (alternative to Fig. 4.68 for cubic and FCC) in Fig. 4.69. These arrangements of points are essentially simple $(K+1)$-point stencils to be used for discrete Laplacians, where K is the kissing number of the lattice. Indeed, these stencils pertain to $\delta_{\Delta,(1,0,0)}$, $\delta_{\Delta,(1,1,0)}$, and $\delta_{\Delta,(1,1,1)}$, as seen previously (in Fig. 4.47).

[58] We use the terms "grid" and "lattice" interchangeably here, but strictly speaking, a lattice can be defined as a set of vectors that are closed under addition, and thus a lattice must contain the origin [58].



(a) simple cubic (b) face-centered cubic (c) body-centered cubic

Figure 4.69: Illustrations of nearest-neighbouring points to the origin in simple cubic, face-centered cubic (FCC), and body-centered cubic (BCC) lattices.

There are multiple ways to define, or construct, the FCC and BCC grids, including the above-mentioned method of tiling space with the traditional cubic unit cells. For example, one can define the FCC grid as a hyperslice of the four-dimensional integer lattice [58], and subsequently one can define the BCC grid as four shifted FCC lattices [58]. This construction is one way to relate the FCC and BCC grids to the 2-D hexagonal grid. Specifically, this construction comes from the A_d root lattice and its dual [58]. In other words, the FCC lattice is the A_3 root lattice and its dual is the BCC lattice [58]. The hexagonal lattice is the A_2 root lattice, but it is self-dual [58]—this is why both FCC and BCC grids are analogous to the hexagonal grid.[59]

Another, perhaps more practical, construction is to view the FCC and BCC lattices as subsets of a Cartesian grid.[60] In a similar fashion, we will define the FCC and BCC grids as "subgrids" on the Cartesian grid using the Cartesian grid spacing h. For this, we first define *index sets* for FCC and BCC grids in the following manner:

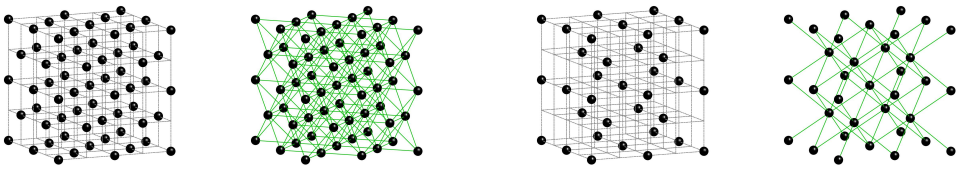
$$\mathcal{I}^{(\mathrm{f})} = \{\boldsymbol{i} \in \mathbb{Z}^3 : (i_x + i_y + i_z) \bmod 2 = 0\} \tag{4.233a}$$

$$\mathcal{I}^{(\mathrm{b})} = \{\boldsymbol{i} \in \mathbb{Z}^3 : (i_x \bmod 2) = (i_y \bmod 2) = (i_z \bmod 2)\} \tag{4.233b}$$

Then we define the FCC and BCC grids respectively as follows:

$$\mathbb{G}_h^{(\mathrm{f})} = \left\{\boldsymbol{i}h \in \mathbb{R}^3 : \boldsymbol{i} \in \mathcal{I}^{(\mathrm{f})}\right\}, \quad \mathbb{G}_h^{(\mathrm{b})} = \left\{\boldsymbol{i}h \in \mathbb{R}^3 : \boldsymbol{i} \in \mathcal{I}^{(\mathrm{b})}\right\} \tag{4.234}$$

These grid constructions are illustrated in Fig. 4.70. It is important to point out (again) that the Cartesian grid spacing h is *not* equal to the distance between nearest-neighbouring points on the FCC and BCC subgrids. The correct inter-nodal distances, under these constructions, are $h' = \sqrt{2}h$ and $h' = \sqrt{3}h$, respectively on the FCC and BCC grids.



(a) FCC (b) BCC

Figure 4.70: Some points in the FCC and BCC lattices as subsets of the cubic lattice, and on right in (a) and (b), adjacent nearest-neighbours are connected by green lines. The origin is the black point at the center of each figure.

[59] See [98] for more discussion on these relationships.

[60] This construction comes from the D_d root lattice, which includes the Cartesian grid in 2-D (but not the 2-D hexagonal lattice) [58].

Any regular lattice can be defined by a set of basis vectors [104], and accordingly, we could also define these FCC and BCC grids respectively as:

$$\mathbb{G}_h^{(\mathrm{f})} = \{\mathbf{V}_\mathrm{f}\boldsymbol{m}h : \boldsymbol{m} \in \mathbb{Z}^3\}\,, \quad \mathbb{G}_h^{(\mathrm{b})} = \{\mathbf{V}_\mathrm{b}\boldsymbol{m}h : \boldsymbol{m} \in \mathbb{Z}^3\} \tag{4.235}$$

where $\mathbf{V}_\mathrm{f}$ and $\mathbf{V}_\mathrm{b}$ are *generator matrices* of the lattices:

$$\mathbf{V}_\mathrm{f} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad \mathbf{V}_\mathrm{b} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \tag{4.236}$$

and $\boldsymbol{m}$ is taken to be a column-vector.

As mentioned previously for use in hexagonal schemes, multidimensional sampling in the frequency domain is characterised by the wavenumber cell of the lattice [191] (the first Brillouin zone [104]), and the wavenumber cell of the chosen sampling lattice is the Voronoi cell of its *reciprocal* or *dual* lattice. In general, the reciprocal of a lattice is generated by the inverse of the generator matrix, and scaled appropriately for the frequency quantity of interest; here, the quantity of interest is $\boldsymbol{\beta}$ in rad/m, so we further scale by $2\pi/h$. The cubic lattice is treated in a simple manner because its generator matrix is the identity, and is thus self-dual. As for the FCC and BCC grids, we respectively define their dual grids as:

$$\widehat{\mathbb{G}_h^{(\mathrm{f})}} = \left\{(2\pi/h)\mathbf{V}_\mathrm{f}^{-1}\boldsymbol{m} \in \mathbb{R}^3 : \boldsymbol{m} \in \mathbb{Z}^3\right\} \tag{4.237a}$$

$$\widehat{\mathbb{G}_h^{(\mathrm{b})}} = \left\{(2\pi/h)\mathbf{V}_\mathrm{b}^{-1}\boldsymbol{m} \in \mathbb{R}^3 : \boldsymbol{m} \in \mathbb{Z}^3\right\} \tag{4.237b}$$

The FCC and BCC grids are reciprocal (respective duals), since $2\mathbf{V}_\mathrm{f}^{-1} = \mathbf{V}_\mathrm{b}$. We could also define these dual grids as:

$$\widehat{\mathbb{G}_h^{(\mathrm{f})}} = \left\{(\pi/h)\boldsymbol{m} \in \mathbb{R}^3 : \boldsymbol{m} \in \mathcal{I}^{(\mathrm{b})}\right\}, \quad \widehat{\mathbb{G}_h^{(\mathrm{b})}} = \left\{(\pi/h)\boldsymbol{m} \in \mathbb{R}^3 : \boldsymbol{m} \in \mathcal{I}^{(\mathrm{f})}\right\} \tag{4.238}$$

By the respective Voronoi cells of the FCC and BCC grids, the wavenumber cell of the FCC lattice is a *truncated octahedron* (fourteen-sided), and the wavenumber cell of the BCC lattice is a *rhombic dodecahedron* (twelve-sided) [58]. We denote these wavenumber cells as $\mathbb{B}_\mathrm{f}$ and $\mathbb{B}_\mathrm{b}$, respectively, and we define them by the following systems of linear inequalities:

$$\mathbb{B}_\mathrm{f} = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{A}_\mathrm{f}\mathbf{x}| \leq \pi\mathbf{b}_\mathrm{f}\} \tag{4.239a}$$

$$\mathbb{B}_\mathrm{b} = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{A}_\mathrm{b}\mathbf{x}| \leq \pi\mathbf{b}_\mathrm{b}\} \tag{4.239b}$$

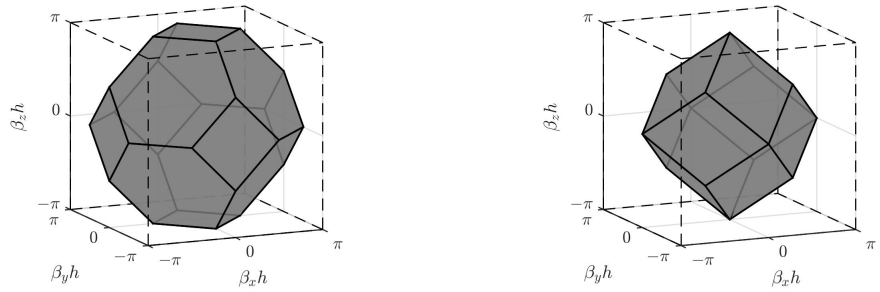
where

$$\mathbf{A}_\mathrm{f} = \begin{bmatrix} 1 & 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & -1 \end{bmatrix}^\mathrm{T}, \qquad \mathbf{A}_\mathrm{b} = \begin{bmatrix} 1 & 0 & 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & 0 & 1 & -1 \end{bmatrix}^\mathrm{T} \tag{4.240a}$$

$$\mathbf{b}_\mathrm{f} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & 1 & 3/2 & 3/2 & 3/2 & 3/2 \end{bmatrix}^\mathrm{T}, \quad \mathbf{b}_\mathrm{b} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^\mathrm{T} \tag{4.240b}$$

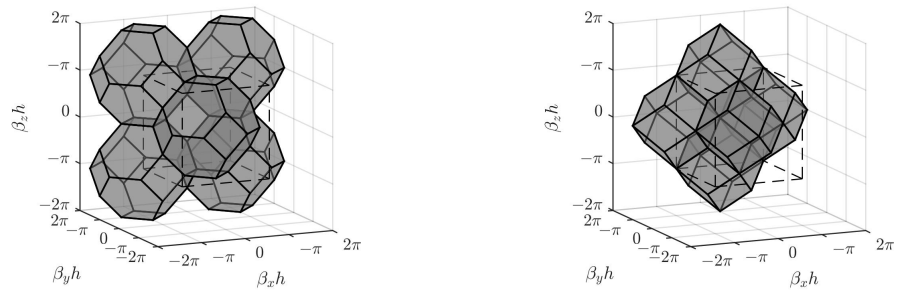
These FCC and BCC wavenumber cells are illustrated in Fig. 4.71, inscribed within the usual cubic

wavenumber cell $\mathbb{B} = [-\pi, \pi]^3$ in the normalised wavenumber domain $\boldsymbol{\beta}_h \in \mathbb{R}^3$, where h is the Cartesian grid spacing. It can be seen that these wavenumber cells only cover a portion of the cubic wavenumber cell $\mathbb{B}$, but it is important to remember that the way these grids have been defined, their grid densities are only a fraction of that of the Cartesian grid. The spatial densities, along with other measures of interest for these grids are listed in Table 4.15. These include the *inradius*, *circumradius*, and volume of each respective wavenumber cell. The circumradius is the largest wavenumber that can be discretised on the grid (only in specific directions), and the inradius of the wavenumber cell determines the largest wavenumber that can be discretised irrespective of wave direction. Fig. 4.72 illustrates how these wavenumber cells tile the wavenumber space. These tilings (in the spatial domain) will be revisited in Chapter 6.



(a) FCC, $\mathbb{B}_\text{f}$ (b) BCC, $\mathbb{B}_\text{b}$

Figure 4.71: Wavenumber cell of FCC and BCC lattices for $\boldsymbol{\beta}_h$ with Cartesian grid spacing h.



(a) FCC, $\mathbb{B}_\text{f}$ (b) BCC, $\mathbb{B}_\text{b}$

Figure 4.72: Wavenumber cells of FCC and BCC lattices, each with four neighbouring cells. The Cartesian wavenumber cell is superimposed, in dashed lines.

Also found in Table 4.15 is the *sampling efficiency* of each grid, which is defined as the volume of the wavenumber cell divided by the volume of the largest sphere inscribed in the wavenumber cell [191]. This is essentially a measure of the relative portion of the wavenumber cell can be used for isotropic sampling. It turns out that the FCC and BCC grids are more efficient for 3-D sampling than the Cartesian grid, with the BCC grid being the optimal one [191]. More specifically, the FCC and BCC grids are respectively 30% and 41% more efficient than the Cartesian grid for 3-D sampling, which also means that on FCC and BCC grids, one can sample an isotropic signal using 23% and 29% fewer samples, respectively, than what is required on a Cartesian grid.

Table 4.15: Some measures for grids and associated 3-D normalised wavenumber cells, in terms of Cartesian grid spacing h, and where h' is the inter-nodal distance on the grid (the distance between nearest-neighbouring grid points).

grid	h'	density of grid	wavenumber cell	inradius	circumradius	volume	sampling eff.
cubic	h	h^{-3}	$\mathbb{B}_{\mathrm{c}}$ (cube)	π	$\sqrt{3}\pi$	$8\pi^3$	52.3%
FCC	$\sqrt{2}h$	$(1/2)h^{-3}$	$\mathbb{B}_{\mathrm{f}}$ (truncated octahedron)	$\sqrt{3/4}\pi$	$\sqrt{5/4}\pi$	$4\pi^3$	68.0%
BCC	$\sqrt{3}h$	$(1/4)h^{-3}$	$\mathbb{B}_{\mathrm{b}}$ (rhombic dodecahedron)	$\sqrt{1/2}\pi$	π	$2\pi^3$	74.0%

We can also associate temporal sampling constraints to the spatial sampling grids, via the relation $\omega_{h/c} = \pm|\boldsymbol{\beta}_h|$. If the spatial bandwidth is limited to $\beta_{h,\max}$, which is between the inradius and circumradius of the wavenumber cell, then we have (4.8) as a constraint on k and λ. If we choose $\beta_{h,\max} = \beta_{\max} h$ as the inradii of the wavenumber cells (as we should for an isotropic sampling), then we should choose h (the Cartesian grid spacing) and k as:

$$\text{FCC:} \quad h = \sqrt{\frac{3}{4}} \left(\frac{\pi}{\beta_{\max}} \right) \quad \Longrightarrow \quad k = \sqrt{\frac{3}{4}} \left(\frac{\lambda}{2 f_{\max}} \right) \tag{4.241a}$$

$$\text{BCC:} \quad h = \frac{1}{\sqrt{2}} \left(\frac{\pi}{\beta_{\max}} \right) \quad \Longrightarrow \quad k = \frac{1}{\sqrt{2}} \left(\frac{\lambda}{2 f_{\max}} \right) \tag{4.241b}$$

where $f_{\max} = \frac{c\beta_{\max}}{2\pi}$. As such, we end up with the following optimal Courant numbers for spatiotemporal sampling:

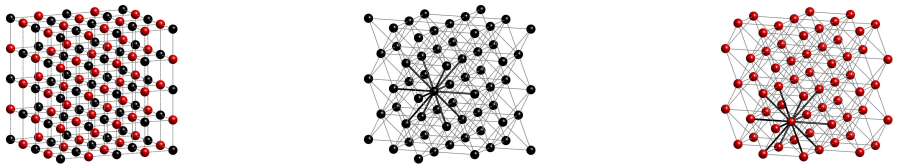
$$\text{FCC:} \quad \lambda = \sqrt{4/3}\,, \qquad \text{BCC:} \quad \lambda = \sqrt{2} \tag{4.242}$$

It is important to remember that the above Courant numbers are defined as $\lambda = ck/h$ with Cartesian grid spacing h, rather than with the inter-nodal distances h'. As such, let us also define the rescaled Courant numbers $\lambda' = ck/h'$ where h' is the inter-nodal distance on the FCC or BCC grid (see Table 4.15). In terms of this rescaled Courant number, the optimal Courant number for spatiotemporal sampling is $\lambda' = \sqrt{2/3}$, equally FCC and BCC grids.

If we compound these temporal sampling constraints with the spatial sampling efficiencies, we can calculate *spatiotemporal sampling efficiencies*, in terms of the spatiotemporal densities required to isotropically sample some solution to the 3-D wave equation with a limiting wavenumber $\beta_{h,\max}$. In doing so, we find that the FCC and BCC grids require 37% and 42% fewer spatiotemporal samples than a Cartesian-based spatiotemporal sampling scheme.

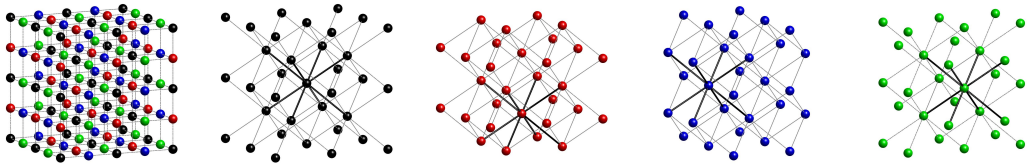
Interleaved FCC and BCC subgrids

Before moving on, we should point out that, in addition to the FCC and BCC subgrids defined above (which include the origin), one finds *shifted* FCC and BCC subgrids on the remaining points in the Cartesian grid. This is illustrated by the checkerboard colouring of the Cartesian grid in Fig. 4.73 for the FCC case (this happens to be the 3-D analogue to Fig. 4.15). Fig. 4.74 illustrates the BCC case, where there are three additional shifted BCC lattices to be found on the Cartesian grid, apart from the BCC subgrid that includes the origin. This is an important detail that leads to spatial decoupling in some Cartesian schemes that were considered earlier, as will be seen shortly.



(a) Cartesian grid (b) FCC subgrid #1 (c) FCC subgrid #2

Figure 4.73: Checkerboard (even-odd) colouring of some points in a Cartesian grid, demonstrating two interleaved FCC subgrids.



(a) Cartesian grid (b) BCC subgrid #1 (c) BCC subgrid #2 (d) BCC subgrid #3 (e) BCC subgrid #4

Figure 4.74: Some Cartesian grid points in four colours, demonstrating four interleaved BCC subgrids.

4.6.2 Discrete Laplacians on FCC and BCC grids

We can adapt the parametrised approximations to the 3-D Laplacian provided in Section 4.5.1 to the FCC and BCC subgrids simply by restricting the triplets $\boldsymbol{q} \in \mathbb{Q}$ to associated stencils of points that are part of those subgrids. When such a stencil fits on the FCC or BCC grid, we will say that the stencil and the associated discrete Laplacian operator are "native" to that grid.

Stencils that are native to the FCC and BCC grids have, respectively, $\boldsymbol{q} \in \mathbb{Q}_\mathrm{f} \subset \mathbb{Q}$ and $\boldsymbol{q} \in \mathbb{Q}_\mathrm{b} \subset \mathbb{Q}$, where $\mathbb{Q}_\mathrm{f}$ and $\mathbb{Q}_\mathrm{b}$ are defined as:

$$\mathbb{Q}_\mathrm{f} := \{\boldsymbol{q} \in \mathbb{Q} : (q_1 + q_1 + q_3) \bmod 2 = 0\} \tag{4.243a}$$

$$\mathbb{Q}_\mathrm{b} := \{\boldsymbol{q} \in \mathbb{Q} : (q_1 \bmod 2) = (q_2 \bmod 2) = (q_3 \bmod 2)\} \tag{4.243b}$$

For example, $\delta_{\Delta,(1,1,0)}$ is native to the FCC grid, and $\delta_{\Delta,(1,1,1)}$ is native to the BCC grid (shown in Figs. 4.75(a) and 4.75(b)), as their stencils pick out the K nearest neighbours on the associated grid.

(a) Stencil for $\delta_{\Delta,(1,1,0)}$, on FCC grid (b) Stencil for $\delta_{\Delta,(1,1,1)}$, on BCC grid

Figure 4.75: Stencils for simplest discrete Laplacians on FCC and BCC grids.

From the discussion about the 2-D rotated five-point scheme in Section 4.3.4, and from the interleaved FCC and BCC subgrids shown earlier (Fig. 4.74), it should be apparent that *spatial decoupling* will be a concern whenever a discrete Laplacian that is native to one of these subgrids is used on the full Cartesian grid.[61] Since the CCP and OCTA schemes (operating over the entire Cartesian grid, af-

[61] Note, this is not exclusive to FCC and BCC subgrids, as one could also define Cartesian subgrids within a Cartesian grid, and find discrete Laplacians that are "native" to those Cartesian subgrids. An example would be $\delta_{\Delta,(2,0,0)}$ with a Cartesian subgrid whose inter-nodal distance is $h' = 2h$.

ter [136]) use $\delta_{\Delta,(1,1,0)}$ and $\delta_{\Delta,(1,1,1)}$, respectively, spatial decouplings are present in those schemes, in a manner similar to what was seen in the rotated five-point scheme in 2-D. This will be discussed in more detail shortly.

Compact discrete Laplacians on the FCC and BCC grids

Analogous to the compact 27-point discrete Laplacian, which uses the first three shells of points on the Cartesian grid, we can define a (compact) 43-point discrete Laplacian using the first three shells of points on the FCC grid:

$$\delta_{\Delta,\Upsilon_{\mathrm{f},43}} = \alpha_{(1,1,0)}\delta_{\Delta,(1,1,0)} + \alpha_{(2,0,0)}\delta_{\Delta,(2,0,0)} + \alpha_{(2,1,1)}\delta_{\Delta,(2,1,1)} \tag{4.244}$$

where $\Upsilon_{\mathrm{f},43} = \{(1,1,0),(2,0,0),(2,1,1)\}$. Similarly, we can define a compact 27-point discrete Laplacian using the first three shells of points on the BCC grid:

$$\delta_{\Delta,\Upsilon_{\mathrm{b},27}} = \alpha_{(1,1,1)}\delta_{\Delta,(1,1,1)} + \alpha_{(2,0,0)}\delta_{\Delta,(2,0,0)} + \alpha_{(2,2,0)}\delta_{\Delta,(2,2,0)} \tag{4.245}$$

where $\Upsilon_{\mathrm{b},27} = \{(1,1,1),(2,0,0),(2,2,0)\}$. For these discrete Laplacians, we define the three-vectors: $\boldsymbol{\alpha}_{\mathrm{f}} = (\alpha_{(1,1,0)},\alpha_{(2,0,0)},\alpha_{(2,1,1)})$ and $\boldsymbol{\alpha}_{\mathrm{b}} = (\alpha_{(1,1,1)},\alpha_{(2,0,0)},\alpha_{(2,2,0)})$, and for consistency we need:

$$\boldsymbol{\alpha}_{\mathrm{f}} \cdot (1,1,1) = 1\,, \qquad \boldsymbol{\alpha}_{\mathrm{b}} \cdot (1,1,1) = 1 \tag{4.246}$$

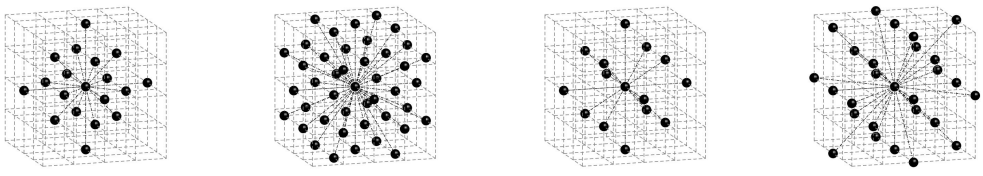
From Taylor expansions of these discrete Laplacians we can obtain constraints for isotropy to fourth-order. For the FCC discrete Laplacian $\delta_{\Delta,\Upsilon_{\mathrm{f},43}}$, the constraint for isotropy is:

$$\alpha_{(2,0,0)} = \frac{3 - 2\alpha_{(1,1,0)}}{11} \neq 0 \tag{4.247}$$

The isotropy constraint for the BCC discrete Laplacian $\delta_{\Delta,\Upsilon_{\mathrm{b},27}}$ is simply:

$$\alpha_{(2,0,0)} = 1/3 \tag{4.248}$$

The choices $\boldsymbol{\alpha}_{\mathrm{f}} = (8/9,1/9,0)$ and $\boldsymbol{\alpha}_{\mathrm{b}} = (2/3,1/3,0)$ lead to isotropic discrete Laplacians with $K = 18$ and $K = 14$, respectively, i.e., using only the first two shells of points (analogous to the 19-point Cartesian isotropic Laplacian). Some of the stencils that can arise from these three-parameter families of discrete Laplacians, other than the special cases $\boldsymbol{\alpha}_{\mathrm{f}} = (1,0,0)$ and $\boldsymbol{\alpha}_{\mathrm{b}} = (1,0,0)$, are illustrated in Fig. 4.76.



(a) $\boldsymbol{\alpha}_{\mathrm{f}} = (\cdot,\cdot,0)$, $K = 18$ (b) $\boldsymbol{\alpha}_{\mathrm{f}} = (\cdot,\cdot,\cdot)$, $K = 42$ (c) $\boldsymbol{\alpha}_{\mathrm{b}} = (\cdot,\cdot,0)$, $K = 14$ (d) $\boldsymbol{\alpha}_{\mathrm{b}} = (\cdot,\cdot,\cdot)$, $K = 26$

Figure 4.76: $(K+1)$-point discrete Laplacians $\delta_{\Delta,\boldsymbol{q}}$ native to FCC and BCC grids for various choices of $\boldsymbol{\alpha}_{\mathrm{f}} = (\alpha_{(1,1,0)},\alpha_{(2,0,0)},\alpha_{(2,1,1)})$ and $\boldsymbol{\alpha}_{\mathrm{b}} = (\alpha_{(1,1,1)},\alpha_{(2,0,0)},\alpha_{(2,2,0)})$. In the captions, "$\cdot$" is a placeholder for a non-zero value.

4.6.3 The simplest non-Cartesian schemes on FCC and BCC grids

In this section we will reconsider the simplest schemes on the FCC and BCC grids, which we respectively called "CCP" and "OCTA" in the Cartesian setting, in order to present their numerical dispersion relations in more appropriate non-cubic wavenumber cells.

The simplest scheme on the FCC grid

The simplest scheme on the FCC grid, henceforth the "13-point FCC scheme", can be written as:

$$\delta_{tt}\underline{u}_{\boldsymbol{i}}^n = c^2\delta_{\Delta,(1,1,0)}\underline{u}_{\boldsymbol{i}}^n\,, \quad \boldsymbol{i} \in \mathcal{I}^{(\mathrm{f})} \tag{4.249}$$

This scheme is essentially the "CCP" Cartesian grid–scheme, but here we restrict $\boldsymbol{i} \in \mathcal{I}^{(\mathrm{f})}$ for the FCC subgrid. Here, $\underline{u}_{\boldsymbol{i}}^n$ is only defined for $\boldsymbol{i} \in \mathcal{I}^{(\mathrm{f})}$, so this scheme uses half the memory (and computational density) of the CCP scheme defined on the Cartesian grid.

In terms of an update recursion, the 13-point FCC scheme can be expressed as:

$$\underline{u}_{\boldsymbol{i}}^{n+1} = (2-3\lambda^2)\underline{u}_{\boldsymbol{i}}^n + \frac{\lambda^2}{4}Q_{\boldsymbol{i},(1,1,0)}^n - \underline{u}_{\boldsymbol{i}}^{n-1}\,, \quad \boldsymbol{i} \in \mathcal{I}^{(\mathrm{f})} \tag{4.250}$$

Considering the Courant number $\lambda = ck/h$ with the Cartesian grid spacing h, this scheme has the stability condition: $\lambda \leq 1$ [21]. In terms of the Courant number $\lambda' = ck/h'$ where $h' = \sqrt{2}h$ is the inter-nodal distance on this FCC grid, we have a stability condition of $\lambda' \leq \sqrt{1/2}$ [196]. Under the choice of $\lambda' = \sqrt{1/3}$, one has the "dodecahedral DWM" [144, 1, 48], wherein the first term on the right-hand side of the update equation cancels out.[62]

The simplest scheme on the BCC grid

The simplest scheme on the BCC grid, which uses the 9-point discrete Laplacian $\delta_{\Delta,(1,1,1)}$, can be expressed as:

$$\delta_{tt}\underline{u}_{\boldsymbol{i}}^n = c^2\delta_{\Delta,(1,1,1)}\underline{u}_{\boldsymbol{i}}^n\,, \quad \boldsymbol{i} \in \mathcal{I}^{(\mathrm{b})} \tag{4.251}$$

This scheme is essentially the "OCTA" Cartesian grid–scheme, but here $\boldsymbol{i}$ is restricted to a BCC subgrid, i.e., $\boldsymbol{i} \in \mathcal{I}^{(\mathrm{b})}$, so this scheme uses *a quarter* of the memory and computational density of the OCTA scheme defined on the Cartesian grid.

The update recursion for this scheme is simply:

$$\underline{u}_{i_x,i_y,i_z}^{n+1} = (2-2\lambda^2)\underline{u}_{i_x,i_y,i_z}^n + \frac{\lambda^2}{4}Q_{\boldsymbol{i},(1,1,0)}^n - \underline{u}_{\boldsymbol{i}}^{n-1}\,, \quad \boldsymbol{i} \in \mathcal{I}^{(\mathrm{b})} \tag{4.252}$$

and for $\lambda = 1$, one has the "octahedral DWM" [21, 48], wherein the first term on the right-hand side cancels out. This scheme permits a concretely passive waveguide mesh implementation for all $\lambda \leq 1$, which is also $\lambda' \leq \sqrt{1/3}$, for $\lambda' = ck/h'$ where $h' = \sqrt{3}h$ is the inter-nodal distance of this BCC grid.

[62] We note that more generally, $\lambda' \leq \sqrt{1/3}$ is the condition required for a concretely passive waveguide mesh implementation using self-loops; see [21]. For $\lambda' = \sqrt{1/2}$, a concretely passive waveguide mesh is in fact possible—indirectly, via two iterations of the tetrahedral DWM [94] (work of the current author). It would be interesting to know if a concretely passive waveguide implementation exists for $\sqrt{1/3} < \lambda' < \sqrt{1/2}$.

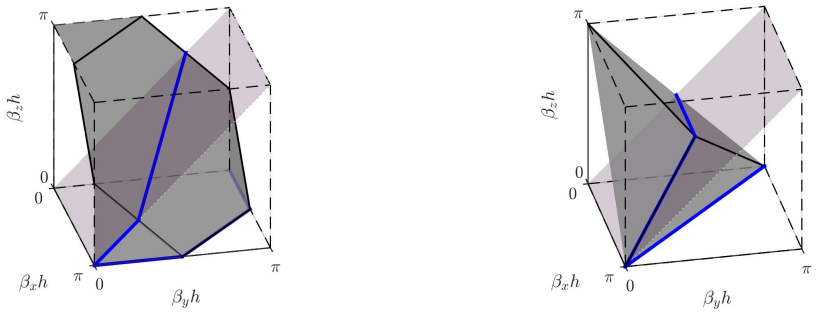
On spatial decoupling

We should note again that that these schemes can be employed on a full Cartesian grid (for $\boldsymbol{i} \in \mathbb{Z}^3$), but this leads to the kind of spatial decoupling discussed for the rotated five-point scheme in 2-D. Considering the previous discussion about interleaved FCC and BCC grids within the Cartesian grid, as illustrated by Figs. 4.73 and 4.74, there is an even-odd decoupling in the CCP scheme. Similarly, the 9-point scheme—native to the BCC grid—decouples into what can be seen as *four* independent solutions. Of course, one is free to employ the CCP and OCTA schemes in this fashion, but it is also wasteful in memory.[63] Such spatial decouplings were not taken into consideration in [131, 134, 136], and as such, the relative computational efficiencies of the CCP and OCTA schemes reported therein can be improved, as will be seen shortly.

Analysing numerical dispersion in FCC and BCC grid–based schemes

We will now analyse the numerical dispersion errors in these non-Cartesian schemes, in terms of normalised wavenumbers $\boldsymbol{\beta}_h = \boldsymbol{\beta}h$, where h is the Cartesian grid spacing. Since we are considering only the FCC and BCC subgrids, we should make use of the wavenumber cells $\mathbb{B}^{(\mathrm{f})}$ and $\mathbb{B}^{(\mathrm{b})}$ defined in (4.239).

It will be informative to start the analysis of dispersion errors by reconsidering the numerical dispersion relations themselves. For the most part, we have considered a 2-D slice of $\boldsymbol{\beta}_h \in [0, \pi]^3$ for which $\beta_z = 0$, and a 2-D side-diagonal slice for which $\beta_y = \beta_z$. These slices of the respective wavenumber cells (positive octants only) are illustrated by Fig. 4.77.



(a) FCC, $\mathbb{B}_{\mathrm{f}}$ (b) BCC, $\mathbb{B}_{\mathrm{b}}$

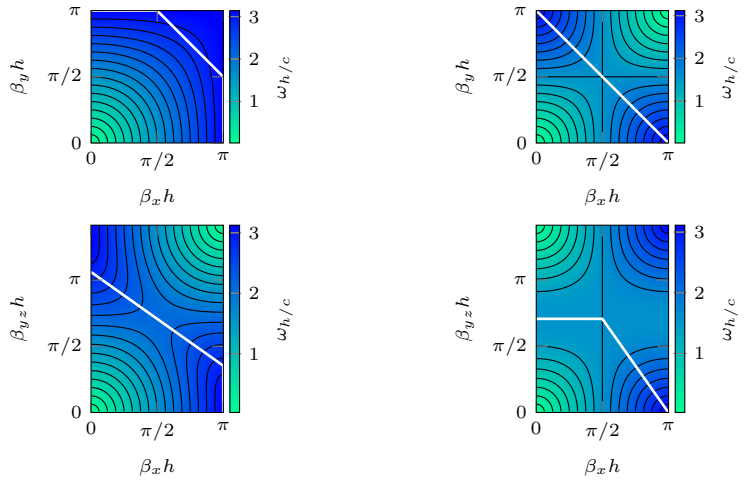
Figure 4.77: Positive octants of wavenumbers cells of FCC and BCC lattices (for $\boldsymbol{\beta}_h$ with Cartesian grid spacing h) and 2-D slice along plane passing through origin with normal vector $\hat{\mathbf{n}} = (0, 1/\sqrt{2}, -1/\sqrt{2})$ intersected with these cells.

The numerical dispersion relations $\omega_{h/c}(\boldsymbol{\beta}_h)$ are shown in Fig. 4.78 for these schemes (with $\lambda = 1$), and the boundaries of the respective wavenumber cells are delimited in white. In this figure it becomes more clear how these dispersion relations follow the periodic tiling of the wavenumber cells.[64] Also, we note that within these more appropriate wavenumber cells, the dispersion relations are monotonically

[63] The 13-point CCP scheme was employed in this fashion in [36], and despite a factor of two redundancy in the CCP scheme, it was concluded that the 13-point CCP scheme outperformed the 27-point IWB scheme for the author's GPU implementations [36]. For more on practical implementations of the 13-point scheme using the native FCC grid, see [101] (by the current author), [304], and [192].

[64] One might notice that the top right corner of the axial slice in Fig. 4.78(a) does not appear to follow a clear periodicity like the other slices. Indeed, it is part of a well-defined periodicity, but this only becomes apparent upon closer examination, starting with the way in which truncated octahedra stack together to tile 3-D space, see Fig. 4.72.

increasing with wavenumber along each wave direction.



(a) FCC 13-point scheme

(b) BCC 9-point scheme

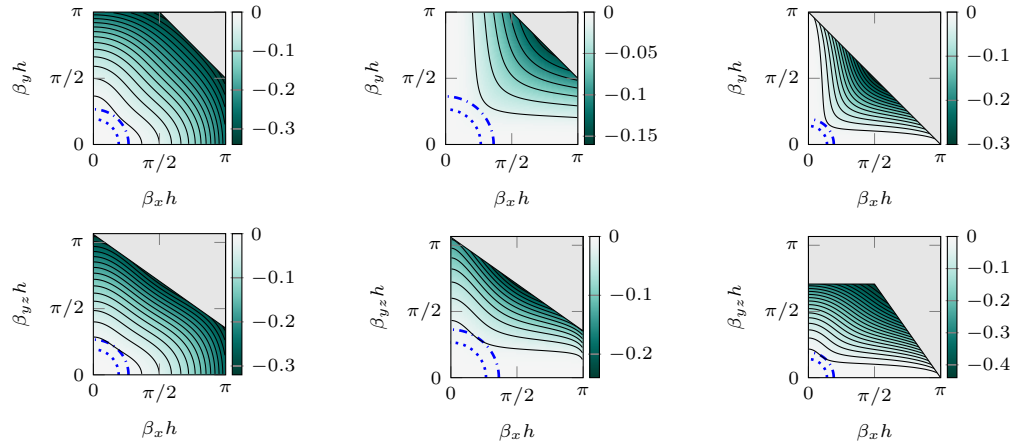
Figure 4.78: Contour plots of numerical dispersion relations for FCC and BCC simplest schemes with $\lambda = 1$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$. The white line denotes the limit of the respective and appropriate wavenumber cells.

The relative dispersion errors for these schemes, along the aforementioned 2-D slices, are shown in Fig. 4.79, wherein the limits of the respective wavenumber cells are denoted by blue lines. We note that Figs. 4.79(b) and 4.79(c) are essentially identical to Figs. 4.53(a) and 4.53(b), except that here we only consider the domains of the respective wavenumber cells. The dispersion errors that are found beyond the limits of these non-Cartesian wavenumber cells do not need to be considered, as they pertain to wavenumbers that cannot be discretised (without leading to aliasing) on the respective subgrids.

We have also displayed the FCC scheme with the choice of Courant number $\lambda = \sqrt{2/3}$, as used in the dodecahedral DWM [144, 1, 82] ($\lambda' = \sqrt{1/3}$). It can be seen that the choice of $\lambda = \sqrt{2/3}$ leads to greater dispersion errors than $\lambda = \lambda_{\max}$, which is as expected for such a $(2, 2)$-accurate scheme.

Comparisons with 27-point compact explicit schemes

It is worth reinvestigating the relative computational efficiencies of these schemes in comparison to the compact 27-point Cartesian schemes. This was initially shown in Fig. 4.56, but here we additionally consider the CCP and OCTA special cases on their native grids; see Fig. 4.80. In this figure, the red dashed line shows the relative computational efficiency (RCE) of the 13-point scheme on the FCC grid (for which $\mu = 1/2$). The green dashed line shows the RCE of this nine-point scheme on the BCC grid (with $\mu = 1/4$). $\mu = 1$ denotes Cartesian grid–based schemes. We can see now that the 13-point FCC scheme is more efficient than all other Cartesian schemes in the compact explicit 27-point family, provided that the error threshold is less than 8%. It is also worth noting that the nine-point BCC scheme is more efficient than the seven-point Cartesian scheme (SLF)—yet still less efficient than other special cases within that family.



(a) FCC 13-point scheme, $\lambda = \sqrt{2/3}$ (b) FCC 13-point scheme, $\lambda = 1$ (c) BCC 9-point scheme, $\lambda = 1$

Figure 4.79: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for FCC and BCC simplest schemes. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$. The grey shaded region is outside of the limits of the respective wavenumber cells.

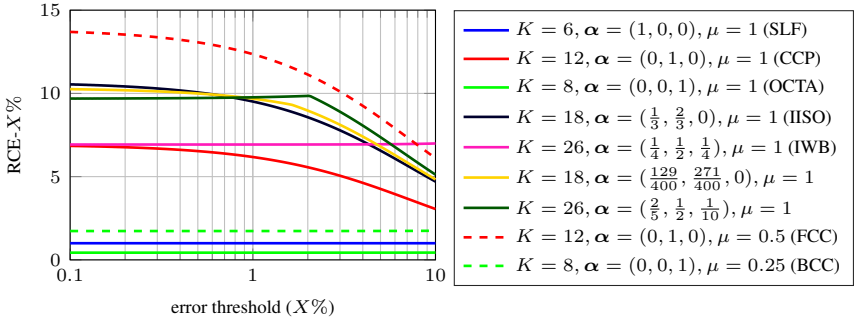


Figure 4.80: Relative computational efficiencies for various choices of $\boldsymbol{\alpha} = (\alpha_{(1,0,0)}, \alpha_{(1,1,0)}, \alpha_{(1,1,1)})$ within compact explicit 27-point family on Cartesian grid ($\mu = 1$), additionally considering non-Cartesian grid densities ($\mu \neq 1$) for the FCC and BCC special cases.

4.6.4 Compact explicit schemes on FCC and BCC grids

At this point we will consider two families of parametrised schemes, using the first three shells of points on the FCC and BCC grids, analogous to the three-parameter family of compact 27-point stencils on the Cartesian grid.

On the FCC grid, a family of schemes can be written as:

$$\delta_{tt} \underline{u}_{\boldsymbol{i}}^n = c^2 \delta_{\Delta, \Upsilon_{\mathrm{f},43}} \underline{u}_{\boldsymbol{i}}^n \,, \quad \boldsymbol{i} \in \mathcal{I}^{(\mathrm{f})} \tag{4.253}$$

where the discrete Laplacian $\delta_{\Delta, \Upsilon_{\mathrm{f},43}}$, employing up to 42 neighbouring points, is parametrised by $\boldsymbol{\alpha}_{\mathrm{f}} = (\alpha_{(1,1,0)}, \alpha_{(2,0,0)}, \alpha_{(2,1,1)})$. The update equation for this scheme can be written as:

$$\begin{aligned} \underline{u}_{\boldsymbol{i}}^{n+1} = {} & \left(2 - \lambda^2 \left(3\alpha_{(1,1,0)} + \frac{3}{2}\alpha_{(2,0,0)} + \alpha_{(2,1,1)}\right)\right) \underline{u}_{\boldsymbol{i}}^n \\ & + \lambda^2 \left(\frac{\alpha_{(1,1,0)}}{4} Q^n_{\boldsymbol{i},(1,1,0)} + \frac{\alpha_{(2,0,0)}}{4} Q^n_{\boldsymbol{i},(2,0,0)} + \frac{\alpha_{(2,1,1)}}{24} Q^n_{\boldsymbol{i},(2,1,1)}\right) - \underline{u}_{\boldsymbol{i}}^{n-1} \,, \quad \boldsymbol{i} \in \mathcal{I}^{(\mathrm{f})} \end{aligned} \tag{4.254}$$

and $\alpha_{(2,1,1)} = 1 - \alpha_{(1,1,0)} - \alpha_{(2,0,0)}$.

On the BCC grid, we consider the following scheme:

$$\delta_{tt}\underline{u}_{\boldsymbol{i}}^n = c^2\delta_{\Delta,\Upsilon_{\mathrm{b},27}}\underline{u}_{\boldsymbol{i}}^n\,, \quad \boldsymbol{i} \in \mathcal{I}^{(\mathrm{b})} \tag{4.255}$$

which is parametrised by $\boldsymbol{\alpha}_{\mathrm{b}} = (\alpha_{(1,1,1)}, \alpha_{(2,0,0)}, \alpha_{(2,2,0)})$. The update recursion for this scheme is:

$$\begin{aligned}\underline{u}_{\boldsymbol{i}}^{n+1} = &\left(2 - \lambda^2\left(2\alpha_{(1,1,1)} + \frac{3}{2}\alpha_{(2,0,0)} + \frac{3}{4}\alpha_{(2,2,0)}\right)\right)\underline{u}_{\boldsymbol{i}}^n \\ &+ \lambda^2\left(\frac{\alpha_{(1,1,1)}}{4}Q_{\boldsymbol{i},(1,1,1)}^n + \frac{\alpha_{(2,0,0)}}{4}Q_{\boldsymbol{i},(2,0,0)}^n + \frac{\alpha_{(2,2,0)}}{16}Q_{\boldsymbol{i},(2,2,0)}^n\right) - \underline{u}_{\boldsymbol{i}}^{n-1}\,, \quad \boldsymbol{i} \in \mathcal{I}^{(\mathrm{b})}\end{aligned} \tag{4.256}$$

and for consistency we require that $\alpha_{(2,2,0)} = 1 - \alpha_{(1,1,1)} - \alpha_{(2,0,0)}$.

Numerical stability

Conditions that are *sufficient* for stability with this family of schemes are straightforward to derive. For the FCC and BCC schemes, we arrive at the following stability (concrete passivity) conditions:

$$\text{FCC:} \quad \lambda \leq \sqrt{\frac{4}{2 + 4\alpha_{(1,1,0)} + \alpha_{(2,0,0)}}}\,, \qquad \text{BCC:} \quad \lambda \leq \sqrt{\frac{8}{3 + 5\alpha_{(1,1,1)} + 3\alpha_{(2,0,0)}}} \tag{4.257}$$

with $\boldsymbol{\alpha}_{\mathrm{f}} \geq 0$ and $\boldsymbol{\alpha}_{\mathrm{f}} \geq 0$ (all coefficients positive).

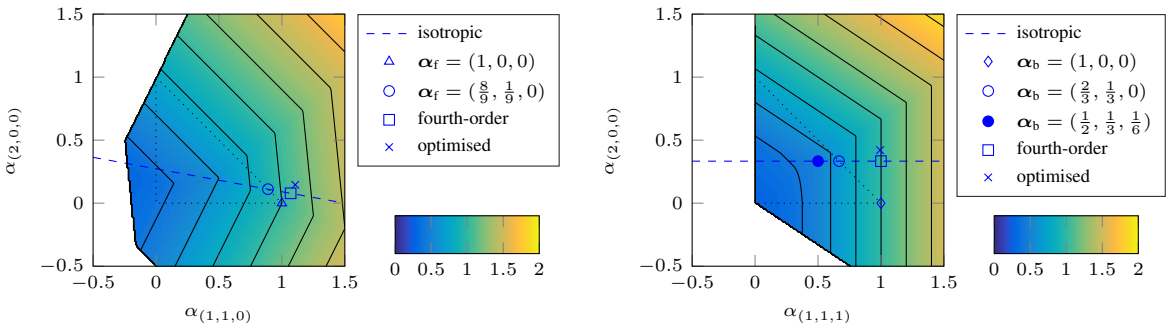
However, we expect that the interesting cases will be found at Courant numbers that go above sufficient (passivity) conditions, as in many special cases of the Cartesian 27-point family [21] and also, as in the case of the 13-point FCC scheme (for which (4.257) is too strict). Furthermore, high-order accuracy is only possible provided that at least one free parameter in the discrete Laplacian is negative (for such an explicit scheme). It would thus be desirable to have necessary and sufficient conditions for these families of schemes (as functions of $\boldsymbol{\alpha}_{\mathrm{f}}$ or $\boldsymbol{\alpha}_{\mathrm{b}}$).

For these schemes, $G(\boldsymbol{\beta}_h)$ is more complicated to analyse than it was for the analogous Cartesian family of schemes, since here, $G(\boldsymbol{\beta}_h)$ is not multilinear in general, and related to this, global extrema in $G(\boldsymbol{\beta}_h)$ are not guaranteed to be found at particular locations on the boundaries of associated wavenumber cells (such as vertices or face-centers).[65] The derivation of necessary and sufficient conditions is currently out of reach, but for now we can still investigate $G(\boldsymbol{\beta}_h)$ numerically. With that said, $\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)$ as a function of $\boldsymbol{\alpha}_{\mathrm{f}}$ and $\boldsymbol{\alpha}_{\mathrm{b}}$ are shown in Fig. 4.81 (calculated numerically), masked by $\min_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h) \geq 0$. We see that, contrary to the Cartesian schemes in Figs. 4.51, stability regions are not entirely divided into convex regions.[66]

Despite not being able to immediately obtain a general stability condition, there are many interesting choices of $\boldsymbol{\alpha}_{\mathrm{f}}$ and $\boldsymbol{\alpha}_{\mathrm{b}}$ for which global extrema are indeed found at vertices of the respective wavenumber cells (and for which we can derive necessary and sufficient stability conditions). We present a few such cases (labelled in Fig. 4.81), starting with the isotropic second-order special cases that are listed in Table 4.16.

[65] Two such examples we have found are $\boldsymbol{\alpha}_{\mathrm{f}} = (0, 0, 1)$ and $\boldsymbol{\alpha}_{\mathrm{b}} = (0.25, 0.15, 0.6)$.

[66] In fact, there are complex details in these stability regions not apparent from this figure; we only give a cursory view for brevity.



(a) FCC (b) BCC

Figure 4.81: Plots of $\max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)$ (evaluated numerically) for FCC and BCC compact schemes, each as a function of the two free parameters and masked by $\min_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h) \geq 0$. Special case schemes are labelled and isotropic schemes appear along the dashed line. The concave regions to the left (not coloured) have $\min_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h) \leq 0$ and are therefore unstable. Contour lines denote increments of 0.2. Concretely passive schemes (following (4.257)) are possible within the triangular region demarcated by the dotted lines.

Table 4.16: Isotropic special cases with $K \leq 26$.

grid	K	$\alpha_{(1,0,0)}$	$\alpha_{(1,1,0)}$	$\alpha_{(1,1,1)}$	$\lambda_{\max}$	$\lambda'_{\max}$
FCC	18	$8/9$	$1/9$	0	$\sqrt{12/11}$	$\sqrt{6/11}$
BCC	14	$2/3$	$1/3$	0	$\sqrt{3/2}$	$\sqrt{3/4}$
BCC	26	$1/2$	$1/3$	$1/6$	$\sqrt{2}$	$\sqrt{2/3}$

Isotropic $(K+1)$-point schemes with $K \leq 26$

We have identified three interesting isotropic schemes within this family of FCC and BCC compact explicit schemes, whose parameters are listed in Table 4.16. These include the unique isotropic schemes that make use of the first two shells of points on each respective grid (analogous to the 19-point isotropic (IISO) scheme on the Cartesian grid), and one 27-point isotropic scheme on the BCC grid, which has $\lambda_{\max}$ set to the optimal value for spatiotemporal sampling (as discussed in Section 4.6.1). The last column of Table 4.16 gives the maximal Courant numbers in terms of the inter-nodal distances of the respective grids, such that $\lambda' = ck/h' \leq \lambda'_{\max}$ is also the stability condition.

The relative dispersion errors of these schemes (with $\lambda = \lambda_{\max}$) are shown in Fig. 4.82. In order to get a better idea of how these schemes compare with each other (having different Courant numbers) and how they compare to the Cartesian schemes from the compact 27-point family, we plot the relative computational efficiencies of these schemes in Fig. 4.83. In that figure, it can be seen that the special cases identified here are significantly more efficient than analogous Cartesian schemes. Notably, the 19-point isotropic scheme on the FCC grid is most efficient for less than 3% dispersion errors, above which the 27-point isotropic scheme on the BCC grid is most efficient (considering only up to 10% errors). It is also worth mentioning that this 27-point scheme on the BCC grid is more than two times as efficient than the 27-point IWB scheme for less than 7% dispersion errors.

We also list minimal cutoff frequencies and critical wavenumbers for the special cases within these families of non-Cartesian schemes in Table 4.17. It is important to remember that the maximal wavenumber that can be isotropically discretised on these non-Cartesian grids is not simply π/h (or even π/h').

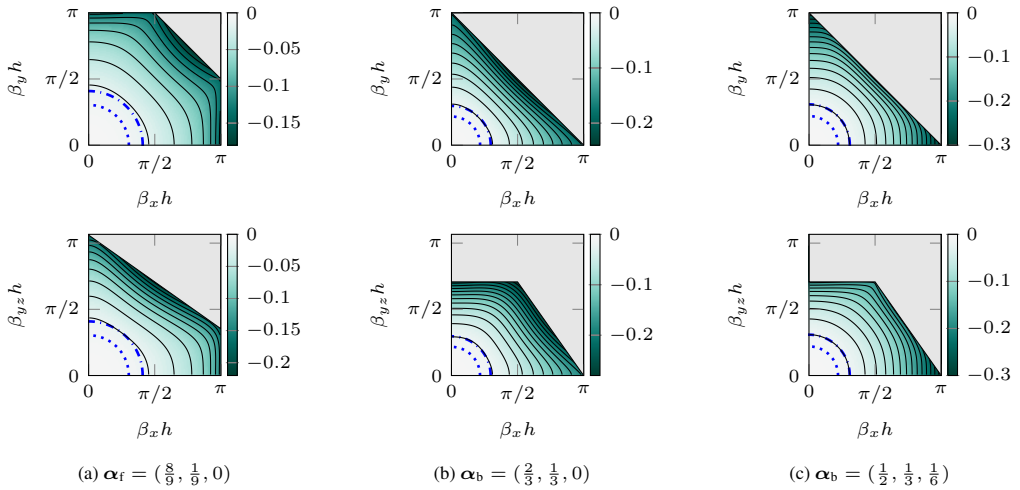
As such, we have provided dimensionless critical wavenumbers, defined as:

$$\beta'_{h,X\%} = \frac{\beta_{h,X\%}}{\text{inradius of wavenumber cell}} \tag{4.258}$$

and for any $X\%$, $\beta'_{h,X\%}$ can be at most unity.

Table 4.17: Special cases of compact schemes on FCC and BCC grids with $K \leq 26$.

grid	K	$\alpha_{(1,0,0)}$	$\alpha_{(1,1,0)}$	$\alpha_{(1,1,1)}$	$\lambda_{\max}$	$\omega_{k,\mathrm{m}}$	$\beta_{h,1\%}$	$\beta_{h,2\%}$	$\beta'_{h,1\%}$	$\beta'_{h,2\%}$
FCC	12	1	0	0	1	0.667π	0.261π	0.358π	0.302	0.414
FCC	18	0	1	0	$\sqrt{12/11}$	0.720π	0.301π	0.408π	0.348	0.471
BCC	8	0	0	1	1	0.500π	0.134π	0.189π	0.190	0.268
BCC	14	2/3	1/3	0	$\sqrt{3/2}$	0.667π	0.213π	0.293π	0.302	0.414
BCC	26	1/2	1/3	1/6	$\sqrt{2}$	0.814π	0.218π	0.306π	0.309	0.432



(a) $\boldsymbol{\alpha}_{\mathrm{f}} = (\frac{8}{9}, \frac{1}{9}, 0)$ (b) $\boldsymbol{\alpha}_{\mathrm{b}} = (\frac{2}{3}, \frac{1}{3}, 0)$ (c) $\boldsymbol{\alpha}_{\mathrm{b}} = (\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$

Figure 4.82: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for parametrised compact schemes on FCC and BCC grids. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$. The grey shaded region is outside of the limits of the respective wavenumber cells.

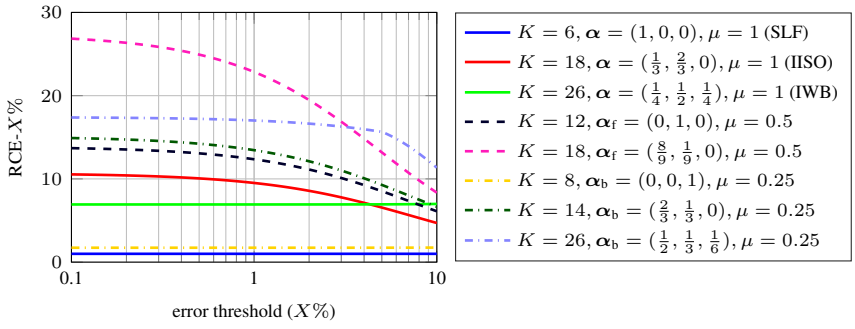


Figure 4.83: Relative computational efficiencies for parametrised compact $(K+1)$-point schemes on Cartesian and non-Cartesian grids. $\mu = 1$ denotes Cartesian grid–based schemes, $\mu = 0.5$ denotes FCC grid–based schemes, and $\mu = 0.25$ denotes BCC grid–based schemes.

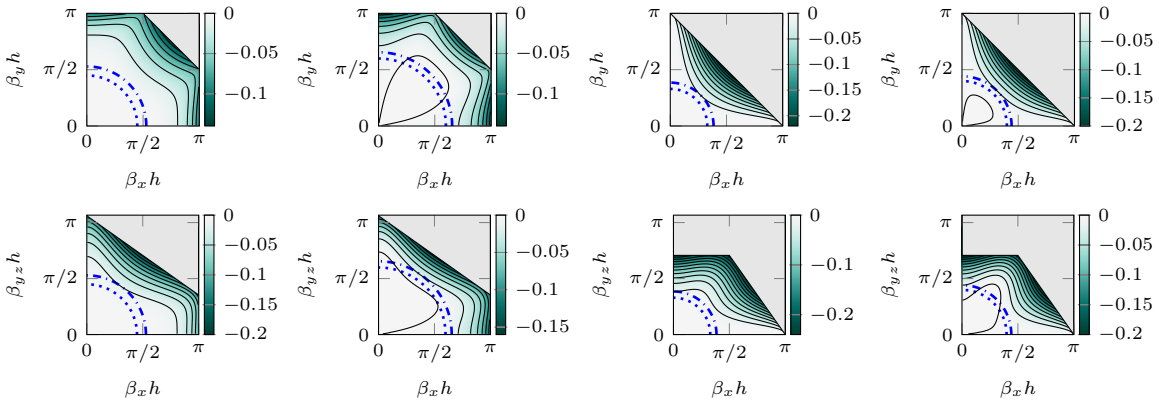
High-order accurate and optimised schemes

At this point we will consider the other special cases that are labelled in Fig. 4.81, namely schemes with fourth-order accuracy (in both space and time) and optimised schemes. The coefficients for the high-order and optimised schemes are listed in Table 4.18. Coefficients for optimised schemes are rounded to three decimal places for brevity, and associated Courant numbers are truncated. Fourth-order accuracy was achieved following the usual approach (imposing isotropy and then using Eqs. (4.145) and (4.146), followed by a stability analysis to identify $\lambda = \lambda_{\max}$). The optimised schemes, optimised for one-percent critical wavenumbers (and taking spatiotemporal densities into account) were found using the same approach as the Cartesian high-order schemes. The relative dispersion errors for these schemes are shown in Fig. 4.84. It is worth pointing out that the fourth-order BCC scheme features exact dispersion along the axial directions.

Also listed in Table 4.18 are the one- and two-percent critical wavenumbers, for which significant improvements can be seen over the schemes in Table 4.17. Notably, the 43-point optimised scheme on the FCC grid has less than two-percent error for 75% of the spatial bandwidth that can be sampled isotropically on the grid, which is comparable in performance to the 43-point optimised scheme previously found on the Cartesian grid (see Table 4.14).

Table 4.18: Fourth-order and optimised schemes on FCC and BCC grids. Note: $\sqrt{\frac{17-\sqrt{73}}{9}} \approx 0.939$.

grid	K	OoA	$\alpha_{(1,1,0)}$	$\alpha_{(1,1,1)}$	$\alpha_{(2,0,0)}$	$\alpha_{(2,1,1)}$	$\alpha_{(2,2,0)}$	λ	$\beta_{h,1\%}$	$\beta_{h,2\%}$	$\beta'_{h,1\%}$	$\beta'_{h,2\%}$
FCC	42	$(4,4)$	$\frac{137+11\sqrt{73}}{216}$	—	$\frac{17-\sqrt{73}}{108}$	$\frac{5-\sqrt{73}}{24}$	—	$\sqrt{\frac{17-\sqrt{73}}{9}}$	0.448π	0.528π	0.517	0.609
FCC	42	$(2,2)$	1.106	—	0.146	-0.252	—	0.953	0.599π	0.651π	0.691	0.752
BCC	26	$(4,4)$	—	1	$1/3$	—	$-1/3$	1	0.322π	0.384π	0.456	0.543
BCC	26	$(2,2)$	—	0.993	0.421	—	-0.414	1.003	0.393π	0.441π	0.555	0.624



(a) fourth-order, FCC (b) optimised, FCC (c) fourth-order, BCC (d) optimised, BCC

Figure 4.84: Contour plots of $\tilde{v}_{\mathrm{p}}(\boldsymbol{\beta}_h) - 1$ for fourth-order accurate and optimised second-order accurate schemes on FCC and BCC grids, whose parameters are listed in Table 4.18. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$. The grey shaded region is outside of the limits of the respective wavenumber cells.

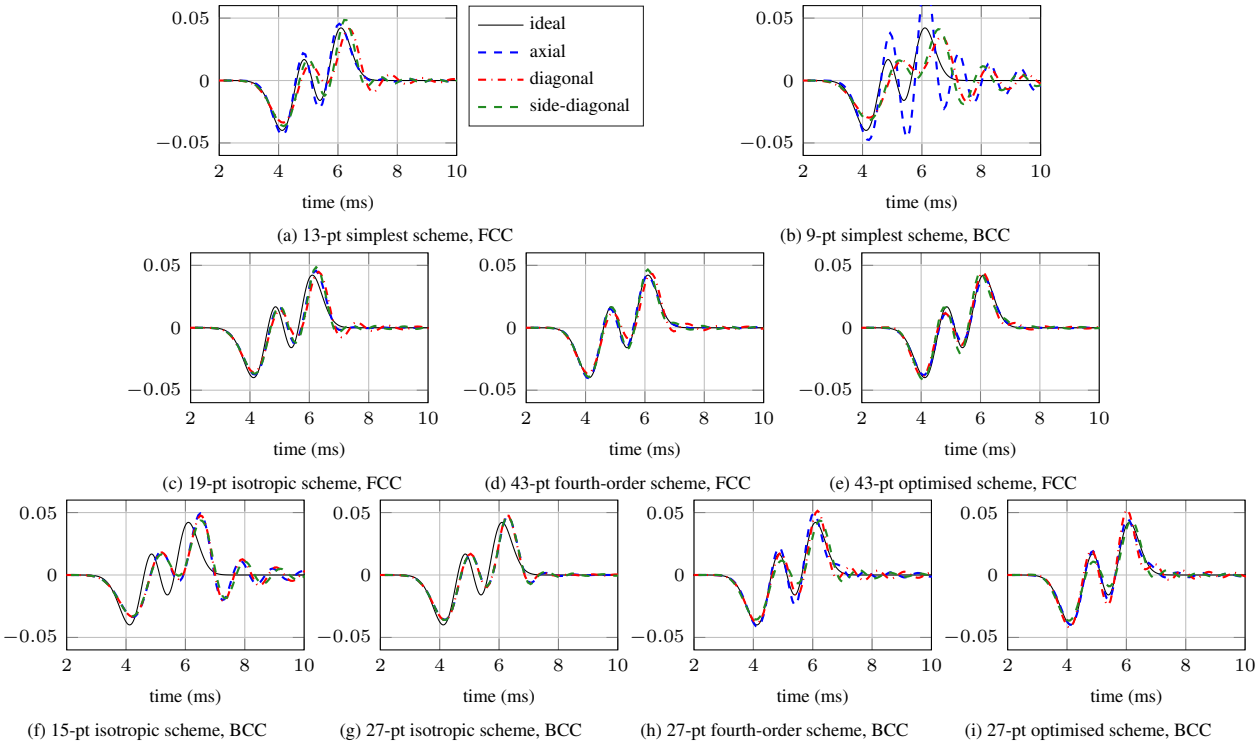
4.6.5 Numerical example

In order to compare the behaviour of these schemes with the previously-seen Cartesian schemes, we consider the same problem that was used for Cartesian schemes in Section 4.5.6. Here we consider the same two scenarios: fixed memory costs (equivalent to those used for Cartesian schemes, using $h = \mu^{1/3}h^\star$ with $h^\star = 14\,\text{cm}$), and fixed spatiotemporal densities (using $h = (\mu/\lambda)^{1/4}h^*$ with $h^* = 14\,\text{cm}$). Recall, the spatial densities are $\mu = 1/2$ for the FCC grid and $\mu = 1/4$ for the BCC grid. The resulting outputs of these simulations are shown in Figs. 4.85 and 4.87, and the associated error signals are shown in Figs. 4.86 and 4.88.

Discussion

We can compare the outputs from this numerical to those previously obtained with Cartesian schemes. Considering fixed memory costs, it is apparent by comparing Figs. 4.85(a) and 4.85(b) to Figs. 4.64(e) and 4.64(b) that the simplest schemes on FCC and BCC grids are most efficient on their native non-Cartesian grids, rather than as CCP and OCTA Cartesian grid–based schemes. Considering fixed spatiotemporal costs, we arrive at the same conclusion by comparing Figs. 4.87(a) and 4.87(b) to Figs. 4.65(e) and 4.65(b) (which are identical to Figs. 4.64(e) and 4.64(b)). We can also note that the simplest BCC scheme only offers results that are comparable in performance to the simplest Cartesian scheme.

Considering the 19-point isotropic scheme on the FCC scheme, we find that there is a notable improvement over all previous 19-point Cartesian schemes (leggy and compact), with the exception of the 19-point optimised leggy scheme with $\lambda = \lambda_{1\%,3}$ under fixed memory costs. As for the 43-point FCC schemes, we find that they are comparable in accuracy to the 43-point compact Cartesian schemes, and slightly better when only memory costs are a concern. We also find that the 27-point BCC schemes are similar in performance to the 25-point compact Cartesian schemes.

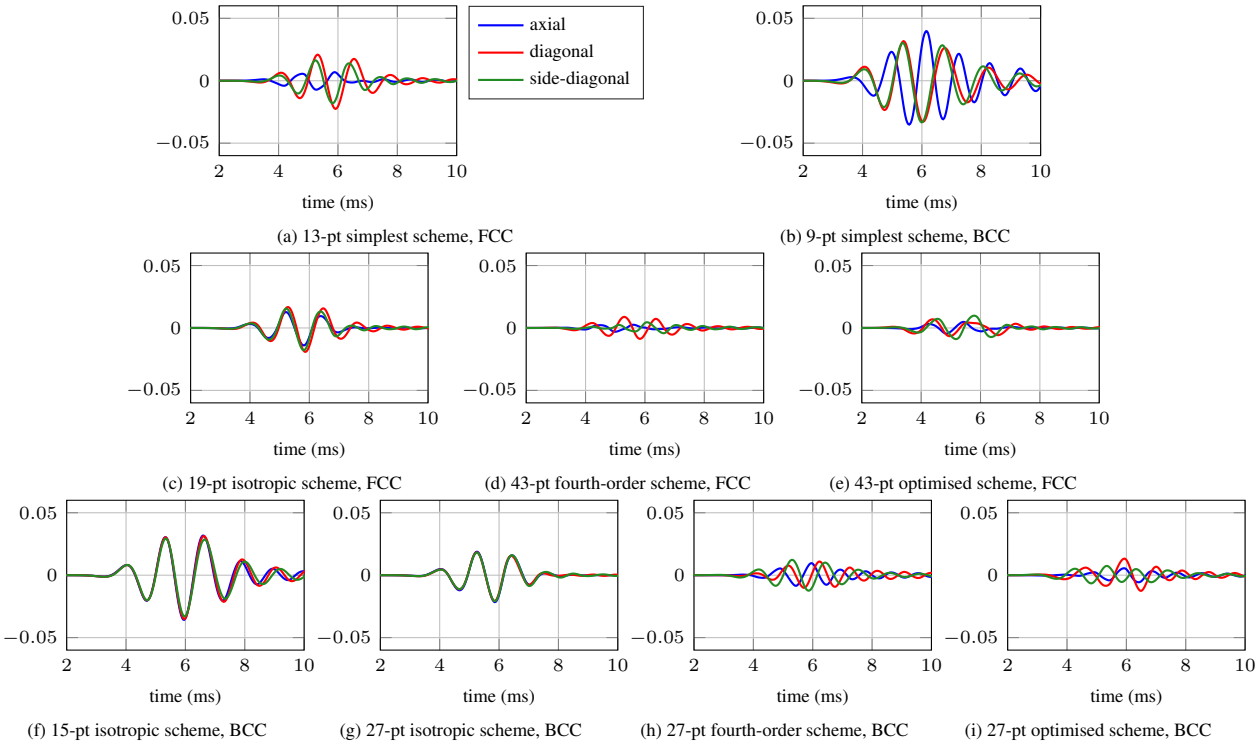


(a) 13-pt simplest scheme, FCC (b) 9-pt simplest scheme, BCC

(c) 19-pt isotropic scheme, FCC (d) 43-pt fourth-order scheme, FCC (e) 43-pt optimised scheme, FCC

(f) 15-pt isotropic scheme, BCC (g) 27-pt isotropic scheme, BCC (h) 27-pt fourth-order scheme, BCC (i) 27-pt optimised scheme, BCC

Figure 4.85: Time-domain plots of outputs at distance $\sqrt{3}$ m from origin, along axial, side-diagonal, and diagonal directions. Here, the various schemes have the same grid spacing (fixed memory costs), and memory costs are equivalent to those used in Fig. 4.64.

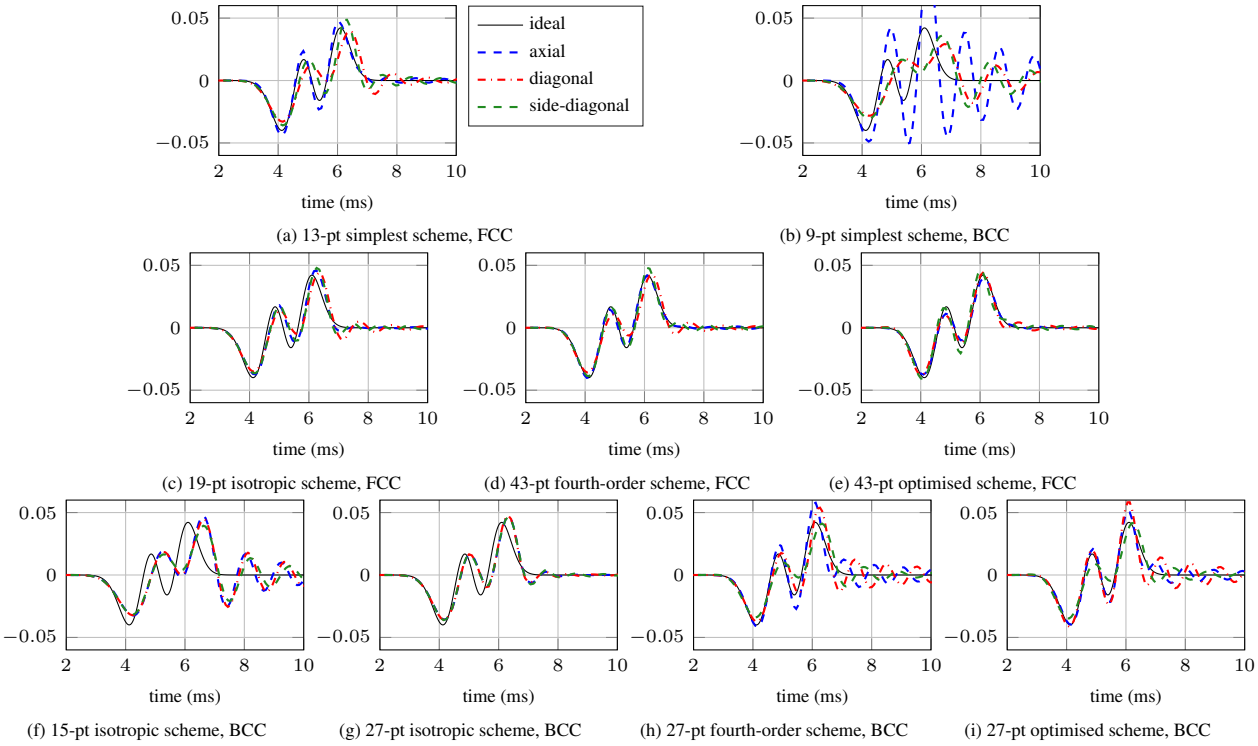


(a) 13-pt simplest scheme, FCC (b) 9-pt simplest scheme, BCC

(c) 19-pt isotropic scheme, FCC (d) 43-pt fourth-order scheme, FCC (e) 43-pt optimised scheme, FCC

(f) 15-pt isotropic scheme, BCC (g) 27-pt isotropic scheme, BCC (h) 27-pt fourth-order scheme, BCC (i) 27-pt optimised scheme, BCC

Figure 4.86: Plots of error signals from Fig. 4.85 (fixed memory costs).

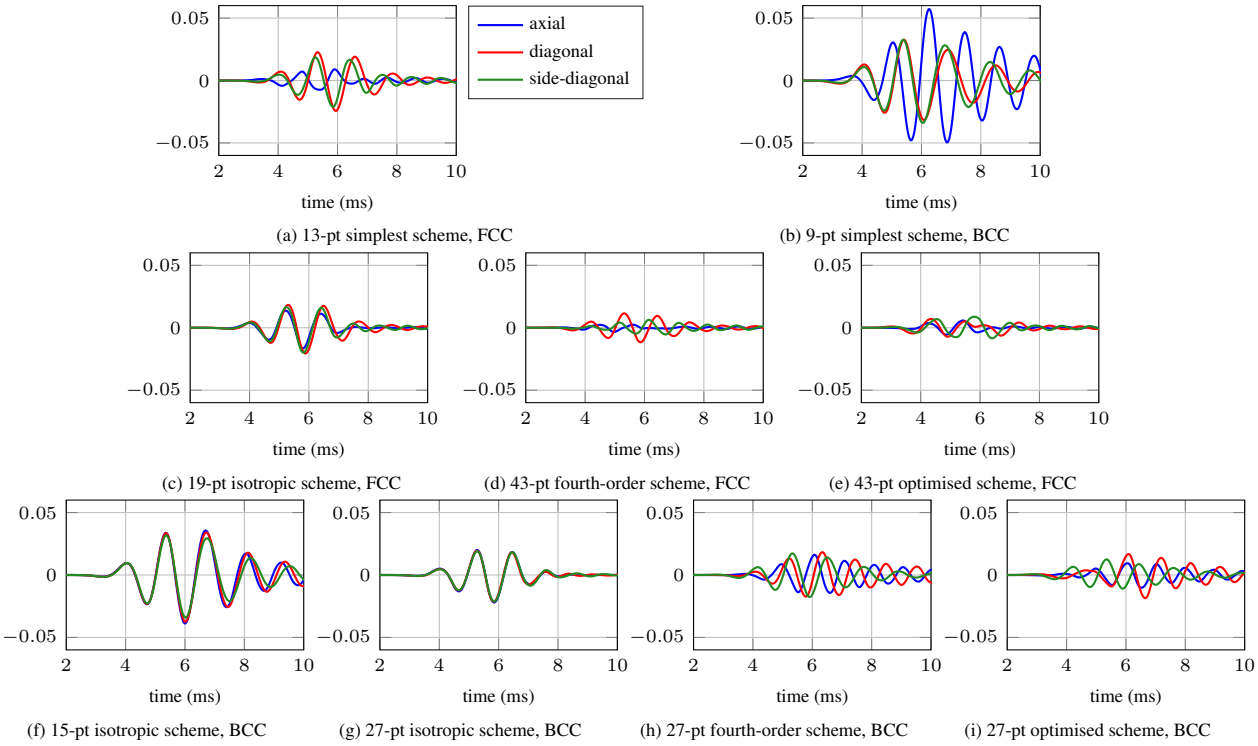


(a) 13-pt simplest scheme, FCC (b) 9-pt simplest scheme, BCC

(c) 19-pt isotropic scheme, FCC (d) 43-pt fourth-order scheme, FCC (e) 43-pt optimised scheme, FCC

(f) 15-pt isotropic scheme, BCC (g) 27-pt isotropic scheme, BCC (h) 27-pt fourth-order scheme, BCC (i) 27-pt optimised scheme, BCC

Figure 4.87: Time-domain plots of outputs at distance $\sqrt{3}$ m from origin, along axial, side-diagonal, and diagonal directions. Here, the various schemes are normalised for spatiotemporal densities (equivalent to those used in Fig. 4.66), which also means similar operation counts for similar stencil sizes.



