Title
Spatial Audio Fields

Author
Weitzner, Nathan

Publication Date
2022

Peer reviewed|Thesis/dissertation

University of California
Santa Barbara

Spatial Audio Fields

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Media Arts and Technology

by

Nathan David Weitzner

Committee in charge:

Dr. Andrés Cabrera, Co-Chair
Professor Curtis Roads, Co-Chair
Professor JoAnn Kuchera-Morin

March 2022

The Dissertation of Nathan David Weitzner is approved.

JoAnn Kuchera-Morin

Curtis Roads, Committee Co-Chair

Andrés Cabrera, Committee Co-Chair

March 2022

Spatial Audio Fields

To my family.

Acknowledgments

I would like to take this opportunity to acknowledge the many contributions of numerous people to this dissertation. I am incredibly grateful to have Dr. Cabrera, Professor Roads, and Professor Kuchera-Morin as my committee. Their knowledge, expertise, and work has long been a point of inspiration. I want to thank Dr. Cabrera in particular, not only for his role as co-chair of my committee, but also for his role as my advisor. He has provided an extraordinary amount of guidance and I give him my most sincere thanks. I also want to thank the faculty, staff, students, and other members of the Media Arts and Technology community. Finally, I would like to thank the people that helped proofread this dissertation: Tori Prado, Larry Weitzner, Anna Curtis, and Candice Frodeman.

Curriculum Vitæ
Nathan David Weitzner

Education

2022	Ph.D. in Media Arts and Technology (Expected), University of California, Santa Barbara
2012	M.S. in Music Technology, Georgia Institute of Technology
2009	B.A. in Music, Furman University

Experience

2015 - 2021	Teaching Assistant, Media Arts and Technology, University of California, Santa Barbara
2018	Virtual Reality Intern, Unite To Light, WorldViz, Santa Barbara
2015	Student Assistant, Media Arts and Technology, University of California, Santa Barbara
2012	Music Technology Consultant, Atlanta based startup company, Atlanta
2012	Graduate Representative, Student Music Advisory Board, Georgia Institute of Technology, Atlanta
2011 - 2012	Teaching Assistant, Survey of Music Technology (summer terms), Georgia Institute of Technology, Atlanta
2011 - 2012	Graduate Student Researcher, Georgia Tech Center for Music Technology, Georgia Institute of Technology, Atlanta

Awards

2013	Department Grant, Media Arts and Technology, University of California, Santa Barbara
2006	Sonya Weber Scholarship, Department of Music, Furman University, Greenville
2004	Strings Scholarship, Department of Music, Furman University, Greenville

Publications

Sajadieh, S., and Weitzner, N. “A Theatrical Turing Test: An Encounter of Telematic Bodies.” In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 297-300. 2016.

Freeman, J., Xie, S., Tsuchiya, T., Shen, W., Chen, Y. L., and Weitzner, N. “Using massMobile, a flexible, scalable, rapid prototyping audience participation framework, in large-scale live musical performances.” Digital Creativity 26, no. 3-4 (2015): 228-244.

Wu, J., Weitzner, N., Yeh, Y., Abel, J., Michon, R., and Wright, M. “Tibetan singing prayer wheel: a hybrid musical-spiritual instrument using gestural control.” In International Conference on New Interfaces and Musical Expression (NIME-15). 2015.

Weitzner, N., Freeman, J., Chen, Y. L., and Garrett, S. “massMobile: towards a flexible framework for large-scale participatory collaborations in live performances.” Organised Sound 18, no. 1 (2013): 30-42.

Weitzner, N., Freeman, J., Garrett, S., and Chen, Y. L. “massMobile-an Audience Participation Framework.” In NIME, vol. 12, pp. 21-23. 2012.

Abstract

Spatial Audio Fields

by

Nathan David Weitzner

When a sound source is spatialized over loudspeakers, the source undergoes a series of transformations before it ultimately reaches the listener. The source is first transformed by panning and other spatialization algorithms based on its spatial parameters in what this dissertation calls the *virtual field*. Next, the source is projected by the *loudspeaker array* into the *acoustic field* further transforming the source. The virtual field, loudspeaker array, and acoustic field form the *spatial audio field* (SAF) which significantly influences how a spatial composition is perceived.

This dissertation presents the theory of spatial audio fields and culmination of research and development of new compositional techniques based on manipulation of loudspeaker gain curves in the virtual field. The majority of current spatial techniques focus on manipulation of a source's spatial parameters and parameters of the spatial image. While sources in spatial audio compositions are assigned positions and trajectories that often change throughout the work, the spatialization algorithms of the virtual field are usually fixed, the loudspeaker array does not change, and the acoustic field rarely changes during the performance. Thus the specific SAF used in a given work, whether constructed deliberately as part of the compositional process or constructed arbitrarily, is usually static.

The SAF is largely unexplored for new compositional techniques on short musical timescales. As the loudspeakers and acoustic field are physical components of the SAF, changes to each are only practical on larger timescales from musical phrases to form. How-

ever, as the virtual field is usually constructed digitally, it can be changed on timescales down to the sample rate. Therefore, the focus of this research is on developing compositional techniques in the virtual field.

To carry out this research, an experimental software program was developed that not only allows for realtime spatialization of multiple audio sources, but also allows for realtime changes to the virtual field. This research has led to the development of new compositional elements, methods of synthesis, and spatial extensions to traditional compositional techniques. These methods and techniques are not limited to spatial audio using loudspeakers in a space, but can also be used in spatial audio over headphones where the acoustic field is almost entirely simulated. This opens up possibilities for developing new compositional techniques based on changes to the simulated acoustic field on shorter musical timescales in conjunction with the techniques in the virtual field developed in this research.

Contents

Chapter 1

Introduction

1.1 Spatial Audio Fields

The *spatial audio field* (SAF) consists of virtual and acoustic subfields where the effects of the virtual field on the source are projected into the acoustic field by the loudspeakers. The "field" in spatial audio fields is analogous to a field in physics where each point in space has a specific value or set of values. The input to the SAF is the source waveform and spatial parameters (such as position) and the output of the field is sound that is perceived by the listener. The SAF represents a series of transformations to the source from the virtual field, loudspeakers, and acoustic field.

Virtual Field

The virtual field of the SAF primarily consists of loudspeaker gain values that are typically generated by a specific panning algorithm. Other values besides gain can be defined, such as effect parameters, to further shape the virtual field. In most spatial compositions, the virtual field (especially the panning algorithm) is static. However, changing the virtual field over time opens up several compositional possibilities.

Loudspeakers

The source having been subjected to the influence of the virtual field and algorithms next is projected into the acoustic field through the loudspeakers. Individual loudspeaker properties such as position, orientation, frequency response, and directionality further transform the source. Properties of the loudspeaker array such as density, configuration, and dimensionality determine how many wavefronts representing a single source are projected into the field and the originating position of these wavefronts. Unlike the virtual field, the physical loudspeaker properties are difficult to change on short timescales.

Acoustic Field

The final field is the acoustic field in which the loudspeakers project the source's audio after having been transformed by the virtual field and loudspeakers. The acoustic field further transforms the source based on the materials of the sound propagating medium, materials of the boundary, and the enclosure of the boundary. Like the loudspeaker field, the acoustic field is also difficult to change on short timescales.

Research Scope

The purpose of defining the spatial audio field is to create a contextual framework to guide this research. While changes to the physical loudspeakers and acoustic field are discussed, the focus of this research is on the virtual field and specifically the loudspeaker gain curves of the virtual field.

1.2 Problem Statement

The spatial audio field (SAF) contains a virtual field, loudspeaker array, and acoustic field. The SAF, and especially the virtual field, are largely unexplored for new spatial compositional techniques on short musical timescales. While sources in spatial audio compositions are assigned positions and trajectories that change throughout the work, the spatialization algorithms are usually fixed, the loudspeaker array does not change, and the acoustic field rarely changes during the performance. Thus the specific SAF used in a given work, whether constructed deliberately as part of the compositional process or constructed arbitrarily, is usually static.

However, the SAF greatly influences how the listener ultimately perceives the work. The extent of this influence is apparent to anyone who has ever composed a spatial work in one space (such as a studio), and performed the work in another (such as a concert hall). The SAF is necessarily different in different spaces. As difference is a fundamental principal of composition, how can the SAF be changed and how can these changes be used as compositional elements?

Although the SAF has always existed in any spatial audio implementation, this research develops a theory of the spatial audio field in order to identify changes to the field that can be used compositionally. As the loudspeakers and acoustic field are physical components of the SAF, changes to each are only practical on larger timescales from musical phrases to form. However, when the virtual field is constructed digitally, it can be changed in short musical timescales down to the sample rate.

This dissertation presents the theory of spatial audio fields and the research and implementation focuses on the virtual field. The research has led to new methods of synthesis, spatial extensions to traditional compositional techniques, and the unification of separate panning algorithms. These methods and techniques are not limited to spatial audio using loudspeakers in a space, but can also be used in spatial audio over headphones where the acoustic field is almost entirely simulated. This opens up possibilities for developing new compositional techniques based on changes to the simulated acoustic field on shorter musical timescales in conjunction with the techniques in the virtual field developed in this research.

1.3 Dissertation Overview

Chapter 2 contextualizes the dissertation by giving a brief history of spatialization in music, examples of contemporary spaces used for spatial audio research, spatial audio techniques, and spatial audio compositions. Next, the theory of spatial audio fields is discussed in Chapter 3. The theory provides a general overview of the three components of the spatial audio field and discusses the different aspects of each component.

After the theory is presented, Chapter 4 discusses composing and modifying the virtual field. This includes methods of changing the density of the loudspeaker array in addition to changing individual loudspeaker gain curves without causing discontinuities in the output signal. The end of Chapter 4 generalizes the gain curve as a parameter curve of a spatial object in the virtual field. Several other types of spatial objects are also discussed.

Chapter 5 discusses the compositional applications of the methods introduced in Chapter 4. These applications include synthesis, rhythm, and spatial extensions to Phase Music. Synthesis methods include amplitude modulation, power modulation, granular, and pulsar synthesis. Rhythm is created by a source moving through gain curves that have been modified. Three spatial extensions to Phase Music are introduced and then discussed in terms of how changes to the virtual field can alter how the phasing process is perceived.

Chapter 6 discusses the original software created to develop and validate the methods and applications presented in the previous chapters. The software allows for realtime spatialization of virtual sources and realtime control of loudspeaker density and gain curves. This chapter also discusses current limitations and plans for future development of the software.

The last two chapters present the conclusion and future work. Chapter 7 concludes this dissertation by readdressing the problem statement and summarizing the findings of this research. Chapter 8 discusses future work that includes using the methods and techniques developed on longer timescales, applications to participatory works, and applications to spatial audio using headphones.

Finally, Appendix A includes a detailed discussion of the example loudspeaker array used throughout this dissertation. The graphs of the gain curves are also discussed in detail to make their meaning clear and to reduce redundant explanations in the body of the dissertation.

Chapter 2

Background

2.1 Spatialization in Music

While the loudspeaker era and computer era contributed to advancements in spatial audio, spatialization in music has roots that extend well before these periods. Polychoral music is written for two or more spatially separated choirs - *cori spezzati* - and was used by composers in the 16th century and earlier [1] [2]. Composers of orchestral works experimented with spatially separated orchestras or spatially separated sections of an orchestra sometimes located offstage. Mozart's *Serenade No. 6 for Orchestra in D major K. 239* (1776) [3] was composed for two orchestras that are thought to have been spatially separated in performance [4]. Berlioz's *Requiem, Op. 5* (1837) [5] has four offstage brass ensembles and Mahler's *Symphony No. 2* (1895) [6] is written for offstage brass ensembles as well. After these orchestral compositions, there is little documented use of spatial techniques in composition until the post World War II era [7, p. 452].

The use of loudspeakers enabled new forms of music such as *acousmatic music*, *musique concrete*, and *elektronische musik* in addition to the evolution of spatialization using loudspeakers which began with manual distribution of the audio signal, then started to include electronic control followed by computer control of the distribution. In the early 1950's, Pierre Schaeffer used a space potentiometer to distribute sound live to one of four loudspeakers [8]. In 1958, Varese's *Poeme Electronique* was presented in the Philips Pavillon at the Brussels world fair. The audio consisted of three tracks and was projected over subgroups of a 350-425 loudspeaker system [8] [9]. Salvatore Martirano's *Sal-Mar Construction* was presented in the 1970's and routed sound to one or more of 24 loudspeakers [10].

The *AUDIUM* sound system consists of 136 loudspeakers placed around the performance space [11]. In this system, volume, direction, and speed are controlled through a customized board and sound is routed to the loudspeakers through a series of binary and ternary switches. The 1970's saw some of the first uses of computer controlled audio spatialization. Chowning's *Turenas* (1972) used a computer to control the movement of sound through space. This piece projected sound through four loudspeakers situated around the listening space and explored distance through the use of artificial reverberation and the Doppler effect [12]. The *HYBRID* synthesizer distributed synthesized sound to 16 loudspeakers controlled by a computer [8]. The computer based system *Trails* (1987) was capable of both preprogramed and realtime distribution of sound up to 32 channels [7].

Digital audio enabled further developments in spatial audio techniques. Using digital audio, amplitudes for each loudspeaker channel can be calculated by a panning algorithm. Many algorithms have been developed such as Ambisonics [13], Vector Base Amplitude Panning (VBAP) [14], Distance Based Amplitude Panning (DBAP) [15], and Wave-Field Synthesis (WFS) [16].

The rise in mobile technology such as cellphones in the late 20th century and smartphones in the early 21st century has allowed for the use of the audiences' devices as a loudspeaker array [17]. In *Dialtones: A Telesymphony*, audience members' ringtones are played by calling their cellphones during the concert [18]. The emergence of smartphones enabled the "mobile phone orchestra", which is not necessarily more spatial than a traditional orchestra, however it can be extended by including audience members as performers in works such as *echobo* [19].

2.2 Spaces

There are many institutions that have dedicated spaces for spatial audio systems and the uses of these spaces range from scientific investigation to composition and performance of spatial audio works. While ad hoc spatial audio systems are very common and used for a single performance or exhibition and then taken down, permanent and semi-permanent spatial audio systems allow for a reproducible spatial audio field. These systems typically include a high-density loudspeaker array (HDLA) with the ability to use several different spatialization algorithms. In several of these systems, physical properties such as loudspeaker placement, room size, and room materials can be changed. There are many examples of spatial audio systems in use and the following will discuss just a few of these examples.

The *Cube* at Virginia Tech has 150 loudspeakers arranged in four rectangular layers with the top layer being a grid layout [20]. It uses an IP-based network to distribute audio and the three main spatialization methods used in this system are 3D Wave-Field Synthesis, Higher-Order Ambisonics (HOA), and VBAP. The *AlloSphere* is a multiuser, interactive platform for visual and aural composition and performance at University of California, Santa Barbara [21]. It includes a 3D 54.1 spatial audio system with loudspeakers arranged in three rings on the outside of the sphere. Spatialization methods include VBAP, DBAP, LBAP, Ambisonics, and WFS. *The Klangdom* [22] is a spatialization system at ZKM consisting of 39 loudspeakers on sliding tracks in four layers. The tracks allow the loudspeakers to be easily reconfigured for different performance setups. The system uses spatialization methods such as Ambisonics and VBAP, and extensions to Ambisonics and VBAP called Sound Surface Panning (SSP) which adds a width parameter to point sources [22]. *Espace de Projection* hall at IRCAM has a 350 loudspeaker array [23] and the materials of the room and room size can be changed. On the walls there are 171 groups of independently controlled rotatable prisms with three different surfaces. The height of the ceiling can be varied and the room can be subdivided into different sections with curtains. The loudspeakers are arranged in four horizontal arrays for WFS, and also in a hemisphere layout for 3D higher-order ambisonics. Additional spatialization methods include VBAP and HRTF. *The Sonic Laboratory* at Queen's University Belfast has a 48-channel spatialization system surrounding the audience [24]. Loudspeakers on box trusses can be raised or lowered and the acoustics can be changed through the positioning of numerous sound absorbers.

Other spaces outside of institutions allow for spatial audio compositions to be experienced by a larger audience. Movie theaters typically use spatial audio systems and there are a variety of formats and standards available that allow for compositions to be performed in this space. There are also many spatial audio systems available to the consumer that use technologies such as 5.1 surround sound, THX Spatial Audio [25], and Dolby Atmos [26].

There are several things to note about the above spaces in relation to this dissertation. First, there are a number of spatialization methods available. Changing the spatialization method changes the gain curves thus changing the virtual field. Second, several of the spaces have configurable loudspeakers. Different loudspeaker configurations change how the effects of the virtual field on the source are projected into the acoustic field. Third, several of the spaces include the ability to change the acoustic field by changing the materials of the walls, changing the room dimensions, and using sound absorbers.

Changing the spatialization method, loudspeaker configuration, or acoustic field are all examples of changes to the broader spatial audio field.

2.3 Spatial Audio Composition Techniques

One of the most basic and effective compositional techniques for spatial audio is the positioning of a source in space. This is either accomplished through direct routing of a source to a loudspeaker, or through the positioning of a virtual source. Source positioning is used extensively by composers and is commonly done through the audio software's built-in panning or through potentiometers on mixing consoles [27] . Positioning of multiple sources can create geometric forms such as a point, line, plane, polyhedron, or cloud known as a *spatial chord* [28]. Creating the illusion of a moving source by assigning it a trajectory is another tool used in spatial audio composition. Chowning's *Turenas* (1972) is a notable example of the use of source trajectories.

Another common technique is the manipulation of the dimensional and immersive attributes of the spatial image. The dimensional attributes include source direction, distance, and extent (depth, width, and height), and the immersive attributes include presence, and envelopment [29].

Perceptual differentiation and unification of audio streams is another technique used in spatial audio composition. The perceptual differentiation of complex auditory scenes is increased by separating sounds spatially [30]. For example, a quiet tone is masked by a loud tone when mixed together, however when both tones are spatially separated, a listener can hear both simultaneously [30]. With spatial audio, the listener is able to interact with the auditory scene by moving their head, or by moving around the listening space [30].

Spatially separating what would normally be together in a sound's spectrum or in a sequence of related events is another technique that uses perceptual differentiation. In *timbre spatialization* [31], a sound's spectrum can be subdivided using various methods such as bandpass filters or sinusoidal decomposition, then these parts can be assigned separate spatial positions. A sound can also be segmented in the time domain, categorized based on spectral or other features, and then spatialized [32]. Individual notes of a melody or melodic ornamentation such as a trill can also be separated spatially as described in [30]. Spatial separation can also be based on pitch-class, duration, dynamics, articulation, or any other property of a sound.

Spatialization of microsound is another technique used in spatial audio composition. Sound particle positioning can be static, move from one location to another, move from a fixed position to a random position, or can be positioned randomly [33].

Spatial modulation synthesis (SM) uses high velocity sources with periodic orbits to synthesize sound based on simulated Doppler effect and distance based attenuation [34]. With a high enough velocity, Doppler shift applied to a source can produce audio-rate frequency modulation called Doppler FM [34]. Similarly, distance based attenuation applied to a high velocity source orbiting with extreme changes in distance can produce audio-rate amplitude modulation [34]. SM is closely related to the work in this dissertation as SM is created through the interaction of a source with the virtual field to produce sound synthesis. In SM, the virtual field (as it is called in this dissertation) not only contains the gain curves for panning a sound among the loudspeakers, but also contains a gain curve for distance based attenuation and a definition of how the source frequency changes based on its radial velocity relative to the listener.

The compositional techniques mentioned above primarily deal with positioning and trajectories of sources to be used in a specific instance of a SAF. All of the techniques can be realized in a static SAF and this dissertation adds new compositional techniques based on changes to the virtual field of the SAF.

Chapter 3

Spatial Audio Fields

3.1 Introduction

This chapter discusses the theory of spatial audio fields. It begins with an overview of the SAF, then discusses the virtual field, loudspeakers, and acoustic field individually. The theory is developed to serve as the research's contextual framework and it assists in identifying parts of the field that can be changed on short timescales.

3.2 The Spatial Audio Field

The spatial audio field (SAF) exists in any spatial audio implementation and is independent of a source or listener. Figure 3.1 shows a diagram of the three components of a spatial audio field: virtual field, loudspeakers, and acoustic field. The SAF can be constructed as an intentional compositional choice, however it is often constructed based on available software and hardware in the performance space.

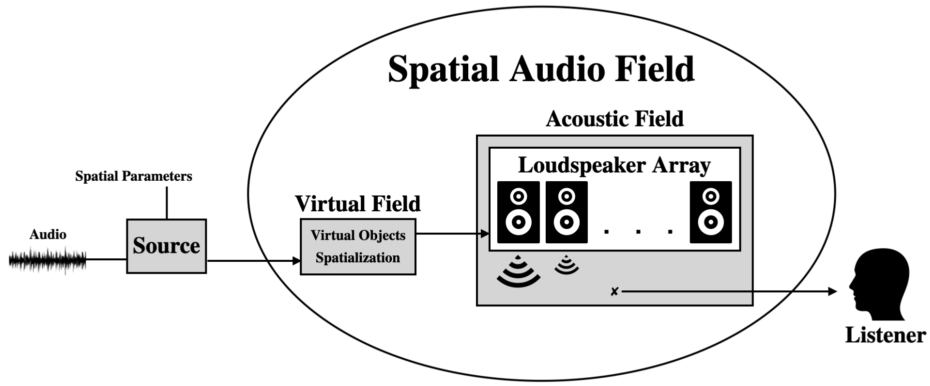


Figure 3.1: The spatial audio field.

3.3 Virtual Field

The virtual field is the first component of the SAF that the source encounters. It is where a source undergoes transformation by virtual spatial objects positioned in the field. Because the virtual field is constructed digitally, it can be changed on short musical timescales down to the sample rate. The virtual field is composed using virtual spatial objects and spatialization algorithms in the field's domain.

Virtual Spatial Objects

Virtual spatial objects are virtual representations of the sound sources, modifiers, and triggers or zones that are positioned in the virtual field. These objects include a position at the very least, but can also include an area of influence (field) around its position. The most common modifier object is the loudspeaker object and its gain is determined by the proximity of a source. Virtual spatial objects and modifier objects that change the loudspeaker array's gain curves are discussed in more detail after discussing basic gain curve modification in Chapter 4.

Spatialization Algorithms

Spatialization algorithms determine how a source signal is processed based on the source's spatial parameters from Figure 3.1. Position, velocity, and attributes of the spatial image are examples of spatial parameters that may or may not be handled by the spatialization algorithms. While a specific virtual field does not need to handle all possible spatial parameters, the handling of source position is usually defined by the panning algorithm.

Algorithms such as linear panning, constant power panning, Ambisonics, and VBAP calculate the loudspeaker gain curves. In Chapter 4, these gain curves will be generalized as parameter curves whose target is loudspeaker gain. Additional algorithms and equations are often used in conjunction with basic positioning and panning algorithms, such as distance based attenuation and reverberation. Other spatialization algorithms use parameters besides position. To simulate the Doppler effect, a source's radial velocity relative to the listener is used to modulate the frequency of the source content. Source width, height, and depth are other parameters that are often handled in the virtual field. The choice of specific spatialization algorithms and whether or not to handle other spatial parameters of the source shape the contents of the virtual field.

