Toward three-dimensional computation of the acoustic energy-stress tensor

Aidan Meacham

Sorbonne Université

École doctorale 391 Sciences Mécaniques, Acoustique, Électronique et Robotique

Institut Jean Le Rond ∂'Alembert, Sorbonne Université, CNRS, UMR 7190

Toward Three-Dimensional Computation of the Acoustic Energy-Stress Tensor

---

Par **Aidan Meacham**

Thèse de doctorat d'Acoustique

Dirigée par Jean-Dominique POLACK, Professeur

et Roland BADEAU, Professeur

Présentée et soutenue publiquement le 09/12/2021

devant un jury composé de :

Alain LE BOT	Directeur de Recherche	Rapporteur
Peter SVENSSON	Professeur	Rapporteur
Thomas HÉLIE	Directeur de Recherche	Examinateur
Philippe DEPALLE	Professeur	Examinateur
Régis MARCHIANO	Professeur	Examinateur
Jean-Christophe VALIÈRE	Professeur	Examinateur
Jean-Dominique POLACK	Professeur	Co-directeur de Thèse
Roland BADEAU	Professeur	Co-directeur de Thèse

Abstract

The energy-stress tensor method is a theoretical approach to approximating the energetic contours of the "stochastic reverberation" using continuity equations between acoustical energy density, sound intensity, and momentum flux. These quantities may all be collected in a single tensor called the energy-stress tensor, which represents the connection between invariants in the acoustic field and the conservation equations relating its terms. This approach was previously demonstrated to be capable of representing the diffuse field in spaces that could be characterized as being one- or two-dimensional with respect to the characteristic length of the tensor quantities, such as a hallway or a floor of an office building. The present thesis aims to extend these findings by a number of avenues: checking the frequency validity of previously derived models, introducing a source term and redefining the model in terms of a finite volume time domain (FVTD) approach, auralizing the results in the context of a hybrid acoustical model, and finally, providing a framework to explore the behavior of the tensor in larger three-dimensional spaces. Temporally and spatially averaged values of the tensor may be calculated directly in terms of the acoustic velocity potential in a pressure-velocity FVTD simulation, giving preliminary information regarding possible assumptions for future models without requiring exhaustive physical measurement. This may be used in future work to characterize simulation parameters in terms of geometry and materials properties rather than physical measurements.

Résumé

La méthode du tenseur énergie impulsion est une approche théorique permettant d'approcher les contours énergétiques de la "réverbération stochastique" à l'aide d'équations de conservation basées sur le processus de diffusion acoustique. Dans ce formalisme, les termes termes décrivant la densité d'énergie, l'intensité sonore et le flux de quantité de mouvement peuvent tous être rassemblés en un seul tenseur appelé le tenseur énergie impulsion, qui fournit la base des équations de conservation capables de caractériser la réverbération stochastique. Cette approche, appelée "méthode du tenseur énergie impulsion" ou méthode EST (après l'anglais), s'est avérée capable de représenter le champ diffus dans des espaces où tous les éléments de simulation touchent au moins une limite, c'est-à-dire des espaces qui peuvent être caractérisés comme étant principalement unidimensionnels ou bidimensionnels, tels qu'un couloir ou un étage d'un immeuble de bureaux. La présente thèse vise à étendre ces résultats par un certain nombre de voies : vérifier la validité fréquentielle des modèles précédemment dérivés, introduire des termes sources et redéfinir le modèle en termes d'approche dans la méthode temporelle des volumes finis (FVTD), auraliser les résultats dans le contexte d'un modèle acoustique hybride, et enfin, fournir un cadre pour explorer le comportement du tenseur de contrainte d'onde dans les espaces tridimensionnels afin de fournir une base pour les équations de volume dans les espaces avec des régions en espace libre, loin de toute frontière qui pourrait être utilisée pour la réduction dimensionnelle. Le calcul des éléments du tenseur énergie impulsion partout dans un espace a été réalisé en dérivant les valeurs instantanées en termes de potentiel de vitesse acoustique pour une simulation FVTD pression-vitesse. Une fois moyennée dans le temps et dans l'espace, cette représentation peut être utilisée pour explorer le comportement de la réverbération stochastique sans mesure physique exhaustive dans des espaces présentant des caractéristiques de diffusion variables, et finalement, pour réaliser la simulation directe des termes du tenseur aux frontières et en espace libre. Cela pourrait être utilisé dans des travaux futurs pour caractériser les paramètres de simulation en termes de géométrie et de propriétés des matériaux plutôt que de mesures physiques.

Résumé Long

La méthode du tenseur énergie impulsion est une approche théorique permettant d'approcher les contours énergétiques de la "réverbération stochastique" à l'aide d'équations de conservation basées sur le processus de diffusion acoustique. La réverbération stochastique est la partie de la réponse impulsionnelle d'une salle qui se produit après le temps de mélange, lorsque l'oreille humaine ne peut plus discerner les réflexions individuelles, lorsque le champ acoustique est devenu diffus. Un champ diffus signifie que l'énergie sonore se propage de manière égale dans toutes les directions et qu'elle est d'intensité égale en tous points d'une salle. Cet état est présumé être un processus gaussien, ce qui signifie qu'il peut être synthétisé avec du bruit sans perte de qualité perceptive.

La diffusion est le processus acoustique facilité par la géométrie de la salle, par lequel l'énergie d'une source commence en champ libre, subit un nombre de réflexions quadratiquement croissant et devient un champ diffus. Des méthodes permettant de caractériser ce processus ont été utilisées pour modéliser la propagation de l'énergie sonore (à savoir, la partie diffuse) dans une salle avec des conditions aux limites basées sur les bilans énergétiques observés aux murs, et sont collectivement appelées "méthode de l'équation de diffusion" ou DEM. La distribution de la densité d'énergie dans la salle, variable dans le temps et dépendant de la fréquence, correspond à la réverbération stochastique dans chaque bande de fréquence, de sorte qu'elle peut être synthétisée par l'application d'une enveloppe à un bruit à bande limitée.

L'un des inconvénients de ces approches est que, bien que les hypothèses de bilan énergétique utilisées dans leur construction garantissent la conservation de l'énergie, il n'en va pas de même pour la conservation de la quantité de mouvement, ce qui nécessite des modifications heuristiques pour obtenir des résultats de simulation précis. En acoustique, le terme qui représente la quantité de mouvement des particules est l'intensité sonore, ou flux de densité d'énergie. Lorsqu'il est examiné instantanément, ce champ vectoriel peut être traité comme un indicateur de la direction d'arrivée des ondes, mais il communique également des informations sur le transfert d'énergie diffuse entre les sections d'une salle. Assurer la conservation de l'intensité sonore nécessite l'introduction d'un terme tensoriel appelé tenseur de contrainte d'onde dans les équations du champ constitutif. Dans le cadre

du DEM, l'intensité sonore variée lentement dans le temps, ne variant qu'en fonction de l'emplacement parce qu'elle fluctue sur une échelle de temps beaucoup plus courte que la densité d'énergie, ce qui signifie que le tenseur de contrainte d'onde est diagonal avec les trois composantes égales, dont la somme correspond à la densité d'énergie globale. Dans certains cas, cependant, cette hypothèse n'est pas valable, et il faut alors considérer le tenseur de contrainte d'onde complet. Ces termes décrivant la densité d'énergie, l'intensité sonore et le flux de quantité de mouvement peuvent tous être rassemblés en un seul tenseur appelé le tenseur énergie impulsion, dont la conservation est une illustration du théorème de Noether pour l'acoustique qui exprime la conservation de la densité d'énergie et l'intensité sonore.

Les équations de conservation résultantes constituent un système qui peut être utilisé pour simuler la réverbération stochastique avec une hypothèse supplémentaire d'équilibre du quantité de mouvement aux frontières du domaine sans correction heuristique. Cette approche, appelée "méthode du tenseur énergie impulsion" ou méthode EST (d'après l'anglais), s'est avérée capable de représenter le champ diffus dans des espaces où tous les éléments de simulation touchent au moins une limite, c'est-à-dire des espaces qui peuvent être caractérisés comme étant principalement unidimensionnels ou bidimensionnels, tels qu'un couloir ou un étage d'un immeuble de bureaux. Afin de simplifier le calcul des éléments hors diagonale du tenseur de contrainte d'onde, la dimensionnalité de ces exemples d'espaces a été utilisée pour contraindre les équations régissant le transfert d'énergie volumétrique. La similarité des équations de conservation ainsi modifiées permis une réduction finale à une forme équivalente aux équations du télégraphe qui peuvent être résolues numériquement.

La présente thèse vise à étendre ces résultats par un certain nombre de voies : vérifier la validité fréquentielle des modèles précédemment dérivés, introduire des termes sources et redéfinir le modèle en termes d'approche par la méthode temporelle des volumes finis (FVTD), auraliser les résultats dans le contexte d'un modèle acoustique hybride, et enfin, fournir un cadre pour explorer le comportement du tenseur de contrainte d'onde dans les espaces tridimensionnels afin de fournir une base pour les équations de volume dans les espaces avec des régions en espace libre, loin de toute frontière qui pourrait être utilisée pour la réduction dimensionnelle.

La réverbération stochastique est supposée avoir lieu au-dessus de la fréquence de Schroeder, en dessous de laquelle le comportement modal est dominant, même à des échelles de temps longues. Les résultats concernant les régions de fréquence pouvant être représentées par la méthode EST en témoignent, et soulignent l'importance de modéliser un espace présentant des caractéristiques de diffusion et de dispersion suffisantes.

L'introduction de termes de source dans le modèle permet l'injection variable dans le temps d'énergie dans le champ diffus en termes de vitesse volumique connue, ce qui permet, par exemple, de représenter des systèmes qui incluent une rétroaction. Ces développements ont été vérifiés avec la même procédure que l'étude de validité ci-dessus, démontrant la

mise en œuvre d'une source dipôle et la représentabilité de la réverbération stochastique résultante. En outre, la méthode EST bénéficie d'une recontextualisation dans le formalisme FVTD, qui est fondé sur la conservation du flux à travers les surfaces de cellules adjacentes, fournissant un mécanisme naturel pour examiner l'intensité sonore entre les éléments du volume. Le développement de ce schéma à la fois pour la méthode EST et pour la modélisation traditionnelle de la pression et de la vitesse était important à la fois pour le schéma d'auralisation et pour la caractérisation des termes du tenseur de contrainte d'onde à partir des champs de pression.

Comme nous l'avons déjà mentionné, la réverbération stochastique peut être synthétisée par l'application d'enveloppes énergétiques variant dans le temps à un bruit à bande limitée ; cependant, il ne s'agit que d'une partie de la réponse impulsionnelle complète d'une salle. Pour démontrer l'applicabilité de la méthode EST dans un contexte de réverbération en temps réel, un modèle hybride composé de la réverbération stochastique, des réflexions précoces de la méthode des sources images et du comportement modal à basse fréquence d'une simulation FVTD pression-vitesse a été assemblé en un seul schéma d'auralisation. L'étalonnage du niveau d'énergie entre les méthodes dans les bandes de temps et de fréquence a été examiné, ainsi que les autres avantages de l'utilisation d'un schéma volumétrique pour la réverbération stochastique plutôt que des réverbérateurs non physiques ou d'autres méthodes de calcul des enveloppes d'énergie dépendant de la fréquence à partir de méthodes stochastiques telles que le traçage de rayons.

Enfin, le calcul des éléments du tenseur énergie impulsion partout dans un espace a été réalisé en dérivant les valeurs instantanées en termes de potentiel de vitesse acoustique pour une simulation FVTD pression-vitesse. Une fois moyennée dans le temps et dans l'espace, cette représentation peut être utilisée pour explorer le comportement de la réverbération stochastique sans mesure physique exhaustive dans des espaces présentant des caractéristiques de diffusion variables, et finalement, pour réaliser la simulation directe des termes du tenseur aux frontières et en espace libre. En outre, le pavage de l'espace de Riemann est exploré comme un moyen d'expliquer géométriquement le processus de diffusion qui se produit dans les salles polyédriques. Cela pourrait être utilisé dans des travaux futurs pour caractériser les paramètres de simulation en termes de géométrie et de propriétés des matériaux plutôt que de mesures physiques.

Acknowledgments

This thesis would not have been possible without the support and encouragement of a globe-spanning network of mentors, teachers, colleagues, friends, and family, all of whom played a role in my development as a researcher and as a person.

To begin, I must thank Lauri Savioja, Sara Martín Román, and Julius O. Smith III for introducing me to room acoustics during my master's degree. They, along with the CCRMA faculty and staff, have educated and shaped me into the acoustician that I am today. Additional thanks to Lauri and Sara are due for their guidance regarding the pursuit of doctoral research, as well as their openness and support during my studies. Without them, I certainly could not have accomplished what I have to date.

Excellent mentorship is a rare and valuable thing, and I am lucky to have had received the best from Andy Unruh during my time at Knowles Electronics / Audience Inc. Beyond his guidance as a manager, Andy encouraged me to improve myself and to seek excellence in all that life has to offer, be it in acoustics and signal processing, sports and fitness, or adventure both at home and abroad. Further thanks go to my colleagues who are too numerous to name, but especially to those who inspired me to pursue a PhD, including Pratik, Ram, Mathieu, Meng, Ki Won, and Shandor.

I would like to thank my reviewers, Alain Le Bot and Peter Svensson, for all of their hard work in helping me to make this manuscript what it is today. I am indebted to your assistance and good grace, and cannot thank you enough for giving me the opportunity to present my research in its best possible light. Additionally, thank you to all of the members of the jury, especially to those traveling from afar, for navigating the complications of a thesis defense amidst an uncertain world. I am fortunate to have had the occasion to work alongside some of you, and to have benefited from all of your academic work which guided and informed my own research. Finally, I would like to acknowledge Régis Marchiano for his effort in leading my second and third year doctoral progress committees, which were invaluable in clarifying my academic and personal goals, and always brought new light onto our research problems.

Je me tourne ensuite vers ceux à qui je dois m'adresser en français, en commençant, bien sûr, par mes collègues de l'Institut Jean le Rond ∂'Alembert. Merci à tous d'avoir été

si accueillants et d'avoir partagé tant de bons moments ensemble. J'aurais aimé que nous ayons plus de temps à passer ensemble, mais nous en aurons davantage à l'avenir, où que ce soit dans le monde. Je tiens à remercier tout particulièrement l'équipe 317, dont je suis très heureux d'avoir fait la connaissance des doctorants actuels et anciens : Alexis, Alice, Antoine L, Antoine W, Antonio, Cécile, Clément, Franck, Jeanne, Manuel, Toufik, et Yann. Je pourrais évidemment écrire un livre ici, mais parfois il faut se contenter de dire merci. Aux américains, dans le laboratoire et ailleurs, notamment Debak, Elliot, Kelly, Nolan, Sam, et Sarabeth, quelle chance nous avons de nous retrouver ici. Je suis vraiment reconnaissant d'avoir (re-)croisé le chemin de chacun d'entre vous, et j'attends avec impatience la prochaine rencontre. Je dois faire une mention spéciale pour Lucille, qui défie toutes les catégories ci-dessus, et m'a fait vraiment sentir chez moi en France. Finalement, Ikram, merci pour ton amitié et ta compagnie tout au long de ce voyage, and Serena, thank you for your uncompromising spirit that sustained us all. Pour tous les autres que je ne peux pas citer ici mais que je ne peux pas oublier, votre esprit de convivialité, au bureau et à Paris, a été un plus dans les meilleurs moments et essentiel dans les plus difficiles, et pour cela, je vous remercie tous. Je n'aurais vraiment pas pu le faire sans vous.

Une grande partie de ce qui fait de la science un défi qui vaut la peine d'être relevé est la communication, pour laquelle je suis si heureux d'avoir eu Roland Badeau comme exemple. Pour ta perspicacité et ta patience en tant que directeur, et de m'avoir tant aidé au cours de ce doctorat, merci.

Enfin, je suis reconnaissant d'avoir eu la chance d'avoir Jean-Dominique Polack comme directeur pour ce projet. J'ai tellement appris de toi, et sans aucun doute, je continuerai à apprendre du travail que nous avons fait. De m'avoir accueilli et avoir relevé tous les défis de la thèse et les complications de la vie, de m'avoir guidé dans des recherches utiles et intéressantes, et de m'avoir donné une nouvelle perspective sur ma carrière et mes passions, je ne saurais jamais assez te remercier.

To my family, who always encouraged my curiosity, musicianship, and the pursuit of a dream, thank you. As my studies have gone further and further afield, your unwavering confidence has helped to keep me grounded wherever in the world I may be.

Finally, to Emma. Thank you for your flexibility, support, and for all of the fun that we've had along the way. I could not have asked for a better partner in writing this thesis and in life. Here's to the next adventure.

Notation

$\underline{\underline{T}}$	Energy-stress tensor
E	Energy density
$\underline{\underline{E}}$	Wave-stress tensor
$\mathbf{I}, \mathbf{J}$	Sound intensity
$\partial_i, \partial_{ii}$	First & second partial derivatives of coordinate i
Ψ	Velocity potential
$\mathbf{v}$	Particle velocity vector
p	Sound pressure
∇	Gradient operator
Δ	Laplacian operator
ρ	Air density
c	Speed of sound
$\mathcal{Q}$	Source volume velocity

Acronyms

DOA	Direction of Arrival
EST	Energy-stress tensor
DEM	Diffusion equation method
FVTD	Finite volume time domain
FDTD	Finite difference time domain

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The acts of analyzing, simulating, and auralizing room acoustics have long been segmented into domains in time and frequency where distinct phenomena take place, despite our understanding of a unifying underlying physical model. This can be attributed to the necessity of approximating particular phenomena given computational constraints, or as a result of psychoacoustic evidence relaxing the precision required to model a specific acoustic behavior with enough detail to satisfy the human auditory system. What may be an effective approach to modeling acoustical processes in one domain may not necessarily be for another. Often, a modeling assumption that leads to a simplification in one domain may be computationally untenable when extended to account for the entirety of the acoustic behavior in a space, delineating the boundary where a different model may be more appropriate.

In considering a typical room impulse response, one such division occurs temporally between the region dominated by distinct arrivals of wavefronts (including both the direct path and early reflections) and later reflections that are so numerous as to be indistinguishable. The earlier, temporally distinct events are efficiently modeled by geometric approaches to acoustics that treat propagating sound in an optical fashion. Rather than modeling the spatial progression of an entire wavefront emanating from a source through a fluid medium, geometrical acoustics approximates a portion of that wavefront as a quasi-plane wave traveling along a ray that undergoes reflection in a specular fashion. While we consider the wavefront to be sufficiently far from the source in order to approximate it with a planar surface, we must still take into account the geometrical attenuation based on the source radiation. This approach is ideally suited to representing the discrete specular events that make up the most salient acoustic content at the beginning of an impulse response, and is reviewed in detail in Savioja and Svensson (2015). The so-called late reflections, which become more and more numerous as time goes on, approach a diffuse field in many rooms, and while geometric methods are capable of reproducing such a field, the basis of those methods

are in representing expanding wave fronts individually, meaning a proliferation of reflections requires a corresponding expansion in the computational resources to model them. In fact, Monte Carlo ray tracing approaches have been shown to be functionally equivalent to energy models as in Le Bot and Bocquillet (2000), but there remain compelling reasons to choose non-geometric approaches, especially for large or coupled room volumes. Depending on definitions, the point in time where the soundfield is characterized as becoming diffuse, meaning it may be modeled stochastically, is typically named the "mixing time" or the "transition time." In this thesis, we consider mixing to be an asymptotic property of particles in a geometric space fulfilling certain requirements, as discussed below, and therefore will use the latter term to describe the moment of change from distinct to diffuse.

Similarly, in the frequency domain, the line is often drawn at Schroeder's frequency, below which distinct modes may be observed, and may be dominant in the overall room response. As in the temporal case, geometrical approaches are capable of modeling this behavior; however, it requires a great number of individual wavefronts to be simulated in order to approach the modal distribution in space. Modeling this constructive and destructive interference in such a piecemeal fashion can be more computationally demanding than a volumetric simulation in the same frequency range. Conversely, while modal excitation models are very efficient below Schroeder's frequency, even for very large spaces, as the number of modes grows, the superposition principle leads to the same computational challenges as geometric models encounter with the growth in echo density.

Comparable lines are often drawn in terms of diffraction effects as well. In most cases, the direct path can be modeled satisfactorily using only distance and taking losses from air absorption into account, but the presence of any interfering object (or even a complete lack of a visual path between source and receiver) introduces frequency-dependent diffraction based on the geometry. This behavior, which is most prominent in the distinct-arrival, low-frequency region, can be handled via extensions to geometrical approaches, and is modeled implicitly in wave-based schemes, meaning that it does not require special consideration. Nonetheless, choosing an approach that excels in this region means that high frequency behavior must be addressed some other way, either by hybridization with geometric techniques above some cutoff frequency and the ensuing complications regarding equalization and alignment between the two techniques, or by taking on the additional computational load of a wideband wave-based approach, even if the domain may be artificially limited in time or space.

As is probably apparent, even though direct simulation of wave phenomena appears to be a consistent solution to the simplifications proposed above, even that approach comes with challenges of both a theoretical and computational nature. For some simple rooms, it is plausible to derive an analytic solution to the acoustic wave equation, but even small changes in boundary conditions or geometry can put closed form solutions out of reach, illustrating

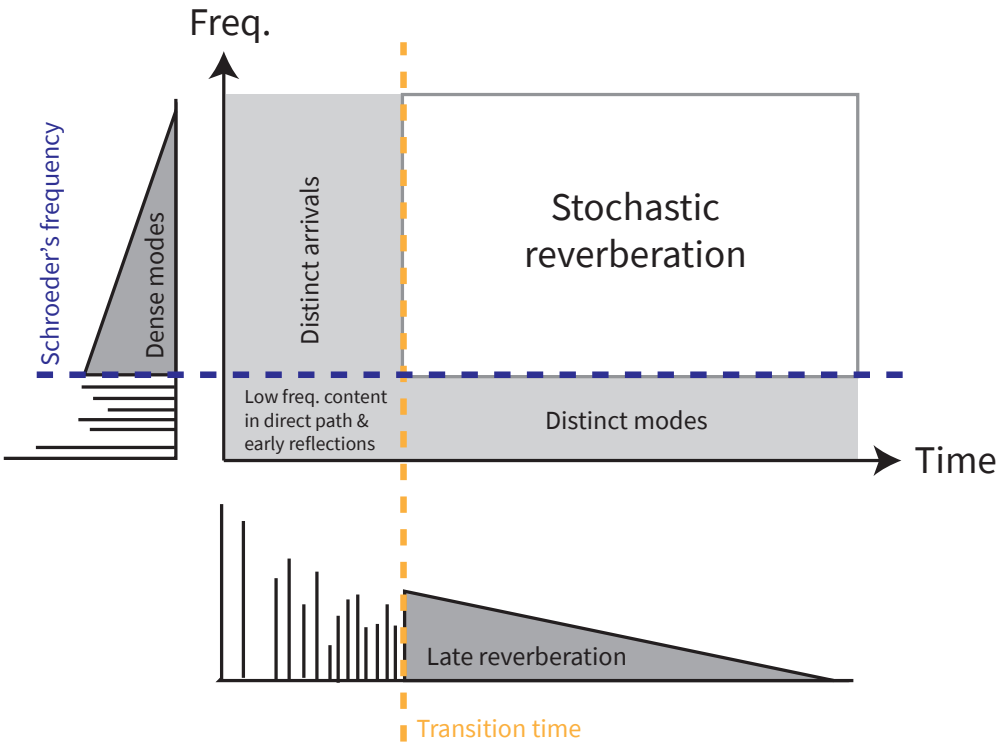

Figure 1.1: Validity domain for stochastic reverberation, after Jot et al. (1997) and Badeau (2019)

the chaotic nature of the underlying physical processes. Accurately representing boundary conditions is also difficult in numerical simulations of the wave equation, whether performed in the frequency or time domain, and since stability of these schemes is often predicated on a mesh size that decreases with the highest frequency to be modeled, they often suffer from extreme computational requirements for even modest sampling rates. Wideband audio rates have only recently become possible with the advent of schemes optimized for parallelization on consumer graphics cards, and while the results in terms of accuracy are exciting, real-time simulation remains out of reach for the time beingLai and Hamilton (2020).

While these delineations appear at first glance to be limitations, they are rather a testament to our shared understanding of the individual phenomena that play important roles in our perception of the acoustics in a particular room. This is perhaps best evidenced by the popularity of hybrid models that combine the best parts of each of these approaches to achieve rapid and accurate simulation of even very complicated spaces.

1.1 Reverberation

Nonetheless, the domain commonly referred to as the "stochastic reverberation," defined as the portion of an impulse response that can be modeled by a stochastic process, remains difficult to predict without performing exhaustive simulations, regardless of approach. This portion of the impulse response is also commonly referred to as the "late reverberation," in contrast to the early reflections that provide most of the spatial information for a listener. In the context of this thesis, however, we will use the term late reverberation in the sense given by Figure 1.1, where it represents the portion of an impulse response falling after the transition time, regardless of frequency. The stochastic reverberation, while perhaps less critical in terms of spatialization, is still an important part of the overall coloration of a particular impulse response, and especially for rooms with long reverberation times, if modeled poorly, can sound artificial, as described in Schroeder et al. (1962). Recent mathematical descriptions of this phenomena include Polack et al. (2019) and Badeau (2019).

The statistical nature of the stochastic reverberation can be thought of as a result of the increasing echo density throughout time, which grows at least quadratically as noted by Kuttruff (2016), and once exceeding the threshold of audibility, results in a perceptually diffuse field. Alternatively, one may consider the threshold of discretization, the point at which there is more than one echo per time sample. Characterization of the transition time in terms of perceptual, theoretical, or signal processing-based thresholds is a subject of ongoing research and is covered in greater detail in Chapter 2. To reach a particular threshold, in any case, a space must be ergodic: that is, a particular wavefront (perhaps as described by a ray) must spend, on average, equal time at any position in any direction within the room. Related is the concept of a space that is mixing, meaning that initially adjacent wavefronts

observed at a much later time will be completely separated, and spread out over the entire room. While there are theoretical spaces that do not fulfill these conditions, including regular polyhedral rooms, which only permit a discrete number of directions of travel from an initial condition, in practice, imperfections in the construction of real rooms, as well as the presence of other interfering objects, means that most non-theoretical spaces have these characteristics. The time scale at which these qualities become apparent, however, depends on the amount of imperfection or scattering material in the room, and is closely related to the measure of transition time mentioned above. These criteria in the context of stochastic reverberation are covered in great detail in Polack (1992) and Polack (1993).

The advantage when it comes to modeling this region is that a description of the field's statistical characteristics is mathematically interchangeable with the field itself. Since the human auditory system is incapable of perceiving differences in the chaotically-changing nature of a room impulse response from one moment to the next, whether as a result of temperature fluctuations or subtle changes in listening position, so too should we feel comfortable exchanging one essentially random sequence for another with the same statistics. In practice, this means that matching the acoustical indices of a particular noise sequence to measured data is a very practical way of modeling the stochastic reverberation without having to perform a simulation with geometrical or wave-based models.

This approach has long been used in room acoustics, from simply applying known frequency-dependent exponential decays to noise in a filterbank-style implementation, as in Moorer (1979), to collecting information from a stochastic ray tracing simulation to modify a noise sequence of increasing echo density, as in Schröder and Vorländer (2011). Of course, the validity of this region is limited, so it is almost exclusively used as one module in a hybrid approach, regardless of how the direct path and early reflections are modeled. In both of these cases, however, certain information must already be known about the space being modeled: while they are appropriate for real-time reproduction of the acoustic field, they rely on either preexisting measurements of a space (which could limit the source-receiver pairs available) or on performing sufficiently-long geometrical acoustics simulations (which may not account for some wave phenomena).

As previously described, since stochastic reverberation is defined to take place above Schroeder's frequency, wideband wave-based simulations that cover this region are unfortunately not yet feasible at real-time simulation update rates. Even though they would obviate the need for noise modeling entirely, in terms of computational efficiency, using a wave-based method in such a fashion could still be seen as wasteful if an alternative method for discovering the statistical properties of the stochastic reverberation were available. Finally, as mentioned before, due to the chaotic nature of the stochastic reverberation, small changes in boundary conditions, atmospheric changes, or perturbations of source and receiver positions can lead to differences in the fine structure of an impulse response that are apparent upon

visual inspection, but are ultimately inaudible. An exhaustive but deterministic modeling approach may not reflect this reality, giving a false impression of accuracy or consistency over time that does not exist.

1.2 Energy-based methods

The focus of this thesis is on directly predicting the characteristics of the stochastic reverberation using an energy-based approach. Energetic methods, which eschew direct simulation of acoustic pressure waves in lieu of modeling the flow of sound energy, have primarily seen use in noise-related applications, as predicting noise levels does not require the degree of resolution that high-fidelity room acoustical modeling does.

The advantage of energetic approaches is that because the envelope of sound energy changes much more slowly temporally and spatially than that of a sound pressure field sampled at a particular point, the numerical methods used to approximate solutions can use far coarser (and therefore, far fewer) elements as well as lower sample rates to discretize a particular problem domain. In cases of outdoor noise studies where typical room acoustics approaches are infeasible due to the size of the area under study, this increase in speed makes prediction of certain acoustical indices tractable.

Energy-based methods have also historically been used to analyze and synthesize spatial impulse responses. Typically, the quantities of interest are the energy density (a scalar at the point of measurement) and the sound intensity (a vector indicating energy flow). Between these two quantities, the progression of the soundfield toward an ideally diffuse state as well as the direction of arrival (DOA) of particular packets of non-diffuse energy can be predicted as a function of time, as demonstrated in Merimaa and Pulkki (2005) and Pulkki and Merimaa (2006).

This thesis also begins with an analysis-synthesis approach, where measurements and simulated soundfields are linked to one another, but the ultimate goal is the direct prediction of diffuse energy from room geometry and materials.

1.3 Motivation

As mentioned before, the advantage of focusing on the stochastic reverberation is that it can be modeled as a noise sequence with statistical properties matching those of the stochastic reverberation in a particular impulse response.

In room acoustics, auralization of the late reverberation has often been accomplished either by performing a simulation that is capable of reproducing the reverberant tail (e.g. ray tracing with thousands of rays or a wave-based simulation that runs beyond the reverberation time of the space) and convolving the impulse response with source material, or

by using a reverberator structure to artificially reproduce the late reverberation in terms of frequency-dependent exponential decay rates. With the advent of fast convolution techniques, pioneered by Reilly and McGrath (1995), both of these approaches are viable for real-time reproduction of known soundfields.

The downside of the reverberator approach is that it requires fitting to a known or measured impulse response, and modification depending on position is not straightforward. Even recent extensions by Alary et al. (2019) to traditional reverberator structures that capture spatial details are best suited to modeling a single spatial source-receiver pair, despite the fact that they allow for accurate representation of input and output directionality.

Thus, for rapid, physically-informed prediction of the stochastic reverberation that directly and deterministically represents acoustical properties such as scattering in a room's evolution toward the diffuse field, wave-based energetic approaches make for an attractive option in balancing realism and efficiency.

1.4 Outlook

Over the course of this thesis, we present a method for modeling the stochastic reverberation based on the acoustic energy-stress tensor, or EST, which represents the energy of sound waves as well as the changes in their momenta and wave stress in a conservative fashion. This means that for particular spaces, we may model the flow of sound energy throughout the space directly, without having to model the individual pressure waves that make up the soundfield. We will cover the theory of the constitutive equations, implementation of a previous and a new discretization of the system including sources, the auralization of the resulting numerical models in a hybrid scheme, and finally, a direct computation of the terms of the energy-stress tensor from a pressure-based simulation with well-known boundary conditions to learn more about the behavior of the tensor without relying on diffuse field assumptions to form a tractable system of equations.

In essence, each of these four main research outputs regarding the EST method each serve a portion of extending the model toward predictively modeling large three-dimensional spaces.

Chapter 3 is a study of a preexisting finite difference time domain (FDTD) solution to a 1-dimensional version of the EST, which simplifies to the telegraph equations. The goal of this study was to address the solution space of the method for all possible combinations of the two boundary condition parameters introduced. To that end, in comparison with two real hallways, we proposed combinations of coefficients for which the method produced a soundfield with equivalent acoustical measures in terms of specific frequency bands. By analyzing the frequency ranges for which this was possible, as well as the differences between the two hallways and the resulting model validity, we attempt to verify the diffuse field

assumptions of the model as well as setting up an approach for performing bandlimited synthesis of the stochastic reverberation where the method is applicable.

Chapter 4 focuses on recasting the method from the previous FDTD discretization to a finite volume time domain (FVTD) approach, and derives the inclusion of source terms for the model. This formalism directly characterizes energy flows between discretized cells as well as providing an implementation of boundary conditions that is more easily generalizable to higher dimensions. As in the previous chapter, comparison of the solution space of the model to physical measurements in a real hallway provides a useful way to confirm the frequency-dependent validity of the model.

Chapter 5 presents an approach to auralize the energy density field resulting from the FVTD model introduced in the previous chapter in the context of a hybrid methodology using geometric and wave-based approaches for the specular and modal regions of synthesized impulse responses, respectively. In this section, we introduce in greater detail the FVTD discretization approach as it applies to a pressure-velocity wave equation, which results in minor changes to the difference equations as well as requiring different boundary conditions. Finally, we present an interface for comparing the simulated hybrid impulse responses to measurements from the real space with both auralizations and spectrograms as a function of distance from a source. Notes on informal listening tests highlight some of the successes and challenges arising from the hybrid model.

Finally, Chapter 6 describes the use of a high-frequency FVTD simulation in the pressure domain, the same as in the previous chapter but with a much higher sample rate, to directly compute and average the time-varying terms of the energy-stress tensor. This allows us to examine the behavior of the tensor in distinct regions of the room that was not possible with previous 1-dimensional numerical approaches or without exhaustive roboticized measurement of a real space. Finally, the results are examined through the lens of validating the diffuse field assumptions inherent in all of the preceding EST models. We present preliminary evidence supporting these claims with the hope that it may enable direct characterization of diffusing spaces with respect to their geometry and materials properties in the future.

In general, the goal of this thesis is to present the EST method as a viable candidate for calculating and synthesizing the stochastic reverberation in known rooms at present, and in the future, to predictively evaluate spaces that are not tractable under the dimensionality assumptions previously required to eliminate the complications of the wave-stress tensor, but rather to numerically solve for it directly. In the course of these studies, demonstrations of the advantages and drawbacks of the method, both computationally, theoretically, and in practical usage, will be discussed.

Part I

Energy-Stress Tensor Method

Chapter 2

Background

2.1 Introduction

Statistical models of reverberation are perhaps some of the oldest in acoustics. For example, Sabine's reverberation equation is still used today to model the stochastic portion of spaces of varying volumes, even if as an approximation, it is based only on rate of decay in terms of volume. Since then, our understanding of reverberation has grown, and with it, the aformentioned partitioning of impulse responses into domains that largely exhibit a particular acoustic phenomenon has allowed us to derive a more complete picture of how to model each of those regions.

In this case, we are interested in the "diffuse field" portion of the late reverberation. Typically, this is defined as a field where energy flows equally in all directions and is evenly distributed throughout a space. Though it has been shown that any absorbing room cannot support an ideal diffuse field, and recent studies such as Nolan et al. (2019) have further highlighted the anisotropic nature of soundfields in absorbing conditions, it is nonetheless a reasonable approximation to make. Given that most non-pathological rooms contain furniture or other scattering surfaces, it seems reasonable to presume that they achieve a diffuse field, even if it is not within the perceptual limits compared to an ideal diffuse field established by Romblom et al. (2016). Lindau et al. (2010) describes a procedure for finding the perceptual transition time, compared to statistical approaches focusing on echo density such as Abel and Huang (2006), Huang and Abel (2007), and Defrance et al. (2009).

Focusing on the diffuse field means that there is no need to accurately represent the early portion of the impulse response nor any modal behavior, though as will become clear, consideration of how these portions are exposed in modeling the diffuse field through energy-based techniques is an important part of understanding how to use such an approach to synthesize a portion of a wideband impulse response.

2.2 Statistical acoustics

Historically, as mentioned before, modeling reverberation has typically been statistical in nature. Even as understanding of the connection between the characteristics of reflections within a room and the resulting acoustical indices grew, before it was plausible to model such interactions, reverberation could be described statistically in terms of the room's attributes. Sabine's empirically derived equation relating room volume, total absorption, and reverberation time, can be regarded as the beginning of statistical approaches to room acoustical modeling.

Later, Eyring's approach illustrated how the loss of energy at each successive discrete reflection of a single ray related exponential decay and absorption. While his formula does not specifically model individual reflections, still essentially neglecting the shape of the room and individual travel times from reflection to reflection, it is the overall concept of a decaying soundfield that provided the basis for methods that could approach the treatment of sound energy in a statistical manner.

2.2.1 Energy Density and Sound Intensity

More recent approaches considered the mathematical and statistical behavior of soundfields themselves. The energy density of a field that satisfies the acoustic wave equation is defined in terms of the kinetic and potential energy at a given location and point in time resulting from the compression and rarefaction of gas. Energy density is not equivalent to sound pressure, but acts in a similar fashion, reflecting the tendency of energy throughout a space to return to equilibrium. Derivation of these quantities in the acoustic context is given in Morse and Feshbach (1953) and Morse and Ingard (1968). Though energy density and flux, which in acoustics we call the sound intensity, cannot be used to directly model pressure, application of these concepts in general acoustical fields such as that of Schiffrer and Stanzial (1994) encouraged interest in energetic approaches to acoustics outside of noise control studies, and as we show later on, can still be used for the synthesis of perceptually equivalent representations of the stochastic reverberation. One particularly popular use of the sound intensity is as a predictor of average direction of arrival for individual pressure wavefronts in a spatial impulse response, as described by Merimaa and Pulkki (2005).

Concurrently with the development of energy-based acoustics, statistical approaches based on the trajectories of particles of sound were being developed. The concept is similar to ray tracing, but the goal is not to collect discrete rays at a listener position, but rather, to understand the statistical properties of such a system throughout time. As demonstrated in Polack (1992), under this formalism, the idea of the transition time could be tied to development of the diffuse field, leading to the concept of diffusion as the mechanism for modeling

the stochastic reverberation. From these ideas, a model of diffuse soundfields based on particle diffusion was developed by Picaut et al. (1997), which would come to be known as the diffusion equation method, or DEM. Later improvements by Jing and Xiang (2007) were a result of computing the energy balance at the boundaries assuming an isotropic distribution of incidence, making it applicable for a wider variety of absorption coefficients. The use of the DEM continues in recent research, especially for large structures as in Sü Gül et al. (2019).

The thread that ties these two somewhat distinct concepts together is that of the diffuse soundfield itself, which comes with assumptions about the nature of the mean energy flow within a space. Thus, predicting the stochastic reverberation with the DEM is inherently tied to the behavior of the energy density and sound intensity within the space. The main issue with the method is that the resulting conservation equation for sound intensity does not explicitly involve time due to the assumption that the energy density in a space is nearly isotropic, resulting from the observation that sound intensity tends to zero much more quickly than energy density. Working from Morse and Feshbach (1953), it turns out that explicitly introducing time variance to the sound intensity requires the consideration of the wave-stress tensor, which generalizes the energy density and sound intensity and allows for conservation of sound intensity to be defined, but also greatly complicates computation of the resulting field. In the next section, the derivation of these relationships is performed, and the collection of the resulting terms into a single conserved quantity, the energy-stress tensor, is demonstrated.

2.3 Energy-Stress Tensor Method

2.3.1 Theory

We begin with a common model of 3-dimensional wave motion in room acoustics:

$$\frac{1}{c^2}\partial_{tt}\Psi - \Delta\Psi = 0, \tag{2.1}$$

where Ψ, the velocity potential of the field, is defined in terms of the particle velocity vector $\mathbf{v} = -\nabla\Psi$ and the sound pressure $p = \rho\partial_t\Psi$, where ∇ is the gradient operator, Δ the Laplacian operator, ρ the air density, and c the speed of sound. Finally, we notate the first and second partial derivatives according to coordinate i as ∂_i and ∂_{ii}, respectively.

We are interested in the energy stored in the field, that is, the energy resulting from the compression of the medium itself and the motion of the waves traveling through it. Normally, given a pressure p and particle velocity $\mathbf{v}$, we can compute the difference between kinetic and potential energy densities with $\frac{\rho}{2}|\mathbf{v}|^2 - p^2/(2\rho c^2)$. Equivalently, using the vector

potential and following the sign convention of Morse and Ingard (1968) p. 168, we may write the Lagrangian L as

$$\begin{aligned} L &= \frac{-\rho}{2}\left[\frac{-1}{c^2}(\partial_t\Psi)^2 + (\partial_x\Psi)^2 + (\partial_y\Psi)^2 + (\partial_z\Psi)^2\right] \\ &= \frac{\rho}{2}\left[\frac{1}{c^2}(\partial_t\Psi)^2 - |\nabla\Psi|^2\right]. \end{aligned} \tag{2.2}$$

By definition, the energy density of the system is given by

$$\begin{aligned} E_{tt} &= \partial_t\Psi(\partial_{(\partial_t\Psi)}L) - L \\ &= \partial_t\Psi(\partial_{(\partial_t\Psi)}L) + \frac{\rho}{2}\left[\frac{-1}{c^2}(\partial_t\Psi)^2 + |\nabla\Psi|^2\right] \\ &= \partial_t\Psi(\partial_{(\partial_t\Psi)}L) - \frac{1}{2c^2}\rho(\partial_t\Psi)^2 + \frac{\rho}{2}|\nabla\Psi|^2 \\ &= \partial_t\Psi(\frac{\rho}{c^2}\partial_t\Psi) - \frac{1}{2c^2}\rho(\partial_t\Psi)^2 + \frac{\rho}{2}|\nabla\Psi|^2 \\ &= \frac{\rho}{c^2}(\partial_t\Psi)^2 - \frac{1}{2c^2}\rho(\partial_t\Psi)^2 + \frac{\rho}{2}|\nabla\Psi|^2 \\ &= \frac{\rho}{2c^2}(\partial_t\Psi)^2 + \frac{\rho}{2}|\nabla\Psi|^2. \end{aligned} \tag{2.3}$$

Similarly, the three terms of the sound intensity $\mathbf{I} = (E_{xt}, E_{yt}, E_{zt})$ are

$$\begin{aligned} E_{xt} &= \partial_t\Psi(\partial_{(\partial_x\Psi)}L) = -\rho\partial_t\Psi\partial_x\Psi, \\ E_{yt} &= \partial_t\Psi(\partial_{(\partial_y\Psi)}L) = -\rho\partial_t\Psi\partial_y\Psi, \\ E_{zt} &= \partial_t\Psi(\partial_{(\partial_z\Psi)}L) = -\rho\partial_t\Psi\partial_z\Psi, \end{aligned} \tag{2.4}$$

or, all together,

$$\mathbf{I} = -\rho\partial_t\Psi\nabla\Psi. \tag{2.5}$$

We confirm that these terms satisfy an equation of continuity, such that any change in the energy density is due to a change in the sound intensity. By the use of the general

Leibniz rule, $\partial_i(\partial_i\Psi)^2 = 2\partial_i\Psi\partial_{ii}\Psi$, we have:

$$\begin{aligned}
&\partial_t E_{tt} + \nabla \cdot \mathbf{I} \\
&\quad = \frac{\rho}{c^2}\partial_t\Psi\partial_{tt}\Psi + \rho\partial_x\Psi\partial_t\partial_x\Psi + \rho\partial_y\Psi\partial_t\partial_y\Psi + \rho\partial_z\Psi\partial_t\partial_z\Psi \\
&\quad - \rho\partial_x(\partial_t\Psi\partial_x\Psi) - \rho\partial_y(\partial_t\Psi\partial_y\Psi) - \rho\partial_z(\partial_t\Psi\partial_z\Psi) \\
&\quad = \frac{\rho}{c^2}\partial_t\Psi\partial_{tt}\Psi + \rho\partial_x\Psi\partial_t\partial_x\Psi + \rho\partial_y\Psi\partial_t\partial_y\Psi + \rho\partial_z\Psi\partial_t\partial_z\Psi \\
&\quad - \rho\partial_x\partial_t\Psi\partial_x\Psi - \rho\partial_y\partial_t\Psi\partial_y\Psi - \rho\partial_z\partial_t\Psi\partial_z\Psi \\
&\quad - \rho\partial_t\Psi\partial_{xx}\Psi - \rho\partial_t\Psi\partial_{yy} - \Psi\partial_t\Psi\partial_{zz}\Psi \\
&\quad = \rho\partial_t\Psi[\frac{1}{c^2}\partial_{tt}\Psi - \Delta\Psi] = 0.
\end{aligned} \tag{2.6}$$

where we recognize the wave equation in the penultimate expression.

Additionally, we define the wave-stress tensor, $\underline{\underline{E}}$, whose components E_{ij} fulfill the second vector continuity equation $\partial_t\mathbf{I} + \nabla \cdot \underline{\underline{E}} = 0$, which will be proven next.

For $i, j = x, y, z$, and with $\alpha_{ij} = 1$ when $i = j$ or -1 otherwise:

$$\begin{aligned}
E_{ii} &= L - \partial_i\Psi\partial_{(\partial_i\Psi)}L \\
&= \frac{\rho}{2}(\frac{1}{c^2}|\partial_t\Psi|^2 + \sum_j \alpha_{ij}|\partial_j\Psi|^2), \\
E_{ij} &= -\partial_i\Psi\partial_{(\partial_j\Psi)}L \\
&= -\rho\partial_i\Psi\partial_j\Psi.
\end{aligned} \tag{2.7}$$

Note that by the symmetry of products of derivatives, $E_{xy} = E_{yx}$, $E_{xz} = E_{zx}$, and $E_{yz} = E_{zy}$.

Thus, the symmetric wave-stress tensor may be written in full as

$$\underline{\underline{E}} = \begin{pmatrix} E_{xx} & E_{xy} & E_{xz} \\ E_{yx} & E_{yy} & E_{yz} \\ E_{zx} & E_{zy} & E_{zz} \end{pmatrix}. \tag{2.8}$$

With this definition, we may write the continuity equations for sound intensity.

$$\begin{aligned}
&\frac{1}{c^2}\partial_t E_{xt} + \partial_x E_{xx} + \partial_y E_{xy} + \partial_z E_{xz} = 0,\\
&\frac{1}{c^2}\partial_t E_{yt} + \partial_x E_{yx} + \partial_y E_{yy} + \partial_z E_{yz} = 0,\\
&\frac{1}{c^2}\partial_t E_{zt} + \partial_x E_{zx} + \partial_y E_{zy} + \partial_z E_{zz} = 0.
\end{aligned} \tag{2.9}$$

Taking the first equation as an example and again recognizing the wave equation in the penultimate expression, we have

$$\begin{aligned}
&\frac{1}{c^2}\partial_t E_{xt} + \partial_x E_{xx} + \partial_y E_{xy} + \partial_z E_{xz}\\
&\quad = \frac{-\rho}{c^2}\partial_x\Psi\partial_{tt}\Psi - \frac{\rho}{c^2}\partial_t\Psi\partial_{tx}\Psi + \frac{\rho}{c^2}\partial_t\Psi\partial_{xt}\Psi\\
&\quad + \rho\partial_x\Psi\partial_{xx}\Psi - \rho\partial_y\Psi\partial_{xy}\Psi - \rho\partial_z\Psi\partial_{xz}\Psi\\
&\quad + \rho\partial_x\Psi\partial_{yy}\Psi + \rho\partial_y\Psi\partial_{xy}\Psi + \rho\partial_x\Psi\partial_{zz}\Psi + \rho\partial_z\Psi\partial_{xz}\Psi\\
&\quad = -\rho\partial_x\Psi[\frac{1}{c^2}\partial_{tt}\Psi - \Delta\Psi] = 0.
\end{aligned} \tag{2.10}$$

The second and third lines of the vector equation proceed in the same manner, showing that the system as a whole, with the inclusion of the continuity equation for energy density, is conservative.

Energy-Stress Tensor

Because of this fact, all of the terms may be collected into a single conserved tensor, which we call the energy-stress tensor, or EST, for short:

$$\underline{\underline{T}} = \begin{pmatrix} E_{tt} & E_{xt} & E_{yt} & E_{zt} \\ E_{xt} & E_{xx} & E_{xy} & E_{xz} \\ E_{yt} & E_{yx} & E_{yy} & E_{yz} \\ E_{zt} & E_{zx} & E_{zy} & E_{zz} \end{pmatrix}. \tag{2.11}$$

Specifically, this tensor distills the continuity equations for energy density, sound intensity, and the energy momentum flux (as expressed in the wave-stress tensor). In fact, the conservation of the system presented in Equations 2.6 and 2.9 is an expression of Noether's theorem, which relates the presence of invariances in a given system to conservation laws. In the context of general relativity, the theorem was used to explain the relationship of the conservation laws of linear momentum, angular momentum, and energy to symmetries in

translation, rotation, and time, respectively.

EST in lower dimensions

In the 3-dimensional case, we have four coupled equations, and because the tensor is symmetric, the free variables are every member of the upper triangular part, for a total of ten.

Furthermore, it can be seen that the 1- and 2-dimensional cases are simply subsets of these equations, where eliminating the last term in each equation and the final equation itself results in a system one dimension lower. Thus, for the 2-dimensional case, there are three equations and six unknowns, and the 1-dimensional case has two equations and three unknowns.

For example, in the 1-dimensional case, we have the following set of equations:

$$\begin{aligned} \frac{1}{c^2}\partial_t E_{tt} + \partial_x E_{xt} &= 0, \\ \partial_t E_{xt} + \partial_x E_{xx} &= 0. \end{aligned} \tag{2.12}$$

Performing the same computations as in the 3-dimensional case demonstrates that this system is also conservative, and thus represents an energy-stress tensor of rank 2. In this case, direct computation of E_{tt} and E_{xx} shows that they are equal, resolving this system of two equations and two unknowns, however, in the 2- or 3-dimensional case, this is no longer the case, leaving these systems underdetermined.

We will now proceed to describe the assumptions and strategies used to resolve these systems using the theory of diffuse fields in a particular space by the specification of boundary conditions, which has been the focus of past and present work on the topic. We will call the use of this approach the energy-stress tensor method, or EST method for short.

2.3.2 Scattering and diffusion

Introducing boundary conditions into the EST method relies on the presence of scattering that occurs at the walls or other surfaces within a space. These reflections can be specular or diffuse depending on the relative scale of the surface's features and the frequency of the incident wave, as shown in Figure 2.1. Note that in this context, the term "diffuse reflection" is used in the optical sense, meaning that incident waves are scattered rather than the mirror-like behavior of a specular reflection. In other words, the name should not be taken to mean a reflection that is occurring after the soundfield has become diffuse.