(a) 13-pt simplest scheme, FCC (b) 9-pt simplest scheme, BCC

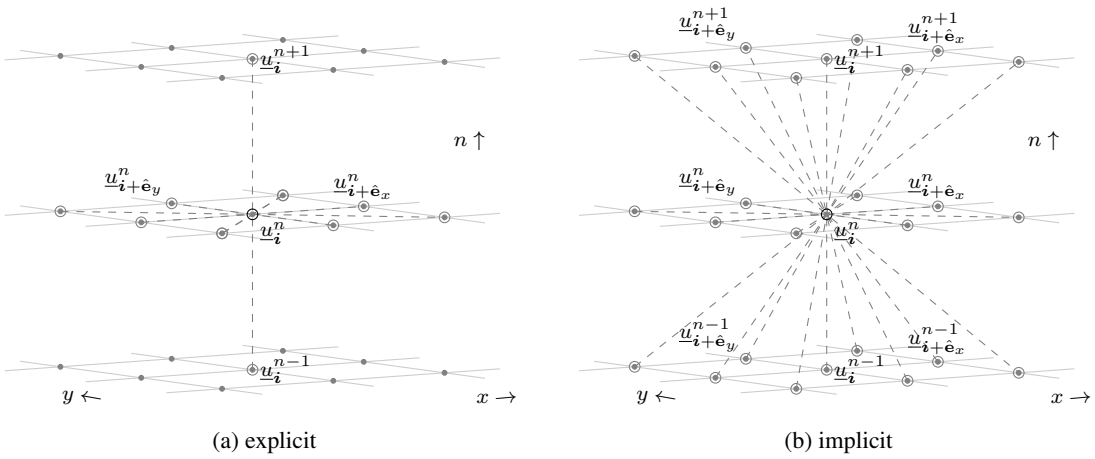
(c) 19-pt isotropic scheme, FCC (d) 43-pt fourth-order scheme, FCC (e) 43-pt optimised scheme, FCC

(f) 15-pt isotropic scheme, BCC (g) 27-pt isotropic scheme, BCC (h) 27-pt fourth-order scheme, BCC (i) 27-pt optimised scheme, BCC

Figure 4.88: Plots of error signals from Fig. 4.87 (normalised for spatiotemporal densities).

4.7 Implicit finite difference schemes for the wave equation

In this section we conduct a short investigation of *implicit* finite difference schemes for the wave equation in 2-D and 3-D. In contrast to explicit schemes, the time integration used in implicit schemes requires a linear system solution over the spatial domain, so they demand more computational resources for each time-step update. On the other hand, implicit formulations can often achieve higher orders of (global) accuracy than explicit formulations using the same spatial stencil. For example, we have seen that the compact nine-point explicit scheme on the 2-D Cartesian grid is limited to second-order global accuracy (in Section 4.3.4), but the same spatial stencil combined with an implicit time integration scheme can achieve fourth-order global accuracy [77]; this is ultimately due to the fact that the implicit formulation takes into account more points in space and time for each point-wise update; see Fig. 4.89.



(a) explicit (b) implicit

Figure 4.89: Illustration of explicit and implicit spatiotemporal stencils for compact scheme with nine-point 2-D spatial stencil.

To keep this section brief, we will consider only the simplest fourth-order accurate implicit schemes on Cartesian and hexagonal grids in 2-D, and on 3-D Cartesian, FCC, and BCC grids, under a unified notation. The 1-D and 2-D Cartesian schemes are well-known, and while the rest are similarly derived, they have yet to be presented in the literature (to our knowledge[67]). However, for illustrative purposes, we start with the classic implicit scheme in 1-D, which is due to von Neumann [185].

4.7.1 Preliminaries

von Neumann's 1-D implicit scheme

von Neumann's implicit scheme for the 1-D wave equation [185] can be written as:

$$\delta_{tt}\underline{u}_i^n = c^2(1 + \alpha k^2 \delta_{tt})\delta_{xx}\underline{u}_i^n \tag{4.259}$$

where $\alpha \in \mathbb{R}$ is a free parameter and $\underline{u}_i^n \approxeq u(ih, nk)$. In fact, this scheme is only implicit for $\alpha \neq 0$; for $\alpha = 0$ this scheme reduces to the simple explicit scheme. It is straightforward to check that this scheme is consistent and second-order accurate, since, under an abuse of notation, $\delta_{tt} = \partial_t^2 + \mathcal{O}(k^2)$

[67] Aside from published work involving the current author [100, 31].

and $\delta_{xx} = \partial_x^2 + \mathcal{O}(h^2)$ imply that:

$$(1 + k^2\alpha\delta_{tt})\delta_{xx} = \left(1 + \alpha k^2\partial_t^2 + \mathcal{O}(k^2)\right)\left(\partial_x^2 + \mathcal{O}(h^2)\right) = \partial_x^2 + \mathcal{O}(h^2) + \mathcal{O}(k^2) + \mathcal{O}(h^2k^2) \quad (4.260)$$

So for $\lambda = ck/h > 0$ fixed and c fixed, and making the substitution $h = ck/\lambda$, we have:

$$\delta_{tt} - c^2(1 + k^2\alpha\delta_{tt})\delta_{xx} = \partial_t^2 - c^2\partial_x^2 + \mathcal{O}(h^2) \quad (4.261)$$

The requirement for a linear system solution is seen by expanding the scheme in full:

$$\begin{aligned}&\underline{u}_i^{n+1} = 2\underline{u}_i^n - \underline{u}_i^{n-1} + \lambda^2\left(\underline{u}_{i+1}^n - 2\underline{u}_i^n + \underline{u}_{i-1}^n\right)\\ &+ \alpha\lambda^2\left(\left(\underline{u}_{i+1}^{n+1} - 2\underline{u}_i^{n+1} + \underline{u}_{i-1}^{n+1}\right) - 2\left(\underline{u}_{i+1}^n - 2\underline{u}_i^n + \underline{u}_{i-1}^n\right) + \left(\underline{u}_{i+1}^{n-1} - 2\underline{u}_i^{n-1} + \underline{u}_{i-1}^{n-1}\right)\right) \quad (4.262)\end{aligned}$$

We see that the solution at $\underline{u}_i^{n+1}$ depends on $\underline{u}_{i+1}^{n+1}$ and $\underline{u}_{i-1}^{n+1}$, which are also unknown, so we must simultaneously solve for all spatial grid points at time $n + 1$. This can be achieved using standard iterative techniques, and sometimes with exact approaches (e.g., using the Thomas algorithm [77]). We will not go into detail about such techniques in this section. We should also note that the storage requirements of this scheme are higher than the explicit two-step form; here three states must be kept in memory, and an additional temporary state is typically required for iterative linear system solution techniques.

Typically, the main interest in implicit schemes is their stability properties. These can be derived through a von Neumann analysis [185] or using energy methods [148], and by now the following condition is well-known so we omit its derivation. The stability condition for this scheme is:

$$\lambda^2 \leq \frac{1}{1 - 4\alpha} \quad \text{when } \alpha < 1/4 \quad (4.263)$$

and for $\alpha \geq 1/4$, the scheme is *unconditionally stable*. As such, for $\alpha \geq 1/4$, the time-step k can be set arbitrarily high for a fixed h, c. At first, this might seem to be an advantage over the explicit scheme, since the explicit scheme is limited to $k \leq h/c$, and a larger k means that fewer time-steps need to be computed for a given duration of time. However, we should recall that for a given h, c, there is an upper limit to the valid bandwidth in a Cartesian grid–based scheme (as explained in Section 4.1.1) and that $\lambda > 1$ would open the door to *temporal aliasing*. So there is actually a limit to this "freedom" in choosing the time-step arbitrarily large, at least in the case of the wave equation.[68] But the ability to formulate unconditionally stable schemes is not actually the interest here; rather, we are interested in using implicit schemes in order to achieve high orders of accuracy.

An equivalent form

Although less conventional, we can also rewrite this scheme in an equivalent form (after, e.g., [22]):

$$(1 + \zeta h^2)\delta_{tt}\underline{u}_i^n = c^2\delta_{xx}\underline{u}_i^n \quad (4.264)$$

where $\zeta \in \mathbb{R}$ is now the free parameter. It is straightforward to see that this scheme is (4.259), with:

$$\zeta = -\alpha\lambda^2 \quad (4.265)$$

[68] Not only that, the simplest explicit scheme is exact for the special case of $\lambda = 1$, so the implicit formulation cannot offer any improvement.

In this case, the stability condition (4.263) becomes:

$$\lambda^2 \leq \frac{\lambda^2}{\lambda^2 + 4\zeta} \tag{4.266}$$

Note that $\alpha \geq 1/4 \implies \lambda^2 + 4\zeta \leq 0$, in which case the above condition is automatically satisfied (unconditional stability). For the conditional stability region ($\alpha < 1/4$)—the region of interest—we have $\lambda^2 + 4\zeta > 0$, and the above simplifies to:

$$\lambda^2 \leq 1 - 4\zeta\,, \quad \zeta < 1/4 \tag{4.267}$$

A general implicit scheme

Let us now generalise von Neumann's implicit scheme, under the alternative form (4.264) with ζ as the free parameter. If we were to consider only Cartesian grid–based schemes (and FCC and BCC forms in 3-D), we could define our grid function in the usual way, i.e., $\underline{u}_{\boldsymbol{i}}^n \approxeq u(\boldsymbol{i}h, nk)$, where $\boldsymbol{i} \in \mathbb{Z}^d$ with grid spacing h and time-step k. However, we need to also hexagonal grids in 2-D, so we define our grid function in a slightly modified form as: $\underline{u}_{\boldsymbol{i}}^n \approxeq u(\mathbf{V}\boldsymbol{i}h, nk)$, where $\mathbf{V} = \mathbf{V}_{\mathrm{H}}$ for hexagonal grids (as defined by (4.157)), otherwise $\mathbf{V}$ is the identity matrix (and $\boldsymbol{i}$ is taken to be a column-vector).

We can then write a general implicit scheme as:

$$(1 + \zeta h^2 \delta_\Delta)\delta_{tt}\underline{u}_{\boldsymbol{i}}^n = c^2 \delta_\Delta \underline{u}_{\boldsymbol{i}}^n \tag{4.268}$$

where δ_Δ is an arbitrary discrete Laplacian operating over the grid of interest, and possibly having free parameters, but it must be at least second-order accurate. Note that for $\zeta = 0$ we have the canonical two-step explicit scheme (4.20). For $\zeta \neq 0$, this implicit scheme will generally require four states in memory: the three states ($\underline{u}_{\boldsymbol{i}}^{n+1}$, $\underline{u}_{\boldsymbol{i}}^n$, $\underline{u}_{\boldsymbol{i}}^{n-1}$) and one temporary for the use of iterative techniques, such as, e.g., Jacobi iteration [170].

Clearly, there are many possibilities for implicit schemes in the family (4.268). Since this constitutes a generalisation of two-step explicit schemes, every explicit scheme we have analysed thus far has an analogous family of implicit schemes. To narrow our investigation, we will restrict our attention to the simplest isotropic discrete Laplacian native to each grid-type considered thus far, leaving only ζ and λ as free parameters in the schemes, which will be constrained for numerical stability.

Before deriving stability conditions, we should introduce yet another equivalent form, which will be used in Chapter 6. By the identity: $\delta_{tt} = \frac{4}{k^2}(\mu_{tt} - 1)$ (and $\mu_{tt} = (e_{t+} + 2 + e_{t-})/4$), this scheme can also be written as:

$$\delta_{tt}\underline{u}_{\boldsymbol{i}}^n = c^2(\theta + (1-\theta)\mu_{tt})\delta_\Delta \underline{u}_{\boldsymbol{i}}^n \tag{4.269}$$

where $\theta = 1 + 4\zeta/\lambda^2$.

Numerical stability

For a von Neumann stability analysis, we consider a plane-wave trial solution of the form: $\underline{u}_{\boldsymbol{i}}^n = \mathrm{e}^{\mathrm{j}(n\omega_k + \boldsymbol{i}\cdot\boldsymbol{\beta}_h)}$, and after insertion into the scheme (4.268) we obtain the following dispersion relation:

$$\omega_k(\boldsymbol{\beta}_h) = 2\sin^{-1}\left(\pm\lambda\left(\frac{1}{G(\boldsymbol{\beta}_h)} - 4\zeta\right)^{-1/2}\right) \tag{4.270}$$

where $G(\boldsymbol{\beta}_h)$ is specific to the discrete Laplacian. Note that the above simplifies to (4.26) for $\zeta = 0$.

Constraining $\omega_k \in \mathbb{R}$ we obtain the stability condition:

$$\lambda \leq \lambda_{\max} = \sqrt{\frac{1}{\Lambda} - 4\zeta}\,, \quad \zeta < \frac{1}{4\Lambda} \tag{4.271}$$

where $\Lambda = \max_{\boldsymbol{\beta}_h} G(\boldsymbol{\beta}_h)$. Of course, this is just a generalisation of (4.267) for the 1-D scheme, wherein $\Lambda = 1$.

4.7.2 Fourth-order accurate and optimised implicit schemes

Let us now apply the method of modified equations to this scheme [301], in order to derived conditions for fourth-order accuracy. Subsequently we will consider optimised second-order accurate schemes. It is assumed that the discrete Laplacian is isotropic to fourth-order, i.e., that:

$$\delta_\Delta = \Delta + Ch^2\Delta^2 + \mathcal{O}(h^4) \tag{4.272}$$

where C is a coefficient that will depend on the discrete Laplacian (to be specified). Considering the discrete d'Alembertian $\delta_\Box = \delta_{tt} + \zeta h^2 \delta_\Delta \delta_{tt} - c^2\delta_\Delta$, we have for λ fixed to some value:

$$\delta_\Box \underline{u} = \Box\underline{u} + \left(\frac{k^2}{12}\partial_t^4 + \zeta\partial_t^2\Delta - Ch^2c^2\Delta^2\right)\underline{u} + \mathcal{O}(k^4) + \mathcal{O}(h^4) \tag{4.273a}$$

$$= \Box\underline{u} + \frac{h^2}{c^2}\underbrace{\left(\frac{\lambda^2}{12}\partial_t^4 + \zeta c^2\partial_t^2\Delta - Cc^4\Delta^2\right)}_{F}\underline{u} + \mathcal{O}(h^4) \tag{4.273b}$$

and we note that the term labelled F can be factored into a product of two terms plus a residual:

$$F = \frac{1}{12}\left(\left(\partial_t^2 - c^2\Delta\right)\left(\lambda^2\partial_t^2 + \lambda^2c^2\Delta + 12\zeta c^2\Delta\right) + \left(\lambda^2 + 12\zeta - 12C\right)\Delta^2\right) \tag{4.274}$$

We can eliminate the residual (last term on the right) with the constraint:

$$\lambda^2 = 12C - 12\zeta \tag{4.275}$$

Then we have:

$$\delta_\Box \underline{u} = \Box(1 + \mathcal{O}(h^2))\underline{u} + \mathcal{O}(h^4) \tag{4.276a}$$

$$= \Box\underline{u} + \mathcal{O}(h^4) \tag{4.276b}$$

Thus, (4.275) is our condition for a fourth-order local truncation error in the scheme. It remains to check if this condition can be satisfied for a stable λ, which amounts to checking C and Λ for each scheme under consideration. We expect $\lambda = \lambda_{\max}$ to be optimal for these $(4,4)$-accurate schemes, so it suffices to check the intersection of (4.275) and (4.271), which for $\lambda = \lambda_{\max}$ can be written as:

$$\zeta = \zeta^\star := \frac{3}{2}C - \frac{1}{8}\Lambda^{-1} \tag{4.277}$$

and the stability condition under this choice of ζ and λ is only that $4C\Lambda \leq 1$. Thus, the symbol $\zeta^\star$ represents the optimal coefficient ζ for a fourth-order accurate implicit scheme (with $\lambda = \lambda_{\max}$). We list these parameters in Table 4.19 for the various schemes under consideration. It can be seen that stability is ensured ($4C\Lambda < 1$) for all of these optimal fourth-order schemes.

Table 4.19: Parameters of implicit schemes using $(K+1)$-point isotropic discrete Laplacians

d	grid	K	C	Λ^{-1}	$\zeta^\star$	$\zeta_{1\%\text{-opt.}}$
2	Cartesian (square)	8	$1/12$	$3/4$	$1/32$	0.0466
2	hexagonal	6	$1/16$	$2/3$	$1/96$	0.0202
3	Cartesian (cubic)	18	$1/12$	$3/4$	$1/32$	0.0466
3	face-centered cubic	18	$1/9$	$12/11$	$1/33$	0.0460
3	body-centered cubic	14	$1/6$	$3/2$	$1/16$	0.0935

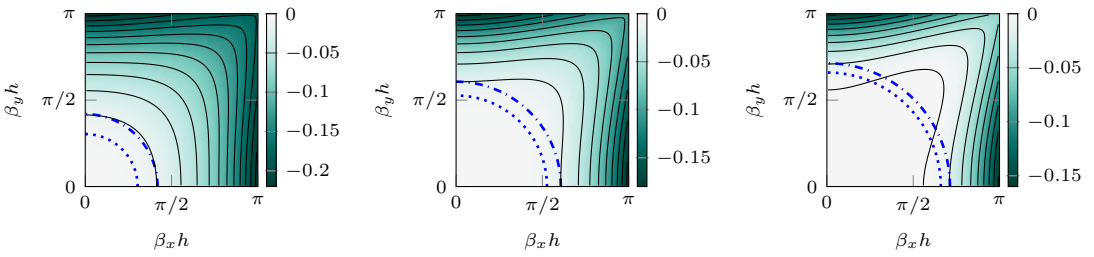
In the last column of Table 4.19, labelled $\zeta_{1\%\text{-opt.}}$, are parameters for the implicit schemes which are optimised for a one-percent critical wavenumber, also taking into account spatiotemporal densities. These optimised values have been truncated to three decimal places, for brevity, and they were found by simple means (the non-linear optimisation is only over one free parameter).

The relative dispersion errors for the optimal fourth-order accurate implicit schemes and the optimised second-order accurate implicit schemes for each respective grid-type are shown in Figs. 4.90–4.94. The isotropic explicit counterparts, which appeared previously in this chapter, are also displayed. In each case, it is clear that improvements in the numerical dispersion are obtained from the explicit case ($\zeta = 0$) to the fourth-order implicit case ($\zeta = \zeta^\star$), and additional improvements are obtained in the one- and two-percent critical wavenumbers with the optimised implicit cases ($\zeta = \zeta_{1\%\text{-opt.}}$).

Towards the goal of identifying the best among these choices of schemes in either 2-D or 3-D, and to compare with explicit schemes, relative computational efficiencies could also be calculated. These will be left out for brevity, although it should be remember that one must also somehow take into account the cost of linear system solutions required at each time-step by such implicit schemes. In all of these schemes, the "implicitness" (measured, e.g., by the magnitude of ζ) is rather small, and it turns out that simple iterative techniques like Jacobi's method [170] work rather well and lead to parallelisable updates—see, e.g., related work from the current author on this topic considering specifically the 3-D Cartesian 19-point implicit scheme [100].

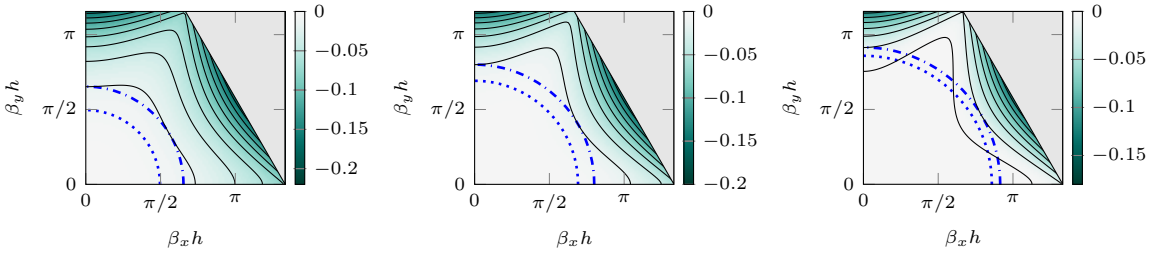
4.8 Summary

In this chapter we have given a detailed overview of finite difference schemes for the wave equation in free space. We considered Cartesian and non-Cartesian (hexagonal, FCC, BCC) grids, using explicit and implicit time integration to derive schemes with high orders of global accuracy and schemes optimised for wideband accuracy. Numerical dispersion was analysed for these schemes taking into account important details related to the use of non-Cartesian (lattice-based) grids, not previously considered in the room acoustics literature. The various schemes presented in this chapter were compared in terms of dispersion errors, taking into account measures of computational costs, and through numerical simulations. Many of the schemes presented in this chapter, particular the high-order accurate and optimised schemes, demonstrate significant improvements in terms of numerical dispersion errors over more conventional schemes found in the literature.



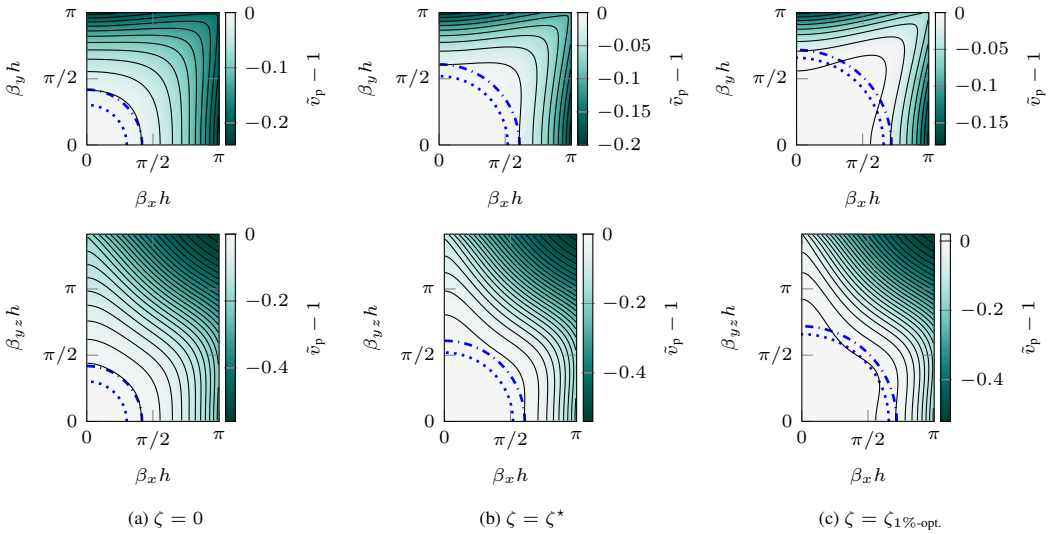
(a) $\zeta = 0$ (b) $\zeta = \zeta^\star$ (c) $\zeta = \zeta_{1\%\text{-opt.}}$

Figure 4.90: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for nine-point implicit/explicit isotropic 2-D schemes with ζ as indicated and $\lambda = \lambda_{\max} = \sqrt{3/4 - 4\zeta}$. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$. The scheme with $\zeta = \zeta^\star$ is fourth-order accurate.



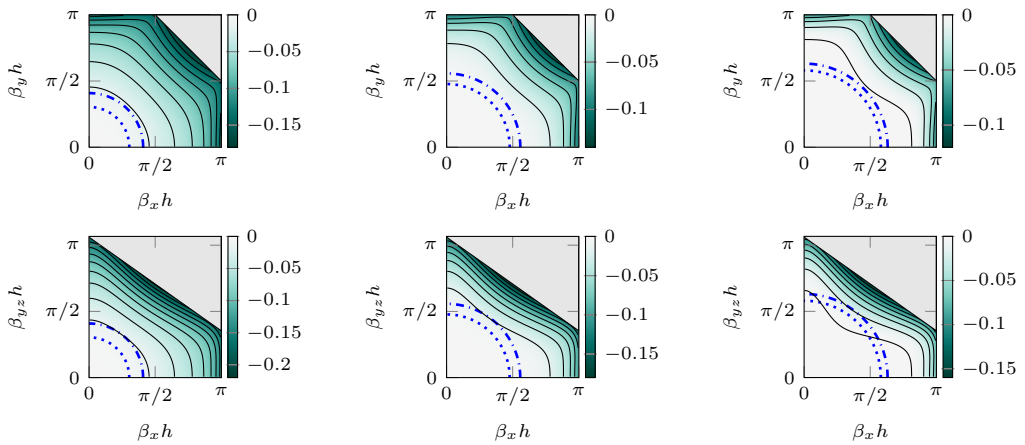
(a) $\zeta = 0$ (b) $\zeta = \zeta^\star$ (c) $\zeta = \zeta_{1\%\text{-opt.}}$

Figure 4.91: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for implicit/explicit seven-point (isotropic) 2-D schemes on the hexagonal grid, with ζ as indicated and $\lambda = \lambda_{\max} = \sqrt{2/3 - 4\zeta}$, for $\boldsymbol{\beta}_h$ in the positive quadrant of $\mathbb{B}_\mathrm{H}$. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$. The scheme with $\zeta = \zeta^\star$ is fourth-order accurate.



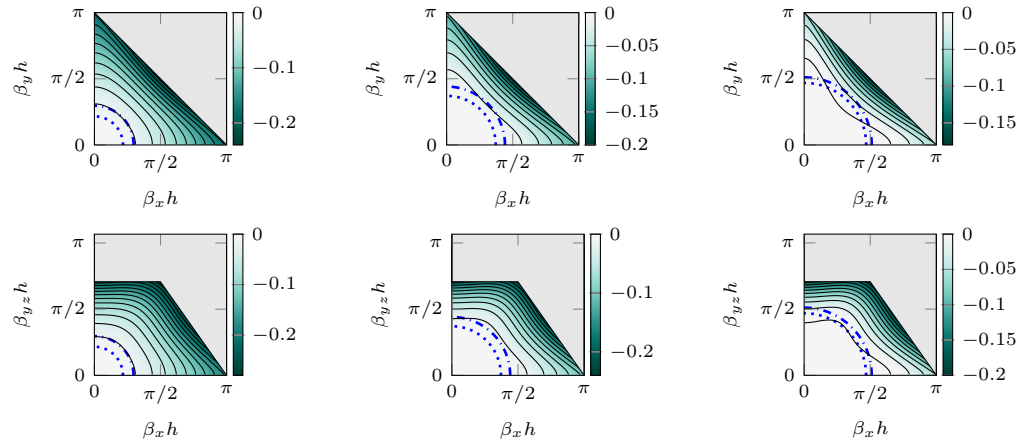
(a) $\zeta = 0$ (b) $\zeta = \zeta^\star$ (c) $\zeta = \zeta_{1\%\text{-opt.}}$

Figure 4.92: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for implicit/explicit 19-point isotropic 3-D Cartesian schemes, with ζ as indicated and $\lambda = \lambda_{\max} = \sqrt{3/4 - 4\zeta}$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$. The scheme with $\zeta = \zeta^\star$ is fourth-order accurate.



(a) $\zeta = 0$ (b) $\zeta = \zeta^\star$ (c) $\zeta = \zeta_{1\%\text{-opt.}}$

Figure 4.93: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for implicit/explicit 19-point isotropic 3-D schemes on FCC grids, with ζ as indicated and $\lambda = \lambda_{\max} = \sqrt{12/11 - 4\zeta}$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$. The grey shaded region is outside of the limits of the respective wavenumber cells. The scheme with $\zeta = \zeta^\star$ is fourth-order accurate..



(a) $\zeta = 0$ (b) $\zeta = \zeta^\star$ (c) $\zeta = \zeta_{1\%\text{-opt.}}$

Figure 4.94: Contour plots of $\tilde{v}_\mathrm{p}(\boldsymbol{\beta}_h) - 1$ for implicit/explicit 15-point isotropic 3-D schemes schemes on the BCC grid, with ζ as indicated and $\lambda = \lambda_{\max} = \sqrt{3/2 - 4\zeta}$. Top row: 2-D slice through $\beta_z = 0$ plane. Bottom row: 2-D slice through $\beta_y = \beta_z$ (side-diagonal) plane. Contour lines denote 2% increments. The dashed-dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,2\%}$ and the dotted line denotes $|\boldsymbol{\beta}_h| = \beta_{h,1\%}$. The grey shaded region is outside of the limits of the respective wavenumber cells. The scheme with $\zeta = \zeta^\star$ is fourth-order accurate.

Chapter 5

Reinvestigating finite difference boundary conditions for room acoustics

In this chapter we revisit boundary conditions for the simplest Cartesian scheme, investigating them in more detail using energetic and matrix analyses for numerical stability. The purpose of this chapter is to identify and discuss some of the challenges in the derivation of stable boundary conditions in the simplest scheme—challenges which carry on to more general schemes for room acoustics.

We start this chapter by rewriting the simplest scheme in d spatial dimensions with boundary conditions that are of the "non-centered" type (using operators of the form $\delta_{x\pm}$) derived from an energy analysis, followed by a modified energy analysis that leads to "centered" boundary conditions (using operators of the form $\delta_{x\cdot}$) over flat-walls and box domains aligned with grid axes. This is followed by a comparison of centered and non-centered boundary types, in terms of their accuracies in approximating modal frequencies and wall reflection coefficients. Next, we investigate the problems that can arise when naively extending centered conditions to more general domains (beyond flat walls and boxes), ultimately revealing that some seemingly harmless approaches to finite difference boundary modelling can lead to instabilities. These analyses will provide motivation for the use of finite volume methods in Chapter 6.

In this chapter we consider the wave equation and Cartesian grid in d spatial dimensions, calling for a slightly different notation than in previous chapters. We will write $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ to represent the spatial position vector,[1] and $u = u(\mathbf{x}, t)$ will be the dependent variable of interest (a scalar field), which could represent, e.g., a pressure or an acoustic velocity potential. The Laplacian operator in d dimensions will be expressed as:

$$\Delta = \sum_{j=1}^{d} \partial_{x_j}^2 \tag{5.1}$$

and the initial value problem expressed over the interior domain $\Omega \subseteq \mathbb{R}^d$ will be:

$$u_{tt} = c^2 \Delta u \qquad t > 0\,, \quad \mathbf{x} \in \Omega \tag{5.2a}$$

$$u = f_0\,,\; u_t = g_0 \qquad t = 0\,, \quad \mathbf{x} \in \Omega \tag{5.2b}$$

[1] For $d = 2$ and $d = 3$ we may also use the conventions: $x_1 = x$, $x_2 = y$ and $x_3 = z$.

In terms of boundary conditions over the boundary $\Gamma = \partial\Omega$, we consider three simple types (for $\mathbf{x} \in \Gamma$):

$$u = 0 \qquad \text{(lossless, homogeneous Dirichlet)} \tag{5.3a}$$

$$\hat{\mathbf{n}} \cdot \nabla u = 0 \qquad \text{(lossless, homogeneous Neumann)} \tag{5.3b}$$

$$\hat{\mathbf{n}} \cdot \nabla u + \frac{\gamma}{c} u_t = 0 \qquad \text{(lossy, frequency-independent)} \tag{5.3c}$$

where $\hat{\mathbf{n}}$ is the outward normal vector over the boundary of the domain and $\gamma \in \mathbb{R}, \gamma = \gamma(\mathbf{x}) \geq 0$. We will refer to (5.3a) and (5.3b) simply as "Dirichlet" and "Neumann" conditions, respectively, and they represent "soft" and rigid walls in the case that u is an acoustic velocity potential or pressure.[2] The third boundary condition, (5.3c), represents a frequency-independent lossy condition. The variable γ is defined as a *specific wall admittance* (and $1/\gamma$ a *specific wall impedance*) in the case that u represents an acoustic velocity potential (it can also be interpreted as such when u represents a pressure, as in [132]). The focus will mostly be on the latter two boundary conditions ((5.3b) and (5.3c)), since these form the basis for room acoustics boundary models, and incidentally, constitute some of the main challenges in finite difference boundary modelling. Finally, we note that for the numerical examples presented in this chapter we will omit units of metres and seconds for h, k, c for brevity.

5.1 Preliminaries

5.1.1 The simplest scheme in d dimensions

We begin by quickly revisiting the simplest scheme with a notation in d dimensions (repeating some of the preliminaries in Section 3.1.1, but, here, generalised to d dimensions), formulated first on the full space, and to be followed by closed domains. We use the Cartesian grid function $\underline{u}^n_{\boldsymbol{i}} \approxeq u(\boldsymbol{i}h, nk)$ where $\boldsymbol{i} \in \mathbb{R}^d$ and we define the infinite Cartesian grid $\mathbb{Z}^d_h := \{\boldsymbol{i}h : \boldsymbol{i} \in \mathbb{Z}^d\}$. Next we define the spatial shift operator:

$$e_{x_j \pm} \underline{u}^n_{\boldsymbol{i}} := \underline{u}^n_{\boldsymbol{i} \pm \hat{\mathbf{e}}_j} \tag{5.4}$$

where $\hat{\mathbf{e}}_j$ are the standard unit vectors in $\mathbb{R}^d$. From these spatial shift operators we can construct the spatial difference operators:

$$\delta_{x_j \pm} := \pm \frac{e_{x_j \pm} - 1}{h}, \quad \delta_{x_j \cdot} := \frac{e_{x_j +} - e_{x_j -}}{2h}, \quad \delta_{x_j x_j} := \frac{e_{x_j +} - 2 + e_{x_j -}}{h^2} \tag{5.5}$$

The simplest $(2d+1)$-point discrete Laplacian in d dimensions can then be written as:

$$\delta_\Delta := \sum_{j=1}^{d} \delta_{x_j x_j} \tag{5.6}$$

For convenience, let us also define $\delta^h_\Delta = h^2 \delta_\Delta$, and $Q^n_{\boldsymbol{i}}$ as:

$$Q^n_{\boldsymbol{i}} = \sum_{j=1}^{d} (\underline{u}^n_{\boldsymbol{i} + \hat{\mathbf{e}}_j} + \underline{u}^n_{\boldsymbol{i} - \hat{\mathbf{e}}_j}) \tag{5.7}$$

[2] In the case that u is a transverse displacement (of a string or a membrane), these conditions are known, respectively, as "fixed" and "free" conditions.

The simplest scheme, generalised to d dimensions, is then simply:

$$\delta_{tt}\underline{u}_{\boldsymbol{i}}^n = c^2\delta_\Delta \underline{u}_{\boldsymbol{i}}^n \tag{5.8}$$

and the update recursion for this scheme is:

$$\underline{u}_{\boldsymbol{i}}^{n+1} = (2 + \lambda^2\delta_\Delta^h)\underline{u}_{\boldsymbol{i}}^n - \underline{u}_{\boldsymbol{i}}^{n-1} \tag{5.9}$$

and a useful representation in practice is:

$$\underline{u}_{\boldsymbol{i}}^{n+1} = (2 - 2d\lambda^2)\underline{u}_{\boldsymbol{i}}^n - \underline{u}_{\boldsymbol{i}}^{n-1} + \lambda^2 Q_{\boldsymbol{i}} \tag{5.10}$$

In order to start the recursion from $n = 0$, the initial two states $\underline{u}_{\boldsymbol{i}}^0, \underline{u}_{\boldsymbol{i}}^1$ can be calculated from the initial conditions of the model problem and a discrete initialisation scheme (such as, e.g., (4.156)).

von Neumann stability and a matrix equation

The stability condition for this scheme is by now obvious ($\lambda \leq \sqrt{1/d}$), but let us briefly consider a von Neumann stability analysis, expressed in a form that will have some connections to matrix equations to appear throughout this chapter. As such, we consider the ansatz:

$$\underline{u}_{\boldsymbol{i}}^n = z^n \mathrm{e}^{\mathrm{j}\boldsymbol{\beta}\cdot\boldsymbol{i}h} \tag{5.11}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_d) \in \mathbb{R}^d$ is the vector wavenumber, and $z \in \mathbb{C}$. Let us write two temporal states of this trial solution as the column vector $\mathbf{w}_{\boldsymbol{i}}^n = \left[z^n \mathrm{e}^{\mathrm{j}\boldsymbol{\beta}\cdot\boldsymbol{i}h}, z^{n-1}\mathrm{e}^{\mathrm{j}\boldsymbol{\beta}\cdot\boldsymbol{i}h}\right]^{\mathrm{T}}$. The scheme, as represented by (5.9), can then be written as the following matrix equation:

$$\mathbf{w}_{\boldsymbol{i}}^{n+1} = \widehat{\mathbf{A}}\mathbf{w}_{\boldsymbol{i}}^n \tag{5.12}$$

where $\widehat{\mathbf{A}}$ is a 2×2 matrix defined as:

$$\widehat{\mathbf{A}} = \begin{bmatrix} 2 - 4\lambda^2 G(\boldsymbol{\beta}, h) & -1 \\ 1 & 0 \end{bmatrix} \tag{5.13}$$

and

$$G(\boldsymbol{\beta}, h) = \sum_{j=1}^{d} \sin^2(\beta_j h/2) \tag{5.14}$$

where this $G(\boldsymbol{\beta}, h)$ is by now familiar, and clearly, $0 \leq G(\boldsymbol{\beta}, h) \leq d$.

It follows that the state $\mathbf{w}_{\boldsymbol{i}}^n$ at any time n could be written in terms of powers of $\widehat{\mathbf{A}}$:

$$\mathbf{w}_{\boldsymbol{i}}^{n+1} = \widehat{\mathbf{A}}^n \mathbf{w}_{\boldsymbol{i}}^1 \tag{5.15}$$

Thus, the evolution of the system with respect to the initial conditions is entirely described by powers of the matrix $\widehat{\mathbf{A}}$, and any growth in $\widehat{\mathbf{A}}^n$ (in some matrix norm) translates to growth in the solution. For this reason, $\widehat{\mathbf{A}}$ can also be called the *amplification matrix* of the scheme [207].

The behaviour of $\widehat{\mathbf{A}}^n$ can be described by its eigenvalues, and the eigenvalues of $\widehat{\mathbf{A}}$, which we also

denote as z, are given by the following characteristic equation:

$$z^2 + (4\lambda^2 G(\boldsymbol{\beta}, h) - 2)z + 1 = 0 \tag{5.16}$$

which is identically derived by substituting our trial solution directly into (5.9). Restricting the eigenvalues of this matrix to $|z| \leq 1$, in order to limit exponential growth in the solution, leads to the conditions $0 \leq \lambda^2 G(\boldsymbol{\beta}, h) \leq 1$, which leads to the well-known von Neumann stability condition [86]:

$$\lambda \leq \lambda_{\max} = \sqrt{1/d} \tag{5.17}$$

5.1.2 Matrix methods for numerical stability analysis

In order to investigate some of the challenges in ensuring numerical stability under boundary termination in such finite difference schemes, it will help to represent the finite difference update equations under analysis as *matrix equations*. In a manner analogous to the treatment of the amplification matrix $\widehat{\mathbf{A}}$, the eigenvalues of the matrices involved in these equations can be analysed to gain insights into numerical stability, and in some cases, derive instructive stability conditions.

For the remainder of this chapter, we consider the wave equation on a closed domain $\overline{\Omega} = (\Omega \cup \Gamma) \subset \mathbb{R}^d$ where Ω is the interior and Γ the boundary, and we consider our grid function $\underline{u}_{\boldsymbol{i}}^n$ on the finite Cartesian grid and associated index set, respectively defined as:

$$\mathbb{G}_h := \mathbb{Z}_h^d \cap \overline{\Omega}\,, \quad \mathcal{I} := \{\boldsymbol{i} \in \mathbb{Z}^d : \boldsymbol{i}h \in \overline{\Omega}\} \tag{5.18}$$

Matrix update equations

In the case of a finite spatial grid, it is always possible to represent $\underline{u}_{\boldsymbol{i}}^n$ as a column vector through a *vectorisation* of $\underline{u}_{\boldsymbol{i}}^n$. For now, we are not concerned with the chosen vectorisation, since the set of eigenvalues of an associated matrix is insensitive to this choice. For any given vectorisation, we can write $\underline{u}_{\boldsymbol{i}}^n$ as a $M \times 1$ column vector $\mathbf{u}^n$, such that the mth element ($m = 0, \ldots, M-1$) of $\mathbf{u}^n$ corresponds to a unique grid point labelled $\boldsymbol{i}_m$ and $\boldsymbol{i}_m \in \mathcal{I}$ and $M = |\mathcal{I}|$. In the case of Dirichlet conditions, boundary points are typically removed so we can also have $\boldsymbol{i}_m \in \mathcal{I}_{\mathrm{i}}$ and $M = |\mathcal{I}_{\mathrm{i}}|$, where $\mathcal{I}_{\mathrm{i}}$ represents the set of regular interior points (to be defined shortly, in Section 5.2).

Under boundary conditions of the three types considered in this chapter (i.e., (5.3)), we will be able to represent the finite difference update for the entire scheme as:

$$\mathbf{u}^{n+1} = (\lambda\mathbf{B} + \mathbf{I})^{-1} \left((2\mathbf{I} + \lambda^2 \mathbf{L})\mathbf{u}^n + (\lambda\mathbf{B} - \mathbf{I})\mathbf{u}^{n-1} \right) \tag{5.19}$$

Here, the matrices of size $M \times M$ are as follows: $\mathbf{I}$ is the identity matrix, $\mathbf{B}$ is a sparse diagonal matrix with non-negative entries pertaining to boundary losses, and $\mathbf{L}$ is a sparse matrix that represents the normalised Laplacian operator δ_Δ^h after boundary conditions have been imposed. Specific constructions for the matrices $\mathbf{L}$ and $\mathbf{B}$ will be specified later.

In the case of lossless boundary conditions ($\mathbf{B} = \mathbf{0}$, where $\mathbf{0}$ is a matrix of zeros), the update reduces to a simpler form:

$$\mathbf{u}^{n+1} = (2\mathbf{I} + \lambda^2 \mathbf{L})\mathbf{u}^n - \mathbf{u}^{n-1} \tag{5.20}$$

which is essentially the matrix representation of (5.9) on a closed domain (with Neumann or Dirichlet boundary conditions).

It is worth also considering the related one-step state-space form. Consider the concatenation of the grid function over two time-steps $\mathbf{w}^n = [\mathbf{u}^n, \mathbf{u}^{n-1}]^{\mathrm{T}}$. The update equation can then be written as:

$$\mathbf{w}^{n+1} = \mathbf{A}\mathbf{w}^n \tag{5.21}$$

where $\mathbf{A}$ is a state-space matrix which takes the form:

$$\mathbf{A} = \begin{bmatrix} (\lambda\mathbf{B} + \mathbf{I})^{-1}(2\mathbf{I} + \lambda^2\mathbf{L}) & (\lambda\mathbf{B} + \mathbf{I})^{-1}(\lambda\mathbf{B} - \mathbf{I}) \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \tag{5.22}$$

and in the lossless case we have simply:

$$\mathbf{A} = \begin{bmatrix} 2\mathbf{I} + \lambda^2\mathbf{L} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \tag{5.23}$$

It follows that any future state can be calculated from the initial state and powers of the matrix $\mathbf{A}$ via:

$$\mathbf{w}^{n+1} = \mathbf{A}^n\mathbf{w}^1 \tag{5.24}$$

where $\mathbf{w}^1$ represents the initial state, derived from the initial conditions. Thus, as in the analogous frequency-domain matrix equation (5.15), the evolution of the system with respect to the initial conditions is entirely described by powers of the matrix $\mathbf{A}$.

Matrix method for stability using the one-step space form

From the matrix update equations (5.20) and (5.21), conditions for stability can be derived through what can be called "matrix methods" for stability analysis [207, 244]. The one-step update (5.21) leads to a *weak* stability condition for the boundary-value problem [207, 244]:

$$\|\mathbf{A}\| \leq 1 \tag{5.25}$$

where $\|\cdot\|$ is an appropriate matrix norm. It is true that $\hat{\rho}(\mathbf{A}) \leq \|\mathbf{A}\|$ under any matrix norm [244, 170], thus a sufficient condition for weak stability is [260]:

$$\hat{\rho}(\mathbf{A}) \leq 1 \tag{5.26}$$

where $\hat{\rho}(\mathbf{A})$ denotes the spectral radius of $\mathbf{A}$ (the maximum eigenvalue of $\mathbf{A}$ in magnitude).

It is important to remark that, in general, the above condition should be seen more as a "stability check" rather than as a "stability condition", since, as explained in [260], this condition on the matrix $\mathbf{A}$ provides no guidance on choosing appropriate boundary conditions for numerical stability. Also, it provides little to no guidance on setting the Courant number, or on dealing with other types of instabilities that may be experienced. Furthermore, estimating $\hat{\rho}(\mathbf{A})$ is not feasible in practice (for large-scale room acoustics simulations) since eigenvalues for this update matrix tend to be distributed evenly around or inside the unit circle in the complex plane (such as, e.g., in the lossless case), so one generally has to solve for the *complete* spectrum of $\mathbf{A}$ in order to identify the largest eigenvalue in magnitude.[3]

[3] On the other hand, it can be useful to analyse the spectrum of $\mathbf{A}$ for smaller grid sizes in order to derive insights into some of the effects of truncation errors in finite precision arithmetic for these types of schemes, as shown in [44].

Matrix method for stability using the two-step lossless form

In the case of lossless (Dirichlet or Neumann) boundary conditions, we can arrive at more instructive stability conditions from (5.26), ultimately focussing on the eigenvalues of the matrix $\mathbf{L}$ [244].

Consider a test solution $\mathbf{u}^n = \boldsymbol{\phi} z^n$, $z^n \in \mathbb{C}$, where $\boldsymbol{\phi}$ is an eigenvector of $\mathbf{L}$, i.e.,

$$\mathbf{L}\boldsymbol{\phi} = \mu\boldsymbol{\phi} \tag{5.27}$$

and $\mu \in \sigma(\mathbf{L})$ is an eigenvalue of $\mathbf{L}$, and $\sigma(\mathbf{L})$ denotes the *spectrum* (the set of eigenvalues) of $\mathbf{L}$. Inserting this trial solution into (5.20) leads to a characteristic equation in z (similar to (5.16)):

$$\lambda^2\mathbf{L}\boldsymbol{\phi} = (z - 2 + z^{-1})\boldsymbol{\phi} \tag{5.28}$$

Letting $z = \mathrm{e}^{\mathrm{j}\omega nk}$ we have:

$$\lambda^2\mathbf{L}\boldsymbol{\phi} = -4\sin^2(\omega k/2)\boldsymbol{\phi} \tag{5.29}$$

and comparing with (5.27) we have:

$$-4\sin^2(\omega k/2) = \lambda^2\mu \quad \Longrightarrow \quad \omega k = 2\sin^{-1}\left(\lambda\sqrt{-\mu/4}\right) \tag{5.30}$$

For this second-order scheme to be stable, we require that $\omega \in \mathbb{R}$ (z on the unit circle).[4] By recognising that the inverse sine function $\sin^{-1}(x)$ is real only for the real interval $x \in [-1, 1]$, we arrive at conditions on the spectrum of $\mathbf{L}$:

$$\Im\{\mu\} = 0\,, \quad \mu \in \sigma(\mathbf{L}) \tag{5.31a}$$

$$\mu \leq 0\,, \quad \mu \in \sigma(\mathbf{L}) \tag{5.31b}$$

which we will respectively rewrite (under a slight abuse of notation) throughout this chapter as:

$$\Im\{\sigma(\mathbf{L})\} = 0\,, \qquad \sigma(\mathbf{L}) \leq 0 \tag{5.32}$$

Also, we have the following condition on the Courant number:

$$\lambda \leq \frac{2}{\sqrt{\hat{\rho}(\mathbf{L})}} \tag{5.33}$$

It is important to note that due to the nature of $\sigma(\mathbf{L})$, the estimation of $\hat{\rho}(\mathbf{L})$ is significantly simpler, whether by hand or by numerical means (e.g., using [149]), than the estimation of $\hat{\rho}(\mathbf{A})$.

In order to see how these conditions relate to stability conditions typically derived from a von Neumann analysis (in the free-space case) we can rewrite (5.31b) and (5.33) as:

$$-4 \leq \lambda^2\mu \leq 0 \quad \Longrightarrow \quad 0 \leq (-\mu/4)\lambda^2 \leq 1\,, \quad \mu \in \sigma(\mathbf{L}) \tag{5.34}$$

Thus, the function $G(\boldsymbol{\beta}, h)$ that appears in a von Neumann analysis is analogous to $-\frac{1}{4}\sigma(\mathbf{L})$. On the other hand, (5.31a) is a condition that does not typically appear in von Neumann analyses, since it is often trivially the case that the frequency-domain version of a discrete Laplacian ($-\frac{4}{h^2}G(\boldsymbol{\beta}, h)$ for this scheme) is real. For the discrete boundary value problem, this is a condition that can be overlooked when

[4] We should note that z is also an eigenvalue of $\mathbf{A}$, and essentially we want to bound $|z| \leq 1$, which implies that we are bounding $\hat{\rho}(\mathbf{A}) \leq 1$, but since z appears in pairwise products equalling one, we require $|z| = 1$.

$\mathbf{L}$ is constructed as a symmetric matrix, as is often the case in 1-D and simple geometrical domains in 2-D and 3-D [207, 171, 260, 24]. However, as will be seen in Section 5.4, there are some commonly-used approaches to finite difference boundary modelling in the literature for which the resulting $\mathbf{L}$ is not necessarily symmetric. In such cases, this condition can be violated, leading to instabilities under Neumann and lossy conditions.

5.2 Deriving boundary conditions through energy methods

In this section we derive boundary conditions for the simplest scheme in d dimensions using energy methods, first using a simple inner product that leads to boundary conditions involving non-centered spatial differences. This is simply a d-dimensional generalisation of the frequency-independent case that appears in Section 3.2, so the derivation will be kept brief. Subsequently, we consider a slightly more involved energy analysis—closely-related to the finite volume approaches to appear in Chapter 6—using a modified inner product over a box-shaped domain that leads to boundary conditions employing spatial differences that are centered spatially about boundary points.

Indicator functions and some index sets

In a similar fashion to Chapter 3, let us define indicator functions for this analysis:

$$q_{\boldsymbol{i},j\pm} = \begin{cases} 1 & \text{if } \boldsymbol{i} \pm \hat{\mathbf{e}}_j \in \mathcal{I} \\ 0 & \text{otherwise} \end{cases}, \quad \overline{q}_{\boldsymbol{i},j\pm} = \begin{cases} 0 & \text{if } \boldsymbol{i} \pm \hat{\mathbf{e}}_j \in \mathcal{I} \\ 1 & \text{otherwise} \end{cases} \tag{5.35}$$

with the symmetry property $q_{\boldsymbol{i},j\pm} = e_{x_j \pm} q_{\boldsymbol{i},j\mp}$.

Let us also define index set $\mathcal{I}_\text{i}$ for points with all interior nearest-neighbours ("regular" interior points) and $\mathcal{I}_\text{b}$ for boundary points:

$$\mathcal{I}_\text{i} := \{\boldsymbol{i} \in \mathbb{Z}^d : q_{x_1\pm,\boldsymbol{i}} = \cdots = q_{x_d\pm,\boldsymbol{i}} = 1\} \tag{5.36a}$$

$$\mathcal{I}_\text{b} := \mathcal{I} \setminus \mathcal{I}_\text{i} \tag{5.36b}$$

An example non-trivial 2-D domain, illustrating these index sets and $\mathcal{I}$, is shown in Fig. 5.1.

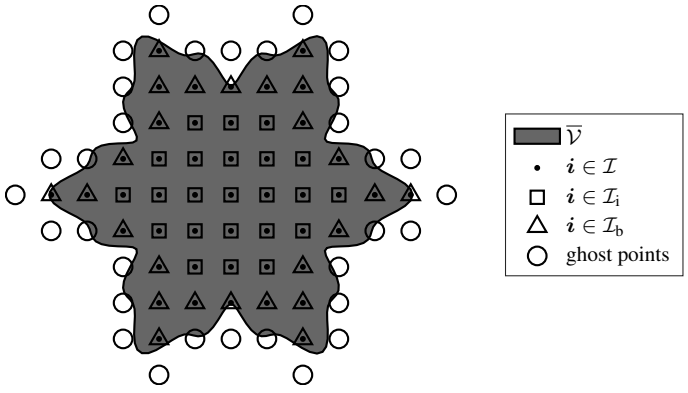
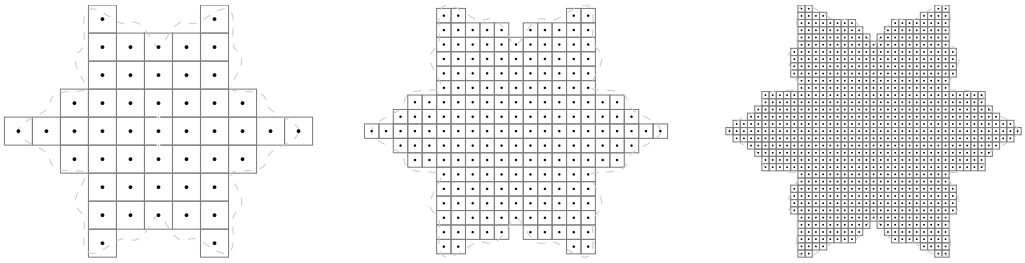


Figure 5.1: Illustration of 2-D Cartesian grid on non-trivial closed domain. This domain is defined in polar coordinates (r, θ) by the polar equation: $r = 4 + \cos^3(6\theta)$ for $\theta = [0, 2\pi]$. The above grid is discretised with $h = 1$.

It is also worth defining here a continuous staircase region $\overline{\Omega}_h$ associated to the finite grid:

$$\overline{\Omega}_h := \bigcup_{\boldsymbol{i}\in\mathcal{I}} \{\boldsymbol{i}h + \mathbf{x}' : \mathbf{x}' \in \mathbb{R}^d, |\mathbf{x}' \cdot \hat{\mathbf{e}}_j| \leq h/2, j = 1, \ldots, d\} \tag{5.37}$$

which is simply the union of hypercubes centered about grid points inside the domain. This is also known as a *Gauss digitization* [129]. Such staircased regions are illustrated by Fig. 5.2.



(a) $\overline{\Omega}_h, h = 1$ (b) $\overline{\Omega}_h, h = 0.5$ (c) $\overline{\Omega}_h, h = 0.25$

Figure 5.2: Implied staircase geometries for the domain illustrated in Fig. 5.1 (and defined in the caption) under various grid spacings, as indicated.

Inner products and summation-by-parts properties

We define inner products on $\mathbb{G}_h$ as:

$$\langle \underline{u}^n, \underline{v}^n \rangle_{\mathbb{G}_h} := \sum_{\boldsymbol{i}\in\mathcal{I}} h^d (\underline{u}^n_{\boldsymbol{i}})(\underline{v}^n_{\boldsymbol{i}}), \quad \|\underline{u}^n\|^2_{\mathbb{G}_h} := \langle \underline{u}^n, \underline{u}^n \rangle_{\mathbb{G}_h} \tag{5.38}$$

where $\underline{u}^n_{\boldsymbol{i}}$ and $\underline{v}^n_{\boldsymbol{i}}$ are grid functions on $\mathbb{G}_h$. Also, we have the following summation-by-parts (SBP) property:

$$\langle \underline{u}^n, \delta_{x_j-}\underline{v}^n \rangle_{\mathbb{G}_h} = -\langle q_{j+}\delta_{x_j+}\underline{u}^n, \underline{v}^n \rangle_{\mathbb{G}_h} + \sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^{d-1} \left(\overline{q}_{\boldsymbol{i},j+}(\underline{u}^n_{\boldsymbol{i}})(\underline{v}^n_{\boldsymbol{i}}) - \overline{q}_{\boldsymbol{i},j-}(\underline{u}^n_{\boldsymbol{i}})(e_{x_j-}\underline{v}^n_{\boldsymbol{i}}) \right) \tag{5.39}$$

5.2.1 Energy analysis for general staircased domain

We will now present an energy analysis to arrive at various staircase boundary conditions for the simplest scheme. This analysis will mirror the 3-D case seen in Section 3.2.2, but the "energy" calculated here is not necessarily a physical energy, rather a positive semi-definite function of the state, since $\underline{u}^n_{\boldsymbol{i}}$ is just an arbitrary scalar grid function in this chapter.

Proceeding with the energy analysis, we multiply both sides of (5.8) by $\frac{1}{c^2}\delta_{t\cdot}\underline{u}^n_{\boldsymbol{i}}$, and take the inner product to obtain:

$$\frac{1}{c^2}\langle \delta_{t\cdot}\underline{u}^n, \delta_{tt}\underline{u}^n \rangle_{\mathbb{G}_h} = \sum_{j=1}^{d} \langle \delta_{t\cdot}\underline{u}^n, \delta_{x_jx_j}\underline{u}^n \rangle_{\mathbb{G}_h} \tag{5.40}$$

and using $\delta_{x_jx_j} = \delta_{x_j-}\delta_{x_j+}$, and by the SBP property (5.39), we have:

$$\frac{1}{c^2}\langle \delta_{t\cdot}\underline{u}^n, \delta_{tt}\underline{u}^n \rangle_{\mathbb{G}_h} + \sum_{j=1}^{d} \langle q_{j+}\delta_{t\cdot}\delta_{x_j+}\underline{u}^n, \delta_{x_j+}\underline{u}^n \rangle_{\mathbb{G}_h} = -\mathfrak{b}^n \tag{5.41}$$

where

$$\mathfrak{b}^n = \sum_{\boldsymbol{i}\in\mathcal{I}_\mathfrak{b}} h^{d-1} \sum_{j=1}^{d} \left(\overline{q}_{\boldsymbol{i},j-} (\delta_{t\cdot} \underline{u}_{\boldsymbol{i}}^n)(\delta_{x_j-} \underline{u}_{\boldsymbol{i}}^n) - \overline{q}_{\boldsymbol{i},j+} (\delta_{t\cdot} \underline{u}_{\boldsymbol{i}}^n)(\delta_{x_j+} \underline{u}_{\boldsymbol{i}}^n) \right) \tag{5.42}$$

and where we have made use of $\delta_{x_j\pm} = e_{x_j\pm}\delta_{x_j\mp}$. Next, using identities (3.43a) and (3.43b), we have:

$$\delta_{t+}\mathfrak{h}^n = -\mathfrak{b}^n\,, \quad \mathfrak{h}^n = \mathfrak{v}^n + \mathfrak{t}^n \tag{5.43}$$

where

$$\mathfrak{v}^n = \frac{1}{2c^2}\|\delta_{t-}\underline{u}^n\|_{\mathbb{G}_h}^2\,, \qquad \mathfrak{t}^n = \frac{1}{2}\sum_{j=1}^{d} \langle q_{j+}\delta_{x_j+}\underline{u}^n, e_{t-}\delta_{x_j+}\underline{u}^n\rangle_{\mathbb{G}_h} \tag{5.44}$$

and making use of identity (3.43e), we rewrite $\mathfrak{t}^n$ as:

$$\mathfrak{t}^n = \frac{1}{2}\sum_{j=1}^{d} \left(\|q_{j+}\delta_{x_j+}\mu_{t-}\underline{u}^n\|_{\mathbb{G}_h}^2 - \frac{k^2}{4}\|q_{j+}\delta_{x_j+}\delta_{t-}\underline{u}^n\|_{\mathbb{G}_h}^2 \right) \tag{5.45}$$

Looking for a condition such that $\mathfrak{h} \geq 0$, we have:

$$\mathfrak{t}^n \geq -\frac{k^2}{8}\sum_{j=1}^{d} \|q_{j+}\delta_{x_j+}\delta_{t-}\underline{u}^n\|_{\mathbb{G}_h}^2 \tag{5.46}$$

and we can make use of the following bound, which is not necessarily tight:

$$\|q_{j\pm}\delta_{x_j\pm}\underline{v}^n\|_{\mathbb{G}_h}^2 \leq \frac{4}{h^2}\|\underline{v}^n\|_{\mathbb{G}_h}^2 \quad \Longrightarrow \quad \sum_{j=1}^{d} \|q_{j\pm}\delta_{x_j\pm}\underline{v}^n\|_{\mathbb{G}_h}^2 \leq \frac{4d}{h^2}\|\underline{v}^n\|_{\mathbb{G}_h}^2 \tag{5.47}$$

which holds for any grid function $\underline{v}_{\boldsymbol{i}}^n$. This allows us to arrive at:

$$\mathfrak{t}^n \geq -\frac{dk^2}{2h^2}\|\delta_{t-}\underline{u}^n\|_{\mathbb{G}_h}^2 \tag{5.48}$$

which leads to a bound on $\mathfrak{h}^n$:

$$\mathfrak{h}^n \geq \frac{1}{2c^2}\left(1 - d\lambda^2\right)\|\delta_{t-}\underline{u}^n\|_{\mathbb{G}_h}^2 \tag{5.49}$$

and the usual condition on the Courant number:

$$\lambda \leq \frac{1}{\sqrt{d}} \quad \Longrightarrow \quad \mathfrak{h}^n \geq 0 \tag{5.50}$$

With regards to the boundary term, we will only consider frequency-independent boundary conditions in this chapter, so we rewrite $\mathfrak{b} = \mathfrak{q}$ where $\mathfrak{q}$ is a power outflow as there will be no stored energy on the boundary. As such, we additionally require $\mathfrak{q} \geq 0$ for the system to be stable as a whole. A sufficient condition is then to have (concretely) passive boundary conditions, which simply means that each term in the summation of $\mathfrak{b}$ remains non-negative. Thus, we need that:

$$\mp\,(\delta_{t\cdot}\underline{u}_{\boldsymbol{i}}^n)(\delta_{x_j\pm}\underline{u}_{\boldsymbol{i}}^n) \geq 0\,, \quad \boldsymbol{i} \in \mathcal{I}_\mathfrak{b}\,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},j\pm} = 1 \tag{5.51}$$

From here it is straightforward to arrive at various boundary conditions that will remain stable. A few

examples are:

$$\underline{u}^n_{\boldsymbol{i}} = 0\,, \qquad \boldsymbol{i} \in \mathcal{I}_\mathrm{b} \qquad \text{(Dirichlet)} \tag{5.52a}$$

$$\delta_{x_j \pm} \underline{u}^n_{\boldsymbol{i}} = 0\,, \qquad \boldsymbol{i} \in \mathcal{I}_\mathrm{b}\,, (\text{when } \overline{q}_{\boldsymbol{i},j\pm} = 1) \qquad \text{(Neumann)} \tag{5.52b}$$

$$\mp \delta_{x_j \pm} \underline{u}^n_{\boldsymbol{i}} = \frac{\gamma_{\boldsymbol{i},j\pm}}{c} \delta_{t\cdot} \underline{u}^n_{\boldsymbol{i}}\,, \quad \boldsymbol{i} \in \mathcal{I}_\mathrm{b}\,, (\text{when } \overline{q}_{\boldsymbol{i},j\pm} = 1)\,, \gamma_{\boldsymbol{i},j\pm} \geq 0 \qquad \text{(frequency-independent lossy)} \tag{5.52c}$$

where the term $\gamma_{\boldsymbol{i},j\pm}$ is simply $\gamma(\mathbf{x})$ projected onto the point $(\boldsymbol{i} \pm \hat{\mathbf{e}}_j/2)h$, and $\gamma_{\boldsymbol{i},j\pm} = 0$ when $\overline{q}_{\boldsymbol{i},j\pm} = 0$. The frequent-independent lossy condition is essentially (3.93) for $d = 3$.

The implementation of the latter two conditions is accomplished in practice by eliminating grid function values at ghost points ($\boldsymbol{i}h \notin \overline{\Omega}$) that appear in the update of a boundary node ($\boldsymbol{i} \in \mathcal{I}_\mathrm{b}$) using the following substitution:

$$\underline{u}^n_{\boldsymbol{i} \pm \hat{\mathbf{e}}_j} = \underline{u}^n_{\boldsymbol{i}} + \frac{\gamma_{\boldsymbol{i},j\pm}}{2\lambda} (\underline{u}^{n+1}_{\boldsymbol{i}} - \underline{u}^{n-1}_{\boldsymbol{i}})\,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},j\pm} = 1 \tag{5.53}$$

Substituting any of the boundary conditions (5.52) into the regular update (5.9), an update can be derived for the entire scheme that is modified according to the geometry of the domain and the chosen boundary condition. In order to express these updates, we first define:

$$K_{\boldsymbol{i}} = \sum_{j=1}^{d} (q_{\boldsymbol{i},j+} + q_{\boldsymbol{i},j-})\,, \quad \overline{K}_{\boldsymbol{i}} = 2d - K_{\boldsymbol{i}}\,, \quad \tilde{Q}^n_{\boldsymbol{i}} = \sum_{j=1}^{d} (q_{\boldsymbol{i},j+} \underline{u}^n_{\boldsymbol{i}+\hat{\mathbf{e}}_j} + q_{\boldsymbol{i},j-} \underline{u}^n_{\boldsymbol{i}-\hat{\mathbf{e}}_j}) \tag{5.54}$$

where $\overline{K}_{\boldsymbol{i}}$ represents the number of ghost-point neighbours (e.g., $\overline{K}_{\boldsymbol{i}} = 1$ at a wall node, $\overline{K}_{\boldsymbol{i}} = 2$ at a corner node in 2-D, etc.) and where $\tilde{Q}^n_{\boldsymbol{i}}$ is a partial summation whose spatial stencil varies over the grid (specialised at the boundary nodes, and $\tilde{Q}^n_{\boldsymbol{i}} = Q^n_{\boldsymbol{i}}$ for $\boldsymbol{i} \in \mathcal{I}_\mathrm{i}$). We note also that $K_{\boldsymbol{i}} < 2d$ for $\boldsymbol{i} \in \mathcal{I}_\mathrm{b}$ and $\overline{K}_{\boldsymbol{i}} = 0$ for $\boldsymbol{i} \in \mathcal{I}_\mathrm{i}$.

The update equations expressed for $\boldsymbol{i} \in \mathcal{I}$, including updates for boundary nodes ($\boldsymbol{i} \in \mathcal{I}_\mathrm{b} \subset \mathcal{I}$) under the three types of boundary conditions considered here, are then:

$$\text{(Dirichlet):} \qquad \underline{u}^{n+1}_{\boldsymbol{i}} = (2 - 2d\lambda^2) \underline{u}^n_{\boldsymbol{i}} - \underline{u}^{n-1}_{\boldsymbol{i}} + \lambda^2 \tilde{Q}^n_{\boldsymbol{i}} \tag{5.55a}$$

$$\text{(Neumann):} \qquad \underline{u}^{n+1}_{\boldsymbol{i}} = (2 - K_{\boldsymbol{i}}\lambda^2) \underline{u}^n_{\boldsymbol{i}} - \underline{u}^{n-1}_{\boldsymbol{i}} + \lambda^2 \tilde{Q}^n_{\boldsymbol{i}} \tag{5.55b}$$

$$\text{(lossy):} \qquad \underline{u}^{n+1}_{\boldsymbol{i}} = \frac{1}{1 + \lambda\gamma'_{\boldsymbol{i}}/2} \left((2 - K_{\boldsymbol{i}}\lambda^2) \underline{u}^n_{\boldsymbol{i}} + (\lambda\gamma'_{\boldsymbol{i}}/2 - 1) \underline{u}^{n-1}_{\boldsymbol{i}} + \lambda^2 \tilde{Q}^n_{\boldsymbol{i}} \right) \tag{5.55c}$$

where $\gamma'_{\boldsymbol{i}} = \sum_{j=1}^{d} (\overline{q}_{\boldsymbol{i},j+} \gamma_{\boldsymbol{i},j+} + \overline{q}_{\boldsymbol{i},j-} \gamma_{\boldsymbol{i},j-})$. Under the simplification $\gamma_{\boldsymbol{i},j\pm} \equiv \gamma_{\boldsymbol{i}}$, we have $\gamma'_{\boldsymbol{i}} = \overline{K}_{\boldsymbol{i}} \gamma_{\boldsymbol{i}}$ and the update of the lossy scheme is:

$$\underline{u}^{n+1}_{\boldsymbol{i}} = \frac{1}{1 + \lambda \overline{K}_{\boldsymbol{i}} \gamma_{\boldsymbol{i}}/2} \left((2 - \lambda^2 K_{\boldsymbol{i}}) \underline{u}^n_{\boldsymbol{i}} + (\lambda \overline{K}_{\boldsymbol{i}} \gamma_{\boldsymbol{i}}/2 - 1) \underline{u}^{n-1}_{\boldsymbol{i}} + \lambda^2 \tilde{Q}^n_{\boldsymbol{i}} \right) \tag{5.56}$$

and we note that the above update equation handles the case of Neumann conditions if we set $\gamma_{\boldsymbol{i}} \equiv 0$.

While these update equations encapsulate the entire scheme under any of the choices in (5.52), for practical purposes, it is useful to remember that for regular interior nodes (for $\boldsymbol{i} \in \mathcal{I}_\mathrm{i}$), the update of the scheme is simplified to the regular "free-space" update (5.10).

Aside – Practical use of a halo ghost points

It is worth mentioning briefly a technique for use with non-centered conditions that can simplify practical implementations [302]. First, let us define the set of ghost points $\mathcal{I}_{\mathrm{g}}$ (i.e., the set of exterior points that are nearest neighbours to boundary points) as:

$$\mathcal{I}_{\mathrm{g}} := \{\boldsymbol{i} \in \mathbb{Z}^d : \boldsymbol{i} \notin \mathbb{G}_h, \exists(\boldsymbol{i} \pm \hat{\mathbf{e}}_j) \in \mathbb{G}_h, j = 1, \ldots, d\} \tag{5.57}$$

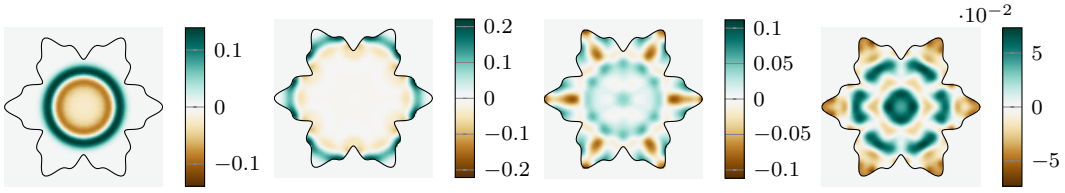
One practical way to proceed with (5.56) is to include a "halo" of fictitious ghost points in the simulation, and to set them to zero for all time, i.e., $\underline{u}_{\boldsymbol{i}}^n = 0, \boldsymbol{i} \in \mathcal{I}_{\mathrm{g}}$. Then one can make use of the modified update:

$$\underline{u}_{\boldsymbol{i}}^{n+1} = \frac{1}{1 + \lambda \overline{K}_{\boldsymbol{i}} \gamma_{\boldsymbol{i}}/2} \left((2 - K_{\boldsymbol{i}} \lambda^2) \underline{u}_{\boldsymbol{i}}^n + (\lambda \overline{K}_{\boldsymbol{i}} \gamma_{\boldsymbol{i}}/2 - 1) \underline{u}_{\boldsymbol{i}}^{n-1} + \lambda^2 Q_{\boldsymbol{i}}^n \right), \quad \boldsymbol{i} \in \mathcal{I} \tag{5.58}$$

where the simplification is that the term $\tilde{Q}_{\boldsymbol{i}}^n$ (whose stencil-shape varies over the boundary nodes) in (5.56) is replaced with the more-uniform $Q_{\boldsymbol{i}}^n$, which is a useful property for parallel implementations [302].

A numerical example

In order to demonstrate the properties of energy conservation in the scheme, we present a numerical example using the domain illustrated in Fig. 5.1, using a finer grid than what is illustrated ($h = 1$ in Fig. 5.1, and here $h = 0.1$). For initial conditions, f_0 is a Gaussian with variance 0.1 and $g_0 = 0$, and lossy boundary conditions are applied with $\gamma_{\boldsymbol{i}} = 0.1$ at boundary nodes. The evolution of the numerical solution over time is shown in Fig. 5.3.



(a) $n = 30$ (b) $n = 55$ (c) $n = 80$ (d) $n = 100$

Figure 5.3: Snapshots of simulations of 2-D wave equation $\lambda = 1/\sqrt{2}$ with f_0 a Gaussian and $g_0 = 0$, after n time-steps as indicated.

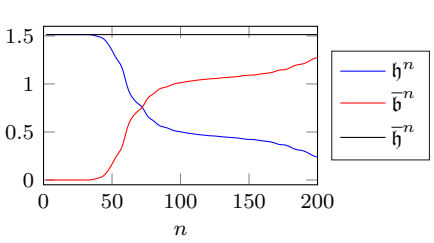
In order to demonstrate the properties of conservation of numerical energy, the energy is calculated according to (5.43) and (5.44), and the boundary term is calculated as:

$$\mathfrak{b}^n = \sum_{\boldsymbol{i} \in \mathcal{I}_{\mathfrak{b}}} h \sum_{j=1}^{2} \frac{\gamma_{\boldsymbol{i}}}{c} (\overline{q}_{\boldsymbol{i},j-} + \overline{q}_{\boldsymbol{i},j+})(\delta_{t\cdot} \underline{u}_{\boldsymbol{i}}^n)^2 \tag{5.59}$$

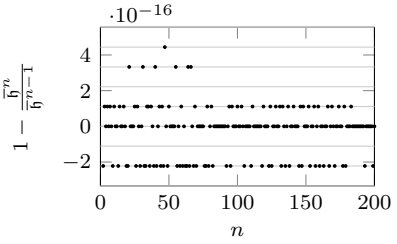
The variation of energetic quantities over time can be seen in Fig. 5.4, confirming conservation of numerical energy (including accumulated boundary losses) to machine accuracy.

Matrix constructions

At this point it is worth presenting constructions for the matrices $\mathbf{L}$ and $\mathbf{B}$ under the boundary conditions we have just derived via energy methods. For any chosen vectorisation the mth element of $\mathbf{u}^n$



(a) Energetic quantities



(b) Relative variations in numerical energy

Figure 5.4: Plots demonstrating conservation of numerical energy. The quantity $\overline{\mathfrak{h}}^n$ represented the total energy including accumulated losses from the boundary, represented by $\overline{\mathfrak{b}}^n$ ($\overline{\mathfrak{h}}^n$ and $\overline{\mathfrak{b}}^n$ defined in (3.83)). The variations in energy per time-step are integer multiples of double precision machine epsilon, as denoted by grey lines.

corresponds to $\boldsymbol{i}_m \in \mathcal{I}$. If Dirichlet conditions are employed and boundary points are removed, then $\boldsymbol{i}_m \in \mathcal{I}_\mathrm{i}$.