Virtual Domain

The virtual domain consists of all possible positions of spatial objects in the virtual field. Essentially, a point is part of the domain if it is handled by the software. Figure 3.2 shows four different domains superimposed on a physical loudspeaker layout. In 3.2a, the virtual domain is restricted to a single point corresponding to the position of the loudspeaker. Positions of any spatial objects outside of this position are ignored. Figure 3.2b shows the domain extended beyond the loudspeaker's position, but not to the left

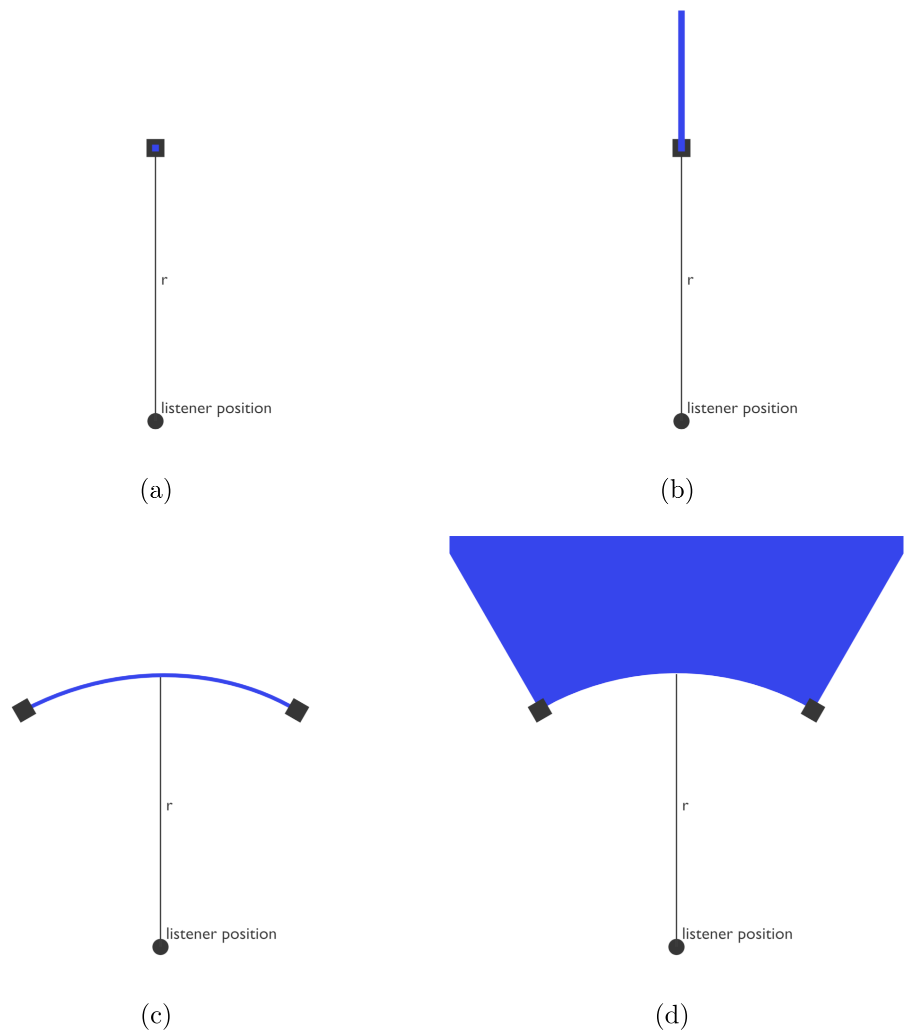


Figure 3.2: Top-down perspective of mono and stereo loudspeaker layouts with different virtual domains where r is the distance from the listener to the loudspeakers. The areas indicated show the allowable spatial object positions in the different domains.

or right. In this case, while the software allows the positions indicated by the domain, it will sound exactly the same as Figure 3.2a unless something is done to modify the source based on its position. A common modification is to apply distance based attenuation and reverberation to imply distance beyond the loudspeaker. Figure 3.2c shows a virtual domain that includes positions between two loudspeakers, and Figure 3.2d shows the same but with an additional distance parameter.

The virtual domain can be used as an element of composition. It can contract and expand, entire sections can be removed, or holes can be created in the virtual field. Limiting the domain creates positions inaccessible to the source even if the source position includes these areas. While the same result can be achieved by carefully composing the source position over time, many types of compositions use aleatoric or stochastic processes to generate source positions. It is also common for parameters to be determined by data not originally intended to be used as source position. In any of these cases, changing the domain is a method of restricting the source's movement from within the virtual field.

Perceptual Domain

The perceptual domain is not part of the virtual field, however it is dependent on the virtual field, loudspeakers, acoustic field, and listener position. It is mentioned here to discuss the use of decorrelation in the virtual field and to discuss mapping in the following section. While the virtual domain can include any position, the perception and localization of source positions may not coincide with the intended source position (see [35]). Thus the perceptual domain, or the set of all possible positions that a source is localized, is often unintentionally different from the virtual domain.

With the listener in the "sweet spot", the perceptual source domain would look similar to source domains in Figure 3.2, however if the listener is significantly closer to one loudspeaker, the perceptual domain can collapse to the location of the nearest loudspeaker due to the precedence effect. The audio signals coming from the active loudspeakers can be decorrelated [36] to increase the similarity of the virtual and perceptual domains. Decorrelation can be turned on and off or only used in parts of the virtual field to take advantage of differences in localization with and without decorrelation.

Virtual and Perceptual Domain Mapping

The final component of the virtual field is the mapping of the virtual domain to the perceptual domain. In an ideal system attempting to accurately position a virtual source in space, the virtual domain directly maps to the perceptual domain. However, this does not necessarily need to be the case and any mapping can be used. Remapping the virtual domain can be used to change a source's movement without altering the source's position or trajectory. This is especially useful in instances where the source position is pre-recorded or comes from other data.

3.4 Loudspeakers

Having been transformed by the virtual field, the source's waveform is further transformed by the loudspeakers as it transitions from the virtual field to the acoustic field. The way in which the source is transformed is determined by the properties of the loudspeakers which are subdivided into individual loudspeaker properties and properties of the loudspeaker array. Like the virtual field, the construction of the loudspeaker component of the SAF may be an intentional compositional choice or it may be determined by whatever is available in the performance space.

Individual Loudspeaker Properties

Individual loudspeaker properties include frequency response, amplitude response, position, directionality, and orientation. The frequency response of a loudspeaker is a fixed property determined by its construction, but different frequency responses can be simulated. The position of a loudspeaker can be physically changed as is done in *The Klangdom* [22] spatialization system or the positions can be simulated using virtual loudspeaker locations as described in [37]. Loudspeaker orientation can also be changed by mounting them on a rotating platform or rotating loudspeakers can be simulated through digital signal processing [7].

Loudspeaker Array Properties

The loudspeaker array properties include uniformity, density, and configuration. The uniformity refers to the similarity between loudspeakers. An entirely uniform array has loudspeakers that are identical. A mixed array such as a 5.1 surround sound loudspeaker array has five identical or similar loudspeakers plus an additional loudspeaker that handles lower frequencies. On the other hand, an *Acousmonium* [38] has several different types of loudspeakers with various characteristics. As with many of the physical properties of the SAF, the uniformity cannot be easily changed on short timescales. However, individual loudspeaker characteristics can be simulated as is done in the *Virtualmonium* [37] and used to instantly change the uniformity of the array.

The density of the loudspeaker array further contributes to the spatial audio field. Density can vary from a single loudspeaker to large numbers of loudspeakers found in a high-density loudspeaker array (HDLA). The density of the array has implications for localization and perception of smoothness of a moving source. Higher density arrays typically result in improved localization accuracy and lower density arrays can have a better perception of trajectory smoothness. The density of the array can be changed in realtime to take advantage of high and low densities depending on the circumstance.

Different loudspeaker array layouts / configurations change the spatial audio field. Common configurations include stereo, quadrophonic, octophonic, ring, spherical, panel, and scattered. Configurations can also be changed in realtime, but the variability of the configuration depends on the top-level configuration. For example, a spherical layout such as the layout used in the *Allosphere* [21] is reduced to a ring layout by disabling all loudspeakers outside of the horizontal plane. Furthermore, a ring layout is reduced to a quadrophonic, stereo, or mono layout by decreasing the number of enabled loudspeakers. Changing the loudspeaker layout by disabling loudspeakers can be done as fast as the sample rate and will be discussed in detail later. However, changing the layout by repositioning loudspeakers is more difficult but can be done by mounting loudspeakers on sliding tracks [22] or moveable platforms [24]. Positions of loudspeakers can also be simulated as is done in the *Virtualmonium* [37].

3.5 Acoustic Field

The third and final component of the SAF is the acoustic field and is the physical environment where sound is spatialized. Of the three components of the SAF, the acoustic field and its properties are the most difficult to change on short timescales because most of the properties are physical properties. On a general level, the physical properties of the space include its *enclosure* and *materials*, while the virtual properties include simulations of how a sound is transformed by simulated physical parameters.

Enclosure

Enclosure refers to the boundaries of the space and ranges from a free field to completely enclosed. Examples within the range of enclosure include partially enclosed spaces, or spaces with openings. Enclosure determines properties of the acoustic field such as reverberation and resonant frequencies.

If the space has some level of enclosure, the dimensions of a space help determine properties such as reverberation time and resonant frequencies. If it was possible to manipulate the space's dimensions in realtime, natural reverberation and resonant frequencies would be a controllable parameter of the acoustic field similar to how brass instruments can vary their lengths through valves or slides. In terms of a room, however, realtime dimensional changes are not practical on short timescales. On the other hand, there are many examples of spaces used for spatial audio that can change their dimensions through moveable walls and ceilings, and subdividing the space with curtains as mentioned in Chapter 2.

Materials

A space's physical materials can be divided into the materials of the sound propagation medium and the materials of the boundary. The material through which sound propagates is a major component of the space in a SAF. This material directly determines the speed of sound and in the context of spatial audio that we normally experience, the material is typically air at a pressure around one atmosphere. Changing the speed of sound disrupts the perception of spatial sound. For example, the interaural time difference of a sound propagating through air is longer than the interaural time difference of the same sound propagating through water despite coming from the same location relative to the listener (see [39]). In addition to the speed of sound, the doppler effect and distance attenuation is also determined by the propagation medium. While modulating the speed of sound by changing the propagation medium to disrupt a listener's localization mechanisms is interesting to consider hypothetically, it is currently not practical and leaves the sound propagation medium as a static component of the acoustic field.

The materials used for the space's boundary determine how much sound is absorbed versus how much sound is reflected. In an anechoic chamber, the materials used for the boundaries are designed to prevent reflection by ideally absorbing all sound. In contrast, dense materials with smooth surfaces reflect a greater proportion of the sound's energy resulting in a more reverberant space. Similarly to the dimensions of a space, the materials of the space's boundaries are not practical to change on short timescales. However, there are examples of spaces where the materials of the boundary are configurable as mentioned in Chapter 2.

Simulated Acoustic Fields

Changing the acoustic field in the context of SAF is primarily done through simulation. Properties of another space can be brought into the space being used or properties can be simulated artificially. Another space can be brought into the actual space through an impulse response being convolved with the source audio, or sound field recording can be used for the source and distributed to the loudspeakers. However, the properties of the acoustic field cannot be canceled out and the listener will ultimately receive the transformation of the source from the simulated space and from the acoustic field.

Compositional Elements of the Acoustic Field

Changes to the acoustic field, such as dramatically changing the dimensions of a concert hall over the course of 500ms, or continuously varying the sound reflectivity of the walls from completely reflective to completely absorptive, are not practical on short timescales. However, changes to the acoustic field can be used compositionally on larger timescales. A spatial composition might have multiple movements with each taking place in a different space for example. On the other hand, changes to the acoustic field can be used compositionally on shorter timescales when the field is simulated.

When a SAF is implemented using headphones, the acoustic field is mostly simulated but still includes an acoustic field. With over-ear headphones, the acoustic field includes the listener's pinna and ear canal. With in-ear headphones or "earbuds", the pinna is bypassed and external sounds can be mostly blocked using noise canceling or acoustic isolation. This permits a completely simulated acoustic field which can be manipulated using signal processing and used compositionally. Another compositional possibility arises from using loudspeakers in addition to headphones. The level of the sound from the physical acoustic field can be modulated by changing the amount of noise canceling allowing for an interplay between physical and simulated acoustic fields in an auditory "mixed reality".

3.6 Summary

Changes to the physical properties of the SAF on short musical timescales are either difficult or impractical. However, many properties of the SAF can be simulated such as room acoustics, loudspeaker properties, and loudspeaker array properties. As the virtual field is usually the part of the spatial audio field that is constructed digitally, the virtual field presents a unique opportunity for developing new compositional techniques.

Chapter 4

The Virtual Field: Loudspeaker Density, Gain Curves, and Spatial Objects

4.1 Introduction

The previous chapter defined the SAF and discussed the field's three main components: virtual field, loudspeakers, and acoustic field. Because a source encounters the virtual field first, changes to the virtual field propagate throughout the rest of the SAF. This chapter will give a detailed discussion of two parts of the virtual field that are seldom changed during the performance of spatial audio works: loudspeaker density and loudspeaker gain curves. This chapter will also further discuss virtual spatial objects in relation to loudspeaker density and gain curves.

This chapter and the following chapters will reference an example loudspeaker array. The array has 16 channels and is configured in a 2-dimensional ring layout. Appendix A gives a detailed description of the layout.

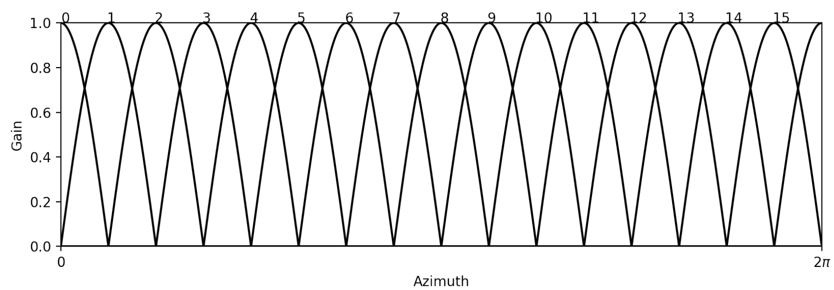
4.2 Loudspeaker Density

There are benefits to using a high density of loudspeakers such as improving localization and reducing the precedence effect. However, just because a large number of loudspeakers are available does not mean that they all must be used. Loudspeakers can be removed from the array without removing them physically by ignoring them in the panning algorithm. Changing the loudspeaker density programmatically by disabling or enabling loudspeakers allows for density changes to occur instantaneously.

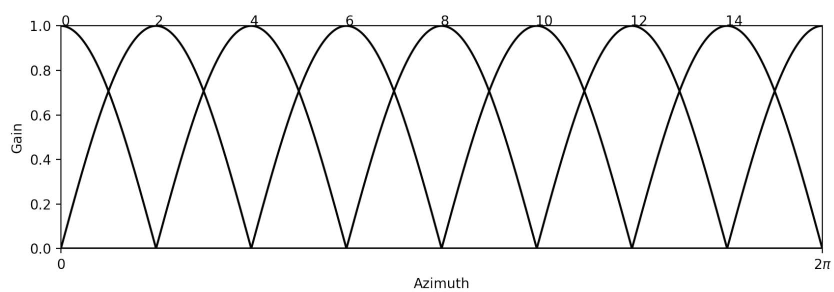
When a loudspeaker is enabled or disabled, the gain curves of the remaining enabled loudspeakers are recalculated (Figure 4.1). This is in contrast to simply unmuting or muting loudspeakers which leave adjacent loudspeaker gain curves the same. Changing loudspeaker density is used to improve trajectory smoothness and to control amplitude modulation resulting from fast moving sources. Density changes can also produce irregular densities in an otherwise regular loudspeaker arrangement such as the example layout. The following section will discuss why variable loudspeaker densities might be used in addition to methods of changing loudspeaker densities while avoiding discontinuities in the audio signal.

Trajectory Smoothness

Panning of slow moving sources sound more smooth with a higher loudspeaker density. On the other hand, at higher rates of speed, a less dense layout results in a smoother sounding movement being observed. This observation has not been confirmed through perceptual studies, however it is suspected that slight discrepancies in the physical loudspeaker layout versus the virtual representation, directionality of the loudspeakers, listener position, and effects of amplitude modulation all contribute to the perception of smoothness as a source moves throughout the space. With the ability to change the



(a) all enabled



(b) every other disabled

Figure 4.1: Loudspeaker gain curves of the example layout comparing two different densities. The loudspeaker channel is indicated at the apex of that loudspeaker's gain curve.

loudspeaker array's density, different densities can be used depending on the velocity of the source.

Amplitude Modulation

Amplitude modulation occurs when a bipolar carrier signal is multiplied by a unipolar modulator signal resulting in periodic variation of the carrier signal's amplitude [7]. When the modulation frequency is lower than approximately 20Hz, the result is tremolo while higher modulation frequencies produce difference tones.

In amplitude panning, virtual sources undergo amplitude modulation as they move through a space. The amplitude of a sound emitted by a loudspeaker is modulated by the gain curve of that loudspeaker based on the sound's position. At higher speeds, this amplitude modulation can begin to produce audible frequencies. In the example ring layout, as a source moves around the circle, at a certain angular velocity the amplitude modulation applied by each loudspeaker will begin to produce audible amplitude modulation. Referring back to Figure 4.1, the graphs show the gain curves for a density that includes all loudspeakers and a density with every other loudspeaker disabled. In Figure 4.1b, the width of the remaining loudspeaker's gain curves increase which reduces the modulation frequency and therefore the threshold at which amplitude modulation is perceived.

While the effects of amplitude modulation can produce interesting results, it may be undesirable in some cases. Decreasing loudspeaker density allows one to raise the threshold for when amplitude changes reach an audible rate thereby allowing a source to move faster before amplitude modulation can be heard.

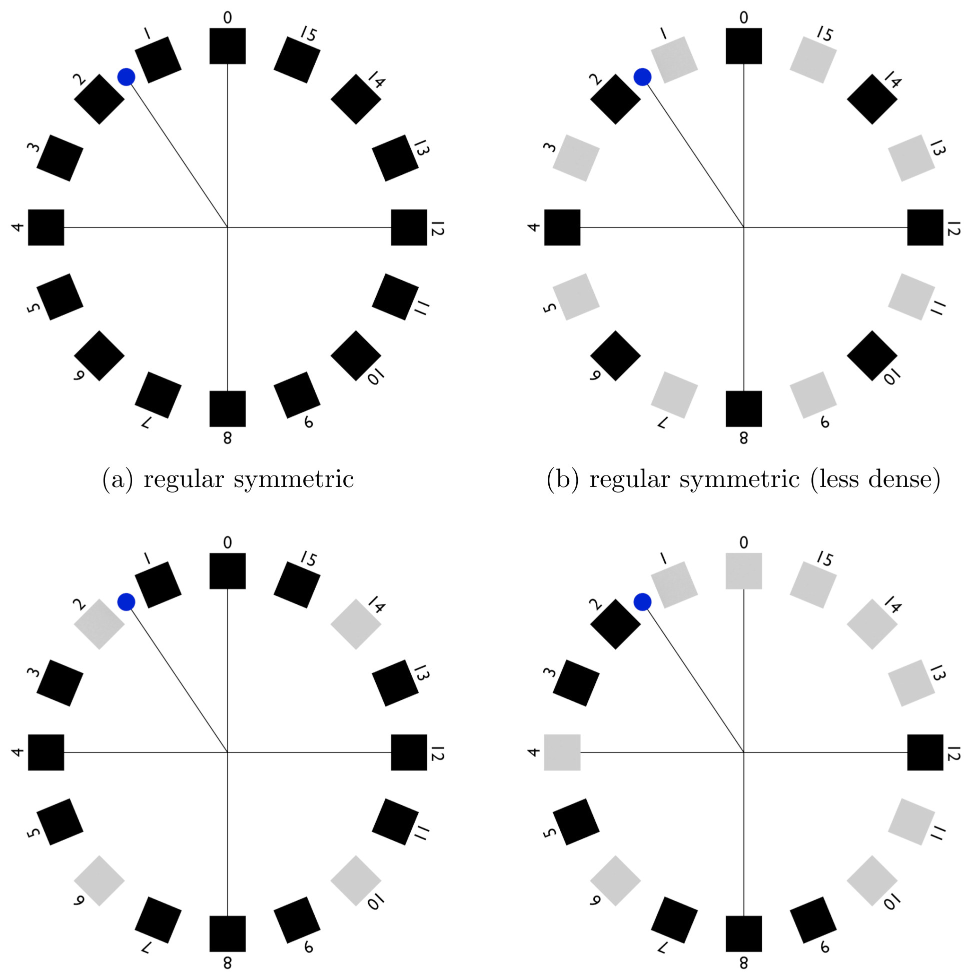
Irregular Densities

Just as the density of a loudspeaker layout does not need to use all of the available loudspeakers, the density of a given layout does not need to be regular. Because the density changes the spatial image and localization, irregular densities result in spatially dependent variations. HDLAs are often configured with some sort of symmetry of loudspeaker distribution. Disabling loudspeakers can produce irregular densities that are either symmetric or asymmetric. Irregular symmetric layouts have varying densities as a source moves through the space, but the variations in the density are symmetric along an axis. By contrast, asymmetric layouts do not have a symmetric density along any axis. Figure 4.2 shows the example layout with different loudspeakers disabled.

Figures 4.2a and 4.2b show regular symmetric density layouts. In Figure 4.2a, all loudspeakers are enabled and the layout is at its maximum density. In Figure 4.2b, every other loudspeaker is disabled and the panning algorithm treats the disabled loudspeakers as if they do not exist. Given the virtual source position in Figure 4.2b, only channels 0 and 2 will produce audio with their gains being calculated based on the virtual source position between these two channels. Figures 4.2c and 4.2d show irregular density layouts with the first being symmetric and the second asymmetric. Figure 4.2c alternates between a higher and lower density while remaining symmetric about two axes. Figure 4.2d is irregular, asymmetric, and the virtual source at the position in the figure is panned between loudspeakers 2 and 12.

4.2.1 Methods

The main challenge to changing loudspeaker densities is avoiding discontinuities in the audio signal that can cause unwanted clicks and pops. For example, Figure 4.3a shows the gain curves of the example layout with the virtual source positioned at the



(a) regular symmetric

(b) regular symmetric (less dense)

(c) irregular symmetric

(d) irregular asymmetric

Figure 4.2: The example layout showing different regularities and symmetries. Disabled loudspeakers are shown in grey.

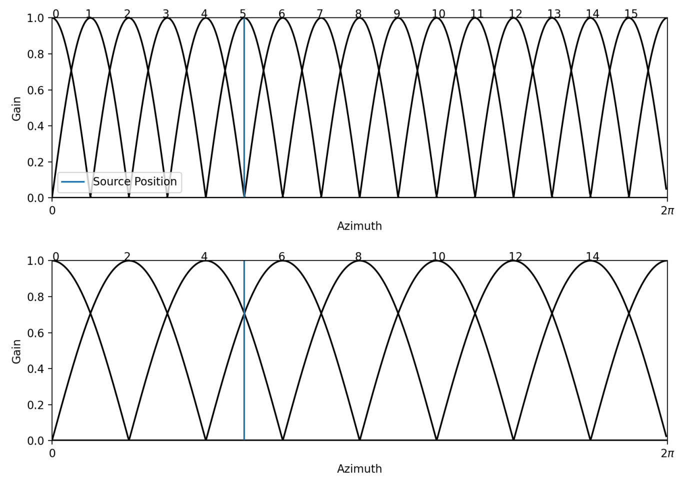
location of loudspeaker 5 before and after the density change. Figure 4.3b shows the waveform of the source distributed among channels 4, 5, and 6 with the density change occurring at time $t = 0$.

Notice that before the density change only channel 5 is producing audio. When the density is changed to the second layout with every other loudspeaker disabled, channel 5 abruptly stops producing audio and the audio is distributed to channels 4 and 5 which abruptly start producing audio. While clicks and pops can be avoided simply by producing no audio or fading out / in the virtual source's audio during the transition, this section proposes three methods of changing the density without interrupting the audio signal.

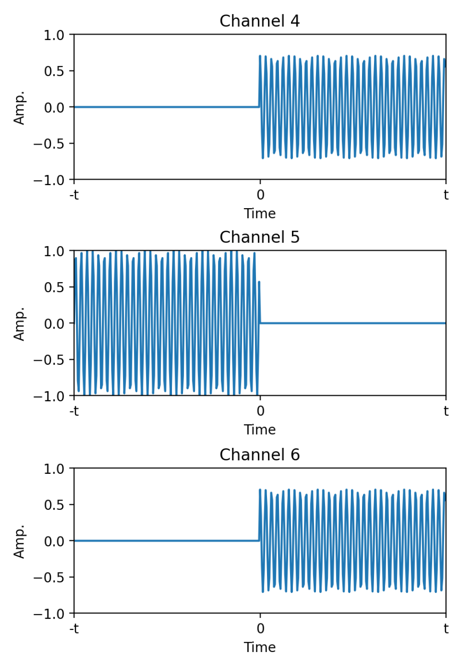
Colocation of the Source and Loudspeaker

In amplitude panning based on loudspeaker pairs or triplets, when the source position equals a loudspeaker position, only that loudspeaker produces sound. If this loudspeaker is enabled in both the current and target density layouts, the density can be changed without causing discontinuities in the audio signal. When using this method, a positional tolerance should be used because the calculated source position may not ever be exactly the same as the loudspeaker position. Figure 4.4 shows tolerance angle t around loudspeaker channel 1. The tolerance should be large enough to detect high velocity sources where the difference in the position calculated from one sample to the next is larger. On the other hand, the tolerance used should be small enough that any change in the gain curve of the loudspeaker in question will result in an amplitude change that does not cause audible discontinuities.