For typical rooms, the region above Schroeder's frequency often demonstrates a sufficiently high proportion of diffuse reflections to cause the distribution of energy throughout the room to become increasingly isotropic and eventually uniform due to the stochastic

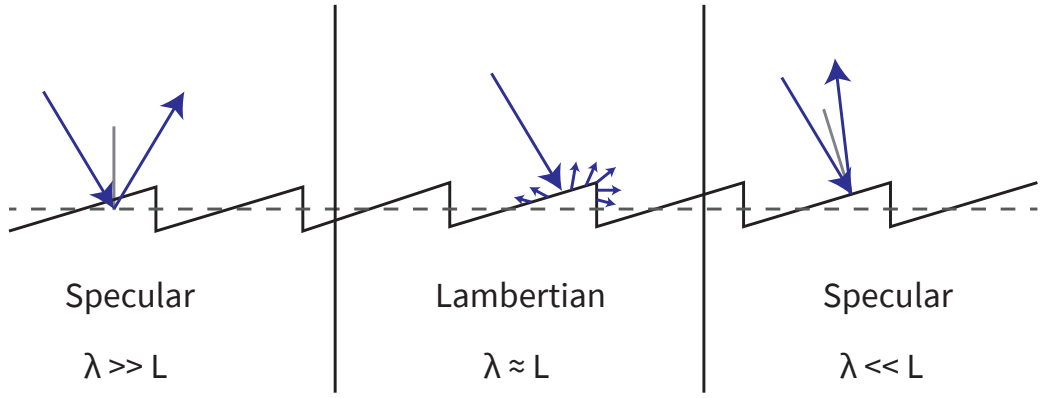

Figure 2.1: Frequency dependence of reflection scattering on surface with characteristic dimension L, after Cox et al. (2006)

nature of the reflected wave directions. This process is referred to as diffusion, and is the driving force behind the temporal evolution of an impulse response from the specular early reflections to the late reverberation.

Study of scattering processes has led to advances in architectural acoustical modeling, both in geometrical approaches such as ray tracing, as well as the energetic methods that are the subject of interest in this research. Additionally, in the practical sense, understanding the design and use of "diffusers" or other scattering surfaces has been important in the design of concert halls and recording studios since their introduction by Schroeder (1975) and summarized in Cox and D'Antonio (2016).

For geometric methods, the importance of modeling diffuse reflections has long been understood, as evidenced in Kleiner et al. (1993), and most modern approaches integrate a frequency-dependent scattering coefficient to approximate these effects within their stochastic frameworks, such as Schröder (2011).

In energy-based modeling, the focus instead has been on the development of the diffusion equation and the integration of a statistical diffusion coefficient. This idea was first proposed by Ollendorff (1969), and was later refined by Picaut et al. (1997), leading to the DEM discussed above.

2.3.3 Dimensional reduction

In the case of the EST method, integrating over surfaces perpendicular to the direction of energy density and sound intensity wave motion and introducing energy and momentum balances on the walls was demonstrated to be capable of modeling 1- and 2-dimensional rooms in Dujourdy et al. (2017) and Dujourdy et al. (2019). In this case, the "dimensionality" of the room refers to the number of cardinal directions that are much longer than the

characteristic wavelength of energy density waves, which are on the order of a meter. This is related to the frequency of the diffuse field, which is dominated by the rate of decay, and has been experimentally shown to be below 20 Hz. This characteristic is described in more detail in Polack et al. (1984a) and Polack et al. (1984b).

Since this is typically very slow in comparison to the frequency of pressure waves, which are often modeled up to 20 kHz to match the human auditory system, in the case in a long corridor or an open office, the problem can be effectively treated as being 1- or 2-dimensional, respectively. Put another way, integrating the EST continuity equations over any dimension that is of the same characteristic length as the energy density (meaning it is not a primary direction of energy propagation) and postulating an energy balance and a momentum balance on the corresponding surfaces allows us to rewrite the sound intensity and wave-stress tensor terms involving derivatives of that coordinate in terms of the remaining directions. In the 1-dimensional case, this amounts to integrating over two dimensions and redefining the four associated terms (two sound intensity directions and two momentum fluxes) in terms of the remaining energy density and sound intensity. As will be seen below, this results in a tractable form resembling the transmission line equations.

In the two dimensional case, such as an open office, only one dimension is able to be eliminated with this strategy, leaving its two associated members of the EST defined in terms of the two remaining sound intensities and four momentum fluxes. Since this process may only remove one continuity equation, the resulting system is underdetermined, and requires further assumptions to close the system. In Dujourdy et al. (2019), for example, the Z-direction was eliminated, resulting in the following system of equations:

$$\frac{1}{c}\partial_t E + \partial_x J_x + \partial_y J_y + \frac{A}{\lambda}E = 0,$$

$$\frac{1}{c}\partial_t \mathbf{J} + \frac{D}{\lambda}\mathbf{J} + (\partial_x, \partial_y)\begin{pmatrix} E_{xx} & E_{xy} \\ E_{xy} & E_{yy} \end{pmatrix} = 0,$$

where $\mathbf{J} = (J_x, J_y) = \mathbf{I}/c$. We will ignore A, D, and λ for the time being, taking them only as arbitrary constants. In this situation, there are 6 free variables ($E, J_x, J_y, E_{xx}, E_{xy}$, and E_{yy}) and only 3 equations.

The authors proposed two further assumptions to close the system. The first is equipartition of energy, meaning that the total energy is evenly split between the two directions of propagation, or in short, that $E_{tt} = E_{xx} + E_{yy}$. The second assumption is decorrelation, which means that the soundfield is isotropic, and the direction of propagation of waves in the field is equally distributed. Under this assumption, an important part of defining the diffuse field, the energy flux from one dimension to another must be zero, as any change would imply that one direction was favored. This eliminates the off-diagonal elements of the

wave-stress tensor, such that $E_{xy} = E_{yx} = 0$. While these assumptions are not necessary for the 1-dimensional cases we will study over the course of this thesis, we mention them here to reflect back upon in the final chapter, when a computational study of a soundfield may allow us to confirm or deny whether these assumptions are valid.

Now, we proceed with a derivation of the reduction of the 3-dimensional EST to a 1-dimensional solution, following from Dujourdy et al. (2017). A hallway is a good candidate for a 1-dimensional reduction, as its primary axis is much longer than either its width or height. We further presume that the width and height are on the order of the characteristic length of the EST, meaning that our discretization need not exceed one sample, and allowing us to treat the entire space according to linear samples along its length.

To begin, we consider an arbitrary hallway as a rectangular solid of dimensions $l_x \times l_y \times l_z$, where l_x is the length, l_y the width and l_z the height of the corridor, on average, implying a cross-sectional area $S = l_y l_z$. Note that despite the fact that we are interested in reducing the problem to modeling only the propagation of sound along the length of the corridor, we still must consider the 3-dimensional nature of the hallway, and thus, begin from the 3-dimensional version of the energy-stress tensor.

Energy balance

We return first to the continuity equation for energy density, Equation 2.6:

$$\partial_t E_{tt} + \partial_x E_{xt} + \partial_y E_{yt} + \partial_z E_{zt} = 0.$$

We hypothesize that E_{tt} and E_{xt} are constant along the cross-sectional area of the corridor, that E_{yt} is independent of z, and that E_{zt} is independent of y.

Then, we may integrate over the cross-sectional area in order to find a single continuity expression for that point along the corridor. With dy and dz elements of l_y and l_z, we have

$$\begin{aligned} 0 &= \frac{1}{c}\left[\iint \partial_t E_{tt} dxdy + \iint \partial_x E_{xt} dxdy + \iint \partial_y E_{yt} dxdy + \iint \partial_z E_{zt} dxdy\right] \\ &= \frac{1}{c}\partial_t ES + \partial_x J_x S + (J_y^+ - J_y^-)l_z + (J_z^+ - J_z^-)l_y. \end{aligned} \tag{2.13}$$

where $J = I/c$.

This relation introduces the mean sound intensity in each axial direction for each of the four walls, indicated by + and - in terms of their coordinate system, such that $J_i^+ = -J_i^-$. Now, we propose an energy balance on the walls relating the energy density and the incident, reflected, and absorbed portion of the sound intensity for a given section.

Given an incident and reflected normal sound intensity, the absorbed sound intensity is,

of course, the difference between the two:

$$J_{\text{abs}} = J_{\text{inc}} - J_{\text{ref}}. \tag{2.14}$$

The absorbed sound intensity, in accordance with Sabine theory, is

$$J_{\text{abs}} = \alpha J_{\text{inc}}, \tag{2.15}$$

where α is the Sabine absorption coefficient.

Now, the task is to determine expressions for the incident and reflected sound intensity in terms of the average sound intensity and energy in front of the wall. Jing and Xiang (2007) argued that one quarter of the total sound energy E enters and leaves each wall, due to its equal distribution. Furthermore, half of the sound intensity in front of each specific wall enters and leaves the wall. As is apparent from the fact that the absorbed sound intensity is proportional to the incident sound intensity, and equal to the difference between the incident and reflected sound intensity, the sign of the sound intensity terms must be opposite. Thus,

$$\begin{aligned} J_{inc} &= \frac{E}{4} + \frac{J}{2}, \\ J_{ref} &= \frac{E}{4} - \frac{J}{2}. \end{aligned} \tag{2.16}$$

This gives us a relationship between the mean sound intensity in front of each wall and the total energy shared by all four walls:

$$J = \frac{\alpha}{2(2-\alpha)} E. \tag{2.17}$$

With $A = \frac{\alpha}{1-\frac{\alpha}{2}}$, this reduces to

$$J = \frac{A}{4} E. \tag{2.18}$$

Then, returning to Equation 2.13 and taking all four walls into account, the expression becomes

$$\frac{1}{c}\partial_t ES + \partial_x J_x S + AE\frac{l_y + l_z}{2} = 0. \tag{2.19}$$

Finally, we develop an expression for the mean free path in the hallway, keeping in mind

that $l_x \gg l_y, l_z$ and allowing us to neglect terms that become very small.

$$\begin{aligned}\lambda &= \frac{4V}{S_{\text{tot}}} \\ &= \frac{4l_x l_y l_z}{2l_x(l_y + l_z) + 2l_y l_z} \\ &= \frac{4l_y l_z}{2(l_y + l_z) + 2l_y l_z / l_x} \\ &= \frac{4S}{2(l_y + l_z)} \\ &= \frac{2S}{l_y + l_z}.\end{aligned} \tag{2.20}$$

Ultimately, we can now write an equation for the energy density and axis-aligned sound intensity:

$$\frac{1}{c}\partial_t E + \partial_x J_x = -\frac{A}{\lambda}E. \tag{2.21}$$

Momentum balance

A similar method may be used to perform a dimensional reduction of Equation 2.9. The assumptions required for this process are that J_x and E_{xx} are constant on the cross-sectional area of the hallway, and like before, E_{xy} is independent of y and E_{xz} is independent of z. In order to form a momentum balance on the walls, diffuse field theory is once again used to define the redistribution of energy from the direction of propagation to the off-axis directions. Acoustically, this redistribution has a physical explanation in scattering, where wave energy is redirected from one direction to another; therefore, the degree of scattering is what ultimately determines the relationship between the incident sound intensity and the wave stress flux in a particular region. In this manner, the off-diagonal components of the wave-stress tensor may be accounted for in terms of the previously used tensor values.

We integrate again over the cross-sectional area of the corridor, as in 2.13, but using the X component of the continuity equation for sound intensity.

$$\begin{aligned}0 &= \frac{1}{c^2}\iint \partial_t E_{xt} dxdy + \iint \partial_x E_{xx} dxdy + \iint \partial_y E_{xy} dxdy + \iint \partial_z E_{xz} dxdy \\ &= \frac{1}{c}\partial_t J_x S + \partial_x E_{xx} S + (E_{xy}^{+} - E_{xy}^{-})l_z + (E_{xz}^{+} - E_{xz}^{-})l_y.\end{aligned} \tag{2.22}$$

The $E_{ij}^{\pm}$ terms are the wave stress (or, equivalently, momentum flux) in front of their respective walls, as noted previously.

Dujourdy et al. (2017) proposed a momentum balance on the walls, such that the scattered sound momentum is proportional to the incident momentum, and similar to the energy balance as before, the scattered portion is equal to the difference between the entering and outgoing momentum. In other words, where the momentum flux M represents an off-diagonal term of the wave-stress tensor such as E_{xy} or E_{xz},

$$\begin{aligned} M_{\text{scat}} &= \beta M_{\text{ent}} \\ &= M_{\text{ent}} - M_{\text{out}}. \end{aligned} \tag{2.23}$$

As before, the challenge is to determine the ingoing and outgoing momentum flux in terms of the wave stress and sound intensity in front of each segment. Since in this case J_x is parallel to the wall, it must contribute equally to each momentum flux. Conversely, according to the momentum balance, half of the wave stress enters and leaves the wall, but with opposite signs.

$$\begin{aligned} M_{ent} &= \frac{J_x}{4} + \frac{M_{\text{scat}}}{2}, \\ M_{out} &= \frac{J_x}{4} - \frac{M_{\text{scat}}}{2}. \end{aligned} \tag{2.24}$$

In combination with the momentum balance, this results in

$$\begin{aligned} M_{\text{scat}} &= \beta(\frac{M}{2} + \frac{J_x}{4}) \\ &= \frac{\beta}{2(2-\beta)} J_x. \end{aligned} \tag{2.25}$$

We define the modified scattering coefficient in a similar fashion as the modified absorption coefficient above:

$$D = \frac{\beta}{1 - \beta/2}. \tag{2.26}$$

Then, all of the wave stress terms in front of the wall may be written as

$$E^+_{xy} = -E^-_{xy} = E^+_{xz} - = E^-_{xz} = \frac{D}{4} J_x. \tag{2.27}$$

At this point, we have completed our coverage of the terms of the EST as we have defined J_y, J_z, E_{xy}, and E_{xz} solely in terms of E, J_x, and E_{xx}. By symmetry, we also have expressions for E_{yx} and E_{zx}, and because no energy is presumed to travel in the Y and Z directions, we may eliminate E_{yy}, E_{yz}, E_{zy}, and E_{zz}. The last term that remains is E_{xx}, however, as it is the only remaining energy density term, it must be equal to the total energy, E_{tt}. Therefore, replacing terms in Equation 2.22 and using the expression

for the mean free path in the hallway, the resulting dimensional reduction of the intensity conservation equation is

$$\frac{1}{c}\partial_t J_x + \partial_x E = -\frac{D}{\lambda} J_x. \tag{2.28}$$

Together, Equations 2.21 and 2.28 constitute a 1-dimensional reduction of the EST system for a long hallway similar to the transmission line equations. As described in Dujourdy et al. (2017), it is possible to relax some of these assumptions by considering averages for each of the integrals performed; however, this does not change the underlying geometric requirements, namely that the room is much longer in one dimension, and that energy flows primarily along that axis.

Discussion

To summarize, the energy and momentum balances mentioned before are effectively hypotheses regarding the absorption and scattering of energy density, sound intensity, and wave stress at the walls based on Sabine's theory of reverberation and diffuse field theory. It is useful to note here that this approach to dimensional reduction and the assumptions implicit in it remains relevant for the rest of the thesis. Further development of this specific form of the 1-dimensional EST is limited to the following chapter and the study of the numerical scheme used to simulate the time-varying energy density in the hallway, however, we will return to the development of the energy and momentum balances once again when we switch to a different discretization scheme for the model. As will shortly be demonstrated, these coupled equations can further be reduced to a form involving only the energy density called the telegraph equation. The development of boundary conditions on the surfaces not included in the volume equations above (essentially the ends of the hallway) is specific to this form, which is why we do not address them here. In later chapters, we will use a finite volume approach to discretization of the system, reusing the ideas from this dimensional reduction, but recast in a form that treats all boundary conditions in the same fashion, rather than integrating some of the assumptions of the physics occurring at the domain edges into the volume equations.

Furthermore, while we will not cover it in detail here as it covers some of the same ground, the studies Dujourdy et al. (2017) and Dujourdy et al. (2019) from which these derivations originate demonstrated the existence of boundary conditions that produce numerical results that match a physical space in the 1000 Hz frequency band. In the following chapter, we will extend these results to cover a wider frequency range, again by comparison with measurements of a real hallway, validating the treatment of such a space as 1-dimensional.

2.3.4 Limitations

While the strategy of dimensional reduction is valid and useful for modeling the stochastic reverberation in rooms with at least one dimension on the order of the characteristic length of energy density, the extension to rooms of 3 dimensions (where none of the lengths are on the order of the meter) using the same strategy is not tractable. Finding an alternative strategy for the development of the energy-stress tensor in 3-dimensional spaces is the focus of the rest of this thesis.

In both the 1- and 2-dimensional cases, the assumptions required to eliminate the off-diagonal elements and reduce the system to the solvable telegraph equation form are a direct result of diffuse field assumptions from the diffusion equation approaches discussed above, and therefore, the solutions are only valid when the field is sufficiently diffuse. This is expected, but nonetheless constrains the use of the method to spaces that scatter enough energy to benefit from modeling the stochastic reverberation in this fashion: a room that is not diffusing, even at high frequencies, may not be representable under these assumptions. Nonetheless, many rooms that do not mix very well may still exhibit sufficient diffusion at high frequencies such that at least a portion of their response may be modeled stochastically.

Furthermore, while the modified absorption and scattering coefficients introduced to parameterize materials at the room boundaries are based on well-known acoustic properties, they are nonetheless predicated on the diffuse field assumptions regarding the acoustic behavior near the boundaries, and can be difficult to assign purely based on geometry or materials properties, particularly when it comes to the characterization of scattering. As with some other simulation methodologies, an adjustment of model procedure (a term borrowed, in this case, from Ewins (2000) and the field of modal analysis) must be carried out to ensure that the method reproduces measured characteristics. While such a method can capably model a particular space, the extension to other spaces where the same assumptions may be invalid, whether due to changes in materials or geometry, may not hold.

As mentioned before, in eschewing the direct description of acoustic pressure and particle velocity, many of the most important aspects of room reverberation are not feasible to represent with the energy-stress tensor method: the direct sound and early reflections contain perhaps the most salient perceptual details, but cannot be modeled with this approach. Nonetheless, the stochastic reverberation still makes important contributions to the overall impression of a given space, and can be difficult to model in a physical fashion without computational difficulty as described before, making it a worthwhile research topic to accelerate the generation of realistic reverberation.

2.3.5 Advantages

One practical advantage of the EST approach is that the sample rates used can be very low owing to the low modulation frequency of the energy density. This is particularly advantageous as the frequencies we actually intend to model with the method are those above Schroeder's frequency, and extending to a full band representation above that. This crossover frequency corresponds to wavelengths that are of comparable lengths to the dimensions of a room, and typically falls below 200 Hz, even for small rooms. Performing a wave-based simulation above this frequency range can be challenging as well as computationally expensive due to the fact that the size of the elements required is inversely proportional to the highest representable frequency. This means that a doubling of the sample rate leads to an 8-fold increase in algorithmic complexity and storage requirements. Conversely, the sample rate required for the EST method depends only on the rate of decay of the energy density, and numerical stability is possible with elements on the order of a meter, independent of the frequency band in consideration. This comes with the added advantage that discretization of the space does not change for different frequency bands; in effect, a single meshing is sufficient for all of the regions of interest, eliminating complications common to wave-based methods such as adaptive meshing or higher order methods.

This quadratic improvement in speed is notable for large spaces, especially those that enclose volume in all 3 dimensions, as the number of elements grows particularly quickly compared to long or flat spaces. That a simulation of the stochastic reverberation above a given frequency might be accomplished with a single, very coarse meshing thus represents an extreme economization over other wave-based methods that are highly dependent on element size.

Another result of the low sample rate requirement is that the corresponding spatial discretization can result in a similar mesh size as other low-frequency wave-based simulations. In a hybrid real-time auralization scheme that can accurately represent modal phenomena, these low-frequency pressure simulations would be required, necessitating their own meshing step. Because of the large spatial discretization, however, a single meshing of the problem domain could be utilized for both a pressure-based room acoustic simulation as well as for the EST method, resulting in increased computational efficiency in cases with changing geometry as well as reduced complexity in the implementation of a resulting hybrid acoustical scheme.

Chapter 3

Frequency Dependence and Validity of a 1D Model

3.1 Introduction

Due to the assumptions that underlie the domain of applicability of stochastic reverberation, specifically the diffuse field criteria, we expect there to be a dependence on frequency regarding the validity of the EST model as well as a minimum threshold of diffusion required to be able to qualify the soundfield as stochastic. These characteristics are expected to change from room to room, as depending on the geometry, absorption, and presence of scattering surfaces, the relative composition of an impulse response may be dominated by modal or specular phenomena to the point where attempting to fit a stochastic model may not make sense. To that end, validating the frequency ranges where the EST method can be expected to function is an important part of understanding its limitations.

Previously, the method was demonstrated to function in the 1000 Hz frequency band for both 1- and 2-dimensional spaces: a long hallway with some recesses and a chicane, and an open-plan office with a number of columns and other diffusing objects. We chose to refocus on the 1-dimensional case of a hallway as the boundary conditions and analysis are somewhat simplified.

The full document referenced in this chapter (Meacham et al. (2019b)) is reproduced in Appendix A.1.

3.2 Parameter fitting

This chapter is an exploration of the relative effects of the two boundary condition parameters in the 1-dimensional EST model, α and β. As mentioned previously, these parameters represent the Sabine absorption and the so-called scattering coefficient, respectively, though their role in the 1-dimensional solution of the EST model is somewhat intertwined as will be shown in this section. Up to this point, these coefficients have been represented in the volume equations for the EST, playing a role in the definition of the boundary of a hallway, even if this is not the same as a typical boundary condition for a system of differential equations. Shortly, the ends of the hallway will be closed with a traditional boundary condition at the extrema of the 1-dimensional problem.

Furthermore, we would like to relate these parameters to frequency-dependent phenomena in real rooms in order to assess if it is possible to reproduce the soundfields, and, if so, to learn more about the required combinations of coefficients. As we are comparing directly to real spaces, we furthermore propose direct observation of the parameters by exhaustively simulating many combinations and observing the resulting changes in the predicted time-varying energy density. For this reason, we will use the finite difference time domain (FDTD) discretization described in Dujourdy et al. (2017), and detailed below in Section 3.2.3.

We assume that if a combination of parameters produces an energy density that by some measure matches the soundfield in a real space, then the real soundfield is representable with the EST approach. As previously mentioned, we expect the real soundfields to be more or less diffuse depending on frequency, therefore we will also evaluate these matches as they pertain to a particular octave band. We propose two specific measures in Section 3.3 below.

3.2.1 Telegraph equation

Now, we revisit the original appearance of α and β in the context of the 1D EST equations in order to understand their context within the present experiment. These terms are a result of the energy balance and momentum balance defined in the dimensional reduction by integration as described in Section 2.3.3. As these two coefficients are the only free variables in the 1-dimensional EST approach, by definition, they completely define the boundary conditions, making the analysis of their effects on the resulting energy density simulations also a complete accounting of the soundfields the method is capable of representing (for the specific geometry at hand).

As a reminder, we previously arrived at a pair of coupled equations illustrating the

conservation of energy density and sound intensity, that is, Equations 2.21 and 2.28:

$$\begin{aligned} \frac{1}{c}\partial_t E + \partial_x J_x &= -\frac{A}{\lambda}E, \\ \frac{1}{c}\partial_t J_x + \partial_x E &= -\frac{D}{\lambda}J_x. \end{aligned} \tag{3.1}$$

Furthermore, we defined the modified absorption and scattering coefficients A and D as follows:

$$\begin{aligned} A &= \frac{\alpha}{1-\frac{\alpha}{2}}, \\ D &= \frac{\beta}{1-\frac{\beta}{2}}, \end{aligned} \tag{3.2}$$

where α and β are the Sabine absorption coefficient and the scattering coefficient, respectively.

From the form given in Equation 3.1, Dujourdy et al. (2017) developed the system by solving for the partial spatial derivatives with respect to the energy density and sound intensity.

$$\begin{aligned} (\frac{1}{c}\partial_t + \frac{A}{\lambda})E &= -\partial_x J_x, \\ (\frac{1}{c}\partial_t + \frac{D}{\lambda})J_x &= -\partial_x E. \end{aligned} \tag{3.3}$$

Since in the first equation, we have an expression in terms of the spatial derivative of the sound intensity, and in the second equation, we have an instance of the sound intensity, we notice that we may take a spatial derivative of the second equation and use the result to eliminate the sound intensity in lieu of a second order expression in energy density.

That is,

$$\begin{aligned} -\partial_x J_x(\frac{1}{c}\partial_t + \frac{D}{\lambda}) &= \partial_{xx}E \\ &= (\frac{1}{c}\partial_t + \frac{D}{\lambda})(\frac{1}{c}\partial_t + \frac{A}{\lambda})E. \end{aligned} \tag{3.4}$$

With some reorganization, and then expanding terms, we arrive at a form called the

telegraph equation:

$$\begin{aligned} 0 &= (\frac{1}{c}\partial_t + \frac{D}{\lambda})(\frac{1}{c}\partial_t + \frac{A}{\lambda})E - \partial_{xx}E \\ &= \frac{1}{c^2}\partial_{tt}E - \partial_{xx}E + \frac{A+D}{\lambda c}\partial_t E + \frac{AD}{\lambda^2}E. \end{aligned} \tag{3.5}$$

Note here again the symmetry of the modified absorption and scattering coefficients. This provides a limiting case in terms of the two coefficients, because even if their appearance in the telegraph equation is identical, we must keep in mind that the definitions of α and β permit only a certain range of values. In the case of α, the range is from 0 to 1, and in the case of β, the range is from 0 to 2. It is for this reason that the smaller of the two coefficients drives the absorption in the volume equations. Conversely, in terms of D, when it is very small, the momentum flux is effectively zero, meaning that most of the energy of the system is transported by the conservation of energy density, a case that resembles purely specular reflection. To the contrary, when D is very large, it does not permit any sound intensity flow along the boundary. This is clearly not realistic for any physical space, however, in between the two extremes, it points to the role of the modified scattering coefficient in opposing the diffusion of energy along the length of the corridor by redirecting it into the off-axis directions. Since those are not represented in this formulation as a result of the dimensional reduction, it appears as a loss of energy density, as with absorption, a point that would need to be revisited in the higher-dimensional case.

To put this statement another way, because no energy density or sound intensity is stored or transfered in the non-longitudinal directions, in this formulation, the effect of scattering is effectively the same as absorption, as any scattered energy is effectively lost. While this seemingly strange behavior comes from our knowledge of the frequency of the EST, leading to our consideration that its terms are effectively constant over the width and height of the hallway, one possible assumption that may help to justify this claim is the relative degree of scattering or diffracting surfaces that transform energy traveling along the length of the space into the other directions compared to the amount of scattering that transverse energy would undergo, again becoming longitudinal. Furthermore, transverse energy may undergo far more reflections over time, and perhaps be absorbed much more quickly. Thus, while we will continue with this approach during the course of the thesis, we note that it may be possible to preserve these terms, even in the 1-dimensional case of a hallway, and to examine the energy density stored in the transverse directions, an idea that will be revisited in Chapter 6, and may be further revisited in future work.

3.2.2 Boundary conditions on ends of the hallway

The final element that is needed in order to fully represent the EST in the hallway are the boundary conditions at the ends. The process of dimensional reduction by integration meant that the boundary conditions along the floor, ceiling, and two walls were all integrated into the volume equations written above as the telegraph equation, however, the physical behavior of the EST at the ends of the hallways remains undefined. Thus, we must introduce conditions for the ends of the hallway in terms of the energy balance defined earlier.

As before, Dujourdy et al. (2017) contend that the absorption at the wall is given by the energy density in front of it and the normal sound intensity incident on it:

$$J_x = \frac{AE}{4} = A_r E, \tag{3.6}$$

where A is again the modified absorption coefficient, and A_r denoting this special instance of the coefficient to distinguish it in the final system of equations.

Introducing the sign of the normal $n = \pm 1$ at each boundary according to the x coordinate, we may rewrite the second member of Equation 3.3 and replace the sound intensity with the expression above.

$$\begin{aligned} -n\partial_x E &= (\frac{1}{c}\partial_t + \frac{D}{\lambda})nJ \\ &= (\frac{1}{c}\partial_t + \frac{D}{\lambda})A_r E. \end{aligned} \tag{3.7}$$

This is a Neumann boundary condition as it is specified in terms of the spatial derivative of the energy density at the boundary.

At this point, we may proceed to a discretization of the continuous equations.

3.2.3 Finite difference time domain discretization

Dujourdy et al. (2017) discretizes the telegraph equation (Equation 3.5) with regularly spaced spatial and temporal samples. Choosing a given spatial sample step Δx and time sample step Δt means that the time evolution of the energy density in a particular space may be modeled with a grid of sample values, where a space of length l implies $l/\Delta x$ spatial samples, and where the number of temporal samples N_s is chosen such that the simulation runs from $t = 0$ to a termination time at $t = N_s \Delta t$. We refer to a specific spatio-temporal sample of the energy density, then, as E_i^n, where i indexes the spatial samples and n indexes the temporal samples.

The second order continuous spatial and temporal derivatives are approximated with second order central finite differences in time and space. Ignoring truncation error results in the following expressions that may be substituted into the telegraph equation:

$$\begin{aligned}\partial_{tt}E|_x^t &= \frac{E_i^{n+1} - 2E_i^n + E_i^{n-1}}{\Delta t^2},\\ \partial_{xx}E|_x^t &= \frac{E_i^n - 2E_i^n + E_i^n}{\Delta x^2}.\end{aligned} \tag{3.8}$$

First order derivatives are also approximated with a central finite difference. Again, ignoring the truncation error, the approximations are:

$$\begin{aligned}\partial_x E|_x^t &= \frac{E_{i+1}^n - E_{i-1}}{2\Delta x},\\ \partial_t E|_x^t &= \frac{E_i^{n-1} - E_i^{n+1}}{2\Delta t}.\end{aligned} \tag{3.9}$$

The first of these expressions is sufficient to discretize the spatial derivative that defines the boundary conditions at the ends of the domain in Equation 3.7. By defining the energy density at a so-called "ghost node" one spatial step beyond each end of the hallway that fulfills the Neumann condition in terms of the energy density at each of the two boundary nodes, the next time step may be computed.

Direct replacement of the continuous derivatives in Equations 3.5 and 3.7 with the approximations above results in a fully-defined explicit FDTD scheme for the interior and boundary samples. The final task is to solve for the energy density at the next time step for each element in the domain and the aforementioned ghost nodes in terms of the current and previous time steps as well as each node's neighbors. We may simplify the resulting system by defining C_r as the Courant-Friedrichs-Lewy coefficient, $c\Delta t/\Delta x$, and collecting common terms with $a = (A + D)c\Delta t/(2\lambda)$ and $b = AD(c\Delta t/\lambda)^2$.

We reproduce the resulting system of equations directly from Dujourdy et al. (2017), where E_1 and E_{nx} are the interior nodes at the boundary of the domain, and $2 < i < nx$ denotes all of the remaining interior nodes:

$$\begin{aligned}E_i^{n+1} &= \frac{E_i^{n-1}(a-1) + E_i^n(2(1-C_r^2) - b) + C_r^2(E_{i+1}^n + E_{i-1}^n)}{(a+1)},\\ E_1^{n+1} &= \frac{E_1^n(2[1 - C_r^2(1 + \frac{A_r D\Delta x}{\lambda})] - b) + 2C_r^2 E_2^n + E_1^{n-1}(a - 1 + A_r C_r)}{(1 + a + A_r C_r)},\\ E_{nx}^{n+1} &= \frac{E_{nx}^n(2[1 - C_r^2(1 + \frac{A_r D\Delta x}{\lambda})] - b) + 2C_r^2 E_{nx-1}^n + E_{nx}^{n-1}(a - 1 + A_r C_r)}{(1 + a + A_r C_r)}.\end{aligned} \tag{3.10}$$

3.2.4 Discussion

With this numerical approximation of the 1-dimensional EST, which upon application of boundary conditions along the length of the hallway (implicit in the the volume equation) and on the two extrema (explicitly solved in terms of a ghost node satisfying a Neumann boundary condition) reduces to a simulation of energy density throughout the space, we may proceed with our experiment.

As a final review of the previous work performed with this model, we note that stability conditions were derived in terms of α, β, the sampling steps Δx and Δt, and all other variable dimensional and physical factors, such as the mean free path and the speed of sound. We will not reproduce the Von Neumann analysis here, instead simply opting to always choose a higher sample rate than necessary for stability regardless of simulation parameters. As is noted in Table 1 of Dujourdy et al. (2017), for $\Delta x = 1$ meter, a mean free path $lambda = 2$ meters, each of which correspond to our simulation setup and the hallway under test, and considering maximal values for α and β of 0.9 and 1.9 respectively, the maximum time step Δt is 1.32 ms. Thus, as long as we stay within those parameters, we may simulate the system with any greater sample rate and be guaranteed stability.

Due to the choice of a spatial step of $\Delta x = 1$ meter in accordance with the characteristic wavelength of the EST, we have very few samples to compute, even for a very long hallway. By comparison, for example, a pressure simulation with cubic elements and the typical Von Neumann stability condition for second order 3-dimensional wave equation schemes requires a spatial step between elements $\Delta = \frac{c\sqrt{3}}{f_s}$, such that a simulation capable of representing a highest frequency of 3 kHz would require a sample rate of twice that, and thus necessitate over 10^3 finite elements to simulate a cubic meter of air, an even smaller volume than what is covered by a single element for a hallway of this width. Of course, the resolution of such a simulation is entirely different from that of the EST, but the point is that we should have no issue choosing a sample rate of $f_s = 1$ kHz or higher with so few elements, even if it gives us no further information. This efficiency is even more pronounced because this discretization should hold even for frequency bands that would be extremely demanding in the pressure-velocity domain: while the frequency-dependent parameters such as absorption and scattering may change, the modulation frequency of the EST is constant, meaning that it should be possible to directly model the stochastic reverberation at high frequencies without a change in the temporal or spatial discretization of the problem domain.

With specification of initial conditions, we have everything necessary to generate 1-dimensional numerical approximations of the energy density in the spaces covered by the dimensional reduction assumptions. For all cases in this chapter, the initial conditions were specified as a temporal Gaussian at the first spatial sample, corresponding (as will soon be noted) with the source position in the real hallways with which we compare the model. This

noticeably reduced ringing caused by introduction of a delta Dirac on the first sample as described in the previous studies. At $f_s = 1$ kHz, a Gaussian of length 7 samples results in a magnitude response with more than 50 dB of attenuation at the Nyquist limit.

Now we proceed to the exploration of the solution space covered by this numerical model in comparison with real spaces to determine the model validity as a function of frequency.

3.3 Measurements

The goal of this study was to understand if and how two different hallways could be modeled using the EST method for particular frequency bands. To that end, we decided to characterize the hallways as well as the numerical results in terms of two separate decay measures derived from impulse responses collected in situ or simulated with the EST method. As we have two free variables in the EST boundary conditions, we knew that two measures would be required to disambiguate them. Given that α was introduced as an absorption coefficient and β was introduced as a scattering coefficient, the measures were chosen to reflect the expected influence of each term; specifically, the change in reverberation time as a result of changing absorption, and the change in the distribution of energy throughout a space as a result of changing the amount of scattering.

To that end, we made physical measurements of impulse responses in multiple locations in each of the two hallways, and simulated impulse responses in two numerical models corresponding to each hallway's geometry. From those impulse responses The first measure, the temporal decay of an impulse response, is the well-known frequency-dependent reverberation time, or T_{30}. Thus, for each measurement position and each frequency band, whether physical or numerical, using Schroeder's reverse integration approach, as in Schroeder (1965), a linear fit to the impulse response's energy decay curve was used to determine the temporal decay rate, regardless of the overall level of the response at a given position. The second measure we compared we called the spatial decay, which is the slope of the total energy of each bandlimited impulse response in terms of the distance of the measurement along the length of the hallway. The point of this measure is to describe the distribution of energy in the hallway, but it cannot disambiguate entirely the effect of absorption compared to scattering.

For example, given two geometrically identical hallways with differing absorption, we hypothesize that the space with more absorption would exhibit a larger spatial decay, as each reflection would lose a greater portion of its energy as it travels down the hallway. Conversely, in the case of two hallways that are identical in absorptive properties, but differ in the amount of scattering, we hypothesize that increasing scattering would also increase the magnitude of spatial decay, as a greater portion of the sound energy would be scattered back toward its original angle of incidence, thus concentrating the strength of the field near

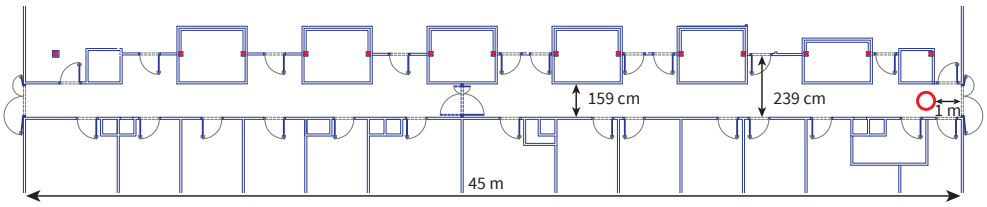

Figure 3.1: Floorplan, alcove hallway

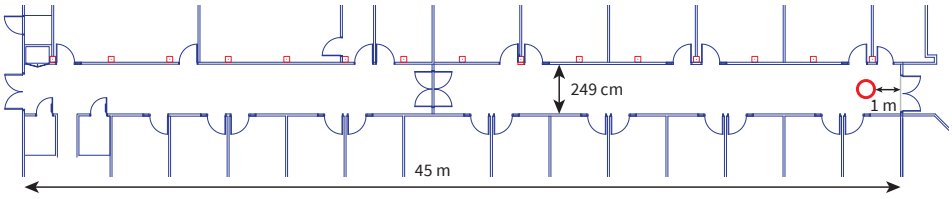

Figure 3.2: Floorplan, plain hallway

the source and weakening it further away. Nonetheless, we expect the two measures to be somewhat independent, allowing us to probe the individual effects of α and β.

3.3.1 Geometry

Both hallways under study were of length 45 meters. The first hallway was 159 cm wide and 237.5 cm tall, shown in plan view in Figure 3.1. Furthermore, it contained alcoves along the north wall of depth 80 cm and height 220 cm. The second corridor was 249 cm wide and 248 cm tall at the apex of its slightly curved ceiling, falling to 228 cm at each side wall. It is shown in plan view in Figure 3.2.

3.3.2 Details

The physical measurements were performed using a *SoundField* ST250 microphone, an *Outline* GRS omnidirectional speaker, and a *MOTU* Traveler sound card. The source was positioned 1 meter away from the end of the hall in each case, 1.5 meters above the ground, and centered between the two walls. Beginning 1 meter from the source, a spacing of 1 meter was used out to 10 meters to correspond with the simulation discretization given below, followed by a spacing of every 2 meters to the end of each hallway for a total of 26 sampling locations. Recordings were made using the the swept sine method (Farina (2000), Farina (2007)) as implemented in the *Adobe Audition* plugin *Aurora*.

For the corresponding simulations, the spatial sampling step was chosen to be $\Delta x = 1$ meter based on the known modulation frequency of the diffuse field. For stability, as previously discussed, the temporal sampling step was set to 1 ms. The the full range up to

Table 3.1: T_{30}s and spatial decays, alcove hallway

Frequency [Hz]	62.5	125	250	500	1000	2000	4000	8000
T_{30} [S]	2.29	0.35	0.38	0.39	0.43	0.35	0.33	0.30
Spatial Decay [dB/m]	0.63	0.49	0.88	0.94	0.66	0.54	0.66	0.84

Table 3.2: T_{30}s and spatial decays, plain hallway

Frequency [Hz]	62.5	125	250	500	1000	2000	4000	8000
T_{30} [S]	2.31	1.45	0.64	0.73	0.73	0.45	0.34	0.25
Spatial Decay [dB/m]	0.04	0.19	0.07	0.10	0.40	0.36	0.54	0.69

the stability criteria of both the absorption and scattering coefficients α and β was simulated, such that α ranged from 0 to 0.9, and β ranged from 0 to 1.9. Nonetheless, Figure 3.10 and Figure 3.11 display only the region with the most relevant combinations of coefficients for each hallway. In the hallway with alcoves, the absorption coefficient α ranges from 0.01 to 0.45, and the scattering coefficient β ranges from 0.01 to 0.5, whereas in the plain hallway, the absorption coefficient α ranges from 0.01 to 0.2, and the scattering coefficient β ranges from 0.01 to 0.25.

3.4 Results

3.4.1 Physical hallways

The measured T_{30}s and Spatial Decays for each hallway, diffusing and not, are given in Tables 3.1 and 3.2, respectively. From this table, Schroeder's frequency (Schroeder (1996)) can be computed for each hallway. For the hallway with alcoves, $f_{\text{Schroeder}} = 2000\sqrt{\frac{T_{30}}{V}}$, where V is the volume of the hallway in cubic meters, predicts a crossover frequency of 81 Hz, whereas the plain hallway's crossover frequency is predicted to be 146 Hz.

The energy sums at each measurement position are given in Figures 3.3 and 3.5, respectively, and the T_{30}s are shown in Figures 3.4 and 3.6 to demonstrate their relative consistency along the length of the hall. In each case, multiple measurements were averaged at each position to ensure accuracy.

3.4.2 Numerical model

An example simulation of the energy density in the alcove hallway is given in Figure 3.7.

Figures 3.8 and 3.9 show the T_{30}s and spatial decays for all of the combinations of α and β that produced physically meaningful results. That is, extremely long temporal decays as

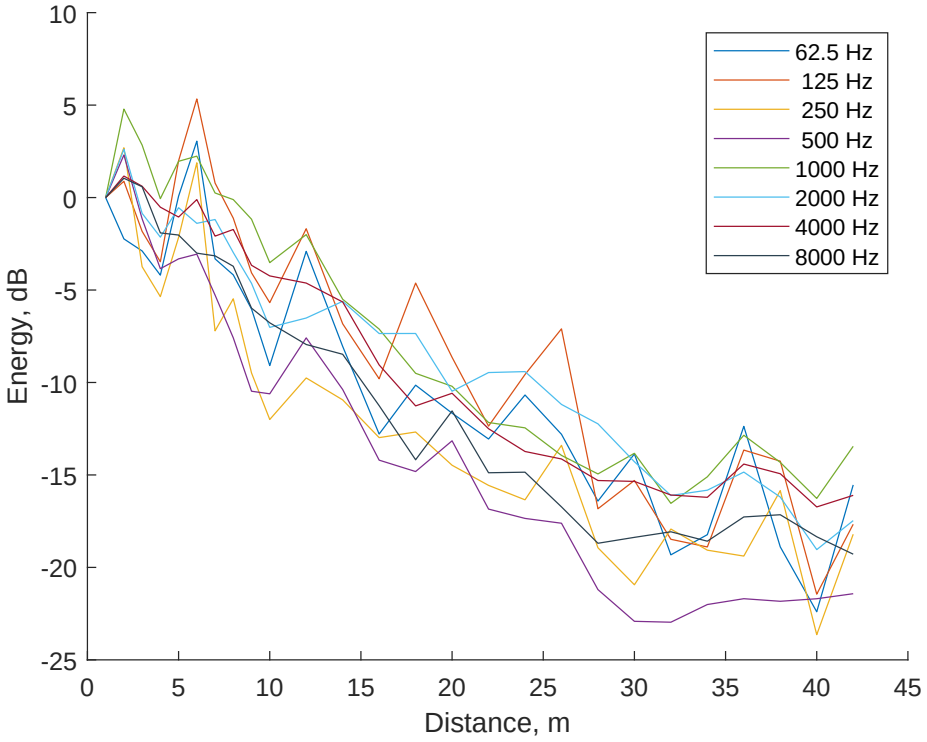

Figure 3.3: Energy measurements, alcove hallway, relative to first (closest) energy summation

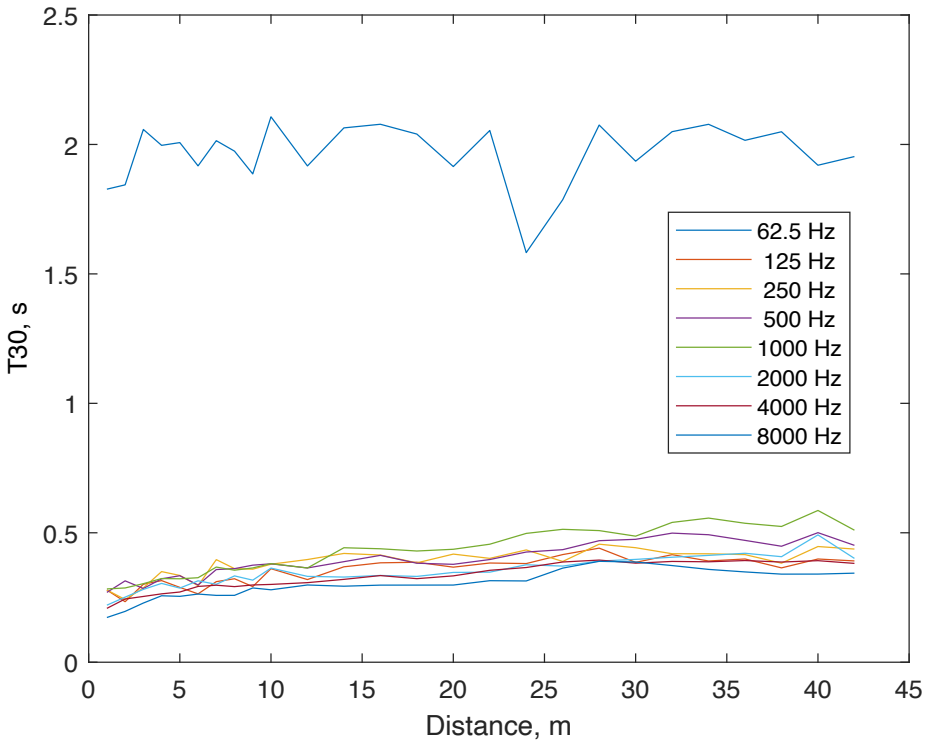

Figure 3.4: T_{30} measurements, alcove hallway

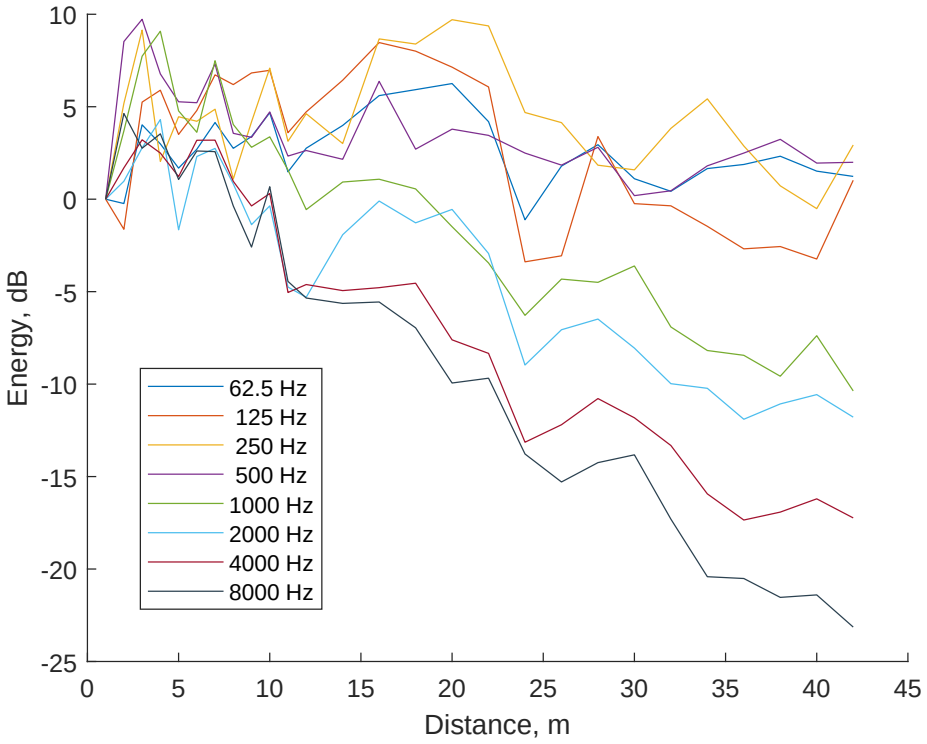

Figure 3.5: Energy measurements, plain hallway, relative to first (closest) energy summation

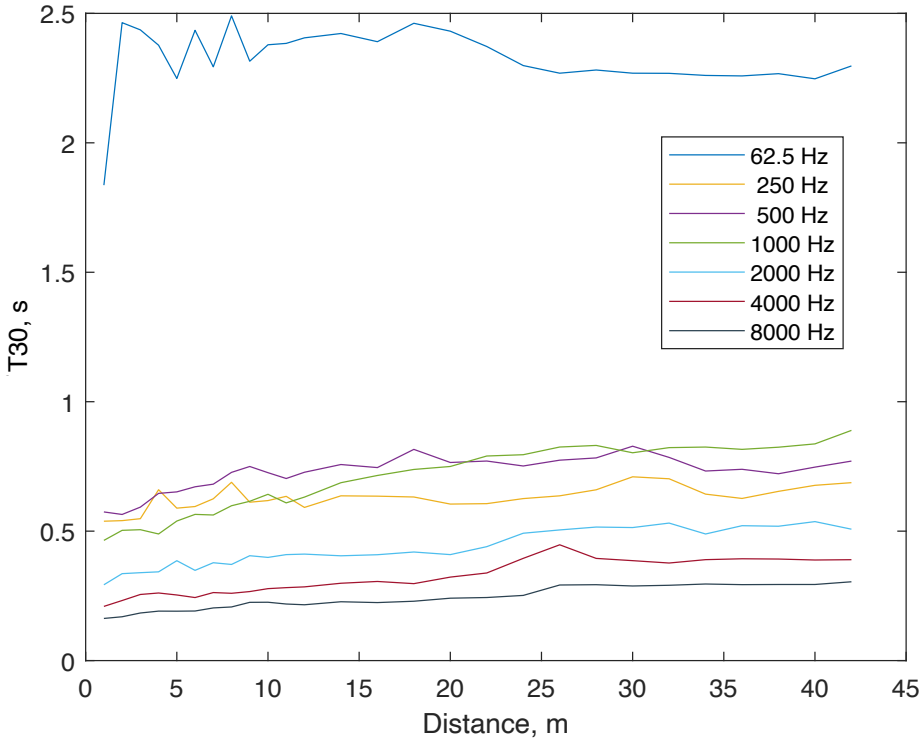

Figure 3.6: T_{30} measurements, plain hallway (note that the 125 Hz octave band has been removed due to spurious values, but was confirmed to adhere to the value given in Table 3.2)

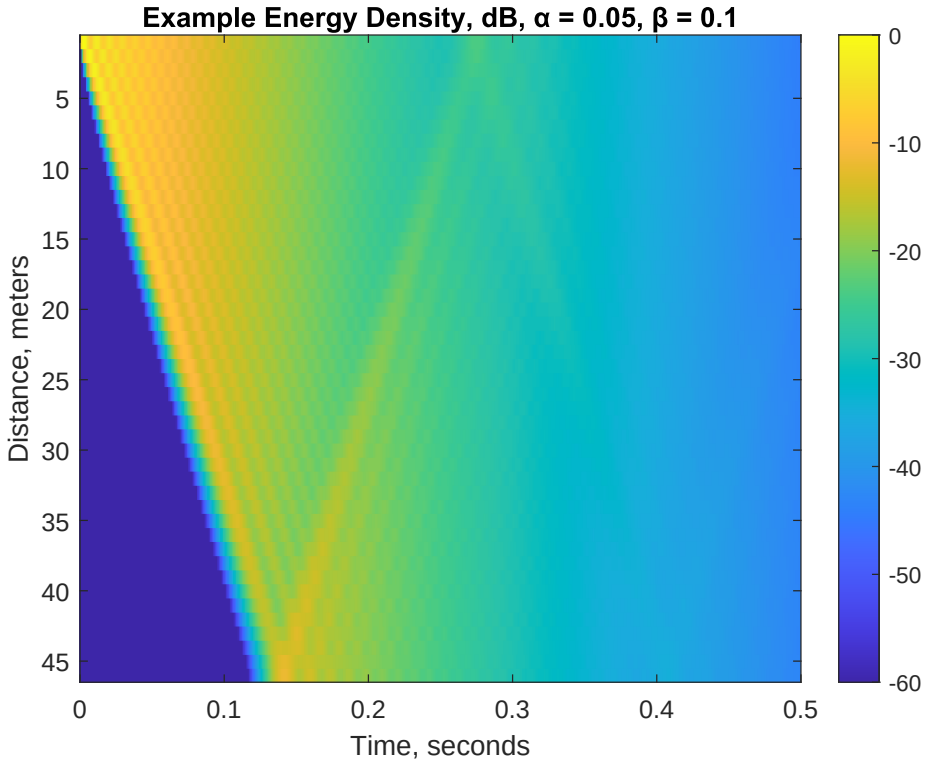

Figure 3.7: Example numerical result with energy density in dB, scaled to the initial value

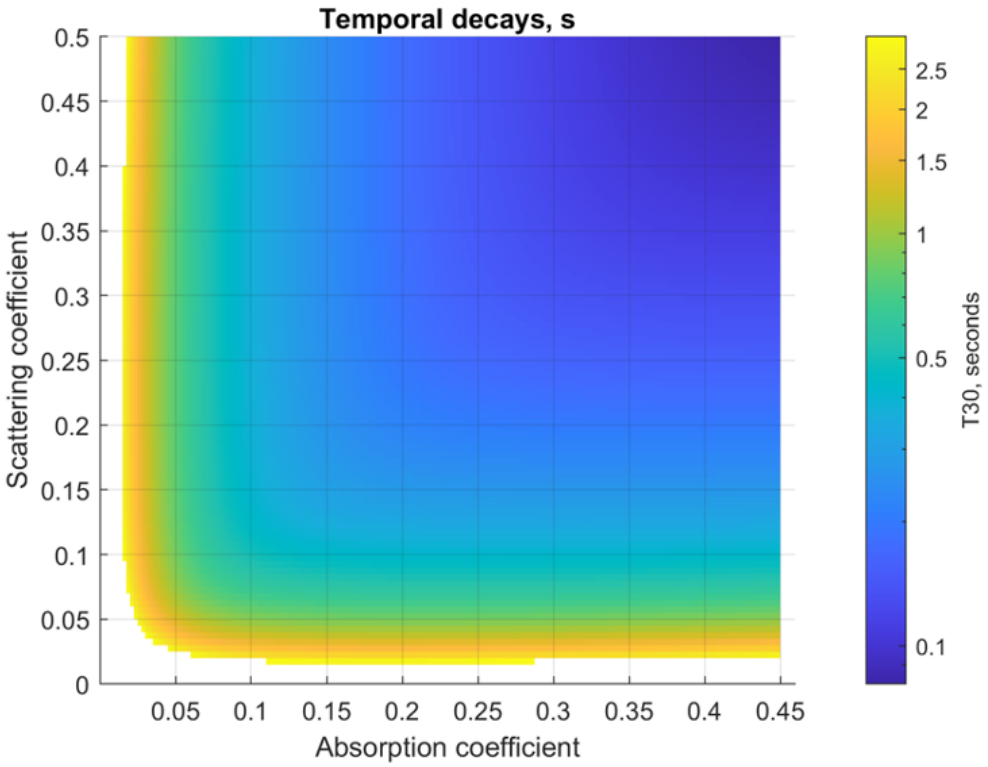

Figure 3.8: Contours of temporal decay for relevant combinations of absorption and scattering coefficients, alcove hallway; the scale is reverberation time in seconds

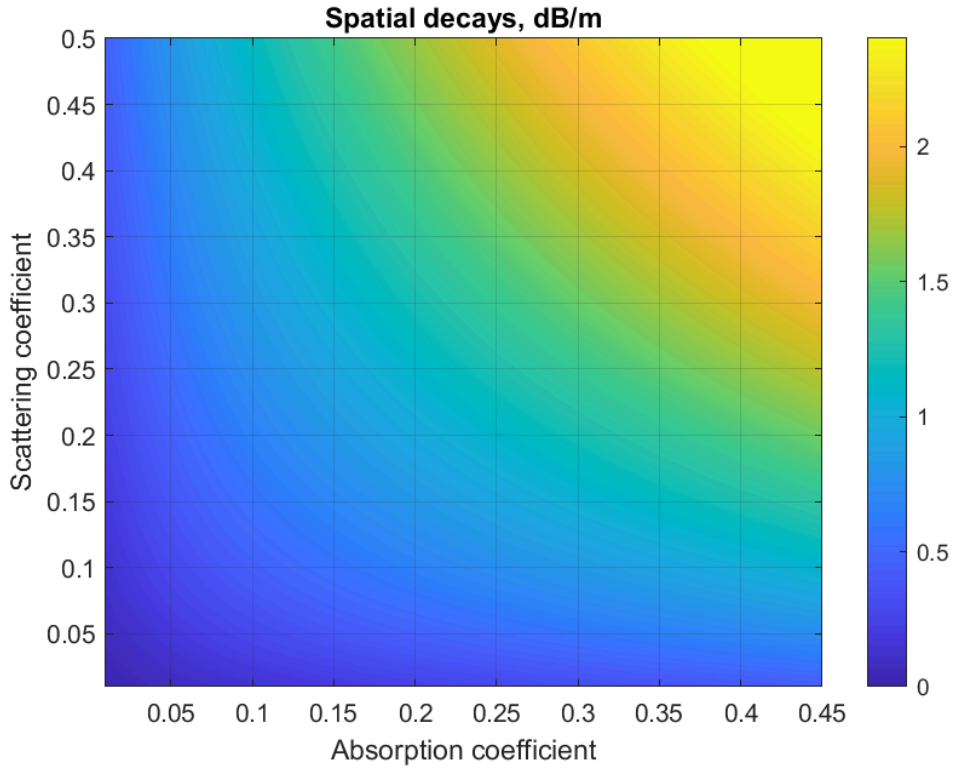

Figure 3.9: Contours of spatial decay for relevant combinations of absorption and scattering coefficients, alcove hallway; the scale is decay slope in dB / meter

a result of the model becoming increasingly lossless are discarded. In Figure 3.8, we have truncated the longest T_{30} to 3 seconds for readability, as they increase exponentially as the smaller of the two coefficients approaches zero.