Under Dirichlet conditions the matrix $\mathbf{L}$ takes the form:

$$\mathbf{L}_{m_1 m_2} = \begin{cases} 1 & \text{if } \|\boldsymbol{i}_{m_1} - \boldsymbol{i}_{m_2}\| = 1 \\ 0 & \text{otherwise} \end{cases}, \quad m_1 \neq m_2 \tag{5.60a}$$

$$\mathbf{L}_{m_1 m_1} = -2d \tag{5.60b}$$

where m_1, m_2 are indices for elements of the matrix, and the notation $\mathbf{L}_{m_1 m_2}$ denotes the element of the matrix $\mathbf{L}$ in the m_1th column and m_2th row.

Under Neumann or lossy conditions, the Laplacian matrix is represented by:

$$\mathbf{L}_{m_1 m_2} = \begin{cases} 1 & \text{if } \|\boldsymbol{i}_{m_1} - \boldsymbol{i}_{m_2}\| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{5.61}$$

$$\mathbf{L}_{m_1 m_1} = -\sum_{m_2 \neq m_1} \mathbf{L}_{m_1 m_2} \tag{5.62}$$

Meanwhile, the diagonal matrix $\mathbf{B}$ takes the form:

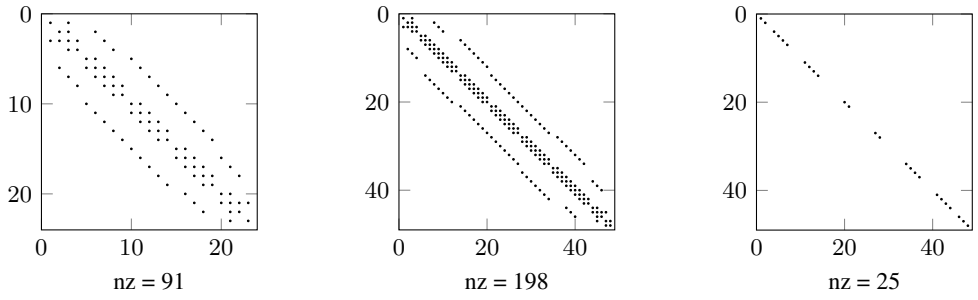
$$\mathbf{B}_{mm} = \frac{1}{2} \sum_{j=1}^{d} (\overline{q}_{\boldsymbol{i}_m, j+} \gamma_{\boldsymbol{i}_m, j+} + \overline{q}_{\boldsymbol{i}_m, j-} \gamma_{\boldsymbol{i}_m, j-}) \tag{5.63}$$

In the simplified case of one admittance value per grid point ($\gamma_{\boldsymbol{i},j\pm} = \gamma_{\boldsymbol{i}}$), we can write $\mathbf{B} = \mathbf{G}\mathbf{Q}$, where $\mathbf{G}$ is a diagonal matrix with $\gamma_{\boldsymbol{i}}$ along the diagonal and $\mathbf{Q}$ is simply:

$$\mathbf{Q}_{mm} = \frac{1}{2} \sum_{j=1}^{d} (\overline{q}_{\boldsymbol{i}_m, j+} + \overline{q}_{\boldsymbol{i}_m, j-}) \tag{5.64}$$

The sparsity patterns of these matrices are shown in Fig. 5.5 for the example 2-D domain and grid illustrated in Fig. 5.1, under one possible vectorisation of the grid function, under Dirichlet conditions (boundary points are removed), and under Neumann (or lossy) conditions.

Clearly, in all of these cases $\mathbf{L}$ is a (real) symmetric matrix, and furthermore, the application of GCT gives the bound $\hat{\rho}(\mathbf{L}) \leq 4d$, which in turn gives the sufficient stability condition $\lambda \leq \sqrt{1/d}$. We can



(a) $\mathbf{L}$, Dirichlet conditions (b) $\mathbf{L}$, Neumann or lossy conditions (c) $\mathbf{Q}$

Figure 5.5: Sparsity patterns of Laplacian matrices $\mathbf{L}$, and $\mathbf{Q}$ matrix, for irregular 2-D domain shown in Fig. 5.1. In this context, "nz" refers to the number of non-zero elements of a sparse matrix. For the matrix appearing in Fig. 5.5(a), boundary points have been removed, so it is smaller than the other two matrices appearing above.

compare these bounds to estimations of $\hat{\rho}(\mathbf{L})$ using MATLAB's `eigs` command [164] in order to see if they are tight for this choice of h and domain. For the Laplacian matrices corresponding to Fig. 5.5 (with $h = 1$) under Dirichlet and Neumann conditions, $\hat{\rho}(\mathbf{L})$ is:

$$\hat{\rho}(\mathbf{L}) \approx 7.5540 \quad \text{(Neumann)}\,, \qquad \hat{\rho}(\mathbf{L}) \approx 7.3976 \quad \text{(Dirichlet)} \tag{5.65}$$

We see that the bound obtained from GCT, $\hat{\rho}(\mathbf{L}) \leq 8$ is not tight at this grid resolution ($h = 1$) and the Courant numbers can be set slightly higher than $\lambda = \sqrt{1/2}$ for the coarse discretisation of the domain illustrated in Fig. 5.1.[5] However, it is important to note that the von Neumann bound is indeed approached as the grid is refined. For example, with the grids used in the above numerical example (with $h = 0.1$), we have:

$$\hat{\rho}(\mathbf{L}) \approx 7.9953 \quad \text{(Neumann)}\,, \qquad \hat{\rho}(\mathbf{L}) \approx 7.9951 \quad \text{(Dirichlet)} \tag{5.66}$$

For the types of 3-D domains encountered in large room acoustics simulations, such deviations generally become negligible and the von Neumann condition can be seen as being a necessary one. It is also worth pointing out that the Laplacian matrix under Dirichlet conditions tends to have a slightly smaller spectral radius than that under Neumann conditions (for the same domain and grid).

Energy analysis in matrix form

We will now rewrite this energy analysis in a matrix representation, as this will reveal some of the underlying structure of the associated matrices, and it will also provide tighter bounds for the energy-based stability conditions.

For this analysis, we assume the grid function $\underline{u}_{\boldsymbol{i}}^n$ is represented by the $M \times 1$ vector $\mathbf{u}^n$ under some

[5] This phenomenon was brought to our attention from a 3-D example in [44]; more will be said in Section 5.3.2 about the 3-D example in [44].

vectorisation (not specified). We define the matrix $\mathbf{D}_{x_j\pm}$ as:

$$\left[\mathbf{D}_{x_j\pm}\right]_{m_1m_2} = \begin{cases} \pm 1 & \text{if } \boldsymbol{i}_{m_2} - \boldsymbol{i}_{m_1} = \pm\hat{\mathbf{e}}_j \\ 0 & \text{otherwise} \end{cases}, \quad m_1 \neq m_2 \tag{5.67a}$$

$$\left[\mathbf{D}_{x_j\pm}\right]_{m_1m_1} = -\sum_{m_2\neq m_1} \left[\mathbf{D}_{x_j\pm}\right]_{m_1m_2} \tag{5.67b}$$

where $\left[\mathbf{D}_{x_j\pm}\right]_{m_1m_2}$ denotes the matrix element in row m_1 and column m_2. The significance of this matrix is that the operation $\delta_{x_j\pm}\underline{u}^n_{\boldsymbol{i}}$ applied over the interior of the grid (when $q_{\boldsymbol{i},j\pm} = 1$) will be represented by the matrix-vector product $\frac{1}{h}\mathbf{D}_{x_j\pm}\mathbf{u}^n$, under any choice of a valid vectorisation. In other words:

$$\frac{1}{h}\mathbf{D}_{x_j\pm}\mathbf{u}^n \quad \Longleftrightarrow \quad \delta_{x_j\pm}\underline{u}^n_{\boldsymbol{i}}\,, \quad \text{whenever } \; q_{\boldsymbol{i},j\pm} = 1 \text{ for } \boldsymbol{i} \in \mathcal{I} \tag{5.68}$$

In 1-D, when the grid function is decomposed as $\mathbf{u}^n = [u^n_0, u^n_1, \ldots, u^n_{M-1}]^{\mathrm{T}}$, these matrices take on a particularly simple form:

$$\mathbf{D}_{x+} = \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \\ & & & 0 \end{bmatrix}, \quad \mathbf{D}_{x-} = \begin{bmatrix} 0 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \tag{5.69}$$

and they have the properties that:

$$-\mathbf{D}^{\mathrm{T}}_{x+}\mathbf{D}_{x+} = -\mathbf{D}^{\mathrm{T}}_{x-}\mathbf{D}_{x-} = (\mathbf{D}_{x+} - \mathbf{D}_{x-}) = \mathbf{D}^{(\pm)}_{xx} := \begin{bmatrix} -1 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & 1 & -2 & 1 & \\ & & & 1 & -1 \end{bmatrix} \tag{5.70}$$

The above matrix, $\mathbf{D}^{(\pm)}_{xx}$, is well-known [171, 259]; it represents a second-order difference under "non-centered" Neumann boundary conditions.[6] It follows that analogous properties hold true in the higher dimensions, since under any vectorisation, $\mathbf{D}_{x_j\pm}$ will have the same underlying structure as (5.69) (populated with additional zero-rows). Thus, we have:

$$-\mathbf{D}^{\mathrm{T}}_{x_j+}\mathbf{D}_{x_j+} = -\mathbf{D}^{\mathrm{T}}_{x_j-}\mathbf{D}_{x_j-} = (\mathbf{D}_{x_j+} - \mathbf{D}_{x_j-}) \tag{5.71}$$

and we would have a matrix $\mathbf{D}^{(\pm)}_{x_jx_j} = (\mathbf{D}_{x_j+} - \mathbf{D}_{x_j-})$ pertaining to a second-order difference along the spatial coordinate direction x_j with non-centered Neumann conditions at boundaries.

We can now express the energetic quantities using these matrices. For convenience in the following calculations, we will drop the superscript n on $\mathbf{u}^n$ and energetic quantities. In terms of matrices and vectors, the energetic quantity $\mathfrak{v}$ can be written as:

$$\mathfrak{v} = \frac{h^d}{2c^2}\|\delta_{t-}\mathbf{u}\|^2 \tag{5.72}$$

[6] Here, the term "non-centered" refers to the discrete boundary conditions being non-centered with respect to the boundary grid points (but not necessarily with respect to Γ). This does not necessarily preclude a second-order accuracy at the boundaries.

which is clearly non-negative. Meanwhile, the quantity $\mathfrak{t}$ can be written as:

$$\mathfrak{t} = \frac{h^d}{2h^2} \sum_{j=1}^{d} \mathbf{u}^{\mathrm{T}} (\mathbf{D}_{x_j+})^{\mathrm{T}} \mathbf{D}_{x_j+} e_{t-} \mathbf{u} = \frac{h^d}{2h^2} \mathbf{u}^{\mathrm{T}} \left(\underbrace{\sum_{j=1}^{d} (\mathbf{D}_{x_j+})^{\mathrm{T}} \mathbf{D}_{x_j+}}_{\mathbf{L}^\star} \right) e_{t-} \mathbf{u} \tag{5.73}$$

The matrix that appears in this energetic term, denoted $\mathbf{L}^\star$, has a close connection to the finite difference Laplacian matrix used in the scheme (represented by $\mathbf{L}$); this will be discussed after the energy analysis.

For a bound on the sign-indeterminate $\mathfrak{t}$, we can rewrite (5.46) as:

$$\mathfrak{t} \geq -\frac{h^d k^2}{8h^2} (\delta_{t-} \mathbf{u})^{\mathrm{T}} \mathbf{L}^\star (\delta_{t-} \mathbf{u}) \tag{5.74}$$

We can then bound $\mathfrak{t}$ by:

$$\mathfrak{t} \geq -\frac{h^d k^2}{8h^2} \hat{\rho}(\mathbf{L}^\star) \|\delta_{t-} \mathbf{u}\|^2 \tag{5.75}$$

For the total numerical energy, we have the bound:

$$\mathfrak{h} \geq \frac{1}{2c^2} \left(1 - \frac{\hat{\rho}(\mathbf{L}^\star)}{4} \lambda^2 \right) \|\delta_{t-} \mathbf{u}\|^2 \tag{5.76}$$

and a stability condition on the Courant number is then:

$$\lambda \leq \frac{2}{\sqrt{\hat{\rho}(\mathbf{L}^\star)}} \quad \implies \quad \mathfrak{h} \geq 0 \tag{5.77}$$

Provided that $\mathfrak{q} \geq 0$, the above is a *sufficient* condition for stability.

At this point it is worth discussing the connection between the matrix $\mathbf{L}^\star$, which appears in (5.77), and the finite difference Laplacian matrix $\mathbf{L}$, which appears in the update of the scheme (5.19) and in the "matrix method" stability condition (5.33). In the case of Neumann or lossy conditions, it is straightforward to see that $\mathbf{L} = -\mathbf{L}^\star$ and thus, $\frac{1}{h^2}\mathbf{L}^\star$ can be seen as the negative of the Laplacian matrix approximating Δ with Neumann conditions. As such, the stability conditions (5.77) and (5.33) are equivalent in the case of Neumann boundary conditions, and condition (5.77) becomes necessary and sufficient for the Neumann problem. However, when Dirichlet conditions are employed, $\mathbf{L} \neq -\mathbf{L}^\star$, since the application of $\underline{u}_{\boldsymbol{i}}^n = 0, \boldsymbol{i} \in \mathcal{I}_\mathrm{b}$ is not taken into account in $\mathbf{L}^\star$. As such, (5.77) is only a sufficient condition for the Dirichlet problem (whereas (5.33) is necessary and sufficient), but it is slightly less restrictive than (5.50).

A convenient tool for bounding the extreme eigenvalues of the matrix $\mathbf{L}^\star$ (or $\mathbf{L}^\star$) is Gershgorin's circle theorem (GCT) [170]. Among other things, GCT says that for a matrix $\mathbf{C}$ whose eigenvalues are real, its spectral radius is lower-bounded by $\min_i \left(c_{ii} - \sum_{i \neq j} |c_{ij}| \right)$ and upper-bounded by $\max_i \left(c_{ii} + \sum_{i \neq j} |c_{ij}| \right)$, where c_{ij} are the entries of $\mathbf{C}$ [170]. Applying GCT to $\mathbf{L}^\star$ gives:

$$\hat{\rho}(\mathbf{L}^\star) \leq 4d \tag{5.78}$$

This gives the usual condition on the Courant number: $\lambda \leq 1/\sqrt{d}$. In fact, the bound (5.47) and the use of GCT are closely-related. We note that

$$\|\mathbf{D}_{x_j \pm} \mathbf{v}\|^2 = \mathbf{v}^{\mathrm{T}} (\mathbf{D}_{x_j \pm})^{\mathrm{T}} \mathbf{D}_{x_j \pm} \mathbf{v} \leq \hat{\rho} \left((\mathbf{D}_{x_j \pm})^{\mathrm{T}} \mathbf{D}_{x_j \pm} \right) \|\mathbf{v}\|^2 \tag{5.79}$$

where $\mathbf{v}$ is an arbitrary grid function $\underline{v}_{\boldsymbol{i}}^n$ in vectorised form, and by the application of GCT we have:

$$\hat{\rho}\left((\mathbf{D}_{x_j\pm})^{\mathrm{T}}\mathbf{D}_{x_j\pm}\right) \leq 4 \tag{5.80}$$

As such, we have:

$$\|\tfrac{1}{h}\mathbf{D}_{x_j\pm}\mathbf{v}\|^2 \leq \frac{4}{h^2}\|\mathbf{v}\|^2 \quad \implies \quad \sum_{j=1}^{d}\|\tfrac{1}{h}\mathbf{D}_{x_j\pm}\mathbf{v}\|^2 \leq \frac{4d}{h^2}\|\mathbf{v}\|^2 \tag{5.81}$$

which is simply (5.47) in matrix form.

Aside - connection to graph Laplacians

There is an interesting connection between the matrix $\mathbf{L}^\star$ and graph theory worth mentioning. We can interpret the Cartesian grid points in $\mathbb{G}_h$ as a graph; i.e., as network of nodes connected by edges defined by the nearest-neighbour adjacencies in the grid. The *Laplacian of a graph* [167, 87] is constructed by taking the *adjacency matrix* of the graph and subtracting it from the diagonal matrix composed of the node degrees (the number of edges per node). This is essentially the matrix construction for the matrix $-\mathbf{L}$ using (5.60), which is equivalent to $\mathbf{L}^\star$. Thus, the matrix $\mathbf{L}^\star$ constitutes a *graph Laplacian*. It is indeed from this connection to a finite difference Laplacian (and the Laplacian operator it approximates) that the graph Laplacian derives its name, and the sign-inversion is simply a convention in the graph theory literature. Each matrix $\mathbf{D}_{x_j\pm}^{\mathrm{T}}\mathbf{D}_{x_j\pm}$ is also a graph Laplacian, for a graph with adjacencies only along a given coordinate direction.

The importance of identifying this connection is that there are many known properties of Laplacian graphs that can be useful in these finite difference problems [167, 87]. For example, given some graph Laplacian $\mathbf{L}'$, some useful properties are:

$$\mathbf{v}^{\mathrm{T}}\mathbf{L}'\mathbf{v} \geq 0 \qquad \text{(positive semi-definite)} \tag{5.82a}$$

$$\min\sigma(\mathbf{L}') = 0 \qquad \text{(has zero eigenvalue)} \tag{5.82b}$$

$$\mathbf{v}^{\mathrm{T}}\mathbf{L}'\mathbf{v} = \sum_{i\sim j}(v_i - v_j)^2 \qquad \text{(a particular quadratic form)} \tag{5.82c}$$

$$\hat{\rho}(\mathbf{L}') \leq 2\left(\max_i\left|[\mathbf{L}']_{i,i}\right|\right) \qquad \text{(a bound on spectral radius via GCT)} \tag{5.82d}$$

where $\mathbf{v}$ is some vector $\mathbf{v} = (v_1, \ldots, v_M)$, and where $i \sim j$ denotes that node i is adjacent to node j in the graph associated to the matrix $\mathbf{L}'$.

5.2.2 A modified energy analysis for box-shaped domains

At this point we will consider boundary conditions beyond the "non-centered" type (i.e., employing spatial differences of the form $\delta_{x_j\pm}$), leading to more involved energy stability analyses. We start by looking at a domain where grid points might line up *exactly* on the boundary surface, such that one would be inclined to use boundary conditions of the "centered" type (i.e., employing spatial differences of the form $\delta_{x_j\cdot}$), as have been used in other boundary models proposed in the literature [133, 137]. For energy-based stability, the derivation of such conditions requires a modified energy analysis [24], which is closely-related to more general energy stability approaches to be seen in the next chapter.

For simplicity, we start by considering a box-shaped domain with dimensions that are integer multiples of the grid spacing h. More specifically, we consider a d-dimensional box-shaped domain $\mathcal{V}_{\text{box}} =$

$\prod_{j=1}^{d}(0, L_j)$ with $L_j = M_j h$ and integer M_j for $j = 1, \ldots, d$. The total number of points in the grid is thus $M = \prod_{j=1}^{d}(M_j + 1)$. An example scenario is illustrated in 2-D in Fig. 5.6. In this figure, the dotted line demarcates the domain Ω_h implied by the use of non-centered conditions, which in this case does quite not match the domain being modelled, motivating the use of boundary conditions that are centered in space about the boundary points.[7]

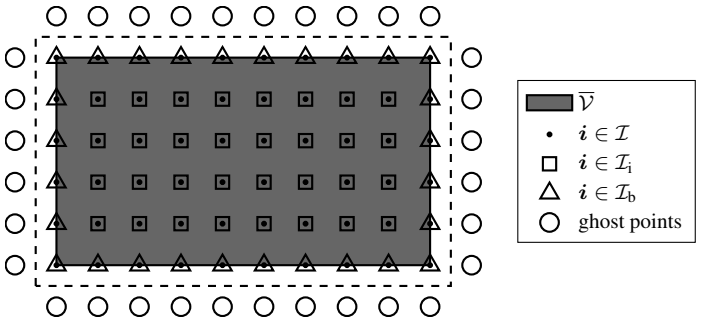


Figure 5.6: Illustration of 2-D Cartesian grid on box-shaped domain. The dotted line denotes the domain for which non-centered conditions are in fact centered about the boundary.

Since the grid points line up exactly with the boundary (i.e., $\boldsymbol{i} \in \Gamma$ for $\boldsymbol{i} \in \mathcal{I}_{\mathrm{b}}$) for the above-defined domain, it makes sense to use a weighted inner product where boundary nodes are weighted appropriately [24]:[8]

$$\langle \underline{u}^n, \underline{v}^n \rangle^{\star}_{\mathbb{G}_h} := \sum_{\boldsymbol{i} \in \mathcal{I}} h^d w_{\boldsymbol{i}} (\underline{u}^n_{\boldsymbol{i}})(\underline{v}^n_{\boldsymbol{i}}) \,, \quad \|\underline{u}^n\|^{\star 2}_{\mathbb{G}_h} := \langle \underline{u}^n, \underline{u}^n \rangle^{\star}_{\mathbb{G}_h} \tag{5.83}$$

where $w_{\boldsymbol{i}}$ and $w_{\boldsymbol{i},j}$ are weightings defined as:

$$w_{\boldsymbol{i}} = \prod_{j=1}^{d} w_{\boldsymbol{i},j} \,, \qquad w_{\boldsymbol{i},j} = \begin{cases} 1 & \text{if } (\boldsymbol{i} + \hat{\mathbf{e}}_j \text{ and } \boldsymbol{i} - \hat{\mathbf{e}}_j) \in \mathcal{I} \\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{5.84}$$

In other words, we apply a weighting factor of one-half for every ghost point neighbour to $\boldsymbol{i}$ ($w_{\boldsymbol{i}} = 1/2$ at a wall node, $w_{\boldsymbol{i}} = 1/4$ at a corner node in 2-D, $w_{\boldsymbol{i}} = 1/8$ at a corner node (with three ghost points) in 3-D, etc.). We also note that over this box-domain (but not in general), these weightings have the symmetry property:

$$w_{\boldsymbol{i}} w^{-1}_{\boldsymbol{i},j} = w_{\boldsymbol{i} \pm \hat{\mathbf{e}}_j} w^{-1}_{\boldsymbol{i} \pm \hat{\mathbf{e}}_j, j} \,, \quad (\boldsymbol{i} \pm \hat{\mathbf{e}}_j) \in \mathcal{I} \tag{5.85}$$

Under this weighted norm in 1-D, and grid point indexed by $i = 0, \ldots, M$, it is possible to derive the following SBP property [24]:

$$\sum_{i=0}^{M} h w_i (\underline{u}^n_i)(\delta_{x-} \underline{v}^n_i) = -\sum_{i=0}^{M} h q_{x+,i} (\delta_{x+} \underline{u}^n_i)(\underline{v}^n_i) + (\underline{u}^n_M)(\mu_{x-} \underline{v}^n_M) - (\underline{u}^n_0)(\mu_{x-} \underline{v}^n_0) \tag{5.86}$$

where $\mu_{x_j \pm} = \frac{1}{2}(1 \pm e_{x_j \pm})$. Extending this condition to the multi-dimensional box-shaped domain

[7] Of course, one could simply rescale h in order to have "non-centered" conditions better fit a box-shaped domain, but this is beside the point of the following analysis, which is to illustrate the energy techniques that can lead to centered conditions.

[8] The $\star$ superscript is used to distinguish this inner product from the usual inner product (5.38).

under consideration is straightforward, leading to the following SBP property:

$$\langle \underline{u}^n, \delta_{x_j-}\underline{v}^n\rangle^\star_{\mathbb{G}_h} = -\langle w_j^{-1} q_{j+}\delta_{x_j+}\underline{u}^n, \underline{v}^n\rangle^\star_{\mathbb{G}_h} + \sum_{\boldsymbol{i}\in\mathcal{I}_\mathrm{b}} w_{\boldsymbol{i}} h^{d-1} w_{\boldsymbol{i},j}^{-1} \left(\overline{q}_{\boldsymbol{i},j+}(\underline{u}^n_{\boldsymbol{i}})(\mu_{x_j-}\underline{v}^n_{\boldsymbol{i}}) - \overline{q}_{\boldsymbol{i},j-}(\underline{u}^n_{\boldsymbol{i}})(\mu_{x_j-}\underline{v}^n_{\boldsymbol{i}})\right) \tag{5.87}$$

Equipped with this SBP property we can proceed with an energy analysis. We take a weighted inner product of (5.8) with $\frac{1}{c^2}\delta_{t\cdot}\underline{u}^n_{\boldsymbol{i}}$, and after applying (5.87), we arrive at an energy balance of the form (5.43), where:

$$\mathfrak{v}^n = \frac{1}{2c^2}\|\delta_{t-}\underline{u}^n\|^{\star 2}_{\mathbb{G}_h}\,, \quad \mathfrak{t}^n = \frac{1}{2}\sum_{j=1}^{d} \underbrace{\langle w_j^{-1} q_{j+}\delta_{x_j+}\underline{u}^n, e_{t-}\delta_{x_j+}\underline{u}^n\rangle^\star_{\mathbb{G}_h}}_{\Theta_j} \tag{5.88}$$

$$\mathfrak{b}^n = \sum_{\boldsymbol{i}\in\mathcal{I}_\mathrm{b}} w_{\boldsymbol{i}} h^{d-1} \sum_{j=1}^{d} w_{\boldsymbol{i},j}^{-1} \left(\overline{q}_{\boldsymbol{i},j-}(\delta_{t\cdot}\underline{u}^n_{\boldsymbol{i}})(\mu_{x_j-}\delta_{x_j+}\underline{u}^n_{\boldsymbol{i}}) - \overline{q}_{\boldsymbol{i},j+}(\delta_{t\cdot}\underline{u}^n_{\boldsymbol{i}})(\mu_{x_j-}\delta_{x_j+}\underline{u}^n_{\boldsymbol{i}})\right) \tag{5.89}$$

For a bound on the sign-indeterminate $\mathfrak{t}^n$, we note that:

$$\Theta_j = \langle w_j^{-1} q_{j+}\delta_{x_j+}\underline{u}^n, q_{j+}e_{t-}\delta_{x_j+}\underline{u}^n\rangle^\star_{\mathbb{G}_h} \geq -\frac{k^2}{4}\langle w_j^{-1} q_{j+}\delta_{t-}\delta_{x_j+}\underline{u}^n\rangle^{\star 2}_{\mathbb{G}_h} \tag{5.90}$$

by the identity (3.43e), which is valid for any inner product. Then it can be shown that the following bound holds for this domain (we omit the derivation):

$$\langle w_j^{-1} q_{j+}\delta_{x_j+}\underline{v}^n, q_{j+}\delta_{x_j+}\underline{v}^n\rangle^\star_{\mathbb{G}_h} \leq \frac{4}{h^2}\|\underline{v}^n\|^{\star 2}_{\mathbb{G}_h} \tag{5.91}$$

and we arrive at:

$$\mathfrak{h}^n \geq \frac{1}{2c^2}(1-d\lambda^2)\|\delta_{t-}\underline{u}^n\|^{\star 2}_{\mathbb{G}_h} \tag{5.92}$$

which leads to the usual condition on the Courant number, $\lambda \leq \sqrt{1/d}$, for $\mathfrak{h}^n \geq 0$.

As for the boundary term $\mathfrak{b}^n$, we can make use of the identity $\mu_{x_j\pm}\delta_{x_j\mp} = \delta_{x_j\cdot}$ to arrive at:

$$\mathfrak{b}^n = \sum_{\boldsymbol{i}\in\mathcal{I}_\mathrm{b}} w_{\boldsymbol{i}} h^{d-1} \sum_{j=1}^{d} w_{\boldsymbol{i},j}^{-1} \left(\overline{q}_{\boldsymbol{i},j-}(\delta_{t\cdot}\underline{u}^n_{\boldsymbol{i}})(\delta_{x_j\cdot}\underline{u}^n_{\boldsymbol{i}}) - \overline{q}_{\boldsymbol{i},j+}(\delta_{t\cdot}\underline{u}^n_{\boldsymbol{i}})(\delta_{x_j\cdot}\underline{u}^n_{\boldsymbol{i}})\right) \tag{5.93}$$

which leads to boundary conditions that can employ centered differences $\delta_{x_j\cdot}$ in the cases of Neumann and lossy boundary types. Considering just Neumann and frequency-independent lossy conditions,[9] we rewrite $\mathfrak{b} = \mathfrak{q}$ (representing a power outflow) and we require $\mathfrak{q} \geq 0$ for stability as a whole. As such, the following choices are available:

$$\delta_{x_j\cdot}\underline{u}^n_{\boldsymbol{i}} = 0\,, \qquad \boldsymbol{i}\in\mathcal{I}_\mathrm{b}\,, (\text{when } \overline{q}_{\boldsymbol{i},j\pm} = 1) \qquad \text{(Neumann)} \tag{5.94a}$$

$$\mp\delta_{x_j\cdot}\underline{u}^n_{\boldsymbol{i}} = \frac{\gamma_{\boldsymbol{i},j\pm}}{c}\delta_{t\cdot}\underline{u}^n_{\boldsymbol{i}}\,, \qquad \boldsymbol{i}\in\mathcal{I}_\mathrm{b}\,, (\text{when } \overline{q}_{\boldsymbol{i},j\pm} = 1)\,, \gamma_{\boldsymbol{i},j\pm}\geq 0 \qquad \text{(lossy)} \tag{5.94b}$$

Stability of the scheme under the above boundary conditions is also ensured for domains that can be composed as Cartesian products of 1-D intervals (e.g., half-spaces aligned with grid axes, quarter-spaces aligned with grid axes, etc.).

In order to express the final boundary updates, one could make use of the following rule to eliminate

[9] Here, there is no distinction between Dirichlet boundary conditions of the "non-centered" or "centered" type, since $\underline{u}^n_{\boldsymbol{i}} = 0$ when $\boldsymbol{i}\in\mathcal{I}_\mathrm{b}$ are also given by (5.52).

ghost points:

$$\underline{u}^n_{\boldsymbol{i}\pm\hat{\mathbf{e}}_j} = \underline{u}^n_{\boldsymbol{i}\mp\hat{\mathbf{e}}_j} + \frac{\gamma_{\boldsymbol{i},j\pm}}{\lambda}(\underline{u}^{n+1}_{\boldsymbol{i}} - \underline{u}^{n-1}_{\boldsymbol{i}})\,, \quad \text{whenever} \quad \overline{q}_{\boldsymbol{i},j\pm} = 1 \tag{5.95}$$

Under Neumann conditions ($\gamma_{\boldsymbol{i},j\pm} \equiv 0$), the update equation can be written as:

$$\underline{u}^{n+1}_{\boldsymbol{i}} = (2 - 2d\lambda^2)\underline{u}^n_{\boldsymbol{i}} - \underline{u}^{n-1}_{\boldsymbol{i}} + \lambda^2 \dot{Q}^n_{\boldsymbol{i}} \tag{5.96}$$

where the partial summation $\dot{Q}^n_{\boldsymbol{i}}$ is

$$\dot{Q}^n_{\boldsymbol{i}} = \sum_{j=1}^{d} w^{-1}_{\boldsymbol{i},j}\left(q_{\boldsymbol{i},j+}\underline{u}^n_{\boldsymbol{i}+\hat{\mathbf{e}}_j} + q_{\boldsymbol{i},j-}\underline{u}^n_{\boldsymbol{i}-\hat{\mathbf{e}}_j}\right) \tag{5.97}$$

which, for this domain, has at most d non-zero summands.

Under lossy conditions, the update equation would be:

$$\underline{u}^{n+1}_{\boldsymbol{i}} = \frac{1}{1+\lambda\gamma'_{\boldsymbol{i}}}\left((2 - 2d\lambda^2)\underline{u}^n_{\boldsymbol{i}} + (\lambda\gamma'_{\boldsymbol{i}} - 1)\underline{u}^{n-1}_{\boldsymbol{i}} + \lambda^2\dot{Q}^n_{\boldsymbol{i}}\right) \tag{5.98}$$

where $\gamma'_{\boldsymbol{i}} = \sum_{j=1}^{d} \gamma_{\boldsymbol{i},j\pm}$.

In the simplified case of one admittance per grid point ($\gamma_{\boldsymbol{i},j\pm} \equiv \gamma_{\boldsymbol{i}}$), we would have:

$$\underline{u}^{n+1}_{\boldsymbol{i}} = \frac{1}{1+\lambda\overline{K}_{\boldsymbol{i}}\gamma_{\boldsymbol{i}}}\left((2 - 2d\lambda^2)\underline{u}^n_{\boldsymbol{i}} + (\lambda\overline{K}_{\boldsymbol{i}}\gamma_{\boldsymbol{i}} - 1)\underline{u}^{n-1}_{\boldsymbol{i}} + \lambda^2\dot{Q}^n_{\boldsymbol{i}}\right) \tag{5.99}$$

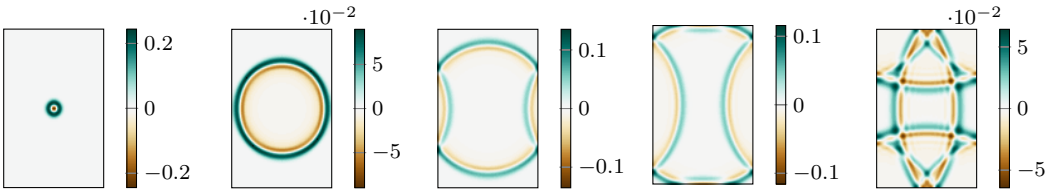
with $\overline{K}_{\boldsymbol{i}}$ representing the number of ghost-point neighbours. Eq. (5.99) is also sufficient to represent Neumann conditions if we let $\gamma_{\boldsymbol{i}} \equiv 0$.

In regards to other boundary conditions in the literature, these conditions are equivalent to the frequency-independent cases proposed in [133, 137] for a 2-D and 3-D box-domain expressed using specific impedances (the inverse of the specific admittance) under the simplification of $\gamma_{\boldsymbol{i},j\pm} = \gamma_{j\pm}$ (one admittance/impedance per wall). We note that spatially-varying impedances over the boundary surface are not permitted by the frequency-domain stability analyses used in [133, 137].

However, in comparison to the frequency-independent cases proposed in [133, 137], an important distinction must be made here. We have restricted our attention to this d-dimensional box-domain, whereas the boundary models in [133, 137] make more liberal use of the substitution rule (5.95) in order to eliminate ghost points over more general domains—this is an approach that is not guaranteed to be stable, as will be discussed in more detail in Section 5.4.

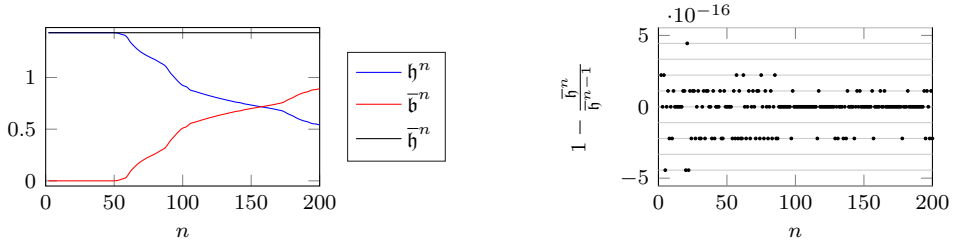
A numerical example

In order to demonstrate energetic conservation under these centered boundary conditions, we consider a numerical example using a rectangular 2-D defined by $(L_x, L_y) = (0.8, 1.2)$, with $c = 1$, $\lambda = \sqrt{1/2}$, $h = 0.01$, and with $\gamma_{\boldsymbol{i}} = 0.1$ at boundaries. The initial conditions are: $u(\mathbf{x}, 0)$ is a spatial Gaussian with variance 0.04, and $u_t(\mathbf{x}, 0) = 0$. The evolution of the numerical solution is shown in Fig. 5.7, and the numerical energy can be seen in Fig. 5.8. We can calculate the energy variations (including accumulated losses), and they appear as multiples of machine epsilon, confirming the conservation of numerical energy under these centered conditions.



(a) $n = 10$ (b) $n = 50$ (c) $n = 70$ (d) $n = 90$ (e) $n = 150$

Figure 5.7: Snapshots of simulations of 2-D wave equation $\lambda = 1/\sqrt{2}$ with f_0 a Gaussian and $g_0 = 0$, after n time-steps as indicated. Centered conditions are employed at boundaries in this example.



(a) Energetic quantities (b) Relative variations in numerical energy

Figure 5.8: Plots demonstrating conservation of numerical energy, associated to numerical example in Fig. 5.7. Here, the quantity $\overline{\mathfrak{h}}^n$ represented the total energy including accumulated losses from the boundary, represented by $\overline{\mathfrak{b}}^n$ ($\overline{\mathfrak{h}}^n$ and $\overline{\mathfrak{b}}^n$ defined in (3.83)). The variations in energy per time-step are integer multiples of double precision machine epsilon, as denoted by grey lines.

Matrix constructions

Let us now consider the matrices pertaining to the use of centered conditions, in terms of matrix stability analyses and matrix forms of the energy analysis.

Under these centered conditions, the Laplacian matrix derived from (5.96) or (5.99) can be composed as the product of two matrices:

$$\mathbf{L} = \mathbf{W}^{-1}\mathbf{Z} \tag{5.100}$$

where $\mathbf{Z}$ is a symmetric matrix and $\mathbf{W}$ is a diagonal matrix, respectively composed as:

$$\mathbf{Z}_{m_1 m_2} = \begin{cases} w_{\boldsymbol{i}_{m_1}} w^{-1}_{\boldsymbol{i}_{m_1},j} & \text{when } \boldsymbol{i}_{m_2} - \boldsymbol{i}_{m_1} = \pm\hat{\mathbf{e}}_j \\ 0 & \text{otherwise} \end{cases}, \quad m_1 \neq m_2 \tag{5.101a}$$

$$\mathbf{Z}_{m_1 m_1} = -\sum_{m_2 \neq m_1} \mathbf{Z}_{m_1 m_2} \tag{5.101b}$$

$$\mathbf{W}_{mm} = w_{\boldsymbol{i}_m} \tag{5.101c}$$

Although not apparent from the above notation, the symmetry in $\mathbf{Z}$ stems from the symmetry property (5.85), which implies that:

$$w_{\boldsymbol{i}_{m_1}} w^{-1}_{\boldsymbol{i}_{m_1},j} = w_{\boldsymbol{i}_{m_2}} w^{-1}_{\boldsymbol{i}_{m_2},j}\,, \quad \boldsymbol{i}_{m_2} - \boldsymbol{i}_{m_1} = \pm\hat{\mathbf{e}}_j \tag{5.102}$$

The diagonal matrices $\mathbf{B}$ and $\mathbf{Q}$ are now given by:

$$\mathbf{B}_{mm} = \sum_{j=1}^{d} (\overline{q}_{\boldsymbol{i}_m,j+}\gamma_{\boldsymbol{i}_m,j+} + \overline{q}_{\boldsymbol{i}_m,j-}\gamma_{\boldsymbol{i}_m,j-})\,, \quad \mathbf{Q}_{mm} = \sum_{j=1}^{d} (\overline{q}_{\boldsymbol{i}_m,j+} + \overline{q}_{\boldsymbol{i}_m,j-}) \tag{5.103}$$

where $\mathbf{B} = \mathbf{GQ}$ in the simplified case that $\gamma_{\boldsymbol{i},j\pm} = \gamma_{\boldsymbol{i}}$, with the diagonal matrix $\mathbf{G}$ having $\gamma_{\boldsymbol{i}}$ along its diagonal.

The Laplacian matrix under centered conditions could also be constructed as follows:

$$\begin{aligned}\mathbf{L}_{m_1 m_2} &= \begin{cases} (q_{\boldsymbol{i}_{m_1},j\pm} + \overline{q}_{\boldsymbol{i}_{m_1},j\mp}) & \text{when } \boldsymbol{i}_{m_2} - \boldsymbol{i}_{m_1} = \pm\hat{\mathbf{e}}_j \\ 0 & \text{otherwise} \end{cases}, \quad m_1 \neq m_2 \\ \mathbf{L}_{m_1 m_1} &= -2d \end{aligned} \tag{5.104}$$

From this construction it may be apparent that $\mathbf{L}$ is not a symmetric matrix. In order to see this more clearly, we give the 1-D form below:

$$\mathbf{L} = \begin{bmatrix} -2 & 2 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 2 & -2 \end{bmatrix} \tag{5.105}$$

As such, there is some question as to whether the necessary condition (5.31a) is satisfied.

A similarity transform

Let us now consider the "matrix method" stability conditions (5.31a), (5.31b) and (5.33) with respect to the above matrix constructions for $\mathbf{L}$ (using centered conditions).

While $\mathbf{W}$ and $\mathbf{Z}$ are symmetric matrices, $\mathbf{L}$ is not symmetric in this case (as exemplified above), and thus its spectrum does not trivially satisfy the condition (5.31a) laid out by the matrix method for stability analysis. On the other hand, we can show that $\mathbf{L}$ has the same spectrum as a symmetric matrix in order to satisfy (5.31a). Essentially what follows is a *similarity transform* (to be explained shortly).

Let μ be an eigenvalue of $\mathbf{L}$ and let the associated eigenvector be $\boldsymbol{\phi}$. We have then:

$$\mathbf{L}\boldsymbol{\phi} = \mathbf{W}^{-1}\mathbf{Z}\boldsymbol{\phi} = \mu\boldsymbol{\phi} \quad \Longrightarrow \quad \mathbf{W}^{1/2}\mathbf{W}^{-1}\mathbf{Z}\boldsymbol{\phi} = \mu\mathbf{W}^{1/2}\boldsymbol{\phi} \tag{5.106}$$

We can define $\mathbf{y} = \mathbf{W}^{1/2}\boldsymbol{\phi}$ and we have:

$$\mathbf{M}\mathbf{y} = \mu\mathbf{y}\,, \quad \mathbf{M} = \mathbf{W}^{-1/2}\mathbf{Z}\mathbf{W}^{-1/2} \tag{5.107}$$

where μ is an eigenvalue of $\mathbf{M}$ with eigenvector $\mathbf{y}$. Thus, $\sigma(\mathbf{M}) = \sigma(\mathbf{L})$, and clearly, $\mathbf{M} = \mathbf{M}^{\mathrm{T}}$, so $\Im\{\sigma(\mathbf{L})\} = 0$.

In terms of the similarity transform, we have:

$$\mathbf{M} = \mathbf{W}^{1/2}\mathbf{L}\mathbf{W}^{-1/2} \tag{5.108}$$

This states that $\mathbf{M}$ and $\mathbf{L}$ are *similar* matrices, and similar matrices share the same spectra [112].

So in terms of the "matrix method" stability conditions, we have the condition (5.31a) satisfied, and (5.31b) is trivially satisfied by GCT, and GCT also provides a bound on $\hat{\rho}(\mathbf{L})$, which leads to the stability condition $\lambda \leq \sqrt{1/d}$.

We can check if this GCT-based stability condition is indeed tight for the Laplacian matrix corresponding to the application of centered boundary conditions to the coarse grid illustrated in Fig. 5.5. For

that case $\hat{\rho}(\mathbf{L})$ is calculated to be (using MATLAB [164])

$$\hat{\rho}(\mathbf{L}) = -8 \tag{5.109}$$

down to numerical precision, so the aforementioned bound is tight for that example. In Section 5.3.2 it will be seen that, in general, $\hat{\rho}(\mathbf{L}) = 4d$ for $\mathbf{L}$ matrices constructed with (5.100).

Energy analysis in matrix form

It is also worth revisiting the modified energy analysis in the context of these matrices.

The stored energetic terms in (5.88) can be written in matrix form as follows (dropping the superscript n):

$$\mathfrak{v} = \frac{h^d}{2c^2}(\delta_{t-}\mathbf{u})^{\mathrm{T}}\mathbf{W}(\delta_{t-}\mathbf{u})\,, \quad \mathfrak{t} = \frac{h^d}{2h^2}\mathbf{u}^{\mathrm{T}}\mathbf{Z}^{\star}e_{t-}\mathbf{u} \tag{5.110}$$

where $\mathbf{W}$ and $\mathbf{W}_j$ are diagonal matrices with entries $w_{\boldsymbol{i}}$ and $w_{\boldsymbol{i},j}$, respectively, and $\mathbf{Z}^{\star}$ is a matrix of the form:

$$\mathbf{Z}^{\star} = \sum_{j=1}^{d} \mathbf{W}\mathbf{W}_j^{-1}(\mathbf{D}_{x_j+})^{\mathrm{T}}\mathbf{D}_{x_j+} \tag{5.111}$$

and thus we have that $\mathbf{Z}^{\star} = -\mathbf{Z}$.[10]

In order to bound the total "energy" for the system, we can write:

$$\mathfrak{t} \geq -\frac{h^d k^2}{8h^2}\mathbf{v}^{\mathrm{T}}\mathbf{Z}^{\star}\mathbf{v} \tag{5.112}$$

where we define $\mathbf{v} = \delta_{t-}\mathbf{u}$ for convenience. Then we have:

$$\mathfrak{h} \geq \frac{1}{2c^2}\mathbf{v}^{\mathrm{T}}\left(\mathbf{W} - \frac{\lambda^2}{4}\mathbf{Z}^{\star}\right)\mathbf{v} \tag{5.113}$$

We require that $\mathfrak{h} \geq 0$, but if we define $\tilde{\mathbf{v}} = \mathbf{W}^{-1/2}\mathbf{v}$, then we only need to satisfy:

$$\tilde{\mathbf{v}}^{\mathrm{T}}\left(\mathbf{I} - \frac{\lambda^2}{4}\mathbf{W}^{-1/2}\mathbf{Z}^{\star}\mathbf{W}^{-1/2}\right)\tilde{\mathbf{v}} \geq 0 \tag{5.114}$$

and we and have that:

$$\tilde{\mathbf{v}}^{\mathrm{T}}\left(\mathbf{I} - \frac{\lambda^2}{4}\mathbf{W}^{-1/2}\mathbf{Z}^{\star}\mathbf{W}^{-1/2}\right)\tilde{\mathbf{v}} \geq \left(1 - \frac{\lambda^2}{4}\hat{\rho}\left(\mathbf{W}^{-1/2}\mathbf{Z}^{\star}\mathbf{W}^{-1/2}\right)\right)\|\tilde{\mathbf{v}}\|^2 \tag{5.115a}$$

since $\mathbf{W}^{-1/2}\mathbf{Z}^{\star}\mathbf{W}^{-1/2}$ is positive semi-definite. Consequently, we know that $\hat{\rho}\left(\mathbf{W}^{-1/2}\mathbf{Z}^{\star}\mathbf{W}^{-1/2}\right) = \hat{\rho}(\mathbf{M}) = \hat{\rho}(\mathbf{W}^{-1}\mathbf{Z})$, and we arrive at the stability condition:

$$\lambda \leq \frac{2}{\sqrt{\hat{\rho}(\mathbf{W}^{-1}\mathbf{Z})}} \quad \Longrightarrow \quad \mathfrak{h} \geq 0 \tag{5.116}$$

Recalling that $\mathbf{L} = \mathbf{W}^{-1}\mathbf{Z}$, we realise that this is equivalent to (5.33). Applying GCT to obtain a bound on $\hat{\rho}(\mathbf{W}^{-1}\mathbf{Z})$, we get $\lambda \leq \sqrt{1/d}$ as a sufficient condition for stability (necessary and sufficient for Neumann conditions).

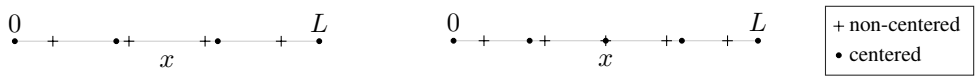
[10] In the context of graph theory, $\mathbf{Z}^{\star}$ can be seen as a *weighted* graph Laplacian (with edges weighted), carrying with it properties (5.82a) and (5.82b).

5.3 Comparing centered and non-centered boundary conditions

So far we have considered two approaches to finite difference boundary modelling: using "non-centered" conditions or centered conditions.[11] However, we have not given a clear motivation for the use of one approach over the other. Indeed, the use of centered spatial difference seems attractive at first, since $\delta_{x_j \cdot} = \partial_{x_j} + \mathcal{O}(h^2)$ whereas $\delta_{x_j \pm} = \partial_{x_j} + \mathcal{O}(h)$, but it is important to remember that the operators $\delta_{x_j \pm}$ are also second-order accurate about a half-shifted position in space. In order to provide some insight into the choice between these two conditions, in this section we compare these two sets of conditions, in terms of numerical modal frequencies (under Neumann boundary conditions) on box-domains, and in terms of numerical reflection coefficients on half-spaces with lossy conditions.

5.3.1 Modal frequencies for the Neumann problem in 1-D

As a simple first example for comparison purposes, let us consider the modal frequencies that arise from the wave equation on a one-dimensional domain $\overline{\Omega} = [0, L]$, approximated using the simple scheme with non-centered and centered Neumann (free) conditions. In this case, it is straightforward to write the grid function $\underline{u}_i^n$ as a $M \times 1$ column vector $\mathbf{u}^n = [u_0^n, u_1^n, \ldots, u_{M-1}^n]^{\mathrm{T}}$. When centered conditions are employed, it is appropriate to choose h such that $M = L/h + 1$. Similarly, when "non-centered" conditions are employed, it is appropriate to choose h such that $M = L/h$, and to stagger those grid points by $h/2$ (from the origin). The resulting grid points are illustrated by Fig. 5.9. Indeed, under this setup, "non-centered" boundary conditions would in fact be centered about the boundaries of interest. However, for lack of a better term, we continue to call them "non-centered conditions".



(a) $M = 4$ (b) $M = 5$

Figure 5.9: $\overline{\Omega} = [0, L]$ and sampling points under various boundary conditions with $M = 4$ and $M = 5$.

In the case of non-centered conditions, we have $\mathbf{L} = \mathbf{D}_{xx}^{(\pm)}$, and under centered Neumann conditions, we have $\mathbf{L} = \mathbf{D}_{xx}^{(\cdot)}$, where $\mathbf{L} = \mathbf{D}_{xx}^{(\pm)}$ is given by (5.70), and where $\mathbf{D}_{xx}^{(\cdot)}$ is defined as:

$$\mathbf{D}_{xx}^{(\cdot)} = \begin{bmatrix} -2 & 2 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 2 & -2 \end{bmatrix} \tag{5.117}$$

It is well-known that the eigenvalues of $\mathbf{D}_{xx}^{(\pm)}$ are [171, 259, 277]:

$$\mu_m = -4 \sin^2 \left(\frac{m\pi}{2M} \right), \quad m = 0, 1, \ldots, M-1 \tag{5.118}$$

[11] More specifically, considering still just a box-domain aligned with grid axes, we have, so far, two approaches to discretising box-shaped domains (and Cartesian products of 1-D intervals), with non-centered or centered discrete (Neumann or frequency-independent lossy) boundary conditions.

and that the eigenvalues of $\mathbf{D}_{xx}^{(\cdot)}$ are:

$$\mu_m = -4\sin^2\left(\frac{m\pi}{2(M-1)}\right), \quad m = 0, 1, \ldots, M-1 \tag{5.119}$$

From these eigenvalues, we can obtain the discrete modal frequencies from (5.30), giving:

$$-4\sin^2(\omega_m k/2) = \lambda^2\mu_m \quad \implies \quad \omega_m k = 2\sin^{-1}\left(\lambda\sqrt{-\mu_m/4}\right), \quad m = 0, \ldots, M-1 \tag{5.120}$$

So for non-centered and centered conditions with $\lambda = 1$, we have the respective modal frequencies:

$$-4\sin^2(\omega_m k/2) = -4\sin^2\left(\frac{m\pi}{2M}\right) \quad \overset{M=L/h}{\Rightarrow} \quad \omega_m = m\frac{c\pi}{L}, \quad m = 0, 1, \ldots, M-1 \tag{5.121a}$$

$$-4\sin^2(\omega_m k/2) = -4\sin^2\left(\frac{m\pi}{2(M-1)}\right) \quad \overset{M=L/h+1}{\Rightarrow} \quad \omega_m = m\frac{c\pi}{L}, \quad m = 0, 1, \ldots, M-1 \tag{5.121b}$$

We find that both types of boundary conditions lead to identical sets of modal frequencies, which also happen to be exact. As such, there is no advantage to one set of boundary conditions over the other for this problem, provided that the grid points are properly spaced in the domain.

5.3.2 Modal frequencies for a box in d dimensions ($d > 1$)

Next, let us consider an extension of this problem to d dimensions. However, we should note that for $d > 1$ it will only be possible to have both centered and "non-centered" conditions centered about the boundary for a box-domain with equal edge-lengths. Nevertheless, we will formulate this problem more generally for unequal edge-lengths so that we can discuss some important properties in regards to stability, after we consider modal frequencies.

We consider a grid function $\underline{u}_{\boldsymbol{i}}^n$ for $\boldsymbol{i} \in \prod_{j=1}^d [0, \ldots, M_j - 1]$, where M_j is a positive integer. For now, we assume that $M_1 = \cdots = M_j = M$, and as such we define a box domain

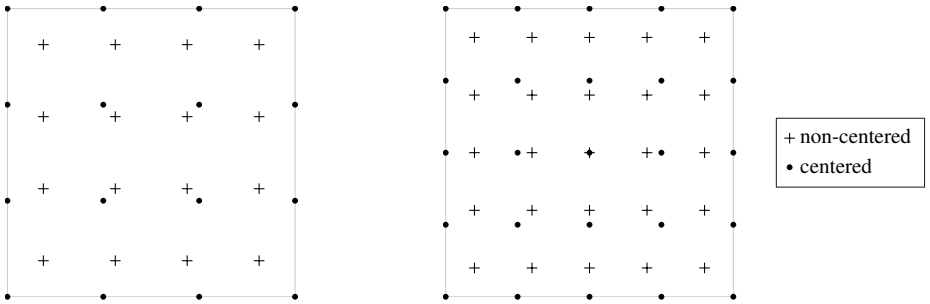
$$\mathcal{V}_{\text{box}} = \prod_{j=1}^{d} (0, L) \tag{5.122}$$

where L is the edge-length. As in the 1-D case, we consider different grid spacings for centered and non-centered conditions. For centered-conditions, we choose h such that $h = L/(M+1)$, and for non-centered conditions we choose $h = L/M$. Furthermore, let the grid for non-centered conditions be staggered by $h/2$ in each coordinate direction, as is appropriate with "non-centered" conditions. These grid points are illustrated in Fig. 5.10.

In this case, the grid function $\underline{u}_{\boldsymbol{i}}^n$ can be decomposed into a vector $\mathbf{u}^n$ of size $M \times 1$ with $M = \prod_{j=1}^d M_j$ using a "column-major" vectorisation, which has the first dimension of $\underline{u}_{\boldsymbol{i}}^n$ stored contiguously in computer memory.[12] In a column-major vectorisation, the mth element of the vector $\mathbf{u}^n$ corresponds to $\underline{u}_{\boldsymbol{i}}^n$ via the mapping:

$$m = \sum_{j=1}^{d} i_j \prod_{l=1}^{j-1} M_j, \quad \boldsymbol{i} = (i_1, \ldots, i_d) \tag{5.123}$$

[12] This is the decomposition provided by the semi-colon operator `(:)` in MATLAB [164].



(a) $M = 4$ (b) $M = 5$

Figure 5.10: $\overline{\Omega} = [0, L]^2$ and sampling points under various boundary conditions with $M = 4$ and $M = 5$.

This vectorisation is illustrated in Fig. 5.11 for a 4×4 grid in 2-D.

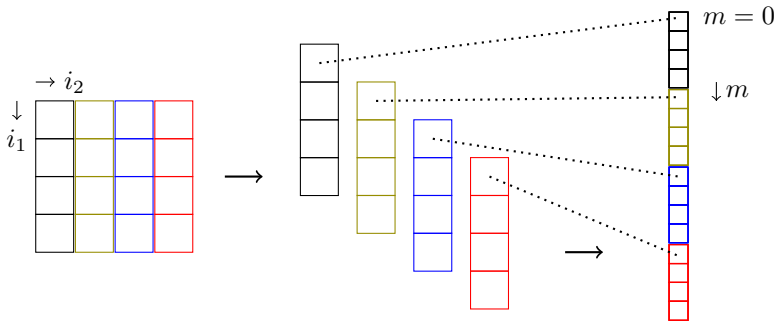


Figure 5.11: Illustration of column-major decomposition of 4×4 grid.

For these box-domains (possibly with unequal edge-lengths), the Laplacian matrices can be represent in compact forms that follow from the update equations derived previously. First, we must introduce the Kronecker products and sum operators [112]. The Kronecker product ($\otimes$) of two arbitrary matrices $\mathbf{A}$ and $\mathbf{B}$, respectively of sizes $M_1 \times N_1$ and $M_2 \times N_2$, is defined as:

$$\mathbf{A}_{\{M_1 \times N_1\}} \otimes \mathbf{B}_{\{M_2 \times N_2\}} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}_{\{M_1 M_2 \times N_1 N_2\}} \tag{5.124}$$

and the Kronecker sum ($\oplus$) is defined for two square matrices as:

$$\mathbf{A}_{\{M_1 \times M_1\}} \oplus \mathbf{B}_{\{M_2 \times M_2\}} = \mathbf{I}_{M_2} \otimes \mathbf{A} + \mathbf{B} \otimes \mathbf{I}_{M_1} \tag{5.125}$$

where $\mathbf{I}_M$ is the identity matrix of size $M \times M$. It is also worth noting the property [112]:

$$\mu_{m_1} \in \sigma(\mathbf{B}_1)\,, \quad \mu_{m_2} \in \sigma(\mathbf{B}_2) \quad \Longrightarrow \quad \mu_{m_1} + \mu_{m_2} \in \sigma(\mathbf{B}_1 \oplus \mathbf{B}_2) \tag{5.126}$$

Let us now relabel the second-order difference matrices $\mathbf{D}_{xx}^{(\cdot)}$ and $\mathbf{D}_{xx}^{(\pm)}$ as $\mathbf{D}_{x_j x_j}^{(\cdot)}, \mathbf{D}_{x_j x_j}^{(\pm)}$, which are now of respective sizes $M_j \times M_j$. The Laplacian matrix corresponding to the application of non-centered conditions can then be written as:

$$\mathbf{L}^{(\pm,d)} = \mathbf{D}_{x_d x_d}^{(\pm)} \oplus \cdots \oplus \mathbf{D}_{x_1 x_1}^{(\pm)} \tag{5.127}$$

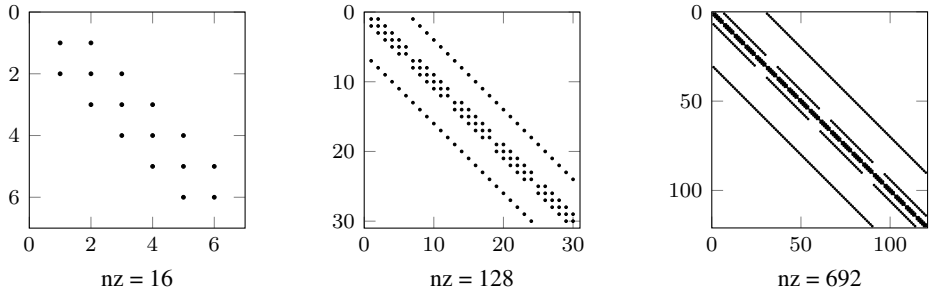
where $\oplus$ represents the Kronecker sums (see [112]).[13] Meanwhile the Laplacian matrix corresponding to centered conditions can be expressed as:

$$\mathbf{L}^{(\cdot,d)} = \mathbf{D}^{(\cdot)}_{x_d x_d} \oplus \cdots \oplus \mathbf{D}^{(\cdot)}_{x_1 x_1} \tag{5.128}$$

The sparsity patterns of such matrices are illustrated in Fig. 5.12. As an aside, it is also worth noting that, for the simplified case of $\gamma_{\boldsymbol{i},j\pm} = \gamma_{\boldsymbol{i}}$, where $\mathbf{B} = \mathbf{G}\mathbf{Q}$, the matrix $\mathbf{Q}$ can be composed as:

$$\mathbf{Q} = \Xi\Big(\operatorname{diag}(\underbrace{[1,0,\ldots,0,1]}_{1\times M_d}) \oplus \cdots \oplus \operatorname{diag}(\underbrace{[1,0,\ldots,0,1]}_{1\times M_1}) \Big) \tag{5.129}$$

where $\Xi = 1/2$ for non-centered conditions, and $\Xi = 1$ for centered conditions, and where $\operatorname{diag}()$ is a function that forms a diagonal matrix from the argument vector.[14]



(a) $\mathbf{L}^{(\cdot,1)}$ (b) $\mathbf{L}^{(\cdot,2)}$ (c) $\mathbf{L}^{(\cdot,3)}$

Figure 5.12: Sparsity patterns of the matrices $\mathbf{L}^{(\cdot,1)}$, $\mathbf{L}^{(\cdot,2)}$, $\mathbf{L}^{(\cdot,3)}$ for $(M_1, M_2, M_3) = (6, 5, 4)$. The sparsity patterns are unchanged for non-centered conditions.

Under these constructions it is straightforward to arrive at an expression for the eigenvalues of these matrices using the property (5.126). For non-centered conditions, we have the eigenvalues:

$$\mu_{\boldsymbol{m}} = -4 \sum_{j=1}^{d} \sin^2\left(\frac{m_j \pi}{2M_j}\right), \quad \boldsymbol{m} \in \prod_{j=1}^{d} [0,\ldots,M_j - 1] \tag{5.130}$$

where $\boldsymbol{m}$ is the vector modal index $\boldsymbol{m} = (m_1,\ldots,m_d)$ with non-negative integer entries.

Meanwhile, the eigenvalues of $\mathbf{L}$ in the case of centered conditions are:

$$\mu_{\boldsymbol{m}} = -4 \sum_{j=1}^{d} \sin^2\left(\frac{m_j \pi}{2(M_j - 1)}\right), \quad \boldsymbol{m} \in \prod_{j=1}^{d} [0,\ldots,M_j - 1] \tag{5.131}$$

The modal frequencies associated to these eigenvalues are given by (5.30). Using the standard choice of $\lambda = \sqrt{1/d}$, and for the box-domain $\mathcal{V}_{\text{box}}$ with appropriately spaced grid points for centered and non-centered conditions, we again find that the two sets of approximate modal frequencies are equivalent. However, as opposed to the case of $d = 1$, these approximate modal frequencies are no longer exact, since there is unavoidable numerical dispersion for $d > 1$, which causes a deviation of numerical modes from their ideal positions.

[13] We should note that these Laplacian matrix constructions are well-known in terms of Kronecker products and regular sums, see, e.g., [277].

[14] Like the command `diag()` in MATLAB [164].

Implications for numerical stability

There are some interesting implications of the above eigenvalue expressions for numerical stability. For this discussion, we do not restrict our attention to equal edge-lengths for $\mathcal{V}_{\text{box}}$.

From the expression (5.131), we have that $\hat{\rho}(\mathbf{L}) = 4d$ for centered conditions, as alluded to earlier, so $\lambda \leq \sqrt{1/d}$ is a tight stability condition for centered conditions over a box domain. Interestingly, for non-centered conditions we have that $\hat{\rho}(\mathbf{L}) < 4d$. As such, for small grid sizes one can set the Courant number slightly above $\lambda \leq \sqrt{1/d}$, but for any appreciable size (such as, for room acoustics) the effect is negligible. It was reported in [44]—from an eigenvalue analysis of the one-step state-space matrix—that the Courant number may be set above the von Neumann limit with non-centered conditions. More specifically, it was reported that for a $7 \times 8 \times 9$ grid with non-centered conditions, the Courant number could be set 2% above the standard condition of $\lambda \leq \sqrt{1/3}$. Indeed, inserting $M_1 = 7, M_2 = 9, M_3 = 9$ into (5.130), we find that $\hat{\rho}(\mathbf{L}^{(\pm,3)}) \approx 11.53 \Rightarrow \lambda_{\max} \approx 1.02\sqrt{1/3}$. It is important to remember, however, that by (5.130), $\hat{\rho}(\mathbf{L}) \to 4d$ as $h \to 0$ for any fixed box-domain size.

Restricting our attention again to the box-domain of equal edge-lengths, we note that—barring the effect of increasing the Courant number slightly beyond $\sqrt{1/d}$ with non-centered conditions, which slightly improves numerical dispersion for $d > 1$—there is no difference between the two choices of conditions for the calculation of modal frequencies for this special-case box domain.

5.3.3 Analysing numerical reflection coefficients

While there is essentially no difference in the use of non-centered versus centered conditions for the Neumann problem (on a box-shaped domain), the story is somewhat different in the case of lossy conditions. The frequency-dependent lossy condition involves the centered operator $\delta_{t\cdot}$ so the pairing with the centered operator $\delta_{x_j\cdot}$ renders the discrete boundary condition *spatiotemporally* centered. In comparison to non-centered lossy conditions, this leads to increased accuracy at the boundary in terms of reflection coefficients, provided that the boundary node lies on the boundary, as we will see in the following analysis.

In Section 2.1.3 we derived a reflection coefficient ($R = \frac{1-\gamma}{1+\gamma}$) at normal incidence for frequency-independent boundary conditions with specific wall admittance γ. In the same way, we can analyse these discrete boundary conditions at normal incidence and derive *numerical* reflection coefficients, similar to the analyses appearing in [133, 137]. Without loss of generality, here we will consider the case of one spatial dimension for a half-space domain: $x \geq 0$ (closed, unbounded) and its boundary at $x = 0$. Analogous to the continuous case ((2.42)), we can assume a trial solution of the form:

$$\underline{u}^n_{i_x} = A\mathrm{e}^{\mathrm{j}(n\omega_k + i_x\beta_h)} + B\mathrm{e}^{\mathrm{j}(n\omega_k - i_x\beta_h)} \tag{5.132}$$

where $\omega_k = \omega k \in \mathbb{R}$ and $\beta_h = \beta h \in \mathbb{R}$. Inserting this trial solution into the centered boundary condition, $\delta_{x\cdot}\underline{u}^n_0 = \frac{\gamma}{c}\delta_{t\cdot}\underline{u}^n_0$, we can arrive at an expression for the numerical reflection coefficient (B/A):

$$R^{(\cdot)}(\beta_h, \gamma) = \frac{\lambda \sin(\beta_h) - \gamma \sin(\omega_k)}{\lambda \sin(\beta_h) + \gamma \sin(\omega_k)} \tag{5.133}$$

Note that this reflectance is purely real, which is a consequence of using centered operators throughout.

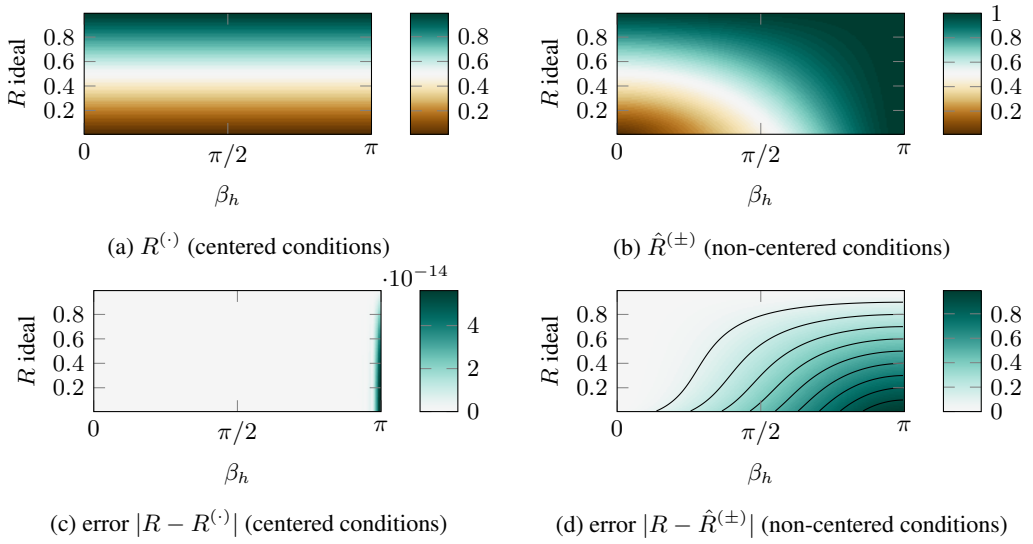
By the same approach, we can find an expression for the numerical reflection coefficient in the

non-centered case. For the boundary condition: $\delta_{x_j+}\underline{u}_0^n = \frac{\gamma}{c}\delta_{t\cdot}\underline{u}_0^n$, we have:

$$R^{(\pm)}(\beta_h,\gamma) = \frac{2\lambda \mathrm{e}^{-\mathrm{j}\beta_h}\sin(\beta_h/2) - \gamma\sin(\omega_k)}{2\lambda \mathrm{e}^{+\mathrm{j}\beta_h}\sin(\beta_h/2) + \gamma\sin(\omega_k)} \tag{5.134}$$

In this case, the reflection coefficient is complex, which is ultimately due to the use of a non-centered difference operator, but it is also indicative of the delay induced by the extra half grid spacing in the implied staircase geometry. For this analysis, we consider only the reflection magnitude, which we denote as $\hat{R}^{(\pm)}(\beta_h,\gamma) = |R^{(\pm)}(\beta_h,\gamma)|$; i.e., we ignore this half-step delay for the proceeding analysis.[15]

Let us now compare these two numerical reflection coefficients for $\lambda = 1$, where it can be assumed that numerical dispersion is not a factor (the interior scheme being exact). We plot the numerical reflection coefficients as a function of the ideal R (solving for γ with $\gamma = \frac{1-R}{1+R}$) and normalised wavenumber β_h in Fig. 5.13. It can be seen that for $\lambda = 1$, the centered conditions are exact down to machine precision. Indeed, they are fully *exact*, which is confirmed by inserting $\omega_k = \beta_h$ and $\lambda = 1$ into (5.133).[16] The non-centered conditions behave as desired in low wavenumbers, as they are consistent (and thus, valid), but they exhibit reflectance errors that generally increase with the wavenumber. Clearly then, in the 1-D case with lossy boundaries, centered conditions are preferred.



(a) $R^{(\cdot)}$ (centered conditions)

(b) $\hat{R}^{(\pm)}$ (non-centered conditions)

(c) error $|R - R^{(\cdot)}|$ (centered conditions)

(d) error $|R - \hat{R}^{(\pm)}|$ (non-centered conditions)

Figure 5.13: Numerical wall reflection coefficients, and absolute errors, for centered and non-centered lossy conditions at normal incidence with $\lambda = 1$. Contour lines denote increments of 0.1.

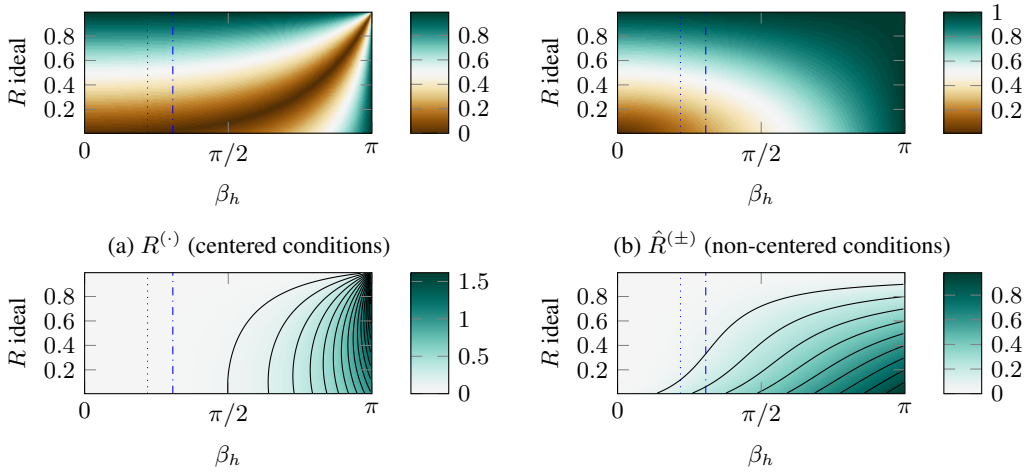
Let us now consider the 2-D case, which, for normal incidence, is encapsulated by $R^{(\cdot)}$ and $R^{(\pm)}$ under the choice of $\lambda = \sqrt{1/2}$. For this two-dimensional case, we have to relate ω_k and β_h using the appropriate numerical dispersion relation. The numerical reflection coefficients pertaining to this 2-D case are shown in Fig. 5.14. In this case both boundary conditions exhibit reflectance errors. Indeed, in higher dimensions, numerical dispersion comes into play, and centered conditions can no longer provide exact reflectance. In comparison to the centered conditions, the non-centered conditions have slightly more error in low wavenumbers, and slightly less in high wavenumbers. It is also worth pointing out, again, that both conditions are exact at DC (consistent). Furthermore, they are exact for $R = 1$ (Neumann conditions), and as such they tend to be more accurate for higher values of R.

[15] Of course, this underlying delay is important for accuracy considerations (i.e., order of accuracy), but in practice a half-grid spacing delay will be imperceptible at any appreciable audio rate.

[16] This is also reflected in [24, Problem 6.6].

To put these errors into perspective with the levels of dispersion errors present in the scheme, the critical wavenumbers $\beta_{h,1\%}$ and $\beta_{h,2\%}$ (defined by (4.33)), which demarcate the regions under which dispersion errors are less than one- or two-percent, are denoted with blue lines in Fig. 5.14. We see that in the regions where dispersion errors are small, reflectance errors are also small. In other words, both boundary conditions are reliable in the region where the scheme itself can be seen as reliable. The analogous 3-D case ($\lambda = \sqrt{1/3}$) is shown in Fig. 5.15, and where it can be seen that reflectance errors are slightly worse in both conditions, but generally they exhibit a similar behaviour to the 2-D cases.

From these analyses, we can conclude that centered conditions will generally perform better for half-space boundaries, but non-centered conditions are relatively accurate in the region of wavenumbers where dispersion errors are also relatively accurate. Then again, it is important to remember that this analysis only applies to half-plane boundaries aligned with grid axes. For more general domains, numerical stability becomes another concern in the use of centered conditions, as will be seen shortly.