There are several things to consider when choosing this method. First of all, this method only works for moving sources or static sources that are positioned within the loudspeaker's tolerance. Next, higher numbers of sources decrease the opportunity for a



(a) gain curves



(b) waveforms

Figure 4.3: (a) Gain curves before and after a density change for a source positioned at loudspeaker 5. (b) Waveforms of loudspeaker channels producing audio with the density change occurring at time $t = 0$.

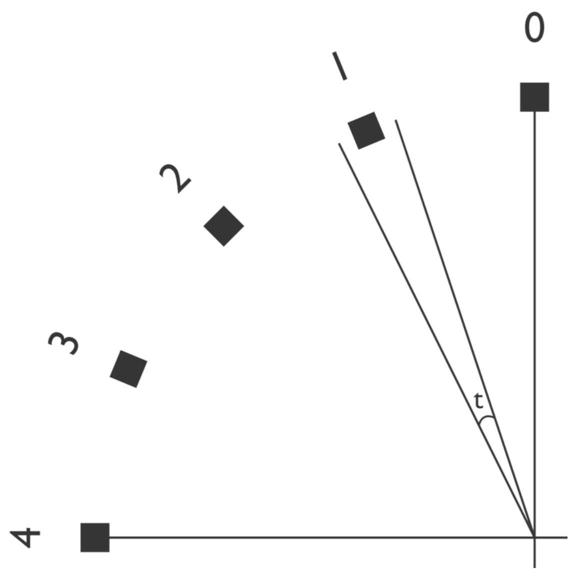
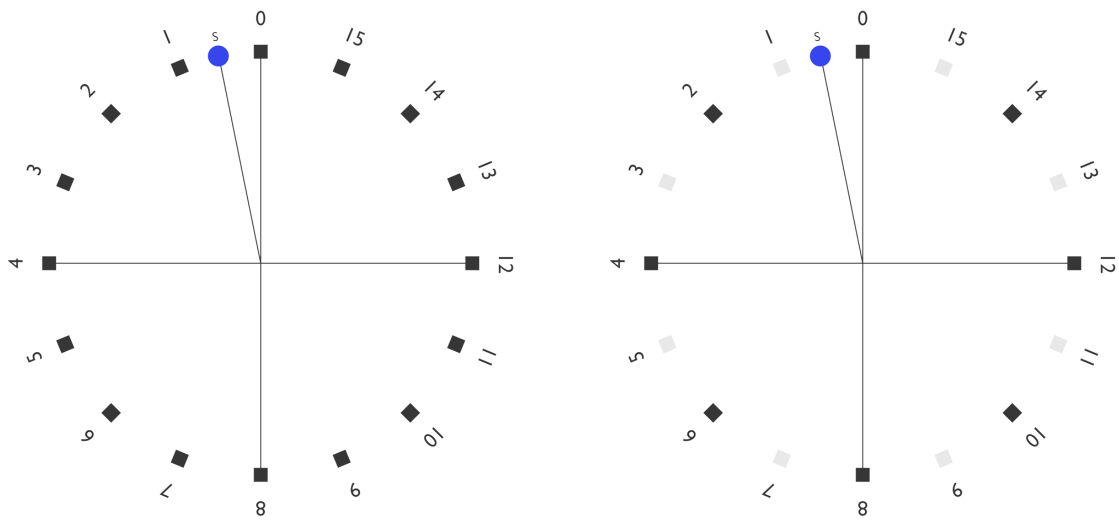


Figure 4.4: Tolerance t shown for loudspeaker channel 1.

density change because there must be a time where all sources are within the tolerance of a loudspeaker enabled in both layouts. Another thing to consider when changing density using colocation is that the current and target layouts must contain at least one loudspeaker in common. Figure 4.5 shows an example of a current and target layout. If source s is rotating counterclockwise, the density change cannot take place when s reaches loudspeaker 1 because it is disabled in the target layout. Changing the density here would abruptly cause the output of loudspeaker 1 to go to zero resulting in an audible click. Loudspeaker 2 is enabled in both the current and target layouts and the density can be changed when s is within loudspeaker 2's tolerance.

Finally, there exists the possibility that the calculated source position will never fall within the tolerance of a loudspeaker enabled in both density layouts preventing the density from ever changing. While this occurrence should be rare in practice, increasing the tolerance, changing the rotational frequency of the source, using sample-wise calculation of the position, or increasing the sample rate can help prevent this.



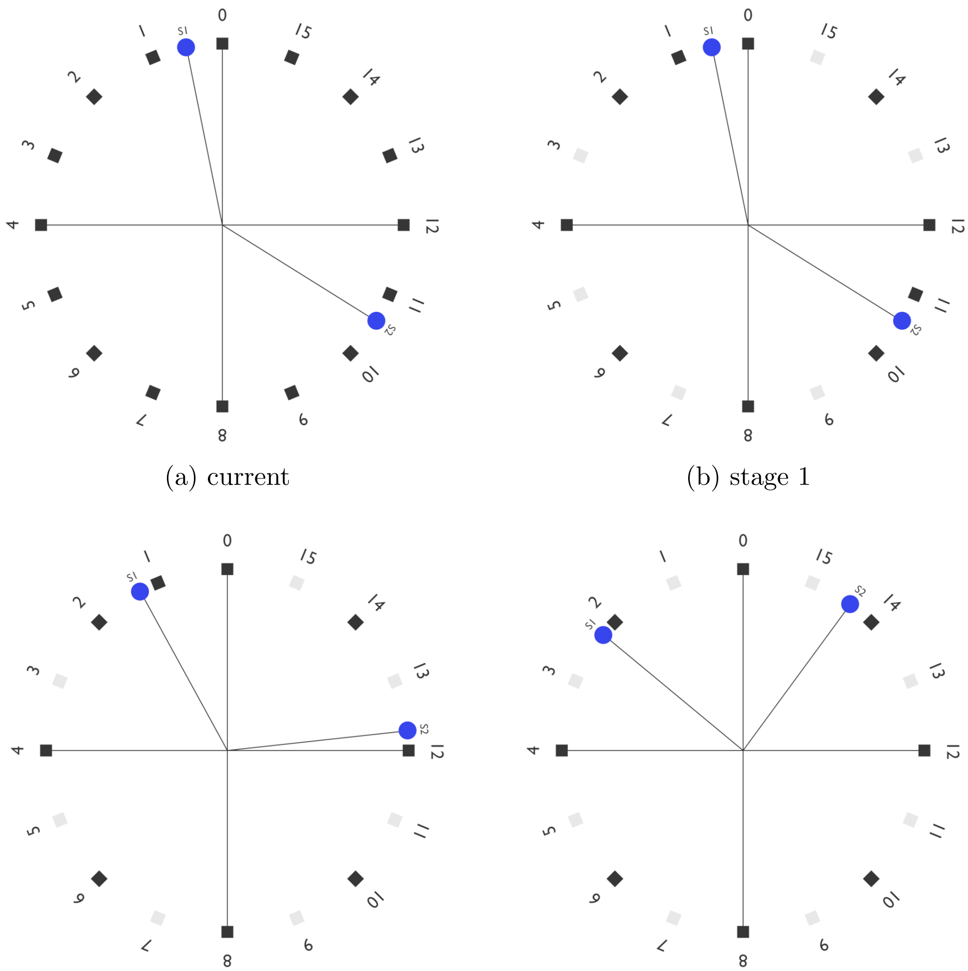
(a) current (b) target

Figure 4.5: Virtual source position s in the current and target densities. Disabled loudspeakers are shown in grey.

Changing Density in Stages

Changing from one density layout to another does not need to be done instantly and can be done in stages while avoiding discontinuities. In contrast to changing the density based on colocation of the sources with loudspeakers enabled in the current and target densities, this method allows loudspeakers to be enabled or disabled when sources are not within their field. A loudspeaker can be added or removed if there are no virtual sources within the loudspeaker pairs of both the current and target loudspeaker layouts. With a single source, changing the density in stages will have no audible difference from changing the density using the previous method. When using multiple sources however, changing the density through several stages allows sources to begin using the target density during this transition. Figure 4.6 shows the intermediary stages of a transition from all loudspeakers enabled to every other loudspeaker disabled.

In the first stage of the transition (Figure 4.6b), all of the loudspeakers disabled in the target can be disabled except for loudspeakers 1 and 11 because they are still



(a) current

(b) stage 1

(c) stage 2

(d) target

Figure 4.6: Transition in stages from the current to target density layout involving two virtual sources.

producing audio from the sources. In Figure 4.6c, loudspeaker 11 is disabled because $s2$ has rotated past loudspeaker 12 and is no longer within loudspeaker 11's gain curve. However, loudspeaker 1 cannot be disabled until $s1$ has rotated past loudspeaker 2 as in Figure 4.6d.

As with the previous method, there are several things to consider. This method only works for moving sources beacuse there must be a time when a source is not within the field of a loudspeaker pair in the current and target layout. Also, the time it takes to complete the density change depends on the speed and position of the virtual source. Therefore, calculating the time is only possible with knowledge of a source's future position. Finally, density changes may be difficult with larger numbers of virtual sources. Depending on the number and movement of the sources, there are situations where a layout will never complete its density change. This can be mitigated by having different intermediary layouts for each source, with the trade off being added computational complexity and requiring more processing resources.

Crossfading Two Densities

Changing densities can also be accomplished by cross-fading the current and target density layouts. Any discontinuities that would result from the layout change will be smoothed out by the crossfade. This method not only works for static sources, but also works in situations with larger numbers of sources eliminating possibility of never being able to complete the density change described in the previous methods. Additionally, the time it takes to change densities is known and is equal to the duration of the cross-fade.

Two things should be considered when using this method. First of all, when the loudspeaker gain curves are transitioning to the new density layout, the overall power of the loudspeakers representing the source can vary during this transition. Secondly, up to four loudspeakers can be producing audio for a given source during the change in density.

This increases artifacts due to phase interactions of the four loudspeakers.

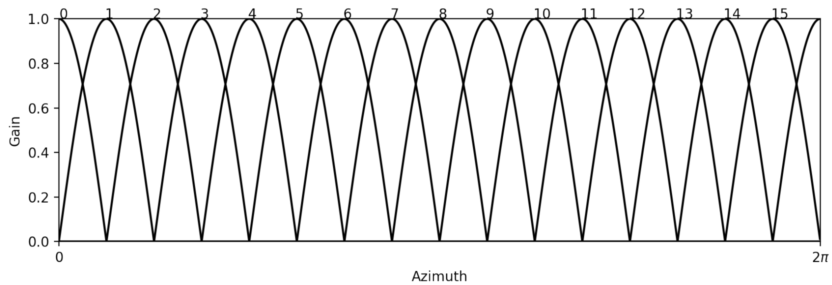
4.3 Loudspeaker Gain Curves

Changing the density of the loudspeaker array described in the previous section is done by adding or removing loudspeakers and changing the gain curve widths of adjacent loudspeakers. This section details methods of changing gain curves for other reasons besides density. In the context of SAFs, a loudspeaker's gain curve is treated as the parameter curve controlling gain of a loudspeaker object. A gain curve's properties include its shape, width, height, symmetry, and position which are all typically determined by the chosen panning algorithm and are typically fixed for a given composition. The different ways that gain curves can be modified are discussed in this section and the compositional application of these modifications are discussed in Chapter 5.

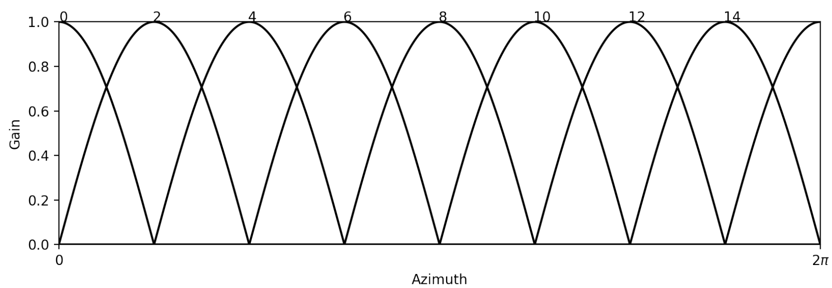
4.3.1 Muting

Disabling a loudspeaker causes adjacent gain curves to increase width to cover the space left by the disabled loudspeaker. However, when a loudspeaker is muted, adjacent gain curves remain unchanged leaving silent gaps in the virtual field. Figure 4.7 shows the gain curves of the example layout with all loudspeakers enabled 4.7a , every other loudspeaker disabled 4.7b, and every other loudspeaker muted 4.7c. With every other loudspeaker muted, the silent gaps only occur at the exact position of the muted loudspeaker. However, larger gaps are created by muting additional loudspeakers.

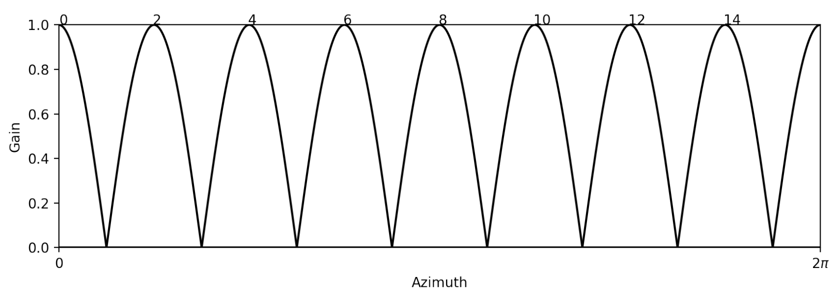
In practice, muting a loudspeaker can be done digitally in the audio processing callback of the software by simply returning a gain of zero if the loudspeaker is muted. While loudspeakers can also be muted in the analog stage by reducing the channel's gain to zero on a mixing console for example, muting them digitally can be done as fast as the



(a) all loudspeakers enabled



(b) every other loudspeaker disabled



(c) every other loudspeaker muted

Figure 4.7: Loudspeaker gain curves of the example layout comparing all enabled, disabled, and muted. Note that the gain curve widths only change when loudspeakers are disabled and they do not change when loudspeakers are muted.

sample rate.

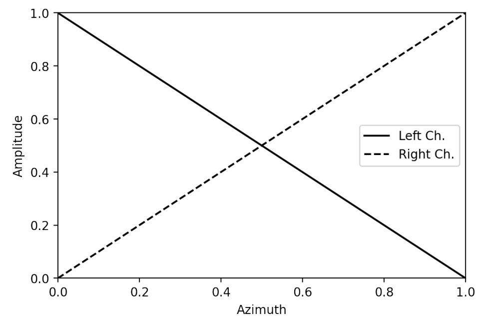
As with any change to a loudspeaker's gain curve, muting a loudspeaker while it is producing audio can cause discontinuities in the audio signal. If the muting is done digitally, a scheduled muting of a loudspeaker can be delayed until all virtual sources are outside of its area of influence in the same way discussed for changing loudspeaker density.

4.3.2 Gain Curve Shape

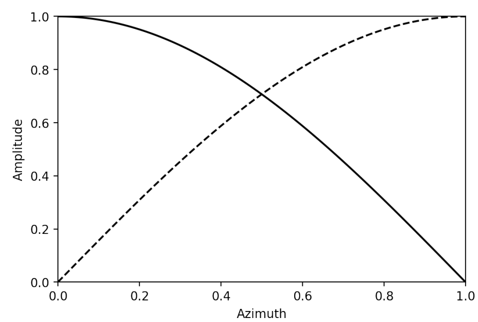
The shape of the gain curve is usually determined by the panning algorithm. Figure 4.8 shows three different inter-loudspeaker gain curve shapes calculated using linear panning, constant power panning, and inverted (constant power gain curve flipped vertically).

Different gain curve shapes result in differences in loudness perception as a source moves from one loudspeaker to another. With linear gain curves in 4.8a, the listener perceives a dip in loudness toward the middle of the loudspeakers. As constant power panning seeks to limit the perceived dip in loudness, its gain curve shape shown in 4.8b reduces this effect. However, a smooth source transition from one loudspeaker to another may not be the intended goal and one might want to *increase* the perceived dip in loudness for compositional purposes or other reasons. The gain curves in 4.8c make the dip in loudness more pronounced.

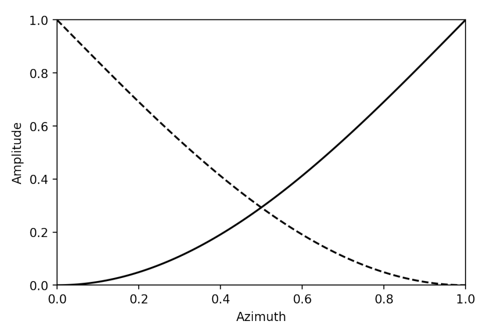
Figure 4.8 only shows three examples of gain curve shapes and any arbitrary shape can be used. For example, the inter-loudspeaker gain curves can contain multiple local maxima, or a loudspeaker's gain curve shape can be different for the loudspeaker to its left compared to the loudspeaker to its right. Regardless of the shape, it is important that the gain curve is continuous and that the tails reach zero gain to avoid discontinuities.



(a) linear



(b) constant power



(c) inverted

Figure 4.8: Various gain curve shapes between a loudspeaker pair.

4.3.3 Gain Curve Widths

Creating variations in loudness in the virtual field can also be achieved by changing the gain curve width. Here, changing the width is regarded as scaling the gain curve shape as opposed to changing it. Figure 4.9 shows the gain curves of a loudspeaker pair using three different gain curve widths and their corresponding power.

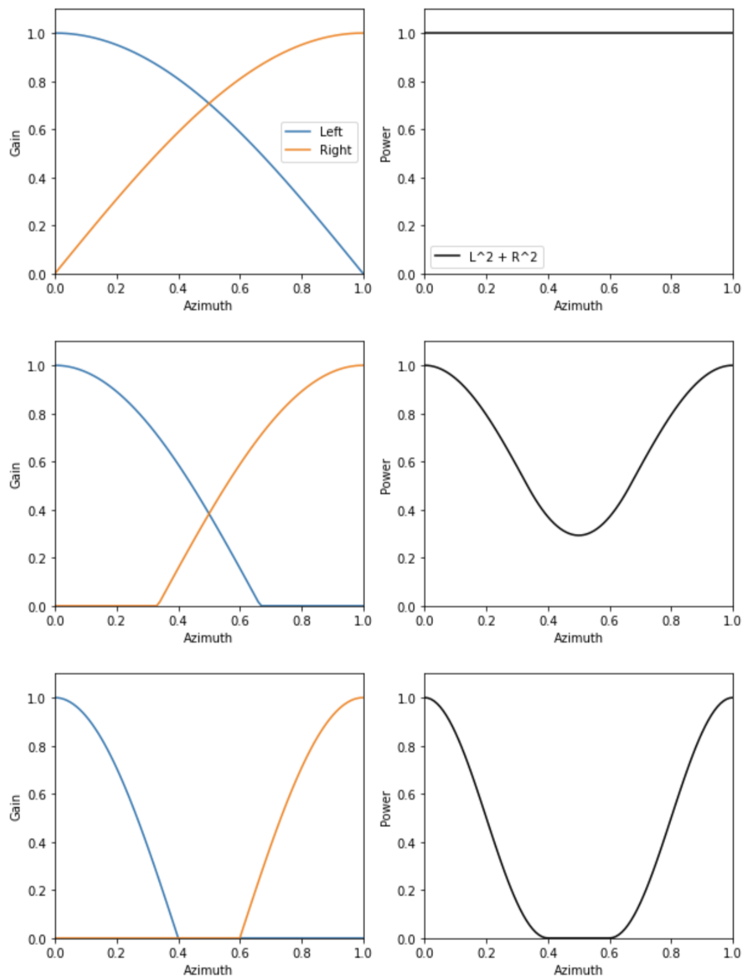


Figure 4.9: Gain curves of a loudspeaker pair showing three different gain curve widths (left column) and their corresponding power (right column).

In the top row of Figure 4.9, the gain curve width of both loudspeakers are at the maximum which produces normal constant power panning as a source moves from one loudspeaker to another. In the middle row of Figure 4.9, the widths are decreased but still overlap. As the source moves through the middle row's gain curves, the sound output is continuous but with a variation in loudness. When decreasing the width, there is a point where the loudspeaker pair's gain curves no longer overlap as in the bottom row of Figure 4.9 resulting in silent gaps in the virtual field.

Changing the gain curve width is done by moving the endpoint of the calculation. In normal amplitude panning, this endpoint is the position of the adjacent loudspeaker. Instead of using the adjacent loudspeaker position, the position of the gain curve's end is found by adding or subtracting the width to its position. If a source is outside of a loudspeaker's gain curve, the gain value is zero for that source, otherwise the loudspeaker gain is calculated. Changing the width can be done in both trigonometric and vector based approaches, however there is an increase in the number of calculations because loudspeaker gain values are being calculated individually and not as a pair.

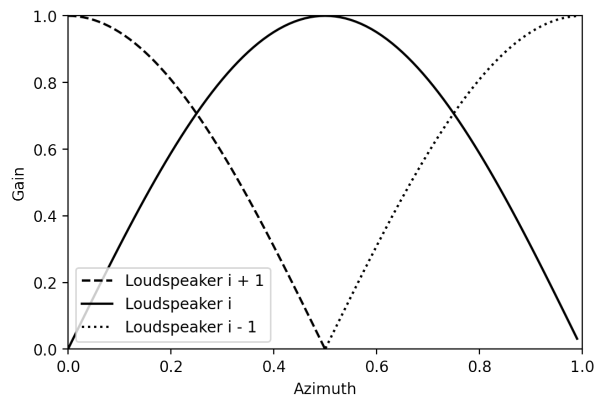
A simpler alternative method can be used to calculate the gain curves. This involves calculating a loudspeaker's gain based on the distance to the source instead of the distance to the adjacent loudspeaker or end of the gain curve. While this is panning based on distance from the source, this method differs from DBAP in that the source must be within a distance determined by the width parameter. Using this method also allows for gain curve widths to extend beyond adjacent loudspeakers.

Abrupt changes to gain curve width has the potential to create audible clicks caused by a discontinuity in the audio signal in a similar way to that which is described in Section 4.2.1 about changing density. The same methods of changing the density without discontinuities can be used to change the width without creating discontinuities. Alternatively, changing the width slowly over time will also prevent unwanted audible discontinuities.

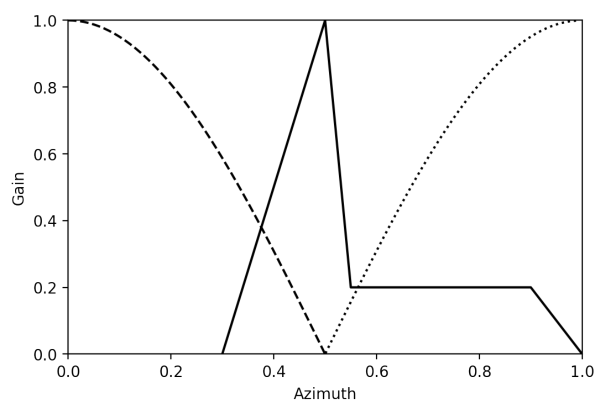
4.3.4 Asymmetric Gain Curves

When calculating typical gain curves, the gain curve is only symmetric about the position of the loudspeaker if adjacent loudspeakers are equidistant. If the distance from the loudspeaker to the left is different from the distance to the loudspeaker on the right, the gain curve is asymmetric. Alternatively, the symmetry of the gain curve can be modified by changing the shape and width of the curve to the left or right of the loudspeaker.