Since we want to represent one frequency band with a single combination of coefficients, we can then compare these values with the desired temporal and spatial decays from each of the measured hallways. In this regard, we are essentially looking for equipotentials on the surfaces within a particular threshold from the target value, which shows the combinations of coefficients which give a "valid" result for that desired characteristic. Thus, for a given frequency band in a given hallway, we have two regions of validity in terms of α and β: one with the corresponding temporal decay, and one with the corresponding spatial decay. Finally, one may superimpose these two regions, and if any overlap is found, then this implies that there exists a pair of coefficients that is capable of representing both spatio-temporal characteristics for that frequency band.

Figures 3.10 and 3.11 show the regions corresponding to combinations of coefficients that produced the same acoustic indices as the measured hallways, for both the diffusing and plain corridors, respectively, as described above. For example, in Figure 3.10, 250 Hz, the blue band demonstrates the range of absorption and scattering coefficients for the simulation that produced a T_{30} within 10% of the measured value, the red band shows the spatial decays that were within 10% of measurements, and the green regions are the combinations of coefficients that fulfill both criteria and could therefore be considered as valid coefficients for representing the room in that frequency band.

3.4.3 Alcove hallway

We found that in the hallway with alcoves, every frequency band including and above the octave centered at 250 Hz had at least one region of validity where the combination of coefficients reproduced the measured spatial and temporal decay rates.

In bands where the regions of validity for each individual measure cross, resulting in two disconnected regions or one significantly overlapped region of model validity, the appropriate combination of coefficients can be determined without the need of a threshold. One exception is the 2000 Hz band, which is similar in some regards to the 125 Hz band, but falls within our specified error, indicating that even though we expect the field to be highly diffuse in that frequency range, we may not capture all of the behavior with the two measures we have chosen.

One further observation regarding the validity patterns is that the two coefficients are nearly symmetric, but not exactly, likely as a result of the small difference in the boundary conditions. Otherwise, in every valid frequency band for the alcove hallway, α and β are effectively interchangeable due to the fact that the smaller of the two drives losses in the

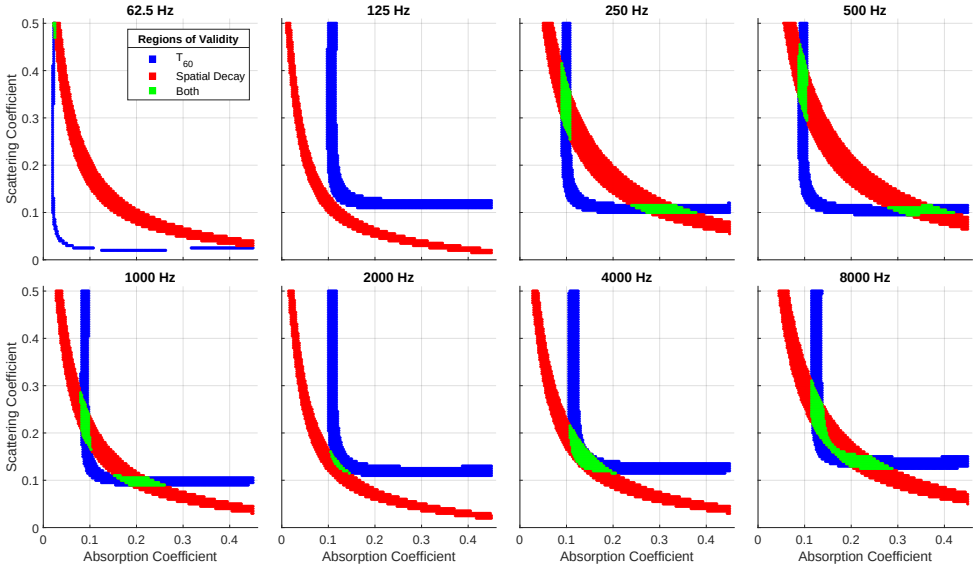

Figure 3.10: Simulated and measured data agreement, alcove hallway

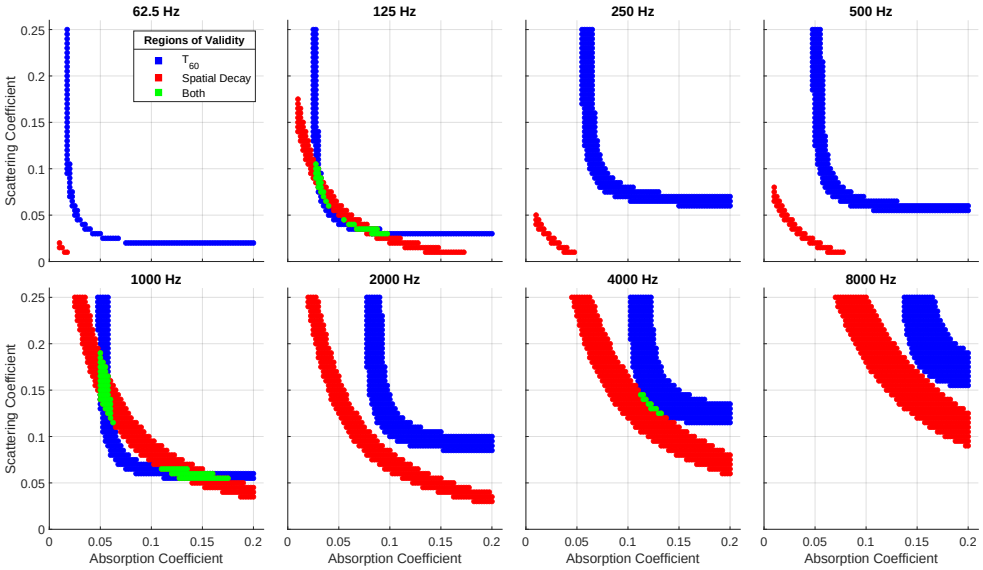

Figure 3.11: Simulated and measured data agreement, plain hallway

simulation, whether by absorption or redistribution.

3.4.4 Plain hallway

In the case of the plain corridor, only the octave bands centered at 125 Hz, 1000 Hz, and 4000 Hz exhibited the same model validity criteria. Since the model was not capable of matching both indices in a number of bands, regardless of the combination of coefficients, the soundfield must not be representable by the EST method. One possible explanation for the specific mechanism of invalidity is that this hallway did not demonstrate a consistent rate of spatial decay along the length of the hallway in some bands, suggesting more modal behavior. Since diffusing surfaces often act to break up strong, specular wavefronts, it could be that the lack of diffusion in this space reduced the amount of sound energy directed into the walls to be absorbed, resulting in a soundfield that was dominated by a modal effect.

This is particularly noticeable when comparing Figures 3.3 and 3.5. Every band in the corridor with alcoves exhibits spatial decay, evidenced by the decreasing energy compared to distance for all of the curves. For the non-diffusing hallway, however, certain octave bands exhibit a lack of decay, notably those at 62.5, 250, and 500 Hz, as well as 125 Hz to a lesser extent. One may argue that in fact, in the higher frequency bands, the validity of the model actually becomes quite marginal, rather than being rejected outright, suggesting that the model may still be capable of representing the feeble stochastic portion. In essence, while the stochastic reverberation was present at higher frequencies, the decay may have been overshadowed by the direct energy in the chosen measures due to a lack of temporal windowing.

3.5 Discussion

Aside from the specific question of validity, a number of observations can be made regarding the results of the simulations and fitting procedure. Often in room acoustics, temporal decays are the primary measure of interest; however, this focus can hide the relevance and behavior of other measures. For example, one takeaway from Figure 3.10 is the shape of the T_{30} curves themselves with regard to the scattering coefficient. If one ignores the spatial decay requirement and focuses purely on the temporal decay, one may observe that for the particular value of absorption coefficient that admits the widest range of scattering coefficients, that there is a minimum value for the scattering coefficient such that the temporal decay is satisfactory, but beyond that point, any scattering coefficient produces effectively the same temporal behavior.

This phenomenon has been noted in the context of simulation packages implementing

scattering, such as CATT-Acoustic and ODEON. This illustrates the idea that some minimum amount of scattering is necessary to accurately predict the reverberation, but as with the EST model, above that value, the temporal behavior of the response does not significantly change. What this analysis may be lacking, however, is the change in spatial characteristics that is shown in these examples, demonstrating that it is still critical to identify the correct scattering coefficient and not to overshoot it while focusing solely on the reverberation time at a particular receiver position. Instead, as shown in the results above, examining the spatial decay curves demonstrates that changes in scattering coefficient affect the spread of energy throughout a room (particularly for the stochastic reverberation) much more strongly across the range of values, meaning that it cannot be chosen arbitrarily.

3.6 Future work

As evidenced by the difference in behavior between the diffusing and non-diffusing hallways, it is important to note that the present analysis was performed using the entirety of the impulse responses rather than attempting to window in on a portion known to have more stochastic behavior. While in the case of the diffusing hallway, the stochastic reverberation was sufficiently dominant to allow the analysis to succeed above Schroeder's frequency, it remains to be seen if eliminating the direct path and early reflections from the analysis would allow matching indices even in rooms where modal or specular behavior is more prominent. It may even be possible to eliminate coherent reflections using matching pursuit approaches, as in Defrance et al. (2009) or Gribonval et al. (1996), leaving only the diffuse portion to analyze.

3.7 Conclusion

In this section, the validity of the EST method was demonstrated, finding a lower bound on the frequency range that it can be used in, at least for sufficiently diffusing 1-dimensional spaces. Furthermore, the strategy of evaluating bandlimited regions of the stochastic reverberation, which has implications for an analysis / synthesis approach to auralization of EST results, was presented and shown to be effective in the regions of validity for the corridor under study. Later in the thesis, we will use these results, both in terms of the frequency bands of validity as well as the derived coefficients, to inform an auralization of stochastic reverberation. Finally, while the approach regarding the transformation of the transmission line equations to the telegraph equation, subsequent solving for the boundary conditions at the ends of the hallways, and discretization with a centered time-and-space scheme is common in numerical methods, next we explore a different formalism that allows us to represent boundary conditions incorporating absorption and scattering without having to integrate

them into the volume equations for the 1-dimensional case.

Part II

Finite Volume Approaches

Chapter 4

Sources and Finite Volume Formulation

4.1 Introduction

In this chapter, we present a formulation of the EST including a source term and present a different numerical approach to solving the system. We have two goals in comparison to the previous chapter driving our development. First, while the definition of an initial value problem (or Cauchy problem) is often sufficient for room acoustics as it allows for the computation of an impulse response, the inclusion of a source term allows the representation of time-varying source characteristics such as directionality or movement. Of course, many of these qualities could be simulated by simulating many initial value problems and interpolating between them, however, sometimes such precomputation is not possible or preferable if a straightforward computation of the resulting field can be accomplished without much more computation. The second goal is the discretization of the EST equations with the finite volume time domain (FVTD) formalism, which has certain advantages over the more common finite difference time domain (FDTD) approach as used in the previous chapter with the telegraph equation. Furthermore, in doing so, we are interested in seeing if it is possible to define boundary conditions for the problem without integrating them directly into the volume equations, but defining them strictly at boundary surfaces to allow for easier representation of spatially-varying absorption or scattering coefficients.

FVTD approaches have drawn research interest in recent years because the formalism allows derivation of stability conditions directly from conservation of energy in the entire problem domain, which is guaranteed to machine precision, compared to Von Neumann analysis for FDTD approaches which fail on some irregular meshings. The primary difference between these is that a FVTD approach is formulated in terms of fluxes through surfaces of

adjacent cells (the "finite volumes" in question), enabling both a straightforward approach to implementing unstructured meshes, that is, domain meshings that are irregularly patterned, as well as the aforementioned conservative properties as a result of tracking the changes between adjacent cells or the boundary quantities. As demonstrated in Bilbao et al. (2016), the use of unstructured meshes is particularly important in a room acoustical context for refinement of the domain boundaries, not for general voxelization of a particular space (which may be tessellated regularly as in an FDTD approach). In this case, since the space under study is entirely rectilinear, the FVTD approach is, practically speaking, equivalent to a standard FDTD approach; however, the formulation is nonetheless particularly useful, as will be seen in Chapter 6.

To put the difference another way, while FDTD approaches are almost always defined over a regular meshing, that is, with a constant spatial step, FVTD approaches allow for a much more flexible meshing of the problem domain, which can play an important role in accurately reproducing soundfields when curved surfaces are involved. To the contrary, the so-called "staircase approximation" necessitated by a purely regular meshing can cause issues with regards to absorption due to the fact that even as element size decreases, the surface area of the regular mesh boundaries is decoupled from the true surface area of the domain to be modeled. While mitigations for this problem exist, like that of a weighting factor based on the voxelization error as presented in Hamilton (2021), it is nonetheless preferable to simply match the boundary where possible. A secondary side benefit is that because these refinements are only required to occur at the boundaries, the interior of the problem may be regularly meshed, as in the FDTD case, which greatly simplifies post-hoc analysis (even if it comes with numerical challenges such as dispersion). This makes it straightforward to parallelize the computation on the interior with a GPU, for example, while leaving the irregular boundary components to be handled in a serial fashion on a coprocessor more suited to the problem.

Regardless of whether a system of differential equations is solved with an FDTD or FVTD approach, they are commonly presented as initial value problems, whereby specifying initial conditions and letting the unforced system evolve throughout time produces the desired solution. This is particularly useful as it is ideally suited for generating impulse responses, which could be modeled as simply as a perturbation at a single cell, or with a more complex distribution derived from a measured or simulated directional response, as in Bilbao and Ahrens (2020). When considering the forced case, however, phenomena such as feedback or time-varying source conditions benefit from the inclusion of a source term in the wave equation and subsequent discretizations. To that end, we return to Equations 2.6 and 2.9 in order to develop sources in a finite volume scheme for the EST method.

The full document referenced in this chapter (Meacham et al. (2019a)) is reproduced in Appendix A.2.

4.2 Sources

4.2.1 1-dimensional EST

Before proceeding to the 3-dimensional definition of sources for the EST, let us first return to a 1-dimensional system with Lagrangian given by Equation 2.2:

$$L = \frac{\rho}{2c^2}(\partial_t \Psi)^2 - \frac{\rho}{2}(\partial_x \Psi)^2, \tag{4.1}$$

where Lagrange's equation gives

$$\frac{d}{dt}(\partial_{(\partial_t \Psi)} L) + \frac{d}{dx}(\partial_{(\partial_x \Psi)} L) = 0. \tag{4.2}$$

In Chapter 2, we demonstrated the continuity equations in three dimensions, and here we repeat the process in the 1-dimensional case to ensure our system is consistent. By definition, we may define the EST in terms of the Lagrangian

$$\begin{aligned} \underline{\underline{T}} &= \begin{pmatrix} \partial_t \Psi \partial_{(\partial_t \Psi)} L - L & \partial_x \Psi \partial_{(\partial_t \Psi)} L \\ \partial_t \Psi \partial_{(\partial_x \Psi)} L & \partial_x \Psi \partial_{(\partial_x \Psi)} L - L \end{pmatrix} \\ &= \begin{pmatrix} T_{tt} & T_{tx} \\ T_{xt} & T_{xx} \end{pmatrix}. \end{aligned} \tag{4.3}$$

Writing out the terms in full, we have

$$\begin{aligned} T_{tt} &= \frac{\rho}{2c^2}(\partial_t \Psi)^2 + \frac{\rho}{2}(\partial_x \Psi)^2, \\ T_{tx} &= \frac{\rho}{c^2}\partial_t \Psi \partial_x \Psi, \\ T_{xt} &= -\rho \partial_x \Psi \partial_t \Psi, \\ T_{xx} &= \frac{-\rho}{2c^2}(\partial_t \Psi)^2 + \frac{-\rho}{2}(\partial_x \Psi)^2. \end{aligned} \tag{4.4}$$

First, we check if the continuity equation $\nabla \cdot \underline{\underline{T}} = 0$ is validated. Keeping in mind the partial derivatives of the Lagrangian with respect to time and space using the chain rule,

$$\begin{aligned} \partial_t L &= \partial_{tt} \Psi \partial_{(\partial_t \Psi)} L + \partial_{tx} \Psi \partial_{(\partial_x \Psi)} L, \\ \partial_x L &= \partial_{tx} \Psi \partial_{(\partial_t \Psi)} L + \partial_{xx} \Psi \partial_{(\partial_x \Psi)} L, \end{aligned} \tag{4.5}$$

we can tackle the first line of the tensor.

$$\begin{aligned}
\partial_t T_{tt} + \partial_x T_{xt} =& \partial_t(\partial_t\Psi\partial_{(\partial_t\Psi)}L - L) + \partial_x(\partial_t\Psi\partial_{(\partial_x\Psi)}L) \\
=& \partial_{tt}\Psi\partial_{(\partial_t\Psi)}L + \partial_t\Psi\frac{d}{dt}(\partial_{(\partial_t\Psi)}L) - \partial_{tt}\Psi\partial_{(\partial_t\Psi)}L \\
& - \partial_{tx}\Psi\partial_{(\partial_x\Psi)}L + \partial_{tx}\Psi\partial_{(\partial_x\Psi)}L + \partial_t\Psi\frac{d}{dx}(\partial_{(\partial_x\Psi)}L) \\
=& \partial_t\Psi\frac{d}{dt}(\partial_{(\partial_t\Psi)}L) + \partial_t\Psi\frac{d}{dx}(\partial_{(\partial_x\Psi)}L) \\
=& \partial_t\Psi(\frac{d}{dt}(\partial_{(\partial_t\Psi)}L) + \frac{d}{dx}(\partial_{(\partial_x\Psi)}L)) \\
=& 0.
\end{aligned} \tag{4.6}$$

where we recognize that the penultimate expression is the wave equation.

Similarly for the second line, we again recognize that the result is the wave equation, validating the continuity equation for the 1D EST.

$$\begin{aligned}
\partial_t T_{tx} + \partial_x T_{xx} =& \partial_t(\partial_x\Psi\partial_{(\partial_t\Psi)}L) + \partial_x(\partial_x\Psi\partial_{(\partial_x\Psi)}L - L) \\
=& \partial_{tx}\Psi\partial_{(\partial_t\Psi)}L + \partial_x\Psi\frac{d}{dt}(\partial_{(\partial_t\Psi)}L) + \partial_{xx}\Psi\partial_{(\partial_x\Psi)}L \\
& + \partial_x\Psi\frac{d}{dx}(\partial_{(\partial_x\Psi)}L) - \partial_{tx}\Psi\partial_{(\partial_t\Psi)}L - \partial_{xx}\Psi\partial_{(\partial_x\Psi)}L \\
=& \partial_x\Psi\frac{d}{dt}(\partial_{(\partial_t\Psi)}L) + \partial_x\Psi\frac{d}{dx}(\partial_{(\partial_x\Psi)}L) \\
=& \partial_x\Psi(\frac{d}{dt}(\partial_{(\partial_t\Psi)}L) + \frac{d}{dx}(\partial_{(\partial_x\Psi)}L)) \\
=& 0.
\end{aligned} \tag{4.7}$$

Now, we introduce a source term and repeat the process:

$$\frac{d}{dt}(\partial_{(\partial_t\Psi)}L) + \frac{d}{dx}(\partial_{(\partial_x\Psi)}L) = Q. \tag{4.8}$$

As the left hand side remains the same, we can see that the divergence with respect to each line is unchanged, except for the fact that the wave equation does not reduce to zero, but to Q. Therefore, by inspection, we may begin from the penultimate line of Equations

4.6 and 4.7:

$$\begin{aligned}
\partial_t T_{tt} + \partial_x T_{xt} &= \partial_t(\partial_t\Psi\partial_{(\partial_t\Psi)}L - L) + \partial_x(\partial_t\Psi\partial_{(\partial_x\Psi)}L) \\
&= \partial_t\Psi(\frac{d}{dt}(\partial_{(\partial_t\Psi)}L) + \frac{d}{dx}(\partial_{(\partial_x\Psi)}L)) \\
&= \partial_t\Psi Q = pQ, \\
\partial_t T_{tx} + \partial_x T_{xx} &= \partial_x\Psi(\frac{d}{dt}(\partial_{(\partial_t\Psi)}L) + \frac{d}{dx}(\partial_{(\partial_x\Psi)}L)) \\
&= \partial_t(\partial_x\Psi\partial_{(\partial_t\Psi)}L) + \partial_x(\partial_x\Psi\partial_{(\partial_x\Psi)}L - L) \\
&= \partial_x\Psi Q = v_x Q.
\end{aligned} \tag{4.9}$$

The extension to the 3-dimensional case follows with the same method, and remembering that $J = \frac{1}{c}E_{xt}$ as in Section 2.3.3 results in the following system of equations:

$$\begin{aligned}
\frac{1}{c}\partial_t E_{tt} + \nabla\cdot\mathbf{J} &= \mathcal{P} = \frac{pQ}{c}, \\
\frac{1}{c}\partial_t\mathbf{J} + \nabla\cdot\begin{pmatrix} E_{xx} & E_{xy} & E_{xz} \\ E_{yx} & E_{yy} & E_{yz} \\ E_{zx} & E_{zy} & E_{zz} \end{pmatrix} &= \mathbf{Q} = \rho\mathbf{v}Q.
\end{aligned} \tag{4.10}$$

We define our source Q to take the form of an arbitrary distribution of volume velocity emanating from a moving membrane into the room. Then, with p and $\mathbf{v}$ the pressure and particle velocity, we want to ensure that the signs for the directions of each quantity are chosen to facilitate a multipole expansion of the source. For example, a monopole source implies that Q and p have the same sign on both sides of the source, therefore Q and p are positive when the membrane is moving into the volume of the room. To the contrary, the net effect of $\mathbf{v}$ should be zero, which demonstrates that $\mathbf{v}$ is positive when the normal of the membrane pointing into the room is oriented in the positive direction of each axis. A dipole source, on the other hand, implies that Q and p are each positive on one side and negative on the other, whereas $\mathbf{v}$ does not change sign, validating our choice of directions.

Thus, for a monopole source, $\mathcal{P}$ is positive and $|\mathbf{Q}|$ is zero, whereas for a dipole, both $\mathcal{P}$ and $\mathbf{Q}$ are positive when $\mathbf{v}$ is positive, with $|\mathbf{Q}| = \mathcal{P}$.

For any arbitrary source, these quantities may be computed as a multipole expansion of the radiation pattern, decomposing its effects into $\mathcal{P}$ and $\mathbf{Q}$ along the primary axes. Modeling of this sort is common in acoustics, where radiation patterns resulting from geometry are often predicted with finite element approaches or edge diffraction models, as in Martin

and Svensson (2018), for example.

In this case, we disregard the coupling of the source and room volume and take an anechoic approach when considering input to the system. This is justified by the fact that we only expect the EST model to represent the stochastic reverberation, and in the frequency range where that assumption holds, namely where the room's eigenmodes are overlapping, the free-field radiation impedance is dominant. With this system, we are able to describe and simulate the stochastic reverberation from a large class of source types and distributions. Because we are able to use such large cells in the numerical simulation of the system, it is plausible for the system to accommodate the computation of changes to the source distribution in near real-time, while taking into account the envelope of $\mathcal{P}$ and $\mathbf{Q}$ from an audio rate input.

4.3 Finite volume model

By restricting our analysis to the same hallway described in Chapter 3, we can once again form a tractable system by dimensional reduction. With the previously derived energy and momentum balances, which continue to be valid given that our diffuse field assumptions nor room geometry have changed, we have the following set of assumptions that are valid everywhere within the space:

$$E_{tt} = E_{xx} = E,$$

$$E_{yy} = E_{zz} = 0,$$

and the following set which are valid specifically at the boundaries:

$$J_y = J_z = \frac{A}{4}E,$$

$$E_{xy} = E_{yx} = E_{xz} = E_{zx} = \frac{D}{4}J_x.$$

Thus, by beginning with the EST continuity equations, applying the assumptions on the interior of the problem domain, and introducing the source terms as previously derived, we have

$$\begin{aligned} \partial_t E + \partial_x J_x + \partial_y J_y + \partial_z J_z &= \mathcal{P}, \\ \partial_t J_x + \partial_x E + \partial_y E_{yx} + \partial_z E_{zx} &= \mathbf{Q}, \end{aligned} \tag{4.11}$$

keeping in mind that we only retain the intensity in the X-direction except at the boundaries.

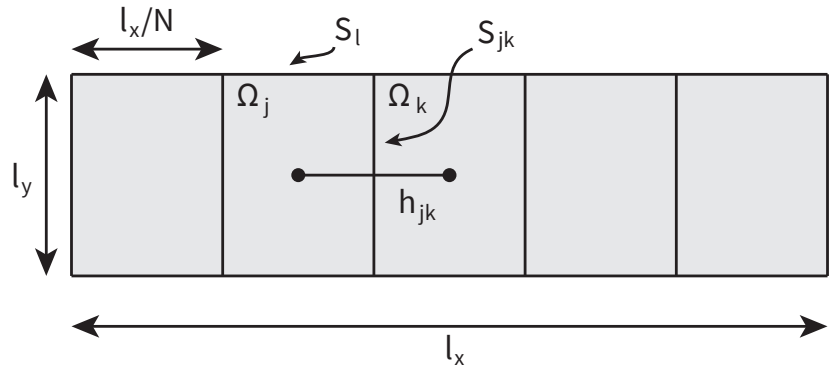

Figure 4.1: Plan view of an example finite volume discretization showing geometry subdivision, intercellular dimensions, and boundary surfaces

In order to close this system by rewriting the remaining terms with respect to the boundary assumptions, we must integrate the equations on discrete portions of the walls. In this case, however, rather than integrating over the cross-sectional area of the hallway, we instead choose to integrate over volume elements, as in finite volume time domain (FVTD) approaches, and then proceeding with discretization of the problem domain in space and time.

As with the FDTD approach used in Chapter 3, this process involves two steps: the subdivision of the domain into regions of finite extent, and secondly, the approximation of the temporal derivatives with difference equations. In this thesis, we use Bilbao et al. (2016) as the model for our discretization approach. The main idea is that rather than sampling points in the domain as with FDTD, the definition of finite regions allows the conversion of volume integrals to surface integrals by the divergence theorem, relating the quantities averaged over a particular volume to the in- and outgoing flows between neighboring volumes. With respect to room acoustics, this means that we may convert conservation equations to a summation of flows through neighboring volume connections. In this chapter, we will describe in detail the discretization process and apply it to our EST system, and in the next chapter, we will describe the boundary conditions for a pressure-velocity simulation, also formulated with the FVTD approach, which will be used for auralization and in the final chapter's direct computation of EST terms.

4.3.1 Spatial discretization

Now, we discretize the hallway into N sequential rectangular solids that will form the cells in our FVTD formulation. These cells Ω_j (of volume $V = l_x l_y l_z/N$, where the dimensions of the hallway are l_x, l_y, and l_z.Furthermore, additional dimensional quantities are computed

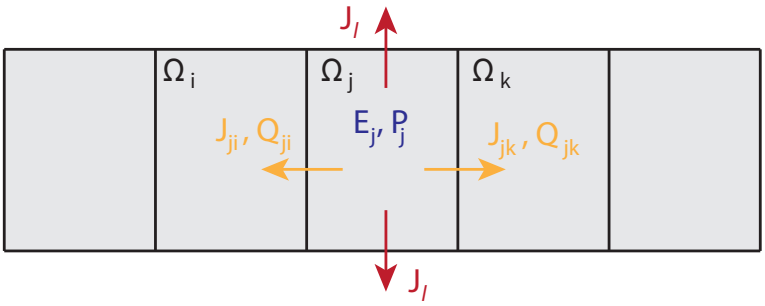

Figure 4.2: Plan view of an example finite volume discretization showing the quantities and directions corresponding to cell Ω_j

in terms of the domain dimensions: h, the distance between cell centroids, $S = l_y l_z$, the surface areas of cell-cell faces, S_l, the cell-boundary face surface areas, and V, the volume of each cell. For an example of such a discretization for a series of rectangular cells, a plan view is given with the various quantities in Figures 4.1 and 4.2.

Integrating Equation 4.11 on cell Ω_j allows us to associate an average energy density E_j with the cell, as well as outward sound intensities J_{ji}, J_{jk}, and J_l, depending if the intensity in question is incident upon another cell or a boundary, in terms of the corresponding surface areas, S_{ji}, S_{jk}, and S_l.

By inspection, since every J_l term corresponds to the incident sound intensity on a given boundary, we can see that they may all be redefined using the energy balance, that is, $J_l = \frac{A}{4}E$, corresponding to the system above. Similarly, as the second line of 4.11 must in fact be defined on a small volume enclosing face S_{jk}, the momentum flux terms may also be written in terms of their balance on the boundary, that is, $E_{yx} = E_{zx} = \frac{D}{4}J_x$. For symmetry, we consider this enclosing volume to be the elementary volume of the cells, V. Finally, by integrating $\mathcal{P}$ over each cell and averaging $\mathbf{Q}$ over each outward surface, the indexed quantities $\mathcal{P}_j$ and Q_{jk} are also associated with each cell and its neighbors.

Thus, by application of the divergence theorem, we may directly write the spatially discretized EST:

$$\begin{aligned} \frac{V}{c}\frac{dE_j}{dt} + \sum_{k=1}^{N}\beta_{jk}SJ_{jk} + \sum_{l=1}^{N_b}\gamma_{jl}S_l\frac{A}{4}E_j &= \mathcal{P}_j, \\ \frac{1}{c}\frac{dJ_{jk}}{dt} + \frac{1}{h}(E_k - E_j) + \sum_{k,l=1}^{N,N_b}\zeta_{jkl}\frac{S_l}{V}\frac{D}{4}J_{jk} &= Q_{jk}, \end{aligned} \tag{4.12}$$

where β_{jk} and γ_{jl} are indicator functions that are 1 if a given cell Ω_j shares a face with another cell Ω_k or face on the boundary S_l, and 0 otherwise. Additionally, ζ_{jkl} is an indicator function for a cell, boundary, and a neighboring cell that has a boundary that shares vertices with the first ("neighboring" boundaries, so to speak). This unique configuration, which differs from indicators found in Bilbao et al. (2016), is a result of the need to define a surface over which scattering may take place according to a parallel flow, as is the case with the contributions of sound intensity and normal stress on a given boundary. Finally, we note that this form is slightly simplified given a constant S_l, whereas the general case requires an average over the boundary surface areas for the final term of the second equation (the momentum balance expression).

While it is initially surprising that the summations run over all possible cells and boundaries, in practice, due to the indicator functions, the result is equivalent to a typical stencil-based algorithm where the computation at a particular cell is given in terms of its neighbors. The difference between this statement and a typical 7-point scheme (as would be the case for a meshing of regular hexahedral solids such as cubes) is that it has the flexibility to handle polyhedra with any number of faces, and in any configuration, including arbitrary unstructured tessellations. Of course, in this case, we will be using a rectilinear voxelization, so despite the change in appearance from a regular FDTD scheme, one may think of these terms in the same fashion as a stencil, giving the surface area and velocity between a cell and all of its neighbors on the interior as well as its neighboring boundary faces.

This approach has the advantage that it is very straightforward to apply different absorption and scattering coefficients at all boundaries without having to solve for a different boundary condition at the ends of the hallway, a fact that we will return to later, nor compute mean absorption and scattering coefficients.

4.3.2 Time domain discretization

Next, the continuous time functions can be discretized by replacing temporal derivatives with discrete approximations based on differences at successive time steps. A detailed description of this procedure is given in Bilbao et al. (2016) Section IV.

First, a discrete time approximation of the continuous time series $f(t)$ by sampling at a fixed time step T is proposed such that $f^n = f(nT)$. Then, the forward and backward shift, difference, and averaging operators are

$$\begin{gathered} e_+ f^n = f^{n+1}, e_- f^n = f^{n-1}, \\ \delta_+ = (e_+ - 1)/T, \delta_- = (1 - e_-)/T, \text{ and} \\ \mu_+ = (e_+ + 1)/2, \mu_- = (e_- + 1)/2, \end{gathered} \tag{4.13}$$

respectively.

We define the temporal approximations to energy density and sound intensity to be offset by half a sample, that is, E_j^n and $J_{jk}^{n+1/2}$, and where $\mathcal{P}_j$ and Q_{jk} are aligned in time with E_j and J_{jk} respectively. We replace the continuous derivatives in Equation 4.12 to find a fully discrete representation of the system.

$$\begin{aligned} \frac{V}{c}\delta_+ E_j + \sum_{k=1}^{N} \beta_{jk} S J_{jk} + \sum_{l=1}^{N_b} \gamma_{jl} S_l \frac{A}{4} \mu_+ E_j = \mathcal{P}_j, \\ \frac{1}{c}\delta_- J_{jk} + \frac{1}{h}(E_k - E_j) + \sum_{l=1}^{N_b} \zeta_{jkl} \frac{S_l}{V} \frac{D}{4} \mu_- J_{jk} = Q_{jk}. \end{aligned} \tag{4.14}$$

Temporal averaging is applied to preserve the time alignment and differential relationship in each equation.

Then, using the notation $e_\pm f = f^\pm$, expanding the temporal operators, and solving for E_j^+ and J_{jk} gives

$$\begin{aligned} E_j^+ = \frac{E_j(1 - \frac{cT}{V}\sum_{l=1}^{N_b} \gamma_{jl} S_l \frac{A}{8}) + \frac{cT}{V}\mathcal{P}_j - \frac{cT}{V}\sum_{k=1}^{N} \beta_{jk} S J_{jk}}{1 + \frac{cT}{V}\sum_{l=1}^{N_b} \gamma_{jl} S_l \frac{A}{8}}, \\ J_{jk} = \frac{J_{jk}^-(1 - \sum_{l=1}^{N_b} \zeta_{jkl} \frac{S_l}{V} cT \frac{D}{8}) + cTQ_{jk} + \frac{cT}{h}(E_j - E_k)}{1 + \sum_{l=1}^{N_b} \zeta_{jkl} \frac{S_l}{V} cT \frac{D}{8}}. \end{aligned} \tag{4.15}$$

This is a realizable two-step FVTD scheme.

4.4 Evaluation and commentary

The resulting scheme was compared to the alcove hallway as in Chapter 3 to determine whether the implementation was capable of representing the previously measured spatial and temporal decays, and if so, whether the regions of validity occurred with the same combinations of coefficients. Additionally, to test the source implementation, not only were the previous measurements with a monopole loudspeaker considered, but an approximation of a dipole source using only the drivers of the *Outline* GRS omnidirectional speaker aligned along the X-axis in opposite phase were measured and compared to simulations. The measurement procedure was identical to that in Section 3.3, with the exception of the measurement distances, which were 4, 8, 12, and 16 meters from the source. The simulations were also performed in a similar manner as before, with a brute force sampling of all possible modified absorption and scattering coefficients, the only difference being the inclusion of a

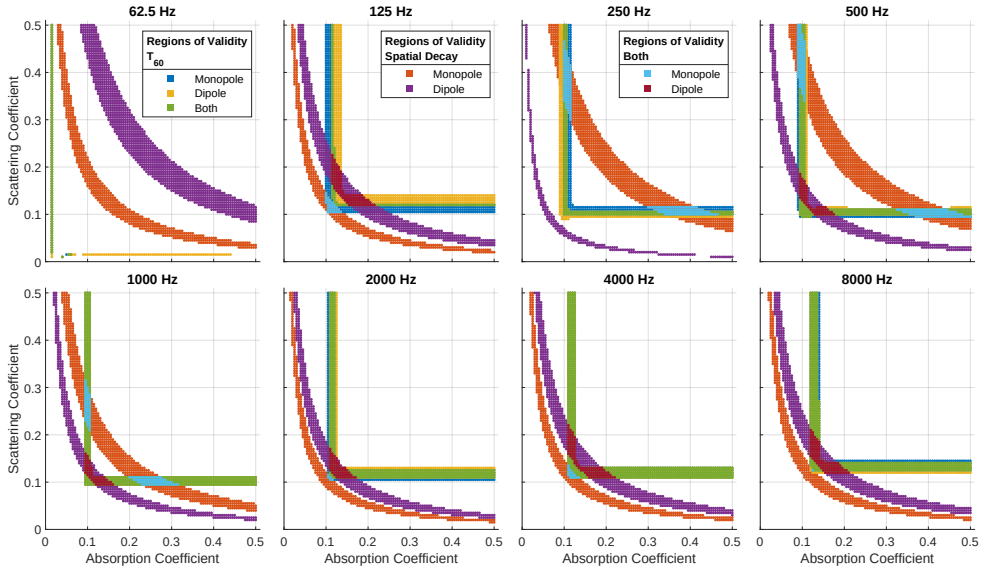

Figure 4.3: Simulated and measured data agreement, monopole and dipole configurations, where legend colors apply to all subgraphs to illustrate regions of validity for all combinations of configurations

dipole source term in the second case. Then, the same processing to compute temporal and spatial decay rates was followed for both the measured and simulated data. The results are shown in Figure 4.3. Note that the legends describing colors representing regions of validity apply to all frequency bands regardless of the figure they appear in. As before, regions where both measures are valid, whether for the monopole or dipole case, are highlighted by the light blue and maroon sections, respectively. Examination of the graphs shows that, with the exception of the 250 Hz octave band, where the dipole does not exhibit a region of overlap due to the lower spatial decay, both sources are valid including and above the 125 Hz octave band. Furthermore, comparison with Figure 3.10 shows that the regions of validity for the monopole source occur at similar pairs of coefficients, but are not exactly the same.

One possibility for this discrepancy is the formulation of boundary conditions in terms of the distribution of energy in front of a set of walls. In the original FDTD approximation of the telegraph equations, there were always strictly speaking 4 surfaces being integrated over: in Section 2.3.3, when it is asserted that a quarter of the energy density in the integrated section impinges upon each wall, the reason for that assertion is that there are four walls. At cells at the ends of the hallway, however, that is not the case, as the absorption of 5 surfaces must be taken into account. On the other hand, while the boundary conditions used in this section are treated as being the same on all surfaces including the end caps, and

the assumption that energy propagation is primarily along the length of the hallway was borne out by the configuration of cells, it may be the case that due to the 1-dimensional propagation assumption that a different boundary condition is in fact necessary for the surfaces at the ends of the hallway.

4.5 Conclusion

In this chapter, the EST method was reformulated in terms of an FVTD approach based on the velocity potential of a typical pressure-based wave equation. This enables differing geometrical topologies in the case of primarily one-dimensional spaces, like long hallways, as well as illustrating the difference in the effects of the modified absorption and scattering coefficients when implemented at all of the 3-dimensional boundaries without incorporation into the volume equations directly. Furthermore, by introducing a source term to the equations and describing its parameterization in terms of a common model for describing physical sources, we expand the EST method's capabilities in terms of the types of problems it may admit as well as clarifying the relationship between the sound intensity and the directionality of sources, even in this 1-dimensional case. Finally, while in most room acoustics situations an initial value problem is sufficient to represent time-invariant phenomena, as is the case with a single impulse response, interpolation between precomputed results of many source locations in a wave-based scheme may still be less efficient than recomputation of a new source, especially given the efficiency of the element sizes for the EST. Though we do not take advantage of it at this point, the inclusion of a source term in the formulation also admits the possibility of simulating moving or otherwise time-varying sources in a natural manner. With this new formulation, we now consider how it might be applied in terms of synthesizing stochastic reverberation from computed energy density profiles. While in this and the preceding chapter, matching acoustical indices such as rate of decay and spatial distribution of energy was the main source of information regarding the choice of simulation parameters in a given frequency band, we now wish to verify that those measures in fact satisfy our perception in terms of representing stochastic reverberation.

Chapter 5

Auralization

5.1 Introduction

In this chapter, a hybrid strategy for the auralization of EST-derived energy envelopes is proposed. The purpose of such an auralization tool is both pedagogical and practical, as it allows simultaneous verification of the methodology and implementation, while also serving as an example of the strengths of the EST approach.

The full document referenced in this chapter (Meacham et al. (2020)) is reproduced in Appendix A.3.

5.2 Hybrid Model

As discussed in Chapter 2, room acoustical models capable of synthesis are often well-suited for a particular frequential-temporal region of an impulse response. While this can be a challenge when it comes to reproducing an entire impulse response with a single method, it naturally leads to the idea of combining synthesized regions from different methodologies. Of course, this can lead to new challenges, including ensuring temporal alignment or guaranteeing agreement between energy levels for multiple different methods.

This implies a tradeoff between complexity and speed or accuracy. For example, while it would be ideal to simply run a wave-based simulation beyond the T_{30} and up to the highest audible frequency in order to properly capture modal behavior and diffraction effects, depending on the size of the problem, such a simulation is unlikely to be finished in less than a few hours, even on powerful hardware. One recent example of such a study is Fratoni (2021), whose collaboration with the University of Edinburgh resulted in results up to 4 kHz at a cost of an hour per second of calculated impulse response, albeit for very large halls. In contrast, even for the same length of impulse response, limiting the frequency range of such a

Frequency [Hz]	250	500	1000	2000	4000	8000
T_{30} [S]	0.38	0.39	0.43	0.35	0.33	0.30
Spatial Decay [dB/m]	-0.88	-0.94	-0.66	-0.54	-0.66	-0.84
Mod. Abs. Coef. A	0.051	0.051	0.047	0.055	0.058	0.067
Mod. Sca. Coef. D	0.54	0.59	0.38	0.25	0.32	0.41

Table 5.1: T_{30}s, spatial decays, and simulation parameters, alcove hallway

simulation to Schroeder's frequency (for example) and using much larger numerical elements could allow for simulations to be performed quickly enough to be considered "real-time," such that moving sources or other changes to simulation parameters could be recomputed and presented to a user without perceived latency. The question remains how to fill in the missing high-frequency data, for which another method suited to that region may be selected, often geometrical in nature.

In such a fashion, combinations of methodologies, carefully calibrated to work with each other in the regions they are individually best suited for, is an attractive way to create wideband impulse responses much more quickly and accurately than would be possible with a single method. Many recent auralization approaches, including Murphy et al. (2008), Oxnard and Murphy (2013), and Poirier-Quinot et al. (2017), use this approach as the basis for improving accuracy and speed. In this case, we propose a hybrid model integrating the EST method to represent and synthesize the stochastic reverberation such that the complementary wave-based and geometric methods are not required to generate data at high frequencies or with an exponentially growing number of reflections, respectively.

As in the previous chapters, we focused again on the diffusing hallway introduced in Chapter 3.

5.2.1 EST method

For this study, we used the formulation of the EST defined in Chapter 4 in order to determine its suitability within a hybrid context. As this portion had previously been validated for monopole and dipole sources in a long hallway with alcoves, we decided to reuse the coefficients determined in the fitting process for the hybrid simulation of the same space. These values can be seen in Table 5.1.

To convert the results from the EST simulation into an auralizable pressure domain signal, the resulting energy density in every cell can be used as an envelope for a noise process. Because the simulation is repeated for each frequency band, these envelopes can be applied to bandlimited noise in each of the specified bands, and summed up to create a wideband stochastic reverberation representation at each point in the hallway. This is similar to the process followed in Luizard et al. (2013), where the DEM was used, but for a

single uncoupled volume and within a hybrid context.

Because the EST method sample rate is much slower than audio rate, it is necessary to resample the envelopes to the final hybrid sample rate in order to ensure smoothness in the time domain. In the spatial domain, it is also important to interpolate between the cells as they are quite large. In practice, when sources are unchanging, the results can be cached, making it simple to move the receiver position.

5.2.2 Low-frequency reverberation

For low-frequency acoustic modeling, we implemented the finite volume time domain (FVTD) method from Bilbao et al. (2016) as in Chapter 4, but as in the original paper, in the pressure and velocity domain. The code for this portion of the hybrid response can be found online, detailed in Appendix B.2. As at the time the implementation was undertaken, there was no open-source implementation of the method available, so a pedagogical rather than performance-oriented approach was taken to assist others who are following along with the original article. Because of the work involved in implementing the method from scratch, we will detail it here with some specificity, giving an overview of the method itself as well as details on our implementation.