(a) $R^{(\cdot)}$ (centered conditions)

(b) $\hat{R}^{(\pm)}$ (non-centered conditions)

(c) error $|R - R^{(\cdot)}|$ (centered conditions)

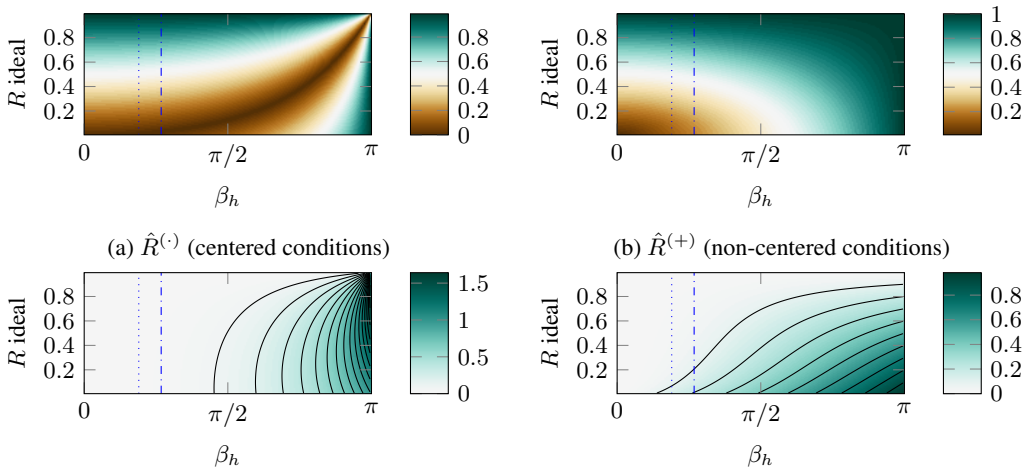
(d) error $|R - \hat{R}^{(\pm)}|$ (non-centered conditions)

Figure 5.14: Numerical wall reflection coefficients, and absolute errors, for centered and non-centered lossy conditions at normal incidence with $\lambda = \sqrt{1/2}$ (2-D). Contour lines denote increments of 0.1. The dashed-dotted line denotes $\beta_{h,2\%}$ and the dotted line denotes $\beta_{h,1\%}$ for this scheme.

Remarks on the use of numerical reflection coefficients for stability analysis

It is worth noting that numerical reflection coefficients can, in special cases, serve to analyse the stability of boundary conditions from which they were derived. For example, the above trial solution can be used for a uniform wall impedance over a planar boundary aligned with grid axes (a semi-infinite half-space). In order to carry out such an analysis, one would let $\omega_k \in \mathbb{C}$ and solve for conditions such that $|R(\omega_k)| \leq 1$. This amount to, essentially, a simplified form of GKSO analysis [109], and has been used to some extent for the simple and interpolated schemes in 2-D and 3-D [133, 137, 135, 136]. In the case of a semi-infinite half-space, such an analysis would lead to the usual conditions on the Courant numbers and the condition $\gamma \geq 0$.

It is important not to rely on such frequency domain approaches beyond their limited scope. Strictly speaking, such stability analyses are only valid on the domains for which the trial solution is valid. In other words, the trial solution must be a solution to both the interior and the boundary under consideration for such stability analyses to remain valid. If there are multiple boundary updates to consider (such as near a corner, or under spatially-varying wall impedances) another suitable (and general) trial



(a) $\hat{R}^{(\cdot)}$ (centered conditions)

(b) $\hat{R}^{(+)}$ (non-centered conditions)

(c) error $|R - \hat{R}^{(\cdot)}|$ (centered conditions)

(d) error $|R - \hat{R}^{(+)}|$ (non-centered conditions)

Figure 5.15: Numerical wall reflection coefficients, and absolute errors, for centered and non-centered lossy conditions at normal incidence with $\lambda = \sqrt{1/3}$ (3-D). Contour lines denote increments of 0.1.

solution must be devised.[17] General domains which do not align neatly with grid axes, and domains that are not convex, are not easily treated by such analyses, since the formulation of suitable trial solutions becomes unwieldy.

In the next section we investigate some boundary conditions that have been proposed in the literature that were previously thought to be numerically stable—via the aforementioned frequency-domain analyses—for general (staircase) domains [133, 137, 135].

5.4 On the stability of some discrete boundary models

In a sense, the implementation of boundary conditions in a finite difference framework can be seen as a problem of eliminating ghost points. Through the use of energy analyses, and under guarantees of stability, we have considered approaches to eliminating ghost points on general staircase domains using non-centered conditions (with (5.53)), and approaches to eliminating ghost points on box-shaped domains (aligned with grid axes) using centered conditions (with (5.95)).

In this section we discuss some of the issues that can arise in extending the use of centered conditions to more general domains—motivated by, e.g., the above comparison of numerical reflection coefficients obtained with centered versus non-centered lossy conditions, or motivated by conventional notions of order of accuracy associated to such conditions (which may not necessarily hold on general domains). One possible approach to extending the use of centered conditions would be to simply apply the substitution rule (5.95) whenever a ghost point is encountered on a finite Cartesian grid. Although not necessarily stable, this approach underpins two frequency-dependent (and frequency-independent) impedance boundary models proposed by Kowalczyk and van Walstijn [133, 137, 135, 136], which have since become widely used in the room acoustics literature [222, 36, 160, 238, 215, 255, 19].

It is worth briefly describing the boundary models proposed in [133, 137, 135, 136]. In these boundary models, each wall is characterised by a complex specific impedance $\xi(z), z \in \mathbb{C}$, expressed as a

[17] The trial solution of two plane waves, with wave directions related by image principles, can possibly be extended to quarter-spaces aligned with grid axes (and so on) under the assumption that the same image principles apply in the finite difference boundary conditions.

parallel mass-spring-damper system [133] (similar to the model shown in Section 2.1.3, also see Appendix A) or as a digital impedance filter (DIF) [137, 135, 136]. The DIF boundary model is formulated with digital filters of the form [137, Eq. (17)]:

$$\xi(z) = \frac{b_0 + B(z)}{a_0 + A(z)} \tag{5.135}$$

where $B(z)$ and $A(z)$ are polynomials in z^{-1} with real coefficients, and $b_0, a_0 \in \mathbb{R}^+$. Using frequency-domain stability analyses (analyses of numerical reflectance coefficients), the authors concluded that the resulting schemes would be stable as a whole, provided that $\xi(z)$ is positive-real and provided that the von Neumann stability condition for the scheme is satisfied (e.g., $\lambda \leq \sqrt{1/2}$ for the simplest 2-D scheme) [137, 135, 136].

However, as we have just remarked, frequency-domain stability analyses have a limited scope, and consequently, the limited stability analyses found in [133, 137, 135, 136], cannot be regarded as general proofs of numerical stability. Then again, instabilities in these boundary models had not (until recently) been reported, despite their wide use. Recently, it has been confirmed that such boundary models are indeed prone to instabilities under simple terminations, in [95] (by the current author) and in [44]. Furthermore, it has been shown that such instabilities are not due to round-off errors [44].[18]

In this section we investigate the stability properties of this approach to boundary modelling, beyond the preliminary observations found in [95, 44], in order to explain why such approaches to boundary modelling fail to be stable in general. In terms of finite difference schemes, we will consider the simple five-point scheme in 2-D for simplicity,[19] but we will also briefly consider generalisations to 2-D compact nine-point schemes [135]. As we are only concerned with the underlying approach to eliminating ghost points, it will suffice to consider the frequency-independent simplification of the DIF boundary model [137, 135] ($B(z) = A(z) = 0$, $\xi \in \mathbb{R}^+$), which also happens to be a special case of the boundary models proposed in [133]. As will be seen in this section, the source of such instabilities can be traced to the underlying Laplacian matrices resulting from the general application of centered conditions via (5.95), rather than from, e.g., the implementation of digital impedance filters at boundaries.[20] Since these underlying Laplacian matrices will be unchanged for a given scheme and problem with complex impedance $\xi \neq 0$ in the DIF model (or the model of [133]), the following analyses apply to the more general frequency-dependent models in [133, 137, 135, 136].

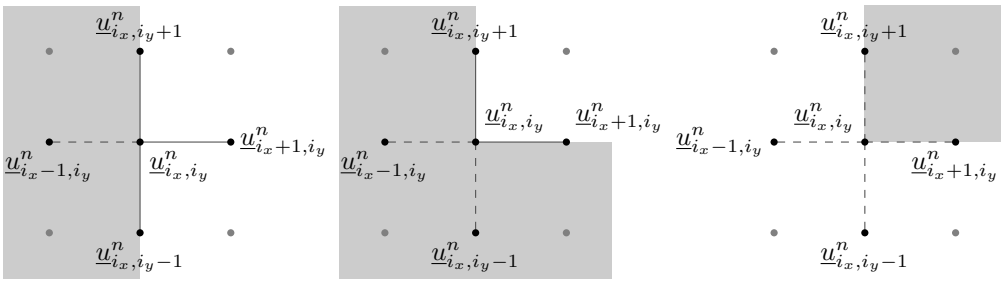
Boundary node types in 2-D for a five-point stencil

In order to illustrate how one might proceed in eliminating ghost points using centered conditions for the five-point scheme in 2-D, let us consider the three types of boundary nodes illustrated in Fig. 5.16. These stencil configurations are commonly known as walls, corners, and "re-entrant corners" (or "inner corners" [137]). The illustrated wall and corner boundaries have, respectively, ghost points at $\underline{u}^n_{i_x-1,i_y}$, and $\underline{u}^n_{i_x-1,i_y}$ and $\underline{u}^{n+1}_{i_x,i_y}$. Before eliminating these ghost points with discrete boundary conditions, the

[18] We should note that Borrel-Jensen [36] first reported instabilities in boundary conditions related to the models presented in [136], although the instabilities reported were with respect to a 3-D boundary stencil configuration devised in [36] (following the approach of [136]), but not one specifically presented in [136]. Indeed, these initial reports of instabilities motivated the investigations that were subsequently presented in [95, 44] for simpler boundary configurations.

[19] Provided that the 2-D scheme exhibits instabilities, it follows that 3-D and higher-dimensional generalisations will also exhibit instabilities since the 2-D scheme can be regarded as a special case of higher-dimensional generalisations.

[20] Of course, improper discretisation of such filters could lead to other instabilities at boundaries, but this will not be investigated here.



(a) wall (b) corner (c) re-entrant corner

Figure 5.16: Illustration of wall, corner, and re-entrant corner boundary nodes in 2-D, with grid points involved in the simple five-point stencil.

finite difference update for $\underline{u}^{n+1}_{i_x,i_y}$ would be as follows:

$$\underline{u}^{n+1}_{i_x,i_y} = -\underline{u}^{n-1}_{i_x,i_y} + 2\underline{u}^{n-1}_{i_x,i_y} + \lambda^2(\underline{u}^n_{i_x+1,i_y} + \underline{u}^n_{i_x-1,i_y} + \underline{u}^n_{i_x,i_y+1} + \underline{u}^n_{i_x,i_y-1} - 4\underline{u}^n_{i_x,i_y}) \tag{5.136}$$

Now let us assume that the left-wall has a wall impedance ξ_x and the bottom wall has a wall impedance ξ_y. First, we note that the special case $\xi_x = \xi_y = 0$ is taken to represent "soft wall" (Dirichlet) conditions, which are implemented by setting ghost points and/or boundary nodes to be zero.[21] As for the case $\xi_x > 0, \xi_y > 0$, we could apply the following centered lossy conditions:

$$\delta_{t\cdot}\underline{u}^n_{i_x,i_y} = c\xi_x\delta_{x\cdot}\underline{u}^n_{i_x,i_y}\,, \qquad \delta_{t\cdot}\underline{u}^n_{i_x,i_y} = c\xi_y\delta_{y\cdot}\underline{u}^n_{i_x,i_y} \tag{5.137}$$

Isolating the ghost points, we have:

$$\underline{u}^n_{i_x+1,i_y} = \underline{u}^n_{i_x-1,i_y} + \frac{1}{\lambda\xi_x}(\underline{u}^{n-1}_{i_x,i_y} - \underline{u}^{n+1}_{i_x,i_y})\,, \qquad \underline{u}^n_{i_x,i_y+1} = \underline{u}^n_{i_x,i_y-1} + \frac{1}{\lambda\xi_y}(\underline{u}^{n-1}_{i_x,i_y} - \underline{u}^{n+1}_{i_x,i_y}) \tag{5.138}$$

and these expressions can be substituted back into (5.136) to arrive at wall and corner boundary updates, whose final forms will be given shortly. As for the re-entrant (inner) corner pictured in Fig. 5.16(c), there are no ghost points to be eliminated, so one might proceed under the assumption that this node-type can be treated like an interior node and updated with (5.136), following [133, 137, 135]. The resulting boundary updates (including the re-entrant corner node) would then be summarised by:

$$\text{wall:} \quad \underline{u}^{n+1}_{i_x,i_y} = \frac{1}{1+\lambda\xi_x^{-1}}\Big((\lambda\xi_x^{-1}-1)\underline{u}^{n-1}_{i_x,i_y} + 2\underline{u}^{n-1}_{i_x,i_y} + \lambda^2(2\underline{u}^n_{i_x+1,i_y} + \underline{u}^n_{i_x,i_y+1} + \underline{u}^n_{i_x,i_y-1} - 4\underline{u}^n_{i_x,i_y})\Big) \tag{5.139a}$$

$$\text{corner:} \quad \underline{u}^{n+1}_{i_x,i_y} = \frac{1}{1+\lambda(\xi_x^{-1}+\xi_y^{-1})}\Big((\lambda(\xi_x^{-1}+\xi_y^{-1})-1)\underline{u}^{n-1}_{i_x,i_y} + 2\underline{u}^{n-1}_{i_x,i_y} + \lambda^2(2\underline{u}^n_{i_x+1,i_y} + 2\underline{u}^n_{i_x,i_y+1} - 4\underline{u}^n_{i_x,i_y})\Big) \tag{5.139b}$$

$$\text{re-entrant corner:} \quad \underline{u}^{n+1}_{i_x,i_y} = -\underline{u}^{n-1}_{i_x,i_y} + 2\underline{u}^{n-1}_{i_x,i_y} + \lambda^2(\underline{u}^n_{i_x+1,i_y} + \underline{u}^n_{i_x-1,i_y} + \underline{u}^n_{i_x,i_y+1} + \underline{u}^n_{i_x,i_y-1} - 4\underline{u}^n_{i_x,i_y}) \tag{5.139c}$$

It is straightforward to work out the other possible orientations for walls and corner nodes, so we leave them out for brevity. The special case $\xi_x = \xi_y = \infty$ models rigid-walls (Neumann conditions), and

[21] This special case is in fact (5.52a), and is thus stable by the previous energy analyses.

it is handled by the above updates by setting $\xi_x^{-1} = \xi_y^{-1} = 0$. What we have described so far is the frequency-independent special case of the discrete boundary models proposed in [133, 137, 135] for the simplest 2-D scheme.

The analogous boundary updates resulting from the application of non-centered conditions ($\delta_{t\cdot}\underline{u}^n_{i_x,i_y} = c\xi_x\delta_{x-}\underline{u}^n_{i_x,i_y}$ and $\delta_{t\cdot}\underline{u}^n_{i_x,i_y} = c\xi_y\delta_{y-}\underline{u}^n_{i_x,i_y}$) would be:

$$\text{wall:} \quad \underline{u}^{n+1}_{i_x,i_y} = \frac{1}{1+\lambda\xi_x^{-1}/2}\Big((\lambda\xi_x^{-1}/2-1)\underline{u}^{n-1}_{i_x,i_y} + 2\underline{u}^{n-1}_{i_x,i_y} + \lambda^2(\underline{u}^n_{i_x+1,i_y} + \underline{u}^n_{i_x,i_y+1} + \underline{u}^n_{i_x,i_y-1} - 3\underline{u}^n_{i_x,i_y})\Big) \tag{5.140a}$$

$$\text{corner:} \quad \underline{u}^{n+1}_{i_x,i_y} = \frac{1}{1+\lambda(\xi_x^{-1}+\xi_y^{-1})/2}\Big((\lambda(\xi_x^{-1}+\xi_y^{-1})/2-1)\underline{u}^{n-1}_{i_x,i_y} + 2\underline{u}^{n-1}_{i_x,i_y} + \lambda^2(\underline{u}^n_{i_x+1,i_y} + \underline{u}^n_{i_x,i_y+1} - 2\underline{u}^n_{i_x,i_y})\Big) \tag{5.140b}$$

$$\text{re-entrant corner:} \quad \underline{u}^{n+1}_{i_x,i_y} = -\underline{u}^{n-1}_{i_x,i_y} + 2\underline{u}^{n-1}_{i_x,i_y} + \lambda^2(\underline{u}^n_{i_x+1,i_y} + \underline{u}^n_{i_x-1,i_y} + \underline{u}^n_{i_x,i_y+1} + \underline{u}^n_{i_x,i_y-1} - 4\underline{u}^n_{i_x,i_y}) \tag{5.140c}$$

These non-centered boundary updates, combined with (5.136), are given by (5.55c) ($\xi = 1/\gamma$), and thus they are energy-stable provided that $\lambda \leq \sqrt{1/2}$ and $\xi_x \geq 0, \xi_y \geq 0$.

Implied staircase geometries

Centered conditions are generally employed under the assumption that boundary nodes lie on the boundary surface of interest, but (clearly) regularly spaced grid points cannot line up with any boundary of interest. Consequently, the use of these centered conditions invariably leads to staircase approximations over general domains (like the use of non-centered conditions).[22]

It is worth briefly illustrating the type of staircase geometry that is implied by this approach to using centered boundary conditions. For an illustrative example, let us define a rotated 2-D box region as:

$$\Omega_{\diamond,\theta} = \{\mathbf{R}_\theta\mathbf{x} \in \mathbb{R}^2 : \mathbf{x} \in (-L_x/2, L_x/2) \times (-L_y/2, L_y/2)\} \tag{5.141}$$

where $\mathbf{R}_\theta$ is a rotation matrix:

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{5.142}$$

and θ is an angle in radians. Let us consider the case where $\Omega = \Omega_{\diamond,\theta}$ with $L_x = L_y = 1$ (a square of unit side-length) and $\theta = \pi/10$ rad.

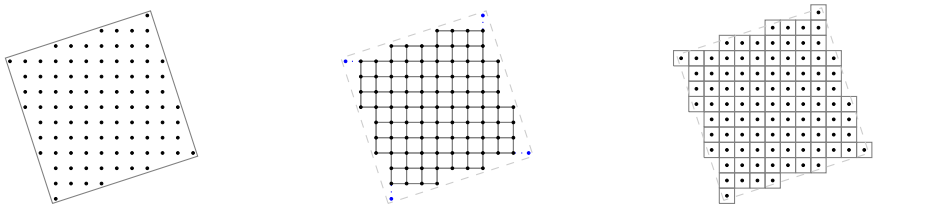
If we were to discretise this rotated square domain with a Cartesian grid $\overline{\Omega} \cap \mathbb{Z}_h^2$ and $h = 0.1$, we would have the set of points illustrated in Fig. 5.17(a).

Using only the three boundary node types illustrated by Fig. 5.16 (and rotations by multiples of $\pi/4$ rad) and the associated centered boundary updates (5.139), an implied staircase geometry could be as illustrated in Fig. 5.17(b). Let us define this implied staircased geometry as the union of square cells that are interconnected by edges and whose vertices are grid points in $\overline{\Omega} \cap \mathbb{Z}_h^2$, and we denote it by $\overline{\Omega}_h^{(\cdot)}$.[23]

It is important to point out that not all of the grid points in $\overline{\Omega} \cap \mathbb{Z}_h^2$ are necessarily employed under the limitations of only three essential boundary stencils, as seen in Fig. 5.17(b). In particular, the grid

[22] Consequently, the order of accuracy of "centered" conditions are reduced to that of non-centered conditions for general domains, and there is essentially no advantage to the use of "centered" conditions over non-centered conditions.

[23] In d dimensions, this would generalise to $\overline{\Omega}'_h$ being the union of hypercubic cells connected by $(d-1)$-dimensional faces, where vertices of the cells are grid points in $\overline{\Omega} \cap \mathbb{Z}_h^d$.



(a) $\overline{\Omega} \cap \mathbb{Z}_h^2$ (b) $\overline{\Omega}_h^{(\cdot)}$, centered conditions (c) $\overline{\Omega}_h$, non-centered conditions

Figure 5.17: (a) Cartesian grid of points, and implied staircase geometries for rotated square under (b): centered and (c): non-centered boundary conditions.

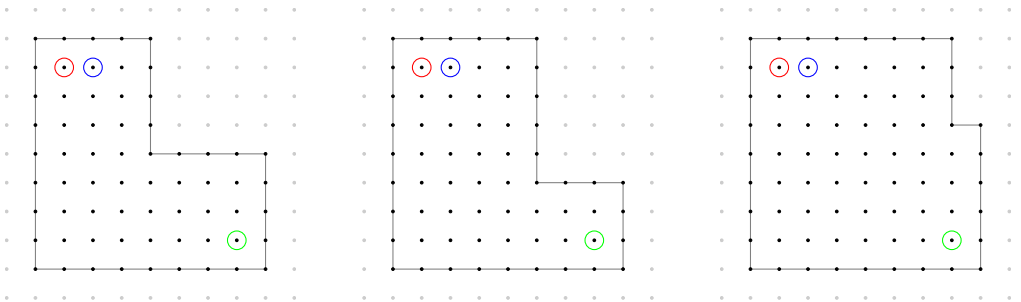
points denoted by blue circles, which are inside the domain, are discarded from $\overline{\Omega}_h^{(\cdot)}$. In fact, it is not entirely necessary to discard such points under the use of centered conditions, but this is the convention that has been taken in [133, 137, 36]. Such stray grid points have typically been viewed as being "one-dimensional" [36], and thereby not consistent with the multidimensional wave equation. In contrast to this, such boundary node configurations are not ambiguous when using non-centered conditions, as they occupy *volume* in the implied staircase geometry $\overline{\Omega}_h$ (see Fig. 5.17(c)), and are thus treated in a straightforward manner.

5.4.1 Testing the stability of centered boundary conditions on L-shaped domains

At this point we will test numerical stability under the use of centered conditions for the simple 2-D scheme (summarised by (5.139)) on a non-trivial domain with rigid walls. We know already that a rectangular domain $\overline{\Omega}_h^{(\cdot)}$ would be stable (by the energy analyses in Section 5.2.2), so we must consider a non-rectangular domain $\overline{\Omega}_h^{(\cdot)}$; the simplest of which is probably an L-shaped domain, since it involves only one re-entrant corner node.

Let us characterise a 2-D discrete L-shaped domain by four integers N_x, N_y, c_x, c_y, where N_x, N_y are the dimensions of the bounding box, and (c_x, c_y) is the location of the re-entrant corner. We also impose the constraint $0 < c_x < N_x$ and $0 < c_y < N_y$. Then we define a discrete L-shaped region as:

$$\mathcal{D}^{(\mathrm{L})}_{N_x,N_y,c_x,c_y} = \{(i_x, i_y) \in \mathbb{Z}^{2+} : i_x \leq N_x, i_y \leq N_y\} \backslash \{(i_x, i_y) \in \mathbb{Z}^{2+} : c_x < i_x \leq N_x, c_y < i_y \leq N_y\} \tag{5.143}$$

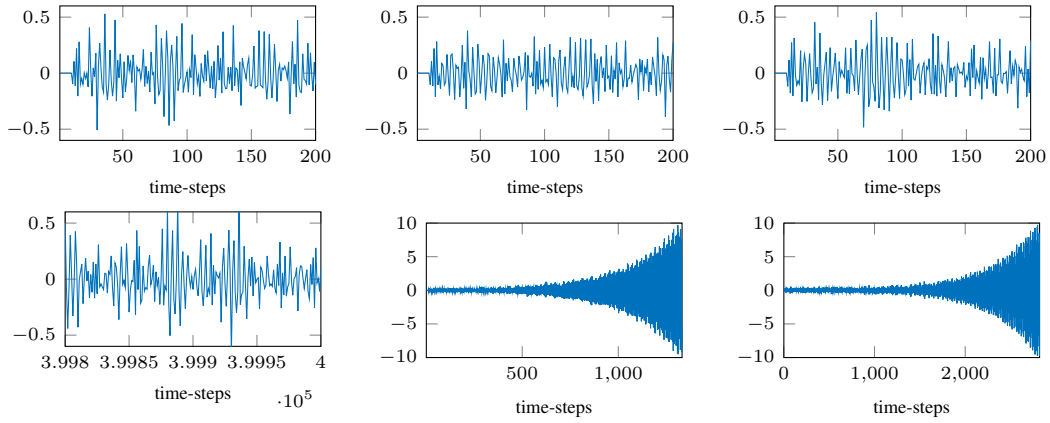


(a) $(N_x, N_y, c_x, c_y) = (8, 8, 4, 4)$ (b) $(N_x, N_y, c_x, c_y) = (8, 8, 5, 3)$ (c) $(N_x, N_y, c_x, c_y) = (8, 8, 7, 4)$

Figure 5.18: Example L-shaped domains $\mathcal{D}^{(\mathrm{L})}_{N_x,N_y,c_x,c_y}$ characterised by (N_x, N_y, c_x, c_y) and (5.143), and surrounding ghost points. The respective origin is at lower-left corners of L-shaped domains.

As test cases, we consider the three L-shaped domains illustrated in Fig. 5.18, with N_x, N_y, c_x, c_y as indicated. For initial conditions, we use $g_0 = 0$ and f_0 is zero everywhere except at the nodes marked

with red and blue circles in Fig. 5.18, which are respectively set to 1 and -1. Using the simplest 2-D scheme with $\lambda = \sqrt{1/2}$ and the centered boundary updates in (5.139) with rigid-walls ($\xi_x^{-1} = \xi_y^{-1} = 0$), simulations are run for up to 40,000 time-steps. Fig. 5.19 shows the outputs of these simulations, at time-steps as indicated, read from the nodes marked by green circles in Fig. 5.18. For the first configuration (illustrated by Fig. 5.18(a)), the simulation is stable for 40,000 time-steps, as seen in Fig. 5.19. The other two configurations give rise to instabilities, i.e., we observe the type of solution growth that should not be present in a numerically stable simulation, indicating that the boundaries proposed in [133, 137, 135] are not as stable as originally claimed.



(a) $(N_x, N_y, c_x, c_y) = (8, 8, 4, 4)$ (b) $(N_x, N_y, c_x, c_y) = (8, 8, 5, 3)$ (c) $(N_x, N_y, c_x, c_y) = (8, 8, 7, 4)$

Figure 5.19: Simulation outputs (at time-steps as indicated) from tests with L-shaped domains (characterised by (N_x, N_y, c_x, c_y) and (5.143)), using centered boundary conditions at corners and walls, and treating the re-entrant corners as an interior node.

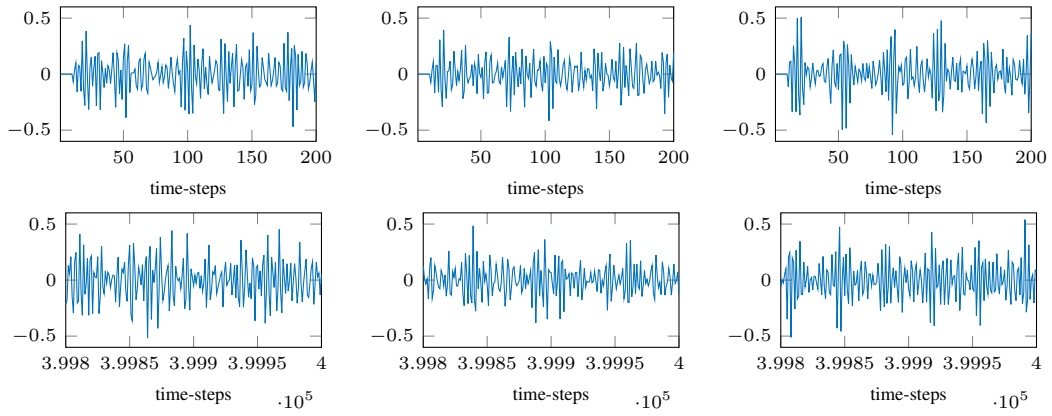
We should note that these instabilities are apparent after $\mathcal{O}(10^2)$–$\mathcal{O}(10^3)$ time-steps, which is relatively late in a simulation of this size. As such, these instabilities are of a different nature than the usual instabilities experienced when the stability condition on the Courant number is violated. Of course, it is possible for finite precision round-off errors to also cause late-time instabilities [44], but if this were the case, we would have seen instabilities in all three configurations. Furthermore, in double-precision arithmetic (as used here) such instabilities should only appear after $\mathcal{O}(10^9)$ time-steps, according to [44]. As such, it is safe to assume that round-off error is not the cause of these instabilities.

Before investigating these instabilities in more detail, it is worth presenting the outputs of analogous simulations using non-centered conditions on the same L-shaped domains, as displayed in Fig. 5.20. Indeed, we see that the obtained outputs for three configurations are free from instabilities (for as many as 40,000 time-steps), as predicted by the energy-based stability analyses.

Matrix analyses – complex eigenvalues

So far we have shown that instabilities can arise in the centered updates (5.139), and this has also been reported in [95] (by the current author) and in [44]. At this point we will explain why these instabilities occur, by identifying which condition is violated (aside from $\hat{\rho}(\mathbf{A}) \leq 1$).

For the rigid-wall special case of the centered boundary conditions summarised by (5.139), the entire update can be written as the matrix recursion (5.20), so the behaviour of the system (barring numerical round-off effects) is completely described by the spectra of the associated matrices $\mathbf{L}$. The associated Laplacian matrices—for these 2-D L-shaped domains and for some vectorisation of $\underline{u}_{\boldsymbol{i}}^n$—can



(a) $(N_x, N_y, c_x, c_y) = (8, 8, 4, 4)$ (b) $(N_x, N_y, c_x, c_y) = (8, 8, 5, 3)$ (c) $(N_x, N_y, c_x, c_y) = (8, 8, 7, 4)$

Figure 5.20: Simulation outputs from tests with L-shaped domains (characterised by (N_x, N_y, c_x, c_y) and (5.143)), using non-centered boundary conditions at corners and walls, and treating re-entrant corners as interior nodes.

be constructed using:

$$\begin{aligned} \mathbf{L}_{m_1 m_2} &= \begin{cases} (q_{\boldsymbol{i}_{m_1}, j\pm} + \overline{q}_{\boldsymbol{i}_{m_1}, j\mp}) & \text{when } \boldsymbol{i}_{m_2} - \boldsymbol{i}_{m_1} = \pm\hat{\mathbf{e}}_j \text{ for } j = 1, 2 \\ 0 & \text{otherwise} \end{cases}, \quad m_1 \neq m_2 \\ \mathbf{L}_{m_1 m_1} &= -\sum_{m_2 \neq m_1} \mathbf{L}_{m_1 m_2} = -4 \end{aligned} \tag{5.144}$$

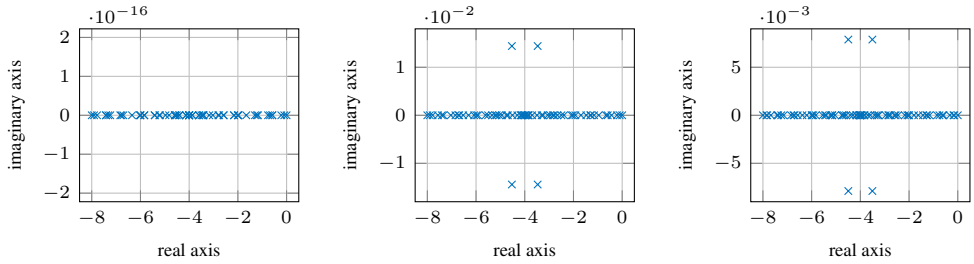
Before analysing the resulting spectra, it is worth noting that under the above construction, the conditions $0 \leq \lambda^2 \sigma(\mathbf{L}) \leq 8$ would be automatically satisfied if it were true that $\Im\{\sigma(\mathbf{L})\} = 0$ and $\lambda \leq \sqrt{1/2}$, simply by the application of Gershgorin's circle theorem. However, the above matrix $\mathbf{L}$ is not symmetric, and a similarity transform that symmetrizes $\mathbf{L}$—such as the one used in Section 5.2.2—is not apparent, so we cannot immediately deduce the nature of $\Im\{\sigma(\mathbf{L})\}$.

As such, let us analyse $\sigma(\mathbf{L})$ numerically, which we can estimate (in double precision arithmetic) using standard eigenvalue solvers, such as MATLAB's `eig` command [164]. In Fig. 5.21 we plot the spectra of the matrices $\mathbf{L}$ associated to the L-shaped domains under test. It can be seen that the condition $\Im\{\sigma(\mathbf{L})\} = 0$ is satisfied (down to machine precision) for the configuration that did not exhibit instabilities. On the other hand, complex eigenvalues are observed for the configurations that exhibited instabilities. Indeed, by (5.30), these complex eigenvalues, which appear in conjugate-pairs, give rise to unstable modes.[24] It is important to point out that such instabilities are *independent* of the Courant number, and as such, they will be present even if $\lambda \leq \sqrt{1/2}$. Later, we will investigate such instabilities in the presence of boundary dissipation, which can sometimes compensate for complex eigenvalues of the underlying Laplacian matrix.

Remarks on the appearance of complex eigenvalues in the spectrum of L

Given these results, one question that naturally arises is whether it is possible to identify staircase geometries $\overline{\Omega}_h^{(\cdot)}$ that only lead to real eigenvalues (such as the first configuration above) in the underlying Laplacian matrix for a given scheme with centered boundaries, following [133, 137] or (5.139).

[24] Each conjugate pair gives rise to four modes, two experiencing exponential growth (over time) and the others exponential decay.



(a) $(N_x, N_y, c_x, c_y) = (8, 8, 4, 4)$ (b) $(N_x, N_y, c_x, c_y) = (8, 8, 5, 3)$ (c) $(N_x, N_y, c_x, c_y) = (8, 8, 7, 4)$

Figure 5.21: Eigenvalues of Laplacian matrices from these test cases on L-shaped domains (characterised by (N_x, N_y, c_x, c_y) and (5.143)), using centered conditions.

To partially answer this question, it is worth noting that the eigenvalues of these finite difference Laplacian matrices are essentially the roots of a polynomial with a degree less than or equal to $|\mathcal{I}|$. It is well-known that there is no general formula for polynomial roots beyond degree four. As such, and considering the grid sizes involved in large-scale room acoustics, identifying such special cases in general would be a challenge (to say the least).

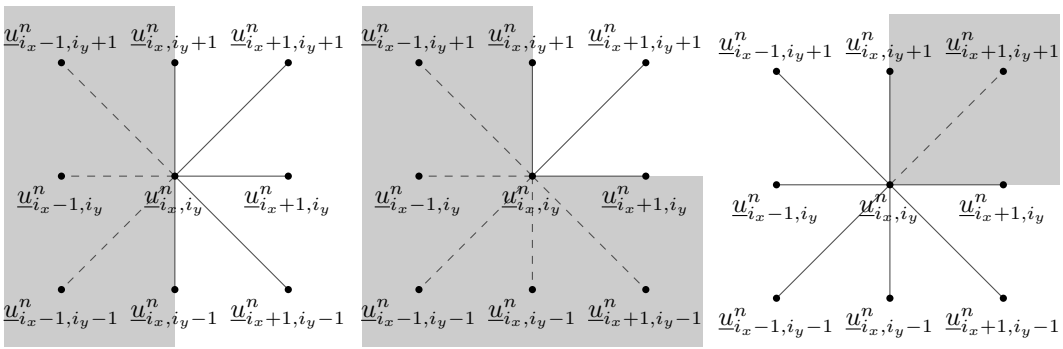
5.4.2 DIF boundaries with boundary losses for an interpolated scheme

While the unstable cases seen earlier for the simple five-point scheme essentially show that the DIF boundary model of [137, 135] is prone to instabilities, and thus cannot be regarded as being numerically stable (apart from the pathological case of Dirichlet boundaries), it is worth also testing the more general DIF "interpolated" boundary conditions that have been proposed for the more general family of compact nine-point schemes [135] in order to see if such instabilities persist in such schemes.

The interior update equation for the compact explicit nine-point schemes can be written as:

$$\begin{aligned}\underline{u}^{n+1}_{i_x,i_y} = -\underline{u}^{n-1}_{i_x,i_y} + (2+\lambda^2 w_0)\underline{u}^{n}_{i_x,i_y} + \lambda^2 w_1(\underline{u}^{n}_{i_x+1,i_y} + \underline{u}^{n}_{i_x-1,i_y} + \underline{u}^{n}_{i_x,i_y+1} + \underline{u}^{n}_{i_x,i_y-1}) \\ + \lambda^2 w_2(\underline{u}^{n}_{i_x+1,i_y+1} + \underline{u}^{n}_{i_x-1,i_y+1} + \underline{u}^{n}_{i_x-1,i_y+1} + \underline{u}^{n}_{i_x-1,i_y-1}) \quad (5.145)\end{aligned}$$

where $w_0 = -2(1+\alpha)$, $w_1 = \alpha$, and $w_2 = (1-\alpha)/2$, and where α is the free parameter.[25] We recall that the von Neumann stability conditions for this family of schemes are: $\lambda \leq 1/\sqrt{2\alpha}$ and $\alpha \geq 0$ [21].



(a) wall (b) corner (c) re-entrant corner

Figure 5.22: Illustration of wall, corner, and re-entrant corner boundary node configurations with the compact nine-point stencil.

[25] Under the notation of [135], we have $\alpha = 1 - 2b$ where b is the parameter that appears in [135].

For the compact nine-point stencil on this L-shaped domain, we have three boundary configurations to consider again: walls, corners, and re-entrant corners, as illustrated in Figure 5.22. Under the simplification of a uniform frequency-independent boundary specific impedance $\xi \geq 0$, the wall, corner, and re-entrant corner DIF boundary updates proposed by Kowalczyk and van Walstijn [135, Eqs. (28),(38),(41)] for the configurations illustrated above are, respectively:

$$\text{wall:} \quad \underline{u}_{i_x,i_y}^{n+1} = \frac{1}{1+\lambda/\xi}\Bigg[(\lambda/\xi - 1)\,\underline{u}_{i_x,i_y}^{n-1} + (2+\lambda^2 w_0)\underline{u}_{i_x,i_y}^{n} + \lambda^2 w_2(2\underline{u}_{i_x+1,i_y+1}^{n} + 2\underline{u}_{i_x-1,i_y-1}^{n}) + \lambda^2 w_1(2\underline{u}_{i_x+1,i_y}^{n} + \underline{u}_{i_x,i_y+1}^{n} + \underline{u}_{i_x,i_y-1}^{n})\Bigg] \tag{5.146a}$$

$$\text{corner:} \quad \underline{u}_{i_x,i_y}^{n+1} = \frac{1}{1+2\lambda/\xi}\Bigg[(2\lambda/\xi - 1)\,\underline{u}_{i_x,i_y}^{n-1} + (2+\lambda^2 w_0)\underline{u}_{i_x,i_y}^{n} + \lambda^2 w_2(4\underline{u}_{i_x+1,i_y+1}^{n}) + \lambda^2 w_1(2\underline{u}_{i_x+1,i_y}^{n} + 2\underline{u}_{i_x,i_y+1}^{n})\Bigg] \tag{5.146b}$$

$$\text{re-entrant corner:} \quad \underline{u}_{i_x,i_y}^{n+1} = -\underline{u}_{i_x,i_y}^{n-1} + (2+\lambda^2 w_0)\underline{u}_{i_x,i_y}^{n} + \lambda^2 w_2(2\underline{u}_{i_x-1,i_y+1}^{n} + 2\underline{u}_{i_x-1,i_y+1}^{n}) + \lambda^2 w_1(\underline{u}_{i_x+1,i_y}^{n} + \underline{u}_{i_x-1,i_y}^{n} + \underline{u}_{i_x,i_y+1}^{n} + \underline{u}_{i_x,i_y-1}^{n}) \tag{5.146c}$$

where w_0, w_1, w_2 are unchanged from (5.145). The remaining wall, corner, and re-entrant corner configurations can be worked out by symmetry, so we leave them out for brevity. For $\alpha = 1 \Rightarrow w_0 = -4, w_1 = 1, w_2 = 0$, these reduce to the centered boundary updates considered earlier ((5.139)) for $\xi_x = \xi_y = \xi$.

At this point we can test the "digital impedance filter" (DIF) boundaries (in the zeroth-order case) for the nine-point scheme with $\alpha = 1/2$ and $\lambda = 1$, i.e., the so-called "interpolated wideband" (IWB) scheme. In this case, we will test *nearly-rigid* wall conditions, with the choice $\xi = 10000 \Rightarrow R \approx 0.9998$; this is an impedance value that has been tested over flat-wall boundaries in [135].

We carry out simulations using the updates (5.145) and (5.146), on the three L-shaped configurations considered earlier. The outputs of these simulations are shown in Fig. 5.23. It can be seen that for the first configuration the numerical solution decays, as would be expected by the presence of boundary losses.[26] On the other hand, the latter two configurations give rise to exponential growth in the numerical solution (instabilities), despite the presence of some boundary losses. The eigenvalues of the associated matrices $\mathbf{L}$ for these examples are displayed in Fig. 5.24, and indeed, we find that complex eigenvalues are present for the two configurations that give rise to instabilities.

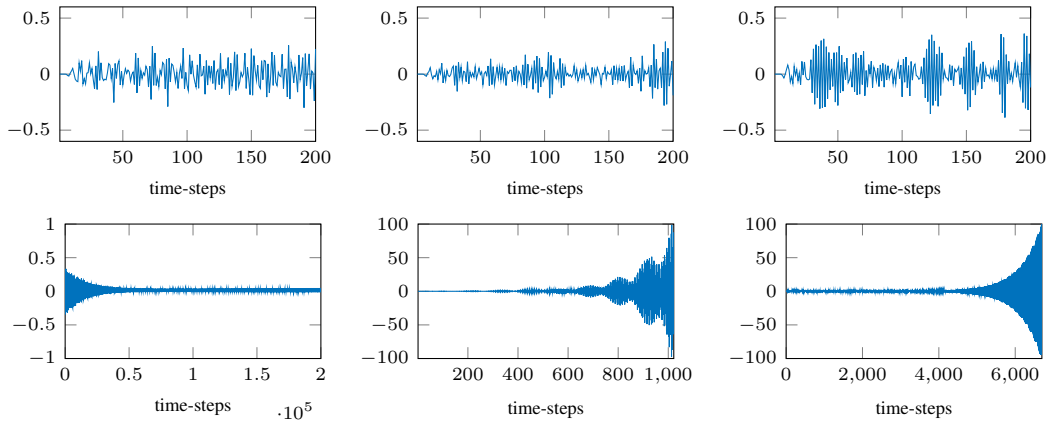
Analysing the spectrum of the state-space system

For another point of view of the stability properties of these examples, we can look at the spectrum of the state-space matrix $\mathbf{A}$ (as in [44]). For this DIF boundary model, the state-space matrix is of the form (5.19) where $\gamma = 1/\xi$, and the matrix $\mathbf{Q}$ can be derived from the update equations (5.146). Solving for $\sigma(\mathbf{A})$, we can then investigate how the eigenvalues $z \in \sigma(\mathbf{A})$ vary in magnitude for various choices of wall impedances ξ.

Fig. 5.25 shows the spectra $\sigma(\mathbf{A})$ for three choices of ξ over the second L-shaped configuration ($(N_x, N_y, c_x, c_y) = (8, 8, 5, 3)$) and using the IWB scheme (and the updates (5.146)).[27] It can be seen

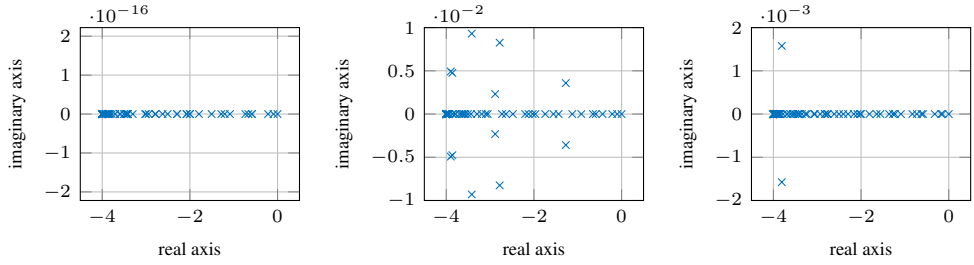
[26] Although, the solution does not completely decay for the Nyquist mode, this is a phenomenon that can also be seen in [44, Fig. 4].

[27] It is worth noting there is a significant clustering of eigenvalues near the Nyquist—this is due to the effects of numerical



(a) $(N_x, N_y, c_x, c_y) = (8, 8, 4, 4)$ (b) $(N_x, N_y, c_x, c_y) = (8, 8, 5, 3)$ (c) $(N_x, N_y, c_x, c_y) = (8, 8, 7, 4)$

Figure 5.23: Simulation outputs (at time-steps as indicated) from tests with L-shaped domains (characterised by (N_x, N_y, c_x, c_y) and (5.143)), using the DIF rigid-wall boundary model; i.e., centered boundary conditions at corners and walls (using (5.146)).



(a) $(N_x, N_y, c_x, c_y) = (8, 8, 4, 4)$ (b) $(N_x, N_y, c_x, c_y) = (8, 8, 5, 3)$ (c) $(N_x, N_y, c_x, c_y) = (8, 8, 7, 4)$

Figure 5.24: Eigenvalues of Laplacian matrices from these test cases on L-shaped domains (characterised by (N_x, N_y, c_x, c_y) and (5.143)), using DIF rigid-wall conditions (i.e., using (5.146)).

that the nearly rigid case ($\xi = 10000$) displays eigenvalues outside the unit circle (z with magnitude greater than one). For $\xi = 100$, most of the eigenvalues move inside the unit circle, yet many unstable modes persist ($|z| > 1$). The unstable modes subside by incorporating enough boundary losses, and in this case the choice of $\xi = 75$ appears to be sufficient.

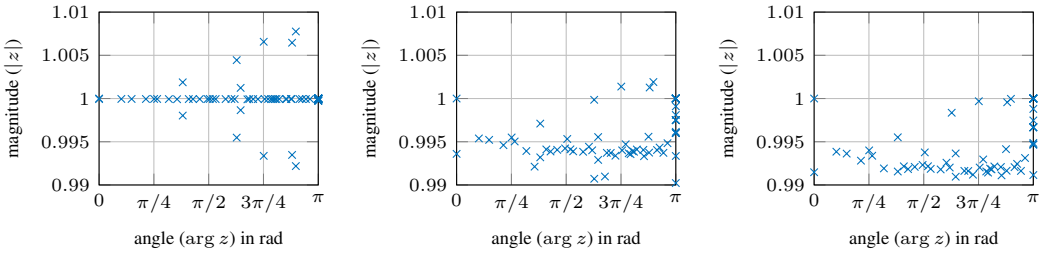
A rotated square domain

The L-shaped geometries we have tested here are some of the simplest examples of unstable configurations, as they contain a single re-entrant corner node. The appearance of complex eigenvalues becomes much more frequent in the general case where one can encounter multiple re-entrant corners.[28] For an example of a slightly more complex geometry, we can consider the rotated square domain illustrated in Fig. 5.17(b), which comprises approximately the same number of grid points as the L-shaped configuration illustrated in Fig. 5.18(c). For this analysis we will consider only the simple five-point scheme.

Fig. 5.26(a) displays the spectrum of the matrix $\mathbf{L}$ which corresponds to DIF centered conditions applied to the simple five-point scheme ($\alpha = 1$ in (5.146)) for this rotated square domain. Several complex eigenvalues appear in this spectrum and they are distributed across the entire spectrum. For another point of view, Figs. 5.26(b) and 5.26(c) show the eigenvalues of the state-space matrix $\mathbf{A}$ under

dispersion in the IWB scheme, which can also be seen in [44, Fig. 12] for the 3-D IWB scheme in the lossless case.

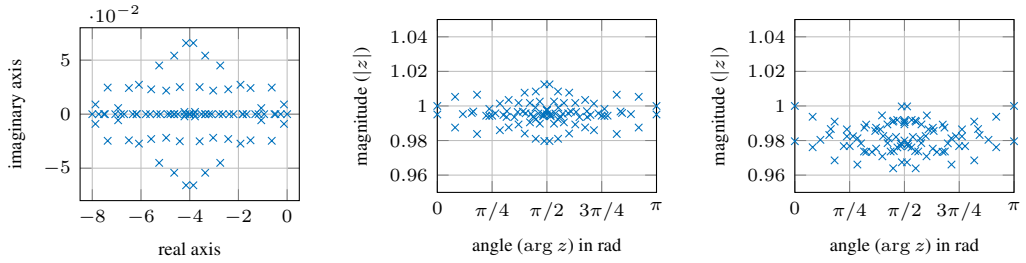
[28] Although we should note that some special cases with multiple re-entrant corners can also end up being stable. Empirically, we have found this to be the case when there is a large degree of symmetry in the domain (with respect to grid axes).



(a) $\xi = 10000$ ($R \approx 0.9998$) (b) $\xi = 100$ ($R \approx 0.9802$) (c) $\xi = 75$ ($R \approx 0.9737$)

Figure 5.25: Eigenvalues of state-space matrices ($z \in \sigma(\mathbf{A})$) for DIF boundary model applied to the L-shaped domain with $(N_x, N_y, c_x, c_y) = (8, 8, 5, 3)$, with wall impedance ξ as indicated. Only one half of the spectrum is plotted due to symmetry of z about the real axis.

two choices of ξ. It can be seen that the choice $\xi = 75$, which was sufficient to compensate for instabilities caused by one re-entrant corner, is no longer sufficient to dampen unstable modes. However, we find that $\xi = 18$ provides enough loss to stabilise the scheme in this case.[29]



(a) $\sigma(\mathbf{L})$ (b) $z \in \sigma(\mathbf{A})$ with $\xi = 75$ ($R \approx 0.97$) (c) $z \in \sigma(\mathbf{A})$ with $\xi = 18$ ($R \approx 0.89$)

Figure 5.26: (a) Eigenvalues of the Laplacian matrix for the DIF boundary model applied to Fig. 5.17(b) using the simplest scheme. (b) Eigenvalues of the associated state-space matrices with ξ as indicated.

Remarks on compensating for unstable modes using boundary losses

Considering that losses in the system can compensate for growth due to unstable modes that arise when using the DIF boundary model, the stability of the resulting finite difference scheme would seem to be conditional on the parameters (in particular, the parameters giving rise to losses) of the continuous model, in a manner that does not mimic the well-posedness of the underlying continuous model (i.e., ξ being non-negative). However, we should point out that even if such stability conditions existed (which must hold in the limiting case of small grid sizes), it would be out of reach to derive such conditions, since one cannot predict which configurations of $\overline{\Omega}_h^{(\cdot)}$ give rise to complex eigenvalues in the underlying Laplacian matrix for a given scheme with DIF boundaries.[30] Similarly, it would be difficult to predict or bound the magnitude of the imaginary parts of complex eigenvalues when they do appear (in order to compensate for them with internal or boundary losses). We must also note that it becomes prohibitively expensive to compute $\sigma(\mathbf{A})$ for large grid sizes, so one cannot rely on direct analyses of $\sigma(\mathbf{A})$ in

[29] Analogous analyses for the 3-D simple scheme can be obtained using the MATLAB [164] scripts provided with [44].

[30] To add to the complexity of the behaviour of these instabilities, we should mention that, although not shown here, there are configurations where rigid-wall DIF boundaries are stable for the five-point scheme and unstable for the IWB scheme (e.g., $(N_x, N_y, c_x, c_y) = (8, 8, 6, 4)$), and cases where the converse is true (e.g., $(N_x, N_y, c_x, c_y) = (8, 8, 5, 4)$). Furthermore, between nine-point variants there seem to be certain L-shaped configurations where one choice of $\alpha \notin \{0, 1\}$ leads to complex eigenvalues, while another $\alpha \notin \{0, 1\}$ does not.

practice. It is thus imperative to formulate boundary models such that the condition $\Im\{\sigma(\mathbf{L})\} = 0$ is guaranteed by construction.

5.5 Summary

In this chapter we looked at boundary conditions for the simplest Cartesian scheme in some detail, focussing mostly on stability properties. We (re)derived non-centered conditions using energy analyses and subsequently we derived centered conditions for box-shaped domains using a modified energy analysis. Matrix representations of these schemes were provided, and we were able to establish links between stability conditions obtained from energy techniques and those obtained from "matrix methods" for stability. Additionally, we compared non-centered conditions to centered conditions, in terms of their accuracies in approximating modal frequencies of box domains, and in terms of approximating numerical reflection coefficients for a half-space boundary. Lastly, we investigated the stability of certain approaches to boundary modelling using centered conditions. It was shown that some widely used boundary models from the room acoustics literature—previously thought to be stable—are in fact prone to instabilities, and we traced those instabilities to complex eigenvalues that can appear unpredictably in the underlying finite difference Laplacian matrices. It was shown that such instabilities were also present with analogous boundary conditions that have been proposed for compact explicit nine-point schemes in 2-D.

Remarks moving forward

Moving on to the next chapter, there are two topics that we could address. The first topic would be the the derivation of stable boundary conditions (beyond simple shoebox domains) for some of the many parametrised schemes seen in Chapter 4. However, it turns out that this would be hasty, since with the set of tools we currently have at our disposable (i.e., energy techniques and matrix methods), we would, at best, end up with boundary conditions of the "non-centered" type (or of the centered type on box domains and/or half-spaces), leading to another form of staircase approximation.

This brings us to the second topic, which is to address some of the errors, yet to be presented, that come with staircase approximations. Such staircasing errors, as will be seen in the next chapter, do not disappear in the limit of small grid spacings, indicating a pervasive inconsistency in staircase-based approaches to room acoustics boundary modelling. As such, the next chapter focusses on a different, yet related, framework that leads to a generalisation of some finite difference schemes, allowing for better adaptation to irregular boundary geometries while still permitting the use of energy techniques for stability analyses.

Chapter 6

Finite volume methods for room acoustics modelling

In this chapter we consider the use of finite volume methods towards the goal of modelling room geometries that do not necessarily adapt to regular Cartesian grids, while ensuring numerical stability via energy methods. Also in this chapter, we investigate some relationships between finite volume and finite difference methods, digital waveguide mesh topologies, and boundary models that have previously appeared in the literature. Additionally, we consider the problem of staircasing (at boundaries) and its effect on the approximation of room modal frequencies and decay times, ultimately revealing important shortcomings of regular grid–based methods (such as finite difference), and demonstrating the advantages of a more general framework that allows for the use of unstructured grids (such as finite volume). Finally, we consider viscothermal losses and general impedance boundary conditions under this finite volume framework, and we present various examples modelling sound propagation over non-trivial domains and with realistic frequency-dependent wall materials, while demonstrating conservation of numerical energy to machine precision.

6.1 Preliminaries

For this chapter we will continue to work with a d-dimensional wave equation, but the variable of interest will now be $\Psi(\mathbf{x}, t)$ with $\mathbf{x} \in \mathbb{R}^d$, representing the acoustic velocity potential over the full space or some closed (and usually bounded) domain of interest: $\overline{\Omega} = (\Omega \cup \Gamma) \subset \mathbb{R}^d$. As such, the PDE of interest is (2.7), but where Δ is the d-dimensional Laplacian. Boundary conditions will be of the various types presented in Chapter 2.

We will also relate these second-order problems to a first-order system (2.1), but in this chapter, we let p and $\mathbf{v}$ represent scalar and vector velocity fields in d spatial dimensions, respectively ($\mathbf{v} = (v_1, \ldots, v_d)$). $\nabla\cdot$ is the d-dimensional divergence operator: $\nabla \cdot \mathbf{v} = \sum_{j=1}^{d} \partial_{x_j} v_j$ and ∇ is the d-dimensional gradient: $\nabla p = (\partial_{x_1} p, \ldots, \partial_{x_d} p)$. The variable p can be seen as a pressure (in Newtons per unit surface area in d-D), $\mathbf{v}$ is the particle velocity (in m/s), and ρ is a density (in kg per unit volume in d-D). Unless otherwise specified, we set $c = 340\,\text{m/s}$ and $\rho = 1.2\,\text{kg/m}^\text{d}$.

6.1.1 Notation for tiling and cells

To present finite volume methods we need to specify a notation for a *tiling* (or a *tessellation*) of the domain of interest by *cells* (or *tiles*). We should remark that the following is a somewhat involved notation for finite volume tiling and cells, but it is chosen for flexibility and generality.

First, we must make a distinction between the domain of interest ($\overline{\Omega}$) and the domain being tiled, since we want to also consider the use of approximate tilings (e.g., staircase approximations). As such, the domain that is being tiled by cells (the "tiled domain") will be represented by $\Omega_\triangle \subseteq \mathbb{R}^d$. When this domain is closed its boundary will be denoted by $\Gamma_\triangle$, and $\overline{\Omega}_\triangle = \Omega_\triangle \cup \Gamma_\triangle$.[1] In the closed case, it is ultimately desirable to have some tiling such that $\Omega_\triangle = \Omega$ and $\Gamma_\triangle = \Gamma$ where $\overline{\Omega} = \Omega \cup \Gamma$ is the closed domain of interest.

For the notation to follow, it will be helpful to refer to the illustration in Fig. 6.1. It should also be noted that this description of a tiling is for d spatial dimensions, and in this general setting it is common to use the terms "volume", "surface area", and "face" as they would be used in 3-D, but generalised to d-dimensions; hence "volume" refers to the d-dimensional volume, which is a length in 1-D, an area in 2-D, etc.. Also, it will be helpful to note that a "polygon" (2-D) and "polyhedron" (3-D) are also known more generally as "polytopes" in d-D [274].

We denote a closed and bounded d-dimensional cell as $\overline{\mathcal{C}}_i$, with the associated index i. This cell has an interior $\mathcal{C}_i$ and boundary $\partial\mathcal{C}_i$, such that $\overline{\mathcal{C}}_i = \mathcal{C}_i \cup \partial\mathcal{C}_i$. A tiling of $\Omega_\triangle$ is a set of cells $\overline{\mathcal{C}}_i$ such that:

$$\bigcup_{i\in I} \overline{\mathcal{C}}_i = \Omega_\triangle \,, \qquad \mathcal{C}_i \cup \mathcal{C}_j = \{\} \,, i \neq j \tag{6.1}$$

where I is the index set of the cells in the tiling. The above essentially states that the domain is tiled by cells without overlaps (the interiors are pair-wise disjoint) and without gaps. Any cell that touches the boundary of the domain (i.e., $\overline{\mathcal{C}}_i \cap \Gamma_\triangle \neq \{\}$) will be called a *boundary cell*. Let us also define an index set for such boundary cells $I^{(\mathrm{b})}$. Any cell that is not a boundary cell will be called an *interior cell*, with the associated index set $I^{(\mathrm{i})}$.

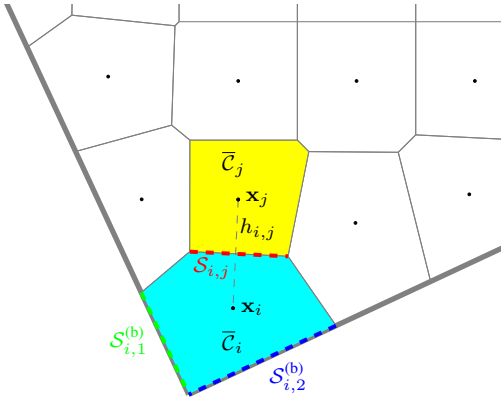


Figure 6.1: A portion of an unstructured tiling. Cell $\overline{\mathcal{C}}_i$ has three neighbouring cells and two boundary faces, $\mathcal{S}^{(\mathrm{b})}_{i,1}$ and $\mathcal{S}^{(\mathrm{b})}_{i,2}$. The neighbouring cell $\overline{\mathcal{C}}_j$ is an interior cell and it has five neighbours, three of which are boundary cells. The inter-cell distance between $\overline{\mathcal{C}}_i$ and $\overline{\mathcal{C}}_j$, which is denoted by $h_{i,j}$, is defined by the associated grid points $\mathbf{x}_i$ and $\mathbf{x}_j$.

[1] We should note that we use the symbol "$\triangle$" in the subscript of $\Omega_\triangle$ because a common example of a cell would be a triangle (or, more generally, a simplex). However, the following notation will not be limited to simplicial cells, nor to cells that are necessarily convex, polygonal, or polyhedral.

Cell faces, neighbouring cells, and boundary faces

The intersection between two closed cells: $\mathcal{S}_{i,j} = \overline{\mathcal{C}}_i \cap \overline{\mathcal{C}}_j = \mathcal{S}_{j,i}$ is called a *face*.[2] Any cell $\overline{\mathcal{C}}_j$ for which $\mathcal{S}_{i,j} \neq \{\}$ is said to be an *adjacent* cell, or a *neighbouring* cell to $\overline{\mathcal{C}}_i$. For convenience, the notation $i \sim j$ will be used to denote that cells $\overline{\mathcal{C}}_i$ and $\overline{\mathcal{C}}_j$ are adjacent. For each cell $\overline{\mathcal{C}}_i$ we define a *neighbour set* N_i which has the indices of neighbouring cells. A *boundary face* is defined as $\mathcal{S}_i^{(\mathrm{b})} = \overline{\mathcal{C}}_i \cap \Gamma_\triangle$. We can further subdivide $\mathcal{S}_i^{(\mathrm{b})}$ into its own tiling, and we express this as: $\mathcal{S}_i^{(\mathrm{b})} = \bigcup_{l=1}^{N_i^{(\mathrm{b})}} \mathcal{S}_{i,l}^{(\mathrm{b})}$, where $\mathcal{S}_{i,l}^{(\mathrm{b})}$ are the $N_i^{(\mathrm{b})} \geq 0$ boundary faces (pairwise disjoint) of a given cell $\overline{\mathcal{C}}_i$. A tiling of convex polytopes can be called "face-to-face" when the faces of the tilings are also faces of convex polytopes.

Associated grid of points

We will also define a grid of points, denoted $\mathbb{G}_{h,\triangle}$, to associate to this tiling, such that each grid point $\mathbf{x}_i \in \mathbb{G}_{h,\triangle}$ is inside at most one cell, i.e., $\mathbf{x}_i \in \overline{\mathcal{C}}_i \setminus \bigcup_{i \in I} \mathcal{S}_{i,j}$. For the most part, the grid point of a cell will be strictly inside the cell, i.e.,: $\mathbf{x}_i \in \mathcal{C}_i$, but it may be desirable to relax that constraint for boundary cells, to allow the grid point of cell to be on a boundary face, i.e., $\mathbf{x}_i \in \mathcal{S}_i^{(\mathrm{b})}$.[3] For the most part, we will consider tilings for which interior cells are *Voronoi cells*, which means that $\mathcal{C}_i$ is the set of points in $\mathbb{R}^d$ that are closer to $\mathbf{x}_i$ than to any other grid point $\mathbf{x}_j \in \mathbb{G}_{h,\triangle}$. Voronoi cells on the interior are convex polytopes in a face-to-face tiling. Also, for each pair of grid points $(\mathbf{x}_i, \mathbf{x}_j)$, we define the inter-cell distance: $h_{i,j} = \|\mathbf{x}_j - \mathbf{x}_i\|$, which can be interpreted as a grid spacing for $i \sim j$.

Cell volumes and face surface areas

In order to eventually arrive at a discrete approximation, we need to specify volumes for cells and surface areas for faces. To each cell $\overline{\mathcal{C}}_i$ we associate a d-dimensional volume $V_i = |\overline{\mathcal{C}}_i|$, and to each face $\mathcal{S}_{i,j}$ we associate a d-dimensional surface area $S_{i,j} = |\mathcal{S}_{i,j}|$.[4] Similarly, we have boundary surface areas $S_i^{(\mathrm{b})} = |\mathcal{S}_i^{(\mathrm{b})}|$ and $S_{i,l}^{(\mathrm{b})} = |\mathcal{S}_{i,l}^{(\mathrm{b})}|$. For 1-D cells, face areas take on binary values.

Locally irregular tilings

For the use of a completely unstructured tiling, in practice one generally needs to store information about each cell ($S_{i,j}$, $h_{i,j}$, and V_i) and its neighbour set N_i.[5] However, for room acoustics simulations, the minimal memory requirements dictated by sampling considerations can already be quite substantial for state variables alone. In order to keep additional memory costs at a minimum, it will be useful to employ tilings that are "locally irregular" (after [107]),[6] meaning that the tiling is a regular[7] configuration of

[2] We should note that a face of a tiling is not necessarily the same as the face of a cell in isolation. For example, one can tile space with rectangles in a distinctive tiling of bricks (rows slightly shifted), but then each rectangular cell has *six* neighbouring cells, and thus six faces in the tiling (rather than, intuitively, four). These two interpretations of "face" coincide in a "face-to-face" tiling [88].

[3] We should note that this is where our definition of a tiling differs slightly from that in [26], which defines inter-cell distances based on cell centroids. The tiling defined here (slightly more general) eventually allows for finite volume interpretations of "centered" conditions in finite difference boundaries, as will be seen in Section 6.5.1.

[4] $S_{i,j}$ is a $(d-1)$-dimensional volume.

[5] Although, by skew-symmetry in adjacent cells, only 50% of that information actually needs to be stored.

[6] This approach is called "local grid deformation" in [39].

[7] We use the term "regular" loosely here—this does not necessarily imply a "regular tiling" or a "regular polytope" cell, as one would find in the geometry literature (e.g., [59]). In 2-D the "regular polytopes" are the "regular polygons", and in 3-D they are the Platonic solids [59]. More generally, a regular polytope in d dimensions has $(d-1)$-dimensional facets (vertices, edges, faces, etc.) that are themselves regular and congruent. A "regular" or "uniform" tiling is a tiling (of the full space) by such regular polytopes. Some tilings we consider to be "regular arrangements of cells" include the Voronoi tilings of the FCC and BCC lattices, which are composed of polyhedral cells that are not regular in the strict sense. The Voronoi cell of the BCC lattice, the truncated octahedron, is both "semi-regular" (regular facets of more than one type) and an "Archimedean solid" (highly symmetric). The Voronoi cell of the FCC lattice, the rhombic dodecahedron, is an "Catalan solid", i.e., *dual* to an Archimedean polyhedron (the

cells (such as a face-to-face tiling of squares or cubes) on the interior, and departs from this regular configuration near the boundary surface. In many such cases, the finite volume update equations for the majority of interior cells simplify to well-known finite difference updates, as will be seen shortly.

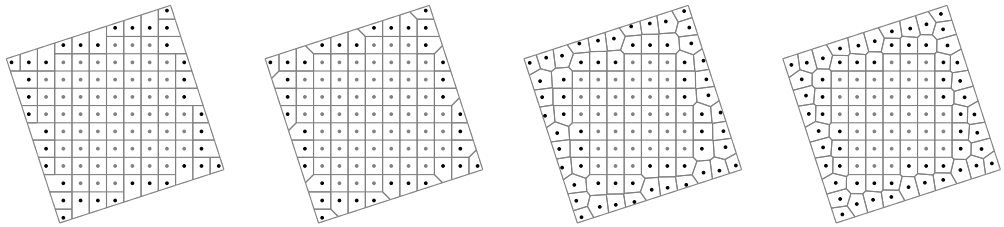


Figure 6.2: Rotated square domain (unit side length, rotated $\pi/10$ rad) with four possible locally irregular tilings. Regular interior cells are denoted by grey grid points and have $h_{i,j} = 0.1$ for $i \sim j$. Irregular cells are denoted by black grid points. Compare with the staircase approximations of the same domain ($\overline{\Omega}_h$) found in Fig. 5.17.

There are many approaches to deriving locally irregular tilings for a given domain and given grid spacing on interior cells.[8] Fig. 6.2 shows four possible examples of locally irregular tilings for a rotated square domain with congruent square cells making up the majority of the interior cells. In these cases, $\overline{\Omega}_\triangle = \overline{\Omega}$ where $\overline{\Omega}$ is the domain of interest. It is worth comparing these with the implied staircase approximations $\overline{\Omega}_h$ illustrated in Fig. 5.17(c). Staircase approximations like those can be seen as tilings of congruent square cells, i.e., $\overline{\Omega}_\triangle = \overline{\Omega}_h$, and clearly in this example $\overline{\Omega}_h \neq \overline{\Omega}$. More will be said about staircase approximations in Section 6.6.

Finally, we note that since locally irregular tilings will play a major role in this chapter, henceforth the symbol h will denote the inter-cell distance of regular interior cells on a locally irregular tiling, and consequently, $\lambda = ck/h$ can be interpreted as the Courant number on the interior of the grid. We should also note that near the boundaries, $h_{i,j}$ will generally vary for fitted tilings, so one cannot define a single Courant number for boundary cells.

6.1.2 Integral forms

Finite volume methods are formulated to approximate *integral forms* of the PDE, so we need to introduce integral forms for the wave equation and the equation describing conservation of mass (2.1a).[9]

Integral form for the second-order wave equation

Considering a cell $\overline{\mathcal{C}}_i \subset \overline{\Omega}$, we can write an integral form for the second-order wave equation as:

$$\frac{1}{c^2} \int_{\overline{\mathcal{C}}_i} \partial_t^2 \Psi \, \mathrm{d}\mathbf{x} = \int_{\overline{\mathcal{C}}_i} \Delta \Psi \, \mathrm{d}\mathbf{x} \tag{6.2}$$

and then using the identity $\Delta\Psi = \nabla \cdot \nabla\Psi$, and subsequently applying the divergence theorem over $\mathcal{C}_i$, we have:

$$\frac{1}{c^2} \int_{\overline{\mathcal{C}}_i} \partial_t^2 \Psi \, \mathrm{d}\mathbf{x} = \int_{\partial\mathcal{C}_i} \hat{\mathbf{n}}_i \cdot \nabla\Psi \, \mathrm{d}\sigma \tag{6.3}$$

cuboctahedron) [59].

[8] This topic, which is generally known as "meshing", cannot be reasonably addressed in this work. See e.g., [270].

[9] It is perhaps worth noting that the continuous integral equations to follow are not necessarily evaluated in practice, but lead to practically realisable fully-discrete schemes.

where $\hat{\mathbf{n}}_i$ is the outward normal vector on the boundary of $\mathcal{C}_i$. By the fact that $\partial\mathcal{C}_i = (\bigcup_{j\in I} \mathcal{S}_{i,j}) \cup \mathcal{S}_i^{(\mathrm{b})}$, we arrive at the following integral form, ready for discretisation by finite volume methods:

$$\frac{1}{c^2}\int_{\overline{\mathcal{C}}_i} \partial_t^2 \Psi \,\mathrm{d}\mathbf{x} = \sum_{j\in N_i} \int_{\mathcal{S}_{i,j}} \hat{\mathbf{n}}_{i,j}\cdot\nabla\Psi \,\mathrm{d}\sigma - \int_{\mathcal{S}_i^{(\mathrm{b})}} v_\perp \,\mathrm{d}\sigma \tag{6.4}$$

where $\hat{\mathbf{n}}_{i,j}$ is the outward normal vector on the face $\mathcal{S}_{i,j}$, and $v_\perp = -\hat{\mathbf{n}}\cdot\nabla\Psi$ for outward normal vector $\hat{\mathbf{n}}$ on $\Gamma_\triangle$.

Integral form for the conservation of mass

Similarly, we can write an integral form for the equation of conservation of mass (2.1a). Taking the integral over a cell $\overline{\mathcal{C}}_i$, we have:

$$\frac{1}{\rho c^2}\int_{\overline{\mathcal{C}}_i} \partial_t p \,\mathrm{d}\mathbf{x} = -\int_{\overline{\mathcal{C}}_i} \nabla\cdot\mathbf{v} \,\mathrm{d}\mathbf{x} \tag{6.5}$$

and after making use of the divergence theorem, we have:

$$\frac{1}{\rho c^2}\int_{\overline{\mathcal{C}}_i} \partial_t p \,\mathrm{d}\mathbf{x} = -\sum_{j\in N_i} \int_{\mathcal{S}_{i,j}} v_{i,j} \,\mathrm{d}\sigma - \int_{\mathcal{S}_i^{(\mathrm{b})}} v_\perp \,\mathrm{d}\sigma \tag{6.6}$$

where we also have $v_\perp = \hat{\mathbf{n}}\cdot\mathbf{v}$. Note that if we substitute in $p = \rho\Psi_t$ and $v_{i,j} = -\hat{\mathbf{n}}_{i,j}\cdot\nabla\Psi$, we recover (6.4). Also, we should note that an integral form for the equation describing conservation of momentum is not strictly necessary for a finite volume discretisation of the system (although it can be employed, as in, e.g., [37]).

6.1.3 Discrete field quantities

Numerical approximations in finite volume methods are generally expressed as spatial averages over cells and cell faces. For example, we will use the following cell-average approximation to $\Psi(\mathbf{x}, t)$:

$$\underline{\Psi}_i^n \approxeq \frac{1}{V_i}\int_{\overline{\mathcal{C}}_i} \Psi(\mathbf{x}, nk) \,\mathrm{d}\mathbf{x} \tag{6.7}$$

where $i \in I$. Time is discretised in the usual manner, with $n \geq 0$ and k the time-step. Considering the grid of points associated to the tiling, the set of $\underline{\Psi}_i^n$ for $i \in I$ can also be seen as a grid function (a set of *pointwise* approximations) in the sense that $\underline{\Psi}_i^n \approxeq \Psi(\mathbf{x}_i, nk)$. It should be noted that this grid function is not necessarily defined over a regular grid of points.