A source passing through the gain curve of a loudspeaker has the same effect as applying a window to a source positioned at that loudspeaker. The window is the gain curve and the duration depends on the source velocity. Thinking of the gain curve as windowing a moving source, one could also consider the gain curve as an envelope whose duration is a function of the gain curve width and source velocity. A basic envelope with an attack, decay, sustain, and release (ADSR) specifies a signal's gain for each part of the envelope in addition to corresponding times for each part. Instead of calculating the gain curve of a loudspeaker using a panning algorithm, the gain curve can be represented by this ADSR envelope, or any other type of envelope. Figure 4.10a shows the gain curves for a loudspeaker triplet calculated using constant power panning, and Figure 4.10b shows the same curves except loudspeaker i is replaced with a gain curve in the shape of an ADSR envelope. Narrowing the widths of the adjacent loudspeaker's gain curves, or muting them entirely, makes the effect of the envelope more pronounced as a source moves through. Variations on a single envelope can be added by swapping the left part of the envelope with the right, or by reversing the direction of the source. Small variations in an envelope can be applied to successive loudspeakers to cause a source to evolve as it moves through the array. Many additional compositional possibilities open up when one considers the gain curves of a loudspeaker array as an envelope, some of



(a)



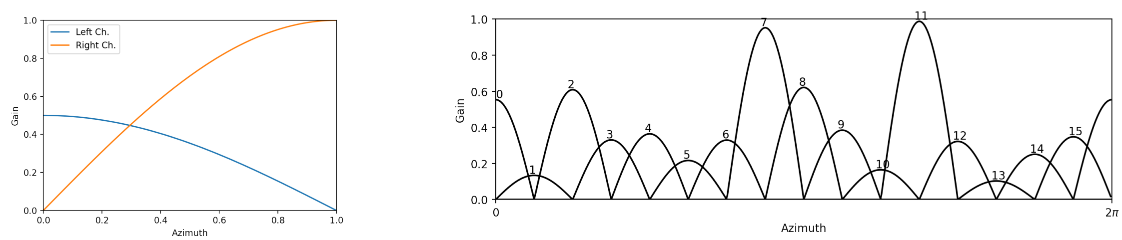
(b)

Figure 4.10: Gain curves of a loudspeaker triplet calculated using constant power panning (a) and an ADSR envelope for loudspeaker i (b).

which are described in the next chapter.

4.3.5 Gain Curve Height

The height of a gain curve refers to the vertical scaling of the gain in a graph of gain over azimuth. Similarly to changing the width, the height of a loudspeaker's gain curves can also be changed. While the effects of changing gain curve widths and heights are similar, the width changes the inter-loudspeaker response while the height changes both the inter-loudspeaker response and the response at the location of the loudspeaker. Figure 4.11a shows the partial gain curves for a loudspeaker pair in the example layout with height $h = 0.5$ for the left loudspeaker and $h = 1.0$ for the right loudspeaker. Figure 4.11b shows the gain curves for the entire array with the heights set to a random value.



(a) loudspeaker pair (b) loudspeaker array

Figure 4.11: Gain curves at different heights.

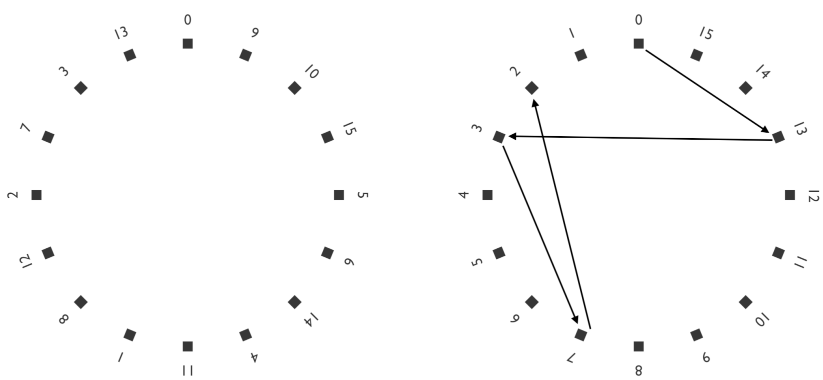
Varying the gain curve heights produces an irregular amplitude response that is spatially dependent. This can be used to create areas that are quieter than other areas, or to tune the response of the array to a specific acoustic field. It was observed by the author that changing the gain curve height from what is originally calculated to produce an even transition from one loudspeaker to another, to a different height changes how the movement was perceived. Taking Figure 4.11a, if a source is moving from the right loudspeaker to the left loudspeaker at a constant velocity, the perceived source position seemed biased towards the right loudspeaker. With a source moving through the gain curves of 4.11b, there was a sense that the source velocity was changing despite being constant.

Implementing the ability to change gain curve height is straightforward. Simply multiplying the calculated loudspeaker's gain curve by a scaling factor (usually 0 - 1) changes this height. However, changes in gain curve height poses the same issue of potentially causing discontinuities in the audio signal as the other types of gain curve modification. The methods described in Section 4.2.1 apply equally to changing gain curve height.

In analog signal processing, where the gain curve is determined by panning circuitry, gain curve height is adjusted by changing the channel's gain on a mixing console or changing the gain on the loudspeaker's amplifier itself. However, changing the gains programmatically allows for faster and more complex changes to the gain curve's heights. Manipulation of gain curve height is a simple yet effective technique that is not only used to fine tune a loudspeaker array's response in a particular space, but also can be used compositionally and will be discussed in the next chapter.

4.3.6 Decoupling Virtual Source Position and Perceived Source Position

In audio spatialization, the virtual source position usually corresponds to a desired location in the physical space where the source is intended to be localized. In this case, the virtual domain directly maps to the perceptual domain (ideally). However, this need not be the case and the mapping can be changed to any desired mapping. One way to change the mapping is to assign false angles [40] or false positions to the loudspeakers. A false position or angle occurs when a loudspeaker is represented in the virtual field that does not correspond to its physical position. If the difference between the false and actual loudspeaker positions are small and do not change the loudspeaker pairings, this causes a moving source to seem as if its velocity is varying even if the velocity is constant. When false positions vary greatly from the actual positions, loudspeaker pairings are changed causing a source to pan through non-adjacent loudspeakers. Figure 4.12a shows the example ring layout with false positions assigned to the loudspeakers. For a virtual source rotating counterclockwise from $0 - \frac{\pi}{2}$ in Figure 4.12a, the path the source takes in the physical layout is shown in Figure 4.12b. For each segment of the path



(a) loudspeakers at false positions

(b) actual loudspeaker positions

Figure 4.12: A virtual source rotating counterclockwise in (a) takes the path indicated in (b).

in Figure 4.12b, the source pans from the loudspeaker at the start of the segment to the loudspeaker at the end of the segment, skipping any loudspeakers in between. Complex patterns of perceived source movement can be created by using false positions. If a source is rotating in a ring layout with false positions, the pattern of movement repeats through every rotation of the source.

The loudspeaker positions do not necessarily need to be swapped with one another as in 4.12 and any position can be used to create different patterns and apparent velocity changes of the source despite the source moving at a constant rotational velocity. Later in this chapter, the loudspeaker's representation in the virtual field is generalized as a type of spatial object with a parameter curve (gain in this case).

4.3.7 Table Based Approach

While one way to use loudspeaker gain curves is to calculate the loudspeaker's gain value every buffer or sample according to the chosen panning algorithm, an alternative method is to use a lookup table for each loudspeaker where every position in the virtual source domain will have a gain value. Figure 4.13 shows the gain curve for loudspeaker 5 where N is the table size.

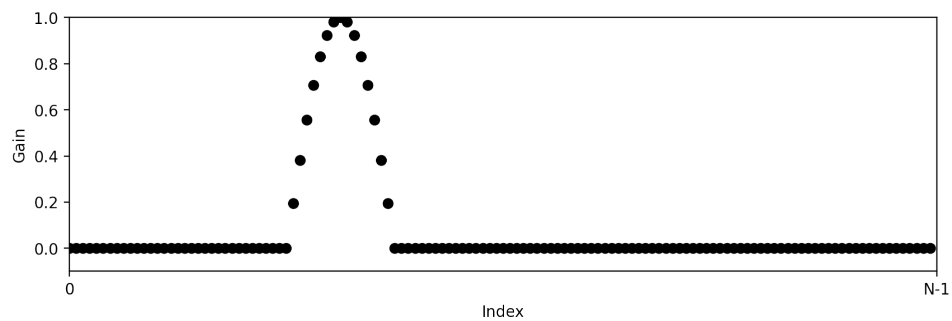


Figure 4.13: Table values for loudspeaker 5's gain curve.

Table values can be generated through a variety of methods. The most straightforward method is to generate these values by a panning algorithm. Other methods include drawing the values a graphic interface, using other functions, or the tables can be populated by other data. As long as the table values avoid large differences between adjacent values (differences that would cause audible discontinuities in the audio signal), any type of data can be used.

Figure 4.14 shows a loudspeaker gain table with additional peaks. Not only will the loudspeaker produce audio normally for source positions at and around the loudspeaker position, but the loudspeaker will also produce audio when the source is at other positions. If a source is rotating at a constant velocity, this will have a similar effect to adding a delay line to the source sound.

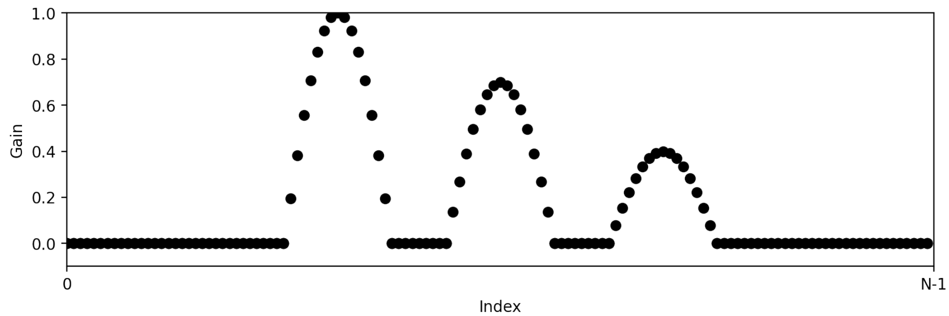


Figure 4.14: Table values for loudspeaker 5's gain curve showing additional peaks.

Loudspeakers can also use the gain table of another loudspeaker. If each loudspeaker has a pointer to a gain table, that pointer can easily be swapped to a different table. For example, a group of loudspeakers can all point to the same table making them all respond identically to the position of the virtual source. Also, reassigning loudspeaker gain tables can have the same effect as assigning false positions to loudspeakers as discussed previously. Using this method instead of false positions can be faster because the loudspeaker's gain curves do not need to be recalculated.

When using tables, values can easily be shifted by adding an offset to the read position. This also has the same effect as using false positions for the loudspeaker position, except that the shape of the table values will necessarily remain the same where they may not in arrays with varying distances between loudspeakers. The width of the gain value shape can be changed by varying the read position step relative to the position of the loudspeaker. The height can be changed by multiplying the table values by a scaling factor. Changing the width and height in this way has the same effect as changing the gain curve height as discussed in the previous sections.

Table based implementations allow for other methods of changing width, height, and shape. Instead of shifting and scaling the table mentioned above, multiple tables can be used for each loudspeaker. The active table can be instantly changed, or interpolation can be used to transition between two tables. When the table values represent gain curves for a current and target density layout, interpolation from the current to the target table avoids the requirement that a source must be moving in order to ensure the density change can be completed as discussed in Section 4.2.1. Additionally, using interpolation between a current and target gain curve table makes the time for the transition known. With the methods discussed in Section 4.2.1, the time it takes to switch gain curves depends on when a moment arises where the gain curves of a loudspeaker's current and target are not producing audio. When interpolating between current and target gain curve tables, the time it takes to transition is simply the interpolation time and does not depend on the position of the sources.

Using gain curve tables also opens up the possibility for interpolating between different panning algorithms. As different panning algorithms produce various gain curves, transitioning from one panning algorithm to another is only a matter of interpolating between two tables.

4.4 Virtual Spatial Objects

A spatial audio field contains many different types of objects. Some of these objects are physical such as loudspeakers and boundaries of the space, while others are virtual. A virtual spatial object refers to an object that exists in the virtual field. A virtual source is a type of spatial object that exists in the virtual field. This section will discuss several more types of virtual spatial objects that are positioned and moved in the same way as a virtual source. Additionally, this section will generalize the decoupling of loudspeaker position and its gain curve discussed previously as the decoupling of a spatial object and its parameter curve or curves. Spatial objects and their parameters including position, width, trajectories, and parameter curves have the potential to create dynamic and complex virtual fields.

4.4.1 Spatial Modifiers

Spatial modifiers directly or indirectly change the source in some way when the source is within the field of the modifier. Direct changes include the the amplitude scaling of a source from the gain curves and other signal processing such as effects. Indirect changes occur when a modifier changes another modifier object.

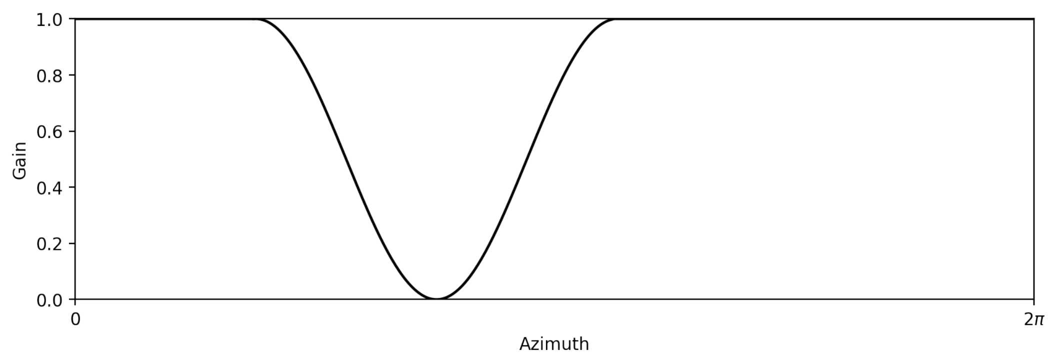
Anti-Loudspeaker

The *anti-loudspeaker* is a spatial object that has a position, width, and can be moved in the same manner as a virtual source. When a loudspeaker is within the anti-loudspeaker's field, it is disabled causing the density of the loudspeaker array to change. Thus the anti-loudspeaker position and width determine the part of the virtual field where the loudspeaker density is low. Because the anti-loudspeaker results in a binary change to a loudspeaker's enabled status, it does not have a parameter curve unlike other modifiers.

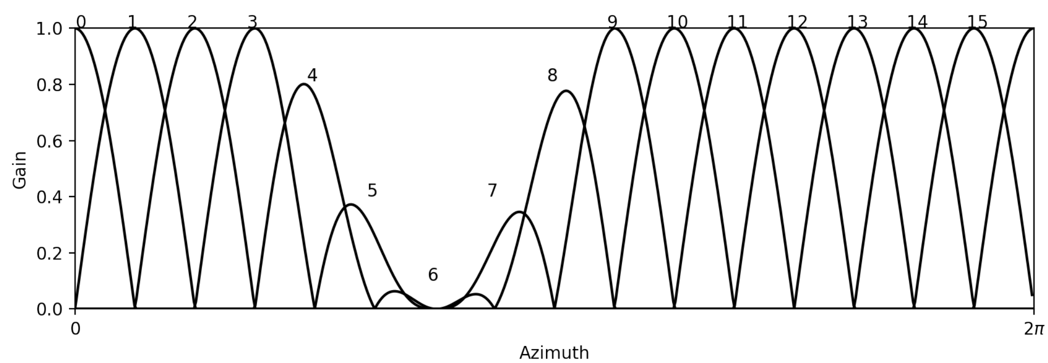
Anti-source / Virtual Drain

As a virtual source adds sound into the field, a virtual *anti-source*, acting as the counterpart to a virtual source, removes audio from the virtual field before it can be projected by the loudspeakers into the acoustic field. The most basic type of anti-source mutes a loudspeaker in the same way that the anti-loudspeaker disables a loudspeaker. The width of an anti-source that mutes loudspeakers within its field can be increased by muting adjacent loudspeakers. However, the width only changes the gap in the virtual field in fixed increments based on adjacent loudspeaker positions.

Finer control of the anti-source's effect, and higher level control of the loudspeaker gain curve heights can be established using what will be called here a *virtual drain*. Figure 4.15 shows a virtual drain applied to the gain curves of the example loudspeaker array. The drain has several parameters including the position, width, depth, and shape.



(a)



(b)

Figure 4.15: Virtual drain (a) applied to the gain curves of the example loudspeaker array (b).

Positioning a virtual drain in the field creates areas where the source is attenuated or silenced completely. Adding a trajectory to the drain, or driving the position, width, or depth of the drain with a function, can create complex interactions between a moving source and the field.

Gain Curve Windowing

While the anti-source attenuates the loudspeaker gain curves within its field, a window can be applied to the loudspeaker array's gain curves to do the opposite. As with the virtual drain, the gain curve window has a position, width, and shape. Although here instead of depth the window has a gain/amplitude. Unlike the virtual drain however, a gain curve window affects the entire array because all gain curves are zeroed outside the positive values of the window. Alternatively, if the desired result is to only window a portion of the loudspeaker array, the tails of the window can be curved back to 1.0 creating a drain-window hybrid.

Effects

The anti-loudspeaker, anti-source, and gain curve window are types of modifiers that attenuate the source signal by changing the loudspeaker gain curves within their field (or outside of their field in the case of a gain curve window). However, any type of signal processing or effect such as a filter, delay, pitch shift, distortion, and similar can be used. Just like a virtual source and modifiers to the gain curves mentioned above, the spatial effect modifier has a position, width and one or more parameter curves that span the width of the effect object's field.

Triggers and Zones

Trigger and zone objects are used to perform an action or change state. A trigger spatial object performs an action when a source passes through the trigger. This can be used for changing loudspeaker densities, gain curves, muting or unmuting the source, or any other action to be completed when the source passes through the trigger. A zone on the other hand is a binary switch where the zone is in one state when the source is within the zone, and another when it is outside.

4.4.2 Decoupling Spatial Object Position and its Parameter Curves

The position and width of spatial objects (such as anti-sources and spatial modifiers), determine which area in the source domain the modification will be applied to the source. In the example layout, if a spatial modifier such as pitch shift covers an area from loudspeaker 1 to loudspeaker 4, pitch shift will only be applied when the source is between loudspeakers 1 and 4. The parameter curve (pitch shift in this case) does not necessarily need to be in the same position as the modifier. While the spatial modifier might be located at loudspeakers 1 to 4, the pitch shift parameter curve could be located at loudspeakers 8 to 11. Separating the modifier position and the modifier's parameter curves spatially only makes sense with multiple sources. If there is only one source, the pitch shift value will change as the source moves through loudspeakers 8 to 11, but the pitch shift is only applied when the source is between loudspeakers 1 to 4 and a change will never be heard. With multiple sources, if source 1 is between loudspeakers 1 and 4, and source 2 is moving through loudspeakers 8 and 11, the position of source 2 will change the value of the pitch shift being applied to source 1.

Two things should be considered when decoupling modifier position and its parameter curve. With multiple sources, two or more sources could be simultaneously trying to modify the same parameter value. One way to handle this situation could be to simply average the values together. The other thing to consider is that a source might be in the middle of a parameter curve while another source enters the area of the modifier causing an abrupt application of the modifier. A simple solution to this scenario is to add an additional parameter curve to the modifier that gradually mixes the unaltered signal with the altered signal at the ends of the modifier's field.

4.5 Conclusion

This chapter has discussed methods of changing the virtual field through manipulation of loudspeaker gain curves and virtual spatial objects. The density of the loudspeaker array can be changed by disabling loudspeakers and recalculating the remaining loudspeaker's gain curves. This allows one to use the tradeoff between localization accuracy and perception of smoothness of a source as a compositional element. Three methods of changing density were also discussed. Density changes manipulate loudspeaker gain curves relative to adjacent loudspeakers. This chapter also introduced techniques of manipulating individual loudspeaker gain curve's shape, width, and height. Additionally, the concept of treating loudspeaker gain curves as envelopes was discussed. Loudspeaker gain values are often calculated by the panning algorithm every buffer or every sample. This chapter proposed a table based implementation and discussed possibilities that arise using this approach. The last section in this chapter frames the virtual representation of a loudspeaker and its gain curve in the context of spatial objects and also discusses several other types of objects that are used to compose the virtual field. The focus of this chapter has been on methods of creating and manipulating the virtual field. The next chapter will discuss compositional applications of these methods.

Chapter 5

Creating Compositional Elements with Gain Curves

5.1 Introduction

This chapter discusses the compositional applications of the methods introduced in the previous chapter. These methods include extensions to synthesis, the use of gain curves to create rhythm, extensions to granular and pulsar synthesis, and spatial extensions of Phase Music.

5.2 Spatial Modulation Synthesis

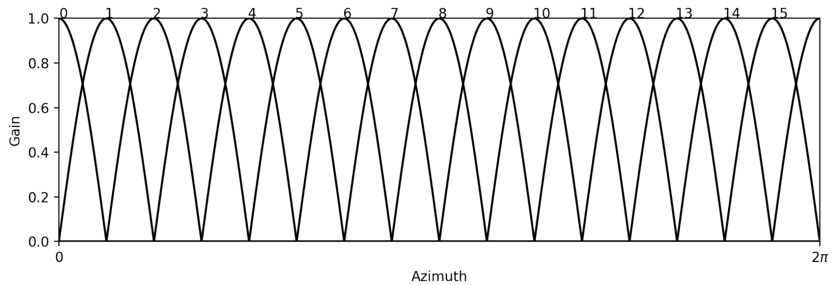
Spatial modulation synthesis utilizes fast moving sound sources simulating the physical changes to frequency and amplitude as a sound moves in the real world [34]. In [34] the focus is on spatial modulation synthesis using the Doppler shift of high velocity sound sources to produce frequency modulation (Doppler FM) while generating envelopes and amplitude modulation from distance based gain attenuation.

In addition to simulating distance based gain attenuation to produce amplitude modulation, the inherent amplitude modulation produced as a result of a virtual source moving in a spatial audio field at a fixed distance can be utilized to create amplitude modulation. This section will describe two methods of producing amplitude modulation resulting from the interaction of a moving source with the gain curves calculated by the panning algorithm.

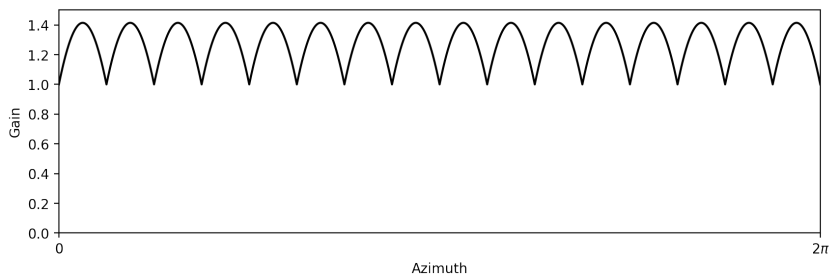
Virtual Source Rotation / Power Modulation

In traditional amplitude modulation, a carrier signal is multiplied by a unipolar modulator [7]. If we consider the example loudspeaker array with a single source rotating counter clockwise at a constant rotational velocity, the amplitude of the source is being modulated by the loudspeaker gain curves. In effect, the loudspeaker array's gain curves are serving as a modulation signal. Figure 5.1 shows the example layout's gain curves and the sum of the loudspeaker gains at each azimuth. However, each gain curve only applies to one loudspeaker and the amplitudes do not typically sum the way that is shown in Figure 5.1b. Amplitude modulation that produces additional frequency content was not observed to be produced by adjacent loudspeakers. On the other hand, with high velocity virtual sources, a single loudspeaker's output through several source rotations produce additional frequency content similar to pulsar synthesis (discussed later).

However, power modulation can be described in terms of the modulation occurring among adjacent loudspeakers. In an ideal setup, power is constant using constant power panning, resulting in the illusion of a sound moving from one loudspeaker to the other. Figure 5.2 shows several graphs of the loudspeaker array power at every azimuth at several different gain curve widths. Thus narrowing the loudspeaker array's gain curve width results in power modulation. The modulation depth is determined by the gain curve width and the modulation frequency is a function of loudspeaker density and source



(a) gain curves of the example layout



(b) sum of loudspeaker gains

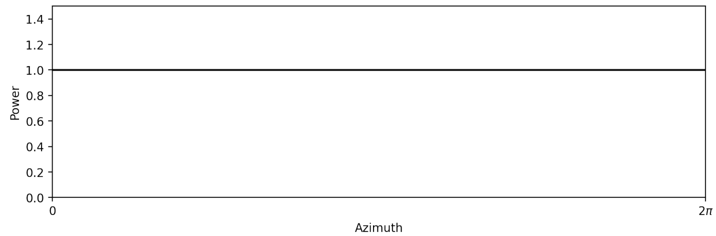
Figure 5.1: Gain curves of the example layout (a) and the sum of all loudspeaker amplitudes per azimuth (b).

rotational velocity.