As mentioned before, while the primary advantage of a FVTD model is for accommodating fitted boundaries rather than a regular meshing for the entire problem domain (as is typically the case with FDTD models), here, its usefulness is due to the convenience of expressing certain aspects of the EST method within the formalism. Because the problem domain was entirely rectilinear, no edge fitting was required, and we accepted the small mismatch between the simulation mesh and the true dimensions of the hallway for simplicity's sake. Often, discretizing a particular space with cubic or rectangular solids is performed by generating a mesh for the bounding box of the problem domain, then including a cell if its centroid falls within the problem domain and excluding it if it does not. When a curved surface is discretized in such a fashion, the result is the so-called "staircase approximation," named for the stepwise appearance of the surface. As discussed in Bilbao et al. (2016), even as cells become smaller and the volumetric approximation becomes closer, the surface area approximation is somewhat divergent. Typically, the solution is to adopt fitted elements at the edge of the problem domain, an approach particularly well-suited to the FVTD formalism. In this case, however, while many spaces such as concert halls or theaters include large curved surfaces with important acoustical consequences, the geometry of the hallway under discussion has no such complications, and therefore we are comfortable eschewing edge fitting to eliminate staircase effects.

A secondary advantage of the approach is that by characterizing the flux on every cell face, a total summation of the energy may be performed such that the entire simulation is

provably lossless (to machine precision). Additionally, because of the shared architecture of the code for both the EST and pressure versions, engineering effort spent on one could be applied to the other.

Pressure-velocity FVTD derivation

$$\begin{aligned} \frac{1}{\rho c^2}\partial_t p + \nabla \cdot \mathbf{v} = 0, \\ \rho \partial_t \mathbf{v} + \nabla p = \mathbf{0}. \end{aligned} \tag{5.1}$$

The derivation of the low-frequency FVTD model begins from the pressure and velocity equations, equivalent to Equation 2.1. While in the context of the EST model, we started from the second order velocity potential version, here, it makes sense to begin in terms of pressure and velocity, as we would like to parameterize boundary conditions and cell-to-cell interactions in terms of these values to start.

Rather than seeking conservation equations for energy density and sound intensity as in Chapter 2, the total energy of the system can be defined in terms of pressure and velocity everywhere in the space, as well as the portion impinging on the problem boundaries, where no motion is expected to take place, but where dissipation will nonetheless occur. The passivity of terminating impedances and local reactivity is then used to develop a one-port network representation of boundary admittances that can be fit to empirical data in a straightforward fashion following familiar circuit synthesis procedures.

This system provides the basis for a conservative simulation of the acoustic field which is then discretized into a numerical method. As the procedure in Section 4.3.1 follows essentially the same process, but with differing boundary conditions, we risk some repetition here, but due to the difference in the handling of the dimensional reduction that was required previously, we consider it appropriate to cover the development again in terms of pressure and velocity.

Boundary conditions

Bilbao et al. (2016) use a one-port circuit model to match desired boundary admittances that relate the sound pressure and normal velocity everywhere at the edge of the problem domain in a locally reactive manner. The boundary admittance model must be capable of storing, reflecting, and dissipating energy, but cannot be allowed to add energy, making a passive network of resistors, capacitors, and inductors an appropriate model. In the simplest case of the circuit model, it reduces to a mass-spring-damper system common in boundary admittance representation, however, by adding M branches of series RLC circuits, more complicated admittances may be specified. This topology is shown in Figure 5.1. Adding the requirement that all components must be nonnegative means that at least one capacitor

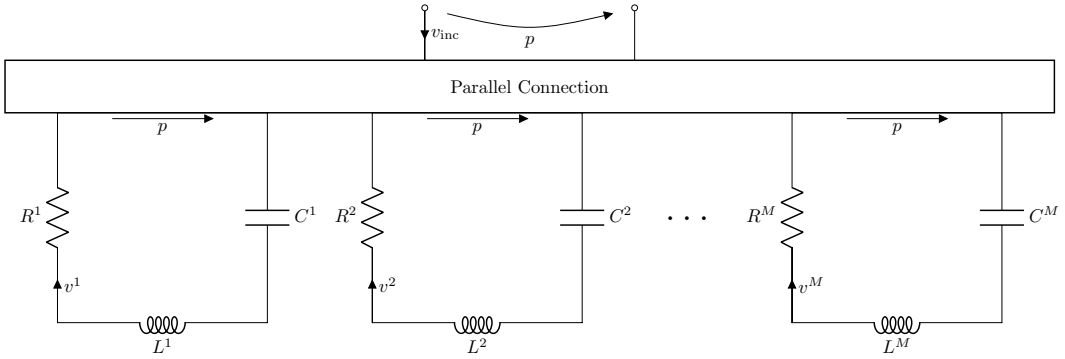

Figure 5.1: Boundary admittance model, after Bilbao et al. (2016), consisting of a one port parallel connection across the boundary pressure p with current equal to the boundary normal velocity v_{inc}.

ensures that the admittance at DC is zero, which prevents the possibility of a mean flow in the direction of the wall.

The admittance of such a network is

$$Y(\mathbf{x}, s) = \sum_{m=1}^{M} \frac{s}{L^m(\mathbf{x})s^2 + R^m(\mathbf{x})s + \frac{1}{C^m(\mathbf{x})}}, \tag{5.2}$$

where $\mathbf{x}$ describes all boundary coordinates and s is the typical Laplace transform variable. This admittance allows us to relate the sound pressure and normal velocity on the boundary, where all quantities are Laplace transformed to the s domain, as

$$\hat{v}_{\text{inc}}(\mathbf{x}, s) = Y\hat{p}(\mathbf{x}, s). \tag{5.3}$$

This formulation, with v^m as the current in the mth RLC branch, can be summarized as follows:

$$\begin{aligned} v_{\text{inc}} &= \sum_m = 1^M v^m, \\ p &= L^m \frac{d}{dt} v^m + R^m v^m + \frac{1}{C^m} g^m, \\ \frac{d}{dt} g^m &= v^m, \end{aligned} \tag{5.4}$$

where the additional variable g^m represents the energy storage in the capacitor. In short, the boundary conditions everywhere in the space may be implemented given an appropriate spatially varying admittance model and a time domain discretization of the temporal derivatives regarding the mth branch current and capacitor reactance.

Spatial discretization

As before, the problem domain is discretized with some freedom into N non-overlapping cells Ω_j, each with an average pressure p_j, and with outward normal velocities incident on another cell or a boundary, v_{jk} or v_l. All of the other cellular quantities (volume, face surface areas, intercellular distances, etc.) are identical to those previously presented, though they remain indexed on a cell-to-cell basis as they appear in the original paper for full generality. Then, using the same indicator function notation as before, where β_{jk} and γ_{jl} are 1 if a given cell Ω_j shares a face with another cell Ω_k or face on the boundary S_l, and 0 otherwise, the first order acoustic wave equation system may be written as

$$\begin{aligned} \frac{V_j}{\rho c^2}\frac{dp_j}{dt} + \sum_{k=1}^{N}\beta_{jk}S_{jk}v_{jk} + \sum_{l=1}^{N_b}\gamma_{jl}S_l v_l = 0, \\ \rho\frac{dv_{jk}}{dt} + \frac{1}{h_{jk}}(p_k - p_j) = 0. \end{aligned} \tag{5.5}$$

Temporal discretization

The same time domain discretization notation and interleaving scheme from Chapter 4.3.2, including Equation 4.13, is also used here to replace continuous derivatives with sampled approximations thereof. This results in the fully discrete system

$$\begin{aligned} \frac{V_j}{\rho c^2}\delta_+ p_j + \sum_{k=1}^{N}\beta_{jk}S_{jk}v_{jk} + \sum_{l=1}^{N_b}\gamma_{jl}S_l v_l = 0, \\ \rho\delta_- v_{jk} + \frac{1}{h_{jk}}(p_k - p_j) = 0. \end{aligned} \tag{5.6}$$

Reintroducing the velocity potential notation $\mathbf{v} = -\nabla\Psi$ and $p = \rho\partial_t\Psi$ allows rewriting of the discrete time system above as the commonly-recognized two-step FVTD scheme:

$$\delta_+\delta_-\Psi_j + \frac{c^2}{V_j}\sum_{k=1}^{N}\frac{\beta_{jk}S_{jk}}{h_{jk}}(\Psi_j - \Psi_k) + \frac{c^2}{V_j}\sum_{l=1}^{N_b}\gamma_{jl}S_l v_l = 0. \tag{5.7}$$

Stability and boundary conditions

Finally, the integral energy balance and boundary admittances are used to write a discrete summation of the energy stored on the interior of the problem, within the capacitances at the boundary, as well as that dissipated by resistance at the boundaries. This is a complete accounting of the energy in the problem domain, and the energetic (and numerical) stability of the scheme as a whole is predicated on non-negativity of the internal energy, which gives

the following condition relating the simulation time step to cell geometry:

$$\frac{1}{V_j}\sum_{k=1}^{N}\frac{\beta_{jk}c^2T_s^2S_{jk}}{2h_{jk}} \leq 1. \tag{5.8}$$

Furthermore, this accounting also gives the update equations for the boundary terms that allows for the full simulation of the interior and admittances.

In the original article, a further procedure is performed to fit the admittance model to estimates of materials properties. In the case of our hybrid model, we directly used some of the calculated coefficients after confirming that the overall absorption presented by the surfaces was in line with the T_{30}s observed in the hallway. While an exhaustive fitting of the real boundary conditions of the hallway was out of scope for the hybrid model experiment, and following standardized measurement procedures as established in Fratoni (2021) for this type of acoustic model would improve the results in this and the following chapter, a similar fitting procedure was used with this method to characterize and examine the behavior of acoustical pots at the *caveau phonocamptique* of Noyon Cathedral, similar to past studies of acoustical pots, such as Valière et al. (2013). This study was also used as an opportunity to explore and test the possibility of using and evaluating a completely unstructured mesh for pressure-velocity simulations due to the importance of curved surfaces within the space, specifically the vault ceiling and support columns. This work is detailed in the master's thesis of Duval (2020), which was performed under the present author's co-supervision. We also used the same spatial Gaussian approach to set the initial conditions as specified in the article, as the low-passed nature of the impulse reduced the ringing of the source at higher frequencies.

Finally, as demonstrated in the article, one of the advantages of the FVTD approach is that numerical energy conservation is preserved to machine accuracy. We agree with the authors' assertion that this property is useful as a debugging tool while writing the code itself. As shown in Figures 5.2 and 5.3, the sum of energetic quantities was constant throughout time to within machine precision.

Implementation details

Our implementation of the scheme was created in MATLAB for pedagogical purposes. As initially the expectation for the code was to support only the version dealing with the EST method as well as the low-frequency pressure simulations described here, a vectorized implementation was pursued that preserved some readability regarding both problem setup as well as the solver itself. Specifically, our solver was a direct implementation of Equations 39-41b, rather than optimizing for a matrix-multiplication time-stepping scheme, or implementation on a GPU.

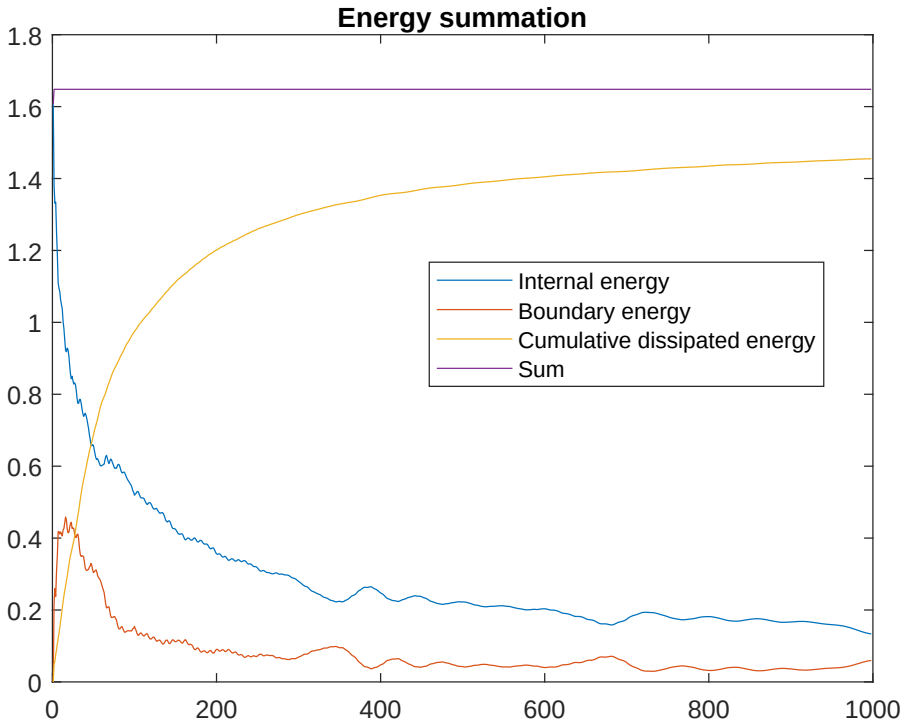

Figure 5.2: Summation of energy of internal cells, stored at the boundary impedances, and the cumulative dissipated energy as a result of resistive terminations for the low-frequency simulation

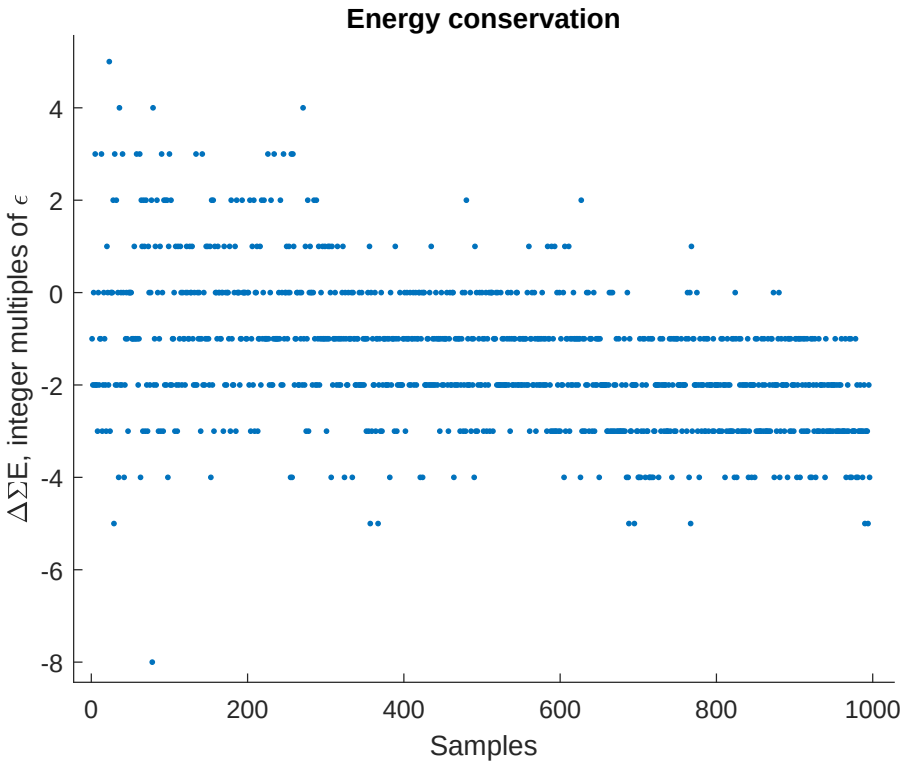

Figure 5.3: Demonstration of energy conservation to machine precision for the low-frequency simulation

As a result, for particularly large problems, the lack of object-orientation leads to fairly extreme inefficiencies in terms of storage and computation of RLC admittances and cell-indexed properties such as volume or intercellular distances. Later versions improved the speed by ameliorating some of these problems, however, the gain in performance is sorely offset by the complications required for implementation, and are not yet included in the code repository. While the indexing scheme used in the implementation is relatively straightforward and allows for a fairly natural expression of the mathematical operations detailed in the article, some of the manipulations required to perform certain operations are nonetheless quite slow under the current scheme, both due to repetition of unnecessary operations as well as lack of optimization for cache alignment. Despite the relatively straightforward nature of the vector expressions themselves, it is nonetheless difficult to exercise the granularity of control necessary to ensure that SIMD or other pipelining routines are used to accelerate the highly homogeneous computations. It would be advantageous for future implementations of the scheme to consider ways of reusing of particular repeated computations that are unavoidable with the naive implementation, or to switch to a numerical environment with stronger control over parallelism for computation of the velocity potential everywhere in the domain at a particular time step.

Hybrid simulation parameters

The simulation sample rate was set according to the highest desired frequency, up to the top of lowest octave band in the EST method (to ensure overlap in order to perform calibration later on), but in order to minimize the effects of dispersion error, we oversampled by a factor of 7 per the findings of Southern et al. (2011a), leading to a final sample rate of 4900 Hz. The resulting soundfield everywhere in the space was then lowpassed to the desired maximum frequency and resampled to the final audio sample rate. As with the EST method, when sources are relatively static, spatial downsampling and interpolation may also be employed to reduce the overall memory usage, further speeding recomputation when the receiver moves. While in this case, the FVTD pressure simulation is only run up to a maximum of about 700 Hz, we use the same approach at a higher sample rate in Chapter 6 as a basis for further exploring EST behavior in the hallway.

5.2.3 Direct path and early reflections

Finally, for the wideband specular portion of the impulse response, the image source method (ISM) from Allen and Berkley (1979) was used. There is some freedom to choose the order of images that will be used, which will be discussed in greater detail later, but the gist of the balance comes between representing all of the most prominent specular reflections without unnecessarily repeating the less-significant late reflections (at high computational cost) that

are already covered by the EST method.

5.3 Calibration

5.3.1 Between simulation methodologies

Southern et al. (2011b) and Southern et al. (2013) were used as a model for calibration between the disparate simulation methods. The sound energy at a distance of 1 meter from the source was used as a point of reference for all three methods in the octave band centered at 500 Hz. An arbitrary level was set in each individual simulation type, and then gain factors were used to match the energy of each signal. Then, upon combining the three signals into a hybrid by superposition in the time domain, the FVTD method (which is presumed to be the most accurate) was used as-is, whereas the high-frequency methods are high-passed above the crossover frequency in order to ensure the non-duplication of energy.

5.3.2 Between simulated results and measurements

Similar to the calibration between simulation methodologies, an energy match at the point of reference, 1 meter from the source, was performed. The uniform application of the correction gain means that the spatial decay is preserved, such that a comparison of the sound energy can be compared as a function of distance in the corridor.

5.4 Evaluation

Evaluation of the simulated impulse responses along the length of the hallway as compared to measured impulse responses was effected using a GUI shown in Figure 5.4. Inspection could be carried out in three ways: the comparison of spectrograms, the direct auralization of the impulse responses, or by playback of source material convolved with each impulse response. A slider at the bottom of the GUI provided a convenient interface for changing the listener position from 1 meter in front of the source to 42 meters (the maximum observation distance measured).

5.5 Results

Informal listening illustrated various characteristics of the auralization system design. Beginning with the interface, feedback to changing the distance parameter was relatively responsive, and could certainly be improved using an interface better suited to real-time interaction. Even with the computational delay triggered upon moving, general trends regarding the appearance of the spectrograms were apparent and consistent with expectations.

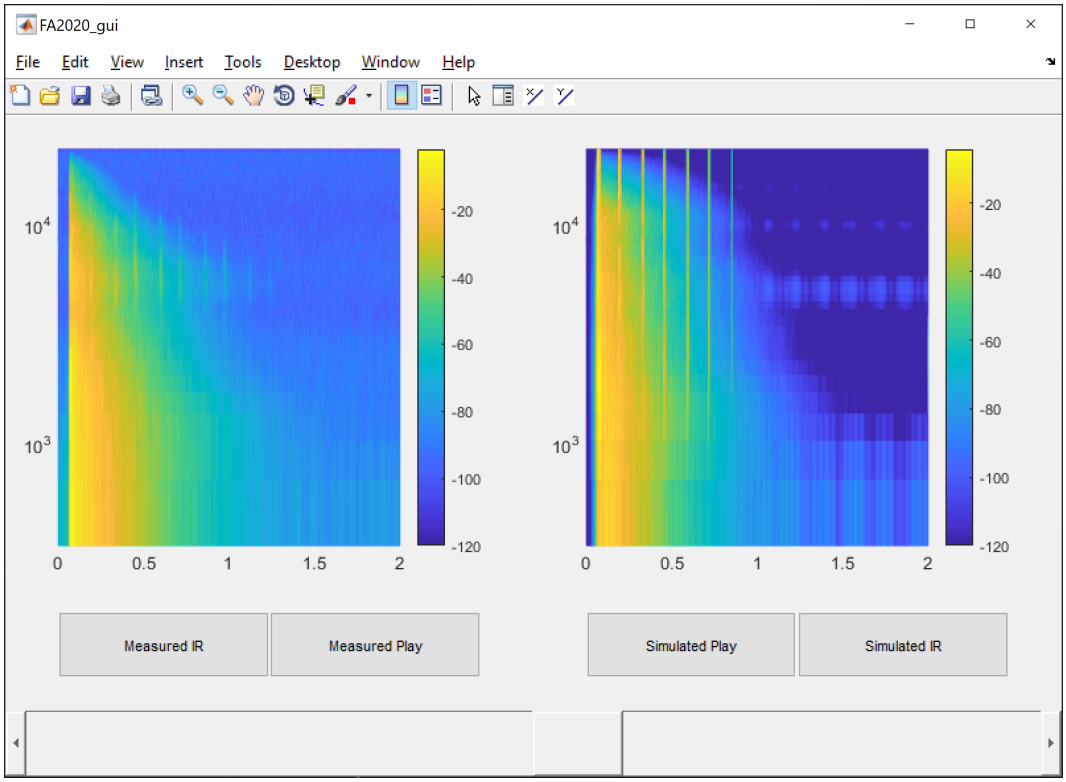

Figure 5.4: A screenshot of the MATLAB GUI used to compare simulated and measured impulse responses as a function of distance along the hall, both visually with spectrograms and aurally through the playback of the IRs themselves, or convolved with source material. In this case, the listener position was slightly past the middle point of the hallway, indicated by the scrollbar at the bottom of the interface.

Acoustically, the simulated and measured impulse responses did not diverge drastically at any point in the hallway, though at no point did they sound identical. The time arrivals of early reflections were accurately rendered, as were the overall decay times and absolute levels of the stochastic reverberation, meaning that on the whole, the impression of the hallway was approximately the same between both auralizations. Some of the most significant differences were in the decay of the ISM and the low-frequency FVTD portion along the length of the hallway, leading to slightly different balances between the saliency of the early and late portions of the impulse response when moving further from the source. These differences appear to be a result of the simulation parameters differing from those used in the EST method, the octave-width calibration used, and the lack of tuning at high frequencies, where viscothermal losses would have more of an effect on the image source reflections. As a result, the simulated impulse responses have a brighter tone throughout the hallway, as well as sounding too diffuse near the source.

One difficulty with the ISM in this specific space was balancing the order of sources against the need to represent the particularly important echoes corresponding to the length of the hallway. For the transverse reflections, a fairly low order was sufficient to give a good spatial impression, but because of the strength of repeating echoes along the length of the corridor, a higher order was needed to continue representing these strong reflections even as the stochastic reverberation became more and more dominant in time.

5.6 Future work

A number of improvements are possible within this scheme, as well as for extensions to other possible hybrid methodologies beyond the approach described here.

While in this case the ISM was chosen for simplicity, it could be changed out for a variety of other geometrical acoustics approaches better suited to a particular type of geometry. In this case, with a long hallway, choosing the proper order was more difficult than expected, whereas a ray-based or digital waveguide approach may have been more successful at reproducing the infrequent longitudinal reflections in an efficient manner, though the same cannot be said for volumes of all shapes or sizes.

A variety of computational optimizations were eschewed during development, but in preparing a true real-time hybrid model, would need to be taken into consideration. These include, but are not limited to, caching of particular simulation results, acceleration of the low-frequency FVTD portion with GPUs, lazily evaluating the resampling operations required between various methods in the region of the receiver (rather than everywhere), and further refinement of the assembly of filters from the constituent acoustic information.

One final note is with respect to the element sizes of the EST method and the pressure-domain FVTD results. It may be possible that a single meshing of the problem domain

could be utilized for both a pressure-based room acoustic simulation as well as for the EST method, given the proper scale of discretization for both techniques. In cases with changing geometry this could prove extremely useful, as well as unifying the storage required and simplifying any voxelization steps.

5.7 Conclusion

In summary, due to its region of validity, the EST method must be used in a hybrid context for synthesizing realistic room impulse responses, and this chapter provides a framework for doing so with the ISM and low-frequency FVTD pressure simulations taking on the non-stochastic portion of the soundfield. A simple interface for examining spectrograms and initiating playback of synthetic and measured impulse responses allowed informal testing of the design during development as well as ultimately facilitating the evaluation of differences between the real and virtual scene. While further theoretical and practical improvements are necessary for use of the method as a general replacement for current stochastic reverberation techniques used in real-time auralization systems, as a drop-in replacement when a specific space has been measured, the EST method appears to be an effective and accurate technique in terms of the measures used to match simulation parameters to the acoustic behavior of the spaces under study. Extending the methodology to higher-dimensional spaces nonetheless remains a challenge due to the lack of information regarding the simplifications that may be presumed regarding the EST terms in real spaces. While thus far, our analysis has relied on the reduction of the system to a 1-dimensional case in order to characterize the members of the wave-stress tensor and the off-axis sound intensity directions, we would like more information about the behavior of those terms in particular spaces to understand what, if any, assumptions we may make when directly modeling them in the future.

Chapter 6

Energy-Stress Tensor Quantities

6.1 Introduction

This chapter proposes an approach for characterizing the behavior of the EST in particular regions of a room without exhaustive physical measurement. In opposition to earlier portions of the thesis, rather than directly simulating members of the EST, instead, we focus on simulating the velocity potential, from which the relevant terms may be computed.

6.2 Toward 3-dimensional prediction

The 1- and 2-dimensional versions of the EST are tractable because of the dimensional reduction that is performed, using energy and momentum balances that result from the geometric and diffuse field assumptions discussed in Section 2.3.3 in order to find boundary condition relationships between the energy density, sound intensity, and the wave-stress tensor. In each of the lower-dimensional cases, these boundary conditions are present on at least some portion of every cell in the numerical approximation, meaning that they can be integrated directly into the propagation equations in a uniform manner, with the "end boundaries" (those occurring at the extremities of the primary axes of the domain) being treated specially. In approaching the time-varying 3-dimensional EST, we must first understand its behavior further away from the boundaries, but given the difficulty in evaluating diffuse fields with physical measurements, we decided instead to pursue a computational approach.

6.3 Pressure domain simulation

In Chapter 5, a finite volume time domain (FVTD) scheme was utilized to perform the simulation of low-frequency acoustic waves. This chapter utilizes the same approach to generate wideband pressure signals that are the basis of the analysis to come, requiring a much higher sample rate and smaller element sizes. Whereas before, the highest frequency to be simulated was determined in concert with the crossover point to stochastic reverberation and the desire to perform simulations as close to real-time as possible, the much higher computational requirements for generating oversampled results that are valid through the same region as the EST method means that the simulations are now effectively offline only, though porting our CPU-based implementation to utilize GPUs (as in Hamilton et al. (2016) would greatly improve runtime. Nonetheless, the FVTD approach is particularly convenient as the parallels between the pressure domain formulation from Bilbao et al. (2016) and the EST formulation from Chapter 4 mean that information about the unknown terms in the energy-stress tensor can be derived in terms of the velocity potential of the pressure domain simulations.

In this case, we performed a simulation of the hallway with alcoves at $f_S = 16000$ Hz. With 8x oversampling, this implied the highest valid frequency was 1000 Hz, which as previously demonstrated overlaps the region of stochastic reverberation. Otherwise, we used the same boundary conditions and approach described in Chapter 5, just with a higher frequency limit. Because of the much larger simulation requirements imposed by our somewhat unsuitable implementation as described in Chapter 5.2.2, both in terms of time and memory, a remote computation toolbox was very useful in managing asynchronous launching and collection of simulations and their results. More details on this setup can be found in Appendix B.3.

6.3.1 Ambisonic microphone approach

Performing the aforementioned pressure domain simulation using the FVTD method over a collection of axis aligned cubic cells gives a convenient representation for computing the components of the energy-stress tensor. Because the faces of the cells are perpendicular to a particular axis, the unit vectors of the velocity fluxes that pass through the faces are also axis aligned. While in the case of an unstructured mesh it would be possible to extract the projections of each face onto the axial vectors, simply using cubic elements greatly simplifies the analysis. Furthermore, as explained in Chapter 5, we are comfortable treating the space as rectilinear, which eschews the need for edge fitting and a more complicated computational scheme, as the lack of any curved surfaces implies that the staircase effect will not be present for the problem domain.

At the smallest scale of a single cell, extracting the pressure and velocity flux from

the local velocity potential is akin to making a measurement with a first-order ambisonic microphone. This makes it straightforward to compare to real measurements to ensure the veracity of the pressure simulation results. In this case, using the standard nomenclature for B-Format audio, the average pressure in the cell corresponds to the W channel, or the omnidirectional output, and the net velocity through the cell (resulting from each axis aligned pair of faces) corresponds to the X, Y, and Z channels, or the linear directional signals.

Of course, the terms of the energy-stress tensor can also be computed at this scale, but local variation arising from the high modulation frequency of pressure waves means that the extracted values at a particular point in space may not be representative of the local behavior more generally. In order to account for this, as with many energy based methods, spatial and temporal averaging are used to examine global trends in the components of the energy-stress tensor.

The quantities of interest can also be averaged over an amalgamation of cells in order to successively approach the large cells that are "properly sized" for the modulation frequency of the EST method.

6.4 Derivation of EST terms in FVTD formalism

One possible approach to calculating the terms of the energy-stress tensor in terms of a computed velocity potential field is to simply refer directly to their definitions as described in Morse and Ingard (1968), and averaging over collections of nodes.

Thus, by definition from Equations 2.3, 2.4 and 2.7, we have:

$$E_{tt} = \frac{\rho}{2}(\frac{1}{c^2}|\partial_t\Psi|^2 + |\nabla\Psi|^2),$$

$$E_{ti} = -\rho\partial_t\Psi\partial_i\Psi,$$

$$E_{ii} = \frac{\rho}{2}(\frac{1}{c^2}|\partial_t\Psi|^2 + \sum_j \alpha_{ij}|\partial_j\Psi|^2),$$

$$E_{ij} = \rho\partial_i\Psi\partial_j\Psi,$$

for $i, j = x, y, z$ with $\alpha_{ij} = 1$ when $i = j$ or -1 otherwise.

Each of these equations can be evaluated on a cell-by-cell basis and averaged over a collection to arrive at an approximation of the EST terms in a particular region.

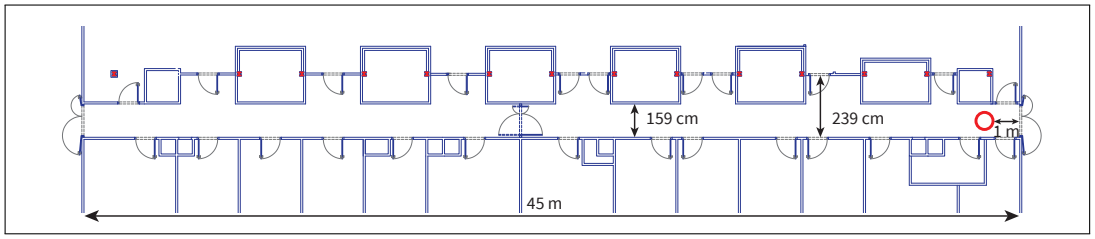

Figure 6.1: Corridor floorplan with dimensions

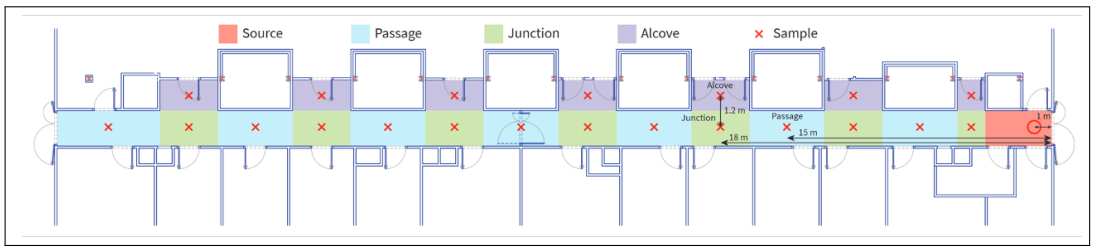

Figure 6.2: Corridor floorplan with sections, source, and measurement positions, with the subjects of Figures 6.5, 6.7, and 6.9 labeled respectively

6.5 Measurements

While the point of simulating an approximate soundfield is to be able to exhaustively enumerate its behavior everywhere in the space, nonetheless, some connection to reality must be established through comparison with measured data. In this case, a lack of roboticization, or indeed, a space dedicated to measurements (as the corridor itself is a regular workplace) meant that it was neither possible nor necessarily desirable to collect an extreme number of measurements within the space. Instead, to facilitate comparison with the simulations, a series of measurements were made along the length of the corridor, as well as within each alcove, corresponding to the volumes to be averaged as discussed above.

The equipment and software used to perform these measurements were as follows. Since we are concerned with the spatial impulse response, a first-order ambisonic microphone, the Core Sound TetraMic, was used to capture not only the pressure field at each sampling position, but also its gradient, allowing inference of the velocity at the measurement position, consistent with the response of a single cell in the FVTD simulation. The exponential sine sweep method Farina (2000, 2007) was used to capture the impulse responses, generated using the Adobe Audition plugin Aurora, played back over an Outline GRS omnidirectional speaker, with a MOTU Traveler sound card as the input and output interface. The sweep length was 20 seconds at a sample rate of 44100 Hz. A digital source level gain was chosen to maximize the signal-to-noise ratio without introducing distortion at the initial recording position.

Table 6.1: Spatially averaged T_{30}s.

Frequency [Hz]	62.5	125	250	500	1000
T_{30} [S]	2.29	0.35	0.38	0.39	0.43

The source was positioned 1 m away from the right end of the hall in Figure 6.1, 1.5 m above the ground, and centered between the two walls. The microphone was also placed 1.5 m above the ground, and was moved to each measurement location on the same plane. The horizontal "centroid" of each section of hallway beginning with the section in front of the first alcove was the targeted measurement position, where the sections are defined using the two side walls of each alcove and the upper wall to divide the hallway into 22 volumes (one of which contained the source). These sections were classified into three types, either hallway, junction, or alcove. These distinctions, as well as the measurement positions, can be visualized in Figure 6.2. As the microphone was moved further from the source, its gain was occasionally digitally readjusted to ensure full dynamic range in the recordings, and these changes were recorded in order to recover the true level for each measurement location.

Converting the A-format signals (the individual channels of each microphone capsule) to B-format as well as applying the individual microphone's calibration file was accomplished using the VVAudio VST plugin VVEncode. To facilitate the processing of the measurements in bulk while reducing opportunities to introduce human error, the batch converter feature of Cockos REAPER was essential, as a GUI-enabled host was required for the plugin. In a similar fashion, rather than using the Aurora plugin to perform the deconvolution for individual recordings directly in Audition, the generated inverse sweep was instead saved and used in a MATLAB script that produced impulse responses that were bit-accurate to plugin-processed versions and could be evaluated rapidly in parallel. While Audition 3.0 (the version used with the Aurora plugin) does have batch processing capabilities, it was deemed too brittle, leading to the development of the MATLAB processing script.

The measurements, plotted according to their linear distance along the hallway, can be seen in Figure 6.3. The measurements taken within an alcove are colored red, all other sample positions are blue, and an estimated time of arrival according to distance (assuming $c = 343$ m/s) is provided in black.

From these measurements, octave band T_{30}s and Schroeder's frequency could be computed. T_{30}s were consistent along the length of the hallway, and are given in Table 6.1. Schroeder's frequency, commonly considered the cutoff below which the room response is dominated by modal effects, is given by $f_{\text{Schroeder}} = 2000\sqrt{\frac{T_{30}}{V}}$, where V is the volume of the hallway in cubic meters, predicting a crossover frequency of 81 Hz.

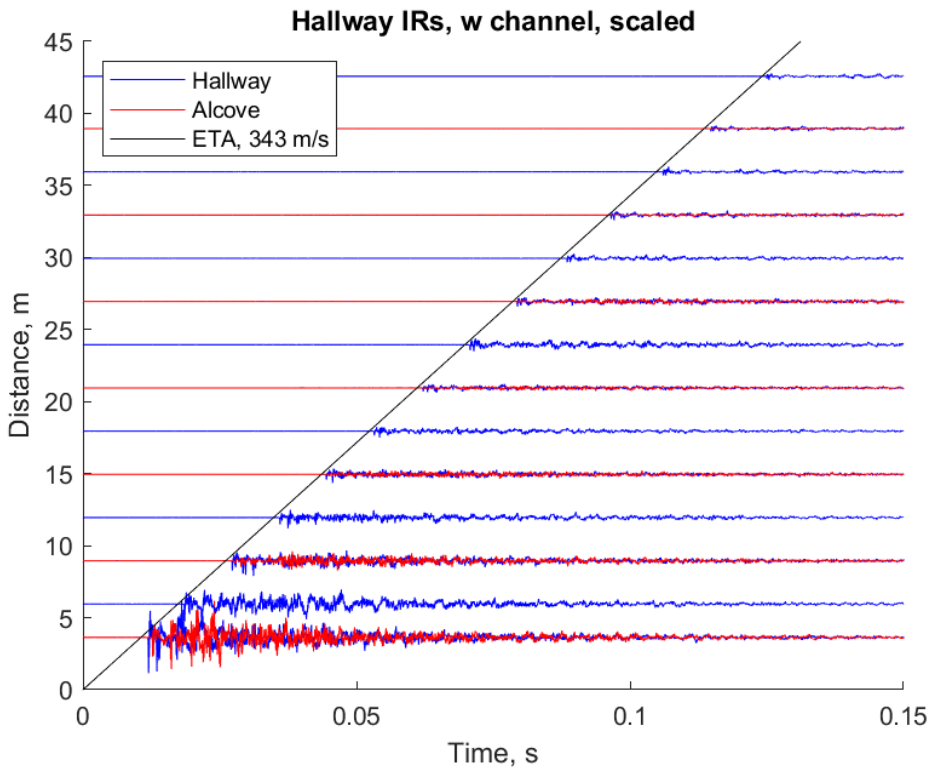

Figure 6.3: Measured impulse responses along the length of the hallway

6.6 Preliminary results

In this case, we subdivided the hallway into medium-sized cells of various configurations in order to examine the coupling between regions near the boundaries and in open space. For clarity in discussing results, we recall that the simulation and mathematical coordinate system was chosen such that the length of the hallway was in the X direction, the width in the Y direction, and height in the Z direction. The hallway was defined as a collection of rectangular solids, with a line of volumes representing the main space of the hallway, delineated at each alcove edge, as well as an adjacent volume for each alcove. From this basic voxelization, subdivisions into other configurations is relatively straightforward, and provides a natural interface for examining behavior in particular regions. For these preliminary results, a mesh division of 3 cells per side (for a total of 27 cells per volume) was undertaken, which resulted in a layer of cells that touch the boundaries and a central series of cells that is completely in free space.

The pressure simulation was carried out for half a second, resulting in the definition of the velocity potential everywhere in the space. As in the previous simulations, the source was a spatio-temporal Gaussian centered 1 meter from the end of the hallway and 1.5 meters off of the ground with a variance of $0.2m^2$. The source signal was also Gaussian in time, a 7-sample window. As the sample rate after downsampling was 2000 Hz, there are 1000 temporal samples for each of the signals. Then, according to the definitions above, the EST values were calculated by averaging over the contiguous regions created by the subdivision of each section of the hallway.

The resulting time series for each term in three selected regions, each of which was the "central" amalgamation of cells from each of their respective volumes, are shown in Figures 6.5, 6.7, and 6.9. We have grouped the energy density with the diagonal terms of the wave-stress tensor, the sound intensity, and the off-diagonal terms of the wave-stress tensor each on separate plots for clarity.

The "passage" region depicted in Figure 6.5 was centered vertically and horizontally in the hallway, with a centroid 14.95 m from the end, an average of 9450 cells. This implies a round-trip travel time to and from the far end of the hallway of 0.176 seconds, about 353 samples at the current sample rate, and 0.088 seconds or 176 samples round-trip to the nearer (source) end of the hallway. The "junction" region depicted in Figure 6.7 was 3 m further down the hallway, an average of 7560 cells, making its long round-trip time slightly shorter at 0.159 seconds or 318 samples, and its short round-trip slightly longer for 0.106 seconds or 211 samples. Finally, the "alcove" region in Figure 6.9, which was the third alcove from the end of the hallway, was located at the same distance as the "junction" region but was offset horizontally by a distance of 1.2 m, an average of 3840 cells. Given this small change in horizontal offset, the distinct behavior in this region is remarkable.

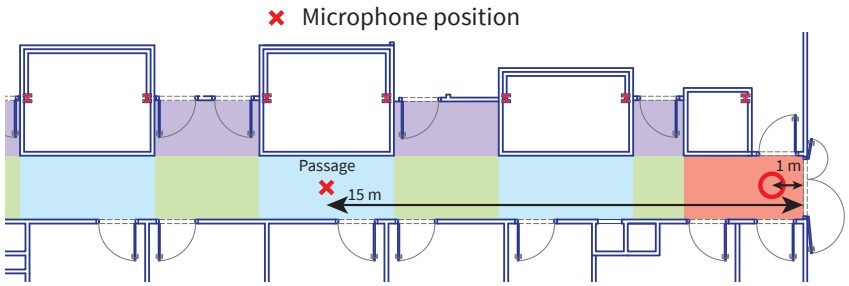

Figure 6.4: Measurement location for Figure 6.5

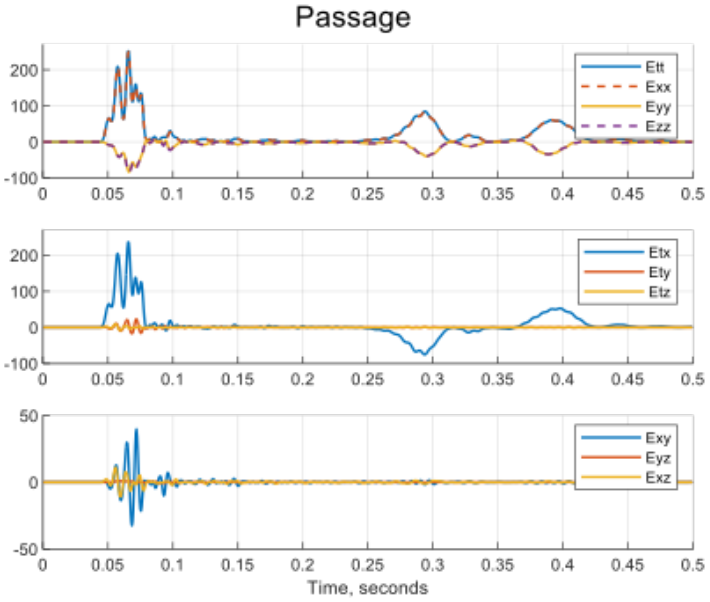

Figure 6.5: Energy-stress tensor terms in the center of a section of corridor between alcoves labeled in Figure 6.4

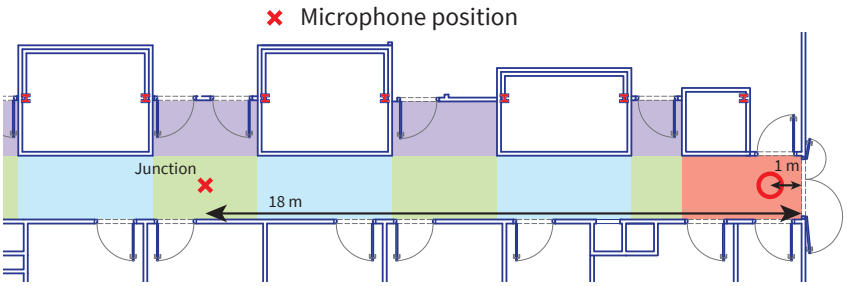

Figure 6.6: Measurement location for Figure 6.7

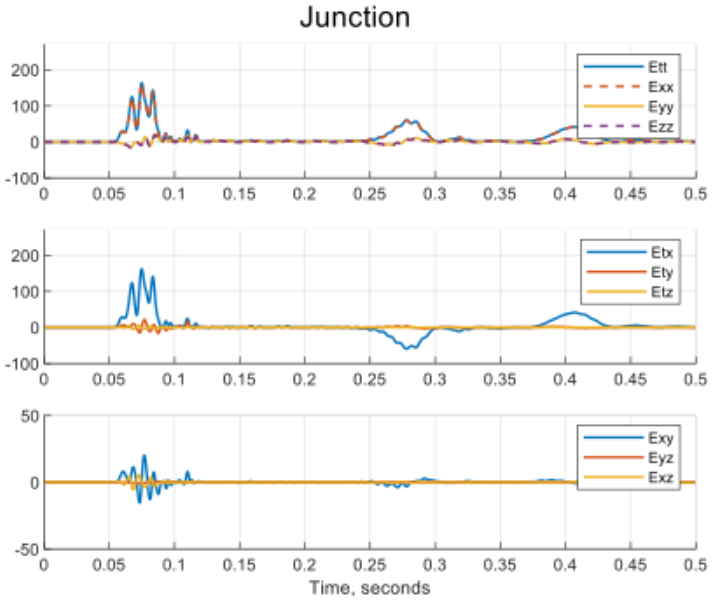

Figure 6.7: Energy-stress tensor terms in the center of a section of corridor in front of an alcove labeled in Figure 6.6

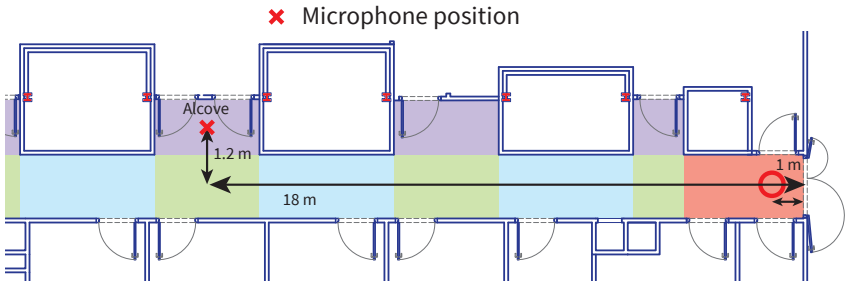

Figure 6.8: Measurement location for Figure 6.9

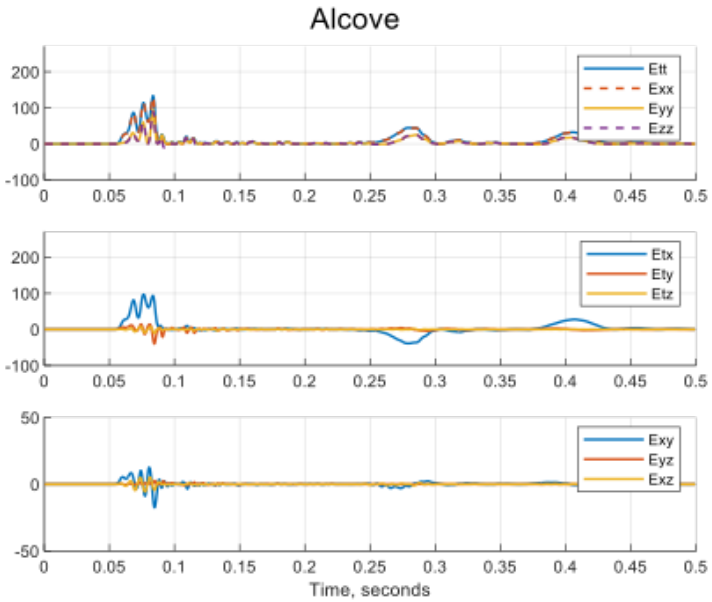

Figure 6.9: Energy-stress tensor terms in the center of an alcove labeled in Figure 6.8

6.6.1 Passage sections

A number of phenomena are apparent from the computed results. Beginning with the passage section in Figure 6.5 as the base case, we can immediately identify some of the most relevant acoustical events in the impulse response. First is the presence of three major peaks in energy: the direct sound (which also likely subsumes the first reflection off of the end of the hallway behind the source) and two reflected wavefronts. As expected, when examining the sound intensity in the x direction (E_{tx}), the direction of travel of each of these wavefronts can clearly be seen, switching with each reflection at the end of the hallway.

The next point of note is the relative agreement between E_{tt} and E_{xx}, implying that most of the energy density in the hallway is driven by the x-aligned component. To the contrary, the E_{yy} and E_{zz} terms are also grouped fairly closely, and for the most part, oppose the motion of the energy density. This implies that in these sections, kinetic energy is the dominant form of energy storage, a surprising result. (Note that the close agreement of these pairs of terms may result in the appearance of two lines upon first inspection of the plot.)

As expected, all of the off-diagonal terms of the wave-stress tensor are non-zero, but are nonetheless much smaller in magnitude than the diagonal terms. Interestingly, the increased power in the E_{xy} component suggests a much stronger coupling between the X and Y dimensions than that of X and Z or of Y and Z; that is to say, more energy is transferred from the longitudinal waves into crosswise waves. This makes sense as the alcoves themselves were oriented along the width of the hall, and, being nearly the full height as the main ceiling, primarily presented surfaces that would cause diffraction in the direction of the alcove.

Second smallest is the E_{xz} component, illustrating some transfer of energy from longitudinal to vertical waves, likely due to scattering off of the ceiling, which was subjectively rougher than the smooth walls. The E_{yz} component is minimal in the entirety of the examined region, demonstrating that the small amount of non-longitudinal energy tended not to change from crosswise to vertical, likely due to the lack of diffracting surfaces oriented in a fashion that would facilitate such transfer.

6.6.2 Junction and alcove sections

Comparison of the passage region with a region adjacent to an alcove as well as an alcove itself also reveals differences, primarily in terms of the E_{yy}, E_{zz}, and off-diagonal components of the wave-stress tensor. Of course, one of the most obvious differences is the change in arrival times for the direct sound as well as the increased spacing of the reflections due to being closer to the middle of the hallway's length. In the junction volume, the most distinct change comes from the non-X components of the wave-stress tensor which do not express

the same degree of opposition to the energy density and E_{xx} terms as in the hallway section. This is closer to our expectation of reality. To the contrary, however, within the alcove, these terms again shy away from zero in the positive direction, equally as surprisingly.

Nonetheless, there are other interesting phenomena occurring at the junction and within the alcove. One observation is that at the junction, E_{xy} is much stronger upon returning reflections, whereas it is quite weak in the hallway section. This is even more pronounced within the alcove, suggesting the strength of the presence of the alcove in curvature from the length of the hallway to the width presented in both directions. Furthermore, the sound intensity in the Y-direction is fairly similar in the hallway and junction, but points much more strongly within the alcove, suggesting the diffracted wavefronts in the region.