Similarly, for the first-order system we can use the following approximation to the pressure field:

$$\underline{p}_i^{n+\frac{1}{2}} \approxeq \frac{1}{V_i}\int_{\overline{\mathcal{C}}_i} p(\mathbf{x}, (n+1/2)k) \,\mathrm{d}\mathbf{x} \tag{6.8}$$

For the particle velocity field $\mathbf{v}$, we will approximate the vector components $v_{i,j} = \hat{\mathbf{n}}_{i,j}\cdot\mathbf{v}$ where $\hat{\mathbf{n}}_{i,j}$ is the outward normal vector for the cell face $\mathcal{S}_{i,j}$. These are approximated with averages over cell faces:

$$\underline{v}_{i,j}^n \approxeq \frac{1}{S_{i,j}}\int_{\mathcal{S}_{i,j}} v_{i,j}(\mathbf{x}, nk) \,\mathrm{d}\sigma \tag{6.9}$$

and we have the important skew-symmetry property:

$$\underline{v}^n_{i,j} = -\underline{v}^n_{j,i} \tag{6.10}$$

When a cell has a non-zero boundary face, we will approximate $v_\perp$ with:

$$\underline{v}^n_{\perp,i} \approxeq \frac{1}{S^{(\mathrm{b})}_i} \int_{\mathcal{S}^{(\mathrm{b})}_i} v_\perp(\mathbf{x}, nk)\,\mathrm{d}\sigma \tag{6.11}$$

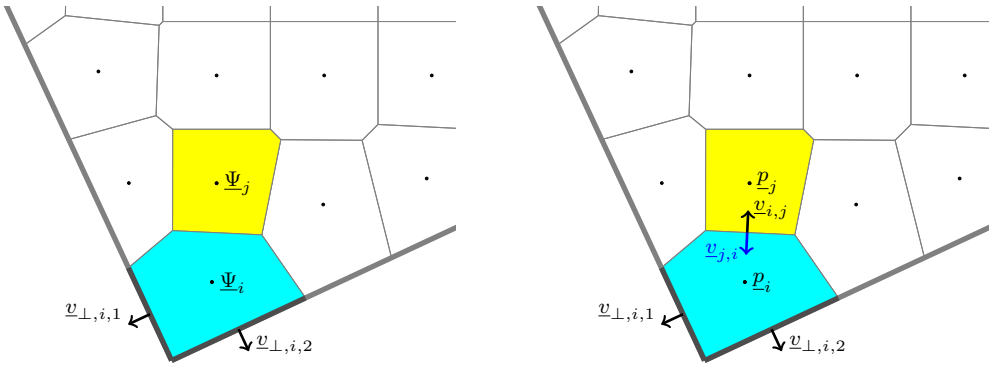
and if so desired, we can approximate the boundary velocity over each subdivided boundary face $\mathcal{S}^{(\mathrm{b})}_{i,l}$ via:

$$\underline{v}^n_{\perp,i,l} \approxeq \frac{1}{S^{(\mathrm{b})}_{i,l}} \int_{\mathcal{S}^{(\mathrm{b})}_{i,l}} v_\perp(\mathbf{x}, nk)\,\mathrm{d}\sigma \tag{6.12}$$

and since $\mathcal{S}^{(\mathrm{b})}_i = \bigcup_{l=1}^{N^{(\mathrm{b})}_i} \mathcal{S}^{(\mathrm{b})}_{i,l}$, we can relate these quantities with:

$$S^{(\mathrm{b})}_i \underline{v}^n_{\perp,i} = \sum_{l=1}^{N^{(\mathrm{b})}_i} S^{(\mathrm{b})}_{i,l} \underline{v}^n_{\perp,i,l} \tag{6.13}$$

These discrete field quantities are illustrated in Fig. 6.3.



(a) (b)

Figure 6.3: Illustration of discrete field quantities used in finite-volume approximations for (a): the second-order wave equation, and (b): the first-order conservation equations.

6.2 Finite volume approximation for linear acoustics equations

We can now write out fully-discrete finite volume approximations to the second-order wave equation and to the first-order system of conservation equations, after [26] but also closely related to [37].

6.2.1 Finite volume approximation to the second-order wave equation

We begin from (6.4) for a finite volume approximation to the second-order wave equation. We introduce the following difference operator, approximating a directional derivative:

$$\delta_{i,j}\underline{\Psi}^n_i = \frac{1}{h_{i,j}}\left(\underline{\Psi}^n_j - \underline{\Psi}^n_i\right) \approx \hat{\mathbf{n}}_{i,j} \cdot \nabla\Psi(\mathbf{x}_i + \hat{\mathbf{n}}_{i,j}h_{i,j}/2, nk) \tag{6.14}$$

It is important to point out that this operator has the skew-symmetry property:

$$\delta_{i,j}\underline{\Psi}_i^n = -\delta_{j,i}\underline{\Psi}_j^n \tag{6.15}$$

Substituting into (6.4) the discrete field quantities $\underline{\Psi}_i^n$ and $\underline{v}_i^n$ for continuous analogues, and using the approximation $\delta_{tt}\underline{\Psi}_i^n \approx \partial_t^2\Psi(\mathbf{x}_i, nk)$, we arrive at the fully discrete scheme:

$$\frac{V_i}{c^2}\delta_{tt}\underline{\Psi}_i^n = \sum_{j\in N_i} S_{i,j}\delta_{i,j}\underline{\Psi}_i^n - \sum_{l=1}^{N_i^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})}\underline{v}_{\perp,i,l}^n \tag{6.16}$$

where the boundary velocity quantities $\underline{v}_{\perp,i,l}^n$ are to be specified by discrete boundary conditions. The simplest choice of boundary conditions to incorporate into (6.16) are rigid walls, $\underline{v}_{\perp,i,l}^n = 0$, which can also seen as (homogeneous) Neumann boundary conditions in $\underline{\Psi}_i^n$.

For interior cells, and also for boundary cells under rigid wall conditions, the finite volume scheme can be operated as the following two-step recursion:

$$\underline{\Psi}_i^{n+1} = 2\underline{\Psi}_i^n - \underline{\Psi}_i^{n-1} + \frac{c^2k^2}{V_i}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}\left(\underline{\Psi}_j^n - \underline{\Psi}_i^n\right) \tag{6.17}$$

with the initial two steps approximated from initial conditions. This second-order-in-time update equation requires the usual two states of memory for $\underline{\Psi}_i^n$ and $\underline{\Psi}_i^{n-1}$ ($\underline{\Psi}_i^{n+1}$ can overwrite $\underline{\Psi}_i^{n-1}$ in place), and some additional storage may be required for $\frac{S_{i,j}}{V_i h_{i,j}}$ and N_i, depending on the type of tiling used. For locally irregular tilings, the amount of additional storage required for these constants scales with the surface area of the domain boundary rather than the volume of the domain.

6.2.2 Finite volume approximation to conservation equations

A standard finite volume approximation to the conservation equations (2.1) can be written as follows. First, we define the approximations:

$$\delta_{t-}\underline{p}_i^{n+\frac{1}{2}} \approx \rho\partial_t p\,, \quad \delta_{t+}\underline{v}_{i,j}^n \approx \partial_t v_{i,j}\,, \quad \delta_{i,j}\underline{p}_i^{n+\frac{1}{2}} \approx \hat{\mathbf{n}}_{i,j}\cdot\nabla p \tag{6.18}$$

Then we have the fully-discrete first-order system:

$$\frac{V_i}{\rho c^2}\delta_{t-}\underline{p}_i^{n+\frac{1}{2}} = -\sum_{j\in N_i} S_{i,j}\underline{v}_{i,j}^n - S_i^{(\mathrm{b})}\underline{v}_{\perp,i}^n \tag{6.19a}$$

$$\rho\delta_{t+}\underline{v}_{i,j}^n = -\delta_{i,j}\underline{p}_i^{n+\frac{1}{2}} \tag{6.19b}$$

Note that we recover the second-order form (6.16) from (6.19a) using the relations:

$$\underline{p}_i^{n+\frac{1}{2}} = \rho\delta_{t+}\underline{\Psi}_i^n\,, \qquad \underline{v}_{i,j}^n = -\delta_{i,j}\underline{\Psi}_i^n \tag{6.20}$$

Generally, the second-order form will be preferred in practice, as it requires fewer variables to be stored per cell. We should also note that (6.19a) is also derived in [37], under the assumption of the use of Voronoi cells in a quasi-Cartesian grid,[10] and (6.19b) is also derived in [37], through an integral form of (2.1b).

[10]Loosely speaking, a quasi-Cartesian grid would be a structured grid that resembles a regular Cartesian grid, but has irregular grid spacings between points; see [37]. We should note, however, that the notation used in [37] is not restricted to such grids.

6.2.3 Simplification to the basic Cartesian scheme

It is worth reviewing how this finite volume scheme relates to the simplest Cartesian finite difference schemes. This is a rather well-known relationship, shown, e.g., in [37] for the first-order system and in [26] for the second-order system. The 3-D case was treated briefly in Section 3.1.2, but we revisit it here for d spatial dimensions. Before we discuss these relationships, we must note that at this point of intersection between finite difference and finite volume we are faced with the problem of having used the symbols i and j in the notation for both frameworks. At the risk of confusion in this short section (but for simplicity elsewhere), for this discussion $\boldsymbol{i}' \in \mathbb{Z}^d$ will be the Cartesian grid index and j' will be the index for Cartesian spatial coordinates in d dimensions, whereas i and j remain cell-indices.

With that said, consider a face-to-face tiling of hypercubes (squares in 2-D, cubic in 3-D) of the full space or some subset of $\overline{\Omega}_\triangle$, with associated grid points at centres of cells. Then for the congruent interior cells (such as the interior cells with grey grid points in Fig. 6.2) we have:

$$V_i = h^d\,, \quad S_{i,j} \overset{i\sim j}{=} h^{d-1}\,, \quad h_{i,j} \overset{i\sim j}{=} h\,, \quad |N_i| = 2d \tag{6.21}$$

and for the $2d$ neighbouring cells, with associated grid points $\mathbf{x}_j$ and outward normal vectors $\hat{\mathbf{n}}_{i,j}$, we have:

$$\hat{\mathbf{n}}_{i,j} \in \{\pm\hat{\mathbf{e}}_{j'}, j' = 1, \ldots, d\} \quad (\text{when } i \sim j) \tag{6.22}$$

where $\hat{\mathbf{e}}_{j'}$ are the standard unit vectors in $\mathbb{R}^d$. Furthermore, let $\mathbf{x}_i \in \mathbb{Z}_h^d$, such that we can write $\mathbf{x}_i = \boldsymbol{i}'h$.

Under the above assumptions and notation, the second-order finite volume update (6.16) can be written as:

$$\underline{\Psi}_{\boldsymbol{i}'}^{n+1} = (2 - 2d\lambda^2)\underline{\Psi}_{\boldsymbol{i}'}^{n} - \underline{\Psi}_{\boldsymbol{i}'}^{n-1} + \lambda^2 \sum_{j'=1}^{d} (\underline{\Psi}_{\boldsymbol{i}'+\hat{\mathbf{e}}_{j'}}^{n} + \underline{\Psi}_{\boldsymbol{i}'-\hat{\mathbf{e}}_{j'}}^{n}) \tag{6.23}$$

where $\lambda = ck/h$. As such, we recover the simplest Cartesian finite difference scheme for the wave equation (the above can be compared to (5.10)). Clearly, under such a simplification, the storage of $V_i, h_{i,j}, S_{i,j}$ and N_i is not necessary for congruent interior cells. The relationship between neighbouring cubic cells in 3-D and the underlying finite difference stencil is illustrated by Fig. 6.4.

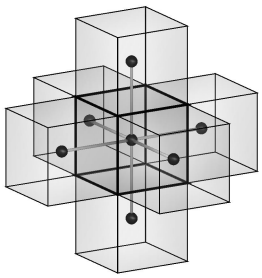


Figure 6.4: Seven-point Cartesian stencil in 3-D and surrounding cubic cells.

Relation to Cartesian staggered finite difference scheme

This finite volume scheme also has a relationship to the staggered finite difference scheme [38] (as already considered for the 3-D case in Section 3.1.2). Let us first generalise the basic Cartesian staggered finite difference scheme to d spatial dimensions, using $\boldsymbol{i}' \in \mathbb{Z}^d$ as the Cartesian grid index and $j' \leq d$ as the Cartesian coordinate index such that $\hat{\mathbf{e}}_{j'}$ are the standard unit vectors in $\mathbb{R}^d$.

We define the d-dimensional Cartesian staggered grid functions:

$$\underline{p}_{\boldsymbol{i}'}^{n+\frac{1}{2}} \approxeq p(\boldsymbol{i}'h, (n+1/2)k)\,, \quad \underline{v}_{j',\boldsymbol{i}'+\hat{\mathbf{e}}_{j'}/2}^{n} \approxeq \hat{\mathbf{e}}_{j'} \cdot \mathbf{v}((\boldsymbol{i}'+\hat{\mathbf{e}}_{j'}/2)h, nk) \tag{6.24}$$

A staggered Cartesian finite difference scheme for (2.1) in d dimensions can then be written as:

$$\frac{1}{\rho c^2}\delta_{t-}\underline{p}_{\boldsymbol{i}'}^{n+\frac{1}{2}} = -\frac{1}{h}\sum_{j'=1}^{d}(\underline{v}_{j',\boldsymbol{i}'+\hat{\mathbf{e}}_{j'}/2}^{n} - \underline{v}_{j',\boldsymbol{i}'-\hat{\mathbf{e}}_{j'}/2}^{n}) \tag{6.25a}$$

$$\rho\delta_{t+}\underline{v}_{j',\boldsymbol{i}'+\hat{\mathbf{e}}_{j'}/2}^{n} = -\frac{1}{h}(\underline{p}_{\boldsymbol{i}'+\hat{\mathbf{e}}_{j'}}^{n+\frac{1}{2}} - \underline{p}_{\boldsymbol{i}'}^{n+\frac{1}{2}})\,, \quad j' = 1, \ldots, d \tag{6.25b}$$

This is equivalent to the first-order finite volume system (6.19) (for congruent interior cells) if we let:

$$\underline{p}_{i}^{n+\frac{1}{2}} = \underline{p}_{\boldsymbol{i}'}^{n+\frac{1}{2}}\,, \qquad v_{i,j} = \pm\underline{v}_{j',\boldsymbol{i}'\pm\hat{\mathbf{e}}_{j'}/2}^{n} \text{ for } \hat{\mathbf{n}}_{i,j} = \pm\hat{\mathbf{e}}_{j'} \tag{6.26}$$

In other words, the finite volume scheme reduces to the staggered finite difference scheme if we consider the quantities $\underline{p}_{i}^{n+\frac{1}{2}}$ and $\underline{v}_{i,j}^{n}$—respectively, average quantities over cell volumes and cell-faces that are naturally staggered in space—to be the staggered point-wise approximations:

$$\underline{p}_{i}^{n+\frac{1}{2}} \approxeq p(\mathbf{x}_i, (n+1/2)k)\,, \quad \underline{v}_{i,j}^{n} \approxeq \hat{\mathbf{n}}_{i,j} \cdot \mathbf{v}(\mathbf{x}_i + \hat{\mathbf{n}}_{i,j}h/2, nk) \tag{6.27}$$

6.2.4 Energy stability analysis and frequency-independent boundary conditions

At this point we can carry out an energy analysis in order to derive stability conditions and boundary conditions. We should note that the energy analysis presented here closely follows that found in [26], only that here we write it in terms of $\underline{\Psi}$ (instead of $\underline{p}$ and $\underline{v}$), and with slight differences in notation. For this analysis, we will use the following identities and inequalities [24]:

$$\delta_{t+}(f^n e_{t-} f^n) = 2(f^n)(\delta_{t\cdot} f^n) \tag{6.28a}$$

$$\delta_{t+}\,(\delta_{t-} f^n)^2 = 2(\delta_{t\cdot} f^n)(\delta_{tt} f^n) \tag{6.28b}$$

$$f^n e_{t-} f^n = (\mu_{t-} f^n)^2 - \frac{k^2}{4}(\delta_{t-} f^n)^2 \tag{6.28c}$$

$$\delta_{t+}(f^n)^2 = 2(\mu_{t+} f^n)(\delta_{t+} f^n) \tag{6.28d}$$

$$\delta_{t+}(\mu_{t-} f^n)^2 = 2(\delta_{t\cdot} f^n)(\mu_{tt} f^n) \tag{6.28e}$$

$$\mu_{tt} = \frac{k^2}{4}\delta_{tt} + 1 \tag{6.28f}$$

where f^n is an arbitrary time-series. Also we will make use of the following inequality (for $a, b \in \mathbb{R}$):

$$(a-b)^2 \leq 2(a^2+b^2) \tag{6.29}$$

Proceeding with the energy analysis, we start by multiplying both sides of (6.16) by $\rho\delta_{t\cdot}\underline{\Psi}_i^n$ and taking the sum over all cells $i \in I$ to get:

$$\sum_{i\in I}\frac{\rho V_i}{c^2}(\delta_{t\cdot}\underline{\Psi}_i^n)(\delta_{tt}\underline{\Psi}_i^n) = \underbrace{\rho\sum_{i\in I}\sum_{j\in N_i} S_{i,j}(\delta_{t\cdot}\underline{\Psi}_i^n)(\delta_{j,i}\underline{\Psi}_j^n)}_{-W} - \underbrace{\rho\sum_{i\in I^{(\mathfrak{b})}}(\delta_{t\cdot}\underline{\Psi}_i^n)\sum_{l=1}^{N_i^{(\mathfrak{b})}} S_{i,l}^{(\mathfrak{b})}\underline{v}_{\perp,i,l}^{n}}_{\mathfrak{b}} \tag{6.30}$$

Focussing on the term labelled W, we can write:

$$W = -\frac{\rho}{2}\left(\sum_{i\in I}\sum_{j\in N_i} S_{i,j}(\delta_{t\cdot}\underline{\Psi}_i^n)(\delta_{i,j}\underline{\Psi}_i^n) + \sum_{j\in I}\sum_{i\in N_j} S_{j,i}(\delta_{t\cdot}\underline{\Psi}_j^n)(\delta_{j,i}\underline{\Psi}_j^n)\right) \tag{6.31a}$$

$$= -\frac{\rho}{2}\sum_{i\in I}\left(\sum_{j\in N_i} S_{i,j}(\delta_{t\cdot}\underline{\Psi}_i^n)(\delta_{i,j}\underline{\Psi}_i^n) - \sum_{j\in N_i} S_{i,j}(\delta_{t\cdot}\underline{\Psi}_j^n)(\delta_{i,j}\underline{\Psi}_i^n)\right) \tag{6.31b}$$

$$= \frac{\rho}{2}\sum_{i\in I}\sum_{j\in N_i} S_{i,j}h_{i,j}(\delta_{i,j}\underline{\Psi}_i^n)(\delta_{t\cdot}\delta_{i,j}\underline{\Psi}_i^n) \tag{6.31c}$$

where we have used the facts that $S_{i,j} = S_{j,i}, h_{i,j} = h_{j,i}$ and the skew-symmetry property (6.15). Then, using the identities (6.28a) and (6.28b), we arrive at the energy balance:

$$\delta_{t+}\mathfrak{h}_\mathrm{i} = -\mathfrak{b} \tag{6.32}$$

where

$$\mathfrak{h}_\mathrm{i} = \sum_{i\in I}\frac{\rho V_i}{2c^2}(\delta_{t-}\underline{\Psi}_i^n)^2 + \frac{\rho}{4}\sum_{i\in I}\sum_{j\in N_i} S_{i,j}h_{i,j}(\delta_{i,j}\underline{\Psi}_i^n)(e_{t-}\delta_{i,j}\underline{\Psi}_i^n) \tag{6.33}$$

Then using (6.28c) we have:

$$\mathfrak{h}_\mathrm{i} = \sum_{i\in I}\frac{\rho V_i}{2c^2}(\delta_{t-}\underline{\Psi}_i^n)^2 + \frac{\rho}{4}\sum_{i\in I}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}\left(\mu_{t-}\underline{\Psi}_j^n - \mu_{t-}\underline{\Psi}_i^n\right)^2 - \frac{\rho k^2}{16}\sum_{i\in I}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}\left(\delta_{t-}\underline{\Psi}_j^n - \delta_{t-}\underline{\Psi}_i^n\right)^2 \tag{6.34}$$

and

$$\mathfrak{h}_\mathrm{i} \geq \sum_{i\in I}\frac{V_i}{2c^2}(\delta_{t-}\underline{\Psi}_i^n)^2 - \frac{k^2}{16}\sum_{i\in I}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}\left(\delta_{t-}\underline{\Psi}_j^n - \delta_{t-}\underline{\Psi}_i^n\right)^2 \tag{6.35}$$

and using (6.29) we have:

$$\mathfrak{h}_\mathrm{i} \geq \sum_{i\in I}\frac{V_i}{2c^2}(\delta_{t-}\underline{\Psi}_i^n)^2 - \frac{k^2}{8}\sum_{i\in I}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}\left((\delta_{t-}\underline{\Psi}_i^n)^2 + (\delta_{t-}\underline{\Psi}_j^n)^2\right) \tag{6.36}$$

which, by symmetry, reduces to:

$$\mathfrak{h}_\mathrm{i} \geq \sum_{i\in I}\frac{V_i}{2c^2}(\delta_{t-}\underline{\Psi}_i^n)^2 - \frac{k^2}{4}\sum_{i\in I}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}(\delta_{t-}\underline{\Psi}_i^n)^2 \tag{6.37}$$

Sufficient conditions for $\mathfrak{h}_\mathrm{i} \geq 0$ are then:

$$\frac{V_i}{2c^2} \geq \frac{k^2}{4}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}, \quad i \in I \tag{6.38}$$

which reduces to the following condition on the time-step:

$$k \leq \frac{1}{c}\sqrt{\min_{i\in I}\frac{2V_i}{\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}}} \tag{6.39}$$

which is a sufficient condition for stability, provided that $\mathfrak{b} \geq 0$. For example, under rigid-wall conditions we have $\mathfrak{b} = 0$ and the above is sufficient for stability.

It is worth noting that for regular Cartesian cells on the interior (i.e., under the simplifications (6.21)

that lead to the simple Cartesian scheme), the condition on the cell geometries (6.38) reduces to the usual condition on the Courant number, $\lambda \leq \sqrt{1/d}$ [26].

Frequency-independent boundary conditions

As a first example of energy-stable boundary conditions, we can make use of a frequency-independent lossy boundary condition. In this finite volume setting we would choose:

$$\underline{v}^n_{\perp,i,l} = \frac{\gamma_{i,l}}{c} \delta_t . \underline{\Psi}^n_i \tag{6.40}$$

where $\gamma_{i,l} \geq 0$ is $\gamma(\mathbf{x})$ averaged over the boundary face $\mathcal{S}^{(\mathrm{b})}_{i,l}$. We rewrite the boundary term as $\mathfrak{b} = \mathfrak{q}$ to represent a power outflow, and we have:

$$\mathfrak{q} = \frac{\rho}{c} \sum_{i \in I^{(\mathrm{b})}} \sum_{l=1}^{N^{(\mathrm{b})}_i} S^{(\mathrm{b})}_{i,l} \gamma_{i,l} (\delta_t . \underline{\Psi}^n_i)^2 \geq 0 \tag{6.41}$$

and the scheme remains stable as a whole under the condition (6.39) and $\gamma_{i,l} \geq 0$. Note that $\gamma_i \equiv 0$ gives the rigid-wall special case $\underline{v}^n_{\perp,i} = 0 \Rightarrow \mathfrak{q} = 0$ (homogeneous Neumann conditions in Ψ).

In this case, the update equation for the scheme becomes:

$$\underline{\Psi}^{n+1}_i = \frac{1}{1+a_i} \left(2\underline{\Psi}^n_i + (a_i - 1)\underline{\Psi}^{n-1}_i + \frac{c^2 k^2}{V_i} \sum_{j \in N_i} S_{i,j} \delta_{i,j} \underline{\Psi}^n_i \right) , \quad a_i = \frac{ck}{2V_i} \sum_{l=1}^{N^{(\mathrm{b})}_i} S^{(\mathrm{b})}_{i,l} \gamma_{i,l} \tag{6.42}$$

It is worth pointing out that this includes the "non-centered" staircase boundaries for the simple Cartesian scheme as a special case. More specifically, in the special case of congruent Cartesian cells throughout the entire tiling, i.e., when $\overline{\Omega}_\vartriangle = \overline{\Omega}_h$, we have $V_i = h^d, S_{i,j} \overset{i \sim j}{=} h^{d-1}, h_{i,j} \overset{i \sim j}{=} h$, and we recover the finite difference boundary update (5.55c). Additionally, for $\gamma_{i,l} = \gamma_i$, we recover the simplified form (5.56). For $d = 3$, we recover (3.97) or (3.101). More will be said about these boundary conditions in Section 6.5.1.

6.2.5 Matrix forms

As we did for finite difference schemes in Chapter 5, it will help to express these finite volume updates as matrix equations, in order to get a better idea of how these schemes relate to finite difference methods, and to derive tighter bounds for stability.

Without loss of generality, let the index set be $I = \{1, \ldots, M\}$. Then we define the diagonal matrix $\mathbf{V}$ with entries $\mathbf{V}_{ii} = V_i$, and the symmetric matrix $\mathbf{S}$ with entries:

$$\mathbf{S}_{ij} = \begin{cases} \frac{S_{i,j}}{h_{i,j}} & i \sim j \\ 0 & \text{otherwise} \end{cases} , \qquad \mathbf{S}_{ii} = -\sum_{i \sim j} \mathbf{S}_{ij} \tag{6.43}$$

Furthermore, let us define the diagonal matrix $\mathbf{C}$ with entries:

$$\mathbf{C}_{ii} = \sum_{l=1}^{N^{(\mathrm{b})}_i} S^{(\mathrm{b})}_{i,l} \gamma_{i,l} \tag{6.44}$$

Then we can write the update equation (6.42) as:

$$\mathbf{u}^{n+1} = \left(\frac{ck}{2}\mathbf{V}^{-1}\mathbf{C} + \mathbf{I}\right)^{-1} \left(\left(2\mathbf{I} + c^2k^2\mathbf{V}^{-1}\mathbf{S}\right)\mathbf{u}^n + \left(\frac{ck}{2}\mathbf{V}^{-1}\mathbf{C} - \mathbf{I}\right)\mathbf{u}^{n-1}\right) \tag{6.45}$$

where $\mathbf{u}^n$ is a column vector (size $M \times 1$) containing $\underline{\Psi}_i^n$ as its entries.

These matrices can be related to the finite difference matrix update equation (5.19) with:

$$\mathbf{B} = \frac{h}{2}\mathbf{V}^{-1}\mathbf{C}\,, \qquad \mathbf{L} = h^2\mathbf{V}^{-1}\mathbf{S} \tag{6.46}$$

In regards to a matrix-type numerical stability for the finite difference update (5.19), in the case that $\mathbf{B} = \mathbf{0}$ we require conditions (5.31a) and (5.31b) to be satisfied. The condition (5.31b) is satisfied by the application of GCT, and for (5.31a) we note that

$$\mathbf{V}^{1/2}\mathbf{L}\mathbf{V}^{-1/2} = h^2\mathbf{V}^{-1/2}\mathbf{S}\mathbf{V}^{-1/2} \tag{6.47}$$

and thus $\mathbf{L} = h^2\mathbf{V}^{-1}\mathbf{S}$ is similar to a symmetric matrix, which means that the condition (5.31a) is satisfied (similar to what was seen in Section 5.2.2).

Considering a Cartesian staircase tiling of cells $\overline{\Omega}_\triangle = \overline{\Omega}_h$, we have that $\mathbf{V} = h^d\mathbf{I}$ and $\mathbf{S} = h^{d-2}\mathbf{L}$, where $\mathbf{L}$ is the matrix defined by (5.62), i.e., the Laplacian matrix derived from non-centered finite difference boundary conditions in the previous chapter. For the energy-stable centered boundary finite difference conditions from the previous chapter, we have that $\mathbf{V} = h^d\mathbf{W}$ and $\mathbf{S} = h^{d-2}\mathbf{Z}$, where $\mathbf{W}$ and $\mathbf{Z}$ are defined by (5.101). This generalisation of centered finite difference boundary conditions will be illustrated in Section 6.5.1.

In terms of matrix stability conditions, by (5.33), and provided that $\mathfrak{b} \geq 0$, a sufficient condition for stability is:

$$k \leq \frac{1}{c}\sqrt{\frac{4}{\hat{\rho}(\mathbf{V}^{-1}\mathbf{S})}} \tag{6.48}$$

where $\hat{\rho}(\mathbf{V}^{-1}\mathbf{S})$ denotes the spectral radius of $\mathbf{V}^{-1}\mathbf{S}$. The above is the necessary and sufficient stability condition (permitting linear growth) in the case that rigid-wall conditions are employed.

We can also arrive at this condition from the energy analysis. To show this, it is first worth noting that:

$$\frac{1}{2}\sum_{i\in I}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}\left(\underline{\Psi}_j^n - \underline{\Psi}_i^n\right)^2 = \mathbf{u}^{\mathrm{T}}\mathbf{S}\mathbf{u} \tag{6.49}$$

where we have dropped the superscript n in $\mathbf{u}^n$. Then from (6.35) we have:

$$\mathfrak{h}_\mathrm{i} \geq \frac{\rho}{2}(\delta_{t-}\mathbf{u})^{\mathrm{T}}\left(\frac{1}{c^2}\mathbf{V} - \frac{k^2}{4}\mathbf{S}\right)(\delta_{t-}\mathbf{u}) \tag{6.50}$$

and by the same procedure used in Eqs. (5.113)–(5.116), it follows that (6.48) is a sufficient condition for stability ($\mathfrak{h}_\mathrm{i} \geq 0$). Finally, we note that the application of GCT to (6.48) gives (6.39). In other words, by GCT it must always be the case that

$$\min_{i\in I}\frac{2V_i}{\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}} \leq \frac{4}{\hat{\rho}(\mathbf{V}^{-1}\mathbf{S})} \tag{6.51}$$

From the point of view of optimising numerical dispersion, it is useful to know that the calculation of $\hat{\rho}(\mathbf{V}^{-1}\mathbf{S})$ can provide for a higher time-step than (6.39), since (6.39) is often too strict for certain cells

and tilings to be seen in the coming sections (in particular, using non-Cartesian cells).

6.3 An implicit finite volume scheme for the wave equation

It is straightforward to generalise this explicit finite volume scheme to an implicit formulation. Analogous to the implicit finite difference schemes that appeared in Section 4.7, we can make use of the time-averaging operator $\mu_{tt} = \frac{1}{4}(e_{t+} + 2 + e_{t-})$ to obtain an implicit finite volume scheme:

$$\frac{V_i}{c^2}\delta_{tt}\underline{\Psi}_i^n = (\theta + (1-\theta)\mu_{tt}) \sum_{j \in N_i} S_{i,j}\delta_{i,j}\underline{\Psi}_i^n - \sum_{l=1}^{N_i^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})}\underline{v}_{\perp,i,l}^n \tag{6.52}$$

where θ is a free parameter. Note that for $\theta = 1$, this scheme is the fully-explicit scheme (6.16). For this scheme we will make use of the frequency-independent lossy boundary condition (6.40). The update recursion for this scheme, in an appropriate matrix form, will be provided shortly.

Energy stability analysis

It is also straightforward to obtain a stability condition for this scheme using energy methods. Proceeding as in the explicit case, but now using identities (6.28e) and (6.28f), we can arrive at an energy balance of the form (6.32). However, in this case the stored energy becomes:

$$\mathfrak{h}_{\mathrm{i}} = \sum_{i \in I} \frac{\rho V_i}{2c^2}(\delta_{t-}\underline{\Psi}_i^n)^2 + \frac{\rho}{4}\sum_{i \in I}\sum_{j \in N_i} S_{i,j}h_{i,j}\left(\mu_{t-}\delta_{i,j}\underline{\Psi}_i^n\right)^2 - \frac{\rho\theta k^2}{16}\sum_{i \in I}\sum_{j \in N_i} S_{i,j}h_{i,j}\left(\delta_{t-}\delta_{i,j}\underline{\Psi}_i^n\right)^2 \tag{6.53}$$

which reduces to (6.34) for $\theta = 1$. Since we have chosen the boundary condition (6.40), we have $\mathfrak{b} = \mathfrak{q} \geq 0$, with $\mathfrak{q}$ defined by (6.41). For stability then, we only need to show that $\mathfrak{h}_{\mathrm{i}} \geq 0$.

Clearly, for $\theta < 0$, we have $\mathfrak{h}_{\mathrm{i}} \geq 0$, regardless of the time-step. In other words, the scheme is unconditionally stable for $\theta < 0$ (a feature that is not of interest here). As for $\theta \geq 0$, we can bound $\mathfrak{h}_{\mathrm{i}}$ as in the explicit case; making use of (6.29) we arrive at the stability condition:

$$k \leq \frac{1}{c\sqrt{\theta}}\sqrt{\min_{i \in I} \frac{2V_i}{\sum_{j \in N_i} \frac{S_{i,j}}{h_{i,j}}}} \tag{6.54}$$

Matrix forms and implicit update

It will be appropriate to use the matrix forms derived earlier in order to see the form of the update recursion. After applying the boundary condition (6.40) to (6.52) we can arrive at the matrix recursion:

$$\mathbf{M}_1\mathbf{u}^{n+1} = \mathbf{M}_2\mathbf{u}^n + \mathbf{M}_3\mathbf{u}^{n-1} \tag{6.55}$$

where

$$\mathbf{M}_1 = \frac{ck}{2}\mathbf{V}^{-1}\mathbf{C} + \mathbf{I} + \frac{c^2k^2}{4}(\theta - 1)\mathbf{V}^{-1}\mathbf{S} \tag{6.56a}$$

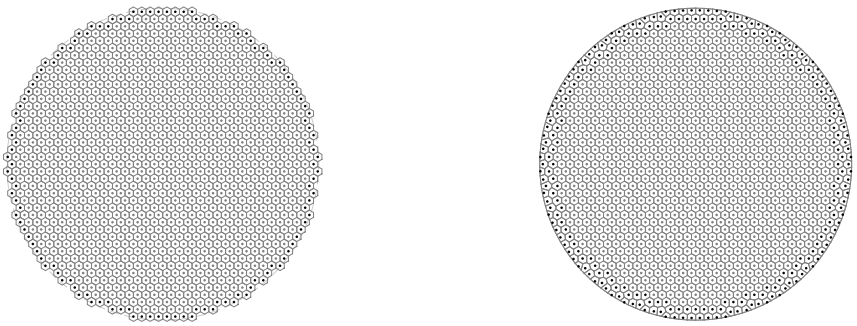
$$\mathbf{M}_2 = 2\mathbf{I} + \frac{c^2k^2}{4}(\theta + 3)\mathbf{V}^{-1}\mathbf{S} \tag{6.56b}$$

$$\mathbf{M}_3 = \frac{ck}{2}\mathbf{V}^{-1}\mathbf{C} - \mathbf{I} + \frac{c^2k^2}{4}(\theta - 1)\mathbf{V}^{-1}\mathbf{S} \tag{6.56c}$$

For $\theta \neq 1$, i.e., when $\mathbf{M}_1$ is not diagonal, this update recursion requires a linear system solution, but this is straightforward to accomplish using iterative methods [170, 213] when θ is chosen to be a small value. Jacobi iteration is also a good candidate for parallelising this process on a larger scale (see, e.g., [101], work by the current author). Finally, we note that a matrix stability condition can also be derived following Section 6.2.5. This gives:

$$k \leq \frac{1}{c\sqrt{\theta}} \sqrt{\frac{4}{\hat{\rho}(\mathbf{V}^{-1}\mathbf{S})}} \tag{6.57}$$

which reduces to (6.48) for $\theta = 1$.



(a) staircase (b) fitted

Figure 6.5: Staircase and fitted hexagonal mesh of circular region with one-metre radius and inter-cell distance of $h = 0.054\,\text{m}$ for interior hexagonal cells.

A numerical example

For a numerical example demonstrating the operation of this implicit/explicit scheme, consider the staircase and fitted (locally irregular) tilings of a circular region illustrated in Fig. 6.5. This circular region has a radius of one metre, and $h = 0.054\,\text{m}$ is used on the interior domain as an inter-cell distance for regular hexagonal cells.[11] For the tilings illustrated in Fig. 6.5 we have $\hat{\rho}(\mathbf{V}^{-1}\mathbf{S}) \lessapprox \frac{6}{h^2}$, so we can use a time-step k that adheres to the von Neumann condition of the analogous hexagonal finite difference scheme:

$$k = (\sqrt{2/3 - 4\zeta})h/c\,, \quad \zeta = \frac{\lambda^2(\theta - 1)}{4} \tag{6.58}$$

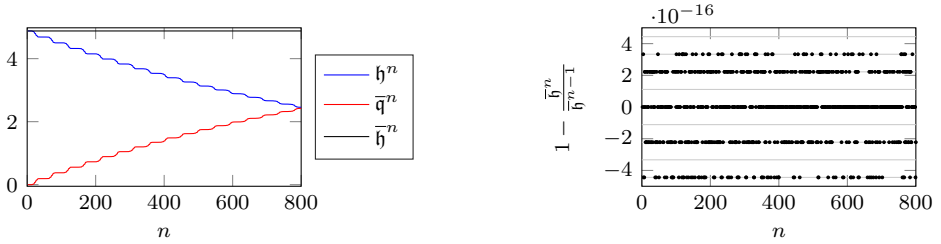
For the following example, we will use $\zeta \in \{0, 1/96, 0.0202\}$, following the analyses of analogous finite difference schemes; as derived in Section 4.7.2, $\zeta = 1/96$ leads to a fourth-order finite difference scheme, and $\zeta = 0.0202$ was chosen for an optimised one-percent critical wavenumber. We must note, however, that $\zeta = 1/96$ is not necessarily fourth-order accurate for this boundary value problem, as the fourth-order accurate update only applies over the interior domain. Nevertheless, we can expect improvements over the simple explicit form.

For the boundary condition (6.41), we set the wall admittance to be $\gamma = 0.01$, and the $\Psi(\mathbf{x}, 0)$ is a Ricker wavelet centered about the origin, with $\sigma = \sqrt{0.02}\,\text{m}$ (see (4.155)) and $\Psi_t(\mathbf{x}, 0) = 0$, and we set $c = 340\,\text{m/s}$. Outputs from these simulations, read at the centre of the circular region, are displayed

[11] We should have noted that over the interior, this finite volume scheme reduces to the implicit/explicit seven-point hexagonal scheme, which is due to the fact that when the distance between barycenters of neighbouring hexagons is h, the regular hexagonal cell has a volume of $\sqrt{3}h^2/2$ and side length $h/\sqrt{3}$. Inserting these values into (6.16) or (6.52) and considering six neighbouring cells, one has the factor $\frac{2}{3h^2}$ that appears in the simple seven-point discrete Laplacian on the hexagonal grid (defined in (4.160)). More will be said about this in the following section.

in Fig. 6.7. Along with these approximations, an "exact solution"—obtained using the fourth-order scheme and a hexagonal fitted mesh with $h = 0.01\,\mathrm{m}$—is displayed in each plot. It can be seen that the "optimised" scheme ($\zeta = 0.0202$) gives the closest agreement to the reference solution, and that the fourth-order scheme is a significant improvement to the usual explicit scheme, which demonstrates erroneous ripples due to numerical dispersion over the interior. Finally, the staircase mesh with an explicit time-integration gives the worse results, as expected, since it features staircasing effects at boundaries as well as numerical dispersion over the interior. More will be said about staircasing effects in Section 6.6.

Finally, in Fig. 6.6 we plot the energetic quantities pertaining to the implicit example illustrated by Fig. 6.7(c) ($\zeta = 1/96$). It can be seen that even in the presence of boundary losses, and although the implicit scheme requires an iterative solution, numerical energy can be conserved to machine accuracy.



(a) Energetic quantities (in Joules)

(b) Relative variations in numerical energy

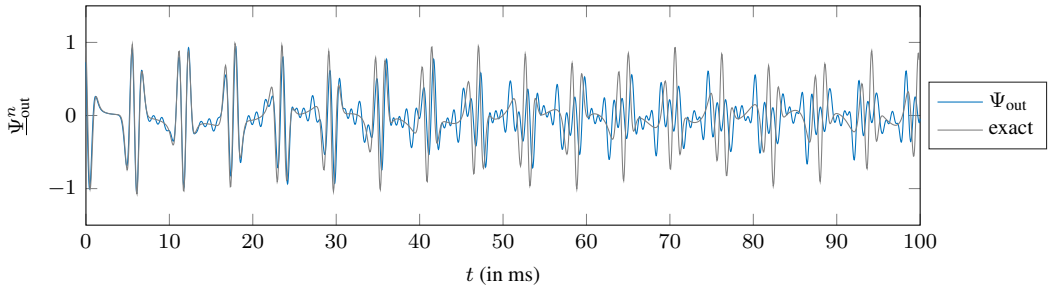
Figure 6.6: Plots demonstrating conservation of numerical energy for implicit hexagonal finite volume scheme with fitted mesh, corresponding to the output in Fig. 6.7(c) ($\zeta = 1/96$). The quantity $\overline{\mathfrak{h}}^n$ represents the total energy, which includes the stored energy on the interior ($\mathfrak{h}^n$) and accumulated losses over the boundary ($\overline{\mathfrak{q}}^n$). The variations in the total energy are integer multiples of double precision machine epsilon, as denoted by grey lines.

6.4 Finite volume cell types related to non-Cartesian finite difference schemes and digital waveguide mesh topologies

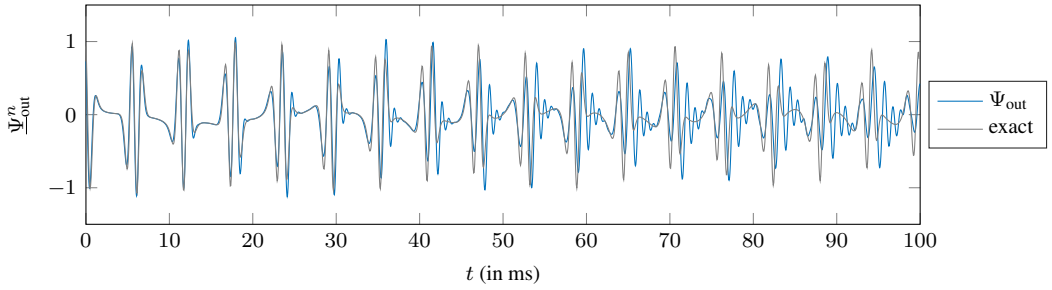
In this section we consider some cell-types that simplify to a well-known finite difference–like update that appears in digital waveguide mesh schemes [286, 288]. By deriving this update from a finite volume framework, we can unify some non-Cartesian finite difference schemes that have been seen in Chapter 4, along with others that have appeared in the digital waveguide mesh literature [286, 80, 288, 21, 144, 1, 48]. This unified view ultimately reduces to a set of constraints on cells in the tiling.

The first of these constraints is that we consider only bounded cells in Voronoi tilings. As bounded Voronoi cells are (by definition) convex polytopes, they can be decomposed into a set of d-dimensional pyramids[12] with a shared apex and bases that are the faces of the convex polytope. Consider then a Voronoi cell enclosing an interior point $\mathbf{x}_i$ with neighbours $\mathbf{x}_j, j \in N_i$, and let us define $K_i = |N_i|$. This cell can be decomposed into K_i d-dimensional pyramids with a shared apex $\mathbf{x}_i$ and with respective bases $\mathcal{S}_{i,j}$ and heights $h_{i,j}/2$. This decomposition is illustrated for a cubic cell in Fig. 6.8. By this pyramidal decomposition, and by the formula for the volume of a pyramid in d dimensions [126] (a

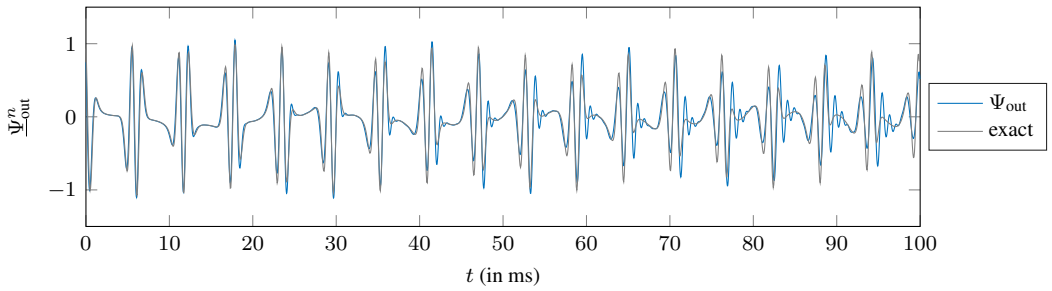
[12] A pyramid in 2-D is just a triangle, and a pyramid in 3-D can have any kind of polygonal base (not just square). A pyramid in 4-D has a polyhedral base, etc.; see [252, 126].



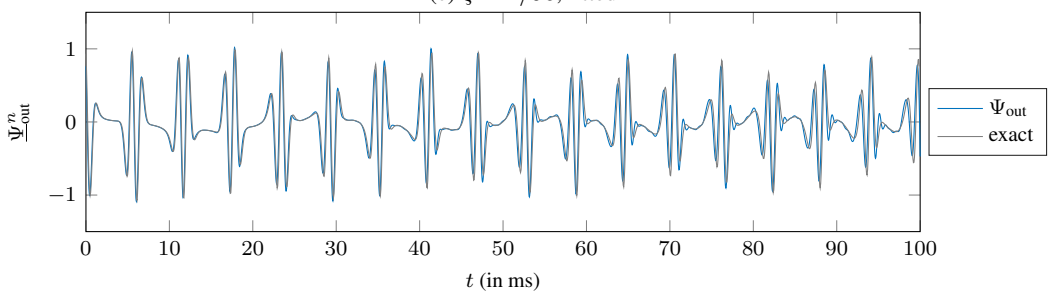
(a) $\zeta = 0$, staircase



(b) $\zeta = 0$, fitted



(c) $\zeta = 1/96$, fitted



(d) $\zeta = 0.0202$, fitted

Figure 6.7: Outputs, read at the origin, from implicit/explicit simulations using hexagonal finite volume grids over a unit radius circular region with frequency-independent lossy walls ($\gamma = 0.01$). The initial condition $\Psi(\mathbf{x}, 0)$ is a Ricker wavelet with $\sigma = \sqrt{0.02}\,\text{m}$ and $\Psi_t(\mathbf{x}, 0) = 0$, and $c = 340\,\text{m/s}$. The inter-cell distance for interior hexagonal cells is $h = 0.05\,\text{m}$ and $k = (\sqrt{2/3 - 4\zeta})h/c$.

generalisation of the formula for the area of a triangle), the volume of the cell is given by:

$$V_i = \sum_{j \in N_i} \frac{S_{i,j} h_{i,j}}{2d} \tag{6.59}$$

where each summand is the volume of an individual pyramid in d dimensions [126] (for $d = 2$, we have "base times height over two", i.e., the area of a triangle).

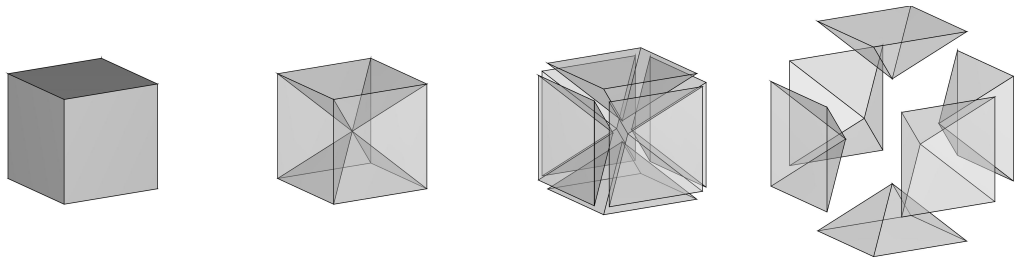


Figure 6.8: Illustration of cubic cell decomposed into six congruent pyramids.

The second constraint is that $S_{i,j}$ and $h_{i,j}$ need to be constant for $j \in N_i$. Let us assume that this is true for some cell $\overline{C}_i$ in the tiling, and then we can write $S_{i,j} = S_i$ and $h_{i,j} = h_i$ (for $j \in N_i$). We have then for this cell:

$$\frac{S_{i,j}}{V_i h_{i,j}} = \frac{2d}{K_i h_i^2} \tag{6.60}$$

and the second-order finite volume update equation then simplifies to:

$$\underline{\Psi}_i^{n+1} = 2\underline{\Psi}_i^n - \underline{\Psi}_i^{n-1} + \lambda_i^2 \kappa_i \sum_{j \in N_i} \left(\underline{\Psi}_j^n - \underline{\Psi}_i^n \right) \tag{6.61}$$

where $\kappa_i = \frac{2d}{K_i}$ and $\lambda_i = ck/h_i$ can be seen as a local Courant number for the cell. In terms of stability conditions, under these simplifications, the sufficient condition (6.38) reduces to:

$$\lambda_i \leq \sqrt{1/d} \tag{6.62}$$

and when this local Courant number is chosen to be $\lambda_i = \sqrt{1/d}$ (i.e., when the above condition is satisfied with equality), (6.61) simplifies to:

$$\underline{\Psi}_i^{n+1} = -\underline{\Psi}_i^{n-1} + \frac{2}{K_i} \sum_{j \in N_i} \underline{\Psi}_j^n \tag{6.63}$$

which is an update recursion that is traditionally used in digital waveguide mesh (DWM) schemes [286, 223, 21].

Since DWM schemes typically have K_i constant over the interior we will limit our attention to Voronoi cells that can tile space in a monohedral tiling (a tiling using only one type of congruent cell). Considering a collection of such cells that are congruent and tile some portion of space, we write $\kappa_i = \kappa$, $K_i = K$, $h_i = h$, and $\lambda_i = \lambda = ck/h$. The Voronoi tessellation of the regular Cartesian grid is one such example, with $K = 2d \Rightarrow \kappa = 1$, and the usual $\lambda \leq \sqrt{1/d}$.

Before listing the non-Cartesian cell types that adhere to these constraints, it is also worth mentioning that when *boundary cells* are also congruent to interior cells (i.e., when $V_i = V, S_{i,j} \overset{i \sim j}{=} S, h_{i,j} \overset{i \sim j}{=} h$), the scheme with frequency-independent lossy boundary conditions is encapsulated by the following

update equation:

$$\underline{\Psi}_i^{n+1} = \frac{1}{1+a_i}\left(2\underline{\Psi}_i^n + (a_i-1)\underline{\Psi}_i^{n-1} + \lambda^2\frac{2d}{K}\sum_{j\in N_i}\left(\underline{\Psi}_j^n - \underline{\Psi}_i^n\right)\right), \quad a_i = \frac{\lambda}{2}\sum_{l=1}^{N_i^{(\mathrm{b})}}\gamma_{i,l} \tag{6.64}$$

which is ultimately a generalisation of the "non-centered" Cartesian staircase boundaries. More generally, we would simply have a "staircase" approximation using non-Cartesian cells. Finally, in the simplified case that $\gamma_{i,l} = \gamma_i$, the scheme can be updated with:

$$\underline{\Psi}_i^{n+1} = \frac{1}{1+\lambda N_i^{(\mathrm{b})}\gamma_i/2}\left(2\underline{\Psi}_i^n + (\lambda N_i^{(\mathrm{b})}\gamma_i/2 - 1)\underline{\Psi}_i^{n-1} + \lambda^2\frac{2d}{K}\sum_{j\in N_i}\left(\underline{\Psi}_j^n - \underline{\Psi}_i^n\right)\right) \tag{6.65}$$

and here $N_i^{(\mathrm{b})}$ can be interpreted as the number of "ghost points" in a finite difference context.

6.4.1 Simplifications to known non-Cartesian DWM/finite difference schemes

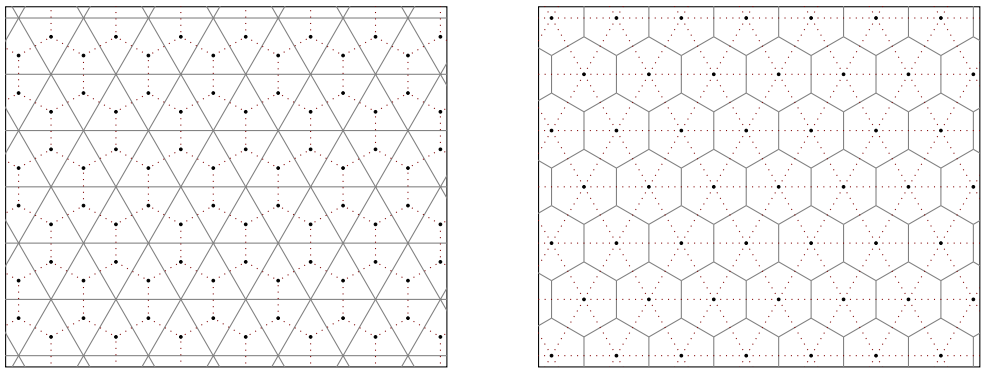
At this point we will list the non-Cartesian cell types that permit the simplification we have just discussed, and we will relate them to DWM and finite difference schemes that have been previously presented in the literature.

Other than Cartesian cells (squares, cubes, etc.), well-known examples of tilings and cells that permit these simplifications in 2-D are the equilateral triangle (in a triangular tiling) ($K = 3$) and the regular hexagon (in a hexagonal tiling) ($K = 6$). These tilings are, respectively, the Voronoi tessellations of the so-called "honeycomb lattice" [11] (which is not a lattice in the strict sense [11, 58]) and the hexagonal lattice [58], as illustrated by Fig. 6.9.[13]

For $\lambda = \sqrt{1/2}$, the resulting finite volume schemes are, respectively, equivalent to the "hexagonal digital waveguide mesh" ($K = 3, \kappa = 4/3$) [288, 21] and the "triangular digital waveguide mesh" ($K = 6, \kappa = 2/3$) [80, 21]. Leaving λ as a free parameter, these finite volume schemes are equivalent to the more general finite difference formulations found in [21, Appendix A]. As may be deduced from Fig. 6.9, the honeycomb scheme has a stencil that appears in two orientations, and each orientation pertains to a shifted hexagonal lattice. It has been shown that at the stability limit of $\lambda = \sqrt{1/2}$, the honeycomb scheme, expressed over two iterations, simplifies to the seven-point hexagonal scheme with grid spacing $\sqrt{3}h$ at its respective stability limit [21]. The von Neumann stability limit of the seven-point hexagonal scheme (with grid spacing h) is $\lambda = \sqrt{2/3}$, which is more relaxed than (6.62).

In 3-D, apart from the cube, the rhombic dodecahedron honeycomb is a well-known tiling [59] that permits such a simplification, and it is the Voronoi tessellation of the FCC lattice. The manner in which this cell would tile $\mathbb{R}^3$ is illustrated by Fig. 6.10. For $\lambda = \sqrt{1/3}$, the resulting finite volume scheme is equivalent to the "dodecahedral DWM" [144, 1], and more generally, this finite volume scheme generalises the 13-point scheme on the FCC grid [101]. It is important to note that here h is taken to represent the inter-nodal distance (which was h' in Chapter 4) rather than a Cartesian grid spacing.

[13] The triangular and hexagonal tilings are said to be *duals* [59], which should be apparent from Fig. 6.9. It is important not to confuse this the notion of a "dual lattice", since the hexagonal lattice is *self-dual*, and the "honeycomb lattice" does not have a dual in the strict sense [59].



(a) triangular tiling and honeycomb lattice (b) hexagonal tiling and hexagonal lattice

Figure 6.9: Non-Cartesian regular tilings in 2-D. Dotted lines denote cell adjacencies. In a traditional digital waveguide mesh, bi-directional unit delay lines would appear along dotted lines; see [286, 80, 288, 21].

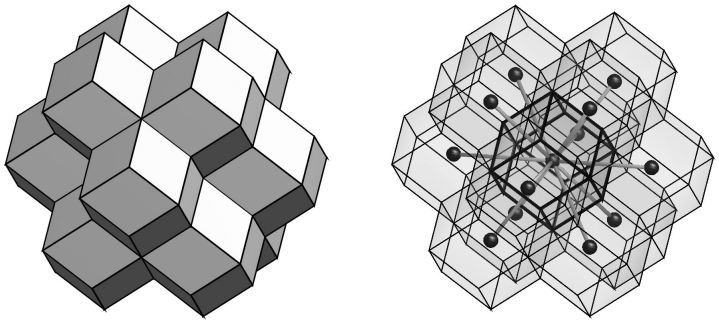


Figure 6.10: Rhombic dodecahedral cell, surrounding neighbouring cells, and underlying stencil.

Beyond the passivity limit — hexagonal and FCC grids

For the hexagonal scheme (using hexagonal cells) and for the 13-point FCC scheme (using rhombic dodecahedral cells), it is well-known that the von Neumann stability limits—respectively $\lambda^2 \leq 2/3$ [280, 21] and $\lambda^2 \leq 1/2$ [21, 196]—are more relaxed than (6.62), which is also known as the *passivity limit* [21]. In fact, we can achieve the same von Neumann conditions for these schemes from an energy analysis, using techniques for bounding eigenvalues of graph Laplacians.

First, we note that, considering congruent cells that permit a simplification to (6.63), the energetic inequality (6.35) can be rewritten as:

$$\mathfrak{h}_{\mathrm{i}} \geq \sum_{i \in I} \frac{1}{2c^2} (\upsilon_i)^2 - \frac{dk^2}{8K} \sum_{i \in I} \sum_{j \in N_i} (\upsilon_j - \upsilon_i)^2 \tag{6.66}$$

where for brevity, $\upsilon_i = \delta_{t-} \underline{\Psi}_i^n$, and where I is possibly an infinite set and $K \equiv K_i$ is the number of faces of the cell, or the kissing number of the underlying grid, and recall that $\frac{S_{i,j}}{V_i h_{i,j}} = \frac{2d}{K_i}$. Furthermore, by symmetry we can write this as:

$$\mathfrak{h}_{\mathrm{i}} \geq \sum_{i \in I} \frac{1}{2c^2} (\upsilon_i)^2 - \frac{dk^2}{4K} \underbrace{\sum_{i \sim j} (\upsilon_j - \upsilon_i)^2}_{W} \tag{6.67}$$

where now the term W represents the quadratic form for a graph Laplacian over the underlying mesh network. A technique in bounding the eigenvalues of such quadratic forms is to reconsider them as a network of smaller graph Laplacians (sub-structures). An inequality that we have used throughout the energy analyses, (6.29), is equivalent to viewing the entire graph Laplacian as interconnected *edge Laplacians* having maximum eigenvalues of two. Since each node is connected to K such edge Laplacians, we have:

$$W \leq 2K \sum_{i \in I} (v_i)^2 \tag{6.68}$$

which leads to the usual bound (6.62), which also holds over a finite grid (of congruent cells).

Let us now consider the hexagonal scheme over the full space. The underlying adjacency framework can be seen as a non-overlapping (and not space-filling) set of triangles, wherein each grid point is a vertex of three triangles. Each such triangle, interpreted as individual graphs, is an example of a *complete graph* (a complete graph with $d+1$ nodes is described by the vertices and edges of a d-dimensional simplex). Then we can use the well-known fact that the graph Laplacian of a complete graph with $d+1$ nodes has a maximum of eigenvalue of $d+1$ [6]. Thus, we have for the hexagonal scheme:

$$W \leq 9 \sum_{i \in I} (v_i)^2 \tag{6.69}$$

and using $d = 2$ and $K = 6$ in (6.67) we have:

$$\lambda \leq \sqrt{2/3} \quad \Longrightarrow \quad \mathfrak{h}_\mathrm{i} \geq 0 \tag{6.70}$$

and we have recovered the von Neumann stability condition for the seven-point hexagonal scheme. We can use the same technique for the FCC case, wherein the adjacency mesh network can be seen as a non-overlapping (and not space-filling) set of regular tetrahedra (simplices in three dimensions) with each node touching four such tetrahedra. Thus, for the FCC case we have:

$$W \leq 16 \sum_{i \in I} (v_i)^2 \tag{6.71}$$

and for $d = 3$ and $K = 12$, (6.67) gives:

$$\lambda \leq \sqrt{1/2} \quad \Longrightarrow \quad \mathfrak{h}_\mathrm{i} \geq 0 \tag{6.72}$$

which is also the von Neumann stability condition for the 13-point FCC scheme.

6.4.2 On the tetrahedral and octahedral DWMs

It is also worth mentioning two other interesting digital waveguide mesh topologies that are closely related to (6.61), yet not necessarily generalised by (6.61). First is the "tetrahedral DWM" [288], which is expressed on a grid commonly known as the so-called "diamond lattice" [11] (strictly speaking, not a lattice [58]), as illustrated in Fig. 6.11. The stencil in this DWM topology—expressed more generally as a finite difference scheme in [21]—reaches out to four neighbouring points and appears in two orientations. These neighbouring points are in the arrangement of the vertices of a regular tetrahedron. In many ways, this scheme can be seen as the 3-D analogue of the 2-D honeycomb scheme (i.e., the "hexagonal DWM" [288, 21]).

The regular tetrahedron would be the natural cell to consider for a finite volume counterpart to the

"tetrahedral scheme", since $K = 4$ and $d = 3$ in (6.61) would give the point-wise update equation for the scheme (stencil weights $\frac{3}{2}\lambda^2$). However, like the regular octahedron, the regular tetrahedron does not tile space in a monohedral tiling, and furthermore, the Voronoi cell of the diamond lattice is the "triakis truncated tetrahedron" [59], which does not satisfy our requirements for a DWM scheme. This cell, illustrated in Fig. 6.11, has four hexagonal faces and 12 triangular (isosceles) faces. Thus, the tetrahedral DWM would not seem to have a direct finite volume counterpart. On the other hand, by rewriting the tetrahedral DWM (with $\lambda = \sqrt{1/3}$) over two iterations [287], one finds that it is equivalent to the 13-point FCC scheme with $\lambda = \sqrt{1/2}$, as shown in [94] (by the current author, left out for brevity). As such, it is indirectly related to the rhombic dodecahedral finite volume scheme.

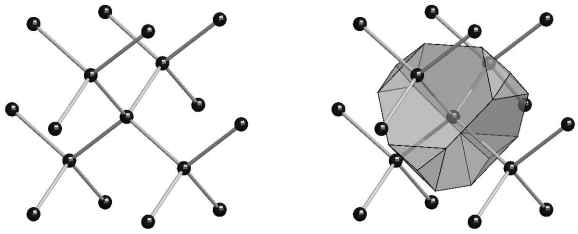


Figure 6.11: Illustration of diamond lattice topology and its Voronoi cell.

The second known DWM topology that is not necessarily generalised by a finite volume counterpart is the "octahedral DWM", or the simplest scheme on the BCC grid [21]. This scheme makes use of a stencil that reaches out to eight neighbouring points, each with stencil weight $\frac{3}{4}\lambda^2$. A natural finite volume cell to fit this stencil would be the regular octahedron (eight faces, congruent)—as illustrated in Fig. 6.12. The values $K = 8$ and $d = 3$ would give $\kappa = \frac{3}{4}$, and thus, (6.61) would describe the point-wise update of the scheme. However, it is well-known that the regular octahedron does not tile space in a monohedral tiling.[14] Meanwhile, the Voronoi cell of the BCC grid is a truncated octahedron (14 faces, six square and eight hexagonal). As such, the "octahedral scheme" on the BCC grid does not appear to have a direct finite volume analogue.

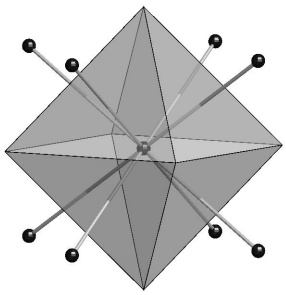


Figure 6.12: Nine-point stencil and hypothetical regular octahedral finite volume cell.

On the other hand, there is indeed a finite volume scheme associated to the Voronoi tessellation of the BCC grid: the truncated octahedral honeycomb [59]. A truncated octahedral cell and fourteen surrounding neighbours are illustrated in Fig. 6.13. In this tiling there are two inter-cell distances, depending on the face through which neighbouring cells are adjacent. Letting the inter-cell distance through hexagonal faces be $h_{i,j} = h'$, the truncated octahedral Voronoi cell has volume $V_i = 4h'^3/(3\sqrt{3})$, and the

[14] Of course, the regular octahedron could appear with other cells in a Voronoi tessellation. One such tiling is the Voronoi tessellation of the set of vertices of the rhombic dodecahedral . This is illustrated in [94, Fig. 2(b)].

eight hexagonal faces have $S_{i,j} = \sqrt{3}h'^2/4$. Meanwhile, the six square faces have $S_{i,j} = h'^2/6$ with the associated inter-cell distance $h_{i,j} = 2h'/\sqrt{3}$. The resulting stability condition, given by (6.39), is: $\lambda' \leq \sqrt{16/45}$, where $\lambda' = ck/h'$. This scheme is also equivalent to a 15-point compact scheme on the BCC grid, described by (4.255) with $\alpha_{1,1,1} = 3/4$ and $\alpha_{2,0,0} = 1/4$.[15] This scheme, however, is not an interesting special case from point of view of its numerical dispersion.[16]

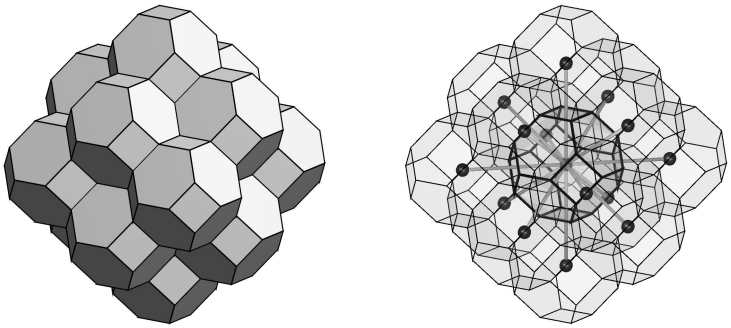


Figure 6.13: Truncated octahedral cell, surrounding neighbouring cells, and underlying stencil.

6.4.3 Two alternative digital waveguide mesh topologies

Although not necessarily of practical use, it is worth mentioning two other tilings which allow for a simplification of the form (6.61), and thus, constitute additional (novel) DWM topologies.

The first new DWM topology comes from the "rhombille" tiling [59], which is a tiling by rhombi with interior angles of $\pi/3$ and $2\pi/3$ rad and congruent faces (equal length edges), as illustrated by in Fig. 6.14. This tiling is dual to the "trihexagonal tiling" (a tiling by regular triangles and regular hexagons) and it is also the Voronoi tiling of the vertices in the trihexagonal tiling (the "Kagome lattice" [59]).[17] The finite volume update for each cell in this scheme is of the form (6.61) with $\kappa = 1/2$. This tiling could also be expressed as a traditional digital waveguide mesh scheme, with bi-directional unit digital lines along the edges of a trihexagonal tiling, and with the usual $2/K$ stencil weights applied to its K neighbours ($K = 4$). As such, it could also be called the "trihexagonal DWM". In the associated finite difference scheme (on the Kagome lattice), a five-point stencil appears in three possible orientations.

This tiling also has an analogue in three spatial dimensions, leading to another DWM topology—presented only out of interest and not recommended for practical use (for reasons to be seen shortly). First we note that the 2-D rhombille tiling can be constructed by dividing each regular hexagon in a hexagonal tiling into three equal-sized rhombi. The 3-D tiling of interest results from an analogous operation applied to the rhombic dodecahedron in a rhombic dodecahedral tiling. A rhombic dodecahedron can be divided into four congruent *rhombohedra*, each having four congruent rhombic faces; Conway calls these "obcubes" (oblique cubes) [59]. This decomposition is illustrated in Fig. 6.15.

Applying this decomposition to every cell in the rhombic dodecahedral tiling leads to another tiling, which Conway calls the *oblate cubille tiling*, or the "obcubille" tiling for short [59]. The obcubille tiling

[15] In this case, (4.257) gives the same sufficient condition for stability: $\lambda' \leq \sqrt{16/45}$, under the rescaling $\lambda' = \lambda/\sqrt{3}$ (i.e., with $\lambda = ck/h$ and letting h represent the Cartesian grid spacing $h = h'/\sqrt{3}$).

[16] A full analysis of the numerical dispersion of this scheme is left out for brevity.

[17] The *dual* of a tiling is not necessarily the *Voronoi tiling* of its vertices. For example, the Voronoi tiling of the vertices in a rhombille tiling is a hexagonal tiling, which is not the same as its dual (the trihexagonal tiling).

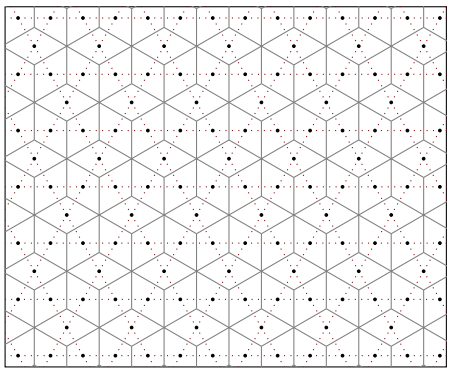


Figure 6.14: Rhombille tiling and underlying grid of points. Dotted lines denote cell adjacencies, comprising the edges trihexagonal tiling.

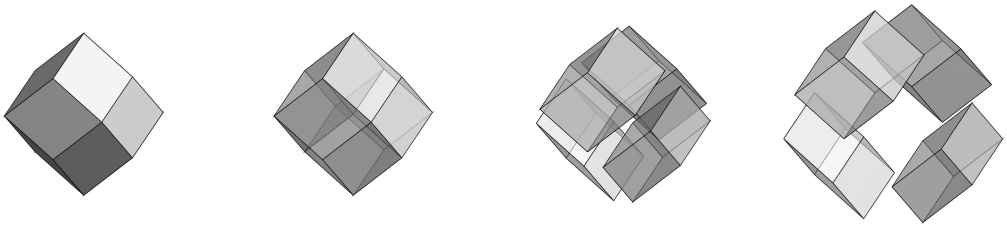


Figure 6.15: Illustration of rhombic dodecahedral cell decomposed into four obcubes.

is dual to the *truncated tetrahedrille* tiling (*trunctetrahedrille*) [59], which is the 3-D analogue of the 2-D trihexagonal tiling. Also, like the 2-D case, the obcubille tiling is the Voronoi tiling of the vertices of its dual tiling. As such, we have another finite volume scheme that reduces to (6.61) with $K = 6$ and $\kappa = 1/3$, and yet another DWM topology: the 3-D analogue of the 2-D trihexagonal DWM. In the associated finite difference scheme (operating over the vertices of the *truncated tetrahedrille* tiling), a seven-point stencil appears in four orientations. This mesh topology is illustrated by Fig. 6.16. While the resulting scheme would be stable, and consistent with the 3-D wave equation, admittedly it requires a rather complicated mesh topology (and thus is not recommended for practical use).

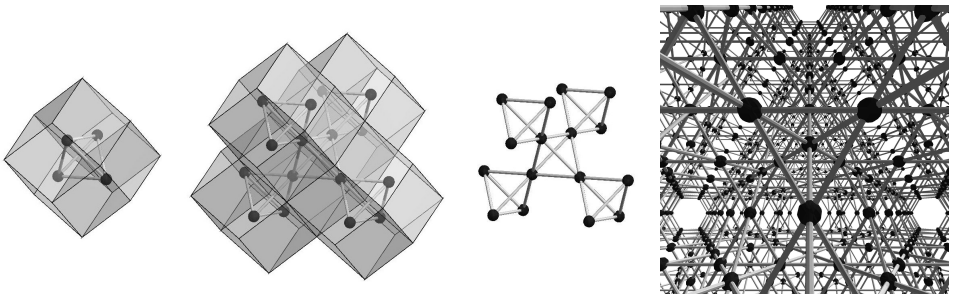


Figure 6.16: Illustrations of digital waveguide mesh topology related to obcubille tiling.

6.5 Boundary conditions in finite difference schemes from a finite volume perspective

In this section we will illustrate some of the relationships between finite volume tilings and boundary conditions associated to regular grid finite difference (and DWM) schemes, demonstrating some shortcomings of previous approaches to boundary modelling.

6.5.1 Non-centered and centered boundaries for the simple Cartesian scheme

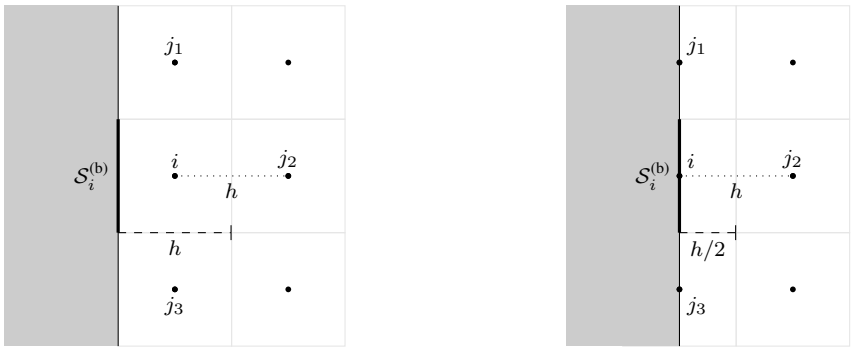
We start with the special cases of "non-centered" and "centered" conditions for the simple Cartesian scheme. Consider the partial tilings in 2-D illustrated in Fig. 6.17, where the cells are labelled by their indices at associated grid points. Let us focus on the cells labelled i, with neighbouring cells indexed by $N_i = \{j_1, j_2, j_3\}$. In Fig. 6.17(a) we have a boundary cell that is a square cell, with $V_i = h^2$, and in Fig. 6.17(b), at the boundary we have a rectangular cell that is effectively a square cell that has been cut in half, giving $V_i = h^2/2$. For the remainder of this short discussion, we will call the square cell a "full cell", and the rectangular cell a "half cell". Note also that for this half cell, we have $\mathbf{x}_i \in \mathcal{S}_i^{(\mathrm{b})}$ (the underlying grid point lies on the boundary). The rest of the cell-measures for these configurations are given in the caption of Fig. 6.17.

It is straightforward to work out from Eq. (6.42) that the boundary updates for the boundary configurations in Figs. 6.17(a) and 6.17(b) are, respectively:

$$\text{full cell:} \quad \underline{\Psi}_i^{n+1} = \frac{1}{1+\lambda\gamma_i/2}\left(2\underline{\Psi}_i^n + (\lambda\gamma_i/2-1)\underline{\Psi}_i^{n-1} + \lambda^2(\underline{\Psi}_{j_1} + \underline{\Psi}_{j_2} + \underline{\Psi}_{j_3} - 3\underline{\Psi}_i)\right) \tag{6.73a}$$

$$\text{half cell:} \quad \underline{\Psi}_i^{n+1} = \frac{1}{1+\lambda\gamma_i}\left(2\underline{\Psi}_i^n + (\lambda\gamma_i-1)\underline{\Psi}_i^{n-1} + \lambda^2(\underline{\Psi}_{j_1} + 2\underline{\Psi}_{j_2} + \underline{\Psi}_{j_3} - 4\underline{\Psi}_i)\right) \tag{6.73b}$$

where $\lambda = ck/h$ and γ_i represents a specific wall admittance. Considering the associated figure, these should be familiar as the "non-centered" and "centered" updates we have seen previously for the simple 2-D finite difference scheme. More specifically, they are essentially the "wall" updates in (5.140a) and (5.139a), respectively, with $\gamma_i = \xi_x^{-1}$. Note that in the "centered" case, the associated grid point appears on the boundary; the associated finite difference update would be different if inter-cell distances were only dictated by cell centroids.



(a) full-cell wall boundary

(b) half-cell wall boundary

Figure 6.17: Two boundary configurations at a wall oriented with underlying grid axes. Left: a full-cell wall boundary, with $V_i = h^2$, $S_{i,j_1} = S_{i,j_2} = S_{i,j_3} = h$. Right: a half-cell wall boundary, with $V_i = h^2/2$, $S_{i,j_1} = S_{i,j_3} = h/2$ and $S_{i,j_3} = h$. In both cases, $S_i^{(\mathrm{b})} = h$, $N_i \in \{j_1, j_2, j_3\}$ and $h_{i,j} = h$ for $j \in N_i$.