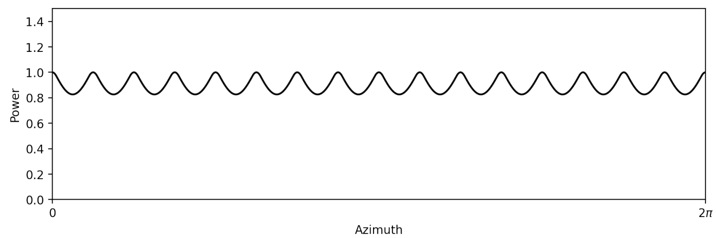
Function Driven Source

The previous section focuses on power modulation as a virtual source rotates around the example layout. There, the position of the virtual source is essentially being modulated by a sawtooth wave with a range from $0 - 2\pi$ whose frequency determines the rotational frequency of the virtual source. Simple translations and scaling of this position signal can produce different patterns of source movement. Using waveforms such as a sine wave for the position signal result in constantly varying source velocity.

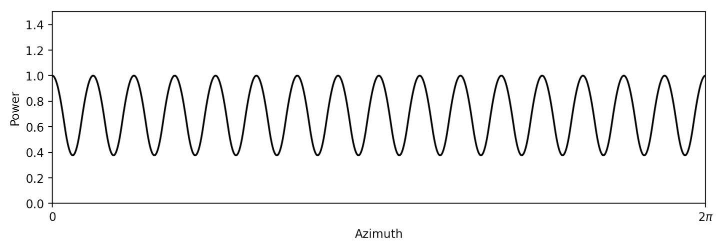
The parameters for the position signal are as follows: center position c_p, amplitude a_p, frequency f_p , and waveform. The center position is the azimuth at which the virtual source oscillates. In the example layout where the position is defined by azimuth, a_p



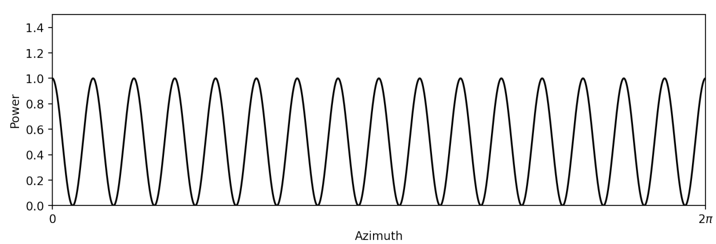
(a) width = 1.0



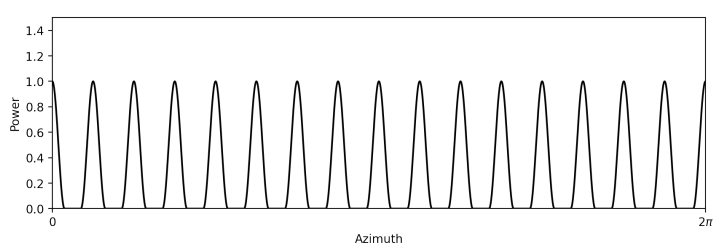
(b) width = 0.9



(c) width = 0.7



(d) width = 0.5

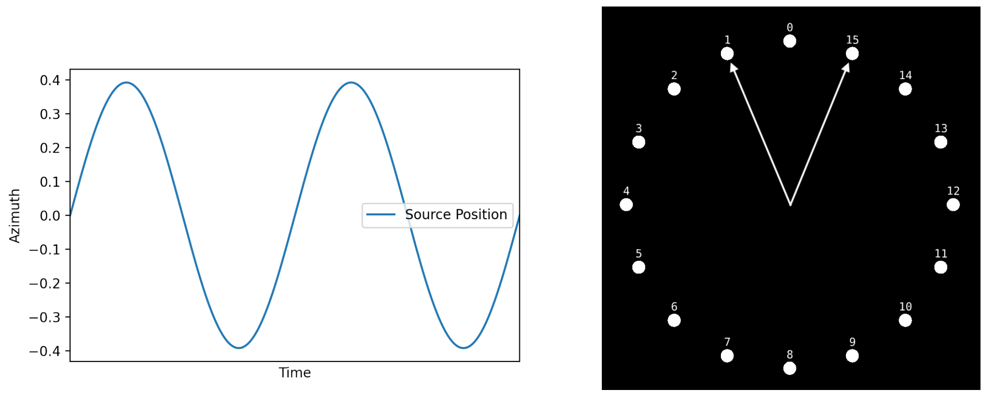


(e) width = 0.3

Figure 5.2: Power calculated for several gain curve widths.

determines how much the azimuth changes around the center position. c_p is the repetition rate of the periodic change in azimuth and the waveform can be any continuous signal. It is important that the signal is continuous in order to avoid discontinuities in the source position which would lead to discontinuities in the audio output signal.

Figure 5.3 shows a sine function determining the position of the virtual source in the example loudspeaker layout. Here, $c_p = 0$ and $a_p = azimuth = \frac{\pi}{8}$



(a) (b)

Figure 5.3: Source position signal (a) and a top-down view of the example layout indicating limits of the source position's movement (b).

$2a_p$ of 5.3a determines the arc length whose endpoints are indicated in 5.3b. If a_p is increased, the source will continue to oscillate around $c_p = 0$ (loudspeaker 0) but will go beyond loudspeakers 1 and 15. Shifting the graph in 5.3a vertically by changing c_p rotates the arc that the source follows while keeping the arc length the same.

The carrier signal is independent of the spatialization, and its parameters (frequency f_c, amplitude a_c, and waveform) can be configured at will. However, a modulator signal is created by the interactions between the source position and the loudspeaker's gain curves over time. Figure 5.4 shows the gain curves of the loudspeaker triplet from 5.3b.

With the source's position being driven by the function in 5.3a, it passes through the gain curves in 5.4 producing a modulation signal for each loudspeaker. This modulation

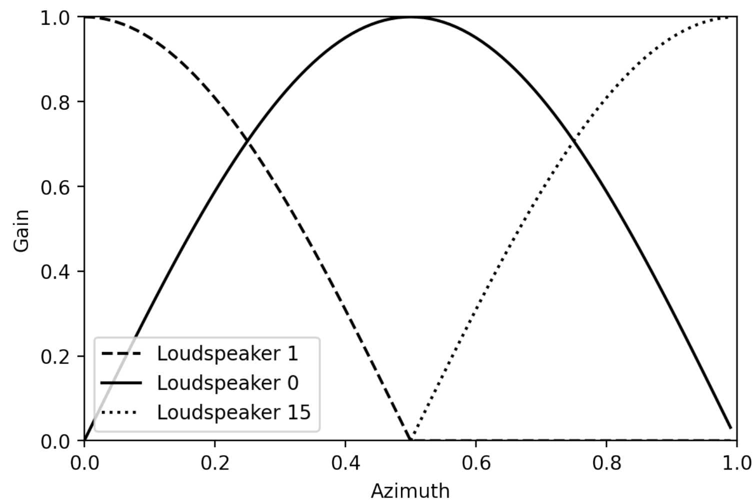


Figure 5.4: Gain curves of loudspeakers 1, 0, and 15 (azimuth normalized).

signal is then applied to the carrier signal (source content) resulting in amplitude modulation synthesis if the frequency of the source position signal is large enough. Using the position signal from 5.3a, Figure 5.5 shows the modulation signal for each of the three loudspeakers in the lefthand column and the corresponding output signal in the righthand column. The output signal shown is the modulation signal multiplied by an arbitrary waveform (sine) representing the source content.

In basic amplitude modulation synthesis where the carrier is a sine wave and the modulator is a sine wave with amplitude from 0 - 1, sidebands are produced at $f_c + f_m$ and $f_c - f_m$. When the waveform is a sine, only two sidebands are produced. The spectrum produced using basic amplitude modulation is highly dependent on the modulator signal's waveform. Figure 5.6 shows the frequency spectrum for loudspeakers 0, 1, and 15 from the previous figures. In this figure, $f_p = 50Hz$, $a_p = \frac{\pi}{8}$, $c_p = 0$, and the waveform is a sine with $f_c = 500Hz$ and $a_c = 1$. In this example, the sidebands produced in channel 0's audio output are:

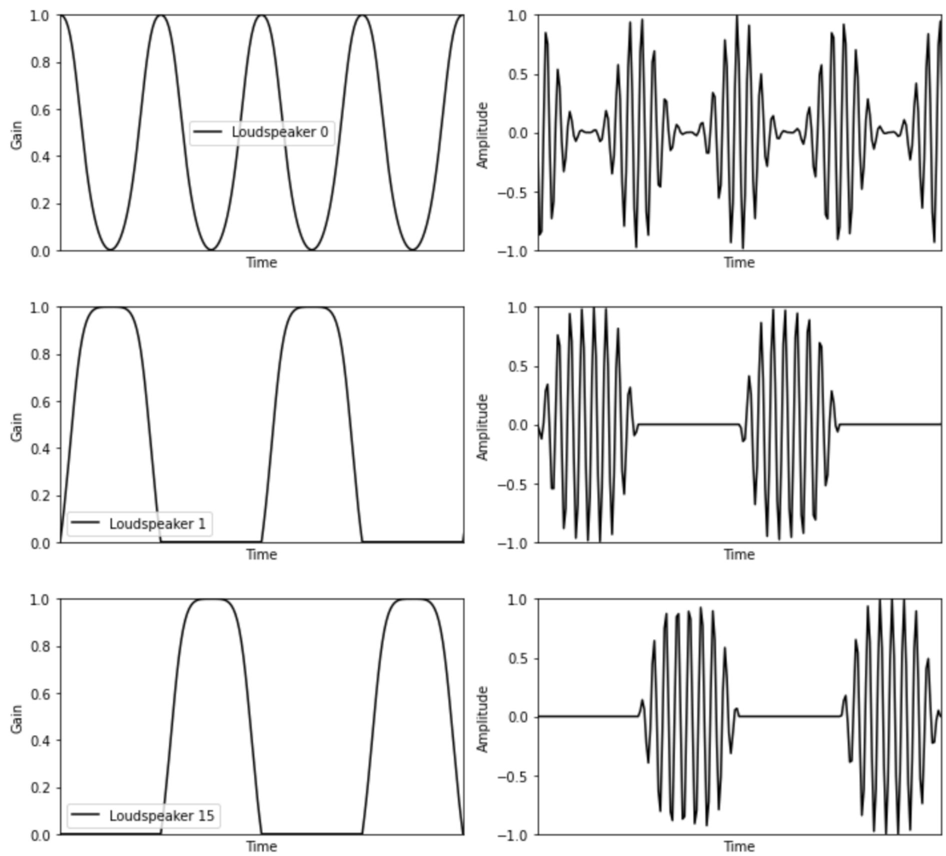
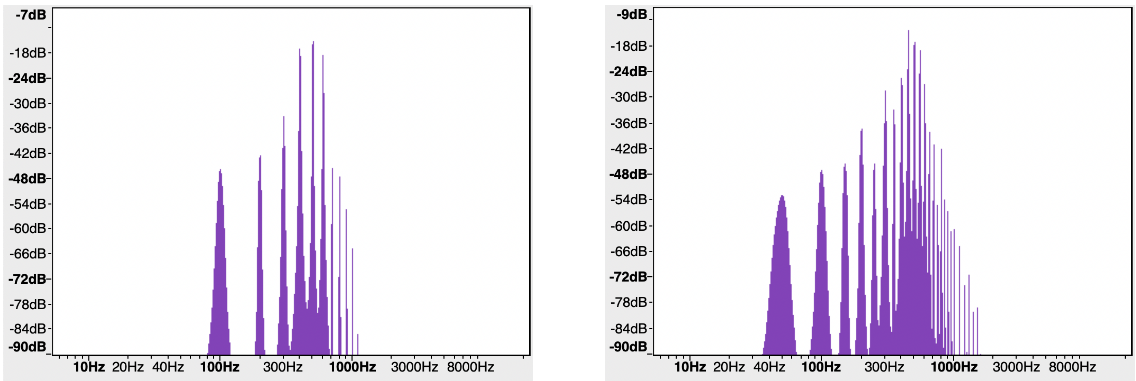


Figure 5.5: Gain over time of three adjacent loudspeakers (left) and the corresponding audio output (right) of the function driven source described in Figure 5.3.

$$\text{upper sideband} = f_c + n2f_p$$

$$\text{lower sideband} = f_c - n2f_p$$

In the above equation, 2 is used because the position signal is not the modulator signal, but the position signal *produces* the modulator signal. With the position signal centered around loudspeaker 0 ($c_p = 0$ radians), the source passes through loudspeaker 0 twice in one cycle. On the other hand, because $a_p = \frac{\pi}{8}$, the source passes into half of loudspeaker 1's gain curve, then reverses direction. This results in one modulation of the



(a) Channel 0.

(b) Channels 1 and 15.

Figure 5.6: Frequency spectra of the loudspeaker triplet. The spectrum for channels 1 and 15 are identical.

source being output on channel 1 per cycle of the source position function and produces the spectrum with more closely spaced sidebands. Given the position signal parameters used for this example, the same is true for loudspeaker 15. Loudspeakers 1 and 15 have an identical spectrum but their waveforms are out of phase. Any waveform can be used to drive the source position. For example, a triangle wave produces a modulator signal that is directly based on the gain curves of the loudspeakers, while a sine wave as in the previous example produces slightly distorted modulator signal as can be seen in Figure 5.5.

Translating and scaling the position signal also produces more complex modulator signal waveforms. In the example layout, as the amplitude of the position signal increases, more loudspeakers become involved adding complexity to the resulting sound. Different offsets of the position signal's center position and different position signal amplitudes can be used to create complex patterns of pulses coming from individual loudspeakers. In addition to complexity that can be created through manipulation of the position signal, additional complexity and variation can be created by changing the spatial audio field using the techniques described in previous chapters.

5.3 Rhythm

Regular layouts containing a moving source suggest repetition and rhythm - both fundamental components of music. If we again think of the loudspeaker gain curves as a windowing function applied to the source, we can use the absence or presence (muted / unmuted) of a gain curve, or the gain curve width and shape to create rhythmic patterns.

Notes and Articulation

Changing loudspeaker gain curve widths can produce different styles of articulation analogous to styles of traditional music articulation. The example layout with 16 channels suggests 4/4 or similar time that has 16 16^{th} notes per measure. With the gain curves calculated normally to produce a smoothly rotating source, this would look like a whole note or 16 individual notes with a *slur* marking. If the gain curve widths are slightly narrowed, but still overlapping, the source will appear to pulse as it moves through the loudspeaker array in a similar fashion as a *portato* style of articulation. When the gain curves no longer overlap, this results in a more *staccato* style of articulation. Figure 5.7 shows the loudspeaker array with several different gain curve widths and the corresponding notation.

Rests and Dynamics

While changing loudspeaker gain curve widths create different note lengths, muting a loudspeaker replaces a note with a rest. Rhythmic patterns are created by muting patterns of loudspeakers. Changing the widths of the unmuted loudspeakers create further note-length variations. Additionally, the heights of the unmuted loudspeakers correspond to accents and dynamics. Variations in gain curve heights further add rhythmic variation to the measure. Figure 5.8 shows a rhythmic pattern produced by variations in gain curve

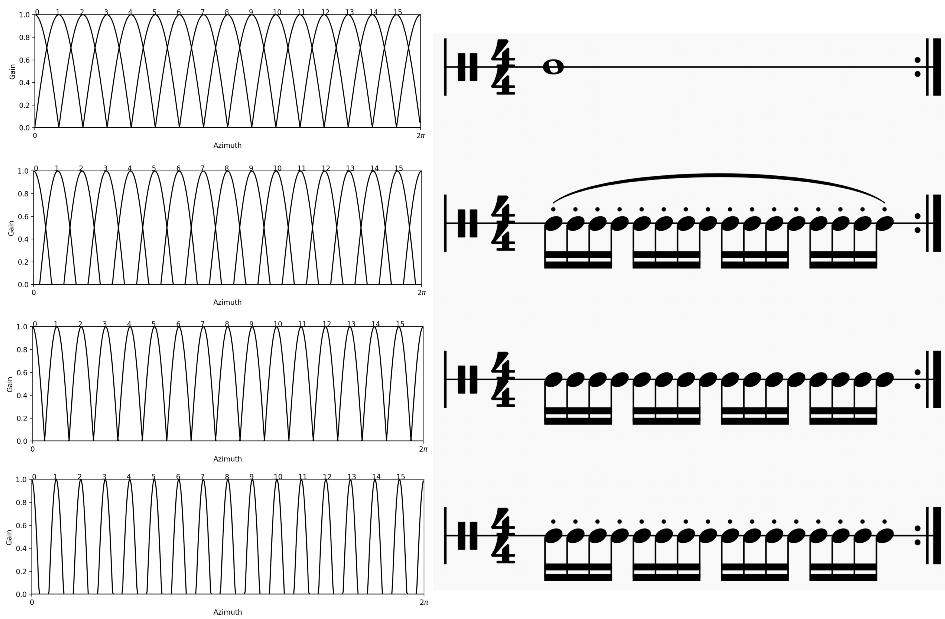
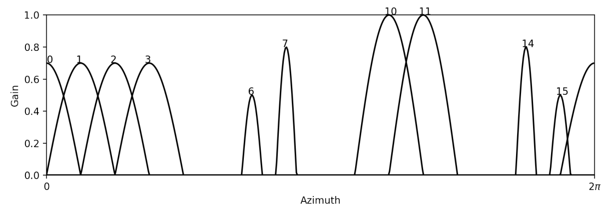
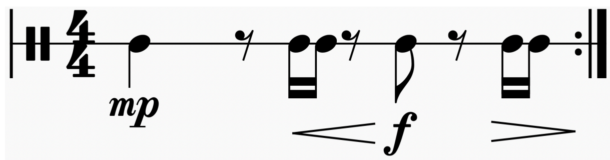


Figure 5.7: Loudspeaker array gain curves with different widths (left) and the corresponding notation (right).

widths, heights, and muting. In the figure, loudspeakers 4, 5, 8, 9, 12, and 13 are muted resulting in the rests in Figure 5.8b. With a rotating source in the example layout with the gain curves in the figure, the source will be slightly attenuated but will pan normally between loudspeakers 0 - 3. Loudspeakers 6 and 7 have narrow, non-overlapping widths with different heights forming a crescendo up to loudspeakers 10 and 11. Loudspeakers 14 and 15 have the same non-overlapping widths as 6 and 7 but the height values are reversed. Finally, note that loudspeaker 0's gain curve is present at the end of the measure. This is because it is assumed here that the beginning of a note occurs at the apex of the loudspeaker's gain curve. The lefthand part of loudspeaker 0's gain curve can be narrowed to produce a faster attack at the beginning of the measure when the measure repeats.



(a)



(b)

Figure 5.8: (a)Loudspeaker array gain curves with different widths, heights, and muting. (b) Approximation of (a) in standard notation.

Continuous Variation

If one rotation of the source corresponds to one measure, multiple rotations represent multiple measures. One approach to adding variation each measure is to change the gain curves after every rotation of the source. For example, the rhythmic pattern in 5.8 can go from producing the rhythm to a single note. By slightly increasing the gain curve widths to their maximum, unmuting muted loudspeakers, and bringing all of the loudspeaker heights to the same value slowly over time can create a transition from rhythm to a single note moving around the space.

Sequencing Rhythmic Changes

Changes to the loudspeaker layout can also be sequenced. Each measure can have a predefined pattern of muted loudspeakers, loudspeaker widths, and loudspeaker heights. Advancement of the sequence can be triggered by a clock, or more appropriately by a trigger placed in the virtual source domain. When the sources passes through the trigger the sequence advances one step.

Source Rotational Velocity

Further rhythmic variation can be added by changing the source's rotational velocity. Discrete changes to the velocity can be used to lengthen or shorten the duration of a measure, or the velocity can be reversed to reverse the rhythmic pattern. Multiplying the source rotational frequency by low integer ratios creates variations on a rhythmic pattern that will line up at specific points in time. With multiple sources, polyrhythm can be created where both sources positions meet at the beginning of the measure after common multiples of the fundamental rotational frequency. Velocity changes can be sequenced in a similar way as sequencing different rhythmic patterns mentioned above and can occur either at the beginning of the measure or at points within the measure.

In addition to discrete changes to a source's rotational velocity, continuous changes result in *accelerando* / *ritardando*. Thus with a limited amount of rhythmic material such as in Figure 5.8, many rhythmic variations can be created that are all related to the original rhythm.

The techniques described in 5.2 about driving source position with a function can be used to repeat subsections of the measure. If the function used will produce a discontinuity in the source position (e.g., sawtooth) an additional window such as a Hann window should be added to the source to prevent discontinuities in the output signal. If the function will not cause a positional discontinuity, the loudspeaker gain curves prevent output signal discontinuities naturally.

Gain Curve Derived Amplitude Envelope

The "notes" in the rhythmic patterns produced by muted loudspeakers and gain curve widths / heights are spatialized according to the loudspeaker positions. If the mapping of the virtual domain corresponds to the physical loudspeaker positions, then the position of each note will follow the previous note in space. However, by extracting the gain curves of the array and applying them directly to the source as an envelope (Figure 5.9), the spatial dependance of the source on the rhythmic pattern is removed. In other words, the source can remain at one location and still produce the rhythmic pattern. This opens up the possibility for the loudspeaker array to present a new pattern of gain curves while previous patterns are directly applied to the source. Layering two different patterns where the rates are set at a ratio can create complex polyrhythms, hemiola, or phasing if the rates are only slightly different.

Extracting the array's gain curves is as simple as sampling the loudspeaker gains at every position and putting that value into a table that is used as the amplitude envelope of the source. It is important to consider how the gains are combined because adding the gains can result in values over 1. The normally calculated gain curves in constant power panning are created to produce a constant power, not a constant sum of the amplitudes. Therefore, one method of creating the envelope from the array's gain curve is to add the square of each gain value at every position as is done in Figure 5.9.

With the source's amplitude envelope derived from the array's gain curves now defined in a table, the envelope can be applied to the source at different rates. Setting the rate at certain multiples of the base rate creates rhythms that will align metrically with the rhythm of the gain curve envelopes. An additional way to create rhythmic variation is to read the table backwards which reverses the rhythm of the source envelope, or to read subsections of the table. Similarly, the source's velocity can be reversed to reverse the spatial envelope. While the source's envelope can be created in other ways to produce a specific rhythmic pattern, deriving the source's envelope from the spatial envelope allows for the ability to musically reference that pattern in a single location or spread out over several locations.

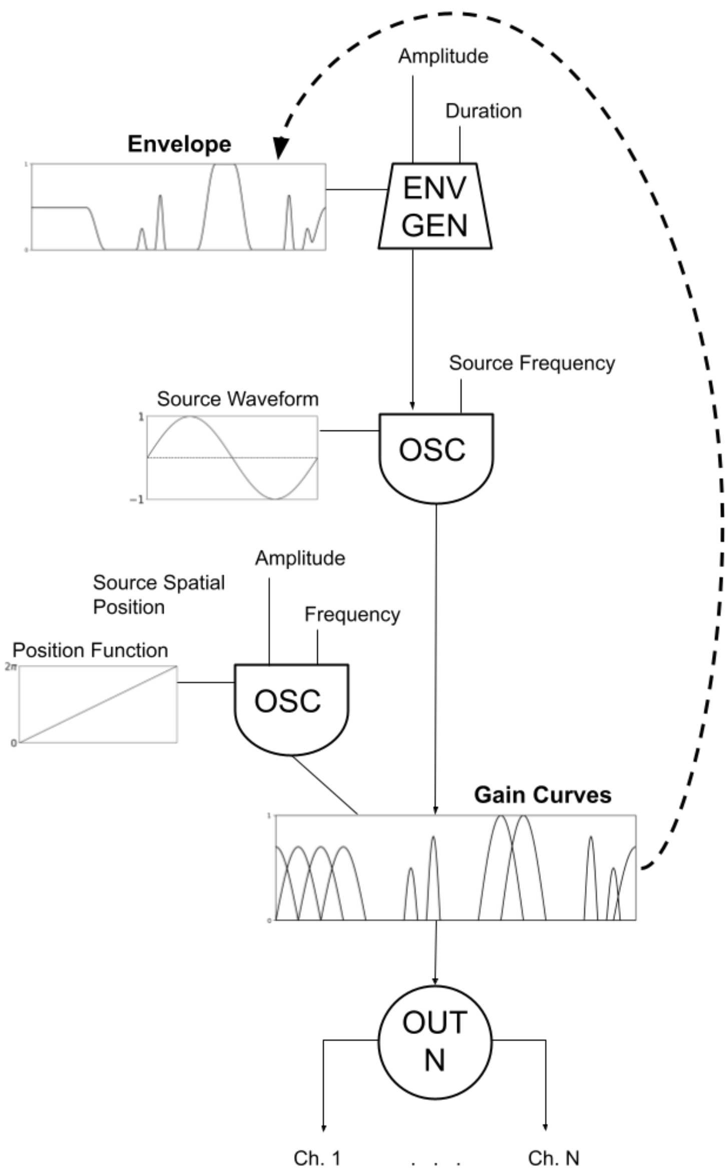


Figure 5.9: Gain curve extraction.