6.6.3 Discussion

First and foremost, noting the general trend that E_{tt} and E_{xx} are similar in magnitude throughout the hallway speaks to the idea that energy transfer is primarily driven by the sound intensity in the X direction, validating the first of the assumptions made in previous dimensional reduction approaches.

Furthermore, given the relatively small magnitude of the momentum flux terms, E_{xy}, E_{xz}, and E_{yz} particularly in the alcove means that in some cases the decorrelation assumption may hold.

Having access to spatially averaged versions of the EST terms for the first time is a powerful tool, but also calls into question how we may learn from these observations. One desirable outcome would be definition of scattering or diffusion coefficients that could directly inform simulation of the EST based on the ratio of energy density, sound intensity, and wave-stress tensor components. This points to the evaluation of other interesting or modifiable room layouts, perhaps as part of a large dataset, to extract more information regarding the behavior on average of the various terms in regions of free space as well as near strongly diffusing geometry, whether reflectors, diffusers, or other room features.

Further analysis of this data is required, specifically regarding the analysis of other regions and surfaces within the hallway. For example, the behavior of the EST near boundaries, specifically along long sections of the hallway as well as around corners going into alcoves, may be particularly revealing. Additionally, while this preliminary analysis found that the behavior of the individual "types" of regions within the hallway (that is, passages, junctions, and alcoves) had relatively consistent behavior corresponding to the descriptions given above, further characterization is necessary to quantify this relationship as well as to evaluate other possible regions of similar behavior. While analysis of the major sound energy packets traveling the hallway has already provided much fodder for analysis, the relationship of these terms to the representation of the stochastic reverberation is an important goal

that should not be forgotten. Particularly in between regions of strong reflections, further characterization of the decay from passing wavefronts may reveal further insights regarding the relationships between energy density, sound intensity, and the wave-stress tensor, and may help to define how to discern between the early reflections and diffuse field in both analysis and synthesis contexts.

6.7 Riemannian tessellation

Previous efforts to couple the known boundary conditions near the edge of the problem domain through a "translation layer" to the interior free-field solution to the EST were shown to be incorrect in Polack (2020). The reason for this was a result of the energy balances under consideration failing to admit absorption.

This provided the impetus for recent investigation into geometric characterization of diffusion, which we will examine now as an alternative approach to direct parameterization of the EST method for a given space. Detailed in Polack et al. (2021) as a preprint, the main idea is to approach the theory of scattering based on image sources calculated by tessellation of a given space, that is, by repeated tilings of the room geometry depending on the angles at its vertices. While this idea is well known in the case of rectangular rooms where it has long been used with respect to the image source method, for example, it is more challenging to define in polyhedral rooms, which leads to the use of Riemannian geometry to define the space while preserving the familiar idea of tessellations.

In geometrically mixing spaces, diffusion occurs regardless of the scattering coefficient on any given surface, but rather, as a result of scattering on curved surfaces within the room. This is distinct from geometrically imperfect real rooms whose shape would suggest that they would not be mixing, meaning that one might expect only a few directions of travel to be supported, but achieve it nonetheless due to construction imprecision and surface scattering. The goal of this approach is to characterize the diffusion process in these types of mixing rooms in terms of their geometry. It is important to note that perfect polygonal or polyhedral rooms are not mixing, due to the discretization of ray direction that occurs in them - to the contrary, only spaces with at least one curved surface on the boundary have been proven to be mixing, such as the stadium introduced by Bunimovich (1979) in dynamical billiard theory.

In short, the behavior of reflections for particular dihedral angles are analyzed in terms of the image sources they produce. Successive reflections in a particular room imply a tessellation of the space, similar to the familiar rectangular case, but in this case, with the possibility of other shapes (though the concept of ray direction is still preserved). It is demonstrated that for right angles, the resulting Riemannian space is Euclidean, which corresponds to the non-mixing nature of rectangular rooms and the linear increase in image

sources at every order. For obtuse angles, however, the curvature is negative, and the number of image sources grows exponentially, as is the case with mixing rooms, such that we may refer to polyhedral spaces with this type of vertices as "pseudo-mixing." Using this approach, the rate of increase in image sources can be computed based on the number of faces, edges, and vertices in the room - a purely geometrical way to determine the mixing abilities of a given space.

Finally, by considering each image source as a wavefront, frequency dependence may be taken into account, demonstrating diffraction in a natural way. This process is performed by considering the conservation equations from the EST method across each reflection angle of the Riemannian tessellation, each of which introduces acceleration to the energy-stress tensor.

With this formalism, it should be possible to compute position- and direction-dependent scattering coefficients based only on the geometry of a particular space. Then, coupling a structured interior mesh implementing the existing EST method with a boundary layer that introduces scattering appears to be a promising path to representing geometric diffusion effects without physical or virtual measurements and an adjustment of model procedure. This idea is related to the stated range of the scattering coefficient briefly discussed in Chapter 3. In Dujourdy et al. (2017) section 5.2, the meaning of scattering coefficients greater than 1 is briefly discussed, and is shown to be indicative of energy staying in a particular region as β tends toward 2. While non-physical, this is related to the idea of directionally-dependent coefficients discussed above, where non-symmetric surfaces may facilitate transfer of energy in one direction, but could reflect a much greater proportion of incident energy in another, meaning a single bidirectional scattering coefficient may be insufficient. Hypotheses regarding the definition of these coefficients in terms of room geometry thus remain an interesting opportunity for future work.

6.8 Discussion

While the analysis of these results is not yet complete, the strategy points to two possibilities for further characterization of the EST method. First is for the constitutive equations relating the wave-stress tensor to the known energy density and sound intensity in free space, which would allow for direct simulation of the full energy-stress tensor in a space with boundary conditions defined only in terms of the modified absorption and scattering coefficients currently defined on the boundary. Second is the empirical relationship between room geometry and the behavior of the diffuse field in a room. For example, while characterization of the presence or prevalence of diffusing surfaces in a particular room is accomplished heuristically through the modified scattering coefficient D, predictively modeling a space requires some sense (beyond simple scattering) of the geometry in order to determine how quickly

it diffuses energy. While this is difficult to approach using only the EST method due to the size of the elements and the difficulty of modeling the introduction of occluding surfaces within such a volume, the ability to accurately model a space in the pressure domain and use the averages of particular volumes to determine a more meaningful diffusion coefficient is now within reach.

6.9 Future work

With these techniques, the relationship between the Sabine scattering coefficient and the EST scattering coefficient could now be clarified based on simulations of a particular space. Whereas physically verifying the agreement between the two would require a carefully calibrated reverberation chamber or other physical acoustics expertise, this approach provides a basis similar to past work on the EST where simulation parameters may be matched directly to measured data (in this case, measured from the pressure simulation) by brute force. In the case where the room geometry is unchanging, a relation between changing absorption or scattering coefficients in the pressure model may be correlated with a proportional change in EST parameters that produce the same stochastic reverberation directly.

Edge fitting, long known to be important for finite difference and finite volume approaches, is also a critical future implementation target for the method, and would allow modeling of non-rectilinear spaces as well as admitting more complicated interior geometry that, in the pressure domain model, would result in increased diffusion. Thus, further studies may attempt to relate the presence and density of thin reflectors in a reverberation-style chamber, which would be directly modeled in the pressure simulation case, to changes in the absorption and scattering coefficients in the EST simulation in cells that contain (but do not physically represent) diffracting and reflecting surfaces.

Thus, between characterization of the changes in simulation parameters in the cases where materials properties on boundaries change but the geometry is unchanging, and vice versa, a more complete picture of the behavior of the EST terms as well as that of the diffuse field itself may be better understood.

6.10 Conclusion

A method for directly calculating EST terms from a pressure-based velocity potential simulation was introduced with appropriate averaging in order to examine the behavior of terms in the EST. The takeaways from this experiment are preliminary, but may be able to inform the direct simulation of the energy density in 3-dimensional spaces in the future through specification of assumptions regarding the quantities of the wave-stress tensor. While the

current results are limited based on the geometry of the measured and simulated acoustic space, the computational strategy is equally valid for many other types of spaces, and therefore presents a promising avenue for future research on the topic. As much more well-known wave-based pressure simulation techniques and the corresponding standards for the characterization of their boundary conditions in terms of real-world materials properties and geometry continue to mature, this approach will also only improve in accuracy of its representation of the stochastic soundfield. This in turn would provide opportunities to explore the comportment of EST terms in 3-dimensional spaces, especially for large or open-air spaces that remain computationally challenging and would benefit from direct and efficient modeling of the EST.

Part III

Conclusions

Chapter 7

Conclusions

This thesis has focused on the development of our understanding of the energy-stress tensor from the perspective of theory, measurements, and computational approaches. Previously, while the mathematics of the tensor were well understood due to its importance in other physical disciplines, the practical application to acoustics remained limited due to the challenges of introducing additional constraints to the system consistent with physical observations of the time evolution of its terms, especially for 3-dimensional spaces. This speaks to the challenging nature of comprehensive measurement of energy fields as well as the definition and verification of assumptions that simplify the complex relationships in the EST. Furthermore, it highlights the innovation of solutions to 1- and 2-dimensional spaces based on a momentum balance involving scattering at the domain boundaries, which built upon the assumptions of the popular diffusion equation method, even if the exact relationship between common materials property measures and boundary conditions remains difficult to evaluate.

In analyzing and extending the results from the 1-dimensional case, this thesis has proposed an improved understanding of the solution space presented by the boundary parameters and its relationship to reverberation characteristics in real spaces on a frequency-by-frequency basis. This demonstrated some of the requirements for particular spaces to be representable with the approach, as well as clarifying the role of the parameters in terms of the physical behavior of the system. With this knowledge, it began to become clear how such an approach could be used to rapidly model the stochastic reverberation in a space across a wide frequency range, provided that the proper coefficients could be found from measurements or other predictions. Furthermore, due to the size of the elements, even a summation of multiple EST simulations with band-specific parameters may still be a more efficient way to represent stochastic reverberation than other wave-based or even geometric approaches to room acoustics, especially when considering the possibility of moving or

directionally time-varying sources made possible by the inclusion of a source term.

This aspect was particularly relevant in ultimately auralizing the stochastic reverberation, which we hope will make the EST method more approachable to future researchers. The straightforward application of time-frequency envelopes to noise sequences means that aural evaluation of results is an effective way to evaluate simulation results for future models which may not be predicated on matching preestablished acoustical indices, but instead directly predict boundary condition parameters based on room geometry and materials properties as the final arbiter of quality. The hybrid model also highlighted further questions regarding the relationship between stochastic reverberation and other high-frequency acoustic phenomena that were not immediately apparent upon examination of solely energy-based results. As demonstrated in the final chapter, some EST terms include wave packets arriving around the same time as strong reflections, perhaps indicating the energy they carry. While we are confident in the model's capability of matching spatio-temporal decays on the scale of an entire impulse response, verifying that the model properly represents the transformation of these individual arrivals of characteristically non-diffuse energy into a more general diffuse field would be an important finding for the assembly of hybrid auralizations, regardless of architecture.

Finally, we hope that the strategy presented in the last chapter, directly computing the EST terms from a high-frequency wave-based model, may be useful in characterizing better assumptions for the method in free space and in regions near boundaries. Continued interest in these sound pressure models integrating improvements in boundary condition modeling, viscothermal effects, and implementation on highly parallel hardware architectures mean that the results from such analysis can only improve. Furthermore, this approach likely surpasses the quality of measurement that would be possible with human-driven real-world recordings, and the fact that a single simulation computes the field everywhere in the space means that it may be competitive with serialized robotic collection of data in terms of raw speed. We hope to have more definitive and expansive results regarding the evaluation of the data in the studied space soon, and hope that the template allows for similar experiments in other acoustical spaces of interest in the future.

7.1 Review

This thesis has covered four topics on the subject of the energy-stress tensor method: frequency validity, a finite volume formulation with sources, auralization of results within a hybrid model, and computation of tensor quantities from pressure-based simulations. Taken as a group, these advancements point toward the usefulness of the EST method for simulating the stochastic reverberation in a variety of spaces and under varying conditions. Furthermore, the advantages in terms of speed and computational complexity, as well as

possible synergies with existing methodologies for other portions of room acoustical simulations, have been highlighted both in theory and in practice.

In terms of frequency validity, the model was demonstrated to function well above Schroeder's frequency in a long hallway for two different source types, monopole and dipole. This served as a confirmation of the region of applicability for the model as well as demonstrating the approach for synthesis that would be undertaken later. Some of the main findings in this section were the illustration of the effects and interpretation of interchangeability of the two simulation coefficients, α and β, nominally characterizing the absorption and scattering on surfaces in the room. Additionally, the difference between hallways that were "sufficiently diffusing" and not was examined, further substantiated the relationship of the EST method to the presence of a diffuse field.

The finite volume formulation of the method was useful in two contexts: for the inclusion of sources in the formalism as well as eventual comparison with acoustic pressure simulations in the same schema. The similarities and differences between the update equations for the two approaches were highlighted, demonstrating the overall effect of diffusion, which is to oppose or slow the flow of diffuse energy throughout a space. While a truly diffuse field must evenly fill a volume, the evolution of an initial introduction of energy to a diffuse state is a more complicated process, relying on specular reflections to distribute initial energy quickly throughout a space and then facilitated by diffuse reflections in particular regions. Previously derived dimensional reduction assumptions on the energy and momentum balances at the boundaries allowed for the characterization of hallways, as before, to check the formulation, but also allowed for additional flexibility in problem domain definition.

Auralization is an important practical aspect of room acoustical modeling, essential for evaluation of results as well as encouraging uptake in contexts that would benefit from its speed or physically-informed model of stochastic reverberation. While the role the EST method plays in a hybrid auralization scheme was clear, difficulties in the temporal transition from early reflections to the stochastic reverberation were highlighted, even as other perceptual characteristics were adequately reproduced. An interface developed to compare simulated and measured impulse responses and convolved source material proved to be a powerful way to evaluate changes to the design of the hybrid method, giving instant feedback on both successes and mistakes. Ultimately, the effort provided a framework for synthesis of the diffuse field from energy density envelopes and a guide for inclusion of the model in future hybrid schemes.

Finally, an approach to quantifying the behavior of the EST terms from high-frequency pressure simulations, facilitated by the development of the FVTD formulation in earlier work, opens an avenue for future characterization of the off-diagonal terms in free space. In lieu of exhaustive physical measurements of a space, a strategy for empirically determining the relationship between materials properties and EST method simulation coefficients for a

particular space was proposed, while also providing insight into how fully three-dimensional modeling of the wave-stress tensor may proceed.

7.2 Future work

Parallel research into the suitability of coupling boundary problem and free field solutions of the EST equations showed that such an approach was incorrect. A new approach based on Riemannian tessellation may illustrate the path forward for directly simulating the EST based on scattering that occurs as a result of boundary geometry instead of wavelength-scale roughness. In this formalism, the uncertainty principle ties wavelength to the density of diffracting locations in an image-source-like decomposition of room geometry, facilitating a measure of the degree of diffusion as a function of frequency based purely on the shape of a room. Future development of this idea including absorption at the boundaries is promising in terms of characterizing anisotropic and frequency-dependent diffusion. Furthermore, it will help clarify how to numerically approach the direct simulation of the EST in terms of its conservation equations at the boundaries and in free space.

Secondly, further attention to the characterization and analysis of the stochastic portion of measured impulse responses is warranted. While this thesis primarily relied on a standard measure of reverberation time, T_{30}, in order to characterize frequency-dependent temporal decay, many rooms that exhibit multi-slope decays are not adequately represented by this measure, and additionally, it may not be the best measure for characterizing the decay of the stochastic reverberation, especially depending on the overall strength of the direct sound. Therefore, both in the general sense, where an improved understanding of how to extract the incoherent portion of an impulse response directly informs our ability to measure the stochastic reverberation, and in the specific case where those measurements are used to generate or evaluate model parameters, future work on the topic is certain to bear fruit.

While a foothold has been gained in the pursuit of this thesis toward the application of the EST method in large three-dimensional spaces, prediction of simulation parameters *a priori* from domain geometry and materials properties requires further research. These two subjects, that of room geometry and characterization of the effects of scattering on the development of the stochastic reverberation, are both intimately tied to the degree of diffusivity in a particular space, or put another way, the rate at which an acoustic field becomes diffuse. Better understanding of how these aspects interact, as well as how to evaluate existing spaces in terms of signal processing, may allow better characterization of diffusing spaces as well as enable more powerful application of the EST method in appropriate contexts.

7.3 Perspectives

While this thesis is coming to a close on the subject of the acoustic energy-stress tensor, it is our hope that it will generate further interest on the topic given the promise of the approach for modeling stochastic reverberation in an efficient and physically relevant manner. Furthermore, as it represents the synthesis of a number of acoustical themes, including reverberation, mixing, scattering, and statistical acoustics, even venturing to soundfield analysis as it relates to the equipartition or decorrelation of diffuse energy as well as anisotropy, we hope to inspire collaboration on the subject as it pertains to our theoretical understanding of room acoustics beyond the practical advantages of the approach.

As outlined at the beginning of this thesis, statistical models of acoustics have long been used for their efficiency and perceptual effectiveness, but it is only now, as increasingly accurate and performant models of the early reverberation are becoming more accessible, that improving our ability to characterize and represent the late reverberation again presents an important avenue for acoustical research. Somewhere between the art of reverberator design and the computational complexity of wave-based acoustics, we hope that the energy-stress tensor approach will one day become an efficient and deterministic solution for the rapid evaluation and reproduction of late soundfields, perhaps bringing a degree of tractability to some of the most challenging acoustic environments in real time acoustical simulations.

Part IV

Appendices

Appendix A

Publications

This appendix contains publications pertaining to the thesis. Please note that as a byproduct of their inclusion in this document that all hyperlinks have been stripped. Please consider acquiring an original version if this functionality is desired.

PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

Lower bound on frequency validity of energy-stress tensor based diffuse sound field model

Aidan MEACHAM[(1)], Roland BADEAU[(2)], Jean-Dominique POLACK[(1)]

[(1)]Sorbonne Université, UMR CNRS 7190, Institut Jean le Rond ∂'Alembert, France
[(2)]LTCI, Télécom ParisTech, Université Paris-Saclay, France
[(1)]aidan.meacham@sorbonne-universite.fr

Abstract

A lower bound on the frequency validity limit is established for an energetic wave equation derived from the energy-stress tensor, examined in the one-dimensional case [Dujourdy et al, Acta Acustica united with Acustica 103:480-491, 2017]. The method efficiently models diffuse sound fields that dominate reverberation at higher frequencies and larger distances. Initially noted in the course of an exhaustive search of the solution space of all valid model parameters, the low-frequency cutoff has implications for the utility of the method in a hybridization context. In practice, the bound is encountered when determining the absorption and diffusion coefficients by iteratively approaching the temporal and spatial decay of measured data. As the test frequency decreases, the ranges of coefficient combinations that result in less than 10% variation from each decay measure can diverge until the region where both measures are satisfactory (the intersection of the two domains) disappears. Further evidence for the bound is provided through comparison with measurements of a long hallway, and stability concerns in the cases where both coefficients are very small are addressed.

Keywords: Room acoustics, finite difference methods, diffuse field

1 INTRODUCTION

While geometric and wave equation based approaches to room acoustics have dominated theoretical research, in the realm of practical implementations, hybrid methods combining both strategies are increasingly popular. Since each approach has strengths and weaknesses depending on the frequency range that is being simulated, it has long been the goal of practicioners to leverage different types of simulations in a cohesive manner to create full-band predictions with efficiency and accuracy that would not be possible with a single type alone. For example, geometric acoustics simulations such as ray tracing (1) or the image source method (2) efficiently model the early part of an impulse response, even at high frequencies, but cannot account for diffraction in a physically meaningful way, an important part of the low frequency response, and become inefficient when considering complex enclosure geometry or when directly modeling the late reverberation of a space. On the other hand, wave equation based methods (such as finite element or volume methods in both the time and frequency domains) excel in the low frequency regime, since they simulate the sound field everywhere in the enclosure, including diffraction and directional effects, with relatively direct implementations of sources and receivers (3, 4). Extending wave modeling methods to higher frequencies, however, requires increasingly fine meshes that become computationally untenable for real-time applications, even with the advent of modern GPU implementations (5).

From this brief overview, it is not difficult to predict how hybrid approaches could combine these methods, utilizing each method in its most effective domain in order to create a full-band model. In practice, however, there are a number of difficulties that arise from combining separate methods in this fashion, not the least of which is the maintenance of machinery (including but not limited to numerical code, meshing tools, automatic matching of boundary conditions across simulation types, room geometry simplifications, and post-processing DSP) for each separate part of the simulation. In this paper, we consider how an energetic wave equation based method mitigates some of these concerns while considering its limitations for future hybridization.

This method, developed by Dujourdy et al. (6), has been found to accurately model the statistical late reverberation, or diffuse field, resulting from an initial perturbation, in both 1- and 2-dimensional contexts (7). Originating from Ollendorff and Picaut's work (8, 9), this method aims to take advantage of an additional degree of freedom given by the inclusion of a diffusion coefficient in addition to the traditional absorption coefficient. Previous work on this topic considered only the 1000 Hz frequency band, but in order to consider its future fitness for hybrid modeling, we will consider in greater detail the frequency ranges it is capable of modeling. Such information provides bounds on the crossover frequencies that would be required in a hybrid context, implying the necessary ranges that would need to be simulated with other methods in order to create a full-band model, and thus whether or not the diffuse field strategy can be useful in a practical architectural acoustics context.

2 ENERGETIC WAVE EQUATION METHOD

For this study, we directly used the theory and implementation of the numerical schemes in (6). For the sake of brevity, we will not reproduce those findings here, but encourage those desiring a full treatment to refer to the original work.

In brief, the main idea of the energetic wave equation is the development of two conservation laws, one each for energy density E and sound intensity J, as defined by Morse and Feshback in (10), rather than the typical energy balance based on Gauss's theorem found in typical finite difference or finite volume formulations. Using these two laws, it is possible to derive a system of coupled equations relating the energy density, the sound intensity, and the wave-stress symmetric tensor $\underline{\underline{E}}$ (11). In the case where one length dominates a space and the cross-sectional area is relatively constant (as is the case in a long hallway), this system of equations can be reduced to one dimension by introducing modified absorption and scattering coefficients and integrating (with further energy and momentum balance hypotheses) over the minor axis walls.

After dimensional reduction, the system resembles the telegrapher's equations, and by inspection can be transformed into a linear second-order hyperbolic equation with a single dependent variable (in this case, the energy density E). Finally, boundary conditions can be derived using the aforementioned hypotheses, and the entire wave propagation system can be discretized with common finite difference time domain strategies.

The main advantage of this approach is that the propagation of diffuse energy allows for very large spatial discretization. The reason for this is that the modulation frequency of the late energy decay that we are concerned about is very low. Typical acoustic wave simulations that are concerned with high frequency content have a very high modulation frequency (since they specifically want to resolve the individual pressure waves), but since we are primarily interested in the stochastic decay of the diffuse high frequency energy, we can assume that the decay itself is not changing very rapidly, and thus can accommodate relatively large spatial sampling. As an example, if we are only concerned with the diffuse energy level every tenth of a second, then the spatial sampling rate can be as coarse as 3 meters, with the temporal sampling rate chosen to satisfy the scheme's stability conditions.

3 FREQUENCY VALIDITY ANALYSIS

3.1 Room geometry

A similar corridor to the original was used for this study. The hallway had an overall length of 45 meters, with a width of 159 cm and a height of 237.5 cm. In the main narrow portion of the hallway, the ceiling height was measured to a fine metal grating suspended below the actual height of the corridor, which carried lighting and conduit. The hard ceiling was 326 cm, with a decrease every 1.5 meters for metal support beams to 280 cm.

In the recesses, all of which were of uniform length and depth (except for the doorway furthest on the right of the plan), the width increased to 239 cm and the height decreased to 220 cm, and ceiling was masonry rather than the grating mentioned above. In some of the recesses, there were glass display cases or small pieces of

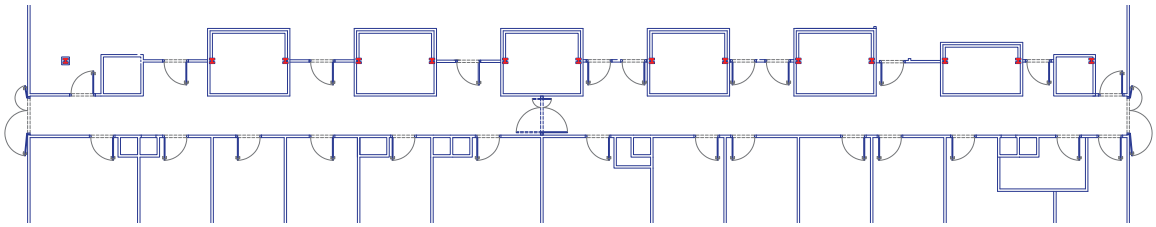

Figure 1. Floorplan for the corridor under consideration.

furniture that contributed to the diffusive effects of the recesses themselves. The entire floor was linoleum, and for the most part, the walls were wooden panel and masonry. There were occasional metal gratings on the flat wall for HVAC, which was audible but not distracting. All doors entering the hallway were closed, and the doors in the center of the hallway were fully open.

3.2 Measurements

Table 1. T_{60}s and Spatial Decays

Frequency [Hz]	62.5	125	250	500	1000	2000	4000	8000
T_{60} [S]	2.29	0.35	0.38	0.39	0.43	0.35	0.33	0.30
Spatial Decay [dB/m]	-0.63	-0.49	-0.88	-0.94	-0.66	-0.54	-0.66	-0.84

Impulse responses were collected using a *SoundField* ST250 microphone and an *Outline* GRS omnidirectional speaker, with a *MOTU* Traveler sound card. As before, the source was positioned 1 meter away from the end of the hall, 1.5 meters above the ground, and centered between the two walls. Beginning 1 meter from the source, measurements were collected with the microphone's X-axis aligned along the length of the hallway. A spacing of 1 meter was used out to 10 meters, which corresponds exactly to the discretization distance in the numerical simulations. Then, recordings were made every 2 meters until the end of the hallway for a total of 26 sampling locations. Recordings were made using the the swept sine method (12, 13) as implemented in the *Adobe Audition* plugin *Aurora*. The sweep length was 20 seconds, and the source level was adjusted digitally to maximize the signal-to-noise ratio without clipping as the microphone was moved further and further from the source. These gains were recorded in order to recover the true measured energy level for each measurement location. The sweep responses were then post-processed by convolution with the inverse sweep to recover impulse responses for each location.

The relevant frequency bands were selected by filtering each impulse response with a standard octave-band filterbank, and then calculating the desired metrics for each resulting bandlimited response. The T_{60} (a measure of temporal decay) was calculated by a linear fit to the Energy Decay Curve in dB for each frequency band and at each receiver location. This is sometimes known as Schroeder's reverse integration (14). Next, each receiver location was averaged to arrive at a single T_{60} for the entire hallway in each band. Finally, spatial decays (similar to strength of sound, or G) were calculated by a linear fit to the sum of energy for each band-limited response across each receiver location. Both sets of calculated values are presented in Table 1.

3.3 Simulations

In order to determine the frequency bands the model was able to represent, we ran simulations for a sampled subset of all combinations of absorption and scattering coefficients α and β. In a typical scenario, a scheme to converge on the desired spatial and temporal characteristics would be used to match a single frequency band, but here, we want to visualize the entire problem space.

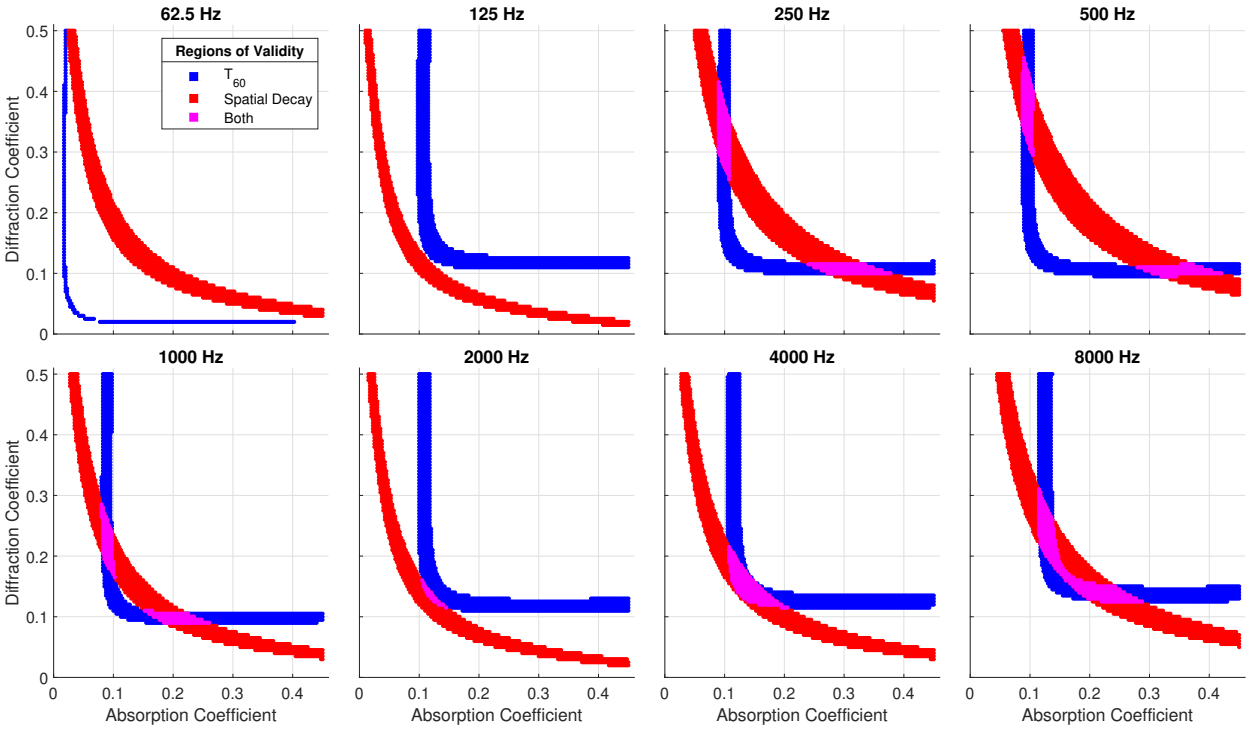

Figure 2. Agreement between simulated and measured data for the hallway under consideration.

For this study, as was the case before, the spatial sampling step was chosen to be $\Delta x = 1$ meter.

The initial conditions were chosen to be a temporal Gaussian, centered at the same location as the source, in order to minimize spurious numerical oscillations. Afterward, the same acoustical indices as with the *in situ* measurements were extracted from each result.

3.4 Valid domains

Finally, we compared the measured and simulated results for each index. For each frequency band (presented here in ascending order), we classified which simulations were within 10% of each desired metric, the spatial and temporal decay rates. To preserve legibility, we have represented in Figure 2 only the region with the most relevant combinations of coefficients for this hallway. The absorption coefficient α ranges from 0.01 to 0.45, and the diffusion coefficient β ranges from 0.01 to 0.5. This region corresponds well to the typical ranges where both T_{60} and spatial energy decays are valid.

As expected, the two coefficients have different effects on the simulations, given that the validity patterns are not symmetric, confirming the observation in the original paper. Furthermore, the regions of validity for each individual measure appear to be smoothly varying. Perhaps most importantly, for some low frequencies, there is no region where both indices are valid, implying a lower bound on the frequency that can be represented by the model for this hallway.

At the lowest frequency band, 62.5 Hz, there are regions of validity outside of the chosen region, but we have chosen to exclude these as either their absorption or diffusion coefficients are unrealistically small or large. Since the regions of validity appear to smoothly vary, it may be the case that with a particular frequency in between this band and 125 Hz, there is a point where the spatial decay and T_{60} curves "switch" places and become valid. We also note the similarity of the 125 Hz band and the 2000 Hz band, since the region of validity is determined by the 10% threshold. In other bands, the regions cross, and therefore can be uniquely

determined to match both measures to arbitrary precision without the need of a threshold. Both of these topics related to a more continuous perspective on the methodology for determining validity deserve closer examination in future work.

The case where the individual metric regions cross resulting in two disconnected regions of validity, up to the lack of symmetry discussed above, can be explained by the observation by Dujourdy et al. (6) that the two coefficients can typically be exchanged, where the smaller of the two acts as the absorption coefficient. While in general this situation is preferable, it implies that the optimization problem, at least in its current formulation, is not convex, and furthermore, more information about the problem will be required in order to uniquely determine a preferred combination of coefficients.

One possible explanation for the lack of a valid region at the lowest frequencies is the mismatch between long T_{60} and relatively low spatial decays. At low frequencies, it is common to have more reverberation, and similarly, the energy may be more evenly distributed as a result of diffusion. At higher frequencies, the opposite is the case, where increased absorption of high-frequency energy results in shorter reverberation times and steeper spatial decay, both of which appear to be more readily represented by this model.

In the case where either coefficient is very small, generally speaking, the simulations become closer and closer to the lossless case, extending the response times and magnifying small perturbations. As a result, simulations with some combinations of α and β can veer into instability. This would further constrain the region of validity, but in practice, instability only occurs when approaching the most extreme values of T_{60} or spatial decay. Thus, such instability can simply serve as a indicator that careful attention should be paid to the physical measurements that are being matched. It appears that since the region of validity at the lowest frequencies is reasonably well balanced between the two coefficients, for this type of hallway, instability should not present a major concern in the context of hybridization since the most dangerous simulations can simply be discarded.

3.5 Implications

Since the lowest frequency that contains a convincing overlap is 250 Hz, for this corridor, that is the lowest bound on the validity of the model. This makes sense, seeing as the late reverberation is more statistical at higher frequencies. For small concert halls, 250 - 500 Hz is a very convenient cutoff frequency for a potential hybridization since a complementary pressure wave time domain approach would be capable of modeling the region in real time (5). Thus, at least for this case, it is worthwhile to consider the model as a possible candidate for accelerating the synthesis of the late tail in a full-band audio simulation.

3.6 Other hallways

In order to confirm these results, the procedure was followed using the recordings from the previous hallway. In the interest of space, we will not fully reproduce the measurements as we have with the current study, but the findings were consistent with the conclusions we have drawn. The corridor from the initial work exhibited regions of validity down to 125 Hz. Given that the two were relatively similar in terms of shape and length, this is perhaps unsurprising, but a useful confirmation nonetheless.

A third unrelated hallway of a similar length but with fewer diffusing surfaces was also measured and was only representable (with “crossed” validity regions) by the model in the 125 Hz and 1000 Hz bands. Upon initial review of these results, it appears that the reason for the lack of validity is primarily due to the reduced spatial decays across all bands. Because the hallway did not have as many absorptive surfaces, the sound energy was much more evenly distributed throughout or even modal, and therefore presented difficulty for the model, which is predicated on the assumption of spatial decay. One possible interpretation is that this incompatibility between temporal and spatial decays is a manifestation of a truly uniform reverberant field in regions far from the source, where the so-called diffuse field theory is valid. Thus, the case where the energy-based model fails may actually be a criterion for spaces where the Sabine equation holds.

Figures displaying the regions of validity for both of these corridors can be found in the appendix.

4 CONCLUSIONS AND FUTURE WORK

In this paper, we have discussed our findings regarding the lower frequency bound for which an energy-based late-reverberation model can accurately represent the stochastic soundfield. The current study focused on a one-dimensional model of a long hallway for which the lower bound was 250 Hz. This means that above this cutoff frequency, the model was able to recreate a diffuse soundfield with acoustical indices consistent with measurements made in a real hallway within 10% error.

The investigation of this lower frequency bound brings to mind a number of questions for future research. In this case, we have examined a single hallway with a given length and certain acoustical indices, but it should be possible to extend the analysis to one dimensional problems of varying lengths, T_{60}, and spatial decays, such that any problem dominated by its length with a sufficient amount of diffusion could be verifiably modeled above a certain frequency. Furthermore, given the discovery that in some frequency bands there are multiple regions of validity, it should be possible to examine more optimal methods for converging on desired coefficients for a given physical measurement in an interative fashion. In such a fashion, it may be possible to better characterize the contours of the underlying search space in order to accelerate convergence.

In terms of bringing the theoretical model into practical application, it is imperative to begin experimenting with a sonification procedure. While the theoretical aspects of the method are interesting in and of themselves, the real value of such an approach is the additional accuracy and efficiency that would be enjoyed upon integration with a full-band hybrid auralization system. To that end, the next step in developing the model is extending the current 1- and 2-dimensional models to 3 dimensions.

Finally, given the large spatial discretization step in this method, it may be the case that the energy model and a pressure wave approach could use the same grid, reducing the meshing that must be completed before simulations can be performed. Verifying that meshes (or subsets thereof) could be trivially reused for both methods in a hybrid context would be especially relevant in the cases where room geometry is time varying, a possibility that deserves closer inspection.

REFERENCES

1. Krokstad A, Strom S, and Sørsdal S. Calculating the Acoustical Room Response by the Use of a Ray Tracing Technique. Journal of Sound and Vibration 1968;8:118–25.
2. Allen JB and Berkley DA. Image Method for Efficiently Simulating Small-room Acoustics. The Journal of the Acoustical Society of America 1979;65:943–50.
3. Savioja L, Rinne TJ, and Takala T. Simulation of Room Acoustics with a 3-D Finite Difference Mesh. In: *Proc. ICMC*. 1994.
4. Botteldooren D. Acoustical Finite-difference Time-domain Simulation in a Quasi-Cartesian Grid. The Journal of the Acoustical Society of America 1994;95:2313–9.
5. Savioja L. Real-Time 3D Finite-Difference Time-Domain Simulation of Low- and Mid-Frequency Room Acoustics. In: *DAFx*. 2010:8.
6. Dujourdy H, Pialot B, Toulemonde T, and Polack JD. An Energetic Wave Equation for Modelling Diffuse Sound Fields – Application to Corridors. Acta Acustica united with Acustica 2017;103:480–91.
7. Dujourdy H, Pialot B, Toulemonde T, and Polack JD. An Energetic Wave Equation for Modelling Diffuse Sound Fields—Application to Open Offices. Wave Motion. Innovations in Wave Modelling II 2019;87:193–212.
8. Ollendorff F. Statistical Room-Acoustics as a Problem of Diffusion (A Proposal). Acta Acustica united with Acustica 1969;21:236–45.
9. Picaut J, Simon L, and Polack JD. A Mathematical Model of Diffuse Sound Field Based on a Diffusion Equation. Acta Acustica united with Acustica 1997;83:614–21.

10. Morse PM and Feshbach H. Methods of Theoretical Physics. Mc Graw-Hill Book Company, 1953.
11. Morse PM and Ingard KU. Theoretical Acoustics. Princeton University Press, 1968. 949 pp.
12. Farina A. Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. In: Audio Engineering Society Convention 108. Audio Engineering Society, 2000.
13. Farina A. Advancements in Impulse Response Measurements by Sine Sweeps. In: Audio Engineering Society Convention 122. Audio Engineering Society, 2007:21.
14. Schroeder MR. New Method of Measuring Reverberation Time. The Journal of the Acoustical Society of America 1965;37:409–12.

APPENDIX

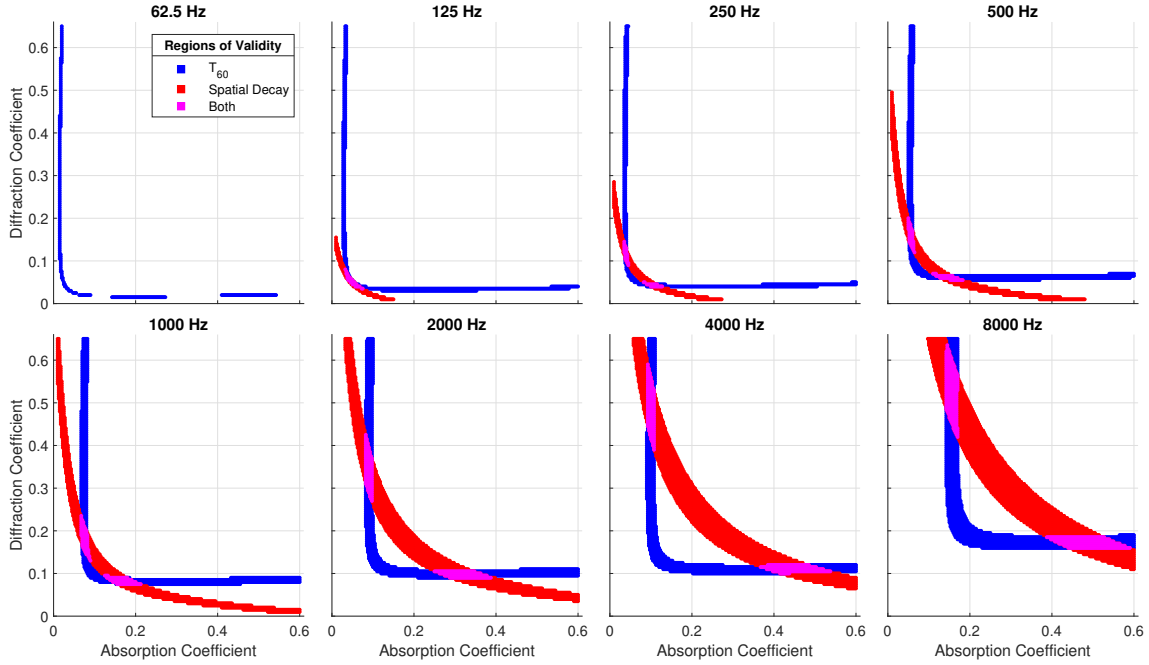

Figure 3. Agreement between simulated and measured data for the hallway studied in Dujourdy et al.

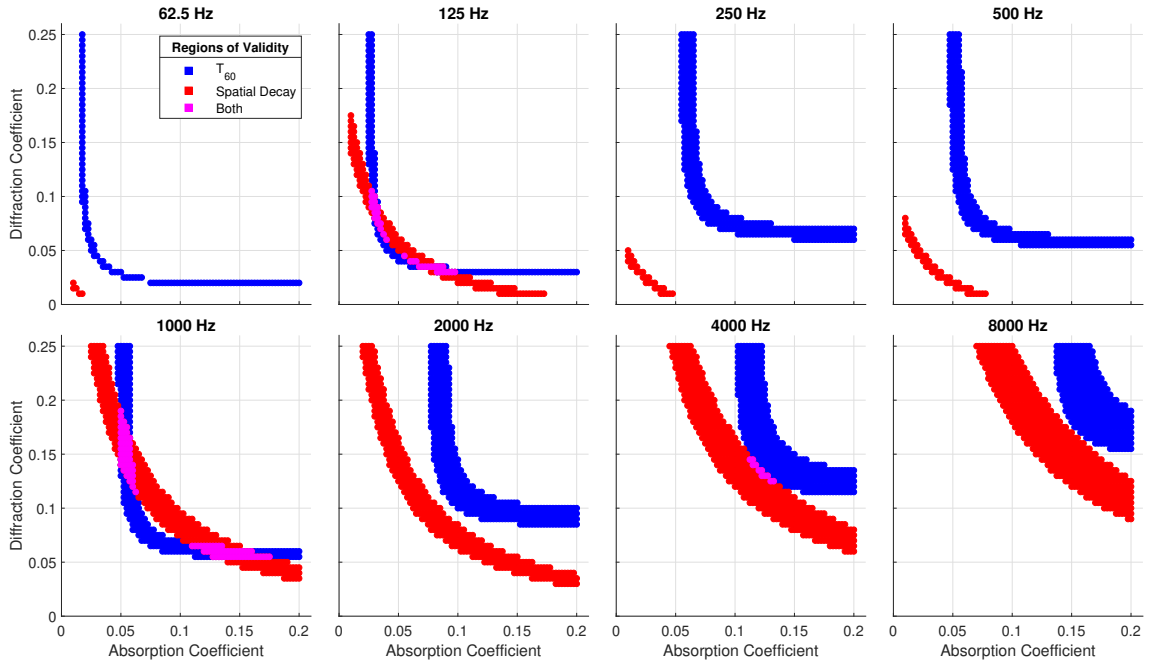

Figure 4. Agreement between simulated and measured data for a hallway with fewer diffusing surfaces.

ISRA 2019 AMSTERDAM

PROCEEDINGS of the
International Symposium on Room Acoustics

15 to 17 September 2019 in Amsterdam, Netherlands

NAG

Implementation of Sources in an Energy-Stress Tensor Based Diffuse Sound Field Model

Aidan MEACHAM[1]; Roland BADEAU[2]; Jean-Dominique POLACK[1]

[1]Sorbonne Université, UMR CNRS 7190, Institut Jean le Rond ∂'Alembert, France

[2]LTCI, Télécom Paris, Institut Polytechnique de Paris, France

ABSTRACT

An implementation of acoustic sources is developed in the context of an energetic wave equation derived from the energy-stress tensor, examined in the one-dimensional case [Dujourdy et al, Acta Acustica united with Acustica 103:480-491, 2017]. The method efficiently models diffuse sound fields that dominate reverberation at higher frequencies and larger distances. Monopole and dipole electroacoustical sources are considered. Using loudspeaker models rather than idealized distributions of sound energy allows for a convenient structure to evaluate directional dependence and frequency dependence for a variety of source types. Compared to initial condition formulations, an explicit source term enables realistic modeling of complex sound sources with the possibility of spatial changes in time. A finite volume time domain (FVTD) approach is utilized to lay the groundwork for future extensions to three dimensions. The spatially invariant model parameters are determined iteratively by comparison with in situ measurements of a long hallway for both the monopole and dipole case in order to verify the validity of the framework.

Keywords: Room acoustics, finite difference methods, diffuse field

1. INTRODUCTION

Previous work, focused on an energetic wave equation based method for simulating room acoustics, has relied upon bespoke dimensional reduction for tractability. While this approach developed by Dujourdy et al. (1, 2) has been successful in 1- and 2-dimensional contexts, the extension to 3 dimensions presents a challenge.

This energetic wave equation has been found to accurately model the statistical late reverberation, or diffuse field, resulting from an initial perturbation. Originating from Ollendorff and Picaut's work (3, 4), it takes advantage of an additional degree of freedom given by the inclusion of a diffusion coefficient in addition to the traditional absorption coefficient resulting from deriving a wave equation based on the energy density and sound intensity rather than pressure and velocity. The main reason to consider this model over more traditional finite difference schemes is that it is highly efficient, both because the element sizes can be very large while still capturing relevant behavior, and because the late reverberation as predicted by more fine-grained models tends to be sensitive to perturbations that result from non-physical modeling phenomena such as dithering.

These solutions used finite difference time domain (FDTD) techniques to numerically model the resulting systems, a common formalism for discretizing partial differential equations. Recently, finite volume time domain (FVTD) approaches have become more popular since they allow for unstructured meshes and also provide convenient machinery to confirm that conserved values are in fact accounted for in terms of storage and dissipation (5). Furthermore, they provide a straightforward programmatic framework that can accommodate 1-, 2-, or 3-dimensional problems with little change in structure.

In many acoustic FDTD or FVTD models, it is common to examine an unforced system where the evolution of acoustical phenomena through time is determined only by the initial distribution of pressure and velocity. This is a convenient formulation in the case where the model will be used to generate an impulse response, as an idealized omnidirectional source can be simply represented, often with a spatial Gaussian. There are some cases where it is useful to consider the forced system,

[1]aidan.meacham@sorbonne-universite.fr

Peer reviewed

however, where source terms are included in the wave equation. Some examples of behavior that is more readily modeled by this approach include feedback or cases when the source position or directivity change over time.

In this study, we lay the groundwork for extending the energetic wave equation model to 3 dimensions in the FVTD formalism by recontextualizing a 1-dimensional problem, while also considering sources of acoustic energy other than initial conditions in order to accommodate many types of sources.

2. THEORY

2.1 Definitions

Following directly from Dujourdy et al. (1), we begin from a common acoustical model, $\frac{1}{c^2}\partial_{tt}\Psi - \Delta\Psi = 0$. The velocity potential Ψ is defined by $\mathbf{v} = -\nabla\Psi$ and $p = \rho\partial_t\Psi$. In this system, the sound pressure and particle velocity vector are given by p and $\mathbf{v}$ respectively, ∇ is the gradient operator, Δ the Laplacian operator, ρ the air density, and c the speed of sound. Finally, we notate ∂_i and ∂_{ii} the first and second derivatives according to coordinate i respectively.

2.2 Volume velocity sources

In order to introduce sources to the energetic wave equation, we can proceed directly to Equation 8 from Dujourdy et al. (1), with the divergence operator $(\nabla\cdot)$:

$$\frac{1}{c}\partial_t E_{tt} + \nabla\cdot\mathbf{J} = 0,$$
$$\frac{1}{c}\partial_t\mathbf{J} + \nabla\cdot\begin{pmatrix} E_{xx} & E_{yx} & E_{zx} \\ E_{xy} & E_{yy} & E_{zy} \\ E_{xz} & E_{yz} & E_{zz} \end{pmatrix} = 0.$$

This system presents a relation between the energy density $E = \frac{\rho}{2}(\frac{1}{c^2}|\partial_t\Psi|^2 + |\nabla\Psi|^2)$ and the sound intensity $\mathbf{J} = -\rho\partial_t\Psi\nabla\Psi/c$ as defined by Morse and Feshback (6), and the wave-stress symmetric tensor as defined by Morse (7) with components $E_{ii} = \frac{\rho}{2}(\frac{1}{c^2}|\partial_t\Psi|^2 + \sum_j \alpha_{ij}|\partial_j\Psi|^2)$ or $E_{ij} = \rho\partial_i\Psi\partial_j\Psi$ for $i,j = x,y,z$ with $\alpha_{ij} = 1$ when $i = j$ or -1 otherwise.

Dimensional analysis shows that we can represent sources in the wave equation with the inclusion of source terms on the right hand side of each part that approximate the effects of a moving membrane in the acoustic space.

If Q is the volume velocity of a source, and p and $\mathbf{v}$ are the pressure and particle velocity immediately in front of that source, we can write

$$\frac{1}{c}\partial_t E_{tt} + \nabla\cdot\mathbf{J} = \mathcal{P} = \frac{pQ}{c},$$
$$\frac{1}{c}\partial_t\mathbf{J} + \nabla\cdot\begin{pmatrix} E_{xx} & E_{yx} & E_{zx} \\ E_{xy} & E_{yy} & E_{zy} \\ E_{xz} & E_{yz} & E_{zz} \end{pmatrix} = \mathbf{Q} = \rho\mathbf{v}Q. \tag{1}$$

When the surface of the source moves into the volume of the room, Q and p are positive, and $\mathbf{v}$ (which is in the same direction as Q by continuity) on the surface is counted positively when the inward normal of the source surface into the room is orientated toward the positive X-axis. In order to account for a variety of source types and topologies, the pressure in front of a given source can be calculated in terms of frequency dependent characteristics such as the Thiele-Small parameters, the radiation impedance of a single driver, or the mutual impedance of an enclosure with multiple drivers. We note that this approach is an anechoic approximation to the impedance problem, as the reverberation within the space or other sources also contributes to the pressure in front of a given radiating structure, creating a coupled system. One justifying explanation is that in the frequency range we are expecting to model (generally above the Schroeder frequency), the free-field radiation impedance dominates, and thus we are free to treat sources individually and without considering the local acoustic conditions.