A corner cell in 2-D

Next we consider the analogous corner boundary cases, as illustrated in Fig. 6.18. Here we have a full-cell corner ($V_i = h^2$) in Fig. 6.18(a), and a corner cell with $V_i = h^2/4$ (a "quarter cell") in Fig. 6.18(b). It is straightforward to work out the respective boundary update equations:

$$\text{full cell:} \qquad \underline{\Psi}_i^{n+1} = \frac{1}{1+\lambda\gamma_i'/2}\left(2\underline{\Psi}_i^n + (\lambda\gamma_i'/2 - 1)\underline{\Psi}_i^{n-1} + \lambda^2(\underline{\Psi}_{j_1} + \underline{\Psi}_{j_2} - 2\underline{\Psi}_i)\right) \tag{6.74a}$$

$$\text{quarter cell:} \qquad \underline{\Psi}_i^{n+1} = \frac{1}{1+\lambda\gamma_i'}\left(2\underline{\Psi}_i^n + (\lambda\gamma_i' - 1)\underline{\Psi}_i^{n-1} + \lambda^2(2\underline{\Psi}_{j_1} + 2\underline{\Psi}_{j_2} - 4\underline{\Psi}_i)\right) \tag{6.74b}$$

where $\gamma_i' = \gamma_{i,1} + \gamma_{i,2}$. These are essentially the corner updates in (5.139b) and (5.140b), with $\gamma_{i,1} = \xi_y^{-1}$ and $\gamma_{i,2} = \xi_x^{-1}$. The stability of these configurations under the usual $\lambda \leq \sqrt{1/2}$ is guaranteed by (6.38). Analogous cases for the 3-D scheme, leading to wall, edge, and corner updates of the centered type, are left out for brevity.

(a) full-cell corner boundary

(b) quarter-cell corner boundary

Figure 6.18: Two boundary configurations at a corner, oriented with underlying grid axes. Left: a full-cell corner boundary, with $V_i = h^2$, $S_{i,j_1} = S_{i,j_2} = h$ and $S_{i,1}^{(b)} = S_{i,2}^{(b)} = h$. Right: a half-cell corner boundary, with $V_i = h^2/4$, $S_{i,j_1} = S_{i,j_2} = h/2$ and $S_{i,1}^{(b)} = S_{i,2}^{(b)} = h/2$. In both cases, $N_i \in \{j_1, j_2\}$ and $h_{i,j} = h$ for $j \in N_i$.

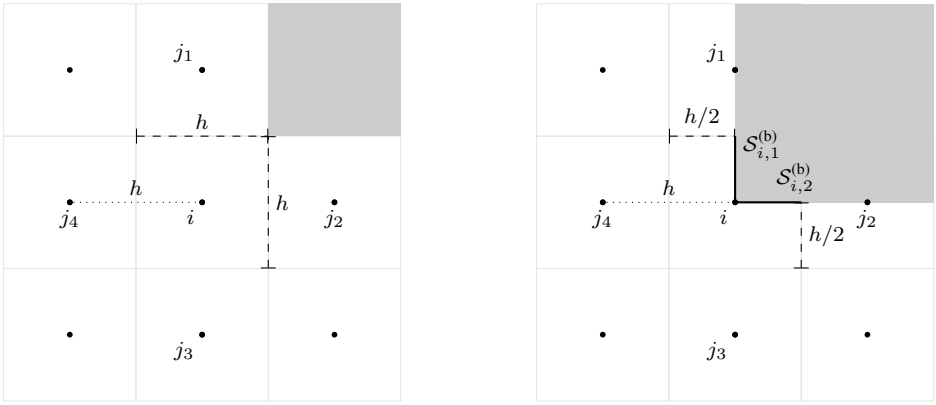
6.5.2 Revisiting the re-entrant corner problem

In addition to being to able to derive boundary conditions of the "centered" type (from a finite volume perspective), we can now combine them *locally* with other cell-types, beyond those encountered in box-domains aligned with grid axes, and we can accomplish this in a consistent and stable framework.

For example, let us reconsider the case of re-entrant corners in 2-D. First, consider the use of "non-centered" conditions, which are equivalent to full cells. In this case, the re-entrant corner case would be illustrated by Fig. 6.19(a). Clearly, in the "non-centered" case, the re-entrant corner cell is just an interior cell (does not possess a boundary face) and can be treated as such (i.e., updated with the regular interior update).

Next, let us reconsider the case that was illustrated in Fig. 5.16(c), which is related to the use of "centered conditions". Recall, from a finite difference point of view, and only considering the "ghost points" that are present (or not present), one might treat this re-entrant corner as an interior node with the usual regular update while employing centered conditions elsewhere, as in [133, 137, 135], but this approach is prone to instabilities, as shown in Section 5.4.

The same re-entrant corner is illustrated in Fig. 6.19(b), but here with a well-defined finite volume



(a) (b)

Figure 6.19: Left: a re-entrant corner configuration using only full square cells. Right: a re-entrant corner configuration with $V_i = \frac{3}{4}h^2$, $N_i \in \{j_1, j_2, j_3, j_4\}$ and $h_{i,j} = h$ for $j \in N_i$, $S_{i,j_1} = S_{i,j_2} = h/2$ and $S_{i,j_3} = S_{i,j_4} = h$, and $S^{(\mathrm{b})}_{i,1} = S^{(\mathrm{b})}_{i,2} = h/2$.

tiling. In this tiling, the re-entrant corner cell is labelled i, and the neighbours j_1, j_2 are half-cells, to which one would apply an update analogous to (6.73b). Meanwhile, the re-entrant corner node, which can be called a "three-quarter cell", can be updated with:

$$\underline{\Psi}_i^{n+1} = \frac{1}{1+\lambda\gamma_i'/3}\left(2\underline{\Psi}_i^n + (\lambda\gamma_i'/3 - 1)\underline{\Psi}_i^{n-1} + \frac{2\lambda^2}{3}(\underline{\Psi}_{j_1} + \underline{\Psi}_{j_2} + 2\underline{\Psi}_{j_3} + 2\underline{\Psi}_{j_4} - 6\underline{\Psi}_i)\right) \quad (6.75)$$

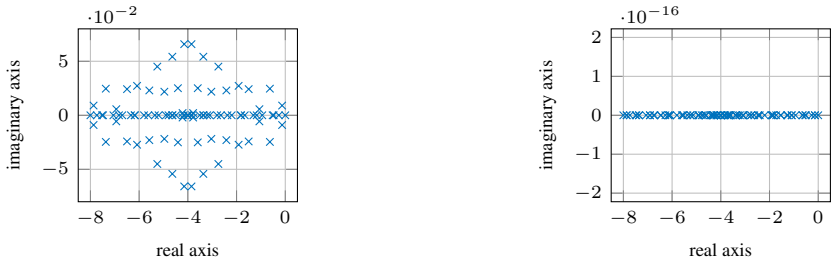
where $\gamma_i' = \gamma_{i,1} + \gamma_{i,2}$, and (6.38) is satisfied for $\lambda \leq \sqrt{1/2}$ with this configuration (i.e., it is stable). It is important to reiterate that this re-entrant corner cell is treated as a *boundary node* as it has non-empty boundary faces, so this approach is fundamentally different from the approach used for the boundary models of Kowalczyk and van Walstijn [133, 137]; their approach does not retain an underlying symmetry between adjoining cell faces, and this symmetry can lead to $\Im\{\sigma(\mathbf{L})\} = 0$.

Using this modified re-entrant corner update, one could freely apply centered conditions to eliminate ghost points throughout the rest of the domain with a guarantee of stability (to numerical precision), thereby eliminating the problem of complex eigenvalues in the underlying Laplacian matrix for the boundary models of Kowalczyk and van Walstijn [133, 137, 135] (for the five-point scheme).[18] For example, Fig. 6.20 shows the eigenvalues for a Laplacian matrix derived from the use of centered rigid wall conditions—including this modified re-entrant corner update—applied to the grid illustrated in Fig. 5.17(b), which lead to multiple unstable modes for the centered approach to eliminating ghost points.

However, despite the ability to "fix" the instabilities that can appear in such centered approaches to eliminating ghost points, it must be remembered that these boundary updates result in staircase approximations, and there is ultimately no advantage over the use of "non-centered" staircase approximations. As we will see in Section 6.6 (and has been shown in [26], and more recently in [30]), a proper tiling of the domain of interest (such as the four possible examples in Fig. 6.2) can provide significant improvements to staircase approximations.

It is also worth mentioning that, while this three-quarter cell update is certainly the finite volume configuration that would pertain to the *implied* staircase geometry $\overline{\Omega}_h^{(\cdot)}$ (defined in Chapter 5), it is not

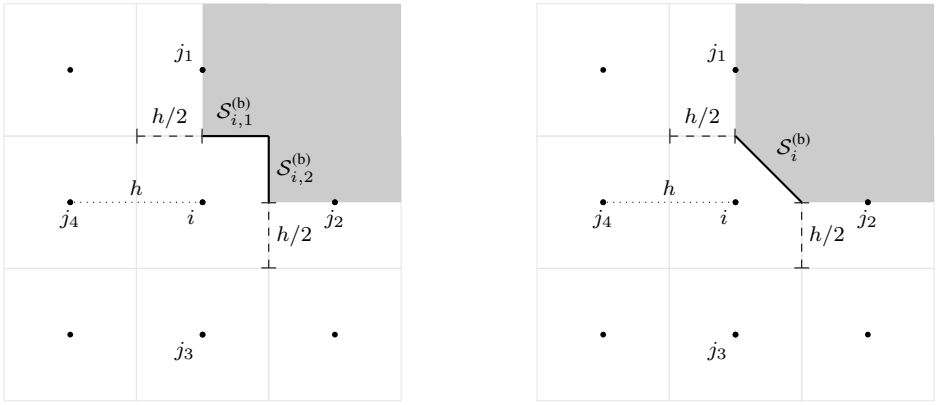
[18] Analogous updates are easily derived for re-entrant edges and corners in the 3-D simple scheme, but are left out for brevity. Instabilities for the re-entrant edge case in 3-D are discussed in [44, Section VI].



(a) (b)

Figure 6.20: Eigenvalues of Laplacian matrices for the simplest scheme with centered Neumann conditions applied to the grid illustrated in Fig. 5.17(b), treating re-entrant corners with (a) regular interior update (as recommended by [133, 137, 135]), and (b) modified re-entrant corner update (6.75).

the only modified update that would change the underlying Laplacian matrix of the DIF model (for the 2-D simple scheme) to a form that is similar to a symmetric matrix. Fig. 6.21 shows two other possible re-entrant cells that would accomplish this task, leaving surrounding cells unchanged in their respective "centered" boundary updates and guaranteeing stability with $\lambda \leq \sqrt{1/2}$ via (6.38).



(a) (b)

Figure 6.21: Two other possible re-entrant corner configurations, leading to modified boundary updates that would fix the instabilities in the frequency-independent DIF boundary model for the 2-D five-point scheme seen in the previous chapter.

6.5.3 On boundary terminations in digital waveguide mesh schemes

It is worth discussing another type of boundary model that has been used in the past for multi-dimensional digital waveguide mesh (DWM) schemes [223, 227, 180], which is derived from the one-dimensional scheme.[19] These DWM boundary models have equivalent finite volume interpretation which can shed light on some inconsistencies that have been observed with their use (e.g., in [133]).

Let us derive the basic one-dimensional boundary update. Consider the discrete node-quantity $\underline{p}_i^n \approx p(x_i, nk)$ located on a boundary with wall admittance $\gamma \geq 0$, and with interior neighbour $\underline{p}_j^n \approx p(x_j, nk)$ with $\|x_j - x_i\| = h$. For the simple 1-D scheme with centered (half-cell) lossy

[19] DWM schemes traditionally operate on abstract "scattering networks" [247], and these networks are associated to some spatial grid of points in order to establish consistency with the wave equation [286, 21, 79]. However, it is important to note that adjacencies between nodes in DWM networks are not necessarily determined by the underlying grid.

boundary conditions, the update equation for the boundary node $\underline{p}_i^n$ would be:

$$\underline{p}_i^{n+1} = \frac{1}{1+\lambda\gamma}\left(2\underline{p}_i^n + (\lambda\gamma - 1)\underline{p}_i^{n-1} + \lambda^2(2\underline{p}_j^n - 2\underline{p}_i^n)\right) \tag{6.76}$$

and for $\lambda = 1$, this reduces to:

$$\underline{p}_i^{n+1} = \frac{1}{1+\gamma}\left(2\underline{p}_j^n + (\gamma - 1)\underline{p}_i^{n-1}\right) \tag{6.77}$$

or

$$\underline{p}_i^{n+1} = (1+R)\underline{p}_j^n - R\,\underline{p}_i^{n-1}, \quad R = \frac{1-\gamma}{1+\gamma} \tag{6.78}$$

where R represents a reflection coefficient. This is the simple 1-D DWM boundary termination [286], upon which d-dimensional DWM boundary terminations are based [286, 227].

Consider now a d-dimensional DWM, where $\underline{p}_i^n$ is an approximation to $p(\mathbf{x}_i, nk)$ for $\mathbf{x}_i \in \mathbb{R}^d$. Traditionally for DWMs, the update equation for interior nodes is of the form:

$$\underline{p}_i^{n+1} = -\underline{p}_i^{n-1} + \frac{2}{K}\sum_{j\in N_i}\underline{p}_j^n \tag{6.79}$$

where here N_i denotes the set of adjacent nodes in the DWM network and $K = |N_i|$ for interior nodes, and the underlying Courant number has been fixed to $\lambda = \sqrt{1/d}$ (as is typical in DWM schemes, but not strictly necessary [21]). As we have seen, and as has been shown in the past [286, 287, 80, 21], this update equation is consistent with the multidimensional wave equation for many cell types and/or mesh network topologies.

An early convention for boundary termination in the DWM literature was to assume that boundary nodes in multi-dimensional waveguide meshes had at most one interior neighbour [286, 227], and as such, boundary nodes were treated simply by employing (6.78). This approach to boundary termination is illustrated by Fig. 6.22 for a quarter-plane domain with the 2-D Cartesian DWM.

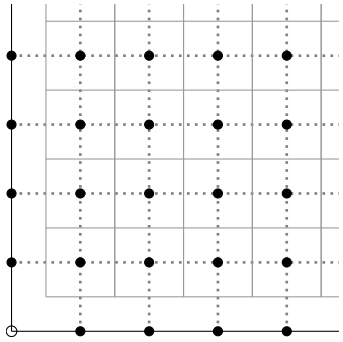


Figure 6.22: Cartesian DWM modelling of a quarter-plane. DWM nodes are denoted by solid black dots, and adjacencies are demonstrated with dotted lines. The black line denotes the boundary. Boundary nodes are not interconnected, and the corner node (hollow circle) is not connected to any interior nodes, by this DWM convention.

We rewrite this "one-dimensional" DWM boundary update as:

$$\underline{p}_i^{n+1} = \frac{1}{1+\gamma^\star}\left(2\underline{p}_j^n + (\gamma^\star - 1)\underline{p}_i^{n-1}\right) \tag{6.80}$$

where here $\underline{p}_i^n$ is a DWM boundary node adjacent to the interior DWM neighbour $\underline{p}_j^n$, with $\|\mathbf{x}_j - \mathbf{x}_i\| =$

h, and $\gamma^\star$ is meant to represent the wall admittance γ. Since this update equation is not necessarily consistent with the underlying problem, the relationship between $\gamma^\star$ and γ is left to be determined. In the DWM literature, this is known as a "one-port boundary termination" [182].

Some authors have interpreted this type of boundary condition as being "locally one-dimensional", with the Courant number being set to $\lambda = 1$ locally at the boundary, resulting in a smaller spatial step near the boundary, and constituting a "1-D ambiguity" problem [133]. Indeed, this approach is not necessarily consistent with the underlying boundary value problem, but as will be seen, there is another explanation that *can be* consistent with the multidimensional wave equation through an equivalent finite volume interpretation.

Let us now determine, in a roundabout way, if this boundary update has a finite volume analogue. First, we will manipulate (in a valid way) the finite volume update (6.42) to get in the form of (6.80), then we will try to identify a finite volume geometry that fully or partially satisfies the resulting update equation. With that said, we consider a uniform boundary wall admittance γ and some boundary cell $\overline{\mathcal{C}}_i$, whose shape is undetermined, and we assume that this cell is adjacent to an interior cell $\overline{\mathcal{C}}_j$. From (6.42) we would then arrive at the boundary update:

$$\underline{p}_i^{n+1} = \frac{1}{1+a_i}\left(2\underline{p}_i^n + (a_i-1)\underline{p}_i^{n-1} + \frac{c^2k^2S_{i,j}}{V_i h_{i,j}}(\underline{p}_j^n - \underline{p}_i^n)\right), \quad a_i = \frac{S_i^{(\mathrm{b})}ck\gamma}{2V_i} \tag{6.81}$$

As there is only one adjacent interior node, we write $S_{i,j} = S$ and $h_{i,j} = h$, and since we assume this to be a traditional DWM mesh (the adjacent interior node is updated with (6.79)), we assume $h = \sqrt{d}ck$ ($\lambda = \sqrt{1/d}$), giving:

$$\underline{p}_i^{n+1} = \frac{1}{1+a_i}\left(\left(2 - \frac{Sh}{dV_i}\right)\underline{p}_i^n + (a_i-1)\underline{p}_i^{n-1} + \frac{Sh}{dV_i}\underline{p}_j^n)\right), \quad a_i = \frac{S_i^{(\mathrm{b})}h\gamma}{2V_i\sqrt{d}} \tag{6.82}$$

The above is equal to the "one-dimensional" DWM boundary update (6.80) provided that:

$$\frac{Sh}{dV_i} = 2\,, \quad \gamma^\star = \sqrt{d}\gamma S_i^{(\mathrm{b})}/S \tag{6.83}$$

For this finite volume update to be equivalent in general with $\gamma = \gamma^\star$ for $\gamma \geq 0$, we would need to find a boundary cell that satisfies the following two conditions:

$$V_i = \frac{Sh}{2d} \tag{6.84a}$$

$$S_i^{(\mathrm{b})} = S/\sqrt{d} \tag{6.84b}$$

For $d = 1$, both conditions are satisfied by a 1-D half-cell, since, by convention, $S_i^{(\mathrm{b})} = S = 1$ and $V_i = h/2$. On the other hand, (6.84b) cannot be satisfied for $d > 1$, since it cannot be the case that $S_i^{(\mathrm{b})} \leq S$ for $d > 1$, given that the face adjacent to an interior cell is already of area S.[20] As such, we should discard (6.84b), which means that $\gamma \neq \gamma^\star$ in general.

At this point we must find a boundary cell that satisfies (6.84a), but this is simply the formula for a d-dimensional pyramid with apex $\mathbf{x}_i$, base $\mathcal{S}_{i,j}$ and height $h/2$ [126]. This implied boundary cell geometry is illustrated by Fig. 6.23 for the quarter-plane example featured in Fig. 6.22. Notice that this implied geometry preserves the non-interconnectedness of boundary nodes, as per the convention of the "1-D" DWM boundary model. But for this implied geometry, $\gamma^\star \neq \gamma$ in general, and thus the DWM

[20] For the boundary cell to enclose some d-dimensional volume, the boundary face must be greater than S.

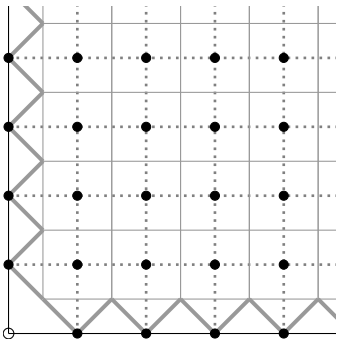


Figure 6.23: Cartesian DWM modelling of a quarter-plane and implied geometry. Note that boundary cells are not adjacent (do not share an adjoining face), and the corner node (hollow circle) has no boundary cell; these properties agree with this "1-D" DWM convention.

boundary model is only consistent with this implied finite volume geometry, if $\gamma = 0$ or $d = 1$. As such, gross errors in wall reflections should be observed from this DWM boundary model, even in the limit of small grid spacings; and indeed, such errors have been reported in the literature [133].

The many-port DWM boundary termination

Later in the development of the DWM method for room and musical acoustics, more advanced boundary terminations were proposed [1, 79, 122, 180]. In particular, the one-port boundary termination was generalised to handle many connections to interior nodes (with boundary nodes still not interconnected), which follows directly from the DWM wave scattering expressions [79]. In [180], this "many-port" boundary termination, which allows for $K_i < K$ connections to interior nodes, was written in an explicit finite difference formulation (known as "Kirchhoff-DWM" in the DWM literature [122]), and it is rewritten here as:[21]

$$\underline{p}_i^{n+1} = \frac{1}{1+\gamma^\star/K_i}\left(\frac{2}{K_i}\sum_{j\in N_i}\underline{p}_j^n + (\gamma^\star/K_i - 1)\underline{p}_i^{n-1}\right) \tag{6.85}$$

Here, N_i is the set of K_i adjacent (interior) nodal indices in the DWM network (the neighbour set). For example, a re-entrant corner boundary node with two interior connections is illustrated in Fig. 6.24.

It is straightforward to extend our finite volume interpretation to this many-port boundary termination. In this case we would start with the following finite volume update for a boundary cell:

$$\underline{p}_i^{n+1} = \frac{1}{1+a_i}\left(2\underline{p}_i^n + (a_i-1)\underline{p}_i^{n-1} + \frac{c^2k^2}{V_i}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}(\underline{p}_j^n - \underline{p}_i^n)\right), \quad a_i = \frac{S_i^{(\mathrm{b})}ck\gamma}{2V_i} \tag{6.86}$$

then, assuming each node $j \in N_i$ is updated with (6.79), we can set $S_{i,j} = S, h_{i,j} = h$, and $h = \sqrt{d}ck$. We recover the many-port boundary update (6.85) provided that the following conditions can be satisfied at the boundary cell:

$$V_i = K_i\frac{Sh}{2d} \tag{6.87a}$$

$$\gamma^\star = \sqrt{d}\gamma S_i^{(\mathrm{b})}/S \tag{6.87b}$$

[21] This expression is not exactly as appears in [180], but it is equivalent to [180, Eq. (21)] with $Y_i = Y$ and $Y_B = \gamma^\star Y$.

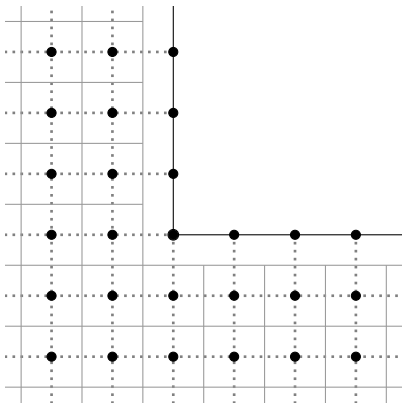


Figure 6.24: Cartesian DWM modelling of a re-entrant corner using many-port boundary termination (after [180]). DWM nodes are denoted by solid black dots, and adjacencies are demonstrated with dotted lines. The black line denotes the boundary. Boundary nodes are not interconnected by this DWM convention.

We note that by (6.87b), $\gamma^\star = \gamma$ for $\gamma > 0$ and $d > 1$ would require that (6.84b) be satisfied, which is not possible (at least in Euclidean geometry). As such, we must assume that $\gamma^\star \neq \gamma$ for $\gamma > 0$, in which case (6.87b) describes the inconsistent (yet well-defined) relationship between $\gamma^\star$ and γ. A boundary cell that would satisfy (6.87a) with K_i connections to interior nodes is simply the union of K_i d-dimensional pyramids with shared apex $\mathbf{x}_i$, bases $\mathcal{S}_{i,j}$ for $j \in N_i$ and heights $h/2$. For the re-entrant corner case, this would be illustrated by Fig. 6.25 (note the re-entrant corner cell).

So by these finite volume interpretations we have shown that the "many-port" (or one-port) DWM boundary models are inconsistent for any underlying $\overline{\Omega}_\triangle$ when $\gamma > 0$ and $d > 1$. However, it is important to note that these boundary terminations are *concretely passive* (by construction), and thus stable.[22] This stability also follows from (6.38), which is satisfied for any such DWM boundary cell (see [95]).

Finally, we note that while we have only illustrated the 2-D Cartesian DWM, this analysis holds for all DWM schemes that are expressed by (6.79), excluding the tetrahedral and octahedral schemes, since they do not have direct finite volume analogues. For example, a naive DWM boundary termination scheme for the "triangular DWM" applied to a quarter plane (following, e.g., [180]) and its underlying finite volume geometry would be as illustrated in Fig. 6.26.

Special case: 45-degree inclined boundary

There is an interesting special case geometry where the boundary updates obtained from three distinct (yet related) boundary formulations (finite volume, centered finite difference, and DWM) coincide. It is worth discussing this example briefly, as it reveals inconsistencies for the centered finite difference boundary terminations proposed in [133, 137, 135, 136], similar to those seen in the DWM boundary terminations.

Consider, as illustrated by Fig. 6.27, a 2-D half-space with a boundary that is at a 45° angle to the grid axes, with grid points aligned on the boundary. Here, the boundary cell $\overline{\mathcal{C}}_i$ has two interior neighbours, labelled j_1, j_2 with $h_{i,j_1} = h_{i,j_2} = h$ for $j \in \{j_1, j_2\}$, and the cell measures are $V_i = h^2/2$, $S_i^{(\mathrm{b})} = \sqrt{2}h$, $S_{i,j_1} = S_{i,j_2} = h$. We will derive a boundary update for this boundary node with three approaches: using finite volume, using finite difference centered conditions (after [133]), and

[22] It has been suggested otherwise in [133].

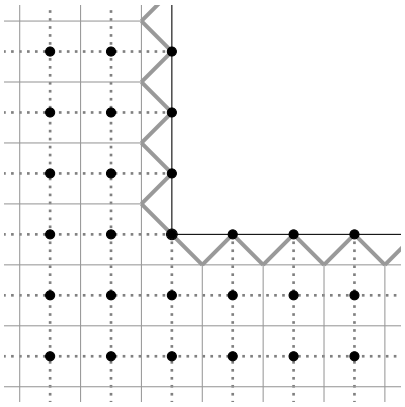


Figure 6.25: Cartesian DWM modelling of a re-entrant corner using many-port boundary termination (after [180]) and implied finite volume geometry at boundary nodes.

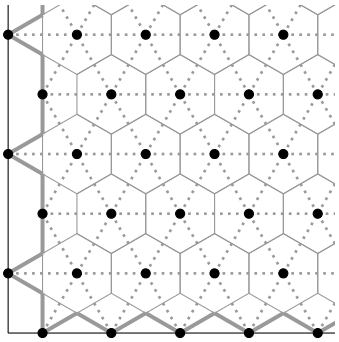


Figure 6.26: Triangular DWM modelling of a quarter-plane using many-port boundary termination (after [180]) and implied finite volume geometry at boundary nodes.

using the many-port DWM boundary. In each case we will use a different specific wall admittance and subsequently we will establish relationships between these admittance values.

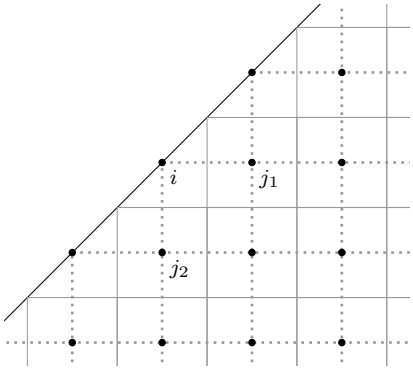


Figure 6.27: Illustration of a half-space boundary at 45° angle to Cartesian grid axes

The pressure update for this node, derived from a finite volume approach with wall admittance γ_0, would be:

$$\underline{p}_i^{n+1} = \frac{1}{1+\sqrt{2}\lambda\gamma_0}\left((2-2\lambda^2)\underline{p}_i^n + (\sqrt{2}\lambda\gamma_0 - 1)\underline{p}_i^{n-1} + \lambda^2(2\underline{p}_{j_1}^n + 2\underline{p}_{j_2}^n)\right) \tag{6.88}$$

Meanwhile, the use of "centered" boundary conditions in a finite difference approach (as in [133, 137]) would lead to the following boundary update:

$$\underline{p}_i^{n+1} = \frac{1}{1+2\lambda\gamma_1}\left((2-2\lambda^2)\underline{p}_i^n + (2\lambda\gamma_1 - 1)\underline{p}_i^{n-1} + \lambda^2(2\underline{p}_{j_1}^n + 2\underline{p}_{j_2}^n)\right) \tag{6.89}$$

where γ_1 is the wall admittance being modelled. Comparing Eqs. (6.88) and (6.89), we can see they are similar in formulation, and ultimately equivalent in the two cases: $\gamma_1 = \gamma_0/\sqrt{2}$ and $\gamma_1 = \gamma_0 = 0$.[23]

If we take the finite volume approach to be the consistent ("correct") approach, then the boundary models of Kowalczyk and van Walstijn [133, 137] would under-approximate the wall admittance, suggesting that they are inconsistent for this configuration (and thus in general). Experimental evidence for this claim can be found in [36, Section 7.1] for the analogous 3-D case shoebox-room rotated 45° about one axis, wherein it was reported that reflectance errors did not decrease as the grid was refined.[24]

In regards to the many-port DWM boundary conditions for this case, we should first point out that Fig. 6.27 is the finite volume geometry implied by their use for the Cartesian DWM scheme. As for the update equation that is obtained, we fix $\lambda^2 = 1/2$ and consider a wall admittance $\gamma^\star$. The "many-port" boundary update is then:

$$\underline{p}_i^{n+1} = \frac{1}{1+\gamma^\star/2}\left((\gamma^\star/2 - 1)\underline{p}_i^{n-1} + \underline{p}_{j_1}^n + \underline{p}_{j_2}^n\right) \tag{6.90}$$

Equating this update to the analogous finite volume update gives the relation $\gamma^\star = 2\gamma_0$ (for $\lambda^2 = 1/2$).

[23] We should point out that the neighbouring cells pictured in Fig. 6.27 are full cells and thus treated with the regular update; and they would also be treated as such in the boundary models in [133, 137]. Thus, while this is not necessarily the *implied* staircase geometry for the "centered approach", which is $\overline{\Omega}_h^{(\cdot)}$ as defined previously (see, e.g., Fig. 5.17(b)), this can be seen as the underlying finite volume geometry for that approach to eliminating ghost points (in this special case). By this underlying finite volume tiling, symmetry is preserved between "re-entrant corner" nodes and adjacent "corner" nodes in this special case for the boundary models of Kowalczyk and van Walstijn [133, 137], and as such, this configuration *would not* contribute to complex eigenvalues in the underlying Laplacian matrix (considering this as a portion of a larger, finite domain).

[24] Further experimental investigation of this special case is left as an exercise for the reader.

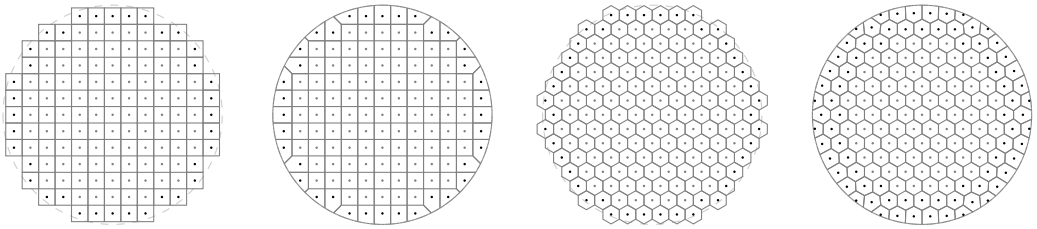
Thus, the DWM approach would over-approximate the wall admittance. Finally, for the case of $\lambda = \sqrt{1/2}$ and $\gamma_0 = \gamma_1 = \gamma^\star = 0$, the three models are in fact equivalent for this special case geometry.

6.6 On staircase approximations

In this section we illustrate some of the advantages to using finite volume approximations with tilings that fit the domain of interest (i.e., $\overline{\Omega}_\triangle = \overline{\Omega}$), rather than staircase approximations that can be obtained from a finite difference perspective (e.g., $\overline{\Omega}_\triangle = \overline{\Omega}_h \neq \overline{\Omega}$).

6.6.1 A circular membrane with free boundaries

As a first example, we consider a circular region of unit radius with free (Neumann) boundary conditions. This is an obvious domain that cannot conform to a grid of regular cells, and as such, it reveals staircasing errors for regular grid schemes (where $\overline{\Omega}_\triangle \neq \overline{\Omega}$). For this example we can compare to known analytical modal frequencies, which can be obtained through a Bessel function expansion. This example was used in [26] with fitted approximations derived from Cartesian meshes; here we also consider the use of hexagonal grids. In accordance with [26], here we set $c = 340\,\text{m/s}$.



(a) Cartesian, staircase (b) Cartesian, fitted (c) Hexagonal, staircase (d) Hexagonal, fitted

Figure 6.28: Locally irregular tilings of a circular region with one-metre radius, with $h = 0.15\,\text{m}$, using regular square and regular hexagonal cells on the interior.

For this example, we consider the use of staircase and fitted tilings using regular Cartesian cells and regular hexagons. The tilings used are illustrated in Fig. 6.28 in both staircase and fitted forms, with inter-cell distances for regular interior cells of $h = 0.15\,\text{m}$. We expect that fitted meshes will provide better approximations than the staircase meshes. For the Cartesian schemes, we choose $k = (\sqrt{1/2})h/c$, and this is stable for both the staircase and fitted meshes by (6.39). As for the hexagonal schemes, we set $k = (\sqrt{2/3})h/c$, which is permitted by (6.48) for the hexagonal meshes illustrated in Fig. 6.28.

The approximate modal frequencies, obtained from eigenvalue analyses of the state-space matrices, are compared to exact frequencies (calculated to four decimal places) in Table 6.1 (Cartesian) and Table 6.2 (hexagonal). It can be seen that, as expected, the fitted approximations provide significantly better matches to theoretical frequencies than the staircase approximations. Apart from the fifth exact modal frequency (at 207.3 Hz), the exact modal frequencies appear in degenerate mode-pairs. These mode-pairs are generally split in the numerical approximations, and such splits are more pronounced for the staircase approximations.

Also provided in the tables are approximate modal frequencies obtained from staircase approximations with three times–refined grids ($h = 0.05\,\text{m}$). Even with a grid density three times higher, approximate frequencies obtained from staircase meshes generally exhibit larger deviations from exact

modal frequencies than those obtained from fitted meshes. We can also note that while the fitted hexagonal mesh uses slightly more points ($\approx 15\%$ more) than the fitted Cartesian mesh, modal frequency approximations are *significantly* improved with the hexagonal scheme.

Approximating modal frequencies with an implicit/explicit hexagonal finite volume scheme

Finally, it is worth considering the implicit/explicit form with the fitted hexagonal mesh illustrated in Fig. 6.28. Table 6.3 shows approximate modal frequencies calculated with three choices of ζ. The first column in the table pertains to the explicit scheme (also shown in Table 6.2), and the second and third columns pertain to implicit forms.

Comparing the explicit case ($\zeta = 0$) to the implicit case of $\zeta = 1/96$, we see that modal frequencies are generally improved with the implicit form, apart from the first mode-pair. As for the third column, there is again a good agreement with exact modal frequencies. We note that the parameter $\zeta = 0.015$ is somewhat of an arbitrary choice, but it can serve to demonstrate how the performance of such implicit schemes can vary with $\zeta \neq 0$. In comparison to the case of $\zeta = 1/96$, modal frequencies are not as well approximated in the lower frequencies (below 287 Hz), but the higher modal frequencies tend to be better approximated. This shows that by choosing ζ freely, one has the ability to trade accuracy in the low frequencies for improved accuracy in the high frequencies.

Table 6.1: Exact and approximate modal frequencies (in Hz) for unit radius circular membrane, approximated using Cartesian meshes with h as indicated (in m), $k = (\sqrt{1/2})h/c$, and $c = 340$ m/s

	Staircase $h = 0.15$		Fitted, $h = 0.15$		Staircase, $h = 0.05$	
exact freq.	approx. freq.	% error	approx. freq.	% error	approx. freq.	% error
99.6315	98.6599	−0.9848	99.6938	0.0625	99.5463	−0.0856
99.6315	98.6599	−0.9848	99.6938	0.0625	99.5463	−0.0856
165.2728	159.4011	−3.6836	165.0306	−0.1468	163.7527	−0.9283
165.2728	162.0531	−1.9869	165.7053	0.2610	164.9914	−0.1706
207.3440	208.6826	0.6415	207.3720	0.0135	208.2173	0.4194
227.3376	216.5196	−4.9963	227.2598	−0.0342	225.0922	−0.9975
227.3376	216.5196	−4.9963	227.2598	−0.0342	225.0922	−0.9975
287.7468	257.8297	−11.6034	285.8184	−0.6747	281.3238	−2.2832
287.7468	282.3708	−1.9039	288.4157	0.2319	286.1135	−0.5709
288.4988	288.7235	0.0778	288.4360	−0.0218	289.4683	0.3349
288.4988	288.7235	0.0778	288.4360	−0.0218	289.4683	0.3349
347.1661	318.1941	−9.1051	345.2388	−0.5582	340.5917	−1.9303
347.1661	318.1941	−9.1051	345.2388	−0.5582	340.5917	−1.9303
362.8869	342.5979	−5.9221	360.0371	−0.7915	362.9028	0.0044
362.8869	357.9912	−1.3675	364.3788	0.4094	364.0396	0.3166

Table 6.2: Exact and approximate modal frequencies (in Hz) for unit radius circular membrane, approximated using hexagonal meshes, with h as indicated (in m), $k = (\sqrt{2/3})h/c$, and $c = 340$ m/s

	Staircase $h = 0.15$		Fitted, $h = 0.15$		Staircase, $h = 0.05$	
exact freq.	approx. freq.	% error	approx. freq.	% error	approx. freq.	% error
99.6315	97.6421	−2.0374	99.6328	0.0013	98.9798	−0.6585
99.6315	97.6421	−2.0374	99.6331	0.0016	98.9798	−0.6585
165.2728	159.7151	−3.4798	165.1975	−0.0456	163.3731	−1.1628
165.2728	159.7151	−3.4798	165.1982	−0.0452	163.3731	−1.1628
207.3440	205.5372	−0.8790	207.1313	−0.1027	207.1258	−0.1053
227.3376	209.6835	−8.4194	226.9650	−0.1641	221.2190	−2.7659
227.3376	225.3504	−0.8818	227.0824	−0.1124	226.3277	−0.4462
287.7468	271.6116	−5.9406	286.9349	−0.2830	281.9899	−2.0416
287.7468	271.6116	−5.9406	286.9662	−0.2720	281.9899	−2.0416
288.4988	283.9290	−1.6095	287.6924	−0.2803	287.8867	−0.2126
288.4988	283.9290	−1.6095	287.6976	−0.2785	287.8867	−0.2126
347.1661	323.0096	−7.4786	345.5213	−0.4760	338.7799	−2.4754
347.1661	323.0096	−7.4786	345.5953	−0.4545	338.7799	−2.4754
362.8869	356.6351	−1.7530	360.9372	−0.5402	361.8318	−0.2916
362.8869	356.7805	−1.7115	360.9383	−0.5399	361.8318	−0.2916

Table 6.3: Exact and approximate modal frequencies (in Hz) for unit radius circular membrane, approximated using hexagonal implicit/explicit scheme with ζ as indicated, $h = 0.15$ m, $k = (\sqrt{2/3 - 4\zeta})h/c$, and $c = 340$ m/s

	$\zeta = 0$		$\zeta = 1/96$		$\zeta = 0.015$	
exact freq.	approx. freq.	% error	approx. freq.	% error	approx. freq.	% error
99.6315	99.6328	0.0013	99.6591	0.0277	99.6707	0.0393
99.6315	99.6331	0.0016	99.6594	0.0280	99.6710	0.0396
165.2728	165.1975	−0.0456	165.3171	0.0268	165.3698	0.0586
165.2728	165.1982	−0.0452	165.3178	0.0272	165.3705	0.0591
207.3440	207.1313	−0.1027	207.3663	0.0107	207.4700	0.0607
227.3376	226.9650	−0.1641	227.2736	−0.0281	227.4099	0.0318
227.3376	227.0824	−0.1124	227.3915	0.0237	227.5279	0.0837
287.7468	286.9349	−0.2830	287.5544	−0.0669	287.8285	0.0284
287.7468	286.9662	−0.2720	287.5859	−0.0560	287.8601	0.0393
288.4988	287.6924	−0.2803	288.3167	−0.0631	288.5930	0.0326
288.4988	287.6976	−0.2785	288.3220	−0.0613	288.5983	0.0345
347.1661	345.5213	−0.4760	346.5946	−0.1649	347.0706	−0.0275
347.1661	345.5953	−0.4545	346.6693	−0.1433	347.1456	−0.0059
362.8869	360.9372	−0.5402	362.1578	−0.2013	362.6996	−0.0517
362.8869	360.9383	−0.5399	362.1589	−0.2010	362.7007	−0.0513

6.6.2 A box rotated in 3-D: modal frequencies and decay times

The following two test cases will serve to further illustrate staircasing effects. For these examples, we consider a 3-D box domain under various rotations with respect to the coordinate axes. The first test case will be to calculate the modal frequencies of this box domain with rigid boundaries, for which analytical frequencies are well-known (and are insensitive to rotations of the domain). The second test case will consider the same box domain, but opposing faces along one box-dimension will have non-rigid boundaries.[25]

Preliminaries

We start by defining a 3-D box domain as:

$$\mathcal{V}_{\text{box}} = [-L_x/2, L_x/2] \times [-L_y/2, L_y/2] \times [-L_z/2, L_z/2] \tag{6.91}$$

Let us also define the following 3-D rotation matrices:

$$\mathbf{R}_1 = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{R}_2 = \begin{bmatrix} 1/2 & -1/2 & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/2 & 1/2 & 1/\sqrt{2} \end{bmatrix} \tag{6.92}$$

Here, $\mathbf{R}_1$ is the rotation matrix for a $\pi/4$-rad rotation about the z-axis, and $\mathbf{R}_2$ is the rotation matrix for a $\pi/4$-rad rotation about the z-axis followed by a $\pi/4$-rad rotation about the y-axis. With these rotation matrices, we define the following rotated box domains:

$$\mathcal{V}_{\text{box}}^{(\text{R1})} = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \mathbf{R}_1\mathbf{x}', \mathbf{x}' \in \mathcal{V}_{\text{box}}\} \tag{6.93a}$$

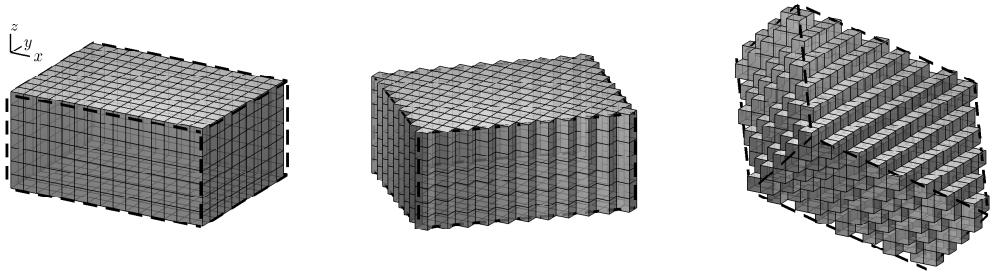
$$\mathcal{V}_{\text{box}}^{(\text{R2})} = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \mathbf{R}_2\mathbf{x}', \mathbf{x}' \in \mathcal{V}_{\text{box}}\} \tag{6.93b}$$

and for convenience we will also denote the unrotated case as $\mathcal{V}_{\text{box}}^{(\text{R0})} = \mathcal{V}_{\text{box}}$. The dimensions of the box are chosen such that L_y and L_z are irrational multiples of L_x,[26] namely $(L_x, L_y, L_z) = (4\sqrt{5}, 4\sqrt{3}, 4)\,\text{m}$. The grid spacing (inter-cell distance) h is set to $h = ck\sqrt{3}$ where $c = 340\,\text{m/s}$ and k is to be specified.

For these tests, we consider staircase and fitted approximations to these rotated box domains, obtained from the use of Cartesian and FCC grids, as well as *fitted* approximations using Cartesian and FCC grids on the interior (i.e., locally irregular). Example staircase domains using Cartesian grids ($\overline{\Omega}_h$) are illustrated in Fig. 6.29 for $h \approx 59\,\text{cm}$ ($1/k = 1\,\text{kHz}$ with $c = 340\,\text{m/s}$). Considering the use of staircased FCC grids, and thus, full rhombic dodecahedral cells in the finite volume setting, analogous staircased approximations are shown in Fig. 6.30. The analogous fitted tilings, derived from the use of cubic and rhombic dodecahedral cells are illustrated in Figs. 6.31 and 6.32, respectively.

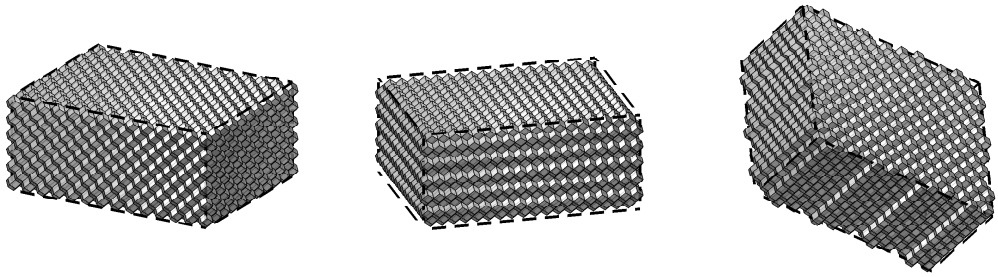
[25] Elements of the following work, carried out by the current author, have been published in [30].

[26] As such, even in the unrotated case, a regular Cartesian grid cannot fit the domain for a given h.



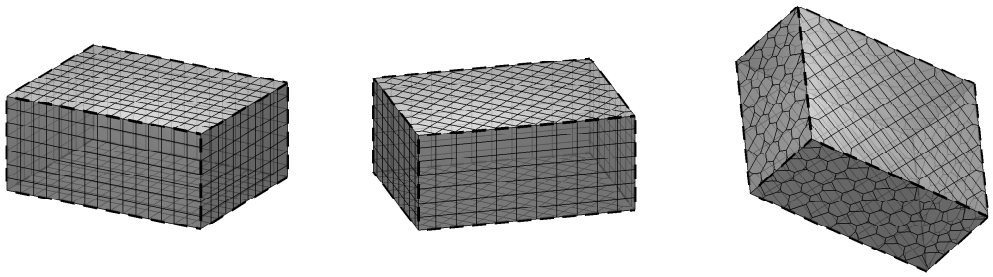
(a) R0 (b) R1 (c) R2

Figure 6.29: Surface meshes of staircased domains using cubic cells pertaining to rotated boxes $\mathcal{V}_{\text{box}}^{(\text{R0})}$, $\mathcal{V}_{\text{box}}^{(\text{R1})}$, and $\mathcal{V}_{\text{box}}^{(\text{R2})}$ with dimensions $(L_x, L_y, L_z) = (4\sqrt{5}, 4\sqrt{3}, 4)$ m. In this figure we have $h \approx 59$ cm, and the underlying box-regions are denoted by dashed lines.



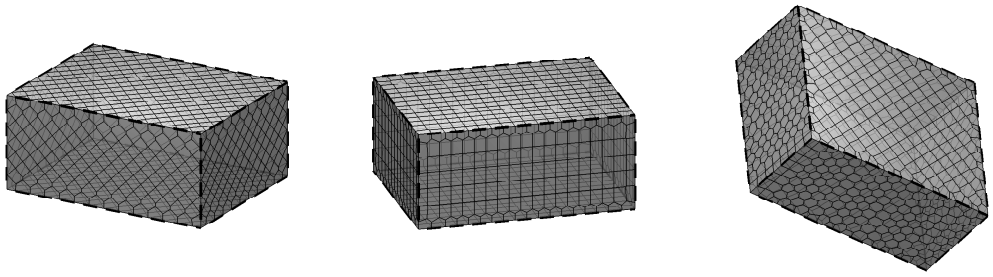
(a) R0 (b) R1 (c) R2

Figure 6.30: Surface meshes of staircased domains, using rhombic dodecahedral cells (FCC grids), pertaining to rotated boxes $\mathcal{V}_{\text{box}}^{(\text{R0})}$, $\mathcal{V}_{\text{box}}^{(\text{R1})}$, and $\mathcal{V}_{\text{box}}^{(\text{R2})}$ with dimensions $(L_x, L_y, L_z) = (4\sqrt{5}, 4\sqrt{3}, 4)$ m. In this figure we have $h \approx 59$ cm, and the underlying box-regions are denoted by dashed lines.



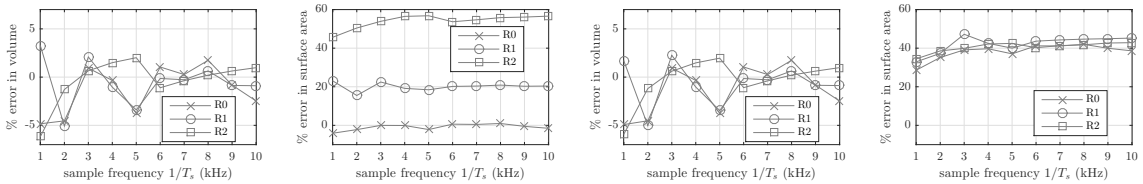
(a) R0 (b) R1 (c) R2

Figure 6.31: Surface meshes of fitted tilings, derived from the use of cubic cells on the interior with $h \approx 59$ cm.



(a) R0 (b) R1 (c) R2

Figure 6.32: Surface meshes of fitted tilings, derived from the use of rhombic dodecahedral cells (FCC grids) on the interior with $h \approx 59$ cm.



(a) Volume errors, Cartesian (b) Surface area errors, Cartesian (c) Volume errors, FCC (d) Surface area errors, FCC

Figure 6.33: Percent error in volumes and surface areas of staircase meshings using Cartesian and FCC grids, for rotated box domains $\mathcal{V}_{\text{box}}^{(\text{R0})}$, $\mathcal{V}_{\text{box}}^{(\text{R1})}$, and $\mathcal{V}_{\text{box}}^{(\text{R2})}$. In these figures, $T_s = k$.

The errors in the volumes and boundary surface areas for the staircased approximations are plotted in Fig. 6.33. It can be seen that the volume tends to be well-approximated in the limit of small grid spacings (high sample rates for λ fixed) in both Cartesian and FCC cases. Indeed, that the total volume of the staircase approximation converges to the correct value is a well-known property of such *digitizations* [129]. On the other hand, the boundary surface areas *do not* converge to the correct value in the limit of small grid spacings, apart from the trivial case of the Cartesian grid matched to the unrotated box. Again, it is relatively well-known that such digitizations do not necessarily provide consistent approximations of the surface area [129, 142].[27] The volumetric and surface area errors for the fitted tilings are virtually non-existent (below 1e-9% relative errors) and thus are not plotted.[28]

We should also mention that for the choice of grid ratio ($ck/h = \sqrt{1/3}$) is verified to be stable by (6.39) for the tilings used here and their associated underlying grids of points (not displayed).

Approximations to modal frequencies

The modal frequencies of this box domain are given by (2.41), and these are independent of the chosen rotation. With $c = 340$ m/s, the first four modes of this box-domain are approximately: 19.0 Hz, 24.5 Hz, 31.0 Hz, and 38.0 Hz. The first approximate modal frequencies from the finite volume meshes (staircase and fitted), obtained from an interior eigenvalue analysis [149] of the associated one-step state matrices $\mathbf{A}$, are displayed in Fig. 6.34, for sample rates $1/k \in \{1, 2, \ldots, 10\}$ kHz. It can be seen that for the staircase grids, modal frequencies are poorly approximated for small sample rates (coarse grids) and they tend to be better approximated in the limit of high sample rates. Meanwhile, as expected, the fitted tilings provide excellent approximations even at low sample rates (i.e., at coarse grids for λ fixed).

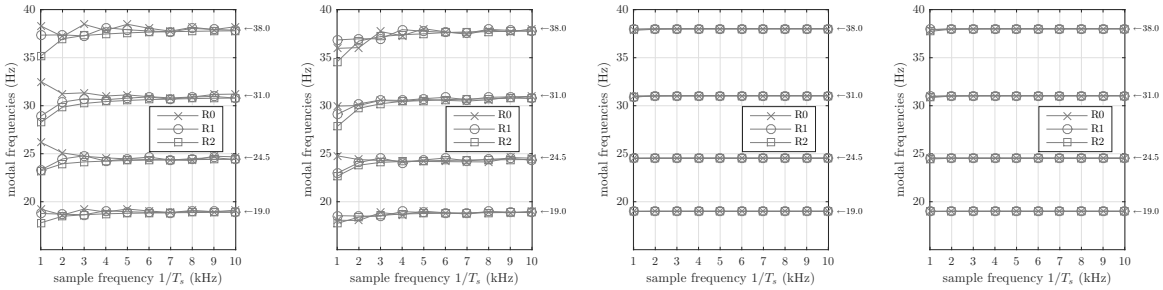
Approximations to decay times for a family of modes with non-rigid opposing faces

For the second test-case we consider the same box domains, but with opposing faces along one box-dimension being non-rigid (lossy). More specifically, we use a wall admittance of $\gamma = 0.05$ for the faces with outward normal vectors $\pm\hat{\mathbf{e}}_x$ when unrotated, and $\gamma = 0$ elsewhere. For this problem, an analytic expression for a family of modes which are constant along one direction (the x-direction when unrotated) is available.[29] These modes have time-dependent parts of the form: z_m^t with $z_m = \mathrm{e}^{\varsigma}\mathrm{e}^{j\omega_m}$,

[27] Other well-known examples include the Gauss digitizations of circles and spheres [129]. For a Gauss-digitized circle, the perimeter area converges to a value of $4/\pi$ times the value for the circle itself. Similarly, for the Gauss digitization of a sphere, the surface area converges to 1.5 times that of the sphere itself.

[28] An interesting observation from Fig. 6.33 is that the volumetric errors for staircase approximations are remarkably similar (though not entirely equivalent) at each sample rate for cubic and rhombic dodecahedral cells.

[29] This is obtained by the method of separation of variables.



(a) Cubic grid, staircase (b) FCC grid, staircase (c) Cubic grid, fitted (d) FCC grid, fitted

Figure 6.34: Frequencies of first four modes from simulated rotated box under Neumann boundary conditions under three rotations (R0,R1,R2) using staircase and fitted approximations. Analytical modal frequencies are marked on right. In these figures, $T_s = k$.

where $\omega_m = 2\pi f_{m,0}$, with $f_{m,0}$ given by (2.41), and ς given by:

$$\varsigma = \frac{c}{L_x} \log\left(\left|\frac{1-\gamma}{1+\gamma}\right|\right) \tag{6.94}$$

from which specific decay times can be obtained. For example, the time for such modes to decay by 60 dB (in intensity) is:

$$\tau_m = \frac{6L_x \log 10}{c \log\left(\left|\frac{1-\gamma}{1+\gamma}\right|\right)}, \tag{6.95}$$

which would be approximately 3.63 seconds in this example.[30]

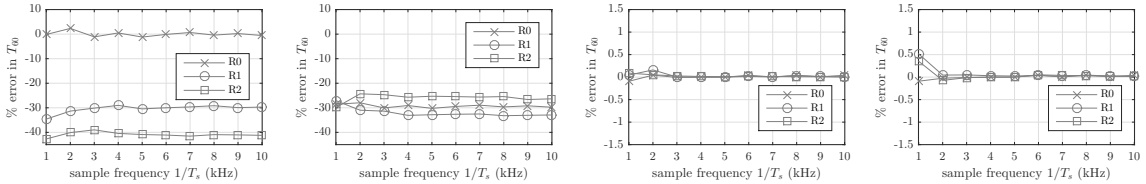
Relative errors in approximate T_{60} decay times for the first two modes in this family (19.0 Hz and 38.0 Hz) are shown in Fig. 6.35 for staircase and fitted approximations, using Cartesian (cubic) and FCC grids. It can be seen that T_{60} times are accurately reproduced with fitted meshings, even with relatively coarse grids. On the other hand, the staircase approximations exhibit significant errors that do not disappear in the limit of small cell sizes, aside from the special case of the unrotated box with a Cartesian mesh. This indicates that staircase approximations are generally *inconsistent* when boundary losses are included in the system. In effect, the resulting T_{60} times are under-approximated for the staircase grids because surface areas, and thus absorptive surfaces, are over-approximated.

It is also worth mentioning that standard frequency-domain analyses of these schemes would indicate that numerical dispersion errors are not significant sources of error, at least at the highest sample rates considered here. For example, for the cubic scheme with a 10 kHz sample rate and $\lambda = \sqrt{1/3}$, dispersion errors are less than 0.001% for the two modes considered here.[31] It is therefore safe to say that dispersion error is not a factor in the exhibited T_{60} errors. In other words, the wave equation is being accurately simulated on the interior, and the sources of errors are indeed staircasing effects.[32]

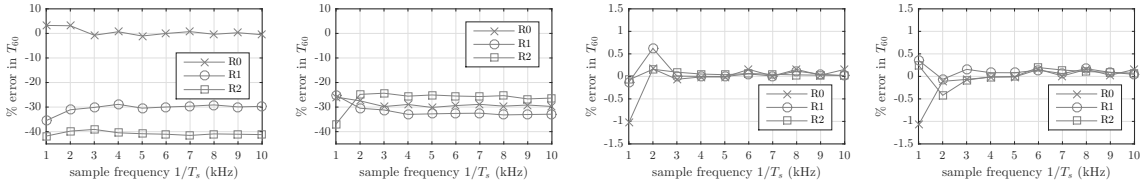
[30] It may be preferable to define the decay time in terms amplitude rather than power. The particular definition of the decay time (power, amplitude; 30 dB, 60 dB) is not important, since we will investigate *relative errors* in the decay times.

[31] This corresponds to a discretisation at 150 points per wavelength for this mode.

[32] Such a conclusion is consequently applicable to other methods that accurately solve the wave equation over the interior yet demonstrate gross staircasing effects at boundaries, such as spectral and pseudospectral methods that operate over box-partitions, e.g., [199, 165, 113].



(a) Cubic, staircase, 19.0 Hz mode (b) FCC, staircase, 19.0 Hz mode (c) Cubic, fitted, 19.0 Hz mode (d) FCC, fitted, 19.0 Hz mode



(e) Cubic, staircase, 38.0 Hz mode (f) FCC, staircase, 38.0 Hz mode (g) Cubic, fitted, 38.0 Hz mode (h) FCC, fitted, 38.0 Hz mode

Figure 6.35: Percent error in T_{60} times of two modes from simulated rotated box under resistive (lossy) boundary conditions ($\gamma = 0.05$) and three rotations (R0,R1,R2) with respect to the cubic and FCC grids. In these figures, $T_s = k$.

6.7 Frequency-dependent losses: viscothermal effects and impedance boundary conditions

In this section we move towards a more detailed room acoustics model, ultimately generalising the discrete schemes given in Chapter 3 for the 3-D Cartesian scheme to the finite volume framework. Here we consider frequency-dependent losses in the form of viscothermal effects (on the interior) and (complex) impedance boundary conditions.[33]

6.7.1 Sound absorption in air with viscothermal effects

An explicit finite volume discretisation of the integral form of (2.63) is:[34]

$$\frac{V_i}{c^2}\delta_{tt}\underline{\Psi}_i^n = (1+\tau_\eta\delta_{t-})\sum_{j\in N_i} S_{i,j}\delta_{i,j}\underline{\Psi}_i^n - \sum_{l=1}^{N_i^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})}\underline{v}_{\perp,i,l}^n \tag{6.96}$$

where $\tau_\eta = \eta/c$ is the relaxation time associated to viscothermal effects, and the underlying model considers only $\omega \ll \tau_\eta^{-1}$ and standard room acoustics indoor conditions (as discussed in Section 2.1.4), in which case τ_η is on the order of 10^{-9} s. In this viscothermal model, we redefine $\underline{v}_{\perp,i,l}^n$ to be:

$$\underline{v}_{\perp,i,l}^n \approxeq \frac{1}{S_{i,l}^{(\mathrm{b})}}\int_{\mathcal{S}_{i,l}^{(\mathrm{b})}} -(1+\tau_\eta\partial_t)\Psi(\mathbf{x},nk)\,\mathrm{d}\sigma \tag{6.97}$$

which will be used to specify impedance boundary conditions shortly; for now we consider the case of rigid walls, i.e., $\underline{v}_{\perp,i,l}^n \equiv 0$.

The update equation for this (rigid wall) scheme is:

$$\underline{\Psi}_i^{n+1} = 2\underline{\Psi}_i^n - \underline{\Psi}_i^{n-1} + (1+\tau_\eta\delta_{t-})\frac{c^2k^2}{V_i}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}\left(\underline{\Psi}_j^n - \underline{\Psi}_i^n\right) \tag{6.98}$$

[33] Elements of this section have been published in [28] and [29]. The following numerical examples were carried out by the present author, and theoretical development of the schemes was carried out cooperatively with Bilbao.

[34] See [28] for an implicit generalisation of this viscothermal scheme.

This update requires three full states ($\underline{\Psi}_i^{n+1}, \underline{\Psi}_i^n, \underline{\Psi}_i^{n-1}$) to be stored in memory, since $\underline{\Psi}_i^{n+1}$ cannot overwrite $\underline{\Psi}_i^{n-1}$ in place.

Energy analysis

A stability condition for this scheme can be derived from an energy analysis. For this analysis we make use of the identities defined in (6.28) and the following additional identities and inequality:

$$\delta_{t-} = \delta_{t\cdot} - \frac{k}{2}\delta_{tt} \tag{6.99a}$$

$$\sum_{i\in I}\sum_{j\in N_i} S_{i,j}(\underline{f}_i^n)(\delta_{i,j}\underline{g}_i^n) = -\frac{1}{2}\sum_{i\in I}\sum_{j\in N_i} S_{i,j}h_{i,j}(\delta_{i,j}\underline{f}_i^n)(\delta_{i,j}\underline{g}_i^n) \tag{6.99b}$$

$$\sum_{i\in I}\sum_{j\in N_i} S_{i,j}(\delta_{i,j}\underline{f}_i^n)^2 \leq 4\sum_{i\in I}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}(\underline{f}_i^n)^2 \tag{6.99c}$$

where $\underline{f}_i^n, \underline{g}_i^n$ are arbitrary grid functions (or cell-averages). We note that (6.99b) was derived in (6.31) (following [26]). Essentially, the identity (6.99b) and the inequality (6.99c) summarise parts of the energy analysis that appeared in Section 6.2.4 (and previously in [26, 30]).

Assuming rigid walls (for now) and taking the inner product of the scheme with $\rho\delta_{t\cdot}\Psi_i$, we have:

$$\sum_{i\in I}\frac{\rho V_i}{c^2}(\delta_{t\cdot}\underline{\Psi}_i^n)(\delta_{tt}\underline{\Psi}_i^n) = -\rho\sum_{i\in I}\sum_{j\in N_i} S_{i,j}(\delta_{t\cdot}\underline{\Psi}_i^n)((1+\tau_\eta\delta_{t-})\delta_{i,j}\underline{\Psi}_i^n) \tag{6.100}$$

Then, using (6.99a), (6.99b), and (6.28b) we arrive at the energetic balance:

$$\delta_{t+}\mathfrak{h}_\mathfrak{i} = -\mathfrak{q}_\mathfrak{i} - \mathfrak{b} \tag{6.101}$$

where

$$\mathfrak{h}_\mathfrak{i} = \sum_{i\in I}\frac{\rho V_i}{2c^2}(\delta_{t-}\underline{\Psi}_i^n)^2 + \frac{\rho}{4}\sum_{i\in I}\sum_{j\in N_i} S_{i,j}h_{i,j}(\delta_{i,j}\underline{\Psi}_i^n)(e_{t-}\delta_{i,j}\underline{\Psi}_i^n) - \frac{\rho\tau_\eta k}{8}\sum_{i\in I}\sum_{j\in N_i} S_{i,j}h_{i,j}(\delta_{t-}\delta_{i,j}\underline{\Psi}_i^n)^2 \tag{6.102a}$$

$$\mathfrak{q}_\mathfrak{i} = \frac{\rho\tau_\eta}{2}\sum_{i\in I}\sum_{j\in N_i} S_{i,j}h_{i,j}(\delta_{t\cdot}\delta_{i,j}\underline{\Psi}_i^n)^2 \tag{6.102b}$$

$$\mathfrak{b} = \rho\sum_{i\in I^{(\mathrm{b})}}(\delta_{t\cdot}\underline{\Psi}_i^n)\sum_{l=1}^{N_i^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})}\underline{v}_{\perp,i,l}^n \tag{6.102c}$$

Considering the case of rigid walls, i.e., $\underline{v}_{\perp,i,l}^n = 0 \implies \mathfrak{b} = 0$, we only need to show that $\mathfrak{h}_\mathfrak{i} \geq 0$ for stability. This is accomplished through the use of identity (6.28c) and subsequently (6.99c), giving:

$$\mathfrak{h}_\mathfrak{i} \geq \sum_{i\in I}\frac{\rho V_i}{2c^2}(\delta_{t-}\underline{\Psi}_i^n)^2 - \frac{\rho}{4}\sum_{i\in I}\sum_{j\in N_i} S_{i,j}h_{i,j}\left(\frac{k^2}{4}+\frac{k\eta}{2}\right)(\delta_{t-}\delta_{i,j}\underline{\Psi}_i^n)^2 \tag{6.103a}$$

$$\geq \sum_{i\in I}\frac{\rho V_i}{2c^2}(\delta_{t-}\underline{\Psi}_i^n)^2 - \rho\sum_{i\in I}\sum_{j\in N_i}\frac{S_{i,j}}{h_{i,j}}\left(\frac{k^2}{4}+\frac{k\eta}{2}\right)(\delta_{t-}\underline{\Psi}_i^n)^2 \tag{6.103b}$$

A sufficient condition for stability ($\mathfrak{h}_i \geq 0$), which can be checked for each cell in the tiling, is then:

$$\frac{V_i}{2c^2} \geq \left(\frac{k^2}{4} + \frac{k\eta}{2}\right) \sum_{i \in I} \sum_{j \in N_i} \frac{S_{i,j}}{h_{i,j}}, \quad i \in I \tag{6.104}$$

and we note that this reduces to (6.38) when $\eta = 0$. For a sufficient condition on the time-step, we have:

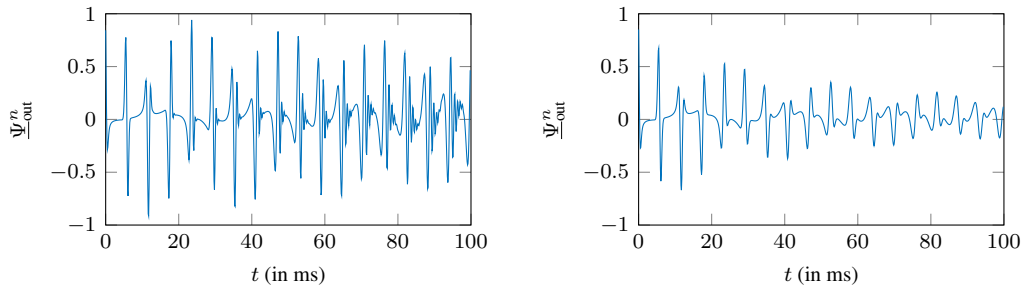
$$k \leq \sqrt{\tau_\eta^2 + \min_{i \in I} \frac{2V_i}{c^2 \sum_{j \in N_i} \frac{S_{i,j}}{h_{i,j}}}} - \tau_\eta \tag{6.105}$$

which also reduces to (6.39) for $\tau_\eta = 0$. In the case of regular Cartesian cells, the above reduces to (3.156). One can also obtain a bound involving $\hat{\rho}(\mathbf{L})$ for the associated underlying Laplacian matrix. It follows from (6.103a) and (6.49) that the following is also a sufficient stability condition:

$$k \leq \sqrt{\tau_\eta^2 + \frac{4}{c^2 \hat{\rho}(\mathbf{L})}} - \tau_\eta \tag{6.106}$$

where $\mathbf{L} = \mathbf{V}^{-1}\mathbf{S}$ and $\mathbf{V}$ and $\mathbf{S}$ are defined in (6.43).

Numerical examples using a hexagonal (fitted) mesh of a circular region are shown in Fig. 6.36, with and without viscothermal losses. We should note that for this example the viscothermal losses are exaggerated for demonstration purposes. Fig. 6.37 demonstrates numerical conservation of energy to machine accuracy, including accumulated viscothermal losses on the interior.



(a) $\eta = 0$ (lossless)

(b) $\eta =$ 1e-3 m (lossy)

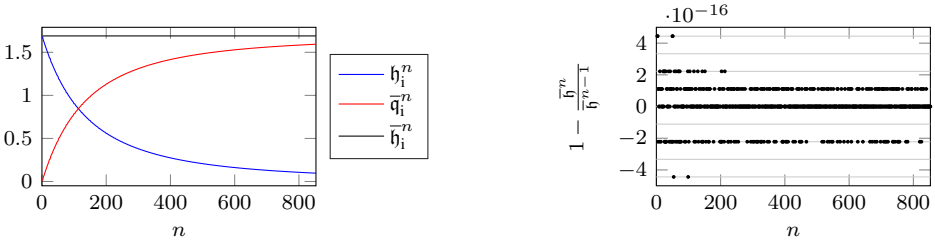
Figure 6.36: Outputs, read at the origin, from simulations using a hexagonal fitted mesh of a unit radius circular region, with rigid walls, with $\Psi(\mathbf{x}, 0)$ a Gaussian with variance 0.1 m^2, $\Psi_t(\mathbf{x}, 0) = 0$, $c = 340$ m/s, and viscothermal parameter η as indicated. The inter-cell distance for interior hexagonal cells is $h = 0.05$ m and the sample rate is set to approximately 8.3 kHz, according to (6.106).

6.7.2 Impedance boundary conditions

At this point we consider the case of frequency-dependent losses at the boundary. We make use of a finite volume generalisation of the boundary condition that appears in (3.128) for the simple Cartesian scheme, which reads:

$$\delta_{t\cdot} \underline{\Psi}_i^n = c \left(D_{i,l} \delta_{t+} \underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2}} + E_{i,l} \mu_{t+} \underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2}} + F_{i,l} \mu_{t+} \underline{g}_{i,l}^{n-\frac{1}{2}} \right) \tag{6.107}$$

where $D, E, F \geq 0$ are defined as in (2.56) and $D_{i,l}, E_{i,l}, F_{i,l} \geq 0$ are their boundary-face-averaged counterparts. Also, we have introduced the auxiliary quantities $\underline{g}_{i,l}^{n-\frac{1}{2}}$ and $\underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2}}$, which are related to



(a) Energetic quantities (in Joules)

(b) Relative variations in numerical energy

Figure 6.37: Plots demonstrating conservation of numerical energy in the presence of viscothermal losses with $\eta = 1\text{e-}3\,\mathrm{m}$, corresponding to the simulation in Fig. 6.36(b). The quantity $\overline{\mathfrak{h}}^n$ represents the total energy, which includes the stored energy on the interior ($\mathfrak{h}_\mathrm{i}^n$) and accumulated losses over the interior ($\overline{\mathfrak{q}}_\mathrm{i}^n$). The variations in the total energy are integer multiples of double precision machine epsilon, as denoted by grey lines.

$\underline{v}_{\perp,i,l}^n$ through:

$$\delta_{t+}\underline{g}_{i,l}^{n-\frac{1}{2}} = \mu_{t+}\hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2}} = \underline{v}_{\perp,i,l}^n \tag{6.108}$$

Energy analysis

The energy analysis for this boundary condition continues from the preceding analysis, but now we consider a non-zero boundary term $\mathfrak{b}$:

$$\mathfrak{b} = \rho c \sum_{i\in I}\sum_{l=1}^{N_i^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})}\left(D_{i,l}\delta_{t+}\hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2}} + E_{i,l}\mu_{t+}\hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2}} + F_{i,l}\mu_{t+}\underline{g}_{i,l}^{n-\frac{1}{2}}\right)\left(\mu_{t+}\hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2}}\right) \tag{6.109}$$

Using identity (6.28d), we can write $\mathfrak{b} = \delta_{t+}\mathfrak{h}_\mathrm{b} + \mathfrak{q}_\mathrm{b}$ and we arrive at an energy balance of the form:

$$\delta_{t+}(\mathfrak{h}_\mathrm{i} + \mathfrak{h}_\mathrm{b}) = -(\mathfrak{q}_\mathrm{i} + \mathfrak{q}_\mathrm{b}) \tag{6.110}$$

where

$$\mathfrak{h}_\mathrm{b} = \frac{\rho c}{2}\sum_{i\in I}\sum_{l=1}^{N_i^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})}\left(D_{i,l}(\hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2}})^2 + F_{i,l}(\underline{g}_{i,l}^{n-\frac{1}{2}})^2\right) \geq 0 \tag{6.111a}$$

$$\mathfrak{q}_\mathrm{b} = \rho c \sum_{i\in I}\sum_{l=1}^{N_i^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})} E_{i,l}(\mu_{t+}\hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2}})^2 \geq 0 \tag{6.111b}$$

and $\mathfrak{h}_\mathrm{i} \geq 0$, $\mathfrak{q}_\mathrm{i} \geq 0$ are as before. Numerical stability is ensured since every term is non-negative, provided that (6.105) is satisfied.

Explicit update for the scheme

The update for this scheme follows from the analogous case seen in Section 3.3.3 for the simplest scheme. First we define the quantities:

$$a_i = \frac{ck}{4V_i}\sum_{l=1}^{N_i^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})}\alpha_{i,l}\,, \qquad \alpha_{i,l} = \left(\frac{D_{i,l}}{k} + \frac{E_{i,l}}{2} + \frac{kF_{i,l}}{4}\right)^{-1} \tag{6.112}$$

The recursion of the scheme, involving the above quantities, is explicit when carried out in the following order:

$$\underline{\Psi}_i^{n+1} = \frac{1}{1+a_i}\Bigg(2\underline{\Psi}_i^n + (a_i-1)\underline{\Psi}_i^{n-1} + (1+\tau_\eta\delta_{t-})\frac{c^2k^2}{V_i}\sum_{j\in N_i} S_{i,j}\delta_{i,j}\underline{\Psi}_i^n \\ - \frac{c^2k}{V_i}\sum_{l=1}^{N_i^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})}\alpha_{i,l}\Big(D_{i,l}\hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2}} - \frac{kF_{i,l}}{2}\underline{g}_{i,l}^{n-\frac{1}{2}}\Big)\Bigg) \tag{6.113a}$$

$$\hat{\underline{v}}_{\perp,i,l}^{n+\frac{1}{2}} = \alpha_{i,l}\left(\frac{1}{2ck}(\underline{\Psi}_i^{n+1} - \underline{\Psi}_i^{n-1}) + \left(\frac{D_{i,l}}{k} - \frac{E_{i,l}}{2} - \frac{kF_{i,l}}{4}\right)\hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2}} - F_{i,l}\underline{g}_{i,l}^{n-\frac{1}{2}}\right) \tag{6.113b}$$

$$\underline{g}_{i,l}^{n+\frac{1}{2}} = \underline{g}_{i,l}^{n-\frac{1}{2}} + \frac{k}{2}\left(\hat{\underline{v}}_{\perp,i,l}^{n+\frac{1}{2}} + \hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2}}\right) \tag{6.113c}$$

It is important to note that for interior cells $i \in I^{(\mathrm{i})}$, the update is simply (6.98).