5.4 Field Based Granular Synthesis

Traditional granular synthesis creates sound objects by combining many microacoustic events (grains) over time [33]. This section discusses *field based granular synthesis* (FBGS) - a method of granular synthesis that constructs and spatializes grains through manipulation of loudspeaker gain curves in the virtual field.

In traditional granular synthesis, an individual grain is created by applying an envelope to a waveform with a duration from around 1ms to 100ms (see [7] and [33] for a comprehensive discussion on granular synthesis). Figure 5.10 shows a diagram of a basic grain generator. An envelope generator, consisting of the envelope shape, overall amplitude, and a duration modulates the amplitude of the oscillator in the figure. Note

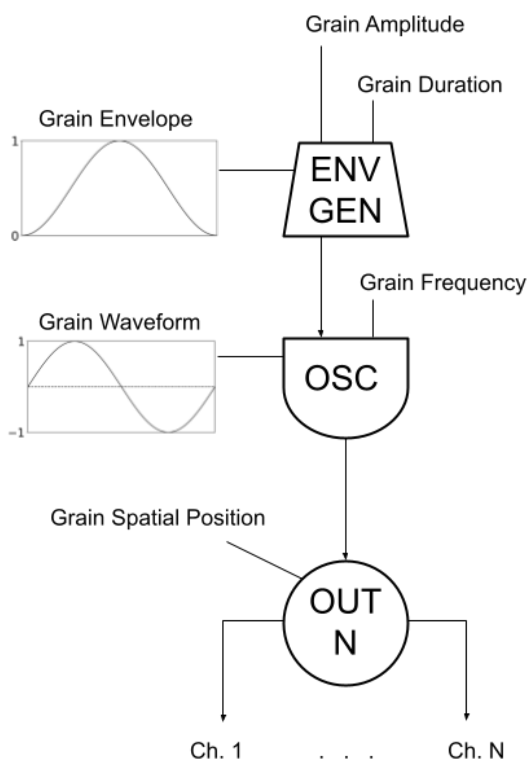


Figure 5.10: Diagram of a basic grain generator [33, p. 91 Adapted with permission].

that in this figure the grain is first constructed and then assigned a spatial position.

In FBGS, the construction of the grain and spatialization of the grain are interdependent. As discussed in Section 4.3, the height, width, and shape of loudspeaker gain curves can be easily changed. When gain curve widths are narrowed to a point where they no longer overlap, the field contains perforations or positions where no audio is produced. Given a rotating source in a perforated field, the source content is being windowed by the loudspeaker gain curves. In the context of FBGS, the loudspeaker gain curves are acting here as the grain envelope.

Field based granular synthesis uses parameters of the SAF to construct and generate grains. Figure 5.11 shows how grains are generated in FBGS and can be compared to the grain generator in 5.10. The first oscillator in 5.11(top) provides the source content.

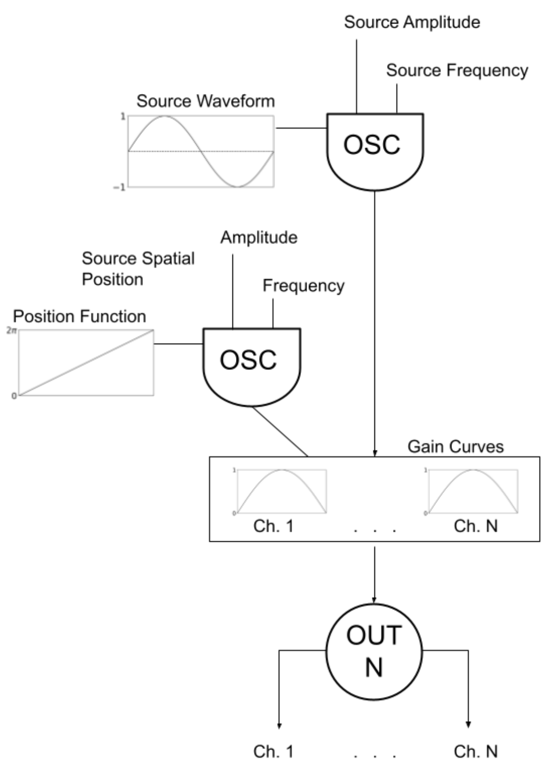
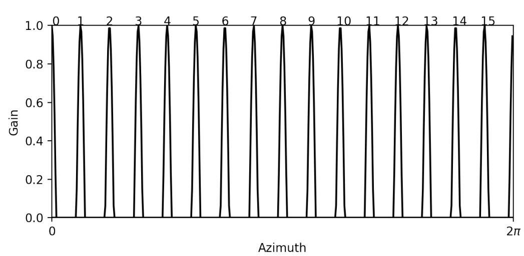


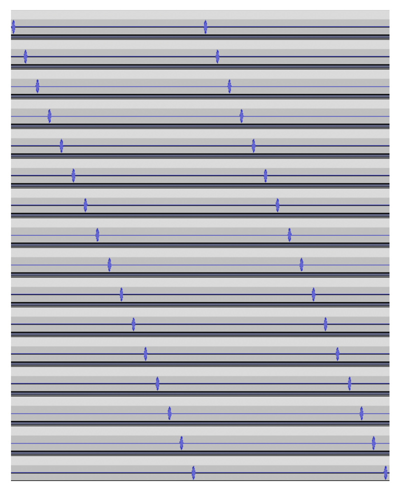
Figure 5.11: Block diagram of the grain generator in field based granular synthesis.

Note that an envelope generator is not connected to this oscillator as it is in Figure 5.10 and that the amplitude is a fixed value because the amplitude is modulated later by the gain curves as the source moves through. The second oscillator in the figure determines the source's position. A sawtooth waveform with range from $0 - 2\pi$ is shown here causing the source to rotate around the full layout. The waveform, frequency, and amplitude of the position function can be changed as described in previous chapters to create different trajectories of the source.

Figure 5.12 shows the audio output for a source rotating in a ring layout where each gain curve does not overlap the gain curve of the adjacent loudspeakers.



(a) Loudspeaker array gain curves with non-overlapping widths.



(b) Audio output of channel 0 (top) through channel 15 (bottom).

Figure 5.12: Gain curves and audio output of the example layout. The source is rotating counterclockwise at a constant angular velocity and the source's content is a sine wave. Two full rotations of the source are shown.

Grain Envelope

The envelope in FBGS is the loudspeaker gain curve. As the source rotates around the layout, the source's amplitude is modulated by the loudspeaker's gain curve which is considered here as the grain's envelope. Chapter 4 discusses several methods of changing the gain curve shape, width and height in addition to creating asymmetric gain curves and treating the gain curve as an envelope.

Grain Duration

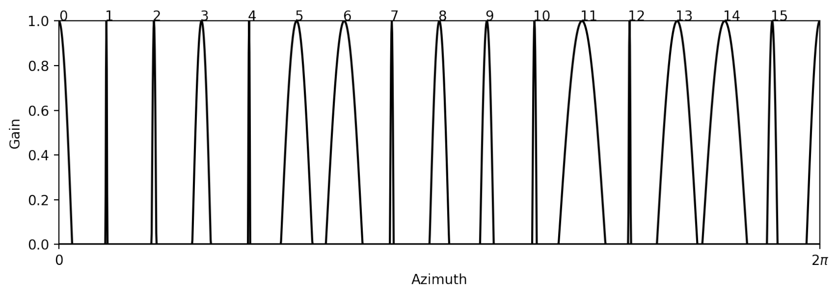
In traditional granular synthesis, grain duration is a parameter applied to a grain before being positioned spatially. In FBGS, grain duration is a function of source rotational velocity and gain curve width. Figure 5.13a shows an example of gain curve widths set to random values. Increasing source velocity makes the source pass through the gain curves more quickly, decreasing individual grain duration. Decreasing the gain curve width while keeping the source velocity constant also decreases the grain's duration. Grain duration can also remain fixed while a source changes velocity by changing the gain curve width.

Grain Amplitude

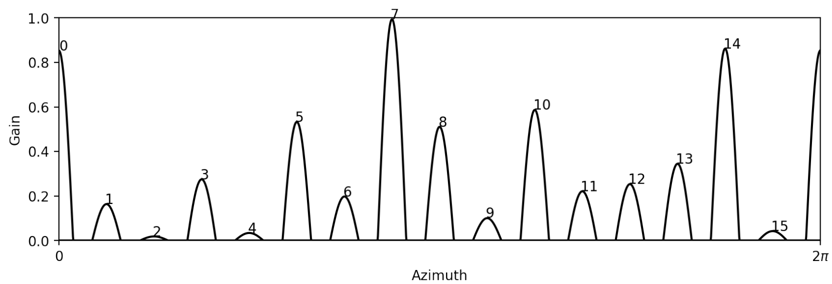
In Figure 5.10 of the basic grain generator, the grain's overall amplitude is determined by the grain amplitude parameter of the envelope generator. Just as the grain's envelope is the gain curve in FBGS, the grain's overall amplitude is the gain curve's height. Figure 5.13b shows the gain curves with a constant width and randomly assigned heights.

Grain Density

The grain density, or grain rate, is the number of grains per unit of time. In FBGS, the grain rate is a function of the loudspeaker density and how fast the source is moving (angular velocity in a ring layout). Decreasing the grain rate by decreasing loudspeaker density using the methods discussed in Chapter 4 results in the widths of the gain curves changing. When using that method, the grain rate decreases but the grain duration



(a)



(b)

Figure 5.13: Gain curves of the example array with random widths (a) and random heights (b).

increases. On the other hand, the grain rate can be decreased while keeping the grain duration the same by muting loudspeakers and keeping the gain curves of adjacent loudspeakers the same. Changing the grain rate through source velocity is straightforward. A faster source velocity equals a higher grain rate because the source is moving through the loudspeakers more quickly. Source velocity can be increased and the grain rate can be decreased to a certain extent depending on the number of loudspeakers in the array.

5.4.1 Synchronicity

In [33, p. 93], *synchronous granular synthesis* and *quasi-synchronous granular synthesis* are discussed. In FBGS, these types of synthesis depend on the symmetry and gain curves in the loudspeaker layout. With a regular symmetric ring layout, constant source velocity produces synchronous granular synthesis where each grain is produced one after another at a constant inter-onset-interval (IOI). Irregular layouts on the other hand produce quasi-synchronous granulation where the interval between grains varies. Modifying the SAF by changing loudspeaker density or muting loudspeakers allows one to go from regular layouts to irregular layouts changing the synchronicity of the granulation. Variations in IOI can also be created by modulating the source's velocity. When driving the source position with a function, the shape, amplitude, and frequency of the function determine the changes in acceleration which cause variations in grain IOI.

Repetition

In the paradigm of a source rotating at a constant angular velocity in the example layout, patterns of grain envelopes (i.e., the gain curves) are repeated throughout each rotation of the source. The spatial position and envelope of a single grain in the pattern, in addition to the sequence of the grains spatially, will remain the same each time the source rotates through the array. The patterns can be changed after each source rotation or after a number of source rotations. Additionally, changing the direction of the source will reverse the pattern adding additional variation.

Source Variations

Parameters of the source can be varied in the same way as is done in traditional granular synthesis. The source frequency and amplitude can be changed to values set randomly, stochastically, or based on other tables and functions. Another possibility is that frequency and amplitude values can be set up to be spatially dependent. In this case the values would be determined by the source's position.

When a source's velocity is low, a listener can track the sequence of grains spatially as a source rotates around the layout. As the source velocity increases, there is a point at which the listener can no longer track the position of the source. This transition from being able to track the source to not being able to track the source is similar to the transition from tremolo to amplitude modulation, or the transition from vibrato to frequency modulation. Even though the listener cannot track the position of the source, the spatialization of grains from a rotating source will sound different from a random spatialization of grains. Additionally, when the gain curves do not overlap, each grain is guaranteed to have only a single wavefront in FBGS making the precedence effect not an issue.

5.5 Field Based Pulsar Synthesis

In *pulsar synthesis*, a pulsar consists of a burst of energy (*pulsaret*) followed by a period of silence (see [33, p. 138]). The pulsaret is usually created by applying an envelope to any type of waveform and pulsar trains are created by repeating a pulsar. Just as field based granular synthesis uses the movement of a source through loudspeaker gain curves, field based pulsar synthesis uses the gain curves in the same way.

Creating a Pulsar With Gain Curves

In the example layout, a pulsaret is created by the loudspeaker gain curves as a source moves in the virtual field. Figure 5.14 shows an example where all loudspeakers are muted except for channel 1. Figure 5.15 shows the creation of a pulsar using channel 1's gain curve only. In Figure 5.15a, a_1 is the angle corresponding to loudspeaker 1's gain curve width and a_2 is the angle where loudspeaker 1's gain curve values are zero. A source rotating around the example layout with a constant angular velocity produces the envelope shown in 5.15b where d is the duration of the pulsaret, s is the interval of silence following the pulsaret, and p is the period of the pulsar. Unlike traditional pulsar synthesis, d, s, and p depend here on the source's rotational velocity and are related as

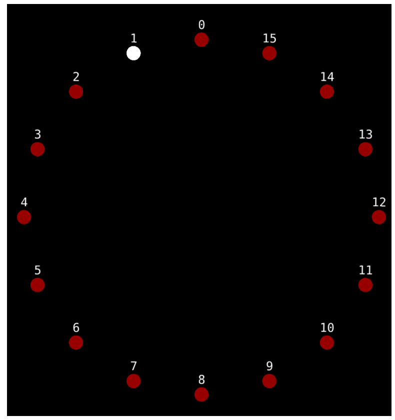


Figure 5.14: The example layout with all loudspeakers muted except for channel 1.

follows:

$$d = \frac{a_1(p_r)}{2\pi}$$

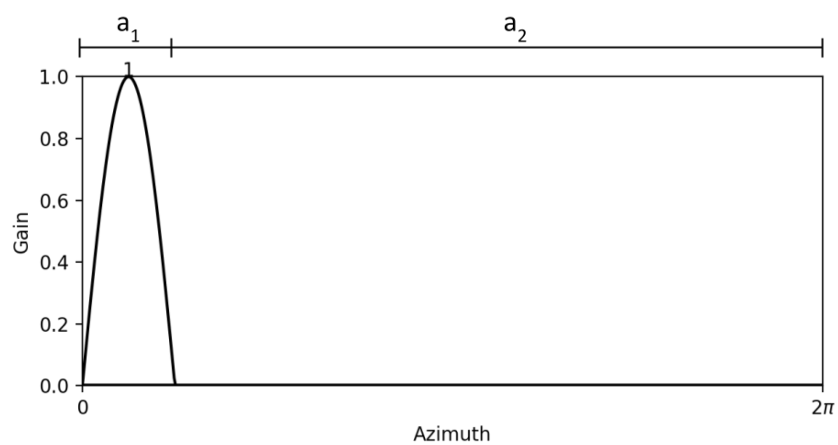
$$s = \frac{a_2(p_r)}{2\pi}$$

$$p = d + s$$

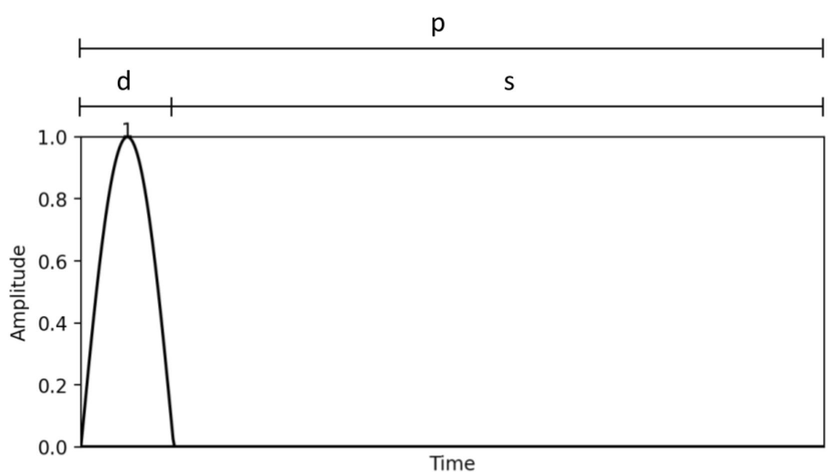
Where p_r is the period of one rotation of the source. Lastly, Figure 5.15c shows the pulsar created by the source's rotation around the example layout with a sine wave as the source content.

Modifying the Pulsar

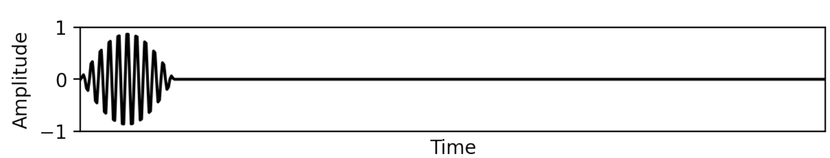
Because the gain curve is functioning as the pulsaret envelope, changes to the gain curve's width, height, and shape directly change the pulsaret envelope. The width of channel 1's gain curve is variable as usual and can be changed over time. With a source rotating at a constant angular velocity, changing the gain curve width changes the pul-



(a) gain curves



(b) pulsar envelope



(c) pulsar

Figure 5.15: Creation of a pulsar using gain curves. (a) All loudspeakers are muted except for channel 1. (b) Envelope of channel 1 produced by a rotating source with a constant angular velocity. (c) Resulting pulsar created from (b) with a sine wave as the source's content.

saret period while the pulsar period remains the same. Continuously narrowing or widening the gain curve results in continuous changes to the pulsar's spectrum. The gain curve height determines the amplitude of the pulsar and can be changed over time to produce increases or decreases in the pulse train's amplitude, or the height can be varied randomly or by any pattern. Finally, the shape can also be varied over time to produce variations in the pulsar's spectrum. The shape can be changed instantaneously to another shape (as long as the source is not within the loudspeaker's gain curve), or interpolation from one shape to another can be used to gradually change the spectrum over time.

Apart from changing the loudspeaker gain curves, changes to the source can be used to modify the pulsar. Increasing or decreasing the source's rotational velocity not only changes the pulsaret period as happens with changes to the gain curve width, but also changes the pulsar period. Instead of using a rotating source, driving the source position with a function can be used to change the emission pattern of the pulsar. For example, if the source's position function is a sine wave with center position at loudspeaker 1's position, the emission pattern will be constant. However, if the source position function's center position is not at the loudspeaker's position, the pulsar period will not be constant but the pattern of different pulsar periods will repeat if the source position function's parameters do not change. Varying the source position function's parameters results in variations in the pulsaret and pulsar periods.

The discussion so far has focused on the creation of a single pulsar through manipulations of a loudspeaker gain curve in the SAF. Essentially, there is little difference between this pulsar and a traditionally created pulsar that is positioned spatially except for the interdependence of parameters and the guarantee of a single wavefront.

5.6 Spatial Phase Music

Phasing in music is a process by which identical or related material is continuously repeated in two or more voices at different speeds or positions. Some of the precursors to phasing in music include the round where identical melodies are performed starting at different times. Additionally, rhythmic elements of phasing can be seen in various techniques such as hemiola and polyrhythm.

Traditional phase music uses two or more voices containing identical or related content performed at different speeds or positions in the loop. Despite limited musical material, many rhythmic, harmonic, and melodic variations emerge as a result of phasing. Phase music can be extended by spatializing the individual voices and further extended using the techniques of manipulating the SAF.

Terry Riley used repetition and the juxtaposition of different fragments of material in several compositions of both acoustic and electronic music. *In C* consists of 53 musical fragments of varying length to be played by any number of performers [41]. The fragments are to be played in order, however each performer can decide how many times they want to repeat any given fragment. Steve Reich is considered to be one of the pioneers of phase music and explains his interest and aesthetic in terms of musical processes in *Music as a Gradual Process* [42]. The tape piece *Come Out* [43] is an early example of continuous phasing applied to speech. *Piano Phase* [44] applies the phasing technique to acoustic music. Another example of phasing occurs in *Pendulum Music* [45]. Several microphones suspended over loudspeakers are pulled back and released by the performers. The sound that occurs is a result of the feedback generated as the microphones pass over the loudspeakers. In this piece, the phasing occurs because of slight differences in physical properties of each pendulum such as length, weight, release height, and release time.

Several spatial extensions to phasing will be discussed in this section in addition to how changes to the virtual field can be used in conjunction with these extensions.

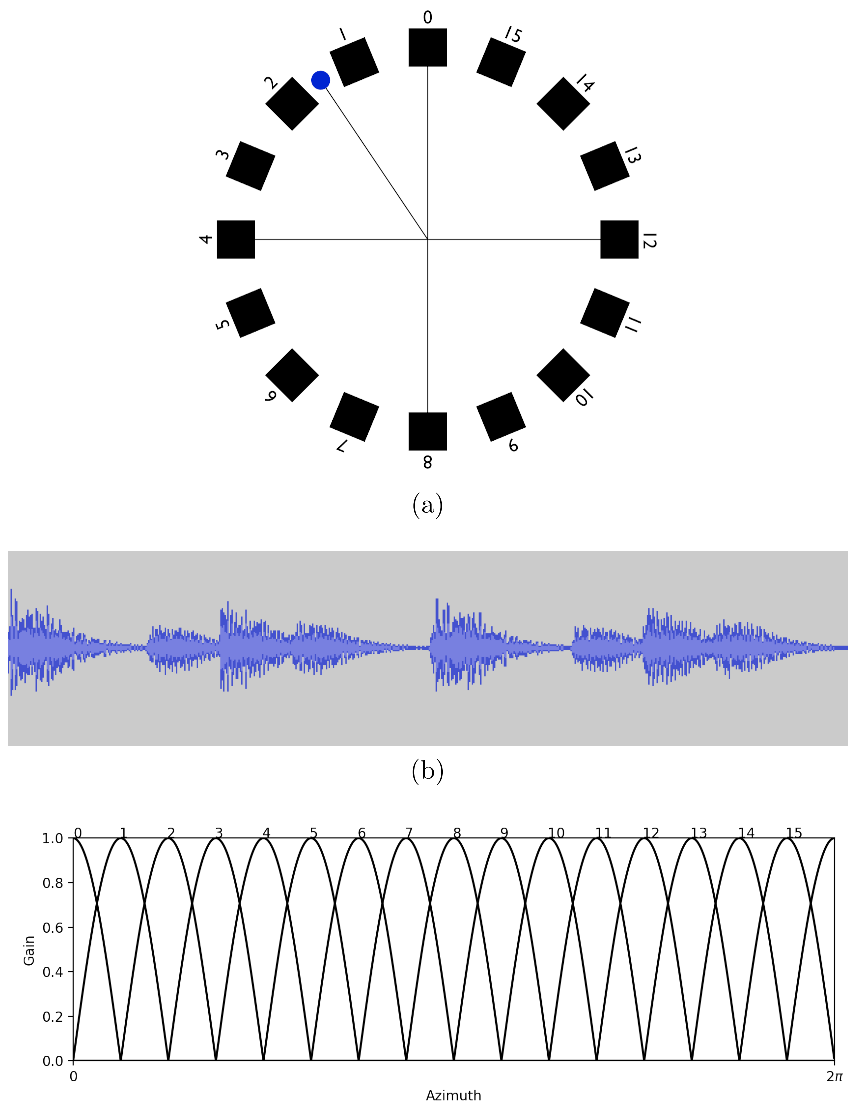
5.6.1 Loop Duration Versus Rotational Period

In traditional phase music, the phasing aspect refers to the position (phase) of one loop compared to another. In contrast, here we are comparing the repetition frequency of a loop compared to its position in space. As the most basic example, consider the example layout with a single virtual source rotating with a constant angular velocity (Figure 5.16a).