For a monopole source, $\mathcal{Q}$ and p have the same sign on both sides of the source, so $\mathcal{P}$ is positive. On the other hand, $\mathbf{v}$ points in opposite direction on either side of the source such that $|\mathbf{Q}|$ is zero.

For a dipole source such as an open baffle loudspeaker, $\mathcal{Q}$, p, and $\mathbf{v}$ all have opposite signs on each side of the radiating surface. As a result, both $\mathcal{P}$ and $\mathbf{Q}$ are positive, with $|\mathbf{Q}| = \mathcal{P}$. Of course, dipole sources also imply particular frequency response effects that can be applied according to the frequency range of a given simulation.

For other types of non-idealized directional sources, directional radiation along the primary axes can be introduced according to near-field analysis of the source topology and geometry. Thus, modeling a sealed enclosure loudspeaker with a single driver, for example, would account for diffractive effects around the cabinet via contributions to $\mathbf{Q}$ corresponding to the principal axes of the wave-stress tensor.

Favorably, this description allows for a wide variety of sources, and can accommodate any system that can be parametrized in terms of its volume velocity output. While these extensions will not be addressed in this work, it is ostensibly possible to include arbitrary loudspeaker designs or other environmental sound sources. Thus, any collection of electromechanical devices can be represented in the model, which can be especially useful in predicting the behavior of sound reinforcement systems in enclosed spaces, for example.

2.3 FVTD model - spatial discretization

In order to proceed with a spatial discretization of this model, we note the similarity of these equations, which are denoted in terms of energy density and sound intensity, with the wave equation as written in terms of pressure and velocity:

$$\begin{aligned} \frac{1}{\rho c^2}\partial_t p + \nabla \cdot \mathbf{v} = 0, \\ \rho \partial_t \mathbf{v} + \nabla p = \mathbf{0}. \end{aligned}$$

In the present case, we explore a quasi-one-dimensional case in order to utilize previously suggested dimensional reductions to account for the off-diagonal terms in the wave-stress tensor $\underline{\underline{E}}$ for agreement with the scalar p. Otherwise, to physical constants, the systems exhibit identical behavior, so we can directly use a generalized FVTD formulation by inspection, replacing p and $\mathbf{v}$ with E and $\mathbf{J}$. In this work, we refer specifically to the straightforward development in Bilbao et al. (8).

The space to be analyzed is a long hallway aligned on the X-axis of length l_x, with width and height $l_y, l_z \ll l_x$. We discretize the hallway into N identical rectangular solids Ω_j (of volume $V = l_x l_y l_z / N$) associated with average energy densities E_j and outward sound intensities J_{jk} and J_l, depending if the intensity in question is incident upon another cell or a boundary. The distance between centroids of each cell is thus $h = l_x/N$, and since the hallway model has a constant section, the surfaces between cells are identical (of area $S = l_y l_z$) and the N_b boundary surfaces take on one of three areas depending on their normal direction ($l_x l_y/N$, $l_x l_z/N$, or $l_y l_z$, indexed as S_l below).

Finally, we can integrate the total $\mathcal{P}$ over each cell and average $\mathbf{Q}$ over each outward surface to arrive at indexed quantities $\mathcal{P}_j$ and Q_{jk}.

Rewriting Bilbao et al. (8) Equation 15 directly in terms of energy density and sound intensity, introducing terms for the wave-stress tensor, and including the previously derived source terms from Equation 1 results in

$$\begin{aligned} \frac{V}{c}\frac{dE_j}{dt} + \sum_{k=1}^{N} \beta_{jk} S J_{jk} + \sum_{l=1}^{N_b} \gamma_{jl} S_l J_l = \mathcal{P}_j, \\ \frac{1}{c}\frac{dJ_{jk}}{dt} + \frac{1}{h}(E_k - E_j) + \sum_{k,l=1}^{N,N_b} \zeta_{jkl}\frac{S_l}{V} E_j = Q_{jk}, \end{aligned} \tag{2}$$

where β_{jk} and γ_{jl} are indicator functions that are 1 if a given cell Ω_j shares a face with another cell Ω_k or face on the boundary S_l, and 0 otherwise. Additionally, ζ_{jkl} is an indicator function for a cell, boundary, and a neighboring cell that has a boundary that shares vertices with the first ("neighboring" boundaries, so to speak). In the configuration of cells under study, this indicator function adds eight terms for each cell that is not at an end - four times each for its neighbor in the positive and negative X direction - since all four of its boundaries (walls, ceiling, and floor) share vertices with

its neighbors' boundaries. In general, this term acts to bidirectionally damp the transfer of energy between cells through the sound intensity terms as an approximation of diffusion. One interpretation of this effect is that it acts to reduce the specularity of reflections that typically facilitates energy transfer in the direction of wave propagation. In other words, sound intensity directed from one cell to another, where each has boundaries that share vertices with their common surface, acts to reduce the energy transfer through that common surface.

We assume, as before, that sound is propagating in the X direction, where E and J_x are constant on a given section of the corridor, J_y is independent of z, and J_z is independent of y. This assumption is borne out by the decision to discretize the hallway into a single line of cells, since the energy can only transfer in between neighboring cells and the boundaries. Then, using the energy balance $J = AE/4$ and the momentum balance $E_{xy} = DJ/4$ from Equations 12 and 21 in Dujourdy et al. (1) with the modified absorption and diffusion coefficients A and D, we can replace the outward normal sound intensities incident on a boundary and the energy densities in front of diffracting boundaries to arrive at

$$\begin{aligned} \frac{V}{c}\frac{dE_j}{dt} + \sum_{k=1}^{N} \beta_{jk} S J_{jk} + \sum_{l=1}^{N_b} \gamma_{jl} S_l \frac{A}{4} E_{jl} &= \mathcal{P}_j, \\ \frac{1}{c}\frac{dJ_{jk}}{dt} + \frac{1}{h}(E_k - E_j) + \sum_{k,l=1}^{N,N_b} \zeta_{jkl} \frac{S_l}{V}\frac{D}{4} J_{jk} &= Q_{jk}, \end{aligned} \tag{3}$$

where E_{jl} indicates the energy density of a cell in front of a particular boundary surface.

One advantage of this approach, over integrating the boundary conditions on the walls and ceilings into a one-dimensional difference equation along the X-axis and solving for more complicated boundary conditions at the ends of the hallway, is that it allows us to apply simple boundary conditions at every discretized boundary while observing the same behavior as the previous model, with the side benefit of simplifying the implementation of spatial variance in the hallway (even though we will not consider such cases in this study).

Typically, FVTD approaches are used in cases where the boundaries are irregular, resulting in the need for unstructured elements. Of course, that is not the case here, such that the eventual implementation is a typical FDTD scheme over a structured grid. Nonetheless, it is convenient to proceed with the FVTD formulation since it provides a convenient formalism for examining the conservation of quantities that correspond to energy in a classic pressure-velocity wave equation. This context allows us to reason about the physical meaning of exchanges between energy density and sound intensity.

2.4 FVTD model - temporal discretization

From the formulation above, we can discretize the elements that are continuous in time by approximating the temporal derivatives with temporal differences evaluated at a given sampling rate. The main advantage of the energetic wave equation model is that the modulation frequency of the late energy decay is very low, which implies that the spatial sampling step can be very large and the temporal sample rate can be low as well. This is in contrast to typical pressure and velocity wave equation approaches, where the goal is to resolve the highest frequency, leading to smaller discretization steps temporally and spatially, and thus, to simulations that are difficult to run in real time, even with accelerated processing (that may take place on a GPU, for example).

In terms of implementation, these conclusions justify the earlier decision to discretize the hallway into very large cells that do not require multiple divisions along the width and height, and can therefore utilize the one-dimensional approximations discussed earlier. Were this not the case, the difference equations relating neighboring cells would need to be more complicated in order to deal with the full wave-stress tensor, and the simplifications made for the boundary conditions would also not be valid.

With a discrete time step T such that $f^n = f(nT)$ is a discrete time approximation of the continuous time series $f(t)$, we introduce the forward and backward shift operators $e_+ f^n = f^{n+1}, e_- f^n = f^{n-1}$, the forward and backward difference operators $\delta_+ = (e_+ - 1)/T, \delta_- = (1 - e_-)/T$, and the forward and backward averaging operators $\mu_+ = (e_+ + 1)/2, \mu_- = (e_- + 1)/2$. Then, by replacing the continuous derivatives in Equation 3 with interleaved approximations of the energy density and sound intensity (as in Bilbao et al. (8) Section IV), a fully discrete version of Equation 3 including

the source terms can be written as

$$\begin{aligned} \frac{V}{c}\delta_+ E_j + \sum_{k=1}^{N} \beta_{jk} S J_{jk} + \sum_{l=1}^{N_b} \gamma_{jl} S_l \frac{A}{4} \mu_+ E_{jl} &= \mathcal{P}_j, \\ \frac{1}{c}\delta_- J_{jk} + \frac{1}{h}(E_k - E_j) + \sum_{l=1}^{N_b} \zeta_{jkl} \frac{S_l}{V} \frac{D}{4} \mu_- J_{jk} &= Q_{jk}, \end{aligned} \tag{4}$$

where $\mathcal{P}_j$ and Q_{jk} are aligned in time with J_{jk} and E_j respectively, and we introduce temporal averaging to preserve the time alignment and differential relationship in each equation.

Then, using the notation $e_\pm f = f^\pm$, expanding the temporal operators, and solving for E_j^+ and J_{jk}:

$$\begin{aligned} E_j^+ &= \frac{E_j(1 - \frac{cT}{V}\sum_{l=1}^{N_b} \gamma_{jl} S_l \frac{A}{8}) + \frac{cT}{V}\mathcal{P}_j - \frac{cT}{V}\sum_{k=1}^{N} \beta_{jk} S J_{jk}}{1 + \frac{cT}{V}\sum_{l=1}^{N_b} \gamma_{jl} S_l \frac{A}{8}}, \\ J_{jk} &= \frac{J_{jk}^-(1 - \sum_{l=1}^{N_b} \zeta_{jkl} \frac{S_l}{V} cT \frac{D}{8}) + cT Q_{jk} + \frac{cT}{h}(E_j - E_k)}{1 + \sum_{l=1}^{N_b} \zeta_{jkl} \frac{S_l}{V} cT \frac{D}{8}}. \end{aligned} \tag{5}$$

This is a realizable two-step FVTD scheme.

Thus, in order to integrate arbitrary source material, the output from the loudspeaker model is bandpass filtered to the selected octave, converted to acoustic power, decimated to fit the much lower simulation sample rate, and injected into the cells corresponding to the position of the source. By repeating the process for multiple octave bands, a full-band response can be synthesized, though it is important to note that an auralization scheme for these results has not yet been realized. Of course, the inclusion of the source in these equations does not preclude also accounting for initial conditions, which may be set just as before by directly specifying the energy density and sound intensity at the first two time steps.

3. FITTING MODEL PARAMETERS TO MEASUREMENTS

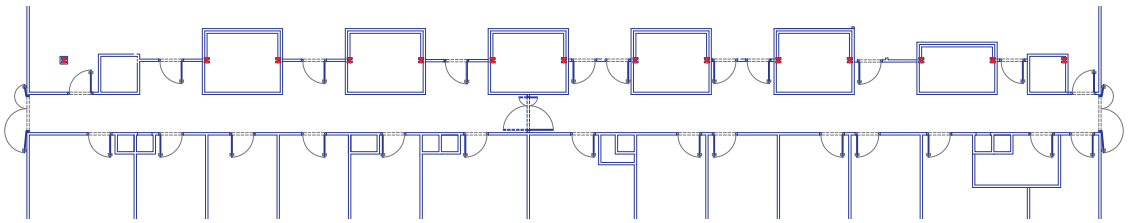

Figure 1 – Floorplan for the corridor under consideration.

3.1 Room geometry

The hallway we examined for this study had an overall length of 45 m, with a width of 1.59 m and a height of 2.375 m. A plan of the corridor can be seen in Figure 1. In the main narrow portion of the hallway, the ceiling height was measured to a fine metal grating suspended below the actual height of the corridor, which carried lighting and conduit. The hard ceiling was 3.26 m, with a decrease every 1.5 m for metal support beams to 2.8 m.

In the recesses, all of which were of uniform length and depth (except for the doorway furthest on the right of the plan), the width increased to 2.39 m and the height decreased to 2.2 m, and ceiling was masonry rather than the grating mentioned above. In some of the recesses, there were glass display cases or small pieces of furniture that contributed to the diffusive effects of the recesses themselves. The entire floor was linoleum, and for the most part, the walls were wooden panel and masonry. There were occasional vertical metal gratings on the flat wall over shallow indentations. The HVAC was audible but not distracting, and did not present an obstacle for the calculation of acoustical indices. All doors entering the hallway were closed, and the doors in the center of the hallway were fully open.

3.2 Measurements

Impulse responses were collected using a *SoundField* ST250 microphone and an *Outline* GRS omnidirectional speaker, with a *MOTU* Traveler sound card. The source was positioned 1 meter away from the right end of the hall in Figure 1, 1.5 m above the ground, and centered between the two walls. Measurements were collected with the microphone's X-axis aligned along the length of the hallway at distances of 4, 8, 12, and 16 m from the source. For one set of recordings, all of the speaker drivers were active, approximating a point source, and for a second set, only the sets of 3 speakers aligned along the X-axis were activated (in groups of 3), with opposite phases, to approximate a dipole source. Recordings were made using the swept sine method (9, 10) as implemented in the *Adobe Audition* plugin *Aurora*. The sweep length was 20 seconds, and the source level was adjusted digitally to maximize the signal-to-noise ratio without clipping. All gains were recorded in order to recover the relative energy level at each measurement location. The sweep responses were then post-processed by convolution with the inverse sweep to recover impulse responses for each location and source type.

In order to determine the frequency ranges that could be fit by the model, each impulse response was filtered into bandlimited responses using a standard octave-band filterbank, and then the desired metrics were calculated for each resulting frequency band. The T_{60} (a measure of temporal decay) was calculated by a linear fit to the Energy Decay Curve in dB for each frequency band and at each receiver location. This is sometimes known as Schroeder's reverse integration (11). Next, each receiver location was averaged to arrive at a single T_{60} for the entire hallway in each band. Finally, spatial decays (similar to strength of sound, or G) were calculated by a linear fit to the sum of energy for each band-limited response across each receiver location.

3.3 Simulation results

In order to verify the reformulation of the model in the FVTD formalism, we ran initial-value problem simulations for all combinations of modified absorption and scattering coefficients A and D. We discretely sampled each variable with enough resolution to confidently resolve the changes in the derived temporal and spatial measures mentioned above. The initial conditions were chosen to be a temporal Gaussian, centered at the same location as the source, in order to minimize spurious numerical oscillations. The sample rate was chosen to fit the stability criterion given in Bilbao et al. (8).

We then compared the results to the physically collected bandlimited impulse responses. We suggest that the subset of simulations whose measures fall within 10% of the *in situ* indices sufficiently model the desired acoustical characteristic. For each measure, those combinations of coefficients that fall within this range form a contiguous band in the search space which we will refer to as a "region of validity."

In the interest of space, we have presented the results of fitting both the monopole and dipole source on the same set of graphs in Figure 2, organized by octave band. Each color indicates a region of validity for the specified measure. Thus, we have highlighted the regions where the T_{60} was within 10%, the regions where the spatial decay was within 10%, and the regions where both spatial and temporal characteristics were within 10% of the measured data. Simulations that satisfy this final category indicate whether the model is ultimately successful for each source type and octave band, and is the intersection of the regions of validity for both T_{60} and spatial decay. In short, those combinations of coefficients are capable of successfully representing the stochastic reverberant field in the space for the given octave band.

For this set of measurements, it can be seen that the monopole case is valid above 125 Hz and the dipole case is valid above 500 Hz (since it could not be matched at 250 Hz). As expected, neither case is valid at 62.5 Hz. Interestingly, in every other band, many simulations were able to satisfy T_{60} for both the monopole and dipole cases, but in no case was a single simulation capable of modeling the spatial decay for both source types.

By verifying that the model reproduces the desired behavior in the impulsive case, the plausibility of steady-state and time varying cases are also confirmed. Therefore, by using the derived parameters, it is possible to approximate the playback of arbitrary source material in the space for combinations of monopole and dipole sources with the appropriate values of $\mathcal{P}$ and $\mathbf{Q}$. It is important to note that the findings of this exercise are not proof that the model is capable of representing an arbitrary space, nor can the derived coefficients be used for other spaces, but serve as validation of the model in this specific instance, and for spaces similar to the selected hallway in terms of geometric, absorptive, and diffusive characteristics.

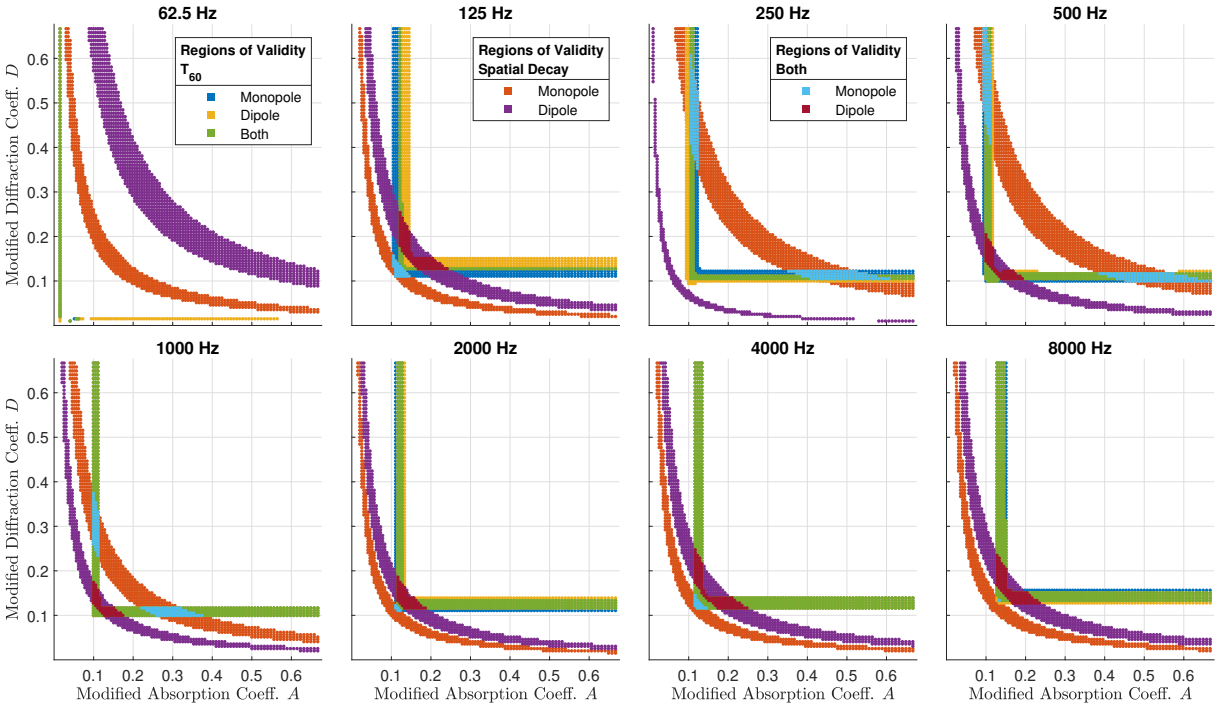

Figure 2 – Agreement between simulated and measured data for the monopole and dipole configurations in each frequency band.

4. CONCLUSIONS AND FUTURE WORK

In this work, we have reimplemented a one-dimensional solution for the wave-stress tensor based wave equation with the addition of arbitrary sources. We compared a range of simulations to physical measurements in a long hallway to show that it is possible to fit boundary conditions that model the acoustic behavior of both monopole and dipole sources.

For the most part, the point of this work has been to prepare a framework that could facilitate extension of the model to three dimensions. While the acoustic behavior immediately in front of boundaries is well represented by the model, wave propagation in free space is not, and therefore must be addressed with a separate method, such as ray tracing or a different FDTD/FVTD scheme. Doing so implies questions about the process of acoustic transfer between the two systems, as well as the necessary thickness of a translation layer between the two problem domains. Furthermore, it is important to consider how the model deals with more realistic cases where diffusion and absorption are not spread evenly over all surfaces.

As mentioned earlier, we omitted acoustic feedback in the source representation for simplicity. Future work may consider the relationship of the energy density to the calculation of radiation impedance in order to realize a more accurate coupled model without the anechoic approximation, including effects from multiple sources.

Though this reformulation of the one-dimensional problem gives similar results to the previous Telegrapher's equation solution, when the modified absorption and diffusion coefficients A and D are close in value, the simulations show oscillations that complicated the evaluation of temporal and spatial measures. Fortunately, it was possible to extract indices for all relevant simulations to facilitate the evaluation of the method, but this behavior should be examined more closely.

While all of the implementations until now have required calibration against real measurements in order to model a space, the model can not be used as a prediction tool in the general case until a systematized method for generating simulation parameters from spatial geometry and materials properties can be demonstrated. What such a method would require in order to capture the conditions for a given space (in either the homogeneous or non-homogeneous case) is as of yet unknown.

Finally, further experiments can be performed to evaluate the real-time performance of the method for predicting reverberation given arbitrary source material or more complex loudspeaker configurations given the significant computational advantages of the large element sizes and low sample rate requirements.

REFERENCES

1. Dujourdy H, Pialot B, Toulemonde T, and Polack JD. An Energetic Wave Equation for Modelling Diffuse Sound Fields – Application to Corridors. Acta Acustica united with Acustica 2017;103:480–91.
2. Dujourdy H, Pialot B, Toulemonde T, and Polack JD. An Energetic Wave Equation for Modelling Diffuse Sound Fields—Application to Open Offices. Wave Motion. Innovations in Wave Modelling II 2019;87:193–212.
3. Ollendorff F. Statistical Room-Acoustics as a Problem of Diffusion (A Proposal). Acta Acustica united with Acustica 1969;21:236–45.
4. Picaut J, Simon L, and Polack JD. A Mathematical Model of Diffuse Sound Field Based on a Diffusion Equation. Acta Acustica united with Acustica 1997;83:614–21.
5. Bilbao S. 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 2013;21:1524–33.
6. Morse PM and Feshbach H. Methods of Theoretical Physics. Mc Graw-Hill Book Company, 1953.
7. Morse PM and Ingard KU. Theoretical Acoustics. Princeton University Press, 1968. 949 pp.
8. Bilbao S, Hamilton B, Botts J, and Savioja L. Finite Volume Time Domain Room Acoustics Simulation under General Impedance Boundary Conditions. IEEE/ACM Transactions on Audio, Speech, and Language Processing 2016;24:161–73.
9. Farina A. Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. In: Audio Engineering Society Convention 108. Audio Engineering Society, 2000.
10. Farina A. Advancements in Impulse Response Measurements by Sine Sweeps. In: Audio Engineering Society Convention 122. Audio Engineering Society, 2007:21.
11. Schroeder MR. New Method of Measuring Reverberation Time. The Journal of the Acoustical Society of America 1965;37:409–12.

AURALIZATION OF A HYBRID SOUND FIELD USING AN ENERGY-STRESS TENSOR BASED MODEL

Aidan Meacham[1] **Roland Badeau**[2] **Jean-Dominique Polack**[1]

[1] Sorbonne Université, Institut d'Alembert, CNRS UMR 7190, Paris, France
[2] LTCI, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France
aidan.meacham@sorbonne-universite.fr

ABSTRACT

A hybrid approach to room impulse response synthesis and auralization is developed in the context of an energy-stress tensor based model of stochastic reverberation. This method for efficiently computing spatially varying energy envelopes has been demonstrated to represent the sound field in a sufficiently-diffusing 1-dimensional hallway above 250 Hz. To synthesize a realistic impulse response from the computed decay curves, the direct path, early reflections, and low frequency portion of the sound field must be calculated separately and then combined with the stochastic field to form a hybrid scheme. In this work, we propose one strategy for generating the stochastic field from the aforementioned energy envelopes and suggest the use of a typical pressure-velocity wave-based scheme and the image source method to generate the other necessary sound field components. Because of the efficiency of the energy-stress tensor based method and the reduced demands on the secondary simulation technique, such a hybridization presents a promising architecture for future real-time auralization in large spaces that may be difficult to model using only a single method.

1. INTRODUCTION

Combining multiple synthesis or simulation techniques in room acoustics is a common way to exploit the advantages of each method while also avoiding their pitfalls. For example, while wave-based simulation techniques accurately represent acoustic phenomena like modal behavior by construction, their computational complexity often grows with the cube of the highest frequency that one wishes to model, making them untenable for near-real-time audio tasks. Conversely, geometric approaches to acoustics are efficient at modeling the early part of sound propagation across the entire frequency band, but often require non-physical hacks to approximate more complex phenomena like edge diffraction. Finally, there are a number of methods focused on the so-called "late reverberation" that can generally be categorized as either energy-based, of which diffusion approaches are the most common, or tunable reverberation schemes, such as feedback delay networks. In both cases, it can be difficult to properly model the direct path and early reflections, and calibration of a particular space can be challenging, but the result is often an extremely efficient way to synthesize the tail part of sound propagation.

One alternative approach to late reverberation is an energy-stress tensor based method (or EST, for short) which is capable of modeling the sound field in sufficiently diffusing spaces [1]. It has been demonstrated that this theoretical approach is valid in primarily unidimensional spaces above the 250 Hz octave band, and is quite efficient due to the large cell sizes and low sample rates that are required to generate spatially and temporally accurate energy envelopes for the late reverberation. Up until this point, however, the integration of such an approach with an auralization method has not been realized. In order to determine how it may perform in a more general real-time acoustics tool, the hallway that has been characterized and replicated numerically in [2] makes for a useful test case.

It is clear that it would be advantageous to use each method in the region it is best suited for: wave-based for low frequencies, geometric for the early reflections, and the EST scheme for the late part. In fact, many modern room impulse response synthesis tools use approaches like this to accelerate computation times and are an area of active research in the academic community [3–6]. Often referred to as hybrid models, these approaches make a trade off between approximation, complexity, and computational speed, since it is often simpler to simulate a scenario entirely with one method, thus avoiding the need to intelligently pick cutoff frequencies or times, or to calibrate the different methodologies such that there is a reasonable expectation that they are modeling the same source or room characteristics. However, while recent advances in computational acoustics on graphics processing units (GPUs) has accelerated the speed at which wave-based or raytraced results can be generated, particularly with the use of clusters [7], on typical consumer hardware, the ability to synthesize full-band impulse responses in real time (and thus allowing for moving sources or receivers) is still the domain of hybrid approaches.

Therefore, we propose one possible method for rapidly generating full-band hybrid impulse responses in much the same way as in previous real-time acoustics engines, but introducing the computation of the late reverberation with the EST model to better take into account room geometry and motivate further development of the approach.

2. BACKGROUND

In this paper, we will mainly focus on the auralization of the late, high-frequency portion of the impulse response and the formation of the hybrid scheme, as the fundamentals of wave-based and geometric methods have been treated extensively in the literature. "Late reverberation" is the temporal region of a response with a sufficient density of reflections such that it can be represented as a stochastic process, rather than the distinct events of the direct sound and early reflections [8]. Similarly, at low frequencies, reverberation is typically dominated by behavior from a few isolated modes, but as the number of modes increases with frequency, they become sufficiently dense such that they can also be represented stochastically. The time and frequency cutoffs for these two criteria are commonly called the transition time and Schroeder's frequency, respectively. We will refer to this region of the time-frequency domain as the "stochastic reverberation," which is the main focus of the proposed hybrid model.

Because of this stochastic structure, this region has traditionally been modeled with reverberation algorithms that are significantly more efficient than wave-based or geometric schemes. One simple approach is directly filtering noise with the goal of approximating measured reverberation times in a frequency-dependent manner. More recent approaches in this vein use feedback delay networks (FDNs), a collection of delay lines that mix initial sound energy over time through feedback to produce natural sounding decay [9]. With a particular choice of absorption coefficients and delay line lengths, certain room characteristics can be reflected in the response of such a network, including (to some extent) the arrival times of early reflections or frequency dependent behavior. One downside of such schemes is that they often require complicated iterative tuning to match a specific chamber, since only some of their parameters are physically motivated.

One alternative approach that takes geometry into account more directly is the class of energy-based methods. Unlike most algorithms that model pressure and velocity waves, this category generally models the flow of sound energy in spaces with diffusing boundaries, and can be considered an extension of classical statistical reverberation. Because of their efficiency, they are often used to model very large spaces, especially when there are coupled volume effects present [10, 11]. One approach is the diffusion equation method (DEM), where sound energy diffuses throughout a space, much as in a heat diffusion problem [12–14]. Previously, models of this type have been auralized [15], however, little work has been done to unite them carefully with other methods in a hybrid model.

In contrast to DEM, the EST method supports a wave equation, which makes it possible to recover sound fields that build up in time depending on room geometry. Conveniently, though, since the end result is still a time series of energy envelopes, we can apply the energy envelopes to filtered noise in much the same fashion as previously studied with the DEM. In this paper, we implement both of the refinements suggested by Luizard et al. [15], taking advantage of multiple simulations for different frequency bands as well as the including the direct path and early reflections. The purpose of auralization is not only practical, in the sense that results from the EST method are bound to be used in applications where auralization is desired, but also pedagogical, as it provides a way to rapidly check that the theoretical framework as well as the numerical implementation are working as expected. Furthermore, it provides a framework for instant auditory feedback on future work on the model, particularly with respect to the development of a model for three dimensional spaces or checking that phenomena mathematically representable with the EST approach are actually audible.

In the interest of space, we will only present a brief overview of the theory here, and we will eschew the numerical treatment entirely, referring those interested readers to the publications in question [1, 16].

The main strategy in the EST model is the development of conservation laws for energy density E and sound intensity $\mathbf{J}$ [17]. In contrast, typically, wave-based acoustic simulations use an energy balance based on Gauss's theorem in terms of a pressure field p and a velocity field $\mathbf{v}$. Using these laws, one can derive a system of coupled equations relating the energy density, the sound intensity, and the wave-stress symmetric tensor $\underline{\underline{E}}$ [18]. In the case where one length dominates a space and the cross-sectional area is relatively constant (as is the case in a long hallway), dimensional reduction can eliminate the off-diagonal terms in the wave-stress tensor $\underline{\underline{E}}$ such that it resembles the scalar pressure field p, and a similar connection can be drawn between the vector velocity field $\mathbf{v}$ and $\mathbf{J}$. Using this formalism, one can follow the development of numerical models for common acoustic models with a few small modifications to the boundary conditions rather than requiring a domain-specific finite difference solution.

The main advantage of the EST method is that the propagation of diffuse energy allows for very large spatial discretization and low sample rates. Since we presume that the stochastic decay rate of the diffuse high frequency energy is not changing very rapidly in time and space, we can average over large cells and step through time at a relatively slow rate.

3. ROOM GEOMETRY AND MEASUREMENTS

For this work, we used the same hallway and impulse response dataset detailed in the frequency validity study [2]. A plan of the corridor can be seen in Figure 1.

The narrow portion of the hallway had an overall length of 45 m, with a width of 1.59 m, a suspended metal grating ceiling at 2.375 m, and a hard ceiling at 3.26 m, with a decrease to 2.8 m for metal support beams every 1.5 m along its length. The width increased to 2.39 m and the height decreased to 2.2 m in the recesses, some of which contained glass display cases or small pieces of furniture. The floor was linoleum and the walls were wooden panel and masonry. The heating, ventilation, and air conditioning systems were audible near vents but did not present an obstacle for the calculation of acoustical indices. All doors

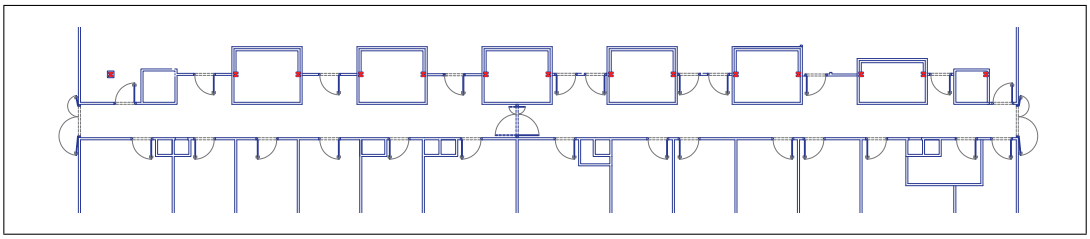

Figure 1. Floorplan for the corridor under consideration.

entering the hallway were closed, and the doors in the center of the hallway were fully open.

Impulse responses were collected using a *SoundField* ST250 microphone and an *Outline* GRS omnidirectional speaker with a *MOTU* Traveler sound card. The source was positioned 1 m away from the right end of the hall in Figure 1, 1.5 m above the ground, and centered between the two walls.

Beginning 1 m from the source, measurements were collected with the microphone's X-axis aligned along the length of the hallway. A spacing of 1 m was used out to 10 m, which corresponds exactly to the discretization distance in the numerical simulations. Then, recordings were made every 2 m until the end of the hallway for a total of 26 sampling locations. Recordings were made using the the swept sine method as implemented in the *Adobe Audition* plugin *Aurora* [19, 20]. The sweep length was 20 seconds at a sample rate of 44100 Hz, and the source level was adjusted digitally to maximize the signal-to-noise ratio without clipping as the microphone was moved further and further from the source. These gains were recorded in order to recover the true measured energy level for each measurement location. The sweep responses were then post-processed by convolution with the inverse sweep to recover impulse responses for each location. The collected impulse responses were time-aligned to the moment playback began so as to preserve the time-of-flight to the receiver, then trimmed to two seconds in length.

4. HYBRID MODEL

To facilitate direct comparison with the measured results, the synthesized impulse responses were also two seconds long at the same sample rate. For clarity, this does not imply that each of the methods used in the hybrid were run at the auralization sample rate, but in order to optimize for computation time, each was run at the lowest possible sample rate required for twice the number of samples in a second. These results were then upsampled (with the necessary filtering) in order to match the length of the measured impulse responses.

4.1 Stochastic Reverberation

The simulation of the stochastic reverberation for the hybrid method in this work comes from the finite volume time domain (FVTD) formulation of the EST model developed previously [16]. In short, the model uses a typical two-step scheme to solve for the energy density and sound intensity fields in a collection of cells. Since we are interested in modeling these fields in multiple frequency bands, we run the simulation multiple times independently with frequency-dependent modified absorption and scattering coefficients.

$$
\begin{aligned}
E_j^+ = &\left[E_j\left(1-\frac{cT}{V}\sum_{l=1}^{N_b}\gamma_{jl}S_l\frac{A}{8}\right)\right.\\
&\left.+\frac{cT}{V}\mathcal{P}_j-\frac{cT}{V}\sum_{k=1}^{N}\beta_{jk}SJ_{jk}\right]\\
&\Bigg/\left[1+\frac{cT}{V}\sum_{l=1}^{N_b}\gamma_{jl}S_l\frac{A}{8}\right],\\
J_{jk} = &\left[J_{jk}^-\left(1-\sum_{l=1}^{N_b}\zeta_{jkl}\frac{S_l}{V}cT\frac{D}{8}\right)\right.\\
&\left.+cTQ_{jk}+\frac{cT}{h}(E_j-E_k)\right]\\
&\Bigg/\left[1+\sum_{l=1}^{N_b}\zeta_{jkl}\frac{S_l}{V}cT\frac{D}{8}\right].
\end{aligned}
\tag{1}
$$

The update equation for this scheme is given in Eqn. (1), where E and J are the discretized average energy densities and sound intensities, j, k, and l are indices for a cell, another cell, and a boundary respectively, $\pm$ exponents indicate the next and previous discrete time steps for time series variables, c is the speed of sound, T is the chosen temporal sampling step, h is the inter-cell distance (with respect to centroids), S and S_l are the areas of boundaries between a cell and a neighboring cell or a boundary, V is the volume of a cell, A and D are the so-called modified absorption and scattering coefficients, γ, β, and ζ are indicator functions that select for relationships between particular cells and boundaries, and $\mathcal{P}$ and Q are source terms used to implement time-varying source behavior. In this case, h, S, S_l, V, A, and D are constant, but in general, could also be indexed on a cell-by-cell or boundary-by-boundary basis. Additionally, we omit the frequency index on A, D, E, J, Q, or $\mathcal{P}$ for clarity.

In this case, since we are interested in assembling an impulse response rather than processing audio directly, it is sufficient to observe the response of the system to an initial stimulus. It is equivalent to do so using the source terms directly, or to set the source terms to zero and use the initial conditions of the energy density field to excite a response in the space. As before, we use measurements

made in the real hallway to inform the parameters used for the EST model for each frequency band. In order to match the measurements during the model calibration, the virtual source is also placed 1 m from the end of the hallway. Then, a brute force tuning search over the model parameter space is performed to see which combinations of modified absorption and scattering coefficients give similar acoustical indices in each frequency band as the real recordings. The two target indices were spatial decays, or the slope of the decrease in sound energy as a function of distance from the source, and temporal decays, essentially T_{60}, both of which must fall within 10% (or better) of the observed value. The corresponding parameters in each frequency band are noted in Tab. 1 and were subsequently used in the hybrid model. More detail on the parameter fitting procedure can be found in [2].

After the model has been tuned for a given frequency band, it is straightforward to change the initial conditions in order to move the source location or add other sources. Given the large cell sizes and low sample rate, energy density envelopes can be rapidly generated when this occurs. Finally, the results are resampled to the audio sample rate that the different methods will be composited at.

In order to auralize the energy density envelopes, white noise of the length desired at the final sample rate is generated and split into octave bands using a filterbank. Then, the envelope associated with each octave band is multiplied with the bandlimited noise to create a signal with the same time of arrival and decay rate as in the measured hall. Finally, all of the resulting bandlimited responses are added in the time domain to create a single stochastic reverberation signal.

In the case where the source is presumed to be stationary, the results at every cell in the hallway can be cached and interpolated given a particular receiver location. It is only when the source distribution changes that the simulation and computation needs to be brought up to date.

4.2 Low-Frequency Reverberation

Modal behavior and diffraction that cannot be modeled by the EST method is handled by a FVTD simulation up through the octave band centered at 500 Hz [21]. Since we don't need to perform any boundary adjustments, the method is equivalent to a typical finite difference approach. In order to ensure that the dispersion error is minimized, the simulation sample rate f_{FVTD} (and by convention the element size) is chosen such that the maximum desired frequency occurs at $0.15 \times f_{\text{FVTD}}$, or about 7x oversampling [22]. We assigned boundary conditions in the model using the example materials given in the original work that best corresponded to each hallway surface. Since in this case the highest desired frequency is about 700 Hz, the sample rate was set at 4900 Hz. The resulting pressure field is then low-passed, spatially downsampled as needed to reduce the total memory footprint, and temporally resampled to match the final auralization sample rate.

Similarly to the stochastic reverberation, the simulation for this contribution only needs to be rerun in the case where the source distribution changes, but otherwise, can be cached in order to interpolate for a particular receiver location.

4.3 Direct Path and Early Reflections

The earliest arriving sound energy is computed with the image source method directly at the final sample rate [23]. There is some freedom to choose the maximum order of sources that will be included in the computation, constrained by the requirement of sufficiently representing the early reflections on one hand and without expending computation time unnecessarily on the multitude of reflections that will be modeled stochastically on the other.

For this method, the contribution must be recomputed any time the source or receiver position changes.

5. CALIBRATION

5.1 Hybrid Model

To calibrate the level between the separate simulations, we use an energy equalization in a frequency band where crossover occurs, using a distance of 1 m from the source as reference [24, 25]. Since all three methods have a valid portion of their frequency response in the octave band centered at 500 Hz, we can set an arbitrary level and find a gain for each simulation type that matches its energy to the target energy in that region. Then, when combining the methods by superposition in the time domain, the FVTD part is used directly as it is the most accurate, and the other methods are high-passed in order to avoid duplication of energy in the band of overlap.

Hybrid strategies are often delineated by frequency, and the point where two methods meet is chosen in order to optimize for speed of simulation at a given quality trade off. For other hybrid models, the topic of crossover frequency is well studied [26, 27]. In this case, however, because the EST method is not applicable at all below the Schroeder frequency, the cutoff frequency between the FVTD portion and the higher frequency models is determined in advance. Fortunately, it is not very high, and therefore the low-frequency FVTD portion can be simulated without too much difficulty. With a GPU, real-time simulation of even small concert halls has long been feasible at the frequency ranges under discussion [28]. Furthermore, given the speed at which the image source and stochastic reverberation portions can be computed, we believe that an optimized version of this approach could approach real-time performance with proper tuning.

5.2 Measurements & Hybrid

To calibrate levels between the measured and simulated impulse responses, a reference distance of 1 m is chosen, and the energy of the two signals are matched at that point. The resulting gain is then applied regardless of where the measurement or simulation is sampled so as to preserve the relative decay as a function of distance from the source.

Frequency [Hz]	250	500	1000	2000	4000	8000
T_{60} [S]	0.38	0.39	0.43	0.35	0.33	0.30
Spatial Decay [dB/m]	-0.88	-0.94	-0.66	-0.54	-0.66	-0.84
Mod. Abs. Coef. A	0.0516	0.0512	0.0473	0.0551	0.0577	0.0671
Mod. Sca. Coef. D	0.5359	0.5920	0.3818	0.2483	0.3210	0.4114

Table 1. Measured characteristics and resulting simulation parameters.

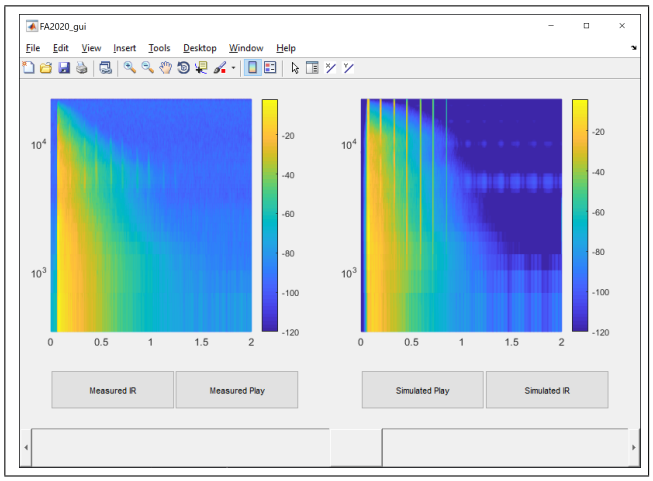

Figure 2. The interface used to compare measured and simulated impulse responses.

6. EVALUATION

The overall behavior of the system was inspected using a graphical interface that allowed visual examination of spectrograms (taken with the same parameters and with the same color scale for spectral power), direct auralization of the resulting impulse responses, as well as convolution of the IRs with source material. A slider ranging from 1 m to 42 m (the maximum measured observation distance from the source) was included, and moving the slider triggered an update of the receiver position, causing any relevant interpolation or re-simulation. With this scheme, it is straightforward to assess the changes in time of arrival, overall spectral envelope, and informally, the perceptual differences between the simulated and measured auditory presentations throughout the hallway as a function of distance.

In Fig. 2, the spectrogram of the measured impulse response is presented on the left, and the spectrogram of the simulated impulse response is presented on the right, corresponding to a receiver distance of around 23 m from the source.

7. RESULTS

Brief informal listening sessions revealed some of the advantages and shortcomings of the hybrid model.

The overall decay times and absolute levels at any given position are fairly similar between the simulated and measured impulse responses, and the arrival times of the most notable reflections are accurately rendered. In the cases where there are differences in decay times, they appear to be mostly due to the stochastic reverberation model's octave-band calibration using the temporal and spatial information from the measurements, which in some cases, was not sufficiently narrow. The FVTD and ISM portions also do not reproduce the measured decay as a function of distance perfectly, resulting in some spatial variation in the prominence of particular portions of the hybrid as a result of the constant level calibration gain. This is likely a result of the lack of tuning for the simulation parameters in both cases, as well as not taking viscothermal losses into account for the high frequency portion of the ISM, which is particularly prominent in the example simulation spectrogram in Fig. 2 above the 8 kHz octave band.

As a result, near the source, the stochastic portion is somewhat overrepresented, resulting in impulse responses that sound too diffuse relative to the strong direct path and early reflections present in the measurements. Furthermore, along the entire length of the hallway, the simulated impulse responses seem to have a slight tonal coloration shifted toward the higher frequencies, possibly resulting from the lack of tuning for the stochastic model above the band centered at 8 kHz.

One of the most apparent phenomena in the recorded impulse responses is the presence of a strong echo near the far end of the hallway from the source as a result of proximity to the reflecting surface. In the simulated impulse responses, this effect appears courtesy of the ISM portion. While the overall envelope of the EST method reflects the corridor-length delay behavior, the fine-structure detail that gives that area of the hallway its characteristic sound is decidedly not a part of the diffuse field, and therefore it cannot be expected that it would be precisely reproduced by the large elements and low sample rate. The second and third recurrences in time of this phenomenon are apparent in the overall effect of the measured impulse responses, if not directly audible as a flutter depending on position. Using a low order for the ISM in an attempt to avoid a proliferation of image sources eliminates these salient features from the simulated version entirely, so it is important to consider the tradeoff between computational efficiency and the relevance of particular specular phenomena.

Overall, while the speed and general characteristics of the simulation were promising, additional improvements must be made in order to provide a satisfactory facsimile of the original hallway.

8. FUTURE WORK

When the source distribution is stationary, the results of the stochastic reverberation contribution can be evaluated everywhere and interpolated on the fly to improve speed. However, if the source distribution changes frequently, it is most efficient to only upsample and do the computations at the points nearest the receiver. Thus, to balance between the two behaviors, it should be possible to begin the computation with only the necessary points, and if the source distribution ceases to change, to then fill in the rest of the possible receiver positions as able.

While the origin of the tonal differences between the simulated and measured impulse responses is unclear, the use of a third-octave filterbank in the design of the stochastic response as well as changing the crossover frequency may help in diagnosing where the issue is located. Additionally, moving to a 3-dimensional model may help model the differences between the narrow portions and alcoves along the hallway, which may have an effect on the overall frequency response at a given position.

In terms of salient spatial features that are not captured by the stochastic reverberation (such as the echo mentioned earlier), it is difficult to predict in the general case how to efficiently generate these effects under the current simulation strategy. In the specific case of the hallway, it is of course possible to generate a higher order of longitudinal image sources, which would give a reasonable approximation of the behavior in the full-band response; however, that is not the case in an arbitrary space where it is less obvious which image sources may be perceptually relevant without calculating higher orders in the first place. While the ISM approach in this case was relatively simple, the wide range of improvements that have been proposed in the literature are fair game for implementation, including frequency-dependent boundary conditions or viscothermal losses.

Recent advances in late reverberation techniques account for directionality in the simulated impulse responses [29]. Simulation of anisotropy is also possible with the EST method, and may present a useful framework for considering non-homogeneous diffusion coefficients as well.

Finally, the prediction of boundary conditions for the EST from material and geometry measurements remains an important avenue of research to eliminate the calibration step and enable simulation of unseen spaces.

9. CONCLUSIONS

In this paper, we have detailed ongoing work on a hybrid auralization scheme utilizing the EST based model for the stochastic reverberation combined with ISM and FVTD models to recreate and listen to the sound field in a long hallway. Using filtered noise and temporal envelopes on an octave-band by octave-band basis, we formed a portion of the impulse response valid above Schroeder's frequency and after the transition time. The gains used for the each method were calibrated using an octave band where all simulations were valid, then for the summation, the ISM and EST method were high-passed above the highest frequency present in the FVTD model, which was used as-is.

An interface to explore the acoustic space and compare the simulated and measured impulse responses was developed and used to derive some initial impressions of the auralization scheme. Informal results from this hybridization were promising in terms of computational speed and reproduction of the calibrated acoustical parameters, however, additional work remains in order to better match other salient spatial cues, and theoretical challenges remain regarding the temporal transition between the high-frequency methods. Nonetheless, the structure presented is an encouraging first step toward geometry-aware stochastic reverberation simulation strategies for real-time applications.

10. REFERENCES

[1] H. Dujourdy, B. Pialot, T. Toulemonde, and J.-D. Polack, "An Energetic Wave Equation for Modelling Diffuse Sound Fields – Application to Corridors," *Acta Acustica united with Acustica*, vol. 103, pp. 480–491, May 2017.

[2] A. Meacham, R. Badeau, and J.-D. Polack, "Lower Bound on Frequency Validity of Energy-Stress Tensor Based Diffuse Sound Field Model," in *Proceedings of the 23rd International Congress on Acoustics*, (Aachen), Sept. 2019.

[3] D. Murphy, S. Shelley, M. Beeson, A. Moore, and A. Southern, "Hybrid room impulse response synthesis in digital waveguide mesh based room acoustics simulations," in *DAFx*, 2008.

[4] D. Schröder and M. Vorländer, "RAVEN: A real-time framework for the auralization of interactive virtual environments," in *Proceedings of Forum Acusticum*, pp. 1541–1546, Jan. 2011.

[5] S. Oxnard and D. Murphy, "Room impulse response synthesis based on a 2D multi-plane FDTD hybrid acoustic model," in *2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, NY, USA), pp. 1–4, IEEE, Oct. 2013.

[6] D. Poirier-Quinot, B. Katz, and M. Noisternig, "EVERTims: Open source framework for real-time auralization in architectural acoustics and virtual reality," in *DAFx-17*, (Edinburgh), Sept. 2017.

[7] H. Lai and B. Hamilton, "Computer Modeling of Barrel-Vaulted Sanctuary Exhibiting Flutter Echo with Comparison to Measurements," *Acoustics*, vol. 2, pp. 87–109, Mar. 2020.

[8] R. Badeau, "Common mathematical framework for stochastic reverberation models," *The Journal of the Acoustical Society of America*, vol. 145, pp. 2733–2745, Apr. 2019.

[9] J.-M. Jot and A. Chaigne, "Digital Delay Networks for Designing Artificial Reverberators," in *Audio Engineering Society Convention 90*, Audio Engineering Society, Feb. 1991.

[10] Z. Sü Gül, N. Xiang, and M. Çalışkan, "Investigations on sound energy decays and flows in a monumental mosque," *The Journal of the Acoustical Society of America*, vol. 140, pp. 344–355, July 2016.

[11] Z. Sü Gül, E. Odabaş, N. Xiang, and M. Çalışkan, "Diffusion equation modeling for sound energy flow analysis in multi domain structures," *The Journal of the Acoustical Society of America*, vol. 145, pp. 2703–2717, Apr. 2019.