A numerical example

A numerical example is presented in order to demonstrate conservation of energy properties, incorporating frequency-dependent losses over the interior and boundaries. For this example, we consider a domed-room with an elliptical base; this is illustrated in Fig. 6.38 with a fitted tiling using rhombic dodecahedral cells on the interior. The wall impedances are set with $D = 2\text{e-}4$ (in s^{-1}), $E = 0.7$, $F = 2\text{e}3$ (in s), the initial condition $\Psi(\mathbf{x}, 0)$ is a normalised Gaussian with variance $0.05\,\text{m}^2$ and $\Psi_t(\mathbf{x}, 0) = 0$, and the viscothermal parameter is set to $\eta = 2\text{e-}6\,\text{m}$. The grid spacing for interior rhombic dodecahedral cells is $h = 12\,\text{cm}$, with the sample rate $(1/k)$ set to approximately 4 kHz (according to (6.106)). Snapshots of the solution are shown in Fig. 6.39 along two 2-D slices of the domain. Numerical energetic quantities are displayed in Fig. 6.40, demonstrating conservation to machine accuracy.

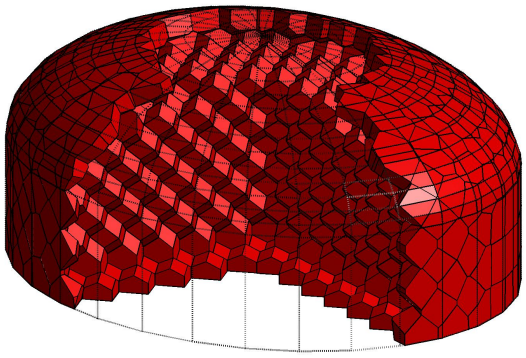
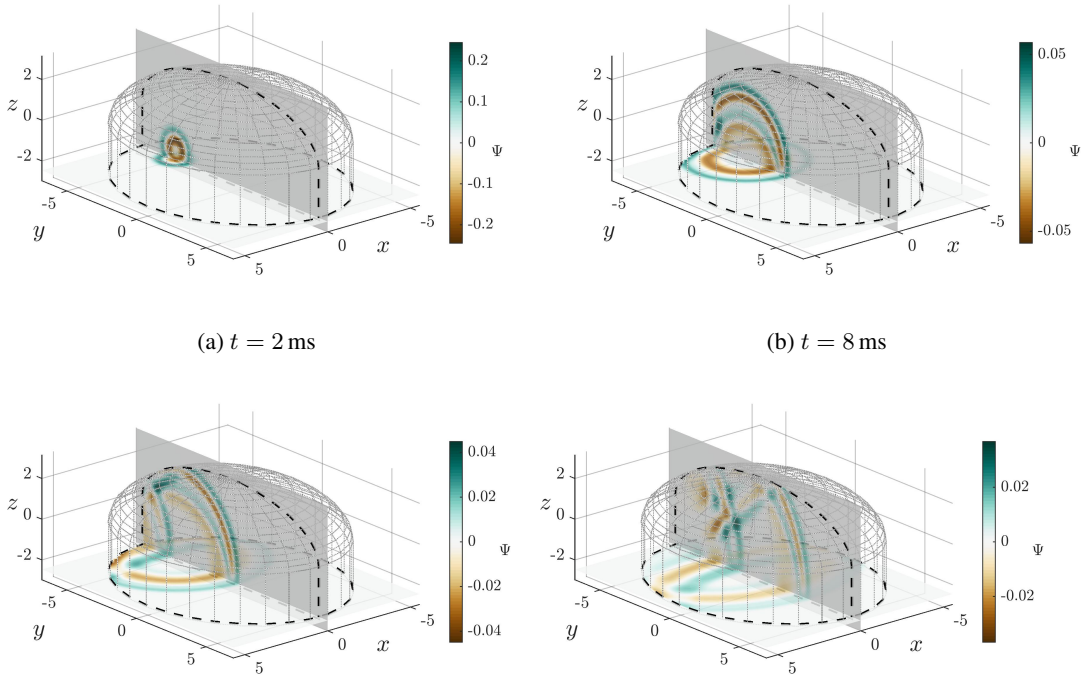


Figure 6.38: Illustration of a (partially filled) fitted tiling of a non-trivial room geometry, using regular rhombic dodecahedra over the interior and irregular cells near the boundary. The room has an elliptical base that measures 12.6 m along the major axis. In this mesh, the inter-cell distance for regular interior cells is $h \approx 64\,\text{cm}$.

6.8 General impedance boundary conditions

In this section we consider a generalisation of the preceding impedance boundary conditions, to the circuit model first presented in [26] and featured in Section 2.1.3. For a quick review, in terms of Ψ and



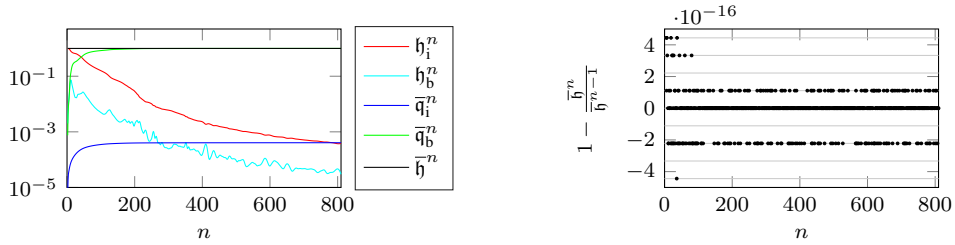
(a) $t = 2\,\text{ms}$

(b) $t = 8\,\text{ms}$

(c) $t = 12\,\text{ms}$

(d) $t = 20\,\text{ms}$

Figure 6.39: Snapshots of the time evolution of the acoustic field, at times as indicated, subject to an initial pulse. The sample rate here is 4 kHz and the FCC grid spacing is $h = 12\,\text{cm}$. Axes are in units of metres.



(a) Energetic quantities (normalised)

(b) Relative variations in numerical energy

Figure 6.40: Plots demonstrating conservation of numerical energy in the presence of frequency-dependent viscothermal and boundary losses, corresponding to the simulation in Fig. 6.39. The quantity $\overline{\mathfrak{h}}^n$ represents the total energy, which includes stored energy on the interior ($\mathfrak{h}_\text{i}^n$) and at boundaries ($\mathfrak{h}_\text{b}^n$), and accumulated losses over the interior ($\overline{\mathfrak{q}}_\text{i}^n$) and over the boundary ($\overline{\mathfrak{q}}_\text{b}^n$). The variations in the total energy are integer multiples of double precision machine epsilon, as denoted by grey lines.

$v_\perp$, this boundary condition (for $\mathbf{x} \in \Gamma$) is written as:

$$\rho\Psi_t = L^{(m)}\partial_t v_\perp^{(m)} + R^{(m)} v_\perp^{(m)} + \frac{1}{C^{(m)}} g^{(m)}\,, \quad \partial_t g^{(m)} = v_\perp^{(m)}\,, \quad m = 1, \ldots, M^{(\mathrm{b})} \tag{6.114}$$

where $L^{(m)}, R^{(m)}, C^{(m)} \geq 0$ are series inductance, resistance, and capacitance values in the mth parallel branch of the circuit (see Fig. 2.5).[35] The outward normal velocity at the boundary is decomposed into $M^{(\mathrm{b})}$ branch velocities:

$$v_\perp = \sum_{m=1}^{M^{(\mathrm{b})}} v_\perp^{(m)} \tag{6.115}$$

The number of branches $M^{(\mathrm{b})}$, and $L^{(m)}, R^{(m)}, C^{(m)}$, can vary over the boundary surface. For $M = 1$ and $(D, E, F) = \frac{1}{\rho c}(L^{(1)}, R^{(1)}, 1/C^{(1)})$, we have the boundary condition discretised in Section 6.7.2.

For this discrete boundary condition, we associate $M_{i,l}^{(\mathrm{b})}$ branch velocities $\underline{v}_{\perp,i,l,m}^n$ to each boundary face $\mathcal{S}_{i,l}^{(\mathrm{b})}$, such that:

$$\underline{v}_{\perp,i,l}^n = \sum_{m=1}^{M_{i,l}^{(\mathrm{b})}} \underline{v}_{\perp,i,l,m}^n \tag{6.116}$$

and for each branch we use the discrete boundary condition:

$$\rho\delta_{t\cdot}\underline{\Psi}_i^n = \left(L_{i,l}^{(m)} \delta_{t+} \underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} + R_{i,l}^{(m)} \mu_{t+} \underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} + \frac{1}{C_{i,l}^{(m)}} \mu_{t+} \underline{g}_{i,l}^{n-\frac{1}{2},(m)} \right) \tag{6.117}$$

where $L_{i,l}^{(m)}, R_{i,l}^{(m)}, C_{i,l}^{(m)} \geq 0$ are $L^{(m)}, R^{(m)}, C^{(m)}$ averaged over the boundary face $\mathcal{S}_{i,l}^{(\mathrm{b})}$. The auxiliary quantities $\underline{g}_{i,l}^{n-\frac{1}{2},(m)}$ and $\underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)}$ are defined by:

$$\delta_{t+} \underline{g}_{i,l}^{n-\frac{1}{2},(m)} = \mu_{t+} \underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} = \underline{v}_{\perp,i,l,m}^n \tag{6.118}$$

Energy analysis

The energy analysis considering the above boundary condition is a straightforward generalisation of the boundary condition we considered in Section 6.7.2. In this case, the boundary term $\mathfrak{b}$ is:

$$\mathfrak{b} = \sum_{i\in I} \sum_{l=1}^{N_i^{(\mathrm{b})}} \sum_{m=1}^{M_{i,l}^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})} \left(L_{i,l}^{(m)} \delta_{t+} \underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} + R_{i,l}^{(m)} \mu_{t+} \underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} + \frac{1}{C_{i,l}^{(m)}} \mu_{t+} \underline{g}_{i,l}^{n-\frac{1}{2},(m)} \right) \left(\mu_{t+} \underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} \right) \tag{6.119}$$

and applying identity (6.28d), we have $\mathfrak{b} = \delta_{t+}\mathfrak{h}_\mathfrak{b} + \mathfrak{q}_\mathfrak{b}$ and an energy balance of the form (6.110), with:

$$\mathfrak{h}_\mathfrak{b} = \frac{1}{2} \sum_{i\in I} \sum_{l=1}^{N_i^{(\mathrm{b})}} \sum_{m=1}^{M_{i,l}^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})} \left(R_{i,l}^{(m)} \left(\underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} \right)^2 + \frac{1}{C_{i,l}^{(m)}} \left(\underline{g}_{i,l}^{n-\frac{1}{2},(m)} \right)^2 \right) \geq 0 \tag{6.120a}$$

$$\mathfrak{q}_\mathfrak{b} = \sum_{i\in I} \sum_{l=1}^{N_i^{(\mathrm{b})}} \sum_{m=1}^{M_{i,l}^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})} R_{i,l}^{(m)} \left(\mu_{t+} \underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} \right)^2 \geq 0 \tag{6.120b}$$

and $\mathfrak{h}_\mathfrak{i} \geq 0$, $\mathfrak{q}_\mathfrak{i} \geq 0$ are as before. Thus, the scheme will be stable under the condition on the time-step (6.105), or (6.106).

[35] Zero admittance at DC is enforced with $C^{(m)} > 0$ [26], ruling out the "soft wall" special case.

Explicit update recursion

For the update recursion, we first define:

$$a_i = \frac{\rho c^2 k}{4V_i} \sum_{l=1}^{N_i^{(\mathrm{b})}} \sum_{m=1}^{M_{i,l}^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})} \alpha_{i,l,m} \,, \qquad \alpha_{i,l}^{(m)} = \left(\frac{L_{i,l}^{(m)}}{k} + \frac{R_{i,l}^{(m)}}{2} + \frac{k}{4C_{i,l}^{(m)}} \right)^{-1} \tag{6.121}$$

The explicit update recursion for the scheme, including viscothermal losses, is carried out in the following order:

$$\underline{\Psi}_i^{n+1} = \frac{1}{1+a_i} \Bigg(2\underline{\Psi}_i^n + (a_i - 1)\underline{\Psi}_i^{n-1} + (1 + \tau_\eta \delta_{t-}) \frac{c^2 k^2}{V_i} \sum_{j \in N_i} S_{i,j} \delta_{i,j} \underline{\Psi}_i^n \\ - \frac{c^2 k}{V_i} \sum_{l=1}^{N_i^{(\mathrm{b})}} \sum_{m=1}^{M_{i,l}^{(\mathrm{b})}} S_{i,l}^{(\mathrm{b})} \alpha_{i,l}^{(m)} \left(L_{i,l}^{(m)} \hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} - \frac{k}{2C_{i,l}^{(m)}} \underline{g}_{i,l}^{n-\frac{1}{2},(m)} \right) \Bigg) \tag{6.122a}$$

$$\hat{\underline{v}}_{\perp,i,l}^{n+\frac{1}{2},(m)} = \alpha_{i,l}^{(m)} \left(\frac{\rho}{2k} (\underline{\Psi}_i^{n+1} - \underline{\Psi}_i^{n-1}) + \left(\frac{L_{i,l}^{(m)}}{k} - \frac{R_{i,l}^{(m)}}{2} - \frac{k}{4C_{i,l}^{(m)}} \right) \hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} - \frac{1}{C_{i,l}^{(m)}} \underline{g}_{i,l}^{n-\frac{1}{2},(m)} \right) \tag{6.122b}$$

$$\underline{g}_{i,l}^{n+\frac{1}{2},(m)} = \underline{g}_{i,l}^{n-\frac{1}{2},(m)} + \frac{k}{2} \left(\hat{\underline{v}}_{\perp,i,l}^{n+\frac{1}{2},(m)} + \hat{\underline{v}}_{\perp,i,l}^{n-\frac{1}{2},(m)} \right) \tag{6.122c}$$

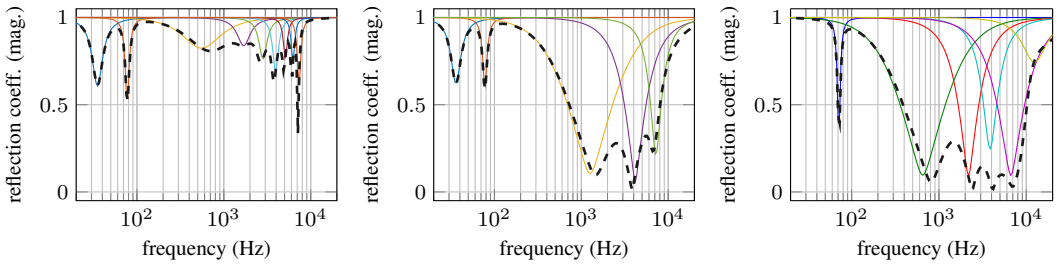
Specifying L, R, C coefficients

For the following two numerical examples, the set of L, R, C coefficients listed in Table 6.4 for three different materials will be used.[36] These three materials, which use up to nine parallel L, R, C branches ($M^{(\mathrm{b})} \leq 9$), are meant to model layered wall structures and treatments that might be found in a typical room. More specifically, Material #1 models a double curtain in front of a wall made up of two porous layers, Material #2 models a porous treatment mounted in front of a gypsum wall, and Material #3 models seat or cushion materials [30]. The frequency-dependent reflection coefficients pertaining to these materials (at normal incidence) are displayed in Fig. 6.41.

Table 6.4: Table of L, R, C coefficients

Material #1			Material #2			Material #3		
$L^{(m)}$	$R^{(m)}$	$C^{(m)}$	$L^{(m)}$	$R^{(m)}$	$C^{(m)}$	$L^{(m)}$	$R^{(m)}$	$C^{(m)}$
24.4272	1672.6264	8.4126e-7	26.0297	1749.91	7.49563e-7	22.7431	750.5833	2.1842e-7
23.2179	1210.5657	1.8279e-7	26.6955	1583.26	1.57614e-7	8.1499e-2	336.6442	7.3053e-7
9.6630e-2	39.3922	8.4863e-7	4.1593e-2	330.129	3.83857e-7	6.7680e-2	491.0421	7.8999e-8
9.4285e-2	35.8868	9.3606e-8	3.2549e-2	457.115	4.47654e-8	6.4599e-2	676.5088	2.6469e-8
9.4588e-2	55.0224	3.3616e-8	4.2137e-2	629.982	1.15506e-8	1.6493e-2	493.5131	3.4710e-8
9.8014e-2	76.1128	1.6608e-8	—	—	—	7.1131e-3	60.0617	2.3215e-8
1.0520e-1	54.2331	9.2427e-9	—	—	—	—	—	—
1.0114e-1	58.0354	6.4229e-9	—	—	—	—	—	—
8.8486e-2	99.5043	5.3150e-9	—	—	—	—	—	—

[36] These coefficients were obtained by Jonathan Botts and are presented in [30]. The problem of estimating L, R, C coefficients for a given impedance model of wall material will not be addressed here; see [30, Section V].



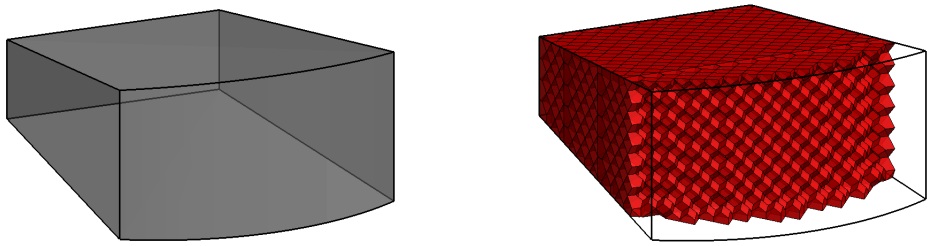
(a) Material #1 (b) Material #2 (c) Material #3

Figure 6.41: Reflection coefficient magnitudes (at normal incidence) for three materials described by L, R, C coefficients in Table 6.4. The dashed lines represent the reflection coefficients of the impedance circuit networks, and the coloured lines illustrate reflection coefficients for individual branch components.

6.8.1 A numerical example demonstrating conservation of energy

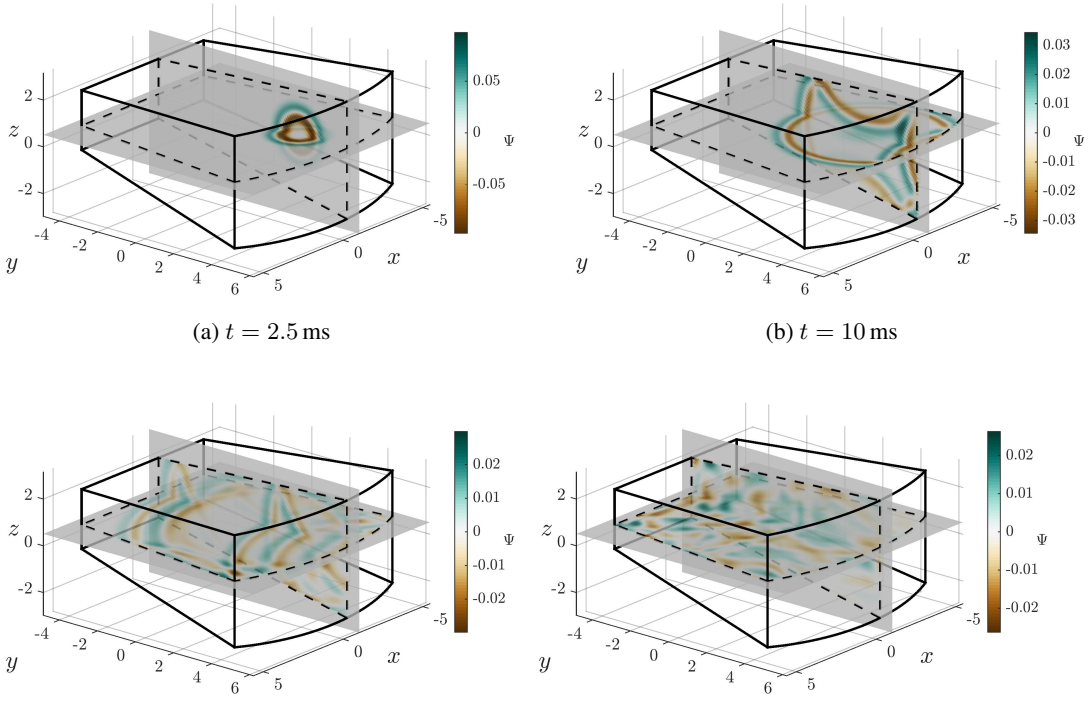
The following numerical example is intended to demonstrate conservation of numerical energy under the general impedance model with the L, R, C coefficients specified in Table 6.4, and with viscothermal losses (using $\eta = 2\text{e-}6\,\text{m}$). For this example, a fan-shaped (theatre-like) room is considered (as illustrated in Fig. 6.42(a)), for which an example fitted tiling, using rhombic dodecahedral cells over the interior, is shown in Fig. 6.42(b).

The room illustrated in Fig. 6.42(a) can be seen as having six faces: five flat and one curved (piecewise-polygonal). In terms of the materials that are associated to each face, Material #1 is used for the curved face, Material #3 is used for the largest flat face (the bottom face) and Material #2 for the remaining four flat faces. The initial condition $\Psi(\mathbf{x}, 0)$ is a Gaussian with variance $0.05\,\text{m}^2$ and $\Psi_t(\mathbf{x}, 0) = 0$, and for this example, a fitted mesh derived from an FCC grid with $h = 0.12\,\text{m}$ is used and the sample rate is set to $4.05\,\text{kHz}$ ($\lambda \approx 0.7$). Snapshots of the numerical solution are shown in Fig. 6.43, and energetic quantities and conservation of numerical energy can be seen in Fig. 6.44.



(a) $\overline{\Omega}$ (b) $\overline{\Omega}_\triangle$ (partially full)

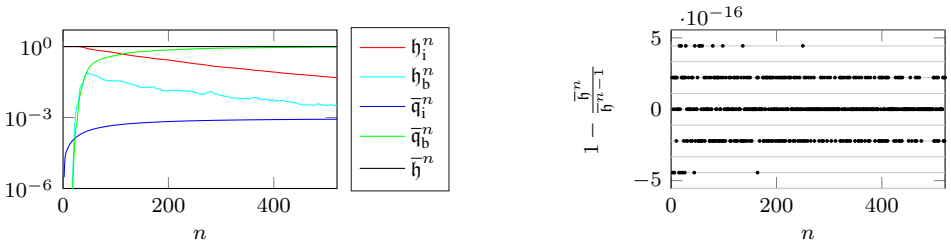
Figure 6.42: Illustration of fitted volumetric meshing (partially full) to fan-shaped domain $\overline{\Omega}$, using rhombic dodecahedral cells on the interior with inter-cell distance $h = 48\,\text{cm}$.



(a) $t = 2.5\,\mathrm{ms}$ (b) $t = 10\,\mathrm{ms}$

(c) $t = 20\,\mathrm{ms}$ (d) $t = 40\,\mathrm{ms}$

Figure 6.43: Snapshots of the time evolution of the acoustic field, at times as indicated, subject to an initial pulse. The sample rate here is 4 kHz and the FCC grid spacing is $h = 12\,\mathrm{cm}$.

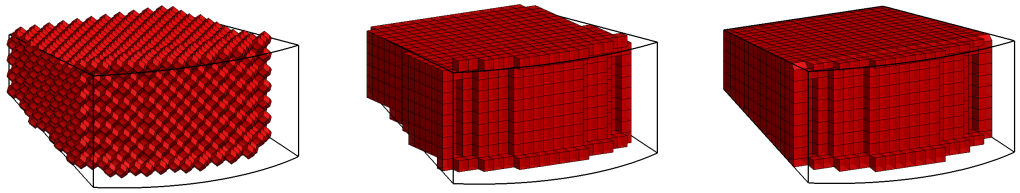


(a) Energetic quantities (normalised) (b) Relative variations in numerical energy

Figure 6.44: Plots demonstrating conservation of numerical energy in the presence of viscothermal losses and general impedance boundary losses, corresponding to the simulation in Fig. 6.43. The quantity $\overline{\mathfrak{h}}^n$ represents the total energy, which includes stored energy on the interior ($\mathfrak{h}_\mathrm{i}^n$) and at boundaries ($\mathfrak{h}_\mathrm{b}^n$), and accumulated losses over the interior ($\overline{\mathfrak{q}}_\mathrm{i}^n$) and over the boundary ($\overline{\mathfrak{q}}_\mathrm{b}^n$). The variations in the total energy are integer multiples of double precision machine epsilon, as denoted by grey lines.

6.8.2 A numerical example comparing staircase and fitted approximations

As a final numerical example, we compare numerical solutions obtained from staircase and fitted meshes of the fan-shaped domain illustrated in Fig. 6.42 with Cartesian and FCC cells. Staircase meshes for the FCC and cubic grids are illustrated in Fig. 6.45, along with a fitted tiling with interior cubic cells. Additionally for this test, we consider 3-D rotations of the domain with respect to the underlying Cartesian and FCC grids. These rotations are given by the rotation matrices $\mathbf{R}_1$ and $\mathbf{R}_2$, and analogous fitted and staircase meshes are derived for these rotated domains. Here we are testing how the numerical solution varies with respect to the choice of grid and with respect to rotations of the domain. Ideally, the numerical solution would be insensitive to such choices, and we expect this to be the case with fitted approximations.



(a) FCC, staircase (b) Cartesian, staircase (c) Cartesian, fitted

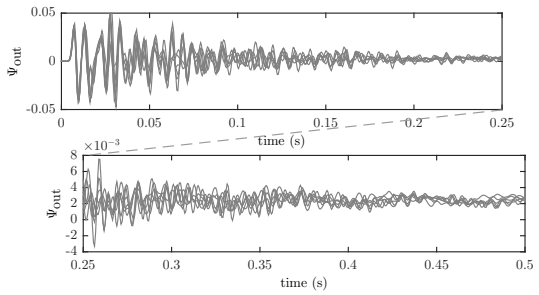
Figure 6.45: Illustrations of (partially full) staircase meshings of fan-shaped domain, using (a): Cartesian cells and (b): rhombic dodecahedral cells, respectively pertaining to the simple Cartesian and FCC finite difference schemes over the interiors. (c): A fitted mesh using Cartesian cells over the interior. For these figures, inter-cell distances on the interior are $h = 48\,\text{cm}$.

For these tests, the initial condition $\Psi(\mathbf{x}, 0)$ is a spatial Gaussian with variance $0.16\,\text{m}^2$ and shifted from the centre of the domain, and $\Psi_t(\mathbf{x}, 0) = 0$. Boundary conditions are set by the wall materials (using the general impedance model) as in the previous example, and $\eta = 0$ (no viscosity). For the Cartesian meshes, the inter-cell distance (on interior cells) is set to $h = \sqrt{3}ck$, where $1/k$ will be specified, and $h = 1.4ck$ is used for the FCC cases ($\lambda \approx 0.7$).

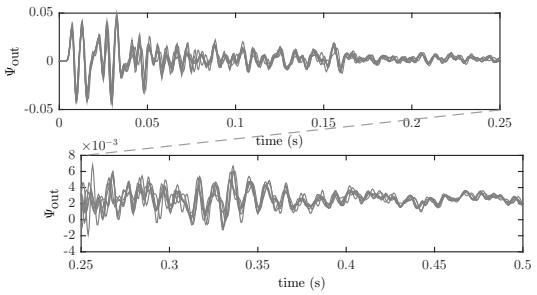
Fig. 6.46 shows the outputs of the simulations, read from the centre of the domain, and at the indicated sample rates. Superimposed within each of the plots in Fig. 6.46 are *six* outputs, pertaining to the three rotations (two rotated and one unrotated) and two grid types. For staircase approximations, the figure shows that there is a large degree of variability between the outputs obtained, and there is little improvement as the grid is refined (as the sample rate is increased). On the other hand, for fitted approximations, the numerical solutions demonstrate an invariance with respect to rotations and the grid types. Furthermore, and most importantly, the fitted approximations appear to converge to a unique solution as the grid is refined.

6.9 Summary

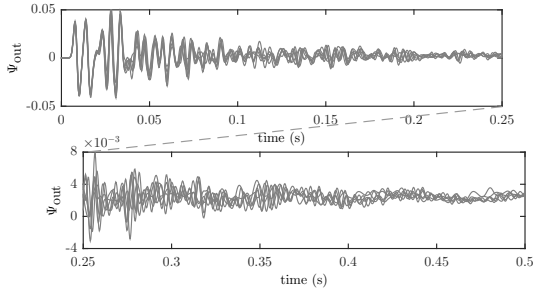
In this chapter we approached room acoustics simulations using finite volume methods operating over regular grids of cells throughout domain-interiors, reducing to known implicit/explicit finite difference schemes in Cartesian and non-Cartesian forms, and using specialised cells fitted to boundaries of non-trivial domains. We investigated the relationships between finite volume methods and various finite difference and digital waveguide mesh schemes. We also considered boundary conditions used for finite difference/DWM schemes under equivalent finite volume interpretations, revealing inconsistencies at non-trivial boundaries in some cases. Staircasing effects were investigated with various test cases,



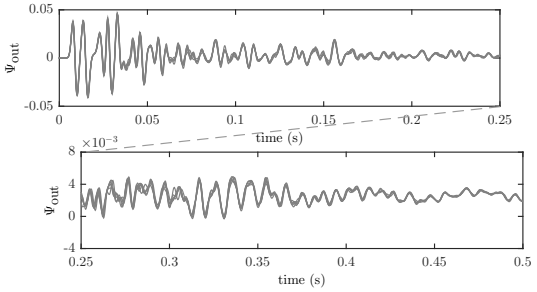
(a) Staircase approximations, $1/k = 2\,\text{kHz}$



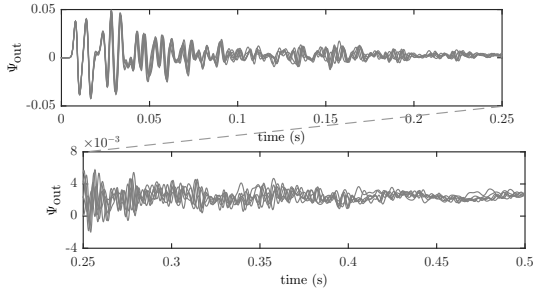
(b) Fitted approximations, $1/k = 2\,\text{kHz}$



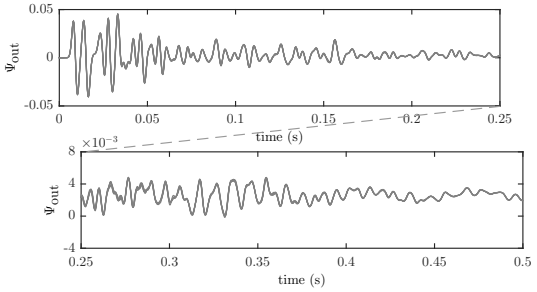
(c) Staircase approximations, $1/k = 4\,\text{kHz}$



(d) Fitted approximations, $1/k = 4\,\text{kHz}$



(e) Staircase approximations, $1/k = 8\,\text{kHz}$



(f) Fitted approximations, $1/k = 8\,\text{kHz}$

Figure 6.46: Plots of Ψ_{out} for staircase and fitted approximations at various sample rates, as indicated. The duration $t \in [0, 0.5\,\text{s}]$ is split in half in order to show detail. Superimposed within each plot are *six* outputs, pertaining to two grid types (cubic and rhombic dodecahedral interior cells) and three rotations.

and it was shown that, as expected, fitted tilings (fitted meshes) lead to significant improvements in the estimation of modal frequencies and in approximating underlying unique solutions. It was shown that decay times are well approximated with fitted meshes, while staircasing effects can lead to gross overestimations of decay times that does not disappear with grid refinement (i.e., inconsistencies in absorptive boundary modelling). Frequency-dependent losses over the interior (viscothermal) and boundaries were presented and tested numerically, using rhombic dodecahedral and cubic cells in 3-D for meshings of non-trivial domains, under various rotations in 3-D, and with general impedance boundaries approximating real-world wall materials.

Chapter 7

Conclusions and final remarks

7.1 Summary

The goals of this thesis have been to develop accurate and stable numerical modelling of room acoustics over non-trivial domains with frequency-dependent losses over the interior and boundaries, while maintaining computational efficiency. What we have presented in this thesis, using time-domain finite difference and finite volume methods, has in some part achieved these goals.

As a brief summary of this thesis, in Chapter 2 we started with detailed PDE models of room acoustics, and in Chapter 3 we presented a discretisation of those equations with the simplest possible 3-D scheme. In Chapter 4 we considered the simulation of free-space wave propagation with finite difference schemes, with the goal of tackling the computational costs associated with mitigating dispersion errors, leading to various new and old schemes (and comparisons thereof) with high-order accuracy and/or optimised for wideband accuracy. In Chapter 5 we investigated the accuracy and stability of two types of finite difference boundary conditions, as well some of the problems that can arise from certain approaches to boundary modelling that are not accompanied by appropriate stability analyses. In Chapter 6 we used finite volume methods to tackle the problems of staircasing and to ensure stability at boundaries using energy methods, resulting in a full discretisation of the PDE models of room acoustics from Chapter 2, in an accurate and consistent manner.

In the final numerical example of Chapter 6, finite volume simulations over a non-trivial 3-D domain with general frequency-dependent impedance boundary conditions modelling three different types of real-world wall materials were presented. These simulations were able to reliably solve for a unique solution to this boundary value problem, and six different simulations (using two different grid types and under three rotations of the domain in 3-D) gave the same result, and this was accomplished while conserving numerical energy and stability to machine accuracy. This ability to reliably extract the solution from a complicated wave-based model of room acoustics is something that has not been previously achieved with other time-domain wave-based methods in the room acoustics literature (to our knowledge). The significance of which cannot be overstated; at this point we can start to close the gap between the mathematical models of wave-based room acoustics and the real-world measurements, by reliably eliminating the numerical approximation errors that exist between these two realms.

7.2 Problems to be solved

There are various problems that could not be fully solved in this thesis. Two areas which seem to be most important are discussed below for potential future work.

On high-order accurate and optimised (wide-stencil) schemes

In Chapter 4 we investigated schemes that reach out to more than the usual nearest neighbours on a Cartesian grid, with high orders of accuracy in space and/or time. Aside from the simplest hexagonal and simplest FCC schemes, most of these schemes have yet to be supplemented with boundary conditions for use in room acoustics models (in a stable and accurate manner), and this constitutes an important challenge. If finite volume methods are to be used as generalisations of wide-stencil finite difference schemes, equivalent volumetric interpretations for the various stencils involved will be necessary—this will require some creativity. It is worth recalling that even the simple nine-point stencil on the BCC grid (the "octahedral scheme") did not have a clear finite volume interpretation, since the seemingly-appropriate volumetric cell, the regular octahedron, does not tile space. Numerical stability for high-order accurate schemes will also be a challenge, since, apart from the implicit schemes presented in Section 4.7, high-order accuracy (or even optimised accuracy) generally requires coefficients that alternate in sign (at least for the second-order wave equation), departing from concretely passive formulations that readily admit simple energy inequalities. There is certainly work in this area from which we can take inspiration (e.g., [55]), but we must also extend such energy estimates to boundaries. Penalty-techniques for imposing boundary conditions weakly [261] may be of use in this regard.

On meshing for room acoustics simulations

Another important topic that played a central role in Chapter 6, but which could not be discussed in detail, is the meshing step required for the proposed uses of finite volume methods. Outside of the room acoustics community, meshing is certainly a well-researched topic, but the meshing constraints for room acoustics simulation may be significantly different from those that appear in other domains. In particular, to keep memory costs to a minimum, and to fully exploit the parallel nature of such methods, room acoustics problems should benefit from meshes employing "regular cells" (including rhombic dodecahedral cells) over the homogeneous interior (corresponding regular finite difference updates), and specialised cells would, ideally, only appear near boundaries of the domain. This may not be a normal mode of operation for conventional meshing algorithms and software packages, which typically make use of tetrahedral or hexahedral elements that vary continuously over the entire 3-D domain [270] rather than only locally at the boundary. Furthermore, rhombic dodecahedral cells, which can lead to favourable numerical dispersion properties, are relatively uncommon to finite volume and finite element methods, and to the associated meshing algorithms and software packages.[1] When meshing for room acoustics problems, one must also keep in mind the constraints of numerical stability and the effect it can have on numerical dispersion over the interior. More specifically, the use of specialised boundary cells should, ideally, not constrain the time-step on the interior beyond the known von Neumann conditions for the associated finite difference schemes, such that they can operate as efficiently as possible (in terms of minimising numerical dispersion) over the interior.

[1] In fact, this work constitutes one of the first uses of rhombic dodecahedral cells for finite volume modelling of non-trivial domains—not only in the field of room acoustics.

7.3 Other aspects of room acoustics modelling

There are many other aspects of room acoustics modelling which could not be treated in this thesis. It is worth mentioning some of these details, as they are ripe for integration into this general finite volume framework, and themselves may benefit from detailed physical modelling.

One obvious topic that has yet to be properly explored with finite volume methods is the simulation of binaural audio, using head and torso models within wave-based simulations. Preliminary work has been conducted with various techniques (e.g., finite difference, finite element, boundary element, and domain-decomposition Fourier methods [123, 172, 305, 239, 169, 197]), but so far not with finite volume methods (involving locally irregular tilings/meshes). Furthermore, it would be desirable to keep such cells to a reasonable size such that they can be integrated into large-scale room acoustic simulations conducted at audio rates. This is an exciting prospect, yet it is a challenging problem that will require more insight into the aforementioned meshing problem.

Another area of investigation is the modelling of realistic sound sources and receivers. Beyond simple point sources, it would be useful to apply finite volume methods to model sound sources and receivers that occupy space (larger than a single cell), have directionality, and possibly move through space throughout the simulation.

Appendix A

Relationship between various impedance boundary conditions

The purpose of this appendix, which relates to a discussion on pg. 82, is to illustrate the relationship between discrete impedance boundary conditions proposed in [133] for the simple 2-D or 3-D Cartesian scheme, and those given in (3.128) (adapted from [26]). We will begin by summarising the discretisation chosen in [133] and then we will show that it is similar, and sometimes equivalent, to the form in (3.128).

The boundary model proposed in [133] for a boundary wall with outward normal vector $\hat{\mathbf{n}} = \hat{\mathbf{e}}_x$ is written, under a notation similar to that used in [133], as follows:

$$p_t = -c\left(\frac{M}{\rho c}p_{xt} + \frac{R}{\rho c}p_x + \frac{K}{\rho c}\int_{-\infty}^{t} p_x \mathrm{d}\tau\right) \tag{A.1}$$

For the discretisation, we use the grid function $\underline{p}^n_{\boldsymbol{i}} \approxeq p(\boldsymbol{i}h, nk)$ and let the $\boldsymbol{i}h$ be a point on (or near) the boundary of interest. Both 2-D and 3-D cases are treated here, since we can have $\boldsymbol{i} \in \mathbb{Z}^2$ or $\boldsymbol{i} \in \mathbb{Z}^3$. The essentials of the discretisation chosen by Kowalczyk and van Walstijn are described by [133, Eqs. (30)–(32) and (40)–(43)]. The bilinear transform (trapezoidal integration/differentiation) is applied to the boundary condition for temporal derivatives and integrals, and the spatial derivatives are discretised with centered operators $\delta_{x\cdot}$ (defined by (3.41d)). The resulting discrete boundary condition is then [133, Eq. (30)]:

$$\underline{p}^{n+1}_{\boldsymbol{i}} - \underline{p}^{n-1}_{\boldsymbol{i}} = \lambda\left(-a_M S^n_M + a_R(\underline{p}^n_{\boldsymbol{i}-\hat{\mathbf{e}}_x} - \underline{p}^n_{\boldsymbol{i}+\hat{\mathbf{e}}_x}) + a_K S^n_K\right) \tag{A.2}$$

where

$$a_M = \frac{2}{k}\frac{M}{\rho c}, \quad a_R = \frac{R}{\rho c}, \quad a_K = \frac{k}{2}\frac{K}{\rho c}, \quad a = a_M + a_R + a_S \tag{A.3}$$

and where S^n_K, S^n_M are auxiliary variables that result from the application of the bilinear transform to treat temporal derivatives and integration. They are expressed by the recursions [133, Eqs. (31), (32), (41), and (42)]:

$$S^n_K = (\underline{p}^n_{\boldsymbol{i}-\hat{\mathbf{e}}_x} - \underline{p}^n_{\boldsymbol{i}+\hat{\mathbf{e}}_x}) + (\underline{p}^{n-1}_{\boldsymbol{i}-\hat{\mathbf{e}}_x} - \underline{p}^{n-1}_{\boldsymbol{i}+\hat{\mathbf{e}}_x}) + S^{n-1}_K \tag{A.4a}$$

$$S^n_M = (\underline{p}^n_{\boldsymbol{i}+\hat{\mathbf{e}}_x} - \underline{p}^n_{\boldsymbol{i}-\hat{\mathbf{e}}_x}) - (\underline{p}^{n-1}_{\boldsymbol{i}+\hat{\mathbf{e}}_x} - \underline{p}^{n-1}_{\boldsymbol{i}-\hat{\mathbf{e}}_x}) - S^{n-1}_M \tag{A.4b}$$

Rearranging these equations leads to an explicit update for the ghost point $\underline{p}^n_{\boldsymbol{i}+\hat{\mathbf{e}}_x}$ [133, Eqs. (33) and

(43)]:

$$\underline{p}^n_{\boldsymbol{i}+\hat{\mathbf{e}}_x} = \underline{p}^n_{\boldsymbol{i}-\hat{\mathbf{e}}_x} + \frac{1}{\lambda a}(\underline{p}^{n-1}_{\boldsymbol{i}} - \underline{p}^{n+1}_{\boldsymbol{i}}) + \frac{a_K - a_M}{a}(\underline{p}^{n-1}_{\boldsymbol{i}-\hat{\mathbf{e}}_x} - \underline{p}^{n-1}_{\boldsymbol{i}+\hat{\mathbf{e}}_x}) + \frac{a_K}{a}S^{n-1}_K + \frac{a_M}{a}S^{n-1}_M \quad \text{(A.5)}$$

which finally leads to an explicit update for the boundary node $\underline{p}^{n+1}_{\boldsymbol{i}}$ that involves the ghost point and auxiliary variables at time $n-1$ [133, Eqs. (34) and (40)], but we leave these out for brevity.

In summary, one proceeds by updating the boundary node with [133, Eqs. (34) and (40)], followed by the ghost point with (A.5), followed by the auxiliary variables with (A.4a) and (A.4b). At edges and corners there is more than one ghost point to update, and each ghost point carries its own pair of auxiliary variables S^n_M and S^n_K [133]. Thus, the boundary update requires three additional states per ghost point, each having a separate update. Updating boundary nodes in the manner proposed in [133] requires the storage of three additional states per ghost point, but only two are updated at each time-step.

At this point we will demonstrate the relationship the above conditions have with (3.127). First, we note that (A.1) is equivalent to:

$$p_t = -c\left(\frac{M}{\rho c}p_{xt} + \frac{R}{\rho c}p_x + \frac{K}{\rho c}g_x\right), \quad g_t = p \quad \text{(A.6)}$$

which is (3.127) for $\hat{\mathbf{n}} = \hat{\mathbf{e}}_x$ after making the change of variable $p \to \Psi$, and $(D, E, F) = Y_0(M, R, K)$, where $Y_0 = \frac{1}{\rho c}$. Then we can rewrite (A.2) in an equivalent form:

$$\delta_{t\cdot}\underline{p}^n_{\boldsymbol{i}} = -cY_0\left(M\overset{\star}{S}{}^n_M + R\underline{p}^n_{x+,\boldsymbol{i}} + K\overset{\star}{S}{}^n_K\right) \quad \text{(A.7)}$$

where

$$\overset{\star}{S}{}^n_M = \frac{1}{hk}S^n_M, \quad \overset{\star}{S}{}^n_K = -\frac{k}{4h}S^n_K \quad \text{(A.8)}$$

and

$$\underline{p}^n_{x+,\boldsymbol{i}} = \delta_{x\cdot}\underline{p}^n_{\boldsymbol{i}} \quad \text{(A.9)}$$

Then we can rewrite (A.4a) and (A.4b) as:

$$\delta_{t-}\overset{\star}{S}{}^n_K = \mu_{t-}\underline{p}^n_{x+,\boldsymbol{i}}, \quad \mu_{t-}\overset{\star}{S}{}^n_M = \delta_{t-}\underline{p}^n_{x+,\boldsymbol{i}} \quad \text{(A.10)}$$

which represent trapezoidal integration and differentiation, respectively. If we then define the auxiliary variables:

$$\mu_{t-}\underline{\hat{p}}^{n+\frac{1}{2}}_{x+,\boldsymbol{i}} = \delta_{t-}\underline{g}^{n+\frac{1}{2}}_{x+,\boldsymbol{i}} = \underline{p}^n_{x+,\boldsymbol{i}} \quad \text{(A.11)}$$

Then we have:

$$\overset{\star}{S}{}^n_K = \mu_{t-}\underline{g}^{n+\frac{1}{2}}_{x+,\boldsymbol{i}}, \quad \overset{\star}{S}{}^n_M = \delta_{t-}\underline{\hat{p}}^{n+\frac{1}{2}}_{x+,\boldsymbol{i}} \quad \text{(A.12)}$$

and the boundary condition can be written as:

$$\delta_{t\cdot}\underline{p}^n_{\boldsymbol{i}} = -cY_0\left(M\delta_{t-}\underline{\hat{p}}^{n+\frac{1}{2}}_{x+,\boldsymbol{i}} + R\mu_{t-}\underline{\hat{p}}^{n+\frac{1}{2}}_{x+,\boldsymbol{i}} + K\mu_{t-}\underline{g}^{n+\frac{1}{2}}_{x+,\boldsymbol{i}}\right) \quad \text{(A.13)}$$

Finally, using the identities $\delta_{t-} = e_{t-}\delta_{t+}$ and $\mu_{t-} = e_{t-}\mu_{t+}$ we can rewrite this as:

$$\delta_{t\cdot}\underline{p}^n_{\boldsymbol{i}} = -c\left(D\delta_{t+}\underline{\hat{p}}^{n-\frac{1}{2}}_{x+,\boldsymbol{i}} + E\mu_{t+}\underline{\hat{p}}^{n-\frac{1}{2}}_{x+,\boldsymbol{i}} + F\mu_{t+}\underline{g}^{n-\frac{1}{2}}_{x+,\boldsymbol{i}}\right) \quad \text{(A.14)}$$

which is now in an equivalent form that is very similar to (3.128).

Comparing the above with (3.128), we note that the choice of the dependent variable of interest is different ($\underline{p}_{\boldsymbol{i}}^n$ rather than $\underline{\Psi}_{\boldsymbol{i}}^n$), but this is simply a question of replacing $\underline{p}_{\boldsymbol{i}}^n$ with $\underline{\Psi}_{\boldsymbol{i}}^n$. The main difference between the two conditions of interest is that (3.128) uses non-centered spatial differences, i.e., we would have $\underline{p}_{-x+,\boldsymbol{i}}^n = \delta_{x+}\underline{p}_{\boldsymbol{i}}^n$ rather than (A.9). This has important consequences for numerical stability over non-trivial domains (discussed at length in Section 5.4). Also, we should note that spatially-varying impedance parameters were not considered in [133], but the parameters D, E, F are permitted to vary along the boundary surface with (3.128).

Bibliography

[1] M.-L. Aird, "Musical instrument modelling using digital waveguides," Ph.D. thesis, University of Bath, 2002.

[2] R. M. Alford, K. R. Kelly, and D. M. Boore, "Accuracy of finite-difference modeling of the acoustic wave equation," *Geophysics*, vol. 39, no. 6, pp. 834–842, 1974.

[3] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *Journal of the Acoustical Society of America*, vol. 65, p. 943, 1979.

[4] Z. Alterman and F. C. Karal, "Propagation of elastic waves in layered media by finite difference methods," *Bulletin of the Seismological Society of America*, vol. 58, no. 1, pp. 367–398, 1968.

[5] W. F. Ames, *Numerical Methods for Partial Differential Equations*, 2nd ed. Academic Press, 1977.

[6] W. N. Anderson Jr. and T. D. Morley, "Eigenvalues of the Laplacian of a graph," *Linear and Multilinear Algebra*, vol. 18, no. 2, pp. 141–145, 1985.

[7] L. Anné, P. Joly, and Q. H. Tran, "Construction and analysis of higher order finite difference schemes for the 1D wave equation," *Computational Geosciences*, vol. 4, no. 3, pp. 207–249, 2000.

[8] L. Anné, Q. H. Tran, and W. W. Symes, "Dispersion and cost analysis of some finite difference schemes in one-parameter acoustic wave modeling," *Computational Geosciences*, vol. 1, no. 1, pp. 1–33, 1997.

[9] A. Arakawa, "Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I," *Journal of Computational Physics*, vol. 1, no. 1, pp. 119–143, 1966.

[10] M. Aretz, *Combined Wave and Ray Based Room Acoustic Simulations of Small Rooms*. Logos Verlag Berlin GmbH, 2012, vol. 12.

[11] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Saunders College, Philadelphia, 1976.

[12] R. Bacon and J. Bowsher, "A discrete model of a struck string," *Acta Acustica united with Acustica*, vol. 41, no. 1, pp. 21–27, 1978.

[13] M. R. Bai, "Study of acoustic resonance in enclosures using eigenanalysis based on boundary element methods," *Journal of the Acoustical Society of America*, vol. 91, no. 5, pp. 2529–2538, 1992.

[14] B. Baliga and S. Patankar, "A new finite-element formulation for convection-diffusion problems," *Numerical Heat Transfer*, vol. 3, no. 4, pp. 393–409, 1980.

[15] G. Barton, *Elements of Green's Functions and Propagation: Potentials, Diffusion, and Waves*. Oxford University Press, 1989.

[16] H. Bass, L. Sutherland, A. Zuckerwar, D. Blackstock, and D. Hester, "Atmospheric absorption of sound: Further developments," *Journal of the Acoustical Society of America*, vol. 97, no. 1, pp. 680–683, 1995.

[17] E. Baysal, D. D. Kosloff, and J. W. Sherwood, "Reverse time migration," *Geophysics*, vol. 48, no. 11, pp. 1514–1524, 1983.

[18] J. Bensa, S. Bilbao, R. Kronland-Martinet, and J. O. Smith III, "The simulation of piano string vibration: From physical models to finite difference schemes and digital waveguides," *Journal of the Acoustical Society of America*, vol. 114, no. 2, pp. 1095–1107, 2003.

[19] N. Bertin, S. Kitić, and R. Gribonval, "Joint estimation of sound source location and boundary impedance with physics-driven cosparse regularization," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Shanghai, China, 2016, pp. 6340–6344.

[20] Z. Bi, K. Wu, C. Wu, and J. Litva, "A new finite-difference time-domain algorithm for solving Maxwell's equations," *IEEE Microwave and Guided Wave Letters*, vol. 1, no. 12, pp. 382–384, 1991.

[21] S. Bilbao, "Wave and scattering methods for the numerical integration of partial differential equations," Ph.D. thesis, Stanford University, 2001.

[22] ——, "Parameterized families of finite difference schemes for the wave equation," *Numerical Methods for Partial Differential Equations*, vol. 20, no. 3, pp. 463–480, 2004.

[23] ——, *Wave and Scattering Methods for Numerical Simulation*. Wiley, 2004.

[24] ——, *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics*. Wiley, 2009.

[25] ——, "Optimized FDTD schemes for 3-D acoustic wave propagation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 5, pp. 1658–1663, 2012.

[26] ——, "Modeling of complex geometries and boundary conditions in finite difference/finite volume time domain room acoustics simulation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 7, pp. 1524–1533, Jul. 2013.

[27] S. Bilbao and B. Hamilton, "Construction and optimization techniques for high order schemes for the two-dimensional wave equation," in *Proceedings of the International Congress on Acoustics (ICA)*, Montréal, Canada, 2013.

[28] ——, "Finite volume modeling of viscothermal losses and frequency-dependent boundaries in room acoustics simulations," in *AES 60th International Conference*, Leuven, Belgium, Feb. 2016, to appear.

[29] ——, "Wave-based room acoustics simulation: Explicit/implicit finite volume modeling of viscothermal losses and frequency-dependent boundaries," *Journal of the Audio Engineering Society*, 2017, accepted for publication.

[30] S. Bilbao, B. Hamilton, J. Botts, and L. Savioja, "Finite volume time domain room acoustics simulation under general impedance boundary conditions," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 1, pp. 161–173, 2016.

[31] S. Bilbao, B. Hamilton, R. Harrison, and A. Torin, *Springer Handbook of Systematic Musicology*. Springer-Verlag, 2017, ch. Finite Difference Schemes in Musical Acoustics: A Tutorial, (to appear).

[32] S. Bilbao and J. O. Smith III, "Finite difference schemes and digital waveguide networks for the wave equation: Stability, passivity, and numerical dispersion," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 3, pp. 255–266, 2003.

[33] S. Bilbao and C. J. Webb, "Timpani drum synthesis in 3D on GPGPUs," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, York, UK, 2012.

[34] G. Boole and J. F. Moulton, *A Treatise on the Calculus of Finite Differences*. MacMillan and Company, 1880.

[35] J. Borish, "Extension of the image model to arbitrary polyhedra," *Journal of the Acoustical Society of America*, vol. 75, no. 6, pp. 1827–1836, 1984.

[36] N. Borrel-Jensen, "Real-time auralisation of the lower frequency sound field using numerical methods on the GPU," M.Sc. thesis, RWTH Aachen University and University of Copenhagen, 2012.

[37] D. Botteldooren, "Acoustical finite-difference time-domain simulation in a quasi-Cartesian grid," *Journal of the Acoustical Society of America*, vol. 95, pp. 2313–2319, 1994.

[38] ——, "Finite-difference time-domain simulation of low-frequency room acoustic problems," *Journal of the Acoustical Society of America*, vol. 98, pp. 3302–3308, 1995.

[39] D. Botteldooren and H. Devos, "Non-uniform grids for acoustical FDTD-simulations," *International Congress on Air- and Structure-borne Sound and Vibration*, pp. 849–856, 1994.

[40] D. Botteldooren, "Feasibility study of active control of noise perceived by operators of large agricultural machines." *Noise Control Engineering Journal*, vol. 40, no. 3, pp. 221–229, 1993.

[41] ——, "Numerical model for moderately nonlinear sound propagation in three-dimensional structures," *Journal of the Acoustical Society of America*, vol. 100, no. 3, pp. 1357–1367, 1996.

[42] ——, "Vorticity and entropy boundary conditions for acoustical finite-difference time-domain simulations," *Journal of the Acoustical Society of America*, vol. 102, no. 1, pp. 170–178, 1997.

[43] J. Botts and L. Savioja, "Integrating finite difference schemes for scalar and vector wave equations," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, 2013, pp. 171–175.

[44] ——, "Spectral and pseudospectral properties of finite difference models used in audio and room acoustics," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 9, pp. 1403–1412, Sept 2014.

[45] ——, "Extension of a spectral time-stepping domain decomposition method for dispersive and dissipative wave propagation," *Journal of the Acoustical Society of America*, vol. 137, no. 4, pp. EL267–EL273, 2015.

[46] M. Bruneau and T. Scelo, *Fundamentals of Acoustics*. ISTE, 2006.

[47] M. J. Buckingham, "Causality, Stokes' wave equation, and acoustic pulse propagation in a viscous fluid," *Physical Review E*, vol. 72, no. 2, p. 026610, 2005.

[48] G. R. Campos and D. M. Howard, "On the computational efficiency of different waveguide mesh topologies for room acoustic simulation," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 5, pp. 1063–1072, 2005.

[49] A. Chaigne and A. Askenfelt, "Numerical simulations of piano strings. I. A physical model for a struck string using finite difference methods," *Journal of the Acoustical Society of America*, vol. 95, no. 2, pp. 1112–1118, 1994.

[50] J. G. Charney, R. Fjörtoft, and J. von Neumann, "Numerical integration of the barotropic vorticity equation," *Tellus*, vol. 2, no. 4, pp. 237–254, 1950.

[51] O. Chiba, T. Kashiwa, H. Shimoda, S. Kagami, and I. Fukai, "Analysis of sound fields in three dimensional space by the time-dependent finite-difference method based on the leap frog algorithm," *Journal Acoustical Society of Japan (J)*, vol. 49, pp. 551–562, 1993.

[52] M. Ciment and S. H. Leventhal, "Higher order compact implicit schemes for the wave equation," *Mathematics of Computation*, vol. 29, no. 132, pp. 985–994, 1975.

[53] G. Cohen, *Higher-order Numerical Methods for Transient Wave Equations*. Springer-Verlag, 2002.

[54] G. Cohen and P. Joly, "Description and analysis of schemes, fourth-order in time and space, for 1-D and 2-D acoustic equations," in *Society of Exploration Geophysicists Annual Meeting*, 1987.

[55] ——, "Construction analysis of fourth-order finite difference schemes for the acoustic wave equation in nonhomogeneous media," *SIAM Journal on Numerical Analysis*, vol. 33, no. 4, pp. 1266–1302, 1996.

[56] L. Collatz, *The Numerical Treatment of Differential Equations*, 3rd ed. Berlin: Springer, 1960.

[57] ——, *The Numerical Treatment of Differential Equations*, 3rd ed. Berlin: Springer, 1966, 2nd printing.

[58] J. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. Springer-Verlag, 1991.

[59] J. H. Conway, H. Burgiel, and C. Goodman-Strauss, *The Symmetry of Things*. A K Peters, 2008.

[60] R. Courant, K. Friedrichs, and H. Lewy, "Über die partiellen differenzengleichungen der mathematischen physik," *Mathematische Annalen*, vol. 100, no. 1, pp. 32–74, 1928.

[61] R. Courant and D. Hilbert, *Methods of Mathematical Physics*. Interscience, 1962, vol. 2.

[62] R. Courant, "Variational methods for the solution of problems of equilibrium and vibrations," *Bulletin of the American Mathematical Society*, vol. 49, no. 1, pp. 1–23, 1943.

[63] T. J. Cox, B.-I. L. Dalenbäck, P. D'Antonio, J.-J. Embrechts, J. Y. Jeon, E. Mommertz, and M. Vorländer, "A tutorial on scattering and diffusion coefficients for room acoustic surfaces," *Acta Acustica united with Acustica*, vol. 92, no. 1, pp. 1–15, 2006.

[64] A. Craggs, "A finite element method for the free vibration of air in ducts and rooms with absorbing walls," *Journal of Sound and Vibration*, vol. 173, no. 4, pp. 568–576, 1994.

[65] M. S. Cramer, "Numerical estimates for the bulk viscosity of ideal gases," *Physics of Fluids*, vol. 24, no. 6, p. 066102, 2012.

[66] M. A. Dablain, "The application of high-order differencing to the scalar wave equation," *Geophysics*, vol. 51, no. 1, pp. 54–66, 1986.

[67] C. de Groot-Hedlin, "Finite-difference time-domain synthesis of infrasound propagation through an absorbing atmosphere," *Journal of the Acoustical Society of America*, vol. 124, no. 3, pp. 1430–1441, 2008.

[68] D. G. Duffy, *Greens Functions with Applications*. CRC Press, 2015.

[69] B. Engquist and A. Majda, "Absorbing boundary conditions for numerical simulation of waves," *Proceedings of the National Academy of Sciences*, vol. 74, no. 5, pp. 1765–1766, 1977.

[70] J. T. Etgen, "Accurate wave equation modeling," *SEP-60: Stanford Exploration Project*, vol. 131, p. 148, 1989.

[71] J. T. Etgen and M. J. O'Brien, "Computational methods for large-scale 3D acoustic finite-difference modeling: A tutorial," *Geophysics*, vol. 72, no. 5, pp. SM223–SM230, 2007.

[72] V. Etienne, T. Tonellot, P. Thierry, V. Berthoumieux, and C. Andreolli, "Optimization of the seismic modeling with the time-domain finite-difference method," in *SEG Annual Meeting*, 2014.

[73] L. Euler, "De la propagation du son," *Histoire de l'Académie royale des sciences et belles lettres pour l'année 1759*, vol. 16, pp. 185–209, 1766.

[74] ——, "Supplément aux recherches sur la propagation du son," *Histoire de l'Académie royale des sciences et belles lettres pour l'année 1759*, vol. 16, pp. 210–240, 1766.

[75] L. C. Evans, *Partial Differential Equations*. American Mathematical Society, 1998.

[76] J. C. Fabero, A. Bautista, and L. Casasús, "An explicit finite differences scheme over hexagonal tessellation," *Applied Mathematics Letters*, vol. 14, no. 5, pp. 593–598, 2001.

[77] G. Fairweather and A. Mitchell, "A high accuracy alternating direction method for the wave equation," *IMA Journal of Applied Mathematics*, vol. 1, no. 4, pp. 309–316, 1965.

[78] G. Fischer, "On a finite difference scheme for solving the non-linear primitive equations for a barotropic fluid with application to the boundary current problem," *Tellus*, vol. 17, no. 4, pp. 405–412, 1965.

[79] F. Fontana, "Physics-based models for the acoustic representation of space in virtual environments," Ph.D. thesis, University of Verona, 2003.

[80] F. Fontana and D. Rocchesso, "A new formulation of the 2D-waveguide mesh for percussion instruments," in *Proceedings of the XI Colloqium on Musical Informatics*, Bologna, Italy, Nov. 1995.

[81] ——, "Signal-theoretic characterization of waveguide mesh geometries for models of two-dimensional wave propagation in elastic media," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 2, pp. 152–161, 2001.

[82] F. Fontana, D. Rocchesso, and E. Apollonio, "Using the waveguide mesh in modelling 3d resonators," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Verona, Italy, Dec. 2000, pp. 229–232.

[83] B. Fornberg, "On a Fourier method for the integration of hyperbolic equations," *SIAM Journal on Numerical Analysis*, vol. 12, no. 4, pp. 509–528, 1975.

[84] ——, "Generation of finite difference formulas on arbitrarily spaced grids," *Mathematics of computation*, vol. 51, no. 184, pp. 699–706, 1988.

[85] ——, *A Practical Guide to Pseudospectral Methods*. Cambridge University Press, 1998, vol. 1.

[86] G. E. Forsythe and W. R. Wasow, *Finite-difference Methods for Partial Differential Equations*. New York: Wiley, 1960.

[87] C. Godsil and G. F. Royle, *Algebraic Graph Theory*. Springer Science & Business Media, 2013, vol. 207.

[88] J. E. Goodman, J. O'Rourke, and K. H. Rosen, *Handbook of Discrete and Computational Geometry*, 2nd ed. CRC Press, 2004.

[89] A. Gourlay and A. R. Mitchell, "A classification of split difference methods for hyperbolic equations in several space dimensions," *SIAM Journal on Numerical Analysis*, vol. 6, no. 1, pp. 62–71, 1969.

[90] B. Gustafsson and P. Wahlund, "Time compact high order difference methods for wave propagation, 2D," *Journal of Scientific Computing*, vol. 25, no. 1-2, pp. 195–211, 2005.

[91] B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods*. Wiley, 1995, vol. 67.

[92] H. Hachabiboglu, B. Gunel, and A. Kondoz, "On the accuracy of first-order numerical derivatives in multidimensional digital waveguide mesh topologies," *IEEE Signal Processing Letters*, vol. 15, pp. 9–12, 2008.

[93] J. Häggblad and B. Engquist, "Consistent modeling of boundaries in acoustic finite-difference time-domain simulations," *Journal of the Acoustical Society of America*, vol. 132, no. 3, pp. 1303–1310, 2012.

[94] B. Hamilton, "Sampling on a diamond grid and the tetrahedral digital waveguide mesh," *IEEE Signal Processing Letters*, vol. 20, no. 10, pp. 925–928, Oct. 2013.

[95] ——, "Finite volume perspectives on finite difference schemes and boundary formulations for wave simulation," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Erlangen, Germany, Sep. 2014.

[96] B. Hamilton and S. Bilbao, "Fourth-order and optimised finite difference schemes for the 2-D wave equation," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Maynooth, Ireland, 2013.

[97] ——, "Hexagonal vs. rectilinear grids for explicit finite difference schemes for the two-dimensional wave equation," in *Proceedings of the International Congress on Acoustics (ICA)*, Montréal, Canada, 2013.

[98] ——, "On finite difference schemes for the 3-D wave equation using non-Cartesian grids," in *Proceedings of the Sound and Music Computing (SMC) Conference*, Stockholm, Sweden, 2013, pp. 592–599.

[99] B. Hamilton, S. Bilbao, and C. J. Webb, "Improved finite difference schemes for a 3-D viscothermal wave equation on a GPU," in *Proceedings of Forum Acusticum*, Krakow, Poland, 2014.

[100] ——, "Revisiting implicit finite difference schemes for 3-D room acoustics simulations on GPU," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Erlangen, Germany, 2014.

[101] B. Hamilton and C. J. Webb, "Room acoustics modelling using GPU-accelerated finite difference and finite volume methods on a face-centered cubic grid," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Maynooth, Ireland, Sep. 2013, pp. 336–343.

[102] B. Hamilton, C. J. Webb, N. D. Fletcher, and S. Bilbao, "Finite difference room acoustics simulation with general impedance boundaries and viscothermal losses in air: Parallel implementation on multiple GPUs," in *Proceedings of the International Symposium on Music and Room Acoustics (ISMRA)*, Buenos Aires, Argentina, Sep. 2016.

[103] B. Hamilton, C. J. Webb, A. Gray, and S. Bilbao, "Large stencil operations for GPU-based 3-D acoustics simulations," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Trondheim, Norway, Dec. 2015.

[104] C. Hammond, *The Basics of Crystallography and Diffraction*. Oxford University Press, USA, 2009.

[105] J. A. Hargreaves, "Time domain boundary element method for room acoustics," Ph.D. dissertation, University of Salford, 2007.

[106] J. A. Hargreaves and Y. W. Lam, "An energy interpretation of the Kirchhoff-Helmholtz boundary integral equation and its application to sound field synthesis," *Acta Acustica united with Acustica*, vol. 100, no. 5, pp. 912–920, 2014.

[107] B. Heinrich, "Boundary value problems and irregular networks," in *Finite Difference Methods on Irregular Networks*. Springer, 1987, pp. 17–39.

[108] J. S. Hesthaven, S. Gottlieb, and D. Gottlieb, *Spectral Methods for Time-dependent Problems*. Cambridge University Press, 2007, vol. 21.

[109] R. L. Higdon, "Numerical absorbing boundary conditions for the wave equation," *Mathematics of Computation*, vol. 49, no. 179, pp. 65–90, 1987.

[110] L. Hiller and P. Ruiz, "Synthesizing musical sounds by solving the wave equation for vibrating objects: Parts I and II," *Journal of the Audio Engineering Society*, vol. 19, no. 7, pp. 542–551, 1971.

[111] O. Holberg, "Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena," *Geophysical Prospecting*, vol. 35, no. 6, pp. 629–655, 1987.

[112] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[113] M. Hornikx, W. De Roeck, and W. Desmet, "A multi-domain Fourier pseudospectral time-domain method for the linearized Euler equations," *Journal of Computational Physics*, vol. 231, no. 14, pp. 4759–4774, 2012.

[114] M. Hornikx and D. Steeghs, "Acoustic propagation modeled by the curvilinear Fourier pseudospectral time-domain method," in *Proceedings of the International Congress on Acoustics (ICA)*, Montréal, Canada, 2013.

[115] J. Huopaniemi, L. Savioja, and M. Karjalainen, "Modeling of reflections and air absorption in acoustical spaces a digital filter design approach," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. IEEE, 1997, pp. 4–pp.

[116] F. Ihlenburg, *Finite Element Analysis of Acoustic Scattering*. Springer, 1998, vol. 132.

[117] International Organization for Standardization, "Acoustics – Attenuation of sound during propagation outdoors. Part 1: Calculation of the absorption of sound by the atmosphere," Geneva, Switzerland, Standard ISO 9613–1, 1993.

[118] H. Jeong, I. Drumm, B. Horner, and Y. Lam, "The modelling of frequency dependent boundary conditions in FDTD simulation of concert hall acoustics," in *Proceedings of the International Congress on Acoustics (ICA)*, 2007.

[119] P. B. Johns, "Simulation of electromagnetic wave interactions by transmission-line modelling (TLM)," *Wave Motion*, vol. 10, no. 5, pp. 597–610, 1988.

[120] A. Kamel, P. Sguazzero, and M. Kindelan, "Cost-effective staggered schemes for the numerical simulation of wave propagation," *International Journal for Numerical Methods in Engineering*, vol. 38, no. 5, pp. 755–773, 1995.

[121] L. V. Kantorovich and V. I. Krylov, *Approximate Methods of Higher Analysis*. Interscience Publishers, 1958, translated by Curtis D. Benster.

[122] M. Karjalainen and C. Erkut, "Digital waveguides versus finite difference structures: Equivalence and mixed modeling," *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 978–989, 2004.

[123] B. F. Katz, "Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation," *Journal of the Acoustical Society of America*, vol. 110, no. 5, pp. 2440–2448, 2001.

[124] J. B. Keller, "The scope of the image method," *Communications on Pure and Applied Mathematics*, vol. 6, no. 4, pp. 505–512, 1953.

[125] ——, "Geometrical theory of diffraction," *Journal of the Optical Society of America*, vol. 52, p. 116, 1962.

[126] M. Kendall, *A Course in the Geometry of n Dimensions*. Hafner, New York, 1961.

[127] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics*, 4th ed. John Wiley & Sons, 2000.

[128] M. Kleiner, B.-I. Dalenbäck, and P. Svensson, "Auralization-an overview," *Journal of the Audio Engineering Society*, vol. 41, no. 11, pp. 861–875, 1993.

[129] R. Klette and A. Rosenfeld, *Digital Geometry: Geometric Methods for Digital Picture Analysis*. Elsevier, 2004.

[130] D. D. Kosloff and E. Baysal, "Forward modeling by a Fourier method," *Geophysics*, vol. 47, no. 10, pp. 1402–1412, 1982.

[131] K. Kowalczyk, "Boundary and medium modelling using compact finite difference schemes in simulations of room acoustics for audio and architectural design applications," Ph.D. dissertation, Queen's University Belfast, 2008.

[132] K. Kowalczyk and M. van Walstijn, "Formulation of a locally reacting wall in finite difference modelling of acoustic spaces," in *Proceedings of the International Symposium on Room Acoustics*, 2007, pp. 1–6.

[133] K. Kowalczyk and M. van Walstijn, "Formulation of locally reacting surfaces in FDTD/K-DWM modelling of acoustic spaces," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 891–906, 2008.

[134] ——, "A comparison of nonstaggered compact FDTD schemes for the 3D wave equation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, Texas, 2010, pp. 197–200.

[135] ——, "Wideband and isotropic room acoustics simulation using 2-D interpolated FDTD schemes," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 1, pp. 78–89, 2010.

[136] ——, "Room acoustics simulation using 3-D compact explicit FDTD schemes," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 1, pp. 34–46, 2011.

[137] K. Kowalczyk and M. v. Walstijn, "Modeling frequency-dependent boundaries as digital impedance filters in FDTD and K-DWM room acoustics simulations," *Journal of the Audio Engineering Society*, vol. 56, no. 7/8, pp. 569–583, 2008.

[138] A. Krokstad, S. Strom, and S. Sørsdal, "Calculating the acoustical room response by the use of a ray tracing technique," *Journal of Sound and Vibration*, vol. 8, no. 1, pp. 118–125, 1968.

[139] A. Kumar, "Isotropic finite-differences," *Journal of Computational Physics*, vol. 201, no. 1, pp. 109–118, 2004.

[140] H. Kuttruff, *Room Acoustics*. CRC Press, 2009.

[141] K. H. Kuttruff, "Auralization of impulse responses modeled on the basis of ray-tracing results," *Journal of the Audio Engineering Society*, vol. 41, no. 11, pp. 876–880, 1993.

[142] J.-O. Lachaud and B. Thibert, "Properties of gauss digitized shapes and digital surface integration," *Journal of Mathematical Imaging and Vision*, pp. 1–19, 2015.

[143] S. Laine, S. Siltanen, T. Lokki, and L. Savioja, "Accelerated beam tracing algorithm," *Applied Acoustics*, vol. 70, no. 1, pp. 172–181, 2009.

[144] J. A. Laird, "The physical modelling of drums using digital waveguides," Ph.D. thesis, University of Bristol, 2001.

[145] P. D. Lax, "Weak solutions of nonlinear hyperbolic equations and their numerical computation," *Communications on Pure and Applied Mathematics*, vol. 7, no. 1, pp. 159–193, 1954.

[146] P. D. Lax and R. D. Richtmyer, "Survey of the stability of linear finite difference equations," *Communications on Pure Applied Mathematics*, vol. 9, pp. 267–293, 1956.

[147] J. le Rond d'Alembert, "Recherches sur la courbe que forme une corde tenduë mise en vibrations," *Histoire de l'Académie royale des sciences et belles lettres pour l'année 1747*, vol. 3, pp. 214–249, 1750.

[148] M. Lees, "von Neumann difference approximation to hyperbolic equations," *Pacific Journal of Mathematics*, vol. 10, pp. 213–222, 1960.

[149] R. R. B. Lehoucq, D. D. C. Sorensen, and C.-C. Yang, *ARPACK User's Guide: Solution of Large-Scale Eigenvalue Problems With Implicitly Restarted Arnoldi Methods*. SIAM, 1998, vol. 6.

[150] S. K. Lele, "Compact finite difference schemes with spectral-like resolution," *Journal of Computational Physics*, vol. 103, no. 1, pp. 16–42, 1992.

[151] R. LeVeque, *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*. SIAM, 2007.

[152] R. J. LeVeque, *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002, vol. 31.

[153] D. Lilly, "A proposed staggered-grid system for numerical integration of dynamic equations," *Monthly Weather Review*, vol. 89, pp. 59–65, 1961.