The rotational period of the source will be t and the duration of the loop the source is playing will be d. When $d = t$, the source loop is in phase with its position in space and every sample will occur in the exact same spatial position. On the other hand, if t is slightly greater than or less than d, the spatial position of the loop will begin to rotate clockwise or counterclockwise. While the musical interest of this basic example is limited, modifications to the loudspeaker gains can extend this technique.

Figure 5.17 shows the effects of muting every other loudspeaker on the loudspeaker array's gain curves. In effect, the loop is being windowed, and spatialized by the movement of the source through the gain curves. The window length can be modified by changing the gain curve widths of individual loudspeakers. As the loop rotates, different rhythmic elements can be heard as a result of the interaction between the loop content and the windowing of the gain curves.

At higher rotational velocities of the source, the amplitude modulation applied by each active loudspeaker begins to be heard as an audible frequency changing the timbre of the sound. The relationship between the loop length, source rotational period, loudspeaker density, loudspeaker density regularity, and window length can produce many timbral variations that can be used compositionally.

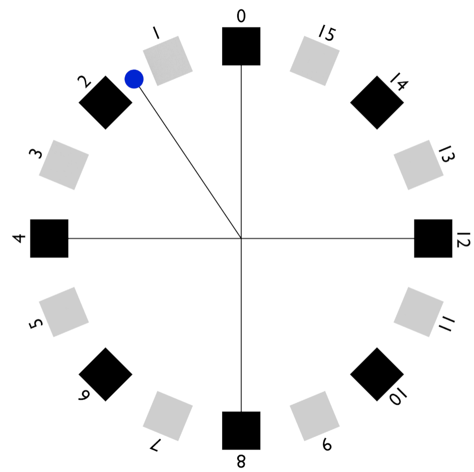


(a)

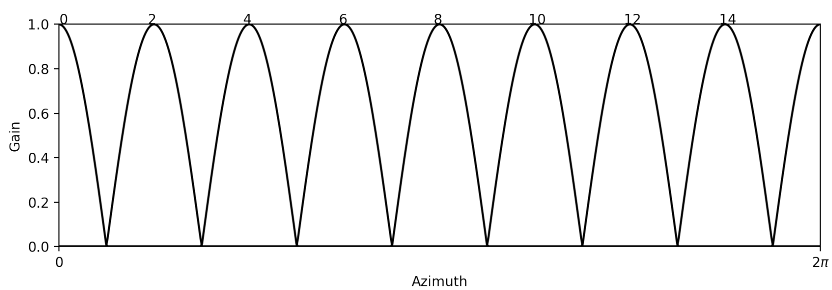
(b)

(c)

Figure 5.16: (a) Top-down view of the example layout showing a single virtual source. (b) Arbitrary waveform of duration d. (c) Gain curves of the loudspeaker array.



(a)



(b)

Figure 5.17: (a) Every other loudspeaker muted and the resulting gain curve graph of the array (b).

At a constant source rotational period, changing the loop duration will change the phase of the loop in respect to its spatial position. The loop duration can be changed by speeding up or slowing down the loop which changes the pitch. Alternatively, the pitch can be kept the same by time-stretching the loop using techniques of phase vocoding. Changing the loop duration can also be done by simply moving the start and end points of the loop.

5.6.2 Proximity

The proximity of two virtual sources playing the same loop at slightly different rates can be changed to modify how the phasing effect is perceived. Consider the example layout in Figure 5.18 with two virtual sources separated by angle *theta* playing identical loops at slightly different speeds. When $theta = 0$, the two sources are perceived as a

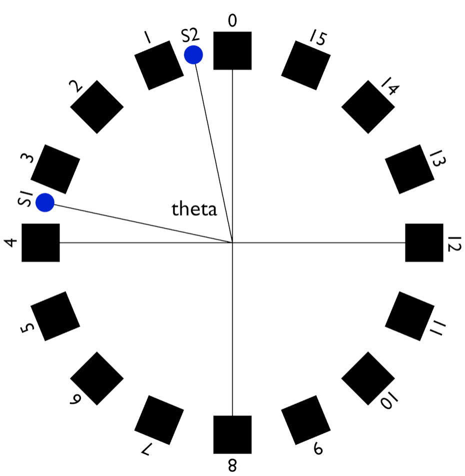


Figure 5.18: Virtual sources S1 and S2 in the example layout.

single auditory stream as in many traditional phase pieces. In the case where the source content is speech, there are times where the original words or phrase cannot be perceived due to effects of masking, or are perceived in different orders than the original. As theta increases, there is a point in which the sources are perceived as two separate auditory streams and the original phrase can again be heard. Conversely, decreasing theta will cause the streams to merge back into one perceptual stream. From a compositional point of view, separating the sources can serve as a way to reintroduce the original "theme".

5.6.3 Spatialized Windowed Phasing

In the methods discussed so far, an audio clip is looped in its entirety. However, a subset of the audio clip can be played by changing the start and end points of the clip or adding an envelope or window. In this method, two voices play a windowed subset of the content incrementing the window's position after each repetition. Figure 5.19 diagrams a Hann window applied to an audio clip for a single source.

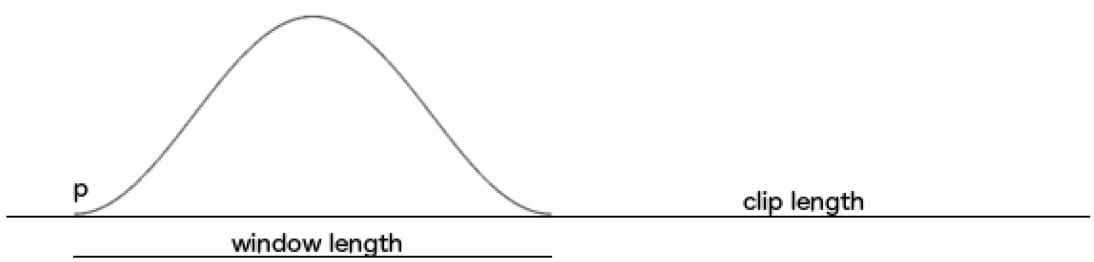


Figure 5.19: Hann window applied to a portion of an audio clip where p is the window's start position.

Parameters in this method of phasing include the window start position, window length, start position increment, and source position. The increment determines how many samples to shift the window after each repetition. An additional parameter determines whether to loop over the window or to loop over the length of the clip. Looping over the window produces continuous sound while looping over the clip will produce periods of silence.

This method of phasing has the potential to produce interesting effects especially when the sources are in different positions. For example, with a 50% overlap of the windows of two sources, a sound essentially pans between the source positions skipping loudspeakers in between. Additionally, if two sources have the same start position and window length, but one source has an increment greater than 0, the phasing effect will begin sounding like traditional phasing. However, as one source advances, the material of the static window will be recombined with different parts of the audio clip producing variations different from traditional phasing.

5.6.4 Phasing and the Virtual Field

Changes to the virtual field serve to highlight the types of phasing discussed in this section. In the first method of phasing with the loop duration versus a source's rotational period, the change in the loop's phase relative to the source's positional phase is difficult to follow. However, modifying gain curves at different positions create points of articulation. When a source passes through these points, the difference in the loop's position from the previous time it passed through the the points is more noticeable. When using the proximity of looping sources or the windowed method, changes to the gain curves height can attenuate a source shifting the focus to the other source. Additionally, a moving virtual drain will attenuate spatially separated sources at different times causing the listener's focus to change. Besides a virtual drain, other spatial objects can also be used to change how the phasing effect is perceived. Moreover, the phase of rotating virtual objects can be used in a similar manner as the first method discussed.

5.7 Conclusion

All of the compositional elements presented in this chapter are created from a source moving through loudspeaker gain curves. The techniques for manipulating gain curves discussed in the previous chapter are the means in which these elements are shaped and controlled over time. When the gain curves are calculated normally, their function is to create the illusion that a source is moving between two loudspeakers. However, when the gain curves are narrowed to the point in which they no longer overlap, they function as an envelope. This creates a continuum where the gain curve widths can be changed over time to go from the first function to the second or vice versa. Unlike traditional synthesis methods based on envelopes where the envelope duration is independently controllable, here it depends on the gain curve width and the velocity of the source.

Two types of source movement were discussed. The first and most basic type of movement is continuous rotation around the example layout. Low versus high source velocity determines if the gain curves produce rhythmic patterns or produce additional frequencies due to amplitude modulation. The second type of movement discussed is an oscillating source position. Here the parameters of the function driving the source's position can be used to keep the source in a specific area of the field and to expand or contract this area over time.

Even with basic changes to loudspeaker gain curves and basic source movement, this chapter has discussed several ways to create compositional elements and variation with simple modifications to the gain curves of the virtual field.

Chapter 6

Implementation

This section will discuss the software developed to implement and evaluate the concepts discussed in the previous chapters.

6.1 Introduction

This system is an experimental software application for realtime spatialization of virtual sources written to implement the broad concepts discussed in this dissertation. It is written in C++ using Allolib [46] and Gamma [47]. The software was originally conceived as a means to compare different panning algorithms through A-B comparison. This was then extended to comparing different loudspeaker densities, loudspeaker gain curves, and experimentation with new forms of composition and synthesis. Source content can consist of recorded audio, live audio, or generated waveforms. Each source can be positioned either individually or relative to other sources and spatialization is achieved through one of several panning algorithms including a point source with a specified width. Decorrelation can also be used to improve localization in several circumstances. This system also allows unique control over the loudspeakers used. The loudspeaker layout can be changed in realtime allowing for different loudspeaker densities and individual loudspeaker gains can also be changed in realtime. The system can also be used in realtime for live spatialization, a composition can be composed programmatically using the API, or the program's output can be recorded to a multichannel sound file for later playback or further assembling in an audio editor. The following sections will discuss the system's design.

6.2 Virtual Sources

The software includes several independent sources that can be configured in realtime. The number of sources is only limited by available processing power. The following will discuss the capabilities of the sources in terms of the source's content, spatial parameters, and other options.

6.2.1 Source Content

Three categories of source material are currently supported. The first consists of elementary waveforms (sine, saw, square, impulse) and noise. While plenty of interesting things can be accomplished with this category alone, elementary waveforms and noise are especially useful for testing and listening to frequency dependent variations in different panning and decorrelation methods. The second category is recorded audio. A sound file can be loaded into a source and played back in a loop. A subset of the loop can be played by specifying a starting point and a duration. Additionally, the playback rate of the sound file can be specified which is needed for phasing. The third category uses audio inputs to the program and each input can be mapped to separate sources. Using audio inputs allows one to generate and manipulate audio in another application before sending it to this application.

6.2.2 Spatial Parameters

Source Positioning

Sources can be positioned either individually or as a group. Specifying the positions of the virtual sources is done in relation to a source's azimuth in 2D. While position can be specified as cartesian coordinates, azimuth is more intuitive for ring layouts. To position sources as a group forming a spatial chord, each source has an azimuth offset from a certain position. These offsets can also be scaled allowing the group to expand or contract to and from a single location.

Source Movement and Trajectories

Source movement is currently limited to continuous rotation of the source around a ring or semicircular layout, or by a function driving the source position in the same layout as discussed previously in this document. For rotation, the azimuth is continually increased or decreased by an amount derived from the rotational frequency parameter and the values are wrapped around $0 - 2\pi$. For setting source movement from a function such as a sine function, the source's center position along with the function's frequency and amplitude are all controllable parameters.

Processing of a source's position is done at the sample rate. The software originally supported both sample rate and buffer rate updating of the source's position. However, with moving sources at high velocities, the "zipper" effect makes buffer rate processing of the source's position unusable and a decision was made to remove this option.

The distances of the virtual sources are fixed at the distance of the loudspeakers from the listening position. It is also assumed that the listening position is at the center or the "sweet spot". Distance can still be simulated by applying distance effects such as distance attenuation and reverberation, but these must be applied to the source content in another application before using it in this software. Additionally, and because the sources are at a fixed distance, Doppler shift has not been implemented. With the listener at the “sweet spot”, there is practically no radial velocity of a source rotating in a ring layout or similar layout with equidistant loudspeakers relative to the listener.

Decorrelation

Each source is decorrelated independently from the next and the user can choose which sources to decorrelate. Decorrelation of the signal is done once per audio callback and there are two available methods used that are part of Allolib. The first method is the Kendall method [36] and the user can set the max jump and phase factor. Alternatively, the Zotter method [48] can be used and has several additional parameters such as delta frequency, max frequency deviation, max tau, start phase, and phase deviation also provided by Allolib. As with most parameters, the decorrelation method and its parameters can be changed in realtime allowing for easy comparison of different decorrelation configurations. However, changes to decorrelation cause a brief discontinuity in the audio output and this has not yet been resolved as of the time of writing.

6.3 Spatialization Algorithms

There are several available spatialization algorithms in the software and the user can swap algorithms in realtime. The first step regardless of the algorithm is to sort and group the loudspeakers into pairs. This is done for two reasons: It eliminates unnecessary computation while the program is running, and groups are necessary for algorithms like VBAP. During runtime, the specific loudspeaker pair is found, then individual gains of the loudspeakers are calculated for the given source position.

In linear panning, for each loudspeaker pair, if the source position is equal to one of the loudspeaker positions, that loudspeaker gets a gain of 1 and the other loudspeaker 0. If the source is between the loudspeakers, the loudspeakers' gains are set based on a linear relation. If the position is not equal to one of the loudspeaker positions, or not in between the loudspeakers, the next pair is tested.

The process is similar for the VBAP implementation. For a 2D layout, a loudspeaker pair is the active pair if the source position is equal to one of the loudspeaker positions, or if the matrix calculations for the loudspeaker pair and the source produces gains that are greater than or equal to zero (see [14]).

Another method uses a point source with a width or spread parameter. For this method, the spread is specified in radians from the source's center position. Unlike the implementation of linear and VBAP, source spread does not rely on groups. For each loudspeaker, it is determined if the loudspeaker lies within the spread of the source. Loudspeakers within this distance can either all have the same gain or the user can choose to smooth out the loudspeaker gains by applying a roll off based on that loudspeaker's distance to the source's center position. The latter approach allows sources to be moved around smoothly because loudspeakers do not suddenly start or stop emitting sound. If the source spread is less than the distance between adjacent loudspeakers, there will be silent gaps as a source moves around the field. Thus changing the source spread on the virtual source is a way of narrowing loudspeaker gain curve widths. Because the source spread is applied to every loudspeaker, it essentially changes all loudspeaker gain curve widths to the same value. When using this method, as a source moves around or changes spread, the number of loudspeakers included in the spread will vary. To compensate for this, the user can choose to normalize the loudspeaker gains included in the spread to maintain constant loudness. Another issue is comb filtering that results from nearby loudspeakers playing identical waveforms. In this case, decorrelating the source not only improves the comb filtering effect but also reduces the impact of the proximity effect.

The final spatialization method currently available positions the source at the nearest loudspeaker. In this method, the nearest loudspeaker has a gain of one and all other loudspeakers have gains of zero. This method works best for static sources and eliminates the precedence effect entirely because only one loudspeaker is producing audio.

6.4 Loudspeakers

Each loudspeaker has its enabled status and gain as controllable parameters. Enabling and disabling loudspeakers and recalculating the gain curves is done in realtime. Each loudspeaker also has an individual gain control that scales the gain calculated by the panning algorithm. This gain adjustment is useful for fine-tuning nuances of a particular loudspeaker setup, or for modifying the panning algorithm for a particular effect. The loudspeaker's gain control is the parameter that corresponds to gain curve height.

Higher level control of loudspeaker parameters is available through the use of loudspeaker groups, the virtual drain, and presets. Figure 6.1 shows the loudspeaker groups window. A loudspeaker group is simply a set of loudspeakers whose parameters can all

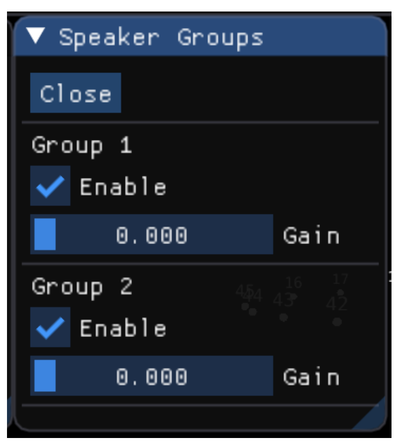


Figure 6.1: Screenshot of the Speaker Groups window.

be set to the same value. This is useful for control of loudspeaker density layouts where a predefined group of loudspeakers can be enabled or disabled together. In addition, the gains of the group can be changed together. Even though a loudspeaker might be in a group, individual parameters of that loudspeaker can still be changed in the main loudspeaker window, however, subsequent changes to the loudspeaker group's parameters override individual loudspeaker parameter values.

The virtual drain interface is shown in Figure 6.2. The drain depth determines how

Figure 6.2: Virtual Drain and Loudspeakers window.

much the gain curves are attenuated and the drain azimuth determines the position of the drain (the example layout is used here therefore the position can be specified in azimuth alone). The drain width determines the extent of the drain's field. Below the virtual drain window in Figure 6.2 is the loudspeaker window showing the effects of the virtual drain's parameter values on the gains of the loudspeakers. The shape of the virtual drain is an inverted Hann function.

In addition to loudspeaker groups and the virtual drain, there are presets for different loudspeaker density layouts and different configurations of muted loudspeakers. This allows the density and muting to be rapidly changed. The presets differ from using loudspeaker groups to change these parameters in that changes to groups are combined while presets override previous preset or group parameter values.

6.5 Time-Based Automation

The ability to create time-based gestures and automation allows the user to create more complex compositions from within the program. Currently, there is a rudimentary API for the user to create automations programmatically. Automations are grouped into *events* which are simply a group of parameter settings and automations. Events can be triggered individually or sequenced using a *master event*. Figure 6.3 gives an example of creating four events with one event serving to sequence the other three.

In line 1 of Figure 6.3, a pointer to a source is created. Any source or combination of sources can be used, but source 0 is used here. Next, four events are created and the event name is passed to the constructor. Automations are then added to event 1 in lines 10 - 14. Line 10 shows adding a breakpoint function where the arguments include a reference to the parameter to automate, the initial value of the parameter, and a list of target value / time pairs where the time is in seconds.

CreatingEvents.cpp

```
1  VirtualSource *vs = sources[0];
2
3  //Create some events
4  auto *event1 = new Event("Event 1");
5  auto *event2 = new Event("Event 2");
6  auto *event3 = new Event("Event 3");
7  auto *eventSequencer = new Event("Sequence");
8
9  //Add components to event1
10 event1->addBPF(&vs->sourceGain,0.0,{1.0,1.0,1.0,1.0,0.5,1.0});
11 event1->addBPF(&vs->centerAzi,0.0,{2.0,1.0,1.0,1.0,0.0,1.0});
12 event1->setParameter(&vs->enabled,1.0);
13 event1->setMenu(&vs->fileMenu,"shortCount.wav");
14 event1->setMenu(&vs->panMethod,"LBAP");
15
16 //Add components to event2
17 event2->setMenu(&vs->panMethod,"Source Spread");
18 event2->setParameter(&vs->scaleSrcWidth,1.0);
19 event2->addBPF(&vs->sourceWidth,0.0,{2.0,3.0});
20
21 //Create some speaker groups
22 auto *speakerGroup = new SpeakerGroup("Group 1");
23 auto *speakerGroup2 = new SpeakerGroup("Group 2");
24
25 speakerGroup->addSpeakersByChannel({0,2,4,6,8,10});
26 speakerGroup2->addSpeakersByChannel({1,3,5,7,9,11,13,15,17});
27
28 //Register the speaker groups
29 speakerGroups.push_back(speakerGroup);
30 speakerGroups.push_back(speakerGroup2);
31
32 //Add gain automation of speaker groups 1 and 2 to event 3
33 event3->addBPF(&speakerGroup->gain,1.0,{0.0,5.0});
34 event3->addBPF(&speakerGroup2->gain,1.0,{0.0,3.0,1.0,3.0});
35
36 //Add events to eventSequencer with the start time in seconds
37 eventSequencer->addEvent(event1,0.0);
38 eventSequencer->addEvent(event2,4.0);
39 eventSequencer->addEvent(event3,5.0);
40
41 //Register the events
42 events.push_back(event1);
43 events.push_back(event2);
44 events.push_back(event3);
45 events.push_back(eventSequencer);
```

Figure 6.3: Example code for creating automations with events.

Lines 12 - 14 show adding initial values to event 1 by passing a reference to the parameter followed by the parameter's value. Next, lines 22 - 34 show the creation of loudspeaker groups and adding gain automation to event 3. After this, events 1, 2, and 3 are added to the sequence event in lines 37 - 39 by passing the event variable followed by the event's start time in seconds. These time values are all relative to the start of the sequence event and not relative to the previous event. Finally, the events are registered with the system in lines 42 - 45.

Figure 6.4 shows the *event window* and *speaker groups window* generated by the code in Figure 6.3. An event can be triggered by clicking on its corresponding button in the event window or by sending a message via OSC [49] from an external program. Because of the way the sequence event is set up in Figure 6.3 lines 37 - 39, triggering this event will trigger event 1 immediately, event 2 after four seconds, and event 3 after five seconds.

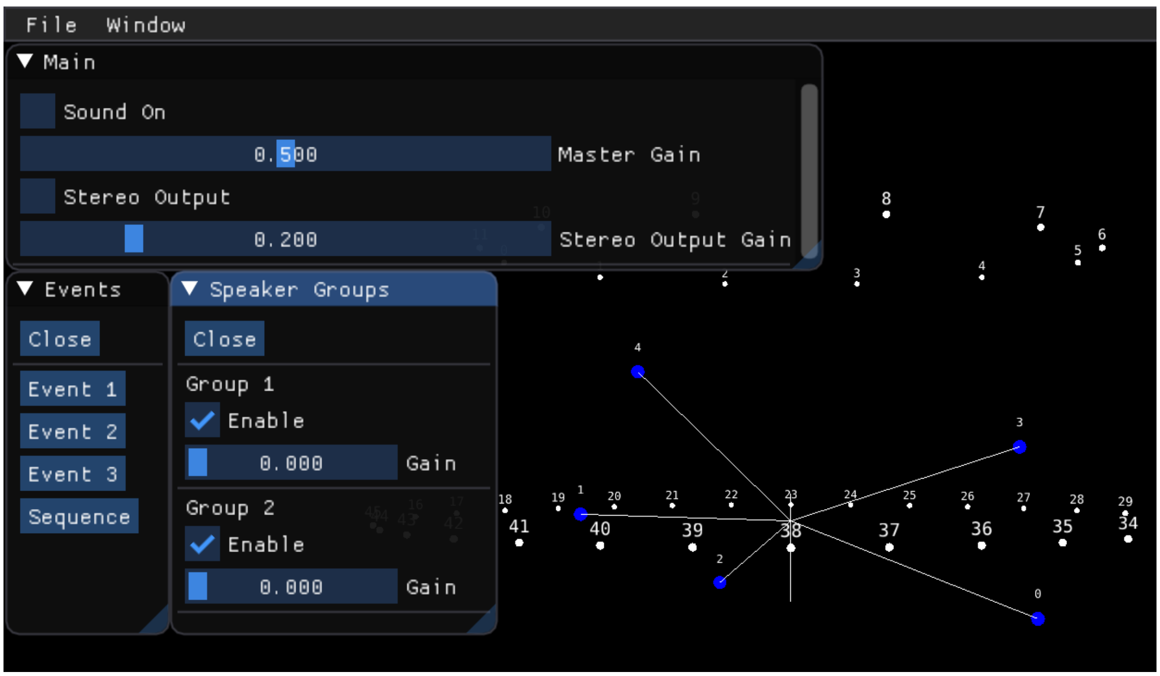


Figure 6.4: Screenshot showing the Events and Speaker Groups windows.

When creating events using the above method, the user must keep two things in mind. First of all, events can be run simultaneously. This presents an issue of parameter conflicts if two events are automating the same parameter. In this case, the parameter value of the latter event registered will override the former. Secondly, the user must keep in mind the behavior of parameters not specifically referenced in an event. Parameters that are not set in an event keep their value. For this reason, an event can play out differently if unused parameters are changed between runnings of that event. If this behavior is problematic, the user simply needs to add the offending parameter to the event's list of initial parameter values.