[12] A. Billon, V. Valeau, A. Sakout, and J. Picaut, "On the use of a diffusion model for acoustically coupled rooms," *The Journal of the Acoustical Society of America*, vol. 120, pp. 2043–2054, Oct. 2006.

[13] Y. Jing and N. Xiang, "A modified diffusion equation for room-acoustic predication," *The Journal of the Acoustical Society of America*, vol. 121, pp. 3284–3287, June 2007.

[14] N. Xiang, Y. Jing, and A. C. Bockman, "Investigation of acoustically coupled enclosures using a diffusion-equation model," *The Journal of the Acoustical Society of America*, vol. 126, pp. 1187–1198, Sept. 2009.

[15] P. Luizard, J.-D. Polack, and B. F. G. Katz, "Auralization of coupled spaces based on a diffusion equation model," in *Proc. Sound and Music Computing Conference (SMAC/SMC)*, (Stockholm, Sweden), pp. 616–621, 2013.

[16] A. Meacham, R. Badeau, and J.-D. Polack, "Implementation of Sources in an Energy-Stress Tensor Based Diffuse Sound Field Model," in *Proceedings of the International Symposium on Room Acoustics*, (Amsterdam), Sept. 2019.

[17] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*. Mc Graw-Hill Book Company, 1953.

[18] P. M. Morse and K. U. Ingard, *Theoretical Acoustics*. Princeton University Press, 1968.

[19] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique," in *Audio Engineering Society Convention 108*, Audio Engineering Society, Feb. 2000.

[20] A. Farina, "Advancements in Impulse Response Measurements by Sine Sweeps," in *Audio Engineering Society Convention 122*, Audio Engineering Society, May 2007.

[21] S. Bilbao, B. Hamilton, J. Botts, and L. Savioja, "Finite Volume Time Domain Room Acoustics Simulation under General Impedance Boundary Conditions," *IEEEACM Trans. Audio Speech Lang. Process.*, vol. 24, pp. 161–173, Jan. 2016.

[22] A. Southern, T. Lokki, L. Savioja, and D. Murphy, "The Perceptual Effects of Dispersion Error on Room Acoustic Model Auralization," in *Proceedings of Forum Acusticum*, (Aalborg), pp. 1553–1558, Jan. 2011.

[23] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, pp. 943–950, Apr. 1979.

[24] A. Southern, S. Siltanen, and L. Savioja, "Spatial Room Impulse Responses with a Hybrid Modeling Method," in *Audio Engineering Society Convention 130*, Audio Engineering Society, May 2011.

[25] A. Southern, S. Siltanen, D. T. Murphy, and L. Savioja, "Room Impulse Response Synthesis and Validation Using a Hybrid Acoustic Model," *IEEE Trans. Audio Speech Lang. Process.*, vol. 21, pp. 1940–1952, Sept. 2013.

[26] S. Oxnard, *Efficient Hybrid Virtual Room Acoustic Modelling*. PhD thesis, University of York, Sept. 2016.

[27] C. Rougier, *Influence of Crossover Frequency on a Hybrid Acoustic Model for Room Impulse Response Synthesis*. Masters thesis, University of York, 2018.

[28] L. Savioja, "Real-time 3D finite-difference time-domain simulation of low- and mid-frequency room acoustics," in *DAFx*, 2010.

[29] B. Alary, A. Politis, S. Schlecht, and V. Välimäki, "Directional Feedback Delay Network," *JAES*, vol. 67, pp. 752–762, Oct. 2019.

Riemannian space tessellation with polyhedral room images

Jean-Dominique Polack[a], Aidan Meacham[b], Roland Badeau[c], Jean-Christophe Valière[d]

a. Institut Jean Le Rond D'Alembert (IJLRA), Sorbonne Université, CNRS UMR 7190, France
jean-dominique.polack@sorbonne-universite.fr
b. Institut Jean Le Rond D'Alembert (IJLRA), Sorbonne Université, CNRS UMR 7190, France
aidan.meacham@sorbonne-universite.fr
c. LTCI, Télécom Paris, Institut Polytechnique de Paris, France
roland.badeau@telecom-paris.fr
d. Institut PPRIME, Université de Poitiers, CNRS UPR3346, France jean.christophe.valiere@univ-poitiers.fr

September 26, 2021

Abstract

We show that Riemannian geometry is the natural setting for developing polyhedral rooms of arbitrary shapes into their image rooms, and therefore counting the image sources. This new setting makes it also possible to account for scattering on particular edges, called hinges, characterized by negative deficit dihedral angles created by reflections on the adjacent faces. Using energy conservation, we show that sound rays are then deviated by the hinges, depending on their frequencies and the distances they pass by.

Keywords: Riemannian geometry, polyhedral rooms, scattering, stress-energy tensor, room acoustic modelling.

1 Introduction

Computing the number of image sources for a rectangular enclosure is an easy task that acousticians routinely carry out [CM78]. Indeed, as all the images of the room tesselate the Euclidean space, the computation simply amounts to dividing the volume of a sphere of radius ct, where c is the speed of sound and t the time elapsed since the source emited, by the volume of the original room, as each image room, or cell, only contains one image source. Thus one obtains a number of image sources that increases with the square of the time elapsed since the source emitted and is inversely proportional to the volume of the room, and most acousticians consider that this approximation is also

valid for rooms of arbitrary shapes. As a consequence, most simulation codes routinely use the mean number of image sources for rectangular rooms when they fix the complexity of the algorithms that compute the late reverberation, independently of the actual room shape [NE93].

Few authors have questionned the validity of this approximation for rooms of arbitrary shapes. [Pol92] has argued that this number should increase exponentially for mixing rooms, as a consequence of the conservation of the phase space measure. Indeed, in mixing rooms, any small volume of the phase space, for example the initial volume around the source, exponentially expands in at least one phase dimension as it propagates with time, and exponentially decreases in at least one dimension, so that any volume decays into exponentially thin stripes. Any elementary cell of the phase space thus intersects exponentially many stripes, corresponding to exponentially many image sources.

Polyhedral rooms, on the other hand, are not mixing rooms since two neighbouring rays in the phase space almost certainly separate linearly with time, and not exponentially. Thus, the preceding argument cannot be used to evaluate the number of image sources. The present paper therefore addresses the non trivial issue of computing the number of images sources for arbitrary polyhedral enclosures. It first examines the reflection and scattering on non rectangular dihedral angles and shows that successive reflections on the adjacent faces always add up to a total angle in excess of 2π, so that taking into account reflections on all the faces of a polyhedron leads to the tessellation of a Riemaniann space with negative curvature where edges take the role of the "hinges" of Regge's discretization of Riemannian spaces [Reg61]. Afterwards, it moves to computing the number of image sources, and proposes a computational scheme based on the number of faces, edges and vertices, that is, the isometry group of the space. Then, it addresses the question of the visible image sources that are linked with the receiver by rays that do not cross any edge of image rooms, and derives from it the notion of visible horizon. Finally, scattering is taken into account by introducing the wave nature of the sound field in a room, and the conservation of its energy and intensity.

2 Scattering on non rectangular dihedral angles

The proper setting for computing scattering on non rectangular dihedral angles is Riemannian geometry. We therefore first expose the principles of this geometry and the properties of its metric tensor.

2.1 Riemannian geometry

We consider an n-dimensional space with its positive-definite metric tensor g_{ij} and the volume element $dV = \sqrt{g}dx^1 \dots dx^n$, where $g = \det g_{ij}$ is positive [Lin 5]. The infinitesimal distance element is given by:

$$ds^2 = g_{ij}dx^i dx^j \tag{1}$$

and we note g^{ij} the inverse matrix of g_{ij}. ∇_i is the covariant derivation with respect to x^i, which differs from the usual partial derivation ∂_i in a way that depends on the tensor rank. For example, for a function Φ:

$$\nabla_j \Phi = \partial_j \Phi = \Phi_j$$

but $\nabla_j \Phi_i = \partial_j \Phi_i - \Gamma^k_{ji} \Phi_k$ and $\nabla_j X^i = \partial_j X^i + \Gamma^i_{jk} X^k$, where X^i are the (contravariant) components of vector X and Γ^k_{ji} the Christoffel symbols linked to the derivatives of the elements of the metric tensor g_{ij}:

$$\Gamma^k_{ji} = \frac{1}{2} g^{kl} \left(\partial_j g_{il} + \partial_i g_{lj} - \partial_l g_{ji} \right) \tag{2}$$

Note that, unlike ordinary differentiations, covariant derivations do not commute. Their commutators is given by the curvature of the time-space [Lin 5]:

$$\left(\nabla_i \nabla_j - \nabla_j \nabla_i \right) X^k = R^k_{mij} X^m \tag{3}$$

where the R^k_{mij} are the elements of the Riemann tensor. However, for a function, $\nabla_i \partial_j \Phi = \nabla_j \partial_i \Phi$; and by construction, all covariant derivatives of the elements of the metric tensor are null. In other words, the contravariant derivation ∇^i is defined by:

$$\nabla^i = g^{ij} \nabla_j = \nabla_j g^{ij}$$

One calls *vectors* tensors with one upper index, such as X^i; and *covectors* tensors with one lower index, such as Φ_i.

The Ricci curvature tensor can be computed from the derivative of the Christoffel symbols, by contraction of the Riemann curvature tensor on two indices. One obtains successively:

- the Riemann tensor: $R^l_{kij} = \partial_i \Gamma^l_{jk} - \partial_j \Gamma^l_{ik} + \Gamma^m_{jk} \Gamma^l_{im} - \Gamma^m_{ik} \Gamma^l_{jm}$
- the Ricci tensor: $R_{ij} = R^l_{ilj}$
- the scalar curvature: $R = g^{ij} R_{ij}$; and the local curvature is obtained by integrating half the scalar curvature over a small space element, that is, $\frac{1}{2} \int R \sqrt{g} dx^1 \dots dx^n$

In this Section and the following one, we shall only consider 2- and 3-dimensional Riemannian spaces, which are embedded in 3- or 4-dimensional Euclidean spaces in order to derive the metric tensor.

2.2 2-dimensional scattering on obtuse angles

We first consider the obtuse angle β of Fig. 1 in a 2-dimensional Euclidean space. A sound source S (black star) emitting inside the angle emits rays in all directions. Some rays impinge on the left arm Ox of the angle and are reflected (red arrows). As the position of impact moves clockwise toward the apex, the reflected ray gradually moves upwards and eventually hits the right arm on which it is reflected once more. Finally, the ray impacts the apex of the angle (black upwards broken-line arrow).

In a similar fashion, some rays will impinge on the right arm Oy and be reflected (green arrows). As the position of impact moves anticlockwise toward the apex, the reflected ray gradually moves upwards and eventually hits the left arm on which it is reflected once more. Despite the continuity of the impinging rays around the apex, there is no continuity of the reflected rays, and this creates scattering.

In order to visualise the scattering, one needs to consider the images of the sources by reflection on the two arms of the angle. Let's call S'_l the image

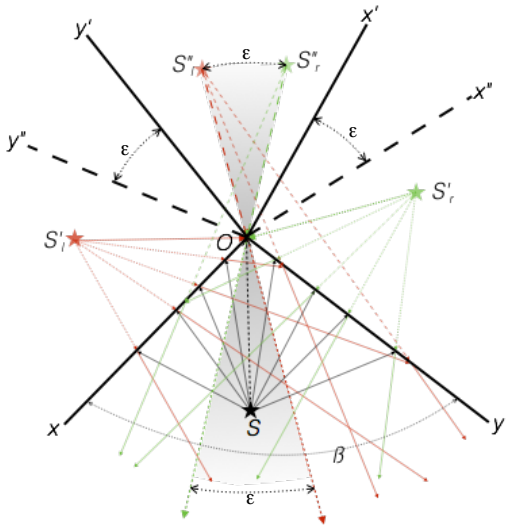

Figure 1: Reflection and scattering of sound rays on obtuse angle β. ε is the excess angle, since second order image angles overlap by ε.

of the source on the left arm (red star), and S'_r the image on the right arm (green star); S''_l the left-most second order image (dim red star), and S''_r the right-most second order image (dim green star). Reflected rays on the arms of the angle are first emitted from the first order image sources S'_l and S'_r, then from the second order sources S''_l and S''_r when the reflected rays hit the opposite arm, until the rays emitted from the first-order image sources reach the apex (red and green dotted-line arrows). In that position, the rays emitted from the second order image sources do not coincide in direction, since they are emitted from two different image sources at angle ε with respect to the apex. Scattering therefore comes in to *fill the gap* between these two directions, and in fact beyond them. In other words, one must consider a continuum of image sources along the sector between S''_l and S''_r (grey sector), that is, one must rotate the second order image angle by ε from position $y''0x'$ to position $y'0x''$.

In fact, when rays rotate *clockwise* around the original source S, the reflected rays rotate *anticlockwise* around the first order source S''_l, and *clockwise* around the second order source S''_r. So, when the secondary source moves *clockwise* from S''_l to S''_r on Fig. 1, the diffracted rays rotate *clockwise* around the apex, thus filling the grey sector with continuity of rotations at its boundaries.

In order to make the second order image angles coincide, one needs to embed Fig. 1 in a 3-dimensional space as in Fig. 2. We call x_1, y_1, and z_1 the three Cartesian coordinates, and keep notations Ox and Oy for the arms of the original flat angular sector, which is now slanted so that angle β exactly projects on a right angle. As a consequence, Ox is elevated by angle α above Ox_1 and Oy is lowered by angle α below Oy_1; similarly, $O(-x)$ is elevated by angle α above $O(-x_1)$ and $O(-y)$ is lowered by angle α below $O(-y_1)$, so that the apex angle remains equal to β. In such a way, we obtain a *locally flat* space where rays are free to cross the borders between subsequent angular sectors without changing their direction. Further, the embedding makes it possible to define the metric

tensor of this space.

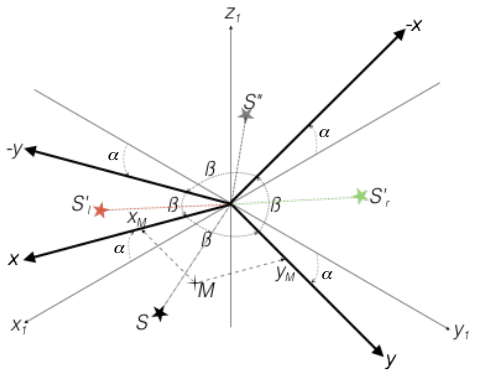

Figure 2: Embedding scattering on obtuse angle in 3-dimensional space. Angles α are vertical and measure deviations of angle arms from horizontal plane. Note that the two second order image sources coalesce in one single image source.

When the angle α is not null, that is, when $\beta > \frac{\pi}{2}$, the position of any point M in the system is given by its coordinates x and y, that are no longer orthogonal. In order to compute the corresponding infinitesimal distance element, one must project coordinates x and y on the Ox_1y_1 plane, then compute the vertical coordinate with equation:

$$z_1 = \tan\alpha(\mid x_1 \mid - \mid y_1 \mid)$$

valid for the four sectors of the system. With the convention $\mathrm{sgn}(0) = 0$, one obtains everywhere, even on the arms of the angles:

$$\begin{aligned} ds^2 &= dx_1^2 + dy_1^2 + dz_1^2 \\ &= (1+\tan^2\alpha)dx_1^2 + (1+\tan^2\alpha)dy_1^2 - 2\mathrm{sgn}(x)\mathrm{sgn}(y)\tan^2\alpha dx_1 dy_1 \end{aligned}$$

with $dx = \sqrt{1+\tan^2\alpha}dx_1$ and $dy = \sqrt{1+\tan^2\alpha}dy_1$, that is:

$$\begin{aligned} ds^2 &= dx^2 + dy^2 - 2\mathrm{sgn}(x)\mathrm{sgn}(y)\frac{\tan^2\alpha}{1+\tan^2\alpha}dxdy \\ &= dx^2 + dy^2 - 2\mathrm{sgn}(x)\mathrm{sgn}(y)\sin^2\alpha\, dxdy \end{aligned}$$

Simple projection on the plane Ox_1y_1 shows that $-\sin^2\alpha = \cos\beta$. Note that $\cos\beta < 0$ since $\frac{\pi}{2} < \beta < \pi$ for obtuse angles. The last equation can therefore be written as:

$$ds^2 = dx^2 + dy^2 + 2\mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta\, dxdy \tag{4}$$

Using the embedding in the 3-dimensional space, distance elements ds are obviously continuous along a line that crosses one of the arms, implying that dx and dy are modified when the line crosses the arms. From the expression of ds^2, it is easy to deduce the metric tensor, its inverse, and its determinant:

$$g_{ij} = \begin{pmatrix} 1 & \mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta \\ \mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta & 1 \end{pmatrix} \tag{5}$$

$$g^{ij} = \frac{1}{\sin^2\beta}\begin{pmatrix} 1 & -\mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta \\ -\mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta & 1 \end{pmatrix}$$
$$g = \sin^2\beta \tag{6}$$

Note that g_{ij} and g^{ij} reduce to the identity matrix on the arms ($x = 0$ or $y = 0$) because of the convention $\mathrm{sgn}(0) = 0$, with $g = 1$.

Successive derivations lead to:

- the Christoffel symbols are all equal to 0, but for:

$$\begin{aligned} \Gamma^y_{xx} &= g^{yy}\partial_x g_{xy} = 2\delta(x)\mathrm{sgn}(y)\cos\beta \\ \Gamma^x_{yy} &= g^{xx}\partial_y g_{xy} = 2\mathrm{sgn}(x)\delta(y)\cos\beta \end{aligned} \tag{7}$$

  where δ is the Dirac distribution, and where we have used the form for g^{xy} valid on the arms because $\delta(x)$, resp. $\delta(y)$, is null everywhere except on the arms $x = 0$, resp. $y = 0$;

- the Ricci tensor has elements:

$$\begin{aligned} R_{xx} &= R_{yy} = 4\delta(x)\delta(y)\cos\beta \\ R_{xy} &= R_{yx} = 0 \end{aligned} \tag{8}$$

- the scalar curvature and the local curvature are resp. $R = 8\delta(x)\delta(y)\cos\beta < 0$ and $\frac{1}{2}\int R\sqrt{g}dxdy$.

It is easy to see that the curvature is null everywhere but at the apex O where it is equal to $4\cos\beta$, which tends toward $4(\frac{\pi}{2} - \beta)$ for small values of $\frac{\pi}{2} - \beta$. Indeed, Regge [Reg61] has shown that the curvature is equal to the total deficit angle at the apex $-\varepsilon = 2\pi - 4\beta$.

2.3 3-dimensional scattering on obtuse dihedral angles

For a 3-dimensional obtuse dihedral angle, Fig 1 represents a projection of the rays on a plane perpendicular to the apical edge. But no equivalent of Fig 2 can be drawn, as the embedding takes place in a 4-dimensional space.

Introducing a new coordinate z, perpendicular to the two coordinates x and y of the Riemannian plane of Fig 2, the infinitesimal distance element (4) must now be completed into:

$$ds^2 = dx^2 + dy^2 + dz^2 + 2\mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta\, dxdy \tag{9}$$

with the metric tensor and its inverse now given by

$$g_{ij} = \begin{pmatrix} 1 & \mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta & 0 \\ \mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{10}$$

$$g^{ij} = \frac{1}{\sin^2\beta}\begin{pmatrix} 1 & -\mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta & 0 \\ -\mathrm{sgn}(x)\mathrm{sgn}(y)\cos\beta & 1 & 0 \\ 0 & 0 & \sin^2\beta \end{pmatrix}$$

with determinant g still given by (6); and g_{ij} and g^{ij} reduce to identity matrices on the planes $x = 0$ and $y = 0$, which we still call arms, with $g = 1$. All

Christoffel symbols are equal to 0 but Γ^y_{xx} and Γ^x_{yy} still given by eq. (7), and all elements of the Ricci tensor are 0 except $R_{xx} = R_{yy}$ still given by eq. (8). The scalar curvature therefore remains $R = 8\delta(x)\delta(y)\cos\beta < 0$, but the local curvature, obtained by integrating half the scalar curvature over a small space element, is now equal to $\frac{1}{2}\int R\sqrt{g}dxdydz$. It remains null everywhere - flat space - except on the apical edge. In the limit where β tends toward a right angle, the curvature around the apical edge is equal to $(2\pi - 4\beta)\ell$, where ℓ is the length of the apical edge, called *hinge* by [Reg61]. Note that hinges are subspaces of co-dimension 2 where excess angle is non null.

2.4 Scattering on reflex angles

The case of a reflex angle, both in 2 and 3 dimensions (see Fig. 3) can be handled in a similar fashion. But in this case, we only need to consider the first order images of the angle to obtain excess angles at the apex, and rays are never reflected twice.

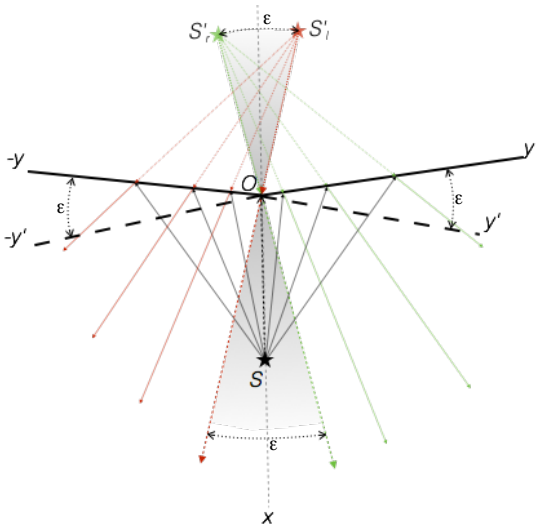

Figure 3: Reflection and scattering of sound rays on reflex angle. ε is the excess angle, since image angles overlap by ε.

As for obtuse angle, when rays rotate *clockwise* around the source, the reflected rays rotate *anticlockwise*, first around the image source S'_l, then around S'_r. In order to fill the gap at the apex, the image source must move *anticlockwise* from S'_l to S'_r. It can be seen on Fig. 3 that the diffracted rays then rotate *anticlockwise* around the apex, thus filling the grey sector with continuity of rotations at its boundaries. One considers again a continuum of image sources along the sector between S'_l and S'_r (grey sector) by rotating the image angle by ε, from position $(-y)Oy'$ to position $(-y')Oy$.

Introducing the axis Ox along the bisector of the reflex angle, Fig. 3 (2-dimensional case) can still be embedded in a 3-dimensional space, as displayed in Fig. 4. The notations are the same as in Sect. 2.2 and 2.3 for the 2-dimensional and 3-dimensional cases respectively. As a consequence, curvatures only occur

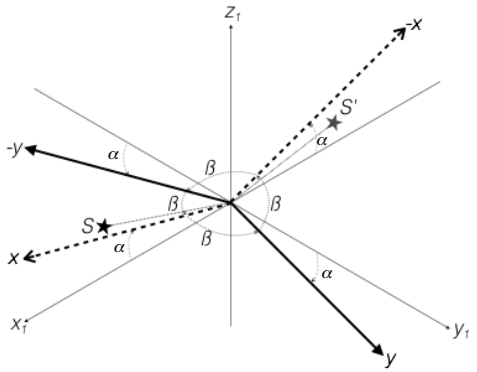

Figure 4: Embedding scattering on reflex angle in 3-dimensional space. Axis Ox is introduced as the bisector of the angle. Angles α and β have the same meaning as in Figs. 1 and 2. This time, the two first order image sources coalesce in one single image source S'.

on the "hinge" and keep the same values, and so do the metric tensors, the Christoffel symbols, and the Ricci tensors.

Note that only one image angle is created by scattering on a reflex angle, making them similar to flat edges or faces, except for the curvature localized on the hinges and the corresponding scattering.

2.5 Scattering on acute angles

The construction of Figs. 1 and 2 can easily be extended to any acute angle. In this case, the number of image angles must be increased, while keeping it odd, until one obtains an excess apical angle (see Fig. 5).

As for the two previous cases, when rays rotate *clockwise* around the original source, the reflected rays rotate *anticlockwise* around the first order image source S'_r, then *clockwise* around he second order source S''_l until the reflected ray reaches the apex. Continuity of rotation requires that this second order image source rotates *clockwise*, but it does not fill the gap between the two second order sources: image source S''_l moves away from S''_r until it reaches a second position, marked with a red broken line in the bottom panel of Fig. 5, at which further rotation makes it disappear. However, just before it disappears, a *third order* image source S''', not represented in the bottom panel of Fig. 5, emerges on the left. This third order image source fills the whole angle with diffracted rays while it keeps on rotating *clockwise* around the apex, until it disappears to the right. Shortly afterwards, the second order image source S''_r appears on the left at the position marked with the green broken line, and further *clockwise* rotation brings it to its original position in the top panel. Thus, this time, the diffracted rays fill the whole angle, plus the two sectors outside the grey sector on Fig. 5.

In order to correctly embed scattering on an acute angle in a 3-dimensional space, adapted sets of coordinates must be introduced; but curvature remains localized at the apical hinge and is equal to the deficit angle $-\varepsilon$ (not represented in Fig. 5), time the length of the hinge in the dihedral case. In fact, since the

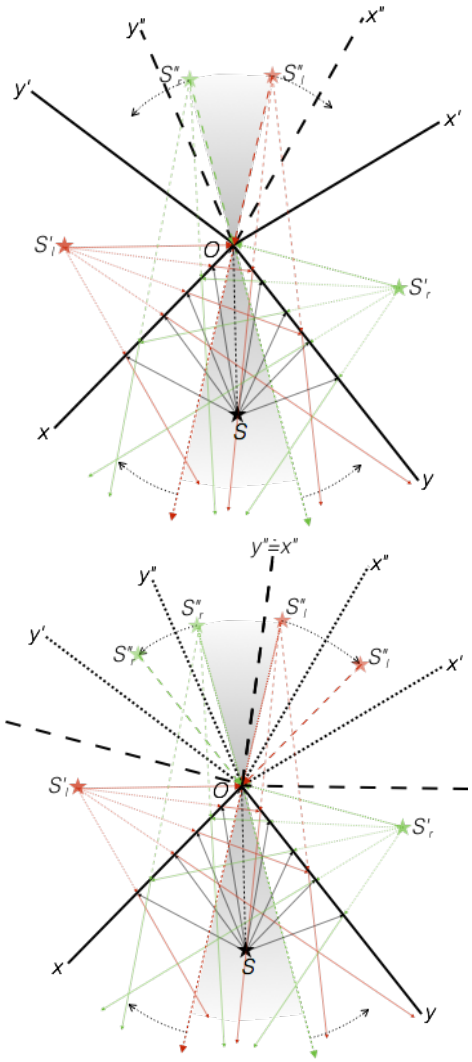

Figure 5: Reflection and scattering of sound rays on acute angle. Top: the visible rays; bottom: rotating S''_l clockwise makes a supplementary image source S''' appear and fill in the whole angle with diffracted rays (not represented), until S''_r eventually becomes visible on the left hand side. Note that for some positions no image source diffracts in the angle, whilst two sources can simultaneously diffract for some other positions.

space is Euclidean everywhere but on the apex, any directions can be chosen, provided that the angle between them is equal to $\frac{\pi}{2} - \frac{\varepsilon}{4} = \beta$. In Fig 6, the angle is located in the front sector, and its 5 images are displayed around the hinge O.

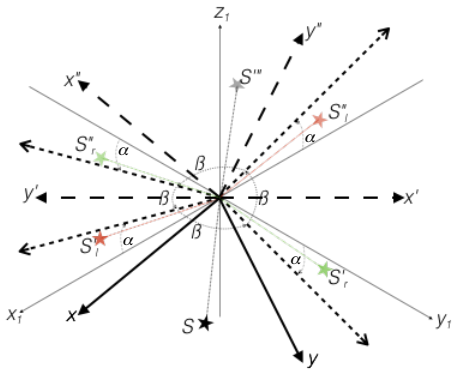

Figure 6: Embedding scattering on acute angle in 3-dimensional space. Short-tick axes represent the original axes of the warped plan; acute angle is inserted in the front sector. Angles α and β have same meanings as in previous Figures. The third order image sources coalesce in one single image source S'''.

Scattering is now created by the non overlapping highest order image sources radiating through the hinge.

3 Number of image sources

Let us now consider polygonal (2-dimensional case) or polyhedral rooms (3-dimensional case). We impose the restriction, that will be discussed at the end of this Section, that all the internal angles of the polygons are right or obtuse in order to obtain the 4 sectors of Figs. 1 and 2. In the polyhedral case, beside keeping all internal dihedral angles right or obtuse for the same reason, we also impose that all vertices are shared by three faces only. Typical examples are pentagons in 2-dimensions, and dodecahedra in 3-dimensions.

3.1 Convex polygonal rooms

For convex polygonal rooms with n edges, the n first order images are obtained by reflection on the edges. However, the order of reflection is not relevant for counting the images: layers are much more relevant. Thus, on Fig. 7, the numbers correspond to the successive layers around the original room.

Accordingly, the first layer is composed of the n images on the edges, to which n supplementary images, one at each vertex, are added. For the next layers, it is more efficient to compute separately the number of free edges and the number of free vertices, that is, edges and vertices that are not common to two adjacent images: one adds one image for each free edge, and one for each free vertex, exactly as for the first layer. It can be seen on Fig. 7 that edge images create $(n-3)$ free edges and $(n-4)$ free vertices; and vertex images

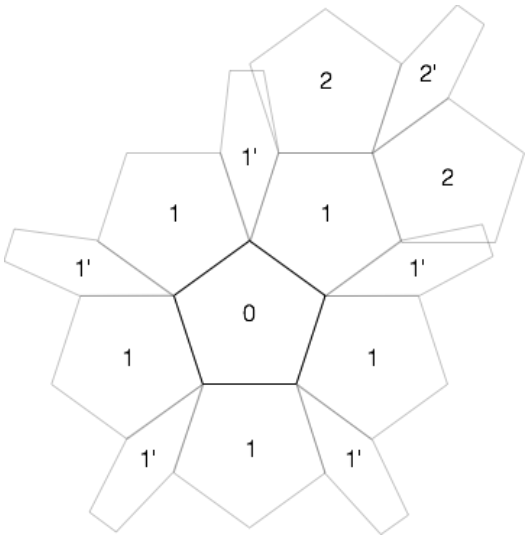

Figure 7: Image rooms of pentagon. Primed images are skewed in order to fit in the picture. Numbers correspond to successive layers of images, with 0 denoting the original room and primes one supplementary order of reflection (see text). Note that one supplementary image must be added each time image rooms overlap (e.g. between 1′and 2).

crate $(n-2)$ free edges and $(n-3)$ free vertices. Note that $n \geqslant 4$ if all angles are right or obtuse. Let's call e_i the number of free edges and v_i the number of free vertices of layer i. One obtains the recurrence formula:

$$\begin{pmatrix} e_i \\ v_i \end{pmatrix} = \begin{pmatrix} n-3 & n-2 \\ n-4 & n-3 \end{pmatrix} \begin{pmatrix} e_{i-1} \\ v_{i-1} \end{pmatrix} = \Lambda^i \begin{pmatrix} e_0 \\ v_0 \end{pmatrix} \tag{11}$$

where $\Lambda = \begin{pmatrix} n-3 & n-2 \\ n-4 & n-3 \end{pmatrix}$ and $\det \Lambda = 1$. The eigenvalues of matrix Λ are respectively $\lambda_1 = \frac{1}{2}(\sqrt{n-2}+\sqrt{n-4})^2 \geqslant 1$ and $\lambda_2 = \frac{1}{2}(\sqrt{n-2}-\sqrt{n-4})^2 \leqslant 1$.

Case $n = 4$: If $n = 4$, we obtain rectangular rooms, for which the number of image sources increases linearly with the layer order. Indeed, in this case, the matrix is upper triangular:

$$\Lambda = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

with $\lambda_1 = \lambda_2 = 1$, and the computation of eq. (11) is straightforward:

$$\Lambda^i = \begin{pmatrix} 1 & 2i \\ 0 & 1 \end{pmatrix}$$

As expected, we obtain for rectangular rooms $e_i = 4(2i+1)$ and $v_i = 4$: the number of corners remains constant and equal to 4, with a linear increase of the number of edge images. And the total number of image sources $\mathfrak{N}_i$ of layer i is $\mathfrak{N}_i = e_i + v_i = 8(i+1)$, that is, increases linearly with i.

Case $n > 4$: If $n \geqslant 5$, then $\lambda_1 > 1$ and $\lambda_2 < 1$, and the eigenvectors are given by:

$$X_1 = \begin{pmatrix} \sqrt{n-2} \\ \sqrt{n-4} \end{pmatrix},\ X_2 = \begin{pmatrix} \sqrt{n-2} \\ -\sqrt{n-4} \end{pmatrix}$$

that is, Λ can be factorized into $\Lambda = S^{-1}DS$ with:

$$S = \begin{pmatrix} \frac{1}{2\sqrt{n-2}} & \frac{1}{2\sqrt{n-4}} \\ \frac{1}{2\sqrt{n-2}} & -\frac{1}{2\sqrt{n-4}} \end{pmatrix}$$

,

$$D = \begin{pmatrix} \frac{1}{2}(\sqrt{n-2}+\sqrt{n-4})^2 & 0 \\ 0 & \frac{1}{2}(\sqrt{n-2}-\sqrt{n-4})^2 \end{pmatrix}$$

and therefore:

$$S^{-1} = \begin{pmatrix} \sqrt{n-2} & \sqrt{n-2} \\ \sqrt{n-4} & -\sqrt{n-4} \end{pmatrix}$$

For large values of the layer number i, $\lambda_2^i \longrightarrow 0$, and eq. (11) reduces to:

$$\begin{pmatrix} e_i \\ v_i \end{pmatrix} \approx \frac{(\sqrt{n-2}+\sqrt{n-4})^{2i}}{2^{i+1}} \begin{pmatrix} 1 & \frac{\sqrt{n-2}}{\sqrt{n-4}} \\ \frac{\sqrt{n-4}}{\sqrt{n-2}} & 1 \end{pmatrix} \begin{pmatrix} e_0 \\ v_0 \end{pmatrix}$$

with $e_0 = v_0 = n$, that is:

$$e_i \approx \frac{n(\sqrt{n-2}+\sqrt{n-4})^{2i+1}}{2^{i+1}\sqrt{n-4}}$$
$$v_i \approx \frac{n(\sqrt{n-2}+\sqrt{n-4})^{2i+1}}{2^{i+1}\sqrt{n-2}}$$

and the total number of image sources $\mathfrak{N}_i$ of layer i is given by the sum of the two contributions, that is:

$$\begin{aligned} \mathfrak{N}_i = e_i + v_i &\approx \frac{n(\sqrt{n-2}+\sqrt{n-4})^{2i+1}}{2^{i+1}} \left(\frac{1}{\sqrt{n-4}} + \frac{1}{\sqrt{n-2}} \right) \\ &= \frac{n(\sqrt{n-2}+\sqrt{n-4})^{2(i+1)}}{2^{i+1}\sqrt{(n-2)(n-4)}} \end{aligned} \tag{12}$$

Example: For a **pentagon**, $n = 5$ and the eigenvalues are $\lambda_1 = \frac{1}{2}(\sqrt{3}+1)^2 \approx 3.73$ and $\lambda_2 = \frac{1}{2}(\sqrt{3}-1)^2 \approx\approx 0.27$. Matrix Λ is now equal to:

$$\Lambda = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$$

which reduces to:

$$\Lambda = \frac{1}{2} \begin{pmatrix} \sqrt{3} & \sqrt{3} \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2+\sqrt{3} & 0 \\ 0 & 2-\sqrt{3} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & 1 \\ \frac{1}{\sqrt{3}} & -1 \end{pmatrix}$$

and the total number of image sources $\mathfrak{N}_i$ of layer i is given by eq. (12):

$$\mathfrak{N}_i = e_i + v_i \approx \frac{5\left(\sqrt{3}+1\right)^{2(i+1)}}{2^{i+1}\sqrt{3}}$$

In other words, the number of image sources increases exponentially with the order of the layer, a very different behaviour than for rectangular rooms, but similar to mixing rooms [Pol92].

3.2 Convex polyhedral rooms

For convex polyhedral rooms, it is not sufficient to only consider the number N of faces. We must also consider the number n_i of edges of each face i. We thus obtain:

- the number of faces: $F = \sum_i 1 = N$
- the number of edges: since one edge is common to 2 faces, this number is $E = \frac{1}{2} \sum_i n_i$
- the number of vertices: with the assumption that vertices are shared by 3 faces only, this number is $V = \frac{1}{3} \sum_i n_i$

With the help of Euler's polyhedron formula $F - E + V = 2$, valid for convex polyhedra, we obtain:

$$F - E + V = \sum_i \left[1 - \frac{n_i}{2} + \frac{n_i}{3}\right] = \sum_i \left[1 - \frac{n_i}{6}\right] = 2$$

that is, introducing the mean number of edges per face $\bar{n} = \frac{1}{N} \sum_i n_i$:

$$(6 - \bar{n})N = 12$$

As a consequence, the mean number of edges per face, the total number of edges, and the total number of vertices are given by:

$$\bar{n} = \frac{6(N-2)}{N},\ E = 3(N-2),\ V = 2(N-2) \tag{13}$$

Note that the assumption that vertices are shared by 3 faces only does not introduce any restriction, as it is easy to "regularize" vertices shared by $m > 3$ faces by cutting-off the vertex according to Fig. 8, thus increasing by 1 the number of faces, by m the number of edges edges, and by $m - 1$ the number of vertices, that is, keeping the Euler characteristic $F - E + V$ equal to 2.

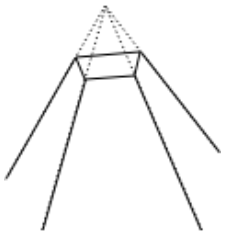

Figure 8: Regularizing vertex shared by 4 faces, by cutting-off the tip of the pyramid, thus removing one vertex but introducing 4 new edges and vertices.

As in Sect. 3.1, we consider successive layers of image rooms built around the original rooms, and we compute separately the number of free faces, edges and vertices belonging to the images created by free faces, edges an vertices. Simple enumeration leads to:

- each free face i creates:

- $N - (n_i + 1)$ free faces,
- $E - 2n_i - \sum_j \epsilon_{ij}(n_j - 3)$ free edges,
- $V - 2n_i - \sum_j \epsilon_{ij}(n_j - 4)$ free vertices,

where $\epsilon_{ij} = 1$ if faces i and j share one edge, 0 otherwise, so that $\sum_i \epsilon_{ij} = n_j$;

- each free edge k creates:
  - $N - 4$ free faces,
  - $E - 5 - \sum_j \gamma_{kj}(n_j - 3) - \sum_{lj} \delta_{klj}(n_j - 2)$ free edges,
  - $V - 6 - \sum_j \gamma_{kj}(n_j - 4) - \sum_{lj} \delta_{klj}(n_j - 3)$ free vertices,

  where $\gamma_{kj} = 1$ if edge k belongs to face j and 0 otherwise, with $\sum_j \gamma_{kj} = 2$; and where $\delta_{klj} = 1$ if edge k belongs to face j and vertex l to edge k, 0 otherwise, with $\sum_{lj} \delta_{klj} = 2$;
- each free vertex l creates:
  - $N - 3$ free faces,
  - $E - 3 - \sum_j \alpha_{lj}(n_j - 2)$ free edges,
  - $V - 4 - \sum_j \alpha_{lj}(n_j - 3)$ free vertices,

  where $\alpha_{lj} = 1$ if vertex l belongs to face j, 0 otherwise, so that $\sum_j \alpha_{lj} = 3$.

Let's now call f_m, e_m and v_m the numbers of free faces, edges and vertices respectively in layer m. We obtain the recurrence formulae:

$$
\begin{aligned}
f_{m+1} =& \sum_{i=1}^{f_m} [N - (n_i + 1)] + \sum_{k=1}^{e_m} [N - 4] + \sum_{l=1}^{v_m} [N - 3] \\
e_{m+1} =& \sum_{i=1}^{f_m} \left[E - 2n_i - \sum_j \epsilon_{ij}(n_j - 3) \right] \\
&+ \sum_{k=1}^{e_m} \left[E - 5 - \sum_j \gamma_{kj}(n_j - 3) - \sum_{lj} \delta_{klj}(n_j - 2) \right] \\
&+ \sum_{l=1}^{v_m} \left[E - 3 - \sum_j \alpha_{lj}(n_j - 2) \right] \\
v_{m+1} =& \sum_{i=1}^{f_m} \left[V - 2n_i - \sum_j \epsilon_{ij}(n_j - 4) \right] \\
&+ \sum_{k=1}^{e_m} \left[V - 6 - \sum_j \gamma_{kj}(n_j - 4) - \sum_{lj} \delta_{klj}(n_j - 3) \right] \\
&+ \sum_{l=1}^{v_m} \left[V - 4 - \sum_j \alpha_{lj}(n_j - 3) \right]
\end{aligned}
$$

Assuming that the number of image rooms increases exponentially with the layer order m, it is legitimate to replace numbers n_i and n_j by their mean $\bar{n}$, leading to the approximation:

$$\begin{aligned}
f_{m+1} \approx & f_m\,[N-\bar{n}-1] + e_m\,[N-4] + v_m\,[N-3] \\
e_{m+1} \approx & f_m\left[E-2\bar{n}-\bar{n}^2+3\bar{n})\right] + e_m\,[E-5-2(\bar{n}-3)-2(\bar{n}-2)] \\
& + v_m\,[E-3-3(\bar{n}-2)] \\
v_{m+1} \approx & f_m\left[V-2\bar{n}-\bar{n}^2+4\bar{n})\right] + e_m\,[V-6-2(\bar{n}-4)-2(\bar{n}-3)] \\
& + v_m\,[V-4-3(\bar{n}-3)]
\end{aligned}$$

that is, to the recurrence formula:

$$\begin{aligned}
\begin{pmatrix} f_{m+1} \\ e_{m+1} \\ v_{m+1} \end{pmatrix} &= \begin{pmatrix} N-\bar{n}-1 & N-4 & N-3 \\ \left(\frac{N}{2}+1\right)\bar{n}-\bar{n}^2 & \left(\frac{N}{2}-4\right)\bar{n}+5 & \left(\frac{N}{2}-3\right)\bar{n}+3 \\ \left(\frac{N}{3}+2\right)\bar{n}-\bar{n}^2 & \left(\frac{N}{3}-4\right)\bar{n}+8 & \left(\frac{N}{3}-3\right)\bar{n}+5 \end{pmatrix} \begin{pmatrix} f_m \\ e_m \\ v_m \end{pmatrix} \\
&= \Lambda^{m+1} \begin{pmatrix} f_0 \\ e_0 \\ v_0 \end{pmatrix}
\end{aligned} \tag{14}$$

with matrix Λ given by:

$$\Lambda = \begin{pmatrix} N-\bar{n}-1 & N-4 & N-3 \\ \left(\frac{N}{2}+1\right)\bar{n}-\bar{n}^2 & \left(\frac{N}{2}-4\right)\bar{n}+5 & \left(\frac{N}{2}-3\right)\bar{n}+3 \\ \left(\frac{N}{3}+2\right)\bar{n}-\bar{n}^2 & \left(\frac{N}{3}-4\right)\bar{n}+8 & \left(\frac{N}{3}-3\right)\bar{n}+5 \end{pmatrix}$$

A lengthy computation leads to the following characteristic equation for eigenvalues λ:

$$\begin{aligned}
-\lambda^3 + \lambda^2\left[\left(\frac{5N}{6}-8\right)\bar{n}+(N-9)\right] - \lambda\left[\frac{4N}{3}\bar{n}^2 - \left(\frac{26N}{3}-8\right)\bar{n}+(10N-9)\right] \\
+\left[-\frac{N}{6}\bar{n}^2+(N-1)\right] = 0
\end{aligned}$$

which, when replacing $\bar{n}$ by its value given in eq. (13), simply reduces to:

$$-\lambda^3 + \lambda^2\left[6N-49+\frac{96}{N}\right] - \lambda\left[6N-49+\frac{96}{N}\right] + 1 = 0$$

or more simply to:

$$-\lambda^3 + \lambda^2\left[\frac{6(N-4)^2}{N}-1\right] - \lambda\left[\frac{6(N-4)^2}{N}-1\right] + 1 = 0 \tag{15}$$

Case $N = 6$: If $N = 6$, we obtain rectangular parallelepiped rooms, for which eq. (13) reduces to $n_i = \bar{n} = 4$ for all i, $e_0 = 12$, and $v_0 = 8$ with $f_0 = N = 6$. In this case, the matrix is upper triangular:

$$\Lambda = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

Since all diagonal terms are equal to 1, the three eigenvalues are equal to 1 and the computation of eq. (11) is straightforward:

$$\Lambda^i = \begin{pmatrix} 1 & 2i & 3i^2 \\ 0 & 1 & 3i \\ 0 & 0 & 1 \end{pmatrix}$$

We obtain for rectangular parallelepiped rooms $f_i = 6(2i+1)^2$, $e_i = 12(2i+1)$ and $v_i = 8$: as expected, the number of corners remains constant and equal to 8, with a linear increase of the number of edge images and a quadratic increase of the number of the face images. And the total number of image sources $\mathfrak{N}_i$ of layer i is $\mathfrak{N}_i = f_i + e_i + v_i = 24(i+1)^2 + 2$, that is, increases quadratically with i.

Case $N > 6$: The direct solution of eq. (15) gives then the three eigenvalues:

- $\lambda_0 = 1$,
- $\lambda_1 = \left[\frac{3(N-4)^2}{N} - 1\right] + \frac{N-4}{N}\sqrt{3[(N-6)(3N-8)]} > 1$,
- $\lambda_2 = \left[\frac{3(N-4)^2}{N} - 1\right] - \frac{N-4}{N}\sqrt{3[(N-6)(3N-8)]} < 1$.

and matrix Λ reduces to:

$$\Lambda = \begin{pmatrix} \frac{(N-2)(N-6)}{N} + 1 & N-4 & N-3 \\ \frac{3(N-2)(N-4)(N-6)}{N^2} & \frac{3(N-4)^2}{N} - 1 & \frac{3(N-3)(N-4)}{N} \\ \frac{2(N-2)(N-6)^2}{N^2} & \frac{2(N-4)(N-6)}{N} & \frac{2(N-3)(N-6)}{N} + 1 \end{pmatrix}$$

A simple matrix manipulation (see Appendix) easily gives the 3 eigenvectors associated to the 3 eigenvalues:

$$X_0 = \begin{pmatrix} (N-3) \\ 0 \\ -\frac{(N-2)(N-6)}{N} \end{pmatrix},\ X_1 = \begin{pmatrix} \sqrt{3[(N-6)(3N-8)]} \\ +\frac{3(N-6)(3N-8)}{N} \\ \frac{2(N-6)}{N}\sqrt{3[(N-6)(3N-8)]} \end{pmatrix},$$

$$X_2 = \begin{pmatrix} \sqrt{3[(N-6)(3N-8)]} \\ -\frac{3(N-6)(3N-8)}{N} \\ \frac{2(N-6)}{N}\sqrt{3[(N-6)(3N-8)]} \end{pmatrix}$$

Thus, matrix Λ can be factorized into $\Lambda = S^{-1}DS$, with the column of matrix S^{-1} equal to the eigenvectors:

$$S^{-1} = \begin{pmatrix} (N-3) & \sqrt{3[(N-6)(3N-8)]} & \sqrt{3[(N-6)(3N-8)]} \\ 0 & \frac{3(N-6)(3N-8)}{N} & -\frac{3(N-6)(3N-8)}{N} \\ -\frac{(N-2)(N-6)}{N} & \frac{2(N-6)}{N}\sqrt{3[(N-6)(3N-8)]} & \frac{2(N-6)}{N}\sqrt{3[(N-6)(3N-8)]} \end{pmatrix}$$

and matrix D given by the eigenvalues:

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}\left[\sqrt{\frac{3}{N}}(N-4) + \sqrt{\frac{(N-6)(3N-8)}{N}}\right]^2 & 0 \\ 0 & 0 & \frac{1}{2}\left[\sqrt{\frac{3}{N}}(N-4) - \sqrt{\frac{(N-6)(3N-8)}{N}}\right]^2 \end{pmatrix}$$

where we have used an alternative form for the eigenvalues λ_1 and λ_2, leading by inversion to matrix S:

$$S = \frac{1}{2}\begin{pmatrix} \frac{4}{3N-8} & 0 & -\frac{2N}{(N-6)(3N-8)} \\ \frac{N-2}{(3N-8)\sqrt{3[(N-6)(3N-8)]}} & \frac{N}{3(N-6)(3N-8)} & \frac{N(N-3)}{(N-6)(3N-8)} \\ \frac{N-2}{(3N-8)\sqrt{3[(N-6)(3N-8)]}} & -\frac{N}{3(N-6)(3N-8)} & \frac{N(N-3)}{(N-6)(3N-8)} \end{pmatrix}$$

.

For large values of the layer number i, $\lambda_2^i \longrightarrow 0$, and eq. (14) is approximated by:

$$\begin{pmatrix} f_i \\ e_i \\ v_i \end{pmatrix} \approx \frac{\left[\sqrt{\frac{3}{N}}(N-4) + \sqrt{\frac{(N-6)(3N-8)}{N}}\right]^{2i}}{2^{i+1}(3N-8)}$$

$$\begin{pmatrix} N-2 & \frac{N\sqrt{3[(N-6)(3N-8)]}}{3(N-6)} & \frac{N(N-3)\sqrt{3[(N-6)(3N-8)]}}{3(N-6)} \\ \frac{(N-2)\sqrt{3[(N-6)(3N-8)]}}{N} & 3N-8 & (N-3)(3N-8) \\ \frac{3(N-2)(N-6)}{N} & 2\sqrt{3[(N-6)(3N-8)]} & 2(N-3)\sqrt{3[(N-6)(3N-8)]} \end{pmatrix}$$

$$\begin{pmatrix} f_0 \\ e_0 \\ v_0 \end{pmatrix} \tag{16}$$

Example: For a **dodecahedron**, $N = 12$ with $\bar{n} = 5$, $e_0 = 30$, and $v_0 = 20$ with $f_0 = N = 12$. The eigenvalues are $\lambda_0 = 1$, $\lambda_1 = 15 + 4\sqrt{14} \approx 29.97$ and $\lambda_2 = 15 - 4\sqrt{14} = 0.03$. Matrix Λ is now equal to:

$$\Lambda = \begin{pmatrix} 9 & 8 & 9 \\ 10 & 15 & 18 \\ 5 & 8 & 10 \end{pmatrix}$$

which reduces to:

$$\Lambda = \frac{1}{2}\begin{pmatrix} 9 & 6\sqrt{14} & 6\sqrt{14} \\ 0 & 28 & -28 \\ -5 & 6\sqrt{14} & 6\sqrt{14} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 15+4\sqrt{14} & 0 \\ 0 & 0 & 15-4\sqrt{14} \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{7} & 0 & -\frac{1}{7} \\ \frac{5}{84\sqrt{14}} & \frac{1}{84} & \frac{9}{28} \\ \frac{5}{84\sqrt{14}} & -\frac{1}{84} & \frac{9}{28} \end{pmatrix}$$

and the total number of image sources $\mathfrak{N}_i$ of layer i is given by the sum of the 3 terms in eq. (16), that is:

$$\mathfrak{N}_i = f_i + e_i + v_i \approx \frac{\left(15+4\sqrt{14}\right)^i}{46}\Big(5\left[4+\sqrt{14}\right]f_0 + 2\left[15+2\sqrt{14}\right]e_0 + 37\left[7+4\sqrt{14}\right]v_0\Big) = \frac{10(29.97)^i\left(316+157\sqrt{14}\right)}{23}$$

In other words, the number of image sources increases exponentially with the order of the layer, as is the case for polygonal rooms with more than 4 edges.