[154] Q. Liu, "The PSTD algorithm: A time-domain method requiring only two cells per wavelength," *Microwave and Optical Technology Letters*, vol. 15, no. 3, pp. 158–165, 1997.

[155] Q. H. Liu, "The pseudospectral time-domain (PSTD) algorithm for acoustic waves in absorptive media," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 45, no. 4, pp. 1044–1055, 1998.

[156] Y. Liu and M. K. Sen, "A new time–space domain high-order finite-difference method for the acoustic wave equation," *Journal of Computational Physics*, vol. 228, no. 23, pp. 8779–8806, 2009.

[157] ——, "A new time–space domain high-order finite-difference method for the acoustic wave equation," *Journal of Computational Physics*, vol. 228, no. 23, pp. 8779–8806, 2009.

[158] ——, "Time–space domain dispersion-relation-based finite-difference method with arbitrary even-order accuracy for the 2D acoustic wave equation," *Journal of Computational Physics*, vol. 232, no. 1, pp. 327–345, 2013.

[159] T. Lokki, A. Southern, and L. Savioja, "Studies on seat dip effect with 3D FDTD modeling," in *Proceedings of Forum Acusticum*, Aalborg, Denmark, 2011.

[160] T. Lokki, A. Southern, S. Siltanen, and L. Savioja, "Acoustics of Epidaurus–studies with room acoustics modelling methods," *Acta Acustica united with Acustica*, vol. 99, no. 1, pp. 40–47, 2013.

[161] A. M. Lyapunov, "The general problem of the stability of motion," *International Journal of Control*, vol. 55, no. 3, pp. 531–534, 1992.

[162] N. Macon and A. Spitzbart, "Inverses of Vandermonde matrices," *American Mathematical Monthly*, pp. 95–100, 1958.

[163] S. Marchand and P. Depalle, "Generalization of the derivative analysis method to non-stationary sinusoidal modeling," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Espoo, Finland, 2008, pp. 281–288.

[164] *MATLAB version 8.5.0.197613 (R2015a)*, The Mathworks, Inc., Natick, Massachusetts, 2015.

[165] R. Mehra, N. Raghuvanshi, L. Savioja, M. C. Lin, and D. Manocha, "An efficient GPU-based time domain solver for the acoustic wave equation," *Applied Acoustics*, vol. 73, no. 2, pp. 83–94, 2012.

[166] J. G. Meloney and K. E. Cummings, "Adaptation of FDTD techniques to acoustic modeling," in *11th Annual Review of Progress in Applied Computational Electromagnetics*, vol. 2, 1995, pp. 724–731.

[167] R. Merris, "Laplacian matrices of graphs: a survey," *Linear algebra and its applications*, vol. 197, pp. 143–176, 1994.

[168] R. M. Mersereau, "The processing of hexagonally sampled two-dimensional signals," *Proceedings of the IEEE*, vol. 67, no. 6, pp. 930–949, 1979.

[169] A. Meshram, R. Mehra, and D. Manocha, "Efficient HRTF computation using adaptive rectangular decomposition," in *IEEE Mixed & Augmented Reality (ISMAR)*, Munich, Bavaria, Germany, 2014.

[170] C. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: SIAM, 2000, vol. 2.

[171] A. Mitchell, *Computational Methods in Partial Differential Equations*. Wiley New York, 1969, vol. 7, no. 9.

[172] P. Mokhtari, H. Takemoto, R. Nishimura, and H. Kato, "Computer simulation of HRTFs for personalization of 3D audio," in *International Symposium on Universal Communication (ISUC)*. IEEE, 2008, pp. 435–440.

[173] N. M. Morales, R. Mehra, and D. Manocha, "Methods, systems, and computer readable media for utilizing parallel adaptive rectangular decomposition (ARD) to perform acoustic simulations," Jun. 2015, US Patent Application 14/743,598.

[174] P. M. Morse and K. U. Ingard, *Theoretical Acoustics*. Princeton University Press, 1968.

[175] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*. New York: McGraw-Hill, 1953.

[176] P. M. Morse, *Vibration and Sound*. McGraw-Hill New York, 1948, vol. 2.

[177] E. Motuk, R. Woods, and S. Bilbao, "Implementation of finite difference schemes for the wave equation on FPGA," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3. IEEE, 2005, pp. iii–237.

[178] R. P. Muñoz and M. Hornikx, "A hybrid Fourier PSTD/DG method to solve the linearized Euler equations: optimization and accuracy," in *22nd AIAA/CEAS Aeroacoustics Conference*, Lyon, France, 2016.

[179] D. T. Murphy, "Digital waveguide mesh topologies in room acoustics modelling," Ph.D. thesis, University of York, 2000.

[180] D. T. Murphy and M. Beeson, "The KW-boundary hybrid digital waveguide mesh for room acoustics applications," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 2, pp. 552–564, 2007.

[181] D. T. Murphy and D. M. Howard, "Digital waveguide mesh topologies in room acoustics modelling," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Verona, Italy, 2000.

[182] D. T. Murphy, A. Kelloniemi, J. Mullen, and S. Shelley, "Acoustic modeling using the digital waveguide mesh," *IEEE Signal Processing Magazine*, vol. 24, no. 2, pp. 55–66, 2007.

[183] S. Ničkovic, M. Gavrilov, and I. Tošic, "Geostrophic adjustment on hexagonal grids," *Monthly Weather Review*, vol. 130, no. 3, pp. 668–683, 2002.

[184] M. Noisternig, B. F. Katz, S. Siltanen, and L. Savioja, "Framework for real-time auralization in architectural acoustics," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 1000–1015, 2008.

[185] G. G. O'Brien, M. A. Hyman, and S. Kaplan, "A study of the numerical solution of partial differential equations," *Journal of Mathematical Physics*, vol. 29, pp. 223–251, 1950.

[186] T. Okuzono, T. Yoshida, K. Sakagami, and T. Otsuru, "An explicit time-domain finite element method for room acoustics simulations: Comparison of the performance with implicit methods," *Applied Acoustics*, vol. 104, pp. 76–84, 2016.

[187] S. A. Orszag and M. Israeli, "Numerical simulation of viscous incompressible flows," *Annual Review of Fluid Mechanics*, vol. 6, no. 1, pp. 281–318, 1974.

[188] M. Patra and M. Karttunen, "Stencils with isotropic discretization error for differential operators," *Numerical Methods for Partial Differential Equations*, vol. 22, no. 4, pp. 936–953, 2006.

[189] J. Peiró and S. Sherwin, "Finite difference, finite element and finite volume methods for partial differential equations," in *Handbook of Materials Modeling*. Springer, 2005, pp. 2415–2446.

[190] 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*, vol. 4, pp. 113–138, 2014.

[191] D. P. Petersen and D. Middleton, "Sampling and reconstruction of wave-number-limited functions in N-dimensional Euclidean spaces," *Information and Control*, vol. 5, no. 4, pp. 279–323, 1962.

[192] K. Petkov, F. Qiu, Z. Fan, A. E. Kaufman, and K. Mueller, "Efficient LBM visual simulation on face-centered cubic lattices," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 5, pp. 802–814, 2009.

[193] H. Phillips and N. Wiener, "Nets and the Dirichlet problem," *Journal of Mathematics and Physics*, vol. 2, p. 105, 1923.

[194] A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications*. McGraw-Hill New York, 1989.

[195] G. W. Platzman, "The lattice structure of the finite-difference primitive and vorticity equations," *Monthly Weather Review*, vol. 86, no. 8, pp. 285–292, 1958.

[196] M. Potter, M. Lamoureux, and M. Nauta, "An FDTD scheme on a face-centered-cubic (FCC) grid for the solution of the wave equation," *Journal of Computational Physics*, vol. 230, no. 15, pp. 6169–6183, 2011.

[197] S. Prepeliă, M. Geronazzo, F. Avanzini, and L. Savioja, "Influence of voxelization on finite difference time domain simulations of head-related transfer functions," *Journal of the Acoustical Society of America*, vol. 139, no. 5, pp. 2489–2504, 2016.

[198] A. Quarteroni, A. M. Quarteroni, and A. Valli, *Numerical Approximation of Partial Differential Equations*. Springer, 2008, vol. 23.

[199] N. Raghuvanshi, R. Narain, and M. C. Lin, "Efficient and accurate sound propagation using adaptive rectangular decomposition," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 5, pp. 789–801, 2009.

[200] R. Renaut-Williamson, "Full discretizations of $u_{tt} = u_{xx}$ and rational approximations to $\cosh \mu z$," *SIAM Journal on Numerical Analysis*, vol. 26, no. 2, pp. 338–347, 1989.

[201] M. Reshef, D. Kosloff, M. Edwards, and C. Hsiung, "Three-dimensional elastic modeling by the Fourier method," *Geophysics*, vol. 53, no. 9, pp. 1184–1193, 1988.

[202] L. Rhaouti, A. Chaigne, and P. Joly, "Time-domain modeling and numerical simulation of a kettledrum," *Journal of the Acoustical Society of America*, vol. 105, pp. 3545–3562, 1999.

[203] L. F. Richardson, "The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam," *Royal Society of London Philosophical Transactions Series A*, vol. 210, pp. 307–357, 1911.

[204] L. F. Richardson and J. A. Gaunt, "The deferred approach to the limit. Part I. Single lattice. Part II. Interpenetrating lattices," *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, pp. 299–361, 1927.

[205] L. Richardson, *Weather Prediction by Numerical Process*. Cambridge University Press, 1922.

[206] R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-value Problems*, 1st ed. Interscience Publishers, 1957.

[207] ——, *Difference Methods for Initial-value Problems*, 2nd ed. Interscience Publishers, 1967.

[208] S. W. Rienstra and A. Hirschberg, "An introduction to acoustics," Lecture notes, Eindhoven University of Technology, Jan. 2015.

[209] C. Roads, *The Computer Music Tutorial*. MIT press, 1996.

[210] N. Röber, M. Spindler, and M. Masuch, "Waveguide-based room acoustics through graphics hardware," in *Proceedings of the International Computer Music Conference (ICMC)*, 2006.

[211] T. Rossing, *Springer Handbook of Acoustics*, 2nd ed. Springer, 2007.

[212] P. Ruiz, "A technique for simulating the vibrations of strings with a digital computer," Master's thesis, University of Illinois, Champaign-Urbana, Illinois, 1969.

[213] Y. Saad, *Iterative Methods for Sparse Linear Systems*. SIAM, 2003.

[214] J. Saarelma, J. Botts, B. Hamilton, and L. Savioja, "Audibility of dispersion error in room acoustic finite-difference time-domain simulation as a function of simulation distance," *Journal of the Acoustical Society of America*, vol. 139, no. 4, pp. 1822–1832, 2016.

[215] J. Saarelma and L. Savioja, "An open source finite difference time-domain solver for room acoustics using graphics processing units," in *Proceedings of Forum Acusticum*, Krakow, Poland, 2014.

[216] S. Sakamoto and H. Tachibana, "Calculation of impulse response in 3-D sound field with absorptive boundary by the finite difference method," in *Proceedings of INTERNOISE*, vol. 3. Noise Control Foundation, 1997, pp. 1597–1600.

[217] S. Sakamoto, "Phase-error analysis of high-order finite difference time domain scheme and its influence on calculation results of impulse response in closed sound field," *Acoustical Science and Technology*, vol. 28, no. 5, pp. 295–309, 2007.

[218] S. Sakamoto, H. Nagatomo, A. Ushiyama, and H. Tachibana, "Calculation of impulse responses and acoustic parameters in a hall by the finite-difference time-domain method," *Acoustical Science and Technology*, vol. 29, no. 4, pp. 256–265, 2008.

[219] T. Sakuma, S. Sakamoto, and T. Otsuru, *Computational Simulation in Architectural and Environmental Acoustics*. Springer, 2014.

[220] A. Saleh and P. Blanchfield, "Analysis of acoustic radiation patterns of array transducers using the TLM method," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 3, no. 1, pp. 39–56, 1990.

[221] L. Savioja, "Improving the three-dimensional digital waveguide mesh by interpolation," *Proceedings of the Nordic Acoustical Meeting*, pp. 265–268, 1998.

[222] ——, "Real-time 3D finite-difference time-domain simulation of low-and mid-frequency room acoustics," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, vol. 1, Graz, Austria, 2010, p. 75.

[223] L. Savioja, T. J. Rinne, and T. Takala, "Simulation of room acoustics with a 3-D finite difference mesh," in *Proceedings of the International Computer Music Conference (ICMC)*, Danish Institute of Electroacoustic Music, Denmark, 1994, pp. 463–466.

[224] L. Savioja and V. Välimäki, "Improved discrete-time modeling of multi-dimensional wave propagation using the interpolated digital waveguide mesh," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 1, Munich, Germany, 1997, pp. 459–462.

[225] ——, "Interpolated rectangular 3-D digital waveguide mesh algorithms with frequency warping," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 6, pp. 783–790, 2003.

[226] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen, "Creating interactive virtual acoustic environments," *Journal of the Audio Engineering Society*, vol. 47, no. 9, pp. 675–705, 1999.

[227] L. Savioja, M. Karjalainen, and T. Takala, "DSP formulation of a finite difference method for room acoustics simulation," in *Proceedings of IEEE Nordic Signal Processing Symposium*, 1996, pp. 455–458.

[228] L. Savioja and U. P. Svensson, "Overview of geometrical room acoustic modeling techniques," *Journal of the Acoustical Society of America*, vol. 138, no. 2, pp. 708–730, 2015.

[229] G. P. Scavone, "An acoustic analysis of single-reed woodwind instruments with an emphasis on design and performance issues and digital waveguide modeling techniques," Ph.D. dissertation, Stanford University, 1997.

[230] J. Schneider and C. Wagner, "FDTD dispersion revisited: Faster-than-light propagation," *IEEE Microwave and Guided Wave Letters*, vol. 9, no. 2, pp. 54–56, 1999.

[231] J. B. Schneider, C. L. Wagner, and R. J. Kruhlak, "Simple conformal methods for finite-difference time-domain modeling of pressure-release surfaces," *Journal of the Acoustical Society of America*, vol. 104, no. 6, pp. 3219–3226, 1998.

[232] M. R. Schroeder, "Digital simulation of sound transmission in reverberant spaces," *Journal of the Acoustical Society of America*, vol. 47, no. 2A, pp. 424–431, 1970.

[233] ——, "Computer models for concert hall acoustics," *American Journal of Physics*, vol. 41, no. 4, pp. 461–471, 1973.

[234] P. Sguazzero, M. Kindelan, and A. Kamel, "Dispersion-bounded numerical integration of the elastodynamic equations with cost-effective staggered schemes," *Computer Methods in Applied Mechanics and Engineering*, vol. 80, no. 1, pp. 165–172, 1990.

[235] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[236] J. Sheaffer and B. Fazenda, "FDTD/K-DWM simulation of 3D room acoustics on general purpose graphics hardware using compute unified device architecture (CUDA)," *Proceedings of the Institute of Acoustics*, vol. 32, no. 5, 2010.

[237] ——, "From source to brain: Modelling sound propagation and localisation in rooms," Ph.D. dissertation, University of Salford, 2013.

[238] ——, "Wavecloud: an open source room acoustics simulator using the finite difference time domain method," in *Proceedings of Forum Acusticum*, Krakow, Poland, 2014.

[239] J. Sheaffer, C. J. Webb, and B. M. Fazenda, "Modelling binaural receivers in finite difference simulation of room acoustics," in *Proceedings of the International Congress on Acoustics (ICA)*, Montréal, Canada, 2013.

[240] W. Sheppard, "Central-difference formulae," *Proceedings of the London Mathematical Society*, vol. 1, no. 1, pp. 449–488, 1899.

[241] G. Shubin and J. Bell, "A modified equation approach to constructing fourth order methods for acoustic wave propagation," *SIAM Journal on Scientific and Statistical Computing*, vol. 8, no. 2, pp. 135–151, 1987.

[242] T. Shuku and K. Ishihara, "The analysis of the acoustic field in irregularly shaped rooms by the finite element method," *Journal of Sound and Vibration*, vol. 29, no. 1, pp. 67–IN1, 1973.

[243] S. Siltanen, T. Lokki, and L. Savioja, "Rays or waves? Understanding the strengths and weaknesses of computational room acoustics modeling techniques," in *Proceedings of the International Symposium on Room Acoustics*, 2010.

[244] G. D. Smith, *Numerical Solution of Partial Differential Equations: with Exercises and Worked Solutions*. Oxford University Press, 1965.

[245] J. O. Smith III, "A new approach to digital reverberation using closed waveguide networks," in *Proceedings of the International Computer Music Conference (ICMC)*, Vancouver, 1985, pp. 47–53.

[246] ——, "Efficient simulation of the reed-bore and bow-string mechanisms," in *Proceedings of the International Computer Music Conference (ICMC)*, The Hague, 1986, pp. 275–280.

[247] ——, "Waveguide digital filters," *CCRMA Technical Report*, 1987.

[248] ——, "Physical modeling using digital waveguides," *Computer Music Journal*, pp. 74–91, 1992.

[249] ——, *Introduction to Digital Filters: With Audio Applications*. W3K Publishing, 2007, vol. 2.

[250] ——, *Physical Audio Signal Processing*. W3K Publishing, 2010.

[251] A. J. W. Sommerfeld, *Partial Differential Equations in Physics*. Academic press, 1949.

[252] D. M. Y. Sommerville, *An Introduction to the Geometry of n Dimensions*. London, 1929, vol. 2.

[253] A. Southern, S. Siltanen, D. Murphy, and L. Savioja, "Room impulse response synthesis and validation using a hybrid acoustic model," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 9, Sep. 2013.

[254] C. Spa, "Time-domain numerical methods in room acoustics simulations," Ph.D. thesis, Universitat Pompeu Fabra, 2009.

[255] C. Spa, A. Rey, and E. Hernandez, "A GPU implementation of an explicit compact FDTD algorithm with a digital impedance filter for room acoustics applications," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 8, pp. 1368–1380, 2015.

[256] D. Speray and S. Kennon, "Volume probes: Interactive data exploration on arbitrary grids," *Computer Graphics*, vol. 24, no. 5, 1990.

[257] W. F. Spotz and G. F. Carey, "A high-order compact formulation for the 3D Poisson equation," *Numerical Methods for Partial Differential Equations*, vol. 12, no. 2, pp. 235–243, 1996.

[258] G. G. Stokes, "Sound attenuation due to viscosity," *Transactions of the Cambridge Philosophical Society*, vol. 8, pp. 75–102, 1845.

[259] G. Strang, "The discrete cosine transform," *SIAM Review*, vol. 41, no. 1, pp. 135–147, 1999.

[260] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*. Philadelphia, PA: SIAM, 2004.

[261] M. Svärd and J. Nordström, "Review of summation-by-parts schemes for initial–boundary-value problems," *Journal of Computational Physics*, vol. 268, pp. 17–38, 2014.

[262] P. Svensson and U. R. Kristiansen, "Computational modelling and simulation of acoustic spaces," in *AES Conference on Virtual, Synthetic, and Entertainment Audio*. Audio Engineering Society, 2002.

[263] U. P. Svensson, R. I. Fred, and J. Vanderkooy, "An analytic secondary source model of edge diffraction impulse responses," *Journal of the Acoustical Society of America*, vol. 106, no. 5, pp. 2331–2344, 1999.

[264] A. Taflove, "Application of the finite-difference time-domain method to sinusoidal steady-state electromagnetic-penetration problems," *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-22, no. 3, pp. 191–202, 1980.

[265] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*, 1st ed. Artech House, 1995.

[266] H. Tal-Ezer, "Spectral methods in time for hyperbolic equations," *SIAM Journal on Numerical Analysis*, vol. 23, no. 1, pp. 11–26, 1986.

[267] S. Thampi, S. Ansumali, R. Adhikari, and S. Succi, "Isotropic discrete Laplacian operators from lattice hydrodynamics," *Journal of Computational Physics*, 2012.

[268] J. W. Thomas, *Numerical Partial Differential Equations: Finite Difference Methods*. Springer, 1995, vol. 1.

[269] V. Thomée, "From finite differences to finite elements: A short history of numerical analysis of partial differential equations," *Journal of Computational and Applied Mathematics*, vol. 128, no. 1, pp. 1–54, 2001.

[270] J. Thompson, B. Soni, and N. Weatherill, *Handbook of Grid Generation*. CRC Press, 1998.

[271] J. G. Tolan and J. B. Schneider, "Locally conformal method for acoustic finite-difference time-domain modeling of rigid surfaces," *Journal of the Acoustical Society of America*, vol. 114, no. 5, pp. 2575–2581, 2003.

[272] A. Torin, "Percussion instrument modelling in 3D: Sound synthesis through time domain numerical simulation," Ph.D. thesis, University of Edinburgh, 2016.

[273] A. Torin and S. Bilbao, "A 3D multi-plate environment for sound synthesis," in *Proceedings of the 16th International Digital Audio Effects Conference (Maynooth, Ireland)*, 2013.

[274] C. D. Toth, J. O'Rourke, and J. E. Goodman, *Handbook of Discrete and Computational Geometry*. CRC press, 2004.

[275] L. N. Trefethen, "Group velocity in finite difference schemes," *SIAM Review*, vol. 24, no. 2, pp. 113–136, 1982.

[276] ——, "Wave propagation and stability for finite difference schemes," Ph.D. thesis, Stanford University, 1982.

[277] ——, *Spectral Methods in MATLAB*. SIAM, 2000, vol. 10.

[278] N. Tsingos, T. Funkhouser, A. Ngan, and I. Carlbom, "Modeling acoustics in virtual environments using the uniform theory of diffraction," in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 2001, pp. 545–552.

[279] T. Tsuchiya, Y. Iwaya, and M. Otani, "Large-scale sound field rendering with graphics processing unit cluster for three-dimensional audio with loudspeaker array," in *Proceedings of the International Congress on Acoustics (ICA)*, Montréal, Canada, 2013.

[280] J. Tuomela, "Fourth order schemes for wave equation, Maxwell's equation and linearized elastodynamic equations," Institut National de Recherche en Informatique et en Automatique (INRIA), Tech. Rep., 1990.

[281] ——, "Fourth-order schemes for the wave equation, Maxwell equations, and linearized elastodynamic equations," *Numerical Methods for Partial Differential Equations*, vol. 10, no. 1, pp. 33–63, 1994.

[282] ——, "On the construction of arbitrary order schemes for the many dimensional wave equation," *BIT Numerical Mathematics*, vol. 36, no. 1, pp. 158–165, 1996.

[283] ——, "A note on high order schemes for the one dimensional wave equation," *BIT Numerical Mathematics*, vol. 35, no. 3, pp. 394–405, 1995.

[284] V. Välimäki, J. D. Parker, L. Savioja, J. O. Smith III, and J. S. Abel, "Fifty years of artificial reverberation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 5, pp. 1421–1448, 2012.

[285] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy, "Physical modeling of plucked string instruments with application to real-time sound synthesis," in *Audio Engineering Society Convention*. Audio Engineering Society, 1995.

[286] S. A. van Duyne and J. O. Smith III, "Physical modeling with the 2-D digital waveguide mesh," in *Proceedings of the International Computer Music Conference (ICMC)*, Tokyo, Japan, 1993.

[287] ——, "The tetrahedral digital waveguide mesh," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, 1995, pp. 234–237.

[288] ——, "The 3D tetrahedral digital waveguide mesh with musical applications," in *Proceedings of the International Computer Music Conference (ICMC)*, Hong Kong, 1996, pp. 9–16.

[289] C. Van Loan, *Computational Frameworks for the Fast Fourier Transform*. SIAM, 1992.

[290] J. van Mourik and D. Murphy, "Explicit higher-order FDTD schemes for 3D room acoustic simulation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 12, pp. 2003–2011, 2014.

[291] M. van Walstijn and K. Kowalczyk, "On the numerical solution of the 2D wave equation with compact FDTD schemes," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Espoo, Finland, 2008, pp. 205–212.

[292] M. van Walstijn and M. Campbell, "Discrete-time modeling of woodwind instrument bores using wave variables," *Journal of the Acoustical Society of America*, vol. 113, no. 1, pp. 575–585, 2003.

[293] R. S. Varga, *Matrix Iterative Analysis*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

[294] R. Vichnevetsky and J. B. Bowles, *Fourier Analysis of Numerical Approximations of Hyperbolic Problems*. Philadelphia, PA: SIAM, 1982.

[295] J. Virieux, "SH-wave propagation in heterogeneous media: velocity-stress finite-difference method," *Geophysics*, vol. 49, no. 11, pp. 1933–1942, 1984.

[296] J. von Neumann and R. D. Richtmyer, "A method for the numerical calculation of hydrodynamic shocks," *Journal of Applied Physics*, vol. 21, no. 3, pp. 232–237, 1950.

[297] M. Vorländer, "Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm," *Journal of the Acoustical Society of America*, vol. 86, p. 172, 1989.

[298] ——, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. Springer Science & Business Media, 2007.

[299] ——, "Computer simulations in room acoustics: Concepts and uncertainties," *Journal of the Acoustical Society of America*, vol. 133, no. 3, pp. 1203–1213, 2013.

[300] C. L. Wagner and J. B. Schneider, "An acoustic finite-difference time-domain algorithm with isotropic dispersion," *Journal of Computational Acoustics*, vol. 13, no. 02, pp. 365–384, 2005.

[301] R. Warming and B. Hyett, "The modified equation approach to the stability and accuracy analysis of finite-difference methods," *Journal of Computational Physics*, vol. 14, no. 2, pp. 159–179, 1974.

[302] C. J. Webb and S. Bilbao, "Computing room acoustics with CUDA - 3D FDTD schemes with boundary losses and viscosity," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Prague, Czech Republic, 2011, pp. 317–320.

[303] C. J. Webb and A. Gray, "Large-scale virtual acoustics simulation at audio rates using three dimensional finite difference time domain and multiple GPUs," in *Proceedings of the International Congress on Acoustics (ICA)*, Montréal, Canada, 2013.

[304] C. J. Webb, "Parallel computation techniques for virtual acoustics and physical modelling synthesis," Ph.D. thesis, University of Edinburgh, 2014.

[305] C. J. Webb and S. Bilbao, "Binaural simulations using audio rate FDTD schemes and CUDA," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, York, UK, 2012, pp. 97–100.

[306] L. Weinberg, *Network Analysis and Synthesis*. RE Krieger Pub. Co., 1975.

[307] N. Whitehead and A. Fit-Florea, "Precision & performance: Floating point and IEEE 754 compliance for NVIDIA GPUs," NVIDIA Corporation, White paper, 2011.

[308] *Mathematica version 8.0*, Wolfram Research, Inc., Champaign, Illinois, 2010.

[309] M. Woodward and F. Muir, "Hexagonal finite difference operators and 3-D wave equation migration," *Stanford Exploration Project*, vol. 38, pp. 195–206, 1984.

[310] T. Xiao and Q. H. Liu, "Finite difference computation of head-related transfer function for human hearing," *Journal of the Acoustical Society of America*, vol. 113, no. 5, pp. 2434–2441, 2003.

[311] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Transactions on Antennas and Propagation*, vol. 14, no. 3, pp. 302–307, 1966.

[312] T. Yokota, S. Sakamoto, and H. Tachibana, "Visualization of sound propagation and scattering in rooms." *Acoustical Science and Technology*, vol. 23, no. 1, pp. 40–46, 2002.

[313] O. Zienkiewicz and E. Onate, "Finite elements versus finite volumes. Is there really a choice?" *Nonlinear Computational Mechanics–State of the Art*, 1991.

[314] O. C. Zienkiewicz, R. L. Taylor, O. C. Zienkiewicz, and R. L. Taylor, *The Finite Element Method*. McGraw-Hill London, 1977, vol. 3.

[315] U. Zölzer, *DAFX: Digital Audio Effects*. John Wiley & Sons, 2011.

Glossary

List of Abbreviations and Acronyms

°C	degrees Celsius
dB	decibels
kg	kilograms
m	metres
Pa, kPa	Pascals, kiloPascals
rad	radians
s	seconds
w.r.t.	with respect to
BCC	body-centered cubic
BEM	boundary element method
BVP	boundary value problem
CCP	cubic close-packed
CFL	Courant-Friedrichs-Lewy
CPU	central processing unit
DC	direct current (zero wavenumber or frequency)
DCT	discrete cosine transform
DFT	discrete Fourier transform
DIF	digital impedance filter
DWM	digital waveguide mesh
FCC	face-centered cubic
FDTD	finite difference time domain
FEM	finite element method
FFT	fast Fourier transform
FLOP	floating-point operation
FMA	fused multiply-add
FPGA	field programmable gate array
FVTD	finite volume time domain
GCT	Gershgorin's circle theorem
GKSO	Gustaffson-Kreiss-Sundstrom-Oliger
GPU	graphics processing unit
IIR	infinite impulse response
IISO	interpolated isotropic
IVP	initial value problem

IWB	interpolated wideband
KIE	Kirchhoff integral equation
KHIE	Kirchhoff-Helmoltz integral equation
LBM	lattice Boltzmann method
MAC	multiply-accumulate
OoA	order of accuracy
OCTA	octahedral
ODE	ordinary differential equation
PDE	partial differential equation
PML	perfectly matched layer
PPW	points per wavelength
RCE	relative computational efficiency
RIR	room impulse response
RLC	resistor (R) inductor (L) capacitor (C)
SBP	summation-by-parts
SLF	standard leapfrog
nz	number of non-zero elements (in a sparse matrix)

List of Symbols for Chapter 2

x, y, z	spatial coordinates in $\mathbb{R}^3$
$\mathbf{x}$	spatial position vector in $\mathbb{R}^3$, $\mathbf{x} = (x, y, z)$
t	time, in seconds
p	pressure field, in kg/(m·s^2)
$\mathbf{v}$	particle vector velocity field, $\mathbf{v} = (v_x, v_y, v_z)$
v_x, v_y, v_z	particle velocity vector field components in m/s
c	speed of sound in air, in m/s
ρ	ambient density of air, in kg/m^3
κ	bulk modulus, in kg/(m·s^2)
∂_t	partial derivative w.r.t. time
$\partial_x, \partial_y, \partial_y$	partial derivative w.r.t. spatial coordinates x, y, z
∇	3-D gradient operator, $\nabla = (\partial_x, \partial_y, \partial_y)^{\mathrm{T}}$
$\nabla\cdot$	3-D divergence operator, $\nabla\cdot = (\partial_x, \partial_y, \partial_y)$
$\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y}$	differential volume elements
$\mathrm{d}\sigma$	differential surface element
$\mathcal{H}$	stored energy, in Joules
$\mathcal{V}$	potential energy, in Joules
$\mathcal{T}$	kinetic energy, in Joules
Ψ	acoustic velocity potential, in m^2/s
Δ	3-D Laplace operator (Laplacian) $\Delta = \nabla \cdot \nabla = \partial_x^2 + \partial_y^2 + \partial_z^2$
p^+, p^-	right- and left-going pressure wave ($d = 1$)
r	radial position
A, B	plane-wave amplitudes
$\boldsymbol{\beta}$	wave vector, $\boldsymbol{\beta} = (\beta_x, \beta_y, \beta_z)$

$\beta_x, \beta_y, \beta_z$	wave vector components, in rad/m
β	wavenumber ($\beta = \lvert\boldsymbol{\beta}\rvert$), in rad/m
ω	angular frequency, in rad/s
f	temporal frequency, in Hz
w	position in m, along wave direction $\hat{\boldsymbol{\beta}}$
$\delta(t)$	Dirac delta in time, in s^{-1}
$\delta(x)$	Dirac delta along x-direction, in m^{-1}
$\delta(\mathbf{x})$	Dirac delta in 3-D space, in m^{-3}
q_s	source volume velocity, in m^3/s
$\mathbf{x}_0$	source position
$\mathbf{x}_1$	receiver position
D	Euclidean distance from $\mathbf{x}_0$, in metres
Ψ'	free-space Green's function
$q_s^{(m)}$	mth source volume velocity
$\mathbf{x}_0^{(m)}$	mth source position
D_m	distance from mth source position
$M^{(\mathrm{s})}$	number of source volume velocities
$F(\mathbf{x}, t)$	source velocity (forcing term), in s^{-1}
$\mathfrak{F}$	power term associated to source velocity
$\overline{\Omega}$	closed domain, $\overline{\Omega} = \Omega \cup \Gamma$
Ω	interior domain
Γ	boundary of Ω
$\mathcal{B}$	energetic term associated to boundary (e.g., power outflow)
$v_\perp$	outward normal component of particle velocity
$\hat{\mathbf{n}}$	outward normal vector at point on boundary
Y	admittance, in m^2s/kg
Y_0	characteristic admittance of air
γ	specific surface admittance (dimensionless)
Z	impedance, in $kg/(m^2s)$
Z_0	characteristic impedance of air
ζ	specific surface impedance (dimensionless)
$\phi(\mathbf{x})$	mode shape
$\psi(t)$	mode temporal function
L_x, L_y, L_z	dimensions of shoebox, in m
$\boldsymbol{m}$	modal vector index
m_x, m_y, m_z	modal vector index components
$f_{\boldsymbol{m}}$	modal frequency
R	reflection coefficient
θ	angle of incidence
s	complex number, $s = \sigma + j\omega$
σ	real part of s, damping/growth factor
e	exponential function
j	imaginary unit $\mathrm{j} = \sqrt{-1}$
$R(\omega)$	frequency-dependent reflection coefficient
$Y(\omega)$	frequency-dependent admittance

A	"inductance" coefficient of frequency-dependent admittance boundary condition, in s^{-1}
B	"resistance" coefficient of frequency-dependent admittance boundary condition
C	"inverse-capacitance" coefficient of frequency-dependent admittance boundary condition, in s
D	"inductance"coefficient of frequency-dependent impedance boundary condition, in s^{-1}
E	"resistance"coefficient of frequency-dependent impedance boundary condition
F	"inverse-capacitance"coefficient of frequency-dependent impedance boundary condition, in s
g	extra variable on boundary, ($\partial_t g = p$)
$\mathcal{H}_\text{i}$	stored energy over interior
$\mathcal{H}_\text{b}$	stored energy over boundary
$\mathcal{Q}$	power outflow
$L^{(m)}$	"inductance" coefficient of mth branch of general impedance boundary condition, in kg/m^2
$R^{(m)}$	"resistance" coefficient of mth branch of general impedance boundary condition, in $kg/(m^2 s)$
$C^{(m)}$	"capacitance" coefficient of mth branch of general impedance boundary condition, in $(ms)^2/kg$
$M^{(\text{b})}$	number of RLC circuit branches in general impedance boundary condition
η	length associated to viscothermal losses, in metres
τ_η	relaxation time associated to viscothermal losses, in seconds
α	temporal absorption coefficient, in Nepers/s
α'	spatial absorption coefficient, in Nepers/m
μ	dynamic viscosity coefficient
μ_b	bulk viscosity coefficient
l_v	characteristic length associated to viscous effects, in m
l_h	characteristic length associated to thermal effects, in m
κ'	coefficient of thermal conductivity
γ'	ratio of specific heats
C_p	heat coefficient at constant pressure for a constant unit mass
α_vib	absorption coefficient associated to vibrational modes of molecules, in Nepers/m
$\alpha_\text{vib,O}$	absorption coefficient associated to vibrational modes of oxygen, in Nepers/m
$\alpha_\text{vib,N}$	absorption coefficient associated to vibrational modes of nitrogen, in Nepers/m
α_tot	total air absorption coefficient, in Nepers/m
f_r	relaxation frequency, in Hz
$f_\text{r,O}$	relaxation frequency for oxygen
$f_\text{r,N}$	relaxation frequency for nitrogen
T	temperature, in degrees Celsius
h_r	relative humidity (%)
$\mathcal{Q}_\text{i}$	power outflow through interior
$\mathcal{Q}_\text{b}$	power outflow through boundary
$h(t)$	impulse response of linear time-invariant system
$x(t)$	input to linear time-invariant system
$y(t)$	output of linear time-invariant system
$\hat{Y}(\omega)$	Fourier transform of $y(t)$
$\hat{X}(\omega)$	Fourier transform of $x(t)$
$\hat{H}(\omega)$	Fourier transform of $h(t)$
n	discrete time index
k	time-step, in s
N	number of grid points

$\mathbf{u}^n$	$N \times 1$ grid vector at time-step n
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	$N \times N$ matrices

List of Symbols for Chapter 3

x, y, z	spatial coordinates in $\mathbb{R}^3$, in m
$\mathbf{x}$	spatial position vector in $\mathbb{R}^3$, $\mathbf{x} = (x, y, z)$
t	time, in seconds
Ψ	Acoustic velocity potential, in m^2/s
p	pressure field, in kg/(m·s^2)
$\mathbf{v}$	particle vector velocity field, $\mathbf{v} = (v_x, v_y, v_z)$
v_x, v_y, v_z	particle velocity vector field components in m/s
c	speed of sound in air, in m/s
ρ	ambient density of air, in kg/m^3
κ	bulk modulus, in kg/(m·s^2)
∂_t	partial derivative w.r.t. time
$\partial_x, \partial_y, \partial_y$	partial derivative w.r.t. spatial coordinates x, y, z
w	a coordinate axis, i.e., $w \in \{x, y, z\}$, or $w \in \pm\{x, y, z\}$
∂_w	partial derivative w.r.t. spatial coordinate $w \in \{x, y, z\}$
∇	3-D gradient operator, $\nabla = (\partial_x, \partial_y, \partial_y)^{\mathrm{T}}$
$\nabla\cdot$	3-D divergence operator, $\nabla\cdot = (\partial_x, \partial_y, \partial_y)$
f_0, g_0	initial conditions for 3-D wave equation
Δ	3-D Laplacian operator $\Delta = \nabla \cdot \nabla = \partial_x^2 + \partial_y^2 + \partial_z^2$
n	discrete time index
F_s	sample rate, in Hz
k	time-step, in s
h	spatial step (Cartesian grid spacing), in m
$\boldsymbol{i}$	discrete spatial vector index, $\boldsymbol{i} = (i_x, i_y, i_z)$
i_x, i_y, i_z	discrete spatial indices along coordinate axes
$\mathbb{Z}^+$	non-negative integers
$\mathbb{Z}_h^3$	3-D Cartesian grid
$\underline{\Psi}_{\boldsymbol{i}}^n$	Cartesian grid function, approximation to acoustic velocity potential, $\underline{\Psi}_{\boldsymbol{i}}^n = \underline{\Psi}_{i_x, i_y, i_z}$
$\hat{\mathbf{e}}_x, \hat{\mathbf{e}}_y, \hat{\mathbf{e}}_z$	standard unit vectors in $\mathbb{R}^3$
$e_{t\pm}$	temporal shifts
$e_{x\pm}, e_{y\pm}, e_{z\pm}$	spatial shifts along coordinate axes
$e_{w\pm}$	spatial shifts along coordinate axis $w \in \{x, y, z\}$
δ_{tt}	second-order spatial difference in time
δ_{ww}	second-order spatial difference along coordinate axis $w \in \{x, y, z\}$
δ_Δ	seven-point discrete Laplacian, $\delta_\Delta = \delta_{xx} + \delta_{yy} + \delta_{zz}$
$Q_{\boldsymbol{i}}^n$	summation over six spatial neighbours of $\underline{\Psi}_{\boldsymbol{i}}^n$, $Q_{\boldsymbol{i}}^n = Q_{i_x, i_y, i_z}^n$
λ	Courant number, $\lambda = ck/h$
g_{rc}	3-D raised cosine
$\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}}$	staggered pressure grid function, $\underline{p}_{\boldsymbol{i}}^{n+\frac{1}{2}} = \underline{p}_{i_x, i_y, i_z}^{n+\frac{1}{2}}$
$\underline{v}_{x, \boldsymbol{i}+\hat{\mathbf{e}}_x/2}^n$	staggered particle velocity grid function, x-component, $\underline{v}_{x, \boldsymbol{i}+\hat{\mathbf{e}}_x/2}^n = \underline{v}_{x, i_x+\frac{1}{2}, i_y, i_z}^n$

$\underline{v}^n_{y,\boldsymbol{i}+\hat{\mathbf{e}}_y/2}$	staggered particle velocity grid function, y-component, $\underline{v}^n_{y,\boldsymbol{i}+\hat{\mathbf{e}}_y/2} = \underline{v}^n_{y,i_x,i_y+\frac{1}{2},i_z}$
$\underline{v}^n_{z,\boldsymbol{i}+\hat{\mathbf{e}}_z/2}$	staggered particle velocity grid function, z-component, $\underline{v}^n_{z,\boldsymbol{i}+\hat{\mathbf{e}}_z/2} = \underline{v}^n_{z,i_x,i_y,i_z+\frac{1}{2}}$
$\underline{\mathbf{v}}^n_{\boldsymbol{i}+\frac{1}{2}}$	staggered particle velocity vector grid function, $\underline{\mathbf{v}}^n_{\boldsymbol{i}+\frac{1}{2}} = (\underline{v}^n_{x,\boldsymbol{i}+\hat{\mathbf{e}}_x/2}, \underline{v}^n_{y,\boldsymbol{i}+\hat{\mathbf{e}}_y/2}, \underline{v}^n_{z,\boldsymbol{i}+\hat{\mathbf{e}}_z/2})^{\mathrm{T}}$
$\delta_{t\pm}$	first-order time differences (forwards/backwards)
$\delta_{w\pm}$	first-order spatial differences along coordinate axis $w \in \{x, y, z\}$ (forwards/backwards)
$\boldsymbol{\delta}_{\mathbf{x}\pm}$	discrete 3-D gradient, $\boldsymbol{\delta}_{\mathbf{x}\pm} = (\delta_{x\pm}, \delta_{y\pm}, \delta_{z\pm})^{\mathrm{T}}$
$\underline{u}^n_{\boldsymbol{i}}, \underline{v}^n_{\boldsymbol{i}}$	arbitrary Cartesian grid functions
Z_0	characteristic impedance of air, $Z_0 = \rho c$
Y_0	characteristic admittance of air, $Y_0 = 1/Z_0$
$e^{1/2}_{t\pm}$	half-step temporal shifts
$e^{1/2}_{w\pm}$	half-step spatial shifts along coordinate axis $w \in \{x, y, z\}$
$\mathrm{d}\mathbf{x}$	differential volume element
$\mathrm{d}\sigma$	differential surface element
$\mathcal{C}$	cubic cell
$\mathcal{S}_{w\pm}$	square faces of cubic cell on sides associated to directions $w \in \{x, y, z\}$
$v_{w\pm}$	particle velocity vector components associated to face $\mathcal{S}_{w\pm}$
$\delta_{t\cdot}$	first-order time difference (centered)
$\mu_{t\pm}$	temporal averaging operators (forwards/backwards)
μ_{tt}	temporal averaging operator (centered)
$\delta_{w\cdot}$	first-order spatial difference (centered), along coordinate axis $w \in \{x, y, z\}$
$\mu_{w\pm}$	spatial averaging operator (forwards/backwards), along coordinate axis $w \in \{x, y, z\}$
$\mathfrak{h}$	stored discrete energy
$\mathfrak{v}$	potential discrete energy
$\mathfrak{t}$	kinetic discrete energy
ω	angular frequency, in rad/s
$\boldsymbol{\beta}$	wave vector, $\boldsymbol{\beta} = (\beta_x, \beta_y, \beta_z)$
$\beta_x, \beta_y, \beta_z$	wave vector components, in rad/m
e	exponential function
j	imaginary unit $\mathrm{j} = \sqrt{-1}$
$\tilde{v}_{\mathrm{p}}$	relative phase velocity of scheme
$\boldsymbol{\beta}_h$	normalised wave vector, $\boldsymbol{\beta}_h = \boldsymbol{\beta} h$
β_{yz}	side-diagonal wave vector component, along $(\hat{\mathbf{e}}_y + \hat{\mathbf{e}}_z)\sqrt{2}$, in rad/m
$\overline{\Omega}$	closed domain, $\overline{\Omega} = \Omega \cup \Gamma$
Ω	interior domain
Γ	boundary of Ω
L_x, L_y, L_z	dimensions of shoebox, in m
$\mathcal{V}_{\mathrm{box}}$	box-shaped region
$\mathbb{G}_h$	Cartesian grid for closed domain, $\mathbb{G}_h = \mathbb{Z}^3_h \cap \overline{\Omega}$
$\mathcal{I}$	vector index set for $\mathbb{G}_h$
$\overline{\Omega}_h$	staircased region, approximating $\overline{\Omega}$
Ω_h	staircased interior region, approximating Ω
Γ_h	staircased boundary surface region, approximating Γ
$q_{\boldsymbol{i},w\pm}$	indicator function
$\overline{q}_{\boldsymbol{i},w\pm}$	complementary indicator function

$\mathcal{I}_{\mathfrak{i}}$	vector index set for regular interior points
$\mathcal{I}_{\mathfrak{b}}$	vector index set for "boundary points"
N_w	dimension of Cartesian box-grid along $w \in \{x, y, z\}$ coordinate axis
σ^2	variance of spatial Gaussian
$\delta(\mathbf{x})$	Dirac delta in 3-D space, in m^{-3}
$\mathbf{x}_0$	source position
$\mathbf{x}_1$	receiver position
γ	specific surface admittance (dimensionless)
$\mathbb{R}^+$	non-negative real numbers
$\mathfrak{b}$	discrete energetic term associated to boundary (e.g., power outflow)
$\overline{\mathfrak{h}}$	accumulated stored energy
$\overline{\mathfrak{b}}$	accumulated dissipated power at boundary
$\gamma_{\boldsymbol{i},w\pm}$	admittance at boundary node $\boldsymbol{i}$ associated to outward direction $w \in \pm\{x, y, z\}$
$\gamma'_{\boldsymbol{i}}$	admittance associated to boundary node $\boldsymbol{i}$
$K_{\boldsymbol{i}}$	number of interior neighbours to $\boldsymbol{i}$
$\overline{K}_{\boldsymbol{i}}$	number of ghost-point neighbours to $\boldsymbol{i}$
$\tilde{Q}^n_{\boldsymbol{i}}$	summation over interior neighbours of $\underline{\Psi}^n_{\boldsymbol{i}}$
$v_\perp$	outward normal component of particle velocity
$\underline{v}^n_{w,\boldsymbol{i}}$	discrete outward normal velocity component, pointing outward with direction $w \in \pm\{x, y, z\}$
A	"inductance" coefficient of frequency-dependent admittance boundary condition, in s^{-1}
B	"resistance" coefficient of frequency-dependent admittance boundary condition
C	"inverse-capacitance" coefficient of frequency-dependent admittance boundary condition, in s
$A_{\boldsymbol{i},w\pm}$	projection of $A(\mathbf{x})$ onto Γ_h
$B_{\boldsymbol{i},w\pm}$	projection of $B(\mathbf{x})$ onto Γ_h
$C_{\boldsymbol{i},w\pm}$	projection of $C(\mathbf{x})$ onto Γ_h
$\mathfrak{h}_{\mathfrak{i}}$	stored discrete energy over interior
$\mathfrak{h}_{\mathfrak{b}}$	stored discrete energy over boundary
$\mathfrak{q}_{\mathfrak{b}}$	discrete power outflow through boundary
$A_{\boldsymbol{i}}$	projection of $A(\mathbf{x})$ onto $\mathbb{G}_h$
$B_{\boldsymbol{i}}$	projection of $B(\mathbf{x})$ onto $\mathbb{G}_h$
$C_{\boldsymbol{i}}$	projection of $C(\mathbf{x})$ onto $\mathbb{G}_h$
$\overline{\mathfrak{q}}_{\mathfrak{b}}$	accumulated discrete power outflow at boundary
D	"inductance"coefficient of frequency-dependent impedance boundary condition, in s^{-1}
E	"resistance"coefficient of frequency-dependent impedance boundary condition
F	"inverse-capacitance"coefficient of frequency-dependent impedance boundary condition, in s
$D_{\boldsymbol{i},w\pm}$	projection of $D(\mathbf{x})$ onto Γ_h
$E_{\boldsymbol{i},w\pm}$	projection of $E(\mathbf{x})$ onto Γ_h
$F_{\boldsymbol{i},w\pm}$	projection of $F(\mathbf{x})$ onto Γ_h
$D_{\boldsymbol{i}}$	projection of $D(\mathbf{x})$ onto $\mathbb{G}_h$
$E_{\boldsymbol{i}}$	projection of $E(\mathbf{x})$ onto $\mathbb{G}_h$
$F_{\boldsymbol{i}}$	projection of $F(\mathbf{x})$ onto $\mathbb{G}_h$
$\hat{\underline{v}}^{n-\frac{1}{2}}_{w,\boldsymbol{i}}$	auxiliary variable related to $\underline{v}^n_{w,\boldsymbol{i}}$
$\underline{g}^{n-\frac{1}{2}}_{w,\boldsymbol{i}}$	auxiliary variable related to $\underline{v}^n_{w,\boldsymbol{i}}$
η	length associated to viscothermal losses, in metres
η'	dimensionless parameter associated to viscothermal losses $\eta' = \eta/h$

$W_{\boldsymbol{i}}^n$	terms in viscothermal update that are additional to lossless update (free-space)
$\tilde{W}_{\boldsymbol{i}}^n$	terms in viscothermal update that are additional to lossless update (on closed domain)
l_v	characteristic length associated to viscous effects, in m
τ_η	relaxation time associated to viscothermal losses, in seconds
τ'_η	dimensionless parameter associated to viscothermal losses $\tau'_\eta = \tau_\eta / k$
$\overline{\mathrm{q}}_{\mathrm{i}}$	accumulated discrete power outflow over interior

List of Symbols for Chapter 4

d	number of spatial dimensions
u	scalar field (solution to wave equation) in $\mathbb{R}^d$, $u = u(\mathbf{x}, t)$
x, y, z	spatial coordinates, in m
$\mathbf{x}$	spatial position vector, $\mathbf{x} = x$ when $d = 1$, $\mathbf{x} = (x, y)$ when $d = 2$, $\mathbf{x} = (x, y, z)$ when $d = 3$
c	wave speed, in m/s
∂_t	partial derivative w.r.t. time
$\partial_x, \partial_y, \partial_y$	partial derivative w.r.t. spatial coordinates x, y, z
Δ	Laplace operator (Laplacian), $\Delta = \partial_x^2$ when $d = 1$, $\Delta = \partial_x^2 + \partial_y^2$ when $d = 2$, and $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$ when $d = 3$
f_0, g_0	initial conditions for d-D wave equation
$\Box$	wave operator (d'Alembertian)
F_s	sample rate, in Hz
k	time-step, in s
h	Cartesian (d-D) and hexagonal (2-D) grid spacing, in m
n	discrete time index
$\hat{\boldsymbol{\beta}}$	wave direction
$\beta_x, \beta_y, \beta_z$	wave vector components, in rad/m
$\boldsymbol{\beta}$	wave vector, $\boldsymbol{\beta} = \beta$ when $d = 1$, $\boldsymbol{\beta} = (\beta_x, \beta_y)$ when $d = 3$, $\boldsymbol{\beta} = (\beta_x, \beta_y, \beta_z)$ when $d = 3$
$\boldsymbol{\beta}_h$	normalised wave vector, $\boldsymbol{\beta}_h = \boldsymbol{\beta} h$
$\mathbb{B}$	Cartesian wavenumber cell
β_h	normalised wavenumber
ω	angular frequency, in rad/s
ω_k	normalised angular frequency
λ	Courant number, $\lambda = ck/h$
$\lambda_{\max}$	maximum Courant number for stability of a given scheme
$\beta_{\max}$	maximum wavenumber of interest, in rad/m
$\beta_{h,\max}$	maximum wavenumber of interest (normalised)
$f_{\max}$	maximum frequency of interest, in Hz
ρ_{SD}	spatial density of scheme
ρ_{STD}	spatiotemporal density of scheme
$\mathbb{R}^+$	non-negative real numbers
$\underline{u}_i^n$	1-D spatiotemporal grid function
$\underline{u}_{\boldsymbol{i}}^n$	spatiotemporal grid function, $d > 1$
$\underline{u}$	continuous approximation to u
$\delta_\Box$	discrete d'Alembertian

δ_{tt}	second-order spatial difference in time
δ_Δ	discrete Laplacian
δ_Δ^h	normalised discrete Laplacian ($\delta_\Delta^h = h^2\delta_\Delta$)
$\hat{\mathbf{e}}_x, \hat{\mathbf{e}}_y, \hat{\mathbf{e}}_z$	standard unit vectors
K	number of neighbouring spatial points involved in stencil of a discrete Laplacian
$\boldsymbol{v}_l$	stencil vectors in discrete Laplacian ($l = 0, \ldots, K$)
ξ_l	stencil weights of discrete Laplacian ($l = 0, \ldots, K$)
G	discrete Laplacian in spatial frequency domain (and scaled by $-h^2/4$)
$\omega_{h/c}$	rescaled normalised angular frequency ($\omega_{h/c} = \omega_k/\lambda$)
$\tilde{v}_\mathrm{p}$	relative phase velocity of scheme
$\varepsilon_{\text{rel. err.}}$	relative error in phase velocity
$\beta_{h,X\%}$	critical wavenumber, under which $\lvert\varepsilon_{\text{rel. err.}}(\boldsymbol{\beta}_h)\rvert \leq X\%$
μ	grid density when $h = 1$ ($\mu \neq 1$ for non-Cartesian grids)
$\beta_{h^\star,X\%}$	critical wavenumber, normalised for spatial grid densities
$\beta_{h^*,X\%}$	critical wavenumber, normalised for spatiotemporal grid densities
$\beta_{h^*,X\%,\text{ref.}}$	normalised critical wavenumber for reference scheme
RCE-$X\%$	relative computational efficiency (in meeting $X\%$ absolute error threshold)
RCE-$X\%$-opts	relative computational efficiency, rescaled for required FLOP instructions
$e_{t\pm}$	temporal shifts
$e_{x\pm}, e_{y\pm}, e_{z\pm}$	spatial shifts along coordinate axes
p	a positive integer
$\delta_{xx}^{(p)}$	three-point 1-D discrete Laplacian, reaching out to neighbours p steps away
α_p	weighting parameters applied to $\delta_{xx}^{(p)}$
$\delta_{xx}^{[M]}$	$2M$th-order accurate, $(2M+1)$-point 1-D discrete Laplacian
$\lambda_{\max,M}$	maximum Courant number for stability of $(2dM+1)$-point"leggy" schemes
$\alpha_{M,m}$	weighting parameters applied to $\delta_{xx}^{(p)}$ to construct $\delta_{xx}^{[M]}$ ($m = 1, \ldots, M$)
$\boldsymbol{\alpha}$	vector of weights for parametrised discrete Laplacians (e.g., $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_P)$)
$\mu_{b,x}, \mu_{b,y}, \mu_{b,z}$	centered spatial averages along x, y, z coordinate axes, reaching out b steps
$\boldsymbol{q}$	vector representing integer pairs or triplets
$\mathbb{Q}$	constrained set of integer pairs or triplets
$\delta_{\Delta,\boldsymbol{q}}$	2-D or 3-D discrete Laplacian, parametrised by integer pair or triplet $\boldsymbol{q}$
$Q_{\boldsymbol{i},\boldsymbol{q}}^n$	sum of neighbours for $\delta_{\Delta,\boldsymbol{q}}\underline{u}_{\boldsymbol{i}}^n$
Υ	set of P integer pairs or triplets $\boldsymbol{q}_p$
$\delta_{\Delta,\Upsilon}$	general parametrised 2-D or 3-D discrete Laplacian
$\delta_{\Delta,\Upsilon_M^+}$	$2M$th-order accurate $(2dM+1)$-point "leggy" discrete Laplacians
$\lambda_{1\%,M}$	Courant number that maximises $\beta_{h,1\%}$ in $(2dM+1)$-point leggy schemes
$\delta_{\Delta,\Upsilon_9}$	compact discrete Laplacian on 2-D Cartesian grid, using up to nine points
$\hat{\boldsymbol{\beta}}$	wave direction
$\omega_{k,\mathrm{d}}(\hat{\boldsymbol{\beta}})$	directional cutoff frequency
$\omega_{k,\mathrm{m}}$	minimum cutoff frequency
$\omega_{k,\mathrm{g}}$	global cutoff frequency
h'	grid spacing (inter-nodal distance) for rotated Cartesian grid ($d = 2$)
h'	grid spacing (inter-nodal distance) for non-Cartesian grids ($d = 3$)
λ'	rescaled Courant number for rotated Cartesian grid ($d = 2$), $\lambda' = ck/h'$
λ'	rescaled Courant number for non-Cartesian grids ($d = 3$), $\lambda' = ck/h'$

$\mathbb{G}_h^{(\mathrm{H})}$	hexagonal grid
$\mathbf{V}_{\mathrm{H}}$	generator matrix for hexagonal grid
x', y', z'	redundant axes for hexagonal grid system
$\hat{\mathbf{e}}_{x'}, \hat{\mathbf{e}}_{y'}, \hat{\mathbf{e}}_{z'}$	unit vectors along x', y', z', for hexagonal grid system
$\boldsymbol{m}$	hexagonal grid index vector, $\boldsymbol{m} = (m_{x'}, m_{y'})^{\mathrm{T}}$
$\underline{u}^n_{\boldsymbol{m}}$	hexagonal grid function
$s_{x'\pm}, s_{y'\pm}, s_{y'\pm}$	shift operators on hexagonal grid
$\partial_{x'}, \partial_{y'}, \partial_{z'}$	partial derivative w.r.t. spatial coordinates x', y', z' (for hexagonal grid system)
$\mathbb{B}_{\mathrm{H}}$	wavenumber cell for hexagonal grid
$\delta_{x'x'}, \delta_{y'y'}, \delta_{z'z'}$	second-order spatial differences along x', y', z' axes
$\delta_{\Delta}^{(\mathrm{H})}$	simplest seven-point discrete Laplacian on hexagonal grid
$\mu_{(a,b),(x',y')}$, etc.	spatial averaging operator for hexagonal grid system
$\delta_{\Delta,\boldsymbol{q}}^{(\mathrm{H})}$	discrete Laplacian on hexagonal grid, parametrised by integer pair $\boldsymbol{q} = (q_1, q_2)$
$\delta_{\Delta,\Upsilon}^{(\mathrm{H})}$	general parametrised discrete Laplacian on hexagonal grid
$Q^n_{\boldsymbol{m},\boldsymbol{q}}$	sum of neighbours for $\delta_{\Delta,\boldsymbol{q}}^{(\mathrm{H})} \underline{u}^n_{\boldsymbol{m}}$
$\delta_{\Delta,\Upsilon_{13}}^{(\mathrm{H})}$	compact discrete Laplacian on 2-D hexagonal grid, using up to 13 points
$\delta_{\Delta,\Upsilon_{27}}$	compact discrete Laplacian on 3-D Cartesian grid, using up to 27 points
β_{yz}	side-diagonal wave vector component, along $(\hat{\mathbf{e}}_y + \hat{\mathbf{e}}_z)\sqrt{2}$, in rad/m
$\mathcal{I}^{(\mathrm{f})}$	index set for FCC grid
$\mathcal{I}^{(\mathrm{b})}$	index set for BCC grid
$\mathbb{G}_h^{(\mathrm{f})}$	FCC grid
$\mathbb{G}_h^{(\mathrm{f})}$	BCC grid
$\mathbb{B}_{\mathrm{f}}$	wavenumber cell for FCC grid
$\mathbb{B}_{\mathrm{b}}$	wavenumber cell for BCC grid
$\mathbb{Q}_{\mathrm{f}}$	constrained set of integer triplets for FCC grid
$\mathbb{Q}_{\mathrm{b}}$	constrained set of integer triplets for BCC grid
$\delta_{\Delta,\Upsilon_{\mathrm{f},43}}$	compact discrete Laplacian on FCC grid, using up to 43 points
$\delta_{\Delta,\Upsilon_{\mathrm{f},27}}$	compact discrete Laplacian on FCC grid, using up to 27 points
$\boldsymbol{\alpha}_{\mathrm{f}}$	vector of three weighting parameters within $\delta_{\Delta,\Upsilon_{\mathrm{f},43}}$
$\boldsymbol{\alpha}_{\mathrm{b}}$	vector of three weighting parameters within $\delta_{\Delta,\Upsilon_{\mathrm{b},27}}$
μ_{tt}	temporal averaging operators (centered)
ξ, θ	parameters for implicit schemes
C	constant in Taylor expansion of a discrete Laplacian
$\xi^\star$	optimal coefficient for fourth-order implicit scheme
$\xi_{1\%\text{-opt.}}$	optimised coefficient for maximising $\beta_{h,1\%}$

List of Symbols for Chapter 5

d	number of spatial dimensions
u	scalar field (solution to wave equation) in $\mathbb{R}^d$, $u = u(\mathbf{x}, t)$
$x_1, x_2, \ldots, x_d$	spatial coordinates in $\mathbb{R}^d$
∂_{x_j}	partial derivative w.r.t. spatial coordinates $x_j, j = 1, \ldots, d$
$\overline{\Omega}$	closed domain, $\overline{\Omega} = \Omega \cup \Gamma$
Ω	interior domain

Γ	boundary of Ω
Δ	Laplacian in d dimensions
f_0, g_0	initial conditions for d-D wave equation
$\hat{\mathbf{n}}$	outward normal vector at point on boundary
γ	specific surface admittance (dimensionless)
∇	d-D gradient operator, $\nabla = (\partial_{x_1}, \ldots, \partial_{x_d})^{\mathrm{T}}$
$\mathbb{Z}_h^d$	d-D Cartesian grid
c	wave speed, in m/s
k	time-step, in s
h	Cartesian grid spacing, in m
n	discrete time index
$\underline{u}_{\boldsymbol{i}}^n$	Cartesian grid function
$\boldsymbol{i}$	discrete spatial vector index, $\boldsymbol{i} = (i_1, \ldots, i_d)$
$e_{t\pm}$	temporal shifts
$e_{x_j\pm}$	spatial shifts along coordinate axis x_j
$\hat{\mathbf{e}}_j$	standard unit vectors in $\mathbb{R}^d$
δ_Δ	discrete Laplacian
δ_Δ^h	normalised discrete Laplacian ($\delta_\Delta^h = h^2\delta_\Delta$)
$Q_{\boldsymbol{i}}^n$	summation over $2d$ spatial neighbours of $\underline{u}_{\boldsymbol{i}}^n$
$\delta_{x_j\pm}$	forward/backward first-order spatial difference along x_j
$\delta_{x_j\cdot}$	centered first-order spatial difference along x_j
$\delta_{x_j x_j}$	centered second-order spatial difference along x_j
δ_{tt}	second-order spatial difference in time, $\delta_{tt} = (e_{t+} - 2 + e_{t-})/k^2$
λ	Courant number, $\lambda = ck/h$
$\lambda_{\max}$	maximum Courant number for stability
z	complex number
ω	angular frequency, in rad/s
$\boldsymbol{\beta}$	wave vector, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_d)$
β_j	wave vector components, in rad/m
β	wavenumber ($\beta = \lvert\boldsymbol{\beta}\rvert$), in rad/m
G	discrete Laplacian in spatial frequency domain (and scaled by $-h^2/4$)
$\widehat{\mathbf{A}}$	spectral amplification matrix
$\mathbb{G}_h$	Cartesian grid for closed domain, $\mathbb{G}_h = \mathbb{Z}_h^d \cap \overline{\Omega}$
M	number of grid points in $\mathbb{G}_h$
$\mathcal{I}$	index set of Cartesian grid points (in $\mathbb{G}_h$)
$\mathbf{u}^n$	$\underline{u}_{\boldsymbol{i}}^n$ in column-vector form (vectorisation)
$\boldsymbol{i}_m$	grid point $\boldsymbol{i}$ corresponding to mth element of $\mathbf{u}^n$
$\mathbf{I}$	$M \times M$ identity matrix
$\mathbf{L}$	$M \times M$ normalised finite difference Laplacian matrix
$\mathbf{B}$	$M \times M$ diagonal matrix associated to boundary losses
$\mathbf{A}$	one-step update matrix
$\mathbf{0}$	empty matrix (matrix of zeros)
$\mathbf{w}$	state vector
$\hat{\rho}(\mathbf{A})$	spectral radius of $\mathbf{A}$ (maximum eigenvalue of $\mathbf{A}$ in absolute value)
$\boldsymbol{\phi}$	eigenvector of $\mathbf{L}$

μ	eigenvalue of $\mathbf{L}$
$\sigma(\mathbf{L})$	set of eigenvalues (spectrum) of $\mathbf{L}$
$q_{\boldsymbol{i},j\pm}$	indicator function
$\overline{q}_{\boldsymbol{i},j\pm}$	complement of $q_{\boldsymbol{i},j\pm}$
$\mathcal{I}_{\mathrm{i}}$	index set of regular interior grid points
$\mathcal{I}_{\mathfrak{b}}$	index set of boundary points
$\overline{\Omega}_h$	continuous staircased region approximating Ω
$\delta_{t\cdot}$	first-order time difference (centered), $\delta_{t\cdot} = (e_{t+} - e_{t-})/(2k)$
$\mathfrak{h}$	stored discrete energy
$\mathfrak{v}$	potential discrete energy
$\mathfrak{t}$	kinetic discrete energy
$\mathfrak{b}$	discrete energetic term associated to boundary (e.g., power outflow)
$\mathfrak{q}$	power outflow
$\gamma_{\boldsymbol{i},j\pm}$	$\gamma(\mathbf{x})$ projected onto $(\boldsymbol{i} \pm \hat{\mathbf{e}}_j/2)h$
$K_{\boldsymbol{i}}$	number of interior neighbours to $\boldsymbol{i}$
$\overline{K}_{\boldsymbol{i}}$	number of ghost-point neighbours to $\boldsymbol{i}$
$\tilde{Q}^n_{\boldsymbol{i}}$	summation over $K_{\boldsymbol{i}} \leq 2d$ interior neighbours of $\underline{u}^n_{\boldsymbol{i}}$
$\mathcal{I}_{\mathrm{g}}$	index set of ghost points
$\overline{\mathfrak{h}}$	accumulated stored energy
$\overline{\mathfrak{b}}$	accumulated dissipated power at boundary
$\mathbf{L}_{m_1 m_2}$	element in m_1th row and m_2th column of matrix $\mathbf{L}$
$\mathbf{D}_{x_j\pm}$	matrix associated to first-order spatial finite difference
$\left[\mathbf{D}_{x_j\pm}\right]_{m_1 m_2}$	element in m_1th row and m_2th column of matrix $\mathbf{D}_{x_j\pm}$
$\mathbf{D}^{(\pm)}_{xx}$	matrix associated to δ_{xx} with non-centered Neumann conditions
$\underline{v}^n_{\boldsymbol{i}}$	arbitrary Cartesian grid function
$\mathbf{v}$	vectorisation of $\underline{v}^n_{\boldsymbol{i}}$
$w_{\boldsymbol{i}}$	scaling factor involved in modified inner product
$w_{\boldsymbol{i},j}$	scaling factor involved in modified inner product
$\dot{Q}^n_{\boldsymbol{i}}$	summation over $K_{\boldsymbol{i}} \leq 2d$ interior neighbours of $\underline{u}^n_{\boldsymbol{i}}$, weighted by $w_{\boldsymbol{i},j}$
$\mathbf{D}^{(\cdot)}_{xx}$	matrix associated to δ_{xx} with centered Neumann conditions
L_j	dimensions of d-D box-domain, in m
M_j	dimensions of d-D box-domain, in grid points
$\mathcal{V}_{\mathrm{box}}$	d-D box-domain
R	reflection coefficient
β_h	normalised wavenumber
ω_k	normalised angular frequency
ξ	specific surface impedance (dimensionless)
$\mathbf{R}_\theta$	rotation matrix
θ	rotation angle

List of Symbols for Chapter 6

d	number of spatial dimensions
Ψ	acoustic velocity potential, in m^2/s

$\overline{\Omega}$	closed domain, $\overline{\Omega} = \Omega \cup \Gamma$
Ω	interior domain
Γ	boundary of Ω
$\mathbf{v}$	particle vector velocity field in d dimensions, $\mathbf{v} = (v_1, \ldots, v_d)$, in m/s
p	pressure field in d dimensions, in Newtons per unit d-dimensional surface area
ρ	density of medium, in kg per unit d-dimensional volume
Δ	Laplacian in d dimensions
∇	d-D gradient operator
$\nabla\cdot$	d-D divergence operator
$\Omega_\triangle$	tiled domain (open)
$\Gamma_\triangle$	closure of $\Omega_\triangle$
$\overline{\Omega}_\triangle$	tiled domain (closed), $\overline{\Omega}_\triangle = \Omega \cup \Gamma_\triangle$
$\overline{\mathcal{C}}_i$	closed, bounded cell
$\mathcal{C}_i$	open interior of $\overline{\mathcal{C}}_i$
$\partial\mathcal{C}_i$	boundary of $\overline{\mathcal{C}}_i$
$\mathcal{S}_{i,j}$	cell face adjoining $\mathcal{C}_i$ and $\mathcal{C}_j$
$I^{(\mathrm{b})}$	index set for cells that touch boundary $\Gamma_\triangle$
$I^{(\mathrm{i})}$	index set for cells that do not touch boundary $\Gamma_\triangle$
N_i	neighbour set (of indices) for cell $\mathcal{C}_i$
$\mathcal{S}_i^{(\mathrm{b})}$	boundary face of cell $\mathcal{C}_i$
$\mathcal{S}_{i,l}^{(\mathrm{b})}$	boundary faces (subdivided) of cell $\mathcal{C}_i$
$N_i^{(\mathrm{b})}$	number of subdivided boundary faces of cell $\mathcal{C}_i$
$\mathbb{G}_{h,\triangle}$	grid of points associated to finite volume tiling
$\mathbf{x}_i$	grid of point associated to cell $\overline{\mathcal{C}}_i$
$h_{i,j}$	inter-cell distance between cells $\mathcal{C}_i$ and $\mathcal{C}_j$
V_i	d-dimensional volume of cell $\overline{\mathcal{C}}_i$
$S_{i,j}$	d-dimensional surface area of face $\mathcal{S}_{i,j}$
$S_i^{(\mathrm{b})}$	surface area of boundary face $\mathcal{S}_i^{(\mathrm{b})}$
$S_{i,l}^{(\mathrm{b})}$	surface area of boundary face $\mathcal{S}_{i,l}^{(\mathrm{b})}$
$\overline{\Omega}_h$	staircased region, approximating $\overline{\Omega}$
λ	Courant number, $\lambda = ck/h$
∂_t	partial derivative w.r.t. time
$\hat{\mathbf{n}}_i$	outward normal vector over the boundary of $\overline{\mathcal{C}}_i$
$\hat{\mathbf{n}}_{i,j}$	outward normal vector over the face $\mathcal{S}_{i,j}$
$v_{i,j}$	outward normal velocity component over face $\mathcal{S}_{i,j}$
$v_\perp$	outward normal component of particle velocity
$\hat{\mathbf{n}}$	outward normal vector at point on boundary of domain
$\mathrm{d}\mathbf{x}$	differential volume element in d-D
$\mathrm{d}\sigma$	differential surface element in d-D
$\underline{\Psi}_i^n$	cell-average approximation to Ψ over cell $\overline{\mathcal{C}}_i$ at time $t = nk$
$\underline{p}_i^{n+\frac{1}{2}}$	cell-average approximation to p over cell $\overline{\mathcal{C}}_i$ at time $t = (n + 1/2)k$
$\underline{v}_{i,j}^n$	face-average approximation to $v_{i,j}$ over face $\mathcal{S}_{i,j}$ at time $t = nk$
$\underline{v}_{\perp,i}^n$	face-average approximation to $v_\perp$ over $S_i^{(\mathrm{b})}$ at time $t = nk$
$\underline{v}_{\perp,i,l}^n$	face-average approximation to $v_\perp$ over $S_{i,l}^{(\mathrm{b})}$ at time $t = nk$
$\delta_{i,j}$	directional difference, operating across cells $\overline{\mathcal{C}}_i$ and $\overline{\mathcal{C}}_j$

$\delta_{t\pm}$	first-order time differences (forwards/backwards)
δ_{tt}	second-order spatial difference in time
$i \sim j$	$\overline{\mathcal{C}}_i$ adjacent to $\overline{\mathcal{C}}_j$
$\mathfrak{h}$	stored discrete energy
$\mathfrak{v}$	potential discrete energy
$\mathfrak{t}$	kinetic discrete energy
$\mathfrak{b}$	discrete energetic term associated to boundary (e.g., power outflow)
$\mathfrak{q}$	power outflow
μ_{tt}	temporal averaging operator (centered)
ξ, θ	parameters for implicit schemes
K_i	number of adjacent neighbours to cell $\overline{\mathcal{C}}_i$
$\mathcal{V}_{\text{box}}$	3-D box region, aligned with Cartesian axes (also $\mathcal{V}_{\text{box}}^{(\text{R0})}$)
$\mathcal{V}_{\text{box}}^{(\text{R1})}, \mathcal{V}_{\text{box}}^{(\text{R1})}$	rotated box domains
η	length associated to viscothermal losses, in metres
τ_η	relaxation time associated to viscothermal losses, in seconds
$\mu_{t\pm}$	temporal averaging operators (forwards/backwards)
$\mathfrak{q}_{\mathfrak{b}}$	discrete power outflow at boundary
$\mathfrak{q}_{\mathfrak{i}}$	discrete power outflow over the interior
$\overline{\mathfrak{q}}_{\mathfrak{b}}$	accumulated discrete power outflow at boundary
$\overline{\mathfrak{q}}_{\mathfrak{i}}$	accumulated discrete power outflow over the interior
D	"inductance"coefficient of frequency-dependent impedance boundary condition, in s^{-1}
E	"resistance"coefficient of frequency-dependent impedance boundary condition
F	"inverse-capacitance"coefficient of frequency-dependent impedance boundary condition, in s
$D_{i,l}, E_{i,l}, F_{i,l}$	boundary face–averaged counterparts of D, E, F
g	extra variable on boundary, ($\partial_t g = p$)
$\underline{\hat{v}}_{\perp,i,l}^{n-\frac{1}{2}}$	auxiliary variable related to $\underline{v}_{\perp,i,l}^{n}$
$\underline{g}_{i,l}^{n-\frac{1}{2}}$	auxiliary variable related to $\underline{v}_{\perp,i,l}^{n}$
$L^{(m)}$	"inductance" coefficient of mth branch of general impedance boundary condition, in $\text{kg/m}^{d\text{-}1}$
$R^{(m)}$	"resistance" coefficient of mth branch of general impedance boundary condition, in $\text{kg/(m}^{d\text{-}1}\text{s)}$
$C^{(m)}$	"capacitance" coefficient of mth branch of general impedance boundary condition, in $\text{m}^{d\text{-}1}\text{s}^2\text{/kg}$
$M^{(\text{b})}$	number of parallel series-RLC circuit branches