6.6 Program I/O

Inputs to the application are done programmatically, through the GUI, and optionally through UDP using OSC. Loudspeaker positions, number of sources, initial parameter values, events, and loudspeaker groups are all set up programmatically before the software is compiled and run. During runtime, the GUI can be used for changing parameter values. Alternatively, all parameters are exposed to external control through the use of OSC messages generated in another application. This application also has audio inputs that can be used as the sources' audio.

The output of the software program is an n-channel WAV file where n is the number of loudspeakers. The WAV file can be used to further edit the output in an audio editor, or it can be used to playback the output without having to use this software.

6.7 Future Development

This section describes the application's current limitations and plans for future development. Currently, each source has a panning algorithm option which determines the loudspeaker gain curves. This is contrary to the idea of SAFs where the source is separate from the field. Spatial objects such as effects, zones, and triggers are not implemented in the program itself, but can be created in an indirect way. For example, effect objects are created by sending the source position to another program generating the source content. When the position is within the effect object's field, the effect is applied to the waveform and sent to this program through its audio inputs. Additionally, loudspeaker positions are currently specified in the code. In order to change loudspeaker positions, for assigning false positions for example, the program needs to be recompiled.

Generalizing the source, loudspeakers, and modifiers as instances of a spatial object will make the program more flexible and easier to understand and use. The virtual domain can be thought as a canvas where spatial objects are created and positioned. The process of creating objects in the virtual field could be as simple as selecting "New Object" from a menu that instantiates a new spatial object. Then the user can select its type and configure its position and parameter curves.

There are several concepts discussed in this dissertation that are not directly implemented in the software. Currently, gain curve width is set globally and changes all of the gain curves to the same value. Gain curve shape can only be determined by the different panning algorithms provided and arbitrary shapes and ADSR style envelopes are not currently supported. Restructuring the code and generalizing the components of the virtual field to spatial objects will make integrating this functionality possible and intuitive to use. Also, the table base approach to gain curve manipulation is not yet implemented but is planned.

The biggest and perhaps most challenging improvement to the software is the extension to 3D loudspeaker arrays. 3D VBAP is already implemented in the code, but it is disabled because it conflicted with the way gain curves are modified in 2D. The current challenge is how to extend the concepts and techniques of manipulating gain curves in 2D to 3D.

There are several other general improvements to the software that are planned. First of all, parameter value interpolation needs to be added in some circumstances. Parameter values used for calculating waveforms, positions, and gain curves are updated at the sample rate. When these parameters are changed using continuous functions, potential discontinuities to the audio signal are generally not a problem. However, when the parameters are changed through the GUI, or externally from OSC messages, the change from the previous value to the next can cause audible discontinuities. Another planned improvement is the addition of persistence. Currently, options and parameters in the program can be initialized in the source code, or they can be sent to the program through OSC messages. Adding the ability to save and recall configurations from within the program will improve its usability. Another improvement planned is the visualization of the virtual field. Loudspeaker positions, source positions, and the size of the loudspeaker corresponding to the RMS value of the signal being sent to that loudspeaker are the only visualizations in the current version. Visualizations of the gain curves and parameter curves of spatial objects can assist in composing the virtual field in addition to assisting the listener in understanding the interactions of the spatial objects (if that is desired).

Chapter 7

Conclusion

When a source is spatialized it undergoes a series of transformations in the SAF before it reaches the listener. The SAF has a large impact on how the listener perceives a work. The example in Section 1.2 of performing a work in a studio versus performing the same work in a concert hall is an example of how different SAFs can drastically change a work. The SAF in a given performance is usually static. Given the impact of the SAF of the composition, how can it be changed and used compositionally?

7.1 Theory

To address the previous question, a theory of spatial audio fields was developed in order to identify parts of the field that can be manipulated on short timescales. While a SAF exists in any spatial audio implementation, it is defined here in terms of the virtual field, loudspeaker array, and acoustic field.

The virtual field, as it is constructed digitally, presents the greatest opportunity for developing new techniques on short timescales. The virtual field is the space where sound sources are transformed and processed based on the source's spatial parameters. It is constructed by the position and trajectories of spatial objects in the virtual field's domain. The spatial objects include the representation of the source and modifiers that transform the source when it is within their area of influence. Modifiers have a position and one or more parameter curves that produce values based on the sources position within the curve. The virtual representation of a loudspeaker and its gain curve is generalized as a type of spatial object that is not restricted to its physical position in the loudspeaker array. The mapping of the virtual domain to the perceptual domain is changed by moving the loudspeakers' positions in the virtual field.

The next step in developing the theory was to identify properties of the physical loudspeakers that modify the source as it is projected into the acoustic field. These properties are categorized as individual loudspeaker properties and properties of the loudspeaker array. Individual properties such as position and orientation can be changed when they are mounted on motorized tracks or gimbals. Changes to position, orientation, response, and directionality can also be simulated through signal processing carried out in the virtual field. Similarly, changes to the loudspeaker array's properties such as density, configuration, and dimensionality can be changed in the virtual field by manipulating the virtual loudspeaker objects.

The last step in developing the theory was to identify properties of the acoustic field that can be changed and used compositionally. These properties are broadly categorized as properties of enclosure and properties of the materials. Enclosure ranges from free field to completely enclosed and impact the resonance and reverberation of the space. The materials of the sound propagation medium determine the speed of sound and the materials of the boundary determine how much sound is absorbed versus how much sound is reflected back into the space. While the acoustic field can be constructed deliberately as a compositional choice, changing the acoustic field on short timescales is difficult. However, simulating an acoustic field is a common practice and changes can be carried out as fast as the sample rate.

7.2 Techniques

This research has resulted in the development of techniques for manipulating loudspeaker gain curves that include changes to the loudspeaker array density and changes to individual loudspeaker gain curves.

Several compositional techniques were developed based on changing the loudspeaker density by disabling or enabling loudspeakers and recalculating their gain curves. This allows one to use loudspeaker density for compositional choices. The tradeoff between the perceived increase smoothness of a moving source in a lower density versus the improved localization accuracy in a higher density can be used intentionally. The threshold at which amplitude modulation due to a source passing through a loudspeaker's gain curve is perceived can be changed by changing the loudspeaker density. Also, irregular densities can be created in an otherwise regular loudspeaker array. In order to use density changes on short musical timescales, methods of changing density while avoiding discontinuities in the output signal are described.

Apart from density, manipulation of individual loudspeaker gain curves create further opportunities for compositional elements. Changing the gain curve's shape, width, and height results in spatially dependent differences in how a source is altered by the virtual field. Furthermore, changing the gain curve shape has led to the proposed unification of separate panning algorithms through gain curve interpolation.

7.3 Applications

Using the above techniques, several compositional applications are described. By manipulating gain curves, spatially driven extensions to synthesis are described including power modulation, amplitude modulation, granular, and pulsar synthesis. The gain curves are also used to create rhythm as a source moves through the array and as a spatial extension to phase music.

The synthesis techniques developed based on manipulation of the virtual field are additions to and extensions of techniques developed by others. Additions to spatial modulation synthesis [34] include power modulation where the modulation depth is controlled by the width of the loudspeaker gain curves, and amplitude modulation that is based on a source passing through a loudspeaker's gain curve.

With a source rotating in a ring layout, rhythmic patterns are created by combinations of muting and changing loudspeaker gain curves. In this case, the gain curves are being used less for panning, but more as envelopes to individual notes. Even with a single rhythmic pattern, many variations can be derived by changing the source's rotational velocity and direction. Additionally, rhythmic patterns created by the gain curves can be extracted and applied directly to the source.

Using loudspeaker gain curves as envelopes also allows for spatially based granular and pulsar synthesis. In traditional methods, grains and pulsarets are created before being positioned in space. Here however, they are created by the source passing through the gain curves of the loudspeakers. Granular synthesis created in this way often involves high velocity sources that cannot be tracked by the listener. While it is possible to construct grains in the traditional way by applying different envelopes for each grain and then carefully sequencing their position, constructing grains by their movement through gain curves encourages thinking of a grain as dependent on its position. Additionally, grain generation based on gain curves encourages the creation of patterns of envelopes and emission over time. Many variations of a particular pattern can be created simply by changing the source's velocity. At high enough source velocities, the amplitude modulation due to individual gain curves begins to produce audible frequencies. With non-overlapping gain curves, pulsars are formed and controlled through the interaction of the moving source with the gain curves.

Finally, three spatial extensions to phase music are described. The proximity of two phasing sources is used to perceptually segregate or unify the audio streams. Another technique uses a single rotating source where phasing refers to the phase difference between the loop's duration and the source's rotational period. The last method uses two or more sources where a moving window is applied to the loop. In each of these extensions, modifying the virtual field is used to shape how the phasing is perceived by the listener.

Chapter 8

Future Work

The focus of the research in this dissertation is on the virtual field component of the SAF. The question that follows is how can the loudspeaker array and acoustic field components of the SAF be used to create compositional elements on short musical timescales. Furthermore, how can these elements be used in conjunction with the methods developed in this dissertation. This work has described many compositional possibilities from the construction of the virtual field to manipulation of loudspeaker gain curves. The next step is to combine these elements on larger timescales of phrases and form.

Another question that arises is how can these elements be used to increase the effectiveness of spatial gestures such as a high velocity source quickly decelerating to a stop. Perhaps reducing gain curve width to the point in which they no longer overlap during the deceleration might increase the listener's ability to track the source because there is no precedence effect and the source is rearticulated at every loudspeaker location.

In this dissertation, spatial objects only interact with other objects through the mapping of their parameter curves. Adding physical simulations such as attractive and repulsive forces to the spatial objects can be used to create complex interactions. In this case, part of the compositional process would be defining the parameters of the physical simulations.

The use of the SAF has implications for works involving audience interaction and participation. Audience members can control parameters of the virtual field using their mobile devices or their mobile devices can serve as the loudspeaker array of the SAF. Apart from using mobile devices, the audience members' positions can be tracked and used as the position of virtual spatial objects. In this case, an audience member can interact with the SAF in a way that is perhaps more transparent to them than interacting with the field using a mobile device.

While this dissertation focuses on spatialization over loudspeakers, the theory of SAF and compositional techniques developed can be applied to spatialization over headphones. Compositions intended for spatialization over headphones inevitably reach a wider audience due to the ubiquity of headphones. With headphones, the loudspeaker array is coupled to the listener and the acoustic field becomes almost entirely simulated in the virtual field. There is an abundance of research and implementation of simulated acoustic fields in addition to compositions using simulated acoustic fields. However, the use of simulated SAFs with headphones can be used in combination with the SAF as described in this dissertation to create an auditory mixed reality. The use of filters and noise cancelation in varying amounts opens up an intriguing area of research for new compositional possibilities.

Appendix A

The Example Layout

A.1 Details

For consistency and to aid in explanation of theory and concepts, a simplified loudspeaker array is referenced throughout this dissertation. This section provides a detailed explanation of this layout and the graphs produced from it.

A.1.1 Layout

The example layout is a 16-channel 2D ring layout with a single source at a fixed distance from the center equal to the loudspeaker distances A.1.

A spherical coordinate system is used, and because the layout is 2D and the source is always at a fixed distance, positions can be specified in terms of azimuth alone. In Figure A.1, the vector that points from the center to loudspeaker 0 has an azimuth equal to 0 and the azimuth increases counterclockwise from $0 - 2\pi$. The loudspeakers are equally spaced around the center at positions $\frac{2\pi n}{N}$ where n and N are the loudspeaker index and number of loudspeakers respectively. The loudspeaker index and channel number are the same for simplicity.

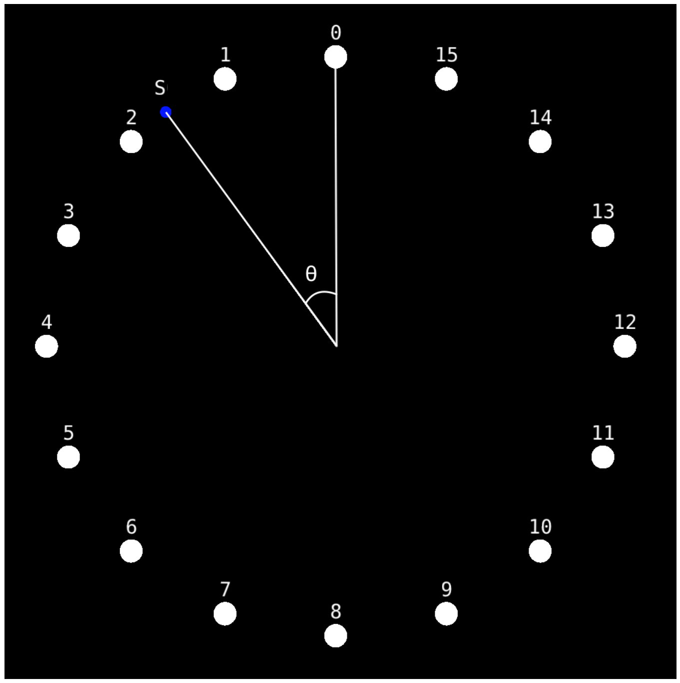
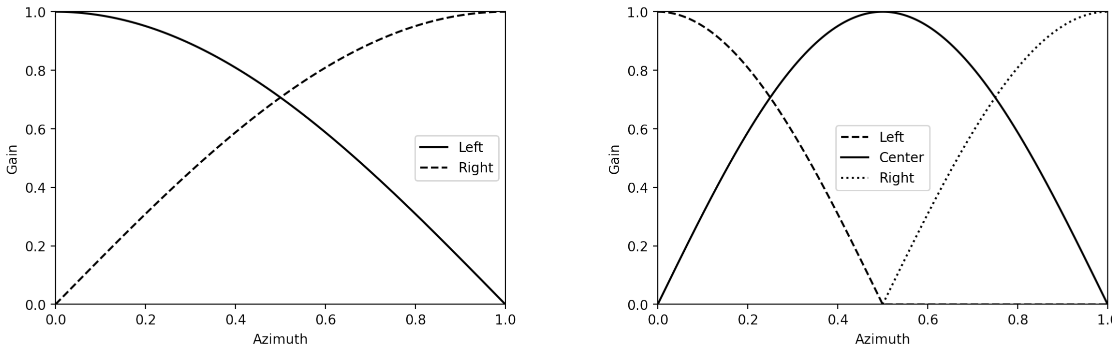


Figure A.1: Top-down view of the example layout.

A.1.2 Gain Curves

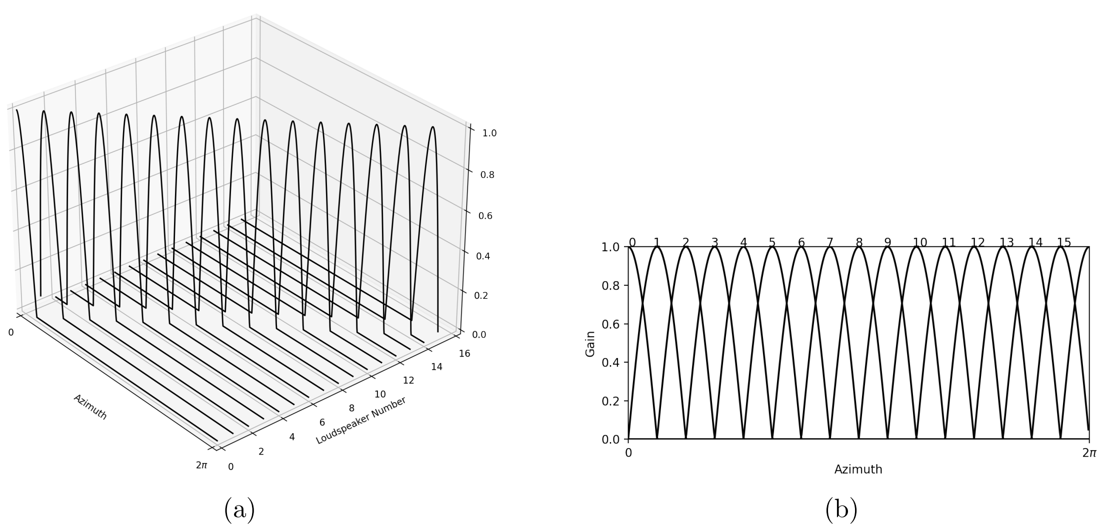
The gain curve of a loudspeaker refers to the loudspeaker's gain based on the source's azimuth. The dissertation refers to the gain curves of loudspeaker pairs, loudspeaker triplets, and the entire loudspeaker array. Loudspeaker pair gain curves are intended to show the per channel gain of each loudspeaker for source positions between the pair (Figure A.2a). These graphs only show half of each loudspeaker's gain curve. Loudspeaker triplet gain curves are intended to show a loudspeakers gain curve in relation to its adjacent loudspeakers (Figure A.2b).

Figure A.3 shows two views of the gain curves calculated using constant power panning for all loudspeakers in the example array. In Figure A.3a, the gain curves are all shown separately. Figure A.3b shows the same gain curves but here the curves are drawn on top of each other with the loudspeaker channel numbers indicated at the apex of its corresponding gain curve. Gain curve graphs of the loudspeaker array used in the



(a) (b)

Figure A.2: Gain curves for a loudspeaker pair (a) and triplet (b). In each graph, the azimuth is normalized to the positions of the left and right loudspeakers.



(a) (b)

Figure A.3: Two perspectives of the loudspeaker array gain curves.

dissertation are represented as shown in A.3b for clarity.

Bibliography

[1] A. F. Carver, *Polychoral music: A venetian phenomenon?*, *Proceedings of the Royal Musical Association* **108** (1981) 1–24.

[2] G. d'Alessi, *Precursors of adriano willaert in the practice of "coro spezzato"*, *Journal of the American Musicological Society* **5** (1952) 187–210.

[3] W. A. Mozart and R. Gerber, "Serenade no. 6. serenata notturna d dur für 2 kleine orchester, das eine bestehend aus 2 solo-violinen, viola und kontrabass, das andere aus streichquartett und pauken. köchel-verzeichnis no. 239. / rev. und mit vorwort versehen von rudolf gerber.."

[4] K. Hudspeth, P. Bakanas, P. Halpin, J. Balmer, and J. Armitage, *mconduct: Gesture transmission and reconstruction for distributed performance*, *In ECLAP 2012 Conference on Information Technologies for Performing Arts, Media Access and Entertainment* (2012) 107–112.

[5] H. Berlioz and R. Fiske, *Requiem: grande messe des morts: op. 5 / hector berlioz; foreword by roger fiske.*, 1969.

[6] G. Mahler, *Symphony no. 2*, 1952.

[7] C. Roads and J. Strawn, *The computer music tutorial.* MIT press, 1996.

[8] C. Roads, *Composing electronic music: a new aesthetic.* Oxford University Press, USA, 2015.

[9] V. Zouhar, R. Lorenz, T. Musil, J. Zmoelnig, and R. Hoeldrich, *Hearing varese's poeme electronique inside a virtual philips pavilion*, 2005.

[10] A. P. Cuervo, *Preserving the electroacoustic music legacy: A case study of the sal-mar construction at the university of illinois*, *Notes* **68** (2011) 33–47.

[11] G. Loy, *About audium-a conversation with stanley shaff*, *Computer Music Journal* **9** (1985) 41–48.

[12] J. Chowning, *Turenas : the realization of a dream*, *Proceedings of The Journées d'Informatique Musicale Journées d'Informatique Musicale* (2011).

[13] F. Zotter and M. Frank, *Ambisonics: A Practical 3D Audio Theory for Recording Studio Production Sound Reinforcement and Virtual Reality.* Springer, 2019.

[14] V. Pulkki, *Virtual sound source positioning using vector base amplitude panning*, *J. Audio Eng. Soc.* **45** (1997) 456–466.

[15] T. Lossius, P. Baltazar, and T. de la Hogue, *Dbap - distance-based amplitude panning*, *Proceedings of the 2009 International Computer Music Conference* (2009).

[16] A. J. Berkhout, D. de Vries, and P. Vogel, *Acoustic control by wave field synthesis*, *The Journal of the Acoustical Society of America* **93** (1993) 2764–2778.

[17] B. Taylor, *A history of the audience as a speaker array*, *Proceedings of the International Conference on New Interfaces for Musical Expression* (2017) 481–486.

[18] G. Levin, *Dialtones-a telesymphony*, 2001.

[19] S. W. Lee and J. Freeman, *echobo : A mobile music instrument designed for audience to play*, *New Interfaces for Musical Expression* (2013).

[20] E. Lyon, T. Caulkins, D. Blount, I. I. Bukvic, C. Nichols, M. Roan, and T. Upthegrove, *Genesis of the cube: The design and deployment of an hdla-based performance and research facility*, *Computer Music Journal* (2016).

[21] A. Cabrera, J. Kuchera-Morin, and C. Roads, *The evolution of spatial audio in the allosphere*, *Computer Music Journal* **40** (12, 2016) 47–61.

[22] C. Ramakrishnan, J. Goßman, and L. Brümmer, *The zkm klangdom*, *NIME '06 Proceedings of the 2006 conference on New interfaces for musical expression* (2006) 140–143.

[23] T. Carpentier, N. Barrett, R. Gottfried, and M. Noisternig, *Holophonic sound in ircam's concert hall: Technological and aesthetic practices*, *Computer Music Journal* **40** (2016) 14–34.

[24] M. Alcorn, C. Corrigan, and P. Rebelo, *Sarc: Studio report*, 2005.

[25] "What is thx spatial audio? how does it help consumers?." https://www.thx.com/spatial-audio. Accessed February 17, 2022.

[26] "Dolby atmos - official site - dolby." https://www.dolby.com/technologies/dolby-atmos. Accessed February 22, 2022.

[27] N. Peters, G. Marentakis, and S. McAdams, *Current technologies and compositional practices for spatialization: A qualitative and quantitative analysis*, *Computer Music Journal* (2011).

[28] C. Roads and S. Ramakrishnan, *Spatialchords : an interface for projection of multichannel spatial geometries live in concert*, .

[29] G. S. Kendall, *Spatial perception and cognition in multichannel audio for electroacoustic music*, *Organised Sound* **15** (2010) 228–238.

[30] L. Brümmer, *Composition and perception in spatial audio*, *Computer Music Journal* (2017).

[31] R. Normandeau, *Timbre spatialisation: The medium is the space*, *Organised Sound* **14** (2009) 277–285.

[32] A. Einbond and D. Schwarz, *Spatializing timbre with corpus-based concatenative synthesis*, pp. 72–75, Michigan Publishing, 2010.

[33] C. Roads, *Microsound*. MIT press, 2004.

[34] R. M. McGee, *Scanning spaces: Paradigms for spatial sonification and synthesis*, 2015.

[35] G. S. Kendall and A. Cabrera, *Why things don't work: What you need to know about spatial audio*, *ICMC* (2011).

[36] G. S. Kendall, *The decorrelation of audio signals and its impact on spatial imagery*, *Computer Music Journal* **19** (1995) 71–87.

[37] N. Barrett, *A musical journey towards permanent high-density loudspeaker arrays*, *Computer Music Journal* (2016).

[38] F. Bayle, *Musique acousmatique propositions–positions*, .

[39] A. Shupak, Z. Sharoni, Y. Yanir, Y. Keynan, Y. Alfie, and P. Halpern, *Underwater hearing and sound localization with and without an air interface*, *Otology and Neurotology* **26** (1, 2005) 127–130.

[40] K. L. Hagan, *Textural composition: Aesthetics, techniques, and spatialization for high-density loudspeaker arrays*, *Computer Music Journal* (2017).

[41] R. Carl, *Terry Riley's In C*. Oxford University Press, 2009.

[42] S. Reich, *Music as a gradual process*, *Writings on music* (1965) 34–36.

[43] S. Reich, *Come Out*. 1966.

[44] S. Reich, *Piano Phase*. 1967.

[45] S. Reich, *Pendulum Music*. 1968.

[46] "Github - allosphere-research-group/allolib: Library for interactive multimedia application development." https://github.com/AlloSphere-Research-Group/allolib. Accessed February 16, 2022.

[47] "Github - allosphere-research-group/gamma: Generic (sound) synthesis library." https://github.com/AlloSphere-Research-Group/Gamma. Accessed February 2, 2022.

[48] F. Zotter, M. Frank, G. Marentakis, and A. Sontacchi, *Phantom source widening with deterministic frequency dependent time delays*, *Proceedings of the 14th International Conference on Digital Audio Effects* (2011).

[49] "Osc index." https://ccrma.stanford.edu/groups/osc/index.html. Accessed March 5, 2022.