Note that the case of vertices shared by more than 3 faces is obtained by first regularising them according to Fig. 8, then moving the intersecting face toward the original vertex: the number of image rooms remains constant. Therefore, assuming that all vertices are shared by 3 faces only does not introduce any restriction.

3.3 Case of reflex angles

The case of reflex angles can now be handled by noticing that reflex angles do not introduce new images. For polygons, the two adjacent edges must be considered as a unique edge, thus diminishing the number of edges and vertices by one. For polyhedra, the two adjacent faces becomes one unique face, thus diminishing the number of faces and edges by one. However, when two reflex dihedral angles share a common vertex between two common adjacent edges, the two adjacent edges also coalesce into one single edge, diminishing both the number of edges and the number of vertices by one, thus keeping Euler characteristic constant and equal to 2.

3.4 Case of acute angles

The case of acute angles can also be handled by noticing that they introduce several new images, depending on the angle. However, both for polygons and for polyhedra, each case must be handled specifically.

4 Visible image sources

In the previous section, we have demonstrated that, except for rectangular rooms, the number of image sources increases exponentially, most of them creating diffracted rays, due to the excess angles at the hinges (see Sec. 2). They correspond to what is usually called "hidden image sources", that is, image sources that are not visible from the receiver. We therefore now address the question of the visible images, and the related question of their *horizon of visibility*, which we define as the distance at which visible image sources smear within the swarm of image sources because the intensity they radiate into the original room starts decreasing exponentially with distance. In doing so, we must handle separately the cases of regular polygons and polyhedra, and the case of irregular ones.

4.1 Regular polygons and polyhedra

The most striking feature of space tessellation by regular polygons or polyhedra (Fig. 9) is the fact that any grouping of one cell and its image on any of its edges or faces displays parallel edges or faces. As a consequence, images sources are visible at long range along some specific discrete directions. Nevertheless, every now and then, one image source becomes hidden behind some hinge, even in these specific directions (red rays in Fig. 9), so that one should rather talk of *channels* - in green in Fig. 9.

It can be seen in Fig. 9 that channels are defined by the repetition of patterns along some discrete directions. The number of these directions is numerable, since there is a numerable number of image sources; and the widths of

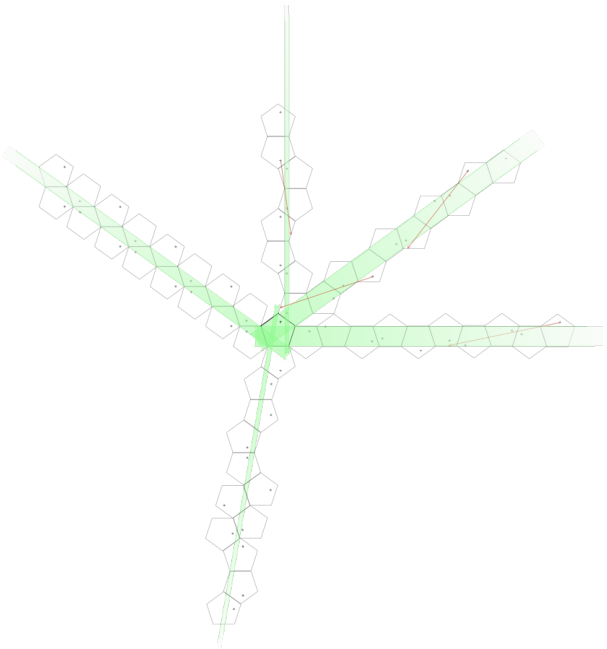

Figure 9: Visible (green rays) and invisible (red rays) images for a regular pentagon. Notice the long-range visibility in some direction.

the channels decrease with the length of the repetition pattern. In any case, at some distance, the widths of the channels become narrow enough so that path differences for all rays joining the image sources to the original room must be taken into account, as they eventually reduce the intensity of the sources when they become narrower that the first Fresnel zones around the direct rays. As a consequence, all distant visible sources eventually radiate an exponentially decreasing intensity, no longer inversely proportional to the square of the travelled distance. In that respect, distant sources do not differ from hidden sources; but the distances at which these transitions operate are frequency dependant, increasing with frequency.

The number of visible images at any range increases at most with distance in the polygon case, and squared distance in the polyhedra case. This is a consequence of Fig. 9, since the cells corresponding to visible images constitute a subset of the Euclidean plane in which the original room lays, bordered by bold lines in the Figure. As the average number of cells in any annulus, resp. any shell, of inner and outer radii d and $d + \Delta d$ is overestimated by dividing its surface by the surface S of the cell, resp. its volume by the volume V of the cell, an overestimation for the number of visible images at distance d is therefore $\frac{2\pi d \Delta d}{S}$, resp. $\frac{4\pi d^2 \Delta d}{V}$, which increases at most with the distance for a polygon, and with the squared distance for a polyhedron. Thus, the number of hidden sources will quickly outnumber the number of visible sources, thus diluting their strength.

Can we evaluate the *horizon for visible sources*? From the above discussion and definition, the maximum distance of visibility can be estimated from the width w of the channel and the wavelength λ of the frequency under consider-

ation as a few times the distance H defined by:

$$H \approx \frac{w^2}{\lambda} \tag{17}$$

Note that no horizon exists for rectangular rooms as no scattering occurs on the vertices, resp. the edges, since no excess angle exists at vertices or edges through reflections.

4.2 Irregular polygons and polyhedra

Space tessellation by irregular polygons or polyhedra (Fig. 10) does not exhibit parallel edges or faces after reflections, and no specific patterns are repeated along any direction. Indeed, as can be seen in Fig. 10, all channels eventually split in narrower ones as they cross some hinges, a process which is signalled by blue arrows in Fig. 10. As a consequence, the horizon of visibility of the image sources is much shorter than in the regular case.

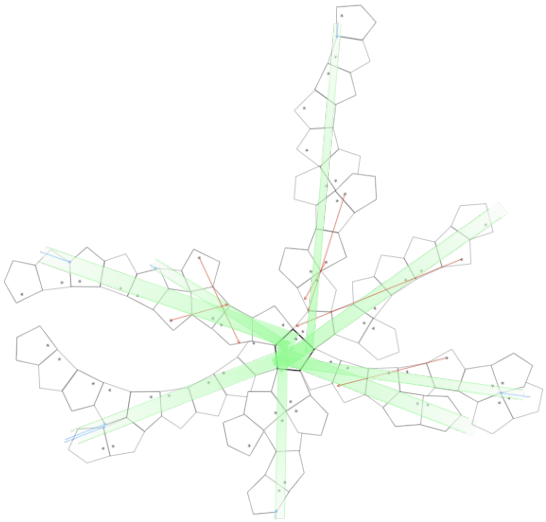

Figure 10: Visible (green rays) and invisible (red rays) images for an irregular pentagon. Blue arrows signal channel splitting at hinges. Notice that visibility quickly vanishes due to random distribution of hinges (here, all vertices).

As can be seen in Fig. 10, the orientations of the image rooms gradually become random with distance, as the angles between non-adjacent, opposing edges or faces are no longer rational ratios of 2π. This creates random distribution of the hinges, and thus the splitting of the channels and eventually scattering. As a consequence, the horizon of visibility of the image sources can be estimated as a few times the characteristic sizes of the cells, at most 10 times in Fig. 10.

5 Taking scattering in account: conservation of stress-energy tensor

Up to now, we have not considered the nature of the signals emitted by the sources. As we are interested in sound waves, the wave properties of the field has to be taken into account. Thus, a time dimension must be added to the space of Sec. 2.1 in order to introduce the wave equation.

5.1 Wave equation and stress-energy tensor

From now on, we consider a 4-dimensional time-space with its metric tensor g_{ij} and the volume element $dV = \sqrt{|g|}dx^0 \dots dx^3$, where $g = \det g_{ij}$ [Lin 5]. But now, g is *negative* as the first eigenvalue of the metric tensor is negative and equal to $-c^2$, where c is the speed of sound: its eigenvector corresponds to the time direction dx^0. The infinitesimal distance element is still given by eq. 1:

$$ds^2 = g_{ij}dx^i dx^j \tag{18}$$

Therefore, covariant derivation, Christoffel symbols, Riemann and Ricci tensor as well as the scalar curvature are still defined as in Sec. 2.1. More precisely, according to Sec. 2.3, we have:

$$g_{ij} = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & \mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta & 0 \\ 0 & \mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{19}$$

$$g^{ij} = \begin{pmatrix} -c^{-2} & 0 & 0 & 0 \\ 0 & \frac{1}{\sin^2\beta} & -\frac{\mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta}{\sin^2\beta} & 0 \\ 0 & -\frac{\mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta}{\sin^2\beta} & \frac{1}{\sin^2\beta} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where β is one fourth of the total dihedral angle around the hinge, and with determinant $g = -c^2 \sin^2\beta < 0$. g_{ij} and g^{ij} reduce to diagonal matrices on the arms, now denoted ($x^1 = 0$ and $x^2 = 0$), with all diagonal elements equal to 1 except $g_{00} = -c^2$ and $g^{00} = -c^{-2}$, and with $g = -c^2$. As for the Christoffel symbols, they are all equal to 0, but for $\Gamma^2_{11} = 2\delta(x^1)\mathrm{sgn}(x^2)\cos\beta$ and $\Gamma^1_{22} = 2\mathrm{sgn}(x^1)\delta(x^2)\cos\beta$ (eq 7); and all elements of the Ricci tensor are 0 except $R_{11} = R_{22} = 4\delta(x^1)\delta(x^2)\cos\beta < 0$ (eq. 8). The scalar curvature therefore remains $R = 8\delta(x^1)\delta(x^2)\cos\beta < 0$, but the local curvature, obtained by integrating half the scalar curvature over a small space element, is now equal to $\frac{1}{2}\int R\sqrt{|g|}dx^0dx^1dx^2dx^3$. It remains null everywhere - flat space - except on the hinges, now given by the apical edges extended by the time laps. In the limit where β tends toward a right angle, the curvature around any hinge is equal to $c(2\pi - 4\beta)\ell\Delta t < 0$, where ℓ is the length of the corresponding apical edge and Δt the time lag.

We now introduce the wave equation:

$$\Box\Phi = \nabla_i g^{ij}\partial_j\Phi = 0 \tag{20}$$

where Φ is the velocity potential. Pressure p and particle velocity v are obtained from the velocity potential as covectors:

$$p = \rho \partial_0 \Phi$$

$$v_i = -\partial_i \Phi$$

where i takes value 1, 2 or 3 (space variables), and where ρ is the density of the fluid in which the waves propagate. Note that it is more useful to use the vector formulations for these two quantities:

$$v^i = -\nabla^i \Phi = -g^{ij} \partial_j \Phi \tag{21}$$

meaning that:

$$v^0 = -\nabla^0 \Phi = -g^{00} \partial_0 \Phi = c^{-2} \partial_0 \Phi = \frac{p}{\rho c^2}$$

In general, the velocity potential Φ is a complex function. We therefore consider the product $\partial_k \Phi^* \Box \Phi$. Differentiation rules lead to:

$$\begin{aligned} \partial_k \Phi^* \Box \Phi = \partial_k \Phi^* \nabla_i g^{ij} \partial_j \Phi &= \nabla_i g^{ij} \left[\partial_k \Phi^* \partial_j \Phi \right] - \left[\nabla_i \partial_k \Phi^* \right] g^{ij} \partial_j \Phi \\ &= \nabla_i g^{ij} \left[\partial_k \Phi^* \partial_j \Phi \right] - \left[\nabla_k \partial_i \Phi^* \right] g^{ij} \partial_j \Phi \\ &= \nabla^j \left[\partial_k \Phi^* \partial_j \Phi \right] - \left[\nabla_k \partial_i \Phi^* \right] g^{ij} \partial_j \Phi = 0 \end{aligned}$$

As i and j are mute indices, keeping only the real part of the preceding equation leads to:

$$\nabla^j \left(\partial_j \Phi^* \partial_k \Phi + \partial_j \Phi \partial_k \Phi^* \right) = \nabla_k \left(\partial_i \Phi^* g^{ij} \partial_j \Phi \right) \tag{22}$$

that is, to:

$$\nabla^i T_{ij} = 0 \tag{23}$$

where T_{ij} is the *symmetrical* stress-energy tensor, defined by:

$$T_{ij} = \frac{\partial_i \Phi^* \partial_j \Phi + \partial_i \Phi \partial_j \Phi^*}{2} - \frac{1}{2} g_{ij} \left(\partial_i \Phi^* g^{ij} \partial_j \Phi \right) \tag{24}$$

It is easy to recognise that the eq. (23) corresponds to the contravariant conservation of the stress-energy tensor.

5.2 Conservation of stress-energy tensor

The conservation of the stress-energy tensor takes a simpler form for T^{ij} than for T_{ij}, which still is symmetric. Indeed, Eq (23) can be written as:

$$g^{lj} \nabla^i T_{il} = g^{ik} \nabla_k T_{il} g^{lj} = \nabla_i T^{ij} = \partial_i T^{ij} + \Gamma^i_{ik} T^{kj} + \Gamma^j_{ik} T^{ik} = 0$$

that is ([Lin 5] p.54):

$$\frac{1}{\sqrt{|g|}} \partial_i (\sqrt{|g|} T^{ij}) + \Gamma^j_{ik} T^{ik} = 0$$

or, after integration on a small 4-dimensional volume V with border ∂V:

$$\int_{\partial V} n_i T^{ij} dS + \int_V \Gamma^j_{ik} T^{ik} dV = 0 \tag{25}$$

In the last equation, n_i is the outgoing normal covector to the boundary, normalized by $n_i g^{ij} n_j = \pm 1$, with a negative sign for time boundaries. If part of the border ∂V is defined by equation $f(x^0 \dots x^3) = 0$, then n_i is given by:

$$n_i = \frac{\pm \partial_i f}{\sqrt{\pm \partial_i f g^{ij} \partial_i f}}$$

where the same sign is used in the numerator and the denominator. And dS is the volume element of the border induced by the metric on ∂V. In other words, it includes the term $\sqrt{|g|}$.

In eq. (25), $T^{ij} = \frac{1}{2}\left[g^{ik}(\Phi_k \Phi_l^* + \Phi_k^* \Phi_l) g^{lj} - g^{ij}\left(\partial_i \Phi^* g^{ij} \partial_j \Phi\right)\right]$ can also be written:

$$T^{ij} = \frac{1}{2}\left[(\Phi^i \Phi^{j*} + \Phi^{i*} \Phi^j) - g^{ij}\left(\Phi^{i*} g_{ij} \Phi^j\right)\right]$$

where the Φ_i, resp. the Φ^j, are the partial covariant derivatives, resp. partial contravariant derivatives, of the velocity potential Φ. Note that the last form is preferred, as it makes use of vectors instead of covectors.

$\left(\Phi^{i*} g_{ij} \Phi^j\right)$ is then given by:

$$\begin{aligned}\left(\Phi^{i*} g_{ij} \Phi^j\right) = \big[&-c^2|\Phi^0|^2 + |\Phi^1|^2 + |\Phi^2|^2 + |\Phi^3|^2 \\ &+ 2\mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta \Re(\Phi^1 \Phi^{2*})\big]\end{aligned}$$

and T^{ij} by:

$$\begin{pmatrix}
\frac{c^{-2}}{2}\left[c^2|\Phi^0|^2 + |\Phi^1|^2 + |\Phi^2|^2 + |\Phi^3|^2 + 2\mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta\Re(\Phi^1\Phi^{2*})\right] & \Re(\Phi^0\Phi^{1*}) & \Re(\Phi^0\Phi^{2*}) & \Re(\Phi^0\Phi^{3*}) \\
\Re(\Phi^0\Phi^{1*}) & \frac{\sin^{-2}\beta}{2}\left[c^2|\Phi^0|^2 - \cos 2\beta|\Phi^1|^2 - |\Phi^2|^2 - |\Phi^3|^2 - 2\mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta\Re(\Phi^1\Phi^{2*})\right] & \frac{\Re(\Phi^1\Phi^{2*})}{\sin^4\beta} + \frac{\mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta}{2\sin^2\beta}\left[-c^2|\Phi^0|^2 + |\Phi^1|^2 + |\Phi^2|^2 + |\Phi^3|^2\right] & \Re(\Phi^1\Phi^{3*}) \\
\Re(\Phi^0\Phi^{2*}) & \frac{\Re(\Phi^1\Phi^{2*})}{\sin^4\beta} + \frac{\mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta}{2\sin^2\beta}\left[-c^2|\Phi^0|^2 + |\Phi^1|^2 + |\Phi^2|^2 + |\Phi^3|^2\right] & \frac{\sin^{-2}\beta}{2}\left[c^2|\Phi^0|^2 - |\Phi^1|^2 - \cos 2\beta|\Phi^2|^2 - |\Phi^3|^2 - 2\mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta\Re(\Phi^1\Phi^{2*})\right] & \Re(\Phi^2\Phi^{3*}) \\
\Re(\Phi^0\Phi^{3*}) & \Re(\Phi^1\Phi^{3*}) & \Re(\Phi^2\Phi^{3*}) & \frac{1}{2}\left[c^2|\Phi^0|^2 - |\Phi^1|^2 - |\Phi^2|^2 + |\Phi^3|^2 - 2\mathrm{sgn}(x^1)\mathrm{sgn}(x^2)\cos\beta\Re(\Phi^1\Phi^{2*})\right]
\end{pmatrix}$$

As for the conservation of the stress-energy tensor, it reduces to:

$$\begin{aligned}
&\int_{\partial V} n_i T^{i0} dS = 0 \\
&\int_{\partial V} n_i T^{i1} dS + \int_V \Gamma^1_{22} T^{22} dV = 0 \\
&\int_{\partial V} n_i T^{i2} dS + \int_V \Gamma^2_{11} T^{11} dV = 0 \\
&\int_{\partial V} n_i T^{i3} dS = 0
\end{aligned}$$

that is to:

$$\begin{aligned}
&\int_{\partial V} n_i T^{i0} dS = 0 \\
&\int_{\partial V} n_i T^{i1} dS = -2\cos\beta \int_V \mathrm{sgn}(x^1)\delta(x^2) T^{22} \sqrt{|g|} dx^0 dx^1 dx^2 dx^3 \\
&\int_{\partial V} n_i T^{i2} dS = -2\cos\beta \int_V \delta(x^1)\mathrm{sgn}(x^2) T^{11} \sqrt{|g|} dx^0 dx^1 dx^2 dx^3 \\
&\int_{\partial V} n_i T^{i3} dS = 0
\end{aligned} \tag{26}$$

Note that the expression of T^{11} and T^{22} on the arms must be used on the left hand side of eq. (26), that is:

$$\begin{aligned}
T^{11} &= \frac{1}{2}\left[c^2|\Phi^0|^2 + |\Phi^1|^2 - |\Phi^2|^2 - |\Phi^3|^2\right] \\
T^{22} &= \frac{1}{2}\left[c^2|\Phi^0|^2 - |\Phi^1|^2 + |\Phi^2|^2 - |\Phi^3|^2\right]
\end{aligned} \tag{27}$$

which are positive.

In other words, when the stress-energy tensor crosses the arm Ox^2 of the angle, it receives a positive acceleration proportional to T^{22} in the x^1 direction; and when it crosses the arm Ox^1 of the angle, it receives a positive acceleration proportional to T^{11} in the x^2 direction. However, it does not correspond to a deviation of the direction of propagation since the space is flat everywhere but on the hinge. It only takes into account the change of coordinates across the arms of the angle.

5.3 Ray scattering

Any ray issued from a sound source will have some thickness. This is simply due to the uncertainty principle (see for example [Ste10]), which states that position and direction cannot be both determined with infinite precision. Note that this is not the case with the geometrical constructions of the previous Sections, where both are simultaneously defined with infinite precision.

According to Hadamard [Had03], waves are discontinuities that move through space. These discontinuities can be infinitesimal, and they are defined by a function $f(x^0, x^1, x^2, x^3) = 0$ on the time-space variables. As a consequence, the velocity potential Φ can be expressed as a function of f, and the wave can be defined by the equation $\left(\Phi^{i*} g_{ij} \Phi^j\right) = 0$, or equivalently $\left(f^{i*} g_{ij} f^j\right) = 0$.

We now consider a wave packet, that is, a wave of finite extension, both along the direction of propagation defined by the vector $\nabla^i f$, and laterally along $f(x^i) = 0$. The uncertainty principle then states that frequency and wave numbers are defined as distribution, the standard deviations of which verify $\Delta x^i \Delta k_i \geqslant \frac{1}{2}$, with $k_0 = \omega = 2\pi f$ the radian frequency.

We then introduce new coordinates ξ^i, defined by $\xi^0 = cx^0$, $\xi^1 = f(x^i)$ (along $\nabla^i f$), ξ^2 and ξ^3 on the surfaces $f(x^i) = 0$ at time ξ^0. As distance must be independent of coordinate systems:

$$g_{ij} dx^i dx^j = \gamma_{kl} d\xi^k d\xi^l = \gamma_{kl} \frac{\partial \xi^k}{\partial x^i} \frac{\partial \xi^l}{\partial x^j} dx^i dx^j$$

that is, $g_{ij} = \gamma_{kl} \frac{\partial \xi^k}{\partial x^i} \frac{\partial \xi^l}{\partial x^j}$, where γ_{kl} is the metric tensor associated with the new coordinates ξ^i. As a consequence, $|g| = |\gamma| |\frac{\partial \xi^k}{\partial x^i}|^2$, where g, γ and $|\frac{\partial \xi^k}{\partial x^i}|$ are determinants. We can further write that $d\xi^1 = c_i dx^i$, with $c_0 = -c$ and c_i, $i = 1, 2, 3$ the direction cosines of the wave packet. The wave packet is therefore completely defined by its distribution on any hyper-surface that crosses its trajectory. On the other hand, on any hyper-surface parallel to the trajectory, the distribution will depend on the distance to the trajectory.

We now can apply the stress-energy conservation eq. (26) to our wave packet. We choose for integration a time interval over which the wave packet moves from one side of the angle arms to the other. Two cases are depicted in Fig. 11.

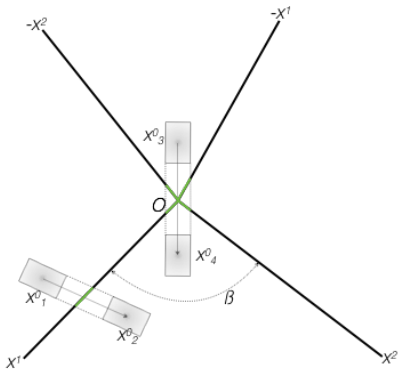

Figure 11: Wave packets crossing angle arms. Packet on the left crosses arm Ox^1 between x_1^0 and x_2^0; packet in the middle symmetrically crosses the two arms. Green lines mark the areas where deviations occur.

The first wave packet is located in sector $(-x^2)Ox^1$ at time x_1^0. It is here defined by the projection of its distribution on the plane of Fig. 11, where it is represented by a rectangle area with levels of grey proportional to the probability of presence of the wave packet. The size of the rectangle is chosen such that the components of the stress-energy tensor are negligible on its boundaries. As time increases, the wave packet moves at angle θ with respect to direction x^2 toward sector x^1Ox^2 in which it is located at time x_2^0. In the sector $(-x^2)Ox^1$,

at time x_1^0, we have therefore:

$$\begin{aligned}\left(\Phi^{i*} g_{ij} \Phi^j\right) &= \left[-c^2|\Phi^0|^2 + |\Phi^1|^2 + |\Phi^2|^2 + |\Phi^3|^2 - 2\cos\beta\Re(\Phi^1\Phi^{2*})\right] = 0\\ T^{00} &= \frac{1}{2c^2}\left[c^2|\Phi^0|^2 + |\Phi^1|^2 + |\Phi^2|^2 + |\Phi^3|^2 - 2\cos\beta\Re(\Phi^1\Phi^{2*})\right]\\ T^{01} &= \Re(\Phi^0\Phi^{1*}),\ T^{02} = \Re(\Phi^0\Phi^{2*}),\ T^{03} = \Re(\Phi^0\Phi^{3*})\end{aligned}$$

and these components remain constant as long as the wave packet remains in sector $(-x^2)Ox^1$. Note that the first equation, which expresses that Φ corresponds to a wave packet, can be rewritten as:

$$c^2|\Phi^0|^2 = |\Phi^1|^2 + |\Phi^2|^2 + |\Phi^3|^2 - 2\cos\beta\Re(\Phi^1\Phi^{2*}) \tag{28}$$

It crosses arm Ox^1 ($x^2 = 0$) in-between times x_1^0 and x_2^0, where it is deviated according to the second line of eq. (26). In other words, the components of T^{ij} do not change, but for component T^{01} which becomes, according to eq. (26):

$$\begin{aligned}\int_{\partial V_{x_2^0}} n_0 T^{01} dS = -2\cos\beta \int_V \mathrm{sgn}(x^1)\delta(x^2) T^{22}\sqrt{|g|}dx^0dx^1dx^2dx^3\\ + \int_{\partial V_{x_1^0}} n_0 T^{01} dS\end{aligned} \tag{29}$$

where $\partial V_{x_1^0}$, resp. $\partial V_{x_2^0}$ is the boundary at time x_1^0, resp. x_2^0, that is, the volumes of the wave packet at x_1^0 and x_2^0, and the components of T^{ij} are negligible on the boundaries of the wave packet. Simple calculations then shows that $n_0 = -c$ and $dS = \sin\beta dx^1dx^2dx^3$ on $\partial V_{x_1^0}$ and $\partial V_{x_2^0}$, with $dV = c\sin\beta dx^0dx^1dx^2dx^3$ which reduces to $dV_{x^2=0} = cdx^0dx^1dx^2dx^3$ on the arm $x^2 = 0$.

Similarly, on the arm $x^2 = 0$, T^{22} is given by eq. (27), that is, with the help of eq. (28):

$$T^{22} = |\Phi^2|^2 - \cos\beta\Re(\Phi^1\Phi^{2*})$$

According to eq. (21), we have $\Phi^i = -v^i$, so that the previous equation becomes:

$$T^{22} = |v^2|^2 - \cos\beta\Re(v^1v^{2*})$$

with the v^i represented in Fig. 12, except for v^3 which is perpendicular to the plane of the figure. Also note that the figure presents the projections of cv_1^0 and cv_2^0 on the plane Ox^1x^2, respectively called v_1 and v_2.

With the help of Fig. 12, it can easily be seen that eq. (28) is equivalent to the vector equation $c\vec{v^0} = \vec{v^1} + \vec{v^2} + \vec{v^3}$, that is, by projecting on v^2:

$$c\vec{v^0}\cdot\vec{v^{2*}} = \vec{v^1}\cdot\vec{v^{2*}} + \vec{v^2}\cdot\vec{v^{2*}} + 0 = cv^0v^{2*}\cos\theta$$

with $\vec{v^1}\cdot\vec{v^{2*}} = v^1v^{2*}\cos(\pi-\beta)$. We therefore recover:

$$\begin{aligned}c\Re(\vec{v^0}\cdot\vec{v^{2*}}) &= |v^2|^2 - \Re(v^1v^{2*})\cos\beta\\ &= |\Phi^2|^2 - \cos\beta\Re(\Phi^1v^{\Phi*}) = c\cos\theta\Re(\Phi^0\Phi^{2*})\end{aligned}$$

where $\cos\theta = c_2$ is the direction cosine of vector $\vec{v^0}$ with respect to axis Ox^2. In Fig. 12, it corresponds to the projection of c_2 on the plane Ox^1x^2. All in all, we obtain on the arm $x^2 = 0$:

$$T^{22} = c_2\Re(c\Phi^0\Phi^{2*}) \tag{30}$$

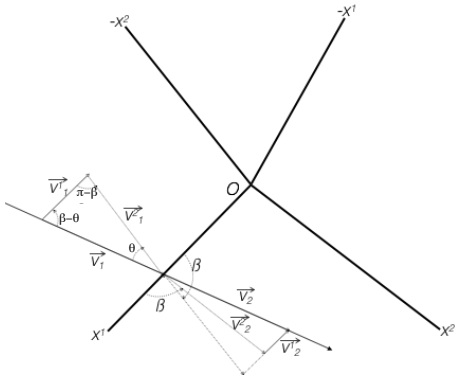

Figure 12: Wave packets travelling at angle δ with respect to arm Ox^2. v_j^i are the components of the particle velocity vector, with v_1 and v_2 the projections of cv_1^0 and cv_2^0 on the plane Ox^1x^2. Note that $\theta < 0$ since v_j^1 are negative.

We must carry out the integration $\int_V \mathrm{sgn}(x^1)\delta(x^2)T^{22}\sqrt{|g|}dx^0dx^1dx^2dx^3$ on arm Ox^2. Taking into account that $\frac{\partial\xi^1}{\partial x^2} = c_2$, we obtain $\delta(x^2) = \frac{\delta(\xi^1-\xi)}{c_2}$, where ξ is the value taken by ξ^1 on the axis Ox^2. Then using the relation between old and new coordinates x^i and ξ^i, we obtain:

$$\begin{aligned}
&\int_V \mathrm{sgn}(x^1)\delta(x^2)T^{22}\sqrt{|g|}dx^0dx^1dx^2dx^3 \\
&\qquad = \int_V \frac{\delta(\xi^1-\xi)}{c_2} c_2 \Re(c\Phi^0\Phi^{2*})\sqrt{|\gamma|}d\xi^0d\xi^1d\xi^2d\xi^3 \\
&\qquad = \int_V \Re(c\Phi^0\Phi^{2*})\sqrt{|\gamma|}d\xi^0d\xi^2d\xi^3
\end{aligned}$$

and with the help of eq. (29):

$$\begin{aligned}
-c\overline{T^{01}}_{\partial V_{x_2^0}} &= \int_{\partial V_{x_2^0}} n_0 T^{01} dS \\
&= -2\cos\beta \int_V \mathrm{sgn}(x^1)\delta(x^2)T^{22}\sqrt{|g|}dx^0dx^1dx^2dx^3 + \int_{\partial V_{x_1^0}} n_0 T^{01} dS \\
&= -2\cos\beta \int_{\partial V_{x^2=0}} \Re(c\Phi^0\Phi^{2*})\sqrt{|\gamma|}d\xi^0d\xi^2d\xi^3 - c\overline{T^{01}}_{\partial V_{x_1^0}} \\
&= -2c\cos\beta\overline{\Re(\Phi^0\Phi^{2*})}_{x^2} - c\overline{T^{01}}_{\partial V_{x_1^0}}
\end{aligned}$$

or simply:

$$\overline{T^{01}}_{\partial V_{x_2^0}} = \overline{T^{01}}_{\partial V_{x_1^0}} + 2\cos\beta\overline{\Re(\Phi^0\Phi^{2*})}_{x^2} \tag{31}$$

In other word, Φ^1 is "*augmented*" with $2\cos\beta\Phi^2$, as is visible in Fig. 12 where $\cos\beta < 0$.

This is not the case for the second wave packet, located in sector $(-x^1)O(-x^2)$ at time x_3^0 and moving to sector x^1Ox^2 at time x_4^0. It symmetrically crosses the arms $x^1 = 0$ and $x^2 = 0$ on its way. Since the function sgn changes sign around

the origin O, the negative deviation compensates for the positive one on each axis, and no deviations occurs. However, if a larger part of the packet passes on one side of the origin, some deviation occurs in proportion of the offset. This case is not represented in Fig. 11.

In both cases, direct calculation shows that T^{00} does not change when crossing arm Ox^1, as expected for energy conservation.

Note that, in the case of a wave packet impinging at near grazing incidence on a reflex angle, the above formalism recovers the sound particle diffraction model of [Ste10]. However, it takes into account the angle of the scattering wedge and the scattering of the image wave packet, both of which are absent from the sound particle diffraction model.

6 Conclusion

We have presented a geometrical theory that naturally accounts for scattering on the boundaries of a room. It introduces Riemannian spaces with negative curvature, which constitute the proper setting for the distribution of images created by non-rectangular rooms with obtuse angles, that is, created by irregular polyhedra. The crucial factor is the excess angle that arises around specific edges, called hinges, when first and second order images are considered, as it pilots the metric tensor of the space and all its geometrical properties, including its curvature. In the case of reflex and acute dihedral angles around edges, less, resp. more, reflection orders must be taken into account to properly tessellate the Riemannian space.

From this Riemannian tessellation, we have proposed a scheme for counting the number of image sources. Here, the parameter is not the order of reflection, but counting the layers of images around the original room. Only free faces, edges and vertices are taken into account to build the layers, and it makes it possible to give a close form formula for the number of image sources in case all dihedral angles are obtuse: the number of images increases exponentially, making polyhedral rooms similar to mixing rooms in this respect. We did not explicitly solved for the cases when some dihedral angles are reflex or acute, but gave some indications as how to handle them.

In the case of regular polygons and polyhedra, we have also shown that image sources are regularly distributed along channels delimited by repetitive distributions of hinges and becoming narrower when the repetitive patterns of image rooms become longer. In the case of irregularly shaped rooms, the channels split randomly since the distribution of hinges is also random. In both cases, we have defined the horizon of visibility as the distance at which the width of the channel becomes narrower than the first Fresnel zone, thus creating an exponential decrease of the image sources with distance.

Lastly, using the curvature on the hinges and complementing it with the uncertainty principle, we were able to describe the scattering of wave packets around dihedral angles. The scattering is proportional to the excess angle, and is best described in terms of the stress-energy tensor, that is, in terms of energy conservation. The basic elements for computing the scattering are given, and must be adapted to each case at hand in order to derive formulae.

The present theory must now be developed to derive scattering coefficients from the distribution of hinges around a room. Most certainly, the excess angles

are the main factors, complemented by the lengths of the hinges. But wave length also plays a rôle, as demonstrated in Sections 4 and 5.3. Proper definition of scattering coefficients has long been missing for the application of the diffusion equation in Room Acoustics [PPS97], despite some recent attempts [GD̈21].

There remain to introduce absorption in the present theory. Due to its geometrical nature, the present theory cannot account for losses in its present form. The classical way to introduce absorption, which consider intensity flows inside the boundaries, cannot be used here without some adaptation. But accounting for losses should not proove difficult since absorption is easily described with the stress-energy tensor used in the last Section, as shown in [DPTP17, DPTP18].

Appendix: factorization of matrix Λ

Remembering that Λ is given by:

$$\Lambda = \begin{pmatrix} \frac{(N-2)(N-6)}{N} + 1 & N-4 & N-3 \\ \frac{3(N-2)(N-4)(N-6)}{N^2} & \frac{3(N-4)^2}{N} - 1 & \frac{3(N-3)(N-4)}{N} \\ \frac{2(N-2)(N-6)^2}{N^2} & \frac{2(N-4)(N-6)}{N} & \frac{2(N-3)(N-6)}{N} + 1 \end{pmatrix}$$

it is easy to see that, if one subtracts the eigenvalue 1 from all its line, the first and last line become proportional with a factor $\frac{2(N-6)}{N}$. Then subtracting $\frac{3(N-4)}{N}$ time the first line from the second one, the second coordinate of the corresponding eigenvector must be null. The fist line then give a relation between the first and the last coordinates, from which the first eigenvector is derived:

$$X_0 = \begin{pmatrix} (N-3) \\ 0 \\ -\frac{(N-2)(N-6)}{N} \end{pmatrix}$$

Using the form $\left[\frac{3(N-4)^2}{N} - 1\right] \pm \frac{N-4}{N}\sqrt{3[(N-6)(3N-8)]}$ for the two other eigenvalues, they are easily subtracted from the diagonal terms of matrix Λ. Then subtracting $\frac{2(N-6)}{N}$ times the first line from the last one gives a relation between the first and the last coordinates: the last coordinates is equal to $\frac{2(N-6)}{N}$ times the first one. Introducing this relationship into the second line gives a first coordinate proportional to $\sqrt{3[(N-6)(3N-8)]}$, from which the other coordinates are obtained. In the end, the corresponding eigenvectors is given by:

$$X = \begin{pmatrix} \sqrt{3[(N-6)(3N-8)]} \\ \pm\frac{3(N-6)(3N-8)}{N} \\ \frac{2(N-6)}{N}\sqrt{3[(N-6)(3N-8)]} \end{pmatrix}$$

with sign + on the second line corresponding to λ_1 and sign − to λ_2.

References

[CM78] L. Cremer and H.A. Müller. *Die wissenschaftlichen Grundlagen der Raumakustik*, volume Band I, page 25. Hirzel Verlag, 1978.

[DPTP17] H. Dujourdy, B. Pialot, T. Toulemonde, and J.D. Polack. An energetic wave equation for modelling diffuse sound fields – application to corridors. *Acta Acustica*, 103:480–491, 2017.

[DPTP18] H. Dujourdy, B. Pialot, T. Toulemonde, and J.D. Polack. An energetic wave equation for modelling diffuse sound fields – application to open offices. *Wave Motion*, 2018.

[GÜ21] Z.S. Gül. Exploration of room acoustics coupling in Hagia Sophia of Istanbul for its different states. *J. Acoust. Soc. Am.*, 149(1):320–339, January 2021.

[Had03] J. Hadamard. *Leçons sur la propagation des ondes et les équations de l'hydrodynamique*. Librairie Scientifique A. Hermann, Paris, 1903.

[Lin 5] B. Linet. Notes de cours de relativité générale, 2004-5.

[NE93] G. Naylor Ed. Special issue on computer modelling and auralisation of sound fields in rooms. *Applied Acoustics*, 38(2-4), 1993.

[Pol92] J.D. Polack. Modifying chambers to play billiards: the foundations of reverberation theory. *Acustica*, 76:257–272, 1992.

[PPS97] J. Picaut, J.D. Polack, and L. Simon. A mathematical diffuse field model based on a diffusion equation. *Acustica united with Acta Acustica*, 83:614–621, 1997.

[Reg61] T. Regge. General Relativity without Coordinates. *Il Nuovo Cimento*, XIX:558–571, Feb. 1961.

[Ste10] U.M. Stephenson. An energetic Approach for the Simulation of Diffraction within Ray Tracing Based on the Uncertainty Relation. *Acta Acustica united with Acustica*, 96:516–535, 2010.

Appendix B

Code listing

B.1 Introduction

This section serves as a reference for portions of code developed over the course of the thesis. Most works can be found on Github under permissive licenses.

B.2 FVTD

`https://github.com/1ceaham/AcousticFVTD_GeneralImpedance/`

B.3 Remote code execution

`https://github.com/1ceaham/sendToRemote`

Bibliography

Abel, J. S. and Huang, P. (2006). A Simple, Robust Measure of Reverberation Echo Density. In *Audio Engineering Society Convention 121*. Audio Engineering Society.

Alary, B., Politis, A., Schlecht, S., and Välimäki, V. (2019). Directional Feedback Delay Network. *Journal of the Audio Engineering Society*, 67(10):752–762.

Allen, J. B. and Berkley, D. A. (1979). Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America*, 65(4):943–950.

Badeau, R. (2019). Common mathematical framework for stochastic reverberation models. *The Journal of the Acoustical Society of America*, 145(4):2733–2745.

Bilbao, S. and Ahrens, J. (2020). Modeling continuous source distributions in wave-based virtual acoustics. *The Journal of the Acoustical Society of America*, 148(6):3951–3962.

Bilbao, S., Hamilton, B., Botts, J., and Savioja, L. (2016). Finite Volume Time Domain Room Acoustics Simulation under General Impedance Boundary Conditions. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(1):161–173.

Bunimovich, L. A. (1979). On the ergodic properties of nowhere dispersing billiards. *Communications in Mathematical Physics*, 65(3):295–312.

Cox, T. and D'Antonio, P. (2016). *Acoustic Absorbers and Diffusers: Theory, Design and Application*. CRC Press, Boca Raton, third edition.

Cox, T. J., Dalenback, B.-I. L., D'Antonio, P., Embrechts, J. J., Jeon, J. Y., Mommertz, E., and Vorländer, M. (2006). A Tutorial on Scattering and Diffusion Coefficients for Room Acoustic Surfaces. *Acta Acustica united with Acustica*, 92(1):1–15.

Defrance, G., Daudet, L., and Polack, J.-D. (2009). Using Matching Pursuit for Estimating Mixing Time Within Room Impulse Responses. *Acta Acustica united with Acustica*, 95(6):1071–1081.

Dujourdy, H., Pialot, B., Toulemonde, T., and Polack, J.-D. (2017). An Energetic Wave Equation for Modelling Diffuse Sound Fields – Application to Corridors. *Acta Acustica united with Acustica*, 103(3):480–491.

Dujourdy, H., Pialot, B., Toulemonde, T., and Polack, J.-D. (2019). An energetic wave equation for modelling diffuse sound fields—Application to open offices. *Wave Motion*, 87:193–212.

Duval, H. (2020). Modélisation Acoustique du Caveau Phonocamptique de la Cathédrale de Noyon. Master's thesis, Sorbonne Université.

Ewins, D. (2000). *Modal Testing: Theory, Practice and Application*. Engineering Dynamics Series. Wiley.

Farina, A. (2000). Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. In *Audio Engineering Society Convention 108*. Audio Engineering Society.

Farina, A. (2007). Advancements in Impulse Response Measurements by Sine Sweeps. In *Audio Engineering Society Convention 122*. Audio Engineering Society.

Fratoni, G. . (2021). *Standardization of Procedures and Calculation Models for the Numerical Simulation of Acoustics of Enclosed Spaces*. Doctoral Thesis, Alma Mater Studiorum - Università di Bologna.

Gribonval, R., Bacry, E., Mallat, S., Depalle, P., and Rodet, X. (1996). Analysis of Sound Signals with High Resolution Matching Pursuit. In *Proc. IEEE Symp. Time-Freq. and Time-Scale Anal. (TFTS'96)*, pages 125–128, Paris, France. IEEE.

Hamilton, B. (2021). PFFDTD.

Hamilton, B., Webb, C. J., Fletcher, N. D., and Bilbao, S. (2016). 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*, page 11.

Huang, P. and Abel, J. S. (2007). Aspects of Reverberation Echo Density. In *Audio Engineering Society Convention 123*. Audio Engineering Society.

Jing, Y. and Xiang, N. (2007). A modified diffusion equation for room-acoustic predication. *The Journal of the Acoustical Society of America*, 121(6):3284–3287.

Jot, J.-M., Cerveau, L., and Warusfel, O. (1997). Analysis and Synthesis of Room Reverberation Based on a Statistical Time-Frequency Model. In *Audio Engineering Society Convention 103*. Audio Engineering Society.

Kleiner, M., Dalenbäck, B.-I., and Svensson, P. (1993). Auralization-An Overview. *Journal of the Audio Engineering Society*, 41(11):861–875.

Kuttruff, H. (2016). *Room Acoustics*. CRC Press.

Lai, H. and Hamilton, B. (2020). Computer Modeling of Barrel-Vaulted Sanctuary Exhibiting Flutter Echo with Comparison to Measurements. *Acoustics*, 2(1):87–109.

Le Bot, A. and Bocquillet, A. (2000). Comparison of an integral equation on energy and the ray-tracing technique in room acoustics. *The Journal of the Acoustical Society of America*, 108(4):1732–1740.

Lindau, A., Kosanke, L., and Weinzierl, S. (2010). Perceptual evaluation of physical predictors of the mixing time in binaural room impulse responses. In *Audio Engineering Society Convention 128*, page 17.

Luizard, P., Polack, J.-D., and Katz, B. F. G. (2013). Auralization of coupled spaces based on a diffusion equation model. In *Proc. Sound and Music Computing Conference (SMAC/SMC)*, pages 616–621, Stockholm, Sweden.

Martin, S. R. and Svensson, U. P. (2018). Modeling Sound Sources with Non-Convex Shapes Using an Edge Diffraction Approach. In *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*. Audio Engineering Society.

Meacham, A., Badeau, R., and Polack, J.-D. (2019a). Implementation of Sources in an Energy-Stress Tensor Based Diffuse Sound Field Model. In *Proceedings of the International Symposium on Room Acoustics*, Amsterdam.

Meacham, A., Badeau, R., and Polack, J.-D. (2019b). Lower Bound on Frequency Validity of Energy-Stress Tensor Based Diffuse Sound Field Model. In *Proceedings of the 23rd International Congress on Acoustics*, Aachen.

Meacham, A., Badeau, R., and Polack, J.-D. (2020). Auralization of a Hybrid Sound Field using a Wave-Stress Tensor Based Model. In *Forum Acusticum*, page 523.

Merimaa, J. and Pulkki, V. (2005). Spatial Impulse Response Rendering I: Analysis and Synthesis. *Journal of the Audio Engineering Society*, 53(12):1115–1127.

Moorer, J. A. (1979). About This Reverberation Business. *Computer Music Journal*, 3(2):13–28.

Morse, P. M. and Feshbach, H. (1953). *Methods of Theoretical Physics*. Mc Graw-Hill Book Company.

Morse, P. M. and Ingard, K. U. (1968). *Theoretical Acoustics*. Princeton University Press.

Murphy, D., Shelley, S., Beeson, M., Moore, A., and Southern, A. (2008). Hybrid room impulse response synthesis in digital waveguide mesh based room acoustics simulations. In *Proceedings of the 11th International Conference on Digital Audio Effects*.

Nolan, M., Verburg, S. A., Brunskog, J., and Fernandez-Grande, E. (2019). Experimental characterization of the sound field in a reverberation room. *The Journal of the Acoustical Society of America*, 145(4):2237–2246.

Ollendorff, F. (1969). Statistical Room-Acoustics as a Problem of Diffusion (A Proposal). *Acta Acustica united with Acustica*, 21(4):236–245.

Oxnard, S. and Murphy, D. (2013). Room impulse response synthesis based on a 2D multi-plane FDTD hybrid acoustic model. In *2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 1–4, New Paltz, NY, USA. IEEE.

Picaut, J., Simon, L., and Polack, J.-D. (1997). A Mathematical Model of Diffuse Sound Field Based on a Diffusion Equation. *Acta Acustica united with Acustica*, 83(4):614–621.

Poirier-Quinot, D., Katz, B., and Noisternig, M. (2017). EVERTims: Open source framework for real-time auralization in architectural acoustics and virtual reality. In *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx 17)*, Edinburgh.

Polack, J.-D. (1992). Modifying Chambers to play Billiards: The Foundations of Reverberation Theory. *Acta Acustica united with Acustica*, 76(6):256–272.

Polack, J.-D. (1993). Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics. *Applied Acoustics*, 38(2):235–244.

Polack, J.-D. (2020). Courbure des rayons sonores, tenseur impulsion-energie, et calcul de Regge en acoustique des salles.

Polack, J.-D., Alrutz, H., and Schroeder, M. R. (1984a). The Modulation Transfer Function of Music Signals and its Applications to Reverberation Measurement. *Acta Acustica united with Acustica*, 54(5):257–265.

Polack, J.-D., Alrutz, H., and Schroeder, M. R. (1984b). Une nouvelle méthode pour décrire la réverbération. *C.R. Acad. Sc. Paris*, pages 21–24.

Polack, J.-D., Meacham, A., and Badeau, R. (2019). Generalized formulation of acoustics. In *Proceedings of the Congrès Français de Mécanique*.

Polack, J.-D., Meacham, A., Badeau, R., and Valiere, J.-C. (2021). Riemannian space tessellation with polyhedral room images. Technical report.

Pulkki, V. and Merimaa, J. (2006). Spatial Impulse Response Rendering II: Reproduction of Diffuse Sound and Listening Tests. *Journal of the Audio Engineering Society*, 54(1/2):3–20.

Reilly, A. and McGrath, D. (1995). Convolution Processing for Realistic Reverberation. In *Audio Engineering Society Convention 98*. Audio Engineering Society.

Romblom, D., Guastavino, C., and Depalle, P. (2016). Perceptual thresholds for non-ideal diffuse field reverberation. *The Journal of the Acoustical Society of America*, 140(5):3908–3916.

Savioja, L. and Svensson, U. P. (2015). Overview of geometrical room acoustic modeling techniques. *The Journal of the Acoustical Society of America*, 138(2):708–730.

Schiffrer, G. and Stanzial, D. (1994). Energetic properties of acoustic fields. *The Journal of the Acoustical Society of America*, 96(6):3645–3653.

Schröder, D. (2011). *Physically Based Real-Time Auralization of Interactive Virtual Environments*. PhD thesis, RWTH Aachen, Berlin.

Schröder, D. and Vorländer, M. (2011). RAVEN: A real-time framework for the auralization of interactive virtual environments. In *Proceedings of Forum Acusticum*, pages 1541–1546.

Schroeder, M., Atal, B., and Bird, C. (1962). Digital computers in room acoustics. In *Proceedings of the 4th International Congress on Acoustics*, volume 21, Copenhagen.

Schroeder, M. R. (1965). New Method of Measuring Reverberation Time. *The Journal of the Acoustical Society of America*, 37(3):409–412.

Schroeder, M. R. (1975). Diffuse sound reflection by maximum-length sequences. *The Journal of the Acoustical Society of America*, 57(1):149–150.

Schroeder, M. R. (1996). The "Schroeder frequency" revisited. *The Journal of the Acoustical Society of America*, 99(5):3240–3241.

Southern, A., Lokki, T., Savioja, L., and Murphy, D. (2011a). The Perceptual Effects of Dispersion Error on Room Acoustic Model Auralization. In *Proceedings of Forum Acusticum*, pages 1553–1558, Aalborg.

Southern, A., Siltanen, S., Murphy, D. T., and Savioja, L. (2013). Room Impulse Response Synthesis and Validation Using a Hybrid Acoustic Model. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(9):1940–1952.

Southern, A., Siltanen, S., and Savioja, L. (2011b). Spatial Room Impulse Responses with a Hybrid Modeling Method. In *Audio Engineering Society Convention 130*. Audio Engineering Society.

Sü Gül, Z., Odabaş, E., Xiang, N., and Çalışkan, M. (2019). Diffusion equation modeling for sound energy flow analysis in multi domain structures. *The Journal of the Acoustical Society of America*, 145(4):2703–2717.

Valière, J.-C., Palazzo-Bertholon, B., Polack, J.-D., and Carvalho, P. (2013). Acoustic Pots in Ancient and Medieval Buildings: Literary Analysis of Ancient Texts and Comparison with Recent Observations in French Churches. *Acta Acustica united with Acustica*, 99(1):70–81.